

RNA MODIFICATIONS AND EPITRANSCRIPTOMICS

EDITED BY: Xiao Han, Jia Meng and Giovanni Nigita

PUBLISHED IN: Frontiers in Cell and Developmental Biology and
Frontiers in Genetics



frontiers

Frontiers eBook Copyright Statement

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

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

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

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

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

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

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

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

ISSN 1664-8714

ISBN 978-2-88976-106-7

DOI 10.3389/978-2-88976-106-7

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

RNA MODIFICATIONS AND EPITRANSCRIPTOMICS

Topic Editors:

Xiao Han, Biotechnology Research Institute, Chinese Academy of Agricultural Sciences, China

Jia Meng, Xi'an Jiaotong-Liverpool University, China

Giovanni Nigita, The Ohio State University, United States

Citation: Han, X., Meng, J., Nigita, G., eds. (2022). RNA Modifications and Epitranscriptomics. Lausanne: Frontiers Media SA.
doi: 10.3389/978-2-88976-106-7

Table of Contents

- 06 *The Potential Roles of RNA N⁶-Methyladenosine in Urological Tumors***
Yang Li, Yu-zheng Ge, Luwei Xu, Zheng Xu, Quanliang Dou and Ruipeng Jia
- 20 *Transcriptome-Wide m⁶A Methylation in Skin Lesions From Patients With Psoriasis Vulgaris***
Ya-Nan Wang and Hong-Zhong Jin
- 33 *RNA-Binding Protein HuR Suppresses Inflammation and Promotes Extracellular Matrix Homeostasis via NKRF in Intervertebral Disc Degeneration***
Zhenxuan Shao, Zhuolong Tu, Yifeng Shi, Sunlong Li, Aimin Wu, Yaosen Wu, Naifeng Tian, Liaojun Sun, Zongyou Pan, Linwei Chen, Weiyang Gao, Yifei Zhou, Xiangyang Wang and Xiaolei Zhang
- 47 *Expression and Prognostic Characteristics of m⁶A RNA Methylation Regulators in Breast Cancer***
Bo Zhang, Yanlin Gu and Guoqin Jiang
- 61 *Structural Insights Into m⁶A-Erasers: A Step Toward Understanding Molecule Specificity and Potential Antiviral Targeting***
Mahmoud Bayoumi and Muhammad Munir
- 75 *A Novel m⁶A Gene Signature Associated With Regulatory Immune Function for Prognosis Prediction in Clear-Cell Renal Cell Carcinoma***
Siteng Chen, Ning Zhang, Encheng Zhang, Tao Wang, Liren Jiang, Xiang Wang and Junhua Zheng
- 86 *Alteration of mRNA 5-Methylcytosine Modification in Neurons After OGD/R and Potential Roles in Cell Stress Response and Apoptosis***
Huan Jian, Chi Zhang, ZhangYang Qi, Xueying Li, Yongfu Lou, Yi Kang, Weimin Deng, Yigang Lv, Chaoyu Wang, Wei Wang, Shenghui Shang, Mengfan Hou, Hengxing Zhou and Shiqing Feng
- 99 *Profiling of RNA N⁶-Methyladenosine Methylation Reveals the Critical Role of m⁶A in Chicken Adipose Deposition***
Bohan Cheng, Li Leng, Ziwei Li, Weijia Wang, Yang Jing, Yudong Li, Ning Wang, Hui Li and Shouzhi Wang
- 113 *m⁶A RNA Methylation Regulators Act as Potential Prognostic Biomarkers in Lung Adenocarcinoma***
Hongbo Wang, Xiangxuan Zhao and Zaiming Lu
- 123 *Identification of m⁶A-Associated RNA Binding Proteins Using an Integrative Computational Framework***
Yiqian Zhang and Michiaki Hamada
- 133 *Spliceosomal snRNA Epitranscriptomics***
Pedro Morais, Hironori Adachi and Yi-Tao Yu
- 149 *HN-CNN: A Heterogeneous Network Based on Convolutional Neural Network for m⁷G Site Disease Association Prediction***
Lin Zhang, Jin Chen, Jiani Ma and Hui Liu
- 157 *The Roles of CircRNAs in Regulating Muscle Development of Livestock Animals***
Zhenguo Yang, Tianle He and Qingyun Chen

- 169 ***Deciphering Epitranscriptome: Modification of mRNA Bases Provides a New Perspective for Post-transcriptional Regulation of Gene Expression***
Suresh Kumar and Trilochan Mohapatra
- 191 ***Identifying RNA N6-Methyladenine Sites in Three Species Based on a Markov Model***
Cong Pian, Zhixin Yang, Yuqian Yang, Liangyun Zhang and Yuanyuan Chen
- 197 ***Dynamics of m6A RNA Methylome on the Hallmarks of Hepatocellular Carcinoma***
Enakshi Sivasudhan, Neil Blake, Zhi-Liang Lu, Jia Meng and Rong Rong
- 211 ***Prognostic Significance and Tumor Immune Microenvironment Heterogeneity of m5C RNA Methylation Regulators in Triple-Negative Breast Cancer***
Zhidong Huang, Junfan Pan, Helin Wang, Xianqiang Du, Yusheng Xu, Zhitang Wang and Debo Chen
- 226 ***DeepOMe: A Web Server for the Prediction of 2'-O-Me Sites Based on the Hybrid CNN and BLSTM Architecture***
Hongyu Li, Li Chen, Zaoli Huang, Xiaotong Luo, Huiqin Li, Jian Ren and Yubin Xie
- 235 ***Co-expression Network Revealed Roles of RNA m⁶A Methylation in Human β -Cell of Type 2 Diabetes Mellitus***
Cong Chen, Qing Xiang, Weilin Liu, Shengxiang Liang, Minguang Yang and Jing Tao
- 246 ***Role of RNA N6-Methyladenosine Modification in Male Infertility and Genital System Tumors***
Shuai Liu, Yongfeng Lao, Yanan Wang, Rongxin Li, Xuefeng Fang, Yunchang Wang, Xiaolong Gao and Zhilong Dong
- 259 ***Emerging Role of m⁶A Methylome in Brain Development: Implications for Neurological Disorders and Potential Treatment***
Godwin Sokpor, Yuanbin Xie, Huu P. Nguyen and Tran Tuoc
- 283 ***REW-ISA V2: A Biclustering Method Fusing Homologous Information for Analyzing and Mining Epi-Transcriptome Data***
Lin Zhang, Shutao Chen, Jiani Ma, Zhaoyang Liu and Hui Liu
- 293 ***N6-Methyladenosine RNA Modification in Inflammation: Roles, Mechanisms, and Applications***
Jiahui Luo, Tao Xu and Kai Sun
- 306 ***The MicroRNA Family Gets Wider: The IsomiRs Classification and Role***
Luisa Tomasello, Rosario Distefano, Giovanni Nigita and Carlo M. Croce
- 321 ***m5C-Related lncRNAs Predict Overall Survival of Patients and Regulate the Tumor Immune Microenvironment in Lung Adenocarcinoma***
Junfan Pan, Zhidong Huang and Yiquan Xu
- 335 ***N6-Methyladenosine Modification and Its Regulation of Respiratory Viruses***
Qianyu Feng, Hongwei Zhao, Lili Xu and Zhengde Xie
- 341 ***RESIC: A Tool for Comprehensive Adenosine to Inosine RNA Editing Site Identification and Classification***
Dean Light, Roni Haas, Mahmoud Yazbak, Tal Elfand, Tal Blau and Ayelet T. Lamm

- 353** *Quantitative Analysis of Methylated Adenosine Modifications Revealed Increased Levels of N⁶-Methyladenosine (m⁶A) and N⁶,2'-O-Dimethyladenosine (m⁶Am) in Serum From Colorectal Cancer and Gastric Cancer Patients*
Yiqiu Hu, Zhihao Fang, Jiayi Mu, Yanqin Huang, Shu Zheng, Ying Yuan and Cheng Guo
- 362** *RNA Motifs and Modification Involve in RNA Long-Distance Transport in Plants*
Tao Wang, Xiaojun Li, Xiaojing Zhang, Qing Wang, Wenqian Liu, Xiaohong Lu, Shunli Gao, Zixi Liu, Mengshuang Liu, Lihong Gao and Wenna Zhang
- 372** *m⁶A Modification Mediates Mucosal Immune Microenvironment and Therapeutic Response in Inflammatory Bowel Disease*
Yongyu Chen, Jing Lei and Song He



The Potential Roles of RNA N6-Methyladenosine in Urological Tumors

Yang Li[†], Yu-zheng Ge[†], Luwei Xu, Zheng Xu, Quanliang Dou and Ruipeng Jia^{*}

Department of Urology, Nanjing First Hospital, Nanjing Medical University, Nanjing, China

OPEN ACCESS

Edited by:

Jia Meng,
Xi'an Jiaotong-Liverpool University,
China

Reviewed by:

Tao P. Wu,
Baylor College of Medicine,
United States
Huilin Huang,
Sun Yat-sen University Cancer Center
(SYSUCC), China

*Correspondence:

Ruipeng Jia
ruipengj@163.com

[†]These authors have contributed
equally to this work

Specialty section:

This article was submitted to
Epigenomics and Epigenetics,
a section of the journal
Frontiers in Cell and Developmental
Biology

Received: 03 July 2020

Accepted: 24 August 2020

Published: 09 September 2020

Citation:

Li Y, Ge Y, Xu L, Xu Z, Dou Q and
Jia R (2020) The Potential Roles
of RNA N6-Methyladenosine
in Urological Tumors.
Front. Cell Dev. Biol. 8:579919.
doi: 10.3389/fcell.2020.579919

N6-methyladenosine (m⁶A) is regarded as the most abundant, prevalent and conserved internal mRNA modification in mammalian cells. M⁶A can be catalyzed by m⁶A methyltransferases METTL3, METTL14 and WTAP (writers), reverted by demethylases ALKBH5 and FTO (erasers), and recognized by m⁶A -binding proteins such as YTHDF1/2/3, IGF2BP1/2/3 and HNRNPA2B1 (readers). Emerging evidence suggests that m⁶A modification is significant for regulating many biological and cellular processes and participates in the pathological development of various diseases, including tumors. This article reviews recent studies on the biological function of m⁶A modification and the methylation modification of m⁶A in urological tumors.

Keywords: N6-methyladenosine (m⁶A), writers, erasers, readers, urological tumors

INTRODUCTION

In past decades, epigenetic modification has been identified to be involved in diverse biological processes and disease progression, attracting more and more attention. Epigenetics is a study of reversible, inheritable phenotypes that do not involve changes in nuclear DNA sequences (Mohammad et al., 2019), and primarily includes RNA interference, histone modification, chromatin rearrangement, DNA methylation and RNA modification (Arguello et al., 2019; McGee and Hargreaves, 2019).

RNA modification was previously regarded as occurring in high-abundance RNA species, while emerging evidence indicates that it is characterized in lowly abundant species of RNA such as non-coding RNAs and Mrna (Dominissini, 2014; Li X. et al., 2016). Among them, RNA methylation has attracted accumulating attention in recent years and N6-methyladenosine (m⁶A) is the most prevalent RNA methylation sites (Pan, 2013). M⁶A modification was firstly reported to be interrelated to the regulation of gene expression, growth and development in 1970s (Desrosiers et al., 1974; Perry et al., 1975; Chandola et al., 2015; Hsu et al., 2017), and it has been regarded as one of the most common mRNA modifications recently. Many researches have revealed that m⁶A modification mainly occurred in the consensus sequence RRACH sequence (R = A, G; H = A, C, U) (Li L. J. et al., 2018), which is enriched in stop codons, 3' untranslated region (UTR) and the last exon in non-coding RNA (Dominissini et al., 2012; Meyer et al., 2012). Besides, m⁶A is widespread in RNA of bacteria, viruses and eukaryotes (Desrosiers et al., 1974; Wang Y. et al., 2014; Deng et al., 2015; Fu et al., 2015; Greer et al., 2015; Zhang et al., 2015; Liu J. et al., 2016; Zhu et al., 2018).

M⁶A modification is reversible and catalyzed by many relevant enzymes (Batista, 2017; Dai et al., 2018). Studies have shown that m⁶A is involved in various biological and disease processes via regulating target gene expression (Chen X.Y. et al., 2019; Lan et al., 2019). M⁶A modification is associated with various diseases, such as neurological diseases (Liu E. Y et al., 2017;

Salta and De Strooper, 2017) and cancers. In this review, we provide a broad overview of the relationship between RNA m⁶A methylation and urological tumors. We further highlight the possible uses in diagnostic, prognostic and therapeutic applications of m⁶A modifications for urological tumors.

REGULATORS OF M⁶A

Similar to histone modification and DNA methylation, m⁶A modification is reversible and dynamic, and influences biological functions that are primarily mediated by three types of regulators: methyltransferases (“writers”), demethylases (“erasers”) and m⁶A binding proteins (“readers”). The methyltransferase complex (MTC) can catalyze m⁶A, demethylase can remove m⁶A, while RNA reader proteins can recognize m⁶A and bind to the RNA. These proteins play an essential biological role in m⁶A modifications (Table 1, Figure 1). Cross-talk among writers, erasers and readers of m⁶A is involved in the development and progression of tumors (Deng et al., 2018; Panneerdoss et al., 2018).

Methyltransferases (“Writers”)

MTC has been identified to regulate the installation of m⁶A and Methyltransferase-like 3 (METTL3), METTL14, and Wilms tumor 1-associated protein (WTAP) have been proved as the core components of this complex (Ping et al., 2014; Schwartz et al., 2014; Zhou J. et al., 2015). METTL3 is an Sadenosyl methionine (SAM)-binding protein and regarded as a major catalytic enzyme with functions reminiscent of the N6-adenine methyltransferase system (Barbieri et al., 2017). Besides, METTL3 is highly conserved in eukaryotes from yeast to humans (Bokar et al., 1997). WTAP can also increase the binding ability of METTL3, thus regulating recruitment of the complex to mRNA targets (Ping et al., 2014). METTL14 could form a stable complex with METTL3 and both of them contain a SAM-binding motif. With the help of WTAP, METTL3-METTL14 could colocalize in nuclear speckles and form a heterodimer, so as to participate in catalytic activity (Liu J. et al., 2014; Zhao X. et al., 2014). Besides, VIRMA, RBM15, ZC3H13 and KIAA1429 are the new components of the m⁶A “writer” complex (Moindrot et al., 2015; Wang X. et al., 2016; Deng et al., 2018; Wen et al., 2018).

METTL3

The writer METTL3 has been identified to be involved in various biological processes. METTL3 can enhance the BAT-mediated adaptive thermogenesis and suppress obesity and systemic insulin resistance via targeting the 3' UTR of the PRDM16, PPARG, and UCP1 transcript to install the m⁶A modification (Wang Y. et al., 2020). The ablation of METTL3 in germ cells severely inhibited spermatogonial differentiation and blocked the initiation of meiosis (Xu et al., 2017). Besides, METTL3 was also shown to be upregulated in various solid tumors and associated with poor prognosis. In oral squamous

cell carcinoma (OSCC), METTL3 can facilitate tumor growth and metastasis through making an increment in m⁶A modification and expression of c-Myc transcript (Zhao W. et al., 2020). In colorectal cancer (CRC), METTL3 stabilizes HK2 and GLUT1 expression via a m⁶A -IGF2BP2/3-dependent mechanism (Shen et al., 2020). Additionally, METTL3 might affect tumor metastasis through promoting the maturation of pri-miR-1246 (Peng et al., 2019). METTL3 enhances the splicing of precursor miR-143-3p and facilitates its biogenesis, thereby promoting the brain metastasis of lung cancer (LC) (Wang H. et al., 2019). Moreover, METTL3 induces non-small cell lung cancer (NSCLC) drug resistance and metastasis by promoting Yes-associated protein (YAP) mRNA translation via a m⁶A -YTHDF1/3/eIF3b-dependent mechanism (Jin D. et al., 2019). In gastric cancer (GC), overexpression of METTL3 can promote the stability of ZMYM1, thereby enhancing epithelial mesenchymal transformation (EMT) process and tumor metastasis (Yue et al., 2019). In addition, upregulated METTL3 facilitates GC growth and liver metastasis through installing m⁶A modifications of HDGF transcript (Wang Q. et al., 2020).

METTL14

Studies have demonstrated that METTL14 is associated with a lower risk for development of neoplasms. In CRC, METTL14 acts as a tumor-suppressor to inhibit cell growth and metastasis *in vitro* and *in vivo*. Mechanistical study demonstrated that downregulated METTL14 substantially abolishes m⁶A modifications of XIST and augments XIST expression (Yang X. et al., 2020). In addition, METTL14 can inhibit CRC cell proliferation, migration and invasion via the miR-375-YAP1/SP1 signal axis (Chen X. et al., 2020). Although both of METTL3 and METTL14 could act as m⁶A “writer”, METTL3 might promote the progression of CRC, while METTL14 functions as a tumor suppressor in CRC. METTL14 can also assume an oncogenic role in triple-negative breast cancer (TNBC) (Shi et al., 2020), pancreatic cancer (Kong et al., 2020) and leukemia (Weng et al., 2018). Moreover, METTL14 is significantly upregulated in Epstein-Barr virus (EBV) latently infected cells. METTL14 can lead to oncogenesis via increasing m⁶A modifications of the indispensable EBV latent antigen EBNA3C and thus facilitating its stability and expression. Interestingly, EBNA3C can also enhance stability and expression of METTL14 (Lang et al., 2019).

METTL16

METTL16 has been recently shown to have distinct target RNAs for m⁶A modification. Studies have revealed that METTL16 can bind a subset of mRNAs and methylate U6 small nuclear RNA (U6 snRNA) and long non-coding RNA (lncRNA) (Brown et al., 2016; Fitzsimmons and Batista, 2019). Moreover, the UACAGAGAA sequence is essential for METTL16-mediated-methylation and the Nterminal module of METTL16 is required for RNA binding (Doxader et al., 2018; Mendel et al., 2018). METTL16 is involved in catalyzing m⁶A in A43 of the U6

TABLE 1 | Functions of m⁶A regulators in RNA metabolism.

Type	m ⁶ A Regulators	Function	References
m ⁶ A writer	METTL3	Catalyzes m ⁶ A modification	Schwartz et al., 2014 Zhou J. et al., 2015
–	METTL14	Forms a stable complex with METTL3	Schwartz et al., 2014
–	–	–	Zhou J. et al., 2015
–	METTL16	Catalyzes m ⁶ A modification	Warda et al., 2017
–	WTAP	Contributes to the localization of METTL3-METTL14 heterodimer to the nuclear speckle	Ping et al., 2014
–	RBM15	Binds the m ⁶ A complex and recruit it to special RNA site	Moindrot et al., 2015
–	VIRMA	Recruits the m ⁶ A complex to the special RNA site and interacts with polyadenylation	Wang T. et al., 2020
–	–	Cleavage factors CPSF5 and CPSF6	–
–	–	–	–
–	ZC3H13	Bridges WTAP to the mRNA-binding factor Nito	Wen et al., 2018
m ⁶ A eraser	FTO	Mediates demethylation of both hm ⁶ A and f ⁶ A in mRNA	Basak et al., 2019
–	–	–	–
–	ALKBH5	Removes m ⁶ A modification	Tang et al., 2018
m ⁶ A reader	YTHDF1	Facilitates mRNA translation efficiency	Liu J. et al., 2020
–	–	–	–
–	YTHDF2	Promotes mRNA degradation	Zhou J. et al., 2015
–	YTHDF3	Enhances translation and degradation by interacting with YTHDF1 and YTHDF2	Shi et al., 2017
–	–	–	Li A. et al., 2017
–	YTHDC1	Recruits the RNA splicing and controls the nuclear export	Roundtree et al., 2017b
–	YTHDC2	Interacts with RNA helicase and increases the translation efficiency of target RNA	Mao et al., 2019
–	–	–	–
–	IGF2BPs	Recruits RNA stabilizers	Huang H. et al., 2018
–	HNRNPA2B1	Mediates mRNA splicing and primary microRNA processing	Alarcon et al., 2015
–	–	–	–
–	HNRNPC	Influences alternative splicing and mRNA localization	Guichard et al., 2012
–	EIF3	Facilitates cap-independent translation	Meyer et al., 2015

small nuclear RNA (Warda et al., 2017). Under loss-of-SAM conditions, METTL16 can induce the splicing of a retained intron, thereby enhancing level of MAT2A and expression of SAM, while down-regulation of METTL16 and YTHDC1 can abolish SAM-responsive regulation of MAT2A (Pendleton et al., 2017; Shima et al., 2017). While the specific role of METTL16 in solid tumors remain to be further explored.

Demethylases (“Erasers”)

The reversible and dynamic m⁶A modification can be mediated by obesity-associated protein (FTO) and alkB homolog 5 (ALKBH5) (m⁶A “erasers”) (Jia et al., 2011; Zheng et al., 2013). Both FTO and ALKBH5 are members of the ALKB family of dioxygenases. As the first reported demethylase, FTO can also mediate demethylation of both N⁶-hydroxymethyladenosine (hm⁶A) and N⁶-formyladenosine (f⁶A) in mRNA (Basak et al., 2019). ALKBH5 plays an essential role in mRNA export and RNA metabolism (Tang et al., 2018).

FTO

As an m⁶A eraser, FTO is associated with the initiation and development of various cancers including hepatocellular carcinoma (HCC), melanoma, breast cancer and glioma. In HCC, SIRT1 destabilizes FTO and thus steering the m⁶A of downstream elements and consecutive mRNA expression in tumorigenesis (Liu X. et al., 2020). In melanoma, FTO can impair IFN γ -induced killing via augmenting CXCR4, PD-1 and SOX10 expression via repressing YTHDF2-mediated degradation and suppress response to anti-PD-1 blockade immunotherapy (Yang S. et al., 2019). In breast cancer, FTO enhances breast cancer cell growth, colony formation and metastasis. Mechanistical study demonstrated that FTO can mediate m⁶A demethylation of BNIP3 transcript and induce its degradation via an YTHDF2 independent mechanism (Niu et al., 2019). The ethyl ester form of meclofenamic acid (MA2) inhibits FTO and enhances the effect of the chemotherapy drug

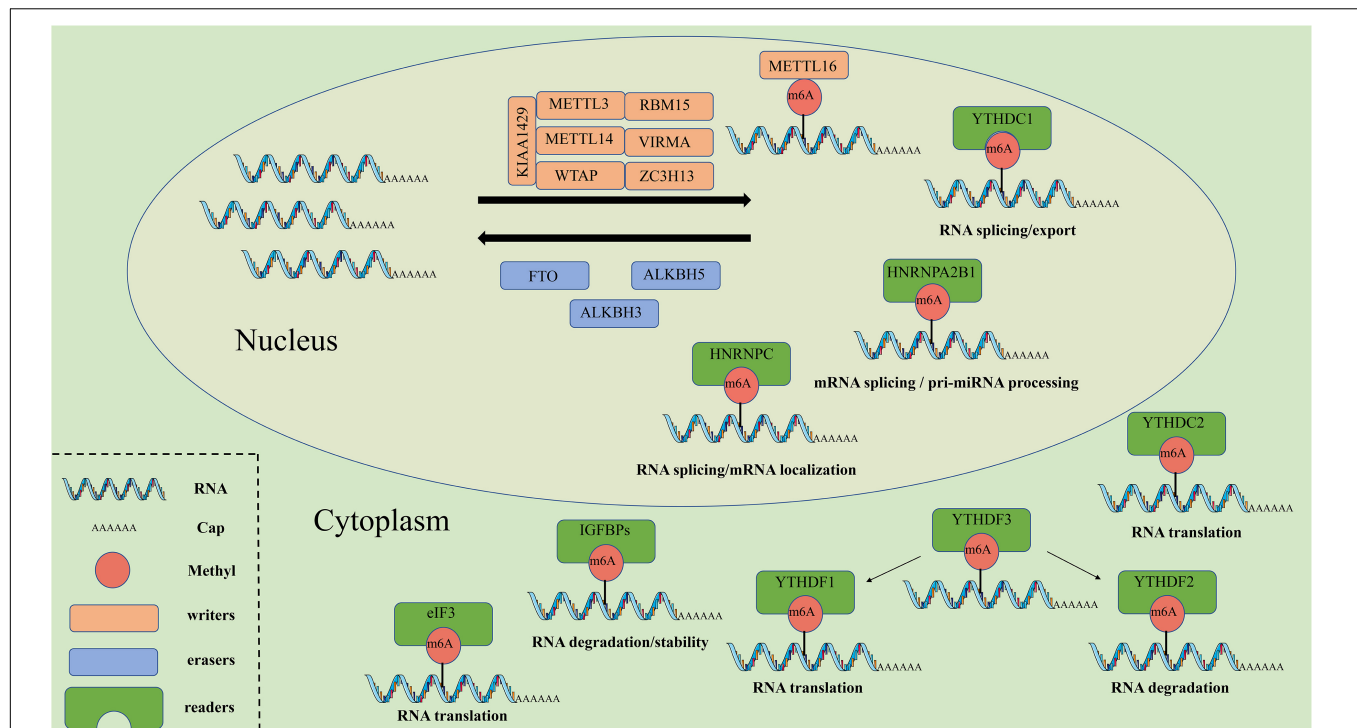


FIGURE 1 | The molecular mechanism of m⁶A. M⁶A can be installed by “writers” (METTL3/14/16, WTAP, RBM15, VIRMA, KIAA1429, and ZC3H13), removed by “erasers” (FTO, ALKBH5, and ALKBH3), and recognized by “readers” (YTHDF1/2/3, YTHDC1/2, IGF2BPs, HNRNPs, and eIF3). METTL3/14/16, methyltransferase like 3/14/16; WTAP, WT1 associated protein; RBM15, RNA binding motif protein 15; VIRMA, vir like m⁶A methyltransferase associated; ZC3H13, zinc finger CCHH-type containing 13; FTO, FTO alpha-ketoglutarate dependent dioxygenase; ALKBH5, alkB homolog 5, RNA demethylase; ALKBH3, alkB homolog 3, RNA demethylase; YTHDF1/2/3, YTH N6-methyladenosine RNA binding protein 1/2/3; YTHDC1/2, YTH domain containing 1/2; IGF2BPs, insulin like growth factor 2 mRNA binding proteins; HNRNPs, heterogeneous nuclear ribonucleo proteins; eIF3, eukaryotic translation initiation factor 3 subunit.

temozolomide (TMZ) on suppressing proliferation of glioma cells (Xiao et al., 2020).

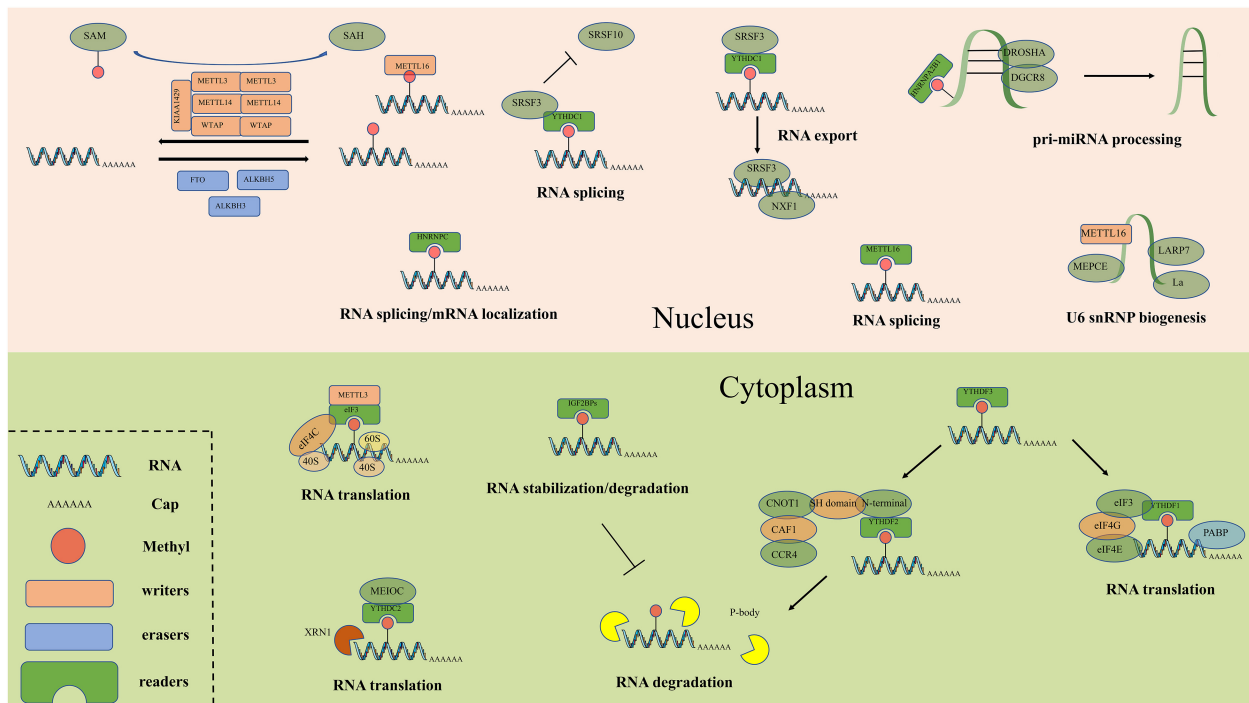
ALKBH5

ALKBH5 has been regarded as a tumor suppressor in many cancers. In NSCLC, ALKBH5 suppresses cell growth and metastasis both *in vitro* and *in vivo* via repressing miR-107/LATS2-mediated YAP activity and YTHDFs-mediated YAP expression (Jin D. et al., 2020). In pancreatic cancer, downregulated ALKBH5 predicts poor prognosis and knockdown of ALKBH5 markedly facilitates tumor growth and metastasis (Tang et al., 2020). In HCC, ALKBH5 is characterized as a tumor suppressor and could attenuate the expression of LYPD1 via an m⁶A-dependent manner in HCC cells (Chen Y. et al., 2020). In addition, ALKBH5 can augment steady-state CYR61 mRNA expression via an m⁶A -dependent mechanism, thereby repressing trophoblast invasion (Li X. C. et al., 2019).

m⁶A Binding Proteins (“Readers”)

M⁶A readers can recognize and bind to m⁶A sites and regulate target RNA translation, splicing, nuclear export and decay (Figure 2). In YTH (YT521-B homology) domain family, the evolutionarily conserved YTH domain acts as the module

for directly binding to m⁶A. YTHDF1–3 and YTHDC1–2 are the main five YTH domain proteins. YTHDF1 can bind to m⁶A sites around the stop codon and thus facilitating mRNA translation efficiency (Liu X. et al., 2020). YTHDF2 can accelerate degradation and deadenylation of the transcripts by bringing m⁶A-modified translatable mRNAs to mRNA decay sites and recruiting CCR4-NOT deadenylase complex (Zhou J. et al., 2015). YTHDF3 can, respectively, promote RNA translation through associating with YTHDF1 and enhance RNA degradation by interacting with YTHDF2 (Li A. et al., 2017; Shi et al., 2017). In contrast to the prevailing model, where each DF paralog binds to distinct subsets of mRNAs, Zaccara and Jaffrey show that the DF paralogs bind proportionately to each m⁶A site throughout the transcriptome (Zaccara and Jaffrey, 2020). YTHDC1 recruits the RNA splicing and control the nuclear export (Roundtree et al., 2017b). YTHDC2 interacts with RNA helicase and increases the translation efficiency of target RNA (Mao et al., 2019). The insulin-like growth factor 2 mRNA binding protein (IGF2BP) family proteins, including IGF2BP1–3, can recognize m⁶A containing transcripts. IGF2BPs exert their functions via recruiting RNA stabilizers (Huang H. et al., 2018). Eukaryotic initiation factor 3 (EIF3) can facilitate cap-independent translation (Meyer et al., 2015). Heterogeneous nuclear ribo nucleo protein (HNRNP) family proteins include hnRNP, hnRNPG and hnRNPA2B1. HnRNP



September 2020 | Volume 8 | Article 579919

recognize m⁶A-modified GAS5 and induce decay of it (Ni et al., 2019). YTHDF3 can serve as a negative regulator to enhance the translation of FOXO3 mRNA, thereby maintaining host antiviral immune function and preventing inflammatory response (Zhang Y. et al., 2019).

YTHDC1

YTHDC1 and YTHDC2 have conserved m⁶A binding domain and preferentially bind to m⁶A-modified RNA in RRM6ACH consensus sequence (Roundtree et al., 2017a). YTHDC1 is involved in processing of pre-mRNA transcripts of F6, SRSF3, and SRSF7 in the oocyte nucleus, and it may play a crucial role in fetal development (Kasowitz et al., 2018). MAT2A mRNA can be methylated by METTL16 and YTHDC1 can bind to the m⁶A modification site of MAT2A 3'-UTR. Downregulation of METTL16 and YTHDC1 might effectively abolish SAM-responsive regulation of MAT2A (Shima et al., 2017). The m⁶A modification site of long non-coding RNA X-inactive specific transcript (XIST) can be preferentially read by YTHDC1 and it's required for XIST function (Patil et al., 2016). Recent study shows that the ability of the YTH domain of YTHDC1 binding to ssDNA is stronger than in an RNA context. However, the YTH domains of YTHDF2 and YTHDF1 exhibit the opposite effect (Woodcock et al., 2020).

YTHDC2

YTHDC2 could bind mitotic transcripts, specific piRNA precursors and interact with RNA granule components, licensing the proper progression of germ cells through meiosis (Bailey et al., 2017). YTHDC2 results in colon cancer metastasis through augmenting translation of HIF-1 α , it may be a potential diagnostic marker and therapy target in colon cancer (Tanabe et al., 2016). YTHDC2 binds to the mRNA of lipogenic genes and participates in the regulation of hepatic lipogenesis and TG homeostasis (Zhou B. et al., 2020).

IGF2BPs

IGFBPs could use common RNA binding domains to recognize m⁶A containing transcripts and play a significant role in many diseases. In breast cancer, FGF13-AS1 can reduce the half-life of c-Myc (Myc) mRNA by binding IGF2BPs, thus suppressing cell proliferation, migration and invasion (Ma et al., 2019). In ovarian cancer, IGF2BP1 enhances cell proliferative and invasive ability by antagonizing miRNA-impaired gene expression, the elevate expression of IGF2BP1 is correlated to poor prognosis (Muller et al., 2018). IGF2BP1 could function as an adaptor protein to recruit the CCR4-NOT complex, so as to initiate the degradation of the lncRNA highly up-regulated in liver cancer (HULC) (Hammerle et al., 2013). In pancreatic cancer, IGF2BP2 could promote cell growth through activating the PI3K/Akt signaling pathway and be negatively regulated by miR-141 (Xu

et al., 2019). In addition, IGF2BP2 enhances cancer stemness-like properties and promotes tumorigenesis by acting as a reader for m⁶A modified DANCER (Hu et al., 2020). In gastric cancer, miR-34a directly targets IGF2BP3, overexpression of IGF2BP3 promotes cell proliferation and invasion (Zhou Y. et al., 2017). IGF2BP3 could interact with RNA-binding protein Lin28b and thereby promotes stability and expression of target mRNAs such as B-cell regulators Pax5 and Arid3a, so as to participate in the fetal–adult hematopoietic switch (Wang S. et al., 2019).

EIF3

EIF3 is crucial for specialized translation initiation via interacting with the 5' cap region, resulting in assemblage of translation initiation complexes on eIF3-specialized mRNA (Lee et al., 2016). Study has proved that YTHDF1 might promote the translation of EIF3 via recognizing the m⁶A-modified sites of EIF3 mRNA and simultaneously augments the overall translational output, thus facilitating tumorigenesis and metastasis in ovarian cancer (Liu X. et al., 2020). In renal cell carcinoma (RCC), knockdown of EIF3 dramatically decreases cell viability with sunitinib treatment. Mechanistically, EIF3 could interact with GRP78 and enhance protein stability by blocking the ubiquitin-mediated degradation of GRP78 (Huang H. et al., 2019). In gallbladder cancer (GBC), EIF3 can stabilize GRK2 protein through blocking ubiquitin-mediated degradation, wherefore activating PI3K/Akt signal pathway and enhancing tumor growth and metastasis (Zhang et al., 2017). All above studies demonstrate that EIF3 is a vital role in the progression of various cancers.

ROLES OF RNA M⁶A IN UROLOGICAL TUMORS

Accumulating evidence indicates that RNA m⁶A modification is related to the tumorigenesis, development and progression of urological tumors. Therefore, we summarize these latest advances of m⁶A modification in urological tumors (Table 2, Figure 3).

Renal Cell Carcinoma

Renal cell carcinoma (RCC) is derived from renal epithelium and is one of the most common cancers worldwide, making up nearly 2% to 3% of all adult malignancies (Rini et al., 2009). Li and collaborators demonstrated that METTL3 was a potential prognostic marker of RCC, and the expression levels of METTL3 are interrelated to tumor size and histological grade. Inhibition of METTL3 could obviously promote cell proliferation, migration and invasion, and make cell cycle arrest (Li X. et al., 2017). In addition, METTL3 knockdown might activate oncogenic PI3K/Akt/mTOR signaling pathway. Hence, METTL3 might function as a tumor suppressor in the tumorigenesis of RCC. Gong and co-workers found that the expression level of METTL14 is decreased in RCC (Gong et al., 2019). Additionally, the mRNA level of METTL14 is associated with RCC patients' overall survival. Knockdown of METTL14 promotes the mRNA and protein expression levels of P2RX6, while P2RX6 could

TABLE 2 | The roles of RNA m⁶A in urological tumors.

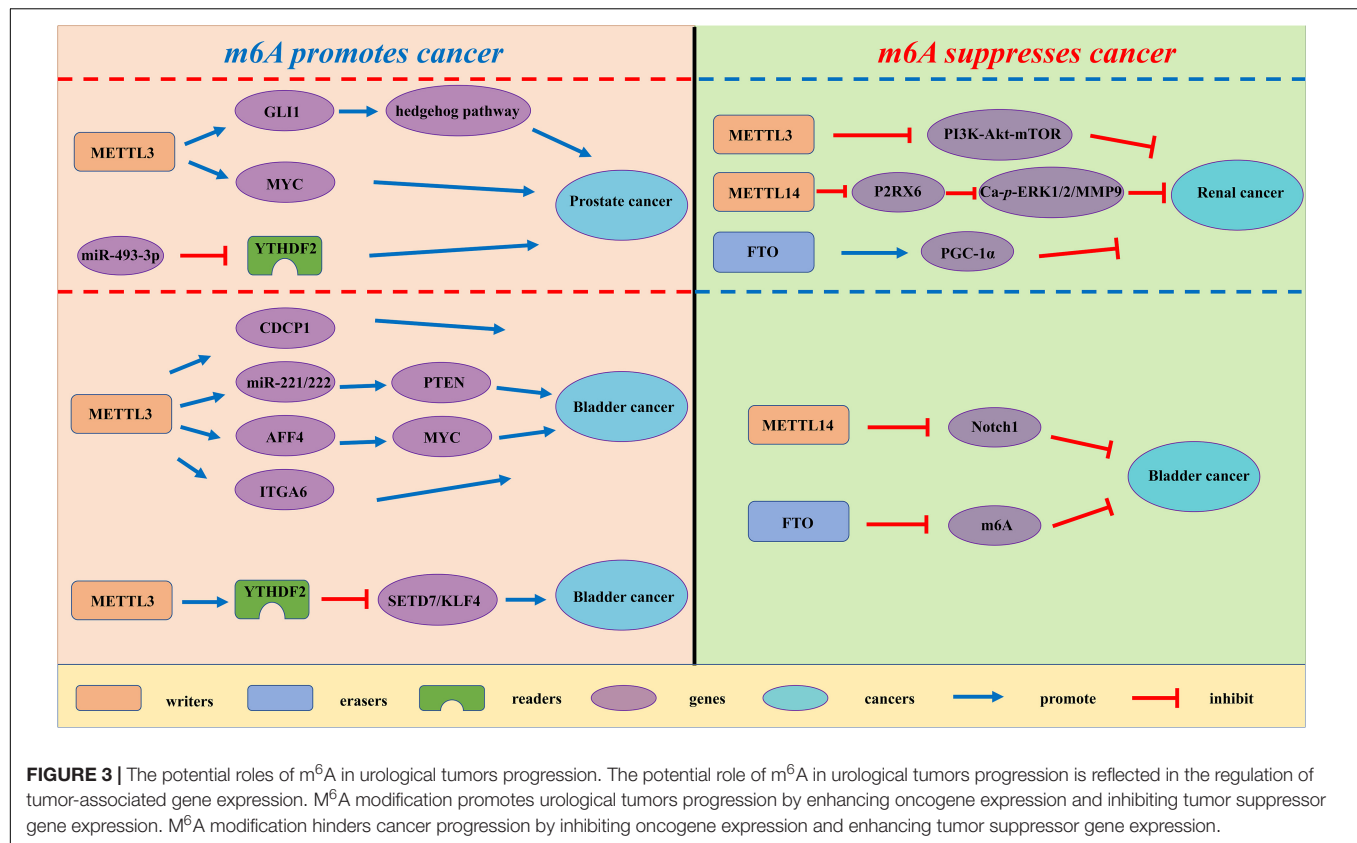
Cancer	m ⁶ A Regulators	Role in cancer	Biological function	Mechanism	References
Renal cancer	METTL3	Suppressor gene	Suppresses RCC proliferation, migration and invasion	Regulates EMT and PI3K-Akt-mTOR pathways	Rini et al., 2009
–	METTL14	Suppressor gene	Suppresses RCC migration and invasion	Down-regulates P2RX6 protein translation	Li X. et al., 2017
–	–	–	–	–	–
–	–	–	–	–	–
–	FTO	Suppressor gene	Suppresses RCC growth	Promotes PGC-1 α expression by reducing m ⁶ A levels	Gong et al., 2019
Prostate cancer	METTL3	Oncogene	Promotes PCa growth and metastasis	Regulates hedgehog pathway	Siegel et al., 2020
–	–	–	–	–	–
–	METTL3	Oncogene	Promotes PCa proliferation, migration and invasion	Promotes MYC expression by increasing m ⁶ A levels	Cai et al., 2019
–	YTHDF2	Oncogene	Promotes PCa proliferation and migration	/	Yuan et al., 2020
–	–	–	–	–	–
Bladder cancer	METTL3	Oncogene	Promotes BC growth	Promotes CDCP1 mRNA modification and translation	Bray et al., 2018
–	METTL3	Oncogene	Promotes BC proliferation	Interacts with the microprocessor protein DGCR8 and positively modulates the pri-miR221/222 process	Yang F. et al., 2019
–	–	–	–	–	–
–	METTL3	Oncogene	Promotes BC growth and metastasis	Regulates AFF4/NF- κ B/MYC signaling network	Han et al., 2019
–	METTL3	Oncogene	Promotes BC growth and metastasis	promote the translation of ITGA6 mRNA	Cheng et al., 2019
–	–	–	–	–	–
–	METTL3/YTHDF2	Oncogene	Promotes BC growth and metastasis	METTL3/YTHDF2 may mediate the mRNA decay of tumor suppressors SETD7 and KLF4	Jin H. et al., 2019
–	METTL14	Suppressor gene	Promotes the proliferation, self-renewal, metastasis and tumor initiating capacity of bladder TICs	METTL14 knockdown may enhance the RNA stability of Notch1 mRNA	Xie et al., 2020
–	–	–	–	–	–
–	–	–	–	–	–
–	FTO	Suppressor Gene	Inhibits BC proliferation and migration	/	Gu et al., 2019

further regulate the Ca²⁺-mediated p-ERK1/2/MMP9 signal pathway promote cell migration and invasion. Zhuang C. et al. (2019) found that PGC-1 α underwent m⁶A methylation in RCC. As an m⁶A demethylase, FTO could recognize the m⁶A sites of PGC-1 α and reduce its methylation level, therefore leading to the increases in mitochondria biogenesis and oxidative phosphorylation and the decreases in tumor growth of RCC (Zhuang C. et al., 2019).

Prostate Cancer

Prostate cancer (PCa) has been regarded as the most common cancer among men and the second cancer-related deaths in

the men in 2019 (Siegel et al., 2020). Despite recent advances in many therapies, the 5 years' survival rate for prostate cancer patients remains low. Cai et al. found that METTL3 is overexpressed in PCa tissues and cell lines (Cai et al., 2019). Elevated expression of METTL3 could promote cell proliferation, survival, colony formation, and invasion. Moreover, knockdown of METTL3 could decrease the m⁶A modification and expression of GLI1, thereby regulating hedgehog pathway. Yuan et al. (2020) demonstrated that the mRNA expression level of METTL3 was increased in prostate cancer tissues. Additionally, the expression level of METTL3 is associated with the deterioration of PCa patients' condition. Mechanistically, METTL3 could enhance



MYC (c-myc) expression via elevating m⁶A levels of MYC mRNA transcript, so as to facilitate the proliferative, migrative and invasive ability of cancer cells. Li found that YTHDF2, an m⁶A reader, was upregulated in prostate cancer tissues and cell lines (Li J. et al., 2018). Knockdown of YTHDF2 led to decreased levels of m⁶A and impaired proliferation and migration of PCa cells. Therefore, YTHDF2 played a vital role in the initiation and progression of PCa.

Bladder Cancer

Bladder cancer (BCa) is the most common urogenital and the 10th most common cancer worldwide, with an estimated 549 000 new cases and 200 000 deaths in 2018 (Bray et al., 2018). Despite the improvement of clinical diagnosis and therapies, BCa is regarded as a major cause of cancer-interrelated morbidity and mortality. In the study of Yang F. et al. (2019) the expression levels of METTL3 were elevated in BCa patient samples. The increase in METTL3 expression was proven to be correlated with BCa growth and progression *in vitro* and *in vivo*. Moreover, METTL3 could positively regulate CDCP1 process based on an m⁶A -dependent mode, bringing about elevated expression of CDCP1. Han et al. (2019) demonstrated that the expression level of METTL3 in BCa was significantly up-regulated and associated with poor prognosis of BCa patients. They found that METTL3 might interact with the microprocessor protein DGCR8 and positively modulate the pri-miR221/222 process through an m⁶A -dependent mechanism. Cheng and coworkers

elucidated that METTL3 was obviously up-regulated in BCa tissues and significantly promoted growth and metastasis of BCa (Cheng et al., 2019). Mechanistically, METTL3 might promote BCa progression via AFF4/NF-κB/MYC signaling pathway. Jin and coworkers demonstrated that METTL3 and ALKBH5 can alter cell adhesion via regulating ITGA6 expression in BCa (Jin H. et al., 2019). Increased m⁶A methylation enhanced the translation of ITGA6 mRNA by binding of YTHDF1 and YTHDF3 and promoted malignant phenotypes in BCa. Xie and coworkers found that knockout of METTL3 impaired tumor growth and metastasis, METTL3/YTHDF2 m⁶A axis could directly degrade the mRNA expression of the tumor suppressors SETD7 and KLF4, leading to the development and progression of BCa (Xie et al., 2020). Gu et al. (2019) found a decrease of N6-methyladenosine in BCa and bladder tumor initiating cells (TICs). In addition, METTL14 is down-regulated in BCa and bladder TICs and it could promote the proliferation, metastasis, self-renewal and enhance tumor initiating capacity of bladder TICs. Mechanistically, METTL14 might regulate Notch1 expression in an m⁶A-dependent manner. Wen demonstrated that knockdown of FTO could accelerate the progression of BCa (Wen et al., 2020), while the potential mechanism remains unknown.

Testicular Germ Cell Tumors

Testicular germ cell tumors (TGCTs) are the most common solid neoplasm among men aged between 14 and 44 years

(Cheng et al., 2018). Despite the advanced prognosis of localized TGCTs, approximately 20–30% of patients may experience disease recurrence during surveillance (Mortensen et al., 2016). Lobo and coworkers demonstrated that abundance of m⁶A and expression of VIRMA/YTHDF3 were different among TGCTs subtypes, with higher levels in seminomas (SEs), suggesting a contribution to SE phenotype maintenance (Lobo et al., 2019). However, the potential biological roles of VIRMA/YTHDF3 remain to be further explored.

Wilms Tumor

Wilms tumor (WT) is the most prevalent childhood kidney tumor characterized by the disorganized and dysregulated development of a kidney (Davidoff, 2009; Servaes et al., 2019). Hua et al. (2020) found an obvious relationship between ALKBH5 rs1378602 AG/AA genotypes and decreased Wilms tumor risk in children in clinical stage I diseases. However, the observed association should be further validated in another well-designed analysis with other larger ethnicities.

POTENTIAL APPLICATION OF RNA m⁶A IN UROLOGICAL TUMORS RNA

RNA m⁶A as Biomarker in Urological Tumors

Mounting evidence has indicated that m⁶A regulators have the potential to be superior diagnostic and prognostic biomarkers for urological tumors patients. Strick et al. conducted qRT-PCR to detect the gene expressions of ALKBH5 and FTO were studied in 166 ccRCC and 106 normal renal tissues. They found that the expression level of ALKBH5 and FTO were obviously decreased in ccRCC tissues (Strick et al., 2020). Declined mRNA levels of ALKBH5 and FTO were related to a shortened overall and cancer-specific survival following nephrectomy. Therefore, ALKBH5 and FTO could be used as prognostic biomarkers for RCC. Zhao Y. et al. (2020) demonstrated that METTL14 mRNA expression negatively correlated with the RCC stages and positively correlated with RCC patients' overall survival, it might be a potential biomarker of RCC. Yuan et al. performed the qRT-PCR to detect the mRNA expression level of METTL3 in 84 clinical human PCa specimens and 32 corresponding adjacent normal specimens. The results showed that a significant positive association between METTL3 expression was observed with tumor stage and metastasis. Moreover, the expression level of METTL3 had remarkable prognostic value for overall survival and disease-free survival (Yuan et al., 2020); hence, METTL3 might play a vital role in PCa progression and metastasis. Chen et al. concluded that m⁶A regulators were related to malignant clinicopathological features of BCa and a risk signature with FTO, WTAP and YTHDC3 might play vital roles in diagnosis and prognosis of BCa patients (Chen M. et al., 2019). In TGCTs, VIRMA and YTHDF3 might be prognostic factors (Lobo et al., 2019).

RNA m⁶A as Therapeutic Targets in Urological Tumors

The critical roles of m⁶A in urological tumors suggest that it has the potential to be involved in tumor therapy. A number of studies have indicated that m⁶A modification is significant in therapies of urological tumors, especially in targeted treatment. Zhuang et al. found that the Von Hippel-Lindau (VHL) -deficient cells expressing FTO might restore mitochondrial activity, induce oxidative stress and ROS production and suppressed tumor growth, via promoting PGC-1 α expression by decreasing m⁶A levels in its mRNA transcripts (Zhuang C. et al., 2019). Therefore, the m⁶A methylation and m⁶A-related regulators, and uncovers an essential FTO-PGC-1 α axis might play a vital role in the treatment of RCC. Gong and coworkers found that ATP could enhance cell migration and invasion via regulating P2RX6 expression in RCC (Gong et al., 2019). Mechanistically, ATP-P2RX6 could modulate the Ca²⁺-mediated p-ERK1/2/MMP9 signal pathway, while METTL14 might down-regulate P2RX6 protein translation in an m⁶A-dependent manner. Herein, the further exploration of regulation of METTL14 expression might contribute to develop a new approach to repress RCC progression. Li et al. suggested that METTL3 expression is higher in PCa than in normal prostate tissues, especially in PCa with bone metastasis (Li E. et al., 2020). METTL3 regulates the expression of Integrin β 1 (ITGB1) through m⁶A-HuR-dependent mechanism, which affects the binding of ITGB1 to Collagen I and tumor cell motility, so as to promote the bone metastasis of PCa. Therefore, METTL3 might act as a therapeutic target for PCa bone metastasis. Wen et al. found that knockdown of FTO could enhance cell proliferation and migration and protect BCa cells from cisplatin-induced cytotoxicity (Wen et al., 2020). Hence, targeting the m⁶A modification of FTO may be beneficial to the treatment of BCa.

DISCUSSION

Recently, RNA epigenetics is emerging as a hot topic. Among them, m⁶A modification has become a new layer of post-transcriptional regulation of gene expression. The implications of m⁶A modifications in human carcinogenesis have been verified in many kinds of cancers, including urological tumors. In this review, we summarized the potential biological effects of m⁶A-related regulators, and particularly focused on the impacts of m⁶A modification on different tumors in the urinary system. M⁶A can be installed by the methyltransferase, while these modifications may be removed by m⁶A eraser demethylases. Furthermore, m⁶A readers could specifically recognize the m⁶A methylation sites and thus regulating mRNA splicing, translation, degradation, nuclear export, and other cellular processes. Besides, m⁶A methylation and its related regulatory factors are reported to be involved in the processing and the biological function of non-coding coding RNAs (Coker et al., 2019; Huang H. et al., 2020).

However, m⁶A methylation seems to serve as a double-edged sword due to the specific mechanism for m⁶A in cancers remains unknown. Some genes may lead to cancer progression after m⁶A

methylation, while removal of m⁶A modification can result in the progression of other tumors. For example, in HCC, sumo1 modification of METTL3 can promote tumor progression via regulating snail mRNA homeostasis (Xu et al., 2020), while in glioblastoma, lncRNA SOX2OT can facilitate temozolomide resistance through promoting SOX2 expression via ALKBH5-mediated epigenetic regulation (Liu X. et al., 2020). In addition, the same m⁶A-associated regulator may play crucial roles in the same type of cancer via targeting different downstream genes. For instance, in CRC, METTL3 can promote tumor progression through enhancing the expression of either MYC (Xiang et al., 2020) or CCNE1 (Zhu W. et al., 2020). Additionally, researches have reported conflicting findings in the same type of cancer; for instance, in CRC, METTL3 and METTL14 play totally opposite roles in tumor initiation and progression (Li T. et al., 2019; Yang X. et al., 2020). Overall, all above studies show that m⁶A methylation and its related regulatory networks are complex and need to be further explored. Moreover, Han et al. found that METTL3 can enhance tumor growth of BCa through accelerating pri-miR221/222 maturation based on m⁶A-dependent mode (Han et al., 2019), while Gu et al. (2019) reported that METTL14 can inhibit bladder tumorigenesis through N⁶-methyladenosine of Notch1. The above discrepancy may result from several factors such as case sample size and different related regulatory genes. Furthermore, studies have identified the therapeutic potential of m⁶A modification. METTL3 might induce NSCLC drug resistance and metastasis via modulating the MALAT1-miR-1914-3p-YAP axis (Jin D. et al., 2019). In glioma, METTL3 can promote glioma radioresistance and stem-like cell maintenance (Visvanathan et al., 2018). In melanoma, FTO can act as an m⁶A demethylase to promote melanoma tumorigenesis and anti-PD-1 resistance (Yang S. et al., 2019). R-2HG can inhibit FTO activity and thus elevating m⁶A mRNA modification in R-2HG-sensitive leukemia cells, thereby generating anti-leukemia effects (Su et al., 2018). In cervical squamous cell carcinoma (CSCC), FTO can regulate the chemo-radiotherapy resistance by targeting β -catenin through mRNA demethylation (Zhou S. et al., 2018).

The advanced development of m⁶A modification study marks a novel insight in the diagnosis and therapy of various diseases. Nevertheless, we believe that future prospects on m⁶A modification need to be further explored. Firstly,

several databases (such as GEPIA, TCGA et al.) were used in many studies to explore the prognostic significance of m⁶A regulators expression in OS and DFS of urological tumors patients. Hence, expansion of the sample size and screening factors are essential for early diagnosis and prognosis; while the specificity and sensitivity of m⁶A-related regulators also need to be discussed. Secondly, more and more clinical practice are urgent for confirming the therapeutic potential of m⁶A regulatory factors and related pathways. Thirdly, it's significant to construct a complex and specific regulatory network model of m⁶A and its associated modifiers in a single cancer. Fourthly, exploring other components of m⁶A methylation and demethylation and effectors is necessary.

CONCLUSION

Urological tumors are major public health concern with growing prevalence. Studies have showed that m⁶A methylation plays a significant role in prevention, treatment and management of various urological tumors; however, more endeavors and more multi-center and large-scale research are urgent for exploring the relationship between m⁶A modification and urological tumors.

AUTHOR CONTRIBUTIONS

YL and YG collected the related manuscript and finished the manuscript and figures. RJ gave constructive guidance and made final approval. LX, ZX, and QD participated in the design of this review. All authors read and approved the final manuscript.

FUNDING

This work was supported by grants from the National Natural Science Foundation of China (NSFC) (Grant Numbers: 81570613, 81370853, and 81802531), Jiangsu Provincial Social Development Project (BE2017615), and Jiangsu Provincial Medical Innovation Team (2016).

REFERENCES

- Alarcon, C. R., Goodarzi, H., Lee, H., Liu, X., Tavazoie, S., and Tavazoie, S. F. (2015). HNRNPA2B1 is a mediator of m(6)A-dependent nuclear RNA processing events. *Cell* 162, 1299–1308. doi: 10.1016/j.cell.2015.08.011
- Arguello, A. E., Leach, R. W., and Kleiner, R. E. (2019). In vitro selection with a site-specifically modified RNA library reveals the binding preferences of N(6)-methyladenosine reader proteins. *Biochemistry* 58, 3386–3395. doi: 10.1021/acs.biochem.9b00485
- Bailey, A. S., Batista, P. J., Gold, R. S., Chen, Y. G., de Rooij, D. G., Chang, H. Y., et al. (2017). The conserved RNA helicase YTHDC2 regulates the transition from proliferation to differentiation in the germline. *eLife* 6:e26116. doi: 10.7554/eLife.26116
- Barbieri, I., Tzelepis, K., Pandolfi, L., Shi, J., Millan-Zambrano, G., Robson, S. C., et al. (2017). Promoter-bound METTL3 maintains myeloid leukaemia by m(6)A-dependent translation control. *Nature* 552, 126–131. doi: 10.1038/nature24678
- Basak, E. A., Koolen, S. L. W., Hurkmans, D. P., Schreurs, M. W. J., Bins, S., Oomen-de Hoop, E., et al. (2019). Correlation between nivolumab exposure and treatment outcomes in non-small-cell lung cancer. *Eur. J. Cancer* 109, 12–20. doi: 10.1016/j.ejca.2018.12.008
- Batista, P. J. (2017). The RNA modification N(6)-methyladenosine and its implications in human disease. *Genomics Proteomics Bioinformatics* 15, 154–163. doi: 10.1016/j.gpb.2017.03.002
- Bokar, J. A., Shambaugh, M. E., Polayes, D., Matera, A. G., and Rottman, F. M. (1997). Purification and cDNA cloning of the AdoMet-binding subunit of the human mRNA (N⁶-adenosine)-methyltransferase. *RNA* 3, 1233–1247.
- Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R. L., Torre, L. A., and Jemal, A. (2018). Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* 68, 394–424. doi: 10.3322/caac.21492

- Brown, J. A., Kinzig, C. G., DeGregorio, S. J., and Steitz, J. A. (2016). Methyltransferase-like protein 16 binds the 3'-terminal triple helix of MALAT1 long noncoding RNA. *Proc. Natl. Acad. Sci. U.S.A.* 113, 14013–14018. doi: 10.1073/pnas.1614759113
- Cai, J., Yang, F., Zhan, H., Situ, J., Li, W., Mao, Y., et al. (2019). RNA m(6)A Methyltransferase METTL3 promotes the growth of prostate cancer by regulating hedgehog pathway. *Onco Targets Ther.* 12, 9143–9152. doi: 10.2147/OTT.S226796
- Chandola, U., Das, R., and Panda, B. (2015). Role of the N6-methyladenosine RNA mark in gene regulation and its implications on development and disease. *Brief. Funct. Genomics* 14, 169–179. doi: 10.1093/bfpg/elu039
- Chen, M., Nie, Z. Y., Wen, X. H., Gao, Y. H., Cao, H., and Zhang, S. F. (2019). m6A RNA methylation regulators can contribute to malignant progression and impact the prognosis of bladder cancer. *Biosci. Rep.* 39:BSR20192892. doi: 10.1042/BSR20192892
- Chen, X., Xu, M., Xu, X., Zeng, K., Liu, X., Sun, L., et al. (2020). METTL14 suppresses CRC progression via regulating N6-methyladenosine-dependent primary miR-375 processing. *Mol. Ther.* 28, 599–612. doi: 10.1016/j.ymthe.2019.11.016
- Chen, X. Y., Zhang, J., and Zhu, J. S. (2019). The role of m(6)A RNA methylation in human cancer. *Mol. Cancer* 18:103. doi: 10.1186/s12943-019-1033-z
- Chen, Y., Zhao, Y., Chen, J., Peng, C., Zhang, Y., Tong, R., et al. (2020). ALKBH5 suppresses malignancy of hepatocellular carcinoma via m(6)A-guided epigenetic inhibition of LYPD1. *Mol. Cancer* 19:123. doi: 10.1186/s12943-020-01239-w
- Cheng, L., Albers, P., Berney, D. M., Feldman, D. R., Daugaard, G., Gilligan, T., et al. (2018). Testicular cancer. *Nat. Rev. Dis. Primers* 4:29. doi: 10.1038/s41572-018-0029-0
- Cheng, M., Sheng, L., Gao, Q., Xiong, Q., Zhang, H., Wu, M., et al. (2019). The m(6)A methyltransferase METTL3 promotes bladder cancer progression via AFF4/NF-kappaB/MYC signaling network. *Oncogene* 38, 3667–3680. doi: 10.1038/s41388-019-0683-z
- Coker, H., Wei, G., and Brockdorff, N. (2019). m6A modification of non-coding RNA and the control of mammalian gene expression. *Biochim. Biophys. Acta Gene Regul. Mech.* 1862, 310–318. doi: 10.1016/j.bbagr.2018.12.002
- Dai, D., Wang, H., Zhu, L., Jin, H., and Wang, X. (2018). N6-methyladenosine links RNA metabolism to cancer progression. *Cell Death Dis.* 9:124. doi: 10.1038/s41419-017-0129-x
- Davidoff, A. M. (2009). Wilms' tumor. *Curr. Opin. Pediatr.* 21, 357–364. doi: 10.1097/MOP.0b013e32832b323a
- Deng, X., Chen, K., Luo, G. Z., Weng, X., Ji, Q., Zhou, T., et al. (2015). Widespread occurrence of N6-methyladenosine in bacterial mRNA. *Nucleic Acids Res.* 43, 6557–6567. doi: 10.1093/nar/gkv596
- Deng, X., Su, R., Weng, H., Huang, H., Li, Z., and Chen, J. (2018). RNA N(6)-methyladenosine modification in cancers: current status and perspectives. *Cell Res.* 28, 507–517. doi: 10.1038/s41422-018-0034-6
- Desrosiers, R., Friderici, K., and Rottman, F. (1974). Identification of methylated nucleosides in messenger RNA from Novikoff hepatoma cells. *Proc. Natl. Acad. Sci. U.S.A.* 71, 3971–3975. doi: 10.1073/pnas.71.10.3971
- Dominissini, D. (2014). Genomics and proteomics. Roadmap to the epitranscriptome. *Science* 346:1192. doi: 10.1126/science.aaa1807
- Dominissini, D., Moshitch-Moshkovitz, S., Schwartz, S., Salmon-Divon, M., Ungar, L., Osenberg, S., et al. (2012). Topology of the human and mouse m6A RNA methylomes revealed by m6A-seq. *Nature* 485, 201–206. doi: 10.1038/nature11112
- Doxtader, K. A., Wang, P., Scarborough, A. M., Seo, D., Conrad, N. K., and Nam, Y. (2018). Structural basis for regulation of METTL16, an S-Adenosylmethionine homeostasis factor. *Mol. Cell* 71, 1001–1011.e4. doi: 10.1016/j.molcel.2018.07.025
- Fitzsimmons, C. M., and Batista, P. J. (2019). It's complicated. m(6)A-dependent regulation of gene expression in cancer. *Biochim. Biophys. Acta Gene Regul. Mech.* 1862, 382–393. doi: 10.1016/j.bbagr.2018.09.010
- Fu, Y., Luo, G. Z., Chen, K., Deng, X., Yu, M., Han, D., et al. (2015). N6-methyldeoxyadenosine marks active transcription start sites in *Chlamydomonas*. *Cell* 161, 879–892. doi: 10.1016/j.cell.2015.04.010
- Gong, D., Zhang, J., Chen, Y., Xu, Y., Ma, J., Hu, G., et al. (2019). The m(6)A-suppressed P2RX6 activation promotes renal cancer cells migration and invasion through ATP-induced Ca(2+) influx modulating ERK1/2 phosphorylation and MMP9 signaling pathway. *J. Exp. Clin. Cancer Res.* 38:233. doi: 10.1186/s13046-019-1223-y
- Greer, E. L., Blanco, M. A., Gu, L., Sendinc, E., Liu, J., Aristizabal-Corrales, D., et al. (2015). DNA Methylation on N6-Adenine in *C. elegans*. *Cell* 161, 868–878. doi: 10.1016/j.cell.2015.04.005
- Gu, C., Wang, Z., Zhou, N., Li, G., Kou, Y., Luo, Y., et al. (2019). Mettl14 inhibits bladder TIC self-renewal and bladder tumorigenesis through N(6)-methyladenosine of Notch1. *Mol. Cancer* 18:168. doi: 10.1186/s12943-019-1084-1
- Guichard, C., Amaddeo, G., Imbeaud, S., Ladeiro, Y., Pelletier, L., Maad, I. B., et al. (2012). Integrated analysis of somatic mutations and focal copy-number changes identifies key genes and pathways in hepatocellular carcinoma. *Nat. Genet.* 44, 694–698. doi: 10.1038/ng.2256
- Hammerle, M., Gutschner, T., Uckelmann, H., Ozgur, S., Fiskin, E., Gross, M., et al. (2013). Posttranscriptional destabilization of the liver-specific long noncoding RNA HULC by the IGF2 mRNA-binding protein 1 (IGF2BP1). *Hepatology* 58, 1703–1712. doi: 10.1002/hep.26537
- Han, J., Wang, J. Z., Yang, X., Yu, H., Zhou, R., Lu, H. C., et al. (2019). METTL3 promote tumor proliferation of bladder cancer by accelerating pri-miR221/222 maturation in m6A-dependent manner. *Mol. Cancer* 18:110. doi: 10.1186/s12943-019-1036-9
- Hou, J., Zhang, H., Liu, J., Zhao, Z., Wang, J., Lu, Z., et al. (2019). YTHDF2 reduction fuels inflammation and vascular abnormalization in hepatocellular carcinoma. *Mol. Cancer* 18:163. doi: 10.1186/s12943-019-1082-3
- Hsu, P. J., Shi, H., and He, C. (2017). Epitranscriptomic influences on development and disease. *Genome Biol.* 18:197. doi: 10.1186/s13059-017-1336-6
- Hu, X., Peng, W. X., Zhou, H., Jiang, J., Zhou, X., Huang, D., et al. (2020). IGF2BP2 regulates DANCER by serving as an N6-methyladenosine reader. *Cell Death Differ.* 27, 1782–1794. doi: 10.1038/s41418-019-0461-z
- Hua, R. X., Liu, J., Fu, W., Zhu, J., Zhang, J., Cheng, J., et al. (2020). ALKBH5 gene polymorphisms and Wilms tumor risk in Chinese children: a five-center case-control study. *J. Clin. Lab. Anal.* 34:e23251. doi: 10.1002/jcla.23251
- Huang, H., Gao, Y., Liu, A., Yang, X., Huang, F., Xu, L., et al. (2019). EIF3D promotes sunitinib resistance of renal cell carcinoma by interacting with GRP78 and inhibiting its degradation. *EBioMedicine* 49, 189–201. doi: 10.1016/j.ebiom.2019.10.030
- Huang, H., Weng, H., and Chen, J. (2020). m(6)A modification in coding and non-coding RNAs: roles and therapeutic implications in cancer. *Cancer Cell* 37, 270–288. doi: 10.1016/j.ccell.2020.02.004
- Huang, H., Weng, H., Sun, W., Qin, X., Shi, H., Wu, H., et al. (2018). Recognition of RNA N(6)-methyladenosine by IGF2BP proteins enhances mRNA stability and translation. *Nat. Cell Biol.* 20, 285–295. doi: 10.1038/s41556-018-0045-z
- Huang, T., Liu, Z., Zheng, Y., Feng, T., Gao, Q., and Zeng, W. (2020). YTHDF2 promotes spermatogenesis through modulating MMPs decay via m(6)A/mRNA pathway. *Cell Death Dis.* 11:37. doi: 10.1038/s41419-020-2235-4
- Jia, G., Fu, Y., Zhao, X., Dai, Q., Zheng, G., Yang, Y., et al. (2011). N6-methyladenosine in nuclear RNA is a major substrate of the obesity-associated FTO. *Nat. Chem. Biol.* 7, 885–887. doi: 10.1038/nchembio.687
- Jin, D., Guo, J., Wu, Y., Du, J., Yang, L., Wang, X., et al. (2019). m(6)A mRNA methylation initiated by METTL3 directly promotes YAP translation and increases YAP activity by regulating the MALAT1-miR-1914-3p-YAP axis to induce NSCLC drug resistance and metastasis. *J. Hematol. Oncol.* 12:135. doi: 10.1186/s13045-019-0830-6
- Jin, D., Guo, J., Wu, Y., Yang, L., Wang, X., Du, J., et al. (2020). m(6)A demethylase ALKBH5 inhibits tumor growth and metastasis by reducing YTHDFs-mediated YAP expression and inhibiting miR-107/LATS2-mediated YAP activity in NSCLC. *Mol. Cancer* 19:40. doi: 10.1186/s12943-020-01161-1
- Jin, H., Ying, X., Que, B., Wang, X., Chao, Y., Zhang, H., et al. (2019). N(6)-methyladenosine modification of ITGA6 mRNA promotes the development and progression of bladder cancer. *EBioMedicine* 47, 195–207. doi: 10.1016/j.ebiom.2019.07.068
- Kasowitz, S. D., Ma, J., Anderson, S. J., Leu, N. A., Xu, Y., Gregory, B. D., et al. (2018). Nuclear m6A reader YTHDC1 regulates alternative polyadenylation and splicing during mouse oocyte development. *PLoS Genet.* 14:e1007412. doi: 10.1371/journal.pgen.1007412
- Kong, F., Liu, X., Zhou, Y., Hou, X., He, J., Li, Q., et al. (2020). Downregulation of METTL14 increases apoptosis and autophagy induced by cisplatin in pancreatic

- cancer cells. *Int. J. Biochem. Cell Biol.* 122:105731. doi: 10.1016/j.biocel.2020.105731
- Lan, Q., Liu, P. Y., Haase, J., Bell, J. L., Huttelmaier, S., and Liu, T. (2019). The Critical Role of RNA m(6)A Methylation in Cancer. *Cancer Res.* 79, 1285–1292. doi: 10.1158/0008-5472.CAN-18-2965
- Lang, F., Singh, R. K., Pei, Y., Zhang, S., Sun, K., and Robertson, E. S. (2019). EBV epitranscriptome reprogramming by METTL14 is critical for viral-associated tumorigenesis. *PLoS Pathog.* 15:e1007796. doi: 10.1371/journal.ppat.1007796
- Lee, A. S., Kranzusch, P. J., Doudna, J. A., and Cate, J. H. (2016). eIF3d is an mRNA cap-binding protein that is required for specialized translation initiation. *Nature* 536, 96–99. doi: 10.1038/nature18954
- Li, A., Chen, Y. S., Ping, X. L., Yang, X., Xiao, W., Yang, Y., et al. (2017). Cytoplasmic m(6)A reader YTHDF3 promotes mRNA translation. *Cell Res.* 27, 444–447. doi: 10.1038/cr.2017.10
- Li, E., Wei, B., Wang, X., and Kang, R. (2020). METTL3 enhances cell adhesion through stabilizing integrin beta1 mRNA via an m6A-HuR-dependent mechanism in prostatic carcinoma. *Am. J. Cancer Res.* 10, 1012–1025.
- Li, J., Meng, S., Xu, M., Wang, S., He, L., Xu, X., et al. (2018). Downregulation of N(6)-methyladenosine binding YTHDF2 protein mediated by miR-493-3p suppresses prostate cancer by elevating N(6)-methyladenosine levels. *Oncotarget* 9, 3752–3764. doi: 10.18632/oncotarget.23365
- Li, L. J., Fan, Y. G., Leng, R. X., Pan, H. F., and Ye, D. Q. (2018). Potential link between m(6)A modification and systemic lupus erythematosus. *Mol. Immunol.* 93, 55–63. doi: 10.1016/j.molimm.2017.11.009
- Li, T., Hu, P. S., Zuo, Z., Lin, J. F., Li, X., Wu, Q. N., et al. (2019). METTL3 facilitates tumor progression via an m(6)A-IGF2BP2-dependent mechanism in colorectal carcinoma. *Mol. Cancer* 18:112. doi: 10.1186/s12943-019-1038-7
- Li, X., Tang, J., Huang, W., Wang, F., Li, P., Qin, C., et al. (2017). The M6A methyltransferase METTL3: acting as a tumor suppressor in renal cell carcinoma. *Oncotarget* 8, 96103–96116. doi: 10.18632/oncotarget.21726
- Li, X., Xiong, X., and Yi, C. (2016). Epitranscriptome sequencing technologies: decoding RNA modifications. *Nat. Methods* 14, 23–31. doi: 10.1038/nmeth.4110
- Li, X. C., Jin, F., Wang, B. Y., Yin, X. J., Hong, W., and Tian, F. J. (2019). The m6A demethylase ALKBH5 controls trophoblast invasion at the maternal-fetal interface by regulating the stability of CYR61 mRNA. *Theranostics* 9, 3853–3865. doi: 10.7150/thno.31868
- Liu, B., Zhou, J., Wang, C., Chi, Y., Wei, Q., Fu, Z., et al. (2020). LncRNA SOX2OT promotes temozolomide resistance by elevating SOX2 expression via ALKBH5-mediated epigenetic regulation in glioblastoma. *Cell Death Dis.* 11:384. doi: 10.1038/s41419-020-2540-y
- Liu, E. Y., Cali, C. P., and Lee, E. B. (2017). RNA metabolism in neurodegenerative disease. *Dis. Model. Mech.* 10, 509–518. doi: 10.1242/dmm.028613
- Liu, J., Li, K., Cai, J., Zhang, M., Zhang, X., Xiong, X., et al. (2020). Landscape and regulation of m(6)A and m(6)Am Methylome across human and mouse tissues. *Mol. Cell* 77, 426–440.e6. doi: 10.1016/j.molcel.2019.09.032
- Liu, J., Yue, Y., Han, D., Wang, X., Fu, Y., Zhang, L., et al. (2014). A METTL3-METTL14 complex mediates mammalian nuclear RNA N6-adenosine methylation. *Nat. Chem. Biol.* 10, 93–95. doi: 10.1038/nchembio.1432
- Liu, J., Zhu, Y., Luo, G. Z., Wang, X., Yue, Y., Wang, X., et al. (2016). Abundant DNA 6mA methylation during early embryogenesis of zebrafish and pig. *Nat. Commun.* 7:13052. doi: 10.1038/ncomms13052
- Liu, T., Wei, Q., Jin, J., Luo, Q., Liu, Y., Yang, Y., et al. (2020). The m6A reader YTHDF1 promotes ovarian cancer progression via augmenting EIF3C translation. *Nucleic Acids Res.* 48, 3816–3831. doi: 10.1093/nar/gkaa048
- Liu, X., Liu, J., Xiao, W., Zeng, Q., Bo, H., Zhu, Y., et al. (2020). SIRT1 regulates N(6)-methyladenosine RNA modification in hepatocarcinogenesis by inducing RANBP2-dependent FTO SUMOylation. *Hepatology*. doi: 10.1002/hep.31222 [Epub ahead of print].
- Lobo, J., Costa, A. L., Cantante, M., Guimaraes, R., Lopes, P., Antunes, L., et al. (2019). m(6)A RNA modification and its writer/reader VIRMA/YTHDF3 in testicular germ cell tumors: a role in seminoma phenotype maintenance. *J. Transl. Med.* 17:79. doi: 10.1186/s12967-019-1837-z
- Ma, F., Liu, X., Zhou, S., Li, W., Liu, C., Chadwick, M., et al. (2019). Long non-coding RNA FGF13-AS1 inhibits glycolysis and stemness properties of breast cancer cells through FGF13-AS1/IGF2BPs/Myc feedback loop. *Cancer Lett.* 450, 63–75. doi: 10.1016/j.canlet.2019.02.008
- Mao, Y., Dong, L., Liu, X. M., Guo, J., Ma, H., Shen, B., et al. (2019). m(6)A in mRNA coding regions promotes translation via the RNA helicase-containing YTHDC2. *Nat. Commun.* 10:5332. doi: 10.1038/s41467-019-13317-9
- McGee, S. L., and Hargreaves, M. (2019). Epigenetics and exercise. *Trends Endocrinol. Metab.* 30, 636–645. doi: 10.1016/j.tem.2019.06.002
- Mendel, M., Chen, K. M., Homolka, D., Gos, P., Pandey, R. R., McCarthy, A. A., et al. (2018). Methylation of structured RNA by the m(6)A Writer METTL16 is essential for mouse embryonic development. *Mol. Cell* 71, 986–1000.e11. doi: 10.1016/j.molcel.2018.08.004
- Meyer, K. D., Patil, D. P., Zhou, J., Zinoviev, A., Skabkin, M. A., Elemento, O., et al. (2015). 5' UTR m(6)A promotes cap-independent translation. *Cell* 163, 999–1010. doi: 10.1016/j.cell.2015.10.012
- Meyer, K. D., Saletore, Y., Zumbo, P., Elemento, O., Mason, C. E., and Jaffrey, S. R. (2012). Comprehensive analysis of mRNA methylation reveals enrichment in 3' UTRs and near stop codons. *Cell* 149, 1635–1646. doi: 10.1016/j.cell.2012.05.003
- Mohammad, H. P., Barbash, O., and Creasy, C. L. (2019). Targeting epigenetic modifications in cancer therapy: erasing the roadmap to cancer. *Nat. Med.* 25, 403–418. doi: 10.1038/s41591-019-0376-8
- Moindrot, B., Cerase, A., Coker, H., Masui, O., Grijzenhout, A., Pintacuda, G., et al. (2015). A Pooled shRNA Screen Identifies Rbm15, Spen, and Wtap as Factors Required for Xist RNA-Mediated Silencing. *Cell Rep.* 12, 562–572. doi: 10.1016/j.celrep.2015.06.053
- Mortensen, M. S., Lauritsen, J., Kier, M. G., Bandak, M., Appelt, A. L., Agerbaek, M., et al. (2016). Late relapses in stage I testicular cancer patients on surveillance. *Eur. Urol.* 70, 365–371. doi: 10.1016/j.eururo.2016.03.016
- Muller, S., Bley, N., Glass, M., Busch, B., Rousseau, V., Misiak, D., et al. (2018). IGF2BP1 enhances an aggressive tumor cell phenotype by impairing miRNA-directed downregulation of oncogenic factors. *Nucleic Acids Res.* 46, 6285–6303. doi: 10.1093/nar/gky229
- Ni, W., Yao, S., Zhou, Y., Liu, Y., Huang, P., Zhou, A., et al. (2019). Long noncoding RNA GAS5 inhibits progression of colorectal cancer by interacting with and triggering YAP phosphorylation and degradation and is negatively regulated by the m(6)A reader YTHDF3. *Mol. Cancer* 18:143. doi: 10.1186/s12943-019-1079-y
- Niu, Y., Lin, Z., Wan, A., Chen, H., Liang, H., Sun, L., et al. (2019). RNA N6-methyladenosine demethylase FTO promotes breast tumor progression through inhibiting BNIP3. *Mol. Cancer* 18:46. doi: 10.1186/s12943-019-1004-4
- Pan, T. (2013). N6-methyl-adenosine modification in messenger and long non-coding RNA. *Trends Biochem. Sci.* 38, 204–209. doi: 10.1016/j.tibs.2012.12.006
- Panneerdoss, S., Eedunuri, V. K., Yadav, P., Timilsina, S., Rajamanickam, S., Viswanadhapalli, S., et al. (2018). Cross-talk among writers, readers, and erasers of m(6)A regulates cancer growth and progression. *Sci. Adv.* 4:eaar8263. doi: 10.1126/sciadv.aar8263
- Patil, D. P., Chen, C. K., Pickering, B. F., Chow, A., Jackson, C., Guttman, M., et al. (2016). m(6)A RNA methylation promotes XIST-mediated transcriptional repression. *Nature* 537, 369–373. doi: 10.1038/nature19342
- Pendleton, K. E., Chen, B., Liu, K., Hunter, O. V., Xie, Y., Tu, B. P., et al. (2017). The U6 snRNA m(6)A Methyltransferase METTL16 Regulates SAM Synthetase Intron Retention. *Cell* 169, 824–835.e14. doi: 10.1016/j.cell.2017.05.003
- Peng, W., Li, J., Chen, R., Gu, Q., Yang, P., Qian, W., et al. (2019). Upregulated METTL3 promotes metastasis of colorectal Cancer via miR-1246/SPRED2/MAPK signaling pathway. *J. Exp. Clin. Cancer Res.* 38:393. doi: 10.1186/s13046-019-1408-4
- Perry, R. P., Kelley, D. E., Friderici, K., and Rottman, F. (1975). The methylated constituents of L cell messenger RNA: evidence for an unusual cluster at the 5' terminus. *Cell* 4, 387–394. doi: 10.1016/0092-8674(75)90159-2
- Pi, J., Wang, W., Ji, M., Wang, X., Wei, X., Jin, J., et al. (2020). YTHDF1 promotes gastric carcinogenesis by controlling translation of FZD7. *Cancer Res.* doi: 10.1158/0008-5472.CAN-20-0066 [Epub ahead of print].
- Ping, X. L., Sun, B. F., Wang, L., Xiao, W., Yang, X., Wang, W. J., et al. (2014). Mammalian WTAP is a regulatory subunit of the RNA N6-methyladenosine methyltransferase. *Cell Res.* 24, 177–189. doi: 10.1038/cr.2014.3
- Rini, B. I., Campbell, S. C., and Escudier, B. (2009). Renal cell carcinoma. *Lancet* 373, 1119–1132. doi: 10.1016/S0140-6736(09)60229-4

- Roundtree, I. A., Evans, M. E., Pan, T., and He, C. (2017a). Dynamic RNA modifications in gene expression regulation. *Cell* 169, 1187–1200. doi: 10.1016/j.cell.2017.05.045
- Roundtree, I. A., Luo, G. Z., Zhang, Z., Wang, X., Zhou, T., Cui, Y., et al. (2017b). YTHDC1 mediates nuclear export of N(6)-methyladenosine methylated mRNAs. *eLife* 6:e31311. doi: 10.7554/eLife.31311
- Salta, E., and De Strooper, B. (2017). Noncoding RNAs in neurodegeneration. *Nat. Rev. Neurosci.* 18, 627–640. doi: 10.1038/nrn.2017.90
- Schwartz, S., Mumbach, M. R., Jovanovic, M., Wang, T., Maciag, K., Bushkin, G. G., et al. (2014). Perturbation of m6A writers reveals two distinct classes of mRNA methylation at internal and 5' sites. *Cell Rep.* 8, 284–296. doi: 10.1016/j.celrep.2014.05.048
- Servaes, S. E., Hoffer, F. A., Smith, E. A., and Khanna, G. (2019). Imaging of Wilms tumor: an update. *Pediatr. Radiol.* 49, 1441–1452. doi: 10.1007/s00247-019-04423-3
- Shen, C., Xuan, B., Yan, T., Ma, Y., Xu, P., Tian, X., et al. (2020). m(6)A-dependent glycolysis enhances colorectal cancer progression. *Mol. Cancer* 19:72. doi: 10.1186/s12943-020-01190-w
- Shi, H., Wang, X., Lu, Z., Zhao, B. S., Ma, H., Hsu, P. J., et al. (2017). YTHDF3 facilitates translation and decay of N(6)-methyladenosine-modified RNA. *Cell Res.* 27, 315–328. doi: 10.1038/cr.2017.15
- Shi, Y., Zheng, C., Jin, Y., Bao, B., Wang, D., Hou, K., et al. (2020). Reduced Expression of METTL3 Promotes Metastasis of Triple-Negative Breast Cancer by m6A Methylation-Mediated COL3A1 Up-Regulation. *Front. Oncol.* 10:1126. doi: 10.3389/fonc.2020.01126
- Shima, H., Matsumoto, M., Ishigami, Y., Ebina, M., Muto, A., Sato, Y., et al. (2017). S-Adenosylmethionine Synthesis Is Regulated by Selective N(6)-Adenosine Methylation and mRNA Degradation Involving METTL16 and YTHDC1. *Cell Rep.* 21, 3354–3363. doi: 10.1016/j.celrep.2017.11.092
- Siegel, R. L., Miller, K. D., and Jemal, A. (2020). Cancer statistics, 2020. *CA Cancer J. Clin.* 70, 7–30. doi: 10.3322/caac.21590
- Strick, A., von Hagen, F., Gundert, L., Klumper, N., Tolkach, Y., Schmidt, D., et al. (2020). The N(6)-methyladenosine (m(6)A) erasers alkylation repair homologue 5 (ALKBH5) and fat mass and obesity-associated protein (FTO) are prognostic biomarkers in patients with clear cell renal carcinoma. *BJU Int.* 125, 617–624. doi: 10.1111/bju.15019
- Su, R., Dong, L., Li, C., Nachtergaele, S., Wunderlich, M., Qing, Y., et al. (2018). R-2HG Exhibits anti-tumor activity by targeting FTO/m(6)A/MYC/CEBPA signaling. *Cell* 172, 90–105.e23. doi: 10.1016/j.cell.2017.11.031
- Tanabe, A., Tanikawa, K., Tsunetomi, M., Takai, K., Ikeda, H., Konno, J., et al. (2016). RNA helicase YTHDC2 promotes cancer metastasis via the enhancement of the efficiency by which HIF-1 α mRNA is translated. *Cancer Lett.* 376, 34–42. doi: 10.1016/j.canlet.2016.02.022
- Tang, B., Yang, Y., Kang, M., Wang, Y., Wang, Y., Bi, Y., et al. (2020). m(6)A demethylase ALKBH5 inhibits pancreatic cancer tumorigenesis by decreasing WIF-1 RNA methylation and mediating Wnt signaling. *Mol. Cancer* 19:3. doi: 10.1186/s12943-019-1128-6
- Tang, C., Klukovich, R., Peng, H., Wang, Z., Yu, T., Zhang, Y., et al. (2018). ALKBH5-dependent m6A demethylation controls splicing and stability of long 3'-UTR mRNAs in male germ cells. *Proc. Natl. Acad. Sci. U.S.A.* 115, E325–E333. doi: 10.1073/pnas.1717794115
- Visvanathan, A., Patil, V., Arora, A., Hegde, A. S., Arivazhagan, A., Santosh, V., et al. (2018). Essential role of METTL3-mediated m(6)A modification in glioma stem-like cells maintenance and radioresistance. *Oncogene* 37, 522–533. doi: 10.1038/onc.2017.351
- Wang, H., Deng, Q., Lv, Z., Ling, Y., Hou, X., Chen, Z., et al. (2019). N6-methyladenosine induced miR-143-3p promotes the brain metastasis of lung cancer via regulation of VASH1. *Mol. Cancer* 18:181. doi: 10.1186/s12943-019-1108-x
- Wang, Q., Chen, C., Ding, Q., Zhao, Y., Wang, Z., Chen, J., et al. (2020). METTL3-mediated m(6)A modification of HDGF mRNA promotes gastric cancer progression and has prognostic significance. *Gut* 69, 1193–1205. doi: 10.1136/gutjnl-2019-319639
- Wang, S., Chim, B., Su, Y., Khil, P., Wong, M., Wang, X., et al. (2019). Enhancement of LIN28B-induced hematopoietic reprogramming by IGF2BP3. *Genes Dev.* 33, 1048–1068. doi: 10.1101/gad.325100.119
- Wang, X., Feng, J., Xue, Y., Guan, Z., Zhang, D., Liu, Z., et al. (2016). Structural basis of N(6)-adenosine methylation by the METTL3-METTL14 complex. *Nature* 534, 575–578. doi: 10.1038/nature18298
- Wang, Y., Gao, M., Zhu, F., Li, X., Yang, Y., Yan, Q., et al. (2020). METTL3 is essential for postnatal development of brown adipose tissue and energy expenditure in mice. *Nat. Commun.* 11:1648. doi: 10.1038/s41467-020-15488-2
- Wang, T., Kong, S., Tao, M., and Ju, S. (2020). The potential role of RNA N6 methyladenosine in Cancer progression. *Mol. Cancer* 19:88. doi: 10.1186/s12943-020-01204-7
- Wang, Y., Li, Y., Toth, J. I., Petroski, M. D., Zhang, Z., and Zhao, J. C. (2014). N6-methyladenosine modification destabilizes developmental regulators in embryonic stem cells. *Nat. Cell Biol.* 16, 191–198. doi: 10.1038/ncb2902
- Warda, A. S., Kretschmer, J., Hackert, P., Lenz, C., Urlaub, H., Hobartner, C., et al. (2017). Human METTL16 is a N(6)-methyladenosine (m(6)A) methyltransferase that targets pre-mRNAs and various non-coding RNAs. *EMBO Rep.* 18, 2004–2014. doi: 10.15252/embr.201744940
- Wen, J., Lv, R., Ma, H., Shen, H., He, C., Wang, J., et al. (2018). Zc3h13 Regulates Nuclear RNA m(6)A Methylation and Mouse Embryonic Stem Cell Self-Renewal. *Mol. Cell* 69, 1028–1038.e6. doi: 10.1016/j.molcel.2018.02.015
- Wen, L., Pan, X., Yu, Y., and Yang, B. (2020). Down-regulation of FTO promotes proliferation and migration, and protects bladder cancer cells from cisplatin-induced cytotoxicity. *BMC Urol.* 20:39. doi: 10.1186/s12894-020-00612-7
- Weng, H., Huang, H., Wu, H., Qin, X., Zhao, B. S., Dong, L., et al. (2018). METTL14 Inhibits Hematopoietic Stem/Progenitor Differentiation and Promotes Leukemogenesis via mRNA m(6)A Modification. *Cell Stem Cell* 22, 191. doi: 10.1016/j.stem.2017.11.016
- Woodcock, C. B., Horton, J. R., Zhou, J., Bedford, M. T., Blumenthal, R. M., Zhang, X., et al. (2020). Biochemical and structural basis for YTH domain of human YTHDC1 binding to methylated adenine in DNA. *Nucleic Acids Res.* doi: 10.1093/nar/gkaa604 [Epub ahead of print].
- Xiang, S., Liang, X., Yin, S., Liu, J., and Xiang, Z. (2020). N6-methyladenosine methyltransferase METTL3 promotes colorectal cancer cell proliferation through enhancing MYC expression. *Am. J. Transl. Res.* 12, 1789–1806.
- Xiao, L., Li, X., Mu, Z., Zhou, J., Zhou, P., Xie, C., et al. (2020). FTO inhibition enhances the anti-tumor effect of temozolomide by targeting MYC-miR-155/23a cluster-MXI1 feedback circuit in glioma. *Cancer Res.* doi: 10.1158/0008-5472.CAN-20-0132 [Epub ahead of print].
- Xie, H., Li, J., Ying, Y., Yan, H., Jin, K., Ma, X., et al. (2020). METTL3/YTHDF2 m(6)A axis promotes tumorigenesis by degrading SETD7 and KLF4 mRNAs in bladder cancer. *J. Cell. Mol. Med.* 24, 4092–4104. doi: 10.1111/jcmm.15063
- Xu, H., Wang, H., Zhao, W., Fu, S., Li, Y., Ni, W., et al. (2020). SUMO1 modification of methyltransferase-like 3 promotes tumor progression via regulating Snail mRNA homeostasis in hepatocellular carcinoma. *Theranostics* 10, 5671–5686. doi: 10.7150/thno.42539
- Xu, K., Yang, Y., Feng, G. H., Sun, B. F., Chen, J. Q., Li, Y. F., et al. (2017). Mettl3-mediated m(6)A regulates spermatogonial differentiation and meiosis initiation. *Cell Res.* 27, 1100–1114. doi: 10.1038/cr.2017.100
- Xu, X., Yu, Y., Zong, K., Lv, P., and Gu, Y. (2019). Up-regulation of IGF2BP2 by multiple mechanisms in pancreatic cancer promotes cancer proliferation by activating the PI3K/Akt signaling pathway. *J. Exp. Clin. Cancer Res.* 38:497. doi: 10.1186/s13046-019-1470-y
- Yang, F., Jin, H., Que, B., Chao, Y., Zhang, H., Ying, X., et al. (2019). Dynamic m(6)A mRNA methylation reveals the role of METTL3-m(6)A-CDCP1 signaling axis in chemical carcinogenesis. *Oncogene* 38, 4755–4772. doi: 10.1038/s41388-019-0755-0
- Yang, S., Wei, J., Cui, Y. H., Park, G., Shah, P., Deng, Y., et al. (2019). m(6)A mRNA demethylase FTO regulates melanoma tumorigenicity and response to anti-PD-1 blockade. *Nat. Commun.* 10:2782. doi: 10.1038/s41467-019-10669-0
- Yang, X., Zhang, S., He, C., Xue, P., Zhang, L., He, Z., et al. (2020). METTL14 suppresses proliferation and metastasis of colorectal cancer by down-regulating oncogenic long non-coding RNA XIIST. *Mol. Cancer* 19:46. doi: 10.1186/s12943-020-1146-4
- Yang, Z., Li, J., Feng, G., Gao, S., Wang, Y., Zhang, S., et al. (2017). MicroRNA-145 Modulates N(6)-Methyladenosine Levels by Targeting the 3'-Untranslated mRNA Region of the N(6)-Methyladenosine Binding YTH Domain Family 2 Protein. *J. Biol. Chem.* 292, 3614–3623. doi: 10.1074/jbc.M116.749689

- Yu, R., Li, Q., Feng, Z., Cai, L., and Xu, Q. (2019). m⁶A Reader YTHDF2 Regulates LPS-Induced Inflammatory Response. *Int. J. Mol. Sci.* 20:1323. doi: 10.3390/ijms20061323
- Yuan, Y., Du, Y., Wang, L., and Liu, X. (2020). The M⁶A methyltransferase METTL3 promotes the development and progression of prostate carcinoma via mediating MYC methylation. *J. Cancer* 11, 3588–3595. doi: 10.7150/jca.42338
- Yue, B., Song, C., Yang, L., Cui, R., Cheng, X., Zhang, Z., et al. (2019). METTL3-mediated N⁶-methyladenosine modification is critical for epithelial-mesenchymal transition and metastasis of gastric cancer. *Mol. Cancer* 18:142. doi: 10.1186/s12943-019-1065-4
- Zaccara, S., and Jaffrey, S. R. (2020). A Unified Model for the Function of YTHDF Proteins in Regulating m(6)A-Modified mRNA. *Cell* 181, 1582.e18–1595.e18. doi: 10.1016/j.cell.2020.05.012
- Zhang, C., Huang, S., Zhuang, H., Ruan, S., Zhou, Z., Huang, K., et al. (2020). YTHDF2 promotes the liver cancer stem cell phenotype and cancer metastasis by regulating OCT4 expression via m⁶A RNA methylation. *Oncogene* 39, 4507–4518. doi: 10.1038/s41388-020-1303-7
- Zhang, F., Xiang, S., Cao, Y., Li, M., Ma, Q., Liang, H., et al. (2017). EIF3D promotes gallbladder cancer development by stabilizing GRK2 kinase and activating PI3K-AKT signaling pathway. *Cell Death Dis.* 8:e2868. doi: 10.1038/cddis.2017.263
- Zhang, G., Huang, H., Liu, D., Cheng, Y., Liu, X., Zhang, W., et al. (2015). N⁶-methyladenine DNA modification in *Drosophila*. *Cell* 161, 893–906. doi: 10.1016/j.cell.2015.04.018
- Zhang, Y., Wang, X., Zhang, X., Wang, J., Ma, Y., Zhang, L., et al. (2019). RNA-binding protein YTHDF3 suppresses interferon-dependent antiviral responses by promoting FOXO3 translation. *Proc. Natl. Acad. Sci. U.S.A.* 116, 976–981. doi: 10.1073/pnas.1812536116
- Zhao, W., Cui, Y., Liu, L., Ma, X., Qi, X., Wang, Y., et al. (2020). METTL3 Facilitates Oral Squamous Cell Carcinoma Tumorigenesis by Enhancing c-Myc Stability via YTHDF1-Mediated m(6)A Modification. *Mol. Ther. Nucleic Acids* 20, 1–12. doi: 10.1016/j.omtn.2020.01.033
- Zhao, X., Yang, Y., Sun, B. F., Shi, Y., Yang, X., Xiao, W., et al. (2014). FTO-dependent demethylation of N⁶-methyladenosine regulates mRNA splicing and is required for adipogenesis. *Cell Res.* 24, 1403–1419. doi: 10.1038/cr.2014.151
- Zhao, Y., Tao, Z., and Chen, X. (2020). Identification of a three-m⁶A related gene risk score model as a potential prognostic biomarker in clear cell renal cell carcinoma. *PeerJ* 8:e8827. doi: 10.7717/peerj.8827
- Zheng, G., Dahl, J. A., Niu, Y., Fedorcsak, P., Huang, C. M., Li, C. J., et al. (2013). ALKBH5 is a mammalian RNA demethylase that impacts RNA metabolism and mouse fertility. *Mol. Cell* 49, 18–29. doi: 10.1016/j.molcel.2012.10.015
- Zhou, B., Liu, C., Xu, L., Yuan, Y., Zhao, J., Zhao, W., et al. (2020). N(6)-methyladenosine Reader Protein Ythdc2 Suppresses Liver Steatosis via Regulation of mRNA Stability of Lipogenic Genes. *Hepatology*. doi: 10.1002/hep.31220 [Epub ahead of print].
- Zhou, J., Wan, J., Gao, X., Zhang, X., Jaffrey, S. R., and Qian, S. B. (2015). Dynamic m(6)A mRNA methylation directs translational control of heat shock response. *Nature* 526, 591–594. doi: 10.1038/nature15377
- Zhou, S., Bai, Z. L., Xia, D., Zhao, Z. J., Zhao, R., Wang, Y. Y., et al. (2018). FTO regulates the chemo-radiotherapy resistance of cervical squamous cell carcinoma (CSCC) by targeting beta-catenin through mRNA demethylation. *Mol. Carcinog.* 57, 590–597. doi: 10.1002/mc.22782
- Zhou, Y., Huang, T., Siu, H. L., Wong, C. C., Dong, Y., Wu, F., et al. (2017). IGF2BP3 functions as a potential oncogene and is a crucial target of miR-34a in gastric carcinogenesis. *Mol. Cancer* 16:77. doi: 10.1186/s12943-017-0647-2
- Zhu, S., Beaulaurier, J., Deikus, G., Wu, T. P., Strahl, M., Hao, Z., et al. (2018). Mapping and characterizing N⁶-methyladenine in eukaryotic genomes using single-molecule real-time sequencing. *Genome Res.* 28, 1067–1078. doi: 10.1101/gr.231068.117
- Zhu, W., Si, Y., Xu, J., Lin, Y., Wang, J. Z., Cao, M., et al. (2020). Methyltransferase like 3 promotes colorectal cancer proliferation by stabilizing CCNE1 mRNA in an m⁶A-dependent manner. *J. Cell. Mol. Med.* 24, 3521–3533. doi: 10.1111/jcmm.15042
- Zhuang, C., Zhuang, C., Luo, X., Huang, X., Yao, L., Li, J., et al. (2019). N⁶-methyladenosine demethylase FTO suppresses clear cell renal cell carcinoma through a novel FTO-PGC-1 α signalling axis. *J. Cell. Mol. Med.* 23, 2163–2173. doi: 10.1111/jcmm.14128
- Zhuang, M., Li, X., Zhu, J., Zhang, J., Niu, F., Liang, F., et al. (2019). The m⁶A reader YTHDF1 regulates axon guidance through translational control of Robo3.1 expression. *Nucleic Acids Res.* 47, 4765–4777. doi: 10.1093/nar/gkz157.reA

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 Li, Ge, Xu, Xu, Dou and Jia. 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.



Transcriptome-Wide m⁶A Methylation in Skin Lesions From Patients With Psoriasis Vulgaris

Ya-Nan Wang and Hong-Zhong Jin*

Department of Dermatology, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing, China

OPEN ACCESS

Edited by:

Giovanni Nigita,
The Ohio State University,
United States

Reviewed by:

Natalia Pinello,
Royal Prince Alfred Hospital, Australia
Haobo Li,
Harvard Medical School,
United States

*Correspondence:

Hong-Zhong Jin
jinhongzhong@263.net

Specialty section:

This article was submitted to
Epigenomics and Epigenetics,
a section of the journal
Frontiers in Cell and Developmental
Biology

Received: 05 August 2020

Accepted: 14 October 2020

Published: 05 November 2020

Citation:

Wang Y-N and Jin H-Z (2020)
Transcriptome-Wide m⁶A Methylation
in Skin Lesions From Patients With
Psoriasis Vulgaris.
Front. Cell Dev. Biol. 8:591629.
doi: 10.3389/fcell.2020.591629

N⁶-methyladenosine (m⁶A) methylation, as the most prevalent internal RNA modification, has been revealed to play critical roles in various biological functions. In this study, we performed m⁶A transcriptome-wide profiling in three kinds of skin tissue: involved psoriatic skin (PP), uninvolved psoriatic skin (PN), and healthy control skin samples (NN). The findings revealed that transcripts of PP contained the fewest m⁶A peaks and lowest m⁶A peak density. The greatest differences of m⁶A methylation were observed in the PP vs. NN and PP vs. PN comparisons. Intriguingly, in these comparisons, hypermethylated m⁶A was mainly enriched within the CDSs and 3'UTRs, while hypomethylated m⁶A was not only enriched within CDSs and 3'UTRs, but also within 5'UTRs. GO and KEGG pathway analyses indicated that hypermethylated transcripts in PP were particularly associated with response-associated terms, cytokine production, and olfactory transduction. Meanwhile, hypomethylated transcripts in PP were mainly associated with development-related processes and the Wnt signaling pathway. In addition, we discovered that 19.3–48.4% of the differentially expressed transcripts in psoriasis vulgaris were modified by m⁶A, and that transcripts with lower expression were more preferentially modified by m⁶A. Moreover, upregulation of gene expression was often accompanied by upregulation of m⁶A methylation, suggesting a regulatory role of m⁶A in psoriasis vulgaris gene expression.

Keywords: N⁶-methyladenosine, m⁶A methylation, psoriasis vulgaris, MeRIP-seq, RNA modification

INTRODUCTION

Psoriasis is a chronic, systemic, inflammatory disease, affecting about 2% of the world's population (Greb et al., 2016). The most common subtype is psoriasis vulgaris, accounting for 80–90% of all cases (Greb et al., 2016). In 2014, the WHO adopted a resolution that defines psoriasis as a chronic, non-communicable, painful, disfiguring, and disabling disease for which there is no cure (World Health Organization, 2014). The skin lesions of psoriasis vulgaris are characterized by excessive proliferation and abnormal differentiation of keratinocytes, accompanied by significant infiltration of inflammatory cells, which originate as a result of dysregulation of immunity triggered by environmental and genetic stimuli (Boehncke and Schon, 2015; Hugh and Weinberg, 2018). Situated between heredity and the environment are epigenetic markers, which are another layer of biological information (Pollock et al., 2017). Extensive epidemiological and molecular analyses

have provided evidence for the involvement of epigenetics in psoriasis vulgaris (Pollock et al., 2017; Rendon and Schäkel, 2019). The effects of pharmacological inhibitors of epigenetic-modifying enzymes in psoriasis vulgaris have been studied and verified (Bovenschen et al., 2011; Hammitzsch et al., 2015). Epigenetic mechanisms modify gene expression without changing the sequence of the genome, such as long non-coding RNA (lncRNA) and microRNA (miRNA) silencing, and DNA methylation, but RNA modifications in psoriasis vulgaris have not yet been reported (Rendon and Schäkel, 2019).

In the past, the study of epigenetic changes in human diseases mainly focused on DNA methylation and histone modification. Although post-transcriptional modifications of RNA have been known for more than 50 years, the effects of these modifications on the regulation of gene expression has only begun to be explored in recent years due to the previous lack of sensitive detection techniques (He, 2010). Against the background of the rapid development of second-generation sequencing technology, as early as 2012, the sequencing technology of methylated RNA immunoprecipitation with next-generation sequencing (MeRIP-Seq) was developed (Dominissini et al., 2013). m⁶A (N⁶-methyladenosine), involving methylation at the 6th position nitrogen atom of adenine (A) of RNA, is the most prevalent RNA modification in messenger RNA (mRNA) and lncRNA of higher eukaryotes (Wang et al., 2017b; Wang et al., 2020). m⁶A methylation is highly conserved and occurs widely in eukaryotic species ranging from yeast, plant, and drosophila to mammals, as well as viruses (Bi et al., 2019). m⁶A methylation can be “written” by methyltransferases [e.g., methyltransferase-like 3 (METTL3), methyltransferase-like 14 (METTL14), and Wilms tumor 1-associated protein (WTAP)] and “erased” by demethylases [e.g., fat-mass and obesity-associated protein (FTO) and alkB homolog 5 (ALKBH5)], which confirms that m⁶A methylation is dynamic and reversible (Bi et al., 2019). In addition, the discovery of binding proteins [e.g., YTH domain family 1–3 (YTHDF1–3)] confirmed that m⁶A methylation has a wide range of biological effects and significance (Fu et al., 2014). In recent years, m⁶A methylation has emerged as a critical post-transcriptional regulator of gene expression programs, and it can regulate various aspects of RNA, including splicing, transport, translation, and stability (Fu et al., 2014; Frye et al., 2018). It has been found that m⁶A methylation is closely related to metabolism, carcinogenesis, nervous system, mental, oral, and immune diseases (Church et al., 2010; Engel and Chen, 2018; Pan et al., 2018). However, to the best of our knowledge, no reports have been published on m⁶A methylation in the context of psoriasis vulgaris.

In this study, we established a transcriptome-wide m⁶A methylome profile of psoriasis vulgaris, as assessed by MeRIP-Seq. In addition, we used RNA-Seq data to perform a combined analysis of m⁶A methylation and mRNA levels.

MATERIALS AND METHODS

Patient Samples

Four patients with psoriasis vulgaris were recruited from the outpatient clinics of the Peking Union Medical College Hospital

(Supplementary Table 1 and Supplementary Figure 1). Prior to recruitment, none of the patients had received systemic tretinoin, glucocorticoid, immunosuppressant, or biological agents, nor had they received PUVA/solarium/UV treatment for at least 3 months or topical therapy for at least 2 weeks before the study start. Infectious diseases, tumors, autoimmune diseases, and other immune-related diseases were ruled out. The subjects had no coagulatory disorders or other diseases making them unsuitable to undergo a surgical operation. Skin biopsies were obtained from the patients with psoriasis vulgaris by minimally invasive surgery under aseptic conditions with local anesthesia (0.5% lidocaine). Two types of skin sample were collected from each patient: one from active lesions and the other from skin not exhibiting any of the macroscopic changes related to psoriatic lesions, at least 3 cm away from the active lesions. Four age- and sex-matched healthy controls without a personal or family history of psoriasis vulgaris were enrolled from the Department of Plastic Surgery at Peking Union Medical College Hospital. Healthy skin tissues were obtained from these volunteers. This study was approved by the Ethics Committee of Peking Union Medical College Hospital and informed consent was obtained from all patients and unaffected individuals.

MeRIP-Seq, RNA-Seq, and Data Analysis

The MeRIP-Seq was performed by Cloudseq Biotech Inc. (Shanghai, China), in accordance with the published procedure with slight modifications (Meyer et al., 2012). Briefly, four biological replicates were used for the PP, PN, and NN groups. Total RNA was extracted from the three groups of skin tissue using TRIzol reagent (Life Technologies, Carlsbad, CA, United States). The quality and quantity of total RNA were assessed using NanoDrop ND-2000 (Thermo Fisher Scientific, Waltham, MA, United States). The RNA integrity was measured using denaturing agarose gel electrophoresis. Seq-StarTM poly(A) mRNA Isolation Kit (Arraystar, Rockville, MD, United States) was used to isolate mRNA from total RNA. mRNA was randomly fragmented to 200 nt by RNA Fragmentation Reagents (Ambion) (Invitrogen, Carlsbad, CA, United States). A total of 5 µg of fragmented mRNA was saved as input control for RNA-Seq, while 500 µg of fragmented mRNA was used to perform m⁶A RNA immunoprecipitation with GenSeqTM m⁶A-MeRIP Kit (GenSeq, Beijing, China) [including anti-m⁶A polyclonal antibody (Synaptic Systems, Goettingen, Germany)]. Both the m⁶A IP sample and the input sample (without immunoprecipitation) were used for library preparation with NEBNext[®] Ultra II Directional RNA Library Prep Kit (New England Biolabs, Ipswich, MA, United States). The library quality was evaluated with BioAnalyzer 2100 system (Agilent Technologies, Santa Clara, CA, United States). Library sequencing was performed on an Illumina HiSeq instrument with 150 bp paired-end reads.

Paired-end reads were harvested from the Illumina HiSeq 4000 sequencer and were subjected to quality control by Q30. After 3' adaptor-trimming and the removal of low-quality reads by cutadapt software (v1.9.3) (Martin, 2011), the reads were aligned to the reference genome (UCSC HG19) with Hisat2 software (v2.0.4) (Kim et al., 2015). Methylated sites on RNAs

(peaks) were identified using Model-based Analysis of ChIP-Seq (MACS) software (Zhang et al., 2008). Peak numbers refer to peaks present in all four samples in each group. Differentially methylated sites (fold change > 2 and $P < 0.00001$) on mRNAs were identified by diffReps (Shen et al., 2013). Genes of interest were visualized in the IGV (Integrative Genomics Viewer) software (v2.3.68) (Thorvaldsdóttir et al., 2013). Those peaks identified by both software programs overlapping with exons of mRNA were determined and selected using custom-made scripts. The GO analysis and pathway enrichment analysis were performed on the differentially methylated protein-coding genes by using GO¹ and KEGG databases². Sequence motifs were identified using Homer (Heinz et al., 2010). Gene expression was calculated by Cufflinks (Trapnell et al., 2012) using sequencing reads from input samples. Cuffdiff (Trapnell et al., 2012) was used to find DE genes.

MeRIP-RT-qPCR

RNA fragmentation and m⁶A-immunoprecipitation were performed as described above. Briefly, total RNA was extracted using TRIzol (Life Technologies) and fragmented by RNA fragmentation reagents (Thermo Fisher Scientific) or not. After saving 50 ng of the total RNA as input, the remaining RNA (2 µg) was used for m⁶A-immunoprecipitation with anti-m⁶A polyclonal antibody (Synaptic Systems) in 500 µL of IP buffer (150 mM NaCl, 0.1% NP-40, 10 mM Tris, pH 7.4, 100 U RNase inhibitor) to obtain the m⁶A pull-down portion (m⁶A IP portion). m⁶A RNAs were immunoprecipitated with Dynabeads® Protein A (Thermo Fisher Scientific) and eluted twice with elution buffer (5 mM Tris-HCL pH 7.5, 1 mM EDTA pH 8.0, 0.05% SDS, 20 mg/ml Proteinase K). m⁶A IP RNAs were recovered by ethanol precipitation and RNA concentration was measured with Qubit® RNA HS Assay Kit (Thermo Fisher Scientific). Then, 2 ng of the total RNA and m⁶A IP RNA were used to synthesize complementary DNA by using the iScript cDNA Synthesis Kit (Bio-Rad, CA, United States). Real-time PCR was subsequently performed, using an SYBR Premix Ex Taq (Takara, Liaoning, China) and ABI 7500 Sequence Detection System (Thermo Fisher Scientific). All procedures were performed in accordance with the manufacturer's protocols. The sequences of the primers used are presented in **Supplementary Table 2**.

Statistical Analyses

Experiments were performed at least three times and representative results are shown. Statistical analysis was performed using GraphPad Prism Version 8.0 software. Differences between individual groups were analyzed using the chi-squared test and Student's *t*-test (two-tailed and unpaired) with triplicate or quadruplicate sets. Pearson's correlation was adopted to carry out the correlation analysis. A value of $P < 0.05$ was considered statistically significant.

¹ www.geneontology.org

² www.genome.jp/kegg

RESULTS

Overview of m⁶A Methylation in Psoriasis Vulgaris mRNA

To obtain the transcriptome-wide m⁶A map of psoriasis vulgaris, we examined three kinds of skin tissue, namely, involved psoriatic skin (PP), uninvolved psoriatic skin (PN), and healthy control skin samples (NN), using m⁶A-targeted antibody coupled with high-throughput sequencing (i.e., MeRIP-Seq). Psoriasis Area and Severity Index (PASI) scores for psoriatic patients ranged from 10.8 to 16.7. Using Illumina HiSeq 4000, we acquired 68,678,097, 62,087,334, and 62,121,508 reads from PP, PN, and NN, respectively. After end-trimming and quality filtering, 57,334,158, 45,130,912, and 49,749,336 high-quality reads (83.11, 72.65, and 80.04% of the total reads) from PP, PN, and NN, respectively, were mapped to the human reference genome (UCSC HG19) (**Supplementary Table 3**).

The fewest m⁶A peaks (sites) were identified in PP, while the most were identified in PN. Specifically, 16,868 m⁶A peaks were identified from 16,520 genes in PP, 22,144 m⁶A peaks were identified from 17,665 genes in PN, and 20,408 m⁶A peaks were identified from 17,358 genes in NN (**Figure 1A**, **Supplementary Figure 2**, **Supplementary Table 4**, and **Supplementary Data 1**). We next compared the shared peaks between groups; we found that transcripts of PP and NN carried the fewest shared m⁶A peaks (a total of 15,176 m⁶A peaks from 10,604 genes), while transcripts of PN and NN carried the largest number of shared m⁶A peaks (a total of 17,678 m⁶A peaks from 11,816 genes), suggesting that the difference between PP and NN was greater than that between PN and NN ($P < 0.001$) (**Figure 1B**).

m⁶A Peak Density and Distribution Pattern in Psoriasis Vulgaris mRNA

Based on the above results, we estimated that there were 0.62–0.75 m⁶A peaks per 1,000 nucleotides or 1.02–1.25 m⁶A peaks per identified transcript in PP, PN, and NN. Among them, transcripts of PP contained the lowest m⁶A peak density (0.62/1k nt or 1.02/gene), while transcripts of PN contained the highest (0.75/1k nt or 1.25/gene) (**Figure 1C** and **Supplementary Table 4**). The numbers of m⁶A peaks varied widely among individual genes, and most of them contained a single m⁶A peak. Specifically, 59.68, 52.95, and 55.25% of the methylated transcripts contained a single m⁶A peak in PP, PN, and NN, while 26.07, 28.42, and 27.15% of the methylated transcripts contained two m⁶A peaks per mRNA, respectively. For three m⁶A peaks, the percentages were further reduced to 8.88, 11.13, and 10.46% in PP, PN, and NN, respectively. Only 5.38, 7.50, and 7.14% of the methylated transcripts containing more than three m⁶A peaks (**Figure 1D**). To determine whether the identified m⁶A peaks were enriched at consensus sequences of RRACH (R represents purine, A is m⁶A, and H is a non-guanine base), we performed motif analysis and found consistent results in PP, PN, and NN (**Figure 1E** and **Supplementary Figure 3**). To analyze the distribution pattern of m⁶A in psoriasis vulgaris, the metagene profiles of all identified m⁶A peaks were investigated in the transcriptomes of PP, PN, and NN, which revealed that m⁶A peaks were highly enriched

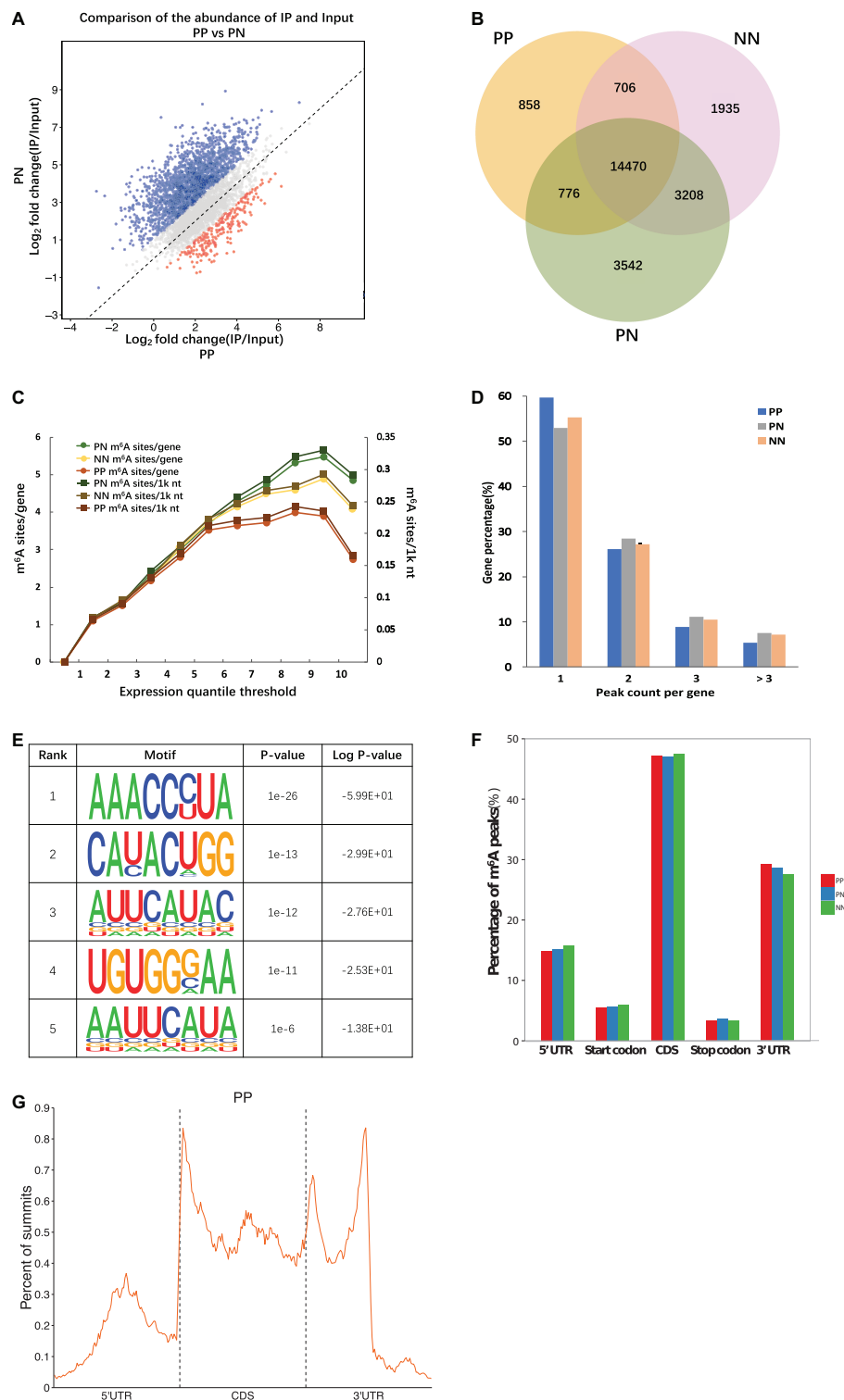


FIGURE 1 | Overview of m⁶A methylome in psoriatic vulgaris skin. **(A)** Comparison of the number of m⁶A peaks identified in PP and PN. **(B)** Numbers of group-specific and common m⁶A peaks. **(C)** Estimation of m⁶A density on transcripts. Transcripts with different expression level are divided into 10 groups, and the m⁶A density of PP, PN, and NN samples is calculated separately. **(D)** Percentage of m⁶A methylated genes with different m⁶A peak number. **(E)** The top five motifs enriched across m⁶A peaks identified from involved psoriatic skin samples. **(F)** Distribution of m⁶A peaks across the length of mRNAs. 5'UTRs, CDS, and 3'UTRs of PP mRNAs are individually binned into regions spanning 1% of their total length, and the percentage of m⁶A peaks that fall within each bin is determined. **(G)** The distribution of m⁶A peak from PP along a metagene.

within the coding sequences (CDSs) and 3' untranslated regions (3'UTRs) (Figures 1E,G and Supplementary Figure 4). These features of m⁶A methylation indicate its high conservation in psoriasis vulgaris.

Analysis of Differentially Methylated RNAs Among PP, PN, and NN Samples

We next analyzed the differentially methylated RNAs (DMRs). The greatest differences were observed in the PP vs. NN and PP vs. PN comparisons (Figure 2A, Supplementary Table 5, and Supplementary Data 2). In the PP vs. NN and PP vs. PN comparisons, the number of hypomethylated m⁶A peaks was much greater than that of hypermethylated m⁶A peaks in PP ($P < 0.001$). Specifically, compared with NN, transcripts of PP contained more hypomethylated m⁶A peaks (1,719 m⁶A peaks from 1,113 genes) than hypermethylated ones (1,470 m⁶A peaks from 1,127 genes). When compared with PN, transcripts of PP also contained more hypomethylated m⁶A peaks (2,316 m⁶A peaks from 1,568 genes) than hypermethylated ones (2,024 m⁶A peaks from 1,362 genes) (Figure 2A and Supplementary Table 5). However, when compared with NN, PN contained slightly more hypermethylated m⁶A peaks (914 m⁶A peaks from 691 genes) than hypomethylated ones (755 m⁶A peaks from 537 genes), but this difference was relatively small compared with those of PP vs. NN and PP vs. PN ($P < 0.001$) (Figure 2A and Supplementary Table 5). Furthermore, upon analyzing the m⁶A peak distribution pattern of these DMRs in PP vs. NN, PP vs. PN, and PN vs. NN, we found that the hypermethylated m⁶A peaks were mainly enriched within CDSs and 3'UTRs, while hypomethylated m⁶A peaks were not only enriched within CDSs and 3'UTRs, but also highly enriched within 5'UTRs (Table 1).

Gene Ontology Analysis of Differentially Methylated RNAs Among PP, PN, and NN Samples

To deduce the potential biological significance of m⁶A methylation in psoriasis vulgaris, we analyzed these DMRs by performing Gene Ontology (GO) analysis. GO analysis revealed that, compared with PN or NN, hypermethylated DMRs in PP were particularly associated with response-related items [e.g., defense response, response to other organisms, inflammatory response, and response to (external) biotic stimulus] and cytokine-related items (e.g., regulation of cytokine production, cytokine production, and interleukin-6 production), suggesting that these hypermethylated DMRs may be involved in the response to environmental stimulations at the initial stage of psoriasis vulgaris (Figure 2B, Supplementary Figure 5A, and Supplementary Data 3). Meanwhile, hypomethylated DMRs in PP were mainly enriched in development-related items (e.g., multicellular organism development, system development, anatomical structure development, developmental process, and nervous system development) and cell–cell signaling, which suggested that the function of hypomethylated DMRs differed from that of hypermethylated DMRs. Hypomethylated DMRs in PP were mainly focused on cell and tissue development processes (Figure 2C, Supplementary Figure 5B, and Supplementary Data 3). In addition to these shared gene

ontological terms, DMRs in the PP vs. NN and PP vs. PN comparisons were also associated with some specific terms. For example, in PP vs. NN, hypermethylated DMRs in PP were specifically associated with the G-protein-coupled receptor signaling pathway (Figure 2B and Supplementary Data 3), including cannabinoid receptors that are involved in the proliferation/differentiation and immune activity of keratinocytes, indicating that these hypermethylated DMRs may participate in critical processes of psoriasis vulgaris (Tóth et al., 2019). Meanwhile, hypomethylated DMRs in PP were specifically associated with the polyol metabolic process (Figure 2C and Supplementary Data 3), which may be related to the previous epidemiological finding that psoriasis patients tended to suffer from metabolic diseases (Hu et al., 2019). When compared with PN, hypermethylated DMRs in PP were specifically associated with cell cycle-related items (e.g., mitotic cell cycle process, mitotic cell cycle, cell cycle, and mitotic nuclear division) (Supplementary Figure 5A and Supplementary Data 3), while hypomethylated DMRs in PP were specifically associated with negative regulation of cellular component movement, locomotion, and learning (Supplementary Figure 5B and Supplementary Data 3). Given that the cell cycle-related genes were dysregulated in keratinocytes in psoriasis vulgaris, m⁶A methylation should play an important role in dysregulation of the cell cycle in keratinocytes in psoriasis vulgaris (Pasquali et al., 2019). Besides, PN and NN exhibited various differences, and the DMRs were mainly associated with system process and multicellular organismal process (Supplementary Figures 5C,D and Supplementary Data 3). All of these results suggest that m⁶A methylation participates in various pathophysiological aspects of psoriasis vulgaris.

Kyoto Encyclopedia of Genes and Genomes Pathway Analysis of Differentially Methylated Genes Among PP, PN, and NN Samples

Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis of DMRs was also conducted for both hypermethylated and hypomethylated genes. This analysis showed that, compared with NN or PN, hypermethylated DMRs in PP were mainly associated with cytokine–cytokine receptor interaction (Figure 2D, Supplementary Figure 5E, and Supplementary Data 4); hypomethylated DMRs in PP were mainly associated with the Wnt signaling pathway (Figure 2E, Supplementary Figure 5F, and Supplementary Data 4). We used quantitative reverse-transcription PCR (RT-qPCR) to validate the important genes WNT5A, DIF1, and DKK2 in Wnt signaling pathway, as well as TNF, IL17A, HIF1A, and SOCS1/3, which are associated with the pathophysiology of psoriasis; all of them showed significant enrichment in immunoprecipitation (IP) pull-down samples (Figure 3 and Supplementary Figure 6). In addition to these shared pathways, DMRs in the PP vs. NN and PP vs. PN comparisons were also associated with some specific pathways. For example, in PP vs. NN, hypermethylated DMRs in PP were specifically associated with olfactory transduction (Figure 2D and Supplementary Data 4), which may explain why psoriasis

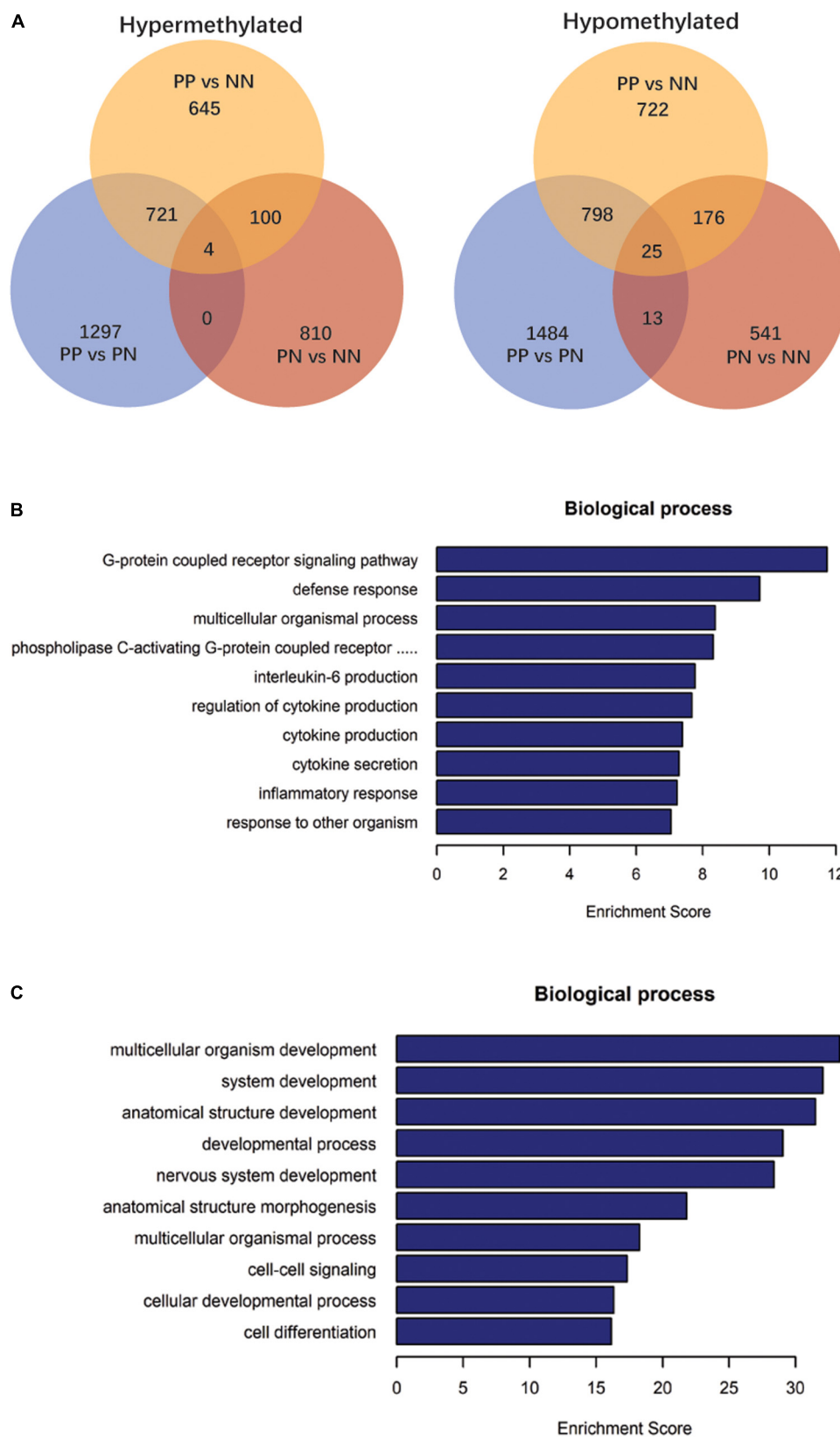


FIGURE 2 | Continued

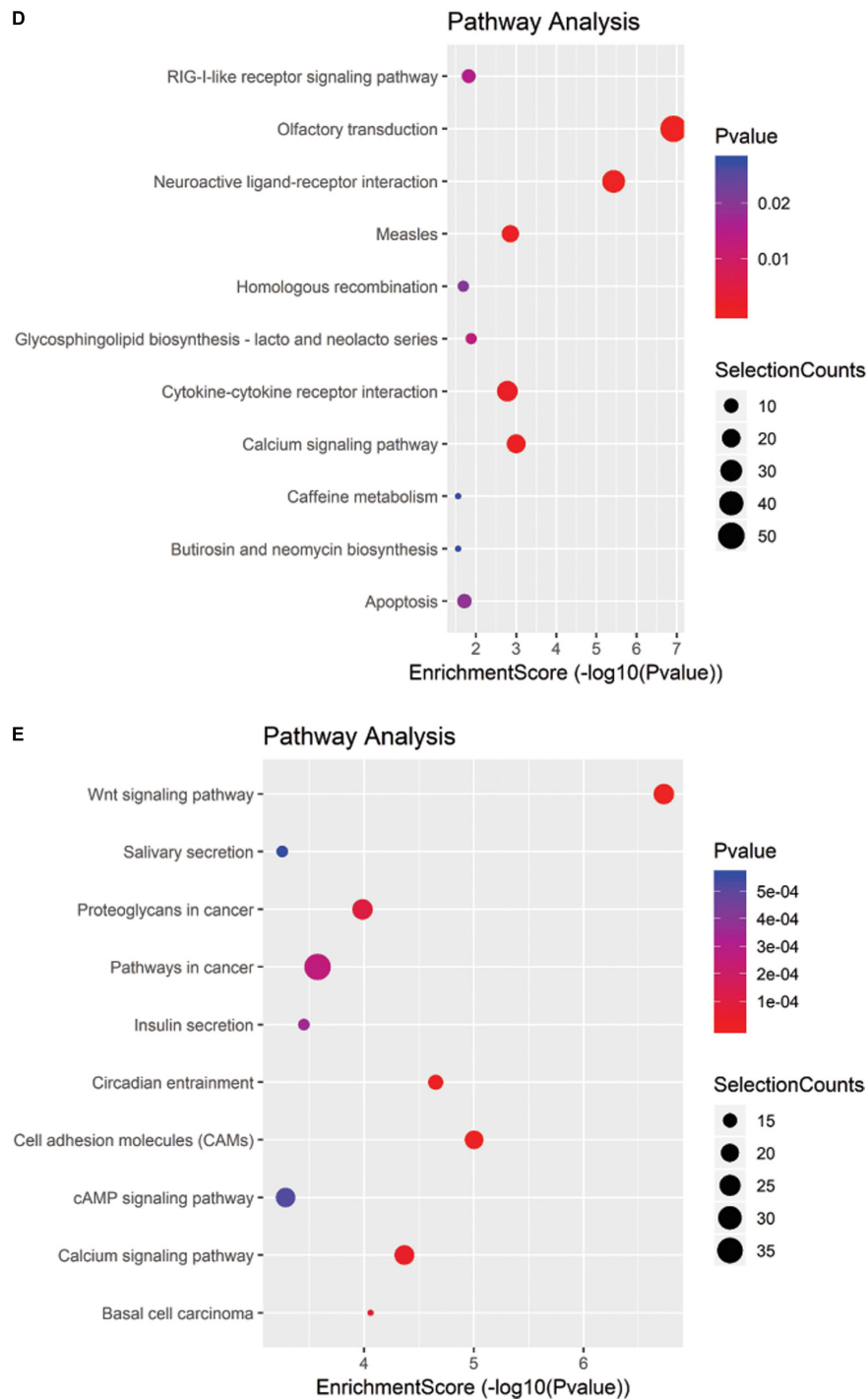


FIGURE 2 | Gene ontology and Kyoto Encyclopedia of Genes and Genomes analyses of coding genes containing altered m⁶A peaks. **(A)** Comparison of the number of hypermethylated and hypomethylated m⁶A peaks identified in PP, PN, and NN samples. **(B)** The top 10 gene ontology terms were significantly enriched for the hypermethylated genes in PP vs. NN. **(C)** The top 10 gene ontology terms were significantly enriched for the hypomethylated genes in PP vs. NN. **(D)** The top 10 significantly enriched pathways for the hypermethylated genes in PP vs. NN. **(E)** The top 10 significantly enriched pathways for the hypomethylated genes in PP vs. NN.

patients have an abnormal sense of smell (Aydın et al., 2016); meanwhile, hypomethylated DMRs in PP were specifically associated with the MAPK signaling pathway (Figure 2E and

Supplementary Data 4), which is involved in almost every aspect of psoriasis vulgaris, including keratinocyte proliferation, differentiation, migration, T-helper-cell differentiation, and

TABLE 1 | Number of m⁶A peaks among PP, PN, and NN samples.

	PP vs. NN (up-regulated)	PP vs. NN (down-regulated)	PP vs. PN (up-regulated)	PP vs. PN (down-regulated)	PN vs. NN (up-regulated)	PN vs. NN (down-regulated)
5'UTR	92	208	139	407	77	76
Start codon	36	73	54	126	34	37
CDS	628	802	882	973	367	329
Stop codon	70	56	112	81	36	29
3'UTR	382	286	572	304	220	190

angiogenesis (Mavropoulos et al., 2013). In the PP vs. PN comparison, DMRs were specifically associated with the cell cycle and gap junction (Supplementary Figures 5E,F and Supplementary Data 4). Besides, from the comparison with NN, DMRs in PN were mainly associated with vascular smooth muscle contraction and osteoclast differentiation (Supplementary Figures 5G,H and Supplementary Data 4).

Relationship Between m⁶A Peaks and mRNA Level

We next performed MeRIP-Seq and RNA-Seq combined analysis to explore whether the extent of m⁶A methylation was associated with the mRNA level of the differentially expressed genes (DEGs). Specifically, upon comparison with NN, 511 highly expressed DEGs were identified in PP, of which 116 (22.7%) transcripts were modified by m⁶A, 87.9% (102/116) of which were hypermethylated. Meanwhile, 773 transcripts expressed at lower levels were identified in PP, of which 340 (44.0%) were modified by m⁶A, 97.6% (332/340) of which were hypomethylated (Table 2, Figure 4A, and Supplementary Data 5). The comparison of PP vs. PN and PN vs. NN showed a similar pattern to the comparisons of PP vs. NN (Table 2, Supplementary Figures 7A,B, and Supplementary Data 5). Based on these results, we estimated that 19.3–48.4% of the DEGs (transcripts) were chemically modified by m⁶A. In the PP vs. NN and PP vs. PN comparisons, the proportion of transcripts expressed at lower levels that were modified by m⁶A was higher than that of highly expressed transcripts (46.1 vs. 23.7%, $P < 0.001$), suggesting that transcripts expressed at lower levels were more preferentially modified by m⁶A. These results also showed that the upregulation of gene expression was often accompanied by the upregulation of m⁶A methylation ($r = 1$, $P < 0.05$). Furthermore, we classified these m⁶A-containing DEGs by their peak position and examined how their mRNA expression levels correlated with the locations of m⁶A peaks. Our analysis showed that, irrespective of whether the m⁶A peaks were located at a 5'UTR, CDS, intron, or 3'UTR, their extent of methylation was always positively correlated with the mRNA expression levels ($P < 0.001$) (Figure 4B, Supplementary Figure 8, and Supplementary Data 6).

To compare the mRNA expression levels of methylated and unmethylated transcripts, we divided the whole transcriptome genes according to whether they were modified by m⁶A. However, we found that, in the PP vs. NN comparison, among m⁶A methylated genes, more were upregulated than downregulated; among non-m⁶A methylated genes, more were

downregulated than upregulated. This trend was consistent with the PP vs. PN and PN vs. NN comparisons (Figure 4C and Supplementary Figure 9).

DISCUSSION

Extensive m⁶A-related studies have been performed in malignant tumors, nervous system diseases, hematopoietic diseases, metabolic diseases, viral diseases, and injuries, but no reports have been published on m⁶A profiling in patients with psoriasis vulgaris (Brocard et al., 2017; Zhang et al., 2017; Sun et al., 2019). In this study, two sequencing libraries, namely, m⁶A-Seq library (IP) and RNA-Seq library (input), were constructed for high-throughput MeRIP-Seq to examine transcriptome-wide m⁶A methylation patterns in a total of 12 skin samples, four each from PP, PN, and NN biopsies. Our data showed that m⁶A is highly conserved across psoriasis vulgaris and healthy controls. Nevertheless, there are differences among PP, PN, and NN, with ~5,624 fewer m⁶A peaks identified in PP samples than in PN samples.

m⁶A-modified nucleotides were previously shown to be widely distributed in human tissues including the liver, kidney, brain, lung, and heart (Dominissini et al., 2012; He et al., 2019; Lan et al., 2019; Mathiyalagan et al., 2019). Here, we detected 1.021, 1.254, and 1.176 m⁶A peaks per gene in skin samples of PP, PN, and NN, respectively. These results further affirm that m⁶A is a universal form of RNA modification in human tissues. However, the levels of m⁶A methylation in skin samples of PP, PN, and NN were lower than the previous estimations of approximately ~1.7 m⁶A residues per gene in a human hepatocellular carcinoma cell line (HepG2) and ~3 m⁶A residues per average mRNA transcript in mammalian cells (Dominissini et al., 2012). The differences in proportions of m⁶A-modified transcripts may be due to the different tissue types. The pattern of adenosine methylation in mRNAs of PP, PN, and NN is consistent with that reported in mouse brain, in that a few of the m⁶A peaks are clustered together while most of them are single (Meyer et al., 2012). However, the proportions of single peaks of PP, PN, and NN (59.68, 52.95, and 55.25%) are even higher than that of mouse brain (46%) (Meyer et al., 2012). We assumed that this difference might be associated with the different tissue types and gene structure characteristics.

Comparative analysis among groups showed that PP and NN contained the fewest shared m⁶A peaks, while PN and NN carried the most. In addition, DMR analysis revealed that the greatest differences were observed in the PP vs. NN and PP vs.

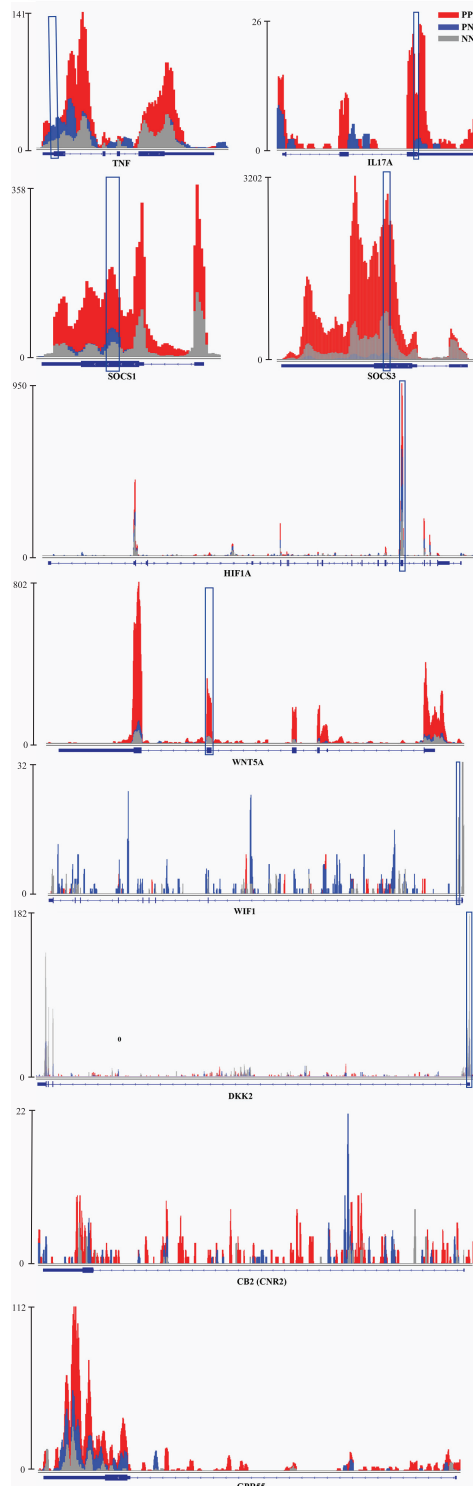


FIGURE 3 | Examples of genes with m⁶A peaks in PP, PN, and NN. Integrative Genomics Viewer (IGV) tracks showed MeRIP-seq reads distribution in TNF, IL-17A, SOCS1/3, HIF1A, WNT5A, WIF1, DKK2, CNR2, and GPR55 mRNA of PP, PN, and NN. Different colors illustrate the accumulation of m⁶A-IP reads. The blue squares indicate the verified sequence fragments.

PN comparisons, instead of PN and NN, suggesting that the difference in m⁶A methylation between PP and NN was greater than that between PN and NN. Compared with PN or NN, the number of hypomethylated DMRs in PP was much greater than that of hypermethylated DMRs, suggesting that DMRs in PP were preferentially hypomethylated. These hypermethylated m⁶A peaks in PP were mainly enriched in CDSs and 3'UTRs, which is consistent with previous studies (Meyer et al., 2012). However, the hypomethylated m⁶A peaks in PP were not only enriched in CDSs and 3'UTRs, but also highly enriched in 5'UTRs. Because 5'UTRs play a major role in controlling translation efficiency and shaping the cellular proteome (Hinnebusch et al., 2016), we assume that the hypomethylated DMRs in PP may be involved in protein translation and shaping.

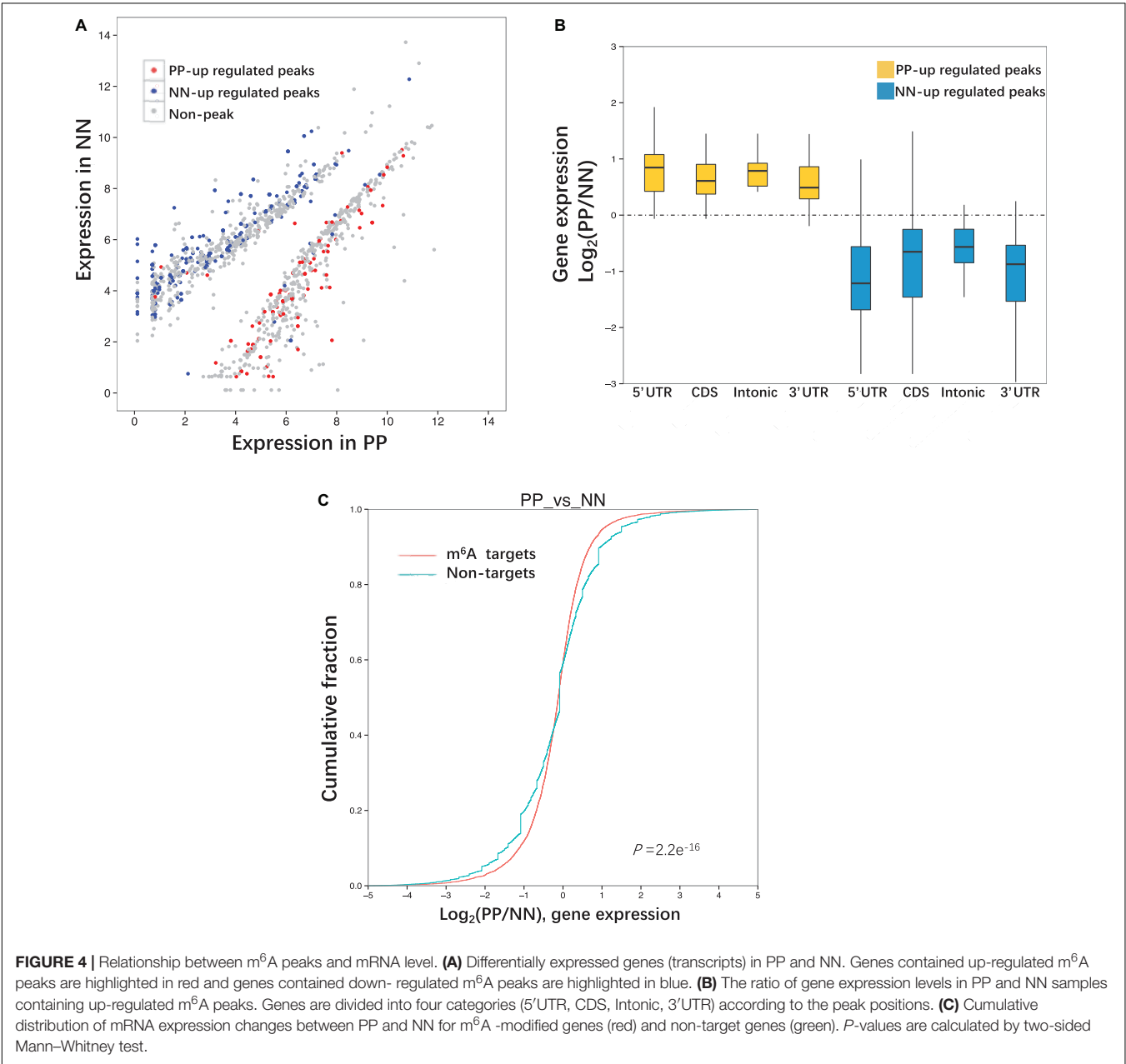
GO analysis revealed that, compared with PN or NN, hypermethylated DMRs in PP were mainly linked to response-associated terms (e.g., response to other organisms, inflammatory response, and response to external/biotic stimulus) and cytokine production and regulation. Because psoriasis is often triggered by internal or external environmental stimuli and persists due to cross-talk between keratinocytes and immunocytes, which mediates the production of cytokines, chemokines, and growth factors (Hugh and Weinberg, 2018), it is suggested that m⁶A methylation may be involved in the important pathogenic processes of psoriasis, including triggering the disease and maintaining inflammation. Distinct from hypermethylated DMRs in PP, the hypomethylated DMRs in PP were mainly associated with development-related terms, suggesting that hypomethylated DMRs in PP may participate in development-related processes. In addition, compared with NN, highly methylated DMRs in PP were specifically enriched in the G-protein-coupled receptor signaling pathway. Studies have reported that cannabinoid type 2 receptor (CB2) and G-protein-coupled receptor 55 (GPR55), as G-protein-coupled receptors, were both increased in psoriasis vulgaris and could attenuate oxidative stress and act in an anti-inflammatory process (Ambrozewicz et al., 2018; **Figure 3**). Another study reported that cannabinoid receptors control the proliferation/differentiation and immune activity of keratinocytes (Tóth et al., 2019). While both CB2 and GPR55 were modified by m⁶A, there was no significant difference among PP, PN and NN in our study (**Figure 3**). Meanwhile, hypomethylated DMRs in PP were particularly associated with the polyol metabolic process, which is one of the major biochemical pathways involved in the development of diabetic macroangiopathy and peripheral neuropathy (Katakami, 2018). It has been shown that there is a significant correlation between psoriasis/diabetes and the polyol metabolic process (Mamizadeh et al., 2019), so we consider that these hypomethylated DMRs in PP may provide a clue to explain why psoriasis is related to diabetes.

Interestingly, KEGG pathway analysis revealed that, compared with NN, highly methylated DMRs in PP were mainly associated with olfactory receptor genes. Ahn et al. (2016) reported that modules linked to psoriasis were mainly associated with olfactory signaling, as assessed by RNA-Seq. Olfactory receptors are known to be expressed not only in nasal tissue but also in skin tissue and specifically in keratinocytes, dendritic cells, and melanocytes

TABLE 2 | Relationship between m⁶A peaks and mRNA level.

	PP vs. NN	PP vs. PN	PN vs. NN
Highly expressed DEGs (transcripts)	511	1,330	409
Highly expressed transcripts modified by m ⁶ A	22.7% (116/511)	24.1% (320/1,330)	23.2% (95/409)
Hypermethylated highly expressed transcripts	87.9% (102/116)	81.9% (262/320)	94.7% (90/95)
Lower expressed DEGs (transcripts)	773	696	498
Lower expressed transcripts modified by m ⁶ A	44.0% (340/773)	48.4% (337/696)	19.3% (96/498)
Hypomethylated lower expressed transcripts	97.6% (332/340)	94.3% (318/337)	91.7% (88/96)

DEGs, differentially expressed genes.



(Ahn et al., 2016). In addition, Aydın et al. (2016) found that the olfactory function of patients with psoriasis was significantly worse than that of healthy controls (*P* < 0.001). These results

suggest that m⁶A methylation may be involved in the mechanism of anosmia in psoriasis vulgaris. In addition, our studies showed that hypomethylated transcripts in PP were mainly associated

with the Wnt signaling pathway. The Wnt gene family is known to encode a set of highly conserved secreted signaling proteins that participate in and control cell differentiation, cell proliferation, and immune-mediated inflammatory cascade in psoriasis vulgaris (Gudjonsson et al., 2010; Dou et al., 2017; Wang et al., 2017a). This suggests that m⁶A methylation may control the critical pathogenic processes in psoriasis vulgaris, including cell differentiation, cell proliferation, and immune-mediated inflammation, by modifying expression of the Wnt gene family (Irrera et al., 2020). Besides, we found that IL-17A and TNF- α , two of the key genes of the TNF- α /IL-23/Th17 axis in psoriasis, are upregulated > 20-fold and > 3.5-fold in m⁶A methylation levels in PP vs. NN, respectively (Figure 3). This axis is widely regarded as the core process of the pathogenesis of psoriasis vulgaris, which could induce the proliferation of keratinocytes and form the inflammatory plaque of psoriasis vulgaris (Rendon and Schäkel, 2019). The rapid and efficient therapeutic response of various monoclonal antibodies against TNF- α or IL-17A strongly supports the TNF- α /IL-23/Th17 axis as being a key factor for the expansion of inflammation and the aggravation of skin lesions in psoriasis vulgaris (Rendon and Schäkel, 2019). Therefore, these findings support the notion that m⁶A may be a critical denominator that controls keratinocyte proliferation, cell differentiation, and inflammation in psoriasis vulgaris.

RNA-Seq data were used for the combined analysis of m⁶A peaks and mRNA level. Upon comparison with PN or NN, about 20% of the highly expressed genes in PP (PP-high) were chemically modified by m⁶A, while about 40% of the genes expressed at lower levels in PP (PP-low) were chemically modified by m⁶A, which was close to the findings in a previous report (over one-third) on the human brain (Dominissini et al., 2012). The estimated difference between PP-high and PP-low suggested that genes expressed at lower levels were preferentially modified by m⁶A. We extended this analysis to the whole transcriptome. However, among genes expressed at a low level in PP, the proportion of m⁶A targets was smaller than that of non-m⁶A targets; meanwhile, among highly expressed genes in PP, the proportion of m⁶A targets was greater than that of non-m⁶A targets. The reason for the results being opposite between the two analyses may be the different data ranges. The former analysis only analyzed DEGs, while the latter analysis analyzed the whole transcriptome. In previous studies on *A. thaliana*, most transcripts with a low expression level were more likely modified by m⁶A in both leaf and flower chloroplasts (Wan et al., 2015; Wang et al., 2017c). However, the root amyloplast presented the methylation feature that the moderately expressed transcripts were more likely to be methylated, and those expressed at the two extremes were less methylated by m⁶A (Wan et al., 2015; Wang et al., 2017c). These results indicated that the relationship between the level of mRNA and whether there is m⁶A modification differs among different species, organs and tissues. Furthermore, we analyzed the relationship between m⁶A level and mRNA level, and found that the upregulation of gene expression was often accompanied by the upregulation of m⁶A methylation regardless of the peak position, suggesting a possible positive relationship between the extent of m⁶A methylation and the mRNA levels. Actually, many researches have revealed that

m⁶A level was positively or negatively correlated with mRNA level (Zhao et al., 2017). Mostly, m⁶A methylation was shown to be negatively correlated with mRNA level when m⁶A methylation was linked to accelerate degradation of target transcripts (Liu et al., 2014). For example, silencing of m⁶A writers (METTL3, METTL14, or WTAP) in mammalian cells has been shown to lead to increases in the abundance of their respective target transcripts (Liu et al., 2014). Overexpression of m⁶A reader (YTHDF2) in human embryonic kidney 293T cells has been shown to lead to decreases in the abundance of the target transcripts (Wu et al., 2020). Besides, m⁶A methylation was shown to be positively correlated with mRNA level when m⁶A methylation was linked to enhance mRNA expression (Zhao et al., 2017). For example, METTL3-mediated m⁶A methylation enhanced ZMYM1 mRNA expression in gastric cancer (Yue et al., 2019). The role of m⁶A methylation in transcriptional regulation needs to be elucidated by further studies in the future.

DATA AVAILABILITY STATEMENT

The datasets generated for this study can be found in the online repositories. The names of the repository/repositories and accession number(s) can be found below: <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE155702>.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Ethics Committee of Peking Union Medical College Hospital. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

Y-NW conceived the project, performed the experiments, performed data analyses, and wrote the manuscript. H-ZJ reviewed the manuscript and made substantial contributions to the drafting process. All authors agreed to be accountable for the content of the work.

FUNDING

This study was funded by the National Natural Science Foundation of China (82073450), the National Key Research and Development Program of China (No. 2016YFC0901500), and the Chinese Academy of Medical Sciences (CAMS) Initiative for Innovative Medicine (2017-I2M-B&R-01).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fcell.2020.591629/full#supplementary-material>

REFERENCES

- Ahn, R., Gupta, R., Lai, K., Chopra, N., Arron, S. T., and Liao, W. (2016). Network analysis of psoriasis reveals biological pathways and roles for coding and long non-coding RNAs. *BMC Genomics* 17:841. doi: 10.1186/s12864-016-3188-y
- Ambrozewicz, E., Wójcik, P., Wroński, A., Łuczaj, W., Jastrząb, A., Żarković, N., et al. (2018). Pathophysiological Alterations of Redox Signaling and Endocannabinoid System in Granulocytes and Plasma of Psoriatic Patients. *Cells* 7:159. doi: 10.3390/cells7100159
- Aydın, E., Tekeli, H., Karabacak, E., Altunay, Y.K., Aydın, Ç., Çerman, A. A., et al. (2016). Olfactory functions in patients with psoriasis vulgaris: correlations with the severity of the disease. *Arch. Dermatol. Res.* 308, 409–414. doi: 10.1007/s00403-016-1662-7
- Bi, Z., Liu, Y., Zhao, Y., Yao, Y., Wu, R., Liu, Q., et al. (2019). A dynamic reversible RNA N(6)-methyladenosine modification: current status and perspectives. *J. Cell Physiol.* 234, 7948–7956. doi: 10.1002/jcp.28014
- Boehncke, W. H., and Schon, M. P. (2015). Psoriasis. *Lancet* 386, 983–994. doi: 10.1016/s0140-6736(14)61909-7
- Bovenschen, H. J., van de Kerkhof, P. C., van Erp, P. E., Woestenrenk, R., Joosten, I., and Koenen, H. J. (2011). Foxp3+ regulatory T cells of psoriasis patients easily differentiate into IL-17A-producing cells and are found in lesional skin. *J. Invest. Dermatol.* 131, 1853–1860. doi: 10.1038/jid.2011.139
- Brocard, M., Ruggieri, A., and Locker, N. (2017). m6A RNA methylation, a new hallmark in virus-host interactions. *J. Gen. Virol.* 98, 2207–2214. doi: 10.1099/jgv.0.000910
- Church, C., Moir, L., McMurray, F., Girard, C., Banks, G. T., Teboul, L., et al. (2010). Overexpression of Fto leads to increased food intake and results in obesity. *Nat. Genet.* 42, 1086–1092. doi: 10.1038/ng.713
- Dominissini, D., Moshitch-Moshkovitz, S., Salmon-Divon, M., Amariglio, N., and Rechavi, G. (2013). Transcriptome-wide mapping of N(6)-methyladenosine by m(6)A-seq based on immunocapturing and massively parallel sequencing. *Nat. Protoc.* 8, 176–189. doi: 10.1038/nprot.2012.148
- Dominissini, D., Moshitch-Moshkovitz, S., Schwartz, S., Salmon-Divon, M., Ungar, L., Osenberg, S., et al. (2012). Topology of the human and mouse m6A RNA methylomes revealed by m6A-seq. *Nature* 485, 201–206. doi: 10.1038/nature11112
- Dou, J., Zhang, L., Xie, X., Ye, L., Yang, C., Wen, L., et al. (2017). Integrative analyses reveal biological pathways and key genes in psoriasis. *Br. J. Dermatol.* 177, 1349–1357. doi: 10.1111/bjd.15682
- Engel, M., and Chen, A. (2018). The emerging role of mRNA methylation in normal and pathological behavior. *Genes Brain Behav.* 17:e12428. doi: 10.1111/gbb.12428
- Frye, M., Harada, B. T., Behm, M., and He, C. (2018). RNA modifications modulate gene expression during development. *Science* 361, 1346–1349. doi: 10.1126/science.aau1646
- Fu, Y., Dominissini, D., Rechavi, G., and He, C. (2014). Gene expression regulation mediated through reversible m(6)A RNA methylation. *Nat. Rev. Genet.* 15, 293–306. doi: 10.1038/nrg3724
- Greb, J. E., Goldminz, A. M., Elder, J. T., Lebowitz, M. G., Gladman, D. D., Wu, J. J., et al. (2016). Psoriasis. *Nat. Rev. Dis. Prim.* 2:16082. doi: 10.1038/nrdp.2016.82
- Gudjonsson, J. E., Johnston, A., Stoll, S. W., Riblett, M. B., Xing, X., Kochkodan, J. J., et al. (2010). Evidence for altered Wnt signaling in psoriatic skin. *J. Invest. Dermatol.* 130, 1849–1859. doi: 10.1038/jid.2010.67
- Hammitzsch, A., Tallant, C., Fedorov, O., O'Mahony, A., Brennan, P. E., Hay, D. A., et al. (2015). CBP30, a selective CBP/p300 bromodomain inhibitor, suppresses human Th17 responses. *Proc. Natl. Acad. Sci. U.S.A.* 112, 10768–10773. doi: 10.1073/pnas.1501956112
- He, C. (2010). Grand challenge commentary: RNA epigenetics? *Nat. Chem. Biol.* 6, 863–865. doi: 10.1038/nchembio.482
- He, Y., Xing, J., Wang, S., Xin, S., Han, Y., and Zhang, J. (2019). Increased m6A methylation level is associated with the progression of human abdominal aortic aneurysm. *Ann. Transl. Med.* 7:797. doi: 10.21037/atm.2019.12.65
- Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y. C., Laslo, P., et al. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol. Cell* 38, 576–589. doi: 10.1016/j.molcel.2010.05.004
- Hinnebusch, A. G., Ivanov, I. P., and Sonenberg, N. (2016). Translational control by 5'-untranslated regions of eukaryotic mRNAs. *Science* 352, 1413–1416. doi: 10.1126/science.aad9868
- Hu, Y., Zhu, Y., Lian, N., Chen, M., Bartke, A., and Yuan, R. (2019). Metabolic syndrome and skin diseases. *Front. Endocrinol.* 10:788. doi: 10.3389/fendo.2019.00788
- Hugh, J. M., and Weinberg, J. M. (2018). Update on the pathophysiology of psoriasis. *Cutis* 102, 6–12.
- Irrera, N., Bitto, A., Vaccaro, M., Mannino, F., Squadrito, V., Pallio, G., et al. (2020). PDRN, a Bioactive natural compound, ameliorates imiquimod-induced psoriasis through NF-κB Pathway Inhibition and Wnt/β-catenin signaling modulation. *Int. J. Mol. Sci.* 21:1215. doi: 10.3390/ijms21041215
- Katakami, N. (2018). Mechanism of development of atherosclerosis and cardiovascular disease in diabetes Mellitus. *J. Atheroscler. Thromb.* 25, 27–39. doi: 10.5551/jat.RV17014
- Kim, D., Langmead, B., and Salzberg, S. L. (2015). HISAT: a fast spliced aligner with low memory requirements. *Nat. Methods* 12, 357–360. doi: 10.1038/nmeth.3317
- Lan, Q., Liu, P. Y., Haase, J., Bell, J. L., Hüttelmaier, S., and Liu, T. (2019). The critical role of RNA m(6)A methylation in cancer. *Cancer Res.* 79, 1285–1292. doi: 10.1158/0008-5472.Can-18-2965
- Liu, J., Yue, Y., Han, D., Wang, X., Fu, Y., Zhang, L., et al. (2014). A METTL3-METTL14 complex mediates mammalian nuclear RNA N6-adenosine methylation. *Nat. Chem. Biol.* 10, 93–95. doi: 10.1038/nchembio.1432
- Mamizadeh, M., Tardeh, Z., and Azami, M. (2019). The association between psoriasis and diabetes mellitus: a systematic review and meta-analysis. *Diabetes Metab. Syndr.* 13, 1405–1412. doi: 10.1016/j.dsx.2019.01.009
- Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J* 17, 10–12.
- Mathiyalagan, P., Adamiak, M., Mayourian, J., Sassi, Y., Liang, Y., Agarwal, N., et al. (2019). FTO-Dependent N(6)-methyladenosine regulates cardiac function during remodeling and repair. *Circulation* 139, 518–532. doi: 10.1161/circulationaha.118.033794
- Mavropoulos, A., Rigopoulou, E. I., Liaskos, C., Bogdanos, D. P., and Sakkas, L. I. (2013). The role of p38 MAPK in the aetiopathogenesis of psoriasis and psoriatic arthritis. *Clin. Dev. Immunol.* 2013:569751. doi: 10.1155/2013/569751
- Meyer, K. D., Saletore, Y., Zumbo, P., Elemento, O., Mason, C. E., and Jaffrey, S. R. (2012). Comprehensive analysis of mRNA methylation reveals enrichment in 3' UTRs and near stop codons. *Cell* 149, 1635–1646. doi: 10.1016/j.cell.2012.05.003
- Pan, Y., Ma, P., Liu, Y., Li, W., and Shu, Y. (2018). Multiple functions of m(6)A RNA methylation in cancer. *J. Hematol. Oncol.* 11:48. doi: 10.1186/s13045-018-0590-8
- Pasquali, L., Srivastava, A., Meisgen, F., Das Mahapatra, K., Xia, P., Xu Landén, N., et al. (2019). The keratinocyte transcriptome in psoriasis: pathways related to immune responses, cell cycle and keratinization. *Acta Derm Venereol.* 99, 196–205. doi: 10.2340/00015555-3066
- Pollock, R. A., Abji, F., and Gladman, D. D. (2017). Epigenetics of psoriatic disease: a systematic review and critical appraisal. *J. Autoimmun.* 78, 29–38. doi: 10.1016/j.jaut.2016.12.002
- Rendon, A., and Schäkel, K. (2019). Psoriasis pathogenesis and treatment. *Int. J. Mol. Sci.* 20:1475. doi: 10.3390/ijms20061475
- Shen, L., Shao, N. Y., Liu, X., Maze, I., Feng, J., and Nestler, E. J. (2013). diffReps: detecting differential chromatin modification sites from ChIP-seq data with biological replicates. *PLoS One* 8:e65598. doi: 10.1371/journal.pone.0065598
- Sun, T., Wu, R., and Ming, L. (2019). The role of m6A RNA methylation in cancer. *Biomed. Pharmacother.* 112:108613. doi: 10.1016/j.biopha.2019.108613
- Thorvaldsdóttir, H., Robinson, J. T., and Mesirov, J. P. (2013). Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief Bioinform.* 14, 178–192. doi: 10.1093/bib/bbs017
- Tóth, K. F., Ádám, D., Bíró, T., and Oláh, A. (2019). Cannabinoid signaling in the skin: therapeutic potential of the "C(ut)annabinoid" System. *Molecules* 24:918. doi: 10.3390/molecules24050918
- Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D. R., et al. (2012). Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat. Protoc.* 7, 562–578. doi: 10.1038/nprot.2012.016

- Wan, Y., Tang, K., Zhang, D., Xie, S., Zhu, X., Wang, Z., et al. (2015). Transcriptome-wide high-throughput deep m(6)A-seq reveals unique differential m(6)A methylation patterns between three organs in *Arabidopsis thaliana*. *Genome Biol.* 16:272. doi: 10.1186/s13059-015-0839-2
- Wang, W., Yu, X., Wu, C., and Jin, H. (2017a). IL-36 γ inhibits differentiation and induces inflammation of keratinocyte via Wnt signaling pathway in psoriasis. *Int. J. Med. Sci.* 14, 1002–1007. doi: 10.7150/ijms.20809
- Wang, X., Huang, J., Zou, T., and Yin, P. (2017b). Human m(6)A writers: two subunits, 2 roles. *RNA Biol.* 14, 300–304. doi: 10.1080/15476286.2017.1282025
- Wang, Y. N., Yu, C. Y., and Jin, H. Z. (2020). RNA N(6)-Methyladenosine Modifications and the Immune Response. *J. Immunol. Res.* 2020:6327614. doi: 10.1155/2020/6327614
- Wang, Z., Tang, K., Zhang, D., Wan, Y., Wen, Y., Lu, Q., et al. (2017c). High-throughput m6A-seq reveals RNA m6A methylation patterns in the chloroplast and mitochondria transcriptomes of *Arabidopsis thaliana*. *PLoS One* 12:e0185612. doi: 10.1371/journal.pone.0185612
- World Health Organization. (2014). *Resolution WHA67.9 Psoriasis Sixty-Seventh World Health Assembly, Resolutions and Decisions*. Geneva: WHO.
- Wu, C., Chen, W., He, J., Jin, S., Liu, Y., Yi, Y., et al. (2020). Interplay of m(6)A and H3K27 trimethylation restrains inflammation during bacterial infection. *Sci. Adv.* 6:eaba0647. doi: 10.1126/sciadv.aba0647
- Yue, B., Song, C., Yang, L., Cui, R., Cheng, X., Zhang, Z., et al. (2019). METTL3-mediated N6-methyladenosine modification is critical for epithelial-mesenchymal transition and metastasis of gastric cancer. *Mol. Cancer* 18:142.
- Zhang, C., Chen, Y., Sun, B., Wang, L., Yang, Y., Ma, D., et al. (2017). m(6)A modulates haematopoietic stem and progenitor cell specification. *Nature* 549, 273–276. doi: 10.1038/nature23883
- Zhang, Y., Liu, T., Meyer, C. A., Eeckhoutte, J., Johnson, D. S., Bernstein, B. E., et al. (2008). Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* 9:R137. doi: 10.1186/gb-2008-9-9-r137
- Zhao, B. S., Roundtree, I. A., and He, C. (2017). Post-transcriptional gene regulation by mRNA modifications. *Nat. Rev. Mol. Cell Biol.* 18, 31–42. doi: 10.1038/nrm.2016.132

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 Wang and Jin. 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.



RNA-Binding Protein HuR Suppresses Inflammation and Promotes Extracellular Matrix Homeostasis via NKRF in Intervertebral Disc Degeneration

Zhenxuan Shao^{1,2,3†}, Zhuolong Tu^{4†}, Yifeng Shi^{1,2,3†}, Sunlong Li^{1,2,3}, Aimin Wu^{1,2,3}, Yaosen Wu^{1,2,3}, Naifeng Tian^{1,2,3}, Liaojun Sun^{1,2,3}, Zongyou Pan⁵, Linwei Chen⁵, Weiyang Gao^{1,2,3}, Yifei Zhou^{1,2,3*}, Xiangyang Wang^{1,2,3*} and Xiaolei Zhang^{1,2,3,6*}

OPEN ACCESS

Edited by:

Jia Meng,
Xi'an Jiaotong-Liverpool University,
China

Reviewed by:

Yukun Zhang,
Huazhong University of Science
and Technology, China
Wenbin Hua,
Huazhong University of Science
and Technology, China

*Correspondence:

Xiaolei Zhang
zhangxiaolei@wmu.edu.cn
Xiangyang Wang
xiangyangwang@wmu.edu.cn
Yifei Zhou
yifeizhou@wmu.edu.cn

[†]These authors have contributed
equally to this work

Specialty section:

This article was submitted to
Epigenomics and Epigenetics,
a section of the journal
Frontiers in Cell and Developmental
Biology

Received: 28 September 2020

Accepted: 26 October 2020

Published: 25 November 2020

Citation:

Shao Z, Tu Z, Shi Y, Li S, Wu A,
Wu Y, Tian N, Sun L, Pan Z, Chen L,
Gao W, Zhou Y, Wang X and Zhang X
(2020) RNA-Binding Protein HuR
Suppresses Inflammation
and Promotes Extracellular Matrix
Homeostasis via NKRF
in Intervertebral Disc Degeneration.
Front. Cell Dev. Biol. 8:611234.
doi: 10.3389/fcell.2020.611234

¹ Department of Orthopaedics, The Second Affiliated Hospital and Yuying Children's Hospital of Wenzhou Medical University, Wenzhou, China, ² Key Laboratory of Orthopaedics of Zhejiang Province, Wenzhou, China, ³ The Second School of Medicine, Wenzhou Medical University, Wenzhou, China, ⁴ Department of Burn, The First Affiliated Hospital of Wenzhou Medical University, Wenzhou, China, ⁵ Department of Orthopaedics, The Second Affiliated Hospital of Zhejiang University School of Medicine, Hangzhou, China, ⁶ Chinese Orthopedic Regenerative Medicine Society, Hangzhou, China

Intervertebral disc degeneration (IVDD) has been reported to be a major cause of low back pain. Studies have demonstrated that IVDD may be dysregulated at the transcriptional level; however, whether post-transcriptional regulation is involved is still unknown. The current study aimed to illustrate the role of Human antigen R (HuR), an RNA binding protein involved in post-transcriptional regulation, in IVDD. The results showed that the expression of HuR was decreased in degenerative nucleus pulposus (NP) tissues as well as in TNF- α -treated NP cells. Downregulation of HuR may lead to increased inflammation and extracellular matrix (ECM) degradation in TNF- α -treated NP cells; however, these effects were not reversed in HuR overexpressed NP cells. Inhibition of the NF- κ B signaling pathway attenuates inflammation and ECM degradation in HuR-deficient NP cells. A mechanism study showed that HuR prompted NKRF mRNA stability via binding to its AU-rich elements, and upregulation of NKRF suppressed inflammation and ECM degradation in HuR-deficient NP cells. Furthermore, we found that NKRF, but not HuR, overexpression ameliorated the process of IVDD in rats *in vivo*. In conclusion, HuR suppressed inflammation and ECM degradation in NP cells via stabilizing NKRF and inhibiting the NF- κ B signaling pathway; NKRF, but not HuR, may serve as a potential therapeutic target for IVDD.

Keywords: intervertebral disc degeneration, HuR, inflammation, extracellular matrix, NKRF

INTRODUCTION

Low back pain is a common disorder in the population, and up to 25% of people in the United States suffer from back and neck pain (Martin et al., 2008; Hoy et al., 2010; Zheng et al., 2018). Severe low back pain can lead to disability, which not only affects the quality of life for individuals but also causes a serious burden to society; the socio-economic burden caused by low back pain has been reported to exceed 85 billion dollars (Martin et al., 2008). Studies have shown that up to 40% of lower back pain is related to intervertebral disc degeneration (IVDD) (Freemont, 2009).

The intervertebral disc is the soft tissue located between the bony vertebrae, and it absorbs the load transmitted by the vertebral body and provides flexibility for the spine. The intervertebral disc is composed of three parts: cartilage endplate, annulus fibrosus (AF), and nucleus pulposus (NP). The jelly-like NP tissue is the major tissue of intervertebral disc to combat against loading, while the dysfunction of NP cells is considered to be the initiating cause of IVDD (Sakai and Grad, 2015; Gorth et al., 2020).

The function of cells can be regulated at various levels, such as transcriptional and post-transcriptional (Uchida et al., 2019). In the previous studies, we found that transcription factor BRD4 is overexpressed in degenerated NP cells, and it may promote matrix metalloproteinase 13 (MMP13) expression *via* MAPK and NF- κ B signaling pathway (Wang et al., 2019). We also showed that the nuclear localization of transcription factor EB (TFEB) was declined in NP cells during IVDD, which caused blockage of autophagy flux and increased apoptosis (Zheng et al., 2019). Also, dysfunction of SOX9 and FOXO has been implicated in IVDD (Gruber et al., 2005; Alvarez-Garcia et al., 2017). These studies reveal that the pathogenesis of IVDD may occur at the transcriptional level; however, whether post-transcriptional is involved is still unclear.

Post-transcriptional regulation refers to the regulation of gene expression at the post-transcriptional level, such as RNA stabilization, RNA cleavage, RNA editing, RNA splicing, and translation (Zhao et al., 2017). RNA-binding protein is a type of post-transcriptional regulation protein, which can bind to RNA and participate in the functional regulation (Glisovic et al., 2008). Human antigen R (HuR), encoded by the ELAVL1 gene, is one of the first identified RNA binding proteins. HuR can bind to the adenine-uracil-rich elements (ARE) at the 3'UTR of the target mRNA molecule and regulates a variety of biological processes (Schultz et al., 2020). Pan et al. (2019) found that HuR may regulate extracellular matrix (ECM) gene expression and pH homeostasis in physiological hypoxia condition in NP cells; however, the role of HuR in pathogenesis of IVDD is still unknown.

In the current study, we evaluated the expression of HuR in NP cells during IVDD; we also upregulated and downregulated its expression *via* lentivirus to explore the role of HuR in IVDD; inflammatory genes were screened to demonstrate its working mechanism; HuR and NKR1F overexpressing lentiviruses were injected into NP tissues to assess their therapeutic effect on IVDD.

MATERIALS AND METHODS

Ethics Statement

All surgical interventions, treatments, and postoperative animal care procedures were performed in strict accordance with the Guide for the Care and Use of Laboratory Animals of the National Institutes of Health and the Animal Care and Use Committee of Wenzhou Medical University (WYDW2020-0560). Human tissue collection and treatments were also permitted by the Second Affiliated Hospital and Yuying

Children's Hospital of Wenzhou Medical University Ethics Committee (LCKY2020-135).

Human NP tissues of different grades were isolated from patients, to compare the expression of ELAVL1 by quantitative PCR (qPCR). The work was given official approval by the Ethics Committee of the Second Affiliated Hospital of Wenzhou Medical University, and we obtained written informed consent from patients or relatives prior to tissue collection.

Reagents and Antibodies

Dimethyl sulfoxide (DMSO), citrate buffer, and collagenase type II were purchased from Sigma-Aldrich (St. Louis, MO, United States). The cell culture reagents were obtained from Gibco (Grand Island, NY, United States). Primary antibodies against IL-1, TNF- α , p65, GAPDH, and Lamin B were obtained from Proteintech (Wuhan, China); primary antibodies against matrix metalloproteinase 3 (MMP3), MMP13, IL-6, Aggrecan, COL2, and NKR1F were acquired from Abcam (Cambridge, MA, United States); and primary antibodies against HuR (ELAVL1), p-p65, and I κ B α were purchased from Cell Signaling Technology (Beverly, United States). Goat anti-rabbit and anti-mouse IgG-HRP antibodies were purchased from Bioworld (Nanjing, China). Alexa Fluor® 488- and Alexa Fluor® 594-conjugated secondary antibodies were obtained from Jackson ImmunoResearch (West Grove, PA, United States). 4',6-Diamidino-2-phenylindole (DAPI) was obtained from Beyotime (Shanghai, China).

Culture of NP Cells and Extraction

Rat NP cells were extracted from healthy NP of young Sprague-Dawley rats. NP tissues were cut up into 1 mm³ and washed for three times with phosphate-buffered saline (PBS) before digestion with 0.25% type II collagenase at 37°C for 2 h. After centrifugation, the cell suspensions were cultured in DMEM/F12 (Gibco) with 10% fetal bovine serum (FBS; Gibco) and antibiotics (1% streptomycin/penicillin) in the incubator at 5% CO₂ at 37°C.

Quantitative Real-Time PCR (qPCR)

After treatment, total RNA was extracted from NP cells by the TRIzol method (Invitrogen). One thousand nanograms of total RNA was reverse transcribed to synthesize cDNA (MBI Fermentas, Germany) using the PrimeScript-RT reagent kit (Takara, Japan) and the CFX96 Real-Time PCR System (Bio-Rad Laboratories, Berkeley, CA, United States). Amplification of the cDNA was performed by SYBR Premix Ex Taq using the CFX96 Real-Time PCR System (Bio-Rad Laboratories, Berkeley, CA, United States). The cycle threshold (Ct) values were determined and normalized to the level of a housekeeping gene (GAPDH). The expression of the target genes in different groups was evaluated using the $2^{-\Delta\Delta C_t}$ method. The forward and reverse primer sequences are provided in **Supplementary Table S1**.

Western Blot (WB) Analysis

Proteins from NP cells were lysed in radioimmunoprecipitation assay (RIPA) buffer (Beyotime) with 1 mM phenylmethanesulfonyl fluoride (PMSF) (Beyotime) and then centrifuged for

20 min at 12,000 rpm at 4°C. The protein concentration was measured by the BCA protein assay kit (Beyotime). Forty nanograms of protein was separated via 8–12% (w/v) sodium dodecyl sulfate polyacrylamide gel electrophoresis and blotted onto polyvinylidene fluoride membranes (Bio-Rad, Hercules, CA, United States). After blocking with 5% non-fat milk for 2 h, the membranes were incubated with primary antibodies against HuR (1:1000), IL-1 (1:800), TNF- α (1:800), p65 (1:1000), MMP3 (1:1000), MMP13 (1:1000), IL-6 (1:800), Aggrecan (1:1000), COL2 (1:1000), NKRF (1:1000), p-p65 (1:800), IkBa (1:800), Lamin B (1:2000), and GAPDH (1:3000) overnight at 4°C, and then subsequently incubated with the respective secondary antibodies for 2 h at room temperature. After washing three times with Tris-buffered saline with Tween® 20, the blots were visualized by electrochemiluminescence plus reagent (Invitrogen). The signals were visualized using the ChemiDoc™ XRS + Imaging System (Bio-Rad, Hercules, CA, United States), and the band densities were quantified with Image Lab 3.0 software (Bio-Rad, Hercules, CA, United States).

Immunofluorescence

Rat NP cells were plated in a six-well plate and treated as described above. After treatment, the NP cells were washed with ice-cold PBS, fixed with 4% (v/v) paraformaldehyde for 15 min, and permeated using 0.1% Triton X-100 diluted in PBS for 10 min. Then, the cells were blocked with 5% bovine serum albumin for 1 h at 37°C and then incubated with a primary antibody at 4°C overnight. The next day, the cells were incubated with Alexa Fluor® 488- or Alexa Fluor® 594-conjugated secondary antibodies (1:300) for 1 h at room temperature and labeled with DAPI for 5 min. Finally, five fields from each slide were chosen randomly for microscopic observation with a Nikon ECLIPSE Ti microscope (Nikon, Japan), and the fluorescence intensity was measured using ImageJ software 2.1 (Bethesda, MD, United States) by observers who were blinded to the experimental groups.

Lentivirus and siRNA Transfection

The NP cells reaching 40–60% confluence were transfected using LV-siHuR lentivirus (from Cell-land, China, TACCAGTTTCAATGGTCATAA), at a multiplicity of infection (MOI) of 30. The NP cells reaching 40–60% confluence were transfected using LV-HuR lentivirus (from OBiO, China, NM_001108848), at a MOI of 50. The NP cells reaching 40–60% confluence were transfected using LV-NKRF lentivirus (from GeneChem, China, XM_003752133), at a MOI of 40. After 12 h of transfection, the culture medium was changed every other day. When confluent, the transfected NP cells were passaged for further experiments.

RIP

RNA-binding protein immunoprecipitation (RIP) can be used to identify specific RNA molecules, which are associated with targeted proteins. The immunoprecipitation of endogenous ribonucleoprotein complexes was analyzed using Imprint® RNA Immunoprecipitation Kit (Sigma) according to the manufacturer's instructions.

Puncture-Induced Rat IVDD Model

Adult male Sprague-Dawley rats (200–250 g, $n = 32$) used for this study were randomly obtained from the Experimental Animal Institute of Wenzhou Medical University and were randomly divided into four groups: Control group, IVDD group, IVDD + LV-NKRF group, and IVDD + LV-HuR group. The experimental level rat tail disc (Co7/8) was located by digital palpation on the coccygeal vertebrae and confirmed by counting the vertebrae from the sacral region in a trial radiograph. The tail skin was sterilized with iodinated polyvinylpyrrolidone, and needles (27G) were used to puncture the whole layer of AF through the tail skin. To make sure the needle is not be punctured too deep, the length of the needle was decided according to the AF and NP dimensions, which were measured in the preliminary experiment about 5 mm. Subsequently, 5 μ l of lentivirus (virus type: LV-NC, LV-NKRF, or LV-HuR; virus concentration: 10^9 TU/ml) was injected into the NP cavity by using a microliter syringe through that 27 G needle (Zheng et al., 2019; Chen Y. et al., 2020). All animals were allowed free unrestricted weight bearing and activity and were monitored every day to ensure their well-being.

X-Ray Image Acquisition and Magnetic Resonance Imaging (MRI)

Four weeks after surgery, X-ray images were obtained for all animals. The rats were fixed in a prone position in the X-ray irradiation system (Kubtec Medical Imaging, United States). The disc height index (DHI) was determined using the published method.

The rats' coccyx was analyzed in sagittal T2-weighted images using a 3.0-T clinical magnet (Philips Intera Achieva 3.0 MR). T2-weighted sections in the sagittal plane were obtained with the previous settings (Zheng et al., 2019). The magnetic resonance images (MRIs) were evaluated by a blinded orthopedic researcher using the Pfirrmann MRI grading system (Pfirrmann et al., 2001).

Histopathological Analysis and Immunohistochemical Examination

The rats were killed by an intraperitoneal overdose of 4% pentobarbital, and the tails were collected. The specimens were decalcified, fixed in formaldehyde, dehydrated, and embedded in paraffin. Then, the tissues were cut into 5- μ m sections. Slides of each disc were stained with hematoxylin and eosin (H&E) and Safranin O-fast green (S-O). The cellularity and morphology of the intervertebral disc were examined by a separate group of experienced histological researchers in a blinded manner using a microscope (Olympus Inc., Tokyo, Japan).

After deparaffinization, sections of sample for immunohistochemistry were incubated with 3% H₂O₂ for 10 min and washed by PBS for three times. Then, the sections were incubated with 0.1% trypsin for 20 min and washed by PBS for three times. Sections were blocked with 1% (w/v) goat serum albumin for 1 h at 37°C, followed by primary antibody (1:200) incubation at 4°C overnight. Negative control sections were incubated with non-specific IgG. Next, the sections were washed

three times with PBS and incubated with HRP-conjugated secondary antibodies for 1 h at 37°C. At least three sections from each specimen were observed. The rate of positive cells each section was quantitated by observers who were blinded to the experimental groups.

Statistical Analysis

Data are presented as the means \pm standard deviation (SD) from three independent experiments. The data were analyzed via GraphPad Prism (United States). Statistical differences between two groups were determined using two-tailed unpaired Student's *t* test, or two-tailed non-parametric Mann–Whitney test. In multiple comparisons, multiple groups were performed with one-way analysis of variance (ANOVA) and Tukey's *post hoc* test. *p* values < 0.05 were considered statistically significant.

RESULTS

The Expression of HuR Is Decreased in Degenerative NP Tissues and TNF- α -Treated NP Cells

Post-transcriptional regulation is the control of gene expression at the RNA level, which contributes substantially to gene expression regulation across tissues. In order to explore the post-transcriptional regulators that are involved in the process of IVDD, we screened data from the GEO DataSets¹. A typical set of data (Series Accession: GSE23130) was shown as heatmap in **Figure 1A**. From the heatmap, it was implied that ELAVL1(HuR) was the most differentially expressed post-transcriptional regulator during IVDD process. To further verify the results, two more datasets (GSE59485 and GSE15227) were included. Venn diagram showed that ELAVL1(HuR) was one of commonly downregulated genes during IVDD (**Figure 1B**). Interestingly, it was found from dataset GSE27494 that ELAVL1(HuR) gene level was also depressed in NP cells under inflammatory factor stimulation (**Figure 1C**). In short, the GEO DataSets data showed that ELAVL1(HuR) expression is decreased in degenerative NP tissues and NP cells under inflammatory factor stimulation.

In order to further certify that ELAVL1(HuR) expression is decreased in IVDD, we collected NP tissues from IVDD patients (Pfirschnann grade II, III, and IV) to detect HuR expression by qPCR (**Figure 1D**). The qPCR results showed that ELAVL1(HuR) expression was decreased as the disc degeneration grade rises (**Figure 1D**). In addition, we established a puncture-induced rat IVDD model to detect HuR expression by immunofluorescence, and the results showed that HuR expression was also decreased in rat degenerative NP tissues (**Figure 1F**). In short, these results showed that HuR expression was decreased in human degenerative NP tissues and rat degenerative NP tissues.

Furthermore, we isolated NP cells from normal human NP tissues and treated them with ascending concentrations and duration of TNF- α . The results showed that HuR expression was decreased in a time- and dose-dependent manner (**Figure 1E**).

HuR Regulates Inflammation in TNF- α -Treated NP Cells

To explore the effects of HuR on IVDD *in vitro*, we regulated HuR expression in NP cells by transfecting with lentivirus-shHuR (LV-shHuR) and lentivirus-HuR (LV-HuR), and the lentivirus transfection efficiency was confirmed by PCR and Western blot (WB) (**Supplementary Figures S2A,B**). Although there are two reacting bands of HuR in HuR-overexpressing cells, it may not affect the subsequent experiments; similar phenomenon could be found in previous studies (Katsanou et al., 2005; Kawagishi et al., 2013).

Next, we detected the expressions of IL-6, IL-1 β , TNF- α , and iNOS (by qPCR, WB, and ELISA), which are commonly used as indicators of inflammation, in NP cells. The qPCR results showed that the expression of IL-6, IL-1 β , TNF- α , and iNOS were expectedly increased under TNF- α treatment; interestingly, proinflammatory mediators were higher in the HuR-deficient and TNF- α -treated group (LV-shHuR + TNF- α) than those in the TNF- α -treated alone group (**Figure 2A**). Meanwhile, the WB and ELISA results also confirmed the qPCR results (**Figures 2B,C**). These results indicate that downregulation of HuR promotes inflammation in TNF- α -treated NP cells.

Downregulation of HuR promotes inflammatory response in TNF- α -treated NP cells; however, to our surprise, HuR upregulation could not inhibit inflammation in TNF- α -treated NP cells. The results of qPCR, WB, and ELISA showed that the expressions of IL-6, IL-1 β , TNF- α , and iNOS in HuR overexpression and the TNF- α -treated group (LV-HuR + TNF- α) were not lower than those in the TNF- α group (**Figure 2**).

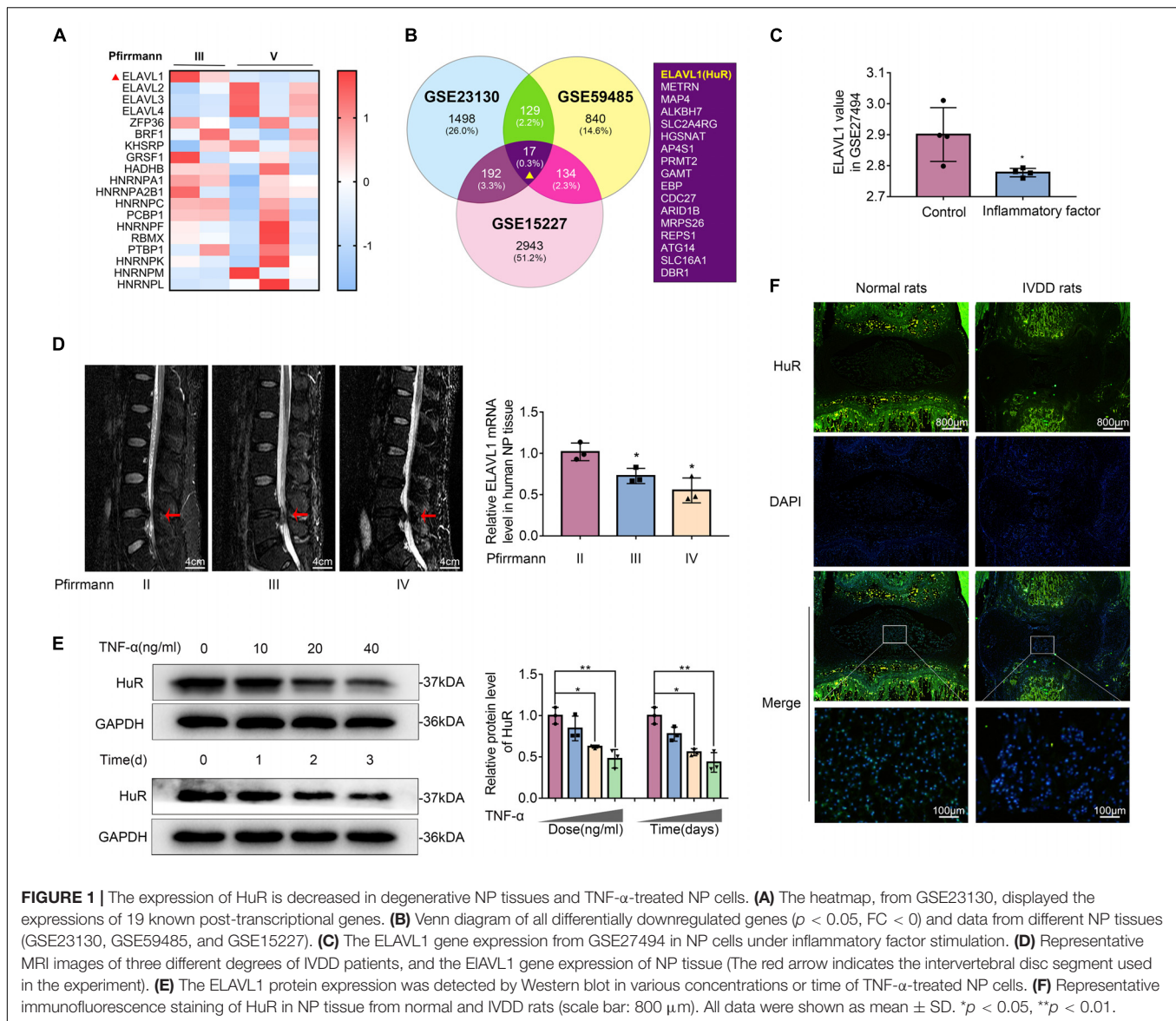
HuR Regulates ECM Metabolism in TNF- α -Treated NP Cells

Extracellular matrix degradation, induced by excessive catabolism and inadequate anabolism, is considered to be one of the most prominent feature in NP cells during IVDD (Hua et al., 2017). We then analyzed Aggrecan and collagen II (COL2) expression to determine ECM anabolism and the expression of MMP3 and MMP13 to determine ECM catabolism in HuR-overexpressing and knockdown NP cells. The qPCR and Western blotting results showed that TNF- α treatment significantly reduced the expression of Aggrecan and COL2, while it increased the expressions of MMP3 and MMP13 (**Figure 2D** and **Supplementary Figure S3**); interestingly, compared with the TNF- α treatment alone group, the expressions of Aggrecan and COL2 were decreased, while the expressions of MMP3 and MMP13 were further increased in the TNF- α + LV-shHuR group. Meanwhile, immunofluorescence and ELISA analyses of COL2 and MMP13 also confirmed the phenomenon (**Figures 2D,F**). To our surprise, HuR upregulation could not maintain ECM homeostasis in the TNF- α -treated NP cell (**Figure 2**).

HuR Deficiency Activates NF- κ B Signaling Pathway in TNF- α -Treated NP Cells

In order to explore the working mechanism of HuR on IVDD in NP cells, we analyzed genes that have been reported to be

¹<https://www.ncbi.nlm.nih.gov/gds>



regulated by HuR in previous studies. The study by Mukherjee et al. (2011) was included into KEGG pathway analysis. The results showed that various pathways were involved, among which NF- κ B signaling pathway is most likely to participate in the regulatory effects of HuR on inflammation in IVDD (Figure 3A). The NF- κ B signaling pathway is a key regulator of inflammation (Liu T. et al., 2017); and evidences demonstrated that inhibition of NF- κ B signaling pathway can delay the progression of IVDD (Liu T. et al., 2017; Chen F. et al., 2020). Thus, we focus on the NF- κ B signaling pathway in further studies.

We evaluated the NF- κ B signaling pathway by WB, and the results showed that the NF- κ B signaling pathway was dramatically activated in TNF- α -treated NP cells, as indicated by the p-p65/p65 ratio, the I κ B α expression in total cell, and the p65 expression in nuclear (Figure 3C); interestingly, the NF- κ B signaling pathway was further activated in the

LV-shHuR + TNF- α group (Figure 3C). These results were also confirmed by immunofluorescence (Figure 3B).

To determine whether inhibiting NF- κ B signaling pathway may suppress inflammation and ECM degradation caused by HuR deficiency, NP cells were pre-treated with the classical NF- κ B inhibitor QNZ (EVP4593). The WB results showed that there were significantly higher levels of inflammation and ECM degradation in HuR-deficient NP cells, while this phenomenon was attenuated by QNZ treatment (Figures 3D,E). The ECM results were also confirmed by immunofluorescence (Figure 3F).

HuR Regulates NF- κ B Signaling Pathway Through NKRF in TNF- α -Treated NP Cells

Next, we asked through which protein does HuR exert its inhibitory effect on NF- κ B signaling pathway. We tested known

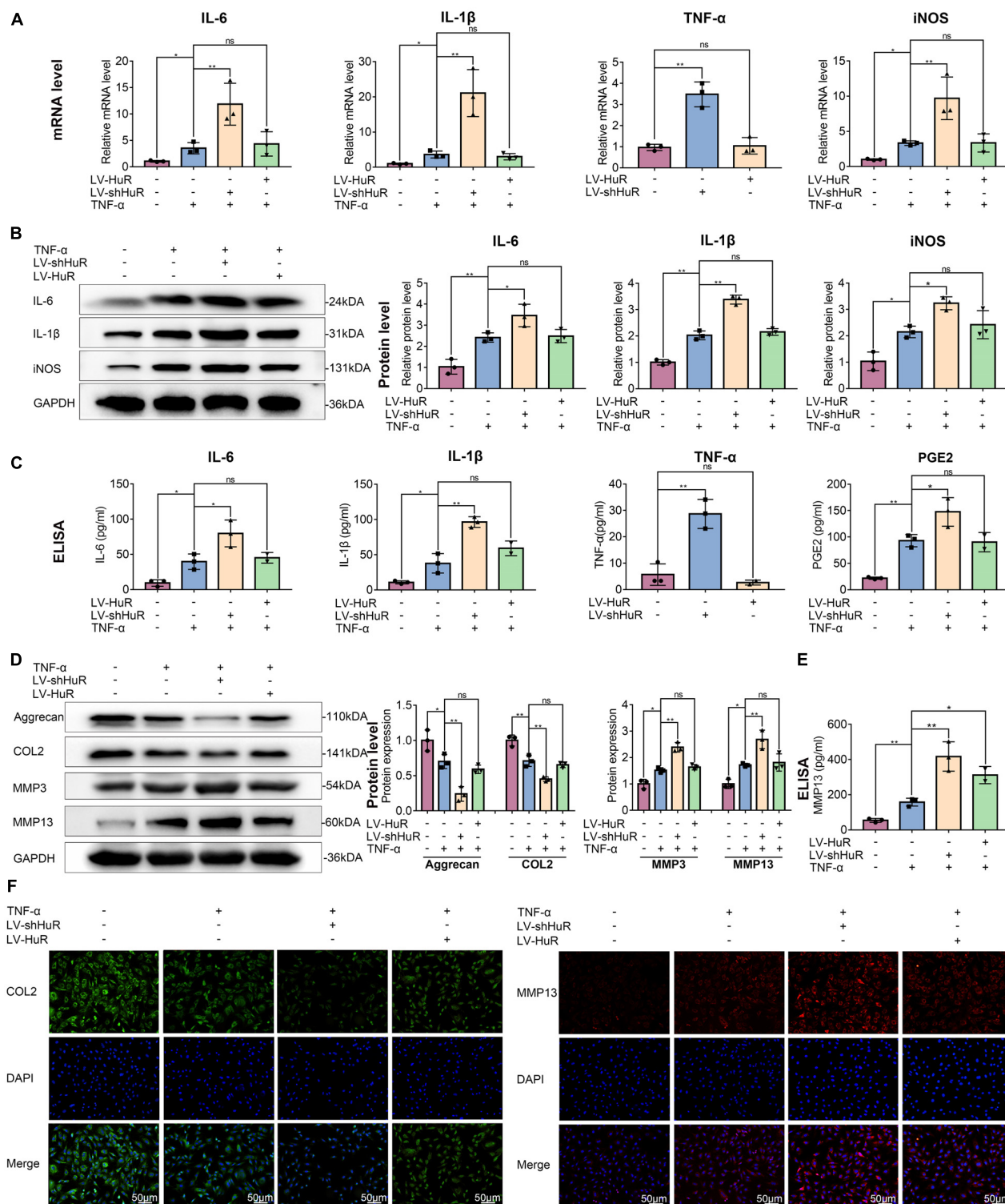


FIGURE 2 | HuR regulates inflammation and ECM degradation in TNF- α -treated NP cells. The cells were transfected with LV-NC, LV-shHuR, or LV-HuR before TNF- α treatment. **(A)** The gene expression of inflammatory cytokines, such as IL-6, IL-1 β , TNF- α , and iNOS, was detected, with or without TNF- α treatment and added lentivirus. **(B,C)** The protein expressions of inflammatory cytokines were detected by Western blot and ELISA. **(D)** The protein expressions of ECM biosynthesis and ECM breakdown proteins, such as Aggrecan, COL2, MMP3, and MMP13, were detected by Western blot. **(E)** The expressions of MMP13 detected by ELISA. **(F)** The COL2 and MMP13 were detected by immunofluorescence combined with DAPI staining for nuclei (scale bar: 50 μ m). All data were shown as mean \pm SD ($n = 3$). * $p < 0.05$, ** $p < 0.01$.

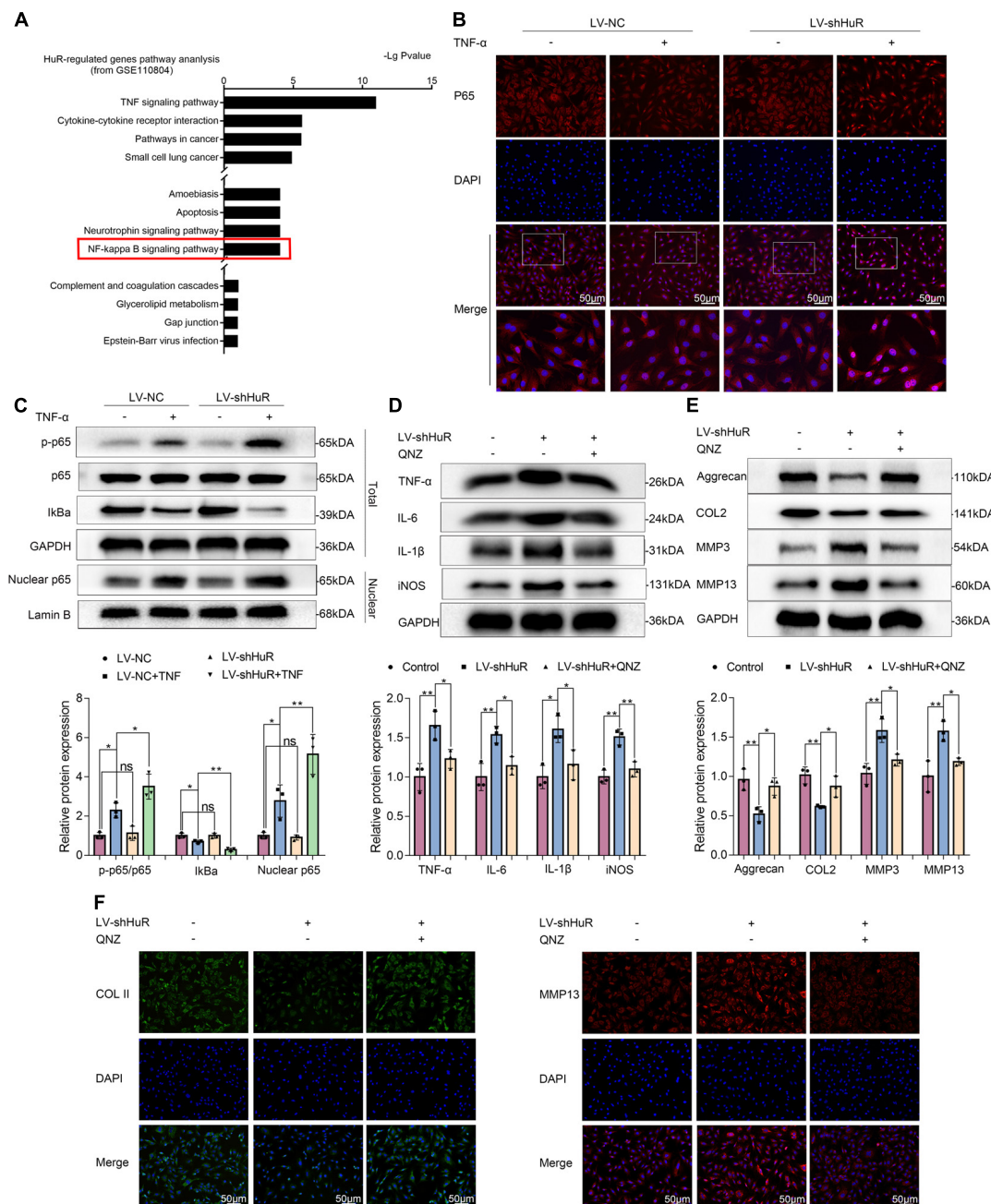
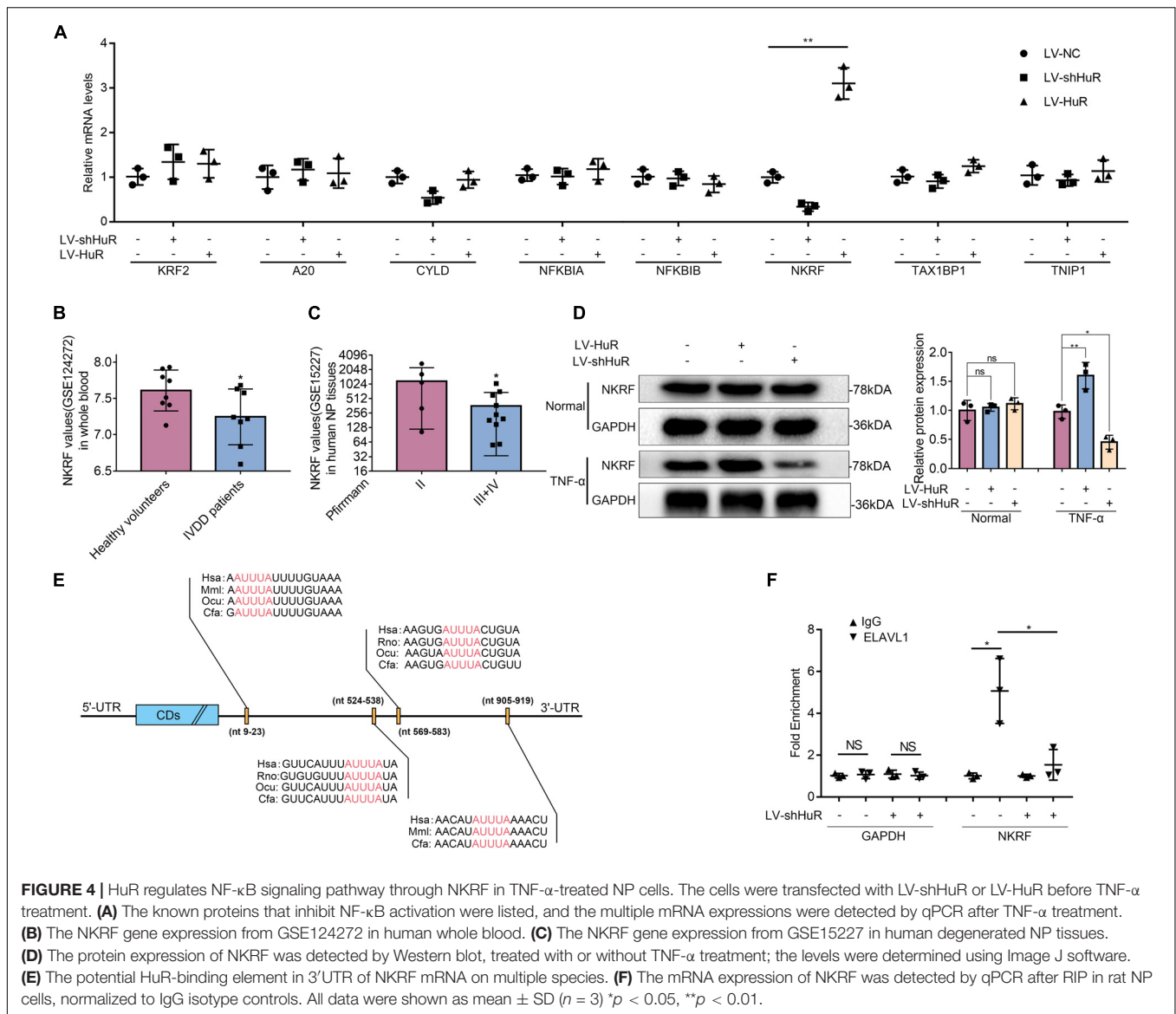


FIGURE 3 | HuR regulates NF- κ B signaling pathway in TNF- α -treated NP cells. The cells were transfected with LV-shHuR before TNF- α treatment. **(A)** The KEGG pathway analysis included HuR-regulated genes (top 1000), from GSE110804, and the regulation of NF- κ B signaling pathway was put forward. **(B)** The p65 was detected by immunofluorescence combined with DAPI staining for the nuclei (scale bar: 50 μ m). **(C)** The protein expression of p-p65, p65, I κ B α , and nucleus p65 was detected by Western blot in the NPCs, with or without TNF- α treatment and added lentivirus. **(D)** The protein expression of IL-6, IL-1 β , TNF- α , and iNOS was detected by Western blot, with or without LV-shHuR lentivirus and added NF- κ B signaling pathway inhibitor QNZ. **(E)** The protein expression of Aggrecan, COL2, MMP3, and MMP13 was detected by Western blot. **(F)** The COL2 and MMP13 were detected by immunofluorescence combined with DAPI staining for nuclei (scale bar: 50 μ m). All data were shown as mean \pm SD ($n = 3$). * $p < 0.05$, ** $p < 0.01$.

proteins that may inhibit NF- κ B activation by qPCR, which include KLF2 (Jha and Das, 2017), A20 (Wertz et al., 2004), Cyld (Trompouki et al., 2003), NFKBIA (Rinkenbaugh and Baldwin, 2011), NFKBIB (Chapman et al., 2007), NKR (Nourbakhsh and Hauser, 1999), Tax1bp1 (Shembade et al., 2011), and

Tnfr1 (Huang et al., 2008). The results showed that the expression of NKR was the most significantly affected by HuR expression (Figure 4A).

The data from the GEO DataSets also showed that the expression of NKR was decreased in both blood and



intervertebral disc tissues of IVDD patients (Figures 4B,C). Meanwhile, we evaluated the expression of NKRF under TNF- α treatment condition in NP cells by WB; it was shown that the expression of NKRF was increased in HuR-overexpressing NP cells, whereas it decreased in HuR knockdown NP cells. We also assessed the expression of NKRF in NP cells without TNF- α treatment; it showed that NKRF was not affected by HuR expression in TNF- α untreated NP cells (Figure 4D).

Previous studies have shown that HuR may exert its effect on gene expression through stabilizing mRNAs in a variety of biological processes (Lopez De Silanes et al., 2004). Bioinformatic study implied that there are multiple HuR binding sites at the 3'UTRs of NKRF mRNA (Figure 4E). In the following study, we verified the binding between HuR and NKRF mRNA by RIP experiments. mRNAs were evaluated by qPCR in HuR pull-down molecules. The results showed that the expression of NKRF mRNA is highly expressed in

HuR pull-down molecules; however, when HuR was knocked down, the expression of NKRF mRNA was significantly decreased (Figure 4F), suggesting that HuR may indeed bind to NKRF mRNA.

NKRF Suppresses NF- κ B Signaling Pathway in HuR-Deficient NP Cells and Suppresses Inflammation and ECM Degradation in TNF- α -Treated NP Cells

NKRF may regulate NF- κ B signaling pathway under physiological conditions; however, whether it may regulate NF- κ B signaling pathway under HuR-deficient conditions is still unknown. Thus, we upregulated NKRF expression in HuR-deficient NP cells by transfecting them with lentivirus-NKRF (LV-NKRF). The WB results showed that the NKRF levels increased by LV-NKRF transfection, implying that we

have successfully established the NKRF upregulation NP cell model (Figure 5A). Next, we assessed whether upregulating NKRF may inhibit NF- κ B signaling pathway induced by HuR overexpression. The WB results showed that the NF- κ B signaling pathway was dramatically activated in the LV-shHuR group, as indicated by the p-p65/p65 ratio, the I κ B α expression in total cell,

and the p65 expression in nuclear, whereas NKRF upregulation can inhibit NF- κ B signaling pathway (Figures 5B,C).

To determine the function of NKRF during the IVDD process, NKRF was overexpressed in TNF- α -treated NP cells, and inflammation and ECM degradation were assessed. The WB results showed that the increased levels of TNF- α , IL-6,

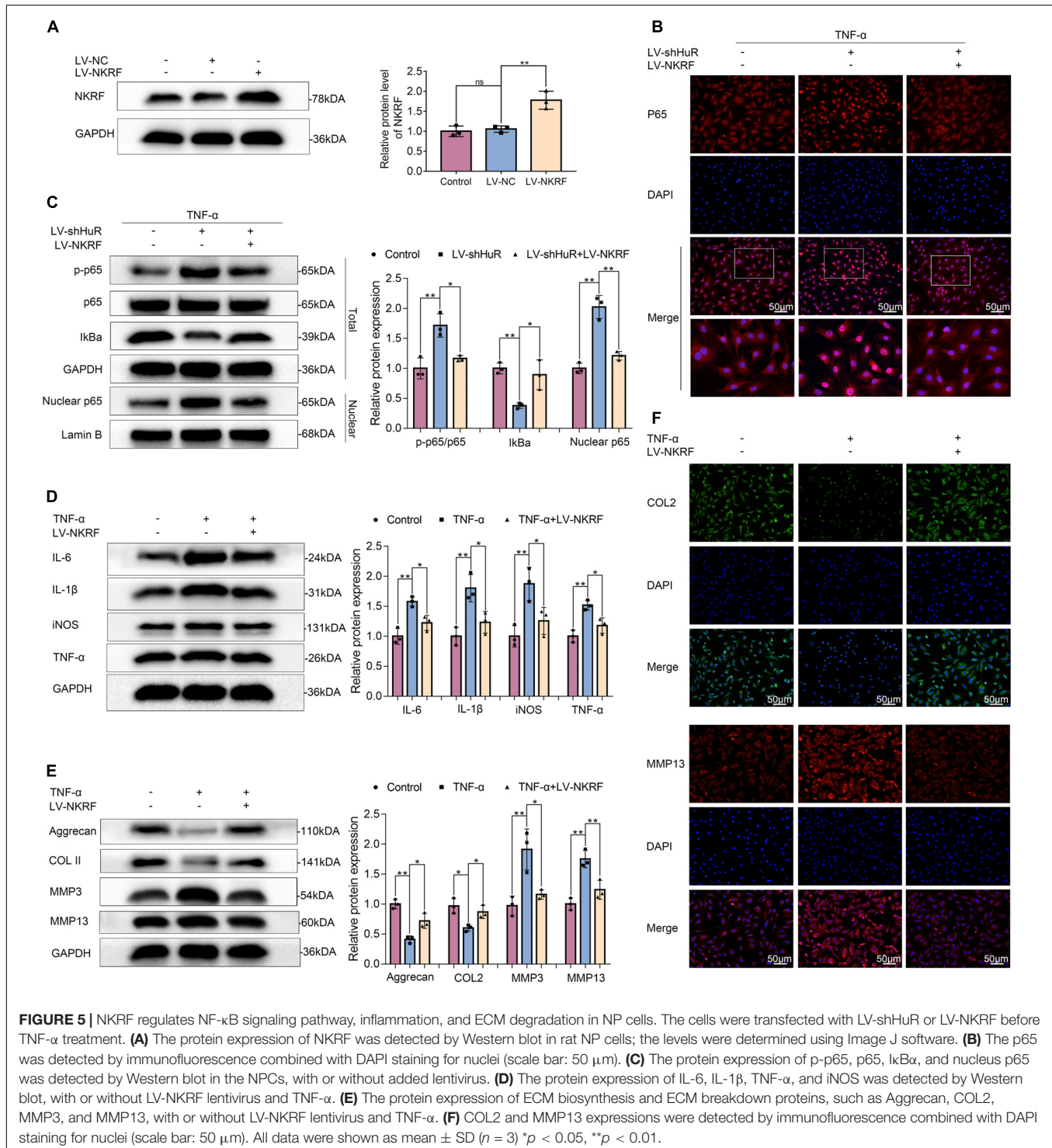


FIGURE 5 | NKRF regulates NF- κ B signaling pathway, inflammation, and ECM degradation in NP cells. The cells were transfected with LV-shHuR or LV-NKRF before TNF- α treatment. **(A)** The protein expression of NKRF was detected by Western blot in rat NP cells; the levels were determined using Image J software. **(B)** The p65 was detected by immunofluorescence combined with DAPI staining for nuclei (scale bar: 50 μ m). **(C)** The protein expression of p-p65, p65, I κ B α , and nucleus p65 was detected by Western blot in the NPCs, with or without added lentivirus. **(D)** The protein expression of IL-6, IL-1 β , TNF- α , and iNOS was detected by Western blot, with or without LV-NKRF lentivirus and TNF- α . **(E)** The protein expression of ECM biosynthesis and ECM breakdown proteins, such as Aggrecan, COL2, MMP3, and MMP13, with or without LV-NKRF lentivirus and TNF- α . **(F)** COL2 and MMP13 expressions were detected by immunofluorescence combined with DAPI staining for nuclei (scale bar: 50 μ m). All data were shown as mean \pm SD ($n = 3$) * $p < 0.05$, ** $p < 0.01$.

IL-1 β , and iNOS induced by TNF- α treatment were attenuated by NKRF overexpression (**Figure 5D**); the degradation of ECM induced by TNF- α was also attenuated by NKRF overexpression (**Figures 5E,F**).

Overexpression of NKRF, but Not HuR, Ameliorates the Process of IVDD *in vivo*

To investigate the protective effects of HuR and NKRF in IVDD *in vivo*, we established puncture-induced rat IVDD model and injected lentivirus into the intervertebral disc to regulate the expression of HuR and NKRF.

The X-ray results showed that disc height was markedly decreased in the IVDD group; however, the disc height was higher in the LV-NKRF group compared to the IVDD group, whereas there was no significant difference between the LV-HuR group and the IVDD group (**Figure 6A**). Meanwhile, we observed that the T2-weighted signal intensity of the intervertebral disc in the LV-NKRF group was brighter than the one in the IVDD group, whereas the MRI T2-weighted signal intensity in the LV-HuR group was as weak as that in the IVDD group (**Figure 6B**). Pfirrmann grade scores were also determined. The scores were lower in the LV-NKRF group, and no significant difference was found in the LV-HuR group compared with the IVDD group (**Figure 6B**).

From H&E staining, we found that the NP tissues were reduced and replaced by the fibrochondrocytes, and the cartilaginous endplate was collapsed in the IVDD group, whereas NKRF upregulation ameliorates these degenerations (**Figure 6C**). According to Safranin O-Fast Green staining, which is sensitive to proteoglycan and hyaluronic acid, we also found that red staining signal in the LV-NKRF group was higher than those in the IVDD group, whereas red staining signal in NP tissues in the LV-HuR group was similar to that in the IVDD group (**Figure 6D**).

Further, the results showed that the inflammation and ECM degradation were prompted in the IVDD group, as indicated by the expression of MMP13 and IL-6 (detected by immunohistochemistry, **Figure 6E**), whereas NKRF overexpression ameliorated the increased inflammation and ECM degradation. In order to investigate the regulatory effect of NKRF on NF- κ B signaling pathway *in vivo*, we detected the expression of p-p65 expressions by immunohistochemical staining. The results showed that the levels of p-p65 expressions were prompted in the IVDD group, whereas NKRF upregulation inhibited the expression of p-p65 expressions (**Figure 6E**). Unfortunately, we also found that HuR overexpression cannot inhibit the expression of MMP13, IL-6, and p-p65, indicating that HuR overexpression may not ameliorate inflammation, ECM degradation, and NF- κ B signaling pathway *in vivo*.

DISCUSSION

Intervertebral disc degeneration has been reported to be a major cause of low back pain. In the current study, we found that HuR may suppress inflammatory response and maintain ECM homeostasis in NP cells *via* inhibiting NF- κ B signaling pathway *in vitro*; a mechanism study revealed that HuR prompts

NKRF mRNA stability *via* binding to the AU-rich element, and upregulation of NKRF inhibits NF- κ B signaling pathway and inflammatory response and maintains ECM homeostasis in HuR-deficient NP cells. *In vivo*, we demonstrated that NKRF overexpression ameliorates the process of IVDD (**Figure 6F**).

Human antigen R is an RNA-binding protein, and HuR stabilizes mRNAs to regulate gene expression, which in turn affects various and contradictory biological processes (Peng et al., 1998; Zhang et al., 2018). Here, we asked whether HuR plays a protective or destructive role in IVDD. Shang et al. (2015) found that HuR promotes the progression of diabetic nephropathy and regulates NADPH oxidase, by regulating the mRNA of NOD2. Zhang and Bowden (2008) found that HuR upregulation promotes non-melanoma skin cancer under UVB by maintaining the stability of COX-2 mRNA. On the other hand, Liu L. et al. (2017) found that HuR upregulation enhances the early recovery of the intestinal epithelium after acute injury via increasing the translation of Cdc42. In this study, we found that the role of HuR in IVDD is complicated; although its overexpression may not alleviate IVDD pathology, its knockdown may exacerbate IVDD process in TNF- α -treated NP cells.

Inflammatory factors play a crucial role in the development of IVDD. However, the regulation of HuR in inflammation is controversial. Mubaid et al. (2019) found that HuR overexpression promotes the inflammation by promoting the translation of STAT3 in muscular dystrophy, and Krishnamurthy et al. (2010) found that HuR knockdown can attenuate the inflammatory response after myocardial infarction in IL-10-deficient mice, by reducing the mRNA expression of TNF- α and TGF- β ; however, Yiakouvakaki et al. (2012) found that high expression of HuR can protect mice from inflammation in pathological enteritis; Ceolotto et al. (2014) found that HuR suppresses chronic inflammation, associated with endothelial injury, by stabilizing the mRNA of SIRT1. In the current study, we found that HuR is negatively related to inflammation in IVDD, because the expression of IL-6, IL-1 β , TNF- α , and iNOS was higher in HuR knockdown and the TNF- α -treated group than in the TNF- α -alone group. At the same time, we found that knockdown of HuR inhibits ECM anabolism (Aggrecan, COL2) and promotes ECM catabolism (MMP3, MMP13).

The NF- κ B signaling pathway plays a significant role in inflammation (Liu T. et al., 2017; Ke et al., 2019). Studies have reported that abnormal activation of the NF- κ B signaling pathway is closely related to the occurrence and development of IVDD (Wuertz et al., 2012), and inhibition of the NF- κ B signaling pathway can delay the progression of IVDD (Liu T. et al., 2017; Taniguchi and Karin, 2018). At the same time, the KEGG pathway analysis of the HuR-regulated gene chip suggests that the NF- κ B signaling pathway plays a role under HuR regulation. However, the regulatory role of HuR on the NF- κ B signaling pathway is controversial. Some studies have shown that HuR activates the NF- κ B signaling pathway. Liu and Yu (2014) found that HuR overexpression in 3T3 cells promotes the expression of I κ B α protein and activates the NF- κ B signaling pathway under heat shock treatment, and Rhee et al. (2010) found that knockdown of HuR with siRNA can inhibit the phosphorylation of NF- κ B in endothelial cells to inhibit the NF- κ B signaling pathway.

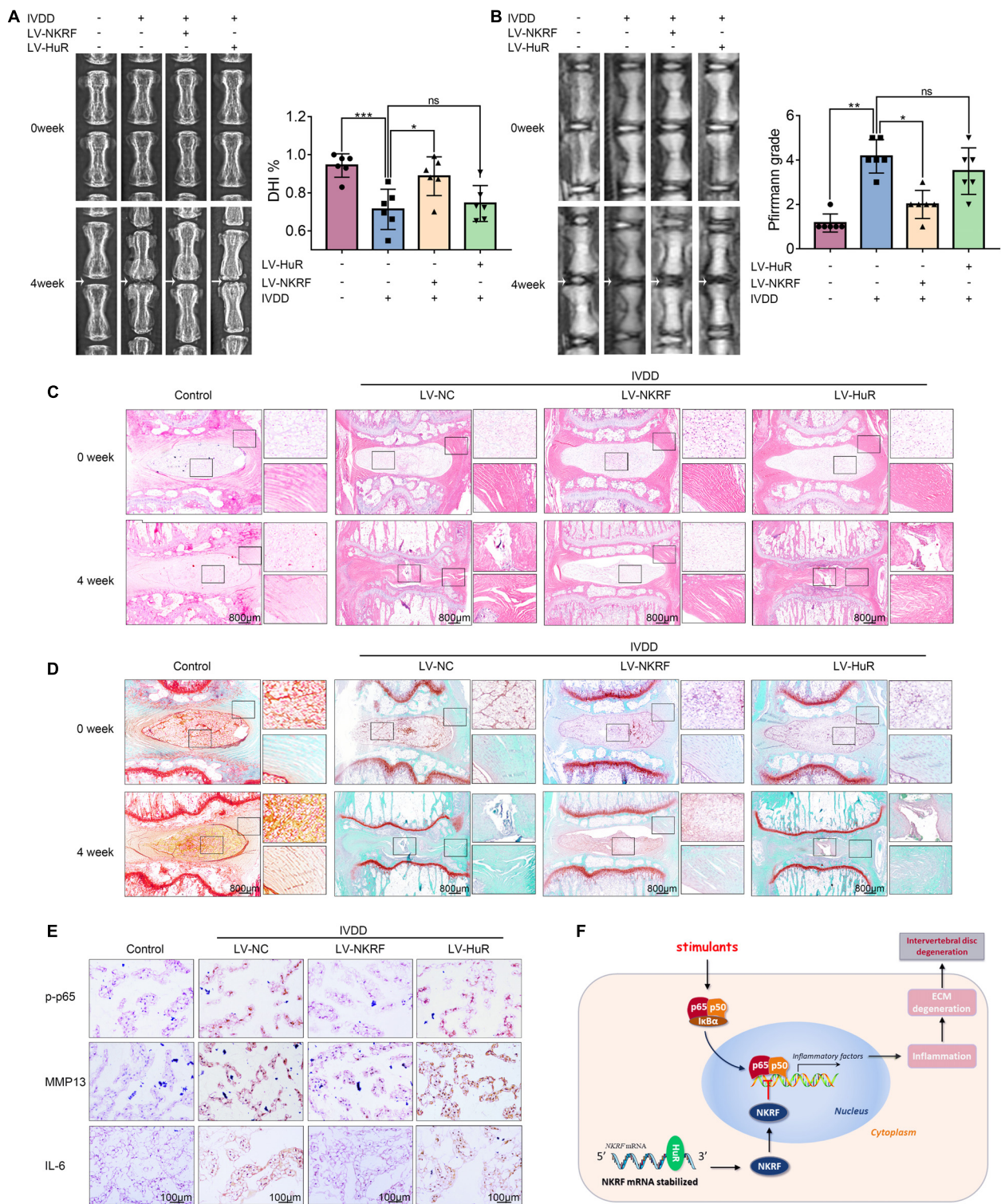


FIGURE 6 | NKRF, but not HuR, overexpression ameliorates the process of IVDD *in vivo*. The rats were injected lentivirus into the intervertebral disc. **(A)** The X-ray of a rat-tail disc at 4 weeks after disc puncture surgery in normal and IVDD rats, with or without lentivirus transfection (white arrows); the disc height index (DHI) of a rat-tail disc. **(B)** T2-weighted MRI of a rat-tail disc at 4 weeks after disc puncture surgery (white arrows); The respective Pfirrmann grade scores of a rat-tail disc. **(C,D)** Representative hematoxylin and eosin staining and Safranin O-Fast Green staining of NP tissues in normal and IVDD rats (bar: 800 μm). **(E)** The immunohistochemical staining of p-p65, MMP13, and IL-6 in intervertebral disc sections (bar: 100 μm). **(F)** Schematic illustration of the research. NKRF inhibits inflammation and ECM degradation in IVDD via NF-κB signaling pathway; mechanism study revealed that HuR prompts NKRF mRNA stability via binding to the AU-rich element. All data were shown as mean ± SD. **p* < 0.05, ***p* < 0.01, ****p* < 0.001.

However, some studies have also shown that HuR may suppress the NF- κ B pathway. Ceolotto et al. (2014) found that in chronic inflammation, HuR inhibits the NF- κ B signaling pathway by stabilizing SIRT1 mRNA, while Hashimoto et al. (2014) found in fibroblasts that the lack of HuR activates the NF- κ B signaling pathway. In our study, the results showed that under TNF- α stimulation conditions, HuR knockdown may promote NF- κ B signaling pathway activation, while NF- κ B pathway inhibitor may alleviate the inflammation induced by HuR knockdown.

Next, we explored the target molecule through which HuR exerts its suppressive effects on NF- κ B signaling pathway. We tested known proteins that inhibit NF- κ B activation and found that the expression of NKRF has the most significant difference. The data in NCBI also showed that the expression of NKRF decreased in both blood and NP tissues of IVDD patients, and the WB results of this study also showed that the expression of NKRF was regulated by the HuR. Furthermore, we found that there are obvious HuR binding sites at the 3'UTRs of NKRF mRNA, and RIP experiments showed that NKRF mRNA can indeed bind to HuR, and the expression of NKRF mRNA is positively correlated to HuR protein expression. In short, NKRF is likely to be the target of HuR.

NKRF, which is abundant in many cells and tissues, is mainly located in the nucleolus and also in the cytoplasm and nucleoplasm (Niedick et al., 2004). NKRF interacts with specific negative regulatory elements to regulate the transcriptional activity of NF- κ B complexes, which inhibits the expression of certain NF- κ B response genes (Feng et al., 2002; Akool El et al., 2003; Gealy et al., 2005). In this research, upregulating the expression of NKRF can inhibit the inflammation and ECM degradation induced by HuR knockdown; meanwhile, we found that the NKRF upregulation can directly inhibit the inflammation induced by TNF- α , reflecting the ability and potential of NKRF to regulate inflammation and ECM degradation directly. However, there are also studies that reported a different role of NKRF in inflammation. Froese et al. (2006) demonstrated that when attacked by pathogens, the survival of NKRF gene-deficient mice showed no difference from wild-type mice; it was proposed that NKRF may be a redundant gene. The contradiction suggests that the role of NKRF in inflammation may be stimulation dependent and cell type dependent.

Based on the function of HuR in NP cells, we tried to upregulate HuR to rescue the TNF- α -induced inflammation and ECM degradation. To our surprise, although the downregulation of HuR caused an increase in inflammation and degradation of ECM in the TNF- α -treated NP cells, upregulation of HuR could not inhibit the inflammation and ECM degradation.

One possible reason is that HuR may have a pleiotropic effect by interacting with different mRNAs (**Supplementary Figure S4**). Gealy et al. (2005) found that HuR promotes the stability of IL-6 mRNA in human fibroblasts when they were infected with human cytomegalovirus; Suzuki et al. (2006) found in rheumatoid arthritis synovium that HuR may promote the stability of TNF- α mRNA; Linker et al. (2005) reported that HuR could promote the expression of iNOS mRNA in human epithelial colon cancer cells. Therefore, although it was shown in our research that HuR may promote NKRF, which, in turn,

inhibits the NF- κ B signaling pathway and helps NP cells to combat inflammation and ECM degradation, the overexpression of HuR may also directly upregulate mRNA that may promote inflammation and ECM degradation. Thus, we propose that HuR may not serve as a therapeutic target for IVDD. Instead, we verified that NKRF overexpression may suppress inflammation and ECM degradation in TNF- α -treated NP cells and ameliorate the process of IVDD *in vivo*. In short, NKRF, but not HuR, may serve as therapeutic target for IVDD (**Figure 6F**).

CONCLUSION

In the current study, we demonstrate that the expression of HuR is decreased in NP cells during the process of IVDD, and downregulation of HuR promotes IVDD *via* activating NF- κ B signaling pathway; a mechanism study reveals that HuR prompts NKRF mRNA stability *via* binding to the AU-rich element. HuR upregulation could not ameliorate the process of IVDD, while upregulation of NKRF inhibits NF- κ B signaling pathway and inflammatory response and maintains ECM homeostasis in HuR-deficient NP cells. Thus, overexpression of NKRF, not HuR, ameliorates the process of IVDD (**Figure 6F**).

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.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Second Affiliated Hospital and Yuying Children's Hospital of Wenzhou Medical University Ethics Committee. The patients/participants provided their written informed consent to participate in this study. The animal study was reviewed and approved by the Animal Care and Use Committee of Wenzhou Medical University.

AUTHOR CONTRIBUTIONS

XW, XZ, and YFZ designed the experiments. ZS, YS, and LC performed the experiments. ZT and SL wrote the manuscript. AW, YW, and NT analyzed the data and prepared all the figures. LS, ZP, and WG constructed the animal models. All authors reviewed and agreed on the manuscript.

FUNDING

This work is supported by the Zhejiang Provincial Natural Science Foundation of China (LGF21H060011, LY17H250002, LQ19H060004, LGF20H060013, and LY18H060012), the Wenzhou Science and Technology Bureau Foundation

(ZY2019014), the National Natural Science Foundation of China (81871806, 81972094, and 81902243), the Zhejiang Provincial Project for Medical and Health Science and Technology (2017KY463), and Lin He's New Medicine and Clinical Translation Academician Workstation Research Fund.

REFERENCES

- Akool El, S., Kleinert, H., Hamada, F. M., Abdelwahab, M. H., Forstermann, U., Pfeilschifter, J., et al. (2003). Nitric oxide increases the decay of matrix metalloproteinase 9 mRNA by inhibiting the expression of mRNA-stabilizing factor HuR. *Mol. Cell Biol.* 23, 4901–4916. doi: 10.1128/mcb.23.14.4901-4916.2003
- Alvarez-Garcia, O., Matsuzaki, T., Olmer, M., Masuda, K., and Lotz, M. K. (2017). Age-related reduction in the expression of FOXO transcription factors and correlations with intervertebral disc degeneration. *J. Orthop. Res.* 35, 2682–2691. doi: 10.1002/jor.23583
- Ceolotto, G., De Kreutzenberg, S. V., Cattelan, A., Fabricio, A. S., Squarcina, E., Gion, M., et al. (2014). Sirtuin 1 stabilization by HuR represses TNF- α and glucose-induced E-selectin release and endothelial cell adhesiveness in vitro: relevance to human metabolic syndrome. *Clin. Sci.* 127, 449–461. doi: 10.1042/cs20130439
- Chapman, S. J., Khor, C. C., Vannberg, F. O., Frodsham, A., Walley, A., Maskell, N. A., et al. (2007). IkappaB genetic polymorphisms and invasive pneumococcal disease. *Am. J. Respir. Crit. Care Med.* 176, 181–187. doi: 10.1164/rccm.200702-1690C
- Chen, F., Jiang, G., Liu, H., Li, Z., Pei, Y., Wang, H., et al. (2020). Melatonin alleviates intervertebral disc degeneration by disrupting the IL-1 β /NF- κ B-NLRP3 inflammasome positive feedback loop. *Bone Res.* 8:10. doi: 10.1038/s41413-020-0087-2
- Chen, Y., Lin, J., Chen, J., Huang, C., Zhang, Z., Wang, J., et al. (2020). Mfn2 is involved in intervertebral disc degeneration through autophagy modulation. *Osteoarthritis. Cartil.* 28, 363–374. doi: 10.1016/j.joca.2019.12.009
- Feng, X., Guo, Z., Nourbakhsh, M., Hauser, H., Ganster, R., Shao, L., et al. (2002). Identification of a negative response element in the human inducible nitric-oxide synthase (hNOS) promoter: The role of NF-kappa B-repressing factor (NRF) in basal repression of the hNOS gene. *Proc. Natl. Acad. Sci. U S A* 99, 14212–14217. doi: 10.1073/pnas.212306199
- Freemont, A. J. (2009). The cellular pathobiology of the degenerate intervertebral disc and discogenic back pain. *Rheumatology* 48, 5–10. doi: 10.1093/rheumatology/ken396
- Froese, N., Schwarzer, M., Niedick, I., Frischmann, U., Koster, M., Kroger, A., et al. (2006). Innate immune responses in NF-kappaB-repressing factor-deficient mice. *Mol. Cell Biol.* 26, 293–302. doi: 10.1128/mcb.26.1.293-302.2006
- Gealy, C., Denson, M., Humphreys, C., Mcsharry, B., Wilkinson, G., and Caswell, R. (2005). Posttranscriptional suppression of interleukin-6 production by human cytomegalovirus. *J. Virol.* 79, 472–485. doi: 10.1128/jvi.79.1.472-485.2005
- Glisovic, T., Bachorik, J. L., Yong, J., and Dreyfuss, G. (2008). RNA-binding proteins and post-transcriptional gene regulation. *FEBS Lett.* 582, 1977–1986. doi: 10.1016/j.febslet.2008.03.004
- Gorth, D. J., Ottone, O. K., Shapiro, I. M., and Risbud, M. V. (2020). Differential Effect of Long-Term Systemic Exposure of TNF α on Health of the Annulus Fibrosus and Nucleus Pulposus of the Intervertebral Disc. *J. Bone Miner. Res.* 35, 725–737. doi: 10.1002/jbmr.3931
- Gruber, H. E., Norton, H. J., Ingram, J. A., and Hanley, E. N., Jr. (2005). The SOX9 transcription factor in the human disc: decreased immunolocalization with age and disc degeneration. *Spine* 30, 625–630. doi: 10.1097/01.brs.0000155420.01444.c6
- Hashimoto, M., Tsugawa, T., Kawagishi, H., Asai, A., and Sugimoto, M. (2014). Loss of HuR leads to senescence-like cytokine induction in rodent fibroblasts by activating NF-kappaB. *Biochim. Biophys. Acta* 5, 3079–3087. doi: 10.1016/j.bbagen.2014.07.005
- Hoy, D., Brooks, P., Blyth, F., and Buchbinder, R. (2010). The Epidemiology of low back pain. *Best Pract. Res. Clin. Rheumatol.* 24, 769–781. doi: 10.1016/j.berh.2010.10.002
- Hua, W. B., Wu, X. H., Zhang, Y. K., Song, Y., Tu, J., Kang, L., et al. (2017). Dysregulated miR-127-5p contributes to type II collagen degradation by targeting matrix metalloproteinase-13 in human intervertebral disc degeneration. *Biochimie* 139, 74–80. doi: 10.1016/j.biochi.2017.05.018
- Huang, L., Verstrepen, L., Heyninck, K., Wullaert, A., Revets, H., De Baetselier, P., et al. (2008). ABINs inhibit EGF receptor-mediated NF-kappaB activation and growth of EGF receptor-overexpressing tumour cells. *Oncogene* 27, 6131–6140. doi: 10.1038/onc.2008.208
- Jha, P., and Das, H. (2017). KLF2 in Regulation of NF- κ B-Mediated Immune Cell Function and Inflammation. *Int. J. Mol. Sci.* 18:2383. doi: 10.3390/ijms18112383
- Katsanou, V., Papadaki, O., Milatos, S., Blackshear, P. J., Anderson, P., Kollias, G., et al. (2005). HuR as a negative posttranscriptional modulator in inflammation. *Mol. Cell.* 14, 777–789. doi: 10.1016/j.molcel.2005.08.007
- Kawagishi, H., Hashimoto, M., Nakamura, H., Tsugawa, T., Watanabe, A., Kontoyiannis, D. L., et al. (2013). HuR maintains a replicative life span by repressing the ARF tumor suppressor. *Mol. Cell Biol.* 33, 1886–1900. doi: 10.1128/mcb.01277-12
- Ke, K., Chen, T. H., Arra, M., Mbalaviele, G., Swarnkar, G., and Abu-Amer, Y. (2019). Attenuation of NF- κ B in Intestinal Epithelial Cells Is Sufficient to Mitigate the Bone Loss Comorbidity of Experimental Mouse Colitis. *J. Bone Miner. Res.* 34, 1880–1893. doi: 10.1002/jbmr.3759
- Krishnamurthy, P., Lambers, E., Verma, S., Thorne, T., Qin, G., Losordo, D. W., et al. (2010). Myocardial knockdown of mRNA-stabilizing protein HuR attenuates post-MI inflammatory response and left ventricular dysfunction in IL-10-null mice. *FASEB J.* 24, 2484–2494. doi: 10.1096/fj.09-149815
- Linker, K., Pautz, A., Fechir, M., Hubrich, T., Greeve, J., and Kleinert, H. (2005). Involvement of KSRP in the post-transcriptional regulation of human iNOS expression-complex interplay of KSRP with TTP and HuR. *Nucleic Acids Res.* 33, 4813–4827. doi: 10.1093/nar/gki797
- Liu, L., Zhuang, R., Xiao, L., Chung, H. K., Luo, J., Turner, D. J., et al. (2017). HuR Enhances Early Restitution of the Intestinal Epithelium by Increasing Cdc42 Translation. *Mol. Cell Biol.* 37, 574–16. doi: 10.1128/mcb.00574-16
- Liu, T., Zhang, L., Joo, D., and Sun, S. C. (2017). NF- κ B signaling in inflammation. *Sign. Transduct. Target. Ther.* 2:17023. doi: 10.1038/sigtrans.2017.23
- Liu, Y., and Yu, W. (2014). Heat shock-mediated regulation of IkappaB-alpha at the post-transcriptional level by HuR. *Mol. Med. Rep.* 9, 553–559. doi: 10.3892/mmr.2013.1820
- Lopez De Silanes, I., Zhan, M., Lal, A., Yang, X., and Gorospe, M. (2004). Identification of a target RNA motif for RNA-binding protein HuR. *Proc. Natl. Acad. Sci. U S A* 101, 2987–2992. doi: 10.1073/pnas.0306453101
- Martin, B. I., Deyo, R. A., Mirza, S. K., Turner, J. A., Comstock, B. A., Hollingworth, W., et al. (2008). Expenditures and health status among adults with back and neck problems. *JAMA* 299, 656–664. doi: 10.1001/jama.299.6.656
- Mubaid, S., Ma, J. F., Omer, A., Ashour, K., Lian, X. J., Sanchez, B. J., et al. (2019). HuR counteracts miR-330 to promote STAT3 translation during inflammation-induced muscle wasting. *Proc. Natl. Acad. Sci. U S A* 116, 17261–17270. doi: 10.1073/pnas.1905172116
- Mukherjee, N., Corcoran, D. L., Nusbaum, J. D., Reid, D. W., Georgiev, S., Hafner, M., et al. (2011). Integrative regulatory mapping indicates that the RNA-binding protein HuR couples pre-mRNA processing and mRNA stability. *Mol. Cell* 43, 327–339. doi: 10.1016/j.molcel.2011.06.007
- Niedick, I., Froese, N., Oumard, A., Mueller, P. P., Nourbakhsh, M., Hauser, H., et al. (2004). Nucleolar localization and mobility analysis of the NF-kappaB repressing factor NRF. *J. Cell. Sci.* 117, 3447–3458. doi: 10.1242/jcs.01129
- Nourbakhsh, M., and Hauser, H. (1999). Constitutive silencing of IFN-beta promoter is mediated by NRF (NF-kappaB-repressing factor), a nuclear inhibitor of NF-kappaB. *EMBO J.* 18, 6415–6425. doi: 10.1093/emboj/18.22.6415

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fcell.2020.611234/full#supplementary-material>

- Pan, H., Strickland, A., Madhu, V., Johnson, Z. I., Chand, S. N., Brody, J. R., et al. (2019). RNA binding protein HuR regulates extracellular matrix gene expression and pH homeostasis independent of controlling HIF-1 α signaling in nucleus pulposus cells. *Matrix Biol.* 77, 23–40. doi: 10.1016/j.matbio.2018.08.003
- Peng, S. S., Chen, C. Y., Xu, N., and Shyu, A. B. (1998). RNA stabilization by the AU-rich element binding protein, HuR, an ELAV protein. *EMBO J.* 17, 3461–3470. doi: 10.1093/emboj/17.12.3461
- Pfirrmann, C. W., Metzendorf, A., Zanetti, M., Hodler, J., and Boos, N. (2001). Magnetic resonance classification of lumbar intervertebral disc degeneration. *Spine* 26, 1873–1878. doi: 10.1097/00007632-200109010-00011
- Rhee, W. J., Ni, C. W., Zheng, Z., Chang, K., Jo, H., and Bao, G. (2010). HuR regulates the expression of stress-sensitive genes and mediates inflammatory response in human umbilical vein endothelial cells. *Proc. Natl. Acad. Sci. U S A* 107, 6858–6863. doi: 10.1073/pnas.1000444107
- Rinkenbaugh, A. L., and Baldwin, A. S. (2011). Monoallelic deletion of NFKBIA in glioblastoma: when less is more. *Cancer Cell* 19, 163–165. doi: 10.1016/j.ccr.2011.01.045
- Sakai, D., and Grad, S. (2015). Advancing the cellular and molecular therapy for intervertebral disc disease. *Adv. Drug Deliv. Rev.* 84, 159–171. doi: 10.1016/j.addr.2014.06.009
- Schultz, C. W., Preet, R., Dhir, T., Dixon, D. A., and Brody, J. R. (2020). Understanding and targeting the disease-related RNA binding protein human antigen R (HuR). *Wiley Interdisc. Rev. RNA* 11:e1581. doi: 10.1002/wrna.1581
- Shang, J., Wan, Q., Wang, X., Duan, Y., Wang, Z., Wei, X., et al. (2015). Identification of NOD2 as a novel target of RNA-binding protein HuR: evidence from NADPH oxidase-mediated HuR signaling in diabetic nephropathy. *Free Radical. Biol. Med.* 79, 217–227. doi: 10.1016/j.freeradbiomed.2014.12.013
- Shembade, N., Pujari, R., Harhaj, N. S., Abbott, D. W., and Harhaj, E. W. (2011). The kinase IKK α inhibits activation of the transcription factor NF- κ B by phosphorylating the regulatory molecule TAX1BP1. *Nat. Immunol.* 12, 834–843. doi: 10.1038/ni.2066
- Suzuki, E., Tsutsumi, A., Sugihara, M., Mamura, M., Goto, D., Matsumoto, I., et al. (2006). Expression of TNF- α , tristetraprolin, T-cell intracellular antigen-1 and Hu antigen R genes in synovium of patients with rheumatoid arthritis. *Int. J. Mol. Med.* 18, 273–278.
- Taniguchi, K., and Karin, M. (2018). NF- κ B, inflammation, immunity and cancer: coming of age. *Nat. Rev. Immunol.* 18, 309–324. doi: 10.1038/nri.2017.142
- Trompouki, E., Hatzivassiliou, E., Tsichritzis, T., Farmer, H., Ashworth, A., and Mosialos, G. (2003). CYLD is a deubiquitinating enzyme that negatively regulates NF- κ B activation by TNFR family members. *Nature* 424, 793–796. doi: 10.1038/nature01803
- Uchida, Y., Chiba, T., Kurimoto, R., and Asahara, H. (2019). Post-transcriptional regulation of inflammation by RNA-binding proteins via cis-elements of mRNAs. *J. Biochem.* 166, 375–382. doi: 10.1093/jb/mvz067
- Wang, J., Hu, J., Chen, X., Huang, C., Lin, J., Shao, Z., et al. (2019). BRD4 inhibition regulates MAPK, NF- κ B signals, and autophagy to suppress MMP-13 expression in diabetic intervertebral disc degeneration. *FASEB J.* 33, 11555–11566. doi: 10.1096/fj.201900703R
- Wert, I. E., O’rourke, K. M., Zhou, H., Eby, M., Aravind, L., Seshagiri, S., et al. (2004). De-ubiquitination and ubiquitin ligase domains of A20 downregulate NF- κ B signalling. *Nature* 430, 694–699. doi: 10.1038/nature02794
- Wuertz, K., Vo, N., Kletsas, D., and Boos, N. (2012). Inflammatory and catabolic signalling in intervertebral discs: the roles of NF- κ B and MAP kinases. *Eur. Cell Mater.* 23, 103–119. doi: 10.22203/ecm.v023a08
- Yiakouvakis, A., Dimitriou, M., Karakasiliotis, I., Eftychi, C., Theocharis, S., and Kontoyiannis, D. L. (2012). Myeloid cell expression of the RNA-binding protein HuR protects mice from pathologic inflammation and colorectal carcinogenesis. *J. Clin. Invest.* 122, 48–61. doi: 10.1172/jci45021
- Zhang, J., and Bowden, G. T. (2008). UVB irradiation regulates Cox-2 mRNA stability through AMPK and HuR in human keratinocytes. *Mol. Carcinog.* 47, 974–983. doi: 10.1002/mc.20450
- Zhang, Z., Yao, Z., Wang, L., Ding, H., Shao, J., Chen, A., et al. (2018). Activation of ferritinophagy is required for the RNA-binding protein ELAVL1/HuR to regulate ferroptosis in hepatic stellate cells. *Autophagy* 14, 2083–2103. doi: 10.1080/15548627.2018.1503146
- Zhao, B. S., Roundtree, I. A., and He, C. (2017). Post-transcriptional gene regulation by mRNA modifications. *Nat. Rev. Mol. Cell Biol.* 18, 31–42. doi: 10.1038/nrm.2016.132
- Zheng, G., Pan, Z., Zhan, Y., Tang, Q., Zheng, F., Zhou, Y., et al. (2019). TFEB protects nucleus pulposus cells against apoptosis and senescence via restoring autophagic flux. *Osteoarthr. Cartil.* 27, 347–357. doi: 10.1016/j.joca.2018.10.011
- Zheng, L., Cao, Y., Ni, S., Qi, H., Ling, Z., Xu, X., et al. (2018). Ciliary parathyroid hormone signaling activates transforming growth factor- β to maintain intervertebral disc homeostasis during aging. *Bone Res.* 6:21. doi: 10.1038/s41413-018-0022-y

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 Shao, Tu, Shi, Li, Wu, Wu, Tian, Sun, Pan, Chen, Gao, Zhou, Wang and Zhang. 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.



Expression and Prognostic Characteristics of m⁶A RNA Methylation Regulators in Breast Cancer

Bo Zhang[†], Yanlin Gu[†] and Guoqin Jiang^{*}

General Surgery Department, The Second Affiliated Hospital of Soochow University, Suzhou, China

OPEN ACCESS

Edited by:

Jia Meng,
Xi'an Jiaotong-Liverpool University,
China

Reviewed by:

Jinhua Wang,
Chinese Academy of Medical
Sciences and Peking Union Medical
College, China
Justin Jong-Leong Wong,
Royal Prince Alfred Hospital, Australia

*Correspondence:

Guoqin Jiang
jiang_guoqin@163.com

[†] These authors have contributed
equally to this work

Specialty section:

This article was submitted to
Epigenomics and Epigenetics,
a section of the journal
Frontiers in Genetics

Received: 09 September 2020

Accepted: 19 November 2020

Published: 10 December 2020

Citation:

Zhang B, Gu YL and Jiang GQ
(2020) Expression and Prognostic
Characteristics of m⁶A RNA
Methylation Regulators in Breast
Cancer. *Front. Genet.* 11:604597.
doi: 10.3389/fgene.2020.604597

Purpose: N6-methyladenosine (m⁶A) is the most prevalent modification in mRNA methylation which has a wide effect on biological functions. This study aims to figure out the efficacy of m⁶A RNA methylation regulator-based biomarkers with prognostic significance in breast cancer.

Patients and Methods: The 23 RNA methylation regulators were firstly analyzed through ONCOMINE, then relative RNA-seq transcriptome and clinical data of 1,096 breast cancer samples and 112 normal tissue samples were acquired from The Cancer Gene Atlas (TCGA) database. The expressive distinction was also showed by the Gene Expression Omnibus (GEO) database. The gene expression data of m⁶A RNA regulators in human tissues were acquired from the Genotype-Tissue Expression (GTEx) database. The R v3.5.1 and other online tools such as STRING, bc-GeneExminer v4.5, Kaplan-Meier Plotter were applied for bioinformatics analysis.

Results: Results from ONCOMINE, TCGA, and GEO databases showed distinctive expression and clinical correlations of m⁶A RNA methylation regulators in breast cancer patients. The high expression of YTHDF3, ZC3H13, LRPPRC, and METTL16 indicated poor survival rate in patients with breast cancer, while high expression of RBM15B pointed to a better survival rate. Both univariate and multivariate Cox regression analyses revealed that age and risk scores were related to overall survival (OS). Univariate analysis also delineated that stage, tumor (T) status, lymph node (N) status, and metastasis (M) status were associated with OS. From another perspective, Kaplan-Meier Plotter platform showed that the relatively high expression of YTHDF3 and LRPPRC and the relatively low expression of RBM15B, ZC3H13, and METTL16 in breast cancer patients had worse Relapse-Free Survival (RFS). Breast Cancer Gene-Expression Miner v4.5 showed that LRPPRC level was negatively associated with ER and PR expression, while METTL16, RBM15B, ZC3H13 level was positively linked with ER and PR expression. In HER-2 (+) breast cancer patients, the expression of LRPPRC, METTL16, RBM15B, and ZC3H13 were all lower than the HER-2 (−) group.

Conclusion: The significant difference in expression levels and prognostic value of m⁶A RNA methylation regulators were analyzed and validated in this study. This signature revealed the potential therapeutic value of m⁶A RNA methylation regulators in breast cancer.

Keywords: N6-methyladenosine (m⁶A) RNA methylation, breast carcinoma, post-transcriptional modification, bioinformatics analysis, prognosis

INTRODUCTION

According to the data GLOBALCAN 2018, breast cancer is the leading cause of cancer-related deaths and the most frequently diagnosed cancer in females with different phenotypes due to genetic and epigenetic diversity (Bray et al., 2018). Benefiting from the early stage diagnosis and treatment, the 5-year relative survival rate of breast cancer patients primarily diagnosed at stage I reaches to nearly 100%, whereas the survival rate of patients first diagnosed at stage IV is only 26% (Miller et al., 2019). Accurate diagnosis, detection and treatment are always effective approaches to improve the prognosis of breast cancer patients. In breast cancer research, epitranscriptome has been focused for its important role in biological functions. The post-transcriptional modifications are widely located in messenger RNA (mRNA) and non-coding RNA (Shi et al., 2019). Until now, over 170 kinds of RNA modifications have been identified (Boccaletto et al., 2018). The RNA methylated modifications occurring on adenosine present in different forms, such as N6-methyladenosine (m⁶A), N6-2'-O-dimethyladenosine (m⁶Am), N1-methyladenosine (m¹A). In the 1970s, N6-methyladenosine (m⁶A) was first identified as the most prevalent modification in mRNA. It has been proved to be a reversible and widespread internal adenosine modification particularly enriched in 3' UTRs codon of mammalian mRNA (Meyer et al., 2012). m⁶A also has a cooperative effect on RNA functions, such as stability, metabolism, transcriptional regulation and intracellular signaling (Roundtree et al., 2017). The biological importance of m⁶A can be embodied in embryonic self-renewal capability (Wang et al., 2014), hematopoietic stem/progenitor cells (HSPCs) emergence (Zhang et al., 2017), circadian control (Fustin et al., 2018), heat shock response (Zhou et al., 2015), neuronal functions (Lence et al., 2016), and tumorigenesis (Lan et al., 2019).

RNA modification is mediated by a series of interplays including “writers” (methyltransferases), “readers” (binding proteins), and “erasers” (demethylases) which regulate a complicated biological regulation process. The “writers” include methyltransferase-like 3 (METTL3), methyltransferase-like 14 (METTL14), methyltransferase-like 16 (METTL16), William tumor 1 associated protein (WTAP), vir like m⁶A methyltransferase associated (VIRMA, also named KIAA1429), zinc finger CCCH domain-containing protein 13 (ZC3H13), RNA binding motif protein 15 (RBM15), and RNA binding motif protein 15B (RBM15B). The “readers” are composed of the YTH family: YTH domain containing 1 (YTHDC1), YTH domain containing 2 (YTHDC2), YTH domain-containing family protein 1 (YTHDF1), YTH domain-containing family

protein 2 (YTHDF2), YTH domain-containing family protein 3 (YTHDF3), heterogeneous nuclear ribonucleoprotein C (HNRNPC), FMRP translational regulator 1 (FMR1), leucine-rich pentatricopeptide-repeat containing (LRPPRC), heterogeneous nuclear ribonucleoprotein A2/B1 (HNRNPA2B1), insulin like growth factor binding proteins (IGFBPs): IGFBP1, IGFBP2, IGFBP3 and heterogeneous nuclear ribonucleoprotein G (HNRNPG, also named RBMX). The “erasers” are fat mass and obesity-associated protein (FTO) and α -ketoglutarate-dependent dioxygenase alkB homolog 5 (ALKBH5) (Warda et al., 2017; Yang et al., 2018). Each m⁶A RNA methylation regulator serves a critical role in the RNA methylation process. Although in the past few years, researches have been done to decipher the interactions coupled with m⁶A RNA methylation regulators, the nature of RNA modifications and their biological functions in breast cancer remain unclear.

To further clarify the expression of various RNA regulators and their impacts on the prognosis of breast cancer, we obtained the expression of 23 m⁶A RNA regulators from ONCOMINE. We downloaded breast cancer datasets from TCGA, which included 1096 breast cancer samples and 112 normal tissue samples to analyze the distinct expressions of 23 RNA methylation regulators and their clinical characteristics in breast cancer. The expressive results were also confirmed by the GEO database. Through the COX regression analysis and the least absolute shrinkage and selection operator (LASSO) regression, we found five m⁶A methylation regulators which were significant to clinical prognosis. Kaplan-Meier Plotter and bc-GenExMiner v4.5 were used to further explore the clinicopathological and prognostic value of the five m⁶A methylation regulators. The Gene Ontology (GO) analysis of 23 regulators and the Gene Set Enrichment Analysis (GSEA) may help to shed light on their biological research value.

MATERIALS AND METHODS

Datasets and Study Cohort ONCOMINE Database

ONCOMINE database¹ is a systematic and comprehensive aggregation of microarray datasets. The 23 m⁶A RNA methylation regulators were initially analyzed through ONCOMINE by evaluating their expression in various cancer types compared with the normal. We searched each regulator for the Cancer vs. Normal Analysis and typed with the threshold of $p < 0.05$ and with the gene ranking at the top 10%.

¹<https://www.oncomine.org/>

TCGA Database

The RNA-seq transcriptome data of 1,096 breast cancer samples and 112 normal tissue samples were downloaded from the TCGA database². Clinical data of breast cancer patients was also downloaded from TCGA. All data were normalized by the expectation-maximization method and converted into Sample IDs by Perl. Patients without survival information were excluded.

GEO Database

The differential expression of m⁶A RNA methylation regulators between breast cancer tissues and adjacent normal tissues was verified by GSE70905. The data was downloaded from the GEO database³. The GSE70905 included 47 breast adenocarcinoma and 47 paired adjacent normal breast tissues.

Kaplan-Meier Plotter

The Kaplan-Meier curves were displayed via Kaplan-Meier Plotter⁴ with the log-rank test. It was performed to provide survival information of m⁶A RNA methylation regulators in breast cancer patients.

Breast Cancer Gene-Expression Miner v4.5

A public statistical mining tool, the Breast Cancer Gene-Expression Miner v4.5⁵, was used to evaluate the correlations between mRNA expression and clinicopathological characteristics, such as ER, PR, and HER-2 status, as well as the Nottingham Prognostic Index (NPI) and the Scarff-Bloom-Richardson (SBR) grading.

Other Online Databases

The Protein-Protein Interaction (PPI) network was used to exhibit the comprehensive functional annotation of related proteins (version11.0)⁶. Besides, the gene expression of m⁶A RNA methylation regulators in human tissues was acquired from the GTEx database⁷. The data was acquired in a way where strict ethical guidelines were followed.

Selection of m⁶A Methylation Regulators

A total of 23 regulators were selected according to the previous study, including ALKBH5, FTO, HNRNPC, HNRNPA2B1, KIAA1429, METTL14, METTL3, METTL16, RBM15, RBMX, RBM15B, WTAP, YTHDC1, YTHDC2, YTHDF1, YTHDF2, YTHDF3, ZC3H13, IGF2BP1, IGF2BP2, IGF2BP3, FMR1, and LRPPRC. Then their expressive correlations and clinicopathological characteristics were analyzed.

Bioinformatic Analysis

The expressive distinction and correlation of 23 regulators were analyzed through “Limma” package by using R v3.5.1⁸ with the cut-off criteria of $p < 0.05$. A heatmap diagram was drawn to

compare the expression of m⁶A RNA methylation regulators between tumors and normal tissues. Vioplot diagrams were constructed to visualize the expression of 23 regulators of breast cancer and normal tissues. Besides, GO analysis of 17 regulators was performed through “GO PLOT” and “Digest” package. Univariate Cox regression analysis and LASSO Cox regression model were used to correlate prognostic m⁶A regulators, and the corresponding results showed that five regulators significantly correlated with the prognosis of breast cancer according to the statistical analysis of log-rank p -value ($p < 0.05$) and the hazard ratio (HR) with 95% confidence intervals. If the $HR > 1$, the gene expression shows a positive correlation with OS, while $HR < 1$ indicates a negative correlation with OS. Additionally, the receiver operating characteristic (ROC) curves were used to evaluate sensitivity and specificity. Next, all the samples selected by LASSO analysis were divided into two groups judged by the risk score (RS). Patients with RS above the mean were assigned to the high-risk group while the rest patients with RS below the mean were assigned to the low-risk groups. The GSEA was performed to study the functions m⁶A RNA methylation regulators by using “kegg.v7.1 symbol.gmt” package.

Statistical Analysis

The expression of m⁶A RNA methylation regulators between breast cancer and normal tissues were compared through the Wilcoxon test. The high-risk group and the low-risk group were classified according to the median risk score. The chi-square test was used to compare the relationship between the clinicopathological variables and risk score. Cox univariate and multivariate analyses were performed to compare the relationship between clinicopathological variables and risk score. The t -test was used to compare the difference between the two groups classified by risk scores. The value $p < 0.05$ was considered to be statistically significant.

RESULTS

Expression of m⁶A RNA Methylation Regulators in Breast Cancer

The ONCOMINE analysis revealed the gene expression of the 23 RNA methylation regulators in different types of cancer compared with normal tissues (Figure 1). From the ONCOMINE database, we had a brief view of the expression of the genes. METTL3, METTL16, ZC3H13, YTHDC1 and FTO were expressed at a low level in breast cancer, while KIAA1429, RBM15, and YTHDF1 were expressed at a high level. Then the m⁶A RNA methylation regulators were analyzed to compare the expression level through the data from the TCGA database. The expression levels in each regulator were compared by the average level of all tissue samples. YTHDF1 ($p < 0.001$), HNRNPA2B1 ($p < 0.001$), HNRNPC ($p < 0.001$), LRPPRC ($p < 0.001$), KIAA1429 ($p < 0.001$), RBM15 ($p < 0.001$), FMR1 ($p < 0.01$), IGF2BP1 ($p < 0.01$), YTHDF2 ($p < 0.05$), and IGF2BP3 ($p < 0.05$) were over-expressed at the mean level in breast cancer tissues compared with normal tissues,

²<http://www.cancergenome.nih.gov/>

³<https://www.ncbi.nlm.nih.gov/geo/>

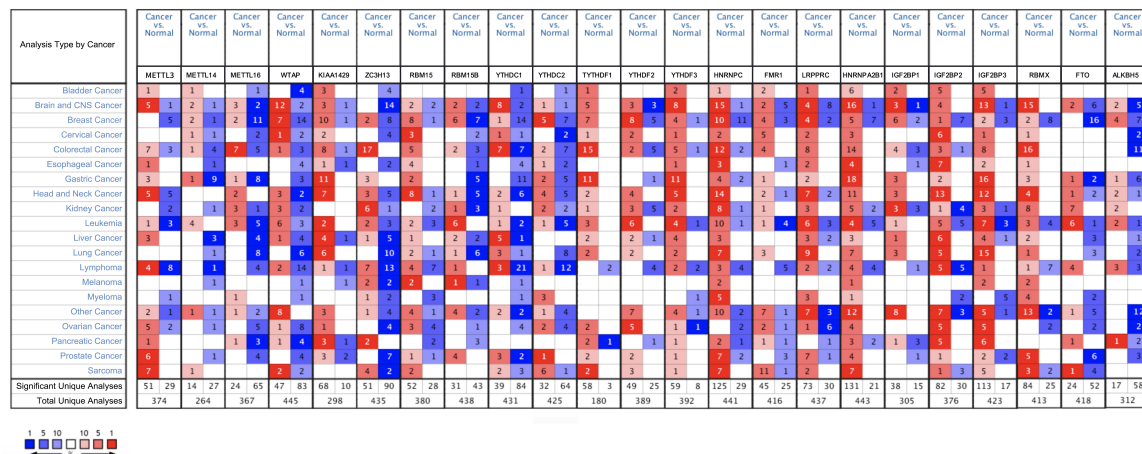
⁴<http://kmplot.com/analysis/>

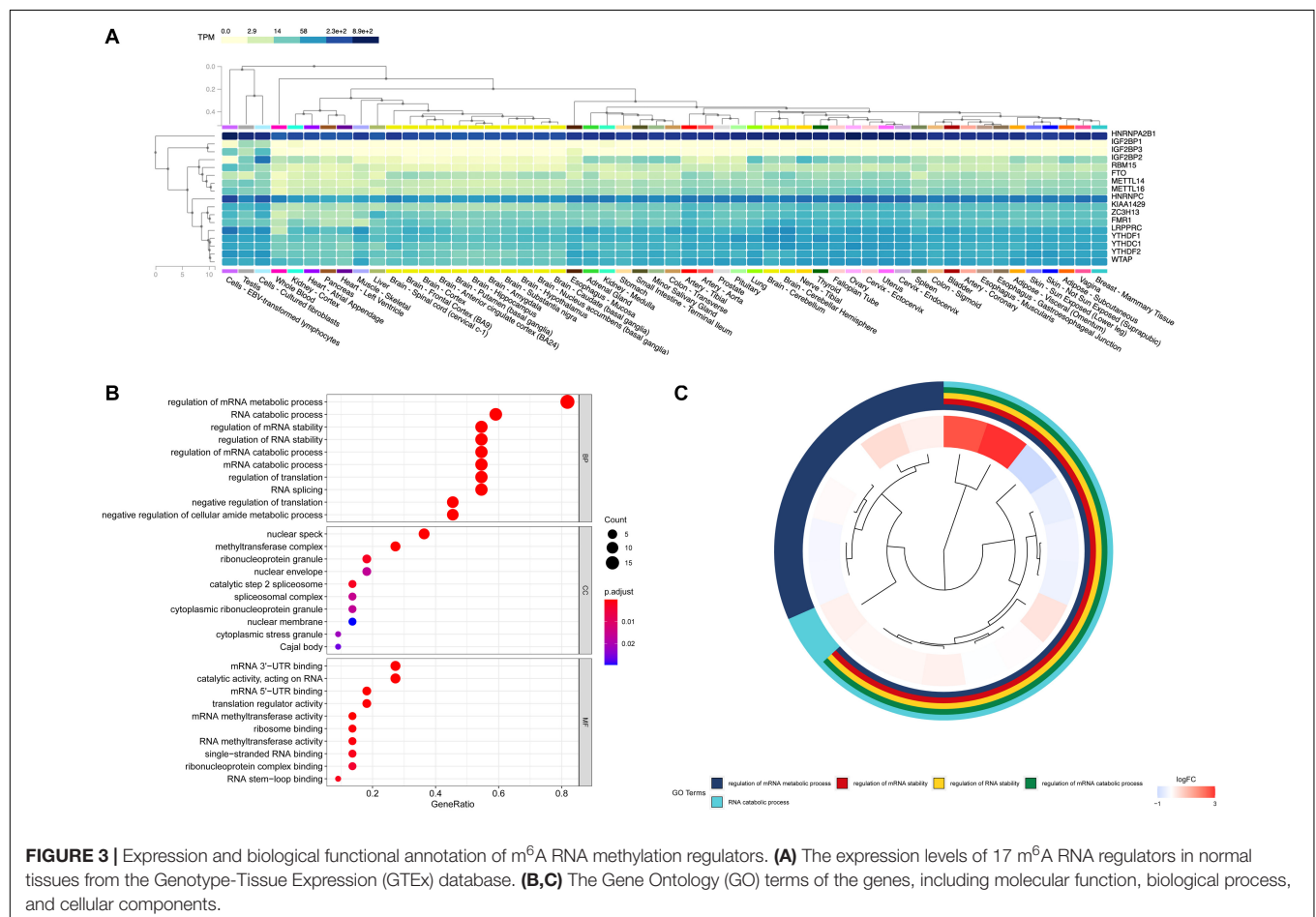
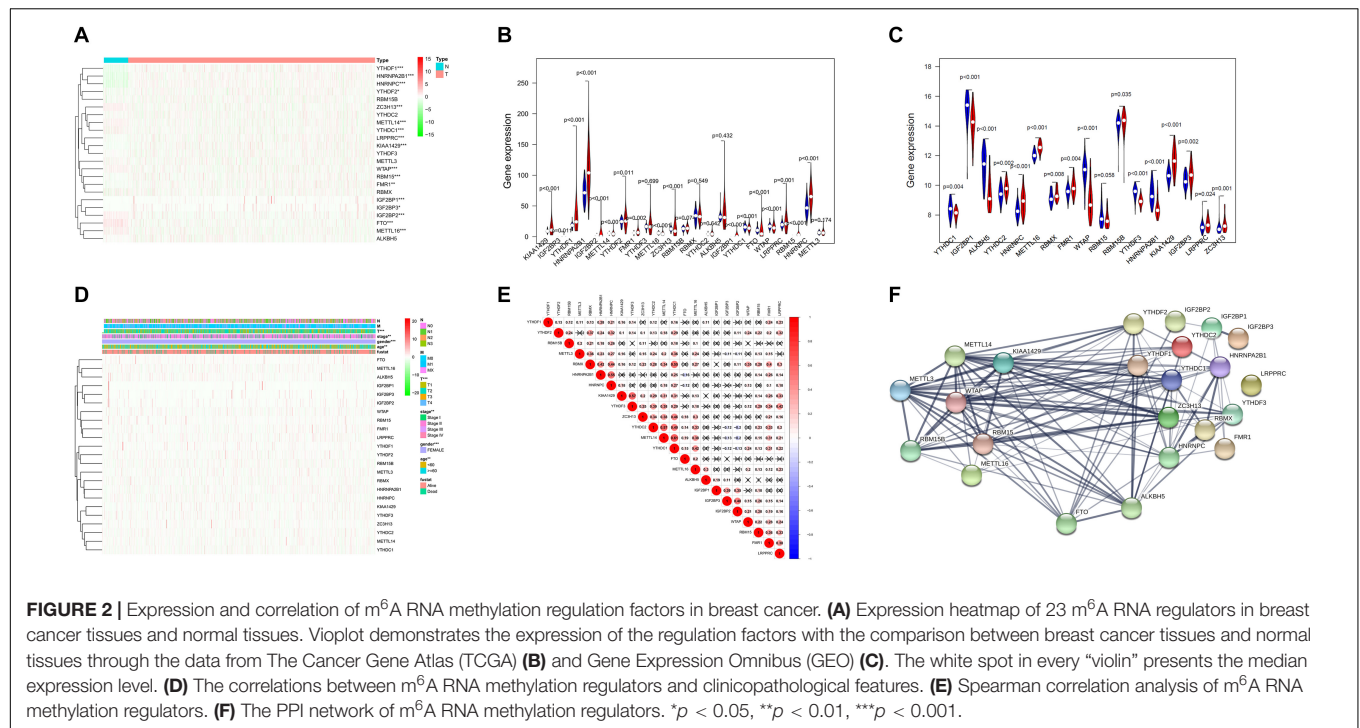
⁵<http://bcgenex.centregauducheau.fr/>

⁶<http://string-db.org>

⁷<https://www.gtexportal.org/>

⁸<http://www.r-project.org/>





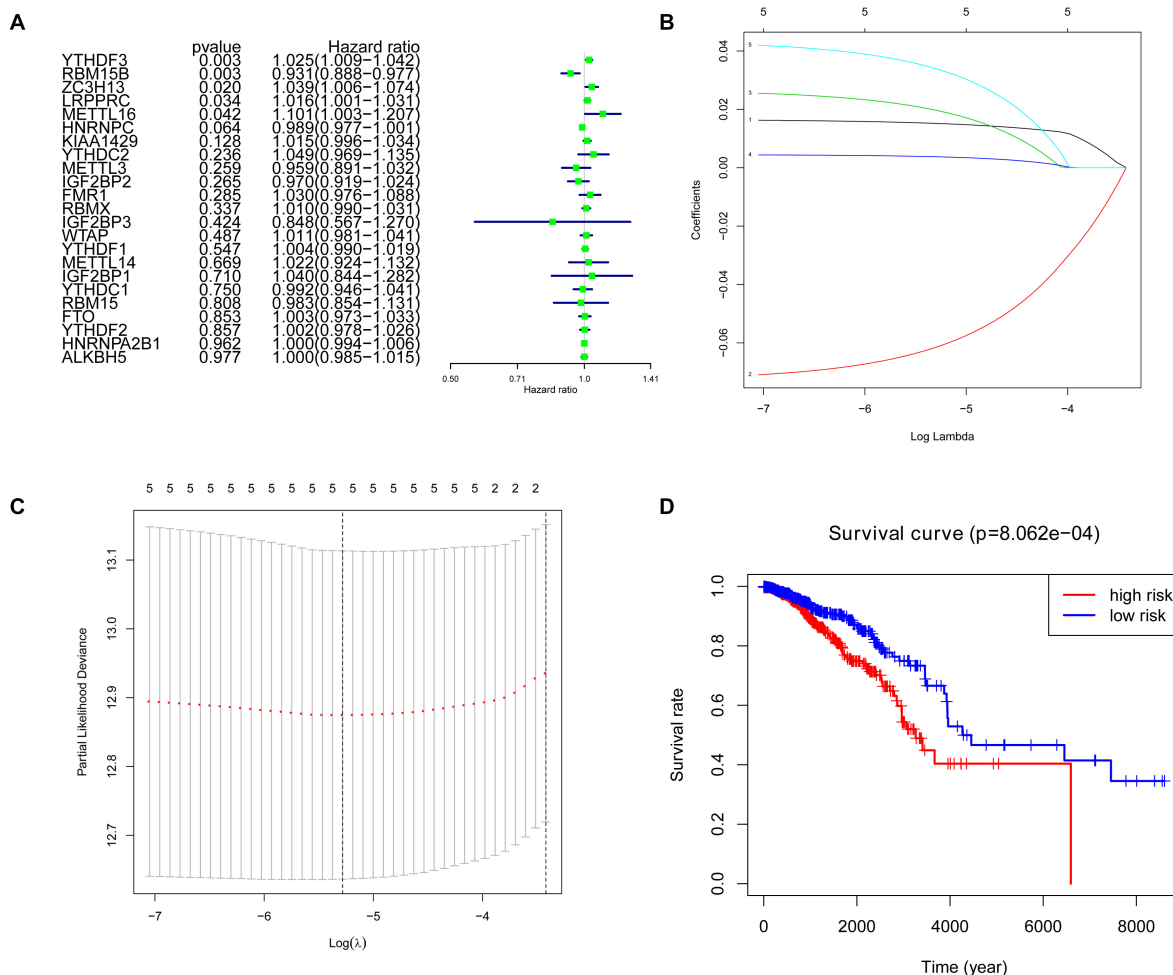
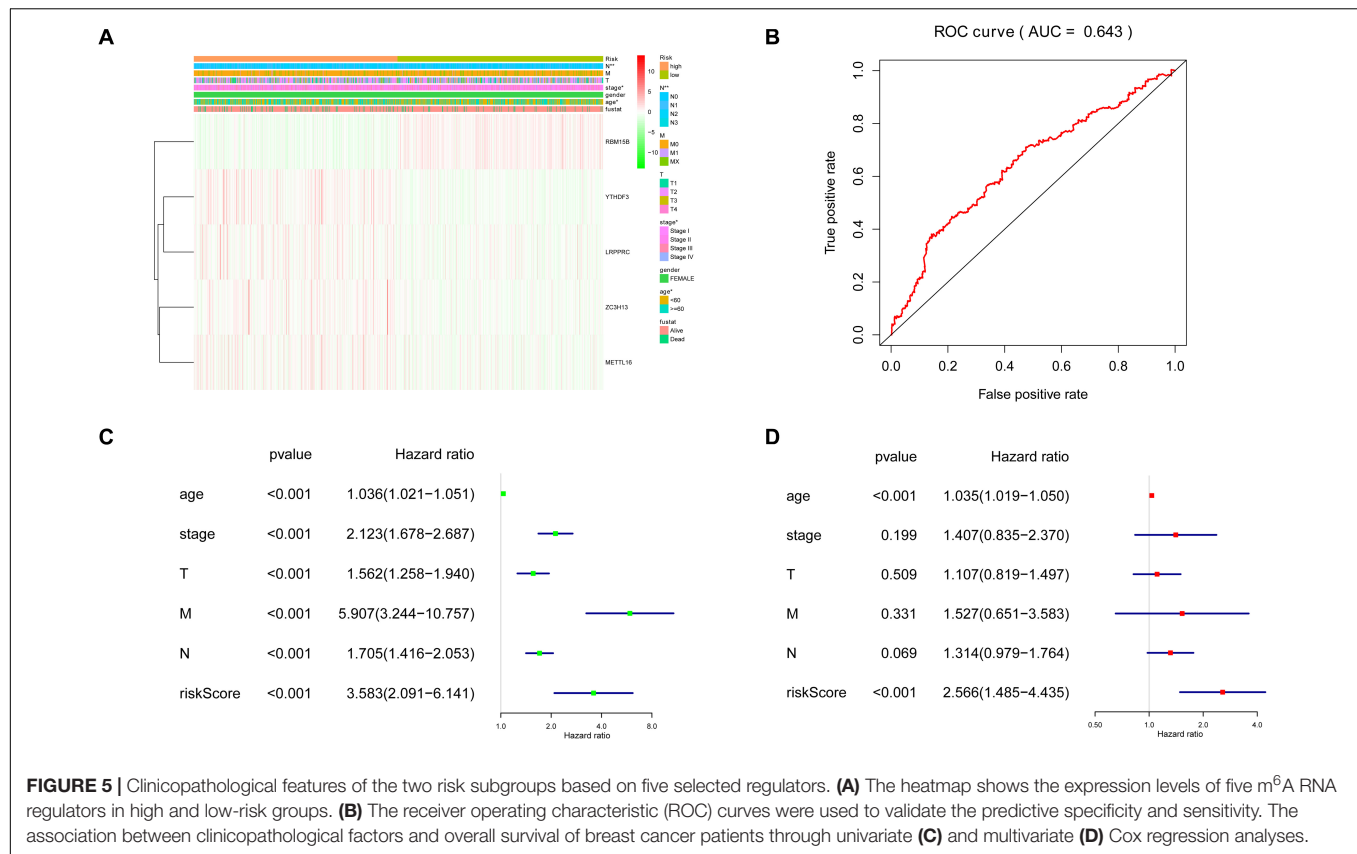


FIGURE 4 | Selection of clinical-pathological m⁶A RNA regulators with prognostic value. **(A)** Univariate Cox regression analysis showed the *p*-value, hazard ratio (HR), and 95% confidence interval (CI) of 23 m⁶A RNA methylation regulators. **(B,C)** Five regulators were selected for risk coefficient calculation through the least absolute shrinkage and selection operator (LASSO) Cox regression. **(D)** Kaplan-Meier Curve showed the overall survival of high and low-risk groups according to the risk score.

(Figure 5A). When the data was divided into high-risk and low-risk groups, it can be observed that RBM15B was in low expression, while YTHDF3, LRPPRC, ZC3H13, and METTL16 were in high expression in the high-risk group. We further performed the receiver operating characteristic (ROC) curve to evaluate the predictive specificity and sensitivity, and the area under the curve (AUC) was determined to be 0.643 (Figure 5B). Both univariate and multivariate Cox regression analyses indicated that age and risk scores were related to OS. Univariate analyses also revealed that stage, T status, N status, and M status were associated with OS (Figures 5C,D). After retrieving the RNA-seq data including the TCGA and SCAN-B databases, we got the comparison outcomes with ER, PR, HER-2 and PAM50 subtypes (Figure 6). LRPPRC was negatively associated with ER and PR expression, while METTL16, RBM15B, and ZC3H13 were positively associated with ER and PR expression. In HER-2 (+) breast cancer patients, LRPPRC, METTL16, RBM15B, and ZC3H13 expressed lower than the HER-2 (−). LRPPRC showed

the highest expression in basal-like breast cancer subtype. METTL16 expressed at a lower level in HER-E and basal-like subtypes. RBM15B expressed higher in Luminal A and normal breast-like subtypes of breast cancer. ZC3H13 expressed lower in basal-like, HER-E and Luminal B subtypes than in Luminal A and basal-like. However, no obvious difference among these five subtypes was observed with regards to YTHDF3 expression. These results were also verified by METABRIC database in bc-GenExMiner v4.5 (Supplementary Figure 1). LRPPRC and YTHDF3 were negatively associated with ER and PR expression, while METTL16 and RBM15B were positively associated with ER and PR expression. In HER-2 (+) breast cancer patients, METTL16 expressed lower than the HER-2 (−), but YTHDF3 expressed higher than the HER-2 (−). And we got similar results of PAM50 subtypes from METABRIC database. As a prognostic factor in breast cancer, the NPI and SBR histological grade is widely applied to predict tumor prognosis. Patients with high NPI and SBR grade tend to have a poor prognosis. Results from



bc-GenExMiner v4.5 showed that patients with higher NPI and SBR grade tended to have higher expression of LRPPRC and lower expression of METTL16, RBM15B, and ZC3H13 (Figure 7). These results were also confirmed by METABRIC database in bc-GenExMiner v4.5 (Supplementary Figure 2). Patients with higher NPI and SBR grade tend to have higher expression of LRPPRC and YTHDF3 and lower expression of METTL16 and RBM15B. The prognostic value of m⁶A RNA methylation regulators on RFS in breast cancer patients was analyzed via the Kaplan-Meier Plotter platform. The results in breast cancer patients showed that the relatively high expression of YTHDF3 and LRPPRC were remarkably associated with worse RFS, whereas relatively high expression of RBM15B, ZC3H13, and METTL16 had better RFS (Figure 8).

Associated Biological Pathways of Five m⁶A RNA Methylation Regulators

The GSEA analysis was performed to identify associated pathways. We selected significantly enriched signaling pathways based on their normalized enrichment score (NES) and normalized *p*-value (Figure 9). YTHDF3 was enriched in mTOR signaling pathway (NES = 1.98, *p* = 0.021), neurotrophin signaling pathway (NES = 1.98, *p* = 0.023), Notch signaling pathway (NES = 2.10, *p* = 0.012), pathways in cancer (NES = 1.84, *p* = 0.047) and Wnt signaling pathway (NES = 1.86, *p* = 0.042). RBM15B was enriched in aminoacyl tRNA biosynthesis (NES = 2.12, *p* = 0.011), mTOR signaling pathway

(NES = 1.94, *p* = 0.023), Notch signaling pathway (NES = 2.10, *p* = 0.008), pathways in cancer (NES = 1.82, *p* = 0.049), and Wnt signaling pathway (NES = 1.86, *p* = 0.044). ZC3H13 was enriched in aminoacyl tRNA biosynthesis (NES = 2.13, *p* = 0.005), lysine degradation (NES = 2.39, *p* < 0.001), mTOR signaling pathway (NES = 1.95, *p* = 0.024), Notch signaling pathway (NES = 2.15, *p* = 0.005), and Wnt signaling pathway (NES = 1.87, *p* = 0.038). LRPPRC was enriched in cell cycle (NES = 2.28, *p* = 0.002), homologous recombination pathway (NES = 1.95, *p* = 0.017), RNA degradation (NES = 2.64, *p* < 0.001), ubiquitin-mediated proteolysis (NES = 2.63, *p* < 0.001), and Wnt signaling pathway (NES = 2.01, *p* = 0.015). LRPPRC was down-regulated in arachidonic acid metabolism (NES = -2.31, *p* = 0.002). METTL16 was enriched in ErbB signaling pathway (NES = 1.96, *p* = 0.010), mismatch repair (NES = 2.01, *p* = 0.008), RNA degradation (NES = 2.59, *p* < 0.001), and Wnt signaling pathway (NES = 2.39, *p* = 0.004). METTL16 was down-regulated in arachidonic acid metabolism (NES = -2.26, *p* = 0.004).

DISCUSSION

The 5-year relative survival rate of breast cancer has been improved significantly due to the improvement of earlier diagnosis and treatment. However, there is still a long way to go for new therapeutic targets for breast cancer treatment. The m⁶A RNA methylation is an important post-transcriptional

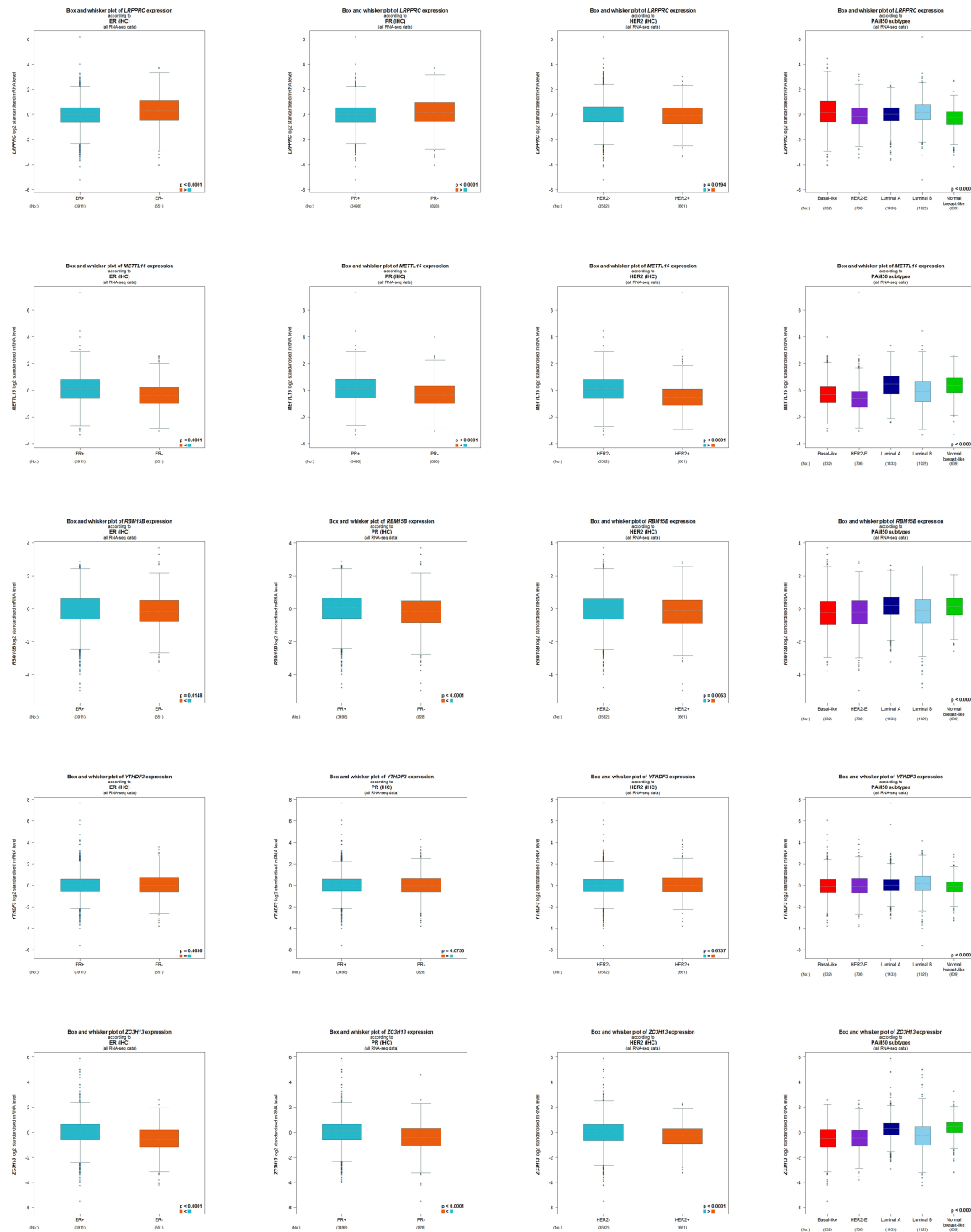


FIGURE 6 | RNA-seq data from bc-GenExMiner v4.5 showing the comparison outcomes with ER, PR, HER-2, and PAM50 subtypes on the expression level of YTHDF3, RBM15B, ZC3H13, LRRPPC, and METTL16.

factor that affects tumor occurrence and development. Previous studies show a close relationship between RNA methylation and breast cancer. Hypoxia-inducible factor (HIF)-1 α -and HIF-2 α -dependent ALKBH5 in breast cancer cells can be

stimulated under hypoxia condition, leading to demethylation of NANOG mRNA. Furthermore, the high expression level of NANOG leads to the enrichment of breast cancer stem cells (BCSCs) (Zhang et al., 2016). Cai et al. (2018) reported that the

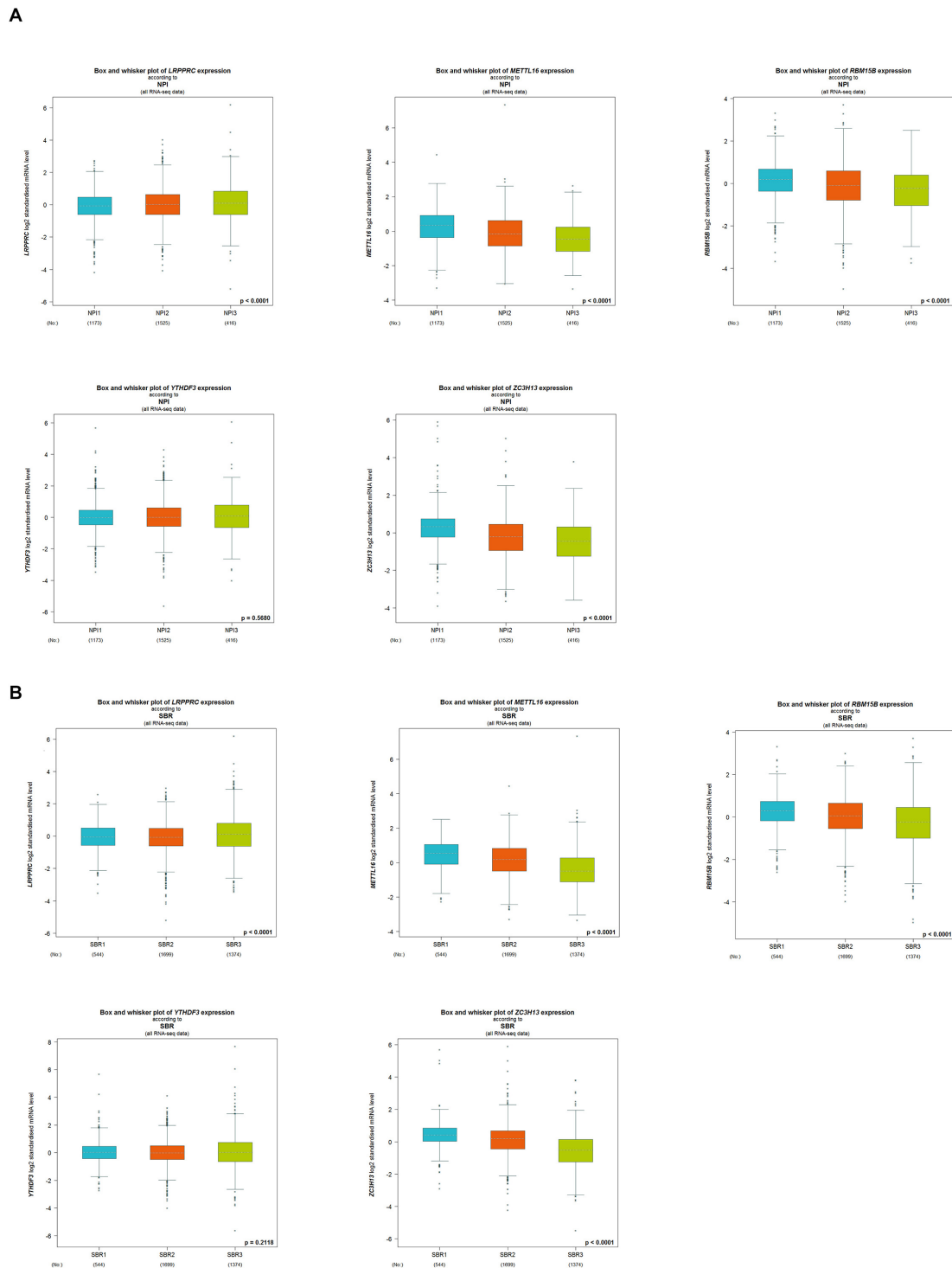
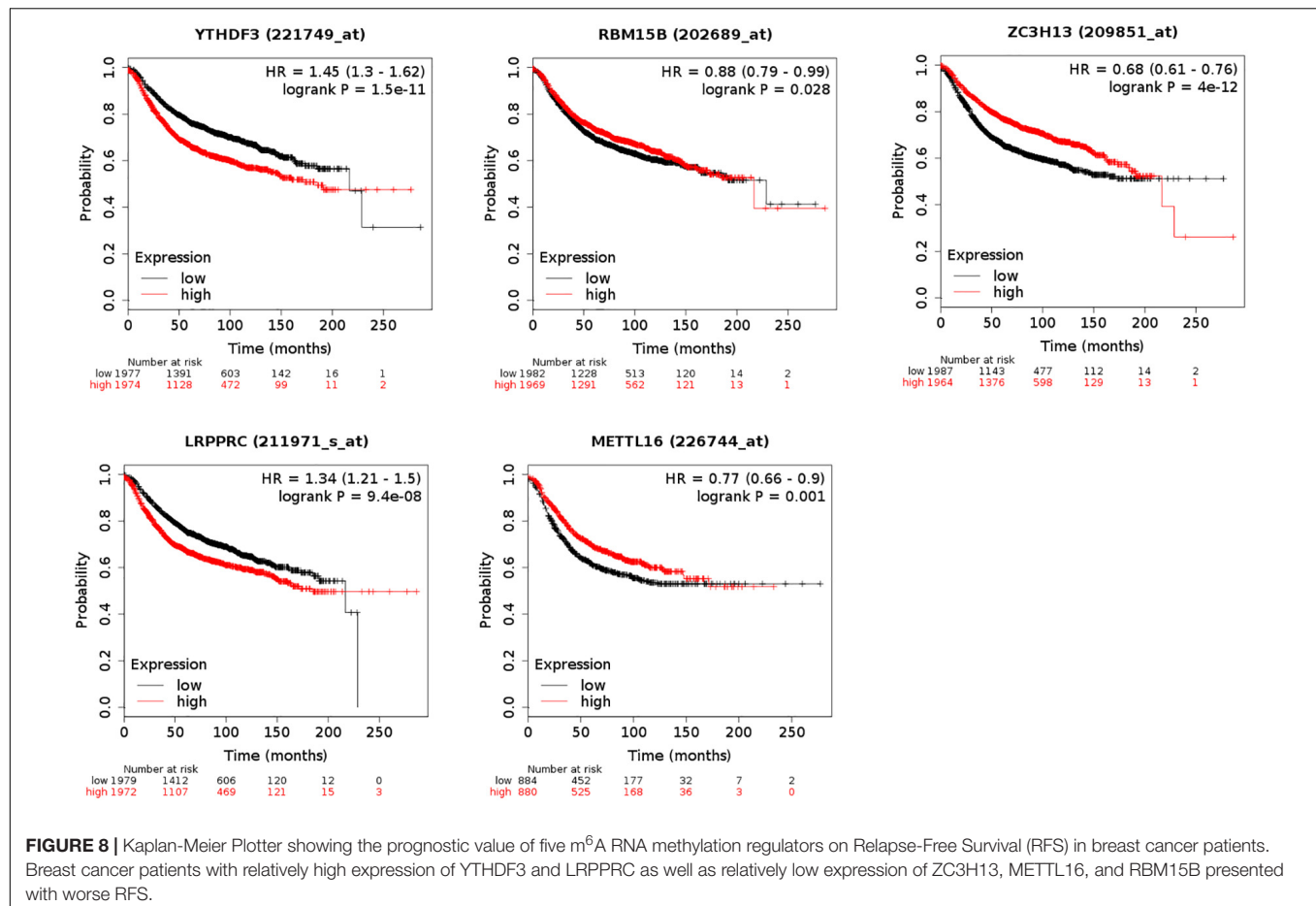


FIGURE 7 | The correlation between the expression level of LRPPRC, METTL16, RBM15B, YTHDF3, and ZC3H13 and the Nottingham prognostic index (NPI) (**A**) and the Scarff-Bloom-Richardson (SBR) (**B**) grading. Patients with higher NPI and SBR grade tend to have higher expression of LRPPRC and lower expression of METTL16, RBM15B, and ZC3H13 ($p < 0.0001$).

mTORC1 activator mammalian hepatitis B X-interacting protein (HBXIP) can active METTL3 by prohibiting the expression of oncogenes-related miRNA let-7g. While the up-regulated

METTL3 can simultaneously promote the HBXIP expression. Thereby, HBXIP, METTL3 and let-7g form a feedback loop accelerating proliferation and invasion of breast cancer. The



“writer” KIAA1429 is confirmed to be up-regulated in breast cancer, and 5'-fluorouracil is effective for exhibiting the expression of KIAA1429 and its methylated target cyclin-dependent kinase 1 (CDK1) (Qian et al., 2019). These studies mostly concentrate on few methylation regulators, such as METTL3, METTL14, KIAA1429, FTO, and ALKBH5. With the continuing discovery of novel m⁶A regulators, there is an ever-increasing demand for in-depth research into RNA methylation about breast cancer.

In our present study, we analyzed the expression of the 23 RNA methylation regulators in the round from ONCOMINE database. Based on 1,096 tumor tissues and 112 normal tissues from the TCGA database, the analytic outcomes showed that 17 out of 23 regulators had statistical significance ($p < 0.05$). Surprisingly, some researches have verified that the repression of HNRNPC can inhibit two breast cancer cell lines (MCF7 and T47D) proliferation through the accumulation of double-stranded RNA (dsRNA), the binding ligand of RIG-I, and anti-proliferation activity induced by the stimulated IFN cytokine (Wu et al., 2018). The silencing of the m⁶A methyltransferase has an effect on the stability of gene expression, such as the interference of the TP53 signaling pathway which is a vital tumor suppressor gene (Dominissini et al., 2012). The p53 protein binds closely with HNRNPC and causes destabilization

of HNRNPC. Under the treatment of doxorubicin, lncRNA SNHG1 and HNRNPC can reach a balance on the mediation of p53 (Shen et al., 2017). In line with the TCGA outcomes, up-regulated HNRNPA2B1 is concluded to promote breast cancer tumorigenic potential by activating extracellular-signal-regulated kinase 1/2 (ERK1/2), and signal transducer and activator of transcription 3 (STAT3). Knockdown of HNRNPA2B1 can prohibit proliferation, accelerate apoptosis and prolong S phase *in vitro*, while restraining tumorigenicity *in vivo* (Hu et al., 2017).

The m⁶A RNA methylation regulators are of great significance in the occurrence, development and prognosis of various types of cancers. Recently, by using bioinformatic databases, researchers have discovered its importance in tumor therapies. The variations and/or mutations of m⁶A RNA methylation regulators occurring in the acute myeloid leukemia (AML) patients present a close relation with the mutations of TP53 may disclose the poor prognosis in AML (Kwok et al., 2017). Patients with high levels of mRNA methylation resulting from overexpression of reader proteins and methyltransferase complexes showed poor survival benefits in prostate cancer (Ji et al., 2020). In the subgroup construction signature of TCGA and GEO database, the m⁶A regulators showed a great value in malignant progression and prognosis of uveal melanoma (UM) (Tang et al., 2020).

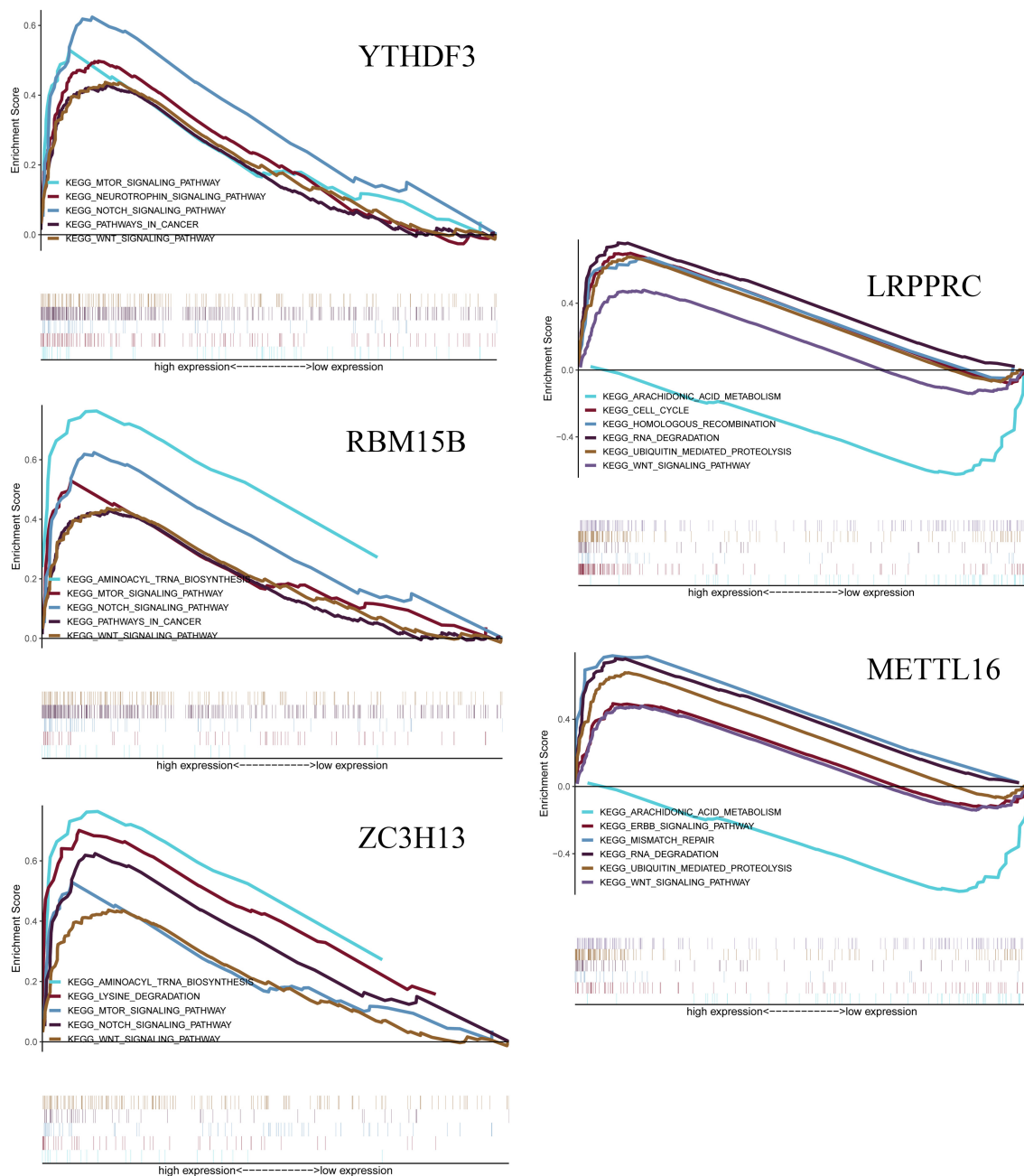


FIGURE 9 | Gene Set Enrichment Analysis (GSEA) shows that the five m⁶A RNA methylation regulators, including YTHDF3, RBM15B, ZC3H13, LRPPRC, METTL16. The results are based on their normalized enrichment score (NES) and normalized *p*-value. Pathways with positive NES are positively correlated with the genes, and pathways with negative NES are negatively correlated with the genes.

According to the clinical database of breast cancer of TCGA, RBM15B, YTHDF3, ZC3H13, METTL16, and LRPPRC were picked up indicating strong associations with clinical characteristics. We used a variety of databases to verify their impact on the survival and prognosis of breast cancer patients and its potential research significance. The outcome revealed a clear direction for further study *in vivo* and *in vitro*. RBM15B is the paralog of RBM15, and they

are confirmed to bind METTL3 relying on the meditation of WTAP (Patil et al., 2016). BRCA1-associated protein-1 (BAP1) is found to be expressed at a low level in breast cancer patients, while RBM15B has a positive correlation with BAP1 in invasive breast cancer (Shahriyari et al., 2019). YTH domain protein family contains several m⁶A-binding proteins playing parts as readers in m⁶A methylation but with different functions. YTHDF1 and YTHDF3 work cooperatively

to promote ribosome loading, while YTHDF2 accelerates the decay of mRNA and YTHDC1 acts as a recruiter for mRNA splicing (Xiao et al., 2016; Hsu et al., 2017; Li et al., 2017). YTHDF3 has been verified to correlate with several kinds of carcinoma, such as colorectal cancer (CRC) (Liu et al., 2019), gastric cancer (Zhang et al., 2019), bladder cancer (Jin et al., 2019), breast cancer (Liu et al., 2019), and so on. In the ONCOMINE and TCGA databases, we didn't notice a considerable difference in the expression of RBM15B and YTHDF3 in breast cancer. Whether it has an impact on the occurrence, development and prognosis of breast cancer needs to be further confirmed by research. ZC3H13 is a novel "writer" mediating m⁶A methylation. The ablation of ZC3H13 can lead to an obvious decrease of m⁶A enrichment, especially at the 3' end of the mRNA. It is also speculated that ZC3H13 is a component of the RBM15/ZC3H13/WTAP/VIRMA/HAIKAI complex, also refers to the m⁶A-METTL-associated complex, which regulates m⁶A methylation either on its own or interacts with METTL3/METTL14 complex (Knuckles et al., 2018). In the triple-negative subtype of inflammatory breast cancer (TN-IBC), ZC3H13 is down-regulated in contrast to TN-non-IBC (Funakoshi et al., 2019). The outcome of this research is consistent with the results of our bioinformatics analysis. METTL16 was first found to have a close correlation with a cancer-promoting long non-coding RNA: metastasis-associated lung adenocarcinoma transcript 1 (MALAT1) (Brown et al., 2016). As a cancer-promoting factor, MALAT1 is closely related to certain biological characteristics such as vascular epithelialization, migration and distant metastasis, postoperative fever, and systemic inflammatory response in breast cancer (Huang et al., 2018; Li et al., 2018; Gomes et al., 2019). METTL16 is also reported in U6 small nuclear (snRNA), pre-mRNAs and other lncRNA (Warda et al., 2017). METTL16 binds the UACAGAGAA sequence on the 3' UTR of MAT2A mRNA sequence for MAT2A splicing, the procedure of which can encode the SAM synthetase expression and form a methylation pattern regulated according to SAM level which is quite different from the way that METTL3-METTL14-WTAP complex acts (Pendleton et al., 2017). Former research has pointed out that the biochemical synthesis of SAM is a unique metabolic "soft spot" of triple-negative BCSCs whose drug-resistance can be eradicated by combining methionine depletion with MAT2A suppression (Strekalova et al., 2019). Although these regulators failed to present desired and consistent outcomes in the analysis of expression levels and prognostic values, it's worth further exploring their biological functions and interactions in breast cancer. Besides, the subtypes of breast cancer may interfere in RNA methylation which indicates new research interests in exploring different subtypes. LRPPRC is an RNA-binding protein regulating mitochondrial DNA-encoded mRNA and also transactivating nuclear DNAs. The immunohistochemical detection of several types of cancers, including lung adenocarcinoma, breast cancer, oesophageal squamous cell cancer, gastric adenocarcinoma, etc., shows that LRPPRC is abundantly expressed in cancer tissues in contrast to the adjacent normal tissues. It can promote apoptosis resistance and invasive activity of cancer cells (Tian et al., 2012). Research

on BCSCs shows that the over-expression of microRNA-1 (miR-1) can target the 3' UTRs of mitochondrial inner membrane organizing system 1 (MINOS1) and glycerol-3-phosphate dehydrogenase 2 (GPD2) under the participation of LRPPRC, which results in mitochondrial damage (Zhang et al., 2019). However, similar effects upon m⁶A RNA methylation need to be proved.

m⁶A RNA methylation is a complex regulative system that acts through methylation and demethylation on different targets. Its effect is also reflected in "promoter" or "suppressor" in tumorigenesis and development. With the improvement of RNA methylation detection technology, an increasing number of regulators are found, which reveals a more complicated regulation system. Further studies are needed to be done to verify these findings and their interactive mechanisms. Besides, the way to control specificity and to strengthen monitoring of RNA methylation regulators in the dynamic post-transcriptional balance is also a great challenge. But we believe, as a promising research area in epigenetics, RNA methylation would bring a new dawn of hope for combating breast cancer.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: <http://www.cancergenome.nih.gov/> and <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE70905>.

AUTHOR CONTRIBUTIONS

GJ supervised and conceptualized the study. BZ searched the articles and designed the study. BZ and YG made the figures and wrote this manuscript. All authors read and revised the final manuscript.

FUNDING

This study was supported by the Suzhou Health Planning Commission's Key Clinical Diagnosis, Treatment Program (LCZX201606), the National Natural Science Foundation of China (81873730), and the Jiangsu Women and Children Health Key Discipline Program (FXK201758).

ACKNOWLEDGMENTS

We acknowledge the reviewers for their constructive comments. We thank Sancy Mary Joyson for English language editing.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2020.604597/full#supplementary-material>

REFERENCES

- Boccalletto, P., Machnicka, M. A., Purta, E., Piatkowski, P., Baginski, B., Wirecki, T. K., et al. (2018). MODOMICS: a database of RNA modification pathways. *Nucleic Acids Res.* 46, D303–D307. doi: 10.1093/nar/gkx1030
- Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R. L., Torre, L. A., and Jemal, A. (2018). Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* 68, 394–424. doi: 10.3322/caac.21492
- Brown, J. A., Kinzig, C. G., Degregorio, S. J., and Steitz, J. A. (2016). Methyltransferase-like protein 16 binds the 3'-terminal triple helix of MALAT1 long noncoding RNA. *Proc. Natl. Acad. Sci. USA* 113, 14013–14018. doi: 10.1073/pnas.1614759113
- Cai, X., Wang, X., Cao, C., Gao, Y., Zhang, S., Yang, Z., et al. (2018). HBXIP-elevated methyltransferase METTL3 promotes the progression of breast cancer via inhibiting tumor suppressor let-7g. *Cancer Lett.* 415, 11–19. doi: 10.1016/j.canlet.2017.11.018
- Dominissini, D., Moshitch-Moshkovitz, S., Schwartz, S., Salmon-Divon, M., Ungar, L., Osenberg, S., et al. (2012). Topology of the human and mouse m⁶A RNA methylomes revealed by m⁶A-seq. *Nature* 485, 201–206. doi: 10.1038/nature11112
- Funakoshi, Y., Wang, Y., Semba, T., Masuda, H., Hout, D., Ueno, N. T., et al. (2019). Comparison of molecular profile in triple-negative inflammatory and non-inflammatory breast cancer not of mesenchymal stem-like subtype. *PLoS One* 14:e0222336. doi: 10.1371/journal.pone.0222336
- Fustin, J. M., Kojima, R., Itoh, K., Chang, H. Y., Ye, S., Zhuang, B., et al. (2018). Two Ckl8 transcripts regulated by m⁶A methylation code for two antagonistic kinases in the control of the circadian clock. *Proc. Natl. Acad. Sci. USA* 115, 5980–5985. doi: 10.1073/pnas.1721371115
- Gomes, C. P., Nobrega-Pereira, S., Domingues-Silva, B., Rebelo, K., Alves-Vale, C., Marinho, S. P., et al. (2019). An antisense transcript mediates MALAT1 response in human breast cancer. *BMC Cancer* 19:771. doi: 10.1186/s12885-019-5962-0
- Hsu, P. J., Zhu, Y., Ma, H., Guo, Y., Shi, X., Liu, Y., et al. (2017). Ythdc2 is an N(6)-methyladenosine binding protein that regulates mammalian spermatogenesis. *Cell Res.* 27:99. doi: 10.1038/cr.2017.99
- Hu, Y., Sun, Z., Deng, J., Hu, B., Yan, W., Wei, H., et al. (2017). Splicing factor hnRNP A2B1 contributes to tumorigenic potential of breast cancer cells through STAT3 and ERK1/2 signaling pathway. *Tumour Biol.* 39:1010428317694318. doi: 10.1177/1010428317694318
- Huang, X. J., Xia, Y., He, G. F., Zheng, L. L., Cai, Y. P., Yin, Y., et al. (2018). MALAT1 promotes angiogenesis of breast cancer. *Oncol. Rep.* 40, 2683–2689. doi: 10.3892/or.2018.6705
- Ji, G., Huang, C., He, S., Gong, Y., Song, G., Li, X., et al. (2020). Comprehensive analysis of m⁶A regulators prognostic value in prostate cancer. *Aging* 12, 14863–14884. doi: 10.18632/aging.103549
- Jin, H., Ying, X., Que, B., Wang, X., Chao, Y., Zhang, H., et al. (2019). N(6)-methyladenosine modification of ITGA6 mRNA promotes the development and progression of bladder cancer. *EBio Med.* 47, 195–207. doi: 10.1016/j.ebiom.2019.07.068
- Knuckles, P., Lence, T., Haussmann, I. U., Jacob, D., Kreim, N., Carl, S. H., et al. (2018). Zc3h13/Flacc is required for adenosine methylation by bridging the mRNA-binding factor Rbm15/Spenito to the m(6)A machinery component Wtap/Fl(2)d. *Genes Dev.* 32, 415–429. doi: 10.1101/gad.309146.117
- Kwok, C. T., Marshall, A. D., Rasko, J. E., and Wong, J. J. (2017). Genetic alterations of m(6)A regulators predict poorer survival in acute myeloid leukemia. *J. Hematol. Oncol.* 10:39. doi: 10.1186/s13045-017-0410-6
- Lan, Q., Liu, P. Y., Haase, J., Bell, J. L., Huttelmaier, S., and Liu, T. (2019). The Critical Role of RNA m(6)A Methylation in Cancer. *Cancer Res.* 79, 1285–1292. doi: 10.1158/0008-5472.Can-18-2965
- Lence, T., Akhtar, J., Bayer, M., Schmid, K., Spindler, L., Ho, C. H., et al. (2016). m(6)A modulates neuronal functions and sex determination in Drosophila. *Nature* 540, 242–247. doi: 10.1038/nature20568
- Li, A., Chen, Y. S., Ping, X. L., Yang, X., Xiao, W., Yang, Y., et al. (2017). Cytoplasmic m(6)A reader YTHDF3 promotes mRNA translation. *Cell Res.* 27, 444–447. doi: 10.1038/cr.2017.10
- Li, Z., Xu, L., Liu, Y., Fu, S., Tu, J., Hu, Y., et al. (2018). LncRNA MALAT1 promotes relapse of breast cancer patients with postoperative fever. *Am. J. Transl. Res.* 10, 3186–3197.
- Liu, L., Liu, X., Dong, Z., Li, J., Yu, Y., Chen, X., et al. (2019). N6-methyladenosine-related Genomic Targets are Altered in Breast Cancer Tissue and Associated with Poor Survival. *J. Cancer* 10:35053. doi: 10.7150/jca.35053
- Liu, X., Liu, L., Dong, Z., Li, J., Yu, Y., Chen, X., et al. (2019). Expression patterns and prognostic value of m(6)A-related genes in colorectal cancer. *Am. J. Transl. Res.* 11, 3972–3991.
- Meyer, K. D., Saletore, Y., Zumbo, P., Elemento, O., Mason, C. E., and Jaffrey, S. R. (2012). Comprehensive analysis of mRNA methylation reveals enrichment in 3' UTRs and near stop codons. *Cell* 149, 1635–1646. doi: 10.1016/j.cell.2012.05.003
- Miller, K. D., Nogueira, L., Mariotto, A. B., Rowland, J. H., Yabroff, K. R., Alfano, C. M., et al. (2019). Cancer treatment and survivorship statistics. *CA A Cancer J. Clin.* 69, 363–385. doi: 10.3322/caac.21565
- Patil, D. P., Chen, C. K., Pickering, B. F., Chow, A., Jackson, C., Guttman, M., et al. (2016). m(6)A RNA methylation promotes XIST-mediated transcriptional repression. *Nature* 537, 369–373. doi: 10.1038/nature19342
- Pendleton, K. E., Chen, B., Liu, K., Hunter, O. V., Xie, Y., Tu, B. P., et al. (2017). The U6 snRNA m(6)A Methyltransferase METTL16 Regulates SAM Synthetase Intron Retention. *Cell* 169, 824–835.e814. doi: 10.1016/j.cell.2017.05.003
- Qian, J. Y., Gao, J., Sun, X., Cao, M. D., Shi, L., Xia, T. S., et al. (2019). KIAA1429 acts as an oncogenic factor in breast cancer by regulating CDK1 in an N6-methyladenosine-independent manner. *Oncogene* 38, 6123–6141. doi: 10.1038/s41388-019-0861-z
- Roundtree, I. A., Evans, M. E., Pan, T., and He, C. (2017). Dynamic RNA Modifications in Gene Expression Regulation. *Cell* 169, 1187–1200. doi: 10.1016/j.cell.2017.05.045
- Shahriyari, L., Abdel-Rahman, M., and Cebulla, C. (2019). BAP1 expression is prognostic in breast and uveal melanoma but not colon cancer and is highly positively correlated with RBM15B and USP19. *PLoS One* 14:e0211507. doi: 10.1371/journal.pone.0211507
- Shen, Y., Liu, S., Fan, J., Jin, Y., Tian, B., Zheng, X., et al. (2017). Nuclear retention of the lncRNA SNHG1 by doxorubicin attenuates hnRNP C-p53 protein interactions. *EMBO Rep.* 18, 536–548. doi: 10.15252/embr.201643139
- Shi, H., Wei, J., and He, C. (2019). Where, When, and How: Context-Dependent Functions of RNA Methylation Writers, Readers, and Erasers. *Mol. Cell* 74, 640–650. doi: 10.1016/j.molcel.2019.04.025
- Stekalova, E., Malin, D., Weisenhorn, E. M. M., Russell, J. D., Hoelper, D., Jain, A., et al. (2019). S-adenosylmethionine biosynthesis is a targetable metabolic vulnerability of cancer stem cells. *Breast Cancer Res. Treat.* 175, 39–50. doi: 10.1007/s10549-019-05146-7
- Tang, J., Wan, Q., and Lu, J. (2020). The prognostic values of m6A RNA methylation regulators in uveal melanoma. *BMC Cancer* 20:674. doi: 10.1186/s12885-020-07159-8
- Tian, T., Ikeda, J., Wang, Y., Mamat, S., Luo, W., Aozasa, K., et al. (2012). Role of leucine-rich pentatricopeptide repeat motif-containing protein (LRPPRC) for anti-apoptosis and tumorigenesis in cancers. *Eur. J. Cancer* 48, 2462–2473. doi: 10.1016/j.ejca.2012.01.018
- Wang, Y., Li, Y., Toth, J. I., Petroski, M. D., Zhang, Z., and Zhao, J. C. (2014). N6-methyladenosine modification destabilizes developmental regulators in embryonic stem cells. *Nat. Cell Biol.* 16, 191–198. doi: 10.1038/ncb2902
- Warda, A. S., Kretschmer, J., Hackert, P., Lenz, C., Urlaub, H., Hobartner, C., et al. (2017). Human METTL16 is a N(6)-methyladenosine (m(6)A) methyltransferase that targets pre-mRNAs and various non-coding RNAs. *EMBO Rep.* 18, 2004–2014. doi: 10.15252/embr.201744940
- Wu, Y., Zhao, W., Liu, Y., Tan, X., Li, X., Zou, Q., et al. (2018). Function of HNRNPC in breast cancer cells by controlling the dsRNA-induced interferon response. *EMBO J.* 37:e99017. doi: 10.15252/embj.201899017
- Xiao, W., Adhikari, S., Dahal, U., Chen, Y. S., Hao, Y. J., Sun, B. F., et al. (2016). Nuclear m(6)A Reader YTHDC1 Regulates mRNA Splicing. *Mol. Cell* 61, 507–519. doi: 10.1016/j.molcel.2016.01.012
- Yang, Y., Hsu, P. J., Chen, Y. S., and Yang, Y. G. (2018). Dynamic transcriptomic m(6)A decoration: writers, erasers, readers and functions in RNA metabolism. *Cell Res.* 28, 616–624. doi: 10.1038/s41422-018-0040-8

- Zhang, C., Chen, Y., Sun, B., Wang, L., Yang, Y., Ma, D., et al. (2017). m⁶A modulates haematopoietic stem and progenitor cell specification. *Nature* 549, 273–276.
- Zhang, C., Samanta, D., Lu, H., Bullen, J. W., Zhang, H., Chen, I., et al. (2016). Hypoxia induces the breast cancer stem cell phenotype by HIF-dependent and ALKBH5-mediated m(6)A-demethylation of NANOG mRNA. *Proc. Natl. Acad. Sci. USA* 113, E2047–E2056. doi: 10.1073/pnas.1602883113
- Zhang, C., Zhang, M., Ge, S., Huang, W., Lin, X., Gao, J., et al. (2019). Reduced m⁶A modification predicts malignant phenotypes and augmented Wnt/PI3K-Akt signaling in gastric cancer. *Cancer Med.* 8, 4766–4781. doi: 10.1002/cam4.2360
- Zhang, S., Liu, C., and Zhang, X. (2019). Mitochondrial Damage Mediated by miR-1 Overexpression in Cancer Stem Cells. *Mol. Ther. Nucleic Acids* 18, 938–953. doi: 10.1016/j.omtn.2019.10.016
- Zhou, J., Wan, J., Gao, X., Zhang, X., Jaffrey, S. R., and Qian, S. B. (2015). Dynamic m(6)A mRNA methylation directs translational control of heat shock response. *Nature* 526, 591–594. doi: 10.1038/nature15377
- 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 Zhang, Gu and Jiang. 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.



Structural Insights Into m⁶A-Erasers: A Step Toward Understanding Molecule Specificity and Potential Antiviral Targeting

Mahmoud Bayoumi^{1,2} and Muhammad Munir^{1*}

¹ Division of Biomedical and Life Sciences, Lancaster University, Lancaster, United Kingdom, ² Virology Department, Faculty of Veterinary Medicine, Cairo University, Giza, Egypt

OPEN ACCESS

Edited by:

Jia Meng,
Xi'an Jiaotong-Liverpool
University, China

Reviewed by:

Sam Thiagalingam,
Boston University, United States
Guifang Jia,
Peking University, China

*Correspondence:

Muhammad Munir
muhammad.munir@lancaster.ac.uk

Specialty section:

This article was submitted to
Epigenomics and Epigenetics,
a section of the journal
Frontiers in Cell and Developmental
Biology

Received: 24 July 2020

Accepted: 26 November 2020

Published: 12 January 2021

Citation:

Bayoumi M and Munir M (2021)
Structural Insights Into m⁶A-Erasers:
A Step Toward Understanding
Molecule Specificity and Potential
Antiviral Targeting.
Front. Cell Dev. Biol. 8:587108.
doi: 10.3389/fcell.2020.587108

The cellular RNA can acquire a variety of chemical modifications during the cell cycle, and compelling pieces of evidence highlight the importance of these modifications in determining the metabolism of RNA and, subsequently, cell physiology. Among myriads of modifications, methylation at the N⁶-position of adenosine (m⁶A) is the most important and abundant internal modification in the messenger RNA. The m⁶A marks are installed by methyltransferase complex proteins (writers) in the majority of eukaryotes and dynamically reversed by demethylases such as FTO and ALKBH5 (erasers). The incorporated m⁶A marks on the RNA transcripts are recognized by m⁶A-binding proteins collectively called readers. Recent epigenetic studies have unequivocally highlighted the association of m⁶A demethylases with a range of biomedical aspects, including human diseases, cancers, and metabolic disorders. Moreover, the mechanisms of demethylation by m⁶A erasers represent a new frontier in the future basic research on RNA biology. In this review, we focused on recent advances describing various physiological, pathological, and viral regulatory roles of m⁶A erasers. Additionally, we aim to analyze structural insights into well-known m⁶A-demethylases in assessing their substrate binding-specificity, efficiency, and selectivity. Knowledge on cellular and viral RNA metabolism will shed light on m⁶A-specific recognition by demethylases and will provide foundations for the future development of efficacious therapeutic agents to various cancerous conditions and open new avenues for the development of antivirals.

Keywords: ALKBH5, antiviral, demethylation, epigenetics, evolution, FTO, structural insights, m⁶A

INTRODUCTION

Epitranscriptome is an emerging area of biology that collectively describes over 100 chemical modifications to various forms of RNAs, including messenger RNA (mRNA), transfer RNA (tRNA), ribosomal RNA, and long non-coding RNAs (lncRNAs). These chemical modifications display an extensive landscape that regulates multiple biological processes (Roundtree et al., 2017). RNA can accept one or more chemical modifications to different bases, including cytosine (m⁵C) (Motorin et al., 2009), adenosine (m¹A) (Li X. et al., 2017; Safra et al., 2017), pseudouridine (Carlile et al., 2014), and inosine (Levanon et al., 2004). However, methylation at the N⁶ position of adenosine (m⁶A) is considered the most prominent modification (Dominissini et al., 2012; Boccaletto et al., 2018).

In addition to cellular RNA, the m⁶A marks are also incorporated into the viral RNA (Krug et al., 1976; Kane and Beemon, 1985; Narayan et al., 1987; Tirumuru et al., 2016; Courtney et al., 2017; Kennedy et al., 2017), hence highlighting unexplored aspects of host–pathogen interactions.

During the physiological regulatory processes, the methylation process is embarked on by the m⁶A methyltransferase complex. Conversely, to reverse the m⁶A marks, the RNA demethylases are required to alleviate the effects of various installed chemical modifications and/or dynamically reverse RNA changes to perform a specified function in cell life cycles (Han et al., 2010). Various mammalian alkylated DNA repair protein (AlkB) homologs share the same basic structure to nine publicly known AlkB protein members (Sundheim et al., 2008; Yang et al., 2008; Aik et al., 2012; Wang et al., 2014). The prototype AlkB gene/protein was firstly identified in *Escherichia coli* strains in the 80s (Kataoka et al., 1983); however, the detailed functions of AlkB proteins in repairing the damage arise from alkylation were described in the 2000s. The bacterial AlkB protein has a broad range of specificity to various nucleobases (Falnes et al., 2002; Delaney and Essigmann, 2004; Delaney et al., 2005; Alemu et al., 2016). Unlike the multifunctional prokaryotic AlkB, the higher-order eukaryotic AlkB homologs, such as ALKBH1–8 and the FTO, have only limited functions with higher substrate specificity for either epigenetic modifications and/or nucleic acids repair function (Falnes et al., 2002).

Human AlkB Homolog-1 (hALKBH1) protein was first documented to repair 3-methylcytosine (3mC) in both DNA and RNA (Westbye et al., 2008). The hALKBH1 was identified to mediate additional lyase activity of DNA at abasic sites in Fe²⁺- or 2-oxoglutarate-independent manner (Müller et al., 2010). Moreover, it has been reported that ALKBH1 regulates post-transcriptional gene expression through promoting methylation reversal of N1-methyladenosine (m¹A) in both cytoplasmic and mitochondrial tRNAs (Liu et al., 2016; Kawarada et al., 2017). Furthermore, mammalian ALKBH1 demethylates m⁵C derivative intermediates on the tRNAs as well in various cellular compartments (Kawarada et al., 2017). More recently, it was confirmed that ALKBH1 could also demethylate N6-methyladenine (m⁶A) on DNA, suggesting dual important epigenomic regulatory roles in DNA and epitranscriptomic roles on various forms of RNAs (Tian et al., 2020; Zhang et al., 2020). Although ALKBH-2 and–3 promote both m¹A and 3-methylcytidine (m³C) demethylation, ALKBH2 efficiently repairs both methylated single-stranded DNA (ssDNA) and double-stranded DNA (dsDNA), whereas ALKBH3 preferentially demethylates single-stranded nucleic acids (Monsen et al., 2010). Recently, ALKBH3 was found to post-transcriptionally regulate protein expression through the demethylation of m¹A on specific cellular transcripts (Woo and Chambers, 2019). Besides this role, ALKBH3 demethylates specific tRNA modifications, including m¹A and m³C, which ultimately promotes cancer progression (Chen Z. et al., 2019). Of all ALKBHs described so far, ALKBH-4 and–7 were found to demethylate preferentially proteins rather than nucleic acids (Li et al., 2013; Wang et al., 2014). Importantly, the widely studied eukaryotic AlkB homologs proteins including ALKBH5 and FTO were found to specifically

demethylate m⁶A, which is the most prevalent internal chemical modification on RNA for epigenetic control of cell life cycles (Jia et al., 2011; Zheng et al., 2013; Feng et al., 2014; Xu et al., 2014). Moreover, ALKBH8 was reported to mediate 5-methoxycarbonylmethyluridin repair through hydroxylation of tRNA (Fu et al., 2010). The detailed function of ALKBH6 has not yet been identified (Hu et al., 2019).

Herein, we aim to provide a comprehensive review of the recent progress made to uncover the structural features of the m⁶A demethylases compared with the rest of the AlkB protein members. Additionally, we aim to draw comparative features between ALKBH5 and FTO for their binding specificity, efficiency, and selectivity along with providing the recent updates of the various regulatory aspects of m⁶A erasers and the promising inhibitors to further guide the development of efficacious therapeutics to target cancers, metabolic disorders, and viruses.

Enzymatic Biochemistry of m⁶A Demethylases

The identification of different nucleobases that had been exposed to oxidative demethylation is deemed essential for understanding the intracellular biological and metabolic functions of the m⁶A-containing substrates. Confined mostly to the nucleus, ALKBH5 utilizes the m⁶A-containing ssRNA as the major substrate for demethylation *via* α -ketoglutaric-dependent oxidase activity (Aik et al., 2014; Feng et al., 2014). The ALKBH5 has also been reported to target the dimethylated adenosine (m²A) in the ribosomal RNA. The m²A is a non-canonical base present in ribosomal RNA as a normal component of the small subunit of the ribosome that assists in the common translation machinery (Ensfielder et al., 2018).

The Schofield group was the first to predict the earliest substrate for FTO, the 3-methylthymine (3mT), *via* bioinformatic analysis (Gerken et al., 2007). Consistent with the functional analysis that exhibited human and murine expressed FTOs repair the 3mT preferentially in ssDNA over dsDNA and favorably demethylate the 3-methyluracil (3mU) in ssRNA over ssDNA (Jia et al., 2008).

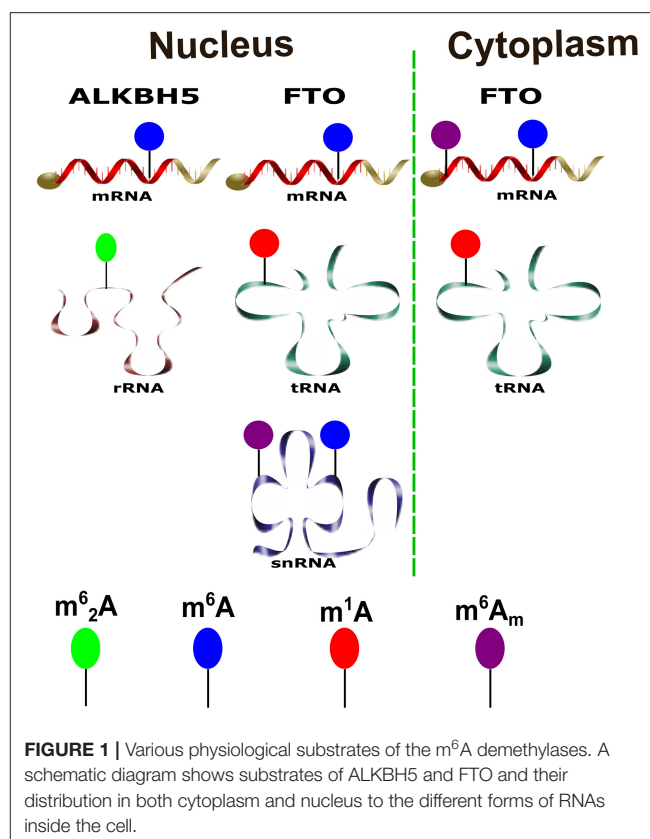
The m⁶A was confirmed to be catalyzed by FTO both *in vivo* and *in vitro* (Jia et al., 2011; Wei et al., 2018; Zhang X. et al., 2019). Furthermore, the +1 position to 5'cap in the polyadenylated RNA was confirmed to be di-methylated at N6 and 2'-O-position (m⁶A_m) as a major substrate for FTO that regulates the 5' mRNA integrity, stability, and resistance to decapping enzyme (e.g., DCP2) (Mauer et al., 2017). Intriguingly, the latter study claimed that the m⁶A_m is the sole physiological substrate for demethylation than m⁶A by FTO. This finding diametrically opposes most compelling evidence stating the relevant substrates of FTO (Jia et al., 2011; Fu et al., 2013; Wei et al., 2018; Zhang X. et al., 2019). It is worth noting that the hepatitis C virus (HCV), an ssRNA virus that belongs to the *Flaviviridae* family, was confirmed to harbor m⁶A marks throughout the entire viral RNA and respond to demethylation activity of FTO despite lacking the 5'cap (Gokhale et al., 2016). Additionally, recent investigations have identified that lacking the m⁶A_m

methyltransferase does not affect the cell growth kinetics and vital cellular processes (Akichika et al., 2019). In contrast, detrimental cellular alterations were observed in FTO knockdown cells (Zhao et al., 2014a; Li Z. et al., 2017). More recently, Sendinc et al. have illustrated that phosphorylated C-terminal domain (CTD) interacting factor-1 is an m⁶A_m methyltransferase and m⁶A_m is an evolutionarily conserved modification to the capped mRNAs. However, no crosstalk between the m⁶A and m⁶A_m was detected in the whole transcriptome mapping. Additionally, m⁶A_m promotes gene regulation mainly through mediating protein translation but not the transcription or mRNA stability (Sendinc et al., 2019). Interestingly, another report emphasizes the non-significant effect of phosphorylated CTD interacting factor-1 on protein translation (Boulas et al., 2019).

Systematically, Wei et al. have investigated the differential FTO substrate preference along with their location in various cell lines. The FTO preferentially mediates methylation reversal of the internal m⁶A in both the cytoplasm and nucleus on the polyadenylated RNAs. The percent of demethylation differs according to the investigated cell line. In contrast, FTO-mediated m⁶A_m-polyA RNA demethylation was confined to the cytoplasm (Wei et al., 2018). Moreover, the biochemical studies have identified additional RNA substrates to FTO in the various forms of RNA, including N1-methyladenosine (m¹A) in tRNA located in both nucleus and cytoplasm. It is important to note that m¹A-demethylated tRNAs have prominent action on translation efficiency (Liu et al., 2016; Wei et al., 2018). Moreover, it was confirmed that both m⁶A and cap m⁶A_m in small nuclear RNAs all found to be substrates for FTO that might control gene expression (Wei et al., 2018). Various physiological substrates for m⁶A-demethylases are summarized in **Figure 1**.

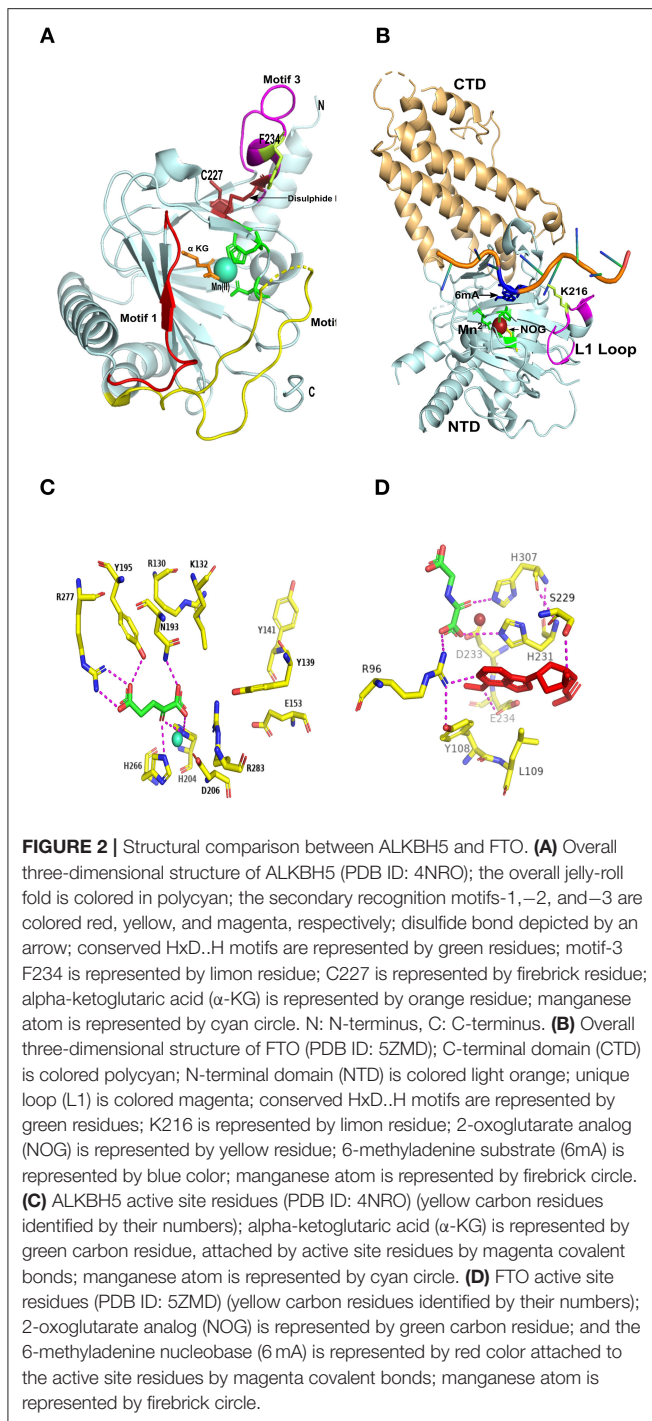
Structural Insights of the m⁶A Demethylases Determine Their Substrate Specificities

Our group has performed a recent comprehensive analysis of the m⁶A demethylases among various orders of animals, especially the avian species (Bayoumi et al., 2020). The study revealed multiple evolutionary changes when compared with *Homo sapiens*. We have revealed that m⁶A-erasers shared the lowest identity percent among the m⁶A-related machinery. However, the overall demethylases' structures were maintained through synonymous structural mutations (Bayoumi et al., 2020). The entire human AlkB-homolog-5 (hALKBH5) protein comprises a polypeptide chain of 394 amino acids (Zheng et al., 2013; Huang and Yin, 2018). Owing to technical challenges that have been experienced in the processing of the apo ALKBH5 enzyme in both *in vitro* enzymatic and crystallographic investigations, including those harboring different ligands, ALKBH5_{66–292} truncated fragment was active for functional and structural studies as well (Aik et al., 2014; Feng et al., 2014). The 65 N-terminus- and 103 C-terminus-residues were not essential for ALKBH5 core oxidative demethylation activity to targeted substrates. However, the C-terminus multiple serine residues were supposed to mediate phosphorylation (Aik et al., 2014).



From the earlier mentioned eukaryotic ALKBH protein family, all shared basic scaffold structure dubbed as jelly-roll [or double-stranded β-helix (DSβH)] fold, which is composed of conserved eight anti-parallel β-sheets in almost all species (Jia et al., 2011; Aik et al., 2012, 2014; Bayoumi et al., 2020). Besides the basic jelly-roll fold, additional secondary structures (nucleotide recognition motifs) were characterized in most ALKBH protein family members. It can be concluded that the basic scaffold has no substrate specificity function, whereas the secondary structures carry some level of specificity. Notably, no secondary structures were identified in both the ALKBH-4 and -7. Therefore, no oxidative methylation activity was detected toward nucleic acids and was only confined to the protein substrates (Li et al., 2013; Wang et al., 2014). From the substrate specificities mentioned earlier, it seems that adenosine (A) is the sole nucleobase to ALKBH5 in ssRNA (Aik et al., 2014; Feng et al., 2014; Xu et al., 2014).

Several groups have worked independently to illustrate the crystallographic analysis of the human ALKBH5 harboring various substrates and inhibitors (Aik et al., 2014; Feng et al., 2014; Xu et al., 2014). All of these groups have identified three unique amino acid motifs (**Figure 2A**). The motif 1, the position of this motif in relation to the active catalytic site, provides a widening surface compared with FTO and ALKBH2 (Feng et al., 2014), which proposes that the ALKBH5 can tolerate bulkier three-dimensional structure substrates for targeted oxidative demethylation (Aik et al., 2014). Additionally, motif 2 was



identified as a long motif that provides flexibility compared with other AlkB proteins (Feng et al., 2014). Notably, motif 3 has been confirmed to impede the double-stranded nucleic acid substrates that confirms ALKBH5 selectivity to an only single-stranded nucleic acid (Feng et al., 2014).

In addition to the conserved active site coordinated residues (HXD...H, motif), the basic residues adjacent to active sites (in motif 1) were also found to be crucial for enzymatic

activity, including K132 (Figure 2C). This was identified to interact with m⁶A and can also accept additional post-translational modifications (e.g., acetylation) that helps the enzymatic oxidative demethylation (Choudhary et al., 2009; Aik et al., 2014). The mutant K132A was identified to severely impair the ALKBH5 activity (Feng et al., 2014). Furthermore, ALKBH5 R130 residue, which was located in the unique motif 1 (Figures 2A,C), was supposed to interact directly with the single-stranded phosphate backbone (Aik et al., 2014). This interaction was confirmed by complete abrogation of the catalytic activity through targeted-mutational studies (Feng et al., 2014).

Likewise, within the long motif 2, unique amino acids were also identified to interact with m⁶A single-stranded substrate that confers substrate specificity, including Q146, K147, and R148 residues. Additionally, this is characterized by reduced demethylation activities (40%) upon their targeted mutations (Sundheim et al., 2006; Yang et al., 2008; Han et al., 2010; Feng et al., 2014). Most importantly, ALKBH5 motif 3 (Figure 2A) was implicated as the main secondary structure in the outer wall of DS β H; however, this motif is also present in other AlkB members (McDonough et al., 2010). The motif specifically flips in a way to impede with double-stranded substrates displaying steric hindrance by covalent disulfide bonding. This bond is conserved among various species of ALKBH5 between C230 and C267 or alternatively connects the C227 thiol group through redox shuffle mechanism generating C227-C267 linkage (Figure 2A). This mechanistically confers single stranded substrates selectivity. Moreover, F234 residue has been found to interact and direct the m⁶A-containing substrate toward the active catalytic site. The residues mentioned earlier were detected to be evolutionary conserved when tested by site-directed mutagenesis to ALKBH5 that specifically ensure strand specificity and secondary structure confirmation. Furthermore, the electrostatic map around the active catalytic site is important for the substrate binding. Mutational analysis found that more basic surfaces mainly to the active sites and the grooves made by the protrusion of the long motif 2 was pivotal for binding with the negative phosphate backbone form single-stranded substrates for optimal oxidative demethylation (Aik et al., 2014; Feng et al., 2014). Collectively, structural insights and the unique motifs and residues could be exploited to provide a better understanding of the substrate- and nucleotide-specificity for upcoming biomedical basic researches and development of ALKBH5 selective inhibitors.

Similar to the AlkB member family, FTO has the conserved jelly-roll motif (DS β H) harboring the active catalytic site in its N-terminal domain (NTD) (1-326). However, a novel fold designated as CTD (from 327 to 498 aa) has been structurally determined and is supposed to strengthen the NTD (Figure 2B). The publicly available crystal structure of FTO lacking the first 31 amino acids still retains the full enzymatic functionality indicating the active site buried in NTD and stabilized by CTD (Han et al., 2010). Likewise, the selectivity of ALKBH5 against the unmethylated strand of double-stranded nucleic acid, FTO, was also identified to harbor an evolutionary stretch of amino acid residues named long loop 1 (L1; residues from 210 to 223) (Figure 2B). We and others have confirmed that the L1 loop is

identified in *H. sapiens* and avian species and unidentified in the rest of AlkB members; this unique loop selectively blocks dsDNA/RNA to serve as a physiological substrate for FTO (Han et al., 2010; Feng et al., 2014; Zhang X. et al., 2019; Bayoumi et al., 2020).

Concerning the putative physiological substrates, it seems that the FTO outperforms the ALKBH5 in the number of physiological substrates to demethylate their methylated nucleobases. FTO promotes oxidative methylation to m⁶A and m⁶A_m in both mRNA and snRNA, and m¹A in tRNA. Furthermore, FTO demethylates 6mA, 3mT, and m³U (Han et al., 2010; Jia et al., 2011; Wei et al., 2018; Zhang X. et al., 2019). This array of substrates toward an AlkB member emphasizes the distinctiveness of the catalytic activity to accommodate various nucleobases. Besides the selectivity to hinder the double-stranded nucleic acids, the L1 loop has been investigated through biochemical and structural analysis to contribute to nucleobases recognition and stabilization of the single-stranded substrate in the FTO active site (Figure 2B; Zhang X. et al., 2019). Comprehensively, the L1 loop represented by K216 from one side and the short loop (residues 86–88) represented by K88 form hydrogen bonds with the phosphate backbones of the nucleotides adjacent to the methylated nucleobase. These lysine residues act as a pincer-like structure in twisting and accommodating the target nucleobase in the catalytic pocket (Figure 2B). Moreover, inside the catalytic pocket, the methylated base is stabilized by the hydrophobic interaction with the surrounding residues: I85, L109, Y108, V228, S229, W230, and H231. In contrast, the N6-methyl group specifically is stabilized in the pocket by the hydrophobic interaction with Y106, L203, and R322 residues (Figure 2D). Importantly, the methylated purine ring interacts with R96 and E234 predominantly by hydrogen bonding, whereas the ribose ring interacts mostly with A229 (Zhang X. et al., 2019). Therefore, the targeted mutations to these hydrogen bonding interacting residues abrogated the demethylation activity (Zhang X. et al., 2019). The same findings were also observed in other ALKBH homologs to residues corresponding to the R96. The site-directed mutation of M61 residue in AlkB and Q112 in ALKBH2 diminished their enzymatic functions (Han et al., 2010), suggesting highly conserved demethylation among various AlkB family members. Albeit, we reported the lowest identity percent of the avian FTO compared with the *H. sapiens*. A higher degree of conservation to the residues surrounding the methylated base in both *H. sapiens* and avian species was noticed, indicating a highly conserved catalytic mechanism even in various organisms exhibiting evolutionary changes. Moreover, we have found that the pincer-like structure in avian species suggests a higher binding affinity with more stabilizing property compared with *H. sapiens* (Bayoumi et al., 2020).

Considering the challenge of the similarity that could affect m⁶A antibody mismatching with m⁶A_m, high-throughput sequencing can differentiate them throughout the transcriptome (Linder et al., 2015). Compared with the m⁶A distribution across the mRNA, the m⁶A_m was documented to be located less frequently (Molinie et al., 2016). At least a 10-fold higher m⁶A level than that of the cap m⁶A_m in mRNA was

confirmed (Wei et al., 2018; Zhang X. et al., 2019). Because the same nucleobase (i.e., adenosine) between m⁶A and m⁶A_m were noticed, FTO superimposition studies exhibited the same oxidative demethylation activity in the same RNA sequence, with no significant effect to the ribose sugar on the enzymatic activity. However, the N6-methyl adenine group was confirmed to surpass other nucleobases to accommodate the active site of FTO, and 3meT was the lowest. Unequivocally, all mentioned substrates contain all the pivotal structural determinants for FTO physiological substrates to accommodate the active site. Moreover, the wide pincer-like structure formed by the unique loop one in FTO can accommodate higher numbers of substrates rather than ALKBH5 with bulkers secondary and tertiary structures such as the cap, stem-loop, and hairpin structures (Zhang X. et al., 2019).

Zou and co-workers have adopted detailed biophysical and biochemical analyses to determine the specificity of m⁶A demethylases in the nucleotide perspectives. They confirmed that both ALKBH5 and FTO do not exhibit strict sequence requirements for their substrates as other m⁶A-recognizing proteins; writers and readers do. Moreover, m⁶A demethylases can recognize and differentiate m⁶A marks in the highly similar nucleotide sequences, even having the same consensus motif, with superiority to the FTO. Notably, erasers can induce different outcomes in different RNA sequences, with different secondary structure conformation (duplex to hairpin transition), concluding that m⁶A itself is considered as a conformational marker (Zou et al., 2016).

Biological Functions of the m⁶A Demethylases

The m⁶A demethylases (ALKBH5 and FTO) modulate various aspects of cell life cycles that can diverge from the regulation of normal metabolic and differentiation functions, which aggravates numerous pathological conditions. In the past few decades, multiple tumor processes were documented across the literature with poor underlying molecular genetic justifications. After that, the field of epigenetics has become a relevant topic to provide possible explanations for several human diseases (Pinello et al., 2018; Chen X. Y. et al., 2019; Huang et al., 2020; Melstrom and Chen, 2020; Zhao et al., 2020).

Pathological Regulatory Aspects of m⁶A Demethylases

Epigenetically, the m⁶A demethylases dictate the fate of various cancerous conditions. In the thoracic cancers, Forristal et al. have investigated the effects of reduced O₂ tension (5%) on the upregulation of certain hypoxia-inducible factors (HIFs) in comparison with human embryonic stem cell control maintained in normoxic condition (20%). They have noticed the translocation of HIFs to the nucleus to reduce O₂ tension condition (Forristal et al., 2010). The translocated HIF-1α protein transcriptionally activates multiple targets as a cellular response to the hypoxia, chief among them was the human ALKBH5 gene (Thalhammer et al., 2011). The ALKBH5 plays an important role in controlling breast cancer progression through the HIF-ALKBH5-dependent pathway. ALKBH5 demethylates m⁶A

marks from NANOG, a master pluripotency factor; the oxidative demethylation activity of the ALKBH5 increases the NANOG transcript and protein expression that enriches breast cancer stem cells in the reduced oxygen tumor microenvironment promoting cancer progression (Zhang et al., 2016a). After that, Zhang et al. have also reported that knockdown of ALKBH5 from breast cancer cells could suppress breast-to-lung metastasis in mice model (Zhang et al., 2016b). Furthermore, FTO contributes to breast cancer development. It has been found that FTO overexpression was associated with a higher incidence of human breast cancer. FTO m⁶A-mediated demethylation of 3'-untranslated region BNIP3 transcript, which is a proapoptotic protein belonging to the Bcl-2 tumor suppressor family, promoting its degradation *via* YTHDF2 independent pathways and specific upregulation of BNIP3 retards breast cancer proliferation and metastasis (Niu et al., 2019). Collectively, it seems that thoracic cancer progression is controlled negatively by specific mRNA methylation reversal (Deng et al., 2018a,c; Mauer and Jaffrey, 2018; Pinello et al., 2018; Rajecka et al., 2019; Melstrom and Chen, 2020). More recently, the elevation of the ALKBH5 level was also confirmed to be involved in lung adenocarcinoma proliferation and invasion under intermittent hypoxia conditions. ALKBH5 demethylates Forkhead box M1 (FOXM1), which is one of the main tumor inducers. Upon m⁶A demethylation, the FOXM1 transcript provides stabilization of the expressed protein (Chao et al., 2020). The ALKBH5 has also been demonstrated to regulate the tumorigenic progression of oral squamous cell carcinoma that antagonizes the utilized chemotherapeutics for the intervention of proliferation and metastasis (Shriwas et al., 2020). The ALKBH5-dependent demethylation of FOXM1 and NANOG transcripts (main regulators of cancer stem cells) promotes chemoresistance of platinum-based drugs through negative regulation of human DEAD-box RNA helicase (DDX3), which are primarily involved in the innate immunity, multiple cell signal processes, and numerous aspects of RNA metabolism (Shriwas et al., 2020).

Despite ALKBH5 has been identified to contribute significantly to physiological osteogenesis (Yu et al., 2020), ALKBH5 mediates osteosarcoma (OS) tumorigenesis *via* demethylation of plasmacytoma variant translocation one, a tumorigenic lncRNA. Mechanistically, ALKBH5 removes the m⁶A marks, increases the stability of mRNA, and enhances the expression of plasmacytoma variant translocation one through inhibiting its YTHDF2 binding, resulting in increased OS cell proliferation rates both *in vitro* and *in vivo* (Int et al., 2020). Similar to the OS tumorigenesis, ALKBH5 possesses a negative regulatory impact in gastric cancer (GC) *via* acting on another lncRNA named nuclear paraspeckle assembly transcript one that results in enhancement of EZH2 expression (a component of the polycomb repressive complex) and ultimately affects the invasion and metastasis in GC tissues (Zhang J. et al., 2019; Zhu et al., 2020). The same fate was identified in FTO overexpression in GC cancer tissues compared with adjacent non-tumorous tissue (Xu et al., 2017; Zhang C. et al., 2019). Taken together, it seems that m⁶A erasers demethylate both mRNA and lncRNA to promote carcinogenesis and have a negative oncogenic signature in multiple cancers.

Likewise, m⁶A demethylases modulate sex-specific tumors. Marked expression of ALKBH5 has been detected in ovarian cancer, which mediates the EGFR-PIK3CA-AKT-mTOR-signaling pathway, a key regulatory pathway in autophagy-induced stress response and nutrient deprivation. Additionally, ALKBH5 enhances the stability of the BCL-2 transcript (which increased in the epithelial ovarian cancer as well) and enhances the interaction between BCL-2 and Beclin1 that inhibit the autophagy from the other side, suggesting that the ALKBH5 controls tumor progression and autophagy flux *via* BCL-2 demethylation (Zhu et al., 2019). In contrast, in males, the ALKBH5 was found to control testicular germ cell tumors type II (Nettersheim et al., 2019).

Not only that soft tissue tumors are controlled epigenetically, but FTO has also been incriminated in the progression of the solid tumor, including melanoma. Two mechanisms were proposed, through single-nucleotide polymorphisms outside of intron one (body mass index-related region), as rs16953002, the variant of intron 8 of FTO that has been reported to be associated with a high risk of melanoma (Iles et al., 2013; Deng et al., 2018b). Additionally, the FTO was identified as a pro-tumorigenic factor in melanoma. The FTO negatively regulates the response to anti-programmed death 1, an immunotherapeutic agent, through the action of melanoma-intrinsic genes including PD-1, C-X-CR-4, and SOX10; those are the major potential gene targets for demethylation by FTO (Yang S. et al., 2019; Melstrom and Chen, 2020; Zhao et al., 2020).

Great focus has been dedicated to deciphering the oncogenic role of FTO in hematopoietic disorders. These include acute myeloid leukemia through promoting leukemogenesis *via* FTO-mediated m⁶A demethylation of core transcripts as ASB2 and RARA mRNAs promoting decreased stability of the target transcripts (Li Z. et al., 2017; Huang et al., 2019; Weng et al., 2019; Zhao et al., 2020). Additionally, ALKBH5 was found to be linked with the devastating malignant brain tumor glioblastoma through the ALKBH5-FOXM1-mediated pathway; in this milieu, ALKBH5 enhances glioblastoma tumorigenesis (Dixit et al., 2017; Zhang et al., 2017; Malacrida et al., 2020).

Unlike the fate of the cancers mentioned earlier, the m⁶A demethylases alleviate the outcome of additional biological processes. ALKBH5 expression was noticed to be downregulated in pancreatic tumors. ALKBH5 targets a lncRNA named KCN15-AS1 *via* direct demethylation and is associated with inhibition of the pancreatic cancer metastasis, which might serve as a potential therapeutic target for pancreatic cancer patients (He et al., 2018). More recently, mechanistic investigations have documented another ALKBH5-mediated inhibition of the most common form of pancreatic cancers, the pancreatic ductal adenocarcinoma, through the ALKBH5 dependent-Wnt inhibitory factor one pathway (Tang et al., 2020). To conclude, ALKBH5 carries suppressive effects on certain tumors to provide mounting evidence to be an excellent new prognostic marker for pancreatic cancers (Cho et al., 2018; Melstrom and Chen, 2020).

Similar findings were noticed with bladder cancer repression through the action of the ALKBH5 and METTL3 in a reciprocal manner on integrin alpha-6 transcript, which enhances various cellular motility and signaling events. The ALKBH5 inhibits the

translation of integrin alpha-6 in the m⁶A-dependent pathway and decreases bladder cancer adhesion, migration, and invasion (Jin et al., 2019). Moreover, colon cancer was suppressed upon overexpression of the ALKBH5 in both cell invasion *in vitro* and metastasis *in vivo* (Yang P. et al., 2019). Thus, ambitious therapeutic candidates have also been proposed in head and neck squamous cell carcinoma *via* overexpression of ALKBH5 and FTO (Pilžys et al., 2019). To conclude, various actions of m⁶A demethylases were noticed to either suppress or enhance cancer development and progression through direct oxidative demethylation on either specific mRNAs or lncRNAs. Additionally, accumulating evidence suggests using m⁶A demethylases or their gene targets for either prognostic and diagnostic markers for specific tumors as indicated earlier, and improving specific inhibitors for future use could open a new frontier in alleviating multiple cancerous conditions.

Metabolic and Physiological Regulatory Roles of m⁶A Demethylases

It is well-documented that m⁶A-containing mRNA regulates various biological processes, including autophagy, which is an evolutionarily conserved degradation pathway in the cell. A critical association between the autophagy from one side and METTL3-ALKBH5 interplay from the other side has been found to control hypoxia/reoxygenation-treated cardiomyocytes (*in vitro* and in an animal model) in which the ALKBH5 acted as a positive regulator in the autophagy *via* regulating m⁶A level on the transcription factor EB mRNA and its subsequent protein expression. Transcription factor EB is the main regulator of autophagy-related genes and ultimately regulates the fate of ischemic heart diseases (Song et al., 2019).

Additionally, the obesity problem in humans has been linked to the FTO. Albeit, obesity is concomitant to various inherited and behavioral determinants that further predisposes to other chronic diseases; the FTO is also incriminated in adipogenesis. FTO single-nucleotide polymorphisms, which are mostly located in intron-1, were found to be linked with obesity in humans (Zhao et al., 2014b). There are multiple proposed mechanistic regulatory roles of FTO in the development and progression of obesity (Gulati et al., 2013). In contrast, others suggested that the FTO gene is under the control of nearby associated genes, chief among them IIRX3 to be the main regulator in obesity (Smemo et al., 2014). However, the obesity–FTO associations are reviewed well elsewhere (Zhou et al., 2017; Deng et al., 2018b; Mauer and Jaffrey, 2018).

Vis-à-vis eraser's physiological roles, the ALKBH5 has been found to play a pivotal role in the regulation of the enrichment of the human placenta during pregnancy *via* the action on trophoblasts that seems to affect the recurrent miscarriage patients. Mechanistic studies have revealed that ALKBH5 mediates the action by affecting the half-life of cysteine-rich angiogenic inducer-61 mRNA that possesses differentiation, migration, and adhesion roles, which are important for normal embryogenesis (Li et al., 2019). Furthermore, FTO was found to be involved in premature ovarian insufficiency-mediated infertility. The reduction of FTO protein expression was concomitant with elevated m⁶A level in ovarian tissue of

premature ovarian insufficiency patients (Ding et al., 2018). A similar finding reported in male mice has a deficiency in ALKBH5. Those mice were identified to have increased levels of m⁶A in their transcripts, consequently impaired fertility and apoptosis along with the ultimate negative effect on the meiotic metaphase stage of the spermatocytes (Zheng et al., 2013). Tang et al. (2017) have unveiled the mechanistic insights of ALKBH5-mediated m⁶A's role in male infertility and revealed that ALKBH5 ensured the production of longer 3'-untranslated region transcripts coupled with correct splicing (Tang et al., 2017). Regarding differentiation functions of demethylases, the ALKBH5 regulates multiple metabolic processes as adipogenesis and myogenesis through modulating the early differentiation markers such as CEBPb and myogenin, respectively (Choi et al., 2019). The FTO was also found to play roles in differentiating the neuronal stem cells in adult mice (Cao et al., 2019). The various pathological and physiological regulatory roles of m⁶A-demethylases are summarized in Table 1.

Viral Regulatory Aspects of m⁶A Demethylases

Similar to cellular transcripts, viral RNA can accept the decoration by m⁶A to regulate/dictate the viral life cycle and outcome of virus–host interactions (Dang et al., 2019). These include multiple viruses of medical importance as well as oncogenic viruses. The m⁶A-demethylase-mediated modification could control viral replication, pathogenesis, infection, and ultimately tumor formation (Imam et al., 2018; Tan et al., 2018; Tsai et al., 2018; Lang et al., 2019). The hepatitis B virus (HBV) is a DNA tumor-causing virus linked with chronic hepatitis, a high risk of liver cirrhosis, and hepatocellular carcinoma (Shepard et al., 2006). HBV intermediate transcripts have been confirmed to bear m⁶A marks from both hepatic tissues of chronic HBV patients and HBV-expressing cells (Imam et al., 2018). Furthermore, m⁶A machinery represented by METTL3, METTL14 from one side, and FTO from the other side mediates two major regulatory functions. Firstly, the viral gene expression and secondly the reverse transcription based on the m⁶A modified site on the epsilon loop of HBV that modulate the fate of HBV in the liver disease pathogenesis and tumor formation (Imam et al., 2018).

Moreover, Kaposi's sarcoma-associated herpesvirus (KSHV) is another salient example of a human oncogenic virus linked with different cancers, including Kaposi's sarcoma and primary effusion lymphoma; KSHV has latent and lytic replication stages in the lifecycle (Ye et al., 2011). Recent advances in epitranscriptome sequencing revealed that m⁶A could modulate the transition between the stages with altered m⁶A methylome, and erasers modulate the lytic gene expression that controls KSHV infection and KSHV-induced oncogenesis. Recent studies have reported that m⁶A modifications play different roles owing to various cell types during lytic replication of KSHV (Ye et al., 2017; Hesser et al., 2018; Tan et al., 2018).

Additionally, Epstein–Barr virus is another example of oncogenic herpes viruses caused by human herpesvirus-4, which is incriminated with 2% of human cancers. Through the interplay of METTL14, YTHDF2, and ALKBH5, Epstein–Barr virus latent protein EBNA3C is responsible for reprogramming the

TABLE 1 | Regulatory aspects of m⁶A demethylases.

m ⁶ A demethylase	Regulatory aspect	Tissue involved	Regulatory Gene(s) & their expression level	References
ALKBH5	Cancer type	Breast cancer	↑NANOG	Zhang et al., 2016a,b
		Glioblastoma	↑FOXM1	Dixit et al., 2017; Zhang et al., 2017; Malacrida et al., 2020
		Lung adenocarcinoma	↑FOXM1	Chao et al., 2020
		Pancreatic cancer	↑KCNK15-AS1	He et al., 2018
			↑WIF- 1	Tang et al., 2020
		Bladder cancer	↓ITGA6	Jin et al., 2019
		Oral squamous cell carcinoma	↑FOXM1 / NANOG	Shriwas et al., 2020
		Osteosarcoma	↑PVT1	Int et al., 2020
		Gastric cancer	↑NEAT1	Zhang J. et al., 2019; Zhu et al., 2020
		Colon cancer		Yang P. et al., 2019
	Metabolic disorder	Ovarian cancer	↑Bcl2	Zhu et al., 2019
		Male germ cell tumor		Nettersheim et al., 2019
		Male infertility	↑Correct spliced/longer transcripts	Tang et al., 2017
	Differentiation	Autophagy (ischemic heart disease)	↑TFEB	Song et al., 2019
		Placenta	↓CYR61	Li et al., 2019
		Adipogenesis	↓CEBPb	Choi et al., 2019
		Myogenesis	↓Myogenin	Choi et al., 2019
FTO	Cancer type	Breast cancer	↓BNIP3	Niu et al., 2019
		Melanoma	↑PD-1 CXCR4 SOX10	Yang S. et al., 2019; Melstrom and Chen, 2020; Zhao et al., 2020
		Acute myeloid leukemia	↓ASB2 and RARA	Li Z. et al., 2017; Huang et al., 2019; Weng et al., 2019; Zhao et al., 2020
	Metabolic disorder	Gastric cancer		Xu et al., 2017; Zhang C. et al., 2019
		Obesity	FTO gene Intron1 IRX3	Zhao et al., 2014b; Smemo et al., 2014
		Premature ovarian insufficiency		Ding et al., 2018
	Differentiation	Neuronal stem cells		Cao et al., 2019

methyloyme that enhances tumorigenesis via the m⁶A-dependent pathway (Lang et al., 2019). Similarly, the Simian virus 40, a DNA oncogenic virus, and HCV, a major RNA tumor-causing virus, are impacted positively or negatively through the m⁶A-dependant pathways, respectively (Gokhale et al., 2016; Tsai et al., 2018). However, mechanistic action of FTO or ALKBH5 for tumor formation remains to be identified.

Additionally, the non-oncogenic viruses are m⁶A decorated as well, and the m⁶A demethylases have an intriguing role in different virus life cycles. During virus infection, the ALKBH5 only induces a regulatory role in virus replication and protein expression as reported previously in human immunodeficiency virus-1 (HIV-1) and vesicular stomatitis virus (VSV) (Lichinchi et al., 2016a; Tirumuru et al., 2016; Liu et al., 2019). In contrast, in others, the FTO only modulates viral infection, including HCV (Gokhale et al., 2016) and enterovirus-71 (Hao et al., 2019). However, in the case of the Zika virus and respiratory syncytial virus, both demethylases have regulatory functions (Lichinchi et al., 2016b; Xue et al., 2019). Cumulatively, it

is plausible that m⁶A demethylases display various regulatory functions in different cell contexts (even those infected with the same virus model), likely through regulating distinct sets of targets, suggesting more detailed analysis for the near future and for designing the correct specific inhibitor. Additionally, future systematic studies will determine the biological function of each of the m⁶A regulatory genes in various cancer settings and the critical target genes to unveil the underlying molecular mechanisms.

m⁶A Demethylases' Inhibitors

Unraveling m⁶A demethylases structures along with a better understanding of their physiological and tumorigenic regulatory roles inspired various groups to develop different types of inhibitors to impede the enzymatic activity. Modulating the m⁶A level inside cells is an ambitious target to control various cancerous condition invasion and metastasis as discussed earlier. Therefore, inhibition of the prototype *E. coli* AlkB was the proof-of-concept to this notion using a natural product named

quercetin (Chen et al., 2012). Importantly, with the availability of the FTO crystallographic structure (Han et al., 2010), a comprehensive mechanistic study to utilize cell-active, natural products (rhein) was confirmed to reversibly bind to the nucleotide-binding pocket *in vitro* and inside cells with reduced cell toxicity. Structurally, the positively charged active site (R316) of FTO was found to interact with the negatively charged carboxyl group of the rhein to hinder m⁶A repair (Figures 3B,D; Chen et al., 2012). Additional wide arrays of FTO small-inhibitor molecules were suggested to abolish FTO catalytic activity *via* either interacting with the nucleotide-binding and/or 2OG binding sites (Figure 3; Aik et al., 2013).

After that, fluorescence polarization studies with chemical displacement have been utilized to validate the use of meclofenamic acid (MA), an anti-inflammatory drug, to provide temporal intervention of mRNA methylation. The MA competes with the m⁶A-binding site (Figures 3A,C) and inhibits FTO over ALKBH5 (Huang et al., 2015). It is worth noticing that MA was reported to be successfully used for inhibition of FTO demethylation activity in the KSHV lifecycle and has been confirmed to enhance the lytic gene expression in comparable results with FTO loss-of-function experiments (Ye et al., 2017). Despite the potent activity of the rhein and MA, inhibiting other essential cellular enzymes were noticed to shut down their activities (Chen et al., 2012; Flanagan et al., 2012; Li et al., 2016; Huang et al., 2019). Wang et al. have utilized the structural similarity between some fluorescein compounds to MA to selectively inhibit FTO activity and provide additional labeling simultaneously (Wang et al., 2015).

Additionally, MO-I-500, a pharmacologically tested FTO inhibitor, was also reported reducing the survival rate of inflammatory breast cancer cell lines selectively (Singh et al., 2016). Moreover, a robust tool has been recognized depending on the difference in both substrate and nucleotide specificities, which provides compounds that occupy both nucleotide and 2OG binding pockets. This method is named the two-component inhibitor tethering strategy (Toh et al., 2015).

Rational drug design through the scaffold hopping approach was also adopted to identify new candidates for FTO inhibitors. These candidates were tested using docking simulations. Structural analysis of MU06-bounded-FTO revealed interaction of R96 and H231 of FTO catalytic pocket with MU06 inhibitor via H bonding (Padariya and Kalathiya, 2016). Recently, fluorescent RNA aptamers were utilized as a tool for studying FTO inhibitors in a high-throughput screening format (Svensen and Jaffrey, 2016). Additional natural compounds were identified as putative FTO inhibitors such as radicicol (Wang et al., 2018) and clausine E (Wang et al., 2019). Other compounds have additional medicinal advantages, such as the anti-epileptic effect (Zheng et al., 2014) and the anti-leukemic activity of the R-2HG (Su et al., 2018).

More recently, promising FTO inhibitors such as FB23 and FB23-2 were selected and tested in an animal model. It was found to impede FTO in a way mimicking FTO depletion in acute myeloid leukemia cell lines. Structurally,

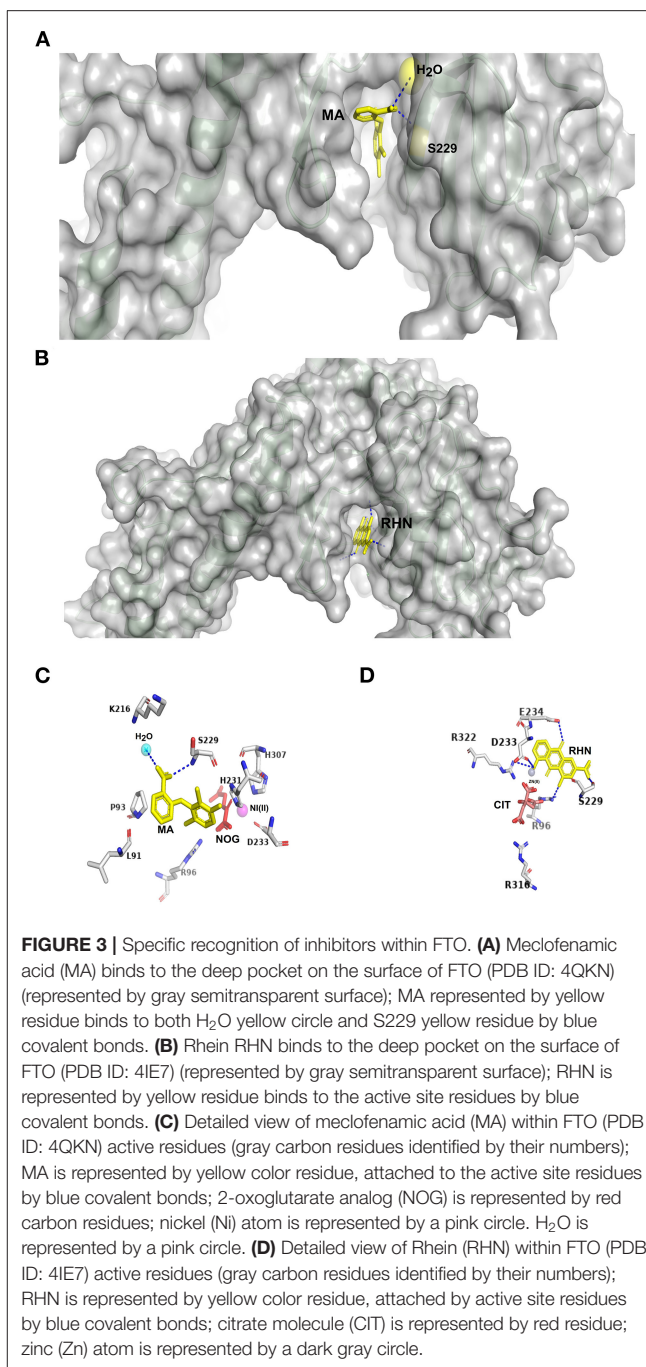


FIGURE 3 | Specific recognition of inhibitors within FTO. **(A)** Meclofenamic acid (MA) binds to the deep pocket on the surface of FTO (PDB ID: 4QKN) (represented by gray semitransparent surface); MA represented by yellow residue binds to both H₂O yellow circle and S229 yellow residue by blue covalent bonds. **(B)** Rhein RHN binds to the deep pocket on the surface of FTO (PDB ID: 4IE7) (represented by gray semitransparent surface); RHN is represented by yellow residue binds to the active site residues by blue covalent bonds. **(C)** Detailed view of meclofenamic acid (MA) within FTO (PDB ID: 4QKN) active residues (gray carbon residues identified by their numbers); MA is represented by yellow color residue, attached to the active site residues by blue covalent bonds; 2-oxoglutarate analog (NOG) is represented by red carbon residues; nickel (Ni) atom is represented by a pink circle. H₂O is represented by a pink circle. **(D)** Detailed view of Rhein (RHN) within FTO (PDB ID: 4IE7) active residues (gray carbon residues identified by their numbers); RHN is represented by yellow color residue, attached by active site residues by blue covalent bonds; citrate molecule (CIT) is represented by red residue; zinc (Zn) atom is represented by a dark gray circle.

these inhibitors have complementarity with the substrate-binding pocket *via* binding with the critical residues in the active site, including S229, R96, and E234 (Huang et al., 2019; Figures 3C,D).

Although most of the compound, as mentioned earlier, can totally or partially inhibit FTO, the MV1035, an imidazobenzoxazin-5-thione, was initially synthesized as a Na⁺ channel blocker, using structural-based *in silico* screening

in the wide-scale proteome. MV1035 was found to interact with ALKBH5 as an off-target molecule. After that, the functional analysis was confirmed to fight the glioblastoma aggressiveness (Malacrida et al., 2020).

CONCLUDING REMARKS: HOW VIRUSES CAN PROVIDE MORE INFORMATION FOR A BETTER UNDERSTANDING OF EPIGENETICS IN THE FUTURE

Methylation of viral RNA has recently been considered as a hallmark in virus–host interactions. Viral epitranscriptome allows us to underpin molecular mechanisms of m⁶A modification and its impact on cellular and viral RNAs behaviors. It has been concluded that the activity of the writers and the readers were associated with restraining the viral replication capacity. In contrast, the demethylases exert an opposite effect in virus-infected cells, suggesting an overall negative regulation of viral replication (Brocard et al., 2017).

Furthermore, m⁶A is proposed to negatively regulate interferon (IFN) responses in virus–host interaction. Significant reduction in various dsDNA viruses (including human cytomegalovirus, HCMV) titers was reported in m⁶A-writers and readers knockout (KO) cells, and marked elevation to viral titers were noticed in FTO- or ALKBH5-KO cells along with the fast turnover of IFN transcripts, hence facilitating viral propagation (Winkler et al., 2019). Mechanistically, cellular RNA helicase (DDX46) inhibits the innate immune response through the DDX46-ALKBH5-dependant pathway, leading to the demethylation of IFN transcripts. Demethylation of these mRNAs enforces their retention in the nucleus and inhibits IFN production and consequently enhances viral propagation (Zheng et al., 2017).

However, this is not the case for all viruses investigated so far. Interestingly, m⁶A has a positive regulatory outcome of certain viruses such as HIV-1. The depletion of the METTL3 and/or METTL14 (writers) has been confirmed to inhibit HIV-1 replication, whereas knockdown of the ALKBH5 enhances the replication (Kennedy et al., 2016, 2017; Tirumuru et al., 2016). The same findings were consistent in enterovirus-71, which is another ssRNA virus. In the enterovirus-71 replication model, the viral RNA copy number and protein expression were regulated mainly by the FTO. Intriguingly, the ALKBH5 does not affect the virus lifecycle (Hao et al., 2019). Moreover, the enhanced viral gene expression and replication have also been reported in the influenza A virus (Courtney et al., 2017) and SV-40 (Tsai et al., 2018). However, the m⁶A demethylases' roles in viral replication have not been investigated in greater detail.

In contrast, the negative impact of m⁶A demethylases was noticed in the HCV (Gokhale et al., 2016), Zika (Lichinchi et al., 2016b), and VSV (Liu et al., 2019). It is important to note that depletion of FTO was concomitant with a reduced infectious virus and HCV RNA release. Interestingly, ALKBH5

does not influence the HCV life cycle (Gokhale et al., 2016). ALKBH5 and FTO enhance the titer and the protein synthesis of the Zika virus, which is another member of the *Flaviviridae* family (Lichinchi et al., 2016b; Tan and Gao, 2018). Notably, it was also confirmed that knockdown of ALKBH5 significantly reduced VSV RNA levels (Liu et al., 2019). Despite intensive studies of epitranscriptome to cellular RNA, the molecular events illustrating virus–cell epitranscriptome interactions are in their infancy, and several fundamental questions need to be answered. Especially, m⁶A demethylases, as the focal point of this review, must understand differences between FTO and ALKBH5 in their pattern of recognition to closely related viral RNA.

Considering the m⁶A mark as a conformational marker, the sequence variation and secondary and tertiary structures between two viruses, which belong to the same family, could be the cause of preferential target to one demethylase compared with another. However, detailed structural and functional studies warrant further investigation that might reveal aspects in understanding the mechanistic action toward viruses to provide efficacious antivirals in the near future.

Moreover, detailed studies of all m⁶A-related proteins (writers, readers, and erasers) could explain the various outcomes against different viruses. This was not surprising, as loss-of-function studies of m⁶A-demethylases have different outcomes in various cancerous conditions, as discussed earlier. Additionally, certain viruses were found to accept the variant of m⁶A modification (i.e., m⁶A_m), which can affect the fate of virus replication (Lichinchi et al., 2016b; Tirumuru et al., 2016; Tan et al., 2018). Interestingly, FTO is the unique demethylase that interacts and responds to m⁶A_m. However, detailed crystallographic analysis of FTO harboring m⁶A and/or m⁶A_m could provide more answers in both cellular and viral epitranscriptomic field.

Considering the splicing function regulated by both demethylases (Zheng et al., 2013; Bartosovic et al., 2017), their role in the splicing process of viruses (DNA viruses, HIV, and influenza A virus) warrants further investigation. Besides, the discrepancies in the various reports in cellular and/or viral epitranscriptome might be owing to variation cell type, site of modifications, the utilized technique for sequencing (Tan et al., 2018; Dang et al., 2019). However, the viral epigenetic is at the stage of infancy and warrants exhaustive research in the near future.

Given the roles of m⁶A-demethylases in multiple virus life cycles and tumorigenic capacity shed light on the future potential use of inhibitors to fight a wide range of biological process simultaneously. The proof-of-concept has been provided from the data described in multiple studies. The use of various FTO inhibitors such as broad-spectrum m⁶A inhibitor named 3-deazaadenosine (DAA) *in vitro* and animal model (Kennedy et al., 2016; Courtney et al., 2017) and the specific FTO inhibitor (MA) in KSHV model (Ye et al., 2017) provide valuable insights. These proof of principle studies underline the applicability of m⁶A “demethylases”

inhibitors in developing next-generation antiviral and cancer therapies.

AUTHOR CONTRIBUTIONS

MM: conceptualization, supervision, writing—review, and editing. MB and MM: formal analysis, investigation, and resources. MB: writing—original draft preparation. All authors contributed to the article and approved the submitted version.

REFERENCES

- Aik, W., Demetriades, M., Hamdan, M. K. K., Bagg, E. A. L., Yeoh, K. K., Lejeune, C., et al. (2013). Structural basis for inhibition of the fat mass and obesity associated protein (FTO). *J. Med. Chem.* 56, 3680–3688. doi: 10.1021/jm400193d
- Aik, W., Scotti, J. S., Choi, H., Gong, L., Demetriades, M., Schofield, C. J., et al. (2014). Structure of human RNA N6-methyladenine demethylase ALKBH5 provides insights into its mechanisms of nucleic acid recognition and demethylation. *Nucleic Acids Res.* 42, 4741–4754. doi: 10.1093/nar/gku085
- Aik, W. S., McDonough, M. A., Thallhammer, A., Chowdhury, R., and Schofield, C. J. (2012). Role of the jelly-roll fold in substrate binding by 2-oxoglutarate oxygenases. *Curr. Opin. Struct. Biol.* 22, 691–700. doi: 10.1016/j.sbi.2012.10.001
- Akchika, S., Hirano, S., Shichino, Y., Suzuki, T., Nishimasu, H., Ishitani, R., et al. (2019). Cap-specific terminal N 6 -methylation of RNA by an RNA polymerase II-associated methyltransferase. *Science* 363, 1–13. doi: 10.1126/science.aav0080
- Alemu, E. A., He, C., and Klungland, A. (2016). ALKBHs-facilitated RNA modifications and de-modifications. *DNA Repair.* 44, 87–91. doi: 10.1016/j.dnarep.2016.05.026
- Bartosovic, M., Molares, H. C., Gregorova, P., Hrossova, D., Kudla, G., and Vanacova, S. (2017). N6-methyladenosine demethylase FTO targets pre-mRNAs and regulates alternative splicing and 3'-end processing. *Nucleic Acids Res.* 45, 11356–11370. doi: 10.1093/nar/gkx778
- Bayoumi, M., Rohaim, M. A., and Munir, M. (2020). Structural and virus regulatory insights into avian N6-methyladenosine (m6A) machinery. *Front. Cell Dev. Biol.* 8:543. doi: 10.3389/fcell.2020.00543
- Boccaletto, P., MacHnicka, M. A., Purta, E., Pitkowski, P., Baginski, B., Wirecki, T. K., et al. (2018). MODOMICS: a database of RNA modification pathways. 2017 update. *Nucleic Acids Res.* 46, 303–307. doi: 10.1093/nar/gkx1030
- Boulias, K., Toczylowska-Socha, D., Hawley, B. R., Liberman, N., Takashima, K., Zaccara, S., et al. (2019). Identification of the m6Am methyltransferase PCIF1 reveals the location and functions of m6Am in the transcriptome. *Mol. Cell* 75, 631–643.e8. doi: 10.1016/j.molcel.2019.06.006
- Brocard, M., Ruggieri, A., and Locker, N. (2017). m6A RNA methylation, a new hallmark in virus-host interactions. *J. Gen. Virol.* 98, 2207–2214. doi: 10.1099/jgv.0.000910
- Cao, Y., Zhuang, Y., Chen, J., Xu, W., Shou, Y., Huang, X., et al. (2019). Dynamic effects of Fto in regulating the proliferation and differentiation of adult neural stem cells of mice. *Hum. Mol. Genet.* 29, 727–735. doi: 10.1093/hmg/ddz274
- Carlile, T. M., Rojas-Duran, M. F., Zinshteyn, B., Shin, H., Bartoli, K. M., and Gilbert, W. V. (2014). Pseudouridine profiling reveals regulated mRNA pseudouridylation in yeast and human cells. *Nature* 515, 143–146. doi: 10.1038/nature13802
- Chao, Y., Shang, J., and Ji, W. (2020). ALKBH5-m6A-FOXO1 signaling axis promotes proliferation and invasion of lung adenocarcinoma cells under intermittent hypoxia. *Biochem. Biophys. Res. Commun.* 521, 499–506. doi: 10.1016/j.bbrc.2019.10.145
- Chen, B., Ye, F., Yu, L., Jia, G., Huang, X., and Zhang, X. (2012). Development of cell-active N 6 -methyladenosine RNA demethylase FTO inhibitor. *J. Am. Chem. Soc.* 134, 17963–17971. doi: 10.1021/ja3064149
- Chen, X. Y., Zhang, J., and Zhu, J. S. (2019). The role of m6A RNA methylation in human cancer. *Mol. Cancer* 18, 1–9. doi: 10.1186/s12943-019-1033-z

FUNDING

This study was funded by the Biotechnology and Biological Sciences Research Council (BB/M008681/1 and BBS/E/I/00001852) and the British Council (172710323 and 332228521). The Ph.D. studies of MB has been financially supported by Newton Mosharafa-Fund (Bureau ID: NMM12/19) and the Egyptian Ministry of Higher Education and Scientific Research, Cultural Affairs and Mission Sector, Egypt.

- Chen, Z., Qi, M., Shen, B., Luo, G., Wu, Y., Li, J., et al. (2019). Transfer RNA demethylase ALKBH3 promotes cancer progression via induction of tRNA-derived small RNAs. *Nucleic Acids Res.* 47, 2533–2545. doi: 10.1093/nar/gky1250
- Cho, S. H., Ha, M., Cho, Y. H., Ryu, J. H., Yang, K., Lee, K. H., et al. (2018). ALKBH5 gene is a novel biomarker that predicts the prognosis of pancreatic cancer: a retrospective multicohort study. *Ann. Hepato-Biliary-Pancreatic Surg.* 22:305. doi: 10.14701/ahbps.2018.22.4.305
- Choi, S. Y., Choi, M. J., Cho, M. Y., and Park, Y. J. (2019). Alkbh5, a RNA demethylase, is involved in fine-tuning of cell differentiation (FS11-07-19). *Curr. Dev. Nutr.* 3:7051274. doi: 10.1093/cdn/nzz037.FS11-07-19
- Choudhary, C., Kumar, C., Gnäd, F., Nielsen, M. L., Rehman, M., Walther, T. C., et al. (2009). Lysine acetylation targets protein complexes and co-regulates major cellular functions. *Science* 325, 834–840. doi: 10.1126/science.1175371
- Courtney, D. G., Kennedy, E. M., Dumm, R. E., Bogerd, H. P., Tsai, K., Heaton, N. S., et al. (2017). Epitranscriptomic enhancement of influenza A virus gene expression and replication. *Cell Host Microbe* 22, 377–386.e5. doi: 10.1016/j.chom.2017.08.004
- Dang, W., Xie, Y., Cao, P., Xin, S., Wang, J., Li, S., et al. (2019). N6-Methyladenosine and viral infection. *Front. Microbiol.* 10:417. doi: 10.3389/fmicb.2019.00417
- Delaney, J. C., and Essigmann, J. M. (2004). Mutagenesis, genotoxicity, and repair of 1-methyladenine, 3-alkylcytosines, 1-methylguanine and 3-methylthymine in alkB *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* 101, 14051–14056. doi: 10.1073/pnas.0403489101
- Delaney, J. C., Smeester, L., Wong, C., Frick, L. E., Taghizadeh, K., Wishnok, J. S., et al. (2005). AlkB reverses etheno DNA lesions caused by lipid oxidation *in vitro* and *in vivo*. *Nat. Struct. Mol. Biol.* 12, 855–860. doi: 10.1038/nsmb996
- Deng, X., Su, R., Feng, X., Wei, M., and Chen, J. (2018a). Role of N 6 -methyladenosine modification in cancer. *Curr. Opin. Genet. Dev.* 48, 1–7. doi: 10.1016/j.gde.2017.10.005
- Deng, X., Su, R., Stanford, S., and Chen, J. (2018b). Critical enzymatic functions of FTO in obesity and cancer. *Front. Endocrinol.* 9:396. doi: 10.3389/fendo.2018.00396
- Deng, X., Su, R., Weng, H., Huang, H., Li, Z., and Chen, J. (2018c). RNA N 6 -methyladenosine modification in cancers: current status and perspectives. *Cell Res.* 28, 507–517. doi: 10.1038/s41422-018-0034-6
- Ding, C., Zou, Q., Ding, J., Ling, M., Wang, W., Li, H., et al. (2018). Increased N6-methyladenosine causes infertility is associated with FTO expression. *J. Cell. Physiol.* 233, 7055–7066. doi: 10.1002/jcp.26507
- Dixit, D., Xie, Q., Rich, J. N., and Zhao, J. C. (2017). Messenger RNA methylation regulates glioblastoma tumorigenesis. *Cancer Cell* 31, 474–475. doi: 10.1016/j.ccell.2017.03.010
- Dominissini, D., Moshitch-Moshkovitz, S., Schwartz, S., Salmon-Divon, M., Ungar, L., Osenberg, S., et al. (2012). Topology of the human and mouse m6A RNA methylomes revealed by m6A-seq. *Nature* 485, 201–206. doi: 10.1038/nature11112
- Ensfelder, T. T., Kurz, M. Q., Iwan, K., Geiger, S., Matheisl, S., Müller, M., et al. (2018). ALKBH5-induced demethylation of mono- and dimethylated adenosine. *Chem. Commun.* 54, 8591–8593. doi: 10.1039/C8CC03980A
- Falnes, P., Johansen, R. F., and Seeberg, E. (2002). AlkB-mediated oxidative demethylation reverses DNA damage in *Escherichia coli*. *Nature* 419, 178–182. doi: 10.1038/nature01048

- Feng, C., Liu, Y., Wang, G., Deng, Z., Zhang, Q., Wu, W., et al. (2014). Crystal structures of the human RNA demethylase alkhh5 reveal basis for substrate recognition. *J. Biol. Chem.* 289, 11571–11583. doi: 10.1074/jbc.M113.546168
- Flanagan, J. U., Yosaatmadja, Y., Teague, R. M., Chai, M. Z. L., Turnbull, A. P., and Squire, C. J. (2012). Crystal structures of three classes of non-steroidal anti-inflammatory drugs in complex with aldo-keto reductase 1C3. *PLoS ONE* 7:e43965. doi: 10.1371/journal.pone.0043965
- Forristal, C. E., Wright, K. L., Hanley, N. A., Oreffo, R. O. C., and Houghton, F. D. (2010). Hypoxia inducible factors regulate pluripotency and proliferation in human embryonic stem cells cultured at reduced oxygen tensions. *Reproduction* 139, 85–97. doi: 10.1530/REP-09-0300
- Fu, Y., Dai, Q., Zhang, W., Ren, J., Pan, T., and He, C. (2010). The AlkB domain of mammalian ABH8 catalyzes hydroxylation of 5-methoxycarbonylmethyluridine at the wobble position of tRNA. *Angew. Chem. Int. Ed.* 49, 8885–8888. doi: 10.1002/anie.201001242
- Fu, Y., Jia, G., Pang, X., Wang, R. N., Wang, X., Li, C. J., et al. (2013). FTO-mediated formation of N6-hydroxymethyladenosine and N 6-formyladenosine in mammalian RNA. *Nat. Commun.* 4:2822. doi: 10.1038/ncomms2822
- Gerken, T., Girard, C. A., Tung, Y. C. L., Webby, C. J., Saudek, V., Hewitson, K. S., et al. (2007). The obesity-associated FTO gene encodes a 2-oxoglutarate-dependent nucleic acid demethylase. *Science* 318, 1469–1472. doi: 10.1126/science.1151710
- Gokhale, N. S., McIntyre, A. B. R., McFadden, M. J., Roder, A. E., Kennedy, E. M., Gandara, J. A., et al. (2016). N6-Methyladenosine in flaviviridae viral RNA genomes regulates infection. *Cell Host Microbe* 20, 654–665. doi: 10.1016/j.chom.2016.09.015
- Gulati, P., Cheung, M. K., Antrobus, R., Church, C. D., Harding, H. P., Tung, Y. C. L., et al. (2013). Role for the obesity-related FTO gene in the cellular sensing of amino acids. *Proc. Natl. Acad. Sci. U.S.A.* 110, 2557–2562. doi: 10.1073/pnas.1222796110
- Han, Z., Niu, T., Chang, J., Lei, X., Zhao, M., Wang, Q., et al. (2010). Crystal structure of the FTO protein reveals basis for its substrate specificity. *Nature* 464, 1205–1209. doi: 10.1038/nature08921
- Hao, H., Hao, S., Chen, H., Chen, Z., Zhang, Y., Wang, J., et al. (2019). N6-methyladenosine modification and METTL3 modulate enterovirus 71 replication. *Nucleic Acids Res.* 47, 362–374. doi: 10.1093/nar/gky1007
- He, Y., Hu, H., Wang, Y., Yuan, H., Lu, Z., Wu, P., et al. (2018). ALKBH5 inhibits pancreatic cancer motility by decreasing long non-coding RNA KCNK15-AS1 methylation. *Cell. Physiol. Biochem.* 48, 838–846. doi: 10.1159/000491915
- Hesser, C. R., Karjolic, J., Dominissini, D., He, C., and Glaunsinger, B. A. (2018). N6-methyladenosine modification and the YTHDF2 reader protein play cell type specific roles in lytic viral gene expression during Kaposi's sarcoma-associated herpesvirus infection. *PLoS Pathog.* 14:e1006995. doi: 10.1371/journal.ppat.1006995
- Hu, B. B., Wang, X. Y., Gu, X. Y., Zou, C., Gao, Z. J., Zhang, H., et al. (2019). N6-methyladenosine (m6A) RNA modification in gastrointestinal tract cancers: roles, mechanisms, and applications. *Mol. Cancer* 18:178. doi: 10.1186/s12943-019-1099-7
- Huang, H., Weng, H., and Chen, J. (2020). m6A modification in coding and non-coding RNAs: roles and therapeutic implications in cancer. *Cancer Cell* 37, 270–288. doi: 10.1016/j.ccell.2020.02.004
- Huang, J., and Yin, P. (2018). Structural insights into N6-methyladenosine (m6A) modification in the transcriptome. *Genomics Proteomics Bioinform.* 16, 85–98. doi: 10.1016/j.gpb.2018.03.001
- Huang, Y., Su, R., Sheng, Y., Dong, L., Dong, Z., Xu, H., et al. (2019). Small-molecule targeting of oncogenic FTO demethylase in acute myeloid leukemia. *Cancer Cell* 35, 677–691. doi: 10.1016/j.ccell.2019.03.006
- Huang, Y., Yan, J., Li, Q., Li, J., Gong, S., Zhou, H., et al. (2015). Meclofenamic acid selectively inhibits FTO demethylation of m6A over ALKBH5. *Nucleic Acids Res.* 43, 373–384. doi: 10.1093/nar/gku1276
- Iles, M. M., Law, M. H., Stacey, S. N., Han, J., Fang, S., Pfeiffer, R., et al. (2013). A variant in FTO shows association with melanoma risk not due to BMI. *Nat. Genet.* 45, 428–432. doi: 10.1038/ng.2571
- Imam, H., Khan, M., Gokhale, N. S., McIntyre, A. B. R., Kim, G. W., Jang, J. Y., et al. (2018). N6-methyladenosine modification of hepatitis b virus RNA differentially regulates the viral life cycle. *Proc. Natl. Acad. Sci. U.S.A.* 115, 8829–8834. doi: 10.1073/pnas.1808319115
- Int, C., Chen, S., Zhou, L., and Wang, Y. (2020). ALKBH5 - mediated - m 6 A demethylation of lncRNA PVT1 plays an oncogenic role in osteosarcoma. *Cancer Cell Int.* 20:34. doi: 10.1186/s12935-020-1105-6
- Jia, G., Fu, Y., Zhao, X., Dai, Q., Zheng, G., Yang, Y., et al. (2011). N6-Methyladenosine in nuclear RNA is a major substrate of the obesity-associated FTO. *Nat. Chem. Biol.* 7, 885–887. doi: 10.1038/nchembio.687
- Jia, G., Yang, C. G., Yang, S., Jian, X., Yi, C., Zhou, Z., et al. (2008). Oxidative demethylation of 3-methylthymine and 3-methyluracil in single-stranded DNA and RNA by mouse and human FTO. *FEBS Lett.* 582, 3313–3319. doi: 10.1016/j.febslet.2008.08.019
- Jin, H., Ying, X., Que, B., Wang, X., Chao, Y., Zhang, H., et al. (2019). N6-methyladenosine modification of ITGA6 mRNA promotes the development and progression of bladder cancer. *EBioMedicine* 47, 195–207. doi: 10.1016/j.ebiom.2019.07.068
- Kane, S. E., and Beemon, K. (1985). Precise localization of m6A in rous sarcoma virus RNA reveals clustering of methylation sites: implications for RNA processing. *Mol. Cell. Biol.* 5, 2298–2306. doi: 10.1128/MCB.5.9.2298
- Kataoka, H., Yamamoto, Y., and Sekiguchi, M. (1983). A new gene (alkB) of *Escherichia coli* that controls sensitivity to methyl methane sulfonate. *J. Bacteriol.* 153, 1301–1307. doi: 10.1128/JB.153.3.1301-1307.1983
- Kawarada, L., Suzuki, T., Ohira, T., Hirata, S., Miyauchi, K., and Suzuki, T. (2017). ALKBH1 is an RNA dioxygenase responsible for cytoplasmic and mitochondrial tRNA modifications. *Nucleic Acids Res.* 45, 7401–7415. doi: 10.1093/nar/gkx354
- Kennedy, E. M., Bogerd, H. P., Kornepati, A. V. R., Kang, D., Ghoshal, D., Marshall, J. B., et al. (2016). Posttranscriptional m6A editing of HIV-1 mRNAs enhances viral gene expression. *Cell Host Microbe* 19, 675–685. doi: 10.1016/j.chom.2016.04.002
- Kennedy, E. M., Courtney, D. G., Tsai, K., and Cullen, B. R. (2017). Viral epitranscriptomics. *J. Virol.* 91, e02263–e02216. doi: 10.1128/JVI.02263-16
- Krug, R. M., Morgan, M. A., and Shatkin, A. J. (1976). Influenza viral mRNA contains internal N6-methyladenosine and 5'-terminal 7-methylguanosine in cap structures. *J. Virol.* 20, 45–53. doi: 10.1128/JVI.20.1.45-53.1976
- Lang, F., Singh, R. K., Pei, Y., Zhang, S., Sun, K., and Robertson, E. S. (2019). EBV epitranscriptome reprogramming by METTL14 is critical for viral-associated tumorigenesis. *PLoS Pathog.* 15:e1007796. doi: 10.1371/journal.ppat.1007796
- Levanon, E. Y., Eisenberg, E., Yelin, R., Nemzer, S., Halleger, M., Shemesh, R., et al. (2004). Systematic identification of abundant A-to-I editing sites in the human transcriptome. *Nat. Biotechnol.* 22, 1001–1005. doi: 10.1038/nbt996
- Li, M. M., Nilsen, A., Shi, Y., Fusser, M., Ding, Y. H., Fu, Y., et al. (2013). ALKBH4-dependent demethylation of actin regulates actomyosin dynamics. *Nat. Commun.* 4:2863. doi: 10.1038/ncomms2863
- Li, Q., Huang, Y., Liu, X., Gan, J., Chen, H., and Yang, C. G. (2016). Rhein inhibits AlkB repair enzymes and sensitizes cells to methylated DNA damage. *J. Biol. Chem.* 291, 11083–11093. doi: 10.1074/jbc.M115.711895
- Li, X., Xiong, X., Zhang, M., Wang, K., Chen, Y., Zhou, J., et al. (2017). Base-resolution mapping reveals distinct m1A methylome in nuclear- and mitochondrial-encoded transcripts. *Mol. Cell* 68, 993–1005.e9. doi: 10.1016/j.molcel.2017.10.019
- Li, X. C., Jin, F., Wang, B. Y., Yin, X. J., Hong, W., and Tian, F. J. (2019). The m6A demethylase ALKBH5 controls trophoblast invasion at the maternal-fetal interface by regulating the stability of CYR61 mRNA. *Theranostics* 9, 3853–3865. doi: 10.7150/thno.31868
- Li, Z., Weng, H., Su, R., Weng, X., Zuo, Z., Li, C., et al. (2017). FTO plays an oncogenic role in acute myeloid leukemia as a N6-methyladenosine RNA demethylase. *Cancer Cell* 31, 127–141. doi: 10.1016/j.ccell.2016.11.017
- Lichinchi, G., Gao, S., Saletore, Y., Gonzalez, G. M., Bansal, V., Wang, Y., et al. (2016a). Dynamics of the human and viral m(6)A RNA methylomes during HIV-1 infection of T cells. *Nat. Microbiol.* 1:16011. doi: 10.1038/nmicrobiol.2016.11
- Lichinchi, G., Zhao, B. S., Wu, Y., Lu, Z., Qin, Y., He, C., et al. (2016b). Dynamics of human and viral RNA methylation during zika virus infection. *Cell Host Microbe* 20, 666–673. doi: 10.1016/j.chom.2016.10.002
- Linder, B., Grozhik, A. V., Olarerin-George, A. O., Meydan, C., Mason, C. E., and Jaffrey, S. R. (2015). Single-nucleotide-resolution mapping of m6A and m6Am throughout the transcriptome. *Nat. Methods* 12, 767–772. doi: 10.1038/nmeth.3453

- Liu, F., Clark, W., Luo, G., Wang, X., Fu, Y., Wei, J., et al. (2016). ALKBH1-mediated tRNA demethylation regulates translation. *Cell* 167, 816–828.e16. doi: 10.1016/j.cell.2016.09.038
- Liu, Y., You, Y., Lu, Z., Yang, J., Li, P., Liu, L., et al. (2019). N6-methyladenosine RNA modification-mediated cellular metabolism rewiring inhibits viral replication. *Science* 365, 1171–1176. doi: 10.1126/science.aax4468
- Malacrida, A., Rivara, M., Di Domizio, A., Cislighi, G., Miloso, M., Zuliani, V., et al. (2020). 3D proteome-wide scale screening and activity evaluation of a new ALKBH5 inhibitor in U87 glioblastoma cell line. *Bioorg. Med. Chem.* 28:115300. doi: 10.1016/j.bmc.2019.115300
- Mauer, J., and Jaffrey, S. R. (2018). FTO, m⁶A, and the hypothesis of reversible epitranscriptomic mRNA modifications. *FEBS Lett.* 592, 2012–2022. doi: 10.1002/1873-3468.13092
- Mauer, J., Luo, X., Blanjoie, A., Jiao, X., Grozhik, A. V., Patil, D. P., et al. (2017). Reversible methylation of m⁶A in the 5' cap controls mRNA stability. *Nature* 541, 371–375. doi: 10.1038/nature21022
- McDonough, M. A., Loenarz, C., Chowdhury, R., Clifton, I. J., and Schofield, C. J. (2010). Structural studies on human 2-oxoglutarate dependent oxygenases. *Curr. Opin. Struct. Biol.* 20, 659–672. doi: 10.1016/j.sbi.2010.08.006
- Melstrom, L., and Chen, J. (2020). RNA N⁶-methyladenosine modification in solid tumors: new therapeutic frontiers. *Cancer Gene Ther.* 27, 625–633. doi: 10.1038/s41417-020-0160-4
- Molinie, B., Wang, J., Lim, K. S., Hillebrand, R., Lu, Z. X., Van Wittenberghe, N., et al. (2016). M6 A-LAIC-seq reveals the census and complexity of the m⁶A epitranscriptome. *Nat. Methods* 13, 692–698. doi: 10.1038/nmeth.3898
- Monsen, V. T., Sundheim, O., Aas, P. A., Westbye, M. P., Sousa, M. M. L., Slupphaug, G., et al. (2010). Divergent β -hairpins determine double-strand versus single-strand substrate recognition of human AlkB-homologues 2 and 3. *Nucleic Acids Res.* 38, 6447–6455. doi: 10.1093/nar/gkq518
- Motorin, Y., Lyko, F., and Helm, M. (2009). 5-methylcytosine in RNA: detection, enzymatic formation and biological functions. *Nucleic Acids Res.* 38, 1415–1430. doi: 10.1093/nar/gkp1117
- Müller, T. A., Meek, K., and Hausinger, R. P. (2010). Human AlkB homologue 1 (ABH1) exhibits DNA lyase activity at abasic sites. *DNA Repair.* 9, 58–65. doi: 10.1016/j.dnarep.2009.10.011
- Narayan, P., Ayers, D. F., Rottman, F. M., Maroney, P. A., and Nilsen, T. W. (1987). Unequal distribution of N6-methyladenosine in influenza virus mRNAs. *Mol. Cell. Biol.* 7, 1572–1575. doi: 10.1128/MCB.7.4.1572
- Nettersheim, D., Berger, D., Jostes, S., Kristiansen, G., Lochnit, G., and Schorle, H. (2019). N6-methyladenosine detected in RNA of testicular germ cell tumors is controlled by METTL3, ALKBH5, YTHDC1/F1/F2, and HNRNPC as writers, erasers, and readers. *Andrology* 7, 498–506. doi: 10.1111/andr.12612
- Niu, Y., Lin, Z., Wan, A., Chen, H., Liang, H., Sun, L., et al. (2019). RNA N6-methyladenosine demethylase FTO promotes breast tumor progression through inhibiting BNIP3. *Mol. Cancer* 18, 1–16. doi: 10.1186/s12943-019-1004-4
- Padariya, M., and Kalathiya, U. (2016). Structure-based design and evaluation of novel N-phenyl-1H-indol-2-amine derivatives for fat mass and obesity-associated (FTO) protein inhibition. *Comput. Biol. Chem.* 64, 414–425. doi: 10.1016/j.compbiolchem.2016.09.008
- PilZys, T., Marcinkowski, M., Kukwa, W., Garbicz, D., Dylewska, M., Ferenc, K., et al. (2019). ALKBH overexpression in head and neck cancer: potential target for novel anticancer therapy. *Sci. Rep.* 9:13249. doi: 10.1038/s41598-019-49550-x
- Pinello, N., Sun, S., and Wong, J. J. L. (2018). Aberrant expression of enzymes regulating m⁶A mRNA methylation: implication in cancer. *Cancer Biol. Med.* 15, 323–334. doi: 10.20892/j.issn.2095-3941.2018.0365
- Rajacka, V., Skalicky, T., and Vanacova, S. (2019). The role of RNA adenosine demethylases in the control of gene expression. *Biochim. Biophys. Acta Gene Regul. Mech.* 1862, 343–355. doi: 10.1016/j.bbagr.2018.12.001
- Roundtree, I. A., Evans, M. E., Pan, T., and He, C. (2017). Dynamic RNA modifications in gene expression regulation. *Cell* 169, 1187–1200. doi: 10.1016/j.cell.2017.05.045
- Safra, M., Sas-Chen, A., Nir, R., Winkler, R., Nachshon, A., Bar-Yaacov, D., et al. (2017). The m¹A landscape on cytosolic and mitochondrial mRNA at single-base resolution. *Nature* 551, 251–255. doi: 10.1038/nature24456
- Sendinc, E., Valle-garcia, D., Dhall, A., Gygi, S. P., Sendinc, E., Valle-garcia, D., et al. (2019). PCIF1 catalyzes m⁶A mRNA methylation to regulate gene expression. *Mol. Cell* 75, 620–630.e9. doi: 10.1016/j.molcel.2019.05.030
- Shepard, C. W., Simard, E. P., Finelli, L., Fiore, A. E., and Bell, B. P. (2006). Hepatitis B virus infection: epidemiology and vaccination. *Epidemiol. Rev.* 28, 112–125. doi: 10.1093/epirev/mxj009
- Shriwas, O., Priyadarshini, M., Samal, S. K., Rath, R., Panda, S., Das Majumdar, S. K., et al. (2020). DDX3 modulates cisplatin resistance in OSCC through ALKBH5-mediated m⁶A-demethylation of FOXM1 and NANOG. *Apoptosis* 25, 233–246. doi: 10.1007/s10495-020-01591-8
- Singh, B., Kinne, H. E., Milligan, R. D., Washburn, L. J., Olsen, M., and Lucci, A. (2016). Important role of FTO in the survival of rare panresistant triple-negative inflammatory breast cancer cells facing a severe metabolic challenge. *PLoS ONE* 11:e0159072. doi: 10.1371/journal.pone.0159072
- Smemo, S., Tena, J. J., Kim, K. H., Gamazon, E. R., Sakabe, N. J., Gómez-Marín, C., et al. (2014). Obesity-associated variants within FTO form long-range functional connections with IIRX3. *Nature* 507, 371–375. doi: 10.1038/nature13138
- Song, H., Feng, X., Zhang, H., Luo, Y., Huang, J., Lin, M., et al. (2019). METTL3 and ALKBH5 oppositely regulate m⁶A modification of TFEB mRNA, which dictates the fate of hypoxia/reoxygenation-treated cardiomyocytes. *Autophagy* 15, 1419–1437. doi: 10.1080/15548627.2019.1586246
- Su, R., Dong, L., Li, C., Nachtergaele, S., Wunderlich, M., Qing, Y., et al. (2018). R-2HG exhibits anti-tumor activity by targeting FTO/m⁶A/MYC/CEBPA signaling. *Cell* 172, 90–105.e23. doi: 10.1016/j.cell.2017.11.031
- Sundheim, O., Talstad, V. A., Vågbo, C. B., Slupphaug, G., and Krokan, H. E. (2008). AlkB demethylases flip out in different ways. *DNA Repair.* 7, 1916–1923. doi: 10.1016/j.dnarep.2008.07.015
- Sundheim, O., Vågbo, C. B., Björås, M., Sousa, M. M. L., Talstad, V., Aas, P. A., et al. (2006). Human ABH3 structure and key residues for oxidative demethylation to reverse DNA/RNA damage. *EMBO J.* 25, 3389–3397. doi: 10.1038/sj.emboj.7601219
- Svensen, N., and Jaffrey, S. R. (2016). Fluorescent RNA aptamers as a tool to study RNA-modifying enzymes. *Cell Chem. Biol.* 23, 415–425. doi: 10.1016/j.chembiol.2015.11.018
- Tan, B., and Gao, S.-J. (2018). RNA epitranscriptomics: regulation of infection of RNA and DNA viruses by N⁶-methyladenosine (m⁶A). *Rev. Med. Virol.* 28, 1–22. doi: 10.1002/rmv.1983
- Tan, B., Liu, H., Zhang, S., Zhang, L., Cui, X., Yuan, H., et al. (2018). Viral and cellular N⁶-methyladenosine (m⁶A) and N⁶, 2'-O-dimethyladenosine (m⁶Am) epitranscriptomes in KSHV life cycle. *Nat. Microbiol.* 3, 108–120. doi: 10.1038/s41564-017-0056-8
- Tang, B., Yang, Y., Kang, M., Wang, Y., Bi, Y., et al. (2020). M6A demethylase ALKBH5 inhibits pancreatic cancer tumorigenesis by decreasing WIF-1 RNA methylation and mediating Wnt signaling. *Mol. Cancer* 19, 1–15. doi: 10.1186/s12943-019-1128-6
- Tang, C., Klukovich, R., Peng, H., Wang, Z., Yu, T., Zhang, Y., et al. (2017). ALKBH5-dependent m⁶A demethylation controls splicing and stability of long 3'-UTR mRNAs in male germ cells. *Proc. Natl. Acad. Sci. U.S.A.* 115, E325–E333. doi: 10.1073/pnas.1717794115
- Thalhammer, A., Bencokova, Z., Poole, R., Loenarz, C., Adam, J., O'Flaherty, L., et al. (2011). Human AlkB homologue 5 is a nuclear 2-oxoglutarate dependent oxygenase and a direct target of hypoxia-inducible factor 1 α (HIF-1 α). *PLoS ONE* 6:e16210. doi: 10.1371/journal.pone.0016210
- Tian, L. F., Liu, Y. P., Chen, L., Tang, Q., Wu, W., Sun, W., et al. (2020). Structural basis of nucleic acid recognition and 6mA demethylation by human ALKBH1. *Cell Res.* 30, 272–275. doi: 10.1038/s41422-019-0233-9
- Tirumuru, N., Zhao, B. S., Lu, W., Lu, Z., He, C., and Wu, L. (2016). N6-methyladenosine of HIV-1 RNA regulates viral infection and HIV-1 Gag protein expression. *Microbiol. Infect. Dis.* 5:e15528. doi: 10.7554/eLife.15528.021
- Toh, J. D. W., Sun, L., Lau, L. Z. M., Tan, J., Low, J. J. A., Tang, C. W. Q., et al. (2015). A strategy based on nucleotide specificity leads to a subfamily-selective and cell-active inhibitor of N6-methyladenosine demethylase FTO. *Chem. Sci.* 6, 112–122. doi: 10.1039/C4SC02554G
- Tsai, K., Courtney, D. G., and Cullen, B. R. (2018). Addition of m⁶A to SV40 late mRNAs enhances viral structural gene expression and replication. *PLoS Pathog.* 14:e1006919. doi: 10.1371/journal.ppat.1006919

- Wang, G., He, Q., Feng, C., Liu, Y., Deng, Z., Qi, X., et al. (2014). The atomic resolution structure of human alkB homolog 7 (ALKBH7), a key protein for programmed necrosis and fat metabolism. *J. Biol. Chem.* 289, 27924–27936. doi: 10.1074/jbc.M114.590505
- Wang, R., Han, Z., Liu, B., Zhou, B., Wang, N., Jiang, Q., et al. (2018). Identification of natural compound radicicol as a potent FTO inhibitor. *Mol. Pharm.* 15, 4092–4098. doi: 10.1021/acs.molpharmaceut.8b00522
- Wang, T., Hong, T., Huang, Y., Su, H., Wu, F., Chen, Y., et al. (2015). Fluorescein derivatives as bifunctional molecules for the simultaneous inhibiting and labeling of FTO protein. *J. Am. Chem. Soc.* 137, 13736–13739. doi: 10.1021/jacs.5b06690
- Wang, Y., Li, J., Han, X., Wang, N., Song, C., Wang, R., et al. (2019). Identification of clausine E as an inhibitor of fat mass and obesity-associated protein (FTO) demethylase activity. *J. Mol. Recognit.* 32:e2800. doi: 10.1002/jmr.2800
- Wei, J., Liu, F., Lu, Z., Fei, Q., Ai, Y., He, P. C., et al. (2018). Differential m⁶A, m⁶A m, and m¹A demethylation mediated by FTO in the cell nucleus and cytoplasm. *Mol. Cell* 71, 973–985.e5. doi: 10.1016/j.molcel.2018.08.011
- Weng, H., Huang, H., and Chen, J. (2019). RNA N⁶-methyladenosine modification in normal and malignant hematopoiesis. *Adv. Exp. Med. Biol.* 1143, 75–93. doi: 10.1007/978-981-13-7342-8_4
- Westbye, M. P., Feyzi, E., Aas, P. A., Vågbo, C. B., Talstad, V. A., Kavli, B., et al. (2008). Human AlkB homolog 1 is a mitochondrial protein that demethylates 3-methylcytosine in DNA and RNA. *J. Biol. Chem.* 283, 25046–25056. doi: 10.1074/jbc.M803776200
- Winkler, R., Gillis, E., Lasman, L., Safran, M., Geula, S., Soyris, C., et al. (2019). m⁶A modification controls the innate immune response to infection by targeting type I interferons. *Nat. Immunol.* 20, 173–182. doi: 10.1038/s41590-018-0275-z
- Woo, H. H., and Chambers, S. K. (2019). Human ALKBH3-induced m¹A demethylation increases the CSF-1 mRNA stability in breast and ovarian cancer cells. *Biochim. Biophys. Acta Gene Regul. Mech.* 1862, 35–46. doi: 10.1016/j.bbaggm.2018.10.008
- Xu, C., Liu, K., Tempel, W., Demetriades, M., Aik, W., Schofield, C. J., et al. (2014). Structures of human ALKBH5 demethylase reveal a unique binding mode for specific single-stranded N⁶-methyladenosine RNA demethylation. *J. Biol. Chem.* 289, 17299–17311. doi: 10.1074/jbc.M114.550350
- Xu, D., Shao, W., Jiang, Y., Wang, X., Liu, Y., and Liu, X. (2017). FTO expression is associated with the occurrence of gastric cancer and prognosis. *Oncol. Rep.* 38, 2285–2292. doi: 10.3892/or.2017.5904
- Xue, M., Zhao, B. S., Zhang, Z., Lu, M., Harder, O., Chen, P., et al. (2019). Viral N⁶-methyladenosine upregulates replication and pathogenesis of human respiratory syncytial virus. *Nat. Commun.* 10:4595. doi: 10.1038/s41467-019-12504-y
- Yang, C. G., Yi, C., Duguid, E. M., Sullivan, C. T., Jian, X., Rice, P. A., et al. (2008). Crystal structures of DNA/RNA repair enzymes AlkB and ABH2 bound to dsDNA. *Nature* 452, 961–965. doi: 10.1038/nature06889
- Yang, P., Wang, Q., Liu, A., Zhu, J., and Feng, J. (2019). ALKBH5 holds prognostic values and inhibits the metastasis of colon cancer. *Pathol. Oncol. Res.* 26, 1615–1623. doi: 10.1007/s12253-019-00737-7
- Yang, S., Wei, J., Cui, Y. H., Park, G., Shah, P., Deng, Y., et al. (2019). m⁶A mRNA demethylase FTO regulates melanoma tumorigenicity and response to anti-PD-1 blockade. *Nat. Commun.* 10:2782. doi: 10.1038/s41467-019-10669-0
- Ye, F., Chen, E. R., and Nilsen, T. W. (2017). Kaposi's sarcoma-associated herpesvirus utilizes and manipulates RNA N⁶-adenosine methylation to promote lytic replication. *J. Virol.* 91, 1–21. doi: 10.1128/JVI.00466-17
- Ye, F., Lei, X., and Gao, S. J. (2011). Mechanisms of kaposi's sarcoma-associated herpesvirus latency and reactivation. *Adv. Virol.* 2011:193860. doi: 10.1155/2011/193860
- Yu, J., Shen, L., Liu, Y., Ming, H., Zhu, X., Chu, M., et al. (2020). The m⁶A methyltransferase METTL3 cooperates with demethylase ALKBH5 to regulate osteogenic differentiation through NF-κB signaling. *Mol. Cell. Biochem.* 463, 203–210. doi: 10.1007/s11010-019-03641-5
- Zhang, C., Samanta, D., Lu, H., Bullen, J. W., Zhang, H., Chen, I., et al. (2016a). Hypoxia induces the breast cancer stem cell phenotype by HIF-dependent and ALKBH5-mediated m⁶A-demethylation of NANOG mRNA. *Proc. Natl. Acad. Sci. U.S.A.* 113, E2047–E2056. doi: 10.1073/pnas.1602883113
- Zhang, C., Zhang, M., Ge, S., Huang, W., Lin, X., Gao, J., et al. (2019). Reduced m⁶A modification predicts malignant phenotypes and augmented Wnt/PI3K-Akt signaling in gastric cancer. *Cancer Med.* 8, 4766–4781. doi: 10.1002/cam4.2360
- Zhang, C., Zhi, W. I., Lu, H., Samanta, D., Chen, I., Gabrielson, E., et al. (2016b). Hypoxia-inducible factors regulate pluripotency factor expression by ZNF217- and ALKBH5-mediated modulation of RNA methylation in breast cancer cells. *Oncotarget* 7, 64527–64542. doi: 10.18632/oncotarget.11743
- Zhang, J., Guo, S., Piao, H. Y., Wang, Y., Wu, Y., Meng, X., et al. (2019). ALKBH5 promotes invasion and metastasis of gastric cancer by decreasing methylation of the lncRNA NEAT1. *J. Physiol. Biochem.* 75, 379–389. doi: 10.1007/s13105-019-00690-8
- Zhang, M., Yang, S., Nelakanti, R., Zhao, W., Liu, G., Li, Z., et al. (2020). Mammalian ALKBH1 serves as an N⁶-m⁶A demethylase of unpairing DNA. *Cell Res.* 30, 197–210. doi: 10.1038/s41422-019-0237-5
- Zhang, S., Zhao, B. S., Zhou, A., Lin, K., Zheng, S., Lu, Z., et al. (2017). m⁶A demethylase ALKBH5 maintains tumorigenicity of glioblastoma stem-like cells by sustaining FOXM1 expression and cell proliferation program. *Cancer Cell* 31, 591–606.e6. doi: 10.1016/j.ccell.2017.02.013
- Zhang, X., Wei, L. H., Wang, Y., Xiao, Y., Liu, J., Zhang, W., et al. (2019). Structural insights into FTO's catalytic mechanism for the demethylation of multiple RNA substrates. *Proc. Natl. Acad. Sci. U.S.A.* 116, 2919–2924. doi: 10.1073/pnas.1820574116
- Zhao, W., Qi, X., Liu, L., Liu, Z., Ma, S., and Wu, J. (2020). Epigenetic regulation of m⁶A modifications in human cancer. *Mol. Ther. Nucleic Acids* 19, 405–412. doi: 10.1016/j.omtn.2019.11.022
- Zhao, X., Yang, Y., Sun, B. F., Shi, Y., Yang, X., Xiao, W., et al. (2014a). FTO-dependent demethylation of N⁶-methyladenosine regulates mRNA splicing and is required for adipogenesis. *Cell Res.* 24, 1403–1419. doi: 10.1038/cr.2014.151
- Zhao, X., Yang, Y., Sun, B. F., Zhao, Y. L., and Yang, Y. G. (2014b). FTO and obesity: mechanisms of association. *Curr. Diab. Rep.* 14:486. doi: 10.1007/s11892-014-0486-0
- Zheng, G., Cox, T., Tribbey, L., Wang, G. Z., Iacoban, P., Booher, M. E., et al. (2014). Synthesis of a FTO inhibitor with anticonvulsant activity. *ACS Chem. Neurosci.* 5, 658–665. doi: 10.1021/cn500042t
- Zheng, G., Dahl, J. A., Niu, Y., Fedorcsak, P., Huang, C. M., Li, C. J., et al. (2013). ALKBH5 is a mammalian RNA demethylase that impacts RNA metabolism and mouse fertility. *Mol. Cell* 49, 18–29. doi: 10.1016/j.molcel.2012.10.015
- Zheng, Q., Hou, J., Zhou, Y., Li, Z., and Cao, X. (2017). The RNA helicase DDX46 inhibits innate immunity by entrapping m⁶A-demethylated antiviral transcripts in the nucleus. *Nat. Immunol.* 18, 1094–1103. doi: 10.1038/ni.3830
- Zhou, Y., Hambly, B. D., and McLachlan, C. S. (2017). FTO associations with obesity and telomere length. *J. Biomed. Sci.* 24, 1–7. doi: 10.1186/s12929-017-0372-6
- Zhu, H., Gan, X., Jiang, X., Diao, S., Wu, H., and Hu, J. (2019). ALKBH5 inhibited autophagy of epithelial ovarian cancer through miR-7 and BCL-2. *J. Exp. Clin. Cancer Res.* 38, 1–15. doi: 10.1186/s13046-019-1159-2
- Zhu, Z., Qian, Q., Zhao, X., Ma, L., and Chen, P. (2020). N⁶-methyladenosine ALKBH5 promotes non-small cell lung cancer progress by regulating TIMP3 stability. *Gene* 731:144348. doi: 10.1016/j.gene.2020.144348
- Zou, S., Toh, J. D. W., Wong, K. H. Q., Gao, Y. G., Hong, W., and Woon, E. C. Y. (2016). N⁶-Methyladenosine: a conformational marker that regulates the substrate specificity of human demethylases FTO and ALKBH5. *Sci. Rep.* 6:25677. doi: 10.1038/srep25677

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 Bayoumi and Munir. 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:

Jia Meng,
Xi'an Jiaotong-Liverpool
University, China

Reviewed by:

Yongwen Luo,
Wuhan University, China
Carmen Jeronimo,
Universidade do Porto, Portugal

*Correspondence:

Xiang Wang
seanw_hs@163.com
Junhua Zheng
zhengjh0471@sina.com

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

Specialty section:

This article was submitted to
Epigenomics and Epigenetics,
a section of the journal
Frontiers in Cell and Developmental
Biology

Received: 13 October 2020

Accepted: 09 December 2020

Published: 21 January 2021

Citation:

Chen S, Zhang N, Zhang E, Wang T,
Jiang L, Wang X and Zheng J (2021)
A Novel m⁶A Gene Signature
Associated With Regulatory Immune
Function for Prognosis Prediction in
Clear-Cell Renal Cell Carcinoma.
Front. Cell Dev. Biol. 8:616972.
doi: 10.3389/fcell.2020.616972

A Novel m⁶A Gene Signature Associated With Regulatory Immune Function for Prognosis Prediction in Clear-Cell Renal Cell Carcinoma

Siteng Chen^{1†}, Ning Zhang^{2†}, Encheng Zhang^{1†}, Tao Wang¹, Liren Jiang³, Xiang Wang^{1*}
and Junhua Zheng^{1*}

¹ Department of Urology, Shanghai General Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China,

² Department of Urology, Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China, ³ Department
of Pathology, Shanghai General Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China

The important role of N⁶-methyladenosine (m⁶A) RNA methylation regulator in carcinogenesis and progression of clear-cell renal cell carcinoma (ccRCC) is poorly understood by now. In this study, we performed comprehensive analyses of m⁶A RNA methylation regulators in 975 ccRCC samples and 332 adjacent normal tissues and identified ccRCC-related m⁶A regulators. Moreover, the m⁶A diagnostic score based on ccRCC-related m⁶A regulators could accurately distinguish ccRCC from normal tissue in the Meta-cohort, which was further validated in the independent GSE-cohort and The Cancer Genome Atlas-cohort, with an area under the curve of 0.924, 0.867, and 0.795, respectively. Effective survival prediction of ccRCC by m⁶A risk score was also identified in the Cancer Genome Atlas training cohort and verified in the testing cohort and the independent GSE22541 cohort, with hazard ratio values of 3.474, 1.679, and 2.101 in the survival prognosis, respectively. The m⁶A risk score was identified as a risk factor of overall survival in ccRCC patients by the univariate Cox regression analysis, which was further verified in both the training cohort and the independent validation cohort. The integrated nomogram combining m⁶A risk score and predictable clinicopathologic factors could accurately predict the survival status of the ccRCC patients, with an area under the curve values of 85.2, 82.4, and 78.3% for the overall survival prediction in 1-, 3- and 5-year, respectively. Weighted gene co-expression network analysis with functional enrichment analysis indicated that m⁶A RNA methylation might affect clinical prognosis through regulating immune functions in patients with ccRCC.

Keywords: N⁶-methyladenosine (m⁶A), clear cell renal cell carcinoma, WGCNA, regulatory immune function, nomogram

INTRODUCTION

It is estimated that there will be 73,750 new cases of renal malignant tumors in the United States in 2020 (Siegel et al., 2020). As one of the most aggressive malignancies, renal cell carcinoma (RCC) accounts for 2–3% of the malignancies (Ljungberg et al., 2015). Clinically, clear-cell RCC (ccRCC) is the most common type of renal carcinoma, representing ~80% of RCC. Although the therapy methods of ccRCC, including the surgical and targeted therapy, have been improved, the poor survival prognosis is also noticed in ccRCC, which is obligated to most metastatic cases of RCC (Reuter, 2006). Even for the localized ccRCC, the tumor recurrence and progression can also be observed in ~25% of patients after primary treatment (De et al., 2014). Therefore, there is still an urgent need to explore effective prognostic markers and fully clarify the molecular mechanism underlying the tumorigenesis of ccRCC.

Currently, the risk stratification and prognostic prediction for the ccRCC patients are mainly conducted using the Fuhrman grade and tumor–node–metastasis stage system. Even patients with similar clinicopathologic characteristics could also suffer from variable survival outcomes. Taking into account the heterogeneity in ccRCC, several potential biomarkers for diagnosis and survival prediction have been reported in recent years, such as carbonic anhydrase 9 (Genega et al., 2010), PBRM1, and BAP1 mutation (Varela et al., 2011). However, few of these potential biomarkers could finally be transformed into clinical practice due to the unsatisfactory accuracy and sensibility.

The modification and epigenetic alteration of RNA have recently been primarily identified, including N⁶-methyladenosine (m⁶A) (Roundtree et al., 2017; Boccaletto et al., 2018). m⁶A RNA methylation acts as one kind of conserved internal modifications in eukaryotic nuclear RNAs (Dubin and Taylor, 1975), which is one of the most common RNA modifications and plays diverse roles in various biological processes (Bokar et al., 1997; Dominissini et al., 2012). The m⁶A status in a cell is regulated by three types of m⁶A RNA methylation genes, including methyltransferases called writers (RBM15, METTL14, KIAA1429, and others), readers (FMR1, IGF2BP1, YTHDC1, and others), and demethylases called erasers (ALKBH5 and FTO) (Li A. et al., 2017). At present, the low expression levels of FTO and ALKBH5 have been proved to be associated with worse cancer-specific survival of RCC patients after nephrectomy (Strick et al., 2020). However, the specific role of m⁶A regulators in carcinogenesis and progression of ccRCC has not been fully understood by now.

In this study, we carried out a comprehensive analysis based on the expressions of m⁶A RNA methylation regulator to identify their important roles in the diagnosis and survival prediction for the ccRCC patients.

MATERIALS AND METHODS

Datasets Source and Screen

Data sets from The Cancer Genome Atlas (TCGA) and Gene Expression Omnibus (GEO) were strictly screened for this study. The inclusion criteria were as follows: studies with more than

60 ccRCC patients and studies with complete clinical data and gene expression data of m⁶A RNA methylation genes. Five GEO datasets, including GSE46699 (Eckel-Passow et al., 2014), GSE53757 (von Roemeling et al., 2014), GSE22541 (Wuttig et al., 2012), GSE17895 (Dalglish et al., 2010), and GSE40435 (Wozniak et al., 2013), were finally selected and downloaded with gene expression data and clinical data. Four datasets (GSE46699, GSE53757, GSE22541, and GSE17895) with microarray data were detected based on the Affymetrix Human Genome U133 Plus 2.0 Array. We firstly performed background adjustment and quantile normalization from the raw CEL data of the four datasets by the Robust Multi-array Average in the R environment (Irizarry et al., 2003). And then, we annotated the probe of each gene by using the hgu133plus2.db in R. Finally, the processed microarray data from the four datasets were merged as the Meta-cohort after batch normalization. The processed expression data of GSE40435 based on the Illumina HumanHT-12 V4.0 was directly downloaded from the GEO database, as it had been normalized in a previous study (Wozniak et al., 2013), which was defined as the GSE-cohort. TCGA-KIRC data set, including normalized RNA-sequencing data and clinical data, was directly downloaded from the TCGA database. Expression data files of 21 m⁶A RNA methylation genes were further extracted from the processed data sets mentioned earlier. Basic clinical characteristics of data sets in this study are shown in **Supplementary Table 1**.

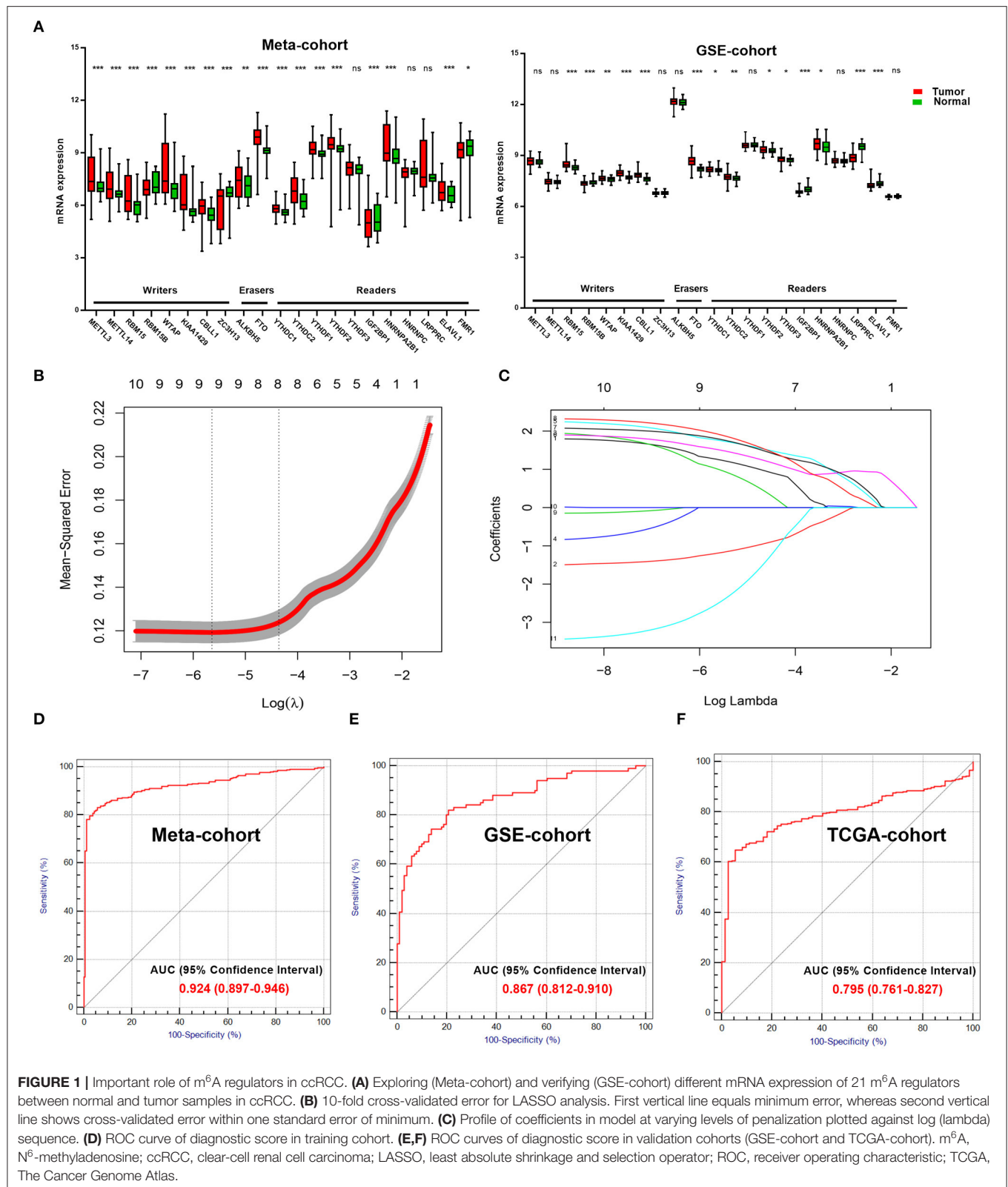
Construction of the Diagnostic and Prognosis Model Based on the m⁶A Regulators for Clear-Cell Renal Cell Carcinoma

Differential expression analysis of the 21 m⁶A RNA methylation genes between normal and tumor samples was performed in the Meta-cohort and the GSE-cohort. Only m⁶A RNA methylation genes differentially expressing in both the two cohorts were further used for the construction of the diagnostic model. We carried out the least absolute shrinkage and selection operator (LASSO) analysis via the *glmnet* package to identify ccRCC-related m⁶A regulators and calculate coefficients of each gene in the Meta-cohort. The number of lambda values in LASSO was set as 1,000 to ensure the robustness of our diagnostic models. The m⁶A diagnostic score was then calculated as follow:

$$\text{m6A diagnostic score} = \sum_{i=1}^n (\text{Coef}_i * D_i)$$

Coef_i means the coefficient of each ccRCC-related m⁶A RNA methylation gene, whereas D_i represents the related messenger RNA (mRNA) expression. The diagnostic model was further verified in two independent cohorts (GSE-cohort and TCGA-cohort).

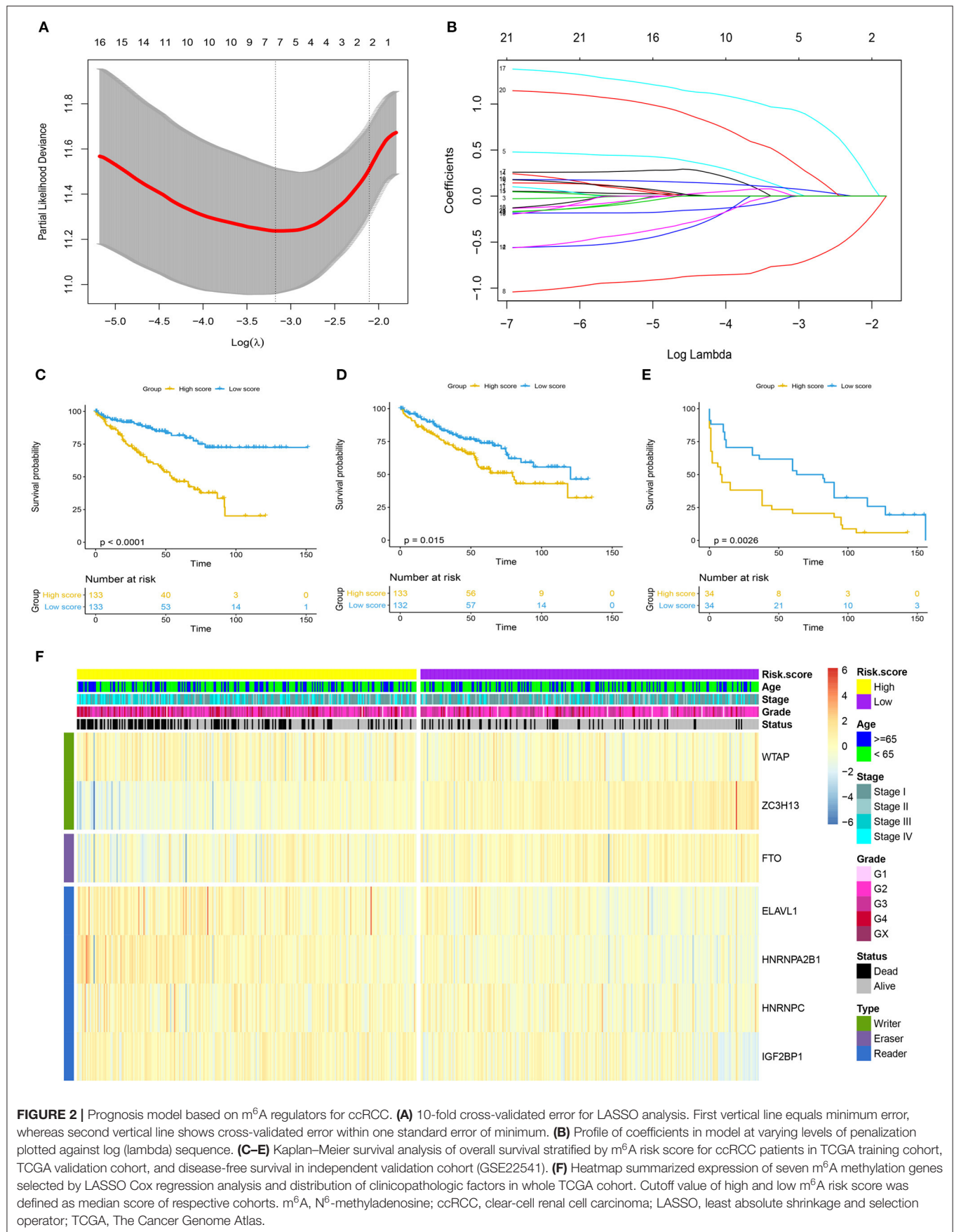
To construct the prognosis model based on the m⁶A regulators for ccRCC, 531 ccRCC patients from the TCGA cohort were randomly assigned to a training cohort and a testing cohort by the random number method. We performed a LASSO Cox regression analysis to identified survival-related m⁶A RNA methylation genes and calculated their coefficients in the TCGA



training cohort. The m⁶A risk score was then calculated as follow:

$$\text{m6A risk score} = \sum_{i=1}^n (\text{Coef}_i * R_i)$$

Coef_i means the coefficient of each survival-related m⁶A RNA methylation gene, whereas R_i represents the related mRNA expression. The prognosis model was further validated in the



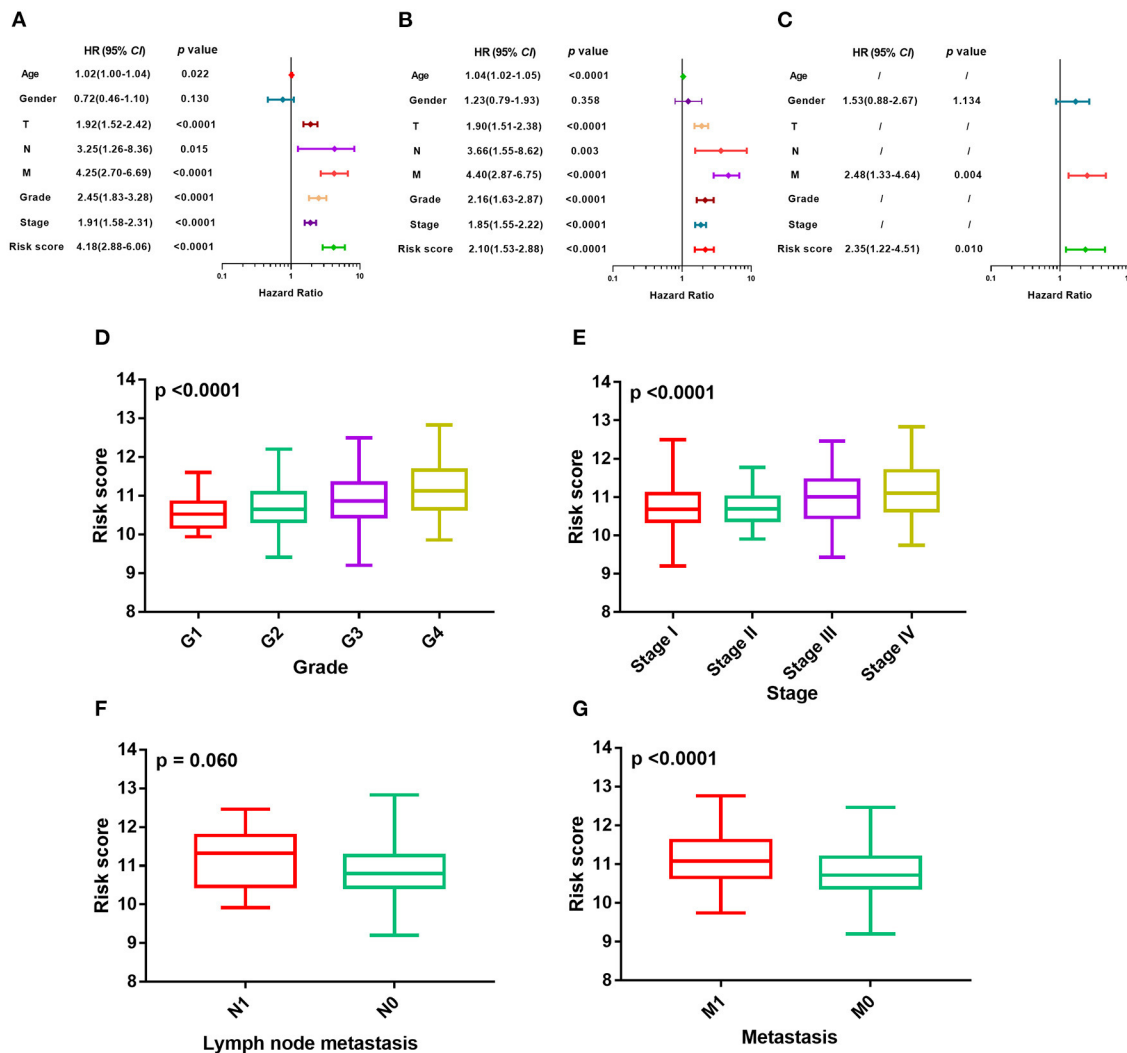


FIGURE 3 | Evaluation of prognosis model based on m⁶A regulators. (A–C) Univariate Cox regression analyses of m⁶A risk score and clinicopathologic factors in TCGA training cohort, TCGA validation cohort, and independent validation cohort. (D–G) Difference in distribution of m⁶A risk score of different tumor stages, tumor grades, metastasis status, and lymph node metastasis status in whole TCGA cohort. m⁶A, N⁶-methyladenosine; TCGA, The Cancer Genome Atlas.

TCGA testing cohort and the independent validation cohort (GSE22541). The cutoff value of high and low m⁶A risk score was defined as the median score of respective cohorts.

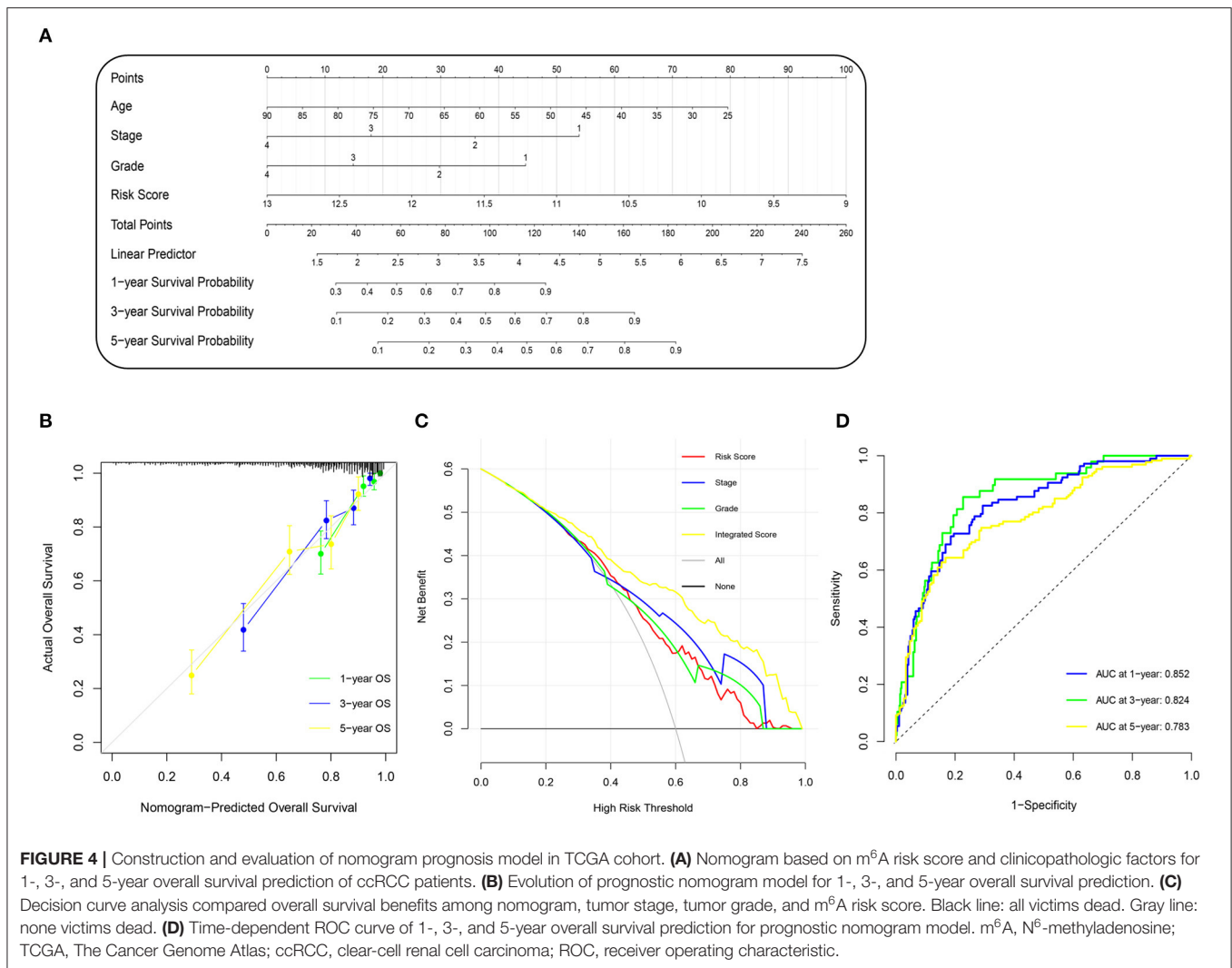
Development and Evaluation of the Prognostic Nomogram Model

We developed a prognostic nomogram model based on m⁶A risk score, age, tumor stage, and tumor grade using the *rms* and *nomogramEx* packages in the R environment. The calibration with bootstrapping was conducted to verify the nomogram-predicted probabilities of the 1-, 3-, and 5-year overall survival (OS) via plotting on the x-axis, with actual OS probabilities plotting on the y-axis. The time-dependent receiver operating characteristic (ROC) curve analysis and the decision

curve analysis were also performed to identify the specificity, sensitivity, and clinical utility of the prognostic model.

Weighted Gene Co-expression Network and Functional Enrichment Analysis

The significantly differentially expressed genes (DEGs) were firstly identified between the ccRCC tissue and the normal renal tissue from GSE40435, GSE53757, and GSE46699. The weighted gene co-expression network analysis (WGCNA) was then performed based on the DEGs by the WGCNA package in R (Langfelder and Horvath, 2008). When the soft-thresholding power of β value was defined as 8, the recruited DEGs were hierarchically clustered into seven gene modules. The correlation between gene module and clinical characteristic was further analyzed to identify the



optimal gene module with the highest correlation. We finally conducted the Kyoto Encyclopedia of Genes and Genomes pathway and the Genetic Ontology analysis by the Metascape (Zhou et al., 2019) to explore the potential biological mechanisms in which the m⁶A risk score might be involved in.

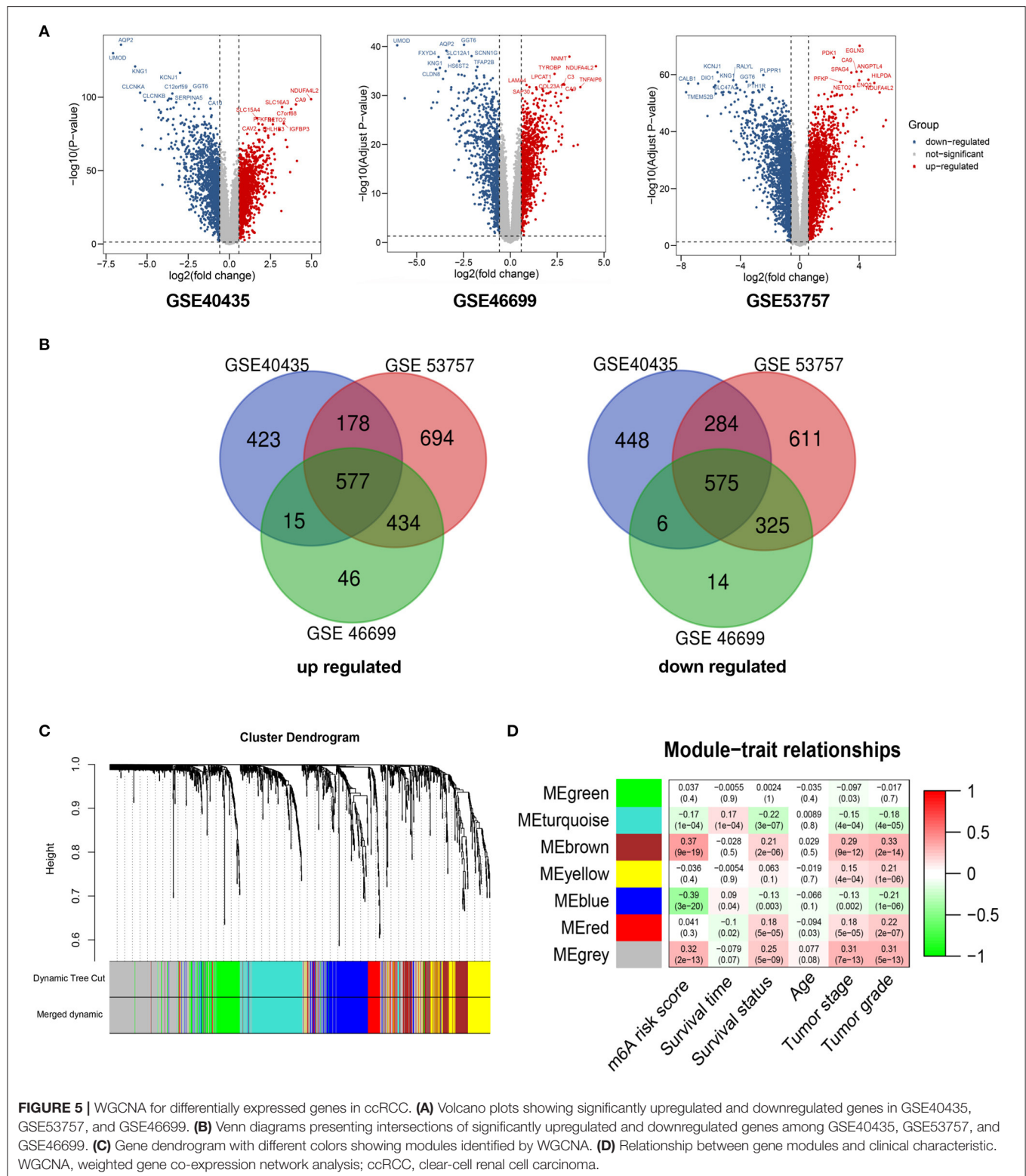
Statistical Analysis

The continuous variable with abnormal distribution was analyzed by non-parametric tests (Mann–Whitney *U* test for comparison between two groups and Kruskal–Wallis test for comparisons among more than two groups). The log-rank test was used for analyzing the Kaplan–Meier curves of OS and disease-free survival. The Cox regression analyses were carried out to identify the m⁶A risk score as a prognostic factor of OS. R (3.6.2) and Statistical Package for Social Sciences 24.0 software (SPSS Inc., Chicago, IL, USA) were used for statistical analysis and data visualization. A two-tailed *P*-value of <0.05 was considered significant.

RESULTS

Important Roles of m⁶A Regulators in Clear-Cell Renal Cell Carcinoma

We explored and verified 21 m⁶A regulators with differential mRNA expressions between the tumor samples and normal samples in the Meta-cohort and the GSE-cohort (Figure 1A). Only 11 m⁶A regulators were significantly differentially expressed in both the two cohorts, including five writers (RBM15, RBM15B, WTAP, KIAA1429, and CBLL1), one eraser (FTO), and five readers (YTHDC1, YTHDC2, YTHDF2, IGF2BP1, and HNRNPA2B1). The 11 m⁶A regulators were further analyzed by the LASSO analysis in the Meta-cohort (Figure 1B). Finally, eight ccRCC-related m⁶A RNA methylation genes, including RBM15, WTAP, CBLL1, FTO, YTHDC1, YTHDC2, YTHDF2, and HNRNPA2B1, were selected for the construction of the m⁶A-related diagnostic model (Figure 1C). The ROC analysis revealed that the area under the curve (AUC) of the diagnostic model reached 0.924 [95% confidence interval (CI): 0.897–0.946] in Meta-cohort (Figure 1D). This diagnostic model was further validated in



two independent cohorts, including GSE-cohort and TCGA-cohort. AUC values reached 0.867 (95% CI: 0.812–0.910) in GSE-cohort (**Figure 1E**) and 0.795 (95% CI: 0.761–0.827) in

TCGA-cohort (**Figure 1F**), which proved the stability of the diagnostic efficiency of our m⁶A-related diagnostic model for the ccRCC patients.

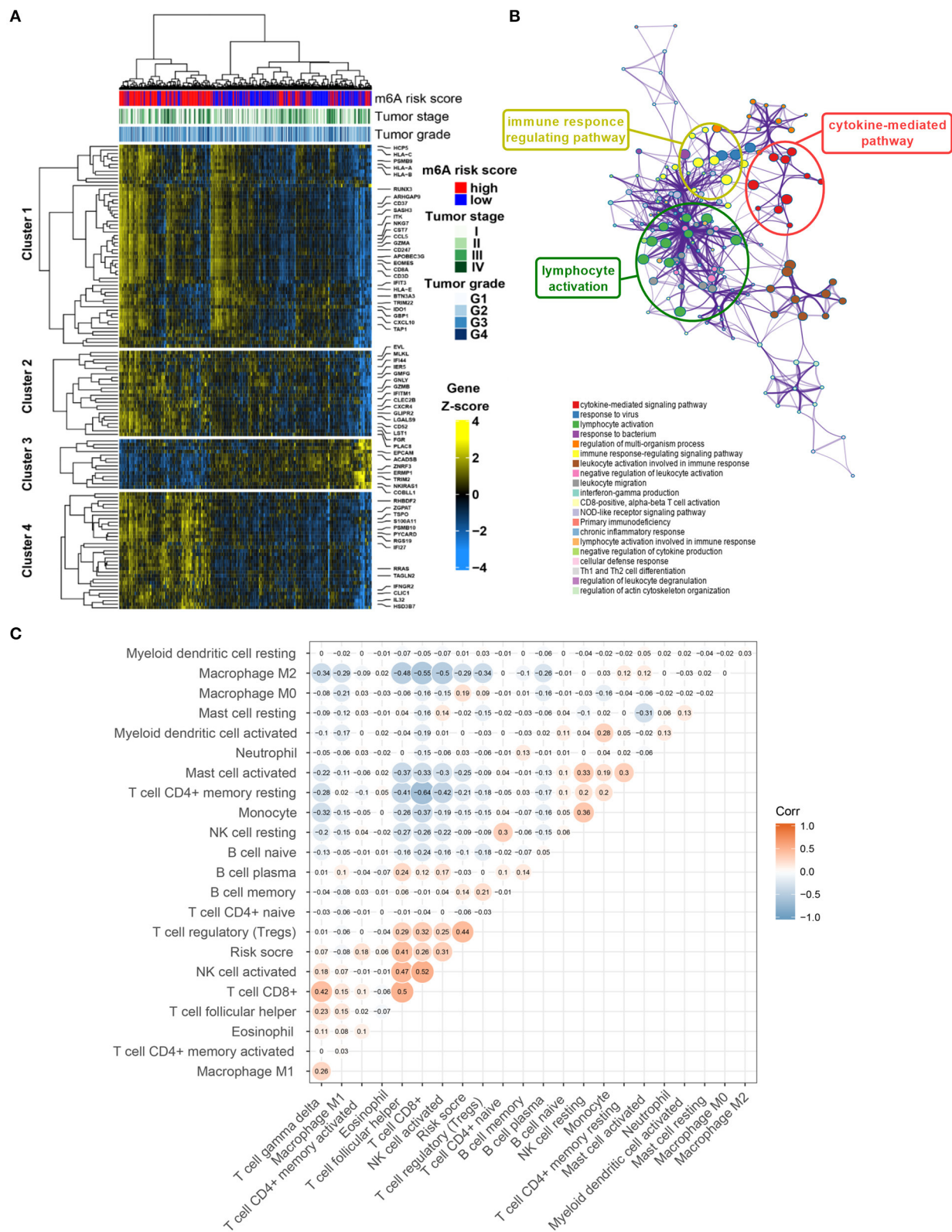


FIGURE 6 | Potential mechanisms of co-expressed genes associated with m⁶A risk score. **(A)** Heatmap visualizing expressions of co-expressed genes in brown module. **(B)** Potentially enriched pathways of co-expressed genes in brown module. **(C)** Correlation analysis of m⁶A risk score and abundance of 22 types of immune cells. m⁶A, N⁶-methyladenosine.

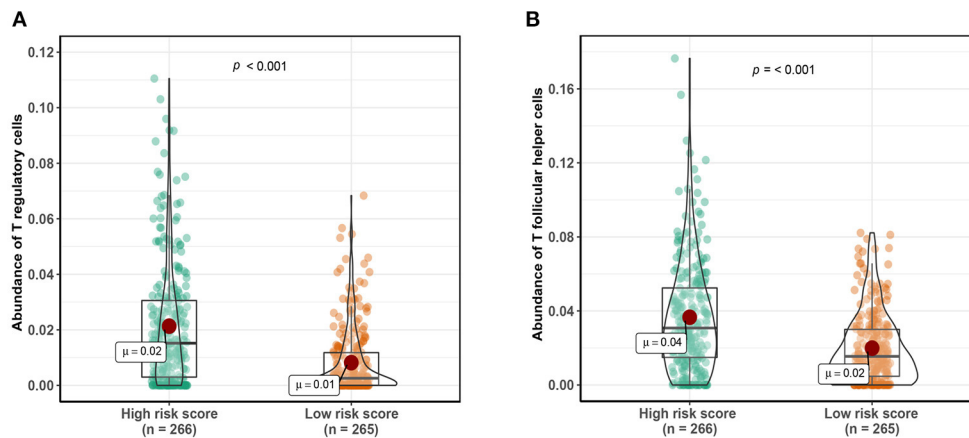


FIGURE 7 | Comparing abundance of T regulatory cells (A) and T follicular helper cells (B) between patients with high-risk scores and low-risk scores. Cutoff value of high- and low-risk scores was defined as median score.

Effective Survival Prediction of Clear-Cell Renal Cell Carcinoma Patients by the m⁶A Risk Score

As shown in **Figure 2A**, the first vertical line pointed at 7 in the TCGA training cohort, which equaled the minimum 10-fold cross-validated error, indicating that seven optional prognosis-related m⁶A regulators, including WTAP, ZC7H13, FTO, ELAVL1, HNRNPA2B1, HNRNPC, and IGF2BP1, were selected by the LASSO Cox regression analysis (**Figure 2B**). A significant difference was observed in the OS [hazard ratio (HR) = 3.474, 95% CI: 2.260–5.342, $p < 0.0001$] between the ccRCC patients with high and low m⁶A risk score in the TCGA training cohort (**Figure 2C**), which was further verified in the TCGA testing cohort (**Figure 2D**) with HR value of 1.679 (95% CI: 1.113–2.532, $p = 0.015$) for the OS and HR value of 2.101 (95% CI: 1.186–3.724, $p = 0.001$) for the disease-free survival in the independent GSE22541 cohort (**Figure 2E**). A higher m⁶A risk score was correlated with the higher tumor stage, higher tumor grade, and death (**Figure 2F**).

The m⁶A risk score was identified as a risk factor for the OS in ccRCC patients by the univariate Cox regression analysis (**Figure 3A**), which was further verified in both the TCGA training cohort (**Figure 3B**) and the independent GSE22541 cohort (**Figure 3C**). There existed significant differences of the m⁶A risk score among different tumor grades ($P < 0.0001$, **Figure 3D**), different tumor stages ($P < 0.0001$, **Figure 3E**), different tumor lymph node metastasis status ($P = 0.06$, **Figure 3F**), and different tumor distant metastasis status ($P < 0.0001$, **Figure 3G**).

Prognostic Accuracy of the m⁶A Risk Score Integrated With the Clinicopathologic Factors

To improve the accuracy of our survival prognostic model, we developed an integrated nomogram by combining the m⁶A risk score and predictable clinicopathologic factors, including age, tumor grade, and tumor stage. The integration nomograms

for the OS prediction are shown in **Figure 4A**. The calibration plots revealed that the 1-, 3-, and 5-year OS probabilities predicted by the integrated nomogram model had an excellent agreement with the actual observations (**Figure 4B**), indicating good performance in predicting the survival status of the ccRCC patients. The decision curve analysis illustrated that when the threshold probability was more than 0.2, the integrated nomogram for the OS prediction could be more favorable than either the m⁶A risk score and the predictable clinicopathologic factors (**Figure 4C**). Further time-dependent ROC curve revealed that the AUC values of the integration nomogram for the OS prediction in 1, 3, and 5 years arrived at 85.2, 82.4, and 78.3%, respectively (**Figure 4D**).

Immune-Related Pathways Were Associated With the m⁶A RNA Methylation Risk Signature

A total of 1,152 DEGs were identified in the GSE40435 cohort, GSE53757 cohort, and GSE46699 cohort (**Figure 5A**), including 577 upregulated genes and 570 downregulated genes (**Figure 5B**). The DEGs were hierarchically clustered into seven gene modules by the WGCNA method (**Figure 5C**), and the brown model was found to perform the highest correlation to the m⁶A RNA methylation risk signature (**Figure 5D**). Significantly different expressions of 128 genes in the brown model were observed between the patients with high and low m⁶A risk scores (**Figure 6A**). As shown in **Figure 6B**, some immune-related pathways, including the immune response regulating pathway, the cytokine-mediated pathway, and the lymphocyte activation-associated pathway, were enriched in genes from the brown model, suggesting that m⁶A RNA methylation might affect clinical prognosis through regulating immune functions of ccRCC patients. Further correlation analysis revealed that the m⁶A risk score was significantly correlated with the abundances of T regulatory cells (Tregs) and T follicular helper cells (Tfh) (**Figure 6C**). Higher abundances of the Tregs and the Tfh were found in the patients with high m⁶A risk scores (**Figures 7A,B**).

DISCUSSION

Due to the high incidence and mortality of ccRCC, the accurate diagnosis and survival prediction of ccRCC patients were urgently needed. Thanks to the innovative developments in high-throughput genetic diagnostic techniques for oncology, there has been a revolutionary improvement in the efficient diagnosis of a malignant tumor at the molecular level (Cancer Genome Atlas Research Network, 2013; Linehan et al., 2016; Zehir et al., 2017).

In this study, we designed a diagnostic model for ccRCC based on the m⁶A regulators. The ccRCC and normal renal tissue could be accurately told out by the m⁶A diagnostic score in Meta-cohort, and the result was further verified in the independent GSE-cohort and TCGA-cohort, with the AUC values of 0.924, 0.867, and 0.795, respectively. Effective survival prediction of ccRCC by the m⁶A risk score was also identified in the TCGA training cohort and validated in the testing cohort and the independent GSE22541 cohort, with the HR values of 3.474, 1.679, and 2.101 for clinical survival, respectively. It was proved that the m⁶A risk score in ccRCC was associated with higher tumor stage, higher tumor grade, and tumor metastasis. All of these results suggested the crucial roles of m⁶A regulators in regulating tumorigenesis and tumor progression, as previously reported (He et al., 2019; Wang et al., 2019; Tian et al., 2020).

An integrated nomogram by combining m⁶A risk score and predictable clinicopathologic factors, including age, tumor grade, and tumor stage, was constructed in this study to improve the accuracy of our survival prognostic model. The time-dependent ROC curve revealed that the AUC values of the integration nomogram for OS prediction in 1, 3, and 5 years arrived at 85.2, 82.4, and 78.3%, respectively, indicating advantageous usability of our survival prediction model in clinical practice.

Different types of m⁶A RNA methylation regulators work differently in tumorigenesis. For example, FTO could promote the progression of lung carcinoma by releasing the m⁶A modification in MZF1 mRNA and strengthening its stability (Liu et al., 2018). However, METTL14 was found to suppress RCC by downregulating P2RX6 protein translation (Gong et al., 2019). Generally, writers could irritate m⁶A modifications in the mRNA of tumor suppressor genes or oncogenes. And then, these modifications could be recognized by readers, resulting in downregulating tumor suppressor or upregulating oncogene expression (Wang et al., 2020).

A total of 128 DEGs associated with the m⁶A RNA methylation risk signature in the brown model were enriched in immune regulated pathways, including the immune response regulating pathway, the cytokine-mediated pathway, and the lymphocyte activation-related pathway. In addition, the m⁶A risk score was significantly correlated with the abundance of Tregs and Tfh. The important role of m⁶A modification in immune cell-related pathogenesis through the m⁶A mediated degradation of Naïve T cell had been recently identified (Li H. B. et al., 2017). Our enrichment results indicated that the m⁶A RNA methylation in ccRCC might affect the prognosis through regulating immune function, especially the functions of Tregs and Tfh.

There are also several defects in this study. Firstly, our main findings were based on integrated bioinformatics analyses, without experimental verification. The experiment verification at the cellular level, including regulation mechanism research and function verification, are still needed. Secondly, all of the data sets analyzed in this study were acquired from the public database, resulting in a potential bias in genetic and clinical data. However, cross-validation among independent datasets has been performed as much as possible to reduce the potential bias. Thirdly, the cutoff value of the high and low m⁶A risk score group was defined as the median score of respective cohorts. Actually, the optional cutoff value of m⁶A risk score to distinguish the patients with high survival risk is supposed to be identified and verified in larger clinical patient cohorts. Finally, our work still requires further clinical validation to verify its application to the clinic in ccRCC patients.

CONCLUSIONS

In summary, we identified the important role of the m⁶A regulator in ccRCC. Immune-related pathways might be involved in the regulation of ccRCC through the m⁶A RNA methylation. The novel m⁶A gene signature might act as an effective biomarker for prognosis prediction of the ccRCC patients, which still requires further clinical validation and experimental verification.

DATA AVAILABILITY STATEMENT

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

AUTHOR CONTRIBUTIONS

JZ, XW, and SC conceived the study, designed the research, and wrote the paper. SC, EZ, NZ, and TW conducted and analyzed experiments. LJ and XW provided samples. JZ supervised the research. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the National Natural Science Foundation of China (81772705, 81972393, and 82002665).

ACKNOWLEDGMENTS

We thank the TCGA, GEO database for gene expression and clinical data.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fcell.2020.616972/full#supplementary-material>

REFERENCES

- Boccalletto, P., Machnicka, M. A., Purta, E., Piatkowski, P., Baginski, B., Wirecki, T. K., et al. (2018). MODOMICS: a database of RNA modification pathways. 2017 update. *Nucleic Acids Res.* 46, 303–307. doi: 10.1093/nar/gkx1030
- Bokar, J. A., Shambaugh, M. E., Polayes, D., Matera, A. G., and Rottman, F. M. (1997). Purification and cDNA cloning of the AdoMet-binding subunit of the human mRNA (N⁶-adenosine)-methyltransferase. *RNA* 3, 1233–1247.
- Cancer Genome Atlas Research Network (2013). Comprehensive molecular characterization of clear cell renal cell carcinoma. *Nature* 499, 43–49. doi: 10.1038/nature12222
- Dagliesh, G. L., Furge, K., Greenman, C., Chen, L., Bignell, G., Butler, A., et al. (2010). Systematic sequencing of renal carcinoma reveals inactivation of histone modifying genes. *Nature* 463, 360–363. doi: 10.1038/nature08672
- De, P., Otterstatter, M. C., Semenciw, R., Ellison, L. F., Marrett, L. D., and Dryer, D. (2014). Trends in incidence, mortality, and survival for kidney cancer in Canada, 1986–2007. *Cancer Causes Control* 25, 1271–1281. doi: 10.1007/s10552-014-0427-x
- Dominissini, D., Moshitch-Moshkovitz, S., Schwartz, S., Salmon-Divon, M., Ungar, L., Osenberg, S., et al. (2012). Topology of the human and mouse m⁶A RNA methylomes revealed by m⁶A-seq. *Nature* 485, 201–206. doi: 10.1038/nature11112
- Dubin, D. T., and Taylor, R. H. (1975). The methylation state of poly A-containing messenger RNA from cultured hamster cells. *Nucleic Acids Res.* 2, 1653–1668. doi: 10.1093/nar/2.10.1653
- Eckel-Passow, J. E., Serie, D. J., Bot, B. M., Joseph, R. W., Cheville, J. C., and Parker, A. S. (2014). ANKS1B is a smoking-related molecular alteration in clear cell renal cell carcinoma. *BMC Urol.* 14:14. doi: 10.1186/1471-2490-14-14
- Genega, E. M., Ghebremichael, M., Najarian, R., Fu, Y., Wang, Y., Argani, P., et al. (2010). Carbonic anhydrase IX expression in renal neoplasms: correlation with tumor type and grade. *Am. J. Clin. Pathol.* 134, 873–879. doi: 10.1309/AJCP57HJMSLZ
- Gong, D., Zhang, J., Chen, Y., Xu, Y., Ma, J., Hu, G., et al. (2019). The m(6)A-suppressed P2RX6 activation promotes renal cancer cells migration and invasion through ATP-induced Ca(2+) influx modulating ERK1/2 phosphorylation and MMP9 signaling pathway. *J. Exp. Clin. Cancer Res.* 38:233. doi: 10.1186/s13046-019-1223-y
- He, L., Li, H., Wu, A., Peng, Y., Shu, G., and Yin, G. (2019). Functions of N⁶-methyladenosine and its role in cancer. *Mol. Cancer* 18:176. doi: 10.1186/s12943-019-1109-9
- Irizarry, R. A., Bolstad, B. M., Collin, F., Cope, L. M., Hobbs, B., and Speed, T. P. (2003). Summaries of Affymetrix GeneChip probe level data. *Nucleic Acids Res.* 31:e15. doi: 10.1093/nar/gng015
- Langfelder, P., and Horvath, S. (2008). WGCNA: an R package for weighted correlation network analysis. *BMC Bioinform.* 9:559. doi: 10.1186/1471-2105-9-559
- Li, A., Chen, Y. S., Ping, X. L., Yang, X., Xiao, W., Yang, Y., et al. (2017). Cytoplasmic m(6)A reader YTHDF3 promotes mRNA translation. *Cell Res.* 27, 444–447. doi: 10.1038/cr.2017.10
- Li, H. B., Tong, J., Zhu, S., Batista, P. J., Duffy, E. E., Zhao, J., et al. (2017). m(6)A mRNA methylation controls T cell homeostasis by targeting the IL-7/STAT5/SOCS pathways. *Nature* 548, 338–342. doi: 10.1038/nature23450
- Linehan, W. M., Spellman, P. T., Ricketts, C. J., Creighton, C. J., Fei, S. S., Davis, C., et al. (2016). Comprehensive molecular characterization of papillary renal-cell carcinoma. *N. Engl. J. Med.* 374, 135–145. doi: 10.1056/NEJMoa1505917
- Liu, J., Ren, D., Du, Z., Wang, H., Zhang, H., and Jin, Y. (2018). m(6)A demethylase FTO facilitates tumor progression in lung squamous cell carcinoma by regulating MZF1 expression. *Biochem. Biophys. Res. Commun.* 502, 456–464. doi: 10.1016/j.bbrc.2018.05.175
- Ljungberg, B., Bensalah, K., Canfield, S., Dabestani, S., Hofmann, F., Hora, M., et al. (2015). EAU guidelines on renal cell carcinoma: 2014 update. *Eur. Urol.* 67, 913–924. doi: 10.1016/j.eururo.2015.01.005
- Reuter, V. E. (2006). The pathology of renal epithelial neoplasms. *Semin. Oncol.* 33, 534–543. doi: 10.1053/j.seminoncol.2006.06.009
- Roundtree, I. A., Evans, M. E., Pan, T., and He, C. (2017). Dynamic RNA modifications in gene expression regulation. *Cell* 169, 1187–1200. doi: 10.1016/j.cell.2017.05.045
- Siegel, R. L., Miller, K. D., and Jemal, A. (2020). Cancer statistics, 2020. *CA Cancer J. Clin.* 70, 7–30. doi: 10.3322/caac.21590
- Strick, A., von Hagen, F., Gundert, L., Klümper, N., Tolkach, Y., Schmidt, D., et al. (2020). The N(6)-methyladenosine (m(6)A) erasers alkylolation repair homologue 5 (ALKBH5) and fat mass and obesity-associated protein (FTO) are prognostic biomarkers in patients with clear cell renal carcinoma. *BJU Int.* 125, 617–624. doi: 10.1111/bju.15019
- Tian, J., Zhu, Y., Rao, M., Cai, Y., Lu, Z., Zou, D., et al. (2020). N(6)-methyladenosine mRNA methylation of PIK3CB regulates AKT signalling to promote PTEN-deficient pancreatic cancer progression. *Gut* 69, 2180–2192. doi: 10.1136/gutjnl-2019-320179
- Varela, I., Tarpey, P., Raine, K., Huang, D., Ong, C. K., Stephens, P., et al. (2011). Exome sequencing identifies frequent mutation of the SWI/SNF complex gene PBRM1 in renal carcinoma. *Nature* 469, 539–542. doi: 10.1038/nature09639
- von Roemeling, C. A., Radisky, D. C., Marlow, L. A., Cooper, S. J., Grebe, S. K., Anastasiadis, P. Z., et al. (2014). Neuronal pentraxin 2 supports clear cell renal cell carcinoma by activating the AMPA-selective glutamate receptor-4. *Cancer Res.* 74, 4796–4810. doi: 10.1158/0008-5472.CAN-14-0210
- Wang, Q., Chen, C., Ding, Q., Zhao, Y., Wang, Z., Chen, J., et al. (2019). METTL3-mediated m(6)A modification of HDGF mRNA promotes gastric cancer progression and has prognostic significance. *Gut* 69, 1193–1205. doi: 10.1136/gutjnl-2019-319639
- Wang, T., Kong, S., Tao, M., and Ju, S. (2020). The potential role of RNA N⁶-methyladenosine in Cancer progression. *Mol. Cancer* 19:88. doi: 10.1186/s12943-020-01204-7
- Wozniak, M. B., Le Calvez-Kelm, F., Abedi-Ardekani, B., Byrnes, G., Durand, G., Carreira, C., et al. (2013). Integrative genome-wide gene expression profiling of clear cell renal cell carcinoma in Czech Republic and in the United States. *PLoS ONE* 8:e57886. doi: 10.1371/journal.pone.0057886
- Wuttig, D., Zastrow, S., Füssel, S., Toma, M. I., Meinhardt, M., Kalman, K., et al. (2012). CD31, EDNRB and TSPAN7 are promising prognostic markers in clear-cell renal cell carcinoma revealed by genome-wide expression analyses of primary tumors and metastases. *Int. J. Cancer* 131, E693–704. doi: 10.1002/ijc.27419
- Zehir, A., Benayed, R., Shah, R. H., Syed, A., Middha, S., Kim, H. R., et al. (2017). Mutational landscape of metastatic cancer revealed from prospective clinical sequencing of 10,000 patients. *Nat. Med.* 23, 703–713. doi: 10.1038/nm.4333
- Zhou, Y., Zhou, B., Pache, L., Chang, M., Khodabakhshi, A. H., Tanaseichuk, O., et al. (2019). Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat. Commun.* 10:1523. doi: 10.1038/s41467-019-09234-6

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 Chen, Zhang, Zhang, Wang, Jiang, Wang and Zheng. 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.



Alteration of mRNA 5-Methylcytosine Modification in Neurons After OGD/R and Potential Roles in Cell Stress Response and Apoptosis

OPEN ACCESS

Edited by:

Jia Meng,
Xi'an Jiaotong-Liverpool University,
China

Reviewed by:

Zhanyang Yu,
Massachusetts General Hospital
and Harvard Medical School,
United States
Jingting Zhu,
The Research Institute at Nationwide
Children's Hospital, United States

*Correspondence:

Shiqing Feng
sqfeng@tmu.edu.cn
Hengxing Zhou
kevin_zyf@126.com

[†]These authors have contributed
equally to this work

Specialty section:

This article was submitted to
Epigenomics and Epigenetics,
a section of the journal
Frontiers in Genetics

Received: 25 November 2020

Accepted: 04 January 2021

Published: 03 February 2021

Citation:

Jian H, Zhang C, Qi Z, Li X, Lou Y,
Kang Y, Deng W, Lv Y, Wang C,
Wang W, Shang S, Hou M, Zhou H
and Feng S (2021) Alteration
of mRNA 5-Methylcytosine
Modification in Neurons After OGD/R
and Potential Roles in Cell Stress
Response and Apoptosis.
Front. Genet. 12:633681.
doi: 10.3389/fgene.2021.633681

Huan Jian^{1,2†}, Chi Zhang^{3,4†}, ZhangYang Qi^{1,2†}, Xueying Li^{3,4,5†}, Yongfu Lou^{1,2}, Yi Kang^{1,2},
Weimin Deng⁵, Yigang Lv^{1,2}, Chaoyu Wang^{1,2}, Wei Wang^{1,2}, Shenghui Shang^{1,2},
Mengfan Hou^{1,2}, Hengxing Zhou^{1,2,3,4*} and Shiqing Feng^{1,2,3,4*}

¹ Department of Orthopaedics, Tianjin Medical University General Hospital, Tianjin, China, ² International Science and Technology Cooperation Base of Spinal Cord Injury, Tianjin Key Laboratory of Spine and Spinal Cord, Tianjin Medical University General Hospital, Tianjin, China, ³ Department of Orthopaedics, Qilu Hospital, Cheeloo College of Medicine, Shandong University, Jinan, China, ⁴ Shandong University Center for Orthopaedics, Cheeloo College of Medicine, Shandong University, Jinan, China, ⁵ Key Laboratory of Immuno Microenvironment and Disease of the Educational Ministry of China, Department of Immunology, Tianjin Medical University, Tianjin, China

Epigenetic modifications play an important role in central nervous system disorders. As a widespread posttranscriptional RNA modification, the role of the m⁵C modification in cerebral ischemia-reperfusion injury (IRI) remains poorly defined. Here, we successfully constructed a neuronal oxygen-glucose deprivation/reoxygenation (OGD/R) model and obtained an overview of the transcriptome-wide m⁵C profiles using RNA-BS-seq. We discovered that the distribution of neuronal m⁵C modifications was highly conserved, significantly enriched in CG-rich regions and concentrated in the mRNA translation initiation regions. After OGD/R, modification level of m⁵C increased, whereas the number of methylated mRNA genes decreased. The amount of overlap of m⁵C sites with the binding sites of most RNA-binding proteins increased significantly, except for that of the RBM3-binding protein. Moreover, hypermethylated genes in neurons were significantly enriched in pathological processes, and the hub hypermethylated genes RPL8 and RPS9 identified by the protein-protein interaction network were significantly related to cerebral injury. Furthermore, the upregulated transcripts with hypermethylated modification were enriched in the processes involved in response to stress and regulation of apoptosis, and these processes were not identified in hypomethylated transcripts. In final, we verified that OGD/R induced neuronal apoptosis *in vitro* using TUNEL and western blot assays. Our study identified novel m⁵C mRNAs associated with ischemia-reperfusion in neurons, providing valuable perspectives for future studies on the role of the RNA methylation in cerebral IRI.

Keywords: m⁵C methylation, epitranscriptome, ischemia-reperfusion injury, oxygen-glucose deprivation/reoxygenation, cell stress, apoptosis

INTRODUCTION

Posttranscriptional RNA modifications have important biological significance in RNA metabolism. The most widely studied RNA posttranscriptional modification is N⁶-methyladenosine (m⁶A). Moreover, “writers,” “erasers,” and “readers” of the m⁶A modification have been identified. Increasing evidence shows that m⁶A exhibits a wide range of effects on mRNA metabolism, including mRNA stability, translation, and splicing (Cao et al., 2016; Meyer and Jaffrey, 2017; Zhao et al., 2017). N¹-methyladenosine (m¹A), 5-methylcytosine (m⁵C), pseudouridine (Ψ), and other modification types also play an important role in the posttranscriptional regulation of genes (Carlile et al., 2014; Li et al., 2015; Dominissini et al., 2016; Amort et al., 2017; Yang et al., 2017). Cytosine methylation is a new type of RNA modification first identified in tRNA and rRNA (Helm, 2006; Schaefer et al., 2009). Using advanced high-throughput techniques combined with next-generation sequencing (NGS), m⁵C modification sites were also found in many mRNAs. The main regulatory proteins that were identified include the NOP2/Sun RNA Methyltransferase Family Member 2 (NSUN2, known as an m⁵C “writer”) and Aly/REF export factor (ALYREF, known as an m⁵C “reader”; Yang et al., 2017). Existing studies have demonstrated that the m⁵C methylation modification of mRNA can regulate key biological processes, such as mRNA nuclear effects, maintenance of mRNA stability, and neural stem cell differentiation (Flores et al., 2017; Yang et al., 2017; Chen et al., 2019b).

Ischemia-reperfusion injury (IRI) could cause extensive tissue and organ damage, and the core mechanism of IRI is blood vessel remodeling after deprivation of the blood supply to a certain area or organ and blood reperfusion after tissue and organ ischemia (Eltzschig and Eckle, 2011). Current studies have demonstrated that ischemia and reperfusion can cause tissue and organ damage through a variety of pathophysiological mechanisms. The main pathological mechanisms include activation of the complement system, calcium overload, reduction in oxidative phosphorylation, endothelial dysfunction, activation of the apoptotic signaling pathways, and an increase in the free radical concentration (Eltzschig and Eckle, 2011; Duehrkop and Rieben, 2014). In brain tissue, ischemia and hypoxia are closely related to many neurological diseases, such as traumatic brain injury (TBI), acute ischemic stroke, and iatrogenic cardiopulmonary bypass surgery (Wiberg et al., 2016; Lopez et al., 2017; Zhao et al., 2018).

In addition, epigenetic effects play an important role in nerve damage caused by IRI. Studies have demonstrated that miR-424 can enhance the levels of DNA methyltransferase 1 (DNMT1) and histone 3 lysine 27 trimethylation through the NFIA/DNMT1 signaling pathway, thereby preventing astrogliosis after cerebral IRI in mice (Zhao et al., 2019). Simultaneously, after cerebral IRI, the m⁶A RNA demethylase AlkB homolog 5 (ALKBH5) selectively demethylates Bcl2 transcripts, thereby preventing the degradation of Bcl2 transcripts and enhancing Bcl2 protein expression (Xu K et al., 2020). Therefore, epigenetic modifications of nerve tissue may play an important role in the protection of nerve function and microenvironment improvement after injury. However,

under cerebral IRI conditions, the landscape and potential functions of m⁵C modifications of mRNA remain unclear.

Thus, in this study, we aimed to gain a deeper understanding of the m⁵C methylation of neurons after IRI. To this end, we established an oxygen-glucose deprivation/reoxygenation (OGD/R) model with primary cerebral neurons to simulate cerebral IRI *in vivo*. Sequencing of bisulfite-treated RNAs (RNA-BS-seq) was performed to analyze whether the methylation modification of mRNA was significantly altered after OGD/R treatment. The results showed that m⁵C methylation modifications were abundant in neurons. Simultaneously, after OGD/R induction, numerous novel m⁵C sites were identified in neuronal mRNAs that were associated with various disease-related pathways. In addition, we examined the potential connections among the m⁵C methylation sites, protein-binding sites and chromosome distribution. In general, our research provided a comprehensive description of the epigenetic mechanism of m⁵C modification in neurons after OGD/R and afforded basic information for further research on the function and specific mechanism of m⁵C after IRI in neural tissues.

MATERIALS AND METHODS

Primary Cerebral Neuron Isolation and Culture

Murine primary cortical neurons were isolated from embryos of pregnant C57/BL6 mice as previously described (Hilgenberg and Smith, 2007; Sciarretta and Minichiello, 2010). Briefly, the cortex of the embryonic mice was dissected in high-glucose DMEM (Gibco, Cat# 31053028) at 4°C. Then, the cortical tissue was cut into 1-mm³ pieces and centrifuged at 800 rpm for 5 min at room temperature. The tissue was incubated with papain solution (10 U/ml; Worthington, Cat# LS 03126) for 15 min at 37°C in a 5% CO₂ incubator and dissociated into single cells by gentle trituration. Then, single cells were resuspended in DMEM-HG containing 10% fetal bovine serum (Gibco, Cat# 10099-141) and 1% penicillin/streptomycin (Invitrogen, Cat# 15140148). Finally, neurons were plated on poly-L-lysine (Sigma-Aldrich, Cat# P4832)-precoated cell culture dishes at 1.0 × 10⁶ cells/ml. After 4 h, the previous medium was replaced with neurobasal medium (Gibco, Cat# 21103049) containing 2 mmol/L glutamine (Gibco, Cat# 25030081), 1% B-27TM Supplement (50X; Gibco, Cat# 17504044), and 1% P/S (Invitrogen, Cat# 15140148). The neuronal media were changed every 3 days.

Oxygen–Glucose Deprivation/Reoxygenation

Neuronal medium was removed, and neurons were washed with phosphate-buffered saline (PBS; Sigma-Aldrich, Cat# D8537) supplemented with 1% P/S several times. To initiate OGD, neuron cells were cultured in glucose-free DMEM (Gibco, Cat# 31053028). Then, neurons were placed in hypoxic conditions (37°C, 94% N₂, 21% O₂, and 5% CO₂) for 3 h. Then, the glucose-free DMEM was changed to normal neuronal culture medium, and neurons were incubated under normal incubatory

conditions (37°C, 5% CO₂) for 24 h. The culture conditions were simulated for reperfusion. Neurons treated without OGD served as a control group.

RNA Extraction

TRIzol reagent (Invitrogen Corporation, CA, United States) was used to extract RNA from the primary cultured neurons following the manufacturer's protocol, and the NEBNext® rRNA Depletion Kit (New England Biolabs) was used to reduce the rRNA content. The RNA concentration of each sample was measured by NanoDrop ND-1000 (Thermo Fisher Scientific). The ratio of OD260/OD280 value was used to assess the purity of the RNA index. When the OD260/OD280 value ranges from 1.8 to 2.1, RNA is considered pure. Electrophoresis was performed on a denaturing agarose gel to evaluate RNA integrity.

RNA-BS-Seq and Identification of m⁵C Sites

Briefly, RNA was bisulfite-converted and purified using the EZ RNA Methylation Kit (Zymo Research). Then, the TruSeq Stranded Total RNA Library Prep Kit (Illumina) was used to construct the RNA libraries. The library quality was assessed by a BioAnalyzer 2100 system (Agilent Technologies, Inc.). Library sequencing was performed on an Illumina HiSeq instrument with 150-bp paired-end reads. Paired-end reads were quality controlled by Q30, and low-quality reads were removed after 3' adaptor trimming by Cutadapt software (v1.9.3). STAR software was used to match clean reads of the input library to the reference genome (version mm10; Dobin et al., 2013), which were obtained from the UCSC database, and meRanGh software (a component of meRanTK) was used to align the clean reads of the bisulfite-treated library to the reference genome (Rieder et al., 2016). meRanCall was used to extract each methylated cytosine (C) site in the genome, and meRanCompare was used for the identification of differentially methylated sites. Finally, the m⁵C sites were considered credible with an m⁵C methylation level ≥ 0.1 and a coverage depth ≥ 10 .

Analysis of the m⁵C Distribution

According to the provided method, the distribution map of the m⁵C locus on the chromosome was drawn according to the m⁵C locus information (Hao et al., 2020). The m⁵C locus information was annotated using BEDTools (Quinlan and Hall, 2010), and m⁵C sites in the mRNA were mapped to five regions: 5'UTR, start codon, CDS, stop codon and 3'UTR. In addition, 21-nt sequences proximal to the m⁵C locus on both sides were extracted with BEDTools, and logo plots were generated using WebLogo. The distribution of methylation peaks was plotted using MetaPlotR software (Olarerin-George and Jaffrey, 2017). Profiles of the RNA-binding protein (RBP) binding sites were downloaded from POSTAR (Hu et al., 2017), and the available binding sites were combined into one BED file and used to analyze the overlap with the m⁵C sites.

Bioinformatics Analysis

To explore the function of the m⁵C modification after OGD/R, the m⁵C sites were selected according to the criteria of the m⁵C methylation level ≥ 0.1 and coverage depth ≥ 10 in all three repetitions. As a comprehensive and excellent biometric analysis website (Zhou et al., 2019), Metascape was used to perform the enrichment analysis of hypermethylated and hypomethylated genes using the following ontology sources: Gene Ontology (GO) Biological Processes and Kyoto Encyclopedia of Genes and Genomes (KEGG) Pathway. Terms with a minimum overlap of 3, p -value < 0.01 , and enrichment factor > 1.5 were considered. Next, we used the STRING database to perform protein-protein interaction (PPI) network analysis of differentially methylated genes and the Molecular Complex Detection (MCODE) plug-in to analyze closely connected network components. Differentially expressed genes (DEGs; $|FC| > 1.5$, $p < 0.01$) methylated by m⁵C were used to further analyze the effect of m⁵C methylation. GO functional analysis was performed using the GO website, and terms with p -values < 0.05 were considered statistically significant. Moreover, KEGG pathway analysis was performed by DAVID (Huang da et al., 2009), and pathways with a p -value < 0.05 were considered to be significant. Gene set enrichment analysis (GSEA) was analyzed by the biological process items, and the mouse gene set data were downloaded from GSKB¹.

Methylated RNA Immunoprecipitation and Quantitative Reverse-Transcription PCR

Methylated RNA immunoprecipitation (MeRIP) was performed by Cloudseq Biotech, Inc. (Shanghai, China). Briefly, RNA was fragmented by incubation with fragmentation buffer at 94°C for 4 min. Stop solution (0.05 M EDTA) was added to stop fragmentation. The anti-m⁵C antibody (Diagenode, Cat# C15200081-100) and protein G beads were incubated with the fragmented RNA in IP buffer for 2 h on a rotating wheel at 4°C to reduce nonspecific binding. After the beads settled to the bottom of the tube, the supernatant was discarded, and IP buffer was added to resuspend the magnetic beads. This procedure was repeated twice, and the tube was placed on ice. The eluent was added and incubated at 4°C for 1 h, and clear supernatant was collected and purified with an RNase MiniElute Kit. The eluted sample was used to obtain RNA after MeRIP. RNA after MeRIP was analyzed by Quantitative Reverse-Transcription PCR (qRT-PCR) together with the corresponding input RNA. The primers used for qRT-PCR analysis are presented in **Table 1**.

Immunocytochemistry

The cultured neurons were fixed with 4% paraformaldehyde for 10 min at room temperature. Then, 0.2% Triton X-100 was used to penetrate the membrane of the neurons, and the neurons were incubated with 10% goat serum (Solarbio, China) for 30 min. Next, neurons were incubated with primary antibodies overnight

¹<http://ge-lab.org/gskb/>

TABLE 1 | MeRIP qRT-PCR primer information.

Name	Primer sequence (5'-3')	
Ngp	Forward: AAGGGGCCAAGAGTGGTAGT	Reverse: TAGTTGTGCAAGGGCCTCAC
Hbb-bs	Forward: TGCACCTGACTGATGCTGAG	Reverse: ACTTCATCGGCGTTCACCTT
Anxa1	Forward: GAGTCTCTCTTCAGTCCCCG	Reverse: GAAACGGGCGCTGCTTGAGG
Glr5	Forward: GCTCTGTAAGCCCTGGAGTG	Reverse: CCACCTTGTCCTTCTCACCA
Dpp4	Forward: AGAGAAGAGGGAGCAGGGAG	Reverse: AGTCTGGCAGTGAACAGCTC
Myo7a	Forward: GGCTCGGAGGAAGAAGGAAC	Reverse: TCCCCAGGAAGCCAAACATC
Prr13	Forward: GTGCGAACCAGACTGAGAA	Reverse: GGAGGCCTTTAAGCATCCGT
Cyba	Forward: AGTGAGGACTTGCGAAGTGG	Reverse: TGTGTGAAACGTCCAGCAGT
Arhgap12	Forward: TGGCCGAGAGAAGTGGAAAG	Reverse: CAGGTCTGACTTGCCACCAG
Ywhah	Forward: ATGGGGGATCGAGAGCAG	Reverse: GGAGGCCATATCGTCGTAGC
Acta2	Forward: AAGAGGAAGACAGCACAGCC	Reverse: GGAGCATCATCACCAGCGAA
Heph	Forward: TTTGCCCTACCAGCTCAGTG	Reverse: TACACACTTCCTTTGCCCCG

at 4°C [anti- β -III tubulin (1:400, Abcam Cat# ab78078) and anti-GFAP (1:400, Abcam Cat# ab7260)]. The next day, the neurons were incubated with the secondary antibody for 1 h at room temperature. Finally, cells were incubated with DAPI (Beyotime Biotechnology, China) for 10 min, and images were taken on a fluorescence microscope (Olympus, Japan).

Western Blot

After the cells were washed with PBS to remove the influence of other protein substances in the medium, RIPA lysate (Solarbio, China) containing protease inhibitors was added to completely lyse the cells at 4°C for 30 min followed by centrifugation at $13,000 \times g$ at 4°C for 10 min. The supernatant was collected, and the Pierce™ BCA Protein Assay Kit (Thermo Scientific) was used to determine the protein concentration. Finally, the proteins were heated in a metal heater at 100°C for 10 min for denaturation. 25 micrograms of protein were separated by 12% SDS-PAGE and transferred to PVDF membranes. Membranes were blocked with 5% skim milk at room temperature for 1 h and then incubated with primary antibodies on a shaker at 4°C overnight using the following antibodies: Anti- β -Actin (1:10000; MBL International, Cat# JM-3598R-100, RRID:AB_2784536), Anti-Caspase-3 (1:1000; Cell Signaling Technology, Cat# 9662, RRID:AB_331439), Anti-Bcl-2 (1:1000; Cell Signaling Technology, Cat# 2870, RRID:AB_2290370), and Anti-Bax (1:1000; Cell Signaling Technology, Cat# 2772, RRID:AB_10695870). After washing in TBST solution contained 0.1% Tween, the immunoblot bands were visualized by using horseradish peroxidase-linked anti-rabbit IgG (1:3000; Cell Signaling Technology, Cat# 7074S, RRID:AB_2099233), and ImageJ software was used for grayscale value analysis.

Statistical Analysis

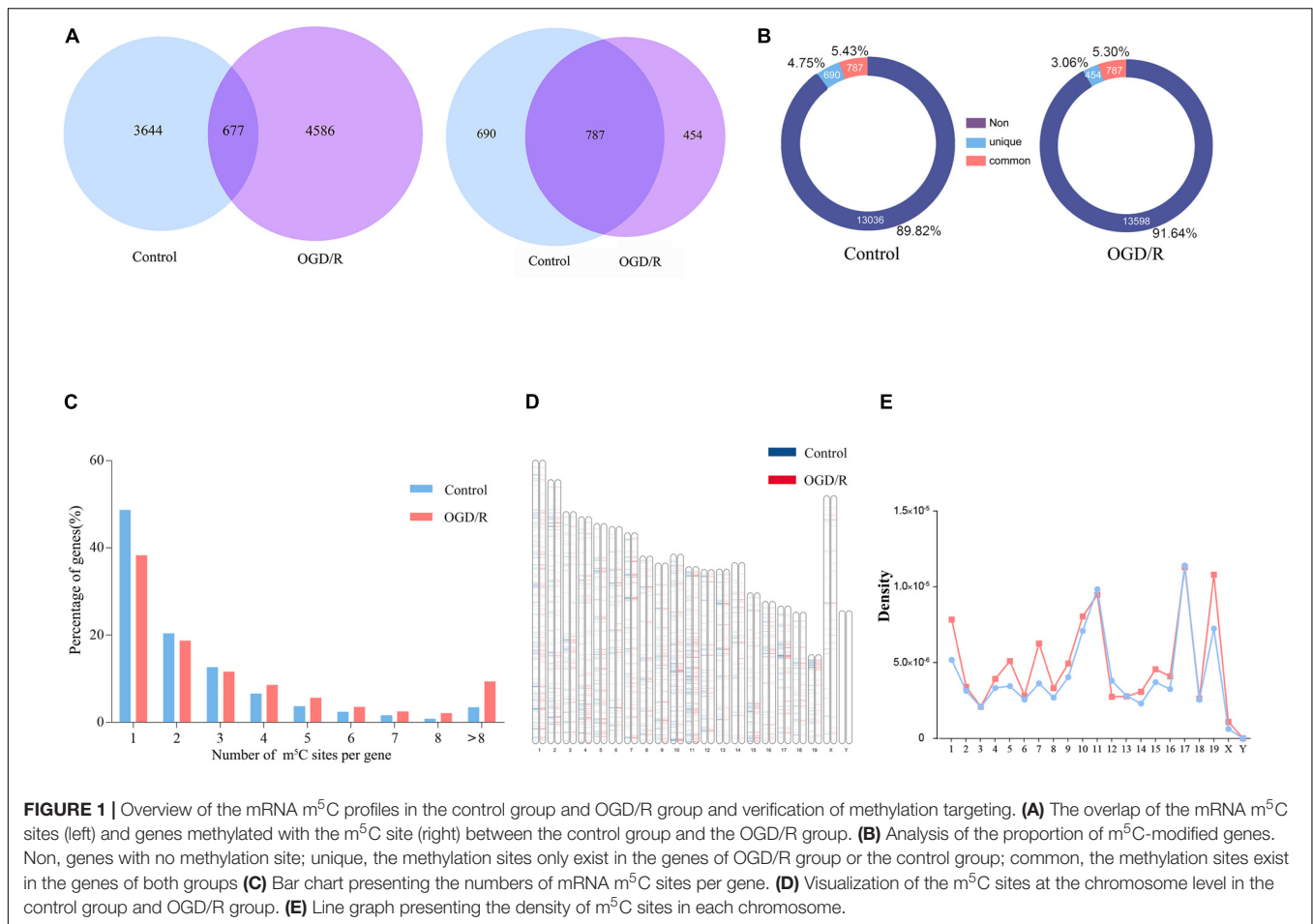
Statistical analysis related to bioinformatics were performed using R software package (unless the method described otherwise) for statistical calculation. The experimental data was analyzed by unpaired *t*-test using GraphPad Prism 8 software, the data were expressed as mean \pm standard deviation (SD) and $P < 0.05$ was considered statistically significant.

RESULTS

mRNA m⁵C Profiling of Neurons and Verification of Methylation Positions

Bisulfite sequencing of neuronal RNAs was performed to obtain an overview of the transcriptome-wide m⁵C profiles (Supplementary Table 1). In general, we discovered 4321 methylation sites in the control group and 5263 methylation sites in the OGD/R group, and most recognition sites were specific in the two groups, corresponding to 84.33% of the control group and 87.14% of the OGD/R group (Figure 1A and Supplementary Figure 1A). Mapping of methylation sites to the mouse mm10 genome showed that the m⁵C sites were identified in neurons located in 1477 (control) and 1241 (OGD/R) annotated genes (Figure 1A), and the methylated genes accounted for 10.18% (control) and 8.36% (OGD/R) of genes with expression identified in neurons, respectively, (Figure 1B). Interestingly, we found that 48.6% of total methylated mRNAs had one m⁵C site in the control group, whereas this proportion was reduced to 38.3% in the OGD/R group ($p = 7.36 \times 10^{-5}$). In addition, mRNAs with more than eight m⁵C sites accounted for 3.4% (control) and 9.3% (OGD/R) of total mRNAs, respectively, ($p = 3.36 \times 10^{-5}$). Compared with that in the control group, the number of m⁵C sites per mRNA in the OGD/R group increased significantly (Figure 1C). The graphics of the methylation sites on the chromosomes were visualized using the R software package according to the reported data and code, indicating that the distribution of m⁵C sites on each chromosome was different between the two groups, especially on chromosomes one, five, seven and nineteen (Figures 1D,E). Moreover, few m⁵C sites were located on sex chromosome X, and no m⁵C sites were noted on chromosome Y in both groups, which seems to be related to the incidence of tissue specificity (Figures 1D,E).

To verify the accuracy of the RNA-BS-seq results, we selected 10 candidate transcripts divided into hypermethylated group and hypomethylated group according to the sequencing data (Supplementary Figure 2), and the m⁵C methylated state was confirmed by methylated RNA immunoprecipitation (MeRIP; Figure 2A). Among the 10 candidate transcripts analyzed, 5



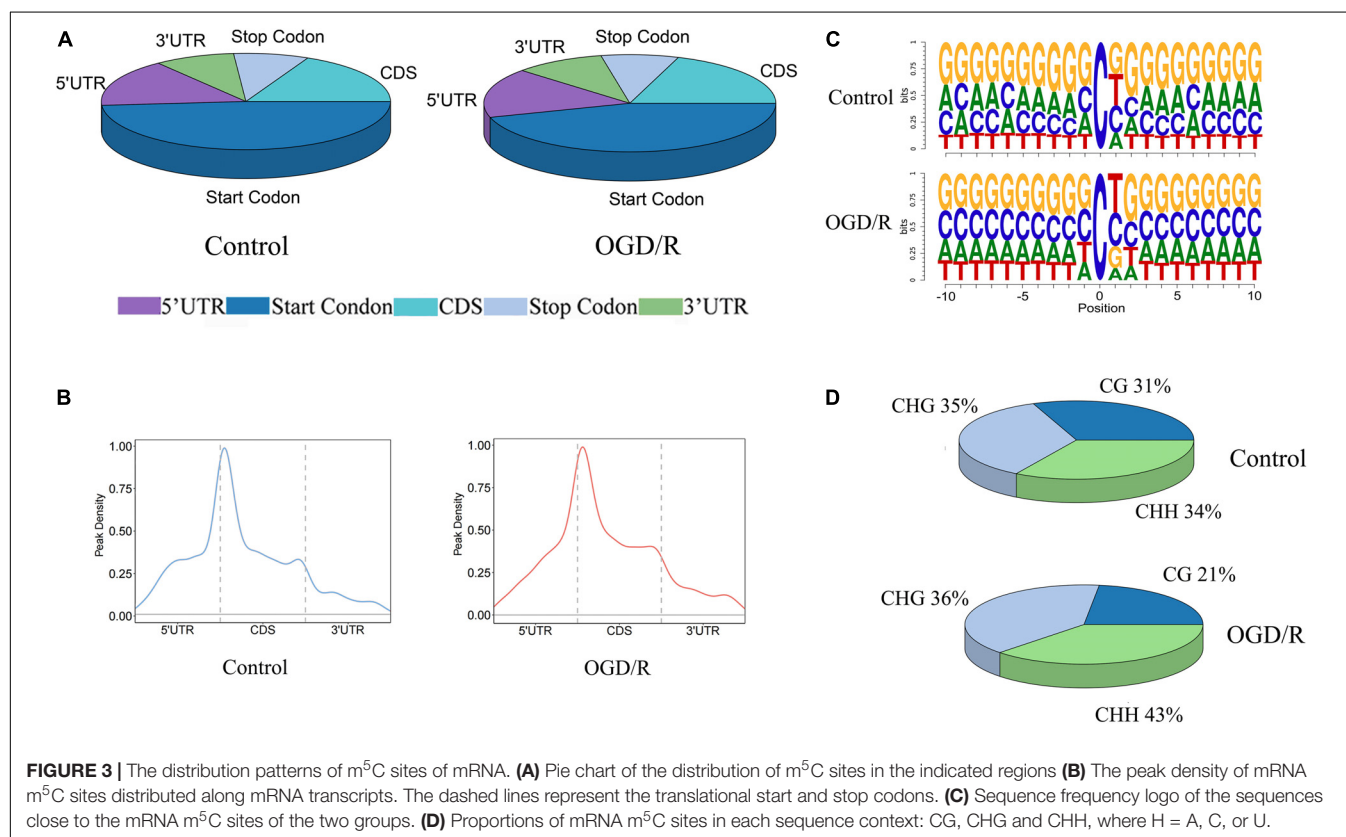
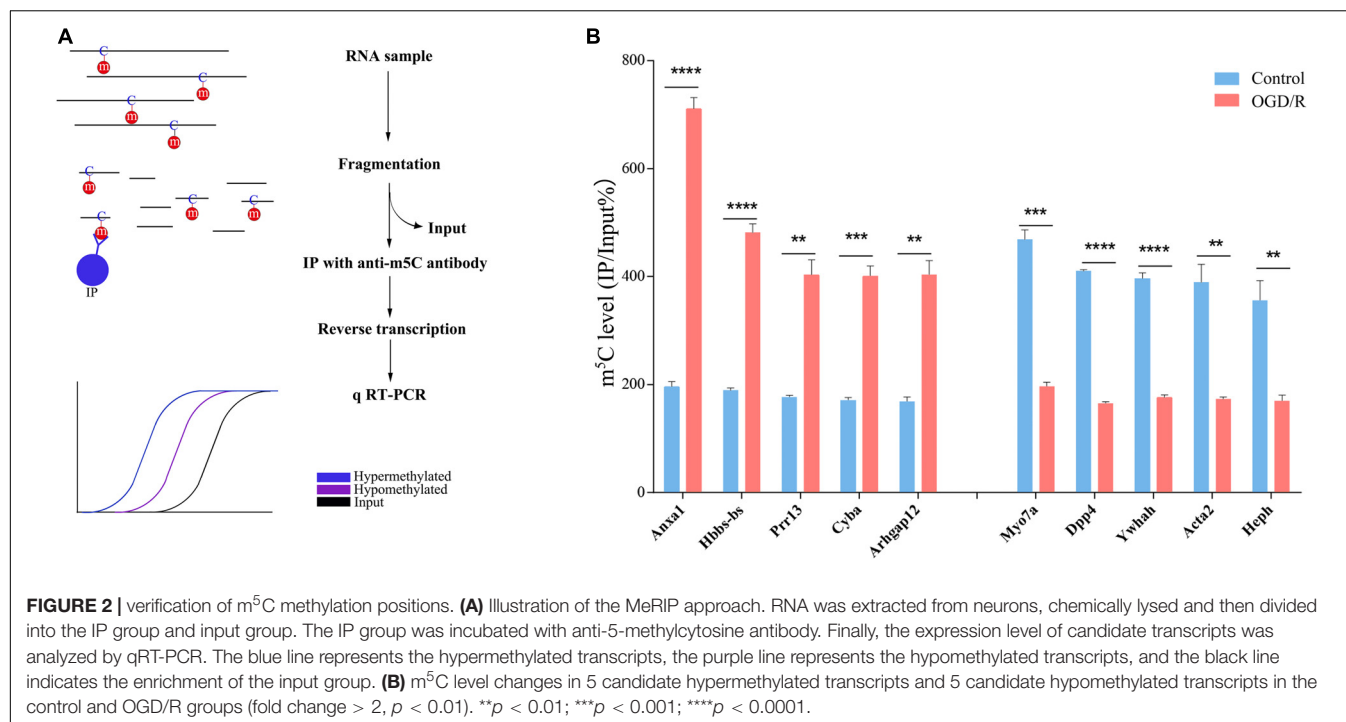
transcripts were hypermethylated (Anxal, Hbbs-bs, Prr15, Cyba, and Arbgap12), and 5 transcripts were hypomethylated after OGD/R (Myo7a, Dpp4, Ywhah, Acta2, and Heph). The MeRIP verification results are completely consistent with the sequencing results; thus, we are confident that our data represent a reliable image of m⁵C modification of neuronal transcripts (Figure 2B).

Common and Distinct Distribution Features of mRNA m⁵C Sites Identified in Neurons Before and After OGD/R Treatment

To further understand the distribution of m⁵C methylation on mRNA, we separately analyzed the distribution of m⁵C sites in the control group and OGD/R group. The positions of m⁵C were divided into five different regions according to their locations in transcripts. The descending order of the degree of m⁵C modification was start codon, coding sequence (CDS), 5'-untranslated region, stop codon and 3'-untranslated region (Figure 3A). It's worth noting that the distribution patterns of m⁵C sites in the control and OGD/R groups were similar. In addition, further analyses indicated that the m⁵C sites were mainly enriched in regions immediately downstream of translation initiation sites in both groups (Figure 3B). Similarly,

a sequence frequency logo showed that m⁵C methylation was embedded in CG-rich environments regardless of whether it had undergone OGD/R treatment (Figure 3C). Moreover, the methylation sites are distributed in three types of special sequence contexts (CG, CHG and CHH, where H = A, C, or U). After OGD/R treatment, the number of CG regions significantly decreased from 31 to 21%, while the number of CHH regions markedly increased from 34 to 43% (Figure 3D).

We next jointly analyzed the mRNA m⁵C site profile and the binding sites of RBPs in the POSTAR database to determine the relationship between the methylation sites and protein-binding sites (Supplementary Table 2). According to the transcriptome sequencing data, 24 RBPs with higher gene expression level were selected, and approximately 18.21% of m⁵C sites in the control group and 22.11% of OGD/R group m⁵C sites overlapped with the mapped RBP binding sites. In particular, RBFOX1, RBFOX2, RBFOX3, APC, CELF4, TARDBP, U2AF2, FUS, and FMR1 correlated with brain tissue (Figures 4A,B). After OGD/R treatment, the overlap probability of m⁵C sites with the binding sites of most RBPs increased, except for the RBM3 binding protein (Figures 4C,D). The analysis showed that OGD/R treatment obviously changed the overlapping probability of RBP-binding sites with m⁵C sites, indicating that m⁵C methylation may indirectly regulate gene function by interacting with RBPs.



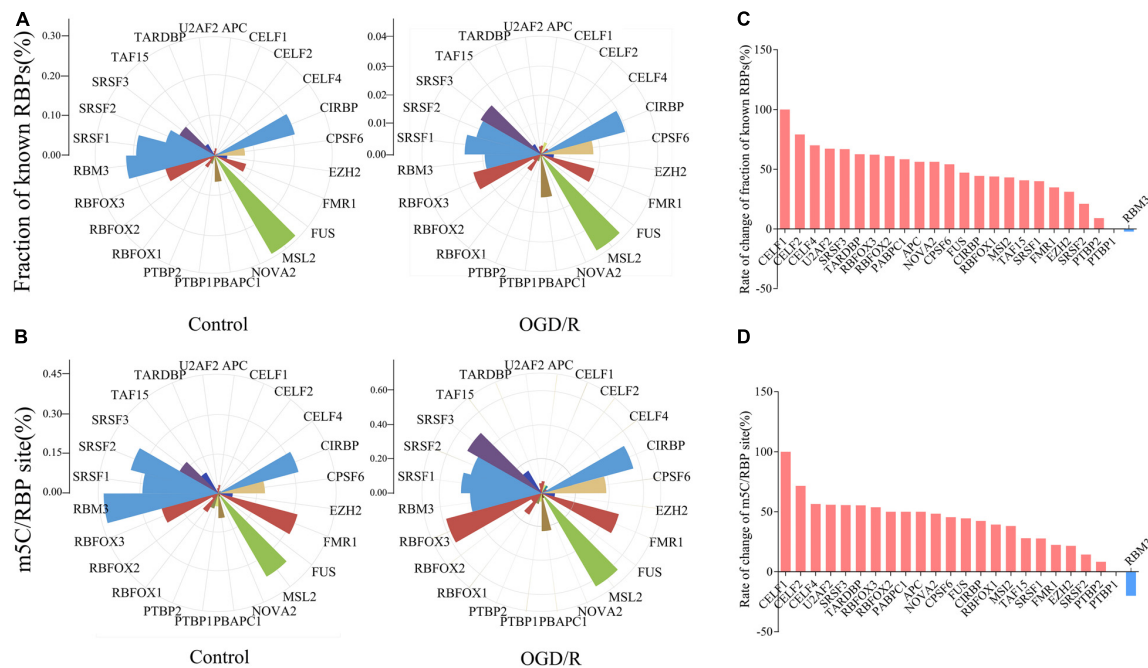


FIGURE 4 | Rose diagrams depict an overlap of mRNA m⁵C sites with RBP-binding sites available in POSTAR. **(A)** Fraction of binding sites overlapping with a m⁵C site for each particular RBP. Fraction of known RBPs = RBP binding sites overlapping with a m⁵C site /RBP binding sites. **(B)** The number of m⁵C sites overlapping with a binding site of a particular protein was normalized against the total number of m⁵C sites. m⁵C/RBP site = m⁵C sites overlapping with a RBP binding site /RBP binding sites. **(C,D)** Histograms present the difference between two groups. The result is the difference between the number of binding sites in the OGD/R group and the control group divided by the number of sites in the OGD/R group. **(C)**, Fraction of known RBPs; **D**, m⁵C/RBP.

Functional Analysis of Differentially Methylated m⁵C-Modified Genes Between the Control and OGD/R Groups

Previous studies have noted that after cerebral IRI, the m⁶A RNA demethylase Alkbh5 can selectively prevent the degradation of the Bcl-2 protein, which may play an important role in reducing nerve apoptosis and protecting nerve function (Xu K et al., 2020). The m⁵C site has been confirmed to be abundant in brain tissue (Amort et al., 2017), whereas the role of the m⁵C RNA modification in cerebral IRI remains unclear. By analyzing the methylation level of methylated transcripts, we strictly filtered the methylation site information (**Supplementary Table 3**) and identified 480 hypermethylated transcripts and 382 hypomethylated transcripts in neurons after OGD/R (**Figure 5A** and **Supplementary Table 4**). To understand the function of differentially methylated genes, we conducted Metascape analyses and found that the hypermethylated genes were mainly enriched in pathological processes, such as Huntington's disease and intrinsic apoptotic signaling pathways, whereas hypomethylated genes were mainly enriched in physiological processes (**Figures 5B,C**). To further confirm the biological function of neurons between the control and OGD/R groups, we performed GSEA and focused on important biological functions, including response to endoplasmic reticulum stress, apoptotic process, negative regulation of cell proliferation, and cell migration (**Figure 5D**). The results indicated that m⁵C modification was more abundant in cellular stress and cell

death-related gene sets after OGD/R treatment, which seemed to be closely related to cerebral injury.

In addition, the PPI network showed that HSPA8, PPP2CA, RPS9, RPL8, RPS14, RAB1B, RPS2, SRSF1, GNB2L1, and RPL31 were the top 10 high-degree nodes of hypermethylated genes, which may play important roles in the pathological processes of the brain (**Figure 5E**). Heat shock protein family A (Hsp70) member 8 (HSPA8) played an important role in the occurrence and development of neurological diseases, such as Alzheimer's disease and Parkinson's disease (Lauterbach, 2013; Silva et al., 2014). Ribosomal protein L8 (RPL8) expression levels were significantly increased in the brain tissue after cerebral hemorrhage, and Ribosomal protein S9 (RPS9) expression was increased significantly in the neurons of the dentate gyrus after acute cerebral ischemia (Kim et al., 2003; Chen et al., 2019a). By analyzing the functions of the main gene modules through MCODE, we found that the main methylation gene modules were significantly related to basic life activities of cells, including gene expression, biosynthetic processes, energy metabolism and transport (**Figure 5F**), and these results were similar to previous studies (Yang et al., 2019).

Integration Analyses of m⁵C-Containing mRNA Methylation and mRNA Transcript Expression

The principal component analysis (PCA) results of transcripts indicated significant differences between the gene expression

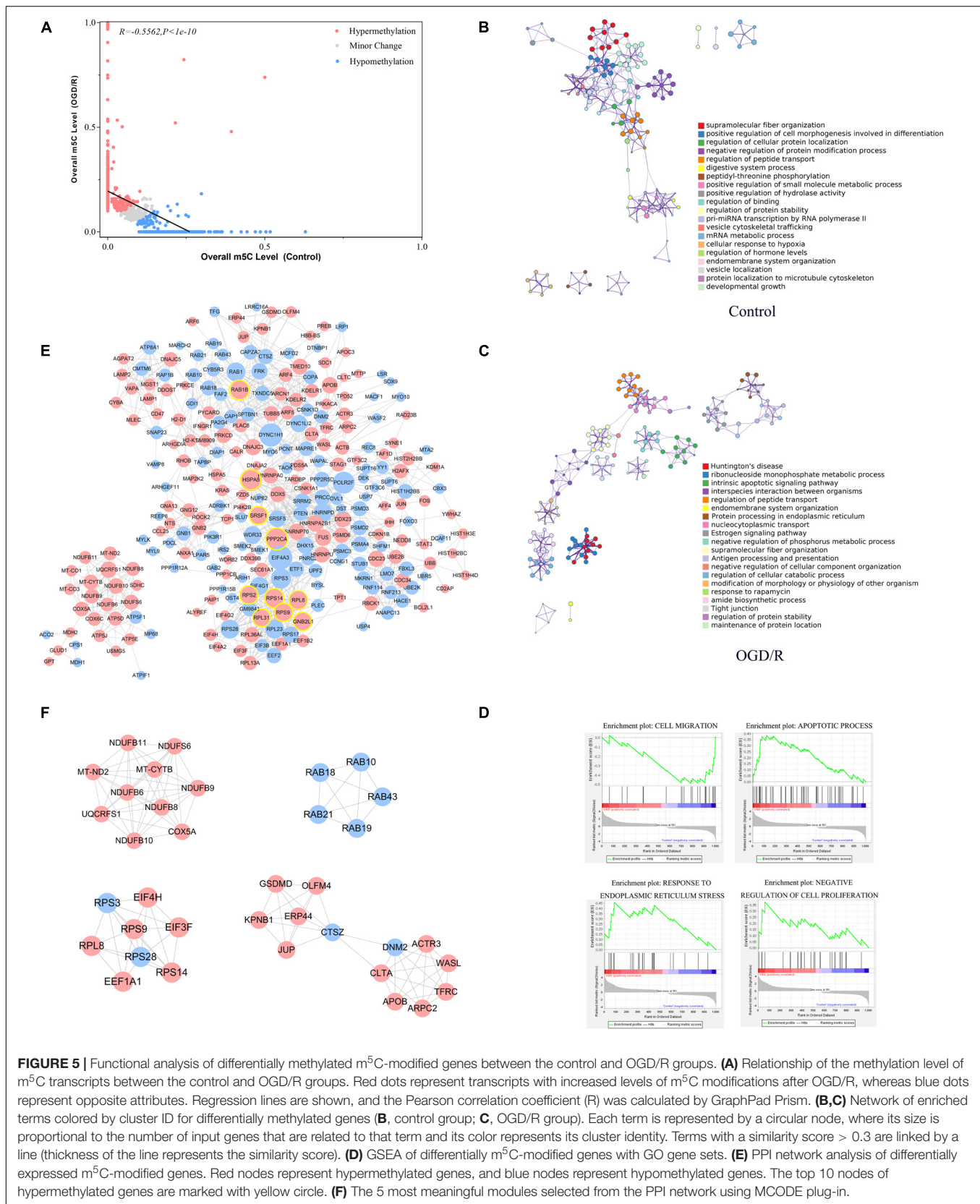


FIGURE 5 | Functional analysis of differentially methylated m⁵C-modified genes between the control and OGD/R groups. **(A)** Relationship of the methylation level of m⁵C transcripts between the control and OGD/R groups. Red dots represent transcripts with increased levels of m⁵C modifications after OGD/R, whereas blue dots represent opposite attributes. Regression lines are shown, and the Pearson correlation coefficient (*R*) was calculated by GraphPad Prism. **(B,C)** Network of enriched terms colored by cluster ID for differentially methylated genes (**B**, control group; **C**, OGD/R group). Each term is represented by a circular node, where its size is proportional to the number of input genes that are related to that term and its color represents its cluster identity. Terms with a similarity score > 0.3 are linked by a line (thickness of the line represents the similarity score). **(D)** GSEA of differentially m⁵C-modified genes with GO gene sets. **(E)** PPI network analysis of differentially expressed m⁵C-modified genes. Red nodes represent hypermethylated genes, and blue nodes represent hypomethylated genes. The top 10 nodes of hypermethylated genes are marked with yellow circle. **(F)** The 5 most meaningful modules selected from the PPI network using MCODE plug-in.

patterns of the two groups, whereas the expression patterns were similar within the same group (**Supplementary Figure 3A**). GO analysis showed that up-regulated DEGs after OGD/R were enriched in specific biological processes (e.g., sensory perception of chemical stimulus, cellular response to stress, and regulation of cell death), molecular functions (e.g., transmembrane signaling receptor activity), and cellular components (e.g., nucleus, non-membrane-bounded organelle, and organelle lumen). However, down-regulated DEGs were enriched in the items closely related to essential neuronal processes and components, including neurogenesis, neuron projection development, generation of neurons, neuron differentiations, axon, and dendritic tree (**Supplementary Figure 3B**). In addition, KEGG pathway analysis indicated that down-regulated DEGs were generally linked to basic neuronal development pathways, whereas up-regulated DEGs were interestingly enriched in some cancer-related pathways (**Supplementary Figure 3C**).

Next, we integrated RNA-Seq data and RNA-BS-seq data to co-analyze the m⁵C-methylated DEGs for the further exploration of the function of m⁵C modification (**Supplementary Table 5**). GO terms of the control group were mainly enriched in physiological processes, such as substance transport, cell location and adhesion signaling, whereas GO terms in the OGD/R group were mainly enriched in the cell stress response and cell death processes (**Figure 6A**). KEGG analysis also showed that genes upregulated after OGD/R were mainly enriched in neurological disease pathways, such as Huntington's disease, Parkinson's disease and Alzheimer's disease, whereas downregulated DEGs were mainly enriched in normal physiological pathways (**Figure 6B**). Finally, the GSEA results indicated that gene sets of apoptotic processes and response to stress were significantly up-regulated (**Figure 6C**). Moreover, our analysis revealed that the level of methylation increased significantly after OGD/R (**Figure 6D**), while the four-quadrant diagram indicated that a strong correlation did not exist between the mRNA m⁵C modification level and the expression level (**Figure 6E**), and similar results were previously reported in systemic lupus erythematosus (SLE; Guo et al., 2020). We further investigated the biological process enrichment of mRNAs that were both upregulated and hypermethylated using GO functional analysis. Particularly, these transcripts were enriched in the following categories: response to chemicals, cellular response to chemical stimuli, regulation of cell death and regulation of apoptotic processes (**Figure 6F**), which indicates that both hypermethylated and upregulated transcripts play an important role in neuronal damage caused by ischemia and hypoxia.

Neuronal Apoptosis Occurs After OGD/R Treatment *in vitro*

To clarify the apoptotic status of neuron after OGD/R treatment, we extracted primary neurons from mice and successfully constructed a neuronal OGD/R model (**Supplementary Figure 4A**). A TUNEL assay was performed to examine apoptosis and the associated protein expression levels between the control and OGD/R models. TUNEL-positive cells detected in the OGD/R group exhibited increased DNA fragmentation,

indicating that the neurons exhibited obvious apoptosis *in vitro* after OGD/R (**Supplementary Figures 4B,C**). In addition, western blot assay revealed that the protein expression levels of Cleaved-Caspase3 and Bax significantly increased after the OGD/R treatment, while the level of Bcl-2 decreased. The experimental results further confirmed that the neuronal damage caused by OGD/R was closely related to neuronal apoptosis (**Supplementary Figures 4D,E**).

DISCUSSION

An increasing number of studies have demonstrated that mRNA methylation is involved in many neural functions and has an important impact on life activities. The absence of Methyltransferase-like 14 (METTL14) in the central nervous system (CNS) can prolong the cell cycle of cortical progenitor cells and reduce the differentiation of radial glial cells (Yoon et al., 2017). YTH domain family 2 (YTHDF2) can regulate neurodevelopment by degrading m⁶A methylation levels of neuronal mRNA (Li et al., 2018). In addition, mRNA methylation also plays an important role in the maturation of oligodendrocytes and the myelination of the CNS (Xu H et al., 2020). To date, most research on mRNA methylation in the nervous system has focused on m⁶A modification. As a new type of RNA methylation, m⁵C has been proven to play important roles in promoting RNA export out of the nucleus, regulating protein translation, and neural stem cell differentiation (Schaefer et al., 2009; Li et al., 2017). Studies have shown that m⁵C methylation is enriched in mouse brain tissues (Amort et al., 2017), whereas the specific distribution and function of mRNA m⁵C in neurons and its role in the important pathological processes of IRI remain unclear.

In this study, we conducted a comparative analysis of cytosine methylation in neuronal mRNA before and after OGD/R treatment, revealing the m⁵C modification of mRNA in neurons for the first time. We also clarified the main differences in neuronal m⁵C modification after OGD/R, including the methylation sequence preference, proportion of neuronal methylated mRNA, mRNA methylation modification level, distribution characteristics of methylated sites in mRNA and the distribution of methylated transcripts in chromosomes. Notably, the total methylation level in a sample depends on the modification level of each transcript and the number of methylated transcripts. In our results, we observed that after OGD/R treatment, the number of methylated sites increased significantly, but the number of methylated transcripts decreased. Compared with that of the control group, the transcript modification level of the OGD/R treatment group was increased, and there were more mRNAs with abundant methylation sites. Motif analysis shows that C sites were enriched in GC-rich regions, which is consistent with the results of previous studies (Yang et al., 2017).

Cytosine methylation accumulates near the translation initiation codon, which is similar to the previously reported distribution pattern of m⁵C in mouse brain tissues. In contrast, neuronal mRNA m⁵C sites are not significantly enriched in the

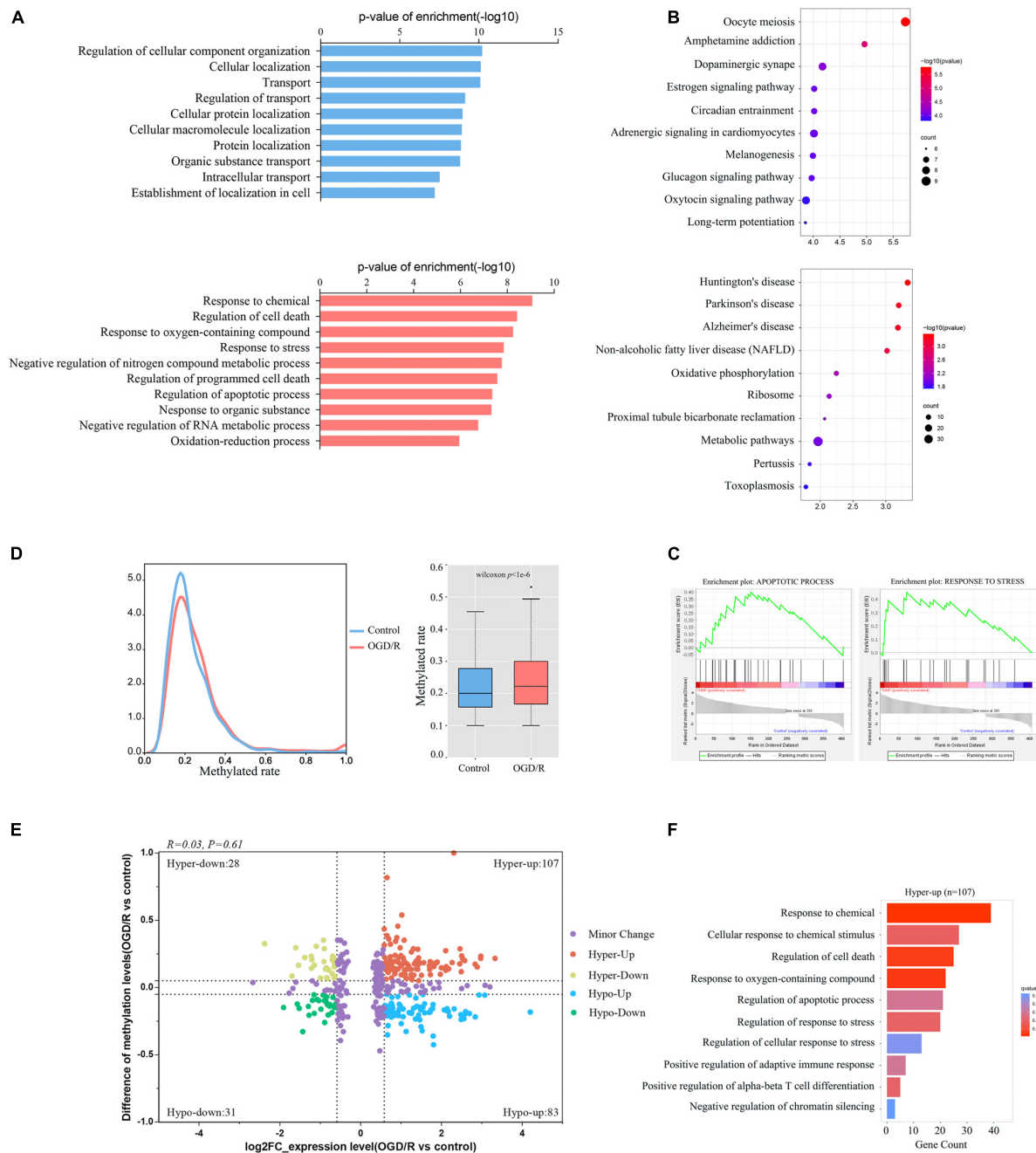


FIGURE 6 | Integration analyses of m⁵C-containing mRNA methylation and mRNA transcript expression. **(A)** GO biological process analyses of DEGs modified by m⁵C in control and OGD/R neurons. **(B)** KEGG analysis of DEGs modified by m⁵C in the control group and OGD/R group. **(C)** GSEA of DEGs modified by m⁵C with Gene Ontology sets. **(D)** Comparison of m⁵C levels in mRNAs between the control group and OGD/R group, density distribution plot (left), and box plot (right). **(E)** mRNA distributions with significant changes in gene expression and m⁵C modification levels between the control group and OGD/R group. **(F)** GO biological process analysis of hypermethylated mRNAs with upregulated genes after OGD/R treatment.

3'UTR region and are relatively more concentrated near the start codon and adjacent 5'UTR and CDS regions. Studies have demonstrated that the distribution patterns of m⁵C in mice and humans are highly conserved, and these sites are concentrated near the start codon, which is similar to the results noted for m⁶A modifications (Dominissini et al., 2012). Moreover, our

results show that the distribution patterns of m⁵C in neuronal mRNA before and after OGD/R treatment are very close, proving that the distribution of m⁵C on mRNA remains stable under different conditions. However, m⁵C in *Arabidopsis* is mainly concentrated in the CDS area (Cui et al., 2017), exhibiting a different distribution pattern and indicating that the m⁵C

modification of mRNA differs between animals and plants. This may be related to regulatory effects of m⁵C methyltransferases and bind proteins. At present, NSUN2 has been found to be the methyltransferases involved in the m⁵C modification of mammalian mRNA, which plays an important role in transcript translation (Tang et al., 2015). In addition, m⁵C methylation-binding nucleoprotein ALYREF can promote mammalian mRNA export and YBX1 promotes mammalian mRNA stability (Yang et al., 2017; Chen et al., 2019b). The main methylase found in plants is TRM4B, which plays an important role in the stability of tRNA and has a negative regulatory effect on mRNA expression (David et al., 2017). Moreover, previous studies discovered that cytosine methylation of mRNA in eukaryotes can effectively promote mRNA translation (Delatte et al., 2016), which may also be related to the distribution of m⁵C enriched around the start codon. RBPs play an important role in the regulation of mRNA expression. Furthermore, studies have shown that suppression of RBP SRSF1 prevents neurodegeneration and motor dysfunction in C9ORF72-related diseases (Hautbergue et al., 2017). After hypoxic-ischemic brain injury, RBM3 promotes neurogenesis in a niche-dependent manner through the IMP2-IGF2 signaling pathway (Zhu et al., 2019). Our study indicated that the overlap sites of SRSF1 and RBM3 with m⁵C are significantly altered before and after OGD/R, which may be related to the pathological processes of hypoxia-reperfusion.

Previous studies have shown that m⁵C-modified mRNA is involved in various biological processes (Flores et al., 2017; Yang et al., 2019; Chen et al., 2019b). In this study, we compared neuronal m⁵C methylation information before and after OGD/R. The results showed that compared with the control group, the OGD/R group had 862 transcripts with different levels of m⁵C methylation, including 480 hypermethylated and 382 hypomethylated transcripts ($p < 0.01$). These transcripts may be related to the pathogenesis of IRI. Previous studies have found that Tet methylcytosine dioxygenase 3 (TET3) has ischemic neuroprotection function by promoting the formation of DNA hydroxymethylation in the brain after IRI, which mainly occurs through biological pathways involved in oxidative stress and DNA repair (Morris-Blanco et al., 2020). Our research found that after OGD/R, the hypermethylated transcript functions were mainly enriched in pathological processes, including Huntington's disease and intrinsic apoptotic signaling pathways, whereas the hypomethylated transcript functions were mainly enriched in physiological processes. Moreover, the hub genes RPS9 and RPL8 of PPI network analysis were significantly increased in the brain tissue after cerebral ischemia and hemorrhage (Kim et al., 2003; Chen et al., 2019a). These findings indicate that m⁵C modification may also be involved in the pathological processes of cerebral injury.

To further explore the role of mRNA m⁵C modification in neurons, we performed bioinformatics analysis of 406 differentially expressed m⁵C-methylated transcripts, including 282 hypermethylated and 124 hypomethylated transcripts ($p < 0.05$). Our analyses did not reveal a strong relationship between the m⁵C methylation level and the corresponding mRNA expression level, which is similar to the results of previous

reports on bladder cancer and SLE (Chen et al., 2019b; Guo et al., 2020), and previous studies have shown that m⁵C methylation plays an important role in promoting the process of nuclear export (Yang et al., 2017). Therefore, the specific regulatory mechanism of m⁵C on mRNA deserves further exploration. Our research found that after OGD/R treatment, neurons are in a state of stress, their responses to chemical and oxygenated compounds are significantly enhanced, and neuronal apoptosis is simultaneously initiated. Methylation modification has been confirmed to play an important role in regulating neuronal death during neuronal oxidative stress (Chen et al., 2019c), but the role of the m⁵C modification in these processes remains unknown. Then, we focused on the up-regulated methylated transcripts and found that up-regulated transcripts with hypermethylated modification were significantly enriched in the processes of apoptosis and neurological diseases. However, these processes were not enriched in hypomethylated transcripts, indicating that m⁵C modification may play an important role in the pathological process of neuronal apoptosis and other neurological diseases. In addition, mRNA m⁵C modification was closely related to the neuronal fate after IRI. However, further functional studies are still needed to clarify the relationship between the neuronal mRNA m⁵C modification and expression level after IRI.

In summary, our study presented the first transcriptome-wide m⁵C methylation map before and after neuronal OGD/R injury and found that the m⁵C modifications are highly conserved in mRNA. Furthermore, our results provide a potential relationship between differential m⁵C mRNA modifications and neuronal damage induced by OGD/R. This new epigenetic modification may provide better insights into the pathogenesis of related neurological diseases and exogenous nerve damage. However, it is unclear how neuronal m⁵C expression can be specifically regulated; for example, the specific role of RNA methyltransferases NSUN2 and DNMT2 in neuronal ischemic injury remains unknown. More research is needed to discover its internal regulation mechanism. Targeting the m⁵C modification will become a promising strategy for the treatment of ischemia/reperfusion injury in the future.

DATA AVAILABILITY STATEMENT

The raw data has been made publicly available. GEO accession: GSE165256.

ETHICS STATEMENT

The animal study was reviewed and approved by Experimental Animal Ethics Committee of Tianjin Medical University.

AUTHOR CONTRIBUTIONS

HJ, CZ, ZQ, and XL conducted experiments and data collection. HJ, YoL, YK, and WD analyzed and interpreted the m⁵C data.

YiL, CW, WW, SS, and MH contributed to statistical analysis. HJ wrote the manuscript. SF and HZ provided the project funding and revised the manuscript. All authors have read and agreed to the published version of the manuscript.

FUNDING

This study was funded by the National Natural Science Foundation of China (81972073), Taishan Scholars Program of Shandong Province-Young Taishan Scholars (tsqn201909197), Young Elite Scientists Sponsorship Program by Tianjin (TJSQNTJ-2017-11), Science and Technology Development Fund of Tianjin Education Commission for Higher Education (2018KJ078), International Cooperation Program of National Natural Science Foundation of China (81620108018), Tianjin Key Research and Development Plan, Key Projects for Science

and Technology Support (19YFZCSY00660), Key Program of Natural Science Foundation of Tianjin (19JCZDJC36300), and National Natural Science Foundation of China (81772342).

ACKNOWLEDGMENTS

We thank NewCore Biotech (Shanghai, China) for bioinformatics support, Cloudseq Biotech, Inc. (Shanghai, China) for the m⁵C bisulfite sequencing service.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2021.633681/full#supplementary-material>

REFERENCES

- Amort, T., Rieder, D., Wille, A., Khokhlova-Cubberley, D., Riml, C., Trixl, L., et al. (2017). Distinct 5-methylcytosine profiles in poly(A) RNA from mouse embryonic stem cells and brain. *Genome Biol.* 18:1.
- Cao, G., Li, H. B., Yin, Z., and Flavell, R. A. (2016). Recent advances in dynamic m6A RNA modification. *Open Biol.* 6:160003. doi: 10.1098/rsob.160003
- Carlile, T. M., Rojas-Duran, M. F., Zinshteyn, B., Shin, H., Bartoli, K. M., and Gilbert, W. V. (2014). Pseudouridine profiling reveals regulated mRNA pseudouridylation in yeast and human cells. *Nature* 515, 143–146. doi: 10.1038/nature13802
- Chen, B., Chen, Z., Liu, M., Gao, X., Cheng, Y., Wei, Y., et al. (2019a). Inhibition of neuronal ferroptosis in the acute phase of intracerebral hemorrhage shows long-term cerebroprotective effects. *Brain Res. Bull.* 153, 122–132. doi: 10.1016/j.brainresbull.2019.08.013
- Chen, X., Li, A., Sun, B. F., Yang, Y., Han, Y. N., Yuan, X., et al. (2019b). 5-methylcytosine promotes pathogenesis of bladder cancer through stabilizing mRNAs. *Nat. Cell Biol.* 21, 978–990. doi: 10.1038/s41556-019-0361-y
- Chen, X., Yu, C., Guo, M., Zheng, X., Ali, S., Huang, H., et al. (2019c). Down-regulation of m6A mRNA methylation is involved in dopaminergic neuronal death. *ACS Chem. Neurosci.* 10, 2355–2363. doi: 10.1021/acschemneuro.8b00657
- Cui, X., Liang, Z., Shen, L., Zhang, Q., Bao, S., Geng, Y., et al. (2017). 5-Methylcytosine RNA methylation in *Arabidopsis thaliana*. *Mol. Plant* 10, 1387–1399. doi: 10.1016/j.molp.2017.09.013
- David, R., Burgess, A., Parker, B., Li, J., Pulsford, K., Sibbritt, T., et al. (2017). Transcriptome-wide mapping of RNA 5-methylcytosine in *Arabidopsis* mRNAs and noncoding RNAs. *Plant Cell* 29, 445–460. doi: 10.1105/tpc.16.00751
- Delatte, B., Wang, F., Ngoc, L. V., Collignon, E., Bonvin, E., Depluis, R., et al. (2016). RNA biochemistry. Transcriptome-wide distribution and function of RNA hydroxymethylcytosine. *Science* 351, 282–285. doi: 10.1126/science.aac5253
- Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., et al. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29, 15–21. doi: 10.1093/bioinformatics/bts635
- Dominissini, D., Moshitch-Moshkovitz, S., Schwartz, S., Salmon-Divon, M., Ungar, L., Osenberg, S., et al. (2012). Topology of the human and mouse m6A RNA methylomes revealed by m6A-seq. *Nature* 485, 201–206. doi: 10.1038/nature11112
- Dominissini, D., Nachtergaele, S., Moshitch-Moshkovitz, S., Peer, E., Kol, N., Ben-Haim, M. S., et al. (2016). The dynamic N(1)-methyladenosine methylome in eukaryotic messenger RNA. *Nature* 530, 441–446. doi: 10.1038/nature16998
- Duehrkop, C., and Rieben, R. (2014). Ischemia/reperfusion injury: effect of simultaneous inhibition of plasma cascade systems versus specific complement inhibition. *Biochem. Pharmacol.* 88, 12–22. doi: 10.1016/j.bcp.2013.12.013
- Eltzschig, H. K., and Eckle, T. (2011). Ischemia and reperfusion—from mechanism to translation. *Nat. Med.* 17, 1391–1401. doi: 10.1038/nm.2507
- Flores, J. V., Cordero-Espinoza, L., Oetzuerk-Winder, F., Andersson-Rolf, A., Selmi, T., Blanco, S., et al. (2017). Cytosine-5 RNA methylation regulates neural stem cell differentiation and motility. *Stem Cell Rep.* 8, 112–124. doi: 10.1016/j.stemcr.2016.11.014
- Guo, G., Wang, H., Shi, X., Ye, L., Yan, K., Chen, Z., et al. (2020). Disease activity-associated alteration of mRNA m(5) C Methylation in CD4(+) T cells of systemic lupus erythematosus. *Front. Cell Dev. Biol.* 8:430. doi: 10.3389/fcell.2020.00430
- Hao, Z., Lv, D., Ge, Y., Shi, J., Weijers, D., Yu, G., et al. (2020). RIdiogram: drawing SVG graphics to visualize and map genome-wide data on the idiograms. *PeerJ Comput. Sci.* 6:e251.
- Hautbergue, G. M., Castelli, L. M., Ferraiuolo, L., Sanchez-Martinez, A., Cooper-Knock, J., Higginbottom, A., et al. (2017). SRSF1-dependent nuclear export inhibition of C9ORF72 repeat transcripts prevents neurodegeneration and associated motor deficits. *Nat. Commun.* 8:16063. doi: 10.1038/ncomms16063
- Helm, M. (2006). Post-transcriptional nucleotide modification and alternative folding of RNA. *Nucleic Acids Res.* 34, 721–733. doi: 10.1093/nar/gkj471
- Hilgenberg, L. G., and Smith, M. A. (2007). Preparation of dissociated mouse cortical neuron cultures. *J. Vis. Exp.* 10:562. doi: 10.3791/562
- Hu, B., Yang, Y. T., Huang, Y., Zhu, Y., and Lu, Z. J. (2017). POSTAR: a platform for exploring post-transcriptional regulation coordinated by RNA-binding proteins. *Nucleic Acids Res.* 45, D104–D114. doi: 10.1093/nar/gkw888
- Huang da, W., Sherman, B. T., and Lempicki, R. A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* 4, 44–57. doi: 10.1038/nprot.2008.211
- Kim, S. Y., Lee, M. Y., Cho, K. C., Choi, Y. S., Choi, J. S., Sung, K. W., et al. (2003). Alterations in mRNA expression of ribosomal protein S9 in hydrogen peroxide-treated neurotumor cells and in rat hippocampus after transient ischemia. *Neurochem. Res.* 28, 925–931. doi: 10.1023/a:1023283628454
- Lauterbach, E. C. (2013). Psychotropics regulate Skp1a, Aldh1a1, and Hspa8 transcription—potential to delay Parkinson's disease. *Prog. Neuropsychopharmacol. Biol. Psychiatry* 40, 236–239. doi: 10.1016/j.pnpbp.2012.08.021
- Li, M., Zhao, X., Wang, W., Shi, H., Pan, Q., Lu, Z., et al. (2018). Ythdf2-mediated m(6)A mRNA clearance modulates neural development in mice. *Genome Biol.* 19:69. doi: 10.1186/s13059-018-1436-y
- Li, Q., Li, X., Tang, H., Jiang, B., Dou, Y., Gorospe, M., et al. (2017). NSUN2-Mediated m5C Methylation and METTL3/METTL14-Mediated m6A Methylation Cooperatively Enhance p21 Translation. *J. Cell. Biochem.* 118, 2587–2598. doi: 10.1002/jcb.25957
- Li, X., Zhu, P., Ma, S., Song, J., Bai, J., Sun, F., et al. (2015). Chemical pulldown reveals dynamic pseudouridylation of the mammalian transcriptome. *Nat. Chem. Biol.* 11, 592–597. doi: 10.1038/nchembio.1836

- Lopez, M. G., Pandharipande, P., Morse, J., Shotwell, M. S., Milne, G. L., Pretorius, M., et al. (2017). Intraoperative cerebral oxygenation, oxidative injury, and delirium following cardiac surgery. *Free Radic. Biol. Med.* 103, 192–198. doi: 10.1016/j.freeradbiomed.2016.12.039
- Meyer, K. D., and Jaffrey, S. R. (2017). Rethinking m(6)A readers, writers, and erasers. *Annu. Rev. Cell Dev. Biol.* 33, 319–342.
- Morris-Blanco, K. C., Chokkalla, A. K., Bertoglat, M. J., and Vemuganti, R. (2020). TET3 regulates DNA hydroxymethylation of neuroprotective genes following focal ischemia. *J. Cereb. Blood Flow Metab.* doi: 10.1177/0271678X20912965 [Epub ahead of print].
- Olarerin-George, A. O., and Jaffrey, S. R. (2017). MetaPlotR: a Perl/R pipeline for plotting metagenes of nucleotide modifications and other transcriptomic sites. *Bioinformatics* 33, 1563–1564. doi: 10.1093/bioinformatics/btx002
- Quinlan, A. R., and Hall, I. M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 26, 841–842. doi: 10.1093/bioinformatics/btq033
- Rieder, D., Amort, T., Kugler, E., Lusser, A., and Trajanoski, Z. (2016). meRanTK: methylated RNA analysis ToolKit. *Bioinformatics* 32, 782–785. doi: 10.1093/bioinformatics/btv647
- Schaefer, M., Pollex, T., Hanna, K., and Lyko, F. (2009). RNA cytosine methylation analysis by bisulfite sequencing. *Nucleic Acids Res.* 37:e12. doi: 10.1093/nar/gkn954
- Sciarretta, C., and Minichiello, L. (2010). The preparation of primary cortical neuron cultures and a practical application using immunofluorescent cytochemistry. *Methods Mol. Biol.* 633, 221–231. doi: 10.1007/978-1-59745-019-5_16
- Silva, P. N., Furuya, T. K., Braga, I. L., Rasmussen, L. T., Labio, R. W., Bertolucci, P. H., et al. (2014). Analysis of HSPA8 and HSPA9 mRNA expression and promoter methylation in the brain and blood of Alzheimer's disease patients. *J. Alzheimers Dis.* 38, 165–170. doi: 10.3233/jad-13-0428
- Tang, H., Fan, X., Xing, J., Liu, Z., Jiang, B., Dou, Y., et al. (2015). NSun2 delays replicative senescence by repressing p27 (KIP1) translation and elevating CDK1 translation. *Aging* 7, 1143–1158. doi: 10.18632/aging.100860
- Wiberg, S., Hassager, C., Schmidt, H., Thomsen, J. H., Frydland, M., Lindholm, M. G., et al. (2016). Neuroprotective effects of the glucagon-like peptide-1 analog exenatide after out-of-hospital cardiac arrest: a randomized controlled trial. *Circulation* 134, 2115–2124. doi: 10.1161/circulationaha.116.024088
- Xu, H., Dzhashiashvili, Y., Shah, A., Kunjamma, R. B., Weng, Y. L., Elbaz, B., et al. (2020). m(6)A mRNA methylation is essential for oligodendrocyte maturation and CNS myelination. *Neuron* 105, 293–309.e295. doi: 10.1016/j.neuron.2019.12.013
- Xu, K., Mo, Y., Li, D., Yu, Q., Wang, L., Lin, F., et al. (2020). N(6)-methyladenosine demethylases Alkbh5/Fto regulate cerebral ischemia-reperfusion injury. *Ther. Adv. Chron. Dis.* 11:2040622320916024. doi: 10.1177/2040622320916024
- Yang, X., Yang, Y., Sun, B. F., Chen, Y. S., Xu, J. W., Lai, W. Y., et al. (2017). 5-methylcytosine promotes mRNA export - NSUN2 as the methyltransferase and ALYREF as an m(5)C reader. *Cell Res.* 27, 606–625. doi: 10.1038/cr.2017.55
- Yang, Y., Wang, L., Han, X., Yang, W. L., Zhang, M., Ma, H. L., et al. (2019). RNA 5-methylcytosine facilitates the maternal-to-zygotic transition by preventing maternal mRNA decay. *Mol. Cell* 75, 1188–1202.e1111. doi: 10.1016/j.molcel.2019.06.033
- Yoon, K. J., Ringeling, F. R., Vissers, C., Jacob, F., Pokrass, M., Jimenez-Cyrus, D., et al. (2017). Temporal control of mammalian cortical neurogenesis by m(6)A methylation. *Cell* 171, 877–889.e817. doi: 10.1016/j.cell.2017.09.003
- Zhao, B. S., Roundtree, I. A., and He, C. (2017). Post-transcriptional gene regulation by mRNA modifications. *Nat. Rev. Mol. Cell Biol.* 18, 31–42. doi: 10.1038/nrm.2016.132
- Zhao, H., Li, G., Wang, R., Tao, Z., Zhang, S., Li, F., et al. (2019). MiR-424 prevents astrogliosis after cerebral ischemia/reperfusion in elderly mice by enhancing repressive H3K27me3 via NFIA/DNMT1 signaling. *FEBS J.* 286, 4926–4936. doi: 10.1111/febs.15029
- Zhao, Z., Lu, C., Li, T., Wang, W., Ye, W., Zeng, R., et al. (2018). The protective effect of melatonin on brain ischemia and reperfusion in rats and humans: *in vivo* assessment and a randomized controlled trial. *J. Pineal Res.* 65:e12521. doi: 10.1111/jpi.12521
- Zhou, Y., Zhou, B., Pache, L., Chang, M., Khodabakhshi, A. H., Tanaseichuk, O., et al. (2019). Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat. Commun.* 10:1523.
- Zhu, X., Yan, J., Bregere, C., Zelmer, A., Goerne, T., Kapfhammer, J. P., et al. (2019). RBM3 promotes neurogenesis in a niche-dependent manner via IMP2-IGF2 signaling pathway after hypoxic-ischemic brain injury. *Nat. Commun.* 10:3983. doi: 10.1038/s41467-019-11870-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 Jian, Zhang, Qi, Li, Lou, Kang, Deng, Lv, Wang, Wang, Shang, Hou, Zhou and Feng. 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.



Profiling of RNA N^6 -Methyladenosine Methylation Reveals the Critical Role of m^6A in Chicken Adipose Deposition

Bohan Cheng^{1,2,3†}, Li Leng^{1,2,3†}, Ziwei Li^{1,2,3}, Weijia Wang^{1,2,3}, Yang Jing^{1,2,3}, Yudong Li^{1,2,3}, Ning Wang^{1,2,3}, Hui Li^{1,2,3} and Shouzhi Wang^{1,2,3*}

¹ Key Laboratory of Chicken Genetics and Breeding, Ministry of Agriculture and Rural Affairs, Harbin, China, ² Key Laboratory of Animal Genetics, Breeding and Reproduction, Education Department of Heilongjiang Province, Harbin, China, ³ College of Animal Science and Technology, Northeast Agricultural University, Harbin, China

OPEN ACCESS

Edited by:

Xiao Han,
Chinese Academy of Agricultural
Sciences, China

Reviewed by:

Xuemei Deng,
China Agricultural University, China
Ying Wang,
University of California, Davis,
United States

*Correspondence:

Shouzhi Wang
shouzhwang@neau.edu.cn

[†]These authors have contributed
equally to this work

Specialty section:

This article was submitted to
Epigenomics and Epigenetics,
a section of the journal
Frontiers in Cell and Developmental
Biology

Received: 01 August 2020

Accepted: 18 January 2021

Published: 05 February 2021

Citation:

Cheng B, Leng L, Li Z, Wang W,
Jing Y, Li Y, Wang N, Li H and Wang S
(2021) Profiling of RNA
 N^6 -Methyladenosine Methylation
Reveals the Critical Role of m^6A in
Chicken Adipose Deposition.
Front. Cell Dev. Biol. 9:590468.
doi: 10.3389/fcell.2021.590468

One of the main objectives of broiler breeding is to prevent excessive abdominal adipose deposition. The role of RNA modification in adipose deposition is not clear. This study was aimed to map m^6A modification landscape in chicken adipose tissue. MeRIP-seq was performed to compare the differences in m^6A methylation pattern between fat and lean broilers. We found that start codons, stop codons, coding regions, and 3'-untranslated regions were generally enriched for m^6A peaks. The high m^6A methylated genes (fat birds vs. lean birds) were primarily associated with fatty acid biosynthesis and fatty acid metabolism, while the low m^6A methylated genes were mainly involved in processes associated with development. Furthermore, we found that the mRNA levels of many genes may be regulated by m^6A modification. This is the first comprehensive characterization of m^6A patterns in the chicken adipose transcriptome, and provides a basis for studying the role of m^6A modification in fat deposition.

Keywords: chicken, fat deposition, adipose tissue, N^6 -methyladenosine, MeRIP-seq

INTRODUCTION

As a result of long-term breeding efforts, the growth rate and meat yield of broilers have significantly improved; however, this has led to excessive body fat (especially abdominal fat) deposition. The accumulation of excess fat in broilers has many undesirable consequences, such as decreased reproductive performance and reduced feed-conversion efficiencies (Zhou et al., 2006; Zhang et al., 2018). Adipose tissue is an important energy storage and endocrine organ and is the cornerstone of energy metabolism homeostasis (McGown et al., 2014; Choe et al., 2016). Adipose tissue development is controlled by a complex network of transcription factors (Farmer, 2006). In addition to transcriptional regulation, evidence suggests that adipose development and fat deposition can be regulated by the epigenetic mechanisms, such as DNA methylation (Zhu et al., 2012), histone modification (Wang et al., 2010), and chromatin remodeling (Siersbaek et al., 2011).

In addition to the chemical modification of DNA and proteins, RNA modification has become a research hotspot in the field of epigenetics in recent years. So far, more than 100 types of chemical modifications of RNA have been identified, with N^6 -methyladenosine (m^6A) methylation being the most pervasive modification in eukaryotes (Yue et al., 2015). m^6A is installed by a multicomponent methyltransferase complex consisting of Methyltransferase Like 3 (METTL3),

METTL14 and Wilms Tumor 1 Associated Protein (WTAP), and erased by m⁶A demethylase fat mass and obesity-associated protein (FTO) and α -ketoglutarate-dependent dioxygenase alkB homolog 5 (ALKBH5) (Yang et al., 2018). m⁶A is involved in many important biological processes through the post-transcriptional regulation of gene expression, including mRNA export, the processing of pri-miRNA, alternative splicing, mRNA degradation and translation (Yang et al., 2018).

In mammals, emerging evidence shows that m⁶A modification plays a critical role in adipose development and hepatic lipid metabolism (Tao et al., 2017; Lu et al., 2018; Wang et al., 2018). Knockdown of METTL3, METTL14, WTAP or FTO inhibited the differentiation of mouse 3T3-L1 preadipocytes (Zhao et al., 2014; Kobayashi et al., 2018). However, whether m⁶A modification is involved in poultry adipose deposition is still largely unknown. Here, we used Northeast Agricultural University broiler lines divergently selected for abdominal fat content (NEAUHLF) as fat and lean animal models to compare the differences in m⁶A topological patterns and functions. We collected abdominal adipose tissue from the two broiler lines for m⁶A methylation profiling with methylated RNA immunoprecipitation (IP) sequencing (MeRIP-seq). Our data showed that the adipose tissue mRNA was extensively methylated with m⁶A to fine-tuning the expression of genes responsible for lipid metabolism and adipogenesis.

MATERIALS AND METHODS

Experimental Birds and Management

Animal studies were conducted according to the guidelines for the care and use of experimental animals established by the Ministry of Science and Technology of the People's Republic of China (approval number: 2006-398) and were approved by the Laboratory Animal Management Committee and the Institutional Biosafety Committee of Northeast Agricultural University (Harbin, China). In total, six male birds (lean line, $n = 3$, and fat line, $n = 3$) from the 23rd generation (G₂₃) of NEAUHLF were used for MeRIP-seq analysis. NEAUHLF has been selected since 1996 using plasma very-low-density lipoprotein concentration and abdominal fat percentage (AFP; abdominal fat weight [AFW]/body weight at 7 weeks [wk] of age [BW₇]) as selection criteria. Details of the breeding procedure have been described previously (Guo et al., 2011; Zhang et al., 2017). All birds used in this study were kept in similar environmental conditions and had free access to feed and water. From hatching to 3 wk of age, all birds received the starter feed (3,100 kcal of ME/kg and 210 g/kg of crude protein [CP]). Then, from 4 to 7 wk of age, all birds were fed a grower diet (3,000 kcal of ME/kg and 190 g/kg of CP).

Tissue Collection

Six male birds (three birds of each broiler line at 7 wk of age) from G₂₃ were slaughtered after fasting for 10 h, and the BW₇ and AFW were measured and used to calculate AFP (Supplementary Figure 1). Abdominal fat tissues were collected, washed with 0.75% NaCl solution, snap-frozen in liquid nitrogen, and stored at -80°C until RNA extraction.

RNA Isolation and Fragmentation

The total RNA from the abdominal adipose tissue was extracted using TRIzol reagent (Invitrogen Co., CA, USA) according to the manufacturer's instructions. The ribosomal-RNA content of the total RNA was reduced using the Ribo-Zero rRNA Removal Kit (Illumina Inc., CA, USA). Then, the RNA was chemically fragmented into fragments of ~ 100 nucleotides in length using fragmentation buffer (Illumina Inc.).

Methylated RNA IP Library Construction and Sequencing

The MeRIP-seq service was provided by Cloudseq Biotech Inc. (Shanghai, China). Briefly, IP of the m⁶A RNA was performed with the GenSeqTM m⁶A RNA IP Kit (GenSeq Inc., China) following the manufacturer's instructions. Both the m⁶A IP samples and the input samples without IP were used for library generation with NEBNext[®] Ultra II Directional RNA Library Prep Kit (New England Biolabs Inc., USA). The quality of the libraries was evaluated with the BioAnalyzer 2100 system (Agilent Technologies Inc., USA). Library sequencing was performed on an Illumina HiSeq instrument with 150 bp paired-end reads.

MeRIP-seq Analysis

Briefly, paired-end reads were harvested from the Illumina HiSeq 4000 sequencer, and Q30 scoring was used for quality control. Cutadapt software (v1.9.3) was employed for 3'-adaptor-trimming and removal of low-quality reads. The clean reads were aligned to the reference chicken genome sequences (galGal6) using Hisat2 software (v2.0.4), and only the uniquely mapped and non-duplicated alignments were further analyzed. The m⁶A-modification peaks were called by MACS software (Zhang et al., 2008), the "effective genome size" parameter was adjusted to the calculated transcriptome size (1.77×10^8) (Dominissini et al., 2012); meanwhile, the input RNA sequencing (RNA-seq) data were used as the background when calling peaks. Peaks that shared at least 50% overlapping lengths were defined as recurrent peaks.

To examine the distribution pattern of the m⁶A peaks throughout different regions of the transcripts, the mRNA transcripts were divided into five non-overlapping segments: the 5'UTR, start codon (100 nucleotides centered on the start codon), CDS, stop codon (100 nucleotides centered on the stop codon), and 3'UTR. Each area was separated into 20 bins (Luo et al., 2014). Bedtools (v2.26.0) was used to count the peak number of each bin, and the counts were employed to plot the patterns by R (v3.4.4).

In the fat and lean groups, the top 1,000 significantly enriched peaks (MACS-assigned fold change > 2 and $P < 0.00001$) within the mRNAs from three biological replicates were combined (Dominissini et al., 2012), and 101 nucleotides centered on the collected peaks of each group were subjected to *de novo* motif analysis using DREME software (Bailey, 2011). The DREME tool in the MEME suite (<http://meme-suite.org/tools/dreme>) was used to discover relatively short (up to 8 bp), ungapped motifs that are enriched within a set of target sequences (m⁶A peak

sequences) relative to a set of control sequences (shuffled m⁶A peak sequences) (Bailey, 2011).

We obtained the common and unique peaks using bedtools intersect (v2.26.0). For a peak to be classified as line-unique, it was assumed not to overlap (<50% overlapping length) any peak of the other line; meanwhile, we defined a peak that appeared in both chicken lines as a common peak (≥50% overlapping length). Line-dynamic methylated peaks (fold change ≥2 and $P < 0.05$), which showing a change of intensity in some of the common peaks, were identified by diffReps software.

The gene expression levels were determined using the input data, and the number of sequenced fragments of each transcript was normalized using the algorithm of Fragments Per Kilobase of Transcript Per Million Fragments Mapped (FPKM) by Cufflinks software (v2.2.1). Differentially expressed transcripts (fold change ≥2 and $P < 0.05$) between fat and lean groups were identified with the Cuffdiff program (v2.2.1). The FPKM of input and IP samples were calculated by Cufflinks software (v2.2.1). NNFPKM (NNFPKM = FPKM_IP/FPKM_INPUT) were used to analyze the m⁶A enrichment level in fat and lean groups. Then FPKM_INPUT and NNFPKM were log₂ transformed and Pearson correlation analysis of mRNA expression levels and m⁶A methylation levels was carried out. Gene Ontology (GO) and pathway analyses were performed using GO (www.geneontology.org) and the Kyoto Encyclopedia of Genes and Genomes (KEGG) database (www.genome.jp/kegg). The thresholds for significant enrichment for GO and KEGG analysis were set at $P < 0.05$.

RESULTS

Transcriptome-Wide Mapping of the M⁶A Methylation Landscape in Chicken Adipose Tissues

MeRIP-seq produced 80,966,354–102,649,284 raw reads from input or IP abdominal fat tissues from lean line (L-AF) and fat line (F-AF) chickens. After filtering out low quality data, more than 71,600,000 high-quality reads from each sample were mapped to the galGal6 genome. More than 86% of the clean reads from all the samples were uniquely mapped to chicken reference genome (Table 1). In the input samples, we detected 9,041 and 9,452 expressed mRNA transcripts in lean line and fat line, respectively. After the methylated-RNA fragments were mapped to the transcriptome, 4,615 m⁶A transcripts (common m⁶A transcripts from three biologic replicates) were identified among the 5,965 coding transcripts (common mRNA transcripts from three biologic replicates) in the L-AF samples and 4,438 m⁶A transcripts among the 6,654 coding transcripts in the F-AF samples (Table 2). The proportion of methylated transcripts were 77 and 67% in L-AF and F-AF, respectively. In addition, we detected 7,097 recurrent m⁶A peaks within 5,965 coding transcripts in L-AF and 6,966 recurrent m⁶A peaks among 6,654 coding transcripts in F-AF (Table 2; Supplementary Data 1). Based on this information, we estimated that the chicken abdominal adipose transcriptome contained 1.54–1.57 m⁶A peaks per methylated transcript and 1.05–1.19 m⁶A peaks per

TABLE 1 | Summary of sequencing data and read-alignment statistics from MeRIP-seq in abdominal adipose in fat and lean broiler lines.

Sample ID	Raw reads	Clean reads	Reads uniquely mapped to genome	Clean reads uniquely mapped (%)
L-AF-IP	93,333,165	93,228,190	80,368,297	86.21
L-AF-input	91,232,892	90,901,841	80,641,851	88.71
F-AF-IP	102,649,284	102,539,687	89,675,215	87.46
F-AF-input	80,966,354	80,583,035	71,641,685	88.81

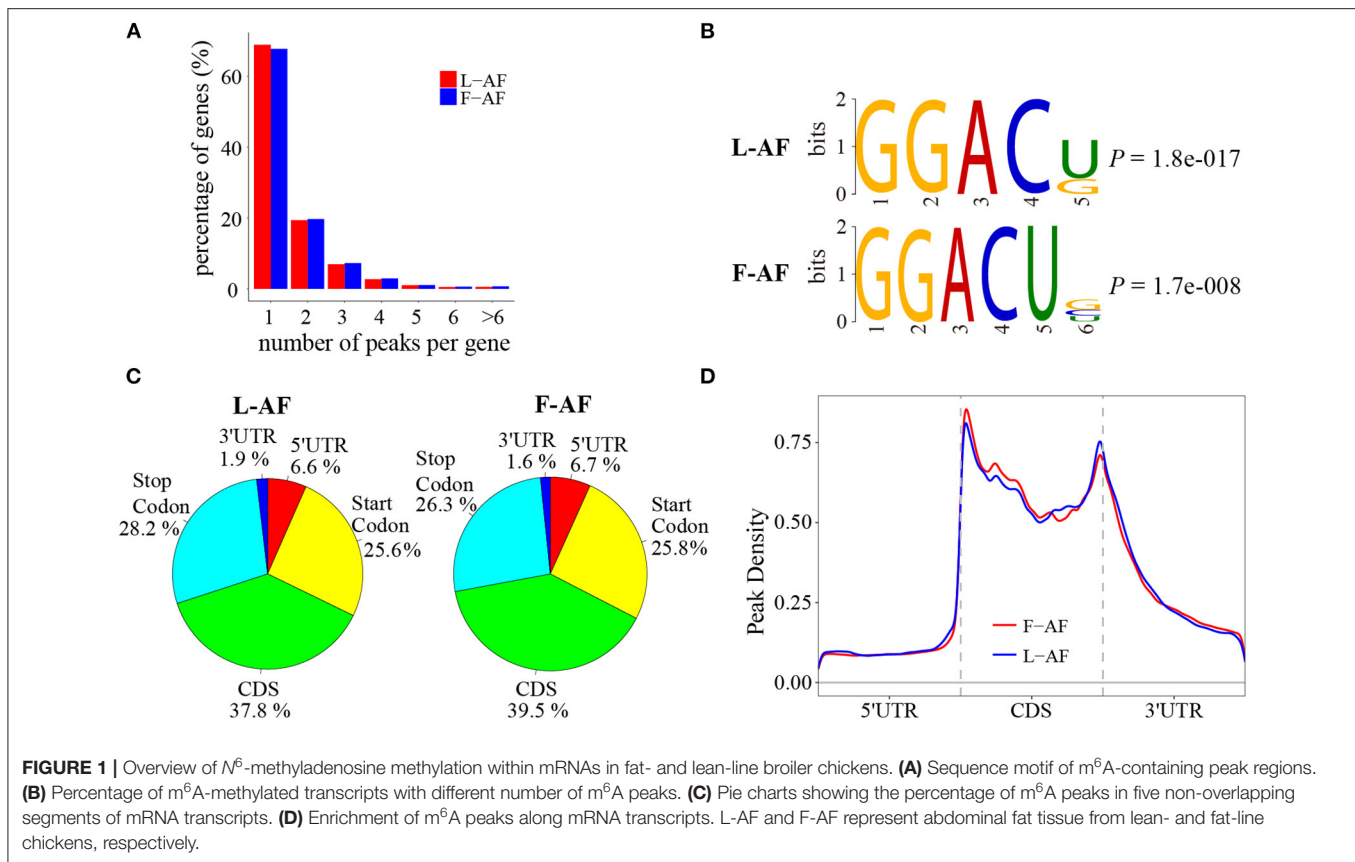
TABLE 2 | Number of m⁶A peaks detected in the abdominal adipose of the two chicken lines.

Sample ID ^a	mRNA transcripts	m ⁶ A mRNA transcripts	Total m ⁶ A peaks	Total m ⁶ A peaks per m ⁶ A transcript	Total m ⁶ A peaks per transcript
L-AF-1	7,639	6,563	12,164	1.85	1.59
L-AF-2	7,965	6,408	11,841	1.85	1.49
L-AF-3	6,914	6,642	12,919	1.95	1.87
L-AF	5,965	4,615	7,097	1.54	1.19
F-AF-1	8,634	6,170	12,015	1.95	1.39
F-AF-2	7,463	6,444	12,958	2.01	1.74
F-AF-3	8,023	6,017	11,307	1.88	1.41
F-AF	6,654	4,438	6,966	1.57	1.05

^aL-AF-1, L-AF-2, and L-AF-3 refer to sample 1, sample 2, and sample 3 of abdominal adipose tissue from lean line, respectively; L-AF means the recurrent peak sample for L-AF-1, L-AF-2, and L-AF-3 (≥50% overlapping lengths); F-AF-1, F-AF-2, and F-AF-3 refer to sample 1, sample 2, and sample 3 of abdominal adipose tissue from fat line, respectively; F-AF means the recurrent peak sample for F-AF-1, F-AF-2, and F-AF-3 (≥50% overlapping lengths).

transcript. These data are similar to those of the chicken ovary and pig *longissimus dorsi* muscle, which have ~1.5 m⁶A peaks per methylated transcript (Fan et al., 2019; Jiang et al., 2019); and mouse liver and pig subcutaneous fatty tissue transcriptomes, which possess 1.3 m⁶A peaks per transcript (Dominissini et al., 2012; Tao et al., 2017). However, our results were lower than that of the mouse L cells, which presenting about 3 m⁶A residues per mRNA transcript (Perry et al., 1975).

To determine how the m⁶A modification was distributed throughout the chicken transcriptome. We classified the methylated transcripts based on the number of m⁶A peaks contained in each transcript, and found that nearly 85% of the methylated transcripts contained one or two m⁶A peaks, and about 5% of the methylated transcripts contained four or more peaks (Figure 1A); this ratio is similar to that previously reported in humans (5.5%) (Dominissini et al., 2012) but is lower than that in pigs (10%) (Wang et al., 2018) and *Arabidopsis thaliana* (17%) (Wan et al., 2015). We then investigated whether the m⁶A peaks we identified share a conservative RRACH motif (where R stands for purine, A represents m⁶A, and H represents a non-guanine base) (Dominissini et al., 2012; Meyer et al., 2012; Luo et al., 2014), we conducted a search for the motifs enriched



in the regions around the m⁶A peaks. The results showed that GGACU was significantly enriched and consistently considered to be the best motif in both broiler lines (**Figure 1B**). To confirm the preferential localization of m⁶A in the transcripts, m⁶A peaks were categorized into five non-overlapping segments: 5'UTR, the start codon segment, coding sequence (CDS), the stop codon segment and 3'UTR. Our results show that m⁶A was most often located in the CDS, and sometimes near the start and stop codons (**Figure 1C**), which is consistent with the patterns identified in the mouse and pig (Tao et al., 2017; Luo et al., 2019). Metagene profiling of the m⁶A peaks showed that they were primarily enriched in CDSs, near the start and stop codons, and close to the beginning of 3'UTRs (**Figure 1D**), which differs from the pattern identified in mammals (Meyer et al., 2012; Tao et al., 2017).

Biological Pathways Associated With Common and Line-Unique m⁶A Genes

To discover the differences in m⁶A modification between the two chicken lines, we first identified the line-unique m⁶A peaks and genes. We found 4,318 peaks (representing 3,325 genes) that were common methylated in L-AF and F-AF, along with 2,783 and 2,656 peaks (representing 1,290 and 1,113 genes, respectively) that were specifically methylated in L-AF and F-AF, respectively (**Figure 2A**; **Supplementary Data 2**). To predict the functions associated with the m⁶A-modified genes, we conducted the gene ontology (GO) biological

process (BP) and KEGG pathway analysis. The m⁶A genes common to both lines were predominantly assigned to lipid metabolism, transcription, protein modification, the Wnt-signaling pathway, and the cytoskeleton ($P < 0.05$, **Figures 2B,C**, **Supplementary Data 3, 4**). In addition, the L-AF-unique m⁶A genes (L-AF UMGs) were significantly involved in development-associated processes, cell junction assembly, ribosome biogenesis, and others ($P < 0.05$, **Figures 3A,B**, **Supplementary Data 3, 4**). However, the F-AF-unique m⁶A genes (F-AF UMGs) were generally involved in the cellular responses to transforming growth factor-beta, mRNA processing, protein localization, and ubiquitin-mediated proteolysis ($P < 0.05$, **Figures 4A,B**, **Supplementary Data 3, 4**).

Involvement of Line-Dynamic m⁶A Genes in Lipogenesis-Related Pathways

In addition to the line-unique m⁶A genes, the line-dynamic m⁶A genes (common to both chicken lines but with different m⁶A peak intensities) were also selected for GO biological process and KEGG pathway analyses. We found 1,504 common m⁶A peaks with remarkably different abundances between the two chicken lines, which represented 1,172 coding genes, of which 71.1% (1,069/1,504) were lower in the fat line compared with the lean line (**Table 3**, **Supplementary Data 5**). **Tables 4, 5** show the top 15 high and low m⁶A peaks of mRNAs (fat line vs. lean line) with the highest fold-change values.

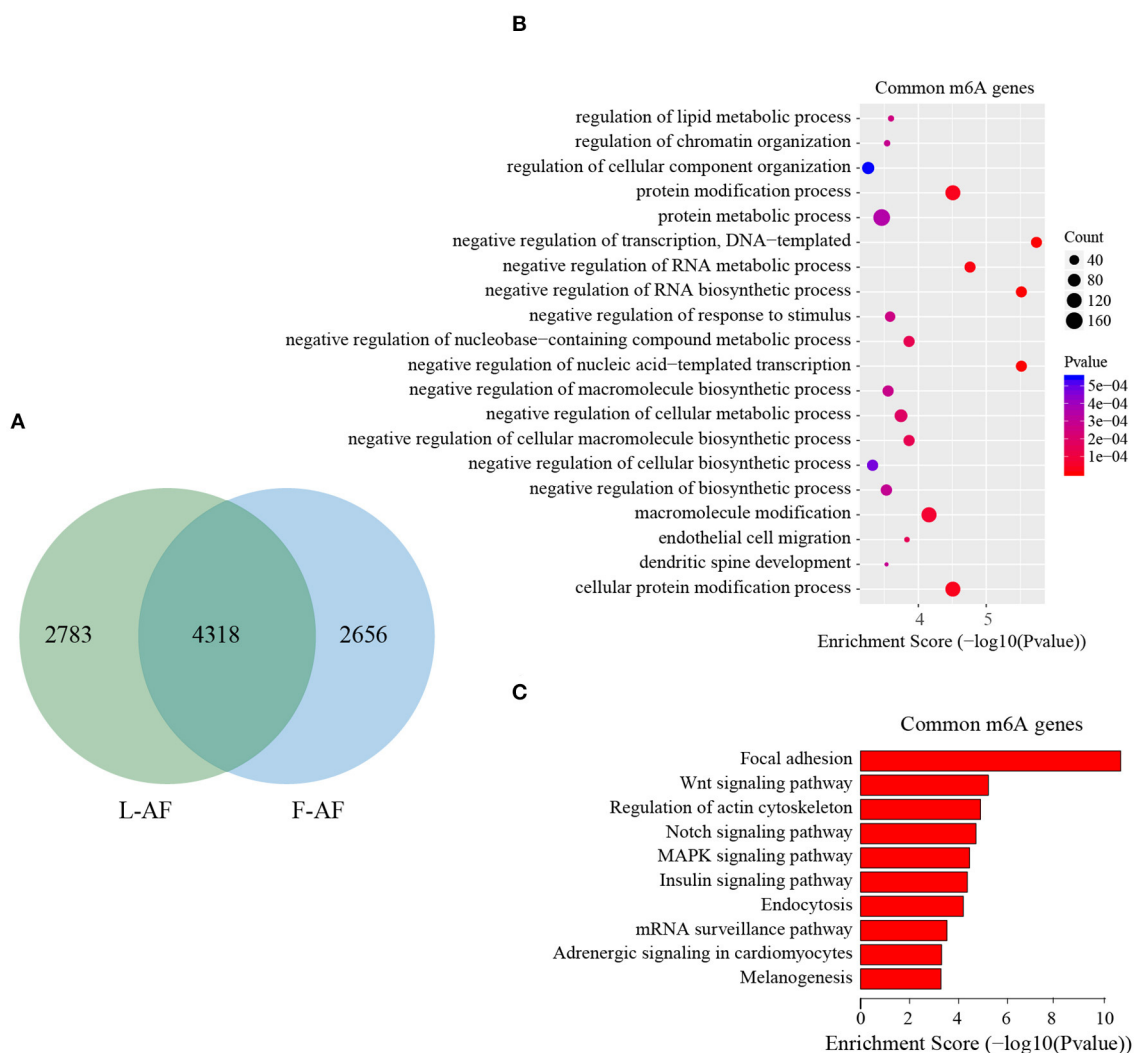


FIGURE 2 | GO biological process and KEGG pathway analyses of common m⁶A genes in broiler chickens. **(A)** Venn diagram showing overlap of the m⁶A peaks from L-AF and F-AF samples. **(B)** GO enrichment analysis of common m⁶A genes ($P < 0.05$). **(C)** Pathway analysis of common m⁶A genes ($P < 0.05$).

We discovered that the genes with high m⁶A peaks were mainly involved in the cellular responses to peptide hormone stimuli and lipogenesis-related pathways, including fatty acid biosynthesis and fatty acid metabolism ($P < 0.05$, **Figures 5A,B, Supplementary Data 6, 7**), while those with low m⁶A peaks were involved with development-associated processes, calcium-signaling pathway, steroid hormone biosynthesis, and others ($P < 0.05$, **Figures 5C,D, Supplementary Data 6, 7**).

Gene mRNA-Level Regulation by m⁶A Modification

To understand whether m⁶A modification can affect gene expression, we used the input RNA-seq data to investigate the differential expression of genes between the two chicken lines. In total, 352 high expression genes and 424 low expression genes in the fat line compared with the lean line were identified (**Supplementary Figure 2, Supplementary Data 8**). Of the 1,172

line-dynamic m⁶A genes in total, 146 (12.5%) showed mRNA-expression differences (**Supplementary Data 9**), indicating that the mRNA levels of these genes may be regulated by m⁶A modification. Among the 146 genes, it should be noted that the mRNA levels of 95% (52/55) of the high m⁶A methylated genes were high and were named “hyper-up” genes. Similarly, the mRNA levels of 92% (84/91) of the low m⁶A methylated genes were low and were named “hypo-down” genes. Only 10 of the 146 genes (7%) showed opposing mRNA expression and m⁶A-methylation trends, and these genes were termed hyper-down or hypo-up genes (**Figure 6A**). Interestingly, several lipogenesis-related genes showed differences in both m⁶A methylation and mRNA expression. For instance, acyl-CoA synthetase long-chain family member 1 (*ACSL1*), which is associated with fatty acid transport, showed significantly higher m⁶A methylation and mRNA levels in the fat birds than in the lean birds (**Figure 6B**); while lipin1 (*LPIN1*), which is associated with

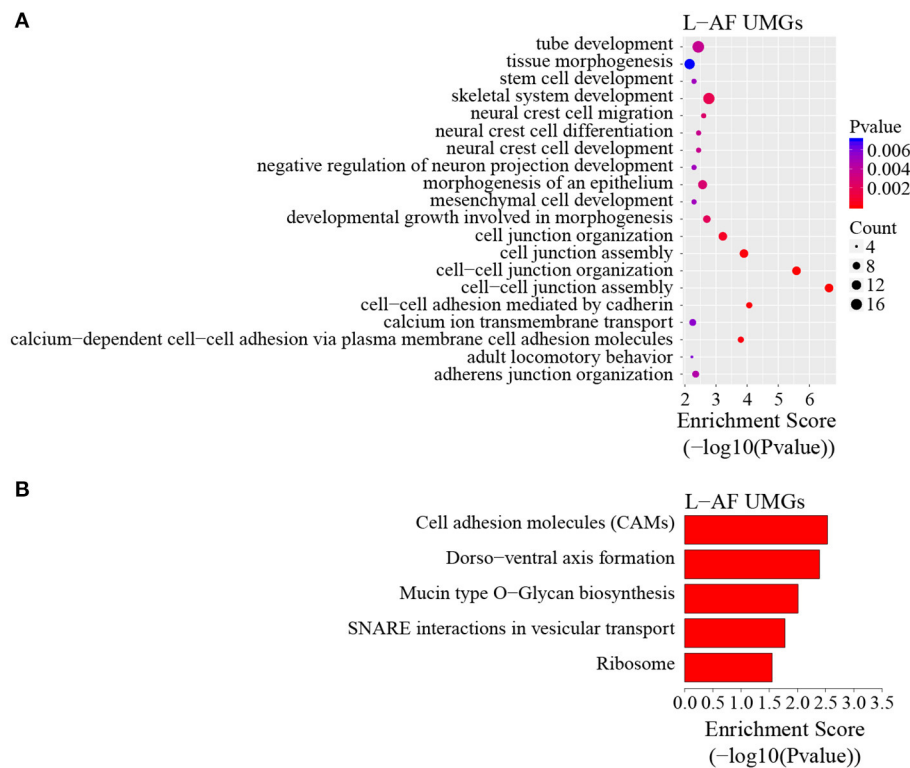


FIGURE 3 | GO biological process and KEGG pathway analyses of unique m⁶A genes in lean-line broiler chickens. **(A)** GO analysis of unique m⁶A genes in the lean line ($P < 0.05$). **(B)** Pathway analysis of unique m⁶A genes in the lean line ($P < 0.05$).

adipocyte differentiation, exhibited lower m⁶A methylation and mRNA expression levels in the fat birds than in the lean birds (Figure 6C). Finally, we further examined whether gene regulation in the chicken adipose tissue is correlated with the m⁶A modification by plotting the abundance of m⁶A peaks with the mRNA expression levels. As shown in Figure 7, the plot of m⁶A peak enrichment level vs. mRNA abundance revealed a negative correlation between global RNA methylation and gene expression in both chicken lines (L-AF: Pearson's $r = -0.9966$, $P < 0.0001$; F-AF: Pearson's $r = -0.9966$, $P < 0.0001$).

DISCUSSION

Adipose tissue is important for energy storage, endocrine functions, and the control of energy metabolism (McGown et al., 2014; Choe et al., 2016). Over the past few decades, the regulatory mechanisms of adipose tissue development and fat deposition, such as transcription factors, DNA methylation and histone modification, have been extensively studied, and a series of important progressions have been made (Farmer, 2006; Wang et al., 2010; Zhu et al., 2012). Recently, various chemical modifications of RNA, such as m⁶A, N¹-methyladenosine (m¹A), and 5-methylcytosine (m⁵C), have been reported to play important roles in many physiological and pathological processes, including embryonic development, spermatogenesis, and the occurrence and development of a variety of diseases

(Lin et al., 2017; Yang et al., 2019; Zhao et al., 2019). However, the role and underlying mechanism of RNA modification in adipose deposition are still uncharted territory. To this end, we conducted m⁶A methylome profiling of chicken adipose tissue using MeRIP-seq. To our knowledge, this is the first comprehensive high-throughput study to explore RNA modification in poultry adipose tissue. Our findings show the differences in m⁶A-modification patterns between the adipose tissue of fat and lean broilers. Further analysis suggested that m⁶A methylation may be an important factor in chicken adipose deposition via the regulation of gene expression.

Nearly 77 and 67% of mRNA transcripts underwent m⁶A methylation in L-AF and F-AF on average, respectively, suggesting that m⁶A plays a major role in adipose tissue development and fat deposition. In addition, the m⁶A peaks were primarily found in the highly conserved sequence motif GGACU (Figure 1B). M⁶A is generated by the binding of m⁶A methyltransferase to a highly conserved consensus sequence, GGACU (Shen et al., 2016). The RNA binding motif of METTL3, METTL14 and WTAP are GGAC, GGAC, and GACU, respectively (Liu et al., 2014). When the highly conserved GAC was mutated to GAU, m⁶A was no longer methylated in Rous sarcoma virus mRNA transcript (Kane and Beemon, 1987). A recent study showed that single nucleotide polymorphisms located at the GGAC positions could affect m⁶A methylation status (Zhang et al., 2020). Despite GGACU motif is important

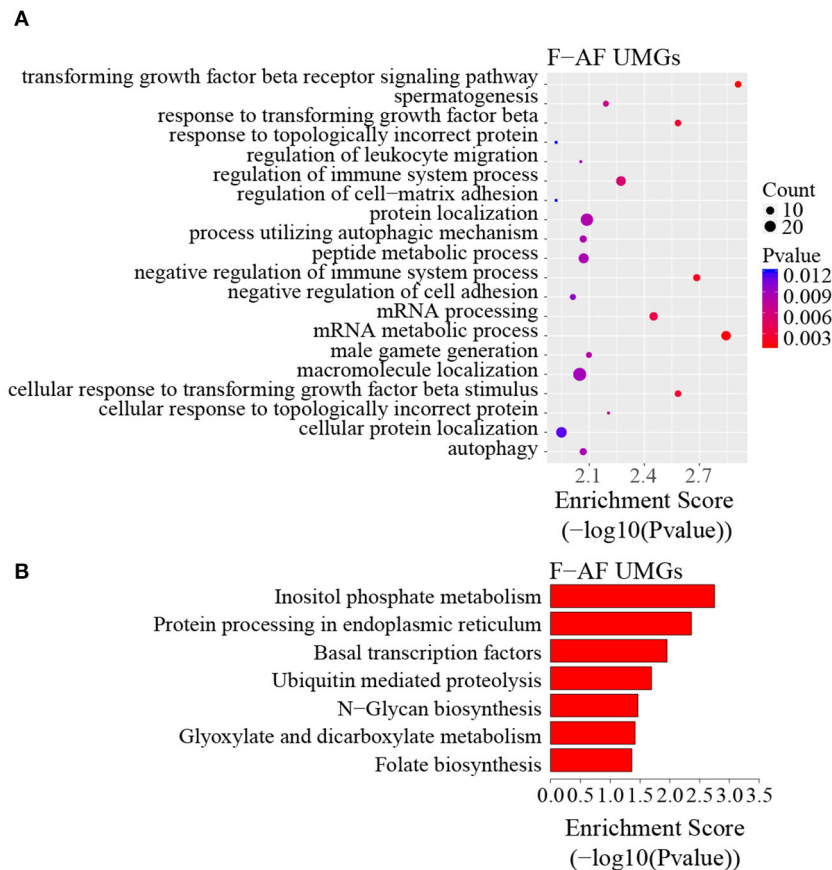


FIGURE 4 | GO biological process and KEGG pathway analyses of unique m⁶A genes in fat-line broiler chickens. **(A)** GO analysis of unique m⁶A genes in the fat line ($P < 0.05$). **(B)** Pathway analysis of unique m⁶A genes in the fat line ($P < 0.05$).

for the recognition by m⁶A methyltransferase, only a portion of GGACU sites are methylated *in vivo* (Gilbert et al., 2016), suggesting that the molecular mechanism regulating m⁶A modification needs to be further explored. In this study, interestingly, the m⁶A peaks were abundant not only in the CDS, stop codons, and 3'UTRs but also near the start codons (Figure 1D). This m⁶A-enrichment pattern is inconsistent with that of mammalian species (Meyer et al., 2012; Tao et al., 2017) but is similar to that of *Xenopus laevis* and *Arabidopsis thaliana* (Luo et al., 2014; Sai et al., 2020). This phenomenon may be attributed to differences in lipogenesis patterns between mammalian and birds (Gondret et al., 2001). It also seems to reflect the unique position of birds in the long evolutionary history of m⁶A modification in animals. In general, the predominance of m⁶A near stop codons and 3'UTRs has been found in most of the mRNAs of mammals, birds, amphibians, and plants (Luo et al., 2014; Tao et al., 2017; Fan et al., 2019; Sai et al., 2020), and this m⁶A-enrichment pattern may be representative of the typical mRNA m⁶A topology of eukaryotes. The high levels of m⁶A methylation in the 3'UTRs or near stop codons may be responsible for mRNA stability and alternative polyadenylation (Shen et al., 2016; Yue et al., 2018). Previous

TABLE 3 | General numbers of line-dynamic methylated peaks and associated genes.

High methylated peak ^a	High methylated gene ^a	Low methylated peak ^a	Low methylated gene ^a
435	334	1,069	838

^aFat line vs. lean line.

studies showed that m⁶A methylation in the CDS is likely to be associated with alternative splicing and translation efficiency (Zhao et al., 2014; Lin et al., 2019). Furthermore, the high m⁶A levels near the start codon may prevent mRNA degradation (Luo et al., 2014). In the present study, a negative relationship was observed between the global mRNA expression level and m⁶A methylation extent in chicken, which indicates m⁶A might affect chicken fat deposition at least in part through the regulation of mRNA stability.

The results of the GO and KEGG analyses in this study showed that m⁶A genes common to both broiler lines were significantly enriched in the processes and pathways associated

TABLE 4 | Top 15 high m⁶A peaks and associated genes.

Chromosome	txStart ^a	txEnd ^b	Gene name	Gene description	Fold change ^c
11	15359661	15360020	CMC2	Cytochrome c oxidase biogenesis	231.6
3	66300750	66301060	RPF2	Ribosomal large subunit assembly	94.2
12	17381938	17382188	CNTN3	Nervous development	87.1
1	51607928	51608248	NCF4	NADPH-oxidase complex assembly	86.4
Z	53573841	53574300	PDE6B	Signal transduction	84.4
8	28154046	28154232	ALG6	Glycosylation of lipids	78.0
5	5892541	5892820	QSER1	Nervous development	73.4
6	11187017	11187280	MYPN	Muscle contraction	57.9
6	18326706	18326912	SCD	Fatty acid biosynthesis	52.3
1	245141	246156	CD69L	Cell adhesion	44.7
23	4068703	4068954	GRIK3	Nervous development	32.7
3	71489501	71489822	PNISR	Pre-mRNA splicing	26.1
3	18088321	18088560	BROX	Protein ubiquitination	22.0
16	2069078	2069235	MICA	Immune response	20.3
7	23307358	23307791	IGFBP2	Signal transduction	20.1

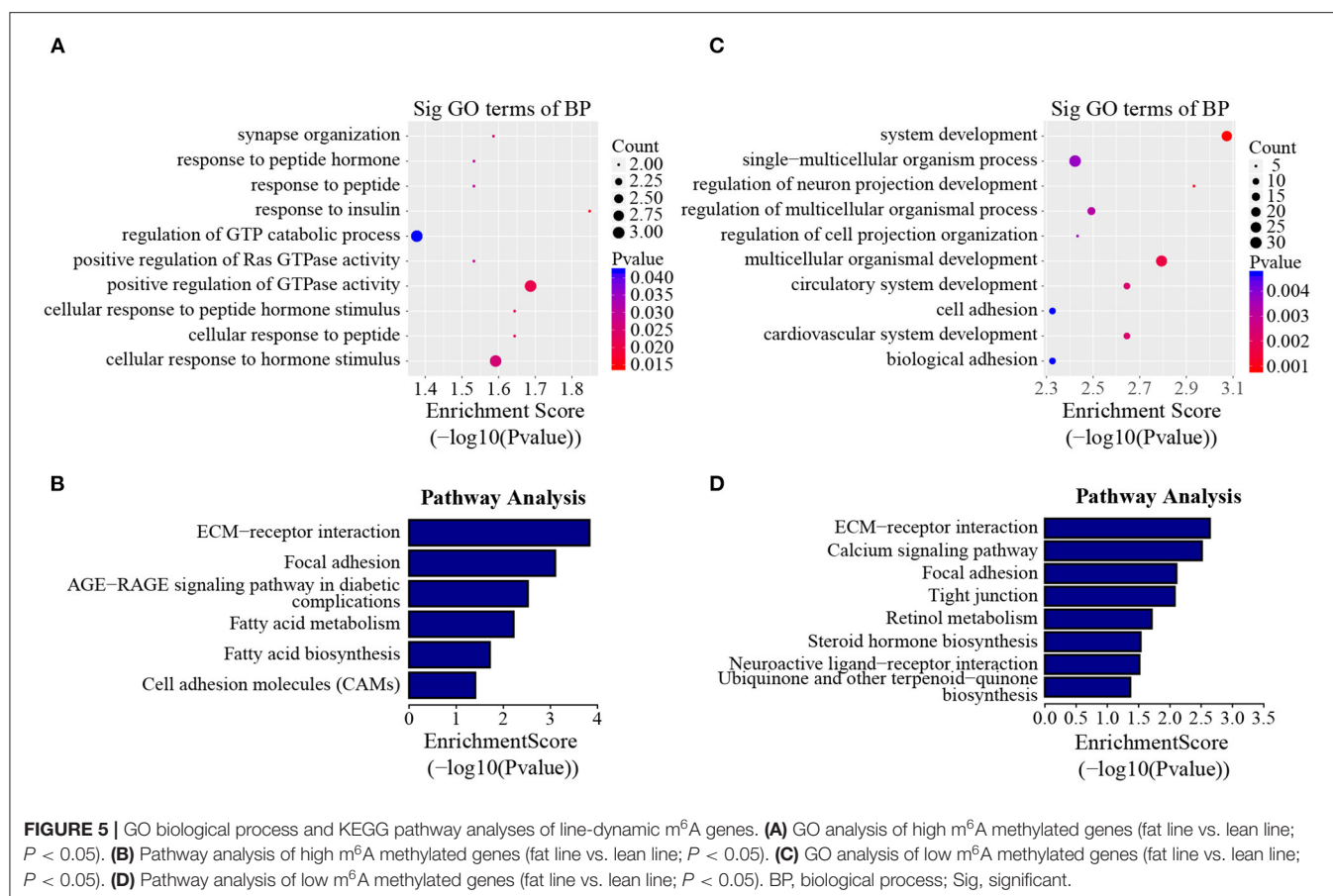
^aStart position of the high m⁶A peaks (fat line vs. lean line).^bEnd position of the high m⁶A peaks (fat line vs. lean line).^cRatio of m⁶A peak intensity in the fat line relative to the lean line.**TABLE 5 |** Top 15 low m⁶A peaks and associated genes.

Chromosome	txStart ^a	txEnd ^b	Gene name	Gene description	Fold change ^c
5	15161823	15162078	MUC2	O-glycan processing	2471.4
5	15157562	15157817	MUC2	O-glycan processing	1162.5
2	149536521	149536599	LOC107050437	Novel gene	1012.1
2	149532830	149533180	LOC107050437	Novel gene	616.0
20	9004330	9004481	EEF1A2	Elongation of translation	507.8
18	616601	616800	MYH1C	Muscle contraction	453.1
5	51897	52030	LOC107051134	Novel gene	435.8
3	22465749	22466142	KCNH1	Ion transport	431.3
27	7062978	7063300	GRB7	Signal transduction	413.1
8	24407796	24408005	TTC39A	Type 2 diabetes	363.5
3	59176501	59177047	SOGA3	Autophagy	362.8
14	6410912	6411084	MSLN	Cell adhesion	343.6
6	32348438	32348614	LOC112532717	Novel gene	275.2
6	5199540	5199624	SFTPA2	Surfactant-related functions	261.0
Z	10819061	10819246	SPEF2	Sperm axoneme assembly	239.2

^aStart position of the low m⁶A peaks (fat line vs. lean line).^bEnd position of the low m⁶A peaks (fat line vs. lean line).^cRatio of m⁶A peak intensity in the lean line relative to the fat line.

with adipose development and fat deposition, such as lipid metabolism and the Wnt-signaling pathway (**Figures 2B,C**), which is consistent with a previous study that showed that the genes commonly methylated in the backfat of both fat (Jinhua) and lean (Landrace) pigs were mainly involved in cellular lipid metabolic processes (Wang et al., 2018). This result supports the findings of the previous study, in which mRNA m⁶A modifications were relevant to tissue-specific functions (Li et al., 2014). In addition, the L-AF-unique m⁶A genes were primarily enriched in developmental-associated processes and,

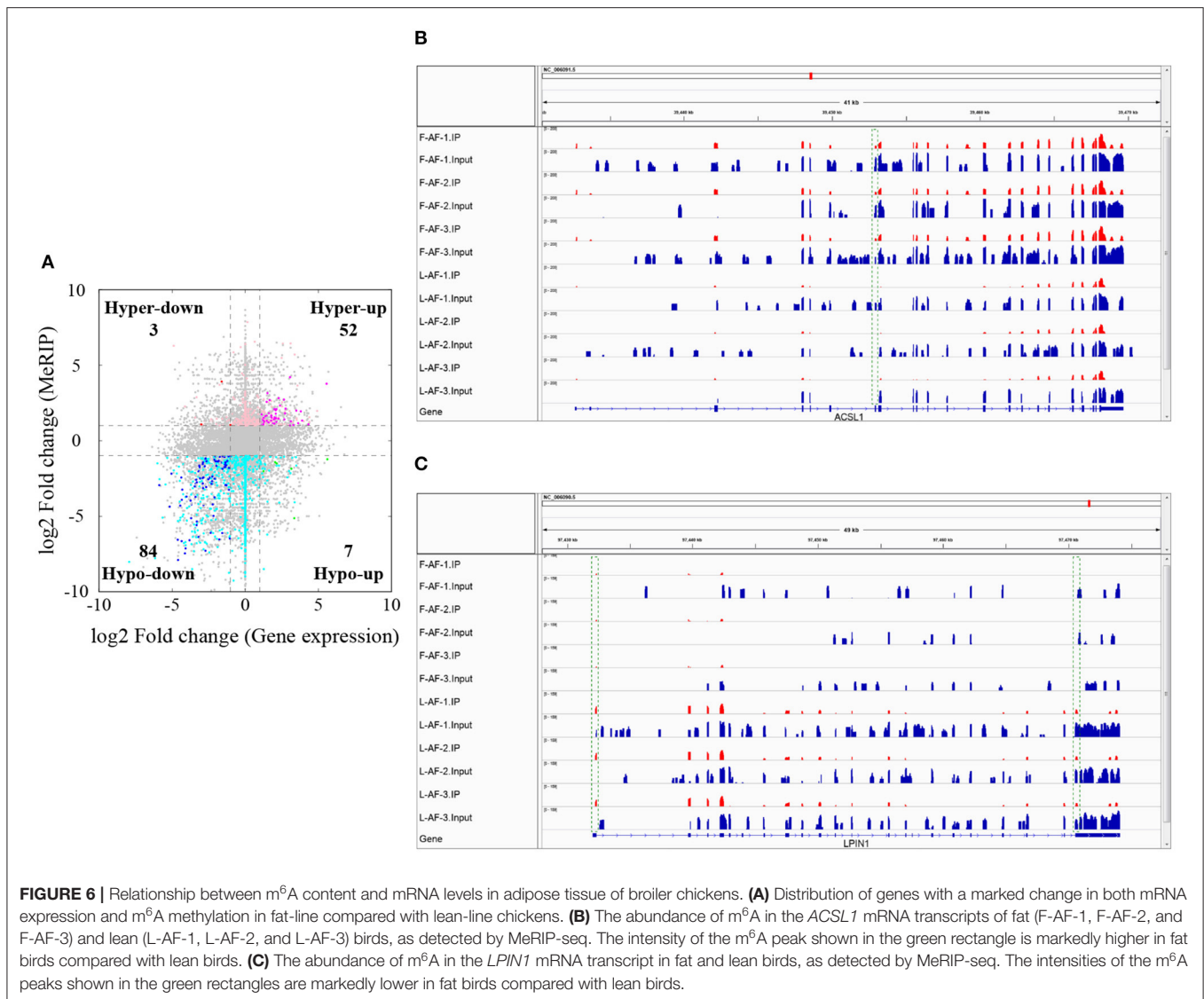
intriguingly, in “ribosome” (**Figures 3A,B**), which includes genes such as ribosomal protein S10 (*RPS10*), ribosomal protein L10a (*RPL10A*), and mitochondrial ribosomal protein L16 (*MRPL16*). This result is different from the previous study, which showed that the unique m⁶A genes in the backfat of fat Jinhua pigs were significantly involved in translational initiation and ribosomal large-subunit biogenesis (Wang et al., 2018). This phenomenon may be attributed to differences in lipogenesis patterns between mammalian and avian species (Gondret et al., 2001). However, F-AF-unique m⁶A genes were significantly



enriched in “mRNA processing” (Figure 4A), which includes genes such as pre-mRNA-processing factor 19 (PRPF19), cleavage- and polyadenylation-specific factor 6 (CPSF6), and CWC22 spliceosome-associated protein homolog (CWC22). The regulation of RNA metabolism by m⁶A modification depends on the recognition and binding of the specific m⁶A-reader proteins to the m⁶A sites (Yang et al., 2018). The m⁶A-reader protein YTH-domain-containing 1 (YTHDC1) has been shown to mediate alternative splicing, and YTH N6-methyladenosine RNA-binding protein 1 (YTHDF1) is responsible for enhancing translation efficiency (Yang et al., 2018). From our results, we speculated that the methylation of the mRNAs associated with mRNA processing and ribosome function might affect the expression of these same mRNAs, resulting in subsequent changes to global pre-mRNA splicing and protein synthesis, which might be another level of regulation involving alternative splicing and translation.

In mammals, an abundance of evidence has shown that m⁶A modification is involved in the regulation of adipose development and fat deposition (Zhang et al., 2015; Wu et al., 2017; Zong et al., 2019). Intriguingly, the results of the GO and KEGG analyses of the genes harboring dynamic methylated peaks showed that the high m⁶A methylated genes (fat line vs. lean line) were mainly involved in processes and pathways associated with lipid metabolism, such as fatty acid metabolism and fatty acid

biosynthesis (Figure 5B), which further supports the importance of m⁶A in obesity. For example, stearoyl coenzyme A desaturase (SCD) is related to fatty acid metabolism and was up-methylated approximately 52-fold in the fat birds compared with the lean birds. SCD is a rate-limiting enzyme that catalyzes the formation of unsaturated fatty acids (Ntambi, 1999). A genome-wide association study in pigs identified SCD as a major gene affecting fatty acid composition and intramuscular fat content (Ros-Freixedes et al., 2016), and a recent study showed that SCD may be important for chicken adipose deposition and metabolism (Mihelic et al., 2020). It worth mentioning that insulin-like growth factor binding protein 2 (IGFBP2) was up-methylated about 20-fold in the fat birds compared with the lean birds (Table 4), although IGFBP2 is not involved with the pathways associated with lipid metabolism. IGFBP2 is a cytokine secreted by differentiating white adipocytes and is regulated by DNA methylation in human abdominal obesity (Zhang et al., 2019). In a previous study, we found that IGFBP2 polymorphism (1196C>A) is significantly associated with AFW and AFP in the NEAUHLF population (Leng et al., 2009). In addition, we also found that the SNP 1196C>A within IGFBP2 3'UTR could influence its expression by affecting the regulation of gga-miR-456-3p (Yu et al., 2014). In this study, the m⁶A peak was found to be located in the CDS of IGFBP2, not 3'UTR. So we considered that the SNP 1196C>A may not affect the m⁶A methylation of

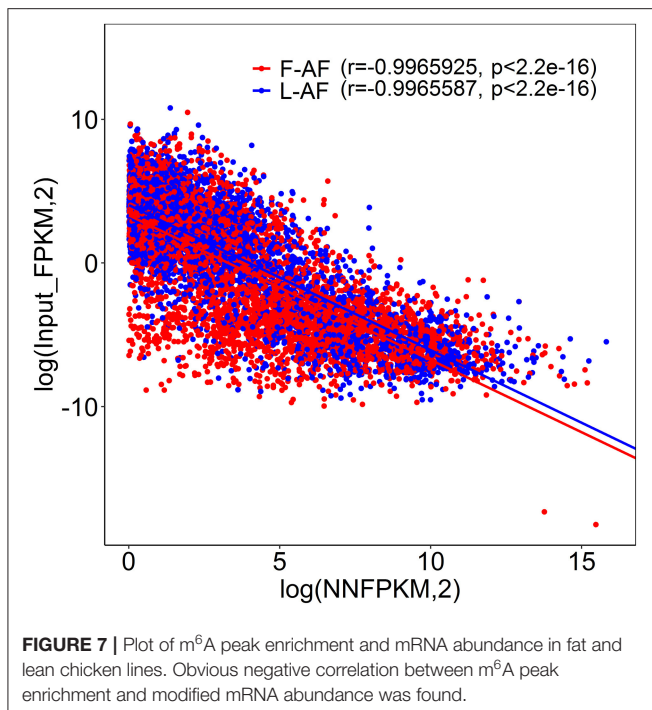


IGFBP2. In contrast to the high m⁶A methylated genes, the genes with low m⁶A peaks (fat line vs. lean line) were primarily related to developmental-associated processes, such as cardiovascular system development (Figure 5C), further reinforcing the theory that there is a close relationship between abdominal obesity and cardiovascular disease (Sahakyan et al., 2015). Therefore, we postulated that the different m⁶A-methylation patterns might reflect significant phenotypic changes between the fat and lean broiler lines.

The mRNA m⁶A modifications are recognized and bound by m⁶A-reader proteins, which include YTH-domain family member 2 (YTHDF2) and insulin-like growth factor 2 mRNA-binding proteins (IGF2BPs) and are involved in the regulation of mRNA stability. YTHDF2 was shown to mediate mRNA decay (Zhu et al., 2014), and IGF2BPs (including IGF2BP1/2/3) are responsible for enhancing mRNA stability (Huang et al., 2018). In the present study, the mRNA expression levels of many genes were found to be affected by their m⁶A levels (Figures 6A, 7). Based on this result, we speculated that the m⁶A sites in the

mRNA transcripts in chicken adipose tissue might be recognized and bound by YTHDF2 or IGF2BPs, thus changing the mRNA stabilities. However, further research is required to validate this hypothesis.

Our results show that several important lipogenic genes, including *ACSL1*, fatty acid synthase (*FASN*), *LPIN1*, and LDL receptor related protein 4 (*LRP4*), showed variations in both m⁶A methylation and mRNA expression. The expansion of adipose tissue mass is the result of an increase in the number of adipocytes and an increase in the size of individual fat cells. The number of adipocytes is determined by adipocyte differentiation (adipogenesis), while the size of the adipocytes is related to triglyceride (TG) accumulation in lipid droplets (Rosen and Spiegelman, 2006). *ACSL1* is an acyl-CoA synthetase and a long-chain fatty acid transport protein that can promote fatty acid uptake by adipocytes (Schaffer and Lodish, 1994; Tong et al., 2006). *FASN* is a key rate-limiting enzyme in *de novo* synthesis of fatty acids (Song et al., 2018). When there is excess energy in the body, most newly synthesized fatty acids are esterified



into TGs for storage (Song et al., 2018). Our findings showed that the m⁶A methylation and mRNA expression of *ACSL1* and *FASN* were higher in the fat birds compared with the lean birds, indicating that hypermethylation of *ACSL1* and *FASN* mRNA in the fat line might promote the formation of TGs by enhancing mRNA stability and, thus, increasing gene-expression levels. Several lines of evidence have shown that adipocytes are integral to energy metabolism regulation (Rondinone, 2006). Preadipocyte differentiation is controlled by a complex network of multiple transcriptional regulators, of which peroxisome proliferator-activated receptor gamma (PPAR γ) is the most important (Farmer, 2006). Research has shown that *LPIN1* interacts with and enhances the transcriptional activity of PPAR γ and promotes the differentiation of 3T3-L1 preadipocytes (Kim et al., 2016). In the current study, the m⁶A methylation and mRNA expression of *LPIN1* were higher in the lean birds compared with the fat birds, which is consistent with a previous study that found *LPIN1* expression levels were increased in the adipose tissue of lean subjects compared with the fat subjects (van Harmelen et al., 2007). This phenomenon might be due to the compensatory increase in *LPIN1* expression to maintain the balance of energy metabolism in the lean birds. We also found 10 genes showing opposing mRNA expression and m⁶A methylation patterns (Figure 6A). Interestingly, one of these, *LRP4*, is related to lipid metabolism; *LRP4* mRNA was low methylated and its expression level was high in the fat line compared with the lean line. *LRP4* is a transmembrane protein of the low-density lipoprotein receptor family (Alrayes et al., 2020). A recent study showed that the mice knockout of *LRP4* gene in adipocytes exhibit a reduction in adipocyte size and improved lipid and glucose homeostasis (Kim et al., 2019),

suggesting that *LRP4* is a positive regulator of adipocyte size. Therefore, it is conceivable that the hypomethylation of *LRP4* in the fat line might increase adipocyte size by promoting *LRP4* mRNA levels through a reduction in YTHDF2-mediated mRNA decay, although, further exploration is needed to shed light on this aspect. We propose that the m⁶A modifications within the mRNAs of *ACSL1*, *FASN*, *LPIN1*, and *LRP4* may be closely involved in adipose deposition and energy homeostasis in chickens. It is notable that, of the 1,172 line-dynamic m⁶A genes, most (87.5%) did not show mRNA-level variations between the fat and lean broiler lines. This phenomenon may be due to two reasons: (1) m⁶A may affect chicken abdominal fat deposition via other mechanisms, such as translation regulation, in addition to the regulation of mRNA stability; (2) Gene expression regulation is complex. Besides m⁶A methylation, the mRNA level of gene is influenced by various transcription and posttranscriptional regulatory factors, such as transcription factors (Farmer, 2006), transcription cofactors (Fabre et al., 2012), DNA methylation (Zhu et al., 2012), histone modification (Wang et al., 2010), chromatin remodeling (Siersbaek et al., 2011), and non-coding RNAs (Li et al., 2016; Losko et al., 2018).

CONCLUSION

In summary, we analyzed the m⁶A methylomes of chicken abdominal adipose tissues and proposed that m⁶A modification may play a key role in regulating the expression of genes contributing to lipid metabolism and adipogenesis. This comprehensive m⁶A map not only provides a basis for studying the roles of m⁶A methylation in chicken fat deposition but also opens a new avenue in the study of RNA epigenetics in adipobiology.

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: [NCBI SRA AND PRJNA657377].

ETHICS STATEMENT

The animal study was reviewed and approved by The Ministry of Science and Technology of the People's Republic of China (Approval Number: 2006-398) and were approved by the Laboratory Animal Management Committee and the Institutional Biosafety Committee of Northeast Agricultural University (Harbin, China).

AUTHOR CONTRIBUTIONS

BC contributed to the design of the experiments, carried out the experiments, performed the statistical analyses, and prepared the manuscript. LL contributed to the design of the experiments and helped with managing the birds. ZL, WW, YL, YJ, NW, and HL contributed to writing the manuscript.

SW conceived and designed the study, participated in data interpretation, and contributed to writing the manuscript. BC and LL contribute equally to this study. All authors gave final approval for publication.

FUNDING

This work was supported by the National Natural Science Foundation (No. 31902142), the National Natural Science Foundation (No. 31572394), and the China Agriculture Research System (No. CARS-41).

REFERENCES

- Alrayes, N., Aziz, A., Ullah, F., Ishfaq, M., Jelani, M., and Wali, A. (2020). Novel missense alteration in LRP4 gene underlies Cenani-Lenz syndactyly syndrome in a consanguineous family. *J. Gene Med.* 22:e3143. doi: 10.1002/jgm.3143
- Bailey, T. L. (2011). DREME: motif discovery in transcription factor ChIP-seq data. *Bioinformatics* 27, 1653–1659. doi: 10.1093/bioinformatics/btr261
- Choe, S. S., Huh, J. Y., Hwang, I. J., Kim, J. I., and Kim, J. B. (2016). Adipose tissue remodeling: its role in energy metabolism and metabolic disorders. *Front. Endocrinol.* 7:30. doi: 10.3389/fendo.2016.00030
- Dominissini, D., Moshitch-Moshkovitz, S., Schwartz, S., Salmon-Divon, M., Ungar, L., Osenberg, S., et al. (2012). Topology of the human and mouse m6A RNA methylomes revealed by m6A-seq. *Nature* 485, 201–206. doi: 10.1038/nature11112
- Fabre, O., Salehzada, T., Lambert, K., Seok, Y. B., Zhou, A., Mercier, J., et al. (2012). RNase L controls terminal adipocyte differentiation, lipids storage and insulin sensitivity via CHOP10 mRNA regulation. *Cell Death Differ.* 19, 1470–1481. doi: 10.1038/cdd.2012.23
- Fan, Y., Zhang, C. S., and Zhu, G. Y. (2019). Profiling of RNA N6-methyladenosine methylation during follicle selection in chicken ovary. *Poult. Sci.* 98, 6117–6124. doi: 10.3382/ps/pez277
- Farmer, S. R. (2006). Transcriptional control of adipocyte formation. *Cell Metab.* 4, 263–273. doi: 10.1016/j.cmet.2006.07.001
- Gilbert, W. V., Bell, T. A., and Schaening, C. (2016). Messenger RNA modifications: form, distribution, and function. *Science* 352, 1408–1412. doi: 10.1126/science.aad8711
- Gondret, F., Ferré, P., and Dugail, I. (2001). ADD-1/SREBP-1 is a major determinant of tissue differential lipogenic capacity in mammalian and avian species. *J. Lipid Res.* 42, 106–113. doi: 10.1016/S0022-2275(20)32341-5
- Guo, L., Sun, B., Shang, Z., Leng, L., Wang, Y., Wang, N., et al. (2011). Comparison of adipose tissue cellularity in chicken lines divergently selected for fatness. *Poult. Sci.* 90, 2024–2034. doi: 10.3382/ps.2010-00863
- Huang, H., Weng, H., Sun, W., Qin, X., Shi, H., Wu, H., et al. (2018). Recognition of RNA N(6)-methyladenosine by IGF2BP proteins enhances mRNA stability and translation. *Nat. Cell Biol.* 20, 285–295. doi: 10.1038/s41556-018-0045-z
- Jiang, Q., Sun, B., Liu, Q., Cai, M., Wu, R., Wang, F., et al. (2019). MTCH2 promotes adipogenesis in intramuscular preadipocytes via an m(6)A-YTHDF1-dependent mechanism. *FASEB J.* 33, 2971–2981. doi: 10.1096/fj.201801393RRR
- Kane, S., and Beemon, K. (1987). Inhibition of methylation at two internal N6-methyladenosine sites caused by GAC to GAU mutations. *J. Biol. Chem.* 262, 3422–3427. doi: 10.1016/S0021-9258(18)61520-0
- Kim, J., Lee, Y. J., Kim, J. M., Lee, S. Y., Bae, M. A., Ahn, J. H., et al. (2016). PPARgamma agonists induce adipocyte differentiation by modulating the expression of Lipin-1, which acts as a PPARgamma phosphatase. *Int. J. Biochem. Cell Biol.* 81, 57–66. doi: 10.1016/j.biocel.2016.10.018
- Kim, S. P., Da, H., Li, Z., Kushwaha, P., Beil, C., Mei, L., et al. (2019). Lrp4 expression by adipocytes and osteoblasts differentially impacts sclerostin's endocrine effects on body composition and glucose metabolism. *J. Biol. Chem.* 294, 6899–6911. doi: 10.1074/jbc.RA118.006769
- Kobayashi, M., Ohsugi, M., Sasako, T., Awazawa, M., Umehara, T., Iwane, A., et al. (2018). The RNA methyltransferase complex of WTAP, METTL3, and METTL14 regulates mitotic clonal expansion in adipogenesis. *Mol. Cell. Biol.* 38, e00116–e00118. doi: 10.1128/MCB.00116-18
- Leng, L., Wang, S., Li, Z., Wang, Q., and Li, H. (2009). A polymorphism in the 3'-flanking region of insulin-like growth factor binding protein 2 gene associated with abdominal fat in chickens. *Poult. Sci.* 88, 938–942. doi: 10.3382/ps.2008-00453
- Li, M., Sun, X., Cai, H., Sun, Y., Plath, M., Li, C., et al. (2016). Long non-coding RNA ADNCR suppresses adipogenic differentiation by targeting miR-204. *Biochim. Biophys. Acta. Gene Regul. Mech.* 1859, 871–882. doi: 10.1016/j.bbaggm.2016.05.003
- Li, Y. L., Wang, X. L., Li, C. P., Hu, S. N., Yu, J., and Song, S. H. (2014). Transcriptome-wide N-6-methyladenosine profiling of rice callus and leaf reveals the presence of tissue-specific competitors involved in selective mRNA modification. *RNA Biol.* 11, 1180–1188. doi: 10.4161/rna.36281
- Lin, X. Y., Chai, G. S., Wu, Y. M., Li, J. X., Chen, F., Liu, J. Z., et al. (2019). RNA m(6)A methylation regulates the epithelial mesenchymal transition of cancer cells and translation of Snail. *Nat. Commun.* 10:2065. doi: 10.1038/s41467-019-09865-9
- Lin, Z., Hsu, P. J., Xing, X. D., Fang, J. H., Lu, Z. K., Zou, Q., et al. (2017). Mettl3-/Mettl14-mediated mRNA N-6-methyladenosine modulates murine spermatogenesis. *Cell Res.* 27, 1216–1230. doi: 10.1038/cr.2017.117
- Liu, J. Z., Yue, Y. N., Han, D. L., Wang, X., Fu, Y., Zhang, L., et al. (2014). A METTL3-METTL14 complex mediates mammalian nuclear RNA N6-adenosine methylation. *Nat. Chem. Biol.* 10, 93–95. doi: 10.1038/nchembio.1432
- Losko, M., Lichawska-Cieslar, A., Kulecka, M., Paziewska, A., Rumienicz, I., Mikula, M., et al. (2018). Ectopic overexpression of MCPIP1 impairs adipogenesis by modulating microRNAs. *Biochim. Biophys. Acta. Mol. Cell Res.* 1865, 186–195. doi: 10.1016/j.bbamcr.2017.09.010
- Lu, N., Li, X. M., Yu, J. Y., Li, Y., Wang, C., Zhang, L. L., et al. (2018). Curcumin attenuates lipopolysaccharide-induced hepatic lipid metabolism disorder by modification of m(6)A RNA methylation in piglets. *Lipids* 53, 53–63. doi: 10.1002/lipd.12023
- Luo, G. Z., MacQueen, A., Zheng, G., Duan, H., Dore, L. C., Lu, Z., et al. (2014). Unique features of the m6A methylome in Arabidopsis thaliana. *Nat. Commun.* 5:5630. doi: 10.1038/ncomms5630
- Luo, Z. P., Zhang, Z. W., Tai, L. N., Zhang, L. F., Sun, Z., and Zhou, L. (2019). Comprehensive analysis of differences of N-6-methyladenosine RNA methylomes between high-fat-fed and normal mouse livers. *Epigenomics* 11, 1267–1282. doi: 10.2217/epi-2019-0009
- McGown, C., Bircard, A., and Younossi, Z. M. (2014). Adipose tissue as an endocrine organ. *Clin. Liver Dis.* 18, 41–58. doi: 10.1016/j.cld.2013.09.012
- Meyer, K. D., Saletore, Y., Zumbo, P., Elemento, O., Mason, C. E., and Jaffrey, S. R. (2012). Comprehensive analysis of mRNA methylation reveals enrichment in 3'UTRs and near stop codons. *Cell* 149, 1635–1646. doi: 10.1016/j.cell.2012.05.003
- Mihelcic, R., Winter, H., Powers, J. B., Das, S., Lamour, K., Campagna, S. R., et al. (2020). Genes controlling polyunsaturated fatty acid synthesis are

ACKNOWLEDGMENTS

The authors would like to thank the members of the poultry breeding group at Northeast Agricultural University for help in managing the birds and collecting the data.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fcell.2021.590468/full#supplementary-material>

- developmentally regulated in broiler chicks. *Br. Poult. Sci.* 61, 508–517. doi: 10.1080/00071668.2020.1759788
- Ntambi, J. M. (1999). Regulation of stearoyl-CoA desaturase by polyunsaturated fatty acids and cholesterol. *J. Lipid Res.* 40, 1549–1558. doi: 10.1016/S0022-2275(20)33401-5
- Perry, R. P., Kelley, D. E., Friderici, K., and Rottman, F. (1975). The methylated constituents of L cell messenger RNA: evidence for an unusual cluster at the 5' terminus. *Cell* 4, 387–394. doi: 10.1016/0092-8674(75)90159-2
- Rondinone, C. M. (2006). Adipocyte-derived hormones, cytokines, and mediators. *Endocrine* 29, 81–90. doi: 10.1385/ENDO:29:1:81
- Rosen, E. D., and Spiegelman, B. M. (2006). Adipocytes as regulators of energy balance and glucose homeostasis. *Nature* 444, 847–853. doi: 10.1038/nature05483
- Ros-Freixedes, R., Gol, S., Pena, R. N., Tor, M., Ibanez-Escriche, N., Dekkers, J. C. M., et al. (2016). Genome-wide association study singles out SCD and LEPR as the two main loci influencing intramuscular fat content and fatty acid composition in duroc pigs. *PLoS ONE* 11:e0152496. doi: 10.1371/journal.pone.0152496
- Sahakyan, K. R., Somers, V. K., Rodriguez-Escudero, J. P., Hodge, D. O., Carter, R. E., Sochor, O., et al. (2015). Normal-weight central obesity: implications for total and cardiovascular mortality. *Ann. Intern. Med.* 163, 827–835. doi: 10.7326/M14-2525
- Sai, L. L., Li, Y., Zhang, Y. C., Zhang, J., Qu, B. P., Guo, Q. M., et al. (2020). Distinct m(6)A methylome profiles in poly(A) RNA from *Xenopus laevis* testis and that treated with atrazine. *Chemosphere* 245:125631. doi: 10.1016/j.chemosphere.2019.125631
- Schaffer, J. E., and Lodish, H. F. (1994). Expression cloning and characterization of a novel adipocyte long chain fatty acid transport protein. *Cell* 79, 427–436. doi: 10.1016/0092-8674(94)90252-6
- Shen, L., Liang, Z., Gu, X., Chen, Y., Teo, Z. W. N., Hou, X., et al. (2016). N6-methyladenosine RNA modification regulates shoot stem cell fate in *Arabidopsis*. *Dev. Cell* 38, 186–200. doi: 10.1016/j.devcel.2016.06.008
- Siersbaek, R., Nielsen, R., John, S., Sung, M. H., Baek, S., Loft, A., et al. (2011). Extensive chromatin remodelling and establishment of transcription factor 'hotspots' during early adipogenesis. *Embo. J.* 30, 1459–1472. doi: 10.1038/emboj.2011.65
- Song, Z. Y., Xiaoli, A. M., and Yang, F. J. (2018). Regulation and metabolic significance of de novo lipogenesis in adipose tissues. *Nutrients* 10:1383. doi: 10.3390/nu10101383
- Tao, X. L., Chen, J. N., Jiang, Y. Z., Wei, Y. Y., Chen, Y., Xu, H. M., et al. (2017). Transcriptome-wide N-6-methyladenosine methylome profiling of porcine muscle and adipose tissues reveals a potential mechanism for transcriptional regulation and differential methylation pattern. *BMC Genomics* 18:336. doi: 10.1186/s12864-017-3719-1
- Tong, F. M., Black, P. N., Coleman, R. A., and DiRusso, C. C. (2006). Fatty acid transport by vectorial acylation in mammals: roles played by different isoforms of rat long-chain acyl-CoA synthetases. *Arch. Biochem. Biophys.* 447, 46–52. doi: 10.1016/j.abb.2006.01.005
- van Harmelen, V., Ryden, M., Sjolin, E., and Hoffstedt, J. (2007). A role of lipid in human obesity and insulin resistance: relation to adipocyte glucose transport and GLUT4 expression. *J. Lipid Res.* 48, 201–206. doi: 10.1194/jlr.M600272-JLR200
- Wan, Y. Z., Tang, K., Zhang, D. Y., Xie, S. J., Zhu, X. H., Wang, Z. G., et al. (2015). Transcriptome-wide high-throughput deep m(6)A-seq reveals unique differential m(6)A methylation patterns between three organs in *Arabidopsis thaliana*. *Genome Biol.* 16:272. doi: 10.1186/s13059-015-0839-2
- Wang, L. F., Jin, Q. H., Lee, J. E., Su, I. H., and Ge, K. (2010). Histone H3K27 methyltransferase Ezh2 represses Wnt genes to facilitate adipogenesis. *Proc. Natl. Acad. Sci. U. S. A.* 107, 7317–7322. doi: 10.1073/pnas.1000031107
- Wang, X., Sun, B., Jiang, Q., Wu, R., Cai, M., Yao, Y., et al. (2018). mRNA m(6)A plays opposite role in regulating UCP2 and PNPLA2 protein expression in adipocytes. *Int. J. Obes.* 42, 1912–1924. doi: 10.1038/s41366-018-0027-z
- Wu, W. C., Feng, J. E., Jiang, D. H., Zhou, X. H., Jiang, Q., Cai, M., et al. (2017). AMPK regulates lipid accumulation in skeletal muscle cells through FTO-dependent demethylation of N-6-methyladenosine. *Sci. Rep.* 7:41606. doi: 10.1038/srep41606
- Yang, Y., Hsu, P. J., Chen, Y. S., and Yang, Y. G. (2018). Dynamic transcriptomic m(6)A decoration: writers, erasers, readers and functions in RNA metabolism. *Cell Res.* 28, 616–624. doi: 10.1038/s41422-018-0040-8
- Yang, Y., Wang, L., Han, X., Yang, W. L., Zhang, M. M., Ma, H. L., et al. (2019). RNA 5-methylcytosine facilitates the maternal-to-zygotic transition by preventing maternal mRNA decay. *Mol. Cell* 75:1188. doi: 10.1016/j.molcel.2019.06.033
- Yu, Y. Y., Qiao, S. P., Sun, Y. N., Song, H., Zhang, X. F., Yan, X. H., et al. (2014). Identification and analysis of a functional SNP 1196C>A in 3'UTR of chicken IGFBP2 gene. *Prog. Biochem. Biophys.* 41, 1163–1172. doi: 10.3724/SP.J.1206.2014.00062
- Yue, Y., Liu, J., and He, C. (2015). RNA N6-methyladenosine methylation in post-transcriptional gene expression regulation. *Genes Dev.* 29, 1343–1355. doi: 10.1101/gad.262766.115
- Yue, Y. A., Liu, J., Cui, X. L., Cao, J., Luo, G. Z., Zhang, Z. Z., et al. (2018). VIRMA mediates preferential m(6)A mRNA methylation in 3'UTR and near stop codon and associates with alternative polyadenylation. *Cell Discov.* 4:10. doi: 10.1038/s41421-018-0019-0
- Zhang, H., Shi, X. R., Huang, T., Zhao, X. N., Chen, W. Y., Gu, N. N., et al. (2020). Dynamic landscape and evolution of m6A methylation in human. *Nucleic Acids Res.* 48, 6251–6264. doi: 10.1093/nar/gkaa347
- Zhang, K., Cheng, B. H., Yang, L. L., Wang, Z. P., Zhang, H. L., Xu, S. S., et al. (2017). Identification of a potential functional single nucleotide polymorphism for fatness and growth traits in the 3'-untranslated region of the PCSK1 gene in chickens. *J. Anim. Sci.* 95, 4776–4786. doi: 10.2527/jas2017.1706
- Zhang, M. Z., Zhang, Y., Ma, J., Guo, F. M., Cao, Q., Zhang, Y., et al. (2015). The demethylase activity of FTO (Fat Mass and Obesity Associated Protein) is required for preadipocyte differentiation. *PLoS ONE* 10:e0133788. doi: 10.1371/journal.pone.0133788
- Zhang, X. L., Gu, H. F., Frystyk, J., Efendic, S., Brismar, K., and Thorell, A. (2019). Analyses of IGFBP2 DNA methylation and mRNA expression in visceral and subcutaneous adipose tissues of obese subjects. *Growth Horm. IGF Res.* 45, 31–36. doi: 10.1016/j.gthir.2019.03.002
- Zhang, X. Y., Wu, M. Q., Wang, S. Z., Zhang, H., Du, Z. Q., Li, Y. M., et al. (2018). Genetic selection on abdominal fat content alters the reproductive performance of broilers. *Animal* 12, 1232–1241. doi: 10.1017/S1751731117002658
- Zhang, Y., Liu, T., Meyer, C. A., Eickhout, J., Johnson, D. S., Bernstein, B. E., et al. (2008). Model-based Analysis of ChIP-Seq (MACS). *Genome Biol.* 9:R137. doi: 10.1186/gb-2008-9-9-r137
- Zhao, X., Yang, Y., Sun, B. F., Shi, Y., Yang, X., Xiao, W., et al. (2014). FTO-dependent demethylation of N6-methyladenosine regulates mRNA splicing and is required for adipogenesis. *Cell Res.* 24, 1403–1419. doi: 10.1038/cr.2014.151
- Zhao, Y. S., Zhao, Q. J., Kholi, P. J., Shen, J., Li, M. X., Wu, X., et al. (2019). m1A regulated genes modulate PI3K/AKT/mTOR and ErbB pathways in gastrointestinal cancer. *Transl. Oncol.* 12, 1323–1333. doi: 10.1016/j.tranon.2019.06.007
- Zhou, H., Deeb, N., Evock-Clover, C. M., Ashwell, C. M., and Lamont, S. J. (2006). Genome-wide linkage analysis to identify chromosomal regions affecting phenotypic traits in the chicken. II. Body composition. *Poult. Sci.* 85, 1712–1721. doi: 10.1093/ps/85.10.1712
- Zhu, J. G., Xia, L., Ji, C. B., Zhang, C. M., Zhu, G. Z., Shi, C. M., et al. (2012). Differential DNA methylation status between human preadipocytes and mature adipocytes. *Cell Biochem. Biophys.* 63, 1–15. doi: 10.1007/s12013-012-9336-3
- Zhu, T. T., Roundtree, I. A., Wang, P., Wang, X., Wang, L., Sun, C., et al. (2014). Crystal structure of the YTH domain of YTHDF2 reveals mechanism for recognition of N6-methyladenosine. *Cell Res.* 24, 1493–1496. doi: 10.1038/cr.2014.152
- Zong, X., Zhao, J., Wang, H., Lu, Z. Q., Wang, F. Q., Du, H. H., et al. (2019). Mettl3 deficiency sustains long-chain fatty acid

absorption through suppressing Traf6-dependent inflammation response. *J. Immunol.* 202, 567–578. doi: 10.4049/jimmunol.1801151

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



m⁶A RNA Methylation Regulators Act as Potential Prognostic Biomarkers in Lung Adenocarcinoma

Hongbo Wang, Xiangxuan Zhao*† and Zaiming Lu**

Department of Radiology, Shengjing Hospital of China Medical University, Shenyang, China

OPEN ACCESS

Edited by:

Giovanni Nigita,
The Ohio State University,
United States

Reviewed by:

Borhane Guezguez,
German Cancer Research Center
(DKFZ), Germany
Xiaoli Ping,
Beckman Research Institute, City of
Hope, United States

*Correspondence:

Xiangxuan Zhao
xiangxuanzhao@163.com
Zaiming Lu
luzaiming@sina.com

†These authors have contributed
equally to this work

Specialty section:

This article was submitted to
Epigenomics and Epigenetics,
a section of the journal
Frontiers in Genetics

Received: 28 October 2020

Accepted: 15 January 2021

Published: 10 February 2021

Citation:

Wang H, Zhao X and Lu Z (2021) m⁶A
RNA Methylation Regulators Act as
Potential Prognostic Biomarkers in
Lung Adenocarcinoma.
Front. Genet. 12:622233.
doi: 10.3389/fgene.2021.622233

N⁶-methyladenosine [m(6)A/m⁶A] methylation is one of the most common RNA modifications in eukaryotic cell mRNA and plays an important regulatory role in mRNA metabolism, splicing, translocation, stability, and translation. Previous studies have demonstrated that the m⁶A modification is highly associated with tumor cell proliferation, migration, and invasion. In the present study, five m⁶A regulatory factors have been revealed, namely heterogeneous nuclear ribonucleoprotein A2/B1 (HNRNPA2B1), heterogeneous nuclear ribonucleoprotein C (HNRNPC), Vir like m⁶A methyltransferase associated protein (KIAA1429/VIRMA), RNA binding motif protein 15 (RBM15) and methyltransferase like 3 (METTL3), which are closely related to the overall survival (OS) of patients with lung adenocarcinoma (LUAD). These five m⁶A regulatory factors exhibited potential prognostic value for the 1, 3, and 5-years survival outcomes of LUAD patients. Our findings revealed that several signaling pathways, such as cell cycle, DNA replication, RNA degradation, RNA polymerase, nucleotide excision repair and basal transcription factors, are activated in the high-risk group of LUAD patients.

Keywords: lung cancer, N⁶-methylAdenosine (m⁶A), prognosis, epitranscriptomics, cancer biomarker

INTRODUCTION

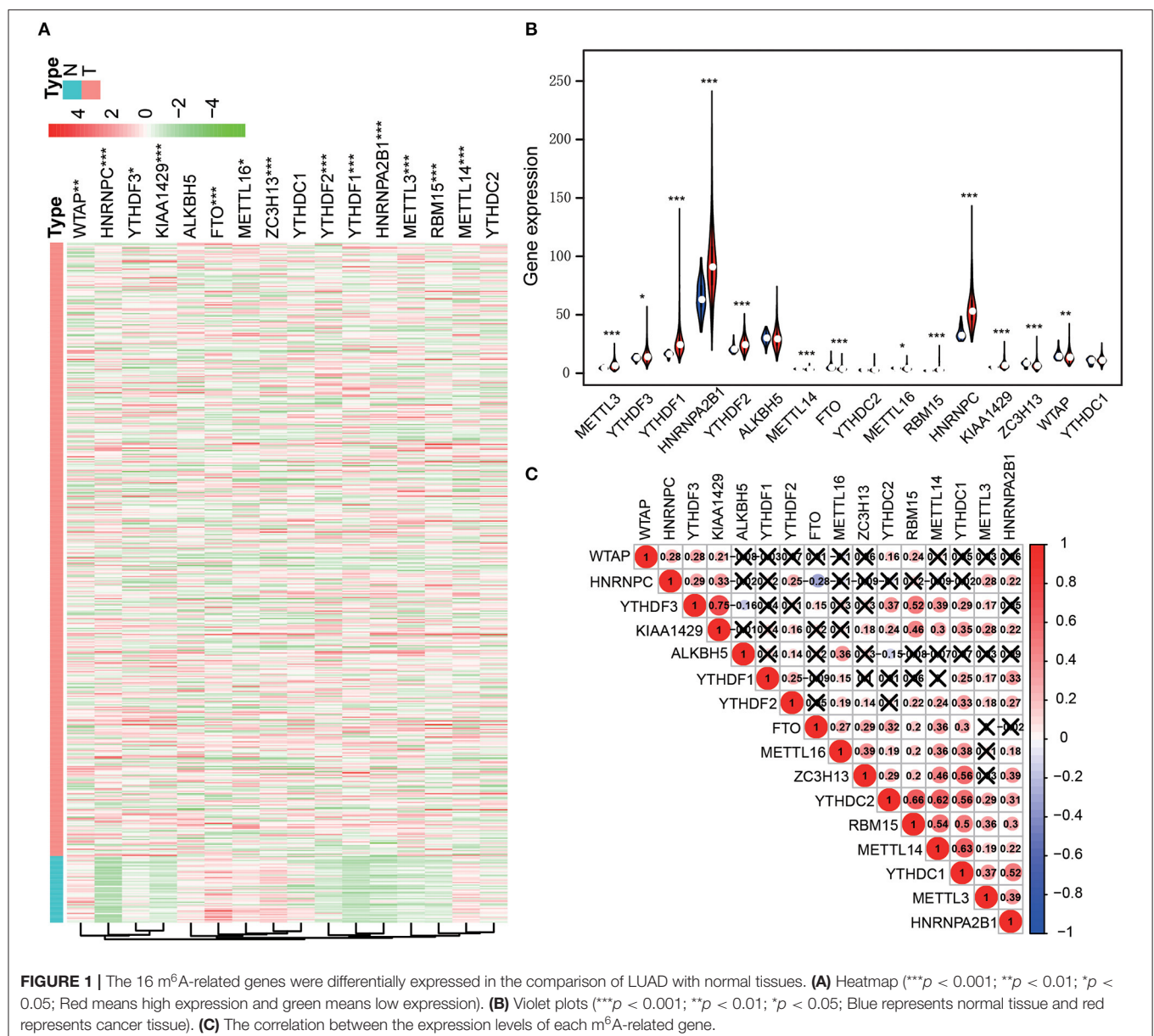
Lung cancer is currently one of the most common malignant tumors presenting the highest fatality rate among all malignancies (Siegel et al., 2020). In opposite to the surgical resection of early lung cancer, advanced lung cancer is mainly treated with radiotherapy and/or chemotherapy, while adjuvant immunotherapy and targeted therapy are also administered (Hirsch et al., 2017). Most patients have already advanced lung cancer by the time of their diagnosis mainly due to the limited knowledge in the pathogenesis of lung cancer, and the 5-year survival rate does not exceed 20% (Siegel et al., 2020). The clinical prognosis of patients is mainly based on tumor stage and other clinical indicators such as tumor node metastasis (TNM) stage. However, huge variation is observed in the final prognosis of the same tumor stage as a result of patients' heterogeneity. Therefore, relying on simple tumor staging may lead to poor prognosis accuracy, greatly affecting patients' further treatment and reducing the overall survival rate (OSR) (Razzouk, 2014; Perakis et al., 2016). The identification of accurate prognostic markers can contribute to the improvement of the treatment of lung cancer patients.

N⁶-methyladenosine (m⁶A) refers to the N⁶ terminal methylation of adenosine, which is a ubiquitous post-transcriptional modification mechanism of RNA in eukaryotic cells (Chen et al., 2019). m⁶A is involved in the RNA metabolism, and more specifically in mRNA translation, degradation, splicing, export, and folding (Liu et al., 2017; Chen et al., 2019; Liu and Gregory, 2019). The completion of m⁶A modification requires the binding of methyltransferase with the

conservative motif RRACH (R=A/G, H=U/A/C) in RNA (Kane and Beemon, 1985; Narayan et al., 1994; Balacco and Solter, 2019). m⁶A often occurs in the stop codons of the 3' untranslated (3'UTRs) and exon regions, respectively (Dominissini et al., 2012; Meyer et al., 2012). m⁶A modification is usually a reversible process that is regulated by various related factors (Jia et al., 2011, 2013). The m⁶A regulators reported so far can be divided into three types. The first type is called Writers including METTL3, METTL14, METTL16, WTAP, KIAA1429, RBM15, and ZC3H13, which are able to recognize RNA and modify m⁶A (Dai et al., 2018; Balacco and Solter, 2019). The second type is Erasers that include fat mass- and obesity-associated protein (FTO) and alk B homolog 5 (ALKBH5). These regulators are mainly responsible for removing m⁶A modifications (Liu et al., 2018;

Pan et al., 2018). The third type is Readers that consist of YTHDF1, YTHDF2, YTHDF3, YTHDC1, YTHDC2, HNRNPC, and HNRNPA2B1. Readers can recognize RNA methylation modifications and further regulate RNA processing, translation, and degradation (Wang et al., 2018; Ma et al., 2019).

Functional analysis has shown that m⁶A is crucial for cell proliferation, cell self-renewal, and apoptosis as it affects many important life processes (Zhou et al., 2019). A large number of studies have confirmed that the aberrant m⁶A modification plays a key role in the occurrence and progression of various tumors including LUAD (Zhou et al., 2019; Yi et al., 2020; Zhang et al., 2020). For instance, m⁶A Reader YTHDF2 can promote the non-small cell lung cancer (NSCLC) progression (Sheng et al., 2019), while the Eraser ALKBH5 can inhibit the metastasis of



NSCLC by inhibiting the miR-107/LATS2-mediated YAP activity (Jin et al., 2020). In addition, m⁶A status can also affect the sensitivity of NSCLC to Afatinib treatment (Meng et al., 2020). However, the potential value of m⁶A for the prognosis of lung cancer treatment still remains unexplored, especially for the prognosis of LUAD. The present study initially confirmed that the expression levels of five m⁶A regulators, including HNRNPA2B1, HNRNPC, KIAA1429, RBM15, and METTL3 were correlated with OS of LUAD patients. m⁶A Writers regulatory factors were also suggested as potential prognostic biomarkers for LUAD.

MATERIALS AND METHODS

Data Acquisition

The LUAD gene expression data and the corresponding clinical data were downloaded from The Cancer Genome Atlas database (TCGA) (<https://cancergenome.nih.gov/>) by using TCGA-assembler in February 2020¹. Gene expression in the downloaded files was normalized using the Fragments Per Kilobase of exon model per Million mapped fragments (FPKM) metric. Data for the m⁶A regulators including METTL3, METTL14, METTL16, WTAP, KIAA1429, RBM15, ZC3H13, FTO, ALKBH5, YTHDF1, YTHDF2, YTHDF3, YTHDC1, YTHDC2, HNRNPC, and HNRNPA2B1, were retrieved by mining the transcriptomics data of LUAD and para-carcinoma tissues. The human tissue expression levels in Genotype-tissue expression (GTEx) database were downloaded in May 2020. The GTEx dataset contains more than 900 organs and tissues of healthy people, with a total of more than 17,000 samples, covering 54 types of tissues in the human body.

Bioinformatics Analysis

Gene expression data of the tumor and control sample were separately sorted. Gene expression data for the 16 m⁶A related genes were then extracted and data with incomplete information were deleted. R software (Version 3.6.1) was used to perform differential expression analysis on the m⁶A regulatory factors in lung tissue samples in a comparison between 497 LUAD tissues (from 467 LUAD patients) and 54 para-carcinoma tissues. Vioplot tool was used to plot violin graphs to visualize the results of the differential expression analysis between LUAD patients and control samples. The Spearman correlation analysis was deployed to study the associations between the expression levels of the 16 m⁶A-related regulatory factors.

The Least Absolute Shrinkage and Selection Operator (LASSO) model is a dimensionality reduction method, which can reduce the number of variables through a penalty mechanism and ultimately achieve the goal of reducing bias. The LASSO model analysis was performed using the glmnet R package. Gene Set Enrichment Analysis (GSEA) software (Version 4.0.3 from Broad Institute official website homepage) was used to analyze the enrichment analysis between high-risk and low-risk groups. A total of

TABLE 1 | Clinical data of LUAD patients in TCGA database.

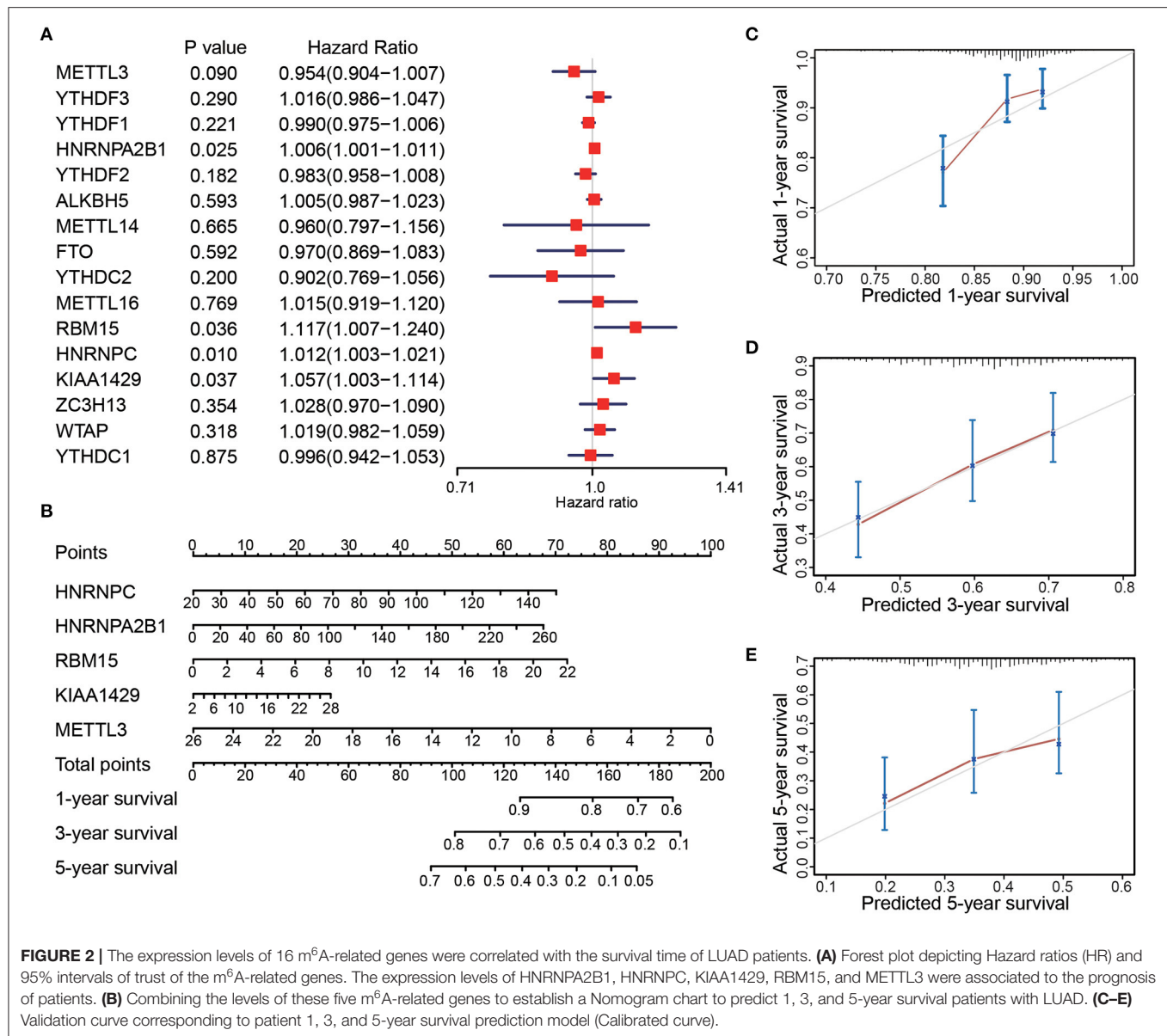
	Features	Numbers	Percentage %
Gender	Female	254	54.39
	Male	213	45.61
Age (Y)	65.01 ± 10.05		
	≤65	227	48.61
	>65	240	51.39
Stage	Stage I	251	53.75
	Stage II	108	23.13
	Stage III	75	16.06
	Stage IV	25	5.35
	Unknown	8	1.71
T	T1	162	34.69
	T2	244	52.25
	T3	39	8.35
	T4	19	4.07
	Unknown	3	0.64
M	M0	314	67.24
	M1	24	5.14
	Unknown	129	27.62
N	N0	300	64.24
	N1	87	18.63
	N2	66	14.13
	N3	2	0.43
	Unknown	12	2.57

55,268 genes were included in the analysis. Moreover, the “c2.cp.kegg.V7.0.symbols.gmt” analysis package was used to study the pathway enrichment. All the LUAD samples were divided into two groups (high-risk and low-risk groups) by median risk score. P-value and False Discovery Rate (FDR) thresholds of 0.05 and 0.25, respectively, were used to infer significant findings.

Statistical Analysis

One-way analysis of variance (Anova) test was used to compare the 16 m⁶A regulatory factors in between 497 LUAD tissues and 54 para-carcinoma tissues. The Spearman correlation was used to clarify the relationship between the m⁶A gene expression level and the basic clinical information (such as age, gender, TMN stage) of LUAD patients. The OSR is defined as the time period from diagnosis to death. Univariate and multivariate COX logistic regression models were used to analyze the prognostic potential of each factor and their ability to predict the survival outcome of LUAD. Kaplan-Meier and receiver operating characteristic (ROC) curves were plotted to demonstrate the prognostic performance of the explored m⁶A related regulators. A prognostic model was established by drawing a Nomogram plot, and calibration curves were used to verify it. P-value threshold of 0.05 was used to infer statistical significance.

¹The cancer genome atlas (TCGA) (2020). <https://portal.gdc.cancer.gov/>



RESULTS

m⁶A Regulator mRNA Levels

After screening the mRNA expression levels of the m⁶A regulators in 497 LUAD and 54 normal control samples, respectively, we analyzed the expressions of 16 m⁶A related regulators: METTL3, METTL14, METTL16, WTAP, KIAA1429, RBM15, ZC3H13, FTO, ALKBH5, YTHDF1, YTHDF2, YTHDF3, YTHDC1, YTHDC2, HNRNPC and HNRNPA2B1. Among them, METTL3, METTL14, KIAA1429, RBM15, ZC3H13, FTO, YTHDF1, YTHDF2, HNRNPC, HNRNPA2B1, WTAP, YTHDF3 and METTL16, were significantly overexpressed in LUAD tissues. The expression of ALKBH5, YTHDC1, and YTHDC2 presented no statistically significant differences (**Figures 1A,B**). Moreover, Pearson

correlation analysis was conducted between the 16 m⁶A related regulators. We found that the positive correlation between the expression levels of YTHDF3 and KIAA1429 was the highest one. Furthermore, significant positive correlations were revealed between YTHDC1, YTHDC2, RBM15, and METTL14 from the same analysis (**Figure 1C**).

m⁶A-Related Gene Expression and LUAD Prognosis

Clinical data of the 467 patients were further analyzed to explore the prognostic potential of the expression levels of the m⁶A regulatory factors in LUAD (**Table 1**). Survival analysis showed that the expression levels of four regulatory factors (HNRNPA2B1, HNRNPC, KIAA1429, and RBM15)

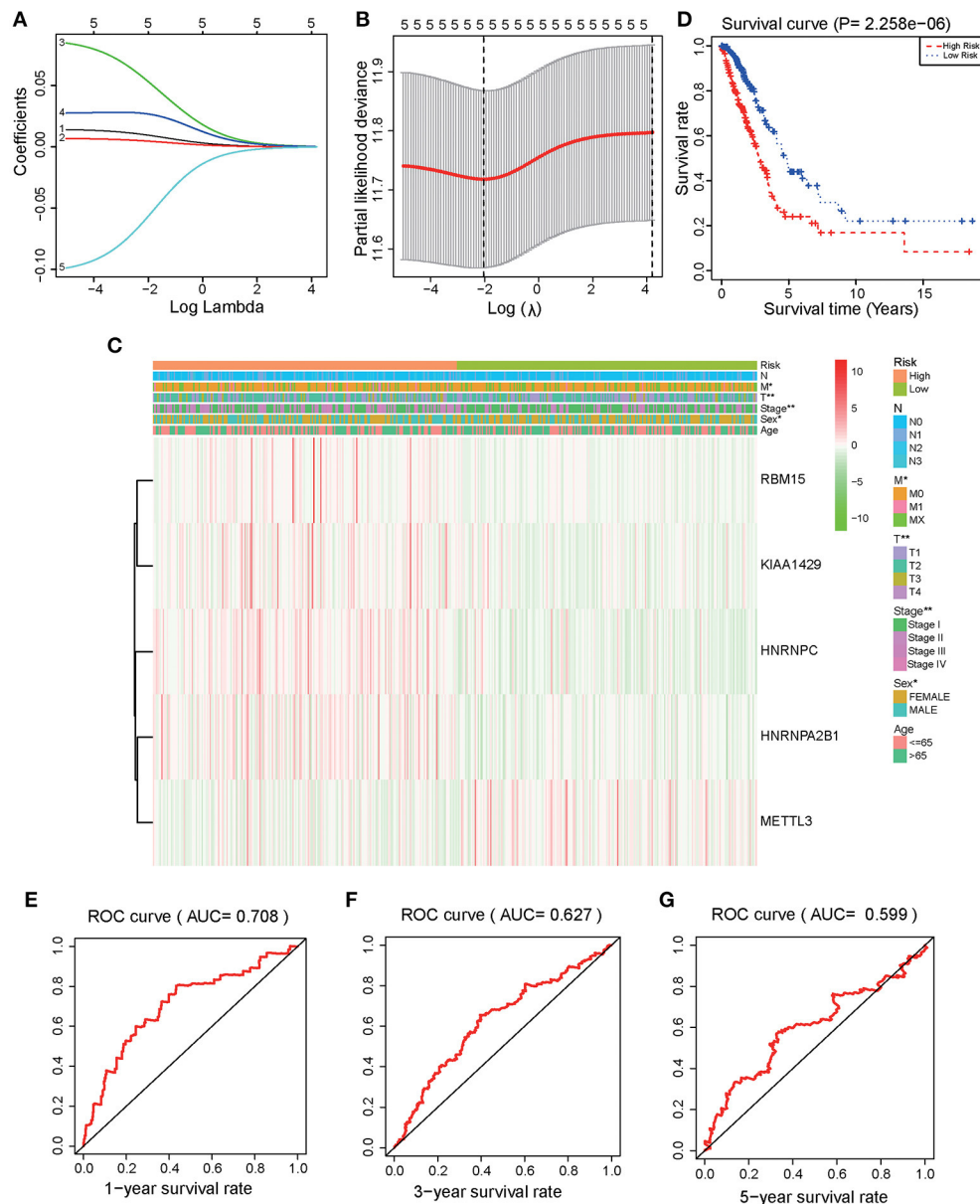


FIGURE 3 | Identification and validation of a risk model to prognose LUAD patients. **(A)** The LASSO analysis model was verified by repeated calculation of the 16 m⁶A genes. **(B)** LASSO coefficient profiles. **(C)** According to LASSO risk factors, LUAD patients were divided into high-risk group and low-risk group. Heatmap demonstrated the expression levels of HNRNPA2B1, HNRNPC, KIAA1429, RBM15, and METTL3 in the two assessed groups. **(D)** Kaplan Meier analysis for the survival of LUAD patients. **(E)** 1-year ($P < 0.001$), **(F)** 3-year ($P < 0.001$), **(G)** 5-year OS ROC curves ($P < 0.001$).

were significantly positively associated with the patient's death risk, while METTL3 was negatively associated with it without reaching statistical significance ($0.05 < p < 0.1$) (Figure 2A). Then, a Nomogram prediction model based on the expression levels of the above five genes and patient's outcomes data were established (Figure 2B). Four hundred and twenty cases were randomly selected from all cases and divided into three groups with each one of them having 140 cases. The results of the calibration curves indicated that the Nomogram prediction model presented

good predictive potential for 1, 3, and 5-year OS of LUAD patients (Figures 2C–E).

LASSO Regression and Risk Co-Efficient

LASSO regression models were used to analyze the risk coefficient and risk value of the expression of the five m⁶A regulators (HNRNPA2B1, HNRNPC, KIAA1429, RBM15, METTL3) in OS prediction (Figures 3A,B). The patients of the present study were divided into high and low-risk groups based on their predicted risk scores. HNRNPA2B1, HNRNPC,

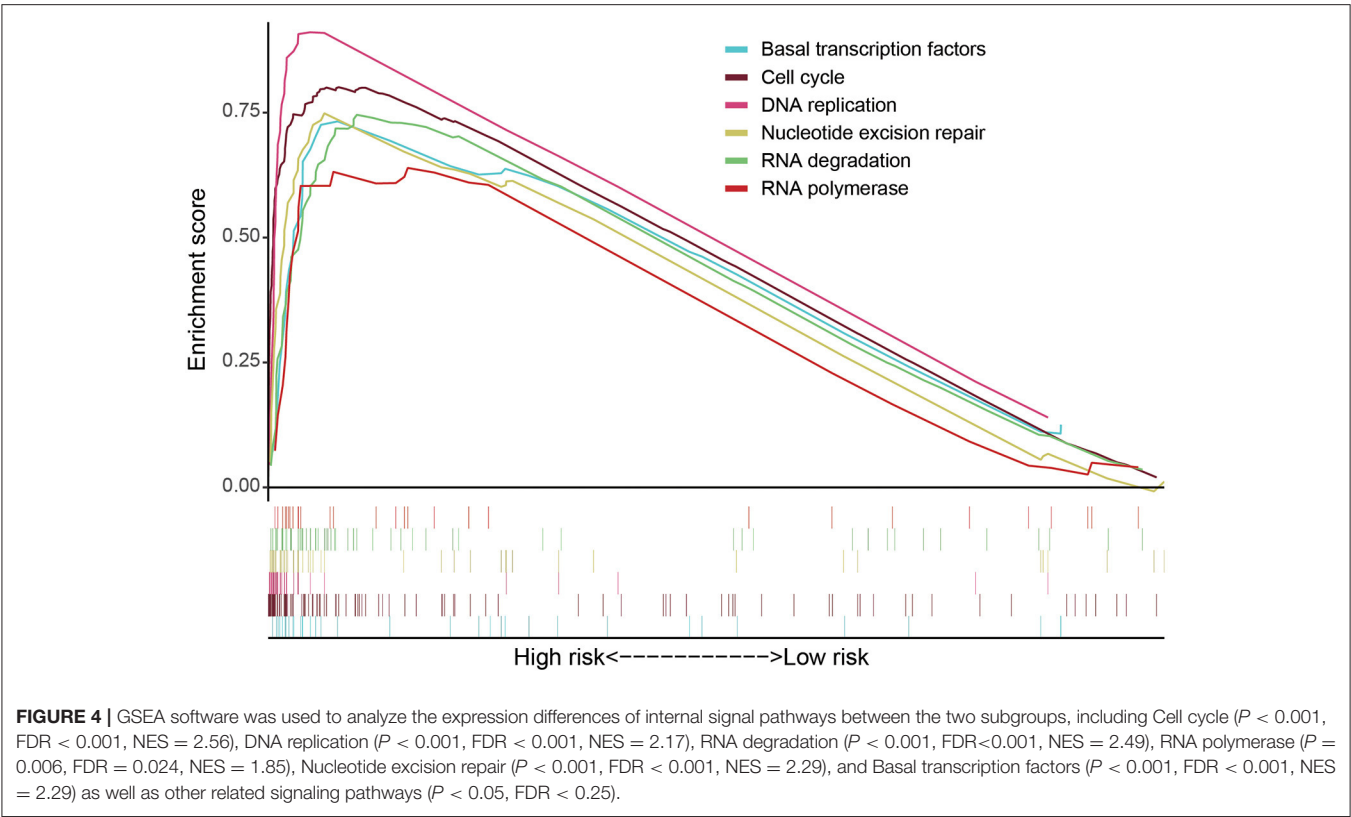


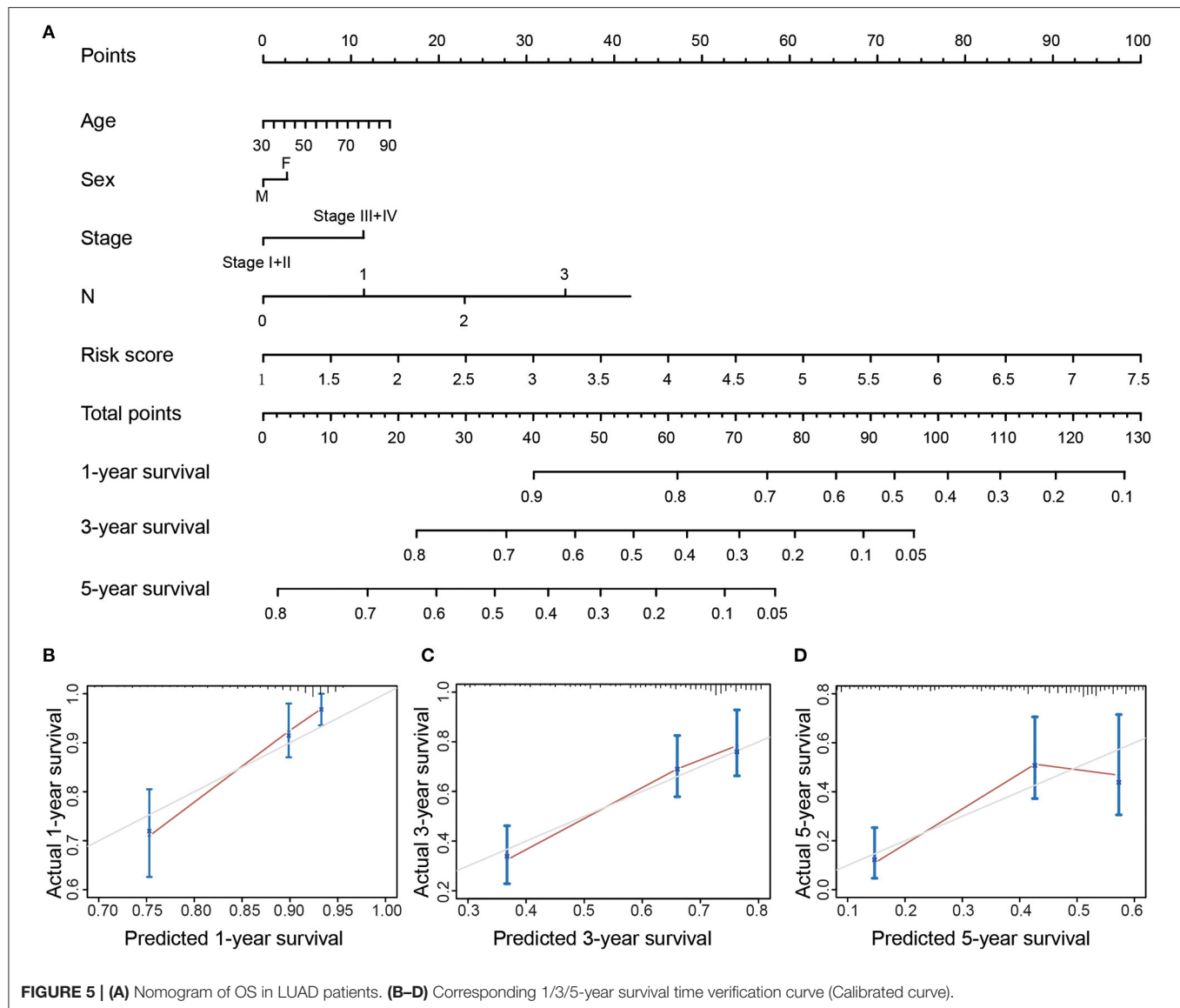
TABLE 2 | Univariate and multivariate COX regression analyses were used to assess the association between the clinical data and risk score of LUAD patients and the prognosis.

	Univariate COX regression		Multivariate COX regression	
	Hazard ratio	P	Hazard ratio	P
Age	1.002 (0.983–1.021)	0.843		
Gender	1.035 (0.717–1.495)	0.852		
Stage	1.654 (1.401–1.951)	<0.001	1.324 (1.056–1.660)	0.015
T	1.632 (1.315–2.024)	<0.001	1.064 (0.836–1.353)	0.615
M	1.757 (0.964–3.203)	0.066		
N	1.790 (1.459–2.196)	<0.001	1.374 (1.046–1.805)	0.022
Risk score	1.793 (1.465–2.195)	<0.001	1.658 (1.331–2.056)	<0.001

KIAA1429, and RBM15 were found to be overexpressed in the high-risk group, while METTL3 was overexpressed in the low-risk group (Figure 3C). Furthermore, survival analysis was conducted using the combined risk value. Results confirmed that the prognosis of patients in the high-risk group was significantly worse than the one in the low-risk group ($P < 0.01$) (Figure 3D). ROC curves for the 1, 3, and 5-year survival prediction demonstrated that the risk value possesses high prognostic accuracy with area under the curve (AUC) of 0.60–0.71 (Figures 3E–G).

KEGG and Multi-Factor Analysis

GSEA software was used to analyze pathway enrichment for the genes that are differentially expressed between high and low-risk patients. The cell pathway was found to be enriched in the set of deregulated factors with a p-value threshold of 0.05 and a FDR threshold of 0.25. The pathway enrichment analysis using KEGG database pathways revealed that Cell cycle ($P < 0.001$, FDR<0001, Normalized Enrichment Score (NES)=2.56), DNA replication ($P < 0.001$, FDR<0001, NES = 2.17), RNA degradation ($P < 0.001$, FDR<0001, NES = 2.49), RNA polymerase ($P = 0.006$, FDR=0.024, NES=1.85), Nucleotide excision repair ($P < 0.001$, FDR<0001, NES=2.29), Basal transcription factors ($P < 0.001$, FDR<0001, NES=2.29), and other related signaling pathways are significantly activated in the high-risk group ($P < 0.05$; FDR<0.25) (Figure 4). Accordingly, a heatmap of the most enriched genes for each identified KEGG pathway by GSEA between high and low-risk groups was presented in Supplementary Figures 1–6 and Supplementary Tables 1–6. Univariate and multivariate COX regression analyses were performed based on existing risk factors and patient clinical information (such as age, gender, and tumor and TMN staging) to evaluate their prognostic potential in LUAD (Table 2). The results of univariate analysis suggested that the stage ($P < 0.001$), T stage ($P < 0.001$), lymph node metastasis stage ($P < 0.001$) and risk score ($P < 0.001$) of LUAD patients are significantly negatively associated with the patient’s OS. The results of

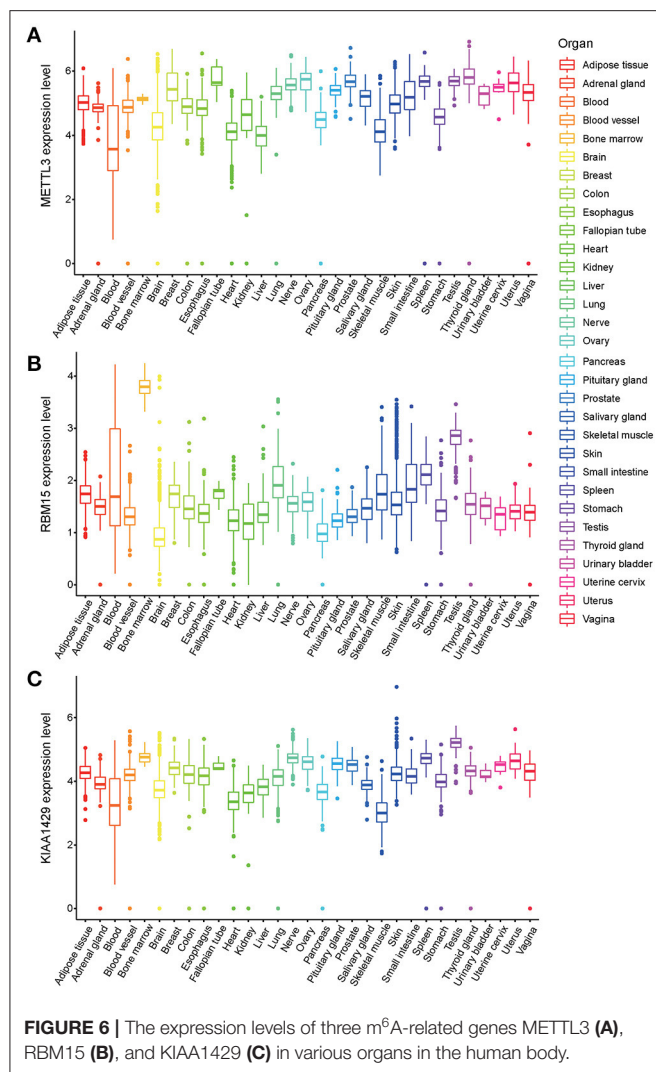


multivariate analysis suggested that the tumor stage ($P = 0.015$), lymph node metastasis stage ($P = 0.022$) and risk score ($P < 0.001$) are significantly negatively associated with the patient's OS.

A Nomogram prognostic analysis model was established based on clinical data such as age, gender, stage, lymph node metastasis and risk value (Figure 5A). The analysis of the 1, 3 and 5-year OS of patients through the Nomogram prediction model has demonstrated that the risk value contributes most to the prediction model followed by lymph node metastasis, age and tumor stage. Three hundred cases were randomly selected from all cases and were then divided into three groups with each one of them having 100 cases. Calibration curves showed that the established Nomogram prediction model presented good predictive potential for the 1, 3, and 5-year OS of LUAD patients (Figures 5B–D).

m⁶A Writers in Normal Human Organ Tissues

METTL3, RBM15, and KIAA1429 have been grouped to the m⁶A Writers type (Wang et al., 2018; Chen et al., 2019; Chen and Wong, 2020). LASSO regression analysis showed that the m⁶A Writers, METTL3, RBM15 and KIAA1429, present higher weight co-efficients (Co-ef) (METTL3: -0.0576562669796008 ; KIAA1429: 0.0269410278179687 ; RBM15: 0.0539704827385957) than those of HNRNPA2B1 (0.00473147643475602) and HNRNPC (0.00964499370244368). The expression levels of three genes (METTL3, RBM15, and KIAA1429) in 54 human normal organ tissues were compared and analyzed using the GTEx database. No significant difference was found in the comparison between the expression levels of METTL3 and KIAA1429 in various tissues of the human body (Figures 6A,B), while RBM15 was overexpressed in bone marrow and testis compared to other organs in the human body (Figure 6C).



DISCUSSION

The present study has discussed the m⁶A regulatory factor-related genes are associated to the overall prognosis of LUAD patients, and the newly introduced prognostic model is proven to accurately prognose the outcomes of LUAD patients. Meanwhile, the m⁶A regulatory factor-related genes are associated with the occurrence and development of LUAD.

m⁶A is one of the most common RNA modifications discovered so far. Accumulating studies have recently shown that the deregulation of the m⁶A RNA modification plays an important role in the occurrence and progression of tumors (Cui et al., 2017; Yang et al., 2017; Dai et al., 2018). For instance, the overexpression of FTO in acute myeloid leukemia (AML) can inhibit the m⁶A levels of Ankyrin repeat and SOCS box containing 2 (ASB2) and of retinoic acid receptor α (RAR α) mRNA, resulting in the occurrence and progression of AML (Li et al., 2017). The low expression of m⁶A regulators METTL14

in HCC (Ma et al., 2017) and the overexpression of the m⁶A regulator ALKBH5 in glioblastoma (Zhang et al., 2017) have been both associated with poor prognosis. Thus, the abnormal modification of m⁶A is closely related to tumor progress, metastasis and survival prognosis.

The expression levels of a variety of genes involved in RNA methylation mechanisms are closely related to the prognosis of lung cancer patients (Sun et al., 2020). m⁶A methylation is an important RNA modification that occurs in various RNA types such as microRNAs (miRs), circRNAs, and lncRNAs. At the same time, a large number of studies have confirmed that m⁶A and tumor progression are strongly correlated (Ma et al., 2019). The levels of m⁶A-related genes are also tightly associated with the prognosis of lung cancer patients. For instance, Liu and colleagues have confirmed that the expression level of m⁶A is weakly correlated with the prognosis of patients with lung squamous cell carcinoma, while it has a strong correlation with the prognosis of patients with LUAD (Liu et al., 2020). A most recent report from Zhuang et al. have reported that the differences in the expression of m⁶A regulators not only have certain diagnostic significance for early lung cancer, but are also closely related to the prognosis of LUAD patients (Zhuang et al., 2020).

In the present study, we have further expanded the number of m⁶A regulatory factors and explored their gene expression levels are related to the prognosis of patients with LUAD. Five m⁶A modification regulators, namely HNRNPA2B1, HNRNPC, KIAA1429, RBM15, and METTL3 are found to be closely related to the prognosis of LUAD patients. LASSO analysis reveals that the outcomes of patients in the high-risk group are significantly worse than the ones in the low-risk group. Both univariate and multivariate analyses conclude that the risk value, stage, and lymph node metastasis are closely related to the patient's prognosis. Finally, a newly developed Nomogram model is able to improve the accuracy of the prognosis of patients compared to conventional risk factors.

The m⁶A writers, METTL3, RBM15, and KIAA1429, were found to be linked with a higher risk in the LASSO regression analysis. Therefore, the abnormal expression of m⁶A writer regulatory factors may affect the prognosis of LUAD. Among them, METTL3 belongs to the class I methyltransferases family and is a predominantly catalytic enzyme in m⁶A modification. METTL3 has been confirmed to be abnormally expressed in a variety of tumors and is believed to be involved in carcinogenesis (Zheng et al., 2019). Prior studies have validated that METTL3 is overexpressed in lung cancer. METTL3 can interact with the transcription factor eIF3h to promote the translation of oncogenes and ultimately to catalyze and accelerate tumor growth and metastasis (Lin et al., 2016; Choe et al., 2018). METTL3 can promote the splicing of the miR-143-3p precursor, which in turn activates the miR-143-3p/VASH1 axis and ultimately leads to the progression and metastasis of lung cancer (Wang et al., 2019). METTL3 has also been shown to increase the expression level of JUNB and promote the occurrence of epithelial-mesenchymal transition (Wanna-Udom et al., 2020). miR-600 (Wei et al., 2019) and miR-33a (Du et al., 2017) is able to inhibit the expression of METTL3, thereby ablating the progression of NSCLC. Thus, METTL3 plays

an important role in the occurrence and development of lung cancer. One of the main contributions of the present study is the validation of the hypothesis that the overexpression of METTL3 is negatively correlated with the prognosis of LUAD. This is consistent with the results of previous studies suggesting that the overexpression of METTL3 often indicates poor prognosis in patients with primary liver cancer (Chen et al., 2018).

m⁶A modification regulatory RBM15 is a member of the split end protein (SPEN) family and can bind with METTL3 and WTAP (Wang et al., 2020). Studies have confirmed that m⁶A plays an important regulatory role in the lncRNA XIST-mediated gene transcription silencing. RBM15 catalyzes the recognition of the m⁶A site on lncRNA XIST by METTL3. This m⁶A methylation process can be blocked by the inhibition of RBM15 (Patil et al., 2016). Recent studies have reported that RBM15 assists ZC3H13 in regulating m⁶A methylation, which is essential for speeding up the progress of glioblastoma multiforme (GBM) (Chow et al., 2017; Knuckles et al., 2018). It has been confirmed that RBM15 can regulate the differentiation of megakaryocytes by modulating the alternative splicing of RNA (Jin et al., 2018). Our results have showed that the overexpression of RBM15 may be related to the prognosis of LUAD by affecting m⁶A.

KIAA1429 can upregulate c-Jun mRNA via m⁶A by increasing its stability and by promoting the proliferation of gastric cancer cells (Miao et al., 2020). KIAA1429 increases the expression of cyclin-dependent kinase 1 (CDK1) mRNA to increase the invasion ability of breast cancer cells (Qian et al., 2019). KIAA1429 has been shown to regulate the m⁶A modification of GATA3 precursor mRNA (Lan et al., 2019) and ID2 mRNA (Cheng et al., 2019) in HCC, thereby promoting the progression and metastasis of HCC. In the present study, we have revealed that the expression level of m⁶A Writer KIAA1429 may act as a prognostic marker for LUAD patients.

REFERENCES

- Balacco, D. L., and Soller, M. (2019). The m(6)a writer: rise of a machine for growing tasks. *Biochemistry* 58, 363–378. doi: 10.1021/acs.biochem.8b01166
- Chen, M., Wei, L., Law, C. T., Tsang, F. H., Shen, J., Cheng, C. L., et al. (2018). RNA N6-methyladenosine methyltransferase-like 3 promotes liver cancer progression through YTHDF2-dependent posttranscriptional silencing of SOCS2. *Hepatology* 67, 2254–2270. doi: 10.1002/hep.29683
- Chen, M., and Wong, C. M. (2020). The emerging roles of N6-methyladenosine (m6A) deregulation in liver carcinogenesis. *Mol. Cancer* 19:44. doi: 10.1186/s12943-020-01172-y
- Chen, X. Y., Zhang, J., and Zhu, J. S. (2019). The role of m(6)A RNA methylation in human cancer. *Mol. Cancer* 18:103. doi: 10.1186/s12943-019-1033-z
- Cheng, X., Li, M., Rao, X., Zhang, W., Li, X., Wang, L., et al. (2019). KIAA1429 regulates the migration and invasion of hepatocellular carcinoma by altering m6A modification of ID2 mRNA. *Onco. Targets Ther.* 12, 3421–3428. doi: 10.2147/OTT.S180954
- Choe, J., Lin, S., Zhang, W., Liu, Q., Wang, L., Ramirez-Moya, J., et al. (2018). mRNA circularization by METTL3-eIF3h enhances translation and promotes oncogenesis. *Nature* 561, 556–560. doi: 10.1038/s41586-018-0538-8
- Chow, R. D., Guzman, C. D., Wang, G., Schmidt, F., Youngblood, M. W., Ye, L., et al. (2017). AAV-mediated direct *in vivo* CRISPR screen identifies functional suppressors in glioblastoma. *Nat. Neurosci.* 20, 1329–1341. doi: 10.1038/nn.4620
- Cui, Q., Shi, H., Ye, P., Li, L., Qu, Q., Sun, G., et al. (2017). m(6)A RNA methylation regulates the self-renewal and tumorigenesis of glioblastoma stem cells. *Cell Rep.* 18, 2622–2634. doi: 10.1016/j.celrep.2017.02.059
- Dai, D., Wang, H., Zhu, L., Jin, H., and Wang, X. (2018). N6-methyladenosine links RNA metabolism to cancer progression. *Cell Death Dis.* 9:124. doi: 10.1038/s41419-017-0129-x
- Dominissini, D., Moshitch-Moshkovitz, S., Schwartz, S., Salmon-Divon, M., Ungar, L., Osenberg, S., et al. (2012). Topology of the human and mouse m6A RNA methylomes revealed by m6A-seq. *Nature* 485, 201–206. doi: 10.1038/nature11112
- Du, M., Zhang, Y., Mao, Y., Mou, J., Zhao, J., Xue, Q., et al. (2017). MiR-33a suppresses proliferation of NSCLC cells via targeting METTL3 mRNA. *Biochem. Biophys. Res. Commun.* 482, 582–589. doi: 10.1016/j.bbrc.2016.11.077
- Hirsch, F. R., Scagliotti, G. V., Mulshine, J. L., Kwon, R., Curran, W. J. Jr., Wu, Y. L., et al. (2017). Lung cancer: current therapies and new targeted treatments. *Lancet* 389, 299–311. doi: 10.1016/S0140-6736(16)30958-8
- Jia, G., Fu, Y., and He, C. (2013). Reversible RNA adenosine methylation in biological regulation. *Trends Genet.* 29, 108–115. doi: 10.1016/j.tig.2012.11.003
- Jia, G., Fu, Y., Zhao, X., Dai, Q., Zheng, G., Yang, Y., et al. (2011). N6-methyladenosine in nuclear RNA is a major substrate of the obesity-associated FTO. *Nat. Chem. Biol.* 7, 885–887. doi: 10.1038/nchembio.687
- Jin, D., Guo, J., Wu, Y., Yang, L., Wang, X., Du, J., et al. (2020). m(6)A demethylase ALKBH5 inhibits tumor growth and metastasis by reducing YTHDFs-mediated YAP expression and inhibiting miR-107/LATS2-mediated YAP activity in NSCLC. *Mol. Cancer* 19:40. doi: 10.1186/s12943-020-01161-1

In conclusion, our findings provide bioinformatics evidence to trigger and support further research on the important role of m⁶A in LUAD. Toward this direction, the validation of the molecular mechanism of m⁶A underlying LUAD occurrence and its association with LUAD prognosis can be further explored by mechanistic experiments with animal models and/or cancer cell lines.

DATA AVAILABILITY STATEMENT

Requests to access the datasets should be directed to Xiangxuan Zhao, xiangxuanzhao@163.com.

AUTHOR CONTRIBUTIONS

XZ and ZL: study design. HW and XZ: data collection. HW, XZ, and ZL: data analysis. XZ and HW: manuscript preparation. All authors read and approved the final manuscript.

FUNDING

This work was partially supported by National Natural Science Foundation of China [81771947] to ZL and [31371425] to XZ; Liaoning Provincial Natural Science Foundation of China [20180551061] to XZ.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2021.622233/full#supplementary-material>

- Jin, S., Mi, Y., Song, J., Zhang, P., and Liu, Y. (2018). PRMT1-RBM15 axis regulates megakaryocytic differentiation of human umbilical cord blood CD34+ cells. *Exp. Therapeut. Med.* 15:2563–2568. doi: 10.3892/etm.2018.5693
- Kane, S. E., and Beemon, K. (1985). Precise localization of m⁶A in Rous sarcoma virus RNA reveals clustering of methylation sites: implications for RNA processing. *Mol. Cell Biol.* 5, 2298–2306. doi: 10.1128/MCB.5.9.2298
- Knuckles, P., Lence, T., Haussmann, I. U., Jacob, D., Kreim, N., Carl, S. H., et al. (2018). Zc3h13/Flacc is required for adenosine methylation by bridging the mRNA-binding factor Rbm15/Spenito to the m⁶A machinery component Wtap/Fl(2)d. *Genes Dev.* 32, 415–429. doi: 10.1101/gad.309146.117
- Lan, T., Li, H., Zhang, D., Xu, L., Liu, H., Hao, X., et al. (2019). KIAA1429 contributes to liver cancer progression through N⁶-methyladenosine-dependent post-transcriptional modification of GATA3. *Mol. Cancer* 18:186. doi: 10.1186/s12943-019-1106-z
- Li, Z., Weng, H., Su, R., Weng, X., Zuo, Z., Li, C., et al. (2017). FTO plays an oncogenic role in acute myeloid leukemia as a N(6)-methyladenosine RNA demethylase. *Cancer Cell* 31, 127–141. doi: 10.1016/j.ccell.2016.11.017
- Lin, S., Choe, J., Du, P., Triboulet, R., and Gregory, R. I. (2016). The m(6)A methyltransferase METTL3 promotes translation in human cancer cells. *Mol. Cell* 62, 335–345. doi: 10.1016/j.molcel.2016.03.021
- Liu, N., Zhou, K. I., Parisien, M., Dai, Q., Diatchenko, L., and Pan, T. (2017). N⁶-methyladenosine alters RNA structure to regulate binding of a low-complexity protein. *Nucleic Acids Res.* 45, 6051–6063. doi: 10.1093/nar/gkx141
- Liu, Q., and Gregory, R. I. (2019). RNAmoD: an integrated system for the annotation of mRNA modifications. *Nucleic Acids Res.* 47, W548–W555. doi: 10.1093/nar/gkz479
- Liu, Y., Guo, X., Zhao, M., Ao, H., Leng, X., Liu, M., et al. (2020). Contributions and prognostic values of m(6) A RNA methylation regulators in non-small-cell lung cancer. *J. Cell Physiol.* 235, 6043–6057. doi: 10.1002/jcp.29531
- Liu, Z.-X., Li, L.-M., Sun, H.-L., and Liu, S.-M. (2018). Link Between m⁶A modification and cancers. *Front. Bioeng. Biotechnol.* 6:89. doi: 10.3389/fbioe.2018.00089
- Ma, J. Z., Yang, F., Zhou, C. C., Liu, F., Yuan, J. H., Wang, F., et al. (2017). METTL14 suppresses the metastatic potential of hepatocellular carcinoma by modulating N(6)-methyladenosine-dependent primary MicroRNA processing. *Hepatology* 65, 529–543. doi: 10.1002/hep.28885
- Ma, S., Chen, C., Ji, X., Liu, J., Zhou, Q., Wang, G., et al. (2019). The interplay between m⁶A RNA methylation and noncoding RNA in cancer. *J. Hematol. Oncol.* 12:121. doi: 10.1186/s13045-019-0805-7
- Meng, Q., Wang, S., Zhou, S., Liu, H., Ma, X., Zhou, X., et al. (2020). Dissecting the m(6)A methylation affection on afatinib resistance in non-small cell lung cancer. *Pharmacogenomics J.* 20, 227–234. doi: 10.1038/s41397-019-0110-4
- Meyer, K. D., Saletore, Y., Zumbo, P., Elemento, O., Mason, C. E., and Jaffrey, S. R. (2012). Comprehensive analysis of mRNA methylation reveals enrichment in 3' UTRs and near stop codons. *Cell* 149, 1635–1646. doi: 10.1016/j.cell.2012.05.003
- Miao, R., Dai, C. C., Mei, L., Xu, J., Sun, S. W., Xing, Y. L., et al. (2020). KIAA1429 regulates cell proliferation by targeting c-Jun messenger RNA directly in gastric cancer. *J. Cell Physiol.* 235:7420–32. doi: 10.1002/jcp.29645
- Narayan, P., Ludwiczak, R. L., Goodwin, E. C., and Rottman, F. M. (1994). Context effects on N⁶-adenosine methylation sites in prolactin mRNA. *Nucleic Acids Res.* 22, 419–426. doi: 10.1093/nar/22.3.419
- Pan, Y., Ma, P., Liu, Y., Li, W., and Shu, Y. (2018). Multiple functions of m(6)A RNA methylation in cancer. *J. Hematol. Oncol.* 11:48. doi: 10.1186/s13045-018-0590-8
- Patil, D. P., Chen, C. K., Pickering, B. F., Chow, A., Jackson, C., Guttman, M., et al. (2016). m(6)A RNA methylation promotes XIST-mediated transcriptional repression. *Nature* 537, 369–373. doi: 10.1038/nature19342
- Perakis, S. O., Thomas, J. E., and Pichler, M. (2016). Non-coding RNAs enabling prognostic stratification and prediction of therapeutic response in colorectal cancer patients. *Adv. Exp. Med. Biol.* 937, 183–204. doi: 10.1007/978-3-319-42059-2_10
- Qian, J. Y., Gao, J., Sun, X., Cao, M. D., Shi, L., Xia, T. S., et al. (2019). KIAA1429 acts as an oncogenic factor in breast cancer by regulating CDK1 in an N⁶-methyladenosine-independent manner. *Oncogene* 38, 6123–6141. doi: 10.1038/s41388-019-0861-z
- Razzouk, S. (2014). Translational genomics and head and neck cancer: toward precision medicine. *Clin. Genet.* 86, 412–421. doi: 10.1111/cge.12487
- Sheng, H., Li, Z., Su, S., Sun, W., Zhang, X., Li, L., et al. (2019). YTH domain family 2 promotes lung cancer cell growth by facilitating 6-phosphogluconate dehydrogenase mRNA translation. *Carcinogenesis* 41:541–50. doi: 10.1093/carcin/bgz152
- Siegel, R. L., Miller, K. D., and Jemal, A. (2020). Cancer statistics, 2020. *CA Cancer J. Clin.* 70, 7–30. doi: 10.3322/caac.21590
- Sun, L., Liu, W. K., Du, X. W., Liu, X. L., Li, G., Yao, Y., et al. (2020). Large-scale transcriptome analysis identified RNA methylation regulators as novel prognostic signatures for lung adenocarcinoma. *Ann. Transl. Med.* 8:751. doi: 10.21037/atm-20-3744
- Wang, H., Deng, Q., Lv, Z., Ling, Y., Hou, X., Chen, Z., et al. (2019). N⁶-methyladenosine induced miR-143-3p promotes the brain metastasis of lung cancer via regulation of VASH1. *Mol. Cancer* 18:181. doi: 10.1186/s12943-019-1108-x
- Wang, S., Chai, P., Jia, R., and Jia, R. (2018). Novel insights on m(6)A RNA methylation in tumorigenesis: a double-edged sword. *Mol. Cancer* 17:101. doi: 10.1186/s12943-018-0847-4
- Wang, T., Kong, S., Tao, M., and Ju, S. (2020). The potential role of RNA N⁶-methyladenosine in Cancer progression. *Mol. Cancer* 19:88. doi: 10.1186/s12943-020-01204-7
- Wanna-Udom, S., Terashima, M., Lyu, H., Ishimura, A., Takino, T., Sakari, M., et al. (2020). The m⁶A methyltransferase METTL3 contributes to transforming growth factor-beta-induced epithelial-mesenchymal transition of lung cancer cells through the regulation of JUNB. *Biochem. Biophys. Res. Commun.* 524, 150–155. doi: 10.1016/j.bbrc.2020.01.042
- Wei, W., Huo, B., and Shi, X. (2019). miR-600 inhibits lung cancer via downregulating the expression of METTL3. *Cancer Manag. Res.* 11, 1177–1187. doi: 10.2147/CMAR.S181058
- Yang, Z., Li, J., Feng, G., Gao, S., Wang, Y., Zhang, S., et al. (2017). MicroRNA-145 modulates N(6)-methyladenosine levels by targeting the 3'-untranslated mRNA region of the N(6)-methyladenosine binding YTH domain family 2 protein. *J. Biol. Chem.* 292, 3614–3623. doi: 10.1074/jbc.M116.749689
- Yi, D., Wang, R., Shi, X., Xu, L., Yilihamu, Y., and Sang, J. (2020). METTL14 promotes the migration and invasion of breast cancer cells by modulating N⁶-methyladenosine and hsa-miR146a5p expression. *Oncol. Rep.* 43, 1375–1386. doi: 10.3892/or.2020.7515
- Zhang, C., Huang, S., Zhuang, H., Ruan, S., Zhou, Z., Huang, K., et al. (2020). YTHDF2 promotes the liver cancer stem cell phenotype and cancer metastasis by regulating OCT4 expression via m⁶A RNA methylation. *Oncogene* 39, 4507–4518. doi: 10.1038/s41388-020-1303-7
- Zhang, S., Zhao, B. S., Zhou, A., Lin, K., Zheng, S., Lu, Z., et al. (2017). m(6)A demethylase ALKBH5 maintains tumorigenicity of glioblastoma stem-like cells by sustaining FOXM1 expression and cell proliferation program. *Cancer Cell* 31:e596. doi: 10.1016/j.ccell.2017.02.013
- Zheng, W., Dong, X., Zhao, Y., Wang, S., Jiang, H., Zhang, M., et al. (2019). Multiple functions and mechanisms underlying the role of METTL3 in human cancers. *Front. Oncol.* 9:1403. doi: 10.3389/fonc.2019.01403
- Zhou, J., Wang, J., Hong, B., Ma, K., Xie, H., Li, L., et al. (2019). Gene signatures and prognostic values of m⁶A regulators in clear cell renal cell carcinoma - a retrospective study using TCGA database. *Aging* 11, 1633–1647. doi: 10.18632/aging.101856
- Zhuang, Z., Chen, L., Mao, Y., Zheng, Q., Li, H., Huang, Y., et al. (2020). Diagnostic, progressive and prognostic performance of m(6)A methylation RNA regulators in lung adenocarcinoma. *Int. J. Biol. Sci.* 16, 1785–1797. doi: 10.7150/ijbs.39046

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 Wang, Zhao and Lu. 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 m⁶A-Associated RNA Binding Proteins Using an Integrative Computational Framework

Yiqian Zhang^{1*} and Michiaki Hamada^{1,2,3,4*}

¹ Department of Electrical Engineering and Bioscience, Faculty of Science and Engineering, Waseda University, Tokyo, Japan, ² AIST-Waseda University Computational Bio Big-Data Open Innovation Laboratory (CBBD-OIL), Tokyo, Japan, ³ Institute for Medical-Oriented Structural Biology, Waseda University, Tokyo, Japan, ⁴ Graduate School of Medicine, Nippon Medical School, Tokyo, Japan

OPEN ACCESS

Edited by:

Jia Meng,
Xi'an Jiaotong-Liverpool University,
China

Reviewed by:

Lin Zhang,
China University of Mining and
Technology, China
Kunqi Chen,
Fujian Medical University, China

*Correspondence:

Yiqian Zhang
z10000507@126.com
Michiaki Hamada
mhamada@waseda.jp

Specialty section:

This article was submitted to
RNA,
a section of the journal
Frontiers in Genetics

Received: 04 November 2020

Accepted: 05 February 2021

Published: 01 March 2021

Citation:

Zhang Y and Hamada M (2021)
Identification of m⁶A-Associated RNA
Binding Proteins Using an Integrative
Computational Framework.
Front. Genet. 12:625797.
doi: 10.3389/fgene.2021.625797

N6-methyladenosine (m⁶A) is an abundant modification on mRNA that plays an important role in regulating essential RNA activities. Several wet lab studies have identified some RNA binding proteins (RBPs) that are related to m⁶A's regulation. The objective of this study was to identify potential m⁶A-associated RBPs using an integrative computational framework. The framework was composed of an enrichment analysis and a classification model. Utilizing RBPs' binding data, we analyzed reproducible m⁶A regions from independent studies using this framework. The enrichment analysis identified known m⁶A-associated RBPs including YTH domain-containing proteins; it also identified RBM3 as a potential m⁶A-associated RBP for mouse. Furthermore, a significant correlation for the identified m⁶A-associated RBPs is observed at the protein expression level rather than the gene expression level. On the other hand, a Random Forest classification model was built for the reproducible m⁶A regions using RBPs' binding data. The RBP-based predictor demonstrated not only competitive performance when compared with sequence-based predictions but also reflected m⁶A's action of repelling against RBPs, which suggested that our framework can infer interaction between m⁶A and m⁶A-associated RBPs beyond sequence level when utilizing RBPs' binding data. In conclusion, we designed an integrative computational framework for the identification of known and potential m⁶A-associated RBPs. We hope the analysis will provide more insights on the studies of m⁶A and RNA modifications.

Keywords: N6-methyladenosine, RNA binding proteins, RNA modification, enrichment analysis, random forest

1. INTRODUCTION

In recent years, RNA modification has emerged as a mode of post-transcriptional gene regulation and has been gaining increasing attention from researchers around the globe. More than 150 types of post-transcriptional modification have been discovered, with N6-methyladenosine (m⁶A) as being one of the most abundant mRNA modification (Roundtree et al., 2017). m⁶A is featured with the DRACH motif (where D = A, G or U; R = A or G; H = A, C or U) and is preferentially located near 3' untranslated regions (3' UTR) (Linder et al., 2015). It has been reported that m⁶A

participates in essential RNA activities including alternative splicing, export, translation, and decay in the nucleus and cytoplasm (Lee et al., 2020).

m⁶A exerts its function through interaction with several RNA binding proteins that can be considered as m⁶A-associated RBPs. There are three main kinds of known m⁶A-associated RBPs that are also known as m⁶A effectors (Shi et al., 2019), they are writer, eraser, and reader. m⁶A writers are methyltransferases like METTL3, METTL14, WTAP, RBM15/15B, while m⁶A erasers are demethylases like FTO, ALKBH5, and m⁶A readers are the proteins that can recognize m⁶A like the YTH domain-containing proteins (YTHDF1/2/3), EIF3 (Lee et al., 2020), FMR1 (Eduvuganti et al., 2017). These m⁶A effectors cooperate with each other to facilitate both temporal and spatial regulation where writers work in the nucleus to introduce the m⁶A modification which is then recognized by various readers in the nucleus and cytoplasm, which can influence activities of their target RNAs.

Furthermore, the roles of m⁶A and m⁶A-associated RBPs in cancer are being a general interest to researchers. The writer METTL3 was early noticed because of its overexpression in acute myeloid leukemia (AML). It was found that m⁶A promotes the translation of oncogenes like c-MYC, BCL2, and PTEN in the human acute myeloid leukemia MOLM-13 cell line (Vu et al., 2017). Because of necessity of METTL3 in the maintain the leukaemic state, it is identified as a potential therapeutic target for AML (Barbieri et al., 2017). Apart from METTL3, a study found that the reader YTHDF2 silenced in HCC cells can provoke inflammation, vascular reconstruction, and metastatic progression (Hou et al., 2019). Besides, m⁶A and m⁶A reader YTHDF1 have been reported to control anti-tumor immunity. YTHDF1 deficient mice had enhanced therapeutic efficacy of PD-1 checkpoint blockade which suggested YTHDF1's potential in anti-cancer immunotherapy (Han et al., 2019). Therefore, the study of m⁶A and m⁶A-associated RBPs enables us to develop a better understanding of gene regulation mechanism and leads to potential therapeutic opportunities.

To unveil m⁶A's regulation mechanism, it is very necessary to study m⁶A-associated RBPs and their target RNAs. High-throughput sequencing technologies like CLIP-seq (Ule et al., 2005) and RIP-Seq (Zhao et al., 2010) make it feasible to study target RNAs of m⁶A effectors at a transcriptome-wide level. Based on the high-throughput sequencing data, a team developed a database for the collection of these target RNAs (Deng et al., 2020), and another team developed a prediction model focused on the targets of m⁶A readers (Zhen et al., 2020). However, computational resources for identification of m⁶A-associated RBPs are still limited. Though there has been a manually-curated database built for the collection of known m⁶A effectors across species (Nie et al., 2020), to identify potential m⁶A-associated RBPs, it needs to develop efficient computational methods. There are some computational methods that have been used to identify m⁶A-associated RBPs. One such method is to build a prediction model based on deep learning and then extract the sequence features (Zhang and Hamada, 2018; Wang and Wang, 2020). However, not all the RBP motifs are available and sequences can not reflect actual binding status, thus limiting their utility in the

identification of m⁶A-associated RBPs. Another group developed an analysis framework to identify cell-specific trans-regulators of m⁶A (An et al., 2020). They identified the association between m⁶A and RBPs but did not take into consideration the interaction such as reading and repelling between them.

In our study we decided to focus on the use of reproducible m⁶A regions for identification of m⁶A-associated RBPs, with consideration of variation among MeRIP-Seq datasets (about 30–60% between studies, even in the same cell type; McIntyre et al., 2020). We aimed to identify m⁶A-associated RBPs from reproducible m⁶A regions using an integrative computational framework. This framework is composed of an enrichment analysis and a classification model. The enrichment analysis allows us to identify RBPs enriched in the m⁶A regions. We were able to identify not only the known m⁶A-associated RBPs like YTH domain-containing proteins, but also a potential m⁶A-associated RBP, RBM3, for mouse. We went on to evaluate the correlation of these m⁶A-associated RBPs with some known m⁶A effectors and compared these to other RBPs. We observed a significant correlation in the protein expression level rather than the gene expression level, which suggested that the m⁶A-associated RBPs participate in potential pathways at the protein-level in gene regulation. On the other hand, we built a Random Forest classification model for the reproducible m⁶A regions using RBPs' binding data in an effort to understand how RBPs contribute to the profiling of m⁶A regions. This RBP-based predictor demonstrated competitive performance when compared with sequence-based methods. Furthermore, the feature importance inferred from this model can be used to reflect m⁶A's action of repelling against RBPs. These results suggested that this framework could enable researchers to infer interaction between m⁶A and m⁶A-associated RBPs beyond sequence level when utilizing RBPs' binding data.

2. MATERIALS AND METHODS

2.1. MeRIP-Seq Data Collection and Processing

To obtain MeRIP-Seq data of which cell lines are also available for RBPs' binding data, we manually searched GEO database (Barrett et al., 2013) and finally collected raw MeRIP-Seq FASTA files from four independent studies using human HEK293T cell line (human embryonic kidney 293 cells) from European Nucleotide Archive with accession numbers SRP090687 (Lichinchi et al., 2016), SRP039397 (Schwartz et al., 2014), SRP007335 (Meyer et al., 2012), and SRP162223. We also collected MeRIP-Seq data from four independent studies using mouse embryonic fibroblasts (MEF) with accession numbers SRP039402 (Schwartz et al., 2014), SRP048596 (Geula et al., 2015), SRP115436 (Zhou et al., 2018), and SRP061617 (Zhou et al., 2015).

We pre-processed MeRIP-Seq data by using FastQC (Andrews et al., 2012) for quality control and Cutadapt (Martin, 2011) for adapter-trimming. Then, we used MoAIMS, a transcriptome-based peak-calling tool, to detect m⁶A regions with steps including mapping, keeping uniquely mapped reads, sorting, and marking duplicates (Zhang and Hamada, 2020). MoAIMS is an

efficient software we developed based on a statistical framework of a mixture negative-binomial distribution. We run MoAIMS with default parameters except that we set `sep_bin_info=F` when analyzing studies with replicates. MoAIMS called enriched regions at 200-bp resolution as default, therefore we obtained

m⁶A-enriched regions with a size of 200 bp for each MeRIP-Seq sample, and then we identified reproducible m⁶A regions using the criteria that regions are called in at least 60% of the replicates in any one study and further in at least three studies.

2.2. The Enrichment Analysis

We retrieved binding site data of RBPs from the POSTAR2 database (Zhu et al., 2019) and identified RBPs enriched in the reproducible m⁶A regions. A permutation test was adopted to assess the significance of RBP's binding in the m⁶A regions. The rest of regions in genes with m⁶A was used as control and then sampled 1,000 times. We kept the ratio of the number of bins in exons to the number of bins spanning exons the same for both m⁶A and control regions to avoid the regions' position being a confounding factor. For each RBP, we calculated the enrichment ratio using the Equation (1) where N_t is the number of m⁶A regions with the RBP and $E(N_c)$ is the average number of control regions with the RBP from 1,000 times of sampling. Then, a p -value was calculated as the proportion of N_c which were equal to or greater than N_t . After that, multiple testing was performed using Benjamini and Hochberg (1995).

$$R = \frac{N_t}{E(N_c)} \quad (1)$$

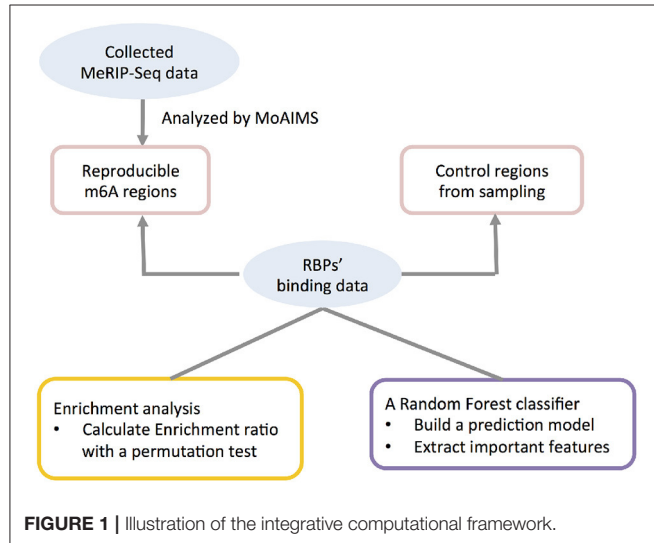


TABLE 1 | RNA binding proteins (RBPs) enriched in reproducible m⁶A regions.

HEK293T	Enrichment ratios*	# m ⁶ A regions with RBPs	p-value**	FDR adjusted p-value
YTHDF2	3.90	6,964	<0.001	<0.003
RBM15	2.73	3,534	<0.001	<0.003
YTHDF3	2.70	52	<0.001	<0.003
YTHDF1	2.49	9,196	<0.001	<0.003
RBM15B	2.32	6,375	<0.001	<0.003
YTHDC1	2.15	7,224	<0.001	<0.003
EIF3D	1.88	593	<0.001	<0.003
NOP58	1.74	159	<0.001	<0.003
HNRNPH1	1.57	47	0.002	0.006
NUDT21	1.48	5,201	<0.001	<0.003
FMR1	1.46	4,443	<0.001	<0.003
DDX3X	1.44	9,470	<0.001	<0.003
EIF3A	1.39	293	<0.001	<0.003
CPSF6	1.34	3,593	<0.001	<0.003
CPSF7	1.31	4,413	<0.001	<0.003
MEF	Enrichment ratio*	# m ⁶ A regions with RBPs	p-value**	FDR adjusted p-value
RBM3	5.81	485	<0.001	<0.001
CREBBP	2.47	24	<0.001	<0.001
SRSF2	2.24	793	<0.001	<0.001
SRSF1	2.13	467	<0.001	<0.001
CPSF6	2.07	94	<0.001	<0.001
CIRBP	1.76	401	<0.001	<0.001

*RBPs are ranked by their enrichment ratios. **P-values were calculated from 1,000 times of permutation. When p-value is zero, it is shown in the table as < 0.001 because it is possible that the p-value is actually <0.001 if times of permutation were increased.

2.3. The Classification Model

We built a Random Forest (RF) classifier to evaluate how much RBPs contribute in discriminating reproducible m⁶A regions. We used the human m⁶A regions with RBPs' binding as the positive data (13,978 in total) and generated 10 sets of control data from the control regions which were set to be an equal data size. We kept the ratio of the number of bins in exons to the number of bins spanning exons the same in both m⁶A and control regions. The binding information (1 for binding, 0 for non-binding) of RBPs was used as the input features. The data was divided into training and test groups at a ratio of 80:20. We implemented the RF classifier using the R package caret (Kuhn, 2008) and randomForest (Liaw and Wiener, 2002) with 5-fold cross validations and “mtry” (the tuning parameters) as 8 (nearly the square root of the number of features). We used the accuracy to measure the performance of the models as shown in the Equation (2) where TP is true positive, TN is true negative, FP is false positive and FN is false negative.

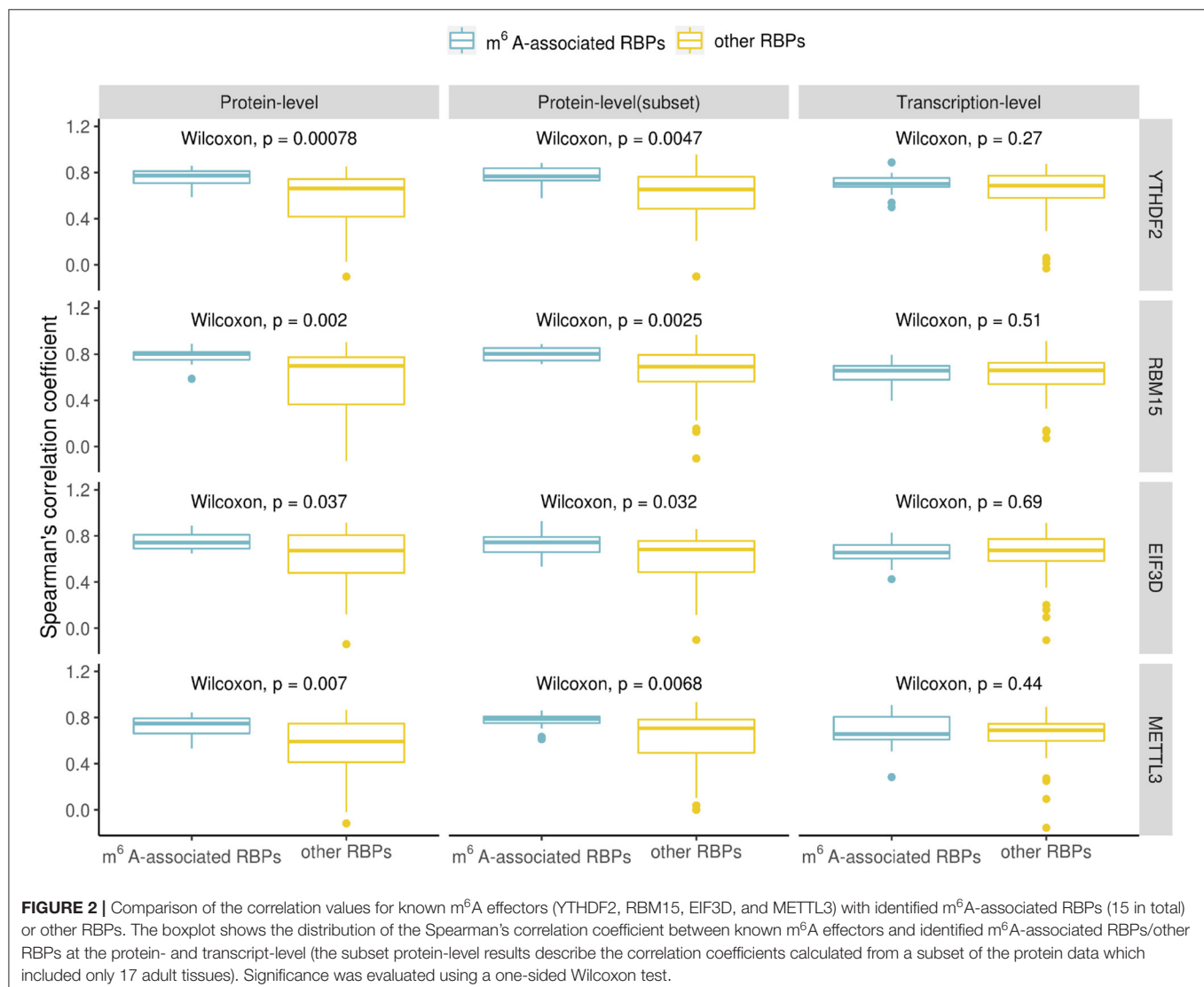
$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (2)$$

A analysis framework including the procedures above was summarized in **Figure 1**.

3. RESULTS

3.1. Identification of m⁶A-Associated RBPs Enriched in Reproducible m⁶A Regions

Because of the considerable variation in the m⁶A datasets (McIntyre et al., 2020), we generated *reproducible* m⁶A regions by collecting MeRIP-Seq data from nine samples of human HEK293T cell line of four independent studies and six samples of mouse MEF cell line of four independent studies. The details of detection of these m⁶A regions are provided in section 2. With a relatively strict criteria, we finally obtained 14,803 reproducible m⁶A regions for HEK293T cell line and 5,576 reproducible m⁶A regions for MEF cell line.



To identify RBPs enriched in m⁶A regions, 71 RBPs for HEK293T/HEK293 and nine RBPs for MEF were retrieved from the POSTAR2 database. For each RBP, we calculated an enrichment score and assessed its significance using a permutation test as described in section 2. When setting the threshold for the enrichment ratio to ≥ 1.3 and FDR (false discovery rate) adjusted p -value to ≤ 0.05 , we obtained enriched RBPs listed in **Table 1**. For HEK293T, we identified several known m⁶A readers including YTH family proteins, FMR1, EIF3, and m⁶A writers RBM15/15B, a component of the WTAP-METTL3 complex (Patil et al., 2016; Lee et al., 2020). For MEF, we found a common RBP, CPSF6, which is enriched for both human and mouse. CPSF6 is a polyadenylation cleavage factor and has been reported to be associated with VIRMA, which mediates preferential m⁶A methylation in the 3' UTR and near stop codon and participates alternative polyadenylation (APA) in human (Yue et al., 2018). Another study found YTHDC1's association with CPSF6 during mouse oocyte development (Kasowitz et al., 2018). In addition, we noticed that RBM3 was highly enriched in m⁶A regions of MEF. RBM3 is an important regulator of circadian gene expression by controlling APA (Liu et al., 2013), therefore we suggest that RBM3 could be associated with m⁶A in the APA regulation process. The full list of enrichment ratios for each of the RBPs is provided in **Supplementary Tables 1, 2**. Besides, for each enriched RBP

(overlap with more than 100 m⁶A regions), we also listed the RBPs that more than 60% of the enriched RBP is overlapped with for HEK293T in **Supplementary Table 3**. As expected, YTHDF1 and DDX3X were shown to have the highest overlapping percentage as they have a considerable overlap with m⁶A regions.

The RBPs in **Table 1** are considered as m⁶A-associated RBPs, therefore we wondered how they are correlated with known m⁶A effectors when compared with other RBPs at both the transcription and the protein expression level. We performed a correlation analysis for all the human RBPs. To do the correlation analysis at the transcription level, we downloaded Illumina Body Map (HBM) (Asmann et al., 2012; Barbosa-Morais et al., 2012; Derrien et al., 2012) from ArrayExpress (Athar et al., 2019) with the accession number E-MTAB-513, which provides gene expression data for 16 human tissues. For the correlation analysis at the protein level, we downloaded mass spectrometry data from Human Proteome Map (HPM) (Kim et al., 2014) for 30 human tissues/cell lines. We checked some known m⁶A effectors including YTHDF2, RBM15, EIF3D which ranked at the top of **Table 1** and METTL3 of which binding data is not available but is a well-known m⁶A writer, and compared their correlation with the identified m⁶A-associated RBPs (15 in total) or with the rest of RBPs (56 in total). Correlation was calculated using the Spearman's correlation coefficient. We observed a similar trend in all the investigated known m⁶A effectors which showed that

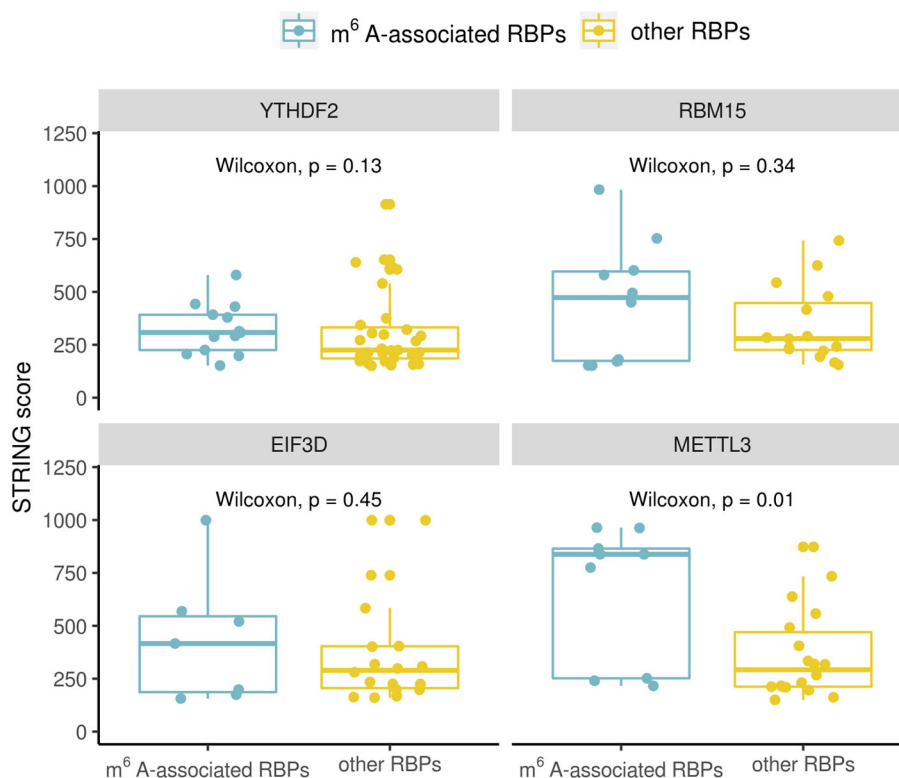
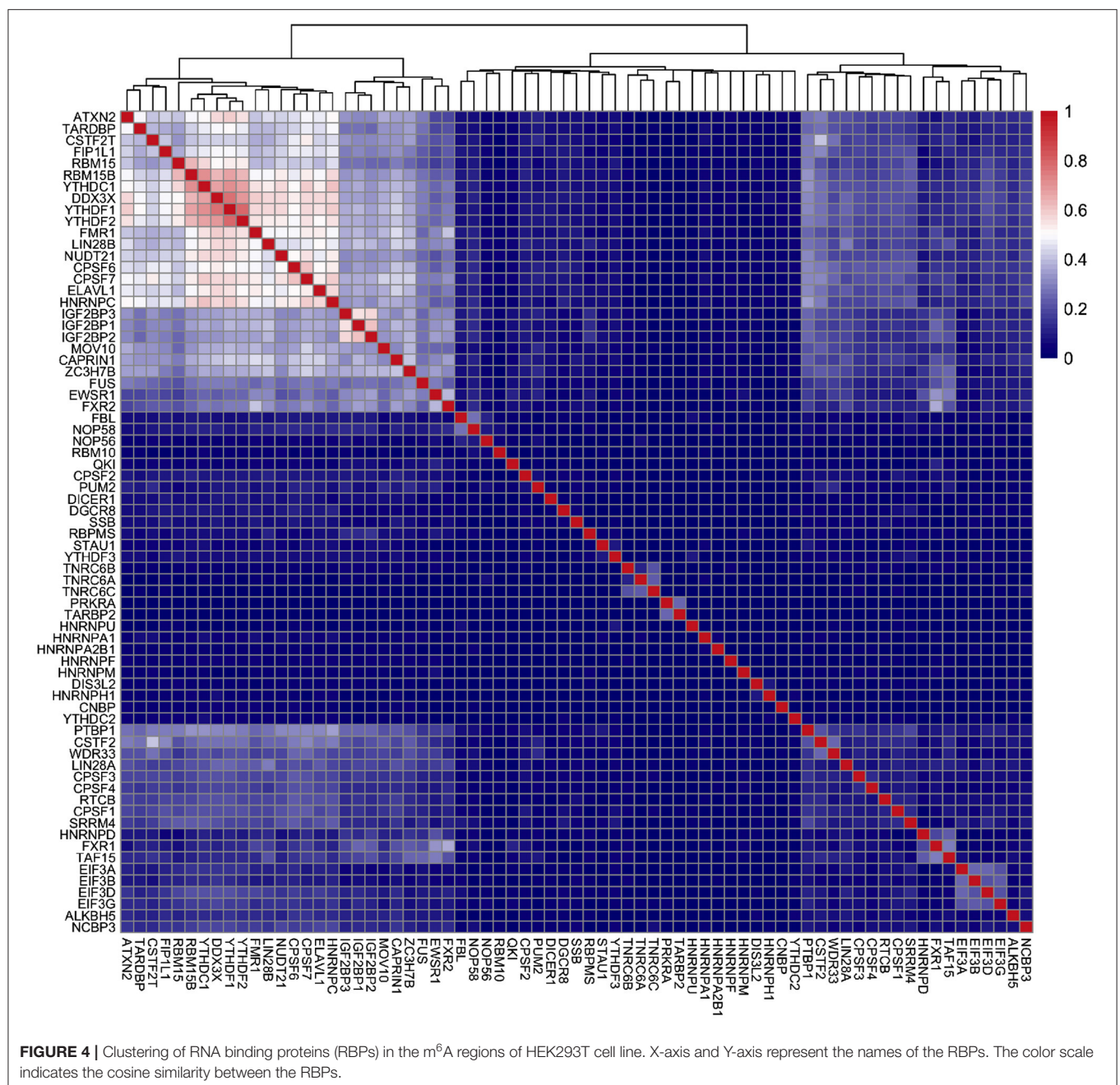


FIGURE 3 | Comparison of protein-protein interactions between known m⁶A effectors (YTHDF2, RBM15, EIF3D, and METTL3) and identified m⁶A-associated RBPs (15 in total)/other RBPs. The boxplot shows the distribution of the interaction scores between known m⁶A effectors and identified m⁶A-associated RBPs/other RBPs. Significance was evaluated using a one-sided Wilcoxon test.

the identified m⁶A-associated RBPs are more correlated with them at the protein-level than the transcription level (**Figure 2**). Because the protein data included more tissues/cell lines than the transcription data, we chose to compare a subset of 17 adult tissues to check the correlation values for avoiding any biased introduced by different dataset sizes. The higher correlation at the protein level was still observed in this subset evaluation as shown in **Figure 2**. Some studies have reported cases of gene regulation with dependency between m⁶A-associated RBPs such as METTL3 and YTHDF2 (Chen et al., 2018; Kasowitz et al., 2018). This observation supports the hypothesis that

m⁶A-associated RBPs are more likely to participate in potential pathways at the protein level. Then, we went on to confirm to if these higher correlation values are the result of protein-protein interactions. To do this we retrieved the protein-protein interaction data from STRING (von Mering et al., 2005). The available interaction scores do not show significant difference between m⁶A-associated RBPs and other RBPs except for METTL3 (**Figure 3**). Because the protein-protein interaction data is still limited, from the available data it is suggested that the higher correlation at the protein-level is marginally related to protein-protein interaction. m⁶A modification is a dynamic



process involving both temporal and spatial regulation between m⁶A effectors, therefore it is expected to have further studies to unveil the regulation mechanism of these proteins.

3.2. Identification of m⁶A-Associated RBPs Contributing to the Classification of m⁶A Regions

After we identified RBPs enriched in the reproducible m⁶A regions, we wanted to develop a more comprehensive

understanding of how RBPs' binding contributes to the profile of m⁶A regions. To do this, we performed a further analysis on the human RBPs. First, we investigated the overall profile of the binding information of RBPs (0 for non-binding and 1 for binding) in the reproducible m⁶A regions. We calculated the pairwise distance between RBPs using cosine similarity and performed clustering (**Figure 4**). The result of the clustering analysis demonstrated the co-occurrence of YTH family proteins and RBM15B which all ranked in the top of the enrichment analysis. Then, we built a Random Forest classifier which incorporated the binding information for each of the RBPs as features. The details of models are described in section 2. The classifier achieved an average accuracy of 0.736 and AUROC (Area Under Receiver Operating Characteristic) of 0.788 as shown in **Figure 5**. We also compared the RBP-based classifier with two sequence-based predictors SRAMP (Zhou et al., 2016) in mature mRNA mode and DeepM6ASeq which showed an accuracy of 0.660 and 0.686, respectively and AUROC of 0.754 for both (**Figure 5**). We plotted top 10 most important features as shown in **Figure 6** and among them found the enriched m⁶A-associated RBPs such as the readers YTHDF1/2, YTHDC1, the writers RBM15/15B. Besides, it is noticed that ELAVL1 also has contribution to the classification of m⁶A regions to some extent. ELAVL1 is reported to have action of being repelled by m⁶A in general that can lead to RNA decay (Wang et al., 2014; Lee et al., 2020). The repelling action of m⁶A against ELAVL1 is consistent with the enrichment results, which show that its enrichment ratio is 0.816. In summary, the RBP-based classifier not only demonstrated competitive performance in the prediction of reproducible m⁶A regions but also helped to infer interaction between m⁶A and m⁶A-associated RBPs beyond sequence level when combined with the results of the enrichment analysis.

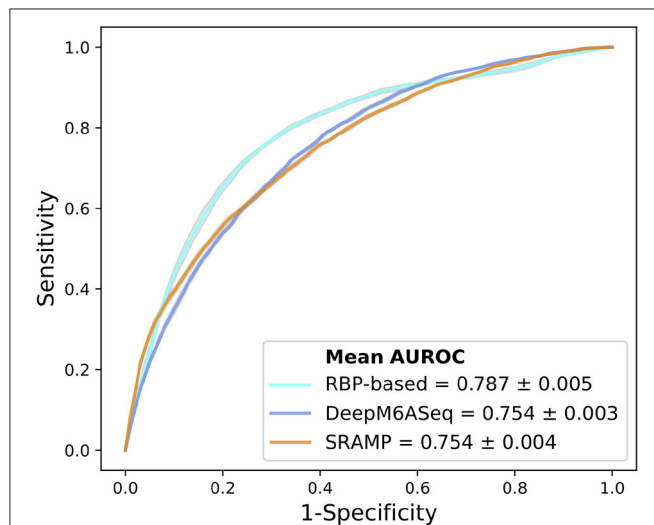


FIGURE 5 | Comparison of AUROC between the RBPs (RNA binding proteins)-based predictor, DeepM6ASeq, and SRAMP in mature mRNA mode for the classification of HEK293T m⁶A regions. The plot represents average ROC from ten times of sampling control regions for each predictor.

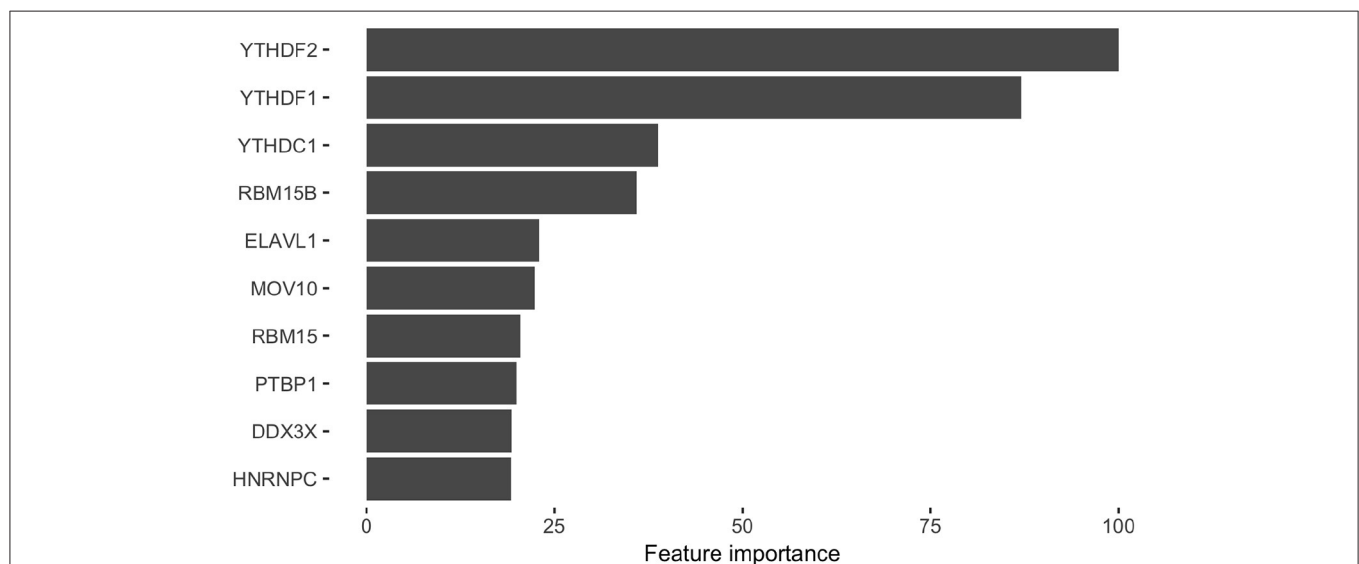


FIGURE 6 | Top 10 RNA binding proteins (RBPs) identified from the classification of the HEK293T reproducible m⁶A regions. The bar graph shows the top 10 RBPs extracted from the classifier for the m⁶A regions. X-axis represents the name of RBPs and Y-axis represents the average importance score from ten times of sampling control regions.

4. DISCUSSION

Utilizing the binding information of RBPs, this computational framework enabled us to identify potential m⁶A-associated RBPs and infer their interaction with m⁶A. This analysis serves as a first step, and future analyses may include some improvements and expansions. First, this framework was designed and tested on a limited number of cell types and organisms. With the increasing amount of data available for m⁶A and RBPs in more cell lines and tissues, this framework could be tested on much larger datasets and may provide valuable insights into the m⁶A regulatory network. Especially, this framework is promising in the application of cancer research. Several studies have identified function of m⁶A effectors like METTL3/14, YTHDF1/2, and IGF2BP1 in multiple cancer types (Cui et al., 2017; Li et al., 2018; Chen et al., 2019; Han et al., 2019; Müller et al., 2019). This framework is expected to provide clues for potential m⁶A effectors and the interaction among them in cancer research. In addition, this framework could be applied to other RNA modifications such as N1-methyladenosine (m1A) (Dominissini et al., 2016), and 5-methylcytosine (m5C) (Amort et al., 2017), which have also been identified as critical RNA modification. Such analyses could help improve experimental design in wet lab applications and help researchers narrow their focus. Third, apart from RBPs, other genomic features like transcription factors and histone modification are worth inspecting for studying the m⁶A regulation networks at multiple layers. These applications highlight the future utility of this framework and its value in the current research climate.

5. CONCLUSION

We designed an integrative computational framework for identification of m⁶A-associated RBPs in reproducible m⁶A regions. This computational framework is composed of an enrichment analysis and a classification model. Using the enrichment analysis, we were able to identify known m⁶A-associated RBPs and several potential ones including RBM3 from mouse. These identified m⁶A-associated RBPs show a significant degree of correlation at their protein level, although this is not seen in their transcriptional profile, which suggests that these m⁶A-associated RBPs participate in potential pathways at the protein-level in gene regulation. On the other hand, we built a classification model for m⁶A regions using a Random

Forest algorithm that uses RBPs' binding information as its input features. The RBP-based predictor not only demonstrated comparable performance to sequence-based predictions but also helped infer interaction between m⁶A and m⁶A-associated RBPs like actions of reading and repelling beyond sequence level. We hope that this analysis framework can assist biologists in their study of RNA modifications.

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

YZ conceived this study, implemented the methods, performed the experiments, and wrote the draft. MH supervised this study and revised the manuscript critically. YZ and MH contributed to the construction of methods and analysis/interpretation of the data. All authors read and approved the final manuscript.

FUNDING

Publication costs are funded by Waseda University [basic research budget]. This study was supported by the Ministry of Education, Culture, Sports, Science and Technology (KAKENHI) [grant numbers JP17K20032, JP16H05879, JP16H01318, and JP16H02484 to MH]. The funding bodies did not play any role in the design of the study or collection, analysis and interpretation of data or in writing the manuscript.

ACKNOWLEDGMENTS

Computation was performed by server facilities at AIST-Waseda University Computational Bio Big-Data Open Innovation Laboratory (CBBD-OIL).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2021.625797/full#supplementary-material>

REFERENCES

- Amort, T., Rieder, D., Wille, A., Khokhlova-Cubberley, D., Riml, C., Trixl, L., et al. (2017). Distinct 5-methylcytosine profiles in poly(A) RNA from mouse embryonic stem cells and brain. *Genome Biol.* 18:1. doi: 10.1186/s13059-016-1139-1
- An, S., Huang, W., Huang, X., Cun, Y., Cheng, W., Sun, X., et al. (2020). Integrative network analysis identifies cell-specific trans regulators of m6A. *Nucleic Acids Res.* 48, 1715–1729. doi: 10.1093/nar/gkz1206
- Andrews, S., Krueger, F., Segonds-Pichon, A., Biggins, L., Krueger, C., and Wingett, S. (2012). *FastQC*. Babraham Institute.
- Asmann, Y. W., Necela, B. M., Kalari, K. R., Hossain, A., Baker, T. R., Carr, J. M., et al. (2012). Detection of redundant fusion transcripts as biomarkers or disease-specific therapeutic targets in breast cancer. *Cancer Res.* 72, 1921–1928. doi: 10.1158/0008-5472.CAN-11-3142
- Athar, A., Füllgrabe, A., George, N., Iqbal, H., Huerta, L., Ali, A., et al. (2019). ArrayExpress update - from bulk to single-cell expression data. *Nucleic Acids Res.* 47, D711–D715. doi: 10.1093/nar/gky964
- Barbieri, I., Tzelepis, K., Pandolfini, L., Shi, J., Millán-Zambrano, G., Robson, S. C., et al. (2017). Promoter-bound METTL3 maintains myeloid leukaemia by m6A-dependent translation control. *Nature* 552, 126–131. doi: 10.1038/nature24678

- Barbosa-Morais, N. L., Irimia, M., Pan, Q., Xiong, H. Y., Gueroussov, S., Lee, L. J., et al. (2012). The evolutionary landscape of alternative splicing in vertebrate species. *Science* 338, 1587–1593. doi: 10.1126/science.1230612
- Barrett, T., Wilhite, S. E., Ledoux, P., Evangelista, C., Kim, I. F., Tomashevsky, M., et al. (2013). NCBI GEO: archive for functional genomics data sets-update. *Nucleic Acids Res.* 41, D991–D995. doi: 10.1093/nar/gks1193
- Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B* 57, 289–300. doi: 10.1111/j.2517-6161.1995.tb02031.x
- Chen, M., Wei, L., Law, C. T., Tsang, F. H., Shen, J., Cheng, C. L., et al. (2018). RNA N6-methyladenosine methyltransferase-like 3 promotes liver cancer progression through YTHDF2-dependent posttranscriptional silencing of SOCS2. *Hepatology* 67, 2254–2270. doi: 10.1002/hep.29683
- Chen, X. Y., Zhang, J., and Zhu, J. S. (2019). The role of m6A RNA methylation in human cancer. *Mol. Cancer* 18:103. doi: 10.1186/s12943-019-1033-z
- Cui, Q., Shi, H., Ye, P., Li, L., Qu, Q., Sun, G., et al. (2017). m6A RNA methylation regulates the self-renewal and tumorigenesis of glioblastoma stem cells. *Cell Rep.* 18, 2622–2634. doi: 10.1016/j.celrep.2017.02.059
- Deng, S., Zhang, H., Zhu, K., Li, X., Ye, Y., Li, R., et al. (2020). M6A2Target: a comprehensive database for targets of m6A writers, erasers and readers. *Brief. Bioinform.* bbaa055. doi: 10.1093/bib/bbaa055
- Derrien, T., Johnson, R., Bussotti, G., Tanzer, A., Djebali, S., Tilgner, H., et al. (2012). The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. *Genome Res.* 22, 1775–1789. doi: 10.1101/gr.132159.111
- Dominissini, D., Nachtgale, S., Moshitch-Moshkovitz, S., Peer, E., Kol, N., Ben-Haim, M. S., et al. (2016). The dynamic N(1)-methyladenosine methylome in eukaryotic messenger RNA. *Nature* 530, 441–446. doi: 10.1038/nature16998
- Edupaganti, R. R., Geiger, S., Lindeboom, R. G. H., Shi, H., Hsu, P. J., Lu, Z., et al. (2017). N6-methyladenosine (m6A) recruits and repels proteins to regulate mRNA homeostasis. *Nat. Struct. Mol. Biol.* 24, 870–878. doi: 10.1038/nsmb.3462
- Geula, S., Moshitch-Moshkovitz, S., Dominissini, D., Mansour, A. A., Kol, N., Salmon-Divon, M., et al. (2015). Stem cells. m6A mRNA methylation facilitates resolution of naïve pluripotency toward differentiation. *Science* 347, 1002–1006. doi: 10.1126/science.1261417
- Han, D., Liu, J., Chen, C., Dong, L., Liu, Y., Chang, R., et al. (2019). Anti-tumour immunity controlled through mRNA m6A methylation and YTHDF1 in dendritic cells. *Nature* 566, 270–274. doi: 10.1038/s41586-019-0916-x
- Hou, J., Zhang, H., Liu, J., Zhao, Z., Wang, J., Lu, Z., et al. (2019). YTHDF2 reduction fuels inflammation and vascular abnormalization in hepatocellular carcinoma. *Mol. Cancer* 18:163. doi: 10.1186/s12943-019-1082-3
- Kasowitz, S. D., Ma, J., Anderson, S. J., Leu, N. A., Xu, Y., Gregory, B. D., et al. (2018). Nuclear m6A reader YTHDC1 regulates alternative polyadenylation and splicing during mouse oocyte development. *PLoS Genet.* 14:e1007412. doi: 10.1371/journal.pgen.1007412
- Kim, M. S., Pinto, S. M., Getnet, D., Nirujogi, R. S., Manda, S. S., Chaekady, R., et al. (2014). A draft map of the human proteome. *Nature* 509, 575–581. doi: 10.1038/nature13302
- Kuhn, M. (2008). Building predictive models in R using the caret package. *J. Stat. Softw.* 28, 1–26. doi: 10.18637/jss.v028.i05
- Lee, Y., Choe, J., Park, O. H., and Kim, Y. K. (2020). Molecular mechanisms driving mRNA degradation by m6A modification. *Trends Genet.* 36, 177–188. doi: 10.1016/j.tig.2019.12.007
- Li, Z., Qian, P., Shao, W., Shi, H., He, X. C., Gogol, M., et al. (2018). Suppression of m6A reader Ythdf2 promotes hematopoietic stem cell expansion. *Cell Res.* 28, 904–917. doi: 10.1038/s41422-018-0072-0
- Liaw, A., and Wiener, M. (2002). Classification and regression by randomforest. *R News* 2, 18–22.
- Lichinchi, G., Zhao, B. S., Wu, Y., Lu, Z., Qin, Y., He, C., and Rana, T. M. (2016). Dynamics of human and viral RNA methylation during Zika virus infection. *Cell Host Microbe* 20, 666–673. doi: 10.1016/j.chom.2016.10.002
- Linder, B., Grozhik, A. V., Olarerin-George, A. O., Meydan, C., Mason, C. E., and Jaffrey, S. R. (2015). Single-nucleotide-resolution mapping of m6A and m6Am throughout the transcriptome. *Nat. Methods* 12, 767–772. doi: 10.1038/nmeth.3453
- Liu, Y., Hu, W., Murakawa, Y., Yin, J., Wang, G., Landthaler, M., et al. (2013). Cold-induced RNA-binding proteins regulate circadian gene expression by controlling alternative polyadenylation. *Sci. Rep.* 3:2054. doi: 10.1038/srep02054
- Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet. J.* 17, 10–12. doi: 10.14806/ej.17.1.200
- McIntyre, A. B. R., Gokhale, N. S., Cerchietti, L., Jaffrey, S. R., Horner, S. M., and Mason, C. E. (2020). Limits in the detection of m6A changes using MeRIP/m6A-seq. *Sci. Rep.* 10:6590. doi: 10.1038/s41598-020-63355-3
- Meyer, K. D., Saletore, Y., Zumbo, P., Elemento, O., Mason, C. E., and Jaffrey, S. R. (2012). Comprehensive analysis of mRNA methylation reveals enrichment in 3' UTRs and near stop codons. *Cell* 149, 1635–1646. doi: 10.1016/j.cell.2012.05.003
- Müller, S., Glaß, M., Singh, A. K., Haase, J., Bley, N., Fuchs, T., et al. (2019). IGF2BP1 promotes SRF-dependent transcription in cancer in a m6A- and miRNA-dependent manner. *Nucleic Acids Res.* 47, 375–390. doi: 10.1093/nar/gky1012
- Nie, F., Feng, P., Song, X., Wu, M., Tang, Q., and Chen, W. (2020). RNAWRE: a resource of writers, readers and erasers of RNA modifications. *Database* 2020:baaa049. doi: 10.1093/database/baaa049
- Patil, D. P., Chen, C. K., Pickering, B. F., Chow, A., Jackson, C., Guttman, M., et al. (2016). m(6)A RNA methylation promotes XIST-mediated transcriptional repression. *Nature* 537, 369–373. doi: 10.1038/nature19342
- Roundtree, I. A., Evans, M. E., Pan, T., and He, C. (2017). Dynamic RNA modifications in gene expression regulation. *Cell* 169, 1187–1200. doi: 10.1016/j.cell.2017.05.045
- Schwartz, S., Mumbach, M. R., Jovanovic, M., Wang, T., Maciag, K., Bushkin, G. G., et al. (2014). Perturbation of m6A writers reveals two distinct classes of mRNA methylation at internal and 5' sites. *Cell Rep.* 8, 284–296. doi: 10.1016/j.celrep.2014.05.048
- Shi, H., Wei, J., and He, C. (2019). Where, when, and how: context-dependent functions of RNA methylation writers, readers, and erasers. *Mol. Cell* 74, 640–650. doi: 10.1016/j.molcel.2019.04.025
- Ule, J., Jensen, K., Mele, A., and Darnell, R. B. (2005). CLIP: a method for identifying protein-RNA interaction sites in living cells. *Methods* 37, 376–386. doi: 10.1016/j.ymeth.2005.07.018
- von Mering, C., Jensen, L. J., Snel, B., Hooper, S. D., Krupp, M., Foglierini, M., et al. (2005). STRING: known and predicted protein-protein associations, integrated and transferred across organisms. *Nucleic Acids Res.* 33, D433–D437. doi: 10.1093/nar/gki005
- Vu, L. P., Pickering, B. F., Cheng, Y., Zaccara, S., Nguyen, D., Minuesa, G., et al. (2017). The N6-methyladenosine (m6A)-forming enzyme METTL3 controls myeloid differentiation of normal hematopoietic and leukemia cells. *Nat. Med.* 23, 1369–1376. doi: 10.1038/nm.4416
- Wang, J., and Wang, L. (2020). Deep analysis of RNA N6-adenosine methylation (m6A) patterns in human cells. *NAR Genomics Bioinform.* 2:lqaa007. doi: 10.1093/nargab/lqaa007
- Wang, Y., Li, Y., Toth, J. I., Petroski, M. D., Zhang, Z., and Zhao, J. C. (2014). N6-methyladenosine modification destabilizes developmental regulators in embryonic stem cells. *Nat. Cell Biol.* 16, 191–198. doi: 10.1038/ncb2902
- Yue, Y., Liu, J., Cui, X., Cao, J., Luo, G., Zhang, Z., et al. (2018). VIRMA mediates preferential m6A mRNA methylation in 3'UTR and near stop codon and associates with alternative polyadenylation. *Cell Discov.* 4:10. doi: 10.1038/s41421-018-0019-0
- Zhang, Y., and Hamada, M. (2018). DeepM6ASeq: prediction and characterization of m6A-containing sequences using deep learning. *BMC Bioinformatics* 19(Suppl. 19):524. doi: 10.1186/s12859-018-2516-4
- Zhang, Y., and Hamada, M. (2020). MoAIMS: efficient software for detection of enriched regions of MeRIP-Seq. *BMC Bioinformatics* 21:103. doi: 10.1186/s12859-020-3430-0
- Zhao, J., Ohsumi, T. K., Kung, J. T., Ogawa, Y., Grau, D. J., Sarma, K., et al. (2010). Genome-wide identification of polycomb-associated RNAs by RIP-seq. *Mol. Cell* 40, 939–953. doi: 10.1016/j.molcel.2010.12.011
- Zhen, D., Wu, Y., Zhang, Y., Chen, K., Song, B., Xu, H., et al. (2020). m6A Reader: epitranscriptome target prediction and functional characterization of N6-methyladenosine (m6A) readers. *Front. Cell. Dev. Biol.* 8:741. doi: 10.3389/fcell.2020.00741
- Zhou, J., Wan, J., Gao, X., Zhang, X., Jaffrey, S. R., and Qian, S. B. (2015). Dynamic m(6)A mRNA methylation directs translational control of heat shock response. *Nature* 526, 591–594. doi: 10.1038/nature15377

- Zhou, J., Wan, J., Shu, X. E., Mao, Y., Liu, X. M., Yuan, X., et al. (2018). N6-methyladenosine guides mRNA alternative translation during integrated stress response. *Mol. Cell* 69, 636–647. doi: 10.1016/j.molcel.2018.01.019
- Zhou, Y., Zeng, P., Li, Y. H., Zhang, Z., and Cui, Q. (2016). SRAMP: prediction of mammalian N6-methyladenosine (m6A) sites based on sequence-derived features. *Nucleic Acids Res.* 44:e91. doi: 10.1093/nar/gkw104
- Zhu, Y., Xu, G., Yang, Y. T., Xu, Z., Chen, X., Shi, B., et al. (2019). POSTAR2: deciphering the post-transcriptional regulatory logics. *Nucleic Acids Res.* 47, D203–D211. doi: 10.1093/nar/gky830

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 and Hamada. 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.



Spliceosomal snRNA Epitranscriptomics

Pedro Morais^{1*}, Hironori Adachi² and Yi-Tao Yu^{2*}

¹ProQR Therapeutics, Leiden, Netherlands, ²Department of Biochemistry and Biophysics, Center for RNA Biology, University of Rochester Medical Center, Rochester, NY, United States

OPEN ACCESS

Edited by:

Jia Meng,
Xi'an Jiaotong-Liverpool University,
China

Reviewed by:

Piyush Khandelia,
Birla Institute of Technology and
Science, India
Olivier Bensaudé,
École Normale Supérieure, France

*Correspondence:

Yi-Tao Yu
yitao_yu@urmc.rochester.edu
Pedro Morais
pmorais@proqr.com

Specialty section:

This article was submitted to
RNA,
a section of the journal
Frontiers in Genetics

Received: 11 January 2021

Accepted: 08 February 2021

Published: 02 March 2021

Citation:

Morais P, Adachi H and Yu Y-T (2021)
Spliceosomal snRNA
Epitranscriptomics.
Front. Genet. 12:652129.
doi: 10.3389/fgene.2021.652129

Small nuclear RNAs (snRNAs) are critical components of the spliceosome that catalyze the splicing of pre-mRNA. snRNAs are each complexed with many proteins to form RNA-protein complexes, termed as small nuclear ribonucleoproteins (snRNPs), in the cell nucleus. snRNPs participate in pre-mRNA splicing by recognizing the critical sequence elements present in the introns, thereby forming active spliceosomes. The recognition is achieved primarily by base-pairing interactions (or nucleotide-nucleotide contact) between snRNAs and pre-mRNA. Notably, snRNAs are extensively modified with different RNA modifications, which confer unique properties to the RNAs. Here, we review the current knowledge of the mechanisms and functions of snRNA modifications and their biological relevance in the splicing process.

Keywords: pre-mRNA splicing, small nuclear RNA, RNA modifications, epitranscriptomics, pseudouridine, 2'-O-methylation, N6-methyladenosine, N2-methylation

INTRODUCTION

Pre-mRNA splicing is, by definition, a co- or post-transcriptional RNA processing reaction by which introns are removed from mRNA precursors, and exons are precisely joined together to form functional mature mRNAs (Berget et al., 1977; Chow et al., 1977; Shi, 2017). The fidelity of this mechanism is critical for correct gene expression as proven by the fact that 10% of all disease-causing single-point mutations in humans generate splicing defects (Krawczak et al., 2007; Scotti and Swanson, 2016). Pre-mRNA splicing occurs via a two-step transesterification reaction pathway (Figure 1; Ruskin et al., 1984). In the first step, the 2'-hydroxyl group (2'-OH) of the branch point nucleotide (adenosine) attacks the phosphate at the 5' exon-intron junction (5' splice site), resulting in the cleavage of the phosphodiester bond between the 5' exon and intron, and the concurrent formation of a new 5'-2' phosphodiester bond between the 5' end of the intron and the branch point adenosine. Thus, a lariat-structured intermediate (lariat intron-3' exon) and a cut-off 5' exon intermediate are produced. In the second step, the 3'-OH group of the cut-off 5' exon attacks the phosphate at the intron-3' exon junction (3' splice site), releasing the lariat intron product and generating the spliced mature mRNA product.

The two chemical reactions of pre-mRNA splicing occur only after the pre-mRNA is assembled into the functional spliceosome, a multi-component complex composed of five small nuclear RNAs (snRNAs U1, U2, U4, U5, and U6), which are present as small nuclear ribonucleoprotein particles (snRNPs, RNA-protein complexes) and a large number of splicing protein factors (Jurica and Moore, 2003).

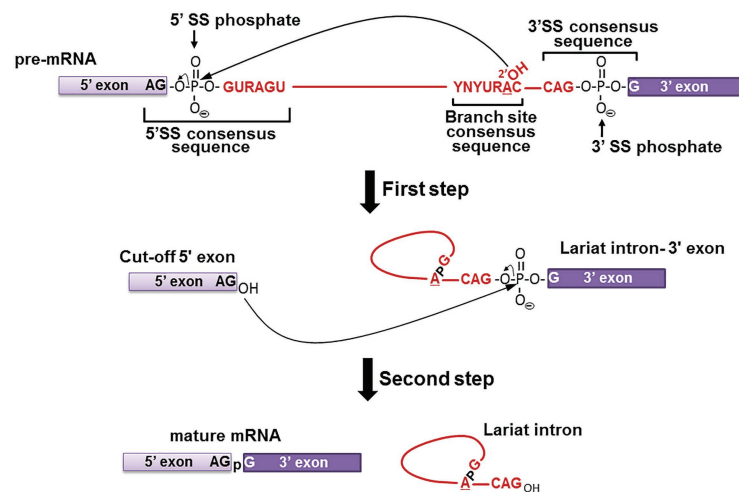


FIGURE 1 | Pre-mRNA splicing pathway. Pre-mRNA splicing takes place via a two-step transesterification reaction pathway. In the first step, the 2'-OH group of the branch site adenosine attacks the phosphate at the 5' exon-intron junction (5' SS phosphate), generating the "Cut-off 5' exon" and the "Lariat intron-3' exon" intermediates. In the second step, the 3'-OH group of the "Cut-off 5' exon" attacks the phosphate at the intron-3' exon junction (3' SS phosphate), yielding the "mature mRNA" and "Lariat intron" products. The consensus sequences at the branch site and the 5' and 3' splice sites are shown (red letters, where Y is pyrimidine and R is purine). The 5' exon (light purple box), 3' exon (dark purple box), and the intron (red line) are also shown.

Mechanistic Role of snRNA in Pre-mRNA Splicing

During spliceosome assembly, spliceosomal snRNPs and splicing factors recognize and interact with the pre-mRNA consensus sequences, facilitating and specifying the transesterification reactions (Figure 2). Specifically, U1 snRNP recognizes the 5' splice site of a pre-mRNA to form a commitment complex (complex E) that commits the pre-mRNA to spliceosome assembly (Kondo et al., 2015). This recognition involves base-pairing interactions between the 10 highly conserved nucleotides at the 5' end of U1 snRNA and the intron sequences of the pre-mRNA at the 5' splice site (G/GUAUGU in yeast or G/GURAGU in vertebrates, where "G" represents the exon-intron junction and R stands for purine; Zhuang and Weiner, 1986). The U2 snRNP then recognizes the pre-mRNA branch site to form a pre-splicing complex called as complex A (Query et al., 1997). This recognition again involves a base-pairing interaction between a highly conserved sequence in U2 snRNA and the pre-mRNA branch site sequence (UACUAAAC in yeast or YNYURAC in vertebrates, where Y, R, N, and the underlined adenosine represent pyrimidine, purine, any nucleotide, and the branch point nucleotide, respectively; Parker et al., 1987; Zhuang et al., 1989). While U1 and U2 snRNPs recognize the 5' splice site and the branch site, respectively, the U2 auxiliary factor (U2AF) recognizes the 3' splice site (YAG/G; in vertebrates, the 3' splice site is preceded by a poly-pyrimidine tract; Ruskin et al., 1988; Wu et al., 1999). After the formation of complex A, the U4/U6.U5 tri-snRNP, in which U4 and U6 snRNAs are extensively base-paired, joins this pre-splicing complex, resulting in the formation of a fully assembled spliceosome (complex B; Boehringer et al., 2004; Nguyen et al., 2016; Wan et al., 2016; Plaschka et al., 2017). In the newly formed spliceosome, U5 snRNA associates with the exon

sequences at the 5' splice site (via non-Watson-Crick nucleotide-nucleotide contact) and possibly interacts with the 3' splice site as well (Newman and Norman, 1992; Wyatt et al., 1992; Sontheimer and Steitz, 1993; Newman et al., 1995; Newman, 1997). Before the first transesterification reaction (first step of splicing) occurs, the spliceosomal RNA-RNA interactions undergo a complex dynamic rearrangement (Wassarman and Steitz, 1992; Nilsen, 1994; Will and Lührmann, 2011). Specifically, U6 snRNA dissociates from U4 snRNA, displaces U1 snRNA in interacting with the 5' splice site (Staley and Guthrie, 1998; Nilsen, 2003), and forms new base-paired duplexes with U2 snRNA (Datta and Weiner, 1991; Wu and Manley, 1991; Madhani and Guthrie, 1992) that are known to be part of the catalytic center (Yean et al., 2000; Wahl et al., 2009; Fica et al., 2013; Zhang et al., 2018). At this point, the first transesterification reaction takes place, leading to the formation of a new complex (complex C), which contains splicing intermediates (Jurica et al., 2004; Will and Lührmann, 2011). After additional conformational changes, the second transesterification reaction (the second step of splicing) occurs, generating matured mRNA and lariat intron products (Will and Lührmann, 2011; Yan et al., 2017). It is important to note that all the interactions occurring in the spliceosome are highly orchestrated, thus allowing for accurate and efficient splicing.

Epitranscriptomics of snRNAs

The emerging field of epitranscriptomics is continually unveiling additional unknown levels of complexity of the transcriptome (Nachtergaele and He, 2017; Roundtree et al., 2017). RNA modifications can have a major impact in RNA folding and function in all types of RNA, including snRNAs. All snRNAs (except for U6) have a 2,2,7-trimethylated 5' guanosine cap (U6 possesses a γ -monomethyl guanosine cap; Singh and Reddy, 1989).

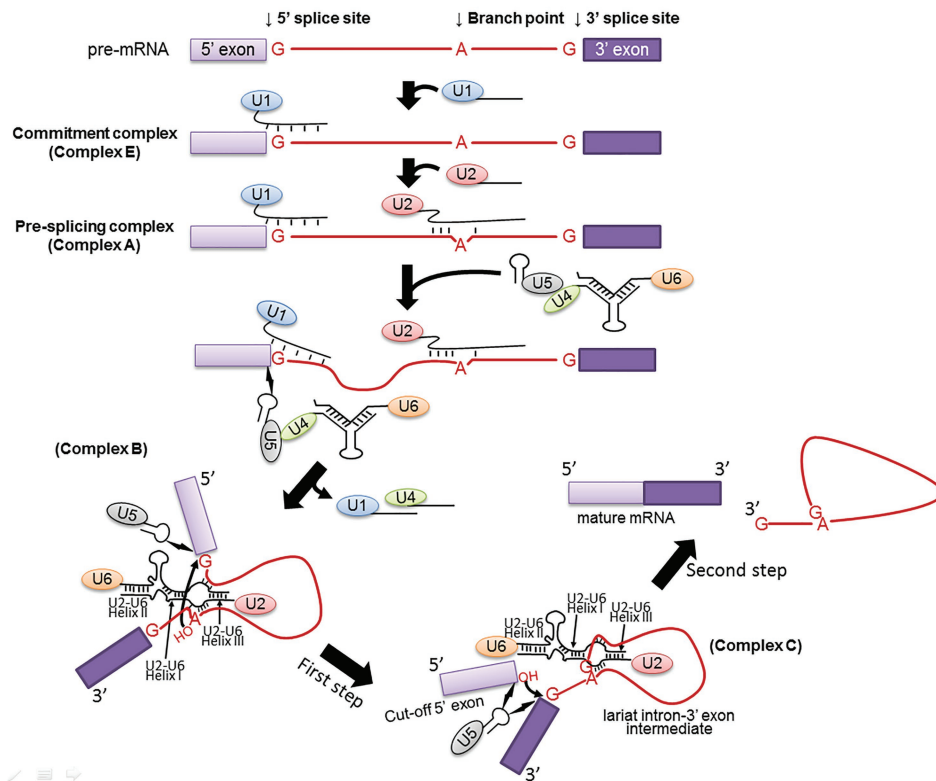


FIGURE 2 | Spliceosome assembly. Spliceosome assembly is a dynamic multi-step process. Shown are several important steps resulting in Complexes E, A, B, and C. The branch site, the 5' and 3' splice sites are also shown. The 5' and 3' exons are in light and dark purple boxes, respectively, and the intron is in red line (and red letters). U1, U2, U4, U5, and U6 snRNPs are also depicted. RNA-RNA interactions, including U1–5' splice site, U2–branch site, U6–5' splice site, U4–U6, and U2–U6 (Helices I, II, and III) are indicated as well. The two curved arrows indicate nucleophilic attacks (transesterification reactions). The lightning bolts indicate non-Watson-Crick nucleotide-nucleotide contacts.

Further, numerous internal nucleotides are modified by pseudouridylation (5-ribosyl isomers of uridine), 2'-O-methylation, and in some cases, base methylation [e.g., N^6 -methyladenosine (m^6A) and N^2 -methylguanosine (m^2G)] (Reddy and Busch, 1988; Massenet et al., 1998; Bohnsack and Sloan, 2018). Notably, the modified nucleotides in spliceosomal snRNAs are remarkably conserved from species to species. For instance, various vertebrate snRNAs contain identical modified nucleotides at identical sites. Although a relatively small number of modified nucleotides have been identified in yeast snRNAs, those that are modified always have counterparts in higher eukaryotic snRNAs (Adachi and Yu, 2014). Furthermore, almost all the modified nucleotides are concentrated in regions that are functionally relevant to the splicing process (Adachi and Yu, 2014). Together, the conservation and the strategic location of these modified nucleotides strongly point to their importance in the process of spliceosome assembly and splicing. It should be pointed out that various post-transcriptional modifications, which are catalyzed by different types of machinery, generate diversity in the snRNAs that likely contribute to pre-mRNA splicing regulation.

In addition to U1, U2, U4, U5, and U6 snRNAs (major), there is also a set of minor spliceosomal snRNA species (U11, U12, U4atac, and U6atac) that participate in the splicing of

a minor class of introns (Tarn and Steitz, 1997; Will and Lührmann, 2005; Turunen et al., 2013). Some of these minor snRNAs are also post-transcriptionally modified. This review will describe spliceosomal snRNA modifications (major and minor classes of snRNAs), focusing on the mechanisms and functions of these modifications.

RNA-DEPENDENT VS. RNA-INDEPENDENT snRNA MODIFICATION MECHANISMS

The most abundant modified nucleotides in snRNAs are pseudouridine (Ψ) and 2'-O-methyl residues, whereas m^6A and m^2G are rarely present in only a few snRNA species (Epstein et al., 1980; Reddy et al., 1981b; Massenet et al., 1998; Bohnsack and Sloan, 2018). RNA modification can be catalyzed by either RNA-dependent or RNA-independent mechanism (De Zoysa and Yu, 2017; Meier, 2017; Wiener and Schwartz, 2020). While the RNA-independent mechanism depends on stand-alone protein enzymes capable of recognizing the substrates and catalyzing the chemical reaction, the RNA-dependent mechanism typically relies on RNA-protein enzyme complexes (RNPs),

each of which is composed of one small RNA and several proteins. In each RNP, the RNA component functions as a guide recognizing the substrate RNA, and one of the protein components has enzymatic activity catalyzing the chemical reaction (Yu et al., 2005).

RNA-Dependent Mechanisms

Both snRNA 2'-O-methylation and pseudouridylation are catalyzed by RNA-dependent mechanisms in high eukaryotes. Specifically, a family of box C/D RNPs is responsible for snRNA 2'-O-methylation (Balakin et al., 1996; Cavaillé et al., 1996; Kiss-László et al., 1996), and another family of RNPs, the box H/ACA RNP family, is accountable for snRNA pseudouridylation (Balakin et al., 1996; Ganot et al., 1997; Ni et al., 1997). Each member of the box C/D RNP family is composed of one unique box C/D RNA and four common core proteins (Fibrillarin, also known as Nop1, Nop56, Nop58, and Snu13). Likewise, members of the box H/ACA RNP family each consist of one unique box H/ACA RNA and a set of four common proteins (Dyskerin, also known as Nap57 or Cbf5, Nhp2, Nop10, and Gar1). Both box H/ACA and C/D snoRNAs are usually intron-encoded in mammals and are matured through splicing and processing (Tycowski et al., 1996; Hirose et al., 2003, 2006).

Mature box C/D RNAs and box H/ACA RNAs both fold into a unique secondary structure (Figure 3). The C/D RNAs, despite their sequence differences, form a signature structure with a terminal stem and two single-stranded sequences sandwiched between box C (RUGAUGA, where R is a purine) and box D' (CUGA) and between box C' (RUGAUGA) and box D (CUGA), respectively (Figure 3). It turns out that the single-stranded sequences serve as guides that base-pair with the substrate RNAs, forming a 10–21-nt duplex and specifying the target nucleotide that is precisely five nucleotides upstream from box D (or box D'). Once the target nucleotide is identified, fibrillarin (Nop1), one of the four box C/D RNP core proteins and a methyltransferase, delivers the methyl group to the target nucleotide at the 2'-O position. The “box D + 5” rule for box C/D RNA-guided snRNA 2'-O-methylation has been verified in various organisms, including *Xenopus*, mouse, and human, indicating that box C/D RNA-guided 2'-O-methylation of snRNA is universal in high eukaryotes (Kiss-László et al., 1996; Karijovich and Yu, 2010). Given that 2'-O-methylation is a sugar-ring modification, it can occur to any nucleotides. Interestingly, to date, no 2'-O-methylated residues have been identified in *Saccharomyces cerevisiae* snRNA.

Similarly, despite their sequence differences, all members of the box H/ACA RNA family fold into a structure known as the “hairpin-hinge (H box, ANANNA, where N represents any nucleotide)-hairpin-tail (ACA box)” structure. In this structure, there are two independent hairpins, in each of which there exists an internal loop (single-stranded) that serves as a guide. In essence, the guide sequences are in two separate segments in the linear RNA that are brought together in internal loops within the hairpins. Base-pairing between the bipartite guide sequence and the snRNA, positions the target uridine at the base of the upper stem of the hairpin, leaving it unpaired within the internal loop (so-called “pseudouridylation pocket”) and

located about 14–16 nucleotides upstream of box H or box ACA (Figure 3). When the target uridine is brought to the pocket, Dyskerin (NAP57 or Cbf5), one of the four box H/ACA RNP core proteins and also a pseudouridylation, converts the uridine to pseudouridine. The box H/ACA RNA-guided pseudouridylation mechanism has been tested and verified in various high eukaryotic systems (Terns and Terns, 2006). Recent analyses have further demonstrated that a minimum number of eight base-pairs between the guide and substrate in the pseudouridylation pocket is required for efficient pseudouridylation (Caton et al., 2018; De Zoysa et al., 2018).

In *S. cerevisiae*, spliceosomal snRNA pseudouridylation is more complex. Both RNA-dependent or RNA-independent mechanisms are used (Massenet et al., 1999; Ma et al., 2003, 2005; Yu et al., 2011). The yeast box H/ACA RNAs can be either encoded in introns of protein-coding genes or in independent transcripts of non-protein-coding genes. In *S. cerevisiae*, snRNA pseudouridylation can also be achieved *via* an RNA-independent (protein-only) mechanism.

RNA-Independent Mechanisms

Stand-alone (protein-only) pseudouridine synthases (Pus) can recognize the substrate and perform the uridine isomerization reaction in a site-specific manner. There are nine Pus enzymes in yeast (Pus1–9) and 11 human homologs (Pus1, Pus3, TruB1, TruB2, RusD1, RusD2, RusD3, RusD4, Pus7, Pus7L, and Pus10; Rintala-Dempsey and Kothe, 2017). In yeast, only Pus1 and Pus7 were identified as capable of RNA-independent pseudouridylation of snRNAs (Massenet et al., 1999; Ma et al., 2003; Basak and Query, 2014; Schwartz et al., 2014). Yeast Pus1 and Pus7 belong to the TruA and TruD families of pseudouridine synthases, respectively. Yeast Pus1 is localized in the nucleus and targets not only snRNAs but also other types of RNA, showing broad substrate specificity (Motorin et al., 1998). In addition to its pseudouridylation activity, Pus1 is also involved in tRNA biogenesis (Simos et al., 1996). Pus7 localizes in the nucleus and cytoplasm of cells and, like Pus1, can also target different RNAs. Pus7 is known for being able to recognize substrates relying on both the sequence (consensus UNUAR motif, where target uridine is underlined) and the secondary structure surrounding the target uridine (Ma et al., 2003; Urban et al., 2009). In yeast, heat shock conditions can further induce the activity of this enzyme (Wu et al., 2011; Schwartz et al., 2014).

Other exotic base methylations, such as m⁶A and m²G, have also been identified in snRNAs. These base modifications are catalyzed by RNA-independent enzymatic machineries. Specifically, the m⁶A modification is catalyzed by the m⁶A methyltransferases, protein-only enzymes known as m⁶A writers (Frye et al., 2018). In this reaction, a methyl group is attached to N6 of an adenosine within a specific RNA motif, resulting in the production of m⁶A methylated RNA (RNA containing an N⁶-methyladenosine). m⁶A has been identified in human U2, U4, and U6 snRNAs, as well as in *S. pombe* U2 and U6 snRNAs. While m⁶A writers METTL4 and METTL16 are responsible for the formation of m⁶A in U2 and U6 snRNAs, respectively, the exact enzyme for the formation of m⁶A in human U4 snRNA (Reddy et al., 1981b) remains unknown.

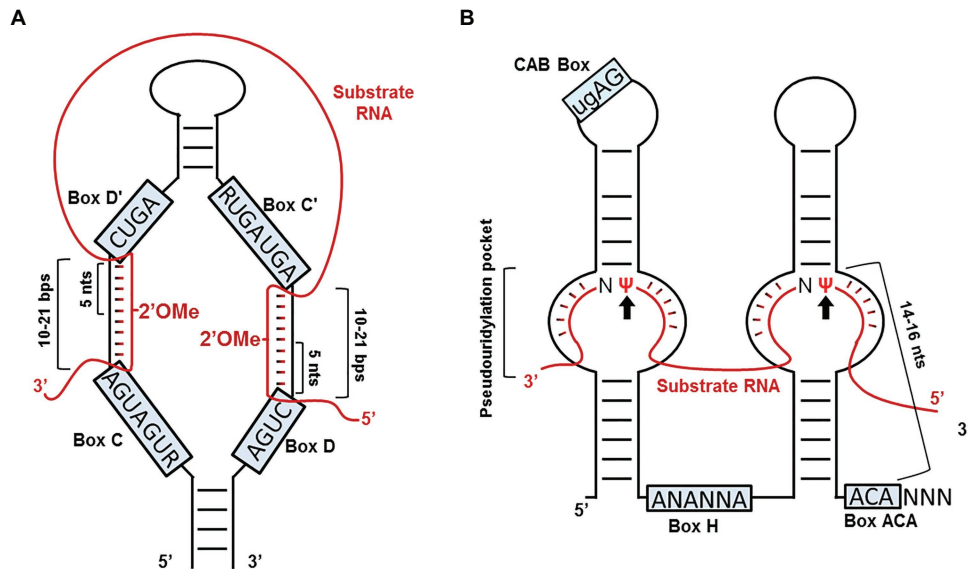


FIGURE 3 | RNA-guided snRNA modifications. **(A)** Box C/D RNA-guided 2'-O-methylation. Box C/D guide RNA (black line) and its substrate RNA (red line) are shown. Boxes C, D', C', and D of the box C/D RNA are indicated. Base-pairing interactions between the two RNAs (usually 10–21 base-pairs) are also depicted. The fifth nucleotide upstream from box D' is 2'-O-methylated (indicated), and so is the fifth nucleotide upstream from box D (also indicated). **(B)** Box H/ACA RNA-guided pseudouridylation. Both box H/ACA RNA (black line) and its substrate RNA (red line) are indicated. Boxes H, ACA, and CAB (a Cajal body localization signal sequence) of the box H/ACA RNA are also indicated. Base-pairing interactions between the guide and substrate RNAs in the pseudouridylation pockets (internal loops) are shown. The modified target uridines (Ψs) are indicated by two arrows. The 14–16 nt distance between the target uridine and box H (or box ACA) is indicated as well.

The m²G modification results from the methylation of N2 of guanine (the guanosyl amino group at the position C2) and was initially identified in tRNAs (Grosjean et al., 1995). It is catalyzed by a class of enzymes known as guanine-(N2)-methyltransferases, which have been identified in several species (Sindhuphak et al., 1985; Sergiev et al., 2006) and seem to have a substrate consensus sequence (UGGC, the target guanosine is underlined). The m²G modification was detected in U6 snRNA decades ago (Epstein et al., 1980), although the nucleotide target sequence in this case (AmGGGA, target guanosine is underlined, and the first nucleotide is 2'-O-methylated) deviates from the consensus. In the context of RNA duplexes, this modification is considered as iso-energetic to guanosine (Rife et al., 1998).

MODIFIED NUCLEOTIDES IN SPLICEOSOMAL snRNAs

It has long been known that mammalian major spliceosomal snRNAs contain a large number of modified nucleotides (Table 1). Specifically, there are two, fourteen, three, three, and three Ψs in mammalian U1, U2, U4, U5, and U6 snRNAs, respectively. There are also three, ten, four, five, and eight 2'-O-methylated residues in mammalian U1, U2, U4, U5, and U6 snRNAs, respectively. In addition, mammalian U2, U4, and U6 snRNAs each contain an m⁶A. Further, mammalian U6 snRNA contains an m²G as well. In contrast, there are a

total of only six constitutively formed Ψs in *S. cerevisiae* snRNAs, including two in U1 snRNA, three in U2 snRNA, and one in U5 snRNA. In addition, *S. cerevisiae* U2 snRNA can be pseudouridylated at two novel sites under stress conditions. A set of minor class spliceosomal snRNAs also exists in mammals including U11, U12, U4atac, and U6atac. These snRNAs contain several Ψs and 2'-O-methylated residues as well. Over the years, the mechanisms (what enzymes are involved) and functions of these modifications have been studied, accumulating a wealth of knowledge.

U1 snRNA

The U1 snRNA is one of the most abundant snRNAs in different species. Only two types of modifications, namely pseudouridylation and 2'-O-methylation, have been detected in mammalian U1 snRNA. Together, there is a total of five modified nucleotides, including Ψ5, Ψ6, Am1, Um2, and Am70 (Reddy et al., 1981a; Reddy and Busch, 1988; Massenet et al., 1999; Kiss et al., 2004; Gu et al., 2005; Krogh et al., 2017; Table 1). In yeast U1 snRNA, only Ψ5 and Ψ6 are identified (Massenet et al., 1999); no 2'-O-methylated residues have been detected. At present, it is still not clear whether the RNA-dependent or RNA-independent mechanism catalyzes pseudouridylation at positions 5 and 6 in yeast U1 snRNA. In human U1 snRNA, pseudouridylation at Ψ5, Ψ6 positions (Branlant et al., 1980) is catalyzed by H/ACA RNP machinery and guided by ACA47 (Kiss et al., 2004) and U109 (Gu et al., 2005), respectively. Mammalian U1 snRNA

TABLE 1 | Yeast and human RNA modifications present in snRNAs with respective guide RNAs (when applicable) and catalyst machinery (Adachi and Yu, 2014).

Species	snRNA	RNA modification	Guide RNA	Catalyst	References
Yeast	U1	Ψ5		Likely non-dependent on box H/ ACA mechanism	Massenet et al., 1999
		Ψ6		Likely non-dependent on box H/ ACA mechanism	Massenet et al., 1999
	U2	Ψ35	-	Pus7p	Massenet et al., 1999; Ma et al., 2003; Schwartz et al., 2014
		Ψ42	snR81	Cbf5	Massenet et al., 1999; Ma et al., 2005; Schwartz et al., 2014
		Ψ44	-	Pus1p	Massenet et al., 1999; Schwartz et al., 2014
		Ψ56 (stress-induced)	-	Pus7p	Wu et al., 2011
		Ψ93 (stress-induced)	snR81	Cbf5	Wu et al., 2011
	U5	Ψ99		Cbf5p	Massenet et al., 1999; Schwartz et al., 2014
	U6	Ψ28 (filamentous growth induced)		Pus1p	Basak and Query, 2014
Human	U1	Ψ5	ACA47	H/ACA RNP	Branlant et al., 1980; Reddy et al., 1981a; Kiss et al., 2004
		Ψ6	U109	H/ACA RNP	Branlant et al., 1980; Reddy et al., 1981a; Gu et al., 2005
	U2	Am1 Um2 Am70 Ψ6	scaRNA7 (U90)	Fibrillarin NR	Krogh et al., 2017 Krogh et al., 2017 Darzacq et al., 2002; Krogh et al., 2017 Dönmez et al., 2004; Deryusheva et al., 2012
		Ψ7	U100	H/ACA RNP	Dönmez et al., 2004; Schattner et al., 2006; Deryusheva et al., 2012
		Ψ15		NR	Dönmez et al., 2004; Deryusheva et al., 2012
		Ψ34	scaRNA8 (U92)	H/ACA RNP or Pus7p	Shibata et al., 1975; Darzacq et al., 2002; Kiss et al., 2004; Deryusheva et al., 2012
		Ψ37	ACA45	H/ACA RNP	Shibata et al., 1975; Kiss et al., 2004; Deryusheva et al., 2012
		Ψ39	ACA26	H/ACA RNP	Shibata et al., 1975; Kiss et al., 2004; Deryusheva et al., 2012
		Ψ41	ACA26	H/ACA RNP	Shibata et al., 1975; Kiss et al., 2004; Deryusheva et al., 2012; Deryusheva and Gall, 2018
		Ψ43	scaRNA8 (U92)	H/ACA RNP or Pus1p	Deryusheva et al., 2012; Deryusheva and Gall, 2017
		Ψ44	scaRNA8 (U92)	H/ACA RNP	Shibata et al., 1975; Darzacq et al., 2002; Kiss et al., 2004; Deryusheva et al., 2012
		Ψ54	U93	H/ACA RNP	Shibata et al., 1975; Reddy and Busch, 1988; Kiss et al., 2002, 2004; Schattner et al., 2006; Deryusheva et al., 2012
		Ψ58 Ψ60 Ψ89	SNORA11 ACA35	H/ACA RNP NR H/ACA RNP	Deryusheva et al., 2012, 2020 Deryusheva et al., 2012 Shibata et al., 1975; Kiss et al., 2004; Deryusheva et al., 2012
		Ψ91 Am1 Um2 Gm11		NR	Deryusheva et al., 2012 Dönmez et al., 2004; Krogh et al., 2017 Krogh et al., 2017
		Gm12	scaRNA2 (HBII-382)	Fibrillarin	Dönmez et al., 2004; Deryusheva et al., 2012; Krogh et al., 2017
		Gm19	SNORD89	Fibrillarin	Dönmez et al., 2004; Deryusheva et al., 2012, 2020; Krogh et al., 2017
		Gm25 Am30	scaRNA9 (mgU2-19/30)	Fibrillarin	Dönmez et al., 2004; Deryusheva et al., 2012; Krogh et al., 2017
		Gm25 Am30	scaRNA2 scaRNA9 (mgU2-19/30)	Fibrillarin Fibrillarin	Deryusheva et al., 2012; Krogh et al., 2017 Deryusheva et al., 2012; Krogh et al., 2017
		Cm40	MBII-19		Hüttenhofer et al., 2001; Deryusheva et al., 2012; Krogh et al., 2017

(Continued)

TABLE 1 | Continued

Species	snRNA	RNA modification	Guide RNA	Catalyst	References
	U4	Um47	scaRNA28	Fibrillarin	Deryusheva et al., 2012; Krogh et al., 2017
		Cm61	scaRNA2 (mgU2–25/61)	Fibrillarin	Tycowski et al., 2004; Deryusheva et al., 2012; Krogh et al., 2017
		m ⁶ Am30		METTL4/Fibrillarin	Mauer et al., 2019; Chen et al., 2020; Goh et al., 2020
		Ψ4		NR	Zerby and Patton, 1997
		Ψ72		NR	Zerby and Patton, 1997
		Ψ79		NR	Zerby and Patton, 1997
		Am1			Krogh et al., 2017
		Gm2			Krogh et al., 2017
		Cm8	scaRNA17 (MBII-119)	Fibrillarin	Krogh et al., 2017
		Am65	scaRNA5 (U87)	Fibrillarin	Darzacq et al., 2002; Krogh et al., 2017
	U5	m ⁶ A100			Reddy et al., 1981b
		Ψ43	ACA57	H/ACA RNP	Krol et al., 1981; Kiss et al., 2004
		Ψ46	U85	H/ACA RNP	Krol et al., 1981; Jády and Kiss, 2001
		Ψ53	scaRNA13 (U93)	H/ACA RNP	Shibata et al., 1975; Reddy and Busch, 1988; Kiss et al., 2002; Schattner et al., 2006
		Am1			Krogh et al., 2017
	U6	Um2			Krogh et al., 2017
		Gm37			Krol et al., 1981; Krogh et al., 2017
		Um41	scaRNA5/6 (U87)	Fibrillarin	Krol et al., 1981; Darzacq et al., 2002; Krogh et al., 2017
		Cm45	scaRNA10 (U85)	Fibrillarin	Krol et al., 1981; Jády and Kiss, 2001; Darzacq et al., 2002; Krogh et al., 2017
		Ψ31	ACA65	H/ACA RNP	Epstein et al., 1980; Schattner et al., 2006
		Ψ40	ACA12	H/ACA RNP	Epstein et al., 1980; Kiss et al., 2004
		Ψ86	ACA65	H/ACA RNP	Epstein et al., 1980; Schattner et al., 2006; Deryusheva et al., 2020
		Am47	SNORD7 (mgU6–47)	Fibrillarin	Epstein et al., 1980; Tycowski et al., 1998; Krogh et al., 2017
		Am53	SNORD8/9 (mgU6–53)	Fibrillarin	Epstein et al., 1980; Ganot et al., 1999; Krogh et al., 2017
		Gm54			Epstein et al., 1980; Krogh et al., 2017
		Cm60	SNORD67 (HBII-166)	Fibrillarin	Epstein et al., 1980; Hüttenhofer et al., 2001; Lestrade and Weber, 2006; Krogh et al., 2017
		Cm62	SNORD94 (U94)	Fibrillarin	Epstein et al., 1980; Vitali et al., 2003; Krogh et al., 2017
		Cm63			Epstein et al., 1980; Krogh et al., 2017
		Am70			Epstein et al., 1980; Krogh et al., 2017
		Cm77	SNORD10 (mgU6–77)	Fibrillarin	Epstein et al., 1980; Tycowski et al., 1998; Krogh et al., 2017
	U12	m ⁶ A43		METTL16	Epstein et al., 1980; Shimba et al., 1995; Aoyama et al., 2020
		m ² G72		Guanine-(N2)-methyltransferases	Epstein et al., 1980
		Ψ19	scaRNA21 (ACA68)	H/ACA RNP	Massenet and Branlant, 1999; Schattner et al., 2006; Deryusheva et al., 2012
		Ψ28	ACA66	H/ACA RNP	Massenet and Branlant, 1999; Schattner et al., 2006; Deryusheva et al., 2012
		Am8			Tycowski et al., 2004; Deryusheva et al., 2012
	U4atac	Gm18			Tycowski et al., 2004; Deryusheva et al., 2012
		Gm22	scaRNA17 (mgU12–22/U4–8)	Fibrillarin	Darzacq et al., 2002; Tycowski et al., 2004; Deryusheva et al., 2012
		Ψ12		NR	Massenet and Branlant, 1999; Deryusheva et al., 2012
	U6atac	Am1			Deryusheva et al., 2012
		Am2			Deryusheva et al., 2012
	U6atac	Gm19			Deryusheva et al., 2012
		Ψ83	SCARNA21	NR	Deryusheva et al., 2012; Massenet and Branlant, 1999; Jorjani et al., 2016

2'-O-methylation at position 70 (Krogh et al., 2017) is likely catalyzed by an RNA-dependent mechanism, given that a box C/D RNA (SCARNA7, also known as U90) has been identified to target this site (Darzacq et al., 2002).

Functions of Ψ s and 2'-O-Methylated Residues Residing in U1 snRNA

Notably, Ψ 5 and Ψ 6 are within the first 10-nucleotide sequence known to base-pair with the 5' splice site of pre-mRNA during splicing. Given that Ψ can affect local RNA structure and enhance base-pairing and base stacking (Ge and Yu, 2013), Ψ 5 and Ψ 6 are believed to be important in the recognition process of the 5' splice site. Indeed, an *in vitro* splicing assay performed to test competitive usage of two 5' splice sites suggested that the two Ψ s in the U1 snRNA could provide a competitive advantage in the 5' splice site selection (Roca, 2005). In another study, it was shown that Ψ 5 or Ψ 6 can be bulged out in certain duplexes consisting of 5' splice sites and U1 snRNA (Roca et al., 2012). Thermodynamic analysis of these duplexes confirmed the stabilization properties of Ψ s (and 2'-O-methylated residues at the first two positions) in this context, possibly by improving the base stacking of the helix. These results are consistent with the results of a previous work showing that a Ψ in a Ψ -G base pair strengthens the interaction between U1 snRNA and the 5' splice site (Freund, 2003). While these conclusions are exciting and make sense, they seem somewhat contradictory to an earlier work of Will et al. (1996), where the authors showed that U1 snRNA-depleted mammalian cell extracts could still be reconstituted for splicing when adding *in vitro*-transcribed (therefore unmodified) U1 snRNA. However, it could well be that the extracts could modify the *in vitro*-transcribed U1 snRNA upon addition. Alternatively, although the unmodified U1 snRNA could still support splicing, it may not be as active as the modified U1 snRNA. The reconstitution assay using unmodified U1 snRNA probably did not reflect the contributions of Ψ 5 and Ψ 6 in 5' splice site recognition. As for the function of 2'-O-methylated residue at position 70 (Am70) in mammalian U1 snRNA, not much is known.

U2 snRNA

U2 snRNA is the most extensively modified among all spliceosomal snRNAs (Shibata et al., 1975; Reddy and Busch, 1988). There are fourteen Ψ s, ten 2'-O-methylated residues, and one m⁶Am residue in vertebrate U2 snRNA (Table 1). Given that a near-complete set of box H/ACA RNAs and a complete set of box C/D RNAs are identified and that they can account for almost all known pseudouridylation and 2'-O-methylation sites, it is believed that the RNA-dependent mechanisms are responsible for the formation of virtually all the Ψ s and 2'-O-methylated residues (except for the first two 2'-O-methylated residues, Am1 and Um2) in vertebrate U2 snRNA (Hüttenhofer et al., 2001; Tycowski et al., 2004; Schattner et al., 2006; Deryusheva and Gall, 2017, 2018; Bohnsack and Sloan, 2018; Deryusheva et al., 2020). There are three Ψ s and no 2'-O-methylated residues identified in *S. cerevisiae* U2 snRNA.

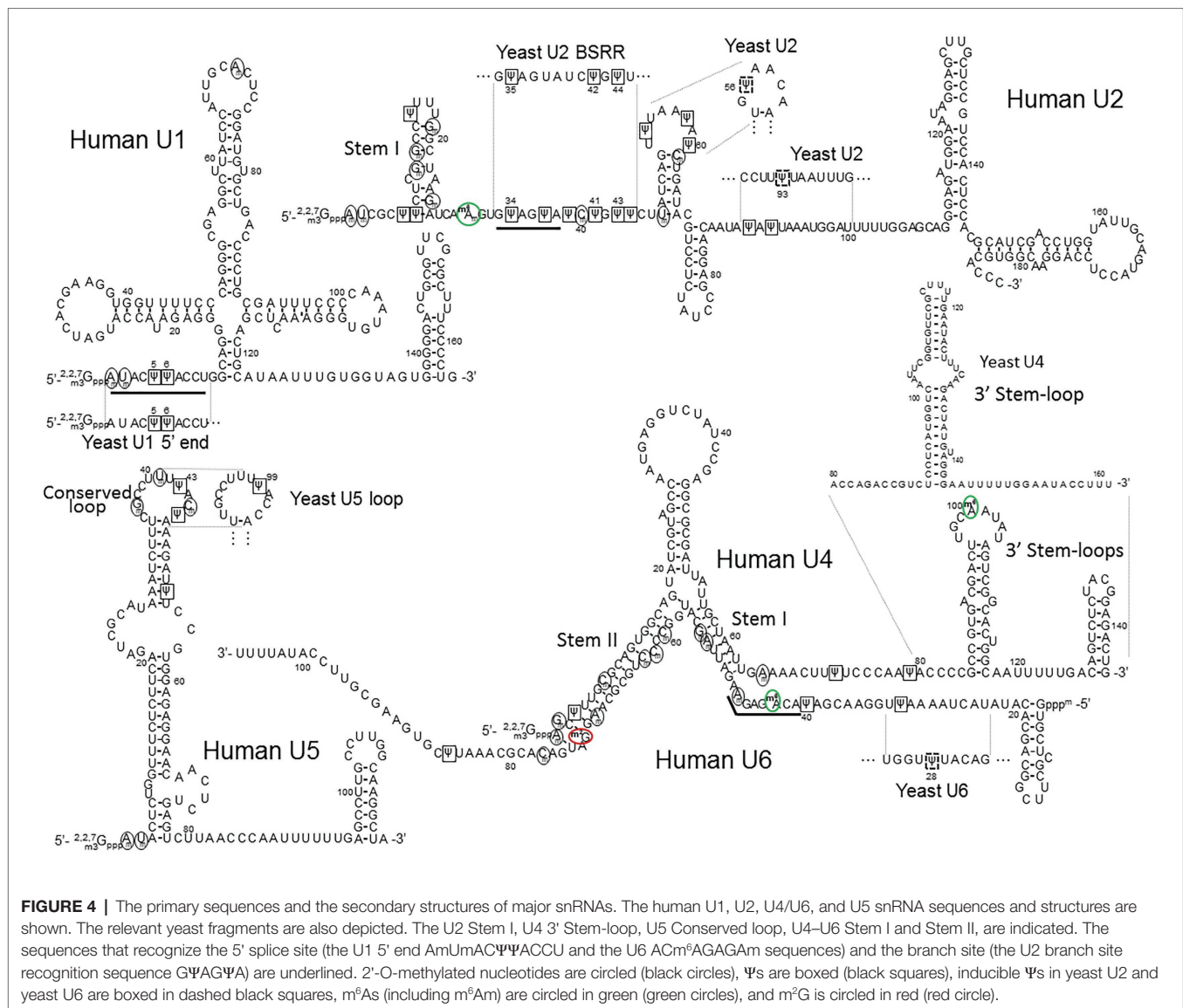
The formation of Ψ at different positions within yeast U2 snRNA can be catalyzed by either RNA-dependent or RNA-independent mechanism (Massenet and Branlant, 1999; Ma et al., 2003, 2005; Schwartz et al., 2014).

Functions of Ψ s and 2'-O-Methylated Residues Residing in Vertebrate U2 snRNA

Many of the U2 snRNA Ψ s and 2'-O-methylated residues have been tested for function, and they are virtually all important for splicing. For example, using the *Xenopus* oocyte microinjection system, Yu et al. showed that only modified U2 snRNA (but not *in vitro* transcribed, unmodified U2 snRNA) was able to restore splicing in U2 snRNA-depleted oocytes, indicating that modified nucleotides of U2 snRNA are crucial for pre-mRNA splicing. Subsequently, they mapped the important modified nucleotides to the 5' end region. They further demonstrated that the modified nucleotides within the 5' end region are essential for the formation of functional U2 snRNP and splicing complexes (Yu et al., 1998). In a different study, Dönmez et al. tested these modified nucleotides individually and demonstrated that three Ψ s (Ψ 6, Ψ 7, and Ψ 15) and five 2'-O-methylated residues (Am1, Um2, Gm11, Gm12, and Gm19), located in the 5'-end region (first 24 nt) of human U2 snRNA, were required for efficient pre-mRNA splicing. While the Ψ s have a cumulative effect in splicing, four of the five 2'-O-methylated residues (Am1, Um2, Gm12, and Gm19) were essential for activity (Dönmez et al., 2004). Soon after, it was shown that the Ψ s in the branch site recognition region (BSRR, Ψ 34, Ψ 37, Ψ 39, Ψ 41, Ψ 43, and Ψ 44) are also essential for pre-mRNA splicing (Zhao and Yu, 2004).

Functions of Ψ s Residing in *Saccharomyces cerevisiae* U2 snRNA

Unlike vertebrate U2 snRNA, there are only three Ψ s that are normally present in *S. cerevisiae* U2 snRNA. They are located at positions 35, 42, and 44 (equivalent to vertebrate U2 snRNA at positions 34, 41, and 43) in the BSRR (Figure 4; De Zoysa and Yu, 2017). Pseudouridylation at these positions, 35, 42, and 44, is catalyzed by Pus7, snR81 RNP, and Pus1, respectively. Among these pseudouridylation enzymes, Pus1 and Pus7 are stand-alone protein pseudouridylases, whereas snR81 RNP is a genuine box H/ACA RNP complex. Several lines of evidence indicate that all these Ψ s contribute to branch site recognition during pre-mRNA splicing. For example, Yang et al. showed that a Pus7-deleted strain exhibited reduced levels of splicing and cell growth in certain conditions (Yang et al., 2005). By analyzing splicing in yeast strains deleted of any of the three pseudouridylases (in all combinations), Wu et al. (2016a) found that the three Ψ s, in coordination with the ATPase Prp5, play an essential role in recognizing the branch site at an early stage during spliceosome assembly. Furthermore, structural studies of U2 snRNA showed the importance of the Ψ 35 in splicing function. Specifically, it was proposed that Ψ 35 affected the local RNA structure to expose the branch site adenosine 2'-OH group, making it available for nucleophilic attack on the 5' splice site – the first transesterification reaction or the



first step of splicing (Newby and Greenbaum, 2001, 2002). However, recent work from Kielkopf's lab (Kennedy et al., 2019) suggested that the role of Ψ35 could be indirect, perhaps more reliant on auxiliary factors.

Inducible Ψ Formation in *Saccharomyces cerevisiae* U2 snRNA

In yeast U2 snRNA, two non-constitutive modifications (Ψ56 and Ψ93) can also be identified in stress conditions (Wu et al., 2011). Pseudouridylation at positions 56 and 93 is catalyzed, respectively, by Pus7, which normally catalyzes the formation of Ψ35, and snR81 RNP, which is responsible for the constitutive formation of Ψ42. The induction of Ψ formation at these two positions is through the Tor-signaling pathway under nutrient deprivation conditions (Wu et al., 2016b). The formation of Ψ56 can also be induced by heat shock. Although different types of machinery, stand-alone protein Pus7, and box H/ACA RNP snR81, catalyze the formation of Ψ56 and Ψ93,

respectively, it appears that the imperfect sequences surrounding the inducible target sites (positions 56 and 93) vs. those flanking the constitutively modified sites (positions 35 and 42) could explain their inducibility (Wu et al., 2011). Induced formation of Ψ56 and Ψ93 plays a role in pre-mRNA splicing, perhaps by helping alter the U2 snRNA structure. Indeed, Ψ56- and Ψ93-mediated U2 snRNA structural change was observed in a structural study (van der Feltz et al., 2018).

A Rare Type of Modification (m⁶Am) in Mammalian U2 snRNA

Besides Ψs and 2'-O-methylated residues, one of the 2'-O-methylated adenosines (Am30) in human U2 snRNA is also base methylated in the N⁶-position (m⁶Am30; Figure 4). This modification is conserved through evolution, from yeast (*S. pombe*) to humans, in the same nucleotide position (Gu et al., 1996). Since m⁶Am30 is located almost immediately upstream of the branch site recognition sequence, it has recently

drawn some attention. Chen et al. (2020) and Goh et al. (2020) have independently identified METTL4 as the methyltransferase responsible for the formation of m⁶Am of mammalian U2 snRNA at position 30. In their study, Chen et al. generated knocked-out METTL4 human cells and observed an effect on splicing in those cells when compared to wild-type cells. However, the direct link between the m⁶Am30 modification in U2 snRNA and splicing was not definitively established. Nonetheless, they demonstrated that METTL4 is the enzyme responsible for the m⁶Am30 modification. Specifically, using recombinant METTL4 and a fragment of U2 snRNA substrate, they carried out an *in vitro* biochemical assay and detected m⁶Am30 formation. However, the 2'-O-methylation of A30 is a prerequisite for the base methylation to occur. Thus, it appears that 2'-O-methylated adenosine (Am30), rather than unmodified adenosine (A30), is the true substrate. Additionally, the level of base methylation could also be severely reduced when changing the 5' and 3' nucleotides, pointing toward sequence recognition by the METTL4. In an independent study, Goh et al. (2020) confirmed METTL4 as the enzyme responsible for the m⁶Am30 modification (with Am being the true substrate). The authors also confirmed the base modification identity with HPLC-MS/MS (Goh et al., 2020). Using transcriptome-wide sequencing, they further showed that this modified nucleotide contributed to splicing regulation. As to the possible mechanism, the authors of this study hypothesized that the modified adenosine (m⁶Am30) could potentially be involved in the recruitment of U2 snRNA to the branch site by U2AF (Zamore et al., 1992; Zhang et al., 1992), a heterodimer that recognizes and binds to the 3' splice site at an early stage of spliceosome assembly (prior to complex A formation), thus affecting the pre-mRNA splicing process (Figure 2).

U4, U5, and U6 snRNAs

There is also a large number of modified nucleotides in mammalian U4, U5, and U6 snRNAs (Table 1). In total, human U4 snRNA has three Ψs (Ψ4, Ψ72, and Ψ79; Zerby and Patton, 1997), four 2'-O-methylated residues (Am1, Gm2, Cm8, and Am65; Krogh et al., 2017), and one m⁶A (m⁶A100; Reddy et al., 1981b). Human U5 contains several Ψs (Ψ43, Ψ46, and Ψ53; Shibata et al., 1975; Krol et al., 1981) and many 2'-O-methylated residues (Am1, Um2, Gm37, Um41, and Cm45; Krol et al., 1981; Krogh et al., 2017). The human U6 snRNA also has a large number of modified nucleotides, including three Ψs (31, 40, and 86), eight 2'-O-methylated residues (Am47, Am53, Gm54, Cm60, Cm62, Cm63, Am70, and Cm77), one m⁶A (m⁶A43), and one m²G (m²G72). Most of these modifications were identified decades ago (Epstein et al., 1980; Reddy and Busch, 1988). Similar to mammalian U2 snRNA modifications, pseudouridylation and 2'-O-methylation (except for the first two methylated residues) of human U4, U5, and U6 snRNAs are likely catalyzed by RNA-guided modification mechanisms (Tycowski et al., 1998; Ganot et al., 1999; Hüttenhofer et al., 2001; Jády and Kiss, 2001; Darzacq et al., 2002; Kiss et al., 2002, 2004; Vitali et al., 2003; Lestrade and Weber, 2006; Schattner et al., 2006; Bohnsack and Sloan, 2018). Interestingly, there is only one Ψ (Ψ99) and no 2'-O-methylated nucleotide

in *S. cerevisiae* U5 snRNA (Massenet et al., 1999). No Ψ nor 2'-O-methylated residues were identified in yeast U4 and U6 snRNAs under normal growth conditions.

Functions of Ψs and 2'-O-Methylated Residues Residing in U4 and U6 snRNAs

While the function of Ψs and 2'-O-methylated residues in yeast and mammalian U4, U5, and U6 snRNAs remains largely unclear, it is speculated that these modifications play a crucial role in splicing. Before participating in spliceosome assembly, U4, U5, and U6 snRNAs assemble into the U4/U6.U5 tri-snRNP particle, in which U4 and U6 snRNAs form an extensive base-pair interaction. The strength of this interaction was empirically determined as a stable one (Brow and Guthrie, 1988). Because of their presence in the base-paired region, Ψs and 2'-O-methylated residues seem to be particularly relevant (Figure 4). Given that Ψs and 2'-O-methylated residues are known to increase base-stacking and enhance base-pairing, it is possible that these modified nucleotides in the U4–U6 helices contribute to stabilizing the interaction. However, the base-pairing between U4 and U6 snRNAs must eventually unwind for the catalytically active spliceosome to form after the U4/U6.U5 tri-snRNP particle enters complex A (pre-mRNA complexed with U1 and U2 snRNPs; see Figure 2). This unwinding event is performed by Brr2, an ATP-dependent helicase with two helicase cassettes in tandem, although only the N-terminal one has unwinding activity (Raghuathan and Guthrie, 1998). Nguyen et al. (2015) obtained the cryo-EM structure of U4/U6.U5 tri-snRNP in yeast and observed that the Brr2 active site is preloaded in the single-stranded region between the stem I of the U4–U6 duplex region and the 3' stem-loop of the U4 snRNA (Figure 4). It is speculated that, since the human U4 snRNA has two Ψs in this single-stranded region (Ψ72 and Ψ79), they could be involved in the recruitment of (or recognition by) the helicase. However, upon deciphering the structure of human U4/U6.U5 tri-snRNP (also by cryo-EM), Agafonov et al. were able to find that Brr2 is located in a different position within the human U4/U6.U5 tri-snRNP complex, approximately 8–10 nm away from the U4/U6 snRNA duplex (Agafonov et al., 2016). To understand the function of Ψ72 and Ψ79 of human U4 snRNA, further research is necessary.

In the course of activation of the spliceosome, or after the unwinding of the U4/U6 snRNA duplex, U1 and U4 snRNAs leave the spliceosome, and the U2, U5, and U6 snRNAs interact with pre-mRNA and with each other (see Figure 2). In particular, U2 and U6 snRNAs form three short base-paired duplexes (helices I, II, and III; Datta and Weiner, 1991; Madhani and Guthrie, 1992; Sun and Manley, 1995; Burke et al., 2012), which are believed to be the catalytic center for splicing (transesterification) reactions. The dynamic formation of U2–U6 snRNA duplexes was also studied in a protein-free system, where U2–U6 snRNA interactions (likely related to the spliceosomal U2–U6 snRNA helices) were detected (Burke et al., 2012; Chu et al., 2020). Notably, there are multiple modified nucleotides (Ψ and 2'-O-methylated residues) in the U2–U6 snRNA duplexes. Using single-molecule fluorescence, Karunatilaka and Rueda further investigated the role of these modified

nucleotides in the dynamics of U2-U6 snRNA interactions. They concluded that the modifications present in the U2 snRNA stem I (**Figure 4**) contribute to the dynamics and conformation of the U2-U6 snRNA complex (Karunatilaka and Rueda, 2014). They also suggested that the modified nucleotides in the U2-U6 snRNA complex might also contribute to protein binding in addition to direct RNA structure stabilization.

Functions of Ψ s and 2'-O-Methylated Residues Residing in U5 snRNA

In the spliceosome, U5 snRNA interacts, through its conserved loop, with the pre-mRNA by directly contacting (non-Watson-Crick pairing) the exon nucleotides at both the 5' and 3' splice sites. Notably, there are two Ψ s and three 2'-O-methylated residues in this conserved loop sequence (GmCCUUmU Ψ ACm Ψ) of human U5 snRNA (Frank et al., 1994; Szkukalek et al., 1995). In *S. cerevisiae* U5, however, there is only one modified nucleotide (Ψ 99), but it is located in the same conserved loop sequence (GCCUUU Ψ AC; **Figure 4**). However, it should be noted that despite the conservation, there is not yet direct evidence indicating that these modified nucleotides contribute to the U5-pre-mRNA interactions.

Other Types of Modifications in U6 snRNA

It was recently reported that METTL16 functions as the methyltransferase responsible for the formation of m⁶A at position 43 within mammalian U6 snRNA (Pendleton et al., 2017; Warda et al., 2017; Aoyama et al., 2020). This modification (m⁶A43; Shimba et al., 1995) could have a direct role in splicing regulation given that it is conserved from *S. pombe* (at position 37; Gu et al., 1996) to human and that it is located in the region which forms base-pairing interactions with the 5' splice site of pre-mRNA before the first step of splicing reaction (transesterification reaction) occurs (**Figures 2, 4**). In this regard, it has been shown that mutations in this region (ACAGAGA), where m⁶A43 (underlined) is located, can be lethal in the yeast organism (Madhani et al., 1990). While these mutations might have directly disrupted the interaction between U6 snRNA and the 5' splice site of pre-mRNA leading to lethality, it is also possible that the mutations prevented the formation of m⁶A43, which is potentially important for 5' splice site recognition in *S. pombe* and higher eukaryotes. Further work is necessary to elucidate the role of m⁶A43. Mammalian U6 snRNA also contains the m²G at position 72 (Epstein et al., 1980). While the exact function of m²G72 in U6 snRNA is still largely unclear, this modified nucleotide is known to base-pair with C3 of U4 snRNA in the U4-U6 duplex structure (see **Figure 4**). Since this modification was determined empirically to have the same thermodynamic stability as unmodified guanosines in the context of a G:C base pair (Rife et al., 1998), one could speculate that its role in splicing could potentially be related to recognition by a splicing factor.

Inducible Ψ Formation in *Saccharomyces cerevisiae* U6 snRNA

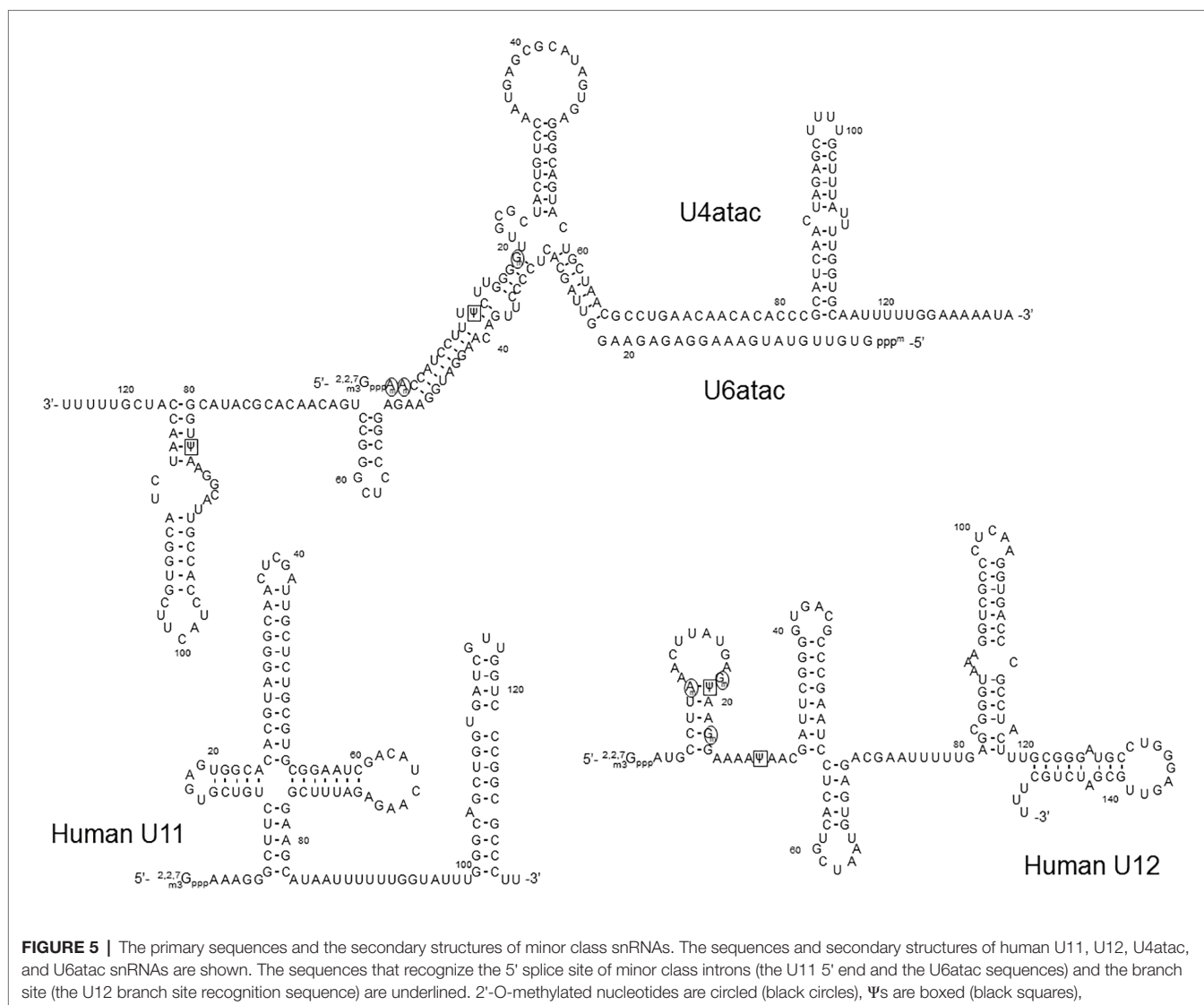
Another study was carried out in the Query lab, focusing on the inducible pseudouridylation of *S. cerevisiae* U6 snRNA at

position 28 (Ψ 28; **Figure 4**). They showed that the formation of Ψ 28 occurred under certain filamentous growth conditions. This growth condition-induced pseudouridylation is catalyzed by Pus1. Subsequent analyses allowed the authors to conclude that Ψ 28 in U6 snRNA directly contributes to filamentous formation (Basak and Query, 2014).

Minor Spliceosomal snRNAs

In addition to the major spliceosomal pathway (U2-dependent, described above), there is a less common (or minor) pathway required for splicing of rare class introns that contain different consensus sequences at the 5' and 3' splice sites and the branch site. Except for U5, which is common for both major and minor splicing pathways, a different set of spliceosomal snRNAs (U11, U12, U4atac, U6atac, and U5) is required for the minor splicing pathway (**Figure 5**). Because it depends on U12 (rather than U2), the minor splicing pathway is also known as the U12-dependent splicing pathway (Montzka and Steitz, 1988; Tarn and Steitz, 1996; Turunen et al., 2013). In the U12-dependent splicing pathway, U11, U12, U4atac, and U6atac snRNAs each play a role that is equivalent to the role of U1, U2, U4, and U6 snRNAs in the major spliceosome, respectively. Expectedly, the secondary structures of the minor class snRNAs are very similar to those of their major class snRNA counterparts (**Figure 5**). RNA-guided nucleotide modifications (pseudouridylation and 2'-O-methylation) in the minor-class snRNAs have also been studied (Jorjani et al., 2016).

In an attempt to map Ψ s in the minor spliceosome snRNAs in HeLa cells, Massenet and Branlant performed the pseudouridylation assay (CMC-modification followed by primer-extension) and identified Ψ s in U12, U4atac, and U6atac snRNAs (Massenet and Branlant, 1999). Surprisingly, no Ψ was detected in U11 snRNA, although its major spliceosome counterpart U1 snRNA has two Ψ s located in the 5'-end region that base-pairs with the 5'-splice site. It should be noted, however, that the 5' 10-nucleotide sequence (AAAAAGGGCU) of U11 snRNA (equivalent to the 5' 10-nucleotide sequence of U1 snRNA, AUAC Ψ Ψ ACCU) lacks the two U-residues (to be pseudouridylated) at the equivalent positions of U1 snRNA (Yu and Steitz, 1997). In the U12 snRNA, the authors detected only two Ψ s at positions 19 and 28. Ψ 19 is equivalent to Ψ 34 of mammalian U2 snRNA that base-pairs with the nucleotide (A or G) within the branch site sequence that is immediately next to the bulged-out nucleotide (branch point adenosine), pointing toward a functional role of this Ψ . Ψ 28 is located in a region that forms a base-pairing helix III with U6atac (equivalent to U2-U6 helix III), which has been shown to have a functional role in the splicing of U2-dependent introns (Sun and Manley, 1995). In the U4atac snRNA, a single Ψ was detected at position 12, located in the region that base-pairs with U6atac, equivalent to U4-U6 stem II in which there are several modified nucleotides in the U4 strand. Finally, Ψ 83 was also identified in the U6atac snRNA 3'-end region. This Ψ could be functionally similar to Ψ 86 in the U6 snRNA 3' terminal region. In a later study, these Ψ s were all confirmed (Deryusheva et al., 2012).



In addition, several 2'-O-methylations were also identified in U12 (at positions 8, 18, and 22) and U4atac snRNAs (at positions 1, 2, and 19, and potentially also position 8, although yet to be confirmed; Deryusheva et al., 2012). However, the exact function of these modified nucleotides remains unknown.

CONCLUDING REMARKS

Understanding the splicing mechanisms at the molecular level is of critical importance not only to fully comprehend gene expression but also to develop new nucleic acid-based therapeutics, such as splice-switching oligonucleotides (Lim and Yokota, 2018), aimed at correcting splicing-associated mutations that lead to aberrant proteins and diseases. pre-mRNA splicing occurs in the spliceosome, an extremely large complex consisting of five snRNAs and a large number of proteins that interact with substrate pre-mRNA in a highly orchestrated manner.

These snRNAs have a critical role in guiding the overall process *via* base-pairing interactions (and nucleotide-nucleotide contact) with the substrate pre-mRNA. Additionally, the snRNAs form dynamic structures that might be crucial for protein recruitment and catalysis. Post-transcriptionally modified nucleotides might contribute significantly in each of these steps during spliceosome assembly and splicing.

snRNA modifications, such as pseudouridine and 2'-O-methylation, have attracted a great deal of attention over the years, and extensive studies of these modifications have provided valuable insights into the mechanism of pre-mRNA splicing regulation. The continuous increase of knowledge of the fine-tuning and subtleties provided by RNA modifications in the spliceosome assembly and splicing processes are benefiting the development of better splicing modulation technologies. With the growing number of clinical trials based on splicing modulation therapies (exon-skipping or exon-inclusion) and the FDA-approved drugs based on this mechanism of action (Stein and Castanotto, 2017; Rüger et al., 2020), the interest

in this field will certainly continue to grow. The novel deep sequencing chemical probing technologies and epitranscriptomics analytical techniques will help us to decipher the yet-to-be discovered code of spliceosomal RNA modifications.

AUTHOR CONTRIBUTIONS

PM, HA, and Y-TY wrote the manuscript and generated the figures. All authors read and approved the final manuscript.

REFERENCES

- Adachi, H., and Yu, Y. -T. (2014). Insight into the mechanisms and functions of spliceosomal snRNA pseudouridylation. *World J. Biol. Chem.* 5, 398–408. doi: 10.4331/wjbc.v5.i4.398
- Agafonov, D. E., Kastner, B., Dybkov, O., Hofe, R. V., Liu, W. -T., Urlaub, H., et al. (2016). Molecular architecture of the human U4/U6.U5 tri-snRNP. *Science* 351, 1416–1420. doi: 10.1126/science.aad2085
- Aoyama, T., Yamashita, S., and Tomita, K. (2020). Mechanistic insights into m6A modification of U6 snRNA by human METTL16. *Nucleic Acids Res.* 48, 5157–5168. doi: 10.1093/nar/gkaa227
- Balakin, A. G., Smith, L., and Fournier, M. J. (1996). The RNA world of the nucleolus: two major families of small RNAs defined by different box elements with related functions. *Cell* 86, 823–834. doi: 10.1016/S0092-8674(00)80156-7
- Basak, A., and Query, C. C. (2014). A pseudouridine residue in the spliceosome core is part of the filamentous growth program in yeast. *Cell Rep.* 8, 966–973. doi: 10.1016/j.celrep.2014.07.004
- Berget, S. M., Moore, C., and Sharp, P. A. (1977). Spliced segments at the 5' terminus of adenovirus 2 late mRNA. *Proc. Natl. Acad. Sci. U. S. A.* 74, 3171–3175. doi: 10.1073/pnas.74.8.3171
- Boehringer, D., Makarov, E. M., Sander, B., Makarova, O. V., Kastner, B., Lührmann, R., et al. (2004). Three-dimensional structure of a pre-catalytic human spliceosomal complex B. *Nat. Struct. Mol. Biol.* 11, 463–468. doi: 10.1038/nsmb761
- Bohsack, M. T., and Sloan, K. E. (2018). Modifications in small nuclear RNAs and their roles in spliceosome assembly and function. *Biol. Chem.* 399, 1265–1276. doi: 10.1515/hsz-2018-0205
- Branlant, C., Krol, A., Ebel, J. P., Lazar, E., Gallinaro, H., and Jacob, M., et al. (1980). Nucleotide sequences of nuclear U1A RNAs from chicken, rat and man. *Nucleic Acids Res.* 8, 4143–4154. doi:10.1093/nar/8.18.4143
- Brow, D. A., and Guthrie, C. (1988). Spliceosomal RNA U6 is remarkably conserved from yeast to mammals. *Nature* 334, 213–218. doi: 10.1038/334213a0
- Burke, J. E., Sashital, D. G., Zuo, X., Wang, Y. -X., and Butcher, S. E. (2012). Structure of the yeast U2/U6 snRNA complex. *RNA* 18, 673–683. doi: 10.1261/rna.031138.111
- Caton, E. A., Kelly, E. K., Kamalampeta, R., and Kothe, U. (2018). Efficient RNA pseudouridylation by eukaryotic H/ACA ribonucleoproteins requires high affinity binding and correct positioning of guide RNA. *Nucleic Acids Res.* 46, 905–916. doi: 10.1093/nar/gkx1167
- Cavaillé, J., Nicoloso, M., and Bachellerie, J. P. (1996). Targeted ribose methylation of RNA in vivo directed by tailored antisense RNA guides. *Nature* 383, 732–735. doi: 10.1038/383732a0
- Chen, H., Gu, L., Orellana, E. A., Wang, Y., Guo, J., Liu, Q., et al. (2020). METTL4 is an snRNA m6Am methyltransferase that regulates RNA splicing. *Cell Res.* 30, 544–547. doi: 10.1038/s41422-019-0270-4
- Chow, L. T., Gelinas, R. E., Broker, T. R., and Roberts, R. J. (1977). An amazing sequence arrangement at the 5' ends of adenovirus 2 messenger RNA. *Cell* 12, 1–8. doi: 10.1016/0092-8674(77)90180-5
- Chu, H., Perea, W., and Greenbaum, N. L. (2020). Role of the central junction in folding topology of the protein-free human U2-U6 snRNA complex. *RNA* 26, 836–850. doi: 10.1261/rna.073379.119
- Darzacq, X., Jádý, B. E., Verheggen, C., Kiss, A. M., Bertrand, E., and Kiss, T. (2002). Cajal body-specific small nuclear RNAs: a novel class of 2'-O-methylation and pseudouridylation guide RNAs. *EMBO J.* 21, 2746–2756. doi: 10.1093/emboj/21.11.2746

FUNDING

The work performed in the Yu lab was supported by grants GM138387 and CA241111 from the US National Institutes of Health and grant CFF YU20G0 from the Cystic Fibrosis Foundation.

ACKNOWLEDGMENTS

We thank members of the Yu lab for insightful discussions.

- Datta, B., and Weiner, A. M. (1991). Genetic evidence for base pairing between U2 and U6 snRNA in mammalian mRNA splicing. *Nature* 352, 821–824. doi: 10.1038/352821a0
- Deryusheva, S., Choleza, M., Barbarossa, A., Gall, J. G., and Bordonne, R. (2012). Post-transcriptional modification of spliceosomal RNAs is normal in SMN-deficient cells. *RNA* 18, 31–36. doi: 10.1261/rna.030106.111
- Deryusheva, S., and Gall, J. G. (2017). Dual nature of pseudouridylation in U2 snRNA: Pus1p-dependent and Pus1p-independent activities in yeasts and higher eukaryotes. *RNA* 23, 1060–1067. doi: 10.1261/rna.061226.117
- Deryusheva, S., and Gall, J. G. (2018). Orchestrated positioning of post-transcriptional modifications at the branch point recognition region of U2 snRNA. *RNA* 24, 30–42. doi: 10.1261/rna.063842.117
- Deryusheva, S., Talhouarne, G. J. S., and Gall, J. G. (2020). “Lost and Found”: snoRNA annotation in the xenopus genome and implications for evolutionary studies. *Mol. Biol. Evol.* 37, 149–166. doi: 10.1093/molbev/msz209
- De Zoysa, M. D., Wu, G., Katz, R., and Yu, Y. -T. (2018). Guide-substrate base-pairing requirement for box H/ACA RNA-guided RNA pseudouridylation. *RNA* 24, 1106–1117. doi: 10.1261/rna.066837.118
- De Zoysa, M. D., and Yu, Y. -T. (2017). Posttranscriptional RNA pseudouridylation. *The Enzymes* 41, 151–167. doi: 10.1016/bs.enz.2017.02.001
- Dönmez, G., Hartmuth, K., and Lührmann, R. (2004). Modified nucleotides at the 5' end of human U2 snRNA are required for spliceosomal E-complex formation. *RNA* 10, 1925–1933. doi: 10.1261/rna.7186504
- Epstein, P., Reddy, R., Henning, D., and Busch, H. (1980). The nucleotide sequence of nuclear U6 (4.7 S) RNA. *J. Biol. Chem.* 255, 8901–8906. doi: 10.1016/S0021-9258(18)43587-9
- Fica, S. M., Tuttle, N., Novak, T., Li, N. -S., Lu, J., Koodathingal, P., et al. (2013). RNA catalyses nuclear pre-mRNA splicing. *Nature* 503, 229–234. doi: 10.1038/nature12734
- Frank, D. N., Roiha, H., and Guthrie, C. (1994). Architecture of the U5 small nuclear RNA. *Mol. Cell. Biol.* 14, 2180–2190. doi: 10.1128/MCB.14.3.2180
- Freund, M. (2003). A novel approach to describe a U1 snRNA binding site. *Nucleic Acids Res.* 31, 6963–6975. doi: 10.1093/nar/gkg901
- Frye, M., Harada, B. T., Behm, M., and He, C. (2018). RNA modifications modulate gene expression during development. *Science* 361, 1346–1349. doi: 10.1126/science.aau1646
- Ganot, P., Bortolin, M. L., and Kiss, T. (1997). Site-specific pseudouridine formation in preribosomal RNA is guided by small nucleolar RNAs. *Cell* 89, 799–809. doi: 10.1016/S0092-8674(00)80263-9
- Ganot, P., Jádý, B. E., Bortolin, M. L., Darzacq, X., and Kiss, T. (1999). Nucleolar factors direct the 2'-O-ribose methylation and pseudouridylation of U6 spliceosomal RNA. *Mol. Cell. Biol.* 19, 6906–6917. doi: 10.1128/MCB.19.10.6906
- Ge, J., and Yu, Y. -T. (2013). RNA pseudouridylation: new insights into an old modification. *Trends Biochem. Sci.* 38, 210–218. doi: 10.1016/j.tibs.2013.01.002
- Goh, Y. T., Koh, C. W. Q., Sim, D. Y., Roca, X., and Goh, W. S. S. (2020). METTL4 catalyzes m6Am methylation in U2 snRNA to regulate pre-mRNA splicing. *Nucleic Acids Res.* 48, 9250–9261. doi: 10.1093/nar/gkaa684
- Grosjean, H., Sprinzl, M., and Steinberg, S. (1995). Posttranscriptionally modified nucleosides in transfer RNA: their locations and frequencies. *Biochimie* 77, 139–141. doi: 10.1016/0300-9084(96)88117-X
- Gu, J., Patton, J. R., Shimba, S., and Reddy, R. (1996). Localization of modified nucleotides in Schizosaccharomyces pombe spliceosomal small nuclear RNAs: modified nucleotides are clustered in functionally important regions. *RNA* 2, 909–918.

- Gu, A. -D., Zhou, H., Yu, C. -H., and Qu, L. -H. (2005). A novel experimental approach for systematic identification of box H/ACA snoRNAs from eukaryotes. *Nucleic Acids Res.* 33:e194. doi: 10.1093/nar/gni185
- Hirose, T., Ideue, T., Nagai, M., Hagiwara, M., Shu, M. -D., and Steitz, J. A. (2006). A spliceosomal intron binding protein, IBP160, links position-dependent assembly of intron-encoded box C/D snoRNP to pre-mRNA splicing. *Mol. Cell* 23, 673–684. doi: 10.1016/j.molcel.2006.07.011
- Hirose, T., Shu, M. -D., and Steitz, J. A. (2003). Splicing-dependent and -independent modes of assembly for intron-encoded box C/D snoRNPs in mammalian cells. *Mol. Cell* 12, 113–123. doi: 10.1016/S1097-2765(03)00267-3
- Hüttenhofer, A., Kieffmann, M., Meier-Ewert, S., O'Brien, J., Lehrach, H., and Bacherle, J. P., et al. (2001). RNomics: an experimental approach that identifies 201 candidates for novel, small, non-messenger RNAs in mouse. *EMBO J.* 20, 2943–2953. doi: 10.1093/emboj/20.11.2943
- Jády, B. E., and Kiss, T. (2001). A small nucleolar guide RNA functions both in 2'-O-ribose methylation and pseudouridylation of the U5 spliceosomal RNA. *EMBO J.* 20, 541–551. doi: 10.1093/emboj/20.3.541
- Jorjani, H., Kehr, S., Jedlinski, D. J., Gumienny, R., Hertel, J., Stadler, P. F., et al. (2016). An updated human snoRNAome. *Nucleic Acids Res.* 44, 5068–5082. doi: 10.1093/nar/gkw386
- Jurica, M. S., and Moore, M. J. (2003). Pre-mRNA splicing: awash in a sea of proteins. *Mol. Cell* 12, 5–14. doi: 10.1016/S1097-2765(03)00270-3
- Jurica, M. S., Sousa, D., Moore, M. J., and Grigorieff, N. (2004). Three-dimensional structure of C complex spliceosomes by electron microscopy. *Nat. Struct. Mol. Biol.* 11, 265–269. doi: 10.1038/nsmb728
- Karijoh, J., and Yu, Y. -T. (2010). Spliceosomal snRNA modifications and their function. *RNA Biol.* 7, 192–204. doi: 10.4161/rna.7.2.11207
- Karunatilaka, K. S., and Rueda, D. (2014). Post-transcriptional modifications modulate conformational dynamics in human U2-U6 snRNA complex. *RNA* 20, 16–23. doi: 10.1261/rna.041806.113
- Kennedy, S. D., Bauer, W. J., Wang, W., and Kielkopf, C. L. (2019). Dynamic stacking of an expected branch point adenosine in duplexes containing pseudouridine-modified or unmodified U2 snRNA sites. *Biochem. Biophys. Res. Commun.* 511, 416–421. doi: 10.1016/j.bbrc.2019.02.073
- Kiss, A. M., Jády, B. E., Bertrand, E., and Kiss, T. (2004). Human box H/ACA pseudouridylation guide RNA machinery. *Mol. Cell. Biol.* 24, 5797–5807. doi: 10.1128/MCB.24.13.5797-5807.2004
- Kiss, A. M., Jády, B. E., Darzacq, X., Verheggen, C., Bertrand, E., and Kiss, T. (2002). A Cajal body-specific pseudouridylation guide RNA is composed of two box H/ACA snoRNA-like domains. *Nucleic Acids Res.* 30, 4643–4649. doi: 10.1093/nar/gkf592
- Kiss-László, Z., Henry, Y., Bacherle, J. P., Caizergues-Ferrer, M., and Kiss, T. (1996). Site-specific ribose methylation of preribosomal RNA: a novel function for small nucleolar RNAs. *Cell* 85, 1077–1088. doi: 10.1016/S0092-8674(00)81308-2
- Kondo, Y., Oubridge, C., van Roon, A. -M. M., and Nagai, K. (2015). Crystal structure of human U1 snRNP, a small nuclear ribonucleoprotein particle, reveals the mechanism of 5' splice site recognition. *eLife* 4:e04986. doi: 10.7554/eLife.04986
- Krawczak, M., Thomas, N. S. T., Hundrieser, B., Mort, M., Wittig, M., Hampe, J., et al. (2007). Single base-pair substitutions in exon-intron junctions of human genes: nature, distribution, and consequences for mRNA splicing. *Hum. Mutat.* 28, 150–158. doi: 10.1002/humu.20400
- Krogh, N., Kongsbak-Wismann, M., Geisler, C., and Nielsen, H. (2017). Substoichiometric ribose methylations in spliceosomal snRNAs. *Org. Biomol. Chem.* 15, 8872–8876. doi: 10.1039/C7OB02317K
- Krol, A., Gallinaro, H., Lazar, E., Jacob, M., and Branlant, C. (1981). The nuclear 5S RNAs from chicken, rat and man. U5 RNAs are encoded by multiple genes. *Nucleic Acids Res.* 9, 769–787. doi: 10.1093/nar/9.4.769
- Lestrade, L., and Weber, M. J. (2006). snoRNA-LBME-db, a comprehensive database of human H/ACA and C/D box snoRNAs. *Nucleic Acids Res.* 34, D158–D162. doi: 10.1093/nar/gkj002
- Lim, K. R. Q., and Yokota, T. (2018). Invention and early history of exon skipping and splice modulation. *Methods Mol. Biol.* 1828, 3–30. doi: 10.1007/978-1-4939-8651-4_1
- Ma, X., Yang, C., Alexandrov, A., Grayhack, E. J., Behm-Ansmant, I., and Yu, Y. -T. (2005). Pseudouridylation of yeast U2 snRNA is catalyzed by either an RNA-guided or RNA-independent mechanism. *EMBO J.* 24, 2403–2413. doi: 10.1038/sj.emboj.7600718
- Ma, X., Zhao, X., and Yu, Y. -T. (2003). Pseudouridylation (psi) of U2 snRNA in *S. cerevisiae* is catalyzed by an RNA-independent mechanism. *EMBO J.* 22, 1889–1897. doi: 10.1093/emboj/cdg191
- Madhani, H. D., Bordonné, R., and Guthrie, C. (1990). Multiple roles for U6 snRNA in the splicing pathway. *Genes Dev.* 4, 2264–2277. doi: 10.1101/gad.4.12b.2264
- Madhani, H. D., and Guthrie, C. (1992). A novel base-pairing interaction between U2 and U6 snRNAs suggests a mechanism for the catalytic activation of the spliceosome. *Cell* 71, 803–817. doi: 10.1016/0092-8674(92)90556-R
- Massenet, S., and Branlant, C. (1999). A limited number of pseudouridine residues in the human atac spliceosomal UsnRNAs as compared to human major spliceosomal UsnRNAs. *RNA* 5, 1495–1503. doi: 10.1017/S1355838299991537
- Massenet, S., Motorin, Y., Lafontaine, D. L. J., Hurt, E. C., Grosjean, H., and Branlant, C. (1999). Pseudouridine mapping in the *Saccharomyces cerevisiae* spliceosomal U small nuclear RNAs (snRNAs) reveals that pseudouridine synthase Pus1p exhibits a dual substrate specificity for U2 snRNA and tRNA. *Mol. Cell. Biol.* 19, 2142–2154. doi: 10.1128/MCB.19.3.2142
- Massenet, S., Mougin, A., and Branlant, C. (1998). "Posttranscriptional modifications in the U small nuclear RNAs" in *Modification and Editing of RNA*. ed. H. Grosjean (Washington, DC: ASM Press), 201–228.
- Mauer, J., Sindelar, M., Despic, V., Guez, T., Hawley, B. R., Vasseur, J. -J., et al. (2019). FTO controls reversible m6Am RNA methylation during snRNA biogenesis. *Nat. Chem. Biol.* 15, 340–347. doi: 10.1038/s41589-019-0231-8
- Meier, U. T. (2017). RNA modification in Cajal bodies. *RNA Biol.* 14, 693–700. doi: 10.1080/15476286.2016.1249091
- Montzka, K. A., and Steitz, J. A. (1988). Additional low-abundance human small nuclear ribonucleoproteins: U11, U12, etc. *Proc. Natl. Acad. Sci.* 85, 8885–8889. doi: 10.1073/pnas.85.23.8885
- Motorin, Y., Keith, G., Simon, C., Foiret, D., Simos, G., Hurt, E., et al. (1998). The yeast tRNA:pseudouridine synthase Pus1p displays a multisite substrate specificity. *RNA* 4, 856–869. doi: 10.1017/S1355838298980396
- Nachtergaele, S., and He, C. (2017). The emerging biology of RNA post-transcriptional modifications. *RNA Biol.* 14, 156–163. doi: 10.1080/15476286.2016.1267096
- Newby, M. I., and Greenbaum, N. L. (2001). A conserved pseudouridine modification in eukaryotic U2 snRNA induces a change in branch-site architecture. *RNA N. Y. N* 7, 833–845. doi: 10.1017/S1355838201002308
- Newby, M. I., and Greenbaum, N. L. (2002). Sculpting of the spliceosomal branch site recognition motif by a conserved pseudouridine. *Nat. Struct. Biol.* 9, 958–965. doi: 10.1038/nsb873
- Newman, A. J. (1997). The role of U5 snRNP in pre-mRNA splicing. *EMBO J.* 16, 5797–5800. doi: 10.1093/emboj/16.19.5797
- Newman, A. J., and Norman, C. (1992). U5 snRNA interacts with exon sequences at 5' and 3' splice sites. *Cell* 68, 743–754. doi: 10.1016/0092-8674(92)90149-7
- Newman, A. J., Teigelkamp, S., and Beggs, J. D. (1995). snRNA interactions at 5' and 3' splice sites monitored by photoactivated crosslinking in yeast spliceosomes. *RNA* 1, 968–980.
- Nguyen, T. H. D., Galej, W. P., Bai, X. -C., Oubridge, C., Newman, A. J., Scheres, S. H. W., et al. (2016). Cryo-EM structure of the yeast U4/U6.U5 tri-snRNP at 3.7 Å resolution. *Nature* 530, 298–302. doi: 10.1038/nature16940
- Nguyen, T. H. D., Galej, W. P., Bai, X., Savva, C. G., Newman, A. J., Scheres, S. H. W., et al. (2015). The architecture of the spliceosomal U4/U6.U5 tri-snRNP. *Nature* 523, 47–52. doi: 10.1038/nature14548
- Ni, J., Samarsky, D. A., Liu, B., Ferbeyre, G., Cedergren, R., and Fournier, M. J. (1997). SnoRNAs as tools for RNA cleavage and modification. *Nucleic Acids Symp. Ser.* 36, 61–63.
- Nilsen, T. W. (1994). RNA-RNA interactions in the spliceosome: unraveling the ties that bind. *Cell* 78, 1–4. doi: 10.1016/0092-8674(94)90563-0
- Nilsen, T. W. (2003). The spliceosome: the most complex macromolecular machine in the cell? *BioEssays* 25, 1147–1149. doi: 10.1002/bies.10394
- Parker, R., Siliciano, P. G., and Guthrie, C. (1987). Recognition of the TACTAAC box during mRNA splicing in yeast involves base pairing to the U2-like snRNA. *Cell* 49, 229–239. doi: 10.1016/0092-8674(87)90564-2
- Pendleton, K. E., Chen, B., Liu, K., Hunter, O. V., Xie, Y., and Tu, B. P., et al. (2017). The U6 snRNA m6a methyltransferase METTL16 regulates SAM synthetase intron retention. *Cell* 169, 824.e14–835.e14. doi: 10.1016/j.cell.2017.05.003
- Plaschka, C., Lin, P. -C., and Nagai, K. (2017). Structure of a pre-catalytic spliceosome. *Nature* 546, 617–621. doi: 10.1038/nature22799
- Query, C. C., McCaw, P. S., and Sharp, P. A. (1997). A minimal spliceosomal complex recognizes the branch site and polypyrimidine tract. *Mol. Cell. Biol.* 17, 2944–2953. doi: 10.1128/MCB.17.5.2944

- Raghuathan, P. L., and Guthrie, C. (1998). RNA unwinding in U4/U6 snRNPs requires ATP hydrolysis and the DEIH-box splicing factor Brr2. *Curr. Biol.* 8, 847–855. doi: 10.1016/S0960-9822(07)00345-4
- Reddy, R., and Busch, H. (1988). “Small nuclear RNAs: RNA sequences, structure, and modifications” in *Structure and function of major and minor small nuclear ribonucleoprotein particles*. ed. M. L. Birnstiel (Heidelberg: Springer-Verlag Press), 1–37.
- Reddy, R., Henning, D., and Busch, H. (1981a). Pseudouridine residues in the 5'-terminus of uridine-rich nuclear RNA I (U1 RNA). *Biochem. Biophys. Res. Commun.* 98, 1076–1083. doi: 10.1016/0006-291x(81)91221-3
- Reddy, R., Henning, D., and Busch, H. (1981b). The primary nucleotide sequence of U4 RNA. *J. Biol. Chem.* 256, 3532–3538.
- Rife, J. P., Cheng, C. S., Moore, P. B., and Strobel, S. A. (1998). N2-Methylguanosine is iso-energetic with guanosine in RNA duplexes and GNRA tetraloops. *Nucleic Acids Res.* 26, 3640–3644. doi: 10.1093/nar/26.16.3640
- Rintala-Dempsey, A. C., and Kothe, U. (2017). Eukaryotic stand-alone pseudouridine synthases – RNA modifying enzymes and emerging regulators of gene expression? *RNA Biol.* 14, 1185–1196. doi: 10.1080/15476286.2016.1276150
- Roca, X. (2005). Determinants of the inherent strength of human 5' splice sites. *RNA* 11, 683–698. doi: 10.1261/rna.2040605
- Roca, X., Akerman, M., Gaus, H., Berdeja, A., Bennett, C. F., and Krainer, A. R. (2012). Widespread recognition of 5' splice sites by noncanonical base-pairing to U1 snRNA involving bulged nucleotides. *Genes Dev.* 26, 1098–1109. doi: 10.1101/gad.190173.112
- Roundtree, I. A., Evans, M. E., Pan, T., and He, C. (2017). Dynamic RNA modifications in gene expression regulation. *Cell* 169, 1187–1200. doi: 10.1016/j.cell.2017.05.045
- Rüger, J., Ioannou, S., Castanotto, D., and Stein, C. A. (2020). Oligonucleotides to the (gene) rescue: FDA approvals 2017–2019. *Trends Pharmacol. Sci.* 41, 27–41. doi: 10.1016/j.tips.2019.10.009
- Ruskin, B., Krainer, A. R., Maniatis, T., and Green, M. R. (1984). Excision of an intact intron as a novel lariat structure during pre-mRNA splicing in vitro. *Cell* 38, 317–331. doi: 10.1016/0092-8674(84)90553-1
- Ruskin, B., Zamore, P. D., and Green, M. R. (1988). A factor, U2AF, is required for U2 snRNP binding and splicing complex assembly. *Cell* 52, 207–219. doi: 10.1016/0092-8674(88)90509-0
- Schattner, P., Barberan-Soler, S., and Lowe, T. M. (2006). A computational screen for mammalian pseudouridylation guide H/ACA RNAs. *RNA* 12, 15–25. doi: 10.1261/rna.2210406
- Schwartz, S., Bernstein, D. A., Mumbach, M. R., Jovanovic, M., Herbst, R. H., León-Ricardo, B. X., et al. (2014). Transcriptome-wide mapping reveals widespread dynamic-regulated pseudouridylation of ncRNA and mRNA. *Cell* 159, 148–162. doi: 10.1016/j.cell.2014.08.028
- Scotti, M. M., and Swanson, M. S. (2016). RNA mis-splicing in disease. *Nat. Rev. Genet.* 17, 19–32. doi: 10.1038/nrg.2015.3
- Sergiev, P. V., Lesnyak, D. V., Bogdanov, A. A., and Dontsova, O. A. (2006). Identification of *Escherichia coli* m2G methyltransferases: II. The ygiO gene encodes a methyltransferase specific for G1835 of the 23 S rRNA. *J. Mol. Biol.* 364, 26–31. doi: 10.1016/j.jmb.2006.09.008
- Shi, Y. (2017). Mechanistic insights into precursor messenger RNA splicing by the spliceosome. *Nat. Rev. Mol. Cell Biol.* 18, 655–670. doi: 10.1038/nrm.2017.86
- Shibata, H., Ro-Choi, T. S., Reddy, R., Choi, Y. C., Henning, D., and Busch, H. (1975). The primary nucleotide sequence of nuclear U-2 ribonucleic acid. The 5'-terminal portion of the molecule. *J. Biol. Chem.* 250, 3909–3920. doi: 10.1016/S0021-9258(19)41485-3
- Shimba, S., Bokar, J. A., Rottman, F., and Reddy, R. (1995). Accurate and efficient N-6-adenosine methylation in spliceosomal U6 small nuclear RNA by HeLa cell extract in vitro. *Nucleic Acids Res.* 23, 2421–2426. doi: 10.1093/nar/23.13.2421
- Simos, G., Tekotte, H., Grosjean, H., Segref, A., Sharma, K., Tollervey, D., et al. (1996). Nuclear pore proteins are involved in the biogenesis of functional tRNA. *EMBO J.* 15, 2270–2284. doi: 10.1002/j.1460-2075.1996.tb00580.x
- Sindhuphak, T., Hellman, U., and Svensson, I. (1985). Site specificities of three transfer RNA methyltransferases from yeast. *Biochim. Biophys. Acta* 824, 66–73. doi: 10.1016/0167-4781(85)90030-2
- Singh, R., and Reddy, R. (1989). Gamma-monomethyl phosphate: a cap structure in spliceosomal U6 small nuclear RNA. *Proc. Natl. Acad. Sci. U. S. A.* 86, 8280–8283. doi: 10.1073/pnas.86.21.8280
- Sontheimer, E. J., and Steitz, J. A. (1993). The U5 and U6 small nuclear RNAs as active site components of the spliceosome. *Science* 262, 1989–1996. doi: 10.1126/science.8266094
- Staley, J. P., and Guthrie, C. (1998). Mechanical devices of the spliceosome: motors, clocks, springs, and things. *Cell* 92, 315–326. doi: 10.1016/S0092-8674(00)80925-3
- Stein, C. A., and Castanotto, D. (2017). FDA-approved oligonucleotide therapies in 2017. *Mol. Ther. J. Am. Soc. Gene Ther.* 25, 1069–1075. doi: 10.1016/j.ymthe.2017.03.023
- Sun, J. S., and Manley, J. L. (1995). A novel U2-U6 snRNA structure is necessary for mammalian mRNA splicing. *Genes Dev.* 9, 843–854. doi: 10.1101/gad.9.7.843
- Szukalek, A., Myslinski, E., Mouglin, A., Luhrmann, R., and Branlant, C. (1995). Phylogenetic conservation of modified nucleotides in the terminal loop 1 of the spliceosomal U5 snRNA. *Biochimie* 77, 16–21. doi: 10.1016/0300-9084(96)88099-0
- Tarn, W. Y., and Steitz, J. A. (1996). A novel spliceosome containing U11, U12, and U5 snRNPs excises a minor class (AT-AC) intron in vitro. *Cell* 84, 801–811. doi: 10.1016/S0092-8674(00)81057-0
- Tarn, W. Y., and Steitz, J. A. (1997). Pre-mRNA splicing: the discovery of a new spliceosome doubles the challenge. *Trends Biochem. Sci.* 22, 132–137. doi: 10.1016/S0968-0004(97)01018-9
- Terns, M., and Terns, R. (2006). Noncoding RNAs of the H/ACA family. *Cold Spring Harb. Symp. Quant. Biol.* 71, 395–405. doi: 10.1101/sqb.2006.71.034
- Turunen, J. J., Niemelä, E. H., Verma, B., and Frilander, M. J. (2013). The significant other: splicing by the minor spliceosome. *Wiley Interdiscip. Rev. RNA* 4, 61–76. doi: 10.1002/wrna.1141
- Tycowski, K. T., Aab, A., and Steitz, J. A. (2004). Guide RNAs with 5' caps and novel box C/D snoRNA-like domains for modification of snRNAs in metazoa. *Curr. Biol. CB* 14, 1985–1995. doi: 10.1016/j.cub.2004.11.003
- Tycowski, K. T., Shu, M. D., and Steitz, J. A. (1996). A mammalian gene with introns instead of exons generating stable RNA products. *Nature* 379, 464–466. doi: 10.1038/379464a0
- Tycowski, K. T., You, Z. H., Graham, P. J., and Steitz, J. A. (1998). Modification of U6 spliceosomal RNA is guided by other small RNAs. *Mol. Cell* 2, 629–638. doi: 10.1016/S1097-2765(00)80161-6
- Urban, A., Behm-Ansmant, I., Branlant, C., and Motorin, Y. (2009). RNA sequence and two-dimensional structure features required for efficient substrate modification by the *Saccharomyces cerevisiae* RNA:[psi]-synthase Pus7p. *J. Biol. Chem.* 284, 5845–5858. doi: 10.1074/jbc.M807986200
- van der Feltz, C., DeHaven, A. C., and Hoskins, A. A. (2018). Stress-induced pseudouridylation alters the structural equilibrium of yeast U2 snRNA stem II. *J. Mol. Biol.* 430, 524–536. doi: 10.1016/j.jmb.2017.10.021
- Vitali, P., Royo, H., Seitz, H., Bachellerie, J. -P., Hüttenhofer, A., and Cavallé, J. (2003). Identification of 13 novel human modification guide RNAs. *Nucleic Acids Res.* 31, 6543–6551. doi: 10.1093/nar/gkg849
- Wahl, M. C., Will, C. L., and Luhrmann, R. (2009). The spliceosome: design principles of a dynamic RNP machine. *Cell* 136, 701–718. doi: 10.1016/j.cell.2009.02.009
- Wan, R., Yan, C., Bai, R., Wang, L., Huang, M., Wong, C. C. L., et al. (2016). The 3.8 Å structure of the U4/U6.U5 tri-snRNP: insights into spliceosome assembly and catalysis. *Science* 351, 466–475. doi: 10.1126/science.1246466
- Warda, A. S., Kretschmer, J., Hackert, P., Lenz, C., Urlaub, H., and Höbartner, C., et al. (2017). Human METTL16 is a N6-methyladenosine (m6A) methyltransferase that targets pre-mRNAs and various non-coding RNAs. *EMBO Rep.* 18, 2004–2014. doi:10.15252/embr.201744940
- Wassarman, D. A., and Steitz, J. A. (1992). Interactions of small nuclear RNAs with precursor messenger RNA during in vitro splicing. *Science* 257, 1918–1925. doi: 10.1126/science.1411506
- Wiener, D., and Schwartz, S. (2020). The epitranscriptome beyond m⁶A. *Nat. Rev. Genet.* 22, 119–131. doi: 10.1038/s41576-020-00295-8
- Will, C. L., and Luhrmann, R. (2005). Splicing of a rare class of introns by the U12-dependent spliceosome. *Biol. Chem.* 386, 713–724. doi: 10.1515/BC.2005.084
- Will, C. L., and Luhrmann, R. (2011). Spliceosome structure and function. *Cold Spring Harb. Perspect. Biol.* 3:a003707. doi: 10.1101/cshperspect.a003707
- Will, C. L., Rümpler, S., and Klein Gunnewiek, J., van Venrooij, W. J., and Luhrmann, R. (1996). In vitro reconstitution of mammalian U1 snRNPs active in splicing: the U1-C protein enhances the formation of early (E)

- spliceosomal complexes. *Nucleic Acids Res.* 24, 4614–4623. doi:10.1093/nar/24.23.4614.
- Wu, G., Adachi, H., Ge, J., Stephenson, D., Query, C. C., and Yu, Y. (2016a). Pseudouridines in U2 snRNA stimulate the ATPase activity of Prp5 during spliceosome assembly. *EMBO J.* 35, 654–667. doi: 10.15252/embj.201593113
- Wu, J. A., and Manley, J. L. (1991). Base pairing between U2 and U6 snRNAs is necessary for splicing of a mammalian pre-mRNA. *Nature* 352, 818–821. doi: 10.1038/352818a0
- Wu, G., Radwan, M. K., Xiao, M., Adachi, H., Fan, J., and Yu, Y. -T. (2016b). The TOR signaling pathway regulates starvation-induced pseudouridylation of yeast U2 snRNA. *RNA* 22, 1146–1152. doi: 10.1261/rna.056796.116
- Wu, S., Romfo, C. M., Nilsen, T. W., and Green, M. R. (1999). Functional recognition of the 3' splice site AG by the splicing factor U2AF35. *Nature* 402, 832–835. doi: 10.1038/45590
- Wu, G., Xiao, M., Yang, C., and Yu, Y. -T. (2011). U2 snRNA is inducibly pseudouridylated at novel sites by Pus7p and snR81 RNP. *EMBO J.* 30, 79–89. doi: 10.1038/emboj.2010.316
- Wyatt, J. R., Sontheimer, E. J., and Steitz, J. A. (1992). Site-specific cross-linking of mammalian U5 snRNP to the 5' splice site before the first step of pre-mRNA splicing. *Genes Dev.* 6, 2542–2553. doi: 10.1101/gad.6.12b.2542
- Yan, C., Wan, R., Bai, R., Huang, G., and Shi, Y. (2017). Structure of a yeast step II catalytically activated spliceosome. *Science* 355, 149–155. doi: 10.1126/science.aak9979
- Yang, C., McPheeters, D. S., and Yu, Y. -T. (2005). ψ 35 in the branch site recognition region of U2 small nuclear RNA is important for pre-mRNA splicing in *Saccharomyces cerevisiae*. *J. Biol. Chem.* 280, 6655–6662. doi: 10.1074/jbc.M413288200
- Yean, S. L., Wuenschell, G., Termini, J., and Lin, R. J. (2000). Metal-ion coordination by U6 small nuclear RNA contributes to catalysis in the spliceosome. *Nature* 408, 881–884. doi: 10.1038/35048617
- Yu, A. T., Ge, J., and Yu, Y. -T. (2011). Pseudouridines in spliceosomal snRNAs. *Protein Cell* 2, 712–725. doi: 10.1007/s13238-011-1087-1
- Yu, Y. T., Shu, M. D., and Steitz, J. A. (1998). Modifications of U2 snRNA are required for snRNP assembly and pre-mRNA splicing. *EMBO J.* 17, 5783–5795. doi: 10.1093/emboj/17.19.5783
- Yu, Y. T., and Steitz, J. A. (1997). Site-specific crosslinking of mammalian U11 and u6atac to the 5' splice site of an AT-AC intron. *Proc. Natl. Acad. Sci. U. S. A.* 94, 6030–6035. doi: 10.1073/pnas.94.12.6030
- Yu, Y. -T., Terns, R. M., and Terns, M. P. (2005). “Mechanisms and functions of RNA-guided RNA modification” in *Fine-tuning of RNA functions by modification and editing*. ed. H. Grosjean (New York: Springer-Verlag Press), 223–262.
- Zamore, P. D., Patton, J. G., and Green, M. R. (1992). Cloning and domain structure of the mammalian splicing factor U2AF. *Nature* 355, 609–614. doi: 10.1038/355609a0
- Zerby, D. B., and Patton, J. R. (1997). Modification of human U4 RNA requires U6 RNA and multiple pseudouridine synthases. *Nucleic Acids Res.* 25, 4808–4815. doi: 10.1093/nar/25.23.4808
- Zhang, X., Yan, C., Zhan, X., Li, L., Lei, J., and Shi, Y. (2018). Structure of the human activated spliceosome in three conformational states. *Cell Res.* 28, 307–322. doi: 10.1038/cr.2018.14
- Zhang, M., Zamore, P. D., Carmo-Fonseca, M., Lamond, A. I., and Green, M. R. (1992). Cloning and intracellular localization of the U2 small nuclear ribonucleoprotein auxiliary factor small subunit. *Proc. Natl. Acad. Sci. U. S. A.* 89, 8769–8773. doi: 10.1073/pnas.89.18.8769
- Zhao, X., and Yu, Y. -T. (2004). Detection and quantitation of RNA base modifications. *RNA* 10, 996–1002. doi: 10.1261/rna.7110804
- Zhuang, Y. A., Goldstein, A. M., and Weiner, A. M. (1989). UACUAAC is the preferred branch site for mammalian mRNA splicing. *Proc. Natl. Acad. Sci. U. S. A.* 86, 2752–2756. doi: 10.1073/pnas.86.8.2752
- Zhuang, Y., and Weiner, A. M. (1986). A compensatory base change in U1 snRNA suppresses a 5' splice site mutation. *Cell* 46, 827–835. doi: 10.1016/0092-8674(86)90064-4

Conflict of Interest: PM is a scientific director at ProQR Therapeutics.

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

Copyright © 2021 Morais, Adachi and Yu. 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.



HN-CNN: A Heterogeneous Network Based on Convolutional Neural Network for m⁷G Site Disease Association Prediction

Lin Zhang^{1,2*}, Jin Chen², Jiani Ma² and Hui Liu^{1,2*}

¹ Engineering Research Center of Intelligent Control for Underground Space, Ministry of Education, China University of Mining and Technology, Xuzhou, China, ² School of Information and Control Engineering, China University of Mining and Technology, Xuzhou, China

OPEN ACCESS

Edited by:

Giovanni Nigita,
The Ohio State University,
United States

Reviewed by:

Zhang Shaowu,
Northwestern Polytechnical University,
China

Quan Zou,

University of Electronic Science
and Technology of China, China

Wei Chen,

North China University of Science
and Technology, China

*Correspondence:

Hui Liu
hui.liu@cumt.edu.cn
Lin Zhang
lin.zhang@cumt.edu.cn

Specialty section:

This article was submitted to
Epigenomics and Epigenetics,
a section of the journal
Frontiers in Genetics

Received: 18 January 2021

Accepted: 15 February 2021

Published: 04 March 2021

Citation:

Zhang L, Chen J, Ma J and Liu H
(2021) HN-CNN: A Heterogeneous
Network Based on Convolutional
Neural Network for m⁷G Site Disease
Association Prediction.
Front. Genet. 12:655284.
doi: 10.3389/fgene.2021.655284

N⁷-methylguanosine (m⁷G) is a typical positively charged RNA modification, playing a vital role in transcriptional regulation. m⁷G can affect the biological processes of mRNA and tRNA and has associations with multiple diseases including cancers. Wet-lab experiments are cost and time ineffective for the identification of disease-related m⁷G sites. Thus, a heterogeneous network method based on Convolutional Neural Networks (HN-CNN) has been proposed to predict unknown associations between m⁷G sites and diseases. HN-CNN constructs a heterogeneous network with m⁷G site similarity, disease similarity, and disease-associated m⁷G sites to formulate features for m⁷G site-disease pairs. Next, a convolutional neural network (CNN) obtains multidimensional and irrelevant features prominently. Finally, XGBoost is adopted to predict the association between m⁷G sites and diseases. The performance of HN-CNN is compared with Naive Bayes (NB), Random Forest (RF), Support Vector Machine (SVM), as well as Gradient Boosting Decision Tree (GBDT) through 10-fold cross-validation. The average AUC of HN-CNN is 0.827, which is superior to others.

Keywords: m⁷G sites, diseases, heterogeneous network, convolutional neural network, XGBoost

INTRODUCTION

N⁷-methylguanosine (m⁷G) is one of the most abundant modifications present in tRNA, rRNA, and mRNA 5' cap and plays critical roles in regulating RNA processing, metabolism, and function (Malbec et al., 2019). As an essential post-transcriptional modification, m⁷G plays an essential role in gene expression, processing and metabolism, protein synthesis, transcription stability and other aspects (Pandolfini et al., 2019). m⁷G is often enriched in the 5'UTR region and AG-enriched contexts. The internal m⁷G modification is dynamically regulated under both H₂O₂ and heat shock treatments, with remarkable accumulations in CDS and 3'UTR regions and functions in promoting mRNA translation efficiency (Malbec et al., 2019). m⁷G₄₆ methylation of specific tRNA is associated with human mutation and the corresponding yeast mutation, which is m⁷G modification at position 46 in tRNA. Reduced m⁷G₄₆ modification causes a growth deficiency phenotype in yeast, which provides a potential mechanism for primordial dwarfism associated with this lesion (Shaheen et al., 2015).

Munns et al. (1985) concluded that a specific autoimmune disorder is associated with the presence of anti-m⁷G autoantibodies in 50 patients' cases. Bradrick (2017) found that mosquito-borne flaviviruses are important human pathogens, and m⁷G of the 5' cap structure is essential for infection. Lin et al. (2018) developed m⁷G methylated tRNA immunoprecipitation sequencing (MeRIP-seq) and tRNA reduction and cleavage sequencing (TRAC-seq) to conform that Mettl1-mediated tRNA m⁷G modification is essential for the proper expression of neural lineage genes. m⁷G methyltransferase complex METTL1/WDR4 causes primordial dwarfism and brain malformation. Thus, m⁷G sites and human diseases may show associations (Enroth et al., 2019). The study of disease-associated m⁷G may reveal the pathogenesis of the disease.

However, there is still a lack of systematic research on RNA modification due to technical limitations. Few studies have systematically explored the association between m⁷G sites and diseases. It is laborious and expensive to find disease-related m⁷G sites by wet-lab experiments. Recently, more and more artificial intelligence methods have been applied in the analysis of biological data. It can be regarded as a classification issue for disease-related m⁷G sites prediction, where the known association is denoted as 1, 0 otherwise. Some classical classifiers can be used to solve this problem, such as Naive Bayesian (NB), Support Vector Machine (SVM), Random Forest (RF), Gradient Boosting Decision Tree (GBDT), and Matrix Factorization (MF). With Bayes theorem, NB is proposed, which has a strong bias for linearity (Ting and Zheng, 2003). The prediction accuracy decreases dramatically in nonlinear scenarios. SVM is known to be suitable in small sample and nonlinear scenarios (Chang and Lin, 2011), which depends on the kernel to map data to a high-dimensional space. The data about disease-related m⁷G sites are high sparsity, so it is not easy to find the appropriate kernel. RF is an essential method in machine learning and has been widely used in many fields (Ham et al., 2005). However, it is not easy to obtain high precision and generalization performance simultaneously. GBDT is suitable for regression analysis, but the computation load is too high (Rao et al., 2019). Consistent with RF, it is also not suitable for sparse data. MF is the classic model of recommendation system (Lee and Seung, 1999). The low-rank matrix can be used to predict the association between m⁷G sites and diseases. But the higher the requirement of a low-rank matrix, the longer the training time.

In this paper, a deep learning framework based on heterogeneous networks and convolutional neural networks is proposed to find disease-associated m⁷G sites. The site-site similarities were calculated according to the chemical structure of m⁷G site, and the disease-disease similarities were achieved by miRNAs based on induced disease sets. Simultaneously, the known associations between the m⁷G site and the disease were incorporated into the heterogeneous network. Then, the convolutional neural network (CNN) was then adopted to extract multidimensional feature, making full use of the sparse data. Finally, XGBoost was used to predict the associations between m⁷G sites and various diseases.

MATERIALS AND METHODS

Datasets

m⁷G DiseaseDB is an m⁷G-disease association database by taking 1218 disease-associated genetic variants as a bridge, which may lead to gain/loss of the m⁷G sites, with implications for disease pathogenesis involving m⁷G RNA methylation (Song et al., 2020). Among them, 768 associations between 741 m⁷G sites and 177 diseases were extracted via 741 variants with high confidence levels in m⁷G DiseaseDB. Specifically, the genomic locations, host genes of those sites were also included for further feature calculation.

In the mathematical view, let $R \in \mathbb{R}^{M \times N}$ be the association matrix consisting of M sites $S = \{s_1, s_2, \dots, s_M\}$ and N diseases $D = \{d_1, d_2, \dots, d_N\}$. If there is an association between m⁷G site s_i and disease d_j , R_{ij} is 1, 0 otherwise.

Heterogeneous Network Based on Convolutional Neural Network

Figure 1 illustrates the framework HN-CNN. A heterogeneous network was constructed with site-site similarity, disease-disease similarity and the known m⁷G-disease associations to generate feature pairs. Then, each feature pair was transformed into a vector with high-dimensional hidden information by CNN. XGBoost predicts the candidate samples lastly, which chooses the regression classification tree as a base learner.

Feature Vector Construction

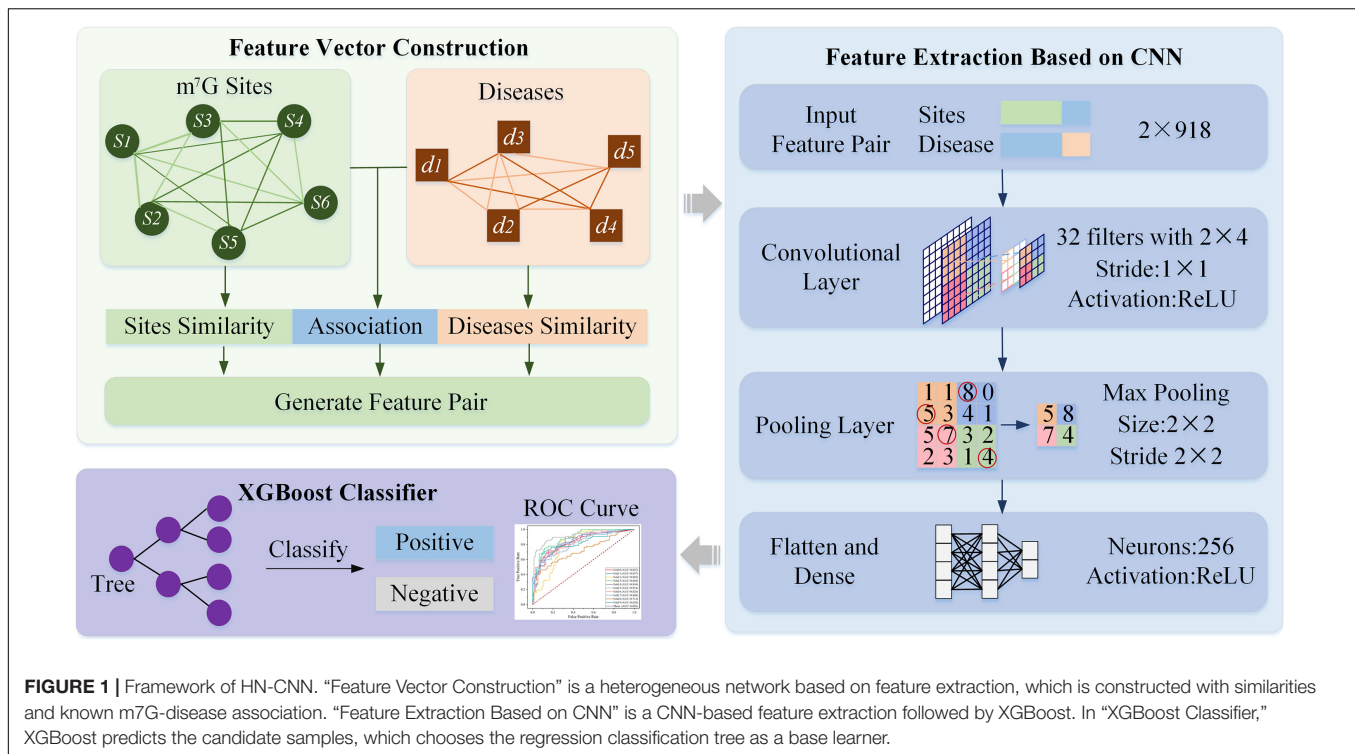
Chemical properties of m⁷G sites were utilized to depict the m⁷G feature just as previously described in similar work (Chen et al., 2019). Based on the chemical features of m⁷G sites, the site similarities were calculated by Jaccard coefficient which is defined as Equation as (1):

$$\text{Jaccard similarity} = \frac{|A \cap B|}{|A \cup B|} = \frac{|A \cap B|}{|A| + |B| - |A \cap B|} \quad (1)$$

where A and B represent the chemical feature of two sites.

In addition, the disease-disease similarity is calculated by DisSetSim (Hu et al., 2017), which is an online system for calculating similarity with diseases names and open source databases. Disease-related genes, functional annotation of genes and the gene functional network of human are involved in calculating disease-disease similarity. Heterogeneous network adopts site-site similarity, disease-disease similarity, combined with the known association between m⁷G sites and diseases, shown directly in Figure 2A.

HN-CNN pays more attention to the latent description of associations of m⁷G sites and diseases. Similarities and association are included in the heterogeneous network. Taking s_5 and d_2 in Figure 2B as an example, vector related to s_5 is selected from the association matrix and site-site similarity, which is different from other sites. Vector related to d_2 is selected from disease-disease similarity and the association matrix to form the vector of d_2 . Those two vectors combine to form the



feature pair about s_5 and d_2 , and each pair is unique. Therefore, the feature pair retains the commonness and the characteristics. Commonness means that the vector representing the same site or disease is invariant. Characteristics means the combination of site-disease is unique, which is different from any other feature pairs. Finally, the feature pair, which is shown in **Figure 2B**, is the connection between heterogeneous network and CNN.

Feature Extraction Based on CNN

Convolutional neural network (CNN) has a deep learning structure, which can mine hidden information. It is superior to the single network in terms of feature extraction and model fitting (Shin et al., 2016). The input layer becomes a multidimensional characteristic surface through the convolutional layer, and the propagation mode between the convolutional layers is shown in Equation (2). Then, features are mapped by pooling, and maximum pooling is shown in Equation (3). Finally, the selected features are flattened to form the final feature vectors:

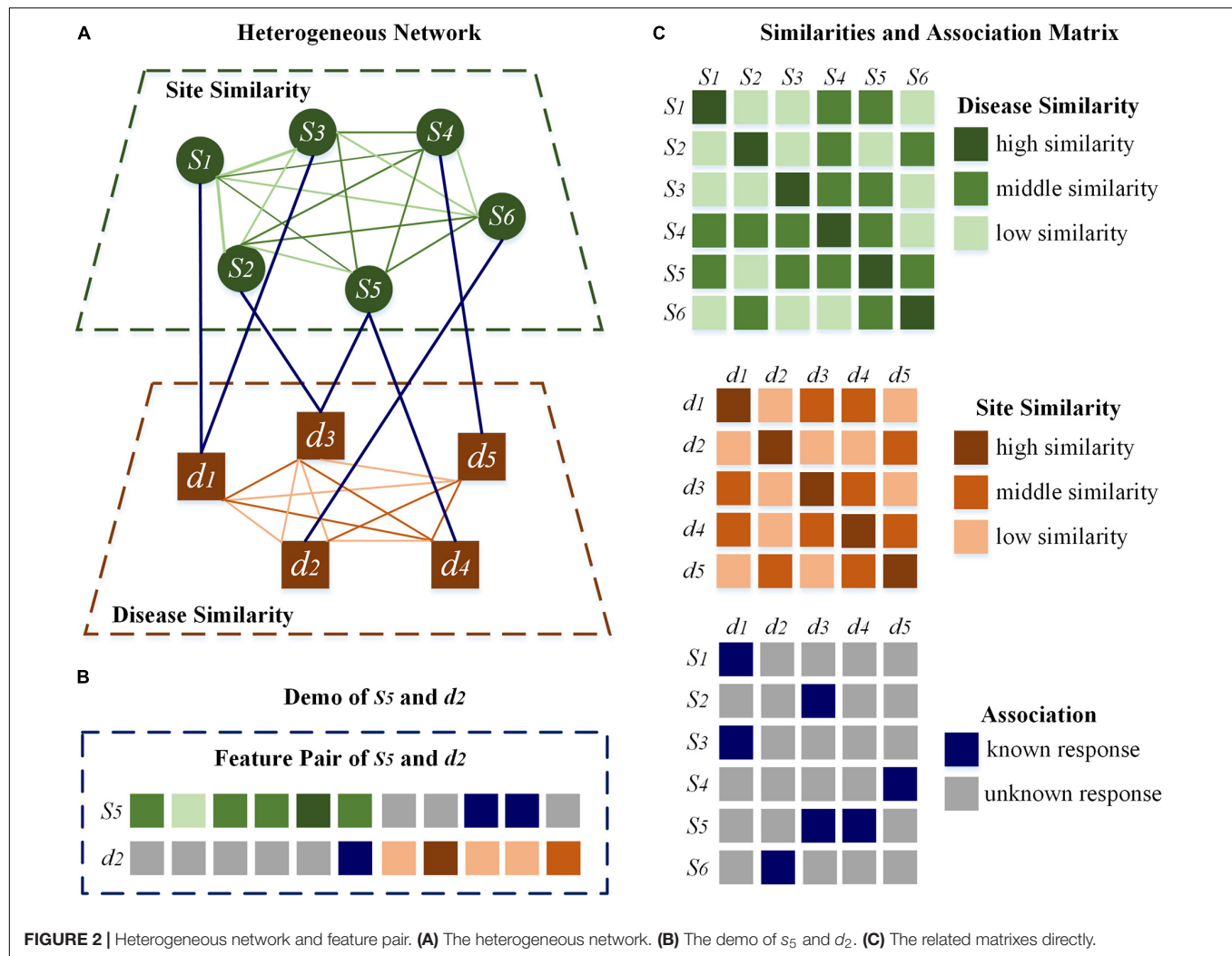
$$H_j^l = \sigma \left(\sum_{i=1}^N H_i^{l-1} * k_{ij}^l + b_j^l \right) \quad (2)$$

where H_j^l is the j -th feature map of the l -th layer, N is the number of the i -th layer's kernels, k_{ij}^l is the j -th element in the i -th convolution kernel at the l layer, b_j^l is the bias parameters, σ is the activation function:

$$\text{Pooling}_j^l = \max_{p \times q} (H_j^l) \quad (3)$$

Where $\max_{p \times q}$ chooses the maximum from H_j^l with the $p \times q$ -size pooling. The Pooling_j^l is the j -th pooling vector in the l -th layer.

Although the feature pairs were achieved in the previous section, the data is sparse with little information. The convolutional layer comprises multiple convolution kernels, which mine different characteristics of feature pairs. Therefore, the generated feature pairs are extracted by CNN. After that, feature vectors are formed, which contain not only various but also different information. In this paper, the associations of adjacent data in feature pairs are weak, so the convolution step size is set as 1 to make full use of each known data and mine each data's hidden information. If the step size becomes bigger, some information will be ignored. The convolution kernel's width was set as 2 to explore the association between m⁷G sites and diseases. To extract more dimensional information and mine the diverse relationships in feature pairs, the more convolution kernels are used, the better performance we have. However, the more computing resources and the longer the computation time are needed with too many kernels, along with the higher repetition rate. Considering high sparsity between the data, such as the sparsity of disease-disease similarity is 72.78%, the number of convolution kernels is set to 32. Meanwhile, the prediction accuracy is the best by experiment. If the number of convolution kernels is reduced, the accuracy will be decreased for mining the information of feature pairs deficiently. When the number of convolution kernels is increased, the accuracy is also decreased for repeated or useless features.



Then, the data are passed into the pooling layer. The pooling layer can reduce the input information dimension, keep the characteristic invariance, select the primary information, and reduce the redundancy information. In this paper, the size of maximum pooling is 2×2 . Length 2 can screen out the data with prominent characteristics between sites and diseases; width 2 can effectively remove the duplicate data and screen out the critical information that has been expanded to the higher dimension.

Finally, feature pairs have been processed into vectors containing various kinds of information, but those vectors contain a large amount of information, with many types. The pooled vectors are compressed by full connection to integrate the feature data. The final feature vectors $V = \{v_1^d, v_2^d, \dots, v_n^d\}$ are formed, where n is the number of known associations, and d is the number of neurons in the full connection layer. In this paper, d is set to 256. When d is less than 256, the performance dramatically decreases due to less information in V . The performance also decreased due to too much or even useless information in V . V contains categorical information, optimizing by cross-entropy, to

make V highly relevant to the original information, and V is used by subsequent classifiers.

XGBoost Classifier

XGBoost classifier is adopted to predict associations between m⁷G site and disease. It retains the feature information better and weakens the influence of parameters on final accuracy. As an integrated learning algorithm that optimizes distributed gradient enhancement, XGBoost has good performance in generalization by regulation and second-order Taylor expansions (Torlay et al., 2017). In this article, the regression classification tree is chosen as a base learner, whose input is V , and output is shown in Equation (4):

$$\hat{y}_i = \sum_{k=1}^K f_k(v_i), f_k \in E \quad (4)$$

where \hat{y}_i is the result, v_i is the i -th vector in eigenvector V , f_k is the k -th decision tree, K is the number of leaf nodes, and E is

the set of classification regression trees. The optimized objective function for XGBoost is shown in Equation (5):

$$L = \sum_{i=1}^n l(\hat{y}_i, y_i) + \sum_{k=1}^K \Omega(f_k) \quad (5)$$

where y_i is the ground truth, and $l(\hat{y}_i, y_i)$ is binary cross-entropy loss and shown in Equation (6):

$$l(\hat{y}_i, y_i) = y_i \ln(1 + e^{-\hat{y}_i}) + (1 - y_i) \ln(1 + e^{\hat{y}_i}) \quad (6)$$

$\Omega(f_k)$ is regularization to prevent overfitting and enhance generalization ability. $\Omega(f_k)$ is shown in Equation (7):

$$\Omega(f) = \gamma T + \frac{1}{2} \lambda \|w\|^2 \quad (7)$$

where γ is the complexity cost by adding new leaf nodes. T is the number of leaves in a tree. $\|w\|^2$ is the sum of the square of each leaf node. λ is the regularization coefficient about the L2 norm $\|w\|^2$.

There are several hyperparameters in XGBoost such as the complexity cost of adding new leaf nodes γ and the regularization coefficient λ . To achieve better AUCs, cross validation is inlaid into XGBoost to find the best parameters with $\gamma \in \{0, 0.2, 0.4, 0.6, 0.8, 1\}$ and $\lambda \in \{0, 0.01, 0.001\}$. Meanwhile, early stopping is adopted to avoid overfitting.

RESULTS

In this paper, HN-CNN is proposed to predict the association between m⁷G sites and diseases, and the performance is evaluated by 10-fold cross-validation. The original correlation matrix only marks the known relationship of m⁷G sites and diseases that can be considered positive, but the unknown does not mean negative. Thus, the same number of the negative is selected from unknown data randomly, and both the positive and the negative constitute the dataset. The set is divided into 10 parts on average, among which nine parts are used for training and the remaining 1 part for testing. The above operation should be repeated 10 times and the AUC should be recorded every time. It should be noted that the test set cannot be repeated in 10 training sets. After 10-folds, the average of 10 AUCs is the final result.

Evaluation Metrics

HN-CNN predicts the positive probability of association between m⁷G sites and diseases. A threshold θ is needed when validation. If the probability is more prominent than θ , the sample is considered as positive. On the other hand, it is identified as negative. True positive rates (TPR) and false positive rates (FPR) are calculated according to the prediction and the truth [Equations (8) and (9)] (Hanczar et al., 2010):

$$TPR = \frac{TP}{TP + FN} \quad (8)$$

$$FPR = \frac{FP}{TN + FP} \quad (9)$$

where TP is true positive, FP is false positive, TN is true negative, and FN is false negative. If θ changes, TPR and FPR will also change. The receiver operating characteristic (ROC) curve is drawn with different $TPRs$ and $FPRs$ (Moses et al., 1993). ROC curve can display the performance of the model intuitively, but it cannot compare models accurately. The area under the ROC curve (AUC) can be used to evaluate the performance of classifier, which ranges from 0 to 1. The more AUC is close to 1, the better performance the classifier has (Fawcett, 2006). So, we choose the ROC curve and AUC to measure the models.

The \overline{AUC} is the mean of m runs of 10-fold cross-validation, which is calculated by Equation (10):

$$\overline{AUC} = \frac{1}{m} \sum_{j=1}^m \left(\frac{1}{10} \sum_{i=1}^{10} AUC_i \right) \quad (10)$$

where m is the number of experiments, AUC_i is the i -th AUC in 10-fold cross-validation. In this paper, $m = 10$.

Comparison With Other Methods

To verify the advantages of CNN in extracting features, features that are not processed by CNN were compared with the features processed by CNN, which are classified with base classifiers such as GBDT, NB, SVM and RF. The result is shown in **Figure 3A**. The ordinate in the figure is the result of 10-fold cross verification, which is the average AUC. All average AUCs are calculated by 10-times of 10-fold cross-validation. The legends “Base Classifier” and “CNN and Base Classifier” are distinguished by whether the feature pair has been processed by CNN. “CNN and Base Classifier” means that feature pairs are processed with CNN, but the models of “Base Classifier” are not, which put feature pairs into classifiers directly.

According to the results in **Figure 3A**, it can be analyzed that the prediction accuracy is significantly improved after CNN extracts the feature with the same parameters and classifiers, which is the most obvious in the RF classifier. Without CNN, the mean AUC is 0.539 by RF. However, the average AUC is 0.698 with CNN, which increased by about 0.16. Besides, it is observed in **Figure 3A** that only the base classifiers without CNN have a greater impact on the prediction results. The average AUC directly predicted by SVM is 0.681, which is about 0.14 higher than that of RF. Classifiers with CNN improve the prediction effect and reduce the gap between classifiers. Therefore, CNN can effectively mine hidden data and improve classification accuracy.

The XGBoost was chosen as the final classifier for two reasons. XGBoost is an integrated machine learning algorithm based on decision trees, and its generalization performance is better than a single classifier. In other words, XGBoost finds the optimal solution within a fixed range of parameters. The results of XGBoost and other methods are shown in **Figure 3B**. CNN+GBDT in X-coordinate means that the features are extracted by CNN and classified by GBDT, and so on. The ordinate is the average AUC of 10-fold cross-validation. It can be analyzed that XGBoost is superior to the base classifiers. The average AUC of HN-CNN is 0.830, which is 0.111 higher than CNN+NB. Therefore, HN-CNN has the advantage in

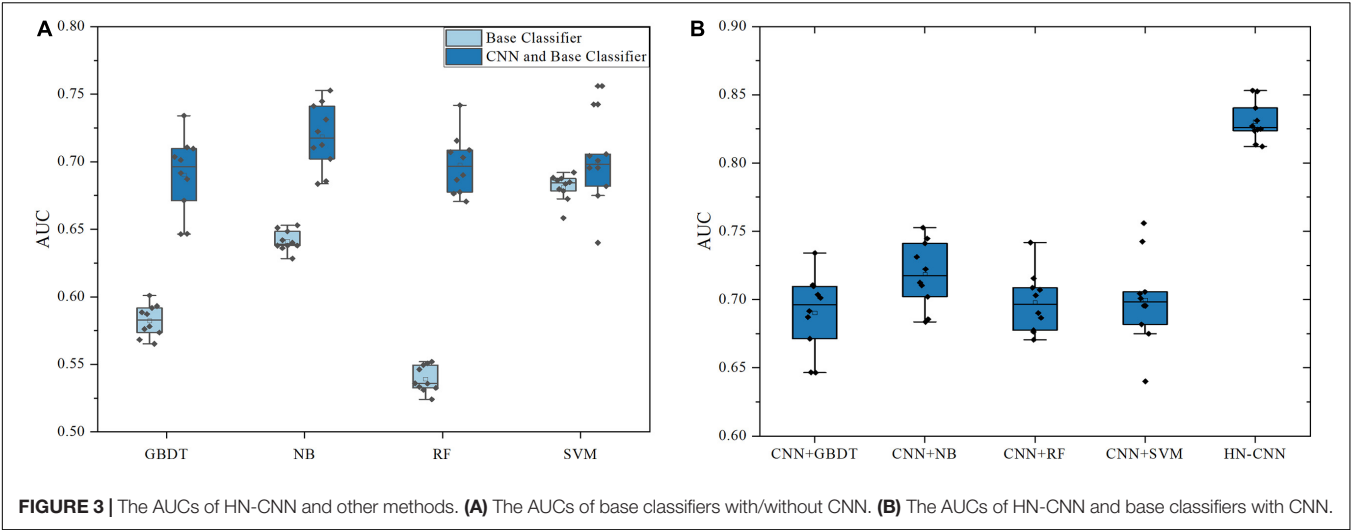


FIGURE 3 | The AUCs of HN-CNN and other methods. (A) The AUCs of base classifiers with/without CNN. (B) The AUCs of HN-CNN and base classifiers with CNN.

TABLE 1 | Case study.

Disease	Gene	GO	p-value	Gene description
Combined oxidative phosphorylation deficiency	FOXRED1	BP	1.03E-04	Mitochondrial respiratory chain complex assembly
Xeroderma pigmentosum	EVC	BP	2.82E-04	Cartilage development
		BP	1.26E-03	Connective tissue development
Moyamoya disease	TPI1	MF	7.09E-04	Isomerase activity
Joubert syndrome	DNAJC5	BP	2.30E-04	Synaptic vesicle exocytosis
		BP	2.98E-04	Synaptic vesicle cycle
		BP	5.08E-04	Vesicle-mediated transport in synapse
		BP	1.23E-03	Neurotransmitter secretion
		BP	1.23E-03	Signal release from synapse
Brody myopathy	PET117	BP	1.03E-04	Mitochondrial respiratory chain complex assembly

feature extraction and classification, which greatly improves the prediction accuracy.

Case Study

The number of known associations is much less than the unknown, which can also be interpreted as the positive is much less than the negative. To weaken the influence of the negative, negative samples equal to the number of positive samples were selected randomly. The highest test accuracy in the 10-fold cross-validation was selected as the final prediction model, which predicts the positive probability of all unknown samples. We selected five of the top 20 to analyze and show the results in Table 1. R. analyzes the related genes with GO based on “clusterProfiler” (Yu et al., 2012). Among the results, CC is short for cellular component, MF is the molecular function, and BP is the biological process. Each gene description is described by *p*-value. If the *p* value is close to 0, the gene description is more obvious.

Combined oxidative phosphorylation deficiency is caused by homozygous or compound heterozygous mutations in the ELAC2 gene, which is a mitochondrial tRNA processing gene (Haack et al., 2013). FOXRED1 can cause complex I deficiency and effect protein function (Calvo et al., 2010). Mitochondrial respiratory chain complex assembly mainly causes mitochondrial

diseases (Deutschmann et al., 2014). There is a high correlation between disease and FOXRED1, in line with the laws of biology.

Xeroderma pigmentosum is a rare genetic disease characterized by extreme photosensitivity, resulting in a higher incidence of cutaneous tumors (Cleaver et al., 1999). EVC is essential for cartilage development (Pacheco et al., 2012). The *p*-value of connective tissue development is 1.26E-03, whose mutations contribute to tumor formation.

Moyamoya disease is a chronic, occlusive cerebrovascular disease with unknown etiology characterized by bilateral stenooclusive changes at the terminal portion of the internal carotid artery and an abnormal vascular network base of the brain (Sakurai et al., 2004). Moyamoya disease is associated with various diseases, like atherosclerosis, autoimmune diseases, Down syndrome. TPI1 is a crucial enzyme in carbohydrate metabolism, negatively associated with tumor size (Jiang et al., 2017). Therefore, TPI1 may inhibit the size of tumors and induce Moyamoya disease.

Inheritance of Joubert syndrome is autosomal and recessive, which is characterized by hypoplasia of the cerebellar vermis (Kendall et al., 1990; Lee et al., 2012). DNAJC5 encodes the cysteine string protein, which is a presynaptic protein implicated in neurodegeneration (Cadieux-Dion et al., 2013). It causes autosomal dominant Kufs disease (Jarrett et al., 2018). One of

Kufs' phenotypes is generalized tonic-clonic seizures, which is similar to related disorders of Joubert syndrome (Chance et al., 1999; Josephson et al., 2001).

Brody myopathy is a rare muscle disorder characterized by exercise-induced impairment of muscle relaxation and stiffness (Odermatt et al., 2000). Pet117 is shown to reside in the mitochondrial matrix, associated with the inner membrane (Taylor et al., 2017). Its gene description hence mitochondrial respiratory efficiency, which is mitochondrial respiratory chain complex assembly (Cogliati et al., 2013). So, it may be further manifested as Brody myopathy symptoms.

DISCUSSION AND CONCLUSION

It is efficient and time-saving to predict the association between m⁷G sites and diseases. HN-CNN integrates diverse information through heterogeneous networks. It adopts CNN to help extract latent relationships in feature pairs, which focuses on personalized associations between m⁷G sites and diseases. At last, XGBoost is used to classify whether there exists association with more generalization. In the 10-fold cross-validation, HN-CNN gets better results than the other methods. The predicted results are analyzed through R to show better demonstrated the reliability of the experimental method in case study. In the future, the data will be updated, and the sparsity will be reduced. HN-CNN will obtain better prediction results in the association prediction due to the amount of data.

REFERENCES

- Bradrick, S. S. (2017). Causes and consequences of flavivirus RNA methylation. *Front. Microbiol.* 8:2374. doi: 10.3389/fmicb.2017.02374
- Cadioux-Dion, M., Andermann, E., Lachance-Touchette, P., Ansorge, O., Meloche, C., Barnabe, A., et al. (2013). Recurrent mutations in DNAJC5 cause autosomal dominant Kufs disease. *Clin. Genet.* 83, 571–575. doi: 10.1111/cge.12020
- Calvo, S. E., Tucker, E. J., Compton, A. G., Kirby, D. M., Crawford, G., Burt, N. P., et al. (2010). High-throughput, pooled sequencing identifies mutations in NUBPL and FOXRED1 in human complex I deficiency. *Nat. Genet.* 42, 851–858. doi: 10.1038/ng.659
- Chance, P. F., Cavalier, L., Satran, D., Pellegrino, J. E., Koenig, M., and Dobyns, W. B. (1999). Clinical nosologic and genetic aspects of Joubert and related <KEYWORDS> syndromes. *J. Child Neurol.* 14, 660–666. doi: 10.1177/088307389901401007
- Chang, C. C., and Lin, C. J. (2011). LIBSVM: a library for support vector machines. *ACM Trans. Intell. Syst. Technol.* 2:27. doi: 10.1145/1961189.1961199
- Chen, W., Feng, P. M., Song, X. M., Lv, H., and Lin, H. (2019). iRNA-m7G: identifying N-7-methylguanosine sites by fusing multiple features. *Mol. Ther. Nucleic Acids* 18, 269–274. doi: 10.1016/j.omtn.2019.08.022
- Cleaver, J. E., Thompson, L. H., Richardson, A. S., and States, J. C. (1999). A summary of mutations in the UV-sensitive disorders: xeroderma pigmentosum, Cockayne syndrome, and trichothiodystrophy. *Hum. Mut.* 14, 9–22. doi: 10.1002/(sici)1098-1004199914:1<9::aid-humu2<3.3.co;2-y
- Cogliati, S., Frezza, C., Soriano, M. E., Varanita, T., Quintana-Cabrera, R., Corrado, M., et al. (2013). Mitochondrial cristae shape determines respiratory chain supercomplexes assembly and respiratory efficiency. *Cell* 155, 160–171. doi: 10.1016/j.cell.2013.08.032
- Deutschmann, A. J., Amberger, A., Zavadil, C., Steinbeisser, H., Mayr, J. A., Feichtinger, R. G., et al. (2014). Mutation or knock-down of 17 beta-hydroxysteroid dehydrogenase type 10 cause loss of MRPP1 and impaired processing of mitochondrial heavy strand transcripts. *Hum. Mol. Genet.* 23, 3618–3628. doi: 10.1093/hmg/ddu072
- Enroth, C., Poulsen, L. D., Iversen, S., Kirpekar, F., Albrechtsen, A., and Vinther, J. (2019). Detection of internal N7-methylguanosine (m7G) RNA modifications by mutational profiling sequencing. *Nucleic Acids Res.* 47:e126. doi: 10.1093/nar/gkz736
- Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognit. Lett.* 27, 861–874. doi: 10.1016/j.patrec.2005.10.010
- Haack, T. B., Kopajtich, R., Freisinger, P., Wieland, T., Rorbach, J., Nicholls, T. J., et al. (2013). ELAC2 mutations cause a mitochondrial RNA processing defect associated with hypertrophic cardiomyopathy. *Am. J. Hum. Genet.* 93, 211–223. doi: 10.1016/j.ajhg.2013.06.006
- Ham, J., Chen, Y. C., Crawford, M. M., and Ghosh, J. (2005). Investigation of the random forest framework for classification of hyperspectral data. *IEEE Trans. Geosci. Remote Sens.* 43, 492–501. doi: 10.1109/Tgrs.2004.842481
- Hanczar, B., Hua, J. P., Sima, C., Weinstein, J., Bittner, M., and Dougherty, E. R. (2010). Small-sample precision of ROC-related estimates. *Bioinformatics* 26, 822–830. doi: 10.1093/bioinformatics/btq037
- Hu, Y., Zhao, L. L., Liu, Z. Y., Ju, H., Shi, H. B., Xu, P. G., et al. (2017). DisSetSim: an online system for calculating similarity between disease sets. *J. Biomed. Semantics* 8:28. doi: 10.1186/s13326-017-0140-2
- Jarrett, P., Easton, A., Rockwood, K., Dyack, S., McCollum, A., Siu, V., et al. (2018). Evidence for cholinergic dysfunction in autosomal dominant kufs disease. *Can. J. Neurol. Sci.* 45, 150–157. doi: 10.1017/cjn.2017.261
- Jiang, H., Ma, N., Shang, Y. R., Zhou, W. T., Chen, T. W., Guan, D. X., et al. (2017). Triosephosphate isomerase 1 suppresses growth, migration and invasion of hepatocellular carcinoma cells. *Biochem. Biophys. Res. Commun.* 482, 1048–1053. doi: 10.1016/j.bbrc.2016.11.156
- Josephson, S. A., Schmidt, R. E., Millsap, P., McManus, D. Q., and Morris, J. C. (2001). Autosomal dominant Kufs' disease: a cause of early onset dementia. *J. Neurol. Sci.* 188, 51–60. doi: 10.1016/s0022-510x(01)00546-9

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

JC and LZ reviewed the resources, wrote the manuscript, and revised the manuscript. JM provided the data and revised the manuscript. HL took the lead in the work and revised the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work is supported by the Fundamental Research Funds for the Central Universities (2019ZDPY15) for support.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2021.655284/full#supplementary-material>

Supplementary Table 1 | The top 50 of most likely site-disease associations.

- Kendall, B., Kingsley, D., Lambert, S. R., Taylor, D., and Finn, P. (1990). Joubert syndrome: a clinico-radiological study. *Neuroradiology* 31, 502–506. doi: 10.1007/bf00340131
- Lee, D. D., and Seung, H. S. (1999). Learning the parts of objects by non-negative matrix factorization. *Nature* 401, 788–791. doi: 10.1038/44565
- Lee, J. E., Silhavy, J. L., Zaki, M. S., Schroth, J., Bielas, S. L., Marsh, S. E., et al. (2012). CEP41 is mutated in Joubert syndrome and is required for tubulin glutamylation at the cilium. *Nat. Genet.* 44, 193–199. doi: 10.1038/ng.1078
- Lin, S., Liu, Q., Lelyveld, V. S., Choe, J., Szostak, J. W., and Gregory, R. I. (2018). Mettl1/Wdr4-Mediated m(7)G tRNA Methylome Is Required for Normal mRNA Translation and Embryonic Stem Cell Self-Renewal and Differentiation. *Mol Cell* 71, 244–255.e5. doi: 10.1016/j.molcel.2018.06.001
- Malbec, L., Zhang, T., Chen, Y. S., Zhang, Y., Sun, B. F., Shi, B. Y., et al. (2019). Dynamic methylome of internal mRNA N(7)-methylguanosine and its regulatory role in translation. *Cell Res.* 29, 927–941. doi: 10.1038/s41422-019-0230-z
- Moses, L. E., Shapiro, D., and Littenberg, B. (1993). Combining independent studies of a diagnostic test into a summary ROC curve: data-analytic approaches and some additional considerations. *Stat. Med.* 12, 1293–1316. doi: 10.1002/sim.4780121203
- Munns, T. W., Liszewski, M. K., Freeman, S. K., and Kaine, J. L. (1985). Detection of human autoantibodies specific for 5'-m⁷GMP and m⁷G(5')ppp(5')N. *Biochem. Biophys. Res. Commun.* 128, 1014–1019. doi: 10.1016/0006-291x(85)90148-2
- Odermatt, A., Barton, K., Khanna, V. K., Mathieu, J., Escobar, D., Kuntzer, T., et al. (2000). The mutation of Pro(789) to Leu reduces the activity of the fast-twitch skeletal muscle sarco(endo)plasmic reticulum Ca²⁺ ATPase (SERCA1) and is associated with Brody disease. *Hum. Genet.* 106, 482–491. doi: 10.1007/s004390000297
- Pacheco, M., Valencia, M., Caparros-Martin, J. A., Mulero, F., Goodship, J. A., and Ruiz-Perez, V. L. (2012). Evc works in chondrocytes and osteoblasts to regulate multiple aspects of growth plate development in the appendicular skeleton and cranial base. *Bone* 50, 28–41. doi: 10.1016/j.bone.2011.08.025
- Pandolfini, L., Barbieri, I., Bannister, A. J., Hendrick, A., Andrews, B., Webster, N., et al. (2019). METTL1 Promotes let-7 MicroRNA Processing via m⁷G Methylation. *Mol Cell* 74, 1278–1290.e9. doi: 10.1016/j.molcel.2019.03.040
- Rao, H., Shi, X. Z., Rodrigue, A. K., Feng, J. J., Xia, Y. C., Elhoseny, M., et al. (2019). Feature selection based on artificial bee colony and gradient boosting decision tree. *Appl. Soft Comput.* 74, 634–642. doi: 10.1016/j.asoc.2018.10.036
- Sakurai, K., Horiuchi, Y., Ikeda, H., Ikezaki, K., Yoshimoto, T., Fukui, M., et al. (2004). A novel susceptibility locus for moyamoya disease on chromosome 8q23. *J. Hum. Genet.* 49, 278–281. doi: 10.1007/s10038-004-0143-6
- Shaheen, R., Abdel-Salam, G. M. H., Guy, M. P., Alomar, R., Abdel-Hamid, M. S., Afifi, H. H., et al. (2015). Mutation in WDR4 impairs tRNA m(7)G(46) methylation and causes a distinct form of microcephalic primordial dwarfism. *Genome Biol.* 16:210. doi: 10.1186/s13059-015-0779-x
- Shin, H.-C., Roth, H. R., Gao, M., Lu, L., Xu, Z., Nogues, I., et al. (2016). Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning. *IEEE Trans. Med. Imaging* 35, 1285–1298. doi: 10.1109/tmi.2016.2528162
- Song, B. W., Tang, Y. J., Chen, K. Q., Wei, Z., Rong, R., Lu, Z. L., et al. (2020). m7GHub: deciphering the location, regulation and pathogenesis of internal mRNA N⁷-methylguanosine (m⁷G) sites in human. *Bioinformatics* 36, 3528–3536. doi: 10.1093/bioinformatics/btaa178
- Taylor, N. G., Swenson, S., Harris, N. J., Germany, E. M., Fox, J. L., and Khalimonchuk, O. (2017). The assembly factor Pet117 couples heme a synthase activity to cytochrome oxidase assembly. *J. Biol. Chem.* 292, 1815–1825. doi: 10.1074/jbc.M116.766980
- Ting, K. M., and Zheng, Z. J. (2003). A study of AdaBoost with naive Bayesian classifiers: weakness and improvement. *Comput. Intell.* 19, 186–200. doi: 10.1111/1467-8640.00219
- Torlay, L., Perrone-Bertolotti, M., Thomas, E., and Baciú, M. (2017). Machine learning-XGBoost analysis of language networks to classify patients with epilepsy. *Brain Inform.* 4, 159–169. doi: 10.1007/s40708-017-0065-7
- Yu, G., Wang, L. G., Han, Y., and He, Q. Y. (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* 16, 284–287. doi: 10.1089/omi.2011.0118

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



The Roles of CircRNAs in Regulating Muscle Development of Livestock Animals

Zhenguo Yang^{*†}, Tianle He[†] and Qingyun Chen

Laboratory for Bio-Feed and Molecular Nutrition, College of Animal Science and Technology, Southwest University, Chongqing, China

OPEN ACCESS

Edited by:

Xiao Han,
Chinese Academy of Agricultural
Sciences, China

Reviewed by:

Rajnish Kumar,
Amity University Uttar Pradesh,
Lucknow Campus, India
Tiande Zou,
Jiangxi Agricultural University, China

*Correspondence:

Zhenguo Yang
guoguo00002@163.com

[†]These authors have contributed
equally to this work

Specialty section:

This article was submitted to
Epigenomics and Epigenetics,
a section of the journal
Frontiers in Cell and Developmental
Biology

Received: 20 October 2020

Accepted: 18 January 2021

Published: 05 March 2021

Citation:

Yang Z, He T and Chen Q (2021)
The Roles of CircRNAs in Regulating
Muscle Development of Livestock
Animals.
Front. Cell Dev. Biol. 9:619329.
doi: 10.3389/fcell.2021.619329

The muscle growth and development of livestock animals is a complex, multistage process, which is regulated by many factors, especially the genes related to muscle development. In recent years, it has been reported frequently that circular RNAs (circRNAs) are involved widely in cell proliferation, cell differentiation, and body development (including muscle development). However, the research on circRNAs in muscle growth and development of livestock animals is still in its infancy. In this paper, we briefly introduce the discovery, classification, biogenesis, biological function, and degradation of circRNAs and focus on the molecular mechanism and mode of action of circRNAs as competitive endogenous RNAs in the muscle development of livestock and poultry. In addition, we also discuss the regulatory mechanism of circRNAs on muscle development in livestock in terms of transcription, translation, and mRNAs. The purpose of this article is to discuss the multiple regulatory roles of circRNAs in the process of muscle development in livestock, to provide new ideas for the development of a new co-expression regulation network, and to lay a foundation for enriching livestock breeding and improving livestock economic traits.

Keywords: circRNAs, livestock animals, muscle development, co-expression regulatory network, transcription and translation

INTRODUCTION

Circular RNAs (circRNAs) are widely found in eukaryotic cells; they are a special type of nucleotide sequence containing conserved microRNA (miRNA) binding sites (Westholm et al., 2014). Studies show that most of the circRNAs that have been found so far are non-coding RNAs; this kind of circRNA does not have the coding function of linear RNA, but it plays a regulatory role in various life activities, including the process of muscle development in livestock animals (Das et al., 2020). Throughout the research results of the past decade, we find that circRNAs are gradually becoming an indispensable part of the gene regulatory network.

Abbreviations: ceRNA: circRNAs-miRNAs-mRNAs; circRNAs: circular RNAs; FoxO1: Forkhead box transcription factor O1; IGF-I: Insulin-like growth factor-I; IGF-II: Insulin-like growth factor-II; MEF2A-2D: myocyte enhancer factor 2A-D; MIRs: Mammalian-Wide Interspersed Repeats; miRNAs: microRNAs; MRF-4: myogenic regulatory factor-4; MSTN: Myostatin; mTOR: Mammalian target of rapamycin; Myf-5: myoblast regulatory factor family myogenic factor-5; Myhc: myosin heavy chain; MyoD: myogenic differentiation antigen; MyoG: myogenin; Pax: paired box; RBP: RNA binding protein; RNA-Seq: RNA sequencing.

Muscle is an essential component of human and most meat animal bodies, and it plays an important role in providing exercise, maintaining posture, and generating heat (Silva et al., 2019; Janssen et al., 2000). Livestock muscle development is a crucial link to individual growth and development and an important research direction in modern animal science. The condition of muscle development directly exerts an economic effect on animals. Therefore, an in-depth understanding of muscle development is of great significance to the development of animal husbandry.

With the rapid development of molecular genetics and the development of molecular biology technology, *in vitro* cell line culture technology, and gene targeting technology, people have become more and more aware of cell growth and development processes at the molecular level (Li X. et al., 2020). For the past few years, scientists have repeatedly pointed out that circRNAs may be an important biomolecule to understand the mechanism of body development. At the same time, circRNAs play a unique role in regulating human and animal muscle development and its related physiological and pathological processes (Das et al., 2020). Liang et al. identify 149 circRNAs that may be related to muscle growth from 3 skeletal muscles of Guizhou miniature pigs (*S. scrofa*); the gene ontology (GO) and KEGG enrichment analysis of the host gene of the circRNAs indicates that these circRNAs are mainly involved in the growth and development of muscle-related signaling pathways (Liang et al., 2017a). The results show that their host genes are closely related to muscle development, chromatin modification, contraction, ATP hydrolysis-coupled proton transport, and cation homeostasis (Liang et al., 2017a). The above studies show that circRNAs are ubiquitous in muscle and play a critical role in the development process. Therefore, in-depth study of the specific mechanism of circRNAs regulating muscle development has become one of the urgent problems for researchers. Here, we focus on the molecular mechanism and mode of action of circRNA as a competitive endogenous RNA in livestock muscle development. The purpose of this study is to expand researchers' understanding of the regulation of muscle development by circRNAs, to collect research data for further improvement of the circRNA-miRNA-mRNA (ceRNA) co-expression regulation network, and to provide theoretical support for the improvement of muscle development and economic traits of livestock.

OVERVIEW OF CIRCRNAS

The Discovery of CircRNAs

CircRNA is a kind of closed circular RNA, which can stably exist in the organism, but it does not have the 5' terminal hat structure and the 3' terminal poly (A) tail structure (Westholm et al., 2014). As early as the 1970s and 1980s, researchers proposed in Nature and PANs that circRNA is a kind of covalently closed circular RNA molecule discovered in plant viroids and eukaryotic cells (Sanger et al., 1976; Hsu and Coca-Prados, 1979; Kos et al., 1986). Later, Danan et al. (2012) found that some non-coding RNAs, snoRNAs, and RNase P RNAs could form circRNAs in archaea. Although these early studies clearly

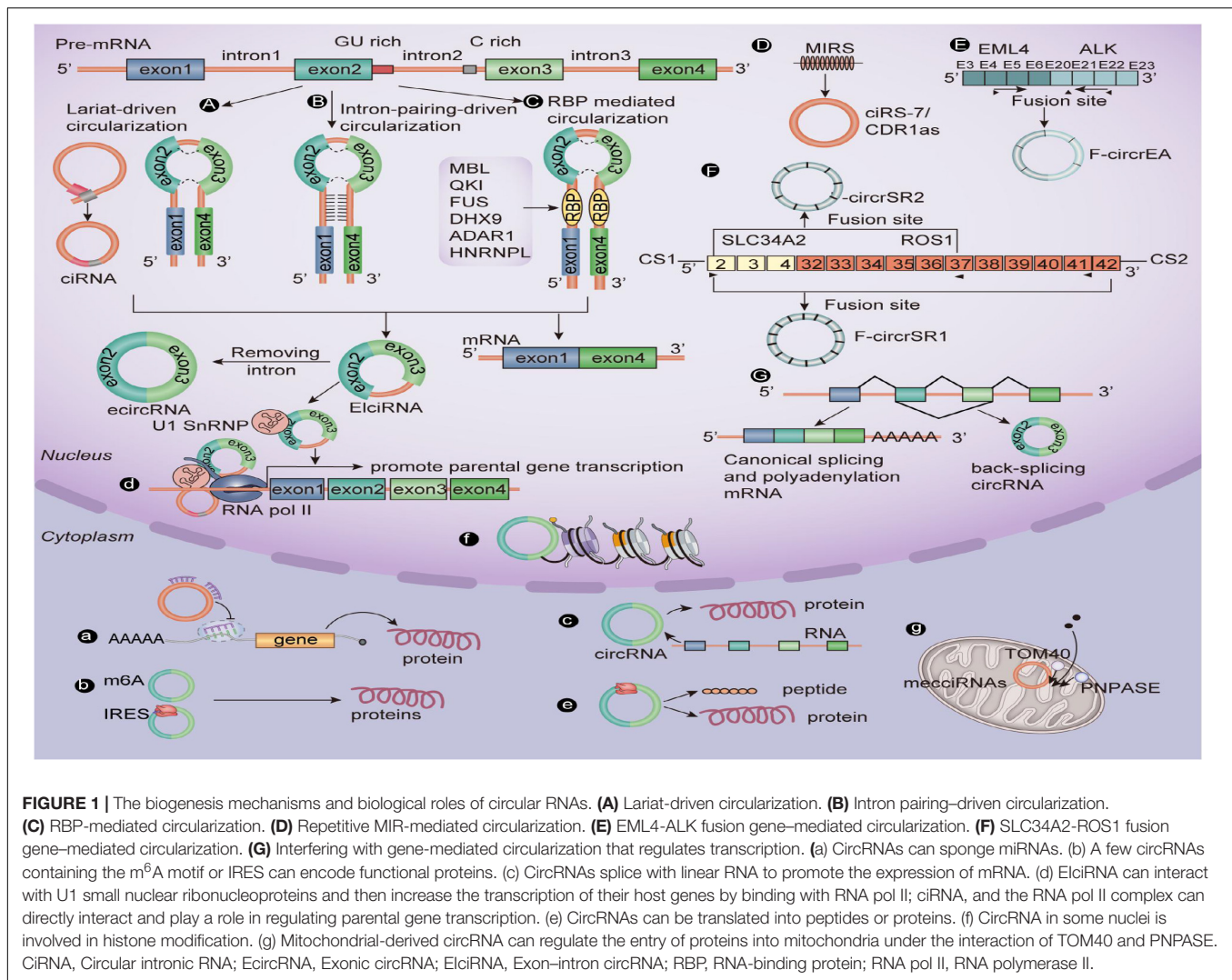
document the existence of circRNA molecules, their potential impact was underappreciated. With the advent of some advanced RNA sequencing techniques and methods for calculating non-polyadenylate RNA transcription, more and more circRNAs have been found in paleontological, nematode, zebrafish, mouse, and human cells (Zhang et al., 2014), and regulation of circRNA levels can lead to a variety of molecular and physiological phenotypic changes, which include effects on growth and development (Liu B. et al., 2019), the nervous system (Irie et al., 2019), innate immunity, microRNAs (Yang et al., 2017), and many disease-related pathways (Du et al., 2018). As a result, circRNA became a hot spot.

The Classification of CircRNAs

In order to carry out follow-up research in a more organized way, scholars depend on the genomic loci and the relationship with the connected parental transcript; circRNAs are categorized into five types: exonic (Danan et al., 2012; Salzman et al., 2012; Memczak et al., 2013; Li X. et al., 2019), intronic (Zhang et al., 2013; Das et al., 2019a,b), sense overlapping (Humphreys et al., 2019; Patop et al., 2019), antisense, and intergenic (Qu et al., 2015; Wang et al., 2016). Cao et al. (2018) found 886 circRNAs in the form of introns and exons in the sheep skeletal muscle circRNA library—most of which interact with muscle-specific miRNA involved in muscle growth and development, especially circ776. It is worth noting that there are few reports on other types of circRNAs regulating muscle development. Therefore, this is also a place where we need further study and breakthroughs.

The Biogenesis of CircRNAs

The difference in structure between circRNA and linear RNA means that they are formed in different ways and have different biological functions (Figure 1). Li X. et al. (2019) prove for the first time that the assembly mechanism of the “intro-definition” and “exon-definition” E complex can exist from the point of view of structural biology. On this basis, they point out that some E complexes assembled on the middle exons of yeast EFM5 or HMRA1 can be chased into circRNA, but this requires exons long enough to achieve this, and most eukaryotic circRNAs are catalyzed by classical splice bodies or group I/II ribozymes (Lasda and Parker, 2014; Chen and Yang, 2015). With reference to the existing research, we summarize and sort out the formation of circRNAs: First, the lasso structure drives cyclization, and the lasso structure is a by-product of exon hopping. After the intron in the lasso structure is removed, exons can be connected to form circRNAs. The second is intron pairing driving cyclization, and there is an intron with a reverse complementary sequence at both ends of the ring-shaped exon. The pairing mediation of the reverse complementary sequence of the intron can make the splicing donor and splice recipient of the exon spliced into a ring closer to each other in space, thus forming circRNAs (Li X. et al., 2019). Besides this, the RNA binding protein (RBP) is an important factor in regulating the production of circRNA. RBP can specifically bind to the flanking introns at both ends of RNA, acting as an RBP while narrowing the distance between the splice recipient and the donor, resulting in the formation of circRNAs (Naqvi et al., 2019;



Paglierini et al., 2020; Zhao et al., 2020). It is well-known that CDR1as, as an antisense transcript of cerebellar degeneration associated protein 1, can be used as a specific circRNA of miR-7, so it is also known as CIRS-7. However, it is surprising that ciRS-7/CDR1as biosynthesis in circRNAs is mediated by mammalian scattered repetitive elements mammalian-wide interspersed repeats (MIRs) (Yoshimoto et al., 2020). In addition, a recent study shows that m⁶A modification can promote the production of circRNAs carrying an open reading frame (ORF) during the development of male germ cells in mice (Tang et al., 2020). Tan et al. (2018) report that the EML4-ALK fusion gene can form F-circrEA. Later, Wu et al. (2019) found that SLC34A2-ROS1 can form F-circrSR1 and FmurcircrSR2 by the fusion gene, and both of them may be used as diagnostic markers of lung cancer. Also, some scholars find that genes that interfere with transcriptional termination contribute to the production of transcriptional read-through products and promote the production of downstream gene-derived circRNAs (Liang et al., 2017; Chen S. et al., 2019). From this point of view, the formation of circRNAs is affected by many biological

factors. Therefore, the formation of circRNAs still needs in-depth study by researchers.

The Biological Characteristics of CircRNAs

Many studies show that circRNA is a structure with a missing 5' end cap and 3' end poly (A) tail, so it is not easy for circRNAs to be degraded by exonuclease RNaseR (Suzuki et al., 2006; Suzuki and Tsukahara, 2014; Li X. et al., 2019). Further, due to circRNAs being highly conserved in many species, such as humans, nematodes, zebrafish, and mice (Salzman et al., 2012; Li X. et al., 2019), only a small number of them can evolve and change rapidly (Zhang et al., 2013, 2014). Some studies point out that the number of circRNAs found in eukaryotic cells has been as high as more than 20,000 (Cocquerelle et al., 1993; Westholm et al., 2014); a few are formed by direct cyclization of introns, and most of them come from exons (Liu et al., 2020a). In addition, as a kind of non-coding RNA, circRNA can only regulate the formation of proteins at the transcriptional or post-transcriptional level (Westholm et al., 2014), and only a few of them can regulate life activities by encoding proteins

(Pamudurti et al., 2017; Yang et al., 2018). Medical researchers believe that circRNAs have certain tissue and disease specificity and have guiding significance for the treatment of many diseases (Cocquerelle et al., 1993).

The Biological Function of CircRNAs

In the past decade, research results have shown that circRNAs can directly or indirectly participate in the process of gene expression, such as RNA translation, miRNA bait, RNA translation, protein–protein interaction, and so on (Westholm et al., 2014; Pamudurti et al., 2017; Yang et al., 2018) (**Figure 1**). Based on the results of previous studies, we summarize the biological functions of circRNAs: (1) CircRNAs can be used as an miRNA sponge to regulate the stability of related mRNAs or protein formation. For example, circHIPK3 acts as a sponge of miR-7 in CRC when the co-expression of miR-7 and circHIPK3 promotes the proliferation of colorectal cancer cells (Zeng et al., 2018). (2) A few circRNAs can encode functional proteins when they contain the m⁶A motif or IRES. For example, studies show that some circRNAs can carry longer ORFs, and the initiation codon of these ORFs is modified by m⁶A and binds to ribosomes to form functional proteins (Zhao et al., 2018; Tang et al., 2020). (3) CircRNAs can be spliced with linear RNA and promote the expression of linear mRNA. For example, complementary pairing of CDR1as and CDR1mRNA can enhance the stability of CDR1mRNA (Rong et al., 2017). (4) EIcircRNA or ciRNA interacts with Pol II and U1 snRNP at the promoter of the parent genes, thus promoting the transcription of the parent genes. For example, circEIF3J and circPAIP2 regulate gene expression by forming complexes with U1 snRNP and Pol II, which bind to the promoter region of the host gene (Li D. et al., 2020). (5) Some circRNAs can be used as translation templates for proteins and peptides. For example, some ribo-circRNAs use the start codon of the host mRNA and bind to membrane-related ribosomes to participate in circRNA translation (Granados-Riveron and Aquino-Jarquín, 2016; Pamudurti et al., 2017). (6) CircRNAs in some nuclei are also involved in histone modification (Burd et al., 2010; Kotake et al., 2011) and RNA maturation (Holdt et al., 2016). CircRNAs (meccRNAs) from humans, and mouse mitochondria can enter mitochondria by interacting with TOM40 and PNPASE (Liu et al., 2020b).

The Degradation of CircRNAs

CircRNAs maintain their cellular homeostasis by highly dynamic and tightly regulated biogenesis and degradation, thereby exerting proper biological functions. Compared with the biogenetic mechanism of circRNA biogenesis, the specific pathway by which cells eventually degrade circRNAs is still in need of continued study. Some studies show that miRNAs may initiate the degradation of circRNAs through Ago2-mediated cleavage. For instance, as the target of CDR1as/ciRS-7, miR-671 can perfectly load Ago2 into CDR1as/ciRS-7, which leads to the cleavage of Ago2 in the nucleus and the subsequent dissolution of RNA in the outer nucleus. However, it is not known whether other miRNAs can regulate the degradation of circRNAs by perfectly matching them with Ago2 (Hansen et al., 2011). A recent study has shown that m⁶A can mediate

circRNA degradation. Park et al. (2019) point out that, when m⁶A carrying circRNA is used as a marker, m⁶A read-write protein YTHDF2 and linker protein HRSP12 can be recruited. As a bridge between YTHDF2 and endoribonuclease RNase P/MRP, HRSP12 binds directly to the GGUUC motif on circRNAs, which leads to the initiation of RNase P/MRP and the gradual degradation of circRNAs. In addition, Fischer et al. (2020) find a degradation mechanism of RNA (including circRNAs). He also points out the specific mechanism of selective degradation of high-structure RNA under normal conditions. High overall structure circRNA decay is regulated globally by two RNA binding proteins, UPF1 and G3BP1 (Chen et al., 2020). Because this pathway perceives the whole RNA structure rather than a linear first-order sequence, it is called structure-mediated RNA decay (SRD). Because mammalian RNA decay pathways are widely linked to translation (Tuck et al., 2020), it is also worth exploring whether circRNAs targeted by SRD have potential encoding peptides. Furthermore, Liu C. X. et al. (2019) discovered an endonuclease RNaseL that can degrade circRNAs in a full range, which seems to increase researchers' understanding of the mechanism of circRNA degradation.

GENERAL SITUATION OF MUSCLE DEVELOPMENT OF LIVESTOCK ANIMALS

Muscle development is a very complex biological process, which mainly depends on the proliferation and hypertrophy of muscle fiber cells (Molkentin and Olson, 1996). Studies show that the number of muscle fibers increases only before birth and does not change much after birth, and the growth of muscle fibers depends on the hypertrophy of muscle fibers (Christ and Brand, 2004). The hypertrophy of muscle fibers includes two aspects: One is the increase of myofibrils, and the other is the increase in the number of nuclei in muscle fibers (Berger et al., 2015). Muscle development is a multistep process regulated by multiple genes, and it is not only regulated by a variety of myogenic regulatory factors, including myogenin (MyoG), myogenic regulatory factor-4 (MRF-4), myosin heavy chain (MyhC), myoblast regulatory factor family myogenic factor-5 (Myf-5), myostatin (MSTN), myocyte enhancer factor2A-D (MEF2A-2D), myogenic differentiation antigen (MyoD), and paired box (Pax) family members Pax 3 and Pax 7 (Chen et al., 2020), but it is also regulated by other related genes, such as insulin-like growth factor-I (IGF-I), insulin-like growth factor-II (IGF-II), forkhead box transcription factor O1 (FoxO1), and mammalian target of rapamycin (mTOR) (Wan et al., 2016). Furthermore, studies confirm that the genes that affect muscle development have two regulatory effects: One is positive promotion; the other is reverse inhibition (McPherron and Lee, 2002; Chen P. R., 2019; Chen et al., 2020). These growth factors play a unique role in regulating muscle development and can regulate cell proliferation, apoptosis, sarcomere activation, and muscle-specific genes at many sites in the muscle (Doynova et al., 2017; Hernández-Hernández et al., 2017) pedigree.

Exploring the regulatory mechanism of muscle cell proliferation and differentiation is one of the research hot spots in developmental biology in recent years. As an important economic trait, the muscle development of livestock has been paid more and more attention by researchers. CircRNAs have been widely studied in the related fields of human medicine and bioinformatics, which provides a new idea for exploring the construction of regulatory networks of circRNAs in animal husbandry. As was mentioned earlier, muscle development is a complex biological process that is affected by many factors. Although researchers have conducted extensive and in-depth studies on circRNAs, some researchers noticed that circRNAs can further guide the process of muscle development by binding to miRNAs or regulating the expression of genes related to muscle development at the transcriptional level (Zhang et al., 2019). Based on previous studies, we summarize the unique role of circRNAs in muscle development (as shown in the **Table 1** and **Figure 2**).

It is reported that the number of circRNAs in animal muscles and muscle cells ranges from 2,000 to 37,000 (Abdelmohsen et al., 2015; Li et al., 2017; Liang et al., 2017b; Wei et al., 2017; Ouyang et al., 2018; Zhang et al., 2018). Li et al. (2017) first used RNA sequencing to detect 6,113 circRNAs from the longissimus dorsi muscle of sheep. The researchers found 12,000 circRNA expressions during muscle aging in monkeys by RNA sequencing (Abdelmohsen et al., 2015). The study found that circRNAs regulated the growth and development of porcine skeletal muscle and the transformation of muscle fiber types at the age of 0–30 days; at the age of 30–240 days, circRNAs regulated the glucose metabolism and calcium signal of porcine skeletal muscle (Liang et al., 2017b). Many reports point out that circRNAs play a unique and irreplaceable role in guiding animal muscle development (Abdelmohsen et al., 2015; Li et al., 2017; Liang et al., 2017b; Wei et al., 2017; Ouyang et al., 2018; Zhang et al., 2018). Liang et al. (2017b) constructed the first miniature pig circRNA database and point out that ssc-ciR-02753, ssc-ciR-03065, ssc-ciR-03066, ssc-ciR-03069, ssc-ciR-04335, ssc-ciR-04348, ssc-ciR-04349, ssc-ciR-04353, and ssc-ciR-04359, can regulate porcine muscle development by affecting cell proliferation and fusion. It is reported that many circRNAs contain binding sites, such as miR-1, miR-133, miR-206, miR-29, miR-378, miR-431 (Ebbesen et al., 2017), miR-7 (Li L. et al., 2019), miR-135a (Greco et al., 2009), miR-1290 (Ng et al., 2015), and miR-876-5p (Cook et al., 2015). Interestingly, these miRNAs are involved in cell proliferation, differentiation, and signal transduction during myogenesis. It lays a theoretical foundation for the construction of a co-expression network of circRNAs regulating the muscle development of livestock. Combined with human medical research, it is not difficult to see that circRNAs and miRNAs can bind and regulate the expression of downstream mRNAs (that is, the ceRNA co-expression regulatory network) in a certain way. A lot of evidence also shows that the muscle development of livestock is indeed regulated by both circRNAs and miRNAs (Abdelmohsen et al., 2015; Li et al., 2017; Liang et al., 2017b; Wei et al., 2017; Ouyang et al., 2018; Zhang et al., 2018). In addition, a very small number of circRNAs can regulate muscle

development in livestock at the transcriptional and translational levels (Shen et al., 2019). However, there is almost no relevant research in this field at present. Therefore, it is necessary for us to pay close attention to the related regulatory mechanisms of circRNAs on muscle development in different ways.

CIRC RNAs REGULATE MUSCLE DEVELOPMENT OF LIVESTOCK ANIMALS

CeRNA Co-expression Network Regulates Muscle Development of Livestock Animals

Muscle development of livestock is an important economic character in the development of animal husbandry. Despite the controversy surrounding the ceRNA hypothesis, large amounts of experimental results show that circRNAs can regulate the expression of mRNAs related to the muscle development of livestock through specific miRNAs (Qian et al., 2017). Therefore, researchers have evaluated multiple co-regulatory relationships during muscle development. Among them, Yue et al. (2020) verify the interaction among circHUWE1, miR-29b, and AKT3 with the help of bioinformatics, double luciferase report analysis, and AGO2-RNA immunoprecipitation (RIP) and point out that circHUWE1 can directly interfere with the ability of miR-29b to release AKT3 inhibition and finally activate the AKT signal pathway, thus promoting the proliferation of bovine myoblasts and inhibiting apoptosis and differentiation of bovine myoblasts. For this reason, Yue et al. (2020) build upon previous research results to construct a ceRNA co-expression regulatory network, which not only provides a good idea for the study of muscle development, but also expands our understanding of the function of circRNAs. KEGG pathway analysis shows that hosting genes of circRNAs are related to the muscle development pathway, including the mammalian target of rapamycin signaling pathway, Wnt signaling pathway, MAPK, and transforming growth factor- β signaling pathway (Li et al., 2017). Yan et al. (2020) identify a total of 5,177 circRNAs in the longissimus dorsi samples of Kazakh cattle and Xinjiang brown cattle by establishing a RNA sequence library; of these, 46 are differentially expressed. The identification of differentially expressed genes shows that the process of muscle development is related to differentially expressed circRNAs. In addition, miRanda predictions show that there are 66 interactions between 65 circRNAs and 14 miRNAs. For this reason, Yan et al. (2020) also establish a co-expression network. Further, some miRNAs known to be involved in myoblast regulation, such as miR-133b and miR-664a, are identified. Sun et al. (2017) study the differentially expressed coding genes, lncRNAs, circRNAs, and miRNAs, in muscle tissue of Lantang and Landrace pigs. The results show that there are 1,401 high expressions of circRNAs and 2,959 low expressions of circRNAs in Lantang pig, of which 236 circRNAs are closely related to muscle development, and 40 circRNAs regulate muscle development by participating in a miRNA-mediated ceRNA

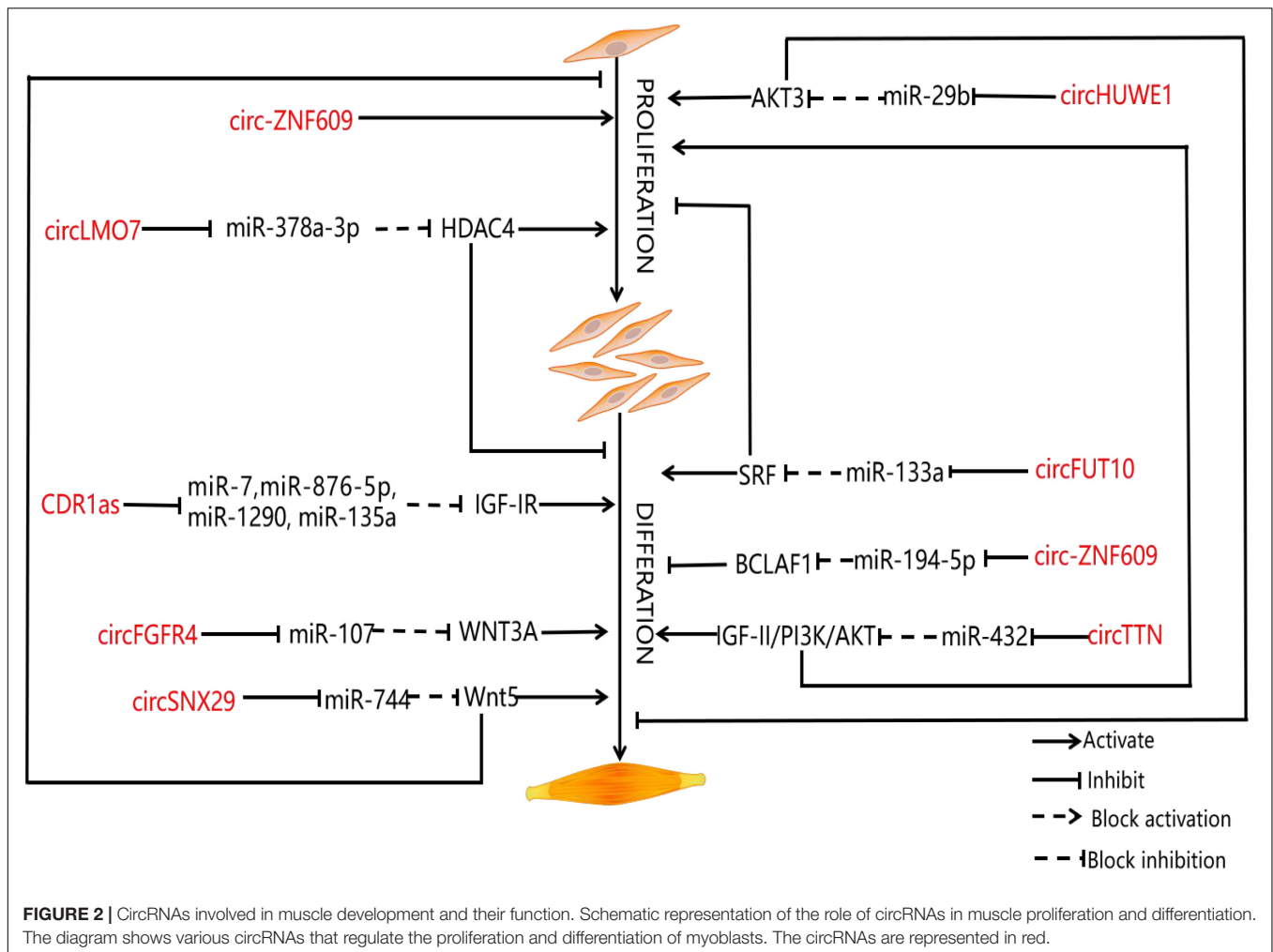
TABLE 1 | CircRNAs involved in muscle development and their function.

circRNAs	Parental gene	Organism	Biological roles	Target miRNA(s)	miRNAs target	References
circ-ZNF609	ZNF609	Homo sapiens	Inhibits differentiation, promotes proliferation; can be translated	miR-194-5p	BCLAF1	Legnini et al., 2017
circLMO7	LMO7	Bovine	Inhibits differentiation, promotes proliferation of primary bovine myoblasts and protects them from apoptosis	miR-378a-3p	HDAC4	Wei et al., 2016
circFUT10	FUT10	Cattle	Reduces proliferation and facilitates differentiation of bovine myoblasts	miR-133a	SRF	Li et al., 2018
circSNX29	SNX29	Bovine	Promote myoblast differentiation and inhibit cell proliferation	miR-744	Wnt5a CaMKII δ and Ca ²⁺	Peng et al., 2019
circFGFR4	FGFR4	Bovine	Promote cell differentiation	miR-107	WNT3A	Li et al., 2018
circHUWE1	HECT, UBA and WWE	Bovine	Promotes myoblast proliferation and inhibits differentiation	miR-29b	AKT3	Yue et al., 2020
bta_circ_03789_1, bta_circ_05453_1	ND	Cattle	May regulate the IGF-IR gene by regulating the miRNAs associated with the longissimus dorsi muscle, and finally regulate muscle development	miR-133b and miR-664a	IGF-IR	Yan et al., 2020
circTTN	TTN	Bovine	Promoted proliferation and differentiation	miR-432	IGF-II, PI3K, AKT	Wang et al., 2019
circ776	ND	Sheep	Involved in muscle cell development and signaling pathway	miR-208	ND	Cao et al., 2018
circ_0001573, circ_0001554, circ_0013564	ND	Pig	Regulation of muscle fiber transformation through mi-499-5p	miR-499-5p	KCNQ1, MRAS and SERTM1	Li B. et al., 2020
CDR1as	Human X Chromosome	Goat	Promotes myoblast differentiation	miR-7, miR-1290, miR-876-5p, miR-135a	IGF-1R	Sun et al., 2017

ND, not determined.

regulatory network. There are 6,113 differentially expressed circRNAs in the longissimus dorsi of Kazakh sheep at both embryonic and adult stages, and the maternal genes of circRNAs are mainly enriched in the signal pathways related to muscle growth and development. Among them, oar_circ_0000385, oar_circ_0000582, and oar_circ_0001099 have multiple binding sites on miRNAs (miR-143, miR-133, and miR-23) related to muscle development (Li et al., 2017). Wei et al. (2017) detect the circRNA expression profile of bovine skeletal muscle at two developmental periods (embryonic and adult longissimus) for the first time and point out that the overexpression of circLMO7 can competitively absorb miR-378a-3p when the expression of miR-378a-3p is downregulated; the target gene hdac4 is activated, thus promoting muscle cell proliferation and inhibiting muscle cell differentiation. After that, Li et al. (2018) find that circFGFR4 and miR-107 are highly expressed in Qinchuan cattle at the embryonic stage (90 days) and adulthood (24 months old) longissimus dorsi, and circFGFR4 could adsorb miR-107. MiR-107 can weaken the expression of Wnt3a by binding to overexpressed circFGFR4. It can be seen that Wnt3a, as the target of miR-107, plays an important role in inhibiting

myotube formation and protecting myoblast apoptosis. To sum up, circFGFR4 can be used as an miR-107 sponge to eliminate the inhibitory effect of miR-107 on the expression of Wnt3a and the differentiation of bovine primary myoblasts. The differential expression of circTitin (circTTN) in bovine skeletal muscle between fetal and adult bovine muscle tissue and the overexpression and inhibition of circTTN induced its promoting effect on the proliferation and differentiation of bovine primary myoblasts because miR-432, the target gene of circTTN, is the regulator of IGF-II. Wang et al. (2019) point out that circTTN can activate the IGF-II/phosphatidylinositol 3-kinase (PI3K)/AKT signal pathway through competitive binding with miR-432, which promotes the proliferation and differentiation of bovine myoblasts. Li et al. (2018) find that circFUT10 in adult bovine muscle can reduce its inhibition on target genes by competitive binding to miR-133a, and the expression of MyHC, MyoD, and MyoG related to muscle development changed synchronously with that of circFUT10 at the mRNA and protein levels. Overexpression of circ-FUT10 can promote MyHC expression, induce myoblast apoptosis, and promote myoblast differentiation. On the contrary, circZfp609 binding to



miR-194-5p inhibits the expression of bcl2-related transcription factor 1 (BCLAF1), which also affects the expression of Myf5 and MyoG, and inhibits myoblast differentiation (Li et al., 2018). CircSNX29 exists widely in bovine primary myoblasts, but its expression level in embryonic skeletal muscle is significantly higher than that in adult skeletal muscle. Peng et al. (2019) find that overexpression of circSNX29 promotes myoblast differentiation and inhibits cell proliferation while interfering with circSNX29 inhibits myoblast differentiation and promotes proliferation. Subsequently, using RNAhybrid for bioinformatics prediction, it was found that circSNX29 may adsorb to miR-744 with 9 potential binding sites. Using a double luciferase report assay, the results show that circSNX29 could directly bind to miR-744 competitively and effectively reverse the inhibitory effect of miR-744 on Wnt5a and CaMKIId. Importantly, through KEGG pathway enrichment analysis, Western blotting, a calcium fluorescence probe, and CamKII activity detection, it is found that overexpression of Wnt5a and circSNX29 activate the non-classical Wnt/Ca²⁺ pathway by increasing the activity of CamKII kinase and the phosphorylation level of PKC and then regulate the proliferation and differentiation of bovine myoblasts. These results are helpful to further understand the role of circRNAs and

miRNAs in myogenesis. The study further shows that muscle development is more efficient in embryo than in adulthood. Interestingly, Von et al. point out that Wnt5a and CaMKIId are the targets of miR-744, and the expression of miR-744 leads to the activation of the atypical Wnt pathway by inhibiting the expression of Wnt5a and CaMKIId (von Maltzahn et al., 2012). From this, we can see that circSNX29 acts as an miR-744 sponge to upregulate the expression of CaMKIId and Wnt5a through activating the Wnt pathway and promotes myoblast differentiation. In addition, it is reported that circRNA9210-miR-23a-MEF2C and circRNA290-miR27b-Foxj3 networks play a unique role in regulating the conversion of muscle fiber types in porcine skeletal muscle (Shen et al., 2019). Li et al. (2017) find a total of 5,086 differentially expressed circRNAs in the RNA sequences of sheep adult longissimus dorsi (LDM-A) and longissimus dorsi (LDM-E) of which 2,146 are downregulated and 2,940 are upregulated. The results of real-time quantitative PCR show that the expression of circRNA 0000552, circRNA 00002456, circRNA 00004666, circRNA00004676, and circRNA 00004690 in LDM-E is relatively higher than that of LDM-A. The expression of circRNA 0003451, circRNA 0005243, circRNA 0005250, and circRNA 0005256 in LDM-A is relatively

higher than that in LDM-E. Thus, the differential expression of circRNAs in sheep muscle is proved. Cao et al. (2018) extracted 75.5 million sequences from the sheep skeletal muscle RNA gene bank. These sequences were mapped to 729 genes in the reference genome of sheep, containing a total of 886 circRNAs. Reverse transcription PCR and DNA sequencing analysis confirm the existence of many kinds of circRNAs and the resistance of sheep circRNAs to RNase R digestion. Finally, Cao et al. (2018) first used RNA-seq to study circRNAs in the longus dorsi muscle of sheep before and after delivery. A total of 6113 circRNAs were detected, of which some circRNAs (circRNA100, circRNA108, circRNA205, circRNA606, circRNA678, circRNA744, and circRNA776) contained at least two conservative targets of miRNAs related to muscle development (miR-29b, miR-133, miR-208, and miR-499, respectively). Thus, it can be seen that most circRNAs interact with muscle-specific miRNAs and then jointly regulate the process of muscle development. The results of GO and KEGG enrichment analysis show that the host gene of circRNAs is involved in muscle cell development and signal transduction (Cao et al., 2018). For this reason, Cao et al. (2018) establish a relatively complete ceRNA network that contains a large number of potential functional circRNA (circRNA 0000385, circRNA 0001099, and circRNA 0000582) and its predicted miRNA targets and downstream regulatory genes. At present, these regulatory networks are an important source of ideas for us to study muscle development. This information may help us to further explore the unique role of circRNAs in muscle development.

As we all know, circRNA CDR1as (CDR1as or CiRS-7) is an antisense transcript of cerebellar degeneration associated protein 1, but in fact, CDR1as is also considered to be related to miRNAs related to muscle development (Geng et al., 2016; Sang et al., 2018). Although CDR1as was identified as the specific circular RNA of miR-7, other miRNAs, such as miR-135a, miR-876-5p, and miR-1290, are also shown to be CDR1 response elements (Geng et al., 2016; Sang et al., 2018). It is worth noting that CDR1as-responsive miRNA and its targeted muscle-derived genes, such as IGF-IR, N-cadherin, ABCG2, WNT5A, EGFR, FAK, and CCNE1, play a key role in normal conditions and muscle diseases, such as DM1, FSHD, and IIM, which makes CDR1 potentially an important regulatory factor in muscle (Geng et al., 2016; Sang et al., 2018; Kyei et al., 2020). First of all, Li L. et al. (2019) discovered that MyoD promotes CDR1as by binding on the CDR1as 5' flank region; however, the overexpression or knockout of CDR1as can significantly induce or hinder the process of muscle differentiation. Second, CDR1as can reduce the downregulation of IGF-IR induced by miR-7 through competitive binding to miR-7, thus activating muscle differentiation in goat metaphase. The above results further indicate that CDR1as plays an irreplaceable role in the regulation of muscle development. At the same time, these potential CDR1as/miRNAs/mRNA regulatory networks provide a basis for further study of the function of CDR1as in muscle development and other life activities and processes.

In general, these results show that circRNA is a key factor that cannot be ignored in the process of muscle development; it can

compete for endogenous miRNAs to form a circRNA-miRNA complex and further relieve the inhibitory effect of miRNAs on mRNA. However, the interaction between circRNAs and endogenous miRNAs needs to be further verified because, in some special cases, the number of miRNA sites that bind to a specific circRNA is limited, and these sites are specific in different species or tissues, so the ceRNA regulatory network is not the whole content of muscle development.

CircRNA Regulates Muscle Development in Livestock Animals at the Transcriptional Level

Because there is no biological original to guide coding protein in the special structure of circRNA, circRNA is considered to be a non-coding RNA in most cases. However, recently, researchers are increasingly finding that a very small number of circRNAs have the ability to encode proteins (Li and Lytton, 1999; Legnini et al., 2017; Yang et al., 2018; Liang et al., 2019). Shen et al. (2019) point out that circRNA41, circRNA69, and circRNA153 differentially expressed in porcine skeletal muscle during oxidation and glycolysis are transcribed from MyH1 (encoding MyHC-2X protein), MyH7 (encoding MyHC- β protein), and MyH2 (encoding MyHC-2A protein) genes, respectively. These results indicate that circRNAs play a unique role in regulating the heterogeneity of muscle fiber types, but at present, the research in this area has not been reported. In addition, previous studies point out that the protein encoded by circRNAs depends on the ORF on the sequence. However, the circRNAs that actually have the function of coding protein should have many necessary conditions at the same time, such as the internal ribosome entry sites (IRES), the biological element activated by translation, and the biological element for detecting protein products (Wang and Wang, 2015; Legnini et al., 2017; Pamudurti et al., 2017). A typical example is circRNA ZNF609, which is one of the earliest endogenous circRNAs that can be translated into a protein driven by RES and regulating myogenesis. Circ-ZNF609 itself does not possess the factors required for cap structure and polyadenosine transcription translation, but it can be involved in muscle cell development by initiating translation under cis-acting elements through a mechanism independent of the cap structure in response to different cellular stresses (Legnini et al., 2017). In addition, a recent study finds that circSamd4 is related to Pura and PURB during muscle development; as myogenic inhibitors, Pura and PURB can inhibit the transcription of the myosin heavy chain (MHC) protein family. Silencing CircSamd4 enhances the binding of PUR protein to the MHC promoter, and overexpression of circSamd4 interferes with the binding of pur protein to the MHC promoter, indicating that circSamd4 could bind to PUR protein and prevent it from interacting with DNA. When using mutant circSamd4 without a PUR binding site, these effects were canceled. In other words, the binding of PUR protein to circSamd4 can promote muscle development by reducing MHC transcription (Pandey et al., 2020). Song et al. (2020) identify 197 differentially expressed circRNAs in the gastrocnemius of Duchenne muscular dystrophy (DMD)

mice and predicted their protein coding ability according to the Nmurine 6-methyladenosine motif and ORF of circRNAs. Among them, 189 circRNAs were predicted to have protein coding potential, and 98 circRNAs may be translated into peptides containing 150 or more amino acids, indicating that circRNAs may play a key role in the pathophysiological mechanism of DMD. CircFAM188B contains an ORF that can be translated as circFAM188B-103aa during the skeletal muscle development of broilers to promote the proliferation of chicken SMSC (Yin et al., 2020). To sum up, although researchers have gradually deepened their understanding of circRNAs, there are still few studies on circRNAs regulating muscle development by encoding proteins.

CircRNAs Directly Regulate Muscle Development Through mRNA

As we mentioned earlier, muscle development in livestock is a complex physiological process, and different kinds of circRNAs play different roles in muscle development. Ling et al. (2020) find that cluster16 circRNAs are highly expressed in the early and late stages of muscle development of Anhui white goat (AWG) embryo and are directly involved in the Wnt signal pathway, AMPK signal pathway, and so on. It can be seen that circRNAs can directly regulate the muscle development of livestock. In addition, circQKI (as well as QKI mRNA) depletion is demonstrated to have a negative effect on myoblast differentiation, indicating that both the circRNA and its linear counterpart cooperate in this process. By contrast, although BNC2 mRNA depletion causes an increase in myotube formation, knockdown of its circular counterpart has no effect on differentiation. Interestingly, circBNC2 expression during myoblast differentiation increases at the expense of the corresponding mRNA, suggesting that circBNC2 could contrast the expression of the anti-differentiative BNC2 mRNA (Legnini et al., 2017). Similarly, circEch1 is the most different circRNA in buffalo and beef muscle *in vitro*. The results of *in vitro* experiments show that the overexpression of circEch1 inhibits the proliferation of bovine myoblasts but promotes differentiation; *in vivo* tests show that circEch1 stimulates skeletal muscle regeneration. In general, circEch1 induces myoblast differentiation and skeletal muscle regeneration (Huang et al., 2021). Although there are few studies in this area at present, we believe that with the development of research technology, the relationship between circRNAs and mRNA related to livestock muscle development will become more and more clear.

CONCLUDING REMARKS AND PERSPECTIVES

CircRNAs are a new regulator of muscle development. However, at present, the functional annotation of circRNAs is mainly to predict and analyze source coding genes or possible miRNA binding sites. In this review, we further discuss the relationship between circRNAs and muscle development of livestock by reviewing the discovery, classification, formation, characteristics,

biological function, and degradation pathway of circRNAs. It is obvious that most studies focus on circRNAs guiding muscle development through the ceRNA co-expression regulation network. In addition, some studies confirm that circRNAs can regulate muscle development in livestock in transcription and translation, and only a few studies show that circRNAs can directly regulate muscle development in livestock. There is no doubt that our point of view plays a directional role in the study of muscle development in domestic animals; at the same time, we believe that there are other ways for circRNAs to regulate muscle development in livestock.

Up to now, the study of intracellular circRNAs made people further realize the complexity of eukaryotic gene expression regulation. In-depth study of the formation and types of circRNAs and the mechanism of action with target genes and exploring its biological function is of great significance for understanding the growth and development of organisms and disease treatment. Due to the variety of circRNAs, the diversity of action modes, and the constraints of research methods, people still need some time to clarify these genes and their regulatory mechanisms. However, as a kind of non-coding RNA discovered in the post-genome era, it effectively enriches the research model of gene expression regulation. We think that, with the rapid development of modern molecular biology technology, new generation sequencing technology, and bioinformation analysis technology, more and more new circRNAs will be discovered in the future, and people will study its function more and more deeply.

AUTHOR CONTRIBUTIONS

ZY and TH: conceptualization and writing-original draft and supervising, reviewing, and editing. ZY, TH, and QC: editing. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the National Key Research and Development Plan of China (Grant No. 2018YFD0501004), the National Natural Science Foundation of China (Grant No. 31902167), the Chongqing Natural Science Foundation (Basic Research and Frontier Exploration Special Project) General Project (Grant No. cstc2019jcyj-msxmX0524), the National Science Foundation for Post-doctoral Scientists of China (Grant No. 2018M640895), the Special Funding for Postdoctoral Research Projects in Chongqing (Grant No. XMT 2081061), and Excellent Youth Exchange Program of China Association for Science and Technology in 2019 [2019(293)-98].

ACKNOWLEDGMENTS

We are very grateful to Prof. Guozhong Dong of Southwest University and Prof. De Wu of Sichuan Agricultural University for their meticulous guidance in the process of writing the article.

REFERENCES

- Abdelmohsen, K., Panda, A. C., De, S., Grammatikakis, I., Kim, J., Ding, J., et al. (2015). Circular RNAs in monkey muscle: age-dependent changes. *Aging (Albany N.Y.)* 7, 903–910. doi: 10.18632/aging.100834
- Berger, J., Hall, T. E., and Currie, P. D. (2015). Novel transgenic lines to label sarcolemma and myofibrils of the musculature. *Zebrafish* 12, 124–125. doi: 10.1089/zeb.2014.1065
- Burd, C. E., Jeck, W. R., Liu, Y., Sanoff, H. K., Wang, Z., and Sharpless, N. E. (2010). Expression of linear and novel circular forms of an INK4/ARF-associated non-coding RNA correlates with atherosclerosis risk. *Plos Genet.* 6:e1001233–e1001247. doi: 10.1371/journal.pgen.1001233
- Cao, Y., You, S., Yao, Y., Liu, Z. J., Hazi, W., Li, C. Y., et al. (2018). Expression profiles of circular RNAs in sheep skeletal muscle. *Asian Austr. J. Anim.* 31, 1550–1557.
- Chen, B., You, W., Wang, Y., and Shan, T. (2020). The regulatory role of Myomaker and Myomixer-Myomerger-Minion in muscle development and regeneration. *Cell. Mol. Life Sci.* 77, 1551–1569. doi: 10.1007/s00018-019-03341-9
- Chen, L. L., and Yang, L. (2015). Regulation of circRNA biogenesis. *RNA Biol.* 12, 381–388. doi: 10.1080/15476286.2015.1020271
- Chen, P. R., Suh, Y., Shin, S., Woodfint, R. M., Hwang, S., and Lee, K. (2019). Exogenous expression of an alternative splicing variant of Myostatin prompts leg muscle fiber hyperplasia in Japanese quail. *Int. J. Mol. Sci.* 20, 4617–4633. doi: 10.3390/ijms20184617
- Chen, S., Huang, V., Xu, X., Livingstone, J., Soares, F., Jeon, J., et al. (2019). Widespread and functional RNA circularization in localized prostate cancer. *Cell* 176, 831–843. doi: 10.1016/j.cell.2019.01.025
- Christ, B., and Brand, S. B. (2004). Limb muscle development. *Int. J. Dev. Biol.* 46, 905–914. doi: 10.1079/IJVP.2001258
- Cocquerelle, C., Mascres, B., Hétiuin, D., and Bailleul, B. (1993). Mis-splicing yields circular RNA molecules. *FASEB J.* 7, 155–160. doi: 10.1096/fasebj.7.1.7678559
- Cook, J. R., MacIntyre, D. A., Samara, E., Kim, S. H., Singh, N., Johnson, M. R., et al. (2015). Exogenous oxytocin modulates human myometrial microRNAs. *Am. J. Obstet. Gynecol.* 213, 65–65. doi: 10.1016/j.ajog.2015.03.015
- Danan, M., Schwartz, S., Edelheit, S., and Sorek, R. (2012). Transcriptome-wide discovery of circular RNAs in Archaea. *Nucleic Acids Res.* 40, 3131–3142. doi: 10.1093/nar/gkr1009
- Das, A., Das, A., Das, D., Abdelmohsen, K., and Panda, A. C. (2020). Circular RNAs in myogenesis. *BBA Gene Regul. Mech.* 1863, 194372–194378. doi: 10.1016/j.bbarm.2019.02.011
- Das, A., Rout, P. K., Gorospe, M., and Panda, A. C. (2019a). Past, present, and future of circRNAs. *EMBO J.* 38, e100836–e100848. doi: 10.3390/ijms20163988
- Das, A., Rout, P. K., Gorospe, M., and Panda, A. C. (2019b). Rolling circle cDNA synthesis uncovers circular RNA splice variants. *Int. J. Mol. Sci.* 20, 3988–3999.
- Doynova, M. D., Markworth, J. F., Cameron-Smith, D., Vickers, M. H., and O'Sullivan, J. M. (2017). Linkages between changes in the 3D organization of the genome and transcription during myotube differentiation in vitro. *Skelet. Muscle* 7, 5–14. doi: 10.1186/s13395-017-0122-1
- Du, W. W., Yang, W., Li, X., Awan, F. M., Yang, Z., Fang, L., et al. (2018). A circular RNA circ-DNMT1 enhances breast cancer progression by activating autophagy. *Oncogene* 37, 5829–5842. doi: 10.1038/s41388-018-0369-y
- Ebbesen, K. K., Hansen, T. B., and Kjems, J. (2017). Insights into circular RNA biology. *RNA Biol.* 14, 1035–1045. doi: 10.1080/15476286.2016.1271524
- Fischer, J. W., Busa, V. F., Shao, Y., and Leung, A. (2020). Structure-mediated RNA Decay by UPF1 and G3BP1. *Mol. Cell* 78, 70–84. doi: 10.1016/j.molcel.2020.01.021
- Geng, H. H., Li, R., Su, Y. M., Xiao, J., Pan, M., Cai, X. X., et al. (2016). The circular RNA Cdr1as promotes myocardial infarction by mediating the regulation of miR-7a on its target genes expression. *PLoS One* 11:e0151753–e0151769. doi: 10.1371/journal.pone.0151753
- Granados-Riveron, J. T., and Aquino-Jarquín, G. (2016). The complexity of the translation ability of circRNAs. *BBA Gene Regul. Mech.* 1859, 1245–1251. doi: 10.1016/j.bbarm.2016.07.009
- Greco, S., De Simone, M., Colussi, C., Zaccagnini, G., Fasanaro, P., Pescatori, M., et al. (2009). Common micro-RNA signature in skeletal muscle damage and regeneration induced by Duchenne muscular dystrophy and acute ischemia. *FASEB J.* 23, 3335–3346. doi: 10.1096/fj.08-128579
- Hansen, T. B., Wiklund, E. D., Bramsen, J. B., Villadsen, S. B., Statham, A. L., Clark, S. J., et al. (2011). miRNA-dependent gene silencing involving Ago2-mediated cleavage of a circular antisense RNA. *EMBO J.* 30, 4414–4422. doi: 10.1038/emboj.2011.359
- Hernández-Hernández, J. M., García-González, E. G., Brun, C. E., and Rudnicki, M. A. (2017). The myogenic regulatory factors, determinants of muscle development, cell identity and regeneration. *Semin. Cell Dev. Biol.* 72, 10–18. doi: 10.1016/j.semcdb.2017.11.010
- Holdt, L. M., Stahringer, A., Sass, K., Pichler, G., Kulak, N. A., Wilfert, W., et al. (2016). Circular non-coding RNA ANRIL modulates ribosomal RNA maturation and atherosclerosis in humans. *Nat. Commun.* 7:12429. doi: 10.1038/ncomms12429
- Hsu, M. T., and Coca-Prados, M. (1979). Electron microscopic evidence for the circular form of RNA in the cytoplasm of eukaryotic cells. *Nature* 280, 339–340. doi: 10.1038/280339a0
- Huang, K. W., Chen, M. J., Zhong, D. D., Luo, X., Feng, T., Song, M. M., et al. (2021). Circular RNA profiling reveals an abundant circEch1 that promotes myogenesis and differentiation of bovine skeletal muscle. *J. Agric. Food Chem.* 69, 592–601. doi: 10.1021/acs.jafc.0c06400
- Humphreys, D. T., Fossat, N., Demuth, M., Tam, P., and Ho, J. (2019). Ularcirc: visualization and enhanced analysis of circular RNAs via back and canonical forward splicing. *Nucleic Acids Res.* 47, e123–e139. doi: 10.1093/nar/gkz718
- Irie, T., Shum, R., Deni, I., Hunkele, A., Le Rouzic, V., Xu, J., et al. (2019). Identification of abundant and evolutionarily conserved opioid receptor circular RNAs in the nervous system modulated by morphine. *Mol. Pharmacol.* 96, 247–258. doi: 10.1124/mol.118.113977
- Janssen, I., Heymsfield, S. B., Wang, Z. M., and Ross, R. (2000). Skeletal muscle mass and distribution in 468 men and women aged 18–88 yr. *J. Appl. Physiol.* 89, 81–88. doi: 10.1152/jappl
- Kos, A., Dijkema, R., Arnberg, A. C., van der Meide, P. H., and Schellekens, H. (1986). The hepatitis delta (δ) virus possesses a circular RNA. *Nature* 323, 558–560. doi: 10.1038/323558a0
- Kotake, Y., Nakagawa, T., Kitagawa, K., Suzuki, S., Liu, N., Kitagawa, M., et al. (2011). Long non-coding RNA ANRIL is required for the PRC2 recruitment to and silencing of p15 INK4B tumor suppressor gene. *Oncogene* 30, 1956–1962. doi: 10.1038/onc.2010.568
- Kyei, B., Li, L., Yang, L., Zhan, S., and Zhang, H. (2020). CDR1as/miRNAs-related regulatory mechanisms in muscle development and diseases. *Gene* 730, 144315–144322. doi: 10.1016/j.gene.2019.144315
- Lasda, E., and Parker, R. (2014). Circular RNAs: diversity of form and function. *RNA* 20, 1829–1842. doi: 10.1261/rna.047126.114
- Legnini, I., Di Timoteo, G., Rossi, F., Morlando, M., Briganti, F., Sthandier, O., et al. (2017). Circ-ZNF609 is a circular RNA that can be translated and functions in myogenesis. *Mol. Cell* 66, 22–37. doi: 10.1016/j.molcel.2017.02.017
- Li, B., Yin, D., Li, P., Zhang, Z., Zhang, X., Li, H., et al. (2020). Profiling and functional analysis of circular RNAs in porcine fast and slow muscles. *Front. Cell Dev. Biol.* 8:322–324. doi: 10.3389/fcell.2020.00322
- Li, C., Li, X., Yao, Y., Ma, Q., Ni, W., Zhang, X., et al. (2017). Genome-wide analysis of circular RNAs in prenatal and postnatal muscle of sheep. *Oncotarget* 8, 97165–97177. doi: 10.18632/oncotarget.21835
- Li, D., Li, Z., Yang, Y., Zeng, X., Li, Y., Du, X., et al. (2020). Circular RNAs as biomarkers and therapeutic targets in environmental chemical exposure-related diseases. *Environ. Res.* 180, 108825–108833. doi: 10.1016/j.envres.2019.108825
- Li, H., Wei, X., Yang, J., Dong, D., Hao, D., Huang, Y., et al. (2018a). circFGFR4 promotes differentiation of myoblasts by binding miR-107 to relieve its inhibition of Wnt3a. *Mol. Ther. Nucleic Acids* 11, 272–283. doi: 10.1016/j.omtn.2018.02.012
- Li, H., Yang, J., Wei, X., Song, C., Dong, D., Huang, Y., et al. (2018b). CircFUT10 reduces proliferation and facilitates differentiation of myoblasts by sponging miR-133a. *J. Cell Physiol.* 233, 4643–4651. doi: 10.1002/jcp.26230
- Li, L., Chen, Y., Nie, L., Ding, X., Zhang, X., Zhao, W., et al. (2019). MyoD-induced circular RNA CDR1as promotes myogenic differentiation of skeletal muscle satellite cells. *BBA Gene Regul. Mech.* 1862, 807–821. doi: 10.1016/j.bbarm.2019.07.001
- Li, X. F., and Lytton, J. (1999). A circularized sodium-calcium exchanger exon 2 transcript. *J. Biol. Chem.* 274, 8153–8160.

- Li, X., Liu, S., Zhang, L., Issaian, A., Hill, R. C., Espinosa, S., et al. (2019). A unified mechanism for intron and exon definition and back-splicing. *Nature* 573, 375–380. doi: 10.1038/s41586-019-1523-6
- Li, X., Xie, S., Qian, L., Cai, C., Bi, H., and Cui, W. (2020). Identification of genes related to skeletal muscle growth and development by integrated analysis of transcriptome and proteome in myostatin-edited Meishan pigs. *J. Proteomics* 213, 103628–103639. doi: 10.1016/j.jprot.2019.103628
- Liang, D., Tatomer, D. C., Luo, Z., Wu, H., Yang, L., Chen, L. L., et al. (2017). The output of protein-coding genes shifts to circular RNAs when the pre-mRNA processing machinery is limiting. *Mol. Cell* 68, 940–954. doi: 10.1016/j.molcel.2017.10.034
- Liang, G., Yang, Y., Niu, G., Tang, Z., and Li, K. (2017a). Circular RNA profiling reveals an abundant circLMO7 that regulates myoblasts differentiation and survival by sponging miR-378a-3p. *Cell Death Dis.* 8, e3153–e3165. doi: 10.1093/dnares/dsx022
- Liang, G., Yang, Y., Niu, G., Tang, Z., and Li, K. (2017b). Genome-wide profiling of *Sus scrofa* circular RNAs across nine organs and three developmental stages. *DNA Res.* 24, 523–535.
- Liang, W. C., Wong, C. W., Liang, P. P., Shi, M., Cao, Y., Rao, S. T., et al. (2019). Translation of the circular RNA circ β -catenin promotes liver cancer cell growth through activation of the Wnt pathway. *Genome Biol.* 20:84. doi: 10.1186/s13059-019-1685-4
- Ling, Y., Zheng, Q., Zhu, L., Xu, L., Sui, M., Zhang, Y., et al. (2020). Trend analysis of the role of circular RNA in goat skeletal muscle development. *BMC Genomics* 21:220. doi: 10.1186/s12864-020-6649-2
- Liu, B., Song, F., Yang, Q., Zhou, Y., Shao, C., Shen, Y., et al. (2019). Characterization of tissue-specific biomarkers with the expression of circRNAs in forensically relevant body fluids. *Int. J. Legal. Med.* 133, 1321–1331. doi: 10.1007/s00414-019-02027-y
- Liu, C. X., Li, X., Nan, F., Jiang, S., Gao, X., Guo, S. K., et al. (2019). Structure and degradation of circular RNAs regulate PKR activation in innate immunity. *Cell* 177, 865–880. doi: 10.1016/j.cell.2019.03.046
- Liu, X., Hu, Z., Zhou, J., Tian, C., Tian, G., He, M., et al. (2020a). Interior circular RNA. *RNA Biol.* 17, 87–97. doi: 10.1080/15476286.2019.1669391
- Liu, X., Wang, X., Li, J., Hu, S., Deng, Y., Yin, H., et al. (2020b). The identification of mecciRNAs and their roles in mitochondrial entry of proteins. *Sci. China Life Sci.* 63, 1429–1449. doi: 10.1007/s11427-020-1631-9
- McPherron, A. C., and Lee, S. J. (2002). Suppression of body fat accumulation in myostatin-deficient mice. *J. Clin. Invest.* 109, 595–601. doi: 10.1172/JCI13562
- Memczak, S., Jens, M., Elefsinioti, A., Torti, F., Krueger, J., Rybak, A., et al. (2013). Circular RNAs are a large class of animal RNAs with regulatory potency. *Nature* 495, 333–338. doi: 10.1038/nature11928
- Molkentin, J. D., and Olson, E. N. (1996). Defining the regulatory networks for muscle development. *Curr. Opin. Genet. Dev.* 6, 445–453. doi: 10.1016/s0959-437x(96)80066-9
- Naqvi, A. S., Asnani, M., Black, K. L., Hayer, K. E., Taylor, D., and Thomas-Tikhonenko, A. (2019). The role of SRSF3 splicing factor in generating circular RNAs. *bioRxiv* [Preprint] 799700. doi: 10.1101/799700
- Ng, P. C., Chan, K. Y., Leung, K. T., Tam, Y. H., Ma, T. P., Lam, H. S., et al. (2015). Comparative MiRNA expression profiles and molecular networks in human small bowel tissues of necrotizing enterocolitis and spontaneous intestinal perforation. *PLoS One* 10:e0135737–e0135753. doi: 10.1371/journal.pone.0135737
- Ouyang, H., Chen, X., Wang, Z., Yu, J., Jia, X., Li, Z., et al. (2018). Circular RNAs are abundant and dynamically expressed during embryonic muscle development in chickens. *DNA Res.* 25, 71–86. doi: 10.1093/dnares/dsx039
- Pagliarini, V., Jolly, A., Bielli, P., Di Rosa, V., De la Grange, P., and Sette, C. (2020). Sam68 binds Alu-rich introns in SMN and promotes pre-mRNA circularization. *Nucleic Acids Res.* 48, 633–645. doi: 10.1093/nar/gkz1117
- Pamudurti, N. R., Bartok, O., Jens, M., Ashwal-Fluss, R., Stottmeister, C., Ruhe, L., et al. (2017). Translation of circRNAs. *Mol. Cell* 66, 9–21. doi: 10.1016/j.molcel.2017.02.021
- Pandey, P. R., Yang, J. H., Tsitsipatis, D., Panda, A. C., and Gorospe, M. (2020). circSamd4 represses myogenic transcriptional activity of PUR proteins. *Nucleic Acids Res.* 48, 3789–3805. doi: 10.1093/nar/gkaa035
- Park, O. H., Ha, H., Lee, Y., Boo, S. H., Kwon, D. H., Song, H. K., et al. (2019). Endoribonucleolytic cleavage of m6A-containing RNAs by RNase P/MRP complex. *Mol. Cell* 74, 494–507. doi: 10.1016/j.molcel.2019.02.034
- Patop, I. L., Wust, S., and Kadener, S. (2019). Past, present, and future of circRNAs. *Embo. J.* 38:e100836. doi: 10.15252/embj.2018100836
- Peng, S., Song, C., Li, H., Cao, X., Ma, Y., Wang, X., et al. (2019). Circular RNA SNX29 sponges miR-744 to regulate proliferation and differentiation of myoblasts by activating the Wnt5a/Ca2+ signaling pathway. *Mol. Ther. Nucleic Acids* 16, 481–493. doi: 10.1016/j.omtn.2019.03.009
- Qian, D. Y., Yan, G. B., Bai, B., Chen, Y., Zhang, S. J., Yao, Y. C., et al. (2017). Differential circRNA expression profiles during the BMP2-induced osteogenic differentiation of MC3T3-E1 cells. *Biomed. Pharmacother.* 90, 492–499. doi: 10.1016/j.biopha.2017.03.051
- Qu, S., Yang, X., Li, X., Wang, J., Gao, Y., and Shang, R. (2015). Circular RNA: a new star of noncoding RNAs. *Cancer Lett.* 365, 141–148. doi: 10.1016/j.canlet.2015.06.003
- Rong, D., Sun, H., Li, Z., Liu, S., Dong, C., Fu, K., et al. (2017). An emerging function of circRNA-miRNAs-mRNA axis in human diseases. *Oncotarget* 8, 73271–73281. doi: 10.18632/oncotarget.19154
- Salzman, J., Gawad, C., Wang, P. L., Lacayo, N., and Brown, P. O. (2012). Circular RNAs are the predominant transcript isoform from hundreds of human genes in diverse cell types. *PLoS One* 7:e30733–e30744. doi: 10.1371/journal.pone.0030733
- Sang, M., Meng, L., Sang, Y., Liu, S., Ding, P., Ju, Y., et al. (2018). Circular RNA ciRS-7 accelerates ESCC progression through acting as a miR-876-5p sponge to enhance MAGE-A family expression. *Cancer Lett.* 426, 37–46. doi: 10.1016/j.canlet.2018.03.049
- Sanger, H. L., Klotz, G., Riesner, D., Gross, H. J., and Kleinschmidt, A. K. (1976). Viroids are single-stranded covalently closed circular RNA molecules existing as highly base-paired rod-like structures. *Proc. Natl. Acad. Sci. U.S.A.* 73, 3852–3856. doi: 10.1073/pnas.73.11.3852
- Shen, L., Gan, M., Tang, Q., Tang, G., Jiang, Y., Li, M., et al. (2019). Comprehensive analysis of lncRNAs and circRNAs reveals the metabolic specialization in oxidative and glycolytic skeletal muscles. *Int. J. Mol. Sci.* 20, 2855–2872. doi: 10.3390/ijms20122855
- Silva, W. J., Graça, F. A., Cruz, A., Silvestre, J. G., Labeit, S., Miyabara, E. H., et al. (2019). miR-29c improves skeletal muscle mass and function throughout myocyte proliferation and differentiation and by repressing atrophy-related genes. *Acta Physiol.* 226, e13278–e13296. doi: 10.1111/apha.13278
- Song, Z. B., Liu, Y. M., Fang, X. B., Xie, M. S., Ma, Z. Y., Zhong, Z. G., et al. (2020). Comprehensive analysis of the expression profile of circRNAs and their predicted protein-coding ability in the muscle of mdx mice. *Funct. Integr. Genomic* 20, 397–407. doi: 10.1007/s10142-019-00724-w
- Sun, J., Xie, M., Huang, Z., Li, H., Chen, T., Sun, R., et al. (2017). Integrated analysis of non-coding RNA and mRNA expression profiles of 2 pig breeds differing in muscle traits. *J. Anim. Sci.* 95, 1092–1103. doi: 10.2527/jas.2016.0867
- Suzuki, H., and Tsukahara, T. (2014). A view of pre-mRNA splicing from RNase R resistant RNAs. *Int. J. Mol. Sci.* 15, 9331–9342. doi: 10.3390/ijms15069331
- Suzuki, H., Zuo, Y., Wang, J., Zhang, M. Q., Malhotra, A., and Mayeda, A. (2006). Characterization of RNase R-digested cellular RNA source that consists of lariat and circular RNAs from pre-mRNA splicing. *Nucleic Acids Res.* 34, e63–e69. doi: 10.1093/nar/gkl151
- Tan, S., Gou, Q., Pu, W., Guo, C., Yang, Y., Wu, K., et al. (2018). Circular RNA F-circEA produced from EML4-ALK fusion gene as a novel liquid biopsy biomarker for non-small cell lung cancer. *Cell Res.* 28, 693–695. doi: 10.1038/s41422-018-0033-7
- Tang, C., Xie, Y., Yu, T., Liu, N., Wang, Z., Woolsey, R. J., et al. (2020). m6A-dependent biogenesis of circular RNAs in male germ cells. *Cell Res.* 30, 211–228. doi: 10.1038/s41422-020-0279-8
- Tuck, A. C., Rankova, A., Arpat, A. B., Liechti, L. A., Hess, D., Iesmantavicius, V., et al. (2020). Mammalian RNA decay pathways are highly specialized and widely linked to translation. *Mol. Cell* 77, 1222–1236. doi: 10.1016/j.molcel.2020.01.007
- von Maltzahn, J., Chang, N. C., Bentzinger, C. F., and Rudnicki, M. A. (2012). Wnt signaling in myogenesis. *Trends Cell Biol.* 22, 602–609. doi: 10.1016/j.tcb.2012.07.008
- Wan, H., Zhu, J., Su, G., Liu, Y., Hua, L., Hu, L., et al. (2016). Dietary supplementation with β -hydroxy- β -methylbutyrate calcium during the early

- postnatal period accelerates skeletal muscle fibre growth and maturity in intra-uterine growth-retarded and normal-birth-weight piglets. *Br. J. Nutr.* 115, 1360–1369. doi: 10.1017/S0007114516000465
- Wang, F., Nazarali, A. J., and Ji, S. (2016). Circular RNAs as potential biomarkers for cancer diagnosis and therapy. *Am. J. Cancer Res.* 6, 1167–1176.
- Wang, X., Cao, X., Dong, D., Shen, X., Cheng, J., Jiang, R., et al. (2019). Circular RNA TTN acts as a miR-432 sponge to facilitate proliferation and differentiation of myoblasts via the IGF2/PI3K/AKT signaling pathway. *Mol. Ther. Nucleic Acids* 18, 966–980. doi: 10.1016/j.omtn.2019.10.019
- Wang, Y., and Wang, Z. (2015). Efficient backsplicing produces translatable circular mRNAs. *RNA Publ RNA Soc.* 21, 172–179. doi: 10.1261/rna.048272.114
- Westholm, J. O., Miura, P., Olson, S., Shenker, S., Joseph, B., Sanfilippo, P., et al. (2014). Genome-wide analysis of drosophila circular RNAs reveals their structural and sequence properties and age-dependent neural accumulation. *Cell Rep.* 9, 966–1980. doi: 10.1016/j.celrep.2014.10.062
- Wei, X., Li, H., Yang, J., Hao, D., Dong, D., Huang, Y., et al. (2017). Circular RNA profiling reveals an abundant circLMO7 that regulates myoblasts differentiation and survival by sponging miR-378a-3p. *Cell Death Dis.* 8, e3153–e3165. doi: 10.1038/cddis.2017.541
- Wei, X., Li, H., Zhang, B., Li, C., Dong, D., Lan, X., et al. (2016). miR-378a-3p promotes differentiation and inhibits proliferation of myoblasts by targeting HDAC4 in skeletal muscle development. *RNA Biol.* 13, 1300–1309. doi: 10.1080/15476286.2016.1239008
- Wu, K., Liao, X., Gong, Y., He, J., Zhou, J. K., Tan, S., et al. (2019). Circular RNA F-circSR derived from SLC34A2-ROS1 fusion gene promotes cell migration in non-small cell lung cancer. *Mol. Cancer.* 18, 98–103. doi: 10.1186/s12943-019-1028-9
- Yan, X. M., Zhang, Z., Meng, Y., Li, H. B., Gao, L., Luo, D., et al. (2020). Genome-wide identification and analysis of circular RNAs differentially expressed in the longissimus dorsi between Kazakh cattle and Xinjiang brown cattle. *PEER J.* 8, 8646–8662. doi: 10.7717/peerj.8646
- Yang, Y. B., Gao, X. Y., Zhang, M. L., Yan, S., Sun, C. J., Xiao, F. Z., et al. (2018). Novel role of FBXW7 circular RNA in repressing Glioma Tumorigenesis. *J. Natl. Cancer Inst.* 110, 304–315. doi: 10.1093/jnci/djx166
- Yang, Z. G., Awan, F. M., Du, W. W., Zeng, Y., Lyu, J., Wu, D., et al. (2017). The circular RNA interacts with STAT3, increasing its nuclear translocation and wound repair by modulating Dnm3a and miR-17 function. *Mol. Ther.* 25, 2062–2074. doi: 10.1016/j.ymthe.2017.05.022
- Yin, H. D., Shen, X. X., Zhao, J., Cao, X. A., He, H. R., Han, S. S., et al. (2020). Circular RNA CircFAM188B encodes a protein that regulates proliferation and differentiation of chicken skeletal muscle satellite cells. *Front. Cell Dev. Biol.* 8:522588. doi: 10.3389/fcell.2020.522588
- Yoshimoto, R., Rahimi, K., Hansen, T. B., Kjems, J., and Mayeda, A. (2020). Biosynthesis of circular RNA ciRS-7/CDR1as is mediated by mammalian-wide interspersed repeats (MIRs). *Iscience* 23, 101345–101359. doi: 10.1016/j.isci.2020.101345
- Yue, B., Wang, J., Ru, W., Wu, J., Cao, X., Yang, H., et al. (2020). The circular RNA circHUWE1 sponges the miR-29b-AKT3 axis to regulate myoblast development. *Mol. Ther. Nucleic Acids* 19, 1086–1097. doi: 10.1016/j.omtn.2019.12.039
- Zeng, K., Chen, X., Xu, M., Liu, X., Hu, X., Xu, T., et al. (2018). CircHIPK3 promotes colorectal cancer growth and metastasis by sponging miR-7. *Cell Death Dis.* 9:417. doi: 10.1038/s41419-018-0454-8
- Zhang, P., Xu, H., Li, R., Wu, W., Chao, Z., Li, C., et al. (2018). Assessment of myoblast circular RNA dynamics and its correlation with miRNA during myogenic differentiation. *Biochem. Cell Biol.* 99, 211–218. doi: 10.1016/j.biocel.2018.04.016
- Zhang, P. P., Chao, Z., Zhang, R., Ding, R. Q., Wang, Y. L., Wu, W., et al. (2019). Circular RNA Regulation of Myogenesis. *Cells* 8, 885–897. doi: 10.3390/cells8080885
- Zhang, X. O., Wang, H. B., Zhang, Y., Lu, X., Chen, L. L., and Yang, L. (2014). Complementary sequence-mediated exon circularization. *Cell* 159, 134–147. doi: 10.1016/j.cell.2014.09.001
- Zhang, Y., Zhang, X. O., Chen, T., Xiang, J. F., Yin, Q. F., Xing, Y. H., et al. (2013). Circular intronic long noncoding RNAs. *Mol. Cell* 51, 792–806. doi: 10.1016/j.molcel.2013.08.017
- Zhao, J., Wu, J., Xu, T., Yang, Q., He, J., and Song, X. (2018). IRESfinder: identifying RNA internal ribosome entry site in eukaryotic cell using framed k-mer features. *J. Genet. Genomics* 45, 403–406. doi: 10.1016/j.jgg.2018.07.006
- Zhao, W., Cui, Y., Liu, L., Qi, X., Liu, J., Ma, S., et al. (2020). Splicing factor derived circular RNA circUHRF1 accelerates oral squamous cell carcinoma tumorigenesis via feedback loop. *Cell Death Differ.* 27, 919–933. doi: 10.1038/s41418-019-0423-5

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 Yang, He 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.



Deciphering Epitranscriptome: Modification of mRNA Bases Provides a New Perspective for Post-transcriptional Regulation of Gene Expression

Suresh Kumar^{1*} and Trilochan Mohapatra²

¹ Division of Biochemistry, ICAR-Indian Agricultural Research Institute, New Delhi, India, ² Indian Council of Agricultural Research, New Delhi, India

OPEN ACCESS

Edited by:

Jia Meng,
Xi'an Jiaotong-Liverpool University,
China

Reviewed by:

Xudong Zhang,
University of California, Riverside,
United States
Sandra Blanco,
University of Salamanca, Spain

*Correspondence:

Suresh Kumar
sureshkumar@iari.res.in;
sureshkumar3_in@yahoo.co.uk
orcid.org/0000-0002-7127-3079

Specialty section:

This article was submitted to
Epigenomics and Epigenetics,
a section of the journal
Frontiers in Cell and Developmental
Biology

Received: 11 November 2020

Accepted: 22 February 2021

Published: 16 March 2021

Citation:

Kumar S and Mohapatra T (2021)
Deciphering Epitranscriptome:
Modification of mRNA Bases Provides
a New Perspective for
Post-transcriptional Regulation of
Gene Expression.
Front. Cell Dev. Biol. 9:628415.
doi: 10.3389/fcell.2021.628415

Gene regulation depends on dynamic and reversibly modifiable biological and chemical information in the epigenome/epitranscriptome. Accumulating evidence suggests that messenger RNAs (mRNAs) are generated in flashing bursts in the cells in a precisely regulated manner. However, the different aspects of the underlying mechanisms are not fully understood. Cellular RNAs are post-transcriptionally modified at the base level, which alters the metabolism of mRNA. The current understanding of epitranscriptome in the animal system is far ahead of that in plants. The accumulating evidence indicates that the epitranscriptomic changes play vital roles in developmental processes and stress responses. Besides being non-genetically encoded, they can be of reversible nature and involved in fine-tuning the expression of gene. However, different aspects of base modifications in mRNAs are far from adequate to assign the molecular basis/functions to the epitranscriptomic changes. Advances in the chemogenetic RNA-labeling and high-throughput next-generation sequencing techniques are enabling functional analysis of the epitranscriptomic modifications to reveal their roles in mRNA biology. Mapping of the common mRNA modifications, including *N*⁶-methyladenosine (m⁶A), and 5-methylcytidine (m⁵C), have enabled the identification of other types of modifications, such as *N*¹-methyladenosine. Methylation of bases in a transcript dynamically regulates the processing, cellular export, translation, and stability of the mRNA; thereby influence the important biological and physiological processes. Here, we summarize the findings in the field of mRNA base modifications with special emphasis on m⁶A, m⁵C, and their roles in growth, development, and stress tolerance, which provide a new perspective for the regulation of gene expression through post-transcriptional modification. This review also addresses some of the scientific and technical issues in epitranscriptomic study, put forward the viewpoints to resolve the issues, and discusses the future perspectives of the research in this area.

Keywords: epitranscriptomics, RNA modification, post-transcriptional regulation, 5-methylcytidine, *N*⁶-methyladenosine, RNA metabolism, mRNA methylation, central dogma

INTRODUCTION

From the genome to proteome, several proficient biological processes regulate cellular growth and functions. Transcription of a gene is a truthful process, as the timing and rate of transcription are subjected to strict regulation, and its accuracy is vital for the vigor and development of the cell (Wang et al., 2018). Because translation of mRNA is a vital process in all living organisms, and assembly of the translational machinery followed by movement along the mRNA consumes ~40% of cellular energy, the process needs to be precisely regulated to conserve energy. The 'Central Dogma of life' describes that genetic information is transformed from DNA to protein through RNA. Both DNA and histone proteins are reversibly modified (epigenetic modifications) to fine-tune the expression of genes/phenotypes (Fu Y. et al., 2014). An analogous process for RNA (epitranscriptomic modification) has been a missing component of the central dogma (**Figure 1**). Reversible biochemical modifications are known now to occur in most of the constituent processes of the central dogma, which dynamically control gene expression. The spectrum of epigenetic base modifications detected so far in DNA is relatively limited (six), about 170 distinct modifications have been identified in RNAs (Boccaletto et al., 2018; Kadumuri and Janga, 2018; Shen et al., 2019; Boo and Kim, 2020; Selmi et al., 2021). RNAs play vital roles in biological systems, not only as structural components [i.e., ribosomal RNAs (rRNAs)], translators [i.e., transfer RNAs (tRNAs)], and messengers (i.e., mRNAs, conveying genetic information to the protein) but also as regulators [i.e., small interfering RNAs (siRNAs), enhancer RNAs (eRNAs)] of several biological processes. The functions of rRNAs, tRNAs, and mRNAs are regulated through co- or post-transcriptional chemical modifications (Boccaletto et al., 2018; Boo and Kim, 2020), the exact role of many of these base modifications remain enigmatic. Although extensive base modifications in rRNAs and tRNAs in terms of the variety/abundance of modifications are well known and have remained undisputed for many decades (Jackman and Alfonzo, 2013), all other classes of RNA are subjected to enzymatic modifications (Xu L. et al., 2017). Several post-transcriptional base modifications in messenger RNA (mRNA) have only recently been identified. Such mRNA base modifications affect different cellular processes like pre-mRNA splicing, mRNA export, translation, and degradation, which shape the cellular transcriptome and proteome. Recent findings indicate that the level of proteins in a cell does not necessarily correspond with the mRNA level (Khan et al., 2013; Wu et al., 2013), which might vary because of various post-transcriptional regulation, including epitranscriptomic modifications affecting mRNA biology. The recent advances in experimental techniques have facilitated the identification of different epitranscriptomic modifications in the coding and untranslated regions (UTRs) of mRNAs (Zhao et al., 2020). While the functions of some of the epitranscriptomic modifications are known, occurrence and function of many other diverse epitranscriptomic modifications are still to be established.

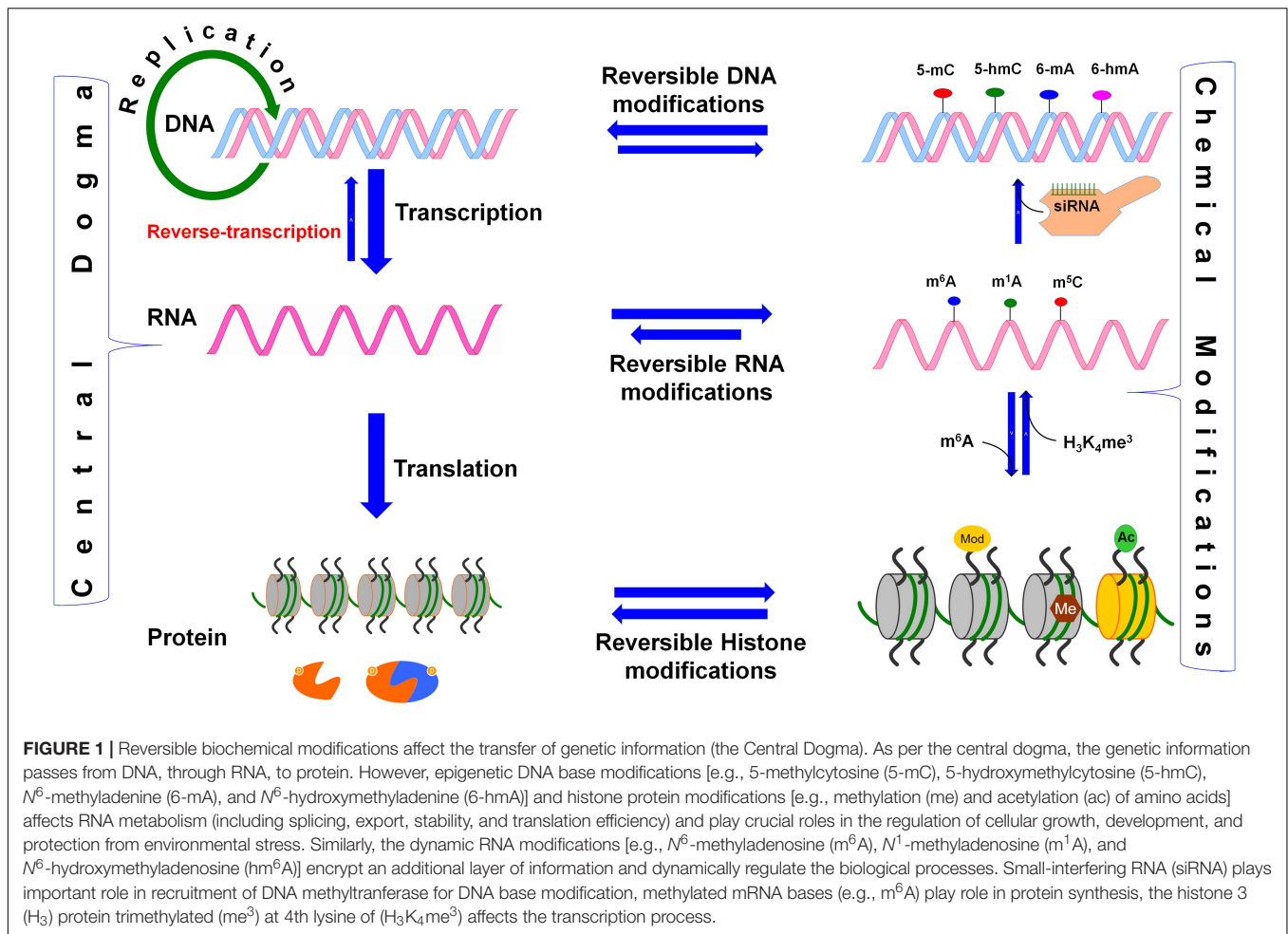
The dynamic and reversible RNA base modifications are catalyzed by distinct enzymes like methyltransferases (writers), and removed by demethylases (erasers). These

modifications are interpreted by a modification-specific binding proteins known as readers. Characterization of writers, readers, and erasers is further advancing our epitranscriptomic understanding of functional genomics. Similar to the epigenetic modifications of DNA bases (Kumar et al., 2018), mRNA base modifications provide another layer of information created by the writers/erasers and interpreted by the readers. Like the reversible nature of DNA base modifications (Wang et al., 2016), some of the mRNA base modifications are known to be reversed by their respective eraser. Although translation process is typically controlled by translation factors and certain non-coding RNAs (ncRNAs), base modifications play equally important role in mRNA metabolism and translation process. Thus, the mRNA base modifications create the epitranscriptomic regulatory machinery that is being elucidated in the animal as well as the plant systems. It is now apparent that mRNA is a dynamic and reversibly modifiable biomolecule (**Figure 2**) that play crucial roles in post-transcriptional regulation of gene expression (Zhao et al., 2017a).

Many of the mRNA base modifications involve attachment of a methyl (CH_3) group at a particular position either on the base [e.g., N^6 -methyladenosine ($m^6\text{A}$), N^1 -methyladenosine ($m^1\text{A}$), 5-methylcytidine ($m^5\text{C}$), 3-methylcytidine ($m^3\text{C}$), N^7 -methylguanosine ($m^7\text{G}$), and 1-methylguanosine ($m^1\text{G}$)], ribose sugar (e.g., 2'-O-methyladenosine), or on both base and sugar [e.g., $N^6,2'$ -O-dimethyladenosine ($m^6\text{Am}$)] (Dominissini et al., 2012; Linder et al., 2015; Dominissini et al., 2016; Molinie et al., 2016; Mauer et al., 2017). Thus, methylation of bases at different position has distinct impact on RNA biology by affecting folding, stability, cellular localization, and/or interaction with other RNAs/proteins (Wu et al., 2016). The $m^6\text{A}$ is one of the most common, reversible epitranscriptomic marks, functionally pertinent in both animal and plant mRNAs (Batista et al., 2014; Shen et al., 2016). Moreover, writers, readers, and erasers for $m^6\text{A}$ are known in animals as well as plants (Bokar et al., 1997; Zhong et al., 2008; Liu et al., 2014; Du et al., 2016; Liu and Pan, 2016; Patil et al., 2016; Roundtree and He, 2016; Martinez-Perez et al., 2017; Zhao et al., 2017b; Arribas-Hernandez et al., 2018; Scutenaire et al., 2018; Wei et al., 2018; Liang et al., 2020). The $m^6\text{A}$ destabilizes $\text{A} = \text{U}$ pairing due to altered energetics/steric hindrance; however, the donor and acceptor in the hydrogen bond remain the same (Roost et al., 2015).

On the other hand, CH_3 of $m^1\text{A}$ in RNA provides a positive charge (which interacts with negatively charged phosphate in the backbone) and it bulges out of the Watson-Crick hydrogen bond resulting in a strong electrostatic interaction (Helm, 2006). Moreover, 2'-O-methylation confers hydrophobicity, which protects the RNA from nucleolytic attack and stabilizes RNA coiling (Kumar et al., 2014). Thus, structure and functions of dynamic RNA modifications during the developmental process and environmental stress, and their effects on gene expression have emerged as a new branch of functional genomics known as 'epitranscriptomics.'

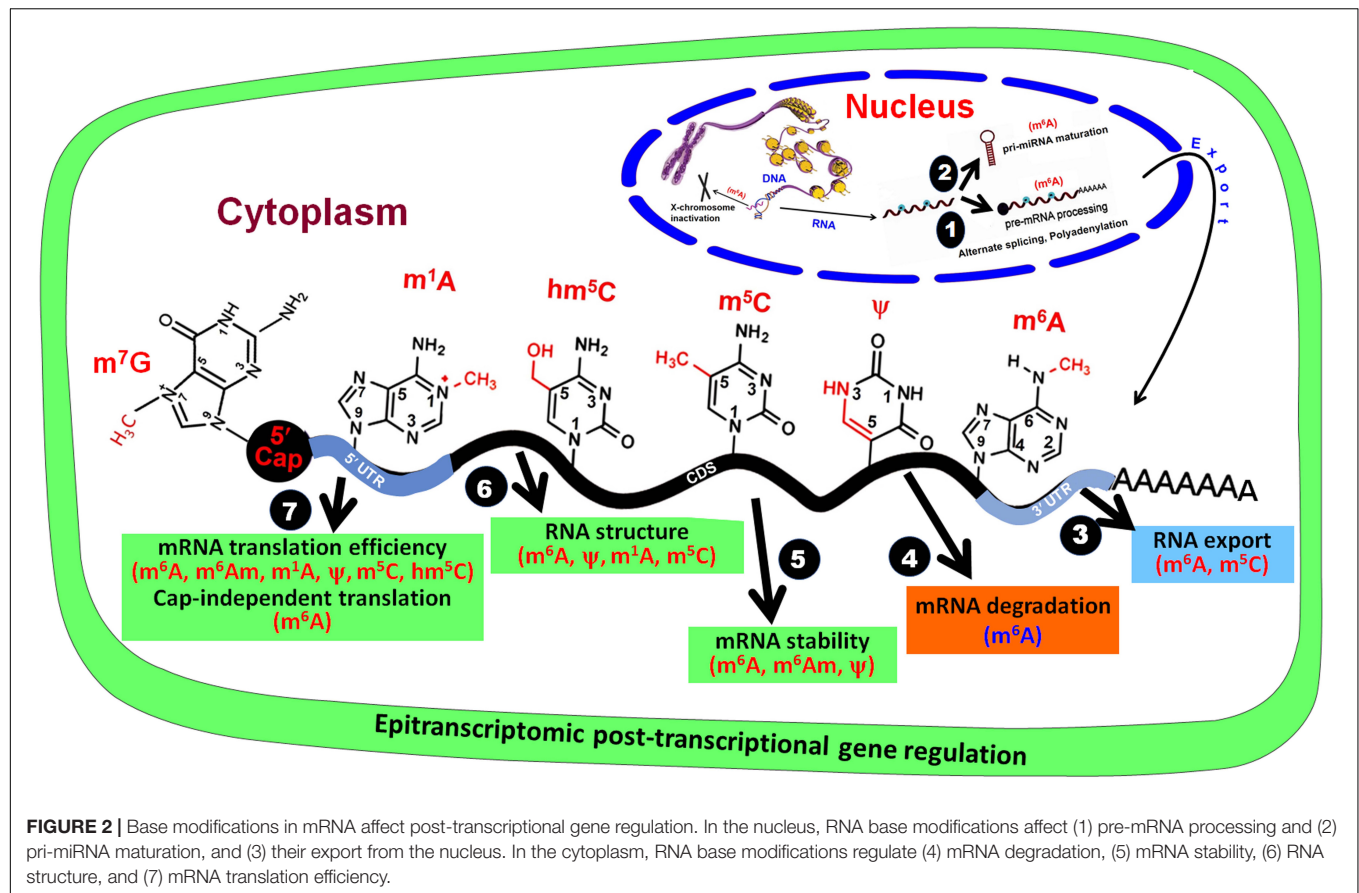
To decipher the biological functions of a modified RNA base, it is vital to identify the writer/reader/eraser that modulates the modification. However, high-throughput detection methods for many of these modifications are still lacking. The recent



advances in high-throughput next-generation sequencing (NGS) together with the novel chemogenetic RNA-labeling techniques have provided unprecedented opportunities to understand the RNA structure and functions. Such advances provide a better understanding of the presence and dynamics of base modifications like m^6A (Zhao et al., 2017b), m^5C (Cui Q. et al., 2017; David et al., 2017; Fang et al., 2020), 5-hydroxymethylcytidine (hm^5C) (Huber et al., 2015; Delatte et al., 2016; Zhang et al., 2016), and m^1A (Dominissini et al., 2016; Li et al., 2016a; Shen et al., 2016; Xiong et al., 2018) in RNAs. Base modifications, such as addition of 5' cap (e.g., N^7 -methylguanosine, m^7G -cap), and RNA editing are vital for mRNA stability (Kiledjian, 2018), translation (Topisirovic et al., 2011; Holstein et al., 2016) and functional diversity (Peng et al., 2018). More importantly, NAD^+ has been reported to be a new/alternative RNA cap in diverse organisms including bacteria, yeast, human (Cahová et al., 2015; Jiao et al., 2017; Walters et al., 2017; Frindert et al., 2018), and plant (Wang et al., 2019). Thousands of transcripts for the protein-coding genes from nuclear and mitochondrial genomes in Arabidopsis were observed to contain NAD^+ cap (Wang et al., 2019). These clearly indicate that NAD^+ cap is one of the evolutionarily conserved caps that affects mRNA metabolic processes. A comprehensive

understanding of the distribution, function, and regulation of RNA base modification will further increase the available knowledge on epitranscriptomic regulation of gene expression.

Epitranscriptomic base modifications have become an interesting topic of research and review, particularly in the animal system (Meyer et al., 2012; Carlile et al., 2014; Fu Y. et al., 2014; Dominissini et al., 2016; Peer et al., 2017; Angelova et al., 2018; Arribas-Hernandez et al., 2018; Khoddami et al., 2019; Leonardi et al., 2020). Now, the epitranscriptomic modifications in plants like Arabidopsis (Luo et al., 2014; Wan et al., 2015; Shen et al., 2016; Zuber et al., 2016; Cui Q. et al., 2017; David et al., 2017; Duan et al., 2017), rice (Li Y. et al., 2014), maize (Luo et al., 2019; Miao et al., 2019), and tomato (Zhou et al., 2019) are also being studied. However, our knowledge of plant epitranscriptomic modifications, except for the 5'-cap and poly-A-tail, is limited to uridylation (de Almeida et al., 2018), m^6A (Li et al., 2018), and m^5C (Cui X. et al., 2017; David et al., 2017). Other types of modifications can also be expected to occur in plant mRNAs but their existence/detection and roles/functions remain to be explored. Considering the crucial and dynamic roles of epitranscriptomic modifications in many biological processes like embryo development, leaf morphogenesis, root development, floral transition, fruit



ripening, and stress tolerance, the importance and future perspectives of epitranscriptomic research in plants are being discussed (Hu et al., 2019; Shen et al., 2019; Liang et al., 2020). The present review focuses on recent developments in base modifications in RNAs, particularly m^6A and m^5C in plant mRNAs, their biochemical properties, and functions. Moreover, the review discusses technological advances in high-throughput detection methods to elucidate epitranscriptomic modifications, as well as the technological limitations. Further advances in the next-generation detection techniques and functional analysis of RNA base modifications might facilitate epitranscriptomic manipulation of the traits of interest.

BIOCHEMISTRY OF ADENOSINE METHYLATION IN MRNA

Methylation of adenosine (A) at N^6 position [in both *syn*- (energetically favored) and *anti*-conformation] results in the formation of m^6A (Zou et al., 2016). The methyltransferase-like 14 (METTL14) complex and Wilm's tumor-associated protein (WTAP) work in cooperation with METTL3, and cofactors KIAA1429, RBM15/RBM15B which constitute a functional methyltransferase to create m^6A in mammalian mRNAs at a consensus sequence of R— m^6A —C—H (where R = A/G, and H = A/C/U) (Patil et al., 2016). Emerging

evidence suggests that VIRMA/KIAA1429 recruits the catalytic core (METTL3/WTAP/METTL14) for a sequence-specific methylation of A to m^6A (Yue et al., 2018). Recent studies suggest that ZC3H13 is another component of the m^6A writer-complex, and it regulates the methylation of A (Knuckles et al., 2018; Wen et al., 2018). Moreover, m^6A mark gets erased by the enzymes like fat mass and obesity-associated protein (FTO) and alkylation repair homolog protein 5 (ALKBH5), which convert it back to A (Jia et al., 2011). FTO oxidatively removes m^6A through N^6 -hydroxymethyladenosine (hm^6A) and N^6 -formyladenosine (f^6A) intermediates (Fu Y. et al., 2014). Thus, m^6A is a reversible epitranscriptomic modification, which functions to regulate gene expression.

m^6A Writer

An RNA methyltransferase complex is comprised of methyltransferase-like 3 (METTL3) (Bokar et al., 1997), METTL14 (Liu et al., 2014), KIAA1429/VIRMA (Schwartz et al., 2014b; Yue et al., 2018), HAKAI (Ruzicka et al., 2017), RNA binding motif protein 15 (RBM15) (Patil et al., 2016), Wilm's tumor 1-associating protein (WTAP) (Ping et al., 2014), and a zinc finger CCCH domain-containing protein 13 (ZC3H13) (Frye et al., 2018; Wen et al., 2018; Yang et al., 2018). It is involved in methylation/modification of adenosine to m^6A in mammals. While METTL3 is known to methylate single-stranded RNAs (ssRNAs) in a sequence-specific (RRACH)

manner, METTL16 methylates structured RNAs having a nonamer sequence (UACAGAGAA; the targeted adenosine for methylation is marked with bold face) (Pendleton et al., 2017). Thus, METTL16 is another m⁶A-specific methyltransferase which targets U6 snRNA and human *MAT2A* mRNA encoding for S-adenosylmethionine (SAM) synthetase (Pendleton et al., 2017). Interestingly, SAM is the methyl group donor for methylation of DNA, RNA, and proteins.

In Arabidopsis, the m⁶A writer complex is composed of adenosine methyltransferase (MTA) (METTL3 ortholog), its homolog MTB (METTL14 ortholog), FKBP12 interacting protein 37 (FIP37) (WTAP ortholog), VIRLIZER/KIAA1229 (VIR), and HAKAI (Ruzicka et al., 2017) (Table 1). Although the components of plant writer complex were observed to be distributed in the nucleoplasm, but FIP37 and VIR do not affect alternative splicing of transcripts (Shen et al., 2016; Ruzicka et al., 2017). While WTAP interacts with METLL3, METTL14, VIRMA, and HAKAI in mammals (Yue et al., 2018), Arabidopsis FIP37 (a WTAP ortholog in mammals) interacts directly with MTA only (Ruzicka et al., 2017). This clearly indicates that the mechanism of adenine methylation (m⁶A) is conserved among the eukaryotes; however, some unique features of m⁶A modification might have been evolved in plants. Most of the constituents of m⁶A writer complex, excluding HAKAI, are needed for the embryonic development. Moreover, m⁶A plays diverse roles in various other developmental processes in plants. Hence even after conserved m⁶A modification machinery in eukaryotes, it appears that individual members of m⁶A writer complex has achieved functional divergence in plants.

m⁶A Reader

Methyladenosine (m⁶A) affects several mRNA metabolic processes in both nucleus and cytoplasm through the recruitment of m⁶A-binding protein (RBP), also known as m⁶A reader (Wang et al., 2015; Yue et al., 2015; Zhou et al., 2015; Xiao et al., 2016; Li A. et al., 2017; Yang et al., 2017; Zhao et al., 2017b; Scutenaire et al., 2018; Wei et al., 2018). Two important classes of m⁶A readers known so far include the YTH domain-containing protein (Zhang et al., 2010) and the heterogeneous nuclear ribonucleo-protein (HNRNP) (Alarcon et al., 2015a). Arabidopsis and rice genomes contain several genes (13 and 12, respectively) for the YTH homolog known as 'evolutionarily conserved C-terminal region' (ECT) (Li D. et al., 2014); however, their role as an m⁶A-reader has only recently been recognized (Arribas-Hernandez et al., 2018; Scutenaire et al., 2018; Wei et al., 2018). These genes exhibit distinct/diverse expression pattern in different organs at different developmental stages, and under different stress conditions. A plant-specific motif URUAW (R = G or A; W = U or A) was reported in the ECT2-binding sites, which is different from the YTH-binding motif observed in human (Xiao et al., 2016; Shen et al., 2019). The binding of ECT2 at m⁶A increases the stability of the transcript responsible for trichome morphogenesis/development in Arabidopsis. ECT2 functions with ECT3 and ECT4 to regulate leaf formation/morphogenesis (Arribas-Hernandez et al., 2018). Structural analysis of the m⁶A binding domain in yeast and mammalian YTH revealed that the it recognizes m⁶A

in mRNA through a hydrophobic aromatic cage containing three conserved tryptophan residues (Luo and Tong, 2014; Theler et al., 2014; Xu et al., 2014). Mutation in the hydrophobic cage of ECT2 and ECT3 was reported to abolish the function of m⁶A recognition (Arribas-Hernandez et al., 2018; Wei et al., 2018; Scutenaire et al., 2018), which suggest that specific-binding of ECT to m⁶A is essential for their functional activity in leaf and trichome development. Occurrence of a number of YTH proteins in Arabidopsis and rice, having very high sequence similarity (Li D. et al., 2014), might help elucidating their roles in interpreting m⁶A epitranscriptome in plants by creating/using multiple knockout mutants.

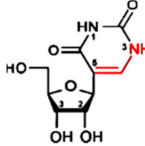
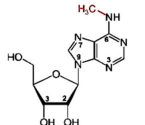
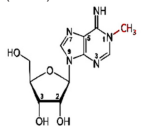
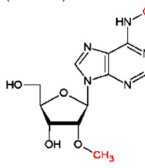
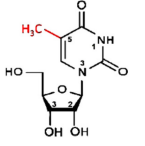
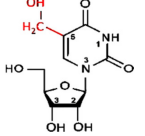
m⁶A Eraser

Since the formation of m⁶A is a reversible process, it is dynamically removed from the mRNA by two ALKBH family m⁶A demethylases namely 'Fat mass and obesity-associated protein' (FTO) (Jia et al., 2011) and α -ketoglutarate-dependent dioxygenase homolog 5 (ALKBH5) (Zheng et al., 2013) in mammals. In Arabidopsis (13) and rice (9) a number of ALKBH family proteins have been reported (Mielecki et al., 2012; Liang et al., 2020). Phylogenetic analysis showed no orthologs of FTO to be present in plants (Liang et al., 2020). But, the existence of multiple copies of ALKBH5 orthologs [six orthologs (ALKBH9A/B/C and ALKBH10A/B/C) in Arabidopsis] suggests redundant functions of these proteins in m⁶A demethylation. They are differentially expressed in different tissues (Duan et al., 2017) with their diverse subcellular localization (Mielecki et al., 2012). This again suggests their role in functional divergence in m⁶A dynamics in plants (Burgess et al., 2016). ALKBH9B, ALKBH10B, and SLALKBH2 (Zhou et al., 2019) remove m⁶A from mRNA in Arabidopsis (Duan et al., 2017; Martinez-Perez et al., 2017). ALKBH10B removes m⁶A from mRNAs for several regulators, which enhances the stability of the transcripts and promotes floral transition. Thus, m⁶A promotes degradation of mRNAs for developmental regulators in Arabidopsis (Duan et al., 2017). This indicates that it might potentially be used as an epitranscriptomic mark for modulating flowering time in crop plants.

Occurrence of m⁶A

Occurrence of m⁶A has been observed across the animals, plants, single-cell organisms (archaea, bacteria, and yeast), and viruses (Zhao et al., 2017a). Three independent studies showed consensus on adenosine methylation (m⁶A) motif RRACH in yeast, mammals, and plants (Dominissini et al., 2012; Schwartz et al., 2013; Luo et al., 2014). It has been detected in mRNAs of many plant species, including Arabidopsis, maize, wheat, oat, and rice (Zhong et al., 2008). In Arabidopsis, m⁶A content varies in different tissues, ranging from 0.4% in seeds to 1.5% in young seedlings (Zhong et al., 2008). Three independent studies reported m⁶A mapping in different ecotypes and tissues of Arabidopsis (Luo et al., 2014; Wan et al., 2015; Shen et al., 2016). m⁶A was reported to be widely distributed in >5,000 transcripts, and accumulated near the start and the stop codons, as well as in the 3' UTR (Luo et al., 2014). However, occurrence of m⁶A near the start codon was not detected in the methylome

TABLE 1 | Modified RNA bases, their modulators, and interpreters.

RNA base modification	Enzymes/proteins						References
	Writer		Eraser		Reader		
	Animal	Plant	Animal	Plant	Animal	Plant	
Pseudouridine (Ψ) 	PUS1, PUS2, PUS3, PUS4, PUS6, PUS7, PUS9, PUS13 DKC1, BoxH/ACA	?	?	?	?	?	Carlile et al., 2014; Lovejoy et al., 2014; Spenkuch et al., 2014; Rintala-Dempsey and Kothe, 2017; Adachi et al., 2019; Khonsari and Klassen, 2020
N^6 -methyladenosine (m6A) 	METTL3, METTL14 METTL16 WTAP RBM15B VIRMA ZC3H13 HAKAI Spenito	MTA, MTB FIP37 VIR HAKAI	ALKBH5 FTO	ALKBH9B ALKBH10B SIALKDH2	YTHDC1 YTHDC2 YTHDF1 YTHDF2 YTHDF3 eIF3 HNRNPC HNRNPA2B1 SRSF2	ECT2 ECT3 ECT4 COSF30L	Zhang et al., 2010; Zheng et al., 2013; Liu et al., 2014; Wang et al., 2015; Du et al., 2016; Patil et al., 2016; Martinez-Perez et al., 2017; Arribas-Hernandez et al., 2018; Pendleton et al., 2017; Scutenaire et al., 2018; Wei et al., 2018
N^1 -methyladenosine (m1A) 	TRMT61B, TRMT10C, and the complex of TRMT6, TRMT61A	?	ALKBH1 ALKBH3	?	?	?	Chujo and Suzuki, 2012; Dominissini et al., 2016; Li et al., 2016a; Liu et al., 2016
$N^6,2'$ -O-dimethyladenosine (m6Am) 	CMTR1 CMTR2 PCIF	?	FTO	?	?	?	Belanger et al., 2010; Jia et al., 2011; Werner et al., 2011; Mauer et al., 2017; Boulias et al., 2019; Sun et al., 2019
5-methylcytidine (m5C) 	NSUN2 DNMT2	TRM4B	?	?	ALYREF YBX1	?	Squires et al., 2012; Hussain et al., 2013; Cui X. et al., 2017; David et al., 2017; Yang et al., 2017; Yang Y. et al., 2019
5-hydroxymethylcytidine (hm5C) 	TET1, TET2, TET3	?	?	?	?	?	Fu L. et al., 2014; Huber et al., 2015; Delatte et al., 2016

Modified RNA bases: Ψ , pseudouridine; 6-mA, N^6 -methyladenosine; 6-mAm, $N^6,2'$ -O-dimethyladenosine; 1-mA, N^1 -methyladenosine; m5C, 5-methylcytidine; hm5C, 5-hydroxymethylcytidine. ALKBH5, AlkB homolog 5; ALYREF, Aly/REF export factor; CMTR1, cap methyltransferase 1; DKC1, Dyskeratosis congenital protein 1; DNMT2, DNA methyltransferase 2; ECT2, Evolutionarily Conserved C Terminal region 2; eIF3, eukaryotic translation initiation factor 3; FIP37, FKBP12 Interacting Protein 37KD; FTO, fat mass and obesity-associated protein; HAKAI, a conserved E3 ubiquitin ligase in Arabidopsis; HNRNPA2B1, HNRNPC-Heterogeneous nuclear ribonucleoproteins A2/B1; KIAA1429, protein virilizer homolog; METTL3, methyltransferase-like 3; MTA, adenosine methyltransferase; MTB, closest homolog of MTA; NSUN2, NOL1/NOP2/Sun RNA methyltransferase family member 2; PUS1–PUS4, Pseudouridine synthase 1–4; RBM15, RNA-binding motif protein 15; SRSF2, serine/arginine-rich splicing factor 2; TET1–TET3, 10–11 translocation protein 1–3; TRM4B, tRNA-specific methyltransferase 4B; TRMT61B, tRNA-1-mA methyltransferase 61B; VIR, VIRLIZER/KIAA1229; WTAP, Wilms' tumor 1 associated protein; YTHDF1–3, YTH domain family proteins 1–3; YTHDC1, YTH domain-containing protein 1; ZC3H13, CCCH-type zinc finger proteins. "?" indicates the unknown writer/eraser/reader.

of leaf, flower, and root of *Arabidopsis* (Wan et al., 2015), probably because of the dynamic nature of the modified m⁶A. Differentially methylated mRNAs were observed in leaf, flower, and root of *Arabidopsis* (Wan et al., 2015), indicating the role of m⁶A in tissue/organ differentiation. The m⁶A writers MTA/MTB, FIP37/VIRILIZER/HAKAI were reported to be involved in embryo and plant development (Shen et al., 2016; Ruzicka et al., 2017; Hu et al., 2019). YTH/ECT and ALKBH, reader and eraser, respectively, play important role in growth, development and flowering in *Arabidopsis* (Duan et al., 2017; Arribas-Hernandez et al., 2018; Scutenaire et al., 2018; Wei et al., 2018). Differential methylation of several transcripts in root, leaf, and flower of *Arabidopsis* (Wan et al., 2015), suggests that m⁶A dynamics of specific transcripts might be an integral part of tissue/organ differentiation in plants (Shen et al., 2019). Another recent work on epitranscriptomic profiling of salt-treated *Arabidopsis* leaf reported m⁶A enrichment in the transcripts for salt- and osmotic-stress responses (Anderson et al., 2018).

In addition to *Arabidopsis*, the enzymes associated with epitranscriptomic modifications have been reported in some of the agronomically important plants like *Nicotiana glauca*, maize, rice, and tomato. The methylases and demethylases have also been reported in plants, and they are evolutionarily conserved. Any change in their expression shows a significant alteration in the m⁶A content in polyadenylated transcriptome, and drastic physiological impacts. Analysis of m⁶A landscape in rice (Li D. et al., 2014) exhibited a similar pattern that was observed in *Arabidopsis*, which indicates a conserved m⁶A distribution in plants. Accumulating evidences also indicate that writers/readers and erasers play important roles in abiotic stress responses in plants (Hu et al., 2019). Zhang F. et al. (2019) identified a panicle-specific m⁶A motif UGWAMH (W = U/A; M = C/A; H = U/A/C) in rice. Despite the progress being made in understanding m⁶A landscape in crop plants, the writers, readers, erasers for m⁶A and its functions in plant growth, development, and survival under the stress are yet to be elucidated. However, the position, pattern, and motif of m⁶A suggest that the writers, readers, and erasers might be conserved across the kingdoms.

Methylation at Other Positions in Adenosine

In addition to the m⁶A, the human epitranscriptome is known to contain other modified/methylated forms of adenosine like m¹A and m⁶Am (Hauenschild et al., 2015; Molinie et al., 2016). Methylation at the N¹ position of adenosine creates N¹-methyladenosine (m¹A), and it has been prevalent in rRNA and tRNA. However, the occurrence of m¹A has also been reported in the human transcriptome (Li X. et al., 2017), which can be erased by ALKBH3 (Li et al., 2016b). The CH₃ group at N¹ position of m¹A interferes with standard base pairing (Hauenschild et al., 2015), which affects mRNA folding around the transcription start site (TSS) and facilitates initiation of translation. Despite the progress in the detection of modified nucleosides, transcriptome-wide distributions of m¹A in plants remain unknown. When adenosine is methylated at the C₂ position of ribose sugar [by 2'-O-methyltransferase (CMTR: Cap methyltransferase) to

form 2'-O-methyladenosine (Am) (Werner et al., 2011) and then it is methylated at the N⁶ position of adenosine [by an unidentified nucleocytoplasmic methyltransferase], it forms N⁶, 2'-O-dimethyladenosine (m⁶Am). The m⁶Am modification is exclusively distributed at the TSS (generally after the m⁷G cap) in certain mRNAs (Linder et al., 2015) at a frequency of 0.003% (Molinie et al., 2016). m⁶Am was reported to be mediated by phosphorylated CTD interacting factor 1 (PCIF1) which catalyzes methylation of m⁶A to m⁶Am at the 5' end of mRNA (Sendinc et al., 2019). Although such epitranscriptomic modifications play important roles in mammals, they are remained to be identified/characterized in plant.

MODIFICATION OF OTHER BASES IN MRNA

Besides the modifications of adenosine, epitranscriptome is known to contain methylation/modification at other bases, for example, m⁵C, hm⁵C, m³C, ac⁴C, m¹G, m⁷G, 8-oxo-G, Uridylation, Pseudouridine (ψ), and Inosine (I), particularly in animal systems (reviewed by Shen et al., 2019; Boo and Kim, 2020). While the occurrence of some of the modified bases (e.g., m⁵C, hm⁵C, m⁷G, and ψ) have been confirmed (Huber et al., 2015; Vandivier et al., 2015; Burgess et al., 2016; Zuber et al., 2016; Cui Q. et al., 2017; Martinez-Perez et al., 2017; Malbec et al., 2019), presence of m¹G has been predicted in *Arabidopsis* epitranscriptome. Many of these epitranscriptomic modifications like m³C, m⁷G, 8-oxoG, and I play important roles in animals (Palladino et al., 2000; Torres et al., 2014; Arimbasseri et al., 2016; Xu L. et al., 2017; Malbec et al., 2019), but their existence/identification and functional characterization remains to be confirmed in plants.

Cytosine Modifications in mRNA

Occurrence of methylcytidine (m⁵C) is common in tRNAs and rRNAs (Squires and Preiss, 2010), but it has also been identified in mRNAs and ncRNAs (Squires et al., 2012). Since m⁵C is less abundant (0.4% of total cytosine, compared to ~1.5% of m⁶A in human transcripts), much less has been researched on its occurrence and functions (Squires et al., 2012; Ke et al., 2015). Detection of m⁵C in mRNAs of different plant species, including *Arabidopsis*, *Medicago*, rice, maize, and foxtail millet, has been reported (Cui Q. et al., 2017). Change in m⁵C level across the tissues in *Arabidopsis*, with a gradual increase during vegetative growth, suggest a dynamic change in m⁵C content during plant growth and development. More than one thousand m⁵C were detected on transcriptome-wide analysis of shoot, root, and siliques of *Arabidopsis*, but only a few dozen of them were commonly present among these tissues (David et al., 2017). m⁵C is generally accumulated in the coding sequence (CDS) of the mRNA in HACCR (where H = A, U or C; R = A or G) and CTYCTYC (Y = U or C) motifs in *Arabidopsis* (Cui Q. et al., 2017). A marginal increase in expression of TRM4B (an m⁵C writer) was observed under cold stress in *Arabidopsis*, but it showed decreased expression under heat stress. However, the expression level of TRM4B was not altered in rice under abiotic

stresses (Zou et al., 2016). TRM4B has been further characterized in plants (David et al., 2017; Cui Q. et al., 2017), and m⁵C was observed to be required for root development and oxidative stress responses (David et al., 2017). TRM4B loss-of-function mutants of *Arabidopsis* exhibited down-regulated expression of short hypocotyl 2 (SHY2) and indoleacetic acid-induced protein 16 (IAA16) genes involved in root development. Stability of the transcripts of such genes was observed to be positively correlated with the m⁵C modification/content (Cui Q. et al., 2017).

Writer, Reader, and Eraser of m⁵C

Formation of m⁵C in human mRNA is catalyzed by methyltransferases such as DNMT2 and NSUN2 (Squires et al., 2012; Bohnsack et al., 2019). NSUN6, a Type II m⁵C site-specific methyltransferase, was reported to negatively correlate m⁵C methylation with translation efficiency (Liu et al., 2020). Recently, Selmi et al. (2021) mapped NSUN6-dependent m⁵C sites in human transcripts, which were located in protein coding RNAs at 3'-UTR within a consensus sequence (CTCCA) motif, and mark translation termination. Eight m⁵C methyltransferases are encoded by *Arabidopsis* genome, two of them are tRNA-specific methyltransferase 4A (TRM4A) and TRM4B (Chen et al., 2010; Cui Q. et al., 2017). While TRM4A is responsible for m⁵C in tRNA, TRM4B targets mRNA for the modification. A recent study demonstrated that an RRM motif-containing ALY protein binds to m⁵C-containing mRNAs in *Arabidopsis* (Pfaff et al., 2018). The *aly* mutants showed shorter primary roots, defective reproductive development including abnormal flowers and reduced seed production (Pfaff et al., 2018). Thus, m⁵C is another important epitranscriptomic mark that affects plant growth, development and adaptive responses in plants. Although m⁵C is reported to be further oxidized to hm⁵C by a family of Ten-eleven translocation (TET) enzymes (Huber et al., 2015; Delatte et al., 2016), varying hm⁵C content in different *Arabidopsis* tissues indicate that it is a dynamic epitranscriptomic mark in plants (Shen et al., 2019). Despite the progress in detecting/distribution of hm⁵C, its oxidation to m⁵C in mRNA is still not fully demonstrated. However, further research would be required to identify m⁵C readers/erasers, and elucidate the mechanisms/functions of m⁵C-mediated regulation of gene expression.

Methylation at Other Positions in Cytosine

Cytosine can also be acetylated at the N⁴ position by an N-acetyltransferase (NAT10) to form ac⁴C. Such modification is commonly found in tRNA, rRNA, but it has also been observed in mRNA (Dong et al., 2016; Arango et al., 2018). ac⁴C was observed distributed in coding and non-coding RNAs in human, abundant near the TSS (Arango et al., 2018). The occurrence of ac⁴C increases mRNA half-life and promotes translation efficiency. NAT10 acts as the primary ac⁴C writer, and NAT10 knocking out reduces ac⁴C content in RNA. In yeast, orphan box C/D snoRNAs complex guides Kre33 (a yeast homolog of human NAT10) to the target sites for ac⁴C modification (Sharma et al., 2017). However, it is still not known whether ac⁴C is a reversible

or not, as neither an ac⁴C reader nor its deacetylation process is known. Moreover, its occurrence in plant and role/function in gene regulation is not yet known.

Modification of Other Bases in mRNA

Uridylation (addition of uridines at the 3' without any template) of mRNA, targeted for degradation, has been reported in both mice and *Arabidopsis* (Shen and Goodman, 2004; Zhang et al., 2017). Uridylation of mRNAs in plants is catalyzed by UTP:RNA uridylyltransferase1 (URT1) and terminal uridylyltransferase (TUTase) (Sement et al., 2013; Lim et al., 2014). Pseudouridine (Ψ), also known as the 5th base of RNA and the first modified RNA base (Davis and Allen, 1957), is a C-glycosidic rotational isomeric form of uridine (U), wherein U is attached to a ribose sugar through a carbon-carbon (instead of a nitrogen-carbon) glycosidic bond. Formation of Ψ in eukaryotes involves an RNA-dependent pseudouridine synthase (PUS) such as Cbf5 which uses a cofactor box (H/ACA ribonucleo-proteins) as a guide. Ψ formation may also occur through an RNA-independent PUS that does not require any cofactor (Carlile et al., 2014; Spenkuch et al., 2014). Ψ may further get methylated by EMG1 at the N¹ position to generate 1-methylpseudouridine (m¹Ψ) (Wurm et al., 2010). Although Ψ is mainly distributed around the CDS and the 3' UTR of the gene (Carlile et al., 2014; Li et al., 2015), its occurrence is yet to be mapped in plant mRNAs.

Oxidation of RNA bases due to excessive reactive oxygen species (ROS, e.g., superoxide, hydroxyl radicals, and hydrogen peroxide) generates different oxidized RNA bases like 8-oxoG, 8-oxo-7,8-dihydroadenosine, 5-hydroxycytidine, cytosine glycol, and 5-hydroxyuridine (Yan and Zaher, 2019). 8-Oxo-7,8-dihydroguanosine (8-oxoG, an oxidized form of guanine base) is one of the most abundant variants of guanosine found in mammalian cells associated with neurodegenerative diseases (Nunomura et al., 2017). This determines the fate of mRNA, including stability and translation (Yan et al., 2019). AU-rich element RBP 1 (AUF1) and Y-box binding protein 1 (YBX1) preferentially bind to 8-oxoG to trigger rapid degradation of 8-oxoG-containing mRNAs (Ishii et al., 2015). Recently poly(C)-binding protein 1 (PCBP1) was identified as an 8-oxoG reader protein. However, the binding of PCBP1 requires two 8-oxoGs located nearby in the RNA, and this is associated with cellular apoptosis under oxidative stress (Ishii et al., 2018). Reversal of 8-oxoG to a normal guanosine base, as observed in the case of many other RNA base modifications, is not yet known. Moreover, the occurrence of such modified RNA base(s) in plant can be expected, particularly under environmental stresses when ROS production increases significantly, but their existence has not yet been reported.

EFFECT OF MODIFIED BASE ON MRNA METABOLISM

Modified bases influence mRNA metabolism, including splicing, export, translation, and degradation of the transcript. Many of the functions of m⁶A in mRNA metabolism in animal system are well-known (Wang et al., 2014; Liu et al., 2015;

Hausmann et al., 2016; Shi et al., 2017; Kasowitz et al., 2018). However, only some of the functions of m⁶A and its readers like ECT2 are known in plants, including the regulation of 3' UTR processing and improving mRNA stability (Wei et al., 2018). Moreover, some other functions of the core components of the methyltransferase complex (MTA, MTB, and FIP37) in plant development and survival under abiotic stresses were deciphered by mutation/knock-out studies in Arabidopsis (Tzafrir et al., 2003; Vespa et al., 2004; Zhong et al., 2008; Bodi et al., 2012). The function of another modified adenosine base, m⁶Am (a close homolog of m⁶A), is yet poorly understood, but it has been reported to improve translation efficiency and mRNA stability in mice by protecting the mRNA from decapping enzymes like DCP2 (Mauer et al., 2017).

m⁵C facilitates binding of ALYREF (an mRNA export adaptor) and removal of NSUN2 (an m⁵C writer), which disrupt mRNA transport from the nucleus (Yang et al., 2017). In Arabidopsis, a reduced ribosomal occupancy was observed in the m⁵C-marked mRNAs, indicating interfering role of m⁵C in binding of translational machinery (Cui Q. et al., 2017). A decreased m⁵C content accelerates mRNA decay, which indicates that it is another important epitranscriptomic mark affecting mRNA stability and translation efficiency in plants (Cui Q. et al., 2017).

Other modified bases, such as ψ , have been depicted to be involved in splicing and undisrupted translation of mRNA in yeast and mammals (Carlile et al., 2014; Karijolich et al., 2015). URT1-dependent uridylation and poly-A binding protein (PABP) in plants was reported to prevent excessive deadenylation, and thus, protects mRNA from degradation in Arabidopsis (Zuber et al., 2016). The occurrence of 8-oxoG considerably inhibits the efficiency of peptide bond formation, which restricts translation and triggers mRNA degradation (Boo and Kim, 2020). Transcription factors (TFs; e.g., ZFP217-dependent METTL3 and HIF-dependent ALKBH5), and miRNAs (e.g., miRNA responsible for RNA-dependent METTL3 activity) may also trigger the expression of writers and erasers of modified bases, demonstrating the feedback activation. This suggests a complex interplay between the modified bases and regulatory pathways. Thus, the stimuli and signaling/regulatory processes that fine-tune the transcription and translation processes of a gene might also affect the activity of writers, readers, and erasers through various RNA modifications. The same signaling pathway may also activate or inactivate the synthesis of readers and erasers through post-translational modifications.

Role of Modified Base on mRNA Translation

Translation process is regulated by the binding of ribosome and initiation factor activities, including phosphorylation of the 'eukaryotic initiation factor 2' (eIF2) (Pavitt, 2018). Translation efficiency was reported to be moderately increased in the METTL3-knockout mutants of mouse embryonic stem cells and embryoid bodies, which suggest a negative regulatory role of m⁶A on translation efficacy (Liu et al., 2015). However, the binding of YTHDF1 (a cytoplasmic m⁶A reader) cooperates with ribosomes and initiation factors to increase translation

efficiency (Wang et al., 2015). Recent studies demonstrate that m⁶A promotes translation efficacy of mRNAs (Li A. et al., 2017; Shi et al., 2017; Weng et al., 2018). Similarly, IGF2BPs (m⁶A-binding proteins) help reinforcing the stability and increase the translation efficiency of m⁶A-containing mRNAs (Huang et al., 2018). Studies also suggested that the presence of m⁶A in 5' UTR of an mRNA promotes initiation of cap-independent translation (Meyer et al., 2015) and the IGF2BPs-mediated translation (Huang et al., 2018). It has also been reported that eIF3 directly binds to m⁶A—harboring 5' UTR and engages the 43S ribosomal complex to begin the translation process, even in the absence of eIF4E (a cap-binding factor) (Meyer et al., 2015). The presence of m⁶A in the coding region of mRNA has been reported to disrupt tRNA boarding and elongation of the translation process *in vitro* (Choi et al., 2016). m⁶A has also been reported to negatively regulate the translation process by serving as a link between transcription and translation processes (Slobodin et al., 2017). All of these findings support the regulatory functions of m⁶A in mRNA translation.

Recent mapping studies indicate that m¹A is abundant in the 5' UTR of mRNA (Dominissini et al., 2016; Li X. et al., 2017), which is associated with higher translational efficiency; however, the underlying mechanism is yet to be discovered. In addition to this, the presence of m⁶Am creates hindrance in the binding of mRNA-decapping enzyme DCP2, which improves the stability of the transcript (Mauer et al., 2017). Moreover, m⁶Am also makes mRNA resistant to microRNA-mediated degradation (Mauer et al., 2017). Similarly, m⁵C has been reported to stabilize RNA secondary structure; hence, it influences translational fidelity (Helm, 2006; Squires and Preiss, 2010). While the presence of m⁵C at the first position in the CCC codon was reported to reduce translational product by ~40% using bacterial whole-cell extract, its presence at the 2nd position of the codon was reported to suppress translation termination (Hoernes et al., 2016). In contrast, hm⁵C has been reported to activate translation in *Drosophila melanogaster* (Delatte et al., 2016). The effects of Ψ on translation efficiency depend on its position in a codon.

Although m⁶A has been known to promote translation efficiency in the animal system (Meyer et al., 2015; Wang et al., 2015; Slobodin et al., 2017), a little is known about its functions in plants where it works differently. In maize, m⁶A was found to be negatively correlated with translation efficiency; however, this depends on the location and content of m⁶A in the gene (Luo et al., 2019). Similarly, m⁵C was also reported to be associated with reduced efficiency of translation in Arabidopsis (Cui Q. et al., 2017). A recent study reports m⁵C to play important role in mRNA stability (Yang L. et al., 2019), which in turn improves translation efficiency. Thus, the role of different methylated bases in mRNA translation needs to be further explored to better understand the epitranscriptomic regulation of gene expression in plants.

Role of Modified Base on mRNA Splicing, Export, and Decay

Transcripts with modified bases get easily exported, translated, and degraded, probably due to the binding of the reader at the modified base. Studies provide convincing evidence for

the regulatory function of m⁶A on processing of pre-mRNA and pri-miRNA (Alarcon et al., 2015b). A family of nuclear hnRNPs, an m⁶A-binding protein accelerate processing of pri-miRNAs through interaction with DGCR8 (Alarcon et al., 2015a). An hnRNPA2B1 modulate alternative splicing of transcripts (Alarcon et al., 2015a). Moreover, hnRNPC plays an important role in the pre-mRNAs processing (Rajagopalan et al., 2012). Pfaff et al. (2018) reported that an RRM motif-containing ALY protein binds to m⁵C-containing mRNAs and helps in mRNA export in Arabidopsis. Reports suggest that controlling RNA modification regulates mRNA stability which ultimately fine tunes the gene expression. Research demonstrates that alternatively spliced mRNAs in animals retain more m⁶A sites and the binding sites for METTL3. Geula et al. (2015) reported that a METTL3-deficient mouse embryonic stem cell retains intron and shows exon skipping. Thus, m⁶A exerts its effect through binding of the reader proteins, particularly a family of proteins containing YTH domain (Xu et al., 2014). YTHDF2 (a well-established m⁶A reader) specifically binds to m⁶A-containing mRNA to deploy CCR4–NOT deadenylase complex (Du et al., 2016) for mRNA transport to the processing bodies (Wang et al., 2014), which promotes degradation of mRNA through translocation of the transcript (Sheth and Parker, 2003). This indicates a linkage between m⁶A and mRNA degradation. m⁶A modification and binding of readers also affect mRNA splicing and alternative polyadenylation (Xiao et al., 2016; Kasowitz et al., 2018). An alternative to 3′–5′ exonucleolytic cleavage on mRNA, endoribonucleolytic cleavage of the m⁶A-containing mRNAs is mediated by interaction among the YTHDF2, heat-responsive protein 12 (HRSP12), and P/MRP (an endoribonuclease RNase) complex (Park et al., 2019). Presence of the 8-oxoG in mRNA causes ribosome stalling followed by no-go decay (Ikeuchi et al., 2018). The roles of modified RNA base in regulation of mRNA stability/decay have recently been reviewed by Boo and Kim (2020).

Effects of Methylated Base on Biological Processes

Complex cellular processes are intricately regulated by mRNA methylation. According to the cellular needs, mRNA export/localization is altered by RNA base methylation (Roundtree et al., 2017; Yang et al., 2017; Chen et al., 2019). The presence of m⁶A in transcripts of pluripotent TFs prompts transcriptomic flexibility in embryonic stem cells of mouse and human (Batista et al., 2014; Geula et al., 2015). Sequestration of METTL3 by ZFP217 indicates a complex interplay between epitranscriptome and TFs (Aguilo et al., 2015). The depletion of m⁶A from glioblastoma stem cells due to METTL3/14 knockdown was reported to promote self-regeneration and tumorigenesis (Cui X. et al., 2017). In Zebrafish (*Danio rerio*), m⁶A coordinates the elimination of maternal mRNAs with the help of Ythdf2 which is essential for maternal-to-zygotic transition (Cui X. et al., 2017). Heat-shock stress suppresses cap-dependent translation and induces adenine methylation (formation of m⁶A) at 5′ UTR of the transcripts (Meyer et al., 2015; Zhou et al., 2015). Although cells can discriminate

between self (modified) and non-self (unmodified) RNAs, epitranscriptome plays an important role in immune responses also (Kariko et al., 2012; Hull and Bevilacqua, 2016). A study on the precursor cells of neurons revealed that m⁵C regulates differentiation and motility of neural stem cells in mice and humans (Flores et al., 2017). Mutants for FIP37 displayed about 85% reduction in m⁶A content and a massive proliferation of apical meristem in the shoot (Shen et al., 2016). Loss of m⁶A in FIP37 mutants of Arabidopsis was reported to be a key regulator of transcripts like WUSCHEL and SHOOTMERISTEMLESS, which results in the accumulation of transcripts due to their decreased decay (Shen et al., 2016).

Advances in epitranscriptomics have revealed several potential biological roles of post-transcriptional mRNA modifications (Zhao et al., 2017a). Reports demonstrate that methylated transcripts have shorter 3′ UTRs and lesser stability than its unmethylated counterpart (Molinie et al., 2016). Thus, methylation of mRNA base, and synthesis/binding of TFs/regulatory proteins get synchronized in response to the development processes and environmental stimuli (Zhao et al., 2017b). In mouse brain, the m⁶A level was reported to increase throughout the lifespan (Meyer et al., 2012). Studies have shown the role of m⁶A accumulation in learning and memory in mouse mediated by Ythdf1 binding in response to stimuli (Shi et al., 2018). Moreover, a recent study suggests the stress-mediated regulation of m⁶A accumulation in patients with depression, indicating that the dysregulation of m⁶A is associated with the development of mental disorders (Engel et al., 2017). An impaired build-up of m⁶A disrupts sex determination in *Drosophila*, and it causes embryonic-lethality in plants (Haussmann et al., 2016). Moreover, reduced accumulation of m⁶A inhibits the differentiation of embryonic stem cells in mammals (Batista et al., 2014; Geula et al., 2015). Some of the studies also suggest that the presence of m⁶A in mRNA plays a crucial role in spermatogenesis in mice (*Mus musculus* L.) (Hsu et al., 2017; Xu K. et al., 2017).

In Arabidopsis, the deficiency of mRNA adenosine methylase enzyme (a homolog of METTL3) has serious effects on plant growth and development (Bodi et al., 2012). Mutation studies on m⁶A methyltransferase core components (MTA, MTB, and FIP37) in Arabidopsis suggest that m⁶A is essential for the survival of the plant (Zhong et al., 2008). Arabidopsis mutants for FIP37 displayed an 85% reduction in m⁶A content and massive proliferation of apical meristem in the shoot (Shen et al., 2016). Knockdown of MTB in Arabidopsis was reported to cause a considerable reduction in height of the plant, while hypomorphic *vir* allele produced defective roots and the *VIR* null mutants were observed to be embryo-lethal (Ruzicka et al., 2017). A distinct pattern of m⁶A accumulation was observed in different organs of *Arabidopsis*, which suggests that m⁶A plays a role in organogenesis and it has tissue-specific functions (Wan et al., 2015). The content of m⁶A in Arabidopsis transcripts is controlled by 13 different ALKBHs (Mielecki et al., 2012) which indicate dynamic expression and diverse subcellular localization of ALKBH in plant. The *alkbh10b* mutants of Arabidopsis showed elevated m⁶A content in >1,000 transcripts and delayed floral transition, indicating that it mediates demethylation (removal

of m⁶A) of regulatory transcripts (Duan et al., 2017). A wild-type *ALKBH10B* could restore the *alkbh10* mutant phenotype, suggesting that m⁶A is an important regulator of flowering time in plants. The regulatory function of ECTs in leaf morphogenesis and trichome development has recently been demonstrated (Arribas-Hernandez et al., 2018; Scutenaire et al., 2018; Wei et al., 2018). ECT2 binds at m⁶A sites in the trichome development-related genes and improves the mRNA stability. Transcripts of the genes in *ect2* mutant get degraded at an accelerated rate and affect trichome branching (Wei et al., 2018), which suggests that m⁶A mediates trichome and leaf development by the recruitment of reader proteins.

A large number of differentially methylated transcripts were observed in leaf, flower, and root of Arabidopsis, while >14,000 transcripts were found to contain m⁶A in rice leaf (Wan et al., 2015). Findings suggest that m⁶A might be involved in tissue-differentiation in plants. *OsMTA2* and *OsFIP* were identified to be the important components of RNA m⁶A methyltransferase complex, and m⁶A is involved in the regulation of sporogenesis, particularly male gametogenesis in rice (Zhang F. et al., 2019). A loss-of-function mutation in *OsFIP* resulted in the early degeneration of microspores, irregular meiosis in prophase I. Tomato *slalkbh2* mutants showed delayed fruit ripening phenotypes and increased m⁶A content compared with the wild-type plants (Zhou et al., 2019). *SLALKBH2* is involved in the demethylation of the *SIDML2* mRNA and regulates its degradation. *SIDML2* encodes a DNA demethylase that regulates the expression of *SLALKBH2* through DNA (5-mC) methylation. This suggests a novel mechanism of gene regulation connecting epigenetics (DNA methylation, 5-mC) and epitranscriptomics (mRNA modification, m⁶A) (Zhou et al., 2019). Loss of function mutation in Arabidopsis for *TRM4B* (an m⁵C writer) resulted in defective root phenotype because of the decreased content of m⁵C in the genes involved in root development (Cui Q. et al., 2017).

mRNA base modifications (m⁶A and m⁵C) are sensitive to environmental changes in plants. The m⁵C content was reported to decrease under drought and heat stress in Arabidopsis (Cui Q. et al., 2017). Similarly, m⁶A content was reported to decrease in drought stress (Zhou et al., 2019), suggesting the epitranscriptomic regulation of stress responses in plants. The findings indicate that m⁶A and m⁵C play important roles in post-transcriptional regulation of gene expression in plants. Several studies on the writers, readers, and erasers in plants demonstrate that mRNA modification is an important molecular mechanism for regulating plant development and environmental responses (Shen et al., 2016; Cui Q. et al., 2017; Duan et al., 2017; Martinez-Perez et al., 2017; Scutenaire et al., 2018). More importantly, most of the above-mentioned functions result due to silencing/over-expressing of gene or due to the combined action of reader/eraser but not only due to the removal/accumulation of any RNA base modification. Thus, the authors agree with the limitations of the studies/reports, and realize the importance of the factor(s) involved. Moreover, the mechanism for synchronized response of writers, readers, and erasers to internal/external stimuli is still elusive. Despite the progress in understanding the functions of

m⁶A and m⁵C in plants, the mode of action of their writers, readers, and erasers are yet to be discovered.

DETECTION OF MODIFIED BASE IN RNA

Post-transcriptional modifications in RNA bases have been reported to play essential roles in various functional RNAs. These modifications alter the structure, processing, and functions of RNAs. A comprehensive understanding of the biochemical modifications in mRNA bases and the changes in accompanying non-covalent interactions is required to gain insights into the functional diversity. m⁶A is the most abundant modified mRNA base and the first epitranscriptomic modification mapped (Dominissini et al., 2012; Meyer et al., 2012). Although marvelous progress has been made in understanding the modified mRNA bases (Liang et al., 2020), in-depth insights into the dynamics, structure, and functions of such epitranscriptomic modification in this fascinating messenger biomolecule are essential. Detection of the modified RNA base helps understanding its dynamics and biological functions. Modern high-throughput technologies together with the conventional methods (Table 2) are expected to advance the field of epitranscriptomics by generating data and discoveries. However, most of the current methods of detecting modified base are specific for a particular modification but recently Khoddami et al. (2019) reported a method (RNA bisulfite sequencing, RBS-seq) for transcriptome-wide detection of multiple base modifications (m⁵C, Ψ, and m¹A) simultaneously at single-base resolution. In this section, we present an overview of the technological advancements in the detection methods, their applications, and their limitations.

Thin-Layer Chromatography

Thin-layer chromatography (TLC) has been one of the conventional methods for revealing RNA nucleobase modification (Keith, 1995). Generally, a modified base differs from its unmodified counterpart in terms of the net charge, polarity, and/or hydrophobicity, which allow their separation through chromatography. TLC separation of bases can be performed in one-dimension (1D) or two-dimensions (2D) using microcrystalline cellulose as a stationary phase. Using this method 2D-TLC maps for several modified RNA bases have been prepared (Grosjean et al., 2004; Barciszewska et al., 2007; Zhong et al., 2008). The sensitivity of the method can be increased by using radioactive (³²P) labeling [site-specific cleavage and radioactive labeling, ligation-assisted extraction, and thin layer chromatography (SCARLET)] to detect the modification within an individual transcript (Liu et al., 2013). However, the TLC-based method fails to provide information about the location/context of the modified base.

High-Performance Liquid Chromatography and MS

The content of modified RNA nucleobases can also be determined in digested mononucleosides by using high/ultra-performance liquid chromatography (UPLC)

TABLE 2 | Techniques for detection of modified RNA base.

Method/technique	Base modification	Detection principle	References
Thin layer chromatography (TLC*), SCARLET	m ⁶ A, m ⁵ C	Difference in the net charge, polarity, and hydrophobicity. Radioactive (³² P) labeling increases sensitivity of the SCARLET technique.	Grosjean et al., 2004; Barciszewska et al., 2007; Liu et al., 2013
HPLC, LC-MS/MS*, Dot-blot*	m ⁶ A, m ¹ A, m ⁵ C, hm ⁵ C	The RNA is digested into mononucleotides and detected on HPLC using UV light or mass spectrometry. In case of LC-MS/MS, modified base is quantified using the nucleoside-to-base ion mass transition. In dot-blot (a semiquantitative method), modified base-specific (e.g., anti-m ⁵ C) antibody is used to detect the modified base.	Jia et al., 2011; Kellner et al., 2014; Huber et al., 2015; Li et al., 2016a; Shen et al., 2016; Thuring et al., 2016; Cui X. et al., 2017; Limbach and Paulines, 2017
Single-molecule real-time (SMRT) technology.	m ⁶ A, m ¹ A	The modified adenine (6-mA) can be discriminated from the unmodified adenine (A).	Vilfan et al., 2013; Dominissini et al., 2016
Chemical pretreatment approach, ICE-Seq	Inosine (I)	Acrylonitrile treatment causes inosine-specific cyanoethylation leading to the truncation of reverse transcription, allowing inosine (I) sites to be detected by subsequent RNA-sequencing.	Sakurai et al., 2014; Suzuki et al., 2015
Modification-specific RT signature technique	Inosine	The modified nucleotide leaves specific signatures in the cDNA sequences, which cause either abortive primer extension and/or misincorporation at or around the modified site.	Levanon et al., 2004
	m ¹ A	Modified nucleotide affects cDNA synthesis either due to its inability to base-pair with its regular partner or by slowing down the rate of cDNA synthesis due to its massive or highly hydrophobic structure.	Hauenschild et al., 2015; Li X. et al., 2017
Biological/chemical induction of modification-specific RT signature	Pseudouridine (ψ),	Pseudouridine reacts with carbodiimide (CMCT) and forms a stable adduct, while U-CMC adducts are removed by alkaline treatment. The resulting ψ-CMC generates RT-arrest, which is detectable in the sequencing profile.	Schaefer et al., 2009
	m ⁵ C	5-mC is RT silent, but it is insensitive to bisulfite deamination. Cytosine (C) residue is deaminated into Uracil due to bisulfite treatment. The presence of C is detected by sequencing, wherein it is replaced by uracil.	Edelheit et al., 2013
	N ⁶ , 2'-O-dimethyladenosine (m ⁶ Am)	Ribose 2'-O-methylation protects the 3'-adjacent phosphodiester bond from alkaline cleavage which is used to identify the 2'-O-methylation site in RNA.	Marchand et al., 2016
Antibody-based method, MeRIP-isolated by crosslinking immunoprecipitation-seq (MeRIP-iCIP), MeRIP-qPCR*, MeRIP-seq*	m ⁶ A, m ⁵ C, hm ⁵ C, m ¹ A	Modification-specific (anti-6-mA) antibody used to immunoprecipitate short RNA fragments, followed by cDNA libraries preparation and sequencing.	Dominissini et al., 2012; Meyer et al., 2012; Chen et al., 2015; Cui et al., 2016; Delatte et al., 2016; Shen et al., 2016; Li X. et al., 2017
Modified bisulfite (BS-seq*) strategy	m ⁵ C	Bisulfite treatment converts unmodified cytosine (C) to uracil, but 5-mC remains unchanged. The presence of C is detected by sequencing, wherein it is replaced by uracil.	Schaefer et al., 2009; Squires et al., 2012; David et al., 2017
N-cyclohexyl-N'-β-methylcarbodiimide (CMC-seq)	Ψ	CMC specifically labels Ψ forming CMC-Ψ adducts which stop RT at one nucleotide 3' to the labeled Ψ site, thereby allows base-resolution detection of Ψ.	Schwartz et al., 2014a
Antibody-free method, MAZTER-seq, m ⁶ A-REF-seq, DART-seq, m ⁶ A-label-seq, m ⁶ A-SEAL*	m ⁶ A	Endoribonuclease-based RNA digestion with m ⁶ A-sensitive RNase (MazF) at unmethylated ACA motif followed by sequencing (MAZTER-seq). In m ⁶ A-SEAL-seq method, DTT-mediated thiol-addition and FTO-mediated oxidation of m ⁶ A to hm ⁶ A as chemical labeling is utilized.	Garcia-Campos et al., 2019; Meyer, 2019; Zhang Z. et al., 2019; Shu et al., 2020; Wang et al., 2020

*Method used for detection of modified base in RNA in plant.

followed by mass spectrometry (MS) (Thuring et al., 2016). The method has been extensively used earlier for the detection and quantification of modified RNA bases. MS coupled to a nano-chromatography system reduces the amount (to picomole) of the sample required. Detection of a modified RNA base through ‘matrix-assisted laser desorption/ionization–time-of-flight’ (MALDI-TOF) is still being optimized (Schwartz and Motorin, 2017). Detection of the modified RNA bases can also be performed using methods like dot-blot and LC-MS/MS (Jia et al., 2011; Delatte et al., 2016; Shen et al., 2016; Cui Q. et al., 2017). Ross et al. (2016) reported LC-MS/MS based, sequence specific detection of modified nucleosides in tRNAs from bacteria and human.

Reverse Transcription-Based Techniques

Reverse transcription (RT)-based techniques use the primer-extension method to reveal the modified RNA bases. The presence of the modified base in mRNA interrupts/inhibits primer extension, which facilitates its context-specific positioning. However, comparative sequence analysis with unmodified RNA transcripts is necessary to eliminate structural RT-stops. The advantage of the RT-based technique includes its applicability and sensitivity to a complex mixture of mRNAs, but it requires pure and concentrated mRNA molecules. Li X. et al. (2017) used a technique named m¹A-MAP to detect the modified base at single-nucleotide resolution to profile m¹A in the human transcriptome. Since m¹A causes truncation and/or misincorporation during cDNA synthesis from the transcript (Hauenschild et al., 2015), a more precise method in detecting the position of m¹A at single-base resolution. Nevertheless, inosine cannot be detected directly by the RT-based technique, as it can base-pair with cytosine. However, inosine-specific cyanoethylation treatment, using acrylonitrile, converts inosine into N¹-cyanoethylinosine (ce¹I) which disable base pairing of inosine with cytosine. This allows inosine sites to be detected by subsequent ‘inosine chemical erasing’ (ICE)-sequencing (ICE-Seq) in mRNA (Suzuki et al., 2015).

NGS Technologies-Based Method for Detection of RNA Modification

The advances in next-generation high-throughput sequencing technologies for the detection of RNA base modifications have considerably improved the epitranscriptomic studies (Li et al., 2016b). Currently, the sequencing technologies employ an amplification step to generate clusters, which provides exceptionally high sequencing output with <0.1% error. However, the read length remains shorter (500–600 nt) in most of the cases. The BS-seq method combines bisulfite conversion followed by NGS to map m⁵C, which has been successfully used for epitranscriptomic analyses in several animals and plants (Squires et al., 2012; David et al., 2017). Although the BS-seq method precisely identifies the site of m⁵C at single-base resolution (**Figure 3A**), it possesses two technical disadvantages. First, the BS-seq fails to distinguish between m⁵C and other modified cytosine bases (e.g., hm⁵C) in mRNA (Nestor et al., 2010). Second, bisulfite treatment

during sample preparation causes degradation of mRNA and thus impedes amplification of m⁵C-containing mRNA which limits the applicability of this method.

The single-molecule sequencing approach uses either of the two NGS principles. While single-molecule real-time (SMRT) technology uses nanowell (zero-mode waveguide, ZMW; Pacific BioSciences) (Vilfan et al., 2013), Nanopore sequencing (Min-ION, Oxford Nanopore Technologies) uses the change in electrical charge to detect the modified base in mRNA passing through nanopore-forming proteins (Liu et al., 2019). One of the advantages of single-molecule sequencing approaches is very long (>10,000 nt) read length, but the accuracy of the sequence is compromised. Such technologies are useful for analyzing modified bases in mRNA, particularly in a context-specific manner, as these methods allow direct sequencing of mRNA without converting it into DNA (such conversion causes the loss of modified base). However, these technologies require extensive optimization for their use in the detection of modified bases in RNA. The first experimentation on using reverse transcriptase (instead of a DNA polymerase) as the enzyme in ZMW of SMRT for direct sequencing of modified bases in mRNA was carried out by Vilfan et al. (2013). The entry of a modified base (m⁶A), present in the template mRNA, into the nanowell (ZMW) causes increased ‘inter-pulse duration’ (IPD) compared to its unmodified counterpart (adenosine). The potential of SMRT sequencing to detect m⁶A was demonstrated by Vilfan et al. (2013). Similarly, Nanopore sequencing has successfully been used to detect m⁶A in native RNA (Liu et al., 2019). Further optimization of the single-molecule sequencing technologies will revolutionize epitranscriptomic research on modified bases in mRNA.

Recently, Fang et al. (2020) reported a new method CRISPR integrated gRNA and reporter sequencing (CIGAR-seq) by combining pooled CRISPR screen and the reporters associated with RNA modification. Using the CIGAR-seq method, they could discover NSUN6 as a novel m⁵C methyltransferase in mRNA. Subsequently, they could demonstrate that this method can be successfully used to identify the regulators of other mRNA modifications such as m¹A.

Antibody-Based Methods for Detection of Modified Bases

RNA base modification, particularly m⁶A, is a widespread epitranscriptomic change that influences nearly every aspect of mRNA biology. Our understanding of the RNA base modification has been facilitated by the recent developments in the use of an antibody to immunoprecipitate RNAs containing modified base, and the high-throughput sequencing technologies. Methyl-RNA-immunoprecipitation-sequencing (MeRIP-seq) (Meyer et al., 2012) and m⁶A-seq (Dominissini et al., 2012) use immunoprecipitation with the help of modification (m⁶A or m⁵C)-specific antibody followed by sequencing (**Figure 3B**) (Yang Y. et al., 2019). Similarly, hMeRIP-seq relies on the anti-hm⁵C antibody to detect hm⁵C in *Drosophila* mRNA

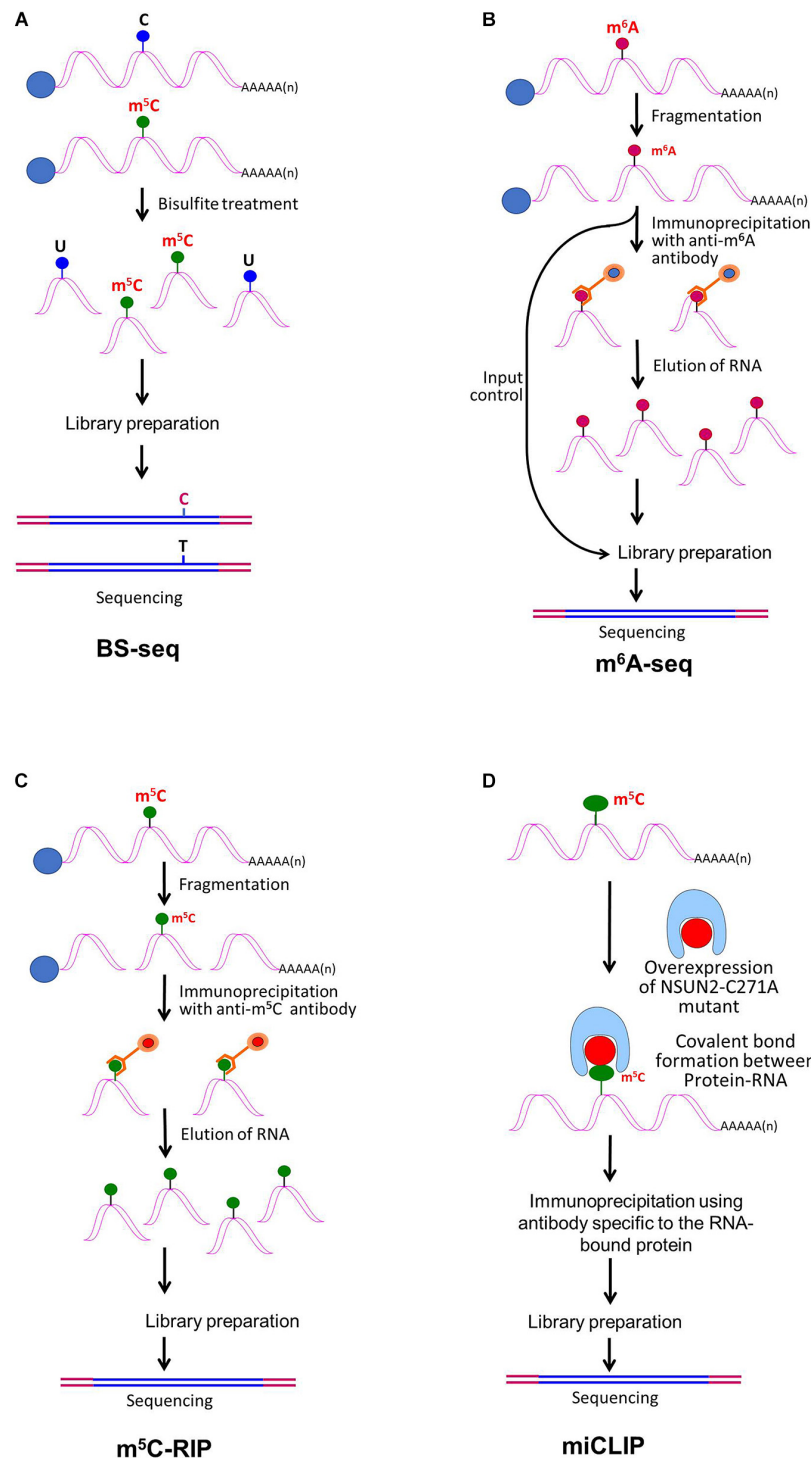


FIGURE 3 | Detection of modified bases in mRNA. **(A)** Bisulfite sequencing (BS-seq) for the detection of 5-methylcytosine (m⁵C). Purified mRNA is fragmented into small (100–200 nt) fragments, and subjected to bisulfite treatment. Bisulfite treatment causes converts cytosine (C) to uracil (U), but m⁵C remains unchanged. Presence of C is detected by sequencing, wherein it is replaced by T. **(B)** Purified mRNAs are fragmented into 100–200 nt, followed by immunoprecipitation using anti-m⁶A antibody to enrich the sample with fragments containing the modified base, library preparation, and high-throughput deep-sequencing for detection of m⁶A. **(C)** Purified mRNAs are fragmented followed by immunoprecipitation using anti-m⁵C antibody of the fragments containing the modified base, library preparation, and sequencing. **(D)** m⁵C individual-nucleotide-resolution crosslinking and immunoprecipitation (m⁵C-miCLIP) exploits catalytic activity of cysteine-to-alanine mutation (C271A) mutant of NSUN2 (methyltransferase) which inhibits release of the enzyme from the protein–RNA complex making stable covalent bond between NSun2 and its RNA targets. Antibody specific to the RNA bound protein is used for immunoprecipitation, followed by library preparation and sequencing. This allows detection of low-abundance methylated RNAs without the need of deep sequencing.

(Delatte et al., 2016). acRIP-Seq uses an ac^4C -specific antibody to identify 4,000 ac^4C in the human transcriptome (Arango et al., 2018). However, these antibody-based detection methods cannot detect hm^5C and ac^4C at single-base resolution, but the success in the single-base resolution of m^6A by sequencing might help to optimize the method (Yuan et al., 2019). This enables studying the dynamics of epitranscriptome, a post-transcriptional regulatory mechanism for gene expression. The modification-specific antibody is used for enrichment/collection of the sample with the fragments containing the modified base. To detect m^5C in mRNA, bisulfite-based technique cannot be used with much success. Hence m^5C RNA immunoprecipitation (m^5C -RIP) was used (Figure 3C) by Edelheit et al. (2013). In this method, an anti- m^5C antibody is used to immunoprecipitate and enrich the modified base containing mRNA fragments, followed by library preparation and sequencing. Methylation individual-nucleotide-resolution crosslinking and immunoprecipitation' (miCLIP) was used (Figure 3D) to identify m^5C in RNA (Hussain et al., 2013; Khoddami and Cairns, 2013). This approach exploits the enzymatic activity of m^5C methyltransferase containing a cysteine-to-alanine mutation (C271A) in NSUN2 which inhibits release of the enzyme from protein-RNA complex. This results in a covalent bond between the enzyme and its RNA targets. Antibody specific to the RNA-bound protein is used to immunoprecipitate the fragments containing the modified base, followed by library preparation and sequencing. The immunoprecipitation allows detection of methylated bases/RNAs in low-abundance without the need of deep sequencing. Subsequently, miCLIP was used to map m^6A at single-base resolution (Linder et al., 2015).

Antibody-Free Sequencing Methods

Many of the RNA base modification detection methods rely on the use of antibodies for immunoprecipitation. However, an antibody may fail to distinguish between two different modified forms of a nucleobase, such as m^6A and m^6Am . Moreover, the methods are dependent on the specificity of the antibody, which emphasizes the desire for the antibody-free method to draw transcriptome-wide atlas of the modified base. Hong et al. (2018) developed an antibody-independent method to detect m^6A at single-nucleotide resolution via 4SedTTP incorporation and FTO demethylation. Since the 4SedTTP stably base pair with A but cause truncation on m^6A -T pairing during reverse transcription. The RT stop signals of RNA with/without FTO treatment is then compared to determine the exact sites of m^6A . Recently, endoribonuclease-based RNA digestion with m^6A -sensitive RNase (*MazF* to cleave RNA at unmethylated ACA motifs) followed by sequencing (MAZTER-seq) (Garcia-Campos et al., 2019), and m^6A -sensitive RNA-endoribonuclease-facilitated sequencing (m^6A -REF-seq) (Zhang Z. et al., 2019) methods were used as antibody-independent methods. Another antibody-free m^6A sequencing (deamination adjacent to RNA modification targets, DART-seq) method was devised, using APOBEC1-YTH (cytidine deaminase fused with m^6A -binding YTH domain) protein which

deaminates C to U at the site adjacent to m^6A . This helps to identify m^6A sites in mRNA (Meyer, 2019). Moreover, two chemical labeling methods viz. m^6A -label-seq (Shu et al., 2020), and m^6A -SEAL (Wang et al., 2020) have also been developed. Wang et al. (2020) combined dithiothreitol (DTT)-mediated thiol-addition reaction [that converts the unstable hm^6A to stable N^6 -dithiolisitolmethyladenosine (dm^6A)] with FTO-mediated enzymatic oxidation of m^6A to hm^6A to develop FTO-assisted m^6A selective chemical labeling (m^6A -SEAL) method for detection of m^6A in mRNA. In a transcriptome-wide m^6A -SEAL-seq analysis, they could identify 8,605 m^6A in human embryonic kidney and 12,297 m^6A in rice leaf. Currently, most of the epitranscriptomic studies employ a detection method with NGS technology for context-specific mapping of the modified base at single-base resolution.

CHALLENGES IN THE DETECTION OF MODIFIED BASES

A major challenge in detection of the modified mRNA base has been the relatively low count of the modified base within the vast mRNA repertoire. Another challenge is the precise quantification and mapping of modified RNA residues at a single-nucleotide level. Additional challenge stems from substantial background signals often present in the maps prepared. The inability of the technique to discriminate between misincorporation/RT-stop due to the modified base and background-pause/misincorporation either due to RNA structure, RT-error, or technical errors of sequencing platform (Schwartz and Motorin, 2017). Since the same antibody can recognize both m^6A (in RNA) and 6-mA (in DNA), the contaminating DNA must be removed to get the real level of the modified base (Liang et al., 2020). Besides, there are many other limitations including intrinsic bias on secondary structures. For example, the m^6A specific-antibody fails to distinguish between m^6A and m^6Am (Schwartz et al., 2013; Linder et al., 2015). Although CMC-based Ψ sequencing has been successful in identifying Ψ at the single-base resolution, it has been associated with the problem of RNA degradation because of the alkaline treatment step (see Zhao et al., 2020). Moreover, current sequencing technologies have not been able to detect hm^5C and m^1A , particularly at the single-base resolution, which limits the functional characterization of these modified bases. Some of the challenges in the detection of modified RNA bases at technological, experimental, and analytical levels are described here.

Sequencing by synthesis approach has many restrictions in detecting the base modification. The specific antibody or chemical required for the detection of a modified base (indirect detection of the modified base) is known for only a limited number of modifications, which may show cross-reactivity. The antibody-based immunoprecipitation (IP) sequencing method (e.g., m^6A -seq or MeRIP) uses a 100–200 nt mapping window which fails to precisely identify m^6A sites (Molinie et al., 2016). The photocrosslinking-assisted m^6A sequencing (PA- m^6A -seq), m^6A individual-nucleotide-resolution crosslinking

and immunoprecipitation (miCLIP), and UV-CLIP techniques suffer from low crosslinking yield and use an indirect method to infer m⁶A sites. The location of m⁶A is inferred near the antibody crosslinking point (the tyrosine residue of antibody and RNA base), but the crosslinking point might be at varying distance from the m⁶A-binding sites, which creates difficulty in precise identification of the m⁶A site, particularly when m⁶A occurs in a cluster (Meyer et al., 2012; Linder et al., 2015). Even the direct detection methods like SMRT face certain challenges such as the ZMW stumbles when a stretch m⁶A gets incorporated; hence, the current throughput level is too low for transcriptome-wide analysis.

Careful selection/inclusion of input controls for base modification mapping is crucial. Mostly, well-known modified bases in rRNA/tRNA serve as intrinsic controls for evaluating the sensitivity and specificity of analytical methods. The conventional transcriptome analysis uses millions of cells from a tissue, epitranscriptome being highly dynamic, the cell-specific analysis would be necessary for the detection/quantification of a particular base modification and its functional characterization (Helm and Motorin, 2017). Furthermore, biological and/or technical replicates (at least 2–3) are very important to filter out falsely-detected sites, as well as to assess the robustness/reproducibility of the detection and quantification method.

FUTURE PERSPECTIVES

Recent studies have provided unprecedented mechanistic insights into RNA base modifications, and NGS-based technologies for detecting RNA base modifications are further improving the scenario. Modern chemical biology tools would be applied to expedite the epitranscriptomic studies. High-throughput technologies to simultaneously identify different modified bases in the same RNA molecule will have considerable applicability, as modified bases may have cumulative effects on regulating biological functions. Plants provide a unique system to elucidate the biological functions of modified RNA bases and their regulatory aspects through investigating epitranscriptomic alterations in higher eukaryotes, which are otherwise difficult to be elucidated using an animal system (Shen et al., 2019). Using a combination of techniques including genetic ablation and NGS-based mapping, the regulatory roles of the epitranscriptome in several developmental processes in plants have been demonstrated. Compared with the writers and erasers, readers for the modified base play a more significant role in responses to environmental stresses. This suggests that deciphering the location/context of epitranscriptomic marks is more important than merely detecting the changes (writing/erasing) in the marks for improved stress adaptation in plants. Therefore, it is important to characterize the role of reader proteins in the epitranscriptomic regulation of gene expression under environmental stresses (Hu et al., 2019). Association between epitranscriptome and stress responses in plants indicates that such epitranscriptomic marks might be utilized in the future as important epimarks for the development of stress-tolerant

crop plants (Vandivier and Gregory, 2018). In the line of the success in the detection of m⁶A at single-base resolution, a similar sequencing method would be optimized for hm⁵C and m¹A mapping. Nevertheless, for functional characterization of epitranscriptomic modifications, quantification of the absolute stoichiometry of RNA modifications is crucial.

However, several questions need to be answered before we can devise appropriate strategies to better utilize the epitranscriptomic information. Some of these include, why only selected mRNAs get modified? Why are only certain adenosine/m⁶A or cytosine/m⁵C at selected sites gets methylated/demethylated? How does the modified base affect downstream mRNA processing? How do different readers recognize their targets? How are the writer, reader, and eraser for a nucleobase get co-ordinately regulated by the developmental/environmental signal? Even if we get answers to some of these questions, several other questions would require to be answered. For example, how do the different mRNA base modifications influence the dynamics/function of each other? Experiments designed to answer some of these questions are underway in laboratories worldwide, and we expect that the next 5 years of research in epitranscriptomics would be more exciting than the past!

CONCLUSION

During the past few years, many RNA modifications and their functional versatility could be discovered due to the advances in chemogenetic RNA labeling techniques, high-throughput NGS, and functional validation. Several other dynamic base modifications in mRNA are also being identified, which would require functional characterization for advances in epitranscriptomics. Numerous other epitranscriptomic modifications may be identified in the future which may show interaction with other modified bases in modulating metabolic pathways. The biological functions of several mRNA base modifications are still poorly understood, their detection at single-base resolution using technological advancements such as nanowell (SMRT) and nanopore (Oxford Nanopore) sequencing is very much promising. However, proper experimental design with a sufficient number of replications, and inclusion of controls would be very important to rule out false-positive results and for the highest confidence level. Moreover, identifying the enzyme(s) involved in modification of RNA base (reader), and replacing it with an unmodified base (eraser) is necessary for devising strategies to manipulate the expression of a gene. However, several fundamental questions remain to be answered, including whether modified bases are conserved among plant species. Answering these questions would substantially improve our knowledge of epitranscriptomics and its effects on plant growth, fitness, and survival under environmental stress. Such investigations, particularly comprehensive studies to demonstrate a linkage between epigenetic and epitranscriptomic regulations, would offer potential new strategies for the manipulation of crop plants with better plasticity/adaptability to the changing climatic conditions. Comprehensive studies on the

correlation between epigenetic and epitranscriptomic regulation of gene expression might provide some newer aspects (Song and Yi, 2017) for the manipulation of a trait through epigenome/epitranscriptome editing to develop climate-smart crop plants for the 21st century (Kumar, 2019).

AUTHOR CONTRIBUTIONS

SK and TM conceived the review. SK prepared the manuscript. SK and TM revised the manuscript and approved the final

draft. Both authors contributed to the article and approved the submitted version.

FUNDING

The epigenomics and epitranscriptomics research are being carried out with financial supports from National Agricultural Science Fund (NASF/ABP-70161/2018-19), and an Extramural Research grant [18(3)/2018-O&P] from the Indian Council of Agricultural Research, Government of India, New Delhi.

REFERENCES

- Adachi, H., De Zoysa, M. D., and Yu, Y. T. (2019). Post-transcriptional pseudouridylation in mRNA as well as in some major types of noncoding RNAs. *Biochem. Biophys. Acta Gene Regul. Mech.* 1862, 230–239. doi: 10.1016/j.bbaggm.2018.11.002
- Aguilo, F., Zhang, F., Sancho, A., Fidalgo, M., Cecilia, S. D., Vashisht, A., et al. (2015). Coordination of m6A mRNA methylation and gene transcription by ZFP217 regulates pluripotency and reprogramming. *Cell Stem Cell* 17, 689–704. doi: 10.1016/j.stem.2015.09.005
- Alarcon, C. R., Goodarzi, H., Lee, H., Liu, X., Tavazoie, S., and Tavazoie, S. F. (2015a). HNRNPA2B1 is a mediator of m6A-dependent nuclear RNA processing events. *Cell* 162, 1299–1308. doi: 10.1016/j.cell.2015.08.011
- Alarcon, C. R., Lee, H., Goodarzi, H., Halberg, N., and Tavazoie, S. F. (2015b). N6-methyladenosine marks primary microRNAs for processing. *Nature* 519, 482–485. doi: 10.1038/nature14281
- Anderson, S. J., Kramer, M. C., Gosai, S. J., Yu, X., Vandivier, L. E., and Nelson, A. D. L. (2018). N6-methyladenosine inhibits local ribonucleolytic cleavage to stabilize mRNAs in Arabidopsis. *Cell Rep.* 25, 1146–1157. doi: 10.1016/j.celrep.2018.10.020
- Angelova, M. T., Dimitrova, D. G., Dinges, N., Lence, T., Worpenberg, L., and Carre, C. (2018). The emerging field of epitranscriptomics in neurodevelopmental and neuronal disorders. *Front. Bioeng. Biotechnol.* 6:46. doi: 10.3389/fbioe.2018.00046
- Arango, D., Sturgill, D., Alhusaini, N., Dillman, A. A., Sweet, T. J., and Hanson, G. (2018). Acetylation of cytidine in mRNA promotes translation efficiency. *Cell* 175, 1872–1886. doi: 10.1016/j.cell.2018.10.030
- Arimbasseri, A. G., Iben, J., Wei, F. Y., Rijal, K., Tomizawa, K., Hafner, M., et al. (2016). Evolving specificity of tRNA 3-methyl-cytidine-32 (m3C32) modification: a subset of tRNAs^{ser} requires N6-isopentenylolation of A37. *RNA* 22, 1400–1410. doi: 10.1261/rna.056259.116
- Arribas-Hernandez, L., Bressendorff, S., Hansen, M. H., Poulsen, C., Erdmann, S., and Brodersen, P. (2018). An m6A-YTH module controls developmental timing and morphogenesis in Arabidopsis. *Plant Cell* 30, 952–967. doi: 10.1105/tpc.17.00833
- Barciszewska, M. Z., Barciszewska, A. M., and Rattan, S. I. S. (2007). TLC-based detection of methylated cytosine: application to aging epigenetics. *Biogerontology* 8, 673–678. doi: 10.1007/s10522-007-9109-3
- Batista, P. J., Molin, B., Wang, J., Qu, K., Zhang, J., Li, L., et al. (2014). m6A RNA modification controls cell fate transition in mammalian embryonic stem cells. *Cell Stem Cell* 15, 707–719. doi: 10.1016/j.stem.2014.09.019
- Belanger, F., Stepinski, J., Darzynkiewicz, E., and Pelletier, J. (2010). Characterization of hMT1, a human Cap1 2'-O-ribose methyltransferase. *J. Biol. Chem.* 285, 33037–33044. doi: 10.1074/jbc.M110.155283
- Boccaletto, P., Machnicka, M. A., Purta, E., Piątkowski, P., Bagiński, B., and Wirecki, T. K. (2018). MODOMICS: a database of RNA modification pathways. 2017 update. *Nucleic Acids Res.* 46, 303–307. doi: 10.1093/nar/gkx1030
- Bodi, Z., Zhong, S., Mehra, S., Song, J., Li, H., Graham, N., et al. (2012). Adenosine methylation in Arabidopsis mRNA is associated with the 3' end and reduced levels cause developmental defects. *Front. Plant Sci.* 3:48. doi: 10.3389/fpls.2012.00048
- Bohnsack, K. E., Höbartner, C., and Bohnsack, M. T. (2019). Eukaryotic 5-methylcytosine (m5C) RNA methyltransferases: mechanisms, cellular functions, and links to disease. *Genes* 10:102. doi: 10.3390/genes10020102
- Bokar, J. A., Shambaugh, M. E., Polayes, D., Matera, A. G., and Rottman, F. M. (1997). Purification and cDNA cloning of the AdoMet-binding subunit of the human mRNA (N6-adenosine)-methyltransferase. *RNA* 3, 1233–1247.
- Boo, S. H., and Kim, Y. K. (2020). The emerging role of RNA modifications in the regulation of mRNA stability. *Expt. Mol. Med.* 52, 400–408. doi: 10.1038/s12276-020-0407-z
- Boulias, K., Toczyłowska-Socha, D., Hawley, B. R., Liberman, N., Takashima, K., and Zaccara, S. (2019). Identification of the m(6)Am methyltransferase PCIF1 reveals the location and functions of m(6)Am in the transcriptome. *Mol. Cell* 75, 631–643.e638. doi: 10.1016/j.molcel.2019.06.006
- Burgess, A., David, R., and Searle, I. R. (2016). Deciphering the epitranscriptome: a green perspective. *J. Integr. Plant Biol.* 58, 822–835. doi: 10.1111/jipb.12483
- Cahová, H., Winz, M. L., Höfer, K., Nübel, G., and Jäschke, A. (2015). NAD captureSeq indicates NAD as a bacterial cap for a subset of regulatory RNAs. *Nature* 519, 374–377. doi: 10.1038/nature14020
- Carille, T. M., Rojas-Duran, M. F., Zinshteyn, B., Shin, H., Bartoli, K. M., and Gilbert, W. V. (2014). Pseudouridine profiling reveals regulated mRNA pseudouridylation in yeast and human cells. *Nature* 515, 143–146. doi: 10.1038/nature13802
- Chen, K., Lu, Z., Wang, X., Fu, Y., and Luo, G. Z. (2015). High-resolution N6-methyladenosine (m6A) map using photo-crosslinking-assisted m6A sequencing. *Angew. Chem. Int. Ed. Engl.* 54, 1587–1590. doi: 10.1002/anie.201410647
- Chen, P., Jäger, G., and Zheng, B. (2010). Transfer RNA modifications and genes for modifying enzymes in *Arabidopsis thaliana*. *BMC Plant Biol.* 10:201. doi: 10.1186/1471-2229-10-201
- Chen, R. X., Chen, X., Xia, L. P., Zhang, J. X., Pan, Z. Z., Ma, X. D., et al. (2019). N6-methyladenosine modification of circNSUN2 facilitates cytoplasmic export and stabilizes HMGA2 to promote colorectal liver metastasis. *Nat. Commun.* 10:4695. doi: 10.1038/s41467-019-12651-2
- Choi, J., Jeong, K. W., Demirci, H., Chen, J., Petrov, A., Prabhakar, A., et al. (2016). N6-methyladenosine in mRNA disrupts tRNA selection and translation-elongation dynamics. *Nat. Str. Mol. Biol.* 23, 110–115. doi: 10.1038/nsmb.3148
- Chujo, T., and Suzuki, T. (2012). Trmt61B is a methyltransferase responsible for 1-methyladenosine at position 58 of human mitochondrial tRNAs. *RNA* 18, 2269–2276. doi: 10.1261/rna.035600.112
- Cui, Q., Shi, H., Ye, P., Li, L., Qu, Q., Sun, G., et al. (2017). m6A RNA methylation regulates the self-renewal and tumorigenesis of glioblastoma stem cells. *Cell Rep.* 18, 2622–2634. doi: 10.1016/j.celrep.2017.02.059
- Cui, X., Liang, X., Shen, L., Zhang, Q., Bao, S., and Geng, Y. (2017). 5-Methylcytosine RNA methylation in *Arabidopsis thaliana*. *Mol. Plant* 10, 1387–1399. doi: 10.1016/j.molp.2017.09.013
- Cui, X., Meng, J., Zhang, S., Chen, Y., and Huang, Y. (2016). A novel algorithm for calling mRNA m6A peaks by modeling biological variances in MeRIP-seq data. *Bioinformatics* 32, 378–385. doi: 10.1093/bioinformatics/btw281
- David, R., Burgess, A., Parker, B., Li, J., Pulsford, K., and Sibbritt, T. (2017). Transcriptome-wide mapping of RNA 5-methylcytosine in Arabidopsis mRNAs and noncoding RNAs. *Plant Cell* 29, 445–460. doi: 10.1105/tpc.16.00751

- Davis, F. F., and Allen, F. W. (1957). Ribonucleic acids from yeast which contain a fifth nucleotide. *J. Biol. Chem.* 227, 907–915. doi: 10.1016/S0021-9258(18)70770-9
- de Almeida, C., Scheer, H., Zuber, H., and Gagliardi, D. (2018). RNA uridylation: a key posttranscriptional modification shaping the coding and noncoding transcriptome. *Wiley Interdiscip. Rev. RNA* 9:e1440. doi: 10.1002/wrna.1440
- Delatte, B., Wang, F., Ngoc, L. V., Collignon, E., Bonvin, E., Deplus, R., et al. (2016). Transcriptome-wide distribution and function of RNA hydroxymethylcytosine. *Science* 351, 282–285. doi: 10.1126/science.aac5253
- Dominissini, D., Moshitch-Moshkovitz, S., Schwartz, S., Salmon-Divon, M., Ungar, L., and Osenberg, S. (2012). Topology of the human and mouse m6A RNA methylomes revealed by m6A-seq. *Nature* 485, 201–206. doi: 10.1038/nature11112
- Dominissini, D., Nachtergaele, S., Moshitch-Moshkovitz, S., Peer, E., Kol, N., Ben-Haim, M. S., et al. (2016). The dynamic N1-methyladenosine methylome in eukaryotic messenger RNA. *Nature* 530, 441–446. doi: 10.1038/nature16998
- Dong, C., Niu, L., Song, W., Xiong, X., Zhang, X., and Zhang, Z. (2016). tRNA modification profiles of the fast-proliferating cancer cells. *Biochem. Biophys. Res. Commun.* 476, 340–345. doi: 10.1016/j.bbrc.2016.05.124
- Du, H., Zhao, Y., He, J., Zhang, Y., Xi, H., Liu, M., et al. (2016). YTHDF2 destabilizes m6A-containing RNA through direct recruitment of the CCR4–NOT deadenylase complex. *Nat. Commun.* 7:12626. doi: 10.1038/ncomms12626
- Duan, H. C., Wei, L. H., Zjang, C., Wang, Y., Chen, L., and Lu, Z. (2017). ALKBH10B is an RNA N6-methyladenosine demethylase affecting Arabidopsis floral transition. *Plant Cell* 29, 2995–3011. doi: 10.1105/tpc.16.00912
- Edelheit, S., Schwartz, S., Mumbach, M. R., Wurtzel, O., and Sorek, R. (2013). Transcriptome-wide mapping of 5-methylcytidine RNA modifications in bacteria, archaea, and yeast reveals m5C within archaeal mRNAs. *PLoS Genet.* 9:e1003602. doi: 10.1371/journal.pgen.1003602
- Engel, M., Roeh, S., Eggert, C., Kaplick, P. M., Tietze, L., and Namendorf, C. (2017). The role of m6A-RNA methylation in stress response regulation. *BioRxiv [Preprint]* doi: 10.1101/200402
- Fang, L., Wang, W., Li, G., Zhang, L., Li, J., Gan, D., et al. (2020). CIGAR-seq, a CRISPR/Cas-based method for unbiased screening of novel mRNA modification regulators. *Mol. Syst. Biol.* 16:e10025. doi: 10.15252/msb.202010025
- Flores, J. V., Cordero-Espinoza, L., Oetzuerk-Winder, F., Andersson-Rolf, A., Selmi, T., Blanco, S., et al. (2017). Cytosine-5 RNA methylation regulates neural stem cell differentiation and motility. *Stem Cell Rep.* 8, 112–124. doi: 10.1016/j.stemcr.2016.11.014
- Frindert, J., Zhang, Y., Nübel, G., Kahloon, M., Kolmar, L., Hotz-Wagenblatt, A., et al. (2018). Identification, biosynthesis, and decapping of NAD-capped RNAs in *B. subtilis*. *Cell Rep.* 24, 1890–1901. doi: 10.1016/j.celrep.2018.07.047
- Frye, M., Harada, B., Behm, M., and He, C. (2018). RNA modifications modulate gene expression during development. *Science* 361, 1346–1349. doi: 10.1126/science.aau1646
- Fu, L., Guerrero, C. R., Zhong, N., Amato, N. J., Liu, Y., and Liu, S. (2014). Tet-mediated formation of 5-hydroxymethylcytosine in RNA. *J. Am. Chem. Soc.* 136, 11582–11585. doi: 10.1021/ja505305z
- Fu, Y., Dominissini, D., Rechavi, G., and He, C. (2014). Gene expression regulation mediated through reversible m6A RNA methylation. *Nat. Rev. Genet.* 15, 293–306. doi: 10.1038/nrg3724
- Garcia-Campos, M. A., Edelheit, S., Toth, U., Safra, M., Shachar, R., and Viukov, S. (2019). Deciphering the "m6A code" via antibody-independent quantitative profiling. *Cell* 3, 731–747. doi: 10.1016/j.cell.2019.06.013
- Geula, S., Moshitch-Moshkovitz, S., Dominissini, D., Mansour, A. A., Kol, N., Salmon-Divon, M., et al. (2015). m6A mRNA methylation facilitates resolution of naïve pluripotency toward differentiation. *Science* 347, 1002–1006. doi: 10.1126/science.1261417
- Grosjean, H., Keith, G., and Droogmans, L. (2004). Detection and quantification of modified nucleotides in RNA using thin-layer chromatography. *Methods Mol. Biol.* 265, 357–391. doi: 10.1385/1-59259-775-0:357
- Hauenschild, R., Tserovski, L., Schmid, K., Thëuring, K., Winz, M. L., Sharma, S., et al. (2015). The reverse transcription signature of N-1-methyladenosine in RNA-Seq is sequence dependent. *Nucleic Acids Res.* 43, 9950–9964. doi: 10.1093/nar/gkv895
- Haussmann, I. U., Bodi, Z., Sanchez-Moran, E., Mongan, N. P., Archer, N., Fray, R. G., et al. (2016). m6A potentiates Sxl alternative pre-mRNA splicing for robust *Drosophila* sex determination. *Nature* 540, 301–304. doi: 10.1038/nature20577
- Helm, M. (2006). Post-transcriptional nucleotide modification and alternative folding of RNA. *Nucleic Acids Res.* 34, 721–733. doi: 10.1093/nar/gkj471
- Helm, M., and Motorin, Y. (2017). Detecting RNA modifications in the epitranscriptome: predict and validate. *Nat. Rev. Genet.* 18, 275–291. doi: 10.1038/nrg.2016.169
- Hoernes, T. P., Clementi, N., Faserl, K., Glasner, H., Breuker, K., and Lindner, H. (2016). Nucleotide modifications within bacterial messenger RNAs regulate their translation and are able to rewire the genetic code. *Nucleic Acids Res.* 44, 852–862. doi: 10.1093/nar/gkv1182
- Holstein, J. M., Anhäuser, L., and Rentmeister, A. (2016). Modifying the 5' cap for click reactions of eukaryotic mRNA and to tune translation efficiency in living cells. *Angew. Chem. Int. Ed.* 55, 10899–10903. doi: 10.1002/anie.201604107
- Hong, T., Yuan, Y., Chen, Z., Xi, K., Wang, T., Xie, Y., et al. (2018). Precise antibody-independent m6A identification via 4SedTTP-involved and FTO-assisted strategy at single-nucleotide resolution. *J. Am. Chem. Soc.* 140, 5886–5889. doi: 10.1021/jacs.7b13633
- Hsu, P. J., Zhu, Y., Ma, H., Ghosh, Y., Shi, X., and Liu, Y. (2017). Ythdc2 is an N6-methyladenosine binding protein that regulates mammalian spermatogenesis. *Cell Res.* 27, 1115–1127. doi: 10.1038/cr.2017.99
- Hu, J., Manduzio, S., and Kang, H. (2019). Epitranscriptomic RNA methylation in plant development and abiotic stress responses. *Front. Plant Sci.* 10:500. doi: 10.3389/fpls.2019.00500
- Huang, H., Weng, H., Sun, W., Qin, X., Shi, H., and Wu, H. (2018). Recognition of RNA N6-methyladenosine by IGF2BP proteins enhances mRNA stability and translation. *Nat. Cell Biol.* 20, 285–295. doi: 10.1038/s41556-018-0045-z
- Huber, S. M., Van Delet, P., Mendil, L., Bachman, M., Smollett, K., and Werner, F. (2015). Formation and abundance of 5-hydroxymethylcytosine in RNA. *Chembiochem* 16, 752–755. doi: 10.1002/cbic.201500013
- Hull, C. M., and Bevilacqua, P. C. (2016). Discriminating self and non-self by RNA: roles for RNA structure, misfolding, and modification in regulating the innate immune sensor PKR. *Acc. Chem. Res.* 49, 1242–1249. doi: 10.1021/acs.accounts.6b00151
- Hussain, S., Sajini, A. A., Blanco, S., Dietmann, S., Lombard, P., Sugimoto, Y., et al. (2013). NSun2-mediated cytosine-5 methylation of vault noncoding RNA determines its processing into regulatory small RNAs. *Cell Rep.* 4, 255–261. doi: 10.1016/j.celrep.2013.06.029
- Ikeuchi, K., Izawa, T., and Inada, T. (2018). Recent progress on the molecular mechanism of quality controls induced by ribosome stalling. *Front. Genet.* 9:743. doi: 10.3389/fgene.2018.00743
- Ishii, T., Hayakawa, H., Igawa, T., Sekiguchi, T., and Sekiguchi, M. (2018). Specific binding of PCBP1 to heavily oxidized RNA to induce cell death. *Proc. Natl Acad. Sci. USA* 115, 6715–6720. doi: 10.1073/pnas.1806912115
- Ishii, T., Hayakawa, H., Sekiguchi, T., Adachi, N., and Sekiguchi, M. (2015). Role of Aul1 in elimination of oxidatively damaged messenger RNA in human cells. *Free Radic. Biol. Med.* 79, 109–116. doi: 10.1016/j.freeradbiomed.2014.11.018
- Jackman, J. E., and Alfonzo, J. D. (2013). Transfer RNA modifications: nature's combinatorial chemistry playground. *Wiley Interdiscip. Rev. RNA* 4, 35–48. doi: 10.1002/wrna.1144
- Jia, G. F., Fu, Y., Zhao, X., Dai, Q., Zheng, G., and Yang, Y. (2011). N6-methyladenosine in nuclear RNA is a major substrate of the obesity-associated FTO. *Nat. Chem. Biol.* 7, 885–887. doi: 10.1038/nchembio.687
- Jiao, X., Doamekpor, S. K., Bird, J. G., Nickels, B. E., Tong, L., Hart, R. P., et al. (2017). 5' End nicotinamide adenine dinucleotide cap in human cells promotes RNA decay through DXO-mediated deNADding. *Cell* 168, 1015–1027. doi: 10.1016/j.cell.2017.02.019
- Kadumuri, R. V., and Janga, S. C. (2018). Epitranscriptomic code and its alterations in human disease. *Trends Mol. Med.* 24, 886–903. doi: 10.1016/j.molmed.2018.07.010
- Karijolic, J., Yi, C., and Yu, Y. T. (2015). Transcriptome-wide dynamics of RNA pseudouridylation. *Nat. Rev. Mol. Cell Biol.* 16, 581–585. doi: 10.1038/nrm4040
- Kariko, K., Muramatsu, H., Keller, J. M., and Weissman, D. (2012). Increased erythropoiesis in mice injected with submicrogram quantities

- of pseudouridine-containing mRNA encoding erythropoietin. *Mol. Ther.* 20, 948–953. doi: 10.1038/mt.2012.7
- Kasowitz, S. D., Ma, J., Anderson, S. J., Leu, N. A., Xu, Y., Gregory, B. D., et al. (2018). Nuclear m6A reader YTHDC1 regulates alternative polyadenylation and splicing during mouse oocyte development. *PLoS Genet.* 14:e1007412. doi: 10.1371/journal.pgen.1007412
- Ke, S., Alemu, E. A., Mertens, C., Gantman, E. C., Fak, J. J., and Mele, A. (2015). A majority of m6A residues are in the last exons, allowing the potential for 30 UTR regulation. *Genes Dev.* 29, 2037–2053. doi: 10.1101/gad.269415.115
- Keith, G. (1995). Mobilities of modified ribonucleotides on two-dimensional cellulose thin-layer chromatography. *Biochimie* 77, 142–144. doi: 10.1016/0300-9084(96)88118-1
- Kellner, S., Ochel, A., Thüring, K., Spenkuch, F., Neumann, J., Sharma, S., et al. (2014). Absolute and relative quantification of RNA modifications via biosynthetic isotopomers. *Nucleic Acids Res.* 42:e142. doi: 10.1093/nar/gku733
- Khan, Z., Ford, M. J., Cusanovich, D. A., Mitrano, A., Pritchard, J. K., and Gilad, Y. (2013). Primate transcript and protein expression levels evolve under compensatory selection pressures. *Science* 342, 1100–1104. doi: 10.1126/science.1242379
- Khoddami, V., and Cairns, B. R. (2013). Identification of direct targets and modified bases of RNA cytosine methyltransferases. *Nat. Biotechnol.* 31, 458–464. doi: 10.1038/nbt.2566
- Khoddami, V., Yerra, A., Mosbrugger, T. L., Fleming, A. M., Burrows, C. J., and Cairns, B. R. (2019). Transcriptome-wide profiling of multiple RNA modifications simultaneously at single-base resolution. *Proc. Natl. Acad. Sci. USA* 116, 6784–6789. doi: 10.1073/pnas.1817334116
- Khonsari, B., and Klassen, R. (2020). Impact of Pus1 pseudouridine synthase on specific decoding events in *Saccharomyces cerevisiae*. *Biomolecules* 10:729.
- Kiledjian, M. (2018). Eukaryotic RNA 5'-end NAD⁺ capping and DeNADding. *Trends Cell Biol.* 28, 454–464. doi: 10.1016/j.tcb.2018.02.005
- Knuckles, P., Lence, T., Haussmann, I. U., Jacob, D., Kreim, N., Carl, S. H., et al. (2018). Zc3h13/Flacc is required for adenosine methylation by bridging the mRNA-binding factor Rbm15/Spenito to the m6A machinery component Wtap/Fl(2)d. *Genes Dev.* 32, 415–429. doi: 10.1101/gad.309146.117
- Kumar, S. (2019). Epigenomics for crop improvement: current status and future perspectives. *J. Genet. Cell Biol.* 3, 128–134.
- Kumar, S., Chinnusamy, V., and Mohapatra, T. (2018). Epigenetics of modified DNA bases: 5-methylcytosine and beyond. *Front. Genet.* 9:640. doi: 10.3389/fgene.2018.00640
- Kumar, S., Mapa, K., and Maiti, S. (2014). Understanding the effect of locked nucleic acid and 2'-O-methyl modification on the hybridization thermodynamics of a miRNA-mRNA pair in the presence and absence of AfPiwi protein. *Biochemistry* 53, 1607–1615. doi: 10.1021/bi401677d
- Leonardi, A., Kovalchuk, N., Yin, L., Endres, L., Evke, S., Nevins, S., et al. (2020). The epitranscriptomic writer ALKBH8 drives tolerance and protects mouse lungs from the environmental pollutant naphthalene. *Epigenetics* 15, 1121–1138. doi: 10.1080/15592294.2020.1750213
- Levanon, E. Y., Eisenberg, E., Yelin, R., Nemzer, S., Hallegger, M., and Shemesh, R. (2004). Systematic identification of abundant A-to-I editing sites in the human transcriptome. *Nat. Biotechnol.* 22, 1001–1005. doi: 10.1038/nbt996
- Li, A., Chen, Y. S., Ping, X. L., Yang, X., Xiao, W., and Yang, Y. (2017). Cytoplasmic m6A reader YTHDF3 promotes mRNA translation. *Cell Res.* 27, 444–447. doi: 10.1038/cr.2017.10
- Li, D., Zhang, H., Hong, Y., Huang, L., Li, X., Zhang, Y., et al. (2014). Genome-wide identification, biochemical characterization, and expression analyses of the YTH domain-containing RNA-binding protein family in Arabidopsis and rice. *Plant Mol. Biol. Rep.* 32, 1169–1186. doi: 10.1007/s11105-014-0724-2
- Li, X., Xiong, X., and Yi, C. (2016b). Epitranscriptome sequencing technologies: decoding RNA modifications. *Nat. Methods* 14, 23–31. doi: 10.1038/nmeth.4110
- Li, X., Xiong, X., Wang, K., Wang, L., Shu, X., Ma, S., et al. (2016a). Transcriptome-wide mapping reveals reversible and dynamic N(1)-methyladenosine methylome. *Nat. Chem. Biol.* 12, 311–316. doi: 10.1038/nchembio.2040
- Li, X., Xiong, X., Zhang, M., Wang, K., Chen, Y., Zhou, J., et al. (2017). Base-resolution mapping reveals distinct m1A methylome in nuclear- and mitochondrial-encoded transcripts. *Mol. Cell* 68, 993–1005. doi: 10.1016/j.molcel.2017.10.019
- Li, X., Zhu, P., Ma, S., Song, J., Bai, J., Sun, F., et al. (2015). Chemical pulldown reveals dynamic pseudouridylation of the mammalian transcriptome. *Nat. Chem. Biol.* 11, 592–597. doi: 10.1038/nchembio.1836
- Li, Y., Wang, X., Li, C., Hu, S., Yu, J., and Song, S. (2014). Transcriptome-wide N6-methyladenosine profiling of rice callus and leaf reveals the presence of tissue-specific competitors involved in selective mRNA modification. *RNA Biol.* 11, 1180–1188. doi: 10.4161/rna.36281
- Li, Z., Shi, J., Yu, L., Zhao, X., Ran, L., and Hu, D. (2018). N6-methyl-adenosine level in *Nicotiana tabacum* is associated with tobacco mosaic virus. *Viro. J.* 15:87. doi: 10.1186/s12985-018-0997-4
- Liang, Z., Riaz, A., Chachar, S., Ding, Y., Du, H., and Gu, X. (2020). Epigenetic modifications of mRNA and DNA in plants. *Mol. Plant* 13, 14–30. doi: 10.1016/j.molp.2019.12.007
- Lim, J., Ha, M., Chang, H., Kwon, S. C., Simanshu, D. K., Patel, D. J., et al. (2014). Uridylation by TUT4 and TUT7 marks mRNA for degradation. *Cell* 159, 1365–1376. doi: 10.1016/j.cell.2014.10.055
- Limbach, P. A., and Paulines, M. J. (2017). Going global: the new era of mapping modifications in RNA. *Wiley Interdiscip. Rev. RNA* 8:e1367. doi: 10.1002/wrna.1367
- Linder, B., Grozhik, A. V., Olarerin-George, A. O., Meydan, C., Mason, C. E., and Jaffrey, S. R. (2015). Single-nucleotide-resolution mapping of m6A and m6Am throughout the transcriptome. *Nat. Methods* 12, 767–772. doi: 10.1038/nmeth.3453
- Liu, F., Clark, W., Luo, G., Wang, X., Fu, Y., and Wei, J. (2016). ALKBH1-mediated tRNA demethylation regulates translation. *Cell* 167, 816–828. doi: 10.1016/j.cell.2016.09.038
- Liu, H., Begik, O., Lucas, M. C., Ramirez, J. M., Mason, C. E., and Wiener, D. (2019). Accurate detection of m6A RNA modifications in native RNA sequences. *Nat. Commun.* 10:4079. doi: 10.1038/s41467-019-11713-9
- Liu, J., Huang, T., Zhang, Y., Zhao, T., Zhao, X., Chen, W., et al. (2020). Sequence- and structure-selective mRNA m5C methylation by NSUN6 in animals. *Natl. Sci. Rev.* nwa273. doi: 10.1093/nsr/nwaa273
- Liu, J., Yue, Y., Han, D., Wang, X., Fu, Y., and Zhang, L. (2014). A METTL3-METTL14 complex mediates mammalian nuclear RNA N6-adenosine methylation. *Nat. Chem. Biol.* 10, 93–95. doi: 10.1038/nchembio.1432
- Liu, N., and Pan, T. (2016). N6-methyladenosine-encoded epitranscriptomics. *Nat. Struct. Mol. Biol.* 23, 98–102. doi: 10.1038/nsmb.3162
- Liu, N., Dai, Q., Zheng, G., He, C., Parisien, M., and Pan, T. (2015). N6-methyladenosine-dependent RNA structural switches regulate RNA-protein interactions. *Nature* 518, 560–564. doi: 10.1038/nature14234
- Liu, N., Parisien, M., Dai, Q., Zheng, G., He, C., and Pan, T. (2013). Probing N6-methyladenosine RNA modification status at single nucleotide resolution in mRNA and long noncoding RNA. *RNA* 19, 1848–1856. doi: 10.1261/rna.041178.113
- Lovejoy, A. F., Riordan, D. P., and Brown, P. O. (2014). Transcriptome-wide mapping of pseudouridines: pseudouridine synthases modify specific mRNAs in *S. cerevisiae*. *PLoS One* 9:e110799. doi: 10.1371/journal.pone.0110799
- Luo, G. Z., Macqueen, A., Zheng, G., Duan, H., Dore, L. C., and Lu, Z. (2014). Unique features of the m6A methylome in *Arabidopsis thaliana*. *Nat. Commun.* 5:5630. doi: 10.1038/ncomms6630
- Luo, J., Wang, Y., Wang, M., Zhang, L., Peng, H., and Zhou, Y. Y. (2019). Natural variation in RNA m6A methylation and its relationship with translational status. *Plant Physiol.* 182, 332–344. doi: 10.1104/pp.19.00987
- Luo, S., and Tong, L. (2014). Molecular basis for the recognition of methylated adenines in RNA by the eukaryotic YTH domain. *Proc. Natl. Acad. Sci. USA* 111, 13834–13839. doi: 10.1073/pnas.1412742111
- Malbec, L., Zhang, T., Chen, Y. S., Zhang, Y., Sun, B. F., Shi, B.-Y., et al. (2019). Dynamic methylome of internal mRNA N7-methylguanosine and its regulatory role in translation. *Cell Res.* 29, 927–941. doi: 10.1038/s41422-019-0230-z
- Marchand, V., Blanloeil-Oillo, F., Helm, M., and Motorin, Y. (2016). Illumina-based RiboMethSeq approach for mapping of 2'-O-Me residues in RNA. *Nucleic Acids Res.* 44:e135. doi: 10.1093/nar/gkw547
- Martinez-Perez, M., Aparicio, F., López-Gresa, M. P., Bellés, J. M., Sánchez-Navarro, J. A., and Pallás, V. (2017). Arabidopsis m6A demethylase activity modulates viral infection of a plant virus and the m6A abundance in its genomic RNAs. *Proc. Natl. Acad. Sci. USA* 14, 10755–10760. doi: 10.1073/pnas.1703139114

- Mauer, J., Luo, X., Blanjoie, A., Jiao, X., Grozhik, A. V., Patil, D. P., et al. (2017). Reversible methylation of m6Am in the 5' cap controls mRNA stability. *Nature* 541, 371–375. doi: 10.1038/nature21022
- Meyer, K. D. (2019). DART-seq: an antibody-free method for global m(6)A detection. *Nat. Methods* 16, 1275–1280. doi: 10.1038/s41592-019-0570-0
- Meyer, K. D., Patil, D. P., Zhou, J., Zinoviev, A., Skabkin, M. A., Elemento, O., et al. (2015). 5' UTR m6A promotes cap-independent translation. *Cell* 163, 999–1010. doi: 10.1016/j.cell.2015.10.012
- Meyer, K. D., Saletore, Y., Zumbo, P., Elemento, O., Mason, C. E., and Jaffrey, S. R. (2012). Comprehensive analysis of mRNA methylation reveals enrichment in 3' UTRs and near stop codons. *Cell* 149, 1635–1646. doi: 10.1016/j.cell.2012.05.003
- Miao, Z., Zhang, T., Qi, Y., Song, J., Han, Z., and Ma, C. (2019). Evolution of the RNA N6-methyladenosine methylome mediated by genomic duplication. *Plant Physiol.* 182, 345–360. doi: 10.1104/pp.19.00323
- Mielecki, D., Zugaj, D. Ł., Muszewska, A., Piwowarski, J., Chojnacka, A., Mielecki, M., et al. (2012). Novel AlkB dioxygenases—alternative models for in silico and in vivo studies. *PLoS One* 7:e30588. doi: 10.1371/journal.pone.0030588
- Molinie, B., Wang, J., Lim, K. S., Hillebrand, R., Lu, Z. X., Van Wittenbergh, N., et al. (2016). m6A-LAIC-seq reveals the census and complexity of the m6A epitranscriptome. *Nat. Methods* 13, 692–698. doi: 10.1038/nmeth.3898
- Nestor, C., Ruzov, A., Meehan, R., and Dunican, D. (2010). Enzymatic approaches and bisulfite sequencing cannot distinguish between 5-methylcytosine and 5-hydroxymethylcytosine in DNA. *Biotechniques* 48, 317–319. doi: 10.2144/000113403
- Nunomura, A., Lee, H. G., Zhu, X., and Perry, G. (2017). Consequences of RNA oxidation on protein synthesis rate and fidelity: implications for the pathophysiology of neuropsychiatric disorders. *Biochem Soc. Trans.* 45, 1053–1066. doi: 10.1042/BST20160433
- Palladino, M. J., Keegan, L. P., O'Connell, M. A., and Reenan, R. A. (2000). A-to-I pre-mRNA editing in drosophila is primarily involved in adult nervous system function and integrity. *Cell* 102, 437–449. doi: 10.1016/S0092-8674(00)00049-0
- Park, O. H., Ha, H., Lee, Y., Boo, S. H., Kwon, D. H., and Song, H. K. (2019). Endoribonucleolytic cleavage of m(6)A-containing RNAs by RNase P/MRP complex. *Mol. Cell* 74, 494–507 e498. doi: 10.1016/j.molcel.2019.02.034
- Patil, D. P., Chen, C. K., Pickering, B. F., Chow, A., Jackson, C., Guttman, M., et al. (2016). m6A RNA methylation promotes XIST-mediated transcriptional repression. *Nature* 537, 369–373. doi: 10.1038/nature19342
- Pavitt, G. D. (2018). Regulation of translation initiation factor eIF2B at the hub of the integrated stress response. *Wiley Interdiscip. Rev. RNA* 9:e1491. doi: 10.1002/wrna.1491
- Peer, E., Rechavi, G., and Dominissini, D. (2017). Epitranscriptomics: regulation of mRNA metabolism through modifications. *Curr. Opin. Chem. Biol.* 41, 93–98. doi: 10.1016/j.cbpa.2017.10.008
- Pendleton, K. E., Chen, B., Liu, K., Hunter, O. V., Xie, Y., Tu, B. P., et al. (2017). The U6 snRNA m(6)A methyltransferase METTL16 regulates SAM synthetase intron retention. *Cell* 169, 824–835. doi: 10.1016/j.cell.2017.05.003
- Peng, X., Xu, X., Wang, Y., Hawke, D. H., Yu, S., Han, L., et al. (2018). A-to-I RNA editing contributes to proteomic diversity in cancer. *Cancer Cell* 33, 817–828. doi: 10.1016/j.ccell.2018.03.026
- Pfaff, C., Ehrnsberger, H. F., Flores-Tornero, M., Sørensen, B. B., Schubert, T., Längst, G., et al. (2018). ALY RNA-binding proteins are required for nucleocytoplasmic mRNA transport and modulate plant growth and development. *Plant Physiol.* 177, 226–240. doi: 10.1104/pp.18.00173
- Ping, X. L., Sun, B. F., Wang, L., Xiao, W., Yang, X., Wang, W. J., et al. (2014). Mammalian WTAP is a regulatory subunit of the RNA N6-methyladenosine methyltransferase. *Cell Res.* 24, 177–189. doi: 10.1038/cr.2014.3
- Rajagopalan, L. E., Westmark, C. J., Jarzembowski, J. A., and Malter, J. S. (2012). hnRNP C tetramer measures RNA length to classify RNA polymerase II transcripts for export. *Science* 335, 1643–1646. doi: 10.1126/science.1218469
- Rintala-Dempsey, A. C., and Kothe, U. (2017). Eukaryotic stand-alone pseudouridine synthases—RNA modifying enzymes and emerging regulators of gene expression? *RNA Biol.* 14, 1185–1196. doi: 10.1080/15476286.2016.1276150
- Roost, C., Lynch, S. R., Batista, P. J., Qu, K., Chang, H. Y., and Kool, E. T. (2015). Structure and thermodynamics of N6-methyladenosine in RNA: a spring-loaded base modification. *J. Am. Chem. Soc.* 137, 2107–2115. doi: 10.1021/ja513080v
- Ross, R., Cao, X., Yu, N., and Limbach, P. A. (2016). Sequence mapping of transfer RNA chemical modifications by liquid chromatography tandem mass spectrometry. *Methods* 107, 73–78. doi: 10.1016/j.ymeth.2016.03.016
- Roundtree, I. A., and He, C. (2016). RNA epigenetics—chemical messages for posttranscriptional gene regulation. *Curr. Opin. Chem. Biol.* 30, 46–51. doi: 10.1016/j.cbpa.2015.10.024
- Roundtree, I. A., Luo, G. Z., Zhang, Z., Wang, X., Zhou, T., Cui, Y., et al. (2017). YTHDC1 mediates nuclear export of N6-methyladenosine methylated mRNAs. *Elife* 6:e31311. doi: 10.7554/eLife.31311.040
- Ruzicka, K., Zhang, M., Campilho, A., Bodi, Z., Kashif, M., Saleh, M., et al. (2017). Identification of factors required for m6A mRNA methylation in Arabidopsis reveals a role for the conserved E3 ubiquitin ligase HAKAI. *New Phytol.* 215, 157–172. doi: 10.1111/nph.14586
- Sakurai, M., Ueda, H., Yano, T., Okada, S., Terajima, H., Mitsuyama, T., et al. (2014). A biochemical landscape of A-to-I RNA editing in the human brain transcriptome. *Genome Res.* 24, 522–534. doi: 10.1101/gr.162537.113
- Schaefer, M., Pollex, T., Hanna, K., and Lyko, F. (2009). RNA cytosine methylation analysis by bisulfite sequencing. *Nucleic Acids Res.* 37:e12. doi: 10.1093/nar/gkn954
- Schwartz, S., Agarwala, S. D., Mumbach, M. R., Jovanovic, M., Mertins, P., Shishkin, A., et al. (2013). High-resolution mapping reveals a conserved, widespread, dynamic mRNA methylation program in yeast meiosis. *Cell* 155, 1409–1421. doi: 10.1016/j.cell.2013.10.047
- Schwartz, S., and Motorin, Y. (2017). Next-generation sequencing technologies for detection of modified nucleotides in RNAs. *RNA Biol.* 14, 1124–1137. doi: 10.1080/15476286.2016.1251543
- Schwartz, S., Bernstein, D. A., Mumbach, M. R., Jovanovic, M., Herbst, R. H., and León-Ricardo, B. X. (2014a). Transcriptome-wide mapping reveals widespread dynamic-regulated pseudouridylation of ncRNA and mRNA. *Cell* 159, 148–162. doi: 10.1016/j.cell.2014.08.028
- Schwartz, S., Mumbach, M. R., Jovanovic, M., Wang, T., Maciag, K., and Bushkin, G. G. (2014b). Perturbation of m6A writers reveals two distinct classes of mRNA methylation at internal and 5' sites. *Cell Rep.* 8, 284–296. doi: 10.1016/j.celrep.2014.05.048
- Scutenaire, J., Deragon, J.-M., Jean, V., Benhamed, M., Raynaud, C., Favory, J. J., et al. (2018). The YTH domain protein ECT2 is an m6A reader required for normal trichome branching in Arabidopsis. *Plant Cell* 30, 986–1005. doi: 10.1105/tpc.17.00854
- Selmi, T., Hussain, S., Dietmann, S., Heiß, M., Borland, K., Flad, S., et al. (2021). Sequence- and structure-specific cytosine-5 mRNA methylation by NSUN6. *Nucleic Acids Res.* 49, 1006–1022. doi: 10.1093/nar/gkaa1193
- Sement, F. M., Ferrier, E., Zuber, H., Merret, R., Alioua, M., Deragon, J. M., et al. (2013). Uridylation prevents 3' trimming of oligoadenylated mRNAs. *Nucleic Acids Res.* 41, 7115–7127. doi: 10.1093/nar/gkt465
- Sending, E., Valle-Garcia, D., Dhall, A., Chen, H., Henriques, T., Navarrete-Perea, J., et al. (2019). PCIF1 catalyzes m6Am mRNA methylation to regulate gene expression. *Mol. Cell* 75, 631–643. doi: 10.1016/j.molcel.2019.05.030
- Sharma, S., Yang, J., van Nues, R., Watzinger, P., Kötter, P., and Lafontaine, D. L. J. (2017). Specialized box C/D snoRNPs act as antisense guides to target RNA base acetylation. *PLoS Genet.* 13:e1006804. doi: 10.1371/journal.pgen.1006804
- Shen, B., and Goodman, H. M. (2004). Uridine addition after microRNA-directed cleavage. *Science* 306, 997–997. doi: 10.1126/science.1103521
- Shen, L., Liang, Z., Gu, X., Chen, Y., Teo, Z. W., Hou, X., et al. (2016). N(6)-methyladenosine RNA modification regulates shoot stem cell fate in Arabidopsis. *Dev. Cell* 38, 186–200. doi: 10.1016/j.devcel.2016.06.008
- Shen, L., Liang, Z., Wong, C. E., and Yu, H. (2019). Messenger RNA modifications in plants. *Trends Plant Sci.* 24, 328–341. doi: 10.1016/j.tplants.2019.01.005
- Sheth, U., and Parker, R. (2003). Decapping and decay of messenger RNA occur in cytoplasmic processing bodies. *Science* 300, 805–808. doi: 10.1126/science.1082320
- Shi, H., Wang, X., Lu, Z., Zhao, B. S., Ma, H., Hsu, P. J., et al. (2017). YTHDF3 facilitates translation and decay of N6-methyladenosine-modified RNA. *Cell Res.* 27, 315–328. doi: 10.1038/cr.2017.15
- Shi, H., Zhang, X., Weng, Y. L., Lu, Z., Liu, Y., Lu, Z., et al. (2018). m6A facilitates hippocampus-dependent learning and memory through YTHDF1. *Nature* 563, 249–253. doi: 10.1038/s41586-018-0666-1

- Shu, X., Cao, J., Cheng, M., Xiang, S., Gao, M., Li, T., et al. (2020). A metabolic labeling method detects m6A transcriptome-wide at single base resolution. *Nat. Chem. Biol.* 16, 887–895. doi: 10.1038/s41589-020-0526-9
- Slobodin, B., Han, R., Calderone, V., Vrielink, J. A. O., Loayza-Puch, F., Elkon, R., et al. (2017). Transcription impacts the efficiency of mRNA translation via co-transcriptional N6-adenosine methylation. *Cell* 169, 326–337. doi: 10.1016/j.cell.2017.03.031
- Song, J., and Yi, C. (2017). Chemical modifications to RNA: a new layer of gene expression regulation. *ACS Chem. Biol.* 12, 316–325. doi: 10.1021/acscchembio.6b00960
- Spenkuch, F., Motorin, Y., and Helm, M. (2014). Pseudouridine: still mysterious, but never a fake (uridine)! *RNA Biol.* 11, 1540–1554. doi: 10.4161/15476286.2014.992278
- Squires, J. E., and Preiss, T. (2010). Function and detection of 5-methylcytosine in eukaryotic RNA. *Epigenomics* 2, 709–715. doi: 10.2217/epi.10.47
- Squires, J. E., Patel, H. R., Nousch, M., Sibbritt, T., Humphreys, D. T., Parker, B. J., et al. (2012). Widespread occurrence of 5-methylcytosine in human coding and non-coding RNA. *Nucleic Acids Res.* 40, 5023–5033. doi: 10.1093/nar/gks144
- Sun, H., Zhang, M., Li, K., Bai, D., and Yi, C. (2019). Cap-specific, terminal N(6)-methylation by a mammalian m(6)Am methyltransferase. *Cell Res.* 29, 80–82. doi: 10.1038/s41422-018-0117-4
- Suzuki, T., Ueda, H., Okada, S., and Sakurai, M. (2015). Transcriptome-wide identification of adenosine-to-inosine editing using the ICE-seq method. *Nat. Protoc.* 10, 715–732. doi: 10.1038/nprot.2015.037
- Theler, D., Dominguez, C., Blatter, M., Boudet, J., and Allain, F. H. (2014). Solution structure of the YTH domain in complex with N6-methyladenosine RNA: a reader of methylated RNA. *Nucleic Acids Res.* 42, 13911–13919. doi: 10.1093/nar/gku1116
- Thuring, K., Schmid, K., Keller, P., and Helm, M. (2016). Analysis of RNA modifications by liquid chromatography-tandem mass spectrometry. *Methods* 107, 48–56. doi: 10.1016/j.ymeth.2016.03.019
- Topisirovic, I., Vitkin, Y. V., Sonenberg, N., and Shatkin, A. J. (2011). Cap and cap-binding proteins in the control of gene expression. *RNA* 2, 277–298. doi: 10.1002/wrna.52
- Torres, A. G., Pineyro, D., Filonava, L., Stracker, T. H., Batlle, E., Ribas, et al. (2014). A-to-I editing on tRNAs: biochemical, biological and evolutionary implications. *FEBS Lett.* 588, 4279–4286. doi: 10.1016/j.febslet.2014.09.025
- Tzafrir, I., Dickerman, A., Brazhnik, O., Nguyen, Q., McElver, J., Frye, C., et al. (2003). The Arabidopsis seedgenes project. *Nucleic Acids Res.* 31, 90–93. doi: 10.1093/nar/gkg028
- Vandivier, L. E., and Gregory, B. D. (2018). New insights into the plant epitranscriptome. *J. Exp. Bot.* 69, 4659–4665. doi: 10.1093/jxb/ery262
- Vandivier, L. E., Campos, R., Kuksa, P. P., Silverman, I. M., Wang, L.-S., and Gregory, B. D. (2015). Chemical modifications mark alternatively spliced and uncapped messenger RNAs in Arabidopsis. *Plant Cell* 27, 3024–3037. doi: 10.1105/tpc.15.00591
- Vespa, L., Vachon, G., Berger, F., Perazza, D., Faure, J.-D., and Herzog, M. (2004). The immunophilin-interacting protein AtFIP37 from Arabidopsis is essential for plant development and is involved in trichome endoreduplication. *Plant Physiol.* 134, 1283–1292. doi: 10.1104/pp.103.028050
- Vilfan, I. D., Tsai, Y. C., Clark, T. A., Wegener, J., Dai, Q., Yi, C., et al. (2013). Analysis of RNA base modification and structural rearrangement by single-molecule real-time detection of reverse transcription. *J. Nanobiotechnol.* 11:8. doi: 10.1186/1477-3155-11-8
- Walters, R. W., Matheny, T., Mizoue, L. S., Rao, B. S., Muhlrud, D., and Parker, R. (2017). Identification of NAD⁺ capped mRNAs in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. USA* 114, 480–485. doi: 10.1073/pnas.1619369114
- Wan, Y., Tang, K., Zhang, D., Xie, S., Zhu, X., Wang, Z., et al. (2015). Transcriptome-wide high-throughput deep m6A-seq reveals unique differential m6A methylation patterns between three organs in *Arabidopsis thaliana*. *Genome Biol.* 16:272. doi: 10.1186/s13059-015-0839-2
- Wang, X., Li, Q., Yuan, W., Kumar, S., and Qian, W. (2016). The cytosolic Fe-S cluster assembly component MET18 is required for the full enzymatic activity of ROS1 in active DNA demethylation. *Sci Rep.* 6:26443. doi: 10.1038/srep26443
- Wang, X., Lu, Z., Gomez, A., Hon, G. C., Yue, Y., Han, D., et al. (2014). N6-methyladenosine-dependent regulation of messenger RNA stability. *Nature* 505, 117–120. doi: 10.1038/nature12730
- Wang, X., Zhao, B. S., Roundtree, I. A., Lu, Z., Han, D., Ma, H., et al. (2015). N6-methyladenosine modulates messenger RNA translation efficiency. *Cell* 161, 1388–1399. doi: 10.1016/j.cell.2015.05.014
- Wang, Y., Ni, T., Wang, W., and Liu, F. (2018). Gene transcription in bursting: a unified mode for realizing accuracy and stochasticity. *Biol. Rev.* 94, 248–258. doi: 10.1111/brv.12452
- Wang, Y., Xiao, Y., Dong, S., Yu, Q., and Jia, G. (2020). Antibody-free enzyme assisted chemical approach for detection of N6-methyladenosine. *Nat. Chem. Biol.* 16, 896–903. doi: 10.1038/s41589-020-0525-x
- Wang, Y., Li, S., Zhao, Y., You, C., Le, B., Gong, Z., et al. (2019). NAD⁺-capped RRNAs are widespread in the *Arabidopsis* transcriptome and can probably be translated. *Proc. Natl. Acad. Sci. USA* 116, 12094–12102. doi: 10.1073/pnas.1903682116
- Wei, L. H., Song, P., Wang, Y., Lu, Z., Tang, Q., Yu, Q., et al. (2018). The m6A reader ECT2 controls trichome morphology by affecting mRNA stability in Arabidopsis. *Plant Cell* 30, 968–985. doi: 10.1105/tpc.17.00934
- Wen, J., Lv, R., Ma, H., Shen, H., He, C., Wang, J., et al. (2018). Zc3h13 regulates nuclear RNA m6A methylation and mouse embryonic stem cell self-renewal. *Mol. Cell* 69, 1028–1038. doi: 10.1016/j.molcel.2018.02.015
- Weng, Y. L., Wang, X., An, R., Cassin, J., Vissers, C., Liu, Y., et al. (2018). Epitranscriptomic m6A regulation of axon regeneration in the adult mammalian nervous system. *Neuron* 97, 313–325. doi: 10.1016/j.neuron.2017.12.036
- Werner, M., Purta, E., Kaminska, K. H., Cymerman, I. A., Campbell, D. A., Mittra, B., et al. (2011). 2'-O-ribose methylation of cap2 in human: function and evolution in a horizontally mobile family. *Nucleic Acids Res.* 39, 4756–4768. doi: 10.1093/nar/gkr038
- Wu, L., Candille, S. I., Choi, Y., Xie, D., Jiang, L., Li-Pook-Than, J., et al. (2013). Variation and genetic control of protein abundance in humans. *Nature* 499, 79–82. doi: 10.1038/nature12223
- Wu, R., Jiang, D., Wang, Y., and Wang, X. (2016). N(6)-methyladenosine (m6A) methylation in mRNA with a dynamic and reversible epigenetic modification. *Mol. Biotechnol.* 58, 450–459. doi: 10.1007/s12033-016-9947-9
- Wurm, J. P., Meyer, B., Bahr, U., Held, M., Frolow, O., Kotter, B., et al. (2010). The ribosome assembly factor Nep1 responsible for bowen-conradi syndrome is a pseudouridine-N1-specific methyltransferase. *Nucleic Acids Res.* 38, 2387–2398. doi: 10.1093/nar/gkp1189
- Xiao, W., Adhikari, S., Dahal, U., Chen, Y. S., Hao, Y. J., Sun, B. F., et al. (2016). Nuclear m6A reader YTHDC1 regulates mRNA splicing. *Mol. Cell* 61, 507–519. doi: 10.1016/j.molcel.2016.01.012
- Xiong, X., Li, X., and Yi, C. (2018). N1-methyladenosine methylome in messenger RNA and non-coding RNA. *Curr. Opin. Chem. Biol.* 45, 179–186. doi: 10.1016/j.cbpa.2018.06.017
- Xu, C., Wang, X., Liu, K., Roundtree, I. A., Tempel, W., Li, Y., et al. (2014). Structural basis for selective binding of m6A RNA by the YTHDC1 YTH domain. *Nat. Chem. Biol.* 10, 927–929. doi: 10.1038/nchembio.1654
- Xu, K., Yang, Y., Feng, G. H., Sun, B. F., Chen, J. Q., Li, Y. F., et al. (2017). Mettl3-mediated m6A regulates spermatogonial differentiation and meiosis initiation. *Cell Res.* 27, 1100–1114. doi: 10.1038/cr.2017.100
- Xu, L., Liu, X., Sheng, N., Oo, K. S., Liang, J., Chionh, Y. H., et al. (2017). Three distinct 3-methylcytidine (m3C) methyltransferases modify tRNA and mRNA in mice and humans. *J. Biol. Chem.* 292, 14695–14703. doi: 10.1074/jbc.M117.798298
- Yan, L. L., and Zaher, H. S. (2019). How do cells cope with RNA damage and its consequences? *J. Biol. Chem.* 294, 15158–15171. doi: 10.1074/jbc.REV119.006513
- Yan, L. L., Simms, C. L., McLoughlin, F., Vierstra, R. D., and Zaher, H. S. (2019). Oxidation and alkylation stresses activate ribosome-quality control. *Nat. Commun.* 10:5611. doi: 10.1038/s41467-019-13579-3
- Yang, L., Perrera, V., Saplaoura, E., Apelt, F., Bahin, M., Kramdi, A., et al. (2019). m5C Methylation guides systemic transport of messenger RNA over graft junctions in plants. *Curr. Biol.* 29, 2465–2476. doi: 10.1016/j.cub.2019.06.042
- Yang, X., Yang, Y., Sun, B. F., Chen, Y. S., Xu, J. W., Lai, W. Y., et al. (2017). 5-methylcytosine promotes mRNA export—NSUN2 as the methyltransferase and ALYREF as an m5C reader. *Cell Res.* 27, 606–625. doi: 10.1038/cr.2017.55

- Yang, Y. Y., Hsu, P. J., Chen, Y. S., and Yang, Y. G. (2018). Dynamic transcriptomic m6A decoration: writers, erasers, readers and functions in RNA metabolism. *Cell Res.* 28, 616–624. doi: 10.1038/s41422-018-0040-8
- Yang, Y., Wang, L., Han, X., Yang, W.-L., Zhang, M., Ma, H.-L., et al. (2019). RNA 5-methylcytosine facilitates the maternal-to-zygotic transition by preventing maternal mRNA decay. *Mol. Cell* 75, 1188–1202. doi: 10.1016/j.molcel.2019.06.033
- Yuan, F., Bi, Y., Siejka-Zielinska, P., Zhou, Y. L., Zhang, X. X., and Song, C. X. (2019). Bisulfite-free and base-resolution analysis of 5-methylcytidine and 5-hydroxymethylcytidine in RNA with peroxotungstate. *Chem. Commun.* 55, 2328–2331. doi: 10.1039/C9CC00274J
- Yue, Y., Hu, J., and He, C. (2015). RNA N6-methyladenosine methylation in post-transcriptional gene expression regulation. *Genes Dev.* 29, 1343–1355. doi: 10.1101/gad.262766.115
- Yue, Y., Liu, J., Cui, X., Cao, J., Luo, G., Zhang, Z., et al. (2018). VIRMA mediates preferential m6A mRNA methylation in 3'UTR and near stop codon and associates with alternative polyadenylation. *Cell Discov.* 4:10. doi: 10.1038/s41421-018-0019-0
- Zhang, F., Zhang, Y.-C., Lio, J.-Y., Yu, Y., Zhou, Y.-F., Feng, Y.-Z., et al. (2019). The subunit of RNA N6-methyladenosine methyltransferase OsFIP regulates early degeneration of microspores in rice. *PLoS Genet.* 15:e1008120. doi: 10.1371/journal.pgen.1008120
- Zhang, H. Y., Xiong, J., Qi, B. L., Feng, Y. Q., and Yuan, B. F. (2016). The existence of 5-hydroxymethylcytosine and 5-formylcytosine in both DNA and RNA in mammals. *Chem. Commun.* 52, 737–740. doi: 10.1039/C5CC07354E
- Zhang, Z., Chen, L. Q., Zhao, Y. L., Yang, C. G., Roundtree, I. A., Zhang, Z., et al. (2019). Single-base mapping of m6A by an antibody-independent method. *Sci. Adv.* 5:eaax0250. doi: 10.1126/sciadv.aax0250
- Zhang, Z., Hu, F., Sung, M. W., Shu, C., Castillo-González, C., Koiwa, H., et al. (2017). RISC-interacting clearing 3'-5' exoribonucleases (RICEs) degrade uridylated cleavage fragments to maintain functional RISC in *Arabidopsis thaliana*. *eLife* 6:e24466. doi: 10.7554/eLife.24466
- Zhang, Z., Theler, D., Kaminska, K. H., Hiller, M., and De La Grange, P. (2010). The YTH domain is a novel RNA binding domain. *J. Biol. Chem.* 285, 14701–14710. doi: 10.1074/jbc.M110.104711
- Zhao, B. S., Roundtree, I. A., and He, C. (2017a). Post-transcriptional gene regulation by mRNA modification. *Nat. Rev. Mol. Cell Biol.* 18, 31–42. doi: 10.1038/nrm.2016.132
- Zhao, B. S., Wang, X., Beadell, A. V., Lu, Z., Shi, H., Kuuspalu, A., et al. (2017b). m6A-dependent maternal mRNA clearance facilitates zebrafish maternal-to-zygotic transition. *Nature* 542, 475–478. doi: 10.1038/nature21355
- Zhao, L. Y., Song, J., Liu, Y., Song, C.-X., and Yi, C. (2020). Mapping the epigenetic modifications of DNA and RNA. *Protein Cell* 11, 792–808. doi: 10.1007/s13238-020-00733-7
- Zheng, G., Dahl, J. A., Niu, Y., Fedorcsak, P., Huang, C.-M., Li, C. J., et al. (2013). ALKBH5 is a mammalian RNA demethylase that impacts RNA metabolism and mouse fertility. *Mol. Cell* 49, 18–29. doi: 10.1016/j.molcel.2012.10.015
- Zhong, S., Li, H., Bodi, Z., Button, J., Vespa, L., Herzog, M., et al. (2008). MTA is an Arabidopsis messenger RNA adenosine methylase and interacts with a homolog of a sex-specific splicing factor. *Plant Cell* 20, 1278–1288. doi: 10.1105/tpc.108.058883
- Zhou, J., Wan, J., Gao, X., Zhang, X., Jaffrey, S., and Qian, S. B. (2015). Dynamic m6A mRNA methylation directs translational control of heat shock response. *Nature* 526, 591–594. doi: 10.1038/nature15377
- Zhou, L., Tian, S., and Qin, G. (2019). RNA methylomes reveal the m6A mediated regulation of DNA demethylase gene SLDML2 in tomato fruit ripening. *Genome Biol.* 20:156. doi: 10.1186/s13059-019-1771-7
- Zou, S., Toh, J. D., Wong, K. H., Gao, Y. G., Hong, W., and Woon, E. C. (2016). N6-Methyladenosine: a conformational marker that regulates the substrate specificity of human demethylases FTO and ALKBH5. *Sci. Rep.* 6:25677. doi: 10.1038/srep25677
- Zuber, H., Scheer, H., Ferrier, E., Sement, F. M., Mercier, P., Stupfler, B., et al. (2016). Uridylation and PABP cooperate to repair mRNA deadenylated ends in Arabidopsis. *Cell Rep.* 14, 2707–2717. doi: 10.1016/j.celrep.2016.02.060

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



Identifying RNA N6-Methyladenine Sites in Three Species Based on a Markov Model

Cong Pian[†], Zhixin Yang[†], Yuqian Yang, Liangyun Zhang and Yuanyuan Chen*

College of Science, Nanjing Agricultural University, Nanjing, China

OPEN ACCESS

Edited by:

Jia Meng,
Xi'an Jiaotong-Liverpool University,
China

Reviewed by:

Kil To Chong,
Jeonbuk National University,
South Korea
Jie Jiang,
Xi'an Jiaotong-Liverpool University,
China

*Correspondence:

Yuanyuan Chen
chenyuanyuan@njau.edu.cn

[†] These authors have contributed
equally to this work

Specialty section:

This article was submitted to
Epigenomics and Epigenetics,
a section of the journal
Frontiers in Genetics

Received: 08 January 2021

Accepted: 03 March 2021

Published: 19 March 2021

Citation:

Pian C, Yang Z, Yang Y, Zhang L
and Chen Y (2021) Identifying RNA
N6-Methyladenine Sites in Three
Species Based on a Markov Model.
Front. Genet. 12:650803.
doi: 10.3389/fgene.2021.650803

N6-methyladenosine (m6A), the most common posttranscriptional modification in eukaryotic mRNAs, plays an important role in mRNA splicing, editing, stability, degradation, etc. Since the methylation state is dynamic, methylation sequencing needs to be carried out over different time periods, which brings some difficulties to identify the RNA methyladenine sites. Thus, it is necessary to develop a fast and accurate method to identify the RNA N6-methyladenosine sites in the transcriptome. In this study, we use first-order and second-order Markov models to identify RNA N6-methyladenine sites in three species (*Saccharomyces cerevisiae*, mouse, and *Homo sapiens*). These two methods can fully consider the correlation between adjacent nucleotides. The results show that the performance of our method is better than that of other existing methods. Furthermore, the codons encoded by three nucleotides have biases in mRNA, and a second-order Markov model can capture this kind of information exactly. This may be the main reason why the performance of the second-order Markov model is better than that of the first-order Markov model in the m6A prediction problem. In addition, we provide a corresponding web tool called MM-m6APred.

Keywords: RNA N6-methyladenine sites, second-order Markov model, codons biases, transfer probability matrix, web tool

INTRODUCTION

To date, more than 160 types of RNA modifications have been discovered (Zhao et al., 2019). In these modifications, N6-methyladenosine (m6A) is the most common and abundant one existing in various species. It is closely associated with diverse biological processes, such as RNA localization and degradation (Wang et al., 2014), RNA structural dynamics (Roost et al., 2015), alternative splicing (Liu et al., 2015), and primary microRNA processing (Alarcón et al., 2015). Thus, identification of m6A sites is of great importance for better understanding their function and mechanisms (Chen et al., 2015). In the past few years, high-throughput experimental methods, such as MERIPP (Geula et al., 2015) and M6ASeq (Meyer et al., 2012), have been used to identify m6A modifications, but these methods have some limitations: (1) The location of the m6A site cannot be accurately located; (2) the cost is high; and (3) they are not applicable for the large-scale identification of m6A sites. Hence, it is highly desirable to develop a fast and accurate computational method for the identification of m6A sites (Dominissini et al., 2012).

Currently, there are several effective methods for predicting m6A sites based on machine learning, mainly including iRNA-Methyl (Chen et al., 2015), SRAMP (Zhou et al., 2016), M6AMRFS (Qiang et al., 2018), M6APred-EL (Wei et al., 2018), pm6A-CNN (Roost et al., 2015),

and iN6-Methyl (Nazari et al., 2019) etc. The above methods actually use the physical and chemical properties of nucleotides in various species, such as the nucleotide frequency at specific locations and the chemical properties of nucleotides, to extract features and predict m6A sites. However, none of these methods can capture the correlation between adjacent nucleotides well, while the Markov model can model this kind of correlation. In fact, Pian et al. (2020) used a first-order Markov model to predict the DNA N6-methyladenine sites. Recently, we proposed a method to predict DNA 4mC sites based on the second-order Markov model (Yang et al., 2020). Later, we found that the second-order Markov model is more suitable for predicting the methylation sites of RNA m6A because of the biases of the triplet codons in mRNA. The main purpose of this article is to provide a more accurate prediction tool of m6A.

Based on this idea, we used a second-order Markov model to identify the m6A sites of RNA. The m6A data of the three species of *Saccharomyces cerevisiae*, mouse, and *Homo sapiens*, were used to evaluate our model. The results show that the prediction performances of the first-order Markov model and the second-order Markov model are significantly better than those of the other four existing prediction tools. In addition, the second-order Markov model outperforms the first-order Markov model, which indicates that the second-order Markov model can capture the codon bias in mRNA well. This suggests that second-order Markov may be able to characterize the codon bias in mRNA.

MATERIALS AND METHODS

Benchmark Datasets

In this study, we used three benchmark datasets from three different species: *S. cerevisiae* (Chen et al., 2015), mouse (Dominissini et al., 2012), and *H. sapiens* (Chen et al., 2017). The corresponding number of positive samples was 1,300, 725, and 1,130. There were as many negative samples as positive samples. Table 1 shows the details of these data. For the three benchmark datasets, the positives were the sequences centered with true m6A sites, while the negatives were the sequences centered with adenines but without any m6A peaks detected. The datasets can be downloaded from the following website¹.

Model Construction

A Markov model is a stochastic process where the next variable depends on only the most recent variable(s) instead of all the previous variables. In this sequence information study, we first model a sequence as a first-order Markov chain, and the

¹ <http://server.malab.cn/M6AMRFS/>

TABLE 1 | Details of benchmark datasets.

Type	Positive	Negative	Total	Length
Yeast cells	1,300	1,300	2,600	51 nt
Mouse	725	725	1,450	41 nt
Homo sapiens	1,130	1,130	2,260	41 nt

current nucleotide depends on the previous nucleotide only. More specifically, for the m6A sequences of positive samples in the training data, we first calculate the initial probability $P_{S_1}^P (P_A^P, P_G^P, P_C^P, P_U^P)$ of the initial state S_1 nucleotide being A, G, C or U, respectively. Then, we need to calculate the transfer probability $P_{S_n-S_{n+1}}^{P_n}$ of the current nucleotide state to the next state individually from the initial state S_1 (for example, $P_{G-A}^{P_2}$ represents the probability that nucleotide G in state S_2 transfers to nucleotide A in state S_3).

Thus, we can obtain the probability of the occurrence of the four nucleotides in the initial state and the transfer probability matrix of each state except the last one. Similarly, for the negative sequences of non-m6A, the probability of the occurrence of the corresponding four nucleotides in the initial state and the transition probability matrix can also be obtained. Therefore, two Markov models are trained based on the m6A sequences and non-m6A sequences in the training dataset.

In the process of prediction, we need to select the probability values according to the nucleotide arrangement of the sequence, including the initial state probability and the corresponding transfer probability from the positive and negative Markov models in the previous step. Then, we calculate the products of positive and negative probability values. Finally, we calculate the ratio of the positive product and negative product. If the ratio is greater than 1, the sequence is considered a m6A sample. Otherwise, it is considered a non-m6A sample.

Since there is a bias in the codon of mRNA (Kurland, 1991; Quax, 2015), we consider using a second-order Markov model to capture this bias. The flowcharts of the training and testing of the second-order Markov model are shown in Figure 1. For the m6A sequences, we first calculate the initial probability $P_{S_1S_2}^P (P_{AA}^P, P_{AG}^P, \dots, P_{UU}^P)$ of the first dinucleotide. Then, we need to calculate the transfer probability $P_{S_nS_{n+1}-S_{n+2}}^{P_n}$ of the current dinucleotide ($S_n S_{n+1}$) to the next nucleotide (S_{n+2}) (for example, $P_{AA-A}^{P_1}$ represents the probability of state S_1S_2 transferring to S_3 , where the nucleotide of state S_1S_2 is AA, and the nucleotide of state S_3 is A). Thus, 39 transfer probability matrices with 16 rows and four columns can be obtained. Similarly, the initial probability and transfer probability can be obtained for non-m6A sequences. Therefore, two Markov models (M_P and M_N) are similarly trained based on the m6A sequences and non-m6A sequences in the training dataset. Taking the sequence “seq = GUAUAUAACUUUUUUCUUAAGGAGCAGGUGUCUGCCUAA” as an example, the probabilities $P(\text{seq}|M_P)$ and $P(\text{seq}|M_N)$ of the sequence “seq” under models M_P and M_N are obtained, respectively. Then, the value of $\text{Ratio} = P(\text{seq}|M_P)/P(\text{seq}|M_N)$ can be used to determine the class of “seq,” where

$$P(\text{seq}|M_P) = P_{GU}^P \times P_{GU-A}^{P_1} \times P_{UA-U}^{P_2} \times \dots \times P_{UA-A}^{P_{39}},$$

and

$$P(\text{seq}|M_N) = P_{GU}^N \times P_{GU-A}^{N_1} \times P_{UA-U}^{N_2} \times \dots \times P_{UA-A}^{N_{39}},$$

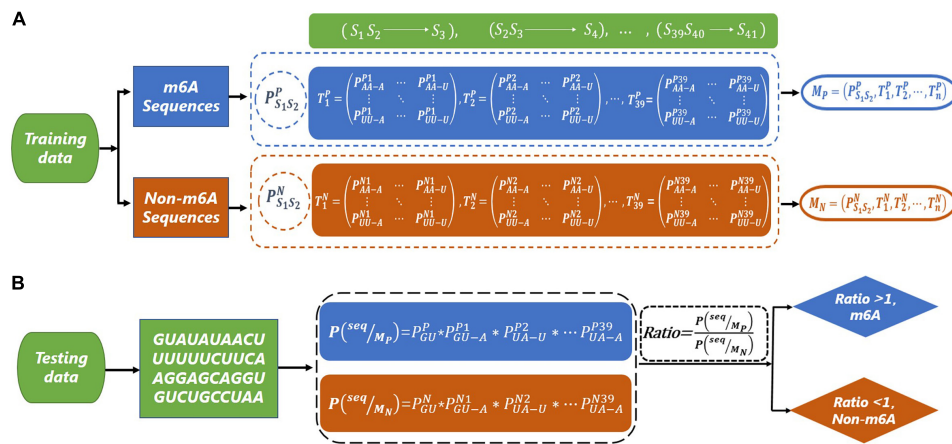


FIGURE 1 | The flow chart of m6A site prediction. **(A)** The construction of second-order Markov model (M_P and M_N) based on m6A sequence and non-m6A sequence. **(B)** The prediction for a test sequence. The sequence “GUUAUAACUUUUUUCUUAAGGAGCAGGUGUCUGCCUAA” is used as an example to explain the prediction process.

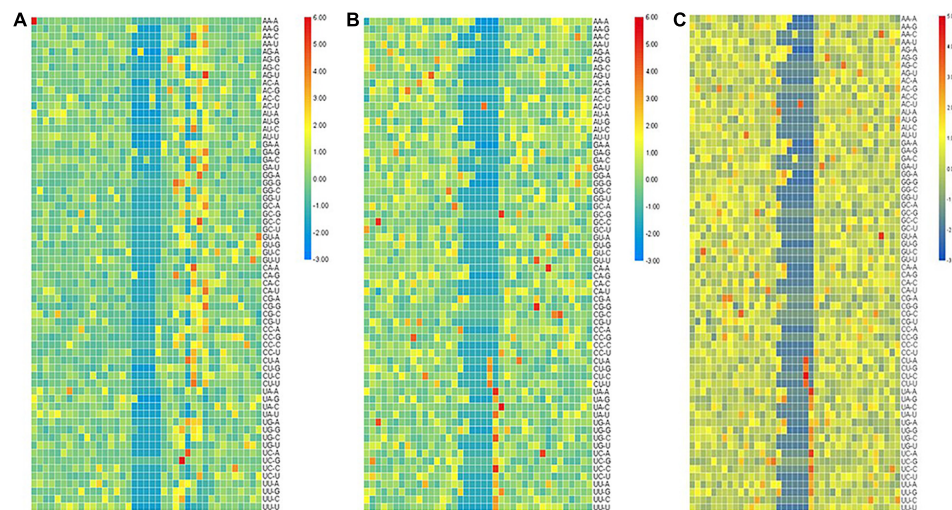


FIGURE 2 | The heat map of standardized quotient of the transfer probabilities of the three types of species. The heat map of standardized quotient of the transfer probabilities of the three types of species. **(A)** *Saccharomyces cerevisiae* cells, **(B)** mouse, and **(C)** *Homo sapiens*.

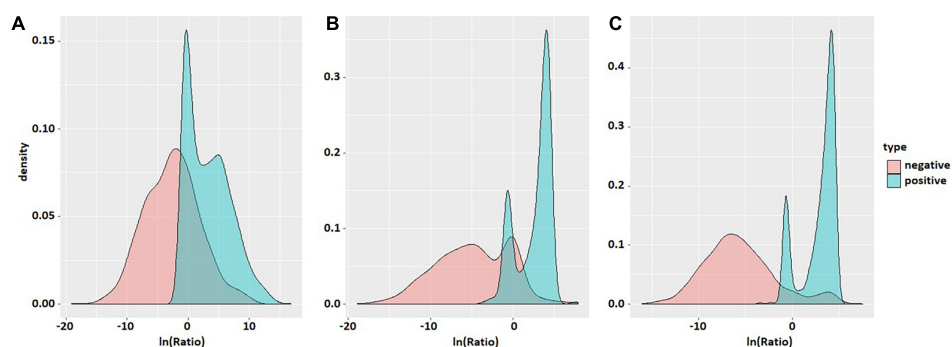


FIGURE 3 | Probability density maps of $\ln(\text{Ratio})$ values of the three species. The three density maps **(A)**, **(B)**, and **(C)** correspond to *S. cerevisiae*, mouse, and *H. sapiens*. Red is negative and blue is positive.

TABLE 2 | Evaluation data comparison table of six methods in (A) *S. cerevisiae*, (B) mouse, and (C) *H. sapiens*.

A	S_n (%)	S_p (%)	ACC (%)	MCC
M6APred-EL	72	72.69	72.34	44.68
SRAMP	71.92	71.38	71.65	43.31
iRNA-Methyl	71.69	73.45	72.57	45.15
M6AMRFS	73.45	72.84	73.14	46.29
First order-MM	73.85	71.69	72.30	49.23
Second order-MM	88.46	98.46	93.46	87.36
B	S_n (%)	S_p (%)	ACC (%)	MCC
M6APred-EL	77.79	1	88.90	79.79
SRAMP	77.79	1	88.90	79.79
iRNA-Methyl	77.66	99.31	88.48	78.84
M6AMRFS	77.79	1	88.90	79.79
First order-MM	79.98	88.88	83.55	74.85
Second order-MM	87.50	88.88	88.29	77.45
C	S_n (%)	S_p (%)	ACC (%)	MCC
M6APred-EL	82.04	99.73	90.89	83.08
SRAMP	79.65	1	89.82	81.35
iRNA-Methyl	80.35	1	90.18	81.95
M6AMRFS	81.95	99.82	90.89	83.11
First order-MM	84.60	87.50	85.00	73.85
Second order-MM	86.46	94.69	90.58	81.43

If the $Ratio > 1$, “seq” is classified as a m6A sequence; otherwise, it is classified as a non-m6A sequence.

Performance Evaluation

Ten-fold cross-validation was used to assess the reliability of the method. In the performance evaluation, the sensitivity (S_n), specificity (S_p), accuracy (ACC), and Mathew's correlation coefficient (MCC) were calculated. They are formulated as follows:

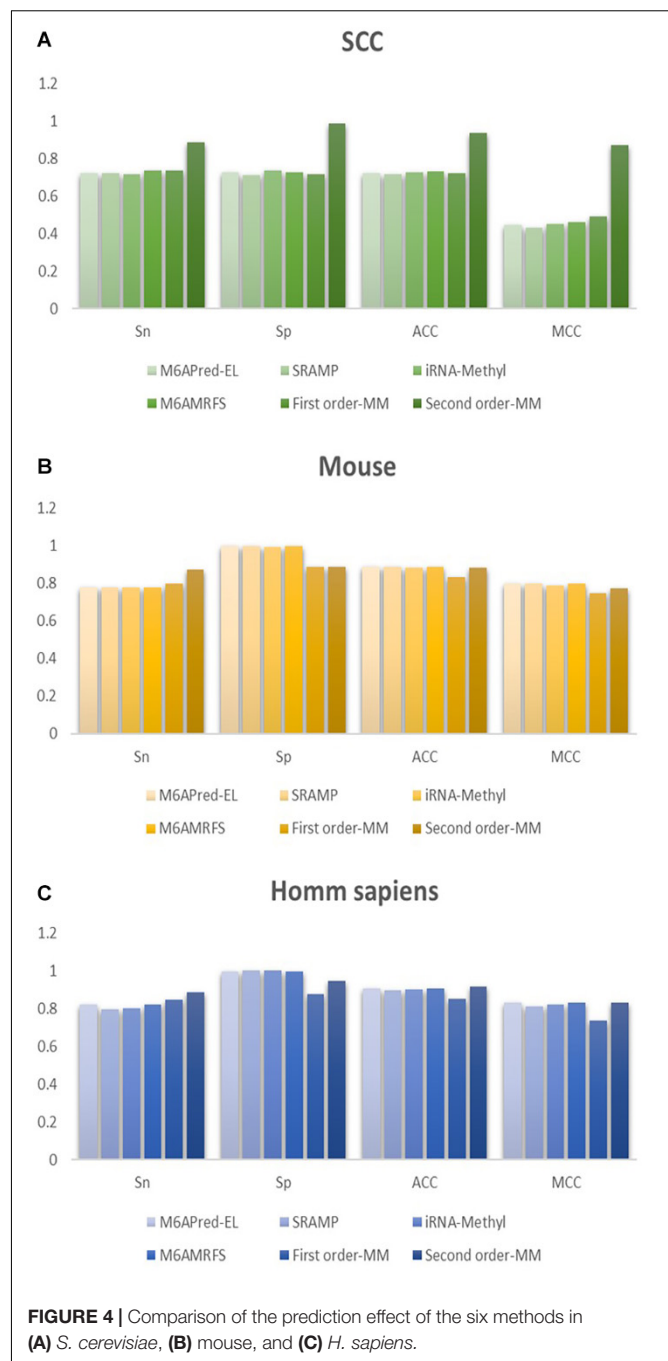
$$S_n = \frac{T_P}{T_P + F_N},$$

$$S_p = \frac{T_N}{T_N + F_P},$$

$$ACC = \frac{T_P + T_N}{T_P + T_N + F_P + F_N},$$

$$MCC = \frac{T_P + T_N - F_P \times F_N}{\sqrt{(T_P + F_P) \times (T_N + F_N) \times (T_P + F_N) \times (T_N + F_P)}}$$

where T_P , T_N , F_P , and F_N denote true positive, true negative, false positive, and false negative, respectively. S_n measures the predictive ability of a predictor for positive samples, while S_p measures the predictive ability of a predictor for negative samples. ACC and MCC are two metrics measuring the overall performance of a predictor.

**FIGURE 4** | Comparison of the prediction effect of the six methods in (A) *S. cerevisiae*, (B) mouse, and (C) *H. sapiens*.

RESULTS AND DISCUSSION

Representation and Illustration of $(P_{S_n S_{n+1} - S_{n+2}}^{P_n} / P_{S_n S_{n+1} - S_{n+2}}^{N_n})$

For the second-order Markov model, the heat map of the quotient matrix $(P_{S_n S_{n+1} - S_{n+2}}^{P_n} / P_{S_n S_{n+1} - S_{n+2}}^{N_n})$ of second-order transfer probability of m6A samples divided by the second-order transfer probability of non-m6A samples is shown in Figure 2. In order to facilitate comparison, the results of

TABLE 3 | Comparison of the prediction effect of m6A in mice based on the m6Avar database.

Method	S_n (%)	S_p (%)	ACC (%)
M6APred-EL	76.42	77.35	75.49
SRAMP	72.03	72.29	71.77
iRNA-Methyl	73.45	74.72	72.18
M6AMRFS	76.58	76.89	76.27
First order-MM	78.15	80.01	80.14
Second order-MM	86.22	87.13	85.32

The best performance in the respective part appears bold in the table.

heat map were standardized. The results show that there is a significant difference in the transfer probability of nucleotides at some positions between the positive and negative samples. This indicated that the second-order Markov chain is informative for predicting sequences containing m6A sites.

We also plotted the line charts of transfer probability of the second-order Markov model (Shown in **Supplementary Material 1**). Similar to the first-order Markov model, the transfer probability of positive samples is significantly different from that of negative samples in the second-order Markov model. Furthermore, the number of significant different sites in the second-order Markov model are obviously greater than that in the first-order Markov model from the line charts in **Supplementary Material 2**. It indicates that more information is provided in the second-order Markov model to help determine the type of sequences.

The Distribution of Ratios in the Positive and Negative Sample Sets

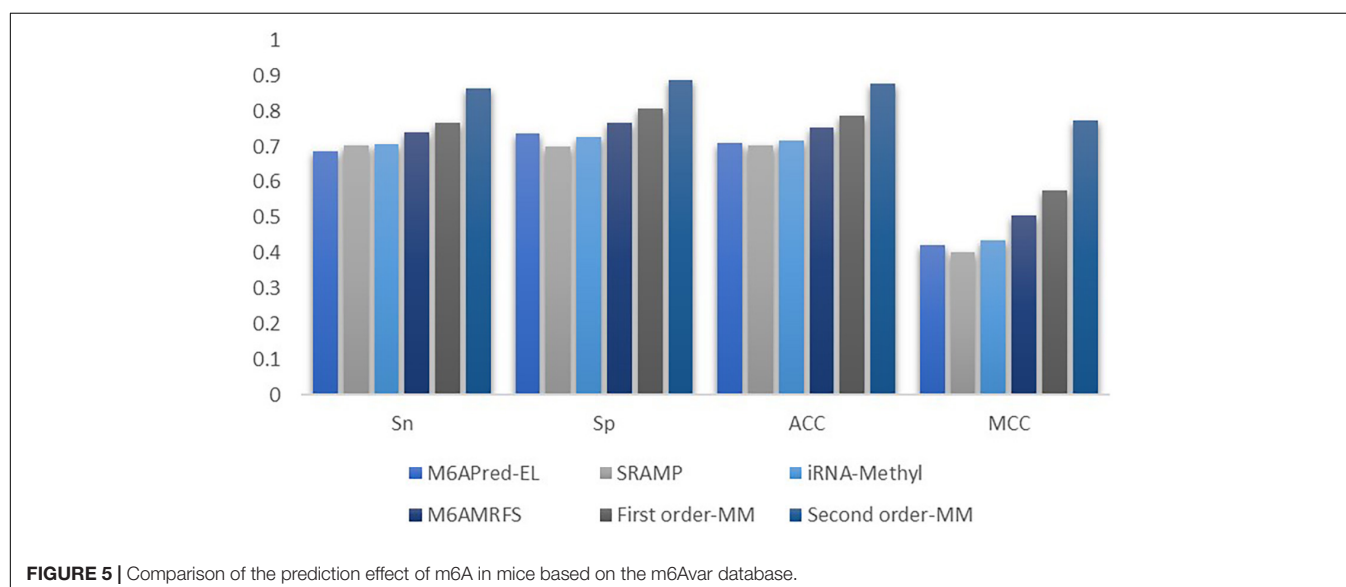
Probability density maps of $\ln(Ratio)$ values for three species based on the second-order transfer probability products are shown in **Figure 3**. It can be found that in each species, the distribution of $\ln(Ratio)$ is very different between positive and

negative samples, except for a small amount of overlap in the probability density graphs. The *Ratio* value of positive samples is significantly greater than that of negative samples, which enables the positive and negative samples to be divided accurately.

Comparison and Analysis

To evaluate our Markov model, we compared the performance of the two methods based on the Markov model with those of other m6A classifiers, including iRNA-Methyl, SRAMP, M6AMRFS, and M6APred-EL. **Table 2** and **Figure 4** show the prediction results of various methods (10-fold cross validation was used in all methods).

It can be found that the two methods based on the Markov model in m6A types of sequence identification had better or equal classification effects than several kinds of classifiers and that the second-order Markov model performed much better than the first-order Markov model in each aspect. It is noteworthy that S_p in several other methods is 100% in the species of mouse and *H. sapiens*, while the S_p of our method is close to 90% on average. Therefore, we checked these non-m6A data and found that the selection of negative sample data in the original literature [12] is unreasonable. The states S_{22} of the negative samples in mouse and human are all C, and the states S_{20} are all A or G. This is the reason why S_p of other methods can reach 100%. To evaluate our method more fairly, we downloaded 725 m6A sequences of mice from the m6Avar database, and the same number of sequences were randomly selected from the non-m6A sequences of the dbSNP database as negative samples. We used these data to retrain new models and carried out 10-fold cross validation in all methods. The performance results of all the above methods are shown in **Table 3** and **Figure 5**. The results indicate that all the performance metrics based on the two Markov model are high. And the second-order Markov model still performed much better than the first-order Markov mode.

**FIGURE 5** | Comparison of the prediction effect of m6A in mice based on the m6Avar database.

Web-Server Implementation

To facilitate the use of the Markov model to identify RNA m6A sites, the user-friendly web server MM-m6APred has been provided. It is freely available at². Our tool can handle RNA sequences of 41 bp or longer. Users can either paste RNA sequences into the text area or upload a FASTA format file.

CONCLUSION

Accurate identification of the m6A site is a necessary step in the study of its biological function. In this study, we used first-order and second-order Markov models to predict the m6A sites of three species. The results show that our method is better than the other four existing prediction tools. This shows that the Markov model can capture the correlation between neighboring nucleotides well. Considering the biases of the codons in mRNA, the second-order Markov model is used to capture these biases. The results show that the prediction performance of the second-order Markov model is significantly better than that of the first-order Markov model. In addition, we also provide the online prediction web tool of m6A, with code available to download (see text footnote 2).

²<http://www.pianlab.cn/m6APred/>

REFERENCES

- Alarcón, C. R., Lee, H., Goodarzi, H., Halberg, N., and Tavazoie, S. F. (2015). N 6-methyladenosine marks primary microRNAs for processing. *Nature* 519, 482–485. doi: 10.1038/nature14281
- Chen, W., Feng, P., Ding, H., Lin, H., and Chou, K.-C. (2015). iRNA-methyl:identifying N6-methyladenosine sites using pseudo nucleotide composition. *Anal. Biochem.* 490, 26–33. doi: 10.1016/j.ab.2015.08.021
- Chen, W., Tang, H., and Lin, H. (2017). MethyRNA: a web server for identification of N6-methyladenosine sites. *J. Biomol. Struct. Dyn.* 35, 683–687.
- Dominissini, D., Moshitch-Moshkovitz, S., Schwartz, S., Salmon-Divon, M., Ungar, L., Osenberg, S., et al. (2012). Topology of the human and mouse m6RNA methylomes revealed by m6A-seq. *Nature* 485, 201–206. doi: 10.1038/nature11112
- Geula, S., Moshitch-Moshkovitz, S., Dominissini, D., Mansour, A. A., Kol, N., Salmon-Divon, M., et al. (2015). m6A mRNA methylation facilitates resolution of naïve pluripotency toward differentiation. *Science* 347, 1002–1006. doi: 10.1126/science.1261417
- Kurland, C. G. (1991). Codon bias and gene expression. *FEBS Lett.* 285, 165–169. doi: 10.1016/0014-5793(91)80797-7
- Liu, B., Liu, F., Wang, X., Chen, J., Fang, L., and Chou, K.-C. (2015). Pse-in-One: a web server for generating various modes of pseudo components of DNA, RNA, and protein sequences. *Nucleic Acids Res.* 43, W65–W71. doi: 10.1093/nar/gkv458
- Meyer, K. D., Saletore, Y., Zumbo, P., Elemento, O., Mason, C. E., and Jaffrey, S. R. (2012). Comprehensive analysis of mRNA methylation reveals enrichment in 3' UTRs Near Stop Codons. *Cell* 149, 1635–1646. doi: 10.1016/j.cell.2012.05.003
- Nazari, I., Tahir, M., Tayara, H., and Chong, K. T. (2019). In6-methyl (5-step): identifying rna n6-methyladenosine sites using deep learning mode via chou's 5-step rules and chou's general pse-knc. *Chemometr. Intell. Lab. Syst.* 193:103811. doi: 10.1016/j.chemolab.2019.103811
- Pian, C., Zhang, G., Li, F., and Fan, X. (2020). Mm-6mapred: identifying dna n6-methyladenine sites based on markov model. *Bioinformatics* 36, 388–392. doi: 10.1093/bioinformatics/btz556

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: <http://server.malab.cn/M6AMRFS>.

AUTHOR CONTRIBUTIONS

CP and ZY contributed equally to this work. All authors read and approved the final manuscript.

ACKNOWLEDGMENTS

We would like to thank Xiao Sun (Southeast University) for his helpful suggestions.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2021.650803/full#supplementary-material>

- Qiang, X., Chen, H., Ye, X., Su, R., and Wei, L. (2018). M6AMRFS: robust prediction of N6-methyladenosine sites with sequence-based features in multiple species. *Front. Genet.* 9:495. doi: 10.3389/fgene.2018.00495
- Quax, T. F. (2015). Codon bias as a means to fine-tune gene expression. *Mol. Cell* 59, 149–161. doi: 10.1016/j.molcel.2015.05.035
- Roost, C., Lynch, S. R., Batista, P. J., Qu, K., Chang, H. Y., and Kool, E. T. (2015). Structure and thermodynamics of n6-methyladenosine in rna: a spring-loaded base modification. *J. Am. Chem. Soc.* 137, 2107–2115. doi: 10.1021/ja513080v
- Wang, X., Lu, Z., Gomez, A., Hon, G. C., Yue, Y., Han, D., et al. (2014). N6-methyladenosine-dependent regulation of messenger RNA stability. *Nature* 505, 117–120. doi: 10.1038/nature12730
- Wei, L., Chen, H., and Su, R. (2018). M6apred-el: a sequence-based predictor for identifying n6-methyladenosine sites using ensemble learning. *Mol. Ther. Nucleic Acids* 12, 635–644. doi: 10.1016/j.omtn.2018.07.004
- Yang, J., Lang, K., Zhang, G., Fan, X., Chen, Y., and Pian, C. (2020). SOMM4mC: a second-order Markov model for DNA N4-methylcytosine site prediction in six species. *Bioinformatics* 36, 4103–4105.
- Zhao, W., Zhou, Y., Cui, Q., and Zhou, Y. (2019). Paces: prediction of n4-acetylcytidine (ac4c) modification sites in mrna. *Sci. Rep.* 9:11112. doi: 10.1038/s41598-019-47594-7
- Zhou, Y., Zeng, P., Li, Y. H., Zhang, Z., and Cui, Q. (2016). SRAMP: prediction of mammalian N6-methyladenosine (m6A) sites based on sequence-derived features. *Nucleic Acids Res.* 44:e91. doi: 10.1093/nar/gkw104

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 Pian, Yang, Yang, Zhang 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.



Dynamics of m6A RNA Methylome on the Hallmarks of Hepatocellular Carcinoma

Enakshi Sivasudhan^{1,2}, Neil Blake², Zhi-Liang Lu^{1,3}, Jia Meng^{1,3} and Rong Rong^{1,3*}

¹ Department of Biological Sciences, Xi'an Jiaotong-Liverpool University, Suzhou, China, ² Department of Clinical Infection, Microbiology and Immunology, Institute of Infection, Veterinary and Ecological Sciences, University of Liverpool, Liverpool, United Kingdom, ³ Institute of Integrative Biology, University of Liverpool, Liverpool, United Kingdom

OPEN ACCESS

Edited by:

Rui Henrique,
Portuguese Oncology Institute,
Portugal

Reviewed by:

Hong Zheng,
Stanford University, United States
Lucia Coscujuela Tarrero,
Italian Institute of Technology (IIT), Italy

*Correspondence:

Rong Rong
Rong.Rong@xjtlu.edu.cn

Specialty section:

This article was submitted to
Epigenomics and Epigenetics,
a section of the journal
Frontiers in Cell and Developmental
Biology

Received: 16 December 2020

Accepted: 23 February 2021

Published: 01 April 2021

Citation:

Sivasudhan E, Blake N, Lu Z-L,
Meng J and Rong R (2021) Dynamics
of m6A RNA Methylome on
the Hallmarks of Hepatocellular
Carcinoma.
Front. Cell Dev. Biol. 9:642443.
doi: 10.3389/fcell.2021.642443

Epidemiological data consistently rank hepatocellular carcinoma (HCC) as one of the leading causes of cancer-related deaths worldwide, often posing severe economic burden on health care. While the molecular etiopathogenesis associated with genetic and epigenetic modifications has been extensively explored, the biological influence of the emerging field of epitranscriptomics and its associated aberrant RNA modifications on tumorigenesis is a largely unexplored territory with immense potential for discovering new therapeutic approaches. In particular, the underlying cellular mechanisms of different hallmarks of hepatocarcinogenesis that are governed by the complex dynamics of m6A RNA methylation demand further investigation. In this review, we reveal the up-to-date knowledge on the mechanistic and functional link between m6A RNA methylation and pathogenesis of HCC.

Keywords: hepatocellular carcinoma, epitranscriptomics, m6A RNA methylation, cancer hallmarks, writers, erasers, readers

INTRODUCTION

The advent of advancements in next-generation sequencing technologies along with the launch of highly specific antibodies capable of identifying chemically modified nucleotides broke new ground for RNA methylation, recently coined “epitranscriptomics,” to gain prominence as a dynamic and reversible modification, analogous to epigenetic regulations. While previous studies have largely focused on the genetic and epigenetic factors that contribute to hepatocellular carcinoma (HCC), research into deciphering the role of epitranscriptomics in triggering liver-related malignancies is still in its infancy. Thus, the dynamics of m6A RNA methylation on the molecular pathogenesis of HCC is an emerging field that requires extensive research, with immense potential to unlock new therapeutic targets to combat hepatocarcinogenesis.

HEPATOCELLULAR CARCINOMA

Hepatocellular carcinoma accounts for over 80% of hepatic neoplasms worldwide, imposing heavy disease burden by being the fourth most common cause of cancer-associated mortality worldwide (El-Serag and Rudolph, 2007; Fitzmaurice et al., 2019). The HCC incidence has been estimated to be more prevalent in males than females and widespread in certain regions including middle and western Africa, eastern and southern Asia, Polynesia, and Melanesia (Ferlay et al., 2010). Several risk

factors contribute to HCC, among which chronic alcohol consumption, cirrhosis, viral hepatitis (due to infection with hepatitis virus B and C), non-alcoholic fatty liver disease (NAFLD), and ingestion and exposure to aflatoxin and aristolochic acid majorly contribute toward the onset of hepatocarcinogenesis (Yang and Roberts, 2010). Chronic hepatitis B and C account for the most frequent etiologies especially in particular geographical locations (such as Oceania, western sub-Saharan Africa, and central Asia) with inadequate medical resources and also the predisposition to gradual hepatic damage and ensued cirrhosis and HCC (Bosch et al., 2004; Stanaway et al., 2016). Pathologically, the regenerating nodules synthesized during cirrhosis create a favorable microenvironment for the transformation of dysplastic hepatocytes to neoplastic lesions ultimately leading to HCC (Alqahtani et al., 2019). In addition, the manifestation of HCC is seldom reported in congenital hepatic fibrosis, ataxia telangiectasia, familial cholestatic cirrhosis, familial polyposis coli, fetal alcohol syndrome, and neurofibromatosis (Leong and Leong, 2005).

Surveillance for HCC, principally in high-risk individuals, includes ultrasonography and biomarker testing (Yang and Roberts, 2010). While emerging research shows several promising biomarkers, pertaining to HCC diagnosis, prognosis, and clinical staging, such as vascular endothelial growth factor (VEGF; Biselli-Chicote et al., 2012), epidermal growth factor (EGF; Kedmi et al., 2015), platelet-derived growth factor (PDGF; Campbell et al., 2005), insulin-like growth factor (IGF; Enguita-Germán and Fortes, 2014), mammalian target of rapamycin (mTOR), and microRNAs (Mínguez and Lachenmayer, 2011; Okada et al., 2015), currently alpha-fetoprotein (AFP) is the only clinically approved serological biomarker (Beudeker and Boonstra, 2020). In fact, high expression levels of EGF and VEGF that promote proliferation and angiogenesis, respectively, have been associated with early recurrence of HCC, while TGF and PDGF receptor protein overexpression has been shown to activate profibrotic pathways that induce liver tumorigenesis (Campbell et al., 2005; Biselli-Chicote et al., 2012; Kedmi et al., 2015). Patients diagnosed with early-stage HCC opt for curative treatment options including surgical removal, orthotopic liver transplantation, or percutaneous ablation, usually performed with radiofrequency ablation or percutaneous alcohol injection. For patients with unremovable tumors, transarterial chemoembolization, carried out by infusing a concoction of chemotherapeutic agents, and transarterial radioembolization, involving treatment with radioactive particles are recommended. Besides, treatment options for patients with advanced stages of cancer include multikinase inhibitors such as sorafenib, lenvatinib, and regorafenib (Llovet et al., 2008; Yang and Roberts, 2010).

The initiation and progression of HCC are facilitated by various genetic and epigenetic alterations, which are built up in hepatocytes, eventually leading to malignant transformation as a result of the conversion of proto-oncogenes to oncogenes and the loss of functional mutation or dosage changes of tumor suppressor genes. Genetic abnormalities include chromosomal translocation, single-nucleotide polymorphisms,

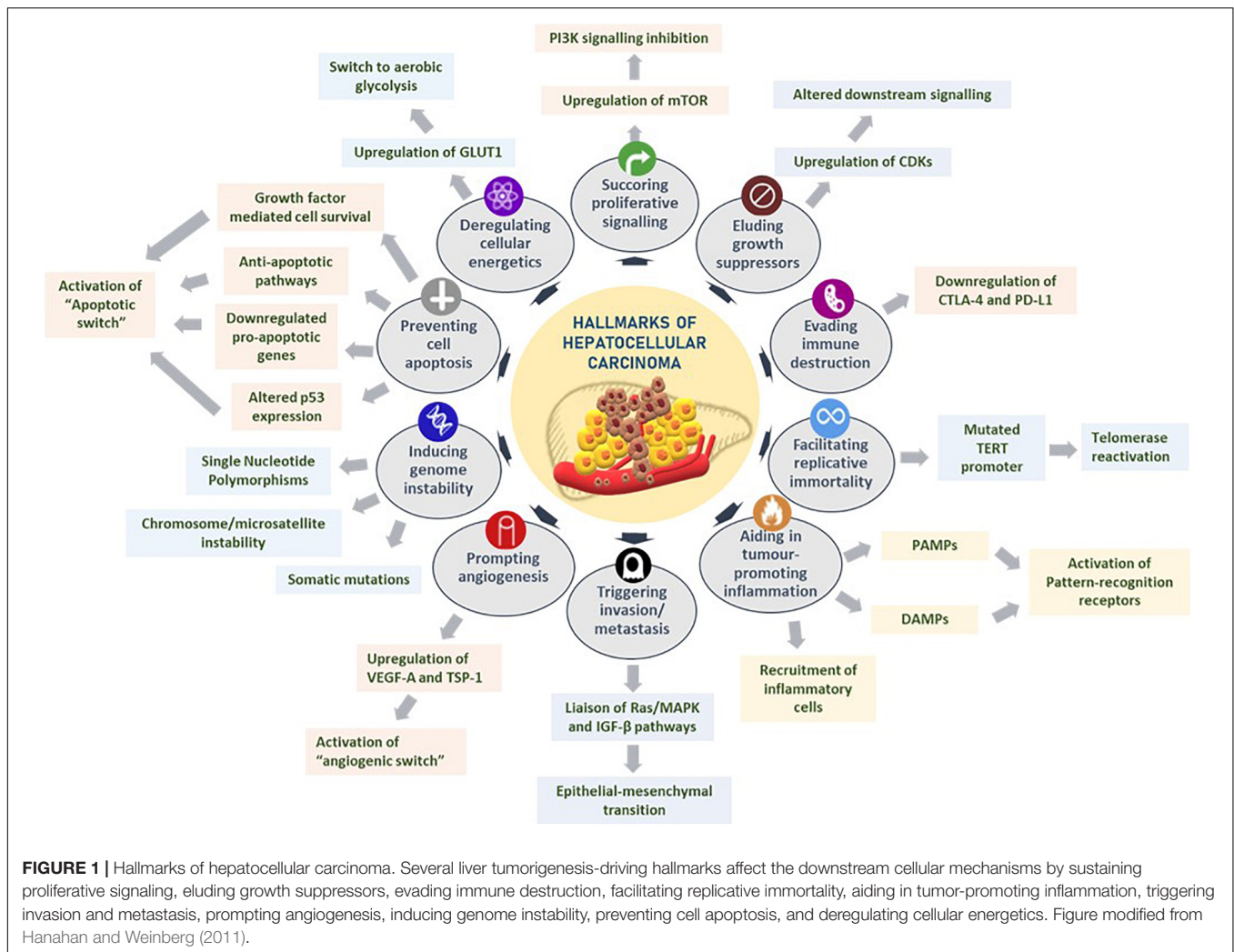
and targeted gene loss and deletion (Singh et al., 2018). Epigenetic changes, on the other hand, inflict no permanent genetic alterations; instead, they affect gene transcription and chromatin integrity. Epigenetic modifications that drive HCC include gene-specific DNA hypo- and hypermethylation, global genomic hypomethylation, aberrant histone modifications, and altered expression of microRNA (Leong and Leong, 2005; Libbrecht et al., 2005). Furthermore, deregulation of signal transduction pathways that govern cell cycle, proliferation, differentiation, and apoptosis, including the Wnt/ β -catenin pathway, Ras/Raf/MAPK pathway, PI3/AKT/mTOR pathway, JAK/STAT pathway, and ubiquitin-proteasome pathway (UPP), can lead to onset of liver tumorigenesis (Alqahtani et al., 2019).

HALLMARKS OF HEPATOCELLULAR CARCINOMA

Hanahan and Weinberg comprehensively presented the exploration of distinct and complementary traits that trigger tumorigenesis and metastatic propagation by logically organizing them into major hallmarks to rationally appreciate the complexities of neoplastic maladies (Hanahan and Weinberg, 2000, 2011). These hallmarks of cancer are as follows: sustaining proliferative signaling, eluding growth suppressors, evading immune destruction, facilitating replicative immortality, aiding in tumor-promoting inflammation, triggering invasion and metastasis, prompting angiogenesis, inducing genomic instability, preventing cell apoptosis, and deregulating cellular energetics (Hanahan and Weinberg, 2011). The hallmarks of cancer with regard to HCC are reviewed below and summarized in **Figure 1**.

Sustaining mitogenic signaling and evading growth suppressors in tumor cells are feasibly the most fundamental characteristics of tumor cells, unlike in normal cells that regulate cell homeostasis, especially pertaining to releasing growth-promoting signals (Hanahan and Weinberg, 2011; Sever and Brugge, 2015). For instance, HCC occurs largely as a result of uninhibited cellular proliferation resulting from a series of dysregulations in normal cell cycle regulators such as cyclin-dependent kinases (CDKs). Given the unique regenerative aptitude of hepatocytes, any reprobate cell proliferation, upregulation of CDKs, or alterations in CDK-related downstream signaling pathways and CDK inhibitors could potentiate the onset of hepatocarcinogenesis (Shen et al., 2019).

The role of the immune system in eradicating certain neoplasia and micrometastases is an area that demands further researching. Owing to the ever-alert surveillance nature of immune cells in eliminating tumor cells, it is worth exploring the potential mechanisms that solid tumors have acquired in successfully evading immunological destruction. In HCC, immune checkpoint inhibitors such as CTLA-4 and PD-L1 regulate the immunosuppression of chronic inflammation brought about by persistent expression of certain cytokines and immune cell recruitment (Makarova-Rusher et al., 2015; Xu et al., 2018). Contrary to previous beliefs that immune responses largely represented an attempt to



eliminate tumorigenesis, an ever-growing assemblage of scientific evidence suggests the paradoxical effect of tumor-induced inflammation in aiding neoplasias (DeNardo et al., 2010; Qian and Pollard, 2010). HCC-associated inflammation could be chiefly attributed to recruitment of inflammatory cells in the tumor microenvironment, extrinsic pathways that activate pattern recognition receptors by pathogen-associated molecule patterns, or damage-associated molecule patterns (DAMPs) released from liver cells undergoing apoptosis (Yu et al., 2018).

Tumor cells ensure continued survival by withstanding two crucial aspects that limit unlimited replicative potential in normal cells: senescence and crisis/apoptosis. It has been observed in cirrhotic liver cells that telomerase, an enzyme that prevents telomere shortening and ensuing cellular senescence, has an impaired activity coupled with subsequent shortening of telomeres implicating senescence of hepatocytes (Nault et al., 2019). To circumvent this and enable replicative perpetuity, telomerase reactivation is elicited through aberrant mutations in TERT promoter, leading to uncontrolled cell proliferation and subsequent HCC development (Donaires et al., 2017; Nault et al., 2019).

The multistep mechanism of invasion and metastasis is broadly regulated by epithelial–mesenchymal transition (EMT), heterotypic involvement of neoplastic stromal cells, and plasticity in the invasive growth properties disseminated by cancer cells (Hugo et al., 2007; Klymkowsky and Savagner, 2009; Egeblad et al., 2010). It has been proposed that in HCC, signaling through the Ras/MAPK pathway could liaise with the TGF- β signal transduction pathway in driving the shift from EMT, rendering tumor cells their mobility (Matsuzaki et al., 2000). Another hallmark of cancer, inducing angiogenesis, involves tumor-associated neovasculature that activates an “angiogenic switch” through the regulation of countervailing inducers and inhibitors, such as VEGF-A and thrombospondin-1, respectively, Hanahan and Weinberg (2011). Given the high invasive nature of HCC, it is not surprising to observe VEGF overexpression in the precancerous stages of dysplastic and cirrhotic liver tissues further to a strong correlation of VEGF expression and tumor grading of HCC (Hamdy et al., 2020).

Acquisition of genomic instability could convene selective advantage on neoplastic cells, enabling them to outgrow and dominate in a tumor microenvironment niche. Premalignant

cells drive tumorigenesis by enhancing their sensitivity to mutagenic agents and compromising the “surveillance systems” that monitor cellular genomic integrity (Jackson and Bartek, 2009). Notably in HCC, genetic alterations are instigated by chromosome and microsatellite instability, accumulated somatic mutations, single-nucleotide polymorphisms, and deregulated signaling pathways (Niu et al., 2016; Rao et al., 2017). Resisting cell death is another crucial hallmark of tumorigenesis. While programmed cell apoptosis, resulting from certain elevated oncogenic signaling mechanisms and hyperproliferation-associated DNA damage, functions as a natural barrier to carcinogenesis, certain tumors eventually progress to high-grade malignancy and induce drug resistance, through an “apoptotic switch” (Adams and Cory, 2007; Hanahan and Weinberg, 2011). Especially in HCC, apoptosis-associated mechanisms are governed by attenuation of p53 function through telomere-induced chromosomal instability (Farazi et al., 2006), downregulation of pro-apoptotic genes such as B cell lymphoma 2 (Mott and Gores, 2007), growth factor-mediated cell survival (Llovet and Bruix, 2008), and overactivation of anti-apoptotic pathways associated with fas pathway inhibitors (Lee et al., 2001).

Tumor energy metabolism, an emerging hallmark, confers the metabolic preferences of cancer cells to favor aerobic glycolysis, famously characterized as the Warburg effect (Vander Heiden et al., 2009). Gluconeogenesis in HCC is driven by upregulating facilitative glucose transporters, notably GLUT1 (Yamamoto et al., 1990), significant elevation of hypoxic regulators such as HIF-1 α (Yang et al., 2014), and expression of rad and myc oncogenes that fuels glycolysis (Tiniakos et al., 1989).

Tumor pathogenesis encompasses a complex network of regulatory pathways involving various genetic, epigenetic, and epitranscriptomic mechanisms. Aberrant epitranscriptomic RNA modifications have also been shown to drive tumorigenesis due to dysregulation of RNA processing, polyadenylation, translation initiation, splicing, stability, and localization, which affect translation of tumor suppressors and oncogenes (Ferlay et al., 2010). Since each cancer type differs from the other, it is not surprising that different tumor promoters may activate different oncogenic pathways, directly or indirectly affecting RNA modifications, such as m6A and thereby their protein levels. It is evidenced that downregulation of RNA m6A modification can promote tumor progression in several types of cancers, such as glioblastoma, endometrial tumors, and leukemia (Yang and Roberts, 2010). Therefore, it is crucial to explore the relationship between aberrant RNA m6A modifications and hepatocarcinogenesis to better understand the disease etiopathogenesis.

m6A RNA MODIFICATION AND ASSOCIATED REGULATORS

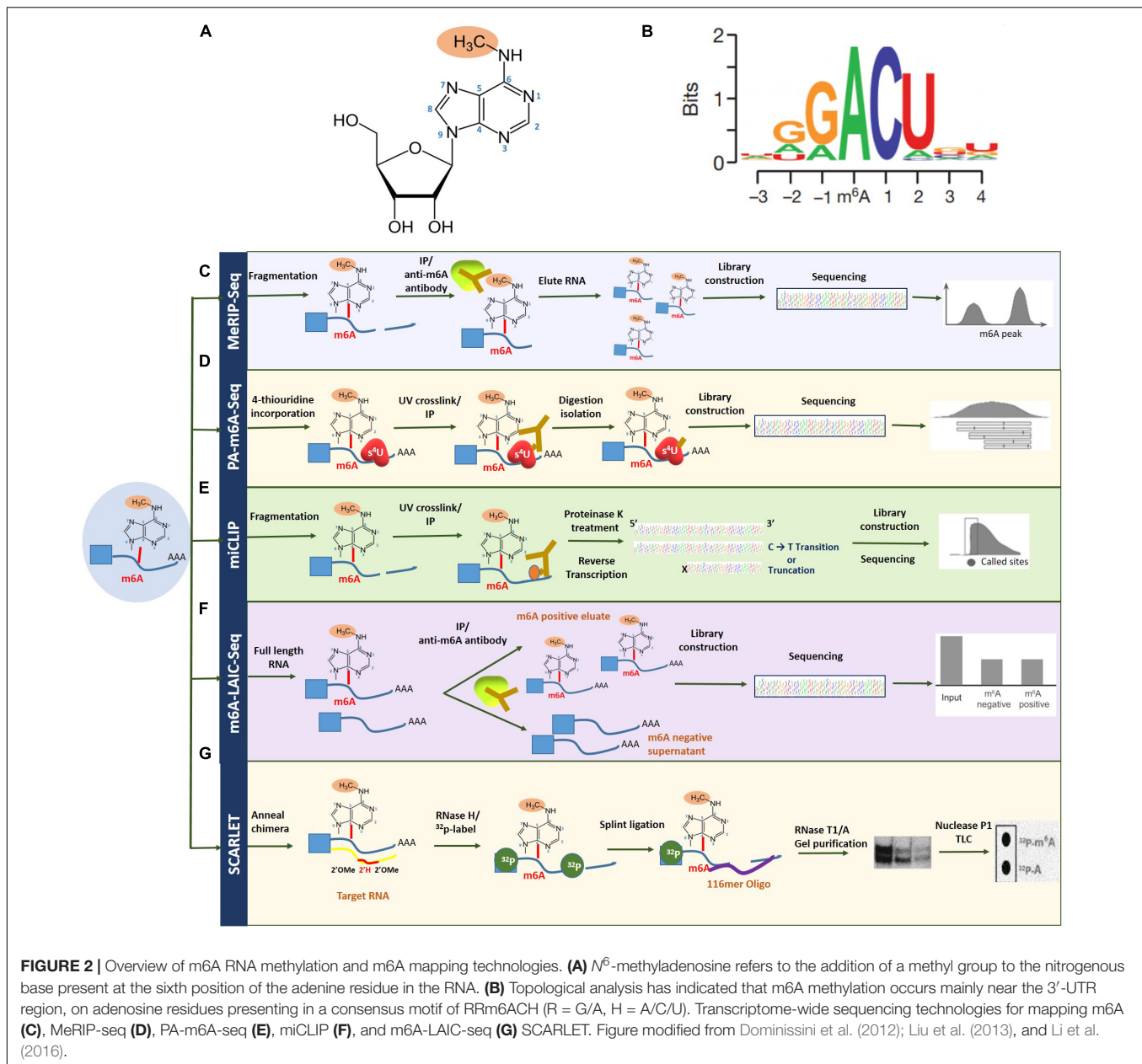
First identified in 1970, N⁶-methyladenosine (m6A) is the most profuse and reversible internal modification omnipresent in eukaryotic mRNA and has been the focus in the emerging field of epitranscriptomics (Wu et al., 2016). While research has

previously explored the crucial role of epigenetic regulation, pertaining to DNA and histone methylation, the dynamic function of m6A RNA modification is a relatively novel and largely unexplored territory, with regard to discovering mechanisms of gene expression regulation (Strahl and Allis, 2000; Suzuki and Bird, 2008; Wu et al., 2016). With the advent of new next-generation sequencing approaches such as MeRIP-seq, over 12,000 highly conserved methylated peaks have been identified in human and mouse transcriptomes, revealing the correlation between m6A abundance and the structural and functional aspects of a specific gene. m6A modification refers to the addition of a methyl group to the nitrogenous base present at the sixth position of the adenine residue in the RNA (Figure 2; Desrosiers et al., 1974). With an approximate estimate of 0.1–0.4% of adenosines subjected to alterations, an average of two to three m6A-modified sites have been predicted to be present in every mRNA transcript (Wei et al., 1975). Topological analysis has indicated that methylation occurs mainly near the 3′-UTR region, on adenosine residues presenting in a consensus motif of RRm6ACH (R = G/A, H = A/C/U; Figure 2; Csepany et al., 1990; Meyer Kate et al., 2012).

Owing to its dynamic and reversible nature, m6A RNA methylation is modulated by a multi-subunit complex of methyltransferase proteins known as “writers,” which add methyl groups to the adenosine; demethylases known as “erasers,” which aid in the removal of methyl groups; and RNA binding proteins known as “readers,” which bind to methylated RNA and regulate discrete downstream mechanisms (Karthiya and Khandelia, 2020).

Writers

The methyltransferase complex is composed of several subunits with key methylases such as METTL3 and METTL14 and regulator proteins such as WTAP, METTL16, ZC3H13, and RBM15/15B (Karthiya and Khandelia, 2020). METTL3 and METTL14, which contain an S-adenosyl methionine (SAM) binding motif, form a core heterodimer and co-localize in nuclear speckles, with the former acting as an enzymatic component and the latter as an allosteric activator (Śledź and Jinek, 2016; Wang et al., 2016). The Wilms tumor 1 (WT1)-associated protein (WTAP), a splicing factor that modulates methylation, regulates the position of the heterodimer while indirectly increasing the catalytic capacity of methyltransferases (Ping et al., 2014; Zhang et al., 2019). RBM15/15B interacts with the methyltransferase complex with the aid of WTAP, further liaises with chromatin remodeling complexes, and modulates cortical development by recruiting the writer complex. ZC3H13 zinc finger protein, on the other hand, acts as a recruiter protein promoting localization of the writer complex in the nucleus. A relatively novel methyltransferase, METTL16, directs deposition of specific RNAs as well as U6 small nuclear RNA in addition to maintaining SAM homeostasis by adding methyl groups to the SAM synthase transcript, thereby gaining control of its stability and splicing mechanisms. Choosing a transcript for methylation is carried out by recruiting methyltransferases to specific promoters by certain transcription factors in addition to being influenced by histone



modifications such as H3K36me3 and H4K20me1 (Kolasinska-Zwiercz et al., 2009; Tilgner et al., 2009; Zaccara et al., 2019).

Erasers

Eraser proteins function primarily by shaping the m6A landscape dynamically. RNA demethylases such as fat mass and obesity-associated protein (FTO; Jia et al., 2011) and AlkB family member 5 (ALKBH5; Zheng et al., 2013) belong to the ALKB family of dioxygenases that appear to have a restricted role under normal physiological conditions, with prominent functions in particular organs such as the testes and in certain ailments. Functionally, FTO performs an indirect role by sequentially oxidizing N^6 -methyladenosine to N^6 -formyladenosine, with an intermediate

hydroxymethyladenosine, while ALKBH5 catalyzes direct removal of methylation (Chen and Wong, 2020).

Readers

The fate of mRNAs containing m6A is predominantly determined by different categories of m6A-binding proteins, termed “readers,” such as YT521-B homology (YTH) domain family, heterogeneous nuclear ribonucleoproteins (hnRNPs), and IGF 2 mRNA-binding proteins (IGFBPs). Such proteins govern the m6A-related downstream cellular mechanisms in tumorigenesis, viral replication, adipogenesis, hemopoiesis, and immune regulation (Zhao et al., 2020).

The YTH domain-containing family (YTHDF) comprises YTHDF1, YTHDF2, YTHDF3, YTHDC1, and YTHDC2

proteins that recognize m6A in a methylation-dependent manner. As the first cytoplasmic reader protein to be discovered, YTHDF2 performs the function of degrading methylated RNA by directly recruiting CCR4–NOT deadenylation complex to the target transcript, inherently reducing its stability and thereafter directing bound mRNA to relevant decay sites such as processing bodies (Wang et al., 2014; Du et al., 2016). YTHDF1 protein significantly augments mRNA translation efficiency through interactions with the translation initiation factor eIF3 and in some cases in an m⁷G-cap-dependent manner (Wang et al., 2015; Liu et al., 2020a). YTHDC1 plays a pertinent role in exon selection during gene splicing (Roundtree and He, 2016). YTHDC2 acts as a putative RNA helicase that governs RNA levels during cell meiosis by forming a complex with meiosis-specific coiled-coil domain-containing protein (MEIOC; Hsu et al., 2017; Jain et al., 2018).

Heterogeneous nuclear ribonucleoprotein protein family encompasses hnRNPs such as A2/B1 (HNRNPA2B1), C (HNRNPC), and G (HNRNPG), which have affinity to structural alterations induced by m6A methylation, commonly known as “m6A switch” (Liu et al., 2015, 2017; Wu et al., 2018). HNRNPA2B1 modulates alternative splicing of mRNA transcripts and processes primary miRNAs through DGCR8-directed interactions (Alarcón et al., 2015). HNRNPC and HNRNPG impact pre-mRNA processing and pre-mRNA alternative splicing, respectively, via interactions with phosphorylated carboxy-terminal domain of enzyme RNA polymerase (Zarnack et al., 2013; Liu et al., 2017). The IGFBP family proteins use a KH RNA binding domain to identify m6A-containing transcripts and exert their function by actively recruiting RNA stabilizers such as HuR to protect mRNA transcripts from degradation (Chen and Wong, 2020).

m6A Mapping Technologies

The field of epitranscriptomics began gaining prominence with the development of methylated RNA immunoprecipitation/m6A sequencing (MeRIP/m6A-seq), capable of conducting a site-specific analysis of m6A modification-based transcriptomic studies. This technique together with ChIP-seq depends on a specific m6A antibody to pull down m6A-containing transcripts that can subsequently be mapped by next-generation sequencing technologies (Zhang and Hamada, 2020). However, to circumvent issues with regard to identifying m6A-modified site, improved technologies such as photo-crosslinking-associated sequencing (PA-m6A-seq), which uses a photo-crosslinking method, and even single-nucleotide resolution m6A mapping are attainable (Linder et al., 2015). Furthermore, with additional UV crosslinking stratagems through techniques such as m6A individual-nucleotide-resolution crosslinking and immunoprecipitation (miCLIP), specific mutations and truncation profiles affected by the presence of m6A can be mapped precisely (Zhang et al., 2019).

Another contemporary technique for high-resolution m6A mapping is the site-specific cleavage and radioactive labeling followed by ligation-assisted extraction and thin-layer chromatography (SCARLET). As implicated by the name, the method involves site-specific cleavage and radiolabeling followed

by splint ligation by DNA ligase, gel purification, and, finally, an analysis using thin-layer chromatography (Liu et al., 2013; Maity and Das, 2016). While this technique cannot employ high-throughput sequencing, newer methods such as m6A level and isoform characterization sequencing (m6A-LAIC-seq), which are compatible with high-throughput sequencing, have lately begun to see the limelight (Molinie et al., 2016). Also, since total RNA is ample without the requirement of enriching the targeted fraction of RNA, this method is deemed suitable for quantifying methylation levels in low abundance RNAs such as tRNAs (Linder et al., 2015). These m6A mapping technologies have been summarized in **Figure 2**. More recently, the third-generation single-nucleotide sequencing technologies such as Nanopore Direct Sequencing, capable of mapping RNA modifications at single base resolution, show promise in identifying m6A sites with improved accuracy (El-Serag and Rudolph, 2007; Fitzmaurice et al., 2019).

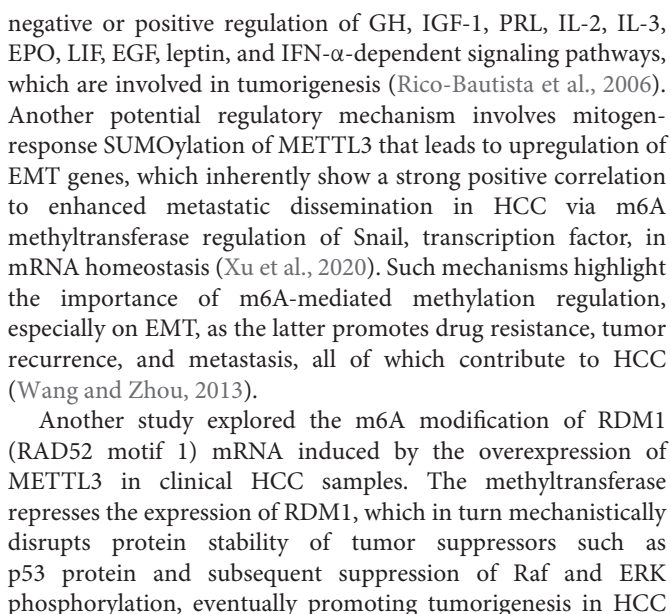
In recent years, several bioinformatics tools and databases addressing different purposes have been developed to organize and integrate complex datasets pertaining to m6A RNA modifications and its associated regulators. An interactive analysis of epitranscriptomic sequencing for m6A site identification can be performed by databases such as deepEA (Zhai et al., 2020) and iMRM (Liu and Chen, 2020), while RNAWRE (Nie et al., 2020), and M6A2Target (Deng et al., 2020) deposit m6A regulator datasets. Furthermore, comprehensive information on reliable m6A methylation sites and peaks from MeRIP-seq data are summarized in m6A-Atlas (Tang et al., 2020) and REPIC (Liu et al., 2020b), respectively. Bioinformatics resources such as RMDisease (Chen K. et al., 2020) and RMVar (Luo et al., 2020) can aid in better understanding the association between various epitranscriptomic modifications and their probable disease relevance.

m6A RNA METHYLATION REGULATORS AND HALLMARKS OF HEPATOCELLULAR CARCINOMA

The impact of m6A writers, readers, and erasers on key pathways that regulate the development of HCC are summarized in **Figure 3** and **Table 1**.

“Writers” in Hepatocellular Carcinoma Hallmarks

In a recent study, Chen et al. (2018) identified suppressor of cytokine signaling 2 (SOCS2) as a downstream target of methyltransferase METTL3-induced m6A modification via the m6A reader protein YTHDF2-dependent pathway (Chen et al., 2018). Previous studies have investigated the function of SOCS2 as a cytokine-inducible negative regulator in Janus kinase/signal transduction and activation of the JAK/STAT pathway (Yoshikawa et al., 2001). SOCS2 downregulation is significantly correlated to advanced Tumor, Node, and Metastasis staging in addition to being a prognostic marker in HCC (Qiu et al., 2013). SOCS2 has also been associated with the



(Chen S. L. et al., 2019). Thus, such cascade of downstream processes leading to cancer could be solely traced back to the oncogenic behavior of METTL3.

Mechanistic studies revealed positive modulation of another subunit of methyltransferase, METTL14, on pri-miR126, a metastasis-inducing miRNA mechanism in a DGCR8-dependent manner (Ma et al., 2017). As a critical constituent of the canonical microprocessor complex for microRNA biogenesis, DGCR8 maintains the upregulation or deregulation of certain tumor-specific miRNAs that intrinsically contribute to enhanced cell proliferation, evasion of apoptosis via angiogenesis, and initiation of invasion and metastatic pathways, especially in HCC (Chu et al., 2014; Deng et al., 2019). Emerging evidence suggests that aberrantly downregulated miR126 boosted the poor overall survival associated with HCC (Bao et al., 2018). Restoring miR126 inhibited cell proliferation in HCC, arrested cell cycle advancement, and induced cell apoptosis (Zhao et al., 2015). Thus, the tumor-suppressive function of METTL14 could be a potential target in developing therapeutics targeting HCC.

TABLE 1 | Role of m6A methylome in hepatocarcinogenesis.

m6A regulator type	m6A regulator	Expression pattern	Function in hepatocellular carcinoma*	References
Writer	METTL3	Upregulated	METTL3 suppresses SOCS2 expression through an m6A-YTHDF2-dependent pathway.	Chen et al., 2018
Writer	METTL3	Upregulated	SUMOylation of METTL3 leads to upregulation of EMT via m6A regulation of Snail transcription factor.	Xu et al., 2020
Writer	METTL3	Upregulated	METTL3 represses the expression of RDM1, which in turn disrupts p53 protein stability.	Chen S. L. et al., 2019
Writer	METTL14	Downregulated	METTL14 liaises with microprocessor protein DGCR8 and positively modifies the microRNA 126 activity in an m6A-dependent manner.	Ma et al., 2017
Writer	WTAP	Upregulated	WTAP drives methylation of ETS1 leading to epigenetic silencing of ETS1 via a HuR-dependent manner.	Chen Y. et al., 2019
Writer	KIAA1429	Upregulated	KIAA1429, with the aid of GATA3-AS, methylates GATA3 pre-mRNA, separating HuR and degrading GATA3 pre-mRNA.	Lan et al., 2019
Eraser	FTO	Upregulated	FTO mechanistically triggers demethylation of PKM2 mRNA and enhances PKM2 translated.	Li et al., 2019
Eraser	FTO	Downregulated	Oncogenic protein SIRT1 downregulates FTO by activating RANBP2 leading to overexpression of m6A + GNAO1.	Liu et al., 2020c
Eraser	ALKBH5	Downregulated	ALKBH5 modulates post-transcriptional inhibition of LY6/PLAUR domain-containing 1 (LYPD1), which in turn is stabilized by m6A reader IGF2BP1.	Chen Y. et al., 2020
Reader	YTHDF1	Upregulated	Modulates PPAR/NOTCH signaling pathways.	Zhao et al., 2018
Reader	YTHDF2	Downregulated	YTHDF2 binds to EGFR 3'-UTR promoting degradation of EGFR mRNA.	Zhong et al., 2019
Reader	YTHDF2	Downregulated	Repressed YTHDF2 activity disrupts tumor vasculature suppression that drives IL11 and SERPINE2 mRNA decay.	Hou et al., 2019
Reader	IGFBP3	Downregulated	IGFBP-3 regulates growth suppression signals via altering TGF- β and/or Rb pathways.	Yumoto et al., 2005
Reader	IGF2BP1	Upregulated	IGF2BP1 stabilizes c-MYC and MKI67 mRNAs and enhances c-Myc and Ki-67 protein translation.	Gutschner et al., 2014
Reader	IGF2BP1	Downregulated	LINC01093 directly binds IGF2BP1, disrupting interactions between IGF2BP1 and GLL1 mRNA leading to the mRNA degradation of the latter.	He et al., 2019

*Based on *in vitro* and *in vivo* studies involving patient samples and hepatocellular carcinoma (HCC) cell lines.

Wilms tumor associated protein, another subunit of the methyltransferase complex that localizes in nuclear speckles, is highly expressed in HCC and serves as an independent predictor of hepatocarcinogenesis (Ping et al., 2014; Chen Y. et al., 2019). WTAP has been shown to promote proliferation and tumorigenesis by epigenetically silencing the ETS1 transcription factor (Chen Y. et al., 2019). Accumulating evidence suggests that ETS proteins regulate various aspects of cancer hallmarks such as proliferation inducing cell signaling, promoting angiogenesis, and evading apoptosis through enhanced nuclear transport (Myers et al., 2005), recruitment of co-repressors (Okamura et al., 2009), increased DNA binding of nuclear proteins (Sharrocks, 2001), and transactivation of certain genes such as VEGF (Tetsu and McCormick, 2017; Fry and Inoue, 2018). For example, ETS1 exhibits preferential binding to wild-type p53, suggesting tumor-suppressive functions in various cancers (Martinez, 2016). ETS1 was also shown to directly upregulate genes necessary for angiogenesis and extracellular matrix remodeling, such as the matrix metalloproteinases MMP-1, MMP-3, and MMP-9 and

integrin β 3 (Oda et al., 1999). Thus, WTAP-mediated epigenetic silencing of ETS1 could indicate a cogent m6A regulator-driven tumorigenesis mechanism at play.

KIAA1429, another crucial component of the m6A methyltransferase complex, is upregulated in HCC cells exhibiting poor prognosis (Lan et al., 2019). Furthermore, based on extensive *in vitro* and *in vivo* studies, they identified GATA binding protein 3 (GATA3) as a direct downstream target of KIAA1429-induced m6A modifications. GATA3 is a transcription factor composed of two zinc fingers at the carboxyl terminus, which has been linked with suppression of metastasis, tumor microenvironment modulation, and promotion of cellular differentiation (Zheng and Blobel, 2010; Chou et al., 2013; Siddiqui et al., 2014). KIAA1429, under the guidance of the long non-coding RNA (lncRNA) GATA3-AS, facilitates m6A methylation on the 3'-UTR of GATA3 pre-mRNA, disrupting the activity of the RNA binding protein HuR to GATA3 pre-mRNA, eventually downregulating GATA3 mRNA expression (Lan et al., 2019). GATA3 targets one of the classical hallmarks

of cancer, the capability of tumor cells to induce invasion and metastasis (Hanahan and Weinberg, 2000; Chou et al., 2010). In fact, a previous study showed that GATA3-AS drives hepatocarcinogenesis via metastasis, particularly in HCC, by suppressing tumor suppressor genes such as p53, PTEN, and key inhibitors, namely, CDKN1A (Luo et al., 2019).

“Erasers” in Hepatocellular Carcinoma Hallmarks

Fat mass and obesity-associated protein demethylase is significantly upregulated in HCC correlating with poor prognosis, while triggering demethylation of pyruvate kinase 2 (PKM2) mRNA (Li et al., 2019). PKM2, a rate-limiting glycolytic muscle isozyme, acts as a catalysis mediator in the irreversible transphosphorylation between adenosine diphosphate and phosphoenolpyruvate, leading to production of pyruvate and ATP (Altenberg and Greulich, 2004). In neoplastic cells, PKM2 is overexpressed under the regulation of oncoproteins such as c-Myc, which activates transcription of hnRNPs, a class of reader proteins that in turn control mRNA m6A regulation (David et al., 2010; Liu et al., 2015). PKM2 thus prompts metabolic reprogramming, a core hallmark of cancer, by facilitating anabolic metabolism in proliferating cells (Ward and Thompson, 2012). Another study that scrutinized the clinicopathological features in HCC patients with high PKM2 levels observed poor prognosis coupled with lower creatinine levels, advanced stage, and higher grade in such groups (Lu et al., 2018). Although Gene Expression Profiling Interactive Analysis (GEPIA) and The Cancer Genome Atlas (TCGA) databases show a strong correlation between the mRNA expression levels of FTO and PKM2, the exact dynamics of m6A demethylases and tumorigenesis inducing altered metabolism portrays a research gap that is yet to be filled.

Sirtuin 1 (SIRT1), also known as NAD-dependent deacetylase SIRT1, is frequently upregulated in HCC, where it regulates chemoresistance and metastasis while maintaining tumorigenicity and self-renewal capabilities of liver cancer stem cells (Wilking and Ahmad, 2015; Farcas et al., 2019). SIRT1 also upregulates oncogenes such as β -catenin (Firestein et al., 2008), HIF-1 α (Laemmle et al., 2012), and c-Myc (Jang et al., 2012), especially in liver cancers. Functioning as an oncogene, SIRT1 downregulates m6A demethylase, FTO, by activating nucleoporin RaBnP2, a protein with a small ubiquitin-related modifier (SUMO) E3 ligase activity. RaBnP2 triggers SUMOylation of FTO at lysine (K)-216 site leading to FTO degradation (Liu et al., 2020c). Furthermore, guanine nucleotide-binding protein G (o) subunit alpha (GNAO1), a tumor suppressor, is an m6A-mediated downstream target of FTO. SIRT1 induces downregulation of FTO, thus leading to degradation of GNAO1, which in turn fuels hepatocarcinogenesis (Liu et al., 2020c).

AlkB family member 5 demethylase is downregulated in HCC with the functional role of suppressing proliferation and invasion capabilities of tumor cells (Chen Y. et al., 2020). Mechanistically, however, ALKBH5-regulated demethylation leads to post-transcriptional inhibition of LY6/PLAUR domain-containing 1

(LYPD1), a neurotransmitter receptor-binding protein involved in the regulation of breast and ovarian cancers, with the potential to act as a prognostic marker (Wang N. et al., 2019). Given that, in general, healthy tissues express relatively low levels of LYPD1 than most peripheral organs (Egerod et al., 2007), downregulation of ALKBH5 leads to significant upregulation of LYPD1, an established oncogenic driver of HCC (Chen Y. et al., 2020). While Chen et al. concluded that an explicit link between LYPD1 and cancer signaling pathways is yet to be established, they were able to provide evidence of the involvement of ALKBH5 in the regulation of P13K/AKT/mTOR and GTPase pathways, major hallmarks of cancer, through gene ontology and ALKBH5/LYPD1 gene knockdown studies (Etienne-Manneville and Hall, 2002; Manning and Toker, 2017; Chen Y. et al., 2020). Potential involvement of ALKBH5/LYPD1-mediated modulation could therefore explain the poor prognosis of HCC.

“Readers” in Hepatocellular Carcinoma Hallmarks

YTH domain-containing family1 m6A reader protein is significantly upregulated in HCC. Based on gene ontology and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analyses, YTHDF1 was found to be associated with p53, NOTCH, and peroxisome proliferator-activated receptors (PPAR) signaling pathways, which are known to aid in HCC progression (Han et al., 2019). PPAR Beta/Delta, ligand-activated transcription factors, have been known to favor pro-tumorigenicity while being central in the interplay of different cancer hallmark capabilities, such as cell proliferation, immune function, induction of angiogenesis, and senescence and replicative immortality (Wagner and Wagner, 2020). For example, in HCC cells, PPAR γ activation leads to cell growth inhibition by overexpression of cell cycle arrest-inducing proteins such as cdc2, p21, p27, and CITED2, in addition to downregulation of cyclin D1, a protein that promotes cell cycle (Hsu and Chi, 2014). Contrarily, oxidative stress imposed by peroxisome proliferators and subsequent induction of PPAR α enables hepatocellular proliferation while inhibiting apoptosis (Reddy et al., 1976; Tachibana et al., 2008). The NOTCH cell fate-regulatory pathway, another probable downstream target of YTHDF1, too, has been shown to be pro-oncogenic, due to its associations with NOTCH coactivator MAML2, a target of genetic alterations, and activation of Sox9- and K19-positive progenitors leading to liver tumorigenesis (Nalesnik et al., 2012; Strazzabosco and Fabris, 2012; Morell and Strazzabosco, 2014). Thus, the relationship between PPAR/NOTCH signaling pathways and the epitranscriptomic role of YTHDF1 is an area that warrants further research.

YTH domain-containing family2 is a novel regulator of tumor-promoting inflammation (Hou et al., 2019), which binds to m6A-containing RNAs and directs them to decay sites for degradation, thereby contributing toward regulating mRNA stability (Du et al., 2016). YTHDF2 inhibits STAT3 phosphorylation and tumorigenesis through interleukin 11 (IL-11) mRNA degradation, which encodes IL-6 family cytokine that triggers HCC with potential for proliferation and metastasis.

The complex role of IL-11 as a pro-inflammatory cytokine in regulating immune response through activation of the JAK-STAT3 signaling pathway, in turn, provides the pro-inflammatory microenvironment required for malignant transformation and tumor progression (Bollrath et al., 2009; Kortylewski et al., 2009). In fact, another recent publication showed that IL-11 levels were significant in postsurgical HCC recurrence due to the associated enhancement of the IL-11-STAT3 signaling (Wang D. et al., 2019). Furthermore, YTHDF2 targets the mRNA degradation of serpin family E member 2 (SERPINE2), a protein present in the extracellular matrix and known to contribute to tumor invasion and metastasis through the oncogenic activation of BRAF, RAS, and MEK, which in turn influence the pro-neoplastic mechanisms of extracellular signal-regulated kinase (ERK) signaling (Bergeron et al., 2010; Yang et al., 2018).

YTH domain-containing family2 represses cell proliferation and activation of MEK and ERK in HCC cells through modification of 3'-UTR site of EGFR, a key factor in epithelial malignancies and directing the subsequent degradation of EGFR mRNA (Zhong et al., 2019). The integrative effects of an enhanced TGF- α -EGFR-MAPK activity on driving neoplasticity in HCC cells (Baek et al., 2010), coupled with a range of cancer hallmarks that EGFR influences, could potentially be negated by the upregulation of m6A regulator YTHDF2.

IGF 2 mRNA-binding protein3 is regarded as a putative tumor suppressor as well as mediator of mechanisms involving growth suppression. Yumoto et al. (2005) postulated that IGFBP3 expression in HCC was concomitant with abnormalities in the TGF- β receptor and/or effectors of its downstream signaling pathway such as Rb (Yumoto et al., 2005). It has been shown that IGFBP3 binds to its putative receptor IGFBP-3R inducing apoptosis, as well as binds to TGF- β receptor, eventually causing Smad activation and thereby inducing apoptosis. This reader protein also activates Stat-1 transcription factor and binds to nuclear receptors like RXR- α , prompting anti-apoptotic and anti-proliferative effects (Zappala et al., 2008; Shahjee and Bhattacharyya, 2014).

IGF 2 mRNA-binding proteins1 is strongly upregulated in HCC, and it stabilizes and upregulates c-Myc and MK167, two major regulators of cell proliferation and apoptosis (Gutschner et al., 2014). This oncofetal reader protein also promotes the expression of serum response factor (SRF), a critical transcription factor that regulates cell adhesion, cytoskeletal regulation, and cell migration, via m6A-mediated impairment of SRF mRNA decay (Descot et al., 2009; Leitner et al., 2011; Esnault et al., 2014). At the post-transcriptional level, IGF2BP1 controls the expression of PDLIM7 and FOXK1, two genes known to promote HCC hallmarks such as tumor proliferation and metastasis (Müller et al., 2019). Recently, another study showed that liver-specific lncRNAs directly bind IGF2BP1, enabling mRNA degradation of transcription factor glioma-associated oncogene homolog 1 (GLI1) mRNA, the latter known to be associated with the hepatocarcinogenesis-inducing Hedgehog pathway (Della Corte et al., 2017; He et al., 2019). Another contemporary work revealed that IGF2BP1 stabilizes the transcript of LINC01138, an oncogenic long intergenic non-coding RNA that promotes tumorigenesis and tumor invasion

(Li et al., 2018). Thus, it can be postulated that IGF2BP1 reader proteins are involved in a variety of complex downstream mechanisms that remarkably influence metastasis hallmark of hepatocarcinogenesis, especially HCC.

CONCLUSION AND FUTURE PERSPECTIVES

In recent years, immense integrative and comprehensive genomic and molecular analyses exploring potential diagnostic and prognostic targets for HCC have brought to light several prominent therapeutic solutions to curb the burden of HCC. While the role of genetic and epigenetic mechanisms on the pathogenesis of hepatocarcinogenesis has been the subject of extensive research, the network of mechanisms that governs epitranscriptomics and its associated RNA modifications, much less its association with liver tumorigenesis, is an emerging field that warrants further investigation. While the field of epitranscriptomics has witnessed a rapid upsurge in research publications in the recent decade, the mechanistic and functional aspects of m6A regulators and methylation levels in hepatocarcinogenesis, particularly HCC, still remains ambiguous, metaphorically signified by the parable of the “blind men and the elephant.” We are yet to find cogent answers to prevailing questions such as the following: How does m6A methylation affect the gene expression regulations in liver tumorigenesis? Which interwoven regulatory networks of pathways contribute toward m6A methylation and the expression of m6A regulator proteins? What m6A-associated mechanistic pathways are modified by external factors such as hepatitis B and C, aflatoxin exposure, and other causal factors of HCC? Do m6A regulators function as oncogenes or tumor suppressors? Only upon identifying the major cancer hallmark influencers of epitranscriptomic regulation can we successfully design therapeutic targets to combat HCC.

AUTHOR CONTRIBUTIONS

ES contributed to manuscript writing, figure preparation, and final editing. NB contributed to critically revising the article and rectifying grammatical errors. ZL, JM, and RR contributed to this manuscript with conception and revision.

FUNDING

This work was supported by the Key Program Special Fund of Xi'an Jiaotong-Liverpool University (KSF-E-23 to RR) and the Research Development Fund of Xi'an Jiaotong-Liverpool University (RDF-17-02-31 to RR).

ACKNOWLEDGMENTS

We apologize to the authors whose excellent works have not been quoted in this review manuscript due to space restriction.

REFERENCES

- Adams, J. M., and Cory, S. (2007). The Bcl-2 apoptotic switch in cancer development and therapy. *Oncogene* 26, 1324–1337. doi: 10.1038/sj.onc.1210220
- Alarcón, C., Goodarzi, H., Lee, H., Liu, X., Tavazoie, S., and Tavazoie, S. (2015). HNRNPA2B1 is a mediator of m6A-dependent nuclear RNA processing events. *Cell* 162, 1299–1308. doi: 10.1016/j.cell.2015.08.011
- Alqahtani, A., Khan, Z., Alloghbi, A., Said Ahmed, T. S., Ashraf, M., and Hammouda, D. M. (2019). Hepatocellular carcinoma: molecular mechanisms and targeted therapies. *Medicina (Kaunas)* 55:526.
- Altenberg, B., and Greulich, K. O. (2004). Genes of glycolysis are ubiquitously overexpressed in 24 cancer classes. *Genomics* 84, 1014–1020. doi: 10.1016/j.ygeno.2004.08.010
- Baek, J. Y., Morris, S. M., Campbell, J., Fausto, N., Yeh, M. M., and Grady, W. M. (2010). TGF-beta inactivation and TGF-alpha overexpression cooperate in an in vivo mouse model to induce hepatocellular carcinoma that recapitulates molecular features of human liver cancer. *Int. J. Cancer* 127, 1060–1071. doi: 10.1002/ijc.25127
- Bao, J., Yu, Y., Chen, J., He, Y., Chen, X., Ren, Z., et al. (2018). MiR-126 negatively regulates PLK-4 to impact the development of hepatocellular carcinoma via ATR/CHEK1 pathway. *Cell Death Dis.* 9:1045.
- Bergeron, S., Lemieux, E., Durand, V., Cagnol, S., Carrier, J. C., Lussier, J. G., et al. (2010). The serine protease inhibitor serpinE2 is a novel target of ERK signaling involved in human colorectal tumorigenesis. *Mol. Cancer* 9:271. doi: 10.1186/1476-4598-9-271
- Beudeker, B. J. B., and Boonstra, A. (2020). Circulating biomarkers for early detection of hepatocellular carcinoma. *Ther. Adv. Gastroenterol.* 13, 1–11.
- Biselli-Chicote, P. M., Oliveira, A. R., Pavarino, E. C., and Goloni-Bertollo, E. M. (2012). VEGF gene alternative splicing: pro- and anti-angiogenic isoforms in cancer. *J. Cancer Res. Clin. Oncol.* 138, 363–370. doi: 10.1007/s00432-011-1073-2
- Bollrath, J., Phesse, T. J., von Burstin, V. A., Putoczki, T., Bennecke, M., Bateman, T., et al. (2009). gp130-Mediated Stat3 activation in enterocytes regulates cell survival and cell-cycle progression during colitis-associated tumorigenesis. *Cancer Cell* 15, 91–102. doi: 10.1016/j.ccr.2009.01.002
- Bosch, F. X., Ribes, J., Díaz, M., and Cléries, R. (2004). Primary liver cancer: worldwide incidence and trends. *Gastroenterology* 127(5 Suppl. 1), S5–S16.
- Campbell, J. S., Hughes, S. D., Gilbertson, D. G., Palmer, T. E., Holdren, M. S., Haran, A. C., et al. (2005). Platelet-derived growth factor C induces liver fibrosis, steatosis, and hepatocellular carcinoma. *Proc. Natl. Acad. Sci. U.S.A.* 102, 3389–3394. doi: 10.1073/pnas.0409722102
- Chen, K., Song, B., Tang, Y., Wei, Z., Xu, Q., Su, J., et al. (2020). RMDisease: a database of genetic variants that affect RNA modifications, with implications for epitranscriptome pathogenesis. *Nucleic Acids Res.* 49, D1396–D1404.
- Chen, M., Wei, L., Law, C. T., Tsang, F. H., Shen, J., Cheng, C. L., et al. (2018). RNA N6-methyladenosine methyltransferase-like 3 promotes liver cancer progression through YTHDF2-dependent posttranscriptional silencing of SOCS2. *Hepatology (Baltimore, Md)* 67, 2254–2270. doi: 10.1002/hep.29683
- Chen, M., and Wong, C.-M. (2020). The emerging roles of N6-methyladenosine (m6A) deregulation in liver carcinogenesis. *Mol. Cancer* 19:44.
- Chen, S.-L., Liu, L. L., Wang, C., Lu, S. X., Yang, X., He, Y., et al. (2019). Loss of RDM1 enhances hepatocellular carcinoma progression via p53 and Ras/Raf/ERK pathways. *Mol. Oncol.* 14, 373–386. doi: 10.1002/1878-0261.12593
- Chen, Y., Peng, C., Chen, J., Chen, D., Yang, B., He, B., et al. (2019). WTAP facilitates progression of hepatocellular carcinoma via m6A-HuR-dependent epigenetic silencing of ETS1. *Mol. Cancer* 18:127.
- Chen, Y., Zhao, Y., Chen, J., Peng, C., Zhang, Y., Tong, R., et al. (2020). ALKBH5 suppresses malignancy of hepatocellular carcinoma via m(6)A-guided epigenetic inhibition of LYPD1. *Mol. Cancer* 19:123.
- Chou, J., Lin, J. H., Brenot, A., Kim, J. W., Provot, S., and Werb, Z. (2013). GATA3 suppresses metastasis and modulates the tumour microenvironment by regulating microRNA-29b expression. *Nat. cell Biol.* 15, 201–213. doi: 10.1038/ncb2672
- Chou, J., Provot, S., and Werb, Z. (2010). GATA3 in development and cancer differentiation: cells GATA have it! *J. Cell. Physiol.* 222, 42–49. doi: 10.1002/jcp.12943
- Chu, R., Mo, G., Duan, Z., Huang, M., Chang, J., Li, X., et al. (2014). miRNAs affect the development of hepatocellular carcinoma via dysregulation of their biogenesis and expression. *Cell Commun. Signal.* 12:45.
- Csepány, T., Lin, A., Baldick, C., and Beemon, K. (1990). Sequence specificity of mRNA N6-adenosine methyltransferase. *J. Biol. Chem.* 265, 20117–20122. doi: 10.1016/s0021-9258(17)30477-5
- David, C. J., Chen, M., Assanah, M., Canoll, P., and Manley, J. L. (2010). HnRNP proteins controlled by c-Myc deregulate pyruvate kinase mRNA splicing in cancer. *Nature* 463, 364–368. doi: 10.1038/nature08697
- Della Corte, C. M., Viscardi, G., Papaccio, F., Esposito, G., Martini, G., Ciardiello, D., et al. (2017). Implication of the hedgehog pathway in hepatocellular carcinoma. *World J. Gastroenterol.* 23, 4330–4340. doi: 10.3748/wjg.v23.i24.4330
- DeNardo, D. G., Andreu, P., and Coussens, L. M. (2010). Interactions between lymphocytes and myeloid cells regulate pro- versus anti-tumor immunity. *Cancer Metastasis Rev.* 29, 309–316. doi: 10.1007/s10555-010-9223-6
- Deng, L., Ren, R., Liu, Z., Song, M., Li, J., Wu, Z., et al. (2019). Stabilizing heterochromatin by DGC8 alleviates senescence and osteoarthritis. *Nat. Commun.* 10:3329.
- Deng, S., Zhang, H., Zhu, K., Li, X., Ye, Y., Li, R., et al. (2020). M6A2Target: a comprehensive database for targets of m6A writers, erasers and readers. *Brief. Bioinform.* 1–11.
- Descot, A., Hoffmann, R., Shaposhnikov, D., Reschke, M., Ullrich, A., and Posern, G. (2009). Negative regulation of the EGFR-MAPK cascade by actin-MAL-mediated Mig6/Errfi-1 induction. *Mol. Cell* 35, 291–304. doi: 10.1016/j.molcel.2009.07.015
- Desrosiers, R., Friderici, K., and Rottman, F. (1974). Identification of methylated nucleosides in messenger RNA from Novikoff hepatoma cells. *Proc. Natl. Acad. Sci. U.S.A.* 71, 3971–3975. doi: 10.1073/pnas.71.10.3971
- Dominissini, D., Moshitch-Moshkovitz, S., Schwartz, S., Salmon-Divon, M., Ungar, L., Osenberg, S., et al. (2012). Topology of the human and mouse m6A RNA methylomes revealed by m6A-seq. *Nature* 485, 201–206. doi: 10.1038/nature11112
- Donaires, F. S., Scatena, N. F., Alves-Paiva, R. M., Podlevsky, J. D., Logeswaran, D., Santana, B. A., et al. (2017). Telomere biology and telomerase mutations in cirrhotic patients with hepatocellular carcinoma. *PLoS One* 12:e0183287. doi: 10.1371/journal.pone.0183287
- Du, H., Zhao, Y., He, J., Zhang, Y., Xi, H., Liu, M., et al. (2016). YTHDF2 destabilizes m6A-containing RNA through direct recruitment of the CCR4–NOT deadenylase complex. *Nat. Commun.* 7:12626.
- Egeblad, M., Nakasone, E. S., and Werb, Z. (2010). Tumors as organs: complex tissues that interface with the entire organism. *Dev. Cell* 18, 884–901. doi: 10.1016/j.devcel.2010.05.012
- Egerod, K., Holst, B., Petersen, P., Hansen, J., Mulder, J., Hokfelt, T., et al. (2007). GPR39 splice variants versus antisense gene LYPD1: expression and regulation in gastrointestinal tract, endocrine pancreas, liver, and white adipose tissue. *Mol. Endocrinol.* 21, 1685–1698. doi: 10.1210/me.2007-0055
- El-Serag, H. B., and Rudolph, K. L. (2007). Hepatocellular carcinoma: epidemiology and molecular carcinogenesis. *Gastroenterology* 132, 2557–2576. doi: 10.1053/j.gastro.2007.04.061
- Enguita-Germán, M., and Fortes, P. (2014). Targeting the insulin-like growth factor pathway in hepatocellular carcinoma. *World J. Hepatol.* 6, 716–737. doi: 10.4254/wjh.v6.i10.716
- Esnault, C., Stewart, A., Gualdrini, F., East, P., Horswell, S., Matthews, N., et al. (2014). Rho-actin signaling to the MRTF coactivators dominates the immediate transcriptional response to serum in fibroblasts. *Genes Dev.* 28, 943–958. doi: 10.1101/gad.239327.114
- Etienne-Manneville, S., and Hall, A. (2002). Rho GTPases in cell biology. *Nature* 420, 629–635.
- Farazi, P. A., Glickman, J., Horner, J., and Depinho, R. A. (2006). Cooperative interactions of p53 mutation, telomere dysfunction, and chronic liver damage in hepatocellular carcinoma progression. *Cancer Res.* 66, 4766–4773. doi: 10.1158/0008-5472.can-05-4608
- Farcas, M., Gavrea, A.-A., Gulei, D., Ionescu, C., Irimie, A., Catana, C. S., et al. (2019). SIRT1 in the development and treatment of hepatocellular carcinoma. *Front. Nutr.* 6:148. doi: 10.3389/fnut.2019.00148
- Ferlay, J., Shin, H. R., Bray, F., Forman, D., Mathers, C., and Parkin, D. M. (2010). Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. *Int. J. Cancer* 127, 2893–2917. doi: 10.1002/ijc.25516

- Firestein, R., Blander, G., Michan, S., Oberdoerffer, P., Ogino, S., Campbell, J., et al. (2008). The SIRT1 deacetylase suppresses intestinal tumorigenesis and colon cancer growth. *PLoS One* 3:e2020. doi: 10.1371/journal.pone.0002020
- Fitzmaurice, C., Abate, D., Abbasi, N., Abbastabar, H., Abd-Allah, F., Abdel-Rahman, O., et al. (2019). Global, regional, and national cancer incidence, mortality, years of life lost, years lived with disability, and disability-adjusted life-years for 29 Cancer Groups, 1990 to 2017: a systematic analysis for the Global Burden of Disease Study. *JAMA Oncol.* 5, 1749–1768.
- Fry, E. A., and Inoue, K. (2018). Aberrant expression of ETS1 and ETS2 proteins in cancer. *Cancer Rep. Rev.* 2.
- Gutschner, T., Hämmerle, M., Pazaitis, N., Bley, N., Fiskin, E., Uckelmann, H., et al. (2014). Insulin-like growth factor 2 mRNA-binding protein 1 (IGF2BP1) is an important protumorigenic factor in hepatocellular carcinoma. *Hepatology* 59, 1900–1911. doi: 10.1002/hep.26997
- Hamdy, M., Shaheen, K., Awad, M., Barakat, E., Shalaby, S., Gupta, N., et al. (2020). Vascular endothelial growth factor (VEGF) as a biochemical marker for the diagnosis of hepatocellular carcinoma (HCC). *Clin. Pract.* 17, 1441–1453.
- Han, D., Liu, J., Chen, C., Dong, L., Liu, Y., Chang, R., et al. (2019). Anti-tumour immunity controlled through mRNA m6A methylation and YTHDF1 in dendritic cells. *Nature* 566, 270–274. doi: 10.1038/s41586-019-0916-x
- Hanahan, D., and Weinberg, R. A. (2000). The hallmarks of cancer. *Cell* 100, 57–70.
- Hanahan, D., and Weinberg, R. A. (2011). Hallmarks of cancer: the next generation. *Cell* 144, 646–674. doi: 10.1016/j.cell.2011.02.013
- He, J., Zuo, Q., Hu, B., Jin, H., Wang, C., Cheng, Z., et al. (2019). A novel, liver-specific long noncoding RNA LINC01093 suppresses HCC progression by interaction with IGF2BP1 to facilitate decay of GLI1 mRNA. *Cancer Lett.* 450, 98–109. doi: 10.1016/j.canlet.2019.02.033
- Hou, J., Zhang, H., Liu, J., Zhao, Z., Wang, J., Lu, Z., et al. (2019). YTHDF2 reduction fuels inflammation and vascular abnormalization in hepatocellular carcinoma. *Mol. Cancer* 18:163.
- Hsu, H.-T., and Chi, C.-W. (2014). Emerging role of the peroxisome proliferator-activated receptor- γ in hepatocellular carcinoma. *J. Hepatocell. Carcinoma* 1, 127–135. doi: 10.2147/jhc.s48512
- Hsu, P. J., Zhu, Y., Ma, H., Guo, Y., Shi, X., Liu, Y., et al. (2017). Ythdc2 is an N(6)-methyladenosine binding protein that regulates mammalian spermatogenesis. *Cell Res.* 27, 1115–1127. doi: 10.1038/cr.2017.99
- Hugo, H., Ackland, M. L., Blick, T., Lawrence, M. G., Clements, J. A., Williams, E. D., et al. (2007). Epithelial–mesenchymal and mesenchymal–epithelial transitions in carcinoma progression. *J. Cell. Physiol.* 213, 374–383. doi: 10.1002/jcp.21223
- Jackson, S. P., and Bartek, J. (2009). The DNA-damage response in human biology and disease. *Nature*. 461, 1071–1078. doi: 10.1038/nature08467
- Jain, D., Puno, M. R., Meydan, C., Lailier, N., Mason, C. E., Lima, C. D., et al. (2018). ketu mutant mice uncover an essential meiotic function for the ancient RNA helicase YTHDC2. *Elife* 7:e30919.
- Jang, K. Y., Noh, S. J., Lehwald, N., Tao, G.-Z., Bellovin, D. I., Park, H. S., et al. (2012). SIRT1 and c-Myc promote liver tumor cell survival and predict poor survival of human hepatocellular carcinomas. *PLoS One* 7:e45119. doi: 10.1371/journal.pone.0045119
- Jia, G., Fu, Y., Zhao, X., Dai, Q., Zheng, G., Yang, Y., et al. (2011). N6-Methyladenosine in nuclear RNA is a major substrate of the obesity-associated FTO. *Nat. Chem. Biol.* 7, 885–887. doi: 10.1038/nchembio.687
- Karthiya, R., and Khandelja, P. (2020). m6A RNA methylation: ramifications for gene expression and human health. *Mol. Biotechnol.* 62, 467–484. doi: 10.1007/s12033-020-00269-5
- Kedmi, M., Ben-Chetrit, N., Körner, C., Mancini, M., Ben-Moshe, N. B., Lauriola, M., et al. (2015). EGF induces microRNAs that target suppressors of cell migration: miR-15b targets MTSS1 in breast cancer. *Sci. Signal.* 8:ra29. doi: 10.1126/scisignal.2005866
- Klymkowsky, M. W., and Savagner, P. (2009). Epithelial-mesenchymal transition: a cancer researcher's conceptual friend and foe. *Am. J. Pathol.* 174, 1588–1593.
- Kolasinska-Zwierz, P., Down, T., Latorre, I., Liu, T., Liu, X. S., and Ahringer, J. (2009). Differential chromatin marking of introns and expressed exons by H3K36me3. *Nat. Genet.* 41, 376–381. doi: 10.1038/ng.322
- Kortylewski, M., Xin, H., Kujawski, M., Lee, H., Liu, Y., Harris, T., et al. (2009). Regulation of the IL-23 and IL-12 balance by Stat3 signaling in the tumor microenvironment. *Cancer Cell* 15, 114–123. doi: 10.1016/j.ccr.2008.12.018
- Laemmle, A., Lechleiter, A., Roh, V., Schwarz, C., Portmann, S., Furer, C., et al. (2012). Inhibition of SIRT1 impairs the accumulation and transcriptional activity of HIF-1 α protein under hypoxic conditions. *PLoS One* 7:e33433. doi: 10.1371/journal.pone.0033433
- Lan, T., Li, H., Zhang, D., Xu, L., Liu, H., Hao, X., et al. (2019). KIAA1429 contributes to liver cancer progression through N6-methyladenosine-dependent post-transcriptional modification of GATA3. *Mol. Cancer* 18:186.
- Lee, S. H., Shin, M. S., Lee, H. S., Bae, J. H., Lee, H. K., Kim, H. S., et al. (2001). Expression of Fas and Fas-related molecules in human hepatocellular carcinoma. *Hum. Pathol.* 32, 250–256. doi: 10.1053/hupa.2001.22769
- Leitner, L., Shaposhnikov, D., Mengel, A., Descot, A., Julien, S., Hoffmann, R., et al. (2011). MAL/MRTF-A controls migration of non-invasive cells by upregulation of cytoskeleton-associated proteins. *J. Cell Sci.* 124, 4318–4331. doi: 10.1242/jcs.092791
- Leong, T. Y., and Leong, A. S. (2005). Epidemiology and carcinogenesis of hepatocellular carcinoma. *HPB* 7, 5–15. doi: 10.1080/13651820410024021
- Li, J., Zhu, L., Shi, Y., Liu, J., Lin, L., and Chen, X. (2019). M6A demethylase FTO promotes hepatocellular carcinoma tumorigenesis via mediating PKM2 demethylation. *Am. J. Transl. Res.* 11, 6084–6092.
- Li, X., Xiong, X., and Yi, C. (2016). Epitranscriptome sequencing technologies: decoding RNA modifications. *Nat. Methods*. 14, 23–31. doi: 10.1038/nmeth.4110
- Li, Z., Zhang, J., Liu, X., Li, S., Wang, Q., Di, C., et al. (2018). The LINC01138 drives malignancies via activating arginine methyltransferase 5 in hepatocellular carcinoma. *Nat. Commun.* 9:1572.
- Libbrecht, L., Desmet, V., and Roskams, T. (2005). Preneoplastic lesions in human hepatocarcinogenesis. *Liver Int.* 25, 16–27. doi: 10.1111/j.1478-3231.2005.01016.x
- Linder, B., Grozhik, A. V., Olarerin-George, A. O., Meydan, C., Mason, C. E., and Jaffrey, S. R. (2015). Single-nucleotide-resolution mapping of m6A and m6Am throughout the transcriptome. *Nat. Methods*. 12, 767–772. doi: 10.1038/nmeth.3453
- Liu, K., and Chen, W. (2020). iMRM: a platform for simultaneously identifying multiple kinds of RNA modifications. *Bioinformatics (Oxford, England)* 36, 3336–3342. doi: 10.1093/bioinformatics/btaa155
- Liu, N., Dai, Q., Zheng, G., He, C., Parisien, M., and Pan, T. (2015). N(6)-methyladenosine-dependent RNA structural switches regulate RNA-protein interactions. *Nature* 518, 560–564. doi: 10.1038/nature14234
- Liu, N., Parisien, M., Dai, Q., Zheng, G., He, C., and Pan, T. (2013). Probing N6-methyladenosine RNA modification status at single nucleotide resolution in mRNA and long noncoding RNA. *RNA* 19, 1848–1856. doi: 10.1261/rna.041178.113
- Liu, N., Zhou, K. I., Parisien, M., Dai, Q., Diatchenko, L., and Pan, T. J. (2017). N6-methyladenosine alters RNA structure to regulate binding of a low-complexity protein. *Nucleic Acids Res.* 45, 6051–6063. doi: 10.1093/nar/gkx141
- Liu, S., Li, G., Li, Q., Zhang, Q., Zhuo, L., Chen, X., et al. (2020a). The roles and mechanisms of YTH domain-containing proteins in cancer development and progression. *Am. J. Cancer Res.* 10, 1068–1084.
- Liu, S., Zhu, A., He, C., and Chen, M. (2020b). REPIC: a database for exploring the N6-methyladenosine methylome. *Genome Biol.* 21:100.
- Liu, X., Liu, J., Xiao, W., Zeng, Q., Bo, H., Zhu, Y., et al. (2020c). SIRT1 regulates N(6)-methyladenosine RNA modification in hepatocarcinogenesis by inducing RANBP2-dependent FTO SUMOylation. *Hepatology (Baltimore, Md)* 72, 2029–2050.
- Llovet, J. M., and Bruix, J. (2008). Molecular targeted therapies in hepatocellular carcinoma. *Hepatology (Baltimore, Md)* 48, 1312–1327.
- Llovet, J. M., Ricci, S., Mazzaferro, V., Hilgard, P., Gane, E., Blanc, J. F., et al. (2008). Sorafenib in advanced hepatocellular carcinoma. *N. Engl. J. Med.* 359, 378–390.
- Lu, D.-H., Lv, W.-W., Li, W.-X., and Gao, Y.-D. (2018). High PKM2 expression is independently correlated with decreased overall survival in hepatocellular carcinoma. *Oncol. Lett.* 16, 3603–3610.
- Luo, X., Li, H., Liang, J., Zhao, Q., Xie, Y., Ren, J., et al. (2020). RMVar: an updated database of functional variants involved in RNA modifications. *Nucleic Acids Res.* 49, D1405–D1412.
- Luo, X., Zhou, N., Wang, L., Zeng, Q., and Tang, H. (2019). Long noncoding RNA GATA3-AS1 promotes cell proliferation and metastasis in hepatocellular

- carcinoma by suppression of PTEN, CDKN1A, and TP53. *Can. J. Gastroenterol. Hepatol.* 2019;1389653.
- Ma, J. Z., Yang, F., Zhou, C. C., Liu, F., Yuan, J. H., Wang, F., et al. (2017). METTL14 suppresses the metastatic potential of hepatocellular carcinoma by modulating N(6)-methyladenosine-dependent primary MicroRNA processing. *Hepatology (Baltimore, Md.)* 65, 529–543. doi: 10.1002/hep.28885
- Maity, A., and Das, B. (2016). N6-methyladenosine modification in mRNA: machinery, function and implications for health and diseases. *FEBS J.* 283, 1607–1630. doi: 10.1111/febs.13614
- Makarova-Rusher, O. V., Medina-Echeverez, J., Duffy, A. G., and Greten, T. F. (2015). The yin and yang of evasion and immune activation in HCC. *J. Hepatol.* 62, 1420–1429. doi: 10.1016/j.jhep.2015.02.038
- Manning, B. D., and Toker, A. (2017). AKT/PKB signaling: navigating the network. *Cell* 169, 381–405. doi: 10.1016/j.cell.2017.04.001
- Martinez, L. A. (2016). Mutant p53 and ETS2, a Tale of Reciprocity. *Front. Oncol.* 6:35. doi: 10.3389/fonc.2016.00035
- Matsuzaki, K., Date, M., Furukawa, F., Tahashi, Y., Matsushita, M., Sakitani, K., et al. (2000). Autocrine stimulatory mechanism by transforming growth factor beta in human hepatocellular carcinoma. *Cancer Res.* 60, 1394–1402.
- Meyer Kate, D., Saletore, Y., Zumbo, P., Elemento, O., Mason Christopher, E., and Jaffrey Samie, R. (2012). Comprehensive Analysis of mRNA Methylation Reveals Enrichment in 3' UTRs and near Stop Codons. *Cell* 149, 1635–1646.
- Mínguez, B., and Lachenmayer, A. (2011). Diagnostic and prognostic molecular markers in hepatocellular carcinoma. *Dis. Markers* 31, 181–190. doi: 10.1155/2011/310675
- Molinie, B., Wang, J., Lim, K. S., Hillebrand, R., Lu, Z.-x., Van Wittenberghe, N., et al. (2016). m6A-LAIC-seq reveals the census and complexity of the m6A epitranscriptome. *Nat. Methods* 13, 692–698. doi: 10.1038/nmeth.3898
- Morell, C. M., and Strazzabosco, M. (2014). Notch signaling and new therapeutic options in liver disease. *J. Hepatol.* 60, 885–890. doi: 10.1016/j.jhep.2013.11.028
- Mott, J. L., and Gores, G. J. (2007). Piercing the armor of hepatobiliary cancer: Bcl-2 homology domain 3 (BH3) mimetics and cell death. *Hepatology (Baltimore, Md.)* 46, 906–911. doi: 10.1002/hep.21812
- Müller, S., Glaß, M., Singh, A. K., Haase, J., Bley, N., Fuchs, T., et al. (2019). IGF2BP1 promotes SRF-dependent transcription in cancer in a m6A- and miRNA-dependent manner. *Nucleic Acids Res.* 47, 375–390. doi: 10.1093/nar/gky1012
- Myers, E., Hill, A. D., Kelly, G., McDermott, E. W., O'Higgins, N. J., Buggy, Y., et al. (2005). Associations and interactions between Ets-1 and Ets-2 and coregulatory proteins, SRC-1, AIB1, and NCoR in breast cancer. *Clin. Cancer Res.* 11, 2111–2122. doi: 10.1158/1078-0432.ccr-04-1192
- Nalesnik, M. A., Tseng, G., Ding, Y., Xiang, G.-S., Zheng, Z.-I., Yu, Y., et al. (2012). Gene deletions and amplifications in human hepatocellular carcinomas: correlation with hepatocyte growth regulation. *Am. J. Pathol.* 180, 1495–1508. doi: 10.1016/j.ajpath.2011.12.021
- Nault, J.-C., Ningarhari, M., Rebouissou, S., and Zucman-Rossi, J. (2019). The role of telomeres and telomerase in cirrhosis and liver cancer. *Nat. Rev. Gastroenterol. Hepatol.* 16, 544–558. doi: 10.1038/s41575-019-0165-3
- Nie, F., Feng, P., Song, X., Wu, M., Tang, Q., and Chen, W. (2020). RNAWRE: a resource of writers, readers and erasers of RNA modifications. *Database (Oxford)* 2020:baaa049.
- Niu, Z.-S., Niu, X.-J., and Wang, W.-H. (2016). Genetic alterations in hepatocellular carcinoma: An update. *World J. Gastroenterol.* 22, 9069–9095. doi: 10.3748/wjg.v22.i41.9069
- Oda, N., Abe, M., and Sato, Y. (1999). ETS-1 converts endothelial cells to the angiogenic phenotype by inducing the expression of matrix metalloproteinases and integrin $\beta 3$. *J. Cell. Physiol.* 178, 121–132. doi: 10.1002/(sici)1097-4652(199902)178:2<121::aid-jcp1>3.0.co;2-f
- Okada, H., Honda, M., Campbell, J. S., Takegoshi, K., Sakai, Y., Yamashita, T., et al. (2015). Inhibition of microRNA-214 ameliorates hepatic fibrosis and tumor incidence in platelet-derived growth factor C transgenic mice. *Cancer Sci.* 106, 1143–1152. doi: 10.1111/cas.12730
- Okamura, K., Yamashita, S., Ando, H., Horibata, Y., Aoyama, C., Takagishi, K., et al. (2009). Identification of nuclear localization and nuclear export signals in Ets2, and the transcriptional regulation of Ets2 and CTP:phosphocholine cytidyltransferase alpha in tetradecanoyl-13-acetate or macrophage-colony stimulating factor stimulated RAW264 cells. *Biochim. Biophys. Acta* 1791, 173–182. doi: 10.1016/j.bbap.2008.12.016
- Ping, X. L., Sun, B. F., Wang, L., Xiao, W., Yang, X., Wang, W. J., et al. (2014). Mammalian WTAP is a regulatory subunit of the RNA N6-methyladenosine methyltransferase. *Cell Res.* 24, 177–189. doi: 10.1038/cr.2014.3
- Qian, B. Z., and Pollard, J. W. (2010). Macrophage diversity enhances tumor progression and metastasis. *Cell* 141, 39–51. doi: 10.1016/j.cell.2010.03.014
- Qiu, X., Zheng, J., Guo, X., Gao, X., Liu, H., Tu, Y., et al. (2013). Reduced expression of SOCS2 and SOCS6 in hepatocellular carcinoma correlates with aggressive tumor progression and poor prognosis. *Mol. Cell. Biochem.* 378, 99–106. doi: 10.1007/s11010-013-1599-5
- Rao, C. V., Asch, A. S., and Yamada, H. Y. (2017). Frequently mutated genes/pathways and genomic instability as prevention targets in liver cancer. *Carcinogenesis* 38, 2–11. doi: 10.1093/carcin/bgw118
- Reddy, J. K., Rao, S., and Moody, D. E. (1976). Hepatocellular carcinomas in acatalasemic mice treated with nafenopin, a hypolipidemic peroxisome proliferator. *Cancer Res.* 36, 1211–1217.
- Rico-Bautista, E., Flores-Morales, A., and Fernández-Pérez, L. (2006). Suppressor of cytokine signaling (SOCS) 2, a protein with multiple functions. *Cytokine Growth Factor Rev.* 17, 431–439. doi: 10.1016/j.cytogfr.2006.09.008
- Roundtree, I. A., and He, C. (2016). Nuclear m(6)A reader YTHDC1 regulates mRNA splicing. *Trends Genet. TIG* 32, 320–321. doi: 10.1016/j.tig.2016.03.006
- Sever, R., and Brugge, J. S. (2015). Signal transduction in cancer. *Cold Spring Harb. Perspect. Med.* 5:a006098.
- Shahjee, H. M., and Bhattacharyya, N. (2014). Activation of various downstream signaling molecules by IGFBP-3. *J. Cancer Ther.* 5, 830–835. doi: 10.4236/jct.2014.59091
- Sharrocks, A. D. (2001). The ETS-domain transcription factor family. *Nat. Rev. Mol. Cell Biol.* 2, 827–837.
- Shen, S., Dean, D. C., Yu, Z., and Duan, Z. (2019). Role of cyclin-dependent kinases (CDKs) in hepatocellular carcinoma: therapeutic potential of targeting the CDK signaling pathway. *Hepatol. Res.* 49, 1097–1108. doi: 10.1111/hepr.13353
- Siddiqui, M. T., Seydafkan, S., and Cohen, C. (2014). GATA3 expression in metastatic urothelial carcinoma in fine needle aspiration cell blocks: A review of 25 cases. *Diagn. Cytopathol.* 42, 809–815. doi: 10.1002/dc.23131
- Singh, A. K., Kumar, R., and Pandey, A. K. (2018). Hepatocellular carcinoma: causes, mechanism of progression and biomarkers. *Curr. Chem. Genom. Transl. Med.* 12, 9–26. doi: 10.2174/2213988501812010009
- Śledź, P., and Jinek, M. (2016). Structural insights into the molecular mechanism of the m(6)A writer complex. *Elife* 5:e18434.
- Stanaway, J. D., Flaxman, A. D., Naghavi, M., Fitzmaurice, C., Vos, T., Abubakar, I., et al. (2016). The global burden of viral hepatitis from 1990 to 2013: findings from the Global Burden of Disease Study 2013. *Lancet (London, England)* 388, 1081–1088.
- Strahl, B. D., and Allis, C. D. (2000). The language of covalent histone modifications. *Nature* 403, 41–45. doi: 10.1038/47412
- Strazzabosco, M., and Fabris, L. (2012). Notch signaling in hepatocellular carcinoma: guilty in association! *Gastroenterology* 143, 1430–1434. doi: 10.1053/j.gastro.2012.10.025
- Suzuki, M. M., and Bird, A. (2008). DNA methylation landscapes: provocative insights from epigenomics. *Nat. Rev. Genet.* 9, 465–476. doi: 10.1038/nrg2341
- Tachibana, K., Yamasaki, D., Ishimoto, K., and Doi, T. (2008). The role of PPARs in cancer. *PPAR Res.* 2008:102737.
- Tang, Y., Chen, K., Song, B., Ma, J., Wu, X., Xu, Q., et al. (2020). m6A-Atlas: a comprehensive knowledgebase for unraveling the N6-methyladenosine (m6A) epitranscriptome. *Nucleic Acids Res.* 49, D134–D143.
- Tetsu, O., and McCormick, F. (2017). ETS-targeted therapy: can it substitute for MEK inhibitors? *Clin. Transl. Med.* 6:16.
- Tilgner, H., Nikolaou, C., Althammer, S., Sammeth, M., Beato, M., Valcárcel, J., et al. (2009). Nucleosome positioning as a determinant of exon recognition. *Nat. Struct. Mol. Biol.* 16, 996–1001. doi: 10.1038/nsmb.1658
- Tiniakos, D., Spandidos, D. A., Kakkanas, A., Pintzas, A., Pollice, L., and Tiniakos, G. (1989). Expression of ras and myc oncogenes in human hepatocellular carcinoma and non-neoplastic liver tissues. *Anticancer Res.* 9, 715–721.
- Vander Heiden, M. G., Cantley, L. C., and Thompson, C. B. (2009). Understanding the Warburg effect: the metabolic requirements of cell proliferation. *Science* 324, 1029–1033. doi: 10.1126/science.1160809

- Wagner, N., and Wagner, K.-D. (2020). PPAR Beta/Delta and the hallmarks of cancer. *Cells* 9:1133. doi: 10.3390/cells9051133
- Wang, D., Zheng, X., Fu, B., Nian, Z., Qian, Y., Sun, R., et al. (2019). Hepatectomy promotes recurrence of liver cancer by enhancing IL-11-STAT3 signaling. *EBioMedicine* 46, 119–132. doi: 10.1016/j.ebiom.2019.07.058
- Wang, N., Zhong, C., Fu, M., Li, L., Wang, F., Lv, P., et al. (2019). Long non-coding RNA HULC promotes the development of breast cancer through regulating LYPD1 expression by sponging miR-6754-5p. *Oncotargets Ther.* 12, 10671–10679. doi: 10.2147/ott.s226040
- Wang, P., Doxtader, K. A., and Nam, Y. (2016). Structural basis for cooperative function of Mettl3 and Mettl14 methyltransferases. *Mol. Cell* 63, 306–317. doi: 10.1016/j.molcel.2016.05.041
- Wang, X., Lu, Z., Gomez, A., Hon, G. C., Yue, Y., Han, D., et al. (2014). N6-methyladenosine-dependent regulation of messenger RNA stability. *Nature* 505, 117–120. doi: 10.1038/nature12730
- Wang, X., Zhao, B. S., Roundtree, I. A., Lu, Z., Han, D., Ma, H., et al. (2015). N(6)-methyladenosine modulates messenger RNA translation efficiency. *Cell* 161, 1388–1399. doi: 10.1016/j.cell.2015.05.014
- Wang, Y., and Zhou, B. P. (2013). Epithelial-mesenchymal transition—a hallmark of breast cancer metastasis. *Cancer Hallm.* 1, 38–49. doi: 10.1166/ch.2013.1004
- Ward, P. S., and Thompson, C. B. (2012). Metabolic reprogramming: a cancer hallmark even warburg did not anticipate. *Cancer Cell* 21, 297–308. doi: 10.1016/j.ccr.2012.02.014
- Wei, C. M., Gershowitz, A., and Moss, B. (1975). Methylated nucleotides block 5' terminus of HeLa cell messenger RNA. *Cell* 4, 379–386. doi: 10.1016/0092-8674(75)90158-0
- Wilking, M. J., and Ahmad, N. (2015). The role of SIRT1 in cancer: the saga continues. *Am. J. Pathol.* 185, 26–28.
- Wu, B., Su, S., Patil, D. P., Liu, H., Gan, J., Jaffrey, S. R., et al. (2018). Molecular basis for the specific and multivalent recognitions of RNA substrates by human hnRNP A2/B1. *Nat. Commun.* 9:420.
- Wu, R., Jiang, D., Wang, Y., and Wang, X. (2016). N6-methyladenosine (m6A) methylation in mRNA with a dynamic and reversible epigenetic modification. *Mol. Biotechnol.* 58, 450–459. doi: 10.1007/s12033-016-9947-9
- Xu, F., Jin, T., Zhu, Y., and Dai, C. (2018). Immune checkpoint therapy in liver cancer. *J. Exp. Clin. Cancer Res.* 37:110.
- Xu, H., Wang, H., Zhao, W., Fu, S., Li, Y., Ni, W., et al. (2020). SUMO1 modification of methyltransferase-like 3 promotes tumor progression via regulating Snail mRNA homeostasis in hepatocellular carcinoma. *Theranostics* 10, 5671–5686. doi: 10.7150/thno.42539
- Yamamoto, T., Seino, Y., Fukumoto, H., Koh, G., Yano, H., Inagaki, N., et al. (1990). Over-expression of facilitative glucose transporter genes in human cancer. *Biochem. Biophys. Res. Commun.* 170, 223–230. doi: 10.1016/0006-291x(90)91263-r
- Yang, J. D., and Roberts, L. R. (2010). Hepatocellular carcinoma: a global view. *Nat. Rev. Gastroenterol. Hepatol.* 7, 448–458. doi: 10.1038/nrgastro.2010.100
- Yang, S. L., Liu, L. P., Jiang, J. X., Xiong, Z. F., He, Q. J., and Wu, C. (2014). The correlation of expression levels of HIF-1 α and HIF-2 α in hepatocellular carcinoma with capsular invasion, portal vein tumor thrombi and patients' clinical outcome. *Jpn J. Clin. Oncol.* 44, 159–167. doi: 10.1093/jjco/hyt194
- Yang, Y., Xin, X., Fu, X., and Xu, D. (2018). Expression pattern of human SERPINE2 in a variety of human tumors. *Oncol. Lett.* 15, breakpgrng4523–4530.
- Yoshikawa, H., Matsubara, K., Qian, G.-S., Jackson, P. E., Groopman, J., Manning, J., et al. (2001). SOCS-1, a negative regulator of the JAK/STAT pathway, is silenced by methylation in human hepatocellular carcinoma and shows growth-suppression activity. *Nat. Genet.* 28, 29–35. doi: 10.1038/ng0501-29
- Yu, L.-X., Ling, Y., and Wang, H.-Y. (2018). Role of nonresolving inflammation in hepatocellular carcinoma development and progression. *NPJ Prec. Oncol.* 2:6.
- Yumoto, E., Nakatsukasa, H., Hanafusa, T., Yumoto, Y., Nouse, K., Matsumoto, E., et al. (2005). IGFBP-3 expression in hepatocellular carcinoma involves abnormalities in TGF-beta and/or Rb signaling pathways. *Int. J. Oncol.* 27, 1223–1230.
- Zaccara, S., Ries, R. J., and Jaffrey, S. R. (2019). Reading, writing and erasing mRNA methylation. *Nat. Rev. Mol. Cell Biol.* 20, 608–624. doi: 10.1038/s41580-019-0168-5
- Zappala, G., Elbi, C., Edwards, J., Gorenstein, J., Rechler, M. M., and Bhattacharyya, N. (2008). Induction of apoptosis in human prostate cancer cells by insulin-like growth factor binding protein-3 does not require binding to retinoid X receptor-alpha. *Endocrinology* 149, 1802–1812. doi: 10.1210/en.2007-1315
- Zarnack, K., König, J., Tajnik, M., Martincorena, I., Eustermann, S., Stévant, I., et al. (2013). Direct competition between hnRNP C and U2AF65 protects the transcriptome from the exonization of Alu elements. *Cell* 152, 453–466. doi: 10.1016/j.cell.2012.12.023
- Zhai, J., Song, J., Zhang, T., Xie, S., and Ma, C. (2020). deepEA: a containerized web server for interactive analysis of epitranscriptome sequencing data. *Plant Physiol.* 185, 29–33.
- Zhang, C., Fu, J., and Zhou, Y. (2019). A review in research progress concerning m6A methylation and immunoregulation. *Front. Immunol.* 10:922. doi: 10.3389/fimmu.2019.00922
- Zhang, Y., and Hamada, M. (2020). MoAIMS: efficient software for detection of enriched regions of MeRIP-Seq. *BMC Bioinformatics.* 21:103. doi: 10.1186/s12859-020-3430-0
- Zhao, C., Li, Y., Zhang, M., Yang, Y., and Chang, L. (2015). miR-126 inhibits cell proliferation and induces cell apoptosis of hepatocellular carcinoma cells partially by targeting Sox2. *Hum. Cell* 28, 91–99. doi: 10.1007/s13577-014-0105-z
- Zhao, X., Chen, Y., Mao, Q., Jiang, X., Jiang, W., Chen, J., et al. (2018). Overexpression of YTHDF1 is associated with poor prognosis in patients with hepatocellular carcinoma. *Cancer Biomarkers* 21, 1–10.
- Zhao, Y., Shi, Y., Shen, H., and Xie, W. (2020). m6A-binding proteins: the emerging crucial performers in epigenetics. *J. Hematol. Oncol.* 13:35.
- Zheng, G., Dahl, J. A., Niu, Y., Fedorcsak, P., Huang, C. M., Li, C. J., et al. (2013). ALKBH5 is a mammalian RNA demethylase that impacts RNA metabolism and mouse fertility. *Mol. Cell* 49, 18–29. doi: 10.1016/j.molcel.2012.10.015
- Zheng, R., and Blobel, G. A. (2010). GATA transcription factors and cancer. *Genes Cancer* 1, 1178–1188. doi: 10.1177/1947601911404223
- Zhong, L., Liao, D., Zhang, M., Zeng, C., Li, X., Zhang, R., et al. (2019). YTHDF2 suppresses cell proliferation and growth via destabilizing the EGFR mRNA in hepatocellular carcinoma. *Cancer Lett.* 442, 252–261. doi: 10.1016/j.canlet.2018.11.006

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

Copyright © 2021 Sivasudhan, Blake, Lu, Meng and Rong. 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.



Prognostic Significance and Tumor Immune Microenvironment Heterogeneity of m5C RNA Methylation Regulators in Triple-Negative Breast Cancer

OPEN ACCESS

Zhidong Huang^{1†}, Junfan Pan^{2†}, Helin Wang^{1†}, Xianqiang Du¹, Yusheng Xu¹, Zhitang Wang¹ and Debo Chen^{1*}

Edited by:

Jia Meng,
Xi'an Jiaotong-Liverpool University,
China

Reviewed by:

Xiaosong Chen,
Shanghai Jiao Tong University, China
Guo Wen Zhi,
Zhengzhou University, China

*Correspondence:

Debo Chen
debochensr@163.com

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

Specialty section:

This article was submitted to
Epigenomics and Epigenetics,
a section of the journal
Frontiers in Cell and Developmental
Biology

Received: 23 January 2021

Accepted: 25 March 2021

Published: 13 April 2021

Citation:

Huang Z, Pan J, Wang H, Du X,
Xu Y, Wang Z and Chen D (2021)
Prognostic Significance and Tumor
Immune Microenvironment
Heterogeneity of m5C RNA
Methylation Regulators
in Triple-Negative Breast Cancer.
Front. Cell Dev. Biol. 9:657547.
doi: 10.3389/fcell.2021.657547

¹ Quanzhou First Hospital of Fujian Medical University, Quanzhou, China, ² Shengli Clinical Medical College of Fujian Medical University, Fuzhou, China

Purpose: The m5C RNA methylation regulators are closely related to tumor proliferation, occurrence, and metastasis. This study aimed to investigate the gene expression, clinicopathological characteristics, and prognostic value of m5C regulators in triple-negative breast cancer (TNBC) and their correlation with the tumor immune microenvironment (TIM).

Methods: The TNBC data, Luminal BC data and HER2 positive BC data set were obtained from The Cancer Genome Atlas and Gene Expression Omnibus, and 11 m5C RNA methylation regulators were analyzed. Univariate Cox regression and the least absolute shrinkage and selection operator regression models were used to develop a prognostic risk signature. The UALCAN and cBioportal databases were used to analyze the gene characteristics and gene alteration frequency of prognosis-related m5C RNA methylation regulators. Gene set enrichment analysis was used to analyze cellular pathways enriched by prognostic factors. The Tumor Immune Single Cell Hub (TISCH) and Timer online databases were used to explore the relationship between prognosis-related genes and the TIM.

Results: Most of the 11 m5C RNA methylation regulators were differentially expressed in TNBC and normal samples. The prognostic risk signature showed good reliability and an independent prognostic value. Prognosis-related gene mutations were mainly amplified. Concurrently, the NOP2/Sun domain family member 2 (*NSUN2*) upregulation was closely related to spliceosome, RNA degradation, cell cycle signaling pathways, and RNA polymerase. Meanwhile, *NSUN6* downregulation was related to extracellular matrix receptor interaction, metabolism, and cell adhesion. Analysis of the TISCH and Timer

databases showed that prognosis-related genes affected the TIM, and the subtypes of immune-infiltrating cells differed between *NSUN2* and *NSUN6*.

Conclusion: Regulatory factors of m5C RNA methylation can predict the clinical prognostic risk of TNBC patients and affect tumor development and the TIM. Thus, they have the potential to be a novel prognostic marker of TNBC, providing clues for understanding the RNA epigenetic modification of TNBC.

Keywords: triple-negative breast cancer, m5C RNA methylation regulator, *NSUN2*, *NSUN6* prognostic risk signature, tumor immune microenvironment

INTRODUCTION

Despite advances in treatment, breast cancer remains a leading cause of cancer-related mortality among women worldwide, accounting for 6.6% of all deaths in 2018 (Bray et al., 2018). Triple-negative breast cancer (TNBC) is often more aggressive than other breast cancer subtypes. Unlike the estrogen receptor (ER)-positive and human epidermal growth factor receptor 2 (HER2) positive subtype, the biology of TNBC includes a high proliferation activity, a high degree of immune infiltration, basal-like or mesenchymal phenotypes, and insufficient homologous recombination (Denkert et al., 2017); its characteristics, including the risk factors, molecular and pathological characteristics, disease course, and sensitivity to chemotherapy, are distinct from those of other breast cancer subtypes (Borri and Granaglia, 2020). Although the basic diagnosis and treatment principles of TNBC are similar to those of other breast cancers, TNBC usually has a poor prognosis due to the lack of expression of targeted hormone receptors and HER2 and an insufficient range of treatment options (Borri and Granaglia, 2020; Cai et al., 2020). Therefore, a comprehensive understanding of the molecular mechanism of TNBC would be beneficial for exploring more effective treatment methods (Ding et al., 2020).

Recent studies have presented RNA modification as an emerging mechanism in gene regulation. This reversible post-transcriptional modification is regulated by “writers” (methyltransferases), “readers,” and “erasers” (demethylase) and affects various molecular functions, such as RNA-protein interaction (Liu et al., 2015), RNA stability (Wang et al., 2014; Yang X. et al., 2017), and translation efficiency (Wang et al., 2015; Schumann et al., 2020). Dysregulation of RNA modifications has been linked to several diseases including cancers, such as leukemia (Paris et al., 2019), breast cancer (Marcel et al., 2013), and prostate cancer (Chen et al., 2017).

5-Methylcytosine (m5C), which exists in mRNAs and ncRNAs, is a common RNA modification in humans (Han et al., 2020). Overall, 95,391 m5C sites in the human genome have been reported to date (Tang et al., 2021), identified with different sequencing methods, including bisulfite sequencing, Aza-IP, and miCLIP-Seq. The writers and readers for m5C modification have been well-studied in previous years. The NOP2/Sun domain family member 2 (*NSUN2*) is considered an important methyltransferase of m5C modification in tRNA, abundant non-coding RNAs, and a small number of mRNAs (Hussain et al., 2013; Blanco et al., 2014; Tang et al., 2021). A recent study showed

that *NSUN6* mediates site-specific deposition of m5C in mRNA to control the translation quality (Selmi et al., 2021, 6). Although the demethylase of m5C is still unknown, the conversion from m5C modification to other types of modifications (for example, hm5C) has been reported (Chen L. et al., 2019). Abnormal m5C has been recently found to play carcinogenic roles in several cancers. For example, *NSUN2* and YBX1 drive the onset of human urothelial carcinoma of the bladder (UCB) by targeting the m5C methylation site in the 3' untranslated region of hepatoma-derived growth factor (HDGF) (Chen X. et al., 2019). In gastric cancer, *NSUN2* acts as an oncogene to repress p57^{Kip2} in an m5C-dependent manner, which in turn, contributes to the development of cancer (Mei et al., 2020). However, the role of m5C in breast cancer, especially in TNBC with the worst prognosis, is still unclear.

The purpose of this study was to explore the differentially expressed genes in m5C regulators in different subtypes of BC, mainly in TNBC; identify prognostic genes; and initially explore the correlation of these genes with the clinicopathological characteristics, cell signaling pathways, and tumor immune microenvironment (TIM). Toward this goal, we analyzed data on the 11 currently reported m5C regulators in different subtypes of BC, mainly in TNBC and adjacent normal tissues from The Cancer Genome Atlas (TCGA) and Gene Expression Omnibus (GEO) databases.

MATERIALS AND METHODS

Data Source

RNA-seq transcriptome data of 99 TNBC samples, 779 Luminal BC samples, and 152 HER2 positive BC samples and the corresponding clinical information of TNBC were downloaded from the TCGA data portal: National Cancer Institute Genome Data Sharing Website¹. RNA-seq data of 212 research samples, including 99 TNBC samples and 113 normal breast tissues, and the corresponding clinical information (Table 1) served as the training cohort. For validation, gene expression profiles from the GEO public dataset² were used. Three datasets, namely, GSE38959, GSE45827, and GSE65194, were selected, with a total of 147 samples. Of these, 112 were TNBC samples, and 35 were normal tissue samples. Besides, the independent GEO

¹<http://cancergenome.nih.gov/>

²<https://www.ncbi.nlm.nih.gov/geo/>

TABLE 1 | The clinical characteristics of triple negative breast cancer patients in the training cohort.

Variable	Patients	Percentage (%)
Age (years)		
≤50	37	36.6
>50	64	63.4
Unknown	0	0.0
Gender		
Male	0	0.0
Female	101	100.0
T stage		
T1	24	23.8
T2	63	62.4
T3	10	9.9
T4	4	4.0
N stage		
N0	63	62.4
N1	25	24.8
N2	11	8.1
N3	2	2.0
Unknown	0	0.0
M stage		
M0	85	84.2
M1	1	1.0
Unknown	15	14.9
Pathological stage		
I	17	16.8
II	63	62.4
III	17	16.8
IV	1	1.0
Unknown	3	3.0
Total	101	100.0

T, tumor; N, nodes; M, metastases

dataset (GSE58812) with 107 TNBC samples was used as a validation group to evaluate the specificity and sensitivity of the risk signature.

Identification of Differentially Expressed Genes in the TCGA Database

A total of 11 m5C RNA methylation regulators, including *NSUN2*, *NSUN3*, *NSUN4*, *NSUN5*, *NSUN6*, *NSUN7*, *DNMT2*, *DNMT3A*, *DNMT3B*, *TET2*, and *ALYREF*, were obtained from published literature (Table 2). The factor expression matrix was used as the clinical information of different subtypes of BC samples and normal breast samples in the TCGA database. Then, the R version (4.0.2) of the limma software package³ was used to identify the differentially expressed m5C RNA methylation regulators between the tumor group and the normal tissue group. Genes with adjusted *P*-values of < 0.05 and $|\log_2(\text{FC})| > 1.0$ were considered as differentially expressed genes (DEGs). Subsequently, heat maps and violin maps were used to show the differential expression of m5C RNA methylation regulators between the two groups.

³<http://www.bioconductor.org/>

TABLE 2 | The list of the RNA modifying proteins involved in m5C.

Regulators	Type
m5C	
NSUN2	"writers"
NSUN3	"writers"
NSUN4	"writers"
NSUN5	"writers"
NSUN6	"writers"
NSUN7	"writers"
DNMT1	"writers"
DNMT3A	"writers"
DNMT3B	"writers"
ALYREF	"readers"
TET2	"erasers"

GEO Database Verification of the Differentially Expressed Genes

First, we integrated all the samples of the three data sets, significantly increasing the number of samples to 147 samples (112 TNBC samples and 35 normal controls). Then, we used the SVA package in the R computing environment to process the batch effect to avoid generating unreliable results. Next, we used the limma package in the R computing environment to analyze the difference in gene expression between tumor tissue and normal tissue (adjusted *P*-value < 0.05 and $|\log_2(\text{FC})| > 1.0$). Subsequently, heat maps and violin maps were generated to show the differential expression of m5C RNA methylation regulators between the two groups.

Protein-Protein Interaction Network Construction and Correlation Analysis

The Search Tool for the Retrieval of Interacting Genes (STRING) database⁴ was designed to analyze the protein-protein interaction (PPI) information. To evaluate the potential PPI relationship, the previously identified DEGs were mapped to the STRING database. The PPI pairs with a combined score of 0.4 were extracted. Subsequently, the PPI network was visualized using Cytoscape software⁵. It should be noted that nodes with a higher degree of connectivity tend to be more essential in maintaining the stability of the entire network. The R software was used to calculate the degree of each protein node.

Construction of the Prognostic Risk Scoring Model

To clarify the relationship between the expression of m5C RNA methylation regulators and overall survival (OS), we performed univariate Cox regression analysis. In the LASSO Cox regression algorithm, the optimal penalty parameter lambda and the corresponding coefficient criteria were determined based on the minimum criteria, through 10-fold cross-validation. Thereafter,

⁴<http://string-db.org/>

⁵www.cytoscape.org/

an ideal prognostic model based on prognosis-related genes was established. The risk score of the predictive model was calculated as follows:

$$\text{Risk score} = \sum_{i=1}^n \text{Coef}_i * x_i$$

where Coef_i is the coefficient, and x_i is the relative expression of each selected gene's z-score conversion. This formula was used to calculate the risk score of each patient in the TCGA database and the GEO dataset (GSE58812).

Database Analyses of Differentially Expressed Genes and Proteins in Normal and Breast Cancer Tissues

UALCAN Database

The UALCAN⁶ database, developed by the TCGA database and the Clinical Proteomic Tumor Analysis Consortium, was used to analyze the differences in the expression of prognosis-related m5C RNA methylation regulators between breast cancer tissues and normal tissues. The analysis was based on clinicopathological parameters, such as the molecular subtypes of breast cancer and tumor histological grade.

Human Protein Atlas

The immunohistochemical images in the Human Protein Atlas (HPA) website⁷ show the expressions of *NSUN2* and *NSUN6* proteins in human normal breast tissue and breast cancer tissue.

cBioPortal Database

The cBioPortal for Cancer Genomics database⁸ was used to analyze the frequency of gene alterations (including mutations, deletions, copy number gains, and amplifications) in prognostic m5C RNA methylation regulators in breast cancer. All searches were performed according to the online instructions of the cBioPortal.

Gene Set Enrichment Analysis

Gene Set Enrichment Analysis (GSEA) was performed in the TNBC cohort to further understand molecular mechanisms of the prognosis-related genes. Gene sets with both a false discovery rate q-val of < 0.25 and a normalized (NOM) p-value of < 0.05 were considered significant. When normalized enrichment score (NES) is a positive value, the gene is positively related to the pathway, and when NES is negative, it means that the gene is negatively related to the pathway.

Tumor Immune Single Cell Hub Database

The Tumor Immune Single Cell Hub (TISCH,⁹), an online database focusing on the tumor microenvironment (TME),

TABLE 3 | The tumor microenvironment including immune cells/inflammatory cells, stromal cells, and malignant cells.

Immune cells	Natural Killer Cells (NK)
	B Cells (B)
	CD4 T Cells (CD4+ T)
	CD8 T Cells (CD8+ T)
	Dendritic Cells (DC)
	Monocytes or Macrophages (Mono/Macro)
	Neutrophils
	Mast Cells (Mast)
	Regulatory T Cells (Treg)
	Proliferative T cells (Tprolif)
Stromal cells	Endothelial Cells (Endothelial)
	Fibroblasts (Fibroblasts)
Malignant cells	Malignant Cells (Malignant)

collected 76 tumor data sets from 27 cancers, including single-cell transcriptome profiles of nearly 2 million cells. In this study, the diverse cell types and cancer types included in the TISCH database were used to systematically investigate the TME (Table 3) heterogeneity.

TIMER Database

The TIMER database¹⁰ provides six main analysis modules, allowing users to interactively explore the relationship between immune infiltrates and various factors, including gene expression, clinical results, somatic mutations, and somatic copy number changes. In this study, the TIMER database was used to evaluate the correlation between DEGs and the level of immune cell infiltration. Data on six types of immune-infiltrating cells, namely B cells, CD4+ T cells, CD8+ T cells, neutrophils, macrophages, and dendritic cells, were collected from the database and analyzed.

Statistical Analysis

All analyses used R v4.0.2 software. The Wilcoxon test was used to compare the expression levels of 11 m5C regulators in different subtypes of BC samples and normal breast samples in the TCGA and GEO database. The Spearman test was used to identify correlations between m5C regulators. The median risk value was used as a cut-off value to divide patients into high-risk and low-risk groups. The Kaplan-Meier method was used to assess the correlation between high-risk and low-risk groups and overall survival. Univariate and multivariate Cox regression analysis identified whether risk score, age, stage, T, and N can be used as independent prognostic factors. *p*-values < 0.05 were considered statistically different.

⁶<http://ualcan.path.uab.edu>

⁷<https://www.proteinatlas.org>

⁸www.cbioportal.org

⁹<http://tisch.comp-genomics.org>

¹⁰<https://cistrome.shinyapps.io/timer/>

RESULTS

Expression of m5C RNA Methylation Regulators in Triple-Negative Breast Cancer

To study the relationship between m5C RNA methylation regulators and BC, we analyzed the expression of m5C RNA methylation regulators in TNBC samples, Luminal BC samples, and HER2 positive BC samples compared with normal tissue samples from the TCGA databases. The TCGA data showed that most of the m5C RNA methylation regulators are abnormally expressed in BC including TNBC (**Figures 1A,C** and **Supplementary Figure 1**). In TNBC tissues *NSUN2* ($p < 0.001$), Aly/REF export factor (*ALYREF*, $p < 0.001$), DNA-methyltransferase (*DNMT3B*) ($p < 0.001$), *DNMT1* ($p < 0.001$),

NSUN5 ($p < 0.001$), and *DNMT3A* ($p < 0.001$) were significantly overexpressed than in normal tissues. Meanwhile, *NSUN3* ($p < 0.005$), *NSUN4* ($p < 0.005$), and Tet methylcytosine dioxygenase 2 (*TET2*) ($p < 0.005$) showed significantly lower expression in TNBC tissues. Consistent results were obtained in the analysis of the GEO database.

Analysis of the three data sets of GSE38959, GSE45827, and GSE65194 showed that compared with those in normal tissues, the expressions of *DNMT1* ($p < 0.001$), *NSUN5* ($p < 0.001$), *NSUN6* ($p < 0.001$), and *NSUN7* ($p < 0.001$) were significantly higher in TNBC tissues. Meanwhile, the expressions of *NSUN3* ($p < 0.001$), *NSUN4* ($p < 0.001$), *NSUN2* ($p < 0.001$), *TET2* ($p < 0.001$), and *DNMT3B* ($p < 0.05$) were significantly lower in TNBC tissues (**Figures 1B,D**). Next, we tried to clarify the relationship among the 11 m5C RNA methylation regulators. Analysis of the String database (**Figure 1E**) showed that *NSUN6*

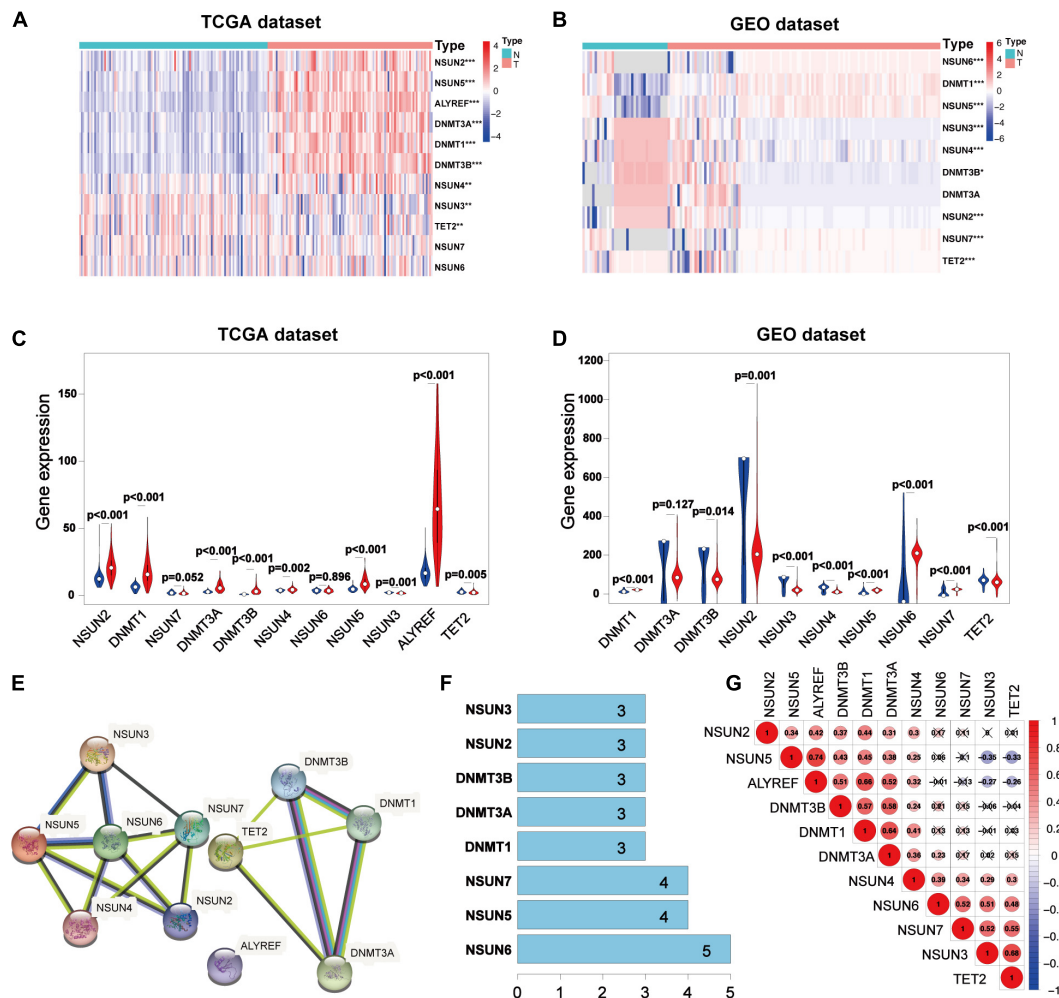


FIGURE 1 | Expression 3D models and correlation of m5C RNA methylation regulators in TNBC. Differential expression heatmap and violin plot of m5C RNA methylation regulators in TNBC and normal tissues from the Human Cancer Gene Atlas (**A,C**) and (**B,D**) Gene Expression Omnibus. The white dot in each “violin” represents the median expression. The PPI network of m5C RNA methylation regulators (**E**), number of interaction nodes between m5C RNA methylation regulators (**F**), and Spearman correlation analysis of m5C RNA methylation regulators (**G**). * $p < 0.05$, ** $p < 0.01$, and *** $p < 0.001$. X indicates $p > 0.05$. N, normal sample; T, tumor sample; blue violin: normal sample; red violin: tumor sample; TNBC, triple-negative breast cancer; PPI, protein-protein interaction.

may be a hub gene of the m5C RNA methylation regulator and interacts with the other five genes (Figure 1F). However, further analysis did not show a strong correlation between *NSUN6* expression and the expression of other m5C RNA methylation regulators. Interestingly, *NSUN5* were positively related with four other genes, and its expression had a strong positive correlation with that of *ALYREF* (Figure 1G). *NSUN5* is highly expressed in TNBC, indicating that this gene may be a key gene in the m5C RNA methylation regulator that affects tumorigenesis and development.

Verification of the Validity of the Risk Signature for Predicting Prognosis

To better understand the prognostic value of m5C RNA methylation regulators in TNBC, we used univariate Cox regression to analyze survival according to the expression of the associated genes in TNBC samples from the TCGA database. The results showed that the two most significant genes influencing OS were *NSUN6* ($p < 0.05$, Figure 2A) and *NSUN2* ($p < 0.2$), with *NSUN6* being a protective factor (hazard ratio, $HR < 1$) and *NSUN2* being an adverse factor ($HR > 1$). Then, with expression profiles of eleven m5C RNA methylation regulators, we conducted the LASSO Cox regression algorithm that identified *NSUN2* and *NSUN6*, and constructed a prognostic model based on these two regulators (Figures 2B,C). Lastly, we determined the coefficients used to calculate the risk score after 10-fold cross-validation. The coefficients of *NSUN6* and *NSUN2* were -0.5714 and 0.024 , respectively (Table 4). The risk score of each TNBC patient was calculated using the following formula: $\text{risk score} = -0.5714 \times \text{NSUN6} + 0.024 \times \text{NSUN2}$. All TNBC patients were then divided into the low-risk and high-risk groups according to the median risk score. As shown in Figure 2D, patients in the high-risk group had significantly lower OS than those in the low-risk group ($p < 0.001$) (Figure 2D). To evaluate the specificity and sensitivity of the risk signature for predicting TNBC prognosis, we performed a time-dependent receiver operating characteristic curve (ROC) analysis in the TCGA database and the GEO dataset (GSE58812), and the area under the ROC curve (AUC) of the 5-year analysis was 0.917 , 0.617 , respectively (Figures 2E,F). The above results indicate that our risk signature can predict the prognosis of TNBC.

Association Between the Clinicopathological Characteristics and Prognostic Risk Score

To further verify the prognostic value of the risk signature, we also explored the correlation between the clinicopathological characteristics of TNBC patients and the risk signature. When the data is divided into high-risk and low-risk groups to draw the heatmap, it is regrettable that no significant difference was seen in clinicopathological characteristics ($p > 0.05$) (Figure 3A). Then we performed the univariate and multivariate Cox regression analysis, showed that riskscore had a significant correlation with OS ($p = 0.002$, 0.028 , respectively). Meanwhile, univariate Cox regression analysis also showed that TNBC tumor size (T), axillary lymph node metastasis (N), and histological stage

were also related to OS ($p < 0.005$) (Figures 3B,C). However, Multivariate Cox regression analysis showed that there was no significant correlation between T, N, M, age, and OS ($p > 0.05$). These results suggest that the risk model established based on the two m5C RNA methylation regulators can be used as an independent prognostic factor for TNBC patients.

Relationship Between Expression of Prognosis-Related Genes and Clinicopathological Features and Their Protein Expression

To determine the mechanism by which *NSUN2* and *NSUN6* affected the prognosis of TNBC patients, we used the UALCAN online database to analyze 1,097 primary breast cancer samples and 114 normal breast samples in TCGA according to molecular subtypes and histological classification. The results showed higher *NSUN2* expression in breast cancer tissues than in normal tissues ($p = 6.38e-09$). Interestingly, in the analysis by molecular subtype, *NSUN2* expression was significantly higher in TNBC samples than in both normal samples and luminal breast cancer samples ($p = 4.12e-10$, $1.11e-6$, respectively). Meanwhile, there was no significant difference in *NSUN2* expression between TNBC and Her2+ breast cancer ($p > 0.05$). Also, *NSUN2* expression increased in a histology-dependent manner, with a significant difference from level 1 to 3 ($p < 0.05$) (Figure 4A).

For *NSUN6*, there was no significant difference in its expression between normal samples and all primary breast cancers ($p > 0.05$), and no significant in molecular subgroup of breast cancers. In the concurrent analysis by histological classification, *NSUN6* expression was found to be significantly lower in histological grades 3 and 4 TNBC than in normal tissues ($p < 0.05$) (Figure 4B). These findings further confirm the validity of our risk signature, with *NSUN2* as a risk factor and *NSUN6* as a protective factor.

The HPA database was then used to further study the protein expression of *NSUN2* and *NSUN6* at the translation level. The immunohistochemical image showed moderate *NSUN2* expression in normal breast glands and high expression in breast cancer tissues. Meanwhile, *NSUN6* showed medium expression in normal breast glands and high expression in breast cancer tissues (Figure 4C).

Gene Alterations and Signaling Pathways of Prognosis-Related Genes in Breast Cancer

Recent studies suggested that gene alterations may affect RNA modifications and disease associations (Song et al., 2020; Chen et al., 2021). Here, we analyzed the alterations in *NSUN2* and *NSUN6* from the cBioportal database, with these alterations mainly including mutations, deletions, copy number gains, and amplifications. The mutation frequencies of *NSUN2* and *NSUN6* were 9% and 6%, respectively, and both were mainly amplified (Figure 5A). We performed GSEA to identify the abnormally activated signaling pathways of *NSUN2* and *NSUN6* that cause their differential expression in breast cancer. Single-gene analysis showed that high *NSUN2* expression was associated

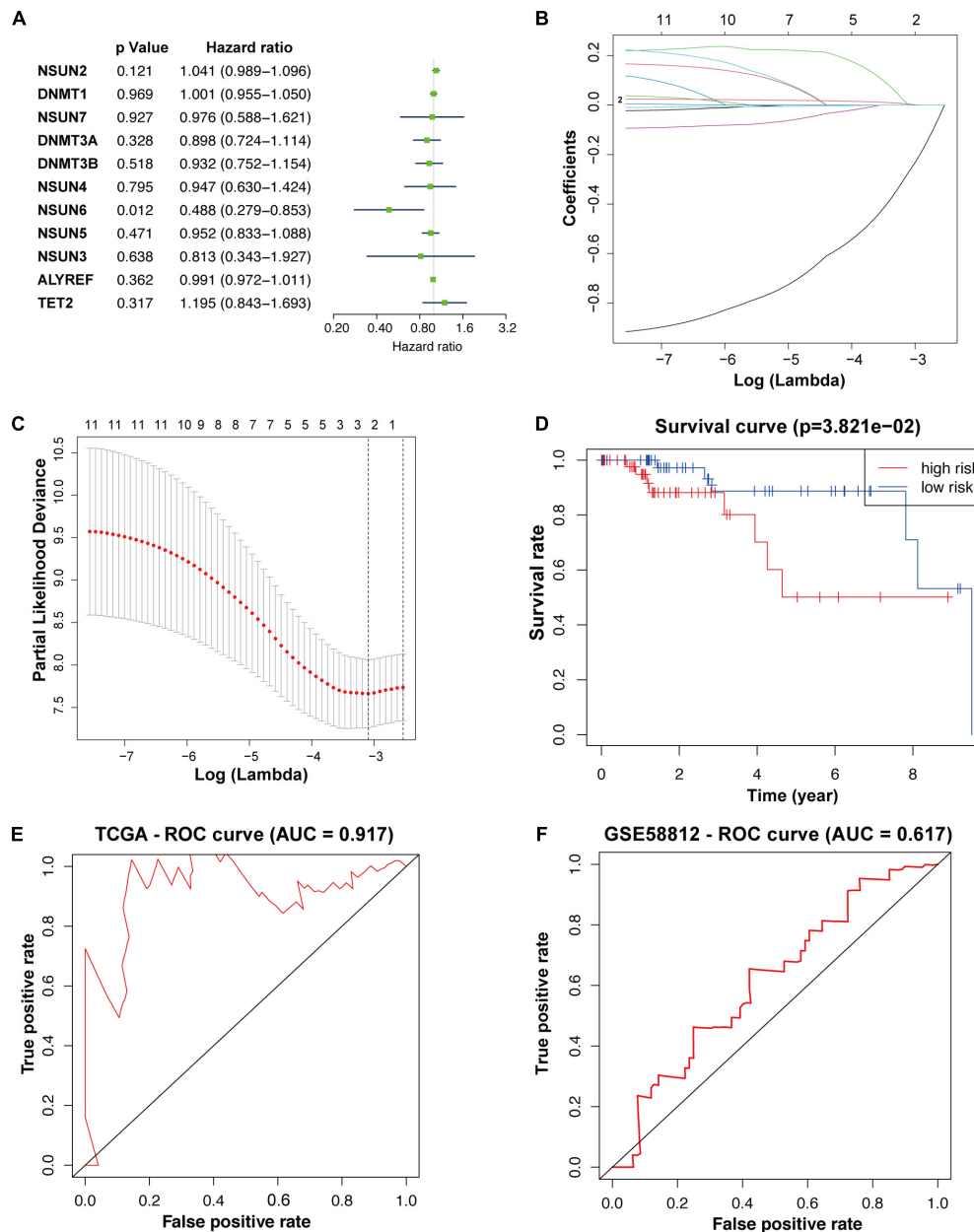


FIGURE 2 | Construction of prognostic risk signature for TNBC patients with two survival-related genes. The p -value, HR value, and 95% confidence interval of the 11 m5C RNA methylators analyzed using univariate Cox regression analysis (A). Using LASSO Cox regression, two m5C RNA methylators were selected for risk coefficient calculation (B,C). Kaplan-Meier overall survival curves based on the risk signature (D). ROC curves verified the specificity and sensitivity of the risk signature prediction in the TCGA database and the GEO dataset (GSE58812) (E,F). TNBC, triple-negative breast cancer; ROC, operating characteristic curve.

with spliceosome (NES = 2.22, $p < 0.001$), RNA degradation (NES = 2.22 $p < 0.001$), cell cycle signaling pathway (NES = 2.12, $p < 0.001$), RNA polymerase (NES = 2.11, $p < 0.001$), and DNA replication (NES = 2.06, $p < 0.001$). Meanwhile, low NSUN6 expression was associated with extracellular matrix receptor interaction (NES = -1.98, $p < 0.005$), fructose and mannose metabolism (NES = -1.85, $p < 0.005$), amino sugar and nucleotide sugar metabolism (NES = -1.77, $p < 0.01$),

complement and coagulation cascades (NES = -1.65, $p < 0.05$), and focal adhesion (NES = -1.64, $p < 0.05$) (Figure 5B).

Correlation Between the Prognosis-Related Genes and the Tumor Microenvironment

Considering the role of TME in tumor occurrence and development and its prognostic impact, we used three

TABLE 4 | Genes selected to build risk signature and the corresponding coefficients.

Genes	Coefficients
NSUN2	0.0248511231827596
NSUN6	-0.571442168838719

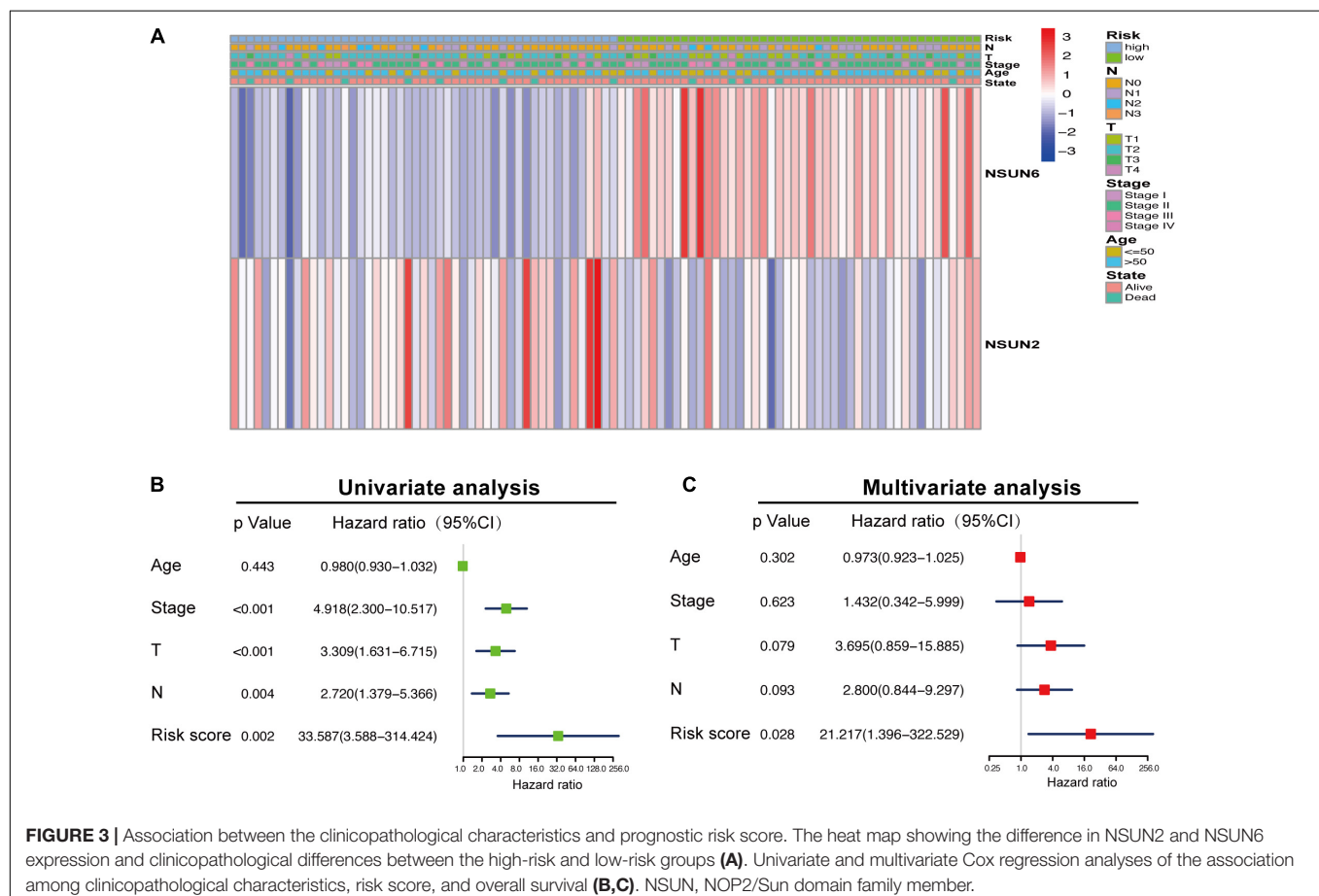
data sets (i.e., GSE114727-inDrop, BRCA_GSE110686, and BRCA_GSE114727_10X) in the TISCH database to analyze the expression of *NSUN2* and *NSUN6* in TME-related cells. We found low to moderate *NSUN2* and *NSUN6* expression in immune cells including B cells, CD4+ T cells, CD8+ T cells, neutrophils, macrophages, dendritic cells, and Tregs (Figures 6A,B). *NSUN2* expression was the highest in monocytes/macrophages, followed by that in proliferative T cells. Meanwhile, *NSUN6* expression was the highest in Tregs. As shown in the figure, *NSUN2* expression is higher than *NSUN6* expression in immune cells.

We then analyzed the GSE114727-inDrop dataset, which is divided into 12 types of cells. Figure 6C shows the number of cells in each cell type, with the distribution and number of various TME-related cells presented (Figure 6D). In this data set, CD4+ T cells were the most abundant immune cells ($n = 5413$). As shown in Figure 6E, the degree of *NSUN2* infiltration in

TME-related cells was higher than that of *NSUN6*, consistent with the results shown in Figure 6A. These results support that m5C regulators are closely related to the TIM in breast cancer.

Correlation Between Prognostic Genes and TIM

To determine whether our two prognosis-related genes can reflect the status of tumor immune infiltration, we further used the TIMER database to analyze the correlation of *NSUN2* and *NSUN6* with immune cell infiltration in TCGA. As shown in Figure 7A, there was no significant change in *NSUN2* expression ($p > 0.05$) as the degree of immune cell (e.g., CD8 + T cells, CD4 + T cells, B cells, neutrophils, macrophages, and dendritic cells) infiltration increased. However, *NSUN6* was positively correlated with the extent of T cell CD4+ infiltration ($p < 0.005$). Meanwhile, it did not correlate with the other five immune-infiltrating cells ($p > 0.05$) (Figure 7B). However, analysis of the correlation of these two genes with immune cell infiltration showed that under high *NSUN2* amplification, the neutrophil count was significantly lower in samples with somatic copy number alterations than that in normal samples ($p < 0.05$) (Figure 7C). Further, macrophage expression level of *NSUN6* in the high amplification state was higher than that in the normal group ($p < 0.05$) (Figure 7D).



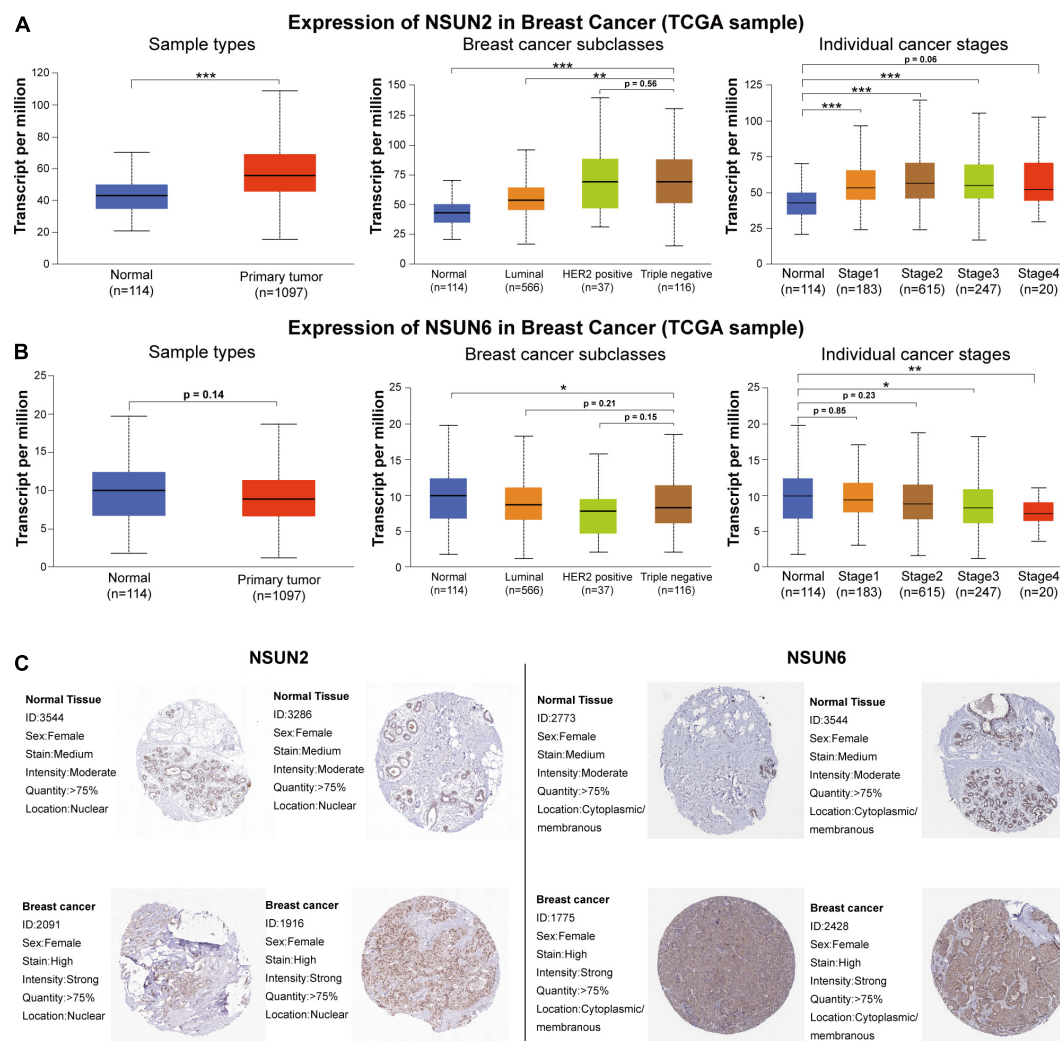


FIGURE 4 | Correlation between the differential expression of the two prognostic genes and clinicopathological patient characteristics and the difference in the expression of these genes at the protein level between normal tissues and breast cancer tissues. mRNA expression of NSUN2 and NSUN6 genes by molecular subtype and histological grade (A,B). The difference in protein expression of NSUN2 and NSUN6 in normal breast tissue and primary breast cancer tissue in the Human Protein Atlas (C). * $p < 0.05$, ** $p < 0.01$, and *** $p < 0.001$. NSUN, NOP2/Sun domain family member.

DISCUSSION

Although RNA modifications have already been reported to be associated with disease pathogenesis and cancer tumorigenesis (Tang et al., 2019; Tian et al., 2020; Zhao et al., 2020), the potential relationships between TNBC and m5C are still unclear. This study found that m5C modification-related regulatory factors were abnormally expressed in TNBC. Concurrently, we found two m5C-modified prognostic genes, and the risk signature constructed based on these two genes showed an independent prognostic value. In addition, these two genes were found to be correlated with the TIM. The value of m5C RNA methylation regulator for predicting the prognosis of TNBC patients and its influence on the TIM of TNBC can be helpful for the diagnosis and treatment of TNBC and ultimately improve patient prognosis. To the best of our knowledge, this

is the first study to report the role of m5C RNA methylation regulator in TNBC.

TNBC accounts for approximately 15–20% of all breast cancer cases (Chen et al., 2018), but the standard treatment modality for this subtype of breast cancer lacks effective treatments. Thus, many studies have attempted to develop specific treatment opportunities for TNBC patients in recent years (André and Zielinski, 2012; Garrido-Castro et al., 2019; Yin et al., 2020). An m6A modification has been recently found to play an important role in various human cancers, including breast cancer, by regulating the expression of oncogenes. Similarly, another internal RNA modification, that is, m5C modification, has also been shown to play an important role in cancer. However, related studies on TNBC and m5C modification are still lacking. The number of studies focused on investigating the value of the m5C RNA methylation regulator as a predictor of tumor prognosis is

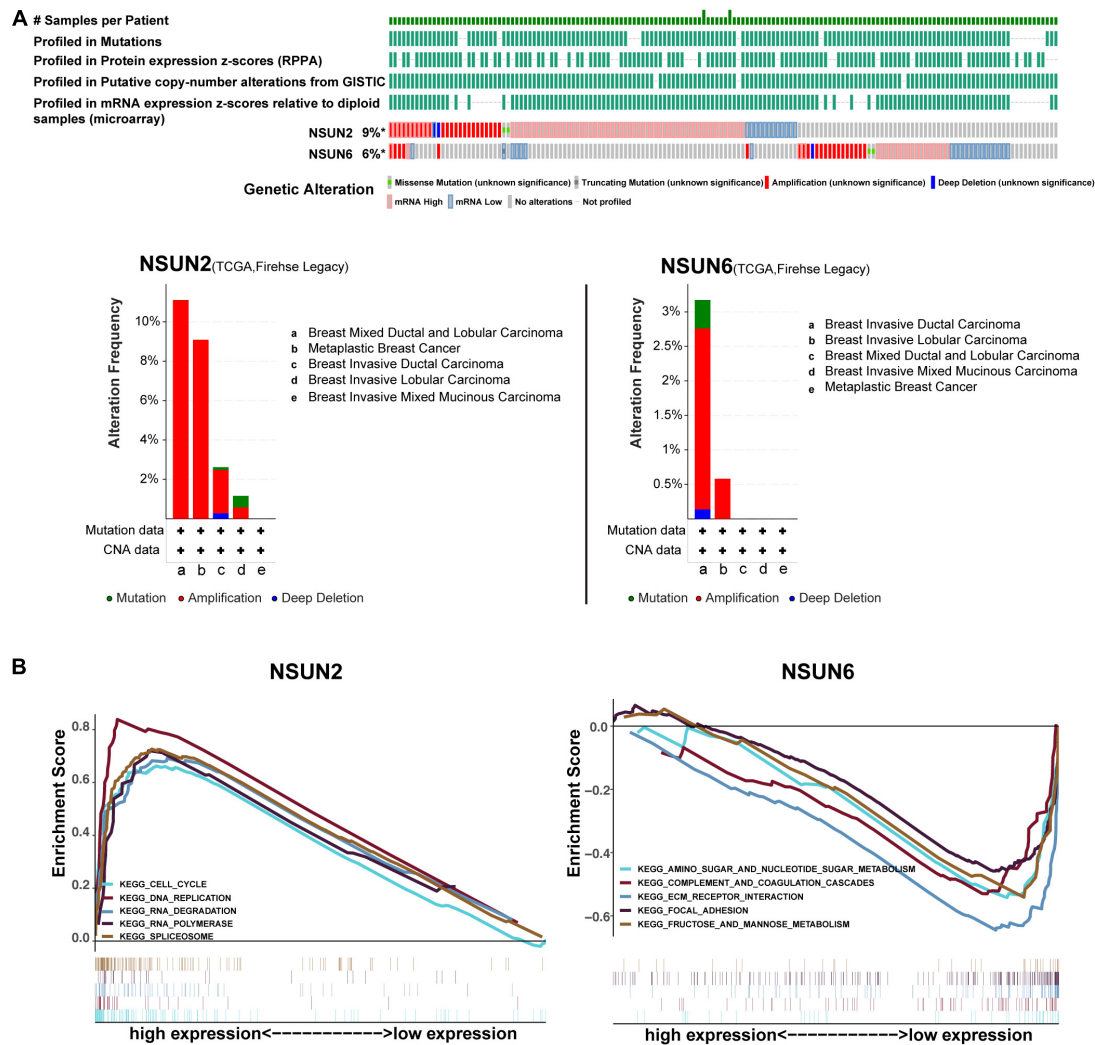


FIGURE 5 | Gene alterations analysis and gene set enrichment analysis of two prognosis-related genes. Analysis of the frequency and distribution of prognosis-related gene alteration in breast cancer based on cBioPortal data (A). GSEA of NSUN2 and NSUN6 in TNBC samples (B). The result is based on NES and NOM *p*-value. GSEA, Gene Set Enrichment Analysis; NSUN, NOP2/Sun domain family member; TNBC, triple-negative breast cancer; NES, normalized enrichment score; NOM, normalized.

increasing. Hypermethylation of m5C mRNA is highly enriched in cancer-related pathways, such as the PI3K-AKT35 and ERK-MAPK36 pathways (Sun et al., 2001; Samatar and Poulidakos, 2014). Also, the m5C transcription factor *NSUN2* and the m5C transcription factor *YBX1* were reported to be abnormally elevated in human UCB. The m5C site-dependent mechanism enhances the stability of HDGF mRNA, thereby promoting the pathogenesis of UCB (Chen X. et al., 2019). In this study, analysis of the expression matrix of 11 m5C RNA methylation regulators in TCGA and GEO databases revealed that they are abnormally expressed in TNBC tissues. This indicated that m5C RNA methylation regulators also play an important role in the pathogenesis of breast cancer.

Methylation regulators of m5C RNA have been verified to be closely related to cancer prognosis. This study found that the m5C RNA methylation regulators *NSUN2* and *NSUN6*

are correlated with the prognosis of TNBC. Concurrently, the prognostic risk signature developed based on these two genes has been proven to reliably predict patient prognosis. Patients identified to be at high risk according to the signature were found to have poor prognosis. These findings indicate that m5C RNA methylation regulatory factors also have a prognostic predictive value, and thus, have the potential to be novel prognostic indicators for TNBC patients.

Many studies have confirmed that *NSUN2* and *NSUN6* are related to cancer tumorigenesis. *NSUN2* is a nuclear RNA methyltransferase that catalyzes the formation of 5-methylcytosine. It has been shown to increase protein production through different mechanisms, including by promoting mRNA stability, affecting miRNA maturation and mRNA nuclear export, altering gene and lncRNA expression, and improving protein synthesis and translation. The highly expressed *NSUN2*

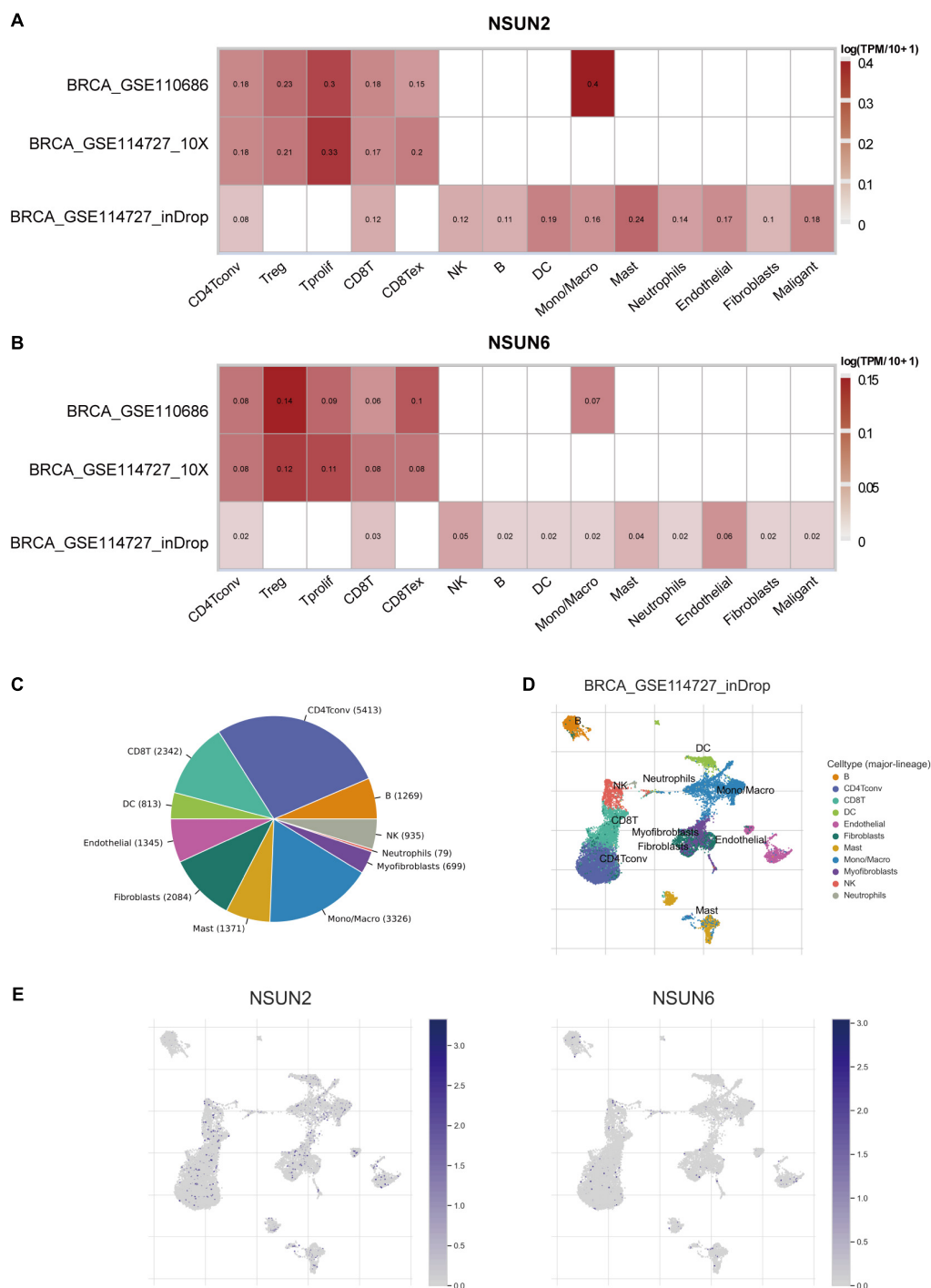


FIGURE 6 | Correlation between the prognostic-related genes and the TME. Correlation analysis between the expressions of NSUN2 and NSUN6 in primary breast cancer tissues and the TME, using the TISCH database (A,B). The cell types and their distribution in the GSE114727_inDrop dataset (C,D). The distribution of NSUN2 and NSUN6 in different cell types was analyzed using single-cell resolution in the GSE114727_inDrop dataset (E). NSUN, NOP2/Sun domain family member; TME; tumor microenvironment; TISCH, Tumor Immune Single Cell Hub.

closely interacts with *RPL6* to promote the proliferation and tumorigenesis of gallbladder cancer cells in both *in vitro* and *in vivo* (Gao et al., 2019). In breast cancer, *NSUN2* is a *MYC* target gene, which is closely related to cell growth and proliferation

(Frye and Watt, 2006; Frye et al., 2010; Yi et al., 2017). Besides *NSUN2* expression is negatively correlated with the status of ER and PR, and is correlated with poor prognosis of BC patients (Yi et al., 2017). *NSUN2* also has a potential prognostic value

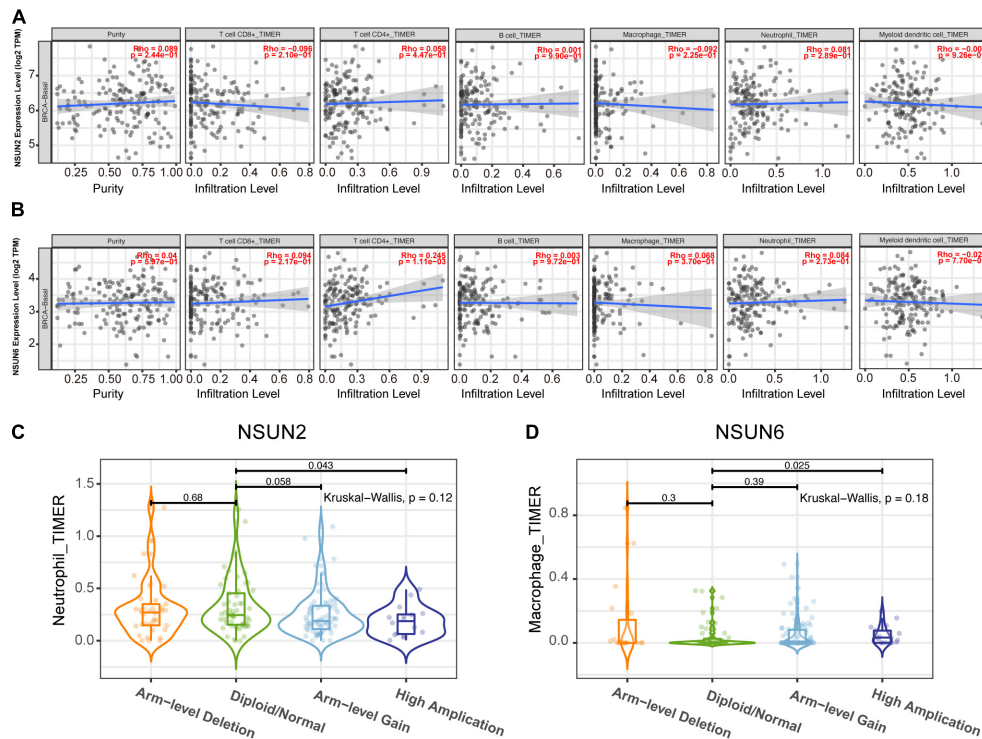


FIGURE 7 | Correlation between prognostic-related genes and TIM. Correlation analysis of NSUN2 and NSUN6 with the infiltration level of the 6 main immune cells after adjusting for the purity (A,B). Analysis according to different groups of somatic copy number alterations showed a significant difference in NSUN2 expression at the neutrophil level (C) and NSUN6 expression at the macrophage level (D) among these groups. TIM, tumor immune microenvironment; NSUN, NOP2/Sun domain family member.

in other cancers, such as lung cancer, ovarian cancer, and head and neck squamous cell carcinoma (Yang J. et al., 2017; Lu et al., 2020). Analysis of the association between clinical pathology and NSUN2 expression in breast cancer samples in the UALCAN database showed that NSUN2 was highly expressed in breast cancer. Further, NSUN2 expression was also higher in TNBC samples than in the luminal subtype samples.

NSUN6 is a methyltransferase that targets mRNA. A recent research has shown that Methylation regulated by NUSN6 mainly occurs in the 3'UTR near the translation termination site downstream of the stop codon, and mainly affects the RNA- and protein-binding factors that regulate mRNA processing and translation. In the testis, ovary and liver tissues, the high expression of NSUN6 is related to the high fidelity of RNA- and protein-binding factors, and a better survival rate of patients. Besides, in the Genotypic Tissue Expression (GTEx) database, the expression level of NSUN6 mRNA in normal breast tissue ranked 9th among 22 kinds of tissues (Selmi et al., 2021). In this study, NSUN6 mRNA expression level was higher in TNBC tissues than in normal tissues. Moreover, we unexpectedly found that NSUN6 is closely related to the stage of breast cancer. It is indicated that in TNBC tissues, the high expression of NSUN6 may also promote the high fidelity of RNA- and protein-binding factors that involve in mRNA processing and translation, and affected the survival of TNBC patients. In addition, analysis using the HPA database showed that NSUN2 and NSUN6 were also abnormally expressed

at the protein level in breast cancer tissues. This further confirms that NSUN2 and NSUN6 are related to the tumorigenesis of breast cancer, consistent with the findings of previous studies (Gao et al., 2019; Selmi et al., 2021).

There have been several reports that gene mutations usually cause phenotypic changes, which are in turn closely related to carcinogenesis and aging (Gonzalo et al., 2017; Kaur et al., 2018; Chen et al., 2021). Concurrently, a meta-analysis of studies on copy number amplification has shown that copy number changes affect genes involved in cell cycle regulation, retinoic acid signaling, complement system, and antigen presentation, which may cause cancer (Brown et al., 2016). The analysis of the cBioPortal database in this study showed that NSUN2 and NSUN6 have a high frequency of gene alterations in breast cancer patients. The main gene changes in NSUN2 and NSUN6 are copy number amplifications. Succeeding GSEA on NSUN2 and NSUN6 in this study showed that these two genes may be involved in important biological processes. NSUN2 upregulation is closely related to the spliceosome, RNA degradation, cell cycle signaling pathways, and RNA polymerase. BUD31 is a c-MYC synthetic lethal gene in human breast epithelial cells. BUD31 has been established to be a component of the core spliceosome required for MYC assembly and catalytic activity (Hsu et al., 2015). In other words, the components of the spliceosome are closely related to aggressive MYC-driven cancers and can be used as therapeutic targets.

In addition, nonsense-mediated RNA decay degrades abnormal RNA and part of normal RNA. RNA degradation can affect various events, including cancer tumorigenesis (Goetz and Wilkinson, 2017). Several studies have shown that cell cycle disorders are the primary driving factor for the immortal proliferation of cancer cells (Hydbring et al., 2016). The mechanisms include activation of cyclin-dependent kinases to drive the entry and development of the cell division cycle, repair of DNA damage, and control of cell death. Cyclin upregulation can lead to cell cycle disorders and uncontrolled cell growth (von Bergwelt-Baildon et al., 2011), indicating that cyclins play a vital role in the pathogenesis of cancer. Regarding RNA polymerase, related studies have shown that certain transforming agents can stimulate the expression of pol III-specific transcription factors TFIIB or TFIIC2. Meanwhile TFIIB is bound and activated by several oncogenic proteins (including c-Myc) and also plays a role in cancer development (White, 2004). Collectively, these findings indicate that the *NSUN2* gene may be involved in important biological processes in cancer tumorigenesis.

Meanwhile, *NSUN6* downregulation is associated with ECM receptor interaction, metabolism, and cell adhesion, and these pathways have been confirmed to be related to cancer. For example, Bao et al. (2019) reported that the ECM-receptor interaction signaling pathway may be related to the development of breast cancer. Research by Akella et al. (2019) showed that changes in metabolism and loss of cell energy are considered hallmarks of all cancers. Also, cell adhesion to the ECM has been recently identified to be a key determinant of drug resistance of cancer cells (Eke and Cordes, 2015). In summary, we speculate that *NSUN6* participates in these pathways in TNBC, ultimately affecting patient prognosis.

The TME comprises stromal cells, bioactive molecules secreted by tumors and stromal cells, the extracellular matrix (ECM), and the lymphatic and vascular systems, all of which play key roles in tumorigenesis and metastasis (Wu and Dai, 2017). TME has been recently widely associated with the prognosis of many cancers and lymph and distant metastasis, including in breast cancer (Zarrilli et al., 2020). A study has found that a variety of transcription factors in the microenvironment induce high expression of microRNA-10b (miR-10b) in metastatic breast cancer cells, promoting cancer cell migration, invasion, and metastasis (Ma et al., 2007). Also, transcription regulators (e.g., nuclear factor-kappa B, mitogen-activated protein kinases, and phosphoinositide-3 kinase/protein kinase-B) are found in thyroid cancer, and these promote the proliferation of oncogenes in the TME, which is in turn related to the thyroid cancer subtype (Ferrari et al., 2020). Immune infiltration in the TME is a current research hotspot. Studies have found that tumor-associated macrophages 1 and 2 and regulatory T cell (Treg) lymphocytes affect the prognosis of ER-positive breast cancer (Hammerl et al., 2018).

However, the impact of TIM on the TNBC progression remains unclear. Compared with other breast cancer subtypes, TNBC is more likely to carry tumor-infiltrating lymphocytes (TILs) (Molinero et al., 2019). Loi et al. (2014) found that for every 10% increase in TILs within the tumor, the risk of death was reduced by 27% in TNBC patients. Further, a high TIL count

was an important predictor of long-term recurrence, with every 10% increase reducing the relative risk of long-term recurrence by 13%. Besides, TNBC is rich in CD8+ T cells and is associated with a better prognosis. Compared with other breast cancer types, TNBC has up-regulated expression of regulatory T cells (Tregs) carrying the surface marker FOXP3+ on CD4+ T cells, which may inhibit the effects of other immune cells, thereby inhibiting effective immune responses (Loi et al., 2014). The above studies support that TILs are a powerful prognostic indicator in TNBC. In this study, we used the TISCH database to analyze the correlation between m5C RNA methylation regulators and the TME. The results showed that *NSUN2* and *NSUN6* are expressed to a certain extent in immune cells and *NSUN6* expression was higher than *NSUN2* expression. *NSUN2* is mainly expressed in monocytes/macrophages and proliferative T cells, whereas *NSUN6* is mainly expressed in Tregs. Additional TIMER database searches were used to show the relationship between two genes with TIM. *NSUN2* and *NSUN6* were correlated with the six major immune cells, with *NSUN6* having the strongest correlation with CD4+T cells.

Monocytes can differentiate into macrophages. Previous studies have shown that tumor-associated macrophages are important immunosuppressive cells that contribute to the growth and metastasis of breast cancer (Yang et al., 2013; Mou et al., 2015). Therefore, the poor prognosis of TNBC may, in part, be attributed to the abnormal expression of *NSUN2* and *NSUN6*, which is in turn, is associated with the upregulation of monocytes/macrophages and Tregs in the TIM.

RNA modification has become an increasingly important field in recent years. To better understand the RNA modifications, several bioinformatics studies were conducted to decipher the functions (Chen K. et al., 2019) and disease association of RNA modifications (Li et al., 2019). However, research on the effect of m5C on breast cancer is still lacking. To the best of our knowledge, this is the first study to conduct an in-depth analysis of the role of m5C RNA regulators in TNBC. However, there are still some limitations in this study. First, there are relatively few TNBC samples ($n = 99$) in the TCGA database, and this may lead to biases in subsequent studies on m5C RNA methylation regulators and clinicopathological characteristics. Second, only one TNBC patient had distant metastases, and thus, it was difficult to analyze the correlation between m5C RNA methylation regulators and M stage. Third, although a correlation analysis between differential genes and immune cancer cells can be performed in the TIMER database, the database cannot correlate clinicopathological characteristics with the degree of immune cell infiltration. Moreover, we were not able to use *in vivo* and *in vitro* experiments to further verify our findings. Fourth, the role of TNBC-related signaling pathways remains unclear. Fifth, we didn't explore the correlation between m5C regulators and prognosis of non-TNBC patients. Sixth, this study had a limited sample size, so more work needs to be done to strengthen and verify the stability of the risk model. Further molecular biology research and experiments, such as with real-time polymerase chain reaction and western blotting are needed. Future research should overcome these problems.

In conclusion, most m5C RNA methylation regulators are abnormally expressed in TNBC. m5C regulatory factors have the value of predicting prognosis. These regulatory factors are closely related to the tumorigenesis of TNBC and affect TIM. Therefore, m5C RNA methylation regulators have the potential to be prognostic markers for TNBC.

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

DC and ZH designed the study. ZH and JP collected and analyzed the data. HW wrote the manuscript. XD generated the figures and tables. YX checked and polished the language. ZW reviewed the manuscript. All authors contributed to this article and approved the submitted version.

REFERENCES

- Akella, N. M., Ciraku, L., and Reginato, M. J. (2019). Fueling the fire: emerging role of the hexosamine biosynthetic pathway in cancer. *BMC Biol.* 17:52. doi: 10.1186/s12915-019-0671-3
- André, F., and Zielinski, C. C. (2012). Optimal strategies for the treatment of metastatic triple-negative breast cancer with currently approved agents. *Ann. Oncol.* 23 (Suppl. 6), vi46–vi51.
- Bao, Y., Wang, L., Shi, L., Yun, F., Liu, X., Chen, Y., et al. (2019). Transcriptome profiling revealed multiple genes and ECM-receptor interaction pathways that may be associated with breast cancer. *Cell. Mol. Biol. Lett.* 24:38. doi: 10.1186/s11658-019-0162-0
- Blanco, S., Dietmann, S., Flores, J. V., Hussain, S., Kutter, C., Humphreys, P., et al. (2014). Aberrant methylation of tRNAs links cellular stress to neurodevelopmental disorders. *EMBO J.* 33, 2020–2039. doi: 10.15252/embj.201489282
- Borri, F., and Granaglia, A. (2020). Pathology of triple negative breast cancer. *Semin. Cancer Biol.* S1044–579X(20)30140-1. doi: 10.1016/j.semcancer.2020.06.005
- Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R. L., Torre, L. A., and Jemal, A. (2018). Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* 68, 394–424. doi: 10.3322/caac.21492
- Brown, T. C., Juhlin, C. C., Healy, J. M., Stenman, A., Rubinstein, J. C., Korah, R., et al. (2016). DNA copy amplification and overexpression of SLC12A7 in adrenocortical carcinoma. *Surgery* 159:250. doi: 10.1016/j.surg.2015.08.038
- Cai, L., Tong, Y., Zhu, X., Shen, K., Zhu, J., and Chen, X. (2020). Prolonged time to adjuvant chemotherapy initiation was associated with worse disease outcome in triple negative breast cancer patients. *Sci. Rep.* 10:7029. doi: 10.1038/s41598-020-64005-4
- Chen, J., Chen, Z., Huang, J., Chen, F., Ye, W., Ding, G., et al. (2018). Bioinformatics identification of dysregulated microRNAs in triple negative breast cancer based on microRNA expression profiling. *Oncol. Lett.* 15, 3017–3023.
- Chen, J., Sun, Y., Xu, X., Wang, D., He, J., Zhou, H., et al. (2017). YTH domain family 2 orchestrates epithelial-mesenchymal transition/proliferation dichotomy in pancreatic cancer cells. *Cell Cycle* 16, 2259–2271. doi: 10.1080/15384101.2017.1380125

FUNDING

This research was supported by the Natural Science Foundation of Fujian Province, PRC (No. 2019J01599) and Startup Fund for Scientific Research of Fujian Medical University (No. 2019QH1253).

ACKNOWLEDGMENTS

The authors thank the TCGA and GEO Network for providing high-quality data.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fcell.2021.657547/full#supplementary-material>

Supplementary Figure 1 | Differential expression heatmap and violin plot of m5C RNA methylation regulators in Luminal BC, and HER2 positive BC from the Human Cancer Gene Atlas. * $p < 0.05$, ** $p < 0.01$, and *** $p < 0.001$. Luminal BC, luminal breast cancer; HER2 positive BC, human epidermal growth factor receptor 2 positive breast cancer.

- Chen, K., Song, B., Tang, Y., Wei, Z., Xu, Q., Su, J., et al. (2021). RMDisease: a database of genetic variants that affect RNA modifications, with implications for epitranscriptome pathogenesis. *Nucleic Acids Res.* 49, D1396–D1404. doi: 10.1093/nar/gkaa790
- Chen, K., Wei, Z., Zhang, Q., Wu, X., Rong, R., Lu, Z., et al. (2019). WHISTLE: a high-accuracy map of the human N6-methyladenosine (m6A) epitranscriptome predicted using a machine learning approach. *Nucleic Acids Res.* 47:e41. doi: 10.1093/nar/gkz074
- Chen, L., Wang, P., Bahal, R., Manautou, J. E., and Zhong, X. (2019). Ontogenic mRNA expression of RNA modification writers, erasers, and readers in mouse liver. *PLoS One* 14:e0227102. doi: 10.1371/journal.pone.0227102
- Chen, X., Li, A., Sun, B.-F., Yang, Y., Han, Y.-N., Yuan, X., et al. (2019). 5-methylcytosine promotes pathogenesis of bladder cancer through stabilizing mRNAs. *Nat. Cell Biol.* 21, 978–990. doi: 10.1038/s41556-019-0361-y
- Denkert, C., Liedtke, C., Tutt, A., and von Minckwitz, G. (2017). Molecular alterations in triple-negative breast cancer—the road to new treatment strategies. *Lancet* 389, 2430–2442. doi: 10.1016/S0140-6736(16)32454-0
- Ding, S., Chen, X., and Shen, K. (2020). Single-cell RNA sequencing in breast cancer: understanding tumor heterogeneity and paving roads to individualized therapy. *Cancer Commun.* 40, 329–344. doi: 10.1002/cac2.12078
- Eke, I., and Cordes, N. (2015). Focal adhesion signaling and therapy resistance in cancer. *Semin. Cancer Biol.* 31, 65–75. doi: 10.1016/j.semcancer.2014.07.009
- Ferrari, S. M., Fallahi, P., Elia, G., Ragusa, F., Ruffilli, I., Paparo, S. R., et al. (2020). Thyroid autoimmune disorders and cancer. *Semin. Cancer Biol.* 64, 135–146. doi: 10.1016/j.semcancer.2019.05.019
- Frye, M., Dragoni, I., Chin, S.-F., Spiteri, I., Kurowski, A., Provenzano, E., et al. (2010). Genomic gain of 5p15 leads to over-expression of Misu (NSUN2) in breast cancer. *Cancer Lett.* 289, 71–80. doi: 10.1016/j.canlet.2009.08.004
- Frye, M., and Watt, F. M. (2006). The RNA methyltransferase Misu (NSUN2) mediates Myc-induced proliferation and is upregulated in tumors. *Curr. Biol.* 16, 971–981.
- Gao, Y., Wang, Z., Zhu, Y., Zhu, Q., Yang, Y., Jin, Y., et al. (2019). NOP2/Sun RNA methyltransferase 2 promotes tumor progression via its interacting partner RPL6 in gallbladder carcinoma. *Cancer Sci.* 110, 3510–3519. doi: 10.1111/cas.14190
- Garrido-Castro, A. C., Lin, N. U., and Polyak, K. (2019). Insights into molecular classifications of triple-negative breast cancer: improving patient selection for treatment. *Cancer Discov.* 9, 176–198. doi: 10.1158/2159-8290.CD-18-1177

- Goetz, A. E., and Wilkinson, M. (2017). Stress and the nonsense-mediated RNA decay pathway. *Cell. Mol. Life Sci.* 74, 3509–3531. doi: 10.1007/s00018-017-2537-6
- Gonzalo, S., Kreienkamp, R., and Askjaer, P. (2017). Hutchinson-gilford progeria syndrome: a premature aging disease caused by LMNA gene mutations. *Ageing Res. Rev.* 33, 18–29.
- Hammerl, D., Smid, M., Timmermans, A. M., Sleijfer, S., Martens, J. W. M., and Debets, R. (2018). Breast cancer genomics and immuno-oncological markers to guide immune therapies. *Semin. Cancer Biol.* 52, 178–188. doi: 10.1016/j.semcancer.2017.11.003
- Han, X., Wang, M., Zhao, Y.-L., Yang, Y., and Yang, Y.-G. (2020). RNA methylations in human cancers. *Semin. Cancer Biol.* S1044-579X(20)30241-8. doi: 10.1016/j.semcancer.2020.11.007
- Hsu, T. Y.-T., Simon, L. M., Neill, N. J., Marcotte, R., Sayad, A., Bland, C. S., et al. (2015). The spliceosome is a therapeutic vulnerability in MYC-driven cancer. *Nature* 525, 384–388. doi: 10.1038/nature14985
- Hussain, S., Sajini, A. A., Blanco, S., Dietmann, S., Lombard, P., Sugimoto, Y., et al. (2013). NSUN2-mediated cytosine-5 methylation of vault noncoding RNA determines its processing into regulatory small RNAs. *Cell Rep.* 4, 255–261. doi: 10.1016/j.celrep.2013.06.029
- Hydbring, P., Malumbres, M., and Sicinski, P. (2016). Non-canonical functions of cell cycle cyclins and cyclin-dependent kinases. *Nat. Rev. Mol. Cell Biol.* 17, 280–292. doi: 10.1038/nrm.2016.27
- Kaur, R. P., Vasudeva, K., Kumar, R., and Munshi, A. (2018). Role of p53 gene in breast cancer: focus on mutation spectrum and therapeutic strategies. *Curr. Pharm. Des.* 24, 3566–3575.
- Li, Y., Xiao, J., Bai, J., Tian, Y., Qu, Y., Chen, X., et al. (2019). Molecular characterization and clinical relevance of m6A regulators across 33 cancer types. *Mol. Cancer* 18:137. doi: 10.1186/s12943-019-1066-3
- Liu, N., Dai, Q., Zheng, G., He, C., Parisien, M., and Pan, T. (2015). N(6)-methyladenosine-dependent RNA structural switches regulate RNA-protein interactions. *Nature* 518, 560–564. doi: 10.1038/nature14234
- Loi, S., Michiels, S., Salgado, R., Sirtaine, N., Jose, V., Fumagalli, D., et al. (2014). Tumor infiltrating lymphocytes are prognostic in triple negative breast cancer and predictive for trastuzumab benefit in early breast cancer: results from the FinHER trial. *Ann. Oncol.* 25, 1544–1550. doi: 10.1093/annonc/mdu112
- Lu, L., Gaffney, S. G., Cannataro, V. L., and Townsend, J. (2020). Transfer RNA methyltransferase gene NSUN2 mRNA expression modifies the effect of T cell activation score on patient survival in head and neck squamous carcinoma. *Oral Oncol.* 101:104554. doi: 10.1016/j.oraloncology.2019.104554
- Ma, L., Teruya-Feldstein, J., and Weinberg, R. A. (2007). Tumour invasion and metastasis initiated by microRNA-10b in breast cancer. *Nature* 449, 682–688. doi: 10.1038/nature06174
- Marcel, V., Ghayad, S. E., Belin, S., Therizols, G., Morel, A.-P., Solano-González, E., et al. (2013). p53 acts as a safeguard of translational control by regulating fibrillar and rRNA methylation in cancer. *Cancer Cell* 24, 318–330. doi: 10.1016/j.ccr.2013.08.013
- Mei, L., Shen, C., Miao, R., Wang, J.-Z., Cao, M.-D., Zhang, Y.-S., et al. (2020). RNA methyltransferase NSUN2 promotes gastric cancer cell proliferation by repressing p57Kip2 by an m5C-dependent manner. *Cell Death Dis.* 11:270. doi: 10.1038/s41419-020-2487-z
- Molinero, L., Li, Y., Chang, C.-W., Maund, S., Berg, M., Harrison, J., et al. (2019). Tumor immune microenvironment and genomic evolution in a patient with metastatic triple negative breast cancer and a complete response to atezolizumab. *J. Immunother. Cancer* 7:274. doi: 10.1186/s40425-019-0740-8
- Mou, W., Xu, Y., Ye, Y., Chen, S., Li, X., Gong, K., et al. (2015). Expression of Sox2 in breast cancer cells promotes the recruitment of M2 macrophages to tumor microenvironment. *Cancer Lett.* 358, 115–123. doi: 10.1016/j.canlet.2014.11.004
- Paris, J., Morgan, M., Campos, J., Spencer, G. J., Shmakova, A., Ivanova, I., et al. (2019). Targeting the RNA m6A Reader YTHDF2 selectively compromises cancer stem cells in acute myeloid leukemia. *Cell Stem Cell* 25, 137–148.e6. doi: 10.1016/j.stem.2019.03.021
- Samatar, A. A., and Poulikakos, P. I. (2014). Targeting RAS-ERK signalling in cancer: promises and challenges. *Nat. Rev. Drug Discov.* 13, 928–942. doi: 10.1038/nrd4281
- Schumann, U., Zhang, H.-N., Sibbritt, T., Pan, A., Horvath, A., Gross, S., et al. (2020). Multiple links between 5-methylcytosine content of mRNA and translation. *BMC Biol.* 18:40. doi: 10.1186/s12915-020-00769-5
- Selmi, T., Hussain, S., Dietmann, S., Heiß, M., Borland, K., Flad, S., et al. (2021). Sequence- and structure-specific cytosine-5 mRNA methylation by NSUN6. *Nucleic Acids Res.* 49, 1006–1022. doi: 10.1093/nar/gkaa1193
- Song, B., Tang, Y., Chen, K., Wei, Z., Rong, R., Lu, Z., et al. (2020). m7GHub: deciphering the location, regulation and pathogenesis of internal mRNA N7-methylguanosine (m7G) sites in human. *Bioinformatics* 36, 3528–3536. doi: 10.1093/bioinformatics/btaa178
- Sun, M., Paciga, J. E., Feldman, R. I., Yuan, Z., Coppola, D., Lu, Y. Y., et al. (2001). Phosphatidylinositol-3-OH Kinase (PI3K)/AKT2, activated in breast cancer, regulates and is induced by estrogen receptor alpha (ERalpha) via interaction between ERalpha and PI3K. *Cancer Res.* 61, 5985–5991.
- Tang, Y., Chen, K., Song, B., Ma, J., Wu, X., Xu, Q., et al. (2021). m6A-Atlas: a comprehensive knowledgebase for unraveling the N6-methyladenosine (m6A) epitranscriptome. *Nucleic Acids Res.* 49, D134–D143. doi: 10.1093/nar/gkaa692
- Tang, Y., Chen, K., Wu, X., Wei, Z., Zhang, S.-Y., Song, B., et al. (2019). DRUM: inference of disease-associated m6A RNA methylation sites from a multi-layer heterogeneous network. *Front. Genet.* 10:266. doi: 10.3389/fgene.2019.00266
- Tian, J., Ying, P., Ke, J., Zhu, Y., Yang, Y., Gong, Y., et al. (2020). ANKLE1 N6-methyladenosine-related variant is associated with colorectal cancer risk by maintaining the genomic stability. *Int. J. Cancer* 146, 3281–3293. doi: 10.1002/ijc.32677
- von Bergwelt-Baildon, M. S., Kondo, E., Klein-González, N., and Wendtner, C. M. (2011). The cyclins: a family of widely expressed tumor antigens? *Expert Rev. Vaccines* 10, 389–395. doi: 10.1586/erv.10.170
- Wang, X., Lu, Z., Gomez, A., Hon, G. C., Yue, Y., Han, D., et al. (2014). N6-methyladenosine-dependent regulation of messenger RNA stability. *Nature* 505, 117–120. doi: 10.1038/nature12730
- Wang, X., Zhao, B. S., Roundtree, I. A., Lu, Z., Han, D., Ma, H., et al. (2015). N(6)-methyladenosine modulates messenger RNA translation efficiency. *Cell* 161, 1388–1399. doi: 10.1016/j.cell.2015.05.014
- White, R. J. (2004). RNA polymerase III transcription and cancer. *Oncogene* 23, 3208–3216. doi: 10.1038/sj.onc.1207547
- Wu, T., and Dai, Y. (2017). Tumor microenvironment and therapeutic response. *Cancer Lett.* 387, 61–68. doi: 10.1016/j.canlet.2016.01.043
- Yang, J., Liao, D., Chen, C., Liu, Y., Chuang, T.-H., Xiang, R., et al. (2013). Tumor-associated macrophages regulate murine breast cancer stem cells through a novel paracrine EGFR/Stat3/Sox-2 signaling pathway. *Stem Cells* 31, 248–258. doi: 10.1002/stem.1281
- Yang, J., Risch, E., Zhang, M., Huang, C., Huang, H., and Lu, L. (2017). Association of tRNA methyltransferase NSUN2/IGF-II molecular signature with ovarian cancer survival. *Future Oncol.* 13, 1981–1990. doi: 10.2217/fon-2017-0084
- Yang, X., Yang, Y., Sun, B.-F., Chen, Y.-S., Xu, J.-W., Lai, W.-Y., et al. (2017). 5-methylcytosine promotes mRNA export – NSUN2 as the methyltransferase and ALYREF as an m5C reader. *Cell Res.* 27, 606–625. doi: 10.1038/cr.2017.55
- Yi, J., Gao, R., Chen, Y., Yang, Z., Han, P., Zhang, H., et al. (2017). Overexpression of NSUN2 by DNA hypomethylation is associated with metastatic progression in human breast cancer. *Oncotarget* 8, 20751–20765.
- Yin, L., Duan, J.-J., Bian, X.-W., and Yu, S. (2020). Triple-negative breast cancer molecular subtyping and treatment progress. *Breast Cancer Res.* 22:61. doi: 10.1186/s13058-020-01296-5
- Zarrilli, G., Businello, G., Dieci, M. V., Paccagnella, S., Carraro, V., Cappellesso, R., et al. (2020). The tumor microenvironment of primitive and metastatic breast cancer: implications for novel therapeutic strategies. *Int. J. Mol. Sci.* 21:8102. doi: 10.3390/ijms21218102
- Zhao, Q., Zhao, Y., Hu, W., Zhang, Y., Wu, X., Lu, J., et al. (2020). m6A RNA modification modulates PI3K/Akt/mTOR signal pathway in gastrointestinal cancer. *Theranostics* 10, 9528–9543. doi: 10.7150/thno.42971

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 Huang, Pan, Wang, Du, Xu, Wang 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.



DeepOMe: A Web Server for the Prediction of 2'-O-Me Sites Based on the Hybrid CNN and BLSTM Architecture

Hongyu Li^{1,2}, Li Chen¹, Zaoli Huang¹, Xiaotong Luo¹, Huiqin Li¹, Jian Ren^{1*} and Yubin Xie^{1*}

¹ School of Life Sciences, Precision Medicine Institute, The First Affiliated Hospital, Sun Yat-sen University, Guangzhou, China, ² School of Computer Science and Engineering, Sun Yat-sen University, Guangzhou, China

OPEN ACCESS

Edited by:

Jia Meng,
Xi'an Jiaotong-Liverpool University,
China

Reviewed by:

Kunqi Chen,
University of Liverpool,
United Kingdom
Lin Zhang,
China University of Mining
and Technology, China

*Correspondence:

Yubin Xie
xieyb6@mail.sysu.edu.cn
Jian Ren
renjian@sysucc.org.cn

Specialty section:

This article was submitted to
Epigenomics and Epigenetics,
a section of the journal
Frontiers in Cell and Developmental
Biology

Received: 28 March 2021

Accepted: 23 April 2021

Published: 14 May 2021

Citation:

Li H, Chen L, Huang Z, Luo X,
Li H, Ren J and Xie Y (2021)
DeepOMe: A Web Server for
the Prediction of 2'-O-Me Sites
Based on the Hybrid CNN
and BLSTM Architecture.
Front. Cell Dev. Biol. 9:686894.
doi: 10.3389/fcell.2021.686894

2'-O-methylations (2'-O-Me or Nm) are one of the most important layers of regulatory control over gene expression. With increasing attentions focused on the characteristics, mechanisms and influences of 2'-O-Me, a revolutionary technique termed Nm-seq were established, allowing the identification of precise 2'-O-Me sites in RNA sequences with high sensitivity. However, as the costs and complexities involved with this new method, the large-scale detection and in-depth study of 2'-O-Me is still largely limited. Therefore, the development of a novel computational method to identify 2'-O-Me sites with adequate reliability is urgently needed at the current stage. To address the above issue, we proposed a hybrid deep-learning algorithm named DeepOMe that combined Convolutional Neural Networks (CNN) and Bidirectional Long Short-term Memory (BLSTM) to accurately predict 2'-O-Me sites in human transcriptome. Validating under 4-, 6-, 8-, and 10-fold cross-validation, we confirmed that our proposed model achieved a high performance (AUC close to 0.998 and AUPR close to 0.880). When testing in the independent data set, DeepOMe was substantially superior to NmSEER V2.0. To facilitate the usage of DeepOMe, a user-friendly web-server was constructed, which can be freely accessed at <http://deepome.renlab.org>.

Keywords: CNN, BLSTM, web service, RNA modification, 2'-O-methylation

INTRODUCTION

To date, hundreds of different RNA modifications have been identified in human transcriptome, and found to be critical in the regulation of various transcriptional events (Behm-Ansmant et al., 2011). Among those, 2'-O-methylation (2'-O-Me) is one of the most abundant RNA modifications, presenting in transfer RNAs (tRNAs) (Somme et al., 2014), ribosomal RNAs (rRNAs) (Rebane and Metspalu, 2002), small nuclear/small nucleolar RNAs (snRNAs/snoRNAs) (Darzacq et al., 2002), microRNAs (miRNAs) (Li, 2005)/Piwi-interacting RNAs (piRNAs) (Yu et al., 2005), and some messenger RNAs (mRNAs) (Dai et al., 2017). The addition of methyl groups on the ribose moiety can affect sterical properties, hydrogen-bonding potential, and structural rigidity of the target RNA (Kierzek et al., 2009; Hengesbach and Schwalbe, 2014), and orchestrating the biogenesis (Ojha et al., 2020), metabolism (Salem et al., 2019), and functions (Choi et al., 2018) of these RNA molecules.

Given its functional importance, the precise detection and functional analysis of 2'-O-Me are important research topics in the community.

Recently, several experimental techniques were developed to pinpoint the precise 2'-O-Me sites. For example, perchloric acid (HClO₄) hydrolysis (Baskin and Dekker, 1967), periodate oxidation hydrolysis (Trim and Parker, 1972), chromatography and mass-spectrometry (Abbate and Rottman, 1972; Sardana and Fuke, 1980; Fenghe and McCloskey, 1999; Kirpekar et al., 2005). At present, high-throughput techniques that established based on deep sequencing were also reported. Typical examples included RiboMethSeq (Krogh et al., 2016; Eraldes et al., 2017; Sharma et al., 2017; Zhou et al., 2017), 2'-OMe-Seq (Incarnato et al., 2017), RibOxi-Seq (Zhu et al., 2017), and Nm-seq (Dai et al., 2017; Hsu et al., 2019).

Although the previous mentioned high-throughput techniques can provide single-nucleotide mapping of 2'-O-Me sites at transcriptome level, the experimental procedure is still expensive and labor-exhausting. Therefore, there is still an urgent need of a computational model to mine the sequence feature of 2'-O-Me sites and identify the 2'-O-Me sites *in silico*. So far, several computational methods such as iRNA-2methyl (Qiu et al., 2017), Deep-2'-O-Me (Mostavi et al., 2018), iRNA-2OM (Yang et al., 2018), NMSEER V2.0 (Zhou et al., 2019), and iRNA-PseKNC (Tahir et al., 2019) have been developed. However, many issues remain in these methods, leaving plenty of room for improvement. Firstly, 2'-O-Me can occur in all types of RNA nucleotides, resulting an extremely imbalanced dataset between positive and negative samples. The traditional classification algorithm, which aims at the overall classification accuracy, pays too much attention to the major class, leading to poor performances in minor class and high false positives. Secondly, previous studies have randomly sampled subsequences near experimentally identified 2'-O-Me sites as negative sequences. This procedure can produce a high degree of similarity between extracted positive and negative sequences in training dataset, which limits the accuracy of traditional sequence-based models. Third, many tools lack a convenient webserver, hindering their widespread use in biological scenario. Therefore, the development of a reliable prediction tool that can not only extract useful features from the primary mRNA sequences but also produce high-precision results is still an important problem to be solved.

The performance of traditional machine learning algorithms relies heavily on data representations. However, features are typically designed by human engineers with extensive domain expertise, and identifying which features are more appropriate for the given task remains difficult. Thanks to the ability of deep learning architectures in automatically extracting high-representation information in the raw data, the application of deep learning framework is a promising way to address the above issues. In recent years, many attempts have been made to apply deep learning algorithms in biological research. For example, DeepBind (Alipanahi et al., 2015) for predicting DNA- and RNA-binding specificity, AlphaFold (Senior et al., 2020) for predicting protein structure, scDeepCluster (Tian et al., 2019) for clustering

single cell RNA-seq data, DeepCpG (Angermueller et al., 2017) for predicting single-cell DNA methylation state. Considering the characteristics of 2'-O-Me, deep learning algorithms are more suitable to analyses the patterns of 2'-O-Me and thus may greatly improve the prediction performance.

In this article, we present DeepOMe, a web server based on a hybrid deep learning architecture for predicting 2'-O-Me sites in Human mRNA. To our best knowledge, our work is the first effort to use the combination of CNN with RNN in the prediction of mRNA modification sites under the sequence-to-sequence mode. Moreover, a webserver was further developed and makes it easier for researchers and experimenters to use our proposed model.

MATERIALS AND METHODS

Dataset Collection

The training and test dataset of DeepOMe was constructed from the recently developed Nm-seq experiment (Dai et al., 2017) which comprised of 4,481 2'-O-Me sites in human transcriptome. The site data were first preprocessed and split into training and independent test set using the scheme presented in **Supplementary Figure 1**. Firstly, 2'-O-Me sites in intergenic region were removed. Due to the reason that our collected data had two coordinates versions (GRCh37 and GRCh38), we next converted original GRCh37 coordinates to GRCh38 coordinates using LiftOver and further mapped it to human transcripts for better transcriptome annotation. Transcript sequences with at least one mapped 2'-O-Me site were extracted according to the corresponding gene set annotation. If the same 2'-O-Me site located in multiple transcripts, the longest transcript were retained in our dataset. Finally, we collected 2,285 RNA sequences with 3,052 2'-O-Me sites as the final data set. We randomly selected 10% of the collected RNA sequences as independent testing set, and the remaining sequences were regarded as training set. As the result, we assembled 2,046 sequences with 2,743 2'-O-Me sites as the training set, and 239 sequences with 309 2'-O-Me sites as the testing set.

Extract Features From Primary mRNA Sequences

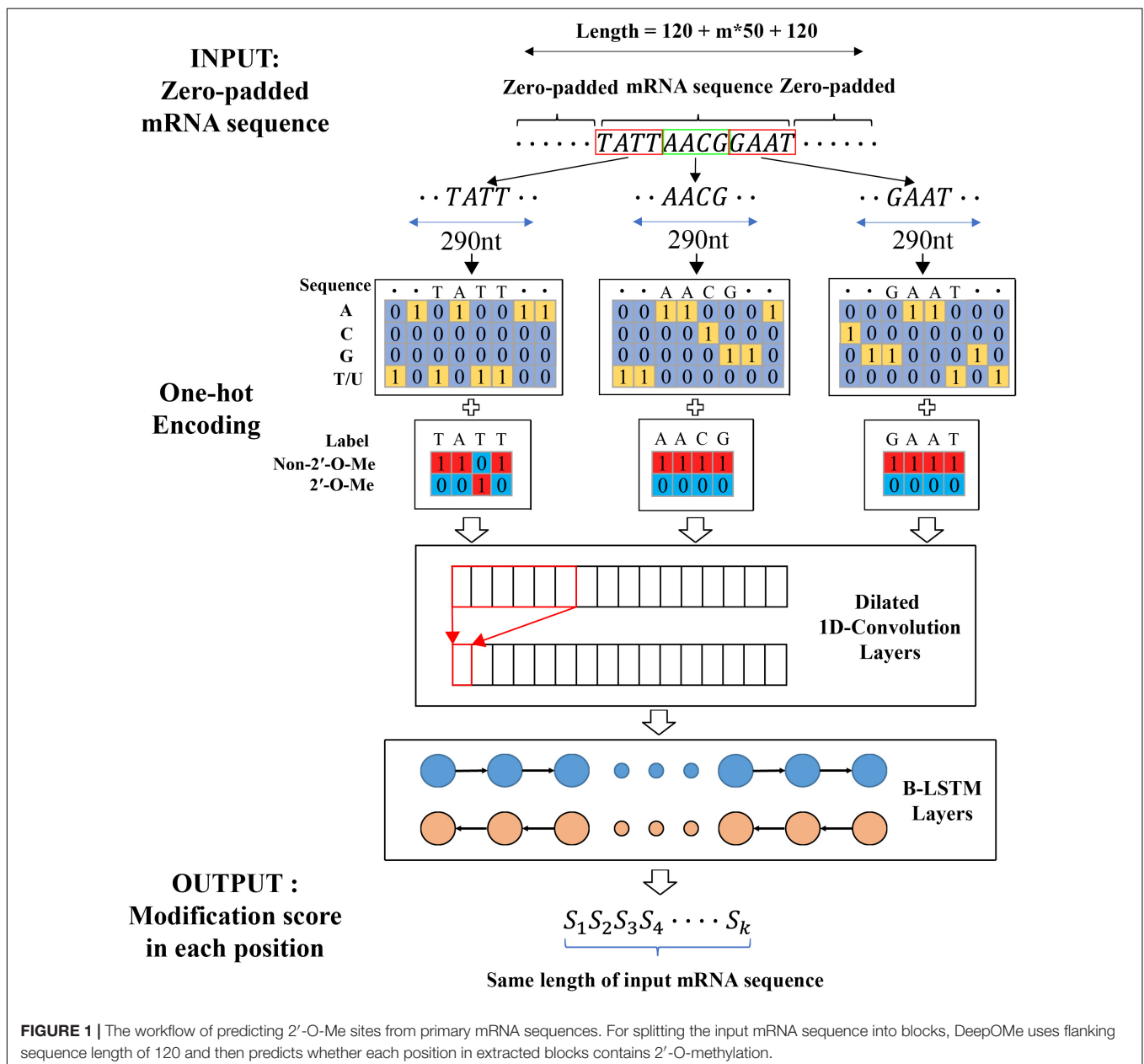
As mentioned above, previous studies extracted the flanking region of specific length around each 2'-O-Me site as the positive sequences for the training process. To create non-2'-O-Me sites or negative sequences, they randomly selected the non-modified RNA sites around known 2'-O-Me sites and captured its surrounding nucleotide sequences as negative sequences. This procedure suffers from several pitfalls.

First of all, the training set would contain overlapping sequences if the randomly selected negative sites were adjacent to 2'-O-Me sites. This would result in high similarity between positive sequences and negative sequence. The similar sequences would generate many redundant sequence-based features and thus make sequence-based machine learning algorithms difficult to train a validity predictor. To avoid such a scenario, a positive-to-negative ratio (1:10 in NmSEER V2.0; 1:1 in iRNA-2OM; 1:4 in Deep-2'-O-Me) in training set should be manually set. Since

in natural transcripts the number of 2'-O-Me sites and non-2'-O-Me sites are highly imbalanced, this kind of operations may always generate many false positives.

To solve these problems, the similar procedure from Jaganathan et al. (2019) was chosen to extract input features and output labels from primary mRNA sequences (as shown in **Figure 1**). Firstly, the transcript sequence was one-hot encoded as follows: A, C, G, T/U mapped to [1,0,0,0], [0,1,0,0], [0,0,1,0], [0,0,0,1], respectively. Next, the one-hot encoded sequence was zero-padded until the length became a multiple of 50 in order to successfully split into non-overlapped blocks of 50 nt. To capture sequence dependent features, such mRNA sequence was further zero-padded at the 5'- and 3'-end with a flanking sequence of

length L . At last, the padded sequence was split into blocks in such a way that the i^{th} block consisted of nucleotide positions from $50(i-1)-L+1$ to $50i+L$. Therefore, the 50nt center regions in the i^{th} block and $(i+1)^{th}$ block had no overlapping sequence in original mRNA sequence. Similarly, the modification output label sequence was one-hot encoded as follows: 2'-O-Me modification and non-2'-O-Me modification were mapped to [0,1] and [1,0] respectively. The one-hot encoded label sequence was zero-padded until the length became a multiple of 50 and then further zero-padded at the start and the end with a flanking sequence of length L . The padded label sequence was split into blocks using the same procedure as described for the inputted mRNA sequence. The extracted one-hot encoded



nucleotides sequences and the corresponding one-hot encoded label sequences were used as inputs and the target outputs to train and evaluate our model.

Architecture of the CNN-BLSTM Model

Figure 2A shows our proposed CNN-BLSTM architecture. DeepOMe is composed of 10 layers of CNN and 2 layers of B-LSTM (Schuster and Paliwal, 1997). The model structure consists of input layer, CNN layers, BLSTM layers, fully connected layer, and the output layer. The input layer can receive one-hot encoded sequence data. In the CNN layers, we first enriched the representation in the Stem Block (**Supplementary Figure 2A**) by computing multiple feature maps with different kernel sizes (Szegedy et al., 2015). Then, we stacked three residual blocks (He et al., 2016) for local feature extraction. The convolution operation in Stem block and ResBlock was 1D-convolution with kernel size of 10 and dilation rate of 2 (**Supplementary Figure 2B**). The CNN layers were used as preprocessing step to extract the deep spatial features from the input sequences. Then, these deep features were fed into two BLSTM layers with 32 units for learning of sequence-dependent features. The last layer in our model is a fully-connected layer with softmax activation, which was used to generate the final prediction score. The detailed architectural information was listed in **Supplementary Table 1**.

Experimental Setup

The proposed model was implemented with the TensorFlow library (Abadi et al., 2016) in Python and trained on an NVIDIA GTX2080 GPU. The proposed model was trained through 100 epochs using batch size of 200. The categorical cross entropy loss between the target and the predicted outputs was minimized using Adam optimizer (Kingma and Ba, 2014). The initial learning rate of 0.001 was used to train the model. Early-stopping (Caruana et al., 2001) was used to control overfitting. We monitored the validation loss at each epoch. When the validation loss has not improved after ten epochs, training is interrupted.

Evaluation Metrics

Testing set was used to validate our proposed model comparing with available prediction tools after cross-validation. The performance was evaluated based on several metrics, namely area under Precision-Recall Curve (AUPR), area under Receiver Operating Characteristic Curve (AUC), sensitivity (Sn), specificity (Sp), precision (Pr), accuracy (Acc), and Matthew's correlation coefficient (Mcc).

When evaluating model's performance in full mRNA sequence, an accuracy metric was largely ineffective since most of the positions in mRNA sequence are not 2'-O-Me sites. The prediction model was like the recommender systems which was to suggest the most proper modification sites in mRNA sequences. Thus, top-k accuracy was more appropriate in such situation. When comparing among different methods, we evaluated the top-k accuracy besides the AUC and AUPR metrics. The top-k accuracy is defined as follows: Suppose the test set has k positions that belong to the right class which is 2'-O-Me site. We choose the threshold so that exactly k test set positions are

predicted as belonging to the right class. The fraction of these k predicted positions that truly belong to the right class is reported as the top-k accuracy.

RESULTS

Flanking Sequence Length Analysis

It is necessary to determine the optimal flanking sequence length L of input sequences for identifying 2'-O-Me sites. Generally speaking, if the flanking sequence around the known 2'-O-Me site is too short, it may not carry enough information for prediction and will lead to poor performance. Otherwise, If the flanking sequence is too long, it may carry too much redundant information, leading to poor generalization. Thus, we first analyzed the averaged AUC and AUPR of the proposed model with different flanking sequence length under 4-fold cross-validation. As shown in **Figures 2B,C** the search step size for flanking sequence length was 10 nt, with a range of 0 and 150. According to the evaluation results, when the flanking sequence length equals to 120 nt and block length equals to 290 nt, the performance generated by our proposed model was the best (AUC = 0.9975, AUPR = 0.8818). Therefore, we selected the flanking sequence with length of 120.

Evaluation of the Prediction Performance

To evaluate the prediction performance of DeepOMe, we performed 4-, 6-, 8-, and 10-fold cross-validation of the training set. **Figures 3A,B** shows the ROC and PR curves of our proposed CNN-BLSTM model under 4-, 6-, 8-, and 10-fold cross-validations with flanking sequence length of 120. As a result, DeepOMe showed an acceptable performance in n -fold cross-validations with the area under the ROC curves (AUROC) close to 0.998 and area under the PR curves (AUPR) close to 0.880.

To rigorously evaluate the prediction and generalizability performance of DeepOMe, we next compared it with other state-of-art predictors using the independent test set. Since only iRNA-2OM, iRNA-2methyl, and NMSEER V2.0 provided web-server or standalone package for usage, the comparison will only perform between them. During the comparison, we further found that there were no responses in the webserver of iRNA-2methyl, hence, the final comparison only performs between DeepOMe, NmSEER V2.0, and iRNA-2OM.

Figures 3C,D presented the comparison results in ROC curves and PR curves. The results showed that the DeepOMe achieved a better performance (AUROC = 0.993, 95%CI:0.993-0.993; AUPR = 0.843) in the testing set than NmSEER V2.0 (AUROC = 0.5969, 95%CI:0.599-0.600; AUPR = 0.00066) and iRNA-2OM (AUROC = 0.6065, 95%CI:0.601-0.612; AUPR = 0.06538). When testing in full mRNA sequences, we further compared the top-k accuracy between DeepOMe, NmSEER V2.0, and iRNA-2OM. The comparison results in **Table 1** suggested that DeepOMe (Top-1 Acc = 0.8602, Top-100 Acc = 0.9563) was more sensitive and robust than NmSEER V2.0 (Top-1 Acc = 0.0, Top-100 Acc = 0.1004) and iRNA-2OM (Top-1 Acc = 0.0, Top-100 Acc = 0.1087).

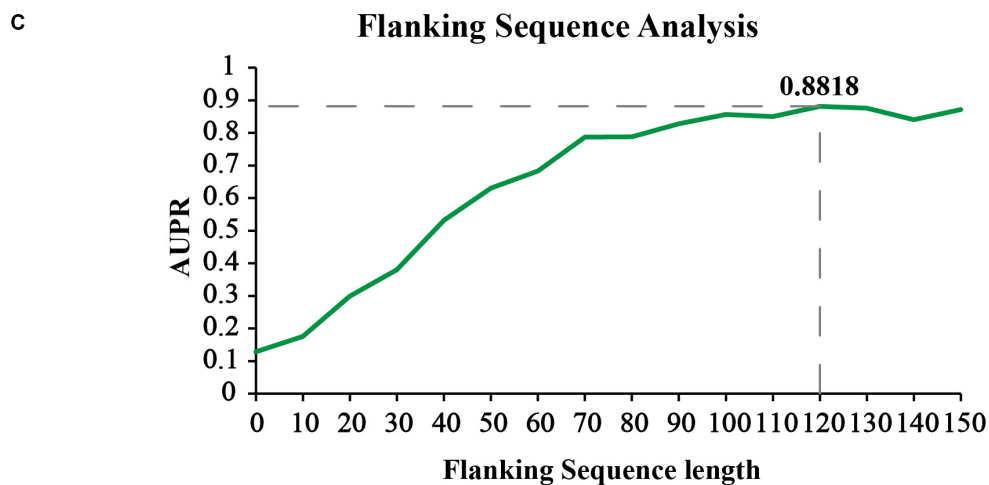
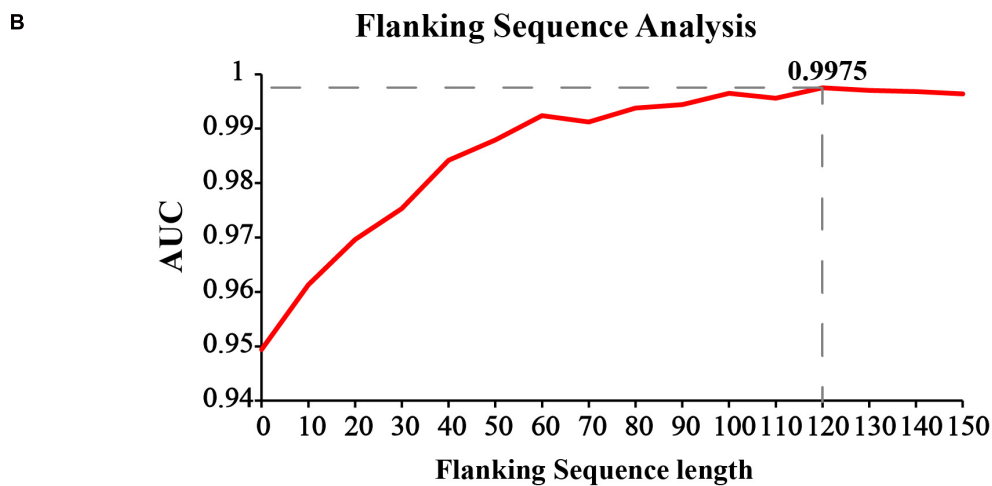
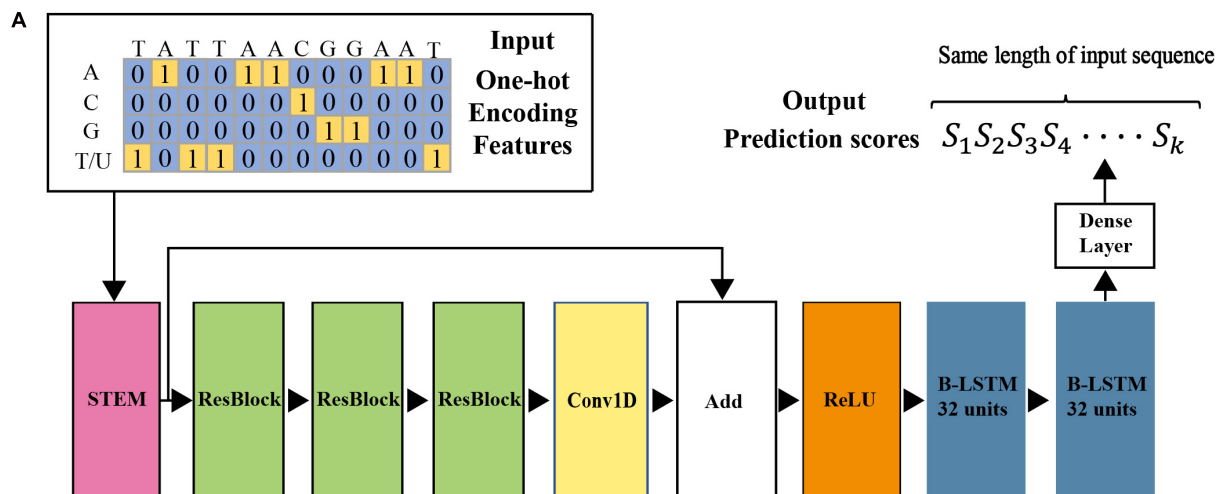


FIGURE 2 | The construction of prediction model in DeepOME. **(A)** Network architecture of the DeepOME prediction model. Flanking sequence selection under 4-fold cross-validation by AUPR **(B)** and AUROC **(C)**.

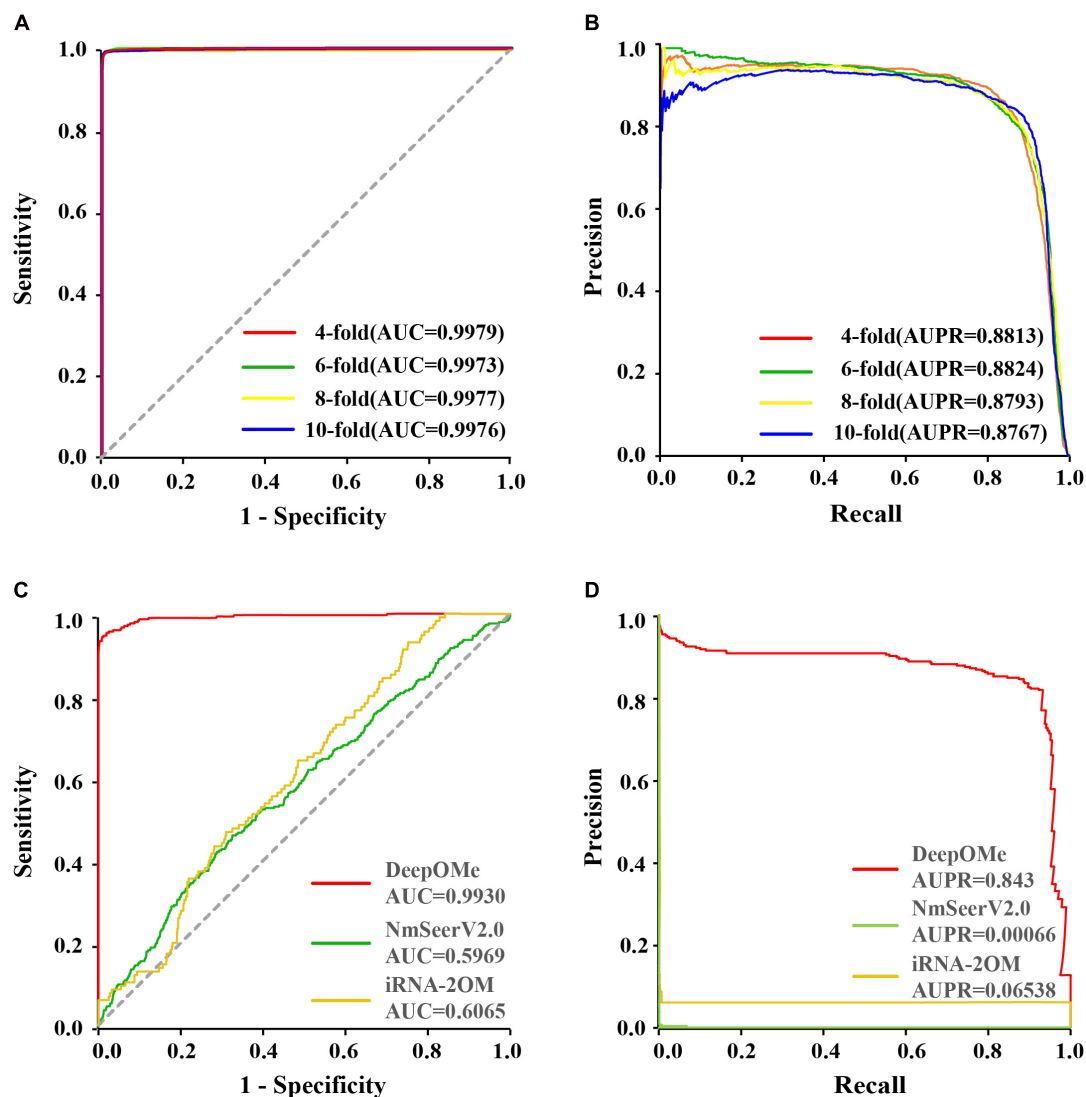


FIGURE 3 | Performance evaluation and comparison. The ROC (A) and PR (B) curves in 4-, 6-, 8-, 10-fold cross-validation. The ROC (C) and PR (D) curves in testing set between DeepOMe, NmSeer V2.0 and iRNA-2OM.

To evaluate the sequence similarity between the predicted sites and the detected sites in transcriptome, sequence logos were generated using WebLogo (Crooks et al., 2004) in training and testing set. **Supplementary Figure 3** presented the graphical representation of sequence similarity. The results showed that the predicted sites under different thresholds were similar to the detected sites both in training set and testing set, proving that our proposed model could precisely identify 2'-O-Me sites.

Web-Server

To facilitate the use of our prediction models, we next developed an online predictor called DeepOMe for the community. The predictor is freely available at <http://deepome.renlab.org>. DeepOMe only requires mRNA sequences to run a prediction. Multiple mRNA sequences can be input into the text area or uploaded as a single FASTA file. For users' convenience, we

selected three thresholds based on the 10-fold cross-validation results (**Figure 4A**), which correspond to the false discovery rate of 0.10, 0.15, and 0.20. The detailed performance values under these three thresholds are shown in **Supplementary Table 2**. Besides, users can select the threshold by setting the false discovery rate in advanced option menu. After the query sequences are submitted to DeepOMe, users can check its running status in the result panel in real time. When the prediction is complete, the button that links out to the result page will be clickable (**Figure 4B**). **Figure 4C** provides a snapshot for the result page of the example mRNA sequence. The prediction position, score and prediction threshold were listed in an interactive table, which allows the users to easily search and sort the results. Remarkably, to facilitate a further analysis of the protein function and RNA structure, we also implemented an automatic pipeline for visualizing the prediction

TABLE 1 | Comparison of Top-k Accuracy between DeepOMe, NmSEER V2.0, and iRNA-2OM in testing set.

Top-k Accuracy	DeepOMe	NmSEER V2.0	iRNA-2OM
Top-1 Accuracy	0.8602	0.0	0.0
Top-3 Accuracy	0.9039	0.0131	0.0
Top-5 Accuracy	0.9082	0.0175	0.0
Top-10 Accuracy	0.9126	0.0175	0.0043
Top-20 Accuracy	0.9257	0.0306	0.0130
Top-30 Accuracy	0.9344	0.0611	0.0130
Top-40 Accuracy	0.9476	0.0699	0.0261
Top-50 Accuracy	0.9520	0.0699	0.0478
Top-60 Accuracy	0.9520	0.0830	0.0696
Top-70 Accuracy	0.9520	0.0917	0.0739
Top-80 Accuracy	0.9563	0.0961	0.0826
Top-90 Accuracy	0.9563	0.1004	0.0870
Top-100 Accuracy	0.9563	0.1004	0.1087

results. By integrating IBS (Liu et al., 2015), InterProScan (Jones et al., 2014), and ViennaRNA (Lorenz et al., 2011) into the web server, DeepOMe can present the graphical representation of the input mRNA sequence together with their predicted sites in the visualization panel. **Figures 4D,E** provide snapshots for the visualization results of RNA secondary structure and protein domain organization. The diagrams can be saved as a vector graphic (SVG) for further analysis.

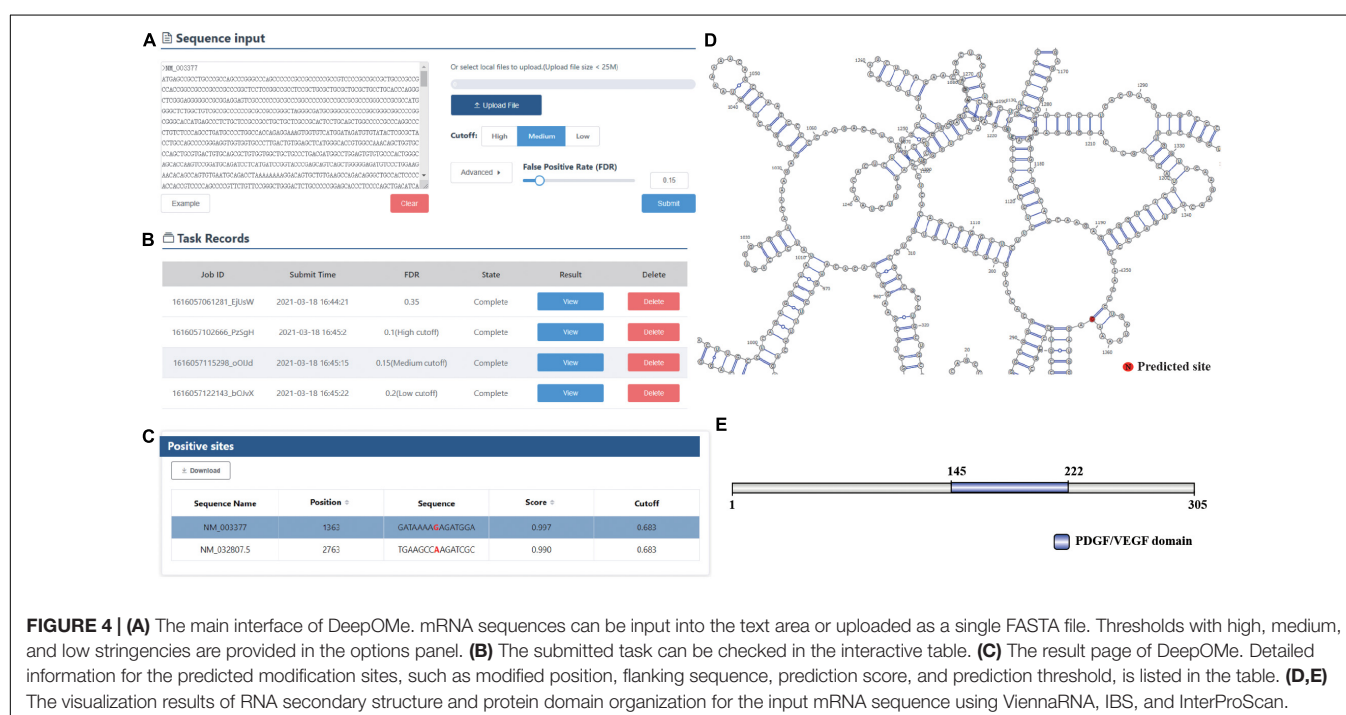
DISCUSSION

2'-O-methylation plays critical roles in regulating gene expressions at the post-transcriptional levels. Thus, proper

identification of the 2'-O-Me site is essential to understand the mechanism of RNA metabolisms. 2'-O-Me can occur in any base on the mRNA sequence. Given a mRNA sequence, we need to get an output sequence with the same length of input sequence. The score in each position of output sequence represents whether this position in the input mRNA contains 2'-O-Me. Therefore, the 2'-O-Me site prediction problem can be considered as a many-to-many prediction problem.

However, previous studies tried to train the prediction model based on a Many-to-One mode. They had to randomly select non-2'-O-Me sites around the known 2'-O-Me sites as negative samples, which resulted in high sequence similarity between positive and negative sequences. Besides, to reduce the degree of imbalance in their training data set, the negative sites were manually down-sampled to obtain a relatively small positive-to-negative ratio. However, in reality, the positive-to-negative ratio in a given RNA sequence was always extremely high, and thus caused their models to have poor generalization ability in unseen data. These were the two main reasons why their models received very poor performance in our testing set.

Unlike the previous works that use handcrafted features for classification, DeepOMe could automatically extract the deep features from primary mRNA sequences by CNN layers. DeepOMe was proven to be more efficient than the available method in terms of all evaluation metrics. We found several factors that may explain the high performance achieved by our proposed model. Firstly, the procedure we used to train and test the models was the many-to-many mode. Thus, there was no need to manually balance the training dataset in our model, allowing to learn sufficient information between 2'-O-Me and non-2'-O-Me sites and achieving lower false positives. Secondly, the use of the dilated 1D CNN compared to the



traditional 1D CNN allowed our model to cover more relevant information by increasing the receptive field of the filters. Additionally, the stacked Resblocks used in our model allowed us to build a deeper network and take advantages of the powerful representational ability of deep neural network. At last, the application of bidirectional LSTM was able to exploit meaningful representations from upstream and downstream sequences. The comparison results suggest that the combination of CNN and BLSTM can successfully capture the key features of the entire mRNA sequences. Although promising performance was obtained in DeepOMe, a number of future improvements are expected. First of all, we have designed only a relatively simple CNN-based neural network model in our current version. Various deeper and wider CNN architectures were awaited exploration in the future. Secondly, attention mechanism will be introduced to achieve better representation for contextual information in the future version. Last but not least, we have now trained the model only based on the experimental 2'-O-Me data for *Homo sapiens*. The prediction models for other species such as *Mus musculus* will be established in the future.

DATA AVAILABILITY STATEMENT

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

REFERENCES

- Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., et al. (2016). "TensorFlow: A system for large-scale machine learning," in *12th USENIX Symposium on Operating Systems Design and Implementation (OSDI 16)*, (California: USENIX Association), 265–283.
- Abbate, J., and Rottman, F. (1972). Gas chromatographic method for determination of 2-O-methylation in RNA. *Anal. Biochem.* 47, 378–388. doi: 10.1016/0003-2697(72)90131-5
- Alipanahi, B., Delong, A., Weirauch, M. T., and Frey, B. J. (2015). Predicting the sequence specificities of DNA- and RNA-binding proteins by deep learning. *Nat. Biotechnol.* 33, 831–838. doi: 10.1038/nbt.3300
- Angermueller, C., Lee, H. J., Reik, W., and Stegle, O. (2017). DeepCpG: accurate prediction of single-cell DNA methylation states using deep learning. *Genome Biol.* 18:67.
- Baskin, F., and Dekker, C. A. (1967). A rapid and specific assay for sugar methylation in ribonucleic acid. *J. Biol. Chem.* 242, 5447–5449. doi: 10.1016/s0021-9258(18)99445-7
- Behm-Ansmant, I., Helm, M., and Motorin, Y. (2011). Use of Specific Chemical Reagents for Detection of Modified Nucleotides in RNA. *J. Nucleic Acids* 2011:408053.
- Caruana, R., Lawrence, S., and Giles, C. L. (2001). *Overfitting In Neural Nets: Backpropagation, Conjugate Gradient, And Early Stopping*, In: *Advances In Neural Information Processing Systems*. Massachusetts: MIT Press, 402–408.
- Choi, J., Indrisunaite, G., Demirci, H., Jeong, K.-W., Wang, J., Petrov, A., et al. (2018). 2'-O-methylation in mRNA disrupts tRNA decoding during translation elongation. *Nat. Struct. Mol. Biol.* 25, 208–216. doi: 10.1038/s41594-018-0030-z
- Crooks, G. E., Hon, G., Chandonia, J.-M., and Brenner, S. E. (2004). WebLogo: a sequence logo generator. *Genome Res.* 14, 1188–1190. doi: 10.1101/gr.849004
- Dai, Q., Moshitch-Moshkovitz, S., Han, D., Kol, N., Amariglio, N., Rechavi, G., et al. (2017). Nm-seq maps 2'-O-methylation sites in human mRNA with base precision. *Nat. Methods.* 14, 695–698. doi: 10.1038/nmeth.4294

AUTHOR CONTRIBUTIONS

HL implemented the DeepOMe algorithm and wrote the manuscript. LC and XL manually collected 2'-O-Me data from published literatures and performed data pre-processing. ZH and HL are respectively responsible for the front-end page display and back-end logic design of the DeepOMe website. JR was responsible for supervision, funding acquisition, and writing-review. YX supervised this work, reviewed and edited the manuscript. All authors have read and approved the manuscript.

FUNDING

This work was supported by the National Natural Science Foundation of China [91753137, 31771462, 81772614, U1611261, 31801105, and 81802438], the National Key R&D Program of China [2017YFA0106700], the Program for Guangdong Introducing Innovative and Entrepreneurial Teams [2017ZT07S096], and the Guangdong Basic and Applied Basic Research Foundation [2018A030313323 and 2020A1515010220].

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fcell.2021.686894/full#supplementary-material>

- Darzacq, X., Jady, B. E., Verheggen, C., Kiss, A. M., Bertrand, E., Kiss, T. (2002). Cajal body-specific small nuclear RNAs: a novel class of 2'-O-methylation and pseudouridylation guide RNA. *EMBO J.* 21, 2746–2756. doi: 10.1093/emboj/21.11.2746
- Erales, J., Marchand, V., Panthu, B., Gillot, S., Belin, S., and Ghayad, S. E. (2017). Evidence for rRNA 2'-O-methylation plasticity: control of intrinsic translational capabilities of human ribosomes. *Proc. Natl. Acad. Sci. U. S. A.* 114, 12934–12939. doi: 10.1073/pnas.1707674114
- Fenghe, Q., and McCloskey, J. A. (1999). Selective detection of ribose-methylated nucleotides in RNA by a mass spectrometry-based method. *Nucleic Acids Res* 27:e20.
- He, K., Zhang, X., Ren, S., and Sun, J. (2016). "Deep Residual Learning for Image Recognition," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, (New York: IEEE), 770–778.
- Hengesbach, M., and Schwalbe, H. (2014). Structural Basis for Regulation of Ribosomal RNA 2-O-Methylation. *Angew. Chem.* 53, 1742–1744. doi: 10.1002/anie.201309604
- Hsu, P. J., Fei, Q., Dai, Q., Shi, H., Dominissini, D., Ma, L., et al. (2019). Single base resolution mapping of 2'-O-methylation sites in human mRNA and in 3' terminal ends of small RNAs. *Methods* 156, 85–90. doi: 10.1016/j.ymeth.2018.11.007
- Incarnato, D., Anselmi, F., Morandi, E., Neri, F., Maldotti, M., and Rapelli, S. (2017). High-throughput single-base resolution mapping of RNA 2'-O-methylated residues. *Nucleic Acids Res.* 45, 1433–1441.
- Jaganathan, K., Kyriazopoulou Panagiotopoulou, S., Mcrae, J. F., Darbandi, S. F., Knowles, D., Li, Y. I., et al. (2019). Predicting Splicing from Primary Sequence with Deep Learning. *Cell* 176:e524.
- Jones, P., Binns, D., Chang, H.-Y., Fraser, M., Li, W., Mcanulla, C., et al. (2014). InterProScan 5: genome-scale protein function classification. *Bioinformatics* 30, 1236–1240. doi: 10.1093/bioinformatics/btu031
- Kierzek, E., Pasternak, A., Pasternak, K., Gdaniec, Z., Yildirim, I., Turner, D. H., et al. (2009). Contributions of stacking, preorganization, and hydrogen bonding

- to the thermodynamic stability of duplexes between RNA and 2'-O-methyl RNA with locked nucleic acids. *Biochemistry* 48, 4377–4387 doi: 10.1021/bi9002056
- Kingma, D., and Ba, J. (2014). "Adam: A Method for Stochastic Optimization," in *International Conference on Learning Representations*. Cornell University: New York
- Kirpekar, F., Hansen, L. H., Rasmussen, A., Poehlsgaard, J., and Vester, B. (2005). The archaeon *Haloarcula marismortui* has few modifications in the central parts of its 23 S ribosomal RNA. *J. Mol. Biol.* 348, 563–573 doi: 10.1016/j.jmb.2005.03.009
- Krogh, N., Jansson, M. D., Häfner, S. J., Tehler, D., Birkedal, U., and Christensen-Dalsgaard, M. (2016). Profiling of 2'-O-Me in human rRNA reveals a subset of fractionally modified positions and provides evidence for ribosome heterogeneity. *Nucleic Acids Res.* 44, 7884–7895. doi: 10.1093/nar/gkw482
- Li, J. (2005). Methylation protects miRNAs and siRNAs from a 3'-end uridylation activity in *Arabidopsis*. *Curr. Biol.* 15, 1501–1507. doi: 10.1016/j.cub.2005.07.029
- Liu, W., Xie, Y., Ma, J., Luo, X., Nie, P., Zuo, Z., et al. (2015). IBS: an illustrator for the presentation and visualization of biological sequences. *Bioinformatics* 31, 3359–3361. doi: 10.1093/bioinformatics/btv362
- Lorenz, R., Bernhart, S. H., Zu Siederdissen, C. H., Tafer, H., Flamm, C., Stadler, P. F., et al. (2011). ViennaRNA Package 2.0. *Algorithms Mol. Biol.* 6, 1–14.
- Mostavi, M., Salekin, S., and Huang, Y. (2018). "Deep-2'-O-Me: Predicting 2'-O-methylation sites by Convolutional Neural Networks. *Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.* 2018, 2394–2397.
- Ojha, S., Malla, S., and Lyons, S. M. (2020). snoRNPs: functions in Ribosome Biogenesis. *Biomolecules* 10:783. doi: 10.3390/biom10050783
- Qiu, W. R., Jiang, S. Y., Sun, B. Q., Xiao, X., Cheng, X., and Chou, K. C. (2017). iRNA-2methyl: identify RNA 2'-O-methylation Sites by Incorporating Sequence-Coupled Effects into General PseKNC and Ensemble Classifier. *Med. Chem.* 13, 734–743.
- Rebane, A., and Metspalu, R. (2002). Locations of several novel 2'-O-methylated nucleotides in human 28S rRNA. *BMC Mol. Biol.* 3:1. doi: 10.1186/1471-2199-3-1
- Salem, E. S., Vonberg, A. D., Borra, V. J., Gill, R. K., and Nakamura, T. (2019). RNAs and RNA-binding proteins in immuno-metabolic homeostasis and diseases. *Front. Cardiovasc. Med.* 6:106. doi: 10.3389/fcvm.2019.00106
- Sardana, M. K., and Fuke, M. (1980). A rapid procedure to determine the content of 2'-O-methylation in RNA by homochromatography. *Anal. Biochem.* 103, 285–288. doi: 10.1016/0003-2697(80)90611-9
- Schuster, M., and Paliwal, K. K. (1997). Bidirectional recurrent neural networks. *IEEE Trans. Signal Proc.* 45, 2673–2681. doi: 10.1109/78.650093
- Senior, A. W., Evans, R., Jumper, J., Kirkpatrick, J., Sifre, L., Green, T., et al. (2020). Improved protein structure prediction using potentials from deep learning. *Nature* 577, 706–710.
- Sharma, S., Marchand, V., Motorin, Y., and Lafontaine, D. (2017). Identification of sites of 2'-O-methylation vulnerability in human ribosomal RNAs by systematic mapping. *Sci. Rep.* 7, 1–15.
- Somme, J., Van Laer, B., Roovers, M., Steyaert, J., Versées, W., and Droogmans, L. (2014). Characterization of two homologous 2'-O-methyltransferases showing different specificities for their tRNA substrates. *RNA* 20, 1257–1271. doi: 10.1261/rna.044503.114
- Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., et al. (2015). "Going deeper with convolutions," in *In: Proceedings Of The Ieee Conference On Computer Vision And Pattern Recognition*, (New York: IEEE), 1–9.
- Tahir, M., Tayara, H., and Chong, K. T. (2019). iRNA-PseKNC(2methyl): identify RNA 2'-O-methylation sites by convolution neural network and Chou's pseudo components. *J. Theor. Biol.* 465, 1–6. doi: 10.1016/j.jtbi.2018.12.034
- Tian, T., Wan, J., Song, Q., and Wei, Z. (2019). Clustering single-cell RNA-seq data with a model-based deep learning approach. *Nat. Mach. Intell.* 1, 191–198. doi: 10.1038/s42256-019-0037-0
- Trim, A. R., and Parker, J. E. (1972). Nucleotide sequence in fourteen dinucleotides, modified by 2'-O-methylation, from yeast ribonucleic acid, determined by periodate degradation and by pentose analysis. *Anal. Biochem.* 46, 482–488. doi: 10.1016/0003-2697(72)90322-3
- Yang, H., Lv, H., Ding, H., Chen, W., and Lin, H. (2018). iRNA-2OM: a Sequence-Based Predictor for Identifying 2'-O-Methylation Sites in Homo sapiens. *J. Comput. Biol.* 25, 1266–1277. doi: 10.1089/cmb.2018.0004
- Yu, B., Yang, Z., Li, J., Minakhina, S., Yang, M., Padgett, R. W., et al. (2005). Methylation as a Crucial Step in Plant microRNA Biogenesis. *Science* 307, 932–935. doi: 10.1126/science.1107130
- Zhou, F., Liu, Y., Rohde, C., Pauli, C., Gerloff, D., Köhn, M., et al. (2017). AML1-ETO requires enhanced C/D box snoRNA/RNP formation to induce self-renewal and leukaemia. *Nat. Cell Biol.* 19, 844–855. doi: 10.1038/ncb3563
- Zhou, Y., Cui, Q., and Zhou, Y. (2019). NmSEER V2.0: a prediction tool for 2'-O-methylation sites based on random forest and multi-encoding combination. *BMC Bioinformatics* 20:690. doi: 10.1186/s12859-019-3265-8
- Zhu, Y., Pirnie, S. P., and Carmichael, G. G. (2017). High-throughput and site-specific identification of 2'-O-methylation sites using ribose oxidation sequencing (RibOxi-seq). *RNA* 23, 1303–1314. doi: 10.1261/rna.061549.117

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, Chen, Huang, Luo, Li, Ren 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.



Co-expression Network Revealed Roles of RNA m⁶A Methylation in Human β -Cell of Type 2 Diabetes Mellitus

Cong Chen, Qing Xiang, Weilin Liu, Shengxiang Liang, Minguang Yang and Jing Tao*

The Institute of Rehabilitation Industry, Fujian University of Traditional Chinese Medicine, Fuzhou, China

OPEN ACCESS

Edited by:

Giovanni Nigita,
The Ohio State University,
United States

Reviewed by:

Yuanhu Xuan,
Shenyang Agricultural University,
China
Xiaofeng Gu,
Institute of Biotechnology (CAAS),
China

*Correspondence:

Jing Tao
taojing01@fjtcu.edu.cn

Specialty section:

This article was submitted to
Epigenomics and Epigenetics,
a section of the journal
Frontiers in Cell and Developmental
Biology

Received: 08 January 2021

Accepted: 07 April 2021

Published: 18 May 2021

Citation:

Chen C, Xiang Q, Liu W, Liang S,
Yang M and Tao J (2021)
Co-expression Network Revealed
Roles of RNA m⁶A Methylation
in Human β -Cell of Type 2 Diabetes
Mellitus.
Front. Cell Dev. Biol. 9:651142.
doi: 10.3389/fcell.2021.651142

RNA m⁶A methylation plays an important role in the pathogenesis of type 2 diabetes mellitus (T2DM). RNA modifications and RNA-modifying regulators have recently emerged as critical factors involved in β -cell function and insulin resistance, including “writers,” “erasers,” and “readers.” However, their key roles in regulating gene expression in T2DM remain unclear. The construction of co-expression network could provide a cue to resolve this complex regulatory pathway. We collected the transcriptome datasets of β -cell in diabetic patients, calculated the partial correlation coefficient, excluded the influence from control variables of diabetes related genes, and identified the genes significantly co-expressed with m⁶A regulators. A total of 985 genes co-expressed with m⁶A regulators (Co-m⁶AR) were identified, which were enriched in metabolic process, MAPK and EGFR signaling pathways. Some of them have been confirmed to play a pivotal role in T2DM, including *CCNL2*, *CSAD*, *COX5A*, *GAB2*, and *MIRLET7I*, etc. Further, we analyzed the m⁶A modification characteristics of Co-m⁶AR in β -cell and identified 228 Co-m⁶AR containing m⁶A methylation sites, involving in several key signaling pathways regulating T2DM. We finally screened out 13 eQTL-SNPs localized in Co-m⁶ARs, and 4 have been reported strongly associated with diabetes, including *GAB2*, *LMNB2*, *XAB2*, and *RBM39*. This co-expression analysis provides important information to reveal the potential regulatory mechanism of RNA m⁶A methylation in T2DM.

Keywords: RNA m⁶A methylation, type 2 diabetes mellitus, co-expression network, RNA m⁶A methyltransferase, insulin

INTRODUCTION

Type 2 diabetes mellitus (T2DM) is considered a major health problem worldwide. This complicated metabolic disorder is characterized by chronic hyperglycemia, which can trigger β -cell function impairment, insulin resistance, and deficiency. However, the pathogenesis of T2DM remains unclear. The critical factor of T2DM is an inadequate functional islet β -cell mass, and a variety of risk factors can lead to islet β -cell dysfunction, either due to a reduced β -cell mass, or to inadequate β -cell phenotype and functional maturity (Dayeh and Ling, 2015). Insulin dysfunction can induce toxicity in different tissues leading to pathological and functional changes

in multiple organs (Singh et al., 2001; Giacco and Brownlee, 2010; Reusch and Manson, 2017). A classical insulin resistance mechanism consists in dysregulated insulin signaling with impaired phosphorylated insulin receptor substrates (IRSs) (Copps et al., 2016) and other altered factors such as insulin-regulated Ser/Thr kinases, AMP-activated protein kinase, and glucose transporter 4.

Epigenetic modifications on DNA, RNA, and proteins may provide the link gene expression into pathological mechanisms of T2DM. Studies have shown different patterns of epigenetic regulation of DNA and proteins, but RNA modifications in T2DM remain poorly understood (He, 2010). N⁶-methyladenosine (m⁶A) is one of the most abundant post-transcriptional modifications on eukaryote RNA. Over 7000 human transcripts contain at least one m⁶A site, enriched in the coding sequence and the 3' untranslated region (3'UTR), near the stop codon of mature polyadenylated mRNAs (Dominissini et al., 2012; Meyer et al., 2012). RNA m⁶A modifications of the eukaryotic transcriptome can modulate mRNA splicing, export, localization, translation, and stability (Wang et al., 2014; Liu et al., 2015). The function of this modification is controlled by a series m⁶A regulator. Intracellular m⁶A regulation is dynamic and reversible, which is composed of methyltransferases (writers) and demethylases (erasers), including *METTL3*, *METTL14*, *WTAP*, *KIAA1429*, and *RBM15* (Bokar et al., 1997; Liu et al., 2014), fat mass and obesity-associated protein (*FTO*) (Jia et al., 2011) and alkylation repair homolog protein 5 (*ALKBH5*) (Zheng et al., 2013; Fedeles et al., 2015). On the other hand, recruiting m⁶A-binding proteins called “readers” can specifically bind to m⁶A-modified transcripts including the nuclear *YTHDC1* and the cytoplasmic *TYHDC2*, *YTHDF1*, *YTHDF2*, and *YTHDF3* (Wang et al., 2014).

Homeostasis of m⁶A modifications on RNA is essential for the regulation of human transcripts expression in T2DM. Demethylase *FTO* is the first gene strongly associated with adipose mass and obesity in β -cells (Scuteri et al., 2007; Taneera et al., 2018), positively correlated with serum glucose (Blauth and Falliner, 1986), and closely associated with metabolic alterations, cardiovascular diseases, and T2DM (Boissel et al., 2009; Church et al., 2010; Lee et al., 2014). *FTO* knockout mice exhibit growth retardation, weight loss, and heart defects (Boissel et al., 2009; Christian et al., 2014), whereas *FTO*-overexpressing mice present obesity, weight gain, and enhanced food intake (Church et al., 2010). Recently, m⁶A sequencing of human T2DM islets revealed that the coefficient of variation of m⁶A methylation regulators is higher in patients with T2DM than in controls, with m⁶A in the patients with T2DM regulating β -cell gene expression to a higher extent than in controls (De Jesus et al., 2019). Further studies have shown that m⁶A methylated RNA and *METTL3* levels are consistently higher in liver tissues of patients with T2DM than in controls; hepatocyte-specific knockout of *METTL3* in mice fed with a high-fat diet (HFD) increases insulin sensitivity and decreases fatty acid synthesis (Xie et al., 2019). *METTL3* significantly modulated HFD-induced metabolic disorders, insulin sensitivity, and hepatogenic diabetes (Li et al., 2020). In addition, m⁶A and RNA sequencing of diabetic hippocampi revealed that the alteration of m⁶A modifications

might cause hippocampal neuron damage and further lead to cognitive impairment in patients with T2DM (Song et al., 2020). Collectively, studies have strongly indicated that m⁶A modifications may play a role in the pathogenesis of T2DM, but the biological role and the underlying mechanism of m⁶A still need to be explored.

Co-expression analysis is an effective bioinformatics method that has not been applied to analyze the mechanism of gene regulation. The construction of gene expression regulatory network by co-expression analysis has been widely used in the regulation of growth and development and the pathogenesis of various diseases. However, this method has not been applied to analyze the regulation of m⁶A regulator on T2DM. In the system composed of multiple variables, when studying the influence or correlation degree of one element on another, the influence of other elements is regarded as constant, that is, the influence of other factors is not considered temporarily, and the degree of the relationship between the two elements is studied separately. The numerical results are partial correlation coefficient. Thus, the utilization of partial correlation coefficient to analysis of co-expression is a powerful method. In the reported m⁶A profiling analysis, they only explore the regulatory genes of m⁶A regulator through differential expression (Scuteri et al., 2007; Taneera et al., 2018). Although this analysis method can display more information, it could not reveal the key influencing factors. Co-expression analysis can mine the most critical related genes from a large sample study. If combined with m⁶A profiling, it is relatively more able to dig out the key regulatory mechanisms.

To investigate the correlation of m⁶A modifications with human transcripts expression in patients with T2DM, we collected methylated RNA immunoprecipitation (MeRIP)-seq and RNA-seq transcriptome samples from human islet β -cells, and found 15 differentially expressed m⁶A genes and used these for analysis, including *METTL3*, *FTO*, and *YTHDF1*, etc. We further identified the potential human transcripts associated with m⁶A genes based on the partial correlation coefficient. Our analyses of GO and KEGG pathway enrichment revealed that most of the identified human transcripts were involved in metabolic pathway and MAPK1/MAPK3 signaling, which is a critical biological process controlling the pathogenesis and development of T2DM, suggesting that mRNA m⁶A methylation plays a crucial role in T2DM.

MATERIALS AND METHODS

Datasets

We downloaded the gene expression profile of GSE50398 from the GEO database, a free and openly available database. The dataset used Agilent GPL6244 Platform and includes the genome-wide mRNA expression data of the 178 samples. We downloaded the processed data, that gene and exon expression normalizations have been carried on by the TMM method (Robinson and Oshlack, 2010) and further normalization has been performed using the adjusting the expression to gene or exon length, respectively.

Partial Correlation

We defined gene expression datasets as random variables G , expression of m⁶A regulators (**Supplementary Table 1**) as random variables M , and the expression matrix of diabetes-related genes (**Supplementary Table 2**) as control variable D . The partial correlation coefficient ρ reflects the association between G and M , where the effects of D are removed. Given the linear regression of G with D and the linear regression of M with D , the ρ is calculated as the linear correlation between the two residuals.

The calculation of partial correlation coefficient, the variable m⁶A factor x , the reference variable diabetes factor Z , the variable gene y to be analyzed, and the partial correlation coefficient between x and y were calculated after excluding the influence of Z variable:

$$r_{xy.Z} = \frac{r_{xy} - r_{xZ}r_{yZ}}{\sqrt{(1 - r_{xZ}^2)(1 - r_{yZ}^2)}}$$

The H0 of partial correlation coefficient test is the partial correlation coefficient of two variables in the population is 0. Using t -test method, the formula is as follows:

$$t = \frac{r\sqrt{n-k-2}}{\sqrt{1-r^2}}$$

Where r is the corresponding partial correlation coefficient, n is the number of sample observations, K is the number of control variables, and $n-k-2$ is the degree of freedom. When $p < 0.05$, the original hypothesis is rejected.

Differential Expression

We analyzed the differentially-expressed genes in pancreatic islet β -cells between T2DM patients and healthy volunteers using GEO2R, based on R language, an online analysis tool for the GEO database. According to the GEO2R criteria, we identified genes as differentially-expressed if $\log_{2}FC > 2$ (upregulated genes) or $\log_{2}FC < -2$ (downregulated genes). Significant differences have been statistically tested and the FDR-adjusted P value has been applied, that FDR is an adjusted P value to trim false positive results. We considered adjusted P values < 0.05 as statistically significant and used the calculated value to decrease the false positive rate.

Gene Ontology and KEGG Pathway Analysis of Differentially Expressed Genes

We used gene ontology (GO) analysis to annotate genes and classify their functions into biological pathways, molecular function, and cellular components (Scuteri et al., 2007). The Kyoto Encyclopedia of Genes and Genomes (KEGG) is a set of databases that disposes biological pathways and genomes related to diseases and drugs. KEGG is a channel promoting an overall and deep understanding of biological systems (Kanehisa, 2002). An FDR-corrected p -value has been utilized, and we set a $p < 0.05$ as the cut-off criterion for statistically significant differences. Cellular components, molecular functions, and biological processes were analyzed using the DAVID online database (Huang da et al., 2009a,b).

Co-expression Network Analysis

We assessed and constructed the co-expression network of differentially-expressed genes. Flat expression patterns across samples in the derivation dataset were filtered out by excluding genes with standard deviations ≤ 0.1 . We calculated the partial correlation coefficient, ρ , for each pair of genes, and we defined all gene pairs with $\rho \geq 0.3$ as gene-gene associations in the network. In the co-expression network, the nodes represent genes and the edges represent the connections with coefficient ≥ 0.3 .

eQTL Analysis

The trans-eQTL were download from eQTLGen database¹. This database incorporates 37 datasets with a total of 31,684 individuals and established to identify the downstream consequences of trait-related genetic variants. the cis-eQTL, trans-eQTL, eQTS and replication results are available on this website. This database contains the trans-eQTL results for about 10,000 known genetic risk variants. The statistically significant trans-eQTLs are browsable, the full results were downloaded for further analysis. The m⁶A peaks were defined as a 100 bp region overlapping the m⁶A position which sequenced by Jesus group (De Jesus et al., 2019).

Software Tools

We used hierarchical cluster analysis to show the volcano plot and heat map of two groups on ImageGP website². The network generation and statistical calculations were analyzed by the (R-platform) Affy package (Gautier et al., 2004). We generated a functional co-expression network annotation using PNATHER (Thomas et al., 2003; Mi et al., 2013). We used the DAVID tool to perform GO enrichment and KEGG analyses, and to estimate network candidate communities on the basis of their associations with functional annotations (Huang da et al., 2009b).

RESULTS

Strategy

Figure 1A depicts our analysis strategy. First, we collected the transcriptome data of islet cells from diabetic patients and controls from published datasets. Then, we analyzed the differential expression of RNA m⁶A methylation regulators, including methyltransferases, demethylases, and methylation readers. Then we used genes associated with diabetes pathways as the control variables to calculate partial correlation coefficients, and we screened the genes co-expressed with RNA m⁶A regulators based on partial correlation coefficient factors. Next, we downloaded the published RNA m⁶A modification profiles of islet beta cells, analyzed the overlapping parts with the above identified genes, and further analyzed the potential biological functions of these genes through the GO enrichment and KEGG pathway analysis. Finally, we screened the trans-eQTL localized in the m⁶A peak of genes in our list.

¹www.eqtlgen.org

²http://www.ehbio.com/Cloud_Platform/front/#/

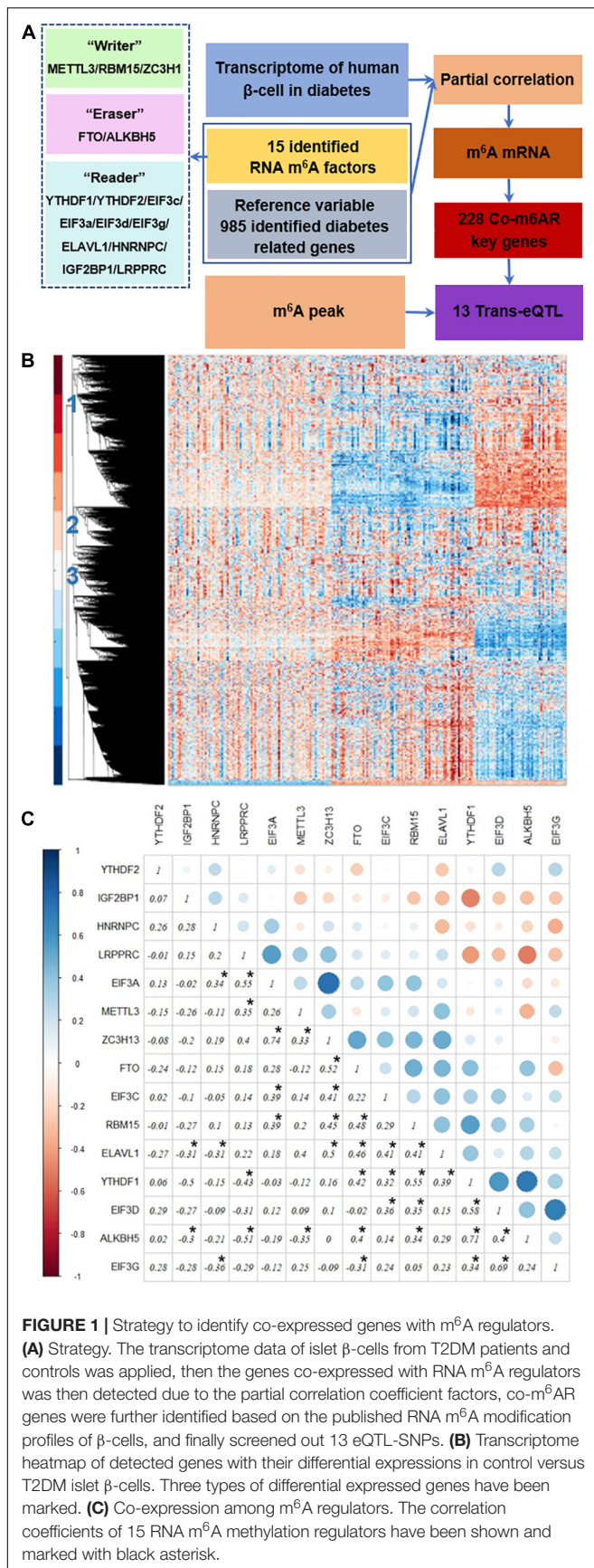


FIGURE 1 | Strategy to identify co-expressed genes with m⁶A regulators.

(A) Strategy. The transcriptome data of islet β -cells from T2DM patients and controls was applied, then the genes co-expressed with RNA m⁶A regulators was then detected due to the partial correlation coefficient factors, co-m⁶AR genes were further identified based on the published RNA m⁶A modification profiles of β -cells, and finally screened out 13 eQTL-SNPs. **(B)** Transcriptome heatmap of detected genes with their differential expressions in control versus T2DM islet β -cells. Three types of differentially expressed genes have been marked. **(C)** Co-expression among m⁶A regulators. The correlation coefficients of 15 RNA m⁶A methylation regulators have been shown and marked with black asterisk.

Co-expression Among RNA m⁶A Regulators in T2DM

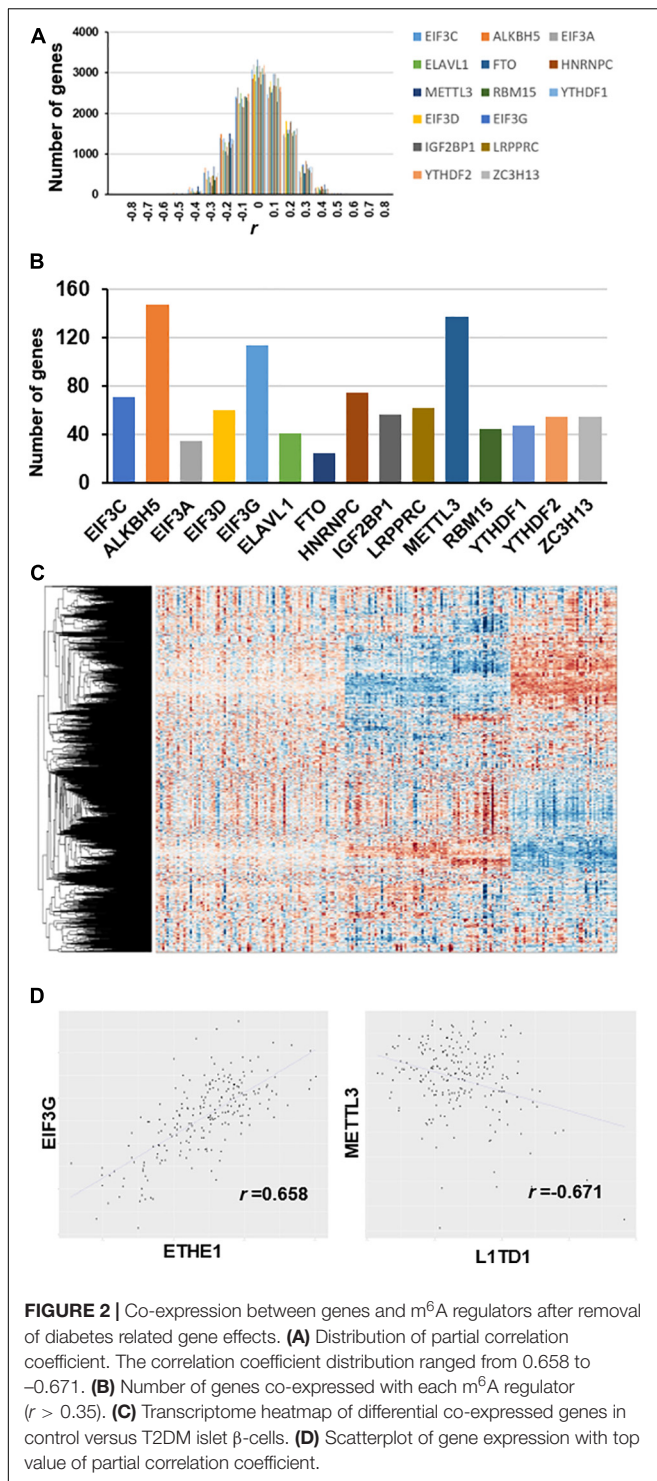
We downloaded the transcriptome datasets of human islet beta cells from 178 T2DM cases. These data had been obtained by gene chip detection, and 12261 genes were detected. The transcriptional heatmap in **Figure 1B** shows detected genes with their differential expressions and classified mainly into three types of expression patterns. We found 15 differentially-expressed genes among 20 RNA m⁶A methylation regulators, including *EIF3C*, *FTO*, *METTL3*, *RBM15*, etc; and 40 differentially-expressed genes among 46 diabetes related factors. We calculated the correlation coefficients of 15 RNA m⁶A methylation regulators by Pearson coefficient, and found a small correlation between them (**Figure 1C**). Among those, the correlation coefficient between *ZC3H13* and *EIF3a* was 0.74, that between *ALKBH5* and *YTHDF1* was 0.71, that between *EIF3G* and *EIF3D* was 0.69, that between *LRPPRC* and *ALKBH5* was -0.51, that between *LRPPRC* and *YTHDF1* was -0.43, and that between *EIF3G* and *HNRNPC* was -0.36.

Genes Co-expressed With RNA m⁶A Regulators (Co-m⁶AR)

We used partial correlation coefficients to calculate the association between genes and 15 RNA m⁶A regulators (**Supplementary Table 3**). In the global analysis, the correlation coefficient distribution approximated the normal distribution, ranging from 0.658 to -0.671 (**Figure 2A**). **Figure 2B** shows that the maximum correlation coefficient between *EIF3G* and *ETHE1* was 0.658, while the correlation coefficient between *METTL3* and *LITD1* was the smallest at 0.671. More than 3000 genes had partial correlation coefficients close to 0. We set the threshold value at 0.35; therefore, co-expressed genes have an absolute value of correlation coefficient greater than 0.35. Our results show that the number of genes associated with the expression of *ALKBH5* is the largest (approximately 150) followed by those associated with *METTL3* (approximately 140 genes), and those associated with *EIF3G* expression (approximately 115 genes) (**Figure 2C**). The number of genes associated with *FTO* expression was the smallest (approximately 30 genes). The heatmap shows the large differences in transcriptional profiles of these clustered genes (**Figure 2D**).

Co-m⁶AR Genes Are Enriched in T2DM-Associated Biological Processes

To acquire a more comprehensive and deep understanding of the identified differentially-expressed genes, we used the DAVID tool to analyze GO function and KEGG pathway enrichments (**Figure 3**). The 985 identified m⁶A-associated genes with high significant enrichment levels were enriched in diverse biological processes, such as metabolic process, *MAPK1/MAPK3* signaling, the T cell receptor signaling pathway, and *EGFR* signaling (**Figure 3B**).



Key Co-m⁶ARs Related to T2DM Revealed by Co-expression Network

We used the co-expression between differentially-expressed genes and the 15 m⁶A methylation factors to generate a network that showed most genes were associated with *METTL3*, *EIF3G*, and *ALKBH5* (Supplementary Figure 1). In Table 1, we show

some genes co-expressed with more than one m⁶A regulator. Ankyrin repeat domain 1 (*ANKRD1*) was co-expressed with *ALKBH5* and *METTL3*. Cyclin L2 (*CCNL2*) was co-expressed with *ALKBH5* and *EIF3G*. LIM domain binding 3 (*LDB3*) and Stromal interaction molecule 1 (*STIM1*) were both co-expressed with other m⁶A regulators. Casein Kinase 1 Alpha 1 Like (*CSNK1A1L*) was co-expressed with both *ALKBH* and *METTL3*, which play important role in the hsa04310:Wnt and the hsa04340:Hedgehog signaling pathways. Cysteine sulfinic acid decarboxylase (*CSAD*) was co-expressed with *ALKBH* and *EIF3C* (involved in hsa00430:Taurine and hypotaur metabolism). Cytochrome c oxidase subunit 5A (*COX5A*) was co-expressed with *ALKBH5* and other m⁶A regulators, which play critical roles in hsa00190:Oxidative phosphorylation, hsa01100:Metabolic pathways, hsa04260:Cardiac muscle contraction, hsa04932:Non-alcoholic fatty liver disease (NAFLD), hsa05010:Alzheimer's disease, hsa05012:Parkinson's disease, and hsa05016:Huntington's disease. *GRB2* associated binding protein 2 (*GAB2*) was co-expressed with *METTL3* and *EIF3G* (an important factor in the hsa04014:Ras and hsa04071:Sphingolipid signaling pathways). miR-let-7i was co-expressed with other m⁶A regulators involved in hsa05206:MicroRNAs in cancer.

mRNA of Co-m⁶AR Genes Have m⁶A Methylation in Islet β -Cells

We also identified 228 Co-m⁶ARs with m⁶A methylation based on the MeRIP-seq of human islet β -cells. As shown in Figure 4A, these genes play function in diverse biological processes (BP) such as signaling and stimulus responses. After the enrichment analysis, we found these genes to be enriched in EGFR, interleukin-2, MAPK1/MAPK3, and PDGF signaling, in the signaling cascades of insulin receptor and RAF/MAP kinase, in Epstein-Barr virus infections, and in mitochondrial translation initiation, VEGFA-VEGFR2, and SCF (Skp2)-mediated degradation pathways (Figure 4B).

Key Co-m⁶ARs With m⁶A Modification in Islet β -Cells Revealed by Co-expression Network

Our co-expression network revealed that most genes were associated with *METTL3*, *EIF3G*, *YTHDF1*, *YTHDF2* and *ALKBH5* (Supplementary Figure 2). In Table 1, we show some genes co-expressed with more than one m⁶A regulator. The gene ubiquitin like and ribosomal protein S30 fusion (*FAU*) was co-expressed with *METTL3* and *EIF3G* (a member of the hsa03010:Ribosome pathway). XPA binding protein 2 (*XAB2*) was co-expressed with *METTL3* and other m⁶A regulators involved in the hsa03040:Spliceosome pathway. We mentioned *GRB2* in this network above.

eQTL Localized in the m⁶A Peaks of Genes Co-expressed With m⁶A Regulators

The trans-eQTL were downloaded from eQTLGen database. We identified 13 eQTL-SNP localized in the m⁶A peak region

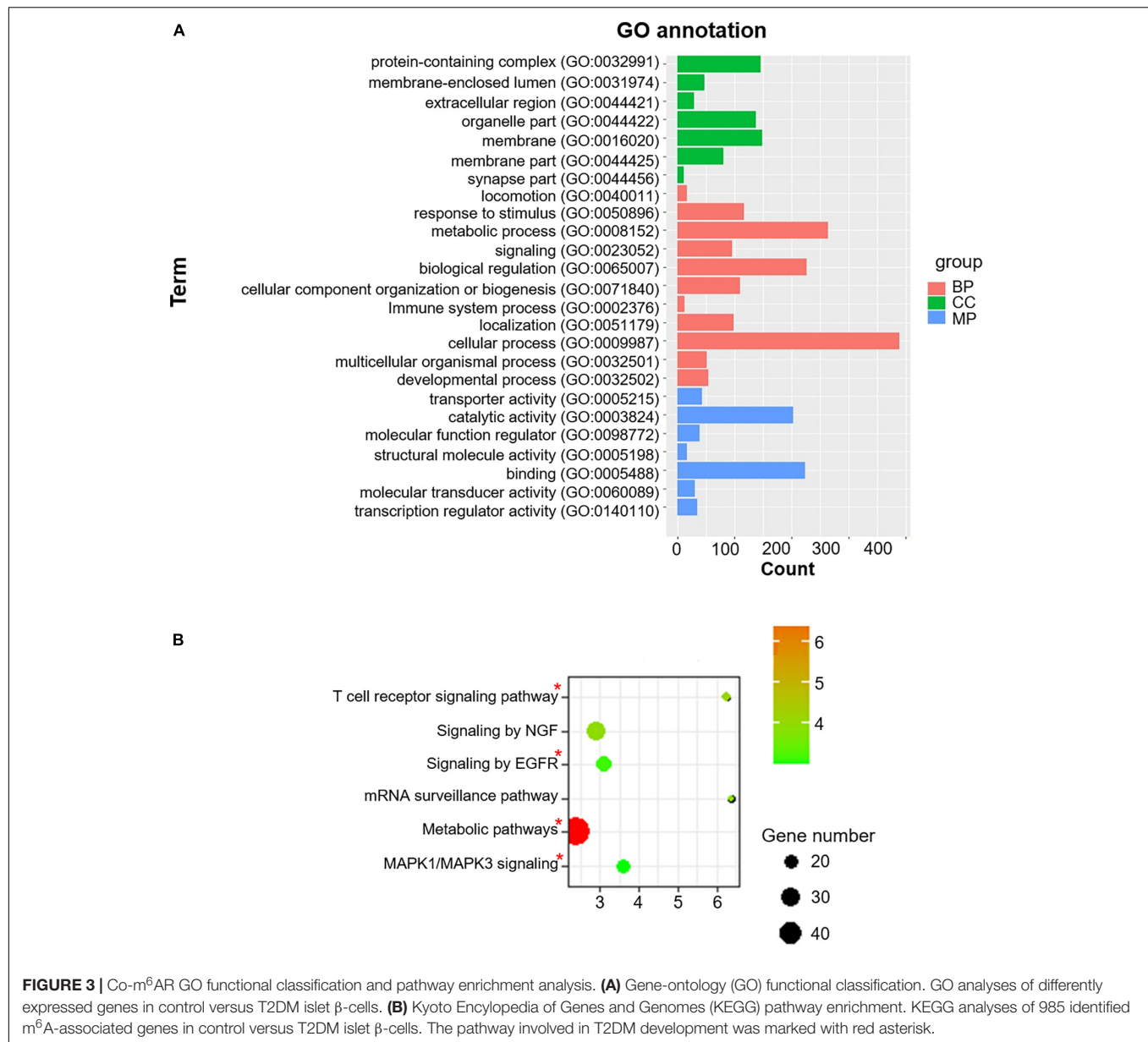


FIGURE 3 | Co-m⁶AR GO functional classification and pathway enrichment analysis. **(A)** Gene-ontology (GO) functional classification. GO analyses of differently expressed genes in control versus T2DM islet β -cells. **(B)** Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment. KEGG analyses of 985 identified m⁶A-associated genes in control versus T2DM islet β -cells. The pathway involved in T2DM development was marked with red asterisk.

(around 100 bp) which is acquired from the m⁶A-MeRIP-seq of human islet beta cells. In **Table 2**, there are three eQTL-SNPs localized in chromosome 1, two in chromosome 11, five in chromosome 19. Chromosome 15, 16, and 20 has only 1, respectively. They are belonging to 10 genes. *PHF13*, *SGTA* and *XAB2* contained 2 eQTL-SNP, respectively. Six eQTL-SNPs are involved in “A” allele. Four genes including *GAB2*, *LMNB2*, *XAB2* and *RBM39* are involved in the regulation of diabetes, fat and insulin signaling, according to text mining.

DISCUSSION

Type 2 diabetes mellitus is a polygenic metabolic disease with a pathogenesis in which altered gene expressions at different

levels play a crucial role. The analysis of gene expression regulation mechanisms in T2DM is helpful to determine potential therapeutic targets and provide new insights for diabetic therapies. Many m⁶A methylation modifications in eukaryotic mRNA participate in various biological processes by affecting mRNA splicing, translocation, degradation, and translation (Yue et al., 2015; Maity and Das, 2016). Thus, m⁶A can cumulatively regulate the expression of key regulatory genes through multiple effects during the diabetic pathogenic process.

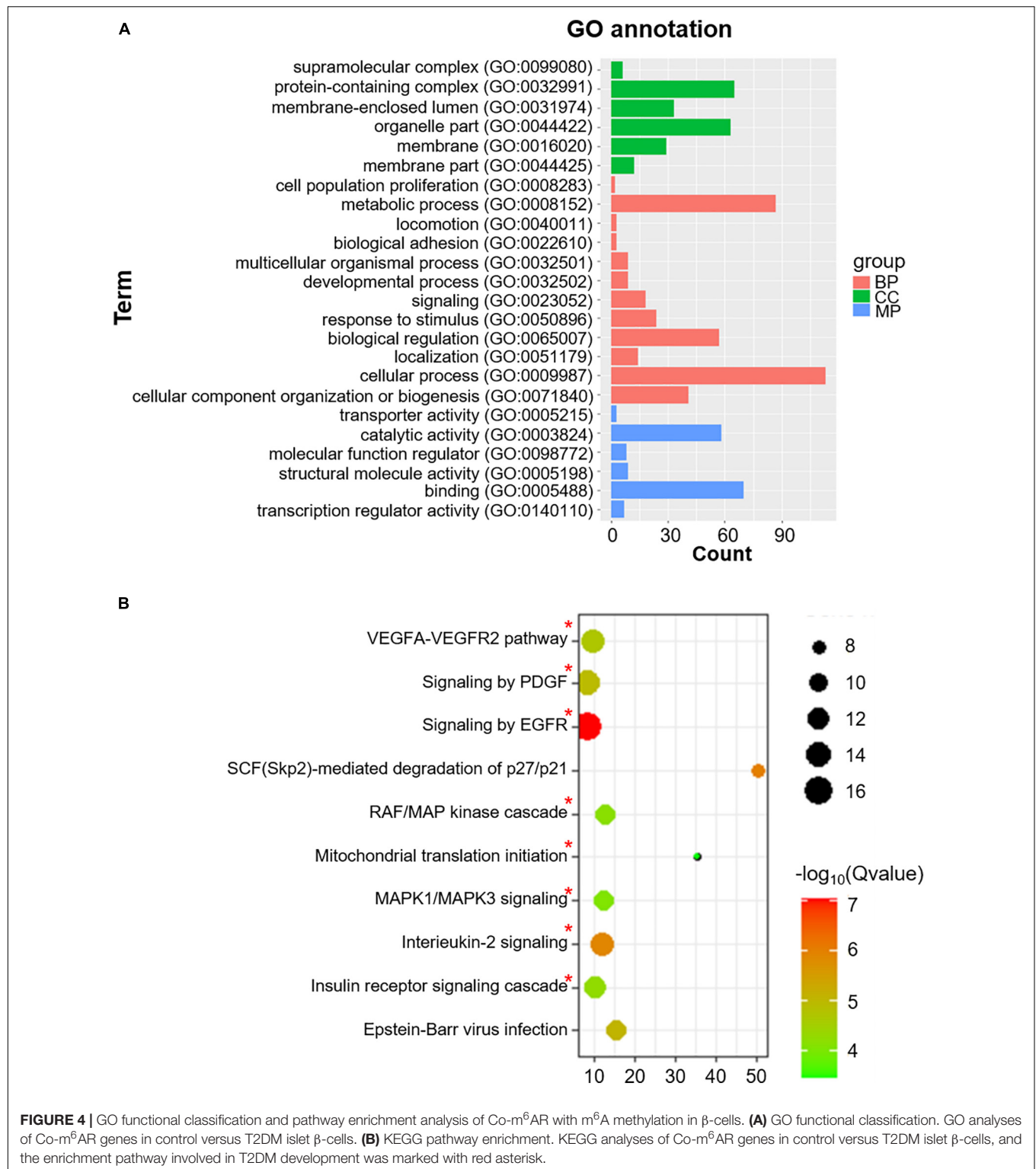
Among the genes corresponding to RNA-seq transcriptomes in pancreatic β -cells, we found 15 m⁶A regulators. Subsequently, we identified the expression of 985 genes significantly correlated with these m⁶A regulators. To investigate the function of these genes, we analyzed their associated pathways. Our GO enrichment analysis assigned the genes to a “metabolic

TABLE 1 | Typical genes co-expressed with m⁶A regulators in T2DM.

Symbol	ALKBH5	METTL3	EIF3G	EIF3C	Others	m ⁶ A	Diabetes	KEGG_PATHWAY
GOLGA8B	Yes	NA	Yes	Yes	NA	NA	NA	
ANKRD1	Yes	Yes	NA	NA	NA	NA	Yes	
BIVM	Yes	Yes	NA	NA	NA	NA	NA	
C19orf25	Yes	Yes	NA	NA	NA	NA	NA	
RBM26	Yes	Yes	NA	NA	NA	NA	NA	
ZFC3H1	Yes	Yes	NA	NA	NA	NA	NA	
CSNK1A1L	Yes	Yes	NA	NA	NA	NA	NA	hsa04310:Wnt signaling pathway, hsa04340:Hedgehog signaling pathway
AHSA2	Yes	NA	Yes	NA	NA	NA	NA	
AP1G2	Yes	NA	Yes	NA	NA	NA	NA	hsa04142:Lysosome
ARGLU1	Yes	NA	Yes	NA	NA	Yes	NA	
CCNL2	Yes	NA	Yes	NA	NA	NA	Yes	
MEG3	Yes	NA	Yes	NA	NA	Yes	NA	
PAN2	Yes	NA	Yes	NA	NA	NA	NA	hsa03018:RNA degradation
CSAD	Yes	NA	NA	Yes	NA	NA	Yes	hsa00430:Taurine and hypotaurine metabolism
COX5A	Yes	NA	NA	NA	Yes	NA	Yes	hsa00190:Oxidative phosphorylation, hsa04260:Cardiac muscle contraction, hsa05010:Alzheimer's disease, hsa05012:Parkinson's disease, hsa05016:Huntington's disease
LRRC47	Yes	NA	NA	NA	Yes	Yes	NA	
OR10A4	Yes	NA	NA	NA	Yes	NA	NA	hsa04740:Olfactory transduction
PRRG2	Yes	NA	NA	NA	Yes	NA	NA	
STIP1	Yes	NA	NA	NA	Yes	Yes	NA	hsa05020:Prion diseases
EZH1	NA	Yes	Yes	NA	NA	NA	NA	
FAU	NA	Yes	Yes	NA	NA	Yes	Yes	hsa03010:Ribosome
KIAA0319L	NA	Yes	Yes	NA	NA	Yes	NA	
PTGR2	NA	Yes	Yes	NA	NA	NA	NA	
SSSCA1	NA	Yes	Yes	NA	NA	NA	NA	
KRTAP3-2	NA	Yes	NA	Yes	NA	NA	NA	
NUP88	NA	Yes	NA	Yes	NA	Yes	NA	hsa03013:RNA transport
CEMIP	NA	Yes	NA	NA	Yes	NA	NA	
XAB2	NA	Yes	NA	NA	Yes	Yes	Yes	hsa03040:Spliceosome
CRYBB3	NA	NA	Yes	NA	Yes	NA	NA	
PTPRH	NA	NA	Yes	NA	Yes	NA	NA	
C15orf62	NA	NA	NA	Yes	Yes	NA	NA	
FBL	NA	NA	NA	Yes	Yes	Yes	NA	hsa03008:Ribosome biogenesis
SPTLC3	NA	NA	NA	Yes	Yes	NA	NA	hsa00600:Sphingolipid metabolism
TRIM22	NA	NA	NA	Yes	Yes	NA	NA	
MIRLET71	NA	NA	NA	NA	Yes	NA	Yes	hsa05206:MicroRNAs in cancer
C21orf90	NA	NA	NA	NA	Yes	NA	NA	
GAB2	NA	NA	NA	NA	Yes	Yes	Yes	hsa04014:Ras signaling pathway, hsa04071:Sphingolipid signaling pathway
GADD45GIP1	NA	NA	NA	NA	Yes	Yes	NA	
LDB3	NA	NA	NA	NA	Yes	NA	Yes	
PID1	NA	NA	NA	NA	Yes	NA	NA	
RSF1	NA	NA	NA	NA	Yes	Yes	NA	
STIM1	NA	NA	NA	NA	Yes	NA	Yes	hsa04020:Calcium signaling pathway, hsa04611:Platelet activation
TEX19	NA	NA	NA	NA	Yes	NA	NA	

pathway” characterized by metabolic abnormalities, abdominal obesity, hypertension, dyslipidemia, and hyperglycemia. Our results indicate that this pathway is closely associated with T2DM in agreement with studies in which metabolic

syndrome is a key etiological factor increasing the risk of T2DM (Hudish et al., 2019; Piche et al., 2020). In addition, other genes were enriched in processes confirmed to be also strongly associated with the pathogenesis and



progression of T2DM such as the “MAPK signaling,” “EGFR signaling,” and “T cell receptor signaling pathway.” The findings suggest that m⁶A methylation factors may be modulating the expression of genes in pancreatic β -cells of patients with T2DM.

In addition, we evaluated the above genes based on MeRIP-seq transcriptomes of human islet β -cells, and we further identified the remaining 228 genes co-expressed with m⁶A regulators. We analyzed their associated processes by GO annotations. They were enriched in “epidermal growth

TABLE 2 | eQTL localized in the m⁶A peaks of genes co-expressed with m⁶A regulators.

SNP	SNP Chr	SNPPos	Assessed allele	Other allele	Gene symbol	Bonferroni P	Reference related to diabetes
rs4908921	1	6613858	T	C	PHF13	0.00037375	NA
rs4908922	1	6613888	T	A	PHF13	0.00017326	NA
rs16833237	1	151404898	C	T	POGZ	9.25E-55	NA
rs7934912	11	78218556	A	T	GAB2	9.51E-26	Yes
rs148861080	11	119116390	T	G	C2CD2L	2.43E-08	NA
rs78810435	15	73116649	A	G	NEO1	8.12E-38	NA
rs138994570	16	57470988	A	G	POLR2C	2.43E-06	NA
rs1049910	19	2430637	G	C	LMNB2	3.15E-24	Yes
rs7009	19	2754792	A	G	SGTA	2.34E-19	NA
rs13282	19	2754810	T	C	SGTA	1.79E-05	NA
rs577145	19	7624377	T	C	XAB2	2.70E-18	Yes
rs541600	19	7624391	C	T	XAB2	3.14E-18	Yes
rs60223674	20	34303255	A	C	RBM39	1.11E-27	Yes

factor receptor (EGFR) signaling,” “insulin receptor signaling,” “MAPK1/MAPK3 signaling,” and “RAF/MAP kinase signaling.” Genetic interaction analyses of enhancers and protein-coding genes suggested that EGFR might be a novel susceptibility gene for T2DM (Yang et al., 2020). In the insulin receptor signaling pathway, PIK3R2 is an important regulatory subunit of PI3K/p85, and it can significantly suppress the activation of the PI3K/Akt pathway participating in the physiological and pathological diabetes processes (Cantley, 2002; Taniguchi et al., 2006). In the MAPK signaling pathway, MAPK3 encodes the extracellular signal regulated kinase 1 (*ERK1*), considered a crucial factor for cell proliferation, regulating the insulin gene expression and β -cell survival (Briaud et al., 2003). *ERK1* showed an increased expression in islet β -cells of diabetic mice (Kanda et al., 2009), and *ERK1* knockout mice are significantly resistant to HFD-induced obesity and insulin resistance as compared with control mice (Jager et al., 2011).

Furthermore, we suggest that the genes co-expressed with multiple m⁶A regulators may be critical factors in the regulation of diabetes by m⁶A. Therefore, we classified these genes and analyzed the published data to explore the potential association between these genes and diabetes. Genome-wide association studies revealed SNPs in *CCN1* and *COX5A* that are significantly associated with the risk of T2DM and a high total urine arsenic content, respectively (Wang et al., 2016; Grau-Perez et al., 2018). Transcriptome studies have found miR-let-7i downregulated in diabetes mice and controlling β -cells (Cheng et al., 2015). The evidence supports that CSAD is involvement in a fulminant type of diabetes (Kawabata et al., 2019). The expression of *STIM* expression was low in islets from T2DM pancreas and was associated with proinflammatory cytokines and palmitate (Kono et al., 2018). As we expected, these identified Co-m⁶AR genes are strongly associated with T2DM processes. However, how m⁶A methylation regulates these genes remains unclear.

The m⁶A alterations in mRNA change gene expression profiles and, thus, regulate the pathogenesis and development of various diseases, including tumors, nervous system diseases, and T2DM. We found *EGFR* mRNA to be affected by the expression level of m⁶A methylation factors. This is consistent

with previous findings on m⁶A-binding protein *YTHDF2* and modulating the location and stability of *EGFR* mRNA at its 3'UTR site, enhancing the degradation of the *EGFR* mRNA, and playing an anti-tumor role in hepatocellular carcinoma (Yu et al., 2019). *YTHDF2* has also been reported to promote the mRNA expression level of inflammatory response in LPS-stimulated monocyte macrophage cells, such as IL-6, IL-2, IL-1 β , and TNF- α (Yu et al., 2019). Another important gene, *ERK*, has also been identified to be regulated by m⁶A, as confirmed by the interaction between *ERK* and *METTL3* and *WTAP*. *METTL3* binds to *USP5* and active ERK-mediates phosphorylation, which makes the m⁶A methyltransferase complex more stable and results in a high level of m⁶A modifications (Prasun, 2020). Taken together, we suggest that the identified genes corresponding to diabetic m⁶A regulators may be crucial for different clinical characteristics of T2DM.

It can be seen that the identified genes are not only co-expressed with m⁶A regulators, but also have m⁶A methylation modifications. Subsequently, we screened these identified genes from the trans-eQTL database, and finally found 13 eQTL-SNPs localized in the m⁶A peak region. Among these eQTL-SNPs, four genes have been reported involved in diabetes. GRB2 associated binding protein 2 (*GAB2*) is an adaptor protein of the insulin receptor substrate 1 family and associated with the downstream signaling from cytokine receptors (Robinson and Oshlack, 2010). XPA binding protein 2 (*XAB2*) exerts as a regulator in hyperglycemia with chronic insulin (Lim et al., 2008). Lamin B2 (*LMNB2*) can regulate fasting blood glucose by increasing insulin secretion or regenerating beta cells (de Toledo et al., 2020). RNA binding motif protein 39 (*RBM39*) were associated with a higher risk of insulin in clinically significant retinopathy of prematurity (Lynch et al., 2016). Obviously, these genes will be of value for studying the prognosis and diagnosis of T2DM, however, the underlying mechanisms of the identified genes in T2DM remain to be elucidated.

In summary, we identified genes co-expressed with m⁶A regulators in human T2DM islets; some of them contained known m⁶A methylation modifications. These genes are enriched in T2DM-related biological processes. These promising genes

provide a novel insight into the progression of T2DM and need to be confirmed in further. The identified transcripts may give helpful information for understanding the effects of m⁶A methylation in the prognosis and diagnosis of T2DM.

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

CC and JT designed the project and wrote the manuscript. SL and MY collected the database. QX and WL did the experiments and analyzed the data. All authors contributed to the article and approved the submitted version.

REFERENCES

- Dayeh, T., and Ling, C. (2015). Does epigenetic dysregulation of pancreatic islets contribute to impaired insulin secretion and type 2 diabetes? *Biochem Cell Biol.* 93, 511–521. doi: 10.1139/bcb-2015-0057
- Reusch, J. E., and Manson, J. E. (2017). Management of Type 2 Diabetes in 2017: Getting to Goal. *JAMA.* 317, 1015–1016. doi: 10.1001/jama.2017.0241
- Giacco, F., and Brownlee, M. (2010). Oxidative stress and diabetic complications. *Circ Res.* 107, 1058–1070.
- Singh, R., Barden, A., Mori, T., and Beilin, L. (2001). Advanced glycation end-products: a review. *Diabetologia.* 44, 129–146.
- Copps, K. D., Hancer, N. J., Qiu, W., and White, M. F. (2016). Serine 302 Phosphorylation of Mouse Insulin Receptor Substrate 1 (IRS1) Is Dispensable for Normal Insulin Signaling and Feedback Regulation by Hepatic S6 Kinase. *J Biol Chem.* 291, 8602–8617. doi: 10.1074/jbc.m116.714915
- He, C. (2010). Grand challenge commentary: RNA epigenetics? *Nat Chem Biol.* 6, 863–865. doi: 10.1038/nchembio.482
- Meyer, K. D., Saletore, Y., Zumbo, P., Elemento, O., Mason, C. E., and Jaffrey, S. R. (2012). Comprehensive analysis of mRNA methylation reveals enrichment in 3' UTRs and near stop codons. *Cell.* 149, 1635–1646. doi: 10.1016/j.cell.2012.05.003
- Dominissini, D., Moshitch-Moshkovitz, S., Schwartz, S., Salmon-Divon, M., Ungar, L., Osenberg, S., et al. (2012). Topology of the human and mouse m⁶A RNA methylomes revealed by m⁶A-seq. *Nature.* 485, 201–206. doi: 10.1038/nature11112
- Wang, X., Lu, Z., Gomez, A., Hon, G. C., Yue, Y., Han, D., et al. (2014). N6-methyladenosine-dependent regulation of messenger RNA stability. *Nature.* 505, 117–120. doi: 10.1038/nature12730
- Liu, N., Dai, Q., Zheng, G., He, C., Parisien, M., and Pan, T. (2015). N(6)-methyladenosine-dependent RNA structural switches regulate RNA-protein interactions. *Nature.* 518, 560–564. doi: 10.1038/nature14234
- Bokar, J. A., Shambaugh, M. E., Polayes, D., Matera, A. G., and Rottman, F. M. (1997). Purification and cDNA cloning of the AdoMet-binding subunit of the human mRNA (N6-adenosine)-methyltransferase. *RNA.* 3, 1233–1247.
- Liu, J., Yue, Y., Han, D., Wang, X., Fu, Y., Zhang, L., et al. (2014). A METTL3-METTL14 complex mediates mammalian nuclear RNA N6-adenosine methylation. *Nat Chem Biol.* 10, 93–95. doi: 10.1038/nchembio.1432
- Jia, G., Fu, Y., Zhao, X., Dai, Q., Zheng, G., Yang, Y., et al. (2011). N6-methyladenosine in nuclear RNA is a major substrate of the obesity-associated FTO. *Nat Chem Biol.* 7, 885–887. doi: 10.1038/nchembio.687

FUNDING

This study was supported by grants from the Scientific Research Foundation for the High-level Talents, Fujian University of Traditional Chinese Medicine (X2020001-talents).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fcell.2021.651142/full#supplementary-material>

Supplementary Figure 1 | Co-m⁶AR co-expression network. Functional protein-protein interaction network of Co-m⁶AR genes in control versus T2DM islet β -cells was shown, the key genes were marked with green asterisk.

Supplementary Figure 2 | Co-expression network of Co-m⁶AR with m⁶A methylation in β -cells. Functional protein-protein interaction network of differently expressed genes with *METTL3*, *EIF3G*, *YTHDF1*, *YTHDF2*, and *ALKBH5* was shown, and these five RNA m⁶A methylation regulators were marked with blue asterisk.

- Zheng, G., Dahl, J. A., Niu, Y., Fedorcsak, P., Huang, C. M., Li, C. J., et al. (2013). ALKBH5 is a mammalian RNA demethylase that impacts RNA metabolism and mouse fertility. *Mol Cell.* 49, 18–29. doi: 10.1016/j.molcel.2012.10.015
- Fedeles, B. I., Singh, V., Delaney, J. C., Li, D., and Essigmann, J. M. (2015). The AlkB Family of Fe(II)/alpha-Ketoglutarate-dependent Dioxxygenases: Repairing Nucleic Acid Alkylation Damage and Beyond. *J Biol Chem.* 290, 20734–20742. doi: 10.1074/jbc.r115.656462
- Taneera, J., Prasad, R. B., Dhaiban, S., Mohammed, A. K., Haataja, L., Arvan, P., et al. (2018). Silencing of the FTO gene inhibits insulin secretion: An in vitro study using GRINCH cells. *Mol Cell Endocrinol.* 472, 10–17. doi: 10.1016/j.mce.2018.06.003
- Scuteri, A., Sanna, S., Chen, W. M., Uda, M., Albai, G., Strait, J., et al. (2007). Genome-wide association scan shows genetic variants in the FTO gene are associated with obesity-related traits. *PLoS Genet.* 3:e115. doi: 10.1371/journal.pgen.0030115
- Blauth, W., and Falliner, A. (1986). [Morphology and classification of cleft hands]. *Handchir Mikrochir Plast Chir.* 18, 161–195.
- Boissel, S., Reish, O., Proulx, K., Kawagoe-Takaki, H., Sedgwick, B., Yeo, G. S., et al. (2009). Loss-of-function mutation in the dioxxygenase-encoding FTO gene causes severe growth retardation and multiple malformations. *Am J Hum Genet.* 85, 106–111. doi: 10.1016/j.ajhg.2009.06.002
- Church, C., Moir, L., McMurray, F., Girard, C., Banks, G. T., Teboul, L., et al. (2010). Overexpression of Fto leads to increased food intake and results in obesity. *Nat Genet.* 42, 1086–1092. doi: 10.1038/ng.713
- Lee, M., Kim, B., and Kim, V. N. (2014). Emerging roles of RNA modification: m(6)A and U-tail. *Cell.* 158, 980–987. doi: 10.1016/j.cell.2014.08.005
- Christian, K. M., Song, H., and Ming, G. L. (2014). Functions and dysfunctions of adult hippocampal neurogenesis. *Annu Rev Neurosci.* 37, 243–262. doi: 10.1146/annurev-neuro-071013-014134
- De Jesus, D. F., Zhang, Z., Kahraman, S., Brown, N. K., Chen, M., Hu, J., et al. (2019). m(6)A mRNA Methylation Regulates Human beta-Cell Biology in Physiological States and in Type 2 Diabetes. *Nat Metab.* 1, 765–774. doi: 10.1038/s42255-019-0089-9
- Xie, W., Ma, L. L., Xu, Y. Q., Wang, B. H., and Li, S. M. (2019). METTL3 inhibits hepatic insulin sensitivity via N6-methyladenosine modification of Fasn mRNA and promoting fatty acid metabolism. *Biochem Biophys Res Commun.* 518, 120–126. doi: 10.1016/j.bbrc.2019.08.018
- Li, Y., Zhang, Q., Cui, G., Zhao, F., Tian, X., Sun, B. F., et al. (2020). m(6)A Regulates Liver Metabolic Disorders and Hepatogenous Diabetes. *Genomics Proteomics Bioinformatics.* **vol page,

- Song, Y., Wang, Q., Li, L., Chen, S., Zhao, Y., and Gao, L. (2020). Comprehensive epigenetic analysis of m⁶A modification in the hippocampal injury of diabetic rats. *Epigenomics*. 12, 1811–1824. doi: 10.2217/epi-2020-0125
- Robinson, M. D., and Oshlack, A. (2010). A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol.* 11, R25.
- Kanehisa, M. (2002). The KEGG database. *Novartis Found Symp* 247, 244–252.
- Huang da, W., Sherman, B. T., and Lempicki, R. A. (2009b). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc.* 4, 44–57. doi: 10.1038/nprot.2008.211
- Huang da, W., Sherman, B. T., and Lempicki, R. A. (2009a). Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res.* 37, 1–13. doi: 10.1093/nar/gkn923
- Gautier, L., Cope, L., Bolstad, B. M., and Irizarry, R. A. (2004). affy-analysis of Affymetrix GeneChip data at the probe level. *Bioinformatics*. 20, 307–315. doi: 10.1093/bioinformatics/btg405
- Mi, H., Muruganujan, A., and Thomas, P. D. (2013). PANTHER in 2013: modeling the evolution of gene function, and other gene attributes, in the context of phylogenetic trees. *Nucleic Acids Res.* 41, D377–D386.
- Thomas, P. D., Campbell, M. J., Kejariwal, A., Mi, H., Karlak, B., Daverman, R., et al. (2003). PANTHER: a library of protein families and subfamilies indexed by function. *Genome Res.* 13, 2129–2141. doi: 10.1101/gr.772403
- Yue, Y., Liu, J., and He, C. (2015). RNA N⁶-methyladenosine methylation in post-transcriptional gene expression regulation. *Genes Dev.* 29, 1343–1355. doi: 10.1101/gad.262766.115
- Maity, A., and Das, B. (2016). N⁶-methyladenosine modification in mRNA: machinery, function and implications for health and diseases. *FEBS J.* 283, 1607–1630. doi: 10.1111/febs.13614
- Piche, M. E., Tchernof, A., and Despres, J. P. (2020). Obesity Phenotypes, Diabetes, and Cardiovascular Diseases. *Circ Res.* 126, 1477–1500. doi: 10.1161/circresaha.120.316101
- Hudish, L. I., Reusch, J. E., and Sussel, L. (2019). beta Cell dysfunction during progression of metabolic syndrome to type 2 diabetes. *J Clin Invest.* 129, 4001–4008. doi: 10.1172/jci.129188
- Yang, Y., Yao, S., Ding, J. M., Chen, W., and Guo, Y. (2020). Enhancer-Genes Interaction Analyses Identified the Epidermal Growth Factor Receptor as a Susceptibility Gene for Type 2 Diabetes Mellitus. *Diabetes Metab J* 45, 241–250. doi: 10.4093/dmj.2019.0204
- Taniguchi, C. M., Kondo, T., Sajan, M., Luo, J., Bronson, R., Asano, T., et al. (2006). Divergent regulation of hepatic glucose and lipid metabolism by phosphoinositide 3-kinase via Akt and PKC λ /zeta. *Cell Metab.* 3, 343–353. doi: 10.1016/j.cmet.2006.04.005
- Cantley, L. C. (2002). The phosphoinositide 3-kinase pathway. *Science*. 296, 1655–1657.
- Briaud, I., Lingohr, M. K., Dickson, L. M., Wrede, C. E., and Rhodes, C. J. (2003). Differential activation mechanisms of Erk-1/2 and p70(S6K) by glucose in pancreatic beta-cells. *Diabetes*. 52, 974–983. doi: 10.2337/diabetes.52.4.974
- Kanda, Y., Shimoda, M., Tawaramoto, K., Hamamoto, S., Tatsumi, F., Kawasaki, F., et al. (2009). Molecular analysis of db gene-related pancreatic beta cell dysfunction; evidence for a compensatory mechanism inhibiting development of diabetes in the db gene heterozygote. *Endocr J.* 56, 997–1008. doi: 10.1507/endocrj.k09e-028
- Jager, J., Corcelle, V., Gremeaux, T., Laurent, K., Waget, A., Pages, G., et al. (2011). Deficiency in the extracellular signal-regulated kinase 1 (ERK1) protects leptin-deficient mice from insulin resistance without affecting obesity. *Diabetologia*. 54, 180–189. doi: 10.1007/s00125-010-1944-0
- Wang, T., Huang, T., Li, Y., Zheng, Y., Manson, J. E., Hu, F. B., et al. (2016). Low birthweight and risk of type 2 diabetes: a Mendelian randomisation study. *Diabetologia*. 59, 1920–1927. doi: 10.1007/s00125-016-4019-z
- Grau-Perez, M., Navas-Acien, A., Galan-Chilet, I., Briangon-Figuero, L. S., Morchon-Simon, D., Bermudez, J. D., et al. (2018). Arsenic exposure, diabetes-related genes and diabetes prevalence in a general population from Spain. *Environ Pollut.* 235, 948–955. doi: 10.1016/j.envpol.2018.01.008
- Cheng, C., Kobayashi, M., Martinez, J. A., Ng, H., Moser, J. J., Wang, X., et al. (2015). Evidence for Epigenetic Regulation of Gene Expression and Function in Chronic Experimental Diabetic Neuropathy. *J Neuropathol Exp Neurol.* 74, 804–817. doi: 10.1097/nen.0000000000000219
- Kawabata, Y., Nishida, N., Awata, T., Kawasaki, E., Imagawa, A., Shimada, A., et al. (2019). Genome-Wide Association Study Confirming a Strong Effect of HLA and Identifying Variants in CSAD/Inc-ITGB7-1 on Chromosome 12q13.13 Associated With Susceptibility to Fulminant Type 1 Diabetes. *Diabetes*. 68, 665–675. doi: 10.2337/db18-0314
- Kono, T., Tong, X., Taleb, S., Bone, R. N., Iida, H., Lee, C. C., et al. (2018). Impaired Store-Operated Calcium Entry and STIM1 Loss Lead to Reduced Insulin Secretion and Increased Endoplasmic Reticulum Stress in the Diabetic beta-Cell. *Diabetes*. 67, 2293–2304. doi: 10.2337/db17-1351
- Yu, R., Li, Q., Feng, Z., Cai, L., and Xu, Q. (2019). m⁶A Reader YTHDF2 Regulates LPS-Induced Inflammatory Response. *Int J Mol Sci.* 20, 1323. doi: 10.3390/ijms20061323
- Prasun, P. (2020). Mitochondrial dysfunction in metabolic syndrome. *Biochim Biophys Acta Mol Basis Dis.* 1866, 165838.
- Lim, J. M., Sherling, D., Teo, C. F., Hausman, D. B., Lin, D., and Wells, L. (2008). Defining the regulated secreted proteome of rodent adipocytes upon the induction of insulin resistance. *J Proteome Res.* 7, 1251–1263. doi: 10.1021/pr7006945
- de Toledo, M., Lopez-Mejia, I. C., Cavelier, P., Pratlong, M., Barrachina, C., Gromada, X., et al. (2020). Lamin C Counteracts Glucose Intolerance in Aging, Obesity, and Diabetes Through beta-Cell Adaptation. *Diabetes*. 69, 647–660. doi: 10.2337/db19-0377
- Lynch, A. M., Wagner, B. D., Mandava, N., Palestine, A. G., Mourani, P. M., McCourt, E. A., et al. (2016). The Relationship of Novel Plasma Proteins in the Early Neonatal Period With Retinopathy of Prematurity. *Invest Ophthalmol Vis Sci.* 57, 5076–5082. doi: 10.1167/iovs.16-19653

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 Chen, Xiang, Liu, Liang, Yang and Tao. 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.



Role of RNA N6-Methyladenosine Modification in Male Infertility and Genital System Tumors

Shuai Liu^{1,2††}, Yongfeng Lao^{1,2††}, Yanan Wang^{1,2†}, Rongxin Li^{1,2}, Xuefeng Fang³,
Yunchang Wang^{4,5}, Xiaolong Gao³ and Zhilong Dong^{1,2*}

¹ Department of Urology, Lanzhou University Second Hospital, Lanzhou, China, ² Gansu Nephro-Urological Clinical Center, Institute of Urology, Department of Urology, Key Laboratory of Urological Disease of Gansu Province, Lanzhou University Second Hospital, Lanzhou, China, ³ Department of Urology, People's Hospital of Jinchang, Jinchang, China, ⁴ Second Clinical Medical College, Lanzhou University, Lanzhou, China, ⁵ Xiangya Hospital, Central South University, Changsha, China

OPEN ACCESS

Edited by:

Jia Meng,
Xi'an Jiaotong-Liverpool University,
China

Reviewed by:

Qingyou Liu,
Guangxi University, China
Hui Li,
Guangxi University, China

*Correspondence:

Zhilong Dong
562691313@qq.com

† Present address:

Shuai Liu,
The Second Hospital of Lanzhou
University, Lanzhou City, China
Yongfeng Lao,
The Second Hospital of Lanzhou
University, Lanzhou City, China

† These authors have contributed
equally to this work

Specialty section:

This article was submitted to
Epigenomics and Epigenetics,
a section of the journal
Frontiers in Cell and Developmental
Biology

Received: 05 March 2021

Accepted: 16 April 2021

Published: 19 May 2021

Citation:

Liu S, Lao Y, Wang Y, Li R, Fang X,
Wang Y, Gao X and Dong Z (2021)
Role of RNA N6-Methyladenosine
Modification in Male Infertility
and Genital System Tumors.
Front. Cell Dev. Biol. 9:676364.
doi: 10.3389/fcell.2021.676364

Epigenetic alterations, particularly RNA methylation, play a crucial role in many types of disease development and progression. Among them, N6-methyladenosine (m6A) is the most common epigenetic RNA modification, and its important roles are not only related to the occurrence, progression, and aggressiveness of tumors but also affect the progression of many non-tumor diseases. The biological effects of RNA m6A modification are dynamically and reversibly regulated by methyltransferases (writers), demethylases (erasers), and m6A binding proteins (readers). This review summarized the current finding of the RNA m6A modification regulators in male infertility and genital system tumors and discussed the role and potential clinical application of the RNA m6A modification in spermatogenesis and male genital system tumors.

Keywords: N6-methyladenosine, RNA methylation, spermatogenesis, male fertility, genital system tumors

INTRODUCTION

Over the last decades, with the rapid development of gene detection technology, epigenetic modification of diseases has become a focus of clinical research. Epigenetics is a branch of genetics that studies heritable and reversible phenotypes of changes in gene expression without alterations in nuclear DNA sequences (Mohammad et al., 2019). Epigenetic processes, including DNA methylation, histone modifications, chromatin rearrangement, and RNA modifications, play a crucial role in the regulation of many physiological and pathological processes, such as embryonic development (Mendel et al., 2018), nervous system development (Li M. et al., 2018), and tumorigenesis (Chen S. et al., 2020; Sun Y. et al., 2020). Therefore, the regulation of these epigenetic processes may be a potential therapeutic intervention. m6A is one of the most common types of eukaryotes in RNA modification (Dominissini et al., 2012; Meyer et al., 2012). Since the pioneering research in the 1970s (Desrosiers et al., 1974; Perry et al., 1975), with the identification of more m6A-related enzymes, the important biological functions played by m6A modification have been gradually revealed around about half a century later.

Since the pioneering research in the 1970s (Desrosiers et al., 1974; Perry et al., 1975), with the identification of more M6A-related enzymes, the important biological functions played by m6A modification have been gradually revealed around about half a century later. With the first m6A demethylase—FTO was found in 2011, more m6A regulators have been found in the past decade which could be divided into three kinds of regulator proteins: methyltransferases (“writer”), such as methyltransferase-like protein 3 (METTL3), methyltransferase-like protein 14 (METTL14), Wilms’ tumor 1-associating protein (WATP) (Liu et al., 2014; Ping et al., 2014; Li Z. et al., 2018), Vir-like m6A methyltransferase-associated (VIRMA; also known as KIAA1429) (Yue et al., 2018);

demethylases (“erasers”), such as obesity-associated protein (FTO) and AlkB family homolog 5 (ALKBH5) (Jia et al., 2011; Zheng et al., 2013); and m6A binding proteins, such as the YTH domain family proteins (YTHDFs) and YTH domain-containing protein 1-2 (YTHDC1-2) (Dominissini et al., 2012; Wang et al., 2014, 2015), The insulin-like growth factor 2 mRNA binding proteins (IGF2BPs) (Huang et al., 2018), heterogeneous nuclear ribonucleoprotein A2B1 (HNRNPA2B1) (Alarcón et al., 2015), and eukaryotic translation initiation factor 3 (eIF3) (Meyer et al., 2015).

With changes in environmental pollution and living habits, the incidence and prevalence of male infertility and andrology tumors increase steadily worldwide and have become a prevalent worldwide problem in recent decades. About 2-15% of couples worldwide are reported to suffer from infertility, of which nearly half are caused by male reproduction (Ostermeier et al., 2002; Ji et al., 2012), a large proportion of male infertility was related to defective spermatogenesis, such as abnormal sperm count, morphology, and function (Guzick et al., 2001). Female infertility can usually be induced by hormone therapy to stimulate the production of oocytes, while male infertility is more difficult to treat due to various reasons for sperm abnormalities (Tournaye et al., 2017). Surgery is the primary treatment option for men with obstructive azoospermia and part of non-obstructive infertility such as varicocele and undescended testicles at birth (Dubin and Amelar, 1971; Verza and Esteves, 2008). Recently, intracytoplasmic sperm injection (ICSI) is a relatively novel approach to assisted reproduction that has brought revolutionary changes in clinical therapy for infertility (Kyono et al., 2001). But there are still a large proportion of patients who cannot meet the wish of fatherhood. Therefore, scientists gradually realize that it is very important to study the causes of male infertility.

Another disease that severely impairs male health and quality of life is genital system tumors. Prostate cancer (PCa) is the most common malignant tumor of the male genital system and its pathogenesis involves multiple factors, which mainly affects elderly men (Siegel et al., 2021). Although the mortality rate of prostate cancer is not high (Tsubokura et al., 2018), due to the molecular mechanisms underlying the development and progression of prostate cancer remain unclear as well as its high incidence, it is worth studying and its treatment is of great significance for the longevity of the elderly. Additionally, the most common reproductive system tumors in men aged 15-35 years old are testicular germ cell tumors with increasing incidence in recent years (Ghazarian et al., 2017). Most testicular germ cell tumors (GCTs) are completely curable with cisplatin-based chemotherapy. However, a small proportion of patients who cannot be cured with cisplatin chemotherapy should be the focus of our research. Therefore, greater insights into the mechanisms regulating spermatogenesis and male genital system tumors will help us found novel molecular targets to develop more effective treatment strategies for the disease. In this review, we have provided an overview of the underlying m6A methylation and the mechanisms of development of spermatogenesis and andrology tumors. We further highlight the potential use of m6A methylation modification as a

biomarker for andrology diseases risk diagnosis, therapeutic applications, and prognosis.

RNA M6A MODIFICATION

The important biological functions played by mRNA m6A dynamic and reversible modification are mainly catalyzed by m6A related enzymes. A large number of studies have shown that the m6A methylation of mRNA modification was catalyzed by three elements (**Figure 1**).

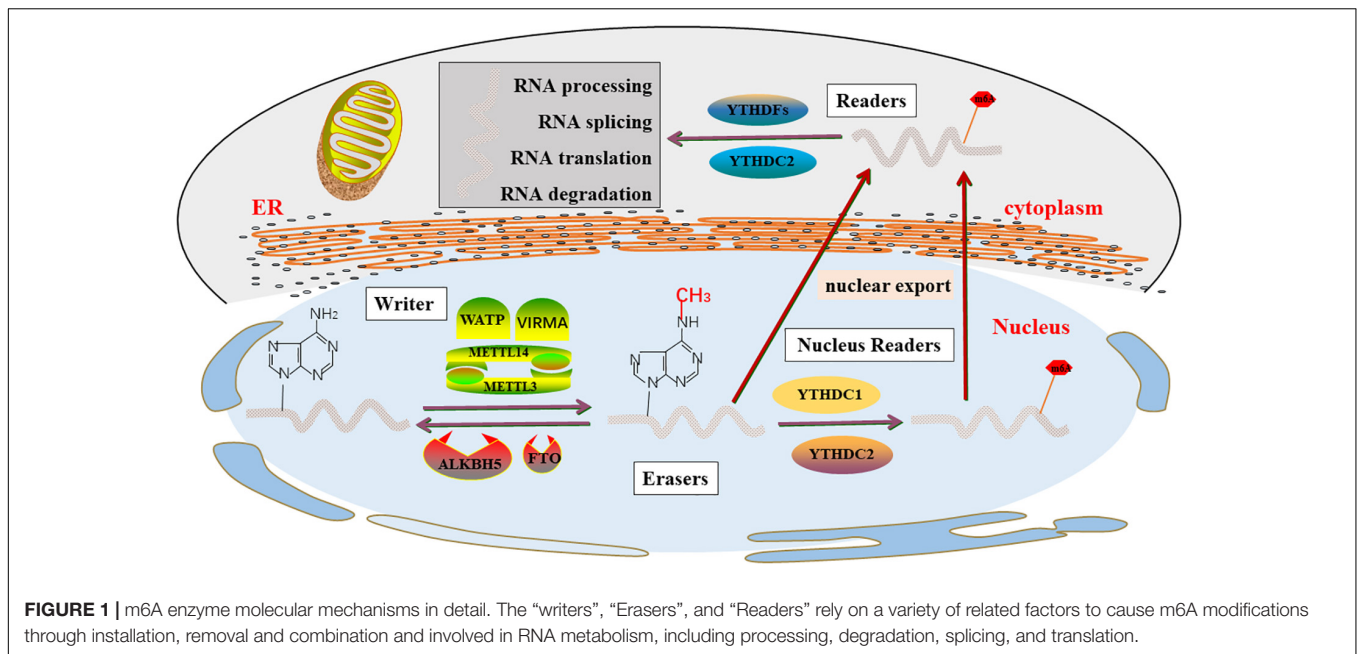
Methyltransferases (“Writers”)

The methylation process is first controlled by a family of enzymes called methyltransferases, also known as a writer, which are proteins that induce specific RNA methylation, which catalyzes the formation of m6A levels. In 1997, METTL3 was first shown to be m6A methylation transferase, and the expression of METTL3 can direct effects the total methylation level of m6A, which has effects on mRNA stability, leading to dysregulated cellular functions (Bokar et al., 1997). Several studies (Śledź and Jinek, 2016; Wang P. et al., 2016; Wang X. et al., 2016) suggest that mRNAs methyltransferases are primarily formed by the METTL3-METTL14 heterodimer, of which METTL3 is a catalytic active subunit and METTL14 plays a key structural role in substrate recognition, maintaining complex integrity, and substrate binding.

As rapid and sensitive detection technology advances, more methyltransferases were found. METTL5 and METTL16 is a methyltransferase that targets pre-mRNAs and various non-coding RNAs, such as rRNA, snRNA (Pendleton et al., 2017; Warda et al., 2017; Wojtas et al., 2017; van Tran et al., 2019; Ignatova et al., 2020). WATP is another m6A “writer” complex, which stabilizes the core complex and targets the METTL3 and METTL14 complexes to their substrates (Ping et al., 2014). VIRMA prioritizes gene methylation modifications near the 3'-UTR and stop codon regions and recruits the m6A complex to specific RNA sites (Yue et al., 2018). ZC3H13 anchors WTAP to the nucleus to promote m6A methylation and regulation (Wen et al., 2018). Recently, another CCHC-containing zinc finger protein, ZCCHC4, has been identified as a novel methyltransferase, which is involved in the modification of 28S rRNA and mediates the subunit distribution and global translation of rRNA ribosomes (Ma et al., 2019; Ren W. et al., 2019; Pinto et al., 2020).

Demethylases (“Erasers”)

RNA m6A modification was mainly removed by demethylases FTO and ALKBH5, which were an essential enzyme in m6A modification so that maintained m6A modification in a dynamic balance, and dynamically regulated developmental and disease processes (Jia et al., 2011; Zhang X. et al., 2020). FTO, originally known as an obesity-susceptibility gene, is strongly associated with obesity risk (Dina et al., 2007; Frayling et al., 2007; Scuteri et al., 2007). A subsequent study (Jia et al., 2011) demonstrated that FTO is the first m6A demethylase of eukaryotic mRNA and the role of FTO in adipogenesis and tumorigenesis is



related to its m6A demethylase activity. The majority of studies (Söderberg et al., 2009; Castillo et al., 2012; Iles et al., 2013; Li et al., 2017) have shown a strong association between FTO and an increased risk of various types of cancer, including breast cancer, prostate cancer, kidney cancer, endometrial cancer, pancreatic cancers, lymphoma, and leukemia. These studies reveal the roles and underlying molecular mechanisms of FTO in cancer pathogenesis. Therefore, the development of selective and effective inhibitors targeting FTO will have the potential to treat cancer, particularly in combination with other therapies to treat cancers that are resistant to currently available therapies. ALKBH5 is another m6A eraser that is localized in the nucleus, which is most highly expressed in the testicles. Therefore, it may be necessary for mouse spermatogenesis and fertility (Zheng et al., 2013; Tang et al., 2018). Available pieces of evidence indicate that ALKBH5 mainly inhibits the development of a variety of cancers.

Taken together, as m6A demethylases, FTO and ALKBH5 have opposite effects in some tumors, and the specific reasons are worth further study. In addition, recent studies have shown that ALKB family homolog 3 (ALKBH3) may be a novel demethylase modified by m6A, and modified mammalian tRNA demethylation promotes protein synthesis in cancer cells (Ueda et al., 2017).

m6A Binding Proteins (“Readers”)

The effect of m6A modification is primarily dependent on downstream RNA-binding proteins, known as m6A “readers”, which prioritize the recognition of m6A modified RNA and combine m6A methylation with RNA processing and biological functions (Li F. et al., 2014; Zhu et al., 2014). Recent studies have shown that m6A modification regulates most RNA processing steps, including mRNA translation (Tanabe et al., 2016; Shi et al., 2017; Mao et al., 2019), mRNA splicing (Hartmann et al., 1999;

Xiao et al., 2016; Kasowitz et al., 2018; Luxton et al., 2019), stability (Du et al., 2016), and transport (Roundtree et al., 2017). In the known m6A readers, most proteins contain a YTH domain that specifically recognizes m6A and A (Stoilov et al., 2002), including YTHDC1 in the nucleus (Hartmann et al., 1999), YTHDFs family in the cytoplasm (Wang et al., 2014), and YTHDC2 in the nucleus and cytoplasm (Wojtas et al., 2017).

YTHDC1 increases Akt phosphorylation by promoting PTEN mRNA degradation, which promotes neuronal survival, especially after ischemia (Zhang Z. et al., 2020). The lack of YTHDC1 in oocytes impedes its development at the primary follicular stage, additionally, result in a large number of selective splicing defects in oocytes (Kasowitz et al., 2018). Specifically, YTHDC1 is required for female oocyte growth and maturation. Chen H. et al. (2020) conducted a seven-center case-control study in Chinese children concluded that YTHDC1 gene polymorphism may have a cumulative effect on the susceptibility of hepatoblastoma oncogene. This study suggests that YTHDC1 might be a potential biomarker and therapeutic target for certain diseases, not only the biological profile of the malignant disease.

According to previous researches, YTHDC2 may play an important role in a variety of diseases genesis, and development. More recent studies suggest that YTHDC2 is a potential candidate gene for pancreatic cancer susceptibility (Fanale et al., 2014) and is associated with immune infiltration in head and neck squamous cell carcinoma (Li Y. et al., 2020). In lung adenocarcinoma, it suppresses the tumorigenesis and development by inhibiting SLC7A11-dependent antioxidant function (Ma et al., 2021). In non-small cell lung cancer, downregulation of m6A reader YTHDC2 promotes tumor progression and predicts poor prognosis (Sun S. et al., 2020). YTHDC2 may promote the metastasis of colon cancer by promoting the translation of HIF-1 α , and YTHDC2 may be a diagnostic marker and target gene for the treatment of colon

cancer (Tanabe et al., 2016). Zeng et al. (2020) suggest that mRNA m6A modification and YTHDC2 expression are crucial to meiotic initiation and progression in female germ cells. And YTHDC2 also regulates the transition from proliferation to differentiation of germline (Bailey et al., 2017).

Another recognized YTH domain-containing protein member is YTHDF1-3, which all have a conserved m6A-binding domain. In 2020, Liu J. et al. (2020) shown that autophagy YTHDF2/3 is required for pluripotent stem cells reprogramming. Shi et al. (2019) found that high expression of YTHDF1 is associated with better hypoxic adaptation suppression of non-small cell lung cancer (NSCLC), however when its depletion causes cancer cells to develop resistance to cisplatin (DDP) therapy via the KEAP1-NRF2-AKR1C1 axis, especially the accumulation of reactive oxygen species (ROS) induced by cisplatin treatment. Also, By stabilizing MAP2K4 and MAP4K4 mRNA transcription, YTHDF2 activates MAPK and NF- κ B signaling pathways promotes the expression of pro-inflammatory cytokines, and aggravates the inflammatory response (Yu et al., 2019).

In eukaryotic cells, in addition to the well-characterized YTH protein, there is also a unique m6A “code reader” protein. These readers recognize and directly and specifically bind to the m6A site, which plays an important role in RNA metabolism. The IGF2BP family is newly reported m6A readers, which are consisted of three members of IGF2BP1-3 (Huang et al., 2018). These proteins are responsible for the stability of targeted mRNAs and are associated with thousands of targets, such as MDR1, MYC, and KRAS (Bell et al., 2013). In short, IGF2BPs recognizes mRNAs modified by m6A and promotes cancer progression by recruiting RNA stabilizers, thus maintaining their stability (Huang et al., 2018; Li et al., 2019; Hanniford et al., 2020).

The primary role of eIF3 is to facilitate translation. Protein translation typically begins with the recruitment of the 43S ribosomal complex to the 5' cap of mRNAs by a cap-binding complex. However, m6A directly binds eukaryotic initiation factor 3 (eIF3) and recruits the 43S complex to initiate translation in the absence of the cap-binding factor eIF4E (Meyer et al., 2015). Recently, Lee et al. (2016) revealed that eIF3d (a subunit of the eIF3 complex) is an mRNA cap-binding protein required for specific translation initiation. DAP5, an eIF4GI homolog that lacks eIF4E binding, which promotes cap-dependent translation by combining directly with eIF3d (de la Parra et al., 2018). (Alarcón et al., 2015) found that the RNA-binding protein hnRNP A2B1 binds to the m6A RNA, whose biochemical footprint matches the m6A common motif, and proposed that hnRNP A2B1 is A nuclear reader of m6A markers, and modulated the effect of this marker on the processing and selective splicing of primary microRNAs to A certain extent.

ROLE OF RNA M6A MODIFICATION IN MALE FERTILITY

Infertility is one of the common health problems in modern society, with about 50 percent of problems being caused by male factors (Drobnis and Johnson, 2015; Whitfield et al., 2015). Sperm

defects are the primary cause of male infertility, including sperm concentration, sperm motility, and morphology (Huynh et al., 2002). In recent years, A growing number of studies have shown that modification of germ cell RNA m6A can lead to defects in the process of spermatogenesis, leading to male infertility, and which can not be cured by assisted reproductive technology (Table 1 and Figure 2). So we will discuss the effects of different enzymes modified by m6A on spermatogenesis, which provides a novel insight for improving the therapeutic effect of this kind of patients, which may be beneficial to the further clinical application of this kind of patients in male infertility treatment.

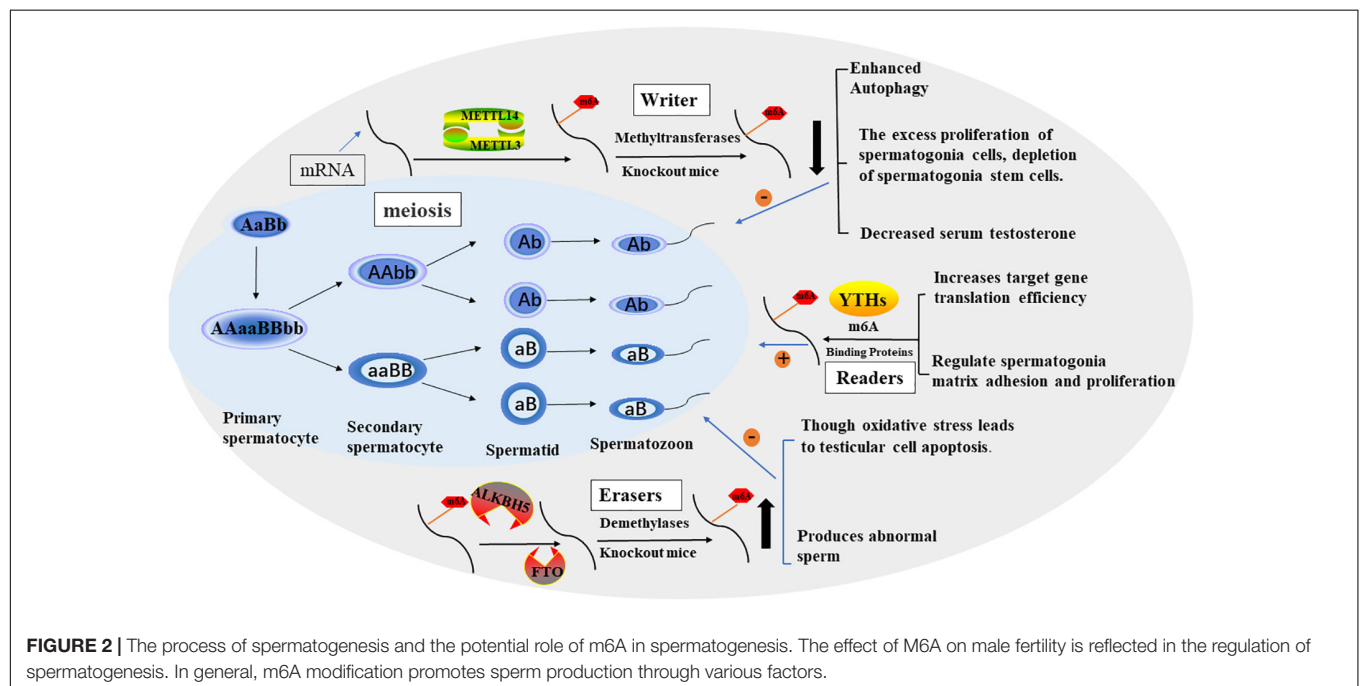
Mammalian spermatogenesis is a highly specialized differentiation process involving multiple regulatory mechanisms, and m6A can influence pre-mRNA splicing, mRNA output, turnover, and translation, which are controlled in the male germline to ensure coordinated gene expression. In 2019, Yang et al. (2016) collected semen samples and tested the content of m6A in sperm ribonucleic acid using liquid mass spectrometry and the expression of m6A-modified proteins using real-time polymerase chain reaction. The results showed that increased M6a content was a risk factor for asthenozoospermia and affected sperm motility. Methyltransferases, especially METTL3, play a key role in increasing the amount of m6A in sperm RNA. Xu et al. (2017) found that METTL3 regulates spermatogonia differentiation and controls meiosis initiation in germ cells. Lin et al. (2017) proposed the m6A mRNA methylomes of mouse spermatogenic cells from the five developmental stages: undifferentiated spermatogonia, A1 type spermatogonia, pre-embryonic spermatogonia, pachytene/diploid spermatogonia, and round spermatogonia. The study highlights the key role of m6A gene modification in germline development and the potential to ensure coordinated translation at different stages of spermatogenesis. YTHDC2 improves the translation efficiency of target genes, affects the testicular volume of male mice, and enables the development of germ cells in mice, which plays a key role in spermatogenesis (Hsu et al., 2017). The sterile mutant “ketu”, caused by YTHDC2 missense mutation, makes the mutant germ cells enter meiosis, but prematurely enter abnormal metaphase and apoptosis. And defective genes that lead from spermatogonia to meiosis (Jain et al., 2018). A recent study (Tang et al., 2020) shows that in mouse spermatogenesis, the start and stop codon of linear ribonucleic acid is usually located around the m6A enriched sites, causing late pachytene spermatocytes to develop into round cells, which then extend into spermatoblast cells.

Methyltransferases (“Writers”) in Spermatogenesis

In 2017, Xu et al. (2017) shown that loss of METTL3 in germ cells severely inhibits spermatogonia differentiation and blocks the onset of meiosis by cultivating germ cell-specific METTL3 knockout mice. Therefore, they demonstrated that METTL3 is necessary for male fertility and spermatogenesis. Consistent with this, Lin et al. (2017) found that the m6A RNA methyltransferase METTL3 and METTL14 proteins colocalize to the nucleus of male germ cells. m6A deficiency caused by

TABLE 1 | Roles of m6A proteins and biological mechanisms exerted in spermatogenesis.

Type	Regulator	role in spermatogenesis	Functional classification	References
Writers	METTL3	Blocked the initiation of meiosis.	Loss of m6A severely inhibited spermatogonia differentiation.	Xu et al., 2017
	METTL3/14	Promoting SSC/progenitor cell proliferation and differentiation.	Loss of m6A leads to dysregulated translation of SSC/progenitor cell and causing SSC depletion.	Lin et al., 2017
	METTL14\ALKBH5	Regulates testosterone synthesis through modulating autophagy in Leydig cells.	Reduced mRNA methylation levels of m6A and enhanced autophagy in LCs.	Chen Y. et al., 2020
	METTL3	Regulating the expression of genes critical for sex hormone synthesis and gonadotropin signaling.	Loss of METTL3 leads to failed gamete maturation and significantly reduced fertility in zebrafish.	Xia et al., 2018
Erasers	ALKBH5	Impaired fertility resulting from apoptosis that affects meiosis metaphase-stage spermatocytes.	ALKBH5 deficiency leads to compromised spermatogenesis in mice.	Zheng et al., 2013
	ALKBH5	Controls splicing and stability of long 3'-UTR mRNAs in male germ cells.	Dysregulation of many genes could contribute to meiotic defects.	Tang et al., 2018
	FTO\YTHDC2	Inhibition of demethylase FTO contributes to MEHP-induced Leydig cell injury.	Leading to aggravated oxidative stress and testicular injury.	Zhao et al., 2020, 2021
	FTO	A associated with reduced semen quality.	Single nucleotide variants could cause a shift in the transcription of the gene.	Landfors et al., 2016
Readers	YTHDC2	Regulates a meiotic in the mammalian germline.	Loss of YTHDC2 results in upregulation of several genes that are normally expressed in the mitotic spermatogonia and downregulation of meiotic genes.	Jain et al., 2018; Wojtas et al., 2017
	YTHDC2	Regulates the transition from proliferation to differentiation in the germline.	The proper progression of germ cells through meiosis is licensed by YTHDC2 through post-transcriptional regulation.	Bailey et al., 2017
	YTHDC2	Regulates mammalian spermatogenesis.	Lacking YTHDC2 are infertile and do not contain germ cells able to develop beyond the zygotene stage of meiotic prophase I.	Hsu et al., 2017
	YTHDC2	Promotes spermatogonia adhesion.	Depletion of YTHDC2 mainly downregulated the expression of MMPs, thus affecting cell adhesion and proliferation.	Huang et al., 2020



germ cell-specific inactivation of METTL3 and METTL14 can lead to excessive proliferation of spermatogonial cells, which would, in turn, result in depletion of spermatogonial stem cells. Immediately after studies proved (Chen Y. et al., 2020) there was

a negative correlation between m6A methyltransferase METTL14 and autophagy in testicular stromal cells. The regulation of autophagy in testicular stromal cells can affect the synthesis of testosterone. So, it provides insights into therapeutic strategies

for azoospermia and oligozoospermia in patients with reduced serum testosterone. Xia et al. (2018) found defects in sperm maturation and sperm motility are significantly reduced in m6A methyltransferase METTL3 mutant zebrafish. Meanwhile, they pointed that m6A METTL3 is playing important role in the expression of genes critical for sex hormone synthesis and gonadotropin signaling.

Demethylases (“Erasers”) in Spermatogenesis

Zheng et al. (2013) were detected the ALKBH5 gene's highest expression levels in testes. m6A gene was increased in male mice with ALKBH5-targeted deletion, and the number of sperm released and incised caudal epididymis were significantly reduced, sperm morphology was abnormal and motility was greatly reduced. The results showed that fertility was impaired due to the abnormal apoptosis and production of a small number of abnormal spermatozoa during meiosis. Tang et al. (2018) came to a similar conclusion, male mice with ALKBH5-targeted deletion testes were approximately half of the size of wild-type controls, the apoptotic germ cells in testis increased, the meiosis process was delayed, and the active germ cells decreased. Meanwhile, they found that m6A tends to label the 3'-UTR of longer mRNAs destined to be degraded during spermatogenesis to keep the stability of long 3'-UTR mRNAs and prevent abnormal splicing to produce shorter transcripts that rapidly degrade.

Landfors et al. (2016) found that FTO genetic variant to be associated with decreased semen quality. In Zhao et al. (2020) conduct a series of studies (Zhao et al., 2020, 2021) demonstrated that Di-(2-Ethylhexyl) phthalate increases m6A RNA modification, deteriorates testicular histology, reduces testosterone concentration, down-regulates spermatogenesis inducer expression, enhances oxidative stress, and increases testicular cell apoptosis by altering the expression of two important RNA methylation regulatory genes, FTO and YTHDC2. These findings link oxidative stress imbalance to the epigenetic effects of DEHP toxicity and provide insights into the testicular toxicity of DEHP from a new perspective of m6A modification.

m6A Binding Proteins (“Readers”) in Spermatogenesis

Several previous studies (Bailey et al., 2017; Hsu et al., 2017; Wojtas et al., 2017) have shown that YTHDC2 improves the translation efficiency of target genes. Unlike other commonly expressed YTH proteins, YTHDC2 is enriched in the testis. The YTHDC2 knockout mice showed defects in spermatogenesis but no other significant developmental defects. Male mice with YTHDC2 knockout were infertile and the germ cells don't develop to the zygotic stage. The sterile mutant “ketu” is a missense mutant of the YTHDC2 gene, which causes defects in the transition of germ cells to the meiotic RNA expression program (Jain et al., 2018). Thus, the regulation of M6A transcription by YTHDC2 is the key to the success of meiosis in the mammalian germline. Recent research (Huang et al., 2020) demonstrated that YTHDF2 also regulates the expression of

MMPs through the m6A/mRNA degradation pathway to regulate spermatogonia cell-matrix adhesion and proliferation.

Therefore, it can be speculated that m6A plays a major regulatory role in sperm production and development, and provides a new direction for the treatment of male infertility, but further exploration is needed.

ROLE OF RNA M6A MODIFICATIONS IN MALE GENITAL SYSTEM TUMORS

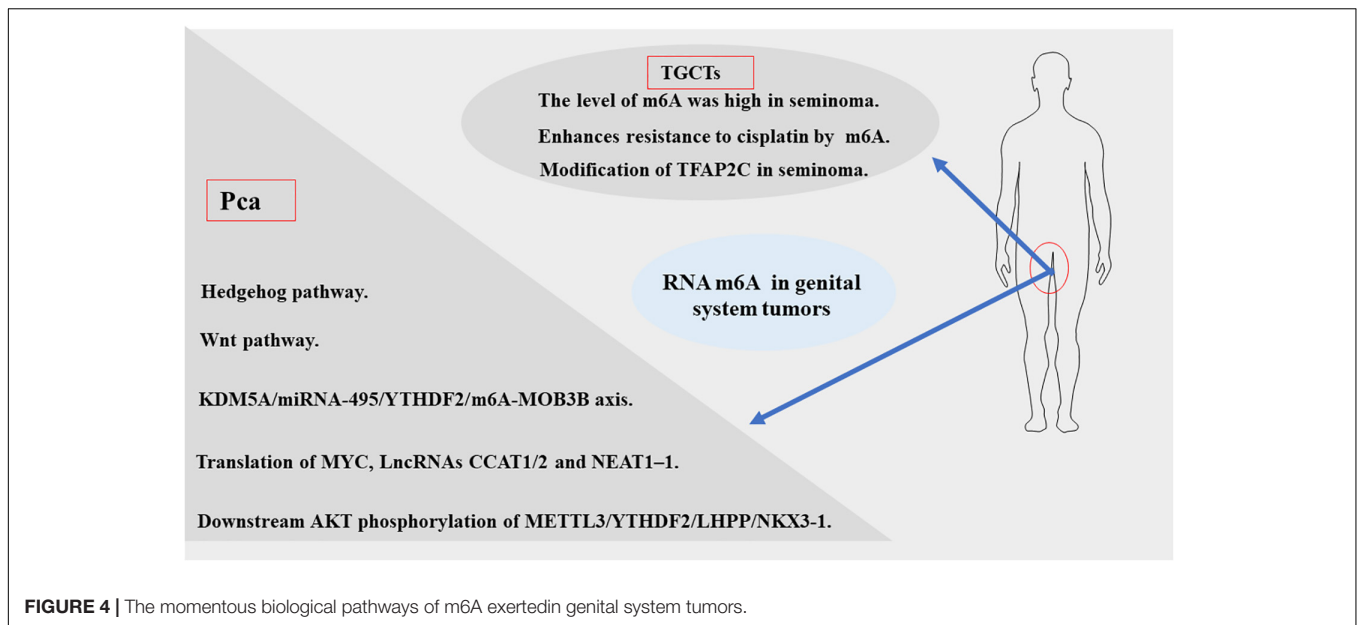
In recent years, breakthroughs have been made in the diagnosis and treatment of male genital system tumors (Nilsson et al., 2009; Chovanec et al., 2018; Lobo et al., 2019b). However, morbidity and mortality rates remain high. This indicates an urgent need to understand the mechanisms of metastasis and drug resistance of tumors. Recent evidence suggests that RNA m6A methylation is closely related to the development and progression of genital system tumors, including carcinogenesis, proliferation, metastasis, and tumor suppression. According to the existing literature, m6A modifications in genital system tumors are mainly concentrated in prostate cancer and testicular germ cell tumors. Therefore, we firstly analyzed the expression of the m6A regulator in prostate cancer and testicular tumors using databases such as The Cancer Genome Atlas (TCGA) dataset¹ (Figure 3). The results also showed that m6A related genes were closely related to male genital system tumors. Next, we briefly review the mechanism of m6A methylation in prostate cancer, testicular germ cell tumors, and seminoma for future treatment (Table 2 and Figure 4).

Prostate Cancer

Globally, about 1.3 million new cases of prostate cancer and 359,000 associated deaths worldwide are reported in 2018, ranking as the second most common cancer and the fifth most leading cause of cancer death in men (Bray et al., 2018). In 2021, in America estimated about 248,530 new prostate cancer cases account for 26% of all incident cases in men, and the second most leading cause of cancer death in the United States (Siegel et al., 2021). Although there are many treatments available, such as androgen-deprivation therapy, a small number of patients still develops resistance (Castration-Resistant Prostate Cancer, CRPC), and progress to lethal metastatic disease (mCRPC). Therefore, identifying novel molecular targets, and understanding the mechanisms driving PCa is of critical importance to treatment. Since 2019, several groups have explored the mechanisms of m6A in prostate cancer.

Cai et al. suggested that upregulation of m6A methyltransferase METTL3 promotes the growth and movement of prostate cancer cells by the Hedgehog pathway for the first time in 2019 (Cai et al., 2019). In 2020, several studies have shown that m6A writer METTL3 and VIRMA serve an oncogenic role to promotes the development and progression of PCa by Wnt pathway (Ma et al., 2020) and promoting the translation of MYC (Yuan et al., 2020), lncRNAs CCAT1/2 (Barros-Silva et al., 2020). Other studies have proposed that

¹<https://portal.gdc.com>



METTL3 regulates the expression of Integrin $\beta 1$ (ITGB1) and a high m6A level of LncRNA NEAT1-1 promotes the bone and lung metastasis of PCa (Li E. et al., 2020; Wen et al., 2020). Du and Li et al. showed that YTHDF2 performed Oncogene functions in PCa through the KDM5A/miRNA-495/YTHDF2/m6A-MOB3B axis and the downstream AKT phosphorylation of METTL3/YTHDF2/LHPP/NKX 3-1, respectively (Du et al., 2020; Li J. et al., 2020). According to a new study in 2021 (Zhu et al., 2021), the m6A demethylase FTO inhibits the invasion and migration of prostate cancer cells by regulating and reducing the total m6A level. Wu et al. found the abnormal expression of m6A modification-related enzymes in PCa leads to the elevated m6A level, which is related to Gleason classification (Wu et al., 2021). Thus, based on the current study, we can be concluded that m6A levels may contribute to the development and progression of PCa. But the specific mechanism still needs to be further explored.

Testicular Germ Cell Tumors

Testicular tumors are relatively rare, accounting for only 1-2% of all tumors in men. The testicular germ cell tumor (TGCTs) accounts for 95% of testicular carcinoma, which is divided into seminoma and non-seminoma histologically (Walsh et al., 2006), while seminoma represents more than 55% of germ cell tumors (GCTs) (Chia et al., 2010).

Nettersheim et al.'s studies (Nettersheim et al., 2019) have shown that RNA levels of m6A were increased during differentiation of GCT cell lines. Lobo et al. (Lobo et al., 2019a) found that the abundance of m6A and the expression of VIRMA/YTHDF3 differed among TGCT subtypes, and the level of m6A was high in seminoma (SE), which accurately distinguished SEs and non-seminomatous tumor (NSTs), forming a new candidate biomarker for patient management. Testicular germ cell tumors are significantly sensitive to anticancer drug cisplatin (Motzer et al., 1991; Einhorn, 1993), which contributes to an overall good prognosis. However,

some patients have developed resistance to platinum-based treatments, the emergence of cisplatin (CDDP) resistance is the main cause of treatment failure and death in patients with testicular germ cell tumors (TGCT), but its biologic background is poorly understood (Motzer et al., 1991; Bagrodia et al., 2016; Bakardjieva-Mihaylova et al., 2019). Recently, Wei et al. (2020) found that METTL3 enhances resistance to cisplatin by m6A modification of TFAP2C in seminoma. According to the above evidence, m6A modification plays an important carcinogenic role in the development and progression of TGCT. Therefore, the m6A related genes can be used as a prognostic indicator of early TGCT and may be a potential therapeutic target to prevent the progression of TGCT.

THE POTENTIAL APPLICATION OF RNA M6A IN CLINICAL

Although there are rarely studies on the use of m6A in andrology, for the treatment of other diseases which points us in the direction of research in the future. The treatment of PD-1 has achieved great success in the clinic. However, only a small proportion of cancer patients benefit from PD-1 blocking therapy, and overcoming resistance to PD-1 blocking has become a priority. METTL3/14 can regulate immune responses to PD-1-treated tumors (Wang et al., 2020). Yang et al. (2019) shown that knocking out FTO made melanoma cells hypersensitive to interferon γ (IFN- γ) and PD-1 therapy in melanoma. Li N. et al. (2020) found that in melanoma, colorectal, loss of the m6A demethylase ALKBH5 makes tumors sensitive to cancer immunotherapy by suppressive immune cell accumulation in the tumor microenvironment. Han et al. (2019) found that binding of YTHDF1 to a transcript encoding a lysosomal protease produces an adequate and persistent anti-tumor immune response. Meanwhile, the therapeutic effect of PD-L1 checkpoint blockade was significantly enhanced in YTHDF1 mice.

In a groundbreaking study, Su et al. (2018) found that the oncometabolite R-2-hydroxyglutarate (R-2HG) inhibits FTO activity. In this study, R-2HG was used to directly inhibit FTO in R-2HG-sensitive acute myeloid leukemia and glioma cells, leading to increased methylation and decreased expression of c-MYC and CEBPA mRNAs, thereby enhancing the anti-tumor effects. A study provides evidence of the non-steroidal anti-inflammatory drug Meclofenamic acid (MA) is a highly selective FTO inhibitor, which leads to increased levels of m6A modification in glioblastoma cells, inhibiting tumor progression and prolonging the life span of glioblastoma stem cell (GSC) transplanted mice (Huang et al., 2015; Cui et al., 2017). Recent studies have developed two FTO inhibitors, FB23 and FB23-2, they have been shown to promote differentiation/apoptosis of acute myeloid leukemia cell lines *in vitro*, and significantly inhibit the proliferation of human acute myeloid leukemia (AML) cell lines and primary AML cells, and the progression of human acute myeloid leukemia in xenografted mice (Huang et al., 2019). Deoxycholic acid reduces the expression of miR-92b-3p by m6A-dependent post-transcriptional modification by promoting the dissociation of METTL3 from METTL3-METTL14-WTAP complex and plays a role as a tumor suppressor in gallbladder cancer (Lin et al., 2020).

Radiotherapy is a common method for the treatment of tumors. However, many cancer patients develop radiation resistance after radiotherapy. The exact reason is not yet clear. Studies have shown that the silencing METTL3 increases the sensitivity of glioma stem-like cells (Visvanathan et al., 2018), colon cancer cells (Zhang et al., 2019), and pancreatic cancer cells (Taketo et al., 2018) to chemotherapy and radiation resistance. FTO modulates β -catenin expression by decreasing the level of m6A in its gene transcriptome, and increasing the activity of excision repair cross complementation group 1 (ERCC1), thereby enhancing radiotherapy tolerance *in vivo* and *in vitro* (Zhou et al., 2018). In Nasopharyngeal Carcinoma (He et al., 2020), YTHDC2 promotes radiotherapy resistance through activating IGF1R/AKT/S6 Signaling Axis.

Chemotherapy is the standard treatment for a variety of cancers, especially for patients who cannot tolerate surgery; however, acquired chemotherapeutic resistance is one of the major causes of treatment failure. Downregulation of m6A demethylase FTO and ALKBH5 increased the modification of the FZD10mRNA gene to reduce the sensitivity of ovarian epithelial cell carcinoma to PARPI by upregulating the Wnt/ β -catenin pathway (Fukumoto et al., 2019). m6A methyltransferase METTL3 mediated autophagy increases the sensitivity of non-small cell lung cancer cells to gefitinib by β -elemene (Liu S. et al., 2020). And METTL3 also directly promotes YAP translation and increases YAP activity by regulating the MALAT1-miR-1914-3p-YAP axis to induce drug resistance and metastasis of non-small cell lung cancer (Jin et al., 2019). m6A-methyladenosine modification modulates the β -catenin signaling pathway to maintain sorafenib resistance in hepatocellular carcinoma (Xu et al., 2020).

According to existing studies, RNA m6 modification not only provides a new idea for male infertility with spermatogenesis disorders but also plays a role in tumorigenesis, showing great potential for early tumor diagnosis, targeted therapy,

and improving the sensitivity of chemotherapy drugs and radiotherapy. However, there may still be limitations in the future clinical application of m6A, such as which patients need m6A assistance; Who can afford to pay; Whether the detection of m6A related proteins involves ethics, etc.

DISCUSSION

As a newly discovered type of post-transcriptional regulation, dynamic reversible m6A modification is the most common type of internal modification for RNA methylation. Although According to previous researches, m6A plays an important role in spermatogenesis, there are still many unsolved problems that need to be further explored, such as the specific mechanism of m6A in spermatogenesis, whether there is a direct link between writer, erasers, and readers. Besides, the majority of the current studies on m6A methylation are focused on animal experiments, which may have specific limitations. What additional external factors can change these enzymes and further affect spermatogenesis. In this review, we found mRNA m6A modifications have been intensely studied in male genital system tumors in the past few years. However, most studies focused on the specific mechanism related to m6A while there is a lack of studies about the application of m6A such as new drug targeting m6A and diagnostic biomarkers based on m6A regulators.

Despite the fruitful results of the preliminary studies, we believe that the following exploratory work still needs to be completed. Above all, more efforts and multi-center, large-scale studies are needed to explore the specific role of m6A in the various pathways that regulate gene expression. Besides, it is still possible that some of the enzymes that modify m6A have not been identified. Such as FMR and HNRNPC may act as a novel m6A binding protein (van Tran et al., 2019). Therefore, the development of new detection methods for m6A will be helpful for the identification of m6A modifiers and the discovery of new regulatory mechanisms. Finally, m6A modifies not only mRNA but also other RNAs including miRNAs (Qi et al., 2019; Thyagarajan et al., 2019), lncRNA (Quinn and Chang, 2016; Yang et al., 2018), and circRNA (Yang et al., 2017; Zhou et al., 2017; Ren C. et al., 2019). Therefore, non-coding RNAs have the potential to become new therapeutic targets for diseases. However, whether there is a potential link between RNA m6A modification and other types of RNA modification remains to be determined.

CONCLUSION

According to the reviewed studies, m6A modification plays an important role in spermatogenesis and male genital system tumor occurrence and development. However, the current research is still only at the basic research level. So, to apply in the clinical as soon as possible, we need more efforts and more multicenter, large-scale studies to further explore the role of m6A gene modification in tumor biology. In a word, our research and understanding of the modification of m6A are still in their infancy.

AUTHOR CONTRIBUTIONS

SL, YL, and YaW collected the related manuscript and finished the manuscript and figures. ZD gave constructive guidance and made final approval. RL, XF, YuW, and XG participated in the design of this review. All authors read and approved the final manuscript.

FUNDING

This work was supported by grants from Funding for Longyuan Young Innovative and Entrepreneurial Talents (842011); The National Innovation and Entrepreneurship

Training Program for Undergraduate (Grant no. 202110730201); Hui-Chun Chin and Tsung-Dao Lee Chinese Undergraduate Research Endowment (LZU-JZH2224); and Cuiying Scientific Training Program for Undergraduates of Lanzhou University Second Hospital (CYXZ2019-06).

ACKNOWLEDGMENTS

We acknowledge and thank our study participants. We apologize for not being able to cite all of relevant contributions, owing to space limitations.

REFERENCES

- Alarcón, C. R., Goodarzi, H., Lee, H., Liu, X., Tavazoie, S., and Tavazoie, S. F. (2015). HNRNPA2B1 Is a Mediator of m^A-Dependent Nuclear RNA Processing Events. *Cell* 162, 1299–1308. doi: 10.1016/j.cell.2015.08.011
- Bagrodia, A., Lee, B. H., Lee, W., Cha, E. K., Sfakianos, J. P., Iyer, G., et al. (2016). Genetic Determinants of Cisplatin Resistance in Patients With Advanced Germ Cell Tumors. *J. Clin. Oncol.* 34, 4000–4007. doi: 10.1200/jco.2016.68.7798
- Bailey, A. S., Batista, P. J., Gold, R. S., Chen, Y. G., de Rooij, D. G., Chang, H. Y., et al. (2017). The conserved RNA helicase YTHDC2 regulates the transition from proliferation to differentiation in the germline. *eLife* 6:e26116. doi: 10.7554/eLife.26116
- Bakardjieva-Mihaylova, V., Skvarova Kramarzova, K., Slamova, M., Svaton, M., Rejlova, K., Zaliava, M., et al. (2019). Molecular Basis of Cisplatin Resistance in Testicular Germ Cell Tumors. *Cancers* 11:1316. doi: 10.3390/cancers11091316
- Barros-Silva, D., Lobo, J., Guimarães-Teixeira, C., Carneiro, I., Oliveira, J., Martens-Uzunova, E. S., et al. (2020). VIRMA-Dependent N⁶-Methyladenosine Modifications Regulate the Expression of Long Non-Coding RNAs CCAT1 and CCAT2 in Prostate Cancer. *Cancers* 12:771. doi: 10.3390/cancers12040771
- Bell, J. L., Wächter, K., Mühlecker, B., Pazaitis, N., Köhn, M., Lederer, M., et al. (2013). Insulin-like growth factor 2 mRNA-binding proteins (IGF2BPs): post-transcriptional drivers of cancer progression? *Cell. Mol. Life Sci.* 70, 2657–2675. doi: 10.1007/s00018-012-1186-z
- Bokar, J. A., Shambaugh, M. E., Polayes, D., Matera, A. G., and Rottman, F. M. (1997). Purification and cDNA cloning of the AdoMet-binding subunit of the human mRNA (N⁶-adenosine)-methyltransferase. *RNA* 3, 1233–1247. doi: 10.1093/rna/3.12.1233
- Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R. L., Torre, L. A., and Jemal, A. (2018). Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* 68, 394–424. doi: 10.3322/caac.21492
- Cai, J., Yang, F., Zhan, H., Situ, J., Li, W., Mao, Y., et al. (2019). RNA m^A Methyltransferase METTL3 Promotes The Growth Of Prostate Cancer By Regulating Hedgehog Pathway. *Onco Targets Ther.* 12, 9143–9152. doi: 10.2147/ott.S226796
- Castillo, J. J., Mull, N., Reagan, J. L., Nemr, S., and Mitri, J. (2012). Increased incidence of non-Hodgkin lymphoma, leukemia, and myeloma in patients with diabetes mellitus type 2: a meta-analysis of observational studies. *Blood* 119, 4845–4850. doi: 10.1182/blood-2011-06-362830
- Chen, H., Li, Y., Li, L., Zhu, J., Yang, Z., Zhang, J., et al. (2020). YTHDC1 gene polymorphisms and hepatoblastoma susceptibility in Chinese children: A seven-center case-control study. *J. Gene Med.* 22:e3249. doi: 10.1002/jgm.3249
- Chen, S., Li, Y., Zhi, S., Ding, Z., Wang, W., Peng, Y., et al. (2020). WTAP promotes osteosarcoma tumorigenesis by repressing HMBOX1 expression in an m^A-dependent manner. *Cell Death Dis.* 11:659. doi: 10.1038/s41419-020-02847-6
- Chen, Y., Wang, J., Xu, D., Xiang, Z., Ding, J., Yang, X., et al. (2020). m^A mRNA methylation regulates testosterone synthesis through modulating autophagy in Leydig cells. *Autophagy* 17, 457–475. doi: 10.1080/15548627.2020.1720431
- Chia, V. M., Quraishi, S. M., Devesa, S. S., Purdue, M. P., Cook, M. B., and McGlynn, K. A. (2010). International trends in the incidence of testicular cancer, 1973–2002. *Cancer Epidemiol. Biomarkers. Prev.* 19, 1151–1159. doi: 10.1158/1055-9965.Epi-10-0031
- Chovanec, M., Albany, C., Mego, M., Montironi, R., Cimdamore, A., and Cheng, L. (2018). Emerging Prognostic Biomarkers in Testicular Germ Cell Tumors: Looking Beyond Established Practice. *Front. Oncol.* 8:571. doi: 10.3389/fonc.2018.00571
- Cui, Q., Shi, H., Ye, P., Li, L., Qu, Q., Sun, G., et al. (2017). m^A RNA Methylation Regulates the Self-Renewal and Tumorigenesis of Glioblastoma Stem Cells. *Cell Rep.* 18, 2622–2634. doi: 10.1016/j.celrep.2017.02.059
- de la Parra, C., Ernlund, A., Alard, A., Ruggles, K., Ueberheide, B., and Schneider, R. J. (2018). A widespread alternate form of cap-dependent mRNA translation initiation. *Nat. Commun.* 9:3068. doi: 10.1038/s41467-018-05539-0
- Desrosiers, R., Friderici, K., and Rottman, F. (1974). Identification of methylated nucleosides in messenger RNA from Novikoff hepatoma cells. *Proc. Natl. Acad. Sci. U.S.A.* 71, 3971–3975. doi: 10.1073/pnas.71.10.3971
- Dina, C., Meyre, D., Gallina, S., Durand, E., Körner, A., Jacobson, P., et al. (2007). Variation in FTO contributes to childhood obesity and severe adult obesity. *Nat. Genet.* 39, 724–726. doi: 10.1038/ng2048
- Dominissini, D., Moshitch-Moshkovitz, S., Schwartz, S., Salmon-Divon, M., Ungar, L., Osenberg, S., et al. (2012). Topology of the human and mouse m⁶A RNA methylomes revealed by m⁶A-seq. *Nature* 485, 201–206. doi: 10.1038/nature11112
- Drobnis, E. Z., and Johnson, M. (2015). The question of sperm DNA fragmentation testing in the male infertility work-up: a response to Professor Lewis' commentary. *Reprod. Biomed. Online* 31, 138–139. doi: 10.1016/j.rbmo.2015.05.004
- Du, C., Lv, C., Feng, Y., and Yu, S. (2020). Activation of the KDM5A/miRNA-495/YTHDF2/m⁶A-MOB3B axis facilitates prostate cancer progression. *J. Exp. Clin. Cancer Res.* 39:223. doi: 10.1186/s13046-020-01735-3
- Du, H., Zhao, Y., He, J., Zhang, Y., Xi, H., Liu, M., et al. (2016). YTHDF2 destabilizes m⁶A-containing RNA through direct recruitment of the CCR4-NOT deadenylase complex. *Nat. Commun.* 7:12626. doi: 10.1038/ncomms12626
- Dubin, L., and Amelar, R. D. (1971). Etiologic factors in 1294 consecutive cases of male infertility. *Fertil. Steril.* 22, 469–474. doi: 10.1016/s0015-028238400-x
- Einhorn, L. H. (1993). General Motors Cancer Research Prizewinners Laureates Lectures. Charles F. Kettering Prize. Clinical trials in testicular cancer. *Cancer* 71, 3182–3184.
- Fanale, D., Iovanna, J. L., Calvo, E. L., Berthezene, P., Belleau, P., Dagorn, J. C., et al. (2014). Germline copy number variation in the YTHDC2 gene: does it have a role in finding a novel potential molecular target involved in pancreatic adenocarcinoma susceptibility? *Expert Opin. Ther. Targets* 18, 841–850. doi: 10.1517/14728222.2014.920324
- Frayling, T. M., Timpson, N. J., Weedon, M. N., Zeggini, E., Freathy, R. M., Lindgren, C. M., et al. (2007). A common variant in the FTO gene is associated with body mass index and predisposes to childhood and adult obesity. *Science* 316, 889–894. doi: 10.1126/science.1141634
- Fukumoto, T., Zhu, H., Nacarelli, T., Karakashev, S., Fatkhutdinov, N., Wu, S., et al. (2019). N-Methylation of Adenosine of FZD10 mRNA Contributes to PARP Inhibitor Resistance. *Cancer Res.* 79, 2812–2820. doi: 10.1158/0008-5472.Can-18-3592
- Ghazarian, A. A., Kelly, S. P., Altekruze, S. F., Rosenberg, P. S., and McGlynn, K. A. (2017). Future of testicular germ cell tumor incidence in the United States: Forecast through 2026. *Cancer* 123, 2320–2328. doi: 10.1002/cncr.30597

- Guzick, D. S., Overstreet, J. W., Factor-Litvak, P., Brazil, C. K., Nakajima, S. T., Coutifaris, C., et al. (2001). Sperm morphology, motility, and concentration in fertile and infertile men. *N. Engl. J. Med.* 345, 1388–1393. doi: 10.1056/NEJMoa003005
- Han, D., Liu, J., Chen, C., Dong, L., Liu, Y., Chang, R., et al. (2019). Anti-tumour immunity controlled through mRNA m⁶A methylation and YTHDF1 in dendritic cells. *Nature* 566, 270–274. doi: 10.1038/s41586-019-0916-x
- Hanniford, D., Ulloa-Morales, A., Karz, A., Berzoti-Coelho, M. G., Moubarak, R. S., Sánchez-Sendra, B., et al. (2020). Epigenetic Silencing of CDR1as Drives IGF2BP3-Mediated Melanoma Invasion and Metastasis. *Cancer Cell* 37, 55–70.e15. doi: 10.1016/j.ccell.2019.12.007
- Hartmann, A. M., Nayler, O., Schwaiger, F. W., Obermeier, A., and Stamm, S. (1999). The interaction and colocalization of Sam68 with the splicing-associated factor YT521-B in nuclear dots is regulated by the Src family kinase p59(fyn). *Mol. Biol. Cell* 10, 3909–3926. doi: 10.1091/mbc.10.11.3909
- He, J. J., Li, Z., Rong, Z. X., Gao, J., Mu, Y., Guan, Y. D., et al. (2020). m⁶A Reader YTHDC2 Promotes Radiotherapy Resistance of Nasopharyngeal Carcinoma via Activating IGF1R/AKT/S6 Signaling Axis. *Front. Oncol.* 10:1166. doi: 10.3389/fonc.2020.01166
- Hsu, P. J., Zhu, Y., Ma, H., Guo, Y., Shi, X., Liu, Y., et al. (2017). Ythdc2 is an N-methyladenosine binding protein that regulates mammalian spermatogenesis. *Cell Res.* 27, 1115–1127. doi: 10.1038/cr.2017.99
- Huang, H., Weng, H., Sun, W., Qin, X., Shi, H., Wu, H., et al. (2018). Recognition of RNA N-methyladenosine by IGF2BP proteins enhances mRNA stability and translation. *Nat. Cell Biol.* 20, 285–295. doi: 10.1038/s41556-018-0045-z
- Huang, T., Liu, Z., Zheng, Y., Feng, T., Gao, Q., and Zeng, W. (2020). YTHDF2 promotes spermatogenic adhesion through modulating MMPs decay via m⁶A/mRNA pathway. *Cell Death Dis.* 11:37. doi: 10.1038/s41419-020-2235-4
- Huang, Y., Su, R., Sheng, Y., Dong, L., Dong, Z., Xu, H., et al. (2019). Small-Molecule Targeting of Oncogenic FTO Demethylase in Acute Myeloid Leukemia. *Cancer Cell* 35, 677–691.e10. doi: 10.1016/j.ccell.2019.03.006
- Huang, Y., Yan, J., Li, Q., Li, J., Gong, S., Zhou, H., et al. (2015). Meclofenamic acid selectively inhibits FTO demethylation of m⁶A over ALKBH5. *Nucleic Acids Res.* 43, 373–384. doi: 10.1093/nar/gku1276
- Huynh, T., Mollard, R., and Trounson, A. (2002). Selected genetic factors associated with male infertility. *Hum. Reprod. Update* 8, 183–198. doi: 10.1093/humupd/8.2.183
- Ignatova, V. V., Stolz, P., Kaiser, S., Gustafsson, T. H., Lastres, P. R., Sanz-Moreno, A., et al. (2020). The rRNA m⁶A methyltransferase METTL5 is involved in pluripotency and developmental programs. *Genes Dev.* 34, 715–729. doi: 10.1101/gad.333369.119
- Iles, M. M., Law, M. H., Stacey, S. N., Han, J., Fang, S., Pfeiffer, R., et al. (2013). A variant in FTO shows association with melanoma risk not due to BMI. *Nat. Genet.* 45, 428–432. doi: 10.1038/ng.2571
- Jain, D., Puno, M. R., Meydan, C., Lailier, N., Mason, C. E., Lima, C. D., et al. (2018). ketu mutant mice uncover an essential meiotic function for the ancient RNA helicase YTHDC2. *eLife* 7:e30919. doi: 10.7554/eLife.30919
- Ji, G., Long, Y., Zhou, Y., Huang, C., Gu, A., and Wang, X. (2012). Common variants in mismatch repair genes associated with increased risk of sperm DNA damage and male infertility. *BMC Med.* 10:49. doi: 10.1186/1741-7015-10-49
- Jia, G., Fu, Y., Zhao, X., Dai, Q., Zheng, G., Yang, Y., et al. (2011). N⁶-methyladenosine in nuclear RNA is a major substrate of the obesity-associated FTO. *Nat. Chem. Biol.* 7, 885–887. doi: 10.1038/nchembio.687
- Jin, D., Guo, J., Wu, Y., Du, J., Yang, L., Wang, X., et al. (2019). m⁶A mRNA methylation initiated by METTL3 directly promotes YAP translation and increases YAP activity by regulating the MALAT1-miR-1914-3p-YAP axis to induce NSCLC drug resistance and metastasis. *J. Hematol. Oncol.* 12:135. doi: 10.1186/s13045-019-0830-6
- Kasowitz, S. D., Ma, J., Anderson, S. J., Leu, N. A., Xu, Y., Gregory, B. D., et al. (2018). Nuclear m⁶A reader YTHDC1 regulates alternative polyadenylation and splicing during mouse oocyte development. *PLoS Genet.* 14:e1007412. doi: 10.1371/journal.pgen.1007412
- Kyono, K., Fukunaga, N., Haigo, K., Chiba, S., and Araki, Y. (2001). Pregnancy achieved following ICSI from a man with Klinefelter's syndrome and spinal cord injury. *Hum. Reprod.* 16, 2347–2349. doi: 10.1093/humrep/16.11.2347
- Landfors, M., Nakken, S., Fusser, M., Dahl, J. A., Klungland, A., and Fedorcsak, P. (2016). Sequencing of FTO and ALKBH5 in men undergoing infertility work-up identifies an infertility-associated variant and two missense mutations. *Fertil. Steril.* 105, 1170–1179.e5. doi: 10.1016/j.fertnstert.2016.01.002
- Lee, A. S., Kranzusch, P. J., Doudna, J. A., and Cate, J. H. (2016). eIF3d is an mRNA cap-binding protein that is required for specialized translation initiation. *Nature* 536, 96–99. doi: 10.1038/nature18954
- Li, E., Wei, B., Wang, X., and Kang, R. (2020). METTL3 enhances cell adhesion through stabilizing integrin β 1 mRNA via an m⁶A-HuR-dependent mechanism in prostatic carcinoma. *Am. J. Cancer Res.* 10, 1012–1025. doi: 10.1016/j.amjcr.2020.01.002
- Li, F., Zhao, D., Wu, J., and Shi, Y. (2014). Structure of the YTH domain of human YTHDF2 in complex with an m⁶A mononucleotide reveals an aromatic cage for m⁶A recognition. *Cell Res.* 24, 1490–1492. doi: 10.1038/cr.2014.153
- Li, J., Xie, H., Ying, Y., Chen, H., Yan, H., He, L., et al. (2020). YTHDF2 mediates the mRNA degradation of the tumor suppressors to induce AKT phosphorylation in N⁶-methyladenosine-dependent way in prostate cancer. *Mol. Cancer* 19:152. doi: 10.1186/s12943-020-01267-6
- Li, M., Zhao, X., Wang, W., Shi, H., Pan, Q., Lu, Z., et al. (2018). Ythdf2-mediated m⁶A mRNA clearance modulates neural development in mice. *Genome Biol.* 19:69. doi: 10.1186/s13059-018-1436-y
- Li, N., Kang, Y., Wang, L., Huff, S., Tang, R., Hui, H., et al. (2020). ALKBH5 regulates anti-PD-1 therapy response by modulating lactate and suppressive immune cell accumulation in tumor microenvironment. *Proc. Natl. Acad. Sci. U.S.A.* 117, 20159–20170. doi: 10.1073/pnas.1918986117
- Li, T., Hu, P. S., Zuo, Z., Lin, J. F., Li, X., Wu, Q. N., et al. (2019). METTL3 facilitates tumor progression via an m⁶A-IGF2BP2-dependent mechanism in colorectal carcinoma. *Mol. Cancer* 18:112. doi: 10.1186/s12943-019-1038-7
- Li, Y., Zheng, J. N., Wang, E. H., Gong, C. J., Lan, K. F., and Ding, X. (2020). The m⁶A reader protein YTHDC2 is a potential biomarker and associated with immune infiltration in head and neck squamous cell carcinoma. *PeerJ* 8:e10385. doi: 10.7717/peerj.10385
- Li, Z., Qian, P., Shao, W., Shi, H., He, X. C., Gogol, M., et al. (2018). Suppression of m⁶A reader Ythdf2 promotes hematopoietic stem cell expansion. *Cell Res.* 28, 904–917. doi: 10.1038/s41422-018-0072-0
- Li, Z., Weng, H., Su, R., Weng, X., Zuo, Z., Li, C., et al. (2017). FTO Plays an Oncogenic Role in Acute Myeloid Leukemia as a N-Methyladenosine RNA Demethylase. *Cancer Cell* 31, 127–141. doi: 10.1016/j.ccell.2016.11.017
- Lin, R., Zhan, M., Yang, L., Wang, H., Shen, H., Huang, S., et al. (2020). Deoxycholic acid modulates the progression of gallbladder cancer through N-methyladenosine-dependent microRNA maturation. *Oncogene* 39, 4983–5000. doi: 10.1038/s41388-020-1349-6
- Lin, Z., Hsu, P. J., Xing, X., Fang, J., Lu, Z., Zou, Q., et al. (2017). Mettl3-/Mettl14-mediated mRNA N-methyladenosine modulates murine spermatogenesis. *Cell Res.* 27, 1216–1230. doi: 10.1038/cr.2017.117
- Liu, J., Gao, M., Xu, S., Chen, Y., Wu, K., Liu, H., et al. (2020). YTHDF2/3 Are Required for Somatic Reprogramming through Different RNA Deadenylation Pathways. *Cell Rep.* 32:108120. doi: 10.1016/j.celrep.2020.108120
- Liu, J., Yue, Y., Han, D., Wang, X., Fu, Y., Zhang, L., et al. (2014). A METTL3-METTL14 complex mediates mammalian nuclear RNA N⁶-adenosine methylation. *Nat. Chem. Biol.* 10, 93–95. doi: 10.1038/nchembio.1432
- Liu, S., Li, Q., Li, G., Zhang, Q., Zhuo, L., Han, X., et al. (2020). The mechanism of m⁶A methyltransferase METTL3-mediated autophagy in reversing gefitinib resistance in NSCLC cells by β -elemene. *Cell Death Dis.* 11:969. doi: 10.1038/s41419-020-03148-8
- Lobo, J., Costa, A. L., Cantante, M., Guimarães, R., Lopes, P., Antunes, L., et al. (2019a). m⁶A RNA modification and its writer/reader VIRMA/YTHDF3 in testicular germ cell tumors: a role in seminoma phenotype maintenance. *J. Transl. Med.* 17:79. doi: 10.1186/s12967-019-1837-z
- Lobo, J., Rodrigues, Á., Guimarães, R., Cantante, M., Lopes, P., Maurício, J., et al. (2019b). Detailed Characterization of Immune Cell Infiltrate and Expression of Immune Checkpoint Molecules PD-L1/CTLA-4 and MMR Proteins in Testicular Germ Cell Tumors Disclose Novel Disease Biomarkers. *Cancers* 11:1535. doi: 10.3390/cancers11101535
- Luxton, H. J., Simpson, B. S., Mills, I. G., Brindle, N. R., Ahmed, Z., Stavrinides, V., et al. (2019). The Oncogene Metadherin Interacts with the Known Splicing Proteins YTHDC1, Sam68 and T-STAR and Plays a Novel Role in Alternative mRNA Splicing. *Cancers* 11:1233. doi: 10.3390/cancers11091233
- Ma, H., Wang, X., Cai, J., Dai, Q., Natchiar, S. K., Lv, R., et al. (2019). N(6)-Methyladenosine methyltransferase ZCCHC4 mediates ribosomal

- RNA methylation. *Nat. Chem. Biol.* 15, 88–94. doi: 10.1038/s41589-018-0184-3
- Ma, L., Chen, T., Zhang, X., Miao, Y., Tian, X., Yu, K., et al. (2021). The mA reader YTHDC2 inhibits lung adenocarcinoma tumorigenesis by suppressing SLC7A11-dependent antioxidant function. *Redox Biol.* 38:101801. doi: 10.1016/j.redox.2020.101801
- Ma, X. X., Cao, Z. G., and Zhao, S. L. (2020). m6A methyltransferase METTL3 promotes the progression of prostate cancer via m6A-modified LEF1. *Eur. Rev. Med. Pharmacol. Sci.* 24, 3565–3571. doi: 10.26355/eurrev_202004_20817
- Mao, Y., Dong, L., Liu, X. M., Guo, J., Ma, H., Shen, B., et al. (2019). mA in mRNA coding regions promotes translation via the RNA helicase-containing YTHDC2. *Nat. Commun.* 10:5332. doi: 10.1038/s41467-019-13317-9
- Mendel, M., Chen, K. M., Homolka, D., Gos, P., Pandey, R. R., McCarthy, A. A., et al. (2018). Methylation of Structured RNA by the mA Writer METTL16 Is Essential for Mouse Embryonic Development. *Mol. Cell.* 71, 986–1000.e11. doi: 10.1016/j.molcel.2018.08.004
- Meyer, K. D., Patil, D. P., Zhou, J., Zinoviev, A., Skabkin, M. A., Elemento, O., et al. (2015). 5' UTR mA Promotes Cap-Independent Translation. *Cell* 163, 999–1010. doi: 10.1016/j.cell.2015.10.012
- Meyer, K. D., Saletore, Y., Zumbo, P., Elemento, O., Mason, C. E., and Jaffrey, S. R. (2012). Comprehensive analysis of mRNA methylation reveals enrichment in 3' UTRs and near stop codons. *Cell* 149, 1635–1646. doi: 10.1016/j.cell.2012.05.003
- Mohammad, H. P., Barbash, O., and Creasy, C. L. (2019). Targeting epigenetic modifications in cancer therapy: erasing the roadmap to cancer. *Nat. Med.* 25, 403–418. doi: 10.1038/s41591-019-0376-8
- Motzer, R. J., Geller, N. L., Tan, C. C., Herr, H., Morse, M., Fair, W., et al. (1991). Salvage chemotherapy for patients with germ cell tumors. The Memorial Sloan-Kettering Cancer Center experience (1979–1989). *Cancer* 67, 1305–1310.
- Nettersheim, D., Berger, D., Jostes, S., Kristiansen, G., Lochnit, G., and Schorle, H. (2019). N6-Methyladenosine detected in RNA of testicular germ cell tumors is controlled by METTL3, ALKBH5, YTHDC1/F1/F2, and HNRNPC as writers, erasers, and readers. *Andrology* 7, 498–506. doi: 10.1111/andr.12612
- Nilsson, J., Skog, J., Nordstrand, A., Baranov, V., Mincheva-Nilsson, L., Breakefield, X. O., et al. (2009). Prostate cancer-derived urine exosomes: a novel approach to biomarkers for prostate cancer. *Br. J. Cancer* 100, 1603–1607. doi: 10.1038/sj.bjc.6605058
- Ostermeier, G. C., Dix, D. J., Miller, D., Khatri, P., and Krawetz, S. A. (2002). Spermatozoal RNA profiles of normal fertile men. *Lancet* 360, 772–777. doi: 10.1016/s0140-673609899-9
- Pendleton, K. E., Chen, B., Liu, K., Hunter, O. V., Xie, Y., Tu, B. P., et al. (2017). The U6 snRNA mA Methyltransferase METTL16 Regulates SAM Synthetase Intron Retention. *Cell* 169, 824–835.e14. doi: 10.1016/j.cell.2017.05.003
- Perry, R. P., Kelley, D. E., Friderici, K., and Rottman, F. (1975). The methylated constituents of L cell messenger RNA: evidence for an unusual cluster at the 5' terminus. *Cell* 4, 387–394. doi: 10.1016/0092-867490159-2
- Ping, X. L., Sun, B. F., Wang, L., Xiao, W., Yang, X., Wang, W. J., et al. (2014). Mammalian WTAP is a regulatory subunit of the RNA N6-methyladenosine methyltransferase. *Cell Res.* 24, 177–189. doi: 10.1038/cr.2014.3
- Pinto, R., Vågbo, C. B., Jakobsson, M. E., Kim, Y., Baltissen, M. P., O'Donohue, M. F., et al. (2020). The human methyltransferase ZCCHC4 catalyses N6-methyladenosine modification of 28S ribosomal RNA. *Nucleic Acids Res.* 48, 830–846. doi: 10.1093/nar/gkz1147
- Qi, L., Gao, C., Feng, F., Zhang, T., Yao, Y., Wang, X., et al. (2019). MicroRNAs associated with lung squamous cell carcinoma: New prognostic biomarkers and therapeutic targets. *J. Cell. Biochem.* 120, 18956–18966. doi: 10.1002/jcb.29216
- Quinn, J. J., and Chang, H. Y. (2016). Unique features of long non-coding RNA biogenesis and function. *Nat. Rev. Genet.* 17, 47–62. doi: 10.1038/nrg.2015.10
- Ren, C., Liu, J., Zheng, B., Yan, P., Sun, Y., and Yue, B. (2019). The circular RNA circ-ITCH acts as a tumour suppressor in osteosarcoma via regulating miR-22. *Artif. Cells Nanomed. Biotechnol.* 47, 3359–3367. doi: 10.1080/21691401.2019.1649273
- Ren, W., Lu, J., Huang, M., Gao, L., Li, D., Wang, G. G., et al. (2019). Structure and regulation of ZCCHC4 in mA-methylation of 28S rRNA. *Nat. Commun.* 10:5042. doi: 10.1038/s41467-019-12923-x
- Roundtree, I. A., Luo, G. Z., Zhang, Z., Wang, X., Zhou, T., Cui, Y., et al. (2017). YTHDC1 mediates nuclear export of N-methyladenosine methylated mRNAs. *eLife* 6:e31311. doi: 10.7554/eLife.31311
- Scuteri, A., Sanna, S., Chen, W. M., Uda, M., Albai, G., Strait, J., et al. (2007). Genome-wide association scan shows genetic variants in the FTO gene are associated with obesity-related traits. *PLoS Genet.* 3:e115. doi: 10.1371/journal.pgen.0030115
- Shi, H., Wang, X., Lu, Z., Zhao, B. S., Ma, H., Hsu, P. J., et al. (2017). YTHDF3 facilitates translation and decay of N-methyladenosine-modified RNA. *Cell Res.* 27, 315–328. doi: 10.1038/cr.2017.15
- Shi, Y., Fan, S., Wu, M., Zuo, Z., Li, X., Jiang, L., et al. (2019). YTHDF1 links hypoxia adaptation and non-small cell lung cancer progression. *Nat. Commun.* 10:4892. doi: 10.1038/s41467-019-12801-6
- Siegel, R. L., Miller, K. D., Fuchs, H. E., and Jemal, A. (2021). Cancer Statistics, 2021. *CA Cancer J. Clin.* 71, 7–33. doi: 10.3322/caac.21654
- Śledź, P., and Jinek, M. (2016). Structural insights into the molecular mechanism of the mA writer complex. *eLife* 5:e18434. doi: 10.7554/eLife.18434
- Söderberg, K. C., Kaprio, J., Verkasalo, P. K., Pukkala, E., Koskenvuo, M., Lundqvist, E., et al. (2009). Overweight, obesity and risk of haematological malignancies: a cohort study of Swedish and Finnish twins. *Eur. J. Cancer* 45, 1232–1238. doi: 10.1016/j.ejca.2008.11.004
- Stoilov, P., Rafalska, I., and Stamm, S. (2002). YTH: a new domain in nuclear proteins. *Trends Biochem. Sci.* 27, 495–497. doi: 10.1016/s0968-000402189-8
- Su, R., Dong, L., Li, C., Nachtergaele, S., Wunderlich, M., Qing, Y., et al. (2018). R-2HG Exhibits Anti-tumor Activity by Targeting FTO/mA/MYC/CEBPA Signaling. *Cell* 172, 90–105.e23. doi: 10.1016/j.cell.2017.11.031
- Sun, S., Han, Q., Liang, M., Zhang, Q., Zhang, J., and Cao, J. (2020). Downregulation of mA reader YTHDC2 promotes tumor progression and predicts poor prognosis in non-small cell lung cancer. *Thorac. Cancer* 11, 3269–3279. doi: 10.1111/1759-7714.13667
- Sun, Y., Li, S., Yu, W., Zhao, Z., Gao, J., Chen, C., et al. (2020). N-methyladenosine-dependent pri-miR-17-92 maturation suppresses PTEN/TMEM127 and promotes sensitivity to everolimus in gastric cancer. *Cell Death Dis.* 11:836. doi: 10.1038/s41419-020-03049-w
- Taketo, K., Konno, M., Asai, A., Koseki, J., Toratani, M., Satoh, T., et al. (2018). The epitranscriptome m6A writer METTL3 promotes chemo- and radioresistance in pancreatic cancer cells. *Int. J. Oncol.* 52, 621–629. doi: 10.3892/ijo.2017.4219
- Tanabe, A., Tanikawa, K., Tsunetomi, M., Takai, K., Ikeda, H., Konno, J., et al. (2016). RNA helicase YTHDC2 promotes cancer metastasis via the enhancement of the efficiency by which HIF-1α mRNA is translated. *Cancer Lett.* 376, 34–42. doi: 10.1016/j.canlet.2016.02.022
- Tang, C., Klukovich, R., Peng, H., Wang, Z., Yu, T., Zhang, Y., et al. (2018). ALKBH5-dependent m6A demethylation controls splicing and stability of long 3'-UTR mRNAs in male germ cells. *Proc. Natl. Acad. Sci. U.S.A.* 115, E325–E333. doi: 10.1073/pnas.1717794115
- Tang, C., Xie, Y., Yu, T., Liu, N., Wang, Z., Woolsey, R. J., et al. (2020). mA-Dependent biogenesis of circular RNAs in male germ cells. *Cell Res.* 30, 211–228. doi: 10.1038/s41422-020-0279-8
- Thyagarajan, A., Tsai, K. Y., and Sahu, R. P. (2019). MicroRNA heterogeneity in melanoma progression. *Semin. Cancer Biol.* 59, 208–220. doi: 10.1016/j.semcancer.2019.05.021
- Tournaye, H., Krausz, C., and Oates, R. D. (2017). Concepts in diagnosis and therapy for male reproductive impairment. *Lancet Diabetes Endocrinol.* 5, 554–564. doi: 10.1016/s2213-858730043-2
- Tsubokura, T., Yamazaki, H., Masui, K., Sasaki, N., Shimizu, D., Suzuki, G., et al. (2018). Comparison of Image-Guided Intensity-Modulated Radiotherapy and Low-dose Rate Brachytherapy with or without External Beam Radiotherapy in Patients with Localized Prostate Cancer. *Sci. Rep.* 8:10538. doi: 10.1038/s41598-018-28730-1
- Ueda, Y., Ooshio, I., Fusamae, Y., Kitae, K., Kawaguchi, M., Jingushi, K., et al. (2017). AlkB homolog 3-mediated tRNA demethylation promotes protein synthesis in cancer cells. *Sci. Rep.* 7:42271. doi: 10.1038/srep42271
- van Tran, N., Ernst, F. G. M., Hawley, B. R., Zorbas, C., Ulryck, N., Hackert, P., et al. (2019). The human 18S rRNA m6A methyltransferase METTL5 is stabilized by TRMT112. *Nucleic Acids Res.* 47, 7719–7733. doi: 10.1093/nar/gkz619
- Verza, S., Jr., and Esteves, S. C. (2008). Sperm defect severity rather than sperm Source is associated with lower fertilization rates after intracytoplasmic sperm injection. *Int. Braz. J. Urol.* 34, 49–56. doi: 10.1590/s1677-55382008000100008
- Visvanathan, A., Patil, V., Arora, A., Hegde, A. S., Arivazhagan, A., Santosh, V., et al. (2018). Essential role of METTL3-mediated mA modification in glioma

- stem-like cells maintenance and radioresistance. *Oncogene* 37, 522–533. doi: 10.1038/onc.2017.351
- Walsh, T. J., Grady, R. W., Porter, M. P., Lin, D. W., and Weiss, N. S. (2006). Incidence of testicular germ cell cancers in U.S. children: SEER program experience 1973 to 2000. *Urology* 68, 402–405. doi: 10.1016/j.urology.2006.02.045
- Wang, L., Hui, H., Agrawal, K., Kang, Y., Li, N., Tang, R., et al. (2020). m A RNA methyltransferases METTL3/14 regulate immune responses to anti-PD-1 therapy. *EMBO J.* 39:e104514. doi: 10.15252/embj.2020104514
- Wang, P., Dostader, K. A., and Nam, Y. (2016). Structural Basis for Cooperative Function of Mettl3 and Mettl14 Methyltransferases. *Mol. Cell.* 63, 306–317. doi: 10.1016/j.molcel.2016.05.041
- Wang, X., Feng, J., Xue, Y., Guan, Z., Zhang, D., Liu, Z., et al. (2016). Structural basis of N-adenosine methylation by the METTL3-METTL14 complex. *Nature* 534, 575–578. doi: 10.1038/nature18298
- Wang, X., Lu, Z., Gomez, A., Hon, G. C., Yue, Y., Han, D., et al. (2014). N6-methyladenosine-dependent regulation of messenger RNA stability. *Nature* 505, 117–120. doi: 10.1038/nature12730
- Wang, X., Zhao, B. S., Roundtree, I. A., Lu, Z., Han, D., Ma, H., et al. (2015). N-methyladenosine Modulates Messenger RNA Translation Efficiency. *Cell* 161, 1388–1399. doi: 10.1016/j.cell.2015.05.014
- Warda, A. S., Kretschmer, J., Hackert, P., Lenz, C., Urlaub, H., Höbartner, C., et al. (2017). Human METTL16 is a N-methyladenosine (mA) methyltransferase that targets pre-mRNAs and various non-coding RNAs. *EMBO Rep.* 18, 2004–2014. doi: 10.15252/embr.201744940
- Wei, J., Yin, Y., Zhou, J., Chen, H., Peng, J., Yang, J., et al. (2020). METTL3 potentiates resistance to cisplatin through m A modification of TFAP2C in seminoma. *J. Cell Mol. Med.* 24, 11366–11380. doi: 10.1111/jcmm.15738
- Wen, J., Lv, R., Ma, H., Shen, H., He, C., Wang, J., et al. (2018). Zc3h13 Regulates Nuclear RNA mA Methylation and Mouse Embryonic Stem Cell Self-Renewal. *Mol. Cell.* 69, 1028–1038.e6. doi: 10.1016/j.molcel.2018.02.015
- Wen, S., Wei, Y., Zen, C., Xiong, W., Niu, Y., and Zhao, Y. (2020). Long non-coding RNA NEAT1 promotes bone metastasis of prostate cancer through N6-methyladenosine. *Mol. Cancer* 19:171. doi: 10.1186/s12943-020-01293-4
- Whitfield, M., Pollet-Villard, X., Levy, R., Drevet, J. R., and Saez, F. (2015). Posttesticular sperm maturation, infertility, and hypercholesterolemia. *Asian J. Androl.* 17, 742–748. doi: 10.4103/1008-682x.155536
- Wojtas, M. N., Pandey, R. R., Mendel, M., Homolka, D., Sachidanandam, R., and Pillai, R. S. (2017). Regulation of mA Transcripts by the 3′→5′ RNA Helicase YTHDC2 Is Essential for a Successful Meiotic Program in the Mammalian Germline. *Mol. Cell* 68, 374–387.e12. doi: 10.1016/j.molcel.2017.09.021
- Wu, Q., Xie, X., Huang, Y., Meng, S., Li, Y., Wang, H., et al. (2021). N6-methyladenosine RNA methylation regulators contribute to the progression of prostate cancer. *J. Cancer* 12, 682–692. doi: 10.7150/jca.46379
- Xia, H., Zhong, C., Wu, X., Chen, J., Tao, B., Xia, X., et al. (2018). Mettl3 Mutation Disrupts Gamete Maturation and Reduces Fertility in Zebrafish. *Genetics* 208, 729–743. doi: 10.1534/genetics.117.300574
- Xiao, W., Adhikari, S., Dahal, U., Chen, Y. S., Hao, Y. J., Sun, B. F., et al. (2016). Nuclear mA Reader YTHDC1 Regulates mRNA Splicing. *Mol. Cell* 61, 507–519. doi: 10.1016/j.molcel.2016.01.012
- Xu, J., Wan, Z., Tang, M., Lin, Z., Jiang, S., Ji, L., et al. (2020). N-methyladenosine-modified CircRNA-SORE sustains sorafenib resistance in hepatocellular carcinoma by regulating β -catenin signaling. *Mol. Cancer* 19:163. doi: 10.1186/s12943-020-01281-8
- Xu, K., Yang, Y., Feng, G. H., Sun, B. F., Chen, J. Q., Li, Y. F., et al. (2017). Mettl3-mediated mA regulates spermatogonial differentiation and meiosis initiation. *Cell Res.* 27, 1100–1114. doi: 10.1038/cr.2017.100
- Yang, D., Qiao, J., Wang, G., Lan, Y., Li, G., Guo, X., et al. (2018). N6-Methyladenosine modification of lincRNA 1281 is critically required for mESC differentiation potential. *Nucleic Acids Res.* 46, 3906–3920. doi: 10.1093/nar/gky130
- Yang, S., Wei, J., Cui, Y. H., Park, G., Shah, P., Deng, Y., et al. (2019). mA mRNA demethylase FTO regulates melanoma tumorigenicity and response to anti-PD-1 blockade. *Nat. Commun.* 10:2782. doi: 10.1038/s41467-019-10669-0
- Yang, Y., Fan, X., Mao, M., Song, X., Wu, P., Zhang, Y., et al. (2017). Extensive translation of circular RNAs driven by N-methyladenosine. *Cell Res.* 27, 626–641. doi: 10.1038/cr.2017.31
- Yang, Y., Huang, W., Huang, J. T., Shen, F., Xiong, J., Yuan, E. F., et al. (2016). Increased N6-methyladenosine in Human Sperm RNA as a Risk Factor for Asthenozoospermia. *Sci. Rep.* 6:24345. doi: 10.1038/srep24345
- Yu, R., Li, Q., Feng, Z., Cai, L., and Xu, Q. (2019). m6A Reader YTHDF2 Regulates LPS-Induced Inflammatory Response. *Int. J. Mol. Sci.* 20:1323. doi: 10.3390/ijms20061323
- Yuan, Y., Du, Y., Wang, L., and Liu, X. (2020). The M6A methyltransferase METTL3 promotes the development and progression of prostate carcinoma via mediating MYC methylation. *J. Cancer* 11, 3588–3595. doi: 10.7150/jca.42338
- Yue, Y., Liu, J., Cui, X., Cao, J., Luo, G., Zhang, Z., et al. (2018). VIRMA mediates preferential mA mRNA methylation in 3′UTR and near stop codon and associates with alternative polyadenylation. *Cell Discov.* 4:10. doi: 10.1038/s41421-018-0019-0
- Zeng, M., Dai, X., Liang, Z., Sun, R., Huang, S., Luo, L., et al. (2020). Critical roles of mRNA mA modification and YTHDC2 expression for meiotic initiation and progression in female germ cells. *Gene* 753:144810. doi: 10.1016/j.gene.2020.144810
- Zhang, X., Wang, F., Wang, Z., Yang, X., Yu, H., Si, S., et al. (2020). ALKBH5 promotes the proliferation of renal cell carcinoma by regulating AURKB expression in an mA-dependent manner. *Ann. Transl. Med.* 8:646. doi: 10.21037/atm-20-3079
- Zhang, Y., Kang, M., Zhang, B., Meng, F., Song, J., Kaneko, H., et al. (2019). mA modification-mediated CBX8 induction regulates stemness and chemosensitivity of colon cancer via upregulation of LGR5. *Mol. Cancer* 18:185. doi: 10.1186/s12943-019-1116-x
- Zhang, Z., Wang, Q., Zhao, X., Shao, L., Liu, G., Zheng, X., et al. (2020). YTHDC1 mitigates ischemic stroke by promoting Akt phosphorylation through destabilizing PTEN mRNA. *Cell Death Dis.* 11:977. doi: 10.1038/s41419-020-03186-2
- Zhao, T., Wang, J., Wu, Y., Han, L., Chen, J., Wei, Y., et al. (2021). Increased m6A modification of RNA methylation related to the inhibition of demethylase FTO contributes to MEHP-induced Leydig cell injury(star). *Environ. Pollut.* 268(Pt A):115627. doi: 10.1016/j.envpol.2020.115627
- Zhao, T. X., Wang, J. K., Shen, L. J., Long, C. L., Liu, B., Wei, Y., et al. (2020). Increased m6A RNA modification is related to the inhibition of the Nrf2-mediated antioxidant response in di-(2-ethylhexyl) phthalate-induced prepubertal testicular injury. *Environ. Pollut.* 259:113911. doi: 10.1016/j.envpol.2020.113911
- Zheng, G., Dahl, J. A., Niu, Y., Fedorcsak, P., Huang, C. M., Li, C. J., et al. (2013). ALKBH5 is a mammalian RNA demethylase that impacts RNA metabolism and mouse fertility. *Mol. Cell* 49, 18–29. doi: 10.1016/j.molcel.2012.10.015
- Zhou, C., Molin, B., Daneshvar, K., Pondick, J. V., Wang, J., Van Wittenberghe, N., et al. (2017). Genome-Wide Maps of m6A circRNAs Identify Widespread and Cell-Type-Specific Methylation Patterns that Are Distinct from mRNAs. *Cell Rep.* 20, 2262–2276. doi: 10.1016/j.celrep.2017.08.027
- Zhou, S., Bai, Z. L., Xia, D., Zhao, Z. J., Zhao, R., Wang, Y. Y., et al. (2018). FTO regulates the chemo-radiotherapy resistance of cervical squamous cell carcinoma (CSCC) by targeting β -catenin through mRNA demethylation. *Mol. Carcinog.* 57, 590–597. doi: 10.1002/mc.22782
- Zhu, K., Li, Y., and Xu, Y. (2021). The FTO mA demethylase inhibits the invasion and migration of prostate cancer cells by regulating total mA levels. *Life Sci.* 271:119180. doi: 10.1016/j.lfs.2021.119180
- Zhu, T., Roundtree, I. A., Wang, P., Wang, X., Wang, L., Sun, C., et al. (2014). Crystal structure of the YTH domain of YTHDF2 reveals mechanism for recognition of N6-methyladenosine. *Cell Res.* 24, 1493–1496. doi: 10.1038/cr.2014.152

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 Liu, Lao, Wang, Li, Fang, Wang, Gao and Dong. 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.



Emerging Role of m⁶A Methylome in Brain Development: Implications for Neurological Disorders and Potential Treatment

Godwin Sokpor^{1†}, Yuanbin Xie^{2†}, Huu P. Nguyen¹ and Tran Tuoc^{1*}

¹ Department of Human Genetics, Ruhr University of Bochum, Bochum, Germany, ² Department of Biochemistry and Molecular Biology, Gannan Medical University, Ganzhou, China

OPEN ACCESS

Edited by:

Giovanni Nigita,
The Ohio State University,
United States

Reviewed by:

Jean-Yves Roignant,
University of Lausanne, Switzerland
Qian Xu,
Xiangya Hospital, Central South
University, China
Jifeng Guo,
Xiangya Hospital, Central South
University, China

*Correspondence:

Godwin Sokpor
Godwin.Sokpor@ruhr-uni-bochum.de
Yuanbin Xie
xyb2003@sina.cn
Tran Tuoc
Tran.Tuoc@ruhr-uni-bochum.de

[†] These authors have contributed
equally to this work

Specialty section:

This article was submitted to
Epigenomics and Epigenetics,
a section of the journal
Frontiers in Cell and Developmental
Biology

Received: 21 January 2021

Accepted: 07 April 2021

Published: 19 May 2021

Citation:

Sokpor G, Xie Y, Nguyen HP and
Tuoc T (2021) Emerging Role of m⁶A
Methylome in Brain Development:
Implications for Neurological
Disorders and Potential Treatment.
Front. Cell Dev. Biol. 9:656849.
doi: 10.3389/fcell.2021.656849

Dynamic modification of RNA affords proximal regulation of gene expression triggered by non-genomic or environmental changes. One such epitranscriptomic alteration in RNA metabolism is the installation of a methyl group on adenosine [N⁶-methyladenosine (m⁶A)] known to be the most prevalent modified state of messenger RNA (mRNA) in the mammalian cell. The methylation machinery responsible for the dynamic deposition and recognition of m⁶A on mRNA is composed of subunits that play specific roles, including reading, writing, and erasing of m⁶A marks on mRNA to influence gene expression. As a result, peculiar cellular perturbations have been linked to dysregulation of components of the mRNA methylation machinery or its cofactors. It is increasingly clear that neural tissues/cells, especially in the brain, make the most of m⁶A modification in maintaining normal morphology and function. Neurons in particular display dynamic distribution of m⁶A marks during development and in adulthood. Interestingly, such dynamic m⁶A patterns are responsive to external cues and experience. Specific disturbances in the neural m⁶A landscape lead to anomalous phenotypes, including aberrant stem/progenitor cell proliferation and differentiation, defective cell fate choices, and abnormal synaptogenesis. Such m⁶A-linked neural perturbations may singularly or together have implications for syndromic or non-syndromic neurological diseases, given that most RNAs in the brain are enriched with m⁶A tags. Here, we review the current perspectives on the m⁶A machinery and function, its role in brain development and possible association with brain disorders, and the prospects of applying the clustered regularly interspaced short palindromic repeats (CRISPR)–dCas13b system to obviate m⁶A-related neurological anomalies.

Keywords: mRNA methylation, mRNA metabolism, N⁶-methyladenosine (m⁶A), cortical development, neurological disorders, clustered regularly interspaced short palindromic repeats (CRISPR)–dCas13b, m⁶A editing

INTRODUCTION

Over 170 chemical modifications of RNA are known to exist in eukaryotes (Boccalletto et al., 2018). These RNA modifications, together referred to as the epitranscriptome, play essential roles in gene expression regulation *via* affecting RNA metabolism: RNA processing, decay, transport, and translation. N⁶-methyladenosine (m⁶A) is among the characterized adenosine methylations of

messenger RNAs (mRNAs) (Engel and Chen, 2018) and the most occurring in mammalian cells (Roundtree et al., 2017a). The m⁶A mRNA methylome is dynamically regulated by factors that install, remove, or bind the m⁶A mark on mRNA. Such dynamism in the m⁶A landscape is known to critically regulate mRNA metabolism to influence gene expression. In essence, m⁶A modification is reported to modulate several biological events, including cell proliferation, differentiation, and embryonic development, and can also lead to disease conditions when dysregulated (Dominissini et al., 2012; Meyer et al., 2012; Ke et al., 2015; Linder et al., 2015; Zhao et al., 2017; Ries et al., 2019).

The impact of m⁶A modification on cell biological processes is notable in nervous tissues (Widagdo et al., 2016; Li et al., 2019). This is because neural cells are known to be enriched with m⁶A-tagged mRNAs. As a result, the developing and adult brain is reported to be ubiquitously enriched with m⁶A modifications (Chang et al., 2017; Zhang F. et al., 2018). The m⁶A level in the brain is temporally regulated in the course of its development such that the adult brain registers the highest level of m⁶A (Meyer et al., 2012). The massive prevalence of m⁶A in the developing and postnatal brain signifies the importance of m⁶A modification in regulating brain morphogenesis and function. Indeed, a chunk of the expanding knowledge indicating the phenomenal role of m⁶A in orchestrating neural development and function includes the proliferation of neural stem cells (NSCs) and other neural precursor cells, neuroprogenitor cell differentiation, gliogenesis, elaboration of neural processes, and synaptic transmission (reviewed in Widagdo and Anggono, 2018; Li et al., 2019; Chokkalla et al., 2020; Dermentzaki and Lotti, 2020; Livneh et al., 2020).

While maintenance of the general m⁶A homeostasis is indispensable for proper brain development and activity, selective hypomethylation and hypermethylation of gene transcripts are critical mechanistic modalities for normal brain neurodevelopment and functionality. Moreover, the targeted binding of m⁶A on transcription factor-encoding mRNAs and disease-risk gene transcripts in the developing brain (Zhang et al., 2020) reflects how m⁶A signaling is critical for brain structure and function in health. It also means that disturbance of the m⁶A RNA methylome can have implications for abnormal anatomy and physiology of the brain, leading to neurological disorders.

In this review, we present the role of the m⁶A methylation machinery in mRNA metabolism, with discussion focused on how the m⁶A landscape regulates brain development and function. Neurodevelopmental, neurodegenerative, and neuropsychiatric disorders of the brain attributable to m⁶A signaling dysregulation are highlighted.

THE N⁶-METHYLADENOSINE MODIFICATION MACHINERY

Early studies revealed that m⁶A is the most substantial posttranscriptional modification in eukaryotic mRNAs (Desrosiers et al., 1974; Perry et al., 1975a,b). It was proposed that nearly 7,000 mRNAs from human and mouse transcriptome

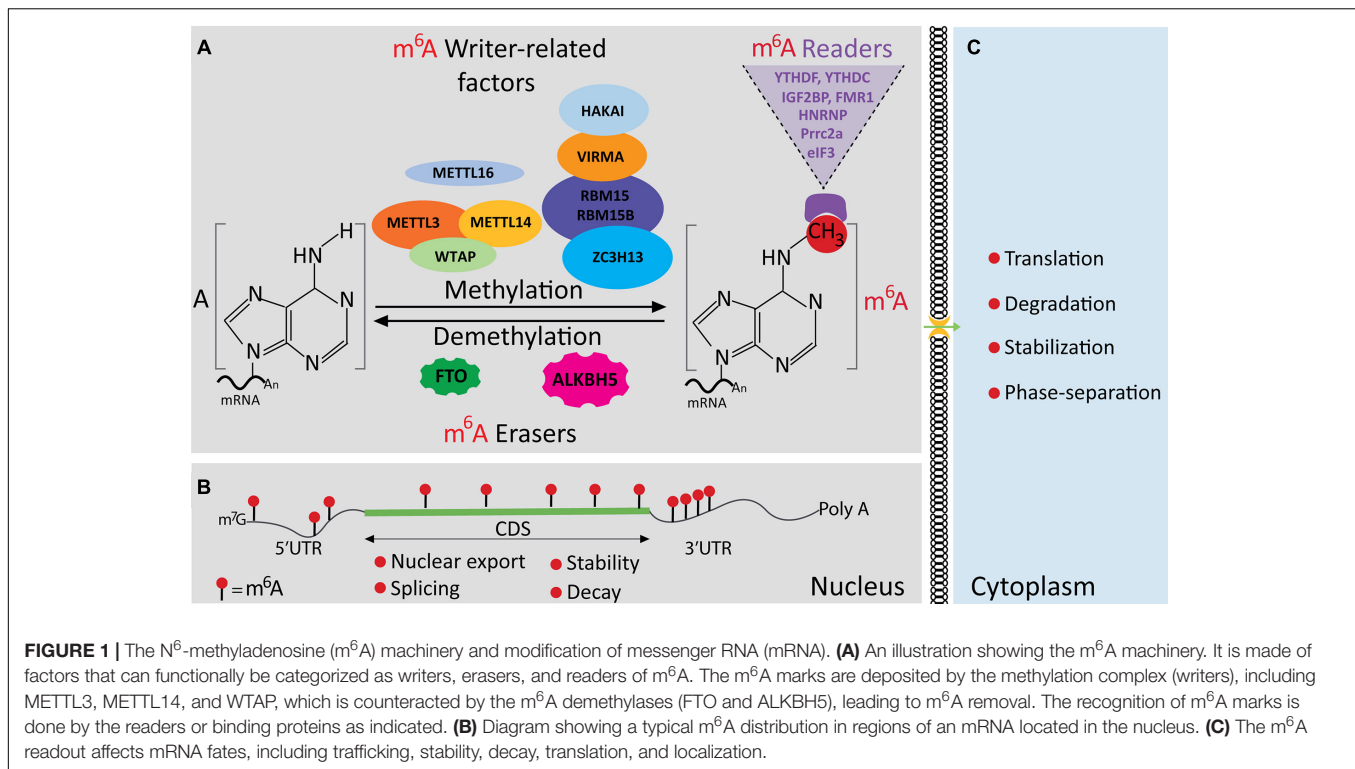
contain m⁶A modification (Dominissini et al., 2012; Meyer et al., 2012). Currently, over 10,000 m⁶A-modified mRNA transcripts have been identified in yeast and mammalian cells (Wang and Zhao, 2016). m⁶A modification and recognition (binding) are achieved by three functional components of the m⁶A machinery, namely, m⁶A methyltransferases (“writers”), m⁶A demethylases (“erasers”), and m⁶A binding or interacting proteins (“readers”) (Figure 1A). Through the specific binding of m⁶A-interacting proteins, m⁶A mRNA methylation is able to play a central role in regulating several aspects of mRNA metabolism, such as transport, splicing, stability, translation, and phase separation (Figure 1). High-throughput data revealed that m⁶A modification typically occurs within the consensus sequence RRACH (R stands for G or A; H stands for A, C, or U) (Dominissini et al., 2012; Meyer et al., 2012). The consensus sequence was recently redefined as DRACH, where D stands for G, A, or U (Linder et al., 2015). As depicted in Figure 1B, m⁶A distribution along mRNA is asymmetric. In general, m⁶A sites are concentrated in the protein coding region (CDS) near stop codons, followed by the 3′ untranslated region (UTR), and in the 5′ UTR (Figure 1B) (Dominissini et al., 2012; Meyer et al., 2012; Ke et al., 2015; Linder et al., 2015).

N⁶-Methyladenosine Writers

The installation of m⁶A is carried out by ~1 MDa m⁶A writer complex composed of the methyltransferase-like protein 3 (METTL3) and METTL14, which heterodimerize (METTL3-METTL14) to function as the enzymatic core of the writer complex (Bokar et al., 1994; Bujnicki et al., 2002; Liu et al., 2014; Iyer et al., 2016). Additionally, other factors are known to interact with the m⁶A writer complex. These include Wilms tumor 1-associated protein (WTAP) (Ping et al., 2014), VIRMA/KIA1429 (Yue et al., 2018), RNA-binding protein 15 (RBM15) (Patil et al., 2016; Huang and Yin, 2018), ZC3H13 (Knuckles et al., 2018), and HAKAI (Yue et al., 2018) (Figure 1A). These cofactors are regulated by the binding of RNA and the catalytic activity of the enzymatic core of the m⁶A writer complex (Bujnicki et al., 2002; Liu et al., 2014; Ping et al., 2014; Iyer et al., 2016; Yue et al., 2018).

METTL3 and METTL14

Bokar et al. (1994) partially purified the m⁶A writer complex using an *in vitro* methylation system and identified MT-70, a 70-kDa sub-complex possessing S-adenosylmethionine-binding methyltransferase capacity (Bokar et al., 1994). Later, it was renamed METTL3 (Narayan and Rottman, 1988; Bokar et al., 1997). Knockout of METTL3 in cells effectively blocks m⁶A modification of mRNAs (Zhong et al., 2008; Agarwala et al., 2012; Geula et al., 2015). On the other hand, METTL14 forms a stable heterodimer with METTL3 to form the methyltransferase core of the m⁶A methylation machinery (Liu et al., 2014; Wang Y. et al., 2014). METTL14, however, lacks enzymatic function and instead acts as an RNA-binding scaffold to augment the enzyme activity of METTL3 by directing the location of SAM methyl group required for the reaction (Śledź and Jinek, 2016; Wang P. et al., 2016; Wang X. et al., 2016). Therefore, METTL3 is the primary enzyme responsible for m⁶A installation on mRNA.



METTL3-METTL14-Associated Adaptors: WTAP, VIRMA (KIAA1429), RBM15/15B, ZC3H13 (KIAA0853), and HAKAI

The core m⁶A writer complex METTL3-METTL4 interacts with other adaptor proteins. It was found that FIP37 (the plant homolog of WTAP) co-localized with MTA (*Arabidopsis* homolog of METTL3) in the nucleus through physical interaction (Zhong et al., 2008). Similar interaction between WTAP and METTL3 was observed in mammalian cells (Liu et al., 2014; Ping et al., 2014; Schwartz et al., 2014). WTAP is key in keeping the METTL3-METTL4 heterodimer in nuclear speckles (Ping et al., 2014). Loss of WTAP leads to the depletion of m⁶A modification in mRNA, indicating that WTAP may orient METTL3-METTL4 onto targets (Ping et al., 2014). However, the detailed mechanism remains elusive. Of note, it was demonstrated that two classes of m⁶A sites exist: WTAP-dependent and WTAP-independent sites (Schwartz et al., 2014). VIRMA is known to also interact with the WTAP-METTL3-METTL4 complex (Figure 1A; Schwartz et al., 2014) and indicates its essentiality for the m⁶A writer complex functionality. Indeed, VIRMA deletion in human cells leads to a significant reduction in mRNA methylation levels, although not as intense as that resulting from WTAP knockdown (Schwartz et al., 2014). Biochemical studies from Yue et al. (2018) demonstrated that VIRMA preferentially regulates m⁶A modification in the 3' UTR proximal to stop codon.

Through proteomic studies, Horiuchi et al. (2013) observed that RBM15 and its paralog RBM15B, together with ZC3H13/KIAA0853, and MTA70 (METTL3), associate with WTAP (Horiuchi et al., 2013; Patil et al., 2016), which

raises the possibility that RBM15 and RBM15B may also play role(s) in m⁶A modification. Indeed, silencing of RBM15 and RBM15B led to a demonstrable decrease in m⁶A levels of mRNA (Patil et al., 2016). Based on Individual-nucleotide resolution UV crosslinking and immunoprecipitation (iCLIP) data, it was proposed that RBM15/15B recruit the m⁶A methylation machinery to perform m⁶A modification through binding to uridine-rich regions near DRACH sites. That notwithstanding, it is not always the case that uridine-rich regions exist near m⁶A sites; therefore, other methylation complex adaptors may mediate the complex binding to such variant m⁶A sites (Patil et al., 2016; Meyer and Jaffrey, 2017).

ZC3H13/KIAA0853 is also an interactor of the m⁶A machinery, and it is demonstrated to be crucial in linking RBM15/15B to WTAP (Horiuchi et al., 2013; Knuckles et al., 2018; Wen et al., 2018). Knockdown of ZC3H13 shifts the localization of the m⁶A adaptors WTAP, Virilizer, and Hakai from nucleus to cytosol in embryonic stem cells and leads to a significant total reduction in m⁶A level on mRNA (Wen et al., 2018). This reflects an essential role of ZC3H13 in the deposition of m⁶A on mRNAs. The E3 ubiquitin ligase HAKAI (CBL1) is another notable factor that interacts with the m⁶A machinery (Figure 1A; Horiuchi et al., 2013; Růžička et al., 2017). However, its function in m⁶A modification of mRNA in mammals is yet to be established.

Erasers (Demethylases) of N⁶-Methyladenosine

m⁶A modification is believed to be a reversible dynamic process premised on the identification of two demethylases: fat

mass and obesity-associated protein (FTO) (Jia et al., 2011) and α -ketoglutarate-dependent dioxygenase alkB homolog 5 (ALKBH5) (**Figure 1A**). However, this important concept has been in controversy due to various supporting data from various studies (Mauer et al., 2017; Darnell et al., 2018; Wei et al., 2018) as discussed below.

Fat Mass and Obesity-Associated Protein

Following an *in vitro* assay, which demonstrated that FTO erases m⁶A methylation of mRNA (Jia et al., 2008), it was further shown that downregulation (knockdown) of *FTO* in HeLa or 293FT cells caused reduction in m⁶A methylation of mRNA (Jia et al., 2011). In support of this observation, it was identified that a small proportion of m⁶A peaks of the whole transcriptome increased in *Fto* knockout mouse (Hess et al., 2013). These evidence consolidates the concept that m⁶A modification can be reversed by FTO functionality. However, this idea was challenged by another study, in which no significant increase in m⁶A level was observed in *Fto* knockout cells (Mauer et al., 2017). Instead, they noticed that FTO exhibits much higher catalytic capacity against m⁶Am than m⁶A. These studies indicate that the preferred substrate of FTO may be m⁶Am (Meyer and Jaffrey, 2017). What could be the explanation behind the discrepancy between these findings? It is worth pointing out that several independent groups have reported that the Kcat/Km of FTO against m⁶A is in the range of 0.6–0.7 min⁻¹ μ M⁻¹ (Jia et al., 2011; Zhu and Yi, 2014; Zou et al., 2016), whereas that from the study of Mauer et al. (2017) is only 0.06 min⁻¹ μ M⁻¹, indicating that most likely there is a technique issue behind quantification of the Kcat/Km of FTO against m⁶A. Additionally, both investigations used different methods to determine the level of m⁶A, noting that the RNase T1 treatment of mRNA combined with thin-layer chromatography can only measure the m⁶A in the case of RGACH, but not RAACH (Mauer et al., 2017). Moreover, a recent study further demonstrates that FTO not only demethylates internal m⁶A but also caps m⁶Am (Wei et al., 2018). The subcellular distribution of FTO varies among cultured cell lines, which indicates that the pattern of FTO demethylation of m⁶A in cytosol or nucleus could be cell lineage-dependent. Consistent with the above studies, it was found that FTO plays a vital role in cell cycle and mitosis regulation in an m⁶A demethylation-dependent manner during spermatogenesis (Huang T. et al., 2018).

Structural studies uncovered that FTO prefers m⁶A-modified nucleobase, and its demethylase activity can be influenced by the primary and the tertiary structure of target RNA (Zhang X. et al., 2019), thus shedding light on the molecular mechanism behind the demethylation function of FTO. Recent findings show that the transcription of *FTO* is regulated by a transcriptional factor *Zfp217* during adipogenesis, and *Zfp217* is critical for *FTO* to associate with m⁶A sites, albeit through competition with YTHDF2 for binding sites (Wei et al., 2019).

ALKBH5

ALKBH5 is another m⁶A factor with demethylase capacity (Jia et al., 2011; Zheng et al., 2013). Manipulating ALKBH5 expression level leads to a slight but significant change in m⁶A levels in the poly(A) region of mRNA. Compared with

FTO, which demethylates m⁶Am and m⁶A, ALKBH5 shows specificity for m⁶A demethylation (Wei et al., 2018). Importantly, m⁶A-mediated conformational change facilitates distinction of substrates with minor sequence by FTO and ALKBH5 (Zou et al., 2016). As a nuclear protein, ALKBH5 is proposed to only erase the m⁶A methylation in the nucleus (Meyer and Jaffrey, 2017). The demethylation capacity of ALKBH5 plays important roles in mRNA splicing, transport, stability, and processing. For instance, spermatogenic transcripts with increased m⁶A levels exhibit increased splicing events in *Alkbh5* KO mice (Tang et al., 2018). Recently, it was reported that METTL3 and ALKBH5 counteractively modulate the m⁶A methylation of *TFEB* transcript to effect regulation of autophagy (Song et al., 2019). The demethylation activity of ALKBH5 can be regulated by DEAD-Box RNA helicase through physical interaction (Shah et al., 2017).

Readers (Binding Proteins) of N⁶-Methyladenosine

The functional significance of m⁶A modification also depends on m⁶A-binding proteins also referred to as m⁶A readers. As described below, we categorize the readers in mammalian cells into two groups: YTH domain-containing proteins, including YTHDC1, YTHDC2, and YTHDF1–3 (Hazra et al., 2019), and Non YTH domain-containing proteins, including eIF3 (Meyer et al., 2015), IGF2BPs (Huang H. et al., 2018), HuR (Dominissini et al., 2012; Wang Y. et al., 2014), FMRP (Zhang F. et al., 2018), hnRNPA2/B1/C (Dominissini et al., 2012; Alarcón et al., 2015), and METTLs (Wang et al., 2015).

YTH Domain Containing N⁶-Methyladenosine-Binding Proteins

This group contains YTHDC1, YTHDC2, and YTHDF1–3 families in mammals. The common YTH domain defines members of this group of m⁶A binding and determines the nature of m⁶A reading (Zhang et al., 2010). However, they are not paralogs. This is because of the non-similarity of other aspects of proteins apart from the common YTH domain (Hazra et al., 2019).

YTHDC1 (YT521-B) is the first identified m⁶A reader, which was found as a protein associated with splicing factors (Imai et al., 1998; Harfmann et al., 1999; Xiao et al., 2016; Hazra et al., 2019). Interestingly, human YTHDC1 shows much greater binding affinity for the m⁶A-modified mRNA region in the context of Gm⁶AC (five-fold to six-fold difference) than Am⁶AC, although the distribution of m⁶A modification found in the consensus sequence is Gm⁶AC (70%) and Am⁶AC (30%) (Xu et al., 2014, 2015). It localizes in various subnuclear bodies close to the nuclear splicing factor (SF) compartments and plays a role in mRNA splicing through physical interaction with splicing factor SRSF3 and SRSF10 (**Figure 2A**; Xiao et al., 2016). Furthermore, YTHDC1 works together with NXF1 and SRSF3 to regulate m⁶A-modified mRNA nuclear export (**Figure 2B**; Roundtree et al., 2017b). Moreover, YTHDC1 has been reported to bind m⁶A-modified MAT2A mRNA. The m⁶A modification results in the

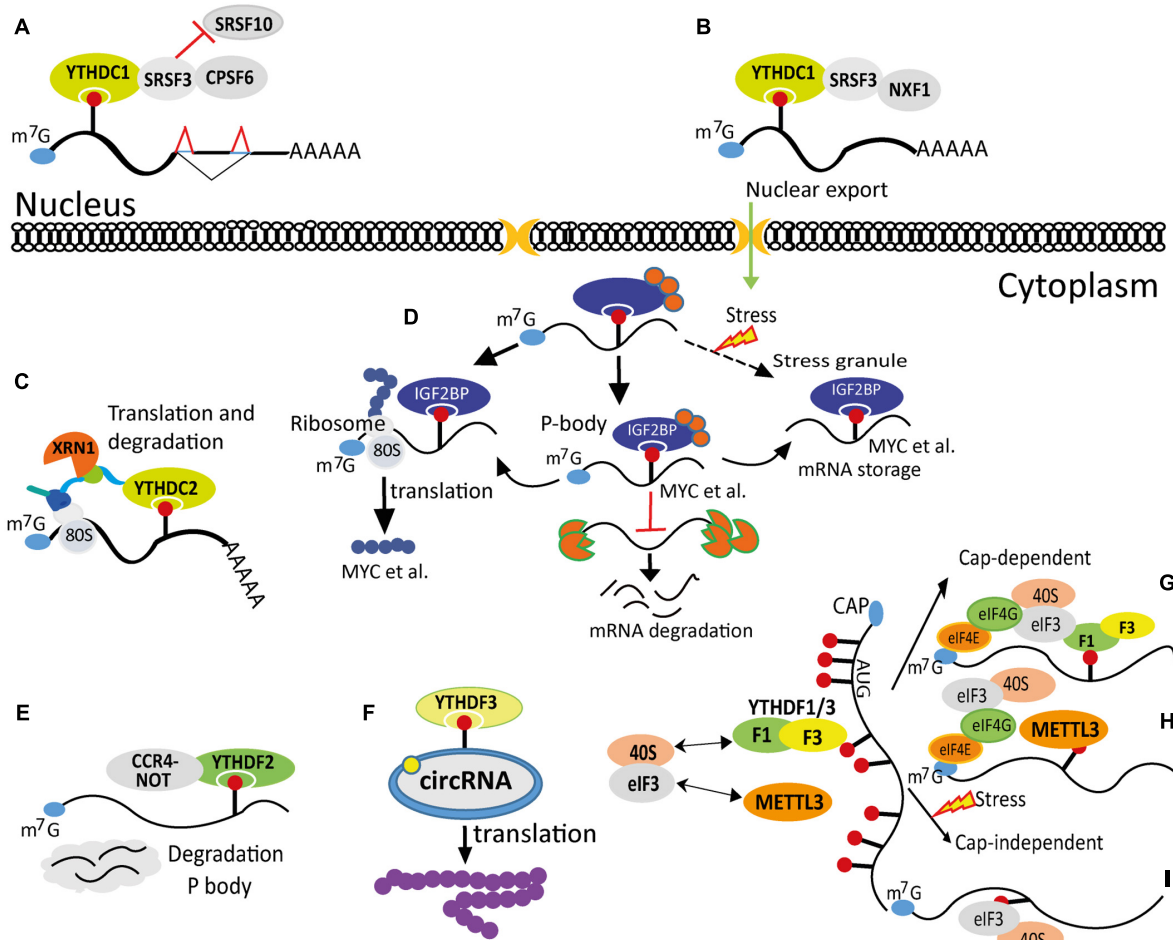


FIGURE 2 | Effects of N^6 -methyladenosine (m^6A) methylation on messenger RNA (mRNA) fate. **(A)** m^6A modification regulates mRNA splicing and polyadenylation via YTHDC1 and its associating factors SRSF3, SRSF10, and CPSF6. **(B)** m^6A modulates mRNA nuclear export through YTHDC1, SRSF3, and NXF1. **(C)** m^6A regulates mRNA translation and stability via YTHDC2-mediated recruitment of the ribosome and the XRN1 exoribonuclease. **(D)** m^6A marks are bound by IGFBPs, which can regulate a subset of mRNA translation, decay in P-body, and storage in stress granules. **(E)** m^6A modification regulates mRNA degradation in P-body through associating with the YTHDF2-CCR4-NOT complex. **(F)** m^6A marks on circRNA modulate its translation via recruiting YTHDF3. **(G)** m^6A marks recruit YTHDF1/YTHDF3 to enhance translation in a Cap-dependent manner. **(H)** METTL3 serves as an m^6A reader and increases translation via recruiting translation initiation complex independent of its methyltransferase activity. **(I)** m^6A directly binds to eIF3 and increases translation in a Cap-independent manner.

degradation of MAT2A mRNA, although the detailed mechanism is not known (Shima et al., 2017).

On the other hand, YTHDC2 is a multi-domain protein and mainly localized in the cytoplasm, but it is also highly expressed in perinuclear compartment. It prefers to bind m^6A -containing RNAs through the YTH domain and enhances RNA degradation. Meanwhile, it also enhances m^6A -modified mRNA translation efficiency (Kretschmer et al., 2018).

Human YTHDF1–3 proteins contain a YTH domain in the C-terminus and a low-complexity domain in the N-terminus. These three members of the YTHDF family share high sequence identity and similarity (65%–80%) (Li et al., 2014; Wang X. et al., 2014; Hazra et al., 2019). As a characterized m^6A modification reader, the human YTHDF2 binds over 3,000 transcripts primarily in their 3' UTRs and around the stop codon. The binding of YTHDF2

leads to degradation of the bound mRNAs in cytoplasmic processing bodies (P-bodies). Knockdown of YTHDF2 leads to an accumulation of m^6A -containing mRNAs (Wang X. et al., 2014). YTHDF2 was also found to associate with CNOT1, the scaffolding component of the CCR4-NOT mRNA deadenylation complex (Figure 2E). This interaction is required for YTHDF2 to localize in P-bodies (Du et al., 2016). Therefore, the main function of YTHDF2 is to control the degradation of m^6A -modified mRNAs (Kang et al., 2014; Hazra et al., 2019).

Unlike YTHDF2, YTHDF1 does not induce the degradation of associated m^6A -containing mRNAs. Instead, but arguably, YTHDF1 increases the translation efficiency of associated mRNAs (about 1,200) in an m^6A -dependent fashion (Wang et al., 2015). This function of YTHDF1 is further supported by the work of Wu et al. (2019), who showed that YTHDF1 targets

m⁶A-modified *Jak2* and regulates its translation (Wu et al., 2019). Recently, another cytoplasmic m⁶A reader protein YTHDF3 was found to interact with YTHDF1 to promote translation, whereas YTHDF3 interacts with YTHDF2 to reinforce mRNA decay (Li A. et al., 2017; Shi et al., 2017). Furthermore, biochemical studies showed that YTHDF3 shares greater than 50% of common m⁶A-modified mRNA targets with YTHDF1 and also with YTHDF2 (Li A. et al., 2017; Shi et al., 2017). In addition, YTHDF3 might also function as m⁶A-modification reader independent of YTHDF1 and YTHDF2 under certain conditions. Oxidative stress induces specific m⁶A modifications in a set of transcripts, and the binding of YTHDF3 to the modifications triggers the mRNA–YTHDF3 complex localization in the stress granules, but without much influence on YTHDF1 and YTHDF2 (Anders et al., 2018). Importantly, YTHDF3 can also enhance translation independent of METTL3-mediated m⁶A modification. For example, YTHDF3 functions together with eIF4G2 and Poly(A)-binding protein 1 (PABP1) to promote the translation of forkhead box protein O3 (FOXO3) (Zhang Y. et al., 2019).

Interestingly, very recent studies have shown evidence indicating functional redundancy of the YTHDFs during mRNA degradation and cellular differentiation. As such, it is only when all three YTHDF homologs (YTHDF1–3) are ablated that mRNA stability and cell differentiation regulation become evident (Kontur et al., 2020; Zaccara and Jaffrey, 2020). This may partly stem from the observations that all three YTHDFs are similar in sequence characteristics and usually have common mRNA binding targets (Zaccara and Jaffrey, 2020). Yet, it has been reported that probably due to variation in its expression, YTHDF2 dominates the m⁶A reader function of all the YTHDFs (Lasman et al., 2020). It was also unraveled that YTHDFs are unable to induce translation in HeLa cells (Zaccara and Jaffrey, 2020). While these new findings present a unified model seeking to define the regulatory functions of YTHDFs in m⁶A modification, they provoke questions that need to be addressed to reconcile the discrepancy between the recent findings and previous observations with respect to the precise role of YTHDFs in mRNA translation.

Non YTH Domain Containing N⁶-Methyladenosine Readers (eIF3, IGF2BPs, HuR, FMRP, HNRNP Proteins, and PRRC2A)

Meyer et al. (2015) characterized the function of eIF3 as an m⁶A reader. eIF3 is preferentially recruited by the m⁶A-modified mRNA over unmethylated mRNA (Meyer et al., 2015). It was shown that about 35% of m⁶A marks in the 5' UTR are also eIF3-binding sites. Depletion of m⁶A through METTL3 loss-of-function decreased the translation of m⁶A-modified mRNA in the 5' UTR, but not the mRNAs bearing m⁶A marks elsewhere (Meyer et al., 2015). Notably, one of the two modes of m⁶A-mediated Cap-independent translation is through direct association of m⁶A in the 5' UTR and eIF3 (Figure 2I; Meyer et al., 2015), while the other mode involves YTHDF1 association with m⁶A mark followed by delivery of eIF3 to the 5' UTR (Figure 2G; Wang et al., 2015). This indicates the correlation between the 5' UTR m⁶A and translation and highlights the involvement of eIF3 in the regulation of mRNA translation.

Currently, it is not known what the detailed mechanisms are in determining the mode of eIF3–5' UTR association.

Insulin-like growth factor-2-binding proteins (IGFBPs), including IGFBP1–3, have been reported as RNA-binding proteins (Bell et al., 2013). Recently, it was demonstrated that IGFBP1–3 bind m⁶A-modified mRNAs with a three-fold to four-fold greater affinity than the m⁶A-unmodified mRNAs (Huang H. et al., 2018). By means of RIP-Seq or PAR-CLIP-Seq, it was found that IGFBP1–3 share 55%–70% RNA targets with preference for binding to the “UGGAC” consensus motif, e.g., *MYC*, *FSCN1*, and *TK1* (Huang H. et al., 2018). Knockdown of METTL14, a critical component of the methylation machinery, dramatically undermined IGFBP binding. Interestingly, knockdown of IGF2BPs reduces mRNA stability (Huang H. et al., 2018). Consistently, IGFBPs were found to associate with three mRNA stabilizing factors, including HuR, MATR3, and PABP1, which can support IGFBPs in stabilizing their mRNA targets (Huang H. et al., 2018).

HuR is an RNA-binding protein with multiple molecular functions. It was first described as a stabilizer of ARE-containing mRNAs (Fan and Steitz, 1998; Peng et al., 1998). It is also known to enhance translation, although it can also exert translation suppression (Hinman and Lou, 2008; Abdelmohsen and Gorospe, 2010). This portrays HuR as both a reader and anti-reader of m⁶A (Dominissini et al., 2012; Wang Y. et al., 2014). However, the underlying mechanism that makes m⁶A modification sites to recruit or block HuR binding is unknown. We think that a sequence-dependent context may be at play in determining the function of HuR in m⁶A interaction. This speculation remains to be investigated.

FMR1 (also known as FMRP1) is an RNA-binding protein and known to associate with hundreds of transcripts to decrease their translation. It binds to m⁶A-modified mRNA in an RNA sequence context-dependent manner. FMR1 selectively binds to the m⁶A marks associated with GGACU RNA sequence (Edupuganti et al., 2017). Bioinformatic analysis revealed that FMR1 and YTHDF1 shared an abundant set of common m⁶A-modified mRNAs, indicating that FMR1 might compete with YTHDF1 for binding of m⁶A-modified mRNAs to downregulate translation (Ascano et al., 2012; Wang et al., 2015). It is possible that the mechanism may underlie the previously reported regulatory function on the translation of mRNA targets.

Heterogeneous nuclear ribonucleoproteins (hnRNPs: hnRNP A2/B1, hnRNP C, and hnRNP G) are RNA-binding proteins that play important roles in pre-RNA processing (Dominissini et al., 2012; Alarcón et al., 2015; Liu et al., 2015, 2017; Xiao et al., 2016). Alarcón et al. (2015) discovered that hnRNP A2/B1 interacts with a group of m⁶A-modified RNAs in the nucleus and regulates their splicing in a comparable pattern as for METTL3. However, the binding of hnRNP A2/B1 to m⁶A is likely indirect and may require an hnRNP C-mediated switch mechanism to do so (Wu et al., 2018). hnRNP C can read m⁶A-modified hairpin and m⁶A-containing RNAs. m⁶A-modification leads to a change in the regional RNA structure and increases the binding of hnRNP C (Liu et al., 2015). Consistently, general reduction in m⁶A marks due to METTL3/L14 knockdown eliminates the association of hnRNP C to the aforementioned

m⁶A-mediated RNA structural modification (Liu et al., 2015). Furthermore, HNRNPG is known to bind m⁶A-modified lncRNA through its C-terminal low-complexity domain (LCD), indicating that LCD domain might be used by some other readers to bind to m⁶A modification (Liu et al., 2017).

Recently, PRRC2a was reported as an m⁶A modification reader (Wu et al., 2019). Through RIP-seq and m⁶A-seq, it was identified that PRRC2a binding peaks within over 2,800 genes in brain samples, and PRRC2a mainly binds to the consensus motif UGGAC in m⁶A-modified transcripts (Wu et al., 2019). PRRC2A was found to be associated with YTHDF2 in granule-like organelles, which may be involved in the regulation of PRRC2A involvement in *Olig2* mRNA stability (Wu et al., 2019). However, since PRRC2A has low tissue expression specificity, it is unclear whether PRRC2A serves as an m⁶A modification reader in other tissues.

Reader Function of METTLs

Besides its role as an m⁶A writer, METTL3 can also bind to m⁶A-modified mRNAs to act as a reader. It was found that METTL3 regulates the translation of some oncogenic m⁶A-modified mRNAs independent of its methyltransferase activity through eIF3 recruitment to the translation initiation complex (Lin et al., 2016). A study from the same group identified a physical interaction between m⁶A-bound METTL3 near the stop codon and eIF3h, providing a mechanism to explain how METTL3 can enhance translation (Choe et al., 2018). The methyltransferase METTL16 also serves as an m⁶A reader in a certain context. When SAM concentrations become low, METTL16 remains bound to m⁶A-modified MAT2A in its 3' UTR hp1 site to enhance MAT2A splicing, resulting in increased MAT2A levels in the cytosol. On the contrary, when SAM levels are high, METTL16 methylates MAT2A and facilitates intron retention (Pendleton et al., 2017).

Deposition of N⁶-Methyladenosine Modification During Transcription

Mechanistically, how m⁶A modification of transcripts is carried out needs elucidation. A recent study uncovered an insightful detail in the installation of m⁶A. Specifically, it was found that H3K36me3 cooperates with METTL3/METTL14 to deposit m⁶A on mRNA (Huang et al., 2019). The study showed that H3K36me3 physically interacts with METTL14, thus recruits the m⁶A methylation machinery to RNA Pol II, and allows the m⁶A methylation machinery to effect m⁶A modification during transcription. Decreasing the level of H3K36me3 through loss-of-function of SETD2, the specific enzyme that converts H3K36me2 or H3K36me0 to H3K36me3, significantly led to the reduction in m⁶A level on RNAs, mimicking the impact of depletion of individual m⁶A writer complex components (Huang et al., 2019).

Impact of N⁶-Methyladenosine Modification on Gene Regulation

The reversible modification of m⁶A exerts functional impact on several aspects of mRNA metabolism, including nuclear

export, polyadenylation, splicing, degradation, and translation (Figure 2). By these means, the m⁶A methylome affords an additional level of gene expression regulation to sculpt the transcriptome (Fu et al., 2014).

N⁶-Methyladenosine Modification Regulates mRNA Splicing

Some factors involved in m⁶A modification of mRNA are known to interact with pre-mRNA splicing factors (SRSFs), indicating a possible role for m⁶A in mRNA splicing (Zhao et al., 2014; Xiao et al., 2016). It has been demonstrated that enrichment of m⁶A modification promotes recruitment of SRSF2 and leads to enhanced exon inclusion of target mRNA (Zhao et al., 2014). It has been further suggested that the m⁶A reader YTHDC1 regulates the association of m⁶A and SRSFs. Indeed, m⁶A-bound YTHDC1 enhances the recruitment of SRSF3 that favors exon inclusion but blocks the recruitment of SRSF10, an exon skipping-related splicing factor (Xiao et al., 2016). Moreover, hnRNPs may also be involved in the regulation of RNA splicing (Liu et al., 2015, 2017). For example, the modification of m⁶A on pre-mRNA favors the binding of hnRNPC (Liu et al., 2015), which could further facilitate splicing through its known function in repressing exon inclusion (Zarnack et al., 2013). Therefore, it is possible that perturbation of the m⁶A machinery components can impair mRNA alternative splicing. This idea is especially supported by the observation that knockdown of METTL3 can antagonize the association of SRSF2 or SRSF3 with m⁶A-modified pre-mRNAs (Zhao et al., 2014; Xiao et al., 2016), and facilitates the expression of the long isoform of MyD88 (MyD88L) via exon skipping attenuation (Feng et al., 2018). Additional evidence is also based on the essential role played by METTL16 in MAT2A-mediated pre-mRNA alternative splicing (Pendleton et al., 2017).

N⁶-Methyladenosine Controls Alternative Polyadenylation

Ke et al. (2015) found that m⁶A modification peaks in the 3' UTR, especially for transcripts that use alternative polyadenylation (APA), and longer last exons exhibit a higher m⁶A density. By comparing the m⁶A density of thousands of mRNA UTRs from liver and brain tissues, it was observed that greater amount m⁶A marks in the last exons are linked to the usage of more distal polyA sites. Indeed, global reduction of m⁶A levels via triple knockdown of METTL3, METTL14, and WTAP changed the polyA sites of one-sixth of the examined 661 mRNAs and promoted the usage of proximal APA sites, indicating that some m⁶A marks inhibit proximal polyadenylation (Ke et al., 2015).

Recently, a mechanism through which m⁶A controls alternative polyadenylation was proposed. VIRMA (Figure 1A) was found to interact with polyA cleavage factors F5 and CPSF6 (Yue et al., 2018). Consistent with an earlier report, knockdown of METTL3 or VIRMA was found to encourage the usage of distal APA sites, thus lengthening the 3' UTR of m⁶A-rich mRNAs. In contrast, CPSF5 knockdown elicits an opposite effect on the length of the 3' UTR of m⁶A-marked mRNAs (Yue et al., 2018).

N⁶-Methyladenosine Promotes Nuclear Export

Considerable amount of nuclear export of mRNAs is regulated by the THO/TREX complex and the nuclear export factor heterodimer NXF1/P15 (Lesbirel and Wilson, 2019). Evidence is accumulating for the role of m⁶A modification in mRNA nuclear export. Knockdown of METTL3 resulted in delayed nuclear export of specific mRNAs of clock genes (Fustin et al., 2013), indicating the requirement of m⁶A methylation for specific mRNA nuclear export. Conversely, knockdown of ALKBH5 increased the cytoplasmic accumulation of polyA mRNAs (Zheng et al., 2013). Moreover, VIRMA was observed to interact with the mRNA export factor ALYREF, and its downregulation led to defective mRNA export (Masuda et al., 2005).

Interestingly, several TREX components associate with the components of the core m⁶A machinery (METTL3-METTL14-WTAP-VIRMA), and TREX also enhances the association of m⁶A reader YTHDC1 with the mRNA. Knockdown of YTHDC1 also resulted in reduced nuclear export of specific mRNAs (Lesbirel et al., 2018). Taken together, the abovementioned literature demonstrates that m⁶A modification factors promote mRNA nuclear transport through physical interaction with the mRNA transport machinery.

N⁶-Methyladenosine Enhances mRNA Degradation

Numerous recent studies suggest that impaired m⁶A writer complex function reduces m⁶A modification levels and raises mRNA stability, indicating that m⁶A methylation drives mRNA degradation (Batista et al., 2014; Schwartz et al., 2014; Wang X. et al., 2014; Wang Y. et al., 2014; Park et al., 2019). Mechanistically, m⁶A-containing mRNA recruits YTHDF2, which is followed by the translocation of the YTHDF2-mRNA complex from the translation machinery to P-bodies, leading to the degradation of YTHDF2-targeted mRNA. As a result, mRNA targets have increased half-life following YTHDF2 knockdown (Wang X. et al., 2014). It has been clearly demonstrated that YTHDF2 enhances m⁶A-modified mRNA decay through recruiting CCR4-NOT deadenylase complex *via* the N-terminus of YTHDF2 and reveals an underlying mechanism by which YTHDF2 regulates degradation of m⁶A-modified mRNAs (Du et al., 2016).

In a recent study, it was reported that some m⁶A-modified mRNAs interact with YTHDF2 to undergo decay in an RNase P/MRP-dependent manner and in which HRSP12 serves as a bridge between YTHDF2 and RNase P/MRP (Park et al., 2019). The interaction of human YTHDF2 and HRSP12 was first hinted by the association between their respective yeast homologs Pho92 and Mmf1 (Krogan et al., 2006). It was found in an immunoprecipitation experiment that HRSP12 links YTHDF2 and RNase P/MRP and that the N-terminus of YTHDF2 is required to interact with HRSP12 (Park et al., 2019). Of note, the subset of m⁶A-modified mRNAs, whose decay depends on YTHDF2-HRSP12-RNase P/MRP complex, contains a specific HRSP12-binding motif proximally upstream of the YTHDF2-binding motif, while the RNase P/MRP cleavage site is downstream and close to the YTHDF2-binding motif (Park et al., 2019). Therefore, this study discloses at least two

mechanisms involved in the degradation of YTHDF2-associated m⁶A RNAs: HRSP12-RNase P/MRP-dependent and CCR4-NOT complex-dependent.

N⁶-Methyladenosine Modulates Translation

The m⁶A reader YTHDF1 enhances translation efficiency *via* interaction with eIF3A/eIF3B, and the YTHDF1-regulated translation likely hinges on eIF4G-dependent loop formation (Wang et al., 2015). According to Meyer et al. (2015), 5' UTR m⁶A elevates cap-independent translation through recruiting the 43S complex to form 48S initiation complex in the absence of the cap-associating complex, eIF4F. This mechanism is important for cells to bypass 5' cap-binding factors to enhance translation under stress conditions (Meyer et al., 2015). Moreover, heat stress-induced cytoplasmic-to-nuclear translocation of YTHDF2 is required for maintaining 5' UTR m⁶A levels *via* competing for binding of the demethylase FTO to m⁶A sites, which further promotes cap-independent translation (Zhou et al., 2015). YTHDF1 preferentially binds to m⁶A marks in 3' UTR of the oncogene *CDCP1* mRNA and promotes translation by increasing the amount of polysome-bound (translationally active) *CDCP1* transcripts (Yang et al., 2019).

Of note, METTL3 is also involved in m⁶A-enhanced mRNA translation through its role as an m⁶A reader in several ways. It promotes mRNA translation *via* physical association with the translation initiation complex (Lin et al., 2016). It was found that promoter-associated METTL3 regulates m⁶A methylation inside the coding region and improves mRNA translation through relief of ribosome stalling (Barbieri et al., 2017).

Besides promoting translation efficiency, m⁶A modification also plays an important role in regulating alternative translation (Zhou J. et al., 2018). It has been reported that widespread alternative translation occurs under various nutrient conditions, but the underlying mechanism is unclear (Gao et al., 2015). Recently, Zhou J. et al. (2018) found that m⁶A modification in the 5' UTR modulates the selection of start codon globally, hence driving alternative translation. As representative examples, *Atf4* depends on decreased m⁶A modification of the upstream open reading frame 2 (uORF2) to improve the translation of the major isoform, and *Gadd45g* heightens the translation of the major isoform by lowering the m⁶A modification of the 5' UTR (Zhou J. et al., 2018).

N⁶-Methyladenosine Methylation Increases the Phase Separation Capacity of mRNA

Only until recently has it become clearer how m⁶A modification drives mRNA fate and why the consequence of m⁶A modifications can vary in various scenarios. According to Ries et al. (2019), the m⁶A readers, YTHDF1–3, experience liquid–liquid phase separation (LLPS). The mRNAs with multiple m⁶A marks serve as a scaffold to bind with YTHDF readers *via* their low-complexity regions (LCRs). The mRNA-YTHDF complexes are then transported into various phase separators, like P-bodies, stress granule, and neuronal granules. The study suggests that the number and allocation of m⁶A modifications in mRNAs remodel the transcriptome of different phase-separated

compartments. The efficacy of m⁶A modification-dependent modulation of an mRNA is likely governed by signals regulating the ability of YTHDF protein involved in LLPS formation (Ries et al., 2019).

N⁶-METHYLADENOSINE MODIFICATION PROMINENTLY REGULATES BRAIN DEVELOPMENT AND FUNCTION

Evidence for the role of m⁶A signaling in modulating the development of the brain and its functions has accumulated in recent years, and the quest for extending the frontier is of great interest. Several investigations have revealed that the various factors that come together to form the m⁶A methylation machinery exert notable effect(s) on specific aspects of brain morphogenesis to permit optimal neural function, as summarized in **Table 1**. Conversely, the dysregulation of the m⁶A methylation machinery is known to elicit perturbations in the neural transcriptome, which have implications for defective development and dysfunction of the brain. The integrity of the m⁶A machinery functionality is of high priority in cells to the extent that simply ablating its cofactors can have significant consequences for brain development disturbance, as exemplified by the importance of Exosc10-mediated regulation of mRNA stability in forebrain development (Ulmke et al., 2021). The sections below discuss how specific factors associated with the m⁶A methylation machinery drive neural development, functional adaptation, and plasticity of the brain (**Figure 3**).

N⁶-Methyladenosine Modification Is Indispensable for Neurogenesis in the Brain

Neurons are produced through the process of neurogenesis, which entails specification and proliferation of NSCs, and the differentiation of such neural progenitors into neuroblasts, which undergo maturation to become functional neurons. It has been shown that the dynamic addition of m⁶A to gene transcripts in the multipotent NSCs greatly influences cortical neuroprogenitor competence and the generation of neurons during brain development (Yao et al., 2016; Boles and Temple, 2017; Yoon et al., 2018; Zhou H. et al., 2018; Rockwell and Hongay, 2019). Dysregulation of writers, erasers, and readers of m⁶A has been reported to cause notable perturbations in the cell cycle progression, proliferation, and differentiation of NSCs in the developing and adult brain.

Effect of N⁶-Methyladenosine Writers on Neurogenesis in the Brain

So far, it has been shown that ablation of the m⁶A writer METTL3 or its cofactor METTL14 in cortical neuroepithelium or isolated cortical NSC results in prolonged cell cycle dynamics of cortical neuroprogenitors and their precocious differentiation into neuronal or neurogenic cells (Batista et al., 2014; Yoon et al., 2017; Wang Y. et al., 2018). Detailed analysis through m⁶A sequencing revealed that gene transcripts involved in the

cell cycle of neural cells, production of neurons, and neuronal differentiation are enriched with m⁶A tags. Interestingly, the decay of such mRNAs is promoted in the absence of METTL3 and METTL14 (Yoon et al., 2017), meaning that METTL3 and METTL14 are key players in driving neurogenesis *via* the stabilization of gene transcripts critical for neurogenesis in the brain. For example, loss of m⁶A due to deletion of METTL3 in mouse cerebellum resulted in overt hypoplasia partly attributable to apoptosis of cerebellar granule cells (Wang C. X. et al., 2018). Key downstream effects of m⁶A on genes important for neurogenesis include the modulation of histone modification in the promoter environment of NSC proliferation- and differentiation-related gene loci (Wang Y. et al., 2018). In the absence of METTL14, the transcription-suppressing histone mark H3K27me3 is upregulated on genes involved in cell proliferation, whereas differentiation-related genes show an increase in the transcription activation histone mark H3K27ac when METTL14 is deficient (Wang Y. et al., 2018). Lack of METTL3 in the developing brain can also cause the aforementioned histone alterations, at least in terms of H3K27me3 enhancement, which can cause transcription repression (Chen J. et al., 2019). This is possible because in the absence of METTL3, which leads to a reduction in m⁶A levels, the polycomb repressor complex becomes hyperactive due to derepression of its core methyltransferase factor Ezh2 (Chen J. et al., 2019).

RBM15, a core component of the m⁶A writer complex (**Figure 1A**), is a potential regulator of cortical neurogenesis due to its distinctive expression in the cortical germinative zone and cortical plate of the developing mouse cortex (Xie et al., 2019). Knockdown of RBM15 in neurons *in vitro* promoted endogenous expression of the chromatin remodeling factor BAF155 (Xie et al., 2019), which is a known key regulator of cortical development (Nguyen et al., 2016, 2018; Narayanan et al., 2018). This profound effect can be linked to a significant reduction in cellular levels of m⁶A due to the inactivation of RBM15 (Knuckles et al., 2018). However, overexpression of RBM15 *in vivo* was found to promote delamination of radial glial cells in the cortical ventricular zone by suppressing the expression BAF155 and, hence, BAF155-dependent gene expression program supportive for cortical development (Xie et al., 2019). The role of RBM15 in cortical neurogenesis further highlights the contribution of m⁶A methyltransferase in brain development.

N⁶-Methyladenosine Erasers Regulate Neurogenesis in the Brain

Erasers of the m⁶A mark (FTO and ALKBH5) can also exert a regulatory effect on the process of neurogenesis given their prominent expression in neurons (Li L. et al., 2017; Yoon et al., 2017; Spychala and Ruther, 2019; Du et al., 2020). Whereas FTO displays the highest expression level late in brain neurogenesis (Li L. et al., 2017; Yoon et al., 2017), ALKBH5 expression decreases in the course of brain development (Du et al., 2020). This may have implications for their roles in the spatiotemporal regulation of neurogenesis during brain development. Indeed, it was reported that FTO deficiency in the adult mouse brain induces signal transducer and activator of transcription (STAT)3

TABLE 1 | m⁶A mRNA methylation factors and their role in brain development and function.

Effector	Experimental manipulation	Phenotype	Mechanisms	References
Neurogenesis				
METTL3	<i>Mettl3^{fl/fl}; Nestin-Cre</i>	Prolongation of the cell cycle of RGCs and protraction of embryonic cortical neurogenesis	m ⁶ A depletion caused increased stability of NSC transcripts	Yoon et al. (2017)
METTL14	<i>Mettl14^{fl/fl}; Nestin-Cre</i>	Reduced NSC proliferation and precocious NSC differentiation; loss of late-born neurons during cortical neurogenesis	Stabilization of CBP and p300 transcripts; H3K27me3-mediated transcription suppression of NSC proliferation genes; upregulation of H3K27ac in differentiation-related genes	Wang Y. et al. (2018)
RBM15	OE of <i>RBM15</i>	NSC delamination	Suppression of BAF155-dependent gene expression	Xie et al. (2019)
FTO	KO of <i>FTO</i>	Decrease in adult NSC proliferation and defective hippocampal neurogenesis	Impairment of BDNF and MAPK signaling	Li L. et al. (2017); Spychala and Ruther (2019)
YTHDF2	<i>Ythdf2^{fl/fl}; Cre (ubiquitously)</i>	Decreased proliferation and differentiation capabilities of NSCs; less Tbr2+ bIPs; Reduced CP thickness	Promotion of m ⁶ A-dependent degradation of neurodevelopment-related transcripts	Li M. et al. (2018)
FMRP	KO of <i>Fmr1</i>	Nuclear retention of neurogenic mRNAs; prolonged cell cycle progression in the postnatal mouse brain	Unknown	Edens et al. (2019)
Exosc10	<i>Exosc10^{fl/fl}; Foxg1-Cre; Emx1-Cre</i>	Apoptosis-mediated cortical agenesis	Mediates degradation of <i>Bbc3</i> and <i>Aen</i> mRNAs, which are effectors of apoptosis	Ulmke et al. (2021)
Gliogenesis				
METTL14	<i>Mettl14^{fl/fl}; Olig2-Cre</i>	Decrease in oligodendrocytes maturation; cortical hypomyelination	Alters alternative splicing and expression of <i>Nfasc155</i>	Xu et al. (2020)
METTL14	<i>Mettl14; Nestin-Cre</i>	Reduced number of s100β+ astrocytes	Unknown	Yoon et al. (2017)
FTO	<i>FTO^{fl/fl}; Olig2-Cre; Nestin-Cre</i>	Loss of OPCs and Sox10+ cells; cortical hypomyelination	Promotes <i>Olig2</i> mRNA degradation	Wu et al. (2019)
PRRC2A	<i>Prrc2a; Nestin-Cre; Olig2-Cre</i>	Loss of OPCs and mature oligodendrocytes; cortical hypomyelination	Gene targeting of <i>Prrc2a</i> by <i>olig2</i> mRNA	Wu et al. (2019)
PRRC2A	<i>Prrc2a; Nestin-Cre;</i>	Reduced proliferation capacity and number of astrocytes	Competitive expression with YTHDF2	Wu et al. (2019)
Axonogenesis, dendritogenesis, synaptic plasticity				
YTHDF1 YTHDF3	KD of <i>Ythdf1</i> and <i>Ythdf3</i>	Abnormal dendritic spine morphology	Inhibition of <i>Apc</i> mRNA translation	Merkurjev et al. (2018)
METTL14	<i>Mettl14^{fl/fl}; D1R-Cre</i>	Abnormal excitability of striatal neurons	Unknown	Koranda et al. (2018)
FTO	KO of <i>FTO</i>	Defective synaptic plasticity	Unknown	Hess et al. (2013)
YTHDF1 YTHDF3	KD of <i>Ythdf1</i> and <i>Ythdf3</i> ; CRISPR/Cas9-based KO of <i>Ythdf1</i>	Suppression of neuronal excitability	Not clear	Merkurjev et al. (2018); Shi et al. (2018)
Learning and behavior				
METTL3	OE of <i>METTL3</i>	Improved long-term memory consolidation	Unknown	Zhang Z. Y. et al. (2018)
METTL14	<i>Mettl14^{fl/fl}; D1R-Cre</i>	Impaired striatum-mediated behavior patterns	Unknown	Koranda et al. (2018)
FTO	CRISPR/Cas9 or shRNA-mediated KD of <i>FTO</i>	Learning disabilities; defective memory processing and verbal fluency	Not clear	Ho et al. (2010); Benedict et al. (2011); Widagdo et al. (2016); Li L. et al. (2017); Walters et al. (2017); Sun et al. (2019)
PRRC2A	<i>Prrc2a; Nestin-Cre; Olig2-Cre</i>	Cognitive defects due to cortical hypomyelination	Unknown	Wu et al. (2019)
YTHDF1	CRISPR/Cas9-based KO of <i>Ythdf1</i> ; KD of <i>Ythdf1</i>	Defective long-term potentiation and synaptic transmission in hippocampus; behavioral defects	Unknown	Shi et al. (2018)

(Continued)

TABLE 1 | Continued

Effector	Experimental manipulation	Phenotype	Mechanisms	References
Circadian clock				
METTL3	KD of <i>Mettl3</i>	Elongation of circadian period	Defective processing of <i>Per2</i> and <i>Arntl</i> (clock genes) mRNAs	Fustin et al. (2013)
Stress response				
FTO	<i>FTO</i> ^{fl/fl} ; <i>Camk2a-Cre</i> ; <i>Nex-CreERT2</i>	Reduced ability to cope with stress	Unknown	Engel et al. (2018)
METTL3	<i>Mettl14</i> ^{fl/fl} ; <i>Camk2a-Cre</i> ; <i>Nex-CreERT2</i>	Reduced ability to cope with stress	Unknown	Engel et al. (2018)

KO, knockout; KD, knockdown; OE, overexpression; fl/fl, double floxed; RGCs, radial glial cells; OPCs, oligodendrocyte precursor cells; NSCs, neural stem cells; CRISPR, clustered regularly interspaced short palindromic repeats; m⁶A, N⁶-methyladenosine.

pathway activation *via* its modulators platelet-derived growth factor receptor (PDGFR) α and suppressor of cytokine signaling (SOCS)5 in an m⁶A-dependent manner (Cao et al., 2020). As a result, a transient increase in the proliferation and differentiation of adult NSCs was observed in the FTO mutant brain, with implications for adult neurogenesis inhibition in the long term (Cao et al., 2020). It was also observed that FTO deletion in adult mouse brain impairs brain-derived neurotrophic factor (BDNF) and mitogen-activated protein kinase (MAPK) signaling pathways, leading to a reduction in adult NSC proliferation and neurogenesis in the hippocampal formation (Li L. et al., 2017; Spychala and Ruther, 2019). Although these studies report diverging effects of FTO loss on adult NSC proliferation, they both show a resultant effect of adult neurogenesis reduction. We think that, while being mindful of the low level of FTO expression in the early developing cortex, conducting an investigation on how FTO regulates corticogenesis in the course of development may lend clarity to how it mechanistically impacts cortical neurogenesis.

Notable N⁶-Methyladenosine Readers in Cortical Neurogenesis

Protein factors that act as readers of the m⁶A mark have also been shown to have a profound effect on neurogenesis in the brain. For instance, the m⁶A reader YTHDF2 has been reported to be indispensable for corticogenesis in mouse. Conditional knockout of YTHDF2 in the mouse neocortical neuroepithelium resulted in a reduction in the proliferation and differentiation of the *Ythdf2*^{-/-} neuroprogenitor cells (Li M. et al., 2018). This phenotype may have mechanistic underpinnings, including abnormal upregulation of genes that inhibit the JAK-STAT signaling pathway, due to increased stability of such gene transcripts in the absence of YTHDF2 (Li M. et al., 2018). Yet, it seems that the induction of neural fate in pluripotent stem cells requires downregulation of YTHDF2, leading to the stability and expression of neural gene transcripts (Heck et al., 2020). We are of the opinion that the functional consequence of the m⁶A reading by YTHDF2 may be contextually variable along the cortical development axis such that reduced dosage may support neural cell fate specification, whereas its increased activity/expression is necessary for later cortical neurodevelopment.

Another m⁶A reader, FMRP, was identified to be critical for neural progenitor cell proliferation. Mice lacking *Fmr1* displayed prolonged cell cycle progression. As a result, proliferation of neural progenitors extended into postnatal stages of brain development (Edens et al., 2019). Of note, it was observed that nuclear export of m⁶A-modified neurogenic mRNAs readable by FMRP is defective, leading to retention of such neurodifferentiation gene transcripts in the nucleus of the *Fmr1*-deficient neural progenitor cells (Edens et al., 2019). Lastly, the m⁶A reader protein Imp (IGF2BP) was identified as a key regulator of NSC proliferation rate through the stabilization of *Myc* mRNA in *Drosophila* brain neuroblasts (Samuels et al., 2020).

Together, the m⁶A machinery has been identified to play critical roles in brain morphogenesis by regulating the proliferation of neural progenitor cells and the production of neurons. As such, hypomethylation due to METTL3 or METTL14 deficiency and aberrant m⁶A reading or erasure in the embryonic or adult brain can precipitate phenotypes, including defective transcriptional prepatterning, abnormal neuroprogenitor pool, impaired neurogenesis, and cortical hypoplasia (Yoon et al., 2017), which can engender deficits in brain structure and function.

N⁶-Methyladenosine Signaling Is Essential for Gliogenesis in the Brain

The process of generating glial cells constitutes gliogenesis. Brain neuroglia include astrocytes and oligodendrocytes, which are derived from the neuroepithelium. During cortical development, a switch from neurogenesis to gliogenesis coincides with a decrease in m⁶A modification of proneural genes (Donega et al., 2018). Although m⁶A enrichment in glial cells is less than that observed in neurons (Chang et al., 2017), a few studies have uncovered the importance of the m⁶A methylome in brain gliogenesis, at least for astrocyte production (astrogenesis) and oligodendrocyte generation (oligodendrogenesis).

Regulation of Glia Production in the Brain by an N⁶-Methyladenosine Writer-Related Factor

It was observed that loss of METTL14-mediated m⁶A writing in the mouse cortex leads to hypomyelination that can be linked to a reduction in the number of (mature) oligodendrocytes (Xu et al., 2020). The loss of oligodendrocytes caused by the absence

of METTL14 in the brain likely did not emanate from abnormal specification or proliferation of oligodendrocyte precursor cells (OPCs) (Xu et al., 2020). Notably, the transcriptome of OPCs and oligodendrocytes is altered following METTL14 deletion, with possible impact on gene expression programs critical for oligodendrocyte lineage progression (Xu et al., 2020). Lack of METTL14 has also been reported to disrupt astrogenesis. Indeed, s100 β -expressing astrocytic progenitors were found to be reduced in the METTL14 knockout mouse cortex at postnatal stage 5 (Yoon et al., 2017). It would be interesting to investigate whether other m⁶A writer-related factors, including METTL3, have roles to play in cortical gliogenesis.

The N⁶-Methyladenosine Eraser FTO Regulates Glia Production in the Brain

m⁶A-mediated RNA methylation dynamics under the guild of FTO is known to influence oligodendrogenesis *via* modulation of the half-life of *Olig2* mRNA (Wu et al., 2019). *Olig2* is a central factor indispensable for oligodendrocyte lineage progression (Liu et al., 2007). Specifically, FTO was reported to regulate the degradation of *Olig2* transcripts *via* removal of m⁶A tags installed on the *Olig2* mRNA. The stability of *Olig2* transcripts in OPCs deficient in FTO was thus seen to increase. In effect, the white matter in FTO mutant mouse brain was characterized by hypomyelination (Wu et al., 2019).

Involvement of N⁶-Methyladenosine Readers in Glia Production in the Brain

The m⁶A reader PRRC2A is known to be essential for oligodendrogenesis. It prominently regulates the specification, proliferation, and differentiation of oligodendroglia and the ability of oligodendrocytes to carry out myelination in the brain (Wu et al., 2019). More specifically, abolishing PRRC2A function in cortical NSCs or precisely in oligodendroglial lineage caused significant loss of OPCs (PDGFR α + cells), Sox10+ cells, and mature oligodendrocytes (CC1+Olig2+), which culminated in hypomyelination in the PRRC2A mutant brain (Wu et al., 2019). Interestingly, deletion of PRRC2A also affects astrogenesis, although slightly. Deficiency of PRRC2A in mouse brain caused a reduction in the proliferative capacity of astrocytes, leading to a reduced number of astrocytes in the mutant mouse brain (Wu et al., 2019). The additional role of PRRC2A in regulating the production of astrocytes in the brain during development may hinge on its interaction with YTHDF2, another m⁶A-binding protein, such that lack of either m⁶A reader augments the expression of the other to influence gliogenesis (Wu et al., 2019).

The competitive relationship between PRRC2A and YTHDF2 makes it complex to explain or reconcile the observation that glial fibrillary acidic protein (GFAP) expression, which can indicate astrocytic cells, was found to be dramatically reduced in neurospheres derived from the E14.5 *Ythdf2*^{-/-} forebrain NSC. Such GFAP+ *Ythdf2*^{-/-} cells also displayed abnormally branched processes (Li M. et al., 2018). Thus, further investigation is required to elucidate the role of YTHDF2 in brain gliogenesis and how the function of PRRC2A features in the regulatory pathway.

N⁶-Methyladenosine Effectors Regulate the Formation of Neural Processes and Synapses

The developing and adult brain is characterized by the outgrowth of dendrites and axons of neurons known to form neural connections called synapses. Interestingly, synapses are enriched with m⁶A, which modulates dendrite formation (dendritogenesis), axonogenesis, and synaptic growth (synaptogenesis) and activity (reviewed in Li et al., 2019; Dermentzaki and Lotti, 2020). m⁶A-based transcriptome profiling of the mouse brain (cortex and cerebellum) showed enrichment of m⁶A modification linked to dendrite and dendritic spine, axon and axon guidance, and synaptogenesis and synaptic transmission (Chang et al., 2017).

Distinctive localization of the YTHDFs, FTO, and METTL14 in dendrites of hippocampal neurons in culture and cortical neurons suggests the involvement of these m⁶A-regulatory factors in the development of neural dendrites. Indeed, *Ythdf1* and *Ythdf3* knockdown in such cultured neurons resulted in abnormal dendritic spine (Merkurjev et al., 2018). Axons are also enriched with FTO, which can be translated locally. As such, FTO ablation in axons resulted in upregulation of m⁶A levels, leading to a reduction in *Gap-43* mRNA translation in axons of cultured dorsal root ganglion neurons (Yu et al., 2018). Yet, GAP-43 is a key factor involved in axon growth in neural tissues (Skene et al., 1986). In effect, the neurons lacking FTO displayed axon elongation repression (Yu et al., 2018). The m⁶A reader YTHDF1 was also reported to influence axon formation by binding and promoting the translation of the axon guidance receptor Robo3.1, which directs spinal commissural axons in crossing the midline, in an m⁶A modification-dependent manner (Zhuang et al., 2019). Together, these observations may have implications for perturbed axonogenesis in the brain lacking optimal m⁶A modification due to ablation of FTO or YTHDF1. At least in the case of the m⁶A-regulatory protein PRRC2A, it was found that axons that form the corpus callosum, a brain midline structure, are hypomyelinated and appeared hypoplastic in the PRRC2A-deleted mouse brain (Wu et al., 2019).

Given the enrichment of m⁶A marks and related proteins in neural processes, it is not surprising that synapses are endowed with m⁶A-modified mRNAs, especially postsynaptic transcripts in the mouse brain (Chang et al., 2017). The high localization of m⁶A-modified mRNAs in synapses reflects the possible impact of the m⁶A epitranscriptome on the structure, maturation, and function of synapses (Chang et al., 2017; Merkurjev et al., 2018; Yu et al., 2018; Zhuang et al., 2019). As a result, selective ablation of YTHDF1 and YTHDF3 in the cultured hippocampal neurons caused excitatory synaptic transmission suppression (Merkurjev et al., 2018; Shi et al., 2018). In addition, synapses formed by neurons lacking YTHDF2 appeared abnormal (Li M. et al., 2018), and synaptic transmission-related transcripts are hypermethylated in dopaminergic neurons with defective synaptic plasticity implication in the FTO-deficient mouse midbrain (Hess et al., 2013). Another indication of

synapse malformation and synaptic plasticity impairment due to m⁶A dysregulation was observed in METTL14-deleted striatal neurons, in which METTL14 abrogation resulted in aberrant neuronal excitability (Koranda et al., 2018). Given that Nito, the *Drosophila* version of RBM15, also regulates synaptic growth through regulation of axonogenesis (Gu et al., 2017), it would be interesting to investigate whether indeed RBM15 is involved in synaptogenesis in the mammalian brain.

Cognition and Behavior Are Modulated by N⁶-Methyladenosine Signaling

The brain's ability to process and store information and form or control behavior patterns has been shown to be greatly regulated by posttranscriptional modification of mRNA involved in brain development (reviewed in Jung and Goldman, 2018; Leighton et al., 2018; Noack and Calegari, 2018). Prominently emerging among these new (epitranscriptomic) levels of brain function regulation is m⁶A modification of mRNA in the brain. Various studies in mouse models have revealed the involvement of the m⁶A machinery-related factors in cognition and behavior (reviewed in Nainar et al., 2016; Chokkalla et al., 2020). The role of m⁶A in the regulation of learning and behavior may be partly explained by the previously discussed role of m⁶A in synaptogenesis and synaptic transmission (Weng et al., 2018).

N⁶-Methyladenosine Writers Involved in Memory and Behavior

In a recent study by Zhang F. et al. (2018), it was found that the enrichment of METTL3 in the mouse hippocampus is supportive for memory consolidation *via* the promotion of neuronal early-response gene translation. Therefore, mice lacking METTL3 in the hippocampus displayed impaired long-term potentiation with attendant reduced ability to consolidate memory. Interestingly, long-term memory consolidation is demonstrably augmented following METTL3 overexpression in the dorsal hippocampus of the wild-type mouse brain (Zhang Z. Y. et al., 2018). The m⁶A writer function of METTL14 is reported to be important for learning and behavior mediated by the striatum. Without affecting the number or morphology of striatal neurons, loss of METTL14 in striatopallidal and striatonigral neurons caused alterations in the transcriptome, eliciting increased neuronal excitability and spike frequency adaptation reduction, which possibly culminated in impairment of striatum-dependent behavior patterns (Koranda et al., 2018).

The N⁶-Methyladenosine Eraser FTO Regulates Learning and Behavior

Accumulation of m⁶A in the brain can affect its learning capacity and behavior. By regulating adult neurogenesis in the mouse hippocampus, FTO has been identified to play a pivotal role in learning (Li L. et al., 2017). Hypermethylation in the mouse brain or hippocampus caused by FTO functional loss was observed to call forth learning disabilities in mice, including increased fear memory consolidation (Widagdo et al., 2016; Walters et al., 2017). Additional evidence indicating the role of FTO in learning

and behavior includes a study in which mice deficient in FTO were reported to exhibit behaviors consistent with depression and anxiety (Sun et al., 2019). Moreover, available data show that memory processing and verbal fluency may be affected in individuals with FTO ablation in the brain (Ho et al., 2010; Benedict et al., 2011).

Readers of N⁶-Methyladenosine Modulate Learning and Memory

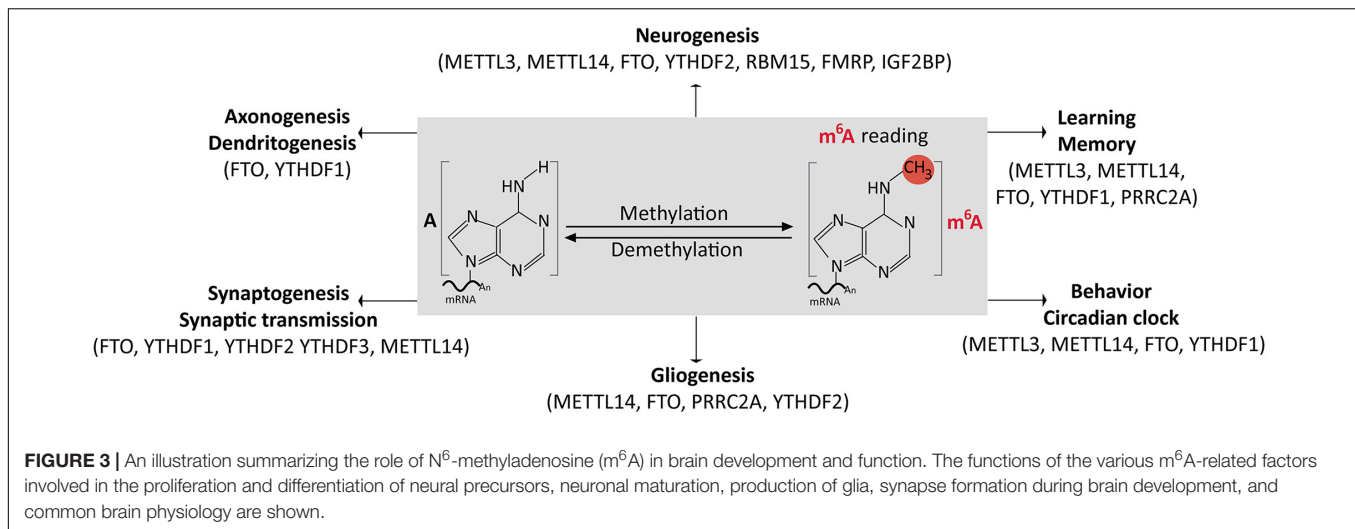
Cognitive deficits have been implicitly linked to lack of function of the m⁶A reader PRRC2A, whose absence caused hypomyelination, leading to the cognitive anomalies in the mouse brain (Wu et al., 2019). Evidence indicating a more direct importance of an m⁶A reader in learning and memory was obtained when YTHDF1 was deleted in the adult mouse brain. It was found that neuronal stimuli can evoke translation of gene transcripts readable by YTHDF1 to facilitate learning and memory (Shi et al., 2018). Hence, silencing of YTHDF1 in the mouse hippocampus resulted in defective long-term potentiation and impaired synaptic transmission in the hippocampus, which did not allow normal learning and memory processing, and the defects were rescuable by YTHDF1 re-expression in the YTHDF1 mutant brain (Shi et al., 2018).

Stress Response Is Regulated by Factors of the N⁶-Methyladenosine Machinery

The brain plays a central role in stress response. In responding to stress, a host of gene expression programs is activated in the brain, leading to the secretion of several neuropeptides (de Kloet et al., 2005). Vulnerability to stressful stimuli and the response mechanism can have implications for neuropsychiatric anomalies under abnormal regulatory conditions. Thus, the transcriptomic stress response system is particularly crucial in maintaining homeostasis following exposure to stress.

Epigenetic mechanisms are known to play central roles in stress response (McEwen et al., 2015), and the epitranscriptome is an emerging gene expression regulation domain for stress modulation (Harvey et al., 2017). A putative role for m⁶A in the regulation of stress response is evidenced by the presence of glucocorticoid response elements upstream the transcription start site of genes that encode for enzymes involved in m⁶A modification (Engel et al., 2018). Additionally, nuclear localization of YTHDF2 precipitated by heat stress results in dynamic methylation of the 5' UTR of newly synthesized mRNAs (Zhou et al., 2015). By limiting FTO, YTHDF2 is able to preserve methylation in the 5' UTR of heat stress-induced mRNAs (Zhou et al., 2015).

In chick, upregulation of FTO in the brain (hypothalamus) may be a mechanism to afford thermoregulation in heat stress conditions (Kisliouk et al., 2020). However, following acute restraint stress, the mouse prefrontal cortex and amygdala displayed m⁶A hypomethylation and hypermethylation, respectively (Engel et al., 2018). Fear-induced stress can cause downregulation of FTO, leading to elevation of m⁶A in the prefrontal cortex and hippocampus of the mouse (Walters et al., 2017). Mice lacking METTL3 or FTO are unable to cope



with stress (Engel et al., 2018). A general effect that may be caused by stress-induced alteration in m⁶A modification is the suppression of mRNAs involved in synaptic plasticity and brain morphogenesis (Engel et al., 2018). Together, the above observations indicate a putative role for m⁶A modulation in the human brain during stressful insults.

NEUROLOGICAL DISORDERS ATTRIBUTABLE TO DEFECTIVE N⁶-METHYLADENOSINE MODIFICATION IN THE BRAIN

Emerging evidence shows that a number of syndromic and non-syndromic neurological disturbances can be linked to m⁶A methylome dysregulation in the brain (Engel and Chen, 2018). This is not surprising, given the previously discussed extensive role of m⁶A in brain neurodevelopment (Figure 3). The m⁶A ubiquity in the brain implies that neural perturbations due to m⁶A dysregulation are likely to be complex and multifactorial in terms of downstream causatives. Neurologic problems so far identified to be caused by genetic variants of m⁶A modification factors can be broadly characterized as neurodevelopmental, neurodegenerative, or neuropsychiatric. Specifically, these include Parkinson's disease (PD), Alzheimer's disease (AD), autism, Smith–Magenis syndrome, schizophrenia, and depression (Table 2). The following subsections discuss the role of m⁶A and associated factors in neurological disorders of the brain.

Fragile X Syndrome

It has been identified that Fragile X syndrome (FXS) is the most common cause of inherited intellectual disorders and usually co-occurs with autism spectrum disorder (ASD). Patients present with features such as poor language development, abnormal behavior, and seizures, which are mainly clinical manifestations of neuronal excitation–inhibition imbalance (Hagerman et al.,

2017; Kaufmann et al., 2017). Silencing of the *FMR1* gene, which leads to lack of FMRP expression, is the cause of FXS (Brown et al., 2001). The role of FMRP in multiple gene expression programs partly accounts for the syndromic nature of FXS (Hagerman et al., 2017). Synaptic abnormalities or loss of neuroplasticity caused by FMRP loss-of-function and perhaps YTHFC2 deficiency is a critical underlying mechanism that contributes to the etiology of FXS and associated ASD (reviewed in Liu et al., 2016; Bagni and Zukin, 2019).

Parkinson's Disease

Parkinson's disease is a complex progressive neurodegenerative disorder mainly associated with death of dopamine-producing neurons in the midbrain (substantia nigra pars compacta) and aggregation of Lewy bodies in various brain regions. The main symptoms of PD include tremor and bradykinesia. Until now, the cause of PD is unknown, as many genetic and environmental risks are involved, making definitive diagnosis and treatment challenging (Kalia and Lang, 2015; Hayes, 2019).

Interestingly, m⁶A methylation deregulation caused by FTO abrogation, in the midbrain or in dopaminergic neurons, has been implicated in PD pathogenesis *via* impairment of neuronal activity and behavior response dependent on dopamine receptor types 2 and 3 (Hess et al., 2013). mRNAs involved in dopaminergic signaling are hypermethylated in the FTO-deficient mouse midbrain and striatum, leading to their decreased translation (Hess et al., 2013). It was found in another study that m⁶A may play a role in loss of dopaminergic neurons, which characterizes PD (Chen X. C. et al., 2019). The study reported that PC12 cells treated with 6-hydroxydopamine (6-OHDA) and the striatum of rat brain with 6-OHDA-induced PD display m⁶A modification downregulation, which is capable of inducing N-methyl-D-aspartate (NMDA) receptor 1 expression, alongside elevated oxidative stress and influx of Ca²⁺, culminating in cell death of dopaminergic neurons. Notably, FTO inhibition, and perhaps inhibition of ALKBH5, can attenuate 6-OHDA-induced PC12 cells apoptosis (Chen X. C. et al., 2019).

TABLE 2 | Brain disorders associated with m⁶A dysregulation.

Neurological disorders	Experimental system	m ⁶ A factor(s) implicated	References
Neurodevelopmental disorders			
Microcephaly	GWAS; KO mice	<i>Fto</i> deletion; <i>METTL5</i> frameshift	Richard et al. (2019)
Fragile X Syndrome	GWAS	SNP in <i>FMRP</i>	Verkerk et al. (1991); Dichtenberg et al. (2008)
Cerebellar ataxia	KO mice; KO <i>Drosophila</i>	Deletion of <i>Ythdc1</i> , <i>Mettl3</i> , <i>Alkbh5</i>	Fernandez-Funez et al. (2000); Ma et al. (2018); Wang C. X. et al. (2018)
Smith–Magenis syndrome	Genetic analysis in mouse	<i>Alkbh5</i> deletion	Ricard et al. (2010)
Intellectual disability	GWAS	<i>METTL5</i> frameshift	Richard et al. (2019)
Autism spectrum disorder	GWAS	Mutations in <i>FMR1</i>	Reddy (2005); Edupuganti et al. (2017); Kaufmann et al. (2017)
Neurodegenerative disorders			
Parkinson's disease	6-OHDA treatment of PC12 cells and rats; KO mice	<i>Fto</i> deletion or inhibition	Hess et al. (2013); Chen X. C. et al. (2019)
Alzheimer's disease	GWAS	SNP in <i>FTO</i>	Ho et al. (2010); Keller et al. (2011); Reitz et al. (2012)
Amyotrophic lateral sclerosis	GWAS	SNP in <i>FTO</i> ; SNP in <i>HNRNP (A2B1 and A1)</i> ; SNP in <i>RBM15</i>	Kim et al. (2013); Cooper-Knock et al. (2017); Mitropoulos et al. (2017)
Cerebellar ataxia	KO mice; KO <i>Drosophila</i>	Deletion of <i>Ythdc1</i> , <i>Mettl3</i> , <i>Alkbh5</i>	Fernandez-Funez et al. (2000); Ma et al. (2018); Wang C. X. et al. (2018)
Multiple sclerosis	GWAS	SNP in <i>METTL1</i>	Mo et al. (2019)
Neuropsychiatric disorders			
Major depressive disorder	GWAS	SNP in <i>ALKBH5</i> ; SNP in <i>FTO</i>	Samaan et al. (2013); Milaneschi et al. (2014); Du et al. (2015)
Schizophrenia	GWAS	SNP in <i>ZC3H13</i>	Oldmeadow et al. (2014)
Attention-deficit/hyperactivity disorder	GWAS	SNP in <i>FTO</i>	Velders et al. (2012); Choudhry et al. (2013)

GWAS, genome-wide association studies; 6-OHDA, 6-hydroxydopamine; SNP, single-nucleotide polymorphism; KO, knockout; m⁶A, N⁶-methyladenosine.

Alzheimer's Disease

The commonest cause of dementia worldwide is AD. It is mainly characterized by progressive (age-dependent) neurodegeneration in brain regions (especially in the temporal and frontal lobes), with key clinical features, including memory loss, behavior abnormalities, and eventual cognitive decline (reviewed in Weller and Budson, 2018; Soria Lopez et al., 2019). Errors in RNA metabolism can have implications for AD. As discussed further, studies in human populations and in mouse models have shown that specific dysregulations in m⁶A mRNA methylation contribute to AD pathogenesis.

Typically, m⁶A levels in various brain regions increase with aging, and this disposition was shown to likely have relevance for AD development (Shafik et al., 2021). Interestingly, while *METTL3* is downregulated in AD brain (hippocampus), it was observed to have accumulated in the postmortem AD brain at levels comparable to the insoluble Tau protein therein (Huang H. et al., 2020). Immunohistochemistry of the entorhinal cortex of patients with AD showed selective deficiency in the expression of another m⁶A factor hnRNP-A/B, which probably underscores the alteration in alternative splicing in the AD brain (Berson et al., 2012). Moreover, *FTO* mis-expression is implicated in the development of AD. Carriers of the *FTO* variant rs9939609 were reported to display systematic deficits in brain volume consistent with brain atrophy in the elderly (Ho et al., 2010). Indeed, a population-based study found an association between

the *FTO* variant rs9939609 and increased risk of AD (Keller et al., 2011). Reitz et al. (2012) reported an increased risk caused by some polymorphisms (rs11075997, rs11075996, rs17219084) in the *FTO* gene in AD cases among some investigated Caribbean Hispanics and Caucasians (Reitz et al., 2012). Reduced verbal fluency in obese and overweight elderly men, with unaffected general cognitive function, was attributed to bearing of the *FTO* A allele. Thus, the (dys)functional effect of *FTO* A allele mainly manifests in the frontal lobe of the brain to constitute AD (Benedict et al., 2011). These observations indicate perturbation of m⁶A signaling as a notable underlying factor in the pathophysiology of AD in humans.

In vitro and *in vivo* experimentations using mouse models have yielded results that further support the involvement of m⁶A mRNA methylation in AD. In one study, it was observed that knockdown of hnRNP A/B impaired alternative splicing in cultured neurons, which resulted in loss of dendrites, and caused memory impairment in mice that can be ascribed to aberrance in the cortical connectome (Berson et al., 2012). The level of hnRNP A/B increases with cholinergic excitation, whereas loss of cholinergic signaling was found to induce AD-like reduction in hnRNP levels in the cortex (Berson et al., 2012).

The AD brain of the APP/PS1 transgenic mouse has elevated levels of m⁶A in the hippocampus and cortex, which may be due to the increased expression of *METTL3* and concurrent downregulation of *FTO* expression (Han et al., 2020). However,

the expression of FTO was identified to be increased in the brain of the triple transgenic AD mouse (Li H. et al., 2018). This gives an impression of the complex nature of the mechanism through which FTO or other m⁶A-associated factors may drive the development of AD. In the case of FTO, a proposed mechanism is that it may promote the phosphorylation of Tau protein by encouraging a methylation scheme leading to stabilization of tuberous sclerosis complex 1 (TSC1) mRNA, which activates the kinase activity of the mammalian target of rapamycin (mTOR) (Li H. et al., 2018).

Interestingly, cognition in a mouse model of AD was observed to improve when FTO was conditionally deleted in neurons in the mouse brain with AD (Li H. et al., 2018). This makes FTO a prospective therapeutic candidate worth further investigation for its potential in slowing down the progression of AD or in remedying related symptoms.

Amyotrophic Lateral Sclerosis

Amyotrophic lateral sclerosis (ALS) is a debilitating neurodegenerative disorder hallmarked by loss of motor neurons leading to skeletal muscle dysfunction and other clinical features, including psychological disorders and respiratory distress (Rowland and Shneider, 2001). It is believed to be idiopathic, with a greater percentage (~90%) of cases being sporadic, while 5%–10% of cases are familial or inheritable (Kiernan et al., 2011). Studies have revealed the prominent role played by pathogenic mutation of factors associated with the RNA methylation machinery (Kim et al., 2013; Cooper-Knock et al., 2017; Mitropoulos et al., 2017).

By means of whole-genome sequencing, it became evident that m⁶A may be involved in the pathogenesis of ALS through FTO function alteration (Mitropoulos et al., 2017). Variants of FTO gene were thus associated with sporadic cases of ALS, which appears to be a founder effect among Greeks (Mitropoulos et al., 2017). In another key study, mutation in the prion-like domain of the m⁶A reader HNRNP (A2B1 and A1) was implicated in the pathogenesis of a familial ALS case (Kim et al., 2013). The work of Cooper-Knock et al. (2017) supports the involvement of RNA-binding protein mutations in ALS. Deleterious variants of RBM15 gene or its paralog RBM15B were found to contribute to the pathogenesis of ALS (Cooper-Knock et al., 2017).

Major Depressive Disorder

Major depressive disorder (MDD) is a common neuropsychiatric condition that is considered a biobehavioral syndrome with clinical characteristics including depressed mood, cognitive dysfunction, neurovegetative disturbance, and diminished interests. Females are known to be more affected by MDD than males. Multiple factors are known to cause MDD. Notable underlying causatives include genetic and environmental factors leading to alteration in the volume of the hippocampus and aberrant brain circuitry (Fava and Kendler, 2000; Flint and Kendler, 2014; Otte et al., 2016).

The m⁶A RNA methylome plays a role in the development of MDD (Engel et al., 2018). Genetic variants of FTO have been implicated in MDD, although heterogeneity in the associated

phenotype is noteworthy (Milaneschi et al., 2014). In particular, it was found in a genome-wide association study that the FTO rs9939609 A variant is associated with a reduced risk of MDD (Samaan et al., 2013). A single-nucleotide polymorphism (rs12936694) in ALKBH5 was also found to likely be the cause of MDD among the Chinese Han population in an association study (Du et al., 2015). Interestingly, by blocking the translocation of ALKBH5 into the nucleus, it was possible to attenuate depression-like behavior in the mouse due to attendant hypermethylation and subsequent degradation of fatty acid amide hydrolase mRNA in astrocytes (Huang R. R. et al., 2020).

THERAPEUTIC PROSPECTS OF CRISPR-Cas13-MEDIATED RNA METHYLATION REGULATION IN N⁶-METHYLADENOSINE-RELATED NEUROLOGICAL DISEASE TREATMENT

While it seems intuitive that a simple strategy of traditional knockdown or overexpression of dysfunctional m⁶A factors in the epitranscriptome can correct pathologic alterations in the RNA methylation program, heterogeneity of the m⁶A methylome and, in some cases, the possible functional duplication or duality of the m⁶A writers, erasers, and readers, possess a challenge for the applicability of such solutions. To circumvent the aforementioned constraints, a system or tool capable of targeting defective m⁶A sites with high specificity should be considered. Such targeted approach to reversing disease-causing m⁶A modification can have therapeutic application if perfected.

The discovery of the Cas13 family of proteins, which are able to target endogenous RNA, has opened up avenues to deliver specific effectors at single sites on gene transcripts (Abudayyeh et al., 2016). By associating clustered regularly interspaced short palindromic repeats (CRISPR) with a catalytically inactive form of Cas13 protein (dCas13), but having preserved RNA binding ability, (m)RNA can be targeted at specific nucleic acid loci with such designed programmable CRISPR-dCas13 system (Figure 4; Wang et al., 2019; Burmistrz et al., 2020). Here, we discuss various salient *in vitro* applications of the CRISPR-dCas13 system to achieve m⁶A editing (Table 3).

Restoring Abnormal Loss of N⁶-Methyladenosine

Gene transcripts that have lost m⁶A because of malfunction of the methyltransferases (METTL3 and/or METTL14) in the m⁶A methylation complex can be repaired using the CRISPR-dCas13 system. This is achievable by fusing dCas13, localized in the nucleus or cytoplasm, with a methyltransferase domain-truncated METTL3 or a modified METTL3:METTL14 complex, respectively. The resultant CRISPR-dCas13 constructs were able to install m⁶A marks, in a site-specific manner, on hypomethylated mRNAs or mRNAs with amendable m⁶A levels, including *Sox2*, *Foxm1*, and *Znf638* in human cells (Wilson et al., 2020). Light-mediated m⁶A editing has also been put forward as another appealing CRISPR-dCas13 system for engineering the

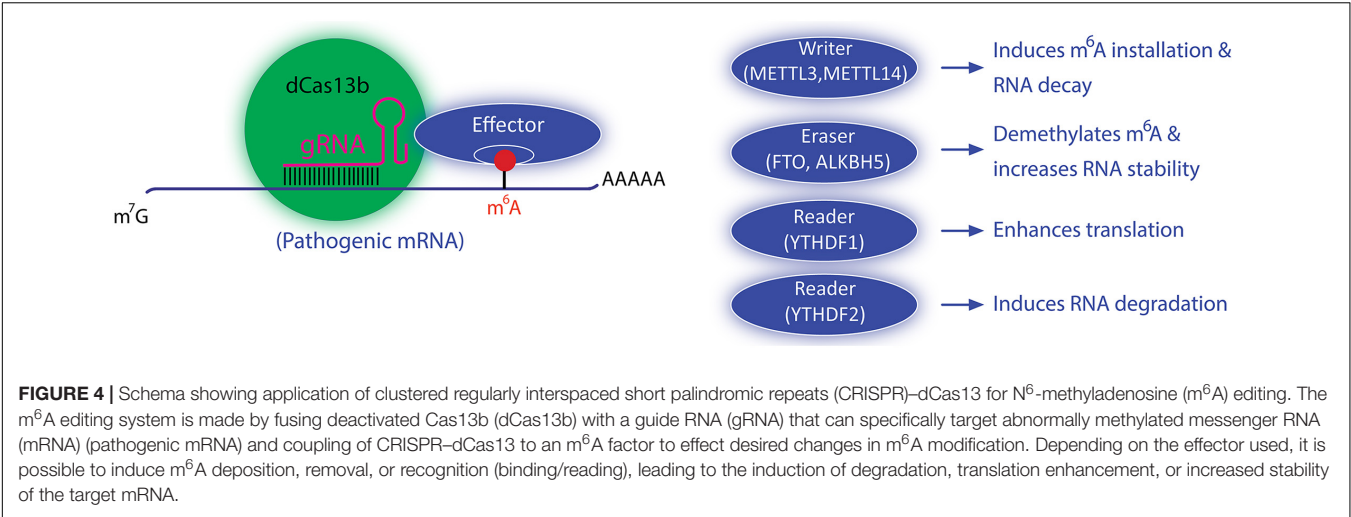


TABLE 3 | Applications based on CRISPR–dCas13b system for targeted manipulation of m⁶A modification and fate of m⁶A-tagged mRNA.

System	Effect	Targeted mRNA	Reference
Targeting m⁶A installation			
dCas13b-METTL3 dCas13b-METTL14	Induces effective m ⁶ A incorporation in endogenous transcripts with increased specificity; regulates m ⁶ A-dependent mechanism for controlling transcript abundance; induces alternative splicing	<i>Actb, Gapdh, Foxm1, Sox2, Brd8, Znf638,</i>	Wilson et al. (2020)
dCas13b-CIBN CRY2PHR-METTL3-METTL14	Effect photoactivatable RNA m ⁶ A level upregulation	<i>TPT1, ACTB, TUG1</i>	Zhao et al. (2020)
Targeting m⁶A removal			
dCas13b-FTO	Demethylates m ⁶ A of targeted mRNAs to enhance their stability	<i>MALAT1, ASB2, HBV, HXB2</i>	Mo et al. (2020)
dCas13b-CIBN CRY2PHR-FTO	Effect photoactivatable RNA m ⁶ A level reduction	<i>MALAT1, Alp, Bglap, RunX2, and Sp7, Pth1r</i>	Zhao et al. (2020)
dCas13b-ALKBH5	Demethylates m ⁶ A of targeted mRNAs to enhance their stability	<i>CYB5A, CTNNB1, EGFR, MYC</i>	Li et al. (2020)
Modulating m⁶A reading			
dCas13b-YTHDF2	Induces RNA degradation	<i>KRAS, PPIB</i>	Rauch et al. (2018)
dCas13b-YTHDF1	Enhances translation with minimal mRNA destabilization effect	Firefly luciferase	Rauch et al. (2018)

CRISPR, clustered regularly interspaced short palindromic repeats; m⁶A, N⁶-methyladenosine.

m⁶A methylome and worth close examination for applicability of the principle in therapeutics. The photoactivatable m⁶A editing CRISPR–dCas13b tool, which has a coupled component made of the methyltransferase domains of METTL3 and METTL14, was effectively employed in adding m⁶A to gene transcripts (*TPT1, ACTB, TUG1*) in human cells (Zhao et al., 2020).

Correcting N⁶-Methyladenosine Removal Incompetence

In the context of m⁶A erasure (demethylation), the m⁶A demethylase FTO and ALKBH5 can be incorporated into the CRISPR–dCas13 system to effect targeted removal of m⁶A on hypermethylated mRNAs or induce hypomethylation as corrective measures. By utilizing a CRISPR–dCas13b–FTO construct, Mo et al. (2020) were able to make mRNAs more stable *via* site-directed demethylation (Table 3) (Mo et al.,

2020). Similarly, by applying a photoactivatable FTO-coupled CRISPR–dCas13b strategy, m⁶A marks were effectively and in a targeted manner removed on endogenous gene transcripts (Zhao et al., 2020). Methylated mRNAs that are preferentially demethylated by ALKBH5 can also be targeted by a CRISPR–dCas13b–ALKBH5 construct to remodel their m⁶A milieu, as applied in reducing the m⁶A levels associated with transcripts like *CYB5A, CTNNB1, EGFR*, and *MYC*, leading to their increased stability and translation (Li et al., 2020).

Rescuing Defective N⁶-Methyladenosine Reading

It is also possible to specifically target m⁶A readers to mRNAs of interest using the CRISPR–dCas13b system. It also implies that m⁶A binding protein dysfunctionality due to mutation can be rectified with a CRISPR–dCas13b construct fused to

an engineered functional version of the relevant defective m⁶A reader. For example, YTHDF1 and/or YTHDF2, the two well-characterized m⁶A readers, can be fused to CRISPR–dCas13b and guided to specific mRNAs for m⁶A reading and subsequent fate alteration. Both CRISPR–dCas13b–YTHDF1 and CRISPR–dCas13b–YTHDF2 constructs were able to effect the native functions of YTHDF1 and YTHDF2, leading to translation enhancement and mRNA degradation in cells, respectively (Rauch et al., 2018).

Targeted N⁶-Methyladenosine Editing in the Diseased Brain as a Promising Treatment Strategy

Based on the intriguing outcomes and specificity of m⁶A editing application *in vitro* (Table 3 and Figure 4), we hereby propose the CRISPR–dCas13 system as a highly efficient tool for precise targeting and repair of aberrant m⁶A-modified mRNAs implicated in the pathophysiology of pertinent neurological disorders. Such a tool can be potentially useful in treating neurological disorders known to have pathologic m⁶A mRNA methylation, demethylation, or reading as the central underlying pathogenesis mechanism (Table 2). Employing high-resolution single-nucleotide binding techniques will be critical for identifying specific nucleotides bearing the abnormal m⁶A modification in the diseased brain. This will improve targeting, leading to the desired effect. An example of a strategy for improving the identification and targeting of nucleotides harboring disease-causing m⁶A marks is by adopting the enhanced crosslinking and immunoprecipitation (eCLIP) technique for robust factor-specific profiling of the m⁶A methylome in the pathologic brain (Van Nostrand et al., 2016). The *in vivo* experimental approach for investigating the potency of the CRISPR–dCas13 system for resolving neurological disorders caused by m⁶A dysregulation would include modeling the disorder in experimental animals and treating them with the rescuing CRISPR–dCas13 construct(s) that would have a target effect in the brain and on the implicated pathogenic mRNA (Figure 4). While the idea of investigating the application of the CRISPR–dCas13 system for rectifying aberrant m⁶A mRNA methylation (Figure 4) implied in neurological disorders sounds interesting, the approach may be fraught with challenges, especially in preventing off-target effects and in rescuing phenotypes of complex syndromic neurological disorders (e.g., ASD, AD, schizophrenia, MDD). Further investigations that can reveal convergent downstream effectors underlying the pathophysiology of polygenic neurological disorders caused by defective m⁶A signaling can help streamline an m⁶A editing-mediated therapeutic strategy.

CONCLUSION

Methylation of mRNA has emerged as a posttranscriptional regulation of gene expression that modulates protein synthesis in cells. Studies have shown that the m⁶A mRNA methylation machinery, composed of writers, erasers, and

readers (Figure 1A), critically and extensively regulates RNA metabolism (trafficking, stability, processing, and translation efficiency) in cells to impact major biological processes. The brain is a hub of m⁶A modification, and the enrichment of m⁶A in the brain is reflective of the essential role it plays in optimal brain morphogenesis and functionality. Hence, the m⁶A interactome is known to regulate several neurodevelopmental processes in the brain, including neurogenesis, gliogenesis, synapse formation, and neuronal activity. Many of the m⁶A factors appear to have multiple functional effects during cortical development and in orchestrating several aspects of brain physiology (Table 1). This may make it challenging to effectively disentangle the rather multifactorial downstream causatives or complex phenotypic effects elicited by the dysregulation of the m⁶A machinery in the brain. As a typical example, whereas the *FTO rs9939609* A variant is a risk factor for brain atrophy in old age (Ho et al., 2010) and AD development (Keller et al., 2011), it appears to be neuroprotective against MDD (Samaan et al., 2013). It also implies that m⁶A signaling is worth considering as a pivotal pathway that can cause novel syndromic neurological disturbances. Indeed, some inherited or acquired defects in the m⁶A RNA methylome are known causes of syndromes such as ASD, Smith–Magenis syndrome, and FXS. It also goes to reason that genetic variants of factors that make the m⁶A machinery pose a risk for certain (novel) non-syndromic neurological anomalies of the brain.

The phenomenal neurodevelopmental role played by m⁶A mRNA methylation and implication for neurological perturbations provoke considerable attention to the emerging involvement of the m⁶A methylome in normal brain structure and function maintenance. Going forward, more robust and advanced probing techniques are required to finely dissect the mechanistic basis of m⁶A-mediated neurodevelopment and its involvement in the pathophysiology of pertinent neurological disorders of the brain. Such sophisticated investigations may uncover therapeutic cues that can potentially fend off the neurological disorders caused by defective RNA methylation in the brain or alleviate the associated symptoms. For now, the application of the CRISPR–dCas13 system to edit m⁶A to restore normality of mRNA state and fate in, say, brain disease conditions is one of the promising approaches for treating abnormal m⁶A signaling-related neurological disorders.

AUTHOR CONTRIBUTIONS

GS, YX, HN, and TT all contributed to writing and editing the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by TU432/1, TU432/3, TU432/6 DFG grants and Schram-Stiftung to TT. We also acknowledge support by the Open Access Publication Funds of the Ruhr-University Bochum.

REFERENCES

- Abdelmohsen, K., and Gorospe, M. (2010). Posttranscriptional regulation of cancer traits by HuR. *Wiley Interdiscipl. Rev. RNA* 1, 214–229. doi: 10.1002/wrna.4
- Abudayyeh, O. O., Gootenberg, J. S., Konermann, S., Joung, J., Slaymaker, I. M., Cox, D. B., et al. (2016). C2c2 is a single-component programmable RNA-guided RNA-targeting CRISPR effector. *Science* 353:aaf5573. doi: 10.1126/science.aaf5573
- Agarwala, S. D., Blitzblau, H. G., Hochwagen, A., and Fink, G. R. (2012). RNA methylation by the MIS complex regulates a cell fate decision in yeast. *PLoS Genet.* 8:e1002732. doi: 10.1371/journal.pgen.1002732
- Alarcón, C. R., Goodarzi, H., Lee, H., Liu, X., Tavazoie, S., and Tavazoie, S. F. (2015). HNRNPA2B1 is a mediator of m(6)A-dependent nuclear RNA processing events. *Cell* 162, 1299–1308. doi: 10.1016/j.cell.2015.08.011
- Anders, M., Chelysheva, I., Goebel, I., Trenkner, T., Zhou, J., Mao, Y., et al. (2018). Dynamic m 6 A methylation facilitates mRNA triaging to stress granules. *Life Sci. Alliance* 1:e201800113. doi: 10.26508/lsa.201800113
- Ascano, M. Jr., Mukherjee, N., Bandaru, P., Miller, J. B., Nusbaum, J. D., Corcoran, D. L., et al. (2012). FMRP targets distinct mRNA sequence elements to regulate protein expression. *Nature* 492, 382–386. doi: 10.1038/nature11737
- Bagni, C., and Zukin, R. S. (2019). A synaptic perspective of fragile X syndrome and autism spectrum disorders. *Neuron* 101, 1070–1088. doi: 10.1016/j.neuron.2019.02.041
- Barbieri, I., Tzelepis, K., Pandolfini, L., Shi, J., Millán-Zambrano, G., Robson, S. C., et al. (2017). Promoter-bound METTL3 maintains myeloid leukaemia by m(6)A-dependent translation control. *Nature* 552, 126–131. doi: 10.1038/nature24678
- Batista, P. J., Molinier, B., Wang, J., Qu, K., Zhang, J., Li, L., et al. (2014). m(6)A RNA modification controls cell fate transition in mammalian embryonic stem cells. *Cell Stem Cell* 15, 707–719. doi: 10.1016/j.stem.2014.09.019
- Bell, J. L., Wächter, K., Mühleck, B., Pazaitis, N., Köhn, M., Lederer, M., et al. (2013). Insulin-like growth factor 2 mRNA-binding proteins (IGF2BPs): post-transcriptional drivers of cancer progression? *Cell. Mol. Life Sci.* 70, 2657–2675. doi: 10.1007/s00018-012-1186-z
- Benedict, C., Jacobsson, J. A., Rönnekaa, E., Sällman-Almén, M., Brooks, S., Schultes, B., et al. (2011). The fat mass and obesity gene is linked to reduced verbal fluency in overweight and obese elderly men. *Neurobiol. Aging* 32:1159.e1–5. doi: 10.1016/j.neurobiolaging.2011.02.006
- Berson, A., Barbash, S., Shaltiel, G., Goll, Y., Hanin, G., Greenberg, D. S., et al. (2012). Cholinergic-associated loss of hnRNP-A/B in Alzheimer's disease impairs cortical splicing and cognitive function in mice. *EMBO Mol. Med.* 4, 730–742. doi: 10.1002/emmm.201100995
- Boccaletto, P., Machnicka, M. A., Purta, E., Piatkowski, P., Baginski, B., Wirecki, T. K., et al. (2018). MODOMICS: a database of RNA modification pathways. 2017 update. *Nucleic Acids Res.* 46, D303–D307. doi: 10.1093/nar/gkx1030
- Bokar, J. A., Rath-Shambaugh, M. E., Ludwiczak, R., Narayan, P., and Rottman, F. (1994). Characterization and partial purification of mRNA N6-adenosine methyltransferase from HeLa cell nuclei. Internal mRNA methylation requires a multisubunit complex. *J. Biol. Chem.* 269, 17697–17704. doi: 10.1016/s0021-9258(17)32497-3
- Bokar, J. A., Shambaugh, M. E., Polayes, D., Matera, A. G., and Rottman, F. M. (1997). Purification and cDNA cloning of the AdoMet-binding subunit of the human mRNA (N6-adenosine)-methyltransferase. *RNA* 3, 1233–1247.
- Boles, N. C., and Temple, S. (2017). Epimetronomics: m6A marks the tempo of corticogenesis. *Neuron* 96, 718–720. doi: 10.1016/j.neuron.2017.11.002
- Brown, V., Jin, P., Ceman, S., Darnell, J. C., O'Donnell, W. T., Tenenbaum, S. A., et al. (2001). Microarray identification of FMRP-associated brain mRNAs and altered mRNA translational profiles in fragile X syndrome. *Cell* 107, 477–487. doi: 10.1016/s0092-8674(01)00568-2
- Bujnicki, J. M., Feder, M., Radlinska, M., and Blumenthal, R. M. (2002). Structure prediction and phylogenetic analysis of a functionally diverse family of proteins homologous to the MT-A70 subunit of the human mRNA:m6A methyltransferase. *J. Mol. Evol.* 55, 431–444. doi: 10.1007/s00239-002-2339-8
- Burmistrz, M., Krakowski, K., and Krawczyk-Balska, A. (2020). RNA-targeting CRISPR-Cas systems and their applications. *Int. J. Mol. Sci.* 21:1122. doi: 10.3390/ijms21031122
- Cao, Y., Zhuang, Y., Chen, J., Xu, W., Shou, Y., Huang, X., et al. (2020). Dynamic effects of Fto in regulating the proliferation and differentiation of adult neural stem cells of mice. *Hum. Mol. Genet.* 29, 727–735. doi: 10.1093/hmg/ddz274
- Chang, M., Lv, H., Zhang, W., Ma, C., He, X., Zhao, S., et al. (2017). Region-specific RNA m(6)A methylation represents a new layer of control in the gene regulatory network in the mouse brain. *Open Biol.* 7:170166. doi: 10.1098/rsob.170166
- Chen, J., Zhang, Y. C., Huang, C., Shen, H., Sun, B., Cheng, X., et al. (2019). m(6)A regulates neurogenesis and neuronal development by modulating histone methyltransferase Ezh2. *Genomics Proteom. Bioinform.* 17, 154–168. doi: 10.1016/j.gpb.2018.12.007
- Chen, X. C., Yu, C. Y., Guo, M. J., Zheng, X. T., Ali, S., Huang, H., et al. (2019). Down-regulation of m6A mRNA methylation is involved in dopaminergic neuronal death. *ACS Chem. Neurosci.* 10, 2355–2363. doi: 10.1021/acscchemneuro.8b00657
- Choe, J., Lin, S., Zhang, W., Liu, Q., Wang, L., Ramirez-Moya, J., et al. (2018). mRNA circularization by METTL3-eIF3h enhances translation and promotes oncogenesis. *Nature* 561, 556–560. doi: 10.1038/s41586-018-0538-8
- Chokkalla, A. K., Mehta, S. L., and Vemuganti, R. (2020). Epitranscriptomic regulation by m(6)A RNA methylation in brain development and diseases. *J. Cereb. Blood Flow Metab.* 40, 2331–2349. doi: 10.1177/0271678X20960033
- Choudhry, Z., Sengupta, S. M., Grizenko, N., Thakur, G. A., Fortier, M. E., Schmitz, N., et al. (2013). Association between obesity-related gene FTO and ADHD. *Obesity* 21, E738–E744. doi: 10.1002/oby.20444
- Cooper-Knock, J., Robins, H., Niedermoser, I., Wyles, M., Heath, P. R., Higginbottom, A., et al. (2017). Targeted genetic screen in amyotrophic lateral sclerosis reveals novel genetic variants with synergistic effect on clinical phenotype. *Front. Mol. Neurosci.* 10:370. doi: 10.3389/fnmol.2017.00370
- Darnell, R. B., Ke, S., and Darnell, J. E. Jr. (2018). Pre-mRNA processing includes N(6) methylation of adenosine residues that are retained in mRNA exons and the fallacy of "RNA epigenetics". *RNA* 24, 262–267. doi: 10.1261/rna.065219.117
- de Kloet, E. R., Joëls, M., and Holsboer, F. (2005). Stress and the brain: from adaptation to disease. *Nat. Rev. Neurosci.* 6, 463–475. doi: 10.1038/nrn1683
- Dermentzaki, G., and Lotti, F. (2020). New insights on the role of N (6)-methyladenosine RNA methylation in the physiology and pathology of the nervous system. *Front. Mol. Biosci.* 7:555372. doi: 10.3389/fmolb.2020.555372
- Desrosiers, R., Friderici, K., and Rottman, F. (1974). Identification of methylated nucleosides in messenger RNA from novikoff hepatoma cells. *Proc. Natl. Acad. Sci. U.S.A.* 71, 3971–3975. doi: 10.1073/pnas.71.10.3971
- Dicthenberg, J. B., Swanger, S. A., Antar, L. N., Singer, R. H., and Bassell, G. J. (2008). A direct role for FMRP in activity-dependent dendritic mRNA transport links filopodial-spine morphogenesis to fragile X syndrome. *Dev. Cell* 14, 926–939. doi: 10.1016/j.devcel.2008.04.003
- Dominissini, D., Moshitch-Moshkovitz, S., Schwartz, S., Salmon-Divon, M., Ungar, L., Osenberg, S., et al. (2012). Topology of the human and mouse m6A RNA methylomes revealed by m6A-seq. *Nature* 485, 201–206. doi: 10.1038/nature11112
- Donega, V., Marcy, G., Lo Giudice, Q., Zweifel, S., Angonin, D., Fiorelli, R., et al. (2018). Transcriptional dysregulation in postnatal glutamatergic progenitors contributes to closure of the cortical neurogenic period. *Cell Rep.* 22, 2567–2574. doi: 10.1016/j.celrep.2018.02.030
- Du, H., Zhao, Y., He, J., Zhang, Y., Xi, H., Liu, M., et al. (2016). YTHDF2 destabilizes m6A-containing RNA through direct recruitment of the CCR4–NOT deadenylase complex. *Nat. Commun.* 7:12626. doi: 10.1038/ncomms12626
- Du, T. F., Li, G. X., Yang, J. L., and Ma, K. L. (2020). RNA demethylase Alkbh5 is widely expressed in neurons and decreased during brain development. *Brain Res. Bull.* 163, 150–159. doi: 10.1016/j.brainresbull.2020.07.018
- Du, T. F., Rao, S. Q., Wu, L., Ye, N., Liu, Z. Y., Hu, H. L., et al. (2015). An association study of the m6A genes with major depressive disorder in Chinese Han population. *J. Affect. Disord.* 183, 279–286. doi: 10.1016/j.jad.2015.05.025
- Edens, B. M., Vissers, C., Su, J., Arumugam, S., Xu, Z., Shi, H., et al. (2019). FMRP modulates neural differentiation through m(6)A-dependent mRNA nuclear export. *Cell Rep.* 28, 845.e5–854.e5. doi: 10.1016/j.celrep.2019.06.072
- Edupuganti, R. R., Geiger, S., Lindeboom, R. G. H., Shi, H., Hsu, P. J., Lu, Z., et al. (2017). N(6)-methyladenosine (m(6)A) recruits and repels proteins to regulate mRNA homeostasis. *Nat. Struct. Mol. Biol.* 24, 870–878. doi: 10.1038/nsmb.3462

- Engel, M., and Chen, A. (2018). The emerging role of mRNA methylation in normal and pathological behavior. *Genes Brain Behav.* 17:e12428. doi: 10.1111/gbb.12428
- Engel, M., Eggert, C., Kaplick, P. M., Eder, M., Roh, S., Tietze, L., et al. (2018). The role of m(6)A/m-RNA methylation in stress response regulation. *Neuron* 99, 389.e9–403.e9. doi: 10.1016/j.neuron.2018.07.009
- Fan, X. C., and Steitz, J. A. (1998). Overexpression of HuR, a nuclear-cytoplasmic shuttling protein, increases the in vivo stability of ARE-containing mRNAs. *EMBO J.* 17, 3448–3460. doi: 10.1093/emboj/17.12.3448
- Fava, M., and Kendler, K. S. (2000). Major depressive disorder. *Neuron* 28, 335–341. doi: 10.1016/s0896-6273(00)00112-4
- Feng, Z., Li, Q., Meng, R., Yi, B., and Xu, Q. (2018). METTL3 regulates alternative splicing of MyD88 upon the lipopolysaccharide-induced inflammatory response in human dental pulp cells. *J. Cell Mol. Med.* 22, 2558–2568. doi: 10.1111/jcmm.13491
- Fernandez-Funez, P., Nino-Rosales, M. L., de Gouyon, B., She, W.-C., Luchak, J. M., Martinez, P., et al. (2000). Identification of genes that modify ataxin-1-induced neurodegeneration. *Nature* 408, 101–106. doi: 10.1038/35040584
- Flint, J., and Kendler, K. S. (2014). The genetics of major depression. *Neuron* 81, 484–503. doi: 10.1016/j.neuron.2014.01.027
- Fu, Y., Dominissini, D., Rechavi, G., and He, C. (2014). Gene expression regulation mediated through reversible m6A RNA methylation. *Nat. Rev. Genet.* 15, 293–306. doi: 10.1038/nrg3724
- Fustin, J. M., Doi, M., Yamaguchi, Y., Hida, H., Nishimura, S., Yoshida, M., et al. (2013). RNA-methylation-dependent RNA processing controls the speed of the circadian clock. *Cell* 155, 793–806. doi: 10.1016/j.cell.2013.10.026
- Gao, X., Wan, J., Liu, B., Ma, M., Shen, B., and Qian, S.-B. (2015). Quantitative profiling of initiating ribosomes in vivo. *Nat. Methods* 12, 147–153. doi: 10.1038/nmeth.3208
- Geula, S., Moshitch-Moshkovitz, S., Dominissini, D., Mansour, A. A., Kol, N., Salmon-Divon, M., et al. (2015). m6A mRNA methylation facilitates resolution of naïve pluripotency toward differentiation. *Science* 347, 1002–1006. doi: 10.1126/science.1261417
- Gu, T. T., Zhao, T., Kohli, U., and Hewes, R. S. (2017). The large and small SPEN family proteins stimulate axon outgrowth during neurosecretory cell remodeling in *Drosophila*. *Dev. Biol.* 431, 226–238. doi: 10.1016/j.ydbio.2017.09.013
- Hagerman, R. J., Berry-Kravis, E., Hazlett, H. C., Bailey, D. B. Jr., Moine, H., Kooy, R. F., et al. (2017). Fragile X syndrome. *Nat. Rev. Dis. Primers* 3:17065. doi: 10.1038/nrdp.2017.65
- Han, M., Liu, Z., Xu, Y., Liu, X., Wang, D., Li, F., et al. (2020). Abnormality of m6A mRNA methylation is involved in alzheimer's disease. *Front. Neurosci.* 14:98. doi: 10.3389/fnins.2020.00098
- Harfmann, A. M., Nayler, O., Schwaiger, F. W., Obermeier, A., and Stamm, S. (1999). The interaction and colocalization of Sam68 with the splicing-associated factor YT521-B in nuclear dots is regulated by the Src family kinase p59(fyn). *Mol. Biol. Cell* 10, 3909–3926. doi: 10.1091/mbc.10.11.3909
- Harvey, R., Dezi, V., Pizzinga, M., and Willis, A. E. (2017). Post-transcriptional control of gene expression following stress: the role of RNA-binding proteins. *Biochem. Soc. Trans.* 45, 1007–1014. doi: 10.1042/bst20160364
- Hayes, M. T. (2019). Parkinson's disease and Parkinsonism. *Am. J. Med.* 132, 802–807. doi: 10.1016/j.amjmed.2019.03.001
- Hazra, D., Chapat, C., and Graille, M. (2019). m6A mRNA destiny: chained to the rYTHm by the YTH-containing proteins. *Genes* 10:49. doi: 10.3390/genes10010049
- Heck, A. M., Russo, J., Wilusz, J., Nishimura, E. O., and Wilusz, C. J. (2020). YTHDF2 destabilizes m(6)A-modified neural-specific RNAs to restrain differentiation in induced pluripotent stem cells. *RNA* 26, 739–755. doi: 10.1261/rna.073502.119
- Hess, M. E., Hess, S., Meyer, K. D., Verhagen, L. A., Koch, L., Brönneke, H. S., et al. (2013). The fat mass and obesity associated gene (Fto) regulates activity of the dopaminergic midbrain circuitry. *Nat. Neurosci.* 16, 1042–1048. doi: 10.1038/nn.3449
- Hinman, M. N., and Lou, H. (2008). Diverse molecular functions of Hu proteins. *Cell. Mol. Life Sci.* 65, 3168–3181. doi: 10.1007/s00018-008-8252-6
- Ho, A. J., Stein, J. L., Hua, X., Lee, S., Hibar, D. P., Leow, A. D., et al. (2010). A commonly carried allele of the obesity-related FTO gene is associated with reduced brain volume in the healthy elderly. *Proc. Natl. Acad. Sci. U.S.A.* 107, 8404–8409. doi: 10.1073/pnas.0910878107
- Horiuchi, K., Kawamura, T., Iwanari, H., Ohashi, R., Naito, M., Kodama, T., et al. (2013). Identification of Wilms' tumor 1-associating protein complex and its role in alternative splicing and the cell cycle. *J. Biol. Chem.* 288, 33292–33302. doi: 10.1074/jbc.M113.500397
- Huang, H., Camats-Perna, J., Medeiros, R., Anggono, V., and Widagdo, J. (2020). Altered expression of the m6A methyltransferase METTL3 in Alzheimer's disease. *eNeuro* 7:ENEURO.0125-20.2020. doi: 10.1523/eneuro.0125-20.2020
- Huang, H., Weng, H., Sun, W., Qin, X., Shi, H., Wu, H., et al. (2018). Recognition of RNA N(6)-methyladenosine by IGF2BP proteins enhances mRNA stability and translation. *Nat. Cell Biol.* 20, 285–295. doi: 10.1038/s41556-018-0045-z
- Huang, H., Weng, H., Zhou, K., Wu, T., Zhao, B. S., Sun, M., et al. (2019). Histone H3 trimethylation at lysine 36 guides m6A RNA modification co-transcriptionally. *Nature* 567, 414–419. doi: 10.1038/s41586-019-1016-7
- Huang, J., and Yin, P. (2018). Structural insights into N(6)-methyladenosine (m(6)A) modification in the transcriptome. *Genomics Proteom. Bioinform.* 16, 85–98. doi: 10.1016/j.gpb.2018.03.001
- Huang, R. R., Zhang, Y., Bai, Y., Han, B., Ju, M. Z., Chen, B. L., et al. (2020). N-6-methyladenosine modification of fatty acid amide hydrolase messenger RNA in circular RNA STAG1-regulated astrocyte dysfunction and depressive-like behaviors. *Biol. Psychiat.* 88, 392–404. doi: 10.1016/j.biopsych.2020.02.018
- Huang, T., Gao, Q., Feng, T., Zheng, Y., Guo, J., and Zeng, W. (2018). FTO knockout causes chromosome instability and G2/M arrest in mouse GC-1 cells. *Front. Genet.* 9:732. doi: 10.3389/fgene.2018.00732
- Imai, Y., Matsuo, N., Ogawa, S., Tohyama, M., and Takagi, T. (1998). Cloning of a gene, YT521, for a novel RNA splicing-related protein induced by hypoxia/reoxygenation. *Brain Res. Mol. Brain Res.* 53, 33–40. doi: 10.1016/s0169-328x(97)00262-3
- Iyer, L. M., Zhang, D., and Aravind, L. (2016). Adenine methylation in eukaryotes: apprehending the complex evolutionary history and functional potential of an epigenetic modification. *BioEssays* 38, 27–40. doi: 10.1002/bies.201500104
- Jia, G., Fu, Y., Zhao, X., Dai, Q., Zheng, G., Yang, Y., et al. (2011). N6-methyladenosine in nuclear RNA is a major substrate of the obesity-associated FTO. *Nat. Chem. Biol.* 7, 885–887. doi: 10.1038/nchembio.687
- Jia, G., Yang, C.-G., Yang, S., Jian, X., Yi, C., Zhou, Z., et al. (2008). Oxidative demethylation of 3-methylthymine and 3-methyluracil in single-stranded DNA and RNA by mouse and human FTO. *FEBS Lett.* 582, 3313–3319. doi: 10.1016/j.febslet.2008.08.019
- Jung, Y., and Goldman, D. (2018). Role of RNA modifications in brain and behavior. *Genes Brain Behav.* 17:e12444. doi: 10.1111/gbb.12444
- Kalia, L. V., and Lang, A. E. (2015). Parkinson's disease. *Lancet* 386, 896–912. doi: 10.1016/S0140-6736(14)61393-3
- Kang, H.-J., Jeong, S.-J., Kim, K.-N., Baek, I.-J., Chang, M., Kang, C.-M., et al. (2014). A novel protein, Pho92, has a conserved YTH domain and regulates phosphate metabolism by decreasing the mRNA stability of PHO4 in *Saccharomyces cerevisiae*. *Biochem. J.* 457, 391–400. doi: 10.1042/bj20130862
- Kaufmann, W. E., Kidd, S. A., Andrews, H. F., Budimirovic, D. B., Esler, A., Haas-Givler, B., et al. (2017). Autism spectrum disorder in fragile X syndrome: cooccurring conditions and current treatment. *Pediatrics* 139, S194–S206. doi: 10.1542/peds.2016-1159F
- Ke, S., Alemu, E. A., Mertens, C., Gantman, E. C., Fak, J. J., Mele, A., et al. (2015). A majority of m6A residues are in the last exons, allowing the potential for 3' UTR regulation. *Genes Dev.* 29, 2037–2053. doi: 10.1101/gad.269415.115
- Keller, L., Xu, W., Wang, H. X., Winblad, B., Fratiglioni, L., and Graff, C. (2011). The obesity related gene, FTO, interacts with APOE, and is associated with Alzheimer's disease risk: a prospective cohort study. *J. Alzheimer's Dis.* 23, 461–469. doi: 10.3233/jad-2010-101068
- Kiernan, M. C., Vucic, S., Cheah, B. C., Turner, M. R., Eisen, A., Hardiman, O., et al. (2011). Amyotrophic lateral sclerosis. *Lancet* 377, 942–955. doi: 10.1016/S0140-6736(10)61156-7
- Kim, H. J., Kim, N. C., Wang, Y. D., Scarborough, E. A., Moore, J., Diaz, Z., et al. (2013). Mutations in prion-like domains in hnRNPA2B1 and hnRNPA1 cause multisystem proteinopathy and ALS. *Nature* 495, 467–473. doi: 10.1038/nature11922
- Kisliouk, T., Rosenberg, T., Ben-Nun, O., Ruzal, M., and Meiri, N. (2020). Early-Life m(6)A RNA demethylation by fat mass and obesity-associated protein

- (FTO) influences resilience or vulnerability to heat stress later in life. *eNeuro* 7, ENEURO.0549-19.2020. doi: 10.1523/ENEURO.0549-19.2020
- Knuckles, P., Lence, T., Haussmann, I. U., Jacob, D., Kreim, N., Carl, S. H., et al. (2018). Zc3h13/Flacc is required for adenosine methylation by bridging the mRNA-binding factor Rbm15/Spenito to the m(6)A machinery component Wtap/Fl(2)d. *Genes Dev.* 32, 415–429. doi: 10.1101/gad.309146.117
- Kontur, C., Jeong, M., Cifuentes, D., and Giraldez, A. J. (2020). Ythdf m(6)A readers function redundantly during Zebrafish development. *Cell Rep.* 33:108598. doi: 10.1016/j.celrep.2020.108598
- Koranda, J. L., Dore, L., Shi, H. L., Patel, M. J., Vaasjo, L. O., Rao, M. N., et al. (2018). Mettl14 is essential for epitranscriptomic regulation of striatal function and learning. *Neuron* 99, 283–292. doi: 10.1016/j.neuron.2018.06.007
- Kretschmer, J., Rao, H., Hackert, P., Sloan, K. E., Höbartner, C., and Bohnsack, M. T. (2018). The m6A reader protein YTHDC2 interacts with the small ribosomal subunit and the 5′–3′ exoribonuclease XRN1. *RNA* 24, 1339–1350. doi: 10.1261/rna.064238.117
- Krogan, N. J., Cagney, G., Yu, H., Zhong, G., Guo, X., Ignatchenko, A., et al. (2006). Global landscape of protein complexes in the yeast *Saccharomyces cerevisiae*. *Nature* 440, 637–643. doi: 10.1038/nature04670
- Lasman, L., Krupalnik, V., Viukov, S., Mor, N., Aguilera-Castrejon, A., Schneir, D., et al. (2020). Context-dependent functional compensation between Ythdf m(6)A reader proteins. *Genes Dev.* 34, 1373–1391. doi: 10.1101/gad.340695.120
- Leighton, L. J., Ke, K., Zajackowski, E. L., Edmunds, J., Spitalo, R. C., and Bredy, T. W. (2018). Experience-dependent neural plasticity, learning, and memory in the era of epitranscriptomics. *Genes Brain Behav.* 17:e12426. doi: 10.1111/gbb.12426
- Lesbirel, S., Viphacone, N., Parker, M., Parker, J., Heath, C., Sudbery, I., et al. (2018). The m(6)A-methylase complex recruits TREX and regulates mRNA export. *Sci. Rep.* 8:13827. doi: 10.1038/s41598-018-32310-8
- Lesbirel, S., and Wilson, S. A. (2019). The m(6)A-methylase complex and mRNA export. *Biochim. Biophys. Acta Gene Regul. Mech.* 1862, 319–328. doi: 10.1016/j.bbargrm.2018.09.008
- Li, A., Chen, Y.-S., Ping, X.-L., Yang, X., Xiao, W., Yang, Y., et al. (2017). Cytoplasmic m6A reader YTHDF3 promotes mRNA translation. *Cell Res.* 27:444. doi: 10.1038/cr.2017.10
- Li, F., Zhao, D., Wu, J., and Shi, Y. (2014). Structure of the YTH domain of human YTHDF2 in complex with an m6A mononucleotide reveals an aromatic cage for m6A recognition. *Cell Res.* 24:1490. doi: 10.1038/cr.2014.153
- Li, H., Ren, Y., Mao, K., Hua, F., Yang, Y., Wei, N., et al. (2018). FTO is involved in Alzheimer's disease by targeting TSC1-mTOR-Tau signaling. *Biochem. Biophys. Res. Commun.* 498, 234–239. doi: 10.1016/j.bbrc.2018.02.201
- Li, J., Chen, Z., Chen, F., Xie, G., Ling, Y., Peng, Y., et al. (2020). Targeted mRNA demethylation using an engineered dCas13b-ALKBH5 fusion protein. *Nucleic Acids Res.* 48, 5684–5694. doi: 10.1093/nar/gkaa269
- Li, J., Yang, X., Qi, Z., Sang, Y., Liu, Y., Xu, B., et al. (2019). The role of mRNA m(6)A methylation in the nervous system. *Cell Biosci.* 9:66. doi: 10.1186/s13578-019-0330-y
- Li, L., Zang, L., Zhang, F., Chen, J., Shen, H., Shu, L., et al. (2017). Fat mass and obesity-associated (FTO) protein regulates adult neurogenesis. *Hum. Mol. Genet.* 26, 2398–2411. doi: 10.1093/hmg/ddx128
- Li, M., Zhao, X., Wang, W., Shi, H., Pan, Q., Lu, Z., et al. (2018). Ythdf2-mediated m(6)A mRNA clearance modulates neural development in mice. *Genome Biol.* 19:69. doi: 10.1186/s13059-018-1436-y
- Lin, S., Choe, J., Du, P., Triboulet, R., and Gregory, R. I. (2016). The m(6)A methyltransferase METTL3 promotes translation in human cancer cells. *Mol. Cell* 62, 335–345. doi: 10.1016/j.molcel.2016.03.021
- Linder, B., Grozhik, A. V., Olaverin-George, A. O., Meydan, C., Mason, C. E., and Jaffrey, S. R. (2015). Single-nucleotide-resolution mapping of m6A and m6Am throughout the transcriptome. *Nat. Methods* 12, 767–772. doi: 10.1038/nmeth.3453
- Liu, J., Yue, Y., Han, D., Wang, X., Fu, Y., Zhang, L., et al. (2014). A METTL3-METTL14 complex mediates mammalian nuclear RNA N6-adenosine methylation. *Nat. Chem. Biol.* 10, 93–95. doi: 10.1038/nchembio.1432
- Liu, K., Ding, Y., Ye, W., Liu, Y., Yang, J., Liu, J., et al. (2016). Structural and functional characterization of the proteins responsible for N(6)-methyladenosine modification and recognition. *Curr. Protein Pept. Sci.* 17, 306–318. doi: 10.2174/1389203716666150901113553
- Liu, N., Dai, Q., Zheng, G., He, C., Parisien, M., and Pan, T. (2015). N(6)-methyladenosine-dependent RNA structural switches regulate RNA-protein interactions. *Nature* 518, 560–564. doi: 10.1038/nature14234
- Liu, N., Zhou, K. I., Parisien, M., Dai, Q., Diatchenko, L., and Pan, T. (2017). N6-methyladenosine alters RNA structure to regulate binding of a low-complexity protein. *Nucleic Acids Res.* 45, 6051–6063. doi: 10.1093/nar/gkx141
- Liu, Z., Hu, X., Cai, J., Liu, B., Peng, X., Wegner, M., et al. (2007). Induction of oligodendrocyte differentiation by Olig2 and Sox10: evidence for reciprocal interactions and dosage-dependent mechanisms. *Dev. Biol.* 302, 683–693. doi: 10.1016/j.ydbio.2006.10.007
- Livneh, I., Moshitch-Moshkovitz, S., Amariglio, N., Rechavi, G., and Dominissini, D. (2020). The m(6)A epitranscriptome: transcriptome plasticity in brain development and function. *Nat. Rev. Neurosci.* 21, 36–51. doi: 10.1038/s41583-019-0244-z
- Ma, C., Chang, M., Lv, H., Zhang, Z. W., Zhang, W., He, X., et al. (2018). RNA m(6)A methylation participates in regulation of postnatal development of the mouse cerebellum. *Genome Biol.* 19:68. doi: 10.1186/s13059-018-1435-z
- Masuda, S., Das, R., Cheng, H., Hurt, E., Dorman, N., and Reed, R. (2005). Recruitment of the human TREX complex to mRNA during splicing. *Genes Dev.* 19, 1512–1517. doi: 10.1101/gad.1302205
- Mauer, J., Luo, X., Blanjoie, A., Jiao, X., Grozhik, A. V., Patil, D. P., et al. (2017). Reversible methylation of m(6)A(m) in the 5′ cap controls mRNA stability. *Nature* 541, 371–375. doi: 10.1038/nature21022
- McEwen, B. S., Bowles, N. P., Gray, J. D., Hill, M. N., Hunter, R. G., Karatsoreos, I. N., et al. (2015). Mechanisms of stress in the brain. *Nat. Neurosci.* 18, 1353–1363. doi: 10.1038/nn.4086
- Merkurjev, D., Hong, W. T., Iida, K., Oomoto, I., Goldie, B. J., Yamaguti, H., et al. (2018). Synaptic N(6)-methyladenosine (m(6)A) epitranscriptome reveals functional partitioning of localized transcripts. *Nat. Neurosci.* 21, 1004–1014. doi: 10.1038/s41593-018-0173-6
- Meyer, K. D., and Jaffrey, S. R. (2017). Rethinking m(6)A readers, writers, and erasers. *Annu. Rev. Cell Dev. Biol.* 33, 319–342. doi: 10.1146/annurev-cellbio-100616-060758
- Meyer, K. D., Patil, D. P., Zhou, J., Zinoviev, A., Skabkin, M. A., Elemento, O., et al. (2015). 5′ UTR m(6)A promotes cap-independent translation. *Cell* 163, 999–1010. doi: 10.1016/j.cell.2015.10.012
- Meyer, K. D., Saletore, Y., Zumbo, P., Elemento, O., Mason, C. E., and Jaffrey, S. R. (2012). Comprehensive analysis of mRNA methylation reveals enrichment in 3′ UTRs and near stop codons. *Cell* 149, 1635–1646. doi: 10.1016/j.cell.2012.05.003
- Milaneschi, Y., Lamers, F., Mbarek, H., Hottenga, J. J., Boomsma, D. I., and Penninx, B. W. (2014). The effect of FTO rs9939609 on major depression differs across MDD subtypes. *Mol. Psychiatry* 19, 960–962. doi: 10.1038/mp.2014.4
- Mitropoulos, K., Merkouri Papadima, E., Xiromerisiou, G., Balasopoulou, A., Charalampidou, K., Galani, V., et al. (2017). Genomic variants in the FTO gene are associated with sporadic amyotrophic lateral sclerosis in Greek patients. *Hum. Genomics* 11:30. doi: 10.1186/s40246-017-0126-2
- Mo, J., Chen, Z., Qin, S., Li, S., Liu, C., Zhang, L., et al. (2020). TRADES: targeted RNA demethylation by SunTag system. *Adv. Sci.* 7:2001402. doi: 10.1002/adv.202001402
- Mo, X. B., Lei, S. F., Qian, Q. Y., Guo, Y. F., Zhang, Y. H., and Zhang, H. (2019). Integrative analysis revealed potential causal genetic and epigenetic factors for multiple sclerosis. *J. Neurol.* 266, 2699–2709. doi: 10.1007/s00415-019-09476-w
- Nainar, S., Marshall, P. R., Tyler, C. R., Spitalo, R. C., and Bredy, T. W. (2016). Evolving insights into RNA modifications and their functional diversity in the brain. *Nat. Neurosci.* 19, 1292–1298. doi: 10.1038/nn.4378
- Narayan, P., and Rottman, F. (1988). An in vitro system for accurate methylation of internal adenosine residues in messenger RNA. *Science* 242, 1159–1162. doi: 10.1126/science.3187541
- Narayanan, R., Pham, L., Kerimoglu, C., Watanabe, T., Hernandez, R. C., Sokpor, G., et al. (2018). Chromatin remodeling BAF155 subunit regulates the genesis of basal progenitors in developing cortex. *Iscience* 4, 109–126. doi: 10.1016/j.isci.2018.05.014
- Nguyen, H., Kerimoglu, C., Pirouz, M., Pham, L., Kiszka, K. A., Sokpor, G., et al. (2018). Epigenetic regulation by BAF complexes limits neural stem cell proliferation by suppressing wnt signaling in late embryonic development. *Stem Cell Rep.* 10, 1734–1750. doi: 10.1016/j.stemcr.2018.04.014

- Nguyen, H., Sokpor, G., Pham, L., Rosenbusch, J., Stoykova, A., Staiger, J. F., et al. (2016). Epigenetic regulation by BAF (mSWI/SNF) chromatin remodeling complexes is indispensable for embryonic development. *Cell Cycle* 15, 1317–1324. doi: 10.1080/15384101.2016.1160984
- Noack, F., and Calegari, F. (2018). Epitranscriptomics: a new regulatory mechanism of brain development and function. *Front. Neurosci.* 12:85. doi: 10.3389/fnins.2018.00085
- Oldmeadow, C., Mossman, D., Evans, T. J., Holliday, E. G., Tooney, P. A., Cairns, M. J., et al. (2014). Combined analysis of exon splicing and genome wide polymorphism data predict schizophrenia risk loci. *J. Psychiatr. Res.* 52, 44–49. doi: 10.1016/j.jpsychires.2014.01.011
- Otte, C., Gold, S. M., Penninx, B. W., Pariante, C. M., Etkin, A., Fava, M., et al. (2016). Major depressive disorder. *Nat. Rev. Dis. Primers* 2:16065. doi: 10.1038/nrdp.2016.65
- Park, O. H., Ha, H., Lee, Y., Boo, S. H., Kwon, D. H., Song, H. K., et al. (2019). Endoribonucleolytic cleavage of m6A-containing RNAs by RNase P/MRP complex. *Mol. Cell* 74, 494.e8–507.e8. doi: 10.1016/j.molcel.2019.02.034
- Patil, D. P., Chen, C. K., Pickering, B. F., Chow, A., Jackson, C., Guttman, M., et al. (2016). m(6)A RNA methylation promotes XIST-mediated transcriptional repression. *Nature* 537, 369–373. doi: 10.1038/nature19342
- Pendleton, K. E., Chen, B., Liu, K., Hunter, O. V., Xie, Y., Tu, B. P., et al. (2017). The U6 snRNA m(6)A methyltransferase METTL16 regulates SAM synthetase intron retention. *Cell* 169, 824.e14–835.e14. doi: 10.1016/j.cell.2017.05.003
- Peng, S. S., Chen, C. Y., Xu, N., and Shyu, A. B. (1998). RNA stabilization by the AU-rich element binding protein, HuR, an ELAV protein. *EMBO J.* 17, 3461–3470. doi: 10.1093/emboj/17.12.3461
- Perry, R. P., Kelley, D. E., Friderici, K., and Rottman, F. (1975a). The methylated constituents of L cell messenger RNA: evidence for an unusual cluster at the 5' terminus. *Cell* 4, 387–394. doi: 10.1016/0092-8674(75)90159-2
- Perry, R. P., Kelley, D. E., Friderici, K. H., and Rottman, F. M. (1975b). Methylated constituents of heterogeneous nuclear RNA: presence in blocked 5' terminal structures. *Cell* 6, 13–19. doi: 10.1016/0092-8674(75)90068-9
- Ping, X. L., Sun, B. F., Wang, L., Xiao, W., Yang, X., Wang, W. J., et al. (2014). Mammalian WTAP is a regulatory subunit of the RNA N6-methyladenosine methyltransferase. *Cell Res.* 24, 177–189. doi: 10.1038/cr.2014.3
- Rauch, S., He, C., and Dickinson, B. C. (2018). Targeted m(6)A reader proteins to study epitranscriptomic regulation of single RNAs. *J. Am. Chem. Soc.* 140, 11974–11981. doi: 10.1021/jacs.8b05012
- Reddy, K. S. (2005). Cytogenetic abnormalities and fragile-X syndrome in autism spectrum disorder. *BMC Med. Genet.* 6:3. doi: 10.1186/1471-2350-6-3
- Reitz, C., Tosto, G., Mayeux, R., and Luchsinger, J. A. (2012). Genetic variants in the Fat and Obesity Associated (FTO) gene and risk of Alzheimer's disease. *PLoS One* 7:e50354. doi: 10.1371/journal.pone.0050354
- Ricard, G., Molina, J., Chrast, J., Gu, W., Gheldof, N., Pradervand, S., et al. (2010). Phenotypic consequences of copy number variation: insights from Smith-Magenis and Potocki-Lupski syndrome mouse models. *PLoS Biol.* 8:e1000543. doi: 10.1371/journal.pbio.1000543
- Richard, E. M., Polla, D. L., Assir, M. Z., Contreras, M., Shahzad, M., Khan, A. A., et al. (2019). Bi-allelic variants in METTL5 cause autosomal-recessive intellectual disability and microcephaly. *Am. J. Hum. Genet.* 105, 869–878. doi: 10.1016/j.ajhg.2019.09.007
- Ries, R. J., Zaccara, S., Klein, P., Orlarier-George, A., Namkoong, S., Pickering, B. F., et al. (2019). m6A enhances the phase separation potential of mRNA. *Nature* 571, 424–428. doi: 10.1038/s41586-019-1374-1
- Rockwell, A. L., and Hongay, C. F. (2019). The m(6)A dynamics of profilin in neurogenesis. *Front. Genet.* 10:987. doi: 10.3389/fgene.2019.00987
- Roundtree, I. A., Evans, M. E., Pan, T., and He, C. (2017a). Dynamic RNA modifications in gene expression regulation. *Cell* 169, 1187–1200. doi: 10.1016/j.cell.2017.05.045
- Roundtree, I. A., Luo, G.-Z., Zhang, Z., Wang, X., Zhou, T., Cui, Y., et al. (2017b). YTHDC1 mediates nuclear export of N(6)-methyladenosine methylated mRNAs. *eLife* 6:e31311. doi: 10.7554/eLife.31311
- Rowland, L. P., and Shneider, N. A. (2001). Amyotrophic lateral sclerosis. *N. Engl. J. Med.* 344, 1688–1700. doi: 10.1056/NEJM200105313442207
- Růžicka, K., Zhang, M., Campilho, A., Bodi, Z., Kashif, M., Saleh, M., et al. (2017). Identification of factors required for m(6) A mRNA methylation in Arabidopsis reveals a role for the conserved E3 ubiquitin ligase HAKAI. *N. Phytol.* 215, 157–172. doi: 10.1111/nph.14586
- Samaan, Z., Anand, S. S., Zhang, X., Desai, D., Rivera, M., Pare, G., et al. (2013). The protective effect of the obesity-associated rs9939609 A variant in fat mass- and obesity-associated gene on depression. *Mol. Psychiatry* 18, 1281–1286. doi: 10.1038/mp.2012.160
- Samuels, T. J., Jarvelin, A. I., Ish-Horowitz, D., and Davis, I. (2020). Imp/IGF2BP levels modulate individual neural stem cell growth and division through myc mRNA stability. *eLife* 9:e51529. doi: 10.7554/eLife.51529
- Schwartz, S., Mumbach, M. R., Jovanovic, M., Wang, T., Maciag, K., Bushkin, G. G., et al. (2014). Perturbation of m6A writers reveals two distinct classes of mRNA methylation at internal and 5' sites. *Cell Rep.* 8, 284–296. doi: 10.1016/j.celrep.2014.05.048
- Shafik, A. M., Zhang, F., Guo, Z., Dai, Q., Pajdzik, K., Li, Y., et al. (2021). N6-methyladenosine dynamics in neurodevelopment and aging, and its potential role in Alzheimer's disease. *Genome Biol.* 22:17. doi: 10.1186/s13059-020-02249-z
- Shah, A., Rashid, F., Awan, H. M., Hu, S., Wang, X., and Chen, L. (2017). The DEAD-Box RNA helicase DDX3 interacts with m(6)A RNA demethylase ALKBH5. *Stem Cell Int.* 2017:8596135. doi: 10.1155/2017/8596135
- Shi, H., Wang, X., Lu, Z., Zhao, B. S., Ma, H., Hsu, P. J., et al. (2017). YTHDF3 facilitates translation and decay of N6-methyladenosine-modified RNA. *Cell Res.* 27:315. doi: 10.1038/cr.2017.15
- Shi, H., Zhang, X., Weng, Y. L., Lu, Z., Liu, Y., Lu, Z., et al. (2018). m(6)A facilitates hippocampus-dependent learning and memory through YTHDF1. *Nature* 563, 249–253. doi: 10.1038/s41586-018-0666-1
- Shima, H., Matsumoto, M., Ishigami, Y., Ebina, M., Muto, A., Sato, Y., et al. (2017). S-Adenosylmethionine synthesis is regulated by selective N6-adenosine methylation and mRNA degradation involving METTL16 and YTHDC1. *Cell Rep.* 21, 3354–3363. doi: 10.1016/j.celrep.2017.11.092
- Skene, J. H. P., Jacobson, R. D., Snipes, G. J., McGuire, C. B., Norden, J. J., and Freeman, J. A. (1986). A protein-induced during nerve growth (Gap-43) is a major component of growth-cone membranes. *Science* 233, 783–786. doi: 10.1126/science.3738509
- Śledź, P., and Jinek, M. (2016). Structural insights into the molecular mechanism of the m(6)A writer complex. *eLife* 5:e18434. doi: 10.7554/eLife.18434
- Song, H., Feng, X., Zhang, H., Luo, Y., Huang, J., Lin, M., et al. (2019). METTL3 and ALKBH5 oppositely regulate m6A modification of TFEB mRNA, which dictates the fate of hypoxia/reoxygenation-treated cardiomyocytes. *Autophagy* 15, 1419–1437. doi: 10.1080/15548627.2019.1586246
- Soria Lopez, J. A., González, H. M., and Léger, G. C. (2019). Alzheimer's disease. *Handbook Clin. Neurol.* 167, 231–255. doi: 10.1016/b978-0-12-804766-8.00013-3
- Spychala, A., and Ruther, U. (2019). FTO affects hippocampal function by regulation of BDNF processing. *PLoS One* 14:e0211937. doi: 10.1371/journal.pone.0211937
- Sun, L., Ma, L., Zhang, H., Cao, Y., Wang, C., Hou, N., et al. (2019). Fto deficiency reduces anxiety- and depression-like behaviors in mice via alterations in gut microbiota. *Theranostics* 9, 721–733. doi: 10.7150/thno.31562
- Tang, C., Klukovich, R., Peng, H., Wang, Z., Yu, T., Zhang, Y., et al. (2018). ALKBH5-dependent m6A demethylation controls splicing and stability of long 3'-UTR mRNAs in male germ cells. *Proc. Natl. Acad. Sci. U.S.A.* 115, E325–E333. doi: 10.1073/pnas.1717794115
- Ulmke, P. A., Xie, Y., Sokpor, G., Pham, L., Shomroni, O., Berulava, T., et al. (2021). Post-transcription regulation by the exosome complex is required for cell survival and forebrain development by repressing P53 signaling. *Development* 148:dev188276. doi: 10.1242/dev.188276
- Van Nostrand, E. L., Pratt, G. A., Shishkin, A. A., Gelboin-Burkhart, C., Fang, M. Y., Sundararaman, B., et al. (2016). Robust transcriptome-wide discovery of RNA-binding protein binding sites with enhanced CLIP (eCLIP). *Nat. Methods* 13, 508–514. doi: 10.1038/nmeth.3810
- Velders, F. P., De Wit, J. E., Jansen, P. W., Jaddoe, V. W., Hofman, A., Verhulst, F. C., et al. (2012). FTO at rs9939609, food responsiveness, emotional control and symptoms of ADHD in preschool children. *PLoS One* 7:e49131. doi: 10.1371/journal.pone.0049131
- Verkerk, A. J., Pieretti, M., Sutcliffe, J. S., Fu, Y. H., Kuhl, D. P., Pizzuti, A., et al. (1991). Identification of a gene (FMR-1) containing a CGG repeat coincident with a breakpoint cluster region exhibiting length variation in fragile X syndrome. *Cell* 65, 905–914. doi: 10.1016/0092-8674(91)90397-h

- Walters, B. J., Mercaldo, V., Gillon, C. J., Yip, M., Neve, R. L., Boyce, F. M., et al. (2017). The role of The RNA demethylase FTO (Fat Mass and Obesity-Associated) and mRNA methylation in hippocampal memory formation. *Neuropsychopharmacology* 42, 1502–1510. doi: 10.1038/npp.2017.31
- Wang, C. X., Cui, G. S., Liu, X., Xu, K., Wang, M., Zhang, X. X., et al. (2018). METTL3-mediated m6A modification is required for cerebellar development. *PLoS Biol.* 16:e2004880. doi: 10.1371/journal.pbio.2004880
- Wang, F., Wang, L., Zou, X., Duan, S., Li, Z., Deng, Z., et al. (2019). Advances in CRISPR-Cas systems for RNA targeting, tracking and editing. *Biotechnol. Adv.* 37, 708–729. doi: 10.1016/j.biotechadv.2019.03.016
- Wang, P., Dostader, K. A., and Nam, Y. (2016). Structural basis for cooperative function of Mettl3 and Mettl14 methyltransferases. *Mol. Cell* 63, 306–317. doi: 10.1016/j.molcel.2016.05.041
- Wang, X., Feng, J., Xue, Y., Guan, Z., Zhang, D., Liu, Z., et al. (2016). Structural basis of N6-adenosine methylation by the METTL3–METTL14 complex. *Nature* 534:575. doi: 10.1038/nature18298
- Wang, X., Lu, Z., Gomez, A., Hon, G. C., Yue, Y., Han, D., et al. (2014). N6-methyladenosine-dependent regulation of messenger RNA stability. *Nature* 505, 117–120. doi: 10.1038/nature12730
- Wang, X., Zhao, B. S., Roundtree, I. A., Lu, Z., Han, D., Ma, H., et al. (2015). N(6)-methyladenosine modulates messenger RNA translation efficiency. *Cell* 161, 1388–1399. doi: 10.1016/j.cell.2015.05.014
- Wang, Y., Li, Y., Toth, J. I., Petroski, M. D., Zhang, Z., and Zhao, J. C. (2014). N6-methyladenosine modification destabilizes developmental regulators in embryonic stem cells. *Nat. Cell Biol.* 16, 191–198. doi: 10.1038/ncb2902
- Wang, Y., Li, Y., Yue, M. H., Wang, J., Kumar, S., Wechsler-Reya, R. J., et al. (2018). N6-methyladenosine RNA modification regulates embryonic neural stem cell self-renewal through histone modifications. *Nat. Neurosci.* 21, 195–206. doi: 10.1038/s41593-017-0057-1
- Wang, Y., and Zhao, J. C. (2016). Update: mechanisms underlying N(6)-methyladenosine modification of eukaryotic mRNA. *Trends Genet.* 32, 763–773. doi: 10.1016/j.tig.2016.09.006
- Wei, H., Peng, J., Lu, J., Song, T., Xie, X., Yang, Y., et al. (2019). Zfp217 mediates m6A mRNA methylation to orchestrate transcriptional and post-transcriptional regulation to promote adipogenic differentiation. *Nucleic Acids Res.* 47, 6130–6144. doi: 10.1093/nar/gkz312
- Wei, J., Liu, F., Lu, Z., Fei, Q., Ai, Y., He, P. C., et al. (2018). Differential m6A, m6Am, and m1A demethylation mediated by FTO in the cell nucleus and cytoplasm. *Mol. Cell.* 71, 973.e5–985.e5. doi: 10.1016/j.molcel.2018.08.011
- Weller, J., and Budson, A. (2018). Current understanding of Alzheimer's disease diagnosis and treatment. *F1000Research* 7:F1000 Faculty Rev–1161. doi: 10.12688/f1000research.14506.1
- Wen, J., Lv, R., Ma, H., Shen, H., He, C., Wang, J., et al. (2018). Zc3h13 regulates nuclear RNA m(6)A methylation and mouse embryonic stem cell self-renewal. *Mol. Cell.* 69, 1028.e6–1038.e6. doi: 10.1016/j.molcel.2018.02.015
- Weng, Y. L., Wang, X., An, R., Cassin, J., Vissers, C., Liu, Y., et al. (2018). Epitranscriptomic m(6)A regulation of axon regeneration in the adult mammalian nervous system. *Neuron* 97, 313.e6–325.e6. doi: 10.1016/j.neuron.2017.12.036
- Widagdo, J., and Anggono, V. (2018). The m6A-epitranscriptomic signature in neurobiology: from neurodevelopment to brain plasticity. *J. Neurochem.* 147, 137–152. doi: 10.1111/jnc.14481
- Widagdo, J., Zhao, Q. Y., Kempen, M. J., Tan, M. C., Ratnu, V. S., Wei, W., et al. (2016). Experience-dependent accumulation of N6-methyladenosine in the prefrontal cortex is associated with memory processes in mice. *J. Neurosci.* 36, 6771–6777. doi: 10.1523/jneurosci.4053-15.2016
- Wilson, C., Chen, P. J., Miao, Z., and Liu, D. R. (2020). Programmable m(6)A modification of cellular RNAs with a Cas13-directed methyltransferase. *Nat. Biotechnol.* 38, 1431–1440. doi: 10.1038/s41587-020-0572-6
- Wu, B., Su, S., Patil, D. P., Liu, H., Gan, J., Jaffrey, S. R., et al. (2018). Molecular basis for the specific and multivalent recognitions of RNA substrates by human hnRNP A2/B1. *Nat. Commun.* 9:420. doi: 10.1038/s41467-017-02770-z
- Wu, R., Li, A., Sun, B., Sun, J.-G., Zhang, J., Zhang, T., et al. (2019). A novel m(6)A reader Prrc2a controls oligodendroglial specification and myelination. *Cell Res.* 29, 23–41. doi: 10.1038/s41422-018-0113-8
- Xiao, W., Adhikari, S., Dahal, U., Chen, Y.-S., Hao, Y.-J., Sun, B.-F., et al. (2016). Nuclear m6A reader YTHDC1 regulates mRNA splicing. *Mol. Cell* 61, 507–519. doi: 10.1016/j.molcel.2016.01.012
- Xie, Y. B., Castro-Hernandez, R., Sokpor, G., Pham, L., Narayanan, R., Rosenbusch, J., et al. (2019). RBM15 modulates the function of chromatin remodeling factor BAF155 through RNA methylation in developing cortex. *Mol. Neurobiol.* 56, 7305–7320. doi: 10.1007/s12035-019-1595-1
- Xu, C., Liu, K., Ahmed, H., Loppnau, P., Schapira, M., and Min, J. (2015). Structural basis for the discriminative recognition of N6-methyladenosine RNA by the human YTH21-B homology domain family of proteins. *J. Biol. Chem.* 290, 24902–24913. doi: 10.1074/jbc.M115.680389
- Xu, C., Wang, X., Liu, K., Roundtree, I. A., Tempel, W., Li, Y., et al. (2014). Structural basis for selective binding of m6A RNA by the YTHDC1 YTH domain. *Nat. Chem. Biol.* 10:927. doi: 10.1038/nchembio.1654
- Xu, H., Dzhashishvili, Y., Shah, A., Kunjamma, R. B., Weng, Y. L., Elbaz, B., et al. (2020). m(6)A mRNA methylation is essential for oligodendrocyte maturation and CNS myelination. *Neuron* 105, 293.e5–309.e5. doi: 10.1016/j.neuron.2019.12.013
- Yang, F., Jin, H., Que, B., Chao, Y., Zhang, H., Ying, X., et al. (2019). Dynamic m6A mRNA methylation reveals the role of METTL3–m6A–CDPC1 signaling axis in chemical carcinogenesis. *Oncogene* 38, 4755–4772. doi: 10.1038/s41388-019-0755-0
- Yao, B., Christian, K. M., He, C., Jin, P., Ming, G. L., and Song, H. (2016). Epigenetic mechanisms in neurogenesis. *Nat. Rev. Neurosci.* 17, 537–549. doi: 10.1038/nrn.2016.70
- Yoon, K. J., Ringeling, F. R., Vissers, C., Jacob, F., Pokrass, M., Jimenez-Cyrus, D., et al. (2017). Temporal control of mammalian cortical neurogenesis by m(6)A methylation. *Cell* 171, 877.e17–889.e17. doi: 10.1016/j.cell.2017.09.003
- Yoon, K. J., Vissers, C., Ming, G. L., and Song, H. (2018). Epigenetics and epitranscriptomics in temporal patterning of cortical neural progenitor competence. *J. Cell Biol.* 217, 1901–1914. doi: 10.1083/jcb.2018.02.117
- Yu, J., Chen, M., Huang, H., Zhu, J., Song, H., Zhu, J., et al. (2018). Dynamic m6A modification regulates local translation of mRNA in axons. *Nucleic Acids Res.* 46, 1412–1423. doi: 10.1093/nar/gkx1182
- Yue, Y., Liu, J., Cui, X., Cao, J., Luo, G., Zhang, Z., et al. (2018). VIRMA mediates preferential m(6)A mRNA methylation in 3'UTR and near stop codon and associates with alternative polyadenylation. *Cell Discov.* 4:10. doi: 10.1038/s41421-018-0019-0
- Zaccara, S., and Jaffrey, S. R. (2020). A unified model for the function of YTHDF proteins in regulating m(6)A-modified mRNA. *Cell* 181, 1582.e18–1595.e18. doi: 10.1016/j.cell.2020.05.012
- Zarnack, K., König, J., Tajnik, M., Martincorena, I., Eustermann, S., Stévant, I., et al. (2013). Direct competition between hnRNP C and U2AF65 protects the transcriptome from the exonization of Alu elements. *Cell* 152, 453–466. doi: 10.1016/j.cell.2012.12.023
- Zhang, F., Kang, Y., Wang, M., Li, Y., Xu, T., Yang, W., et al. (2018). Fragile X mental retardation protein modulates the stability of its m6A-marked messenger RNA targets. *Hum. Mol. Genet.* 27, 3936–3950. doi: 10.1093/hmg/ddy292
- Zhang, L., Du, K., Wang, J., Nie, Y., Lee, T., and Sun, T. (2020). Unique and specific m(6)A RNA methylation in mouse embryonic and postnatal cerebral cortices. *Genes* 11:1139. doi: 10.3390/genes11101139
- Zhang, X., Wei, L.-H., Wang, Y., Xiao, Y., Liu, J., Zhang, W., et al. (2019). Structural insights into FTO's catalytic mechanism for the demethylation of multiple RNA substrates. *Proc. Natl. Acad. Sci. U.S.A.* 116, 2919–2924. doi: 10.1073/pnas.1820574116
- Zhang, Y., Wang, X., Zhang, X., Wang, J., Ma, Y., Zhang, L., et al. (2019). RNA-binding protein YTHDF3 suppresses interferon-dependent antiviral responses by promoting FOXO3 translation. *Proc. Natl. Acad. Sci. U.S.A.* 116, 976–981. doi: 10.1073/pnas.1812536116
- Zhang, Z., Theler, D., Kaminska, K. H., Hiller, M., de la Grange, P., Pudimat, R., et al. (2010). The YTH domain is a novel RNA binding domain. *J. Biol. Chem.* 285, 14701–14710. doi: 10.1074/jbc.M110.104711
- Zhang, Z. Y., Wang, M., Xie, D. F., Huang, Z. H., Zhang, L. S., Yang, Y., et al. (2018). METTL3-mediated N6-methyladenosine mRNA modification enhances long-term memory consolidation. *Cell Res.* 28, 1050–1061. doi: 10.1038/s41422-018-0092-9
- Zhao, B. S., Roundtree, I. A., and He, C. (2017). Post-transcriptional gene regulation by mRNA modifications. *Nat. Rev. Mol. Cell Biol.* 18, 31–42. doi: 10.1038/nrm.2016.132

- Zhao, J., Li, B., Ma, J., Jin, W., and Ma, X. (2020). Photoactivatable RNA N(6)-methyladenosine editing with CRISPR-Cas13. *Small* 16:e1907301. doi: 10.1002/smll.201907301
- Zhao, X., Yang, Y., Sun, B.-F., Shi, Y., Yang, X., Xiao, W., et al. (2014). FTO-dependent demethylation of N6-methyladenosine regulates mRNA splicing and is required for adipogenesis. *Cell Res.* 24, 1403–1419. doi: 10.1038/cr.2014.151
- Zheng, G., Dahl, J. A., Niu, Y., Fedorcsak, P., Huang, C.-M., Li, C. J., et al. (2013). ALKBH5 is a mammalian RNA demethylase that impacts RNA metabolism and mouse fertility. *Mol. Cell* 49, 18–29. doi: 10.1016/j.molcel.2012.10.015
- Zhong, S., Li, H., Bodi, Z., Button, J., Vespa, L., Herzog, M., et al. (2008). MTA is an *Arabidopsis* messenger RNA adenosine methylase and interacts with a homolog of a sex-specific splicing factor. *Plant Cell* 20, 1278–1288. doi: 10.1105/tpc.108.058883
- Zhou, H., Wang, B., Sun, H., Xu, X., and Wang, Y. (2018). Epigenetic regulations in neural stem cells and neurological diseases. *Stem cells Int.* 2018:6087143. doi: 10.1155/2018/6087143
- Zhou, J., Wan, J., Gao, X., Zhang, X., Jaffrey, S. R., and Qian, S. B. (2015). Dynamic m(6)A mRNA methylation directs translational control of heat shock response. *Nature* 526, 591–594. doi: 10.1038/nature15377
- Zhou, J., Wan, J., Shu, X. E., Mao, Y., Liu, X.-M., Yuan, X., et al. (2018). N(6)-methyladenosine guides mRNA alternative translation during integrated stress response. *Mol. Cell* 69, 636.e7–647.e7. doi: 10.1016/j.molcel.2018.01.019
- Zhu, C., and Yi, C. (2014). Switching demethylation activities between AlkB family RNA/DNA demethylases through exchange of active-site residues. *Angew. Chem. Int. Ed.* 53, 3659–3662. doi: 10.1002/anie.201310050
- Zhuang, M., Li, X., Zhu, J., Zhang, J., Niu, F., Liang, F., et al. (2019). The m6A reader YTHDF1 regulates axon guidance through translational control of Robo3.1 expression. *Nucleic Acids Res.* 47, 4765–4777. doi: 10.1093/nar/gkz157
- Zou, S., Toh, J. D. W., Wong, K. H. Q., Gao, Y.-G., Hong, W., and Woon, E. C. Y. (2016). N(6)-Methyladenosine: a conformational marker that regulates the substrate specificity of human demethylases FTO and ALKBH5. *Sci. Rep.* 6:25677. doi: 10.1038/srep25677

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 Sokpor, Xie, Nguyen and Tuoc. 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.



REW-ISA V2: A Biclustering Method Fusing Homologous Information for Analyzing and Mining Epi-Transcriptome Data

Lin Zhang^{1,2}, Shutao Chen^{1,2}, Jiani Ma^{1,2}, Zhaoyang Liu^{1,2} and Hui Liu^{1,2*}

¹ Engineering Research Center of Intelligent Control for Underground Space, China University of Mining and Technology, Ministry of Education, Xuzhou, China, ² School of Information and Control Engineering, China University of Mining and Technology, Xuzhou, China

OPEN ACCESS

Edited by:

Giovanni Nigita,
The Ohio State University,
United States

Reviewed by:

Rui Henriques,
Universidade de Lisboa, Portugal
Jie Jiang,
Xi'an Jiaotong-Liverpool
University, China
Rathipriya R,
Periyar University, India

*Correspondence:

Hui Liu
hui.liu@cumt.edu.cn

Specialty section:

This article was submitted to
Epigenomics and Epigenetics,
a section of the journal
Frontiers in Genetics

Received: 17 January 2021

Accepted: 28 April 2021

Published: 28 May 2021

Citation:

Zhang L, Chen S, Ma J, Liu Z and
Liu H (2021) REW-ISA V2: A
Biclustering Method Fusing
Homologous Information for Analyzing
and Mining Epi-Transcriptome Data.
Front. Genet. 12:654820.
doi: 10.3389/fgene.2021.654820

Background: Previous studies have shown that N6-methyladenosine (m⁶A) is related to many life processes and physiological and pathological phenomena. However, the specific regulatory mechanism of m⁶A sites at the systematic level is not clear. Therefore, mining the RNA co-methylation patterns in the epi-transcriptome data is expected to explain the specific regulation mechanism of m⁶A.

Methods: Considering that the epi-transcriptome data contains homologous information (the genes corresponding to the m⁶A sites and the cell lines corresponding to the experimental conditions), rational use of this information will help reveal the regulatory mechanism of m⁶A. Therefore, based on the RNA expression weighted iterative signature algorithm (REW-ISA), we have fused homologous information and developed the REW-ISA V2 algorithm.

Results: Then, REW-ISA V2 was applied in the MERIP-seq data to find potential local function blocks (LFBs), where sites are hyper-methylated simultaneously across the specific conditions. Finally, REW-ISA V2 obtained fifteen LFBs. Compared with the most advanced biclustering algorithm, the LFBs obtained by REW-ISA V2 have more significant biological significance. Further biological analysis showed that these LFBs were highly correlated with some signal pathways and m⁶A methyltransferase.

Conclusion: REW-ISA V2 fuses homologous information to mine co-methylation patterns in the epi-transcriptome data, in which sites are co-methylated under specific conditions.

Keywords: m⁶A methylation, homologous information, iterative signature algorithm, biclustering, unsupervised learning

INTRODUCTION

At present, researchers have identified more than 170 different chemical modifications in RNA (Frye et al., 2018). N⁶-methyladenine (m⁶A) is the most common and abundant post-transcriptional RNA modification in mRNAs and long non-coding RNAs (Fu et al., 2014), and its methylation occurs at the sixth position of nitrogen atoms of adenosine. Studies have shown

that m⁶A is involved in some RNA metabolic processes such as mRNA transcription, translation, nucleation, splicing and degradation (Ping et al., 2014; Lin et al., 2016; Deng et al., 2018). Besides, m⁶A also plays an important role in the early development of eukaryotic cells, sex determination, antiviral immunity, brain development, and directed differentiation of hematopoietic stem cells (Zhang et al., 2017, 2019a). In addition to the above biological processes, m⁶A modification is also related to many pathological phenomena, such as leukemia, glioma and hepatocellular carcinoma (Lachén-Montes et al., 2016; Chai et al., 2019).

The m⁶A methylation in RNA is a dynamic and reversible process regulated by methyltransferases and demethylases. Since the main role of m⁶A methyltransferases is to catalyze RNA to produce m⁶A methylation modifications, these enzymes are often called “writers.” The most common m⁶A writer is composed of core components METTL3, METTL14, WTAP, and other subunits (Liu et al., 2014; Ping et al., 2014). On the contrary, m⁶A demethylases mainly mediate m⁶A demethylation, so these enzymes are also known as “erasers.” The common erasers are FTO, AKLBH5, and so on (Jia et al., 2011). Studies have shown that m⁶A has a series of biological functions because many RNA binding proteins mediate it. These binding proteins can specifically recognize m⁶A methylated adenosine on RNA, so these proteins are often referred to as “readers.” The common readers include protein YT521-B homologous (YTH) domain family (Meyer et al., 2015), etc. In recent years, with the development of methylated RNA immunoprecipitation sequencing (MeRIP-seq, or m⁶A-seq) technology (Dominiussini et al., 2012; Meng et al., 2014), many m⁶A experimental data continue to emerge, which makes it possible to analyze m⁶A in the whole transcriptome. However, since there are a few enzymes, such as m⁶A writers, erasers and readers only, each enzyme may regulate a large number of m⁶A sites. In other words, the methylation level of m⁶A site regulated by the same enzyme may share the same pattern, which is called the co-methylation pattern of m⁶A.

Till this day, some researchers have used clustering methods to study the co-methylation patterns in epi-transcriptome data, trying to clarify the functional mechanism of m⁶A methylation. Based on MeRIP-seq data, Liu et al. used *k*-means clustering, hierarchical clustering, Bayesian factor regression model and non-negative matrix decomposition to cluster m⁶A sites (Liu et al., 2015). To better fit the distribution of epi-transcriptome data, Zhang et al. proposed an infinite beta binomial mixture model based on Dirichlet Process (DPBBM) to reveal the co-methylation patterns (Zhang et al., 2019b). Besides, our previously proposed RNA Expression Weighted Iterative Signature Algorithm (REW-ISA) (Zhang et al., 2020) applied biclustering to the analysis of epi-transcriptome data for the first time. However, the above methods only used the read counts of the m⁶A sites of the IP sample and the input sample in MeRIP-seq data. They did not fully consider the homologous information of sites and experimental conditions. Homology is a central concept in comparative biology, in which the most basic meaning of homology is to have a common ancestor. The homologous information of MeRIP-seq data can

be divided into two categories: the genes corresponding to the m⁶A sites and the cell lines (or environments) corresponding to the experimental conditions. Appropriate use of the above-mentioned homologous information will help discover potential local functional blocks (LFBs) and better reveal the m⁶A regulatory mechanism. Besides, although some of the most advanced biclustering methods have been developed, such as runibic (Wang et al., 2016; Orzechowski et al., 2018a), EBIC (Orzechowski et al., 2018b), QUBIC2 (Xie et al., 2020) and RecBic (Liu et al., 2020), their goal is to identify the trend-preserving biclusters. However, when mining m⁶A co-methylation pattern, it is expected to obtain locally hyper-methylated biclusters, so these new methods are not applicable.

Therefore, we proposed an improved RNA expression weighted iterative signature algorithm (REW-ISA V2), which fuses the homologous information of sites and experimental conditions in the iterative search for LFBs. Consistent with the previous method, each potential LFB is identified by the row threshold (defined as T_R) and column threshold (defined as T_C) during the LFB searching strategy. It is important to note that REW-ISA V2 updates T_R and T_C 's selection process, optimizing the selection of thresholds through the built-in rich constraint framework. According to the previous study (Henriques et al., 2015, 2017), REW-ISA V2 is a non-deterministic greedy algorithm, which can be used to find hyper-methylated biclusters. Besides, REW-ISA V2 can obtain these overlapping LFBs when there is overlap between the LFBs implied in the input data.

To verify the effectiveness of the fusion of homologous information, REW-ISA V2 was applied to the collected MeRIP-seq data to find potential LFBs. The obtained LFBs were further analyzed by the Gene Ontology (GO) analysis, Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis, and enzyme-specific experiments, in an attempt to reveal the possible regulatory mechanism of m⁶A. As a result, REW-ISA V2 can better find potential LFBs with high methylation levels in the epi-transcriptome data.

METHODS

Pre-processing of Real Data

As is known, MeRIP-Seq data profiles the m⁶A epi-transcriptome by IP and input samples. Thus, we first need to follow (Chen et al., 2019) and (Wu et al., 2019) to quantify the information of m⁶A sites. Specifically, after downloading the sequencing data from Gene Expression Omnibus (GEO) in SRA format, the Tophat2 (Kim et al., 2013) needs to be used to compare the sequencing data reads with the human reference genome, and finally obtain the Fragments Per Kilobase of transcript per Million (FPKM) statistics.

To mine the potential LFBs in the epi-transcriptome data, only the FPKM statistical information of IP and input samples are not enough. It is necessary to calculate the m⁶A methylation level of each m⁶A site under each experimental condition. Let m denote the total number of m⁶A sites and n denote the total number of conditions. Therefore, according to the REW-ISA, the methylation level matrix $P \in \mathbb{R}^{m \times n}$ and the RNA expression level

matrix $W \in \mathbb{R}^{m \times n}$ can be further calculated using the IP sample and the input samples, as shown in (1, 2).

$$p_{ij} = \frac{t_{ij} + \alpha}{t_{ij} + h_{ij} + 2\alpha}, \quad (1)$$

$$w_{ij} = \log(t_{ij} + h_{ij} + 1). \quad (2)$$

In (1) and (2), t_{ij} represents the FPKM of the i -th m⁶A site under the j -th condition in the IP sample, and h_{ij} represents the FPKM of the i -th m⁶A site under the j -th condition in the input sample. Besides, α in (1) is a very small value, aiming to avoid NaN where FPKM of both IP and input samples are zeros. The purpose of introducing the RNA expression level is to provide a confidence level for m⁶A methylation level in further biclustering analysis.

REW-ISA V2

To eliminate the effect of global sites or conditions on P , REW-ISA V2 performs standard normalization on the whole, rows and columns of P in turn to eliminate the global effect, as shown in (3–5). P^{nw} , P^{nr} , and P^{nc} represent the matrices obtained after whole normalization, row normalization, and column normalization, respectively.

$$p_{ij}^{nw} = \frac{p_{ij} - \text{mean}(P)}{\max(P) - \min(P)}, \quad (3)$$

$$p_{ij}^{nr} = \frac{p_{ij}^{nw} - \text{mean}(P_i^{nw})}{\max(P_i^{nw}) - \min(P_i^{nw})}, \quad (4)$$

$$p_{ij}^{nc} = \frac{p_{ij}^{nr} - \text{mean}(P_j^{nr})}{\max(P_j^{nr}) - \min(P_j^{nr})}. \quad (5)$$

In (3–5), $\text{mean}(\cdot)$ represents calculating the mean value, $\max(\cdot)$ represents calculating the maximum value, and $\min(\cdot)$ represents calculating the minimum value. P_i^{nw} represents the i -th row in P^{nw} , and P_j^{nr} represents the j -th column in P^{nr} . Then min-max normalization is performed on the overall data to generate P^t , which will facilitate subsequent combination with RNA expression level.

$$p_{ij}^t = \frac{p_{ij}^{nc} - \min(P^{nc})}{\max(P^{nc}) - \min(P^{nc})}. \quad (6)$$

For the RNA expression level matrix W , since its distribution fluctuates with the MeRIP-seq data, it is necessary to perform the min-max normalization on W to generate W^t , which acts as confidence matrix for P^t .

$$w_{ij}^t = \frac{w_{ij} - \min(W)}{\max(W) - \min(W)}. \quad (7)$$

Suppose that $k-1$ ($2 \leq k \leq K$) LFBs have been found, and the k -th LFB is currently being searched. Assuming that the k -th LFB is B_k , the site indicator ρ_k and the condition indicator κ_k are used to indicate the sites and conditions contained in B_k . Specifically, the site indication ρ_{ik} is one if the i -th site is present in B_k (zero otherwise). The condition indication κ_{jk} is one if the j -th condition is present in B_k (zero otherwise). The average

methylation level μ_k^p and average expression level μ_k^w of B_k can be further calculated, as shown in (8, 9), respectively.

$$\mu_k^p = \frac{\sum_{i=1}^m \sum_{j=1}^n p_{ij}^t \rho_{ik} \kappa_{jk}}{\sum_{i=1}^m \rho_{ik} \sum_{j=1}^n \kappa_{jk}}, \quad (8)$$

$$\mu_k^w = \frac{\sum_{i=1}^m \sum_{j=1}^n w_{ij}^t \rho_{ik} \kappa_{jk}}{\sum_{i=1}^m \rho_{ik} \sum_{j=1}^n \kappa_{jk}}. \quad (9)$$

Each time a LFB is found, the average methylation level and average expression level of the LFB should to be removed from P^t and W^t . The purpose of removing is to prevent the algorithm from falling into a loop looking for a strong LFB. Let residual matrix $P^{(k)}$ represent the methylation level matrix after eliminating the μ^p of the first $k-1$ LFBs,

$$p_{ij}^{(k)} = p_{ij}^t - \sum_{z=1}^{k-1} (\mu_z^p \rho_{iz} \kappa_{jz}). \quad (10)$$

Then, $P^{(k)}$ turns into $P^{R(k)}$ after row min-max normalization and turns into $P^{C(k)}$ after column min-max normalization. Similarly, let $W^{(k)}$ represent the RNA expression level matrix after eliminating the μ^w of the first $k-1$ LFBs,

$$w_{ij}^{(k)} = w_{ij}^t - \sum_{z=1}^{k-1} (\mu_z^w \rho_{iz} \kappa_{jz}). \quad (11)$$

After obtaining the above $P^{R(k)}$, $P^{C(k)}$ and $W^{(k)}$, combined with the homologous information of sites and conditions, the algorithm begins to search for LFBs iteratively. The algorithm running from a randomly selected site's subset U' and updates the conditions' subset V' according to (12).

$$\begin{cases} e_{U'v}^C = \frac{1}{|U'|} \sum_{u \in U'} (w_{uv}^{(k)} \cdot p_{uv}^{R(k)}) & v \in V \\ t_{U'v}^C = \left| \rho(P_{U'v}^t \cdot W_{U'v}^t, \frac{\sum_{b \in H_v^C} (P_{U'b}^t \cdot W_{U'b}^t)}{|H_v^C|}) \right| & v \in H_v^C, H_v^C \in V, \\ V' = \{v \in V : |e_{U'v}^C \cdot t_{U'v}^C - \frac{1}{|V|} \sum_{v \in V} e_{U'v}^C \cdot t_{U'v}^C| > \frac{T_C}{\sqrt{|U'|}}\} \end{cases} \quad (12)$$

where V is the conditions set of P^t , refers to the u -th site under the v -th condition in P^R , is the RNA expression level of the u -th site under v -th condition, H_v^C represents the subset of homologous conditions corresponding to the v -th condition. $\rho(\cdot)$ represents to calculate Pearson similarity, $|\cdot|$ represents to calculate absolute value (or module). Besides, T_C is a hyperparameter, and its function is to select the subset of conditions V' . In (12), $e_{U'v}^C$ is calculated based on $P^{R(k)}$ and $W^{(k)}$, which represents the average methylation level score of the v -th condition combined with the confidence of the expression level. $t_{U'v}^C$ is calculated based on P^t and W^t , representing the average similarity score of the v -th condition relative to its homologous conditions subset. In the process of calculating $e_{U'v}^C$ and, only the sites involved in U' are considered.

Then, the subsets of sites are updated following (13).

$$\begin{cases} e_{uV'}^R = \frac{1}{|V'|} \sum_{v \in V'} (w_{uv}^{(k)} \cdot p_{uv}^{C(k)}) & u \in U \\ t_{uV'}^R = \left| \rho(P_{uV'}^t \cdot W_{uV'}^t, \frac{\sum_{a \in H_u^R} (P_{aV'}^t \cdot W_{aV'}^t)}{|H_u^R|}) \right| & u \in H_u^R, H_u^R \in U, \\ U' = \{u \in U : |e_{uV'}^R \cdot t_{uV'}^R - \frac{1}{|U|} \sum_{u \in U} e_{uV'}^R \cdot t_{uV'}^R| > \frac{T_R}{\sqrt{|V'|}}\} \end{cases} \quad (13)$$

where U is the sites set of P^t , refers to the u -th site under the v -th condition in P^C , H_u^R represents the subset of homologous sites corresponding to the u -th site. Besides, T_R is a hyperparameter, and its function is to update the subset of sites U' . $e_{uV'}^R$ represents the average methylation level score of the u -th site combined with the confidence of the expression level. $t_{uV'}^R$ represents the average similarity score of the u -th site relative to its homologous sites subset. In the process of calculating $e_{uV'}^R$ and $t_{uV'}^R$, only the conditions involved in V' are considered.

Using the preset hyperparameters T_R and T_C , U' and V' are updated iteratively by (12, 13) until convergence is satisfied (or the maximum number of preset iterations is reached). The convergence condition is shown in (14).

$$\frac{|U' \cap U''|}{|U' \cup U''|} \geq \varepsilon \quad (14)$$

where ε is the default convergence criteria, and its value is slightly <1 . U'' represents the site's subset in the previous iteration, and U' represents its subset in the current iteration. If the algorithm converges within the maximum number of iterations, it means that the k -th LFB, $B_k = \{U', V'\}$ has been found. The flow chart of searching for the k -th LFB by REW-ISA V2 is shown in **Figure 1**.

Then the algorithm will return to (8) and continue to look for the next LFB. Conversely, if the convergence condition of (14) is not satisfied when the algorithm reaches the maximum number of iterations, REW-ISA V2 will automatically terminate and output all previously obtained LFBs. We recommend setting ε to 0.99 and the maximum number of iterations not <50 . The closer the value of ε is to 1 and the greater the maximum number of iterations, the more accurate the LFBs obtained by REW-ISA V2. The REW-ISA V2 algorithm based on R language can be downloaded freely from <https://github.com/labiiip/REWISAV2>.

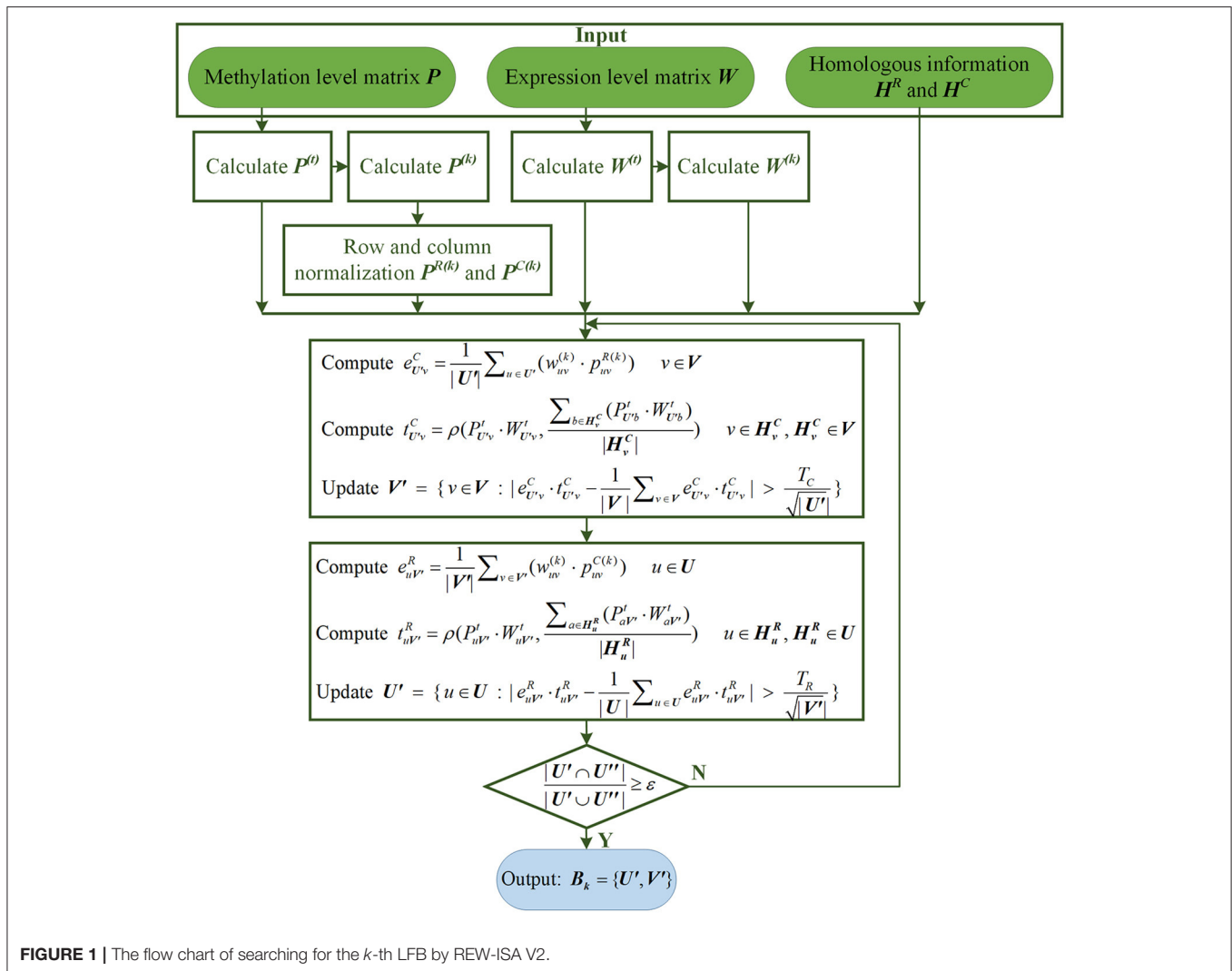


FIGURE 1 | The flow chart of searching for the k -th LFB by REW-ISA V2.

Enrichment Constraint Framework

It can be seen from (12, 13) that the selection of T_R and T_C will greatly affect the biological significance of the obtained LFBs. Therefore, based on Meng et al. (2009), we introduced a grid search-based enrichment constraint framework for the algorithm to optimize T_R and T_C selection further. For LFBs obtained under different T_R and T_C combinations, we need to extract the genes corresponding to the m⁶A sites in each LFB and then perform GO analysis based on “clusterProfiler” (Yu et al., 2012) for each LFB. For the range of T_R and T_C , we recommend setting it between 0 and 3, and the step size is 0.1. On this basis, the range of specific thresholds should be appropriately adjusted according to the input real data. Assuming that a LFB is obtained, the number of genes corresponding to the m⁶A site contained in it is M . The number of GO terms obtained by GO analysis of the LFB is l . Then the weighted enrichment score (WE_score) (Li et al., 2012) of this LFB can be calculated by (15).

$$\begin{cases} s_i = -\log(p_i) \\ \text{WE_score} = \frac{s_1 m_1 / M + s_2 m_2 / M + \dots + s_l m_l / M}{m_1 / M + m_2 / M + \dots + m_l / M + m_{\text{non}} / M} \end{cases} \quad (15)$$

where p_i is the p -value of the i -th GO term, m_i is the number of genes of the i -th GO term enriched, m_{non} is the number of genes covered by LFB but not enriched by any GO term. The higher the WE_score, the stronger the biological significance of this LFB.

However, as the number of genes corresponding to the sites in LFB increases, WE_score will also show an increasing trend, as shown in **Supplementary Figure 1**. Therefore, only using WE_score to evaluate the biological significance of obtained LFBs is not perfect, and the number of genes corresponding to the sites in LFBs also needs to be considered. Assume that the data analyzed contain a total of M_{all} genes, and further assume an obtained LFB is containing M genes and record its WE_score as W_m . We randomly select M genes from all genomes, and their WE_score is recorded as W_{rm} . The relative promotion rate (RPR) of WE_score can be further calculated, as shown in (16).

$$\text{RPR} = \frac{M(W_m - W_{rm})}{M_{\text{all}} W_{rm}}. \quad (16)$$

The larger the RPR is, the larger the area of the obtained LFB is, and the more biological significance of the obtained LFB is. On the one hand, in the actual process of mining LFBs, we hope to get more LFBs. On the other hand, we hope to get LFBs with rich biological significance. Therefore, the number of LFBs obtained by each pair of threshold combinations is obtained by grid search under different T_R and T_C combinations. The threshold combinations corresponding to the maximum number of LFBs are selected. Then, the average RPR of the LFBs is calculated based on the selected combination of T_R and T_C . Finally, the optimal T_R and T_C are the threshold combinations corresponding to the maximum average RPR.

RESULTS

We collected 32 samples from 10 publicly human m⁶A MeRIP-seq datasets (Dominissini et al., 2012; Meyer et al., 2012; Fustin

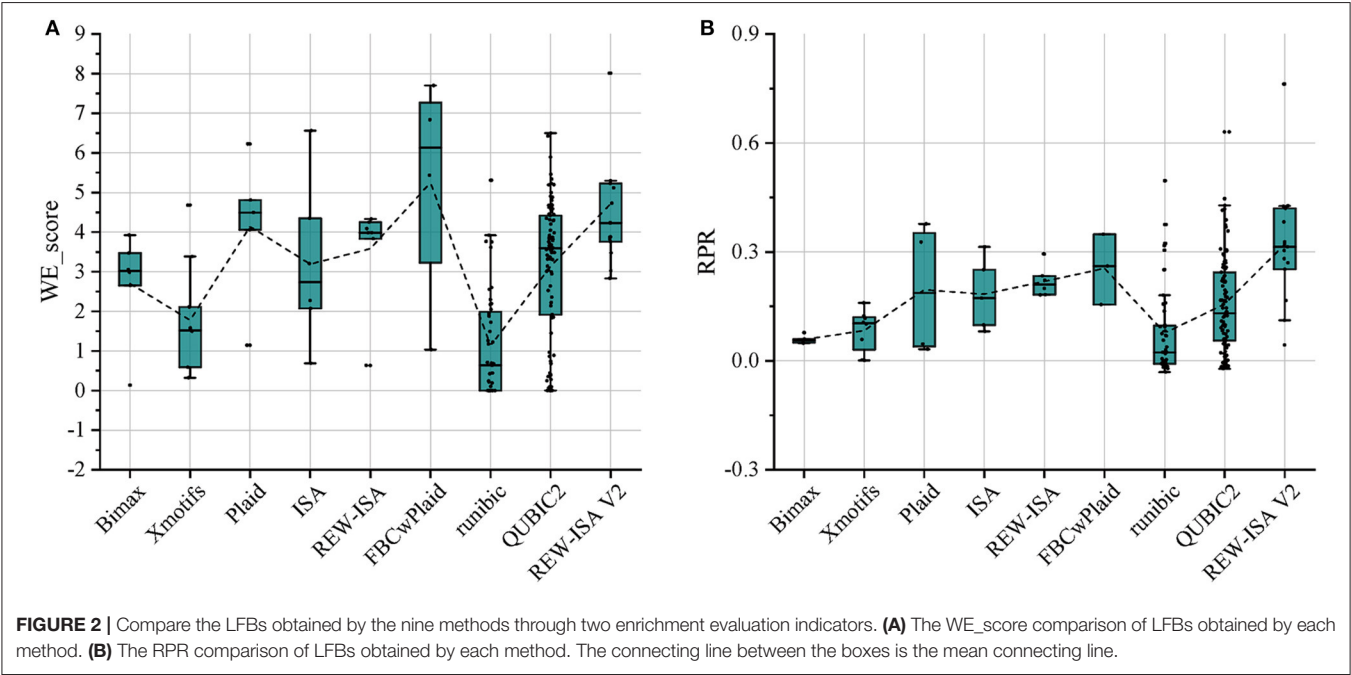
et al., 2013; Batista et al., 2014; Schwartz et al., 2014; Wang et al., 2014; Barbieri et al., 2017; Li et al., 2017; Pendleton et al., 2017) to mine potential LFBs, most of which can be retrieved from the MeT-DB V2.0 database (Liu et al., 2018). **Table 1** summarizes the MeRIP-seq real data set used in this project. Then, calculate the corresponding P and W through (1) and (2), and perform REW-ISA V2. Within the range of T_R being 0.1–2 with step size 0.1, and T_C being 0.1–2 with step size 0.1, T_R and T_C are optimized through the enrichment constraint framework. The experiments were repeated ten times for each parameter setting. Although optimizing T_R and T_C based on the gathered biological significance may produce biased results. However, this process provides guidance for the selection of T_R and T_C . Finally, under the optimal T_R of 0.4 and the optimal T_C of 0.7, a total of fifteen LFBs are obtained. The number of m⁶A sites, the number of genes corresponding to m⁶A sites and the number of conditions contained in these LFBs are shown in **Supplementary Table 1**.

For the above-mentioned real data set, Bimax (Prelić et al., 2006), Xmotifs (Murali and Kasif, 2003), Plaid (Lazzeroni and Owen, 2002), ISA (Bergmann et al., 2003), REW-ISA (Zhang et al., 2020), FBCwPlaid (Chen et al., 2021), runibic (Orzechowski et al., 2018a), and QUBIC2 (Xie et al., 2020) were all included for comparison with REW-ISA V2. To make the LFBs obtained by the above methods have significant biological significance, the parameters of these methods have been appropriately adjusted. For each LFB obtained by each method, the two enrichment indicators, WE_score and RPR, were both calculated for evaluation. The comparison results are shown in **Figures 2A,B**, respectively. As can be seen from **Figure 2A**, the average WE_score of the LFBs obtained by the REW-ISA V2 algorithm is higher than that of ISA and REW-ISA, which indicates that the fusion of homologous information is effective for mining LFBs. Although the average WE_score of LFBs obtained by REW-ISA V2 is lower than that of the FBCwPlaid algorithm, there are significant differences in RPR between the two methods. After further analysis of the LFBs, we found that this was caused by the size of LFBs found by REW-ISA V2 was smaller than that found by the FBCwPlaid algorithm. In other words, the LFBs found by REW-ISA V2 had higher enrichment scores with fewer corresponding genes. Besides, we can find that runibic and QUBIC2 do not perform well in the task of m⁶A hypermethylation pattern recognition. It may be due to the following two points. On the one hand, the two algorithms mainly identify the trend-preserving biclusters, which is different from the hyper-methylation bicluster. On the other hand, the LFBs obtained are generally small. This also reflects the need of developing biclustering methods for epi-transcriptome data. In a word, the average RPR of LFBs inferred by REW-ISA V2 is significantly higher than that of other biclustering algorithms, which means that the LFBs obtained by REW-ISA V2 may be more biologically significant.

To further explore the biological significance of the obtained LFBs, we selected four LFBs with more sites from the fifteen LFBs. As can be seen from **Supplementary Table 1**, for the four selected LFBs, they cover 1,256, 1,619, 824, and 1,148 genes, respectively. An important feature of any biclustering is the

TABLE 1 | MeRIP-seq datasets used in the study.

ID	GEO accession	Cell line	Treatment	Source
1–4	SRR456542–SRR456549, SRR456551–SRR456557	HepG2	UV, HGF, IFN, UT	Dominissini et al., 2012
5–6	SRR903368–SRR903379	U2OS	CTL, DAA	Fustin et al., 2013
7–10	SRR847358–SRR847377	HeLa	Ctrl, METTL3-, METTL14-, WTAP-	Liu et al., 2014
11–12	SRR1182582–SRR1182590	ES/NPC	hNPC, hESC	Schwartz et al., 2014
13–18	SRR1182591–SRR1182596, SRR494613–SRR494618, SRR5080301–SRR50312	HEK293	Ctrl, WTAP-, METTL3-, METTL16-	Meyer et al., 2012; Schwartz et al., 2014; Pendleton et al., 2017
19–21	SRR1182597–SRR1182602	OKMS	D0, D5_WITH_DOX, D5_WO_DOX	Schwartz et al., 2014
22–26	SRR1182603–SRR1182630	A549	Ctrl, METTL3-, METTL14-, WTAP-, KIAA1429-	Schwartz et al., 2014
27–28	SRR3066062–SRR3066069	AML	Ctrl, FTO+	Li et al., 2017
29–30	SRR5239086–SRR5239109	AML2	Ctrl, METTL3-	Barbieri et al., 2017
31–32	SRR1035213–SRR1035224	ESC	T0, T48	Batista et al., 2014



identified subsets of conditions, so the conditions contained in the four selected LFBs are explored in detail, as shown in **Supplementary Table 2**. The methylation level heatmaps of the four selected LFBs are shown in **Figure 3**. For the KEGG pathway analysis, six KEGG pathways known to be regulated by RNA methylation were selected (Dominissini et al., 2012; Xiang et al., 2017), such as apoptosis, DNA repair, fatty acid metabolism, etc. Then, Fisher's exact test was used to verify whether each LFB was significantly enriched in some specific pathways. The output *p*-value shows the correlation between

the four LFBs obtained and six biological pathways, as well as the importance of multiple hypothesis correction. We could see from **Supplementary Table 3** that the four selected LFBs are significantly enriched in the ultraviolet (UV) response up. Although the enrichment degree of LFB2 is lower than that of the other three LFBs in the UV response up, its enrichment in the apoptosis is significantly higher than that of the other three LFBs, indicating that LFB2 may further affect apoptosis through some other m⁶A-related pathways. Besides, LFB1, LFB3, and LFB4 are also significantly enriched in

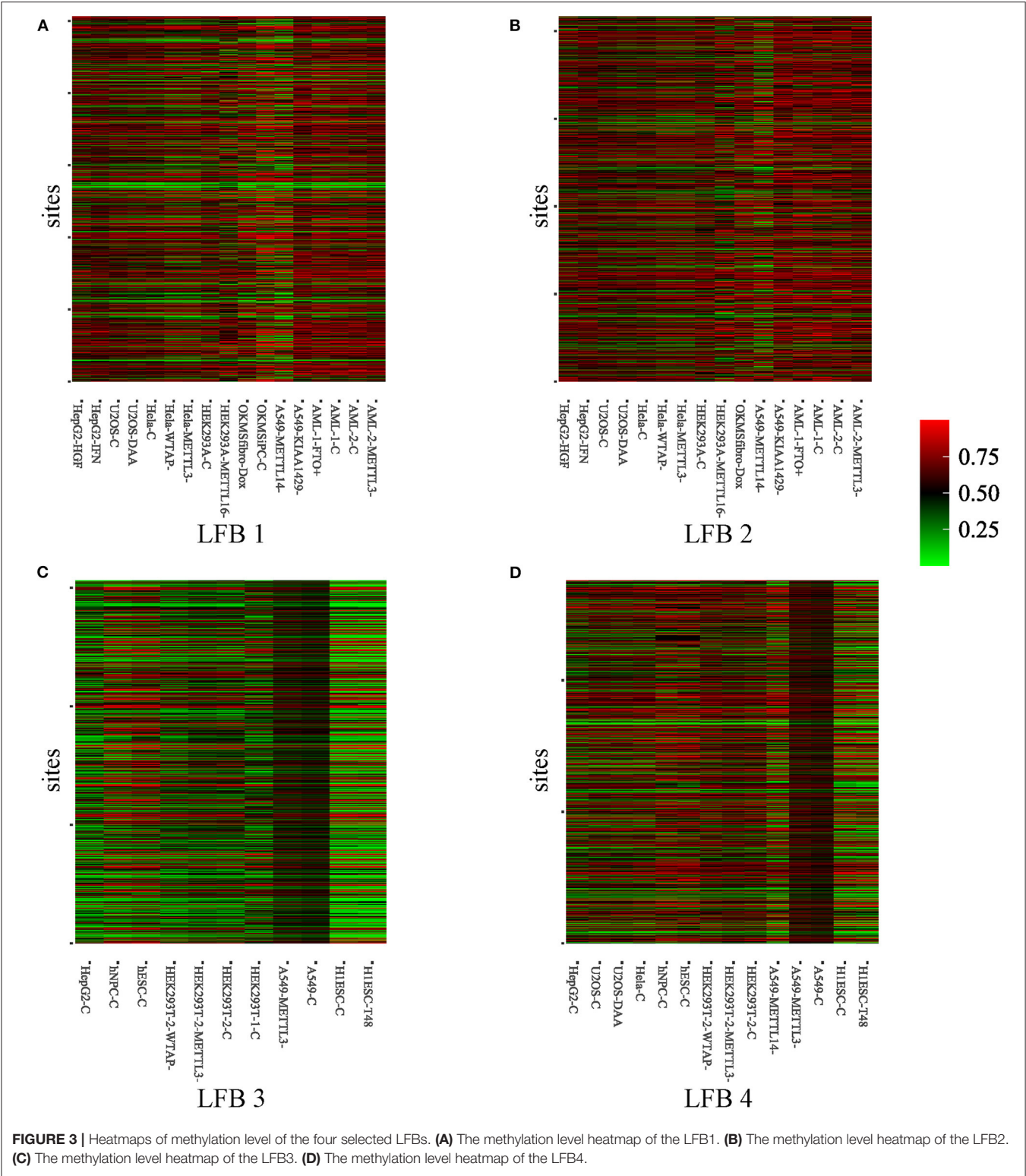


FIGURE 3 | Heatmaps of methylation level of the four selected LFBs. **(A)** The methylation level heatmap of the LFB1. **(B)** The methylation level heatmap of the LFB2. **(C)** The methylation level heatmap of the LFB3. **(D)** The methylation level heatmap of the LFB4.

DNA repair, which may be related to DNA damage caused by ultraviolet radiation. Since m⁶A has been proved to be related to stem cell differentiation and cancer progression (Batista et al., 2014), there is a reasonable explanation for enriching LFB1 and LFB3 in fatty acid metabolism. As the main components of neutral fat, phospholipids and glycolipids, fatty acids can meet various body needs and regulate metabolism, growth and development (Azain, 2004). The p53 pathway enriched

in LFB4 indicates that LFB4 may be related to stress signal, regulation of intracellular homeostasis, chromosome segregation, and cell division (Harris and Levine, 2005). Through the above analysis, it is not difficult to see that the LFBs obtained by REW-ISA V2 have more significant biological significance than the randomly selected LFB. Therefore, an in-depth analysis of the LFBs obtained may help reveal the specific regulatory mechanism of m⁶A.

To check whether the detected LFBs have biological significance, we further conducted the enzymes substrate specificity experiments on the four selected LFBs. Since LFB covers hyper-methylated sites and conditions, the sites and conditions involved in each LFB are more likely to be the target sites of m⁶A methyltransferases. Therefore, we studied the association between each selected LFB and four m⁶A methyltransferases, including METTL3, METTL14, WTAP as well as KIAA1429. For this purpose, 38,845 METTL3 targeted gene sites, 19,099 METTL14 targeted gene sites, 35,144 WTAP targeted gene sites, and 1,784 KIAA1429 targeted gene sites included in the real data were first identified by TREW tool (Liu et al., 2018). After REW-ISA V2, we summarized the distribution of target RNA methylation sites involved in each LFB (**Supplementary Table 4**). Then, the association between the sites in each selected LFB and m⁶A methyltransferases target sites was further evaluated by Fisher's exact test. The experimental enrichment results are shown in **Supplementary Table 5**, where *p*-value indicates the significance of association between sites and methyltransferase target sites. The results showed that the sites contained in the four selected LFBs were significantly enriched in the target sites of the four methyltransferases. This means that under specific conditions, the LFBs obtained by REW-ISA V2 were indeed the collaboratively hyper-methylated sites, which will help biologists to further study the specific regulation mechanism of m⁶A. The detailed analysis process and results can be obtained in the **Supplementary Materials**.

DISCUSSION

Although more and more studies have shown that the modification of m⁶A in RNA is related to many important biological functions, the specific regulatory mechanism of m⁶A is still unclear. To quickly and effectively predict potential functional m⁶A sites from the epi-transcriptome data, it is important to develop some computational algorithms, which will help us have a more comprehensive understanding of m⁶A-related life processes. Based on REW-ISA, in this article, we developed REW-ISA V2 to better reveal the potential local co-methylation patterns across subsets of conditions. REW-ISA V2 was implemented on the real MeRIP-seq data, and a total of 15

LFBs were obtained. Further comparison and analysis show that, compared with other biclustering algorithms, the LFBs obtained by REW-ISA V2 has more significant biological significance.

REW-ISA V2 could obtain reliable biclustering patterns because of the use of homologous information. More specifically, the sites' methylation levels corresponding to the same gene will show a similar trend with a high probability. Similarly, conditions derived from the same cell line will have similar trends in all sites. Therefore, the rational use of homologous information will help to better mine local co-methylation patterns. Of course, REW-ISA V2 still has some deficiencies that need to be improved in the future. First of all, REW-ISA V2 uses simple multiplication to fuse homologous information, which inevitably introduces noise at the same time. Secondly, because the database on which GO analysis depends is incomplete, the enrichment constraint framework designed is prone to human error. Finally, the enrichment constraint framework built into REW-ISA V2 usually takes a long time. In the future, we will use BSig (Henriques and Madeira, 2018) to better evaluate the obtained LFBs and develop a new computational model to overcome these limitations.

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

LZ and SC built the architecture for REW-ISA V2, designed and implemented the experiments, analyzed the result, and wrote the paper. JM conducted the experiments, analyzed the result, and revised the paper. ZL and HL supervised the project, analyzed the result, and revised the paper. All authors read, critically revised, and approved the final manuscript.

FUNDING

This work has been supported by Fundamental Research Funds for the Central Universities (Grant No. 2019ZDPY15 to LZ). The funding body did not play any roles in the design of the study and collection, analysis, and interpretation of data and in writing the manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2021.654820/full#supplementary-material>

REFERENCES

- Azain, M. (2004). Role of fatty acids in adipocyte growth and development. *J. Anim. Sci.* 82, 916–924. doi: 10.1093/ansci/82.3.916

- Barbieri, I., Tzelepis, K., Pandolfini, L., Shi, J., Millán-Zambrano, G., Robson, S. C., et al. (2017). Promoter-bound METTL3 maintains myeloid leukaemia by m⁶A-dependent translation control. *Nature* 552, 126–131. doi: 10.1038/nature24678

- Batista, P. J., Molinie, B., Wang, J., Qu, K., Zhang, J., Li, L., et al. (2014). m6A RNA modification controls cell fate transition in mammalian embryonic stem cells. *Cell Stem Cell* 15, 707–719. doi: 10.1016/j.stem.2014.09.019
- Bergmann, S., Ihmels, J., and Barkai, N. (2003). Iterative signature algorithm for the analysis of large-scale gene expression data. *Phys. Rev. E* 67:031902. doi: 10.1103/PhysRevE.67.031902
- Chai, R., Wu, F., Wang, Q., Zhang, S., Zhang, K., Liu, Y., et al. (2019). m6A RNA methylation regulators contribute to malignant progression and have clinical prognostic impact in gliomas. *Aging* 11, 1204–1225. doi: 10.18632/aging.101829
- Chen, K., Wei, Z., Zhang, Q., Wu, X., Rong, R., Lu, Z., et al. (2019). WHISTLE: a high-accuracy map of the human N6-methyladenosine (m6A) epitranscriptome predicted using a machine learning approach. *Nucleic Acids Res.* 47:e41. doi: 10.1093/nar/gkx074
- Chen, S., Zhang, L., Lu, L., Meng, J., and Liu, H. (2021). FBCwPlaid: a functional bi-clustering analysis of epi-transcriptome profiling data via a weighted plaid model. *IEEE/ACM Trans. Comput. Biol. Bioinform.* doi: 10.1109/TCBB.2021.3049366
- Deng, X., Su, R., Feng, X., Wei, M., and Chen, J. (2018). Role of N6-methyladenosine modification in cancer. *Curr Opin. Genet. Dev.* 48, 1–7. doi: 10.1016/j.gde.2017.10.005
- Dominissini, D., Moshitch-Moshkovitz, S., Schwartz, S., Salmon-Divon, M., Ungar, L., Osenberg, S., et al. (2012). Topology of the human and mouse m6A RNA methylomes revealed by m6A-seq. *Nature* 485, 201–206. doi: 10.1038/nature11112
- Frye, M., Harada, B. T., Behm, M., and He, C. (2018). RNA modifications modulate gene expression during development. *Science* 361, 1346–1349. doi: 10.1126/science.aau1646
- Fu, Y., Dominissini, D., Rechavi, G., and He, C. (2014). Gene expression regulation mediated through reversible m6A RNA methylation. *Nat. Rev. Genet.* 15, 293–306. doi: 10.1038/nrg3724
- Fustin, J.-M., Doi, M., Yamaguchi, Y., Hida, H., Nishimura, S., Yoshida, M., et al. (2013). RNA-methylation-dependent RNA processing controls the speed of the circadian clock. *Cell* 155, 793–806. doi: 10.1016/j.cell.2013.10.026
- Harris, S. L., and Levine, A. J. (2005). The p53 pathway: positive and negative feedback loops. *Oncogene* 24, 2899–2908. doi: 10.1038/sj.onc.1208615
- Henriques, R., Antunes, C., and Madeira, S. C. (2015). A structured view on pattern mining-based biclustering. *Pattern Recognit.* 48, 3941–3958. doi: 10.1016/j.patcog.2015.06.018
- Henriques, R., Ferreira, F. L., and Madeira, S. C. (2017). BicPAMS: software for biological data analysis with pattern-based biclustering. *BMC Bioinformatics* 18:82. doi: 10.1186/s12859-017-1493-3
- Henriques, R., and Madeira, S. C. (2018). BSig: evaluating the statistical significance of biclustering solutions. *Data Min. Knowl. Discov.* 32, 124–161. doi: 10.1007/s10618-017-0521-2
- Jia, G., Fu, Y., Zhao, X., Dai, Q., Zheng, G., Yang, Y., et al. (2011). N6-methyladenosine in nuclear RNA is a major substrate of the obesity-associated FTO. *Nat. Chem. Biol.* 7, 885–887. doi: 10.1038/nchembio.687
- Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., and Salzberg, S. L. (2013). TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol.* 14:R36. doi: 10.1186/gb-2013-14-4-r36
- Lachén-Montes, M., González-Morales, A., De Morentin, X. M., Pérez-Vallderama, E., Ausín, K., Zelaya, M. V., et al. (2016). An early dysregulation of FAK and MEK/ERK signaling pathways precedes the β -amyloid deposition in the olfactory bulb of APP/PS1 mouse model of Alzheimer's disease. *J. Proteom.* 148, 149–158. doi: 10.1016/j.jprot.2016.07.032
- Lazzeroni, L., and Owen, A. (2002). Plaid models for gene expression data. *Stat. Sinica* 12, 61–86. doi: 10.1109/ITW.2002.1115477
- Li, L., Guo, Y., Wu, W., Shi, Y., Cheng, J., and Tao, S. (2012). A comparison and evaluation of five biclustering algorithms by quantifying goodness of biclusters for gene expression data. *BioData Min.* 5:8. doi: 10.1186/1756-0381-5-8
- Li, Z., Weng, H., Su, R., Weng, X., Zuo, Z., Li, C., et al. (2017). FTO plays an oncogenic role in acute myeloid leukemia as a N6-methyladenosine RNA demethylase. *Cancer Cell* 31, 127–141. doi: 10.1016/j.ccell.2016.11.017
- Lin, S., Choe, J., Du, P., Triboulet, R., and Gregory, R. I. (2016). The m6A methyltransferase METTL3 promotes translation in human cancer cells. *Mol. Cell* 62, 335–345. doi: 10.1016/j.molcel.2016.03.021
- Liu, H., Wang, H., Wei, Z., Zhang, S., Hua, G., Zhang, S., et al. (2018). MeT-DB V2.0: elucidating context-specific functions of N6-methyl-adenosine methyltranscriptome. *Nucleic Acids Res.* 46, D281–D287. doi: 10.1093/nar/gkx1080
- Liu, J., Yue, Y., Han, D., Wang, X., Fu, Y., Zhang, L., et al. (2014). A METTL3–METTL14 complex mediates mammalian nuclear RNA N6-adenosine methylation. *Nat. Chem. Biol.* 10, 93–95. doi: 10.1038/nchembio.1432
- Liu, L., Zhang, S., Zhang, Y., Liu, H., Zhang, L., Chen, R., et al. (2015). Decomposition of RNA methylome reveals co-methylation patterns induced by latent enzymatic regulators of the epitranscriptome. *Mol. Biosyst.* 11, 262–274. doi: 10.1039/C4MB00604F
- Liu, X., Li, D., Liu, J., Su, Z., and Li, G. (2020). RecBic: a fast and accurate algorithm recognizing trend-preserving biclusters. *Bioinformatics* 36, 5054–5060. doi: 10.1093/bioinformatics/btaa630
- Meng, J., Gao, S., and Huang, Y. (2009). Enrichment constrained time-dependent clustering analysis for finding meaningful temporal transcription modules. *Bioinformatics* 25, 1521–1527. doi: 10.1093/bioinformatics/btp235
- Meng, J., Lu, Z., Liu, H., Zhang, L., Zhang, S., Chen, Y., et al. (2014). A protocol for RNA methylation differential analysis with MeRIP-Seq data and exomePeak R/Bioconductor package. *Methods* 69, 274–281. doi: 10.1016/j.ymeth.2014.06.008
- Meyer, K. D., Patil, D. P., Zhou, J., Zinoviev, A., Skabkin, M. A., Elemento, O., et al. (2015). 5' UTR m6A promotes cap-independent translation. *Cell* 163, 999–1010. doi: 10.1016/j.cell.2015.10.012
- Meyer, K. D., Saletore, Y., Zumbo, P., Elemento, O., Mason, C. E., and Jaffrey, S. R. (2012). Comprehensive analysis of mRNA methylation reveals enrichment in 3' UTRs and near stop codons. *Cell* 149, 1635–1646. doi: 10.1016/j.cell.2012.05.003
- Murali, T., and Kasif, S. (2003). Extracting conserved gene expression motifs from gene expression data. *Pac. Symp. Biocomput.* 8, 77–88. doi: 10.1142/9789812776303_0008
- Orzechowski, P., Pańszczyk, A., Huang, X., and Moore, J. H. (2018a). runibic: a Bioconductor package for parallel row-based biclustering of gene expression data. *Bioinformatics* 34, 4302–4304. doi: 10.1093/bioinformatics/bty512
- Orzechowski, P., Sipper, M., Huang, X., and Moore, J. H. (2018b). EBIC: an evolutionary-based parallel biclustering algorithm for pattern discovery. *Bioinformatics* 34, 3719–3726. doi: 10.1093/bioinformatics/bty401
- Pendleton, K. E., Chen, B., Liu, K., Hunter, O. V., Xie, Y., Tu, B. P., et al. (2017). The U6 snRNA m6A methyltransferase METTL16 regulates SAM synthetase intron retention. *Cell* 169, 824–835.e14. doi: 10.1016/j.cell.2017.05.003
- Ping, X., Sun, B., Wang, L., Xiao, W., Yang, X., Wang, W., et al. (2014). Mammalian WTAP is a regulatory subunit of the RNA N6-methyladenosine methyltransferase. *Cell Res.* 24, 177–189. doi: 10.1038/cr.2014.3
- Prelić, A., Bleuler, S., Zimmermann, P., Wille, A., Bühlmann, P., Gruissem, W., et al. (2006). A systematic comparison and evaluation of biclustering methods for gene expression data. *Bioinformatics* 22, 1122–1129. doi: 10.1093/bioinformatics/btl060
- Schwartz, S., Mumbach, M. R., Jovanovic, M., Wang, T., Maciag, K., Bushkin, G. G., et al. (2014). Perturbation of m6A writers reveals two distinct classes of mRNA methylation at internal and 5' sites. *Cell Rep.* 8, 284–296. doi: 10.1016/j.celrep.2014.05.048
- Wang, X., Lu, Z., Gomez, A., Hon, G. C., Yue, Y., Han, D., et al. (2014). N6-methyladenosine-dependent regulation of messenger RNA stability. *Nature* 505, 117–120. doi: 10.1038/nature12730
- Wang, Z., Li, G., Robinson, R. W., and Huang, X. (2016). UniBic: Sequential row-based biclustering algorithm for analysis of gene expression data. *Sci. Rep.* 6:23466. doi: 10.1038/srep23466
- Wu, X., Wei, Z., Chen, K., Zhang, Q., Su, J., Liu, H., et al. (2019). m6Acomet: large-scale functional prediction of individual m6A RNA methylation sites from an RNA co-methylation network. *BMC Bioinformatics* 20:223. doi: 10.1186/s12859-019-2840-3
- Xiang, Y., Laurent, B., Hsu, C.-H., Nachtergaele, S., Lu, Z., Sheng, W., et al. (2017). RNA m6A methylation regulates the ultraviolet-induced DNA damage response. *Nature* 543, 573–576. doi: 10.1038/nature21671
- Xie, J., Ma, A., Zhang, Y., Liu, B., Cao, S., Wang, C., et al. (2020). QUBIC2: a novel and robust biclustering algorithm for analyses and interpretation of large-scale RNA-Seq data. *Bioinformatics* 36, 1143–1149. doi: 10.1093/bioinformatics/btz692

- Yu, G., Wang, L., Han, Y., and He, Q. (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. *Omic* 16, 284–287. doi: 10.1089/omi.2011.0118
- Zhang, C., Chen, Y., Sun, B., Wang, L., Yang, Y., Ma, D., et al. (2017). m6A modulates haematopoietic stem and progenitor cell specification. *Nature* 549, 273–276. doi: 10.1038/nature23883
- Zhang, C., Fu, J., and Zhou, Y. (2019a). A review in research progress concerning m6A methylation and immunoregulation. *Front. Immunol.* 10:922. doi: 10.3389/fimmu.2019.00922
- Zhang, L., Chen, S., Zhu, J., Meng, J., and Liu, H. (2020). REW-ISA: unveiling local functional blocks in epi-transcriptome profiling data via an RNA expression-weighted iterative signature algorithm. *BMC Bioinformatics* 21:447. doi: 10.1186/s12859-020-03787-w
- Zhang, L., He, Y., Wang, H., Liu, H., Huang, Y., Wang, X., et al. (2019b). Clustering count-based RNA methylation data using a nonparametric generative model. *Curr. Bioinform.* 14, 11–23. doi: 10.2174/1574893613666180601080008
- 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, Chen, Ma, Liu and Liu. 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.



N6-Methyladenosine RNA Modification in Inflammation: Roles, Mechanisms, and Applications

Jiahui Luo¹, Tao Xu^{2*} and Kai Sun^{3*}

¹ The Center for Biomedical Research, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China, ² Department of Rehabilitation, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China, ³ Department of Orthopedics, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China

OPEN ACCESS

Edited by:

Giovanni Nigita,
The Ohio State University,
United States

Reviewed by:

Gioacchino Paolo Marceca,
University of Catania, Italy
Abhijit Shukla,
Memorial Sloan Kettering Cancer
Center, United States
Natalia Pinello,
Centenary Institute of Cancer
Medicine and Cell Biology, Australia

*Correspondence:

Tao Xu
xutao0101@yeah.net
Kai Sun
1085844308@qq.com

Specialty section:

This article was submitted to
Epigenomics and Epigenetics,
a section of the journal
Frontiers in Cell and Developmental
Biology

Received: 22 February 2021

Accepted: 10 May 2021

Published: 04 June 2021

Citation:

Luo J, Xu T and Sun K (2021)
N6-Methyladenosine RNA
Modification in Inflammation: Roles,
Mechanisms, and Applications.
Front. Cell Dev. Biol. 9:670711.
doi: 10.3389/fcell.2021.670711

N6-methyladenosine (m6A) is the most prevalent internal mRNA modification. m6A can be installed by the methyltransferase complex and removed by demethylases, which are involved in regulating post-transcriptional expression of target genes. RNA methylation is linked to various inflammatory states, including autoimmunity, infection, metabolic disease, cancer, neurodegenerative diseases, heart diseases, and bone diseases. However, systematic knowledge of the relationship between m6A modification and inflammation in human diseases remains unclear. In this review, we will discuss the association between m6A modification and inflammatory response in diseases, especially the role, mechanisms, and potential clinical application of m6A as a biomarker and therapeutic target for inflammatory diseases.

Keywords: N6-methyladenosine, RNA modification, epigenetics, inflammation, inflammatory disease

INTRODUCTION

In recent years, epitranscriptomic modifications have been recognized as essential regulators of various physiological processes and disease progression. RNA modification is an example of dynamic epigenetic regulation. Early RNA modification studies have revealed important functions of RNA modification in translation and splicing. These studies focused on abundant non-coding RNAs, such as rRNA, tRNA, and small nuclear RNA (Jung and Goldman, 2018). The tRNA modification N1-methyladenosine (m1A) regulates the associations between tRNA and polysomes to stabilize tRNA tertiary structure and modify translation (Li et al., 2016). Pseudouridine (Ψ) in snRNA participates in mRNA splicing, whereas Ψ in rRNA ensures translational fidelity (Xiong et al., 2017).

N6-methyladenosine (m6A) is the most prevalent internal mRNA modification, which was first discovered in the early 1970s (Dubin and Taylor, 1975). m6A participates in almost all processes involving mRNA metabolism, including RNA transcription, translation, and degradation (Roundtree et al., 2017a; Nachtergaele and He, 2018). Similar to DNA methylation, RNA m6A modification is a dynamic and reversible process (Zhao et al., 2020). m6A is installed by the methyltransferase complex (MTC) and removed by demethylases, which regulate the post-transcriptional expression of target genes (Batista, 2017; Dai et al., 2018). In molecular processing, m6A is involved in many steps of RNA metabolism, including pre-mRNA splicing, mRNA

translation, nuclear export, mRNA degradation, and non-coding RNA biogenesis (Liu et al., 2017; Shi et al., 2017; Liu and Gregory, 2019).

Inflammation is a complex physiological reaction to microorganisms, autoimmunity, allergies, metabolism, and physical damage, all of which produce different types of inflammatory responses (Hawiger and Zienkiewicz, 2019). In acute inflammation, the initial response of the body to a stimulus is achieved by increasing white blood cell migration and plasma leakage from the blood to the site of injury (Bayarsaihan, 2011; Varela et al., 2018). Chronic inflammation progresses slowly, occurs for a longer duration, and can cause many diseases, including periodontal disease and diabetes (Dunning, 2009). Inflammatory response requires a complex regulatory network to achieve signal- and gene-specific levels of function (Medzhitov and Horng, 2009) for activating specific genes for antimicrobial defense, immune response, and tissue repair and remodeling (Medzhitov, 2008). Emerging evidence suggests that epigenetic modifications are involved in inflammatory response. Macrophages play critical roles in various inflammatory diseases, including obesity and arthritis. Chromatin modifications have been reported to participate in the regulation of macrophage phenotype (Nemat et al., 2009). In addition, DNA methylation and covalent histone modifications of transcription factors, including NF- κ B and the STAT families, have been found to modulate inflammatory genes (Medzhitov and Horng, 2009). However, the role of RNA modification in regulating inflammation and anti-inflammatory gene expression has only been recently elucidated. In this review, we will discuss the association between m6A methylation and inflammatory response in diseases, especially its role, mechanism, and potential clinical application in inflammation-related diseases.

RNA m6A Methylation

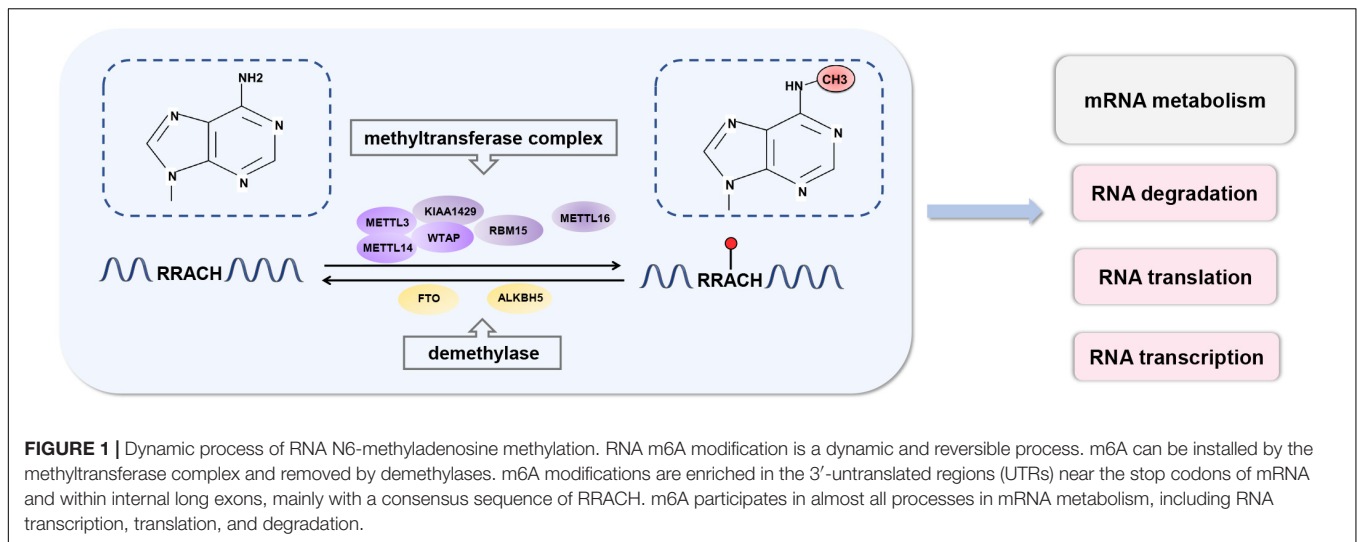
In mammals, approximately 0.1–0.4% of adenosine in mRNA is modified by m6A, and each transcript has an average of 2–3 m6A modification sites (Desrosiers et al., 1974; Wei et al., 1975; Fu et al., 2014). m6A modifications are enriched in the 3'-untranslated regions (UTRs) near the stop codons of mRNA and within internal long exons, mainly with a consensus sequence of RRACH (R = G or A; H = A, C, or U) (Kane and Beemon, 1985; Dominissini et al., 2012). There are three types of regulators involved in m6A modifications and are referred to as “writers,” “erasers,” and “readers” (Figure 1).

“Writers” traditionally refer to a highly conserved mRNA methyltransferase complex that consists of methyltransferase-like 3 (METTL3), methyltransferase-like 14 (METTL14), and Wilms tumor suppressor-1-associated protein (WTAP) (Ping et al., 2014; Schwartz et al., 2014). Both METTL3 and METTL14 contain a S-adenosylmethionine-binding motif (Wang et al., 2016). METTL3 acts as a major catalytic component that modulates m6A modification (Barbieri et al., 2017). METTL14 is a pseudomethyltransferase that serves as an accessory component of METTL3 and participates in substrate recognition (Weng et al., 2018). They form a stable heterodimer and co-localize in nuclear speckles in a 1:1 ratio. WTAP is the main regulatory component and interacts directly with METTL3. It promotes

methylation and ensures the nuclear location of the core writer complex (Liu et al., 2014; Schöller et al., 2018). The other known writers are methyltransferase-like 16 (METTL16) (Pendleton et al., 2017; Warda et al., 2017), KIAA1429 (Schwartz et al., 2014), and RBM15 (Moindrot et al., 2015). METTL16 methylates long non-coding RNA and U6 small nuclear RNA and regulates S-adenosylmethionine homeostasis (Pendleton et al., 2017; Warda et al., 2017). KIAA1429 and RBM15 are also components of methyltransferase complexes. KIAA1429 is required for full methylation in mammals (Schwartz et al., 2014). RBM15 mediates the formation of m6A in X-inactive specific transcripts and cellular mRNAs (Moindrot et al., 2015). However, the detailed functions of these components are still poorly understood, and there are likely other mechanisms that remain to be investigated.

Reversible modifications of m6A can be removed by “erasers,” such as fat mass and obesity-related protein (FTO) and ALKB homolog 5 protein (ALKBH5) (Jia et al., 2011; Zheng et al., 2013). These are ferrous iron- and α -ketoglutarate-dependent demethylases, which can oxidatively remove m6A methylated groups from mRNA (Chen and Wong, 2020). Both FTO and ALKBH5 belong to the ALKB family of dioxygenases (Fedele et al., 2015). FTO sequentially oxidizes m6A to N6-hydroxymethyladenosine and N6-formyladenosine, which can be further hydrolyzed to adenine (Wei et al., 2018). ALKBH5 directly catalyzes the removal of m6A modifications. Recently, another m6A demethylase, ALKBH3, that preferentially act on m6A in tRNA rather than mRNA or rRNA has been reported (Ueda et al., 2017). Given that m6A is a dynamic modification in response to stimuli, demethylases ensure the equilibrium of m6A modifications in the transcriptome.

The last type of m6A regulatory proteins are “readers,” which can recognize and bind to the m6A modification site in RNA and play an important role in different biological functions (Figure 2). The first m6A “reader” proteins identified were YTHDF1, YTHDF2, YTHDF3, YTHDC1, and YTHDC2, which contain a conserved YTH domain (YT521-B homology) (Liao et al., 2018). The YTH domain functions as a module that recognizes modifications and directly binds to m6A-modified RNA in the RRM6ACH consensus sequence (Roundtree et al., 2017a). YTHDF2, the first “reader” identified, accelerates mRNA degradation by binding m6A at the 3' UTR and localizing the targeted mRNA to processing bodies (Ries et al., 2019). In addition, YTHDF2 can induce mRNA degradation through the recruitment of the CCR4-NOT deadenylation machinery (Du et al., 2016). Moreover, YTHDF2 has been found to prevent m6A modification in the 5' UTR of the FTO protein by binding to the m6A site in the nucleus, thus promoting RNA translation in a cap-independent manner (Zhou et al., 2015). YTHDF1 binds to m6A sites around the stop codon and enhances mRNA translation by recruiting the eIF3 translation initiation complex (Wang et al., 2015; Chen and Wong, 2020). YTHDF3 has a fine-tuning effect on the RNA accessibility of YTHDF1 and YTHDF2 (Hu B. B. et al., 2019). YTHDF3 can interact with YTHDF1 to improve RNA translation efficiency and combine with YTHDF2 to promote RNA degradation (Li et al., 2017; Shi et al., 2017). Recent studies have revealed a new model



describing the functions of YTHDF proteins. YTHDF proteins function together to mediate mRNA degradation, and they show identical binding to all m6A sites in mRNA (Zaccara and Jaffrey, 2020). Each YTHDF paralog can compensate for the function of other YTHDF paralogs (Zaccara and Jaffrey, 2020). Another study confirmed the context-dependent functional compensation between YTHDF proteins (Lasman et al., 2020). YTHDC1 regulates RNA splicing and controls the nuclear export of its target RNA (Meyer and Jaffrey, 2017; Roundtree et al., 2017b). YTHDC2 promotes translation elongation by interacting with RNA helicases (Hsu et al., 2017). Three heterogeneous nuclear ribonucleoproteins (hnRNPs) are common “readers,” namely hnRNPC, hnRNPG, and hnRNPA2B1 (Alarcón Claudio et al., 2015; Liu et al., 2015; Meyer Kate et al., 2015). These proteins indirectly bind to transcripts through the “m6A switch” mechanism. This mechanism relies on the ability of m6A to thermodynamically destabilize into short double helices. In this way, the single-stranded hnRNP binding motif is exposed and provides access to RNA-binding proteins. hnRNPC and hnRNPG influence mRNA localization and alternative splicing (König et al., 2010; Liu et al., 2015, 2017). HnRNPA2B1 promotes the transcription of precursor miRNA by binding to m6A-containing primary miRNAs and interacting with the microRNA microprocessor complex (Alarcón Claudio et al., 2015). Recently, the IGFBP family IGFBP1-3 has been identified as a distinct family of m6A “readers” (Li et al., 2019; Müller et al., 2019; Wang S. et al., 2019). The binding motif of these proteins (UGGAC) overlaps with the consensus sequence of m6A. In addition, these proteins contain common RNA-binding domains that recognize m6A-containing transcripts. IGFbps recruit RNA stabilizers that protect m6A-containing mRNA from degradation, thereby promoting the expression of target transcripts (He et al., 2019).

Physiological Role of m6A RNA Modification

Under normal conditions, m6A modification interferes with RNA recognition (Karikó et al., 2005). Viruses can use m6A

modification to escape host immune recognition because of the presence of widely distributed m6A modifications in viral mRNA (Tang et al., 2021). In addition, mounting evidence has demonstrated m6A modifications in viral RNA (Durbin et al., 2016; Lu et al., 2020), host circRNAs (Chen et al., 2019), and normal endogenous ssRNA (Gao et al., 2020). m6A can serve as a key factor for the innate immune system to avoid abnormal immune recognition.

Innate immunity is activated immediately after the RNA recognition process, and it induces the production of cytokines, such as type I interferons (IFNs) (Winkler et al., 2019). m6A modification regulates the innate immune response by targeting IFN. IFNB1 mRNA can be modified by m6A within both the coding sequence and the 3' UTR (Rubio et al., 2018). IFN production triggered by dsDNA or human cytomegalovirus (HCMV) is mediated by m6A regulators. METTL14 depletion reduces virus reproduction and promotes IFNB1 mRNA accumulation, thus affecting pathways related to metabolic reprogramming, stress responses, and aging (Rubio et al., 2018). ALKBH5 depletion decreases IFNB1 mRNA production by regulating antiviral immune responses (Rubio et al., 2018). Moreover, it has been reported that IFNB mRNA is stabilized by repression of METTL3 or YTHDF2 after viral infection or inactivated virus stimulation (Winkler et al., 2019). Together, these findings reveal the important role of m6A as a negative regulator of IFN response, consequently facilitating viral propagation.

Role of RNA m6A in Inflammation: From Cells to Disease Progression

m6A modifications can cause changes in inflammation-related genes during inflammation. A systematic study of m6A modified cross-linking, substrate genes, and modified regulation illustrated the mechanism of action of m6A in inflammation (Figure 3 and Table 1). In the following sections, we will elucidate the interaction between m6A modification and inflammation in the

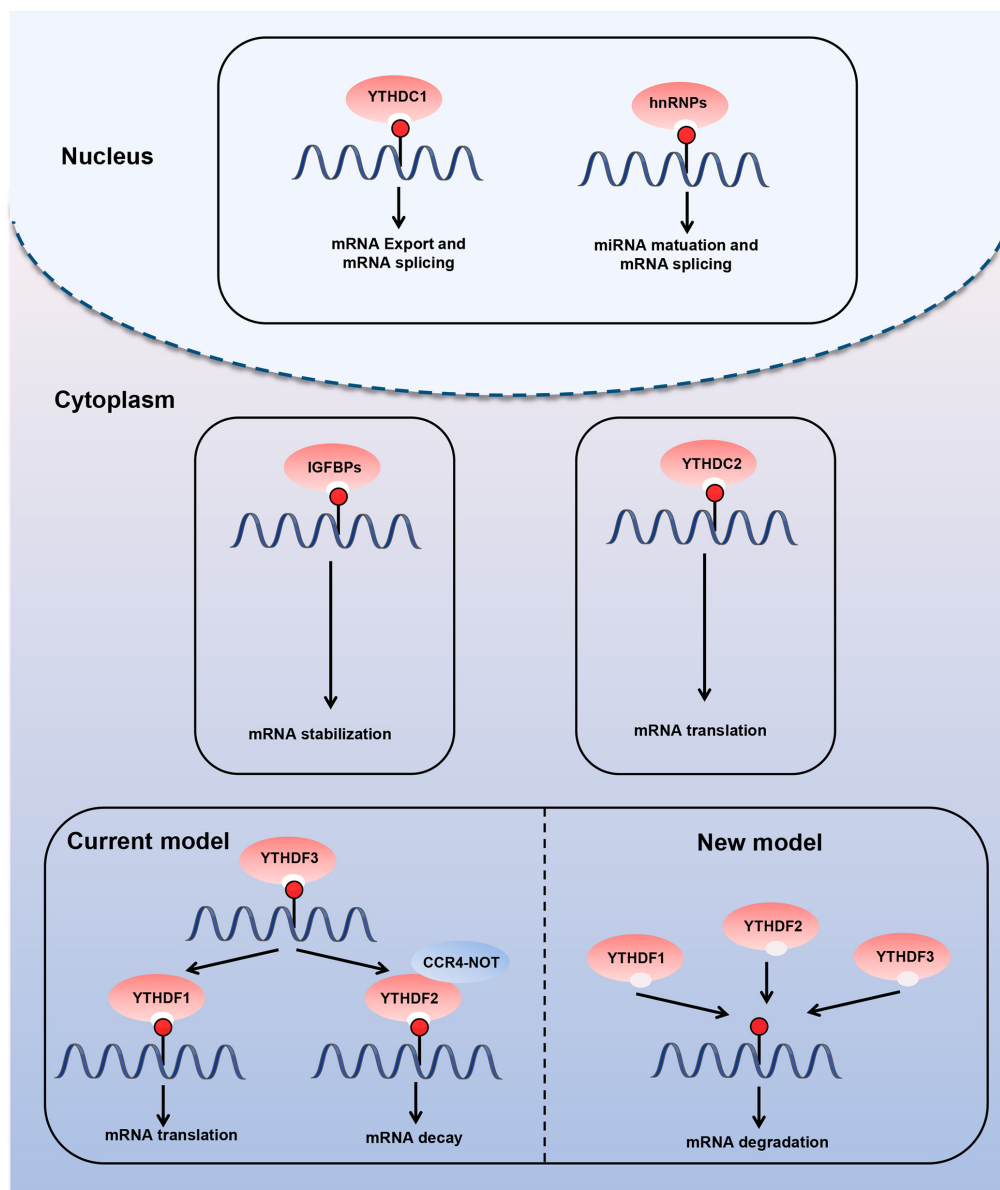


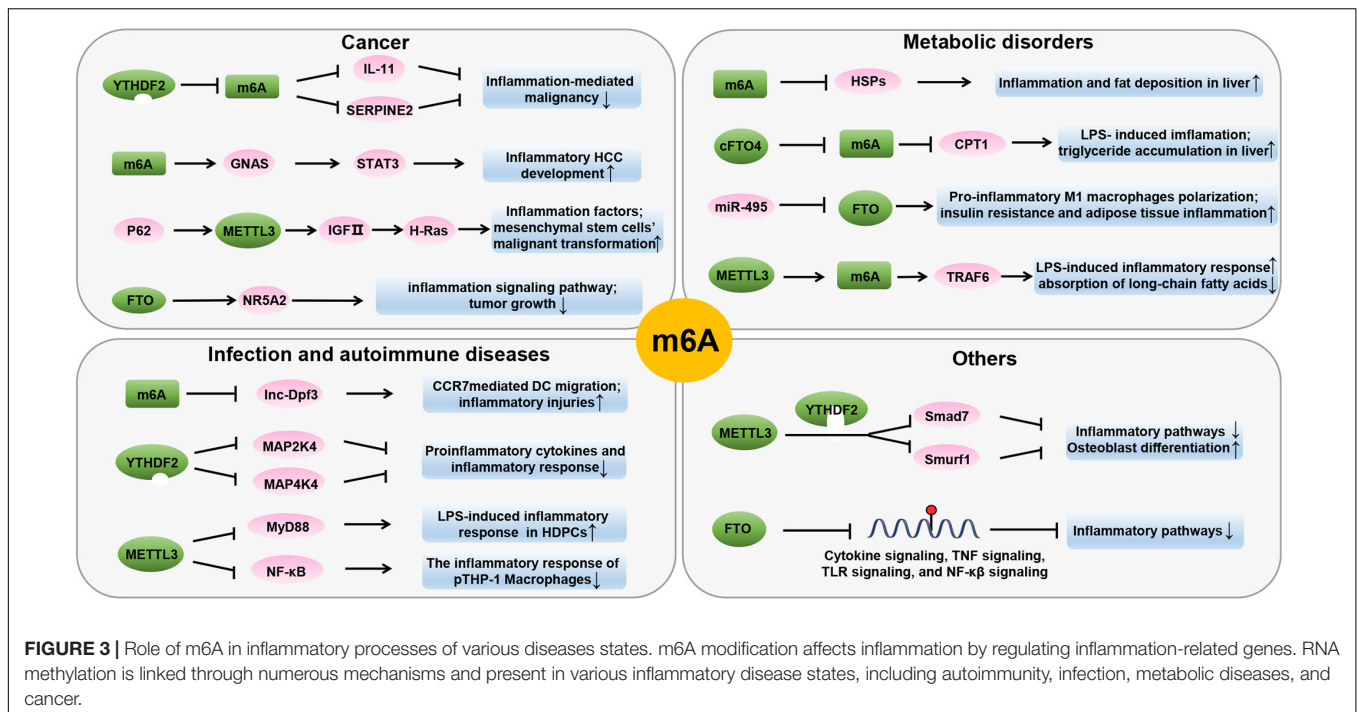
FIGURE 2 | Diverse biological consequences of m6A methylation by different m6A “readers.” m6A “readers” can recognize and bind to the m6A modification sites in RNA. They play an important role in different biological functions. The first m6A “readers” identified were YTHDF1, YTHDF2, YTHDF3, YTHDC1, and YTHDC2, which contain a conserved YTH domain (YT521-B homology). Moreover, three heterogeneous nuclear ribonucleoproteins (hnRNPs) are common “readers,” namely hnRNPC, hnRNPG, and hnRNPA2B1. m6A “readers” are involved in various steps of RNA metabolism, including pre-mRNA splicing, mRNA translation, nuclear export, and mRNA degradation.

pathogenesis of various diseases, including metabolic disorders, autoimmune diseases, and malignant diseases.

Metabolic Disorders

Under normal circumstances, the metabolic state of the human body is maintained under homeostasis. When normal metabolic processes are disrupted, metabolic disorders occur. m6A is involved in the regulation of gene expression and cellular metabolism. Accumulated data demonstrate the significant role of m6A methylation in metabolic homeostasis.

Lipopolysaccharide (LPS) is considered an early trigger of inflammation and metabolic diseases. The *Fto* gene, which encodes m6A demethylase, has been found to be responsive to LPS and play a role in linking inflammation with metabolic responses (Zhang et al., 2016). LPS stimulation reduced the expression of carnitine palmitoyltransferase 1 (CPT1) in the liver, upregulated the truncated cFTO4 protein, and dramatically reduced m6A levels near the translation start site of CPT1 (Zhang et al., 2016). In addition, the core methyltransferase of m6A, METTL3, has been found to participate in the molecular



mechanism of inflammation-induced malabsorption of long-chain fatty acids (LCFAs). METTL3 inhibits LPS-induced inflammatory response by exerting anti-malabsorption of LCFA activity *in vitro* (Zong et al., 2019). The depletion of *Mettl3* reduces the m6A level of *Traf6* mRNA, thereby decreasing the expression of TRAF6 and inhibiting the NF-κB and MAPK signaling pathways (Zong et al., 2019). These data suggest that m6A methylation plays an essential role in cellular metabolic homeostasis.

In obesity, extensive inflammation and stress reactions usually occur, which subsequently develop into chronic, low-grade, local inflammation and disruption of metabolic homeostasis (Ghemrawi et al., 2018). Studies have shown a correlation between inflammation and obesity. Experiments in obese mice have shown that the amounts of proinflammatory cytokines secreted by adipocytes and macrophages in adipose tissues are sufficient to disrupt insulin signaling (Olefsky and Glass, 2010). In addition, activation of the JNK and NF-κB signaling pathways causes obesity-induced inflammation and abnormal insulin action. The *Fto* gene is known to play an important role in the regulation of metabolic homeostasis, mainly owing to its association with body mass index (BMI). FTO overexpression increases food intake and promotes obesity (Church et al., 2010). miR-495 inhibited the expression of its target gene *Fto*, resulting in the promotion of the transformation of macrophages into M1-type pro-inflammatory macrophages, and aggravated insulin resistance and adipose tissue inflammation in type 2 diabetes mellitus (T2D) -mice (Hu et al., 2019a). In contrast, dominant point mutations in mouse *Fto* gene decreased fat mass, promoted energy expenditure, and improved the inflammatory profile of white adipose tissue of mutant mice (Church et al., 2009). Recent scientific literature has shed light on the genetic relationship

between FTO and inflammation markers. Notably, studies have found that FTO is related to C-reactive protein (CRP) levels (Lighthart et al., 2016). Greater adiposity conferred by FTO SNPs leads to higher CRP levels (Welsh et al., 2010). Another study demonstrated that there was a significant positive correlation between hs-CRP, leptin, and broad BMI, but there was no significant difference in the FTO rs9939609 genotype (Zimmermann et al., 2011). Obesity is associated with a shorter telomere length (Zhou et al., 2017). Telomere length attrition may be affected by obesity-related inflammation, oxidative stress, and FTO gene pathways. The increase in BMI is genetically related to telomere length shortening, and low-grade inflammation may result from increased CRP levels (Rode et al., 2014). The association between the *Fto* gene and expression of inflammatory markers is unclear, and the underlying mechanisms remain to be elucidated.

Non-alcoholic fatty liver disease is accompanied by inflammation, which contributes to the development of fibrosis, cirrhosis, and hepatocellular carcinoma (HCC) (Sun et al., 2013). Chronic corticosterone exposure can induce liver inflammation and fibrosis and increase mRNA and m6A methylation of several heat shock proteins (HSPs) (Feng et al., 2020). HSPs are activated during acute stress response to exert a cytoprotective effect (Stolte et al., 2009), which has been confirmed to be related to m6A-mediated post-transcriptional regulation. In addition, the present study demonstrated that FTO expression was significantly increased in the livers of non-alcoholic steatohepatitis (NASH) patients and in a rodent model of NASH (Lim et al., 2016). Genetic silencing of *Fto* had a protective effect on palmitate-induced oxidative stress, endoplasmic reticulum (ER) stress, and apoptosis *in vitro* (Lim et al., 2016). Inflammation is a characteristic of Metabolic syndrome (MetS).

TABLE 1 | Roles of m6A in RNA metabolism and inflammatory processes.

Disease		Aberrant expression of m6A enzymes	Target RNA	Change of target RNA level	Effect of enzyme on target RNA	Role of m6A in diseases	Mechanism	References
Metabolic disorders	Obesity	FTO ↑	–			Increase fat mass, decrease energy expenditure	Promotes inflammatory profile in white adipose tissue; leads to higher CRP levels	Church et al., 2009; Zimmermann et al., 2011; Ligthart et al., 2016
	NASH	FTO ↑	–			Promotes fat accumulation, inflammation, and lipotoxicity in liver	Genetic silencing of FTO protects against palmitate-induced oxidative stress, mitochondrial dysfunction, ER stress, and apoptosis	Lim et al., 2016
Autoimmune diseases And Infection	RA	METTL3 ↑	(p)-NF-κB	↓	Phosphorylation and nucleus translocation	Increases inflammatory response in macrophages	Overexpression of METTL3 significantly attenuated the inflammatory response through the effect on NF-κB	Wang J. et al., 2019
	Dental pulp inflammation	METTL3 ↑	MyD88	↓	Splicing	Increases the expression of inflammatory cytokines	METTL3 regulate alternative splicing of MyD88; METTL3 depletion decreased the expression of inflammatory cytokines and the NF-κB signaling and MAPK signaling pathway	Feng et al., 2018
Cancer	HCC	YTHDF2 ↓	IL-11, SERPINE2	↑	Degradation	Promotes inflammation-mediated malignancy	YTHDF2 processed the decay of m6 A-containing IL11 and SERPINE2 mRNAs; YTHDF2 transcription succumbed to HIF-2α	Hou et al., 2019
	HCC	Unknown	GNAS	↑	Expression	Promotes LPS-induced HCC cell growth and invasion	Increasing m6A methylation of GNAS mRNA, GNAS promotes LPS-induced STAT3 activation in HCC cells through inhibiting long non-coding RNA TPTEP1 interacting with STAT3	Ding et al., 2020
	ICC	FTO ↓	NR5A2	↓	Expression	Regulates inflammatory gene expression and tumor growth	FTO knockdown inhibited the expression of NR5A2 which is associated with PDAC and transcriptionally regulates inflammatory gene expression	Rong et al., 2019
Others	Poststroke	FTO ↓	–			Modulates poststroke brain damage	Promotes cytokine signaling, TNF signaling, TLR signaling, and NF-κB signaling pathways	Chokkalla et al., 2019

FTO may serve as a risk factor for MetS and inflammatory markers (Kraja et al., 2014). Thus, m6A methylation regulators can act as a link between inflammation and metabolic responses. These findings also warrant further investigation.

Autoimmune Diseases and Infections

Inflammation is primarily caused by immune cells and cytokines. Inflammatory cells are composed of various immune cells, including neutrophils, macrophages, lymphocytes, and plasma cells. Cytokines are composed of inflammatory mediators, such as interleukin-1β (IL-1β), IL-6, and tumor necrosis factor-α (TNF-α) (Medzhitov, 2008). The presence of m6A modifications has been reported in both autoimmune diseases and infections.

Peripheral dendritic cells (DCs) mature in response to microbial products or inflammatory signals and then upregulate CC-chemokine receptor 7 (CCR7). m6A modification plays a critical role in DC-dependent inflammatory response. CCR7 stimulation upregulates lnc-Dpf3 by removing m6A modification to prevent RNA degradation (Liu et al., 2019). lnc-Dpf3 feedback restrains CCR7-induced DC migration and inflammatory response. Macrophages play an important role in various chronic diseases. Recent studies have shown that m6A modification is involved in regulating macrophage phenotype. Silencing the m6A “reader” YTHDF2 increases MAP2K4 and MAP4K4 mRNA expression levels by stabilizing mRNA transcription, which in turn activates MAPK and NF-κB signaling pathways, further

inducing the expression of pro-inflammatory cytokines and aggravating inflammatory response in LPS-stimulated RAW 264.7 cells (Yu et al., 2019). Rheumatoid arthritis (RA) is a chronic inflammatory autoimmune disease characterized by the progressive destruction of the articular cartilage and bone (Scott et al., 2010). Studies have shown that m6A modification is involved in RA pathogenesis. It has been reported that METTL3 expression significantly increased in RA patients and is positively correlated with CRP and ESR, the two common markers of RA disease activity (Wang J. et al., 2019). In addition, an *in vitro* study showed that LPS stimulation promoted the expression and biological activity of METTL3 in macrophages. However, overexpression of METTL3 can attenuate LPS-induced inflammation through the NF- κ B signaling pathway (Wang J. et al., 2019).

Commensal bacteria, especially gut microbiota, play a role in critical physiological functions such as host metabolism, immune system development, and even behavior (Sommer and Backhed, 2016). Metabolites and fermentation products from gut microbiota have been reported to mediate intestinal effects in the host by regulating transcription and epigenetic modifications (Koh et al., 2016; Agus et al., 2018). Variations in the gut microbiota correlate with m6A modifications in the cecum. METTL16 is another N6-adenosine-methyltransferase (Pendleton et al., 2017; Shima et al., 2017; Warda et al., 2017). A previous study reported that METTL16 expression decreases in the absence of microbiota. The target mRNA, *Mat2a*, which encodes S-adenosylmethionine synthase, is less methylated (Jabs et al., 2020). FTO deletion in mice resulted in changes in bacterial characteristics associated with reduced inflammation, such as a higher abundance of *Lactobacillus* and a lower content of Porphyromonadaceae and *Helicobacter* (Sun et al., 2019). *Helicobacter pylori* interacts with gastric epithelial cells to induce gastric inflammation and epithelial damage. The expression of WTAP has been confirmed to be downregulated in AGS cells stimulated by *H. pylori* (Kim et al., 2007). Dental pulp inflammation is a common public health problem caused by oral bacterial infections. It has been reported that m6A levels and METTL3 expression are increased in LPS-stimulated human dental pulp cells (HDPCs). METTL3 deletion promotes the expression of MyD88S in HDPCs induced by LPS, reduces the expression of inflammatory cytokines, and inhibits the activation of the NF- κ B and MAPK signaling pathways (Feng et al., 2018). These studies suggest that epitranscriptomic modifications act as an additional level of interaction between commensal bacteria and their host, affecting pathways related to inflammation and antimicrobial responses.

Cancer

Inflammation is associated with the development of a cancer-prone microenvironment. Many types of tumors occur in these inflammatory microenvironments, and the inflammatory state can be pro-tumorigenic (Singh et al., 2017). Studies have shown that a disruption in the balance between pro- and anti-inflammatory mechanisms leads to chronic inflammation, which can induce tumor initiation and progression (Seruga et al., 2008). As inflammation progresses, it promotes the production

of tumor-promoting cytokines (Singh et al., 2017) such as IL-11 (Bollrath et al., 2009), IL-1 β (Tu et al., 2008), and IL-6 (Grivennikov et al., 2009; Singh et al., 2017), which support the survival, growth, and metastasis of tumor cells (Garg et al., 2012). Recently, numerous studies have suggested an association between m6A modifications and pro-inflammatory genes in the tumor microenvironment.

In the complex tumor microenvironment network, mesenchymal stromal cells (MSCs) play a key role in promoting tumor progression by interacting with tumor cells and other stromal cells (Pelizzo et al., 2018; Whiteside, 2018). The inflammation and autophagy-related gene P62 is highly expressed in most human tumor tissues (Moscat and Diaz-Meco, 2009; Ren et al., 2014). It has been demonstrated that P62 synergizes with TNF- α to promote the malignant transformation of human MSCs by forming insulin growth factor II (IGF-II) promoter-enhancer chromatin loops and increasing METTL3 occupancy on IGFII 3' UTR, and upregulates the expression of H-Ras by harboring inflammation-related factors, such as TNFR1, CLYD, EGR1, NF- κ B, TLR4, and PPAR γ (Xin et al., 2018). In HCC, m6A modification levels and mRNA expression are increased. A recent study found that the deletion of the m6A "reader" YTHDF2 in human HCC cells or mouse hepatocytes aggravates inflammation, vascular reconstruction, and metastatic progression (Hou et al., 2019). YTHDF2 promotes the decay of m6A-containing IL-11 and serpin family E member 2 (SERPINE2) mRNAs, which are involved in inflammation-mediated malignant tumors and the destruction of normal blood vessels (Hou et al., 2019). Another study reported that LPS stimulation can increase m6A methylation of G-protein alpha-subunit (GNAS) mRNA to promote GNAS expression in HCC cells. Highly expressed GNAS promotes the growth and invasion of HCC cells by interacting with STAT3 (Ding et al., 2020). Intrahepatic cholangiocarcinoma (ICC) is the second most malignant type of primary liver cancer with a high degree of incidence and a very poor prognosis. FTO was found to regulate the inflammatory signaling pathways in clinical ICC samples. FTO knockdown inhibited the expression of NR5A2, which is a nuclear receptor related to pancreatic adenocarcinoma and transcriptionally regulates the expression of inflammatory genes (Rong et al., 2019). Moreover, the expression level of FTO was decreased in clinical ICC samples and inversely correlated with CA19-9 expression and micro-vessel density (MVD) (Rong et al., 2019). Thus, these results suggest that FTO may be a target for predicting the prognosis of ICC. It has recently been found that the evaluation of m6A modification patterns within individual tumors could predict tumor inflammation stage and prognosis in gastric cancer (Zhang et al., 2020). A low m6A score, characterized by increased mutation burden and activation of immunity, has a higher 5-year survival rate. However, activation of the stroma and lack of effective immune infiltration were observed in the high m6A score subtype and associated with poorer survival. Patients with a lower m6A score demonstrated significant therapeutic and clinical benefits. In addition, m6A modification of non-coding RNAs can affect tumor formation. METTL3 may have an oncogenic role in bladder cancer by positively modulating

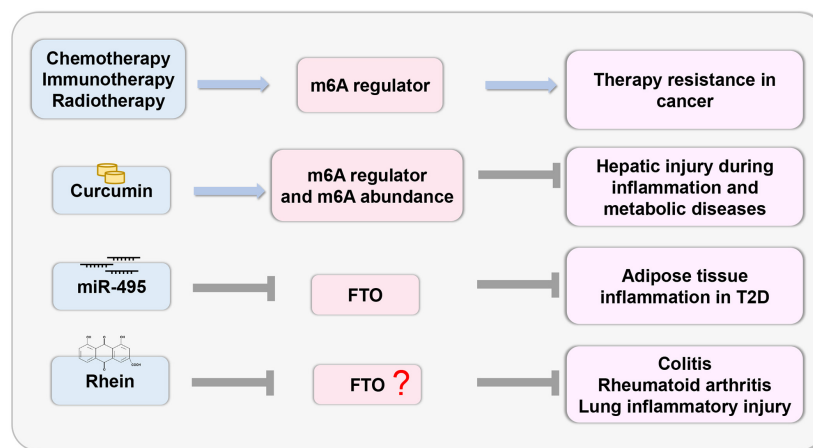


FIGURE 4 | Therapeutic potential based on m6A. The study of epigenetic changes in inflammatory response provides the possibility to develop effective drugs with specific targets based on the m6A regulator.

pri-miR221/222 maturation in an m6A-dependent manner (Han et al., 2019).

Others

Many infectious diseases, such as osteomyelitis, osteoarthritis, and periodontitis, can disrupt bone homeostasis (Josse et al., 2015; Pacios et al., 2015; Maruotti et al., 2017). LPS is the main pathogenic factor in infectious bone destruction. Studies have revealed that METTL3 knockdown stabilizes Smad7 and Smurf1 mRNA transcripts via YTHDF2 to inhibit osteoblast differentiation and Smad-dependent signaling, and regulates MAPK signaling to activate inflammatory response to LPS (Zhang et al., 2019). A previous study emphasized the influence of genetic risk for obesity and osteoarthritis, but the association is only regulated by the effect on BMI, which is consistent with what is known about the biology of the FTO gene (Panoutsopoulou et al., 2014).

Neuronal fate after ischemic stroke is determined by intricate biochemical and molecular events, including oxidative stress, ER stress, apoptosis, and inflammation (Mehta et al., 2007). The brain has been shown to have higher m6A abundance than other mammalian organs (Meyer et al., 2012). Thus, post-transcriptional regulators fine-tune the post-ischemic pathophysiology. An *in silico* analysis demonstrated increased m6A methylation in major inflammatory pathways, including IL-6 cytokine, TNF, TLR, and NF- κ B signaling pathways. The expression of major neuronally localized m6A demethylase FTO was downregulated, which presumably leads to decreased demethylation of m6A tagged RNAs in the post-stroke brain. However, ALKBH5 expression was unaltered in the ischemic brain (Chokkalla et al., 2019). In addition, a decrease in total m6A demethylase activity was reported in the study, indicating that FTO downregulation might be responsible for the decreased m6A demethylation in the post-stroke brain (Chokkalla et al., 2019). Therefore, it is plausible that ischemia-induced loss of FTO might be a modulator of post-ischemic secondary brain damage.

As a major cause of mortality worldwide, cardiovascular and metabolic diseases have a higher risk of occurrence in people with high levels of systemic inflammation and are independent of traditional cardiometabolic risk factors (Danesh et al., 2004). A recent study evaluated the association between inflammation and cardiometabolic phenotypes. The genetic overlap includes FTO, which has been confirmed to have an effect on CRP (Ligthart et al., 2015). Owing to its role in obesity-predisposing genetic variants, FTO has been shown to have nominally significant associations with cardiovascular disease and is referred to as a cardiovascular biomarker (He et al., 2010).

Polycystic ovary syndrome (PCOS) is a highly heterogeneous disease of the reproductive system that is associated with cellular metabolism and chronic inflammation (Ojeda-Ojeda et al., 2013). It has been reported that SNP rs9939609 of FTO is significantly related to PCOS (Zhao et al., 2016). Notably, FTO is a risk factor for PCOS that is independent of its correlation with BMI or obesity (Zhao et al., 2016). Therefore, the FTO gene has a profound effect on PCOS and is considered a candidate gene for this disorder.

Therapeutic Potential

Recent studies have demonstrated that deregulation of m6A regulators is linked to cancer therapy resistance. Chemotherapeutic drugs modulate m6A regulators to stabilize the mRNA of oncogenes and induce chemoresistance (Shriwas et al., 2020). The tumor microenvironment activated by m6A regulators induces resistance to immunotherapy in cancer cells (Yang et al., 2019; Yi et al., 2020). Radiation stabilizes the mRNA of cancer stem cells by modulating the m6A regulator (Taketo et al., 2018; Visvanathan et al., 2018). As m6A plays an important role in therapy resistance, the targeted inhibition of m6A regulators may lead to better outcomes.

The study of epigenetic changes in inflammatory response provides the possibility of developing effective drugs with specific targets (Figure 4). Recently, curcumin, a yellow polyphenolic pigment derived from the spice turmeric (*Curcuma longa*), has

been reported to protect against LPS-induced liver injury and lipid metabolism disorders. Dietary curcumin can regulate the mRNA expression of METTL3, METTL14, ALKBH5, FTO, and YTHDF2 and increase m6A abundance, which has a protective effect on hepatic injury due to inflammation and metabolic diseases (Lu et al., 2018). In T2D, miR-495 inhibited FTO expression, thus promoting the transformation of macrophages into M1-type pro-inflammatory macrophages, and aggravated insulin resistance and adipose tissue inflammation (Hu et al., 2019a). Thus, this research provides a theoretical basis for targeted treatment. Notably, there are few inhibitors that specifically target m6A regulatory proteins. The natural product Rhein, which is identified as a demethylase FTO inhibitor, competitively binds to the FTO active site *in vitro* (Chen et al., 2012). Rhein has been found to suppress ATP-triggered inflammatory responses in fibroblast-like synoviocytes of rheumatoid rats and ameliorate experimental colitis (Hu et al., 2019b; Wu et al., 2020). In addition, Rhein treatment can inhibit inflammatory lung injury induced by the human respiratory syncytial virus (Shen et al., 2019). However, whether these protective effects are mediated by m6A modifications remain to be investigated. Thus, more efficacious medicines and novel therapeutic strategies related to m6A modification should be explored.

PERSPECTIVES AND CONCLUSION

Although m6A has become the focus of numerous studies, the mechanism by which m6A modifications affect inflammation in human diseases remains unclear. For instance, human HCC exhibits a characteristic increase in m6A modification, which is related to inflammation-mediated malignancy (Hou et al., 2019). Meanwhile, LPS stimulation can increase m6A levels and trigger inflammatory cytokine production and pathways in liver tumors and dental pulp inflammation (Feng et al., 2018; Ding et al., 2020). Mechanistically, it has been reported that silencing *Mettl3* could maintain LCFA absorption by blocking

the TRAF6-dependent inflammation response (Zong et al., 2019). However, METTL3 depletion enhances the expression of pro-inflammatory cytokines and promotes the activation of MAPK and NF- κ B signaling pathways in osteoblasts (Zhang et al., 2019). In RA, LPS stimulation upregulates the expression and biological activity of METTL3 in macrophages, while METTL3 overexpression inhibited the inflammatory response (Wang J. et al., 2019). These data suggest that the associations between m6A modifications and inflammatory responses are highly dependent on the context of specific diseases and signaling molecules, especially as many m6A regulators are still being investigated to identify novel functions.

In summary, RNA methylation plays essential roles in various inflammatory disease states, including autoimmunity, infection, metabolic disease, cancer, neurodegenerative disease, heart disease, and bone disease via multiple mechanisms. Nevertheless, our understanding of this phenomenon is still incomplete. A major challenge is to study the precise functions of each m6A regulatory factor in different cells during the inflammatory response at various stages of disease development. In addition, other components of m6A modification remain to be discovered. Importantly, the development of scientific technology, such as the application of m6A mapping methods and m6A editing tools, will be immensely helpful for future research on m6A modification at the single nucleotide level.

AUTHOR CONTRIBUTIONS

TX and KS: concept and design. JL: drafting the manuscript. All authors: revising the manuscript content and approving the final version of the manuscript.

FUNDING

This study was supported by the National Natural Science Foundation of China (no. 82072556).

REFERENCES

- Agus, A., Planchais, J., and Sokol, H. (2018). Gut Microbiota Regulation of Tryptophan Metabolism in Health and Disease. *Cell Host Microbe* 23, 716–724. doi: 10.1016/j.chom.2018.05.003
- Alarcón Claudio, R., Goodarzi, H., Lee, H., Liu, X., Tavazoie, S., and Tavazoie Sohail, F. (2015). HNRNPA2B1 Is a Mediator of m6A-Dependent Nuclear RNA Processing Events. *Cell* 162, 1299–1308. doi: 10.1016/j.cell.2015.08.011
- Barbieri, I., Tzelepis, K., Pandolfini, L., Shi, J., Millán-Zambrano, G., Robson, S. C., et al. (2017). Promoter-bound METTL3 maintains myeloid leukaemia by m(6)A-dependent translation control. *Nature* 552, 126–131. doi: 10.1038/nature24678
- Batista, P. J. (2017). The RNA Modification N(6)-methyladenosine and Its Implications in Human Disease. *Genom. Proteom. Bioinform.* 15, 154–163. doi: 10.1016/j.gpb.2017.03.002
- Bayarsaihan, D. (2011). Epigenetic mechanisms in inflammation. *J. Dent. Res.* 90, 9–17. doi: 10.1177/0022034510378683
- Bollrath, J., Phesse, T. J., von Burstin, V. A., Putoczki, T., Bennecke, M., Bateman, T., et al. (2009). gp130-Mediated Stat3 Activation in Enterocytes Regulates Cell Survival and Cell-Cycle Progression during Colitis-Associated Tumorigenesis. *Cancer Cell* 15, 91–102. doi: 10.1016/j.ccr.2009.01.002
- Chen, B., Ye, F., Yu, L., Jia, G., Huang, X., Zhang, X., et al. (2012). Development of cell-active N6-methyladenosine RNA demethylase FTO inhibitor. *J. Am. Chem. Soc.* 134, 17963–17971. doi: 10.1021/ja3064149
- Chen, M., and Wong, C.-M. (2020). The emerging roles of N6-methyladenosine (m6A) deregulation in liver carcinogenesis. *Mol. Cancer* 19:44. doi: 10.1186/s12943-020-01172-y
- Chen, Y. G., Chen, R., Ahmad, S., Verma, R., Kasturi, S. P., Amaya, L., et al. (2019). N6-Methyladenosine Modification Controls Circular RNA Immunity. *Mol. Cell* 76, 96.e–109.e. doi: 10.1016/j.molcel.2019.07.016
- Chokkalla, A. K., Mehta, S. L., Kim, T., Chelluboina, B., Kim, J., and Vemuganti, R. (2019). Transient Focal Ischemia Significantly Alters the m(6)A Epitranscriptomic Tagging of RNAs in the Brain. *Stroke* 50, 2912–2921. doi: 10.1161/strokeaha.119.026433
- Church, C., Lee, S., Bagg, E. A., McTaggart, J. S., Deacon, R., Gerken, T., et al. (2009). A mouse model for the metabolic effects of the human fat mass and obesity associated FTO gene. *PLoS Genet.* 5:e1000599. doi: 10.1371/journal.pgen.1000599

- Church, C., Moir, L., McMurray, F., Girard, C., Banks, G. T., Teboul, L., et al. (2010). Overexpression of Fto leads to increased food intake and results in obesity. *Nat. Genet.* 42, 1086–1092. doi: 10.1038/ng.713
- Dai, D., Wang, H., Zhu, L., Jin, H., and Wang, X. (2018). N6-methyladenosine links RNA metabolism to cancer progression. *Cell Death Dis.* 9:124. doi: 10.1038/s41419-017-0129-x
- Danesh, J., Wheeler, J. G., Hirschfield, G. M., Eda, S., Eiriksdottir, G., Rumley, A., et al. (2004). C-reactive protein and other circulating markers of inflammation in the prediction of coronary heart disease. *N. Engl. J. Med.* 350, 1387–1397. doi: 10.1056/NEJMoa032804
- Desrosiers, R., Friderici, K., and Rottman, F. (1974). Identification of methylated nucleosides in messenger RNA from Novikoff hepatoma cells. *Proc. Natl. Acad. Sci. U. S. A.* 71, 3971–3975. doi: 10.1073/pnas.71.10.3971
- Ding, H., Zhang, X., Su, Y., Jia, C., and Dai, C. (2020). GNAS promotes inflammation-related hepatocellular carcinoma progression by promoting STAT3 activation. *Cell Mol. Biol. Lett.* 25:8. doi: 10.1186/s11658-020-00204-1
- Dominissini, D., Moshitch-Moshkovitz, S., Schwartz, S., Salmon-Divon, M., Ungar, L., Osenberg, S., et al. (2012). Topology of the human and mouse m6A RNA methylomes revealed by m6A-seq. *Nature* 485, 201–206. doi: 10.1038/nature11112
- Du, H., Zhao, Y., He, J., Zhang, Y., Xi, H., Liu, M., et al. (2016). YTHDF2 destabilizes m6A-containing RNA through direct recruitment of the CCR4–NOT deadenylase complex. *Nat. Commun.* 7:12626. doi: 10.1038/ncomms12626
- Dubin, D. T., and Taylor, R. H. (1975). The methylation state of poly A-containing messenger RNA from cultured hamster cells. *Nucleic Acids Res.* 2, 1653–1668. doi: 10.1093/nar/2.10.1653
- Dunning, T. (2009). Periodontal disease- the overlooked diabetes complication. *Nephrol. Nurs. J.* 36, 489–495.
- Durbin, A. F., Wang, C., Marcotrigiano, J., and Gehrke, L. (2016). RNAs Containing Modified Nucleotides Fail To Trigger RIG-I Conformational Changes for Innate Immune Signaling. *mBio* 7:5. doi: 10.1128/mBio.00833-16
- Fedeles, B. I., Singh, V., Delaney, J. C., Li, D., and Essigmann, J. M. (2015). The AlkB Family of Fe(II)/ α -Ketoglutarate-dependent Dioxygenases: Repairing Nucleic Acid Alkylation Damage and Beyond. *J. Biol. Chem.* 290, 20734–20742. doi: 10.1074/jbc.R115.656462
- Feng, Y., Hu, Y., Hou, Z., Sun, Q., Jia, Y., and Zhao, R. (2020). Chronic corticosterone exposure induces liver inflammation and fibrosis in association with m(6)A-linked post-transcriptional suppression of heat shock proteins in chicken. *Cell Stress Chaper.* 25, 47–56. doi: 10.1007/s12192-019-01034-7
- Feng, Z., Li, Q., Meng, R., Yi, B., and Xu, Q. (2018). METTL3 regulates alternative splicing of MyD88 upon the lipopolysaccharide-induced inflammatory response in human dental pulp cells. *J. Cell. Mol. Med.* 22, 2558–2568. doi: 10.1111/jcmm.13491
- Fu, Y., Dominissini, D., Rechavi, G., and He, C. (2014). Gene expression regulation mediated through reversible m6A RNA methylation. *Nat. Rev. Genet.* 15, 293–306. doi: 10.1038/nrg3724
- Gao, Y., Vasic, R., Song, Y., Teng, R., Liu, C., Gbyli, R., et al. (2020). m(6)A Modification Prevents Formation of Endogenous Double-Stranded RNAs and Deleterious Innate Immune Responses during Hematopoietic Development. *Immunity* 52, 1007–1021.e. doi: 10.1016/j.immuni.2020.05.003
- Garg, A. D., Kaczmarek, A., Krysko, O., Vandenabeele, P., Krysko, D. V., and Agostinis, P. E. R. (2012). stress-induced inflammation: does it aid or impede disease progression? *Trends Mol. Med.* 18, 589–598. doi: 10.1016/j.molmed.2012.06.010
- Ghemrawi, R., Battaglia-Hsu, S.-F., and Arnold, C. (2018). Endoplasmic Reticulum Stress in Metabolic Disorders. *Cells* 7:63. doi: 10.3390/cells7060063
- Grivennikov, S., Karin, E., Terzic, J., Mucida, D., Yu, G.-Y., Vallabhapurapu, S., et al. (2009). IL-6 and Stat3 are required for survival of intestinal epithelial cells and development of colitis-associated cancer. *Cancer Cell.* 15, 103–113. doi: 10.1016/j.ccr.2009.01.001
- Han, J., Wang, J. Z., Yang, X., Yu, H., Zhou, R., Lu, H. C., et al. (2019). METTL3 promote tumor proliferation of bladder cancer by accelerating pri-miR221/222 maturation in m6A-dependent manner. *Mol. Cancer* 18:110. doi: 10.1186/s12943-019-1036-9
- Hawiger, J., and Zienkiewicz, J. (2019). Decoding inflammation, its causes, genomic responses, and emerging countermeasures. *Scand. J. Immunol.* 90:e12812. doi: 10.1111/sji.12812
- He, L., Li, H., Wu, A., Peng, Y., Shu, G., and Yin, G. (2019). Functions of N6-methyladenosine and its role in cancer. *Mol. Cancer* 18:176. doi: 10.1186/s12943-019-1109-9
- He, M., Cornelis, M. C., Franks, P. W., Zhang, C., Hu, F. B., and Qi, L. (2010). Obesity genotype score and cardiovascular risk in women with type 2 diabetes mellitus. *Arterioscl. Thromb. Vasc. Biol.* 30, 327–332. doi: 10.1161/ATVBAHA.109.196196
- Hou, J., Zhang, H., Liu, J., Zhao, Z., Wang, J., Lu, Z., et al. (2019). YTHDF2 reduction fuels inflammation and vascular abnormalization in hepatocellular carcinoma. *Mol. Cancer* 18:163. doi: 10.1186/s12943-019-1082-3
- Hsu, P. J., Zhu, Y., Ma, H., Guo, Y., Shi, X., Liu, Y., et al. (2017). Ythdc2 is an N(6)-methyladenosine binding protein that regulates mammalian spermatogenesis. *Cell. Res.* 27, 1115–1127. doi: 10.1038/cr.2017.99
- Hu, B.-B., Wang, X.-Y., Gu, X.-Y., Zou, C., Gao, Z.-J., Zhang, H., et al. (2019). N(6)-methyladenosine (m(6)A) RNA modification in gastrointestinal tract cancers: roles, mechanisms, and applications. *Mol. Cancer* 18:178. doi: 10.1186/s12943-019-1099-7
- Hu, F., Tong, J., Deng, B., Zheng, J., and Lu, C. (2019a). MiR-495 regulates macrophage M1/M2 polarization and insulin resistance in high-fat diet-fed mice via targeting FTO. *Pflugers Arch.* 471, 1529–1537. doi: 10.1007/s00424-019-02316-w
- Hu, F., Zhu, D., Pei, W., Lee, I., Zhang, X., Pan, L., et al. (2019b). Rhein inhibits ATP-triggered inflammatory responses in rheumatoid rat fibroblast-like synoviocytes. *Int. Immunopharmacol.* 75:105780. doi: 10.1016/j.intimp.2019.105780
- Jabs, S., Biton, A., and Becavin, C. (2020). Impact of the gut microbiota on the m(6)A epitranscriptome of mouse cecum and liver. *Nat. Commun.* 11:1344. doi: 10.1038/s41467-020-15126-x
- Jia, G., Fu, Y., Zhao, X., Dai, Q., Zheng, G., Yang, Y., et al. (2011). N6-methyladenosine in nuclear RNA is a major substrate of the obesity-associated FTO. *Nat. Chem. Biol.* 7, 885–887. doi: 10.1038/nchembio.687
- Josse, J., Velard, F., and Gangloff, S. C. (2015). Staphylococcus aureus vs. Osteoblast: Relationship and Consequences in Osteomyelitis. *Front. Cell. Infect. Microbiol.* 5:85. doi: 10.3389/fcimb.2015.00085
- Jung, Y., and Goldman, D. (2018). Role of RNA modifications in brain and behavior. *Genes Brain Behav.* 17:e12444. doi: 10.1111/gbb.12444
- Kane, S. E., and Beemon, K. (1985). Precise localization of m6A in Rous sarcoma virus RNA reveals clustering of methylation sites: implications for RNA processing. *Mol. Cell. Biol.* 5:2298. doi: 10.1128/MCB.5.9.2298
- Karikó, K., Buckstein, M., Ni, H., and Weissman, D. (2005). Suppression of RNA recognition by Toll-like receptors: the impact of nucleoside modification and the evolutionary origin of RNA. *Immunity* 23, 165–175. doi: 10.1016/j.immuni.2005.06.008
- Kim, N., Park, W. Y., Kim, J. M., Park, Y. S., Lee, D. H., Park, J. H., et al. (2007). Analysis of Gene Expression Profile of AGS Cells Stimulated by *Helicobacter pylori* Adhesion. *Gut. Liver* 1, 40–48. doi: 10.5009/gnl.2007.1.1.40
- Koh, A., De Vadder, F., Kovatcheva-Datchary, P., and Backhed, F. (2016). From Dietary Fiber to Host Physiology: Short-Chain Fatty Acids as Key Bacterial Metabolites. *Cell* 165, 1332–1345. doi: 10.1016/j.cell.2016.05.041
- König, J., Zarnack, K., Rot, G., Curk, T., Kayikci, M., Zupan, B., et al. (2010). iCLIP reveals the function of hnRNP particles in splicing at individual nucleotide resolution. *Nat. Struct. Mol. Biol.* 17, 909–915. doi: 10.1038/nsmb.1838
- Kraja, A. T., Chasman, D. I., North, K. E., Reiner, A. P., Yanek, L. R., Kilpeläinen, T. O., et al. (2014). Pleiotropic genes for metabolic syndrome and inflammation. *Mol. Genet. Metab.* 112, 317–338. doi: 10.1016/j.ymgme.2014.04.007
- Lasman, L., Krupalnik, V., Viukov, S., Mor, N., Aguilera-Castrejon, A., Schneir, D., et al. (2020). Context-dependent functional compensation between Ythdf m(6)A reader proteins. *Genes. Dev.* 34, 1373–1391. doi: 10.1101/gad.340695.120
- Li, A., Chen, Y.-S., Ping, X.-L., Yang, X., Xiao, W., Yang, Y., et al. (2017). Cytoplasmic m6A reader YTHDF3 promotes mRNA translation. *Cell. Res.* 27, 444–447. doi: 10.1038/cr.2017.10
- Li, T., Hu, P.-S., Zuo, Z., Lin, J.-F., Li, X., Wu, Q.-N., et al. (2019). METTL3 facilitates tumor progression via an m(6)A-IGF2BP2-dependent mechanism in colorectal carcinoma. *Mol. Cancer* 18:112. doi: 10.1186/s12943-019-1038-7
- Li, X., Xiong, X., Wang, K., Wang, L., Shu, X., Ma, S., et al. (2016). Transcriptome-wide mapping reveals reversible and dynamic N(1)-methyladenosine methylome. *Nat. Chem. Biol.* 12, 311–316. doi: 10.1038/nchembio.2040

- Liao, S., Sun, H., and Xu, C. Y. T. H. (2018). Domain: A Family of N(6)-methyladenosine (m(6)A) Readers. *Genom. Proteom. Bioinform.* 16, 99–107. doi: 10.1016/j.gpb.2018.04.002
- Ligthart, S., de Vries, P. S., Uitterlinden, A. G., Hofman, A., Group Clw, Franco, O. H., et al. (2015). Pleiotropy among common genetic loci identified for cardiometabolic disorders and C-reactive protein. *PLoS One* 10:e0118859. doi: 10.1371/journal.pone.0118859
- Ligthart, S., Vaez, A., Hsu, Y. H., Stolk, R., Uitterlinden, A. G., Hofman, A., et al. (2016). Bivariate genome-wide association study identifies novel pleiotropic loci for lipids and inflammation. *BMC Genom.* 17:443. doi: 10.1186/s12864-016-2712-4
- Lim, A., Zhou, J., Sinha, R. A., Singh, B. K., Ghosh, S., Lim, K. H., et al. (2016). Hepatic FTO expression is increased in NASH and its silencing attenuates palmitic acid-induced lipotoxicity. *Biochem. Biophys. Res. Commun.* 479, 476–481. doi: 10.1016/j.bbrc.2016.09.086
- Liu, J., Yue, Y., Han, D., Wang, X., Fu, Y., Zhang, L., et al. (2014). A METTL3-METTL14 complex mediates mammalian nuclear RNA N6-adenosine methylation. *Nat. Chem. Biol.* 10, 93–95. doi: 10.1038/nchembio.1432
- Liu, J., Zhang, X., Chen, K., Cheng, Y., Liu, S., Xia, M., et al. (2019). CCR7 Chemokine Receptor-Inducible Inc-Dpf3 Restrains Dendritic Cell Migration by Inhibiting HIF-1 α -Mediated Glycolysis. *Immunity* 50, 600–615.e. doi: 10.1016/j.immuni.2019.01.021
- Liu, N., Dai, Q., Zheng, G., He, C., Parisien, M., and Pan, T. (2015). N6-methyladenosine-dependent RNA structural switches regulate RNA-protein interactions. *Nature* 518, 560–564. doi: 10.1038/nature14234
- Liu, N., Zhou, K. I., Parisien, M., Dai, Q., Diatchenko, L., and Pan, T. (2017). N6-methyladenosine alters RNA structure to regulate binding of a low-complexity protein. *Nucleic Acids Res.* 45, 6051–6063. doi: 10.1093/nar/gkx141
- Liu, Q., and Gregory, R. I. (2019). RNAmoD: an integrated system for the annotation of mRNA modifications. *Nucleic Acids Res.* 47, W548–W555. doi: 10.1093/nar/gkz479
- Lu, M., Zhang, Z., Xue, M., Zhao, B. S., Harder, O., Li, A., et al. (2020). N(6)-methyladenosine modification enables viral RNA to escape recognition by RNA sensor RIG-I. *Nat. Microbiol.* 5, 584–598. doi: 10.1038/s41564-019-0653-9
- Lu, N., Li, X., Yu, J., Li, Y., Wang, C., Zhang, L., et al. (2018). Curcumin Attenuates Lipopolysaccharide-Induced Hepatic Lipid Metabolism Disorder by Modification of m(6) A RNA Methylation in Piglets. *Lipids* 53, 53–63. doi: 10.1002/lipd.12023
- Maruotti, N., Corrado, A., and Cantatore, F. P. (2017). Osteoblast role in osteoarthritis pathogenesis. *J. Cell. Physiol.* 232, 2957–2963. doi: 10.1002/jcp.25969
- Medzhitov, R. (2008). Origin and physiological roles of inflammation. *Nature* 454, 428–435. doi: 10.1038/nature07201
- Medzhitov, R., and Horng, T. (2009). Transcriptional control of the inflammatory response. *Nat. Rev. Immunol.* 9, 692–703. doi: 10.1038/nri2634
- Mehta, S. L., Manhas, N., and Raghubir, R. (2007). Molecular targets in cerebral ischemia for developing novel therapeutics. *Brain Res. Rev.* 54, 34–66. doi: 10.1016/j.brainresrev.2006.11.003
- Meyer, Kate, D., Patil Deepak, P., Zhou, J., Zinoviev, A., Skabkin Maxim, A., Elemento, O., et al. (2015). 5' UTR m6A Promotes Cap-Independent Translation. *Cell* 163, 999–1010. doi: 10.1016/j.cell.2015.10.012
- Meyer, K. D., and Jaffrey, S. R. (2017). Rethinking m6A Readers, Writers, and Erasers. *Annu. Rev. Cell. Devel. Biol.* 33, 319–342. doi: 10.1146/annurev-cellbio-100616-060758
- Meyer, K. D., Saletore, Y., Zumbo, P., Elemento, O., Mason, C. E., and Jaffrey, S. R. (2012). Comprehensive analysis of mRNA methylation reveals enrichment in 3' UTRs and near stop codons. *Cell* 149, 1635–1646. doi: 10.1016/j.cell.2012.05.003
- Moindrot, B., Cerase, A., Coker, H., Masui, O., Grijzenhout, A., Pintacuda, G., et al. (2015). A Pooled shRNA Screen Identifies Rbm15, Spen, and Wtap as Factors Required for Xist RNA-Mediated Silencing. *Cell. Rep.* 12, 562–572. doi: 10.1016/j.celrep.2015.06.053
- Moscat, J., and Diaz-Meco, M. T. (2009). p62 at the crossroads of autophagy, apoptosis, and cancer. *Cell* 137, 1001–1004. doi: 10.1016/j.cell.2009.05.023
- Müller, S., Glaß, M., Singh, A. K., Haase, J., Bley, N., Fuchs, T., et al. (2019). IGF2BP1 promotes SRF-dependent transcription in cancer in a m6A- and miRNA-dependent manner. *Nucleic Acids Res.* 47, 375–390. doi: 10.1093/nar/gky1012
- Nachtergaele, S., and He, C. (2018). Chemical Modifications in the Life of an mRNA Transcript. *Annu. Rev. Genet.* 52, 349–372. doi: 10.1146/annurev-genet-120417-031522
- Nemat, K., Yadollah, S., and Mahdi, M. (2009). Chronic Inflammation and Oxidative Stress as a Major Cause of Age-Related Diseases and Cancer. *Rec. Pat. Inflam. Allerg. Drug Discov.* 3, 73–80. doi: 10.2174/187221309787158371
- Ojeda-Ojeda, M., Murri, M., Insenser, M., and Escobar-Morreale, H. F. (2013). Mediators of low-grade chronic inflammation in polycystic ovary syndrome (PCOS). *Curr. Pharmaceut. Des.* 19, 5775–5791. doi: 10.2174/1381612811319320012
- Olefsky, J. M., and Glass, C. K. (2010). Macrophages, inflammation, and insulin resistance. *Annu. Rev. Physiol.* 72, 219–246. doi: 10.1146/annurev-physiol-021909-135846
- Pacios, S., Xiao, W., Mattos, M., Lim, J., Tarapore, R. S., Alsadun, S., et al. (2015). Osteoblast Lineage Cells Play an Essential Role in Periodontal Bone Loss Through Activation of Nuclear Factor-Kappa B. *Sci. Rep.* 5:16694. doi: 10.1038/srep16694
- Panoutsopoulou, K., Metrustry, S., Doherty, S. A., Laslett, L. L., Maciewicz, R. A., Hart, D. J., et al. (2014). The effect of FTO variation on increased osteoarthritis risk is mediated through body mass index: a Mendelian randomisation study. *Ann. Rheum. Dis.* 73, 2082–2086. doi: 10.1136/annrheumdis-2013-203772
- Pelizzo, G., Veschi, V., Mantelli, M., Croce, S., Di Benedetto, V., D'Angelo, P., et al. (2018). Microenvironment in neuroblastoma: isolation and characterization of tumor-derived mesenchymal stromal cells. *BMC Cancer* 18:1176. doi: 10.1186/s12885-018-5082-2
- Pendleton, K. E., Chen, B., Liu, K., Hunter, O. V., Xie, Y., Tu, B. P., et al. (2017). The U6 snRNA m(6)A Methyltransferase METTL16 Regulates SAM Synthetase Intron Retention. *Cell* 169, 824–835.e. doi: 10.1016/j.cell.2017.05.003
- Ping, X.-L., Sun, B.-F., Wang, L., Xiao, W., Yang, X., Wang, W.-J., et al. (2014). Mammalian WTAP is a regulatory subunit of the RNA N6-methyladenosine methyltransferase. *Cell Res.* 24, 177–189. doi: 10.1038/cr.2014.3
- Ren, F., Shu, G., Liu, G., Liu, D., Zhou, J., Yuan, L., et al. (2014). Knockdown of p62/sequestosome 1 attenuates autophagy and inhibits colorectal cancer cell growth. *Mol. Cell. Biochem.* 385, 95–102. doi: 10.1007/s11010-013-1818-0
- Ries, R. J., Zaccara, S., Klein, P., Olarerin-George, A., Namkoong, S., Pickering, B. F., et al. (2019). m6A enhances the phase separation potential of mRNA. *Nature* 571, 424–428. doi: 10.1038/s41586-019-1374-1
- Rode, L., Nordestgaard, B. G., Weischer, M., and Bojesen, S. E. (2014). Increased body mass index, elevated C-reactive protein, and short telomere length. *J. Clin. Endocrinol. Metabol.* 99, E1671–E1675. doi: 10.1210/jc.2014-1161
- Rong, Z.-X., Li, Z., He, J.-J., Liu, L.-Y., Ren, X.-X., Gao, J., et al. (2019). Downregulation of Fat Mass and Obesity Associated (FTO) Promotes the Progression of Intrahepatic Cholangiocarcinoma. *Front. Oncol.* 9:369. doi: 10.3389/fonc.2019.00369
- Roundtree, I. A., Evans, M. E., Pan, T., and He, C. (2017a). Dynamic RNA Modifications in Gene Expression Regulation. *Cell* 169, 1187–1200. doi: 10.1016/j.cell.2017.05.045
- Roundtree, I. A., Luo, G.-Z., Zhang, Z., Wang, X., Zhou, T., Cui, Y., et al. (2017b). YTHDC1 mediates nuclear export of N(6)-methyladenosine methylated mRNAs. *Elife* 6:e31311. doi: 10.7554/eLife.31311
- Rubio, R. M., Depledge, D. P., Bianco, C., Thompson, L., and Mohr, I. R. (2018). m(6) A modification enzymes shape innate responses to DNA by regulating interferon β . *Genes. Dev.* 32, 1472–1484. doi: 10.1101/gad.319475.118
- Schöller, E., Weichmann, F., Treiber, T., Ringle, S., Treiber, N., Flatley, A., et al. (2018). Interactions, localization, and phosphorylation of the m(6)A generating METTL3-METTL14-WTAP complex. *RNA* 24, 499–512. doi: 10.1261/rna.064063.117
- Schwartz, S., Mumbach Maxwell, R., Jovanovic, M., Wang, T., Maciag, K., Bushkin, G. G., et al. (2014). Perturbation of m6A Writers Reveals Two Distinct Classes of mRNA Methylation at Internal and 5' Sites. *Cell Rep.* 8, 284–296. doi: 10.1016/j.celrep.2014.05.048
- Scott, D. L., Wolfe, F., and Huizinga, T. W. (2010). Rheumatoid arthritis. *Lancet* 376, 1094–1108. doi: 10.1016/s0140-6736(10)60826-4
- Seruga, B., Zhang, H., Bernstein, L. J., and Tannock, I. F. (2008). Cytokines and their relationship to the symptoms and outcome of cancer. *Nat. Rev. Cancer* 8, 887–899. doi: 10.1038/nrc2507

- Shen, C., Zhang, Z., Xie, T., Ji, J., Xu, J., Lin, L., et al. (2019). Rhein Suppresses Lung Inflammatory Injury Induced by Human Respiratory Syncytial Virus Through Inhibiting NLRP3 Inflammasome Activation via NF- κ B Pathway in Mice. *Front. Pharmacol.* 10:1600. doi: 10.3389/fphar.2019.01600
- Shi, H., Wang, X., Lu, Z., Zhao, B. S., Ma, H., Hsu, P. J., et al. (2017). YTHDF3 facilitates translation and decay of N(6)-methyladenosine-modified RNA. *Cell Res.* 27, 315–328. doi: 10.1038/cr.2017.15
- Shima, H., Matsumoto, M., Ishigami, Y., Ebina, M., Muto, A., Sato, Y., et al. (2017). S-Adenosylmethionine Synthesis Is Regulated by Selective N(6)-Adenosine Methylation and mRNA Degradation Involving METTL16 and YTHDC1. *Cell Rep.* 21, 3354–3363. doi: 10.1016/j.celrep.2017.11.092
- Shriwas, O., Mohapatra, P., Mohanty, S., and Dash, R. (2020). The Impact of m6A RNA Modification in Therapy Resistance of Cancer: Implication in Chemotherapy, Radiotherapy, and Immunotherapy. *Front. Oncol.* 10:612337. doi: 10.3389/fonc.2020.612337
- Singh, R., Mishra, M. K., and Aggarwal, H. (2017). Inflammation, Immunity, and Cancer. *Mediat. Inflamm.* 2017:6027305. doi: 10.1155/2017/6027305
- Sommer, F., and Backhed, F. (2016). Know your neighbor: Microbiota and host epithelial cells interact locally to control intestinal function and physiology. *BioEssays* 38, 455–464. doi: 10.1002/bies.201500151
- Stolte, E. H., Chadzinska, M., Przybylska, D., Flik, G., Savelkoul, H. F., and Verburg-van Kemenade, B. M. (2009). The immune response differentially regulates Hsp70 and glucocorticoid receptor expression in vitro and in vivo in common carp (*Cyprinus carpio* L.). *Fish Shellf. Immunol.* 27, 9–16. doi: 10.1016/j.fsi.2008.11.003
- Sun, L., Ma, L., Zhang, H., Cao, Y., Wang, C., Hou, N., et al. (2019). Fto Deficiency Reduces Anxiety- and Depression-Like Behaviors in Mice via Alterations in Gut Microbiota. *Theranostics* 9, 721–733. doi: 10.7150/thno.31562
- Sun, X., Luo, W., Tan, X., Li, Q., Zhao, Y., Zhong, W., et al. (2013). Increased plasma corticosterone contributes to the development of alcoholic fatty liver in mice. *Am. J. Physiol. Gastr. Liver Physiol.* 305, G849–G861. doi: 10.1152/ajpgi.00139.2013
- Taketo, K., Konno, M., Asai, A., Koseki, J., Toratani, M., Satoh, T., et al. (2018). The epitranscriptome m6A writer METTL3 promotes chemo- and radioresistance in pancreatic cancer cells. *Int. J. Oncol.* 52, 621–629. doi: 10.3892/ijo.2017.4219
- Tang, L., Wei, X., Li, T., Chen, Y., Dai, Z., Lu, C., et al. (2021). Emerging Perspectives of RNA N (6)-methyladenosine (m(6)A) Modification on Immunity and Autoimmune Diseases. *Front. Immunol.* 12:630358. doi: 10.3389/fimmu.2021.630358
- Tu, S., Bhagat, G., Cui, G., Takaishi, S., Kurt-Jones, E. A., Rickman, B., et al. (2008). Overexpression of interleukin-1 β induces gastric inflammation and cancer and mobilizes myeloid-derived suppressor cells in mice. *Cancer Cell.* 14, 408–419. doi: 10.1016/j.ccr.2008.10.011
- Ueda, Y., Ooshio, I., Fusamae, Y., Kitae, K., Kawaguchi, M., Jingushi, K., et al. (2017). AlkB homolog 3-mediated tRNA demethylation promotes protein synthesis in cancer cells. *Sci. Rep.* 7:42271. doi: 10.1038/srep42271
- Varela, M. L., Mogildea, M., Moreno, I., and Lopes, A. (2018). Acute Inflammation and Metabolism. *Inflammation* 41, 1115–1127. doi: 10.1007/s10753-018-0739-1
- Visvanathan, A., Patil, V., Arora, A., Hegde, A. S., Arivazhagan, A., Santosh, V., et al. (2018). Essential role of METTL3-mediated m(6)A modification in glioma stem-like cells maintenance and radioresistance. *Oncogene* 37, 522–533. doi: 10.1038/onc.2017.351
- Wang, J., Yan, S., Lu, H., Wang, S., and Xu, D. (2019). METTL3 Attenuates LPS-Induced Inflammatory Response in Macrophages via NF-. *Med. Inflamm.* 2019, 3120391. doi: 10.1155/2019/3120391
- Wang, P., Doxtader, K. A., and Nam, Y. (2016). Structural Basis for Cooperative Function of Mettl3 and Mettl14 Methyltransferases. *Mol. Cell.* 63, 306–317. doi: 10.1016/j.molcel.2016.05.041
- Wang, S., Chim, B., Su, Y., Khil, P., Wong, M., Wang, X., et al. (2019). Enhancement of LIN28B-induced hematopoietic reprogramming by IGF2BP3. *Genes Dev.* 33, 1048–1068. doi: 10.1101/gad.325100.119
- Wang, X., Zhao, B. S., Roundtree, I. A., Lu, Z., Han, D., Ma, H., et al. (2015). N(6)-methyladenosine Modulates Messenger RNA Translation Efficiency. *Cell* 161, 1388–1399. doi: 10.1016/j.cell.2015.05.014
- Warda, A. S., Kretschmer, J., Hackert, P., Lenz, C., Urlaub, H., Höbartner, C., et al. (2017). Human METTL16 is a N(6)-methyladenosine (m(6)A) methyltransferase that targets pre-mRNAs and various non-coding RNAs. *EMBO Rep.* 18, 2004–2014. doi: 10.15252/embr.201744940
- Wei, C.-M., Gershowitz, A., and Moss, B. (1975). Methylated nucleotides block 5' terminus of HeLa cell messenger RNA. *Cell* 4, 379–386. doi: 10.1016/0092-8674(75)90158-0
- Wei, J., Liu, F., Lu, Z., Fei, Q., Ai, Y., He, P. C., et al. (2018). Differential m(6)A, m(6)A(m), and m(1)A Demethylation Mediated by FTO in the Cell Nucleus and Cytoplasm. *Mol. Cell* 71, 973–985.e. doi: 10.1016/j.molcel.2018.08.011
- Welsh, P., Polisecki, E., Robertson, M., Jahn, S., Buckley, B. M., de Craen, A. J., et al. (2010). Unraveling the directional link between adiposity and inflammation: a bidirectional Mendelian randomization approach. *J. Clin. Endocrinol. Metabol.* 95, 93–99. doi: 10.1210/jc.2009-1064
- Weng, H., Huang, H., Wu, H., Qin, X., Zhao, B. S., Dong, L., et al. (2018). METTL14 Inhibits Hematopoietic Stem/Progenitor Differentiation and Promotes Leukemogenesis via mRNA m(6)A Modification. *Cell. Stem Cell.* 22, 191–205.e. doi: 10.1016/j.stem.2017.11.016
- Whiteside, T. L. (2018). Exosome and mesenchymal stem cell cross-talk in the tumor microenvironment. *Sem. Immunol.* 35, 69–79. doi: 10.1016/j.smim.2017.12.003
- Winkler, R., Gillis, E., Lasman, L., Safra, M., Geula, S., Soyris, C., et al. (2019). m(6)A modification controls the innate immune response to infection by targeting type I interferons. *Nat. Immunol.* 20, 173–182. doi: 10.1038/s41590-018-0275-z
- Wu, J., Wei, Z., Cheng, P., Qian, C., Xu, F., Yang, Y., et al. (2020). Rhein modulates host purine metabolism in intestine through gut microbiota and ameliorates experimental colitis. *Theranostics* 10, 10665–10679. doi: 10.7150/thno.43528
- Xin, X., Wang, C., Lin, Z., Xu, J., Lu, Y., Meng, Q., et al. (2018). Inflammatory-Related P62 Triggers Malignant Transformation of Mesenchymal Stem Cells through the Cascade of CUDR-CTCF-IGFII-RAS Signaling. *Mol. Ther. Nucleic Acids* 11, 367–381. doi: 10.1016/j.omtn.2018.03.002
- Xiong, X., Yi, C., and Peng, J. (2017). Epitranscriptomics: Toward A Better Understanding of RNA Modifications. *Genom. Proteom. Bioinform.* 15, 147–153. doi: 10.1016/j.gpb.2017.03.003
- Yang, S., Wei, J., Cui, Y. H., Park, G., Shah, P., Deng, Y., et al. (2019). m(6)A mRNA demethylase FTO regulates melanoma tumorigenicity and response to anti-PD-1 blockade. *Nat. Commun.* 10:2782. doi: 10.1038/s41467-019-10669-0
- Yi, L., Wu, G., Guo, L., Zou, X., and Huang, P. (2020). Comprehensive Analysis of the PD-L1 and Immune Infiltrates of m(6)A RNA Methylation Regulators in Head and Neck Squamous Cell Carcinoma. *Mol. Ther. Nucleic Acids* 21, 299–314. doi: 10.1016/j.omtn.2020.06.001
- Yu, R., Li, Q., Feng, Z., Cai, L., and Xu, Q. (2019). m6A Reader YTHDF2 Regulates LPS-Induced Inflammatory Response. *Int. J. Mol. Sci.* 20:1323. doi: 10.3390/ijms20061323
- Zaccara, S., and Jaffrey, S. R. (2020). A Unified Model for the Function of YTHDF Proteins in Regulating m(6)A-Modified mRNA. *Cell* 181, 1582.e–1595.e. doi: 10.1016/j.cell.2020.05.012
- Zhang, B., Wu, Q., Li, B., Wang, D., Wang, L., and Zhou, Y. L. (2020). m(6)A regulator-mediated methylation modification patterns and tumor microenvironment infiltration characterization in gastric cancer. *Mol. Cancer* 19:53. doi: 10.1186/s12943-020-01170-0
- Zhang, Y., Gu, X., Li, D., Cai, L., and Xu, Q. (2019). METTL3 Regulates Osteoblast Differentiation and Inflammatory Response via Smad Signaling and MAPK Signaling. *Int. J. Mol. Sci.* 21:199. doi: 10.3390/ijms21010199
- Zhang, Y., Guo, F., and Zhao, R. (2016). Hepatic expression of FTO and fatty acid metabolic genes changes in response to lipopolysaccharide with alterations in m(6)A modification of relevant mRNAs in the chicken. *Br. Poultry Sci.* 57, 628–635. doi: 10.1080/00071668.2016.1201199
- Zhao, H., Lv, Y., Li, L., and Chen, Z. J. (2016). Genetic Studies on Polycystic Ovary Syndrome. *Best Pract. Res. Clin. Obstetr. Gynaecol.* 37, 56–65. doi: 10.1016/j.bpobgyn.2016.04.002
- Zhao, W., Qi, X., Liu, L., Ma, S., Liu, J., and Wu, J. (2020). Epigenetic Regulation of m(6)A Modifications in Human Cancer. *Mol. Ther. Nucleic Acids* 19, 405–412. doi: 10.1016/j.omtn.2019.11.022
- Zheng, G., Dahl, J. A., Niu, Y., Fedorcsak, P., Huang, C.-M., Li, C. J., et al. (2013). ALKBH5 is a mammalian RNA demethylase that impacts RNA metabolism and mouse fertility. *Mol. Cell.* 49, 18–29. doi: 10.1016/j.molcel.2012.10.015

- Zhou, J., Wan, J., Gao, X., Zhang, X., Jaffrey, S. R., and Qian, S.-B. (2015). Dynamic m(6)A mRNA methylation directs translational control of heat shock response. *Nature* 526, 591–594. doi: 10.1038/nature15377
- Zhou, Y., Hambly, B. D., and McLachlan, C. S. F. T. O. (2017). associations with obesity and telomere length. *J. Biomed. Sci.* 24:65. doi: 10.1186/s12929-017-0372-6
- Zimmermann, E., Skogstrand, K., Hougaard, D. M., Astrup, A., Hansen, T., Pedersen, O., et al. (2011). Influences of the common FTO rs9939609 variant on inflammatory markers throughout a broad range of body mass index. *PLoS One* 6:e15958. doi: 10.1371/journal.pone.0015958
- Zong, X., Zhao, J., Wang, H., Lu, Z., Wang, F., Du, H., et al. (2019). Mettl3 Deficiency Sustains Long-Chain Fatty Acid Absorption through Suppressing Traf6-Dependent Inflammation Response. *J. Immunol.* 202, 567–578. doi: 10.4049/jimmunol.1801151
- 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 Luo, Xu and Sun. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



The MicroRNA Family Gets Wider: The IsomiRs Classification and Role

Luisa Tomasello*, Rosario Distefano, Giovanni Nigita and Carlo M. Croce

Department of Cancer Biology and Genetics, Comprehensive Cancer Center, The Ohio State University, Columbus, OH, United States

OPEN ACCESS

Edited by:

Giovanni Cenci,
Sapienza University of Rome, Italy

Reviewed by:

Silvio Zaina,
University of Guanajuato, Mexico
Alessandra Tessitore,
University of L'Aquila, Italy

*Correspondence:

Luisa Tomasello
luisa.tomasello@osumc.edu

Specialty section:

This article was submitted to
Epigenomics and Epigenetics,
a section of the journal
Frontiers in Cell and Developmental
Biology

Received: 16 February 2021

Accepted: 29 April 2021

Published: 09 June 2021

Citation:

Tomasello L, Distefano R, Nigita G
and Croce CM (2021) The MicroRNA
Family Gets Wider: The IsomiRs
Classification and Role.
Front. Cell Dev. Biol. 9:668648.
doi: 10.3389/fcell.2021.668648

MicroRNAs (miRNAs or miRs) are the most characterized class of non-coding RNAs and are engaged in many cellular processes, including cell differentiation, development, and homeostasis. MicroRNA dysregulation was observed in several diseases, cancer included. Epitranscriptomics is a branch of epigenomics that embraces all RNA modifications occurring after DNA transcription and RNA synthesis and involving coding and non-coding RNAs. The development of new high-throughput technologies, especially deep RNA sequencing, has facilitated the discovery of miRNA isoforms (named isomiRs) resulting from RNA modifications mediated by enzymes, such as deaminases and exonucleases, and differing from the canonical ones in length, sequence, or both. In this review, we summarize the distinct classes of isomiRs, their regulation and biogenesis, and the active role of these newly discovered molecules in cancer and other diseases.

Keywords: microRNA, variants, isomiRs, microRNA biogenesis, novel microRNA in cancer

INTRODUCTION

MicroRNAs are small non-coding RNA observed for the first time in the early 1990s (Lee et al., 1993) and characterized as a class of functional molecules in *Caenorhabditis elegans*, 10 years later (Reinhart et al., 2000; Lagos-Quintana, 2001). The discovery of microRNAs emphasizes the role of RNA as a functional molecule regulating gene expression at the post-transcriptional level (Huntzinger and Izaurralde, 2011). More than 2,000 (2,654, according to miRBase v22) mature microRNAs have been discovered in *Homo sapiens* (Kozomara et al., 2019). Several studies have elucidated the relevance of these molecules in regulating cellular processes and their steady presence in physiological and disease-related pathways (Friedman et al., 2008).

The microRNA maturation is a multi-step processing event that starts in the nucleus. The RNase III DROSHA, in connection with the Microprocessor complex subunit DGCR8 (DiGeorge syndrome critical region 8), cleaves a primary RNA-transcript into a stem-loop precursor of approximately 70 nucleotides, called pre-miRNA. This RNA-hairpin product is carried into the cytoplasm by Exportin 5 (XPO5) and processed by the RNase III DICER into the mature microRNA (Ha and Kim, 2014).

Once the microRNA biogenesis is complete, the single-strand mature molecule is loaded by the RISC (RNA-induced silencing) complex. The mRNA target recognition occurs through the binding between the short seed region at the 5' of the microRNA (nucleotides 2–8) and a partially or perfectly complementary region on the target gene 3' UTR (Ha and Kim, 2014).

At least 45,000 sequences matching with microRNA seed sequences, the miRNA responsive elements (MRE), were found in 3' UTR of human protein-coding genes (Friedman et al., 2008), indicating that these small RNAs could regulate most of the human proteins (Friedman et al., 2008). Due to the seed-sequence brevity, it is conceivable to predict more than one target for each microRNA. Indeed, hundreds of mRNAs could be controlled by a single microRNA (Friedman et al., 2008).

Epitranscriptomics is the study of post-synthetic modifications involving the RNA chemical structure (Frye et al., 2016). These changes, mediated by a wide range of proteins, including but not limited to RNA-methyltransferases, deaminases, uridylyltransferases, poly(A) RNA polymerases, and exonucleases (De Almeida et al., 2018; Lan et al., 2019; Yu and Kim, 2020), also occur on microRNAs (Alarcón et al., 2015; Nishikura, 2016; Gutiérrez-Vázquez et al., 2017) and could be responsible for their sequence and length changes.

The rise of the high-throughput technology next-generation sequencing (NGS) has recently allowed several novel microRNAs to be detected alongside the well-known sequences. At first, these new molecules were interpreted as sequencing/mapping errors. However, later on, it was widely demonstrated that the percentage of *non-templated* nucleotide additions (%NTA) observed in small RNA sequencing data was significantly higher than the expected rate of sequencing error-rate calculated using small artificial RNAs (Linsen et al., 2009; Wyman et al., 2011). The development of more advanced analysis algorithms has supported these studies in confirming that canonical microRNA sequence modifications are not experimental artifacts but physiological events occurring *in vivo* (Linsen et al., 2009; Wyman et al., 2011). Moreover, there is evidence that isomiRs have a functional role just as their related canonical fragments: microRNA isomers can bind Argonaute (Ago) proteins, as demonstrated by co-immunoprecipitation assay (Cloonan et al., 2011; Londin et al., 2015; Haseeb et al., 2017), and can inhibit the expression of specific targets, as shown by luciferase assay *in vitro* (Cloonan et al., 2011).

In 2015, Londin et al. (2015) identified 3,707 novel microRNAs examining 1,323 samples from 13 different human tissues. The data presented on these newly discovered molecules suggested that, as the canonical microRNAs, the novel isoforms have a tissue-dependent expression (Londin et al., 2015). Their genome distribution is mostly intergenic (57.6%) and intronic (17.4%); moreover, out of the 31 miRNA genomic clusters identified by the authors, 21 involved novel variants, further proving a similar genomic organization with the canonical counterpart (Londin et al., 2015).

MicroRNA isoforms are heterogeneous and can variate for length, sequence, or both. The sequence variants hold more or fewer nucleotides at 5' or a 3' end than the canonical ones. Concurrently, the polymorphic (internal) isomiRs include

different nucleotides within the mature sequence that distinguish these isoforms from the database-annotated microRNAs (Wu et al., 2018).

A recent classification categorizes the microRNAs and their variants into five classes (Figure 1):

- (a) canonical microRNAs, whose mature sequence is the one reported in the microRNA databases;
- (b) 5' isomiRs, with changes in length at the 5' end;
- (c) 3' isomiRs, with changes in length at the 3' end;
- (d) polymorphic isomiRs, with identical length except for changes within the mature sequence, between the first and the last nucleotide; and
- (e) mixed type isomiRs, with changes in length and sequence (Wu et al., 2018).

Small variations in length and sequence of mature microRNA could be responsible for seed modification, potentially resulting in targetome shifting. This molecular event implies that isomiRs could have distinct or divergent functions compared with their related canonical counterparts. IsomiR expression varies among different tissue and cancer types, demonstrating their functional peculiarity and potential role as biomarkers (Telonis et al., 2017). Given the critical gene-regulatory function of these small molecules, their diffuse expression, and their involvement in the control of cellular processes, it is crucial to acquire a comprehensive knowledge of all the microRNAs and their functional isoforms expressed.

5' AND 3' ISOMIRS

MicroRNAs with modifications of the sequence are called 5' or 3' isomiRs, depending on which microRNA end shifts. The 5' isomiR rate is significantly lower than the 3' isomiR one: 5–15% compared with 40–50% (Tan et al., 2014), even if the low percentage of 5' variants can be offset by a high expression of these new isoforms and still have a relevant impact on the regulation of shared or exclusive targets (Chiang et al., 2010). It would be logical to assume that variations occurring at the 5' end of the mature microRNA, which could reasonably affect the seed sequence, should weigh more on the potential targetome shifting than variations involving the 3' end. Nevertheless, it was proven that the pairing between the microRNA 3' end and its target firmly contributes to the interaction stability, maintaining favorable total interaction energy. The microRNA 3' end plays a compensatory role when the presence of mismatches or bubbles between the mRNA target and the microRNA-seed region makes the binding weak (Bail et al., 2010; Moore et al., 2015).

Variations in length could be the consequence of DROSHA and DICER imprecise cleavage during the microRNA biogenesis steps or the action of specific exonucleases that remove nucleotides at its extremities, making the microRNA shorter (Neilsen et al., 2012). In both cases, the resulting isomiRs are classified as *templated* because their sequences match the parental gene (Neilsen et al., 2012). The length differences can also be attributed to the post-transcriptional addition of few nucleotides at the 5' or 3' end of the mature sequence by nucleotidyl

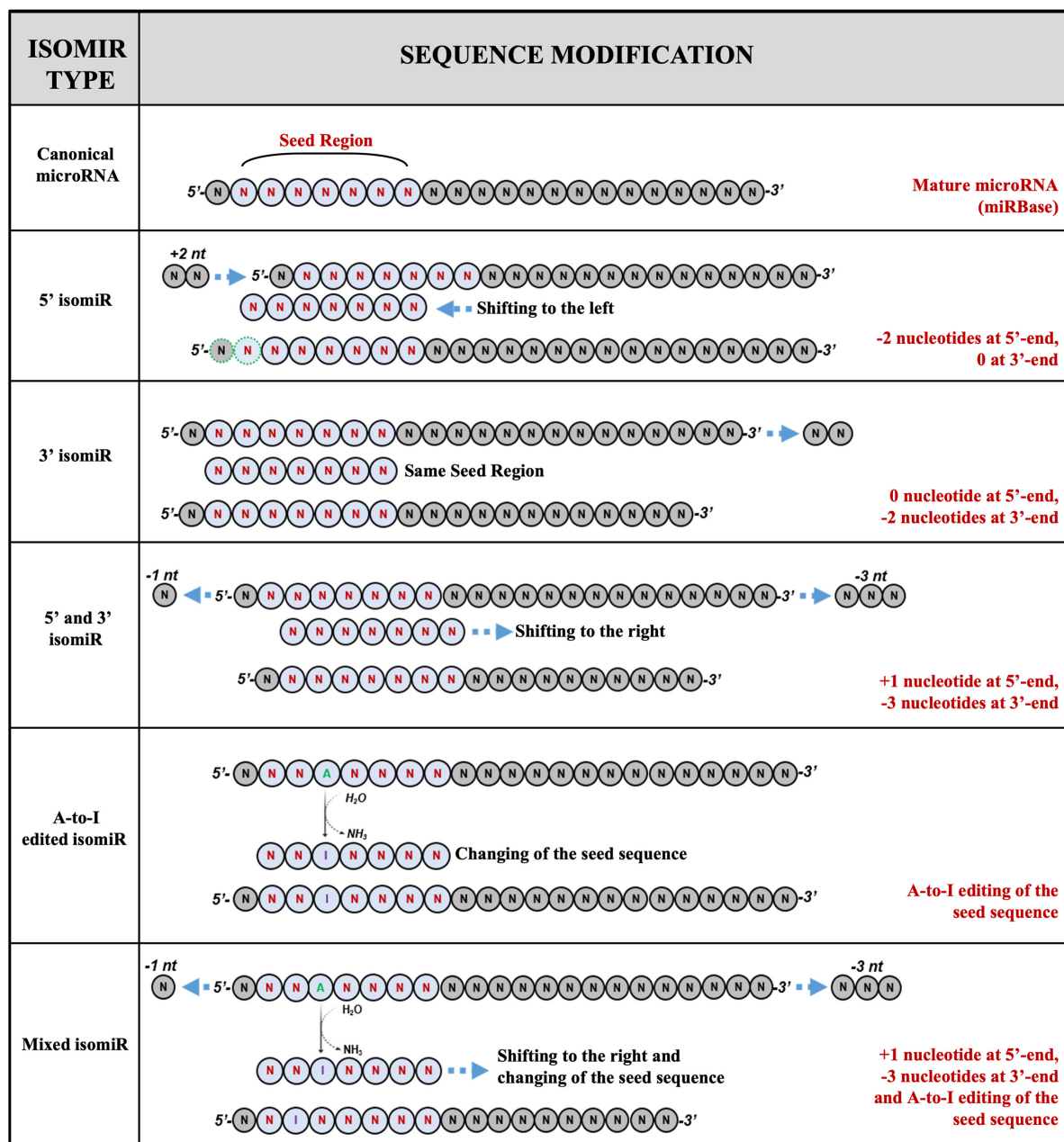


FIGURE 1 | Examples of isomiRs. MicroRNA isoforms can variate for length, sequence, or both. The current classification identified five classes of variants: (1) canonical microRNAs; (2) 5' isomiRs; (3) 3' isomiRs; (4) polymorphic isomiRs; (5) mixed type isomiRs.

transferases (Wyman et al., 2011). These variants are considered *non-templated* because they contain nucleotides not existing in the parental gene sequence (Nielsen et al., 2012).

DROSHA and DICER Alternative Cleavage: From One Pri-miRNA Gene to Several MicroRNA Variants

The biogenesis of microRNAs starts with transcribing a primary structure (pri-miRNA) by RNA polymerase II. The pri-miRNA

consists of a terminal loop, an upper and a lower stem surrounded by two basal single-strand flanking sequences (Ha and Kim, 2014). The RNase III DROSHA (Table 1), aided by DGCR8, processes this molecule in the nucleus and produces the first cut in correspondence of the 5' end of the 5p arm and 3' end of the 3p arm (Ha and Kim, 2014). Han et al. (2006) described how the DROSHA cleavage is always expected to occur 11 bp far away from the junction between the stem and the basal unpaired sequences (ssRNA/dsRNA junction). The precision of this phenomenon induced to hypothesize that DGCR8 could

TABLE 1 | Enzymes affecting microRNA length and sequence.

Name	Type	References
DROSHA	Ribonuclease (RNase) III double-stranded RNA-specific	Han et al., 2006; Starega-Roslan et al., 2015; Boffill-De Ros et al., 2019
DICER1	Ribonuclease (RNase) III double-stranded RNA-specific	MacRae et al., 2007; Gu et al., 2012; Ha and Kim, 2014; Starega-Roslan et al., 2015; Song and Rossi, 2017
Nibbler	3'–5' exonuclease	Han et al., 2011; Liu et al., 2011; Wyman et al., 2011
PARN	3'-exonucleases with a preference for poly(A) substrates	Katoh et al., 2015
TENT2 or PAPD4 or GLD2	Poly(A) RNA polymerase	Katoh et al., 2009; Burroughs et al., 2010; Wyman et al., 2011
TUT4 or ZCCHC11	RNA uridylyltransferase	Jones et al., 2009; Wyman et al., 2011; Thornton et al., 2014; Yang et al., 2020
TUT3 or PAPD5	Poly(A) RNA polymerase	Wyman et al., 2011
MTPAP or TENT6	Mitochondrial poly(A) polymerase	Wyman et al., 2011
PAPOLG	Poly(A) DNA/RNA polymerase	Katoh et al., 2009
TUT1 or TENT1	Terminal uridylyltransferase and nuclear poly(A) polymerase	Wyman et al., 2011
TUT7 or ZCCHC6 or PAD6	Terminal uridylyltransferase	Wyman et al., 2011; Thornton et al., 2014; Yang et al., 2020
ADAR enzymes	Adenosine deaminase RNA specific	Bazak et al., 2014; GTEx Consortium et al., 2017; Wang et al., 2017; Cesarini et al., 2018; Li et al., 2018; Xu et al., 2019; Marceca et al., 2020
APOBEC enzymes	Cytidine deaminase RNA specific	Marceca et al., 2020

act as a “molecular meter,” recognizing and anchoring the pri-miRNA substrate, forming the “pre-cleavage complex,” and preparing the way for DROSHA-mediated catalysis (Han et al., 2006; **Figure 2A**). Making mutant artificial pri-miRNA-30a with modified regions, they demonstrated that the terminal loop does not affect the cut because it weakly interacts with DGCR8 protein. However, modifications in this pri-miRNA area, especially in the loop size, could compromise the catalysis efficiency (Han et al., 2006). Besides, alterations of the stem length and the single-stranded basal segments could undermine the cleavage site recognition from DGCR8, leading to imprecise processing of the pri-miRNA (Han et al., 2006).

Bofill-De Ros et al. (2019) explained how the 3D structural characteristics of pri-miRNA affect the DROSHA cleavage ambiguity (**Figure 2B**). The employment of miR-9 paralogs showed that the pri-miRNA lower-stem flexibility and distortion could play a central role in driving DROSHA cleavage, potentially destabilizing the fidelity of the cut (Bofill-De Ros et al., 2019).

After the pre-miRNA exportation from the nucleus to the cytoplasm by Exportin-5 (XPO5), another RNase III, named DICER (**Table 1**), processes the short hairpin RNA (shRNA) by eliminating the terminal loop and forming a double-strand miRNA/miRNA* (Ha and Kim, 2014). Similarly to the first catalysis, the cleavage precision is essential to generate a specific mature molecule. Indeed, the cut inaccuracy could generate new microRNA variants with altered seed sequences and, reasonably, different targets and roles (Tan et al., 2014).

DICER is an RNase III enzyme holding eight different domains, including an amino-terminal helicase domain, a PAZ (Piwi/Argonaute/Zwille) domain, and 2 RNase III domains (Song and Rossi, 2017). DICER-mediated pre-miRNA catalysis starts

with recognizing the open ends of pre-miRNA and trapping the RNA molecules inside the enzyme catalytic pocket. At this step, the PAZ domain is essential to “measure” the dsRNA from the 3' end of the shRNA to ensure the generation of a mature microRNA duplex with a species-specific length and the typical 2-nucleotides 3' overhang (MacRae et al., 2007).

Gu et al. (2012) employed artificial shRNAs to describe the DICER processing of pre-miRNA. They established a “loop-counting rule” to predict the accuracy of the cut: DICER cleavage fidelity can be maintained if the enzyme recognizes an ssRNA sequence, such as the terminal loop or an internal bulge, precisely situated two nucleotides far away from the cleavage site, previously determined by the PAZ domain “measuring.” In other words, the presence of a single-stranded structure in correspondence with the enzyme helicase domain is required to stabilize the catalytic RNase III domain, thus supporting the correct cleavage (Gu et al., 2012; **Figure 2C**).

The fidelity of DROSHA and DICER cleavage is influenced not only by the pri-miRNA and pre-miRNA structures but also by their sequences. Deep sequencing data on the human cell line HEK293T, embryonic stem cells, and differentiated cells from murine models showed that DROSHA and DICER cleavage sites seldom include G residues on their sequences. Moreover, data highlighted a strong presence of U residues at both the mature microRNA ends (Starega-Roslan et al., 2015).

Although the study shows that the DROSHA cut fidelity seems more influenced by the cleavage site sequence than DICER, it was demonstrated that both the enzymes undergo an adequate sequence-dependent regulation that affects the precision of the cut, involving their RNase domains differently. The DROSHA RIIIA domain produces more heterogeneous molecules than

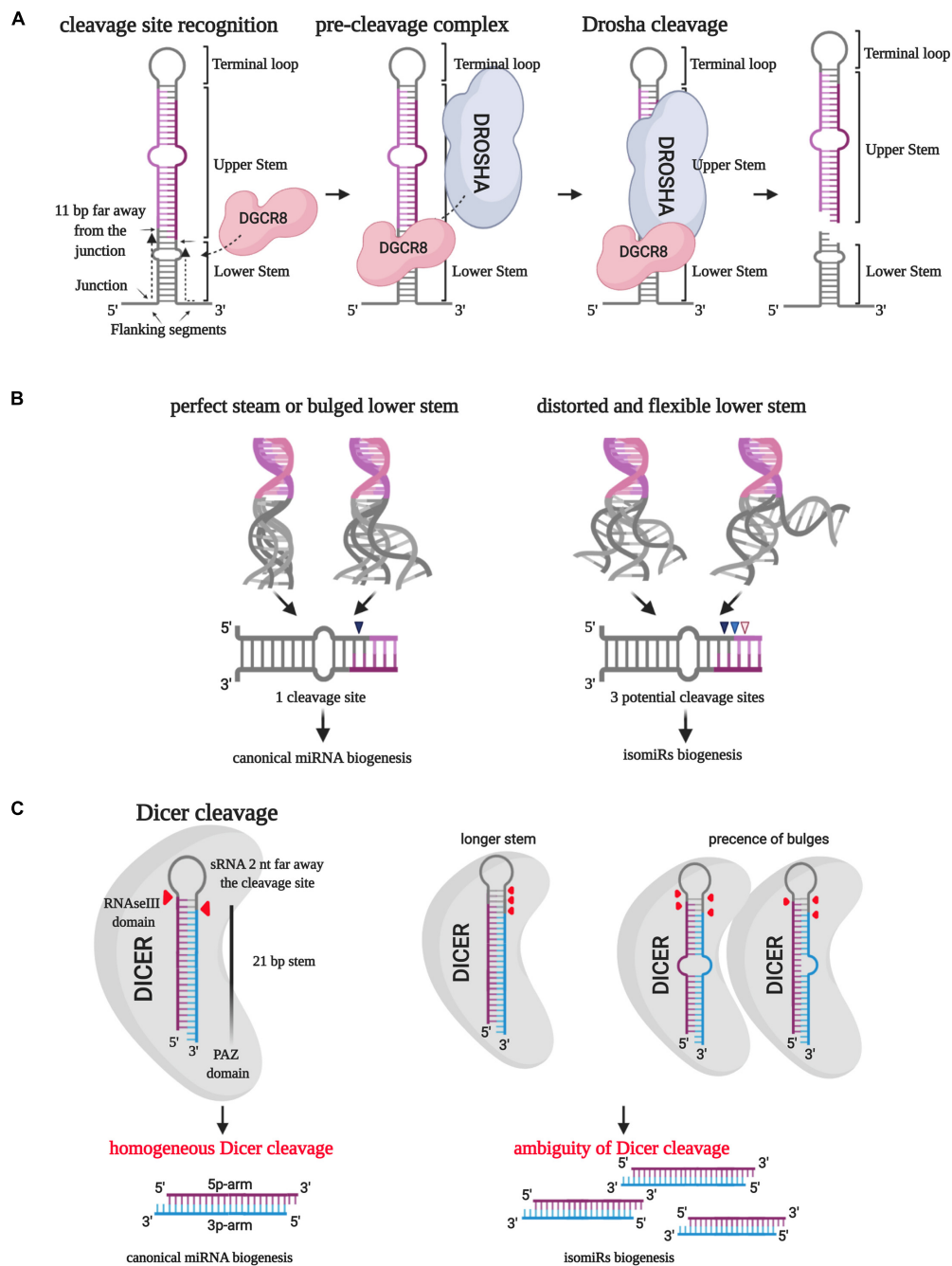


FIGURE 2 | IsomiR biogenesis. **(A)** The RNase III DROSHA, aided by DGCR8, processes pri-miRNA in the nucleus and produces the first cut in correspondence of the 5' end of the 5p arm and 3' end of the 3p arm. DGCR8 acts as a molecular meter and identifies the cleavage site 11 bp far away from the junction point between the lower stem and the basal unpaired sequences. **(B)** The secondary structure of the lower stem of pri-miRNA affects the DROSHA cleavage precision: a perfect or bulged lower stem leads to a homogeneous cleavage site in more than 97% of the cases. On the contrary, a distorted and flexible lower stem creates three potential cleavage sites. **(C)** The RNase III DICER processes the short hairpin RNA (shRNA) by eliminating the terminal loop and forming a double-strand miRNA/miRNA*. Different lengths and the presence of bulges can affect the PAZ domain-mediated "measurement" of the lower stem leading to the selection of multiple cleavage sites.

the DICER RIIIA domain, which cuts more precisely. On the contrary, the DROSHA RIIIB domain catalysis activity is much more specific than the DICER RIIIB domain one (Starega-Roslan et al., 2015).

Summing up, even as we are still used to considering isomiRs generation as the exception to the rule, this is a misconception: it is very infrequent that DROSHA and DICER cleavage produces only one microRNA variant from a single microRNA gene.

IsomiRs Can Arise From Exoribonuclease Nibbling Activity

The trimming action of exoribonucleases (**Table 1**) could also be a source of microRNA variants (Han et al., 2011; Katoh et al., 2015). These enzymes act on mature microRNA ends or microRNA precursors during microRNA biogenesis processes. One of the first indications of this mechanism was observed in *Drosophila melanogaster*, where the 3′–5′ exoribonuclease Nibbler (Nbr) (**Table 1**) contributes to generating a 22-nt-long microRNA after the processing mediated by DICER (Han et al., 2011). Han et al. (2011) studied the case of miR-34 in flies: the maturation of this microRNA runs through the typical multi-step DROSHA/DICER biogenesis process. DICER can generate molecules of 22 or 24 nt, and Ago1 or Ago2 can load both of them. In the former case, the RISC complex constitution led to the post-transcriptional regulation of microRNA targets. In the latter case, the longer molecules bound to Ago1 is available for the Nibbler trimming because of the weaker binding with the Ago1 PAZ domain. The sculpt of the 3′ end and the restoration of a 22-nt-long molecule enhances the activity of miR-34 (Han et al., 2011). *Nibbler* knockout causes the loss of many 3′ isomiRs and a semi-lethal and sterile phenotype in flies (Han et al., 2011; Liu et al., 2011).

Experiments performed on the human cervical carcinoma cell line HeLa have shown a microRNA 3′ variability tracing the one observed in *D. melanogaster*, thus suggesting the presence of a human exoribonuclease homolog of Nibbler (Han et al., 2011).

Katoh et al. (2015) investigated the role of another exoribonuclease named PARN, which interacts with microRNAs, specifically miR-122, in hepatocellular carcinoma cells. CUGBP1, a protein binding UG-rich microRNAs, recruits PARN and leads it to the miR-122. PARN causes deadenylation, with consequent destabilization of the miR-122 3′ end, affecting the cellular level of canonical miR-122 (Katoh et al., 2015).

To date, the definition of the role of exoribonucleases in the human isomiR generation is still at the beginning. Nevertheless, the evidence collected so far suggests the likely presence of mammalian homologs with an active role in isomiRs biogenesis and regulation of mature microRNA stability.

Non-templated microRNA Variant Generation by Nucleotidyl Transferases

The post-transcriptional addition of nucleotides to small RNA 3′ end contributes to the heterogeneity of microRNAs and the generation of new variants. Through next-generation small RNA sequencing experiments, Wyman et al. (2011) defined 39 microRNA modifications ascribable to 3′ nucleotides addition. These modifications are physiological and influenced by biological processes, such as cell differentiation (Berezikov et al., 2006; Wyman et al., 2011). They were recognized in a broad range of species and cell types, in different diseases and biological conditions (Berezikov et al., 2006). The most prevalent modifications identified are adenylation and uridylation. In human and mouse, ~50% of 3′ modifications are mono-adenylation, and ~25% are mono-uridylation (Wyman et al., 2011). The nucleotide additions, mediated by at least

eight nucleotidyl transferases, affect microRNA stability and efficiency undergoing their modification process (Jones et al., 2009; Katoh et al., 2009; Burroughs et al., 2010; Wyman et al., 2011). Typically, microRNA uridylation is associated with molecule degradation, whereas adenylation leads to improved microRNA stability (Rüegger and Großhans, 2012). The principal nucleotidyl transferases identified so far are PAPD4 (TENT2 or GLD2), ZCCHC11 (TUT4), PAPD5 (TUT3), MTPAP (TENT6), PAPOLG, TUT1 (TENT1), and ZCCHC6 (TUT7) (Jones et al., 2009; Katoh et al., 2009; Burroughs et al., 2010; Wyman et al., 2011; **Table 1**). The downregulation of these enzymes contributes to a specific decreased number of microRNA 3′ end modifications. For example, the depletion of TUT1 and ZCCHC6 causes the selective loss of the 3′ U variant of the miR-200a and let-7e, respectively (Wyman et al., 2011). Zcchc11 (TUT4) and Zcchc6 (TUT7) modify, through 3′ uridylation, a specific microRNA group that shares a TUTase recognition sequence motif and targets proteins belonging to the Homeobox family in P19 embryonal carcinoma cells from mouse (Thornton et al., 2014). Thornton et al. (2014) have further demonstrated the importance of these proteins during zebrafish development steps through the regulation of Homeobox proteins, emphasizing that microRNA uridylation is a physiological and finely regulated process. Katoh et al. (2009) described a delicate mechanism, orchestrated by the nucleotidyl transferase GLD2 and exonuclease enzymes. The process stabilizes the microRNA-122 molecule, with specific liver-associated functions in hepatocellular carcinoma cells. After the canonical biogenesis process, the 22-nt variant of miR-122 is stretched at the 3′ end by GLD-2-mediated poly(A) adenylation. Later, this longer variant undergoes cleaving by 5′–3′ exonucleases that restore a molecule long between 21 and 23 nucleotides. This elongation/degradation process “corrects” the microRNA length to produce a stable molecule, not too long, not too short, that can be loaded by Ago2 (Katoh et al., 2009). Similarly, other microRNAs, including but not limited to miR-7, miR-222, and miR-769, are subjected to uridylation by TUT4 and TUT7 when the binding with Ago2 leaves their 3′ end exposed (Yang et al., 2020). Then, oligouridylation microRNAs undergo degradation by exonuclease DIS3L2 (Yang et al., 2020).

These represent a few examples that can describe how the addition of *non-templated* nucleotides regulates the stability or degradation of microRNAs, generating new isomiRs and indirectly changing their mRNA targets expression pattern.

POLYMORPHIC ISOMIRS

SNPs in microRNAs

The frequency of single-nucleotide polymorphisms (SNPs) occurring within microRNA genes is consistently lower than observed in other genomic regions (Saunders et al., 2007). The selective evolutionary pressure on microRNA sequences deters genetic variations on microRNA loci, supporting the conservation of these regions and their functional importance. The SNPs’ density observed in microRNA seed sequences is less than 1% of the total SNPs in the human genome

(Saunders et al., 2007). Despite the rareness of this event, genetic variations in precursor or mature molecule sequences significantly impact microRNAs transcription and biogenesis. The presence of SNPs could block or enhance some microRNA maturation, thus changing canonical and new microRNA variants expression patterns and influencing their gene-silencing regulation (Sun et al., 2009). To cite a few examples, Calin et al. (2005) demonstrated that a germline-specific mutation on the pri-miR-15a/16-1 impairs the tumor suppressor miR-16-1 biogenesis, thus increasing the risk of familial chronic lymphocytic leukemia.

Similarly, the SNP rs895819 on the terminal loop of pre-miR-27a blocks the derived microRNA maturation, whose down-regulation is associated with an attenuating risk of familial breast cancer (Yang et al., 2010). In 2012, a genome-wide study (Gong et al., 2012) compared the canonical pri-miR, pre-miR, and mature microRNA sequences annotated on the miRBase database (Griffiths-Jones, 2006) from 9 species with the classified SNPs collected from the NCBI dbSNP database (Sherry, 2001). Predictions of putative targets for wild-type microRNAs revealed that more than 50% of predicted miRNA-target bindings (55,887) were negatively affected by SNPs in the seed sequence. At the same time, more than 50% of the predicted targets for the new microRNA variants were exclusive for the SNP-isomiRs (Gong et al., 2012). Experimental validation of these data through Luciferase assay showed a partial or total loss of binding for at least four isomiRs and the addition of a new target for miR-627 (Gong et al., 2012). Particularly impressive is the SNP rs3746444 on miR-499-3p that heavily affects the binding between this isomiR and BCL2, which remains a validated target for the canonical molecule (Gong et al., 2012). Another example is the case of miR-124, whose isomiR, holding the SNP rs34059726, completely loses the ability to target ATP6V0E1, restoring the Luciferase activity from 10% in the presence of wild-type microRNA to 80% (Nishikura, 2016).

The importance of SNPs is defined not only by their frequency but also by the functional modifications they could induce. In the case of microRNAs, it has been well established that the presence of SNPs on microRNA genes sequence could affect their biogenesis and, indirectly, the expression of their targets. We have recently started to associate the presence of SNPs within mature microRNAs with the generation of isomiRs. Small RNA sequencing allowed the discovery of several new microRNA variants ascribable to genetic variations. The characterization of these SNP-isomiRs and the potential role that they could assume in specific biological conditions and diseases remains an exciting field to explore in the next future.

MicroRNA Editing

RNA editing is a type of RNA processing that occurs on double-strand RNA molecules at the co-transcriptional or post-transcriptional level (Nishikura, 2016). This process consists of specific bases deamination leading to a modification of the sequence (Nishikura, 2016). Among observed editing types, the adenosine-to-inosine (A-to-I) and the cytosine-to-uracil (C-to-U) (Nishikura, 2016) represent the most common modifications.

The A-to-I modification contributes to almost 90% of all editing events, and it is mediated by proteins belonging to the Adenosine Deaminase Acting on RNA (ADAR) family (Table 1), particularly ADAR (or ADAR1) and ADARB1 (or ADAR2) (Bazak et al., 2014). The interpretation of the inosine as guanine *de facto* results in a functional substitution A-to-G (Bazak et al., 2014).

To date, 2,885 A-to-I and 104 C-to-U unique editing events have been identified on microRNA transcripts, but only 257 have been confirmed by further investigations (Marceca et al., 2020). The consequences of editing modifications on microRNAs could change their expression or function (Li et al., 2018). Editing events involving pri-miRNA, pre-miRNA, and mature sequences could affect the microRNA maturation process, interfering with DROSHA and DICER cleavage, as well as the asymmetric selection of the strand (Li et al., 2018).

The editing contribution to the isomiRs generation is substantial, even if the accurate detection of editing sites in mature microRNAs by small RNA sequencing could be a long and complicated process due to the brevity of these molecules (Li et al., 2018). Li et al. (2018) identified 367 new editing sites in mature microRNAs. Their data on edited pre-miRNAs allowed the development of a custom pre-miRNA database to map newly edited mature microRNAs correctly. The editing sites have been identified throughout the mature molecule sequence, anticipating the potential change of seed sequence in these microRNAs (Li et al., 2018). Indeed, their target prediction analyses demonstrated that canonical and edited microRNAs shared only 10–35% of common targets (Li et al., 2018).

A critical functional effect has been observed following the editing of miR-200b in position 5 (within the seed sequence) in breast and ovarian cancer cells (Wang et al., 2017). In these two cancer models, the editing of miR-200b induces the loss of ability to target ZEB1/ZEB2 (Wang et al., 2017). Concurrently, the gain of new targets, including the metastasis suppressor LIFR, contributes to conferring a new role to the edited miR-200b: a negative regulator of cancer metastasis becomes a promoter of cell invasion and migration in response to ADAR-mediated modifications (Wang et al., 2017).

Similarly, canonical and edited miR-589-3p play two distinct roles in normal brain and glioblastoma tissues (Cesarini et al., 2018). Cesarini et al. (2018) showed that almost 100% of miR-589-3p molecules are edited in normal brain cells. This editing level strongly decreases in tumor cells together with astrocytomas grade of malignancy. Moreover, they demonstrated that the edited miR-589-3p gains the capacity to inhibit ADAM12, a well-characterized oncogene promoting glioblastoma cell aggressiveness, thus explaining its high editing level in normal brain cells (Cesarini et al., 2018).

By contrast, Xu et al. (2019) demonstrated the tumor-suppressive effect of edited miR-379-5p in ovarian, breast, renal, and lung cancer cell lines. Relying on The Cancer Genome Atlas (TCGA) miRNA-Seq data, they observed a lower editing level of miR-379-5p in seven different tumor tissues, significantly correlating the higher expression of the edited variant with better patient survival. Experiments conducted *in vitro* and *in vivo* demonstrated that by acquiring a new group of targets, particularly CD97, edited miR-379-5p induces apoptosis

in cancer cells and, consequently, reduces cell proliferation (Xu et al., 2019).

These findings suggest that microRNA editing is a critical event that could potentially affect the expression or the role of the edited molecules, whether it takes place within the seed sequence or other regions of precursors or mature molecules. Lastly, the editing level could become a predictive factor of risk when it causes loss or gain of microRNA function due to the targetome-shifting in some particular diseases.

THE FUNCTIONAL IMPORTANCE OF ISOMIRS

IsomiRs Bind Ago Proteins

The Argonaute (Ago) proteins are essential mediators of microRNAs regulative action through suppressing protein translation or degrading mRNA by their specific RNase activity (Hutvagner and Zamore, 2002; Liu, 2004; Meister et al., 2004). Eight Ago proteins were discovered in humans, but only four can bind and load microRNAs, and just one, Ago2, holds the endonuclease activity essential to mediate the repressive action of microRNAs on their target (Meister et al., 2004). After the first step of biogenesis, the pre-miRNA hairpin is processed by the RNase DICER and then, as a mature microRNA molecule, loaded by the RISC Loading Complex (RLC), which comprises, in addition to DICER, TRBP and Ago2 (Gregory et al., 2004; Chendrimada et al., 2005).

Ago2 is identified as the RISC complex effector protein since it prevents the expression of the target mRNA via direct degradation or translation process barring (Liu, 2004; Meister et al., 2004). Therefore, it is considered the “slicer” of the RISC complex and the only protein essential for the complex proper functioning (Rand et al., 2004). It can be said that microRNAs exert their function of gene-expression repressors through their association with the Ago2 protein. Consequently, the evidence of microRNA recruitment by Ago2 strongly suggests the truly functional status of the microRNA molecule.

As previously mentioned, Londin et al. (2015) identified 3,707 novel microRNAs analyzing 1,323 samples across 13 different tissues. Crossing these data with 43 Ago CLIP-Seq (10 self-performed on their samples and 33 obtained from available public samples), they found 1,657 (44.7%) newly discovered miRNA sequences and 1,517 (54.7%) miRBase-cataloged microRNAs present in one or more of the samples examined, thus supporting the evidence of a similar microRNA-Ago binding rate for novel and canonical microRNAs.

In 2017, through AGO2 RIP-Seq analysis on normal and osteoarthritis chondrocytes, Haseeb et al. (2017) identified a pool of microRNAs and isomiRs, expressed in human chondrocytes, directly interacting with Ago2. MicroRNA novel variants represented 52% of all sequenced microRNAs (Haseeb et al., 2017). Although the authors detected isomiRs belonging to each of the categories (5′ or 3′ deletion, 5′ or 3′ addition, and internal substitutions), the approximate total of variants (46% out of 52%) was represented by 3′ isomiRs and only 6% by microRNAs with 5′ modifications (Haseeb et al., 2017). A reasonable explanation

for reading this phenomenon is that Ago2 and the other RISC complex proteins bind the mature microRNAs in correspondence of the 5′ end, thus protecting microRNAs from exonucleases nibbling action (Haseeb et al., 2017).

In confirmation of an independent functional role for these Ago-bound isomiR, *in silico* target prediction analyses for the canonical miR-140-3p and one of the most abundant 5′ deletion isoforms have unveiled that they share only 50 targets out of 190 exclusive canonical and 317 exclusive isoform targets, suggesting a potential peculiar role for this 5′ isomiR in chondrocytes (Haseeb et al., 2017).

Not only comprehensive approaches proved the isomiR loading into Ago2 (Ebhardt et al., 2009; Martí et al., 2010). As in the case of 5′-isomiR-101 (Llorens et al., 2013) or the miR-222 isoforms (Yu et al., 2017), the study of specific novel variants has revealed, through Ago2 co-immunoprecipitation assays, the interaction between isomiRs and Ago2 protein.

Ultimately, the evidence of the cooperative binding between isomiRs and Ago2 has been repeatedly verified to support the hypothesis that isomiRs are functional molecules. They are likely to harness the same functional pathways and are loaded by the RISC complex as their canonical counterparts.

The Target Redirecting

The most intriguing question about isomiRs is represented by their ability to repress new and different targets. In other words, are they unique and independent functional molecules? Cloonan et al. (2011) analyzed canonical microRNAs and their isomers expression in 10 different human tissues. The results demonstrated a strong correlation of expression between the two groups, corroborating the hypothesis that isomiRs could have a supportive role in targeting biological pathways already regulated by canonical counterparts (Cloonan et al., 2011). Moreover, their data suggested that canonical microRNAs and isomiRs cooperative targeting action is mainly geared toward key genes belonging to cancer pathways. The participation of more molecules in targeting only one mRNA strongly decreases the off-target effects (Cloonan et al., 2011).

In subsequent years, other groups have confirmed or contradicted these assumptions. Salem et al. (2016) studied the expression and function of miR-140-3p and its 5′ isomiR in breast cancer cells. Both these versions of miR-140-3p have been found upregulated in breast cancer tissues and participate in a tumor-suppressive strategy (Salem et al., 2016). Despite the collaborative repressing role, they affect different pathways: the canonical miR-140-3p controls the stemness of breast cancer cells, although the 5′ isoform, more expressed than the canonical microRNA, causes cell cycle arrest, along with inhibition of proliferation and cell migration. The seed sequence shifting in the 5′ variant allows the gaining of novel targets: COL4A1, ITGA6, and MARCKSL1 (Salem et al., 2016). In other cases, isomiRs and the canonical counterparts could have divergent functions, such as miR-411 and its 5′ isomiR in human vascular fibroblasts and venous tissues (van der Kwast et al., 2020): in response to acute ischemia, the level of 5′ isomiR-411 rapidly decreases, while canonical miR-411 undergoes upregulation (van der Kwast et al., 2020). Moreover,

the seed sequence shifting contributes to the acquisition of exclusive targets for both microRNAs that could justify their opposite expression trend: only the canonical mir-411 represses the expression of TGFB, leading to a pro-angiogenic phenotype, while the 5' isomiR controls F3 and ANGPT1, thus decreasing cell migration and angiogenesis (van der Kwast et al., 2020). Similarly, after neural differentiation in human embryonic stem cells, the 5' isomiR-9-1 gains the capacity to repress two new targets, DNMT3B and NCAM2, concurrently losing the ability to inhibit CDH1, that persists as a canonical miR-9-1 target (Tan et al., 2014).

The study of miRNA expression in human retina samples led to identifying a new 5' isoform of the neuronal-specific miR-124a-3p, representing less than 25% of mir-124a-3p total variants in human retina samples (Karali et al., 2016). Despite the shifting of one single base, the change of the seed sequence in this 5' isoform supports the gain of a new target, CDH11, a gene involved in neuronal differentiation, never identified as a canonical miR-124a-3p target (Karali et al., 2016).

Not only 5' isomiRs but also 3' isomiRs can increase or diversify canonical microRNA functions, as described by Yu et al. (2017) for miR-222 and its longer 3' variants. They have demonstrated how the upregulation of 3' isomiR variations of miR-222 subverts the well-known anti-apoptotic role of canonical miR-222 by inhibiting many members belonging to the PI3K-AKT pathway such as PIK3R3 (Yu et al., 2017).

Together with the already cited examples of edited microRNAs, these findings illustrate that isomiRs are functional molecules that could extend or change the canonical microRNAs' role by acquiring or losing different targets. On one side, the potential impact of 5' isomiRs is readily explained by variations in the seed sequence and the resulting acquisition or loss of some target control. On the other side, the role of 3' isomiRs is less predictable and more complex. Variations on the 3' end of microRNAs could affect the molecule biogenesis or the degradation, the efficiency of the loading process by Ago proteins, and the stability and the strength of the miRNA:mRNA binding (Bofill-De Ros et al., 2020), thus creating or preventing the conditions for targets inhibition. The complicated rules governing the miRNA:mRNA binding could make the study of the isomiRs' targeting properties long and frustrating. Most prediction bioinformatics tools rely on the conventional miRNA-target pairing recognition, considering the miRNA seed sequence solely. Therefore, the improvement of new strategies to discriminate selective targets for these novel microRNA variants is primary. In this context, the miR-CLIP technique (Imig et al., 2015), based on the transfection of pre-miRNAs conjugated with biotin and psoralen to trap mRNA targets in cells, aided the identification of targets differentially regulated by miR-124 and its 5' isoform (Wang et al., 2020). After the transfection of a modified pre-miRNA-124 miR-CLIP probe in HEK293T cells, supported by the subsequent individual transfection of canonical and isomiR-124, the authors identified 16 potential targets (Wang et al., 2020). Out of these 16 selected candidates, 12 were mostly regulated by canonical miR-124, three were regulated by both isoforms, and just one was strongly inhibited exclusively by isomiR-124 (Wang et al., 2020).

In conclusion, it is plausible to assume that isomiRs have an independent targeting activity. They could play as supporters or competitors of canonical microRNAs, making the unbalanced biogenesis promoting an isoform expression instead of another (with identical or opposite functions) a protective cell strategy to strengthen or reduce a specific microRNA inhibition power.

TECHNIQUES FOR ISOMIRS DETECTION

Next-generation (NextGen) sequencing is so far the method of choice for isomiR detection. The nature of the sequence to detect does not affect the efficiency or specificity of this technique because it is not based on the principle of primer or probe annealing, and indeed it is employed to discover new microRNA variants. The main obstacle preventing NextGen sequencing as a daily laboratory routine procedure is the cost and the need for most laboratories to rely on an external service, increasing experimental times.

Since microRNA research has assumed an even more critical role in molecular biology, many protocols for detecting these small RNAs have been developed (Ye et al., 2019). In particular, poly(A) and stem-loop qRT-PCR, with or without the employment of hydrolysis-based probes (Taqman) (Figure 3A), have become the most commonly used commercial techniques because of the high level of specificity, the brevity of experimental times, and the relatively low cost of reagents and machines (Ye et al., 2019). However, these methods have significant limitations in detecting and quantify isomiRs accurately. The annealing of primers and probes requires acknowledging the sequence, thus making the detection of new molecules technically impossible (Schamberger and Orbán, 2014; Magee et al., 2017). Moreover, the discrimination of two sequences that differ by only one or a few nucleotides is not guaranteed by these protocols and must be established empirically for each molecule using customized probes and appropriate controls (Schamberger and Orbán, 2014; Magee et al., 2017). It follows that the quantification of a specific isoform in a sample containing an abundance of the same microRNA variants is far from easy and is strongly affected by the expression of the particular isoform to detect (Avendaño-Vázquez and Flores-Jasso, 2020). Various attempts have been made recently to assess the expression of microRNA variants, including, for example, Dumbbell-PCR (Zhou et al., 2010; Honda and Kirino, 2015) and two-tailed RT-qPCR (Androvic et al., 2017). The first one is based on the employment of a 3'-stem-loop adapter, which acts as the reverse transcription trigger, and a 5'-stem-loop adapter, which contains a stop signal for reverse transcription (Honda and Kirino, 2015). The adapters are ligated specifically to the microRNA two ends by a T4 RNA ligase (Rnl2). Gaps or overlaps due to modifications of the microRNA sequence strongly influence the efficacy of this ligation process (Honda and Kirino, 2015; Figure 3B). Moreover, the employment of a Taqman probe partially complementary to the microRNA and partially to the 3' adapter sequences confers further specificity to this method that should discriminate both 5' and 3' variations during the ligation and amplification steps (Honda and Kirino, 2015; Figure 3B). Two-tailed RT-qPCR is a technique based

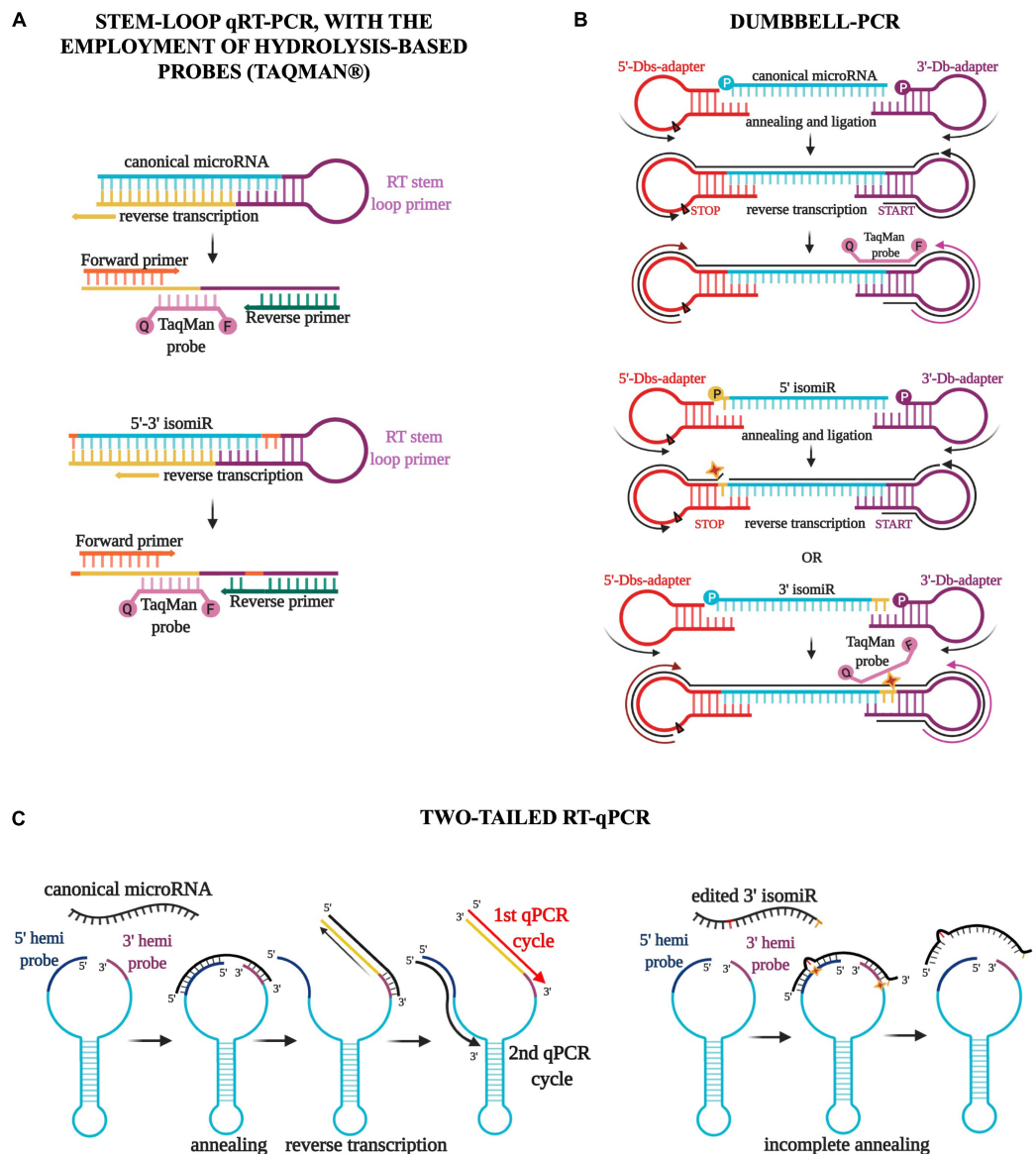


FIGURE 3 | Three different qPCR techniques for the detection of isomiRs. **(A)** The stem-loop qRT-PCR, with the employment of hydrolysis-based probes (Taqman), has become the most commonly used commercial technique. However, this method has significant limitations in detecting and quantifying isomiRs accurately. The discrimination of two sequences that differ by only one or a few nucleotides is not guaranteed by this protocol and must be established empirically for each molecule using customized probes and appropriate controls. **(B)** Dumbbell-PCR employs a 3'-stem-loop adapter, which acts as the reverse transcription trigger, and a 5'-stem-loop adapter, which contains a stop signal for reverse transcription. IsomiR gaps or overlaps strongly impact the efficacy of the ligation process and the annealing of a Taqman probe partially complementary to the microRNA and partially to the 3' adapter sequences. **(C)** Two-tailed RT-qPCR is characterized by the design of a long-structured primer (~50 nucleotides) holding two hemiprobcs complementary to the 5' and 3' ends of the microRNA. The reverse transcription starts from the 3' end, extending the primer sequence with the complementary sequence of the target microRNA and simultaneously detaching the 5' end. The amplification step uses two specific primers, one annealing the microRNA sequence and the other the 5' hemiprobe. The use of two short hemiprobcs increases the sensitivity and specificity of this technique: the brevity of these two sequences makes them more susceptible to possible isomiR mismatches.

on the design of a long structured primer (~50 nucleotides) holding two hemiprobcs complementary, respectively to the 5' and 3' ends of the microRNA to detect (Androvic et al., 2017; **Figure 3C**). After the annealing between the long primer and the microRNA sequence, the reverse transcription starts from the 3' end, extending the primer sequence with the complementary sequence of the target microRNA and simultaneously detaching

the 5' end (Androvic et al., 2017; **Figure 3C**). The following amplification step uses two specific primers, one annealing the microRNA sequence and the other the 5' hemiprobe. The use of two short hemiprobcs increases the sensitivity and specificity of this technique: the brevity of these two sequences makes them more susceptible to possible mismatches in the input sequence, as with isomiRs (Androvic et al., 2017; **Figure 3C**).

A different and more conventional approach has been considered to detect isomiRs containing edited nucleotides or SNPs. A mere Sanger sequencing of the RT-PCR products can determine the presence of modifications on pri-miRNA, pre-miRNA, or mature microRNA sequences (when required, after the addition of a poly(A) tail and a 5' adapter to make the molecules long enough for the amplification step and following sequencing) (Kawahara, 2012). The limit of this technique is the ability to identify edited variants when representing less than 10% of the total. Moreover, it requires previous knowledge of the editing sites or SNPs, thus narrowing the discovery of new isoforms (Kawahara, 2012).

Multiplex Single Base Primer Extension Assay can also detect internal modifications on microRNA sequences (Podini and Vallone, 2009). This protocol starts with multiplex PCR amplification of the regions containing the modifications to identify. Purified PCR products are then amplified again using 5'→3' primers with the last nucleotide at 3' end adjacent to the modification and a fluorescently labeled dideoxynucleotide (ddNTP) corresponding to the modification site (Podini and Vallone, 2009). PCR products are loaded onto capillary electrophoresis, and the resulting electropherograms are analyzed with suitable analysis software (Podini and Vallone, 2009).

In conclusion, the heterogeneity of these molecules makes developing a reliable, easy-to-use, and universal protocol a challenge for molecular biologists. So far, a unique and extensively validated technique to detect isomiRs does not exist. However, although all these methods still suffer from specificity and efficiency limitations, even using nanomaterial and fluorescent-based systems (Ye et al., 2019), many steps forward have been made in this field, and many others will be needed in the following years.

MICRORNA VARIANTS IN CANCER

IsomiRs are functional and independent molecules able to bind Ago proteins and play the role of gene-expression regulators as their canonical counterparts (Cloonan et al., 2011; Londin et al., 2015; Haseeb et al., 2017). Besides, their expression is finely regulated in different tissues and pathological conditions (Linsen et al., 2009; Wyman et al., 2011).

Telonis et al. (2017) performed a comprehensive study of isomiR expression, analyzing miRNA-Seq data from 32 different normal and tumor tissues belonging to The Cancer Genome Atlas (TCGA). The authors showed that differentially expressed isomiRs could discriminate between normal and cancer tissues and different tumor types (Telonis et al., 2017). Using binarized isomiR profiles that individually classify each isomiR as “present” or “absent,” they were able to recognize and cluster several tumor datasets efficiently (Telonis et al., 2017). This study has also highlighted that some isomiRs are ubiquitously expressed while others are tissue specific, such as microRNA variants of miR-9 and miR-219, two microRNAs detected mainly in the nervous system and involved in neuronal development and differentiation, expressed only in low-grade glioma (LGG) datasets, suggesting the potential role for these

cancer-specific molecules as biomarkers of some types of tumor (Telonis et al., 2017).

Other groups carried out isomiR profiling using TCGA miRNA-seq data, focusing mainly on edited microRNA expression.

Wang et al. (2017) reported the presence of 19 editing sites frequently expressed in cancer tissues, called “miRNA-editing-hotspots,” presenting an editing level above 5% and detected in at least 10 samples per cancer tissue. The analysis, conducted on 8,595 TCGA samples from 20 different tumor types, also pointed out an association between these editing hotspots and the expression of some critical oncogenes and tumor suppressors. For instance, the edited variant of miR-200b correlates with TP53 in head and neck, endometrial, and breast cancer, with NRAS and BRAF in thyroid cancer, and with CDH1 in gastric cancer (Wang et al., 2017).

Pinto et al. (2018) analyzed 10,593 miRNA-seq samples from the TCGA dataset representing 32 cancer and normal tissues. Applying stringent filters to avoid selecting molecules with an inconsiderable level of editing, they found 129 new editing sites on mature microRNA molecules, but only 55 showing an average editing level above 1% and three below but very close to 1%, in at least one out of the 32 TCGA examined tissues (Pinto et al., 2018). The expression analysis of these well-represented edited microRNAs displayed a significant lowering of the editing level in 19 of the 22 cancer tissues compared with the corresponding normal controls. This general condition of hypo-miRNA-editing in cancer suggests that miRNA-editing dysregulation could have a role in cancer progression. To confirm these results, they observed 56 patient cases, holding 26 different editing sites in 15 diverse tumor samples, and classifying patients into two groups according to their miRNA-editing levels. In line with the previous observations, better prognosis and, consequently, better patient survival were observed to be associated with higher miRNA-editing levels (Pinto et al., 2018).

In conclusion, an increasing number of papers are revealing the role of isomiRs in cancer. Conversely, the study of individual isomiRs and their specific targets is still an emerging field but of immense importance because the potential of isomiRs as prognostic and diagnostic markers in tumor conditions might be invaluable.

ISOMIRS IN NEURODEGENERATIVE AND METABOLIC DISEASES

Although most of the attention has been paid to the role of isomiRs in cancer, the interest in the behavior of these newly discovered molecules in chronic conditions, such as neurodegenerative and metabolic diseases, begins to catch on.

In 2016, a study involving early- and late-stage Alzheimer's patients revealed a significant change in 5' miRNA isoform level between the two groups of patients (Wang et al., 2016). Interestingly, among the 47 miRNAs showing relevant differences in their 5' variants level through the progression of the disease, 17 are actively involved in Alzheimer's disease pathogenesis (Wang et al., 2016). Similarly, an important dysregulation of canonical

miRNAs and isomiRs was observed in Huntington's patients (Martí et al., 2010). A massively parallel small RNA sequencing analysis, performed on healthy subjects and patients, unveiled that ~80–90% of miRNAs mapped in the human brain showed modification at 3' end, with a predominance of nucleotide addition as the most common modification, and ~35% presented nucleotide substitutions along the sequence (Martí et al., 2010). Moreover, this study demonstrated that the dysregulation of isomiRs presenting modifications at the 5' end significantly alter the expression of critical genes belonging to Huntington's disease canonical pathways (Martí et al., 2010). This observation, combined with a commonly observed co-expression of canonical miRNA and isomiRs in Huntington's patients' samples, suggests a cooperative role for the dysregulated isomiRs and the reference microRNAs in Huntington's disease (Martí et al., 2010).

The identification and study of new miRNA isoforms are starting to take hold also in the field of metabolic diseases. A study on the miR-27 family genes in metabolism has recently described the functional importance of miR-27 isoforms in metabolic processes associated with diseases (Ma et al., 2019). The overexpression of miR-27b-3p and two 3' isoforms in murine hepatocytes demonstrated the different impact of these molecules on the expression of some proteins with a critical role in metabolism: only the canonical miR-27b-3p strongly downregulates PEPCK, FAS, and SREBP1C, as well as the isomiR-27b-3p negatively affects the expression of G6PASE, CPT1A, and BMAL1, thus demonstrating the independent and distinct role of these miRNA variants in liver cells (Ma et al., 2019).

Baran-Gale et al. (2013) studied the miRNA profile in murine insulinoma cells and human beta cell and whole islets, finding an abundance of highly expressed 5'-shifted isomiRs. Then, they selected 10 microRNAs as potential regulatory hubs in type 2 diabetes, three of which are represented by 5' isomiRs: miR-375 + 1, miR-375-1, and miR-183-5p + 1 (Baran-Gale et al., 2013). *In silico* analyses and experimental validations confirmed that *Mtpn* is regulated only by the canonical miR-375, while *Atp6v0c* and *Cdc42* are predominantly repressed by miR-375 + 1 and miR-375-1, thus promoting the functional importance of 5'-shifted isomiRs as molecules able to affect the expression of type 2 diabetes-associated genes independently (Baran-Gale et al., 2013). In contrast to 5' and 3'-shifted isomiRs, the functional role of edited microRNAs in chronic neurodegenerative and metabolic diseases has not been currently assessed, despite the central role of ADAR proteins and RNA editing in these pathological contexts (Gan et al., 2006; Singh et al., 2007; Yang et al., 2012; Gaisler-Salomon et al., 2014; Aizawa et al., 2016; Khermesh et al., 2016). Certainly, taking into account the primary role of miRNA editing in cancer, it could represent an exciting field of study for future investigations also in other diseases.

DISCUSSION

In the past 20 years, microRNA research has become a primary branch of molecular biology. With the increasing employment of NextGen sequencing, the identification of new microRNA

variants, sometimes even more expressed and active than the database-annotated counterparts, has helped to reassess some aspects of microRNA biology so far regarded as established, such as the role of microRNA 3' end in the stability of target-binding and Ago2 microRNA-loading process.

However, notwithstanding the great leaps forward that isomiRs research has taken, there are still many questions to answer, and, among them, the most important: are isomiRs independent functional molecules? Despite the growing number of papers supporting the evidence of a specific role for these newly discovered molecules in assisting or preventing the activity of the canonical microRNAs, the current technical and bioinformatics limitations in predicting new individual targets, especially for non-seed-based nucleotide substitutions isomiRs, imply the idea that many of these new microRNA family members could be unnecessary or repetitive. Nevertheless, it is worth pointing out that the biogenesis of isomiRs is not a casual event but occurs under precise control. Most of these novel variants are well conserved across species and specifically expressed in some tissues and physiological or pathological conditions, including cancer, thus strengthening the hypothesis of an autonomous function.

Moreover, in the past, many researchers had to face the dilemma of a multiple, and commonly opponent, role of the same microRNA in different tissues or conditions (Banzhaf-Strathmann and Edbauer, 2014; Costa-Pinheiro et al., 2015; Wen et al., 2018; Shen et al., 2019; Rezaei et al., 2020; Xiang et al., 2020). It might be interesting to speculate if the presence of differentially expressed isoforms, aggregated and analyzed as a single microRNA, could explain the microRNA duality frequently observed.

Certainly, recognizing these microRNA variants does not question all the previous relevant discoveries about microRNA biology and genetics. Instead, it means that there are still questions to address and regulatory mechanisms to explore in the complex world of these critical regulators.

AUTHOR CONTRIBUTIONS

LT, RD, and GN conceived the structure of the article. LT wrote the article. RD, GN, and CMC read, edited, and approved the article and helped the discussion and correction of English writing. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the NIH Grants R35 CA197706 "Cancer Gene Discovery to Identify Targetable Targets" (to CMC).

ACKNOWLEDGMENTS

Figures 2, 3 were created with BioRender.com.

REFERENCES

- Aizawa, H., Hideyama, T., Yamashita, T., Kimura, T., Suzuki, N., Aoki, M., et al. (2016). Deficient RNA-editing enzyme ADAR2 in an amyotrophic lateral sclerosis patient with a FUSP525L mutation. *J. Clin. Neurosci.* 32, 128–129. doi: 10.1016/j.jocn.2015.12.039
- Alarcón, C. R., Lee, H., Goodarzi, H., Halberg, N., and Tavazoie, S. F. (2015). N6-methyladenosine marks primary microRNAs for processing. *Nature* 519, 482–485. doi: 10.1038/nature14281
- Androvic, P., Valihrach, L., Elling, J., Sjoback, R., and Kubista, M. (2017). Two-tailed RT-qPCR: a novel method for highly accurate miRNA quantification. *Nucleic Acids Res.* 45:e144. doi: 10.1093/nar/gkx588
- Avendaño-Vázquez, S. E., and Flores-Jasso, C. F. (2020). Stumbling on elusive cargo: how isomiRs challenge microRNA detection and quantification, the case of extracellular vesicles. *J. Extracell. Vesicles* 9:1784617. doi: 10.1080/20013078.2020.1784617
- Bail, S., Swerdel, M., Liu, H., Jiao, X., Goff, L. A., Hart, R. P., et al. (2010). Differential regulation of microRNA stability. *RNA* 16, 1032–1039. doi: 10.1261/rna.1851510
- Banzhaf-Strathmann, J., and Edbauer, D. (2014). Good guy or bad guy: the opposing roles of microRNA 125b in cancer. *Cell Commun. Signal.* 12:30. doi: 10.1186/1478-811X-12-30
- Baran-Gale, J., Fannin, E. E., Kurtz, C. L., and Sethupathy, P. (2013). Beta Cell 5'-Shifted isomiRs are candidate regulatory hubs in Type 2 diabetes. *PLoS One* 8:e73240. doi: 10.1371/journal.pone.0073240
- Bazak, L., Haviv, A., Barak, M., Jacob-Hirsch, J., Deng, P., Zhang, R., et al. (2014). A-to-I RNA editing occurs at over a hundred million genomic sites, located in a majority of human genes. *Genome Res.* 24, 365–376. doi: 10.1101/gr.164749.113
- Berezikov, E., van Tetering, G., Verheul, M., van de Belt, J., van Laake, L., Vos, J., et al. (2006). Many novel mammalian microRNA candidates identified by extensive cloning and RAKE analysis. *Genome Res.* 16, 1289–1298. doi: 10.1101/gr.5159906
- Bofill-De Ros, X., Kasprzak, W. K., Bhandari, Y., Fan, L., Cavanaugh, Q., Jiang, M., et al. (2019). Structural differences between Pri-miRNA paralogs promote alternative drosha cleavage and expand target repertoires. *Cell Rep.* 26, 447–459.e4. doi: 10.1016/j.celrep.2018.12.054
- Bofill-De Ros, X., Yang, A., and Gu, S. (2020). IsomiRs: expanding the miRNA repression toolbox beyond the seed. *Biochim. Biophys. Acta BBA Gene Regul. Mech.* 1863:194373. doi: 10.1016/j.bbaggm.2019.03.005
- Burroughs, A. M., Ando, Y., de Hoon, M. J. L., Tomaru, Y., Nishibu, T., Ukekawa, R., et al. (2010). A comprehensive survey of 3' animal miRNA modification events and a possible role for 3' adenylation in modulating miRNA targeting effectiveness. *Genome Res.* 20, 1398–1410. doi: 10.1101/gr.106054.110
- Calin, G. A., Ferracin, M., Cimmino, A., Di Leva, G., Shimizu, M., Wojcik, S. E., et al. (2005). A MicroRNA signature associated with prognosis and progression in chronic lymphocytic leukemia. *N. Engl. J. Med.* 353, 1793–1801. doi: 10.1056/NEJMoa050995
- Cesarini, V., Silvestris, D. A., Tassinari, V., Tomaselli, S., Alon, S., Eisenberg, E., et al. (2018). ADAR2/miR-589-3p axis controls glioblastoma cell migration/invasion. *Nucleic Acids Res.* 46, 2045–2059. doi: 10.1093/nar/gkx1257
- Chendrimada, T. P., Gregory, R. I., Kumaraswamy, E., Norman, J., Cooch, N., Nishikura, K., et al. (2005). TRBP recruits the Dicer complex to Ago2 for microRNA processing and gene silencing. *Nature* 436, 740–744. doi: 10.1038/nature03868
- Chiang, H. R., Schoenfeld, L. W., Ruby, J. G., Auyeung, V. C., Spies, N., Baek, D., et al. (2010). Mammalian microRNAs: experimental evaluation of novel and previously annotated genes. *Genes Dev.* 24, 992–1009. doi: 10.1101/gad.1884710
- Cloonan, N., Wani, S., Xu, Q., Gu, J., Lea, K., Heater, S., et al. (2011). MicroRNAs and their isomiRs function cooperatively to target common biological pathways. *Genome Biol.* 12:R126. doi: 10.1186/gb-2011-12-12-r126
- Costa-Pinheiro, P., Ramalho-Carvalho, J., Vieira, F. Q., Torres-Ferreira, J., Oliveira, J., Gonçalves, C. S., et al. (2015). MicroRNA-375 plays a dual role in prostate carcinogenesis. *Clin. Epigenetics* 7:42. doi: 10.1186/s13148-015-0076-2
- De Almeida, C., Scheer, H., Zuber, H., and Gagliardi, D. (2018). RNA uridylation: a key post-transcriptional modification shaping the coding and noncoding transcriptome: RNA uridylation. *Wiley Interdiscip. Rev. RNA* 9:e1440. doi: 10.1002/wrna.1440
- Ebhardt, H. A., Tsang, H. H., Dai, D. C., Liu, Y., Bostan, B., and Fahlman, R. P. (2009). Meta-analysis of small RNA-sequencing errors reveals ubiquitous post-transcriptional RNA modifications. *Nucleic Acids Res.* 37, 2461–2470. doi: 10.1093/nar/gkp093
- Friedman, R. C., Farh, K. K.-H., Burge, C. B., and Bartel, D. P. (2008). Most mammalian mRNAs are conserved targets of microRNAs. *Genome Res.* 19, 92–105. doi: 10.1101/gr.082701.108
- Frye, M., Jaffrey, S. R., Pan, T., Rechavi, G., and Suzuki, T. (2016). RNA modifications: what have we learned and where are we headed? *Nat. Rev. Genet.* 17, 365–372. doi: 10.1038/nrg.2016.47
- Gaisler-Salomon, I., Kravitz, E., Feiler, Y., Safran, M., Biegon, A., Amariglio, N., et al. (2014). Hippocampus-specific deficiency in RNA editing of GluA2 in Alzheimer's disease. *Neurobiol. Aging* 35, 1785–1791. doi: 10.1016/j.neurobiolaging.2014.02.018
- Gan, Z., Zhao, L., Yang, L., Huang, P., Zhao, F., Li, W., et al. (2006). RNA editing by ADAR2 is metabolically regulated in pancreatic islets and β -cells. *J. Biol. Chem.* 281, 33386–33394. doi: 10.1074/jbc.M604484200
- Gong, J., Tong, Y., Zhang, H.-M., Wang, K., Hu, T., Shan, G., et al. (2012). Genome-wide identification of SNPs in microRNA genes and the SNP effects on microRNA target binding and biogenesis. *Hum. Mutat.* 33, 254–263. doi: 10.1002/humu.21641
- Gregory, R. I., Yan, K., Amuthan, G., Chendrimada, T., Doratotaj, B., Cooch, N., et al. (2004). The Microprocessor complex mediates the genesis of microRNAs. *Nature* 432, 235–240. doi: 10.1038/nature03120
- Griffiths-Jones, S. (2006). miRBase: microRNA sequences, targets and gene nomenclature. *Nucleic Acids Res.* 34, D140–D144. doi: 10.1093/nar/gkj112
- GTEx Consortium, Tan, M. H., Li, Q., Shanmugam, R., Piskol, R., Kohler, J., et al. (2017). Dynamic landscape and regulation of RNA editing in mammals. *Nature* 550, 249–254. doi: 10.1038/nature24041
- Gu, S., Jin, L., Zhang, Y., Huang, Y., Zhang, F., Valdmanis, P. N., et al. (2012). The loop position of shRNAs and Pre-miRNAs is critical for the accuracy of dicer processing in vivo. *Cell* 151, 900–911. doi: 10.1016/j.cell.2012.09.042
- Gutiérrez-Vázquez, C., Enright, A. J., Rodríguez-Galán, A., Pérez-García, A., Collier, P., Jones, M. R., et al. (2017). 3' Uridylation controls mature microRNA turnover during CD4 T-cell activation. *RNA* 23, 882–891. doi: 10.1261/rna.060095.116
- Ha, M., and Kim, V. N. (2014). Regulation of microRNA biogenesis. *Nat. Rev. Mol. Cell Biol.* 15, 509–524. doi: 10.1038/nrm3838
- Han, B. W., Hung, J.-H., Weng, Z., Zamore, P. D., and Ameres, S. L. (2011). The 3'-to-5' exoribonuclease nibbler shapes the 3' ends of MicroRNAs bound to *Drosophila* Argonaute1. *Curr. Biol.* 21, 1878–1887. doi: 10.1016/j.cub.2011.09.034
- Han, J., Lee, Y., Yeom, K.-H., Nam, J.-W., Heo, I., Rhee, J.-K., et al. (2006). Molecular basis for the recognition of primary microRNAs by the Drosha-DGCR8 complex. *Cell* 125, 887–901. doi: 10.1016/j.cell.2006.03.043
- Haseeb, A., Makki, M. S., Khan, N. M., Ahmad, I., and Haqqi, T. M. (2017). Deep sequencing and analyses of miRNAs, isomiRs and miRNA induced silencing complex (miRISC)-associated miRNome in primary human chondrocytes. *Sci. Rep.* 7:15178. doi: 10.1038/s41598-017-15388-4
- Honda, S., and Kirino, Y. (2015). Dumbbell-PCR: a method to quantify specific small RNA variants with a single nucleotide resolution at terminal sequences. *Nucleic Acids Res.* 43:e77. doi: 10.1093/nar/gkv218
- Huntzinger, E., and Izaurralde, E. (2011). Gene silencing by microRNAs: contributions of translational repression and mRNA decay. *Nat. Rev. Genet.* 12, 99–110. doi: 10.1038/nrg2936
- Hutvagner, G., and Zamore, P. D. (2002). A microRNA in a multiple-turnover RNAi enzyme complex. *Science* 297, 2056–2060. doi: 10.1126/science.1073827
- Imig, J., Brunschweiler, A., Brümmer, A., Guennegewig, B., Mittal, N., Kishore, S., et al. (2015). miR-CLIP capture of a miRNA targetome uncovers a lincRNA H19-miR-106a interaction. *Nat. Chem. Biol.* 11, 107–114. doi: 10.1038/nchembio.1713
- Jones, M. R., Quinton, L. J., Blahna, M. T., Neilson, J. R., Fu, S., Ivanov, A. R., et al. (2009). Zcchc11-dependent uridylation of microRNA directs cytokine expression. *Nat. Cell Biol.* 11, 1157–1163. doi: 10.1038/ncb1931
- Karali, M., Persico, M., Mutarelli, M., Carissimo, A., Pizzo, M., Singh Marwah, V., et al. (2016). High-resolution analysis of the human retina miRNome reveals

- isomiR variations and novel microRNAs. *Nucleic Acids Res.* 44, 1525–1540. doi: 10.1093/nar/gkw039
- Katoh, T., Hojo, H., and Suzuki, T. (2015). Destabilization of microRNAs in human cells by 3' deadenylation mediated by PARN and CUGBP1. *Nucleic Acids Res.* 43, 7521–7534. doi: 10.1093/nar/gkv669
- Katoh, T., Sakaguchi, Y., Miyauchi, K., Suzuki, T., Kashiwabara, S.-I., Baba, T., et al. (2009). Selective stabilization of mammalian microRNAs by 3' adenylation mediated by the cytoplasmic poly(A) polymerase GLD-2. *Genes Dev.* 23, 433–438. doi: 10.1101/gad.1761509
- Kawahara, Y. (2012). Quantification of adenosine-to-inosine editing of microRNAs using a conventional method. *Nat. Protoc.* 7, 1426–1437. doi: 10.1038/nprot.2012.073
- Khermish, K., D'Erchia, A. M., Barak, M., Annese, A., Wachtel, C., Levanon, E. Y., et al. (2016). Reduced levels of protein recoding by A-to-I RNA editing in Alzheimer's disease. *RNA* 22, 290–302. doi: 10.1261/rna.054627.115
- Kozomara, A., Birgaoanu, M., and Griffiths-Jones, S. (2019). miRBase: from microRNA sequences to function. *Nucleic Acids Res.* 47, D155–D162. doi: 10.1093/nar/gky1141
- Lagos-Quintana, M. (2001). Identification of novel genes coding for small expressed RNAs. *Science* 294, 853–858. doi: 10.1126/science.1064921
- Lan, Q., Liu, P. Y., Haase, J., Bell, J. L., Hüttelmaier, S., and Liu, T. (2019). The critical role of RNA m6a methylation in cancer. *Cancer Res.* 79, 1285–1292. doi: 10.1158/0008-5472.CAN-18-2965
- Lee, R. C., Feinbaum, R. L., and Ambros, V. (1993). The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. *Cell* 75, 843–854. doi: 10.1016/0092-8674(93)90529-Y
- Li, L., Song, Y., Shi, X., Liu, J., Xiong, S., Chen, W., et al. (2018). The landscape of miRNA editing in animals and its impact on miRNA biogenesis and targeting. *Genome Res.* 28, 132–143. doi: 10.1101/gr.224386.117
- Linsen, S. E. V., de Wit, E., Janssens, G., Heater, S., Chapman, L., Parkin, R. K., et al. (2009). Limitations and possibilities of small RNA digital gene expression profiling. *Nat. Methods* 6, 474–476. doi: 10.1038/nmeth0709-474
- Liu, J. (2004). Argonaute2 is the catalytic engine of mammalian RNAi. *Science* 305, 1437–1441. doi: 10.1126/science.1102513
- Liu, N., Abe, M., Sabin, L. R., Hendriks, G.-J., Naqvi, A. S., Yu, Z., et al. (2011). The exoribonuclease nibbler controls 3' end processing of MicroRNAs in *Drosophila*. *Curr. Biol.* 21, 1888–1893. doi: 10.1016/j.cub.2011.10.006
- Llorens, F., Bañez-Coronel, M., Pantano, L., del Río, J. A., Ferrer, I., Estivill, X., et al. (2013). A highly expressed miR-101 isomiR is a functional silencing small RNA. *BMC Genomics* 14:104. doi: 10.1186/1471-2164-14-104
- Londin, E., Lohrer, P., Telonis, A. G., Quann, K., Clark, P., Jing, Y., et al. (2015). Analysis of 13 cell types reveals evidence for the expression of numerous novel primate- and tissue-specific microRNAs. *Proc. Natl. Acad. Sci. U.S.A.* 112, E1106–E1115. doi: 10.1073/pnas.1420955112
- Ma, M., Yin, Z., Zhong, H., Liang, T., and Guo, L. (2019). Analysis of the expression, function, and evolution of miR-27 isoforms and their responses in metabolic processes. *Genomics* 111, 1249–1257. doi: 10.1016/j.ygeno.2018.08.004
- MacRae, I. J., Zhou, K., and Doudna, J. A. (2007). Structural determinants of RNA recognition and cleavage by Dicer. *Nat. Struct. Mol. Biol.* 14, 934–940. doi: 10.1038/nsmb1293
- Magee, R., Telonis, A. G., Cherlin, T., Rigoutsos, I., and Londin, E. (2017). Assessment of isomiR discrimination using commercial qPCR methods. *Noncoding RNA* 3:18. doi: 10.3390/ncrna3020018
- Marceca, G. P., Distefano, R., Tomasello, L., Lagana, A., Russo, F., Calore, F., et al. (2020). MiREDiBase: a manually curated database of editing events in microRNAs. *bioRxiv* [preprint]. doi: 10.1101/2020.09.04.283689
- Martí, E., Pantano, L., Bañez-Coronel, M., Llorens, F., Miñones-Moyano, E., Porta, S., et al. (2010). A myriad of miRNA variants in control and Huntington's disease brain regions detected by massively parallel sequencing. *Nucleic Acids Res.* 38, 7219–7235. doi: 10.1093/nar/gkq575
- Meister, G., Landthaler, M., Patkaniowska, A., Dorsett, Y., Teng, G., and Tuschl, T. (2004). Human argonaute2 mediates RNA cleavage targeted by miRNAs and siRNAs. *Mol. Cell* 15, 185–197. doi: 10.1016/j.molcel.2004.07.007
- Moore, M. J., Scheel, T. K. H., Luna, J. M., Park, C. Y., Fak, J. J., Nishiuchi, E., et al. (2015). miRNA-target chimeras reveal miRNA 3'-end pairing as a major determinant of Argonaute target specificity. *Nat. Commun.* 6, 8864. doi: 10.1038/ncomms9864
- Neilsen, C. T., Goodall, G. J., and Bracken, C. P. (2012). IsomiRs – the overlooked repertoire in the dynamic microRNAome. *Trends Genet.* 28, 544–549. doi: 10.1016/j.tig.2012.07.005
- Nishikura, K. (2016). A-to-I editing of coding and non-coding RNAs by ADARs. *Nat. Rev. Mol. Cell Biol.* 17, 83–96. doi: 10.1038/nrm.2015.4
- Pinto, Y., Buchumenski, I., Levanon, E. Y., and Eisenberg, E. (2018). Human cancer tissues exhibit reduced A-to-I editing of miRNAs coupled with elevated editing of their targets. *Nucleic Acids Res.* 46, 71–82. doi: 10.1093/nar/gkx1176
- Podini, D., and Vallone, P. M. (2009). "SNP genotyping using multiplex single base primer extension assays," in *Single Nucleotide Polymorphisms: Methods and Protocols*, ed. A. A. Komar (Totowa, NJ: Humana Press), 379–391. doi: 10.1007/978-1-60327-411-1_23
- Rand, T. A., Ginalski, K., Grishin, N. V., and Wang, X. (2004). Biochemical identification of Argonaute 2 as the sole protein required for RNA-induced silencing complex activity. *Proc. Natl. Acad. Sci. U.S.A.* 101, 14385–14389. doi: 10.1073/pnas.0405913101
- Reinhart, B. J., Slack, F. J., Basson, M., Pasquinelli, A. E., Bettinger, J. C., Rougvie, A. E., et al. (2000). The 21-nucleotide let-7 RNA regulates developmental timing in *Caenorhabditis elegans*. *Nature* 403, 901–906. doi: 10.1038/35002607
- Rezaei, T., Amini, M., Hashemi, Z. S., Mansoori, B., Rezaei, S., Karami, H., et al. (2020). microRNA-181 serves as a dual-role regulator in the development of human cancers. *Free Radic. Biol. Med.* 152, 432–454. doi: 10.1016/j.freeradbiomed.2019.12.043
- Rüegger, S., and Großhans, H. (2012). MicroRNA turnover: when, how, and why. *Trends Biochem. Sci.* 37, 436–446. doi: 10.1016/j.tibs.2012.07.002
- Salem, O., Erdem, N., Jung, J., Münstermann, E., Wörner, A., Wilhelm, H., et al. (2016). The highly expressed 5' isomiR of hsa-miR-140-3p contributes to the tumor-suppressive effects of miR-140 by reducing breast cancer proliferation and migration. *BMC Genomics* 17:566. doi: 10.1186/s12864-016-2869-x
- Saunders, M. A., Liang, H., and Li, W.-H. (2007). Human polymorphism at microRNAs and microRNA target sites. *Proc. Natl. Acad. Sci. U.S.A.* 104, 3300–3305. doi: 10.1073/pnas.0611347104
- Schamberger, A., and Orbán, T. I. (2014). 3' IsomiR species and DNA contamination influence reliable quantification of MicroRNAs by stem-loop quantitative PCR. *PLoS One* 9:e106315. doi: 10.1371/journal.pone.0106315
- Shen, K., Cao, Z., Zhu, R., You, L., and Zhang, T. (2019). The dual functional role of MicroRNA-18a (miR-18a) in cancer development. *Clin. Transl. Med.* 8:32. doi: 10.1186/s40169-019-0250-9
- Sherry, S. T. (2001). dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res.* 29, 308–311. doi: 10.1093/nar/29.1.308
- Singh, M., Kesterson, R. A., Jacobs, M. M., Joers, J. M., Gore, J. C., and Emeson, R. B. (2007). Hyperphagia-mediated obesity in transgenic mice misexpressing the RNA-editing enzyme ADAR2*. *J. Biol. Chem.* 282, 22448–22459. doi: 10.1074/jbc.M700265200
- Song, M.-S., and Rossi, J. J. (2017). Molecular mechanisms of Dicer: endonuclease and enzymatic activity. *Biochem. J.* 474, 1603–1618. doi: 10.1042/BCJ20160759
- Starega-Roslan, J., Witkos, T., Galka-Marciniak, P., and Krzyzosiak, W. (2015). Sequence features of Drosha and Dicer cleavage sites affect the complexity of IsomiRs. *Int. J. Mol. Sci.* 16, 8110–8127. doi: 10.3390/ijms16048110
- Sun, G., Yan, J., Noltner, K., Feng, J., Li, H., Sarkis, D. A., et al. (2009). SNPs in human miRNA genes affect biogenesis and function. *RNA* 15, 1640–1651. doi: 10.1261/rna.1560209
- Tan, G. C., Chan, E., Molnar, A., Sarkar, R., Alexieva, D., Isa, I. M., et al. (2014). 5' isomiR variation is of functional and evolutionary importance. *Nucleic Acids Res.* 42, 9424–9435. doi: 10.1093/nar/gku656
- Telonis, A. G., Magee, R., Lohrer, P., Chervoneva, I., Londin, E., and Rigoutsos, I. (2017). Knowledge about the presence or absence of miRNA isoforms (isomiRs) can successfully discriminate amongst 32 TCGA cancer types. *Nucleic Acids Res.* 45, 2973–2985. doi: 10.1093/nar/gkx082
- Thornton, J. E., Du, P., Jing, L., Sjekloca, L., Lin, S., Grossi, E., et al. (2014). Selective microRNA uridylation by Zcchc6 (TUT7) and Zcchc11 (TUT4). *Nucleic Acids Res.* 42, 11777–11791. doi: 10.1093/nar/gku805
- van der Kwast, R. V. C. T., Woudenberg, T., Quax, P. H. A., and Nossent, A. Y. (2020). MicroRNA-411 and Its 5'-IsomiR have distinct targets and functions and are differentially regulated in the vasculature under ischemia. *Mol. Ther.* 28, 157–170. doi: 10.1016/j.ymthe.2019.10.002

- Wang, S., Xu, Y., Li, M., Tu, J., and Lu, Z. (2016). Dysregulation of miRNA isoform level at 5' end in Alzheimer's disease. *Gene* 584, 167–172. doi: 10.1016/j.gene.2016.02.020
- Wang, Y., Sonesson, C., Malinowska, A. L., Laski, A., Ghosh, S., Kanitz, A., et al. (2020). MiR-CLIP reveals iso-miR selective regulation in the miR-124 targetome. *Nucleic Acids Res.* 49, 25–37. doi: 10.1093/nar/gkaa1117
- Wang, Y., Xu, X., Yu, S., Jeong, K. J., Zhou, Z., Han, L., et al. (2017). Systematic characterization of A-to-I RNA editing hotspots in microRNAs across human cancers. *Genome Res.* 27, 1112–1125. doi: 10.1101/gr.219741.116
- Wen, R., Umeano, A. C., Essegian, D. J., Sabitaliyevich, U. Y., Wang, K., and Farooqi, A. A. (2018). Role of microRNA-410 in molecular oncology: a double edged sword. *J. Cell. Biochem.* 119, 8737–8742. doi: 10.1002/jcb.27251
- Wu, C. W., Evans, J. M., Huang, S., Mahoney, D. W., Dukek, B. A., Taylor, W. R., et al. (2018). A comprehensive approach to sequence-oriented IsomiR annotation (CASMiR): demonstration with IsomiR profiling in colorectal neoplasia. *BMC Genomics* 19:401. doi: 10.1186/s12864-018-4794-7
- Wyman, S. K., Knouf, E. C., Parkin, R. K., Fritz, B. R., Lin, D. W., Dennis, L. M., et al. (2011). Post-transcriptional generation of miRNA variants by multiple nucleotidyl transferases contributes to miRNA transcriptome complexity. *Genome Res.* 21, 1450–1461. doi: 10.1101/gr.118059.110
- Xiang, Y., Tian, Q., Guan, L., and Niu, S. (2020). The dual role of miR-186 in cancers: oncomir battling with tumor suppressor miRNA. *Front. Oncol.* 10:233. doi: 10.3389/fonc.2020.00233
- Xu, X., Wang, Y., Mojumdar, K., Zhou, Z., Jeong, K. J., Mangala, L. S., et al. (2019). A-to-I-edited miRNA-379-5p inhibits cancer cell proliferation through CD97-induced apoptosis. *J. Clin. Invest.* 129, 5343–5356. doi: 10.1172/JCI123396
- Yang, A., Shao, T.-J., Bofill-De Ros, X., Lian, C., Villanueva, P., Dai, L., et al. (2020). AGO-bound mature miRNAs are oligouridylylated by TUTs and subsequently degraded by DIS3L2. *Nat. Commun.* 11:2765. doi: 10.1038/s41467-020-16533-w
- Yang, L., Huang, P., Li, F., Zhao, L., Zhang, Y., Li, S., et al. (2012). c-Jun amino-terminal kinase-1 mediates glucose-responsive upregulation of the RNA editing enzyme ADAR2 in pancreatic beta-cells. *PLoS One* 7:e48611. doi: 10.1371/journal.pone.0048611
- Yang, R., Schlehe, B., Hemminki, K., Sutter, C., Bugert, P., Wappenschmidt, B., et al. (2010). A genetic variant in the pre-miR-27a oncogene is associated with a reduced familial breast cancer risk. *Breast Cancer Res. Treat.* 121, 693–702. doi: 10.1007/s10549-009-0633-5
- Ye, J., Xu, M., Tian, X., Cai, S., and Zeng, S. (2019). Research advances in the detection of miRNA. *J. Pharm. Anal.* 9, 217–226. doi: 10.1016/j.jppha.2019.05.004
- Yu, F., Pillman, K. A., Neilsen, C. T., Toubia, J., Lawrence, D. M., Tsykin, A., et al. (2017). Naturally existing isoforms of miR-222 have distinct functions. *Nucleic Acids Res.* 45, 11371–11385. doi: 10.1093/nar/gkx788
- Yu, S., and Kim, V. N. (2020). A tale of non-canonical tails: gene regulation by post-transcriptional RNA tailing. *Nat. Rev. Mol. Cell Biol.* 21, 542–556. doi: 10.1038/s41580-020-0246-8
- Zhou, Y., Huang, Q., Gao, J., Lu, J., Shen, X., and Fan, C. (2010). A dumbbell probe-mediated rolling circle amplification strategy for highly sensitive microRNA detection. *Nucleic Acids Res.* 38:e156. doi: 10.1093/nar/gkq556

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 Tomasello, Distefano, Nigita and Croce. 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.



m5C-Related lncRNAs Predict Overall Survival of Patients and Regulate the Tumor Immune Microenvironment in Lung Adenocarcinoma

Junfan Pan^{1†}, Zhidong Huang^{2†} and Yiquan Xu^{3*}

OPEN ACCESS

Edited by:

Jia Meng,
Xi'an Jiaotong-Liverpool University,
China

Reviewed by:

Kunqi Chen,
University of Liverpool,
United Kingdom
Lian Liu,
Shaanxi Normal University, China

*Correspondence:

Yiquan Xu
xuyiquan@126.com

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

Specialty section:

This article was submitted to
Epigenomics and Epigenetics,
a section of the journal
Frontiers in Cell and Developmental
Biology

Received: 24 February 2021

Accepted: 01 June 2021

Published: 29 June 2021

Citation:

Pan J, Huang Z and Xu Y (2021)
m5C-Related lncRNAs Predict Overall
Survival of Patients and Regulate
the Tumor Immune Microenvironment
in Lung Adenocarcinoma.
Front. Cell Dev. Biol. 9:671821.
doi: 10.3389/fcell.2021.671821

¹ Shengli Clinical Medical College of Fujian Medical University, Fuzhou, China, ² Quanzhou First Hospital, Fujian Medical University, Quanzhou, China, ³ Department of Thoracic Oncology, Fujian Medical University Cancer Hospital, Fujian Cancer Hospital, Fuzhou, China

Long non-coding RNAs (lncRNAs), which are involved in the regulation of RNA methylation, can be used to evaluate tumor prognosis. lncRNAs are closely related to the prognosis of patients with lung adenocarcinoma (LUAD); thus, it is crucial to identify RNA methylation-associated lncRNAs with definitive prognostic value. We used Pearson correlation analysis to construct a 5-Methylcytosine (m5C)-related lncRNAs–mRNAs coexpression network. Univariate and multivariate Cox proportional risk analyses were then used to determine a risk model for m5C-associated lncRNAs with prognostic value. The risk model was verified using Kaplan–Meier analysis, univariate and multivariate Cox regression analysis, and receiver operating characteristic curve analysis. We used principal component analysis and gene set enrichment analysis functional annotation to analyze the risk model. We also verified the expression level of m5C-related lncRNAs *in vitro*. The association between the risk model and tumor-infiltrating immune cells was assessed using the CIBERSORT tool and the TIMER database. Based on these analyses, a total of 14 m5C-related lncRNAs with prognostic value were selected to build the risk model. Patients were divided into high- and low-risk groups according to the median risk score. The prognosis of the high-risk group was worse than that of the low-risk group, suggesting the good sensitivity and specificity of the constructed risk model. In addition, 5 types of immune cells were significantly different in the high- and low-risk groups, and 6 types of immune cells were negatively correlated with the risk score. These results suggested that the risk model based on 14 m5C-related lncRNAs with prognostic value might be a promising prognostic tool for LUAD and might facilitate the management of patients with LUAD.

Keywords: m5C, lncRNA, lung adenocarcinoma, prognostic signature, overall survival, tumor immune microenvironment

INTRODUCTION

Lung cancer is the leading cause of cancer-related deaths worldwide. Approximately, 18 million people are diagnosed with lung cancer each year, and 16 million die because of this disease (Bray et al., 2018). There are 2 main types of lung cancer: non-small cell lung cancer (NSCLC) and small cell lung cancer (SCLC). In particular, NSCLC is the main type of primary lung cancer, accounting for approximately 80–85% of cases, and it can be further classified into adenocarcinoma, squamous cell carcinoma, and large cell carcinoma (Liu et al., 2015; Osmani et al., 2018). Adenocarcinoma is the major type of NSCLC, and most patients are diagnosed when already in an advanced and inoperable metastatic stage, with a 5-y survival rate of only 4.7% (Cohen et al., 2020). Recently, targeted therapy for specific driver gene mutations and the use of immune checkpoint inhibitors have popularized individualized therapy and effectively improved the prognosis of patients with advanced metastatic disease. However, because of the high drug resistance and metastasis rate, more specific biomarkers are needed that could lead to the development of more effective diagnostic and treatment strategies.

Methylation is one among the wide range of RNA post-transcriptional modifications, determining translation efficiency (Li Q. et al., 2017; Zhao et al., 2017). N6-methyladenosine (m6A) epitranscriptional modification is the main modification mode of RNA, and a number of high-throughput experimental methods have been developed to characterize the transcriptome range of modifications of m6A (Chen K. et al., 2019). Besides m6A, 5-methylcytosine (m5C) is another common mRNA modification (Dou et al., 2020). More specifically, m5C, in which the methyl group is attached to the fifth position of the cytosine ring, is catalyzed by RNA methyltransferase. The m5C locus has been reported to be involved in a variety of biological processes, including structural stability and metabolism of RNA, tRNA recognition, and stress response (Zhao et al., 2017). In addition, m5C modification has also been closely related to cancer progression. A recent study has shown that in human urothelial cell carcinoma of the bladder, m5C regulators bound to the 3'UTR of oncogene mRNA, stabilizing its expression, thereby promoting cancer progression (Chen X. et al., 2019). Bioinformatic studies have shown that m5C regulators could be used as a prognostic factor for lung adenocarcinoma (LUAD) (Sun L. et al., 2020), head and neck squamous cell carcinoma (Xue et al., 2020), and hepatocellular carcinoma (He et al., 2020b).

Long non-coding RNAs (lncRNAs), which are the main non-coding RNAs, are transcripts longer than 200 nt. They are known to play a key role in chromatin remodeling, transcription, and post-transcriptional regulation (Kopp and Mendell, 2018). In addition, RNA methylation of lncRNAs has been demonstrated to affect cancer progression. For instance, in hepatocellular carcinoma, the m6A “writer” METTL3 increased the stability of LINC00958, promoting the progression of cancer (Zuo et al., 2020). Similarly, the m6A “eraser” ALKBH5 increased the invasion and metastasis of tumor cells in gastric cancer by

inhibiting the methylation of NEAT1 (Zhang et al., 2019). m5C modification sites are also widely distributed in non-coding RNAs. Studies revealed that progressive conventional bisulfite sequencing in HeLa cells identified new m5C candidate sites, with approximately 1780 of these sites being shown to exist in multiple types. Such non-coding RNA types include lncRNAs (Squires et al., 2012). However, there have been few studies on the regulation of m5C in lncRNAs, and further research is needed.

In recent years, a large number of studies have shown that immune cells in the tumor microenvironment (TME) play a vital role in cancer progression and the therapeutic efficacy of applied treatments (Lei et al., 2020). In the early stage of tumorigenesis, anti-tumor immune cells in the tumor microenvironment tend to target cancer cells. As the tumor progresses, tumor cells will eventually resist the cytotoxic effects of immune cells and escape immune surveillance. lncRNAs have become key regulatory elements in the immune system, playing an important role in guiding the development of a variety of immune cells and controlling the dynamic transcription program. The dynamic transcription program is a sign of immune cell activation (Sun J. et al., 2020). For example, in epithelial cells, lincRNA-Cox2 was reported to participate in the Mi-2/nucleosome remodeling and deacetylase (Mi2/NuRD) complex, facilitating the recruitment of the IL12 β promoter to prevent the expression of IL12 β (Atianand et al., 2017). The chromatin-related CD4⁺ Th1-specific lncRNA lnc-MAF-4 was found to be inversely proportional to the expression of the transcription factor MAF, and its downregulation elicited the differentiation of helper CD4⁺ T-cells to the Th2 phenotype (Ranzani et al., 2015). Recent studies have also emphasized that lncRNAs play a key role in cancer immune regulation. In colorectal cancer, the antisense lncRNA SATB2-AS1 was shown to regulate SATB2 expression, inducing the expression of the CXCL9 and CXCL10 Th1-type chemokines and mediating the transport of effector T-cells (Xu et al., 2019). Based on this result, it was considered to serve as a potential target for colorectal cancer immunotherapy. In addition, a correlation between lncRNAs and immune cell infiltration has also been observed in uveal melanoma (Wang et al., 2018), pancreatic cancer (Qin et al., 2019), and cervical cancer (Guo et al., 2019). However, current research on the correlation between lncRNAs and immune cell infiltration is rare in LUAD, and further research is needed.

Although there are many bioinformatics studies focus on the RNA modifications (Chen K. et al., 2019, 2021; Liu and Chen, 2020; Liu et al., 2020; Song et al., 2020; Chen H. et al., 2021; Tang et al., 2021; Xu et al., 2021), to the best of our knowledge, this is the first in-depth analysis of the role of m5C regulators in LUAD. In this study, we used the expression data of m5C-related lncRNAs in LUAD from the Cancer Genome Atlas (TCGA) dataset, screened out 14 m5C-related lncRNAs with prognostic value, constructed an m5C-related lncRNAs prognostic signature (m5C-LPS), and further analyzed the relationship between m5C-LPS and immune cell infiltration subtypes. We aimed to explore the different gene characteristics, prognostic value, and impact on the tumor immune microenvironment (TIM) of the RNA methylation of

m5C-related lncRNAs in LUAD, so as to provide guidance for the treatment of LUAD.

MATERIALS AND METHODS

Data Acquisition

Transcriptome analysis of raw data and corresponding clinical information of the LUAD cohort were downloaded from TCGA data portal¹. A total of 551 patients with LUAD with lncRNA expression profiles, including 497 LUAD tissues and 54 normal lung tissues, of which 486 cases contained follow-up time and complete clinical case characteristics. **Table 1** lists the detailed clinical characteristics of the 486 patients with LUAD. lncRNA annotation file of Genome Reference Consortium Human Build 38 (GRCh38) was acquired from the GENCODE website for annotation of the lncRNAs in the TCGA dataset. Based on recognizing the Ensemble IDs of the genes, 14,142 lncRNAs were identified in the TCGA dataset.

¹<http://cancergenome.nih.gov/>

TABLE 1 | The clinical characteristics of lung adenocarcinoma patients in the TCGA database.

Variables	No. of Patients	Percentage (%)
Age (years)		
≤65	227	46.7
>65	240	49.4
Unknown	19	3.9
Gender		
Female	264	54.3
Male	222	45.7
Pathological stage		
I	262	53.9
II	112	23.0
III	79	16.3
IV	25	5.1
Unknown	8	1.7
T stage		
T1	163	33.5
T2	260	53.5
T3	41	8.5
T4	19	3.9
Unknown	3	0.6
N stage		
N0	312	64.2
N1	90	18.5
N2	70	14.4
N3	2	0.4
Unknown	12	2.5
M stage		
M0	333	68.5
M1	24	4.9
Unknown	129	26.6

Identification of Differentially Expressed Genes in TCGA Database

A total of 13 m5C regulators were obtained from the published literature, comprising *NSUN2*, *NSUN3*, *NSUN4*, *NSUN5*, *NSUN6*, *NSUN7*, *ALYREF*, *DNMT1*, *DNMT3A*, *DNMT3B*, *TET2*, *TRDMT1*, and *YBX1*. Extract of the expression matrix and clinical data of m5C regulators of 551 patients with LUAD were obtained from TCGA database. Then, the limma software package in R version (4.0.2) was used to identify the differentially expressed m5C regulators between LUAD and normal lung tissues. A p value < 0.05 and $|\log_2(\text{folding change})| > 1$ were considered significantly different. Subsequently, heatmaps were used to show the differential expression of m5C regulators between LUAD and normal lung tissues.

Construction of Protein–Protein Interaction (PPI) Network

We analyzed the differentially expressed 13 m5C regulators using the Search Tool for Interaction Genes (STRING) database² to construct a PPI network (Li et al., 2020). As a threshold of genes at the center of the PPI network was set a minimum gene interaction score < 0.7 .

m5C-Related lncRNAs

The “limma R” package was used to detect m5C-related lncRNAs. The m5C-related lncRNAs were identified by the correlation analysis between the m5C genes and lncRNA expression levels in the LUAD samples. The correlation coefficient > 0.3 and $p < 0.001$ as criteria, and 1094 m5C-related lncRNAs were identified.

Constructing the Prognostic Risk Model of m5C-Related lncRNAs

In order to identify m5C-related lncRNAs with prognostic value, We conducted a univariate Cox analysis based on the standard of $p < 0.01$. Subsequently, multivariable Cox analysis was used to establish a risk score. The risk score of each patient was calculated according to the following formula:

Risk score = coef (lncRNAn) \times expr (lncRNAn), where coef (lncRNAn) and expr (lncRNAn) represent the survival correlation regression coefficient and expression value, respectively, of each m5C-related lncRNA.

It should be noted that coef (lncRNAn) was defined as the correlation between lncRNAs and survival, whereas expr (lncRNAn) was defined as the expression of lncRNAs.

Evaluation of the Risk Model of 14m5C-Related lncRNAs as Independent Prognostic Factor in LUAD

According to the median value of the prognostic risk score, patients with LUAD were divided into high- and low-risk groups. The Kaplan–Meier survival curve was used to compare the overall survival (OS) of patients in the two groups.

²<https://string-db.org/>

Principal component analysis (PCA) was conducted for effective dimensionality reduction, pattern recognition, and exploratory visualization analysis of the whole genome, m5C-related coding genes, and m5C-related lncRNAs expression profiles. Univariate and multivariate Cox regression analyses were used to assess whether the risk score was independent of other clinical variables. The receiver operating characteristic (ROC) curve was used to evaluate the diagnostic and prognostic value of clinicopathological features. A p value < 0.05 was considered statistically significant.

Gene Set Enrichment Analysis (GSEA)

Gene set enrichment analysis was performed in the LUAD cohort to gain insights into the biological pathways of the high- and low-risk subgroups defined by the expression characteristics of the 14 m5C-related lncRNAs. Gene sets with false discovery rate (FDR) < 0.25 and normalized p value < 0.05 were considered significant.

Cell Lines and Reagents

Human normal lung epithelial cell line BEAS-2B, human LUAD cell lines PC-9, H1299 were purchased from the Cell Room of School of Medicine, Central South University. All cells were cultured in RPMI 1640 medium (Gibco, United States) supplemented with 10% fetal bovine serum (10% FBS), and the cells were cultured in a humidified atmosphere at 37°C and 5% CO₂.

Total RNA Extraction and Real-Time Quantitative PCR

In order to evaluate the expression level of m5C-related lncRNAs, we used RNA trizol reagent (Invitrogen, Carlsbad, CA, United States) to isolate total cellular RNA. Reverse transcription was done using PrimeScript™ RT reagent Kit (Takara, Japan). Real-time fluorescent quantitative PCR was performed by using GoTaq® qPCR Master Mix kit (Promega, United States). The expression level of related lncRNAs was calculated using $2^{-\Delta\Delta CT}$, and the related GAPDH mRNA expression was used as an endogenous control. The primer sequences involved in this study are shown in **Supplementary Table 1**. Each PCR reaction was performed in triplicate.

Prediction of m5C Sites on 14 lncRNAs

RNAm5Cfinder³ (Li et al., 2018), iRNA-m5C-PseDNC⁴ (Qiu et al., 2017), iRNA-m5C⁵ (Lv et al., 2020) databases were used to further verify the 14 prognostic-related lncRNAs were m5C modified lncRNAs by predicting m5C modification sites on lncRNAs.

Analysis of Immune Cell Characteristics

CIBERSORT is an analytical tool with a gene expression signature matrix of 547 marker genes used for the quantitative infiltration of immune cell components. LM22, which defines the 22 immune

cell subtypes of the annotation of the gene signature matrix, was downloaded from the CIBERSORT website⁶. We used 100 permutations of the default signature matrix to calculate the CIBERSORT p value and root-mean-square error for each sample file to improve the accuracy of the deconvolution algorithm. Then, we used the CIBERSORT value of $p < 0.05$ and data of LUAD tissues were screened and selected for subsequent analysis. The CIBERSORT algorithm was used to analyze the immune cell composition of LUAD samples in TCGA cohort.

The TIMER Database

In order to assess the risk score and the level of correlation of the infiltration of immune cells, we used the TIMER database⁷. This database contains 10897 TCGA samples of 32 types of cancer. Data were collected on 6 types of immune infiltrates, including B-cells, CD4⁺ T-cells, CD8⁺ T-cells, neutrophils, macrophages, and dendritic cells. In addition, it provides three main analysis modules: Immune, Exploration, and Estimation. The Immune module includes clinical outcomes, somatic mutation, and somatic copy number change, enabling users to comprehensively analyze the relationship between immune cell infiltration and multiple factors (Li T. et al., 2017).

Statistical Analysis

One-way analysis of variance was used to compare the expression levels of 13 m5C regulators in 497 LUAD tissues and 54 normal lung tissues. A coexpression network of 14 prognostic m5C-related lncRNAs-mRNAs was established and visualized using Cytoscape. Gene Expression Profiling Interactive Analysis (GEPIA) was used to evaluate the correlation between m5C-related lncRNAs and m5C regulators. The ggpubr package in R was used to analyze the correlation between the expression of 14 m5C-related lncRNAs and clinicopathological factors. PCA was used to effectively reduce dimensionality, pattern recognition, and exploratory visualization of high-dimensional data for the whole genome, m5C-related coding genes, and 14 m5C-related lncRNAs expression profiles. GSEA was performed for functional annotation. The Kaplan–Meier method was used to compare the OS time of each group. Univariate and multivariate Cox proportional hazards models were used to determine significant prognostic factors. ROC curve was used to evaluate the predictive efficiency of the prognostic risk scoring model. The RT- qPCR results were analyzed using One-way ANOVA. Statistical analysis was performed using R software (version 4.0.2). A $p < 0.05$ was considered statistically different.

RESULTS

Differentially Expressed m5C Regulators Between LUAD and Normal Lung Tissues

To identify the differentially expressed genes and related functions of m5C regulatory factors in LUAD, we used TCGA database to analyze 497 and 54 cases of LUAD and normal lung

³<https://www.rnanut.net/rnam5cfinder/>

⁴<http://www.jci-bioinfo.cn/iRNA-m5C-PseDNC>

⁵<http://lin-group.cn/server/iRNA-m5C/service.html>

⁶<http://CIBERSORT.stanford.edu/>

⁷<http://cistrome.shinyapps.io/timer/>

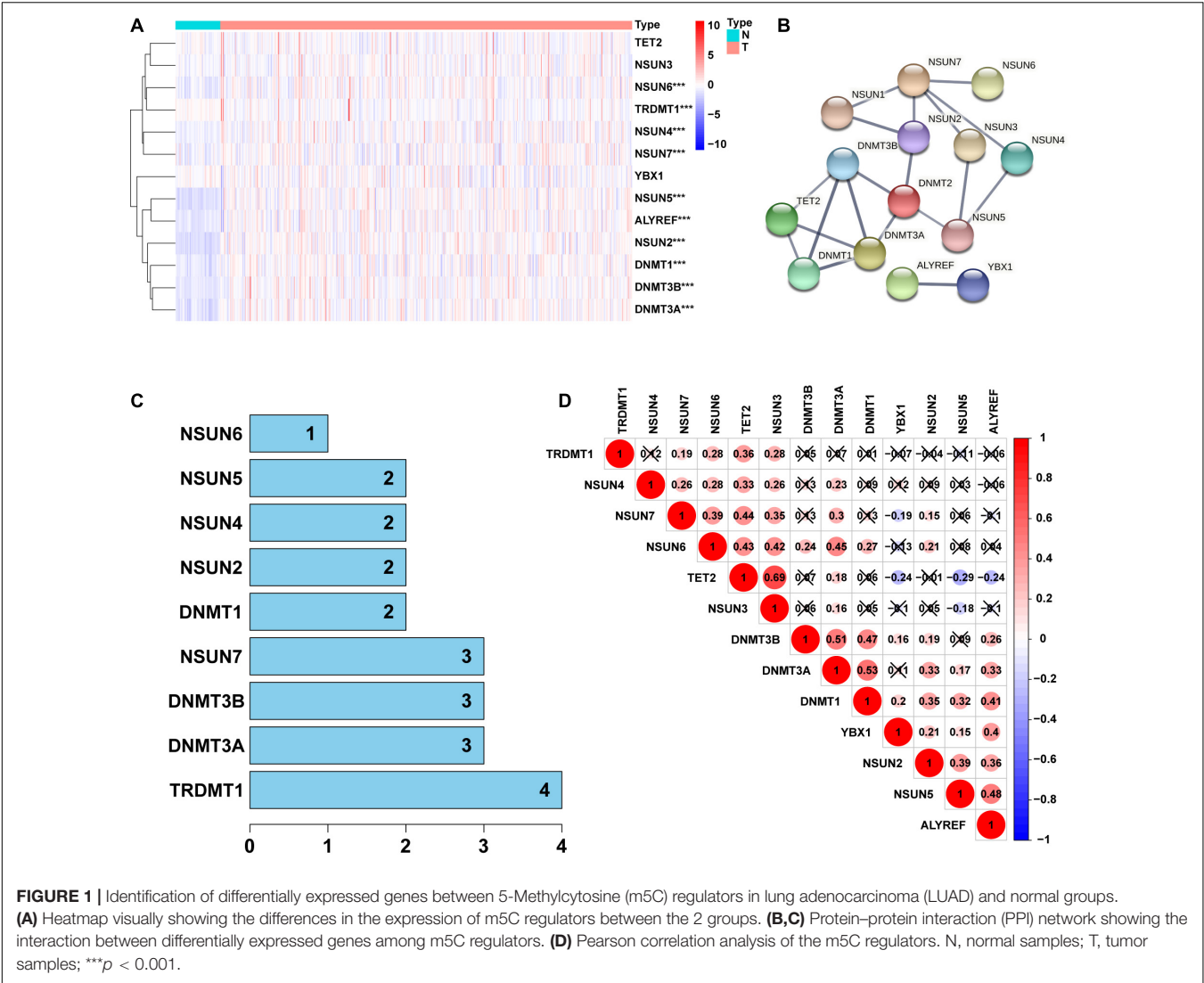
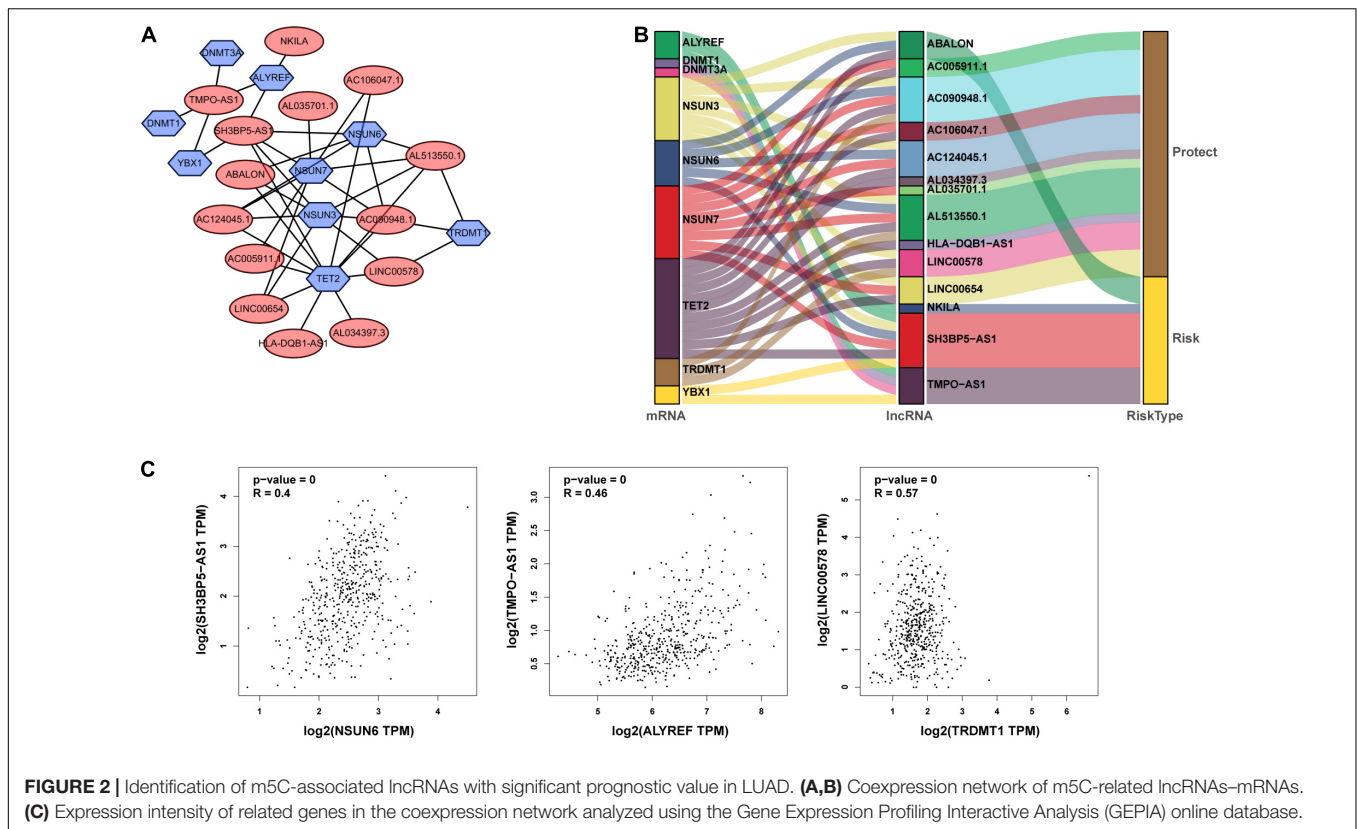


TABLE 2 | The 14 m5C-Related Prognostic lncRNAs.

ID	Coef	HR	HR. 95L	HR. 95H	P value
AC124045.1	−0.487330859	0.614263765	0.404713916	0.932312821	0.022068291
NKILA	0.103614091	1.109172333	1.048491764	1.173364738	0.000306694
AC090948.1	−0.504854345	0.603593486	0.35583685	1.02385432	0.061133855
AL035701.1	−0.108845737	0.89686876	0.78069115	1.030335201	0.124106608
LINC00578	−0.156948523	0.854748057	0.763901987	0.95639788	0.006189461
AC106047.1	−0.304007158	0.737855585	0.542076261	1.004343674	0.05330851
ABALON	0.26512182	1.303589769	0.917226317	1.852701188	0.139348063
HLA-DQB1-AS1	−0.047113201	0.9539794	0.907522841	1.002814094	0.064365554
AC005911.1	−0.339018709	0.71246912	0.482910828	1.051151097	0.087533584
AL513550.1	−0.377572971	0.685523176	0.499373156	0.941063851	0.019504797
AL034397.3	−0.19339043	0.824160135	0.66003064	1.02910363	0.087863805
TMPO-AS1	0.299437423	1.349099622	1.105221645	1.646791662	0.003246416
SH3BP5-AS1	0.239153168	1.270173071	1.078451102	1.495978471	0.004174978
LINC00654	−0.32999152	0.71892983	0.564726421	0.915239807	0.007384148



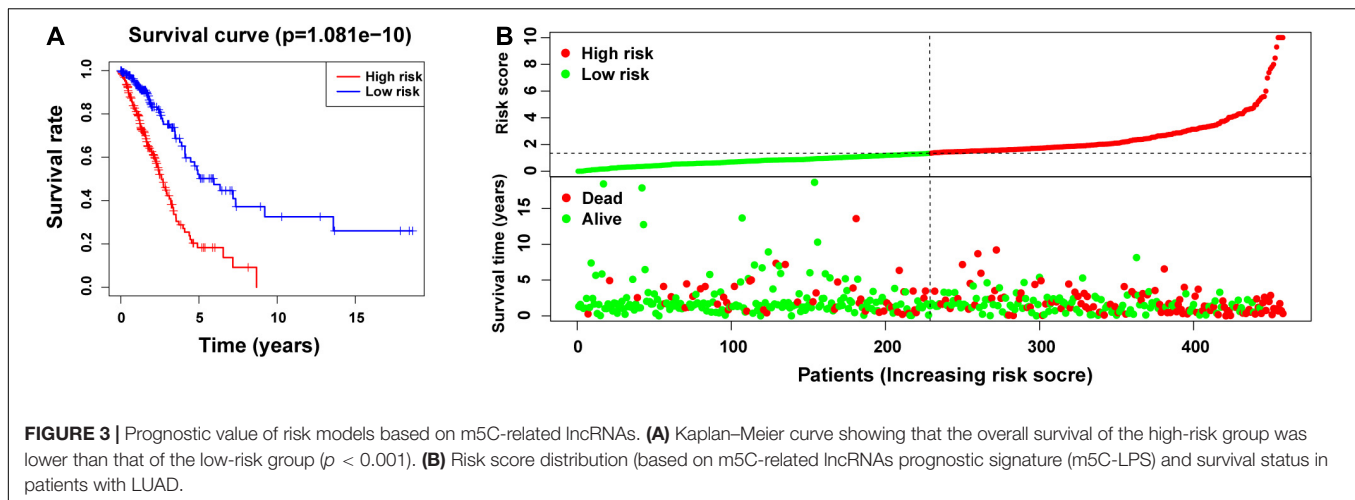
tissues, respectively. Our results showed that the expression of m5C regulators was significantly different between LUAD and adjacent normal tissues (**Figure 1A**). We found that among all genes tested, the expression of *NSUN2*, *NSUN5*, *DNMT3B*, *DNMT3A*, *ALYREF*, *DNMT1*, *NSUN6*, *NSUN4*, and *NSUN7* were significantly higher in LUAD than normal tissues ($p < 0.001$). We also observed that the expression of *TRDMT1* in cancer tissues was significantly lower than that in the normal group ($p < 0.001$), whereas no significant differences were noted between the groups in the expression of *NSUN3*, *YBX1*, and *TET2* (**Supplementary Table 2**). Next, we tried to clarify the relationship between m5C regulators. We used the STRING database to construct a PPI network of 14 m5C-related lncRNAs (**Figure 1B**), and the number of nodes were shown in **Figure 1C**. *TRDMT1* was the core gene of the network and was found to interact with 4 other genes. However, when we performed the correlation analysis, *TRDMT1* did not show a strong correlation with other genes. In addition, we observed that *NSUN4*, *NSUN7*, *NSUN6*, *DNMT3B*, *DNMT3A*, and *DNMT1* had weak to moderate correlation with other genes, whereas *NSUN3* and *TET2* had the strongest correlation (**Figure 1D**). The above results indicated that there was a certain interaction between m5C regulators in LUAD.

Identification of m5C-Related lncRNAs With Important Prognostic Value

From TCGA database, we identified 1094 m5C-related lncRNAs combined with LUAD survival data and performed univariate

Cox regression analysis to determine the m5C-related lncRNAs with important prognostic value. Accordingly, our results revealed 53 m5C-related lncRNAs associated with cancer risk ($p < 0.05$) (**Supplementary Table 3**). Among them, 46 m5C-related lncRNAs were demonstrated to be protective factor ($HR < 1$), whereas 7 of them were high-risk factors ($HR > 1$). Furthermore, we used multivariate Cox regression analysis; of the 53 m5C-related lncRNAs, 14 m5C-related lncRNAs were revealed to have prognostic value (**Table 2**). As shown in **Figures 2A,B**, we constructed a coexpression network for the visualization of 14 m5C-related lncRNAs-mRNAs in LUAD. We detected that the highest number of lncRNAs was coexpressed with *TET2* ($n = 11$), followed by *NSUN7* ($n = 8$) and *NSUN3* ($n = 7$). The highest number of mRNAs was shown to be coexpressed with *SH3BP5-AS1* ($n = 6$), with the coexpressed lncRNAs being the main protective factors. Then, we used the GEPIA online database to analyze the expression intensity of related genes in the coexpression network ($R \geq 0.35$) (**Figure 2C** and **Supplementary Figure 1**). We found that there was a weak correlation between m5C-related lncRNAs and m5C genes, among which the strongest correlation was found between *TRDMT1* and *LINC00578* ($R = 0.57$, $p < 0.001$), followed by *ALYREF* and *TMPO-AS1* ($R = 0.46$, $p < 0.001$), with the expression of m5C-related mRNAs being positively correlated with the expression of lncRNAs.

Subsequently, we divided 1344 patients with LUAD into high- and low-risk groups according to the obtained risk score formula



and median risk values. Kaplan–Meier survival analysis showed that the OS of the high-risk group was poorer than that of the low-risk group ($p < 0.001$) (Figure 3A), indicating that the risk score had a prognostic value. We then drew risk curves and scatter plots to illustrate the relationship between the risk score of patients with LUAD and the corresponding survival status (Figure 3B), which suggested that the higher the risk score, the higher the mortality rate. Hence, we identified m5C-related lncRNAs with important prognostic value and established the prognostic value of m5C-LPS based on the 14 m5C-related lncRNAs.

Correlation Between Differential Expression of m5C-Related lncRNAs and Clinicopathological Variables

As shown in Figure 4A and Supplementary Figure 2, we performed univariate analysis on the 14m5C-related lncRNAs and divided the patients into high- and low-expression groups according to the expression of single genes. We also observed significant differences in the OS of patients. Among them, the OS of patients in the high-expression group of *ABALON*, *AL034397.3*, *NKILA*, and *TMPO-AS1* was demonstrated to be lower than that in the low-expression group ($p < 0.05$). In contrast, the OS of patients in the high-expression group of *AC005911.1*, *AC090948.1*, *AC106047.1*, *AC124045.1*, *AL513550.1*, *HLA-DQB1-AS1*, *LINC00654*, *AL035701.1*, *LINC00578*, and *SH3BP5-AS1* was higher than that in the low-expression group ($p < 0.05$). According to the heatmap, the pathological stage ($p < 0.05$), T stage ($p < 0.001$), and survival status ($p < 0.001$) were significantly different between the high- and low-risk groups of m5C-LPS. However, we did not find any significant differences in gender, age, N stage and M stage (Figure 4B). Next, we performed an in-depth analysis of the correlation between m5C-related lncRNAs and clinicopathological features and found that in T stage, 5 lncRNAs exhibited significant differences across groups. Likewise, 7 lncRNAs were found to show significant differences among different groups in the N stage. Regarding M and S stages, 2 and 6 lncRNAs, respectively, were identified with significant differences among

different groups. Interestingly, we detected that the expression of *AL034397.3* varied across groups in all (T, N, M, and S) stages, suggesting that this might be the core prognostic gene (Figure 4C).

Verification of the Prognostic Model Constructed With the m5C-Related lncRNAs and Construction of the Nomogram

We used univariate and multivariate Cox regression analysis to identify whether m5C-LPS could be used as an independent prognostic factor. In the univariate Cox regression analysis, we obtained an HR = 1.157 and 95% CI: 1.118–1.197 of the risk score ($p < 0.001$), whereas in the multivariate Cox regression analysis, an HR = 1.282 and 95% CI: 1.211–1.356 of the risk score ($p < 0.001$) was achieved, indicating that the risk score as an important prognostic factor could be independent of age, gender, pathological stage, and TNM stage (Figure 5A). In order to evaluate its specificity and sensitivity for predicting the prognosis of patients with LUAD, we evaluated the area under the ROC curve (AUC) value of the risk score. We found that the AUC values of the risk score at 1, 3, and 5 years were 0.760, 0.754, and 0.779, respectively; these were demonstrated to be higher than those of other clinicopathological factors (Figure 5B). The above results indicated that m5C-LPS was a significant independent prognostic factor for patients with LUAD. At the same time, we used the risk score to construct the nomogram, which can be used as a quantitative tool to predict the prognosis of patients in clinical practice (Figure 5C).

Differences in the m5C Status of Low- and High-Risk Groups

We used PCA to detect the different distribution patterns of m5C on the genome-wide expression profile, m5C RNA modification-related gene expression profile, and m5C-LPS related lncRNAs expression profile between the high- and low-risk groups (Figure 6A). Based on m5C-LPS, patients were divided into low- and high-risk groups, which were distributed in different

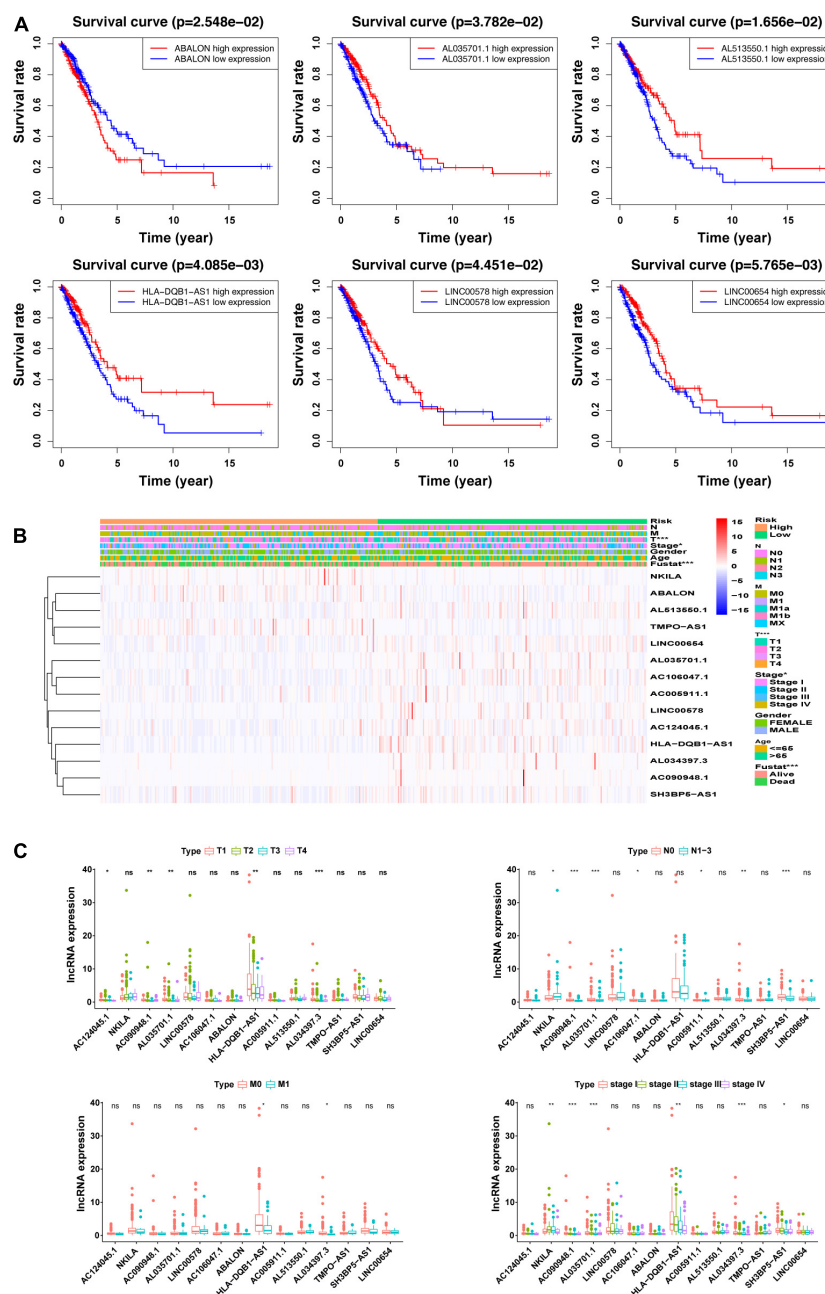


FIGURE 4 | Correlation between the expression of the 14 m5C-related lncRNAs and clinicopathological factors. **(A)** overall survival (OS) analysis of the 14 m5C-related lncRNAs in the GEPIA database between high- and low-expression groups. **(B)** Heatmap showing clinicopathological features and differences in gene expression in the high- and low-risk group. **(C)** Differences in the expression of the 14 m5C-related lncRNAs in T, N, M, and S stage groups. * $p < 0.5$, ** $p < 0.01$, and *** $p < 0.001$. ns, no sense.

directions in a more obvious manner than when the other 2 methods were used. Hence, m5C-LPS could divide patients with LUAD into 2 groups, indicating the sensitivity and specificity of the predictive model. To identify the abnormally activated signal pathways of m5C-related lncRNAs, we conducted a GSEA. the results of our analysis showed that the high expression of m5C-related lncRNAs was related to cell cycle signaling pathways, DNA replication, and P53 signaling pathways (Figure 6B).

Verification of Expression Level *in vitro* and Predicting m5C Modification Sites on 14 Prognostic lncRNAs

In order to verify the expression level of prognostic m5C-related lncRNAs in LUAD cells, we used RT-qPCR analysis to detect BEAS-2B and LUAD cells (Figure 6C). Among them, 9 lncRNAs (AC090948.1, AC124045.1, AL513550.1, AC005911.1,

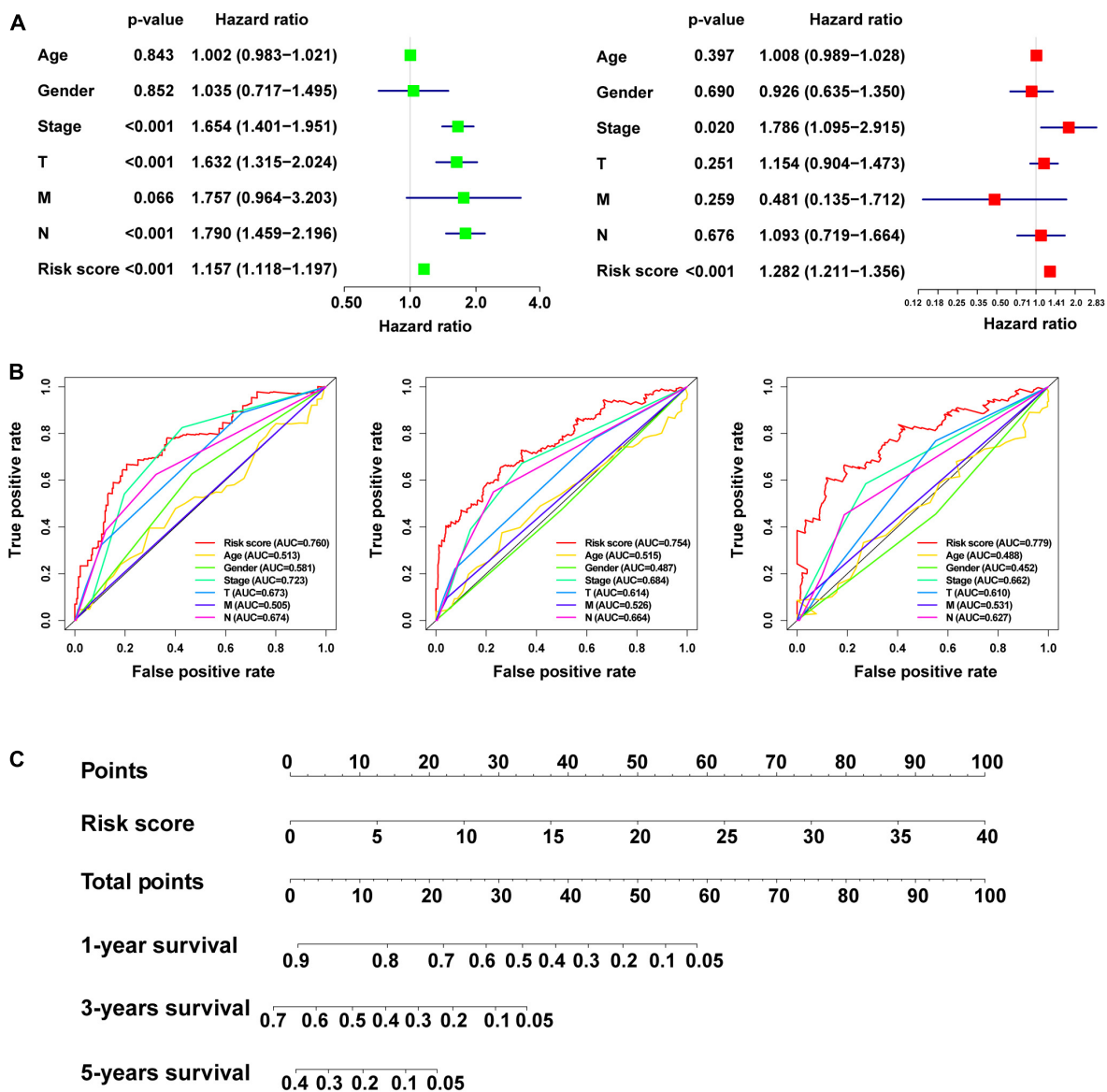


FIGURE 5 | Verification of the risk model and construction of the nomogram. **(A)** Univariate and multivariate Cox regression analysis of the prognostic value of risk scores and clinical features. **(B)** Determination of the area under the ROC curve (AUC) of the risk score and clinical characteristics based on the ROC curve. **(C)** Construction of the nomogram based on the risk score.

AL034397.3, LINC00578, AL035701.1, HLA-DQB1-AS1, and SH3BP5-AS1) were downregulated in both LUAD cells, combined with Figure 4A, Supplementary Figure 2, and Table 2, their high expression were associated with better survival, $HR < 1$, considering that they might play a role as tumor suppressor genes. However, LINC00654, AC106047.1, and NKILA were downregulated in PC-9 cells and upregulated in H1299 cells. ABALON and TMPO-AS1 were downregulated in both LUAD cells, However, their low expression were associated with better survival, $HR > 1$, so the specific mechanism needed to be further studied. We further studied the m5C modification sites on prognostic m5C-related lncRNAs using online databases, and it was found that there were m5C modification sites on 14

lncRNAs (Supplementary Table 4). However, there were some differences in the results of the three databases, which might be caused by different calculation and prediction methods.

Correlation Between m5C-LPS and Tumor Immune Microenvironment of LUAD

In this study, we analyzed the importance of TIM based on the risk characteristics of m5C-LPS. By using CIBERSORT to screen for values of $p < 0.05$, CD4 naive T-cells were not expressed in the high and low risk groups and were excluded (Figure 7A). Therefore, we analyzed the correlation

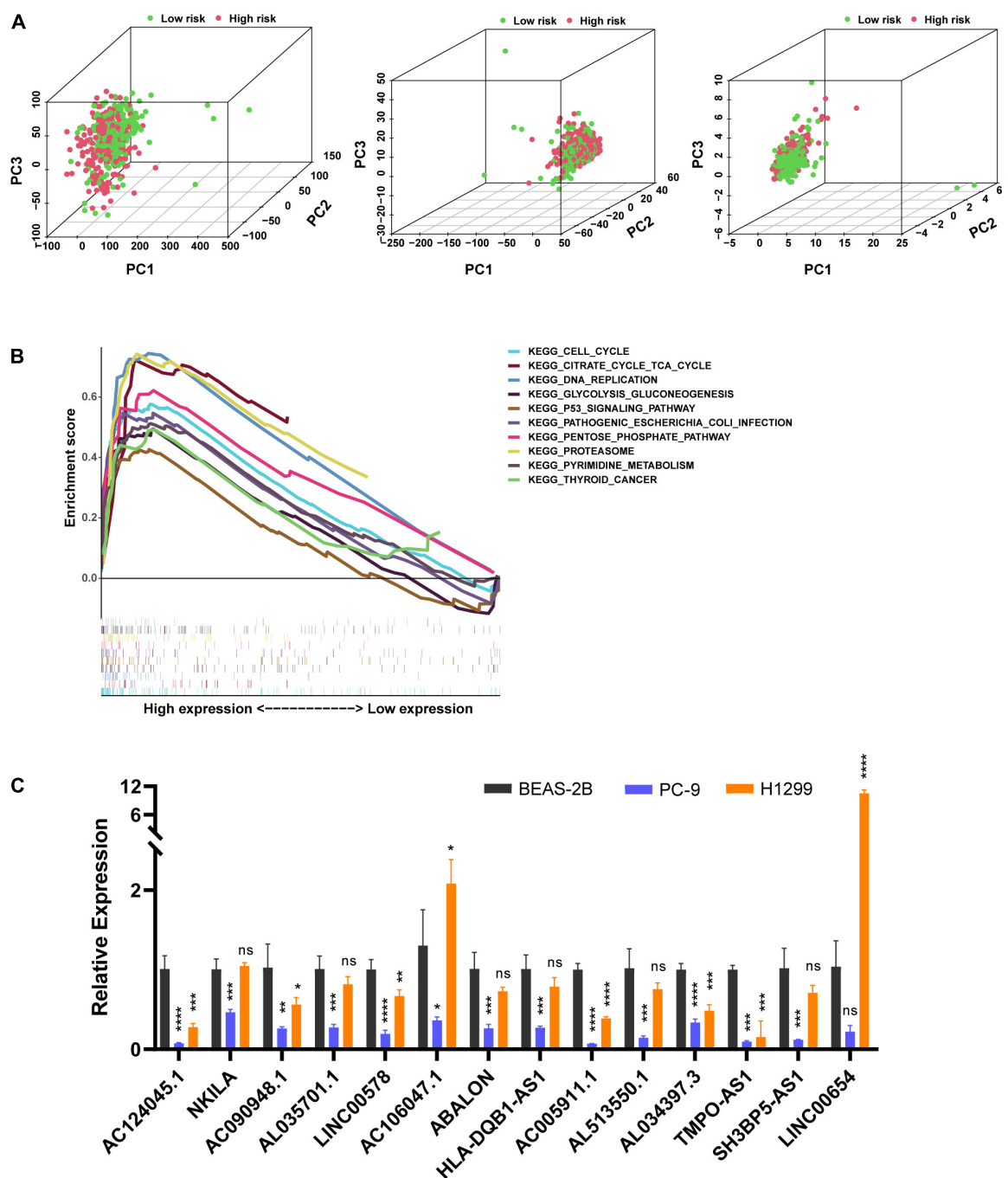


FIGURE 6 | The m5C status was different between the high- and low-risk groups and the expression of m5C-related lncRNAs differed *in vitro*. **(A)** Principal component analysis (PCA) was performed for the low- and high-risk groups based on the whole genome and m5C-related coding genes, and a risk model was constructed using 14 m5C-related lncRNAs. **(B)** Gene set enrichment analysis of m5C-related lncRNAs. **(C)** Expression level of m5C-related lncRNAs in BEAS-2B cells and LUAD cells. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, and **** $p < 0.0001$. ns, no sense.

between 21 tumor-infiltrating immune cells and m5C-LPS. We accordingly found that among them, monocytes, dendritic cell-resting, and mast cell-resting exhibited a higher expression in the high-risk group than in the low-risk group ($p < 0.05$). In contrast, the levels of activated M0 macrophages and mast

cells in the high-risk group were lower than those in the low-risk group ($p < 0.05$) (Figure 7B). In addition, we further analyzed the correlation between 21 types of tumor-infiltrating cells. We identified that activated CD4 memory T-cells had the strongest correlation with CD8 T-cells ($r = 0.55$), followed

by plasma cells and M2 macrophages ($r = -0.39$) (Figure 7C). Finally, we studied the relationship between the m5C-LPS risk score and tumor-infiltrating immune cells. Our results showed that the risk score was negatively correlated with 6 tumor-infiltrating immune cell subtypes (B-cells, CD4⁺T-cells, CD8⁺T-cells, macrophages, neutrophils, and dendritic cells; $p < 0.05$) (Figure 7D). These findings indicated that the risk characteristics of m5C-LPS could distinguish different characteristics of tumor immune cells in LUAD.

DISCUSSION

Although active multimodal therapy (surgery, chemotherapy, radiotherapy, targeted therapy, and immunotherapy) has greatly improved the survival of patients with LUAD, the outcome of treatment remains unsatisfactory. Patients with similar clinical risk factors have very different prognoses and tumor responses to treatment. Thus, finding effective therapeutic targets is of great value for the diagnosis and treatment of LUAD. More than 150

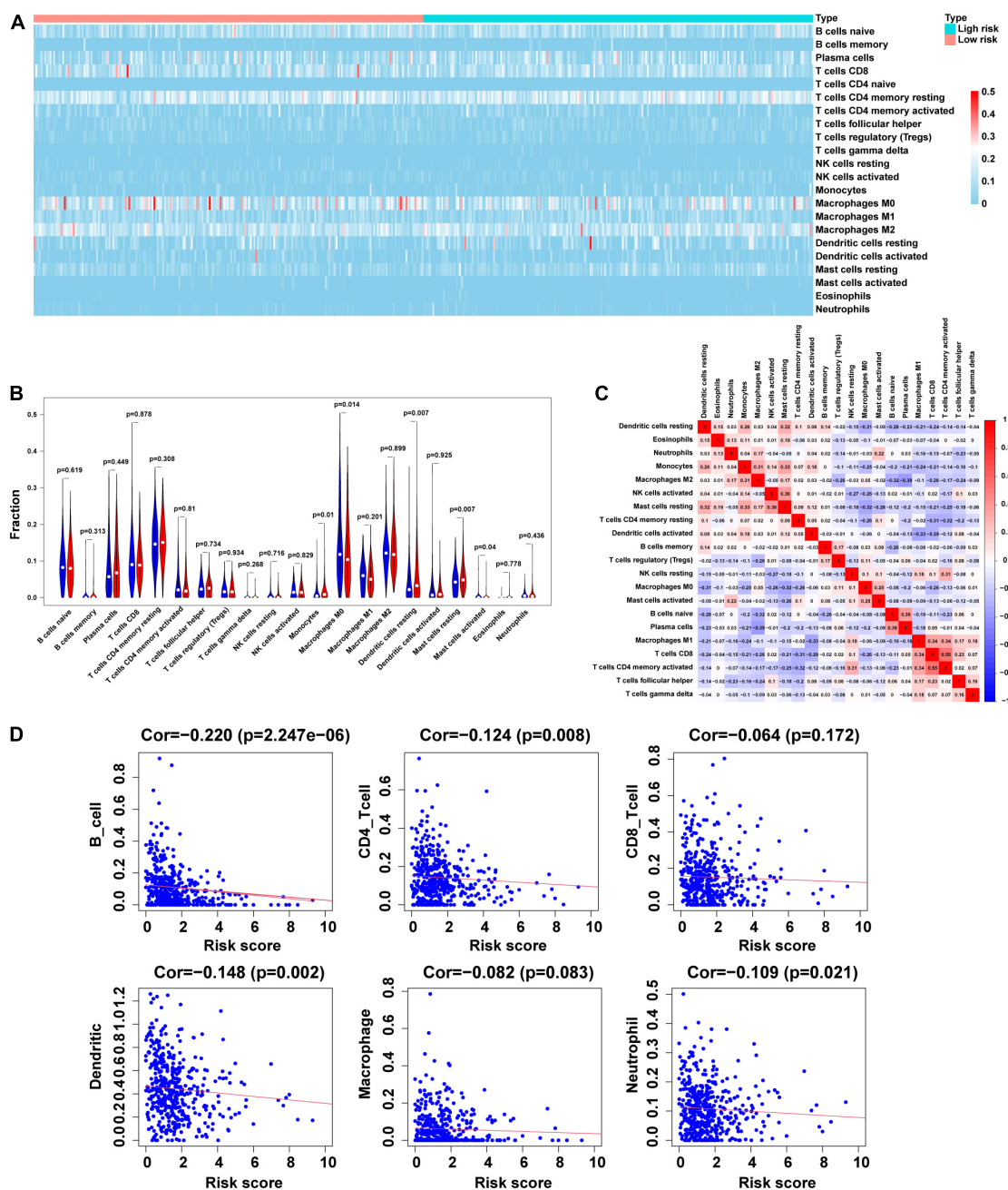


FIGURE 7 | Correlation between tumor-infiltrating immune cells and risk model. (A,B) Heatmap and violin plot of 21 tumor-infiltrating immune cell types in low- and high-risk groups. (C) Spearman correlation analysis of 21 tumor-infiltrating immune cells. (D) Correlation of risk score with 6 tumor-infiltrating immune cell subtypes.

RNA post-transcriptional modifications have been discovered, with m6A, m5C, and m1A being the most common ones. They have been reported to play an important role in regulating gene expression and disease progression. However, abnormal RNA modification has been shown to cause a series of diseases, including cancer (Esteve-Puig et al., 2020). In this study, we used TCGA database and STRING to analyze the differences in the expression of 13 m5C regulators in LUAD and protein interactions, respectively. To the best of our knowledge, this is the first in-depth analysis of the role of m5C regulators in LUAD.

Long non-coding RNAs are widely expressed in human cells, playing a key role in various biological events, such as genome expression and cell differentiation (Wahlestedt, 2013; Luo et al., 2015; Ballantyne et al., 2016). Increasing evidence have indicated that the abnormal expression of lncRNAs might be related to the occurrence and development of a variety of cancers (Dong et al., 2018; Zhuang et al., 2019; Zhang et al., 2020). For example, the m6A “writer” METTL14 was reported to inhibit the proliferation and metastasis of colorectal cancer through the downregulation of the oncogenic XIST lncRNA (Yang et al., 2020). The lncRNA activating regulator of DKK1 (LNCAROD) was shown to be stabilized by m6A methylation in HNSCC cells and suggested to be used as a scaffold to promote the YBX1-HSPA1A PPI and stability of the YBX1 protein, resulting in the cell proliferation and migration of head and neck squamous cell cancer (Ban et al., 2020). At present, few studies exist on m5C-related lncRNAs. Studies have performed quantitative mapping of the m5C sites in *Arabidopsis thaliana* on a transcriptome range, and found more than 1000 m5C sites in mRNA, long non-coding RNA and other non-coding RNAs (David et al., 2017). He et al. (2020a) found that compared with adjacent non-tumor tissues, lncRNA m5C methylation occurs more frequently in Hepatocellular Carcinoma (HCC), and more methylated genes are up-regulated. KEGG pathway enrichment analysis show that hypermethylated genes are closely related to cancer pathways. It is speculated that m5C may be involved in the occurrence and development of HCC. Recent studies have found that H19 lncRNA modified by m5A “writer” NSUN2 might promote the occurrence and development of hepatocellular carcinoma by recruiting the G3BP1 oncoprotein and MYC protein (Sun Z. et al., 2020). However, there have been no reports on the role of m5C-related lncRNA in LUAD.

In this study, we analyzed the correlation between the expression of 1094 m5C-related lncRNAs and the prognosis of LUAD. In addition, 14 m5C-related lncRNAs with prognostic value were identified. Using multivariate Cox and risk scoring methods, we developed a m5C-LPS that divided patients with LUAD into high- and low-risk groups with significantly different OS. In addition, univariate and multivariate Cox analysis and AUC values of ROC curves confirmed that the risk score based on the expression of the 14 m5C-related lncRNAs could predict patient prognosis independently of traditional clinical risk factors and molecular characteristics. The currently constructed LUAD risk scoring model was mainly based on whole-genome sequencing, including lncRNA, miRNA, and mRNA. When exploring the potential of lncRNAs as novel

tumor biomarkers, previous studies have focused on single molecules. However, a lncRNA is not enough to serve as a new type of malignant tumor biomarker. As lncRNAs have been reported to be released in various body fluids, including serum, saliva, and urine, it could be assumed that lncRNAs exist stably in human serum/plasma. A number of studies have reported on the differences in the expression of lncRNAs in serum/plasma as a new type of biomarker in assessing patients. However, this is the first study to report a LUAD risk score model based on 14 m5C-related lncRNAs with prognostic value. In this study, 14 m5C-related lncRNAs were selected, 4 of which have been previously studied in NSCLC (Lu et al., 2017; Zhou et al., 2019; Jin et al., 2020). Bioinformatic analysis showed that HLA-DQB1-AS1 and AL034397.3 were immune-related lncRNAs in LUAD, and the constructed risk model had prognostic value.

Furthermore, lncRNAs are known to play a vital role in the TIM. The value of immune-related lncRNAs has been shown in many cancers, such as hepatocellular carcinoma (HCC), diffuse large B-cell lymphoma, and breast cancer. Currently, a number of studies have focused on 7 immune-related lncRNAs (AC022784-1, NKILA, AC026355-1, AC068338-3, LINC01843, SYNPR-AS1, and AC123595-1) to construct a LUAD risk score model, which could effectively predict clinical prognosis. In this study, we further analyzed the correlation between m5C-LPS and the distribution of tumor-infiltrating immune cells. We found that the risk score was negatively correlated with the degree of these 6 immune cell infiltrates. Our results show that the risk characteristics of m5C-LPS could distinguish different characteristics of tumor-infiltrating immune cells in LUAD. As such, this study is the first to analyze the role of m5C-related lncRNAs in LUAD and their correlation with TIM.

However, we also recognize some limitations of this study. For example, the dataset used for the initial analysis was relatively insufficient. We only downloaded the data from TCGA and failed to retrieve other lncRNA expression levels supporting LUAD, patient clinicopathological characteristics, survival, and follow-up data. In addition, the nomogram we constructed only included the risk score, because the patients of stage I and II were far more than those of stage III and IV in the TCGA, so we excluded the stage and constructed a nomogram using risk score. We only validated the prognostic m5C-related lncRNAs at the cytological level. The constructed risk scoring model needs to be validated by tissue level and *in vivo* experiments to make the prediction results more reliable. Additionally, there was a lack of experiments such as MeRIP-seq, RNA-BisSeq, Aza-IP, miCLIP and other experiments to further confirm m5C modification sites on 14 lncRNAs. We will incorporate this work into future research.

CONCLUSION

The m5C-LPS we constructed was demonstrated to have an independent prognostic value and high reliability and might provide some clues for further research on the mechanism of the m5C post-transcriptional modification of lncRNAs, and at

the same time bestow new resources for a better understanding of the mechanism of immune cell-specific genes involved in the regulation of cancer.

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

JP and YX designed this experiment. JP and ZH were responsible for literature review, data collection, analysis, and writing. YX

was responsible for modification. All authors contributed to this article and approved the submitted version.

ACKNOWLEDGMENTS

The authors thank the TCGA dataset for providing high quality data.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fcell.2021.671821/full#supplementary-material>

REFERENCES

- Atianand, M. K., Caffrey, D. R., and Fitzgerald, K. A. (2017). Immunobiology of Long Noncoding RNAs. *Annu. Rev. Immunol.* 35, 177–198. doi: 10.1146/annurev-immunol-041015-055459
- Ballantyne, R. L., Zhang, X., Nuñez, S., Xue, C., Zhao, W., Reed, E., et al. (2016). Genome-wide interrogation reveals hundreds of long intergenic noncoding RNAs that associate with cardiometabolic trait. *Hum. Mol. Genet.* 25, 3125–3141.
- Ban, Y., Tan, P., Cai, J., Li, J., Hu, M., Zhou, Y., et al. (2020). LNCAROD is stabilized by m6A methylation and promotes cancer progression via forming a ternary complex with HSPA1A and YBX1 in head and neck squamous cell carcinoma. *Mol. Oncol.* 14, 1282–1296. doi: 10.1002/1878-0261.12676
- Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R. L., Torre, L. A., and Jemal, A. (2018). Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* 68, 394–424. doi: 10.3322/caac.21492
- Chen, H., Yao, J., Bao, R., Dong, Y., Zhang, T., Du, Y., et al. (2021). Cross-talk of four types of RNA modification writers defines tumor microenvironment and pharmacogenomic landscape in colorectal cancer. *Mol. Cancer* 20:29. doi: 10.1186/s12943-021-01322-w
- Chen, K., Song, B., Tang, Y., Wei, Z., Xu, Q., Su, J., et al. (2021). RMDisease: a database of genetic variants that affect RNA modifications, with implications for epitranscriptome pathogenesis. *Nucleic Acids Res.* 49, D1396–D1404. doi: 10.1093/nar/gkaa790
- Chen, K., Wei, Z., Zhang, Q., Wu, X., Rong, R., Lu, Z., et al. (2019). WHISTLE: a high-accuracy map of the human N6-methyladenosine (m6A) epitranscriptome predicted using a machine learning approach. *Nucleic Acids Res.* 47:e41. doi: 10.1093/nar/gkz074
- Chen, X., Li, A., Sun, B. F., Yang, Y., Han, Y. N., Yuan, X., et al. (2019). 5-methylcytosine promotes pathogenesis of bladder cancer through stabilizing mRNAs. *Nat. Cell Biol.* 21, 978–990. doi: 10.1038/s41556-019-0361-y
- Cohen, D., Hondelink, L. M., Solleveld-Westerink, N., Uljee, S. M., Ruano, D., Cleton-Jansen, A. M., et al. (2020). Optimizing Mutation and Fusion Detection in NSCLC by Sequential DNA and RNA Sequencing. *J. Thorac. Oncol.* 15, 1000–1014. doi: 10.1016/j.jtho.2020.01.019
- David, R., Burgess, A., Parker, B., Li, J., Pulsford, K., Sibbritt, T., et al. (2017). Transcriptome-Wide Mapping of RNA 5-Methylcytosine in Arabidopsis mRNAs and Noncoding RNAs. *Plant Cell* 29, 445–460. doi: 10.1105/tpc.16.00751
- Dong, H. X., Wang, R., Jin, X. Y., Zeng, J., and Pan, J. (2018). lncRNA DGCR5 promotes lung adenocarcinoma (LUAD) progression via inhibiting hsa-mir-22-3p. *J. Cell Physiol.* 233, 4126–4136. doi: 10.1002/jcp.26215
- Dou, L., Li, X., Ding, H., Xu, L., and Xiang, H. (2020). Prediction of m5C Modifications in RNA Sequences by Combining Multiple Sequence Features. *Mol. Ther. Nucleic Acids* 21, 332–342. doi: 10.1016/j.omtn.2020.06.004
- Esteve-Puig, R., Bueno-Costa, A., and Esteller, M. (2020). Writers, readers and erasers of RNA modifications in cancer. *Cancer Lett.* 474, 127–137. doi: 10.1016/j.canlet.2020.01.021
- Guo, X., Xiao, H., Guo, S., Li, J., Wang, Y., Chen, J., et al. (2019). Long noncoding RNA HOTAIR knockdown inhibits autophagy and epithelial-mesenchymal transition through the Wnt signaling pathway in radioresistant human cervical cancer HeLa cells. *J. Cell Physiol.* 234, 3478–3489. doi: 10.1002/jcp.26828
- He, Y., Shi, Q., Zhang, Y., Yuan, X., and Yu, Z. (2020a). Transcriptome-Wide 5-Methylcytosine Functional Profiling of Long Non-Coding RNA in Hepatocellular Carcinoma. *Cancer Manag. Res.* 12, 6877–6885. doi: 10.2147/cmar.S262450
- He, Y., Yu, X., Li, J., Zhang, Q., Zheng, Q., and Guo, W. (2020b). Role of m5C-related regulatory genes in the diagnosis and prognosis of hepatocellular carcinoma. *Am. J. Transl. Res.* 12, 912–922.
- Jin, D., Song, Y., Chen, Y., and Zhang, P. (2020). Identification of a Seven-lncRNA Immune Risk Signature and Construction of a Predictive Nomogram for Lung Adenocarcinoma. *Biomed Res. Int.* 2020, 1–17. doi: 10.1155/2020/7929132
- Kopp, F., and Mendell, J. T. (2018). Functional Classification and Experimental Dissection of Long Noncoding RNAs. *Cell* 172, 393–407. doi: 10.1016/j.cell.2018.01.011
- Lei, X., Lei, Y., Li, J. K., Du, W. X., Li, R. G., Yang, J., et al. (2020). Immune cells within the tumor microenvironment: biological functions and roles in cancer immunotherapy. *Cancer Lett.* 470, 126–133. doi: 10.1016/j.canlet.2019.11.009
- Li, J., Huang, Y., Yang, X., Zhou, Y., and Zhou, Y. (2018). RNAm5Cfinder: a Web-server for Predicting RNA 5-methylcytosine (m5C) Sites Based on Random Forest. *Sci. Rep.* 8:17299. doi: 10.1038/s41598-018-35502-4
- Li, Q., Li, X., Tang, H., Jiang, B., Dou, Y., Gorospe, M., et al. (2017). NSUN2-Mediated m5C Methylation and METTL3/METTL14-Mediated m6A Methylation Cooperatively Enhance p21 Translation. *J. Cell. Biochem.* 118, 2587–2598. doi: 10.1002/jcb.25957
- Li, T., Fan, J., Wang, B., Traugh, N., Chen, Q., Liu, J. S., et al. (2017). TIMER: a Web Server for Comprehensive Analysis of Tumor-Infiltrating Immune Cells. *Cancer Res.* 77, e108–e110. doi: 10.1158/0008-5472.CAN-17-0307
- Li, T., Fu, J., Zeng, Z., Cohen, D., Li, J., Chen, Q., et al. (2020). TIMER2.0 for analysis of tumor-infiltrating immune cells. *Nucleic Acids Res.* 48, W509–W514. doi: 10.1093/nar/gkaa407
- Liu, J., Cho, S. N., Akkanti, B., Jin, N., Mao, J., Long, W., et al. (2015). Erbb2 Pathway Activation upon Smad4 Loss Promotes Lung Tumor Growth and Metastasis. *Cell Rep.* 10, 1599–1613. doi: 10.1016/j.celrep.2015.02.014
- Liu, K., Cao, L., Du, P., and Chen, W. (2020). im6A-TS-CNN: identifying the N(6)-Methyladenine Site in Multiple Tissues by Using the Convolutional Neural Network. *Mol. Ther. Nucleic Acids* 21, 1044–1049. doi: 10.1016/j.omtn.2020.07.034
- Liu, K., and Chen, W. (2020). iMRM: a platform for simultaneously identifying multiple kinds of RNA modifications. *Bioinformatics* 36, 3336–3342. doi: 10.1093/bioinformatics/btaa155

- Lu, Z., Li, Y., Wang, J., Che, Y., Sun, S., Huang, J., et al. (2017). Long non-coding RNA NKILA inhibits migration and invasion of non-small cell lung cancer via NF-kappaB/Snail pathway. *J. Exp. Clin. Cancer Res.* 36:54. doi: 10.1186/s13046-017-0518-0
- Luo, G., Wang, M., Wu, X., Tao, D., Xiao, X., Wang, L., et al. (2015). Long Non-Coding RNA MEG3 Inhibits Cell Proliferation and Induces Apoptosis in Prostate Cancer. *Cell. Physiol. Biochem.* 37, 2209–2220. doi: 10.1159/000438577
- Lv, H., Zhang, Z. M., Li, S. H., Tan, J. X., Chen, W., and Lin, H. (2020). Evaluation of different computational methods on 5-methylcytosine sites identification. *Brief. Bioinform.* 21, 982–995. doi: 10.1093/bib/bbz048
- Osmani, L., Askin, F., Gabrielson, E., and Li, Q. K. (2018). Current WHO guidelines and the critical role of immunohistochemical markers in the subclassification of non-small cell lung carcinoma (NSCLC): moving from targeted therapy to immunotherapy. *Semin. Cancer Biol.* 52, 103–109. doi: 10.1016/j.semcancer.2017.11.019
- Qin, Y., Liu, X., Pan, L., Zhou, R., and Zhang, X. (2019). Long noncoding RNA MIR155HG facilitates pancreatic cancer progression through negative regulation of miR-802. *J. Cell. Biochem.* 120, 17926–17934. doi: 10.1002/jcb.29060
- Qiu, W. R., Jiang, S. Y., Xu, Z. C., Xiao, X., and Chou, K. C. (2017). iRNA5C-PseDNC: identifying RNA 5-methylcytosine sites by incorporating physical-chemical properties into pseudo dinucleotide composition. *Oncotarget* 8, 41178–41188. doi: 10.18632/oncotarget.17104
- Ranzani, V., Rossetti, G., Panzeri, I., Arrighi, A., Bonnal, R. J., Curti, S., et al. (2015). The long intergenic noncoding RNA landscape of human lymphocytes highlights the regulation of T cell differentiation by linc-MAF-4. *Nat. Immunol.* 16, 318–325. doi: 10.1038/ni.3093
- Song, B., Tang, Y., Chen, K., Wei, Z., Rong, R., Lu, Z., et al. (2020). m7GHub: deciphering the location, regulation and pathogenesis of internal mRNA N7-methylguanosine (m7G) sites in human. *Bioinformatics* 36, 3528–3536. doi: 10.1093/bioinformatics/btaa178
- Squires, J. E., Patel, H. R., Nusch, M., Sibbritt, T., Humphreys, D. T., Parker, B. J., et al. (2012). Widespread occurrence of 5-methylcytosine in human coding and non-coding RNA. *Nucleic Acids Res.* 40, 5023–5033. doi: 10.1093/nar/gks144
- Sun, J., Zhang, Z., Bao, S., Yan, C., Hou, P., Wu, N., et al. (2020). Identification of tumor immune infiltration-associated lncRNAs for improving prognosis and immunotherapy response of patients with non-small cell lung cancer. *J. Immunother. Cancer* 8:e000110. doi: 10.1136/jitc-2019-000110
- Sun, L., Liu, W. K., Du, X. W., Liu, X. L., Li, G., Yao, Y., et al. (2020). Large-scale transcriptome analysis identified RNA methylation regulators as novel prognostic signatures for lung adenocarcinoma. *Ann. Transl. Med.* 8:751. doi: 10.21037/atm-20-3744
- Sun, Z., Xue, S., Zhang, M., Xu, H., Hu, X., Chen, S., et al. (2020). Aberrant NSUN2-mediated m(5)C modification of H19 lncRNA is associated with poor differentiation of hepatocellular carcinoma. *Oncogene* 39, 6906–6919. doi: 10.1038/s41388-020-01475-w
- Tang, Y., Chen, K., Song, B., Ma, J., Wu, X., Xu, Q., et al. (2021). m6A-Atlas: a comprehensive knowledgebase for unraveling the N6-methyladenosine (m6A) epitranscriptome. *Nucleic Acids Res.* 49, D134–D143. doi: 10.1093/nar/gkaa692
- Wahlestedt, C. (2013). Targeting long non-coding RNA to therapeutically upregulate gene expression. *Nat. Rev. Drug Discov.* 12, 433–446. doi: 10.1038/nrd4018
- Wang, L., Felts, S. J., Van Keulen, V. P., Scheid, A. D., Block, M. S., Markovic, S. N., et al. (2018). Integrative Genome-Wide Analysis of Long Noncoding RNAs in Diverse Immune Cell Types of Melanoma Patients. *Cancer Res.* 78, 4411–4423. doi: 10.1158/0008-5472.CAN-18-0529
- Xu, M., Xu, X., Pan, B., Chen, X., Lin, K., Zeng, K., et al. (2019). lncRNA SATB2-AS1 inhibits tumor metastasis and affects the tumor immune cell microenvironment in colorectal cancer by regulating SATB2. *Mol. Cancer* 18:135. doi: 10.1186/s12943-019-1063-6
- Xu, Q., Chen, K., and Meng, J. (2021). WHISTLE: a Functionally Annotated High-Accuracy Map of Human m(6)A Epitranscriptome. *Methods Mol. Biol.* 2284, 519–529. doi: 10.1007/978-1-0716-1307-8_28
- Xue, M., Shi, Q., Zheng, L., Li, Q., Yang, L., and Zhang, Y. (2020). Gene signatures of m5C regulators may predict prognoses of patients with head and neck squamous cell carcinoma. *Am. J. Transl. Res.* 12, 6841–6852.
- Yang, X., Zhang, S., He, C., Xue, P., Zhang, L., He, Z., et al. (2020). METTL14 suppresses proliferation and metastasis of colorectal cancer by down-regulating oncogenic long non-coding RNA XIST. *Mol. Cancer* 19:46. doi: 10.1186/s12943-020-1146-4
- Zhang, J., Guo, S., Piao, H. Y., Wang, Y., Wu, Y., Meng, X. Y., et al. (2019). ALKBH5 promotes invasion and metastasis of gastric cancer by decreasing methylation of the lncRNA NEAT1. *J. Physiol. Biochem.* 75, 379–389. doi: 10.1007/s13105-019-00690-8
- Zhang, X. Z., Mao, H. L., Zhang, S. J., Sun, L., Zhang, W. J., Chen, Q. Z., et al. (2020). lncRNA PCAT18 inhibits proliferation, migration and invasion of gastric cancer cells through miR-135b suppression to promote CLDN11 expression. *Life Sci.* 249:117478. doi: 10.1016/j.lfs.2020.117478
- Zhao, B. S., Roundtree, I. A., and He, C. (2017). Post-transcriptional gene regulation by mRNA modifications. *Nat. Rev. Mol. Cell Biol.* 18, 31–42. doi: 10.1038/nrm.2016.132
- Zhou, W., Liu, T., Saren, G., Liao, L., Fang, W., and Zhao, H. (2019). Comprehensive analysis of differentially expressed long non-coding RNAs in non-small cell lung cancer. *Oncol. Lett.* 18, 1145–1156. doi: 10.3892/ol.2019.10414
- Zhuang, C., Ma, Q., Zhuang, C., Ye, J., Zhang, F., and Gui, Y. (2019). lncRNA GClnc1 promotes proliferation and invasion of bladder cancer through activation of MYC. *FASEB J.* 33, 11045–11059. doi: 10.1096/fj.201900078RR
- Zuo, X., Chen, Z., Gao, W., Zhang, Y., Wang, J., Wang, J., et al. (2020). M6A-mediated upregulation of LINC00958 increases lipogenesis and acts as a nanotherapeutic target in hepatocellular carcinoma. *J. Hematol. Oncol.* 13:5. doi: 10.1186/s13045-019-0839-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 Pan, Huang and Xu. 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.



N6-Methyladenosine Modification and Its Regulation of Respiratory Viruses

Qianyu Feng^{1,2}, Hongwei Zhao^{1,2}, Lili Xu^{1,2*} and Zhengde Xie^{1,2}

¹ Beijing Key Laboratory of Paediatric Respiratory Infection Diseases, Key Laboratory of Major Diseases in Children, National Key Discipline of Paediatrics (Capital Medical University), Ministry of Education, National Clinical Research Center for Respiratory Diseases, National Center for Children's Health, Beijing Paediatric Research Institute, Beijing Children's Hospital, Capital Medical University, Beijing, China, ² Research Unit of Critical Infection in Children, Chinese Academy of Medical Sciences, Beijing, China

OPEN ACCESS

Edited by:

Xiao Han,
Chinese Academy of Agricultural
Sciences, China

Reviewed by:

Lei Gu,
Max Planck Institute for Heart
and Lung Research, Germany
Lei Sun,
Shanghai Jiao Tong University, China
Jianmin Wang,
Institute of Pathogen Biology (CAMS),
China

*Correspondence:

Lili Xu
justinexull26@163.com

Specialty section:

This article was submitted to
Epigenomics and Epigenetics,
a section of the journal
Frontiers in Cell and Developmental
Biology

Received: 25 April 2021

Accepted: 28 June 2021

Published: 23 July 2021

Citation:

Feng Q, Zhao H, Xu L and Xie Z
(2021) N6-Methyladenosine
Modification and Its Regulation
of Respiratory Viruses.
Front. Cell Dev. Biol. 9:699997.
doi: 10.3389/fcell.2021.699997

N6-methyladenosine (m⁶A) is a ubiquitous RNA modification in eukaryotes. It plays important roles in the translocation, stabilization and translation of mRNA. Many recent studies have shown that the dysregulation of m⁶A modification is connected with diseases caused by pathogenic viruses, and studies on the role of m⁶A in virus-host interactions have shown that m⁶A plays a wide range of regulatory roles in the life cycle of viruses. Respiratory viruses are common pathogens that can impose a large disease burden on young children and elderly people. Here, we review the effects of m⁶A modification on respiratory virus replication and life cycle and host immunity against viruses.

Keywords: N6-methyladenosine, RNA modification, respiratory syncytial virus, adenovirus, human metapneumovirus, influenza A virus

INTRODUCTION

RNA posttranscriptional modification is very common in eukaryotes, and more than one hundred types of modifications have been reported to date (Boccaletto et al., 2018). Most of these modifications are deposited on non-coding RNAs, such as ribosomal RNA (rRNA) and transfer RNA (tRNA), which play roles in regulating RNA structure, function, and translation (Cantara et al., 2010; Boccaletto et al., 2018). Among these modifications, m⁶A is the most common modification to poly(a) RNA components, and it is thought to be involved in the processing of messenger RNA (mRNA) (Desrosiers et al., 1974). The m⁶A modification involves a dynamically reversible process that is catalyzed by methyltransferases (writers) and eliminated by erasers. In addition, m⁶A can interact with m⁶A-binding protein (reader) or indirectly modulate the structure of RNA to regulate the interaction between RNA reader proteins (Jia et al., 2011; Liu et al., 2014; Ping et al., 2014). Recent evidence has shown that m⁶A participates in multiple stages of mRNA life, such as RNA folding and structure, maturation, stability, splicing, export, translation, and decay, and it plays an important role in these stages (Liu and Zhang, 2018).

In recent years, increasing numbers of studies have been conducted on the role of the m⁶A modification in different pathogenic viruses, with results revealing its impact on the regulation of the viral life cycle, such as its influence on the expression of specific genes related to the viral life cycle and its possible inhibition or promotion of the replication of various pathogenic viruses (Imam et al., 2020).

With the spread of all kinds of respiratory virus infections worldwide, respiratory virus infections have received increased attention. Respiratory virus infection can cause pneumonia in children, and severe cases can lead to acute respiratory distress syndrome, toxic encephalopathy, heart failure, etc., and in the elderly, it can aggravate underlying lung diseases and be life threatening. Respiratory viruses mainly include influenza virus, respiratory syncytial virus, and adenovirus. Influenza and respiratory syncytial virus cause nearly 300,000 deaths in children under 5 years old every year, while adenovirus and other viruses are associated with high morbidity and mortality rates (Nair et al., 2010; Wang et al., 2020). In particular, several emerging respiratory viruses with potentially epidemic characteristics, such as SARS-CoV-2, which causes coronavirus disease 2019 (COVID-19), threaten the world. The burden of disease caused by respiratory viruses has received increasing attention.

In this paper, the role of the m⁶A modification in the replication of respiratory viruses and its effect on the host immune response are discussed.

RNA m⁶A Modifications and Protein Factors

The m⁶A modification is regulated by methyltransferase, demethylase and m⁶A RNA-binding proteins. Methyltransferases play roles as “writers” and include methyltransferase-like 3 (METTL3), METTL14, and Wilm’s tumor-associated protein (WTAP). Among these proteins, METTL3 and METTL14 form heterodimers and promote intracellular mRNA methylation (Liu et al., 2014), but only METTL3 has methyltransferase activity, with METTL14 involved in substrate recognition (Śledź and Jinek, 2016; Wang P. et al., 2016; Wang X. et al., 2016). In addition, METTL14 can provide RNA-binding scaffolds, participate in allosteric activation and enhance the catalytic function of METTL3 (Śledź and Jinek, 2016). WTAP drives the m⁶A “writer” complex to the splice site by interacting with the METTL3/METTL14 heterodimer (Liu et al., 2014; Wu et al., 2016). Recently, two other members of the m⁶A methyltransferase complex, RBM15, and its homologous partner, RBM15b, have been shown to recruit the METTL3/METTL4 protein complex to target RNA for selective methylation (Patil et al., 2016).

Demethylation enzymes, including fat mass and obesity protein (FTO) and ALKB homologous 5 (ALKBH5), belong to the family of Fe²⁺/α-ketoglutarate-dependent dioxygenases and play an “eraser” role in m⁶A methylation (Tan and Gao, 2018). FTO is located in the nucleus and cytoplasm (Jia et al., 2011; Cheung et al., 2013), and ALKBH5 is located in the nucleus (Zheng et al., 2013). FTO regulates the exon splicing of the adipogenic regulator RUNX1T1 by regulating the level of m⁶A

near splice sites, thereby regulating differentiation (Zhao et al., 2014). The demethylation activity of ALKBH5 significantly affects mRNA output, RNA metabolism and mRNA processing factor assembly in nuclear speckles (Zheng et al., 2013).

The role of m⁶A RNA-binding protein is as a “reader.” The most well-known readers include the three members of the YTH domain family (YTHDF), YTHDF1, YTHDF2 and YTHDF3, located in the cytoplasm (Dominissini et al., 2012; Wang et al., 2015; Shi et al., 2017). In addition, two other readers are YTHDC proteins, in which YTHDC1 is located in the nucleus and YTHDC2 is located in the cytoplasm (Morohashi et al., 2011; Xu et al., 2014). YTHDF1 directly interacts with the translation initiation factor eukaryotic initiation factor 3 (eIF3) to increase the translation efficiency of mRNA targets (Wang et al., 2015), while YTHDF2 directly recruits the CCR4-NOT dehydrogenase complex to accelerate the degradation of m⁶A-modified RNA (Wang et al., 2014). Previous studies have shown that the N-terminal domain of YTHDF2 facilitates YTHDF2-mRNA complex localization in P-bodies, which are involved in mRNA degradation (Du et al., 2016). YTHDF3 promotes the functions of YTHDF1 and YTHDF2, cooperates with YTHDF1 to promote translation, and cooperates with YTHDF2 to promote mRNA activity attenuation (Li et al., 2017; Shi et al., 2017). YTHDC1 can drive the binding of methylated mRNA targets to serine and arginine rich splicing factor 3 (SRSF3) and nuclear export factor 1 (NXF1) to form mRNA-protein complexes (mRNPs), which together promote nuclear export of RNA (Roundtree et al., 2017). In addition, YTHDC1 is competitively bound by SRSF3 and SRSF10 to regulate RNA splicing (Xiao et al., 2016). YTHDC2 interacts with small ribosomal subunits and XRN1 to affect RNA translation and degradation, respectively (Hsu et al., 2017; Wojtas et al., 2017). In addition to members of the YTH family, other proteins have been reported to bind to m⁶A. The combination of an m⁶A-modified 5'UTR and eIF3 complex can directly recruit the 43S preinitiation complex to the 5'UTR of the mRNA, thereby stimulating translation initiation (Meyer et al., 2015). Splicing and microRNA maturation is regulated by the binding of hnRNPA2/B1 to m⁶A-modified RNA (Alarcón et al., 2015). Under normal and stress conditions, insulin-like growth factor 2 mRNA-binding proteins (IGF2BPs) promote the stability and storage of their target mRNAs in an m⁶A-dependent manner, thereby affecting gene expression output (Huang et al., 2018).

The Role of m⁶A During Respiratory Virus Infections

m⁶A and Adenovirus

An adenovirus (AdV) is a DNA virus that relies on the processing of host RNA (Price et al., 2020). It uses cellular RNA polymerase II and spliceosome mechanisms to produce early and late genes from two DNA strands to produce mature mRNA (Sommer et al., 1976). Price et al. (2020) found that after adenovirus infection of A549 cells, the expression levels of other interacting enzymes did not change significantly, except for a slight increase in the levels of FTO and ALKBH5, and the host proteins were concentrated at the sites of early viral RNA synthesis. Through indirect immunofluorescence microscopy, it has been observed that some

TABLE 1 | The regulators of m⁶A in respiratory viruses.

Virus	Molecule	Change	Sample source	Biological function	References
AdV	METTL3/METTL14	Up	A549 cell	Promote viral late gene replication and transcription	Price et al., 2020
	METTL3/WTAP	Up	A549 cell	Promote viral late gene splicing	
	YTHDC1	Up	A549 cell	Promote viral late gene splicing	
RSV	METTL3/METTL14	Up	Hela cell	Promote RSV protein expression	Xue et al., 2019
	YTHDF1-3	Up	Hela cell/A549 cell/Vero cell	Promote RSV replication, mRNA transcription,	
	ALKBH5/FTO	Down	Hela cell	Suppress RSV protein expression	
hMPV	METTL3/METTL14	Up	A549 cell	Promote hMPV replication, viral protein and RNA level	Lu et al., 2020
	YTHDF1-3	Up	A549 cell	Enhance viral protein expression, the release of infectious virus, antigenome and mRNAs	
	ALKBH5/FTO	Down	A549 cell	Suppress virus replication	
IAV	METTL3	Up	A549 cell	Promote IAV replication and virion production	Courtney et al., 2017
	YTHDF2	Up	A549 cell	Promote IAV replication and viral spread	

writer proteins, such as METTL3, METTL14, and WTAP, and reader protein YTHDC1 migrate from areas of dispersion in the nucleus to the site of viral RNA synthesis within 18 h of A549 cells infection with AdV5 (Price et al., 2020).

According to the research of Price, the early and late viral transcripts of adenovirus include METTL3-dependent m⁶A modifications, but m⁶A is not necessary for the early phase of adenovirus infection (Price et al., 2020). This group knocked out METTL3 or METTL14 and infected cells 48 h later, and the early gene replication and transcription products of the virus were found to be basically unaffected after infection, but the production of late RNA, late protein, and infectious progeny was significantly reduced in the METTL3- and METTL14-knockout cells (Price et al., 2020). Moreover, the cytoplasmic m⁶A reader and eraser did not affect the adenovirus infection cycle (Price et al., 2020).

The different effects of m⁶A deletion on early and late viral RNAs are mainly due to the important role that m⁶A plays in regulating the splicing efficiency of late viral RNA (Price et al., 2020). Spliced and unspliced viral RNA was detected by qRT-PCR, and the ratio of the two products was used to determine splicing efficiency (Price et al., 2020). After METTL3 or WTAP knockout, the splicing efficiency of the protein encoded by the E1A gene expressed early during infection did not change, while the splicing efficiency of the protein encoded by the fiber gene expressed late in infection decreased significantly (Price et al., 2020). Similar but much less dramatic results were observed upon YTHDC1 knockout (Price et al., 2020). Moreover, METTL3 knockout is widely downregulated during adenovirus late RNA processing. Short-read sequencing results showed that the overall abundance of all viral transcripts in late infection, except that of L152K, was reduced after METTL3 knockout (Price et al., 2020).

m⁶A and Respiratory Syncytial Virus

Human respiratory syncytial virus (RSV) is a non-segmented negative strand (NNS) RNA virus in which both genomic and anti-genomic mRNAs undergo m⁶A modification (Whelan et al., 2016). Xue et al. (2019) found that RSV replication was

positively regulated by m⁶A reader and writer proteins, but eraser proteins had the opposite, negative regulatory, effect. RT-PCR was used to measure RSV genomic RNA and mRNA in cells. Although overexpression of YTHDF1-3 significantly increased the synthesis of RSV genomic RNA and mRNA, overexpression of YTHDF2 resulted in more active replication of RSV than transcription and more synthesized RNA than mRNA (Xue et al., 2019). When METTL3 and METTL14 were overexpressed in HeLa cells, increased synthesis of viral proteins was observed, and the level of m⁶A marks in viral RNA was significantly higher than that in control groups (Xue et al., 2019). In contrast, eraser proteins downregulated RSV replication and reduced viral RNA m⁶A content (Xue et al., 2019).

Abrogation of m⁶A was shown to result in attenuated RSV infection (Xue et al., 2019). Because m⁶A on the G gene and G mRNA is more abundantly enriched than other genes and mRNA, silent mutation of m⁶A sites in the modified region produced mutated RSV (rgRSVs) expressing recombinant green fluorescent protein (Xue et al., 2019). Compared with parental rgRSV, m⁶A-mutant rgRSVs showed inhibited viral protein synthesis, delayed replication kinetics, and reduced titers (Xue et al., 2019). In addition, in an *in vivo* model of lower respiratory tract infection HAE, replication and spread of m⁶A-site-mutated rgRSVs were defective, and the pathogenicity of these viruses was reduced (Xue et al., 2019). However, the immunogenicity of m⁶A-site-mutant rgRSVs was high in cotton mice (Xue et al., 2019).

m⁶A and Human Metapneumovirus

Human metapneumovirus (hMPV) is a NNS RNA virus. Lu et al. (2020) found that the hMPV genome, antigenome and mRNA all contain m⁶A modifications, and the strongest m⁶A peak appeared in the G gene of the genome and antigenome. The effect of m⁶A on hMPV is similar to that on RSV. According to Lu's research, replication and gene expression of hMPV were promoted by m⁶A modification, and weakening of hMPV infection in cell culture was caused by the abolition of m⁶A (Lu et al., 2020). Both transient and stable overexpression of m⁶A-binding proteins had pro-viral effects in hMPV infection, and the

replication, protein and RNA levels of hMPV were significantly upregulated by overexpression of writer proteins (Lu et al., 2020).

m⁶A and Influenza A Virus

The influenza A virus (IAV) genome consists of eight single-stranded negative-sense RNA segments that replicate in the nucleus (Noda et al., 2012). Courtney et al. (2017) found that the expression of IAV transcripts was prompted by *cis*-action of m⁶A-modified residues. When METTL3 was knocked out or the m⁶A site in IAV haemagglutinin (HA) was removed, a reduction in viral gene expression, replication and pathogenicity was observed (Courtney et al., 2017). In addition, the expression of the IAV gene and the production of virus particles were increased by the overexpression of YTHDF2, but overexpression of YTHDF1 and YTHDF3 had little effect on IAV replication (Courtney et al., 2017).

According to Courtney's research, because the deposition of m⁶A not only enhanced splicing but also enhanced the expression of unspliced IAV gene fragments at equivalent levels; therefore, the m⁶A modification is unlikely regulate splicing (Courtney et al., 2017). In addition, the increase in IAV protein expression was closely related to an increase in mRNA expression; the m⁶A modification is also unlikely to enhance translation (Courtney et al., 2017). The mechanism by which the m⁶A modification increases IAV gene expression and viral replication needs to be further explored.

m⁶A and the Antiviral Response

m⁶A modification not only can affect virus replication but can also regulate the immune response of virus infection, but the influence of the m⁶A modification on the host immune response during virus infection has not been fully elucidated.

Viral RNA genomes and intermediates formed during viral replication often have uncapped RNAs with 5'-triphosphate or 5'-diphosphate groups, which can be detected by cytoplasmic RIG-I (McFadden and Horner, 2020). RIG-I binds non-self RNA and induces signal transduction through mitochondrial antiviral signaling proteins (MAVS) to produce type I and type III interferons (IFNs), thereby activating the antiviral response (McFadden and Horner, 2020). A study by Lu et al. showed that the detection of RNA 5' triphosphorylation by RIG-I was diminished by the m⁶A modification of the NNS genome and antigenome, and three NNS RNA virus families (*Pneumoviridae*, *Paramyxoviridae*, and *Rhabdoviridae*) imitated host RNA through m⁶A modification to avoid RIG-I recognition (Lu et al., 2021). In another study by Lu, hMPV was used as a model to demonstrate that m⁶A acts as a molecular marker to distinguish self- from non-self RNA via RIG-I (Lu et al., 2020). When the viral antigenome and genomic RNA were deficient in m⁶A marks, enhanced activation of the RIG-I signaling pathway was observed (Lu et al., 2020). In contrast, the demethylase ALKBH5 was recruited by the RNA helicase DDX46 to demethylate m⁶A-modified antiviral transcripts, which were then retained in the nucleus, preventing their translation and inhibiting the production of interferon (Zheng et al., 2017).

Multiple studies have shown that METTL3-mediated methylation of viral RNA m⁶A promotes immune escape.

Winkler et al. (2019) demonstrated that the loss of METTL3 led to increased induction of interferon-stimulated immune responses. In addition to METTL3 depletion, the deletion of METTL14 increased IFNB1 mRNA production and reduced viral replication (Rubio et al., 2018). In addition, YTHDF2 deletion led to increased type I interferon levels and greater induction of interferon-stimulated genes (Winkler et al., 2019). YTHDF2 was shown to isolate m⁶A-modified circular RNAs and prevent endogenous circular RNAs from activating the RIG-I antiviral pathway (Lu et al., 2020).

Demethylation of m⁶A viral RNA presents potential antiviral therapeutic opportunities. Methylase inhibitor 3-deazaadenosine (DAA) has an antiviral effect (Gordon et al., 2003; Kennedy et al., 2016; Manners et al., 2019), but because DAA reduces the formation of methyl donor S-adenosylmethionine (SAM) (Fustin et al., 2013) and thus inhibits all types of RNA modification, it is not clear which mechanism of methylation inhibition causes the antiviral effect. Therefore, the development of new drugs specifically targeting the mechanism by which m⁶A modification regulates viral infection will be important to the advancement of antiviral therapy.

CONCLUDING REMARKS

In this review, we discussed the latest developments in the role of m⁶A methylation in the life cycle of respiratory viruses. In AdV, RSV, hMPV, and IAV, m⁶A modification mainly supports virus replication (Table 1). However, in other viruses, such as ZIKV and HCV, m⁶A plays a negative regulatory role in virus replication (Manners et al., 2019). This disparity indicates that m⁶A has a complicated role in the process of viral gene expression.

In addition, m⁶A modification is widely involved in the immune response, but the mechanism by which the m⁶A modification affects antiviral immunity is not fully understood. The role of m⁶A modification in the antiviral immune response needs to be further studied in the future, with the results expected to help in developing new antiviral treatment strategies.

AUTHOR CONTRIBUTIONS

QF and HZ wrote the initial draft of the manuscript. ZX reviewed the manuscript. LX proposed the original idea and reviewed the manuscript. This work represents a collaboration among all of the authors. All authors read and approved the final submitted version of the manuscript.

FUNDING

This work was supported by grants from the National Major S&T Research Projects for the Control and Prevention of Major Infections Diseases in China (2017ZX10103004-004 and 2018ZX10305409-001-004) and the CAMS Innovation Fund for Medical Sciences (CIFMS, 2019-12M-5-026).

REFERENCES

- Alarcón, C. R., Goodarzi, H., Lee, H., Liu, X., Tavazoie, S., and Tavazoie, S. F. (2015). HNRNP2B1 is a mediator of m⁶A-dependent nuclear RNA processing events. *Cell* 162, 1299–1308. doi: 10.1016/j.cell.2015.08.011
- Boccalletto, P., Machnicka, M. A., Purta, E., Piątkowski, P., Bagiński, B., Wirecki, T. K., et al. (2018). MODOMICS: a database of RNA modification pathways. 2017 update. *Nucleic Acids Res.* 46, D303–D307.
- Cantara, W. A., Crain, P. F., Rozenski, J., McCloskey, J. A., Harris, K. A., Zhang, X., et al. (2010). The RNA modification database. RNAMDB: 2011 update. *Nucleic Acids Res.* 39, D195–D201.
- Cheung, M., Gulati, P., O'rahilly, S., and Yeo, G. (2013). FTO expression is regulated by availability of essential amino acids. *Int. J. Obes.* 37, 744–747. doi: 10.1038/ijo.2012.77
- Courtney, D. G., Kennedy, E. M., Dumm, R. E., Bogerd, H. P., Tsai, K., Heaton, N. S., et al. (2017). Epitranscriptomic enhancement of influenza A virus gene expression and replication. *Cell Host Microbe* 22, 377–386. doi: 10.1016/j.chom.2017.08.004
- Desrosiers, R., Friderici, K., and Rottman, F. (1974). Identification of methylated nucleosides in messenger RNA from Novikoff hepatoma cells. *Proc. Natl. Acad. Sci. U.S.A.* 71, 3971–3975. doi: 10.1073/pnas.71.10.3971
- Dominissini, D., Moshitch-Moshkovitz, S., Schwartz, S., Salmon-Divon, M., Ungar, L., Osenberg, S., et al. (2012). Topology of the human and mouse m⁶A RNA methylomes revealed by m⁶A-seq. *Nature* 485, 201–206. doi: 10.1038/nature11112
- Du, H., Zhao, Y., He, J., Zhang, Y., Xi, H., Liu, M., et al. (2016). YTHDF2 destabilizes m⁶A-containing RNA through direct recruitment of the CCR4–NOT deadenylase complex. *Nat. Commun.* 7:12626.
- Fustin, J.-M., Doi, M., Yamaguchi, Y., Hida, H., Nishimura, S., Yoshida, M., et al. (2013). RNA-methylation-dependent RNA processing controls the speed of the circadian clock. *Cell* 155, 793–806. doi: 10.1016/j.cell.2013.10.026
- Gordon, R. K., Ginalski, K., Rudnicki, W. R., Rychlewski, L., Pankaskie, M. C., Bujnicki, J. M., et al. (2003). Anti-HIV-1 activity of 3–deaza–adenosine analogs: inhibition of S–adenosylhomocysteine hydrolase and nucleotide congeners. *Eur. J. Biochem.* 270, 3507–3517. doi: 10.1046/j.1432-1033.2003.03726.x
- Hsu, P. J., Zhu, Y., Ma, H., Guo, Y., Shi, X., Liu, Y., et al. (2017). Ythdc2 is an N⁶-methyladenosine binding protein that regulates mammalian spermatogenesis. *Cell Res.* 27, 1115–1127. doi: 10.1038/cr.2017.99
- Huang, H., Weng, H., Sun, W., Qin, X., Shi, H., Wu, H., et al. (2018). Recognition of RNA N⁶-methyladenosine by IGF2BP proteins enhances mRNA stability and translation. *Nat. Cell Biol.* 20, 285–295. doi: 10.1038/s41556-018-0045-z
- Imam, H., Kim, G.-W., and Siddiqui, A. (2020). Epitranscriptomic (N⁶-methyladenosine) modification of viral RNA and virus-host interactions. *Front. Cell. Infect. Microbiol.* 10:584283. doi: 10.3389/fcimb.2020.584283
- Jia, G., Fu, Y., Zhao, X., Dai, Q., Zheng, G., Yang, Y., et al. (2011). N⁶-methyladenosine in nuclear RNA is a major substrate of the obesity-associated FTO. *Nat. Chem. Biol.* 7, 885–887. doi: 10.1038/nchembio.687
- Kennedy, E. M., Bogerd, H. P., Kornepati, A. V., Kang, D., Ghoshal, D., Marshall, J. B., et al. (2016). Posttranscriptional m⁶A editing of HIV-1 mRNAs enhances viral gene expression. *Cell Host Microbe* 19, 675–685. doi: 10.1016/j.chom.2016.04.002
- Li, A., Chen, Y.-S., Ping, X.-L., Yang, X., Xiao, W., Yang, Y., et al. (2017). Cytoplasmic m⁶A reader YTHDF3 promotes mRNA translation. *Cell Res.* 27, 444–447. doi: 10.1038/cr.2017.10
- Liu, J., Yue, Y., Han, D., Wang, X., Fu, Y., Zhang, L., et al. (2014). A METTL3–METTL14 complex mediates mammalian nuclear RNA N⁶-adenosine methylation. *Nat. Chem. Biol.* 10, 93–95. doi: 10.1038/nchembio.1432
- Liu, Z., and Zhang, J. (2018). Human C-to-U coding RNA editing is largely nonadaptive. *Mol. Biol. Evol.* 35, 963–969. doi: 10.1093/molbev/msy011
- Lu, M., Xue, M., Wang, H.-T., Kairis, E. L., Ahmad, S., Wei, J., et al. (2021). Non-segmented negative-sense RNA viruses utilize N⁶-methyladenosine (m⁶A) as a common strategy to evade host innate immunity. *J. Virol.* 95:e01939-20.
- Lu, M., Zhang, Z., Xue, M., Zhao, B. S., Harder, O., Li, A., et al. (2020). N⁶-methyladenosine modification enables viral RNA to escape recognition by RNA sensor RIG-I. *Nat. Microbiol.* 5, 584–598. doi: 10.1038/s41564-019-0653-9
- Manners, O., Baquero-Perez, B., and Whitehouse, A. (2019). m⁶A: widespread regulatory control in virus replication. *Biochim. Biophys. Acta Gene Regul. Mech.* 1862, 370–381. doi: 10.1016/j.bbagr.2018.10.015
- McFadden, M. J., and Horner, S. M. (2020). N⁶-methyladenosine regulates host responses to viral infection. *Trends Biochem. Sci.* 46, 366–367. doi: 10.1016/j.tibs.2020.11.008
- Meyer, K. D., Patil, D. P., Zhou, J., Zinoviev, A., Skabkin, M. A., Elemento, O., et al. (2015). 5' UTR m⁶A promotes cap-independent translation. *Cell* 163, 999–1010. doi: 10.1016/j.cell.2015.10.012
- Morohashi, K., Sahara, H., Watashi, K., Iwabata, K., Sunoki, T., Kuramochi, K., et al. (2011). Cyclosporin A associated helicase-like protein facilitates the association of hepatitis C virus RNA polymerase with its cellular cyclophilin B. *PLoS One* 6:e18285. doi: 10.1371/journal.pone.0018285
- Nair, H., Nokes, D. J., Gessner, B. D., Dherani, M., Madhi, S. A., Singleton, R. J., et al. (2010). Global burden of acute lower respiratory infections due to respiratory syncytial virus in young children: a systematic review and meta-analysis. *Lancet* 375, 1545–1555.
- Noda, T., Sugita, Y., Aoyama, K., Hirase, A., Kawakami, E., Miyazawa, A., et al. (2012). Three-dimensional analysis of ribonucleoprotein complexes in influenza A virus. *Nat. Commun.* 3:639.
- Patil, D. P., Chen, C.-K., Pickering, B. F., Chow, A., Jackson, C., Guttman, M., et al. (2016). m⁶A RNA methylation promotes XIST-mediated transcriptional repression. *Nature* 537, 369–373. doi: 10.1038/nature19342
- Ping, X.-L., Sun, B.-F., Wang, L., Xiao, W., Yang, X., Wang, W.-J., et al. (2014). Mammalian WTAP is a regulatory subunit of the RNA N⁶-methyladenosine methyltransferase. *Cell Res.* 24, 177–189. doi: 10.1038/cr.2014.3
- Price, A. M., Hayer, K. E., McIntyre, A. B., Gokhale, N. S., Abebe, J. S., Della Fera, A. N., et al. (2020). Direct RNA sequencing reveals m⁶A modifications on adenovirus RNA are necessary for efficient splicing. *Nat. Commun.* 11:6016.
- Roundtree, I. A., Luo, G.-Z., Zhang, Z., Wang, X., Zhou, T., Cui, Y., et al. (2017). YTHDC1 mediates nuclear export of N⁶-methyladenosine methylated mRNAs. *eLife* 6:e31311.
- Rubio, R. M., Depledge, D. P., Bianco, C., Thompson, L., and Mohr, I. (2018). RNA m⁶A modification enzymes shape innate responses to DNA by regulating interferon β . *Genes Dev.* 32, 1472–1484. doi: 10.1101/gad.319475.118
- Shi, H., Wang, X., Lu, Z., Zhao, B. S., Ma, H., Hsu, P. J., et al. (2017). YTHDF3 facilitates translation and decay of N⁶-methyladenosine-modified RNA. *Cell Res.* 27, 315–328. doi: 10.1038/cr.2017.15
- Śledź, P., and Jinek, M. (2016). Structural insights into the molecular mechanism of the m⁶A writer complex. *eLife* 5:e18434.
- Sommer, S., Salditt-Georgieff, M., Bachenheimer, S., Darnell, J., Furuichi, Y., Morgan, M., et al. (1976). The methylation of adenovirus-specific nuclear and cytoplasmic RNA. *Nucleic Acids Res.* 3, 749–766. doi: 10.1093/nar/3.3.749
- Tan, B., and Gao, S.-J. (2018). The RNA epitranscriptome of DNA viruses. *J. Virol.* 92:e00696-18.
- Wang, P., Doxtader, K. A., and Nam, Y. (2016). Structural basis for cooperative function of Mettl3 and Mettl14 methyltransferases. *Mol. Cell* 63, 306–317. doi: 10.1016/j.molcel.2016.05.041
- Wang, X., Feng, J., Xue, Y., Guan, Z., Zhang, D., Liu, Z., et al. (2016). Structural basis of N⁶-adenosine methylation by the METTL3–METTL14 complex. *Nature* 534, 575–578. doi: 10.1038/nature18298
- Wang, X., Li, Y., O'Brien, K. L., Madhi, S. A., Widdowson, M.-A., Byass, P., et al. (2020). Global burden of respiratory infections associated with seasonal influenza in children under 5 years in 2018: a systematic review and modelling study. *Lancet Glob. Health* 8, e497–e510.
- Wang, X., Lu, Z., Gomez, A., Hon, G. C., Yue, Y., Han, D., et al. (2014). N⁶-methyladenosine-dependent regulation of messenger RNA stability. *Nature* 505, 117–120. doi: 10.1038/nature12730
- Wang, X., Zhao, B. S., Roundtree, I. A., Lu, Z., Han, D., Ma, H., et al. (2015). N⁶-methyladenosine modulates messenger RNA translation efficiency. *Cell* 161, 1388–1399. doi: 10.1016/j.cell.2015.05.014
- Whelan, J. N., Reddy, K. D., Uversky, V. N., and Teng, M. N. (2016). Functional correlations of respiratory syncytial virus proteins to intrinsic disorder. *Mol. Biosyst.* 12, 1507–1526. doi: 10.1039/c6mb00122j
- Winkler, R., Gillis, E., Lasman, L., Safrá, M., Geula, S., Soyris, C., et al. (2019). m⁶A modification controls the innate immune response to infection by targeting type I interferons. *Nat. Immunol.* 20, 173–182. doi: 10.1038/s41590-018-0275-z

- Wojtas, M. N., Pandey, R. R., Mendel, M., Homolka, D., Sachidanandam, R., and Pillai, R. S. (2017). Regulation of m⁶A transcripts by the 3' 5' RNA helicase YTHDC2 is essential for a successful meiotic program in the mammalian germline. *Mol. Cell* 68, 374–387.
- Wu, R., Jiang, D., Wang, Y., and Wang, X. (2016). N 6-methyladenosine (m⁶A) methylation in mRNA with a dynamic and reversible epigenetic modification. *Mol. Biotechnol.* 58, 450–459. doi: 10.1007/s12033-016-9947-9
- Xiao, W., Adhikari, S., Dahal, U., Chen, Y.-S., Hao, Y.-J., Sun, B.-F., et al. (2016). Nuclear m⁶A reader YTHDC1 regulates mRNA splicing. *Mol. Cell* 61, 507–519.
- Xu, C., Wang, X., Liu, K., Roundtree, I. A., Tempel, W., Li, Y., et al. (2014). Structural basis for selective binding of m⁶A RNA by the YTHDC1 YTH domain. *Nat. Chem. Biol.* 10, 927–929. doi: 10.1038/nchembio.1654
- Xue, M., Zhao, B. S., Zhang, Z., Lu, M., Harder, O., Chen, P., et al. (2019). Viral N 6-methyladenosine upregulates replication and pathogenesis of human respiratory syncytial virus. *Nat. Commun.* 10:4595.
- Zhao, X., Yang, Y., Sun, B.-F., Shi, Y., Yang, X., Xiao, W., et al. (2014). FTO-dependent demethylation of N6-methyladenosine regulates mRNA splicing and is required for adipogenesis. *Cell Res.* 24, 1403–1419. doi: 10.1038/cr.2014.151
- Zheng, G., Dahl, J. A., Niu, Y., Fedorcsak, P., Huang, C.-M., Li, C. J., et al. (2013). ALKBH5 is a mammalian RNA demethylase that impacts RNA metabolism and mouse fertility. *Mol. Cell* 49, 18–29. doi: 10.1016/j.molcel.2012.10.015
- Zheng, Q., Hou, J., Zhou, Y., Li, Z., and Cao, X. (2017). The RNA helicase DDX46 inhibits innate immunity by entrapping m⁶A-demethylated antiviral transcripts in the nucleus. *Nat. Immunol.* 18:1094. doi: 10.1038/ni.3830
- 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 Feng, Zhao, Xu 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.



RESIC: A Tool for Comprehensive Adenosine to Inosine RNA Editing Site Identification and Classification

Dean Light[†], Roni Haas[†], Mahmoud Yazbak, Tal Elfand, Tal Blau and Ayelet T. Lamm*

Faculty of Biology, Technion – Israel Institute of Technology, Haifa, Israel

OPEN ACCESS

Edited by:

Xiao Han,
Biotechnology Research Institute,
Chinese Academy of Agricultural
Sciences, China

Reviewed by:

Yong-Fang Li,
Henan Normal University, China
Yongfeng Jin,
Zhejiang University, China

*Correspondence:

Ayelet T. Lamm
ayeletla@technion.ac.il

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

Specialty section:

This article was submitted to
RNA,
a section of the journal
Frontiers in Genetics

Received: 28 March 2021

Accepted: 07 June 2021

Published: 23 July 2021

Citation:

Light D, Haas R, Yazbak M,
Elfand T, Blau T and Lamm AT (2021)
RESIC: A Tool for Comprehensive
Adenosine to Inosine RNA Editing Site
Identification and Classification.
Front. Genet. 12:686851.
doi: 10.3389/fgene.2021.686851

Adenosine to inosine (A-to-I) RNA editing, the most prevalent type of RNA editing in metazoans, is carried out by adenosine deaminases (ADARs) in double-stranded RNA regions. Several computational approaches have been recently developed to identify A-to-I RNA editing sites from sequencing data, each addressing a particular issue. Here, we present RNA Editing Sites Identification and Classification (RESIC), an efficient pipeline that combines several approaches for the detection and classification of RNA editing sites. The pipeline can be used for all organisms and can use any number of RNA-sequencing datasets as input. RESIC provides (1) the detection of editing sites in both repetitive and non-repetitive genomic regions; (2) the identification of hyper-edited regions; and (3) optional exclusion of polymorphism sites to increase reliability, based on DNA, and ADAR-mutant RNA sequencing datasets, or SNP databases. We demonstrate the utility of RESIC by applying it to human, successfully overlapping and extending the list of known putative editing sites. We further tested changes in the patterns of A-to-I RNA editing, and RNA abundance of ADAR enzymes, following SARS-CoV-2 infection in human cell lines. Our results suggest that upon SARS-CoV-2 infection, compared to mock, the number of hyper editing sites is increased, and in agreement, the activity of ADAR1, which catalyzes hyper-editing, is enhanced. These results imply the involvement of A-to-I RNA editing in conceiving the unpredicted phenotype of COVID-19 disease. RESIC code is open-source and is easily extendable.

Keywords: SARS-CoV-2, ADAR, epitranscriptome, interferon, hyper-editing

INTRODUCTION

The conversion of adenosine to inosine (A-to-I) in double-stranded RNA regions, by adenosine deaminases (ADARs) enzymes, is the most common form of RNA editing in metazoans (Bazak et al., 2014). This type of RNA editing is crucial for normal development of an organism and has a major role in the innate immune response (Mannion et al., 2014; Ganem and Lamm, 2017; Eisenberg and Levanon, 2018). It was shown that changes in editing events are correlated with several types of diseases; including cancer (Maas et al., 2006; Galeano et al., 2012; Gallo and Locatelli, 2012; Kung et al., 2018). Editing sites may serve as biomarkers for cancer and ADAR enzymes are considered as promising gene therapy agents to fight cancer (Ganem et al., 2017). In addition, ADARs are known to be involved in regulation of innate immune response by blocking the interferon (IFN) response upon viral infection (Quin et al., 2021). For these reasons, A-to-I RNA editing is an extensively studied research field in many organisms, and identification of editing sites is of major interest.

In recent years, many efforts have been invested in developing computational approaches to detect A-to-I RNA editing sites from sequencing data (Pinto and Levanon, 2019). Since inosine is very similar in structure to guanosine, inosine is interpreted as guanosine by polymerases during sequencing. This enables the detection of editing sites by comparing between DNA and RNA sequences, to track adenosine to guanosine (A-to-G) mismatches. However, the detection should be carefully performed to avoid false reports due to sequencing and alignment mistakes, alterations in sequence originated from polymorphism, somatic mutations, or other changes which are not the result of editing events (Pinto and Levanon, 2019). The problem is exacerbated by the fact that editing in humans frequently occurs in repetitive regions (Athanasiadis et al., 2004; Blow et al., 2004; Kim et al., 2004; Levanon et al., 2004, 2005; Barak et al., 2009; Kleinberger and Eisenberg, 2010; Osenberg et al., 2010; Paz-Yaacov et al., 2010; Wu J. et al., 2011), which tend to cause alignment errors (Treangen and Salzberg, 2011).

Several tools developed to detect A-to-I RNA editing sites are based on comparison between RNA-seq reads and DNA reference genome. Among these tools are REDIttools, which suggest simple comparison using samtools (Picardi and Pesole, 2013), and GIREMI that focused on distinguishing between SNPs and editing, relying on existing SNP databases and a given RNA-seq data (Zhang and Xiao, 2015). Some tools support a direct comparison between RNA-seq reads and DNA reads from the same source, allowing editing site identification without the need for previous knowledge (Picardi and Pesole, 2013; Lee et al., 2015; Wang et al., 2016; Piechotta et al., 2017). A major advantage in comparing between DNA and RNA sequences of the same biological sample is the ability to increase accuracy by excluding changes deriving from unpublished SNPs (Pinto and Levanon, 2019). Another way to increase the results accuracy is parallel comparison between several RNA-seq datasets of several individuals, while taking into consideration that true editing sites would appear in all or most samples (Ramaswami et al., 2013; Wang et al., 2016; Goldstein et al., 2017).

Hyper-editing by ADAR enzymes, which is defined as multiple A-to-I editing sites in a proximity, is a widespread phenomenon. Since most tools designed to identify editing sites are based on the detection of a limited number of mismatches in read alignments (to reduce alignment errors and running time), hyper-editing events, which result in multiple mismatches in a single read (SR), are usually unexposed. Therefore, several recent methods were specially oriented to track hyper-editing sites. Wu D. et al. (2011) and Porath et al. (2014) developed methods that are based on the conversion of unmapped read-sequences to a three-base code genome and thus enable identification of hyper-editing sites. Namely, all As are transformed to Gs in the reference genome and in the RNA-seq reads that previously failed to align, and realignment is then carried out. Following reversion to original sequences, hyper-editing sites, which are rich with A-to-G mismatches, can be located. In both studies, conversion to a three-base code was repeated for all possible nucleotide pairs. It was shown that A-to-G editing was enriched over the other editing types.

Despite the efforts to develop computational tools for A-to-I RNA editing site detection from sequencing data, to date there is

not a single platform enabling robust detection of editing sites of different classes and their classification. Here, we present RNA Editing Sites Identification and Classification (RESIC), which enables detection and classification of A-to-I RNA editing sites of different types in a single tool. We expanded the pipeline we previously applied to identify editing sites in repetitive and non-repetitive regions (Goldstein et al., 2017) and adopted the method by Wu D. et al. (2011) and Porath et al. (2014) to find hyper-editing sites. The tool includes an alignment-graph of distinctive architecture and several filtration steps to reduce false identifications. RESIC also enables distinguishing between polymorphism and editing events to increase reliability, by using DNA sequences, ADAR mutant RNA-sequencing datasets, or a SNP database. We demonstrate the utility of RESIC by applying it to mapping A-to-I RNA editing sites in 16 human tissues, from the Illumina Human Body Map project, analyzed for a similar purpose by others (Zhu et al., 2013; Bazak et al., 2014; Porath et al., 2014). Our analysis reproduced known putative editing sites, detected by others and included in the RADAR database (Ramaswami and Li, 2014), and extended the list of known sites.

Since aberrant IFN and cytokine responses were observed in COVID-19 patients (Moore and June, 2020) and ADAR1 was shown to activate the IFN reaction (Baños-Lara et al., 2013), we further interrogate the activity of A-to-I RNA editing upon SARS-CoV-2 infection. We show that in SARS-CoV-2 infected samples, compared to mock, ADAR1 is the only A-to-I RNA editing enzyme that is differentially expressed, and the numbers of A-to-I hyper editing sites are larger.

METHODS AND DEFINITIONS

RNA Editing Sites Identification and Classification Algorithmic Definitions

RNA Editing Sites Identification and Classification enables the user to supply DNA or RNA datasets that should exhibit the desired editing phenomena and DNA or RNA sequencing datasets that should not exhibit the desired editing phenomena. The latter group is used to exclude changes deriving from SNPs. Since nucleotide changes in the former sequencing datasets correspond to positive evidence of that sites undergoing editing and the latter datasets correspond to negative evidence, we term these sets of datasets as positive and negative datasets. RESIC is completely reference agnostic. The users provide whichever reference file they wish to use for the alignment as well.

Ambiguous Read Filtering

For ambiguous read filtering, we adopted the method of Porath et al. (2014). Briefly, we filtered out the reads that meet the next criteria: one or more nucleotides represent over 60% or under 10% of the read sequence, more than 10% Ns (when a base call could not be done), average Phred quality score < 25, and more than 20 repeats of a single nucleotide in a row.

Alignment Scheme

We define an alignment scheme to be a 4-tuple $S = (A, p, f_1, f_2)$ where A is an alignment algorithm, p is a list of alignment parameters for A , f_1 is a preprocessing function of the raw

datasets and f_2 is a postprocessing function for aligned and misaligned reads. These seemingly verbose definitions enable RESIC to decouple the choice of alignment algorithm from the rest of the modules in RESIC.

Let S be an alignment scheme, L be a sequencing dataset and R be a genome reference, we define $P_{S,L,R}$ and $N_{S,L,R}$ to be the aligned and misaligned read fractions resulting from running S on L and R . We define $S(L, R) = (P_{S,L,R}, N_{S,L,R})$.

We say that a scheme is normal if f_1 and f_2 are identity functions in said scheme. Pseudo code for calculating $S(L, R)$:

```
def S(L, R):
    L', R' = f_1(L, R) # preprocessing the
                        sequencing
    datasets
    P', N' = A(L', R', p) # Aligning the sequencing
                        datasets
    P, N = f_2(P', N') # post processing results
    return P, N
```

Graph Aligner

Given a directed acyclic graph $G = (V, E)$ where nodes in V are alignment schemes, L a sequencing dataset and R a reference we define new alignment schemes $G_v(L, R)$ for each $v \in V$ to be defined as follows:

$$G_v(L, R) = \begin{cases} v(L, R) & \text{In}(v) = \emptyset \\ v(L'', R) & L'' = \cap_{u \in V(u, v) \in E} N_{G_u, L, R} \end{cases}$$

We define $G(L, R)$ to be a set of aligned sequencing datasets $\{P_{G_v, L, R} \mid v \in V\}$. Simply put, we perform the alignment scheme of node v on all read fragments that were misaligned in any of v 's ancestors.

3nt Genome Alignment Scheme

Let X and Y be two distinct nucleotides. To be able mapping hyper editing sites, we apply the 3nt alignment scheme by which each appearance of either X or Y is transformed into X in both the sequencing datasets (reads) and the reference genomes. That was similarly done by others (Wu D. et al., 2011; Porath et al., 2014). However, we present an advanced 3nt technique to map hyper antisense reads as was not described elsewhere, to the best of our knowledge. First, for each X and Y nucleotides pairs, we first apply the scheme to the reads at the given node (see section "Graph Aligner") and to the reference sense strand. Next, in order to identify hyper editing sites on the antisense strand, for each X and Y nucleotides pairs, we create the complement reference genome, based on the original reference, and reverse the reads that were unmapped in the previous step, to achieve the 3'-5' direction, same as the created reference. Then we reapply the 3nt alignment scheme while considering the manipulation of the reference genome and reads when recording the mapped reads as aligned to the antisense.

In each step, after mapping the reads, aligned and unaligned reads are reverted to their original sequence *via* custom python scripts. **Supplementary Figure 1** illustrates in details the 3nt genome alignment scheme. To conduct the 3nt scheme we use *awk* (Aho et al., 1996) and *sed* (Free Software Foundation, 2019).

Site Filtering

After performing the graph alignment for each of the given sequencing datasets, *samtools* (Li et al., 2009) is used to convert the files into pileup format. Then, several filtering steps are performed as detailed below. All parameters (l , k_1 , k_2 , u , r , and c) are user defined.

First, sites with no nucleotide changes and sites covered by less than l reads are discarded. We discarded sites from the positive datasets if those same sites appeared in any negative dataset with a nucleotide change.

Editing Percent Filtering

For each positive sequencing dataset, we filter out any site: (1) whose most abundant nucleotide change constitutes less than k_1 percent or more than k_2 percent of the reads mapped to that site, (2) whose other nucleotide changes constitute over u percent of the reads mapped to that site, and (3) whose most abundant nucleotide change is in at least r reads. We term: k_1 , the minimal editing percent threshold, k_2 , the maximal editing percent threshold, r , the editing read threshold, and u , the editing noise threshold.

Unique Site Filtering

We filter all sites that were defined as editing sites at the previous step, under more than one editing category (e.g., non-repetitive and hyper non-repetitive A to C), if they represented more than one type of nucleotide change (e.g., once A to G and the other time A to C).

Hyper Editing Filtering

Deriving from our method, it may be possible that under the hyper editing categories, a non-hyper editing site would be recorded. Namely, for each pair of nucleotides X and Y that we perform the 3nt genome scheme, other nucleotide mismatches than hyper X to Y or Y to X may be noted, enabled by the new conditions created by the 3nt scheme. Therefore, we filter the hyper editing files to include only X to Y or Y to X changes (see an illustration in **Supplementary Figure 1**).

CONSENSUS

We filter out any sites that are not present in over c percent of positive datasets.

A-to-I Editing Identification Pipe

We implemented a hyper editing alignment scheme and built an alignment graph that could target any editing type. Specifically, in the analysis described here we only applied RESIC to A-to-I RNA editing.

A-to-I RNA Editing Alignment Graph

In our screen for A-to-I editing sites, we define two classes of alignment schemes, non-repetitive alignment for reads that map uniquely to the genome and repetitive for repetitive regions or regions that cannot be differentiated by our reads *via* alignment.

Our graph alignment, summarized in **Supplementary Figure 2**, is as follows: we align sequencing datasets using the non-repetitive normal scheme followed by the repetitive normal scheme. Then we branch out and for each pair of distinct nucleotides X and Y, we perform the non-repetitive 3nt genome scheme, and the repetitive 3nt genome scheme.

RNA Editing Profiling of Illumina BodyMap2 Transcriptome

RNA-seq datasets from 16 human tissues (Illumina Human Body Map 2.0 Project; GEO accession number GSE30611) that were sequenced at 75 SR, were downloaded from SRA. FastQC was used to control the read quality and trimming was performed accordingly. Reads were further collapsed and then taken for a RESIC run. For the underlying sequencing algorithm, we used Bowtie (Langmead et al., 2009) alignment tool. For the non-repetitive and repetitive alignments, we configured bowtie to align to fragments if they map to under 2, or 20 different genomic locations, respectively, with at most 3 single base mismatches and to consider matches for a read r as the set of alignment results for r with the smallest alignment score ($-m\ 2\ -n\ 3\ -best\ -strata\ -l\ 50\ -chunkmbs\ 200$, and $-m\ 20\ -n\ 3\ -best\ -strata\ -l\ 50\ -chunkmbs\ 200$, respectively). Similar alignment was used in Goldstein et al. (2017). For the site filtration steps, we choose $l = 2$ to be the coverage per site threshold, $k_1 = 30$ and $k_2 = 99$ for the editing minimal and maximal percent threshold, respectively, $u = 3$ for the noise thresholds, $r = 2$ for the editing read threshold, and $c = 0$ for consensus threshold. The site lists obtained for each tissue were filtered to have only A-to-I sites, namely A-to-G, or T-to-C mismatches in both strands.

The list of obtained editing sites was compared to the entire list from RADAR database (Ramaswami and Li, 2014). We considered as shared editing sites, sites that are included in the RADAR list or sites that have gene annotations similar to the ones appeared in the RADAR list.

Experimental Validation of Novel Editing Sites

We investigated three novel candidate editing sites that were found in brain tissue using RESIC. For validation using Sanger sequencing, RNA from three sections of brain glioblastoma sample (a kind gift from Dedi Meiri and Yaniv Lerenthal) was used. RNA was treated with turbo DNase I (Invitrogen) and then a reverse transcriptase reaction was performed with Maxima First Strand cDNA Synthesis Kit (Thermo Scientific), using primers that surrounded the candidate editing sites (listed in the Supplementary material). The amplification products were directly sequenced by Sanger sequencing.

Profiling of SARS-CoV-2 Infected Calu-3 Cells

Raw RNA-seq data of Calu-3 human Lung adenocarcinoma cells infected with SARS-CoV-2 virus or mock, were downloaded from

SRA, BioProject PRJNA615032. FastQC was used to control the read quality and trimming was performed accordingly. Reads were collapsed and first aligned to the SARS-CoV-2 reference genome version NC_045512.2 using bowtie. Alignment to the SARS-CoV-2 genome was made to exclude reads that are originated from the virus for further analysis, and to validate that in contrast to the mock samples, the SARS-CoV-2 samples are infected with the virus. Indeed, few thousands of reads were mapped to the SARS-CoV-2 genome, only for the SARS-CoV-2 infected samples. We applied RESIC separately on the unaligned reads of the mock and SARS-CoV-2 infected samples (three biological replicates each) to identify changes in RNA editing events upon coronavirus infection. For the underline sequencing algorithm, we used Bowtie (Langmead et al., 2009) alignment tool. For the non-repetitive and repetitive alignments, we configured bowtie to align to fragments if they map to under 2, or 100 different genomic locations, respectively, with at most three single base mismatches ($-m\ 2\ -n\ 3\ -best\ -strata\ -l\ 50\ -chunkmbs\ 200$, and $-m\ 100\ -n\ 3\ -best\ -strata\ -l\ 50\ -chunkmbs\ 200$, respectively). For the site filtration steps, we choose $l = 2$ to be the coverage per site threshold, $k_1 = 30$ and $k_2 = 99$ for the editing minimal and maximal percent threshold, respectively, $u = 3$ for the noise thresholds, and $r = 2$ for the editing read threshold. The consensus module was run with $c = 0.5$. We then filtered the site lists obtained to have only A-to-I sites. Since the RNA library preparation strategy was stranded (the sequenced strand must be from the actual expressed strand), we filtered the files obtained to include only actual A-to-G sites and not T-to-C. To test the difference in the numbers of editing sites, under the non-repetitive and hyper-non-repetitive classes, between all SARS-CoV-2 and mock samples, we performed Two-tailed T -test with equal variances (to determine equal variances Levene Test was performed). For this test, we included the normalized (according to the total read coverage per each class stage) numbers of editing sites of six repeats for each sample type, originated from 3 biological replicas that were evaluated twice for each strand separately.

To perform differential expression analysis (DEA), we mapped the same unaligned reads that were used for RESIC analysis before, to the human transcriptome version GRCh37 (hg19) using bowtie. Gene expression levels were evaluated by read counts. We then compared our created gene counts to the already processed counts downloaded from GEO: GSE147507. Although read count values were not identical, as expected due to the use of different alignment tools, the trend was the same.

We performed DEA, using DESeq2 (Love et al., 2014), with `lfcShrink` function and `apeglm` shrinkage estimator type.

RNA Editing Sites Identification and Classification Code Availability

RNA Editing Sites Identification and Classification is an open source available at our GitHub repository¹.

¹<https://github.com/Lammlab/Resic>

RESULTS AND DISCUSSION

RNA Editing Sites Identification and Classification—A Comprehensive Tool for Identification RNA Editing Sites

To have a complete identification of RNA editing sites, which include sites in non-repetitive regions, sites in repetitive regions, and sites in hyper-editing regions, we generated a novel tool termed “RESIC.” RESIC composed of an enhanced alignment graph model to identify and classify editing sites by their type, multiple-step filtering process to increase result reliability in a flexible manner (**Figures 1A,B**) and plots for data visualization (see an example in **Figure 2**).

To use RESIC, the user should supply positive datasets, i.e., RNA-/DNA sequencing datasets that should exhibit the desired editing phenomena, and a reference genome. The user may also supply negative datasets, DNA or RNA sequencing datasets that should not exhibit the desired editing phenomena, or a SNP database, to contradict editing-site existence.

All datasets are first processed according to the graph alignment (**Figures 1A,B**). The graph alignment was designed to track editing sites of different types by aligning a given set of reads to the reference genome in a specific parameter configuration setup that represents each editing class. Namely, we demanded unique or multiple alignment to detect non-repetitive and repetitive sites, respectively. Next for sequences that did not align, we converted read-sequences to a three-base code to detect hyper editing sites for all possible nucleotide pairs, as described by Wu D. et al. (2011) and Porath et al. (2014). Our alignment-graph distinctive architecture enables the fluent utilization of an unmapped read-set that was discarded in one alignment level for defining editing sites of a different class in the next level (**Figure 1B**). This enables the identification of multiple editing-site classes in a single platform. While RESIC was built to provide a way to consolidate the many ongoing efforts at A-to-I editing site identification, our graph aligner model is general and robust enough to stand on its own and contribute to general identification of nucleotide changes. RESIC was based on algorithms and scripts whose ability to correctly identify editing sites was tested experimentally in Wu D. et al. (2011) and Goldstein et al. (2017). Further validation is also presented below.

Following alignment, the candidate editing sites that were identified are going through strict multi-stage filtering process (**Figure 1A**). The filtering process’ aim is to increase the results’ reliability considering different types of possible errors. In one type, sites in which there is more than one mismatch type or sites showing low change ratio are suspicious as technical errors likely to be formed during sequencing or alignment and discarded due to low reliability. The user may easily modify the limiting thresholds controlling these filtering steps (i.e., minimal coverage per site, minimal change ratio, and maximal noise ratio). For example, filtering out ambiguous reads to reduce alignment errors as well as filtering low covered and noisy (with more than one mismatch) sites. Also, reducing the maximal editing percent threshold to less than 100% can reduce SNPs. In another type, incorrect recognition of SNPs as editing sites can be prevented by excluding sites that show the same nucleotide alterations in

both the DNA and the RNA sequences. The user may choose (but it is not mandatory) to supplying DNA sequencing data of the same individual used to detect editing sites, for enabling the described DNA based exclusion. Another way in which SNPs can be distinguished from editing sites is parallel comparison between several samples of different individuals, by testing the consensus level of editing sites. The rationale behind parallel comparison of various individuals is that true editing sites would appear in all or most samples (Ramaswami et al., 2013; Wang et al., 2016; Goldstein et al., 2017). In addition, biological replicas can eliminate changes that occurred because of sequencing errors. Testing for consensus in editing sites among several samples is a less favorable option to eliminate SNPs, in case DNA sequencing dataset is supplied. The user may choose to neutralize the consensus filtering step or modify the consensus threshold.

Our motivation was to build one tool that envelopes several algorithms and enables prediction of all classes of editing sites with as much flexibility as possible. Most of the current bioinformatics tools as described in the introduction focus on one class of editing sites identification (for example, only on hyper editing sites as in Wu D. et al. (2011), with very limited flexibility on the input data, and on the stringency of the detection.

RNA Editing Sites Identification and Classification Enhanced the Number of Identified A-to-I Editing Sites in Human Tissues

In order to test the utility of the tool, we used RNA-seq datasets from seven human tissues: adipose, adrenal, brain, breast, colon, kidney, and heart (Illumina Human Body Map 2.0 Project; GEO accession number GSE30611) that were sequenced at 75 SR. These datasets were previously screened for editing sites by others (Zhu et al., 2013; Bazak et al., 2014; Porath et al., 2014).

We used the latest GRCh37 SNP database (NCBI) to eliminate changes that are not originated from A-to-I RNA editing, but from genomic polymorphism. All datasets were processed according to the graph alignment and went through all filtration steps (for parameters setup, see “Methods and Definitions”). Since each of the 16 samples is originated from a different tissue, and editing sites may be tissue specific (Picardi et al., 2015), we defined $c = 0$ for consensus threshold. To test the power of RESIC to specifically identify A-to-I editing sites, we compared the output of RESIC (**Supplementary Table 1**) to the collection of A-to-I RNA editing sites, taken from RADAR (Ramaswami and Li, 2014). It is indicated by our comparison (**Table 1**) that over 75% of the non-repetitive sites RESIC identified, and over 65% of non-repetitive hyper sites are also included in the RADAR collection. This large overlap is expected, since RADAR is based, among others, on the same samples analyzed by us, and at the same time strengthening the reliability of RESIC. Since hyper editing sites are less frequently found by traditional tools, dictated by the more common alignment parameter setup (Pinto and Levanon, 2019) and tools that aimed for tracking hyper-editing sites (Wu D. et al., 2011; Porath et al., 2014) are less abundant, it is not surprising that a smaller overlap was obtained for non-repetitive hyper sites, compared to non-repetitive.

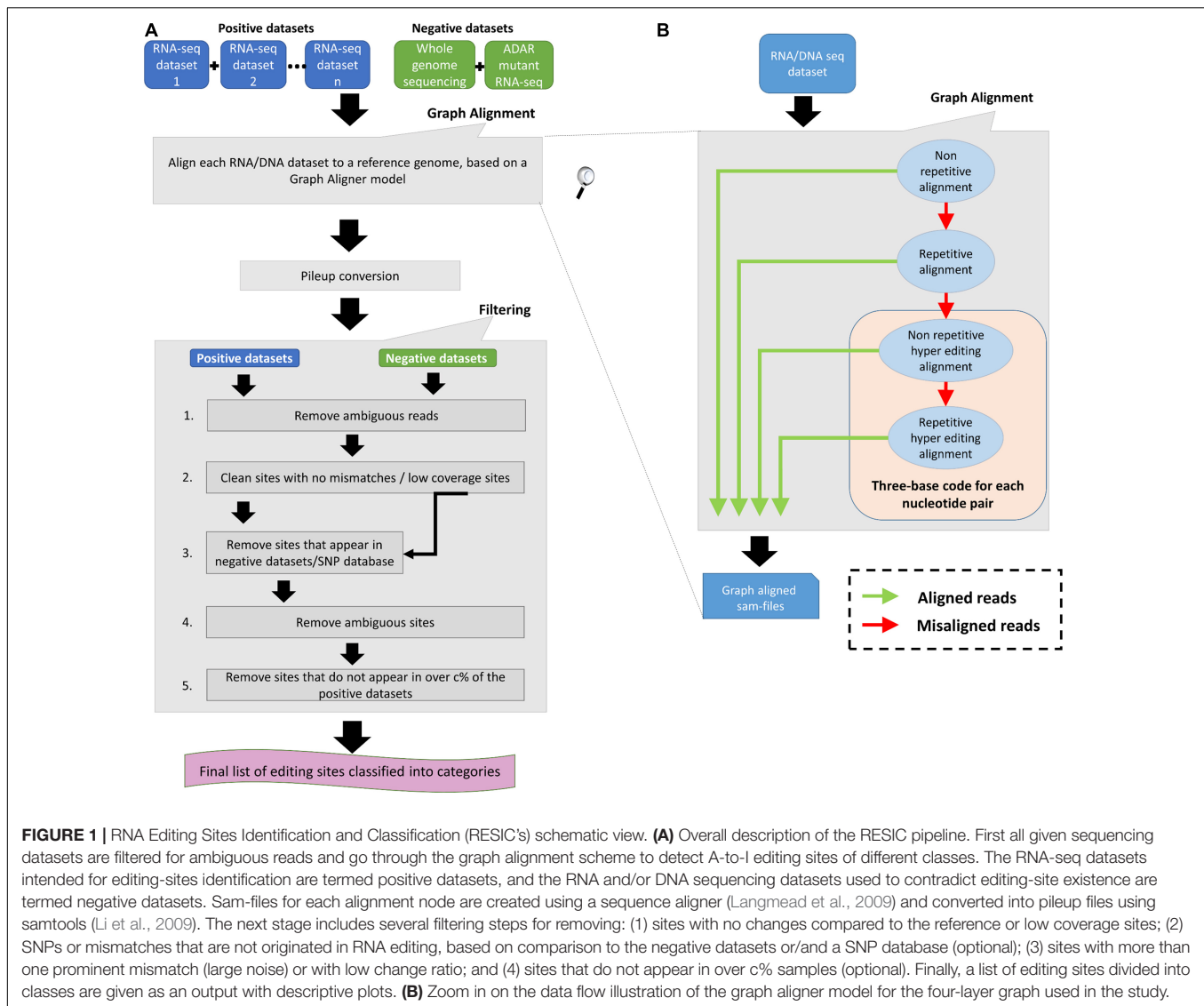


FIGURE 1 | RNA Editing Sites Identification and Classification (RESIC)'s schematic view. **(A)** Overall description of the RESIC pipeline. First all given sequencing datasets are filtered for ambiguous reads and go through the graph alignment scheme to detect A-to-I editing sites of different classes. The RNA-seq datasets intended for editing-sites identification are termed positive datasets, and the RNA and/or DNA sequencing datasets used to contradict editing-site existence are termed negative datasets. Sam-files for each alignment node are created using a sequence aligner (Langmead et al., 2009) and converted into pileup files using samtools (Li et al., 2009). The next stage includes several filtering steps for removing: (1) sites with no changes compared to the reference or low coverage sites; (2) SNPs or mismatches that are not originated in RNA editing, based on comparison to the negative datasets or/and a SNP database (optional); (3) sites with more than one prominent mismatch (large noise) or with low change ratio; and (4) sites that do not appear in over c% samples (optional). Finally, a list of editing sites divided into classes are given as an output with descriptive plots. **(B)** Zoom in on the data flow illustration of the graph aligner model for the four-layer graph used in the study.

Among all classes defined *via* RESIC, the “non-repetitive” class yielded the largest overlap (75.4%, 65.4%, 27.0%, and 30.7%, for non-repetitive, non-repetitive hyper, repetitive, and repetitive hyper, respectively; **Table 1**). For repetitive site classes, smaller overlap was obtained.

A substantial portion of sites detected by RESIC were not identified by others. The explanation for the new identified sites in this study may be the result of the usage of different tools for alignment [i.e., Bowtie in our case and BWA, or a combination of Bowtie, SOAP, and BWA in Zhu et al. (2013) and Porath et al. (2014)], as well as various threshold parameters and filtering criteria being set to consider sites as “editing sites” across tools.

The distribution of the editing events divided into classes can be shown in **Figure 2**, presenting for example the RESIC results for an adrenal tissue sample (the plots obtained from the rest of the samples can be found in the **Supplementary Figures 3–8**. A-to-G and T-to-C are both considered as editing changes because the data is not stranded. Over all classes being

identified according to the graph alignment, A-to-G and T-to-C types were highly enriched, as expected (**Figure 2**). While non A-to-G mismatches are expected to be uncommon (Li et al., 2011; Kleinman and Majewski, 2012), RESIC still identified a certain amount of sites of that type, although in a much lower extent. This may be the result of rare SNPs that are uncovered by the SNP database being used.

Overall, the unique characterization of RESIC enables the detection of different classes of editing events, in one tool. The specificity of RESIC can be seamlessly controlled by modifying the running parameters, and by supplying datasets to exclude SNPs.

Finally, we validated experimentally using sangar sequencing a few of the most likely novel candidate editing sites that were found using RESIC in the brain tissue. These sites, which were not included in the RADAR collection, are in: chromosome 2, position 130737822, *RAB6C* gene; chromosome 14, position 28733993; chromosome 15, position 39889079, *THBS1* gene. For validation, we used RNA from three sections of brain

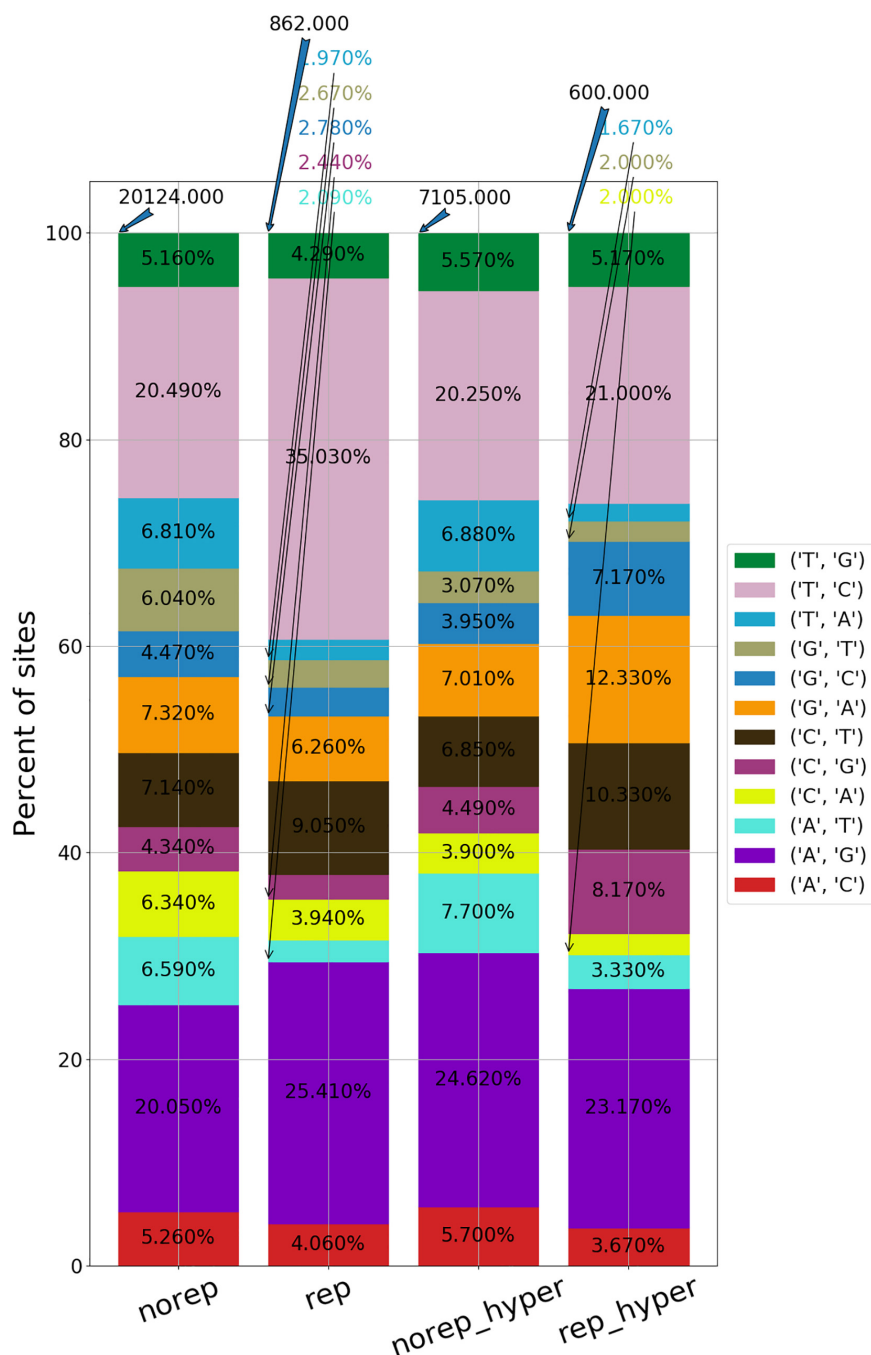


FIGURE 2 | An example for RESIC editing percent distribution plot, obtained for an adrenal tissue sample. Blue arrow at the top of each bar shows the total number of sites being identified for the class. The percentages on the bars present the total number of editing type out of all identified site in the class. The x-axis labels represent the RNA editing classes. Norep: non-repetitive, rep: repetitive, norep_hyper: non-repetitive hyper, rep_hyper: repetitive hyper.

TABLE 1 | Overlap levels of the detected A-to-I editing sites with RADAR database.

RESIC editing class	Non-repetitive	Non-repetitive hyper	Repetitive	Repetitive hyper
Number of A-to-I editing sites by RESIC	23,194	12,496	1,225	895
Shared with RADAR database full list (%) ¹	75.4	65.4	27.0	30.7
Novel sites, non-shared with RADAR database full list (%)	24.6	34.6	73.0	69.3

¹We considered sites as "shared with RADAR" if they were either included in the collection taken from RADAR, or their annotated genes are in RADAR's collection.

glioblastoma tissue. For two of the novel editing sites, in chromosome 2 and chromosome 14, a high editing ratio was clearly observed in sanger sequencing, for all tissue sections (**Supplementary Figure 9**). For the third site on chromosome 15, no editing was observed. However, since not the same cells were used for the bioinformatics analysis and the experimental validation, this test does not disapprove of the existence of a real editing site at that location. In addition, this sample was taken from cancer cells, which were already shown to have differences in editing levels (Ganem et al., 2017). Considering the biological differences between the samples used for the bioinformatics analysis and the experimental validation, the overall support of the validation results in RESIC tool reliability is strong.

SARS CoV-2 Infection Results in an Extensive A-to-I Hyper RNA Editing and Upregulation of ADAR1 Enzyme

Systemic inflammatory responses to viral infection are triggered by IFN-mediated innate immune response (Schneider et al., 2014). Properly orchestrated, this type of immune response leads to inhibition of virus replication, promotion of virus clearance and induction of tissue repair. However, in some people infected with COVID-19, unpredictably, the innate immune response is exaggerated leading to Acute Respiratory Distress Syndrome (ARDS) (Moore and June, 2020). The innate immune response is regulated by ADAR enzymes, which modulate the IFN response to viral infection and reduce the innate immune response. ADAR1 was also shown to prevent Melanoma Differentiation-Associated Protein 5 (MDA5) from sensing dsRNA (Liddicoat et al., 2015) and activating both type I and type III IFNs (Baños-Lara et al., 2013). Therefore, we wished to interrogate the activity of ADAR enzymes following SARS-CoV-2 infection. For this purpose, we analyzed the data of Blanco-Melo et al. (2020), of human Calu3 cells infected with SARS-CoV-2 virus or mock, to examine the differences in A-to-I RNA editing patterns.

To identify and classify RNA editing sites, we applied RESIC on Calu3 cell lines that were infected with SARS-CoV-2 virus or mock (see “Methods and Definitions”). Following RESIC run, we assessed the numbers of the most prevalent classes of A-to-I RNA editing types: non-repetitive, and hyper non-repetitive. We compared between SARS-CoV-2 and mock editing site numbers for each class following normalization, relying on the total read processed in each node. A-to-I non-repetitive hyper editing was significantly more frequent in SARS-CoV-2 infected cells compared to mock (P -value = 0.0371). Overall, the number of hyper editing sites upon SARS-CoV-2 infection was 36.45% greater than in mock (**Figure 3**). For the non-repetitive class, the site numbers were relatively low for both sample types. The number of non-repetitive editing sites was larger as well in SARS-CoV-2, but the difference was not significant (P -value = 0.0765). Overall, the number of non-repetitive sites upon SARS-CoV-2 infection was 28.62% greater than in mock (**Figure 3**).

We next wished to further validate the editing sites that we found in human adenocarcinomic lung epithelial (Calu3)

cells that were infected with SARS-CoV-2, using a different human cell line. For that aim, we used RNA-seq data from human adenocarcinomic alveolar basal epithelial (A549) that were infected with SARS-CoV-2 virus (Blanco-Melo et al., 2020) and searched for RESIC predicted sites in these cells. Only two biological replicates from A549 samples had enough sequencing coverage to search for editing sites. Using these two samples, we were able to detect (in A549 cell line) 52% of the non-repetitive sites, and 37% of the hyper-non-repetitive sites that we found in SARS-CoV-2 Calu3 cells. This large overlap of editing sites, which was obtained despite a low amount of RNA-seq reads among the A549 analyzed samples and the comparison between distinct types of cells (which are expected to have different editing sites), is encouraging.

Another goal was to validate the novel editing sites found by us, using the BodyMap2 dataset, that include seven different tissues: adipose, adrenal, brain, breast, colon, kidney, and heart. For that purpose, we searched for the new identified sites (**Supplementary Table 1**) in mock, A549, and Calu3 cell lines. Out of the full list of non-repetitive new editing sites that we found, 26.9% were also detected under the same category (non-repetitive) in A549 and Calu3 cells, although the cell lines are originated from the lungs that were not represented in our original analysis.

To understand whether the higher editing activity upon SARS-CoV-2 infection is manifested by a larger number of sites in the same genes as in mock, or additional sites located in new genes, we characterized the editing landscape with respect to site locations and gene annotation.

We first tested the ratio of shared sites between SARS-CoV-2 and mock samples. For both non-repetitive and hyper non-repetitive classes, almost ~70% of the sites were unique among samples infected with SARS-CoV-2, and about 60% of the sites were unique for mock samples (**Supplementary Table 2**). We next annotated the unique sites for each sample type, across different classes. It was apparent that the answer for our initial question is that among the unique editing sites following SARS-CoV-2 infection, some of the sites were extended the editing in baseline genes (considering the mock as the baseline), but most of the sites were in new genes (**Supplementary Table 2**).

Interestingly, the gene APOBEC3C became hyper edited following SARS-CoV-2 infection (**Supplementary Table 3**), while in mock samples it was classified under the non-repetitive class. The APOBEC family of enzymes edits C-to-U RNA modifications and known to be involved in regulation of innate immune response (Rosenberg et al., 2011; Schaefer et al., 2017). C-to-U editing of antibody-coding genes in the host's DNA leads to diversification of the repertoire of antibodies produced against viruses, called somatic hypermutation (SHM) (Cogné, 2013). Therefore, hyper editing in APOBEC genes may indicate for their involvement in COVID-19 phenotype, as part of a complex immune regulation system, controlling by A-to-I RNA editing.

To test if the sets of unique edited genes display shared biological processes, we examined their biological process enrichment, using the web-based tool GeneMANIA (Zuberi et al., 2013). Submitting the list of hyper-non-repetitive unique edited genes, upon SARS-CoV-2 infection, resulted in the enrichment

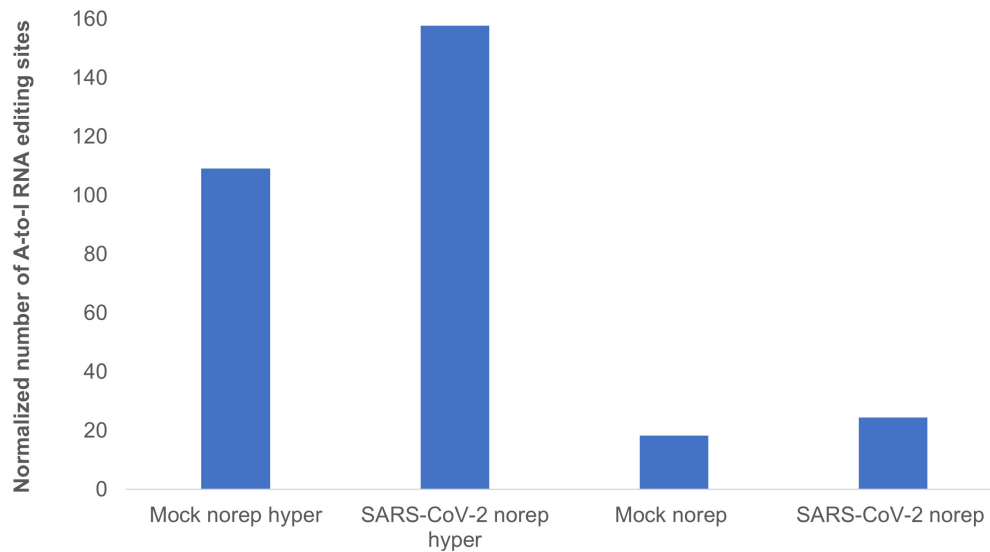


FIGURE 3 | A-to-I hyper editing sites are more frequent in SARS-CoV-2 infected cells compared to mock. Presented here are the numbers of A-to-I RNA editing sites detected in SARS-CoV-2 or Mock samples in total, normalized to total counts. Norep: non-repetitive editing sites, norep hyper: hyper non-repetitive editing sites. The number of non-repetitive hyper editing sites is significantly higher in SARS-CoV-2 samples compared to mock (P -value = 0.0371), and the number of non-repetitive sites is similar in both SARS-CoV-2 and mock (P -value > 0.05).

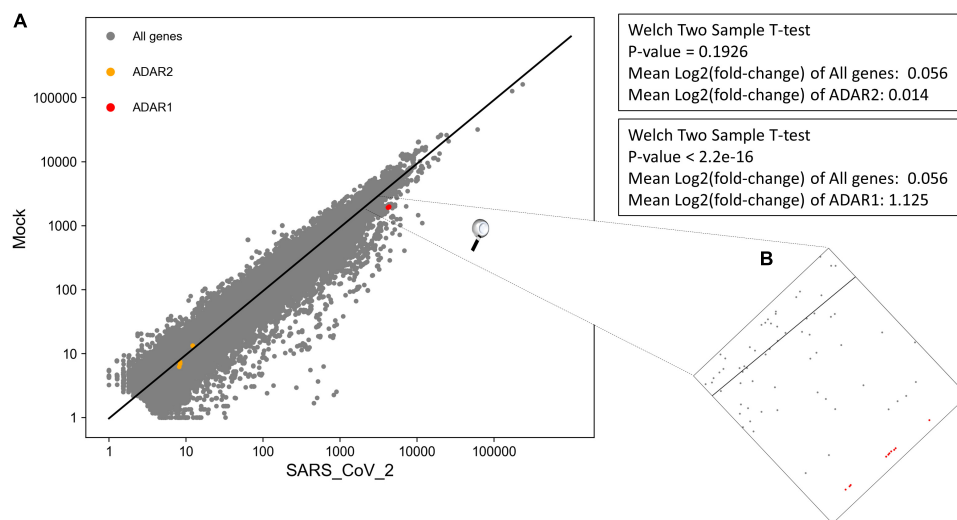


FIGURE 4 | Significant upregulation of ADAR1 isoforms, but not ADAR2, in cells infected with SARS-CoV-2 virus, compared to mock. **(A)** Log scale plot shows normalized gene counts from mock cells against cells infected by SARS-CoV-2 virus. Every dot in the graph represents a gene: ADAR1 isoforms (red), ADAR2 isoforms (orange), and all other genes (gray). The black line is the regression line for all genes. The P -values were obtained using a Welch two-sample T -test on only transcripts with coefficient variation > 1. **(B)** Zoom in on ADAR1 isoforms on the log scale plot. All 10 isoforms are closely located on the plot.

of processes related to the regulation of I-kappaB kinase/NF-kappaB signaling (P -adjusted values of related pathways in the range of 1.91E-03–3.50E-04; **Supplementary Table 3**). This result is intriguing in the light of strong indications suggesting that NF-kappaB pathway signaling has a critical role in controlling an excessive immune activation and ARDS (Kircheis et al., 2020). These indications, together with our result of hyper-editing in genes participating in the NF-kappaB pathway, suggest that A-to-I RNA editing activity may be critical to define the progression of COVID-19 disease and the risk to develop ARDS. We next,

tested an enrichment for genes that were classified under the non-repetitive class and were uniquely edited in SARS-CoV-2 samples. Strong enrichment was obtained for processes related to IFN response (P -adjusted values of related pathways in the range of 1.25E-06–1.03E-12; **Supplementary Table 4**), corroborating previous evidence that ADARs control IFN activation under viral infections (Baños-Lara et al., 2013), and suggests particularly that in COVID-19, ADARs control the level of immune response. We further run GeneMANIA for biological process enrichment with the mock unique gene sets, as a control. No processes related

to IFN or NF-kappaB signaling were enriched in $FDR < 0.05$ (**Supplementary Tables 5, 6**).

Given these observations, we reasoned that interrogating the differences in ADAR RNA expression levels between SARS-CoV-2 and mock treated samples, would help to complete the picture. Therefore, we performed DEA, on the same data of Blanco-Melo et al. (2020) used for RESIC. We first created the read counts for both SARA-CoV-2 and mock Calu3 samples (see “Methods and Definitions”). We then compared our created gene counts to those reported by Blanco-Melo et al. (2020) GEO accession number GSE147507. Although read count values were not identical, as expected due to the use of different alignment tools, the trend was the same.

We performed DEA to identify changes in the expression of ADAR genes and A-to-I RNA editing, following SARS-CoV-2 infection (**Supplementary Table 7**). We found that all 10 ADAR1 isoforms are significantly upregulated **Figure 4**, $P\text{-value} = 2.2e-16$) in SARS-CoV-2 Calu3 infected cells. In comparison, the expression of ADAR2 did not differ between SARS-CoV-2 and mock samples (**Figure 4**, $P\text{-value} > 0.05$). We also found that *IFIH1* (NM_022168.4), that encodes for MDA5 is significantly up-regulated in the SARS-CoV-2 infected samples (**Supplementary Table 7**, $P\text{-adjusted} = 7.46e-137$). This is in line with findings indicating that the IFN response upon SARS-CoV-2 infection is primarily regulated by MDA5 (Yin et al., 2021). Since ADAR1 is known to prevent MDA5 from sensing dsRNA (Liddicoat et al., 2015), this result strengthens the conclusion that ADAR1 is largely involved in the immune response following SARS-CoV-2.

Collectively, we suggest that upon SARS-CoV-2 infection, compared to mock (1) the number of hyper editing sites is increased; and (2) ADAR1 activity is enhanced. The combination between these two observations goes together with the finding that ADAR1 is the enzyme mostly catalyzing hyper editing sites (Porath et al., 2014).

We tested if these results hold true for more *in vitro* SARS-CoV-2 infected cell types created in the same study. For that purpose, we downloaded from GEO the already processed gene count data, for A549 and NHBE cells infected with SARS-CoV-2 high-multiplicity of infection (MOI). We chose downloading the already processed gene count data after validating for Calu3 cells that the gene count values created by us and the downloaded gene count from GEO (accession number GSE147507) are of the same trend (see “Methods and Definitions”). For A549 cells, with a vector expressing human ACE2, indeed ADAR1, but not ADAR2, was significantly upregulated following SARS-CoV-2 infection, corroborating our previous results for Calu3 cells. However, for NHBE cells, both ADAR1 and ADAR2 were not significantly changed after SARS-CoV-2 infection (**Supplementary Table 8**). The non-significant upregulation of ADAR1 in the last case may be because of the different cell types used. In any event, we concluded that unlike ADAR2 the expression of ADAR1 is substantially different upon SARS-CoV-2 infection, at least in some cell types.

Taken together, our results suggest that the catalyzation of hyper editing sites by ADAR1 is enhanced following SARS-CoV-2 infection. These results are intriguing in the context of ADAR1's role to block the IFN response, and particularly the role of hyper editing events to suppress the IFN induction (Vitali and Scadden, 2010). We hypothesize that editing levels might be indicative of the progression of COVID-19 disease and the risk to develop ARDS, as holds true in autoimmune diseases, due to the editing effect on the IFN response. Therefore, these results shed new light on the involvement of A-to-I RNA editing mechanism in COVID-19 disease. We note that supporting experimental validation is required to assess our conclusions. Our analysis encourages further exhaustive study of A-to-I RNA editing role in COVID-19 disease.

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

DL, RH, and ATL conceived and designed the study and wrote the manuscript with input from all authors. DL, RH, MY, TE, and TB implemented and developed RESIC. RH analyzed the data. ATL supervised the work. All authors contributed to the article and approved the submitted version.

FUNDING

This work was funded by the Israeli Centers of Research Excellence (I-CORE) program (Center No. 1796/12 to ATL), the Israel Science Foundation (grant no. 927/18 to ATL), and the NSF-BSF Molecular and Cellular Biosciences (MCB) (grant no. 2018738 to ATL).

ACKNOWLEDGMENTS

We thank Dan Bleiberg for providing technical support and Orna Ben-Naim Zgayer for her help with experimental validation. We also thank David Meiri, Yaniv Lerenthal, and Cannasoul analytics for providing samples.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2021.686851/full#supplementary-material>

REFERENCES

- Aho, A., Kernighan, B., and Weinberger, P. (1996). *Awk A Pattern Scanning and Processing Language*, Second Edn. Murray Hill, NJ: Bell Laboratories.
- Athanasiadis, A., Rich, A., and Maas, S. (2004). Widespread A-to-I RNA editing of Alu-containing mRNAs in the human transcriptome. *PLoS Biol.* 2:e391. doi: 10.1371/journal.pbio.0020391
- Baños-Lara, M. D. R., Ghosh, A., and Guerrero-Plata, A. (2013). Critical role of MDA5 in the interferon response induced by human metapneumovirus infection in dendritic cells and in vivo. *J. Virol.* 87, 1242–1251. doi: 10.1128/JVI.01213-12
- Barak, M., Levanon, E. Y., Eisenberg, E., Paz, N., Rechavi, G., Church, G. M., et al. (2009). Evidence for large diversity in the human transcriptome created by Alu RNA editing. *Nucleic Acids Res.* 37, 6905–6915. doi: 10.1093/nar/gkp729
- Bazak, L., Haviv, A., Barak, M., Jacob-Hirsch, J., Deng, P., Zhang, R., et al. (2014). A-to-I RNA editing occurs at over a hundred million genomic sites, located in a majority of human genes. *Genome Res.* 24, 365–376. doi: 10.1101/gr.164749.113
- Blanco-Melo, D., Nilsson-Payant, B. E., Liu, W.-C., Møller, R., Panis, M., Sachs, D., et al. (2020). SARS-CoV-2 launches a unique transcriptional signature from in vitro, ex vivo, and in vivo systems. *bioRxiv* [Preprint]. doi: 10.1101/2020.03.24.004655
- Blow, M., Futreal, P. A., Wooster, R., and Stratton, M. R. (2004). A survey of RNA editing in human brain. *Genome Res.* 14, 2379–2387. doi: 10.1101/gr.2951204
- Cogné, M. (2013). Activation-induced deaminase in B lymphocyte maturation and beyond. *Biomed. J.* 36, 259–268. doi: 10.1013/2319-4170.113191
- Eisenberg, E., and Levanon, E. Y. (2018). A-to-I RNA editing - immune protector and transcriptome diversifier. *Nat. Rev. Genet.* 19, 473–490. doi: 10.1038/s41576-018-0006-1
- Free Software Foundation (2019). *GNU Operating System. Free Software Foundation, I.* Available online at: <https://www.gnu.org/software/sed/> (accessed August 20, 2019).
- Galeano, F., Tomaselli, S., Locatelli, F., and Gallo, A. (2012). A-to-I RNA editing: the "ADAR" side of human cancer. *Semin. Cell. Dev. Biol.* 23, 244–250. doi: 10.1016/j.semcdb.2011.09.003
- Gallo, A., and Locatelli, F. (2012). ADARs: allies or enemies? The importance of A-to-I RNA editing in human disease: from cancer to HIV-1. *Biol. Rev. Camb. Philos. Soc.* 87, 95–110. doi: 10.1111/j.1469-185X.2011.00186.x
- Ganem, N. S., Ben-Asher, N., and Lamm, A. T. (2017). In cancer, A-to-I RNA editing can be the driver, the passenger, or the mechanic. *Drug Resistance Updates* 32, 16–22. doi: 10.1016/j.drug.2017.09.001
- Ganem, N. S., and Lamm, A. T. (2017). A-to-I RNA editing - thinking beyond the single nucleotide. *RNA Biol.* 14, 1690–1694. doi: 10.1080/15476286.2017.1364830
- Goldstein, B., Agranat-Tamir, L., Light, D., Ben-Naim Zgayer, O., Fishman, A., and Lamm, A. T. (2017). A-to-I RNA editing promotes developmental stage-specific gene and lncRNA expression. *Genome Res.* 27, 462–470.
- Kim, D. D., Kim, T. T., Walsh, T., Kobayashi, Y., Matisse, T. C., Buyske, S., et al. (2004). Widespread RNA editing of embedded alu elements in the human transcriptome. *Genome Res.* 14, 1719–1725. doi: 10.1101/gr.2855504
- Kirchels, R., Haasbach, E., Lueftenegger, D., Heyken, W. T., Ocker, M., and Planz, O. (2020). NF-kappaB pathway as a potential target for treatment of critical stage COVID-19 patients. *Front. Immunol.* 11:598444. doi: 10.3389/fimmu.2020.598444
- Kleinberger, Y., and Eisenberg, E. (2010). Large-scale analysis of structural, sequence and thermodynamic characteristics of A-to-I RNA editing sites in human Alu repeats. *BMC Genomics* 11:453. doi: 10.1186/1471-2164-11-453
- Kleinman, C. L., and Majewski, J. (2012). Comment on "widespread RNA and DNA sequence differences in the human transcriptome". *Science* 335, 1302–1302. doi: 10.1126/science.1209658
- Kung, C. P., Maggi, L. B. Jr., and Weber, J. D. (2018). The role of RNA editing in cancer development and metabolic disorders. *Front. Endocrinol. (Lausanne)* 9:762. doi: 10.3389/fendo.2018.00762
- Langmead, B., Trapnell, C., Pop, M., and Salzberg, S. L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* 10:R25. doi: 10.1186/gb-2009-10-3-r25
- Lee, S. Y., Joung, J. G., Park, C. H., Park, J. H., and Kim, J. H. (2015). RCARE: RNA Sequence Comparison and Annotation for RNA Editing. *BMC Med. Genomics* 8(Suppl. 2):S8. doi: 10.1186/1755-8794-8-S2-S8
- Levanon, E. Y., Eisenberg, E., Yelin, R., Nemzer, S., Hallegger, M., Shemesh, R., et al. (2004). Systematic identification of abundant A-to-I editing sites in the human transcriptome. *Nat. Biotechnol.* 22, 1001–1005. doi: 10.1038/nbt996
- Levanon, K., Eisenberg, E., Rechavi, G., and Levanon, E. Y. (2005). Letter from the editor: Adenosine-to-inosine RNA editing in Alu repeats in the human genome. *EMBO Rep.* 6, 831–835. doi: 10.1038/sj.embor.7400507
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics (Oxford, England)* 25, 278–279. doi: 10.1093/bioinformatics/btp352
- Li, M., Wang, I. X., Li, Y., Bruzel, A., Richards, A. L., Toung, J. M., et al. (2011). Widespread RNA and DNA sequence differences in the human transcriptome. *Science* 333, 53–58. doi: 10.1126/science.1207018
- Liddicoat, B. J., Piskol, R., Chalk, A. M., Ramaswami, G., Higuchi, M., Hartner, J. C., et al. (2015). RNA editing by ADAR1 prevents MDA5 sensing of endogenous dsRNA as nonself. *Science* 349, 1115–1120. doi: 10.1126/science.aac7049
- Love, M. I., 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
- Maas, S., Kawahara, Y., Tamburro, K. M., and Nishikura, K. (2006). A-to-I RNA editing and human disease. *RNA Biol.* 3, 1–9. doi: 10.4161/rna.3.1.2495
- Mannon, N. M., Greenwood, S. M., Young, R., Cox, S., Brindle, J., Read, D., et al. (2014). The RNA-editing enzyme ADAR1 controls innate immune responses to RNA. *Cell Rep.* 9, 1482–1494. doi: 10.1016/j.celrep.2014.10.041
- Moore, B. J. B., and June, C. H. (2020). Cytokine release syndrome in severe COVID-19. *Science* 368, 473–474. doi: 10.1126/science.abb8925
- Osenberg, S., Paz Yaacov, N., Safran, M., Moshkovitz, S., Shtrichman, R., Sherf, O., et al. (2010). Alu sequences in undifferentiated human embryonic stem cells display high levels of A-to-I RNA editing. *PLoS One* 5:e11173. doi: 10.1371/journal.pone.0011173
- Paz-Yaacov, N., Levanon, E. Y., Nevo, E., Kinar, Y., Harmelin, A., Jacob-Hirsch, J., et al. (2010). Adenosine-to-inosine RNA editing shapes transcriptome diversity in primates. *Proc. Natl. Acad. Sci. U.S.A.* 107, 12174–12179. doi: 10.1073/pnas.1006183107
- Picardi, E., Manzari, C., Mastropasqua, F., Aiello, I., D'Erchia, A. M., and Pesole, G. (2015). Profiling RNA editing in human tissues: towards the inosinome Atlas. *Sci. Rep.* 5:14941. doi: 10.1038/srep14941
- Picardi, E., and Pesole, G. (2013). REDIttools: high-throughput RNA editing detection made easy. *Bioinformatics* 29, 1813–1814. doi: 10.1093/bioinformatics/btt287
- Piechotta, M., Wyler, E., Ohler, U., Landthaler, M., and Dieterich, C. (2017). JACUSA: site-specific identification of RNA editing events from replicate sequencing data. *BMC Bioinform.* 18:7. doi: 10.1186/s12859-016-1432-8
- Pinto, Y., and Levanon, E. Y. (2019). Computational approaches for detection and quantification of A-to-I RNA-editing. *Methods (San Diego, Calif.)* 156, 25–31. doi: 10.1016/j.ymeth.2018.11.011
- Porath, H. T., Carmi, S., and Levanon, E. Y. (2014). A genome-wide map of hyper-edited RNA reveals numerous new sites. *Nat. Commun.* 5:4726. doi: 10.1038/ncomms5726
- Quin, J., Sedmik, J., Vukic, D., Khan, A., Keegan, L. P., and O'Connell, M. A. (2021). ADAR RNA modifications, the epitranscriptome and innate immunity. *Trends Biochem. Sci.* doi: 10.1016/j.tibs.2021.02.002 [Epub ahead of print].
- Ramaswami, G., and Li, J. B. (2014). RADAR: a rigorously annotated database of A-to-I RNA editing. *Nucleic Acids Res.* 42, 109–113. doi: 10.1093/nar/gkt996
- Ramaswami, G., Zhang, R., Piskol, R., Keegan, L. P., Deng, P., O'Connell, M. A., et al. (2013). Identifying RNA editing sites using RNA sequencing data alone. *Nat. Methods* 10, 128–132. doi: 10.1038/nmeth.2330
- Rosenberg, B. R., Hamilton, C. E., Mwangi, M. M., Dewell, S., and Papavasiliou, F. N. (2011). Transcriptome-wide sequencing reveals numerous APOBEC1 mRNA-editing targets in transcript 3' UTRs. *Nat. Struct. Mol. Biol.* 18, 230–236. doi: 10.1038/nsmb.1975
- Schaefer, M., Kapoor, U., and Jantsch, M. F. (2017). Understanding RNA modifications: the promises and technological bottlenecks of the 'epitranscriptome'. *Open Biol.* 7:170077. doi: 10.1098/rsob.170077
- Schneider, W. M., Chevillotte, M. D., and Rice, C. M. (2014). Interferon-stimulated genes: a complex web of host defenses. *Annu. Rev. Immunol.* 32, 513–545. doi: 10.1146/annurev-immunol-032713-120231

- Treangen, T. J., and Salzberg, S. L. (2011). Repetitive DNA and next-generation sequencing: computational challenges and solutions. *Nat. Rev. Genet.* 13, 36–46. doi: 10.1038/nrg3117
- Vitali, P., and Scadden, A. D. J. (2010). Double-stranded RNAs containing multiple IU pairs are sufficient to suppress interferon induction and apoptosis. *Nat. Struct. Mol. Biol.* 17, 1043–1050. doi: 10.1038/nsmb.1864
- Wang, Z., Lian, J., Li, Q., Zhang, P., Zhou, Y., Zhan, X., et al. (2016). RES-Scanner: a software package for genome-wide identification of RNA-editing sites. *GigaScience* 5:37. doi: 10.1186/s13742-016-0143-4
- Wu, D., Lamm, A. T., and Fire, A. Z. (2011). Competition between ADAR and RNAi pathways for an extensive class of RNA targets. *Nat. Struct. Mol. Biol.* 18, 1094–1101. doi: 10.1038/nsmb.2129
- Wu, J., Xie, F., Qian, K., Gibson, A. W., Edberg, J. C., and Kimberly, R. P. (2011). FAS mRNA editing in human systemic lupus erythematosus. *Hum. Mutat.* 32, 1268–1277. doi: 10.1002/humu.21565
- Yin, X., Riva, L., Pu, Y., Martin-Sancho, L., Kanamune, J., Yamamoto, Y., et al. (2021). MDA5 governs the innate immune response to SARS-CoV-2 in lung epithelial cells. *Cell Reports* 34:108628. doi: 10.1016/j.celrep.2020.108628
- Zhang, Q., and Xiao, X. (2015). Genome sequence-independent identification of RNA editing sites. *Nat. Methods* 12, 347–350. doi: 10.1038/nmeth.3314
- Zhu, S., Xiang, J.-F., Chen, T., Chen, L.-L., and Yang, L. (2013). Prediction of constitutive A-to-I editing sites from human transcriptomes in the absence of genomic sequences. *BMC Genomics* 14:206. doi: 10.1186/1471-2164-14-206
- Zuberi, K., Franz, M., Rodriguez, H., Montojo, J., Lopes, C. T., Bader, G. D., et al. (2013). GeneMANIA prediction server 2013 update. *Nucleic Acids Res.* 41, W115–W122. doi: 10.1093/nar/gkt533

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 Light, Haas, Yazbak, Elfand, Blau and Lamm. 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 Analysis of Methylated Adenosine Modifications Revealed Increased Levels of N^6 -Methyladenosine (m^6A) and $N^6,2'$ -O-Dimethyladenosine (m^6Am) in Serum From Colorectal Cancer and Gastric Cancer Patients

OPEN ACCESS

Edited by:

Jia Meng,
Xi'an Jiaotong-Liverpool University,
China

Reviewed by:

Tongxing Song,
Huazhong Agricultural University,
China
Quanqing Zhang,
University of California, Riverside,
United States

*Correspondence:

Ying Yuan
yuanying1999@zju.edu.cn
Cheng Guo
cheng_guo@zju.edu.cn

Specialty section:

This article was submitted to
Epigenomics and Epigenetics,
a section of the journal
Frontiers in Cell and Developmental
Biology

Received: 13 April 2021

Accepted: 21 June 2021

Published: 26 July 2021

Citation:

Hu Y, Fang Z, Mu J, Huang Y,
Zheng S, Yuan Y and Guo C (2021)
Quantitative Analysis of Methylated
Adenosine Modifications Revealed
Increased Levels
of N^6 -Methyladenosine (m^6A)
and $N^6,2'$ -O-Dimethyladenosine
(m^6Am) in Serum From Colorectal
Cancer and Gastric Cancer Patients.
Front. Cell Dev. Biol. 9:694673.
doi: 10.3389/fcell.2021.694673

Yiqiu Hu¹, Zhihao Fang¹, Jiayi Mu¹, Yanqin Huang¹, Shu Zheng¹, Ying Yuan^{2*} and Cheng Guo^{1,3*}

¹ Cancer Institute (Key Laboratory of Cancer Prevention and Intervention, China National Ministry of Education), The Second Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou, China, ² Department of Medical Oncology, The Second Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou, China, ³ Cancer Center, Zhejiang University, Hangzhou, China

Colorectal cancer and gastric cancer are the most prevalent gastrointestinal malignancies worldwide, and early detection of these cancers is crucial to reduce their incidence and mortality. RNA methylation plays an important regulatory role in a variety of physiological activities, and it has drawn great attention in recent years. Methylated adenosine (A) modifications such as N^6 -methyladenosine (m^6A), N^1 -methyladenosine (m^1A), $2'$ -O-methyladenosine (Am), $N^6,2'$ -O-dimethyladenosine (m^6Am), and N^6,N^6 -dimethyladenosine (m^6_2A) are typical epigenetic markers of RNA, and they are closely correlated to various diseases including cancer. Serum is a valuable source of biofluid for biomarker discovery, and determination of these adenosine modifications in human serum is desirable since they are emerging biomarkers for detection of diseases. In this work, a targeted quantitative analysis method using hydrophilic interaction liquid chromatography–tandem mass spectrometry (HILIC-MS/MS) was developed and utilized to analyze these methylated adenosine modifications in serum samples. The concentration differences between the healthy volunteers and cancer patients were evaluated by Mann–Whitney test, and receiver operator characteristic (ROC) curve analysis was performed to access the potential of these nucleosides as biomarkers. We demonstrated the presence of the m^6Am in human serum for the first time, and we successfully quantified the concentrations of A, m^6A , m^1A , and m^6Am in serum samples from 99 healthy controls, 51 colorectal cancer patients, and 27 gastric cancer patients. We found that the levels of m^6A and m^6Am in serum were both increased in colorectal cancer or gastric cancer patients, compared to that in healthy controls. These

results indicate that m^6A and m^6Am in serum may act as potential biomarkers for early detection and prognosis of colorectal cancer and gastric cancer. In addition, the present work will stimulate investigations on the effects of adenosine methylation on the initiation and progression of colorectal cancer and gastric cancer.

Keywords: colorectal cancer, gastric cancer, RNA methylation, methylated adenosine, hydrophilic interaction liquid chromatography-tandem mass spectrometry, biomarker

INTRODUCTION

Colorectal cancer and gastric cancer are two common malignancies that are the second and third causes of cancer-related deaths all over the world, respectively (Brenner et al., 2014; Siegel et al., 2017; Bray et al., 2018). The incidence and mortality of these two types of cancer have been rising due to the difficulties in early detection. Detection of these cancers in the early stage has the advantage of improving survival rate and reducing costs of patients. Currently, the screening and diagnosis of colorectal cancer and gastric cancer mainly depend on results of colonoscopy and gastroscopy, respectively, which are both invasive to patients, leading to low compliance. From this point of view, it is essential to discover novel non-invasive biomarkers for early detection of colorectal cancer and gastric cancer to elongate the survival time and decrease the pains of patients.

Posttranscriptional modifications of RNA play crucial regulatory roles in a number of physiological activities (Feinberg, 2018; Bates, 2020). Until now, more than 170 chemical modifications of RNA have been identified, and these modifications are involved in the regulation of RNA functions (Jonkhout et al., 2017; Ontiveros et al., 2019). Among these chemical modifications, RNA methylation modifications have attracted great attention since accumulating evidences have been obtained to confirm RNA methylation as a novel layer of epigenetic alteration. Although methylation modifications in RNA had been identified in the 1970s (Dubin and Taylor, 1975; Perry et al., 1975), understanding of the functions of RNA methylation was limited until the identification of regulatory proteins such as fat mass and obesity-associated protein (FTO) (Jia et al., 2011). Through their unique regulatory proteins including “writers,” “erasers,” and “readers,” RNA methylation modifications display distinct features associated with many key cellular functions such as splicing, stability, and translation of RNA (Boo and Kim, 2020; Shi et al., 2020). And it has been revealed that RNA methylation participates in the initiation and progression of a number of diseases including cancer (Fang et al., 2021).

As a typical product of RNA methylation, N^6 -methyladenosine (m^6A) is the most predominant modification in mRNA (Chen et al., 2019). The level of m^6A is dynamic and reversible, and this is ascribed to the regulation of m^6A methyltransferases and demethylases (Dominissini et al., 2012; Meyer and Jaffrey, 2017). The aberrant level of m^6A modification is closely related to tumorigenesis and development (Lan et al., 2019; Li et al., 2019; Zhang et al., 2019; Liu et al., 2020; Wang et al., 2020a,b; Jiang et al., 2021). In recent years, other methylated

adenosine modifications such as N^1 -methyladenosine (m^1A), $2'$ - O -methyladenosine (Am), $N^6,2'$ - O -dimethyladenosine (m^6Am), and N^6,N^6 -dimethyladenosine (m^6_2A) have also been revealed to play critical roles in the pathogenesis of various cancers (Safra et al., 2017; Sendinc et al., 2019; Chen et al., 2020a; Dong and Cui, 2020; Alriquet et al., 2021). Therefore, these methylated adenosine modifications have great potential to be indicators for early detection of diseases.

Serum contains a large number of biomolecules, and it is a preferred body fluid in the realm of biomarker discovery (Djukovic et al., 2010; Chen et al., 2011; Huang et al., 2018; Ottosson et al., 2019; Stepien et al., 2021). In the past decades, multiple analytical approaches have been utilized for the analysis of modified nucleosides (Beale et al., 2018; Farrokhi et al., 2018; Pero-Gascon et al., 2018; Guo et al., 2019). Compared with other analytical techniques, liquid chromatography tandem mass spectrometry (LC-MS/MS) is more favored to be used for biomarker discovery due to its great advantages in selectivity, sensitivity, accuracy, and high throughput. The analysis time could be largely shortened when ultra-performance liquid chromatography (UPLC) was used, and thus, UPLC-MS/MS becomes the preferred method for the determination of large-scale clinical samples (Guo et al., 2016, 2018a,b, 2020; Chen et al., 2020b; Zheng et al., 2020; Su et al., 2021). In addition, hydrophilic interaction liquid chromatography (HILIC) has the advantages of higher sensitivity because the organic solvent-rich mobile phase is more volatile and can enhance desolvation and ionization efficiency in the ion source. In the present study, a fast, sensitive, simple, and reliable HILIC-MS/MS method for targeted quantitative detection of methylated adenosine modifications including A , m^6A , m^1A , Am , m^6Am , and m^6_2A in human serum was established. By using the developed method, we revealed the presence of m^6Am in human serum for the first time and quantified A , m^6A , m^1A , and m^6Am in serum from 51 colorectal cancer patients, 27 gastric cancer patients, and 99 healthy volunteers. In addition, we carried out statistical analysis to compare the differences of these nucleosides between cancer patients and healthy controls and evaluated the potential of these adenosine modifications as biomarkers for early detection of colorectal cancer and gastric cancer.

MATERIALS AND METHODS

Chemicals and Reagents

Chromatographic-grade methanol was obtained from J.T. Baker (Radnor, PA, United States). Acetonitrile of HPLC grade

was bought from Merck KGaA (Darmstadt, Germany). A, m⁶A, m¹A, Am, m⁶Am, m⁶₂A, ¹³C₅-adenosine ([¹³C₅]A), D₃-N⁶-methyladenosine ([D₃]m⁶A), D₃-N¹-methyladenosine ([D₃]m¹A), and D₃-N⁶-2'-O-dimethyladenosine ([D₃]m⁶Am) were gained from Toronto Research Chemical (Toronto, ON, Canada). Acetic acid (CH₃COOH) was purchased from Fluka (Muskegon, MI, United States). Ammonium acetate and malic acid were bought from Sigma Aldrich (St Louis, MO, United States). Water was purified by using a Milli-Q water purification apparatus (Millipore, Milford, MA, United States).

Instrumentation

Targeted quantitative analyses of adenosine and methylated adenosine were carried out on an Acquity UPLC system (Waters, Milford, MA, United States) coupled with a 4000 QTRAP mass spectrometer (AB SCIEX, Foster City, CA, United States). A Waters BEH HILIC column (2.1 × 100 mm, 1.7 μm) was implemented for chromatographic separation at room temperature. Mass spectrometer was operated in electrospray ionization (ESI) positive ion mode, and data were acquired by using multiple-reaction monitoring (MRM) mode. Data acquisition and processing were controlled by Analyst 1.6.3 software.

Sample Collection

This research was approved by the Ethics Committee of the Second Affiliated Hospital, Zhejiang University School of Medicine (SAHZU). A total of 99 healthy controls (mean age of 50.3 ± 7.8 years, range 33–73 years) without serious infections or cancers, 51 patients with colorectal cancer (mean age of 63.2 ± 10.6 years, range 28–79 years), and 27 patients with gastric cancer (mean age of 65.9 ± 8.5 years, range 48–80 years) were recruited. The information of volunteers is shown in **Table 1**. All the subjects with cancers were confirmed by pathologist and did not undergo any type of therapy. Before participation, informed consent was provided by all volunteers. Then, the serum samples were collected in the early morning and reserved at −80°C until analysis.

Sample Preparation

Serum samples of 100 μl were placed into a 1.5-ml centrifuge tube and spiked with 10 μl of isotope-labeled internal standards (IS) after being thawed in ice. In order to perfectly remove the protein, 330 μl of prechilled methanol/acetonitrile (2:1, v/v) was added, and the mixture was vortexed for 1 min. After being placed at −20°C for 2 h, the obtained mixture was centrifuged at 13,000 rpm at 4°C for 15 min. Subsequently, 352 μl of supernatant was evaporated to dryness under vacuum. Then, the dried samples were redissolved in 80 μl of acetonitrile/water

(9:1, v/v). After vortexing for 10 s, ultrasonication for 15 s, and centrifuging at 13,000 rpm for 15 min at 4°C, in that order, 70 μl of the supernatant fraction was moved into the vial for HILIC-MS/MS detection.

HILIC-MS/MS Analysis

The mobile phases were (A) H₂O containing 10 mM ammonium acetate and 0.2% acetic acid and (B) acetonitrile containing 2 mM ammonium acetate, 0.2% acetic acid, and 0.05 mM malic acid. The desired sample separation was achieved by the optimized LC gradient program as follows: 0 min, 95% B; 3 min, 94% B; 3.5 min, 60% B; 5.5 min, 60% B; 6 min, 94% B; and 12.5 min, 94% B. The injection volume was 5 μl, and each sample was measured twice. The samples were maintained at 4°C. To minimize interference of the mass spectrometer, a switching valve was used, and the eluents from the column during 1.5–3.8 and 4.8–5.6 min were introduced into the ion source. The ion spray voltage was optimized as 5.5 kV. The ion source temperature (TEM) was set at 550°C. The pressure of ion source gas 1 (GS1) and ion source gas 2 (GS2) were both set at 50 psi. The curtain gas (CUR) was set at 40 psi. The dwell time was 45 ms for each ion transition.

For targeted quantitative analysis, the transitions of *m/z* 268.1→136.0 (A), *m/z* 282.1→150.0 (m⁶A and m¹A), *m/z* 282.1→136.0 (Am), *m/z* 296.1→150.0 (m⁶Am), and *m/z* 296.1→164.0 (m⁶₂A) were monitored. For isotope-labeled internal standards, the transitions of *m/z* 273.1→136.0 ([¹³C₅]A), *m/z* 285.1→153.0 ([D₃]m⁶A and [D₃]m¹A), and *m/z* 299.1→153.0 ([D₃]m⁶Am) were monitored. The optimized MRM parameters for these nucleosides including declustering potential (DP), entrance potential (EP), collision energy (CE), and collision cell exit potential (CXP) are listed in **Supplementary Table 1**.

Method Validation

In order to establish the calibration curves, different concentrations of the working solutions of nucleoside standards mixed with internal standards (IS) were made. The final concentrations of IS were as follows: [¹³C₅]A 25 nM, [D₃]m⁶A 2.5 nM, [D₃]m¹A 125 nM, and [D₃]m⁶Am 1.25 nM. By measuring the peak area ratio of the nucleosides to the corresponding IS (y), against the concentration of the analyte (x), the linearities were constructed as $y = ax + b$. In addition, the limit of detection (LOD) and limit of quantification (LOQ) were obtained by analyzing the concentrations of the standard solutions with signal-to-noise ratios equal to 3 and 10, respectively.

For the purpose of evaluating extraction recovery, serum samples were spiked with standards at three different levels of A (1, 5, and 50 nM), m⁶A (1, 5, and 50 nM), m¹A (10, 100, and 500 nM), and m⁶Am (0.1, 1, and 5 nM). After addition of the IS solution, the serum samples were processed and measured as mentioned above. The recovery (R) was estimated by using the following formula: $\text{recovery} = (\text{concentration in spiked sample} - \text{concentration in original sample}) / \text{spiked concentration} \times 100\%$.

In order to evaluate intra-day and inter-day precision, the quality control (QC) samples at three different levels of A (1,

TABLE 1 | The basic information of individuals recruited in this study.

Group	Normal	Colorectal cancer	Gastric cancer
Number of cases	99	51	27
Gender (man/woman)	70/29	28/23	17/10
Age (year)	50.26 ± 7.82	63.16 ± 10.64	65.39 ± 8.53

10, and 75 nM), m^6A (1, 10, and 75 nM), m^1A (10, 250, and 750 nM), and m^6Am (0.1, 2.5, and 20 nM) were measured on the same day and three consecutive days, respectively. The accuracy was acquired by comparing the obtained concentration to the theoretical value.

In consideration of matrix effect evaluation, a slope comparison method was applied. By adding standard solutions and IS to the serum and pure solvent, respectively, different calibration curves were subsequently established, and their slopes were compared. The slope ratio of the calibration curve established in serum to that in pure solvent was described as the matrix effect.

Statistics Analysis

Statistical analyses were performed by utilizing the SPSS statistics 24.0 software (IBM, Armonk, NY, United States). The concentration differences between healthy volunteers and cancer patients were assessed by the Mann–Whitney test, and a p -value less than 0.05 was significant. Receiver operating characteristic (ROC) analysis was used to assess the ability of these nucleosides to discriminate cancer patients from healthy controls.

RESULTS AND DISCUSSION

Optimization of Chromatographic Conditions and Mass Spectrometry Parameters

In order to achieve excellent chromatographic separation and obtain good chromatographic peak shape with appropriate retention, chromatographic conditions including the type of

column and the composition of mobile phase were optimized. The chemical structures of adenosine and its methylated modifications are illustrated in **Figure 1**. Since m^1A is positively charged, the retention of m^1A on a C18 column is very weak. It is eluted from the C18 column quickly, and a symmetric peak shape is hard to obtain. As a complementary tool, HILIC is widely used for the separation of compounds with high polarity, and we found that m^1A can be retained well on the HILIC column. Besides, a high proportion of organic solvent was used for HILIC separation, which could enhance the ionization efficiency of targets, leading to the improvement of detection sensitivity. Therefore, the BEH HILIC column (2.1×100 mm, $1.7 \mu m$) was selected for subsequent analysis. In our previous study, we found that malic acid could enhance the detection of methylated nucleosides in HILIC–MS/MS (Guo et al., 2018a). Hence, malic acid was added into the mobile phase for the analysis of methylated adenosine modifications. As illustrated in **Figure 2**, the analysis can be finished within 6 min, and target nucleosides were perfectly separated, which suggested that the analytical method is quick, has high throughput, and is fit for large clinical samples.

To optimize the MRM parameters, a peristaltic pump was used, and the standard solution was directly infused into the mass spectrometer. For A, $[M + H]^+$ ion at m/z 268.1 was observed in full-scan ESI–MS. Abundant $[M + H]^+$ ions at m/z 282.1 were both observed for m^6A and Am; for m^6Am and m^6_2A , $[M + H]^+$ ions at m/z 296.1 were both observed; and for m^1A , M^+ ion at m/z 282.1 was observed. Then, we carried out collision-induced dissociation (CID) experiments. According to the results of CID, for m^6A , m^1A , and m^6Am , the most abundant fragment ions were all at m/z 150.0. For A and Am, the most abundant fragment ions were both at m/z 136.0, and for m^6_2A , the most abundant

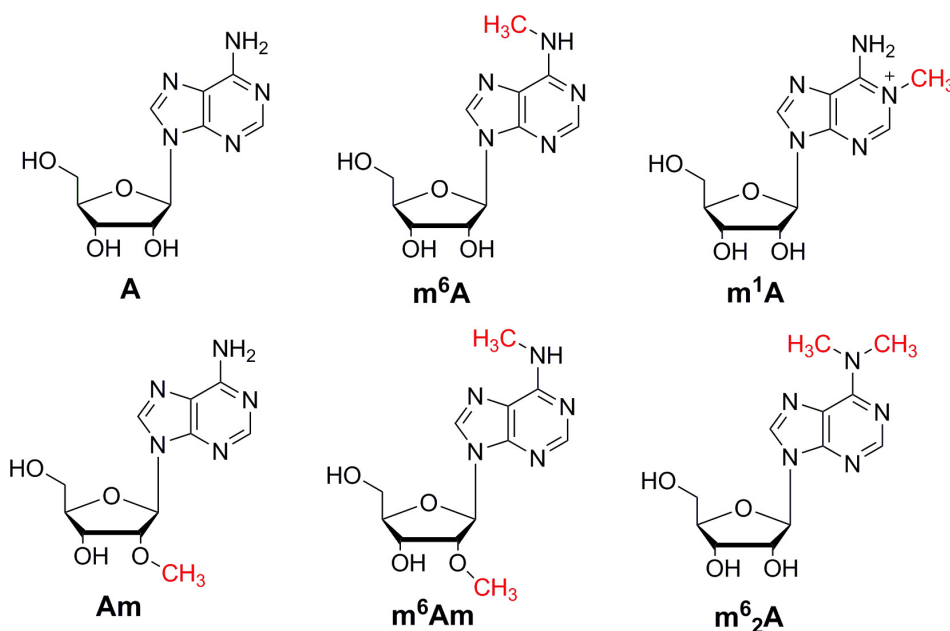
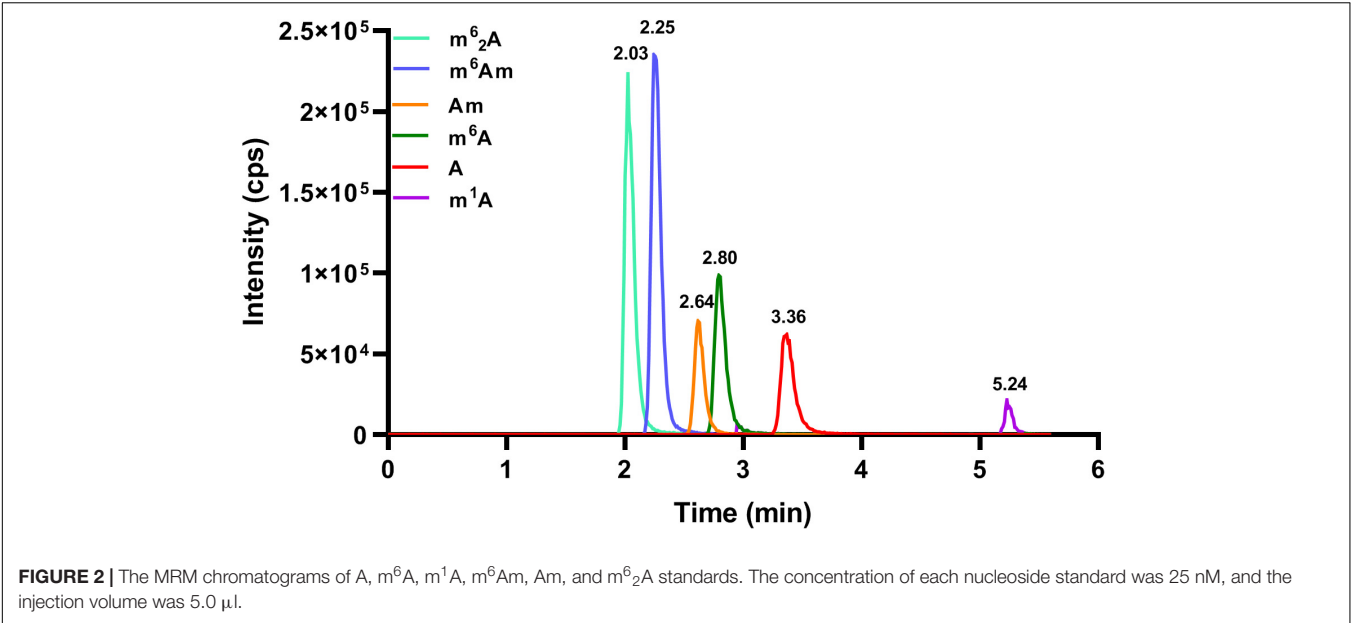


FIGURE 1 | The chemical structures of A, m^6A , m^1A , m^6Am , Am, and m^6_2A .



fragment ion was at m/z 164.0. From this point of view, the ion transitions m/z 268.1→136.0 (A), m/z 282.1→150.0 (m⁶A and m¹A), m/z 282.1→136.0 (Am), m/z 296.1→150.0 (m⁶Am), and m/z 296.1→164.0 (m⁶₂A) were utilized for quantitative determination. Similarly, the ion transitions m/z 273.1→136.0 ([¹³C₅]A), m/z 285.1→153.0 ([D₃]m⁶A and [D₃]m¹A), and m/z 299.1→153.0 ([D₃]m⁶Am) were monitored. Additionally, mass spectrometry parameters containing DP, EP, CE, CXP, ion source temperature (TEM), ion spray voltage, ion source gas 1 (GS1), ion source gas 2 (GS2), and curtain gas (CUR) were optimized to enhance the sensitivity and are summarized in **Supplementary Table 1**. Under these optimized conditions, the limit of detection (LOD) value can reach 0.0025 nM (A), 0.01 nM (m⁶A), 0.25 nM (m¹A), and 0.01 nM (m⁶Am), which are lower than the LOD values reported before (Djukovic et al., 2010; Chen et al., 2011; Guo et al., 2019), indicating that the analytical method has admirable sensitivity.

Validation of Analytical Method

For method validation, we investigated the linearity, recovery, LOD, LOQ, intra-day and inter-day precision, and matrix effects. As shown in **Table 2**, the calibration curve of each analyte showed excellent linearity ($R^2 > 0.999$) within a wide analytical range. The data of LOD and the LOQ of the analytes are also illustrated in **Table 2**, which indicated that the sensitivity of the developed method was superb. As for the matrix effect, calibration curves

were also constructed in serum extracts and in pure solvent. The slope ratio values for A, m⁶A, m¹A, and m⁶Am were 92.5, 96.9, 98.5, and 93.3%, respectively, which implied that the matrix had no interferences in this study.

As shown in **Supplementary Table 3**, recoveries ranged from 98.9 to 109.4%, indicating that satisfactory recovery was obtained. As shown in **Supplementary Table 4**, the intra-day precision value was within 6.0%, and the inter-day precision value was within 8.8%. The accuracy of the intra-day and inter-day analysis was in range of 89.0 to 105.5%. These data demonstrated that satisfactory reproducibility and accuracy were determined.

Moreover, to check the system stability after numbers of injections, QC samples were analyzed every 20 serum samples, and the accuracy and retention time were monitored. The results showed that the equipment system had an outstanding stability during measurement. In a word, all these results mentioned above revealed that the developed HILIC-MS/MS method is quick, sensitive, accurate, reliable, and reproducible. And it can meet the requirement for the targeted quantitative analysis of these nucleosides in serum samples.

Quantification of Methylated Adenosine Modifications in Human Serum

By applying the developed HILIC-MS/MS method, we investigated the contents of these nucleosides in serum

TABLE 2 | Linearities of A, m⁶A, m¹A, and m⁶Am in the HILIC-MS/MS analysis method.

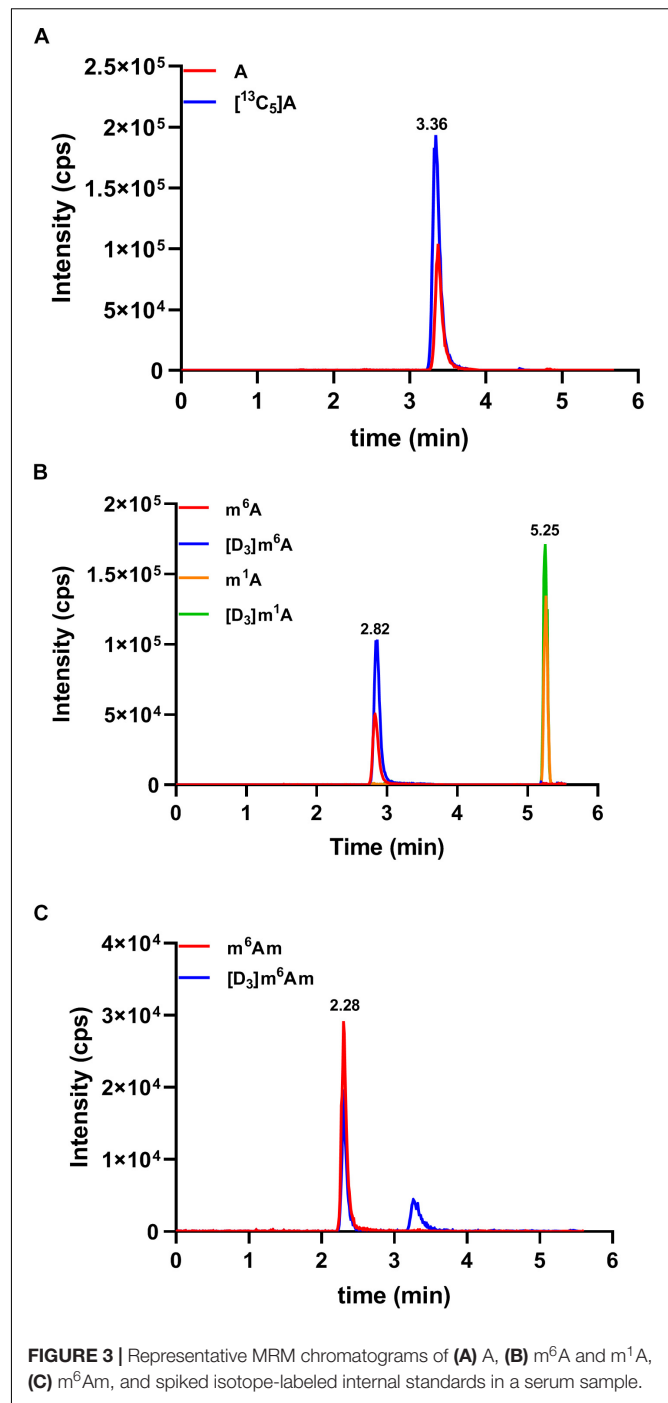
Compound	Linear equation	R ² value	Linear range (nM)	LOD (nM)	LOQ (nM)
A	$y = 0.0302x + 0.0235$	0.9998	0.5–100	0.0025	0.01
m ⁶ A	$y = 0.1205x + 0.0182$	0.9999	0.1–100	0.01	0.1
m ¹ A	$y = 0.0062x + 0.0053$	1	1–1,000	0.25	0.5
m ⁶ Am	$y = 0.7408x - 0.0032$	1	0.05–5	0.01	0.025

samples collected from 51 colorectal cancer patients, 27 gastric cancer patients, and 99 healthy controls. As demonstrated in **Figure 3**, the presence of A, m⁶A, m¹A, and m⁶Am could be unquestionably confirmed in all serum samples since the retention time of these four nucleosides was identical to that of their corresponding internal standard, whereas Am and m⁶2A were not detected in all the serum samples due to their trace amounts in serum. The levels of A, m⁶A, m¹A, and m⁶Am in serum samples were calculated according to the calibration curves, and the detailed concentrations were presented in **Supplementary Table 2**. In summary, the concentrations of A, m⁶A, m¹A, and m⁶Am in human serum ranged from 1.95 to 34.19, 2.24 to 9.74, 115.16 to 215.77, and 0.16 to 2.56 nM, respectively.

Content Change of Methylated Adenosine Modifications in Serum Samples From Cancer Patients

We next evaluated whether there were differences in the concentrations of these nucleosides between healthy controls and cancer patients. In serum samples, the concentrations of A, m⁶A, m¹A, and m⁶Am from healthy volunteers were in the range of 3.20–17.86, 2.24–7.73, 115.16–211.44, and 0.16–2.28 nM, respectively, and the average concentrations were 7.98 ± 3.01 , 4.51 ± 1.08 , 154.58 ± 21.10 , and 0.63 ± 0.37 nM, respectively ($n = 99$). The concentrations of A, m⁶A, m¹A, and m⁶Am in serum from patients with colorectal cancer were in the range of 2.41–27.94, 2.64–9.24, 117.45–215.77, and 0.17–2.56 nM, respectively, and the average concentrations were 8.47 ± 6.30 , 5.57 ± 1.67 , 158.62 ± 24.79 , and 1.13 ± 0.53 nM, respectively ($n = 51$). For patients with gastric cancers, the concentrations of A, m⁶A, m¹A, and m⁶Am in serum were in the range of 1.95–34.19, 4.94–9.74, 116.84–209.92, and 0.24–2.15 nM, respectively, and the average concentrations were 11.53 ± 8.55 , 6.93 ± 1.38 , 156.83 ± 27.07 , and 0.91 ± 0.57 nM, respectively ($n = 27$). As illustrated in **Figure 4**, it is apparent that the levels of m⁶A and m⁶Am in serum were intensely increased in patients with colorectal cancer or gastric cancers, compared to those in healthy controls (colorectal cancer: $p < 0.001$ for m⁶A and $p < 0.0001$ for m⁶Am; gastric cancer: $p < 0.0001$ for m⁶A and $p < 0.05$ for m⁶Am). However, there was no difference in the levels of A and m¹A between healthy controls and colorectal cancer or gastric cancer patients ($p > 0.05$).

Furthermore, ROC analysis was performed to evaluate the ability of these nucleosides to distinguish cancer patients from healthy controls. As demonstrated in **Figure 5**, for healthy volunteers and colorectal cancer patients, the area under the curve (AUC) for m⁶A and m⁶Am was 0.679 and 0.791, respectively. And for healthy controls and gastric cancer patients, the AUC was 0.931 and 0.647 for m⁶A and m⁶Am, respectively. These results indicated a correlation between the levels of m⁶A and m⁶Am in serum and the incidence of colorectal cancer and gastric cancer. Interestingly, we found that the AUC of m⁶Am was higher than that of m⁶A in colorectal cancer, whereas the AUC of m⁶A was higher than



that of m⁶Am in gastric cancer, implying that m⁶Am and m⁶A were more effective indicators of colorectal cancer and gastric cancer, respectively. Colorectal cancer and gastric cancer are of a high incidence worldwide and has awfully high mortality, and early detection is extremely necessary. The results of this study revealed that the increase of m⁶A and m⁶Am in serum might have great potential to be novel non-invasive biomarkers for early detection of colorectal cancer and gastric cancer.

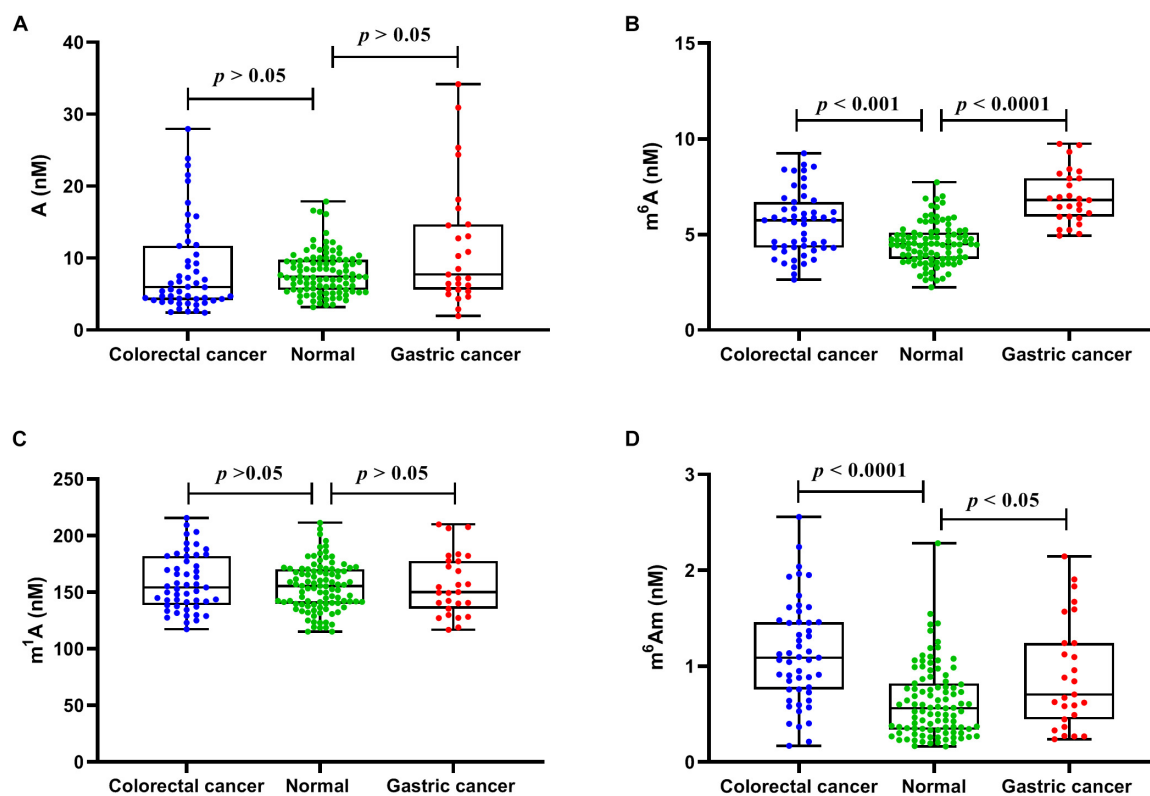


FIGURE 4 | The measured concentrations of (A) A, (B) m^6A , (C) m^1A , and (D) m^6Am in serum samples and statistical analysis.

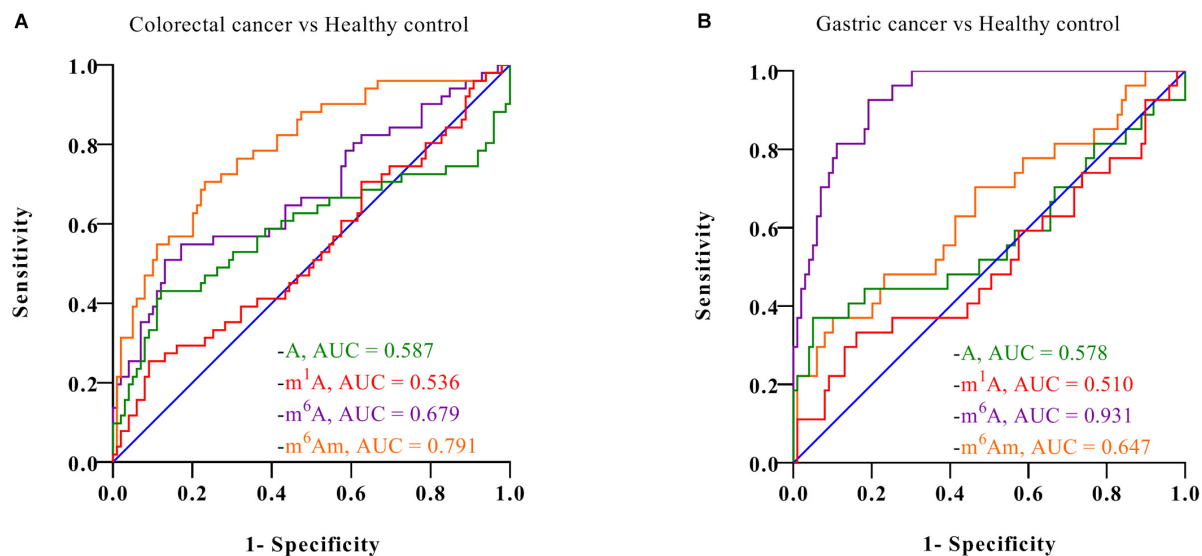


FIGURE 5 | ROC analysis for (A) colorectal cancer vs. healthy controls and (B) gastric cancer vs. healthy controls.

CONCLUSION

In the present study, we developed a robust, sensitive, and trustworthy HILIC-MS/MS method for targeted quantitative analysis of A, m^6A , m^1A , Am, m^6Am , and m^6_2A in human

serum. By applying the established method, we successfully revealed, for the first time, the presence of m^6Am in human serum and quantified the concentrations of A, m^6A , m^1A , and m^6Am in 177 serum samples from three groups, namely, healthy volunteers, colorectal cancer patients, and gastric cancer patients.

We found that the amount of m⁶A and m⁶Am were significantly higher in colorectal cancer or gastric cancer patients than those in healthy volunteers. Our data indicated that m⁶A and m⁶Am may serve as potential non-invasive biomarkers for early detection of colorectal cancer and gastric cancer. Furthermore, these results suggested that these methylated adenosine modifications might play important regulatory roles in the pathogenesis and progression of cancer.

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.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Institutional Review Board of Medical Research, The Second Affiliated Hospital, Zhejiang University School of Medicine. The patients/participants provided their written informed consent to participate in this study.

REFERENCES

- Alriquet, M., Calloni, G., Martinez-Limon, A., Delli Ponti, R., Hanspach, G., Hengesbach, M., et al. (2021). The protective role of m1A during stress-induced granulation. *J. Mol. Cell. Biol.* 12, 870–880. doi: 10.1093/jmcb/mjaa023
- Bates, S. E. (2020). Epigenetic therapies for cancer. *N. Engl. J. Med.* 383, 650–663. doi: 10.1056/NEJMra1805035
- Beale, D. J., Pinu, F. R., Kouremenos, K. A., Poojary, M. M., Narayana, V. K., Boughton, B. A., et al. (2018). Review of recent developments in GC–MS approaches to metabolomics-based research. *Metabolomics* 14:152. doi: 10.1007/s1306-018-1449-2
- Boo, S. H., and Kim, Y. K. (2020). The emerging role of RNA modifications in the regulation of mRNA stability. *Exp. Mol. Med.* 52, 400–408. doi: 10.1038/s12276-020-0407-z
- Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R. L., Torre, L. A., and Jemal, A. (2018). Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA. Cancer. J. Clin.* 68, 394–424. doi: 10.3322/caac.21492
- Brenner, H., Kloor, M., and Pox, C. P. (2014). Colorectal cancer. *Lancet.* 383, 1490–1502. doi: 10.1016/s0140-6736(13)61649-9
- Chen, F., Xue, J., Zhou, L., Wu, S., and Chen, Z. (2011). Identification of serum biomarkers of hepatocarcinoma through liquid chromatography/mass spectrometry-based metabolomic method. *Anal. Bioanal. Chem.* 401, 1899–1904. doi: 10.1007/s00216-011-5245-3
- Chen, H., Gu, L., Orellana, E. A., Wang, Y., Guo, J., Liu, Q., et al. (2020a). METTL4 is an snRNA m6Am methyltransferase that regulates RNA splicing. *Cell. Res.* 30, 544–547. doi: 10.1038/s41422-019-0270-4
- Chen, Q., Hu, Y., Fang, Z., Ye, M., Li, J., Zhang, S., et al. (2020b). Elevated levels of oxidative nucleic acid modification markers in urine from gastric cancer patients: quantitative analysis by ultra performance liquid chromatography-tandem mass spectrometry. *Front. Chem.* 8:606495. doi: 10.3389/fchem.2020.606495
- Chen, Y. G., Chen, R., Ahmad, S., Verma, R., Kasturi, S. P., Amaya, L., et al. (2019). N6-methyladenosine modification controls circular RNA immunity. *Mol. Cell.* 76, 96–109. doi: 10.1016/j.molcel.2019.07.016
- Djukovic, D., Baniyadi, H. R., Kc, R., Hammoud, Z., and Raftery, D. (2010). Targeted serum metabolite profiling of nucleosides in esophageal

AUTHOR CONTRIBUTIONS

CG and YY designed the study. YiH, CG, and ZF performed the experiments. ZF and JM collected the serum samples. YiH and CG analyzed and interpreted the data and wrote the manuscript. YaH and SZ edited the manuscript. All authors commented and approved the final manuscript.

FUNDING

This study was supported by the National Key R&D Program of China (2016YFC1302803), Key R&D Program of Zhejiang Province (2021C03125), Natural Science Foundation of Zhejiang Province (LY19B050007), and National Natural Science Foundation of China (21402172).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fcell.2021.694673/full#supplementary-material>

- adenocarcinoma. *Rapid. Commun. Mass. Spectrom.* 24, 3057–3062. doi: 10.1002/rcm.4739
- Dominissini, D., Moshitch-Moshkovitz, S., Schwartz, S., Salmon-Divon, M., Ungar, L., Osenberg, S., et al. (2012). Topology of the human and mouse m6A RNA methylomes revealed by m6A-seq. *Nature.* 485, 201–206. doi: 10.1038/nature11112
- Dong, Z., and Cui, H. (2020). The emerging roles of RNA modifications in glioblastoma. *Cancers. (Basel)* 12:736. doi: 10.3390/cancers12030736
- Dubin, D. T., and Taylor, R. H. (1975). The methylation state of poly a-containing messenger RNA from cultured hamster cells. *Nucleic. Acids. Res.* 2, 1653–1668. doi: 10.1093/nar/2.10.1653
- Fang, Z., Hu, Y., Hu, J., Huang, Y., Zheng, S., and Guo, C. (2021). The crucial roles of N6-methyladenosine m6A modification in the carcinogenesis and progression of colorectal cancer. *Cell Biosci.* 11:72. doi: 10.1186/s13578-021-00583-8
- Farrokhi, V., Chen, X., and Neubert, H. (2018). Protein turnover measurements in human serum by serial immunoaffinity LC-MS/MS. *Clin. Chem.* 64, 279–288. doi: 10.1373/clinchem.2017.272922
- Feinberg, A. P. (2018). The key role of epigenetics in human disease prevention and mitigation. *N. Engl. J. Med.* 378, 1323–1334. doi: 10.1056/NEJMra1402513
- Guo, C., Chen, Q., Chen, J., Yu, J., Hu, Y., Zhang, S., et al. (2020). 8-Hydroxyguanosine as a possible RNA oxidative modification marker in urine from colorectal cancer patients: evaluation by ultra performance liquid chromatography-tandem mass spectrometry. *J. Chromatogr. B. Analyt. Technol. Biomed. Life. Sci.* 1136:121931. doi: 10.1016/j.jchromb.2019.121931
- Guo, C., Li, X., Wang, R., Yu, J., Ye, M., Mao, L., et al. (2016). Association between oxidative DNA damage and risk of colorectal cancer: sensitive determination of urinary 8-hydroxy-2'-deoxyguanosine by UPLC-MS/MS analysis. *Sci. Rep.* 6:32581. doi: 10.1038/srep32581
- Guo, C., Xie, C., Chen, Q., Cao, X., Guo, M., Zheng, S., et al. (2018a). A novel malic acid-enhanced method for the analysis of 5-methyl-2'-deoxycytidine, 5-hydroxymethyl-2'-deoxycytidine, 5-methylcytidine and 5-hydroxymethylcytidine in human urine using hydrophilic interaction liquid chromatography-tandem mass spectrometry. *Analytica. Chimica. Acta.* 1034, 110–118. doi: 10.1016/j.aca.2018.06.081
- Guo, C., Xie, C., Ding, P., Qin, G., Mo, W., Cao, X., et al. (2018b). Quantification of glycolic acid in human serum by stable isotope dilution ultra performance

- liquid chromatography electrospray ionization tandem mass spectrometry. *J. Chromatogr. B. Analyt. Technol. Biomed. Life. Sci.* 1072, 315–319. doi: 10.1016/j.jchromb.2017.11.037
- Guo, M., Liu, D., Sha, Q., Geng, H., Liang, J., and Tang, D. (2019). Succinic acid enhanced quantitative determination of blood modified nucleosides in the development of diabetic nephropathy based on hydrophilic interaction liquid chromatography mass spectrometry. *J. Pharm. Biomed. Anal.* 164, 309–316. doi: 10.1016/j.jpba.2018.10.042
- Huang, J., Weinstein, S. J., Moore, S. C., Derkach, A., Hua, X., Liao, L. M., et al. (2018). Serum metabolomic profiling of all-cause mortality: a prospective analysis in the alpha-tocopherol, beta-carotene cancer prevention (ATBC) study cohort. *Am. J. Epidemiol.* 187, 1721–1732. doi: 10.1093/aje/kwy017
- Jia, G., Fu, Y., Zhao, X., Dai, Q., Zheng, G., Yang, Y., et al. (2011). N6-methyladenosine in nuclear RNA is a major substrate of the obesity-associated FTO. *Nat. Chem. Biol.* 7, 885–887. doi: 10.1038/nchembio.687
- Jiang, X., Liu, B., Nie, Z., Duan, L., Xiong, Q., Jin, Z., et al. (2021). The role of m6A modification in the biological functions and diseases. *Signal. Transduct. Target. Ther.* 6:74. doi: 10.1038/s41392-020-00450-x
- Jonkhout, N., Tran, J., Smith, M. A., Schonrock, N., Mattick, J. S., and Novoa, E. M. (2017). The RNA modification landscape in human disease. *RNA* 23, 1754–1769. doi: 10.1261/rna.063503.117
- Lan, Q., Liu, P. Y., Haase, J., Bell, J. L., Huttelmaier, S., and Liu, T. (2019). The critical role of RNA m6A methylation in cancer. *Cancer. Res.* 79, 1285–1292. doi: 10.1158/0008-5472.CAN-18-2965
- Li, T., Hu, P. S., Zuo, Z., Lin, J. F., Li, X., Wu, Q. N., et al. (2019). METTL3 facilitates tumor progression via an m6A-IGF2BP2-dependent mechanism in colorectal carcinoma. *Mol. Cancer* 18:112. doi: 10.1186/s12943-019-1038-7
- Liu, T., Yang, S., Sui, J., Xu, S. Y., Cheng, Y. P., Shen, B., et al. (2020). Dysregulated N6-methyladenosine methylation writer METTL3 contributes to the proliferation and migration of gastric cancer. *J. Cell. Physiol.* 235, 548–562. doi: 10.1002/jcp.28994
- Meyer, K. D., and Jaffrey, S. R. (2017). Rethinking m6A readers, writers, and erasers. *Annu. Rev. Cell. Dev. Biol.* 33, 319–342. doi: 10.1146/annurev-cellbio-100616-060758
- Ontiveros, R. J., Stoute, J., and Liu, K. F. (2019). The chemical diversity of RNA modifications. *Biochem. J.* 476, 1227–1245. doi: 10.1042/BCJ20180445
- Ottosson, F., Smith, E., Gallo, W., Fernandez, C., and Melander, O. (2019). Purine metabolites and carnitine biosynthesis intermediates are biomarkers for incident type 2 diabetes. *J. Clin. Endocrinol. Metab.* 104, 4921–4930. doi: 10.1210/je.2019-00822
- Pero-Gascon, R., Sanz-Nebot, V., Berezovski, M. V., and Benavente, F. (2018). Analysis of circulating microRNAs and their post-transcriptional modifications in cancer serum by on-line solid-phase extraction-capillary electrophoresis-mass spectrometry. *Anal. Chem.* 90, 6618–6625. doi: 10.1021/acs.analchem.8b00405
- Perry, R. P., Kelley, D. E., Friderici, K., and Rottman, F. (1975). The methylated constituents of L cell messenger RNA: Evidence for an unusual cluster at the 5' terminus. *Cell* 4, 387–394. doi: 10.1016/0092-8674(75)90159-2
- Safra, M., Sas-Chen, A., Nir, R., Winkler, R., Nachshon, A., Bar-Yaacov, D., et al. (2017). The m1A landscape on cytosolic and mitochondrial mRNA at single-base resolution. *Nature* 551, 251–255. doi: 10.1038/nature24456
- Sendinc, E., Valle-Garcia, D., Dhall, A., Chen, H., Henriques, T., Navarrete-Perea, J., et al. (2019). PCIF1 catalyzes m6Am mRNA methylation to regulate gene expression. *Mol. Cell* 75, 620–630. doi: 10.1016/j.molcel.2019.05.030
- Shi, H., Chai, P., Jia, R., and Fan, X. (2020). Novel insight into the regulatory roles of diverse RNA modifications: re-defining the bridge between transcription and translation. *Mol. Cancer* 19:78. doi: 10.1186/s12943-020-01194-6
- Siegel, R. L., Miller, K. D., Fedewa, S. A., Ahnen, D. J., Meester, R. G. S., Barzi, A., et al. (2017). Colorectal cancer statistics, 2017. *CA. Cancer. J. Clin.* 67, 177–193. doi: 10.3322/caac.21395
- Stepien, M., Keski-Rahkonen, P., Kiss, A., Robinot, N., Duarte-Salles, T., Murphy, N., et al. (2021). Metabolic perturbations prior to hepatocellular carcinoma diagnosis: Findings from a prospective observational cohort study. *Int. J. Cancer* 148, 609–625. doi: 10.1002/ijc.33236
- Su, X., Li, X., Wang, H., and Cai, Z. (2021). Simultaneous determination of methionine cycle metabolites, urea cycle intermediates and polyamines in serum, urine and intestinal tissue by using UHPLC-MS/MS. *Talanta* 224:121868. doi: 10.1016/j.talanta.2020.121868
- Wang, Q., Chen, C., Ding, Q., Zhao, Y., Wang, Z., Chen, J., et al. (2020a). METTL3-mediated m6A modification of HDGF mRNA promotes gastric cancer progression and has prognostic significance. *Gut* 69, 1193–1205. doi: 10.1136/gutjnl-2019-319639
- Wang, T., Kong, S., Tao, M., and Ju, S. (2020b). The potential role of RNA N6-methyladenosine in cancer progression. *Mol. Cancer* 19:88. doi: 10.1186/s12943-020-01204-7
- Zhang, Y., Kang, M., Zhang, B., Meng, F., Song, J., Kaneko, H., et al. (2019). m6A modification-mediated CBX8 induction regulates stemness and chemosensitivity of colon cancer via upregulation of LGR5. *Mol. Cancer* 18:185. doi: 10.1186/s12943-019-1116-x
- Zheng, F., Zhao, X., Zeng, Z., Wang, L., Lv, W., Wang, Q., et al. (2020). Development of a plasma pseudotargeted metabolomics method based on ultra-high-performance liquid chromatography-mass spectrometry. *Nat. Protoc.* 15, 2519–2537. doi: 10.1038/s41596-020-0341-5

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 Hu, Fang, Mu, Huang, Zheng, Yuan and Guo. 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.



RNA Motifs and Modification Involve in RNA Long-Distance Transport in Plants

Tao Wang[†], Xiaojun Li[†], Xiaojing Zhang, Qing Wang, Wenqian Liu, Xiaohong Lu, Shunli Gao, Zixi Liu, Mengshuang Liu, Lihong Gao and Wenna Zhang*

Beijing Key Laboratory of Growth and Developmental Regulation for Protected Vegetable Crops, China Agricultural University, Beijing, China

OPEN ACCESS

Edited by:

Xiao Han,
Biotechnology Research Institute,
Chinese Academy of Agricultural
Sciences, China

Reviewed by:

Li Pu,
Biotechnology Research Institute,
Chinese Academy of Agricultural
Sciences, China
Xiaofeng Gu,
Institute of Biotechnology (CAAS),
China

*Correspondence:

Wenna Zhang
zhangwenna@cau.edu.cn;
wennafhx@163.com

[†] These authors have contributed
equally to this work

Specialty section:

This article was submitted to
Epigenomics and Epigenetics,
a section of the journal
Frontiers in Cell and Developmental
Biology

Received: 09 January 2021

Accepted: 22 February 2021

Published: 01 April 2021

Citation:

Wang T, Li X, Zhang X, Wang Q,
Liu W, Lu X, Gao S, Liu Z, Liu M,
Gao L and Zhang W (2021) RNA
Motifs and Modification Involve
in RNA Long-Distance Transport
in Plants.
Front. Cell Dev. Biol. 9:651278.
doi: 10.3389/fcell.2021.651278

A large number of RNA molecules have been found in the phloem of higher plants, and they can be transported to distant organelles through the phloem. RNA signals are important cues to be evolving in fortification strategies by long-distance transportation when suffering from various physiological challenges. So far, the mechanism of RNA selectively transportation through phloem cells is still in progress. Up to now, evidence have shown that several RNA motifs including Polypyrimidine (poly-CU) sequence, transfer RNA (tRNA)-related sequence, Single Nucleotide Mutation bound with specific RNA binding proteins to form Ribonucleotide protein (RNP) complexes could facilitate RNA mobility in plants. Furthermore, some RNA secondary structure such as tRNA-like structure (TLS), untranslated region (UTR) of mRNA, stem-loop structure of pre-miRNA also contributed to the mobility of RNAs. Latest researchs found that RNA methylation such as methylated 5' cytosine (m5C) played an important role in RNA transport and function. These studies lay a theoretical foundation to uncover the mechanism of RNA transport. We aim to provide ideas and clues to inspire future research on the function of RNA motifs in RNA long-distance transport, furthermore to explore the underlying mechanism of RNA systematic signaling.

Keywords: RNA motif, RNA transport, RNA methylation, RNA structure, TLS

INTRODUCTION

As the main transportation pathway, higher plants' vascular system, including xylem and phloem plays an important role in process of growth and development (Lucas et al., 2013). Sugars, amino acids, phytohormones, and macromolecules such as RNA, proteins are transported from source to distant sink through the vascular system, in which system macromolecules upload from companion cells (CC) to sieve tube (SE) via plasmodesmata (PD; Notaguchi and Okamoto, 2015). Till now, cDNA library and omics profiling have been commendably established to identify a wide range of RNA signaling various plants species, such as cucumber (*Cucumis sativus* L.) (Zhang et al., 2016; Liu et al., 2020), pumpkin (*Cucurbit maxima*) (Omid et al., 2007), *Nicotiana benthamiana*/tomato (*Solanum lycopersicum*) (Xia et al., 2018), *Arabidopsis* (*Arabidopsis thaliana*) (Thieme et al., 2015), and grapes (*Vitis vinifera*) (Yang et al., 2015). Evidence of biological studies *in vivo* have indicated that the regulatory function of RNAs (non-coding RNA and mRNA) were involving plant growth, defense,

biological/economic yield, and certain morphological characteristics (Banerjee et al., 2009; Huang and Yu, 2009).

It has been reported that RNA local mobile from CC to SE are following a selective way. Properties affect RNA movement or upload to SE including transcript length, shape, charge, and genetic sequence or structure motifs (Morris, 2018; Maizel et al., 2020). One assumption about the intercellular and long-distance transport mechanism of RNA is that many RNAs have different structure motifs. They may guide RNAs transport across different types of cell boundaries, or these motifs can be recognized and combined with the transport protein, finally, trigger the transportation in the phloem. Therefore, we focus on the mechanism of long-distance RNA transport and the role of mRNA in regulating transcription and translation. In addition, the potential function of mobile proteins encoding by the mRNAs also needs to be explored.

Here, we summarize some of the findings of RNA transport mechanisms, mainly related to RNA motifs including Polypyrimidine (poly-CU) sequence, transfer RNA (tRNA)-related sequence, Single Nucleotide Mutation. Furthermore, some RNA secondary structures such as tRNA-like structure (TLS), untranslated region (UTR) of mRNA, stem-loop structure of pre-miRNA and RNA methylation such as methylated 5' cytosine (m5C) also contribute to the mobility of RNAs in plants. All the evidence of these RNA motifs that assist RNA transportation provides us a better understanding of the long-distance transport of plant endogenous RNA mechanism.

RNA MOTIF AND STRUCTURE PLAY AN IMPORTANT ROLE IN VIROID TRANSPORT AND PATHOGENICITY

The pathogenicity of viroid in plants was determined by their single-nucleotide motif, secondary, and tertiary structure (Zhong et al., 2008; Steger and Perreault, 2016). Being composed of 246 to 401 nucleotides, viroid was a kind of single-stranded, circular and non-coding RNA. Viroid RNA, including base-paired double-stranded regions and unpaired single circular stranded regions, exists in the form of highly base paired rod-like structure in nature (Góra-Sochacka, 2004; Flores et al., 2005; Ding, 2009).

Firstly, different base sites in RNA loop of *PSTVd* could affect the systemic infection and pathogenicity of *PSTVd*. *PSTVd* consists 359 RNA nucleotides folded into a secondary structure, which contains 26 base pair stems and 27 loops, at least 11 loops were essential for the systemic transport of *PSTVd* in *N. benthamiana* (Owens, 2007; Ding, 2009). For example, loop7 played a key role in mediating the transport of *PSTVd* from bundle sheath to phloem. Loop 6 was very important for the transport of *PSTVd* from palisade mesophyll to spongy mesophyll (Takeda et al., 2011). The bases in loop27 will affect the transport of *PSTVd* from epidermis to palisade mesophyll cells in *N. benthamiana* (Figure 1A) (Wu et al., 2019).

Then, secondary and tertiary structure of viroid may recognize RNA binding protein to facilitate long-distance transport in phloem of plants (Leontis et al., 2002, 2006; Noller, 2005; Zhong et al., 2007). Viroids harboring RNA motifs affected the

reorganization on binding protein and transportation between plant cells (Figure 1B) (Takeda and Ding, 2009). For example, the structure formed by loop27 in *PSTVd* was similar to the stem-loop structure formed of histone mRNA in 3'UTR in animal cells, this region can bind to Stem-loop Binding Protein (SLBP; Skrajna et al., 2017). The interaction of Cucumber Phloem Lectin (CsPP2) with *Hop stunt viroid* (*HSVd*) both *in vitro* and *in vivo* (Gómez and Pallás, 2004) and the interaction of two other proteins CmmLec and a 14 kDa protein with *ASBVd* suggested that the proteins may play a role in viroid RNA transport (Gómez et al., 2005). Tomato Viroid RNA-binding protein 1 (VirP1) binds to the right-terminal region of (+)-*PSTVd*, and it also participates in the transport of *PSTVd* by interacting with other proteins (Gozmanova et al., 2003; Maniataki et al., 2003). More evidence proved that there existed a recognition between circular (+)-*PSTVd* and the largest subunit of Pol II, DNA-dependent RNA polymerase (DdRP; Wang et al., 2016). Base on nuclear magnetic resonance (NMR) and X-ray crystallography, most loops in *PSTVd* were found form three-dimensional (3D) motifs through non-Watson Crick (non-WC) base pairing, base stacking and other base-base interactions (Wu et al., 2019). Therefore, RNA sequence evolution constrained by 3D structure and RNA-protein interactions may act most in regulating viroid RNA transport and the establishment of distinct cellular boundaries.

POLYPYRIMIDINE MOTIF RECOGNIZING RNA BINDING PROTEIN MEDIATING mRNA LONG-DISTANCE TRANSPORT

Endogenous long-distance transmissible mRNA in plants showed conserved sequence motifs such as UAGGUUA and ACUUCU in the 3'UTR region, which may be important RNA sequences for protein recognition (Rosin et al., 2003). A kind of polypyrimidine binding protein (PTB) recognized and combined mRNA forming Ribonucleic Protein (RNP) complexes to keep mRNA stability through Poly-CU motifs, facilitating RNP complex selectively transport (Ham et al., 2009). The 3'UTR of POTH1 mRNA is rich in polypyrimidine sequences, and 3'UTR can bind to polypyrimidine tract RNA-binding protein *in vitro* test. These evidence suggest that 3'UTR probably mediate the long-distance transport of POTH1 mRNA from stem to root (Mahajan et al., 2012; Ham and Lucas, 2017).

PTB proteins such as CmRBP50, StPTB, PbPTB3 detected in the phloem of pumpkin (*Cucurbita maxima*), potato (*Solanum tuberosum*), and pear (*Pyrus betulaefolia*) were confirmed binding the CU-rich motifs of *CmGAIP*, *StBEL5*, and *PbWoxT1*, thereby facilitating long-distance movement of mRNAs (Ham et al., 2009; Li et al., 2011; Cho et al., 2015; Duan et al., 2016). Furthermore, a protein complex containing five conserved domains of WD40, Transparent TESTA Glabra1 (PbTTG1) in pear (*P. betulaefolia*), was identified interaction with PbPTB3. The protein complex PbTTG1 and PbPTB3 promote *PbWoxT1* mRNA transport in grafted pear (Wang et al., 2019). Those findings provided important insights into polypyrimidine motif recognizing RNA binding protein mediating mRNA long-distance transport in phloem.

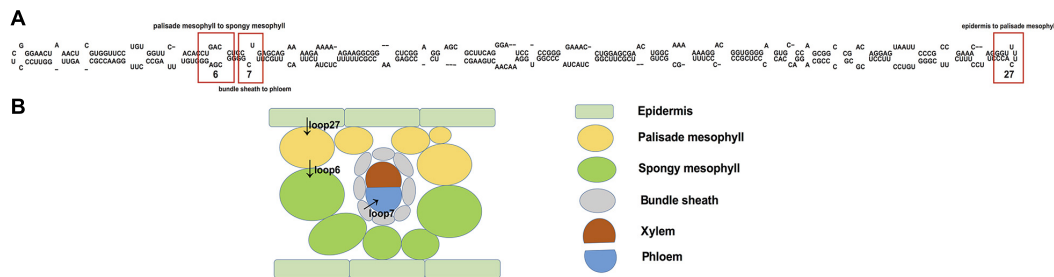


FIGURE 1 | Functional analysis of PSTVd secondary structure and pathways for systemic trafficking of PSTVd in an infected plant. **(A)** PSTVd consists 359 RNA nucleotides folded into a secondary structure, which contains 26 base pair stems and 27 loops. Loop 6, 7, 27 represent different unpaired single circular stranded regions of PSTVd. The nucleosides U43/C318 in PSTVd loop7 played a key role in mediating the transport of PSTVd from bundle sheath to phloem. The six nucleosides of loop 6, G36, A37, C38, C323, G324, and A325, when the mutated nucleosides (G36U/A37C/C38G) close loop 6, its systematic transport was limited, and loop 6 was very important for the transport of PSTVd from palisade mesophyll to spongy mesophyll. The U178G/U179G double mutant in loop27 has replication ability in tobacco cells, but it can only spread in the upper epidermis of inoculated leaves, suggesting that the bases in loop27 will affect the transport of PSTVd from epidermis to palisade mesophyll cells in *N. benthamiana*. **(B)** The black arrows illustrate the distinct steps of PSTVd cell-to-cell transport in leaves. Cell types are indicated on the right in B with different colors.

SECONDARY STRUCTURE OF PRE-miRNA AND UNTRANSLATED REGION IS ONE OF IMPORTANT FACTORS MEDIATING RNA AND PROTEIN MOBILITY

Secondary structure, including the hairpin structure of pre-miRNA and UTR region, was reported trigger RNA mobility. Most of the identified pre-miRNAs were presumed to have stem-loop structures (Lee et al., 2003). Previous studies mainly focused on the function and transportation of mature miRNA in the phloem sap (Kehr and Kragler, 2018), yet the role of precursors of miRNA through the phloem long-distance transportation remain unknown. A recent study reveals a plant BAP-like proteins NtPBL on *N. benthamiana* and figures out its affinity to binding both tRNAs/TLS and pre-miRNA (Atabekova et al., 2017). Therefore, whether the function and the structure of pre-miRNA were similar to TLS and further could mediate the RNA transportation were studied. Pre-miR390b which contain a hairpin flanking with both 200-nucleotides regions of up and downstream was tested. Result indicated that pre-miR390b mediated the transportation of a movement protein deficient *Potato virus X* (PVX) move from lower leaves to upper leaves (Lezzhov et al., 2019). Furthermore, by analyzing the available reads of *C. maxima* phloem sap RNA data, not only sequence of pre-miRNA390b, but also pre-miR390a, pre-miR159 and pre-miR168 were found in the same database. Additionally, secondary structure predictions indicated pre-miRNA390 (a and b) containing imperfect hairpin structures (Lezzhov et al., 2019). Interestingly, Zhang et al. (2016) indicated that the core structures for tRNA-related-sequence trigger mRNA transportation was the hairpin motif which similar to the structure of miRNA precursor. Here, we summarize the recent research evidence and present a possible mRNA transportation scheme with secondary hairpin structures (Figure 2).

The UTR including both 3' UTR and 5' UTR, plays an essential role in post-transcriptional regulation of gene expression and controlling the translation initiation and efficiency of mRNA.

Researches revealed that the sequence and structure of the UTR in the mRNA affect their expression, stability, and localization (Andreassi and Riccio, 2009; Hannapel, 2010; Hogg and Goff, 2010; Leppek et al., 2018) and the phloem-mobility of transcripts (Malka et al., 2017). Previous research has identified that the sequence of mRNA 5'UTR participates in the transport of mRNA from the nucleus to the cytoplasm (Banerjee et al., 2009; Cho et al., 2015). *StBEL5* is a member of the potato (*S. tuberosum*) KNOTTED transcription factor family, which accumulates with the increases of the short-day light cycle. It has proved that 5'UTR of *StBEL5* promote mRNA transport from leaves to the stolon tip of the underground through the phloem, and increase the yield of tubers (Banerjee et al., 2009). Additionally, studies have found that the long-distance transportation of certain mRNAs also related to the sequence of 3'UTR. The 3'UTR of the *GAI* gene can promote its transportability. Deleting or mutating the sequence of 3'UTR abolished mRNA transportation of *GAI* in *A. thaliana* (Huang and Yu, 2009).

Despite the sequence of UTR region, the secondary structures, including stem-loop, hairpin, pseudo-knots or the tertiary structures such as tRNA or TLS in the UTR region also crucial for the mRNA phloem transportation. These structures may have high affinity to RNA-binding proteins which assist mRNA phloem transportation (Tolstyko et al., 2020). Studies have reported that disrupt the secondary structure of *GAI* 3'UTR reduce its transcript transportation ability through the phloem (Huang and Yu, 2009). Furthermore, TLSs were usually embedded in the UTR region to promote mRNAs' transportation (Matsuda et al., 2004; Lezzhov et al., 2019).

tRNA-RELATED SEQUENCE AND tRNA-LIKE STRUCTURE

Studies have found that phloem sap contains a large number of non-coding RNAs (Yoo et al., 2004; Buhtz et al., 2008; Pant et al., 2008; Zhang et al., 2009), including tRNAs and tRNA fragments (Zhang et al., 2009). A research was done on the pumpkin

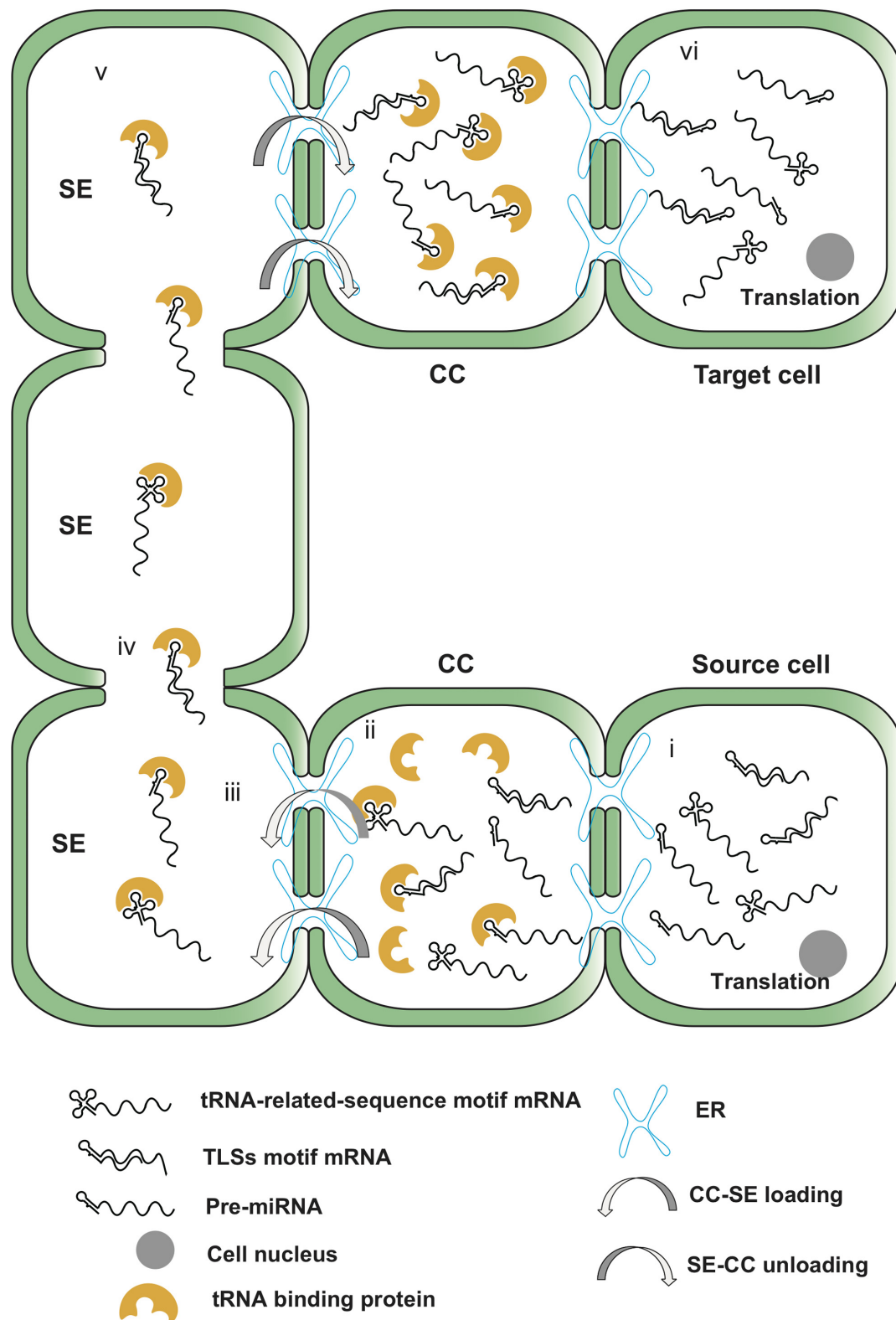


FIGURE 2 | The transportation pathway of mRNA triggered by TLSs and secondary hairpin structure of pre-miRNA. (i) Phloem-mobile mRNAs are transcribed in the source cell and transported from cell nucleus; (ii) The target phloem-mobile mRNA with secondary hairpin structure enters the CC and recognized by binding protein; (iii) The protein carries the target mRNA and pre-miRNA moves through the plasmodesmata assisted by PD receptor protein; (iv) mRNA and pre-miRNA molecules transport in SE; (v) mRNA and pre-miRNA unload from SE to CC; (vi) Finally, target RNAs enter the target cell and translate to functional proteins.

(*C. maxima*) showed that phloem-specific tRNA fragments could interfere with ribosomal activity, further, block the transcription of mRNA during its transportation in the phloem (Zhang et al., 2009; Ren et al., 2019; **Figure 2**).

The TLS was first discovered in the viral genome and existed in the 3' terminal of certain plant viruses' genomes with positive-strand RNA (Fechter et al., 2001; Dreher, 2009). Although viral TLSs sequence is wide apart from the tRNA sequence, they can form an L-shape structure similar to tRNAs (Dreher, 2010). Virus TLS has a function similar to the mRNA 3' poly-A tail, and it contributes to the stability and translation activity of viral RNA and essential for its replication and infectivity (Fechter et al., 2001; Qi et al., 2004; Dreher, 2010; Takeda et al., 2011). In addition, the TLS also can be recognized by tRNA specific proteins, which is very likely related to their influence on viral transport. The 3' UTR region (where the TLS located) of Brome mosaic virus RNA3, Tobacco mosaic virus and Turnip yellow mosaic virus were fused to an unmovable PVX with GFP and agroinfiltrated to *N. benthamiana*. Green fluorescence examination indicated the mobility of these three TLS through the phloem (Lezzhov et al., 2019). These data suggest that viral TLS could mimic real tRNA and be recognized by specific proteins and trigger phloem transportation.

Endogenous tRNA-related sequences were found to be enriched in on the 3' UTR region of mobile mRNA database (Thieme et al., 2015; Yang et al., 2015; Guan et al., 2017). Evidence showed that fusion the unmovable *GUS* with tRNA^{Gly} or tRNA^{Met} resulted in the transportation of *GUS* mRNA. Furthermore, deleted the TLS motif of mRNA-tRNA transcript dissipate the mobility of the mRNA (Zhang et al., 2016). Used the mutant of phosphatidylcholine kinase CK1 gene and tRNA^{Gly} core sequence to form a fusion transcript, CK1: tRNA^{Gly} transcript can be transported bidirectionally between the rootstock and scion, while the CK1 mutant transcript which lacking the tRNA^{Gly} sequence cannot be transported. Moreover, the presence of the full-length tRNA^{Met} sequence would trigger the transportation of DNDMC1 poly (A) transcripts into meiosis tissues (Thieme et al., 2015).

However, not all tRNA-relate sequence or TLS could trigger the mobility of mRNA. In *N. benthamiana*/Tomato heterograft, only 11 out of 174 high expressed mRNA, which contains TLS motif could transport (Xia et al., 2018). Additionally, unlike tRNA^{Gly} and tRNA^{Met} which were found in the phloem sap could mediate *GUS* transportation, tRNA^{Ile} was not found in the phloem sap and cannot mediate *GUS* transportation (Zhang et al., 2016). The different mobility of tRNA-relate sequences also results in line with the finding on pumpkin (Zhang et al., 2009). Subsequently, it was found through sequencing that tRNA^{Gly} has specific m5C modifications, while tRNA^{Ile} does not have base methylation modifications (Burgess et al., 2015). Further research demonstrated that cytosine's methylation modification plays a role in regulating the long-distance movement of transcripts to specific tissues (Yang et al., 2019). Base modifications may change the transcript's stability and function on the combination of cytosine, structure, and intracellular factors (Karijovich et al., 2010; Zhang et al., 2012).

RNA METHYLATION IS INVOLVED IN PROMOTING MRNA TRANSPORT IN PLANTS

More and more studies have shown that RNA methylation is deeply involved in plant developmental regulation and abiotic stress response (Merret et al., 2015; Růžička et al., 2017; Wei et al., 2018; Hu et al., 2019). Recent studies have provided much evidence for the involvement of methylation in plant RNA transport. Compared with other types of methylation, m5C is more likely to participate in systemic mRNA transfer in the phloem of plants. The study reported that the 5mC base-modified mRNA in *Arabidopsis* can move to different organs over graft union. *TUMOR CONTROLLED TRANSLATION PROTEIN 1* (*TCTP1*) and *HEAT SHOCK* homologous protein 70.1 (*HSC70.1*) were insufficiently methylated due to 5mC mutation with reduced transport, 5mC modified *TCTP1* can transport to root cells and increase root growth (Yang et al., 2019). The possible roles of methylation in the systematic transport of plant mRNA are listed as follows:

Methylation and Stability

As we known that mRNA is never naked in the phloem mediated transportation. Except PTB proteins, m5C can also increase the stability of mRNA, such as (tRNA METHYLTRANSFERASE 4 (TRM4)-mediated m5C methylation in *Arabidopsis* delayed the degradation of *SHY2* and *IAA16* transcripts, thereby affecting root development (Cui et al., 2017). Therefore, the ability of methylation to increase mRNA stability may be beneficial for long-distance RNA transport.

Methylation and Sequence Specificity

The mobility of numerous endogenous mRNAs form CC to SE is motif-dependent (Ham et al., 2009; Li et al., 2011; Mahajan et al., 2012; Cho et al., 2015; Duan et al., 2016; Ham and Lucas, 2017). MEME-ChIP recognized that certain consistent motifs exist near the mRNA methylation sites were more than 50% in the methylated mRNA population (Cui et al., 2017). In fact, methylation was necessary for the transport of *Arabidopsis TCTP1* mRNA through SE. In this process, specific motifs in the mRNA were recognized by methyltransferase to achieve methylation of *TCTP1* mRNA (Yang et al., 2019). A segment of *Arabidopsis TCTP1* mRNA selected for the affinity test of pumpkin phloem exudate contains methylation sites and surrounding areas, which proved that this segment of *Arabidopsis TCTP1* mRNA was bound to CmPS1 *in vitro* (Tolstyko et al., 2020).

Methylation and TLS

A methyltransferase TRM4, selectively methylated specific nucleotides on certain tRNAs in yeast, can significantly enhance the stability of tRNAs and avoid tRNA degradation caused by environmental changes (Müller et al., 2019; Xue et al., 2020). As mentioned before, TLS was found be significantly enriched in mobile mRNA datasets of *Arabidopsis*. For example, tRNA^{Gly} or tRNA^{Met} methylated with specific 5mC in *Arabidopsis*,

can trigger the mobility of non-mobile *GUS* mRNA. While tRNA^{lle} with little or no methylated cytosines was non-mobile (Zhang et al., 2016). Such base modifications have the potential to change the structure of the transcript and interact with cytokines, and thus may mediate long-distance movement of the transcript. 5mC methylation increases the stability of TLS and mRNA in *Arabidopsis*, thereby increasing the stability and ability of RNA transport.

Methylation and Translation Activity

To achieve the function of translated mobile mRNA in destination cells, it is necessary to ensure that the mRNA will not be translated in advance during the transport process. Analysis of the results of m5C-RIP-Seq showed that the m5C peak was mostly located in the CDS region, and it was significantly enriched in the regions after start and before stop codon (Cui et al., 2017). The high content of m5C in *Arabidopsis* mRNA was more related to ribosomal subunits (40S and 60S) and monosome (80S) rather than polysome, proving the abundance of m5C was negatively correlated with mRNA translation activity (Cui et al., 2017). This may help to avoid mRNA translation during transport.

DISCUSSION

With more evidence revealed, the passive flow assumption that RNA transport is only about RNA molecular abundance and half-life has been denied (Liu and Chen, 2018). On the one hand, the transport of different RNAs has caused specific phenotype changes (Banerjee et al., 2009; Huang and Yu, 2009). On the other hand, the potential regulatory substances or mechanisms for long-distance RNA transport are constantly being discovered. The mechanism by which RNA binding protein forms an RNA-protein complex to mediate transport is proposed in the research of small RNA (Yoo et al., 2004; Ham et al., 2014) and viroid RNA (Pallas and Gómez, 2013). Without denying this assumption, it has been discovered that RNA motifs and modifications play an important role in transportation. More and more evidences show that at least part of RNA is tightly regulated and plays an important role in the long-distance transportation of plants. Therefore, the improvement of protein-mediated RNA long-distance transport mechanism and the discovery of new mechanisms are attracting researchers in this field.

In our review, first of all, the variation of RNA motifs affect the transport of viroid RNA in plants. The single base mutation may involve the change of 3D conformation of base, thus affecting the recognition and interaction with some protein factors in plants, thus affecting the transport of RNA. The same is true for small sequences, which have a stable conformation and thus bind to specific proteins. The secondary structure of pre-miRNA, such as stem loop, hairpin, pseudo-knots or the tertiary structures such as tRNA or TLS in the UTR region also critical for the mRNA transport, also plays an important role in triggering RNA transport. The common type of methylation, m5C, can mediate the migration of mRNA to different organs through graft union. It is understood that the methylation of m5C increases the stability of mRNA, improves its transport capacity, and

improves the interaction between mRNA and some transport factors. The reversible regulation of methylation gives us reason to specify that methyltransferase and demethylase mediate long-distance transport of mRNA and the switch of translation status. Studies have found that TLS is also one of the reasons for triggering mRNA transport, the specific tRNA fragment found in phloem supports this view. The special structure of TLS may also participate in transport through binding with specific protein factors in phloem sap.

Like DNA, the methylation of RNA also requires writer, reader, eraser, and related enzymes have been found in some plants (Xue et al., 2020). In the process of long-distance transportation to specific target tissues to function, mRNA molecules undergo at least one state switch: transportation and translation. In the process of transportation, RNA molecules are usually required to have the ability to be recognized by loading and transportation factors, high stability, and low translation ability. However, after mRNA undergoes SE-mediated unloading and finally reaches the target cell, the translation ability needs to be restored. The reversible regulation of methylation gives us reason to speculate that methyltransferase and demethylase mediate long-distance transport of mRNA and the switch of translation status. Previous studies have proved that methylation is necessary for the transport of certain mRNAs (Yang et al., 2017, 2019). Plants are accompanied by changes in the level of RNA methylation at different stages of growth and development and in response to stress. In addition, research on plant viruses found that after infecting cells, viral RNA can be transported through the phloem to the far end of the plant under the envelope of protein (Navarro et al., 2019). The TLS structure is also necessary for long-distance transport of certain mRNAs (Zhang et al., 2016), and the pumpkin phloem protein CmPS1 showed an affinity for the RNA stem-loop structure in *in vitro* experiments (Tolstyko et al., 2020). Combining these research results, here we propose a prospective model of methylation-mediated long-distance transport of plant mRNA via SE, hoping to inspire follow-up research.

The systematic transport of mRNA during growth and development of plants or caused by environmental changes is likely to be involved in various stages from the production of source cells to translation in target cells: (i) Methylation mediates mRNA nucleocytoplasmic shuttling and marking as transport mRNA. (ii) The motifs surrounding methylation are recognized to form a protein-RNA complex. (iii) Mediate size-exclusion limit of PD to achieve CC-SE transport. (iv) Stabilize RNA molecules transport in SE and reduce translation ability. (v) Mediates the unloading of CC-SE. (vi) Reduce methylation level and restore translation ability (Figure 3). Nevertheless, the evidence to support this model is quite limited. Methylation regulation may not be so extensively involved in the long-distance transport process of plant mRNA. The relationship between methylation and TLS is considered weak (Yang et al., 2019). There is insufficient evidence that methylation sites mediate the formation of protein-RNA complexes. Some studies believe that methylation abundance is not correlated or negatively correlated with mRNA stability (David et al., 2017; Yang et al., 2019). There is currently no evidence that methylation or demethylation

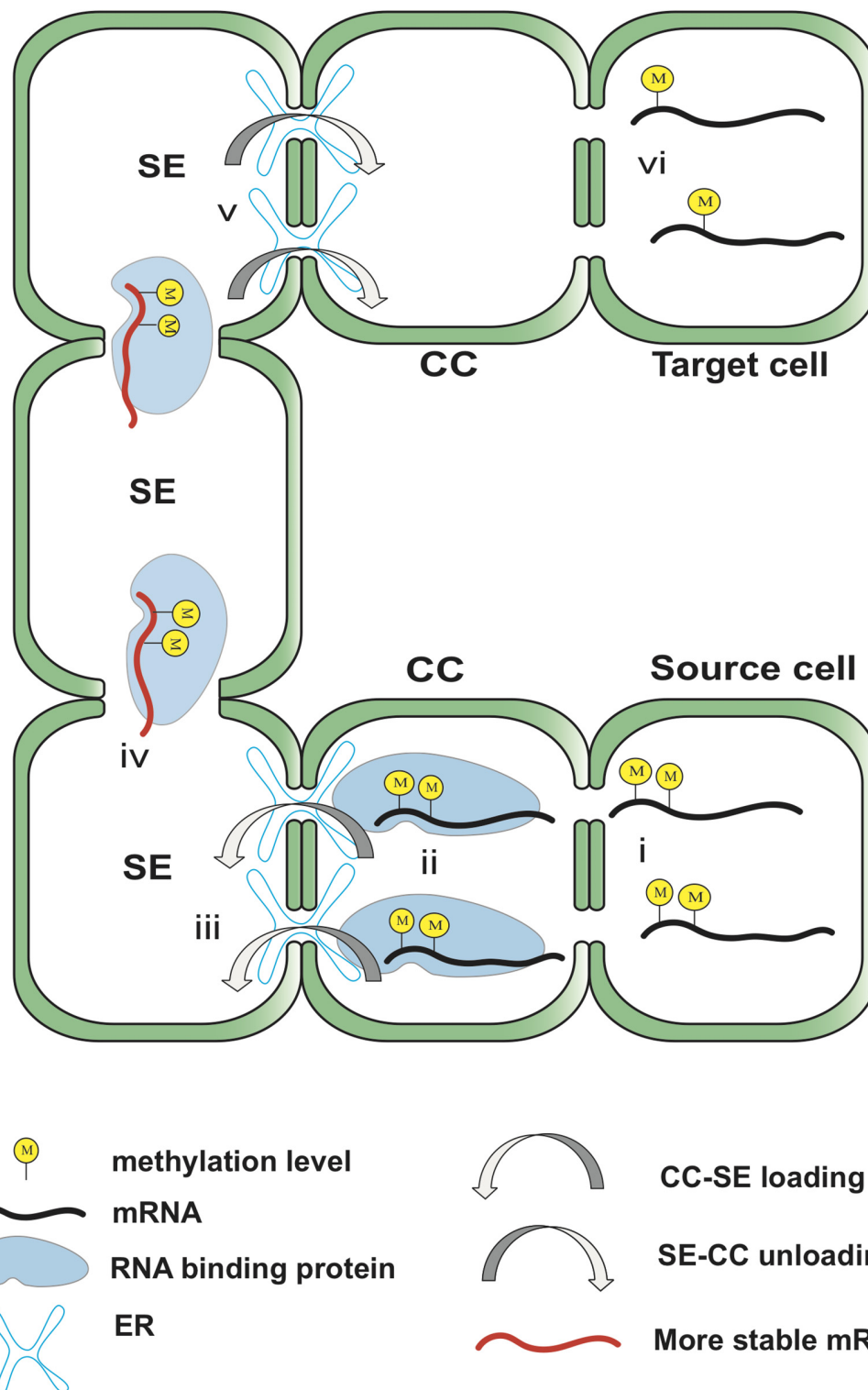


FIGURE 3 | Hypothetical model of the potential involvement of methylation in mRNA transport. The transcribed mRNA is methylated by methyltransferase, which in turn mediates the transport of the mRNA from the nucleus to the cytoplasm. Methylation may reduce the translation ability of the mRNA until it reaches the target cell. RBP recognizes the motif near the methylation site to form an RNA-protein complex. Methylation is involved in regulating the size-exclusion limit of PD to carry out STS loading of mRNA and stabilize mRNA and its TLS in SE. Methylation is involved in identifying unloading locations and completing sieve element unloading of mRNA near the target cell. After reaching the target cell, mRNA reduces the methylation level under the action of demethylase and restores translation ability. The locations of yellow circles in the figure do not indicate the actual methylation sites, but the number of them indicates the level of methylation.

causes changes in mRNA translation ability before and after phloem unloading.

There are still many questions to be answered about the mechanism of RNA transportation. For example, will all RNA be transport? What is the relationship between RNA motifs and transport between different cells during regulated transportation? How these RNA motifs participate in guidance, and whether they have corresponding changes in this process, requires further research? Although we have drawn some insights from the research, there is still a long way to go for the comprehensive answer.

CONCLUSION

In this paper, the possible mechanism of long-distance RNA transport was discussed from four aspects: single nucleotide mutation, secondary structure of pre-miRNA, tRNA-related sequence and TLS, and RNA methylation. These studies lay a theoretical foundation to uncover the mechanism of RNA transport. The importance of this review is not only to explore

the mechanism of RNA signaling, but also to provide an insight for plant growth, development and stress resistance regulation in RNA levels.

AUTHOR CONTRIBUTIONS

LG and WZ revised the final manuscript. TW and ZL wrote viroid and polypyrimidine part. XLi and ML wrote methylation part. XZ, XLu, and WL wrote pre-miRNA and TLS part. QW and SG revised the whole draft. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the National Key Research and Development Program of China (2018YFD1000800 and 2019YFD1000300) funded for WZ and the National Natural Science Foundation of China (31872158 for WZ).

REFERENCES

- Andreassi, C., and Riccio, A. (2009). To localize or not to localize: mRNA fate is in 3'UTR ends. *Trends Cell Biol.* 19, 465–474. doi: 10.1016/j.tcb.2009.06.001
- Atabekova, A. K., Pankratenko, A. V., Makarova, S. S., Lazareva, E. A., Owens, R. A., Solov'yev, A. G., et al. (2017). Phylogenetic and functional analyses of a plant protein related to human B-cell receptor-associated proteins. *Biochimie* 132, 28–37. doi: 10.1016/j.biochi.2016.10.009
- Banerjee, A. K., Lin, T., and Hannapel, D. J. (2009). Untranslated regions of a mobile transcript mediate RNA metabolism. *Plant Physiol.* 151, 1831–1843. doi: 10.1104/pp.109.144428
- Buhtz, A., Springer, F., Chappell, L., Baulcombe, D. C., and Kehr, J. (2008). Identification and characterization of small RNAs from the phloem of *Brassica napus*. *Plant J.* 53, 739–749. doi: 10.1111/j.1365-3113X.2007.03368.x
- Burgess, A. L., David, R., and Searle, I. R. (2015). Conservation of tRNA and rRNA 5-methylcytosine in the kingdom Plantae. *BMC Plant Biol.* 15:199. doi: 10.1186/s12870-015-0580-8
- Cho, S. K., Sharma, P., Butler, N. M., Kang, I.-H., Shah, S., Rao, A. G., et al. (2015). Polypyrimidine tract-binding proteins of potato mediate tuberization through an interaction with StBEL5 RNA. *J. Exp. Bot.* 66, 6835–6847. doi: 10.1093/jxb/erv389
- Cui, X., Liang, Z., Shen, L., Zhang, Q., Bao, S., Geng, Y., et al. (2017). 5-Methylcytosine RNA methylation in *Arabidopsis thaliana*. *Mol. Plant* 10, 1387–1399. doi: 10.1016/j.molp.2017.09.013
- David, R., Burgess, A., Parker, B., Li, J., Pulsford, K., Sibbritt, T., et al. (2017). Transcriptome-Wide Mapping of RNA 5-methylcytosine in Arabidopsis mRNAs and Non-coding RNAs. *Plant Cell* 29, 445–460. doi: 10.1105/tpc.16.00751
- Ding, B. (2009). The biology of viroid-host interactions. *Annu. Rev. Phytopathol.* 47, 105–131. doi: 10.1146/annurev-phyto-080508-081927
- Dreher, T. W. (2009). Role of tRNA-like structures in controlling plant virus replication. *Virus Res.* 139, 217–229. doi: 10.1016/j.virusres.2008.06.010
- Dreher, T. W. (2010). Viral tRNAs and tRNA-like structures: viral tRNAs and TLSs. *Wiley Interdiscip. Rev.* 1, 402–414. doi: 10.1002/wrna.42
- Duan, X., Zhang, W., Huang, J., Hao, L., Wang, S., Wang, A., et al. (2016). *PbWoxT1* mRNA from pear (*Pyrus betulaefolia*) undergoes long-distance transport assisted by a polypyrimidine tract binding protein. *New Phytol.* 210, 511–524. doi: 10.1111/nph.13793
- Fechter, P., Rudinger-Thirion, J., Florentz, C., and Giegé, R. (2001). Novel features in the tRNA-like world of plant viral RNAs. *Cell. Mol. Life Sci.* 58, 1547–1561. doi: 10.1007/PL00000795
- Flores, R., Hernández, C., Martínez, de Alba, A. E., Daròs, J.-A., and Di Serio, F. (2005). Viroids and viroid-host interactions. *Annu. Rev. Phytopathol.* 43, 117–139. doi: 10.1146/annurev.phyto.43.040204.140243
- Gómez, G., and Pallás, V. (2004). A long-distance translocatable phloem protein from cucumber forms a ribonucleoprotein complex *in vivo* with hop stunt viroid RNA. *J. Virol.* 78, 10104–10110. doi: 10.1128/JVI.78.18.10104-10110.2004
- Gómez, G., Torres, H., and Pallás, V. (2005). Identification of translocatable RNA-binding phloem proteins from melon, potential components of the long-distance RNA transport system. *Plant J.* 41, 107–116. doi: 10.1111/j.1365-3113X.2004.02278.x
- Góra-Sochacka, A. (2004). Viroids: unusual small pathogenic RNAs. *Acta Biochim. Pol.* 51, 587–607.
- Gozmanova, M., Denti, M. A., Minkov, I. N., Tsagris, M., and Tabler, M. (2003). Characterization of the RNA motif responsible for the specific interaction of potato spindle tuber viroid (PSTVD) and the tomato protein Virp1. *Nucleic Acids Res.* 31, 5534–5543.
- Guan, D., Yan, B., Thieme, C., Hua, J., Zhu, H., Boheler, K. R., et al. (2017). PlaMoM: a comprehensive database compiles plant mobile macromolecules. *Nucleic Acids Res.* 45, 1021–1028. doi: 10.1093/nar/gkw988
- Ham, B. K., Brandom, J. L., Xoconostle-Cázares, B., Ringgold, V., Lough, T. J., and Lucas, W. J. (2009). A polypyrimidine tract binding protein, pumpkin RBP50, forms the basis of a phloem-mobile ribonucleoprotein complex. *Plant Cell* 21, 197–215. doi: 10.1105/tpc.108.061317
- Ham, B. K., Li, G., Jia, W., Leary, J. A., and Lucas, W. J. (2014). Systemic delivery of siRNA in pumpkin by a plant PHLOEM SMALL RNA-BINDING PROTEIN 1-ribonucleoprotein complex. *Plant J.* 80, 683–694.
- Ham, B. K., and Lucas, W. J. (2017). Phloem-mobile RNAs as systemic signaling agents. *Annu. Rev. Plant Biol.* 68, 173–195. doi: 10.1146/annurev-arplant-042916-041139
- Hannapel, D. J. (2010). A model system of development regulated by the long-distance transport of mRNA. *J. Integr. Plant Biol.* 52, 40–52. doi: 10.1111/j.1744-7909.2010.00911.x
- Hogg, J. R., and Goff, S. P. (2010). Upf1 Senses 3'UTR length to potentiate mRNA Decay. *Cell* 143, 379–389. doi: 10.1016/j.cell.2010.10.005

- Hu, J., Manduzio, S., and Kang, H. (2019). Epitranscriptomic RNA methylation in plant development and Abiotic stress responses. *Front. Plant Sci.* 10:500. doi: 10.3389/fpls.2019.00500
- Huang, N.-C., and Yu, T.-S. (2009). The sequences of Arabidopsis GA-INSENSITIVE RNA constitute the motifs that are necessary and sufficient for RNA long-distance trafficking. *Plant J.* 59, 921–929. doi: 10.1111/j.1365-313X.2009.03918.x
- Karijolich, J., Kantartzis, A., and Yu, Y.-T. (2010). RNA modifications: a mechanism that modulates gene expression. *Methods Mol. Biol.* 629, 1–19. doi: 10.1007/978-1-60761-657-3_1
- Kehr, J., and Kragler, F. (2018). Long distance RNA movement. *New Phytol.* 218, 29–40. doi: 10.1111/nph.15025
- Lee, Y., Ahn, C., Han, J., Choi, H., Kim, J., Yim, J., et al. (2003). The nuclear RNase III Drosha initiates microRNA processing. *Nature* 425, 415–419. doi: 10.1038/nature01957
- Leontis, N. B., Lescoute, A., and Westhof, E. (2006). The building blocks and motifs of RNA architecture. *Curr. Opin. Struct. Biol.* 16, 279–287. doi: 10.1016/j.sbi.2006.05.009
- Leontis, N. B., Stombaugh, J., and Westhof, E. (2002). Motif prediction in ribosomal RNAs Lessons and prospects for automated motif prediction in homologous RNA molecules. *Biochimie* 84, 961–973. doi: 10.1016/S0300-9084(02)01463-3
- Leppeck, K., Das, R., and Barna, M. (2018). Functional 5' UTR mRNA structures in eukaryotic translation regulation and how to find them. *Nat. Rev. Mol. Cell Biol.* 19, 158–174. doi: 10.1038/nrm.2017.103
- Lezzhov, A. A., Atabekova, A. K., Tolstyko, E. A., Lazareva, E. A., and Solovyev, A. G. (2019). RNA phloem transport mediated by pre-miRNA and viral tRNA-like structures. *Plant Sci.* 284, 99–107. doi: 10.1016/j.plantsci.2019.04.005
- Li, P., Ham, B.-K., and Lucas, W. J. (2011). CmRBP50 protein phosphorylation is essential for assembly of a stable phloem-mobile high-affinity ribonucleoprotein complex. *J. Biol. Chem.* 286, 23142–23149. doi: 10.1074/jbc.M111.244129
- Liu, L., and Chen, X. M. (2018). Intercellular and systemic trafficking of RNAs in plants. *Nat. Plants* 4, 869–878.
- Liu, W., Xiang, C., Li, X., Wang, T., Lu, X., Liu, Z., et al. (2020). Identification of long-distance transmissible mRNA between scion and rootstock in cucurbit seedling heterografts. *Int. J. Mol. Sci.* 21:5253.
- Lucas, W. J., Groover, A., Lichtenberger, R., Furuta, K., Yadav, S. R., Helariutta, Y. K., et al. (2013). The plant vascular system: evolution, development and functions. *J. Integr. Plant Biol.* 55, 294–388.
- Mahajan, A., Bhogale, S., Kang, I. H., Hannapel, D. J., and Banerjee, A. K. (2012). The mRNA of a Knotted1-like transcription factor of potato is phloem mobile. *Plant Mol. Biol.* 79, 595–608. doi: 10.1007/s11103-012-9931-0
- Maizel, A., Markmann, K., Timmermans, M., and Wachter, A. (2020). To move or not to move: roles and specificity of plant RNA mobility. *Curr. Opin. Plant Biol.* 57, 52–60.
- Malka, Y., Steiman-Shimony, A., Rosenthal, E., Argaman, L., Cohen-Daniel, L., Arbib, E., et al. (2017). Post-transcriptional 3'-UTR cleavage of mRNA transcripts generates thousands of stable uncapped autonomous RNA fragments. *Nat. Commun.* 8:2029. doi: 10.1038/s41467-017-02099-7
- Maniataki, E., Martinez de Alba, A. E., Sägesser, R., Tabler, M., and Tsagris, M. (2003). Viroid RNA systemic spread may depend on the interaction of a 71-nucleotide bulged hairpin with the host protein VirP1. *RNA* 9, 346–354.
- Matsuda, D., Yoshinari, S., and Dreher, T. W. (2004). EEF1A binding to aminoacylated viral RNA represses minus strand synthesis by TYMV RNA-dependent RNA polymerase. *Virology* 321, 47–56. doi: 10.1016/j.virol.2003.10.028
- Merret, R., Nagarajan, V. K., Carpentier, M.-C., Park, S., Favory, J.-J., Descombin, J., et al. (2015). Heat-induced ribosome pausing triggers mRNA co-translational decay in *Arabidopsis thaliana*. *Nucleic Acids Res.* 43, 4121–4132. doi: 10.1093/nar/gkv234
- Morris, R. J. (2018). On the selectivity, specificity and signalling potential of the long-distance movement of messenger RNA. *Curr. Opin. Plant Biol.* 43, 1–7.
- Müller, M., Samel-Pommerencke, A., Legrand, C., Tuorto, F., Lyko, F., and Ehrenhofer-Murray, A. E. (2019). Division of labour: tRNA methylation by the NSun2 tRNA methyltransferases Trm4a and Trm4b in fission yeast. *RNA Biol.* 16, 249–256. doi: 10.1080/15476286.2019.1568819
- Navarro, J. A., Sanchez-Navarro, J. A., and Pallas, V. (2019). Key checkpoints in the movement of plant viruses through the host. *Adv. Virus Res.* 104, 1–64. doi: 10.1016/bs.aivir.2019.05.001
- Noller, H. F. (2005). RNA structure: reading the ribosome. *Science* 309, 1508–1514. doi: 10.1126/science.1111771
- Notaguchi, M., and Okamoto, S. (2015). Dynamics of long-distance signaling via plant vascular tissues. *Front. Plant Sci.* 6:161. doi: 10.3389/fpls.2015.00161
- Omid, A., Keilin, T., Glass, A., Leshkowitz, D., and Wolf, S. (2007). Characterization of phloem-sap transcription profile in melon plants. *J. Exp. Bot.* 58, 3645–3656.
- Owens, R. A. (2007). Potato spindle tuber viroid: The simplicity paradox resolved? *Mol. Plant Pathol.* 8, 549–560. doi: 10.1111/j.1364-3703.2007.00418.x
- Pallas, V., and Gómez, G. (2013). Phloem RNA-binding proteins as potential components of the long-distance RNA transport system. *Front. Plant Sci.* 4:130. doi: 10.3389/fpls.2013.00130
- Pant, B. D., Buhtz, A., Kehr, J., and Scheible, W.-R. (2008). MicroRNA399 is a long-distance signal for the regulation of plant phosphate homeostasis. *Plant J.* 53, 731–738. doi: 10.1111/j.1365-313X.2007.03363.x
- Qi, Y., Pélissier, T., Itaya, A., Hunt, E., Wassenegger, M., and Ding, B. (2004). Direct role of a viroid RNA motif in mediating directional RNA trafficking across a specific cellular boundary. *Plant Cell* 16, 1741–1752. doi: 10.1105/tpc.021980
- Ren, B., Wang, X., Duan, J., and Ma, J. (2019). Rhizobial tRNA-derived small RNAs are signal molecules regulating plant nodulation. *Science* 365, 919–922. doi: 10.1126/science.aav8907
- Rosin, F. M., Hart, J. K., Horner, H. T., Davies, P. J., and Hannapel, D. J. (2003). Overexpression of a knotted-like homeobox gene of potato alters vegetative development by decreasing gibberellin accumulation. *Plant Physiol.* 132, 106–117. doi: 10.1104/pp.102.015560
- Růžicka, K., Zhang, M., Campilho, A., Bodi, Z., Kashif, M., Saleh, M., et al. (2017). Identification of factors required for m6A mRNA methylation in Arabidopsis reveals a role for the conserved E3 ubiquitin ligase HAKAI. *New Phytol.* 215, 157–172. doi: 10.1111/nph.14586
- Skrajina, A., Yang, X.-C., Bucholc, K., Zhang, J., Hall, T. M. T., Dadlez, M., et al. (2017). U7 snRNP is recruited to histone pre-mRNA in a FLASH-dependent manner by two separate regions of the stem-loop binding protein. *RNA* 23, 938–951. doi: 10.1261/rna.060806.117
- Steger, G., and Perreault, J.-P. (2016). Structure and associated biological functions of Viroids. *Adv. Virus Res.* 94, 141–172. doi: 10.1016/bs.aivir.2015.11.002
- Takeda, R., and Ding, B. (2009). Viroid intercellular trafficking: RNA motifs, cellular factors and broad impacts. *Viruses* 1, 210–221. doi: 10.3390/v1020210
- Takeda, R., Petrov, A. I., Leontis, N. B., and Ding, B. (2011). A three-dimensional RNA motif in Potato spindle tuber viroid mediates trafficking from palisade mesophyll to spongy mesophyll in *Nicotiana benthamiana*. *Plant Cell* 23, 258–272. doi: 10.1105/tpc.110.081414
- Thieme, C. J., Rojas-Triana, M., Stecyk, E., Schudoma, C., Zhang, W., Yang, L., et al. (2015). Endogenous Arabidopsis messenger RNAs transported to distant tissues. *Nat. Plants* 1:15025. doi: 10.1038/nplants.2015.25
- Tolstyko, E. A., Lezzhov, A. A., Morozov, S. Y., and Solovyev, A. G. (2020). Phloem transport of structured RNAs: a widening repertoire of trafficking signals and protein factors. *Plant Sci.* 299:110602. doi: 10.1016/j.plantsci.2020.110602
- Wang, S., Wang, S., Zhang, W., Zhang, Q., Hao, L., Zhang, Y., et al. (2019). PbTTG1 forms a ribonucleoprotein complex with polypyrimidine tract-binding protein PbPTB3 to facilitate the long-distance trafficking of PbWoxT1 mRNA. *Plant Sci.* 280, 424–432. doi: 10.1016/j.plantsci.2019.01.008
- Wang, Y., Qu, J., Ji, S., Wallace, A., Wu, J., Li, Y., et al. (2016). A land plant-specific transcription factor directly enhances transcription of a pathogenic noncoding RNA template by DNA-dependent RNA polymerase II. *Plant Cell* 28, 1094–1107. doi: 10.1105/tpc.16.00100
- Wei, L.-H., Song, P., Wang, Y., Lu, Z., Tang, Q., Yu, Q., et al. (2018). The m6A Reader ECT2 Controls trichome morphology by affecting mRNA stability in Arabidopsis. *Plant Cell* 30, 968–985. doi: 10.1105/tpc.17.00934
- Wu, J., Leontis, N. B., Zirbel, C. L., Bisaro, D. M., and Ding, B. (2019). A three-dimensional RNA motif mediates directional trafficking of Potato spindle tuber viroid from epidermal to palisade mesophyll cells in *Nicotiana benthamiana*. *PLoS Pathog.* 15:e1008147. doi: 10.1371/journal.ppat.1008147
- Xia, C., Zheng, Y., Huang, J., Zhou, X., Li, R., Zha, M., et al. (2018). Elucidation of the mechanisms of long-distance mRNA movement in a *Nicotiana benthamiana*/tomato heterograft system. *Plant Physiol.* 177, 745–758. doi: 10.1104/pp.17.01836

- Xue, C., Zhao, Y., and Li, L. (2020). Advances in RNA cytosine-5 methylation: detection, regulatory mechanisms, biological functions and links to cancer. *Biomark. Res.* 8:43. doi: 10.1186/s40364-020-00225-0
- Yang, L., Perrera, V., Saplaoura, E., Apelt, F., Bahin, M., Kramdi, A., et al. (2019). M5C methylation guides systemic transport of messenger RNA over graft junctions in plants. *Curr. Biol.* 29:2465. doi: 10.1016/j.cub.2019.06.042
- Yang, Y., Huang, M., Qi, L., Song, J., Li, Q., and Wang, R. (2017). Differential expression analysis of genes related to graft union healing in *Pyrus ussuriensis* Maxim by cDNA-AFLP. *Sci. Hortic.* 225, 700–706. doi: 10.1016/j.scienta.2017.07.028
- Yang, Y., Mao, L., Jittayasothorn, Y., Kang, Y., Jiao, C., Fei, Z., et al. (2015). Messenger RNA exchange between scions and rootstocks in grafted grapevines. *BMC Plant Biol.* 15:251. doi: 10.1186/s12870-015-0626-y
- Yoo, B.-C., Kragler, F., Varkonyi-Gasic, E., Haywood, V., Archer-Evans, S., Lee, Y. M., et al. (2004). A systemic small RNA signaling system in plants. *Plant Cell* 16, 1979–2000. doi: 10.1105/tpc.104.023614
- Zhang, S., Sun, L., and Kragler, F. (2009). The phloem-delivered RNA pool contains small non-coding RNAs and interferes with translation. *Plant Physiol.* 150, 378–387. doi: 10.1104/pp.108.134767
- Zhang, W., Thieme, C. J., Kollwig, G., Apelt, F., Yang, L., Winter, N., et al. (2016). tRNA-related sequences trigger systemic mRNA transport in plants. *Plant Cell* 28, 1237–1249. doi: 10.1105/tpc.15.01056
- Zhang, X., Liu, Z., Yi, J., Tang, H., Xing, J., Yu, M., et al. (2012). The tRNA methyltransferase NSun2 stabilizes p16 INK4 mRNA by methylating the 3'-untranslated region of p16. *Nat. Commun.* 3:712. doi: 10.1038/ncomms1692
- Zhong, X., Archual, A. J., Amin, A. A., and Ding, B. (2008). A genomic map of viroid RNA motifs critical for replication and systemic trafficking. *Plant Cell* 20, 35–47. doi: 10.1105/tpc.107.056606
- Zhong, X., Tao, X., Stombaugh, J., Leontis, N., and Ding, B. (2007). Tertiary structure and function of an RNA motif required for plant vascular entry to initiate systemic trafficking. *EMBO J.* 26, 3836–3846. doi: 10.1038/sj.emboj.7601812

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 Wang, Li, Zhang, Wang, Liu, Lu, Gao, Liu, Liu, Gao and Zhang. 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.



m⁶A Modification Mediates Mucosal Immune Microenvironment and Therapeutic Response in Inflammatory Bowel Disease

Yongyu Chen*, Jing Lei and Song He*

Department of Gastroenterology, The Second Affiliated Hospital of Chongqing Medical University, Chongqing, China

OPEN ACCESS

Edited by:

Jia Meng,
Xi'an Jiaotong-Liverpool University,
China

Reviewed by:

Hu Zhang,
Sichuan University, China
Shukui Wang,
Nanjing Medical University, China

*Correspondence:

Yongyu Chen
617585218@qq.com
Song He
hedoctor65@cqmu.edu.cn

Specialty section:

This article was submitted to
Epigenomics and Epigenetics,
a section of the journal
Frontiers in Cell and Developmental
Biology

Received: 07 April 2021

Accepted: 12 July 2021

Published: 06 August 2021

Citation:

Chen Y, Lei J and He S (2021)
m⁶A Modification Mediates Mucosal
Immune Microenvironment
and Therapeutic Response in
Inflammatory Bowel Disease.
Front. Cell Dev. Biol. 9:692160.
doi: 10.3389/fcell.2021.692160

Accumulating evidence links m⁶A modification with immune infiltration. However, the correlation and mechanism by which m⁶A modification promotes intestinal immune infiltration in inflammatory bowel disease (IBD) is unknown. Here, genomic information from IBD tissues was integrated to evaluate disease-related m⁶A modification, and the correlation between the m⁶A modification pattern and the immune microenvironment in the intestinal mucosa was explored. Next, we identified hub genes from the key modules of the m⁶A cluster and analyzed the correlation among the hub genes, immune infiltration, and therapy. We found that IGF2BP1 and IGF2BP2 expression was decreased in Crohn's disease (CD) tissues and that IGF2BP2 was decreased in ulcerative colitis (UC) tissues compared with normal tissues ($P < 0.05$). m⁶A cluster2, containing higher expressions of IL15, IL16, and IL18, was enriched in M0 macrophage, M1 macrophage, native B cells, memory B cells, and m⁶A cluster1 with high expression of IL8 and was enriched in resting dendritic and plasma cells ($P < 0.05$). Furthermore, we reveal that expression of m⁶A phenotype-related hub genes (i.e., NUP37, SNRPG, H2AFZ) was increased with a high abundance of M1 macrophages, M0 macrophages, and naive B cells in IBD ($P < 0.01$). Immune checkpoint expression in the genecluster1 with higher expression of hub genes was increased. The anti-TNF therapeutic response of patients in genecluster1 was more significant, and the therapeutic effect of CD was better than that of UC. These findings indicate that m⁶A modification may affect immune infiltration and therapeutic response in IBD. Assessing the expression of m⁶A phenotype-related hub genes might guide the choice of IBD drugs and improve the prediction of therapeutic response to anti-TNF therapy.

Keywords: m⁶A modification, IBD, immune infiltration, immunotherapy, anti-TNF response

INTRODUCTION

Inflammatory bowel disease (IBD), including Crohn's disease (CD) and ulcerative colitis (UC), is a chronic intestinal inflammatory disease (Torres et al., 2017; Ungaro et al., 2017). The precise etiology of IBD remains unclear, but genetic predisposition, disruption of mucosal immune homeostasis, microbial dysbiosis, and environmental factors are involved (Halfvarson et al., 2017; Neurath, 2017). In particular, the immune system is implicated as a major contributor to IBD progression, predominantly through an imbalance between the anti-inflammatory responses of regulatory T cells and M2 macrophages versus the pro-inflammatory responses of M1 macrophages,

T helper (Th) 1, Th17, and neutrophils (Canavan et al., 2016; de Souza and Fiocchi, 2016; Zhou et al., 2018, 2019). Current treatments are aimed at relieving inflammation through the use of steroids, biological agents (primarily anti-TNF), and immunosuppressive drugs (Torres et al., 2017; Ungaro et al., 2017). However, approximately 30% of patients using anti-TNF agents do not respond to treatment (Belarif et al., 2019). Thus, the potential mechanism of the progression of IBD needs to be elucidated.

N⁶-methyladenosine (m⁶A) modification, as the most prevalent RNA epigenetic modification, occurs transcriptionally (Lan et al., 2019; Huang H. et al., 2020). m⁶A methylation is catalyzed by the MTase complex referred to as “writers,” including methyltransferase-like 3/14/16 (METTL3/14/16), RNA binding motif protein 15 (RBM15), KIAA1429, and Wilms tumor 1-associated protein (WTAP; Chen X. Y. et al., 2019), whereas its demethylation is mediated by “eraser” proteins (ALKBH5 and FTO) (Huang Y. et al., 2019; Jin et al., 2020). The m⁶A modification is interpreted by m⁶A “reader” proteins, including the insulin-like growth factor 2 mRNA-binding protein 1/2/3 (IGF2BP1/2/3), YTH-domain family 1/2/3 (YTHDF1/2/3), and HNRNPA2B1 (Huang H. et al., 2018; Jiang et al., 2021). Accumulating evidence suggests that m⁶A modification plays an important role in inflammation, various tumors, innate immunity, and immunotherapy through different m⁶A regulator-related modifications (Chen X. Y. et al., 2019; Wang J. et al., 2019). Most studies focus on various tumors, single regulator, or single immunity cell types, but the influence of m⁶A modification on disease progression has not been examined in IBD.

Recent studies find that m⁶A modification is associated with disease-related immune infiltration. For example, METTL3-mediated m⁶A methylation promotes the activation of dendritic cells (DCs) (Wang H. et al., 2019), and METTL3 leads to an imbalance between Treg cells and native T cells with Treg cells losing the ability to suppress immune and native T cells losing the ability to induce inflammation (Geula et al., 2015; Li et al., 2017). YTHDF1 binds to lysosomal proteases, increases their translational efficiency in DCs, and enhances the tumor infiltrating CD8 + T cell antitumor response (Han et al., 2019). A recent study finds that METTL14 deficiency in T cells promotes spontaneous colitis in mice (Lu et al., 2020). However, the mechanism connecting IBD-related m⁶A modification and immune cell infiltration has not yet been elucidated. Thus, a comprehensive understanding of the different immune infiltrating characteristics induced by m⁶A regulators will improve our understanding of immune regulation in IBD.

Herein, the genomic information of IBD samples was integrated to assess the m⁶A modification patterns. Next, we explored the correlation between the m⁶A modification pattern and immune infiltration, investigating the functions and related mechanism of m⁶A modification patterns in the immune microenvironment in IBD. Surprisingly, we reveal that m⁶A modification patterns are associated with the disease types and infiltration of multiple immune cells. According to differentially expressed genes in different m⁶A modification patterns, we reveal that m⁶A phenotype-related hub genes affect disease

characteristics and immune cell-infiltrating characteristics in IBD patients and demonstrate that m⁶A phenotype-related hub genes are connected with disease progression, immune cell infiltration, anti-TNF therapeutic response, and immunotherapy. Assessing the expression of m⁶A phenotype-related hub genes might guide the choice of IBD drugs and improve the prediction of therapeutic response to anti-TNF therapy.

MATERIALS AND METHODS

Inflammatory Bowel Disease Data Set Source and Preprocessing

Public data and full clinical annotation were downloaded from the Gene-Expression Omnibus (GEO) database.¹ GSE111889 (Lloyd-Price et al., 2019) (251 samples) was obtained in this study for further analysis (**Supplementary Table 1**). Data preprocessing was performed as follows: the downloaded data set was the expression profile obtained after data homogenization. First, the probes corresponding to the genes were identified, and no-load probes were removed. When multiple probes corresponded to the same gene, the median number was selected as the expression level of the gene. Four samples with zero expression of all the genes were filtered out, and the final expression matrix of 247 samples was used for subsequent analysis. Then, based on gene expression spectrum data and protein-coding genes from the Ensembl database,² 55,765 target gene expression profile data was obtained. Finally, 27 m⁶A methylase regulators (i.e., writers, erasers and readers) in three types were used with reference to previous studies (Deng et al., 2018). Among them, 24 m⁶A methylase regulators were supported by expression profile data (**Supplementary Table 2**).

Sample Classification

Patients with qualitatively different m⁶A modification patterns were classified by using the R package of ConsensusClusterPlus (gapttools). The optimal *k* value was chosen by identifying the inflection point of the sum of the squared error (SSE). The decline slows down after *k* = *i*, and *k* = *i* is selected.

Screening for Differential Genes in m⁶A clusters

Edger analysis (gapttools) was used to identify differentially expressed genes among m⁶A cluster patterns, and 4745 genes and gene expression profiles of differentially expressed genes among m⁶A cluster patterns were obtained. Adjusted *P*-value < 0.05 was considered to be the criteria of differentially expressed genes.

Construction of the Gene Co-expression Network and Identification of the Key Modules

High-connection genes and modules were identified using the R-package “WGCNA.” Each module was constructed by

¹<https://www.ncbi.nlm.nih.gov/gds>

²<https://grch37.ensembl.org/index.html>

calculating the soft threshold power β . The WGCNA algorithm was used to construct the co-expression module after setting the soft threshold power value. Adjacency information was converted into topological overlap, and the network connectivity of genes was measured. Based on the dissimilarity of the topological overlap matrix (TOM), a hierarchical clustering function was performed to classify genes with similar expression profiles into modules (the minimum size of gene dendrogram = 25; merged highly similar modules with height = 0.25). The correlation between the clinical traits and the module eigengenes was used to determine the relevant modules. The correlation of gene expression profiles and the module eigengenes was used to quantitatively measure module membership. Gene significance was defined as the absolute value of the correlation between genes and clinical traits. Then, relevant modules of high importance for clinical traits were identified (Supplementary Table 3).

Functional Enrichment Analysis of Genes in the Key Modules

The R package clusterProfiler was used to perform functional enrichment analysis of genes in the key modules. The results of the molecular function (MF), cellular component (CC), and biological functions (BF) in the Gene Ontology (GO) analysis and the enrichment analysis of Kyoto Encyclopedia of Genes and Genomes (KEGG) were extracted.

Identification of the Hub Genes in the Co-expression Networks

To search for interacting genes, genes in the key modules were imported into a search tool (STRING³; score >900) to predict the PPI network. We defined interaction network hub genes as a degree of nodes ≥ 200 and interaction network non-hub genes as a degree <200 in the PPI (Supplementary Table 4). Genes with significance >0.6 and module membership >0.8 in the modules were defined as key genes (Supplementary Table 5). Finally, hub genes identified in the disease-related module and hub genes in PPI were intersected, and 29 disease-related hub genes were finally obtained (Supplementary Table 6).

Difference Analysis of Immune Infiltration and Therapeutic Effect

The CiberSort algorithm of CiberSort software was performed to calculate the proportion of immune cell infiltration of different types, and a stacking diagram and box plot were drawn for display. The differential expression of genes at the immune examination sites (PD-L1, IDO1, CD86, ICOS, TNFRSF9, and CTLA4) and the response to anti-TNF treatment were analyzed.

Statistical Analysis

Most analyses were performed using R software,⁴ Prism software,⁵ and SPSS software⁶ using two-tailed unpaired Student's

t-test or log-rank test unless otherwise specified. The chi-square test was applied to compare the response rates to therapy. In all analyses, *P*-values were bilateral, and *P* < 0.05 was considered statistically significant.

RESULTS

Landscape of m⁶A Regulators and the Mucosal Immune Microenvironment in IBD

Gene-Expression Omnibus data sets with clinical information and available data (GSE111889, Supplementary Table 1; Lloyd-Price et al., 2019) were enrolled in our study. Based on the previous data (Deng et al., 2018), we identified 24 m⁶A regulators, including nine “writers,” two “erasers,” and 13 “readers” (Supplementary Table 2). The reversible process of m⁶A regulator-mediated methylation is summarized in Figure 1A, and the locations of m⁶A regulators on chromosomes are shown in Figure 1B. To ascertain whether the above m⁶A regulators influenced the disease progression of IBD, we evaluated the mRNA expression of m⁶A regulators in UC, CD, and normal tissues and found that expression of “readers” (i.e., IGF2BP1 and IGF2BP2) was dramatically decreased in CD tissues compared with normal tissues (*P* < 0.01) (Figure 1D). The expression of IGF2BP2 was markedly decreased in UC tissues compared with normal tissues (*P* < 0.001) (Figure 1C). The expression of IGF2BP1, METTL16, YTHDC2, and KIAA1429 was downregulated in CD tissues compared with UC tissues (*P* < 0.05) (Supplementary Figure 1A). These results suggest that m⁶A regulators exhibit high heterogeneity of the expressional alteration landscape among normal, UC, or CD patients, indicating that different m⁶A regulator expression may affect disease progression in different IBD types.

Additionally, we analyzed immune cell-infiltrating characteristics in the intestinal mucosa in IBD and found that UC was characterized by M1 macrophages (*P* < 0.001); M0 macrophages (*P* < 0.001); neutrophils (*P* < 0.001); eosinophils (*P* < 0.01) (Figures 1E,I); and the inflammatory factors IL18, IL8, and IL15 (Figure 1H). M1 macrophages, neutrophils, and M0 macrophages were enriched, but memory B cells were decreased in CD tissues (Figures 1F,J; *P* < 0.001) with the upregulation of IL15, IL8, and IL1 β expression (Figure 1G). Compared with CD tissues, eosinophils and follicular helper T cells were more abundant in UC tissues (Figure 1K; *P* < 0.01). These results indicate that alterations in the mucosal immune microenvironment may be the prominent pathogenic factors of IBD, and local infiltration of M1 macrophages and neutrophils may promote the progression of IBD.

m⁶A Methylation Modification Patterns Are Mediated by 24 m⁶A Regulators in IBD Patients

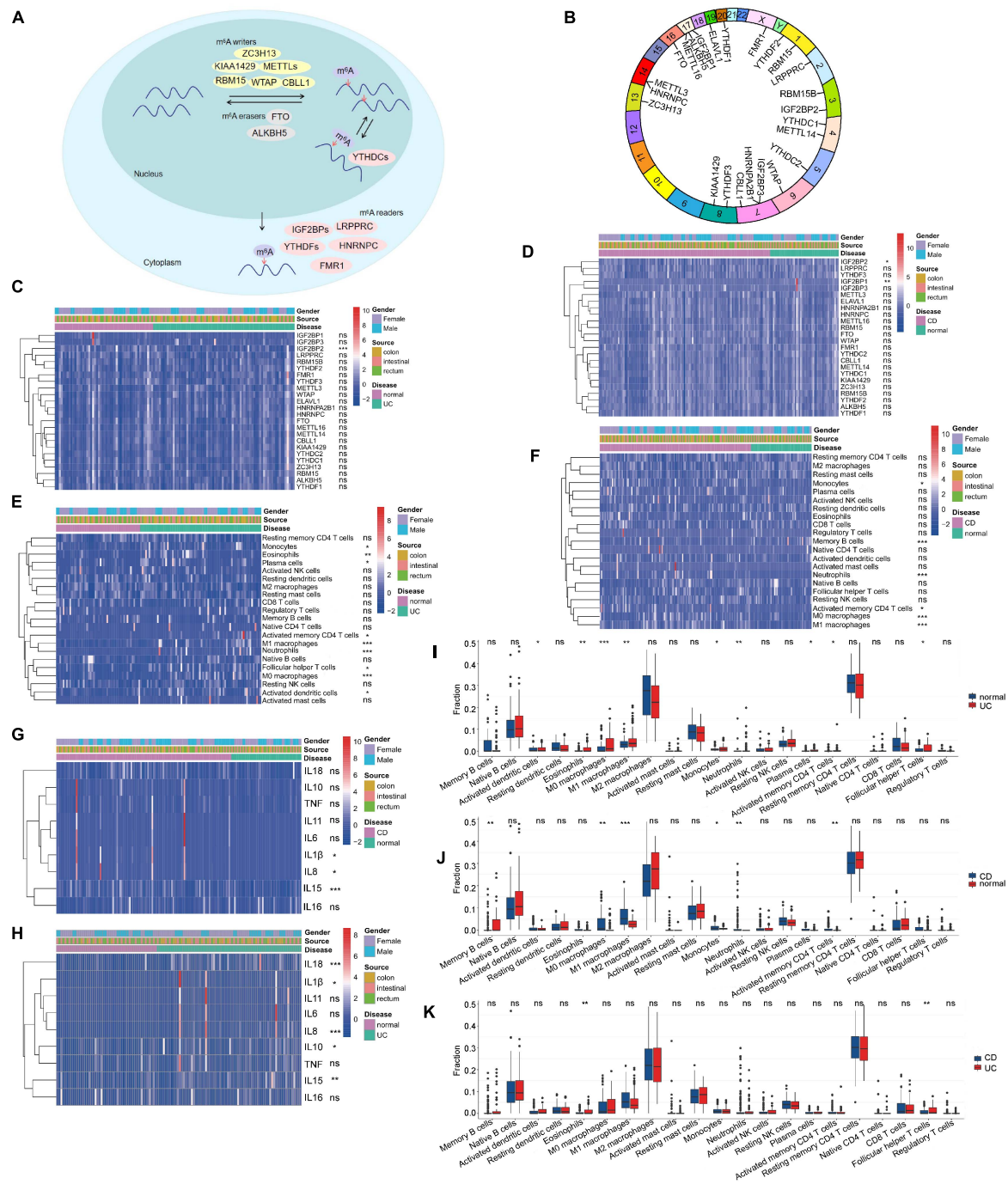
According to the expression of the 24 m⁶A regulators, m⁶A modification patterns were classified using the R package of ConsensusClusterPlus (gapttools). The optimal clustering

³<https://string-db.org/>

⁴<https://www.bioconductor.org/>

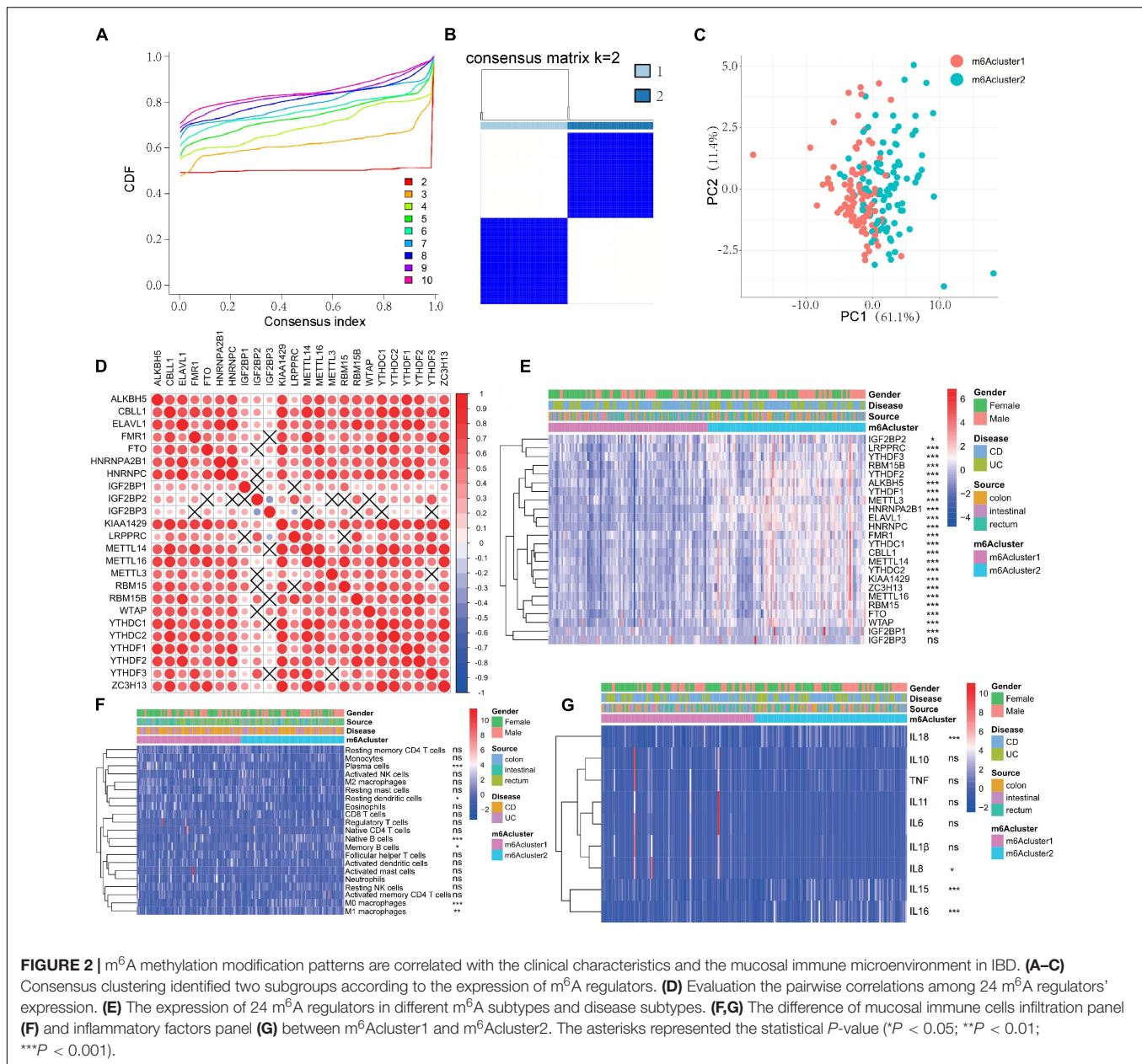
⁵<https://www.graphpad.com/>

⁶<https://www.ibm.com/analytics/spss-statistics-software>



stability ($k = 2-10$) was determined by the similarity of m⁶A regulator expression and the proportion of ambiguous clustering measurement, and $k = 2$ was identified (Figures 2A,B and Supplementary Figure 1B). Based on unsupervised clustering,

we eventually identified two distinct modification patterns, including 99 cases in m⁶Acluster1 and 98 cases in m⁶Acluster2 (Supplementary Table 7). IBD samples could be completely distinguished based on the expression of these 24 m⁶A regulators



(Figure 2C). To explore the correlation among “writers,” “erasers,” and “readers,” we evaluated pairwise correlations among the expression of 24 m⁶A regulators in IBD and revealed that positive correlations were more frequent (Figure 2D). We found that remarkable correlations existed among “writers,” “erasers,” and “readers,” and m⁶A regulators with the same function exhibited significant correlations in expression. IBD with high expression of IGF2BP2 (a “reader” gene) exhibited downregulation of IGF2BP3 (a “reader” gene) although upregulation of IGF2BP2 did not affect expression of other genes except for YTHDF3, LRPPRC, and RBM15B (“writer” genes). The “writer” genes CBL1, KIAA1429, ZC3H13, and METTL16 presented a common trend in FTO (an “eraser” gene) expression. IBD with upregulation of the “writer” gene KIAA1429 displayed

a common trend in ALKBH5 (an “eraser” gene) expression. Additionally, the change of IGF2BPs expression did not affect expression of these “eraser” genes. Therefore, crosstalk among different m⁶A regulators may be crucial for the generation of m⁶A modification patterns in individual IBD patients. After unsupervised clustering, we investigated the expression of 24 m⁶A regulators in different m⁶A subtypes and disease subtypes (Figure 2E) and found that the expression of m⁶A methylation regulators was increased in the m⁶A cluster2 compared with the m⁶A cluster1 except for IGF2BP3 (*P* < 0.05). Next, the clinicopathological features between the two subtypes were compared. However, there were no significant differences between m⁶A cluster1 and m⁶A cluster2 with respect to patient gender, location, or disease subtypes (*P* > 0.05).

m⁶A Modification Patterns Regulate the Mucosal Immune Microenvironment in IBD

We next compared differences in immune cell infiltration between the two m⁶A modification patterns using CIBERSORT. To our surprise, m⁶Acluster2 was prominently associated with immune activation (**Figure 2F**). Subsequent analyses of immune infiltration suggested that m⁶Acluster2 was significantly abundant in innate immune cell infiltration, including M0 macrophage ($P < 0.001$), M1 macrophage ($P < 0.01$), and acquired immune cells (native B cells, memory B cells; $P < 0.05$). m⁶Acluster1 was enriched in innate immune cell infiltration, including resting DCs ($P < 0.05$) and plasma cells ($P < 0.001$) (**Figure 2F**). Furthermore, we found that expression of IL15, IL16, and IL18 was upregulated in m⁶Acluster2 ($P < 0.001$). However, expression of IL8 in m⁶Acluster1 was increased ($P < 0.05$) (**Figure 2G**). The significant differences in the infiltration of immune cells and inflammatory factors between the two m⁶Aclusters indicate that m⁶A methylation modifications may change the mucosal immune microenvironment of IBD.

Biological Characteristics of Key m⁶A Modules

To explore the differentially expressed genes in the two m⁶Aclusters, we used the Edger package in GAPTools. Subsequently, 4737 differentially expressed genes were screened out between the two clusters, among which 2274 genes were upregulated and 2463 genes were downregulated. A weighted gene co-expression network was then constructed (**Supplementary Figure 1C**). After calculating the eigengenes of each module, performing cluster analysis on the modules, and merging the modules close to each other into new modules, we obtained seven modules (**Figures 3A,B**). The turquoise module was the most correlated with the m⁶Acluster classification phenotype (**Figure 3B** and **Supplementary Table 3**; $P = 2e-30$). The R package clusterProfiler was used to analyze the function and pathway information of the top 20 enriched genes in the turquoise module. As shown in **Figure 3C**, genes in the turquoise module were markedly enriched in the biological processes (BP), including ribonucleoprotein complex biogenesis and ribosome biogenesis (**Figure 3C**; $P < 0.05$); CC were involved, including the composition of ribosomal subunit, chromosomal region (**Figure 3D**; $P < 0.05$); MF-related genes were markedly enriched in structural constituent of ribosome (**Figure 3E**; $P < 0.05$); the KEGG pathway was enriched in ribosome and spliceosome (**Figure 3F**; $P < 0.05$).

Twenty-Nine m⁶A Phenotype-Related Hub Genes Are Correlated With the IBD Characteristics

To further validate the correlation of m⁶A phenotype-related hub genes with disease characteristics, we constructed a PPI network to analyze 1463 related genes in the key module (**Figure 4A**), identifying 112 hub genes (**Supplementary Table 4**).

Genes (gene significance > 0.6 ; module membership > 0.8) were identified from the turquoise module as key genes with a total of 321 genes (**Figure 4B** and **Supplementary Table 5**). Finally, the hub genes of the interaction network were intersected with the hub genes identified in the m⁶A-related module, and 29 disease-related hub genes were obtained (**Figure 4C** and **Supplementary Table 6**).

We found that 29 disease-related genes were significantly different among the m⁶Acluster classifications. The expression of 29 genes in m⁶Acluster2 was higher than that in m⁶Acluster1 (**Figure 4D**; $P < 0.001$). Furthermore, we explored differences in the expression of 29 disease-related genes between normal, CD, and UC tissues and found that the expression of NUP37, SNRPG, H2AFZ, MAGOHB, NUP107, ALYREF, CDK2, F3, SNRPD1, SNRPF, and UBE2N was significantly upregulated in the IBD tissues compared with normal tissues (**Figure 4E**; $P < 0.05$) although there was no significant difference in the other 18 genes (**Figure 4E**; $P > 0.05$). Compared with normal tissues, NUP37, H2AFZ, SNRPG, RBM8A, RNPS1, SNRPD3, ALYREF, MAGOHB, U2AF2, and SRRT were increased in CD tissues (**Figure 4F**; $P < 0.05$). Expression of NUP37, RBM8A, RNPS1, H2AFZ, SNRPG, PRPF19, U2AF2, ALYREF, MAGOHB, NHP2L1, SNRPD1, SNRPD3, MAGOH, POLR2K, CPSF3, and SRSF9 was markedly higher in UC tissues than in normal tissues (**Figure 4G**; $P < 0.05$). However, only CPSF3, NUP37, SRSF3, UBE2N, and SRSF1 were upregulated in UC tissues but not in CD tissues (**Figure 4H**; $P < 0.05$). These results suggest that 29 m⁶A phenotype-related hub genes are differentially expressed in different IBD types.

Consensus Clustering for m⁶A Phenotype-Related Hub Genes With the Characteristics of IBD Patients

Based on expression of the 29 m⁶A phenotype-related genes, we classified IBD patients into different genomic subtypes by unsupervised clustering analyses (K-means) and found that the gene clusters could be well distinguished (**Figure 5A** and **Supplementary Figure 1D**). We found that there were marked differences among the 29 key disease genes in the gene cluster classification ($P < 0.001$) with higher gene expression in genecluster1 and lower gene expression in genecluster2 (**Figure 5C**). As shown in **Figure 5B**, almost all patients with the m⁶Acluster2 subtype were classified into genecluster1. However, no significant difference in patient gender, location, or disease subtypes was observed between genecluster1 and genecluster2 (**Figures 5B,C**; $P > 0.05$). The expression of all 29 m⁶A phenotype-related hub genes was significantly higher in genecluster1 than in genecluster2 (**Figure 5D**; $P < 0.001$).

m⁶A Phenotype-Related Hub Genes Influence Immune Infiltration and Therapeutic Response

To explore the effect of m⁶A phenotype-related hub genes on the immune infiltration of IBD, we evaluated immune infiltration between genecluster1 with upregulated m⁶A phenotype-related hub gene expression and genecluster2 with downregulated m⁶A

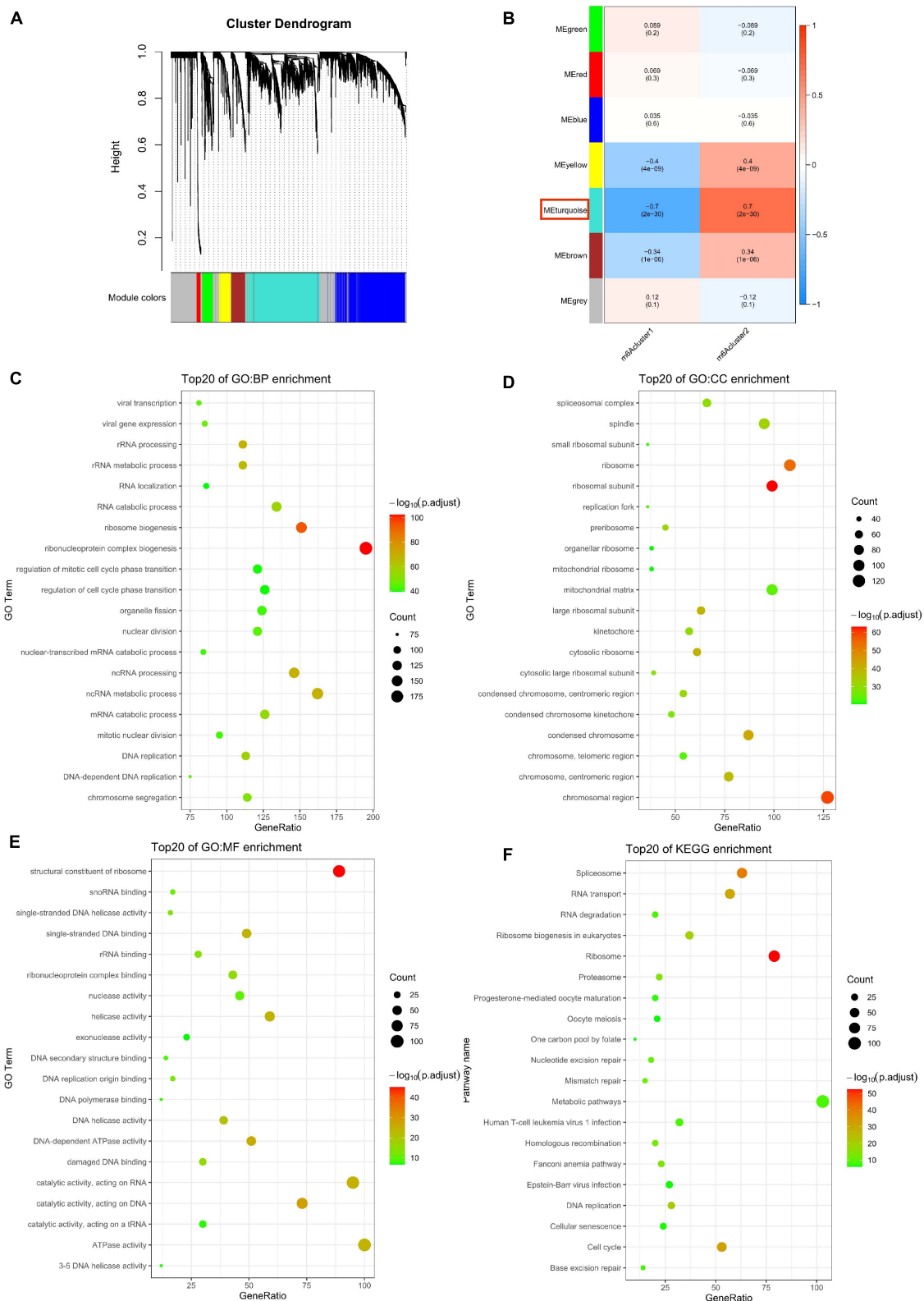


FIGURE 3 | Biological characteristics of key m⁶A module. **(A)** Gene dendrogram obtained by average linkage hierarchical clustering. **(B)** The correlation between the module eigengenes and m⁶A modification patterns. **(C–F)** Functional annotation of the genes in the key m⁶A module by using Gene Ontology (GO) terms of biological processes (BP) panel **(C)**, cellular component (CC) panel **(D)**, molecular function (MF) panel **(E)**, and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway **(F)**.

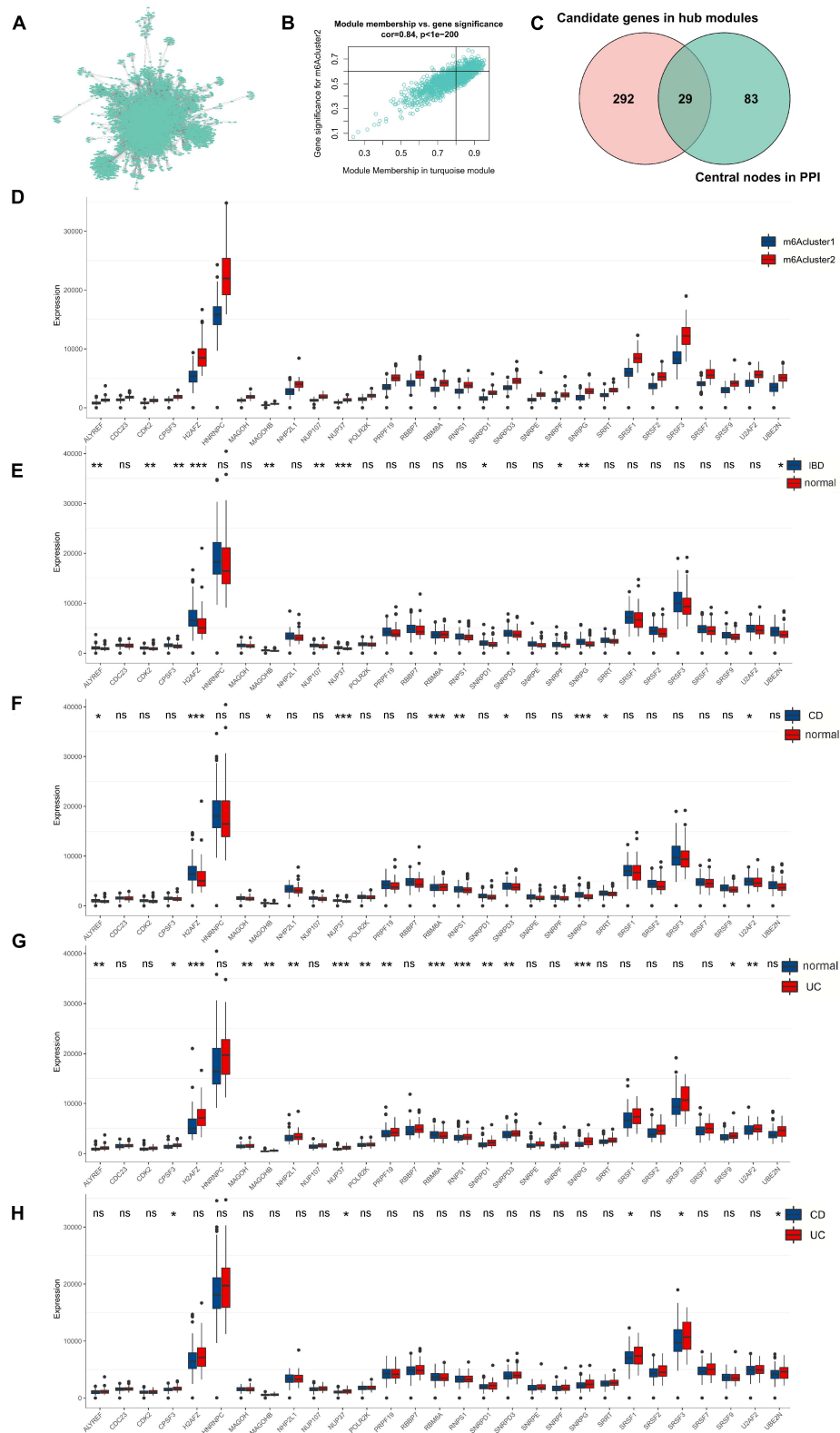


FIGURE 4 | Twenty-nine m⁶A phenotype-related hub genes affect disease characteristics in IBD. **(A)** The PPI network of genes in the key module. **(B)** The significance genes in the turquoise module. **(C)** The hub genes of the interaction network are intersected with the hub genes identified in the m⁶A-related module. **(D)** The 29 genes' expression in different m⁶A cluster classification. **(E–H)** The 29 genes' expression between normal, UC, and CD tissues. The asterisks represented the statistical *P*-value (**P* < 0.05; ***P* < 0.01; ****P* < 0.001).

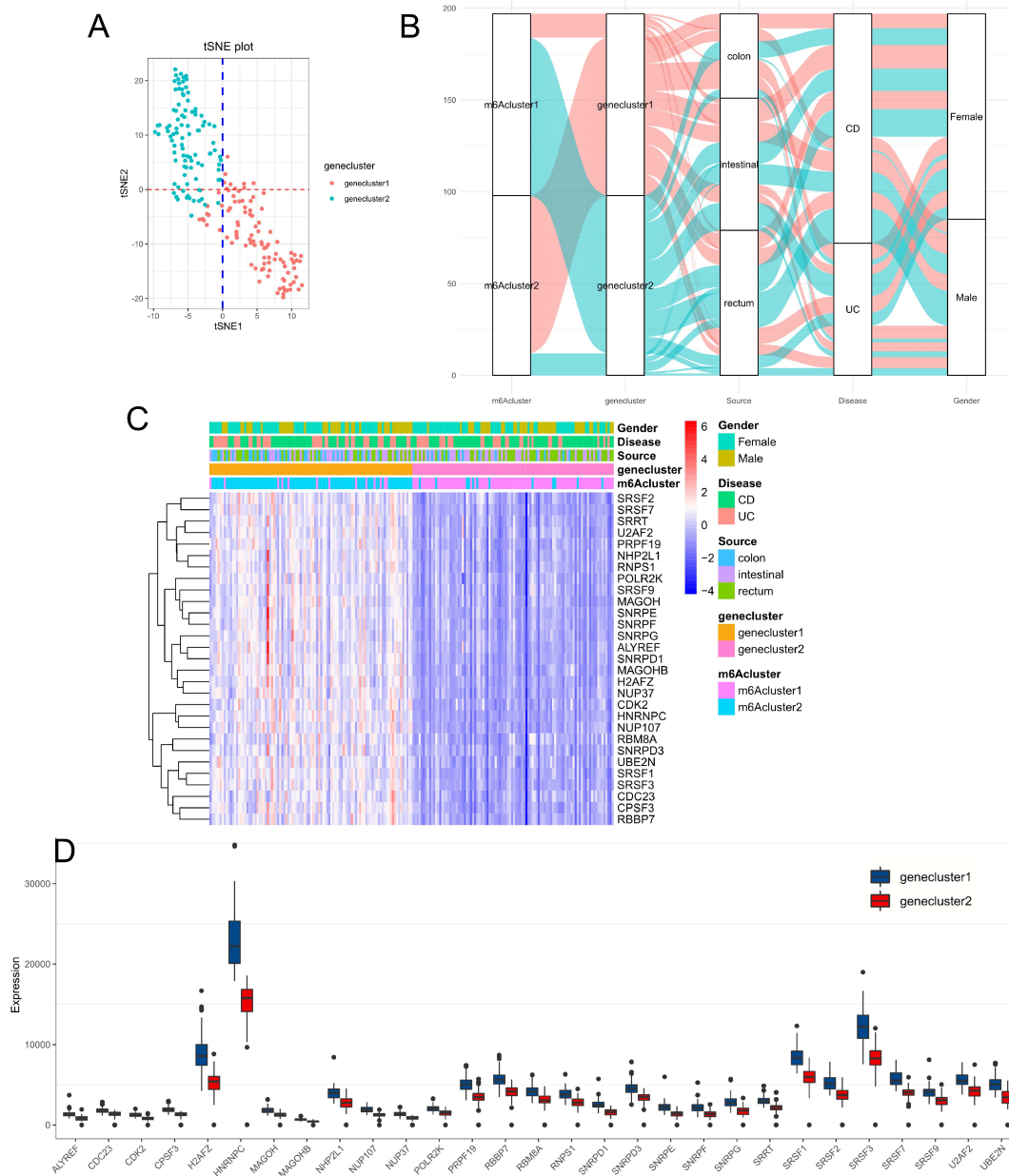


FIGURE 5 | Consensus clustering for m⁶A phenotype-related hub genes with the characteristics of IBD patients. **(A)** Clustering the samples with the tSNE method. **(B)** Alluvial diagram showing the changes of m⁶A clusters, disease subtype, location, gender, and gene clusters. **(C,D)** The expression of 29 hub genes between genecluster1 and genecluster2.

phenotype-related hub gene expression. Genecluster1 displayed higher infiltration levels of M1 macrophages, M0 macrophages, naive B cells, CD4 memory-activated T cells, memory B cells, and activated DCs, whereas genecluster2 was more correlated with resting DCs, M2 macrophages, CD4 memory resting T cells, plasma cells, resting mast cells, and eosinophils (Figures 6A,B; $P < 0.05$).

Expression differences of nine important immune checkpoints supported by expression data among different gene clusters are shown in Figure 6C. We found that the expression of all immune

checkpoints in genecluster1 was significantly higher than that in genecluster2 ($P < 0.01$), indicating that m⁶A phenotype-related hub genes might affect immunotherapy. Current treatments are biological agents anti-TNF (primarily infliximab) for refractory and severe forms of IBD. Thus, we further verified the 29 m⁶A phenotype-related hub genes in GSE16879 (Arijs et al., 2009), which was available data and clinical treatment information. However, the expression of only 26 hub genes was supported by expression data in the GSE16879 data set. Based on the expression of 26 hub genes, we classified patients using the

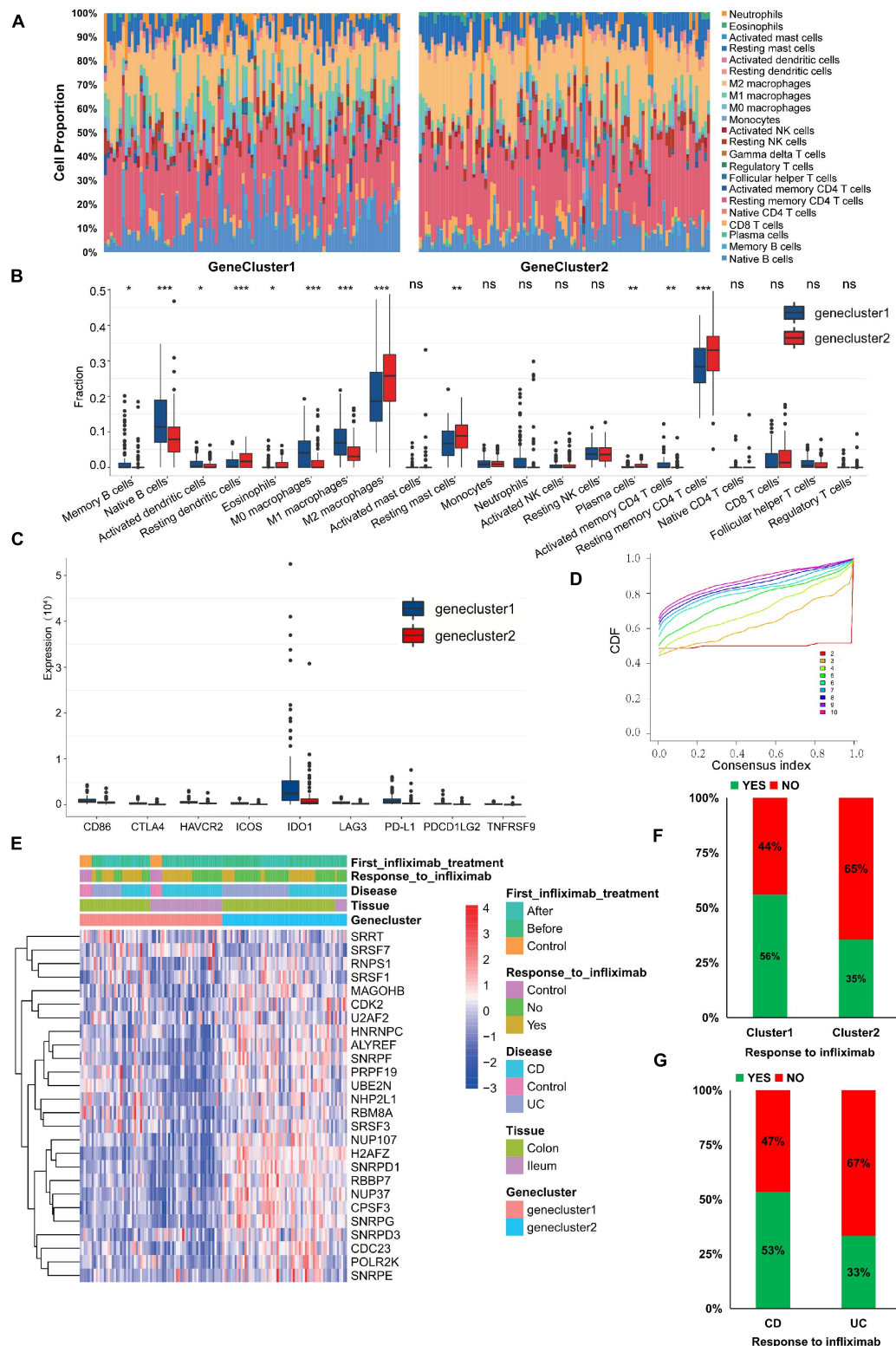


FIGURE 6 | m⁶A phenotype-related hub genes influence the immune infiltration and therapeutic response. **(A,B)** The immune cell infiltrating levels in different gene cluster subtypes in the GSE111889 cohort. **(C)** The expression differences of nine important immune checkpoints among different gene clusters. **(D)** Consensus clustering cumulative distribution function (CDF) for $k = 2-10$. **(E)** The expression of hub genes, anti-TNF therapeutic response, and disease subtype between GeneCluster1 and GeneCluster2. **(F,G)** Infliximab therapeutic response in different gene clusters **(F)** and disease subtypes **(G)**. The asterisks represented the statistical P -value (* $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$).

R package of ConsensusClusterPlus (gaptools), and $k = 2$ was identified (Figure 6D and Supplementary Figure 1E). Significant differences in infliximab therapy existed between the two gene clusters (Figure 6E). The expression of hub genes (H2AFZ, NUP37, SNRPD1, CPSF3, RBBP7, and SNRPG) in genecluster2 was significantly higher than in genecluster1, whereas expression of SRSF7 and SRRT in genecluster2 were lower than in genecluster1 (Figure 6E; $P < 0.05$). The response to infliximab in genecluster1 was better than that in genecluster2 (Figure 6F), and the response to infliximab in CD was better than that in UC (Figure 6G), suggesting that m⁶A-related hub gene expression may affect the therapeutic effect of infliximab. These results indicate that IBD with high expression of some m⁶A phenotype-related hub genes may represent those characterized by resistance to infliximab. Assessing the expression of m⁶A phenotype-related hub genes might guide the choice of IBD drugs and predict the response to anti-TNF therapy.

DISCUSSION

Currently, as most studies on disease-related m⁶A modification are limited to various tumors, single regulators, or single immune cell types, the effects of multiple m⁶A regulators on the overall immune infiltration characteristics and on IBD progression are poorly understood. Clarifying the role of m⁶A modification in the immune cell infiltration of IBD will improve our understanding of disease progression and disease-related local immune infiltration and strengthen more effective therapeutic strategies.

N⁶-methyladenosine regulators are involved in inflammatory development. However, only one study reports an association between m⁶A regulators and disease progression in IBD, and it reports that METTL14 (a “writer” gene) deficiency in T cells promotes spontaneous colitis in mice (Lu et al., 2020). IGF2BPs, a new family of m⁶A “readers,” are composed of two RNA recognition motif domains and four K homology domains (Bell et al., 2013) and promote the processes of various tumors (Huang X. et al., 2018; Huang S. et al., 2019; Hanniford et al., 2020). MYC, ACTIN, and LIN28B are reported as targets of IGF2BPs (Nielsen et al., 1999). Here, consistent with a previous study (Chen et al., 2021), we find that not only remarkable correlations exist among “writers,” “erasers,” and “readers,” but also that m⁶A regulators with the same function display significant correlations in expression. We find that expression of “readers” (i.e., IGF2BP1 and IGF2BP2) was downregulated in CD tissues compared with normal tissues, and expression of the IGF2BP2 was markedly decreased in UC tissues. The m⁶A regulators among normal, UC, and CD patients exhibit high heterogeneity of the expressional alteration landscape. However, the exact molecular mechanisms by which m⁶A regulators modulate IBD (UC and CD) development remain elusive.

As a reversible RNA methylation, m⁶A is added by “writers” and removed by “erasers,” indicating that they may represent opposite roles (Lan et al., 2019). However, it is interesting to find that m⁶A writers and erasers were positively correlated

in our study. The “writer” genes CBLL1, KIAA1429, ZC3H13, and METTL16 presented a common trend in FTO (an “eraser” gene) expression. The “writer” gene KIAA1429 displayed a common trend in ALKBH5 (an “eraser” gene) expression in IBD. Previous studies find that m⁶A writers that increase the RNA m⁶A level are oncogenes (Lan et al., 2019; Qian et al., 2019; Wang X. K. et al., 2021), and FTO and ALKBH5 that decrease the RNA m⁶A level are also oncogenes (Shen et al., 2020; Su et al., 2020). The targets of “erasers” and “writers” that are reported in different studies are different, suggesting different mechanisms and other regulatory factors exist in disease development other than the “writer” and “eraser.” Additionally, the m⁶A “readers” exert post-transcriptional functions, which may partly account for the seemingly contradictory roles between the “eraser” and “writer.” In the future, the effect of m⁶A “writers” and “erasers” on the same RNA simultaneously should be examined to explore whether other regulatory factors are involved in m⁶A modification.

More recently, m⁶A regulators are shown to play crucial roles in disease-related immune cell infiltration. ALKBH5 inhibits antiviral innate responses by erasing their m⁶A modification (Zheng et al., 2017). METTL3, potentially serving as an anti-inflammatory target, methylates STAT1 mRNA and drives M1 macrophage polarization (Liu et al., 2019; Wang J. et al., 2019). LPS-induced inflammatory responses are regulated by the m⁶A reader YTHDF2 in macrophages (Yu et al., 2019). FTO knockdown induces a decrease of PPAR- γ and STAT1 transcripts through accelerating mRNA degradation mediated by YTHDF and inhibiting M1/M2 activation (Gu et al., 2020). It is reported that the immune cells contributing to IBD progression are predominantly M1 macrophages, neutrophils, and Th cells (de Souza and Fiocchi, 2016; Zhou et al., 2018, 2019). The m⁶A “writer” protein METTL3 has important roles in reprogramming naive T cells for proliferation and differentiation (Li et al., 2017). Deficiency of METTL14, another m⁶A “writer,” causes induction of native T cells into induced Treg cells (Lu et al., 2020). In this study, we find that m⁶Acluster2 with a higher expression of m⁶A regulators was significantly abundant in innate immune cell infiltration, including M0 macrophage, M1 macrophage, and acquired immune cells (native B cells and memory B cells), and m⁶Acluster1 with a lower expression of m⁶A regulators was enriched in innate immune cell infiltration, including resting DCs and acquired immune cells (plasma cells), suggesting that m⁶A methylation modification may upregulate innate immune cells associated with inflammation and inhibit the activation of acquired immune B cells during IBD disease progression. In addition, neutrophils and activated memory CD4 T cells were more highly expressed in CD and UC tissues than in normal tissues, but no significant differences in the T cell types or neutrophils exist between the two m⁶A modification patterns, indicating that m⁶A regulators may influence inflammation development primarily by recruiting macrophages. METTL3 deficiency increases the secretion of IL8 and elevates the recruitment of neutrophils (He et al., 2021). In our study, IL8 expression in m⁶Acluster1 was increased, and expression of IL15, IL16, and IL18 was upregulated in m⁶Acluster2, suggesting that m⁶A methylation modification may

change mucosal inflammatory factors in IBD. Nevertheless, our data on m⁶A regulators are confined to bioinformatics analysis, and more research is needed to explore the exact molecular mechanisms by which these disease-related m⁶A regulators induce immune cell infiltration in IBD.

Next, we identified 29 disease-related hub genes from the key module, which was the most correlated with the m⁶A cluster phenotype, and found that expression of NUP37, SNRPG, H2AFZ, and ALYREF was significantly upregulated in the IBD (CD and UC) tissues compared with normal tissues. NUP37, a component of the nuclear pore complex, is both a significantly mutated gene and a tumor-destructive gene in various carcinomas (Chen J. et al., 2019; Huang L. et al., 2020). SNRPG belongs to the small nuclear ribonucleoprotein peptide family, which might participate in tumor chemotherapeutic resistance (Lan et al., 2020). H2AFZ is significantly upregulated in hepatocellular carcinoma and is related to poor prognosis (Tang et al., 2020). ALYREF activates the Wnt/ β -catenin pathway to promote cell proliferation (Wang J. et al., 2021). However, there are few reports on the association of these genes with IBD or other inflammatory diseases. Our study suggests that these hub genes may be downstream targets of m⁶A regulators and can be expected to trigger new therapeutic strategies for IBD. However, further experiments are needed to verify their roles in IBD development.

Furthermore, we classified patients into different genomic subtypes according to expression of the 29 obtained m⁶A phenotype-related hub genes. Genecluster1, with higher expression of m⁶A phenotype-related hub genes, exhibited higher infiltration levels of M1 macrophages, M0 macrophages, naive B cells, CD4 memory-activated T cells, memory B cells, and activated DCs, whereas genecluster2 was more correlated with resting DCs, M2 macrophages, CD4 memory resting T cells, plasma cells, resting mast cells, and eosinophils. Despite these findings, these hub genes could, at least in part, explain how m⁶A regulators stimulate the inflammation-associated immune cell infiltration in IBD.

N⁶-methyladenosine regulators are closely associated with immunotherapy (Yang et al., 2019; Li et al., 2020). For example, inhibiting FTO sensitizes melanoma cells to anti-PD-1 and interferon gamma (IFN γ) treatment (Yang et al., 2019). Knockout of YTHDF1 enhances the therapeutic efficacy of PD-L1 checkpoint blockade in mice (Han et al., 2019). Downregulation of ALKBH5 correlates with a positive response to PD-1 blockade in melanoma patients (Li et al., 2020). In this study, we found that expression levels of all immune checkpoints in genecluster1 with a higher expression of m⁶A phenotype-related hub genes was significantly increased, suggesting that m⁶A phenotype-related hub genes might influence immunotherapy. Furthermore, as current treatments are mostly the biological agent anti-TNF (primarily infliximab) for refractory and severe forms of IBD (Koelink et al., 2020), we verified the m⁶A phenotype-related hub genes in GSE16879 and found that the therapeutic effect of cluster1 with the lower expression of hub genes, including H2AFZ, NUP37, SNRPD1, CPSF3, and RBBP7, was better than that of cluster2. Furthermore, the therapeutic effect in CD was better than that in UC, indicating that IBD with high expression

of some m⁶A phenotype-related hub genes (H2AFZ, NUP37, SNRPD1, CPSF3, and RBBP7) may represent those characterized by resistance to infliximab. Our findings provide new possibilities for enhancing the efficacy of anti-TNF therapy for IBD. However, our study is on bioinformatics and is retrospective in nature, and the predictive value of gene signatures should be further verified.

CONCLUSION

We identified for the first time the alterations and correlations among m⁶A modulators, m⁶A-related immune infiltration, m⁶A phenotype-related hub genes, and therapeutic response in IBD. Assessing the expression of m⁶A phenotype-related hub genes may guide the choice of IBD drugs and improve the prediction of therapeutic response to anti-TNF therapy.

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.

ETHICS STATEMENT

The IBD patient data in our study were downloaded from the publicly available data sets in which informed consent were complete.

AUTHOR CONTRIBUTIONS

SH and YC: study design and drafting manuscript. YC and JL: data analysis and statistical analysis. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by grants from China Postdoctoral Science Foundation (No. 2020M680145).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fcell.2021.692160/full#supplementary-material>

Supplementary Figure 1 | Landscape of m⁶A regulators in IBD subtypes and consensus clustering and gene co-expression network in different IBD cohort. **(A)** The expression of 24 m⁶A regulators between UC and CD tissues. **(B)** Consensus clustering CDF for $k = 2-10$ based on m⁶A regulators' expression in the GSE111889 cohort. **(C)** A weighted gene co-expression network is constructed for candidate gene sets. **(D)** Clustering CDF for $k = 2-10$ based on the expression of hub genes in the GSE111889 cohort. **(E)** Consensus clustering CDF for $k = 2-10$ based on hub genes expression in the GSE16879 cohort. The asterisks represented the statistical P -value ($*P < 0.05$).

REFERENCES

- Arijs, I., De Hertogh, G., Lemaire, K., Quintens, R., Van Lommel, L., Van Steen, K., et al. (2009). Mucosal gene expression of antimicrobial peptides in inflammatory bowel disease before and after first infliximab treatment. *PLoS One* 4:e7984. doi: 10.1371/journal.pone.0007984
- Belarif, L., Danger, R., Kermarrec, L., Nerrière-Daguin, V., Pengam, S., Durand, T., et al. (2019). IL-7 receptor influences anti-TNF responsiveness and T cell gut homing in inflammatory bowel disease. *J. Clin. Invest.* 129, 1910–1925. doi: 10.1172/JCI121668
- Bell, J. L., Wachter, K., Muhleck, B., Pazaitis, N., Kohn, M., Lederer, M., et al. (2013). Insulin-like growth factor 2 mRNA-binding proteins (IGF2BPs): post-transcriptional drivers of cancer progression? *Cell. Mol. Life Sci.* 70, 2657–2675. doi: 10.1007/s00018-012-1186-z
- Canavan, J. B., Scotta, C., Vossenkomper, A., Goldberg, R., Elder, M. J., Shoval, I., et al. (2016). Developing in vitro expanded CD45RA⁺ regulatory T cells as an adoptive cell therapy for Crohn's disease. *Gut* 65, 584–594. doi: 10.1136/gutjnl-2014-306919
- Chen, H., Yao, J., Bao, R., Dong, Y., Zhang, T., Du, Y., et al. (2021). Cross-talk of four types of RNA modification writers defines tumor microenvironment and pharmacogenomic landscape in colorectal cancer. *Mol. Cancer* 20:29. doi: 10.1186/s12943-021-01322-w
- Chen, J., Wo, D., Ma, E., Yan, H., Peng, J., Zhu, W., et al. (2019). Deletion of low-density lipoprotein-related receptor 5 inhibits liver cancer cell proliferation via destabilizing Nucleoporin 37. *Cell Commun. Signal.* 17, 174. doi: 10.1186/s12964-019-0495-3
- Chen, X. Y., Zhang, J., and Zhu, J. S. (2019). The role of m(6)A RNA methylation in human cancer. *Mol. Cancer* 18:103. doi: 10.1186/s12943-019-1033-z
- de Souza, H. S., and Fiocchi, C. (2016). Immunopathogenesis of IBD: current state of the art. *Nat. Rev. Gastroenterol. Hepatol.* 13, 13–27. doi: 10.1038/nrgastro.2015.186
- Deng, X., Su, R., Weng, H., Huang, H., Li, Z., and Chen, J. (2018). RNA N(6)-methyladenosine modification in cancers: current status and perspectives. *Cell Res.* 28, 507–517. doi: 10.1038/s41422-018-0034-6
- Geula, S., Moshitch-Moshkovitz, S., Dominissini, D., Mansour, A. A., Kol, N., Salmon-Divon, M., et al. (2015). Stem cells. m6A mRNA methylation facilitates resolution of naive pluripotency toward differentiation. *Science* 347, 1002–1006. doi: 10.1126/science.1261417
- Gu, X., Zhang, Y., Li, D., Cai, H., Cai, L., and Xu, Q. (2020). N6-methyladenosine demethylase FTO promotes M1 and M2 macrophage activation. *Cell. Signal.* 69:109553. doi: 10.1016/j.cellsig.2020.109553
- Halfvarson, J., Brislawn, C. J., Lamendella, R., Vazquez-Baeza, Y., Walters, W. A., Bramer, L. M., et al. (2017). Dynamics of the human gut microbiome in inflammatory bowel disease. *Nat. Microbiol.* 2:17004. doi: 10.1038/nmicrobiol.2017.4
- Han, D., Liu, J., Chen, C., Dong, L., Liu, Y., Chang, R., et al. (2019). Anti-tumour immunity controlled through mRNA m(6)A methylation and YTHDF1 in dendritic cells. *Nature* 566, 270–274. doi: 10.1038/s41586-019-0916-x
- Hanniford, D., Ulloa-Morales, A., Karz, A., Berzoti-Coelho, M. G., Moubarak, R. S., Sanchez-Sendra, B., et al. (2020). Epigenetic silencing of CDR1as drives IGF2BP3-mediated melanoma invasion and metastasis. *Cancer Cell* 37, 55–70.e15. doi: 10.1016/j.ccell.2019.12.007
- He, J., Zhou, M., Yin, J., Wan, J., Chu, J., Jia, J., et al. (2021). METTL3 restrains papillary thyroid cancer progression via m(6)A/c-Rel/IL-8-mediated neutrophil infiltration. *Mol. Ther.* 29, 1821–1837. doi: 10.1016/j.ymthe.2021.01.019
- Huang, H., Weng, H., and Chen, J. (2020). The biogenesis and precise control of RNA m(6)A methylation. *Trends Genet.* 36, 44–52. doi: 10.1016/j.tig.2019.10.011
- Huang, H., Weng, H., Sun, W., Qin, X., Shi, H., Wu, H., et al. (2018). Recognition of RNA N(6)-methyladenosine by IGF2BP proteins enhances mRNA stability and translation. *Nat. Cell Biol.* 20, 285–295. doi: 10.1038/s41556-018-0045-z
- Huang, L., Wang, T., Wang, F., Hu, X., Zhan, G., Jin, X., et al. (2020). NUP37 silencing induces inhibition of cell proliferation, G1 phase cell cycle arrest and apoptosis in non-small cell lung cancer cells. *Pathol. Res. Pract.* 216:152836. doi: 10.1016/j.prp.2020.152836
- Huang, S., Wu, Z., Cheng, Y., Wei, W., and Hao, L. (2019). Insulin-like growth factor 2 mRNA binding protein 2 promotes aerobic glycolysis and cell proliferation in pancreatic ductal adenocarcinoma via stabilizing GLUT1 mRNA. *Acta Biochim. Biophys. Sin.* 51, 743–752. doi: 10.1093/abbs/gmz048
- Huang, X., Zhang, H., Guo, X., Zhu, Z., Cai, H., and Kong, X. (2018). Insulin-like growth factor 2 mRNA-binding protein 1 (IGF2BP1) in cancer. *J. Hematol. Oncol.* 11:88. doi: 10.1186/s13045-018-0628-y
- Huang, Y., Su, R., Sheng, Y., Dong, L., Dong, Z., Xu, H., et al. (2019). Small-molecule targeting of oncogenic FTO demethylase in acute myeloid leukemia. *Cancer Cell* 35, 677–691.e610. doi: 10.1016/j.ccell.2019.03.006
- Jiang, X., Liu, B., Nie, Z., Duan, L., Xiong, Q., Jin, Z., et al. (2021). The role of m6A modification in the biological functions and diseases. *Signal. Transduct. Target Ther.* 6:74. doi: 10.1038/s41392-020-00450-x
- Jin, D., Guo, J., Wu, Y., Yang, L., Wang, X., Du, J., et al. (2020). m(6)A demethylase ALKBH5 inhibits tumor growth and metastasis by reducing YTHDFs-mediated YAP expression and inhibiting miR-107/LATS2-mediated YAP activity in NSCLC. *Mol. Cancer* 19:40. doi: 10.1186/s12943-020-01161-1
- Koelink, P. J., Bloemendaal, F. M., Li, B., Westera, L., Vogels, E. W. M., van Roest, M., et al. (2020). Anti-TNF therapy in IBD exerts its therapeutic effect through macrophage IL-10 signalling. *Gut* 69, 1053–1063. doi: 10.1136/gutjnl-2019-318264
- Lan, Q., Liu, P. Y., Haase, J., Bell, J. L., Huttelmaier, S., and Liu, T. (2019). The critical role of RNA m(6)A methylation in cancer. *Cancer Res.* 79, 1285–1292. doi: 10.1158/0008-5472.CAN-18-2965
- Lan, Y., Lou, J., Hu, J., Yu, Z., Lyu, W., and Zhang, B. (2020). Downregulation of SNRPG induces cell cycle arrest and sensitizes human glioblastoma cells to temozolomide by targeting Myc through a p53-dependent signaling pathway. *Cancer Biol. Med.* 17, 112–131. doi: 10.20892/j.issn.2095-3941.2019.0164
- Li, H. B., Tong, J., Zhu, S., Batista, P. J., Duffy, E. E., Zhao, J., et al. (2017). m(6)A mRNA methylation controls T cell homeostasis by targeting the IL-7/STAT5/SOCS pathways. *Nature* 548, 338–342. doi: 10.1038/nature23450
- Li, N., Kang, Y., Wang, L., Huff, S., Tang, R., Hui, H., et al. (2020). ALKBH5 regulates anti-PD-1 therapy response by modulating lactate and suppressive immune cell accumulation in tumor microenvironment. *Proc. Natl. Acad. Sci. U.S.A.* 117, 20159–20170. doi: 10.1073/pnas.1918986117
- Liu, Y., Liu, Z., Tang, H., Shen, Y., Gong, Z., Xie, N., et al. (2019). The N(6)-methyladenosine (m(6)A)-forming enzyme METTL3 facilitates M1 macrophage polarization through the methylation of STAT1 mRNA. *Am. J. Physiol. Cell Physiol.* 317, C762–C775. doi: 10.1152/ajpcell.00212.2019
- Lloyd-Price, J., Arze, C., Ananthakrishnan, A. N., Schirmer, M., Avila-Pacheco, J., Poon, T. W., et al. (2019). Multi-omics of the gut microbial ecosystem in inflammatory bowel diseases. *Nature* 569, 655–662. doi: 10.1038/s41586-019-1237-9
- Lu, T. X., Zheng, Z., Zhang, L., Sun, H. L., Bissonnette, M., Huang, H., et al. (2020). A new model of spontaneous colitis in mice induced by deletion of an RNA m(6)A methyltransferase component METTL14 in T cells. *Cell. Mol. Gastroenterol. Hepatol.* 10, 747–761. doi: 10.1016/j.jcmgh.2020.07.001
- Neurath, M. (2017). Current and emerging therapeutic targets for IBD. *Nat. Rev. Gastroenterol. Hepatol.* 14:688. doi: 10.1038/nrgastro.2017.138
- Nielsen, J., Christiansen, J., Lykke-Andersen, J., Johnsen, A. H., Wewer, U. M., and Nielsen, F. C. (1999). A family of insulin-like growth factor II mRNA-binding proteins represses translation in late development. *Mol. Cell. Biol.* 19, 1262–1270. doi: 10.1128/MCB.19.2.1262
- Qian, J. Y., Gao, J., Sun, X., Cao, M. D., Shi, L., Xia, T. S., et al. (2019). KIAA1429 acts as an oncogenic factor in breast cancer by regulating CDK1 in an N6-methyladenosine-independent manner. *Oncogene* 38, 6123–6141. doi: 10.1038/s41388-019-0861-z
- Shen, C., Sheng, Y., Zhu, A. C., Robinson, S., Jiang, X., Dong, L., et al. (2020). RNA demethylase ALKBH5 selectively promotes tumorigenesis and cancer stem cell self-renewal in acute myeloid leukemia. *Cell Stem Cell* 27, 64–80.e69. doi: 10.1016/j.stem.2020.04.009
- Su, R., Dong, L., Li, Y., Gao, M., Han, L., Wunderlich, M., et al. (2020). Targeting FTO suppresses cancer stem cell maintenance and immune evasion. *Cancer Cell* 38, 79–96.e11. doi: 10.1016/j.ccell.2020.04.017
- Tang, S., Huang, X., Wang, X., Zhou, X., Huang, H., Qin, L., et al. (2020). Vital and distinct roles of H2A.Z isoforms in hepatocellular carcinoma. *Onco Targets Ther.* 13, 4319–4337. doi: 10.2147/OTT.S243823
- Torres, J., Mehandru, S., Colombel, J. F., and Peyrin-Biroulet, L. (2017). Crohn's disease. *Lancet* 389, 1741–1755. doi: 10.1016/S0140-6736(16)31711-1
- Ungaro, R., Mehandru, S., Allen, P. B., Peyrin-Biroulet, L., and Colombel, J. F. (2017). Ulcerative colitis. *Lancet* 389, 1756–1770. doi: 10.1016/S0140-6736(16)32126-2

- Wang, H., Hu, X., Huang, M., Liu, J., Gu, Y., Ma, L., et al. (2019). Mettl3-mediated mRNA m(6)A methylation promotes dendritic cell activation. *Nat. Commun.* 10:1898. doi: 10.1038/s41467-019-09903-6
- Wang, J., Li, Y., Xu, B., Dong, J., Zhao, H., Zhao, D., et al. (2021). ALYREF drives cancer cell proliferation through an ALYREF-MYC positive feedback loop in glioblastoma. *Onco Targets Ther.* 14, 145–155. doi: 10.2147/OTT.S286408
- Wang, J., Yan, S., Lu, H., Wang, S., and Xu, D. (2019). METTL3 attenuates LPS-induced inflammatory response in macrophages via NF-kappaB signaling pathway. *Mediators Inflamm.* 2019:3120391. doi: 10.1155/2019/3120391
- Wang, X. K., Zhang, Y. W., Wang, C. M., Li, B., Zhang, T. Z., Zhou, W. J., et al. (2021). METTL16 promotes cell proliferation by up-regulating cyclin D1 expression in gastric cancer. *J. Cell. Mol. Med.* 25, 6602–6617. doi: 10.1111/jcmm.16664
- Yang, S., Wei, J., Cui, Y. H., Park, G., Shah, P., Deng, Y., et al. (2019). m(6)A mRNA demethylase FTO regulates melanoma tumorigenicity and response to anti-PD-1 blockade. *Nat. Commun.* 10:2782. doi: 10.1038/s41467-019-10669-0
- Yu, R., Li, Q., Feng, Z., Cai, L., and Xu, Q. (2019). m6A reader YTHDF2 regulates LPS-induced inflammatory response. *Int. J. Mol. Sci.* 20:1323. doi: 10.3390/ijms20061323
- Zheng, Q., Hou, J., Zhou, Y., Li, Z., and Cao, X. (2017). The RNA helicase DDX46 inhibits innate immunity by entrapping m(6)A-demethylated antiviral transcripts in the nucleus. *Nat. Immunol.* 18, 1094–1103. doi: 10.1038/ni.3830
- Zhou, G., Yu, L., Fang, L., Yang, W., Yu, T., Miao, Y., et al. (2018). CD177(+) neutrophils as functionally activated neutrophils negatively regulate IBD. *Gut* 67, 1052–1063. doi: 10.1136/gutjnl-2016-313535
- Zhou, X., Li, W., Wang, S., Zhang, P., Wang, Q., Xiao, J., et al. (2019). YAP aggravates inflammatory bowel disease by regulating M1/M2 macrophage polarization and gut microbial homeostasis. *Cell Rep.* 27, 1176–1189. doi: 10.1016/j.celrep.2019.03.028

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

Advantages of publishing in Frontiers



OPEN ACCESS

Articles are free to read
for greatest visibility
and readership



FAST PUBLICATION

Around 90 days
from submission
to decision



HIGH QUALITY PEER-REVIEW

Rigorous, collaborative,
and constructive
peer-review



TRANSPARENT PEER-REVIEW

Editors and reviewers
acknowledged by name
on published articles

Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org

Contact us: frontiersin.org/about/contact



REPRODUCIBILITY OF RESEARCH

Support open data
and methods to enhance
research reproducibility



DIGITAL PUBLISHING

Articles designed
for optimal readership
across devices



FOLLOW US

@frontiersin



IMPACT METRICS

Advanced article metrics
track visibility across
digital media



EXTENSIVE PROMOTION

Marketing
and promotion
of impactful research



LOOP RESEARCH NETWORK

Our network
increases your
article's readership