frontiers | Research Topics

Application of epigenomics data to improve human and livestock health

Edited by
Xiao Wang, Ying Yu and Lingzhao Fang

Published in
Frontiers in Genetics
Frontiers in Cell and Developmental Biology

ISSN 1664-8714
ISBN 978-2-8325-2670-5
DOI 10.3389/978-2-8325-2670-5

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

Application of epigenomics data to improve human and livestock health

Topic editors
Xiao Wang — Konge Larsen ApS, Denmark
Ying Yu — China Agricultural University, China
Lingzhao Fang — Aarhus University, Denmark

Topic coordinator
Zhimin Wang — The First Affiliated Hospital of Zhengzhou University, China

Citation
Wang, X., Yu, Y., Fang, L., eds. (2023). *Application of epigenomics data to improve human and livestock health*. Lausanne: Frontiers Media SA.
doi: 10.3389/978-2-8325-2670-5

Author XW was employed by Konge Larsen ApS.
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.

Table of contents

frontiers | Frontiers in Genetics

TYPE Editorial
PUBLISHED 26 May 2023
DOI 10.3389/fgene.2023.1221035

OPEN ACCESS

EDITED AND REVIEWED BY
Michael E. Symonds,
University of Nottingham,
United Kingdom

*CORRESPONDENCE
Xiao Wang,
xiaowangzntc@163.com
Ying Yu,
yuying@cau.edu.cn
Lingzhao Fang,
lingzhao.fang@qgg.au.dk
Zhimin Wang,
fccwangzm@zzu.edu.cn

RECEIVED 11 May 2023
ACCEPTED 19 May 2023
PUBLISHED 26 May 2023

CITATION
Wang X, Yu Y, Fang L and Wang Z (2023), Editorial: Application of epigenomics data to improve human and livestock health. *Front. Genet.* 14:1221035. doi: 10.3389/fgene.2023.1221035

Editorial: Application of epigenomics data to improve human and livestock health

Xiao Wang[1,2]*, Ying Yu[3]*, Lingzhao Fang[4]* and Zhimin Wang[5]*

[1]Institute of Animal Science and Veterinary Medicine, Shandong Academy of Agricultural Sciences, Jinan, China, [2]Konge Larsen ApS, Kongens Lyngby, Denmark, [3]Laboratory of Animal Genetics and Breeding, Ministry of Agriculture and Rural Affairs of China, National Engineering Laboratory of Animal Breeding, College of Animal Science and Technology, China Agricultural University, Beijing, China, [4]Center for Quantitative Genetics and Genomics, Aarhus University, Aarhus, Denmark, [5]Division of Endocrinology and Metabolic Diseases, The First Affiliated Hospital of Zhengzhou University, Zhengzhou, China

KEYWORDS

epigenetics, epigenomics data, health improvement, human, livestock

Editorial on the Research Topic
Application of epigenomics data to improve human and livestock health

With rapid advancements in next-generation sequencing technology, an enormous amount of epigenomic sequencing data is generated and helps us identify the epigenomic biomarkers and interpret biological mechanisms underlying complex health traits in human and livestock.

In Chinese Yorkshire pigs, Wang et al. reported associations of meat quality traits with DNA methylation and identified several candidate genes associated with these traits, such as *NCAM1*, *MED13*, and *TRIM37*. Rodriguez-Casanova et al. identified the promoter hypermethylation of *WNT1* in cfDNA as a potential noninvasive biomarker for luminal B breast cancer that supported the application of Infinium MethylationEpic array to identify new epigenetic noninvasive biomarkers in breast cancer.

Based on the combined RRBS DNA methylome and transcriptome, Huang et al. performed a genome-wide comparison of DNA methylation and gene expression in *Clostridium perfringens* (Cp) type C-infected resistant and susceptible piglets. Such integrative analysis identified 168, 198, and 7 mRNAs, showing inverse correlations between methylation and expression with Cp infection, and revealed that the differentially expressed (DE) genes *LBP*, *TBX21*, and *LCN2* were likely involved in the piglets against Cp infection.

As microRNA (miRNA) plays a key role in gene regulation, Li et al. found that miR-208b expression increased in C2C12 cells but Mettl8 expression decreased significantly, while Myh4 expression decreased and Myh7 expression increased. Zou et al. identified the common targets and the transcript levels of miR-223-3p, miR-122-5p, and miR-93-5p in polycystic ovarian syndrome (PCOS) rat ovaries.

For circulating miRNAs, miR-1-3p participates in myocardial apoptosis, and its upregulation in circulation is a direct and powerful indicator of fetal ventricular septal defect (Yang Y. et al.). Ma et al. reported the expression of circular RNA (circRNA) circ_0059706 in *de novo* acute myeloid leukemia and its association with prognosis. In animals, circRNAs may interact with miRNAs to further regulate mRNA to regulate sperm motility in Yili geese, including 20 circRNAs, 18 miRNAs, and 177 mRNAs targeting ppy-mir-16, hsa-mir-221-3p, gga-mir-499-5p, etc (Wu et al.).

The long non-coding RNAs (lncRNAs) are engaged in vital biological regulatory processes. Jiang et al. established a prognostic risk model with 10 lncRNAs and obtained a good predictive

accuracy for overall survival of breast cancer individuals in both training and validation cohorts. In cows, Yang J. et al. detected 287 DE genes and 70 DE lncRNAs, where lncRNAs adjacent to the somatic cells count and somatic cell score QTLs influenced the mastitis pathogenesis by upregulating the expression of *TLR4*, *NOD2*, *CXCL8*, and *OAS2* genes.

As a dynamic and reversible RNA modification, N6-methyladenine (m^6A) is involved in a wide range of biological and pathological processes. Gu et al. identified three different m^6A sub-types including 27 samples in sub-type C1, 21 samples in sub-type C2, and 58 samples in sub-type C3. Li et al. identified 1,565 upregulated and 542 downregulated m^6A methylation peaks with significant changes.

Histone post-translational modification is an essential epigenetic process controlling a variety of biological activities. Xie et al. hypothesized that lactylation of histones or non-histone appeared to engage in various biochemical processes to influence the biological reactions of the organism, when lactate reaches a specific level under a certain circumstance.

In summary, our Research Topic gathers the findings of identified epigenetic biomarkers (methylated genes, miRNAs, circRNAs, lncRNAs, and m^6A) and reveals biological mechanisms using epigenomics data that could be used for further relevant studies.

Author contributions

XW conceived the idea for this Research Topic. As guest editors, XW, LF, and YY invited authors and supervised manuscript review. ZW coordinated this Research Topic. XW and ZW wrote the editorial with input from LF and YY. All authors contributed to the article and approved the submitted version.

Acknowledgments

The authors would like to thank Dan Hao of the University of Copenhagen for her assistance in designing the theme of this Research Topic.

Conflict of interest

Author XW was employed by Konge Larsen ApS.

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

Publisher's note

References

Yang, J., Tang, Y., Liu, X., Zhang, J., Zahoor, Khan M., Mi, S., Wang, C., and Yu, Y. (2022). Characterization of peripheral white blood cells transcriptome to unravel the regulatory signatures of bovine subclinical mastitis resistance. Front. Genet. 13:949850. doi:10.3389/fgene.2022.949850

Yang, Y., Yang, H., Lian, X., Yang, S., Shen, H., Wu, S., Wang, X., and Lyu, G. (2022). Circulating microRNA: Myocardium-derived prenatal biomarker of ventricular septal defects. Front. Genet. 13:899034. doi:10.3389/fgene.2022.899034

frontiers in Genetics

ORIGINAL RESEARCH
published: 23 February 2022
doi: 10.3389/fgene.2022.820464

MiR-208b Regulates the Conversion of Skeletal Muscle Fiber Types by Inhibiting Mettl8 Expression

Xiang Li†, *Hanfang Bi*†, *Shanshan Xie and Wentao Cui**

Institute of Animal Sciences, Chinese Academy of Agricultural Sciences, Beijing, China

Skeletal muscle, the main source of animal meat products, contains muscle fiber as a key unit. It is well known that transformation takes place between different types of muscle fibers, however, the conversion mechanism is not clear. In a previous study, our lab has demonstrated that there is a decrease in type I muscle fibers and an increase in type IIB muscle fibers in skeletal muscle of myostatin gene-edited Meishan pigs. Very interestingly, we observed the down regulation of miR-208b expression and an increase in expression the predicted target gene Mettl8 (Methyltransferase like 8) in skeletal muscle of MSTN gene-edited Meishan pigs. These results reveal that there is a potential connection between the conversion of skeletal muscle fiber types and miR-208b and Mettl8 expression. In this study, we first explored the expression patterns of miR-208b and Mettl8 in skeletal muscle in Meishan pigs; and then C2C12 cells were used to simulate the development and maturation of muscle fibers. Our results indicated that Myh4 expression level decreased and Myh7 expression level increased following overexpression of miR-208b in C2C12 cells. We therefore speculate that miR-208b can promote the conversion of fast-twitch fibers to slow-twitch fibers. The targeting relationship between Mettl8 and miR-208b was confirmed by results obtained using dual luciferase assay, RT-qPCR, and WB analysis. Following the transfection of Mettl8 siRNA into C2C12 cells, we observed that Mettl8 expression decreased significantly while Myh7 expression increased and Myh4 expression decreased, indicating that Mettl8 promotes the conversion of slow muscle fibers to fast muscle fibers. Additionally, changes in skeletal muscle fiber types are observed in those mice where miR-208b and Mettl8 genes are knocked out. The miR-208b knockout inhibits the formation of slow muscle fibers, and the Mettl8 knockout inhibits the formation of fast muscle fibers. In conclusion, our research results show that miR-208b regulates the conversion of different muscle fiber types by inhibiting Mettl8 expression.

Keywords: skeletal muscle, muscle fiber type conversion, miR-208b, Mettl8, miRNA, myostatin

OPEN ACCESS

Edited by:
Seyed Javad Mowla, Tarbiat Modares University, Iran

Reviewed by:
Kejun Wang, Henan Agricultural University, China
Marcio Duarte, University of Guelph, Canada

***Correspondence:**
Wentao Cui
cuiwentao@caas.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: *23 November 2021*
Accepted: *09 February 2022*
Published: *23 February 2022*

Citation:
Li X, Bi H, Xie S and Cui W (2022) MiR-208b Regulates the Conversion of Skeletal Muscle Fiber Types by Inhibiting Mettl8 Expression. Front. Genet. 13:820464. doi: 10.3389/fgene.2022.820464

INTRODUCTION

Skeletal muscle accounts for 30–40% of body weight in mammals. It is the main source of meat products for meat-producing animals and also an important metabolic organ (Doran et al., 2009; Listrat et al., 2016; Abel, 2018). Skeletal muscle dysfunction is related to a variety of muscle diseases, such as sarcopenia, muscle hypertrophy, amyotrophic lateral sclerosis, and muscle atrophy in

diabetic patients (Fontelonga et al., 2019; Garneau and Aguer, 2019; Deldicque, 2020). There are four types of skeletal muscle fibers: slow contraction oxidative metabolism type (I), fast contraction oxidative metabolism type (IIA), fast contraction glycolytic metabolism type (IIB), and fast contraction oxidative glycolysis facultative metabolism type (IIX). Different composition of muscle fiber types in skeletal muscle is a key factor that determines meat quality (Choi et al., 2007). For example, a high proportion of slow-oxidized muscle fibers improves meat tenderness, juiciness, and color (Lee et al., 2010).

Muscle-related genes such as myogenic regulatory factors (MRFs) and myostatin (MSTN) play important role in regulation of composition of different muscle fiber types. Hennebry et al. (Hennebry et al., 2009) demonstrated that loss of MSTN through gene knockout method resulted in an increase in type IIB muscle fibers and a decrease in type IIA and type I muscle fibers in mouse tibial anterior muscle. In recent years, non-coding RNAs have been shown to be involved in the regulation of muscle fiber transformation. Long-chain non-coding RNAs lnc-Six1 (Tieland et al., 2018) and lnc-mg (Du et al., 2019) can be used as miRNA molecular sponges to indirectly affect the expression of related genes, thereby regulating the transformation between fast and slow muscle fibers. In addition, miRNAs also play a key role in the regulatory network of muscle fiber typing by regulating the expression of target genes.

Micro RNAs (miRNAs) are a class of highly conserved single-stranded, non-coding RNAs with a length of 21–24 base pairs (Bhat et al., 2016; Lu and Rothenberg, 2018). MiRNAs play their regulatory roles by directly degrading target genes or by inhibiting target gene translation. MiR-208b, a member of miR-208 family, is encoded by the introns of the β-cardiac myosin heavy chain protein gene Myh7 and is specifically expressed in myocardium and skeletal muscle (Van Rooij et al., 2008). Recent studies on miR-208b focus mainly on heart disease, but miR-208b's specific mechanism of action to regulate skeletal muscle growth and development is still unclear. Additionally, Methyltransferase like 8 (Mettl8) is a member of the methyltransferase-like protein family (Badri et al., 2008; Tobi et al., 2018), which is mainly involved in cell differentiation, cell component formation, protein metabolism, and phylogenetic processes. Although Mettl8 is widely expressed, there is no report on the specific role of Mettl8 in skeletal muscle growth and development.

Our lab recently produced genetically engineered Meishan pigs containing a ZFN-edited myostatin loss-of-function MKO mutation ($MSTN^{-/-}$) that led to the hypertrophy of skeletal muscles (Qian et al., 2015). Analysis of deep miRNA sequencing data from skeletal muscle predicated that the expression of miR-208b was down-regulated and expression of miR-208b's target gene Mettl8 was up-regulated in $MSTN^{-/-}$ pigs (Xie et al., 2019; Li et al., 2020), implying that miR-208b and Mettl8 may play roles in skeletal muscle growth and development and regulating function of MSTN. We also measured the changes in the muscle fiber composition of the longissimus dorsi muscle in $MSTN^{-/-}$ Meishan pigs. Results indicated that type I fibers were reduced while type IIB fibers increased (Qian et al., 2015), suggesting that the changes in skeletal muscle fiber types in Meishan pigs induced by MSTN knockout may be related to miR-208b and its target gene Mettl8. Therefore, in this study, we first analyzed the expression patterns of miR-208b and Mettl8 in Meishan pigs, and then analyzed the effect of miR-208b on MyHC gene expression during the differentiation process of C2C12 cells by either overexpressing or inhibiting miR-208b. Bioinformatics analysis methods, dual luciferase reporter assay system and RNA interference experiment were used to predict and verify the target gene Mettl8 of miR-208b, and initially explore the effect of Mettl8 on the skeletal muscle fiber transformation process. Finally, we generated miR-208b and Mettl8 gene knockout mice to further analyze its function in the process of skeletal muscle fiber transformation in mice. Our research on the law of muscle fiber transformation may provide a new direction for the selection of excellent varieties and the treatment of skeletal muscle diseases.

MATERIALS AND METHODS

Animals

Both MKO ($MSTN^{-/-}$) and MWT (MSTN wild type, $MSTN^{+/+}$) Meishan pigs were produced using zinc finger nucleases (ZFN) and somatic cell nucleus transfer (SCNT) techniques (Qian et al., 2015). All pigs were maintained in Qingdao animal facility, fed with the same standard diet, and raised under the same conditions. 65 days of embryo development, 4-month old, 8-month-old, 16-month-old male pigs were used in this study. The gene-edited C57BL/6 mice were prepared by the Institute of Zoology, Chinese Academy of Sciences using CRISPR/Cas9 technology and microinjection technology. All experimental mice were 8 weeks old and weighed about 22 g. All samples were quickly collected after the animals were euthanized. The tissues used to extract total RNA or protein were quickly frozen in liquid nitrogen, and the tissues used for paraffin sections were immersed in 4% tissue fixative.

Cell Culture and Transfection

The HEK293T cells used for dual luciferase reporter assay were preserved by our laboratory. C2C12 myoblast cells were purchased from the Cell Resource Center in IBMS in CAMS/PUMC. The cells were cultured in growth medium consisting of Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% FBS and 1% penicillin/streptomycin. Myogenic differentiation was induced by replacing differentiation medium (DMEM supplemented with 2% horse serum and 1% penicillin/streptomycin) when cell confluence reached 60–70%.

Transfection was performed with the Lipofectamine 2000 reagent (Invitrogen) combined with 50 nM miRNA mimics or corresponding control, 100 nM miRNA inhibitor or corresponding control, 50 nM Mettl8 siRNA or corresponding control when cell confluence reached 70–80%.

Plasmids Construction

The region of Mettl8 3′ UTR flanking the miR-208b binding site was amplified from mouse genomic DNA using PCR. The target

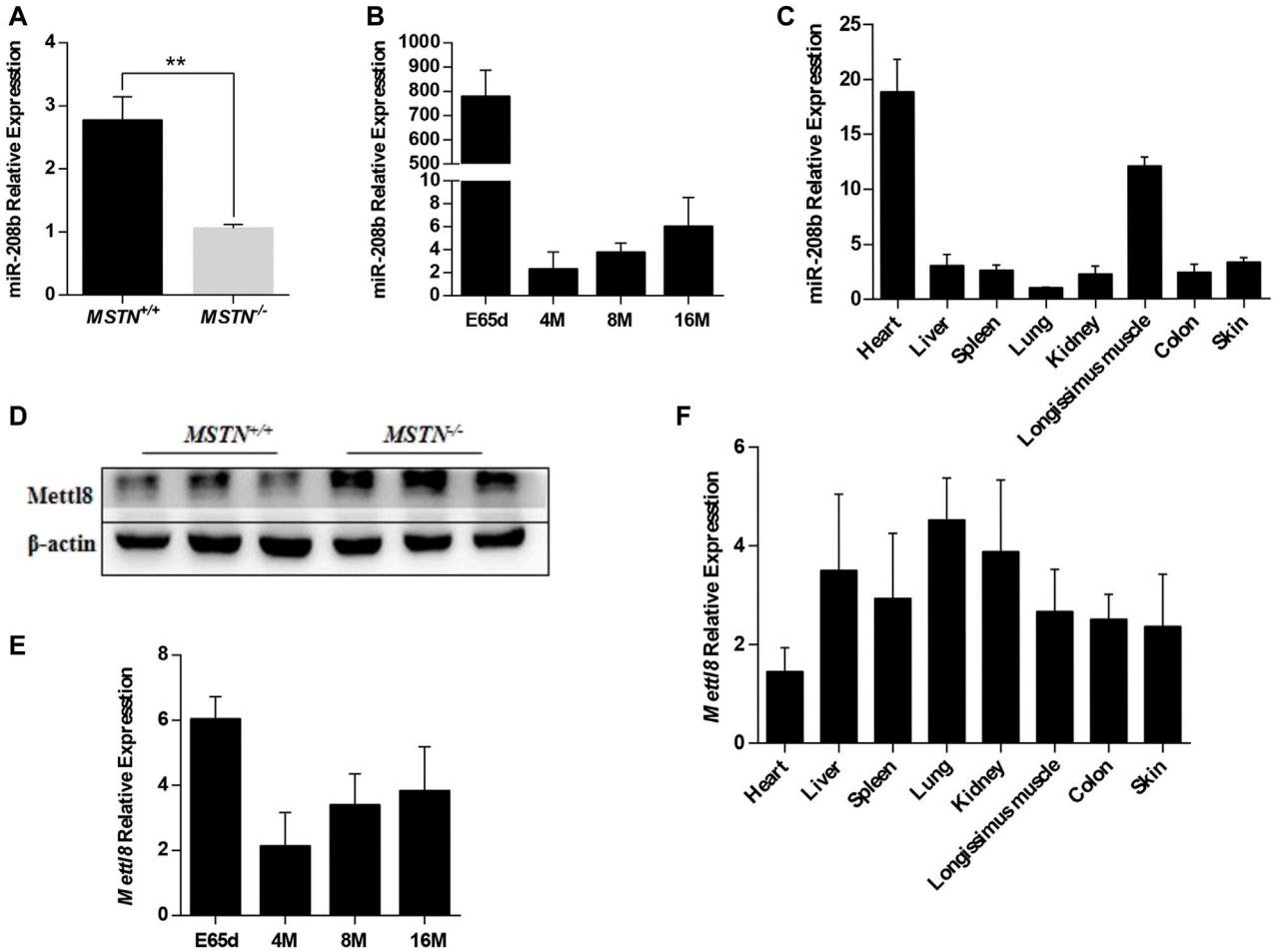


FIGURE 1 | Expression pattern of miR-208b and Mettl8 in Meishan pigs. **(A)** Expression level of miR-208b in the longissimus dorsi from wild type ($MSTN^{+/+}$) and $MSTN^{-/-}$ Meishan pigs on E65d. **(B)** Expression level of miR-208b in the longissimus dorsi from wild type ($MSTN^{+/+}$) Meishan pigs at different ages. **(C)** Expression level of miR-208b in different tissues from wild type ($MSTN^{+/+}$) Meishan pigs on E65d. **(D)** Expression level of Mettl8 protein in the longissimus dorsi muscle from wild type ($MSTN^{+/+}$) and $MSTN^{-/-}$ Meishan pigs on E65d. **(E)** Expression level of Mettl8 protein in the longissimus dorsi muscle from wild type ($MSTN^{+/+}$) at different ages. **(F)** Expression level of Mettl8 protein in different tissues from wild type ($MSTN^{+/+}$) on E65. E65d means pigs on day 65 of embryonic development, 4, 8, 16 M indicate 4, 8, and16 months after birth, respectively. Statistical comparisons are performed using unpaired *t*-tests in **(A,B)** (The data of 4, 8 and 16 M are compared with the data of E65d, respectively), **(E)** (The data of 4, 8 and 16 M are compared with the data of E65d, respectively)and using ordinary one-way ANOVA in **(C,F)**. $^{**}p < 0.01$.

sequence GGGAGCT (800–806 bp) was mutated to CCCTCGA using overlap PCR. Primers were showed in **Supplementary Table S1**. The PCR product was cloned into the vector downstream of the Renilla Luciferase open reading frame using the NotI and XhoI restriction sites. We obtained two pmiR-RB-REPORT vectors (RiboBio) with wide-type and mutant 3′ UTR of Mettl8.

Dual Luciferase Reporter Assay

HEK293T cells were co-transfected with 100 ng of the wide-type or mutant 3′UTR luciferase reporter and 50 nM of the miR-208b mimics or control duplexes using the Lipofectamine 2000 reagent (Invitrogen) in 48-well plates. After transfection for 48 h, cells were harvested by adding 300 μL of a passive lysis buffer. Renilla and firefly luciferase activities were measured with the Dual Luciferase Assay System (Promega, Madison, WI) in a TD-20/20 luminometer (Turner Biosystems, Sunnyvale, CA), and the Renilla luciferase signal was normalized to the firefly luciferase signal. The normalized Renilla luciferase activity was compared with the wild type, miR-208b, and the mutant using the Student's *t* test ($p < 0.05$).

RNA Isolation and Quantitative Real-Time PCR

Total RNA was extracted from muscle samples by using TRIzol (Invitrogen) per manufacturer's instructions. RNA quality was assessed by using the RNA Nano 6000 Assay Kit of the Bioanalyzer 2100 system (Agilent Technologies, CA, United States), agarose gel electrophoresis and NanoDrop (Thermo Fisher). Each sample (1 μg) was reverse transcribed into cDNA by using the RevertAid^{TM} First Strand cDNA

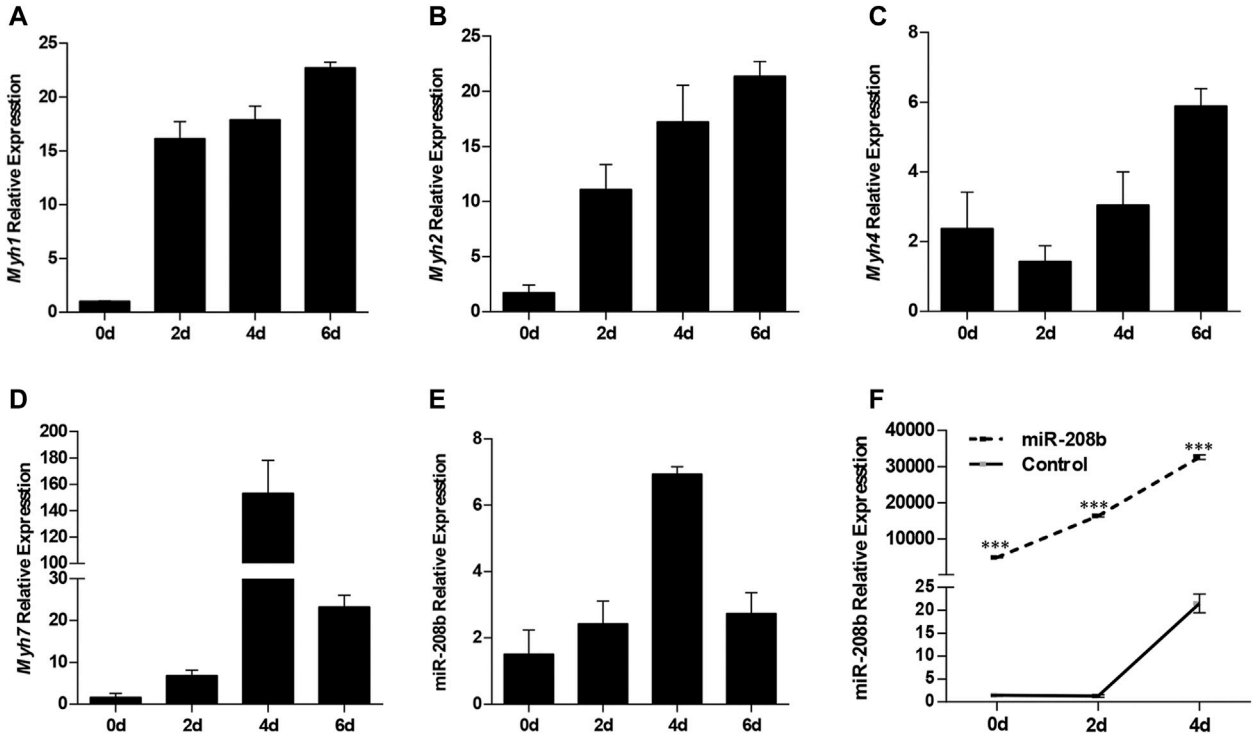


FIGURE 2 | The expression patterns of MyHC and miR-208b during of C2C12 cell myogenic differentiation. **(A–E)**: Expression levels of Myh1 **(A)**, Myh2 **(B)**, Myh4 **(C)**, Myh7 **(D)**, and miR-208b **(E)** during C2C12 myogenic differentiation on days 0, 2, 4, and 6. **(F)** Relative expression level of miR-208b during C2C12 myogenic differentiation on days 0, 2, and 4 post transfections with miR-208b mimics and control vectors, respectively. Statistical comparisons are performed using ordinary one-way ANOVA in **(A–E)** and using multiple *t*-tests in **(F)**.

Synthesis Kit (Fermentas). Real-time PCR was performed in Quant Studio 3 system (Thermo Fisher ABI) using 10 pM of each specific primer (**Supplementary Table S1**) and SYBR Premix ExTagTM (Takara) according to the manufacturer's protocols. The $2^{-\Delta\Delta Ct}$ method was employed to calculate the relative expression levels of mRNAs (Livak and Schmittgen, 2001).

Immunoblotting

For protein extraction, 2 mL of lysis buffer [8 mol/L urea, 2% SDS, × 1 Protease Inhibitor Cocktail (Roche Ltd. Basel, Switzerland)] was added to each sample, followed by sonication on ice and centrifugation at 13,000 rpm for 10 min at 4°C. The supernatant was then transferred to a fresh tube. Total protein concentration was determined using a BCA Quantitative Test Kit (Beyotime). Each protein sample was loaded in equal amount and then separated by 10% or 12.5% SDS PAGE. Following transfer of protein from gel to nitrocellulose (NC) membrane and blocked with 5% skimmed milk for 2 h, immunoblotting was performed using standard method for the following proteins with corresponding detection antibodies: Mettl8 (Rabbit polyclonal antibody, Biorbyt), Myh4 (Monoclonal Antibody, Thermo Scientific), Myh7 (Rabbit polyclonal antibody, Santa Cruz), myosin light chain 9 (MYL9) (Rabbit polyclonal antibody, Abcam), beta-tublin (Rabbit polyclonal antibody, Cell Signaling Technology), beta-actin (Rabbit polyclonal antibody, Abcam). Beta-tublin and beta-actin were used as an internal reference in Western blot. Super Signal West Pico chemiluminescent substrate (Thermo Fisher Scientific) was used to develop color band. Image J software was used to analyze protein band density. Graphd Prism 8.0 software was used to make histograms using protein band density data.

Haematoxylin and Eosin (HE) Staining

The gastrocnemius (GAS) and soleus muscle (SOL) were dissected at the time when mice were euthanized, fixed in 4% paraformaldehyde, and embedded in paraffin. Muscle sections were stained with hematoxylin and eosin, and pictures were taken from four random fields at ×40 magnification. HE staining method are the same as previously described (Cai et al., 2017).

Statistical Analysis

The data of all experimental groups and control groups were statistically analyzed using the Analyze procedure in Graphd Prism 8.0 software. All data were expressed as average ±standard deviation. Unpaired *t*-test, multiple *t*-test (Multiple comparisons between two samples)or one-way variance test (comparison between multiple groups) was used to identify if there was differential expression. $p < 0.05$ is considered statistically significant.

RESULTS

Expression Patterns of miR-208b and Mettl8 in Meishan Pigs

We have found that Mettl8 and miR-208b were related to each other by multi-omics analysis of MSTN gene-edited Meishan pig skeletal muscle (Xie et al., 2019; Li et al., 2020). Sequencing results showed that compared with $MSTN^{+/+}$ group, the expression level of miR-208b was significantly down-regulated while the expression level of Mettl8, was significantly up-regulated in $MSTN^{-/-}$ group (Xie et al., 2019; Li et al., 2020). In this study, the results of RT-qPCR and Western blot were consistent with previous sequencing results (**Figures 1A,D**). We tested miR-208b expression profiles during growth different stages and in different tissues in Meishan pigs (**Figures 1B,C**). The results showed that miR-208b had the highest expression in skeletal muscle on 65 days of embryo development but decreased post birth. It is speculated that miR-208b mainly plays a role in the embryonic stage of Meishan pigs. The expression level of miR-208b is the highest in the myocardium and skeletal muscle of Meishan pigs, indicating that miR-208b is a muscle-specific miRNA and plays its role mainly in the striated muscle of Meishan pigs. We also further tested Mettl8 expression profiles during growth different stages and in different tissues in Meishan pigs (**Figures 1E,F**). The results showed that Mettl8 expression level in the longissimus dorsi of Meishan pigs is higher on 65 days of embryo development in other growth periods. Although Mettl8 is widely expressed in a variety of tissues, its expression in myocardium and skeletal muscle is relatively low, indicating that it may play an opposite role when compared to miR-208b.

MiR-208b Promotes the Conversion of Fast Muscle Fibers to Slow Muscle Fibers During C2C12 Myogenic Differentiation

C2C12 cells were used as a model of myogenic differentiation to simulate the process of muscle fiber formation *in vitro*. Changes in expression levels of different muscle fiber marker genes MyHC and miR-208b during myogenic differentiation were monitored and detected (**Figures 2A–E**). To clarify the specific functions of miR-208b in the transformation of different types of muscle fibers, miR-208b mimics were transfected into C2C12 myoblasts (**Figure 2F**). After successfully overexpressing miR-208b, myogenic differentiation of C2C12 cells was induced and changes in the expression levels of different types of MyHC at the mRNA level were detected (**Figures 3A–D**). Results showed that after increasing the expression of miR-208b, the expression of the slow muscle marker gene Myh7 increased, while the expression of the fast muscle marker gene Myh4 decreased significantly. The expression of intermediate muscle fiber marker genes Myh1 and Myh2 also showed a decreasing trend. On the other hand, following the transfection with miR-208b inhibitors, the opposite results were obtained (**Figures 3F–J**), further indicating that miR-208b can regulate the conversion of fast-twitch fibers to slow-twitch fibers. At the same time, we also observed changes in protein expression of Myh4, Myh7, and MYL9 which is enriched in fast-twitch fibers. The results showed that the expression of Myh4 protein decreased significantly, while the expression of Myh7 protein increased significantly, which is consistent with the transcription level results (**Figure 3E**). The expression level of MYL9 was significantly reduced (**Figure 3E**), indicating that the overexpression of miR-208b may affect the expression of genes related to fast-twitch fiber formation.

MiR-208b Inhibits Expression of Target Gene Mettl8 by Binding to 3′UTR

We used RNAhybird (Rehmsmeier et al., 2004) to predict the binding site of miR-208b in the target gene Mettl8 (**Figure 4A**). To verify the targeting relationship between miR-208b and Mettl8, miR-208b mimics and luciferase vectors were co-transfected in HEK293T cells. The luciferase vectors contained wild type Mettl8 3′UTR sequence or sequence with a point mutation at binding site. We observed that compared with negative control, the luciferase activity of the cells co-transfected with miR-208b and Mettl8 3′UTR sequence was significantly decreased (**Figure 4B**), while the luciferase activity in the cells transfected with mutant sequence did not change significantly (**Figure 4B**), indicating that miR-208b can target to bind Mettl8 and inhibit its expression. We then further explored the effect of overexpressing miR-208b on mRNA level and the protein level of Mettl8 in C2C12 cells. RT-qPCR results showed that after overexpression of miR-208b, the expression of Mettl8 mRNA level in C2C12 cells was not significantly reduced [(**Figure 4C**), but the protein level was significantly reduced (**Figures 4D,E**)], indicating that miR-208b targeted to Mettl8 3′UTR sequence by incomplete complementary pairing to inhibit its protein translation process.

Mettl8 Promotes the Conversion of Slow Muscle Fibers to Fast Muscle Fibers in the Process of C2C12 Myogenic Differentiation

Once the targeting relationship between miR-208b and Mettl8 was determined, we then monitored the expression pattern of Mettl8 during myogenic differentiation (**Figure 5A**). To further explore the effect of Mettl8 on muscle fiber transformation, RNAi technology was used to knock down the expression of C2C12 endogenous Mettl8 protein (**Figures 5B,C**), followed by using RT-qPCR to analyze the expression of different types of muscle fiber type marker genes MyHC. The test results showed (**Figures 5D–G**) that after down regulated Mettl8 expression, Myh4 expression was significantly down-regulated on the second day of differentiation, Myh7 expression was significantly up-regulated on 2–4 days of differentiation. Although Myh1 and Myh2 were downregulated within 4 days of differentiation. To further confirm the translation level, Western blot was used to detect the protein expression of Myh4, Myh7 and MYL9 in the cells on the 4th day of differentiation. The results showed (**Figure 5H**) that the expression level of Myh4 and MYL9 was significantly reduced while expression level of Myh7 increased significantly following the inhibition of Mettl8 protein expression. These results are consistent with those obtained with transfection of miR-208b mimics and further demonstrated that Mettl8 promoted the conversion of slow

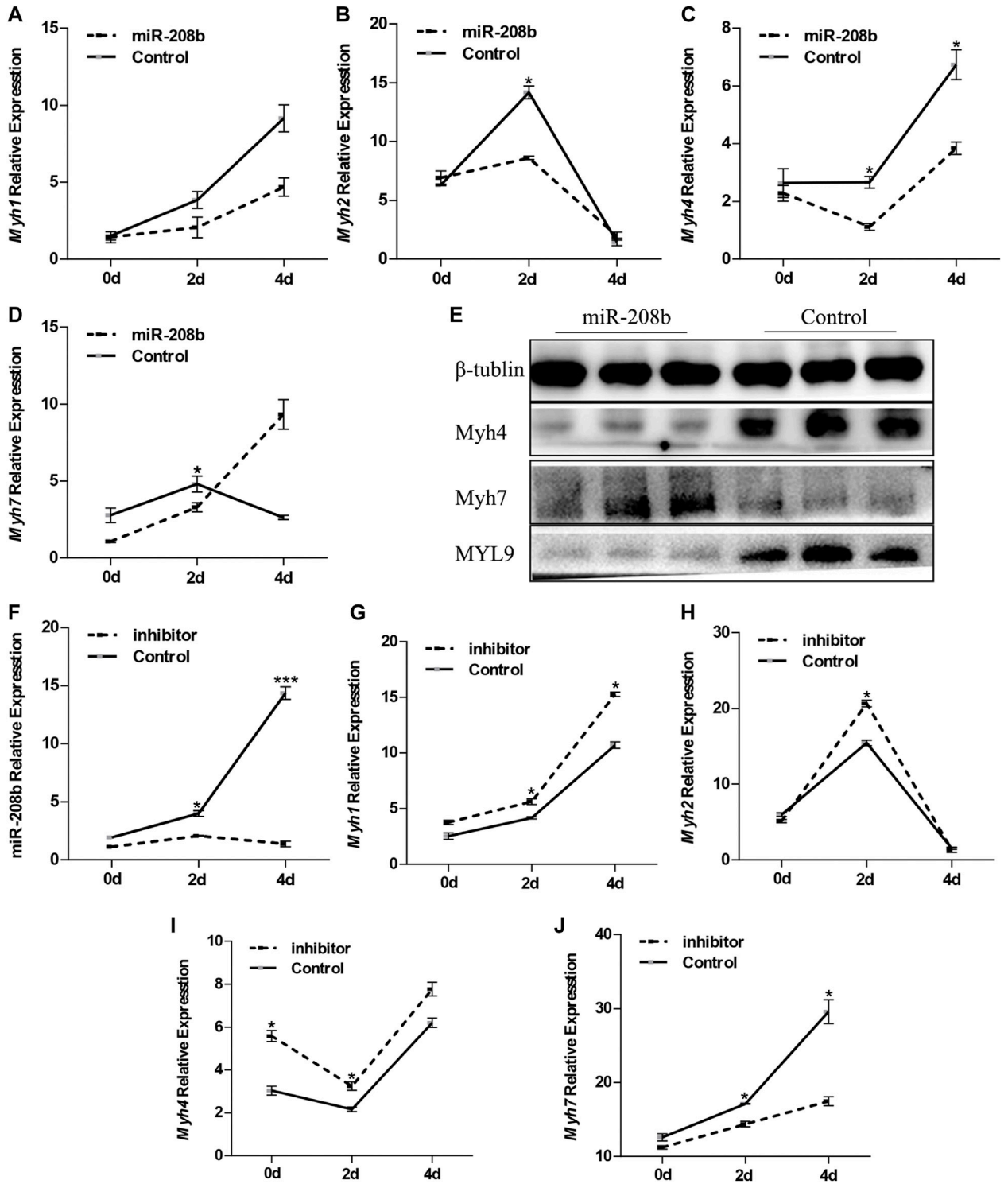


FIGURE 3 | Effect of miR-208b on MyHC gene expression during myogenic differentiation. **(A–D)**: relative mRNA expression levels of Myh1 **(A)**, Myh2 **(B)**, Myh4 **(C)**, and Myh7 **(D)** in C2C12 cells overexpressing miR-208b on days 0, 2, 4 of myogenic differentiation. **(E)** protein expression levels of Myh4, Myh7, MYL9 post 4 days of miR-208b over-expression during C2C12 myogenic differentiation. miR-208b represents the transfection of miR-208b mimics to achieve the effect of over-expression of miR-208b. **(F)** Relative expression level of MyHC after transfection of 100 nM miR-208b inhibitor and control on days 0, 2, 4 during differentiation. **(G–J)** respectively indicate relative expression levels of Myh1 **(G)**, Myh2 **(H)**, Myh4 **(I)**, and Myh7 **(J)** on 0, 2, and 4 days of C2C12 myogenic differentiation after miR-208b expression is inhibited. Statistical comparisons are performed using multiple t-tests.*$p < 0.05$.

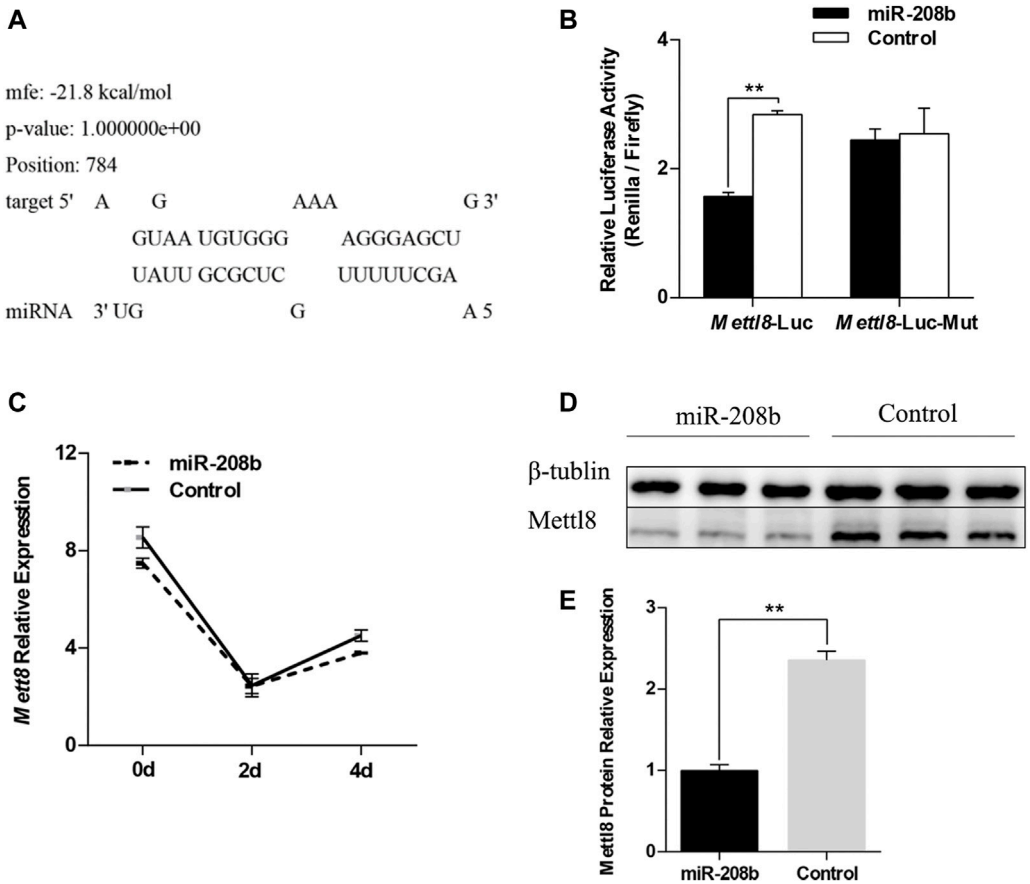


FIGURE 4 | Verification of the targeting relationship between miR-208b and Mettl8. **(A)** The binding site of miR-208b and Mettl8 3′UTR sequence. **(B)** Results of dual luciferase reporter assay. Mettl8-Luc: wild-type plasmid; Mett18-Luc-Mut: Mettl8 mutant plasmid. The ordinate is the relative luciferase activity (renilla luciferase activity/firefly luciferase activity). MiR-208b represents C2C12 cells transfected with miR-208b mimics to achieve the effect of miR-208b over-expression. **(C)** Changes in Mettl8 mRNA expression level after miR-208b overexpression in C2C12 cells during myogenic differentiation. **(D)** Western blot results of Mettl8 protein after miR-208b overexpression. **(E)** Relative expression level of Mettl8 protein calculated using Image J software. Statistical comparisons are performed using multiple *t*-tests in **(B–C)** and using unpaired *t*-tests in **(E)**. $^{**}p < 0.01$.

muscle fibers to fast muscle fibers during the process of C2C12 myogenic differentiation.

Effect of MiR-208b and Mettl8 on the Transformation of Different Types of Muscle Fibers in Mouse Skeletal Muscle

To further explore the functions of miR-208b and Mettl8 *in vivo*, we produced miR-208b and Mettl8 gene knockout mice using CRISPR/Cas9 technology to examine if miR-208b and Mettl8 gene knockout has any effect on muscle fiber morphology. We isolated gastrocnemius and soleus muscles from 8-week-old knockout mice and wild-type mice and performed HE staining. It was observed that, compared with wild type mice, gastrocnemius muscle fibers became larger in miR-208b knockout mice (**Figure 6A**). On the other hand, there was no significant changes in soleus muscle, which contained more slow muscle fibers (**Figure 6A**). Compared with wild-type mice, muscle fibers in gastrocnemius muscle of Mettl8 gene knockout mice became smaller, indicating that the muscle fibers may be transformed into thinner and longer slow muscle fibers (**Figure 6B**). Like observed in miR-208b knockout mice, there is no significant changes in muscle fibers of soleus muscle in Mettl8 knockout mice (**Figure 6B**).

We also used RT-qPCR to determine the expression levels of different muscle fiber marker genes MyHC in mouse skeletal muscle. The results showed (**Figures 6C–E**) that, compared with wild-type mice, the expression level of Myh7 in GAS, TA, and QUA decreased significantly. On the other hand, the expression level of the target gene Mettl8 increased in all three skeletal muscles, being significantly up-regulated in GAS. The expression of Myh1 increased in all three skeletal muscles, reaching to significant level in TA and QUA. The expression of Myh2 decreased in GAS and TA muscle but increased in QUA muscle, reaching a significant decrease level in TA. Myh4 expression increased in all three skeletal muscles, but its change is significant in GAS. Compared with wild-type mice, the expression levels of Myh7 in GAS, TA and QUA increased

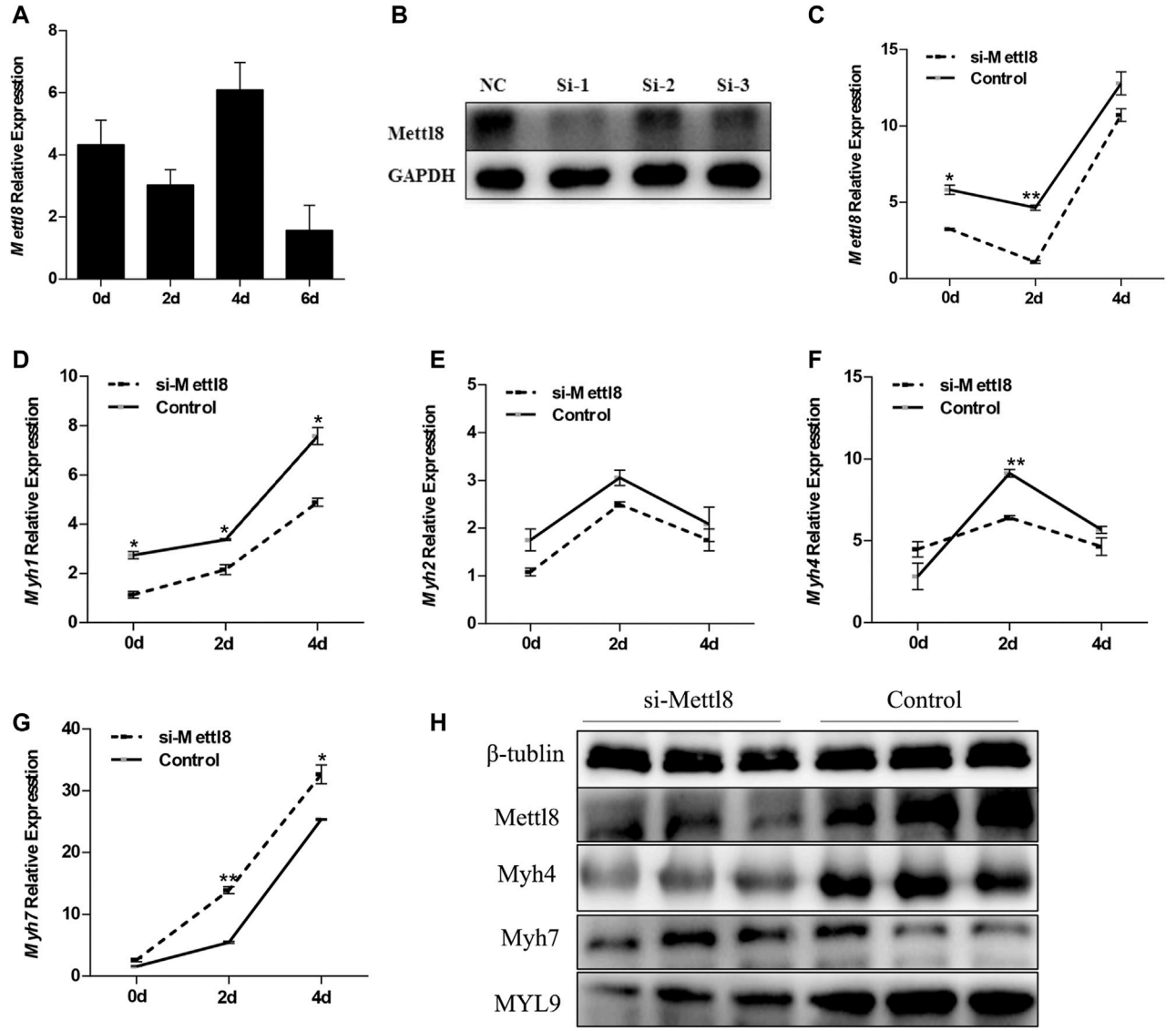


FIGURE 5 | The effect of knocking down Mettl8 on MyHC expression in C2C12 cells during myogenic differentiation. **(A)** The expression of Mettl8 in the process of myogenic differentiation of C2C12 cells. **(B)** The change of Mettl8 protein expression level after transfection of siRNA, NC represents negative control, si-1, si-2, and si-3 represent the three designed Mettl8 interfering RNA. **(C)** Changes in Mettl8 mRNA expression after transfection of siRNA. **(D–G)**: Expression levels of Myh1 **(D)**, Myh2 **(E)**, Myh4 **(F)**, and Myh7 **(G)** in the endogenously Mettl8 knockdown C2C12 cells during the myogenic differentiation on days 0, 2, 4. **(H)** Protein expression levels of Mettl8, Myh4, Myh7, MYL9 in the endogenously Mettl8 knockdown C2C12 cells. si-Mettl8: C2C12 cells transfected with 50 nM Mettl8 interfering RNA to reduce the expression of endogenous Mettl8. Statistical comparisons are performed using ordinary one-way ANOVA in **(A)** and using multiple *t*-tests in **(D–G)**. $^{*}p < 0.05$; $^{**}p < 0.01$.

significantly in Mettl8 gene knockout mice. The expression level of Myh1 in these three skeletal muscles from Mettl8 gene knockout mice decreased, but the decrease was not significant. The expression level of Myh2 increased in these three skeletal muscles from Mettl8 gene knockout mice, and the increase reached a significant level in QUA. The expression level of Myh4 was down-regulated in these three skeletal muscles from Mettl8 gene knockout mice, with significant changes in GAS (**Figures 6F–H**). The above experimental results prove that miR-208b and Mettl8 play the opposite effect in the conversion of fast and slow muscle fibers following their gene knocking out, but the degree of their influence is slightly different in different types of skeletal muscles, likely due to the different composition ratios of different types of skeletal muscles.

DISCUSSION

MiRNA is a short single-stranded RNA that does not encode protein. It has the function of regulating targeted gene expression at the translation level or transcription level. (Filipowicz, 2005; Pu et al., 2019). With the advancement of high-throughput

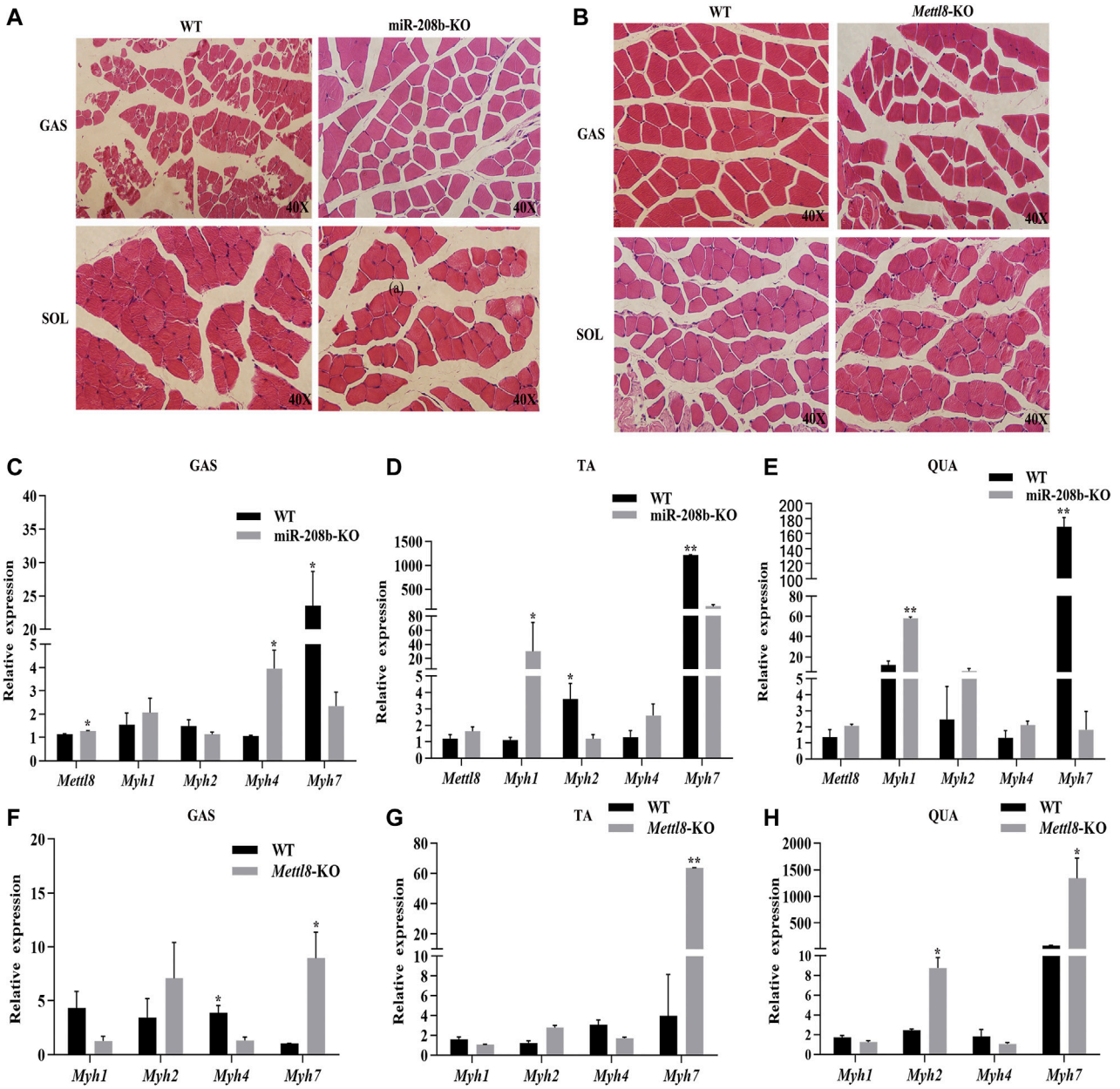


FIGURE 6 | Transformation of different skeletal muscle fiber types in miR-208b and Mettl8 knockout mice. **(A)** HE staining (40×) of skeletal muscle from miR-208b knockout mice and wild-type mice. **(B)** HE staining (40×) of skeletal muscle from Mettl8 knockout mice and wild-type mice. **(C–E)**: Expression levels of Mettl8 and MyHC in GAS **(C)**, TA **(D)**, and QUA **(E)** of miR-208b knockout mice and wild-type mice, respectively. **(F–H)**: Expression levels of Mettl8 and MyHC in GAS **(F)** TA **(G)** and QUA **(H)** of Mettl8 knockout mice and wild-type mice, respectively. WT: wild-type mice; miR-208b-KO: miR-208b gene knockout mice; Mettl8-KO: Mettl8 gene knockout mice; GAS: gastrocnemius muscle; SOL: soleus muscle; QUA: quadriceps muscle. Statistical comparisons are performed using multiple *t*-tests. $^{*}p < 0.05$; $^{**}p < 0.01$.

sequencing technology, more and more miRNAs have been identified to be involved in regulating a variety of biological processes, including skeletal muscle growth and development (O'rourke et al., 2007). MA et al. extracted RNAs from skeletal muscles such as peroneus longus, longissimus dorsi, and psoas major in pigs, and identified different types of differential miRNAs from different types of muscles through transcriptome sequencing (Ma et al., 2015). They speculated that miRNAs may be involved in the regulation of muscle fiber type transformation. Liu et al. observed that an individual knockout of miR-133a-1 or miR-133a-2 did not significantly affect the growth and development of mouse muscles, but simultaneous knockout of both miR-133a-1 and miR-133a-2 led to disease of fast-twitch fibers, mitochondrial damage, and the conversion of fast and slow-twitch fibers (Liu et al., 2008). Wang et al. noticed that the expression of miR-499 in the pig's extensor toe and soleus muscle was negatively correlated with the expression of the key transcription factor Sox6 (SRY-box

transcription factor 6) of fast-contracting muscle fibers (Wang et al., 2017). After overexpression of miR-499 in porcine muscle satellite cells, the expression of MyHC I and MyHC IIA mRNA increased. MSTN is a negative regulator of skeletal muscle growth and development. It is involved in complex cellular signaling pathways and inhibits myogenesis by regulating the expression of target genes and other molecular mechanisms (Cai et al., 2017), but the underlying molecular mechanism of MSTN functions has is still not fully understood yet. Our previous studies showed that the expression level of miR-208b decreased significantly in MKO Meishan pigs, indicating that miR-208b may be involved in regulating skeletal muscle growth and development. Our current study further confirmed the reliability of the previous miRNA sequencing results through molecular and biological experiments. Our results demonstrated that the expression level of miR-208b in the skeletal muscle of MKO Meishan pigs was significantly down-regulated, indicating that miR-208b is involved in the regulation of skeletal muscle growth and development. The proportion of slow-twitch fibers was higher on 65 days of embryonic development than post birth in pigs. Our results showed that miR-208b is highly expressed on 65 days of embryonic development, which suggested that miR-208b was involved in the formation of slow-twitch muscle fibers. We speculate that MSTN may act as an upstream regulator that affects the processing of miR-208b precursor sequence or the transcription of its coding sequence, which is different from the results of MSTN as a downstream target gene of miRNA reported in a previous report (Callis et al., 2009). Our results may provide a new insight to study the mechanism of MSTN action in skeletal muscle growth and development.

C2C12 myoblasts are often used as an *in vitro* model to study of skeletal muscle growth and development. Many previous studies have demonstrated that some regulatory factors, including miRNA, can affect the expression of different muscle fiber type marker genes during C2C12 cell myogenic differentiation. Xu et al. showed that, during the process of C2C12 myogenic differentiation, overexpression of miR-139-5p can down-regulate MyHC I and MyHC IIA by inhibiting the expression of CaN, NFATc1, MEF2C, and MCIP1.4 in the CaN/NFAT signaling pathway. On the other hand, inhibition of miR-139-5p expression led to the opposite result (Xu et al., 2018). Cheng et al. observed that miR-204-5p can significantly reduce the ratio of slow muscle fiber genes in myoblasts by targeting MEF2C and ERRγ with overexpressing or inhibiting the expression of miR-204-5p in C2C12 cells post induction of myogenic differentiation (Cheng et al., 2018). In our study, liposome transfection was used to transfect miR-208b mimics and inhibitors in C2C12 myoblasts and our results confirmed that miR-208b has an effect on the expression of different muscle fiber marker genes at mRNA and protein levels during the process of myogenic differentiation. In summary, our study indicates miR-208b can regulate the transformation of fast and slow muscle fibers.

The analysis of preliminary screening sequence and bioinformatical method along with results from dual luciferase experiment confirmed that Mettl8 is a target gene of miR-208. Mettl8 is a member of the methyltransferase-like protein family and is also a tension-inducing or inhibiting protein (TIP). Mettl8, as a TIP protein, has three subtypes, Tip1, Tip2, Tip3. And these three subtypes contain nucleic acid receptor co-regulators, histone acetyltransferase, and sequence characteristics of the chromatin remodeling factor of histone deacetylase (Heery et al., 1997; Boyer et al., 2002). Previous studies demonstrated that Tip1 and Tip3 are extremely sensitive to the tension of smooth muscle. Under tension, Tip1 promotes myogenic differentiation fate, and under the influence of tension inhibition, Tip3 promotes adipogenic differentiation (Jakkaraju et al., 2005). To date, there has been no report that Mettl8 can participate in the regulation of skeletal muscle growth and development related processes. To verify the biological function of Mettl8 in muscle fiber typing, we used C2C12 cells as a model to explore the changes in Mettl8 expression during myogenic differentiation and found that its expression pattern is similar to Myh4 expression. Then we further conducted an interference experiment with Mettl8 and confirmed that Mettl8 can indeed affect the expression of fast and slow muscle marker genes at both the transcription and protein levels.

We successfully constructed miR-208b gene knockout mice and compared with wild-type mice. In terms of phenotype, there is no significant change in the morphology of muscle fibers of soleus muscle except the fact that the cross-section of single muscle cell of gastrocnemius muscle became larger. However, at the molecular level, there is a significant change in the expression levels of different muscle fiber marker genes MyHC in the skeletal muscle of knockout mice has changed significantly. With the advancement of miRNA research in muscles, it is clear that some miRNAs are found to be specifically expressed in muscles. The muscle specific miRNAs are called MyomiR, such as miR-206, miR-208b and miR-499 that are enriched in type I muscle fibers. Most knocked out MyomiRs in mice have little effect on the phenotype of skeletal muscle (Zhao et al., 2007; Boettger et al., 2014). For example, skeletal muscle-specific knockout of miR-206 did not result in an obvious change in phenotype as evidenced by no significant changes in body weight, soleus muscle weight, or the morphology of muscle fibers (Williams et al., 2009). Additionally, we also generated Mettl8 knockout mice and compared with wild-type mice. Again, there is no obvious change in phenotype except for the increase of gastrocnemius fiber area. At the molecular level, the pattern of change in MyHC expression is just the opposite of miR-208b knockout. This clearly shows that Mettl8 is the target gene of miR-208b and has the opposite effect on muscle fiber transformation compared to miR-208b. We observed that miR-208b regulates the conversion of muscle fibers to slow muscle by targeting and inhibiting Mettl8 while Mettl8 can in turn affect the host gene Myh7 encoding miR-208b, thus forming a regulatory network which is similar to the regulatory network of miR-499 and its target gene Sox6 in a previous study using a mouse model of skeletal muscle atrophy (Mccarthy et al., 2009). Of course, there may be many regulatory factors and signal pathways similar to MSTN involved in this network, and the specific mechanism needs to be investigated in depth in the future.

In this study, we used MTSN-edited Meishan pigs as research animals to successfully confirm the previous sequencing analysis results of miR-208b and Mettl8 and then analyzed expression profiles of miR-208b and Mettl8 in skeletal muscle from different tissues during embryonic development and post birth. Then we

employed C2C12 myoblasts as a model to investigate the effect of overexpression or inhibition of miR-208b on the expression of different marker genes MyHC during the differentiation of myoblasts. Results with C2C12 myoblasts show that miR-208b can promote the production of slow muscle fibers. Use of bioinformatic analysis and dual luciferase experiments further verified the targeting relationship between miR-208b and Mettl8. Through the Mettl8 interference experiment in C2C12 cells, it was confirmed that Mettl8 can affect the differentiation of different marker genes MyHC during myoblast differentiation and its effect is just the opposite of miR-208b. Finally, the miR-208b and Mettl8 gene knockout mice were successfully generated and results from these knockout mice further demonstrated that miR-208b and Mettl8 play opposite roles in the transformation of fast and slow muscle fibers.

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.

ETHICS STATEMENT

The animal study was reviewed and approved by the Institutional Animal Care and Use Committee (IACUC) at Institute of Animal Sciences, Chinese Academy of Agricultural Sciences.

AUTHOR CONTRIBUTIONS

XL conducted most experiments, performed data analysis and wrote the manuscript. HB conducted part of molecular biology and animal experiments. SX contributed to the preparation of the gene-editing Meishan pig and performed some data analysis. WC designed this study, wrote and revised the manuscript.

FUNDING

This study was supported by National Basic Research Project (2015CB943100), National Transgenic Project of China (2014ZX08006-003, 2016ZX08006-001) and The Agricultural Science and Technology Innovation Program (ASTIP-IAS05).

ACKNOWLEDGMENTS

We thank the following: Shulin Yang, Yanfang Wang, Zhonglin Tang, Yulian Mu, and Hong Ao for their guidance and suggestions.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.820464/full#supplementary-material

REFERENCES

Abel, E. D. (2018). Mitochondrial Dynamics and Metabolic Regulation in Cardiac and Skeletal Muscle. *Trans. Am. Clin. Climatol Assoc.* 129, 266–278.

Badri, K. R., Zhou, Y., Dhru, U., Aramgam, S., and Schuger, L. (2008). Effects of the SANT Domain of Tension-Induced/inhibited Proteins (TIPs), Novel Partners of the Histone Acetyltransferase P300, on P300 Activity and TIP-6-Induced Adipogenesis. *Mol. Cel Biol* 28, 6358–6372. doi:10.1128/mcb.00333-08

Bhat, S. S., Jarmolowski, A., and Szweykowska-Kulinska, Z. (2016). MicroRNA Biogenesis: Epigenetic Modifications as Another Layer of Complexity to the microRNA Expression Regulation. *Acta Biochim. Pol.* 63, 717–723. doi:10.18388/abp.2016_1370

Boettger, T., Wüst, S., Nolte, H., and Braun, T. (2014). The miR-206/133b Cluster Is Dispensable for Development, Survival and Regeneration of Skeletal Muscle. *Skeletal Muscle* 4, 23. doi:10.1186/s13395-014-0023-5

Boyer, L. A., Langer, M. R., Crowley, K. A., Tan, S., Denu, J. M., and Peterson, C. L. (2002). Essential Role for the SANT Domain in the Functioning of Multiple Chromatin Remodeling Enzymes. *Mol. Cel* 10, 935–942. doi:10.1016/s1097-2765(02)00634-2

Cai, C., Qian, L., Jiang, S., Sun, Y., Wang, Q., Ma, D., et al. (2017). Loss-of-function Myostatin Mutation Increases Insulin Sensitivity and browning of white Fat in Meishan Pigs. *Oncotarget* 8, 34911–34922. doi:10.18632/oncotarget.16822

Callis, T. E., Pandya, K., Seok, H. Y., Tang, R.-H., Tatsuguchi, M., Huang, Z.-P., et al. (2009). MicroRNA-208a Is a Regulator of Cardiac Hypertrophy and Conduction in Mice. *J. Clin. Invest.* 119, 2772–2786. doi:10.1172/jci36154

Cheng, X., Du, J., Shen, L., Tan, Z., Jiang, D., Jiang, A., et al. (2018). MiR-204-5p Regulates C2C12 Myoblast Differentiation by Targeting MEF2C and ERRγ. *Biomed. Pharmacother.* 101, 528–535. doi:10.1016/j.biopha.2018.02.096

Choi, Y. M., Ryu, Y. C., and Kim, B. C. (2007). Influence of Myosin Heavy- and Light Chain Isoforms on Early Postmortem Glycolytic Rate and Pork Quality. *Meat Sci.* 76, 281–288. doi:10.1016/j.meatsci.2006.11.009

Deldicque, L. (2020). Protein Intake and Exercise-Induced Skeletal Muscle Hypertrophy: An Update. *Nutrients* 12, 2023. doi:10.3390/nu12072023

Doran, P., Donoghue, P., O'connell, K., Gannon, J., and Ohlendieck, K. (2009). Proteomics of Skeletal Muscle Aging. *Proteomics* 9, 989–1003. doi:10.1002/pmic.200800365

Du, J., Zhang, P., Zhao, X., He, J., Xu, Y., Zou, Q., et al. (2019). MicroRNA-351-5p Mediates Skeletal Myogenesis by Directly Targeting Lactamase-β and Is Regulated by Lnc-mg. *FASEB j.* 33, 1911–1926. doi:10.1096/fj.201701394RRR

Filipowicz, W. (2005). RNAi: the Nuts and Bolts of the RISC Machine. *Cell* 122, 17–20. doi:10.1016/j.cell.2005.06.023

Fontelonga, T. M., Jordan, B., Nunes, A. M., Barraza-Flores, P., Bolden, N., Wuebbles, R. D., et al. (2019). Sunitinib Promotes Myogenic Regeneration and Mitigates Disease Progression in the Mdx Mouse Model of Duchenne Muscular Dystrophy. *Hum. Mol. Genet.* 28, 2120–2132. doi:10.1093/hmg/ddz044

Garneau, L., and Aguer, C. (2019). Role of Myokines in the Development of Skeletal Muscle Insulin Resistance and Related Metabolic Defects in Type 2 Diabetes. *Diabetes Metab.* 45, 505–516. doi:10.1016/j.diabet.2019.02.006

Heery, D. M., Kalkhoven, E., Hoare, S., and Parker, M. G. (1997). A Signature Motif in Transcriptional Co-activators Mediates Binding to Nuclear Receptors. *Nature* 387, 733–736. doi:10.1038/42750

Hennebry, A., Berry, C., Siriett, V., O'callaghan, P., Chau, L., Watson, T., et al. (2009). Myostatin Regulates Fiber-type Composition of Skeletal Muscle by Regulating MEF2 and MyoD Gene Expression. *Am. J. Physiology-Cell Physiol.* 296, C525–C534. doi:10.1152/ajpcell.00259.2007

Jakkaraju, S., Zhe, X., Pan, D., Choudhury, R., and Schuger, L. (2005). TIPs Are Tension-Responsive Proteins Involved in Myogenic versus Adipogenic Differentiation. *Developmental Cel* 9, 39–49. doi:10.1016/j.devcel.2005.04.015

Lee, S. H., Joo, S. T., and Ryu, Y. C. (2010). Skeletal Muscle Fiber Type and Myofibrillar Proteins in Relation to Meat Quality. *Meat Sci.* 86, 166–170. doi:10.1016/j.meatsci.2010.04.040

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. doi:10.1016/j.jprot.2019.103628

Listrat, A., Lebret, B., Louveau, I., Astruc, T., Bonnet, M., Lefaucheur, L., et al. (2016). How Muscle Structure and Composition Influence Meat and Flesh Quality. *Scientific World J.* 2016, 1–14. doi:10.1155/2016/3182746

Liu, N., Bezprozvannaya, S., Williams, A. H., Qi, X., Richardson, J. A., Bassel-Duby, R., et al. (2008). microRNA-133a Regulates Cardiomyocyte Proliferation and Suppresses Smooth Muscle Gene Expression in the Heart. *Genes Dev.* 22, 3242–3254. doi:10.1101/gad.1738708

Livak, K. J., and Schmittgen, T. D. (2001). Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the 2−ΔΔCT Method. *Methods* 25, 402–408. doi:10.1006/meth.2001.1262

Lu, T. X., and Rothenberg, M. E. (2018). MicroRNA. *J. Allergy Clin. Immunol.* 141, 1202–1207. doi:10.1016/j.jaci.2017.08.034

Ma, J., Wang, H., Liu, R., Jin, L., Tang, Q., Wang, X., et al. (2015). The miRNA Transcriptome Directly Reflects the Physiological and Biochemical Differences between Red, White, and Intermediate Muscle Fiber Types. *Ijms* 16, 9635–9653. doi:10.3390/ijms16059635

Mccarthy, J. J., Esser, K. A., Peterson, C. A., and Dupont-Versteegden, E. E. (2009). Evidence of MyomiR Network Regulation of *β*-myosin Heavy Chain Gene Expression during Skeletal Muscle Atrophy. *Physiol. Genomics* 39, 219–226. doi:10.1152/physiolgenomics.00042.2009

O'rourke, J. R., Georges, S. A., Seay, H. R., Tapscott, S. J., Mcmanus, M. T., Goldhamer, D. J., et al. (2007). Essential Role for Dicer during Skeletal Muscle Development. *Developmental Biol.* 311, 359–368. doi:10.1016/j.ydbio.2007.08.032

Pu, M., Chen, J., Tao, Z., Miao, L., Qi, X., Wang, Y., et al. (2019). Regulatory Network of miRNA on its Target: Coordination between Transcriptional and post-transcriptional Regulation of Gene Expression. *Cell. Mol. Life Sci.* 76, 441–451. doi:10.1007/s00018-018-2940-7

Qian, L., Tang, M., Yang, J., Wang, Q., Cai, C., Jiang, S., et al. (2015). Targeted Mutations in Myostatin by Zinc-finger Nucleases Result in Double-Muscled Phenotype in Meishan Pigs. *Sci. Rep.* 5, 14435. doi:10.1038/srep14435

Rehmsmeier, M., Steffen, P., Höchsmann, M., and Giegerich, R. (2004). Fast and Effective Prediction of microRNA/target Duplexes. *Rna* 10, 1507–1517. doi:10.1261/rna.5248604

Tieland, M., Trouwborst, I., and Clark, B. C. (2018). Skeletal Muscle Performance and Ageing. *J. Cachexia, Sarcopenia Muscle* 9, 3–19. doi:10.1002/jcsm.12238

Tobi, E. W., Slieker, R. C., Luijk, R., Dekkers, K. F., Stein, A. D., Xu, K. M., et al. (2018). DNA Methylation as a Mediator of the Association between Prenatal Adversity and Risk Factors for Metabolic Disease in Adulthood. *Sci. Adv.* 4, eaao4364. doi:10.1126/sciadv.aao4364

Van Rooij, E., Liu, N., and Olson, E. N. (2008). MicroRNAs Flex Their Muscles. *Trends Genet.* 24, 159–166. doi:10.1016/j.tig.2008.01.007

Wang, X. Y., Chen, X. L., Huang, Z. Q., Chen, D. W., Yu, B., He, J., et al. (2017). MicroRNA-499-5p Regulates Porcine Myofiber Specification by Controlling Sox6 Expression. *Animal* 11, 2268–2274. doi:10.1017/s1751731117001008

Williams, A. H., Valdez, G., Moresi, V., Qi, X., Mcanally, J., Elliott, J. L., et al. (2009). MicroRNA-206 Delays ALS Progression and Promotes Regeneration of Neuromuscular Synapses in Mice. *Science* 326, 1549–1554. doi:10.1126/science.1181046

Xie, S., Li, X., Qian, L., Cai, C., Xiao, G., Jiang, S., et al. (2019). An Integrated Analysis of mRNA and miRNA in Skeletal Muscle from Myostatin-Edited Meishan Pigs. *Genome* 62, 305–315. doi:10.1139/gen-2018-0110

Xu, M., Chen, X., Huang, Z., Chen, D., Yu, B., Chen, H., et al. (2018). MicroRNA-139-5p Suppresses Myosin Heavy Chain I and IIa Expression via Inhibition of the Calcineurin/NFAT Signaling Pathway. *Biochem. Biophysical Res. Commun.* 500, 930–936. doi:10.1016/j.bbrc.2018.04.202

Zhao, Y., Ransom, J. F., Li, A., Vedantham, V., Von Drehle, M., Muth, A. N., et al. (2007). Dysregulation of Cardiogenesis, Cardiac Conduction, and Cell Cycle in Mice Lacking miRNA-1-2. *Cell* 129, 303–317. doi:10.1016/j.cell.2007.03.030

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.

frontiers in Genetics

ORIGINAL RESEARCH
published: 25 March 2022
doi: 10.3389/fgene.2022.803477

Combined Analysis of RRBS DNA Methylome and Transcriptome Reveal Novel Candidate Genes Related to Porcine *Clostridium perfringens* Type C-Induced Diarrhea

Xiaoyu Huang[1], *Qiaoli Yang*[1], *Zunqiang Yan*[1], *Pengfei Wang*[1], *Hairen Shi*[2], *Jie Li*[1], *Xuefeng Shang*[1] *and Shuangbao Gun*[1,3]*

[1]College of Animal Science and Technology, Gansu Agricultural University, Lanzhou, China, [2]Tibet Academy of Agricultural and Animal Husbandry Science, Lasa, China, [3]Gansu Research Center for Swine Production Engineering and Technology, Lanzhou, China

Clostridium perfringens type C (Cp) is one of the principal microorganisms responsible for bacterial diarrhea in neonatal and pre-weaning piglets. To better understand the molecular effects of Cp infection, we performed a genome-wide comparison of the changes in DNA methylation and gene expression in Cp infected resistant and susceptible piglets. We characterized the pattern of changes in methylation and found 6485, 5968, and 6472 differentially methylated regions (DMRs) of piglets infected with Cp in IR vs. IC, IS vs. IC, and IS vs. IR groups, respectively. These methylation changes for genes mainly involved in immune and inflammatory responses, cell adhesion, and activation of transcription factors. Gene ontology and KEGG pathway analyses showed that the differentially methylated genes (DMGs) were associated with negative regulation of transcription, apoptotic processes, protein binding, and kinase activity. In addition, they were enriched in immunity-related pathways, such as MAPK signaling pathway, Toll-like receptor signaling pathway, and NF-kappa B signaling pathway. Integrative analysis identified 168, 198, and 7 mRNAs showing inverse correlations between methylation and expression with Cp infection. Altered DNA methylation and expression of various genes suggested their roles and potential functional interactions upon Cp infection, 14 immune-associated mRNAs with differential methylation and transcriptional repression were identified in IS vs. IR, commonly revealing that the differentially expressed genes (DEGs) *LBP*, *TBX21*, and *LCN2* were likely involved in the piglets against Cp infection. The present results provide further insight into the DNA methylation epigenetic alterations of *C. perfringens* type C infected piglet ileum tissues, and may advance the identification of biomarkers and drug targets for predicting susceptibility to and controlling *C. perfringens* type C-induced piglet diarrhea.

Keywords: DNA methylome, transcriptome, *Clostridium perfringens* type C, diarrhea, pig, resistance, susceptibility

OPEN ACCESS

Edited by:
Sundararajan Jayaraman, University of Illinois, United States

Reviewed by:
Mengjin Zhu, Huazhong Agricultural University, China
Xiangdong Ding, China Agricultural University, China

***Correspondence:**
Shuangbao Gun
gunsbao056@126.com

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

Received: *28 October 2021*
Accepted: *24 February 2022*
Published: *25 March 2022*

Citation:
Huang X, Yang Q, Yan Z, Wang P, Shi H, Li J, Shang X and Gun S (2022) Combined Analysis of RRBS DNA Methylome and Transcriptome Reveal Novel Candidate Genes Related to Porcine Clostridium perfringens Type C-Induced Diarrhea. Front. Genet. 13:803477. doi: 10.3389/fgene.2022.803477

Abbreviations: DMRs, differentially methylated regions; *E. coli*, *Escherichia coli*; GO, gene ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; mRNA, messenger RNA; RNAseq, RNA sequencing; RRBS, reduced representation bisulfite sequencing.

INTRODUCTION

Clostridium perfringens (*C. perfringens*) type C (Cp) frequently causes the severe, acute, and lethal necrotic enteritis (NE) in humans and livestock (Rood and Cole 1991; Lyras and Rood 2014), such as calves, sheep, goats, and pigs (Songer and Uzal 2005), especially in newborn piglets (Petit et al., 1999; Rood et al., 2018). Newborn piglets from the birthday until 3 weeks of age are highly susceptible to the clostridia because of their incompletely developed intestinal immune system, leading to mortality rates up to 100%. The Cp infection spreads rapidly and becomes an important problem worldwide (Posthaus et al., 2020).

C. perfringens type C beta (CPB) toxin is the essential virulence factor (Sayeed et al., 2008; Vidal et al., 2008; Uzal et al., 2009). Usually, the colonization and rapid proliferation of Cp intruded into the piglet's incompletely developed intestine forebode the start of NE disease. Due to trypsin inhibitors in the prevention of degradation of CPB, secretions of CPB toxin leads to initial epithelial damage or irritation, the toxin-induced intestinal damages rapidly cause an increase in permeability of vessels in the lamina propria, even the epithelial barrier further disrupted. The change of luminal environment causes acceleration of toxin production, which is absorbed into the systemic circulation and further causes hemorrhage, tissue necrosis, and even enterotoxemia (Posthaus et al., 2020). The Cp diseases can occur in acute and subacute-to-chronic forms (Hogh 1967; Chean et al., 2014). Piglets with acute disease usually appear with several symptoms such as depression, hemorrhagic diarrhea, and death mainly within the first 3 d after birth (Songer and Uzal 2005). Piglets with more protracted subacute-to-chronic clinical features almost have non-hemorrhagic diarrhea, and appear with hypo-immunity, growth reduction, and emaciation (Songer and Uzal 2005). Stoy et al. found that *C. perfringens* type A infection led to the increased expressions of inflammatory related genes (*CCL5*, *NFKBIA*, *IL8*, *IL1RN*, and *TNFAIP3*) of IPEC-J2 cells, and total count and densities of bacteria were markedly high in pigs NE, showing the significant positive correlation with disease severity (Stoy et al., 2015).

DNA methylation is one of the central epigenetic modifications; in mammalian genomes, it mainly occurs on cytosines at position C5 in CpG dinucleotides (Wang et al., 2019). DNA methylation always participates in numerous immunity and physiology processes, such as genomic imprinting, transcriptional regulation, growth, and developmental, immune, and inflammation regulation (Schübeler 2015). The methylation state normally dynamically changes and serves to regulate expressions of the responsive genes during host responses to environmental stimuli of pathogen infection, drug treatment, pollutants, and immune and inflammatory diseases (Kiga et al., 2014; Jiang et al., 2018; Swathy et al., 2018; Chen et al., 2019). Generally, the DNA occurrence of promoter methylation is often accompanied by transcription inhibition (Lee and Bartolomei 2013). Hypomethylation can promote the increase of transcriptional activity (Lluis and Cosma 2013), abnormal methylation can cause many diseases. Research has reported that bacterial endotoxins have profound impacts on gene expression in intestinal epithelial cells through DNA methylation modifications. The expressions of *FUT1* (Dai et al., 2017) and *FUT2* (Wu et al., 2018) were epigenetically modulated by DNA methylation of their promoters, regulating ETEC F18 resistance in weanling piglets. Other studies also addressed the impact of infection and LPS on the DNA methylation status of immune cells. In human macrophages, LPS-induced specific methylation changes lead to inactivation of pro-inflammatory pathways (Novakovic et al., 2016). However, systematic investigations on the global DNA methylation changes induced by *C. perfringens* type C infection and the methylation pattern of responsive genes are still scant.

This study aimed to explore the genomic regions and distribution of DNA methylation in piglet ileum tissues exposed to Cp infection and screened the potential DNA methylation targets for piglets against Cp infection, combined with RNAseq data of our previous study (Huang X. et al., 2019). This study comprehensively analyzed the effects of RRBS DNA methylome level and transcriptome level, and provided new insights in Cp-induced piglet diarrhea disease, which may contribute to the identification of biomarkers for diarrhea resistance against Cp infection.

MATERIALS AND METHODS

Animal Experiment

Bacterial culture, feeding, and management of piglets were in accordance with the description of Huang et al. (Huang X. et al., 2019; Huang X. Y. et al., 2019); the details were as follows: *Cp* strain (CVCC 2032; China Veterinary Culture Collection Center) was anaerobic shaking cultured 16 h at 37°C in the bouillon culture-medium (HopeBio, Qingdao, China), and an expected concentration of 1×10^9 CFU/ml Cp medium was used to inoculate piglets orally.

The 30 7-day-old experimental piglets (Landrace male × Yorkshire female) tested seronegative for *Escherichia coli* (*E. coli*), *Salmonella,* and *C. perfringens* by commercial enzyme-linked immunosorbent assay (ELISA) kits (Jiancheng Bioengineering Institute, Nanjing, China) from Dingxi City, Gansu Province of China. Then 25 piglets were randomly orally challenged with 1 ml 1×10^9 CFU/ml Cp medium for 5 consecutive days, the 5 remaining piglets were the control group (IC), and all piglets were housed separately and isolated in climate-controlled and fully air-conditioned, receiving water and diets *ad libitum* (Huang X. et al., 2019; Huang X. Y. et al., 2019).

During Cp infection, summing and ranking total scores of each piglet, according to fecal consistency from 0 to 3 grade (Kelly et al., 1990), meanwhile combining with the clinical signs, the top 5 piglets with the highest and lowest fecal scores were designated as susceptibility (IS) and resistance (IR) groups, respectively. Piglets of IR, IS, and IC groups were humanly slaughtered

under barbiturate anesthesia. The ileum tissues were collected to extract DNA (**Supplementary Figure S1**).

Nucleic Acid Isolation

According to the manufacturer's instruction, genomic DNA and total RNA were exacted by using the QIAamp Fast DNA Tissue Kit (Qiagen, Dusseldorf, Germany) and TRIzol reagent (Invitrogen, United States). Qualities and integrities of RNA extracts were assessed using the NanoPhotometer® spectrophotometer (Thermo Scientific, United States) and by 1% agarose gel electrophoresis with RNA Nano6000 Assay Kit of the Bionalyzer 2100 system (Agilent Technologies, United States), which were used for library preparation and subsequent analysis.

Library Preparation and RRBS Sequencing

The 200 to 1,000 bp in length fragmented DNA samples by MspI enzyme (NEB, United States) were then subjected to bisulfite conversion for converting any unmethylated cytosine to Uracil by EZ DNA Methylation-Gold™ Kit (Zymo Research, United States). Further, the Accel-NGS Methyl-Seq DNA Library Kit (Swift, MI) was utilized for attaching adapters to single-stranded DNA fragments. Bead-based SPRI clean-ups were used to remove both oligonucleotides and small fragments, as well as changing enzymatic buffer composition. Finally, we performed the pair-end 2 × 150 bp sequencing on an Illumina Hiseq 4000 platform housed in the LC Sciences (Hangzhou, China).

Data Normalization

Sequencing reads that contained adapter contamination, low quality bases, and undetermined bases were removed using Cutadapt and perl scripts. Quality control was verified using FastQC v0.11.4 (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/) (Cock et al., 2010), then reads were mapped and aligned to *Sus scrofa* 11.1 reference genome using Bismark v0.22.1 (Krueger and Andrews 2011). Further, for each cytosine site (or guanine corresponding to a cytosine on the opposite strand) in the reference genome sequence, the DNA methylation levels were determined by the ratio of numbers of reads supporting C (methylated) to that of total reads (methylated and unmethylated) using per scripts in house and MethPipe (Song et al., 2013). Analysis of differentially methylated regions (DMRs) was calculated by R package-MethylKit (Akalin et al., 2012) with default parameters (1000 bp slide windows, 500 bp overlap, p value < 0.05).

The Conjunctive Analysis of RRBS and RNA-Seq

Gene promoter DNA methylation usually inhibits gene expression. To explore the effect of DNA methylation on gene expression during piglets suffering from Cp infection, we also had conjunctively analyzed the negative relationship between promoter DMGs and differentially expressed genes (DEGs). Overlapping analysis was performed for DMGs and DEGs, screening negative correlation between gene expression levels and methylation levels by Pearson correlation analysis. The methylation level of DMGs ($|\Delta\text{Methylation}| > 5\%$) and expression level of DEGs ($|\Delta\text{expression } \log_2\text{FC}| > 1$) both with the $p < 0.05$ were selected to perform the conjoint analysis.

Bioinformatics Analyses

Gene ontology (GO) analysis of DMG-associated genes was performed using DAVID functional annotation tool (Huang et al., 2009). All annotated genes in *Sus scrofa* genome were used as background for GO analysis. Pathway enrichment analysis was performed using the Kyoto Encyclopedia of Genes and Genomes (KEGG) database available within the DAVID platform, and with WikiPathways database (https://www.wikipathways.org/) (Martens et al., 2021).

Data Statistics and Analysis

The t-test analysis and one-way ANOVA analysis were used to analyze the significance of diarrhea scores and fecal bacteria shedding of piglets in the IR, IS, and IC groups through SPSS 21.0 software. All values were expressed by mean ± standard error (M±SE), the $p < 0.05$ means the significant difference, $p < 0.01$ means the obviously significant difference.

RESULTS

Diarrhea Scores and Fecal Bacteria Shedding of Piglets in the IR, IS, and IC Groups

The average diarrhea score and total diarrhea score of piglets among the IR, IS, and IC groups were statistically analyzed, and the results are shown in **Table 1**. The average diarrhea scores of piglets in the IR and IS groups were 2.01 ± 0.26 and 2.68 ± 0.04, respectively, which were significantly higher than that in the IC group (0.61 ± 0.02) ($p < 0.01$). The total diarrhea scores of piglets in the IR and IS groups were 37.60 ± 0.87 and 67.6 ± 1.21, respectively, which were significantly higher than that in the IC group (11.40 ± 0.51, $p < 0.01$). These results suggested that the average diarrhea scores and total diarrhea scores of the IR and IS groups were significantly increased after Cp infection ($p < 0.01$).

The numbers of piglet feces Cp in the IR, IS, and IC groups were also counted, which are shown in **Figure 1**. As increasing times of Cp infection, the numbers of Cp in feces of piglets in the IR and IS groups were significantly increased, meanwhile, the numbers of feces Cp of piglets in the IS group were also significantly higher than those in the IR group ($p < 0.05$), which were both significantly higher than those in the IC group ($p < 0.01$). The results showed that Cp infection led to increasing numbers of feces bacteria of piglets, the tolerant piglets in the IR group had stronger resistance to bacterial infection, showing numbers of feces bacteria shedding obviously less than those in the IS group.

DNA Methylation Profiles of RRBS

Genome DNA methylation profiles of quintuplicate samples of IR, IS, and IC groups were analyzed. Overall, RRBS yielded an amount of 39–49 million reads per sample. After quality filtering, ranging from 59.6–65.63%, average 61.23% of reads were successfully aligned to the *Sus scrofa* 11 reference genomes. In total, we had identified 1.7–2.2 million CpG sites per sample, of which, the average 1.96 million were covered in all samples, representing 7.9% of total numbers of CpGs in the *Sus scrofa* genome. Raw sequencing data and mapping statistics are summarized in **Supplementary Table S1**.

To detect the DNA methylome changes induced by Cp infection, we compared Cp infected and uninfected piglets to identify methylated enrichment peaks (MEPs) in genomic DNA (**Supplementary Table S2**). Statistically, compared to the IC group, there were 157,833 MEPs identified in the IR group ($p < 0.05$), of which 110,874 were hypermethylated and 46,959 were hypomethylated (**Supplementary Table S3**), and 150,128 MEPs were identified in the IS group ($p < 0.05$), of which 79,818 were hypermethylated and 70,310 were hypomethylated (**Supplementary Table S4**). Mentionable, compared to the IR group, 160,738 MEPs were identified in the IS group ($p < 0.05$), of which 52,995 were hypermethylated and 107,743 were hypomethylated (**Supplementary Table S5**).

The chromosomal distribution of the methylated peaks was determined to assess whether methylation was associated with specific chromosomal features. As shown in the methylation map (**Figure 2**), the distribution of identified methylations almost covered all chromosomal, the methylation density in these regions was distinct among the chromosomes, chromosome MT (mitochondria) in particular, contained a relatively large unmethylated region among samples, which was related to the different degrees and correlation between methylation profiles of the infected IR and IS groups and the uninfected IC group.

Based on the CpG ratio, GC content, and length of the CpG-rich region, we divided gene promoters into three types: High-density CpG promoter (HCP), low-density CpG promoter (LCP), and intermediate density CpG promoter (ICP) (Yu et al., 2013). The information of MEPs in three comparative groups were shown and revealed the relatively uneven distribution across the genome. The majority of MEPs were in the intergenic regions, following in the exon and intron, lowest in the promoter (**Table 2**).

Then distribution of the MEPs in three types of promoter were analyzed. We found that the numbers of MEPs were obviously increased in the IR and IS groups after *C. perfringens* type C infection, which included MEP in LCP types, followed by in HCPs and ICPs. It is worth attention that in promoter CGIs, HCPs had more MEPs in the IS vs. IR group than ICPs or LCPs (**Table 2**).

TABLE 1 | Analyses of the diarrhea scores of piglets among the IR, IS, and IC groups.

Group	IC	IR	IS
Average diarrhea scores	0.61 ± 0.02^{C}	2.01 ± 0.26^{B}	2.68 ± 0.04^{A}
Total diarrhea scores	11.40 ± 0.51^{C}	37.60 ± 0.87^{B}	67.6 ± 1.21^{A}

Methylation Status of Genome CpG Islands

CpG islands (CGIs) obtained particular attention and interest for their role in controlling gene expression through epigenetic modification. We grouped the CGIs into four classes according to their distance to the RefSeq genes: promoter CGIs [from about −2 kb to +0.5 kb around the transcription start site (TSS)], exon and intron CGIs [from +0.5 kb around the TSS to the transcription terminal site (TTS)], and intergenic CGIs (about 2 kb to those that do not fall into either the promoter or the intragenic group) (**Figure 3A**).

We also analyzed the CpG methylation status of different gene segments in the piglet ileum after Cp infection. The numbers of MEPs in the four classes of CGIs and CGIs shore among IR, IS, and IC groups are shown in **Figures 3B,C**. Most MEPs were distributed in the intergenic CGIs among IR, IS, and IC comparative groups. It was worth mentioning that more CGI methylations happened in the IR vs. IC group (**Supplementary Table S6**) and the IS vs. IC group (**Supplementary Table S7**) than in the IS vs. IR group (**Supplementary Table S8**) ($p < 0.05$). Importantly, for the promoter CGIs, intergenic CGIs and 3′ transcript CGI, the numbers of hyper MEPs were also higher than hypo MEPs in the IR vs. IC group and the IS vs. IC group, which was opposite to those in the IS vs. IR group (**Figure 3B**; **Table 3**).

Identification of Differentially Methylated Genes

To explore the DMGs of piglet ileum tissues induced by Cp infection, we subsequently mapped all DMRs to their nearest genomic features and analyzed the DMGs located in promoter and CGI regions. Compared to the IC group, there were 6,485 DMGs having one or more DMRs including 5,186 mRNAs, 14

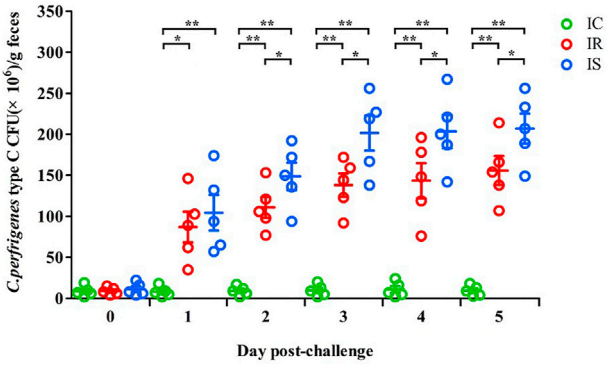


FIGURE 1 | The fecal shedding levels of piglets in the IR, IS, and IC groups after Cp infection. Note: Fecal CFUs were determined by plate count method. The horizontal line represents the mean. Green circles represent the IC group, red circles represent the IR group, and blue circles represent the IS group. An asterisk denotes a significant difference (*$p < 0.05$, **$p < 0.01$).

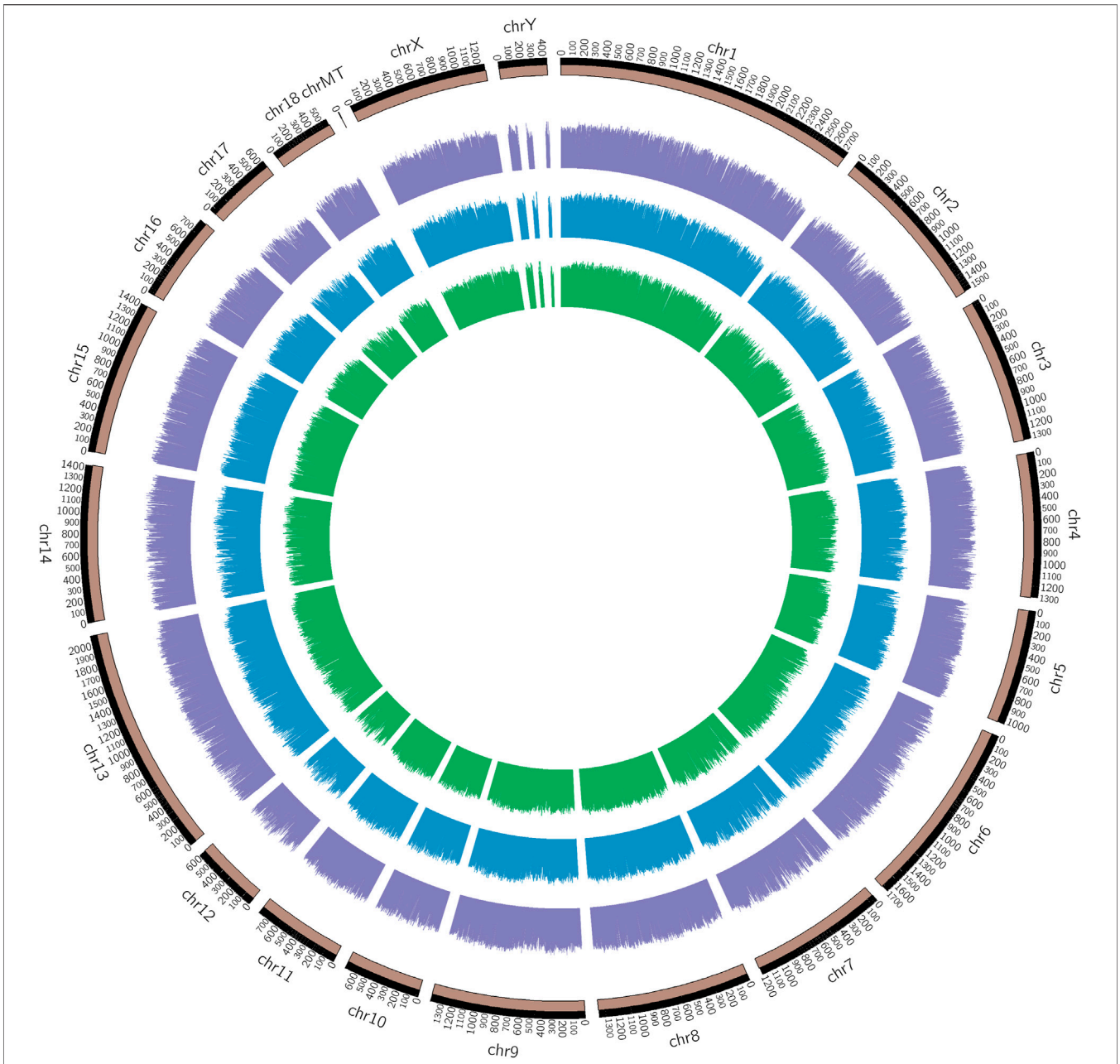


FIGURE 2 | Genomic distribution of DMRs among Cp infected and uninfected piglets. Note: From the inner circle to the outer circle represents the IR vs. IC, IS vs. IC, and IS vs. IR groups.

TABLE 2 | Analysis and distribution of the significant methylated genes of the IR, IS, and IC groups.

	IR vs. IC		IS vs. IC		IS vs. IR	
	Hyper-methyl	**Hypo-methyl**	**Hyper-methyl**	**Hypo-methyl**	**Hyper-methyl**	**Hypo-methyl**
Promoter	15,934	5,396	10,268	9,066	6,402	15,995
Exon	23,456	8,894	16,261	14,012	10,289	22,678
Intergenic	54,161	25,762	41,121	36,791	28,458	52,589
Intron	17,272	6,907	12,168	10,441	7,846	16,481

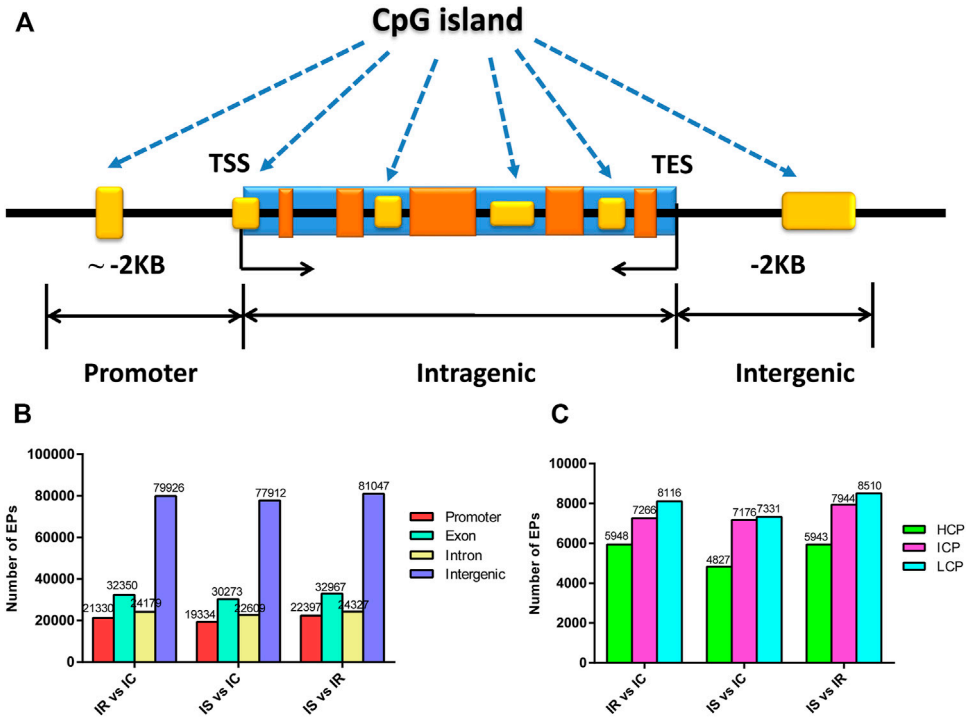


FIGURE 3 | Distribution of DNA methylation enrichment peaks in piglet ileum infected with *C. perfringens* type C infection. **(A)** Generic diagram showing CpG islands (CGIs) relative to gene transcript regions. **(B)** Numbers of MEPs in the CGI region of IR, IS, and IC groups. **(C)** Numbers of MEPs in the promoter CGIs. Note: HCP represents high CpG density promoter; ICP represents intermediate CpG density promoter; LCP represents low CpG density promoter.

TABLE 3 | The distributions of the significant CGIs methylations of the IR, IS, and IC groups.

	IR vs. IC		IS vs. IC		IS vs. IR	
	Hyper-methyl	**Hypo-methyl**	**Hyper-methyl**	**Hypo-methyl**	**Hyper-methyl**	**Hypo-methyl**
Promoter CGI	2,579	759	1,520	1,450	885	2,574
Intragenic CGI	14,535	5,312	9,854	8,597	6,191	13,882
3′ transcript CGI	1,489	435	913	749	505	1,418
Intergenic CGI	28,660	11,363	19,980	17,818	12,822	27,933

Note: CGIs represents CpG islands; CGIs shore represents CpG island shore.

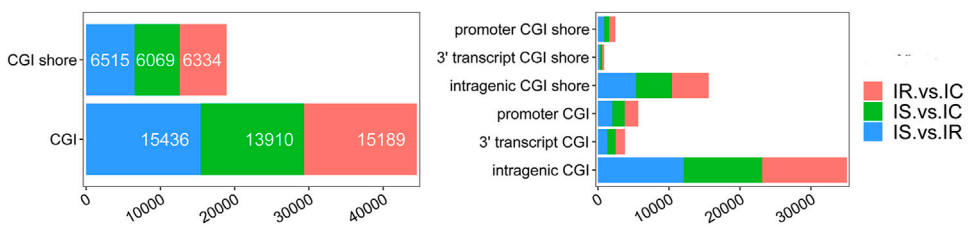


FIGURE 4 | The details of CGIs and CGI shores distributed in the IR vs. IC, IS vs. IC, and IS vs. IR groups.

miRNA, and 989 pseudogenes in the IR group (**Supplementary Table S6**), 5,968 DMGs including 4,819 mRNAs, 6 miRNA, and 886 pseudogenes in the IS group (**Supplementary Table S7**), meanwhile, compared to the IR group, 6,472 DMGs including 5,214 mRNAs, 14 miRNA, and 958 pseudogenes in the IS group (**Supplementary Table S8**). Except for miRNA and pseudogenes, DMGs almost contained an average of 4 DMRs in the gene body and promoter regions, mainly

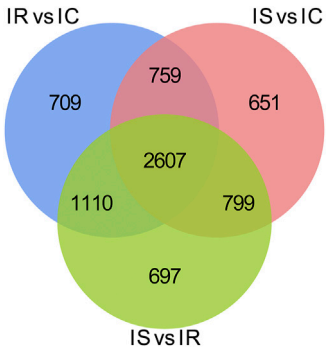


FIGURE 5 | Venn distribution of the differentially methylated mRNAs with CGIs in the IR vs. IC, IS vs. IC, and IS vs. IR groups. GO and KEGG signaling pathway analyses.

distributed in CGI and CGI shore (**Figure 4**). In total, 2607 DMGs were detected among the IR vs. IC, IS vs. IC, and IS vs. IR comparative groups (**Figure 5**).

In order to further characterize DMGs, gene ontology and KEGG signaling pathway analyses were carried out. GO analysis revealed 46, 35, and 38 GO terms significantly enriched in the IR vs. IC, IS vs. IC, and IS vs. IR groups (**Supplementary Table S9**), respectively. We found that Cp infection caused the DMGs mainly enriched in the biological and molecular functions, such as signal transduction, ion channel activity, nucleotide binding, protein transferase activity, immune response, cytoskeleton, GTPase activator activity, histone deacetylase complex, inactivation of MAPK activity, and nucleotide immunoglobulin production (**Figure 6**).

Among the KEGG pathway enrichment, a total of 23, 25, and 29 pathways were significantly enriched in the IR vs. IC, IS vs. IC, and IS vs. IR groups (**Supplementary Table S10**), respectively. The DMRs were enriched by MAPK signaling pathway, NF-kappa B signaling pathway, tight junction, chemokine signaling pathway, calcium signaling pathway, lysosome, PI3K-Akt signaling pathway, autophagy, Toll-like receptor signaling pathway, ECM-receptor interaction, and several metabolism pathways (**Figure 7**).

Correlation Analysis Between DNA Methylation and Gene Expression

DNA methylation of gene promoters is usually involved in inhibiting the expression level of the corresponding genes (Jones 2012). The potential effects of DNA methylation on gene expression were characterized by comparing methylation and RNA expression data (Guo et al., 2019). By analyzing a wide association between transcriptome gene expression and epigenome DNA methylation (in promoter or body), we explored the relationship between methylation changes at the promoter regions and gene expression changes (our previous study) of the IR, IS, and IC groups (Huang X. et al., 2019; Huang X. Y. et al., 2019).

Usually, there is the significant negative association between mean methylation levels of promoter regions and expressions of mRNAs. According to methylation levels of DMGs ($|\Delta\text{Methylation}| > 5\%$) and expression levels of DEGs ($|\Delta\text{expression } \log_2\text{FC}| > 1$) both with the $p < 0.05$, there were 168 mRNAs, 198 mRNAs, and 7 mRNAs screened in the IR vs. IC, IS vs. IC, and IS vs. IR groups, respectively (**Supplementary Table S11**). Importantly, to further reveal the functions of epigenetics and transcriptomics during piglets' resistance to Cp infection, we also analyzed the immune-associated mRNAs in the IS vs. IR group and constructed a differential methylation and transcription network (**Figure 8**). In this network, the screened mRNAs mainly belonged to T cell receptor, Toll like receptor, NF-κB, MAPK, JAK-STAT, IL-17, Th1, and Th2 cell differentiation signaling pathways, such as *NFKB1*, *MAPK14*, *TRIM25*, *TLR6*, *IL21R*, *LBP*, *IRF7*, *TBX21* and so on. Moreover, there were 14 mRNAs, which promoter methylations were obviously inversely related to transcriptional repressions, commonly identified (**Table 4**), in which *LCN2* with down-expressed and hypomethylated, *AKT2* with up-expressed and hypermethylated, were also identified in the IR vs. IC and IS vs. IC groups.

DISCUSSION

In this study, we have revealed the epigenetic alterations in piglet ileum tissues due to infection by Cp from a genome-wide comparative methylome analysis. However, it may only provide limited insights into the biological mechanisms of piglet diarrhea induced by Cp infection. Generally, complementary effects and synergistic interactions between omics in life science can be captured by integrative studies of multiple molecular layers. Building upon the successes in single-omics research, a better understanding of the molecular functions and disease etiologies by multi-omics integrated approaches from different omic levels (e.g., genetics, epigenetics, mRNA transcripts, proteins, and metabolites), as well as their interrelations and combined influences on the disease processes (Sun and Hu 2016). Therefore, we further conducted an integrated analysis of RRBS and RNA sequencing data and identified a subset of genes that was implicated in the piglet response to Cp infection.

The host immune response is crucial for defense against microbial pathogens, it is not different to that found in the complex process of host-pathogen interactions, the genomic expression pattern and program of host reflects responses to pathogens and virulence (Boldrick et al., 2002). Recently, a study reported that epigenetic modulations such as host DNA methylation could be manipulated to influence the host's gene expressions in response to defensing pathogens infection (Paschos and Allday, 2010). In addition, the changes and differences of DNA methylation presumably largely reflected the abilities of host epigenetic responses involved in the immune system against or triggers by pathogens (Tarakhovsky 2010). In this study, we have revealed the epigenetic alterations in

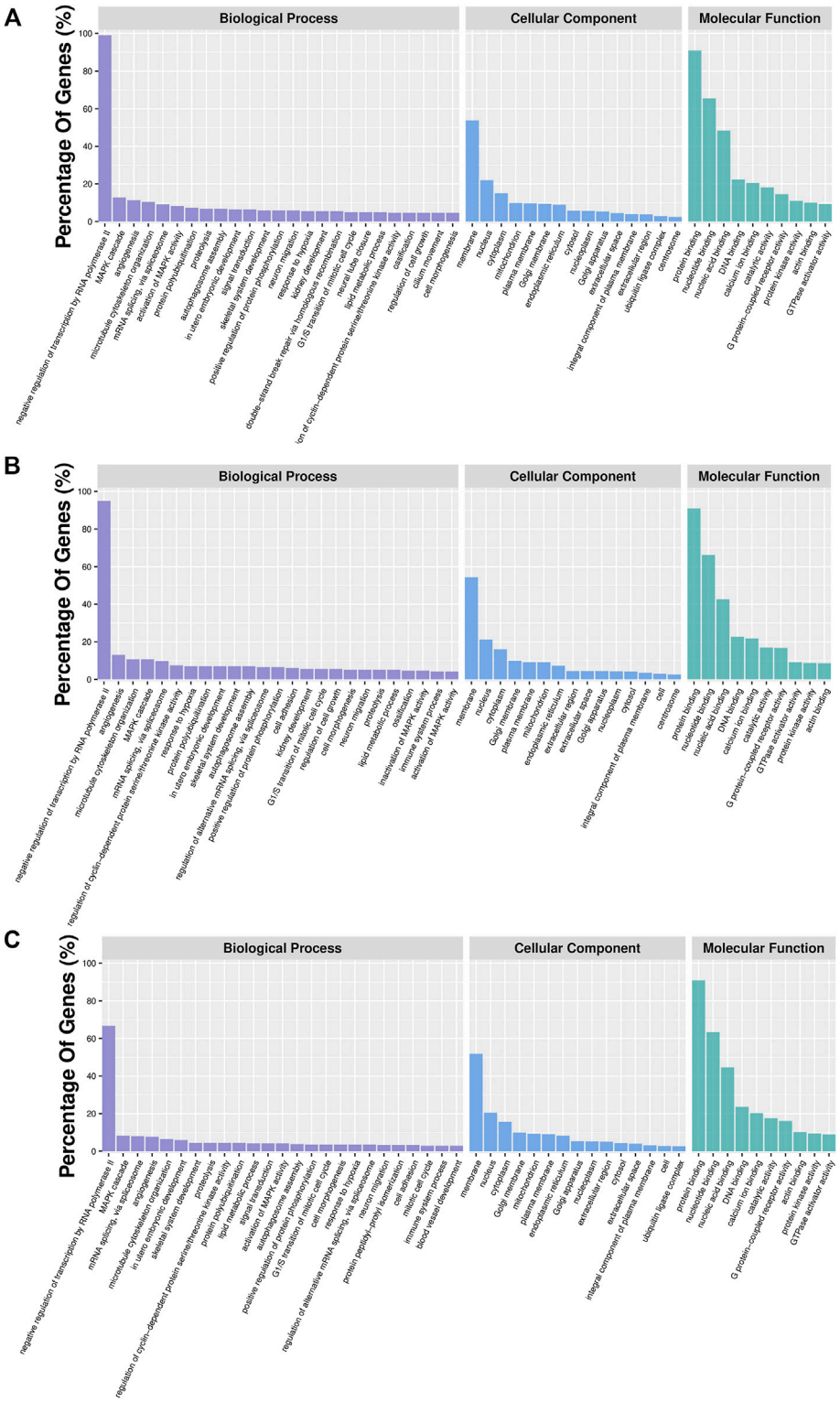


FIGURE 6 | GO analyses of DMGs located in the vicinity of significant differentially methylated regions (DMRs) of the IR vs. IC **(A)**, IS vs. IC **(B)**, and IS vs. IR groups **(C)**. Bar plots display enriched GO terms. The plots show significantly enriched degrees ($p < 0.05$).

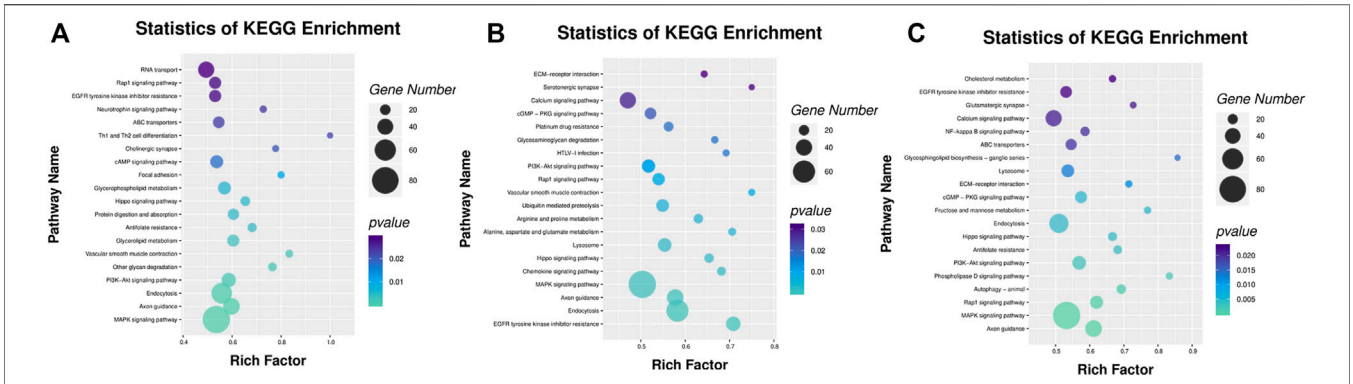


FIGURE 7 | KEGG signal pathway analyses of DMGs located in the vicinity of significant differentially methylated regions (DMRs) of the IR vs. IC **(A)**, IS vs. IC **(B)**, and IS vs. IR groups **(C)**. Bar plots display enriched KEGG pathways. The plots show significantly enriched degrees ($p < 0.05$).

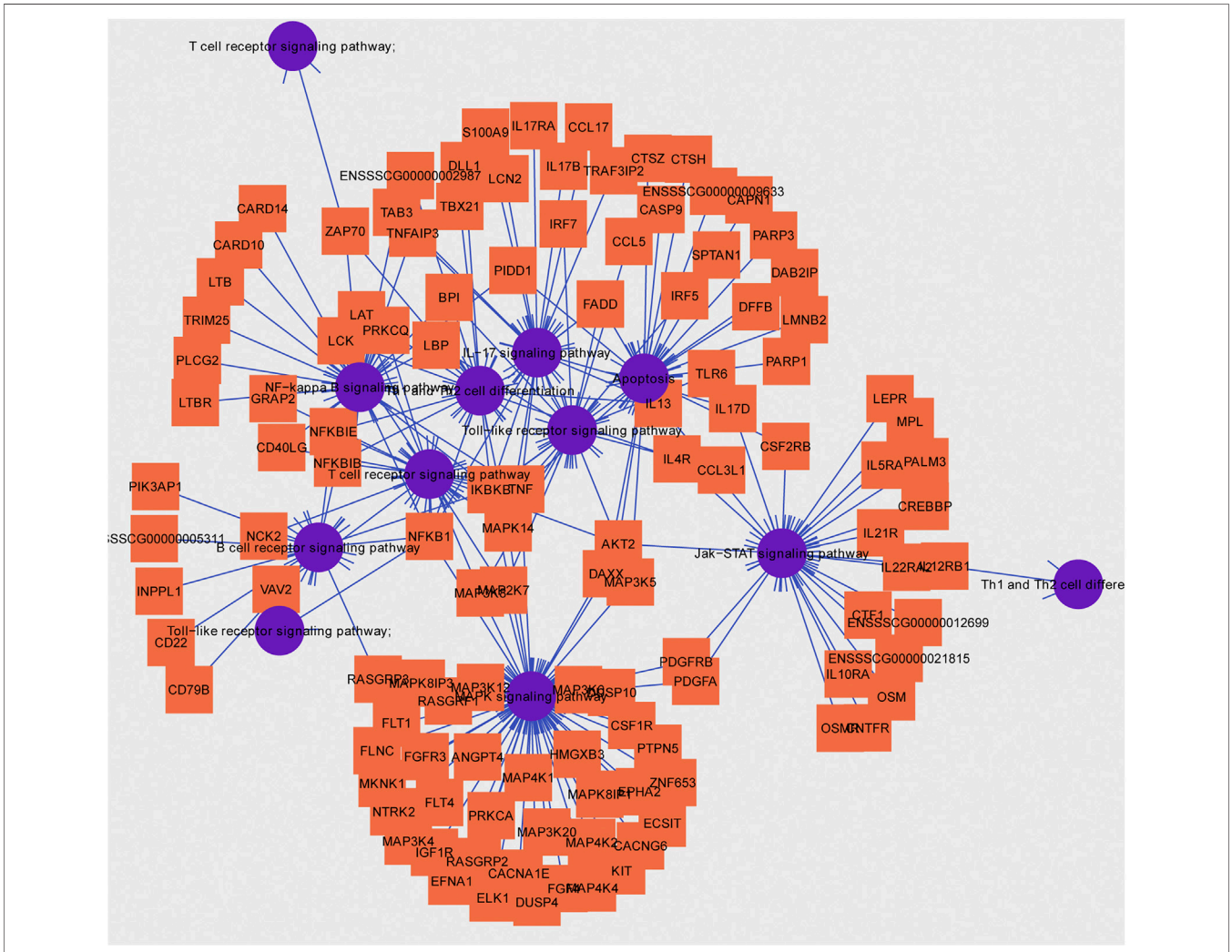


FIGURE 8 | Hub pathways and associated genes in piglets of the IS vs. IR group suffering from *C. perfringens* type C infection.

TABLE 4 | List of differentially methylated and expressed mRNAs in *C. perfringens* type C infected piglets of the IR vs. IS group.

Gene name	Expression change	Methylation change
CCL17	Upregulated	Hypermethylated
LBP	Downregulated	Hypomethylated
ENSSSCG00000002987	Downregulated	Hypomethylated
LCN2	Downregulated	Hypomethylated
FLNC	Upregulated	Hypomethylated
AKT2	Upregulated	Hypomethylated
ENSSSCG00000012699	Upregulated	Hypomethylated
ZAP70	Upregulated	Hypomethylated
NTRK2	Downregulated	Hypermethylated
CD40LG	Upregulated	Hypermethylated
TBX21	Upregulated	Hypermethylated
CASP9	Upregulated	Hypermethylated
IL22RA2	Upregulated	Hypermethylated
PALM3	Upregulated	Hypermethylated

piglet ileum tissues due to infection by Cp from a genome-wide comparative methylome analysis.

Our study found that the methylated peaks DNA methylation almost covered most chromosomal regions and presented obviously distinct methylation densities (**Figure 2**), indicating that Cp infection could trigger changes of DNA methylation of candidates. These changes may be correlated with piglets responding to Cp infection. Analysis of differential methylation genes revealed that 6635 DMGs were identified in piglets after Cp infection, in which 3,366 DMGs were identified both in the IR and IS groups (**Figure 5**), suggesting that the common DMGs identified in piglets of the IR and IS groups had an amount of similar methylation patterns during Cp infection, and further altered expressions of immune-related genes. Meanwhile, 697 DMGs especially identified in the IS vs. IR group (**Figure 5**), illustrating Cp-infected tolerant and sensitive piglets appeared with different methylation patterns induced by the Cp strain. These identified genes involved in the immune system and bacterial infection have been well studied. For example, *JAK1*, *JAK2,* and *JAK3*, found in three comparable groups, were the prototypical members of the JAK family and could play an essential signaling role for cytokines and interferons involved in immunity and antiviral responses (Ferrao et al., 2018). *RASGRF1*, found in IS vs. IC and IS vs. IR comparative groups, mainly participated in the Ras signaling pathway (Manyes et al., 2014). *AIRE* was one of the important regulators of peripheral T cell homeostasis, and played a certain role in control of intestinal tolerance (Kekalainen et al., 2015). Activation of *AKT2* was considered to protect mice from defense against *Salmonella* enterica Typhimurium infection (Kum et al., 2011), even blocking the development of intestinal inflammatory disease (Liu et al., 2019). These differential DMGs may be the crucial resistance candidates of piglet diarrhea.

The candidate mRNAs modified by differential DNA methylation merit greater attention. Further, combining the promoter methylation gene and its mRNA expression level, we had constructed the networks with differential methylations of DMGs and differential expressions of DEGs (**Supplementary Table S11**), and screened many immune related mRNAs, *S100A9*, *TARBP2*, *TRIM25*, *TBX21*, *LBP*, *TLR6*, involving the piglets resistant to Cp-induced diarrhea disease through defense-related signaling pathways of T cell receptor, Toll-like receptor, NF-κB, MAPK, and so on. Furthermore, these screened mRNAs' methylation of their promoters was inversely related to transcriptional repression in piglets after Cp infection. *S100A9* gene is considered to play crucial functions in participating innate immunity and mediating the inflammatory response during infection-induced host inflammation (Ometto et al., 2017). Researchers found S100A8/A9 recombinant attenuated bacterial adherence and invasion (Wang et al., 2018), and inhibited growth of multiple species, including *E. coli*, *S. aureus*, *Salmonella typhimurium*, and *C. perfringens* type C (Wang et al., 2018; Huang X. et al., 2019; Wang et al., 2020). However, the expression of *S100A9* gene was actually controlled by the methylation status of its promoter (Chandra et al., 2018). In our study, *S100A9* gene was down-expressed in transcription and hypermethylated in methylation, and was identified in the Cp infected IR and IS groups, hinting that the hypermethylation *S100A9* gene caused the down expression, which may play some roles in protecting piglets from Cp-induced diarrhea resistance.

Transactivation response element RNA-binding protein *TARBP2* inhibits the catalytic activity of interferon-induced double-stranded RNA (dsRNA)-activated protein kinase (PKR) to regulate stress-induced signaling pathways during viral infections and cell stress, and *TARBP2* is also a regulator of microRNA biogenesis and cellular stress response (Ukhueduan et al., 2021). *TARBP2* has been characterized as a key regulator for promoting or inhibiting cell proliferation and invasion. Research has shown that the *TARBP2* suppresses IFN-beta production and the innate antiviral response, especially in regulating the antiviral signaling pathways of the innate immune system (Ling et al., 2019). *TARBP2* gene, differential expressed and differential DNA methylated, was identified in Cp infected piglets of the IR and IS groups, which may play some important roles in piglet diarrhea resistance.

To further explore the susceptibility and tolerance of piglet resistance to Cp-induced diarrhea, we screened the potential resistant candidates from the IR and IS groups by integrating data of DNA methylation and mRNA expression (**Figure 8**). There were 14 mRNAs for which methylation of their promoters obviously inversely related to transcriptional repression were identified (**Table 4**), such as *LCN2*, *TBX21*, and *LBP*. *LCN2* with up-expressed and hypermethylated was identified both in the IR and IS groups after Cp infection. The changes in its promoter methylation and gene expression indicated that *LCN2* may participate in immune response to *C. perfringens* type C stimulation. We also constructed a network of immune-associated mRNAs both different changes in methylation and transcription in the IS vs. IR group (**Figure 8**), we speculated that these genes may be functionally linked and regulated by promoter methylation level of piglets in response to *C. perfringens* type Cp-induced diarrhea.

TBX21 is an important transcription factor of adaptive immunity that regulates the Th1/Th2 balance and increasing evidence has pointed to the critical roles in regulating innate immunity, cytokine balance, immune dysregulation, and bacterial

infection. A study reported that *pneumococcus* upregulated *TBX21* expression in the respiratory epithelium, knocking down *TBX21* suppressed *pneumococcus*-induced *TLR2* expression (Woo et al., 2014), and a change of *TBX21* may lead to the dysregulation of type I interferon pathways and T cell pathways (Gourh et al., 2009). These studies have proved the adaptive immune regulator *TBX21* participated in regulating host innate immune responses during pathogenic bacterial infections. In this study, differential methylated *TBX21* was identified in piglets in the IS and IR groups after Cp infection, and it was also found that up-regulated and hypermethylated *TBX21* in the IS vs. IR comparative group suggests that hypermethylated *TBX21* gene may cause the down expression in the transcriptional level, while the up-expression of *TBX21* gene may be regulated by other factors, which commonly participate in the process of piglet resistance to Cp-induced diarrhea.

LPS-binding protein (*LBP*) is a plasma protein that transfers LPS to the cell surface CD14 presented on the myeloid lineage, playing the crucial roles in the host innate immune response during the development of inflammatory and infectious-related diseases (Meng et al., 2021). *LBP* is also essential to control bacterial infection. Recently, *LBP* has been shown to potentiate the host immune responses to LPS invade, relieving *Salmonella typhimurium* or *E. coli* induced generation of reactive oxygen species in host macrophages (Sclutt et al., 1999). The LPS-LBP complex can bind to a receptor complex (CD14, MD2, and TLR4) for initiating signal cascade and triggering the secretion of pro-inflammatory cytokines (Gabarin et al., 2021; Won et al., 2021). While accomplishing either blocking *LBP* binding to LPS or binding LPS/LBP complexes to CD14 can protect the host from LPS-induced toxicity (Le Roy et al., 1999). Therefore, in this study, the hyper methylation of *LBP* was significantly negative regulation with its down expression in transcription in susceptible piglets for Cp infection, while the tolerance piglets with high expression may maintain balance of the inflammatory response induced by Cp infection. Considering the functions and expression characteristic of these genes, we believe that *LBP* and *TBX21* are strongly associated with piglet diarrhea induced by Cp infection. We speculated that these genes may be functionally linked and regulated by promoter methylation level of piglets in response to Cp-induced diarrhea, which may play some certain roles in protecting piglet resistance of diarrhea caused by bacterial infection.

In conclusion, we have profiled the landscape of DNA methylation of piglets in response to Cp infection, and analyzed the methylation and transcriptome data to further reveal the potential candidates implicated in the piglet immune response. Our findings have provided insight into the molecular effects of DNA methylation of piglet resistance to *C. perfringens* type C infection, which may contribute to the selection and breeding of piglet diarrhea resistant to *C. perfringens* type C infection.

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 animal study was reviewed and approved by the Animal Care and Use Ethics Committee of Gansu Agricultural University and Gansu Research Center for Swine Production Engineering and Technology. Written informed consent was obtained from the owners for the participation of their animals in this study.

AUTHOR CONTRIBUTIONS

XH and SG conceived of this study and wrote the manuscript. XH, HS and ZY performed the experiments. QY, PW and ZY performed the data analyses. JL and XS participated in the analysis and interpretation of data. All the authors read and approved the manuscript. All authors reviewed and approved the final manuscript.

FUNDING

This research was supported by the Young Science Fund, Gansu Agricultural University (GAU-DK-QNJJ-202112).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.803477/full#supplementary-material

REFERENCES

Akalin, A., Kormaksson, M., Li, S., Garrett-Bakelman, F. E., Figueroa, M. E., Melnick, A., et al. (2012). methylKit: a Comprehensive R Package for the Analysis of Genome-wide DNA Methylation Profiles. *Genome Biol.* 13, R87. doi:10.1186/gb-2012-13-10-R87

Boldrick, J. C., Alizadeh, A. A., Diehn, M., Dudoit, S., Liu, C. L., Belcher, C. E., et al. (2002). Stereotyped and Specific Gene Expression Programs in Human Innate Immune Responses to Bacteria. *Proc. Natl. Acad. Sci.* 99, 972–977. doi:10.1073/pnas.231625398

Chandra, A., Senapati, S., Roy, S., Chatterjee, G., and Chatterjee, R. (2018). Epigenome-wide DNA Methylation Regulates Cardinal Pathological Features of Psoriasis. *Clin. Epigenet* 10, 16. doi:10.1186/s13148-018-0541-9

Chean, R., Kotsanas, D., Francis, M. J., Palombo, E. A., Jadhav, S. R., Awad, M. M., et al. (2014). Comparing the Identification of Clostridium Spp. By Two Matrix-Assisted Laser Desorption Ionization-Time of Flight (MALDI-TOF) Mass Spectrometry Platforms to 16S rRNA PCR Sequencing as a Reference Standard: A Detailed Analysis of Age of Culture and Sample Preparation. *Anaerobe* 30, 85–89. doi:10.1016/j.anaerobe.2014.09.007

Chen, J., Wu, Y., Sun, Y., Dong, X., Wang, Z., Zhang, Z., et al. (2019). Bacterial Lipopolysaccharide Induced Alterations of Genome-wide DNA Methylation

and Promoter Methylation of Lactation-Related Genes in Bovine Mammary Epithelial Cells. *Toxins* 11, 298. doi:10.3390/toxins11050298

Cock, P. J. A., Fields, C. J., Goto, N., Heuer, M. L., and Rice, P. M. (2010). The Sanger FASTQ File Format for Sequences with Quality Scores, and the Solexa/Illumina FASTQ Variants. *Nucleic Acids Res.* 38, 1767–1771. doi:10.1093/nar/gkp1137

Dai, C., Sun, L., Xia, R., Sun, S., Zhu, G., Wu, S., et al. (2017). Correlation between the Methylation of the FUT1 Promoter Region and FUT1 Expression in the Duodenum of Piglets from Newborn to Weaning. *3 Biotech.* 7, 247. doi:10.1007/s13205-017-0880-9

Ferrao, R. D., Wallweber, H. J., and Lupardus, P. J. (2018). Receptor-mediated Dimerization of JAK2 FERM Domains Is Required for JAK2 Activation. *Elife* 7, e38089. doi:10.7554/eLife.38089

Gabarin, R. S., Li, M., Zimmel, P. A., Marshall, J. C., Li, Y., and Zhang, H. (2021). Intracellular and Extracellular Lipopolysaccharide Signaling in Sepsis: Avenues for Novel Therapeutic Strategies. *J. Innate Immun.* 13, 323–332. doi:10.1159/000515740

Gourh, P., Agarwal, S. K., Divecha, D., Assassi, S., Paz, G., Arora-Singh, R. K., et al. (2009). Polymorphisms inTBX21andSTAT4increase the Risk of Systemic Sclerosis: Evidence of Possible Geneâ€"gene Interaction and Alterations in Th1/Th2 Cytokines. *Arthritis Rheum.* 60, 3794–3806. doi:10.1002/art.24958

Guo, Y., van Schaik, T., Jhamat, N., Niazi, A., Chanrot, M., Charpigny, G., et al. (2019). Differential Gene Expression in Bovine Endometrial Epithelial Cells after challenge with LPS; Specific Implications for Genes Involved in Embryo Maternal Interactions. *Plos One* 14, e0222081. doi:10.1371/journal.pone.0222081

Hogh, P. (1967). Necrotizing Infectious Enteritis in Piglets, Caused by *Clostridium perfringens* Type C. II. Incidence and Clinical Features. *Acta Vet. Scand.* 8, 301–323.

Huang, D. 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

Huang, X. Y., Sun, W. Y., Yan, Z. Q., Shi, H. R., Yang, Q. L., Wang, P. F., et al. (2019a). Integrative Analyses of Long Non-coding RNA and mRNA Involved in Piglet Ileum Immune Response to *Clostridium perfringens* Type C Infection. *Front. Cel. Infect. Microbiol.* 9, 130. doi:10.3389/fcimb.2019.00130

Huang, X. Y., Sun, W. Y., Yan, Z. Q., Shi, H. R., Yang, Q. L., Wang, P. F., et al. (2019b). Novel Insights Reveal Anti-microbial Gene Regulation of Piglet Intestine Immune in Response to *Clostridium perfringens* Infection. *Sci. Rep.* 9, 1963. doi:10.1038/s41598-018-37898-5

Jiang, S., Yan, K., Sun, B., Gao, S., Yang, X., Ni, Y., et al. (2018). Long-Term High-Fat Diet Decreases Hepatic Iron Storage Associated with Suppressing TFR2 and ZIP14 Expression in Rats. *J. Agric. Food Chem.* 66, 11612–11621. doi:10.1021/acs.jafc.8b02974

Jones, P. A. (2012). Functions of DNA Methylation: Islands, Start Sites, Gene Bodies and beyond. *Nat. Rev. Genet.* 13, 484–492. doi:10.1038/nrg3230

Kekäläinen, E., Lehto, M.-K., Smeds, E., Pöntynen, N., Pekkarinen, P. T., Ulmanen, I., et al. (2015). Lymphopenia-induced Proliferation in the Absence of Functional Autoimmune Regulator (Aire) Induces Colitis in Mice. *Immunol. Lett.* 167, 17–22. doi:10.1016/j.imlet.2015.06.010

Kelly, D., O'Brien, J. J., and Mccracken, K. J. (1990). Effect of Creep Feeding on the Incidence, Duration and Severity of post-weaning Diarrhoea in Pigs. *Res. Vet. Sci.* 49, 223–228. doi:10.1016/s0034-5288(18)31082-8

Kiga, K., Mimuro, H., Suzuki, M., Shinozaki-Ushiku, A., Kobayashi, T., Sanada, T., et al. (2014). Epigenetic Silencing of miR-210 Increases the Proliferation of Gastric Epithelium during Chronic *Helicobacter pylori* Infection. *Nat. Commun.* 5, 4497. doi:10.1038/ncomms5497

Krueger, F., and Andrews, S. R. (2011). Bismark: a Flexible Aligner and Methylation Caller for Bisulfite-Seq Applications. *Bioinformatics* 27, 1571–1572. doi:10.1093/bioinformatics/btr167

Kum, W. W. S., Lo, B. C., Yu, H. B., and Finlay, B. B. (2011). Protective Role of Akt2 in *Salmonella enterica* Serovar Typhimurium-Induced Gastroenterocolitis. *Infect. Immun.* 79, 2554–2566. doi:10.1128/iai.01235-10

Le Roy, D., Di Padova, F., Tees, R., Lengacher, S., Landmann, R., Glauser, M. P., et al. (1999). Monoclonal Antibodies to Murine Lipopolysaccharide (LPS)-binding Protein (LBP) Protect Mice from Lethal Endotoxemia by Blocking Either the Binding of LPS to LBP or the Presentation of LPS/LBP Complexes to CD14. *J. Immunol.* 162, 7454–7460.

Lee, J. T., and Bartolomei, M. S. (2013). X-inactivation, Imprinting, and Long Noncoding RNAs in Health and Disease. *Cell* 152, 1308–1323. doi:10.1016/j.cell.2013.02.016

Ling, T., Weng, G.-X., Li, J., Li, C., Wang, W., Cao, L., et al. (2019). TARBP2 Inhibits IRF7 Activation by Suppressing TRAF6-Mediated K63-Linked Ubiquitination of IRF7. *Mol. Immunol.* 109, 116–125. doi:10.1016/j.molimm.2019.02.019

Liu, L., Liang, L., Liang, H., Wang, M., Lu, B., Xue, M., et al. (2019). Fusobacterium Nucleatum Aggravates the Progression of Colitis by Regulating M1 Macrophage Polarization via AKT2 Pathway. *Front. Immunol.* 10, 15. doi:10.3389/fimmu.2019.01324

Lluis, F., and Cosma, M. P. (2013). Resetting Epigenetic Signatures to Induce Somatic Cell Reprogramming. *Cell. Mol. Life Sci.* 70, 1413–1424. doi:10.1007/s00018-012-1137-8

Lyras, D., and Rood, J. I. (2014). Preface: ClostPath 2013 Meeting on the Molecular Biology and Pathogenesis of the Clostridia Special Issue. *Anaerobe* 30, 183. doi:10.1016/j.anaerobe.2014.11.006

Manyes, L., Arribas, M., Gomez, C., Calzada, N., Fernandez-Medarde, A., and Santos, E. (2014). Transcriptional Profiling Reveals Functional Links between RasGrf1 and Pttg1 in Pancreatic Beta Cells. *Bmc Genomics* 15, 1019. doi:10.1186/1471-2164-15-1019

Martens, M., Ammar, A., Riutta, A., Waagmeester, A., Slenter, D. N., Hanspers, K., et al. (2021). WikiPathways: Connecting Communities. *Nucleic Acids Res.* 49, D613–D621. doi:10.1093/nar/gkaa1024

Meng, L., Song, Z., Liu, A., Dahmen, U., Yang, X., and Fang, H. (2021). Effects of Lipopolysaccharide-Binding Protein (LBP) Single Nucleotide Polymorphism (SNP) in Infections, Inflammatory Diseases, Metabolic Disorders and Cancers. *Front. Immunol.* 12, 12. doi:10.3389/fimmu.2021.681810

Novakovic, B., Habibi, E., Wang, S.-Y., Arts, R. J. W., Davar, R., Megchelenbrink, W., et al. (2016). β-Glucan Reverses the Epigenetic State of LPS-Induced Immunological Tolerance. *Cell* 167, 1354–1368. e1314. doi:10.1016/j.cell.2016.09.034

Ometto, F., Friso, L., Astorri, D., Botsios, C., Raffeiner, B., Punzi, L., et al. (2017). Calprotectin in Rheumatic Diseases. *Exp. Biol. Med. (Maywood)* 242, 859–873. doi:10.1177/1535370216681551

Paschos, K., and Allday, M. J. (2010). Epigenetic Reprogramming of Host Genes in Viral and Microbial Pathogenesis. *Trends Microbiology* 18, 439–447. doi:10.1016/j.tim.2010.07.003

Petit, L., Gibert, M., and Popoff, M. R. (1999). *Clostridium perfringens*: Toxinotype and Genotype. *Trends Microbiology* 7, 104–110. doi:10.1016/s0966-842x(98)01430-9

Posthaus, H., Kittl, S., Tarek, B., and Bruggisser, J. (2020). Clostridium Perfringenstype C Necrotic Enteritis in Pigs: Diagnosis, Pathogenesis, and Prevention. *J. VET. Diagn. Invest.* 32, 203–212. doi:10.1177/1040638719900180

Rood, J. I., Adams, V., Lacey, J., Lyras, D., McClane, B. A., Melville, S. B., et al. (2018). Expansion of the *Clostridium perfringens* Toxin-Based Typing Scheme. *Anaerobe* 53, 5–10. doi:10.1016/j.anaerobe.2018.04.011

Rood, J. I., and Cole, S. T. (1991). Molecular Genetics and Pathogenesis of *Clostridium perfringens*. *Microbiol. Rev.* 55, 621–648. doi:10.1128/mmbr.55.4.621-648.1991

Sayeed, S., Uzal, F. A., Fisher, D. J., Saputo, J., Vidal, J. E., Chen, Y., et al. (2008). Beta Toxin Is Essential for the Intestinal Virulence of *Clostridium perfringens* Type C Disease Isolate CN3685 in a Rabbit Ileal Loop Model. *Mol. Microbiol.* 67, 15–30. doi:10.1111/j.1365-2958.2007.06007.x

Schutt, C. (1999). Fighting Infection: the Role of Lipopolysaccharide Binding Proteins CD14 and LBP. *Pathobiology* 67, 227–229. doi:10.1159/000028097

Schübeler, D. (2015). Function and Information Content of DNA Methylation. *Nature* 517, 321–326. doi:10.1038/nature14192

Song, Q., Decato, B., Hong, E. E., Zhou, M., Fang, F., Qu, J., et al. (2013). A Reference Methylome Database and Analysis Pipeline to Facilitate Integrative and Comparative Epigenomics. *Plos One* 8, e81148. doi:10.1371/journal.pone.0081148

Songer, J. G., and Uzal, F. A. (2005). Clostridial Enteric Infections in Pigs. *J. VET. Diagn. Invest.* 17, 528–536. doi:10.1177/104063870501700602

Støy, A. C. F., Mølbak, L., Delègue, C. L., Thymann, T., Sangild, P. T., Heegaard, P. M. H., et al. (2015). Necrotizing Enterocolitis in Preterm Pigs Is Associated with

Increased Density of Intestinal Mucosa-Associated Bacteria Including *Clostridium perfringens*. *Neonatology* 108, 188–195. doi:10.1159/000431280

Sun, Y. V., and Hu, Y.-J. (2016). Integrative Analysis of Multi-Omics Data for Discovery and Functional Studies of Complex Human Diseases. *Adv. Genet.* Vol. 93, 147–190. doi:10.1016/bs.adgen.2015.11.004

Swathy, B., Saradalekshmi, K. R., Nair, I. V., Nair, C., and Banerjee, M. (2018). Understanding the Influence of Antipsychotic Drugs on Global Methylation Events and its Relevance in Treatment Response. *Epigenomics* 10, 233–247. doi:10.2217/epi-2017-0086

Tarakhovsky, A. (2010). Tools and Landscapes of Epigenetics. *Nat. Immunol.* 11, 565–568. doi:10.1038/ni0710-565

Ukhueduan, B., Chukwurah, E., and Patel, R. C. (2021). Regulation of PKR Activation and Apoptosis during Oxidative Stress by TRBP Phosphorylation. *Int. J. Biochem. Cel Biol.* 137, 106030. doi:10.1016/j.biocel.2021.106030

Uzal, F. A., Saputo, J., Sayeed, S., Vidal, J. E., Fisher, D. J., Poon, R., et al. (2009). Development and Application of New Mouse Models to Study the Pathogenesis of *Clostridium perfringens* Type C Enterotoxemias. *Infect. Immun.* 77, 5291–5299. doi:10.1128/iai.00825-09

Vidal, J. E., McClane, B. A., Saputo, J., Parker, J., and Uzal, F. A. (2008). Effects of *Clostridium perfringens* Beta-Toxin on the Rabbit Small Intestine and colon. *Infect. Immun.* 76, 4396–4404. doi:10.1128/iai.00547-08

Wang, H., Zong, Q., Wang, S., Zhao, C., Wu, S., and Bao, W. (2019). Genome-Wide DNA Methylome and Transcriptome Analysis of Porcine Intestinal Epithelial Cells upon Deoxynivalenol Exposure. *J. Agric. Food Chem.* 67, 6423–6431. doi:10.1021/acs.jafc.9b00613

Wang, S., Song, R., Wang, Z., Jing, Z., Wang, S., and Ma, J. (2018). S100A8/A9 in Inflammation. *Front. Immunol.* 9, 14. doi:10.3389/fimmu.2018.01298

Wang, W., Zhou, C., Tang, H., Yu, Y., and Zhang, Q. (2020). Combined Analysis of DNA Methylome and Transcriptome Reveal Novel Candidate Genes Related to Porcine *Escherichia coli* F4ab/ac-Induced Diarrhea. *Front. Cel. Infect. Microbiol.* 10, 15. doi:10.3389/fcimb.2020.00250

Won, Y., Yang, J. I., Park, S., and Chun, J. S. (2021). Lipopolysaccharide Binding Protein and CD14, Cofactors of Toll-like Receptors, Are Essential for Low-Grade Inflammation-Induced Exacerbation of Cartilage Damage in Mouse Models of Posttraumatic Osteoarthritis. *Arthritis Rheumatol.* 73, 1451–1460. doi:10.1002/art.41679

Woo, C. H., Shin, S. G., Koh, S. H., and Lim, J. H. (2014). TBX21 Participates in Innate Immune Response by Regulating Toll-like Receptor 2 Expression inStreptococcus Pneumoniaeinfections. *Mol. Oral Microbiol.* 29, 233–243. doi:10.1111/omi.12061

Wu, Z., Feng, H., Cao, Y., Huang, Y., Dai, C., Wu, S., et al. (2018). New Insight into the Molecular Mechanism of the FUT2 Regulating *Escherichia coli* F18 Resistance in Weaned Piglets. *Ijms* 19, 3301. doi:10.3390/ijms19113301

Yu, M., Li, M., Luo, S. H., Zhang, Y., Liu, H., Gao, Y. X., et al. (2013). Folic Acid Stimulation of Neural Stem Cell Proliferation is Associated With Altered Methylation Profile of PI3K/Akt/CREB. *J Nutr Biochem.* 25, 495, 502. doi:10.1016/j.jnutbio.2013.12.010

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.

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 11 August 2022
DOI 10.3389/fgene.2022.899034

OPEN ACCESS

EDITED BY
Xiao Wang,
Konge Larsen ApS, Denmark

REVIEWED BY
Madhu Khullar,
Post Graduate Institute of Medical Education and Research (PGIMER), India
Zhengwei Yuan,
ShengJing Hospital of China Medical University, China

*CORRESPONDENCE
Guorong Lyu,
lgr_feus@sina.com

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

RECEIVED 18 March 2022
ACCEPTED 13 July 2022
PUBLISHED 11 August 2022

CITATION
Yang Y, Yang H, Lian X, Yang S, Shen H, Wu S, Wang X and Lyu G (2022), Circulating microRNA: Myocardium-derived prenatal biomarker of ventricular septal defects. *Front. Genet.* 13:899034. doi: 10.3389/fgene.2022.899034

Circulating microRNA: Myocardium-derived prenatal biomarker of ventricular septal defects

Yiru Yang[1], Hainan Yang[2], Xihua Lian[1,3], Shuping Yang[4], Haolin Shen[4], Shufen Wu[4], Xiali Wang[5] and Guorong Lyu[1,5]*

[1]Department of Ultrasound, The Second Affiliated Hospital of Fujian Medical University, Quanzhou, Fujian, China, [2]Department of Ultrasound, The First Affiliated Hospital of Xiamen University, Xiamen, Fujian, China, [3]Department of Pathology and Biomedical Science, University of Otago, Christchurch, New Zealand, [4]Department of Ultrasound, Zhangzhou Affiliated Hospital of Fujian Medical University, Zhangzhou, Fujian, China, [5]Collaborative Innovation Center for Maternal and Infant Health Service Application Technology, Quanzhou Medical College, Quanzhou, Fujian, China

Background: Recently, circulating microRNAs (miRNAs) from maternal blood and amniotic fluid have been used as biomarkers for ventricular septal defect (VSD) diagnosis. However, whether circulating miRNAs are associated with fetal myocardium remains unknown.

Methods: Dimethadione (DMO) induced a VSD rat model. The miRNA expression profiles of the myocardium, amniotic fluid and maternal serum were analyzed. Differentially expressed microRNAs (DE-microRNAs) were verified by qRT–PCR. The target gene of miR-1-3p was confirmed by dual luciferase reporter assays. Expression of amniotic fluid-derived DE-microRNAs was verified in clinical samples.

Results: MiRNAs were differentially expressed in VSD fetal rats and might be involved in cardiomyocyte differentiation and apoptosis. MiR-1-3p, miR-1b and miR-293-5p were downregulated in the myocardium and upregulated in amniotic fluid/maternal serum. The expression of amniotic fluid-derived DE-microRNAs (miR-1-3p, miR-206 and miR-184) was verified in clinical samples. Dual luciferase reporter assays confirmed that miR-1-3p directly targeted SLC8A1/NCX1.

Conclusion: MiR-1-3p, miR-1b and miR-293-5p are downregulated in VSD myocardium and upregulated in circulation and may be released into circulation by cardiomyocytes. MiR-1-3p targets SLC8A1/NCX1 and participates in myocardial apoptosis. MiR-1-3p upregulation in circulation is a direct and powerful indicator of fetal VSD and is expected to serve as a prenatal VSD diagnostic marker.

KEYWORDS

ventricular septal defect (VSD), circulating microRNA (miRNA), prenatal diagnosis (MeSH), biomarker, fetus [mesh]

1 Introduction

Ventricular septal defects (VSDs) are the most common congenital heart defects (CHDs), accounting for approximately 40% of CHDs (Spicer et al., 2014; Cox et al., 2020). Accurate prenatal diagnosis is helpful for reducing mortality and improving prognosis (van Nisselrooij et al., 2020). However, it was reported that due to factors such as the experience of sonographers and the quality of ultrasound images, approximately half of CHDs were not identified prenatally. Even when the ultrasound image quality is good, 31% of CHD cases are still missed (van Nisselrooij et al., 2020). Therefore, it is of great significance to seek biomarkers for the prenatal diagnosis of VSD to improve the accuracy of diagnosis and pregnancy outcomes.

Human genetics research has identified many genes related to hereditary and sporadic CHD that encode transcription factors that regulate the morphogenesis of the ventricular septum or outflow tract during heart development (Bruneau, 2008). Epigenetic mechanisms, including DNA methylation and noncoding RNA (ncRNA), are involved in the pathogenesis and phenotype of VSD (Grunert et al., 2016; Thomford et al., 2018; Yang et al., 2021). MicroRNAs (miRNAs), as highly conserved ncRNAs, regulate gene expression at the posttranscriptional level by binding to target genes, inducing epigenetic modifications, which are closely related to the cell cycle and mammalian development (Pulignani and Andreassi, 2019; Panni et al., 2020). By regulating VSD-related transcription factors, miRNAs participate in the proliferation and differentiation of cardiomyocytes, and morphogenesis of heart, as well as pathophysiological processes such as myocardial hypoxia and cardiac remodelling, which are interrelated with the occurrence, progression and outcome of VSD (Smith et al., 2015; Islas and Moreno-Cuevas, 2018; Sabour et al., 2018; Meng et al., 2020). MiRNAs can be encapsulated in lipid vesicles or connected with protein or lipoprotein complexes to ensure stability and avoid degradation during extracellular secretion. Therefore, miRNAs have the potential to be diagnostic markers of diseases (Creemers et al., 2012; Mori et al., 2019).

Maternal blood is commonly used for prenatal diagnosis and is easy to obtain, while amniotic fluid has a high content of fetal free nucleic acids and is less likely to be contaminated by maternal nucleic acids (Hui et al., 2012). In recent years, people have attempted to use circulating miRNAs derived from maternal blood and amniotic fluid as biomarkers for CHD diagnosis (Song et al., 2018; Jin et al., 2021; Yang et al., 2021). It is, however, unclear whether circulating miRNAs are associated with fetal heart tissue. Thus, we analyzed the expression profiles of amniotic fluid-derived and maternal serum-derived miRNAs in fetal VSD rats to explore the relationship between myocardium-derived and circulating miRNAs. Furthermore, the possibility of using circulating miRNAs as VSD prenatal biomarkers was verified in the clinical samples to serve as the basis for circulating miRNAs to assist prenatal diagnosis.

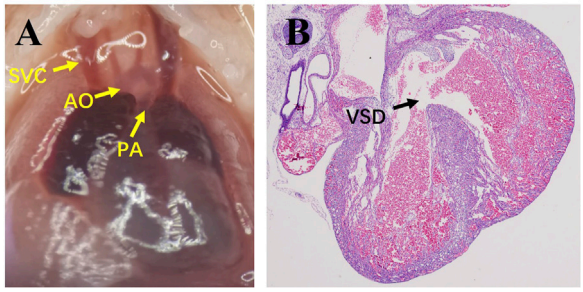


FIGURE 1
Images of microdissection and the pathology of VSD fetuses. **(A)**, Microdissection shows the correct connection of major vessels; **(B)**, pathology shows VSD. SVC, superior vena cava; AO, aorta; PA, pulmonary artery; VSD, ventricular septal defect.

2 Materials and methods

2.1 Construction of a ventricular septal defect rat model

Sprague–Dawley (SD) rats (Shanghai SLAC Laboratory Animal Co., Ltd.) were selected to construct VSD models. Female and male rats in estrus were kept in cages at a ratio of 2:1. The day when vaginal plugs were found was recorded as embryonic day 0 (D0). The dams were separated from male rats and randomized to a negative control (NC) group and a dimethadione (DMO) group. The mean maternal age of rats was 14.44 (SD, 5.17) weeks vs. 15.44 (SD, 4.06) weeks in the NC and DMO groups, respectively.

From 19:00 on D8, the DMO group was given 5 ml/kg DMO (drug concentration, 60 mg/ml) by oral gavage once every 12 h six times, while the NC group was given the same dose of distilled water at the same time. The rats were fed standard food and distilled water *ad libitum* and received humane care.

In the DMO group, excluding fetuses with umbilical hernia, subcutaneous edema or other obvious abnormal appearances, VSD fetuses without malformation of major vessels were selected as the VSD group (Figure 1), and fetuses with intact ventricular septum were selected as the non-VSD group.

2.2 Sample collection and pretreatment

On D19, the pregnant mice were anaesthetized by intraperitoneal injection of pentobarbital (40 mg/kg), and the amniotic fluid was carefully collected with a sterile syringe after the uterus was exposed. The amniotic fluid was immediately

TABLE 1 General characteristics of pregnant women.

	NC group (n = 7)	VSD group (n = 7)	p
Maternal age (years)	30.00 ± 4.90	29.00 ± 5.03	0.713
Gestational age (weeks)	21.80 ± 3.12	24.59 ± 1.35	0.140
Weight (kg)	53.04 ± 7.18	56.92 ± 8.67	0.397
Parity(n)	0.00 (0.00,2.00)	0.00 (0.00,2.00)	1.000
Oligohydramnios or polyhydramnios (%)	0.00	0.00	—

NC, negative control; VSD, ventricular septal defect.

centrifuged at 1,200 g and 4°C for 10 min, and the supernatant was recovered and stored at −80°C.

Blood was withdrawn through cardiac puncture and incubated at room temperature for 1 h. The coagulated blood was centrifuged twice at 4°C (1700 g, 10 min and 2000 g, 10 min), and then the supernatant was stored at −80°C.

The fetus was obtained by caesarean section. Microdissection was used to observe the position and connection of the major blood vessels. After that, the fetal heart was removed and washed in cold PBS solution. The heart tissue was embedded in paraffin according to the routine procedure. Parts of the wax blocks were cut into 3 μm slices and prepared for HE staining to observe the ventricular septum, and other parts were stored for RNA extraction.

2.3 Collection of clinical amniotic fluid samples

From August 2020 to June 2021, women who visited the Second Affiliated Hospital of Fujian Medical University were included when ultrasound-guided amniocentesis met their clinical needs. The amniotic fluid was extracted after informed consent was obtained from all pregnant women. The amniotic fluid of fetuses with a normal chromosome karyotype and without pregnancy complications or other diseases that may affect the growth and development of the fetus was collected. According to the results of follow-up after birth or induction of labor, they were classified into the VSD group and the NC group. General characteristics of pregnant women are displayed in Table 1. This study was carried out in accordance with The Code of Ethics of the World Medical Association (Declaration of Helsinki) and was approved by the Medical Ethics Committee of the Second Affiliated Hospital of Fujian Medical University (2019-233, 2021-73).

2.4 MicroRNA sequencing

Corresponding kits were used to extract total RNA/miRNA from wax blocks of the myocardium (RecoverAll™ Total Nucleic Acid Isolation Kit, Ambion, Thermo Fisher Scientific), amniotic fluid (HiPure Universal RNA Mini Kit, Magen) and maternal serum (miRNeasy Mini Kit, QIAGEN). A cDNA library of amniotic fluid-derived miRNAs (QIAseq miRNA Library Kit, QIAGEN) and myocardium-derived and serum-derived miRNAs (TruSeq Small RNA Library Preparation kit, Illumina) was constructed, respectively. The library preparations were sequenced on the Illumina HiSeq 2500 sequencing system (for myocardium-derived and serum-derived miRNAs) and Illumina NovaSeq 6000 system (for amniotic fluid-derived miRNAs).

2.5 Sequencing data processing

Bcl2fastq (bcl2fastq, RRID:SCR_015058) was used to perform recognition on the original image and convert it into the original sequence. The Fastx-toolkit (FASTX-Toolkit, RRID: SCR_005534) was used to evaluate and filter the quality. The expression was normalized using the transcripts read number per million (TPM) and counts per million (CPM) method. DESeq (DESeq, RRID:SCR_000154) and an R package (LIMMA, RRID: SCR_010943) was used to perform differential expression analysis. The threshold values $p < 0.05$ and fold change (FC) ≥ 2 indicated upregulated differentially expressed microRNA (DE-microRNA), while $p < 0.05$ and FC ≤ 0.5 indicated downregulated DE-microRNA.

2.6 Target gene prediction and bioinformatics analysis of differentially expressed microRNAs

MiRWalk 3.0 (miRWalk, RRID:SCR_016509) predicted the target genes of DE-microRNAs. Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analyzes of target genes whose binding probability > 0.95 were performed by g:Profiler (version e104_eg51_p15_3922dba). String (https://cn.string-db.org/) was used to perform protein–protein interaction (PPI)

analysis and establish PPI networks. The plug-in "cytoHubba" of Cytoscape (Cytoscape, RRID:SCR_003032) screened out the top 20 hub genes in the PPI networks.

2.7 Quantitative real-time polymerase chain reaction

Total RNA was extracted by TRIzol (Invitrogen, Thermo Fisher Scientific) and the RecoverAll™ Total Nucleic Acid Isolation Kit (Ambion, Thermo Fisher Scientific) from amniotic fluid, serum and myocardium wax blocks. MiRNA First Strand cDNA Synthesis (Sangon Biotech) was used to reverse transcribe RNA to cDNA. Polymerase chain reaction (PCR) was performed according to the manual of the TB Green Premix Ex Taq kit (TAKARA). All reactions were performed in triplicate. The result was normalized to U6 (Universal U6 Primer F, Sangon Biotech) and calculated using the $2^{-\Delta\Delta Ct}$ method. The primer sequences are shown in Supplementary Table S1.

2.8 Double luciferase reporter gene assay

The wild type (WT) and mutant (MU) of the target 3′-UTR were cloned and inserted into the pSI-Check2 vector (Promega), and the successful construction of the plasmid was verified by sequencing. Before transfection, 293T cells and the target plasmid were seeded in a 96-well plate. Then, LipoFiter 3.0 (Hanbio) was used to cotransfect the WT and MU plasmids with the miR-1-3p plasmid. The dual luciferase reporter gene assay system (Promega) was used to evaluate the activities of firefly luciferase and Renilla luciferase.

2.9 Statistical analysis

Normally distributed data are expressed as the mean ± standard deviation (mean ± SD); nonnormally distributed data are expressed as the median (lower quartile, upper quartile). The comparison between the two groups was performed by *t* test or Mann–Whitney test. $p < 0.05$ was regarded as significantly different.

3 Results

3.1 Ventricular septal defect group microRNAs expression profile and bioinformatics analysis

Of the myocardium and amniotic fluid samples (divided into NC group, non-VSD group, and VSD group) and maternal serum samples (divided into NC group and VSD group), seven samples were selected for miRNA sequencing in each group.

3.1.1 MicroRNAs expression profile

Compared with the NC group, there were 53 myocardium-derived DE-microRNAs in the VSD group, of which 23 were upregulated (miR-497-3p, miR-7b, etc.) and 30 were downregulated (miR-1-3p, miR-1b, miR-293-5p and miR-3580-3p, etc., Supplementary Table S2; Figures 2A,B). There were 34 DE-microRNAs in the non-VSD group, including 9 upregulated and 25 downregulated microRNAs (Supplementary Table S3).

There were 33 amniotic fluid-derived DE-microRNAs in the VSD group, including 22 upregulated (miR-15b-5p, miR-1b, etc.) and 11 downregulated microRNAs (miR-1843b-5p, miR-299a-5p, etc., Supplementary Table S4; Figures 2C,D). There were 48 DE-microRNAs in the non-VSD group, including 39 upregulated and 9 downregulated microRNAs (Supplementary Table S5).

Twenty-seven maternal serum-derived DE-microRNAs were detected in the VSD group, including 20 upregulated (miR-129-5p, miR-206-3p, miR-293-5p, miR-3580-3p, miR-494-3p, etc.), and 7 downregulated microRNAs (miR-208a-3p, miR-208b-3p, etc., Supplementary Table S6; Figures 2E,F).

3.1.2 Target gene prediction and bioinformatics analysis of differentially expressed microRNAs from different samples

Considering the possible effects of DMO on microRNA expression in the fetus, DE-microRNAs in the myocardium and amniotic fluid of the non-VSD group were excluded from the target gene prediction and bioinformatics analysis, and the unique DE-microRNAs in the VSD group were analyzed.

There were 35 unique DE-microRNAs in the myocardium of the VSD group, of which 20 were upregulated and 15 were downregulated. A total of 10,429 target genes of these DE-microRNAs were identified by miRWalk3.0, of which 207 related to VSD were recorded in disease-related databases (including OMIM, KEGG Disease database and GWAS Catalog database). Twenty unique DE-microRNAs of the VSD group were discovered in amniotic fluid, 10 of which were upregulated and 10 of which were downregulated. A total of 9,352 target genes were predicted, of which 179 were related to VSD. Maternal serum-derived DE-microRNAs predicted a total of 11,317 target genes, of which 231 were related to VSD.

GO analysis of the abovementioned VSD-related target genes showed that the target genes of different samples were significantly enriched in biological process (BP), such as heart development, cardiac chamber development, cardiac chamber morphogenesis, and circulatory system development. Cellular component (CC) analysis showed that they are mainly located in

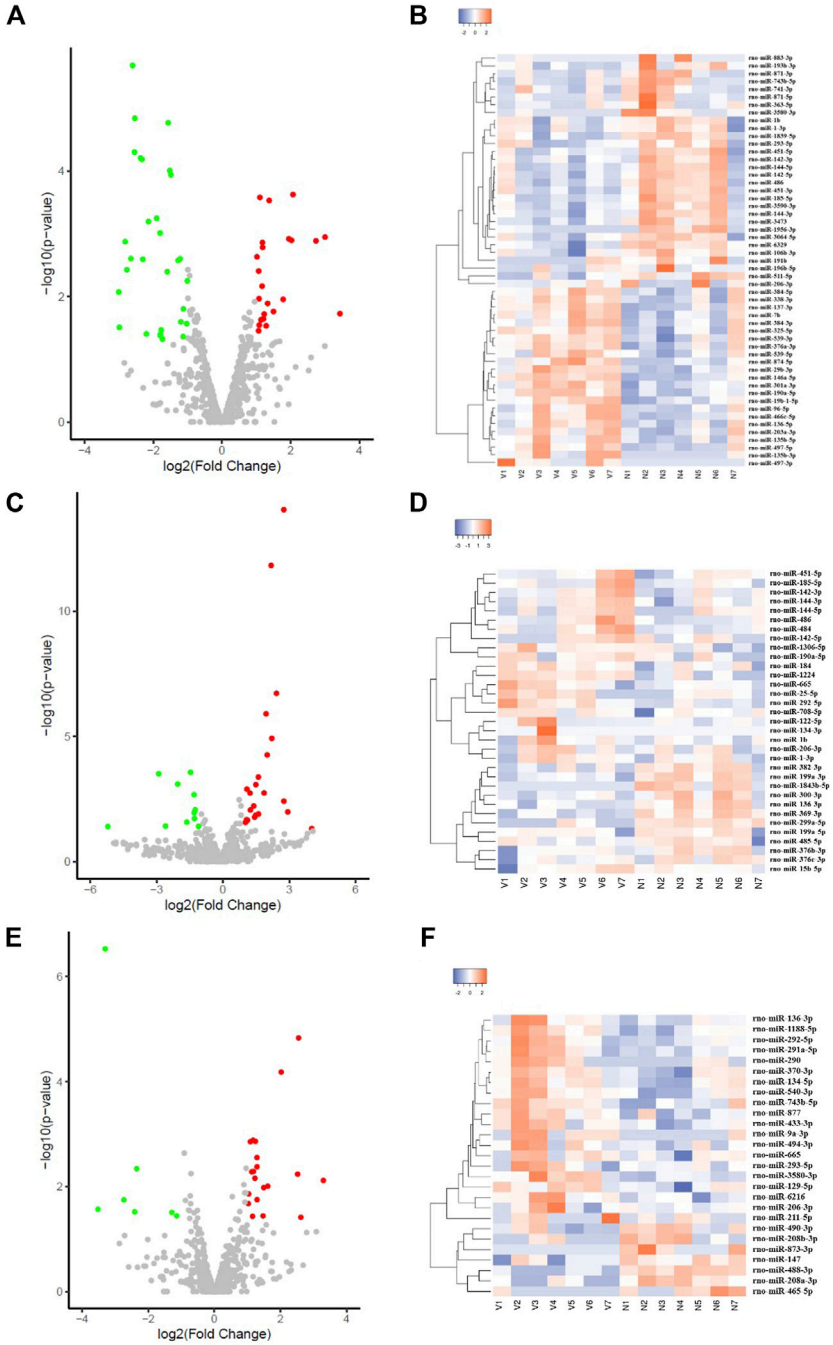


FIGURE 2
Volcano plot and cluster heatmap of DE-microRNAs in the VSD group **(A,C,E)** are all volcano plots; **(B,D,F)** are cluster heatmaps; **(A,B)** represents myocardium-derived DE-microRNAs; **(C,D)** represents amniotic fluid-derived DE-microRNAs; **(E,F)** represents maternal serum-derived DE-microRNAs.

the membrane-bounded organelle, nucleoplasm and nucleus and perform molecular function (MF), such as chromatin binding, transcription factor binding and protein binding. KEGG analysis indicated that in addition to participating in cancer-related pathways, the abovementioned genes were also enriched in pathways that regulate pluripotency of stem cells, Ras related to heart disease, and MAPK related to cardiomyocyte proliferation (Supplementary Figure S1).

TABLE 2 DE-microRNAs related to hub genes.

	Related DE-microRNAs
Myocardium	miR-142-3p, miR-1839-5p, miR-185-5p, miR-301a-3p, miR-3064-5p, miR-325-5p, miR-3473, miR-3580-3p, miR-363-5p, miR-384-5p, miR-497-3p, miR-741-3p, miR-874-5p, miR-96-5p
Amniotic fluid	miR-122-5p, miR-134-3p, miR-184, miR-1843b-5p, miR-199a-5p, miR-299a-5p, miR-665
Serum	miR-3580-3p, miR-370-3p, miR-433-3p, miR-494-3p, miR-6216, miR-665, miR-873-3p, miR-877

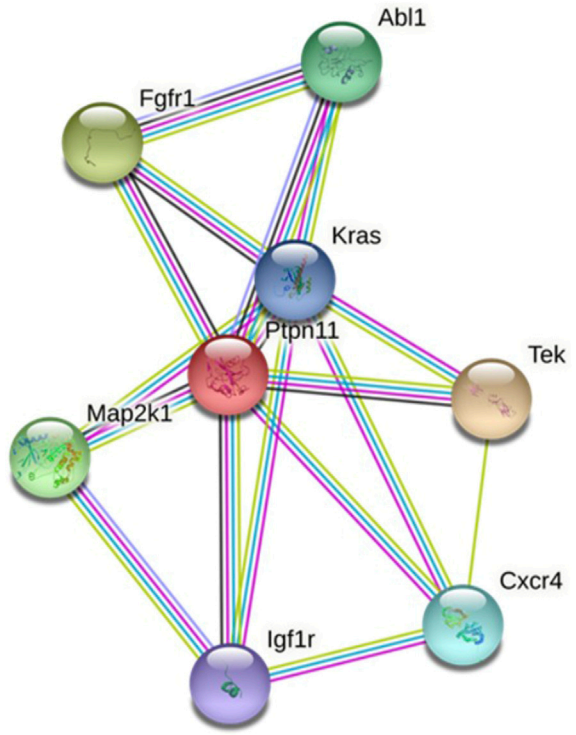


FIGURE 3
PPI network of proteins coded by key target genes.

3.1.3 Construction and analysis of protein–protein interaction network

The PPI network of the protein expression of target genes was constructed. Combining the 7 algorithms of cytoHubba (including MCC, DMNC, MNC, Degree, EPC, Closeness and Radiality), the top 20 hub genes in the PPI network were screened out, and the intersection was assessed. The hub genes in the myocardium are *Kras*, *Map2k1*, *Fgfr1*, *Ptpn11* and *Igf1r*, and the corresponding myocardium-derived DE-microRNAs are miR-3580-3p, miR-497-3p and miR-96-5p, etc. The hub genes in amniotic fluid are *Map2k1*, *Abl1*, *Cxcr4* and *Tek*, corresponding to the amniotic fluid-derived DE-microRNAs are miR-199a-5p and miR-184, etc. The hub genes in maternal serum are *Ptpn11*, *Fgfr1*, *Igf1r* and *Map2k1*, and the corresponding maternal serum-derived DE-microRNAs are miR-3580-3p and miR-494-3p, etc. (Supplementary Table S7; Table 2).

Hub genes were synthesized in three kinds of samples to construct a PPI network, elucidating the relationship and interaction among these proteins (Figure 3).

3.1.4 Quantitative real-time polymerase chain reaction confirmed the expression of differentially expressed microRNAs

The expression of DE-microRNAs, which were differentially expressed in both the myocardium and circulation, or were associated with VSD according to previous studies, or related to hub genes, were verified in myocardium, amniotic fluid and maternal serum (Figure 4). Each group consisted of nine samples. MiR-1-3p, miR-1b, and miR-293-5p are DE-microRNAs differentially expressed in both the myocardium and amniotic fluid/maternal serum. Among them, miR-1-3p and miR-1b are downregulated in the myocardium but upregulated in the amniotic fluid. The expression of miR-293-5p is downregulated in myocardium but upregulated in maternal serum. MiR-3580-3p has a similar expression trend as miR-293-5p in the myocardium and serum, and the difference was not significant. In addition, miR-206-3p was significantly overexpressed in amniotic fluid and maternal serum. MiR-185-5p and miR-96-5p were differentially expressed in myocardium. MiR-15b-5p and miR-184 were differentially expressed in amniotic fluid. MiR-208b-3p, miR-877 and miR-433-3p were differentially expressed in maternal serum. MiR-142-3p, miR-122-5p and miR-134-3p showed differential expression in both VSD group and non-VSD group.

3.2 Expression of amniotic fluid-derived differentially expressed microRNAs in clinical samples

Amniotic fluid-derived DE-microRNAs that were verified by quantitative real-time polymerase chain reaction (qRT–PCR) were selected. BLAST (v2.8.1) was used to identify homologous human miRNAs and confirm conservation (Supplementary Table S8).

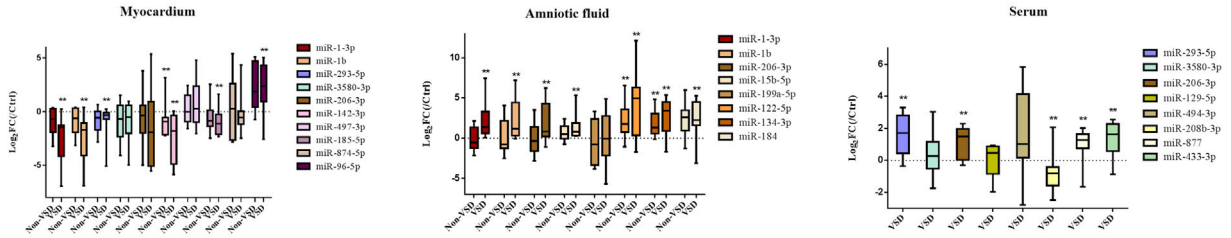


FIGURE 4
qRT–PCR verified the expression of DE-microRNAs. (Compared with NC group: *$p < 0.05$, **$p < 0.01$. $n = 9$)

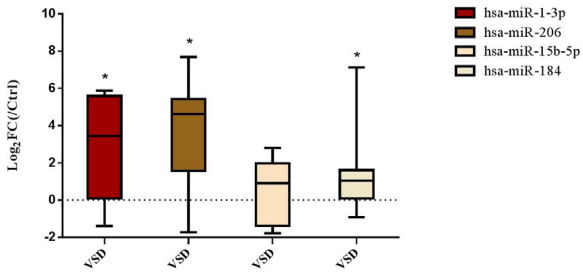


FIGURE 5
Expression of amniotic fluid-derived DE-microRNAs in clinical samples. (Compared with NC group: *$p < 0.05$. $n = 7$)

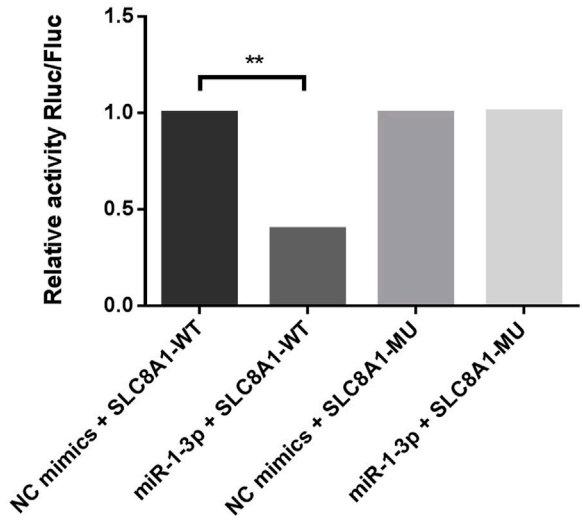


FIGURE 6
Dual luciferase reporter assay verifies that SLC8A1 is the direct target of miR-1-3p. (MiR-1-3p significantly reduced the luciferase activity of plasmids containing the WT SLC8A1 3′UTR. Fluc, firefly luciferase activity; Rluc, Renilla luciferase activity; WT, wild type; MU, mutant. **$p < 0.01$)

Seven human fetuses were diagnosed with perimembranous VSD or muscular VSD by prenatal ultrasound examination, and the diagnosis was confirmed by follow-up after birth. The size of defect ranged from 1.2 to 4.8 mm. It was confirmed in clinical amniotic fluid samples that miR-1-3p, miR-206 and miR-184 were also overexpressed in clinical amniotic fluid samples (Figure 5).

3.3 *SLC8A1* is the target gene of miR-1-3p

MiR-1-3p, which is differentially expressed in both the myocardium and amniotic fluid, was selected for target gene prediction. Solute carrier family 8 member A1 (*SLC8A1*, also known as sodium-calcium exchanger, *NCX1*) may be the target gene of miR-1-3p. The 3′-UTR of *SLC8A1* was cloned and inserted into the pSI-Check2 vector to construct a recombinant plasmid. The dual luciferase reporter system was used to detect the relative luciferase activity. It was demonstrated that miR-1-3p can significantly reduce the luciferase activity of the plasmid containing the wild-type *SLC8A1* 3′UTR but has no significant effect on the plasmid containing the mutant *SLC8A1* 3′UTR, indicating that miR-1-3p can inhibit luciferase activity by binding to the 3′UTR of *SLC8A1*. Consequently, *SLC8A1* is the direct target of miR-1-3p (Figure 6).

4 Discussion

In this study, we detected and analyzed the expression profiles of miRNAs in the myocardium, amniotic fluid and maternal serum of VSD fetal rats and found that miR-1-3p, miR-1b and miR-293-5p were differentially expressed in amniotic fluid/maternal serum and the myocardium. The expression of amniotic fluid-derived DE-microRNAs (miR-1-3p, miR-206 and miR-184) was confirmed in clinical samples. Then, it was predicted and verified that one of the target genes of miR-1-3p is *SLC8A1/NCX1*, which is related to cardiomyocyte apoptosis, indicating that circulating miRNAs involved in the regulation of VSD may be derived from the

myocardium, providing a theoretical basis for the use of circulating miRNAs to assist in the diagnosis of VSD.

MiRNAs participate in heart development, and their dysregulation may be related to CHD (Bruneau, 2008). The absence of the miRNA processing enzyme Dicer can lead to abnormal formation of the heart outflow tract and chambers in mammals, so miRNA is essential for heart development (Saxena and Tabin, 2010). Previous studies have revealed that a variety of miRNAs are involved in processes such as heart development and cardiomyocyte proliferation and differentiation, which are closely related to CHDs, including VSD (Wang et al., 2016; Toni et al., 2020; Zhuang et al., 2020).

MiR-1 is a myocardial and skeletal muscle-specific miRNA (Townley-Tilson et al., 2010). Li et al. (2013) analyzed the expression of miR-1-1 in the human heart and found that the overexpression of *GJA1* and *SOX9* was correlated with the decrease in miR-1-1 in VSD, indicating that miR-1-1 regulates the above target genes related to the pathogenesis of VSD. In patients with tetralogy of Fallot (TOF), miR-1 and miR-133 are significantly downregulated, which is predicted to affect the development and function of the heart by regulating genes such as *KCNJ2*, *FBN2*, *SLC38A3* and *TNNI1* (Grunert et al., 2019). Therefore, the downregulation of miR-1 expression is closely associated with heart development. In contrast, it was reported that the overexpression of miR-1 and miR-133 effectively promotes the reprogramming of fibroblasts to cardiomyocytes, reduces apoptosis and increases the viability of P19 cells differentiated into cardiomyocytes (Liu et al., 2017; Riching et al., 2021). In a study on circulating miRNA, Stoica et al. (2020) detected plasma miR-1 in children with CHDs who underwent surgery and found that high expression of plasma miR-1 was associated with longer intensive care time, more serious cardiovascular events and an increasing ventilation index. Hence, miR-1 can be an indicator to evaluate the prognosis of children after surgery.

Our results supported the above reports. In the myocardium of VSD fetal rats, the expression of miR-1-3p and miR-1b, both of which belong to the mir-1 family, was significantly downregulated but was upregulated in the amniotic fluid of rats and humans. This finding of circulating miR-1 is similar to the findings of the study of Stoica et al. (2020). Furthermore, a dual luciferase reporter system was used to verify that miR-1-3p participates in the regulation of VSD by directly targeting *SLC8A1/NCX1*. *SLC8A1/NCX1* is highly expressed in the myocardium and is responsible for expelling Ca^{2+} from cardiomyocytes during diastole, which is the main mechanism by which cardiomyocytes return to a resting state after excitation (Wakimoto et al., 2000; Fagerberg et al., 2014). Studies have shown that *SLC8A1/NCX1* deficiency results in death and cardiomyocyte apoptosis in mice in the middle embryonic stage, suggesting that *SLC8A1/NCX1* is essential for fetuses and embryonic cardiomyocyte survival (Wakimoto et al., 2000). In addition, *SLC8A1/NCX1* can induce a variety of heart defects, such as VSD and arrhythmia (Raveau et al., 2012). It was reported that *SLC8A1/NCX1* elevated in ductus arteriosus (DA)-dependent CHDs, and play a role in preventing DA functional closure and delaying the anatomical closure process (Li et al., 2017). Previous studies showed that overexpression of *SLC8A1/NCX1* was associated with myocardial hypertrophy (Tyser et al., 2016; Ottolia et al., 2021). Roos et al. (2007) indicated that increased expression of the *SLC8A1/NCX1* directly led to heart hypertrophy, and the magnitude of the hypertrophy and pathology increased with increasing *SLC8A1/NCX1* expression. It is almost always accompanied by loss of myocytes (apoptosis and/or necrosis) in hypertrophy progresses (Roos et al., 2007). He et al. (2015) also proved that *SLC8A1/NCX1* in mouse with myocardial hypertrophy is increased and positively correlated with the cardiomyocyte apoptosis. Therefore, the overexpression of *SLC8A1/NCX1* could lead to hypertrophy, apoptosis and/or necrosis. Our study further confirmed that miR-1-3p directly targets *SLC8A1/NCX1* in the myocardium and participates in the regulation of cardiac electrophysiological activities and cardiomyocyte apoptosis, which is closely related to the pathogenesis of VSD.

This study also verified the upregulation of miR-293-5p, miR-15b-5p, miR-206, miR-184, miR-877 and miR-433-3p and the downregulation of miR-208b-3p in circulation. The miRNAs mentioned above have been shown to regulate the pluripotency of stem cells, the differentiation and apoptosis of cardiomyocytes and other processes related to the occurrence of heart diseases. MiR-293-5p was differentially expressed in both the myocardium and maternal serum. MiR-293-5p is a member of the miR-290–295 cluster that is most abundantly expressed in rat pluripotent stem cells and is involved in the regulation of pluripotency and reprogramming in rats (Sherstyuk et al., 2017). MiR-15b may be upregulated due to apoptosis induced by cardiac ischaemia/reperfusion injury (Liu et al., 2012; Liu et al., 2014). Zhao et al. (2019) found that miR-15b can directly target the 3′-UTR of *SETD3* to inhibit its expression and myoblast differentiation, while downregulation of miR-15b can promote myoblast differentiation. MiR-208b is mainly expressed in the embryonic heart and skeletal muscle. The expression of miR-208b decreases during embryonic development to adulthood, while an increase in the adult heart may be related to pathological remodelling of the heart in dilated cardiomyopathy (Zhou et al., 2017). MiR-208b-3p participated in miR-208b-3p/Med13/Wnt/β-catenin signaling pathway axis and against hypoxia/reoxygenation injury. MiR-206 is significantly decreased in the peripheral blood of TOF patients after surgery (Abu-Halima et al., 2017). MiR-184 was increased in cardiomyocytes suffering from oxidative stress, and inhibition of miR-184 could inhibit cardiomyocyte apoptosis (Zou et al., 2020a). MiR-877 were differentially expressed in the right ventricle of pulmonary arterial hypertension rats (Joshi et al., 2016). MiR-433-3p upregulated in patients with critical coronary stenosis (Infante et al., 2019). Accordingly, our study indicated that DE-microRNAs (miR-293-5p, miR-15b-5p, miR-206, miR-208b-3p, etc.) in the circulation of amniotic fluid and maternal serum are

related to VSD, which is helpful for the prenatal diagnosis of fetal VSD.

Scholars have proposed that circulating miRNAs can be packaged in microvesicles, exosomes, and apoptotic bodies or combined with RNA-binding proteins or lipoprotein complexes to maintain stability and prevent degradation (Creemers et al., 2012; Tsatsaronis et al., 2018; Mori et al., 2019). Previous studies of arrhythmogenic cardiomyopathy, acute myocardial infarction, liver injury, tumours and other diseases have found that miRNAs are underexpressed in tissue but overexpressed in circulation, indicating that circulating DE-microRNAs may originate from damaged tissues or apoptotic cells (Bueno Marinas et al., 2020; Zou et al., 2020b). Apoptosis play an important role in the pathogenesis of VSD. The increase of apoptosis would result in cardiac defects, including VSD (Chen et al., 2009). Studies revealed that increased apoptosis observed in interventricular septum or outflow tract cushions was likely contributing to VSD (Gaussin et al., 2002; Liang et al., 2007). It was reported that apoptotic cells were increased in VSD models. Kumar et al. (2007) found that the apoptotic cells were significantly increased in the ventricular myocardium of the streptozotocin-induced diabetic mice embryos. The increased apoptosis was also confirmed in several genetically mutated models of VSDs (Chen et al., 2009; Huang et al., 2012; Liu et al., 2018). In this study, miR-1-3p, miR-1b and miR-293-5p all showed similar low expression in the myocardium of VSD fetal mice and high expression in amniotic fluid and maternal serum, which supports this hypothesis and indicates that the miRNAs mentioned above may be actively or passively released into the circulation by cardiomyocytes, suggesting myocardial injury or apoptosis. Therefore, the overexpression of miRNAs such as miR-1-3p, miR-1b and miR-293-5p in the circulation is a direct and powerful clue related to fetal VSD.

This study has certain limitations. Some miRNAs considered to be related to VSD in previous studies (such as miR-133, miR-181c, etc.) were not greatly differentially expressed in this study. The diversity may be because the samples in this study were derived from fetal mice and related body fluids, while some of the previous studies selected different samples, such as VSD patients after birth and cells cultured *in vitro*, or different pretreatment methods, such as whole blood and exosomes of amniotic fluid/serum. This study verified that one of the target genes of miR-1-3p is *SLC8A1/NCX1* by the dual luciferase reporter system. The expression and regulatory pathways of miR-1-3p and *SLC8A1/NCX1* in VSD fetus remains to be further explored.

5 Conclusion

Our study comprehensively analyzed the expression profile of miRNAs in the myocardium, amniotic fluid and maternal serum and found that miR-1-3p, miR-1b and miR-293-5p were differentially expressed in both the myocardium and circulation and may be released by necrotic or apoptotic cardiomyocytes and thus appear to be downregulated in the myocardium and upregulated in the circulation. MiR-1-3p targeting *SLC8A1* participates in the regulation of VSD, and the same expression was confirmed in human amniotic fluid, which indicates that miRNA has great potential to become a biomarker for the prenatal diagnosis of VSD. This finding provided a foundation and broadened the horizon of the use of circulating miRNAs to assist in the prenatal diagnosis of VSD. The value in clinical application is expected to be verified in long-term follow-up of new cohorts in the future.

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 GEO - GSE194240.

Ethics statement

The studies involving human participants were reviewed and approved by the Medical Ethics Committee of the Second Affiliated Hospital of Fujian Medical University. The patients/participants provided their written informed consent to participate in this study. The animal study was reviewed and approved by the Medical Ethics Committee of the Second Affiliated Hospital of Fujian Medical University.

Author contributions

YY contributed to conceptualization, formal analysis, investigation, visualization, and writing—original draft. HY and XL contributed to formal analysis and investigation. SY contributed to conceptualization. HS and SW contributed to resources. XW contributed to investigation. GL contributed to conceptualization and writing—review and editing.

Funding

This study was sponsored by the Quanzhou City Science and Technology Program of China (2021C059R) and the Famous Teacher Studio of Quanzhou Medical College. The funding body played no role in the design of the study or collection, analysis, or interpretation of data or in writing the manuscript.

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.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.899034/full#supplementary-material

References

Abu-Halima, M., Meese, E., Keller, A., Abdul-Khaliq, H., and Rädle-Hurst, T. (2017). Analysis of circulating microRNAs in patients with repaired Tetralogy of Fallot with and without heart failure. *J. Transl. Med.* 15, 156. doi:10.1186/s12967-017-1255-z

Bruneau, B. G. (2008). The developmental genetics of congenital heart disease. *Nature* 451, 943–948. doi:10.1038/nature06801

Bueno Marinas, M., Celeghin, R., Cason, M., Bariani, R., Frigo, A. C., Jager, J., et al. (2020). A microRNA expression profile as non-invasive biomarker in a large arrhythmogenic cardiomyopathy cohort. *Int. J. Mol. Sci.* 21, E1536. doi:10.3390/ijms21041536

Chen, Q., Chen, H., Zheng, D., Kuang, C., Fang, H., Zou, B., et al. (2009). Smad7 is required for the development and function of the heart. *J. Biol. Chem.* 284, 292–300. doi:10.1074/jbc.M807233200

Cox, K., Algaze-Yojay, C., Punn, R., and Silverman, N. (2020). The natural and unnatural history of ventricular septal defects presenting in infancy: An echocardiography-based review. *J. Am. Soc. Echocardiogr.* 33, 763–770. doi:10.1016/j.echo.2020.01.013

Creemers, E. E., Tijsen, A. J., and Pinto, Y. M. (2012). Circulating microRNAs: Novel biomarkers and extracellular communicators in cardiovascular disease? *Circ. Res.* 110, 483–495. doi:10.1161/CIRCRESAHA.111.247452

Fagerberg, L., Hallström, B. M., Oksvold, P., Kampf, C., Djureinovic, D., Odeberg, J., et al. (2014). Analysis of the human tissue-specific expression by genome-wide integration of transcriptomics and antibody-based proteomics. *Mol. Cell. Proteomics* 13, 397–406. doi:10.1074/mcp.M113.035600

Gaussin, V., Van de Putte, T., Mishina, Y., Hanks, M. C., Zwijsen, A., Huylebroeck, D., et al. (2002). Endocardial cushion and myocardial defects after cardiac myocyte-specific conditional deletion of the bone morphogenetic protein receptor ALK3. *Proc. Natl. Acad. Sci. U. S. A.* 99, 2878–2883. doi:10.1073/pnas.042390499

Grunert, M., Appelt, S., Dunkel, I., Berger, F., and Sperling, S. R. (2019). Altered microRNA and target gene expression related to Tetralogy of Fallot. *Sci. Rep.* 9, 19063. doi:10.1038/s41598-019-55570-4

Grunert, M., Dorn, C., Cui, H., Dunkel, I., Schulz, K., Schoenhals, S., et al. (2016). Comparative DNA methylation and gene expression analysis identifies novel genes for structural congenital heart diseases. *Cardiovasc. Res.* 112, 464–477. doi:10.1093/cvr/cvw195

He, J., Cai, Y., Luo, L. M., and Wang, R. (2015). Expression of Wnt and NCX1 and its correlation with cardiomyocyte apoptosis in mouse with myocardial hypertrophy. *Asian pac. J. Trop. Med.* 8, 930–936. doi:10.1016/j.apjtm.2015.10.002

Huang, X., Huang, F., Yang, D., Dong, F., Shi, X., Wang, H., et al. (2012). Expression of microRNA-122 contributes to apoptosis in H9C2 myocytes. *J. Cell. Mol. Med.* 16, 2637–2646. doi:10.1111/j.1582-4934.2012.01577.x

Hui, L., Slonim, D. K., Wick, H. C., Johnson, K. L., and Bianchi, D. W. (2012). The amniotic fluid transcriptome: A source of novel information about human fetal development. *Obstet. Gynecol.* 119, 111–118. doi:10.1097/AOG.0b013e31823d4150

Infante, T., Forte, E., Punzo, B., Cademartiri, F., Cavaliere, C., Soricelli, A., et al. (2019). Correlation of circulating miR-765, miR-93-5p, and miR-433-3p to obstructive coronary heart disease evaluated by cardiac computed tomography. *Am. J. Cardiol.* 124, 176–182. doi:10.1016/j.amjcard.2019.04.016

Islas, J. F., and Moreno-Cuevas, J. E. (2018). A MicroRNA perspective on cardiovascular development and diseases: An update. *Int. J. Mol. Sci.* 19, 2075. doi:10.3390/ijms19072075

Jin, Y., Ai, L., Chai, X., Tang, P., Zhang, W., Yang, L., et al. (2021). Maternal circulating exosomal miRNAs as non-invasive biomarkers for the prediction of fetal ventricular septal defect. *Front. Genet.* 12, 717208. doi:10.3389/fgene.2021.717208

Joshi, S. R., Dhagia, V., Gairhe, S., Edwards, J. G., McMurtry, I. F., and Gupte, S. A. (2016). MicroRNA-140 is elevated and mitofusin-1 is downregulated in the right ventricle of the Sugen5416/hypoxia/normoxia model of pulmonary arterial hypertension. *Am. J. Physiol. Heart Circ. Physiol.* 311, H689–H698. doi:10.1152/ajpheart.00264.2016

Kumar, S. D., Dheen, S. T., and Tay, S. S. (2007). Maternal diabetes induces congenital heart defects in mice by altering the expression of genes involved in cardiovascular development. *Cardiovasc. Diabetol.* 6, 34. doi:10.1186/1475-2840-6-34

Li, J., Cao, Y., Ma, X. J., Wang, H. J., Zhang, J., Luo, X., et al. (2013). Roles of miR-1-1 and miR-181c in ventricular septal defects. *Int. J. Cardiol.* 168, 1441–1446. doi:10.1016/j.ijcard.2012.12.048

Li, M., Jiang, C., Ye, L., Wang, S., Zhang, H., Liu, J., et al. (2017). The role of Na+/Ca2+ exchanger 1 in maintaining ductus arteriosus patency. *Sci. Rep.* 7, 9826. doi:10.1038/s41598-017-10377-z

Liang, X., Sun, Y., Schneider, J., Ding, J. H., Cheng, H., Ye, M., et al. (2007). Pinch1 is required for normal development of cranial and cardiac neural crest-derived structures. *Circ. Res.* 100, 527–535. doi:10.1161/01.RES.0000259041.37059.8c

Liu, F., Liu, X., Xu, Z., Yuan, P., Zhou, Q., Jin, J., et al. (2018). Molecular mechanisms of Ellis-van Creveld gene variations in ventricular septal defect. *Mol. Med. Rep.* 17 (1), 1527–1536. doi:10.3892/mmr.2017.8088

Liu, L. F., Liang, Z., Lv, Z. R., Liu, X. H., Bai, J., Chen, J., et al. (2012). MicroRNA-15a/b are up-regulated in response to myocardial ischemia/reperfusion injury. *J. Geriatr. Cardiol.* 9, 28–32. doi:10.3724/SP.J.1263.2012.00028

Liu, L., Yuan, Y., He, X., Xia, X., and Mo, X. (2017). MicroRNA-1 upregulation promotes myocardiocyte proliferation and suppresses apoptosis during heart development. *Mol. Med. Rep.* 15, 2837–2842. doi:10.3892/mmr.2017.6282

Liu, L., Zhang, G., Liang, Z., Liu, X., Li, T., Fan, J., et al. (2014). MicroRNA-15b enhances hypoxia/reoxygenation-induced apoptosis of cardiomyocytes via a mitochondrial apoptotic pathway. *Apoptosis* 19, 19–29. doi:10.1007/s10495-013-0899-2

Meng, X., Zhang, P., and Zhang, L. (2020). Fetal hypoxia impacts on proliferation and differentiation of sca-1(+) cardiac progenitor cells and maturation of cardiomyocytes: A role of MicroRNA-210. *Genes* 11, 328. doi:10.3390/genes11030328

Mori, M. A., Ludwig, R. G., Garcia-Martin, R., Brandão, B. B., and Kahn, C. R. (2019). Extracellular miRNAs: From biomarkers to mediators of physiology and disease. *Cell Metab.* 30, 656–673. doi:10.1016/j.cmet.2019.07.011

Ottolia, M., John, S., Hazan, A., and Goldhaber, J. I. (2021). The cardiac Na(+)-Ca(2+) exchanger: From structure to function. *Compr. Physiol.* 12, 2681–2717. doi:10.1002/cphy.c200031

Panni, S., Lovering, R. C., Porras, P., and Orchard, S. (2020). Non-coding RNA regulatory networks. *Biochim. Biophys. Acta. Gene Regul. Mech.* 1863, 194417. doi:10.1016/j.bbagrm.2019.194417

Pulignani, S., and Andreassi, M. G. (2019). MicroRNAs and congenital heart disease: Where are we now? *Rev. Esp. Cardiol.* 72, 7–9. doi:10.1016/j.rec.2018.06.030

Raveau, M., Lignon, J. M., Nalesso, V., Duchon, A., Groner, Y., Sharp, A. J., et al. (2012). The App-Runx1 region is critical for birth defects and electrocardiographic dysfunctions observed in a Down syndrome mouse model. *PLoS Genet.* 8, e1002724. doi:10.1371/journal.pgen.1002724

Riching, A. S., Danis, E., Zhao, Y., Cao, Y., Chi, C., Bagchi, R. A., et al. (2021). Suppression of canonical TGF-β signaling enables GATA4 to interact with H3K27me3 demethylase JMJD3 to promote cardiomyogenesis. *J. Mol. Cell. Cardiol.* 153, 44–59. doi:10.1016/j.yjmcc.2020.12.005

Roos, K. P., Jordan, M. C., Fishbein, M. C., Ritter, M. R., Friedlander, M., Chang, H. C., et al. (2007). Hypertrophy and heart failure in mice overexpressing the cardiac sodium-calcium exchanger. *J. Card. Fail.* 13, 318–329. doi:10.1016/j.cardfail.2007.01.004

Sabour, D., Machado, R. S. R., Pinto, J. P., Rohani, S., Sahito, R. G. A., Hescheler, J., et al. (2018). Parallel genome-wide profiling of coding and non-coding RNAs to identify novel regulatory elements in embryonic and maturated heart. *Mol. Ther. Nucleic Acids* 12, 158–173. doi:10.1016/j.omtn.2018.04.018

Saxena, A., and Tabin, C. J. (2010). miRNA-processing enzyme Dicer is necessary for cardiac outflow tract alignment and chamber septation. *Proc. Natl. Acad. Sci. U. S. A.* 107, 87–91. doi:10.1073/pnas.0912870107

Sherstyuk, V. V., Medvedev, S. P., Elisaphenko, E. A., Vaskova, E. A., Ri, M. T., Vyatkin, Y. V., et al. (2017). Genome-wide profiling and differential expression of microRNA in rat pluripotent stem cells. *Sci. Rep.* 7, 2787. doi:10.1038/s41598-017-02632-0

Smith, T., Rajakaruna, C., Caputo, M., and Emanueli, C. (2015). MicroRNAs in congenital heart disease. *Ann. Transl. Med.* 3, 333. doi:10.3978/j.issn.2305-5839.2015.12.25

Song, Y., Higgins, H., Guo, J., Harrison, K., Schultz, E. N., Hales, B. J., et al. (2018). Clinical significance of circulating microRNAs as markers in detecting and predicting congenital heart defects in children. *J. Transl. Med.* 16, 42. doi:10.1186/s12967-018-1411-0

Spicer, D. E., Hsu, H. H., Co-Vu, J., Anderson, R. H., and Fricker, F. J. (2014). Ventricular septal defect. *Orphanet J. Rare Dis.* 9, 144. doi:10.1186/s13023-014-0144-2

Stoica, S. C., Dorobantu, D. M., Vardeu, A., Biglino, G., Ford, K. L., Bruno, D. V., et al. (2020). MicroRNAs as potential biomarkers in congenital heart surgery. *J. Thorac. Cardiovasc. Surg.* 159, 1532–1540. doi:10.1016/j.jtcvs.2019.03.062

Thomford, N. E., Dzobo, K., Yao, N. A., Chimusa, E., Evans, J., Okai, E., et al. (2018). Genomics and Epigenomics of congenital heart defects: Expert review and lessons learned in africa. *OMICS* 22, 301–321. doi:10.1089/omi.2018.0033

Toni, L. S., Hailu, F., and Sucharov, C. C. (2020). Dysregulated micro-RNAs and long noncoding RNAs in cardiac development and pediatric heart failure. *Am. J. Physiol. Heart Circ. Physiol.* 318, H1308–H1315. doi:10.1152/ajpheart.00511.2019

Townley-Tilson, W. H., Callis, T. E., and Wang, D. (2010). MicroRNAs 1, 133, and 206: Critical factors of skeletal and cardiac muscle development, function, and disease. *Int. J. Biochem. Cell Biol.* 42, 1252–1255. doi:10.1016/j.biocel.2009.03.002

Tsatsaronis, J. A., Franch-Arroyo, S., Resch, U., and Charpentier, E. (2018). Extracellular vesicle RNA: A universal mediator of microbial communication? *Trends Microbiol.* 26, 401–410. doi:10.1016/j.tim.2018.02.009

Tyser, R. C., Miranda, A. M., Chen, C. M., Davidson, S. M., Srinivas, S., and Riley, P. R. (2016). Calcium handling precedes cardiac differentiation to initiate the first heartbeat. *eLife* 5, e17113. doi:10.7554/eLife.17113

van Nisselrooij, A. E. L., Teunissen, A. K. K., Clur, S. A., Rozendaal, L., Pajkrt, E., Linskens, I. H., et al. (2020). Why are congenital heart defects being missed? *Ultrasound Obstet. Gynecol.* 55, 747–757. doi:10.1002/uog.20358

Wakimoto, K., Kobayashi, K., Kuro, O. M., Yao, A., Iwamoto, T., Yanaka, N., et al. (2000). Targeted disruption of Na+/Ca2+ exchanger gene leads to cardiomyocyte apoptosis and defects in heartbeat. *J. Biol. Chem.* 275, 36991–36998. doi:10.1074/jbc.M004035200

Wang, L., Song, G., Liu, M., Chen, B., Chen, Y., Shen, Y., et al. (2016). MicroRNA-375 overexpression influences P19 cell proliferation, apoptosis and differentiation through the Notch signaling pathway. *Int. J. Mol. Med.* 37, 47–55. doi:10.3892/ijmm.2015.2399

Wang, Z., Yang, Y., Xiong, W., Zhou, R., Song, N., Liu, L., et al. (2020). Dexmedetomidine protects H9C2 against hypoxia/reoxygenation injury through miR-208b-3p/Med13/Wnt signaling pathway axis. *Biomed. Pharmacother. = Biomedecine Pharmacother.* 125, 110001. doi:10.1016/j.biopha.2020.110001

Yang, H., Yang, S., Shen, H., Wu, S., Ruan, J., and Lyu, G. (2021). Construction of the amniotic fluid-derived exosomal ceRNA network associated with ventricular septal defect. *Genomics* 113, 4293–4302. doi:10.1016/j.ygeno.2021.11.003

Zhao, M. J., Xie, J., Shu, W. J., Wang, H. Y., Bi, J., Jiang, W., et al. (2019). MiR-15b and miR-322 inhibit SETD3 expression to repress muscle cell differentiation. *Cell Death Dis.* 10, 183. doi:10.1038/s41419-019-1432-5

Zhou, Q., Schötterl, S., Backes, D., Brunner, E., Hahn, J. K., Ionesi, E., et al. (2017). Inhibition of miR-208b improves cardiac function in titin-based dilated cardiomyopathy. *Int. J. Cardiol.* 230, 634–641. doi:10.1016/j.ijcard.2016.12.171

Zhuang, S., Fu, Y., Li, J., Li, M., Hu, X., Zhu, J., et al. (2020). MicroRNA-375 overexpression disrupts cardiac development of Zebrafish (*Danio rerio*) by targeting notch2. *Protoplasma* 257, 1309–1318. doi:10.1007/s00709-020-01490-4

Zou, J. F., Wu, X. N., Shi, R. H., Sun, Y. Q., Qin, F. J., and Yang, Y. M. (2020). Inhibition of microRNA-184 reduces H2O2-mediated cardiomyocyte injury via targeting FBXO28. *Eur. Rev. Med. Pharmacol. Sci.* 24, 11251–11258. doi:10.26355/eurrev_202011_23614

Zou, X., Zhu, D., Zhang, H., Zhang, S., Zhou, X., He, X., et al. (2020). MicroRNA expression profiling analysis in serum for nasopharyngeal carcinoma diagnosis. *Gene* 727, 144243. doi:10.1016/j.gene.2019.144243

frontiers | Frontiers in Genetics

TYPE Review
PUBLISHED 23 August 2022
DOI 10.3389/fgene.2022.949252

OPEN ACCESS

EDITED BY
Xiao Wang,
Konge Larsen ApS, Denmark

REVIEWED BY
Linchong Sun,
Guangdong Academy of Medical Sciences, China
Xiu Mei Ma,
Macau University of Science and Technology, Macao SAR, China
Laura De Clerck,
Ghent University, Belgium

*CORRESPONDENCE
Wei Huang,
huangwei1212520@163.com
Ling Cao,
lzcaoling@163.comhttp://lzcaoling@163.com

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

RECEIVED 20 May 2022
ACCEPTED 18 July 2022
PUBLISHED 23 August 2022

CITATION
Xie Y, Hu H, Liu M, Zhou T, Cheng X, Huang W and Cao L (2022), The role and mechanism of histone lactylation in health and diseases.
Front. Genet. 13:949252.
doi: 10.3389/fgene.2022.949252

The role and mechanism of histone lactylation in health and diseases

Yumei Xie[1], Hongxia Hu[1], Maoting Liu[1], Tingting Zhou[2], Xi Cheng[2], Wei Huang[2]* and Ling Cao[1]*

[1]Department of Nephrology, Sichuan Clinical Research Center for Nephropathy, The Affiliated Hospital of Southwest Medical University, Luzhou, Sichuan, China, [2]Department of Endocrinology and Metabolism, Metabolic Vascular Diseases Key Laboratory of Sichuan Province, The Affiliated Hospital of Southwest Medical University, Luzhou, Sichuan, China

Whether under anaerobic or aerobic conditions, glycolysis results in production of lactate. Increasing evidence suggests that lactate serves as a multifunctional signaling molecule that develops non-metabolic activities in addition to serving as a key metabolite to link glycolysis and oxidative phosphorylation. Histone posttranslational modification patterns (HPTMs) are essential epigenetic processes controlling a variety of biological activities. Proteomics based on mass spectrometry (MS) has been used to progressively reveal new HPTMs. Recent discoveries of histone lactylation modification mediated by lactate and subsequent research demonstrating its involvement in cancer, inflammation, lung fibrosis, and other conditions suggest that it plays a significant role in immune regulation and homeostasis maintenance. This review provides a brief overview of the complicated control of histone lactylation modification in both pathological and physiological conditions.

KEYWORDS

histone lactylation, epigenetic, gene transcription, inflammation, tumor, novel posttranslational modification, lactate

Introduction

As part of the glycolysis process, lactate dehydrogenase (LDH) catalyzes a specific type of hydroxycarboxylic acid. It is produced by pyruvate breakdown in either anaerobic or aerobic conditions (Fletcher, 1907; Rogatzki et al., 2015). L-lactate and D-lactate are the two enantiomers of lactate that reside in the human body (Oh et al., 1979; Levitt and Levitt, 2020). The former is primarily found in human serum, whereas the latter comes from dietary intake. L-lactate is referred to in this article whenever the word lactate is not explicitly used. For quite a long time, lactate was thought to be a waste product of anaerobic glycolysis. However, emerging data suggest that lactate serves as a multifunctional bio-signaling molecule in addition to being a key metabolite connecting glycolysis and oxidative phosphorylation. On the one hand, lactate regulates intracellular and extracellular metabolic processes across the entire body. On the other hand, it also has a variety of biological effects, including anti-inflammation, immunological regulation, and gene expression, through

receptors expressed in different cells and tissues (Ferguson et al., 2018; Brooks, 2020). Histone undergoes posttranslational alterations in both the C-terminal region and projecting N-terminal tails, which play a crucial role in histone modifications. Numerous histone modifications, including acetylation, methylation, and crotonylation, have been documented before the 2019 discovery of histone lactylation. Recent findings by Zhang et al. state that lactate contributed to epigenetic regulation of genes by lactylating histone lysine residues and that lactate was found to be a precursor to histone lysine lactylation (Kla), which stimulated gene transcription from chromatin (Zhang et al., 2019); these findings demonstrate lactylation's critical role in immune regulation and homeostasis maintenance. Furthermore, lactylation modifications have primarily been studied on histones as of late, but we hypothesize that lactylation also occurs on non-histone proteins, similar to other modifications such as crotonylation. Additionally, the "YnLac" chemical reporter, which has an alkynyl functionalized bioorthogonal structure, may detect new lactylation modification sites in non-histone proteins (Sun et al., 2022). In addition, several microbes and plants have also been identified with lactylated global proteins, including non-histone (Gao et al., 2020; Zhang et al., 2021a; Meng et al., 2021; Sun et al., 2022). On the other hand, there are currently no investigations of non-histone lactylation in mammals, and additional research is required. In this review, we hope to share information on histone lactylation and discuss new findings that demonstrate how this process regulates a variety of pathological conditions.

Histone lactylation

Core histones (H2A, H2B, H3, and H4) and linker histones (H1 and H5), which are alkaline, positively charged proteins, are found in histones. The elementary unit of a chromosome is the nucleosome, which consists of a histone octamer with 200 bp of DNA wrapped around it. Histone posttranslational modification (PTM) is a part of epigenetics. The term "epigenetics" refers to changes in gene expression that are heritable during cell division but do not include changes in the DNA coding sequence. Through covalent modification, different acyl groups can be joined to the amino acid residues on histones. Acyl groups have a variety of impacts on how tightly histone and DNA are bound because of various covalent modification forces (Strahl and Allis, 2000; Tessarz and Kouzarides, 2014; Bowman and Poirier, 2015; Zhang et al., 2021b). This variation will be amplified during the gene expression process, finally resulting in different biological signals that cause transcriptional activation or gene silencing (Table 1). So, histone lactylation will result in different biological responses.

Discovery of lactylation

In 2019, Zhang et al. discovered that in human breast cancer cells, there was a mass shift which is 72.021Da on lysine residues by mass spectrometry (Zhang et al., 2019). Because this mass shift matched that produced by the attachment of a lactyl group to the ε-amino group of a lysine residue, the scientists suspected that it was caused by lactate. So, they created a synthetic peptide to see if it had any chemical characteristics with the peptide obtained *in vivo* that caused the mass change. Additionally, they carried out metabolic labeling studies with isotopic glucose and isotopic lactate to support it, which is in accordance with the conclusion of immunoblotting. The data showed that the unique histone modification known as "lysine lactylation" subsequently was derived directly from either exogenous or endogenous lactate. Zhang et al. made a crucial observation that lysine lactylation differed from lysine acetylation in its kinetics. Under the same conditions, lactylation occurs throughout 24 h, whereas acylations reach a stable state at 6 h. This demonstrates that lactylation and acetylation occur at different times. The scientists genetically removed LDH, which catalyzes the conversion of pyruvate to lactate, and discovered that lysine lactylation was completely abrogated (Zhang et al., 2019). This proved that the kinetics was only affected by lactate. All evidence signifies that lactylation is controlled by lactate and that the process takes longer than that of acetylation.

The enzymes of lactylation

Specific enzymes or enzyme complexes known as "writers" and "erasers" control the addition and removal of histone acylations, activating transcriptional signals that are then read by effector proteins known as "readers" to influence downstream signal pathways and start various biological events (Strahl and Allis, 2000; Figlia et al., 2020). Therefore, lysine lactylation, a novel histone modification, should have some comparable components evolved for it to regulate gene expressions, just as the other histone modifications have a kit of their enzymes. The "writer" and "eraser" is a pair of enzymes with some opposing enzymatic activities to install or remove lactyl groups from modified lysine residues. The "reader" is a type of protein that specifically recognizes this modification and translates it into a variety of functional outcomes within the cell. Currently, it has been shown that "writer" and "eraser" are components of the histone acetyltransferases (HATs) and histone deacetylases (HDACs), respectively (Zhang et al., 2019; Moreno-Yruela et al., 2022a). Additionally, it should possess a class of substrates known as "lactyl-CoA", which directly adds a lactyl group to lysine residues (Zhang et al., 2019). Although there is no evidence to support the presence of "lactyl-CoA", *in vivo*, we hypothesize that the enzyme ACSS2, also known as acetyl-CoA

synthase 2, may be able to produce "lactyl-CoA" since the enzymes found in lactylation are highly coincident with those in acetylation, and in acetylation, ACSS2 produces acetyl-CoA. Zhang et al. showed strong p53-dependent, p300-mediated H3 and H4 lactylation and a commensurate impact on transcription, indicating that p300, an acetyltransferase known for mediating histone lactylation, was responsible for lactate inductions of the histone lactylation (Zhang et al., 2019). After that, Moreno-Yruela, C. et al. demonstrated *in vitro* that two families of deacetylases, namely, HDAC1-3 (histone deacetylase 1–3) and SIRT1-3 (silent information regulator1-3) were lysine delactylases (Figlia et al., 2020; Moreno-Yruela et al., 2022a). HDACs are a subclass of lysine deacetylases that may cleave lactyllysine marks. There are 18 enzymes in total among two families of HDACs (Delcuve et al., 2012). They confirmed that HDACs 1 and 3 in cells had a higher impact on modified histone to remove the lactyl group since HDAC1-3 showed substantial activity toward lysine lactylation (Moreno-Yruela et al., 2022a). So far, we have determined that p300, HDAC1-3, and SIRT1-3 are responsible for setting up and removing histone lactylation. These results, above all, constitute a crucial step toward a thorough characterization of the regulatory components of this pathway, even if the substrate of lactylation has not yet been identified.

It is worth noting that lysine D-lactylation is said to occur by a non-enzymatic acyl transfer. Methylglyoxal is a reactive molecule produced during glycolysis. In the glyoxalase pathway, glyoxalase 1 traps it as D-lactylglutathionethen, and then glyoxalase 2 regenerates glutathione by releasing D-lactate. During the process, D-lactylglutathione is reactive to nucleophiles and can transfer the D-lactyl acyl group onto lysine residues (Gaffney et al., 2020). HDACs also cause the elimination of D-lactylation, even though no enzymes are involved in the acyl installation. D-lactylation and L-lactylation have differing catalytic efficiencies when removed by HDACs (Moreno-Yruela et al., 2022b). The varied distribution of HDACs may be the cause of this. In the cell nucleus, HDAC1–3 may target a few proteins that have been discovered to have D-lactylation modifications (Gaffney et al., 2020; Jennings et al., 2021). Additionally, SIRT2, which is largely found in the cytoplasm, is more likely to break down D-lactylation. However, there are still a lot of questions that require additional research.

Mechanisms of lactylation

Similarly, we use acetylation as a model to determine how lysine lactylation occurs (Figure 1). We hypothesize that cells have a "lactate clock" (Zhang et al., 2019). When exogenous or endogenous lactate accumulates to a certain quantity in cells, the "lactate clock" is activated and initiates lysine lactylation. Then, relevant enzymes start the process. First, the "writer" transfers "lactyl-CoA" as a substrate to histone lysine residues. This generates that the degree to which the changed histone binds to the lactyl group alters how tightly it binds to the DNA molecule, indirectly controlling the expression of genes. Then, the "erasers" arise to complete it, preventing histone lysine from having a lasting impact and stopping the whole cycle of histone lactylation. However, the aforementioned procedure is only capable of reaching the speculation stage; hence, more testing is required. Although the exact mechanism of histone lactylation is unknown, Zhang et al. have identified certain particular locations. From mouse bone marrow–derived macrophages (BMDMs) and human HeLa cells, they found 26 and 16 histone lactylation sites, respectively (Zhang et al., 2019). Additionally, some new sites has been found and all the sites will make it much easier for future researchers to analyze histone lactylation (Figure 2). Interestingly, a newly created predictor called FSL-Klanew may help forecast lactylation sites. According to the model's inventor, the model can generate candidates for further experimental approaches in addition to being a cutting-edge tool for lactylation site profiles (Zhang et al., 2019; Jiang et al., 2021a).

The numerous histone modifications are more or less connected because differences between them are fundamentally existent only in the modification groups. Therefore, we suspect that other types of histone modifications, especially acetylation, intersect with lactylation. Before the discovery of lactylation, acetylation and glucose metabolism were investigated. For instance, Wellen et al. discovered that the acetyl-CoA generated by ATP citrate lyase served as a bridge between the acetylation process and glucose metabolism (Wellen et al., 2009). Latham et al. later discovered that lactate may inhibit the activity of deacetylase, regulating gene expressions (Latham et al., 2012). Furthermore, it was found by Moussaieff et al. and Li, L et al. that glycolysis also influenced the co-factor and substrate levels of acetylation and that, in many glycolysis-dependent cells, acetyl-CoA typically altered simultaneously with lactate (Moussaieff et al., 2015; Li et al., 2020). There should be a connection between acetylation and lactylation. In fact, it is indeed the same. The histone lysine lactyltransferase, p300, was experimentally shown by Zhang et al. In addition, Moreno-Yruela, C et al. discovered that a portion of HDACs had a delactylating effect (Zhang et al., 2019; Moreno-Yruela et al., 2022a). Since acetylation and lactylation are regulated by both HATs and HDACs, and it is reasonable to assume that the two are connected. Although the link between lactylation and acetylation has not yet been defined, we can make assumptions based on several investigations (de Ruijter et al., 2003). For instance, lactate can increase the levels of lactylation and acetylation in macrophage HMGB1 (Yang et al., 2022a). Additionally, when glucose is used as a treatment, the levels of lactylation and acetylation both increase within a certain limit (Zhang et al., 2019). Mice that are exposed to cold also have higher levels of lactylation and acetylation (Lu et al., 2022). Lactylation and acetylation levels under hypoxia, however, differ depending on

the type of cell; in Hela cells, lactylation levels increase, while acetylation levels decrease. In contrast, when lactylation levels increase in murine macrophages, there is no change in acetylation levels (Zhang et al., 2019). Therefore, we believe that it is unreasonable to simply categorize the relationship between lactylation and acetylation or other acylation modifications as synergistic or competitive because the changes in lactylation and acetylation occur differently in diverse cells and respond differently to various stimuli. Although no research has yet provided an explanation for the result, we may assume that they are related to the various subtypes of the pertinent enzymes. As is well-known, various tissues express different subtypes of HDACs differently (de Ruijter et al., 2003). It remains unclear if, even in the presence of the same stimulus, different distribution or activity of a specific subtype of HDACs in cells leads to asynchronous changes in lactylation and acetylation.

Other histone modifications, such as crotonylation and butyrylation, have also been discovered to be connected to lactylation. For instance, a recent discovery indicates that histone lactylation and crotonylation are crucial for epigenetic regulation of brain development. In this work, lactylation and crotonylation were shown to have a synergistic effect on the processes of neural differentiation and cell proliferation (Dai et al., 2022). Lactylation may also be related to butyrylation mediated by butyric acid because butyric acid contributes to an increased lactylation level of whole protein in Hela cells and may prevent lactylation *via* inhibiting a wide range of HDACs. Similar to this, we hypothesize that lactylation may be connected in some unknown ways to other histone modifications, including butyrylation, propionylation, succinylation, glutarylation, beta-hydroxybutyrylation, and 2-hydroxyisobutyrylation (Chen et al., 2007; Xie et al., 2012; Dai et al., 2014; Xie et al., 2016; Bao et al., 2019). But, we do not fully understand all of this right now; thus, it has to be investigated in further research.

Lactylation and diseases

HPTMs are involved in gene activation or gene silencing in cancer and inflammation (Allen et al., 2014; Benayoun et al., 2019; Evans et al., 2020). Belonging to HPTMs, lactylation is also intimately linked to numerous illnesses. An increasing amount of evidence points to the involvement of lactate-mediated histone lactylation in the pathophysiology of several systems. The function of histone lactylation in tumors, inflammation, embryogenesis, neuropathy, and pulmonary fibrosis will be discussed in the following sections.

Tumors

The microenvironment of tumors is crucial to development and growth of malignancies. Various factors, including metabolites such as lactate, can affect it. Lactate is released by tumor cells and detected by macrophages through transporters found on those cells (Colegio et al., 2014; Chen et al., 2017; Jiang et al., 2021b). In the past, lactate was thought to be a biological marker of malignancy. For instance, Martinez-Z et al. discovered that the accumulation of lactate outside the tumor was closely related to both a shorter overall patient survival rate and a greater incidence of metastasis in tumor patients (Martínez-Zaguilán et al., 1996). The German scientist Otto Heinrich Warburg discovered the Warburg effect in the 20th century, which describes how most tumor cells have a high glucose absorption rate. Tumor cells produced and secreted significant quantities of lactate as a result of aggressive glycolysis occurring in aerobic conditions (Warburg, 1928). The Warburg effect hypothesis brought increased attention to the role of glycolysis in cancer. When lactate from a Warburg-type metabolism was discovered, Arg1 expression surged. It has even been associated with protumor in several systems and is also a marker for M2 macrophages (Colegio et al., 2014; Carmona-Fontaine et al., 2017; Arlauckas et al., 2018). In conclusion, it is not difficult to hypothesize that lactate-induced histone lactylation should be a crucial factor in the growth of tumors. Exact histone lysine lactylation was identified in mouse bone marrow–derived macrophages and HeLa cells. Research studies also demonstrated that lactylation could be enhanced under hypoxia, and intracellular or extracellular lactate could affect it (Zhang et al., 2019). The carcinogenic significance of histone lactylation in ocular melanoma was then shown by Jie Yu et al. They demonstrated that YTH domain family protein 2 (YTHDF2), an m6A (N6-methyladenosine) reader protein, is activated by histone lactylation, and this finding provided novel histone lactylation targets for treating ocular melanoma (Yu et al., 2021). It is to be note that it was the first time lactylation of histones was shown to promote oncogene expression and quicken tumor development, indicating that lactylation was involved in tumor development. In tumor immune escape, the cell types known as tumor-infiltrating myeloid cells (TIMs) play a significant role. Recently, Jia et al. discovered that lactylation also controls TIMs. Their findings demonstrated that lactate accumulated in the tumor microenvironment effectively increased methyltransferase-like 3 (METTL3) in TIMs *via* H3K18la. They discovered it was critical for boosting TIMs' immunosuppressive abilities using lactylation-driven METTL3-mediated RNA m6A modification after identifying two lactylation modification sites in the zinc-finger domain of METTL3 (Xiong et al., 2022).

Inflammation

It was well-established that inflammation is essential for onset and development of many illnesses (Liu et al., 2017). Growing evidence has shown several molecular mechanisms,

TABLE 1 Role of histone lactylation in disease models.

Study type	Model species	Intervention and dose	Study site	Result	Mechanism	Ref.
	Ocular melanoma cells–induced nude mice		Pan Kla, H3K18la	Promote YTHDF2's transcription to accelerate tumorigenesis	YTHDF2↑; PER1↓; D TP53↓	[47]
	Bleomycin or TGF-β1-induced mice		Pan-Kla	Lactate induces Kla and pro-fibrotic gene expression mediated by p300	Arg↑; Opn↑; Pdgfa↑; Thbs↑; Vegfa↑	[68]
In						
	CLP polymicrobial sepsis mice	Lactate (0.5 g/kg body weight)	Pan-Kla	Perhaps improves the translocation of HMGB1 from the nucleus to the cytoplasm	Kla of HMGB1↑	[28]
vivo	SDS, ECS, and KCI-induced mice	Oxamate (1 g/kg)	Pan-Kla, H1la	Neuronal excitation increases Kla in the brain	Kla of H1	[66]
	ICR mice; Chinese Small Tail Han ewes		H3K18la	Lactate-induced Kla may contribute to remodeling endometrial receptivity	H3K18la may promote the ratio of GSH/GSSG	[63]
	AD mouse model		Pan-Kla; H4K18la; H4K5la; H4K8la; H3K18la; H3K23la	The H4K12la level is specifically elevated in plaque of 5XFAD mice	Form a positive feedback glycolysis/H4K 12la/ PKM2 loop	[67]
	Bacteria or LPS, IFNγ, and hypoxia-induced cells (MCF-7, MDA-MB-231, HeLa, A549, HepG2, MEF, and RAW 264.7 cells)	LA (25 mM); Glucose (0-25mM); 2-DG (0-10 mM; IL-4 (20ng/mL); Oxamate (0-20mM); Rotenone (0-50nM); DCA (10nM)	Pan-Kla, H3K18la, H4K5la	Histone Kla induces M2-like genes in M1 macrophages	Arg1↑; Mmp9↑;Rtn4↑; Tgm1↑; Spsb4↑; Hsd11b1↑; Nos2↑; IL-6↑; Tnf↑; Cxcl1↑; Ccl4↑; Ccr9↑	[7]
	Sox2, Klf4, and Oct4 (SKO)-induced MEFs	Glis1 or Flag	Pan Kla, H3K18la	Glis1 modulates lactylation during reprogramming	Hk2↑; Pgk1↑; Pfkl↑; Pkm↑; Eno1↑; Ldha↑	[25]
In	Human ocular melanoma cell lines	2-DG(0-10mN); Oxamate (0-20mM); siLDHA; siLDHB; Nala(0-25mM)	Pan Kla, H3K18la	Promote YTHDF2's transcription to accelerate tumorigenesis	YTHDF2↑; PER1↓; TP53↓	[47]
	TGF-β1 or bleomycin-induced lung fibroblast and BALFs	Lactate; si p300	Pan-Kla	Lactate induces Kla and pro-fibrotic gene expression mediated by p300	Arg↑; Opn↑; Pdgfa↑; Thbs1↑; Vegfa↑	[68]
	Bacteria- or LPS-induced $BCP^{-/-}$ BMDMs	Nala (25 mM)	Pan-Kla	BCAP deficiency reduces lactate and lactylation	FOXO1↑; GSK3β↑	[61]
	Hypoxia-induced human lung bronchial epithelial cell and NSCLC cell;	Lactate (5 or 10 mM)	Pan-Kla, H4K8la	Kla medicated by lactate regulates metabolism-related gene	SDH↑; IDH↑; HIF1A↑; HK-1↓; G6PD↓; PKM↓	[42]
	B. cinerea	Lactate (0-10mM)	Pan-Kla	Lactylated proteins participate in fungal pathogenicity	Kla of EIF-5A↑	[9]
	LPS and hypoxia-induced macrophages	Oxamate (20 mM); Lactate (10mM)	Pan-Kla	Perhaps improves the translocation of HMGB1 from the nucleus to the cytoplasm	Kla of HMGB1↑	[28]
vitro	Primary embryonic mice neuron	Lactate(0-25m M), oxamate; 4-CIN	Pan-Kla; H1la	Neuronal excitation increases Kla in the brain	Kla of H1↑	[66]

Abbreviations: ↑, Upregulated; ↓, Downregulated; $^{+}$, and; Kla, lysine lactylation; YTHDF2, YTH N6-methyladenosine RNA-binding protein 2; PER1, period1 gene; TP53, the gene encoding the p53 protein; TGF-β1, transforming growth factor-β1; BALFs, bronchoalveolar lavage fluids; p300, a acetyltransferase; Arg1, arginase1; Opn, osteopontin; Pdgfa, platelet-derived growth factor A; Thbs1, thrombospondin-1; Vegfa, vascular endothelial growth factor A; CLP, cecal ligation and puncture; HMGB1, high mobility box-1; SDS: social defeat; ECS, electroconvulsive stimulation; KCI, potassium chloride; GSH, reduced glutathione; GSSG, oxidized glutathione; AD, Alzheimer's disease; 5XFAD, a genetically modified mouse type with Alzheimer's disease; Pkm2, pyruvate kinase M2; BMDMs, mouse bone marrow derived macrophages; IFNγ, interferon-γ; LPS, lipopolysaccharide; La, lactic acid; 2-DG,2-deoxy-D-glucose; IL-4, interleukin 4; DCA, Sodium dichloroacetate; Mmp9, matrixmetalloproteinase9; Rtn4r, reticulon 4 receptor; Tgm1, transglutaminase 1; Spsb4, SPRY domain- and SOCS box-containing protein 2; Hsd11b1, 11β-hydroxysteroid dehydrogenase type 1; Nos2, nitric oxide synthase 2; IL-6, interleukin 6; Tnf, tumor necrosis factor; Cxcl1, CXC chemokine ligand 1; CcI4, C-C motif chemokine ligand 4; Ccr9, CC chemokine receptor 9; Sox2, sex determining region Y-box 2; Klf4, Krüppel-like factor 4; Oct4, octamer-binding transcription factor 4; MEFs, mouse embryonic fibroblasts; Glis1, Gli-like transcription factor 1; Hk2, hexokinase 2; Pgk1, phosphoglycerate kinase1; Pfk1, phosphofructokinase-1; Eno1, enolase1; Ldha, lactate dehydrogenase; siLDHA, siRNAs for LDHA; siLDHB, siRNAs for LDHB; sip300, siRNAs for p300; Nala, sodium lactate; BCAP, B-cell adapter for PI3K; FOXO1, forkhead box protein O1; GSK3β, glycogen synthase kinase 3β; SDH, succinate dehydrogenase; PKM, pyruvate kinase; NSCLC, non-small cell lung cancer; eIF-5A, translation initiation factor 5A; 4-CIN, α-cyano-4-hydroxycinnamate; GSKA, lactate dehydrogenase activity inhibitor.

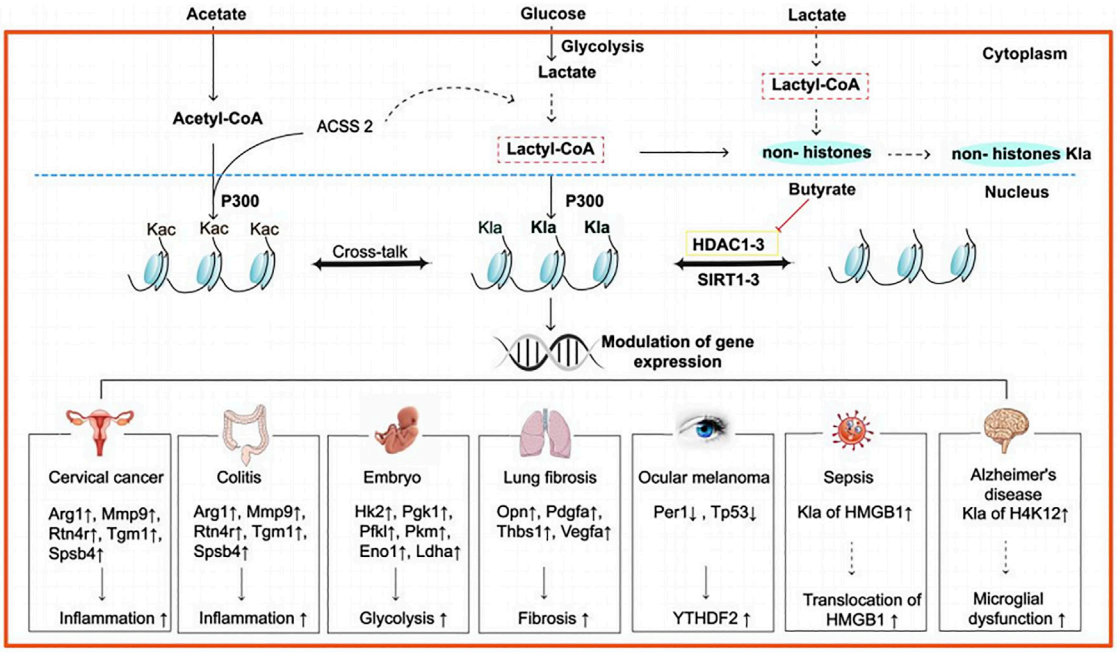


FIGURE 1
Possible mechanism and roles of histone lactylation in some diseases. Lactate derives from the conversion of glucose by glycolysis, and it may synthesize lactyl-CoA, which is hitherto unclear with the ACSS2. Then, the lactyl group is transferred by p300 to Lys lactylations, leading to various physiopathological activities in different diseases. For example, in human cervical cancer and colitis, inflammation-related genes such as Arg1, Mmp9, Rtn4r, Tgm1, and Spsb4 are upregulated; in embryon, some glycolysis-related genes, including Hk2, Pgk1, Pfk1, Pkm, Eno1, and Ldha, are upregulated; in lung fibrosis, fibrosis-related genes such as Opn, Pdgfa, Thbs1, and Vegfa are upregulated; in ocular melanoma, two tumor suppressor genes, namely, Per1 and Tp53 are upregulated; In sepsis, Kla of HMGB1 is upregulated, contributing to the translocation of HMGB1; in AD, Kla of H4K12 is upregulated, resulting in microglial dysfunction. Abbreviations: ACSS2, synthetase short-chain family member 2; p300, an acetyltransferase; Arg1, arginase1; Mmp9, matrixmetalloproteinase9; Rtn4r, reticulon 4 receptor; Tgm1, transglutaminase 1; Spsb4, SPRY domain- and SOCS box-containing protein 3; Hk2, hexokinase 2; Pgk1, phosphoglycerate kinase1; Pfk1, phosphofructokinase-2; Pkm, pyruvate kinase M2; Eno1, enolase1; Ldha, lactate dehydrogenase; Opn, osteopontin; Pdgfa, platelet-derived growth factor A; Thbs 1, thrombospondin-1; Vegfa, vascular endothelial growth factor A; Per1, period circadian regulator 1; Tp53, tumor suppressor gene TP53; the dashed lines mean that the process has not been proved; The solid lines mean that the process has been proved.

and now histone lactylation is also involved in the activation of inflammation (Chen et al., 2019a; Feng et al., 2019; Wenzel et al., 2019; Certo et al., 2021). As a result, histone lactylation may create novel therapeutic strategies for prevention and treatment of various illnesses focused on inflammation (Chen et al., 2019b; Wang et al., 2020a; Wang et al., 2020b; Sousa et al., 2020). Under various disease conditions, such as sepsis, cancer, chronic inflammation, and autoimmune diseases, it has been reported that lactate produced by aerobic glycolysis has immunosuppressive effects in local tissues (Sangsuwan et al., 2020). It has also been suggested that glycolysis at high rates may provide a way to meet the increased demand for biosynthetic precursors, which is used for pro-inflammatory protein synthesis after pro-inflammatory immune cell activation (Pucino et al., 2017). Macrophages can be classified as M1 or M2 functional phenotypes, which act in pro-inflammatory or anti-inflammatory capacities, respectively. As was previously reported, Zhang et al. demonstrated that lactate and histone lactylation significantly contributed to transformation of pro-inflammatory M1 macrophages into anti-inflammatory M2 macrophages (Zhang et al., 2019), and this was further supported in adipose tissue. According to other research, lactylation is a consequence rather than a cause of macrophage activation, but it occurs coincidently with an IL-6-and Arg1-dependent metabolic rewiring under inflammatory duress (Palsson-McDermott and O'Neill, 2013). Clinical trials have shown similar findings. In a study of 13 healthy volunteers and 35 critically ill patients with septic shock, Chu, X et al. found that H3K18la may reflect the severity of critical illness and presence of infection, suggesting that H3K18la may have a significant impact on the balance of inflammation in sepsis (Chu et al., 2021). In addition, Irizarry-C et al. showed that when lacking B-cell adapter for PI3K (BCAP), macrophages show impaired lactate production and also exhibit reduced histone lactylation and lower tissue repair gene expression, which results in blunting their reparative transition (Irizarry-Caro et al., 2020). In short, the fact that histone lactylation plays a role in the development of inflammation is affirmative, suggesting a novel avenue for treatment of inflammation.

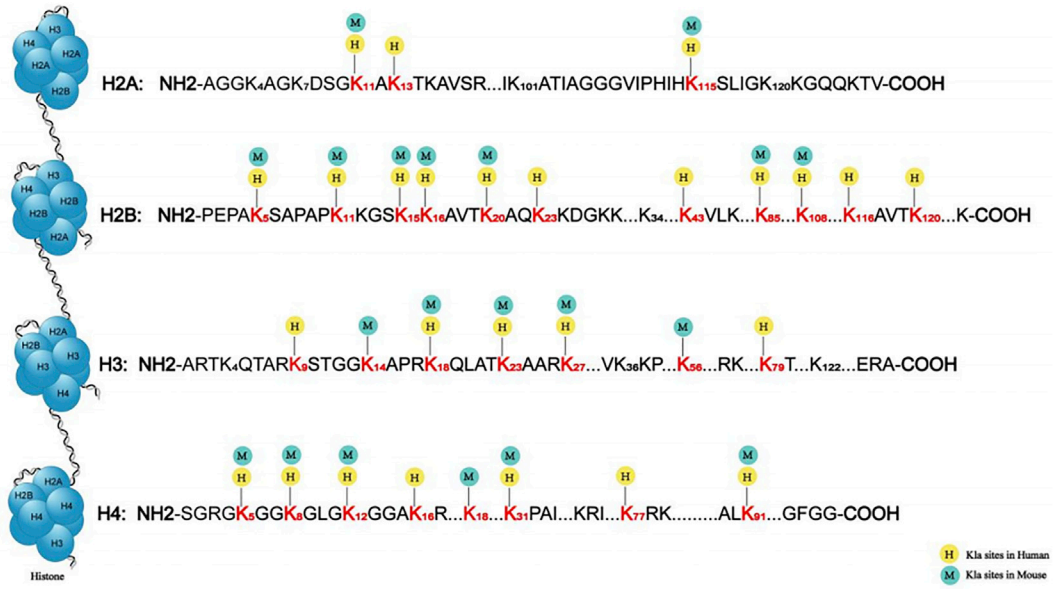


FIGURE 2
Sites of histone lactylation in humans and mice. The lactylation sites identified on the four core histones, namely, H3, H4, H2A, and H2B. The different numbers represent the various sequences of lysine in histones. Some sites, such as H3K18, locate in both humans and mice, shown by both yellow circle icon with "H" and green circle icon with "M". Others locate in only humans or mice, shown by yellow circle icon with "H" and green circle icon with "M", respectively.

Embryogenesis

Recently, it has also been claimed that lactylation affects embryogenesis. Yang et al. discovered that H3K18la, which is a reflection of histone lactylation in the endometrium, participated in remodeling uterine receptivity (Yang et al., 2022b). This finding provided a novel insight for developing potential clinical intervention strategies to increase pregnancy rates following both natural and assisted conception. The effects of hypoxia on the development of mouse pre-implantation embryos were later shown *in vitro*. In hypoxic conditions (2 percent O_2), as compared to atmospheric oxygen content, the lactylation in embryos was significantly less (Yang et al., 2021). Additionally, hypoxia reduces the activity of LDHA, which reduces lactate production. Moreover, another study showed that the accumulation of H3K18la on germline and cleavage embryo genes promotes transcriptional elongation (Tian and Zhou, 2022). The authors found that adding lactate activated the germline genes and cleavage embryo genes in mouse embryonic stem cells (ESCs). So, they proposed that to enhance transcriptional elongation, cofactors are recruited through lactylation, which is stimulated by lactate.

Neuropathy

Numerous studies have shown that astrocyte-derived lactate may be utilized to fuel neurons and can also serve as a signaling molecule by activating lactate receptors on the surface of neuronal cell membranes and changing how those receptors operate. In fact, lactylation, as with lactate, is also closely linked to neural activity. Hagihara et al. discovered that the degree of lactylation and lactate was correlated and that lactylation was controlled by neural excitation and social stress in brain cells (Hagihara et al., 2021). Social defeat stress and cerebral stimulation both increase brain lactate and lactylation levels, with the latter being a chronic process. After identifying 63 candidate lysine-lactylated proteins, additional research showed that lower social behavior is associated with higher histone H1la (Hagihara et al., 2021). A recent study found that lactylation had a role in the etiology of Alzheimer's disease (AD). They discovered that the level of H4K12la was upregulated in Aβ plaque-adjacent microglia; thus, they hypothesized that this increased glycolysis/H4K12la/PKM2 positive feedback loop, which exacerbated microglial dysfunction in AD (Pan et al., 2022). So, this may be a possible therapy for AD.

Pulmonary fibrosis

Fibrosis is a common pathogenic characteristic of many illnesses, particularly interstitial lung disease, including pulmonary fibrosis. Pulmonary fibers may be stimulated by lactate that myofibroblasts and macrophages produce in the extracellular environment. To fulfill its non-metabolic activities, Cui et al. proposed that lactate generated by

myofibroblasts may modify the pathogenic phenotype of alveolar macrophages. Then, they demonstrated that there was markedly elevated lactate in the conditioned media of transforming growth factor -β1 (TGF-1)–induced lung myofibroblasts and in the bronchoalveolar lavage fluids (BALFs) from animals with TGF-1-induced or bleomycin-induced lung fibrosis (Cui et al., 2021). Significantly, both the media and BALFs encouraged macrophages to produce pro-fibrotic mediators. Mechanistically, histone lactylation at the promoters of the pro-fibrotic genes in macrophages is boosted, which is consistent with the increase of lactate change in these cells. Histone lactylation and expression of pro-fibrotic genes were mediated by p300, as shown by the reduced levels of these processes in p300-knockdown macrophages (Cui et al., 2021). These discoveries provide fresh insight into the pathophysiology of the crucial role played by myofibroblast glycolysis in the etiology of lung fibrosis.

Summary and future perspectives

Most somatic cells contain lactate, and histone lactylation alterations may also be common *in vivo*. The impact of lactylation on gene expression is well-established. The mechanism of histone lactylation is still unclear. In our hypothesis, when lactate reaches a specific level under healthy or pathological circumstances, it causes lactylation of histones or non-histone to engage in various biochemical processes in the body, influencing the biological reactions of the organism. How it begins, whether substrate enzymes are present before the transfer, and how the body detects when lactylation should end are all yet unknown at present. We, thus, need to conduct additional investigation. As we already noted, lactylation has been shown to occur in some diseases such as tumors, inflammation, and so on. Whether there is lactinylation in other diseases in which lactate increases within a period of time needs further research. For example, when in ischemic and hypoxic disorders such as cardiac ischemia, cells may produce more lactate, resulting in lactylation. Also, in disorders such as rheumatoid arthritis, atherosclerotic plaques, poisoning, and intervertebral disc illnesses, it will happen, too (Cheng et al., 2022). In addition, lactate level also increases in certain physiological states such as exercises, and whether the lactylation levels change during this process (Plaza-Diaz et al., 2022) or whether the function is the same and lactylation level under physiological conditions and disease conditions is different are all worth exploring. Histone lactylation represents a novel face of histone posttranslational modifications, and more research studies are needed to unravel its mysteries. As previously mentioned, D-lactylation and L-lactylation have differing catalytic efficiencies when removed by HDACs (Gaffney et al., 2020; Moreno-Yruela et al., 2022b). This requires further research. Currently, non-histone modifications are poorly studied and deserve more attention. Last, we hope that this article will help readers and provide them with new research ideas.

Author contributions

LC and WH were responsible for the conception and design of the study. Y-MX, H-XH, M-TL, and T-TZ were responsible for data collection, analysis, and image processing. Y-MX wrote the manuscript, and XC revised the manuscript. Y-MX and H-XH were responsible for the final approval of the version to be submitted. All authors read and approved the final manuscript.

Funding

The authors are thankful for the support of the National Natural Science Foundation of China (No. 82170834, No. 81970676 and No. 81800741), the Office of Science Technology and Talent Work of Luzhou (No. 2020LZXNYDP02 and No.2021LZXNYD-G01) and the project of Affiliated Hospital of Southwest Medical University (No.11160).

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.

References

Allen, B. G., Bhatia, S. K., Anderson, C. M., Eichenberger-Gilmore, J. M., Sibenaller, Z. A., Mapuskar, K. A., et al. (2014). Ketogenic diets as an adjuvant cancer therapy: History and potential mechanism. *Redox Biol.* 2, 963–970. doi:10.1016/j.redox.2014.08.002

Arlauckas, S. P., Garren, S. B., Garris, C. S., Kohler, R. H., Oh, J., Pittet, M. J., et al. (2018). Arg1 expression defines immunosuppressive subsets of tumor-associated macrophages. *Theranostics* 8 (21), 5842–5854. doi:10.7150/thno.26888

Bao, X., Liu, Z., Zhang, W., Gladysz, K., Fung, Y. M. E., Tian, G., et al. (2019). Glutarylation of histone H4 lysine 91 regulates chromatin dynamics. *Mol. Cell* 76 (4), 660–675. e9. doi:10.1016/j.molcel.2019.08.018

Benayoun, B. A., Pollina, E. A., Singh, P. P., Mahmoudi, S., Harel, I., Casey, K. M., et al. (2019). Remodeling of epigenome and transcriptome landscapes with aging in mice reveals widespread induction of inflammatory responses. *Genome Res.* 29 (4), 697–709. doi:10.1101/gr.240093.118

Bowman, G. D., and Poirier, M. G. (2015). Post-translational modifications of histones that influence nucleosome dynamics. *Chem. Rev.* 115 (6), 2274–2295. doi:10.1021/cr500350x

Brooks, G. A. (2020). Lactate as a fulcrum of metabolism. *Redox Biol.* 35, 101454. doi:10.1016/j.redox.2020.101454

Carmona-Fontaine, C., Deforet, M., Akkari, L., Thompson, C. B., Joyce, J. A., and Xavier, J. B. (2017). Metabolic origins of spatial organization in the tumor microenvironment. *Proc. Natl. Acad. Sci. U. S. A.* 114 (11), 2934–2939. doi:10.1073/pnas.1700600114

Certo, M., Elkafrawy, H., Pucino, V., Cucchi, D., Cheung, K. C. P., and Mauro, C. (2021). Endothelial cell and T-cell crosstalk: Targeting metabolism as a therapeutic approach in chronic inflammation. *Br. J. Pharmacol.* 178 (10), 2041–2059. doi:10.1111/bph.15002

Chen, D. Q., Feng, Y. L., Chen, L., Liu, J. R., Wang, M., Vaziri, N. D., et al. (2019). Poricoic acid A enhances melatonin inhibition of AKI-to-CKD transition by regulating Gas6/AxlNFκB/Nrf2 axis. *Free Radic. Biol. Med.* 134, 484–497. doi:10.1016/j.freeradbiomed.2019.01.046

Chen, L., Chen, D. Q., Liu, J. R., Zhang, J., Vaziri, N. D., Zhuang, S., et al. (2019). Unilateral ureteral obstruction causes gut microbial dysbiosis and metabolome disorders contributing to tubulointerstitial fibrosis. *Exp. Mol. Med.* 51 (3), 1–18. doi:10.1038/s12276-019-0234-2

Chen, P., Zuo, H., Xiong, H., Kolar, M. J., Chu, Q., Saghatelian, A., et al. (2017). Gpr132 sensing of lactate mediates tumor-macrophage interplay to promote breast cancer metastasis. *Proc. Natl. Acad. Sci. U. S. A.* 114 (3), 580–585. doi:10.1073/pnas.1614035114

Chen, Y., Sprung, R., Tang, Y., Ball, H., Sangras, B., Kim, S. C., et al. (2007). Lysine propionylation and butyrylation are novel post-translational modifications in histones. *Mol. Cell. Proteomics* 6 (5), 812–819. doi:10.1074/mcp.M700021-MCP200

Cheng, C., Xu, Z. Q., Yang, C., and Wu, X. H. (2022). *Lactylation driven by lactate metabolism in the disc accelerates intervertebral disc degeneration: A hypothesis*. Midlothian, Scotland: Medical Hypotheses, 159.

Chu, X., Di, C., Chang, P., Li, L., Feng, Z., Xiao, S., et al. (2021). Lactylated histone H3K18 as a potential biomarker for the diagnosis and predicting the severity of septic shock. *Front. Immunol.* 12, 786666. doi:10.3389/fimmu.2021.786666

Colegio, O. R., Chu, N. Q., Szabo, A. L., Chu, T., Rhebergen, A. M., Jairam, V., et al. (2014). Functional polarization of tumour-associated macrophages by tumour-derived lactic acid. *Nature* 513 (7519), 559–563. doi:10.1038/nature13490

Cui, H. C., Xie, N., Banerjee, S., Ge, J., Jiang, D., Dey, T., et al. (2021). Lung myofibroblasts promote macrophage profibrotic activity through lactate-induced histone lactylation. *Am. J. Respir. Cell Mol. Biol.* 64 (1), 115–125. doi:10.1165/rcmb.2020-0360OC

Dai, L. Z., Peng, C., Montellier, E., Lu, Z., Chen, Y., Ishii, H., et al. (2014). Lysine 2-hydroxyisobutyrylation is a widely distributed active histone mark. *Nat. Chem. Biol.* 10 (5), 365–370. doi:10.1038/nchembio.1497

Dai, S.-K., Liu, P.-P., Li, X., Jiao, L.-F., Teng, Z.-Q., and Liu, C.-M. (2022). *Dynamic profiling and functional interpretation of histone Kcr and Kla during neural development*. Cambridge, England: Development.

de Ruijter, A. J., van Gennip, A. H., Caron, H. N., Kemp, S., and van Kuilenburg, A. B. P. (2003). Histone deacetylases (HDACs): Characterization of the classical HDAC family. *Biochem. J.* 370 (3), 737–749. doi:10.1042/BJ20021321

Delcuve, G. P., Khan, D. H., and Davie, J. R. (2012). Roles of histone deacetylases in epigenetic regulation: Emerging paradigms from studies with inhibitors. *Clin. Epigenetics* 4, 5. doi:10.1186/1868-7083-4-5

Evans, L. W., Stratton, M. S., and Ferguson, B. S. (2020). Dietary natural products as epigenetic modifiers in aging-associated inflammation and disease. *Nat. Prod. Rep.* 37 (5), 653–676. doi:10.1039/c9np00057g

Feng, Y. L., Cao, G., Chen, D. Q., Vaziri, N. D., Chen, L., Zhang, J., et al. (2019). Microbiome-metabolomics reveals gut microbiota associated with glycine-conjugated metabolites and polyamine metabolism in chronic kidney disease. *Cell. Mol. Life Sci.* 76 (24), 4961–4978. doi:10.1007/s00018-019-03155-9

Ferguson, B. S., Rogatzki, M. J., Goodwin, M. L., Kane, D. A., Rightmire, Z., and Gladden, L. B. (2018). Lactate metabolism: Historical context, prior misinterpretations, and current understanding. *Eur. J. Appl. Physiol.* 118 (4), 691–728. doi:10.1007/s00421-017-3795-6

Figlia, G., Willnow, P., and Teleman, A. A. (2020). Metabolites regulate cell signaling and growth via covalent modification of proteins. *Dev. Cell* 54 (2), 156–170. doi:10.1016/j.devcel.2020.06.036

Fletcher, W. M. (1907). Lactic acid in amphibian muscle. *J. Physiol.* 35 (4), 247–309. doi:10.1113/jphysiol.1907.sp001194

Gaffney, D. O., Jennings, E. Q., Anderson, C. C., Marentette, J. O., Shi, T., Schou Oxvig, A. M., et al. (2020). Non-enzymatic lysine lactoylation of glycolytic enzymes. *Cell Chem. Biol.* 27 (2), 206–213. doi:10.1016/j.chembiol.2019.11.005

Gao, M., Zhang, N., and Liang, W. (2020). Systematic analysis of lysine lactylation in the plant fungal pathogen botrytis cinerea. *Front. Microbiol.* 11, 594743. doi:10.3389/fmicb.2020.594743

Hagihara, H., Shoji, H., Otabi, H., Toyoda, A., Katoh, K., Namihira, M., et al. (2021). Protein lactylation induced by neural excitation. *Cell Rep.* 37 (2), 109820. doi:10.1016/j.celrep.2021.109820

Irizarry-Caro, R. A., McDaniel, M. M., Overcast, G. R., Jain, V. G., Troutman, T. D., and Pasare, C. (2020). TLR signaling adapter BCAP regulates inflammatory to reparatory macrophage transition by promoting histone lactylation. *Proc. Natl. Acad. Sci. U. S. A.* 117 (48), 30628–30638. doi:10.1073/pnas.2009778117

Jennings, E. Q., Ray, J. D., Zerio, C. J., Trujillo, M. N., McDonald, D. M., Chapman, E., et al. (2021). Sirtuin 2 regulates protein LactoylLys modifications. *Chembiochem.* 22 (12), 2102–2106. doi:10.1002/cbic.202000883

Jiang, J., Huang, D., Jiang, Y., Hou, J., Tian, M., Li, J., et al. (2021). Lactate modulates cellular metabolism through histone lactylation-mediated gene expression in non-small cell lung cancer. *Front. Oncol.* 11, 647559. doi:10.3389/fonc.2021.647559

Jiang, P. R., Ning, W., Shi, Y., Liu, C., Mo, S., Zhou, H., et al. (2021). FSL-kla: A few-shot learning-based multi-feature hybrid system for lactylation site prediction. *Comput. Struct. Biotechnol. J.* 19, 4497–4509. doi:10.1016/j.csbj.2021.08.013

Latham, T., Mackay, L., Sproul, D., Karim, M., Culley, J., Harrison, D. J., et al. (2012). Lactate, a product of glycolytic metabolism, inhibits histone deacetylase activity and promotes changes in gene expression. *Nucleic Acids Res.* 40 (11), 4794–4803. doi:10.1093/nar/gks066

Levitt, M. D., and Levitt, D. G. (2020). Quantitative evaluation of D-lactate pathophysiology: New insights into the mechanisms involved and the many areas in need of further investigation. *Clin. Exp. Gastroenterol.* 13, 321–337. doi:10.2147/CEG.S260600

Li, L., Chen, K., Wang, T., Wu, Y., Xing, G., Chen, M., et al. (2020). Glis1 facilitates induction of pluripotency via an epigenome-metabolome-epigenome signalling cascade. *Nat. Metab.* 2 (9), 882–892. doi:10.1038/s42255-020-0267-9

Liu, Y. Z., Wang, Y. X., and Jiang, C. L. (2017). Inflammation: The common pathway of stress-related diseases. *Front. Hum. Neurosci.* 11, 316. doi:10.3389/fnhum.2017.00316

Lu, J., Fu, S., Dai, J., Hu, J., Li, S., Ji, H., et al. (2022). Integrated metabolism and epigenetic modifications in the macrophages of mice in responses to cold stress. *J. Zhejiang Univ. Sci. B* 23 (6), 461–480. doi:10.1631/jzus.B2101091

Martínez-Zaguilán, R., Seftor, E. A., Seftor, R. E., Chu, Y. W., Gillies, R. J., and Hendrix, M. J. (1996). Acidic pH enhances the invasive behavior of human melanoma cells. *Clin. Exp. Metastasis* 14 (2), 176–186. doi:10.1007/BF00121214

Meng, X., Baine, J. M., Yan, T., and Wang, S. (2021). Comprehensive analysis of lysine lactylation in rice (oryza sativa) grains. *J. Agric. Food Chem.* 69 (29), 8287–8297. doi:10.1021/acs.jafc.1c00760

Moreno-Yruela, C., Bæk, M., Monda, F., and Olsen, C. A. (2022). Chiral posttranslational modification to lysine ε-amino groups. *Acc. Chem. Res.* 55 (10), 1456–1466. doi:10.1021/acs.accounts.2c00115

Moreno-Yruela, C., Zhang, D., Wei, W., Bæk, M., Liu, W., Gao, J., et al. (2022). Class I histone deacetylases (HDAC1-3) are histone lysine delactylases. *Sci. Adv.* 8 (3), eabi6696. doi:10.1126/sciadv.abi6696

Moussaieff, A., Rouleau, M., Kitsberg, D., Cohen, M., Levy, G., Barasch, D., et al. (2015). Glycolysis-mediated changes in acetyl-CoA and histone acetylation control the early differentiation of embryonic stem cells. *Cell Metab.* 21 (3), 392–402. doi:10.1016/j.cmet.2015.02.002

Oh, M. S., Phelps, K. R., TraubeM.Barbosa-Saldivar, J. L., Boxhill, C., and Carroll, H. J. (1979). D-lactic acidosis in a man with the short-bowel syndrome. *N. Engl. J. Med.* 301 (5), 249–252. doi:10.1056/NEJM197908023010505

Palsson-McDermott, E. M., and O'Neill, L. A. (2013). The Warburg effect then and now: From cancer to inflammatory diseases. *Bioessays* 35 (11), 965–973. doi:10.1002/bies.201300084

Pan, R. Y., He, L., Zhang, J., Liu, X., Liao, Y., Gao, J., et al. (2022). Positive feedback regulation of microglial glucose metabolism by histone H4 lysine 12 lactylation in Alzheimer's disease. *Cell Metab.* 34 (4), 634–648 e6. doi:10.1016/j.cmet.2022.02.013

Plaza-Diaz, J., Izquierdo, D., Torres-Martos, A., Baig, A. T., Aguilera, C. M., and Ruiz-Ojeda, F. J. (2022). Impact of physical activity and exercise on the epigenome in skeletal muscle and effects on systemic metabolism. *Biomedicines* 10 (1), 126. doi:10.3390/biomedicines10010126

Pucino, V., Bombardieri, M., Pitzalis, C., and Mauro, C. (2017). Lactate at the crossroads of metabolism, inflammation, and autoimmunity. *Eur. J. Immunol.* 47 (1), 14–21. doi:10.1002/eji.201646477

Rogatzki, M. J., Ferguson, B. S., Goodwin, M. L., and Gladden, L. B. (2015). Lactate is always the end product of glycolysis. *Front. Neurosci.* 9, 22. doi:10.3389/fnins.2015.00022

Sangsuwan, R., Thuamsang, B., Pacifici, N., Allen, R., Han, H., Miakicheva, S., et al. (2020). Lactate exposure promotes immunosuppressive phenotypes in innate immune cells. *Cell. Mol. Bioeng.* 13 (5), 541–557. doi:10.1007/s12195-020-00652-x

Sousa, L. P., Pinho, V., and Teixeira, M. M. (2020). Harnessing inflammation resolving-based therapeutic agents to treat pulmonary viral infections: What can the future offer to COVID-19? *Br. J. Pharmacol.* 177 (17), 3898–3904. doi:10.1111/bph.15164

Strahl, B. D., and Allis, C. D. (2000). The language of covalent histone modifications. *Nature* 403 (6765), 41–45. doi:10.1038/47412

Sun, Y., Chen, Y., and Peng, T. (2022). A bioorthogonal chemical reporter for the detection and identification of protein lactylation. *Chem. Sci.* 13 (20), 6019–6027. doi:10.1039/d2sc00918h

Tessarz, P., and Kouzarides, T. (2014). Histone core modifications regulating nucleosome structure and dynamics. *Nat. Rev. Mol. Cell Biol.* 15 (11), 703–708. doi:10.1038/nrm3890

Tian, Q., and Zhou, L. Q. (2022). Lactate activates germline and cleavage embryo genes in mouse embryonic stem cells. *Cells* 11 (3), 548. doi:10.3390/cells11030548

Wang, M., Hu, H. H., Chen, Y. Y., Chen, L., Wu, X. Q., and Zhao, Y. Y. (2020). Novel poricoic acids attenuate renal fibrosis through regulating redox signalling and aryl hydrocarbon receptor activation. *Phytomedicine* 79, 153323. doi:10.1016/j.phymed.2020.153323

Wang, X., Antony, V., Wang, Y., Wu, G., and Liang, G. (2020). Pattern recognition receptor-mediated inflammation in diabetic vascular complications. *Med. Res. Rev.* 40 (6), 2466–2484. doi:10.1002/med.21711

Warburg, O. (1928). The chemical constitution of respiration ferment. *Science* 68 (1767), 437–443. doi:10.1126/science.68.1767.437

Wellen, K. E., Hatzivassiliou, G., Sachdeva, U. M., Bui, T. V., Cross, J. R., and Thompson, C. B. (2009). ATP-citrate lyase links cellular metabolism to histone acetylation. *Science* 324 (5930), 1076–1080. doi:10.1126/science.1164097

Wenzel, U. O., Bode, M., Kurts, C., and Ehmke, H. (2019). Salt, inflammation, IL-17 and hypertension. *Br. J. Pharmacol.* 176 (12), 1853–1863. doi:10.1111/bph.14359

Xie, Z. Y., Dai, J., Dai, L., Tan, M., Cheng, Z., Wu, Y., et al. (2012). Lysine succinylation and lysine malonylation in histones. *Mol. Cell. Proteomics.* 11 (5), 100–107. doi:10.1074/mcp.M111.015875

Xie, Z. Y., Zhang, D., Chung, D., Tang, Z., Huang, H., Dai, L., et al. (2016). Metabolic regulation of gene expression by histone lysine beta-hydroxybutyrylation. *Mol. Cell* 62 (2), 194–206. doi:10.1016/j.molcel.2016.03.036

Xiong, J., He, J., Zhu, J., Pan, J., Liao, W., Ye, H., et al. (2022). Lactylation-driven METTL3-mediated RNA m(6)A modification promotes immunosuppression of tumor-infiltrating myeloid cells. *Mol. Cell* 82 (9), 1660–1677 e10. doi:10.1016/j.molcel.2022.02.033

Yang, K., Fan, M., Wang, X., Xu, J., Wang, Y., Tu, F., et al. (2022). Lactate promotes macrophage HMGB1 lactylation, acetylation, and exosomal release in polymicrobial sepsis. *Cell Death Differ.* 29 (1), 133–146. doi:10.1038/s41418-021-00841-9

Yang, Q., Liu, J., Wang, Y., Zhao, W., Wang, W., Cui, J., et al. (2022). A proteomic atlas of ligand-receptor interactions at the ovine maternal-fetal interface reveals the role of histone lactylation in uterine remodeling. *J. Biol. Chem.* 298 (1), 101456. doi:10.1016/j.jbc.2021.101456

Yang, W., Wang, P., Cao, P., Wang, S., Yang, Y., Su, H., et al. (2021). Hypoxic *in vitro* culture reduces histone lactylation and impairs pre-implantation embryonic development in mice. *Epigenetics Chromatin* 14 (1), 57. doi:10.1186/s13072-021-00431-6

Yu, J., Chai, P., Xie, M., Ge, S., Ruan, J., Fan, X., et al. (2021). Histone lactylation drives oncogenesis by facilitating m(6)A reader protein YTHDF2 expression in ocular melanoma. *Genome Biol.* 22 (1), 85. doi:10.1186/s13059-021-02308-z

Zhang, D., Tang, Z., Huang, H., Zhou, G., Cui, C., Weng, Y., et al. (2019). Metabolic regulation of gene expression by histone lactylation. *Nature* 574 (7779), 575–580. doi:10.1038/s41586-019-1678-1

Zhang, N., Jiang, N., Yu, L., Guan, T., Sang, X., Feng, Y., et al. (2021). Protein lactylation critically regulates energy metabolism in the Protozoan parasite trypanosoma brucei. *Front. Cell Dev. Biol.* 9, 719720. doi:10.3389/fcell.2021.719720

Zhang, Y., Sun, Z., Jia, J., Du, T., Zhang, N., Tang, Y., et al. (2021). Overview of histone modification. *Adv. Exp. Med. Biol.* 1283, 1–16. doi:10.1007/978-981-15-8104-5_1

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 20 September 2022
DOI 10.3389/fgene.2022.949850

OPEN ACCESS

EDITED BY
Ann-Kristin Östlund Farrants,
Stockholm University, Sweden

REVIEWED BY
Peipei Ma,
Shanghai Jiao Tong University, China
M. Nazmul Hoque,
Bangabandhu Sheikh Mujibur Rahman Agricultural University, Bangladesh

*CORRESPONDENCE
Ying Yu,
yuying@cau.edu.cn

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

RECEIVED 21 May 2022
ACCEPTED 26 August 2022
PUBLISHED 20 September 2022

CITATION
Yang J, Tang Y, Liu X, Zhang J, Zahoor Khan M, Mi S, Wang C and Yu Y (2022), Characterization of peripheral white blood cells transcriptome to unravel the regulatory signatures of bovine subclinical mastitis resistance. *Front. Genet.* 13:949850. doi: 10.3389/fgene.2022.949850

Characterization of peripheral white blood cells transcriptome to unravel the regulatory signatures of bovine subclinical mastitis resistance

Jinyan Yang[1], Yongjie Tang[1], Xueqin Liu[1], Jinning Zhang[1], Muhammad Zahoor Khan[2], Siyuan Mi[1], Chuduan Wang[1] and Ying Yu[1]*

[1]Laboratory of Animal Genetics and Breeding, Ministry of Agriculture and Rural Affairs of China, National Engineering Laboratory of Animal Breeding, College of Animal Science and Technology, China Agricultural University, Beijing, China, [2]Department of Animal Sciences, Faculty of Veterinary and Animal Sciences, University of Agriculture, Dera Ismail Khan, Pakistan

Subclinical bovine mastitis is a pathogenic infection of the breast characterized by a marked decrease in milk production and quality. As it has no obvious clinical symptoms, diagnosis and treatment are challenging. Therefore, searching for biomarkers in cows' peripheral white blood cells is valuable for preventing and treating subclinical mastitis. Thus, in this study, the transcriptome of peripheral blood from healthy and subclinical mastitis cows was characterized to find the regulatory signatures of bovine subclinical mastitis using RNA-seq. A total of 287 differentially expressed genes (DEGs) and 70 differentially expressed lncRNAs (DELs) were detected, and 37 DELs were documented near known Quantitative Trait Loci (QTL) associated with the mastitis of cows. Bioinformatic analysis indicated that lncRNAs MSTRG25101.2, MSTRG.56327.1, and MSTRG.18968.1, which are adjacent to the SCS QTL and SCC QTL, may be candidate lncRNAs that influence the pathogenesis of mastitis in cows by up-regulating the expression of genes *TLR4*, *NOD2*, *CXCL8*, and *OAS2*. Moreover, the alternative splicing (AS) pattern of transcriptional sequence differences between healthy cows and subclinical mastitis cows suggested a molecular mechanism of mastitis resistance and susceptibility. A total of 2,212 differential alternative splicing (DAS) events, corresponding to 1,621 unique DAS genes, were identified in both groups and significantly enriched in immune and inflammatory pathways. Of these, 29 DAS genes were subject to regulation by 32 alternative splicing SNPs, showing diverse and specific splicing patterns and events. It is hypothesized that the *PIK3C2B* and *PPRPF8* splice variants associated with AS SNPs (rs42705933 and rs133847062) may be risk factors for susceptibility to bovine subclinical mastitis. Altogether, these key blood markers associated with resistance to subclinical mastitis and SNPs associated with alternative splicing of genes provide the basis for genetic breeding for resistance to subclinical mastitis in cows.

KEYWORDS

bovine subclinical mastitis, lncRNA, mRNA, alternative splicing, single nucleotide polymorphism, peripheral blood transcriptome

1 Introduction

Bovine mastitis is the foremost production and major economic burden confronted by the global dairy industry (Hoque et al., 2019). The disease may be asymptomatic or symptomatic (subclinical or clinical mastitis), persistent (chronic), or nonpersistent. Infectious agents such as *Staphylococcus aureus* (*S. aureus*), *Streptococcus agalactiae*, and *Mycoplasma spp* as well as environmental pathogens such as *Escherichia coli*, *Klebsiella pneumoniae,* and *Enterobacter aerogenes* are the main causes of mastitis in cattle (Gao et al., 2017; Hoque et al., 2020, 2019). In countries such as China, 10–40% of mastitis cases are caused by *S. aureus* (Wang et al., 2018). These bacteria enter the mammary gland and are recognized by the interaction of their pathogen-associated molecular patterns (PAMP), and the resulting induced inflammatory response leads to the shedding of somatic cells (SC) into the milk. It has been demonstrated that both the number of somatic cells count (SCC) and changes in their gene expression are associated with physiological processes in the mammary gland and bacterial infection (Leitner et al., 2000), with 200,000 cells/mL being used as the best cut-off point to distinguish between infected and uninfected (Schepers et al., 1997; Goncalves et al., 2018).

Blood is a mixture of various immune cells, such as lymphocytes, neutrophils, and monocytes, which have an advantage in reflecting traits such as immunity and disease resistance in cows (Chaussabel, 2015). Also, circulating leukocytes in the blood play a key role in the onset, development, and regression of mastitis, as they are the primary source of immune cells attracted to the mammary gland during infection (Cheng et al., 2021). Most importantly, blood is an easily accessible and least damaging tissue sample compared to bovine mammary epithelial cells, which can accurately reflect the physiological status and health of cows during lactation (Bai et al., 2016). Thus, gene expression profiles from blood offer new opportunities to clarify the mechanisms underlying the complex traits of cows. Blood biomarkers play an important role in characterizing the disease state of animals, and many studies have reported altered mRNA and long non-coding RNA abundance in the mammary gland after mastitis, with some stable markers such as lncRNA XIST (Ma et al., 2019a), lncRNA-TUB (Wang H. et al, 2019), LRRC75A-AS1 (Wang et al., 2020b) and lncRNAs PRANCR (Mi et al., 2021) were identified, but the role of gene transcription and complex networks in the blood remains unclear.

The diversity of functional protein products is mostly attributed to gene alternative splicing. Alternative splicing is a fundamental mechanism by which introns in pre-mRNAs are clipped and exons are bonded together in different configurations, leading to alterations in the structure of major transcripts (Baralle and Giudice, 2017). The previous findings showed that well over 60% of human genes undergo substitutive splicing (Modrek et al., 2001) which has a strong association with diseases (Faustino and Cooper, 2003; Urbanski et al., 2018; Yang et al., 2019). In cows, one study found that 4,567 of 21,755 bovine genes are alternatively spliced, and the most common AS event is exon skipping (Chacko and Ranganathan, 2009). These AS events are strongly associated with disease resistance in cows. For example, two differential alternative splicing (DAS) events in the genes *EXOC7* and *KIF2C* affect protein functional domains and are associated with the susceptibility to *Mycobacterium avium* subspecies paratuberculosis (Li et al., 2022). Furthermore, the specific alternative splicing patterns exhibited in genes *SLAMF7* (Ju et al., 2012a), *BOLA-DQA2* (Hou et al., 2012), and *CD46* (Wang et al., 2016) in mammary gland tissues of cows were associated with mastitis resistance in cows.

Moreover, 15% of point mutations on genomic DNA that lead to genetic diseases impact pre-mRNA splicing (Wang et al., 2016). Several splicing-related single nucleotide polymorphisms (SNP) may directly alter the coding region, leading to aberrant alternative splicing that affects the disease phenotype (Wang et al., 2014; Ren et al., 2021). In cattle, a study found that an independent spontaneous splice site variant in *COL2A1* (g.32473300 G > A) was most likely responsible for chondrodysplasia during early fetal development in cows. Similarly, the SNP g.5880C > T, SNP g.18174A > G, and SNP g.10766T > C might affect the binding with splicing-related factors, subsequently causing the production of aberrant splice variant *HMGB3-TV1* (Li et al., 2013; Wang et al., 2014), *NCF4-TV* (Ju et al., 2015), and *NCF1-TV1* (Zhang et al., 2015), thus increasing somatic cell scores in cows.

In the current study, we used RNA-seq to characterize gene expression patterns in the peripheral white blood cells transcriptome of healthy cows and subclinical mastitis cows. Furthermore, we obtained several candidate blood-based transcriptional biomarkers based on the comprehensive analysis of gene function annotation and gene expression patterns. Afterward, we further analyzed alternative splicing events at the transcription level and searched for splicing-associated mutations to elucidate the molecular regulatory mechanisms of bovine subclinical mastitis resistance.

2 Materials and methods

2.1 Sample collection and preparation

Three healthy cows and three subclinical mastitis cows were chosen according to three consecutive months of SCC records and other related phenotypes. In this study, cows with milk SCC values below 100,000 cells/mL were regarded as healthy cows. In contrast, cows with SCC values in the range of 200,000 cells/mL and 500,000 cells/mL were considered to be suffering from subclinical mastitis cows. For similarity in the biological background, the cows in this study had similar lactation days and all calved one litter (Supplementary Table S1).

About 20 mL of anticoagulant blood sample was collected from each cow. About 10 mL of anticoagulant blood sample was centrifuged at 3,500 r/min for 15 min and the middle leukocyte layer was collected and stored in 1 mL of TRIzol at -80 °C for RNA extraction. In addition, about 2 mL blood sample was sent to Jinhaikeyu company for HP testing (Beijing, China). A further amount of blood was used for genotyping at GGP Bull 150K BeadChip (Neogene, Lansing, MI, United States).

2.2 Hematological parameters

A Sysmex K-4500 automated hematology analyzer (Sysmex Corporation, Kobe, Japan) was used to test 24 hematological parameters, including White Blood Cells (WBC), Red Blood Cells (RBC), Hemoglobin (HGB), Red blood cell-specific volume (HCT), Mean Corpuscular Volume (MCV), Mean Corpuscular Hemoglobin (MCH), Mean Corpuscular Hemoglobin Concentration (MCHC), Platelet count (PLT), Neutrophil Ratio (NETU%), Neutrophil count (NETU#), Lymphocyte Ratio (LYMPH%), Lymphocyte count (LYMPH#), Monocyte Ratio (MONO%). Monocyte count (MONO #), Eosinophil Ratio (EO %), Eosinophil count (EO#), Basophil Ratio (BASO%), Basophil count (BASO#), Platelet Distribution Width (PDW), Mean Platelet Volume (MPV), Red cell distribution width - stand error (RDW-SD), Red blood cell distribution width - coefficient of variation (RDW- CV), Platelet-Large Cell Ratio (P-LCR), Platelet cubic metric distribution width (PCT).

2.3 RNA extraction and RNA sequencing

Total RNA of peripheral white blood cells for each cow was extracted with the TRIzol reagent (Invitrogen, Carlsbad, CA, United States) following the manufacturer's procedures. The concentration of RNA was estimated using the NanoDrop 2000 (ThermoFisher Scientific, Waltham, MA, United States) and the RNA Nano 6000 Assay Kit of the Bioanalyzer 2,100 system was used to assess RNA integrity. The RNA quality was checked for contamination and degradation by a 1% agarose gel. Then, qualified RNA was used to construct RNA-seq libraries. Finally, cDNA libraries were sequenced with 150 bp paired-end reads from Illumina NovaSeq 6,000 platform (Novogene, Beijing, China).

2.4 RNA-Seq data analysis

2.4.1 Reads quality control and mapping

The quality of the raw data was evaluated with FastQC version 0.11.8. The clean data were obtained by using Trimmomatic version 0.38 to filter out reads with adapter sequences and low-quality reads from the raw reads (Bolger et al., 2014). Specifically, remove splice sequences, bases with mass less than three at the 5′ and 3′ ends of reads, and all bases with average mass less than 15 in the window; trim reads with sequence length less than 36. Subsequently, paired-end reads from each sample were aligned to the reference genome using STAR version 2.7.7a (Dobin et al., 2013). The generation of SAM files was sorted into BAM files using SAMtools version 1.9 (Li et al., 2009). Subsequently, each sample was assembled and merged using StringTie version 1.3.5 (Pertea et al., 2015).

2.4.2 Prediction novel lncRNA

In the current study, strict filtering conditions were set on the annotated transcripts to obtain novel lncRNA transcripts. Firstly, transcripts shorter than 200bp and less than two exons were discarded. Then, transcripts with the class codes "i" (intronic lncRNA, ilncRNA), "u" (Intervening lncRNAs, lincRNA), and "x" (antisense lncRNA, lncNAT) were preserved. Finally, protein-coding potential predictions were performed using the Coding-Non-Coding Index (CNCI) (Sun et al., 2013), ORF Length and GC content (LGC) (Wang G. et al., 2019), Coding Potential Calculator (CPC) (Kang et al., 2017), and Predictor of Long non-coding RNAs and messenger RNAs based on an improved K-mer scheme (PLEK) (Li et al., 2014), and only those transcripts at the intersection of the four software tools 132 were selected as novel lncRNA transcripts.

2.4.3 Quantification and Identification of Differentially Expressed Genes

Transcripts were quantified using HTSeq-counts version 2.0.1 and results were expressed as read counts (Anders et al., 2015). Subsequently, normalization of reads and differential expression analysis of reads were performed using the DESeq2 (Love et al., 2014). Differentially expressed genes (DEGs) and differentially expressed lncRNAs (DELs) were screened with the criteria of p-value < 0.05 and log2 (fold change) > 1.5.

2.4.4 Prediction of the *cis* and *trans*-target genes of lncRNAs

To explore the functions of lncRNAs, we simultaneously predicted the *cis*- and *trans*-target genes of lncRNAs. Protein-

coding genes near the transcript positions of lncRNAs (100k upstream and downstream) were selected as *cis*-target genes of lncRNAs with the BEDTools version2.1.2 (Quinlan and Hall, 2010). *Trans*-target genes were predicted based on Pearson correlation coefficients between DELs and DEGs calculated by the R package-Hmisc, with screening criteria of $r > 0.98$ and p-value < 0.05.

2.4.5 Gene Alternative Splicing Analysis

In the alternative splicing analysis, alternative splicing events were detected using rMATS version 4.1.1 (Shen et al., 2014). In this study, five alternative splicing patterns, namely exon skipping (ES), retained intron (RI), mutually exclusive exon (MXE), alternative 5′ splice site (A5SS), and alternative 3' splice site (A3SS), were detected from healthy cows and subclinical mastitis cows using rMATS software version 4.1.1. Statistical verification (FDR <0.05) was performed when the differences in isoform ratio (IncLevelDifference) of different alternative splice genes (DAS genes) exceeded the defined threshold of 0.01%. Static images of sarcograms were generated outside of IGV using sarcograms, which is a python tool that is part of the MISO package.

2.4.6 Functional annotation of SNPs

In this study, all animals were genotyped using the GGP Bull 150K BeadChip (Neogene, Lansing, MI, United States) and the effects of all SNPs were annotated by ensemble Variant Effect Predictor (VEP) to screen for SNPs associated with alternative splicing. Subsequently, SNPs were localized to target genes based on positional information to identify candidate SNPs associated with alternative splice genes.

2.5 Functional enrichment analysis

Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis and gene ontology (GO) analysis of DEGs, target genes of DELs, and DAS genes were performed on the KOBAS 3.0 online website (http://kobas.cbi.pku.edu.cn/kobas3/).

2.6 Construction interaction network of DELs, DEGs, and pathways

A regulatory network of DELs, target genes of DELs, and pathways were constructed using Cytoscape version 3.5.1 (Yin et al., 2020)

2.7 Statistical analysis

Linear regression analysis was performed using GraphPad Prism (version nine; GraphPad Software, San Diego, CA, United States) and Student's t-test was used to examine significant differences in the expression levels of lncRNAs and mRNAs between healthy and subclinical mastitis cows. The ggpubr was used to analyze correlations between routine blood test parameters and standardized read counts of genes.

3 Results

3.1 Overview of RNA-seq data

RNA-seq was performed on six cDNA libraries using the Illumina NovaSeq 6,000 platform. The total reads mapped to the bovine genome have been summarized in Table 1. A total of 12.79 million raw reads per library and 12.59 million clean-read pairs were generated from all cDNA libraries. The GC contents of the reads ranged from 50 to 52%. The average alignment rate of clean reads to the bovine genome (version: ARS-UCD1.2) was 95.89%, while 90.91% of clean reads were uniquely mapped reads. Reads that uniquely mapped to the reference genome were used for further analysis (Table 1).

3.2 Identification and characterization of lncRNAs in subclinical mastitis cows vs. healthy cows

The identification of novel lncRNAs transcripts was followed by a series of criteria. Upon transcript filtering and protein-coding potential prediction, 5,510, 5,664, 5,296, and 5,165 lncRNAs transcripts were predicted by CPC, CNCI, LGC, and PLEK, respectively (Figure 1A). Of particular note, 4,424 lncRNAs transcripts were identified as novel lncRNAs transcripts, which were evaluated by four prediction tools. Of the 4,424 identified lncRNAs transcripts, 76.29% had two or three exons, while 0.52% had more than six exons (Figure 1B).

In the peripheral blood of healthy cows and subclinical mastitis cows, a total of 6,365 lncRNAs were screened, including 1,941 known lncRNAs transcripts. These transcripts were widely distributed in all chromosomes of the bovine genome, and the highest number of transcripts was on chromosome five (Figure 1C). lncRNAs were classified according to their genomic location, with intervening lncRNAs having the highest proportion (47.11%) of all novel lncRNAs, followed by antisense lncRNAs (26.92%) and the lowest percentage of intronic lncRNAs (18.85%) (Figure 1D).

3.3 lncRNAs and mRNA expression profile changes in subclinical mastitis cows vs. healthy cows

A total of 287 differentially expressed genes (DEGs) and 70 differentially expressed lncRNAs (DELs) were identified

TABLE 1 Summary of reads mapped to the bovine genome.

Cows	Raw Reads	Clean Reads	Total Mapped Rate (%)	Uniquely Mapped Rate (%)
HY1	20,989,867	20,691,802	96.15	91.51
HY2	19,257,369	18,970,017	96.32	91.39
HY3	24,010,276	23,632,252	96.70	92.13
SM1	21,807,528	21,481,669	96.03	90.67
SM2	23,137,865	22,746,107	96.17	90.44
SM3	18,669,665	18,418,778	93.96	89.31

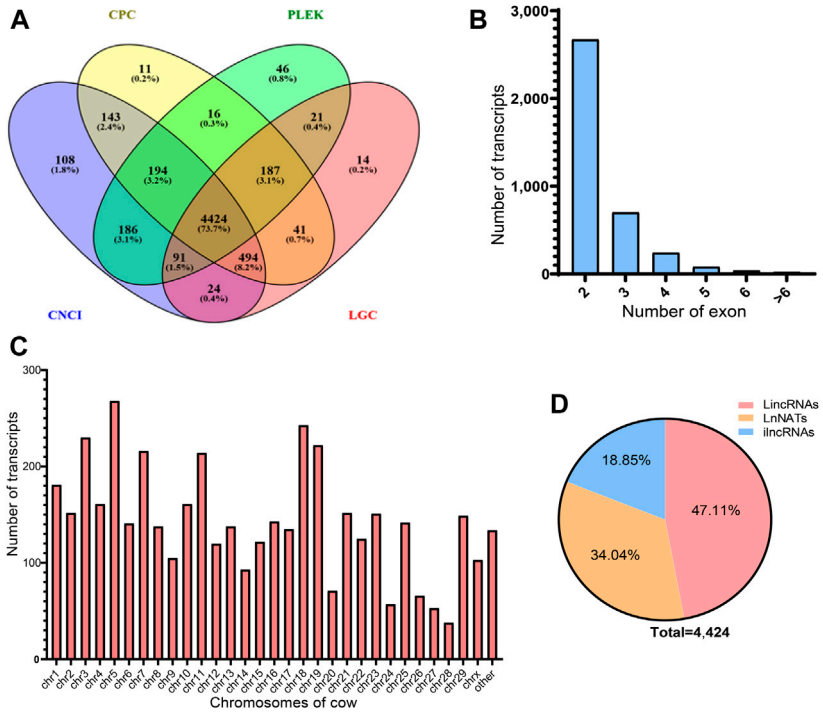


FIGURE 1
Identification and characterization of lncRNAs in cows with subclinical mastitis and healthy cows. **(A)** The Venn diagram of novel lncRNAs. The intersection part showed the novel lncRNAs predicted by the coding potential prediction software CPC, PLEK, CNCI, and LGC together. **(B)** Exon number of identified lncRNAs transcripts. **(C)** Distribution of identified lncRNAs transcripts on all chromosomes. **(D)** Among 4,421 lncRNA transcripts, the proportion of intronic lncRNA (ilncRNA), Intervening lncRNAs (lincRNA), and antisense lncRNA (lncNAT).

between the two groups (p-value < 0.05, log(fold change) > 1.5) (left panel in Figures 2A,B). Compared with healthy cows, 253 genes and 57 lncRNAs were significantly up-regulated, while 34 genes and 13 lncRNAs were significantly down-regulated in subclinical mastitis cows. The clustering results of the heatmaps showed that the expression patterns of lncRNAs and mRNAs were quite different between healthy and subclinical mastitis cows (right panel in Figures 2A,B).

Next, the functions of 287 DEGs were investigated. As shown in Figures 3A, B, 287 DEGs were involved in 301 GO terms (p-value < 0.05), which contained 55 molecular compositions (CC), 190 biological processes (BP), and 56 molecular functions (MF). Furthermore, KEGG results of DEGs showed 17 enriched signaling pathways (p-value < 0.05) (Figure 3B). Among them, the NOD-like receptor signaling pathway was enriched with the highest number of DEGs (Supplementary Table S2). The genes *FCAR*, *PTAFR*, and *SELPLG* are involved in the *S. aureus* infection pathway, among which *SELPLG* is also involved in the cell adhesion signaling pathway (Supplementary Table S2). Notably, the genes *SLC40A1* and *SLC11A2* were involved in the ferroptosis signaling pathway, and *SLC40A1* was up-

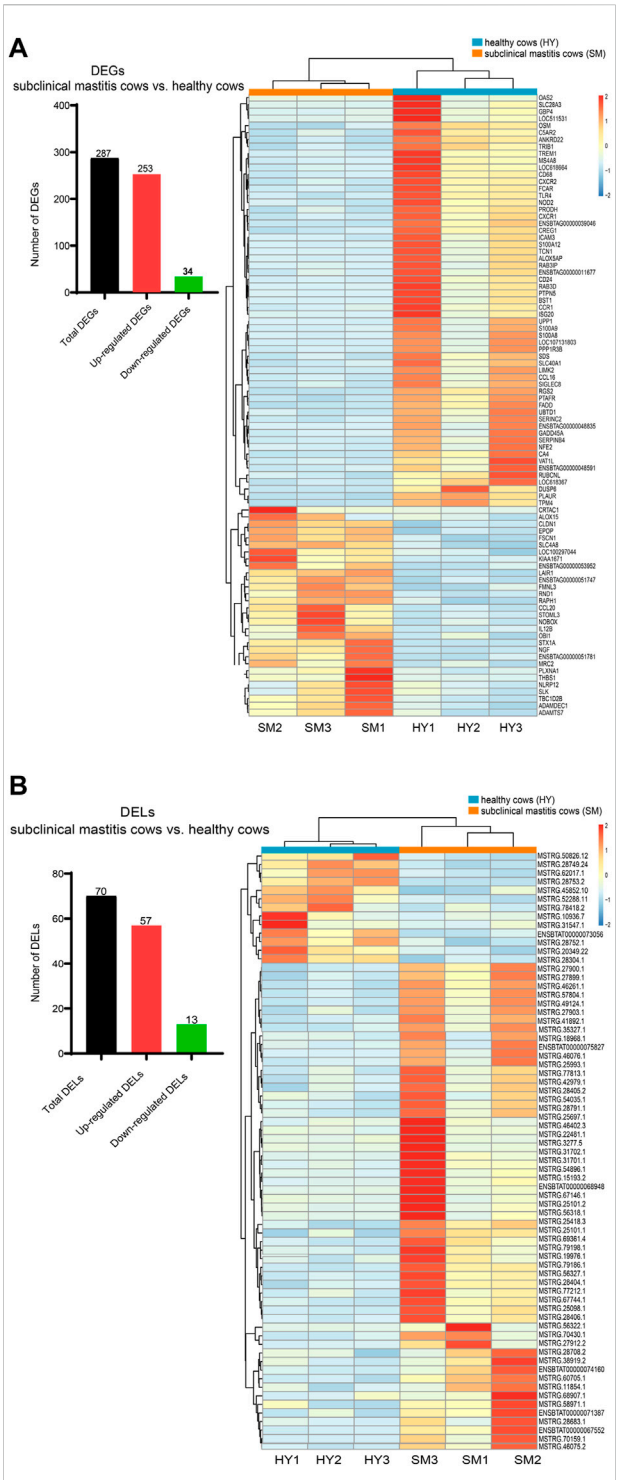


FIGURE 2
lncRNAs and mRNA expression profile changes in subclinical mastitis cows *vs.* healthy cows. The number **(A)** and the heatmap **(B)** differentially expressed genes (DEGs) and differentially expressed lncRNAs (DELs) in subclinical mastitis (SM) cows *vs.* healthy (HY) cows (p-value < 0.05, log2 (fold change) > 1.5).

regulated (Figure 3B) in subclinical mastitis cows *vs.* healthy cows.

We also investigated the relationship between DEGs and routine blood test parameters and found that the standardized read counts of genes *DCK*, *ARGE*, *VNN2*, and *OAS2* were significantly correlated with White Blood Cell (WBC) count (10^9/L) (p-value < 0.05) (Figure 4A). Similarly, the standardized read counts of genes *GBP6* and *TBXAS1* were significantly correlated with the Monocyte Ratio (MONO %) in peripheral blood (p-value < 0.05) (Figure 4B).

3.4 Combined DELs and DEGs of healthy cows and subclinical mastitis cows for exploring their co-expression

Potential target genes of DELs were predicted by co-localization and co-expression. In this study, 40 *cis*-target genes and 124 *trans*-target genes were predicted (Supplementary Table S2). The results of GO analysis showed that target genes of DELs were involved in 197 BP, 44 CC, and 55 MF (Figure 5A). Most of the target genes were located in cytoplasmic vesicles, outside of the plasma membrane and were involved in the positive regulation of interleukin-17 production, inflammatory response, and cell adhesion (Figure 5A).

The results of the KEGG analysis showed 12 pathways (p-value < 0.05) (Figure 5B). Most of these pathways were related to immune responses and inflammation, like NOD-like receptor, p53, IL-17, and *S. aureus* infection pathway. Besides, the genes *FCAR*, *PTAFR*, and *SELPLG* are involved in the *S. aureus* infection pathway, among which *SELPLG* is also involved in the cell adhesion signaling pathway (Supplementary Table S2). Notably, most DELs and their target genes showed the same expression trends (Supplementary Table S2).

3.5 Identification of candidate biomarkers associated with subclinical mastitis

In this study, the functions of DELs were further annotated by comparing the genomic positions of DELs and QTL-related regions in Cattle QTLdb. The results showed that 245 DELs were located adjacent to 481 QTLs in Cattle QTLdb (Figure 5C). These QTLs surrounding DELs were associated with milk production performance and reproductive traits, with more than 75% of QTLs being associated with milk (Supplementary Table S3). More importantly, 35 DELs were located adjacent to three QTLs related to mastitis, such as SCC QTL, somatic cell score (SCS) QTL, and white blood cell count QTL (Supplementary Table S3). What's more, it was found that 32 lncRNAs were adjacent to stillbirth and calving ease QTLs (Supplementary Table S3), of which 15 DELs

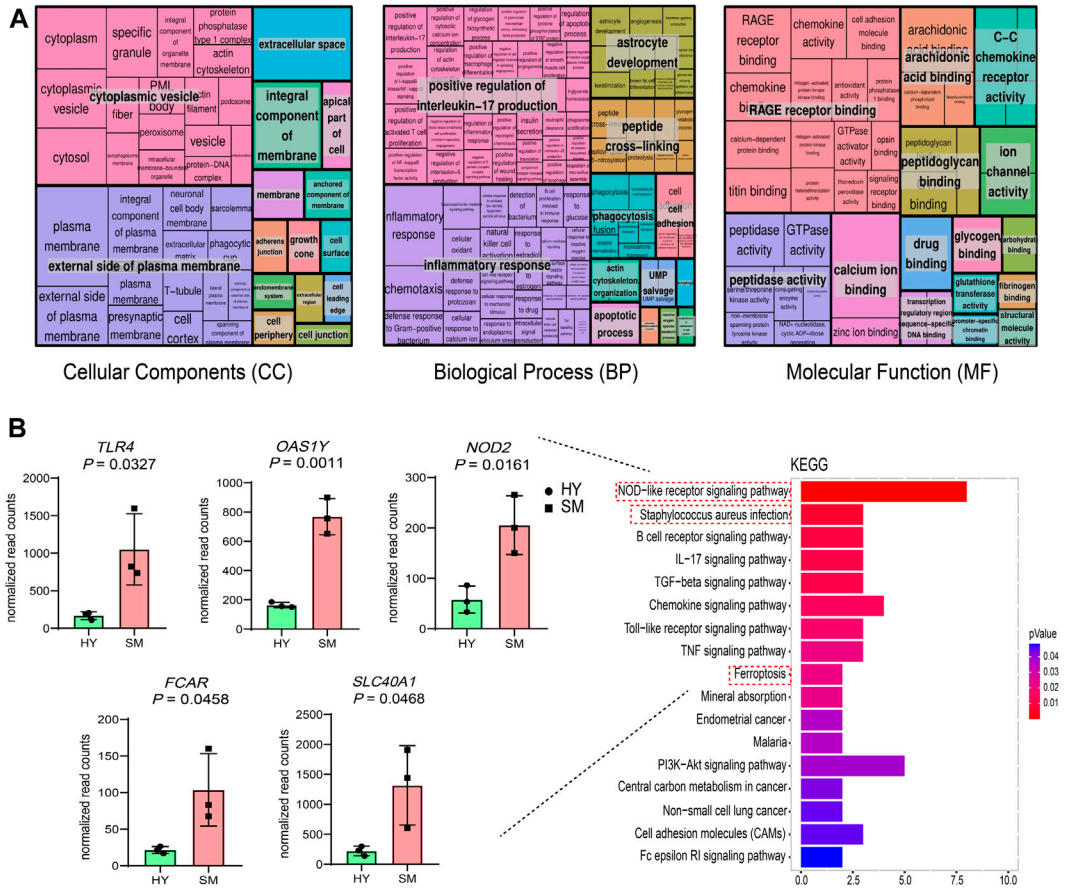


FIGURE 3
Results of GO and KEGG analysis of 287 differentially expressed genes (DEGs) in subclinical mastitis (SM) cows *vs.* healthy (HY) cows. **(A)** REVIGO summary and visualization of the significant Go terms (p-value < 0.05), including 190 GO terms of biological process (BP), 50 GO terms of cellular components (CC), and 56 GO terms of molecular function (MF). **(B)** KEGG pathways of DEGs (p-value < 0.05), and expression differences of several DEGs involved in the *S. aureus* infection, cell adhesion molecules, and ferroptosis pathway.

were also adjacent to white blood cell QTL, SCS QTL, and SCC QTL, such as MSTRG.10101.2, MSTRG.28749.24, MSTRG.46076.1, MSTRG.67501.5, and MSTRG.23805.2, MSTRG.28197.1 (Supplementary Table S3), and so on.

3.6 Co-expression networks revealed the regulatory mechanism of lncRNAs in subclinical mastitis

Previous studies have shown that lncRNAs may impact disease resistance of livestock and poultry by regulating the expression levels of their *cis*-and *trans*-target genes. In this study, five signaling pathways were shared by DEGs and DELs, including NOD-like receptor, *S. aureus* infection, cell adhesion, B cell receptor, and IL-17 signaling pathways, involving 18 DEGs, 37 DELs, and 7 DEGs. As shown in Figure 6, the network included 117 nodes, 133 edges, and lncRNAs MSTRG.25101.1, MSTRG.25098.1, MSTRG.25418.3, MSTRG.56318.1, MSTRG.15193.2, MSTRG.56327.1, MSTRG.67146.1, and MSTRG.54896.1 may play key nodes to regulate the expression levels of core genes *OAS1Y*, *GBP4*, *OAS2*, *FCAR*, *SELPLG*, *DCK*, and *ICAM3*. This result implied the important role of lncRNAs in bovine subclinical mastitis.

3.7 Identification and characterization of total alternative splicing events and differential alternative splicing events

A comprehensive analysis of alternative splicing events in peripheral blood was also performed. As shown in Figure 7A, a

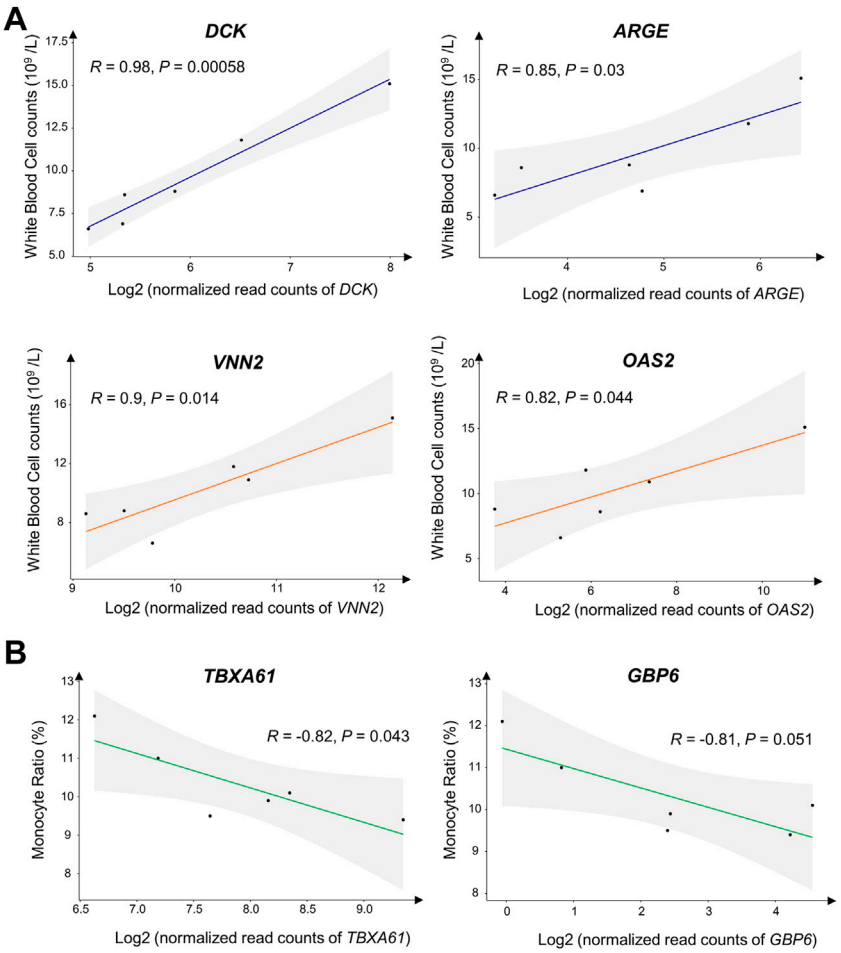


FIGURE 4
Linear analysis of normalized read counts for several differentially expressed genes (DEGs). and routine blood test parameters. **(A)** DEGs are significantly associated with White Blood Cell (WBC) count (10^9/L). **(B)** DEGs are significantly associated with Monocyte Ratio (MONO %).

total of 65,839 alternative splicing (AS) events were found between healthy and subclinical mastitis cows (*FDR* < 0.05), including 2,212 differential alternative splicing (DAS) events. Five patterns of AS consist of exon skipping (ES), retained intron (RI), mutually exclusive exon (MXE), alternative 5′ splice site (A5SS), and alternative 3' splice site (A3SS). A total of 2,212 events were segregated into 82.59% ES events, 11.66% MXE events, 2.44% A3SS and A5SS events, and 3.30% RI events (Figure 7A). More than two types of AS events were presented in 139 genes, 89 genes showed both MXE and SE events, 18 genes showed both SE and RI events (Figure 7B), and the genes *CCDC51*, *PIK3CG*, *MCEMP1*, *STRADA*, and *ATP2A3* presented three types of AS events.

Further analysis revealed that 60 and 13 AS genes presented RI events, 16 and 17 AS genes presented A5SS events, 13 and 8 AS genes presented A3SS events, and the same number of genes presented MXE events in healthy and subclinical mastitis cows, respectively (Figure 7C). Notably, the number of SE events increased two-fold in subclinical mastitis cows compared to healthy cows (Figure 7C).

3.8 Several DAS genes were mastitis-specific and involved in immune, inflammatory response, and reproduction-related pathways

Through the integration of DAS events that belong to the same gene, 2,212 DAS events were found to correspond to 1,621 unique genes (Figure 8A). Gene Ontology and KEGG pathway analysis were used to characterize the primary functions of AS genes. In all three functional categories, AS genes were annotated to different GO terms, including 142 GO terms for CC, 455 GO terms for BP, and 172 GO terms for MF (Figure 8B), affecting insulin binding, regulation of bacterial entry into host cells, and response to

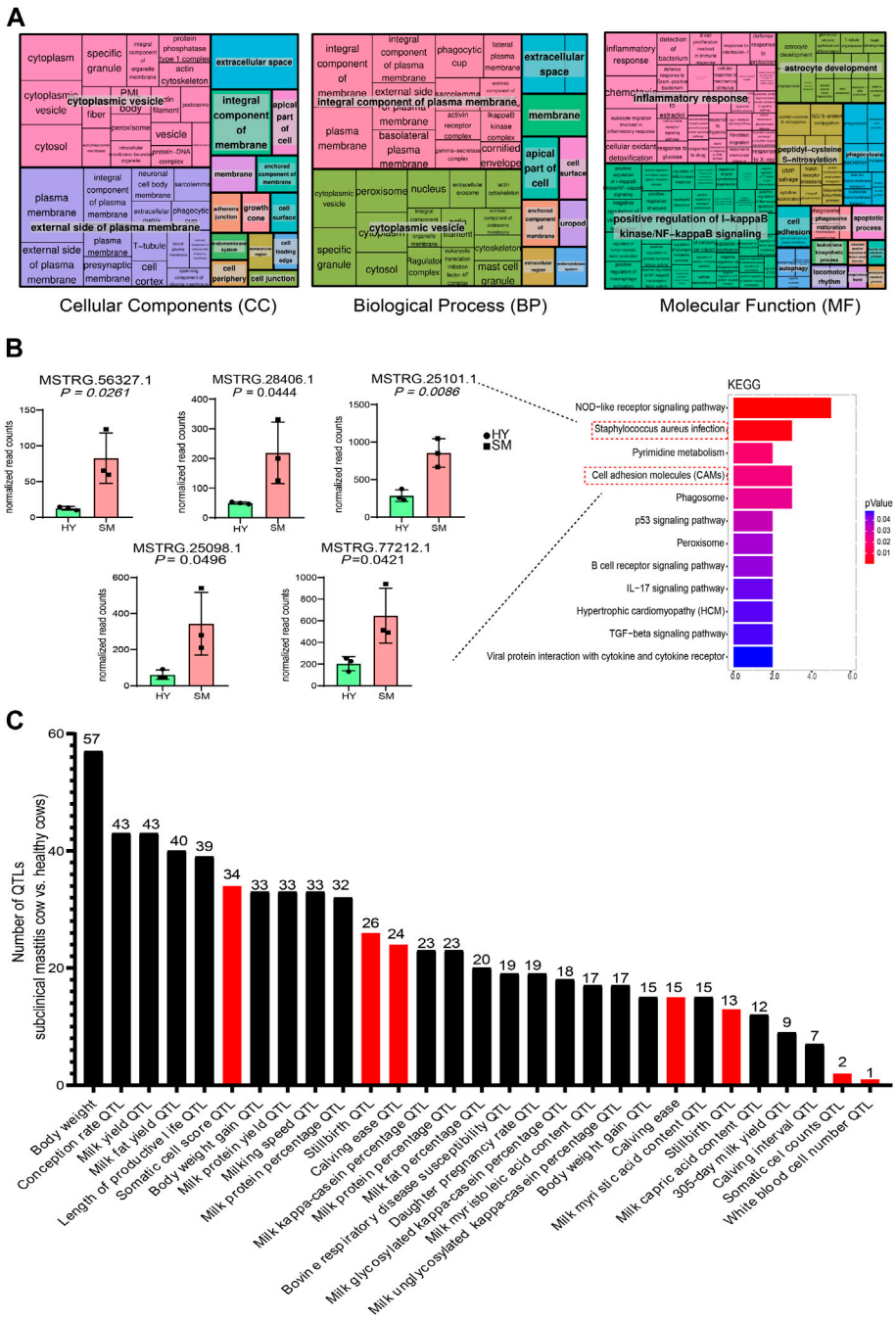


FIGURE 5
Results of GO and KEGG analysis of lncRNAs *cis*-and *trans*-target genes. **(A)** Significant GO terms (*p*-value < 0.05) for *cis*- and *trans*-target genes of lncRNAs, including 197 GO terms for biological processes (BP), 44 GO terms for cellular components (CC) and 55 GO terms for molecular functions (MF). **(B)** Significant KEGG pathways for *cis* and *trans* target genes of lncRNAs (*p*-value < 0.05), and expression differences of the differentially expressed genes (DEGs) involved in the enrichment of the NOD-like receptor signaling, *S. aureus* infection, and cell adhesion pathways. **(C)** Annotation results of DELs in subclinical mastitis cows compared to healthy cows, with red bars showing DELs associated with reproductive traits and mastitis.

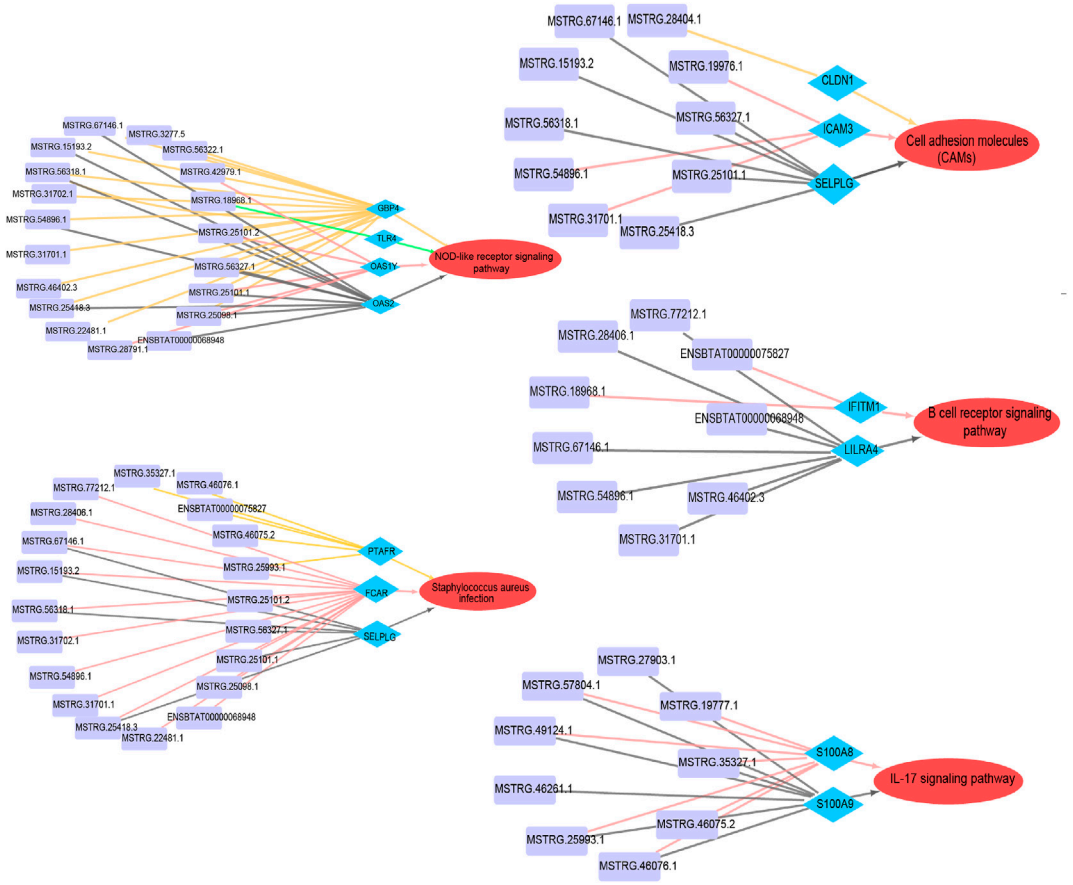


FIGURE 6
Co-expression networks of lncRNAs *cis*- and *trans*- target gene pathways; In particular, these five pathways were shared by differentially expressed genes (DEGs) and differentially expressed lncRNAs (DELs), including the NOD-like receptor, *S. aureus* infection, and cell adhesion (CAMs), B-cell receptor, and IL-17 signaling pathways. Purple rectangles represents DELs, blue diamonds represents DEGs, red ellipses represents critical paths, and different target relationships in the same regulatory network are distinguished by different colored lines.

ionizing radiation. KEGG pathway analysis further revealed that DAS genes were enriched in 154 pathways, which were classified into five categories in the KEGG database.

At one hierarchical level of the KEGG results (Figure 8C), a total of 42 signaling pathways associated with the organism system were enriched (Supplementary Table S4). In particular, the highest proportion of signaling pathways related to the immune system were reported, such as chemokines, Toll-like receptor, NOD-like receptor, hematopoietic cell lineage, natural killer cell-mediated cytotoxicity, IL-17, and leukocyte transendothelial migration signaling pathways (Figure 8C), which comprised 134 unique genes, representing 62.62% of the catalog (Supplementary Table S4). As a matter of note, 32 DAS genes were found involved in reproductive trait-related signaling pathways, such as thermogenesis, oxytocin, estrogen, and progesterone-mediated oocyte maturation signaling pathways (Supplementary Table S4). Hierarchical level two showed that 18 signaling pathways are involved in the regulation of mastitis, including MAPK, Ras, cGMP-PKG, NF-κb, phosphatidylinositol, and mTOR, wnt, and TNF pathways (Figure 8C). At the third level, a total of 19 pathways were identified as being related to cellular processes.

Worthy of note, the genes *SLC40A1*, *ACSL6*, *GCLC*, *ACSL5*, *CYBB*, *VDAC3*, and *FTH1* were also enriched in the ferroptosis signaling pathway (Figure 8C). Of these, *SLC40A1* was up-regulated and also enriched in the ferroptosis signaling pathway as DEGs (Figure 3B).

3.9 SNPs rs42705933 and rs133847062 may impact mastitis resistance by altering the alternative splicing patterns of *PIK3C2B* and *PRPF8*

From the results of the annotation of ensemble Variant Effect Predictor (VEP) on all SNPs, 51.97% of the SNPs were intronic

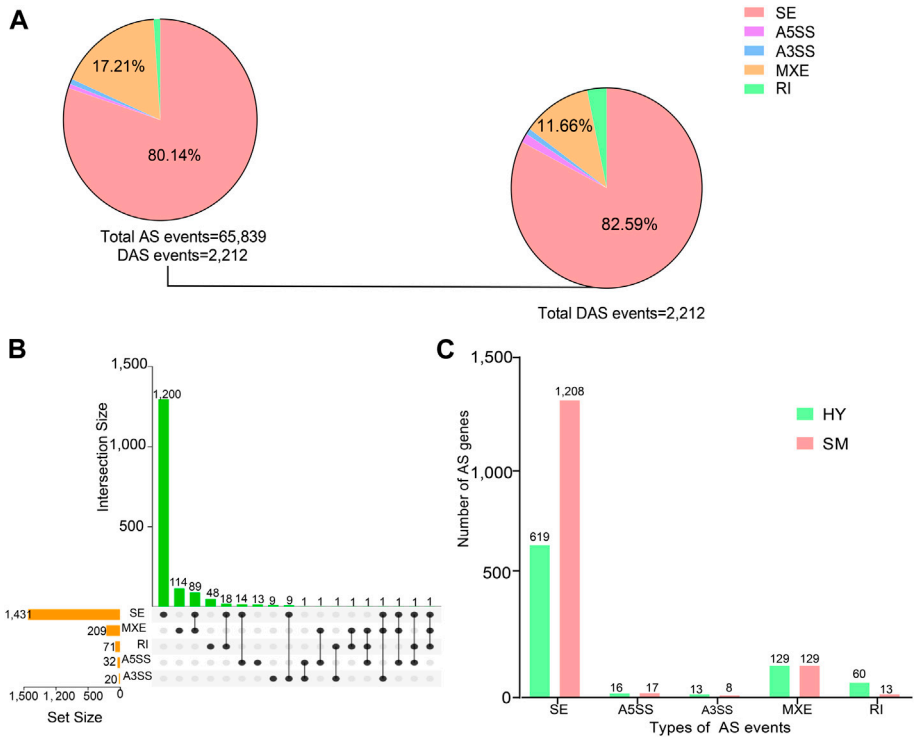


FIGURE 7
Identification and characterization of total alternative splicing (AS) events and differential AS (DAS) events. **(A)** Distribution of five AS types including skipping exon (SE), alternative 5′ splice site (A5SS), alternative 3′ splice site (A3SS), mutually exclusive exon (MXE), and retained intron (RI) among 65,839 AS events and 2,212 unique DAS events (*FDR* < 0.05). **(B)** Statistical results of shared and specific AS events in two groups. The bar graph on the left shows the number of raw AS events. The connections between points indicated the presence of intersections between corresponding data sets. **(C)** The number of five AS event genes in the healthy and subclinical mastitis cows.

variants, 34.00% were intergenic variants, and only 0.02% of the variants were associated with alternative splicing (Supplementary Figure S1), which included spliced donor region variants and spliced polypyrimidine tract variants. Ultimately, a total of 32 SNPs associated with alternative splicing were screened, corresponding to 29 unique alternative splicing genes (Table 2), including 23 genes that exhibited SE events, three genes that exhibited MXE events, and three genes that exhibited A5SS events, A3SS events, and RI events (Table 2).

To explore the mechanism by which alternative splicing variants regulate subclinical mastitis resistance, it was further investigated the genotype of these 32 AS SNPs between healthy and subclinical mastitis cows. Two AS SNPs were associated with subclinical mastitis resistance, and the genotypes of these AS SNPs were completely dissimilar between the healthy and subclinical mastitis cows (Table 3). Specifically, SNP rs42705933 on chromosome 16 at 2,210,365 bp of gene *PIK3C2B* had the genotype AA in healthy cows and AG in subclinical mastitis cows. Furthermore, SNP rs133847062 on chromosome 19 at 22,777,382bp of gene *PRPF8* had the genotype AC in healthy cows and AA in subclinical mastitis cows (Table 3). Notably, the two DAS genes both presented exon skipping compared to the healthy cows (Figure 9A).

To gain more insight into the interactions of these two mastitis resistance-related DAS genes with DEGs, a PPI network was constructed for functional association analysis. As a result, 110 interactions were found to be present in 37 genes (interaction score >0.4). As shown in Figure 9B, some proteins were able to interact, such as *GBP4* could interact with *OAS1Y* and *OAS2*, *MYD88* could interact with *S100A9*, and *TLR4* could interact with *CXCL8*, *NOD2*, and *PTAFR*. On the other hand, one protein interacts with another protein through a third protein. For example, *OAS2* could interact with *TBXAS1* via *PTGS1*, *PIK3C2B* could interact with *AREG* via *EGFR*, and *CXCR1* could interact with *S100A8* via *CXCL6*. The above results suggest that these proteins play a synergistic role in the regulation of bovine subclinical mastitis.

4 Discussion

In this study, the expression profiles of mRNAs and lncRNAs from healthy cows and subclinical mastitis cows

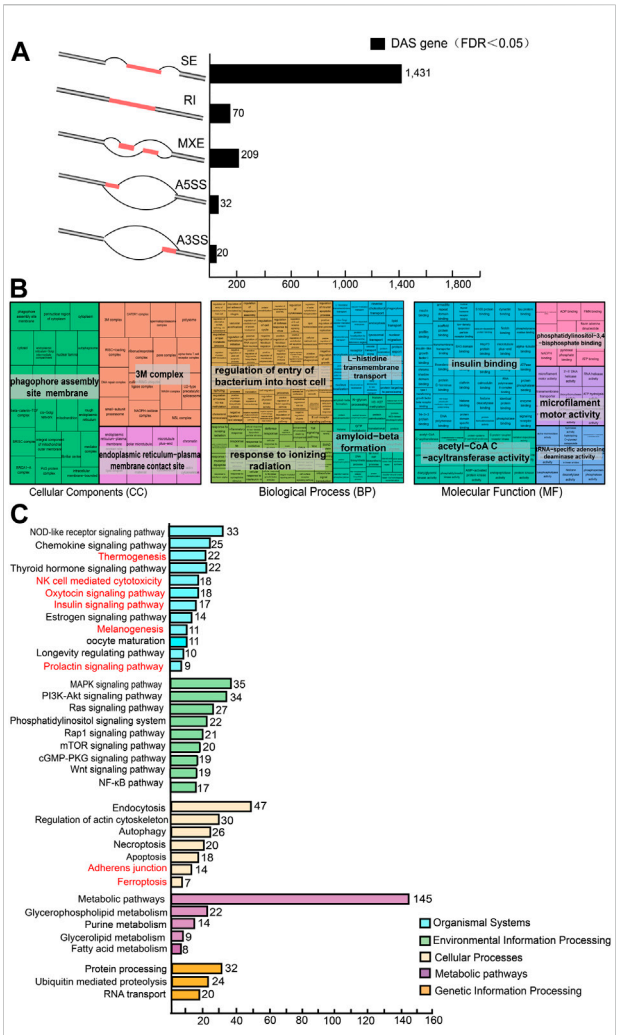


FIGURE 8
Functional characterization of differential alternative splicing (DAS) genes. **(A)** Distribution of the five essential AS events detected from 1,621 DAS genes, and a schematic illustration of the five AS models, including skipping exon (SE), alternative 5′ splice site (A5SS), alternative 3′ splice site (A3SS), mutually exclusive exon (MXE), and retained intron (RI). **(B)** Significant GO terms for 1,621 DAS genes (p-value < 0.05), including 455 GO terms for biological processes (BP), 142 GO terms for cellular components (CC) and 172 GO terms for molecular functions (MF). **(C)** Significant KEGG pathways of 1,621 DAS genes (p-value < 0.05), divided into five types, with different colors representing different types.

were characterized, and the functions of 287 DEGs, 70 DELs, and 1,621 DAS genes were annotated, obtaining a number of significantly enriched pathways. Notably, the NOD-like receptor signaling pathway, cytokine-cell receptor interaction, Toll-like receptor signaling pathway, and *S. aureus* infection pathway were also significantly enriched in the available studies of bovine mammary epithelial (BME) cells (Wang et al., 2016; Wang et al., 2020a; Lin et al., 2021), suggesting that these pathways play important regulatory roles in the pathogenesis of mastitis in blood and breast tissue. Subsequently, based on target gene prediction and QTL database annotation, the results showed that the lncRNAs MSTRG25101.2, MSTRG.56327 0.1, and MSTRG.18968.1, which are adjacent to the SCS QTL and SCC QTL, regulate the expression levels of genes *NOD2*, *CXCL8*, OAS2, and *TLR4*, which are involved in the NOD-like receptor signaling pathway, Cytokine-cytokine receptor interaction, and Toll-like receptor signaling pathway.

The innate immune system is the host's first line of defense against pathogenic microorganisms and relies on pattern recognition receptors (PRRs) (Kawai and Akira, 2010; Celhar et al., 2012; Kawasaki and Kawai, 2014). These PRRs recognize molecular features expressed on microorganisms and the interaction of PRRs with microbe-associated molecular patterns (MAMP) can lead to the expression of pro-inflammatory cytokines and other immunomodulatory molecules (Gilbert et al., 2013). Among several different types of PRRs in cows, Toll-like receptors (TLRs) and NOD-like receptors (NLRs) have been extensively studied (Bhattarai et al., 2018). In the present study, genes *NOD2*, *CXCL8*, *OAS2*, and *TLR4* were differentially expressed in the blood transcriptome of subclinical mastitis cows and involved in the NOD-like receptor, Toll-like receptor pathway and were identified as key candidates for this study. *TLR4* and *NOD2*, as key PRRs, and their recognition of MAMP have been shown to be a key event in the development of mammary inflammation (Goldammer et al., 2004; Aitken et al., 2011; Porcherie et al., 2012). *CXCL8* is an important neutrophil chelator in BME cells (Wellnitz and Bruckmaier, 2012), induces chemotaxis in the target cells, and has a significant change in expression levels upon *TLR4* activation, which could serve as a potential biological for improving the outcome of mastitis markers (Islam et al., 2020).

Ferroptosis is a regulatory form of iron-dependent cell death (Dixon et al., 2012; Zuo et al., 2022), and is characterized by abnormal iron metabolism (Yang and Stockwell, 2016), lethal lipid peroxidation, and reactive oxygen species accumulation. Several studies have demonstrated that ferroptosis plays a vital regulatory role in the pathology of cancers such as lung cancer (Yang and Stockwell, 2016), gastrointestinal cancer (Zhu et al., 2020), and breast cancer (Xu et al., 2021; Zhang et al., 2021). Yet, few studies reported the relationship between ferroptosis and subclinical mastitis of cows, and the precise mechanism of programmed cell death in bovine mastitis is not fully understood. In our study, KEGG results showed that *SLC40A1* (Deng et al., 2021), *SLC11A2* (Weijiao et al., 2021), *ACSL6*, *GCLC* (Chang et al., 2021; Wang et al., 2021), *ACSL5*, *CYBB* (Ping et al., 2022), *VDAC3* (Yang et al., 2020; Zhu et al., 2021), and *FTH1* (Liu et al., 2022) were involved in the ferroptosis signaling pathway (Supplementary Table S4). Among them, *SLC40A1* was up-regulated in subclinical mastitis cows compared to healthy cows (Figure 3B). Moreover, *SLC40A1*, *ACSL6*, *GCLC*,

TABLE 2 AS SNPs of all SNPs and the AS events of their target genes.

SNP ID	Gene symbol	Location	Allele	*p*-Value	*FDR*	AS Events
rs137001652	*MED15*	17:72,433,357	G	1.35E-10	1.70E-07	MXE
rs134643460	*CDC7*	3:51,880,683	C	9.51E-08	8.14E-06	RI
rs109129224	*FGR*	2:125,704,116	C	9.31E-08	5.58E-05	SE
rs111019940	*NDC1*	3:92,418,182	G	3.24E-07	0.000143	SE
rs134557868	*ACSL6*	7:22,527,204	T	7.16E-06	0.001195	SE
rs385381247	*ACSL6*	7:22,502,848	T	7.16E-06	0.001195	SE
rs29014580	*MRE11*	15:1,472,757	G	7.43E-06	0.001221	SE
rs137215894	*ATP8B4*	10:60,229,145	C	8.72E-06	0.00136	SE
rs132674081	*UBR4*	2:133,508,452	A	1.01E-05	0.001517	SE
rs110095083	*CCDC77*	5:107,313,161	C	1.46E-05	0.001932	SE
rs136092614	*DID O 1*	13:54,507,164	A	1.84E-05	0.002255	SE
rs109892136	*MY O 5C*	10:58,100,890	C	2.26E-05	0.003691	MXE
rs133838482	*NEMP1*	5:56,366,065	C	3.85E-05	0.003762	SE
rs133135962	*DBNL*	22:387,756	T	5.06E-05	0.004496	SE
rs43725274	*ABTB1*	22:59,835,232	A	7.91E-05	0.006175	SE
rs41640891	*ITPR1*	22:21,678,074	C	8.00E-05	0.006228	SE
rs110291055	*NUGGC*	8:10,736,676	C	0.000152	0.008051	A3SS
rs109339682	*PIK3C2B*	16:2,195,385	A	0.000138	0.008948	SE
rs42705933	*PIK3C2B*	16:2,210,365	A	0.000138	0.008948	SE
rs110422856	*NCAPG*	6:37,374,718	G	0.00021	0.011918	SE
rs109978478	*ARAP1*	15:52,256,813	A	0.000439	0.020127	SE
rs41645253	*ZNF397*	24:21,783,988	G	0.000961	0.021919	A5SS
rs110464146	*POGZ*	3:19,350,945	G	0.000552	0.023303	SE
rs133847062	*PRPF8*	19:22,777,382	A	0.000597	0.024564	SE
rs384018186	*ERGIC1*	20:4,689,032	T	0.000762	0.028958	SE
rs43349825	*ERGIC1*	20:4,684,346	G	0.000762	0.028958	SE
rs110293454	*MGAM*	4:105,356,030	A	0.000923	0.032992	SE
rs110477374	*ETV6*	5:98,106,200	G	0.000949	0.033582	SE
rs135381754	*ZSCAN29*	21:55,118,729	A	0.00114	0.038199	SE
rs108957555	*TNRC18*	25:38,925,016	T	0.001053	0.047162	MXE
rs41930956	*ABCA6*	19:61,476,059	G	0.00169	0.049079	SE
rs109376798	*GART*	1:1,997,766	T	0.001692	0.049115	SE

Italicized values in Table 2 represent genes with AS SNPs.

TABLE 3 Genotypes of SNP rs42705933 and rs133847062 in two groups.

SNP	REF	ALT	HY1	HY2	HY3	SM1	SM2	SM3
rs42705933	A	G	0/0	0/0	0/0	0/1	0/1	0/1
rs133847062	A	C	0/1	0/1	0/1	0/0	0/0	0/0

Note: 0 represents the reference genomic loci and 1 represents the mutation loci.

ACSL5, *CYBB*, *VDAC3*, and *FTH1* were also DAS genes. Of these, the *SLC40A1* presented SE and RI events and the *FTH1* showed four alternative splicing patterns, including SE and RI, MXE, and A3SS. It is hypothesized that alternative splicing may influence the development of subclinical mastitis by altering the alternative spliced pattern of genes associated with ferroptosis.

RNA splicing is an important regulatory mechanism that links trait-associated variants and complex traits. Alternative splicing generates diverse transcripts with significant roles in disease resistance (Ju et al., 2012a; Asselstine et al., 2022) and the metabolic process (Sun et al., 2021) of the cow. In this study, a total of 2,212 differential splicing events were detected, affecting

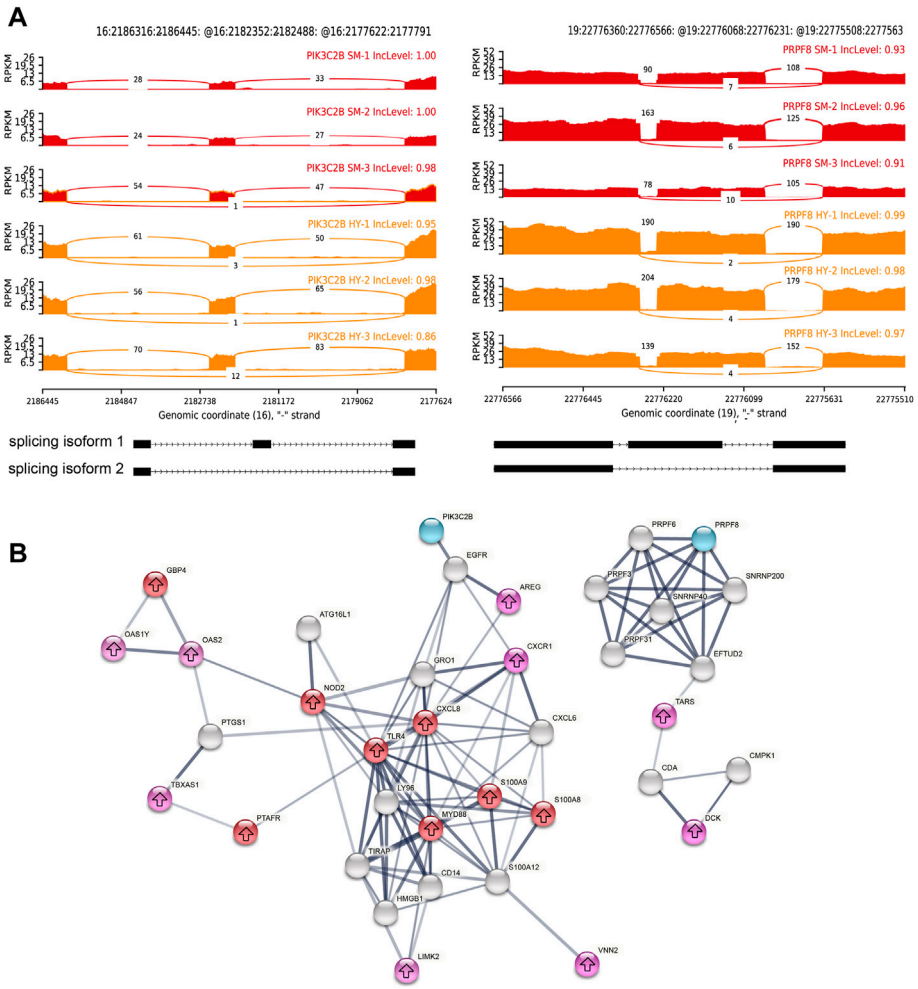


FIGURE 9
Interacting networks of alternative splicing genes, differentially expressed genes (DEGs), and the target genes of differentially expressed lncRNA (DELs). **(A)** differential alternative splicing (DAS) genes *PIK3C2B* and *PRPF8* exhibited skipping exon (SE) events compared to healthy cows. Red represents healthy cows (HY1, HY2, and HY3) and yellow represents subclinical mastitis cows (SM1, SM2, and SM3). Inclevel represents inclusion level for sample replicates, and gene structure graph represents alternative splicing isoforms imputed from GTF files. **(B)** PPI network of subclinical mastitis resistance-associated genes. Purple nodes represents DEGs, red nodes represents the target genes of lncRNAs that were close to somatic cell count (SCC) and somatic cell score (SCS), and white cell count QTLs, blue nodes represents two DAS genes regulated by SNPs, and gray nodes represents the second shell of interactors. The thickness of the line represents the strength of the interaction, and the arrow represents the up-regulated gene.

1,621 different genes, and GO enrichment analysis showed that they mainly affected some genes that are thought to be involved in transcriptional regulation, proteolytic processes, and neurodevelopment (Figure 8C). It was furthermore reported that the two groups exhibited different AS profiles, with a two-fold increase in the number of SE events in subclinical mastitis cows compared to healthy cows (Figure 7C), whereas there was no significant increase in the number of MXE events, A5SS events and A3SS events, or even a decrease in the number of RI events, demonstrating a link between specific alternative splicing patterns and subclinical mastitis in cows, and similar results were obtained in mammary gland tissues infected with *S. aureus* (Sun et al., 2021). Consequently, we hypothesized that mastitis-specific AS patterns may increase susceptibility to mastitis and that gene-specific AS patterns in the mammary gland of healthy cows may positively influence the disease resistance of cows, which requires further study.

Mutations affected by pre-mRNA splicing account for more than 15% of disease-causing mutations (Caminsky et al., 2014). Previous studies have documented numerous loci of disease susceptibility genes associated with certain specific splicing patterns in cows. For instance, SNPs

rs39652267 and rs39631044 in the 3′ flanking region of the gene *JAK2* were dramatically associated with SCC and SCS, and SNP rs43046497 on the intron nine of the gene *STAT5A* was significantly linked to *IL-6* (Usman et al., 2014). These results indicated that the alternative splicing process of mastitis resistance-associated genes is mediated by SNPs. In our study, a total of 32 AS SNPs were identified (Table 2), of which two SNPs rs42705933 and rs133847062 had completely different genotypes between healthy cows and subclinical mastitis cows, which corresponded to SE events of *PIK3C2B* and *PRPF8* (Table 3). Notably, neither of the candidate mastitis genes affected by splice variants in this study, *PIK3C2B* and *PRPF8* were simultaneously regulated at the transcriptional level, indicating independent regulation of genes by alternative splicing, while the PPI interaction network showed that these two alternative splice genes interacted with subclinical mastitis-related genes at the protein level (Figure 9B).

Among them, *PIK3C2B* belongs to class II of the PI3Ks family and performs a pivotal role in the control of membrane trafficking and intracellular signaling (Vanhaesebroeck et al., 2010; Margaria et al., 2019). The splicing factor *PRPF8* is critical for breast cell survival and has potential prognostic value in breast cancer (Cao et al., 2022). What is more, these two SNPs are variants of spliced polypyrimidine fragments (Supplementary Table S5), which are recognized by the polypyrimidine fragment binding protein of the spliceosome complex and are essential for the initial recognition of introns in mammals (Riolo et al., 2021), and the expansion of the polypyrimidine fragment enhances the efficiency of splicing (Romfo and Wise, 1997). Therefore, we hypothesized that the SNPs rs42705933 and rs133847062 may act as pathogenic mutations that cause the genes *PIK3C2B* and *PRPF8* to exhibit exon skipping, which in turn affects the resistance of cows to subclinical mastitis, but the detailed regulatory mechanism remains to be verified in subsequent future studies.

5 Conclusion

Our data suggest that genes *TLR4*, *NOD2*, *CXCL8*, and *OAS2* are key components involved in the host immune response to bovine mastitis and that the lncRNAs MSTRG25101.2, MSTRG.56327.1, and MSTRG.18968.1 adjacent to the SCC QTL and SCS QTL may potentially regulate expression of these candidate genes. In addition, characterization of peripheral white blood cell transcriptome alternative splicing events indicated that specific types of alternative splicing events appear to be associated with mastitis resistance in cows and that single nucleotide polymorphisms in the DAS genes *PIK3C2B* and *PPRF8* may be risk factors for mastitis susceptibility in cows. The genes and lncRNAs highlighted in this study could serve as expression biomarkers for mastitis that can be used for genetic improvement of dairy cattle for resilience to mastitis. In conclusion, our study provides a basis for further investigation of the molecular mechanisms of resistance and susceptibility to subclinical mastitis in cows.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: https://www.ncbi.nlm.nih.gov/, PRJNA839517.

Ethics statement

The animal study was reviewed and approved by the Animal Welfare Committee of China Agricultural University, Beijing, China.

Author contributions

YY and JY contributed to the conception and design of the study. YT, XL, and SM contributed to the data collection. JY, JZ, and XL analyzed and interpreted the data. JY contributed to the writing of the manuscript. YT, MK, and CW made major contributions to the revision and refinement of the article. YY supervised the research and provided funding. All authors contributed to the article and approved the submitted version. All authors contributed to manuscript revision, read, and approved the submitted version.

Funding

This article was financially supported by the National Key R&D Program of China (2021YFD1200900, 2021YFD1200903), NSFC-PSF Joint Project (31961143009), Beijing Dairy Industry Innovation Team (BAIC06), China Agriculture Research System of MOF and MARA, Beijing Natural Science Foundation (6182021), and the Program for Changjiang Scholar and Innovation Research Team in University (IRT-15R62).

Acknowledgments

The authors thank all members of the Laboratory of Animal Molecular and Quantitative Genetics (China Agricultural University) for sample collection work and Gerile Dari, Wenlong Li, Yibing Liu, and others for helpful discussions.

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.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.949850/full#supplementary-material

References

Aitken, S. L., Corl, C. M., and Sordillo, L. M. (2011). Immunopathology of mastitis: Insights into disease recognition and resolution. *J. Mammary Gland. Biol. Neoplasia* 16 (4), 291–304. doi:10.1007/s10911-011-9230-4

Anders, S., Pyl, P. T., and Huber, W. (2015). HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics* 31 (2), 166–169. doi:10.1093/bioinformatics/btu638

Asselstine, V., Medrano, J. F., and Canovas, A. (2022). Identification of novel alternative splicing associated with mastitis disease in Holstein dairy cows using large gap read mapping. *BMC Genomics* 23 (1), 222. doi:10.1186/s12864-022-08430-x

Bai, X., Zheng, Z., Liu, B., Ji, X., Bai, Y., and Zhang, W. (2016). Whole blood transcriptional profiling comparison between different milk yield of Chinese Holstein cows using RNA-seq data. *BMC Genomics* 17 (7), 512. doi:10.1186/s12864-016-2901-1

Baralle, F. E., and Giudice, J. (2017). Alternative splicing as a regulator of development and tissue identity. *Nat. Rev. Mol. Cell Biol.* 18 (7), 437–451. doi:10.1038/nrm.2017.27

Bhattarai, D., Worku, T., Dad, R., Rehman, Z. U., Gong, X., and Zhang, S. (2018). Mechanism of pattern recognition receptors (PRRs) and host pathogen interplay in bovine mastitis. *Microb. Pathog.* 120, 64–70. doi:10.1016/j.micpath.2018.04.010

Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* 30 (15), 2114–2120. doi:10.1093/bioinformatics/btu170

Caminsky, N., Mucaki, E. J., and Rogan, P. K. (2014). Interpretation of mRNA splicing mutations in genetic disease: Review of the literature and guidelines for information-theoretical analysis. *F1000Res.* 3, 282. doi:10.12688/f1000research.5654.1

Cao, D., Xue, J., Huang, G., An, J., and An, W. (2022). The role of splicing factor PRPF8 in breast cancer. *Technol. Health Care.* 30 (S1), 293–301. doi:10.3233/THC-THC228028

Celhar, T., Magalhaes, R., and Fairhurst, A. M. (2012). TLR7 and TLR9 in SLE: When sensing self goes wrong. *Immunol. Res.* 53 (1-3), 58–77. doi:10.1007/s12026-012-8270-1

Chacko, E., and Ranganathan, S. (2009). Genome-wide analysis of alternative splicing in cow: Implications in bovine as a model for human diseases. *BMC Genomics* 10, S11. Suppl 3. doi:10.1186/1471-2164-10-S3-S11

Chang, K., Yuan, C., and Liu, X. (2021). Ferroptosis-Related gene signature accurately predicts survival outcomes in patients with Clear-Cell renal cell carcinoma. *Front. Oncol.* 11, 649347. doi:10.3389/fonc.2021.649347

Chaussabel, D. (2015). Assessment of immune status using blood transcriptomics and potential implications for global health. *Semin. Immunol.* 27 (1), 58–66. doi:10.1016/j.smim.2015.03.002

Cheng, Z., Buggiotti, L., Salavati, M., Marchitelli, C., Palma-Vera, S., Wylie, A., et al. (2021). Global transcriptomic profiles of circulating leucocytes in early lactation cows with clinical or subclinical mastitis. *Mol. Biol. Rep.* 48 (5), 4611–4623. doi:10.1007/s11033-021-06494-8

Deng, S., Zheng, Y., Mo, Y., Xu, X., Li, Y., Zhang, Y., et al. (2021). Ferroptosis suppressive genes correlate with immunosuppression in glioblastoma. *World Neurosurg.* 152, e436–e448. doi:10.1016/j.wneu.2021.05.098

Dixon, S. J., Lemberg, K. M., Lamprecht, M. R., Skouta, R., Zaitsev, E. M., Gleason, C. E., et al. (2012). Ferroptosis: An iron-dependent form of nonapoptotic cell death. *Cell* 149 (5), 1060–1072. doi:10.1016/j.cell.2012.03.042

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., et al. (2013). Star: Ultrafast universal RNA-seq aligner. *Bioinformatics* 29 (1), 15–21. doi:10.1093/bioinformatics/bts635

Faustino, N. A., and Cooper, T. A. (2003). Pre-mRNA splicing and human disease. *Genes Dev.* 17 (4), 419–437. doi:10.1101/gad.1048803

Gao, J., Barkema, H. W., Zhang, L., Liu, G., Deng, Z., Cai, L., et al. (2017). Incidence of clinical mastitis and distribution of pathogens on large Chinese dairy farms. *J. Dairy Sci.* 100 (6), 4797–4806. doi:10.3168/jds.2016-12334

Gilbert, F. B., Cunha, P., Jensen, K., Glass, E. J., Foucras, G., Robert-Granie, C., et al. (2013). Differential response of bovine mammary epithelial cells to *Staphylococcus aureus* or *Escherichia coli* agonists of the innate immune system. *Vet. Res.* 44, 40. doi:10.1186/1297-9716-44-40

Goldammer, T., Zerbe, H., Molenaar, A., Schuberth, H. J., Brunner, R. M., Kata, S. R., et al. (2004). Mastitis increases mammary mRNA abundance of beta-defensin 5, toll-like-receptor 2 (TLR2), and TLR4 but not TLR9 in cattle. *Clin. Diagn. Lab. Immunol.* 11 (1), 174–185. doi:10.1128/cdli.11.1.174-185.2004

Goncalves, J. L., Cue, R. I., Botaro, B. G., Horst, J. A., Valloto, A. A., and Santos, M. V. (2018). Milk losses associated with somatic cell counts by parity and stage of lactation. *J. Dairy Sci.* 101 (5), 4357–4366. doi:10.3168/jds.2017-13286

Hoque, M. N., Istiaq, A., Clement, R. A., Gibson, K. M., Saha, O., Islam, O. K., et al. (2020). Insights into the resistome of bovine clinical mastitis microbiome, a key factor in disease complication. *Front. Microbiol.* 11, 860. doi:10.3389/fmicb.2020.00860

Hoque, M. N., Istiaq, A., Clement, R. A., Sultana, M., Crandall, K. A., Siddiki, A. Z., et al. (2019). Metagenomic deep sequencing reveals association of microbiome signature with functional biases in bovine mastitis. *Sci. Rep.* 9 (1), 13536. doi:10.1038/s41598-019-49468-4

Hou, Q., Huang, J., Ju, Z., Li, Q., Li, L., Wang, C., et al. (2012). Identification of splice variants targeted microRNAs and functional single nucleotide polymorphisms of the BOLA-DQA2 gene in dairy cattle. *DNA Cell Biol.* 31 (5), 739–744. doi:10.1089/dna.2011.1402

Islam, M. A., Takagi, M., Fukuyama, K., Komatsu, R., Albarracin, L., Nochi, T., et al. (2020). Transcriptome analysis of the inflammatory responses of bovine mammary epithelial cells: Exploring immunomodulatory target genes for bovine mastitis. *Pathogens* 9 (3), E200. doi:10.3390/pathogens9030200

Ju, Z., Wang, C., Li, Q., Hou, M., Gao, S., Hou, Q., et al. (2012a). Alternative splicing and mRNA expression analysis of bovine SLAMF7 gene in healthy and mastitis mammary tissues. *Mol. Biol. Rep.* 39 (4), 4155–4161. doi:10.1007/s11033-011-1198-z

Ju, Z., Wang, C., Wang, X., Yang, C., Sun, Y., Jiang, Q., et al. (2015). Role of an SNP in alternative splicing of bovine NCF4 and mastitis susceptibility. *PLoS One* 10 (11), e0143705. doi:10.1371/journal.pone.0143705

Kang, Y. J., Yang, D. C., Kong, L., Hou, M., Meng, Y. Q., Wei, L., et al. (2017). CPC2: A fast and accurate coding potential calculator based on sequence intrinsic features. *Nucleic Acids Res.* 45 (W1), W12–W16. doi:10.1093/nar/gkx428

Kawai, T., and Akira, S. (2010). The role of pattern-recognition receptors in innate immunity: Update on Toll-like receptors. *Nat. Immunol.* 11 (5), 373–384. doi:10.1038/ni.1863

Kawasaki, T., and Kawai, T. (2014). Toll-like receptor signaling pathways. *Front. Immunol.* 5, 461. doi:10.3389/fimmu.2014.00461

Leitner, G., Yadlin, B., Glickman, A., Chaffer, M., and Saran, A. (2000). Systemic and local immune response of cows to intramammary infection with *Staphylococcus aureus*. *Res. Vet. Sci.* 69 (2), 181–184. doi:10.1053/rvsc.2000.0409

Li, A., Zhang, J., and Zhou, Z. (2014). Plek: A tool for predicting long non-coding RNAs and messenger RNAs based on an improved k-mer scheme. *BMC Bioinforma.* 15, 311. doi:10.1186/1471-2105-15-311

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics* 25 (16), 2078–2079. doi:10.1093/bioinformatics/btp352

Li, H., Huang, J., Zhang, J., Gao, Y., Han, B., and Sun, D. (2022). Identification of alternative splicing events associated with paratuberculosis in dairy cattle using Multi-Tissue RNA sequencing data. *Genes (Basel)* 13 (3), 497. doi:10.3390/genes13030497

Li, L., Huang, J., Ju, Z., Li, Q., Wang, C., Qi, C., et al. (2013). Multiple promoters and targeted microRNAs direct the expressions of HMGB3 gene transcripts in dairy cattle. *Anim. Genet.* 44 (3), 241–250. doi:10.1111/age.12007

Lin, C., Zhu, Y., Hao, Z., Xu, H., Li, T., Yang, J., et al. (2021). Genome-Wide analysis of LncRNA in bovine mammary epithelial cell injuries induced by *Escherichia coli* and *Staphylococcus aureus*. *Int. J. Mol. Sci.* 22 (18), 9719. doi:10.3390/ijms22189719

Liu, J., Ren, Z., Yang, L., Zhu, L., Li, Y., Bie, C., et al. (2022). The NSUN5-FTH1/FTL pathway mediates ferroptosis in bone marrow-derived mesenchymal stem cells. *Cell Death Discov.* 8 (1), 99. doi:10.1038/s41420-022-00902-z

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 (12), 550. doi:10.1186/s13059-014-0550-8

Ma, M., Pei, Y., Wang, X., Feng, J., Zhang, Y., and Gao, M. Q. (2019a). LncRNA XIST mediates bovine mammary epithelial cell inflammatory response via NF-κB/NLRP3 inflammasome pathway. *Cell Prolif.* 52 (1), e12525. doi:10.1111/cpr.12525

Margaria, J. P., Ratto, E., Gozzelino, L., Li, H., and Hirsch, E. (2019). Class II PI3Ks at the intersection between signal transduction and membrane trafficking. *Biomolecules* 9 (3), E104. doi:10.3390/biom9030104

Mi, S., Tang, Y., Dari, G., Shi, Y., Zhang, J., Zhang, H., et al. (2021). Transcriptome sequencing analysis for the identification of stable lncRNAs associated with bovine *Staphylococcus aureus* mastitis. *J. Anim. Sci. Biotechnol.* 12 (1), 120. doi:10.1186/s40104-021-00639-2

Modrek, B., Resch, A., Grasso, C., and Lee, C. (2001). Genome-wide detection of alternative splicing in expressed sequences of human genes. *Nucleic Acids Res.* 29 (13), 2850–2859. doi:10.1093/nar/29.13.2850

Pertea, M., Pertea, G. M., Antonescu, C. M., Chang, T. C., Mendell, J. T., and Salzberg, S. L. (2015). StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat. Biotechnol.* 33 (3), 290–295. doi:10.1038/nbt.3122

Ping, S., Wang, S., Zhao, Y., He, J., Li, G., Li, D., et al. (2022). Identification and validation of a ferroptosis-related gene signature for predicting survival in skin cutaneous melanoma. *Cancer Med.* doi:10.1002/cam4.4706

Porcherie, A., Cunha, P., Trotereau, A., Roussel, P., Gilbert, F. B., Rainard, P., et al. (2012). Repertoire of *Escherichia coli* agonists sensed by innate immunity receptors of the bovine udder and mammary epithelial cells. *Vet. Res.* 43, 14. doi:10.1186/1297-9716-43-14

Quinlan, A. R., and Hall, I. M. (2010). BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics* 26 (6), 841–842. doi:10.1093/bioinformatics/btq033

Ren, P., Lu, L., Cai, S., Chen, J., Lin, W., and Han, F. (2021). Alternative splicing: A new cause and potential therapeutic target in autoimmune disease. *Front. Immunol.* 12, 713540. doi:10.3389/fimmu.2021.713540

Riolo, G., Cantara, S., and Ricci, C. (2021). What's wrong in a jump? Prediction and validation of splice site variants. *Methods Protoc.* 4 (3), 62. doi:10.3390/mps4030062

Romfo, C. M., and Wise, J. A. (1997). Both the polypyrimidine tract and the 3' splice site function prior to the first step of splicing in fission yeast. *Nucleic Acids Res.* 25 (22), 4658–4665. doi:10.1093/nar/25.22.4658

Schepers, A. J., Lam, T. J., Schukken, Y. H., Wilmink, J. B., and Hanekamp, W. J. (1997). Estimation of variance components for somatic cell counts to determine thresholds for uninfected quarters. *J. Dairy Sci.* 80 (8), 1833–1840. doi:10.3168/jds.S0022-0302(97)76118-6

Shen, S., Park, J. W., Lu, Z. X., Lin, L., Henry, M. D., Wu, Y. N., et al. (2014). Rmats: Robust and flexible detection of differential alternative splicing from replicate RNA-Seq data. *Proc. Natl. Acad. Sci. U. S. A.* 111 (51), E5593–E5601. doi:10.1073/pnas.1419161111

Sun, H. Z., Zhu, Z., Zhou, M., Wang, J., Dugan, M., and Guan, L. L. (2021). Gene co-expression and alternative splicing analysis of key metabolic tissues to unravel the regulatory signatures of fatty acid composition in cattle. *RNA Biol.* 18 (6), 854–862. doi:10.1080/15476286.2020.1824060

Sun, L., Luo, H., Bu, D., Zhao, G., Yu, K., Zhang, C., et al. (2013). Utilizing sequence intrinsic composition to classify protein-coding and long non-coding transcripts. *Nucleic Acids Res.* 41 (17), e166. doi:10.1093/nar/gkt646

Urbanski, L. M., Leclair, N., and Anczukow, O. (2018). Alternative-splicing defects in cancer: Splicing regulators and their downstream targets, guiding the way to novel cancer therapeutics. *Wiley Interdiscip. Rev. RNA* 9 (4), e1476. doi:10.1002/wrna.1476

Usman, T., Yu, Y., Liu, C., Wang, X., Zhang, Q., and Wang, Y. (2014). Genetic effects of single nucleotide polymorphisms in JAK2 and STAT5A genes on susceptibility of Chinese Holsteins to mastitis. *Mol. Biol. Rep.* 41 (12), 8293–8301. doi:10.1007/s11033-014-3730-4

Vanhaesebroeck, B., Guillermet-Guibert, J., Graupera, M., and Bilanges, B. (2010). The emerging mechanisms of isoform-specific PI3K signalling. *Nat. Rev. Mol. Cell Biol.* 11 (5), 329–341. doi:10.1038/nrm2882

Wang, D., Wei, G., Ma, J., Cheng, S., Jia, L., Song, X., et al. (2021). Identification of the prognostic value of ferroptosis-related gene signature in breast cancer patients. *BMC Cancer* 21 (1), 645. doi:10.1186/s12885-021-08341-2

Wang, G., Yin, H., Li, B., Yu, C., Wang, F., Xu, X., et al. (2019a). Characterization and identification of long non-coding RNAs based on feature relationship. *Bioinformatics* 35 (17), 2949–2956. doi:10.1093/bioinformatics/btz008

Wang, H., Wang, X., Li, X., Wang, Q., Qing, S., Zhang, Y., et al. (2019b). A novel long non-coding RNA regulates the immune response in MAC-T cells and contributes to bovine mastitis. *FEBS J.* 286 (9), 1780–1795. doi:10.1111/febs.14783

Wang, W., Lin, X., Jiang, T., Peng, Z., Xu, J., Yi, L., et al. (2018). Prevalence and characterization of *Staphylococcus aureus* cultured from raw milk taken from dairy cows with mastitis in beijing, China. *Front. Microbiol.* 9, 1123. doi:10.3389/fmicb.2018.01123

Wang, X. G., Ju, Z. H., Hou, M. H., Jiang, Q., Yang, C. H., Zhang, Y., et al. (2016a). Deciphering transcriptome and complex alternative splicing transcripts in mammary gland tissues from cows naturally infected with *Staphylococcus aureus* mastitis. *PLoS One* 11 (7), e0159719. doi:10.1371/journal.pone.0159719

Wang, X., Su, F., Yu, X., Geng, N., Li, L., Wang, R., et al. (2020a). RNA-Seq whole transcriptome analysis of bovine mammary epithelial cells in response to intracellular *Staphylococcus aureus*. *Front. Vet. Sci.* 7, 642. doi:10.3389/fvets.2020.00642

Wang, X., Wang, H., Zhang, R., Li, D., and Gao, M. Q. (2020b). LRRC75A antisense lncRNA1 knockout attenuates inflammatory responses of bovine mammary epithelial cells. *Int. J. Biol. Sci.* 16 (2), 251–263. doi:10.7150/ijbs.38214

Wang, X., Zhong, J., Gao, Y., Ju, Z., and Huang, J. (2014). A SNP in intron 8 of CD46 causes a novel transcript associated with mastitis in Holsteins. *BMC Genomics* 15, 630. doi:10.1186/1471-2164-15-630

Weijiao, Y., Fuchun, L., Mengjie, C., Xiaoqing, Q., Hao, L., Yuan, L., et al. (2021). Immune infiltration and a ferroptosis-associated gene signature for predicting the prognosis of patients with endometrial cancer. *Aging (Albany NY)* 13 (12), 16713–16732. doi:10.18632/aging.203190

Wellnitz, O., and Bruckmaier, R. M. (2012). The innate immune response of the bovine mammary gland to bacterial infection. *Vet. J.* 192 (2), 148–152. doi:10.1016/j.tvjl.2011.09.013

Xu, Z., Jiang, S., Ma, J., Tang, D., Yan, C., and Fang, K. (2021). Comprehensive analysis of Ferroptosis-Related LncRNAs in breast cancer patients reveals prognostic value and relationship with tumor immune microenvironment. *Front. Surg.* 8, 742360. doi:10.3389/fsurg.2021.742360

Yang, Q., Zhao, J., Zhang, W., Chen, D., and Wang, Y. (2019). Aberrant alternative splicing in breast cancer. *J. Mol. Cell Biol.* 11 (10), 920–929. doi:10.1093/jmcb/mjz033

Yang, W. S., and Stockwell, B. R. (2016). Ferroptosis: Death by lipid peroxidation. *Trends Cell Biol.* 26 (3), 165–176. doi:10.1016/j.tcb.2015.10.014

Yang, Y., Luo, M., Zhang, K., Zhang, J., Gao, T., Connell, D. O., et al. (2020). Nedd4 ubiquitylates VDAC2/3 to suppress erastin-induced ferroptosis in melanoma. *Nat. Commun.* 11 (1), 433. doi:10.1038/s41467-020-14324-x

Yin, J., Zeng, X., Ai, Z., Yu, M., Wu, Y., and Li, S. (2020). Construction and analysis of a lncRNA-miRNA-mRNA network based on competitive endogenous RNA reveal functional lncRNAs in oral cancer. *BMC Med. Genomics* 13 (1), 84. doi:10.1186/s12920-020-00741-w

Zhang, K., Ping, L., Du, T., Liang, G., Huang, Y., Li, Z., et al. (2021). A Ferroptosis-Related lncRNAs signature predicts prognosis and immune microenvironment for breast cancer. *Front. Mol. Biosci.* 8, 678877. doi:10.3389/fmolb.2021.678877

Zhang, Z., Wang, X., Li, R., Ju, Z., Qi, C., Zhang, Y., et al. (2015). Genetic mutations potentially cause two novel NCF1 splice variants up-regulated in the mammary gland, blood and neutrophil of cows infected by *Escherichia coli*. *Microbiol. Res.* 174, 24–32. doi:10.1016/j.micres.2015.03.005

Zhu, J. H., De Mello, R. A., Yan, Q. L., Wang, J. W., Chen, Y., Ye, Q. H., et al. (2020). MiR-139-5p/SLC7A11 inhibits the proliferation, invasion and metastasis of pancreatic carcinoma via PI3K/Akt signaling pathway. *Biochim. Biophys. Acta. Mol. Basis Dis.* 1866 (6), 165747. doi:10.1016/j.bbadis.2020.165747

Zhu, T., Liu, B., Wu, D., Xu, G., and Fan, Y. (2021). Autophagy regulates VDAC3 ubiquitination by FBXW7 to promote Erastin-Induced ferroptosis in acute lymphoblastic leukemia. *Front. Cell Dev. Biol.* 9, 740884. doi:10.3389/fcell.2021.740884

Zuo, Y. B., Zhang, Y. F., Zhang, R., Tian, J. W., Lv, X. B., Li, R., et al. (2022). Ferroptosis in cancer progression: Role of noncoding RNAs. *Int. J. Biol. Sci.* 18 (5), 1829–1843. doi:10.7150/ijbs.66917

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 26 September 2022
DOI 10.3389/fgene.2022.970097

OPEN ACCESS

EDITED BY
Xiao Wang,
Konge Larsen ApS, Denmark

REVIEWED BY
Md Zubbair Malik,
Jawaharlal Nehru University, India
Hui Li,
Guangxi University, China
Yahui Gao,
University of Maryland, United States

*CORRESPONDENCE
Haiying Li,
lhy-3@163.com

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

RECEIVED 15 June 2022
ACCEPTED 07 September 2022
PUBLISHED 26 September 2022

CITATION
Wu Y, Li H, Zhao X, Baki G, Ma C, Yao Y, Li J, Yao Y and Wang L (2022), Differential expression of circRNAs of testes with high and low sperm motility in Yili geese.
Front. Genet. 13:970097.
doi: 10.3389/fgene.2022.970097

Differential expression of circRNAs of testes with high and low sperm motility in Yili geese

Yingping Wu, Haiying Li*, Xiaoyu Zhao, Gulnar Baki, Chen Ma, Yingying Yao, Jiahui Li, Yang Yao and Lin Wang

College of Animal Science, Xinjiang Agricultural University, Urumqi, China

The aim of this study was to explore the potential biological function of circular RNAs (circRNAs) in the sperm motility traits of Xinjiang Yili geese, and to provide a reference for analyzing the mechanism of regulation of Yili geese sperm motility. The 10 selected Xinjiang Yili Geese with high or low sperm motility (five for each group) were 3 years old, in good health, and were kept in the same feeding conditions. Yili geese were slaughtered for the collection of testicular tissue and high-throughput sequencing technology was used to screen differentially expressed circRNAs for bioinformatics analysis. Combined with the previously screened miRNAs related to the sperm motility of Yili geese, the circRNAs miRNAs regulatory network was constructed. The results showed that a total of 26,311 circRNAs were obtained from testicular tissues with high and low sperm motility, and 173 DECs were screened between the two groups ($p < 0.05$, |log2Foldchange|>0), of which 82 were up-regulated and 91 were down-regulated. Functional analysis of the source genes of these DECs showed that the source genes were mainly involved in biological processes. KEGG enrichment analysis showed that the source genes of DECs were mainly enriched in autophagy-animal, ubiquinone and other terpenoid-quinone biosynthesis, progesterone-mediated oocyte maturation, regulation of the actin cytoskeleton and other pathways. Furthermore, the visual regulatory network of differential circRNA-miRNA-mRNA was constructed, including 20 circRNAs, 18 miRNAs and 177 mRNAs, and nine core regulatory circRNAs were screened, including novell_circ_0045314, novel_circ_0019994 and novel_circ_0020422, etc., targeting ppy-mir-16, hsa-mir-221–3p, gga-mir-499–5p, etc. The results suggest that circRNAs may interact with miRNAs to further regulate mRNA to regulate sperm motility in Yili geese, so as to provide a reference for analyzing the molecular mechanism of sperm motility regulation.

KEYWORDS

yili geese, sperm motility, testis, circrnas, transcriptomic

1 Introduction

The Yili goose is the only local poultry variety in China that is derived from the herbivorous characteristics of gray geese. As a high-quality local characteristic poultry, the Yili goose has the characteristics of strong adaptability, good flight, heat resistance, cold resistance, rough feeding resistance, strong disease resistance and stress resistance, as well

as good meat quality (Zhao et al., 2020). Reproductive traits are the most important economic traits in poultry production. Improving reproductive performance is always an important goal of poultry genetic improvement. The semen quality of male birds is directly related to the fertilization ability of sperm, which is one of the important indicators affecting reproductive efficiency. According to previous research, the low fertility of male Yili geese and the low fertilization rate of eggs limit the development of the Yili goose industry (Zhao et al., 2019). We also found that the sperm motility of yili geese was significantly different among individuals, and there was a positive correlation between sperm motility and fertilization rate (Wu et al., 2022). Therefore, there is an urgent need to use molecular biotechnology to improve the reproductive performance of male Yili geese and to explore the genetic mechanisms affecting their semen quality.

CircRNAs are closed circular RNA transcripts formed by reverse splicing from a single RNA precursor and are found in all higher eukaryotes, including mammals (Li et al., 2018). It has been found that circRNA plays an important role in organisms. It can be used as competitive endogenous RNA (CeRNA) or as an "miRNA sponge" to inhibit the activity of miRNA, weaken or even relieve the inhibitory effect of miRNA on downstream target genes, and promote the expression of target genes (Zhong et al., 2018). At present, there are many studies on the roles of circRNAs in various aspects of the mammalian reproductive system, such as testicular development (Li et al., 2021; Zhang et al., 2021), follicle development (Guo et al., 2020; Liang et al., 2020), embryonic development (Quan and Li, 2018), etc. However, studies on the molecular mechanisms of circRNAs in the poultry reproductive system have mainly focused on the follicle development of female poultry (Shen M. et al., 2019; Shen M. et al., 2020; Wu et al., 2020). Research on the mechanism of regulation of circRNA with regard to sperm motility in the Yili goose is still lacking. In this study, Yili geese with extreme differences in sperm motility were taken as the objects of research. The expression pattern of circRNA in the testicular tissue of Yili geese with high and low sperm motility was analyzed by RNA-seq technology, and the circRNA related to the sperm motility of male geese was screened and identified, so as to provide a theoretical basis for the genetic and breeding improvement of Yili geese.

2 Materials and methods

2.1 Sample collection

The experimental animals used in this study were provided by Hengxin Industrial Co., Ltd. Emin County, Xinjiang. According to the data on sperm motility and egg fertilization rate, among the 3-year-old Xinjiang Yili goose with similar body weight (3.65±0.40 kg) and the same feeding conditions, 5 geese with high and low sperm motility were selected ($p < 0.01$) (Table 1). After the geese were sacrificed, their testis tissues were collected immediately, and stored in liquid nitrogen, rinsed with PBS 1 to 2 times, and immediately placed in a cryopreservation tube containing RNA preservation solution. The testis samples of the HFR geese (high sperm motility group) were labeled as HFR-1~HFR-5, and the testis samples of the LFR geese (low sperm motility group) were labeled as LFR-1~LFR-5. The samples were labeled and stored at 4 °C overnight and then stored at -80 °C the next day until they were used for the extraction of total RNA.

2.2 Total RNA extraction and illumina sequencing

The total RNA was extracted using Trizol (Invitrogen, Carlsbad, CA, United States), following the manufacturer's protocol, and agarose gels, a Nanodrop ND-1000 spectrophotometer (IMPLEN,CA, United States) and an Agilent 2100 Bioanalyzer (Agilent Technologies, CA, United States) were used to ensure that the quality of the samples was sufficient for transcriptome sequencing. After the samples passed quality control, the strand-specific library was constructed by removing ribosomal RNA (circRNA library building and the linear RNA process). After the libraries were qualified, the library preparations were sequenced on an Illumina Hiseq platform and 150 bp paired-end reads were generated.

2.3 circRNAs identification

Raw data (raw reads) in fastq format were firstly processed through in-house perl scripts. In this step, clean data (clean reads) were obtained by removing reads containing adapter, reads containing ploy-N and low-quality reads from raw data. At the same time, for Q20, Q30 and GC content the clean data were calculated. The index of the reference genome (Yan et al., 2020) was built using Bowtie2 (v2.2.8) (Langmead et al., 2009) and paired-end clean reads were aligned to the reference genome using Bowtie. The circRNA was detected and identified using find_circ (v1.2) (Memczak et al., 2013)and CIRI2 (v2.0.5) (Gao et al., 2015), and the intersection of the two types of software was also used to identify the circRNA.

2.4 circRNAs difference analysis and functional enrichment analysis

HTSeq v0.6.0 (Anders et al., 2015)was used to count the reads numbers mapped to each gene, and the expression amount was normalized with TPM (Zhou et al., 2010). The

TABLE 1 Sperm motility and fertilization rate of Yili geese with high and low sperm motility.

Item	HFR-1 (%)	HFR-2 (%)	HFR-3 (%)	HFR-4 (%)	HFR-5 (%)	Mean ± SD	LFR-1 (%)	LFR-2 (%)	LFR-3 (%)	LFR-4 (%)	LFR-5 (%)	Mean ± SD
Sperm motility (%)	77.31	70.73	63.27	58.40	58.42	65.63%±8.25%	31.67	37.48	32.46	39.49	38.09	35.84%±3.53%
Fertilization rate (%)	98.33	97.50	94.17	93.24	91.77	95.00%±2.81%	40.99	43.45	43.50	49.36	46.67	44.79%±3.25%

input data of circRNA differential expression is readcounts data obtained from circrna expression level analysis. Differential expression analysis of the two groups was performed using the DESeq2 R package (1.10.1) (Michael et al., 2014). circRNAs with a $p < 0.05$ and |log2Foldchange|>0 found by DESeq2 were set as the threshold for significantly differential expression. The GOseq (Release 2.1.2) (Young et al., 2010) software was used for the GO enrichment analysis of differentially expressed circRNA-derived genes, and KOBAS (2.0) (Mao et al., 2005; Kanehisa et al., 2008) was used for Paththe way enrichment analysis. A *P* of <0.05 was set as the threshold for statistically significant results.

miRNA target sites in the exons of circRNA loci were identified using miRanda (Enright et al., 2003). miRNA target genes were predicted to be the intersection of miRanda (Enright et al., 2003) and RNAhybrid (Kruger and Rehmsmeier, 2006). Cytoscape (v3.7.1) (Paul et al., 2003) software was used to construct the circRNA-miRNA-mRNA networks.

2.5 Real-time PCR validation of sequencing results

Randomly selected from the transcriptome sequencing results, 8 DECs related to the HFR and LFR of Xinjiang Yili geese were used for fluorescence-based quantitative validation. Oligo 7.0 software was used to design primers (Supplementary Table S1). SYBR GREEN reagent (TaKaRa) was used to amplify the target gene and internal reference gene (beta-actin) mRNAs on a ROCHE 480 quantitative PCR instrument (Eppendorf, Germany). The PCR reaction system (20 μl) included the 10 μl AceQ Universal SYBR qPCR Master Mix, 0.4 μl upstream primer (10 μmol L/L), 0.4 μl downstream primer (10 μmol L/L), 6.7 μl ddH_2O, and 2.5 μl cDNA. The reaction conditions were as follows: 95°C for 5 min, 95°C for 10 s, 60°C for 30 s, 40 cycles; dissolution curve: 60°C→95°C, with a temperature increase of 0.3°C every 15 s. Quantitative expression results were calculated according to the cross point (CP) values, and the relative expression levels were calculated according to the $2^{-\Delta\Delta Ct}$ method (Bustin et al., 2009).

3 Results

3.1 Screening of yili geese with high and low sperm motility

Based on the semen quality data from six instances and the fertilization rate data from five instances, five individuals in the high sperm motility group and five individuals in the low sperm motility group were screened out. An independent t-test showed that the sperm motility and fertilization rate of the individuals in the high sperm motility group were significantly higher than in the low sperm motility group ($p < 0.01$, Table 1).

3.2 Evaluation of sequencing data quality

It can be seen from Table 2 that 495,401,294 clean reads and 475,402,620 reads were obtained by sequencing Yili geese with high and low sperm motility, respectively. The Q20 and Q30 of each sample were at least 97.49% and 93.06%, respectively, and the GC contents were between 45.40% and 47.40%. These results indicated that the quality of transcriptome sequencing results met the needs of subsequent analyses. According to the statistical results, 91.29%–92.37% of the clean reads were mapped to the reference genome, of which 86.06%–89.40% of the clean reads were uniquely mapped. The results showed that there were enough reads of each sample mapped to the reference genome, and the selected reference genome was suitable.

3.3 circRNAs identification and characteristic analysis

3.3.1 Identification of circRNAs

A total of 26,311 circRNAs were identified in testicular tissues (Figure 1), of which 6,520 and 6,032 were unique of the Xinjiang Yili geese with HFR and LFR, respectively. According to the annotation of the reference genome, it was found that many of the host genes of circRNAs are related to male reproduction (Table 3, Supplementary Table S2), such as spermatogenesis-related *SPA5L*, *SPAT5*, *SPAT6*, *SPT17*, *SPT48*, etc., testis-specific genes *TESK1*, *TEX10*, *TSG10* etc., and flagellar motility genes *SPEF2*, *CF206*, *IF172*, etc., showing that it is

TABLE 2 Evaluation of sequencing data quality.

Sample name	Raw reads	Clean reads	Q20 (%)	Q30 (%)	GC content (%)	Total mapped	Uniquely mapped
HFR_1	86,345,158	85,695,512	97.65	93.43	47.19	78,316,871 (91.39%)	73,750,086 (86.06%)
HFR_2	105,769,898	104,987,610	97.63	93.38	46.50	96,332,847 (91.76%)	92,098,738 (87.72%)
HFR_3	93,937,816	93,265,686	97.49	93.06	46.53	85,686,995 (91.87%)	82,404,051 (88.35%)
HFR_4	105,232,800	104,604,358	97.63	93.39	46.81	96,203,761 (91.97%)	92,639,975 (88.56%)
HFR_5	107,655,208	106,848,128	97.61	93.36	47.22	98,144,265 (91.85%)	94,536,194 (88.48%)
LFR_1	104,869,524	103,850,998	97.60	93.37	47.40	95,013,805 (91.49%)	90,797,347 (87.43%)
LFR_2	82,214,726	81,594,444	97.68	93.48	46.60	74,489,752 (91.29%)	71,049,348 (87.08%)
LFR_3	106,378,240	105,732,972	97.60	93.28	45.40	97,160,534 (91.89%)	94,049,495 (88.95%)
LFR_4	84,278,058	83,586,102	97.53	93.11	45.80	77,212,079 (92.37%)	74,728,549 (89.40%)
LFR_5	101,369,790	100,638,104	97.60	93.28	46.65	92,182,939 (91.60%)	87,737,183 (87.18%)

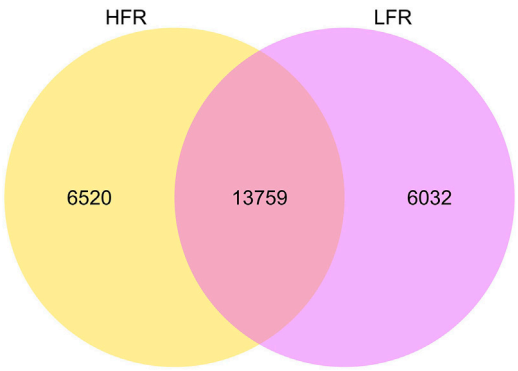


FIGURE 1
circRNAs difference between HFR and LFR. The big circle represents each comparison combination, the sum of the numbers in each big circle represents the total number of circrnas identified by the comparison combination, and the overlapping part of the circle represents the number of circrnas common between the combinations.

feasible to sequence testicular tissue by RNA-seq technology and to screen for biomarkers related to geese reproduction.

3.3.2 Characteristic analysis of circRNAs

According to the comparison results of circRNAs and the reference genome, there were 14,232 circRNAs without corresponding host genes, and other circRNAs were mapped to 12,080 genes. The lengths of these circRNAs ranged from 23 bp to 1,511 bp, with an average length of 357 bp (Figure 2). Most circRNAs were distributed on chromosome 2, accounting for 8.43%, followed by chromosomes 1, 7, and 10, accounting for 6.38%, 5.83%, and 4.87%, respectively (Figure 3). According to their location in the genome, most circRNAs were derived from intergenic regions (34.19%), followed by exonic regions (33.45%) and intronic regions (32.36%) (Figure 4).

3.4 circRNAs differential expression analysis

Compared with the low sperm motility group, a total of 173 DECs (differentially expressed circRNAs) were screened in the high sperm motility group, of which 82 were up-regulated and 91 were down-regulated ($p < 0.05$, |log2Foldchange|>0) (Figure 5), and 77 circRNAs had the source genes. Among them, the circRNAs with the smallest p values were novel_circ_0013769 and novel_circ_0007998, both located in the intergenic region, and both being up-regulated circRNAs. The fold difference was the largest in novel_circ_0042868 (source gene *PDE6C*), novel_circ_0018689 (intergenic_region) and novel_circ_0030568 (source gene *ADAM9*), with a fold difference of 7.6687, 7.1261 and 7.0849, respectively. There were 87 circRNAs with more than a 4-fold difference, accounting for 50.29% of the DECs. Hierarchical clustering analysis was performed on the circRNAs that were significantly differentially expressed in the testes of the two groups, as shown in the heat map (Figure 6), red color and blue color respectively represent the significant increase or decrease in the HFR group compared with the LFR group. It can be seen that the repeatability within the sample group was good, and the difference between the groups was large, which can better reflect the differences between different treatments.

3.5 Functional analysis of circRNA host genes

The functional analysis of the source genes of these DECs (Figure 7; Table 4) showed that 58 significant GO terms were enriched ($p < 0.05$), and biological processes (63.79%) were mainly enriched in cellular component organization, multi-organism processes, and the reproduction of a single-celled organism. The molecular function (31.03%) was mainly enriched in adenyl nucleotide binding, cytoskeletal protein binding, actin binding, etc., and the cellular component (5.17%) was mainly enriched in

TABLE 3 Annotation of host genes associated with gander reproduction for partial circRNAs.

circRNAs	Host genes	Description	position
novel_circ_0013754	*SPA5L*	BOVIN Spermatogenesis-associated protein 5-like protein 1	chr19:22147438–22148362
novel_circ_0008443	*SPAT5*	MOUSE Spermatogenesis-associated protein 5	chr15:12275327–12294121
novel_circ_0007364	*SPAT6*	HUMAN Spermatogenesis-associated protein 6	chr14:23958879–23965428
novel_circ_0000297	*SPT17*	HUMAN Spermatogenesis-associated protein 17	chr10:17374912–17389052
novel_circ_0041302	*SPT48*	HUMAN Spermatogenesis-associated protein 48	chr8:26482947–26486091
novel_circ_0004385	*TESK1*	HUMAN Dual specificity testis-specific protein kinase 1	chr12:25506529–25508863
novel_circ_0041026	*TEX10*	CHICK Testis-expressed protein 10 homolog	chr8:19326415–19339250
novel_circ_0033683	*TSG10*	HUMAN Testis-specific gene 10 protein	chr4:37691305–37693198
novel_circ_0004352	*SPEF2*	RAT Sperm flagellar protein 2	chr12:23705501–23734451
novel_circ_0014569	*CF206*	MACFA Cilia- and flagella-associated protein 206	chr1:20634161–20635541
novel_circ_0014350	*IF172*	MOUSE Intraflagellar transport protein 172 homolog	chr1:141083–142558

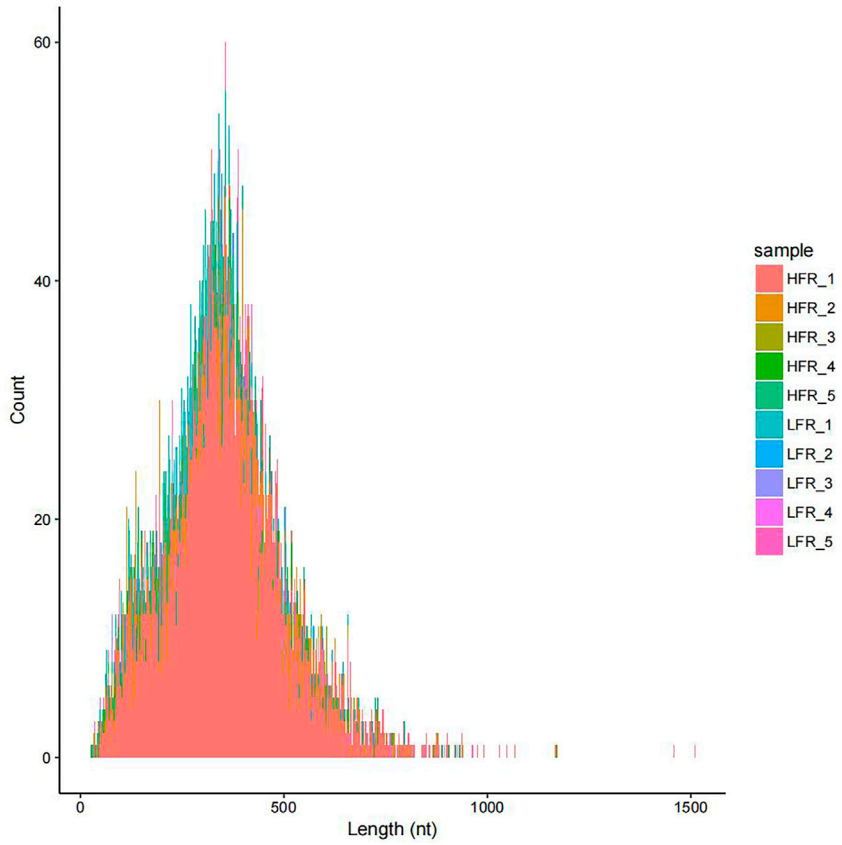


FIGURE 2
Length distribution of circRNAs.

the microtubule organizing center, integral to the Golgi membrane. KEGG analysis was mainly enriched in autophagy-animal, ubiquinone and other terpenoid-quinone biosynthesis, progesterone-mediated oocyte maturation, oocyte meiosis and purine metabolism and other signaling pathways, of which metabolic pathways were enriched. The pathways with the most

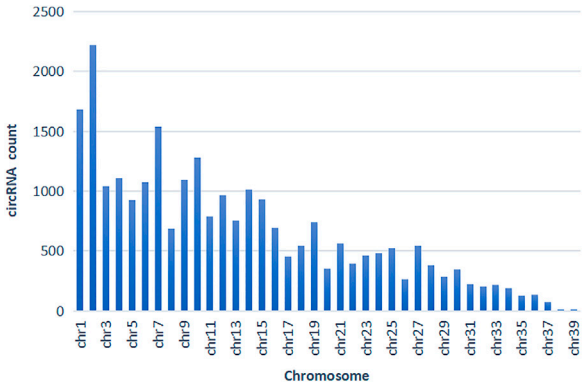


FIGURE 3
circRNA chromosome distribution.

in these two pathways were *Pdpk1* (phosphoinositide-dependent protein kinase 1) and *COQ5* (methoxy-6-polypropylene-1,4-benzoquinone methylase) (Figure 8;Table 5).

3.6 Construction and analysis of circRNA-miRNA-mRNA interaction network

circRNAs can adsorb miRNAs by combining with them, and can act as miRNA sponges (Li et al., 2018). Therefore, miRNA binding site analysis on the identified circRNAs is helpful to further study the function of circRNAs. In this study, the differentially expressed miRNAs in the testis of Yili geese with high and low sperm motility obtained in the previous stage were combined to construct a differential circRNA-miRNA-mRNA interaction network. As shown in Figure 9, the network was enriched with 20 circRNAs,

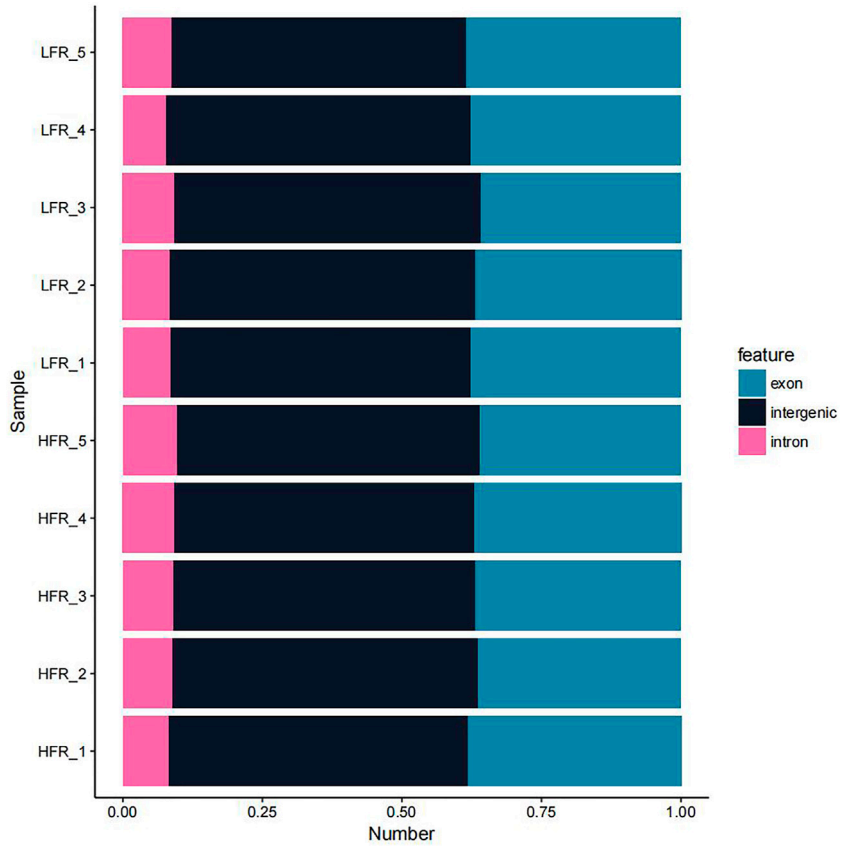


FIGURE 4
circRNA source statistics.

genes, autophagy-animal and ubiquinone and other terpenoid-quinone biosynthesis were the most significantly enriched pathways, and the source genes of differential circRNAs involved

18 miRNAs and 177 mRNAs, and each circRNAs had at least 2 or more miRNA binding sites. Aca-mir-212–5p has 10 target circRNAs and 49 target mRNAs. Hsa-mir-221–3p has 3 target circRNAs and

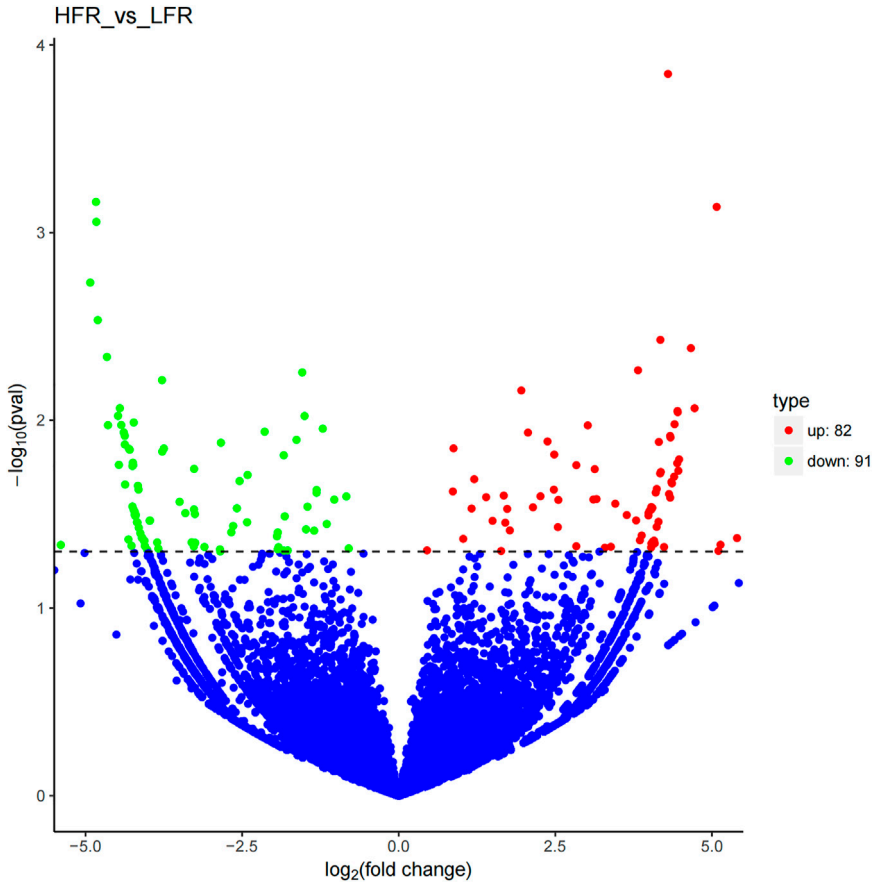


FIGURE 5
Differential circRNA volcano map. Differentially expressed circRNAs were filtered using a $p < 0.05$ as a threshold. Red spots represent up-regulated circRNAs, and green spots indicate down-regulated circRNAs. Blue spots represent circRNAs that did not show obvious changes between the HFR and LFR samples.

10 target mRNAs. Bta-mir-221 has 3 target circRNAs and 11 target mRNAs. In addition, novel_ circ_ 0017590, novel_ circ_0018059 had four miRNA binding sites.

novel_ circ_ 0045314, novel_ circ_ 0019994 and novel_ circ_ 0020422 has two negatively regulated target miRNAs, novell_ circ_0017628, novel_ circ_ 0018615, novel_ circ_ 0021179, novel_circ_ 0029447, novel_ circ_ 0014152 and novel_ circ_ 0010078 has one negatively regulated target miRNA. These results suggested that circRNAs in testis may regulate testicular development and sperm motility by actively participating in binding with miRNAs, acting as competitive endogenous RNAs, regulating the function of target miRNAs, and thus indirectly targeting mRNA levels.

3.7 Fluorescence quantitative polymerase chain reaction

In order to validate the differentially expressed genes identified by transcriptome sequencing, eight circRNAs were randomly selected and confirmed by qRT-PCR using *GAPDH* as the internal reference gene. The results showed that the expression trends of the eight circRNAs were consistent with those of the transcriptome sequencing results. Therefore, the transcriptome sequencing results are reliable and can be further studied and analyzed (Figure 10).

4 Discussion

Semen quality is one of the important indicators to measure the breeding value of male birds. The evaluation indicators of semen quality mainly include semen color, sperm motility, sperm viability, sperm deformity rate, semen volume, pH, etc (Jerysz and Lukaszewicz, 2013; Lukaszewicz et al., 2021). Sperm motility is a reflection of metabolic capacity, and is an important factor in ensuring that the sperm and egg meet and complete fertilization. Studies have shown that sperm motility traits are highly positively correlated with fertilization rate (Nguyen, 2019).

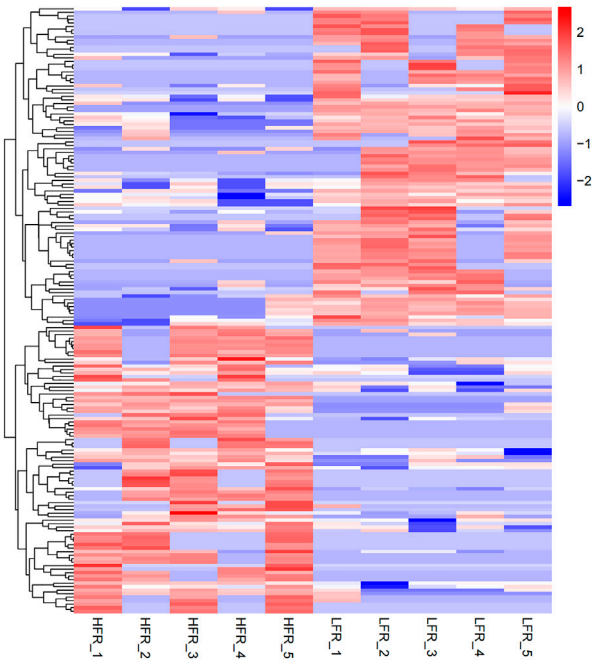


FIGURE 6
Differential circRNA cluster map. The horizontal axis is the sample, and the vertical axis is the differential gene. The left side clusters the genes according to the expression similarity, and the upper part clusters each sample according to the expression profile similarity. The expression level gradually increases from blue to red, and the number is the relative expression level after homogenization.

Moreover, the evaluation of sperm motility traits is relatively simple compared with other semen quality traits, and is easy to generalize. Therefore, sperm motility is the most suitable indicator to measure the reproductive performance of breeding poultry. The testis is an important reproductive organ of male birds, and its main functions are spermatogenesis and endocrine function (Iddrriss et al., 2018). The seminiferous tubules are distributed inside the testis and are the main location of spermatogenesis, making them the most suitable organ for determining male sperm motility.

circRNAs are a new class of RNAs that are different from traditional linear RNAs. They are conserved in different species and are specifically expressed in tissues at different developmental stages. circRNAs can be regulated by affecting the transcription, mRNA conversion and translation of RNA-binding proteins and miRNAs(Panda, 2018; Kristensen et al., 2019). So far, research on circRNAs has mainly focused on humans and model animals, and less research has been carried out on testicular development. Like other vertebrates, the avian testis is the site of spermatogenesis and androgen production, and research on testicular biology in avians mainly focuses on chickens and ducks (Estermann et al., 2021). Research on circRNA in poultry mainly focuses on muscle cell proliferation and differentiation (Ouyang et al., 2018; Shen X. X. et al., 2019), follicle development (Shen M. M. et al., 2020; Wu et al., 2020), disease and immunity (Qiu et al., 2018; Wang et al., 2020) and intramuscular fat deposition (Wang, 2018). However, in the study of the genetic mechanisms regulating sperm motility, the mechanism of generation and the downstream functions of circRNAs remain unclear.

Studies have shown that in mammals, the brain and testis are the tissues with the highest expression levels of circRNAs(Zhou et al., 2018; Mahmoudi and Cairns, 2019). The normal progression of testicular development and spermatogenesis depends on the precise regulation of related genes at the transcriptional and post-transcriptional levels, and ncRNAs can be temporally and spatially refined. At present, research on sperm motility regulation, especially non-coding RNA-mediated gene regulation, is relatively scarce. In this study, based on the data on semen quality and egg fertilization rate, male Yili geese with extreme differences in sperm motility were strictly screened. High-throughput sequencing technology was used for the first time to study the circRNA expression profile of the testicular tissue of Yili geese with high and low sperm motility, and a total of 26,312 circRNAs were identified. Most of these circRNAs were 23 bp-1,511 bp in length, with an average length of 357 bp, which is similar to the length of circRNAs identified in other animals (Liang et al., 2017; Yuan et al., 2018). It is worth noting that in this study, circRNAs were relatively evenly distributed in intergenic, exonic and intronic regions, and most of them were derived from intergenic regions (34.19%). Tang et al, (2020) and other studies have shown that circRNAs are abundantly expressed in male germ cells during spermatogenesis, and circRNA levels increase with the progress of spermatogenesis. This study identified a higher number of circRNAs in high sperm motility testes, indicating that there are more sperm cells in testes with high sperm motility, that is, the spermatogenesis ability is stronger. In addition, compared with the low sperm motility group, a total of 173 DECs were screened, of which 82 were up-regulated and 91 were down-regulated. The source genes of these DECs are widely involved in biological processes such as growth and development, reproduction, and metabolism.

To explore the relationship between these DECs and sperm motility in Yili geese, GO enrichment analysis was performed on the host genes of these circRNAs, and it was found that they were mainly involved in biological processes (63.79%), and were significantly enriched in the reproduction of a single-celled organism, cytoskeletal protein binding, actin binding, adenyl nucleotide binding, microtubule organizing center and other terms. Signaling pathways that play an important role in sperm motility are included in the significantly enriched signaling pathways. Among them, actin is an important part of the cytoskeleton (Hohmann and Dehghani, 2019), and is distributed in mammalian cells, including Sertoli cells in the

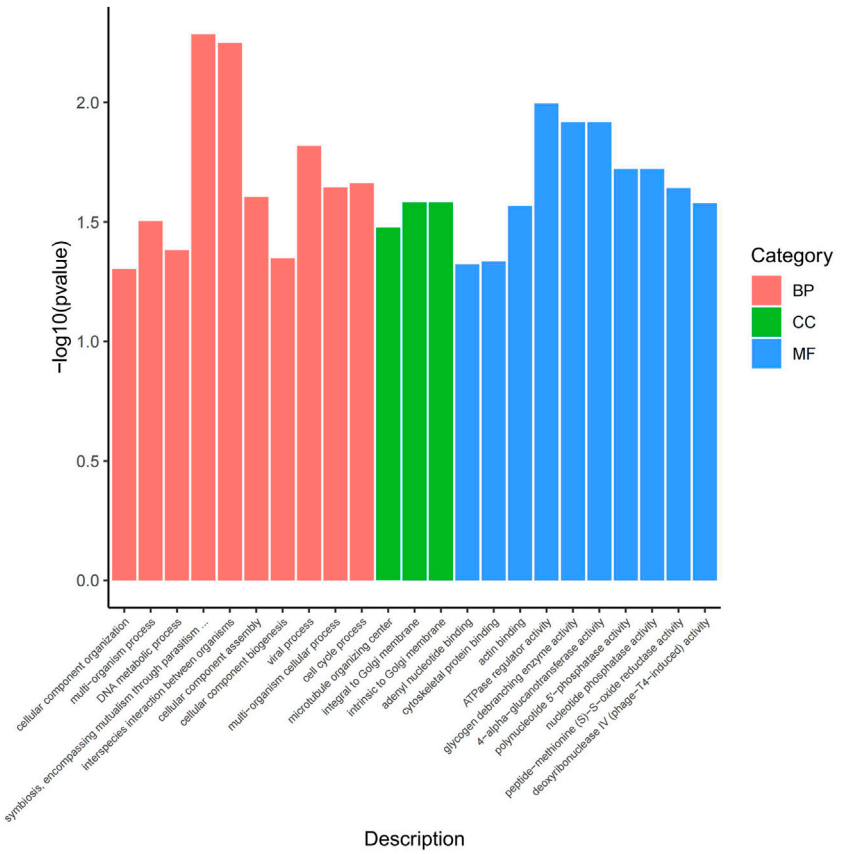


FIGURE 7
GO enrichment map. The horizontal axis represents the GO term of the next level of the three categories of GO, and the vertical axis represents the *p* value ranges.

testis, and actin polymerization in the sperm tail during capacitation regulates motility (Finkelstein et al., 2013). The core of sperm flagella is made up of microtubules, along with tens of thousands of tiny molecular motors called dyneins, which bend these microtubules rhythmically, creating wave after wave of motion, resulting in steering. Sudarshan (Gadadhar et al., 2021) and others found that a special enzymatic modification of tubulin, called glycylation, is the key to keeping sperm moving in a straight line. It is suggested that these circRNAs may be involved in the regulation of sperm motility.

KEGG enrichment analysis showed that autophagy-animal and ubiquinone and other terpenoid-quinone biosynthesis were the most significantly enriched pathways ($p < 0.05$), and the source genes of differential circRNAs involved in these two pathways were *Pdpk1* (phosphomuscle alcohol-dependent protein kinase 1) and *COQ5* (methoxy-6-polyacryl-1,4-benzoquinone methylase). The *PDPK1* gene is expressed in spermatogenic cells, and PDPK1 is localized in the post-acrosomal region of mouse epididymal caudate sperm (Vadnais et al., 2013). *COQ5* is a methyltransferase with adenopicromethionine as a methyl donor. It is located on the stromal side of the inner mitochondrial membrane and plays a crucial role in the stability and activity of the CoQ protein complex (Nguyen et al., 2014). CoQ may not only participate in electron transfer in the mitochondrial respiratory chain of cells, maintaining the functions of cellular respiration and energy metabolism, but also participate in the redox reaction of the body as an antioxidant (Zhong and Zhong, 2014).

Autophagy is a basic process that exists in all eukaryotes, and is involved in the life processes of various cells in the male reproductive system, as well as being involved in the key pathophysiological processes of many diseases of the male reproductive system, such as azoospermia, oligospermia, asthenozoospermia, cryptorchidism, and orchitis (Zhu et al., 2019). Lei et al, (2021) studied the initiation of autophagy in mouse testis and found that after autophagy was blocked in sperm cells, the assembly of a series of key structures during sperm differentiation was disrupted. For example, the deformed scaffold of the sperm–sperm collar (manchette) and the tail shaft (axoneme) to ensure sperm movement at the same time lead to a

TABLE 4 GO enrichment analysis of circRNA host genes.

Item	GO Terms	*P*	Gene number	Gene names	GO_accession
Cellular component	microtubule organizing center	0.033481	2	*GULP1,ANGPT2*	GO:0005815
	integral to Golgi membrane	0.026197	1	*SIPA1L1*	GO:0030173
	intrinsic to Golgi membrane	0.026197	1	*SIPA1L1*	GO:0031228
Molecular function	adenyl nucleotide binding	0.047514	12	*RARS2,PRKCD,USP32,Arl15,NSF,KIF14,EXOC3,TTBK2,Pdpk1,RAB8B,BUB1,GULP1*	GO:0030554
	cytoskeletal protein binding	0.046215	5	*MAP7,Pdpk1,KIF14,ESPN,ANGPT2*	GO:0008092
	actin binding	0.027103	4	*Pdpk1,MAP7,ESPN,ANGPT2*	GO:0003779
	ATPase regulator activity	0.010083	2	*GULP1,STARD13*	GO:0060590
Biological process	cellular component organization	0.049637	9	*GULP1,SIPA1L1,MAP7,Pdpk1,EXOC3,HSF3,ANGPT2,CFAP43,DTNB*	GO:0016043
	multi-organism process	0.031369	7	*STARD13,GULP1,FUT8,HSF3,PLAC9,MAP7,NSF*	GO:0051704
	DNA metabolic process	0.04148	7	*NSF,ADAM9,EXOC3,GULP1,SYDE2,Dmd,STARD13*	GO:0006259
	symbiosis, encompassing mutualism through parasitism	0.0051803	6	*HSF3,FUT8,STARD13,GULP1,NSF,MAP7*	GO:0044403
	interspecies interaction between organisms	0.0056309	6	*FUT8,HSF3,STARD13,GULP1,NSF,MAP7*	GO:0044419

large amount of cytoplasm to remain in the sperm head that ought to have been removed, resulting in morphological deformities and movement disorders. Yefimova et al, (2013) abolished autophagy by germ-cell-specific knockout of *Atg7*, resulting in reduced testicular weight, sperm deformities, and significantly reduced fertility in male mice. Liu et al, (2016) found that Sertoli cell-specific knockout of *Atg5* or *Atg7* disrupted autophagy, resulting in disorders of the vas deferens and deformed sperm heads, which in turn affected the reproductive performance of male mice.

Coenzyme Q (coenzyme Q, CoQ) is a class of quinone substances widely distributed in living organisms. Benzoquinone, also known as ubidearenone and ubiquinone, is a lipid-soluble antioxidant mainly present in the mitochondria, and is also a natural antioxidant and free radical scavenger in mammals (Lafuente et al., 2013; Liu et al., 2017). As an energy promoter and antioxidant, coenzyme Q10 is mainly distributed in the mitochondria in the midsection of sperm, and the availability of the energy required for sperm motility and other energy-dependent processes depends on the availability of coenzyme Q10 (Aby and Lavon, 1997). Studies have shown that sperm concentration, motility and semen parameters are related to *CoQ10* concentration, as CoQ10 can reduce stress oxidation, increase antioxidant enzyme activity, and improve overall antioxidant capacity (Balercia et al., 2009; Lafuente et al., 2013). Studies have confirmed that ubiquinone (Q10) coenzyme is more active in the biosynthesis of the male testis and female follicular fluid, and confirmed that coenzyme Q10 is inseparable from mammalian reproduction (Shetty et al., 2013; Varela-Lopez et al., 2017). Huo Min (Huo, 2018) added different concentrations of *CoQ10* exogenously to a frozen dilution solution, and detected various indicators of cashmere goat sperm after thawing. The results show that compared with the control group, the sperm quality after freezing and thawing was significantly improved ($p < 0.05$). When the concentration of coenzyme Q10 was 400 μg/ml, the sperm motility rate was 60.2%, which was significantly higher than that of the control group ($p < 0.05$), and the intracellular ROS content was significantly decreased ($p < 0.05$); when the concentration of coenzyme Q10 was 40 μg/ml, the sperm plasma membrane integrity rate was 63.0%, the acrosome integrity rate was 76.5%, and the mitochondrial membrane potential was 2.9, which were significantly higher than the levels found in the control group ($p < 0.05$). Hamed et al, (2019) explored the protective effect of coenzyme Q10 (*CoQ10*) and berberine (BB) combined with and without varicocelectomy on sperm

10.3389/fgene.2022.970097

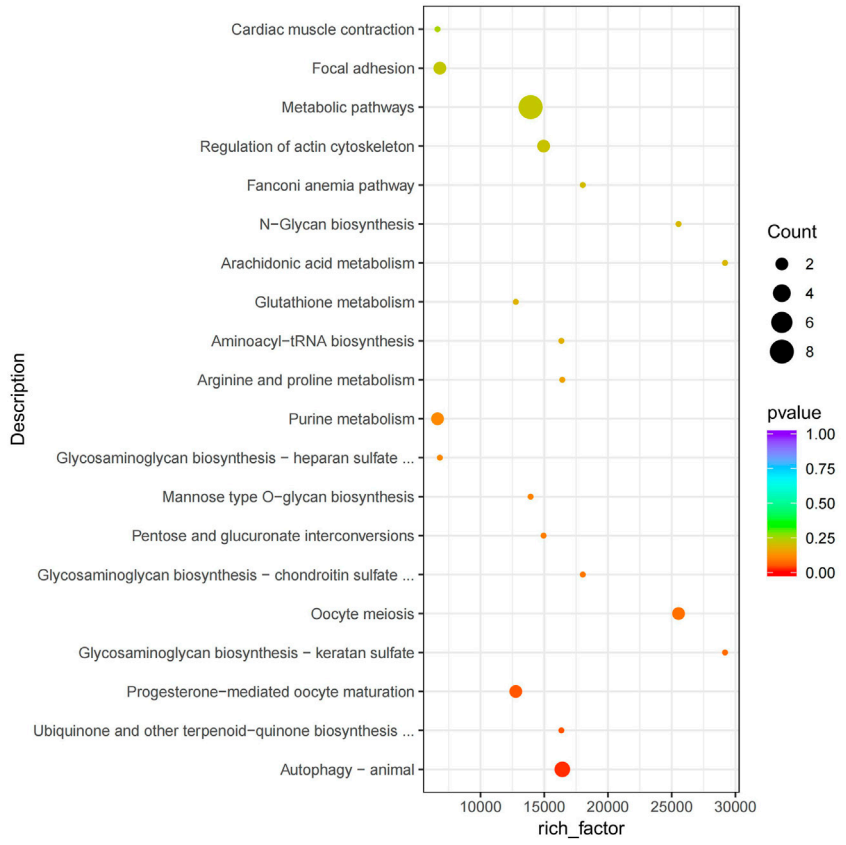


FIGURE 8
KEGG enrichment map. The vertical axis represents the pathway name, the horizontal axis represents rich factor, the size of the point represents the number of differentially expressed genes in this pathway, and the color of the point corresponds to different *p* value ranges.

TABLE 5 KEGG enrichment analysis of circRNA host genes.

Item	KEGG pathway	*P*	Gene number	Gene names	KEGG_ID
KEGG PATHWAY	Autophagy-animal	0.017301176	3	*PRKCD,RB1CC1,Pdpk1*	acyg04140
	Ubiquinone and other terpenoid-quinone biosynthesis	0.046647725	1	*COQ5*	acyg00130
	Progesterone-mediated oocyte maturation	0.051907408	2	*Cpeb3,BUB1*	acyg04914
	Oocyte meiosis	0.079218003	2	Cpeb3,BUB1	acyg04114
	Metabolic pathways	0.223388493	8	*GMPR,ODC1,XYLT1,ALOX5,FUT8,COQ5,ispd,PDE6C*	acyg01100

parameters in rats with postoperative varicocele. They showed that BB and *CoQ10* alone and/or together could improve sperm parameters and reduce sperm DNA damage in varicocele-induced rats. It is worth noting that although the metabolic pathways in the KEGG enrichment were not significantly enriched, the pathways with the most enrichment of source genes were identified. These results indicate that there are cellular differences in testicular tissue in Yili geese with high and low sperm motility, and many of the DECs identified may play a roles in regulating autophagy, anti-oxidation and cell metabolism, thereby participating in the regulation of sperm motility in Yili geese.

The circRNA-miRNA-mRNA interaction network further revealed that cirRNA may regulate the expression of mRNA through the interaction with miRNA to regulate the sperm movement ofYili geese. In the interaction network, miR-

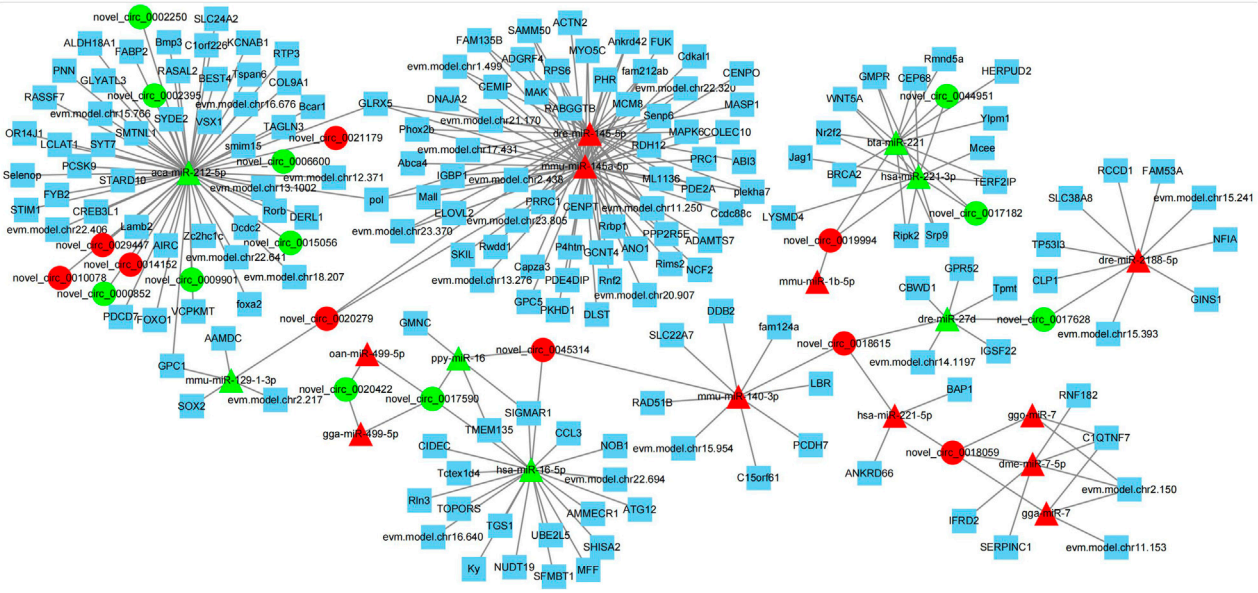


FIGURE 9
Interaction diagram of differential circRNA-miRNA-mRNA network. Triangle nodes represent the targeted miRNAs, circular nodes show circRNAs, and squares nodes show mRNA. Red color represent up-regulated, and green color indicate down-regulated.

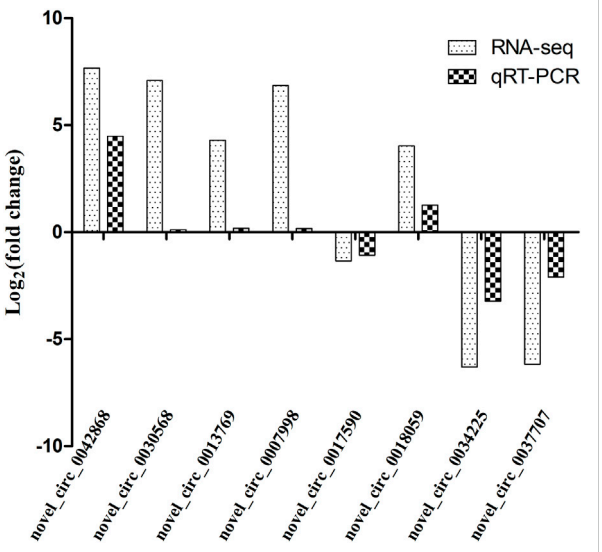


FIGURE 10
Comparison of qRT-PCR and RNA-seq results. Total RNA extracted from the hypothalami, pituitary glands, and ovaries tissues that were measured by qRT-PCR analysis; relative expression levels were calculated according to the $2^{-\Delta\Delta Ct}$ method using *GAPDH* as an internal reference gene.

212–5p, which has the most binding sites for circRNAs, was reported to be developmentally related to the follicles of ewe during FSH stimulation (Hya et al., 2022). Spermatogonial stem-cell-specific miR-221 is involved in the regulation of spermatogenesis in mice (Smorag et al., 2012). Qin et al, (2019) confirmed that circRNA-9119 is a regulatory circRNA involved in testicular inflammation and acts as a sponge for miR26a and miR-136 in Sertoli cells and Leydig cells in response to mimic RNAs, and that the stimulation of viral replication produces poly I:C. These results indicate that circRNAs have potential roles in the regulation of testicular development or sperm motility in Yili geese, and whether these circRNAs really affect testicular development or sperm motility requires further experiments in order to study and verify their mechanisms.

5 Conclusion

In summary, this study analyzed the expression profile of circRNAs in the testis of Yili geese with high and low sperm motility, identified a large number of circRNAs in the testis, revealed the genomic characteristics and length distribution of circRNAs, and constructed the circRNA-miRNA interaction network. The results suggest that circRNAs may regulate the sperm motility of Yili geese through interaction with miRNAs, which provides a solid foundation for identifying and characterizing the key circRNAs involved in testis development or the regulation of sperm motility.

Data availability statement

The data presented in the study are deposited in the BioProjectin NCBI repository (https://www.ncbi.nlm.nih.gov/bioproject/PRJNA856143), accession number is PRJNA856143.

Ethics statement

The animal study was reviewed and approved by Animal Welfare and Ethics Committee of XinJiang Agricultural University (Animal Protocol number: 2020035).

Author contributions

Conceptualization, HL, and YW; methodology, YW; validation, YW and XZ; investigation, YW, XZ, GB, CM, and YY; resources, HL; data curation, YW, JL, YY, LW and GB; writing—original draft preparation, YW and XZ; writing-review and editing, YW; visualization, HL; supervision, HL; project administration, HL; funding acquisition, HL.

Funding

This study was supported by the National Natural Science Foundation of China (Grant No. 32160785), 2022 Xinjiang Agricultural University Graduate Research Innovation Project (XJAUGRI2022004); and 2022 Xinjiang poultry breeding technology integration demonstration and promotion project (2022xjjq-z-01).

Acknowledgments

We highly acknowledge Hengxin Industrial Co., Ltd. of Ermin County, Xinjiang, China, who provided the animals and physical support.

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.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.970097/full#supplementary-material

Reference

Aby, Lewin., and Lavon., Haim (1997). The effect of coenzyme Q10 on sperm motility and function. *Mol. Asp. Med.* 18, s213–s219. doi:10.1016/s0098-2997(97)00036-8

Anders, S., Pyl, P. T., and Huber, W. (2015). HTSeq-a Python framework to work with high-throughput sequencing data. *Bioinformatics* 31 (2), 166–169. doi:10.1093/bioinformatics/btu638

Balercia, G., Buldreghini, E., Vignini, A., Tiano, L., Paggi, F., Amoroso, S., et al. (2009). Coenzyme Q10 treatment in infertile men with idiopathic asthenozoospermia: A placebo-controlled, double-blind randomized trial. *Fertil. Steril.* 91 (5), 1785–1792. doi:10.1016/j.fertnstert.2008.02.119

Bustin, S. A., Benes, V., Garson, J. A., Hellemans, J., Huggett, J., Kubista, M., et al. (2009). The MIQE guidelines: Minimum information for publication of quantitative real-time PCR experiments. *Clin. Chem.* 55, 611–622. doi:10.1373/clinchem.2008.112797

Enright, A. J., John, B., Gaul, U., Tuschl, T., Sander, C., and Marks, D. S. (2003). MicroRNA targets in Drosophila. *Genome Biol.* 5 (1), R1. doi:10.1186/gb-2003-5-1-r1

Estermann, M. A., Major, A. T., and Smith, C. A. (2021). Genetic regulation of avian testis development. *Genes* 12 (9), 1459. doi:10.3390/genes12091459

Finkelstein, M., Megnag, B., Ickowicz, D., and Breitbart, H. (2013). Regulation of sperm motility by PIP2(4, 5) and actin polymerization. *Dev. Biol.* 381 (1), 62–72. doi:10.1016/j.ydbio.2013.06.014

Gadadhar, S., Viar, G. A., Hansen, J. N., Gong, A., Kostarev, A., Ialy-Radio, C., et al. (2021). Tubulin glycylation controls axonemal dynein activity, flagellar beat, and male fertility. *Science* 371 (6525), eabd4914. doi:10.1126/science.abd491410.1126/science.abd4914

Gao, Y., Wang, J., and Zhao, F. (2015). CIRI: An efficient and unbiased algorithm for de novo circular RNA identification. *Genome Biol.* 16 (1), 4. doi:10.1186/s13059-014-0571-3

Guo, T. Y., Huang, L., Yao, W., Du, X., Li, Q. Q., Ma, M. L., et al. (2020). The potential biological functions of circular RNAs during the initiation of atresia in pig follicles. *Domest. Anim. Endocrinol.* 72, 106401. doi:10.1016/j.domaniend.2019.106401

Hamed, N., Hamid, S., Mohammadi, A., Soleimanifar, F., Izadpanah, F., Haddad Kashani, H., et al. (2019). The protective effect of coenzyme Q10 and berberine on sperm parameters, with and without varicocelectomy in rats with surgically induced varicoceles. *Comp. Clin. Path.* 28 (2), 479–485. doi:10.1007/s00580-018-2850-y

Hohmann, T., and Dehghani, F. (2019). The cytoskeleton-A complex interacting meshwork. *Cells* 8 (4), 362. doi:10.3390/cells8040362

Huo, min. (2018). *Effects of Antioxidants on Cryopreservation of Cashmere Semen.* Chinese Dissertation of Inner Mongolia Agricultural University.

Hya, B., Jla, B., Syxa, B., Han, X. Y., Song, X. T., Qi, M. Y., et al. (2022). miRNA expression analysis of the sheep follicle during the prerecruitment, dominant, and mature stages of development under FSH stimulation. *Theriogenology* 181, 161–169. doi:10.1016/j.theriogenology.2022.01.001

Iddrriss, A. R. I., Obese, F. Y., and Robinson, J. E . (2018). Testis size and asymmetry in the Guinea fowl (*Numida meleagris*): A test of the compensation hypothesis. *Avian Biol. Res.* 11 (2), 123–131. doi:10.3184/175815618X15203333340470

Jerysz, A., and Lukaszewicz, E. (2013). Effect of dietary Selenium and Vitamin E on ganders' response to semen collection and Ejaculate characteristics. *Biol. Trace Elem. Res.* 153 (1-3), 196–204. doi:10.1007/s12011-013-9652-5

Kanehisa, M., Araki, M., Goto, S., Hattori, M., Hirakawa, M., Itoh, M., et al. (2008). KEGG for linking genomes to life and the environment. *Nucleic Acids Res.* 36, D480–D484. doi:10.1093/nar/gkm882

Kristensen, L. S., Andersen, M. S., Stagsted, L. V. W., Ebbesen, K. K., Hansen, T. B., and Kjems, J. (2019). The biogenesis, biology and characterization of circular RNAs. *Nat. Rev. Genet.* 20, 675–691. doi:10.1038/s41576-019-0158-7

Kruger, J., and Rehmsmeier, M. (2006). RNAhybrid: microRNA target prediction easy, fast and flexible. *Nucleic Acids Res.* 34, 451–454. doi:10.1093/nar/gkl243

Lafuente, R., Gonzalez-Comadran, M., Sola, I., Lopez, G., Brassesco, M., Carreras, R., et al. (2013). Coenzyme Q10 and male infertility: A meta-analysis. *J. Assist. Reprod. Genet.* 30 (9), 1147–1156. doi:10.1007/s10815-013-0047-5

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 (3), R25–R10. doi:10.1186/gb-2009-10-3-r25

Lei, Y., Zhang, X., Xu, Q., Liu, S., Li, C., Jiang, H., et al. (2021). Autophagic elimination of ribosomes during spermiogenesis provides energy for flagellar motility. *Dev. Cell* 56 (16), 2313–2328.e7. doi:10.1016/j.devcel.2021.07.015

Li, T. T., Luo, R. R., Wang, X., Wang, H., Zhao, X., Guo, Y., et al. (2021). Unraveling stage-dependent expression patterns of circular RNAs and their related ceRNA modulation in ovine postnatal testis development. *Front. Cell Dev. Biol.* 9, 627439. doi:10.3389/fcell.2021.627439

Li, X., Yang, L., and Chen, L. L. (2018). The biogenesis, functions, and challenges of circular RNAs. *Mol. Cell* 71 (3), 428–442. doi:10.1016/j.molcel.2018.06.034

Liang, G., Yan, J., Guo, J., and Tang, Z. (2020). Identification of ovarian circular RNAs and differential expression analysis between Meishan and Large White pigs. *Animals.* 10 (7), 1114. doi:10.3390/ani10071114

Liang, G., Yang, Y., Niu, G., Tang, Z., and Li, K. (2017). Genome-wide profiling of *Sus scrofa* circular RNAs across nine organs and three developmental stages. *DNA Res.* 24 (5), 523–535. doi:10.1093/dnares/dsx022

Liu, C., Wang, H. N., Shang, Y. L., Liu, W., Song, Z., Zhao, H., et al. (2016). Autophagy is required for ectoplasmic specialization assembly in sertoli cells. *Autophagy* 12 (5), 814–832. doi:10.1080/15548627.2016.1159377

Liu, H. T., Cheng, S. B., Huang, Y. C., and Lin, P. T. (2017). Coenzyme Q10 and oxidative stress: Inflammation Status in Hepatocellular carcinoma Patients after Surgery. *Nutrients* 9 (1), 29. doi:10.3390/nu9010029

Lukaszewicz, E., Kowalczyk, A., and Jerysz, A. (2021). Characteristics of semen collected from gander included in the genetic resources conservation program. *Poult. Sci.* 100 (9), 101314. doi:10.1016/j.psj.2021.101314

Mahmoudi, E., and Cairns, M. J. (2019). Circular RNAs are temporospatially regulated throughout development and ageing in the rat. *Sci. Rep.* 9, 2564. doi:10.1038/s41598-019-38860-9

Mao, X., Cai, T., Olyarchuk, J. G., and Wei, L. (2005). Automated genome annotation and pathway identification using the KEGG Orthology (KO) as a controlled vocabulary. *Bioinformatics* 21 (19), 3787–3793. doi:10.1093/bioinformatics/bti430

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 (7441), 333–338. doi:10.1038/nature11928

Michael, I. L., Wolfgang, H., and Simon, A. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15 (12), 550. doi:10.1186/s13059-014-0550-8

Nguyen, T. M. D. (2019). Main signaling pathways involved in the control of fowl sperm motility. *Poult. Sci.* 98, 1528–1538. doi:10.3382/ps/pey465

Nguyen, T. P., Casarin, A., Desbats, M. A., Doimo, M., Trevisson, E., Santos-Ocana, C., et al. (2014). Molecular characterization of the human COQ5 C -methyltransferase in coenzyme Q10 biosynthesis. *Biochim. Biophys. Acta* 1841, 1628–1638. doi:10.1016/j.bbalip.2014.08.007

Ouyang, H. J., Chen, X. L., Wang, Z. J., 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 (1), 71–86. doi:10.1093/dnares/dsx039

Panda, A. C. (2018). "Circular RNAs act as miRNA sponges," in *Circular RNAs. Advances in experimental medicine and biology, vol 1087*. Editor J. Xiao (Singapore: Springer). doi:10.1007/978-981-13-1426-1_6

Paul, S., Andrew, M., Owen, O., Baliga, N. S., Wang, J. T., Ramage, D., et al. (2003). Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Res.* 13 (11), 2498–2504. doi:10.1101/gr.1239303

Qin, L., Lin, J., and Xie, X. (2019). CircRNA-9119 suppresses poly I:C induced inflammation in Leydig and Sertoli cells via TLR3 and RIG-I signal pathways. *Mol. Med.* 25, 28. doi:10.1186/s10020-019-0094-1

Qiu, L., Chang, G., Bi, Y., Liu, X., and Chen, G. (2018). Circular RNA and mRNA profiling reveal competing endogenous RNA networks during avian leukosis virus, subgroup J-induced tumorigenesis in chickens. *PLoS One* 13 (10), e0204931. doi:10.1371/journal.pone.0204931

Quan, G., and Li, J. (2018). Circular RNAs: Biogenesis, expression and their potential roles in reproductionCircular RNAs: Biogenesis, expression and their potential roles in reproduction. *J. Ovarian Res.* 11 (1), 9. doi:10.1186/s13048-018-0381-4

Shen, M., Li, T., Zhang, G., Wu, P., Chen, F., Lou, Q., et al. (2019). Dynamic expression and functional analysis of circRNA in granulosa cells during follicular development in chicken. *BMC Genomics* 20 (1), 96. doi:10.1186/s12864-019-5462-2

Shen, M. M., Wu, P., Li, T. T., Wu, P. F., Chen, F. X., Chen, L., et al. (2020). Transcriptome analysis of circRNA and mRNA in Theca cells during follicular development in chickens. *Genes* 11 (5), 489. doi:10.3390/genes11050489

Shen, M., Wu, P., Li, T., Wu, P., Chen, F., Chen, L., et al. (2020). Transcriptome analysis of circRNA and mRNA in Theca cells during follicular development in chickens. *Genes* 11 (5), 489. doi:10.3390/genes11050489

Shen, X. X., Liu, Z. H., Cao, X. N., He, H., Han, S., Chen, Y., et al. (2019). Circular RNA profiling identified an abundant circular RNA circTMTC1 that inhibits chicken skeletal muscle satellite cell differentiation by sponging miR-128-3p. *Int. J. Biol. Sci.* 15 (10), 2265–2281. doi:10.7150/ijbs.36412

Shetty, R. A., Forster, M. J., and Sumien, N. (2013). Coenzyme Q(10) supplementation reverses age-related impairments in spatial learning and lowers protein oxidation. *Age* 35 (5), 1821–1834. doi:10.1007/s11357-012-9484-9

Smorag, L., Zheng, Y., Nolte, J., Zechner, U., Engel, W., and Pantakani, D. V. K. (2012). MicroRNA signature in various cell types of mouse spermatogenesis: Evidence for stage-specifically expressed miRNA-221, -203 and -34b-5p mediated spermatogenesis regulation. *Biol. Cell* 104 (11), 677–692. doi:10.1111/boc.201200014

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 (3), 211–228. doi:10.1038/s41422-020-0279-8

Vadnais, M. L., Aghajanian, H. K., Lin, A., and Gerton, G. L. (2013). Signaling in sperm: Toward a molecular understanding of the acquisition of sperm motility in the mouse epididymis. *Biol. Reprod.* 89 (5), 127. doi:10.1095/biolreprod.113.110163

Varela-Lopez, A., Ochoa, J. J., Llamas-Elvira, J. M., Lopez-Frias, M., Planells, E., Ramirez-Tortosa, M., et al. (2017). Age-related Loss in Bone Mineral density of rats fed Lifelong on a fish Oil-based diet is Avoided by coenzyme Q10 addition. *Nutrients* 9 (3), 176. doi:10.3390/nu9020176

Wang, L. L., You, Z., Wang, M. Y., Yuan, Y., Liu, C., Yang, N., et al. (2020). Genome-wide analysis of circular RNAs involved in Marek's disease tumourigenesis in chickens. *RNA Biol.* 17 (4), 517–527. doi:10.1080/15476286.2020.1713538

Wang, S. S. (2018). *Integrated analysis to identify candidate functional genes for fat formation and deposition in ducks*. Yangzhou: Yangzhou University, 43–53. doi:10.1038/s41598-017-04178-7

Wu, Y. P., Ding, Y. W., Li, H. Y., Zhao, X. Y., Cao, Y., Pan, L., et al. (2022). Comparative study on semen quality, reproductive performance and testicular development of Yili geese with high and low sperm motility. *Chin. J. Animal Sci.*, 1–9. doi:10.19556/j.0258-7033.20211202-01

Wu, Y., Xiao, H., Pi, J., Zhang, H., Pan, A., Pu, Y., et al. (2020). The circular RNA aplacirc_13267 upregulates duck granulosa cell apoptosis by the apla-miR-1-13/THBS1 signaling pathway. *J. Cell. Physiol.* 235 (18), 5750–5763. doi:10.1002/jcp.29509

Yan, Li., Gao., Guangliang, Yu, Lin., Hu, S., Luo, Y., Wang, G., et al. (2020). Pacific Biosciences assembly with Hi-C mapping generates an improved, chromosome-level goose genome. *GigaScience* 9 (10), giaa114. doi:10.1093/gigascience/giaa114

Yefimova, M. G., Messaddeq, N., Harnois, T., Meunier, A. C., Clarhaut, J., Noblanc, A., et al. (2013). A chimerical phagocytosis model reveals the recruitment by Sertoli cells of autophagy for the degradation of ingested illegitimate substrates. *Autophagy* 9 (5), 653–666. doi:10.4161/auto.23839

Young, M. D., Wakefield, M. J., Smyth, G. K., and Oshlack, A. (2010). Gene ontology analysis for RNA-seq: Accounting for selection bias. *Genome Biol.* 11 (2), R14. doi:10.1186/gb-2010-11-2-r14

Yuan, G., Wu, M., Fan, Y., Li, S., Lai, Z., Huang, Y., et al. (2018). Identification and characterization of circular RNAs in Qinchuan cattle testis. *R. Soc. Open Sci.* 5 (7), 180413. doi:10.1098/rsos.180413

Zhang, F., Zhang, X., Ning, W., Zhang, X., Ru, Z., Wang, S., et al. (2021). Expression analysis of circular RNAs in Y oung and Sexually mature boar testes. *Animals.* 11 (5), 1430. doi:10.3390/ani11051430

Zhao, X. Y., Shaershanbieke, A., Ma, C., Wu, Y. P., Wang, J. H., Li , H. Y., et al. (2020). Effects of dietary Crude protein I evel on Laying raws, reproductive Hormone concentrations and related gene mRNA expression in reproductive Axis of yili geese. *China Anim. Husb. & Veterinary Med.* 47 (06), 1729–1738. doi:10.16431/j.cnki.1671-7236.2020.06.011

Zhao, X. Y., Wu, Y. P., Peng, X., Wang, J. H., Duan, Y. Q., Li, Z. P., et al. (2019). Effects of dietary Crude protein level on performance, Hatching performance and Serum Biochemcial Indexes of yili geese. *Chin. J. Animal Nutr.* 31 (04), 1630–1636.

Zhong, X. L., and Zhong, C. G. (2014). Role of coenzyme Q in xenobiotics-induced mitochondrial damage:its advances[J]. *Chin. J. Pharmacol. Toxicol.* 28 (2), 302–308. doi:10.3867/j.issn.1000-3002.2014.02.027

Zhong, Y. X., Du, Y. J., Yang, X., Mo, Y., Fan, C., Xiong, F., et al. (2018). Circular RNAs function as ceRNAs to regulate and control human cancer progression. *Mol. Cancer* 17 (1), 79. doi:10.1186/s12943-018-0827-8

Zhou, L., Chen, J., Li, Z., Li, X., Hu, X., Huang, Y., et al. (2010). Integrated profiling of microRNAs and mRNAs: microRNAs located on Xq27.3 associate with clear cell renal cell carcinoma. *PLoS One* 305 (12), e15224. doi:10.1371/journal.pone.0015224

Zhou, T., Xie, X., Li, M., Shi, J., Zhou, J., Knox, K., et al. (2018). Rat BodyMap transcriptomes reveal unique circular RNA features across tissue types and developmental stages. *RNA* 24, 1443–1456. doi:10.1261/rna.067132.118

Zhu, Y., Yin, Q., Wei, D., Yang, Z., Du, Y., Ma, Y., et al. (2019). Autophagy in male reproduction. *Syst. Biol. Reprod. Med.* 65 (4), 265–272. doi:10.1080/19396368.2019.1606361

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 21 October 2022
DOI 10.3389/fgene.2022.961142

OPEN ACCESS

EDITED BY
Xiao Wang,
Konge Larsen Aps, Denmark

REVIEWED BY
Jin-Yu Sun,
Nanjing Medical University, China
Denggang Fu,
Indiana University, United States
Ken Mills,
Queen's University Belfast, United Kingdom
Lin Fu,
The Second Affiliated Hospital of Guangzhou Medical University, China

*CORRESPONDENCE
Jiang Lin,
2651329493@qq.com
Jun Qian,
qianjun0007@hotmail.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 04 June 2022
ACCEPTED 10 October 2022
PUBLISHED 21 October 2022

CITATION
Ma J, Wen X, Xu Z, Xia P, Jin Y, Lin J and Qian J (2022), Predicting the influence of Circ_0059706 expression on prognosis in patients with acute myeloid leukemia using classical statistics and machine learning.
Front. Genet. 13:961142.
doi: 10.3389/fgene.2022.961142

Predicting the influence of *Circ_0059706* expression on prognosis in patients with acute myeloid leukemia using classical statistics and machine learning

Jichun Ma[1,2,3†], Xiangmei Wen[1,2,3†], Zijun Xu[1,2,3], Peihui Xia[1,2,3], Ye Jin[4,2,3], Jiang Lin[1,2,3*] and Jun Qian[4,2,3*]

[1]Deparrtment of Central Lab, Affiliated People's Hospital of Jiangsu University, Zhenjiang, China, [2]Zhenjiang Clinical Research Center of Hematology, Affiliated People's Hospital of Jiangsu University, Zhenjiang, China, [3]The Key Lab of Precision Diagnosis and Treatment in Hematologic Malignancies of Zhenjiang City, Affiliated People's Hospital of Jiangsu University, Zhenjiang, China, [4]Deparrtment of Hematology, Affiliated People's Hospital of Jiangsu University, Zhenjiang, China

Background: Various circular RNA (circRNA) molecules are abnormally expressed in acute myeloid leukemia (AML), and associated with disease occurrence and development, as well as patient prognosis. The roles of *circ_0059706*, a circRNA derived from *ID1*, in AML remain largely unclear.

Results: Here, we reported *circ_0059706* expression in *de novo* AML and its association with prognosis. We found that *circ_0059706* expression was significantly lower in AML patients than in controls ($p < 0.001$). Survival analysis of patients with AML divided into two groups according to high and low *circ_0059706* expression showed that overall survival (OS) of patients with high *circ_0059706* expression was significantly longer than that of those with low expression ($p < 0.05$). Further, female patients with AML and those aged >60 years old in the high *circ_0059706* expression group had longer OS than male patients and those younger than 60 years. Multiple regression analysis showed that *circ_0059706* was an independent factor-affecting prognosis of all patients with AML. To evaluate the prospects for application of *circ_0059706* in machine learning predictions, we developed seven types of algorithm. The gradient boosting (GB) model exhibited higher performance in prediction of 1-year prognosis and 3-year prognosis, with AUROC 0.796 and 0.847. We analyzed the importance of variables and found that *circ_0059706* expression level was the first important variables among all 26 factors included in the GB algorithm, suggesting the importance of *circ_0059706* in prediction model. Further, overexpression of *circ_0059706* inhibited cell growth and increased apoptosis of leukemia cells *in vitro*.

Conclusion: These results provide evidence that high expression of *circ_0059706* is propitious for patient prognosis and suggest *circ_0059706* as a potential new biomarker for diagnosis and prognosis evaluation in AML, with high predictive value and good prospects for application in machine learning algorithms.

KEYWORDS
circ_0059706, acute myeloid leukemia, machine learning, prognosis, biomarker

Introduction

Acute myeloid leukemia (AML) is one of the most common hematological malignancies and the most frequent type of acute leukemia in adults (Haubner et al., 2019; Newell and Cook, 2021). In recent years, treatment approaches, including molecular targeted therapy and hematopoietic stem cell transplantation, among others, have led to great progress in improving patient outcomes; however, 5-year survival rates remain low (Huang et al., 2019; Burchert et al., 2020; Vetrie et al., 2020; Du et al., 2021; Pollyea et al., 2021; Zhang et al., 2021). Therefore, research into the molecular mechanisms underlying the occurrence and development of AML is crucial to inform discovery of new clinical markers and therapeutic targets.

Circular RNA (circRNA) is a type of non-coding RNA characterized by a closed ring structure and circRNA molecules are widely distributed in eukaryotes, where they perform complex biological functions (Kristensen et al., 2019; Zhou et al., 2020). As circRNAs do not have 5′ end cap or 3′-polyadenylation tail structures, they cannot easily be recognized and degraded by RNase enzymes (Huang et al., 2020). Hence, circRNAs have high stability, as well as specificity, which contribute to its good potential for application in the field of tumor biomarkers (Xu et al., 2019; Wang et al., 2021; Wei et al., 2020). Some circRNA molecules are abnormally expressed in AML and associated with patient prognosis (Chen et al., 2018; Sun et al., 2019; Lin et al., 2021); for example, *circ-VIM* is significantly up-regulated in AML, and its over-expression is an independent prognostic factor associated with duration of overall and leukemia-free survival of patients with AML (Yi et al., 2019).

Developments in big data and computer hardware and software technologies have led to widespread use of machine learning in medicine (Ngiam and Khor, 2019; Radakovich et al., 2020). Compared with traditional statistics, machine learning has more powerful predictive ability (Lewis and Kemp, 2021; Tran et al., 2021) and its value for application in assisting disease diagnosis and predicting clinical outcomes has also attracted the attention of scholars.

Our previous study has revealed *ID1* transcript level significantly increased in AML and act as an independent risk factor in young non-M3 patients. *Circ_0059706* is a circular RNA, formed by *ID1* during its splicing. In this study, we investigated *circ_0059706* expression in patients with AML, evaluated its clinical significance, and analyzed the predictive ability of *circ_0059706* expression for AML prognosis using machine learning. The aim of the study was to explore the value of *circ_0059706* as a new tumor marker for predicting AML prognosis.

Materials and methods

Patients

This study was approved by the Ethics Committee of the Affiliated People's Hospital of Jiangsu University and included 100 patients newly-diagnosed with AML and 33 healthy controls (K-20190020-Y). All samples were from the sample bank at our hospital and all patients signed informed consent forms. AML was classified according to World Health Organization (WHO) criteria and French-American-British (FAB) classification. Mutations were detected by high-resolution melting analysis (Wen et al., 2015).

Cell culture and transfection

The K562 and THP-1 human leukemia cell line were purchased from ATCC. Cells were cultured in RPMI 1640 medium (Wisent, Nanjing, China) containing 10% fetal calf serum (FCS) (ExCell Bio, Shanghai, China) and 100 U/ml penicillin/streptomycin (Wisent, Nanjing, China) at 37°C in a 5% CO_2 humidified atmosphere. Lentiviruses over-expressing *circ_0059706* were purchased from Shanghai Jikai Biological Co., Ltd. (Shanghai, China) and cell transfection performed according to the manufacturer's instructions.

Cell growth assays

Cells were seeded at 3×10^3 per well in 96-well plates. After culture for 0, 24, 48, and 72 h, 10 μl CCK-8 (Dojindo, Kumamoto, Japan) solution was added to each well. Optical density values were measured at 450 nm absorbance using a microplate reader.

Cell apoptosis assay

Cells (5×10^5) were seeded into 6-well plates containing complete 1640 culture solution, without FCS, for 48 h. Apoptosis rate was determined using an apoptosis detection kit (Annexin V PE/7-AAD, BD Biosciences, Franklin Lakes, NJ, United States),

TABLE 1 Methods used to generate the five derivative variables included in machine learning models.

Variable	Low group	High group
exp-group	Level of *circ_0059706* expression <0.254	Level of *circ_0059706* expression ≥0.254
age-group	Age <60 years	Age ≥60 years
WBC-group	WBC <30 × 10^9/L	WBC ≥30 × 10^9/L
Hb-group	HB < 110 g/L	HB ≥ 110 g/L
Plt-group	PLT <100 × 10^9/L	PLT ≥100 × 10^9/L

and analyzed by flow cytometry on a FACSCalibur platform (Becton Dickinson, San Jose, CA, Unied States).

RNA isolation, reverse transcription, and real-time quantitative PCR (RQ-PCR)

Mononuclear cells were extracted from bone marrow (BM) specimens using gradient centrifugation (TBD Sciences, China). RNA extraction and reverse transcription were conducted based on the instructions in miScript kits (Qiagen, Tilden, Germany). Reverse transcription and RQ-PCR were conducted as previously reported (Ma et al., 2014). The primers for *circ_0059706* were 5′-TGGTAAACTCTCATTCCACGTTC-3' (forward) and 5′-CCA CTGGCGACTTTCATGAT-3' (reverse). The primers used as controls were ABL and sequences were 5′- TCCTCCAGCTGT TATCTGGAAGA -3′ (forward) and 5′-TCCAACGAGCGGCTTCAC -3' (reverse). The primers for miR-326 were 5′-GTCGTATCCAGTGCAGGGTCCGAGGTA TTCGCACTGGATACGA CCTGGAG-3' (forward) and 5′-GCCGAGCCTCTGGGCCCTTC-3' (reverse). The primers for U6 were 5′-CTCGCTTCGGCAGCACA-3' (forward) and 5′-AACGCTTCACGAATTTGCGT-3' (reverse). Relative *circ_0059706* expression levels were calculated using $2^{-\Delta\Delta CT}$ method.

Statistical analysis

Data were analyzed using SPSS 20.0 software. Relative levels of *circ_0059706* expression were calculated using the $2^{-\Delta\Delta CT}$ method. Categorical variables were analyzed using chi square tests and/or Fisher's exact tests. The diagnostic capacity of *circ_0059706* was analyzed using receiver operating characteristic (ROC) curves and area under the curve (AUC) values. Survival was analyzed using the Kaplan-Meier method. Univariate and multivariate Cox regression analyses were conducted. Differences in continuous variables between two groups were compared by Student's t-test. Differences were considered statistically significant at two-tailed $p < 0.05$.

For machine learning, case and survival data, including 26 characteristic variables from 57 cases, were used. Twelve basic variables included in the analysis were: *circ_0059706* expression level, sex, age, white blood cell (WBC) count, hemoglobin (HB) level, platelet (PLT) count, BM blasts, diagnosis, karyotype chromosome abnormalities, chromosome risk group, blast percentage, and granulocyte count. Mutations of nine genes (*CEBPA*, *NPM1*, *FLT3*, *CKIT*, *RAS*, *IDH1/DH2*, *DNMT3A*, *SRSF2*, and *SETBP1*) were also included as variables. Five derivative variables and the methods used to generate them are shown in Table 1. Data from all cases were randomly divided into a training set (75%) for model development and a test set (25%) for performance evaluation. We developed seven types of machine learning algorithm, including: logistic regression (LR), random forest (RF), gradient boosting (GB), neural network (NNK), support vector machine (SVM), k-NearestNeighor (KNN), and Gaussian naïve Bayes (GNB). Parameters of 1-year survival and 3-year survival are as follows: C = 0.01 in LR, n_estimators = 300, random_state = 10 in RF, C = 100 in SVM, MLPClassifier (80, 10), random_state = 100 and random_state = 200 in NNK separately, n_estimators = 500, random_state = 300 and n_estimators = 200, random_state = 280 in GB separately, neighbors = 9 in KNN. Area under the ROC curve (AUROC), sensitivity, and specificity values were used as performance evaluation indicators. Machine learning algorithms were developed using Python software (version 3.7.6.). The "LogisticRegression," "randomforestclassfier," "gradientbootingclassifier," "MLP classifier," "SVC," "Kneighborsclassifier," and "Gaussian NB" functions in the "Sklearn" package (version 1.0.2.) were used for machine learning algorithms, and the "matplotlib" package was applied for machine learning data display.

Results

Circ_0059706 expression and its capacity to distinguish between patients with AML and healthy controls

Levels of *circ_0059706* expression in BM samples from 100 patients with AML and 33 healthy controls were detected

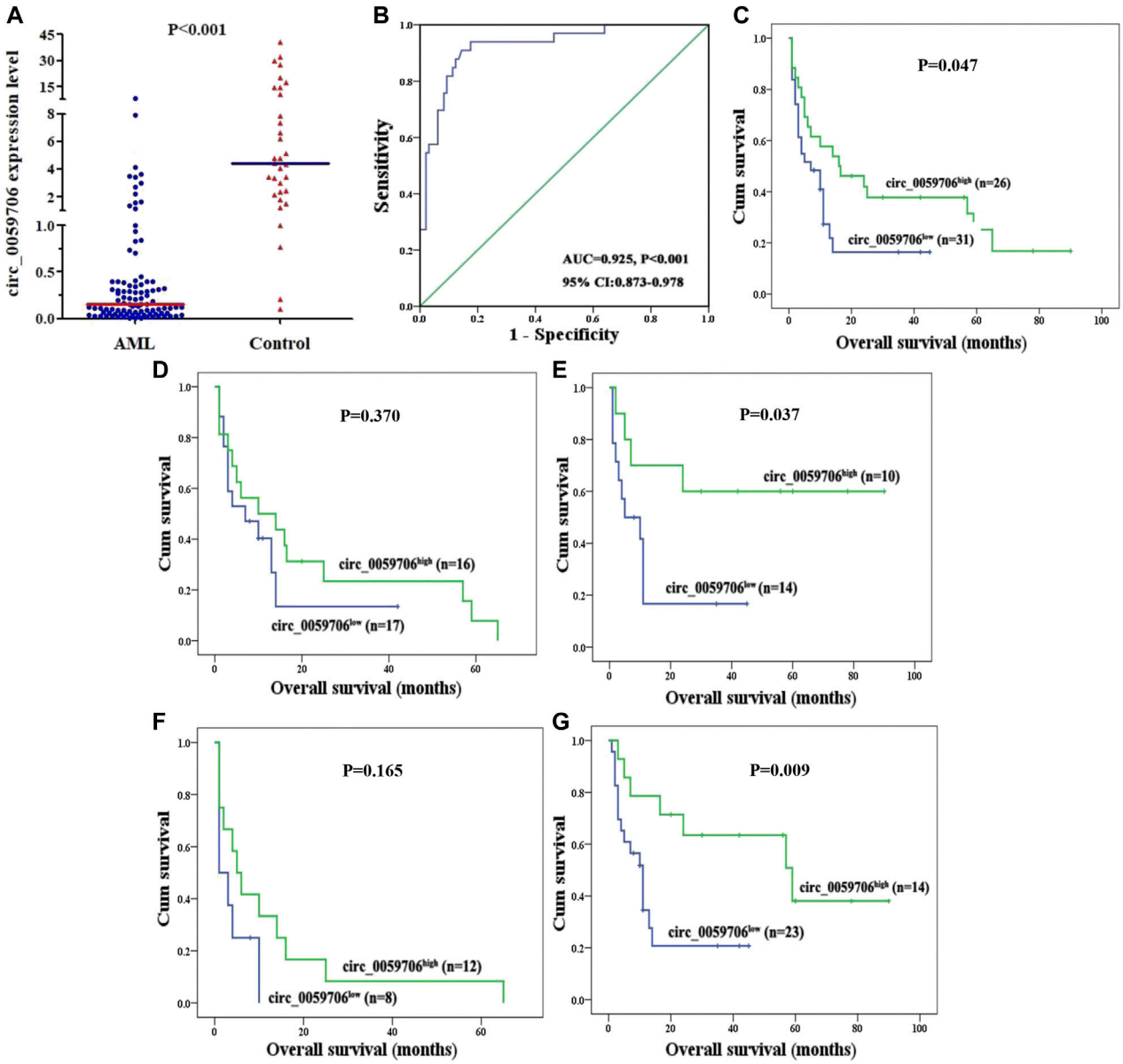


FIGURE 1
Expression of *circ_0059706* in patients with AML and its impact on overall survival. **(A)** Level of *circ_0059706* expression in controls and patients with AML detected by RQ-PCR. **(B)** Discriminative capacity of *circ_0057606* expression in patients with AML determined by ROC curve analysis. **(C–G)** Impact of *circ_0057606* expression on overall survival in: **(C)** total, **(D)** male, **(E)** female, **(F)** age >60 years, and **(G)** age ≤60 years, patients with AML.

by RQ-PCR. Median *circ_0059706* expression levels in healthy controls and patients with AML were 4.581 and 0.153, respectively; *Circ_0059706* expression was significantly lower in AML than in controls ($p < 0.001$) (Figure 1A). The AUC value of *circ_0059706* was 0.925 in AML patients (95% confidence interval: 0.873–0.978, $p < 0.001$) (Figure 1B), indicating that *circ_0059706* expression has potential as biomarker to distinguish AML from control samples.

Associations between *Circ_0059706* expression and patient clinical characteristics

To investigate associations of *circ_0059706* expression with AML clinical characteristics, the total patient group was divided into *circ_0059706*^{high} and *circ_0059706* ^{low} groups, according to median +1/16 standard deviation of *circ_0059706* expression level, using a

10.3389/fgene.2022.961142

TABLE 2 Correlation between circ_0059706 expression and patients parameters.

Patient's parameters	Status of circ_0059706 expression		
	Low (n = 57)	High (n = 40)	*p* value
Sex, male/female	29/28	25/15	0.354
Age, median (range), years$^{\Delta}$	54 (18–84)	57 (10–81)	0.143
WBC, median (range), $\times10^9$/L$^{\Delta}$	37.7 (1.2–207.5)	59.3 (0.8–528.0)	0.124
Hemoglobin, median (range), g/L$^{\Delta}$	84.9 (42–141)	93.2 (32–131)	0.099
Platelets, median (range),$\times10^9$/L$^{\Delta}$	49.5 (4–192)	58.5 (4–382)	0.448
BM blasts, median (range), %$^{\Delta}$	44.5 (0–94.5)	35.0 (0–95)	0.070
FAB classification			0.504
M0	1	0	
M1	2	0	
M2	23	18	
M3	8	5	
M4	13	9	
M5	4	7	
Risk classification			0.965
Low	13	9	
Intermediate	37	25	
High	7	4	
No data	0	2	
Karyotypes			0.661
normal	28	19	
t(8;21)	6	3	
t(16;16)	0	1	
t(15;17)	6	6	
+8	2	2	
t(9;21)	1	0	
−7/7q-	1	0	
Complex	6	3	
Others	7	4	
No data	0	2	
Gene mutations *			
C-KIT (+/−)	2/46	0/31	0.130
FLT3 (+/−)	4/44	5/26	0.147
NPM1(+/−)	12/37	5/28	0.156
C/EBPA (+/−)	5/43	1/31	0.116
N/K-RAS (+/−)	2/36	1/27	0.380
IDH1/2 (+/−)	0/48	1/31	0.115
DNMT3A (+/−)	2/46	2/30	0.343
U2AF1 (+/−)	0/48	0/32	
CR(+/−)	21/20	18/16	0.933

WBC, white blood cells; FAB, French-American-British classification; AML, acute myeloid leukaemia; CR, complete remission.

cutoff value of 0.254 (sensitivity, 91.3%; specificity, 85.6%). Then, clinical parameters were compared between the high and low expression groups (Table 2). No significant differences were detected between the high and low expression groups; however, HB level, WBC count, and platelet count were higher in peripheral blood from the *circ_0059706*^{high} group than from the *circ_0059706*^{low} group, and the proportion of BM blast cells in the *circ_0059706*^{high} group was lower than that of the *circ_0059706*^{low} group. No correlations between gene mutations and *circ_0059706* expression were detected.

TABLE 3 Univariate and multivariate analyses of prognostic factors for overall survival in whole-cohort AML patients.

Prognostic factors	Univariate		Multivariate	
	Hazard ratio (95% CI)	*p* value	Hazard ratio (95% CI)	*p* value
Age	2.557 (1.379–4.739)	0.003	2.519 (1.238–5.125)	0.011
Risk classification	2.265 (1.342–3.820)	0.002	1.527 (0.858–2.720)	0.150
circ_0059706 expression	0.529 (0.273–1.024)	0.059	0.340 (0.166–0.699)	0.003
WBC	2.482 (1.326–4.647)	0.004	2.077 (1.036–4.165)	0.040
Sex	0.628 (0.333–1.187)	0.152	---	---
IDH1/2 mutations (+/−)	4.936 (0.625–38.974)	0.130	---	---
NPM1 mutations (+/−)	1.767 (0.808–3.862)	0.154	---	---

+: positive; −: negative; *, +: bi-allelic mutation; −: mono-allelic mutation or wild type.

TABLE 4 Predictive performance comparison in the six kinds of machine learning algorithms.

	One year survival				Three years survival			
	AUROC	Sebsitivity	Specificity	Accuracy	AUROC	Sebsitivity	Specificity	Accuracy
LR	0.574	0.615	0.5	0.667	0.805	0.857	1	0.867
RF	0.722	0.6	0	0.6	1	0.8	0	0.8
GB	0.796	0.615	0.75	0.733	0.847	0.923	1	0.933
NNK	0.685	0.692	1	0.733	0.75	0.857	1	0.867
SVM	0.593	0.615	0.5	0.6	0.855	0.8	0	0.8
KNN	0.546	0.583	0.333	0.533	0.903	0.833	0	0.8
GNB	0.5	0.5	0.364	0.4	0.333	0.667	0	0.533

Association of *Circ_0059706* expression level and prognosis of patients with AML

To explore association of *circ_0059706* expression with patient prognosis, we analyzed survival data from 57 patients. Although *circ_0059706* had no value for predicting complete remission (CR), patients in the *circ_0059706*^{high} group had significantly longer overall survival (OS) than those in the *circ_0059706*^{low} group (p = 0.047) (Figure 1C). Further, compared with male patients with AML, OS of *circ_0059706*^{high} female patients was longer than that of *circ_0059706*^{low} female patients (p = 0.037) (Figures 1D,E). Compared with older than 60 years, OS of those in the *circ_0059706*^{high} group was significantly longer than that of those in the *circ_0059706*^{low} group who were younger than 60-years-old AML (p = 0.009) (Figures 1F,G). Variables resulting in $p < 0.1$ in univariate analysis (age, WBC, risk classification, and *circ_0059706* expression) were included in multivariate analysis, which demonstrated that *circ_0059706* was an independent factor associated with poor prognosis in the total AML patients (p = 0.020) (Table 3).

Evaluation of prediction ability of *Circ_0059706* in machine learning algorithms

First, LR, RF, GB, NNK, SVM, KNN, and GNB 7 machine learning algorithms were developed using training set data, and their performance evaluated. As shown in Table 4, the GB model had better performance in predicting 1-year prognosis and 3-year prognosis, with AUROC 0.796 and 0.847, and sensitivity 0.615 and 0.923, specificity 0.75 and 1.

The selection of variables for machine learning algorithms is critical. Therefore, we analyzed the importance of variables included in the GB, LR, and RF algorithms, which had good modeling performance. *Circ_0059706* expression level was the first important variable among all 26 features included in the GB

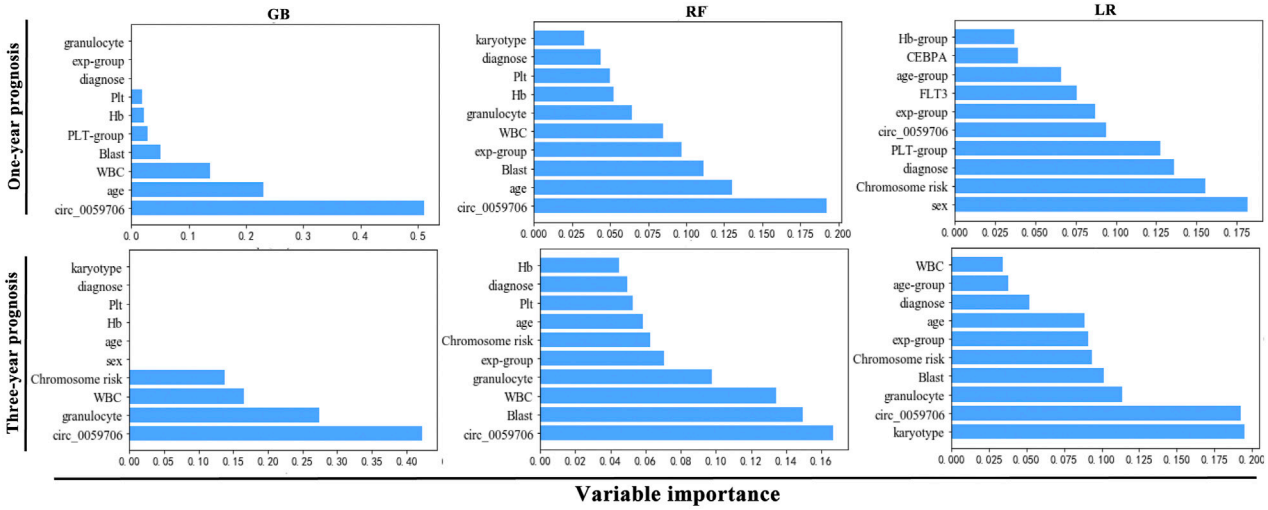


FIGURE 2
Top 10 variables most important in each algorithm.

and RF algorithms and was among the most important in the LR algorithm (Figure 2).

Circ_0059706 inhibits cell growth and increases apoptosis

To study the effect of *circ_0059706* on leukemia cells, we over-expressed it in THP-1 and K562 cells (Figure 3A), and found that cell growth rate was inhibited by *circ_0059706* over-expression (Figure 3B). Moreover, the apoptosis rate of *circ_0059706*-transfected THP-1 and K562 cells was significantly higher than that of the control group ($p < 0.01$) (Figure 3C).

To determine the possible mechanism involved in the functionality of *circ_0059706*, we performed additional analysis. We analyzed the miRNAs that may bind to *circ_0059706* by use of the circinteratome database (https://circinteractome.nia.nih.gov/index.html). Then, the expression levels of these miRNAs in AML patients were analyzed by the GEO database (Datasets: GSE51908). Bioinformatics analysis revealed that *circ_0059706* contains a binding site for miR-326 (Figure 3D). The levels of *miR-326* were lower in AML patients when compared with controls in GSE51908 datasets (Figure 3E). Finally, the expression levels of *miR-326* were found to be up-regulated in cells that overexpressed *circ_0059706* (Figure 3F).

Discussion

CircRNAs are non-coding RNAs that have recently emerged as a potential tumor biomarkers and drug targets, with good prospects for clinical application. *CircPLXNB2* is a valuable predictor of prognosis in patients with AML (Lin, Wang, Bian, Sun, Guo, Kong, Zhao, Guo, Li, Wu, Wang, Wang and Li, 2021). Further, Liu et al. (2022) found that expression of *hsa_circ_0004277* can be restored after chemotherapy in patients with AML, suggesting that its up-regulation is associated with successful treatment; hence, *hsa_circ_0004277* is a potential diagnostic marker and treatment target in AML. Further, *circRNF220* can distinguish AML from ALL and other hematological malignancies, while high *circRNF220* expression is an unfavorable prognostic marker of recurrence, due to its role in sequestration of miR-30a, which increases MYSM1 and IER2 expression and is implicated in AML relapse (Liu et al., 2021).

Circ_0059706 derived from *ID1*. *ID1* is a negative regulator of the HLH transcription factor that plays the role of an oncogene in promoting cell cycle, proliferation and inhibiting apoptosis (Yin et al., 2017; Chen et al., 2020). Our group previously reported that *ID1* expression was up-regulated in AML patients, and the high expression associated with poor prognoses. Here, we found that *circ_0059706* was down-regulated in AML. HB, WBC, and PLT levels were higher in peripheral blood from the *circ_0059706*^{high} group than those in the group with low *circ_0059706* expression. Further, the proportion of BM blast cells in the *circ_0059706*^{high} group was lower than that in the *circ_0059706*^{low} group. Hence, routine blood parameters appeared to be better in the high expression group than those of the low expression group. There were no correlations between gene mutations and *circ_0059706* expression. Survival analysis showed that *circ_0059706* has no value for prediction of CR; however, OS was significantly higher in patients with high *circ_0059706*

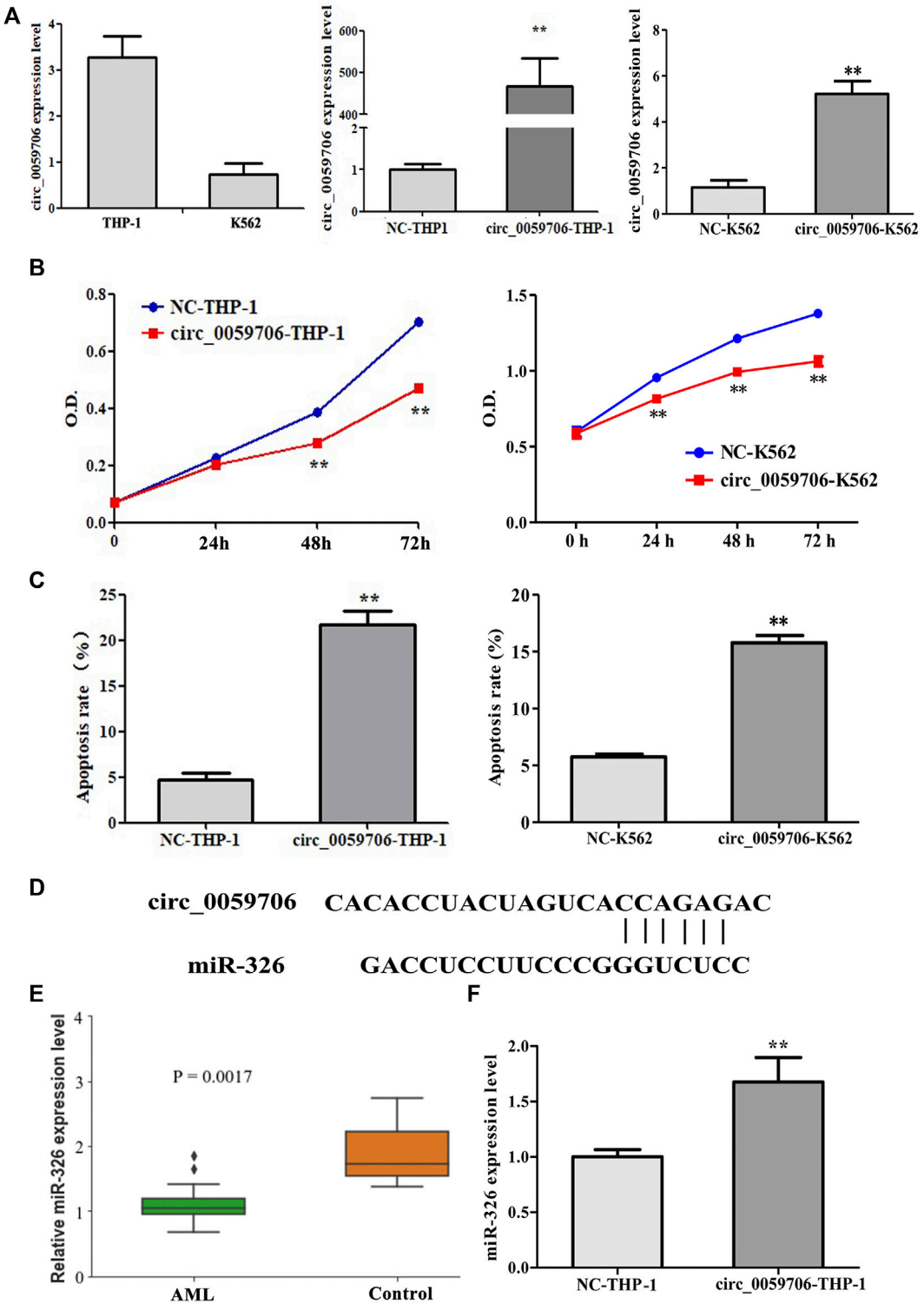


FIGURE 3
Effect of *circ_0059706* on cell growth and apoptosis. **(A)** The level of *circ_0059706* expression was detected by RQ-PCR. **(B)** *Circ_0059706* inhibited cell growth in THP-1 and K562 cells detected by CCK8. **(C)** *Circ_0059706* promoted cellular apoptosis in THP-1 and K562 cells, as detected by flow cytometry. **(D)** Bioinformatics analysis of the binding site between circ_0059706 and miR-326. **(E)** Expression levels of miR-326 in controls and AML patients were analyzed in the GSE51908 datasets. **(F)** The level of miR-326 was detected in *circ_0059706* overexpressed cells by RQ-PCR. **, compared with the NC group, $p < 0.01$.

expression than those with low expression. Longer OS was observed in female patients and those >60-years-old with high *circ_0059706* expression than in male patients <60-years-old. Furthermore, the total patient group was divided into *circ_0059706*^{high} and *circ_0059706*^{low} groups, according to circ_0059706 expression level at quartiles or tertiles. Patients in the *circ_0059706*^{high} group had significantly longer OS than those in the *circ_0059706*^{low} group (p = 0.038, p = 0.027). Multivariate analysis showed that *circ_0059706* low expression was an independent factor associated with poor prognosis of all patients with AML, indicating that *circ_0059706* has potential for application as a new biomarker for diagnosis and prognosis evaluation of AML.

Traditional statistics are generally used to infer relationships between variables, while machine learning models aim to make the most accurate predictions possible, and are increasingly being applied in medical prediction models. Gao et al. predicted a significant association between Luminal and HER2 breast cancer subtypes and estrogen/progesterone and HER2 receptor status, using the DeepCC method (Gao et al., 2019). Lee et al. comprehensively analyzed RNA-seq data and identified a potential role for machine learning in identifying categories of acute leukemia (Lee et al., 2021). Based on traditional statistical analysis, we found that *Circ_0059706* level are closely related with survival, suggesting its potential as a biomarker in patients with AML. Different machine learning algorithms may be optimal for any data set; therefore, to assess the prospects for application of *circ_0059706* levels in AML in machine learning algorithms, we developed seven types of machine learning algorithm, including LR, RF, GB, NNK, SVM, KNN, and GNB. The GB model had better performance in predicting 1-year prognosis and 3-year prognosis. We were unable to predict 5-year survival due to insufficient data. In recent years, scholars have established many risk assessment methods in various disease prognosis models using machine learning, which can provide guidance for the selection of treatment methods and prognosis assessment. For example, Heo applied an NNK algorithm to establish a prediction model for long-term prognosis in patients following ischemic stroke (Heo et al., 2019), while Tian established an early gastric cancer lymph node metastasis prediction model using a regularized dual averaging approach (Tian et al., 2021). Moreover, machine learning achieved acceptable prediction of central lymph node metastasis, with a GB model performing best, which may help to determine the optimal extent of initial surgical treatment for patients with T1–T2 stage, non-invasive, and clinically node-negative papillary thyroid cancer (Zhu et al., 2021).

Variables are crucial to the prediction results generated by machine learning; hence, the key roles of variables included in the machine learning models was also a focus of our attention. We analyzed the importance of variables in the GB, LR, and RF algorithms, which had good modeling performance. *Circ_0059706* expression level was the first important variable among all 26 features included in the GB and RF algorithms, and it ranked highly in the LR algorithm. It indicated that *circ_0059706* has a high predictive value and a good prospect for application in machine learning, supporting the potential of this circRNA as a new biomarker for diagnosis and prognosis evaluation in AML.

Furthermore, we analyzed the effect of *circ_0059706* on cell growth and apoptosis in leukemia cells. The results showed that *circ_0059706* overexpression inhibited cell growth and increased apoptosis, further supporting the hypothesis that the high expression of this circRNA is propitious for patient prognosis. To investigate the mechanisms, we analyzed miRNAs with common binding sites for *circ_0059706* in the circinteratome database. The expression levels of miRNAs were analyzed by datasets GSE51908. Combined with literature reports, we focused *miR-326*, which was downregulated in GSE51908 datasets. P Cheng reported that expression of *miR-326* was downregulated in AML patients compared with that in normal. Overexpression of *miR-326* inhibited proliferation, promoted cell apoptosis and PMA-induced differentiation in AML cells (Cheng et al., 2020). Moreover, *miR-326* down regulated in ALL patients and negative associated with its expression and MDR (Ghodousi and Rahgozar, 2018). These results suggested *miR-326* maybe act as a tumor suppressor miRNA in leukemia and it was upregulated in *circ_0059706* over-expressed cells. *miR-326* was up-regulated in *circ_0059706* overexpression cells, it may be a mechanism of inhibited growth and promoted cell apoptosis. However, more experiments needed to verify, such as luciferase reporter experiment, the effect of up/down-regulation of *miR-326* expression on cell biological function, etc.

Conclusion

Taken together, our results indicate that down-regulation of *circ_0059706* is a frequent event and predicts poor prognosis in patients with *de novo* AML. *Circ_0059706* showed good predictive effects in machine learning models and was among the most important variables in the developed models. In addition, *circ_0059706* overexpression could inhibit cell growth and increase apoptosis. These results demonstrated that *circ_0059706* might act as a potential biomarker for prognosis in *de novo* AML.

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 Ethics Committee of the Affiliated People's Hospital of Jiangsu University. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

Author contributions

JM, XW, and ZX: performed the experiments. PX and YJ collected the data, performed data analyses. JL provided the resources. Jun Qian reviewed and edited the manuscript.

Funding

This study was supported by National Natural Science foundation of China (81900163,81970118), "Liu Ge Yi Gong Cheng" of Jiangsu Province (LGY2018024), China Postdoctoral Science Foundation funded project (2016M601748), youth medical talents project of "Ke Jiao Qiang Wei" project of Jiangsu province (QNRC2016450), Zhenjiang Clinical Research Center of Hematology (SS2018009). Social Development Foundation of Zhenjiang (SH2022065, SH2019067), Scientific Research Project of the Fifth 169 Project of Zhenjiang (21). Project of Zhenjiang First People's Hospital (Y2021010-S).

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

References

Burchert, A., Bug, G., Fritz, L. V., Finke, J., Stelljes, M., Rollig, C., et al. (2020). Sorafenib maintenance after allogeneic hematopoietic stem cell transplantation for acute myeloid leukemia with FLT3-internal tandem duplication mutation (SORMAIN). *J. Clin. Oncol.* 38 (26), 2993–3002. doi:10.1200/JCO.19.03345

Chen, H., Liu, T., Liu, J., Feng, Y., Wang, B., Wang, J., et al. (2018). Circ-ANAPC7 is upregulated in acute myeloid leukemia and appears to target the MiR-181 family. *Cell. Physiol. biochem.* 47 (5), 1998–2007. doi:10.1159/000491468

Chen, S. D., Yang, J. L., Lin, Y. C., Chao, A. C., and Yang, D. I. (2020). Emerging roles of inhibitor of differentiation-1 in alzheimer's disease: Cell cycle reentry and beyond. *Cells* 9 (7), E1746. doi:10.3390/cells9071746

Cheng, P., Lu, P., Guan, J., Zhou, Y., Zou, L., Yi, X., et al. (2020). LncRNA KCNQ1OT1 controls cell proliferation, differentiation and apoptosis by sponging miR-326 to regulate c-Myc expression in acute myeloid leukemia. *Neoplasma* 67 (2), 238–248. doi:10.4149/neo_2018_181215N972

Du, W., Xu, A., Huang, Y., Cao, J., Zhu, H., Yang, B., et al. (2021). The role of autophagy in targeted therapy for acute myeloid leukemia. *Autophagy* 17 (10), 2665–2679. doi:10.1080/15548627.2020.1822628

Gao, F., Wang, W., Tan, M., Zhu, L., Zhang, Y., Fessler, E., et al. (2019). DeepCC: A novel deep learning-based framework for cancer molecular subtype classification. *Oncogenesis* 8 (9), 44. doi:10.1038/s41389-019-0157-8

Ghodousi, E. S., and Rahgozar, S. (2018). MicroRNA-326 and microRNA-200c: Two novel biomarkers for diagnosis and prognosis of pediatric acute lymphoblastic leukemia. *J. Cell. Biochem.* 119 (7), 6024–6032. doi:10.1002/jcb.26800

Haubner, S., Perna, F., Kohnke, T., Schmidt, C., Berman, S., Augsberger, C., et al. (2019). Coexpression profile of leukemic stem cell markers for combinatorial targeted therapy in AML. *Leukemia* 33 (1), 64–74. doi:10.1038/s41375-018-0180-3

Heo, J., Yoon, J. G., Park, H., Kim, Y. D., Nam, H. S., and Heo, J. H. (2019). Machine learning-based model for prediction of outcomes in acute stroke. *Stroke* 50 (5), 1263–1265. doi:10.1161/STROKEAHA.118.024293

Huang, A., Zheng, H., Wu, Z., Chen, M., and Huang, Y. (2020). Circular RNA-protein interactions: Functions, mechanisms, and identification. *Theranostics* 10 (8), 3503–3517. doi:10.7150/thno.42174

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 (4), 677–691. doi:10.1016/j.ccell.2019.03.006

Kristensen, L. S., Andersen, M. S., Stagsted, L. V. W., Ebbesen, K. K., Hansen, T. B., and Kjems, J. (2019). The biogenesis, biology and characterization of circular RNAs. *Nat. Rev. Genet.* 20 (11), 675–691. doi:10.1038/s41576-019-0158-7

Lee, J., Cho, S., Hong, S. E., Kang, D., Choi, H., Lee, J. M., et al. (2021). Integrative analysis of gene expression data by RNA sequencing for differential diagnosis of acute leukemia: Potential application of machine learning. *Front. Oncol.* 11, 717616. doi:10.3389/fonc.2021.717616

Lewis, J. E., and Kemp, M. L. (2021). Integration of machine learning and genome-scale metabolic modeling identifies multi-omics biomarkers for radiation resistance. *Nat. Commun.* 12 (1), 2700. doi:10.1038/s41467-021-22989-1

Lin, L., Wang, Y., Bian, S., Sun, L., Guo, Z., Kong, D., et al. (2021). A circular RNA derived from PLXNB2 as a valuable predictor of the prognosis of patients with acute myeloid leukaemia. *J. Transl. Med.* 19 (1), 123. doi:10.1186/s12967-021-02793-7

Liu, X., Liu, X., Cai, M., Luo, A., He, Y., Liu, S., et al. (2021). CircRNF220, not its linear cognate gene RNF220, regulates cell growth and is associated with relapse in pediatric acute myeloid leukemia. *Mol. Cancer* 20 (1), 139. doi:10.1186/s12943-021-01395-7

Liu, Y., Chen, X., Liu, J., Jin, Y., and Wang, W. (2022). Circular RNA circ_0004277 inhibits acute myeloid leukemia progression through MicroRNA-134-5p/single stranded DNA binding protein 2. *Bioengineered* 13 (4), 9662–9673. doi:10.1080/21655979.2022.2059609

Ma, J., Lin, J., Qian, J., Qian, W., Yin, J., Yang, B., et al. (2014). MiR-378 promotes the migration of liver cancer cells by down-regulating Fus expression. *Cell. Physiol. biochem.* 34 (6), 2266–2274. doi:10.1159/000369669

Newell, L. F., and Cook, R. J. (2021). Advances in acute myeloid leukemia. *BMJ* 375, n2026. doi:10.1136/bmj.n2026

Ngiam, K. Y., and Khor, I. W. (2019). Big data and machine learning algorithms for health-care delivery. *Lancet. Oncol.* 20 (5), e262–e273. doi:10.1016/S1470-2045(19)30149-4

Pollyea, D. A., Bixby, D., Perl, A., Bhatt, V. R., Altman, J. K., Appelbaum, F. R., et al. (2021). NCCN guidelines insights: Acute myeloid leukemia, version 2.2021. *J. Natl. Compr. Canc. Netw.* 19 (1), 16–27. doi:10.6004/jnccn.2021.0002

Radakovich, N., Nagy, M., and Nazha, A. (2020). Machine learning in haematological malignancies. *Lancet. Haematol.* 7 (7), e541–e550. doi:10.1016/S2352-3026(20)30121-6

Sun, Y. M., Wang, W. T., Zeng, Z. C., Chen, T. Q., Han, C., Pan, Q., et al. (2019). circMYBL2, a circRNA from MYBL2, regulates FLT3 translation by recruiting PTBP1 to promote FLT3-ITD AML progression. *Blood* 134 (18), 1533–1546. doi:10.1182/blood.2019000802

Tian, H., Ning, Z., Zong, Z., Liu, J., Hu, C., Ying, H., et al. (2021). Application of machine learning algorithms to predict lymph node metastasis in early gastric cancer. *Front. Med.* 8, 759013. doi:10.3389/fmed.2021.759013

Tran, K. A., Kondrashova, O., Bradley, A., Williams, E. D., Pearson, J. V., and Waddell, N. (2021). Deep learning in cancer diagnosis, prognosis and treatment selection. *Genome Med.* 13 (1), 152. doi:10.1186/s13073-021-00968-x

Vetrie, D., Helgason, G. V., and Copland, M. (2020). The leukaemia stem cell: Similarities, differences and clinical prospects in CML and AML. *Nat. Rev. Cancer* 20 (3), 158–173. doi:10.1038/s41568-019-0230-9

Wang, S., Zhang, K., Tan, S., Xin, J., Yuan, Q., Xu, H., et al. (2021). Circular RNAs in body fluids as cancer biomarkers: The new frontier of liquid biopsies. *Mol. Cancer* 20 (1), 13. doi:10.1186/s12943-020-01298-z

Wei, C. Y., Zhu, M. X., Lu, N. H., Liu, J. Q., Yang, Y. W., Zhang, Y., et al. (2020). Circular RNA circ_0020710 drives tumor progression and immune evasion by regulating the miR-370-3p/CXCL12 axis in melanoma. *Mol. Cancer* 19 (1), 84. doi:10.1186/s12943-020-01191-9

Wen, X. M., Hu, J. B., Yang, J., Qian, W., Yao, D. M., Deng, Z. Q., et al. (2015). CEBPA methylation and mutation in myelodysplastic syndrome. *Med. Oncol.* 32 (7), 192. doi:10.1007/s12032-015-0605-z

Xu, H., Wang, C., Song, H., Xu, Y., and Ji, G. (2019). RNA-Seq profiling of circular RNAs in human colorectal Cancer liver metastasis and the potential biomarkers. *Mol. Cancer* 18 (1), 8. doi:10.1186/s12943-018-0932-8

Yi, Y. Y., Yi, J., Zhu, X., Zhang, J., Zhou, J., Tang, X., et al. (2019). Circular RNA of vimentin expression as a valuable predictor for acute myeloid leukemia development and prognosis. *J. Cell. Physiol.* 234 (4), 3711–3719. doi:10.1002/jcp.27145

Yin, X., Tang, B., Li, J. H., Wang, Y., Zhang, L., Xie, X. Y., et al. (2017). ID1 promotes hepatocellular carcinoma proliferation and confers chemoresistance to oxaliplatin by activating pentose phosphate pathway. *J. Exp. Clin. Cancer Res.* 36 (1), 166. doi:10.1186/s13046-017-0637-7

Zhang, X. H., Chen, J., Han, M. Z., Huang, H., Jiang, E. L., Jiang, M., et al. (2021). The consensus from the Chinese society of Hematology on indications, conditioning regimens and donor selection for allogeneic hematopoietic stem cell transplantation: 2021 update. *J. Hematol. Oncol.* 14 (1), 145. doi:10.1186/s13045-021-01159-2

Zhou, W. Y., Cai, Z. R., Liu, J., Wang, D. S., Ju, H. Q., and Xu, R. H. (2020). Circular RNA: Metabolism, functions and interactions with proteins. *Mol. Cancer* 19 (1), 172. doi:10.1186/s12943-020-01286-3

Zhu, J., Zheng, J., Li, L., Huang, R., Ren, H., Wang, D., et al. (2021). Application of machine learning algorithms to predict central lymph node metastasis in T1-T2, non-invasive, and clinically node negative papillary thyroid carcinoma. *Front. Med.* 8, 635771. doi:10.3389/fmed.2021.635771

frontiers | Frontiers in Cell and Developmental Biology

TYPE Brief Research Report
PUBLISHED 25 October 2022
DOI 10.3389/fcell.2022.1016955

OPEN ACCESS

EDITED BY
Xiao Wang,
Konge Larsen ApS, Denmark

REVIEWED BY
Nagarajan Kannan,
Mayo Clinic, United States
Miodrag Guzvic,
University Medical Center Regensburg, Germany

*CORRESPONDENCE
Rafael López-López,
rafael.lopez.lopez@sergas.es
Angel Díaz-Lagares,
angel.diaz.lagares@sergas.es

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

‡These authors have contributed equally to this work and share senior authorship

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

RECEIVED 11 August 2022
ACCEPTED 12 October 2022
PUBLISHED 25 October 2022

CITATION
Rodriguez-Casanova A, Costa-Fraga N, Castro-Carballeira C, González-Conde M, Abuin C, Bao-Caamano A, García-Caballero T, Brozos-Vazquez E, Rodriguez-López C, Cebey V, Palacios P, Cueva JF, López-López R, Costa C and Díaz-Lagares A (2022), A genome-wide cell-free DNA methylation analysis identifies an episignature associated with metastatic luminal B breast cancer. *Front. Cell Dev. Biol.* 10:1016955. doi: 10.3389/fcell.2022.1016955

A genome-wide cell-free DNA methylation analysis identifies an episignature associated with metastatic luminal B breast cancer

Aitor Rodriguez-Casanova[1,2,3†], Nicolas Costa-Fraga[1,3†], Clara Castro-Carballeira[4], Miriam González-Conde[2,5], Carmen Abuin[2], Aida Bao-Caamano[1,3], Tomás García-Caballero[6], Elena Brozos-Vazquez[7], Carmela Rodriguez-López[7], Victor Cebey[7], Patricia Palacios[7], Juan F. Cueva[5,7], Rafael López-López[2,5,7*], Clotilde Costa[2,5‡] and Angel Díaz-Lagares[1,5*‡]

[1]Epigenomics Unit, Cancer Epigenomics, Translational Medical Oncology Group (ONCOMET), Health Research Institute of Santiago de Compostela (IDIS), University Clinical Hospital of Santiago (CHUS/SERGAS), Santiago de Compostela, Spain, [2]Roche-Chus Joint Unit, Translational Medical Oncology Group (ONCOMET), Health Research Institute of Santiago (IDIS), Santiago de Compostela, Spain, [3]Universidade de Santiago de Compostela (USC), Santiago de Compostela, Spain, [4]Department of Oncology, Marqués de Valdecilla University Hospital, Santander, Spain, [5]Centro de Investigación Biomédica en Red Cáncer (CIBERONC), ISCIII, Madrid, Spain, [6]Department of Morphological Sciences, University of Santiago de Compostela and Xerencia de Xestión Integrada de Santiago (XXIS/SERGAS), Santiago de Compostela, Spain, [7]Translational Medical Oncology Group (ONCOMET), Health Research Institute of Santiago de Compostela (IDIS), University Clinical Hospital of Santiago (CHUS/SERGAS), Santiago de Compostela, Spain

Breast cancers of the luminal B subtype are frequent tumors with high proliferation and poor prognosis. Epigenetic alterations have been found in breast tumors and in biological fluids. We aimed to profile the cell-free DNA (cfDNA) methylome of metastatic luminal B breast cancer (LBBC) patients using an epigenomic approach to discover potential noninvasive biomarkers. Plasma cfDNA was analyzed using the Infinium MethylationEpic array in a cohort of 14 women, including metastatic LBBC patients and nontumor controls. The methylation levels of cfDNA and tissue samples were validated with droplet digital PCR. The methylation and gene expression data of 582 primary luminal breast tumors and 79 nontumor tissues were obtained from The Cancer Genome Atlas (TCGA). We found an episignature of 1,467 differentially methylated CpGs that clearly identified patients with LBBC. Among the genes identified, the promoter hypermethylation of *WNT1* was validated in cfDNA, showing an area under the ROC curve (AUC) of 0.86 for the noninvasive detection of metastatic LBBC. Both paired cfDNA and primary/metastatic breast tumor samples showed hypermethylation of *WNT1*. TCGA analysis revealed significant *WNT1* hypermethylation in the primary tumors of luminal breast cancer patients, with a negative association between *WNT1* methylation and gene expression. In this proof-of-principle study, we discovered an episignature associated with metastatic LBBC using a genome-wide cfDNA methylation

approach. We also identified the promoter hypermethylation of *WNT1* in cfDNA as a potential noninvasive biomarker for luminal breast cancer. Our results support the use of EPIC arrays to identify new epigenetic noninvasive biomarkers in breast cancer.

KEYWORDS

DNA methylation, EPIC Array, cell-free DNA, liquid biopsy, metastasis, luminal B, breast cancer

Introduction

Breast cancer (BC) is the most frequently diagnosed cancer in women worldwide, with 2.3 million new cases (11.7% of all cancer cases) in 2020, representing the leading cause of cancer death in women (Sung et al., 2021). BC is a heterogeneous disease with several distinct clinical characteristics that, according to a gene expression profile, can be divided into four molecular subtypes: luminal, HER2-enriched, basal-like, and normal breast-like (Perou et al., 2000). In addition, luminal tumors can be divided into the luminal A and B subtypes according to the expression profile of the estrogen receptor (ER), progesterone (PR), HER2, and proliferation tumor status (Cheang et al., 2009). The luminal B subtype is a common BC subtype characterized by high proliferation, resistance to standard therapies, risk of early relapse, and poor prognosis (Tran and Bedard, 2011; Cornen et al., 2014; Li et al., 2016). In addition, this tumor subtype is more likely to exhibit local recurrence and single bone metastases than nonluminal BC. However, recent studies have not investigated this tumor subtype as thoroughly as other subtypes (Li et al., 2016). Notably, the incidence of luminal B tumors has increased in recent years in many racial/ethnic and age groups (Acheampong et al., 2020).

Cancer metastasis is characterized by highly variable clinical manifestations and is responsible for over 90% of cancer-related deaths (Gupta and Massague, 2006; Chaffer and Weinberg, 2011). However, despite recent advances, the clinical need to identify biomarkers in metastatic BC disease remains unmet (Gupta et al., 2020). In recent years, liquid biopsy has emerged as a good opportunity to address this clinical need. This noninvasive approach allows for the characterization of the molecular landscape of circulating tumor elements in body fluids, such as epigenetic modifications of cell-free DNA (cfDNA), to obtain biomarkers for the management of cancer patients (Siravegna et al., 2017).

The most well-known epigenetic modification is DNA methylation, which is an important regulator of gene expression originating from the addition of a methyl group (CH_3) to the 5' carbon of cytosines in cytosine–phosphate–guanine (CpG) dinucleotides (Bao-Caamano et al., 2020). The deregulation of this epigenetic mechanism in breast tumor cells has major implications for cancer development, progression, and therapy response (Gomez-Miragaya et al., 2019; Martin-Pardillos et al., 2019; Palomeras et al., 2019; Pineda et al., 2019). Notably, the analysis of DNA methylation in liquid biopsy has shown utility as a potential clinical biomarker for BC patients (Palanca-Ballester et al., 2021).

Recently published studies of other tumor types have shown that epigenomic approaches based on the Infinium MethylationEPIC array (EPIC array) technology, which covers over 850,000 CpG sites along the human genome, could be useful to profile the methylation of cfDNA in biological fluids (Gallardo-Gomez et al., 2018; Herrgott et al., 2022). Therefore, this proof-of-principle study aimed to profile the cfDNA methylome of luminal B breast cancer (LBBC) patients using an EPIC array approach to discover new noninvasive biomarkers. In this study, we identified an epigenetic signature (episignature) based on the methylation of cfDNA associated with metastatic LBBC. Among the genes of this episignature, we confirmed the hypermethylation of *WNT1* in cfDNA and tumor tissues (primary and metastatic) as a potential new biomarker for LBBC patients. The results of our work support the application of the EPIC array technology as a noninvasive tool to identify new biomarkers in breast cancer.

Methods

Study participants

In this retrospective study, 9 luminal B metastatic breast cancer patients and 5 healthy controls (nontumor controls) were recruited between 2016 and 2018 at the Medical Oncology Department at the University Clinical Hospital of Santiago de Compostela (Spain). Most of the metastatic patients of this study (7 out of 9) had been diagnosed in the past at M0 stage. Two patients of our cohort had metastases at the time of primary tumor diagnosis. The study was approved by the Galician Ethical Committee (reference number 2015/772) and conducted in accordance with the guidelines for Good Clinical Practice and the Declaration of Helsinki. All participants included in the study signed the informed consent to participate.

Blood and tissue samples

Blood sample was obtained from all the patients at the time of diagnosis of metastasis and before starting the treatment. Blood samples were collected by phlebotomy into collection tubes

containing K_2EDTA as an anticoagulant. Plasma was isolated within 2 h of collection by initial centrifugation at 1,700 × g for 10 min at room temperature (RT), followed by a second centrifugation at 15,000 × g for 10 min at RT. Isolated plasma was stored at −80°C until analysis. All tumor tissues used were obtained according to standard-of-care (SOC) procedures. We used formalin-fixed and paraffin-embedded (FFPE) primary and/or metastatic tumor and matched nontumor tissue samples available from 4 patients included in the study. Whole slide FFPE tissue sections of 10 µm were obtained.

Isolation of DNA from plasma and tissue samples

We used the QIAamp® Circulating Nucleic Acid Kit (Qiagen) and the vacuum system QIAvac 24 Plus (Qiagen) following the manufacturer´s recommendations to isolate cfDNA from 2 ml of plasma. DNA was also isolated from 10-µm FFPE tissue sections using the AllPrep DNA/RNA FFPE Kit (Qiagen) following the manufacturer's protocol. The quality and quantity of DNA from FFPE tissue sections were evaluated with a NanoDrop (Thermo Fisher), and cfDNA was quantified using the Qubit 1× dsDNA High-Sensitivity Assay Kit and a Qubit 4.0 Fluorometer (Thermo Fisher Scientific). The DNA from FFPE tissue sections and cfDNA from plasma were stored at −80°C until analysis.

Genome-wide cell-free DNA methylation analysis

Fifteen nanograms of each individual sample of plasma cfDNA was bisulfite-converted using the EZ DNA Methylation Lightning Kit (Zymo Research) following the manufacturer's recommendations. Subsequently, the bisulfite-modified cfDNA was then subjected to whole genome amplification (WGA) using the EpiTect Whole Bisulfitome Kit (Qiagen) according to the manufacturer's protocol. Briefly, the bisulfite-modified cfDNA of each individual sample was amplified with a reaction buffer containing REPLI-g Midi DNA Polymerase (Qiagen) at 28°C for 8 h, which was subsequently inactivated at 95 °C for 5 min. After the WGA of cfDNA, the Illumina Infinium HD methylation protocol was followed using MethylationEPIC BeadChips that were analyzed in a HiScan (Illumina). Samples with a mean detection p-value < 0.01 were considered valid for the analysis. The methylation data were processed in the R statistical environment using RnBeads 2.0 (Muller et al., 2019). Raw intensity data files (IDATs) were imported into RnBeads 2.0 for quality control and preprocessing. First, a greedycut algorithm was used to filter out low-quality probes. Probes overlapping with SNPs and probes whose sequences mapped to multiple genomic locations (cross-reactive) were removed. IDATs obtained in the array were normalized using the beta-mixture quantile (BMIQ) method. Hierarchical linear models were used to obtain the methylation differences between groups. p-values were corrected for multiple testing using the Benjamini–Hochberg method, and a false discovery rate (FDR) < 10% was selected for significance. The DNA methylation level was represented as the average β-value, which was calculated as the ratio of the fluorescent signal intensity of the methylated probe to those of total (methylated and unmethylated) probes. Average β-values were used to calculate the mean methylation difference between groups as the Δβ-value (β-value Luminal B – β-value Control). An unsupervised hierarchical clustering heatmap of β-values was generated using the ComplexHeatmap package. Gene ontology (GO) enrichment analysis of biological pathways from the PANTHER database was performed using GENECODIS (Tabas-Madrid et al., 2012).

Methylation and expression analysis from The Cancer Genome Atlas

The DNA methylation (β-values) and expression data of *WNT1* in luminal primary breast tumors and nontumor controls were obtained from the public datasets of The Cancer Genome Atlas (TCGA) Cancer Genome Atlas, N. (2012). The breast cancer subtype of patients was obtained from the clinical information available at TCGA and the classification of these TCGA patients based on PAM50 assay performed by Netanely et al. (2016).

Methylation analysis of the *WNT1* promoter in cell-free DNA by droplet digital PCR

The methylation of the *WNT1* promoter was analyzed by droplet digital PCR (ddPCR) in a QX200 system (Bio-Rad). Twenty nanograms of plasma cfDNA and 30 ng of DNA from FFPE tissue samples were bisulfite converted using the EZ DNA Methylation Lightning Kit (Zymo Research) following the manufacturer's recommendations. A custom Bio-Rad assay to detect the methylation status of *WNT1* (cg27196808) was designed: *WNT1*-M for methylation and *WNT1*-U for unmethylation (Supplementary Table S1). A multiplex preamplification reaction was performed with ~2 ng of bisulfite-converted DNA using SsoAdvanced™ PreAmp Supermix (Bio-Rad), *WNT1*-M, and *WNT1*-U. The PCR conditions were as follows: 3 min at 95°C, 10 cycles of 95°C for 15 s and 50.6°C for 4 min, and a final hold step at 4°C. Next, a multiplex reaction mix was prepared with 2 µL of the preamplification product using ddPCR Supermix for Probes (No dUTP) (Bio-Rad), *WNT1*-M, and *WNT1*-U. The QX200™ Droplet Generator (Bio-Rad) was

used to generate droplets. The thermocycling conditions were as follows: 10 min at 95°C, 40 cycles of 95°C for 15 s and 50.6°C for 30 s, 98°C for 10 min, and a final hold step at 4°C. The temperature ramp increment was 2.5°C/s for all steps. Droplets were counted and analyzed using the QX200™ Droplet Reader system (Bio-Rad), and the QuantaSoft analysis software (Bio-Rad) was used to acquire data. Water was included as a no-template control, and the Human Methylated and Non-Methylated DNA set (Zymo Research) was used as a positive control for methylation and unmethylation. Reactions were performed in triplicate. DNA methylation was expressed according to the following formula: Methylation (%) = [M/(U + M)] x 100, where M represents the copies/µl of methylated cfDNA, and U the copies/µl of unmethylated cfDNA.

Statistical analysis

The Kolmogorov–Smirnov test was used to evaluate the normality of the distribution of the data. The nonparametric Mann–Whitney U test was used to compare methylation data. To assess the diagnostic accuracy, a receiver operating characteristic (ROC) curve was generated. The greatest combination of sensitivity and specificity was obtained using the Youden index (J): J = sensitivity + specificity - 1. The association between DNA methylation and gene expression was evaluated with a Spearman correlation. The GraphPad Prism 6.0 software (GraphPad Software) and the R statistical environment (version 4.2.0) were used for statistical analysis and graphical representation. All expressed *p*-values were calculated with two-tailed tests and were considered significant when the *p*-value < 0.05.

Results

Clinical characteristics of patients

A retrospective cohort of 14 women was included in this study: 9 patients with LBBC at the time of metastatic disease diagnosis and 5 nontumor controls. The mean age of the patients was 66 ± 16 years, whereas the control group had a mean age of 53 ± 10 years. The main clinical characteristics of the analyzed cohort are described in Supplementary Table S2. All patients had distant metastases and invasive ductal carcinoma with a high Ki-67 proliferative index (≥20%), and they were positive for estrogen receptors (ER+). Eight out of nine patients were positive for progesterone receptors (PR+), and two patients had HER2 overexpression (Supplementary Table S2). Six out of the 9 patients (66%) included in the study had lung metastasis, 4 patients (44%) showed bone lesions, and 3 patients (33%) had liver affectation. In addition, 5 of the patients (55%) had multiple metastatic locations.

Genome-wide cell-free DNA methylation analysis of metastatic patients with luminal B breast cancer

The analysis of DNA methylome with the EPIC array methodology usually needs a high amount of DNA, which is difficult to obtain in the clinic from individual plasma samples. As a novelty in our study, to overcome this limitation, we have used small amounts of cfDNA from individual plasma samples, which were genome-wide amplified after bisulfite modification and then analyzed using EPIC arrays. Thus, using this approach we performed a genome-wide cell-free DNA methylation analysis in our cohort of 9 LBBC patients and 5 nontumor controls (Figure 1A). After hybridizing the samples in the EPIC array, 2 LBBC samples showed a mean detection *p*-value > 0.01 and were not considered valid for the analysis. Therefore, we ultimately compared the methylation status of cfDNA in 7 LBBC patients and 5 nontumor controls, leading to 28,799 differentially methylated CpGs (DMCpGs) ($p < 0.05$; FDR < 10%) between LBBC and nontumor controls. These DMCpGs showed a wide distribution throughout all chromosomes of the genome (Figure 1B). Of these DMCpGs, 92% (26,486) were hypomethylated and 8% (2,313) were hypermethylated in LBBC patients with respect to nontumor controls (Figure 1C). Most of the hypomethylated CpGs were distributed in regions with low CpG density (open sea) (Figure 1D) and outside promoter regions (Figure 1E), whereas hypermethylated CpGs were mainly located in CpG islands (CpGIs) (Figure 1F) and promoters (Figure 1G).

Identification of a cell-free DNA episignature in metastatic patients with luminal B breast cancer

The aberrant hypermethylation of CpGI promoters is a very relevant feature that usually occurs in tumor cells (Baylin and Chen, 2005). Therefore, we focused our study on analyzing the methylation profile of cfDNA at the CpGIs of promoters. In these regions of cfDNA, we identified 1,467 DMCpGs ($p < 0.05$; FDR < 10%) with a difference in methylation (Δβ-value) higher than 0.20 (Δβ-value > |0.20|) (Figure 1A). Notably, this epigenetic signature (episignature) of 1,467 DMCpGs was able to clearly differentiate LBBC patients from nontumor controls (Figure 2A). Next, to obtain information related to the functional pathways involved in the identified episignature, we performed a gene ontology (GO) enrichment analysis based on the PANTHER database. This analysis revealed that methylation differences in the cfDNA of LBBC patients and nontumor controls were mainly

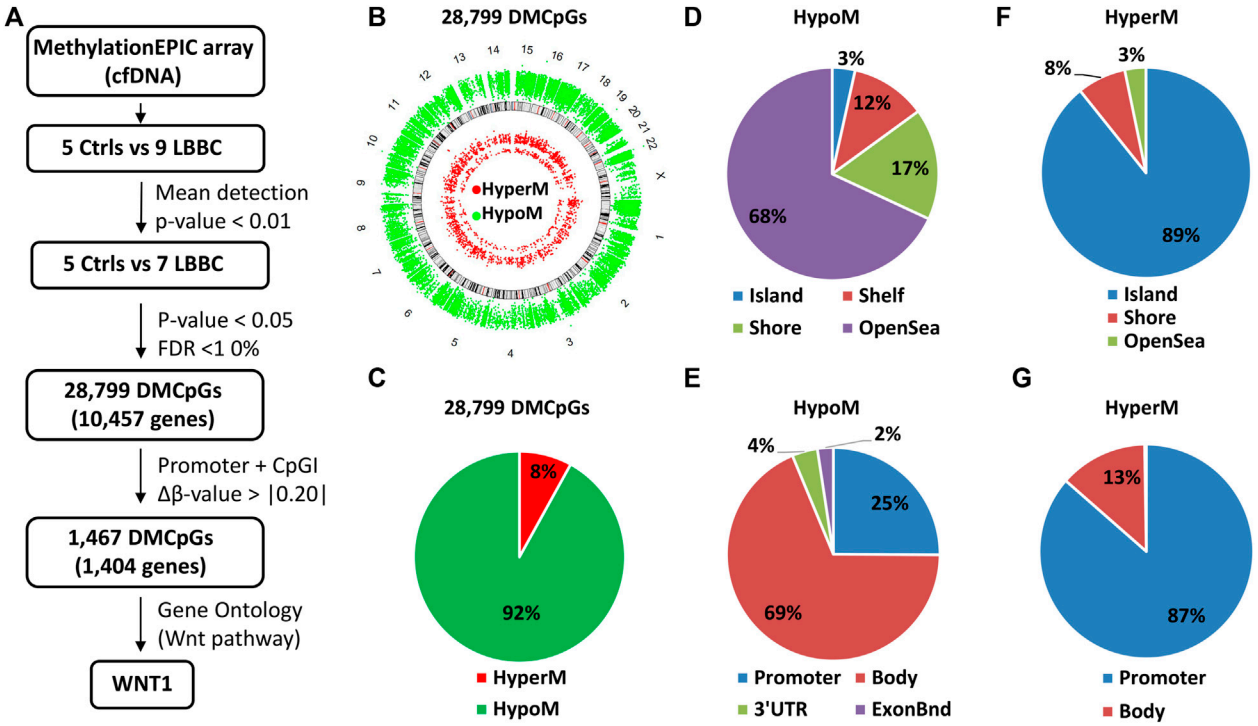


FIGURE 1
Methylation landscape of cell-free DNA in metastatic patients with luminal B breast cancer. **(A)** Flowchart of the cfDNA methylation analysis. Individual cfDNA samples from controls and luminal B breast cancer patients were analyzed with EPIC array. **(B–G)** Description of the 28,799 differentially methylated CpGs (DMCpGs) found in cfDNA of luminal B breast cancer patients according to **(B)** chromosome location, **(C)** methylation status, **(D,F)** CpG context, and **(E,G)** gene location. cfDNA, cell-free DNA; Ctrls, controls; LBBC, luminal B breast cancer; FDR, false discovery rate; HypoM, hypomethylated; HyperM, hypermethylated.

associated with genes related to the Wnt signaling pathway (Figure 2B). Table 1 shows the 34 DMCpGs (corresponding to 24 genes) of the episignature of cfDNA that are associated with the Wnt signaling pathway. Relevantly, the genes of these 34 DMCpGs that are associated with Wnt signaling belonged to a network significantly enriched in protein interactions ($p < 0.001$) according to a STRING analysis (Supplementary Figure S1).

Hypermethylation of the *WNT1* promoter in the cell-free DNA and tumor samples of patients with luminal B breast cancer

Among the DMCpGs in the episignature obtained from the cfDNA of LBBC patients that were associated with the Wnt signaling pathway (Table 1), we found 2 CpGs (cg27196808 and cg02771661) located in *WNT1* that were hypermethylated in the cfDNA of LBBC patients with respect to nontumor controls. To confirm this aberrant methylation, we selected the most DMCpG of *WNT1* (cg27196808), and we analyzed its methylation status in the cfDNA of our cohort using ddPCR. As expected, the methylation of *WNT1* was significantly higher in LBBC patients than in nontumor controls (Figure 3A). In addition, using a ROC curve analysis, the methylation status of the *WNT1* promoter analyzed by ddPCR accurately differentiated LBBC patients from nontumor controls, with an area under the ROC curve (AUC) of 0.86 (95% CI: 0.65–1.00, $p = 0.045$) (Figure 3B), a sensitivity of 78% (CI 95%: 40%–98%), and a specificity of 100% (CI 95%: 40%–100%).

To verify that the hypermethylation of *WNT1* found in cfDNA (cg27196808) was also present in the tumor tissues of patients with LBBC, we assessed its methylation status by ddPCR in the available matched primary and/or metastatic tumor tissue samples ($n = 4$) of our cohort. This assay revealed that *WNT1* hypermethylation was present not only in the cfDNA but also in the paired primary and/or metastatic tumor samples of the LBBC patients analyzed (Figure 3C). Next, we took advantage of public DNA methylation array data from The Cancer Genome Atlas (TCGA) to evaluate whether the hypermethylation of the *WNT1* promoter was a frequent event in LBBC. This analysis showed that the methylation of the *WNT1* promoter (cg27196808) was significantly higher in luminal B primary tumors (stages I–IV) than in nontumor controls (Figure 3D), and this observation was

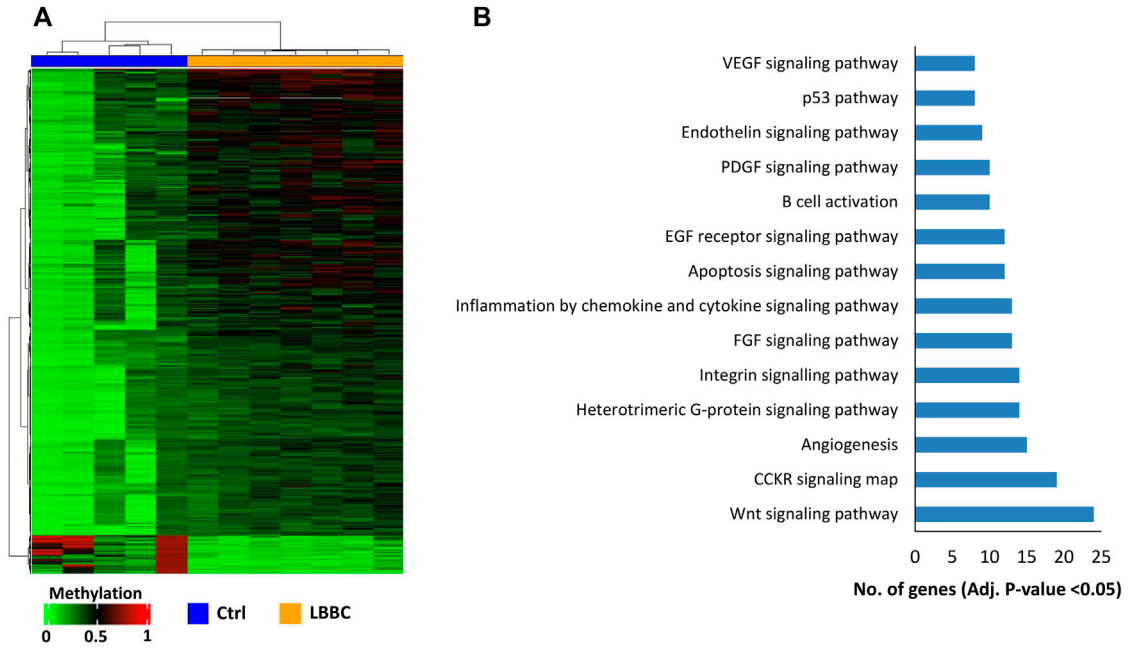


FIGURE 2
Episignature of cell-free DNA in metastatic patients with luminal B breast cancer. **(A)** Unsupervised hierarchical clustering heatmap of the episignature (1,467 DMCpGs) obtained in cfDNA that differentiates LBBC patients (n = 7) from nontumor controls (n = 5). **(B)** Gene Ontology enrichment analysis by the PANTHER database, showing the most representative pathways associated with the episignature of cfDNA in luminal B breast cancer patients. Ctrls, controls; LBBC, luminal B breast cancer; DMCpGs, differentially methylated CpGs; cfDNA, cell-free DNA.

consistent across all TNM tumor stages (Figure 3E). However, the methylation status of *WNT1* did not differ among the tumor stages of LBBC analyzed (Figure 3E). An ROC curve analysis showed that *WNT1* methylation differentiated luminal B primary tumors (stages I-IV) from nontumor controls with high diagnostic accuracy, with an AUC of 0.87 (95% CI: 0.82–0.92, $p < 0.0001$) (Figure 3F). In addition, we also analyzed the *WNT1* expression data (RNAseq) available from LBBC patients and nontumor controls included in TCGA, revealing that *WNT1* was significantly downregulated in this BC subtype (Supplementary Figure S2).

Finally, we also evaluated in breast primary tumors from TCGA whether the hypermethylation of *WNT1* was a specific event of LBBC patients. The methylation levels of *WNT1* were significantly higher in LBBC than in the other breast tumor subtypes (Supplementary Figure S3). Interestingly, we observed that *WNT1* was also significantly hypermethylated in other breast cancer subtypes (Luminal A, triple negative and HER2+) in comparison with nontumor controls.

Discussion

Alterations in epigenetic mechanisms, such as aberrant DNA methylation, are implicated in the development, progression, and therapy response of BC (Palomeras et al., 2019; Pineda et al., 2019; Glodzik et al., 2020). In recent years, the methylation analysis of liquid biopsy samples in BC patients has shown clinical utility as a biomarker for the detection, prognosis, and monitoring of the disease (Hoque et al., 2006; Mastoraki et al., 2018; Palanca-Ballester et al., 2021). However, a clinical need to find new noninvasive biomarkers associated with metastatic BC subtypes persists (Gupta et al., 2020). Herein, we focused our study on patients with advanced LBBC, since it represents a frequent, aggressive and poor prognosis BC subtype (Creighton, 2012). The characterization of liquid biopsy samples using epigenomic tools for genome-wide methylation analyses has been recently proposed as a good approach to discover new noninvasive biomarkers (Shen et al., 2018). Thus, we used a genome-wide DNA methylation approach based on the EPIC array methodology to profile the methylome of cfDNA and discover novel noninvasive biomarkers in metastatic LBBC patients.

Our work revealed that the cfDNA of patients with metastatic LBBC is characterized by the hypomethylation of regions with a low density of CpGs and the site-specific hypermethylation of CpGIs in promoter regions. Importantly, this pattern is similar to the deregulation of DNA methylation that has been previously described in cancer cells (Nishiyama and Nakanishi, 2021), suggesting that the methylation profile in the cfDNA of the

TABLE 1 The 34 CpGs of cfDNA episignature found in metastatic patients with luminal B breast cancer associated with the Wnt signaling pathway.

TargetID[a]	Chr[b]	Position	Gene name	Gene region	Δβ[c]	*p*-value
cg26821418	9	2016890	*SMARCA2*	5′UTR; 5′UTR; 5′UTR	0.35	0.0028
cg27201625	10	6622279	*PRKCQ*	TSS200	0.33	0.0008
cg04351665	10	6622297	*PRKCQ*	TSS200	0.21	0.0057
cg03306374	16	23847325	*PRKCB*	1stExon; 5′UTR; 5′UTR	0.28	0.0024
				1stExon		
cg06931245	8	28351501	*FZD3*	TSS1500	0.26	0.0023
cg18463655	8	28351544	*FZD3*	TSS200; TSS200	0.23	0.0041
cg26631144	8	30670260	*PPP2CB*	5′UTR; 1stExon; 5′UTR	0.28	0.0014
				1stExon		
cg02478409	6	33589019	*ITPR3*	TSS200	0.32	0.0020
cg16490096	1	40367661	*MYCL1*	1stExon; 5′UTR; 5′UTR	0.23	0.0024
				1stExon; 5′UTR; 1stExon		
cg20462899	1	40367831	*MYCL1*	TSS200; TSS200; TSS200	0.29	0.0056
cg02771661	**12**	**49372162**	***WNT1***	**TSS200**	**0.21**	**0.0047**
cg27196808	**12**	**49372281**	***WNT1***	**1stExon;5′UTR**	**0.22**	**0.0066**
cg13469346	3	53195186	*PRKCD*	TSS200; TSS200	0.24	0.0026
cg21950287	19	54385441	*PRKCG*	TSS200	0.23	0.0051
cg13885159	11	62473858	*GNG3*	TSS1500	0.23	0.0064
cg03922588	11	62473871	*GNG3*	TSS1500	0.27	0.0021
cg25220961	17	64298782	*PRKCA*	TSS200	0.27	0.0023
cg11676500	17	64298789	*PRKCA*	TSS200	0.23	0.0026
cg08221093	16	68119222	*NFATC3*	TSS200; TSS200; TSS200	0.25	0.0067
				TSS200		
cg21367137	16	68119381	*NFATC3*	5′UTR; 5′UTR; 5′UTR	0.26	0.0046
				1stExon; 1stExon; 1stExon		
cg22722737	9	82187628	*TLE4*	5′UTR; 5′UTR; 5′UTR; 5′UT; 1stExon; 1stExon; 1stExon; 1stExon; 1stExo; 5′UTR	0.22	0.0050
cg26753733	4	102268824	*PPP3CA*	TSS200; TSS200; TSS200	0.24	0.0039
cg08764167	10	103113933	*BTRC*	5′UTR; 5′UTR; 1stExon	0.21	0.0050
				1stExon		
cg20359285	2	119603969	*EN1*	1stExon	0.21	0.0043
cg00557469	5	133562427	*PPP2CA*	TSS1500	0.25	0.0049
cg18671773	5	141016477	*HDAC3*	TSS200	0.27	0.0060
cg16248329	4	187644739	*FAT1*	5′UTR	0.23	0.0026
cg02968914	19	1955395	*CSNK1G2*	5′UTR	−0.32	0.0042
cg01895482	19	2556145	*GNG7*	5′UTR	−0.35	0.0015
cg07223632	22	46930499	*CELSR1*	1stExon	−0.36	0.0028
cg00875636	22	46931138	*CELSR1*	1stExon	−0.28	0.0030
cg27334938	18	77167042	*NFATC1*	5′UTR	−0.30	0.0051
cg02113385	18	77203443	*NFATC1*	5′UTR	−0.35	0.0060
cg27475132	4	187645120	*FAT1*	TSS200	−0.23	0.0053

[a]CpGs located in CGIs of promoters (TSS1500, TSS200, 5′UTR, 1st exon).
[b]Chromosome.
[c]Δβ-values (β-value Luminal B - β-value Control). CpGs of the gene *WNT1* are indicated in bold.

patients in our cohort mirrors the epigenetic alterations of BC cells.

Specific genes, such as *RASSF1A* and *BRCA1*, have been previously described to exhibit aberrant hypermethylation of their promoter CpGIs in BC (Rice et al., 1998; Dammann et al., 2001). Accordingly, we focused our study on promoter CpGIs and were able to identify a novel noninvasive episignature in cfDNA based on 1,467 CpGs that was associated with LBBC patients. We found that the genes of this episignature were related to relevant biological pathways, mainly Wnt signalling

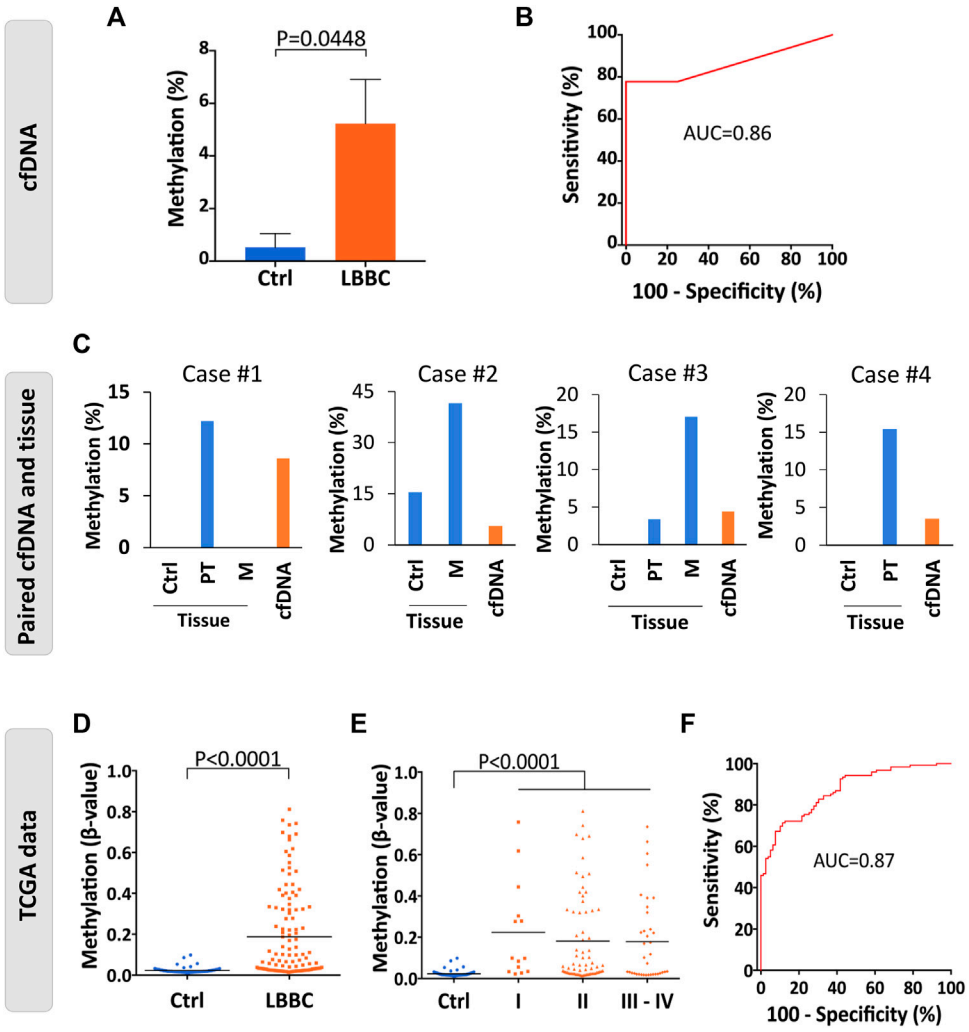


FIGURE 3
Methylation status of the *WNT1* promoter in the cfDNA and tumor samples of patients with luminal B breast cancer. **(A)** Validation of the methylation levels of the *WNT1* promoter (cg27196808) in cell-free DNA (cfDNA) of luminal B breast cancer patients (n = 9) and nontumor controls (n = 4) analyzed by droplet digital PCR (ddPCR). Methylation levels are represented as the mean ± SEM. **(B)** Diagnostic accuracy of the methylation of the *WNT1* promoter using droplet digital PCR (ddPCR) in cfDNA for the detection of metastatic luminal B breast cancer patients (n = 9) with respect to nontumor controls (n = 4). **(C)** Methylation levels of *WNT1* in cfDNA and paired breast primary and/or metastatic tumor samples of 4 luminal B breast cancer patients analyzed by droplet digital PCR (ddPCR). Nontumor tissues from the same patients were used as controls. **(D–E)** Methylation status of *WNT1* in primary tumors of luminal B breast cancer patients (n = 122) and nontumor controls (n = 79) analyzed by 450K methylation array and obtained from TCGA considering **(D)** all TNM stages together (stages I–IV, n = 122) or **(E)** separated according to TNM stage (I, n = 14; II, n = 70; III-IV, n = 37). The horizontal line represents the mean methylation values. **(F)** ROC curve evaluating the methylation of *WNT1* to detect primary tumors of luminal B breast cancer patients (stages I-IV, n = 122) with respect to nontumor controls (n = 79) from TCGA. Ctrl, control; P, p-value; AUC, area under the ROC curve; ROC curve, receiver operating characteristic curve. PT, primary tumor; M, metastasis.

pathway. Among these genes, we focused on *WNT1*, which is involved in the canonical Wnt signaling pathway (also known as Wnt/β-Catenin) in cancer cells (Ayyanan et al., 2006; Mehta et al., 2021). Thus, we confirmed that the promoter CpGI of *WNT1* was hypermethylated in the cfDNA of patients with metastatic LBBC and that this aberrant methylation showed a high diagnostic accuracy to detect this BC subtype, suggesting that the hypermethylation of *WNT1* could be a suitable biomarker for cancer detection and monitoring of metastatic patients. In line with this, it has been recently shown that methylation biomarkers of cfDNA with high diagnostic accuracy are useful not only for diagnosis but also for monitoring tumor burden dynamics under different therapeutic regimens in advanced disease (Barault et al., 2018). Importantly, evaluating prognosis and monitoring tumor response in real time during treatment continues to be an unmet clinical need in advanced BC (Gupta et al., 2020). Wnt signaling is a very relevant pathway in BC whose molecular alterations have clinical implications to establish the prognosis of the disease

(Mukherjee et al., 2012) and has been associated to breast cancer therapy response (Abreu de Oliveira et al., 2022). Therefore, it would be interesting to evaluate in future studies whether the hypermethylation of *WNT1* could be useful for the selection of patients susceptible to systemic therapies (CDK inhibitors for example) in the BC metastatic setting.

Of note, we also found that the promoter hypermethylation of *WNT1* was present not only in cfDNA of LBBC patients but also in their primary and/or metastatic tumors. This finding is in accordance with our previous work and that of other authors showing that the molecular landscape present in liquid biopsy may also be detected in the corresponding tumor tissue of patients (Rahvar et al., 2020; Rodriguez-Casanova et al., 2021). In addition, when we extended our study to the analysis of breast primary tumors using the public TCGA database, we were able to confirm that the hypermethylation of *WNT1* is a frequent event in early and advanced LBBC, suggesting that the epigenetic deregulation of *WNT1* is not a specific biomarker of metastatic disease but rather a biomarker of breast cancer cells. In agreement with our results, the aberrant methylation of other genes involved in the Wnt signaling pathway (e.g., *WNT5A* and *WNT7A*) has previously been described in tumor cells from the BC luminal subtype (Shan et al., 2019) and in other tumor types, such as gastric cancer or chronic lymphocytic leukemia (Liu et al., 2019; Poppova et al., 2022).

Several studies have shown that aberrant promoter hypermethylation is a relevant mechanism that is able to repress the expression of key genes in breast tumor cells (Alvarez et al., 2013; Palomeras et al., 2019). Accordingly, the analysis of luminal breast tumors from the TCGA database also revealed that *WNT1* promoter hypermethylation was associated with a downregulation of its gene expression in primary tumors, suggesting that *WNT1* is epigenetically regulated in luminal BC. The downregulation of *WNT1* observed in particular BC subtypes corroborates the work by Koval and Katanaev (2018), who reported low expression of this gene in primary tumors of nontriple-negative BC patients (ER+/PR + and HER+). Indeed, it has been reported that the deregulation of some Wnt signaling components depends on the BC subtype, with many being downregulated in the luminal B subtype (Smid et al., 2008).

One limitation of our study is that the epigenomic profiling of cfDNA performed is based on a retrospective cohort of patients with a small sample size. Although the results obtained in this work should be taken with caution, they provide the basis for further large, prospective and independent studies that validate the clinical utility of the potential epigenetic biomarkers identified herein. In addition, it would be interesting to evaluate in future works the implications of the epigenetic deregulation of *WNT1* in the development of metastasis.

In summary, in this proof-of-principle study, we discovered an episignature associated with patients with advanced LBBC using a genome-wide cfDNA methylation approach. We also identified the promoter hypermethylation of *WNT1* in cfDNA as a potential noninvasive biomarker for luminal BC. Our results support the use of EPIC array technology to identify new noninvasive biomarkers in BC.

Data availability statement

The methylation data of cfDNA analyzed in this study with EPIC Array are available at the NCBI GEO repository with accession number GSE214344.

Ethics statement

The studies involving human participants were reviewed and approved by Galician Ethical Committee (reference number 2015/772). The patients/participants provided their written informed consent to participate in this study.

Author contributions

Conceptualization, CC, RL-L, and AD-L; resources, TG-C, EB-V, CR-L, VC, PP, JFC, and RL-L; Data curation, AR-C, NC-F, and CA; methodology, AR-C, NC-F, CC-C, MG-C, CA, AB-C, and TG-C; formal analysis, AR-C, NC-F, CC, and AD-L; writing—original draft preparation, AR-C, NC-F, CC, and AD-L; writing—review and editing, AR-C, NC-F, JFC, RL-L, CC, and AD-L; supervision, CC, RL-L, and AD-L. All authors have read and agreed to the published version of the manuscript.

Funding

This research was funded by "Beca FSEOM/FECMA para Proyectos de Investigación en Cáncer de Mama 2017", the Roche-Chus Joint Unit (IN853B 2018/03), "Axencia Galega de Innovación (GAIN), Vicepresidencia Segunda e Consellería de Economía, Empresa e Innovación". AR-C is supported by the Roche-Chus Joint Unit (IN853B 2018/03) funded by GAIN, "Axencia Galega de Innovación (GAIN), Vicepresidencia Segunda e Consellería de Economía, Empresa e Innovación". NC-F is funded by a predoctoral contract from "Xunta de Galicia" (IN606A-2020/004). AB-C is funded by a predoctoral contract PFIS (FI19/00240) from "Instituto de Salud Carlos III" (ISCIII) co-funded by "Fondo Social Europeo" (FSE). MG-C is supported by "Instituto de Salud Carlos III" (ISCIII), with the grant FI19/00140, co-financed by the European Regional Development Fund (FEDER). CC is supported by AECC (INVES211437COST). AD-L is funded by a contract "Juan Rodés" (JR17/00016) from "Instituto de Salud Carlos III" (ISCIII).

Acknowledgments

We would like to thank all the patients and healthy volunteers included in this study for their kind collaboration, and all personnel who provided technical support with sample collection. We also thank María Amalia Jácome, from MODES

Group (CITIC, Faculty of Science, Universidade da Coruña), for the support with DNA methylome analysis.

Conflict of interest

RL-L has received honoraria for participation in Advisory Boards from Roche, AstraZeneca, Merck, Merck Sharp and; Dohme, Bayer, Bristol-Myers Squibb, Novartis, Janssen, Lilly, Pfizer, and Leo; travel, accommodations, and expenses from PharmaMar, Roche, Bristol-Myers Squibb, and Pierre Fabre; research funding from Roche and Merck; and is co-founder and shareholder in Nasasbiotech, S.L., Mtrap Inc.

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

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcell.2022.1016955/full#supplementary-material

References

Abreu de Oliveira, W. A., El Laithy, Y., Bruna, A., Annibali, D., and Lluis, F. (2022). Wnt signaling in the breast: From development to disease. *Front. Cell Dev. Biol.* 10, 884467. doi:10.3389/fcell.2022.884467

Acheampong, T., Kehm, R. D., Terry, M. B., Argov, E. L., and Tehranifar, P. (2020). Incidence trends of breast cancer molecular subtypes by age and race/ethnicity in the US from 2010 to 2016. *JAMA Netw. Open* 3 (8), e2013226. doi:10.1001/jamanetworkopen.2020.13226

Alvarez, C., Tapia, T., Cornejo, V., Fernandez, W., Munoz, A., Camus, M., et al. (2013). Silencing of tumor suppressor genes RASSF1A, SLIT2, and WIF1 by promoter hypermethylation in hereditary breast cancer. *Mol. Carcinog.* 52 (6), 475–487. doi:10.1002/mc.21881

Ayyanan, A., Civenni, G., Ciarloni, L., Morel, C., Mueller, N., Lefort, K., et al. (2006). Increased Wnt signaling triggers oncogenic conversion of human breast epithelial cells by a Notch-dependent mechanism. *Proc. Natl. Acad. Sci. U. S. A.* 103 (10), 3799–3804. doi:10.1073/pnas.0600065103

Bao-Caamano, A., Rodriguez-Casanova, A., and Diaz-Lagares, A. (2020). Epigenetics of circulating tumor cells in breast cancer. *Adv. Exp. Med. Biol.* 1220, 117–134. doi:10.1007/978-3-030-35805-1_8

Barault, L., Amatu, A., Siravegna, G., Ponzetti, A., Moran, S., Cassingena, A., et al. (2018). Discovery of methylated circulating DNA biomarkers for comprehensive non-invasive monitoring of treatment response in metastatic colorectal cancer. *Gut* 67 (11), 1995–2005. doi:10.1136/gutjnl-2016-313372

Baylin, S. B., and Chen, W. Y. (2005). Aberrant gene silencing in tumor progression: Implications for control of cancer. *Cold Spring Harb. Symp. Quant. Biol.* 70, 427–433. doi:10.1101/sqb.2005.70.010

Cancer Genome Atlas, N. (2012). Comprehensive molecular portraits of human breast tumours. *Nature* 490 (7418), 61–70. doi:10.1038/nature11412

Chaffer, C. L., and Weinberg, R. A. (2011). A perspective on cancer cell metastasis. *Science* 331 (6024), 1559–1564. doi:10.1126/science.1203543

Cheang, M. C., Chia, S. K., Voduc, D., Gao, D., Leung, S., Snider, J., et al. (2009). Ki67 index, HER2 status, and prognosis of patients with luminal B breast cancer. *J. Natl. Cancer Inst.* 101 (10), 736–750. doi:10.1093/jnci/djp082

Cornen, S., Guille, A., Adelaide, J., Addou-Klouche, L., Finetti, P., Saade, M. R., et al. (2014). Candidate luminal B breast cancer genes identified by genome, gene expression and DNA methylation profiling. *PLoS One* 9 (1), e81843. doi:10.1371/journal.pone.0081843

Creighton, C. J. (2012). The molecular profile of luminal B breast cancer. *Biologics* 6, 289–297. doi:10.2147/BTT.S29923

Dammann, R., Yang, G., and Pfeifer, G. P. (2001). Hypermethylation of the cpG island of Ras association domain family 1A (RASSF1A), a putative tumor suppressor gene from the 3p21.3 locus, occurs in a large percentage of human breast cancers. *Cancer Res.* 61 (7), 3105–3109.

Gallardo-Gomez, M., Moran, S., Paez de la Cadena, M., Martinez-Zorzano, V. S., Rodriguez-Berrocal, F. J., Rodriguez-Girondo, M., et al. (2018). A new approach to epigenome-wide discovery of non-invasive methylation biomarkers for colorectal cancer screening in circulating cell-free DNA using pooled samples. *Clin. Epigenetics* 10, 53. doi:10.1186/s13148-018-0487-y

Glodzik, D., Bosch, A., Hartman, J., Aine, M., Vallon-Christersson, J., Reutersward, C., et al. (2020). Comprehensive molecular comparison of BRCA1 hypermethylated and BRCA1 mutated triple negative breast cancers. *Nat. Commun.* 11 (1), 3747. doi:10.1038/s41467-020-17537-2

Gomez-Miragaya, J., Moran, S., Calleja-Cervantes, M. E., Collado-Sole, A., Pare, L., Gomez, A., et al. (2019). The altered transcriptome and DNA methylation profiles of docetaxel resistance in breast cancer PDX models. *Mol. Cancer Res.* 17 (10), 2063–2076. doi:10.1158/1541-7786.MCR-19-0040

Gupta, G., Lee, C. D., Guye, M. L., Van Sciver, R. E., Lee, M. P., Lafever, A. C., et al. (2020). Unmet clinical need: Developing prognostic biomarkers and precision medicine to forecast early tumor relapse, detect chemo-resistance and improve overall survival in high-risk breast cancer. *Ann. Breast Cancer Ther.* 4 (1), 48–57. doi:10.36959/739/525

Gupta, G. P., and Massague, J. (2006). Cancer metastasis: Building a framework. *Cell* 127 (4), 679–695. doi:10.1016/j.cell.2006.11.001

Herrgott, G. A., Asmaro, K. P., Wells, M., Sabedot, T. S., Malta, T. M., Mosella, M. S., et al. (2022). Detection of tumor-specific DNA methylation markers in the blood of patients with pituitary neuroendocrine tumors. *Neuro. Oncol.* 24 (7), 1126–1139. doi:10.1093/neuonc/noac050

Hoque, M. O., Feng, Q., Toure, P., Dem, A., Critchlow, C. W., Hawes, S. E., et al. (2006). Detection of aberrant methylation of four genes in plasma DNA for the detection of breast cancer. *J. Clin. Oncol.* 24 (26), 4262–4269. doi:10.1200/JCO.2005.01.3516

Koval, A., and Katanaev, V. L. (2018). Dramatic dysbalancing of the Wnt pathway in breast cancers. *Sci. Rep.* 8 (1), 7329. doi:10.1038/s41598-018-25672-6

Li, Z. H., Hu, P. H., Tu, J. H., and Yu, N. S. (2016). Luminal B breast cancer: Patterns of recurrence and clinical outcome. *Oncotarget* 7 (40), 65024–65033. doi:10.18632/oncotarget.11344

Liu, Y., Qiao, Y., Zhang, H., Li, W., and Zheng, J. (2019). Wnt7a, frequently silenced by CpG methylation, inhibits tumor growth and metastasis via suppressing epithelial-mesenchymal transition in gastric cancer. *J. Cell. Biochem.* 120 (10), 18142–18151. doi:10.1002/jcb.29118

Martin-Pardillos, A., Valls Chiva, A., Bande Vargas, G., Hurtado Blanco, P., Pineiro Cid, R., Guijarro, P. J., et al. (2019). The role of clonal communication and heterogeneity in breast cancer. *BMC Cancer* 19 (1), 666. doi:10.1186/s12885-019-5883-y

Mastoraki, S., Strati, A., Tzanikou, E., Chimonidou, M., Politaki, E., Voutsina, A., et al. (2018). ESR1 methylation: A liquid biopsy-based epigenetic assay for the follow-up of patients with metastatic breast cancer receiving endocrine treatment. *Clin. Cancer Res.* 24 (6), 1500–1510. doi:10.1158/1078-0432.CCR-17-1181

Mehta, S., Hingole, S., and Chaudhary, V. (2021). The emerging mechanisms of Wnt secretion and signaling in development. *Front. Cell Dev. Biol.* 9, 714746. doi:10.3389/fcell.2021.714746

Mukherjee, N., Bhattacharya, N., Alam, N., Roy, A., Roychoudhury, S., and Panda, C. K. (2012). Subtype-specific alterations of the Wnt signaling pathway in

breast cancer: Clinical and prognostic significance. *Cancer Sci.* 103 (2), 210–220. doi:10.1111/j.1349-7006.2011.02131.x

Muller, F., Scherer, M., Assenov, Y., Lutsik, P., Walter, J., Lengauer, T., et al. (2019). RnBeads 2.0: Comprehensive analysis of DNA methylation data. *Genome Biol.* 20 (1), 55. doi:10.1186/s13059-019-1664-9

Netanely, D., Avraham, A., Ben-Baruch, A., Evron, E., and Shamir, R. (2016). Expression and methylation patterns partition luminal-A breast tumors into distinct prognostic subgroups. *Breast Cancer Res.* 18 (1), 74. doi:10.1186/s13058-016-0724-2

Nishiyama, A., and Nakanishi, M. (2021). Navigating the DNA methylation landscape of cancer. *Trends Genet.* 37 (11), 1012–1027. doi:10.1016/j.tig.2021.05.002

Palanca-Ballester, C., Rodriguez-Casanova, A., Torres, S., Calabuig-Farinas, S., Exposito, F., Serrano, D., et al. (2021). Cancer epigenetic biomarkers in liquid biopsy for high incidence malignancies. *Cancers (Basel)* 13 (12), 3016. doi:10.3390/cancers13123016

Palomeras, S., Diaz-Lagares, A., Vinas, G., Setien, F., Ferreira, H. J., Oliveras, G., et al. (2019). Epigenetic silencing of TGFBI confers resistance to trastuzumab in human breast cancer. *Breast Cancer Res.* 21 (1), 79. doi:10.1186/s13058-019-1160-x

Perou, C. M., Sorlie, T., Eisen, M. B., van de Rijn, M., Jeffrey, S. S., Rees, C. A., et al. (2000). Molecular portraits of human breast tumours. *Nature* 406 (6797), 747–752. doi:10.1038/35021093

Pineda, B., Diaz-Lagares, A., Perez-Fidalgo, J. A., Burgues, O., Gonzalez-Barrallo, I., Crujeiras, A. B., et al. (2019). A two-gene epigenetic signature for the prediction of response to neoadjuvant chemotherapy in triple-negative breast cancer patients. *Clin. Epigenetics* 11 (1), 33. doi:10.1186/s13148-019-0626-0

Poppova, L., Pavlova, S., Gonzalez, B., Kotaskova, J., Plevova, K., Dumbovic, G., et al. (2022). Memory B-cell like chronic lymphocytic leukaemia is associated with specific methylation profile of WNT5A promoter and undetectable expression of WNT5A gene. *Epigenetics*, 1–8. doi:10.1080/15592294.2022.2050004

Rahvar, F., Salimi, M., and Mozdarani, H. (2020). Plasma GBP2 promoter methylation is associated with advanced stages in breast cancer. *Genet. Mol. Biol.* 43 (4), e20190230. doi:10.1590/1678-4685-GMB-2019-0230

Rice, J. C., Massey-Brown, K. S., and Futscher, B. W. (1998). Aberrant methylation of the BRCA1 CpG island promoter is associated with decreased BRCA1 mRNA in sporadic breast cancer cells. *Oncogene* 17 (14), 1807–1812. doi:10.1038/sj.onc.1202086

Rodriguez-Casanova, A., Bao-Caamano, A., Lago-Leston, R. M., Brozos-Vazquez, E., Costa-Fraga, N., Ferreiros-Vidal, I., et al. (2021). Evaluation of a targeted next-generation sequencing panel for the non-invasive detection of variants in circulating DNA of colorectal cancer. *J. Clin. Med.* 10 (19), 4487. doi:10.3390/jcm10194487

Shan, M., Zhang, L., Liu, Y., Gao, C., Kang, W., Yang, W., et al. (2019). DNA methylation profiles and their diagnostic utility in BC. *Dis. Markers* 2019, 6328503. doi:10.1155/2019/6328503

Shen, S. Y., Singhania, R., Fehringer, G., Chakravarthy, A., Roehrl, M. H. A., Chadwick, D., et al. (2018). Sensitive tumour detection and classification using plasma cell-free DNA methylomes. *Nature* 563 (7732), 579–583. doi:10.1038/s41586-018-0703-0

Siravegna, G., Marsoni, S., Siena, S., and Bardelli, A. (2017). Integrating liquid biopsies into the management of cancer. *Nat. Rev. Clin. Oncol.* 14 (9), 531–548. doi:10.1038/nrclinonc.2017.14

Smid, M., Wang, Y., Zhang, Y., Sieuwerts, A. M., Yu, J., Klijn, J. G., et al. (2008). Subtypes of breast cancer show preferential site of relapse. *Cancer Res.* 68 (9), 3108–3114. doi:10.1158/0008-5472.CAN-07-5644

Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., et al. (2021). Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *Ca. Cancer J. Clin.* 71 (3), 209–249. doi:10.3322/caac.21660

Tabas-Madrid, D., Nogales-Cadenas, R., and Pascual-Montano, A. (2012). GeneCodis3: A non-redundant and modular enrichment analysis tool for functional genomics. *Nucleic Acids Res.* 40, W478–W483. doi:10.1093/nar/gks402

Tran, B., and Bedard, P. L. (2011). Luminal-B breast cancer and novel therapeutic targets. *Breast Cancer Res.* 13 (6), 221. doi:10.1186/bcr2904

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 04 January 2023
DOI 10.3389/fgene.2022.1028711

OPEN ACCESS

EDITED BY
Xiao Wang,
Konge Larsen ApS, Denmark

REVIEWED BY
Apratim Mitra,
Eunice Kennedy Shriver National Institute of Child Health and Human Development (NIH), United States
Xiangdong Ding,
China Agricultural University, China

*CORRESPONDENCE
Guoqing Tang,
tyq003@163.com

†These authors contributed equally to this work

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

RECEIVED 26 August 2022
ACCEPTED 14 December 2022
PUBLISHED 04 January 2023

CITATION
Wang K, Wang S, Ji X, Chen D, Shen Q, Yu Y, Wu P, Li X and Tang G (2023), Epigenome-wide association studies of meat traits in Chinese Yorkshire pigs highlights several DNA methylation loci and genes.
Front. Genet. 13:1028711.
doi: 10.3389/fgene.2022.1028711

Epigenome-wide association studies of meat traits in Chinese Yorkshire pigs highlights several DNA methylation loci and genes

Kai Wang[1†], Shujie Wang[1†], Xiang Ji[1], Dong Chen[1], Qi Shen[1], Yang Yu[1], Pingxian Wu[1,2], Xuewei Li[1] and Guoqing Tang[1*]

[1]Farm Animal Genetic Resources Exploration and Innovation Key Laboratory of Sichuan Province, Sichuan Agricultural University, Chengdu, China, [2]Chongqing Academy of Animal Science, Chongqing, China

In this study, we aimed to identified CpG sites at which DNA methylation levels are associated with meat quality traits in 140 Yorkshire pigs, including pH at 45 min (pH_{45min}), pH at 24 h (pH_{24h}), drip loss (DL), meat redness value (a*), yellowness (b*) and lightness (L*). Genome-wide methylation levels were measured in muscular tissue using reduced representation bisulfite sequencing (RRBS). Associations between DNA methylation levels and meat quality traits were examined using linear mixed-effect models that were adjusted for gender, year, month and body weight. A Bonferroni-corrected *p*-value lower than 7.79×10^{-8} was considered statistically significant threshold. Eight CpG sites were associated with DL, including CpG sites annotated to *RBM4* gene (cpg301054, cpg301055, cpg301058, cpg301059, cpg301066, cpg301072 and cpg301073) and *NCAM1* gene (cpg1802985). Two CpG sites were associated with b*, including *RNFT1* and *MED13* (cpg2272837) and *TRIM37* gene (cpg2270611). Five CpG sites were associated with L*, including *GSDMA* and *LRRC3C* gene (cpg2252750) and ENSSSCG00000043539 and *IRX1* gene (cpg2820178, cpg2820179, cpg2820181 and cpg2820182). No significant associations were observed with pH_{45min}, pH_{24h} or a*. We reported associations of meat quality traits with DNA methylation and identified some candidate genes associated with these traits, such as *NCAM1, MED13 and TRIM37* gene. These results provide new insight into the epigenetic molecular mechanisms of meat quality traits in pigs.

KEYWORDS

DNA methylation, EWAS, CPGs, pig, gene

1 Introduction

Meat quality are important traits in the pig industry. Meat quality traits mainly include, pH_{45m}, pH_{24h}, water-holding capacity (WHC) or DL, meat color and intramuscular fat content (IMF). In the process of pig genetics and breeding, the production performance of pig has always been the main breeding goal, and has

achieved remarkable results. However, in recent years, the meat quality of pig has attracted more and more attention. Improving meat quality has become a high priority for the pork industry.

A great deal of progress has been made by genome-wide association studies (GWAS) to identify genetic loci for meat quality traits (Zhang et al., 2015; Shen et al., 2016; Verardo et al., 2017; Xing et al., 2019). In recent decades, more than 30,000 quantitative trait loci (QTL) have been released for public access on the Pig QTLdb (release 40 December 2019. http://www.animalgenome.org/cgi-bin/QTLdb/SS/index). Among them, 730 QTLs have been found to affect pork pH and 651 QTLs are associated with meat color.

While tremendous progress has been made in identifying QTLs associated with meat quality traits, epigenetic mechanisms for regulating gene expression are less understood, such as DNA methylation, histone modification and chromatin accessibility. In particular, DNA methylation at CpG sites plays an important role in development, cell differentiation, imprinting and regulation of gene expression. DNA methylation is an annotation system that marks genetic text to guide how and when to read information and control transcription (Dor and Cedar, 2018). DNA methylation has been shown to be related to pig traits, including growth (Jin et al., 2014), reproduction (Bell et al., 2011) and immune response (Wang et al., 2017).

Similar to GWAS, epigenome-wide association studies (EWAS) use epigenetic factors instead of SNP to identify candidate genes for traits (Flanagan, 2015). In recent years, EWAS have identified associations for DNA methylation and complex traits in humans, such as body-mass index (BMI) (Dick et al., 2014; Demerath et al., 2015), obesity (Silva and Garvin, 2009; Klodian et al., 2018) and diseases (Dedeurwaerder et al., 2011; Mathias et al., 2016; Soriano-Tárraga et al., 2016; Meeks et al., 2019). However, up to now, most EWAS studies have been carried out in human but no EWAS studies have been conducted on pigs.

In this study, we aimed to investigate the association between DNA methylation and meat quality traits in Yorkshire pigs by using muscular tissue. We conducted EWAS using RRBS data and then identified 20 significant CpG sites associations with meat quality traits. The results are a step toward realizing the epigenetic molecular mechanisms of meat quality traits and identifying new loci.

TABLE 1 Descriptive statistics of meat quality characteristics of muscular tissue from Yorkshire pigs.

Characteristics	Yorkshire pigs (mean ±[a]SD)
N	140
Gender (male)	51
Weight (kg)	111.71 ± 12.97
$pH_{45\ min}$[b]	6.30 ± .29
[c]pH_{24h}	5.91 ± .33
DL (%)[d]	2.89 ± 3.14
a*[e]	5.09 ± 1.24
b*[f]	2.36 ± 3.66
L*[g]	45.77 ± 2.43

[a]SD, standard deviation.
[b]pH_{45min}, pH at 45 min.
[c]$pH_{24\ h}$, pH at 24 h.
[d]DL, drip loss.
[e]a*, meat redness value.
[f]b, meat yellowness value.
[g]L*, meat lightness value.

2 Results

2.1 Animals and meat quality traits

A total of 140 Yorkshire pigs (51 male and 89 female) were sampled in this study. The characteristics of these pigs were presented in Table 1. The mean value of pH_{45min}, pH_{24h}, DL, a*, b*, and L* were 6.30, 5.91, 2.89%, 5.09, 2.36, and 45.77, respectively.

2.2 DNA methylation

We constructed RRBS libraries from muscular tissue to examine methylation patterns in 140 pigs. We sequencing the libraries using Illumina HiSeq platform and then obtained on an average of 14.22 ± 1.93 Gb raw bases per sample (Supplementary Table S1). After quality control, we obtained on an average of 11.28 ± 1.74 Gb clean bases per sample. Moreover, the average bisulfite conversion of Yorkshire pigs was over 99%. Besides, more than 60% of the reads of Yorkshire pigs were mapped to the porcine reference genome. We filtered ML data for CpG sites with at least 10 × coverage, and present in at least 105 samples, corresponding to 3,083,713 CpG sites for further analysis. Figure 1 showed the distribution of CpG sites and methylation level of CpG sites for Yorkshire pigs in the 18 autosomal chromosomes.

2.3 EWAS

We performed EWAS between CpG ML and 6 meat quality metrics, including pH_{45min}, pH_{24h}, DL, a*, b* and L*. We then used the R package "CpGassoc" to determine associations between DNA methylation and phenotype, as is common practice for GWAS of quantitative traits. Based on the Bonferroni correction for the number of CpG sites tested, associations were deemed significant if *p*-value were below 7.79×10^{-8}.

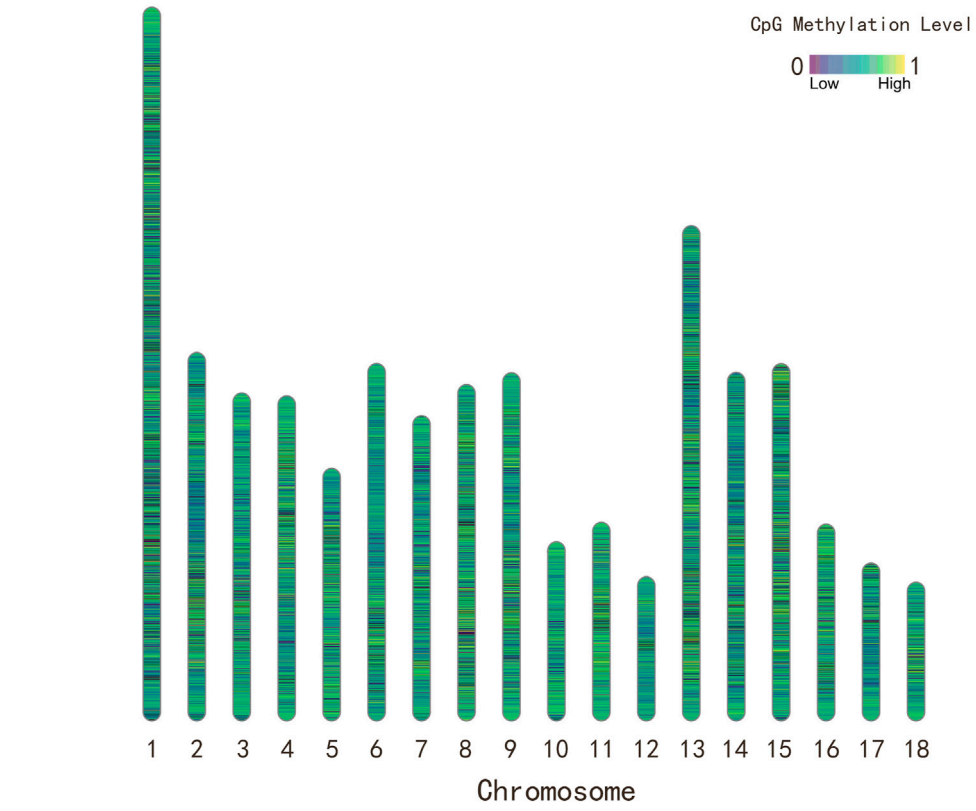


FIGURE 1
The distribution and methylation level of CpG sites for Yorkshire pigs in the 18 autosomal chromosomes. The color of chromosome ideogram represents the methylation level of CpG sites.

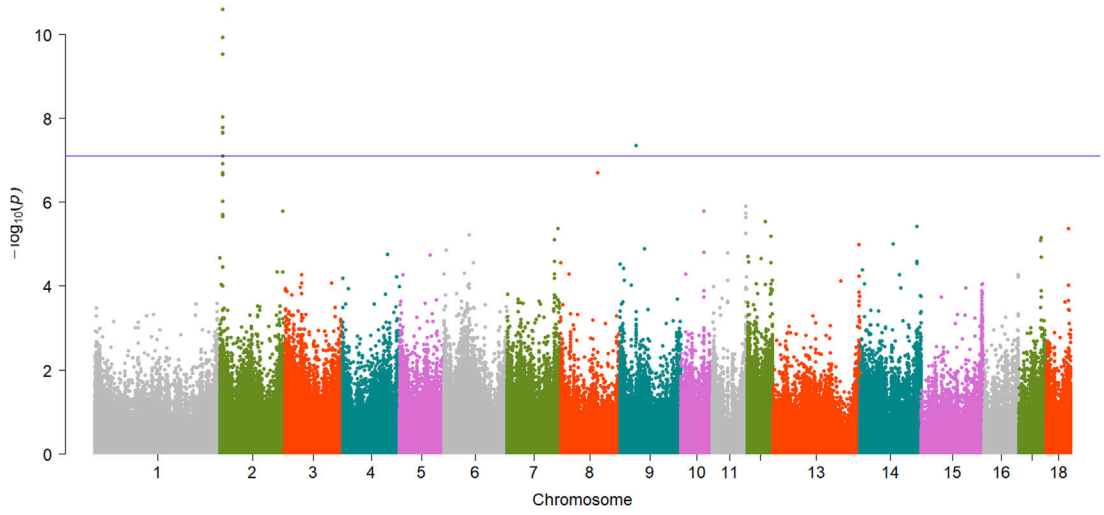


FIGURE 2
Manhattan plot showing EWAS results for DL trait. Each dot represents an CpG site and the blue line represents the threshold.

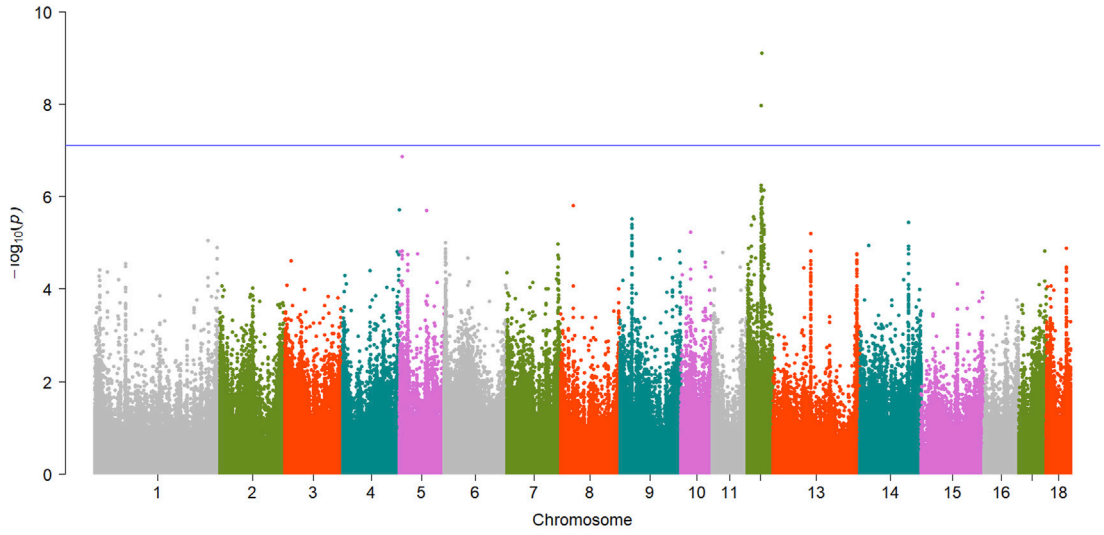


FIGURE 3
Manhattan plot showing EWAS results for b* trait. Each dot represents an CpG site and the blue line represents the threshold.

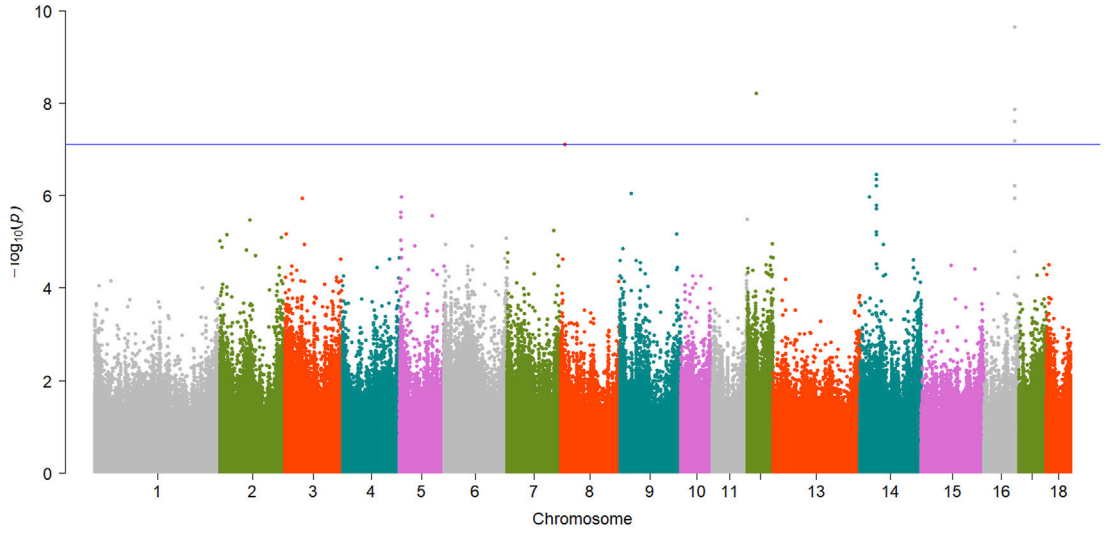


FIGURE 4
Manhattan plot showing EWAS results for L* trait. Each dot represents an CpG site and the blue line represents the threshold.

A total of 15 significant associations were detected, corresponding to 3 unique phenotypes (Figure 2, Figure 3 and Figure 4) where the *p*-value was below 7.79×10^{-8}. Table 2 summarized the significant CpG sites associated with these traits. However, we did not find significant CpG sites ($P < 7.79 \times 10^{-8}$) for pH_{45min} (Figure 5A), pH_{24h} (Figure 5B) and a* (Figure 5C).

As Table 2 shown, of the 15 significant CpG sites, 9 CpG sites were intragenic, and 6 CpG sites were intergenic. The distance between intergenic loci and nearby flanking genes ranged between 2,640 bp and 249,955 bp. The candidate genes listed for each site correspond to the gene itself for intragenic, and the two nearest flanking genes by distance for intergenic, with the distance between the site and each flanking gene listed for intergenic associations.

Subsequently, we calculated the average methylation level (ML) of significant CpG sites (Supplementary Table S2). Of these 15 significant CpG sites, 13 CpG sites (cpg301073, cpg301059,

TABLE 2 Summary of significant CpG sites associated with meat quality traits.

Traits	CpG sites	Chr	Pos	Intra/intergenic	Candidate genes	*p*-value
DL	cpg301073	2	5783171	Intragenic, exon	*RBM4*	2.47e-11
DL	cpg301059	2	5783071	Intragenic, exon	*RBM4*	1.13e-10
DL	cpg301055	2	5783024	Intragenic, exon	*RBM4*	2.85e-10
DL	cpg301058	2	5783070	Intragenic, exon	*RBM4*	9.17e-09
DL	cpg301066	2	5783114	Intragenic, exon	*RBM4*	1.63e-08
DL	cpg301054	2	5783023	Intragenic, exon	*RBM4*	2.14e-08
DL	cpg301072	2	5783170	Intragenic, exon	*RBM4*	2.20e-08
DL	cpg1802985	9	40844891	Intragenic, Intron	*NCAM1*	4.40e-08
b*	cpg2272837	12	36261646	Intergenic, 86,511 bp, 72,095 bp	*RNFT1*	7.72e-10
					MED13	
b*	cpg2270611	12	35340816	Intragenic, Intron	*TRIM37*	1.06e-08
L*	cpg2252750	12	22395823	Intergenic 16,700 bp 2,640 bp	*GSDMA*	6.05e-09
					LRRC3C	
L*	cpg2820181	16	77445768	Intergenic, 216,225 bp, 249,936 bp	ENSSSCG00000043539	2.20e-10
					IRX1	
L*	cpg2820178	16	77445749	Intergenic, 216,206 bp, 249,955 bp	ENSSSCG00000043539	1.3e-08
					IRX1	
L*	cpg2820182	16	77445775	Intergenic, 216,232 bp, 249,929 bp	ENSSSCG00000043539	2.44e-08
					IRX1	
L*	cpg2820179	16	77445750	Intergenic, 216,207 bp, 249,954 bp	ENSSSCG00000043539	6.61e-08
					IRX1	

cpg301055, cpg301058, cpg301066, cpg301054, cpg301072, cpg2272837, cpg2270611, cpg2820181, cpg2820178, cpg2820182, and cpg2820179) were hypermethylation, and only 2 CpG sites (cpg1802985 and cpg2252750) were hypomethylation.

2.4 Candidate genes

Totals of 8 CpG sites reached significant level and were regarded as significant sites for DL trait (Table 2). Of the 8 CpG sites, 7 sites were located within the exon1 of *RBM4* gene (Figure S1), and the top CpG site was cpg301073 (SSC2:5783171, $P = 2.47 \times 10^{-11}$). Besides, only 1 CpG site (cpg1802985, SSC9:40844891, $P = 4.40 \times 10^{-8}$) were located within the *NCAM1* gene (Supplementary Figure S2).

For meat color trait, totals of 2 and 5 CpG sites reached significant level for b* and L*, respectively. For b*, one of them, cpg2272837 (SSC12:36261646, $P = 7.72 \times 10^{-10}$), was located in the intergenic region and the nearest flanking genes were *RNFT1* gene and *MED13* gene (Supplementary Figure S3). Another one was cpg2270611 (SSC12:35240816, $P = 1.06 \times 10^{-8}$) that was located within the *TRIM37* gene (Supplementary Figure S3). In addition, for L*, the nearest flanking genes of the cpg2252750 (SSC12:22395823, $P = 6.05 \times 10^{-9}$) were *GSDMA* and *LRRC3C* (Supplementary Figure S4). A total of 4 CpG sites were located in the intergenic region and the nearest flanking genes were ENSSSCG00000043539 and *IRX1* gene (Supplementary Figure S5). Of them, the top CpG site was cpg2820181 (SSC16: 77445768, $P = 2.20 \times 10^{-10}$).

The main GO terms enriched in 9 candidate genes (Supplementary Table S1) might be related to negative

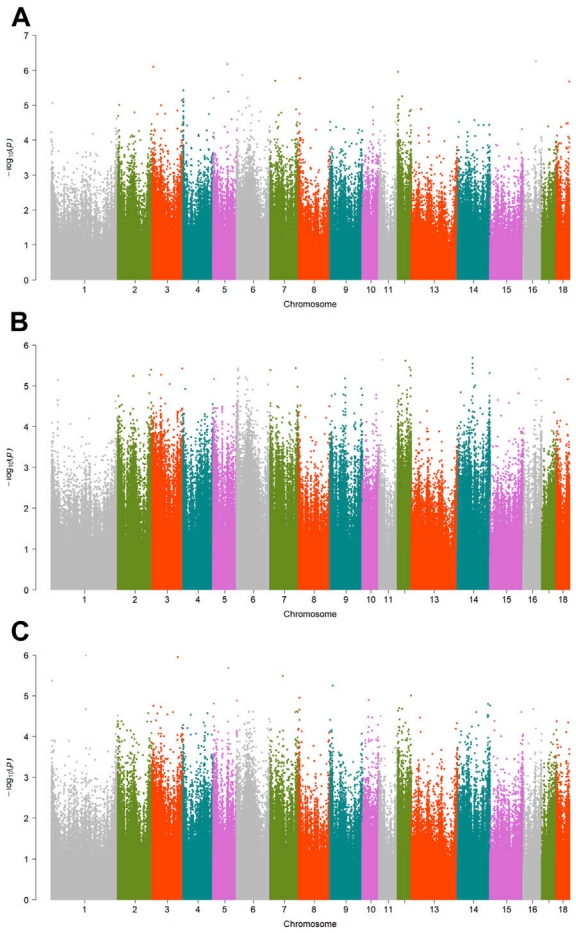


FIGURE 5
Manhattan plots showing EWAS results for other traits. Each dot represents an CpG site. **(A)** Manhattan plot showing EWAS results for pH_{45min}. **(B)** Manhattan plot showing EWAS results for pH_{24h}. **(C)** Manhattan plot showing EWAS results for a*.

regulation of centriole replication ($P = 2.4 \times 10^{-3}$) and protein autoubiquitination ($P = 2.4 \times 10^{-2}$).

3 Discussion

This study reported the results of EWAS of meat quality traits in 140 Yorkshire pigs, including drip loss (DL), meat pH ($pH_{45\ min}$, $pH_{24\ h}$), and meat color (a*, b* and L*). Just as genome-wide association studies (GWAS) grew from the field of genetic epidemiology, so too do EWAS derive from the field of epigenetic epidemiology. Although there have been numerous studies of EWAS in humans, few studies perform EWAS analysis to identified associations between DNA methylation and complex traits in livestock (Flanagan, 2015). In the current study, we found associations between DNA methylation levels and meat quality traits in the muscle tissue of pigs. These findings suggest that it was possible to find significant associations between DNA methylation and traits by using tissues associated with traits of interest. Besides, one of the advantages of using mammalian models, such as pigs, is that tissues that are not readily available in human studies can be collected.

DL trait is used to describe the water-holding capacity of meats. The water-holding capacity of meats is affected by multiple factors (Huff-Lonergan and Lonergan, 2005), including muscle cell structure, muscle contraction, muscle cell protein and genetic factors. We observed that methylation of 7 CpG sites annotated to the *RBM4* gene on SSC2 and 1 CpG site annotated to the *NCAM1* gene on SSC9 were associated with DL trait. The *RBM4* gene encodes RNA Binding Motif Protein 4 (RBM4) which participates in both precursor mRNA splicing regulation and translational control in muscle cells. RBM4 protein promotes the expression of many muscle-specific mRNAs from individual genes *via* its activity in modulating alternative splicing in myoblasts (Lin and Tarn, 2012). The *NCAM1* gene encodes a cell adhesion protein that is a member of cell adhesion molecules (CAMs) family. The CAMs are associated with the binding of a cell to another cell or to the extracellular matrix. They play important role in cell proliferation, differentiation, trafficking, motility, apoptosis and tissue architecture. Up to now, it is not clear how methylation of *RBM4* gene and *NCAM1* gene are involved in DL. Multiple studies have shown that *NCAM1* is associated with denervation and reinnervation, and is often used as a marker of muscle fiber denervation. We cannot definitively know that the *NCAM1* is associated with DL. However, that *NCAM1* is associated with denervation and reinnervation, and is often used as a marker of muscle fiber denervation (Barns et al., 2014; Messi et al., 2016) provides evidence to conclude that *NCAM1* may be involved in WHC of muscle by regulating skeletal muscle fiber.

When consumers choose fresh meat, meat color plays an important visual role. In this study, meat color traits were recorded by a Minolta CR-300 colorimeter. We identified several of significant CpG sites in 2 chromosome regions (SSC12 and SSC16). Yellowness of meat (b*) is mainly influenced by the fat deposits in muscle (Calnan et al., 2017). Usually, yellowness increases with the amount of fat deposited in the muscle. We identified 2 significant CpG sites for b* on SSC12 (Figure S3), which had'n been previously reported. The cpg2272837 is annotated to the upstream of *MED13* gene. The *MED13* is a protein coding gene that encodes a component of the mediator complex. The mediator complex acts as a centralized hub for transcriptional regulation and plays an important role in metabolic control (Amoasii et al., 2018). Previous study reported that *MED13* overexpression enhanced lipid metabolism, insulin sensitivity, and decreased susceptibility to obesity (Baskin et al., 2014). Yellowness of meat is affected by muscle fat deposition. This result could imply that *MED13* may

affect yellowness of meat by participating in lipid metabolism. The cpg2270611 is annotated within the *TRIM37* gene which encodes a peroxisomal protein (TRIM37) with E3 ubiquitin-ligase activity. The mutations in the *TRIM37* gene caused mulibrey (muscle-liver-brain-eye) nanism (MUL). Although the physiological function of TRIM37 *in vivo* is unclear, a study has shown that members of the ubiquitin-proteosome pathway can participate in energy metabolism by affecting the regulation of insulin signaling (Karlberg et al., 2005). Energy metabolism plays an important role in the process of muscle transformation after slaughter. Functional studies are needed to investigate the roles of the *MED13* and *TRIM37* gene in meat color among Yorkshire pigs.

A growing body of evidence supports the contribution of epigenetic modification to phenotypic variation in livestock (Triantaphyllopoulos et al., 2016; Wang and Ibeagha-Awemu, 2020) and supports the potential application of epigenetic biomarkers, particularly DNA methylation in livestock breeding programs. Moreover, a previous study has reported the association between SNPs and differential DNA methylation (Imgenberg-Kreuz et al., 2018). This provided one possible mechanism that SNPs impacts gene expression by altering DNA methylation, thereby suggesting the possible application of epigenetic biomarkers in livestock improvement breeding (Maldonado et al., 2019). Thus, in the development of new breeding methods, the relationship between DNA methylation biomarkers identified by EWAS and production traits can be considered in order to be able to quantify the epigenetic contribution to breeding value prediction. (Wang and Ibeagha-Awemu, 2020). Therefore, more studies are needed to get a better understanding of the epigenetic mechanisms underlying phenotyping variation in pig production.

4 Materials and methods

4.1 Animals and DNA sample

A total of 140 Yorkshire pigs (51 male and 89 female) were used to extracted genomic DNA from muscular tissue. The pigs were raised under the same recommended environment at the conservation farm of Mingxing Agricultural science and Technology Development Co.,Ltd. All individuals were raised to 111.71 Kg (± 12.97 Kg) weight, transported to the slaughterhouse, and were fasting for 24 h before slaughter determination. After the carcass composition traits were determined, meat quality traits were measured using methods previously described in detail (Duan et al., 2009; Ma et al., 2009). All meat quality measurements were taken on the left side of the carcass. Muscle pH values of the longissimus dorsi muscle were measured at 45 min (pH_{45min}) and 24 h (pH_{24h}) using a portable pH meter (model 720A; Orion Research Inc., Boston, MA, United States). Meat color, including lightness (L*), redness (a*), and yellowness (b*) were measured at 45 min using a Minolta CR-300 colorimeter (Minolta Camera, Osaka, Japan). Drip loss (DL) was accessed by the method described by Otto et al. (2004). Briefly, a slice of fresh muscle was placed in a plastic bag on a grid parallel to the fiber direction. The weight loss percentages after 24 h of storage at 4°C were calculated. Muscle pH, meat color, and DL were measured in triplicate for each sample, and the average of the three measurements was used. DNA samples from the muscle tissue were snap-frozen in liquid nitrogen then held at −80°C until analysis.

4.2 DNA methylation data

Briefly, genomic DNA was isolated from flash frozen muscular tissue. Then, the construction of RRBS libraries and paired-end sequencing using Illumina HisSeq analyzer was performed at Novogene technology co., LTD. (Beijing, China). Raw sequencing data were processed by an Illumina base-calling pipeline. Clean reads were aligned to the pig reference genome (Sscrofa11.1) using Bismark (Felix and Andrews, 2011) after removing adaptor sequences. Next, ML were measured using bismark_methylation_extrator program. A quality control procedure was adopted to ensure the high data quality by 1) retaining only CpG cytosines across all samples; 2) removing CpG sites with missing methylation values at >35 samples; 3) removing CpG sites with coverage <10 reads within a sample. A total of 3,083,713 CpG sites were retained for further analysis. The distribution of CpG sites were conducted using R package Rldeogram v.2.2 (Hao et al., 2020).

4.3 Data analysis

We used the R package CpGassoc v2.60 (Barfield et al., 2012) to test for association between methylation and phenotype. We applied the linear mixed-model with the CpG ML score (vary between 0 and 1) as the outcome, and adjusted for sex, year, month, parity and body weight. The following model was used:

$$y = \mu + x\beta + u + e$$

where y is vector of the phenotypes, μ is the mean, x is the vector of CpG, β is CpG effect, μ represent random effects, e is the vector of random residuals.

We corrected for multiple comparisons using a Bonferroni correction for $0.05/3,083,713$ test setting a significant threshold p-value to 7.79×10^{-8}. We constructed Manhattan plots to present the results of epigenome-wide

association analysis using the R package qqman v0.1.8 (Turner, 2018).

4.4 Candidate genes and annotation

Then we identified candidate genes based on the significant CpG sites by the Ensemble biomart (http://www.biomart.org). The candidate genes listed for each site correspond to the gene itself for intragenic, and the two nearest flanking genes by distance for intergenic, with the distance between the site and each flanking gene listed for intergenic associations. The gene lists were then submitted for enrichment analysis using the Database for Annotation, Visualization and Integrated Discovery (DAVID) v6.8 (http://david.ncifcrf.gov/). Significant Gene Ontology (GO) terms and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways were selected after filtering with $P < 0.05$.

Data availability statement

The data presented in the study are deposited in the Figshare repository, accession number 10.6084/m9.figshare.20633217.

Ethics statement

The animal study was reviewed and approved by All experimental procedures were performed in accordance with the Institutional Review Board (IRB14044) and the Institutional Animal Care and Use Committee of the Sichuan Agricultural University under permit number DKY-B20140302. Written informed consent was obtained from the owners for the participation of their animals in this study.

Author contributions

Conceptualization: GT and XL; Methodology: GT and KW; Formal analysis: KW, SW; Writing—original draft preparation: KW and SW; Writing—review and editing: XJ and DC; Funding acquisition: XL; Supervision: QS, YY, KW, and QT. All authors have read and approved the manuscript.

Funding

The study was supported by grants from the Sichuan Science and Technology Program (2020YFN0024), the earmarked fund for the China Agriculture Research System (CARS-35-01 A), the National Key R&D Program of China (2018YFD0501204), the National Natural Science Foundation of China (C170102) and the Sichuan Innovation Team of Pig (sccxtd-2021–08). This study was supported by High-performance Computing Platform of Sichuan Agricultural University.

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.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.1028711/full#supplementary-material

SUPPLEMENTARY FIGURE S1
The Manhattan plot of significant CpG sites on SSC2 for DL trait and the graphic of significant CpG sites in the genomic region. Each dot represents an CpG site and the green dot in the circle represents significant CpG site. The X-axis represents genomic region, and the Y-axis represents average methylation level of CpG sites.

SUPPLEMENTARY FIGURE S2
The Manhattan plot of significant CpG sites on SSC9 for DL trait and the graphic of significant CpG sites in the genomic region. Each dot represents an CpG site and the green dot in the circle represents significant CpG site. The X-axis represents genomic region, and the Y-axis represents average methylation level of CpG sites.

SUPPLEMENTARY FIGURE S3
The Manhattan plot of significant CpG sites on SSC12 for b* and the graphic of significant CpG sites in the genomic region. Each dot represents an CpG site and the green dot in the circle represents significant CpG site. The X-axis represents genomic region, and the Y-axis represents average methylation level of CpG sites.

SUPPLEMENTARY FIGURE S4
The Manhattan plot of significant CpG sites on SSC12 for L* and the graphic of significant CpG sites in the genomic region. Each dot represents an CpG site and the green dot in the circle represents significant CpG site. The X-axis represents genomic region, and the Y-axis represents average methylation level of CpG sites.

SUPPLEMENTARY FIGURE S5
The Manhattan plot of significant CpG sites on SSC16 for L* and the graphic of significant CpG sites in the genomic region. Each dot represents an CpG site and the green dot in the circle represents significant CpG site. The X-axis represents genomic region, and the Y-axis represents average methylation level of CpG sites.

References

Amoasii, L., Olson, E. N., and Bassel-Duby, R. (2018). Control of muscle metabolism by the mediator complex. *Cold Spring Harb. Perspect. Med.* 8, a029843. doi:10.1101/cshperspect.a029843

Barfield, R. T., Kilaru, V., Smith, A. K., and Conneely, K. N. (2012). CpGassoc: an R function for analysis of DNA methylation microarray data. *Bioinformatics* 28, 1280–1281. doi:10.1093/bioinformatics/bts124

Barns, M., Gondro, C., Tellam, R. L., Radley-Crabb, H. G., Grounds, M. D., and Shavlakadze, T. (2014). Molecular analyses provide insight into mechanisms underlying sarcopenia and myofibre denervation in old skeletal muscles of mice. *Int. J. Biochem. Cell Biol.* 53, 174–185. doi:10.1016/j.biocel.2014.04.025

Baskin, K. K., Grueter, C. E., Kusminski, C. M., Holland, W. L., Bookout, A. L., Satapati, S., et al. (2014). MED13-dependent signaling from the heart confers leanness by enhancing metabolism in adipose tissue and liver. *EMBO Mol. Med.* 6, 1610–1621. doi:10.15252/emmm.201404218

Bell, J. T., Pai, A. A., Pickrell, J. K., Gaffney, D. J., Pique-Regi, R., Degner, J. F., et al. (2011). DNA methylation patterns associate with genetic and gene expression variation in HapMap cell lines. *Genome Biol.* 12, R10.R10. doi:10.1186/gb-2011-12-1-r10

Calnan, H. B., Jacob, R. H., Pethick, D. W., and Gardner, G. E. (2017). Selection for intramuscular fat and lean meat yield will improve the bloomed colour of Australian lamb loin meat. *Meat Sci.* 131, 187–195. doi:10.1016/j.meatsci.2017.05.001

Dedeurwaerder, S., Desmedt, C., Calonne, E., Singhal, S. K., Haibe-Kains, B., Defrance, M., et al. (2011). DNA methylation profiling reveals a predominant immune component in breast cancers. *Embo Mol. Med.* 3, 726–741. doi:10.1002/emmm.201100801

Demerath, E. W., Guan, W., Grove, M. L., Stella, A., Michael, M., Yi-Hui, Z., et al. (2015). Epigenome-wide association study (EWAS) of BMI, BMI change and waist circumference in African American adults identifies multiple replicated loci. *Hum. Mol. Genet.* 15, 4464–4479. doi:10.1093/hmg/ddv161

Dick, K. J., Nelson, C. P., Tsaprouni, L., Sandling, J. K., A?Ssi, D., Wahl, S., et al. (2014). DNA methylation and body-mass index: A genome-wide analysis. *Lancet* 383, 1990–1998. doi:10.1016/S0140-6736(13)62674-4

Dor, Y., and Cedar, H. (2018). Principles of DNA methylation and their implications for biology and medicine. *Lancet* 392, 777–786. doi:10.1016/S0140-6736(18)31268-6

Duan, Y. Y., Ma, J. W., Yuan, F., Huang, L. B., Yang, K. X., Xie, J. P., et al. (2009). Genome-wide identification of quantitative trait loci for pork temperature, pH decline, and glycolytic potential in a large-scale White Duroc x Chinese Erhualian resource population. *J. Animal Sci.* 87, 9–16. doi:10.2527/jas.2008-1128

Felix, K., and Andrews, S. R. (2011). Bismark: A flexible aligner and methylation caller for bisulfite-seq applications. *Bioinformatics* 11, 1571–1572. doi:10.1093/bioinformatics/btr167

Flanagan, J. M. (2015). Epigenome-wide association studies (EWAS): Past, present, and future. *Methods Mol. Biol.* 1238, 51–63. doi:10.1007/978-1-4939-1804-1_3

Hao, Z., Lv, D., Ge, Y., Shi, J., Weijers, D., Yu, G., et al. (2020). RIdeogram: Drawing SVG graphics to visualize and map genome-wide data on the idiograms. *PeerJ. Comput. Sci.* 6, e251. doi:10.7717/peerj-cs.251

Huff-Lonergan, E., and Lonergan, S. M. (2005). Mechanisms of water-holding capacity of meat: The role of postmortem biochemical and structural changes. *Meat Sci.* 71, 194–204. doi:10.1016/j.meatsci.2005.04.022

Imgenberg-Kreuz, J., Carlsson Almlöf, J., Leonard, D., Alexsson, A., Nordmark, G., Eloranta, M. L., et al. (2018). DNA methylation mapping identifies gene regulatory effects in patients with systemic lupus erythematosus. *Ann. Rheum. Dis.* 77, 736–743. doi:10.1136/annrheumdis-2017-212379

Jin, L., Jiang, Z., Xia, Y., Lou, P., Chen, L., Wang, H., et al. (2014), Genome-wide DNA methylation changes in skeletal muscle between young and middle-aged pigs. *BMC Genomics* 15, 653. doi:10.1186/1471-2164-15-653

Karlberg, N., Jalanko, H., Kallijarvi, J., Lehesjoki, A. E., and Lipsanen-Nyman, M. (2005). Insulin resistance syndrome in subjects with mutated RING finger protein TRIM37. *Diabetes* 54, 3577–3581. doi:10.2337/diabetes.54.12.3577

Klodian, D., Braun, K. V. E., Jana, N., Trudy, V., Demerath, E. W., Guan, W., et al. (2018). An epigenome-wide association study of obesity-related traits. *Am. J. Epidemiol.* 8, 1662–1669. doi:10.1093/aje/kwy025

Lin, J. C., and Tarn, W. Y. (2012). Multiple roles of RBM4 in muscle cell differentiation. *Front. Biosci. Sch. Ed.* 4, 181–189. doi:10.2741/260

Ma, J., Ren, J., Guo, Y., Duan, Y., Ding, N., Zhou, L., et al. (2009). Genome-wide identification of quantitative trait loci for carcass composition and meat quality in a large-scale White Duroc x Chinese Erhualian resource population. *Anim. Genet.* 40, 637–647. doi:10.1111/j.1365-2052.2009.01892.x

Maldonado, M. B. C., De Rezende Neto, N. B., Nagamatsu, S. T., Carazzolle, M. F., Hoff, J. L., Whitacre, L. K., et al. (2019). Identification of bovine CpG SNPs as potential targets for epigenetic regulation via DNA methylation. *PLoS One* 14, e0222329. doi:10.1371/journal.pone.0222329

Mathias, R. A., David, M., Muhammad, A., Stefan, E., Ek, W. E., Ulf, G., et al. (2016). Epigenome-wide association study reveals differential DNA methylation in individuals with a history of myocardial infarction. *Hum. Mol. Genet.* 21, 4739–4748. doi:10.1093/hmg/ddw302

Meeks, K. a. C., Peter, H., Andrea, V., Juliet, A., Silver, B., Tom, B., et al. (2019). Epigenome-wide association study in whole blood on type 2 diabetes among sub-saharan african individuals: Findings from the RODAM study. *Int. J. Epidemiol.*, 58–70. doi:10.1093/ije/dyy171

Messi, M. L., Li, T., Wang, Z. M., Marsh, A. P., Nicklas, B., and Delbono, O. (2016). Resistance training enhances skeletal muscle innervation without modifying the number of satellite cells or their myofiber association in obese older adults. *J. Gerontol. A Biol. Sci. Med. Sci.* 71, 1273–1280. doi:10.1093/gerona/glv176

Otto, G., Roehe, R., Looft, H., Thoelking, L., and Kalm, E. (2004). Comparison of different methods for determination of drip loss and their relationships to meat quality and carcass characteristics in pigs. *Meat Sci.* 68, 401–409.

Shen, L., Du, J., Xia, Y., Tan, Z., Fu, Y., Yang, Q., et al. (2016). Genome-wide landscape of DNA methylomes and their relationship with mRNA and miRNA transcriptomes in oxidative and glycolytic skeletal muscles. *Sci. Rep.* 6, 32186. doi:10.1038/srep32186

Silva, G. B., and Garvin, J. L. (2009). Akt1 mediates purinergic-dependent NOS3 activation in thick ascending limbs. *Am. J. physiology. Ren. physiology* 297, F646–F652. doi:10.1152/ajprenal.00270.2009

Soriano-Tárraga, C., Jiménez-Conde, J., Giralt-Steinhauer, E., Mola-Caminal, M., Vivanco-Hidalgo, R. M., Ois, A., et al. (2016). Epigenome-wide association study identifies TXNIP gene associated with type 2 diabetes mellitus and sustained hyperglycemia. *Hum. Mol. Genet.* 25, 609–619. doi:10.1093/hmg/ddv493

Triantaphyllopoulos, K. A., Ikonomopoulos, I., and Bannister, A. J. (2016). Epigenetics and inheritance of phenotype variation in livestock. *Epigenetics Chromatin* 9, 31. doi:10.1186/s13072-016-0081-5

Turner, S. D. (2018). qqman: an R package for visualizing GWAS results using Q-Q and manhattan plots. *J. Open Source Softw.* 3, 731. doi:10.21105/joss.00731

Verardo, L. L., Sevón-Aimonen, M. L., Serenius, T., Hietakangas, V., and Uimari, P. (2017). Whole-genome association analysis of pork meat pH revealed three significant regions and several potential genes in Finnish Yorkshire pigs. *Bmc Genet.* 18, 13. doi:10.1186/s12863-017-0482-x

Wang, H., Wang, J., Ning, C., Zheng, X., Fu, J., Wang, A., et al. (2017). Genome-wide DNA methylation and transcriptome analyses reveal genes involved in immune responses of pig peripheral blood mononuclear cells to poly I:C. *Sci. Rep.* 7, 9709. doi:10.1038/s41598-017-10648-9

Wang, M., and Ibeagha-Awemu, E. M. (2020). Impacts of epigenetic processes on the Health and productivity of livestock. *Front. Genet.* 11, 613636. doi:10.3389/fgene.2020.613636

Xing, K., Zhao, X., Ao, H., Chen, S., Yang, T., Tan, Z., et al. (2019). Transcriptome analysis of miRNA and mRNA in the livers of pigs with highly diverged backfat thickness. *Sci. Rep.* 9, 16740. doi:10.1038/s41598-019-53377-x

Zhang, C., Wang, Z., Bruce, H., Kemp, R. A., Charagu, P., Miar, Y., et al. (2015). Genome-wide association studies (GWAS) identify a QTL close to PRKAG3 affecting meat pH and colour in crossbred commercial pigs. *Bmc Genet.* 16, 33. doi:10.1186/s12863-015-0192-1

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 26 January 2023
DOI 10.3389/fgene.2023.1113515

OPEN ACCESS

EDITED BY
Xiao Wang,
Konge Larsen ApS, Denmark

REVIEWED BY
Yueqi Chen,
Army Medical University, China
Chunhui Ma,
Shanghai General Hospital, China
Weiwei Lin,
Zhejiang University, China

*CORRESPONDENCE
Zhirong Wang,
zjgfy_spine_wzr@njucm.edu.cn
Long Xiao,
zjgfy_spine_xl@njucm.edu.cn
Shuaijie Lv,
lvshuaijie1990@zcmu.edu.cn

†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 Genetics

RECEIVED 01 December 2022
ACCEPTED 13 January 2023
PUBLISHED 26 January 2023

CITATION
Gu Y, Wang Z, Wang R, Yang Y, Tong P, Lv S, Xiao L and Wang Z (2023), N6-methyladenine regulator-mediated RNA methylation modification patterns in immune microenvironment regulation of osteoarthritis.
Front. Genet. 14:1113515.
doi: 10.3389/fgene.2023.1113515

N6-methyladenine regulator-mediated RNA methylation modification patterns in immune microenvironment regulation of osteoarthritis

Yong Gu[1,2†], Zhengming Wang[3,4†], Rui Wang[5], Yunshang Yang[1,2], Peijian Tong[5], Shuaijie Lv[5*], Long Xiao[1,2*] and Zhirong Wang[1,2*]

[1]Translational Medical Innovation Center, Zhangjiagang TCM Hospital Affiliated to Nanjing University of Chinese Medicine, Zhangjiagang, China, [2]Department of Orthopedics, Zhangjiagang TCM Hospital Affiliated to Nanjing University of Chinese Medicine, Zhangjiagang, China, [3]Shi's Center of Orthopedics and Traumatology, Shuguang Hospital Affiliated to Shanghai University of Traditional Chinese Medicine, Shanghai, China, [4]Institute of Traumatology and Orthopedics, Shanghai Academy of Traditional Chinese Medicine, Shanghai, China, [5]The First Affiliated Hospital of Zhejiang Chinese Medical University, Zhejiang Provincial Hospital of Chinese Medicine, Hangzhou, China

Background: Osteoarthritis is a common chronic degenerative disease, and recently, an increasing number of studies have shown that immunity plays an important role in the progression of osteoarthritis, which is exacerbated by local inflammation. The role of N6-methyladenine (m^6A) modification in immunity is being explored. However, the role of m^6A modification in regulating the immune microenvironment of osteoarthritis remains unknown. In this study, we sought to discuss the association between the N6-methyladenine (m^6A) modification and the immune microenvironment of osteoarthritis.

Methods: First, the data and gene expression profiles of 139 samples, including 33 healthy samples and 106 osteoarthritis samples, were obtained from the Genetics osteoARthritis and Progression (GARP) study. Then the differences in m^6A regulators between healthy individuals and osteoarthritis patients were analyzed. The correlation between m^6A regulators and immune characteristics was also investigated by single-sample gene set enrichment analysis (ssGSEA). Principal component analysis (PCA), Gene Set Variation Analysis (GSVA) enrichment analysis, weighted gene coexpression network analysis (WGCNA), and Associated R packages were used to identify the m^6A phenotype and its biological functions.

Results: A total of 23 m^6A regulators were involved in this study. We found a close correlation between most m^6A regulators in all samples as well as in osteoarthritis samples. VIRMA and LRPPRC were the most highly correlated m^6A regulators and showed a positive correlation, whereas VIRMA and RBM15B were the most negatively correlated. M^6A regulators are associated with osteoarthritis immune characteristics. For example, MDSC cell abundance was strongly correlated with RBM15B and HNRNPC. Meanwhile, RBM15B and HNRNPC were important effectors of natural killer cell immune responses. IGFBP3 is an important regulator of cytolytic activity immune function. We performed an unsupervised consensus cluster analysis of the osteoarthritis samples based on the expression of 23 m^6A regulators. Three different m^6A subtypes of osteoarthritis were identified, including 27 samples in subtype C1, 21 samples in subtype C2, and 58 samples in subtype C3. Different m^6A subtypes have unique biological pathways and play different roles in the immune microenvironment of osteoarthritis.

Conclusion: The m^{6}A modification plays a crucial role in the diversity and complexity of the immune microenvironment in osteoarthritis.

KEYWORDS

N6-methyladenine (m^{6}A), osteoarthritis, methylation, immune, microenvironment

1 Introduction

Osteoarthritis (OA) is a common chronic degenerative disease that is characterized by joint pain, swelling, and limited activity, resulting in decreased activity and dysfunction of elderly individuals (Abramoff and Caldera, 2020). Patients may endure severe pain with decreased joint mobility, resulting in rising healthcare system costs and decreased work productivity. It was thought in the past that OA is simply produced by mechanical wear and tear and that its mechanism was an imbalance in joint biomechanics (Vincent, 2013). Recently, OA has been understood to result from a complex interplay of local and systemic factors. An increasing number of studies have demonstrated that immune cell infiltration plays an important role in the progression of OA and that local inflammation further aggravates the disease process (Moradi et al., 2015; Rosshirt et al., 2019). The body's natural wound healing response is manifested in osteoarthritic joints, and there is growing interest in how immunity influences disease initiation and progression (Daghestani and Kraus, 2015). Therefore, immunomodulation in osteoarthritis may be key to the new pathological mechanisms behind it and may shed some light on the discovery of novel immunotherapies for osteoarthritis.

Currently, an increasing number of studies are revealing a novel mode of inheritance, epigenetics, which is based on changes in the expression levels of genes caused by non-genetic sequence alterations (Harvey et al., 2018). Among them, as the third layer of epigenetics, more than 150 RNA modifications have been identified, including N1-methyladenosine (m^{1}A), N6-methyladenine (m^{6}A), 5-methylcytosine (m^{5}C), and 7-methylguanosine (m^{7}G). Notably, m^{6}A is the most abundant form and has received substantial attention (Patil et al., 2016). It is a dynamic and reversible RNA modification that is involved in a wide range of biological and pathological processes, such as cancer progression and inflammation (Lan et al., 2019; Zong et al., 2019). m^{6}A is the most common chemical modification of eukaryotic mRNA and is important in the regulation of mRNA stability, splicing, and translation (Cao et al., 2016). Its regulatory proteins include writers (METTL3, METTL14, WTAP, etc.), erasers (FTO, ALKBH5, etc.), and readers (YTHDF1, YTHDF2, YTHDF3, etc.) (Yang et al., 2018).

Recent studies have identified that m^{6}A modification can regulate various aspects of immune function, including immune recognition, activation of innate and adaptive immune responses, and cell fate decisions (Shulman and Stern-Ginossar, 2020). Despite increasing evidence for the regulatory role of m^{6}A in immune responses, current studies focusing on the role of m^{6}A modification in the immune-related pathogenesis of osteoarthritis are still lacking. Existing studies have mainly focused on METTL3 and FTO (Liu et al., 2019; arc et al., 2012; Panoutsopoulou et al., 2014). The correlation between m^{6}A regulators and osteoarthritis remains elusive and requires further exploration. In-depth investigation of immune dysregulation between normal samples and osteoarthritis samples as well as among the various subtypes of osteoarthritis and how m^{6}A regulators act on these changes may shed light on osteoarthritis pathogenesis from a new perspective.

However, previous studies have been limited to a few m^{6}A regulators due to technical limitations. In this study, we systematically evaluated the modification patterns of m^{6}A regulators in osteoarthritis, which furthers our understanding of the immune microenvironment in osteoarthritis. We found that the classification model based on m^{6}A regulators could distinguish osteoarthritis samples from healthy samples. There was a high degree of coordination and correlation between m^{6}A regulators and infiltrating immune cells, immune responses, and immune functions in osteoarthritis. We identified 3 distinct m^{6}A-modified subtypes where different immune characteristics were observed, and we compared the biological functions of these subtypes. In addition, we studied 1175 m^{6}A phenotype-related genes and their biological functions. In conclusion, the effect of m^{6}A modification on the immune microenvironment of osteoarthritis cannot be ignored.

2 Materials and methods

2.1 Dataset sources and preprocessing

The data used in this study consisted of 139 samples, including 33 healthy samples and 106 osteoarthritis samples. These samples were obtained from 139 participants of the Genetics osteoARthritis and Progression (GARP) study, and gene expression profiles were extracted from peripheral blood mononuclear cells (PBMCs) of these participants (The age and gender information are provided in Supplementary Table S1). The sample processing protocol and RNA extraction method were well described in a previous study (Ramos et al., 2014). The dataset was deposited in the Gene Expression Omnibus (GEO) database with the accession number GSE48556. The R/Bioconductor package "GEOquery" (Davis & Meltzer, 2007) was used to extract the GEO dataset, which consisted of the gene expression matrix and clinical features. According to the annotation information of the GPL6947 platform, probe mapping was applied to genes. If multiple probes corresponded to one gene, the average value was taken, and probes corresponding to multiple genes were deleted. Matrix expression values were preprocessed by correction with the "normalizeBetweenArrays" function in the "limma" package (Ritchie et al., 2015).

2.2 Alteration analysis of m^{6}A regulators between healthy individuals and osteoarthritis patients

These 23 m^{6}A regulators involved in the study included 8 writers (METTL3, METTL14, METTL16, WTAP, VIRMA, ZC3H13, RBM15, and RBM15B), 2 erasers (FTO and ALKBH5), and 13 readers

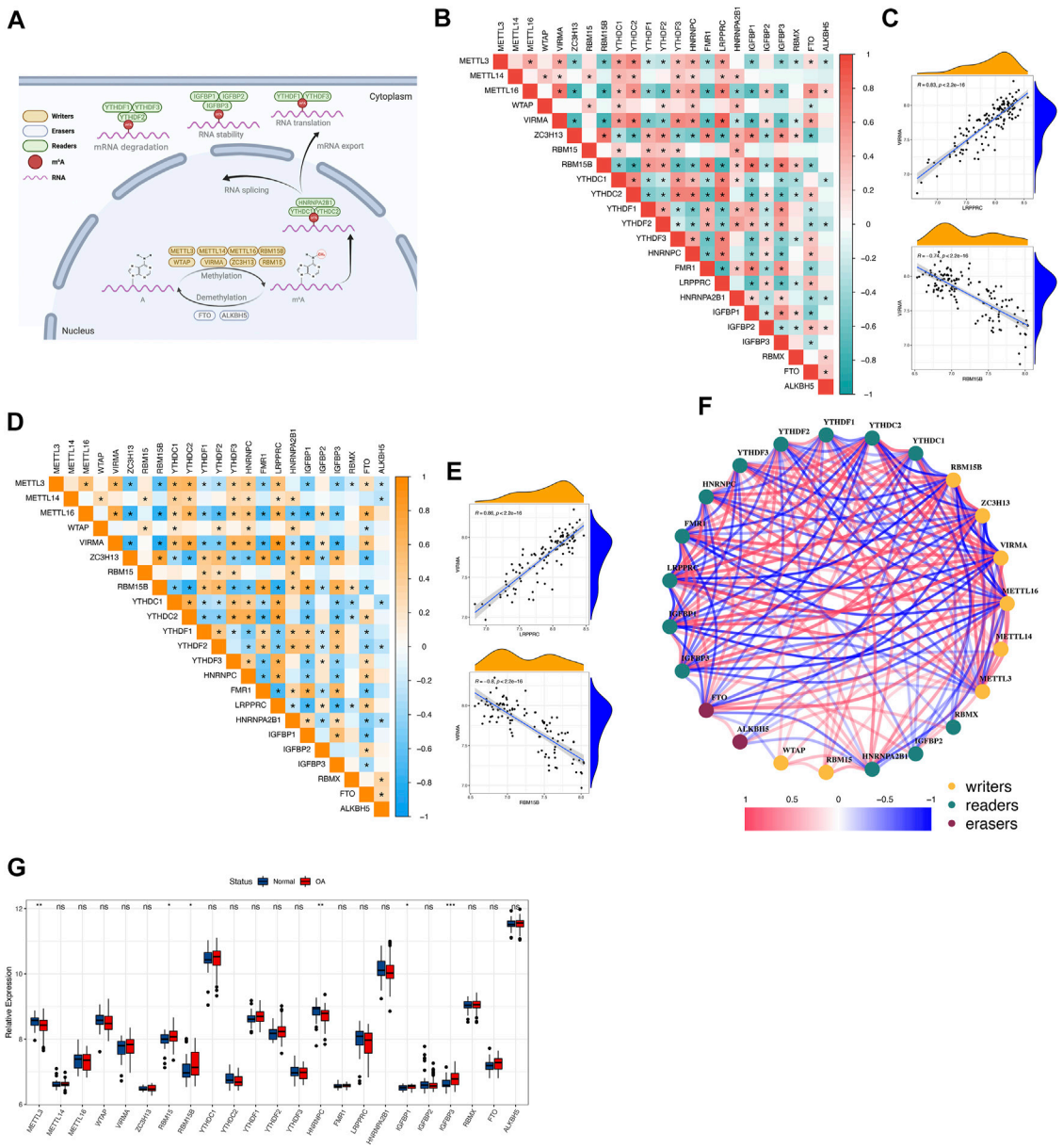


FIGURE 1
The landscape of m^{6}A RNA methylation regulators in osteoarthritis **(A)** Overview of the dynamic reversible process of m^{6}A RNA methylation modification regulated by "writers", "erasers" and "readers" in osteoarthritis and their potential biological functions for RNA. **(B,C)** The correlation of the expression of 23 m^{6}A regulators in all samples. Red indicates a positive correlation, and green indicates a negative correlation. Two scatter plots showed the two most relevant sets of m^{6}A regulators: VIRMA and LRPPRC were the most positively correlated and VIRMA and RBM15B were the most negatively correlated. Above * means $p < 0.05$. **(D,E)** The correlation of the expression of 23 m^{6}A regulators in osteoarthritis samples. Orange indicates a positive correlation, and blue indicates a negative correlation. Two scatter plots showed the two most relevant sets of m^{6}A regulators: VIRMA and LRPPRC were the most positively correlated and VIRMA and RBM15B were the most negatively correlated. Above * means $p < 0.05$. **(F)** The regulatory network of 23 m^{6}A regulators in osteoarthritis: red indicates a positive correlation, and blue indicates a negative correlation. **(G,H)** The boxplot and heatmap show the expression of 23 m^{6}A regulators between healthy and osteoarthritis samples. Above * indicates $p < 0.05$, ** indicates $p < 0.01$, *** indicates $p < 0.001$, and ns indicates that the difference was not statistically significant.

(YTHDC1, YTHDC2, YTHDF1, YTHDF2, YTHDF3, HNRNPC, FMR1, LRPPRC, HNRNPA2B1, IGFBP1, IGFBP2, IGFBP3, and RBMX). The expression relationships among the 23 m^{6}A regulators were evaluated by Spearman correlation analysis in all samples and osteoarthritis samples. Then, we constructed a correlation network of these 23 m^{6}A regulators. The expression differences of the 23 m^{6}A regulators between healthy and osteoarthritis samples were compared by the Wilcoxon test. OA-related m^{6}A regulators were determined by univariate logistic regression with a cutoff criterion of *p*-value <0.2. Least absolute shrinkage and selection operator (LASSO) regression was used for feature selection and dimension reduction. Multivariate logistic regression was used to develop a m^{6}A regulator-associated osteoarthritis classification model and external data sets (Details are provided in Supplementary Table S2) were used for verification.

Receiver operating characteristic (ROC) curve analysis was used to evaluate the discriminatory performance of the model signatures.

2.3 Correlation between m^{6}A regulators and immune characteristics

Single-sample gene set enrichment analysis (ssGSEA) was used to estimate the abundance of specific infiltrating immune cells and the activity of specific immune responses and immune function. It defines an enrichment fraction to express the absolute degree of enrichment of a gene set in each sample (Shen et al., 2019). The immune cell gene set and the immune function gene set were derived from previous studies (Zhang et al., 2020a; Liang et al., 2020). The immune genes and immune response gene sets were obtained from the ImmPort database (http://www.immport.org) (Bhattacharya et al., 2014). The Wilcoxon test was used to compare the abundance of immune cells, immune response, and immune function enrichment scores between healthy and osteoarthritis samples. We analyzed the correlation between the expression of m^{6}A regulators and the immune cell fraction, immune response activity, and immune function activity by the Spearman method.

2.4 Identification of distinct m^{6}A modification patterns by unsupervised clustering

The ConsensusClusterPlus package was applied to classify disease samples into distinct subtypes based on the expression of 23 m^{6}A regulators (Wilkerson and Hayes, 2010). This is an unsupervised clustering analysis method. The Euclidean distance was utilized to calculate the similarity distance between samples, and the K-means algorithm was used to evaluate cluster numbers and robustness (Hartigan and Wong, 2013). The maximum cluster number was set to be 9. Eighty percent of the samples were sampled by the resampling scheme, and resampling was conducted 1000 times. The final cluster number was determined by the consensus matrix and the cluster consensus score (>0.8). Principal component analysis (PCA) was used to further verify the distinct modification patterns of 23 m^{6}A regulators.

2.5 Immune characteristics and biological enrichment analysis of distinct m^{6}A modification subtypes

We compared the differences in the immune cell fraction, immune response activity, and immune function activity among the m^{6}A subtypes by the Kruskal test. To investigate the differences in biological functions and processes between m^{6}A modification patterns, Gene Set Variation Analysis (GSVA) enrichment analysis was applied by the "GSVA" package. GSVA, known as gene set variant analysis, is a non-parametric unsupervised analysis method that transforms the expression matrix of genes across different samples into a pathway activation score matrix and evaluates whether different biological pathways are enriched across samples (Hanzelmann et al., 2013). The HALLMARKS pathway and KEGG pathway are two commonly used pathway gene sets. From the MSigDB database (http://www.gsea-msigdb.org/gsea/msigdb), the "h.all.v7.4.symbols" and "c2.cp.kegg.v7.4.symbols" gene sets were downloaded for running the GSVA analysis. Pathway activation scores were compared between the two groups by the R package "limma", and adjusted *p*-values less than 0.05 were considered statistically significant.

2.6 Identification of m^{6}A phenotype-related genes

To identify genes mediated by m^{6}A regulators, differentially expressed genes (DEGs) between distinct m^{6}A phenotypes were analyzed by the empirical Bayesian method of the "limma" R package, and the cutoff criterion for screening DEGs were set as adjusted *p*-value <0.01. The biological functions of m^{6}A phenotype-related genes were analyzed by GO and KEGG enrichment analysis using the R package "clusterProfiler" (Yu et al., 2012). WGCNA (weighted gene coexpression network analysis) was used to identify the modification pattern-related gene modules through the "WGCNA" package (Langfelder and Horvath, 2008).

3 Results

3.1 Landscape of m^{6}A regulators between healthy and osteoarthritis samples

A total of 23 m^{6}A regulators were involved in this study, including 8 writers, 13 readers, and 2 erasers. An overview of m^{6}A regulators and their functions was given (Figure 1A). By analyzing the transcriptome matrix, we found a close correlation between most m^{6}A regulators in all samples as well as in osteoarthritis samples (Figures 1B, D). This illustrated that these regulators influenced each other, and this correlation was approximately the same across all samples as well as across disease samples. Among them, VIRMA and LRPPRC were the most highly correlated m^{6}A regulators expressed in all samples and osteoarthritis samples and showed a positive correlation, whereas VIRMA and RBM15B were the most negatively correlated (Figures 1C, E). These suggested that they function together. We constructed a correlation network of m^{6}A regulators in osteoarthritis (Figure 1F), again verifying that they generally function as a group. Differential expression analysis identified 6 m^{6}A regulators with altered expression. These factors with altered expression were distributed among the writers and readers, whereas the erasers did not change significantly, suggesting that they might not play an important role in osteoarthritis independently (Figure 1G).

3.2 m^{6}A regulators contribute to the osteoarthritis process

We employed a series of statistical algorithms to explore the impact of m^{6}A modification on osteoarthritis pathogenesis. We found that 12 m^{6}A regulators were associated with osteoarthritis by univariate logistic regression (Figure 2A; see Supplementary Table S3 in the Supplementary Material). LASSO regression was performed on 12 m^{6}A regulators for feature selection and

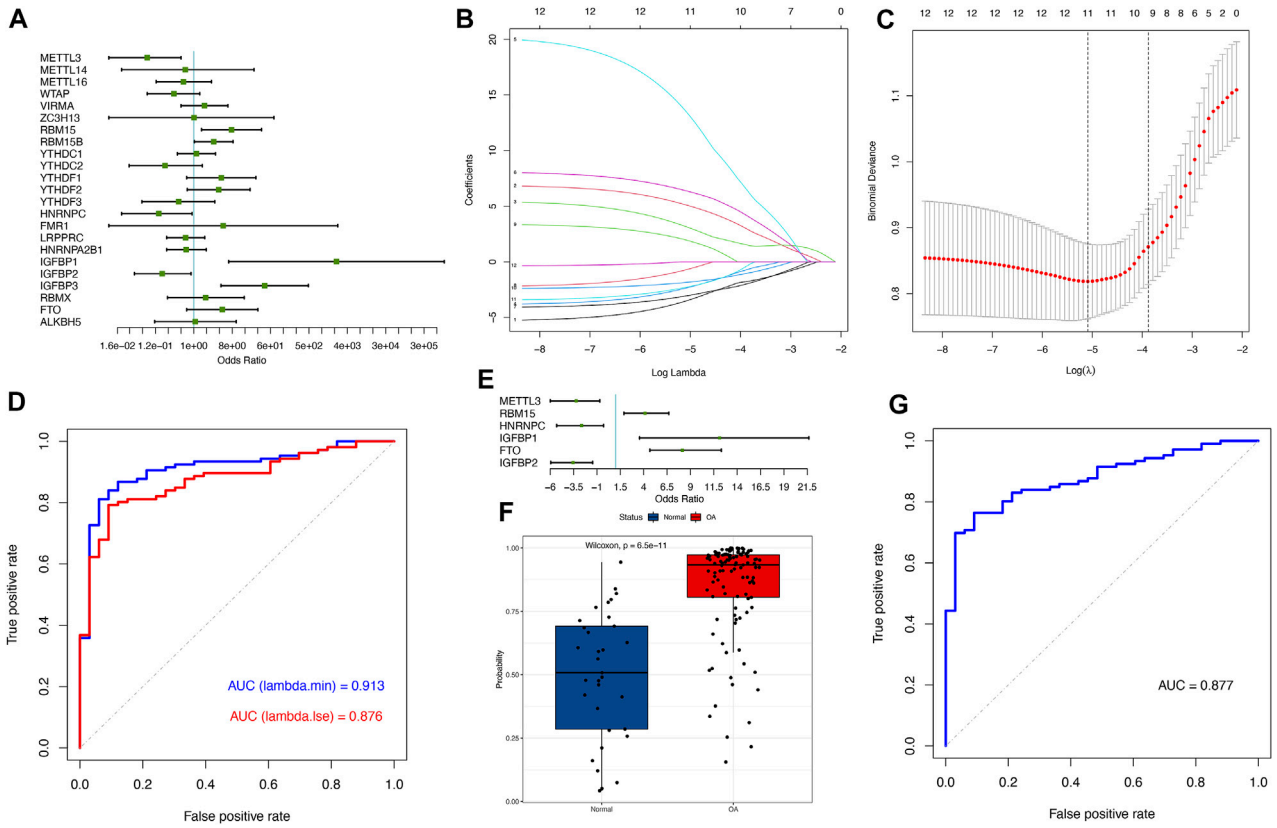


FIGURE 2
m^{6}A regulators have profound effects on the osteoarthritis process. **(A)** Univariate logistic regression presenting the relationship between m^{6}A regulators and osteoarthritis revealed 12 OA-related m^{6}A regulators ($p < 0.2$). **(B)** LASSO coefficient distribution of 12 OA-associated m^{6}A regulators. **(C)** Cross-validation for tuning parameter selection in the LASSO regression. For each λ value, around the mean of the target parameter shown by the red dot, we obtained the confidence interval of one target parameter. Two dashed lines indicate two special λ values: lambda. min was referred to in all λ values, the one giving the minimum target parametric mean, and lambda.1se refers to the one that gave the simplest model with a variance in the lambda. min λ value. **(D)** The ROC curve illustrated that the LASSO model, which includes 12 m^{6}A regulators according to MSE, has better performance in distinguishing between normal and osteoarthritis samples than 1-SE, with an AUC value of 0.913. **(E)** Multivariate logistic regression presenting the relationship between m^{6}A regulators and osteoarthritis revealed 6 osteoarthritis-related m^{6}A regulators ($p < 0.05$). **(F)** Boxplot showing the probability difference between normal and osteoarthritis samples, where osteoarthritis had a much higher probability score than normal samples. **(G)** The ROC curve depicting that the classification model based on 6 m^{6}A regulators has good prediction performance.

dimensionality reduction to exclude non-significant regulators (Figures 2B, C). We found that all 12 m^{6}A regulators were essential for osteoarthritis. When lambda with the minimum squared error (MSE) was employed, 12 variables were obtained, and 8 were obtained when the one-fold standard error (1-SE) was chosen. The ROC curve indicated the higher precision of the former 12 regulators in distinguishing disease from the normal group (AUC = 0.913) (Figure 2D). Multivariate logistic regression was performed to develop a categorical model to discriminate between normal and osteoarthritis samples (Figure 2E; see Supplementary Table S4 in the Supplementary Material). The model consisted of 6 m^{6}A regulators and discriminated well between healthy and osteoarthritis samples based on predicted probability values, where the probability scores of osteoarthritis were significantly higher than those of healthy samples (Figure 2F). The ROC curve illustrated the excellent performance of the 6 m^{6}A regulators in classifying health and osteoarthritis, indicating their diagnostic value for OA (Figure 2G). We found that our model based on the expression of 6 m^{6}A regulators had excellent diagnostic performance in other data sets (Supplementary Figure S1).

3.3 m^{6}A regulators are associated with osteoarthritis immune characteristics

To investigate the biological behaviors between m^{6}A regulators and the immune microenvironment, we performed correlation analyses between m^{6}A regulators and immune infiltration cells, immune responses, and immune functions. First, the results revealed differences in the abundance of 23 infiltrating immune cells between healthy and osteoarthritis samples (See Supplementary Figure S2 in the Supplementary Material for comprehensive image analysis). Some immune cells changed in osteoarthritis, such as myeloid-derived suppressor cells (MDSCs), T follicular helper cells, type 1 T helper cells and type 17 T helper cells, involving innate immunity and adaptive immunity. Correlation analysis identified that m^{6}A regulators were closely associated with these immune cells (Figure 3A). For example, the MDSC cell abundance was positively correlated with RBM15B and negatively correlated with HNRNPC (Figures 3B, C). This showed that infiltrating MDSCs were increased in osteoarthritis, which was closely related to the expression of RBM15B and HNRNPC.

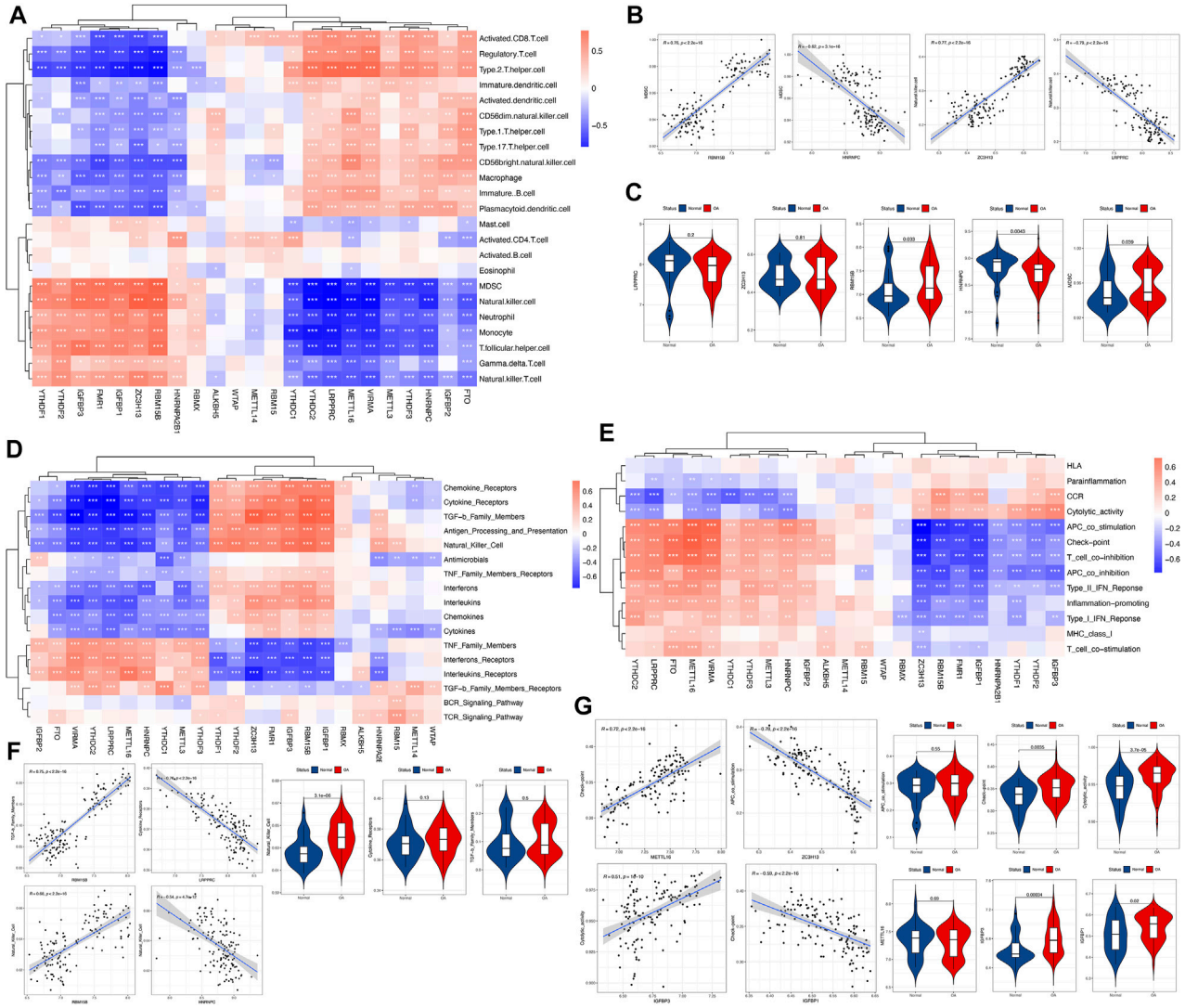


FIGURE 3
Correlations between the immune cell infiltration fraction, immune response gene sets, immune function gene sets, and m^{6}A regulators. **(A)** Heatmap demonstrating the correlations between each immune infiltration cell type and each m^{6}A regulator. Red indicates a positive correlation, and blue indicates a negative correlation. **(B,C)** Scatterplot demonstrating the correlations between the dysregulated immune cell fraction and the m^{6}A regulator. The fraction or expression status is presented as a violin plot, indicating that there were more MDSCs, higher expression of RBM15B, and lower expression of HNRNPC in osteoarthritis. **(D,F)** Heatmap demonstrating the correlations between each immune response gene set and each m^{6}A regulator. For dysregulated m^{6}A regulators, the most positively correlated pair was natural killer cell-RBM15B, and the most negatively correlated pair was natural killer cell-HNRNPC. There was a more active natural killer cell reaction in osteoarthritis, as presented by the violin plot. **(E,G)** Heatmap demonstrating the correlations between each immune function gene set and each m^{6}A regulator. For dysregulated m^{6}A regulators, pairs with a stronger positive correlation were cytolytic activity-IGFBP3, whereas a negative correlation was found for check-point-IGFBP1, and there were stronger check-point and cytolytic activity functions activated in osteoarthritis, as presented by the violin plot.

Similarly, we analyzed the immune response in osteoarthritis. The differences in the activity of each immunoreaction gene set between the healthy and osteoarthritis samples are presented (See Supplementary Figure S3 in the Supplementary Material for comprehensive image analysis). Several immune responses were increased in osteoarthritis, such as natural killer cell activity and TNF family members receptors. Natural killer cell activity was positively correlated with RBM15B, but it was negatively correlated with HNRNPC (Figures 3D, F). This suggested that RBM15B and HNRNPC played important roles in the natural killer cell response of osteoarthritis. We observed that RBM15B had a positive regulatory effect on multiple immune responses, while HNRNPC showed a negative regulatory effect. We also investigated the active state of immune function, in which half of the immune functions were altered in patients with osteoarthritis (See Supplementary Figure S4 in the Supplementary Material for comprehensive image analysis). For example, the check-point, cytolytic activity, and T-cell costimulation scores were higher in osteoarthritis. We found that METTL16 and ZC3H13 were positively and negatively correlated with most immune functions, respectively (Figures 3E, G). Check-point-METTL16 was the most positively correlated pair, and the most negatively correlated pair was APC costimulation-ZC3H13

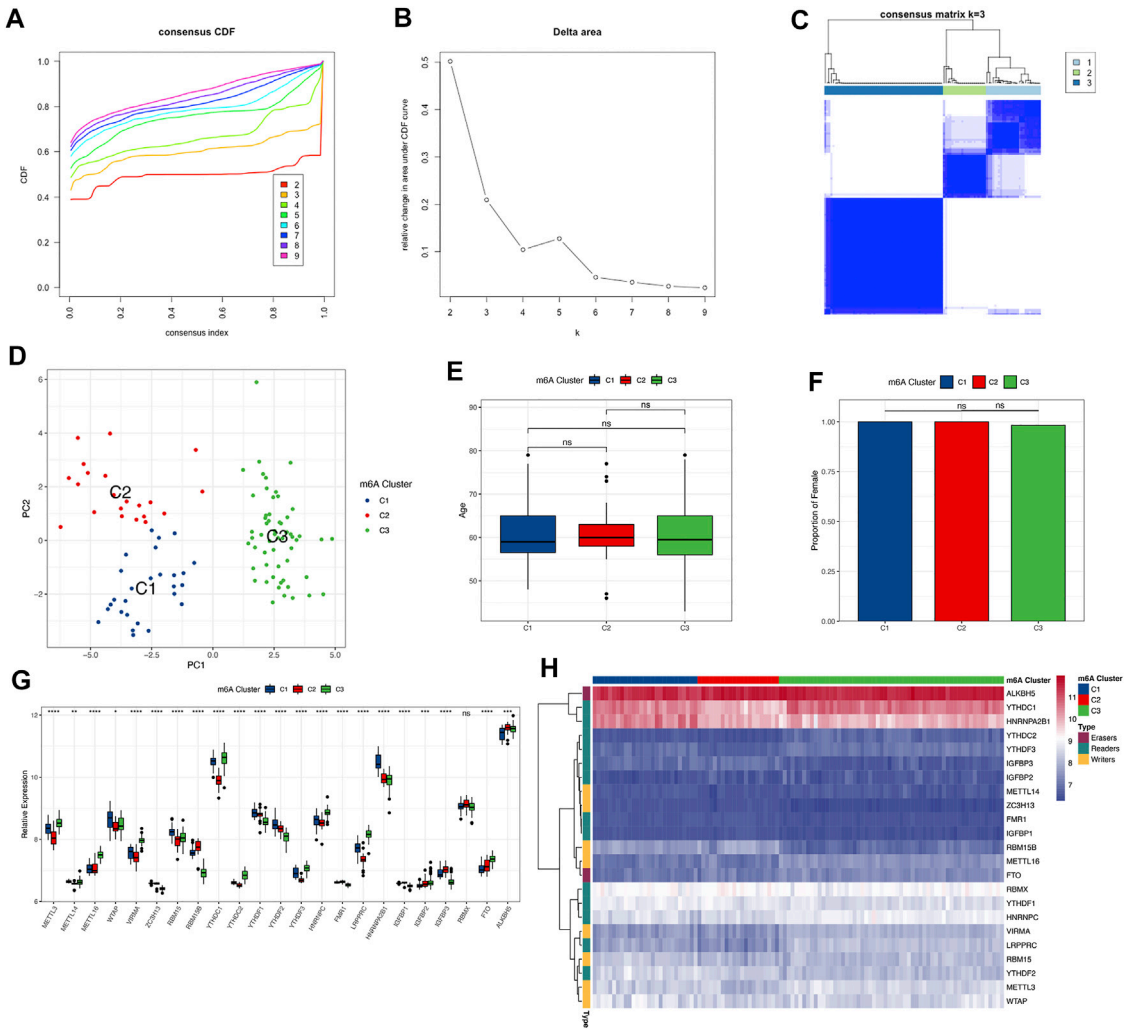


FIGURE 4
Identification of subtypes by unsupervised clustering based on the expression of 23 m^{6}A regulators. **(A)** Consensus clustering cumulative distribution function (CDF) for k = 2–9. **(B)** Relative change in the area under the CDF curve for k = 2–9. **(C)** Heatmap of the consensus matrix for osteoarthritis samples. **(D)** PCA of the transcriptome profiles of 3 m^{6}A subtypes, showing a remarkable difference in the transcriptome between different modification patterns. **(E,F)** Comparison of age and gender. The boxplot illustrated the association of age with the 3 subtypes. The bar plot illustrates the association of age with the 3 subtypes. Above ns means the difference was not statistically significant. **(G)** Expression differences of 23 m^{6}A regulators among the 3 m^{6}A subtypes. Above * indicates $p < 0.05$, ** indicates $p < 0.01$, *** indicates $p < 0.001$, **** indicates $p < 0.0001$, and ns indicates that the difference was not statistically significant. **(H)** Heatmap of the expression status of 23 m^{6}A regulators in the 3 subtypes with unsupervised clustering.

(Figure 3G). However, for dysregulated m^{6}A regulators, pairs with a stronger positive correlation were cytolytic activity-IGFBP3, whereas a negative correlation was with check-point-IGFBP1 (Figure 3G).

3.4 Patterns of m^{6}A methylation modification mediated by 23 regulators in osteoarthritis

To investigate the m^{6}A modification patterns in osteoarthritis, we performed an unsupervised consensus cluster analysis of the osteoarthritis samples based on the expression of 23 m^{6}A regulators. Three different m^{6}A subtypes of osteoarthritis were identified, including 27 samples in subtype C1, 21 samples in subtype C2, and 58 samples in subtype C3 (Figures 4A–C). The results of PCA confirmed that the 23 m^{6}A regulators could discriminate the 3 subtypes well (Figure 4D). The 3 distinct modification patterns differed from the current osteoarthritis classification, with no significant differences in clinical features between the different modification patterns, such as sex and age (Figures 4E, F). Except for RBMX, all m^{6}A regulators showed significant differences in their expression among the 3 m^{6}A subtypes (Figure 4G). The 23 m^{6}A regulators could still be divided into 3 parts according to their expression levels (Figure 4H), verifying the diversity of m^{6}A modification patterns in osteoarthritis.

3.5 Immune characteristics of 3 distinct m^{6}A subtypes

To determine the differences in immune microenvironment features among these different m^{6}A modification patterns, infiltrating immune cells, immune response gene sets, and immune

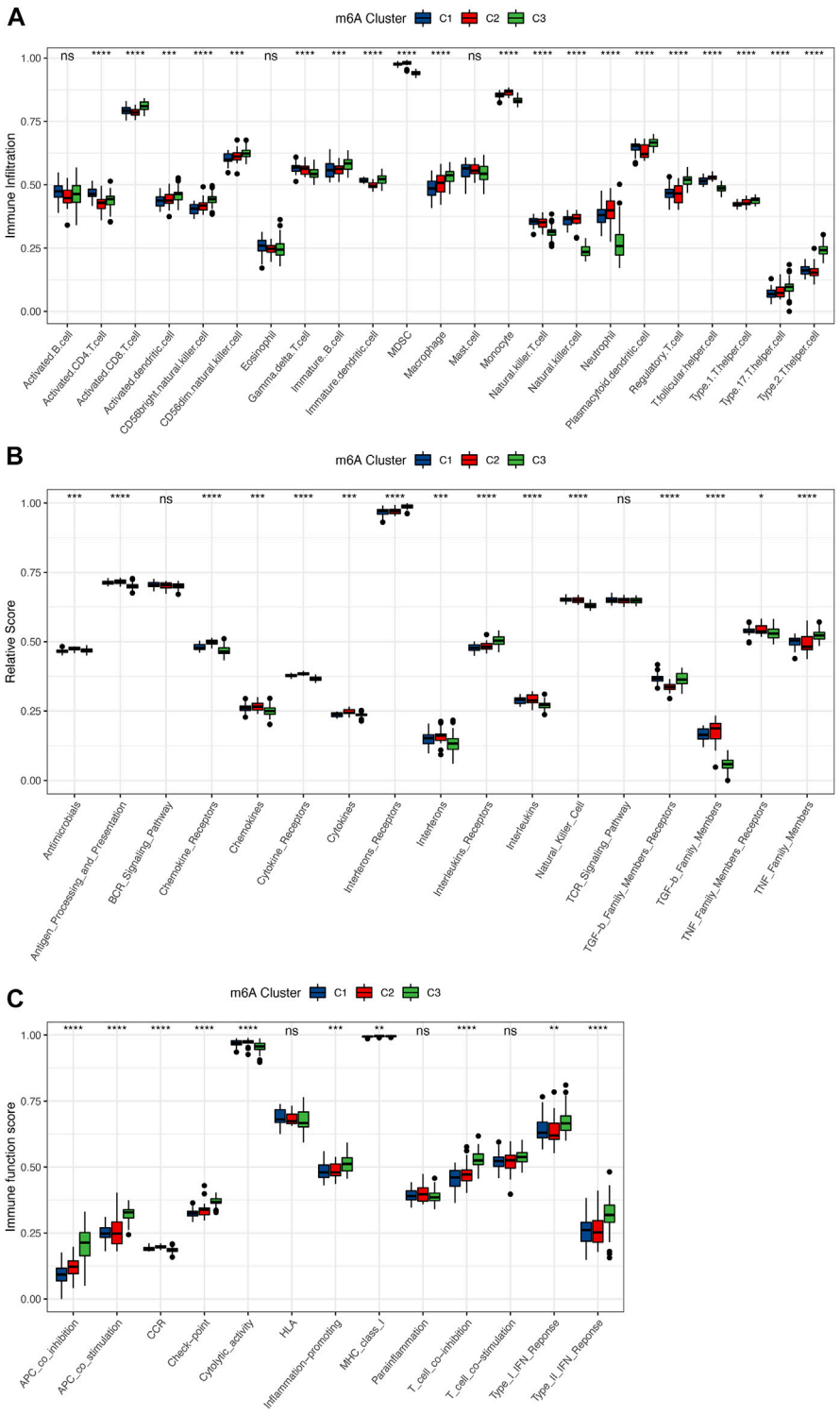


FIGURE 5
Immune microenvironment characteristics among 3 distinct m^{6}A subtypes. **(A)** Differences in the abundance of each immune cell infiltration in 3 m^{6}A subtypes. **(B)** Activity differences of each immune response gene set in 3 m^{6}A subtypes. **(C)** Expression differences of each immune function gene set in 3 m^{6}A subtypes. Above * indicates $p < 0.05$, ** indicates $p < 0.01$, *** indicates $p < 0.001$, **** indicates $p < 0.0001$, and ns indicates that the difference was not statistically significant.

function gene sets were assessed, and we found that the immune features were different among the three groups. The vast majority of the immune cells were distinct in the 3 patterns (Figure 5A). Subtype C3 had relatively higher infiltrating immune cells than subtypes C1 and C2, and the immune cell infiltration status of C1 was closer to that of C2. Subtype C3 had higher levels of activated CD8 T Cells, CD56bright natural killer cells, immune B Cells and macrophages, whereas monocytes, natural killer cells and MDSCs

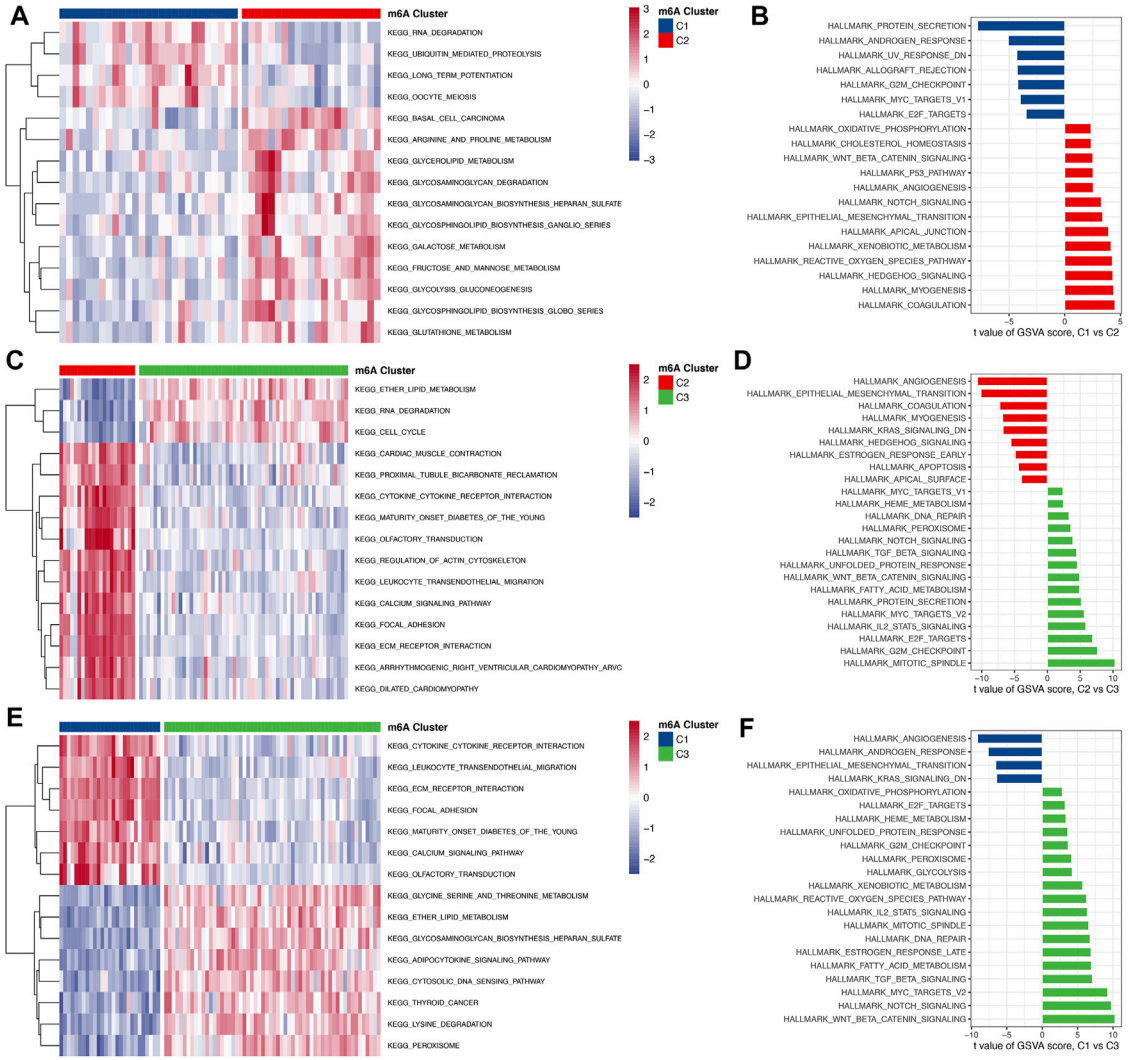


FIGURE 6
The biological function characteristics among the 3 m^{6}A subtypes. **(A,B)** Differences in GSVA enrichment scores for the KEGG pathway and HALLMARK pathway between m^{6}A cluster 1 and cluster 2 (A for the KEGG pathway and B for the HALLMARK pathway). **(C,D)** Differences in GSVA enrichment scores for the KEGG pathway and HALLMARK pathway between m^{6}A cluster 2 and cluster 3 (A for the KEGG pathway and B for the HALLMARK pathway). **(E,F)** Differences in GSVA enrichment scores for the KEGG pathway and HALLMARK pathway between m^{6}A cluster 1 and cluster 3 (A for the KEGG pathway and B for the HALLMARK pathway).

were enriched in subtype C2. A similar pattern in terms of immunoreactivity was observed, with more complex results. The immunoreactivity of subtype C3 differed from those of subtypes C1 and C2, with the status of C1 and C2 more similar, whereas the immunoreactivity of C2 was more active. For example, chemokines, cytokine receptor, and cytokines were more active in subtype C2, while interferon receptors and interleukins were more active in subtype C3, and TGF-b family members were much lower than in C1 and C2 (Figure 5B). Similar trends were also observed in the immune function scores (Figure 5C). The differences in terms of immune function were greater in subtype C3 than in C1 and C2. These results suggested that m^{6}A modification of subtype C3 mediated a unique immune inflammatory response that was distinguished from subtypes C1 and C2, whereas subtypes C1 and C2 also mediated distinct immune responses. The above results once again demonstrated that m^{6}A methylation modification had an important regulatory effect on the formation of different immune microenvironments in osteoarthritis.

3.6 Biological properties of the 3 m^{6}A modification patterns

To investigate the biological responses in the 3 m^{6}A subtypes, we compared KEGG pathways and HALLMARKS pathways between each of them and applied GSVA enrichment analysis to evaluate the activation status of biological pathways. Compared with subtypes C1 and C3, subtype C2 had more enriched pathways, such as the ECM receptor interaction, calcium signaling, cytokine to cytokine interaction receptor, and leukocyte transendothelial migration pathways (Figures 6A–D). Subtypes C1 and C3 had almost the same number of enriched pathways compared with each other

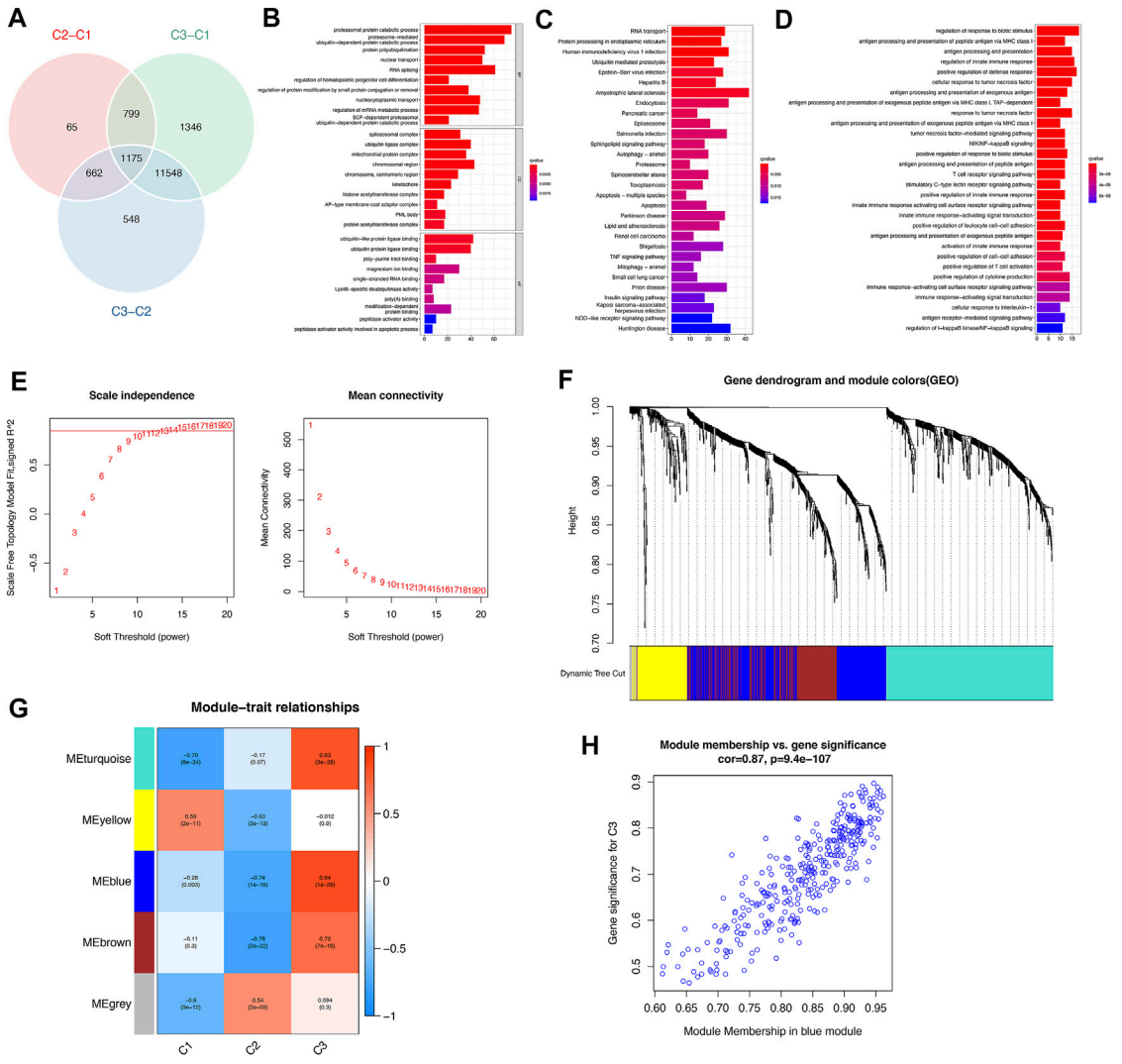


FIGURE 7
Identification and functional analysis of m^{6}A phenotype-related genes in osteoarthritis. **(A)** Venn diagram of 1175 m^{6}A phenotype-related genes. **(B)** GO enrichment analysis revealed the biological function characteristics of m^{6}A phenotype-related genes. **(C)** KEGG enrichment analysis revealed the biological signaling pathways of m^{6}A phenotype-related genes. **(D)** GO-BP enrichment results revealed the biological processes (BPs) of m^{6}A modification-mediated immune genes. **(E)** Analysis of the scale-free topology model fit index and the mean connectivity for various soft-thresholding powers. **(F)** Gene dendrogram obtained by average linkage hierarchical clustering. The color row underneath the dendrogram shows the module assignment determined by the dynamic tree cut, in which 5 modules were identified. **(G)** Heatmap of the correlation between characteristic gene modules and 3 distinct m^{6}A subtypes: red indicates a positive correlation, and blue indicates a negative correlation. **(H)** The scatterplot of gene significance for m^{6}A subtype C3 *vs.* module membership in the blue module. The gene significance and module membership exhibited highly significant correlations, implying that the hub genes of the blue module also tended to be highly correlated with subtype C3.

(Figures 6E, F). To further understand the molecular mechanisms by which genes were involved in the regulation mediated by m^{6}A regulators, we obtained differential gene intersections between m^{6}A phenotypes to obtain phenotype-related genes. As a result, a total of 1175 m^{6}A phenotype-related genes were obtained (Figure 7A), and GO enrichment analysis revealed that they were mainly involved in RNA splicing, nuclear transport, and regulation of mRNA metabolic processes (Figure 7B).

KEGG enrichment analysis showed that signaling pathways mainly involved RNA transport, protein processing in the endoplasmic reticulum, the TNF signaling pathway, and the NOD–like receptor signaling pathway (Figure 7C). We extracted 57 immune genes from these m^{6}A phenotype-related genes (See Supplementary Table S5 in the Supplementary Material), of which the enriched biological processes were remarkably related to the regulation of the response to biotic stimulus, regulation of the innate immune response, cellular response to tumor necrosis factor, and T Cell receptor signaling pathway (Figure 7D). We then constructed a comprehensive gene map associated with m^{6}A modification patterns, and WGCNA identified gene-gene modules associated with distinct m^{6}A subtypes (Figures 7E, F). Five gene modules were identified, and distinct modification patterns matched their related genes (Figure 7G; see Supplementary Table S6 in the Supplementary Material). For example, genes in the blue module were highly correlated with the m^{6}A regulator modification pattern C2 (Figure 7H). These genes were highly correlated not only

with their corresponding modules but also with their corresponding subtypes, further illustrating that genes deserve deep exploration. These results might elucidate the gene expression regulatory network mediated by m^{6}A regulators.

4 Discussion

Osteoarthritis is a chronic degenerative disease with a complex pathological mechanism that has not been clarified thus far. The understanding of this process has gone beyond mechanical wear and tear, in which inflammatory processes and immune responses also exist (Woodell-May and Sommerfeld, 2020). Increasing evidence has confirmed the indispensable role of m^{6}A modification in innate and adaptive immune responses (Zheng et al., 2017). To date, studies have been carried out to explore the role of m^{6}A in immunity, especially in tumor microenvironment infiltrating cells (Han et al., 2019; Yang et al., 2019; Zhang et al., 2019). Therefore, we believe that similar results can be observed in the regulation of the immune microenvironment of osteoarthritis by m^{6}A modification. In this study, we systematically investigated the modification pattern of m^{6}A in the immune microenvironment of osteoarthritis. To clarify how m^{6}A modification shapes the immune cell infiltration, immune response, immune function, and activation pathway of osteoarthritis, we conducted a series of analyses and obtained the following findings.

First, we found that compared with the normal samples, the expression of some m^{6}A regulators was out of balance in osteoarthritis. At the same time, there was a close relationship between the 23 regulators. We constructed a regulatory network of m^{6}A regulators, which indicates that m^{6}A regulators interact with each other and participate in the development of osteoarthritis. We used a series of multiple statistical approaches to screen out the significant m^{6}A regulators involved in osteoarthritis. The disease classification model based on these factors can distinguish healthy and osteoarthritis samples well, which confirms the important role of m^{6}A regulators in osteoarthritis. METTL3, HNRNPC, and IGFBP1 may be the most important among the 23 m^{6}A regulators, and they are of great significance in multivariate analysis. METTL3 has a functional role in mediating osteoarthritis progression by regulating NF-κB signaling and extracellular matrix (ECM) synthesis in chondrocytes (Liu et al., 2019). It has also been further confirmed that the expression of the m^{6}A methylated gene METTL3 is decreased in osteoarthritis and may be involved in osteoarthritis by regulating inflammatory responses (Sang et al., 2021). In addition, studies have shown that in pathological conditions, increased concentrations of IGF-I in joint synovial fluid are accompanied by increased levels of IGFBP-1 and IGFBP-3 (Matsumoto et al., 1996). These results are consistent with our findings in peripheral blood. However, the changes in HNRNPC in osteoarthritis have still not been specifically reported.

Next, we explored the correlations between m^{6}A regulators and immune characteristics in osteoarthritis, including scores for infiltrating immune cells, immune response, and immune function. We found that most m^{6}A regulators were closely associated with these immune characteristics, implying an important role of m^{6}A modification in the regulation of the osteoarthritis immune microenvironment. For example, MDSC abundance was strongly positively correlated with RBM15B and negatively correlated with HNRNPC. MDSCs can inhibit body immune cells to exert regular innate and adaptive immune functions. In the context of innate immunity, MDSCs downregulate the expression of NKG2D by membrane-bound TGF-β, which inhibits the function of NK cells (Li et al., 2009). MDSCs can also induce Treg expansion and promote the negative regulatory effect of Treg on immunity (Serafini et al., 2008). In terms of adaptive immunity, MDSCs can inhibit T Cell immune response responses and proliferation through multiple pathways (Rodríguez and Ochoa, 2008). Studies have found that MDSCs are significantly expanded in arthritic mice and RA patients. The transfer of MDSCs promotes disease progression, and proinflammatory MDSCs with the ability to drive Th17 cell differentiation may be a key pathogenic factor in autoimmune arthritis (Guo et al., 2016). RBM15B is reported to recruit this complex to certain mRNA and lncRNA XIST to promote m^{6}A formation (Coker et al., 2020). HNRNPC plays a cancer-promoting role in adrenocortical carcinoma (ACC) progression, and experiments have demonstrated that HNRNPC promotes the proliferation, migration, and invasion of H295R and SW13 cells and influences the immune microenvironment (Xu et al., 2021). These findings may point to an immunoregulatory mechanism of m^{6}A in osteoarthritis.

Third, unsupervised clustering of osteoarthritis samples based on the expression profiles of 23 m^{6}A regulators identified 3 subtypes with unique m^{6}A modification patterns, each with unique immune characteristics. Among them, subtype C3 had more infiltrating immune cells and more active immune functions than subtype C1 and subtype C2, and a portion of the immune response was more active in C3. We confirmed the reliability of phenotypic classification of different m^{6}A alterations by contrasting immune properties across the subtypes. The inspiration for this approach stems from a recent high-quality study in which a team used this approach to identify 3 distinct novel m^{6}A modification patterns in gastric cancer, gaining a deeper understanding of the tumor microenvironment (Zhang et al., 2020a). Identifying new molecular subtypes will not only unearth new pathogenesis but also enable the development of more precise treatment regimens. For osteoarthritis, Coutinho de Almeida R *et al* performed an unsupervised cluster analysis based on the top 1000 gene expressions deregulated in osteoarthritis, resulting in 2 distinct osteoarthritis subtypes possessing distinct cartilage pathophysiological processes as well as radiological features (Coutinho de Almeida et al., 2021). Thus, the 3 different m^{6}A subtypes in osteoarthritis suggest that the m^{6}A modification patterns present in peripheral blood can indeed be considered another pathobiology-based classification of osteoarthritis, which is related to the phenotypic features of the disease.

Finally, we identified m^{6}A phenotype-associated genes and m^{6}A modification subtype-associated gene modules. The expression regulation of these genes and gene sets is affected by m^{6}A modification, and revealing their biological functions can help illustrate the pathogenesis of osteoarthritis from the perspective of m^{6}A modification. Subtype C3 had more activation in the well-known TGF-β signaling pathway, while decreased NK cell infiltration was seen in subtype C3, and IGFBP1 was downregulated in subtype C3. These results may suggest that IGFBP1, the TGF-β signaling pathway, and NK cells are strongly implicated in osteoarthritis. The results of our study can give many of these similar correlations, and other researchers in the field will be directed to rapidly capture key m^{6}A regulators and immune signatures in osteoarthritis. This is one of the most important scientific implications of our study.

Although there is no consensus on the immune characteristics of OA, more and more studies have shown that immune inflammation is closely related to the pain and pathological progress of OA in recent years (Zhang et al., 2020b; Miller et al., 2020; Woodell-May and Sommerfeld, 2020; Li et al., 2022). And immunoengineering is expected to become the next-generation of arthritis treatment method (Klimak et al., 2021). Our study is the first to systematically analyze the relationship between m^{6}A modification and the immune microenvironment of osteoarthritis, and we are also the first team to introduce the latest m^{6}A mechanisms in osteoarthritis. Through this study, we obtained a wealth of results that can open new directions for studying the immune-related pathogenesis of osteoarthritis from the perspective of m^{6}A modification mechanisms. In addition, we confirmed that m^{6}A modification is involved in the regulation of the immune microenvironment in osteoarthritis. Current correlative studies of m^{6}A regulators in the osteoarthritis field are tenuous, and thus, this osteoarthritis research was seminal. We combined the latest m^{6}A modification and immune microenvironment theory to unravel osteoarthritis pathogenesis, largely complementing the gap in osteoarthritis regarding epigenetic modifications, particularly m^{6}A modification. This study will motivate more researchers to carry out m^{6}A-related research in the field of osteoarthritis, and the numerous results of this research can provide a better direction for them.

However, the study has some drawbacks. Firstly, this study was based on bioinformatics analysis, and many of the results were valid in theory, but their accuracy needs to be verified experimentally. Immune cell fraction were calculated by using currently accepted methods, but single-cell sequencing is still required to obtain the most accurate immune cell count. Secondly, some clinical characteristics were not available, such as Osteoarthritis Research Society International (OARSI) score, visual analog scale (VAS) score, joint range of motion and radiographic staging. All these made it difficult to reveal the correlation between clinical severity or outcome and the diversity of m^{6}A modification patterns. We also could not perform the analyses that associate the m^{6}A-mediated gene expression regulatory network with the progression of OA. In addition, because of the lack of clinical efficacy data, we failed to reflect the advantages of m^{6}A patterns compared with current diagnosis methods and its help for follow-up treatment. We hope to obtain data in the future and analyze them from the perspective of multiomics to obtain more valuable results. Thirdly, it is worth noting that expression level of m^{6}A regulators is not identical to m^{6}A methylation level and clinical samples are required for methylation level detection. Fourthly, Many studies were limited to gene regulation mediated by m^{6}A, but ignored the mechanism of regulating m^{6}A precipitation (Yang et al., 2022). For example, aging and inflammation are not only pathogenic factors, but also the result of m^{6}A dysfunction. Elucidating the complex regulatory mechanism of m^{6}A is helpful for the targeted treatment of bone related diseases. Nevertheless, these results enhance the understanding of the novel pathogenesis and phenotypes of osteoarthritis and provide new ideas for promoting personalized immunotherapy in the future.

5 Conclusion

This study reveals a potential regulatory mechanism of m^{6}A methylation modification in the immune microenvironment of osteoarthritis. The diversity of m^{6}A modification patterns is a factor contributing to the heterogeneity and complexity of the osteoarthritis immune microenvironment that cannot be ignored. Comprehensive analysis of m^{6}A modification patterns in osteoarthritis allows us to gain a deeper understanding of the underlying mechanisms of osteoarthritis immune regulatory networks and guide more effective precision therapies.

Data availability statement

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

Author contributions

YG, LX, and YSY established the concept of this study. YG, ZMW, and PJT performed the literature research, data extraction, statistical analysis, and drafting paper. RW and LX revised this paper. LX, ZRW, and SJL contributed to the review of the paper and supervision of the whole study. YG and ZMW contributed equally to this work. All authors have read and approved the manuscript.

Funding

This study was supported by the National Natural Science Foundation of China (82074473, and 82104892), Natural Science Foundation of Jiangsu province (BK20180001, BK20191201, and BE2020666), Project of Excellent Young Talents of Traditional Chinese Medicine of Zhejiang Province (2019ZQ016) and the "Zhejiang Medical and Health Science and Technology Young Talents Program (2019RC059)" from the Health Commission of Zhejiang Province.

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

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2023.1113515/full#supplementary-material

References

Abramoff, B., and Caldera, F. E. (2020). Osteoarthritis: Pathology, diagnosis, and treatment options. *Med. Clin. North Am.* 104, 293–311. doi:10.1016/j.mcna.2019.10.007

arc, O. C., arc, O. C., Zeggini, E., Panoutsopoulou, K., Southam, L., Rayner, N. W., et al. (2012). Identification of new susceptibility loci for osteoarthritis (arcOGEN): A genome-wide association study. *Lancet* 380, 815–823. doi:10.1016/S0140-6736(12)60681-3

Bhattacharya, S., Andorf, S., Gomes, L., Dunn, P., Schaefer, H., Pontius, J., et al. (2014). ImmPort: Disseminating data to the public for the future of immunology. *Immunol. Res.* 58, 234–239. doi:10.1007/s12026-014-8516-1

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

Coker, H., Wei, G., Moindrot, B., Mohammed, S., Nesterova, T., and Brockdorff, N. (2020). The role of the Xist 5' m6A region and RBM15 in X chromosome inactivation. *Wellcome Open Res.* 5, 31. doi:10.12688/wellcomeopenres.15711.1

Coutinho de Almeida, R., Mahfouz, A., Mei, H., Houtman, E., den Hollander, W., Soul, J., et al. (2021). Identification and characterization of two consistent osteoarthritis subtypes by transcriptome and clinical data integration. *Rheumatol. Oxf.* 60, 1166–1175. doi:10.1093/rheumatology/keaa391

Daghestani, H. N., and Kraus, V. B. (2015). Inflammatory biomarkers in osteoarthritis. *Osteoarthr. Cartil.* 23, 1890–1896. doi:10.1016/j.joca.2015.02.009

Guo, C., Hu, F., Yi, H., Feng, Z., Li, C., Shi, L., et al. (2016). Myeloid-derived suppressor cells have a proinflammatory role in the pathogenesis of autoimmune arthritis. *Ann. Rheum. Dis.* 75, 278–285. doi:10.1136/annrheumdis-2014-205508

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

Hanzelmann, S., Castelo, R., and Guinney, J. (2013). GSVA: Gene set variation analysis for microarray and RNA-seq data. *BMC Bioinforma.* 14, 7. doi:10.1186/1471-2105-14-7

Hartigan, J. A., and Wong, M. A. (2013). A K-means clustering algorithm. *Appl. Stat.* 28, 100–108.

Harvey, Z. H., Chen, Y., and Jarosz, D. F. (2018). Protein-based inheritance: Epigenetics beyond the chromosome. *Mol. Cell* 69, 195–202. doi:10.1016/j.molcel.2017.10.030

Klimak, M., Nims, R. J., Pferdehirt, L., Collins, K. H., Harasymowicz, N. S., Oswald, S. J., et al. (2021). Immunoengineering the next generation of arthritis therapies. *Acta Biomater.* 133, 74–86. doi:10.1016/j.actbio.2021.03.062

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

Langfelder, P., and Horvath, S. (2008). WGCNA: an R package for weighted correlation network analysis. *BMC Bioinforma.* 9, 559. doi:10.1186/1471-2105-9-559

Li, H., Han, Y., Guo, Q., Zhang, M., and Cao, X. (2009). Cancer-expanded myeloid-derived suppressor cells induce anergy of NK cells through membrane-bound TGF-beta 1. *J. Immunol.* 182, 240–249. doi:10.4049/jimmunol.182.1.240

Li, M., Yin, H., Yan, Z., Li, H., Wu, J., Wang, Y., et al. (2022). The immune microenvironment in cartilage injury and repair. *Acta Biomater.* 140, 23–42. doi:10.1016/j.actbio.2021.12.006

Liang, J-Y., Wang, D-S., Lin, H-C., Chen, X-X., Yang, H., Zheng, Y., et al. (2020). A novel ferroptosis-related gene signature for overall survival prediction in patients with hepatocellular carcinoma. *Int. J. Biol. Sci.* 16, 2430–2441. doi:10.7150/ijbs.45050

Liu, Q., Li, M., Jiang, L., Jiang, R., and Fu, B. (2019). METTL3 promotes experimental osteoarthritis development by regulating inflammatory response and apoptosis in chondrocyte. *Biochem. Biophys. Res. Commun.* 516, 22–27. doi:10.1016/j.bbrc.2019.05.168

Matsumoto, T., Gargosky, S. E., Iwasaki, K., and Rosenfeld, R. G. (1996). Identification and characterization of insulin-like growth factors (IGFs), IGF-binding proteins (IGFBPs), and IGFBP proteases in human synovial fluid. *J. Clin. Endocrinol. Metab.* 81, 150–155. doi:10.1210/jcem.81.1.8550744

Miller, R. J., Malfait, A. M., and Miller, R. E. (2020). The innate immune response as a mediator of osteoarthritis pain. *Osteoarthr. Cartil.* 28 (5), 562–571. doi:10.1016/j.joca.2019.11.006

Moradi, B., Rosshirt, N., Tripel, E., Kirsch, J., Barié, A., Zeifang, F., et al. (2015). Unicompartmental and bicompartmental knee osteoarthritis show different patterns of mononuclear cell infiltration and cytokine release in the affected joints. *Clin. Exp. Immunol.* 180, 143–154. doi:10.1111/cei.12486

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

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

Ramos, Y. F., Bos, S. D., Lakenberg, N., Bohringer, S., den Hollander, W. J., Kloppenburg, M., et al. (2014). Genes expressed in blood link osteoarthritis with apoptotic pathways. *Ann. Rheum. Dis.* 73, 1844–1853. doi:10.1136/annrheumdis-2013-203405

Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., et al. (2015). Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* 43, e47. doi:10.1093/nar/gkv007

Rodríguez, P. C., and Ochoa, A. C. (2008). Arginine regulation by myeloid derived suppressor cells and tolerance in cancer: Mechanisms and therapeutic perspectives. *Immunol. Rev.* 222, 180–191. doi:10.1111/j.1600-065X.2008.00608.x

Rosshirt, N., Hagmann, S., Tripel, E., Gotterbarm, T., Kirsch, J., Zeifang, F., et al. (2019). A predominant Th1 polarization is present in synovial fluid of end-stage osteoarthritic knee joints: Analysis of peripheral blood, synovial fluid and synovial membrane. *Clin. Exp. Immunol.* 195, 395–406. doi:10.1111/cei.13230

Sang, W., Xue, S., Jiang, Y., Lu, H., Zhu, L., Wang, C., et al. (2021). METTL3 involves the progression of osteoarthritis probably by affecting ECM degradation and regulating the inflammatory response. *Life Sci.* 278, 119528. doi:10.1016/j.lfs.2021.119528

Serafini, P., Mgebroff, S., Noonan, K., and Borrello, I. (2008). Myeloid-derived suppressor cells promote cross-tolerance in B-cell lymphoma by expanding regulatory T cells. *Cancer Res.* 68, 5439–5449. doi:10.1158/0008-5472.CAN-07-6621

Shen, S., Wang, G., Zhang, R., Zhao, Y., Yu, H., Wei, Y., et al. (2019). Development and validation of an immune gene-set based Prognostic signature in ovarian cancer. *EBioMedicine* 40, 318–326. doi:10.1016/j.ebiom.2018.12.054

Shulman, Z., and Stern-Ginossar, N. (2020). The RNA modification N6-methyladenosine as a novel regulator of the immune system. *Nat. Immunol.* 21, 501–512. doi:10.1038/s41590-020-0650-4

Vincent, T. L. (2013). Targeting mechanotransduction pathways in osteoarthritis: A focus on the pericellular matrix. *Curr. Opin. Pharmacol.* 13, 449–454. doi:10.1016/j.coph.2013.01.010

Wilkerson, M. D., and Hayes, D. N. (2010). ConsensusClusterPlus: A class discovery tool with confidence assessments and item tracking. *Bioinformatics* 26, 1572–1573. doi:10.1093/bioinformatics/btq170

Woodell-May, J. E., and Sommerfeld, S. D. (2020). Role of inflammation and the immune system in the progression of osteoarthritis. *J. Orthop. Res.* 38, 253–257. doi:10.1002/jor.24457

Xu, F., Guan, Y., Ma, Y., Xue, L., Zhang, P., Yang, X., et al. (2021). Bioinformatic analyses and experimental validation of the role of m6A RNA methylation regulators in progression and prognosis of adrenocortical carcinoma. *Aging (Albany NY)* 13, 11919–11941. doi:10.18632/aging.202896

Yang, C., Dong, Z., Ling, Z., and Chen, Y. (2022). The crucial mechanism and therapeutic implication of RNA methylation in bone pathophysiology. *Ageing Res. Rev.* 79, 101641. doi:10.1016/j.arr.2022.101641

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

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

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

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, 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, H., Cai, D., and Bai, X. (2020). Macrophages regulate the progression of osteoarthritis. *Osteoarthr. Cartil.* 28 (5), 555–561. doi:10.1016/j.joca.2020.01.007

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

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

frontiers | Frontiers in Cell and Developmental Biology

TYPE Original Research
PUBLISHED 02 February 2023
DOI 10.3389/fcell.2023.1117637

OPEN ACCESS

EDITED BY
Lingzhao Fang,
University of Edinburgh, United Kingdom

REVIEWED BY
Siyuan Mi,
China Agricultural University, China
Tara G. McDaneld,
Agricultural Research Service (USDA), United States

*CORRESPONDENCE
Ling Guo,
guoling@sysucc.org.cn
Huanxin Lin,
linhx@sysucc.org.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 06 December 2022
ACCEPTED 23 January 2023
PUBLISHED 02 February 2023

CITATION
Jiang T, Wang Y, Chen X, Xia W, Xue S, Gu L, Guo L and Lin H (2023), Neutrophil extracellular traps (NETs)-related lncRNAs signature for predicting prognosis and the immune microenvironment in breast cancer.
Front. Cell Dev. Biol. 11:1117637.
doi: 10.3389/fcell.2023.1117637

Neutrophil extracellular traps (NETs)-related lncRNAs signature for predicting prognosis and the immune microenvironment in breast cancer

Tongchao Jiang[1†], Ying Wang[1†], Xiaoyu Chen[1], Wen Xia[2], Shuyu Xue[1], Liwen Gu[3], Ling Guo[3*] and Huanxin Lin[1*]

[1]State Key Laboratory of Oncology in South China, Guangdong Key Laboratory of Nasopharyngeal Carcinoma Diagnosis and Therapy, Department of Radiotherapy, Collaborative Innovation Center for Cancer Medicine, Sun Yat-sen University Cancer Center, Guangzhou, Guangdong, China, [2]State Key Laboratory of Oncology in South China, Guangdong Key Laboratory of Nasopharyngeal Carcinoma Diagnosis and Therapy, Department of Medical Oncology, Collaborative Innovation Center for Cancer Medicine, Sun Yat-sen University Cancer Center, Guangzhou, Guangdong, China, [3]State Key Laboratory of Oncology in South China, Guangdong Key Laboratory of Nasopharyngeal Carcinoma Diagnosis and Therapy, Department of Nasopharyngeal Carcinoma, Collaborative Innovation Center for Cancer Medicine, Sun Yat-sen University Cancer Center, Guangzhou, Guangdong, China

Background: Neutrophil extracellular traps (NETs) are closely associated to tumorigenesis and development. However, the relationship between NETs-related long non-coding RNAs (lncRNAs) and the characteristics of breast tumor remains an enigma. This study aimed to explore the clinical prognostic value of NETs-related lncRNAs, their correlation with the tumor microenvironment (TME) and their predictive ability of drug sensitivity in patients with breast cancer (BC).

Methods: The expression profiles of RNA-sequencing and relevant clinical data of BC patients were extracted from TCGA database. The co-expression network analysis, univariable, least absolute shrinkage and selection operator (LASSO) and multivariable Cox algorithms were employed to construct the NETs-related lncRNAs signature. A nomogram was established and validated to explore the clinical application. Furthermore, the immune microenvironment and drug sensitivity for BC with different prognostic risks were explored. Finally, the expression pattern of lncRNAs was validated using qRT-PCR in BC tissues and their adjacent non-cancerous tissues.

Results: Based on NETs-related lncRNAs, a prognostic risk model consisted of 10 lncRNAs (SFTA1P, ACTA2-AS1, AC004816.2, AC000067.1, LINC01235, LINC01010, AL133467.1, AC092919.1, AL591468.1, and MIR200CHG) was established. The Kaplan-Meier analysis showed that the overall survival (OS) was significantly better in low-risk BC patients than in high-risk BC patients ($P_{\text{training cohort}} < 0.001$, $P_{\text{validation cohort}} = 0.009$). The nomogram also showed good predictive accuracy for OS of BC individuals in both training and validation cohorts. The function enrichment analysis revealed that high-risk group was mainly enriched in immune-related functions and pathways, and the tumor mutation burden in this group was markedly higher than that in the low-risk group ($p = 0.022$). Moreover, significant differences were observed in immune cells, immune functions and immune checkpoint genes among BC patients at different risks ($p < 0.05$). The response to chemotherapeutic agents and immunotherapy were also closely related with the expression of NETs-related lncRNAs ($p < 0.001$). The expression of

lncRNAs from experimental validation were generally consistent with the bioinformatics analysis results.

Conclusion: Our study provided a novel prognostic model for BC and yielded strong scientific rationale for individualized treatment strategies, elucidating immunotherapy in BC patients.

KEYWORDS

breast cancer, NETs, lncRNA, prognosis, tumor microenvironment, anticancer drugs

Introduction

Breast cancer (BC) is the most prevalent and second most deadly malignancy in women, accounting for 31% of all newly diagnosed cancers (Siegel et al., 2022). At present, despite the advances in effective therapeutic strategies including surgical resection, endocrine therapy, and the combination of surgery with radiotherapy, chemotherapy and immunotherapy, BC still confronted with high morbidity, aggressiveness, metastasis, and recurrence rates (Pondé et al., 2019). Moreover, due to the remarkable tumor heterogeneity, breast cancers with the same subtypes can respond differently to therapy and have different prognosis (Pondé et al., 2019). Thus, specific molecular biomarkers and therapeutic targets for BC are pivotal elements to guide clinical practice.

Neutrophils, the most abundant endogenous immune effector cells, can respond to specific stimulation by releasing neutrophil extracellular traps (NETs), a type of regulated cell death termed "neutrophil extracellular traposis (NETosis)" (Ireland and Oliver, 2020). Primarily described as an antimicrobial mechanism for entrapping, constraining, and killing invading bacteria and other pathogens, NETs are complex extracellular networks composed of nuclear DNA fibers and mitochondria decorated with granular antimicrobial enzymes and histones (Papayannopoulos, 2018; Demkow, 2021). Subsequent studies have shown that NETs, forming a protective shield, have multiple pro-tumor capabilities, including primary growth and metastasis (Ireland and Oliver, 2020; Martins-Cardoso et al., 2020).

The long non-coding RNAs (lncRNAs), non-coding RNA longer than 200 nucleotides, do not directly participate in protein coding in cells, but are engaged in vital biological regulatory processes, including transcriptional regulation, mRNA processing regulation and mRNA post-transcriptional regulation (Yang et al., 2022). Recent finding indicated a role of lncRNA in regulating NETs in lung cancer (Wang et al., 2022). However, the lncRNAs associated with NETs in BC are less studied, and most lncRNAs regulating NETs have not been determined. Potential NETs-related biomarkers and prognostic biomarkers can be identified with the advances in high-throughput sequencing technologies and bioinformatics (Huang da et al., 2009).

Immune checkpoint blockade (ICB) therapy, relying on the immune tumor microenvironment, is an effective therapeutic strategy that blocks the immune checkpoint pathway to keep tumor cells from evading immune surveillance (Franzén et al., 2022). Unfortunately, most breast tumors with less tumor-infiltrating cytotoxic T cells and lower PD-L1 expression are usually considered as immune "cold" tumors and tend to have poorer efficacy with ICB therapy (Tekpli et al., 2019). Currently, a study has revealed that NETs have the ability to suppress T-cell responses in the tumor microenvironment through metabolism and functional exhaustion, thereby affecting immunotherapeutic efficacy (Kaltenmeier et al., 2021). Moreover, lncRNAs act as a key player in reshaping the immune landscape, regulating metabolic reprogramming, and functioning as a bond between tumor metabolism and anti-tumor immunity (Yang et al., 2022). Thus, exploring the interaction between NETs-related lncRNAs and tumor immune microenvironment can help to improve the understanding of the pathogenesis of "cold" breast cancer and offer potential therapeutic strategies for "cold" breast tumor.

In the current study, we first constructed a prognostic risk model composed of 10 NETs-related lncRNAs for BC patients based on public databases, and evaluated this risk model performance. Furthermore, the clinical significance and application value of the model and its effects on immune microenvironment and drug sensitivity were also explored. To the best of our knowledge, no previous studies have investigated the predictive value of NETs-related lncRNAs and their relationship with the immune microenvironment in BC. The present study identified NETs-related lncRNAs that may be potential therapeutic targets and prognostic and predictive markers for BC patients and could be used to further improve the treatment outcome of BC patients through individualized therapy.

Materials and methods

Data acquisition

The transcriptome RNA-seq data [fragments per kilobase of transcript per Million mapped reads (FPKM)] of 1,222 samples were obtained from the TCGA public database, including 1110 BC tissues and 112 normal adjacent tissues (https://portal.gdc.cancer.gov/repository). The corresponding clinicopathological data of BC were also downloaded from TCGA together. The protein-coding genes and lncRNAs were distinguished by applying the ensembl human genome browser GRCh38.p13. A total of 170 NETs-related genes were acquired from previously published studies (as shown in Supplementary Table S1) (Dwyer et al., 2014; Papayannopoulos, 2018). Then the correlation between the expression of NETs-related genes and corresponding lncRNAs was quantified by calculating the Pearson correlation coefficients. The strict criteria were used to identify NETs-related lncRNAs, with $p < 0.001$ and the absolute value of Pearson correlation coefficient more than 0.4 ($|R| > 0.4$).

Determination of differentially expressed NETs-Related LncRNAs

The expression profiles of NETs-related lncRNAs of the 112 normal breast samples and 1110 BC samples were obtained, and the differential expression analysis was carried out with $|\text{Log}_2$ fold change [FC]$|$ >1 and false discovery rate (FDR) <0.05 using the

"limma" and "pheatmap" R package. After deleting patients with incomplete information, 937 patients with breast cancer in total were randomly divided in an 8:2 ratio into training and validation cohorts for constructing and validating the NETs-related lncRNAs signature. Univariate Cox regression analysis for overall survival (OS) was employed to determine prognostic lncRNAs with $p < 0.05$ in the training cohort. The least absolute shrinkage and selection operator (LASSO) Cox regression algorithm was performed to lessen the chance of overfitting. Furthermore, multivariate Cox regression analysis was applied to calculate the regression coefficient of the prognostic risk score model.

Construction of the LncRNA-mRNA co-expression network

The lncRNA-mRNA co-expression network was constructed to demonstrate the correlation between the NETs-related lncRNAs and their corresponding mRNAs, and visualized using the Cytoscape software (version 3.7.2, http://www.cytoscape.org/).

Construction and validation of NETs-related LncRNAs signature

The risk score for each BC patient in both the training and validation cohorts was calculated according to the following equation:

$$Risk\ score = \sum_{i=1}^{\infty} coef i \times xi \quad (1)$$

where coefi is the regression coefficient and xi is the corresponding lncRNA expression level. The BC patients in the training cohort were separated into high- and low-risk groups based on the median risk score. The same median score was used to divide patients in the validation cohort into high- and low-risk groups. The Kaplan-Meier method, risk score heatmap, distribution of risk score and survival status were used to assess the validity of the prognostic risk model by applying the "survivalROC" and "pheatmap" R package.

Univariate and multivariate Cox regression algorithm were conducted to assess the prognostic significance of risk scores based on other clinical parameters (age, stage, and subtype) in the training cohort. The receiver operating characteristic (ROC) curves and C-index were performed to evaluate the predictive power of this signature. Then, a nomogram was constructed based on these parameters as implemented in the "rms" R package. The validity of prognostic model in their accuracy of prediction for 1-, 3-, and 5-years OS was further assessed by the calibration curves and time-dependent receiver operating characteristic (ROC) curves in the training and validation cohorts.

Functional enrichment analysis

To investigate the potential biological functions of the 10 NETs-related lncRNAs in BC, the gene set enrichment analysis (GSEA) was conducted to look for the tumor hallmarks associated with risk scores. The FDR <0.25 and *P* adjusted value <0.05 were adopted as the criteria for statistical significance. An enrichment lot was performed to visualize the top five functions enriched by two groups. After that, the differentially expressed genes between the two risk scores groups were determined by differential expression analysis of two groups with $|\log FC| > 1$ and $p < 0.05$. The pathway enrichment analysis for Gene ontology (GO) and kyoto encyclopedia of genes and genomes (KEGG) on the screened genes were conducted by employing the "clusterProfiler" R package. Then, the "enrichplot" and "ggplot2" R packages were applied to visualize the enrichment results.

Comprehensive analysis of molecular variation and immune infiltration

The 977 single nucleotide variants (SNVs) data were downloaded in the TCGA public database and the "Maftools" R package was applied to calculate the tumor mutation burden (TMB) for each BC patient. The correlation between the risk score and TMB was analyzed using the Spearman's algorithm. The Kaplan-Meier survival analysis was performed for different TMB groups. Furthermore, the CIBERSORT deconvolution algorithm with 1,000 permutations was applied to calculate the abundance of immune cell infiltration in the tumor immune microenvironment of high- and low-risk patients for BC with $p < 0.05$.

Evaluation of the immune cell infiltration, immune function, and immune checkpoint genes

Single-sample Gene Set Enrichment Analysis (ssGSEA) was performed to determine the biological function differences between the high- and low-risk groups by applying the "GSVA" R package. The potential immune checkpoint genes were chosen according to previous published literature. The different expression levels of immune checkpoint genes between risk groups were analyzed by the Wilcoxon test.

Evaluation of drug sensitivity and immunotherapy efficacy

To assess the therapeutic efficacy of chemotherapy and targeted agents in constructing prognostic risk signature of NETs-related lncRNAs, the half inhibitory concentration (IC_{50}) of chemotherapeutic and target therapeutic drugs were calculated using the "ggplot2" and "pRRophetic" R packages. The Wilcoxon signed-rank method was applied to compare the IC_{50} of different risk groups. Then, immunotherapy responsiveness was predicted in different risk groups by applying the Tumor Immune Dysfunction and Exclusion (TIDE) online Tool (http://tide.dfci.harvard.edu/).

Tissue specimens and quantitative real-time polymerase chain reaction

A total of 10 matched pairs of tumor specimens and adjacent normal tissues were obtained from BC patients who underwent tumor resection. All tissue specimens were collected from the

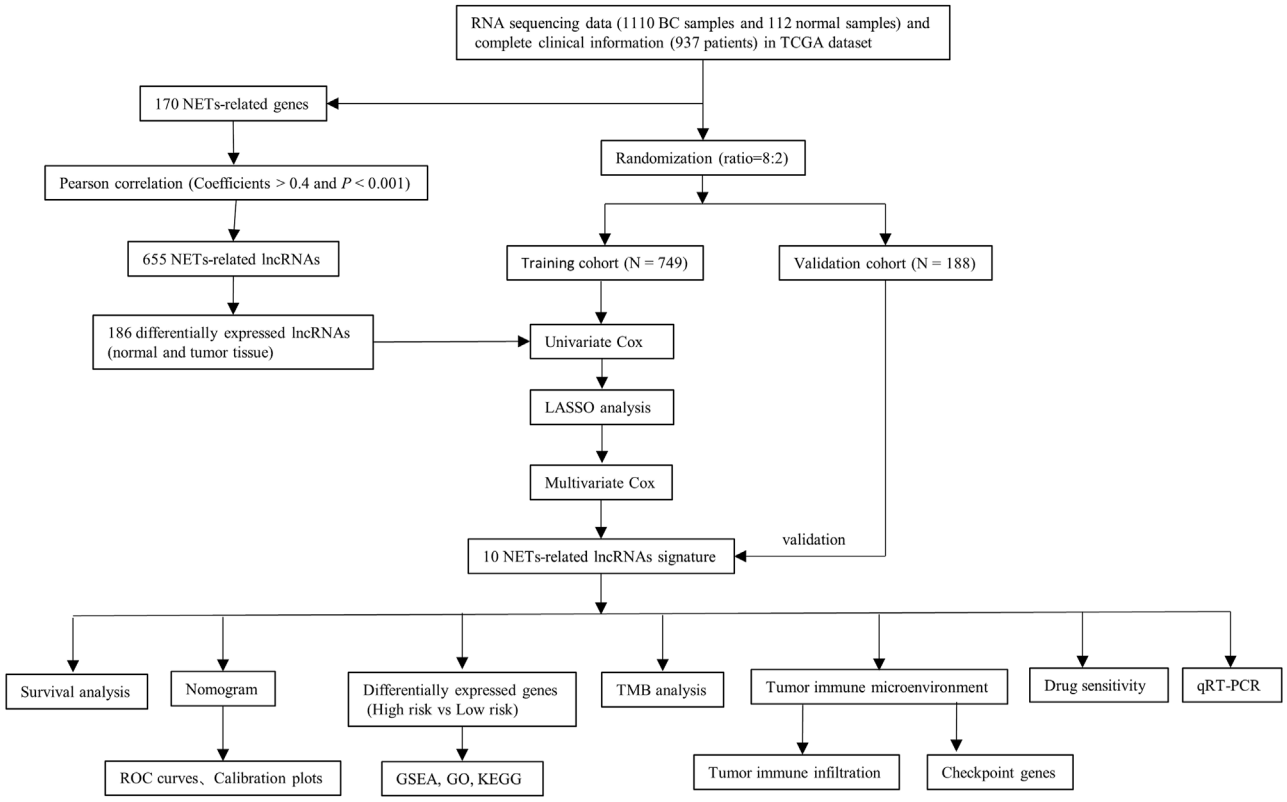


FIGURE 1
Flow diagram of the study.

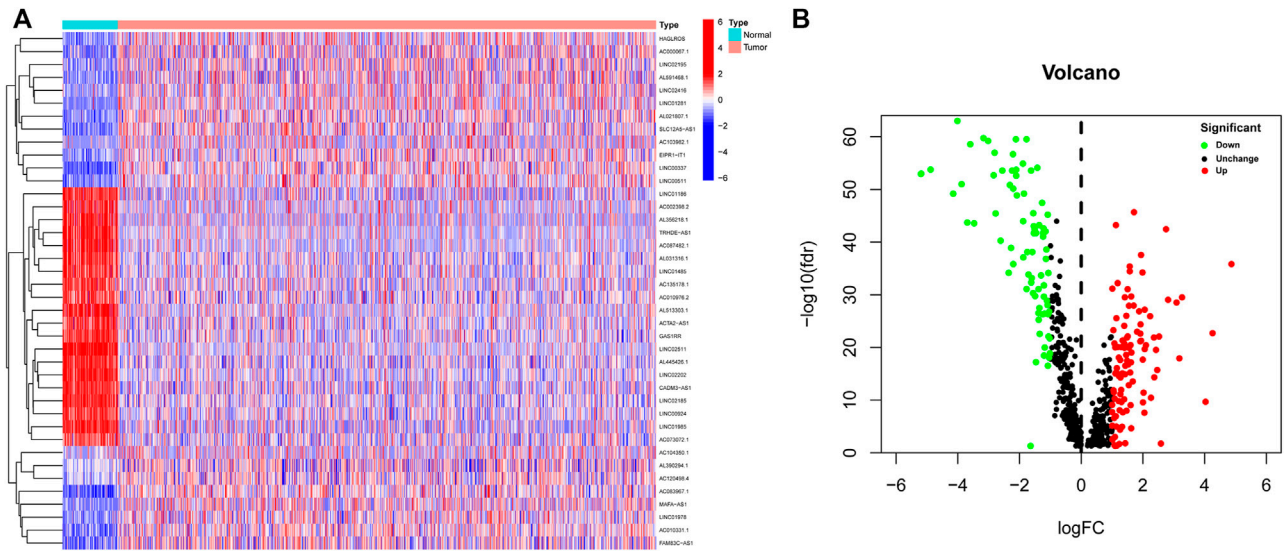


FIGURE 2
Identification of differentially expressed NETs-related lncRNAs in BC. **(A)** Heatmap demonstrating the differential expression of the top 20 up-regulated and 20 down-regulated NETs-related lncRNAs in BC tissues and normal tissues. **(B)** Volcano plot exhibiting 105 up-regulated and 81 down-regulated NETs-related lncRNAs. NETs, neutrophil extracellular traps; lncRNAs, long non-coding RNAs; BC, breast cancer.

Breast Surgery Department of Sun Yat-sen University Cancer Center, Guangzhou, China, with patients' consent and approval from the Medical Ethics Committee of the hospital.

For RNA expression assay, total RNA was extracted using TRIzol reagent (Invitrogen, 15596018) and quantified using a Nanodrop (Thermo Scientific). Then, cDNA was obtained using the

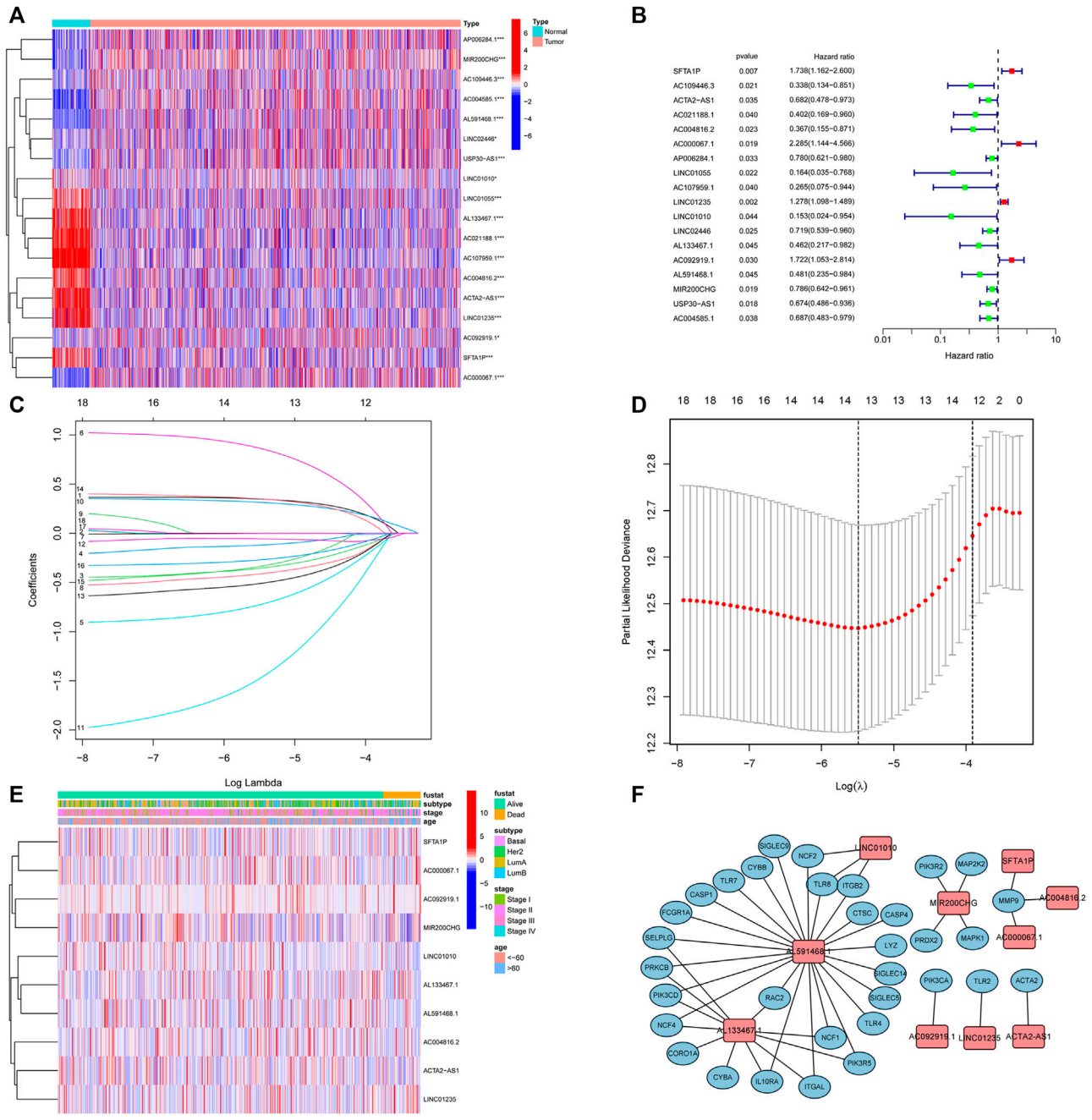


FIGURE 3
Identification of prognostic NETs-related lncRNAs. **(A)** Heatmap of OS-associated differentially expressed NETs-related lncRNAs. **(B)** Forest plot of NETs-related lncRNAs associated with BC prognosis *via* univariate analysis. **(C)** LASSO coefficient profiles of 13 NETs-related lncRNAs. **(D)** Cross-validation for tuning parameter selection in the proportional hazards model. **(E)** Heatmap of the 10 NETs-related lncRNAs *via* multivariate analysis. **(F)** Correlation between 10 NETs-related lncRNAs and NETs genes. NETs, neutrophil extracellular traps; lncRNAs, long non-coding RNAs; BC, breast cancer; OS, overall survival.

PrimeScript™ RT Reagent Kit (Takara, RR036A) according to the manufacturer's instruction. Quantitative real-time polymerase chain reaction (qRT-PCR) was performed with the TB Green™ Premix Ex Taq™ (TaKaRa, RR420A) using SYBR Green (Roche) on a LightCycler 480 (Roche). The relative abundance of each lnRNA, using GAPDH as an endogenous control, was calculated by $2^{-\Delta\Delta CT}$ method. The primers sequences used in this study are listed in Supplementary Table S2. Analysis between the two groups was performed by paired-sample t-tests. $p < 0.05$ was considered statistically significant.

Results

Identification of differentially expressed LncRNAs related to NETs

A flow chart of the study is depicted in Figure 1. A total of 14,142 lncRNA transcripts and 19,658 protein-coding genes were identified from the TCGA database. The gene expression of NETs in BC was screened by matching the mRNA expression matrix of TCGA

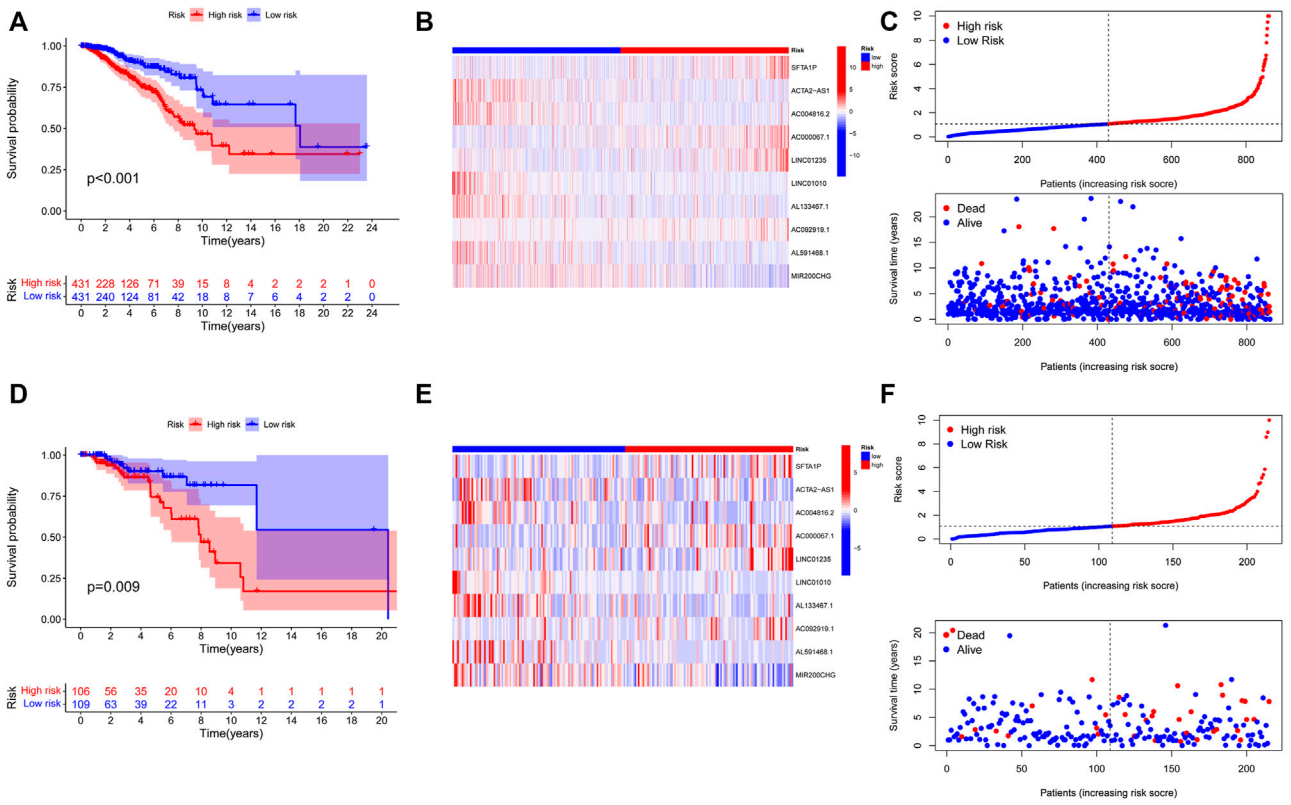


FIGURE 4
Construction and validation of prognostic signatures of NETs-related lncRNAs. **(A-D)** Kaplan–Meier curves of overall survival for BC patients based on the risk score in the training **(A)** and validation **(D)** cohorts. **(B-E)** Heatmap of the signature of 10 NETs-related lncRNAs in the training **(B)** and the validation **(E)** cohorts. **(C-F)** Distribution of risk score, OS, and OS status in the training **(C)** and the validation **(F)** cohorts. NETs, neutrophil extracellular traps; lncRNAs, long non-coding RNAs; BC, breast cancer; OS, overall survival.

and 170 NETs-related genes. Afterwards, 655 lncRNAs were identified to be highly correlated with the expression of NETs-related genes ($|R| > 0.4$ and $p < 0.001$), among which 186 lncRNAs (as shown in Supplementary Table S3) were identified based on differential expression between normal and tumor tissue. A heatmap was performed to demonstrate the differential expression of the top 20 up-regulated and 20 down-regulated lncRNAs (Figure 2A), and volcano plot of differentially expressed lncRNAs showed 105 up-regulated lncRNAs and 81 down-regulated lncRNAs in BC (Figure 2B).

Identification of prognostic NETs-Related LncRNAs and establishment of risk model

Next, to establish a convincing risk predictive model, BC patients were randomly divided in an 8:2 ratio into training ($n = 749$) and validation cohorts ($n = 188$). In the training cohort, a univariate Cox regression algorithm was performed to determine the lncRNAs related to BC patient survival and 18 prognosis-related candidate lncRNAs were filtered out (Figures 3A, B). The 18 lncRNAs were incorporated into the LASSO Cox regression algorithm to improve model accuracy and reduce model overfitting (Figure 3C), and cross-validation was performed and 13 prognostic lncRNAs were screened out (Figure 3D). To further improve clinical utility, 13 lncRNAs signatures were purified by performing stepwise multivariate Cox regression analysis (Figure 3E). Then, a prognostic risk model was constructed based on the correlation coefficient of the expression of lncRNAs in the multivariable Cox regression model. The following formula was applied to determine the risk score: risk score = (0.366 × SFTA1P expression level) + (−0.500 × ACTA2-AS1 expression level) + (−0.907 × AC004816.2 expression level) + (1.013 × AC000067.1 expression level) + (0.361 × LINC01235 expression level) + (−2.069 × LINC01010 expression level) + (−0.685 × AL133467.1 expression level) + (0.387 × AC092919.1 expression level) + (−0.676 × AL591468.1 expression level) + (−0.377 × MIR200CHG expression level). The correlation between the 10 prognostic lncRNAs and NETs-related genes was shown in Figure 3F.

Validation of the prognostic risk model

Initially, The BC patients in the training and validation cohorts were categorized into high- and low-risk groups according to the training cohort's medium risk score. The validation cohort was used for internal validation to weight the predictive capability of the model. As represented by the Kaplan–Meier curves, a significantly inferior OS was observed for those in the high-risk group in comparison with those in the low-risk group in both the training ($p < 0.001$) (Figure 4A) and validation cohorts ($p = 0.009$) (Figure 4D). The visualized

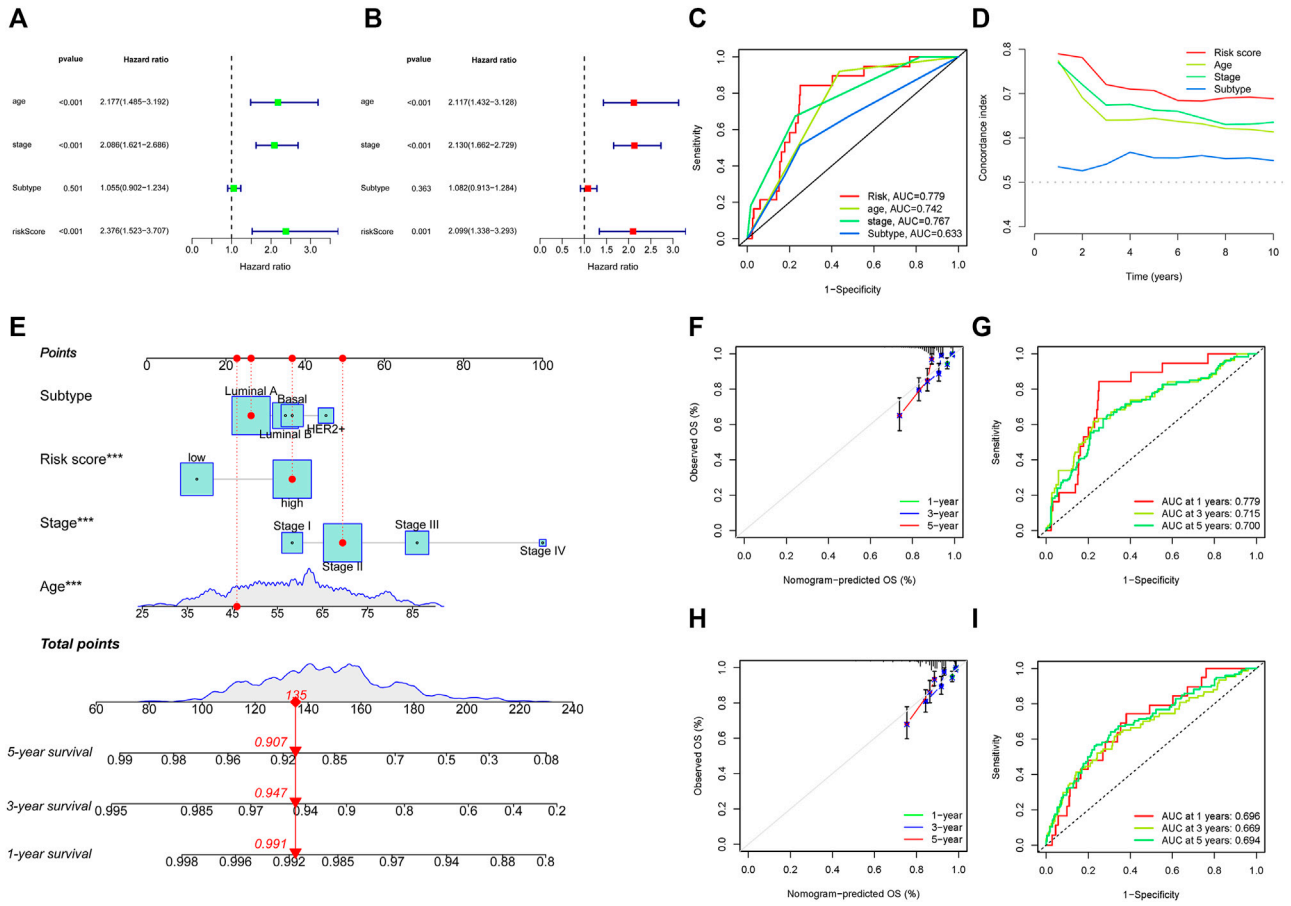


FIGURE 5
Identification of prognostic parameters in BC. **(A)** The univariate Cox regression analysis in training cohort. **(B)** The multivariate Cox regression analysis in training cohort. **(C, D)** ROC curves and C-index for the risk score, age, stage, and subtype in training cohort. **(E)** Construction of nomogram for survival prediction in training cohort. **(F, I)** Calibration plots and ROC curves of the nomogram for predicting the probability of OS at 1-, 3-, and 5- years in training **(F, G)** and validation **(H, I)** cohorts. ROC, receiver operating characteristic; BC, breast cancer.

heatmap revealed differential expression of lncRNAs in high- and low-risk groups in training and validation cohorts (Figures 4B, E). Along with the increase in the risk score, the proportion of patients in the high-risk group also increased, and so did the level of mortality (Figures 4C, F).

Construction and assessment of a clinical prognostic model

To further evaluate the possibility that the risk score could serve as an independent BC prognostic signature, the univariate Cox regression analysis was performed by matching the risk score of BC patients in the training cohort with conventional clinicopathological parameters (age, stage, and subtype). In the group of 749 individuals included in this investigation, age, stage, and risk score were associated with the prognosis of the BC patients (Figure 5A). Considering the impact of molecular subtype on BC survival, a multivariable Cox algorithm was conducted to further screen out four prognostic factors of the BC patients (age, stage, subtype, and risk score) (Figure 5B), which was consistent with the result of the univariate Cox regression analysis. The ROC curves and C-index also revealed that the risk score acted as an important role in predicting BC prognosis in training cohort (Figures 5C, D). Based on the four parameters, a nomogram was conducted to predict an individual's prognosis at 1-, 3-, and 5- years (Figure 5E). The calibration curves showed high consistency between the predicted and the actual 1-, 3- and 5-years survival probabilities (Figure 5F). ROC curves analysis also showed satisfactory AUC values at 1-, 3- and 5- years (0.779, 0.715, and 0.700, respectively) (Figure 5G). In addition, the predictive power and performance of this nomogram was also confirmed in the validation cohort, with better calibration curves and AUC values (Figures 5H, I).

Molecular characteristics of different risk groups

To gain further insight into the specific molecular differences between high- and low-risk groups, differentially expressed genes (as shown in Supplementary Table S4) were identified and functional annotation was conducted by the GSEA. The GSEA with | ES scores | > 0.5, FDR <0.25 and *P* adjusted value <0.05 for the five most significant

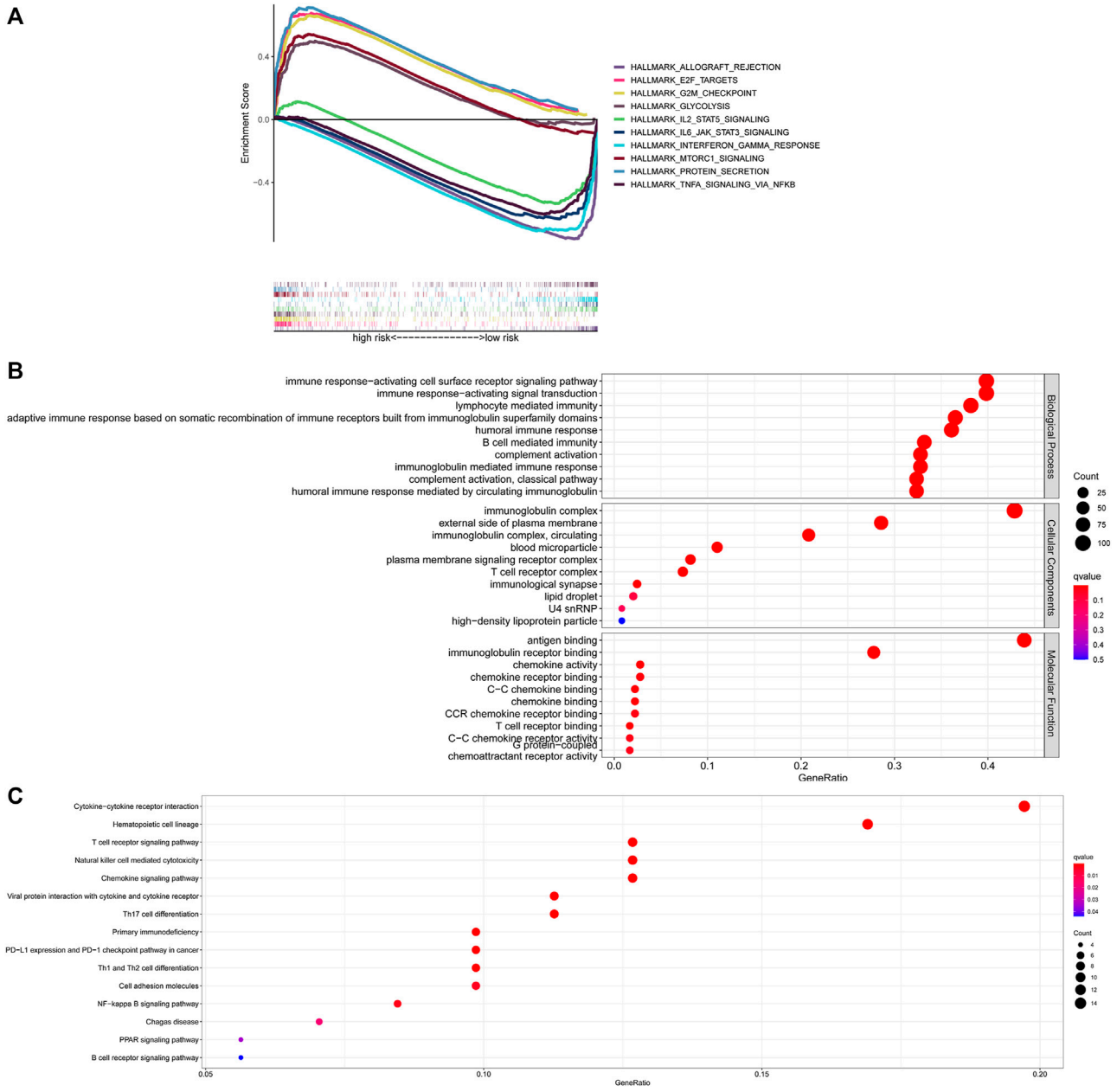


FIGURE 6
Exploration of the possible functions and pathways of NETs-related lncRNAs. **(A)** Tumor hallmarks in high- and low-risk subgroups. **(B)** GO analysis of different expression gene sets between different risk subgroups. **(C)** KEGG analysis of different expression gene sets between different risk subgroups. NETs, neutrophil extracellular traps; lncRNAs, long non-coding RNAs.

pathways showed enrichment of tumor hallmarks in high- and low-risk subgroups. The high-risk group was enriched in HALLMARK_PROTEIN_SECRETION, HALLMARK_GLYCOLYSIS, HALLMARK_G2M_CHECKPOINT, HALLMARK_MTORC1_SIGNALING, HALLMARK_E2F_TARGETS, while the low-risk group was enriched in HALLMARK_ALLOGRAFT_REJECTION, HALLMARK_IL2_STAT5_SIGNALING, HALLMARK_TNFA_SIGNALING_VIA_NFKB, HALLMARK_INTERFERON_GAMMA_RESPONSE, HALLMARK_IL6_JAK_STAT3_SIGNALING (Figure 6A). Then, the potential functions and pathways were identified using GO and KEGG enrichment analyses. GO analysis showed that the differentially expressed genes between the risk score subgroups were enriched in immune response–activating cell surface receptor signaling pathway and immune response–activating signal transduction (Biological Process), immunoglobulin complex and external side of plasma membrane (Cellular Components), antigen binding and immunoglobulin receptor binding (Molecular Function; Figure 6B). KEGG analysis showed that the differentially expressed genes were enriched in cytokine–cytokine receptor interaction, hematopoietic cell lineage, T cell receptor signaling pathway, natural killer cell mediated cytotoxicity, chemokine signaling pathway (Figure 6C). These results indicated that

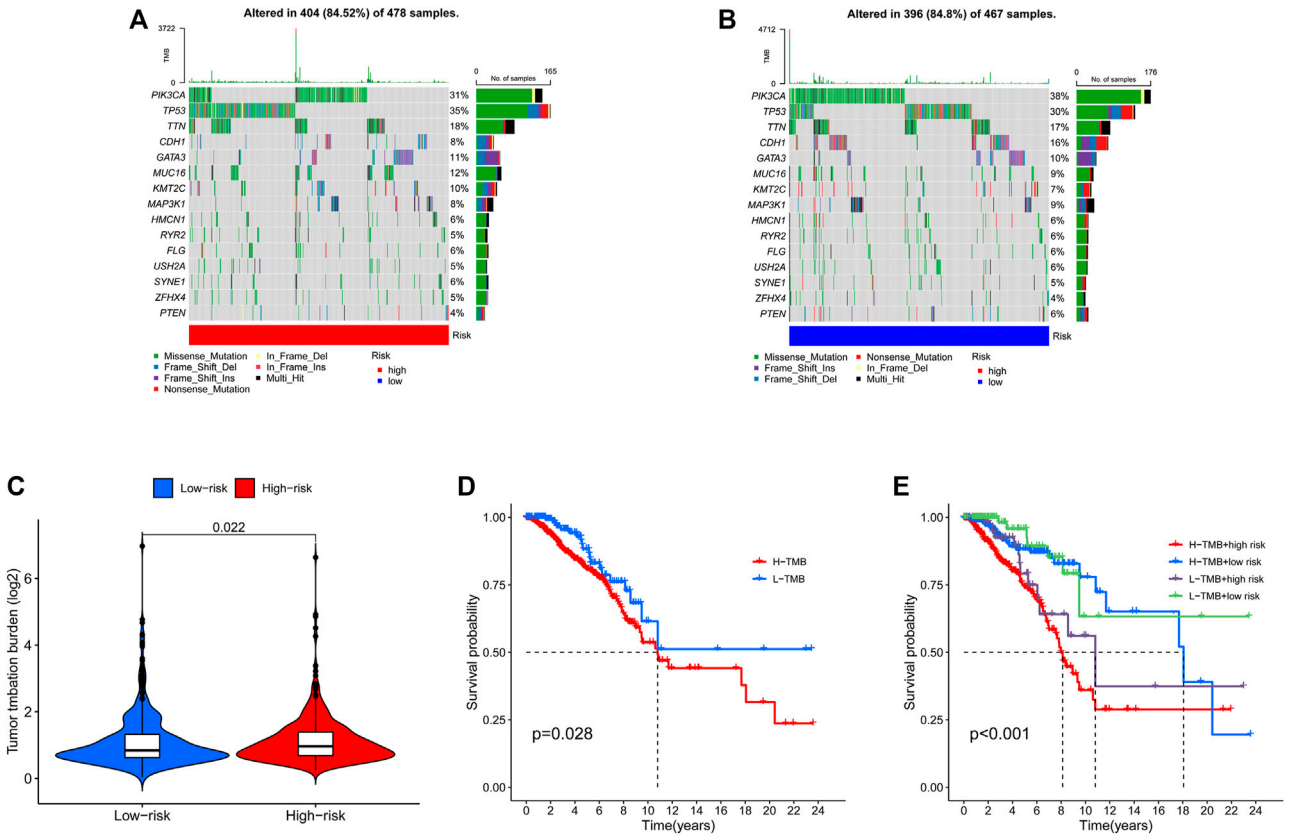


FIGURE 7
Landscape of mutation profiles in high- and low-risk BC patients. **(A)** The TMB in high-risk BC patients. **(B)** The TMB in low-risk BC patients. **(C)** The difference of TMB was significant in BC patients with different risk. **(D)** The lower TMB was positively correlated with the OS in BC patients. **(E)** Comparison of OS in BC patients with different risks and different TMBs. BC, breast cancer; TMB, Tumor mutation burden; OS, overall survival.

enrichment was mainly focused on immune-related functions and pathways.

Relationship between risk score and TMB

The TMB has been implicated as a biomarker for responses to immune checkpoint inhibitor therapy (Chan et al., 2019). The TMB of the high- and low-risk groups of BC patients was shown in Figures 7A, B. The difference in TMB between the two groups was statistically significant (p = 0.022, Figure 7C). Moreover, the patients in the low-TMB group experienced significantly better survival than those in the high-TMB group (p = 0.028, Figure 7D). Moreover, special subgroup analyses stratifying samples were performed according to the combination of TMB status and risk groups. The result showed that some patients with high-risk score in the low- or high-TMB groups had significantly shorter OS than those with low-risk score in the low- or high-TMB groups (p < 0.001, Figure 7E), but no significance could be calculated in the different TMB groups with the same risk scores. Thus, a high level of TMB, commonly considered as a "universal marker," might fail to accurately predict the reactivity of checkpoint inhibitors across all cancer types. The exploration of the tumor microenvironment (TME) based on the risk predictive model was further conducted, and the proportion of 22 types of tumor-infiltrating immune cells in the high-risk and low-risk groups are shown in Figures 8A, B.

Immune status and immune function in the different risk groups

Then, we analyzed the potential correlation between NETs-related lncRNAs and 16 immune cells and the scores of 13 immunological functions using the ssGSEA algorithm. The results revealed that activated dendritic cells (aDCs), B cells, CD8 T cells, dendritic cells (DCs), immature Dendritic Cells (iDCs), mast cells, neutrophils, natural killer cells, plasmacytoid dendritic cells (pDCs), T follicular helper cell, type1 T helper cells, type2 T helper cells, tumor-infiltrating lymphocyte were more predominant in the low-risk group, while regulatory cells (Tregs) were more abundant in the high-risk group (Figure 9A). In addition, immune-related functions were highly enriched in the low-risk group (p < 0.01, Figure 9B). It has also been reported that blocking the immune checkpoint pathway is an extremely promising approach to achieve anti-cancer immunity (Li et al., 2019). Therefore, we made a comparison in the expression discrepancies of checkpoint genes between the high- and low-risk groups. The result was shown in Figure 9C, which indicated that there was significant difference in the expression of all checkpoint genes between the two groups, with all immune checkpoint molecules except CD276 being highly expressed in the high-risk group. These results demonstrate that NETs-related lncRNAs signature can be applied to evaluate the tumor immune microenvironment and the expression of immune checkpoint genes in BC patients.

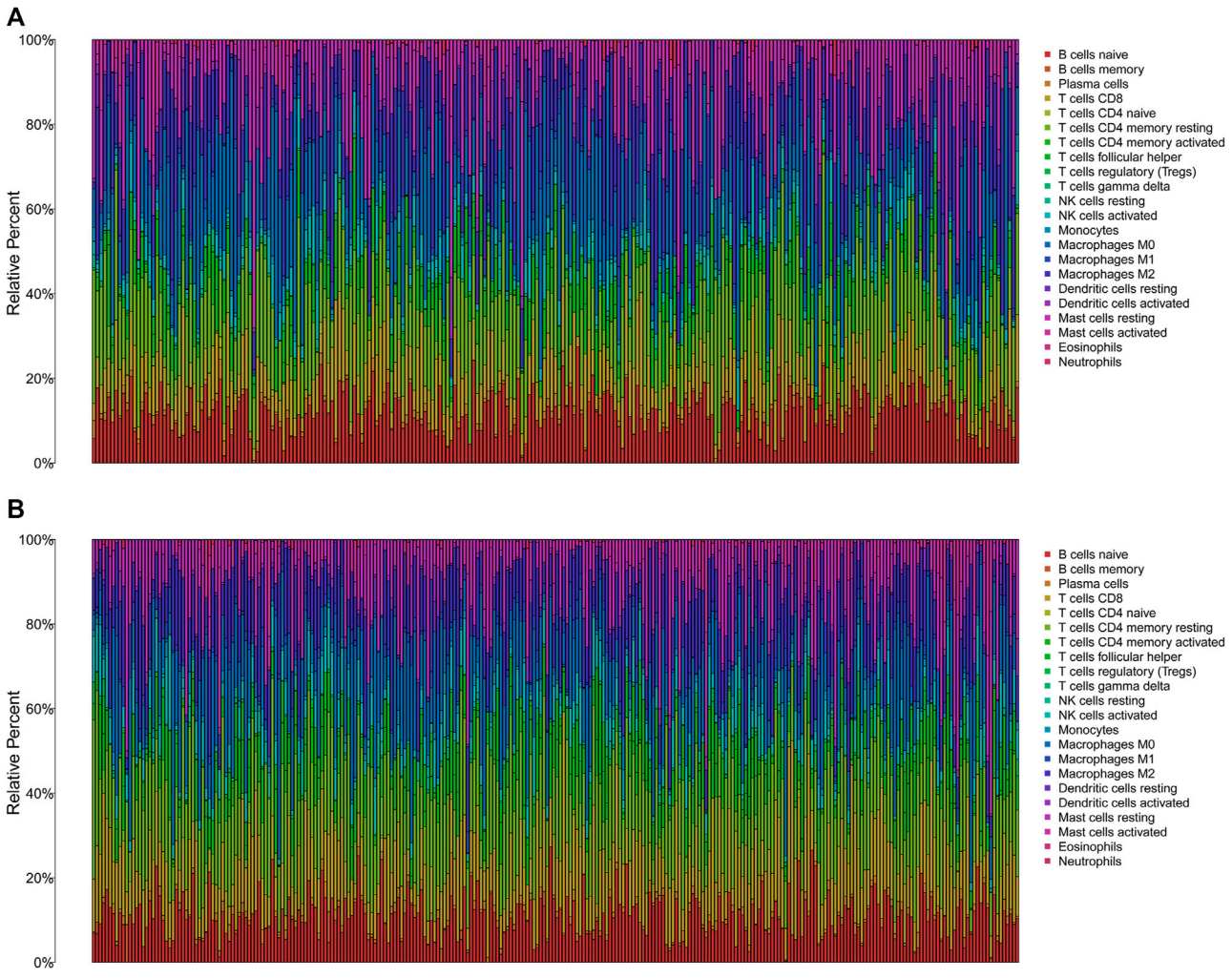


FIGURE 8
The immune infiltration of 22 immune cell types in high- and low-risk patients with BC. **(A)** The immune infiltration of 22 immune cell types in high-risk patients with BC. **(B)** The immune infiltration of 22 immune cell types in low-risk patients with BC. BC, breast cancer.

Prediction of chemotherapy efficacy and immunotherapy response

To assess the prediction performance of NETs-related lncRNAs on drug therapy for BC, we performed an analysis of the relationship between high- and low-risk groups and the efficacy of commonly used therapeutic agents using the "pRRophetic" R package. Our study suggested that the high-risk group showed significant correlation with a higher IC_{50} for chemotherapeutic and targeted agents such as cisplatin, gemcitabine, paclitaxel, vinorelbine, and gefitinib, which might indicate low-risk group was more suitable for these agents (all $p < 0.001$, Figures 10A–E). Conversely, lapatinib is more suitable for the high-risk group with a lower IC_{50} ($p < 0.001$, Figure 10F). In immunotherapy, TIDE scores were applied to assess the efficacy of immune checkpoint (PD-1 and CTLA-4) inhibitors in the high- and low-risk groups. High-risk patients had significantly lower TIDE scores compared to low-risk patients ($p < 0.001$, Figure 10G), indicating that patients in the high-risk group might have better response when receiving immune therapy. In the era of immunotherapy, the focus on treatment efficacy should be accompanied by a focus on the management of immune-related adverse effects. The commonly used immunosuppressive agents, such as methotrexate and rapamycin, differed significantly between the two groups, with the high-risk group having a higher IC_{50} ($p < 0.001$, Figures 10H, I).

Exploration and verification of signature-related LncRNAs expression profiles

First, the paired lncRNAs expression profiles of breast samples were analyzed by TCGA databases. Compared with adjacent normal breast tissues, the lncRNAs expression of SFTA1P, ACTA2-AS1, AC004816.2, LINC01235, and AL133467.1 were lower, while the expression of AC000067.1, AL591468.1, and MIR200CHG was higher in BC tissues (Figure 11A). Only LINC01010 and AC092919.1 were not statistically significant in paired tissues. To further confirm the validity of this model, the expression levels of all lncRNAs in 10 BC tumor tissues and matched normal tissues from our

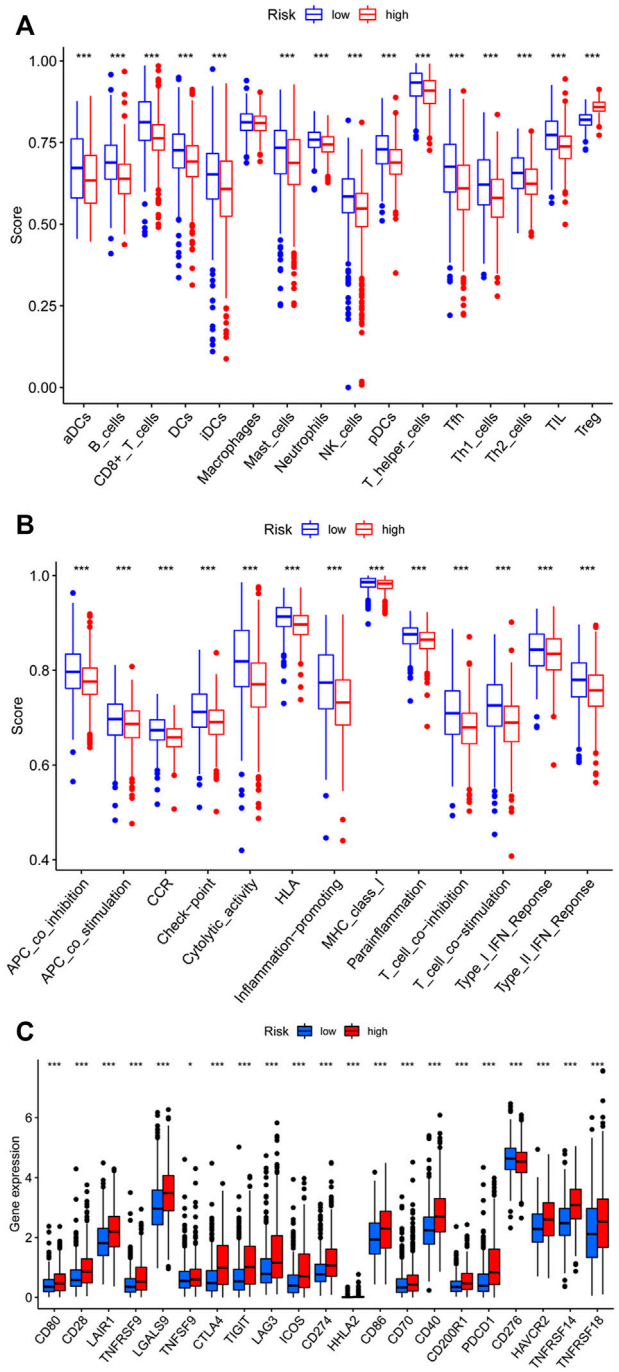


FIGURE 9
Assessment of Immune Cell Infiltration, Immune Function Signature, and Immune Checkpoint Genes in Different risk Groups. **(A)** Differences in the infiltration of immune cells between the high- and low-risk groups. **(B)** Differential expression levels of immune function signature between the high- and low-risk groups. **(C)** Differential expression of immune checkpoint genes between the high- and low-risk groups. ns, not significant, $*p < 0.05$, $**p < 0.01$, $***p < 0.001$.

hospital were detected using qRT-PCR. Except for LINC01010, their expression trend was consistent with the bioinformatic analysis results (Figure 11B). The qRT-PCR results indicated that our bioinformatics analysis was accurate and reliable, reinforcing our conclusions from these data.

Discussion

Recent Studies have shown that the prognosis and therapeutic response in tumor patients can vary according to molecular characteristics, although patients share homogeneous clinicopathological risk variables (Chan et al., 2019). Thus, novel molecular prognostic markers need to be identified to complement the clinical parameters to predict prognosis. NETs play a critical role in the tumor microenvironment and contribute to tumor migration, invasion, and distant metastasis in different ways (Ireland and Oliver, 2020; Martins-Cardoso et al., 2020; Demkow, 2021). Functionally, NETs function as a physical barrier to shield tumor cells from interactions with neighboring anti-tumor immune cells such as NK cells and $CD8^+$ T cells in the tumor immune microenvironment (TME), affecting the tumor immune landscape and tumor response to immunotherapy (Ireland and Oliver, 2020; Teijeira et al., 2020). LncRNAs, as particular type of non-coding RNAs, have been reported to mediate NETs-induced cancer cell metastasis in the TME (Wang et al., 2022). Therefore, the exploration of the relevance of the model based on NETs-related lncRNAs and the immune microenvironment of BC patients is essential.

In this current study, we comprehensively identified NETs-related lncRNAs by constructing a co-expression network based on the correlation analysis between lncRNAs and NETs-related genes. To avoid overfitting and strengthen the clinical practicability, univariate, LASSO, and multivariable logistic regression algorithms were applied to screen out 10 NETs-related lncRNAs for the construction of prognostic risk model. Furthermore, we stratified patients into high-risk and low-risk groups based on this prognostic risk model. Survival analysis showed poorer prognosis in high-risk group. The risk score was an independent risk parameter in a cox regression analysis combining with clinical characteristics (age, stage, and subtype). Functional enrichment analysis of differentially expressed genes indicated that there were significantly different in immune-related functions and pathways between the two groups. Moreover, the tumor immune cells, immune function, immune checkpoint genes, and drug sensitivity in BC based on the prognostic risk model were further analyzed, which all demonstrated the potential predictive utility of the model in immunotherapy of BC patients. Finally, as demonstrated by our validation experiment, we were able to confirm the consistency expression of NETs-related lncRNAs by qRT-PCR in the BC tumor tissue and paired normal tissue.

After witnessing the success of molecular targeted therapies in the clinical application of several solid tumors, there is a growing enthusiasm in studying the impact of lncRNAs on tumors (Yang et al., 2022; Yardim-Akaydin et al., 2022). The lncRNA SFTA1P acts as an oncogene and promotes the growth and invasion of lung (Du et al., 2020; Zhu et al., 2021) and liver cancers (Huang et al., 2020) in various signaling pathways, such as mTOR signaling pathway, AKT signaling pathway. Interestingly, in the GSEA analysis, we also found that the mTOR signaling pathway was enriched in high-risk group. Several studies have reported the function of ACTA2-AS1, enhancing the malignant phenotype of cervical cancer (Luo et al., 2020), while exhibiting anti-tumor effects in liver cancer (Zhou and Lv, 2019), and lung adenocarcinoma (Ying et al., 2020). Additionally, ACTA2-AS1 has been implicated in platinum resistance in ovarian cancer and lung cancer (Lin et al., 2022; Liu et al., 2022). LINC01235 promotes gastric cancer migration and invasion *via* epithelial-mesenchymal

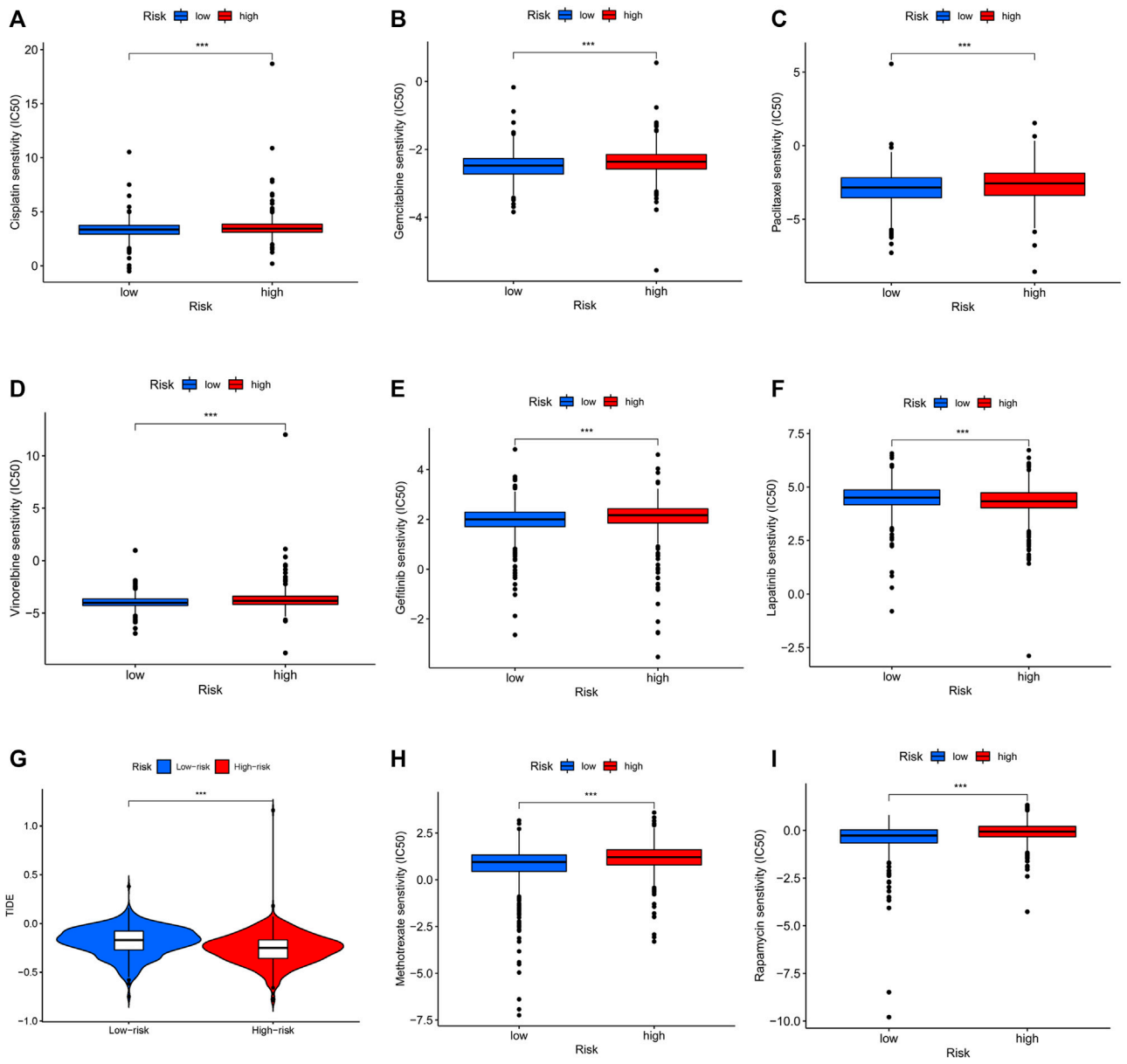


FIGURE 10
Differential sensitivity to chemotherapies and immunotherapy between the high- and low-risk patients with BC. BC, breast cancer. ***$p < 0.001$.

transition (EMT) pathway (Zhang et al., 2021). Of note, NETs can activate the EMT program to drive the pro-metastatic phenotype of human breast cancer cells (Martins-Cardoso et al., 2020). LINC01010 is downregulated in HBV-transgenic hepatocellular carcinoma cell line and is a potential tumor suppressor that inhibits the development of HBV-associated hepatocellular carcinoma (Gan et al., 2021). Tang et al. (2021) reported that MIR200CHG can directly bind to the transcription factor YB-1 and inhibit its ubiquitination and degradation to promote proliferation, invasion, and drug resistance in breast cancer. However, in our current study, MIR200CHG acts as a protective factor in the development of BC. Therefore, further studies are needed to explore its role. For the five remaining NETs-related lncRNAs (AC004816.2, AC000067.1, AL133467.1, AC092919.1, and AL591468.1), there have been no studies exploring their potential roles in the development of cancer at present. Thus, further research in the future is needed to understand the deeper mechanisms.

NETs-related lncRNAs play a multifaceted role in the tumor microenvironment and can alter the interaction and crosstalk between tumor cells and the tumor microenvironment, leading to immunosuppression and therapy resistance, thereby allowing tumor cells to evade immune surveillance (Papayannopoulos, 2018; Yang et al., 2022). Here, the GO and KEGG analyses indicated that the differentially expressed genes between the high- and low-risk groups were mainly enriched in immune-related pathways. This further confirms the role of NETs-related lncRNAs in the tumor microenvironment by mainly influencing the immune function of the body.

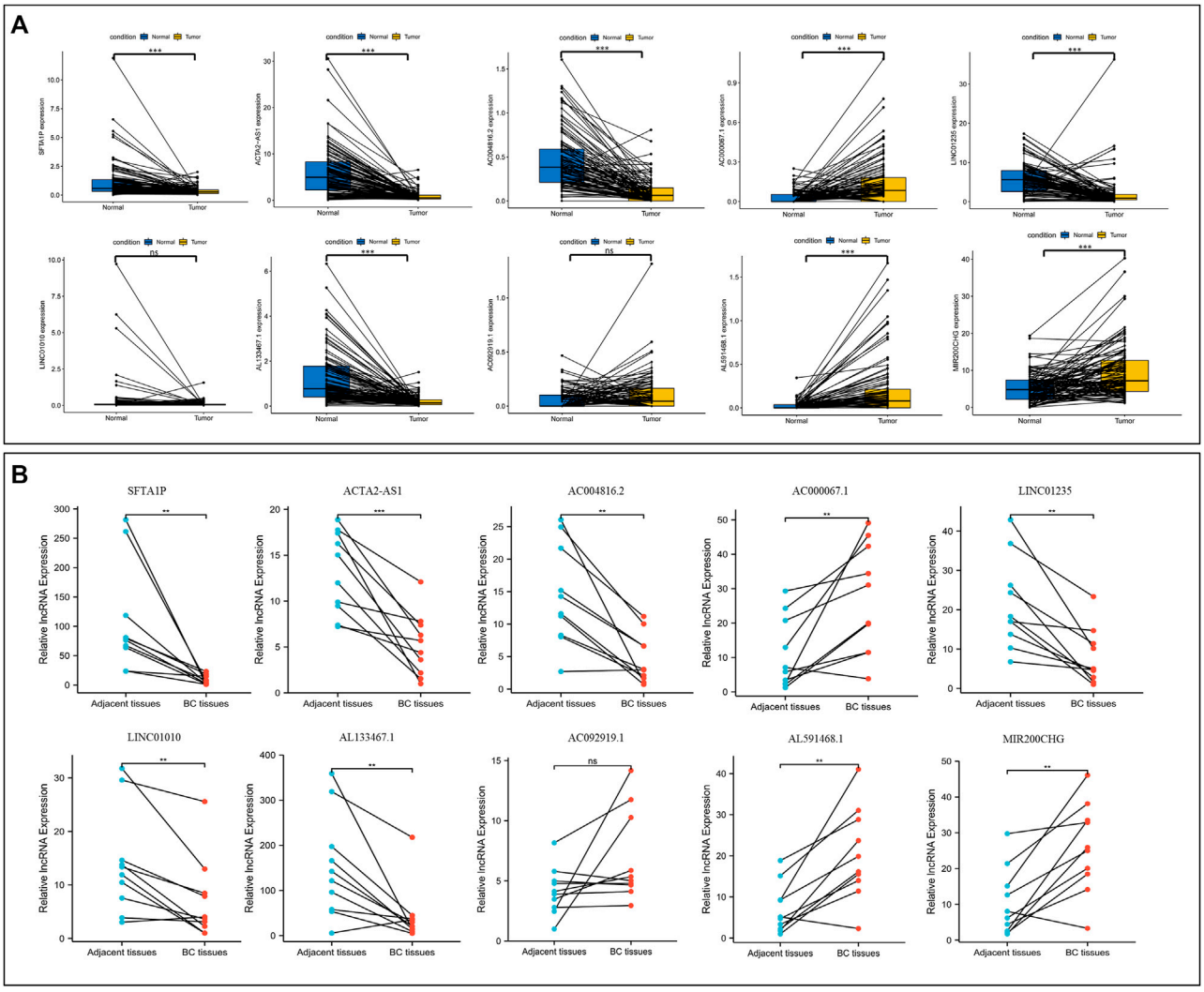


FIGURE 11
Validation the expression level of NETs-related lncRNAs. **(A)** The differential expression of 10 NETs-related lncRNAs between tumors and their matched benign tissue in TCGA RNA-seq data. **(B)** The expression levels of 10 NETs-related lncRNAs in 10 paired samples of BC tumor and adjacent normal tissues were examined by qRT-PCR. ns, not significant, $*p < 0.05$, $**p < 0.01$, $***p < 0.001$. NETs, neutrophil extracellular traps; lncRNAs, long non-coding RNAs; BC, breast cancer.

Genetic mutations are the basis for tumor development, and specific mutations predict the response to therapy and prognosis (Chan et al., 2019). In our current study, PIK3CA mutations occurred most commonly between the two groups and appeared more frequently in the low-risk group than in the high-risk group. Several studies have been conducted on PIK3CA mutations, but results on the prognostic significance of PIK3CA mutations appear to be conflicting (Kalinsky et al., 2009; Mosele et al., 2020). Here, PIK3CA mutation is a positive effect on BC patient survival. Further studies on the effect of PI3K mutations on patient survival need to be investigated. TMB is emerging as a potent biomarker for predicting the efficacy of immunotherapy in cancer patients (Chan et al., 2019; Sha et al., 2020). In this study, the TMB of BC patients was positively correlated with the risk score. High-risk patients had higher levels of immune checkpoint genes expression and better immunotherapy outcomes, which indicated that the prediction result of TMB and NETs-related lncRNAs is consistent. Notably, the risk score appeared to have a greater impact on BC patient survival compared to TMB. Therefore, the risk score of BC patients can be used as a complement to TMB to better predict patient immunotherapy outcomes.

The tumor microenvironment (TME), composed of tumor cells and stromal cells, is associated with tumorigenesis, pathogenesis, and tumor progression, supporting the cancer cells replicative proliferation and affecting the tumors malignant phenotype (Chen and Song, 2022; Franzén et al., 2022). The presence of many immune cells in TME, charactered by a “double-edged sword”, behave as the anti-tumor and pro-tumor cells, protecting us against tumor cells or modulating tumor cells migration, invasion, metastasis, and anticancer drug sensitivity (Li et al., 2020; Mehraj et al., 2021; Peña-Romero and Orenes-Piñero, 2022). In this current study, the differences in immune infiltration, immune function, and

immune checkpoint genes expression between high-risk and low-risk BC patients were elaborated, and we found statistically significant differences in the immune status of the body. Activated dendritic cells, activated CD8 T cells, natural killer cells and mast cells have been reported to be antitumor immune cells (Peña-Romero and Orenes-Piñero, 2022). Here, we found higher levels of tumor antagonistic immune cells in the low-risk group compared to the high-risk group. Also, immune function analysis showed that patients with low-risk scores exhibited more immune activity, which probably explains why low-risk patients have a better prognosis. Historically, breast cancer was not a highly immunogenic tumor due to the low mutation rate and few neoantigens, and selecting the suitable patients for immunotherapy is difficult (Franzén et al., 2022). In the present study, patients with high-risk scores had higher expression of immune checkpoint genes and better immunotherapy outcomes, suggesting that our signature might be used to assess the suitable population for immunotherapy.

Finally, based on the "pRophetic" algorithm and the TEDER program, we assessed the susceptibility of high- and low-risk populations to commonly used chemotherapeutic agents, molecularly targeted therapies, and immunotherapeutic responses. Our results confirm the potential predictive value of NETs-related lncRNAs for chemotherapy susceptibility, targeted therapy and immunotherapy efficacy. BC chemotherapy agents, such as cisplatin, gemcitabine, paclitaxel and vinorelbine, are more sensitive in low-risk groups. Gefitinib, an epidermal growth factor receptor (EGFR) inhibitor, is effective in cancers that have activated mutations in EGFR. However, BC routinely exhibit intrinsic resistance to anti-EGFR therapeutics (You et al., 2021). Here, we identified the sensitivity of low-risk populations to gefitinib, suggesting that anti-EGFR therapy is effective for specific populations. Lapatinib, a reversible inhibitor of intracellular tyrosine kinase activity of HER2 and EGFR1, is used in combination with capecitabine in advanced or metastatic BC (Tesch and Gelmon, 2020). In our study, the IC_{50} value of lapatinib was lower in the high-risk group, indicating that lapatinib is more appropriate for these high-risk populations. As immune checkpoint inhibitors improve overall survival in various cancers, the accompanying immunotherapy-mediated side-effects, such as colitis, hepatitis, and rheumatic diseases, have drawn attention (Kim et al., 2022; Okiyama and Tanaka, 2022). Selective the appropriate immunosuppressive agents for BC treatment are necessary due to the differences in the immune system of individuals and their sensitivity to drugs. Here, we explored the drug sensitivity of two immunosuppressive agents: methotrexate and rapamycin. Methotrexate is commonly used in the treatment of autoimmune diseases, such as rheumatic immune-related adverse events (Kostine et al., 2018). Some studies have reported that the combination of rapamycin and other chemotherapeutic agents can improve the efficacy (Niu et al., 2011; Sun et al., 2021). In addition, the combination with rapamycin and anti-PD-1 synergistically inhibits tumor growth and mitigates immune-related colitis in a mouse melanoma model (Bai et al., 2021). Thus, methotrexate and rapamycin are promising for the treatment of adverse events caused by immunotherapy. In our exploration, the IC_{50} values of methotrexate and rapamycin were higher in high-risk group. Understandably, as the efficacy of immunotherapy increases, so do the side effects that come with it. Thus, our NETs-related lncRNAs signature may help to identify patients who would benefit from immunosuppressive therapy, but the underlying mechanism of action still needs to be clarified.

Certainly, there are several potential limitations to the current study. The in-depth molecular mechanisms used to construct NETs-related lncRNAs prognostic model need to be further validated in experimental studies. Moreover, the study data were based only on the TCGA public databases, which may represent a selection bias. Therefore, further multicenter, large-scale studies are now required to better determine its clinical utility and predictive validity.

Conclusion

To summarize, a novel prognostic model based on 10 NETs-related lncRNAs was successfully constructed, which demonstrated good predictive capacity and effectiveness for BC. Furthermore, our NETs-related lncRNAs signature was significantly correlated with TMB, tumor immune microenvironment, and anti-cancer agents, indicating that these molecular changes might explain individual differences in the treatment effectiveness. These findings may provide provides a new analytical perspective on BC treatment decisions and enhance biological understanding in BC.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: https://portal.gdc.cancer.gov/repository, TCGA Breast Cancer (BRCA).

Ethics statement

The studies involving human participants were reviewed and approved by the Medical Ethics Committee of Sun Yat-sen University Cancer Center. The patients/participants provided their written informed consent to participate in this study.

Author contributions

TJ, LG, and HL contributed to the study conception and design. TJ and YW wrote the manuscript. TJ, XC, WX, SX, and LG collected and analysed the raw data. All authors read and approved the final manuscript.

Funding

This study was supported by the National Natural Science Foundation of China (HL, grant numbers 81773103), the Natural Science Foundation of Guangdong Province (HL, grant numbers 2017A030313617), and the National Science Foundation for Young Scientists of China (WX, grant numbers 82102838). HL contributed to the study conception and design. WX collected and analysed the raw data.

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.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcell.2023.1117637/full#supplementary-material

References

Bai, X., Wang, X., Ma, G., Song, J., Liu, X., Wu, X., et al. (2021). Improvement of PD-1 blockade efficacy and elimination of immune-related gastrointestinal adverse effect by mTOR inhibitor. *Front. Immunol.* 12, 793831. doi:10.3389/fimmu.2021.793831

Chan, T. A., Yarchoan, M., Jaffee, E., Swanton, C., Quezada, S. A., Stenzinger, A., et al. (2019). Development of tumor mutation burden as an immunotherapy biomarker: Utility for the oncology clinic. *Ann. Oncol.* 30, 44–56. doi:10.1093/annonc/mdy495

Chen, X., and Song, E. (2022). The theory of tumor ecosystem. *Cancer Commun. (Lond)* 42, 587–608. doi:10.1002/cac2.12316

Demkow, U. (2021). Neutrophil extracellular traps (NETs) in cancer invasion, evasion and metastasis. *Cancers (Basel)* 13, 4495. doi:10.3390/cancers13174495

Du, D., Shen, X., Zhang, Y., Yin, L., Pu, Y., and Liang, G. (2020). Expression of long non-coding RNA SFTA1P and its function in non-small cell lung cancer. *Pathol. Res. Pract.* 216, 153049. doi:10.1016/j.prp.2020.153049

Dwyer, M., Shan, Q., D'Ortona, S., Maurer, R., Mitchell, R., Olesen, H., et al. (2014). Cystic fibrosis sputum DNA has NETosis characteristics and neutrophil extracellular trap release is regulated by macrophage migration-inhibitory factor. *J. Innate Immun.* 6, 765–779. doi:10.1159/000363242

FranzéN, A. S., Raftery, M. J., and Pecher, G. (2022). Implications for immunotherapy of breast cancer by understanding the microenvironment of a solid tumor. *Cancers (Basel)* 14, 3178. doi:10.3390/cancers14133178

Gan, L., Shangguan, Q., Zhang, F., Tong, X., Qi, D., Zhao, Y., et al. (2021). HBV HBx-downregulated lncRNA LINC01010 attenuates cell proliferation by interacting with vimentin. *Int. J. Mol. Sci.* 22, 12497. doi:10.3390/ijms222212497

Huang Da, W., Sherman, B. T., and Lempicki, R. A. (2009). Bioinformatics enrichment tools: Paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res.* 37, 1–13. doi:10.1093/nar/gkn923

Huang, G., Yang, Y., Lv, M., Huang, T., Zhan, X., Kang, W., et al. (2020). Novel lncRNA SFTA1P promotes tumor growth by down-regulating miR-4766-5p via PI3K/AKT/mTOR signaling pathway in hepatocellular carcinoma. *Onco Targets Ther.* 13, 9759–9770. doi:10.2147/OTT.S248660

Ireland, A. S., and Oliver, T. G. (2020). Neutrophils create an ImpeNETrable shield between tumor and cytotoxic immune cells. *Immunity* 52, 729–731. doi:10.1016/j.immuni.2020.04.009

Kalinsky, K., Jacks, L. M., Heguy, A., Patil, S., Drobnjak, M., Bhanot, U. K., et al. (2009). PIK3CA mutation associates with improved outcome in breast cancer. *Clin. Cancer Res.* 15, 5049–5059. doi:10.1158/1078-0432.CCR-09-0632

Kaltenmeier, C., Yazdani, H. O., Morder, K., Geller, D. A., Simmons, R. L., and Tohme, S. (2021). Neutrophil extracellular traps promote T cell exhaustion in the tumor microenvironment. *Front. Immunol.* 12, 785222. doi:10.3389/fimmu.2021.785222

Kim, S. T., Chu, Y., Misoi, M., Suarez-Almazor, M. E., Tayar, J. H., Lu, H., et al. (2022). Distinct molecular and immune hallmarks of inflammatory arthritis induced by immune checkpoint inhibitors for cancer therapy. *Nat. Commun.* 13, 1970. doi:10.1038/s41467-022-29539-3

Kostine, M., Rouxel, L., Barnetche, T., Veillon, R., Martin, F., Dutriaux, C., et al. (2018). Rheumatic disorders associated with immune checkpoint inhibitors in patients with cancer-clinical aspects and relationship with tumour response: A single-centre prospective cohort study. *Ann. Rheum. Dis.* 77, 393–398. doi:10.1136/annrheumdis-2017-212257

Li, B., Chan, H. L., and Chen, P. (2019). Immune checkpoint inhibitors: Basics and challenges. *Curr. Med. Chem.* 26, 3009–3025. doi:10.2174/0929867324666170804143706

Li, L., Yu, R., Cai, T., Chen, Z., Lan, M., Zou, T., et al. (2020). Effects of immune cells and cytokines on inflammation and immunosuppression in the tumor microenvironment. *Int. Immunopharmacol.* 88, 106939. doi:10.1016/j.intimp.2020.106939

Lin, C., Zheng, M., Yang, Y., Chen, Y., Zhang, X., Zhu, L., et al. (2022). Knockdown of lncRNA ACTA2-AS1 reverses cisplatin resistance of ovarian cancer cells via inhibition of miR-378a-3p-regulated Wnt5a. *Bioengineered* 13, 9829–9838. doi:10.1080/21655979.2022.2061181

Liu, X., Zhang, X., and Du, S. (2022). Long non-coding RNA ACTA2-AS1 inhibits the cisplatin resistance of non-small cell lung cancer cells through inhibiting autophagy by suppressing TSC2. *Cell Cycle* 21, 368–378. doi:10.1080/15384101.2021.2020433

Luo, L., Wang, M., Li, X., Luo, C., Tan, S., Yin, S., et al. (2020). A novel mechanism by which ACTA2-AS1 promotes cervical cancer progression: Acting as a ceRNA of miR-143-3p to regulate SMAD3 expression. *Cancer Cell Int.* 20, 372. doi:10.1186/s12935-020-01471-w

Martins-Cardoso, K., Almeida, V. H., Bagri, K. M., Rossi, M. I. D., Mermelstein, C. S., KöNIG, S., et al. (2020). Neutrophil Extracellular Traps (NETs) Promote Pro-Metastatic Phenotype in Human Breast Cancer Cells through Epithelial-Mesenchymal Transition, *Cancers (Basel)*. 12. 1542. doi:10.3390/cancers12061542

Mehraj, U., Ganai, R. A., Macha, M. A., Hamid, A., Zargar, M. A., Bhat, A. A., et al. (2021). The tumor microenvironment as driver of stemness and therapeutic resistance in breast cancer: New challenges and therapeutic opportunities. *Cell Oncol. (Dordr)* 44, 1209–1229. doi:10.1007/s13402-021-00634-9

Mosele, F., Stefanovska, B., Lusque, A., Tran Dien, A., Garberis, I., Droin, N., et al. (2020). Outcome and molecular landscape of patients with PIK3CA-mutated metastatic breast cancer. *Ann. Oncol.* 31, 377–386. doi:10.1016/j.annonc.2019.11.006

Niu, H., Wang, J., Li, H., and He, P. (2011). Rapamycin potentiates cytotoxicity by docetaxel possibly through downregulation of Survivin in lung cancer cells. *J. Exp. Clin. Cancer Res.* 30, 28. doi:10.1186/1756-9966-30-28

Okiyama, N., and Tanaka, R. (2022). Immune-related adverse events in various organs caused by immune checkpoint inhibitors. *Allergol. Int.* 71, 169–178. doi:10.1016/j.alit.2022.01.001

Papayannopoulos, V. (2018). Neutrophil extracellular traps in immunity and disease. *Nat. Rev. Immunol.* 18, 134–147. doi:10.1038/nri.2017.105

PeñA-Romero, A. C., and Orenes-PiñERO, E. (2022). Dual Effect of Immune Cells within Tumour Microenvironment: Pro- and Anti-Tumour Effects and Their Triggers, *Cancers (Basel)*. 14. 1681. doi:10.3390/cancers14071681

Pondé, N. F., Zardavas, D., and Piccart, M. (2019). Progress in adjuvant systemic therapy for breast cancer. *Nat. Rev. Clin. Oncol.* 16, 27–44. doi:10.1038/s41571-018-0089-9

Sha, D., Jin, Z., Budczies, J., Kluck, K., Stenzinger, A., and Sinicrope, F. A. (2020). Tumor mutational burden as a predictive biomarker in solid tumors. *Cancer Discov.* 10, 1808–1825. doi:10.1158/2159-8290.CD-20-0522

Siegel, R. L., Miller, K. D., Fuchs, H. E., and Jemal, A. (2022). Cancer statistics, 2022. *CA Cancer J. Clin.* 72, 7–33. doi:10.3322/caac.21708

Sun, C. Y., Li, Y. Z., Cao, D., Zhou, Y. F., Zhang, M. Y., and Wang, H. Y. (2021). Rapamycin and trametinib: A rational combination for treatment of NSCLC. *Int. J. Biol. Sci.* 17, 3211–3223. doi:10.7150/ijbs.62752

Tang, L., Wei, D., Xu, X., Mao, X., Mo, D., Yan, L., et al. (2021). Long non-coding RNA MIR200CHG promotes breast cancer proliferation, invasion, and drug resistance by interacting with and stabilizing YB-1. *NPJ Breast Cancer* 7, 94. doi:10.1038/s41523-021-00293-x

Teijeira, Á., Garasa, S., Gato, M., Alfaro, C., Migueliz, I., Cirella, A., et al. (2020). CXCR1 and CXCR2 chemokine receptor agonists produced by tumors induce neutrophil extracellular traps that interfere with immune cytotoxicity. *Immunity* 52, 856–871. doi:10.1016/j.immuni.2020.03.001

Tekpli, X., Lien, T., RøSSEVOLD, A. H., Nebdal, D., Borgen, E., Ohnstad, H. O., et al. (2019). An independent poor-prognosis subtype of breast cancer defined by a distinct tumor immune microenvironment. *Nat. Commun.* 10, 5499. doi:10.1038/s41467-019-13329-5

Tesch, M. E., and Gelmon, K. A. (2020). Targeting HER2 in breast cancer: Latest developments on treatment sequencing and the introduction of biosimilars. *Drugs* 80, 1811–1830. doi:10.1007/s40265-020-01411-y

Wang, Y., Liu, F., Chen, L., Fang, C., Li, S., Yuan, S., et al. (2022). Neutrophil extracellular traps (NETs) promote non-small cell lung cancer metastasis by suppressing lncRNA

MIR503HG to activate the NF-κB/NLRP3 inflammasome pathway. *Front. Immunol.* 13, 867516. doi:10.3389/fimmu.2022.867516

Yang, J., Liu, F., Wang, Y., Qu, L., and Lin, A. (2022). LncRNAs in tumor metabolic reprogramming and immune microenvironment remodeling. *Cancer Lett.* 543, 215798. doi:10.1016/j.canlet.2022.215798

Yardim-Akaydin, S., Karahalil, B., and Baytas, S. N. (2022). New therapy strategies in the management of breast cancer. *Drug Discov. Today* 27, 1755–1762. doi:10.1016/j.drudis.2022.03.014

Ying, K., Wang, L., Long, G., Lian, C., Chen, Z., and Lin, W. (2020). ACTA2-AS1 suppresses lung adenocarcinoma progression via sequestering miR-378a-3p and miR-4428 to elevate SOX7 expression. *Cell Biol. Int.* 44, 2438–2449. doi:10.1002/cbin.11451

You, K. S., Yi, Y. W., Cho, J., and Seong, Y. S. (2021). Dual inhibition of AKT and MEK pathways potentiates the anti-cancer effect of gefitinib in triple-negative breast cancer cells. *Cancers (Basel)* 13, 1205. doi:10.3390/cancers13061205

Zhang, C., Liang, Y., Zhang, C. D., Pei, J. P., Wu, K. Z., Li, Y. Z., et al. (2021). The novel role and function of LINC01235 in metastasis of gastric cancer cells by inducing epithelial-mesenchymal transition. *Genomics* 113, 1504–1513. doi:10.1016/j.ygeno.2021.03.027

Zhou, R. J., and Lv, H. Z. (2019). Knockdown of ACTA2-AS1 promotes liver cancer cell proliferation, migration and invasion. *Mol. Med. Rep.* 19, 2263–2270. doi:10.3892/mmr.2019.9856

Zhu, B., Finch-Edmondson, M., Leong, K. W., Zhang, X., Mitheera, V., Lin, Q. X. X., et al. (2021). LncRNA SFTA1P mediates positive feedback regulation of the Hippo-YAP/TAZ signaling pathway in non-small cell lung cancer. *Cell Death Discov.* 7, 369. doi:10.1038/s41420-021-00761-0

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 16 February 2023
DOI 10.3389/fgene.2023.1097706

OPEN ACCESS

EDITED BY
Xiao Wang,
Konge Larsen ApS, Denmark

REVIEWED BY
Enoch Appiah Adu Gyamfi,
The State University of New York (SUNY), United States
Zhiming Zhao,
Second Hospital of Hebei Medical University, China

*CORRESPONDENCE
Changhong Zhu,
reprodcentre@163.com

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

RECEIVED 14 November 2022
ACCEPTED 06 February 2023
PUBLISHED 16 February 2023

CITATION
Zou L, Feng Q, Xia W and Zhu C (2023), Bioinformatics analysis of the common targets of miR-223-3p, miR-122-5p, and miR-93-5p in polycystic ovarian syndrome.
Front. Genet. 14:1097706.
doi: 10.3389/fgene.2023.1097706

Bioinformatics analysis of the common targets of miR-223-3p, miR-122-5p, and miR-93-5p in polycystic ovarian syndrome

Liping Zou, Qiwen Feng, Wei Xia and Changhong Zhu*

Institute of Reproductive Health, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China

Polycystic ovarian syndrome (PCOS) is one of the most common gynecological endocrine disorders. MicroRNAs (miRNAs) play extensive roles in the pathogenesis of PCOS and can serve as potential diagnostic markers. However, most studies focused on the regulatory mechanisms of individual miRNAs, and the combined regulatory effects of multiple miRNAs remain unclear. The aim of this study was to identify the common targets of miR-223-3p, miR-122-5p, and miR-93-5p; and assess the transcript levels of some of these targets in PCOS rat ovaries. Transcriptome profiles of granulosa cells from PCOS patients were obtained from the Gene Expression Omnibus (GEO) database to identify differentially expressed genes (DEGs). A total of 1,144 DEGs were screened, 204 of which were upregulated and 940 were downregulated. According to the miRWalk algorithm, 4,284 genes were targeted by all three miRNAs at the same time, and intersection with DEGs was used to obtain candidate target genes. A total of 265 candidate target genes were screened, and the detected target genes were subjected to Gene ontology (GO) and KEGG pathway enrichment, followed by PPI network analysis. Then, qRT-PCR was used to determine the levels of 12 genes in PCOS rat ovaries. The expressions of 10 of these genes were found to be consistent with our bioinformatics results. In conclusion, JMJD1C, PLCG2, SMAD3, FOSL2, TGFB1, TRIB1, GAS7, TRIM25, NFYA, and CALCRL may participate in the development of PCOS. Our findings contribute to the identification of biomarkers that may promote the effective prevention and treatment of PCOS in the future.

KEYWORDS

PCOS, miRNA, bioinformatics analysis, biomarkers, granulosa cells

Introduction

PCOS is a common endocrine and metabolic disease in women of reproductive age. According to the Rotterdam diagnostic criteria, it encompasses all combinations of ovulation dysfunction, hyperandrogenemia, and polycystic ovary morphology (Ehrmann, 2005; Rosenfield and Ehrmann, 2016). Patients with PCOS often have a range of other health problems, including infertility, insulin resistance (IR), and obesity (Patel, 2018). Several studies suggest that genetic, environmental, and epigenetic factors may play important roles in the pathogenesis of PCOS (Escobar-Morreale, 2018), but the exact mechanism remains largely unclear.

MiRNAs are small endogenous single-stranded non-coding RNA molecules consisting of 18–25 nucleotides that regulate gene expression at the post-transcriptional level (Fabian et al., 2010). It has been reported that miRNAs are involved in various signaling pathways in PCOS, including amino acid metabolism, hormone regulation, cell differentiation, etc. (Imbar and Eisenberg, 2014; Mu et al., 2021). Differentially expressed miRNAs play important roles in PCOS pathogenesis and serve as potential diagnostic markers (Deswal and Dang, 2020). In addition, miRNAs regulate many biological processes associated with obesity, including adipogenesis, insulin secretion, and glucose uptake (Butler et al., 2020). Inflammation of adipose tissue in obese patients contributes to obesity-related metabolic dysfunction, such as insulin resistance and type 2 diabetes (Wu and Ballantyne, 2020).

Upregulation of miR-223 and miR-93 was detected in adipose tissue of PCOS patients (Chen et al., 2013; Udesen et al., 2020), while miR-93 was also significantly upregulated in the ovarian cortex and follicular fluid of PCOS patients (Jiang et al., 2015; Butler et al., 2019). Interestingly, miR-223 and miR-93 were not only overexpressed in adipose tissue of PCOS patients, but also in control patients with IR, with which it was positively correlated *in vivo* (Chen et al., 2013; Chuang et al., 2015). MiR-223 and miR-93 downregulate GLUT4 expression and inhibit insulin-stimulated glucose uptake in adipocytes, suggesting that they may play an important role in other IR-related diseases such as T2DM and obesity. Increased serum miR-122 levels were found in PCOS patients with impaired glucose metabolism (Jiang et al., 2016). Elevated miR-122 in circulation was positively associated with obesity and IR in young adults (Wang et al., 2015). Furthermore, miR-122 and miR-223 were found to be increased in obesity or hyperglycemia, and their intracellular roles are related to the development of IR (Murri et al., 2018; Udesen et al., 2020).

Experimental studies confirmed that the expression of miR-223-3p, miR-122-5p and miR-93-5p was significantly upregulated in PCOS (Chen et al., 2013; Jiang et al., 2016). This may indicate that these differential miRNAs are key molecules involved in the pathological process of PCOS. However, most studies focused the regulation mechanism of individual miRNAs, while the combined regulatory effect of multiple miRNAs remained unclear.

Each miRNA may affect hundreds of targets, while itself being regulated by several distinct miRNAs. Since the efficacy of single markers is limited, and multi-marker-based models can provide more reliable information for the diagnosis and therapeutic management of PCOS, we aimed to identify the common targets of miR-223-3p, miR-122-5p, and miR-93-5p; and assess the transcript levels of some of these targets in PCOS rat ovaries.

Materials and methods

Microarray data and DEGs screening

The gene expression profile GSE34526 was acquired from the GEO website. GSE34526 contained 3 granulosa cells samples (GSM850527-GSM850529) from female controls and 7 granulosa cells samples (GSM850530-GSM850536) from PCOS patients, based on the GPL570 [HG-U133_Plus_2] Affymetrix Human Genome U133 Plus 2.0 Array. DEGs between PCOS and normal samples were obtained from the GEO database by GEO2R analysis (https://www.ncbi.nlm.nih.gov/geo/geo2r). A p-value <0.05 and $|\log FC| > 2$ as used as the DEGs cut-off criteria. A volcano plot was generated to visualize DEG expression changes using GraphPad Prism 8.0 software (GraphPad, United States).

Prediction of target genes

The miRWalk Version 3.0 (https://mirwalk.umm.uni-heidelberg.de), which is linked to three online databases (TargetScan, miRDB, and miRTarBase), stores predicted data including experimentally verified miRNA-target interactions. The targets genes of miR-223-3p, miR-122-5p, and miR-93-5p were downloaded from the miRWalk 3.0 database, and the intersection genes were selected for further analysis. Then, the overlapping genes among DEGs and targets genes of the miRNAs were obtained using the Venn diagrams tool (https://bioinformatics.psb.ugent.be).

Functional enrichment analysis of target genes

GO functional enrichment and KEGG pathway analyses were performed for the gene overlaps using Metascape (https://metascape.org). Metascape always uses the latest data, which integrates data sources from GO, KEGG, UniProt, and DrugBank to achieve pathway enrichment and biological process annotation. $p < 0.01$ was considered statistically significant.

PPI network analysis of target genes

The Search Tool for the Retrieval of Interacting Genes (STRING) online database (https://string-db.org) was used to predict the protein-protein interaction (PPI) network specific to target genes. Then, PPI networks were visualized using Cytoscape version 3.8.2 (https://cytoscape.org) and the clusters (highly interconnected regions) in the PPI network were identified using the Cytoscape plugin MCODE.

Genes related to obesity with PCOS

The DEGs related to obesity in PCOS were identified in the dataset of GSE80432, including four granulosa cell samples (GSM2127203-GSM2127204, GSM2127215-GSM2127216) from normal weight PCOS patients and four granulosa cell samples (GSM2127209, GSM2127211-GSM2127212, GSM2127214) from obese PCOS patients. The datasets were analyzed on the GPL6244 [HuGene-1_0-st] Affymetrix Human Gene 1.0 ST Array [transcript (gene) version]. The p-value <0.05 and Fold Change >1.5 were set as the cut-off criteria for DEGs. The Venn diagram tool was used to identify the genes shared between DEGs and targets genes of miRNAs.

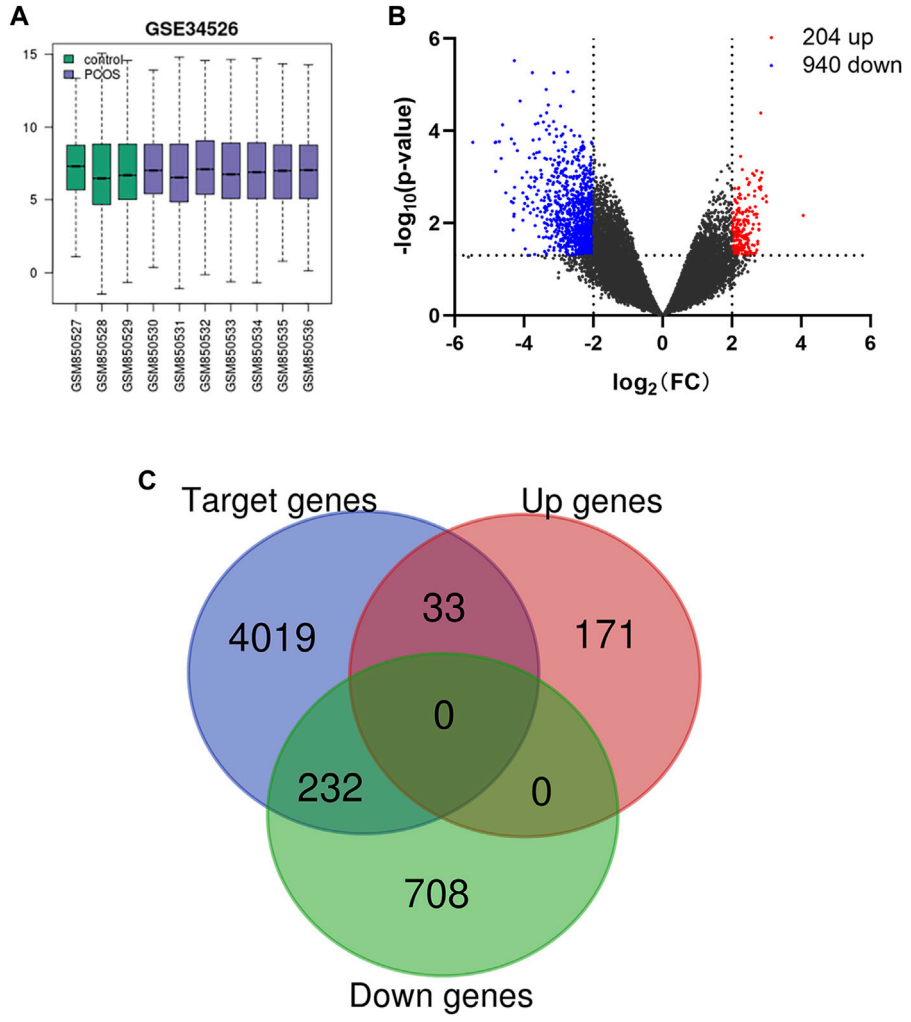


FIGURE 1
(A) Box plot of gene expression profiles after standardization. **(B)** Volcano plot of DEGs in PCOS. Red dots represent upregulation, blue dots represent downregulation, and gray dots represent no differential expression. **(C)** Venn diagram of DEGs and miRNA target genes.

TFs related to miR-223-3p, miR-122-5p and miR-93-5p

The prediction of TFs related to miR-223-3p, miR-122-5p, and miR-93-5p was conducted using TransmiR v2.0 (https://www.cuilab.cn/transmir), which is a database for TF-miRNA regulatory association. Subsequently, the TFs involved in PCOS shared between DEGs and predicted TFs were screened using the Venn diagram tool.

Animal model

Female Sprague–Dawley rats (n = 20; 21 days old) were obtained from the Hubei Provincial Center for Disease Control and Prevention (Wuhan, China). Animal experiments were approved by the Animal Ethics Committee of Tongji Medical College, Huazhong University of Science and Technology (Wuhan, China). PCOS was induced by injection of DHEA (60 mg/kg body weight) dissolved in 0.2 mL sesame oil. The control group was injected with the vehicle (0.2 mL sesame oil). Daily treatment was continued for up to 21 days. In this study, we used a rat model that exhibits reproductive and metabolic abnormalities similar to human PCOS to uncover molecular mechanisms.

Quantitative real-time polymerase chain reaction (qRT-PCR)

Total RNA was extracted from the ovaries of PCOS rats using TRIzol reagent (Ambion). The cDNA as generated using the PrimeScript™ RT reagent kit or miRNA First-Strand Synthesis kit (Takara). The qRT-PCR was performed using qPCR SYBR Green Master Mix (Vazyme). GAPDH and U6 were used as internal controls for mRNA and miRNA expression, respectively. The relative expression of the mRNAs and miRNAs was calculated using the $2^{-(\Delta\Delta Ct)}$ method.

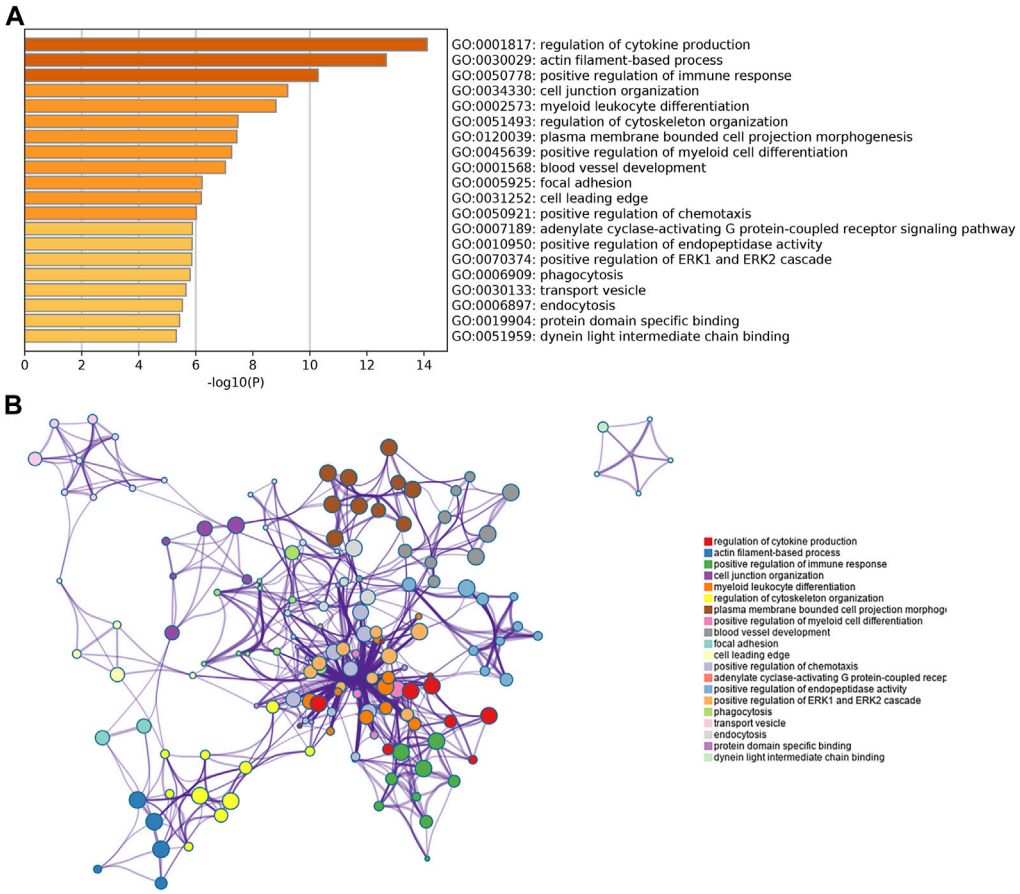


FIGURE 2
GO enrichment analysis of overlapping genes obtained from the intersection between DEGs and miRNA target genes. **(A)** Bar chart of the most highly enriched terms. **(B)** Network of the most highly enriched terms.

Statistical analysis

Data are presented as the means ± SD. Data were analyzed using Student's t-test. Differences with p-values <0.05 were considered statistically significant. The results were analyzed and visualized using GraphPad Prism 8.0 software (GraphPad Inc, United States).

Results

Identification of DEGs and target genes in PCOS

After normalization of the microarray data from GSE34526 (Figure 1A), a total of 1,144 DEGs were identified in PCOS patients compared with female controls, consisting of 204 upregulated genes and 940 downregulated genes (Figure 1B). The downstream target genes associated with miRNAs were identified, including 6,653 for miR-223-3p, 12,132 for miR-122-5p, and 15,170 for miR-93-5p. Among them, 4,284 target genes were shared among all three miRNAs. Finally, a total of 265 overlapping genes shared between DEGs and target genes were obtained, including 33 upregulated genes and 232 downregulated genes (Figure 1C).

GO and KEGG analysis of candidate target genes in PCOS

To further explore the function of the identified target genes, GO term and KEGG pathway enrichment analyses were performed. GO analysis was divided into the MF, CC, and BP categories (Figure 2). The results showed that 265 genes were mainly enriched in the regulation of cytokine production, actin filament-based processes, positive regulation of the immune response, cell junction organization, and myeloid leukocyte differentiation. Moreover, 12 KEGG pathways were overrepresented, including *salmonella* infection, morphine addiction, osteoclast differentiation, cell adhesion molecules, adherens junction, and endocytosis (Figure 3).

PPI network construction and hub gene selection

To explore the interactions among proteins encoded by the identified target genes, a PPI network was constructed, including 265 nodes and 389 edges (Figure 4). The three downregulated genes JMJD1C, PLCG2, and SERPINA1 were selected as hub genes based on enrichment degree ≥2 as the cutoff criterion.

A

Pathway ID	Pathway name	p	Included genes
ko05132	Salmonella infection	0.0002	CDC42,DYNC1H1,IFNGR1,TLR4,WASF2,NLRC4
ko05032	Morphine addiction	0.0002	GRK2,ARRB1,GABRA2,PDE4B,PDE4D,ADCY4
ko04380	Osteoclast differentiation	0.0002	CREB1,CSF1R,FOSL2,IFNGR1,PLCG2,TGFB1,LILRB3
ko05152	Tuberculosis	0.0003	C3,CD74,CREB1,IFNGR1,NFYA,NFYC,TGFB1,TLR4,CD4
ko05142	Chagas disease (American trypanosomiasis)	0.0004	C3,IFNGR1,SMAD3,TGFB1,TLR4,CFLAR,LRP1,PECAM1
hsa04514	Cell adhesion molecules (CAMs)	0.0004	CD4,CD6,CD28,NCAM1,PECAM1,CLDN16,CLDN18,CDC42 ,PLCG2
hsa04520	Adherens junction	0.0006	CDC42,SMAD3,IQGAP1,WASF2,SORBS1
hsa04144	Endocytosis	0.0008	GRK2,ARRB1,CDC42,CSF1R,IGF2R,SMAD3,TGFB1,GIT2, VPS35,BRCA2,PLCG2,PTEN,PTGER3,LPAR2,ADCY4
hsa04142	Lysosome	0.0011	ARSA,IGF2R,LAPTM5,AP1S2,LITAF,ARSG
hsa04666	Fc gamma R-mediated phagocytosis	0.0017	CDC42,INPP5D,PLCG2,WASF2,SCIN,INPP4A,PTEN
hsa05202	Transcriptional misregulation in cancer	0.0033	CSF1R,ZEB1,KDM6A,IL1R2,NFKBIZ,BCL11B,JMJD1C
ko05100	Bacterial invasion of epithelial cells	0.0058	CDC42,GAB1,WASF2,SHC4,PLCG2,PTEN

B

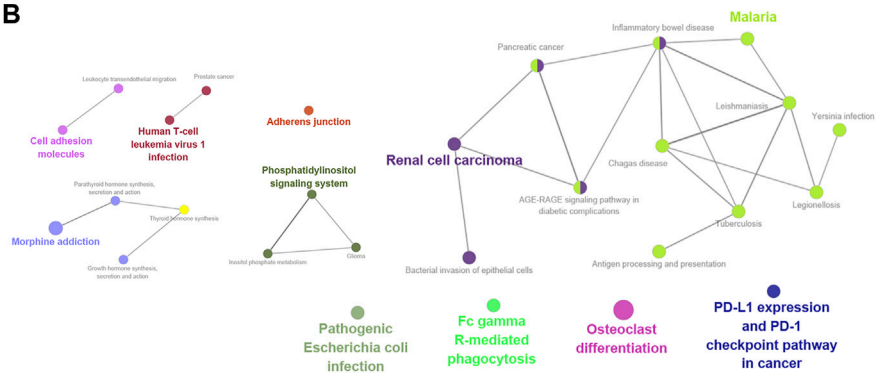


FIGURE 3
KEGG enrichment analysis of overlapping genes obtained from the intersection between DEGs and miRNA target genes. **(A)** Genes involved in enriched signaling pathways. **(B)** Network of the most highly enriched signaling pathways.

TFs associated with miR-223-3p, miR-122-5p and miR-93-5p

A total of 229 TFs associated with miR-223-3p, miR-122-5p, and miR-93-5p were downloaded from the TransmiR database. Compared with normal granulosa cells, 7 TFs were differentially downregulated in PCOS, including TGFB1, SMAD3, FOSL2, JMJD1C, CREB1, TRIM25, and NFYA.

DEGs related to obesity in PCOS

To further clarify the potential mechanisms of miRNAs related to adipogenesis in obese patients with PCOS, the genes that were differentially expressed in granulosa cells and associated with obesity were selected for further analysis. Based on the intersection between 265 genes and DEGs from the GSE80432 dataset, one upregulated and four downregulated genes were obtained, including CALCRL, TGFB1, TRIB1, GAS7, and FOSL2. TGFB1 and FOSL2 are both major transcription factors and closely related to the pathological process of obesity in PCOS.

Validation of the identified DEGs in a rat model of PCOS

To validate the results of bioinformatics analysis, we detected the expression of three miRNAs and 12 predicted target genes using qRT-PCR in PCOS model rats. As shown in Figure 5, miR-223-3p, miR-122-5p, and miR-93-5p were significantly upregulated in the PCOS group compared with control group. JMJD1C, PLCG2, SMAD3, FOSL2, TGFB1, TRIB1, GAS7, TRIM25, and NFYA were downregulated, while CALCRL was highly expressed in PCOS rats. Furthermore, SERPINA1 and CREB1 were not validated in PCOS rats.

Discussion

PCOS is a common endocrine and metabolic disorder that can easily lead to ovulation dysfunction in women of childbearing age, resulting in infertility. The development of oocytes is inseparable from the role of follicular granulosa cells, which provide nutrients for oocytes through gap junctions and regulate oocyte development through paracrine signals (Clarke, 2018). In this study, a total of 265 candidate target genes were screened by integrating the

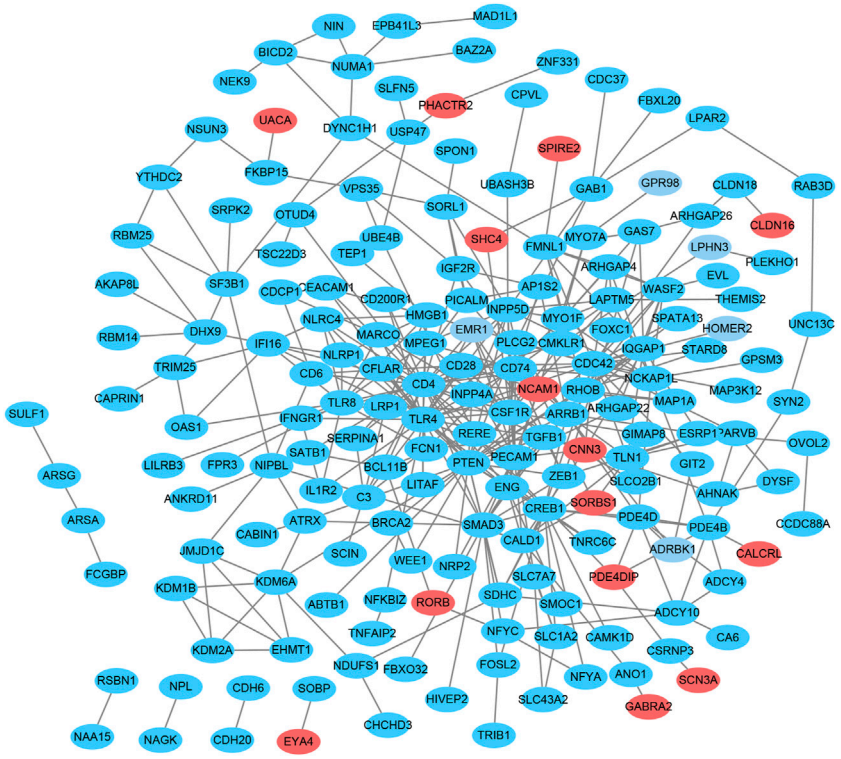


FIGURE 4
Protein-protein interaction network analysis of the overlapping genes. The PPI network consists of 265 nodes and 389 edges. Red indicates upregulation, and blue indicates downregulation.

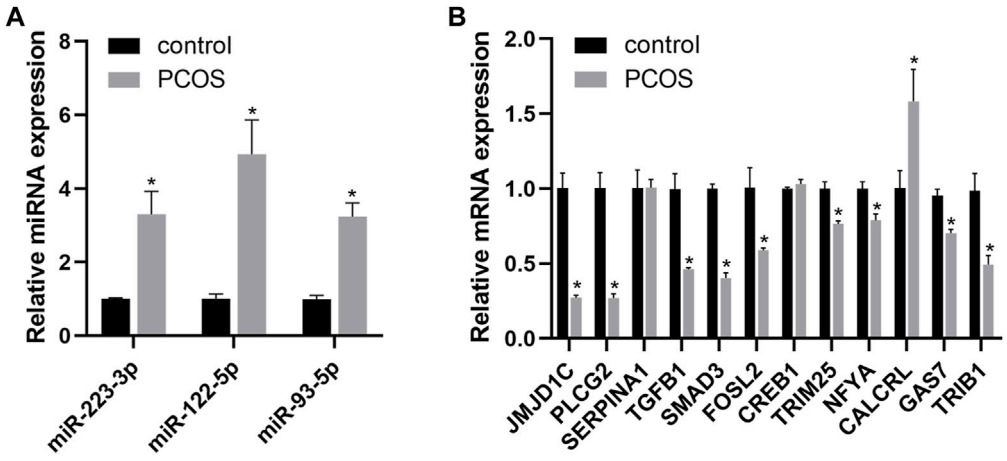


FIGURE 5
The qPCR verification results of three miRNAs **(A)** and twelve candidate target genes **(B)** in the rat model of PCOS. An asterisk indicates a significant difference between the two groups ($*p < 0.05$) according to Student's t-test.

information of DEGs and miRNA target genes in granulosa cells of PCOS patients with insulin resistance, among which 7 TF genes and five target genes closely related to obesity were identified. The function of candidate target genes was analyzed and the enriched GO terms included regulation of cytokine production, actin filament-based process, positive regulation of immune response, cell junction organization, and myeloid leukocyte differentiation.

We selected three DEGs as hub genes, the top three of which were JMJD1C, SERPINA1, and PLCG2. It stands to reason that miRNAs may play a major role in PCOS through these key target genes. We verified their transcriptional levels in a rat model of PCOS and found that the expression of JMJD1C and PLCG2 was downregulated, while the expression of SERPINA1 was unchanged. The mechanism may be that miRNA mediates the attenuation of JMJD1C and PLCG2 mRNA, hindering the

translation of SERPINA1 mRNA without affecting its stability. Among them, JMJD1C and SERPINA1 play important roles in adipogenesis. JMJD1C regulates the induction of adipogenic transcription factors *via* H3K9me2 (Buerger et al., 2017) and promotes adipogenesis *in vivo* to increase liver and plasma triglyceride levels (Viscarra et al., 2020). SERPINA1, which may serve as an important molecular marker of obesity, affects energy expenditure by regulating the AMPK pathway and promotes the development of obesity-related metabolic complications (Mansuy-Aubert et al., 2013). PLCG2 regulates oxidative stress in hypoxic-ischemic encephalopathy, and activated PLCG2 signaling to attenuate oxidative stress-induced neuronal degeneration and apoptosis (Hu et al., 2020). Ovarian oxidative stress imbalance is a key feature of PCOS, suggesting that downregulation of PLCG2 is closely related to PCOS (Murri et al., 2013).

The selected transcription factors included TGFB1, SMAD3, FOSL2, JMJD1C, CREB1, TRIM25, and NFYA. The mRNA levels of TGFB1, SMAD3, FOSL2, TRIM25, and NFYA were decreased in our animal model. TGFB1 partially inhibits adipogenesis through SMAD3 (Tsurutani et al., 2011), and downregulation of the TGFB1 signaling pathway may promote adipogenesis by altering the expression of adipogenic genes to change the PCOS phenotype (Li and Wu, 2020). Furthermore, androgens were found to induce ovarian fibrosis through the TGFB1 signaling pathway in a rat model of PCOS (Wang et al., 2018). It is speculated that TGFB1 mainly regulates adipogenesis in PCOS and has less effect on ovarian fibrosis. FOSL2 is known to be a key regulator of adipokine LEP expression in obese mice and humans (Wrann et al., 2012). TRIM25 enhances the antioxidant defense by activating Nrf2 (Liu et al., 2020), and downregulation of TRIM25 may contribute to the oxidative stress imbalance phenotype of PCOS. NFYA regulates the gene expression of adiponectin in adipose tissue, one of the adipokines secreted by adipocytes that regulate energy homeostasis related to insulin sensitivity (Park et al., 2004).

Although CALCRL gene expression was reported to be lower in the adipose tissue of obese patients (Aguilera et al., 2015; Kim et al., 2020), CALCRL transcription was increased in the rat PCOS model. This difference may be caused by differences between species. TRIB1 knockout mice exhibit obesity and increased lipid accumulation in their livers (Bauer et al., 2015). TRIB1 can promote adipose tissue thermogenesis by regulating mitochondrial function (Zhang et al., 2021). PCOS patients have lower brown adipose tissue (BAT) activity and reduced BAT thermogenesis, which is associated with increased insulin resistance (Robinson et al., 1992; Shorakae et al., 2019). The expression of the GAS7 gene was found to be decreased in the ovaries of obese women, and its transcript was also downregulated in the ovaries of DHT-treated rats (Ruebel et al., 2017), which was consistent with our prediction. These results demonstrate the accuracy of predicting miRNA target genes and demonstrate that miRNAs regulate the pathological process of PCOS.

Although this study revealed the common targets of three miRNAs in PCOS using bioinformatics analyses and validated them using a rat model, there are still some limitations in the present study. Firstly, the GSE34526 and GSE80432 datasets were collected from granulosa cells samples, which may be the reason why the expression of SERPINA1 and CREB1 in PCOS rats was inconsistent with the bioinformatics analyses. In addition, the miRNA–mRNA relationships in PCOS were based on target prediction. However, the fact that the predicted targets are dysregulated in the PCOS rat model does not mean that they were targeted by the three miRNAs, which needs further experimental validation. Thus, a knockdown or overexpression of the miRNAs in granulosa cells and assessment of the mRNA and protein levels of the predicted targets would have yielded more reliable information.

In conclusion, JMJD1C, PLCG2, SMAD3, FOSL2, TGFB1, TRIB1, GAS7, TRIM25, NFYA, and CALCRL were found to be potential common targets of the three miRNAs. These possible targets were downregulated in PCOS rat ovaries, and so may serve as biomarkers for PCOS diagnosis and treatment. Instead of relying solely on the target gene prediction of a single miRNA, we used multiple miRNAs identified in clinical samples to predict common target genes, which provides strong evidence for the involvement of miRNAs in PCOS and broadens our understanding of gene expression changes in PCOS.

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.

Ethics statement

The animal study was reviewed and approved by Animal Ethics Committee of Tongji Medical College, Huazhong University of Science and Technology (Wuhan, China).

Author contributions

LZ conceived the study, analyzed the data, and wrote the manuscript. QF performed experiments. CZ and WX reviewed the paper.

Funding

This work was supported by the grants from the National Key Research and Development Program of China (2022YFC2702705).

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.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2023.1097706/full#supplementary-material

References

Aguilera, C. M., Gomez-Llorente, C., Tofe, I., Gil-Campos, M., Cañete, R., and Gil, Á. (2015). Genome-wide expression in visceral adipose tissue from obese prepubertal children. *Int. J. Mol. Sci.* 16 (4), 7723–7737. doi:10.3390/ijms16047723

Bauer, R. C., Sasaki, M., Cohen, D. M., Cui, J., Smith, M. A., Yenilmez, B. O., et al. (2015). Tribbles-1 regulates hepatic lipogenesis through posttranscriptional regulation of C/EBPα. *J. Clin. Invest.* 125 (10), 3809–3818. doi:10.1172/jci77095

Buerger, F., Müller, S., Ney, N., Weiner, J., Heiker, J. T., Kallendrusch, S., et al. (2017). Depletion of Jmjd1c impairs adipogenesis in murine 3T3-L1 cells. *Biochim. Biophys. Acta Mol. Basis Dis.* 1863 (7), 1709–1717. doi:10.1016/j.bbadis.2017.05.011

Butler, A. E., Ramachandran, V., Hayat, S., Dargham, S. R., Cunningham, T. K., Benurwar, M., et al. (2019). Expression of microRNA in follicular fluid in women with and without PCOS. *Sci. Rep.* 9 (1), 16306. doi:10.1038/s41598-019-52856-5

Butler, A. E., Ramachandran, V., Sathyapalan, T., David, R., Gooderham, N. J., Benurwar, M., et al. (2020). microRNA expression in women with and without polycystic ovarian syndrome matched for body mass index. *Front. Endocrinol. (Lausanne)* 11, 206. doi:10.3389/fendo.2020.00206

Chen, Y. H., Heneidi, S., Lee, J. M., Layman, L. C., Stepp, D. W., Gamboa, G. M., et al. (2013). miRNA-93 inhibits GLUT4 and is overexpressed in adipose tissue of polycystic ovary syndrome patients and women with insulin resistance. *Diabetes* 62 (7), 2278–2286. doi:10.2337/db12-0963

Chuang, T. Y., Wu, H. L., Chen, C. C., Gamboa, G. M., Layman, L. C., Diamond, M. P., et al. (2015). MicroRNA-223 expression is upregulated in insulin resistant human adipose tissue. *J. Diabetes Res.* 2015, 943659. doi:10.1155/2015/943659

Clarke, H. J. (2018). Regulation of germ cell development by intercellular signaling in the mammalian ovarian follicle. *Wiley Interdiscip. Rev. Dev. Biol.* 7 (1). doi:10.1002/wdev.294

Deswal, R., and Dang, A. S. (2020). Dissecting the role of micro-RNAs as a diagnostic marker for polycystic ovary syndrome: A systematic review and meta-analysis. *Fertil. Steril.* 113 (3), 661–669.e2. doi:10.1016/j.fertnstert.2019.11.001

Ehrmann, D. A. (2005). Polycystic ovary syndrome. *N. Engl. J. Med.* 352 (12), 1223–1236. doi:10.1056/NEJMra041536

Escobar-Morreale, H. F. (2018). Polycystic ovary syndrome: Definition, aetiology, diagnosis and treatment. *Nat. Rev. Endocrinol.* 14 (5), 270–284. doi:10.1038/nrendo.2018.24

Fabian, M. R., Sonenberg, N., and Filipowicz, W. (2010). Regulation of mRNA translation and stability by microRNAs. *Annu. Rev. Biochem.* 79, 351–379. doi:10.1146/annurev-biochem-060308-103103

Hu, X., Li, S., Doycheva, D. M., Huang, L., Lenahan, C., Liu, R., et al. (2020). Rh-CSF1 attenuates oxidative stress and neuronal apoptosis via the CSF1R/PLCG2/PKA/UCP2 signaling pathway in a rat model of neonatal HIE. *Oxid. Med. Cell Longev.* 2020, 6801587. doi:10.1155/2020/6801587

Imbar, T., and Eisenberg, I. (2014). Regulatory role of microRNAs in ovarian function. *Fertil. Steril.* 101 (6), 1524–1530. doi:10.1016/j.fertnstert.2014.04.024

Jiang, L., Huang, J., Chen, Y., Yang, Y., Li, R., Li, Y., et al. (2016). Identification of several circulating microRNAs from a genome-wide circulating microRNA expression profile as potential biomarkers for impaired glucose metabolism in polycystic ovarian syndrome. *Endocrine* 53 (1), 280–290. doi:10.1007/s12020-016-0878-9

Jiang, L., Huang, J., Li, L., Chen, Y., Chen, X., Zhao, X., et al. (2015). MicroRNA-93 promotes ovarian granulosa cells proliferation through targeting CDKN1A in polycystic ovarian syndrome. *J. Clin. Endocrinol. Metab.* 100 (5), E729–E738. doi:10.1210/jc.2014-3827

Kim, J., Lee, S. K., Kim, D., Choe, H., Jang, Y. J., Park, H. S., et al. (2020). Altered expression of adrenomedullin 2 and its receptor in the adipose tissue of obese patients. *J. Clin. Endocrinol. Metab.* 105 (1), dgz066. doi:10.1210/clinem/dgz066

Li, S. N., and Wu, J. F. (2020). TGF-β/SMAD signaling regulation of mesenchymal stem cells in adipocyte commitment. *Stem Cell Res. Ther.* 11 (1), 41. doi:10.1186/s13287-020-1552-y

Liu, Y., Tao, S., Liao, L., Li, Y., Li, H., Li, Z., et al. (2020). TRIM25 promotes the cell survival and growth of hepatocellular carcinoma through targeting Keap1-Nrf2 pathway. *Nat. Commun.* 11 (1), 348. doi:10.1038/s41467-019-14190-2

Mansuy-Aubert, V., Zhou, Q. L., Xie, X., Gong, Z., Huang, J. Y., Khan, A. R., et al. (2013). Imbalance between neutrophil elastase and its inhibitor α1-antitrypsin in obesity alters insulin sensitivity, inflammation, and energy expenditure. *Cell Metab.* 17 (4), 534–548. doi:10.1016/j.cmet.2013.03.005

Mu, L., Sun, X., Tu, M., and Zhang, D. (2021). Non-coding RNAs in polycystic ovary syndrome: A systematic review and meta-analysis. *Reprod. Biol. Endocrinol.* 19 (1), 10. doi:10.1186/s12958-020-00687-9

Murri, M., Insenser, M., Fernández-Durán, E., San-Millán, J. L., Luque-Ramírez, M., and Escobar-Morreale, H. F. (2018). Non-targeted profiling of circulating microRNAs in women with polycystic ovary syndrome (PCOS): Effects of obesity and sex hormones. *Metabolism* 86, 49–60. doi:10.1016/j.metabol.2018.01.011

Murri, M., Luque-Ramírez, M., Insenser, M., Ojeda-Ojeda, M., and Escobar-Morreale, H. F. (2013). Circulating markers of oxidative stress and polycystic ovary syndrome (PCOS): A systematic review and meta-analysis. *Hum. Reprod. Update* 19 (3), 268–288. doi:10.1093/humupd/dms059

Park, S. K., Oh, S. Y., Lee, M. Y., Yoon, S., Kim, K. S., and Kim, J. W. (2004). CCAAT/enhancer binding protein and nuclear factor-Y regulate adiponectin gene expression in adipose tissue. *Diabetes* 53 (11), 2757–2766. doi:10.2337/diabetes.53.11.2757

Patel, S. (2018). Polycystic ovary syndrome (PCOS), an inflammatory, systemic, lifestyle endocrinopathy. *J. Steroid Biochem. Mol. Biol.* 182, 27–36. doi:10.1016/j.jsbmb.2018.04.008

Robinson, S., Chan, S. P., Spacey, S., Anyaoku, V., Johnston, D. G., and Franks, S. (1992). Postprandial thermogenesis is reduced in polycystic ovary syndrome and is associated with increased insulin resistance. *Clin. Endocrinol. (Oxf)* 36 (6), 537–543. doi:10.1111/j.1365-2265.1992.tb02262.x

Rosenfield, R. L., and Ehrmann, D. A. (2016). The pathogenesis of polycystic ovary syndrome (PCOS): The hypothesis of PCOS as functional ovarian hyperandrogenism revisited. *Endocr. Rev.* 37 (5), 467–520. doi:10.1210/er.2015-1104

Ruebel, M. L., Cotter, M., Sims, C. R., Moutos, D. M., Badger, T. M., Cleves, M. A., et al. (2017). Obesity modulates inflammation and lipid metabolism oocyte gene expression: A single-cell transcriptome perspective. *J. Clin. Endocrinol. Metab.* 102 (6), 2029–2038. doi:10.1210/jc.2016-3524

Shorakae, S., Jona, E., de Courten, B., Lambert, G. W., Lambert, E. A., Phillips, S. E., et al. (2019). Brown adipose tissue thermogenesis in polycystic ovary syndrome. *Clin. Endocrinol. (Oxf)* 90 (3), 425–432. doi:10.1111/cen.13913

Tsurutani, Y., Fujimoto, M., Takemoto, M., Irisuna, H., Koshizaka, M., Onishi, S., et al. (2011). The roles of transforming growth factor-β and Smad3 signaling in adipocyte differentiation and obesity. *Biochem. Biophys. Res. Commun.* 407 (1), 68–73. doi:10.1016/j.bbrc.2011.02.106

Udesen, P. B., Glintborg, D., Sørensen, A. E., Svendsen, R., Nielsen, N. L. S., Wissing, M. L. M., et al. (2020). Metformin decreases miR-122, miR-223 and miR-29a in women with polycystic ovary syndrome. *Endocr. Connect.* 9 (11), 1075–1084. doi:10.1530/ec-20-0195

Viscarra, J. A., Wang, Y., Nguyen, H. P., Choi, Y. G., and Sul, H. S. (2020). Histone demethylase JMJD1C is phosphorylated by mTOR to activate de novo lipogenesis. *Nat. Commun.* 11 (1), 796. doi:10.1038/s41467-020-14617-1

Wang, D., Wang, W., Liang, Q., He, X., Xia, Y., Shen, S., et al. (2018). DHEA-induced ovarian hyperfibrosis is mediated by TGF-β signaling pathway. *J. Ovarian Res.* 11 (1), 6. doi:10.1186/s13048-017-0375-7

Wang, R., Hong, J., Cao, Y., Shi, J., Gu, W., Ning, G., et al. (2015). Elevated circulating microRNA-122 is associated with obesity and insulin resistance in young adults. *Eur. J. Endocrinol.* 172 (3), 291–300. doi:10.1530/eje-14-0867

Wrann, C. D., Eguchi, J., Bozec, A., Xu, Z., Mikkelsen, T., Gimble, J., et al. (2012). FOSL2 promotes leptin gene expression in human and mouse adipocytes. *J. Clin. Invest.* 122 (3), 1010–1021. doi:10.1172/jci58431

Wu, H., and Ballantyne, C. M. (2020). Metabolic inflammation and insulin resistance in obesity. *Circ. Res.* 126 (11), 1549–1564. doi:10.1161/circresaha.119.315896

Zhang, X., Zhang, B., Zhang, C., Sun, G., and Sun, X. (2021). Trib1 deficiency causes Brown adipose respiratory chain depletion and mitochondrial disorder. *Cell Death Dis.* 12 (12), 1098. doi:10.1038/s41419-021-04389-x

frontiers | Frontiers in Cell and Developmental Biology

TYPE Original Research
PUBLISHED 07 June 2023
DOI 10.3389/fcell.2023.1136096

OPEN ACCESS

EDITED BY
Xiao Wang,
Konge Larsen ApS, Denmark

REVIEWED BY
Qing Lin,
Johns Hopkins University, United States
Ebrahim Mohammadi,
Kurdistan University of Medical Sciences, Iran

*CORRESPONDENCE
Xiaoming Liu,
liuxiaoming26@163.com
Canxia Xu,
xucanxia2000@163.com

†These authors have contributed equally to this work

RECEIVED 02 January 2023
ACCEPTED 13 April 2023
PUBLISHED 07 June 2023

CITATION
Li H, Lin J, Cheng S, Chi J, Luo J, Tang Y, Zhao W, Shu Y, Liu X and Xu C (2023), Comprehensive analysis of differences in N6-methyladenosine RNA methylomes in *Helicobacter pylori* infection. *Front. Cell Dev. Biol.* 11:1136096. doi: 10.3389/fcell.2023.1136096

Comprehensive analysis of differences in N6-methyladenosine RNA methylomes in *Helicobacter pylori* infection

Huan Li[1†], Jiahui Lin[1†], Sha Cheng[1], Jingshu Chi[1], Ju Luo[1], Yu Tang[1], Wenfang Zhao[1], Yufeng Shu[1], Xiaoming Liu[1,2*] and Canxia Xu[1,2*]

[1]Department of Gastroenterology, The Third Xiangya Hospital of Central South University, Changsha, Hunan, China, [2]Hunan Key Laboratory of Non-Resolving Inflammation and Cancer, Central South University, Changsha, Hunan, China

Background: *Helicobacter pylori* (*H.pylori*) infection is an important factor in the occurrence of human gastric diseases, but its pathogenic mechanism is not clear. N6-methyladenosine (m6A) is the most prevalent reversible methylation modification in mammalian RNA and it plays a crucial role in controlling many biological processes. However, there are no studies reported that whether *H. pylori* infection impacts the m6A methylation of stomach. In this study, we measured the overall level changes of m6A methylation of RNA under *H. pylori* infection through *in vitro* and *in vivo* experiment.

Methods: The total quantity of m6A was quantified in gastric tissues of clinical patients and C57 mice with *H. pylori* infection, as well as acute infection model [*H. pylori* and GES-1 cells were cocultured for 48 h at a multiplicity of infection (MOI) from of 10:1 to 50:1]. Furthermore, we performed m6A methylation sequencing and RNA-sequencing on the cell model and RNA-sequencing on animal model.

Results: Quantitative detection of RNA methylation showed that *H. pylori* infection group had higher m6A modification level. M6A methylation sequencing identified 2,107 significantly changed m6A methylation peaks, including 1,565 upregulated peaks and 542 downregulated peaks. A total of 2,487 mRNA was upregulated and 1,029 mRNA was downregulated. According to the comprehensive analysis of MeRIP-seq and RNA-seq, we identified 200 hypermethylation and upregulation, 129 hypermethylation but downregulation, 19 hypomethylation and downregulation and 106 hypomethylation but upregulation genes. The GO and KEGG pathway analysis of these differential methylation and regulatory genes revealed a wide range of biological functions. Moreover, combining with mice RNA-seq results, qRT- PCR showed that m6A regulators, METTL3, WTAP, FTO and ALKBH5, has significant difference; Two key genes, PTPN14 and ADAMTS1, had significant difference by qRT- PCR.

Conclusion: These findings provide a basis for further investigation of the role of m6A methylation modification in *H. pylori*-associated gastritis.

KEYWORDS

gastritis, *Helicobacter pylori*, MeRIP-seq, M6A, N6-methyladenosine

Introduction

Helicobacter pylori (*H. pylori*), a Gram-negative microaerobic bacterium, is closely related to diseases such as gastritis, peptic ulcer and chronic gastritis (Cover and Blaser, 2009; Asano et al., 2016). It can initiate gastric carcinogenesis following the Correa cascade (Correa and Piazuelo, 2012). Once atrophy and intestinal metaplasia occur, there is still a lack of effective therapy to reverse the pathological changes, and some patients still progress to gastric cancer. Therefore, it is of great clinical significance to further explore the molecular mechanism of gastric diseases caused by *H. pylori* infection and find new intervention strategies and targets.

N6-Methyladenosine (m6A), involving methylation at the N6 position of RNA adenine, is the most prevalent RNA modification in eukaryotes (Meyer and Jaffrey, 2014; Huang et al., 2020a). In the 1970s, a study reported that there has m6A modification in mRNA and non-coding RNA of eucaryon (Desrosiers et al., 1974).

M6A is the most prevalent post-transcriptional modification of mRNAs and non-coding RNAs, which determines RNA fate, such as splicing, localization, stabilization, translation efficiency and nuclear export (Guo et al., 2021; Li et al., 2022a; Zhang et al., 2022). Recent years, more and more studies have reported that m6A plays different role during the growth and development of mammals, including embryonic development, circadian rhythm, neurogenesis, stress responses, sex determination and tumorigenesis (Pan et al., 2018; Chokkalla et al., 2020; Jiang et al., 2021; Xiao et al., 2022a). M6A modification mainly involves three enzymes, a family methyltransferase enzymes (writers), which including Methyltransferase Like 3 (METTL3), Methyltransferase Like 14 (METTL14), WT1 Associated Protein (WTAP) and et all, catalyze addition of m6A (Jiang et al., 2021) (Sun et al., 2022) (Sacco et al., 2022). The demethylase enzymes (erasers) that catalyze removal of m6A, such as alpha-ketoglutarate-dependent dioxygenase AlkB homolog 5 (ALKBH5) and fat mass and obesity-associated protein (FTO) (Roignant and Soller, 2017; Jiang et al., 2021). The m6A reader proteins can recognize the m6A-modified RNAs, which are divided into different protein families, such as IGF2 mRNA binding proteins (IGF2BP1/2/3) families, eukaryotic initiation factor (eIF) 3, the proteins contain the YT521-B homology (YTH) domain (YTHDF1/2/3 and YTHDC1/2) and et all (Jiang et al., 2021) (Shi et al., 2018) (Zhou et al., 2022). It is now clear that this reversible post-transcriptional modification is essential for gene regulation.

At present, research on stomach-related diseases m6A is mainly in gastric carcinoma and rarely in non-cancer disease. The role of m6A RNA modifications in diseases associated with *H. pylori* infection has not been investigated. In this study, we used high-throughput sequencing (MeRIP-seq) to identify the potential m6A modification of inflammation in gastric epithelial cells (GES-1) treated with *H. pylori*. Differential methylation genes (DMG), differential expression genes (DEG) and differential methylation and expression genes (DMEG) were analyzed by gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways to reveal the biological significance of genomes. In addition, combining with mice RNA-seq data, we used qRT-PCR tests to observe the expression of five common m6A regulatory and the three key gene, which were consistent in the sequencing results of cell model and animal model. These findings may provide new insights into the molecular mechanisms involved in *H. pylori* infection.

Materials and methods

Bacterial strains and cell lines

H. pylori was isolated from the gastric mucosa of gastric ulcer patient during gastroscopy as described (Xia et al., 2020). It was cultured in Columbia agar containing 10% sheep blood (Nanjing bianzhen Biological Technology Co., LTD., China) and antibiotics (5 mg/L cefsulodin, 5 mg/L amphotericin B, 5 mg/L trimethoprim, 10 mg/L vancomycin) (Oxoid, United Kingdom) at 37°C under microaerophilic conditions (5% O_2, 10% CO_2, and 85% N_2) for 3–5 days. When the value of OD600 was 1, the bacterial concentration was approximately 2×10^8 CFU/mL.

GES-1 cells were obtained from Hybribio Biotech Ltd. (Guangdong, China). The GES-1 cells were cultured in RPMI-1640 medium (Gibco, United States), containing 10% fetal bovine serum (Biological Industries, Israel) and maintained at 37°C in humidified 5% CO_2 incubator.

Clinical specimens

Four *H. pylori*-positive and four *H. pylori*-negative gastric tissues were collected from patients who underwent gastric biopsies at the Xiangya Third Hospital, Central South University (Changsha, China). The diagnoses were based on clinical and histological laboratory examination. All patients had signed informed consent for the study. The clinical information of patients was shown in Supplementary Table S1. This study was approved by the Ethics Committees of the Xiangya Third Hospital, Central South University.

H. pylori -infected animal model

Four to five weeks old male C57 BL/6 (18–22 g) were used. All the experimental animals were foster in the Department of Laboratory Animal Science of Central South University and were housed in an experimental animal room, which meets the specific pathogen-free (SPF)-class Meets the SPF standard, to ensure an environment with 12 h of light and 12 h of darkness. The eight mice were divided into two groups: control group (n = 4) and *H. pylori* infection group (n = 4). The mice were orally gavaged with 0.3 mL *H. pylori* suspension in phosphate buffered saline (PBS) (1×10^9) once daily for 9 days (repeat three times with 1 day off for three consecutive days) according to our previous study (Xia et al., 2020). The mice were only gavaged with sterile PBS in control group. The mice were sacrificed by cervical dislocation under CO_2 narcosis at 2 weeks after last gavage. Rapid urease test (RUT) and Giemsa staining were used to verify whether mice were infected with *H. pylori* (Supplementary Figure S1).

Cell infection model

GES-1 cells were seeded in 6-well plates until the density reached 60%–80% (~3 × 10^5) without *H. pylori* intervention and the cell culture medium containing no antibiotics. *H. pylori* was collected and re-suspended into antibiotic-free cell culture medium. The concentration of *H. pylori* suspension was adjusted to 1× 10^9/mL. Then, *H. pylori* suspension was added to GES-1 cells at a MOI of 10:1–50:1 and incubated for 48 h.

RNA extraction and qRT-PCR

Total RNA in tissues was extracted by the TRIzol reagent (Invitrogen, United States). Moreover, the extracted total RNA dissolved in RNase/DNase-free water. The ReverTra Ace qPCR RT Master Mix with gDNA Remover (Vazyme Biotech Co., Ltd., China) was used to reverse transcribe RNA in accordance with the manual. Primers for qRT-PCR were listed in Supplementary Table S2.

Quantification of the m6A modification

Total RNA was isolated as above. The quality of RNA was analyzed using a NanoDrop1000 (Thermo Fisher, United States). The EpiQuik m6A Methylation Quantification Kit (Epigentek, P-9005-96, United States) was used to measure the global m6A enrichment of mRNA. 200 ng RNA was coated in assay wells from each sample. The m6A levels are colorimetrical quantified at a wavelength of 450 nm absorbance.

RNA-seq and m6A-RNA immunoprecipitation sequencing

Total RNA was isolated from GES-1 cells and gastric tissue of mouse by TRIzol reagent as above. The Poly (A) RNA was purified from 50 μg total RNA using Dynabeads Oligo (dT) (Thermo Fisher, Carlsbad CA, United States) and two rounds of purification were used. Next, a Magnesium RNA Fragmentation Module was used to fragment the captured mRNA at 86°C for 7 min. Cleaved RNA fragments were incubated with m6A-specific antibody (Synaptic Systems GmbH, Goettingen, Germany) for 2 h at 4°C in IP buffer which was consist of 750 mM NaCl, 50 mM Tris-HCl and 0.5% Igepal CA-630. After performing IP, the IP product was synthesized into cDNA using reverse tran-scriptase (Invitrogen SuperScript™ II Reverse Transcriptase, CA, United States). *Escherichia coli* DNA polymerase I (NEB, United States), RNase H (NEB, United States), and dUTP Solution (Thermo Fisher, United States) which assisted the synthesis of the double-stranded DNA and the ends of the double-stranded DNA were repaired to form blunt ends. The two strands were digested with the enzyme UDG (NEB, United States) after adding an A base to both blunt ends and using magnetic beads to screen and purify the fragments according to size. Through PCR experiment, a library with a fragment size of 300 ± 50 bp was established (Supplementary Table S2). Finally, an Illumina NovaSeq™ 6000 (LC- Bio Technology Co., Ltd., Hangzhou, China) was used to sequencing with PE150 (2 bp × 150 bp paired-end) sequencing mode.

Bioinformatics analysis

Fastp (https://github.com/OpenGene/fastp) was used for quality control on the original data and acquire clean data. HISAT2 package (http://daehwankim lab.github.io/hisat2) was used to compare the acquired clean data to the genome (human genome, version: hg19; and *mus musculus* genome, version: GRCm38). The R package exome-Peak (https://bioconductor.org/packages/exome Peak) was used to perform peak calling analysis and peak analysis of genetic difference. The IGV software (http://www.igv.org) visualized the results. HOMER (http://homer.ucsd.edu/homer/motif) and MEME2 (http://meme-suite.org) were used for motif analysis. StringTie (https://ccb.jhu.edu/softw are/stringtie) was used to determine the expression levels of all mRNAs in the input libraries. The different expression of mRNAs was selected according to thresholds of a p-value < 0.05 and a |log2 (fold change)| >1 with the R package edgeR (https://bioconduct or.org/packages/edgeR).

Statistical analysis

SPSS 22.0 and GraphPad Prism 7.0 were used for data processing. The t-test and $\chi 2$ test were used to analyze the differences among different samples. A p-value less than 0.05 was considered to indicate statistical significance (*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, ****$p < 0.0001$).

Results

Establishment of *H. pylori* infection model *in vivo* and *in vitro*

In this study, GES-1 cells were treated with *H. pylori* with a MOI of 10:1 for 48 h. We detected the mRNA expression of VEGF, IL-6 and IL-8 by qRT-PCR (Figure 1A). The results show that the expression levels of these proinflammatory factors were significantly increased in the GES-1 cells treated with *H. pylori* ($p < 0.05$). Next, we observed significant increase in the overall level of m6A methylation in *H. pylori*-infected patients and mice (Figures 1B, C), and mild significant increase in *H. pylori*-infected cells (Figure 1D).

Overview of methylation RNA immunoprecipitation sequencing

In the MeRIP-seq library, the two sets of samples obtained an average of 41,015,657 and 44,650,725 valid reads, while in the RNA-seq library the two groups obtained an average of 37,346,096 and 40,732,363 valid reads (Supplementary Table S3). Among the IP samples, the average matching rate of valid reads in the control group and *H. pylori* group was 97.2% and 97.1%, respectively. The mean matching rates for valid reads in the input samples were 97.5% and 97.7% (Supplementary Table S4). Clean reads that can be matched to the reference genome are defined as exons, introns

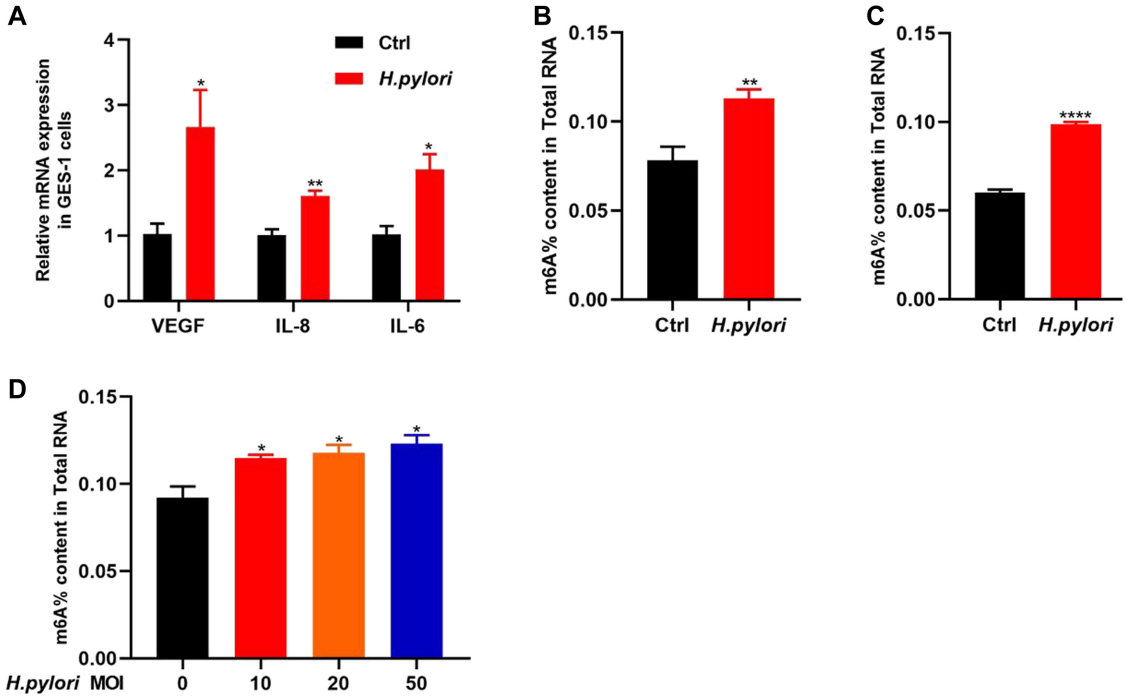


FIGURE 1
Establishment of *H. pylori* infection model and determination of total m6A **(A)** The RNA-level expression of inflammatory factors IL-8, IL-6, and VEGF after *H. pylori* infection of GES-1 cell. **(B)** The total m6A content in *H. pylori* negative and positive patients ($n = 4$). **(C)** The total m6A content in *H. pylori* negative and positive animal ($n = 4$). **(D)** The total m6A content in *H. pylori*-uninfected and *H. pylori*-infected cells in different MOI ($n = 3$). *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$; ****$p < 0.0001$.

and intergenic sequences according to the regional information of the reference genome. The mean rates of IP and exons in the input samples were 97.39% and 97.44% for the control group and 96.1% and 96.86% for the *H. pylori*-infected group, respectively (Supplementary Figure S2).

Profile of the m6A modification in GES-1 cells treated with *H. pylori*

To obtain a map of m6A modifications in gastric epithelial cells infected with *H. pylori*, we used meRIP-seq to performed a transcriptome analysis of m6A modification. Combining all the peak reads, we found that the enrichment of reads was located near the transcription start site (TSS) and the transcription end site (TES) (Figure 2A). To further understand the distribution of the differential peak on the functional elements of the gene, we divided it into three regions: the 5′ untranslated region (5′ UTR), the first exon, the other exons and the 3′ UTR (Figure 2B). Meanwhile, we analyzed the distribution pattern of differential m6A methylation peaks. A total of 24.46% of the m6A methylation peaks were contained in the 5′UTR, 45.5% were enriched in the 3′ UTR, while 11.38% were enriched in the exons (Figure 2C). Under the screening conditions of |log2 (fold change)| > 1 and *p*-value < 0.05, a total of 9,097 peaks were identified in both groups. The results showed 2,107 significantly different peaks compared to the control group, of which 1,565 peaks were upregulated and 542 peaks were downregulated (Figure 2D). The top 20 distinct m6A methylation peaks are shown in Supplementary Table S5.

Differential m6A modification is involved in important biological pathways

To explore the important functions of m6A modification in *H. pylori*-induced gastric epithelial cells, GO and KEGG enrichment analyses were performed for the above m6A differential peak (DMG) genes. The GO results were classified into three categories: cellular component (CC) and biological process (BP) and molecular function (MF) categories. It can be observed that both hypermethylated and hypomethylated genes are associated with "regulation of transcription, DNA template," "signal transduction," "apoptotic process," " regulation of transcription by RNA polymerase II," "cell cycle" and "RNA splicing" (ontology: biological processes); "nucleus," " membrane," "cytoplasm" and "cytoplasm" (ontology: cellular components); and "protein binding," "RNA binding," "metal ion binding" (ontology: molecular function) (Figures 3A, C). In addition, the results of the KEGG signaling pathway analysis showed that the genes upregulated by the m6A peak were mainly enriched in "fatty acid elongation," "primary immunodeficiency," "Epstein-Barr virus infection," "drug

10.3389/fcell.2023.1136096

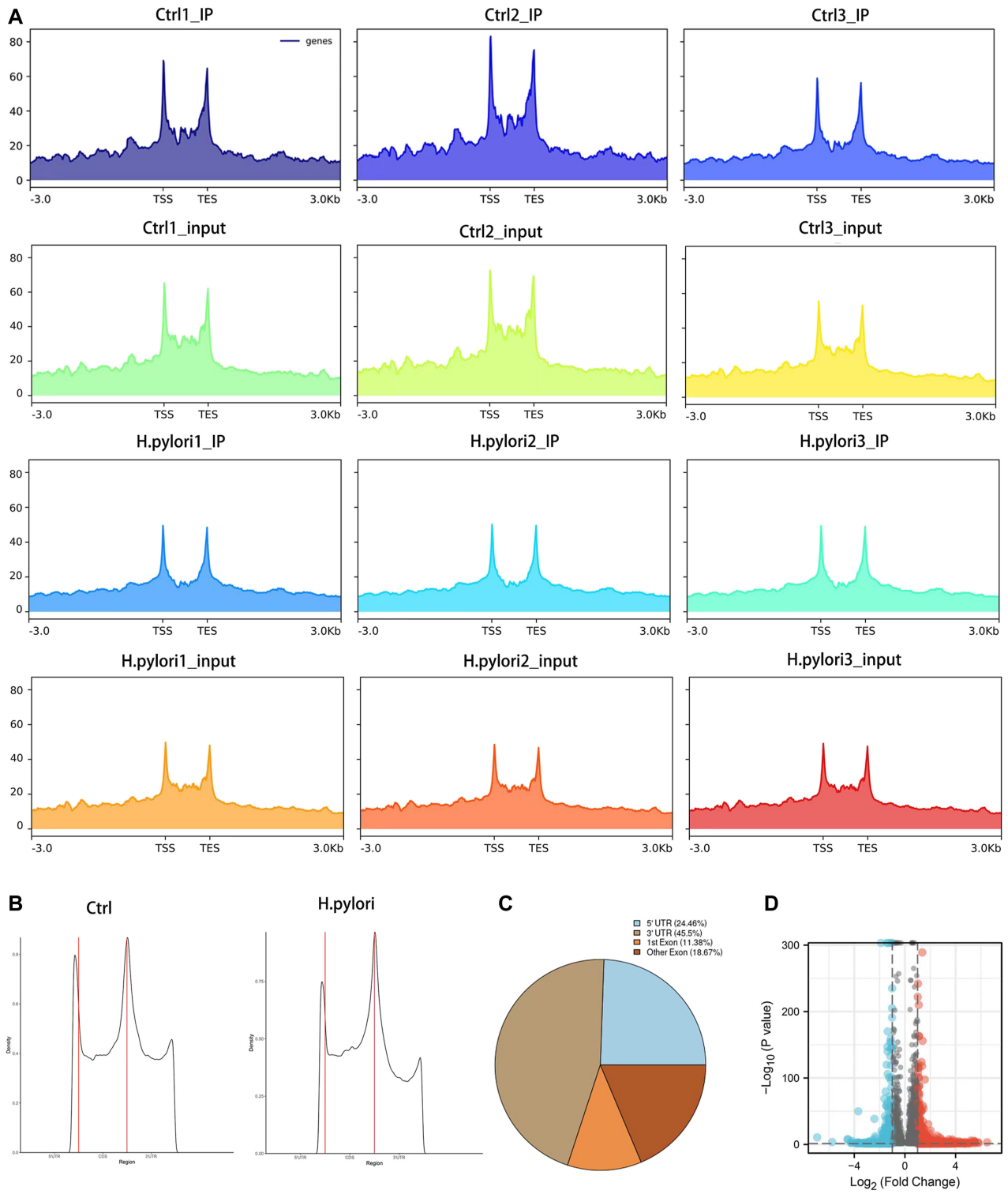


FIGURE 2
Significant dysregulation of the m6A peak during *H. pylori* infection of gastric epithelial cells. **(A)** Enrichment of peaks near the gene transcription start site. **(B,C)** Distribution of differentially methylated m6A peaks in control and *H. pylori* groups **(D)** Volcano plot of genes with differential m6A peaks (|log2(FC)|>1 and *p*-value < 0.05).

metabolism-other enzymes," "NF-κB signaling pathway," and "basic transcription factors" (Figure 3B); The genes downregulated by the m6A peak were mainly concentrated in "natural killer cell mediated cytotoxicity," "basal transcription factors," "cell cycle," "mRNA surveillance pathway," and "pyruvate metabolism" (Figure 3D).

10.3389/fcell.2023.1136096

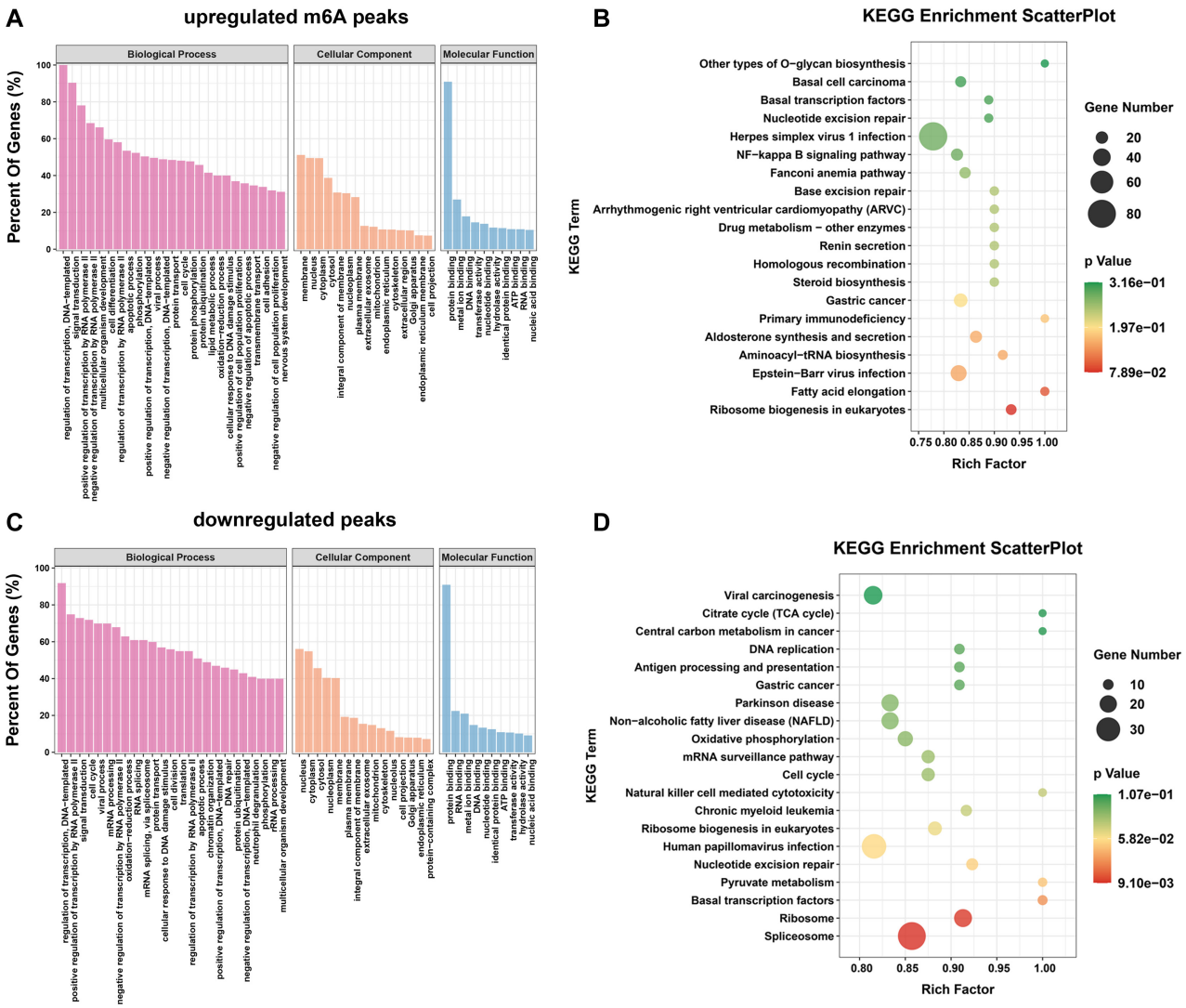


FIGURE 3
Differential m6A modifications are involved in important biological pathways. **(A,B)** GO enrichment and KEGG pathway analysis of the hypermethylated peaks. **(C,D)** GO enrichment and KEGG pathway analysis of the hypomethylated peaks.

Analysis of RNA-seq differential expression genes

To explore the potential relationship between m6A modifications and gene expression, differential expression gene analysis was performed using input sequencing data. By hierarchical clustering of RNA-seq data, we detected significantly different expression between the control and *H. pylori* groups (Figure 4A). We then screened the RNA-seq database for a total of 3,516 differential genes ($|\log2(FC)| > 1$ and p-value < 0.05) compared to control samples. Among them, 2,487 upregulated genes and 1,029 downregulated genes were identified (Figure 4B). These differential expression genes were then used for GO enrichment and KEGG pathway analysis. GO enrichment results showed these genes were significantly related to "translation initiation," "SRP-dependent cotranslation protein targeting membranes," "viral transcription" and "mRNA splicing, via spliceosomes" (Figure 4C). KEGG analysis showed that these genes were mainly enriched in "Notch signaling pathway," "adherens junctions," "Hippo signaling pathway," "protein processing in endoplasmic reticulum," "ubiquitin mediated proteolysis" and "oxidative phosphorylation" (Figure 4D).

Combined analysis between m6A-seq and RNA-seq

To further explore the functional significance of m6A modifications in *H. pylori*-infected gastric epithelial cells, we investigated whether m6A methylation underlies the observed differences in expression. For this purpose, DMGs and DEGs were detected using m6A-seq data and RNA-seq data. Thereafter, a combination of m6A-seq and RNA-seq analysis classified all genes into four major groups: including 200 hypermethylated and upregulated (hypo-up), 129 hypermethylated but downregulated

10.3389/fcell.2023.1136096

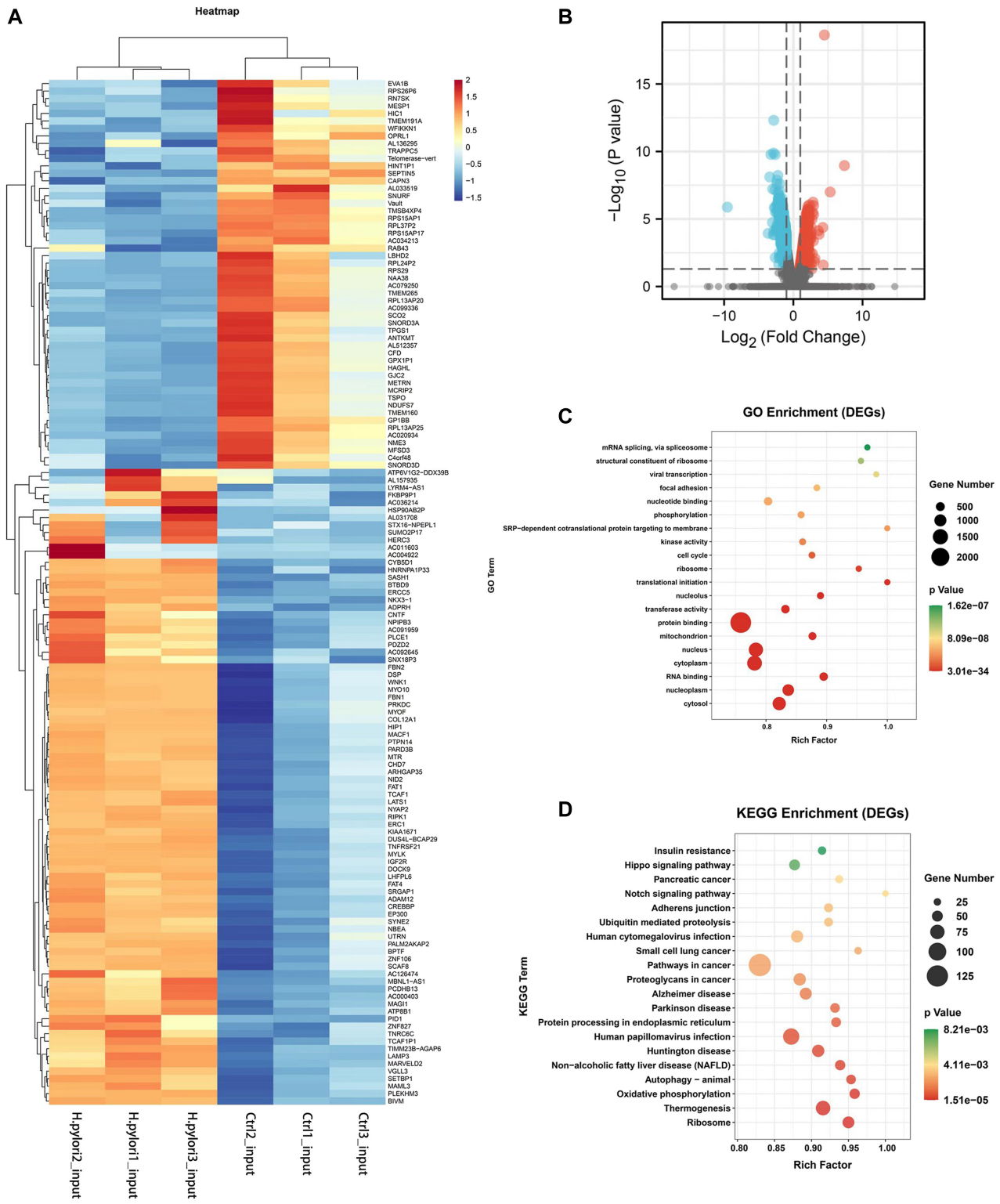


FIGURE 4
Differentially expressed gene analysis by RNA-Seq **(A)** Heat map showing differentially expressed mRNAs in three *H. pylori* input samples and three control input samples. **(B)** Volcano plot showing the differentially expressed mRNAs between *H. pylori* and control groups with statistical significance (fold change ≥ 2.0 and $p < 0.05$). **(C)** GO enrichment analysis of differential genes. **(D)** KEGG pathway analysis of differential genes.

(hypo-down), 19 hypermethylated and downregulated (hypo-down) and 106 hypomethylated but upregulated genes or transcripts (hypo-up) (Figure 5A). Four groups of DMEG were further investigated by KEGG analysis. The results showed that hyper-up genes were mainly enriched in "Adherens junctions," "FoxO signaling pathway" and "Fatty acid degradation" pathways

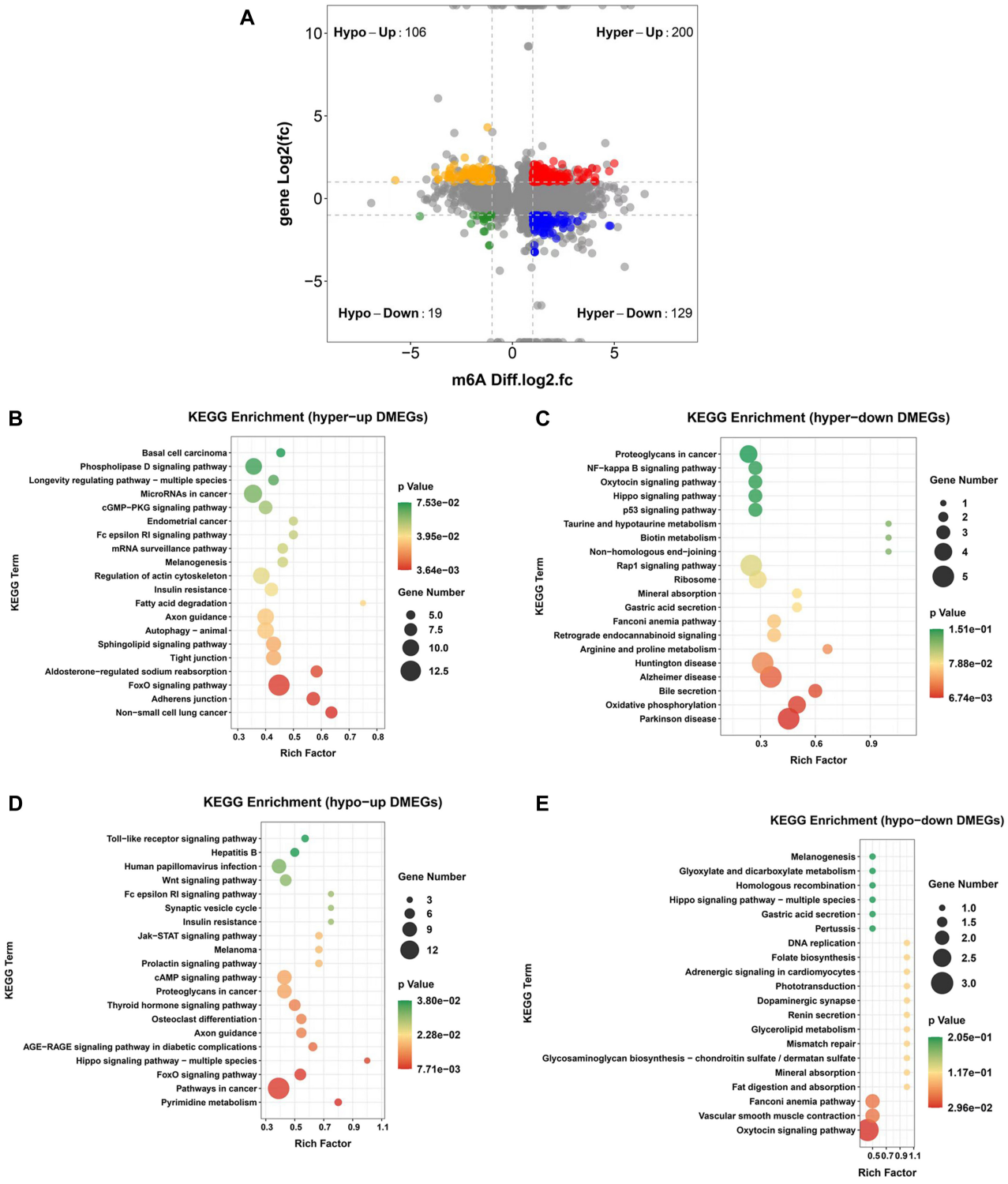


FIGURE 5
Combined analysis between m6A-Seq and RNA-Seq **(A)** Four-quadrant diagram of hyper-up, hyper-down, hypo-up, and hypo-down DMEG. **(B)** KEGG pathway analysis of hyper-up DMEG **(C)** KEGG pathway analysis of hyper-down DMEG **(D)** KEGG pathway analysis of hypo-up DMEG **(E)** KEGG pathway analysis of hypo-down DMEG.

(Figure 5B); in contrast, hyper-down genes were mainly enriched in "Bile secretion," "Gastric acid secretion," "Oxidative phosphorylation" and "NF-κappa B signaling pathway" (Figure 5C). In addition, hypo-up genes were mainly enriched in "Toll-like receptor signaling pathway," "Wnt signaling pathway," "Jak-STAT signaling pathway," "cAMP signaling pathway,"

10.3389/fcell.2023.1136096

TABLE 1 Top 20 transcripts of differential m6A modification and mRNA expression between control group and *H. pylori* group.

Gene name	Change	Seqnames	m6A modification change						mRNA expression change	
			Peak start	Peak end	Width	Peak region	logFC	*p*-value	logFC	*p*-value
PTPN14	Hyper-up	chr1	214,532,254	214,532,432	179	Exon	5	4.57E-02	2.13	1.75E-06
FRY	Hyper-up	chr13	32,298,645	32,298,945	301	Exon	4.75	1.74E-04	1.65	9.74E-04
TRANK1	Hyper-up	chr3	36,857,290	36,857,679	390	Exon	4.14	9.55E-04	1.79	4.63E-03
HSPA12A	Hyper-up	chr10	116,827,654	116,827,984	331	3′ UTR	4.07	1.70E-02	1.02	4.62E-02
ADAM10	Hyper-up	chr15	58,748,838	58,749,018	181	3′ UTR	4.06	5.75E-03	1.29	1.96E-02
BOLA2	Hyper-down	chr16	29,454,886	29,455,005	120	5′ UTR	4.81	3.72E-03	−1.65	5.48E-06
AL136038	Hyper-down	chr14	63,642,540	63,642,600	61	Exon	4.75	4.07E-02	−1.64	3.21E-04
QPCTL	Hyper-down	chr19	45,703,571	45,703,690	120	3′ UTR	3.44	1.74E-03	−1.07	2.95E-04
PIDD1	Hyper-down	chr11	804,682	805,011	330	5′ UTR	3.20	3.24E-03	−1.38	9.14E-05
RPS9	Hyper-down	chr19	54,224,672	54,224,881	210	Exon	2.91	1.17E-02	−1.28	7.32E-05
NSF	Hypo-up	chr17	46,640,090	46,643,106	3,017	3′ UTR	−5.75	2.34E-04	1.10	2.92E-02
ADAMTS1	Hypo-up	chr21	26,843,766	26,844,363	598	5′ UTR	−3.78	2.00E-03	1.57	2.25E-03
MAPRE2	Hypo-up	chr18	34,977,018	34,978,391	1,374	5′ UTR	−3.65	8.91E-03	1.15	1.58E-02
CEP78	Hypo-up	chr9	78,276,633	78,276,783	151	3′ UTR	−3.63	4.37E-03	1.17	1.91E-02
LAMA3	Hypo-up	chr18	23,899,378	23,899,528	151	Exon	−3.31	1.17E-02	1.42	4.32E-03
SLX1B	Hypo- down	chr16	29,457,636	29,458,189	554	3′ UTR	−4.54	1.95E-08	−1.07	2.50E-04
RPS3AP47	Hypo- down	chr15	43,115,761	43,115,907	147	Exon	−2.03	2.88E-03	−1.52	5.21E-05
RPL29P11	Hypo- down	chr3	37,016,898	37,017,014	117	Exon	−1.91	8.71E-06	−1.01	5.54E-04
WTIP	Hypo- down	chr19	34,504,119	34,504,179	61	3′ UTR	−1.73	4.47E-02	−1.04	1.35E-03
PPP1R12C	Hypo- down	chr19	55,092,432	55,093,073	642	Exon	−1.47	5.01E-13	−1.06	2.37E-04

"pathway in cancer" and "Hippo signaling pathway-multi-species" (Figure 5D), while hypo-down genes were mainly enriched in "Vascular smooth muscle contraction" and "DNA replication" (Figure 5E). Moreover, we list the top 20 transcripts of differential m6A modification and mRNA expression between control group and *H. pylori* group based on diff.log2.fc (Table 1).

PPI network and hub genes were identified in DMEG

The PPI network of DMEG is carried out by the STRING database (Figure 6A) and Cytoscape. As above, the network was divided into four clusters, which respectively are hyper-up, hyper-down, hypo-down and hypo-up DMEGs (Figure 6B). GO enrichment analysis was performed for each DMEG cluster to elucidate its biological functions (Figures 6C–F). PPI network interaction data are listed in Supplementary Table S6.

Validation of differential expression genes

In the RNA-seq data of mice, we analyzed the mRNA levels of 28 m6A regulators, except IGF2BP1, 27 of 28 m6A regulators showed an increasing tendency (Figure 7A). Furthermore, qRT-PCR was used to detect the expression levels of five common regulators, including METTL3, METTL14, WTAP, FTO and ALKBH5; The change trends in those genes revealed by qRT-PCR were consistent with the with the RNA-seq results, those gene all did arrive significant difference except METTL14 (Figure 7E).

We used sequencing of mice samples results to validate the expression of top 20 genes of cell sequencing (Table 1), which shows that only 3 genes were consistent, including PTPN14, BOLA2 and ADAMTS1 (Figures 7B, C). We performed IGV visualization for the three genes and all found significantly different m6A levels (Figure 7D). Moreover, except BOLA2, PTPN14 and ADAMTS1 had significant difference by qRT- PCR (Figure 7F).

10.3389/fcell.2023.1136096

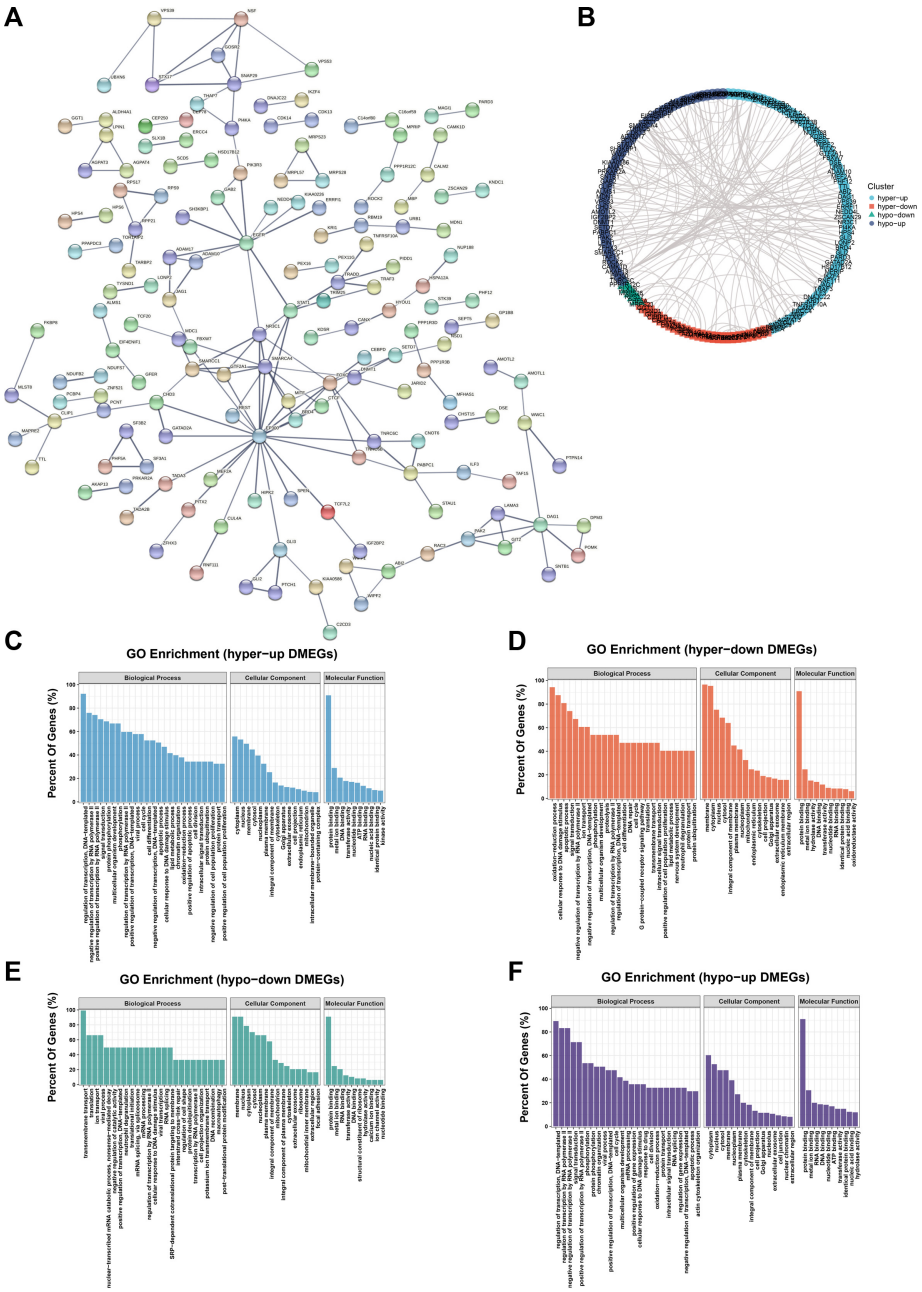


FIGURE 6
PPI networks and hub genes were found in DMEG. **(A)** PPI network of DMEG constructed from STRING database **(B)** Cytoscape was performed divided into four clusters. Blue represents hyper-up genes, orange represents hyper-down genes, green represents hypo-down genes, and purple represents hypo-up genes. **(C)** GO enrichment analysis of the hyper-up cluster in this DMEG. **(D)** GO enrichment analysis of hyper-down clusters in this DMEG. **(E)** GO enrichment analysis of the hypo-down cluster in this DMEG. **(F)** GO enrichment analysis of the hypo-up cluster in this DMEG.

Discussion

H. pylori infection can damage the stomach mucosa, leading to the development of various stomach diseases which involve many pathophysiological changes. Abnormalities in m6A modifying enzymes can cause a series of diseases (Jiang et al., 2021) (Zhang et al., 2019). However, the mechanism of m6A modification in *H. pylori*-induced gastric epithelial infection remains unclear. In this study, the relationship between m6A modification profile and *H. pylori*-induced gastric epithelial infection was analyzed for the first time.

In the beginning, we found that *H. pylori* infection increase the level of m6A modification *in vitro* and *in vivo*. After that, we obtained an overview of m6A modification in *H. pylori* infection associated gastritis through MeRIP-seq. The total peak numbers of m6A revealed significant differences in m6A modification between the control and *H. pylori* groups. Therefore, we assume that m6A modification may be related to *H. pylori* -induced gastritis.

10.3389/fcell.2023.1136096

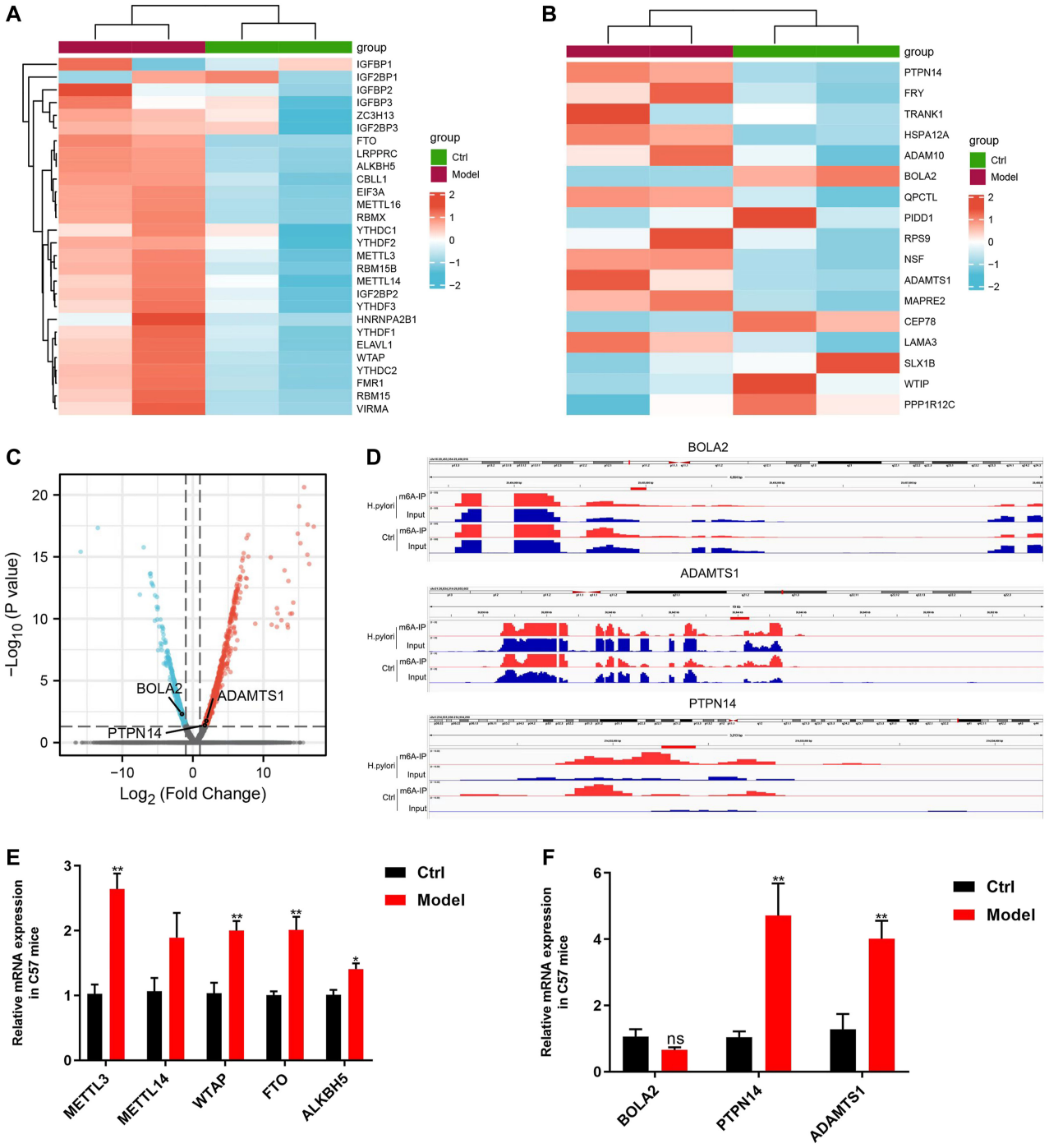


FIGURE 7
Validation of differential expression genes **(A)** Heatmap of m6A regulators in sequencing of mice samples. **(B)** RNA-seq data of mice to validate the top 20 genes from cell sequencing by Heatmap. **(C)** Volcanic map of differentially expressed genes sequencing of mice samples. **(D)** IGV visualization show the three key m6A-modified genes. **(E)** qRT- PCR results for five common m6A regulators. **(F)** qRT- PCR results of the three key m6A-modified genes.

As is known to all, m6A modification of mRNA often affects the occurrence and development of the disease. In this study, we identified 2,107 significantly different peaks compared to the control, of which 1,565 peaks were upregulated and 542 peaks were downregulated. From this, we found that *H. pylori* can alter the methylation peak of GES-1. Therefore, we hypothesized that m6A modification may be associated with *H. pylori*-induced gastric epithelial infection. M6A peaks were mostly enriched near the 3′ UTR region, and these sites were m6A specific and consistent with previous studies (Huang et al., 2020a). The 3′ UTR regulates mRNA stability, localization, expression and translation of mRNA. Multiple RNA-binding proteins bind in this region to perform regulatory functions and regulate the interaction

between proteins (Mayr, 2019). In addition, in *H. pylori*-infected gastric epithelial cells, differential methylation peaks were significantly enriched in "transcriptional regulation," "RNA splicing," "signal transduction," "apoptotic processes" and "cell cycle." Previous studies indicated that *H. pylori* involve the regulation of apoptosis, proliferation and the cell cycle (Hirata et al., 2001; Nozawa et al., 2002; Ding et al., 2008). This suggests a conserved and fundamental role of m6A in the regulation of development and cell fate specification. The hypermethylation peaks were mainly concentrated in "fatty acid elongation," "EBV infection," "drug metabolism," "NF-κB signaling pathway" "EBV infection," "drug metabolism," "NF-κB signaling pathway" and "basic transcription factor" pathways. NF-κB is a key regulator of the immune response against *H. pylori* infection and is known to modulate genes involved in the control of inflammation, cell proliferation and apoptosis (Lamb and Chen, 2010; Chaturvedi et al., 2011; Shu et al., 2022). The hypomethylation peaks were mainly enriched in "natural killer cell-mediated cytotoxicity," "basic transcription factors," "cell cycle," "mRNA surveillance pathway," "basic transcription factors." The mRNA surveillance pathway" and "pyruvate metabolism" pathways. This evidence indicates that m6A modification is probably associated with *H. pylori* -associated gastritis.

In order to clarify the mechanism of m6A affecting the process of *H. pylori* infection in gastric epithelial cells, we combined m6A methylation group with transcription group to find the key signaling pathways affected by m6A modification. Previous studies have shown that during the time course of *H. pylori* infection, *H. pylori* infection destroys the integrity of the gastric mucosa. *H. pylori* induces classical and alternative NF-κB signaling pathways through its effector ADP-L-glycero-β-D-manno-heptose (ADP-heptose), leading to deleterious gastric pathophysiology (Maubach et al., 2022). It has also been shown that *H. pylori* can induce a signaling cascade by activating the Toll-like receptor pathway, which ultimately leads to the transcription of pro-and anti-inflammatory cytokines and type I interferons (Peek et al., 2010). The Hippo signaling pathway appears to be a protective pathway in the host-pathogen conflict that generates an inflammatory environment, cellular injury, and epithelial renewal and differentiation, limiting the loss of gastric epithelial properties prior to adenocarcinoma development, which may be beneficial for *H. pylori* colonization and chronic infection (Molina-Castro et al., 2020). As in previous studies, some classical pathways regarding *H. pylori* causing gastric disease were significantly enriched in the present study. These include NF-κB signaling pathway (Keates et al., 1997; Sasaran et al., 2021), p53 signaling pathway (Cai et al., 2021; Imai et al., 2021), Hippo signaling pathway, Toll-like receptor signaling pathway (Lam et al., 2022), and Wnt signaling pathway (Abdi et al., 2021). This evidence also suggests that m6A modification may be associated with *H. pylori*-induced gastric disease.

In the past few years, numerous studies have illustrated the biological effects of m6A modification on RNA. On the one hand, the m6A methylation process is reversible, and this mark on RNA can be written or erased under various stimuli and biological factors (Feng et al., 2022; Liu et al., 2022; Wang et al., 2022). On the other hand, m6A can affect RNA processing and metabolism through a variety of mechanisms, including selective polyadenylation, selective splicing, RNA stability, RNA export, RNA degradation, and translation (Wang et al., 2014; Zhao et al., 2014; Coots et al., 2017; Hong et al., 2022). Thus, m6A up- or downregulates gene expression in a complex and context-dependent manner. For this reason, we observed four groups of DMEGs in the present study, which are hyper-up, hyper-down, hypo-up, and hypo-down. Our functional enrichment analysis showed that these four groups of DMEGs are associated with essential and different biological processes. Many previous studies reported that m6A modification is involved in different biological processes, such as transcriptional regulation, signal transduction, and the DNA damage response (Jia et al., 2011; Zheng et al., 2013; Hong et al., 2022). Our results are corresponded to previously these published studies.

In the RNA-seq data of mice, we found numerous m6A regulators were found to have changes. Except IGF2BP1, other m6A regulators showed an increasing tendency. To further verify the results of sequencing, we observed the expression of five common regulators by qRT- PCR which showed similar results to sequencing. Those genes (METTL3, WTAP, FTO and ALKBH5) all did arrive significant difference except METTL14. METTL3, as one of the core components of the m6A methyltransferase complex, has been found to be closely related to multiple signaling pathways, such as the JAK/STAT (Yao et al., 2019), MAPK/NF-κB (Li et al., 2020a), PI3K/AKT (Bi et al., 2021), and Wnt/β-catenin pathways (Cui et al., 2020). WTAP has been reported to be associated with a number of signaling pathways, such as TGFβ (Li et al., 2020b), hippo (Hu et al., 2020), NF-κB (Li et al., 2021), and Hedgehog pathways (Wei et al., 2022). FTO is associated with various signaling pathways, for example, PKA/CREB (Hu et al., 2022), TNF-α (Li et al., 2022b), ERK (Xiao et al., 2022b), WNT (Kim et al., 2022) and JAK2/STAT3 pathways (Shen et al., 2021). ALKBH5 is involved in many signaling pathways, including WNT (Lin et al., 2022), PTEN/AKT (He et al., 2021), NF-κB (Qu et al., 2022), AKT (Wang et al., 2020). Interestingly, many studies have shown that *H. pylori* infection is closely related to these signaling pathways. *H. pylori* can active the expression of STAT1 and PD-L1 which may prevent immune surveillance in the gastric mucosa, allowing premalignant lesions to progress to gastric cancer (Li et al., 2022c). *H. pylori* can induce injuries to the stomach through MAPK/NF-κB pathway (Shu et al., 2022). *H. pylori* can induce the occurrence of gastric carcinogenesis at the early stage by activating the PI3K/Akt signaling pathway (Xu et al., 2018). *H. pylori* infection activated WNT/β-catenin signaling pathway by upregulating to induce gastritis (Zuo et al., 2022). Judging from these, m6A regulators may also involve in regulating different signaling pathways in *H. pylori*-associated gastritis.

Moreover, the results of sequencing of mice samples were used to validate the expression of top 20 genes of cell sequencing, which found three genes, PTPN14, BOLA2 and ADAMTS1, that are consistent. Furthermore, qRT- PCR showed PTPN14 and ADAMTS1 had significant difference. Although there was no significant difference in BOLA2 expression, there was a downward trend. *H. pylori* was able to significantly upregulate PTPN14 and ADAMTS1 mRNA expression levels. At present, there are no studies relationship between these three genes and *H. pylori*-infected diseases. Many studies showed that PTPN14 has different function, such as suppressing the occurrence and development of tumor (Hatterschide et al., 2022), blunting the formation of atherosclerosis (Yang et al., 2021a) and promoting inflammation and fibrosis (Fu et al., 2020; Lin et al., 2021). Previous

studies confirmed that ADAMTS1 is involved in inhibiting the proliferation, polarization and migration of tumor (Li et al., 2015; de Assis Lima et al., 2021), affecting the quality of oocytes and embryonic development potential (Yang et al., 2021b) and promoting collagen production (Toba et al., 2016). In addition, a recent study revealed that YTHDF2 inhibited ADAMTS1 expression and promoted sperm adhesion through m6A/mRNA pathway (Huang et al., 2020b). Above all, these results suggest that m6A is likely to exhibit an as-yet- unknown function in the process by *H. pylori* -induced gastritis.

In summary, we can infer that m6A methylation was shown to play a role in *H. pylori*-induced gastritis. Though the mechanism of m6A-regulated gastritis is not clearly understood, we provide the first m6A transcriptome profile of gastritis and an initial map revealing the function of m6A modification in gastritis using advanced technologies, thereby contributing critical insights for further research on the role of m6A in gastritis. These findings provide a basis for further investigation of the role of m6A methylation modification in *H. pylori* infection of the gastric mucosa. However, the finding into clinical scenario may be limited by the lack of verification of the expression and distribution of m6A related regulatory molecules in clinical samples of *H. pylori*-associated gastritis, which should be further studied in future investigations.

Data availability statement

The data presented in the study are deposited in the GEO repository, accession number GSE230869 and GSE231337.

Ethics statement

The studies involving human participants were reviewed and approved by the Ethics Committee of the Third Xiangya Hospital of Central South University. The patients/participants provided their written informed consent to participate in this study. The animal study was reviewed and approved by the Ethics Committee of the Third Xiangya Hospital of Central South University.

Author contributions

All of the authors contributed to the conception of the article. The main experimental conception and design: CX and XL; performed the experiments: HL and JiL; analyzed the data and contributed reagents: SC, JC, and JuL; collected samples: YT, WZ, and YS; writing the manuscript: HL and JiL. All authors read and approved the final manuscript.

Funding

This study was supported by Changsha Science and Technology Project (kq2202118), Natural Science Foundation of Hunan Province (2022JJ30906) and Hunan Provincial Innovation Foundation For Postgraduate (QL20220061).

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

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcell.2023.1136096/full#supplementary-material

References

Abdi, E., Latifi-Navid, S., Abedi Sarvestani, F., and Esmailnejad, M. H. (2021). Emerging therapeutic targets for gastric cancer from a host-Helicobacter pylori interaction perspective. *Expert Opin. Ther. Targets* 25 (8), 685–699. [published Online First: 2021/08/20]. doi:10.1080/14728222.2021.1971195

Asano, N., Imatani, A., Watanabe, T., Fushiya, J., Kondo, Y., Jin, X., et al. (2016). Cdx2 expression and intestinal metaplasia induced by *H. pylori* infection of gastric cells is regulated by NOD1-mediated innate immune responses. *Cancer Res.* 76 (5), 1135–1145. [published Online First: 2016/01/14]. doi:10.1158/0008-5472.CAN-15-2272

Bi, X., Lv, X., Liu, D., Guo, H., Yao, G., Wang, L., et al. (2021). METTL3-mediated maturation of miR-126-5p promotes ovarian cancer progression via PTEN-mediated PI3K/Akt/mTOR pathway. *Cancer Gene Ther.* 28 (3-4), 335–349. [published Online First: 2020/09/18]. doi:10.1038/s41417-020-00222-3

Cai, Q., Shi, P., Yuan, Y., Peng, J., Ou, X., Zhou, W., et al. (2021). Inflammation-associated senescence promotes Helicobacter pylori-induced atrophic gastritis. *Cell. Mol. Gastroenterol. Hepatol.* 11 (3), 857–880. [published Online First: 2020/11/09]. doi:10.1016/j.jcmgh.2020.10.015

Chaturvedi, M. M., Sung, B., Yadav, V. R., Kannappan, R., and Aggarwal, B. B. (2011). NF-κB addiction and its role in cancer: 'one size does not fit all'. *Oncogene* 30 (14), 1615–1630. [published Online First: 2010/12/21]. doi:10.1038/onc.2010.566

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 (12), 2331–2349. [published Online First: 2020/09/25]. doi:10.1177/0271678X20960033

Coots, R. A., Liu, X. M., Mao, Y., Dong, L., Zhou, J., Wan, J., et al. (2017). m6A Facilitates eIF4F-Independent mRNA Translation. *Mol. Cell.* 68 (3), 504–514. [published Online First: 2017/11/07]. doi:10.1016/j.molcel.2017.10.002

Correa, P., and Piazuelo, M. B. (2012). The gastric precancerous cascade. *J. Dig. Dis.* 13 (1), 2–9. [published Online First: 2011/12/23]. doi:10.1111/j.1751-2980.2011.00550.x

Cover, T. L., and Blaser, M. J. (2009). *Helicobacter pylori* in health and disease. *Gastroenterology* 136 (6), 1863–1873. [published Online First: 2009/05/22]. doi:10.1053/j.gastro.2009.01.073

Cui, X., Wang, Z., Li, J., Zhu, J., Ren, Z., Zhang, D., et al. (2020). Cross talk between RNA N6-methyladenosine methyltransferase-like 3 and miR-186 regulates hepatoblastoma progression through Wnt/β-catenin signalling pathway. *Cell. Prolif.* 53 (3), e12768. [published Online First: 2020/01/23]. doi:10.1111/cpr.12768

de Assis Lima, M., da Silva, S. V., Serrano-Garrido, O., Hülsemann, M., Santos-Neres, L., Rodríguez-Manzaneque, J. C., et al. (2021). Metalloprotease ADAMTS-1 decreases

cell migration and invasion modulating the spatiotemporal dynamics of Cdc42 activity. *Cell. Signal* 77, 109827. [published Online First: 2020/11/09]. doi:10.1016/j.cellsig.2020.109827

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 (10), 3971–3975. [published Online First: 1974/10/01]. doi:10.1073/pnas.71.10.3971

Ding, S. Z., Smith, M. F., Jr., and Goldberg, J. B. (2008). *Helicobacter pylori* and mitogen-activated protein kinases regulate the cell cycle, proliferation and apoptosis in gastric epithelial cells. *J. Gastroenterol. Hepatol.* 23 (7), e67–e78. [published Online First: 2008/08/16]. doi:10.1111/j.1440-1746.2007.04912.x

Feng, H., Yuan, X., Wu, S., Yuan, Y., Cui, L., Lin, D., et al. (2022). Effects of writers, erasers and readers within miRNA-related m6A modification in cancers. *Cell. Prolif.* 56, e13340. [published Online First: 2022/09/27]. doi:10.1111/cpr.13340

Fu, B., Yin, S., Lin, X., Shi, L., Wang, Y., Zhang, S., et al. (2020). PTPN14 aggravates inflammation through promoting proteasomal degradation of SOCS7 in acute liver failure. *Cell. Death Dis.* 11 (9), 803. [published Online First: 2020/09/27]. doi:10.1038/s41419-020-03014-7

Guo, J., Zheng, J., Zhang, H., and Tong, J. (2021). RNA m6A methylation regulators in ovarian cancer. *Cancer Cell. Int.* 21 (1), 609. [published Online First: 2021/11/20]. doi:10.1186/s12935-021-02318-8

Hatterschide, J., Castagnino, P., Kim, H. W., Sperry, S. M., Montone, K. T., Basu, D., et al. (2022). YAP1 activation by human papillomavirus E7 promotes basal cell identity in squamous epithelia. *Elife* 11, e75466. [published Online First: 2022/02/17]. doi:10.7554/eLife.75466

He, Y., Yue, H., Cheng, Y., Ding, Z., Xu, Z., Lv, C., et al. (2021). ALKBH5-mediated m(6)A demethylation of KCNK15-AS1 inhibits pancreatic cancer progression via regulating KCNK15 and PTEN/AKT signaling. *Cell. Death Dis.* 12 (12), 1121. [published Online First: 2021/12/03]. doi:10.1038/s41419-021-04401-4

Hirata, Y., Maeda, S., Mitsuno, Y., Akanuma, M., Yamaji, Y., Ogura, K., et al. (2001). *Helicobacter pylori* activates the cyclin D1 gene through mitogen-activated protein kinase pathway in gastric cancer cells. *Infect. Immun.* 69 (6), 3965–3971. [published Online First: 2001/05/12]. doi:10.1128/IAI.69.6.3965-3971.2001

Hong, J., Xu, K., and Lee, J. H. (2022). Biological roles of the RNA m(6)A modification and its implications in cancer. *Exp. Mol. Med.* 54 (11), 1822–1832. [published Online First: 2022/11/30]. doi:10.1038/s12276-022-00897-8

Hu, C., Yu, M., Li, C., Wang, Y., Li, X., Ulrich, B., et al. (2020). miR-550-1 functions as a tumor suppressor in acute myeloid leukemia via the hippo signaling pathway. *Int. J. Biol. Sci.* 16 (15), 2853–2867. [published Online First: 2020/10/17]. doi:10.7150/ijbs.44365

Hu, Y., Chen, J., Wang, Y., Sun, J., Huang, P., Feng, J., et al. (2022). Fat mass and obesity-associated protein alleviates $A\beta_{1\text{-}40}$ induced retinal pigment epithelial cells degeneration via PKA/CREB signaling pathway. *Cell. Biol. Int.* 47, 584–597. [published Online First: 2022/11/16]. doi:10.1002/cbin.11959

Huang, H., Weng, H., and Chen, J. (2020). The biogenesis and precise control of RNA m(6)A methylation. *Trends Genet.* 36 (1), 44–52. [published Online First: 2019/12/08]. doi:10.1016/j.tig.2019.10.011

Huang, T., Liu, Z., Zheng, Y., Feng, T., Gao, Q., and Zeng, W. (2020). YTHDF2 promotes spermagonial adhesion through modulating MMPs decay via m(6)A/mRNA pathway. *Cell. Death Dis.* 11 (1), 37. [published Online First: 2020/01/22]. doi:10.1038/s41419-020-2235-4

Imai, S., Ooki, T., Murata-Kamiya, N., Komura, D., Tahmina, K., Wu, W., et al. (2021). *Helicobacter pylori* CagA elicits BRCAness to induce genome instability that may underlie bacterial gastric carcinogenesis. *Cell. Host Microbe* 29 (6), 941–958.e10. [published Online First: 2021/05/15]. doi:10.1016/j.chom.2021.04.006

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 (12), 885–887. [published Online First: 2011/10/18]. 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 (1), 74. [published Online First: 2021/02/22]. doi:10.1038/s41392-020-00450-x

Keates, S., Hitti, Y. S., Upton, M., and Kelly, C. P. (1997). *Helicobacter pylori* infection activates NF-kappa B in gastric epithelial cells. *Gastroenterology* 113 (4), 1099–1109. [published Online First: 1997/10/10]. doi:10.1053/gast.1997.v113.pm9322504

Kim, H., Jang, S., and Lee, Y. S. (2022). The m6A(m)-independent role of FTO in regulating WNT signaling pathways. *Life Sci. Alliance* 5 (5). [published Online First: 2022/02/17]. doi:10.26508/lsa.202101250

Lam, S. Y., Mommersteeg, M. C., Yu, B., Broer, L., Spaander, M. C. W., Frost, F., et al. (2022). Toll-like receptor 1 locus Re-examined in a genome-wide association study update on anti-Helicobacter pylori IgG titers. *Gastroenterology* 162 (6), 1705–1715. [published Online First: 2022/01/16]. doi:10.1053/j.gastro.2022.01.011

Lamb, A., and Chen, L. F. (2010). The many roads traveled by *Helicobacter pylori* to NFκB activation. *Gut Microbes* 1 (2), 109–113. [published Online First: 2011/02/18]. doi:10.4161/gmic.1.2.11587

Li, M., Liu, L., Zang, W., Wang, Y., Du, Y., Chen, X., et al. (2015). miR-365 overexpression promotes cell proliferation and invasion by targeting ADAMTS-1 in breast cancer. *Int. J. Oncol.* 47 (1), 296–302. [published Online First: 2015/05/23]. doi:10.3892/ijo.2015.3015

Li, D., Cai, L., Meng, R., Feng, Z., and Xu, Q. (2020). METTL3 modulates osteoclast differentiation and function by controlling RNA stability and nuclear export. *Int. J. Mol. Sci.* 21 (5), 1660. [published Online First: 2020/03/04]. doi:10.3390/ijms21051660

Li, L., Chen, Y. X., Yang, B., Liao, J. Y., Peng, J. W., and Zhu, S. (2020). The crosstalk between RNA m6A epitranscriptome and TGFβ signaling pathway contributes to the arrest of cell cycle. *Gene* 738, 144483. [published Online First: 2020/02/20]. doi:10.1016/j.gene.2020.144483

Li, Q., Wang, C., Dong, W., Su, Y., and Ma, Z. (2021). WTAP facilitates progression of endometrial cancer via CAV-1/NF-κB axis. *Cell. Biol. Int.* 45 (6), 1269–1277. [published Online First: 2021/02/10]. doi:10.1002/cbin.11570

Li, P., Wang, Y., Sun, Y., Jiang, S., and Li, J. (2022). N (6)-methyladenosine RNA methylation: From regulatory mechanisms to potential clinical applications. *Front. Cell. Dev. Biol.* 10, 1055808. [published Online First: 2022/11/22]. doi:10.3389/fcell.2022.1055808

Li, B., Du, M., Sun, Q., Cao, Z., and He, H. (2022). m6 A demethylase Fto regulates the TNF-alpha-induced inflammatory response in cementoblasts. *Oral Dis.* [published Online First: 2022/10/14]. doi:10.1111/odi.14396

Li, X., Pan, K., Vieth, M., Gerhard, M., Li, W., and Mejías-Luque, R. (2022). JAK-STAT1 signaling pathway is an early response to *Helicobacter pylori* infection and contributes to immune escape and gastric carcinogenesis. *Int. J. Mol. Sci.* 23 (8), 4147. [published Online First: 2022/04/24]. doi:10.3390/ijms23084147

Lin, Y., Shao, Z., Zhao, M., Li, J., and Xu, X. (2021). PTPN14 deficiency alleviates podocyte injury through suppressing inflammation and fibrosis by targeting TRIP6 in diabetic nephropathy. *Biochem. Biophys. Res. Commun.* 550, 62–69. [published Online First: 2021/03/09]. doi:10.1016/j.bbrc.2020.12.030

Lin, C., Wang, Y., Dong, Y., Lai, S., Wang, L., Weng, S., et al. (2022). N6-methyladenosine-mediated SH3BP5-AS1 upregulation promotes GEM chemoresistance in pancreatic cancer by activating the Wnt signaling pathway. *Biol. Direct* 17 (1), 33. [published Online First: 2022/11/19]. doi:10.1186/s13062-022-00347-5

Liu, Z., Zou, H., Dang, Q., Xu, H., Liu, L., Zhang, Y., et al. (2022). Biological and pharmacological roles of m(6)A modifications in cancer drug resistance. *Mol. Cancer* 21 (1), 220. [published Online First: 2022/12/15]. doi:10.1186/s12943-022-01680-z

Maubach, G., Vieth, M., Boccellato, F., and Naumann, M. (2022). Helicobacter pylori-induced NF-κB: Trailblazer for gastric pathophysiology. *Trends Mol. Med.* 28 (3), 210–222. [published Online First: 2022/01/12]. doi:10.1016/j.molmed.2021.12.005

Mayr, C. (2019). What are 3' UTRs doing? *Cold Spring Harb. Perspect. Biol.* 11 (10), a034728. [published Online First: 2018/09/06]. doi:10.1101/cshperspect.a034728

Meyer, K. D., and Jaffrey, S. R. (2014). The dynamic epitranscriptome: N6-methyladenosine and gene expression control. *Nat. Rev. Mol. Cell. Biol.* 15 (5), 313–326. [published Online First: 2014/04/10]. doi:10.1038/nrm3785

Molina-Castro, S. E., Tiffon, C., Giraud, J., Boeuf, H., Sifre, E., Giese, A., et al. (2020). The hippo kinase LATS2 controls Helicobacter pylori-induced epithelial-mesenchymal transition and intestinal metaplasia in gastric mucosa. *Cell. Mol. Gastroenterol. Hepatol.* 9 (2), 257–276. [published Online First: 2019/11/02]. doi:10.1016/j.jcmgh.2019.10.007

Nozawa, Y., Nishihara, K., Peek, R. M., Nakano, M., Uji, T., Ajioka, H., et al. (2002). Identification of a signaling cascade for interleukin-8 production by *Helicobacter pylori* in human gastric epithelial cells. *Biochem. Pharmacol.* 64 (1), 21–30. [published Online First: 2002/07/11]. doi:10.1016/s0006-2952(02)01030-4

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 (1), 48. [published Online First: 2018/03/29]. doi:10.1186/s13045-018-0590-8

Peek, R. M., Jr., Fiske, C., and Wilson, K. T. (2010). Role of innate immunity in Helicobacter pylori-induced gastric malignancy. *Physiol. Rev.* 90 (3), 831–858. [published Online First: 2010/07/29]. doi:10.1152/physrev.00039.2009

Qu, J., Hou, Y., Chen, Q., Chen, J., Li, Y., Zhang, E., et al. (2022). RNA demethylase ALKBH5 promotes tumorigenesis in multiple myeloma via TRAF1-mediated activation of NF-κB and MAPK signaling pathways. *Oncogene* 41 (3), 400–413. [published Online First: 2021/11/12]. doi:10.1038/s41388-021-02095-8

Roignant, J. Y., and Soller, M. (2017). m(6 A in mRNA: An ancient mechanism for fine-tuning gene expression. *Trends Genet.* 33 (6), 380–390. [published Online First: 2017/05/14]. doi:10.1016/j.tig.2017.04.003

Sacco, M. T., Bland, K. M., and Horner, S. M. (2022). WTAP targets the METTL3 m(6)a-methyltransferase complex to cytoplasmic hepatitis C virus RNA to regulate infection. *J. Virol.* 96 (22), e0099722. [published Online First: 2022/11/01]. doi:10.1128/jvi.00997-22

Sasaran, M. O., Melit, L. E., and Dobru, E. D. (2021). MicroRNA modulation of host immune response and inflammation triggered by *Helicobacter pylori*. *Int. J. Mol. Sci.* 22 (3), 1406. [published Online First: 2021/02/13]. doi:10.3390/ijms22031406

Shen, Z., Liu, P., Sun, Q., Li, Y., Acharya, R., Li, X., et al. (2021). FTO inhibits UPR(mt)-induced apoptosis by activating JAK2/STAT3 pathway and reducing m6A level in adipocytes. *Apoptosis* 26 (7-8), 474–487. [published Online First: 2021/07/03]. doi:10.1007/s10495-021-01683-z

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 (7730), 249–253. [published Online First: 2018/11/08]. doi:10.1038/s41586-018-0666-1

Shu, C., Tian, J., Si, X., and Xie, X. (2022). Blueberry anthocyanin extracts protect against Helicobacter pylori-induced peptic epithelium injuries both *in vitro* and *in vivo*: The key role of MAPK/NF-κB pathway. *Eur. J. Nutr.* 61 (5), 2749–2759. [published Online First: 2022/03/16]. doi:10.1007/s00394-022-02830-1

Sun, Z., Chen, W., Wang, Z., Wang, S., Zan, J., Zheng, L., et al. (2022). Matr3 reshapes m6A modification complex to alleviate macrophage inflammation during atherosclerosis. *Clin. Immunol.* 245, 109176. [published Online First: 2022/11/12]. doi:10.1016/j.clim.2022.109176

Toba, H., de Castro Bras, L. E., Baicu, C. F., Zile, M. R., Lindsey, M. L., and Bradshaw, A. D. (2016). Increased ADAMTS1 mediates SPARC-dependent collagen deposition in the aging myocardium. *Am. J. Physiol. Endocrinol. Metab.* 310 (11), E1027–E1035. [published Online First: 2016/05/05]. doi:10.1152/ajpendo.00040.2016

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 (7481), 117–120. [published Online First: 2013/11/29]. doi:10.1038/nature12730

Wang, H. F., Kuang, M. J., Han, S. J., Wang, A. B., Qiu, J., Wang, F., et al. (2020). BMP2 modified by the m(6)A demethylation enzyme ALKBH5 in the ossification of the ligamentum flavum through the AKT signaling pathway. *Calcif. Tissue Int.* 106 (5), 486–493. [published Online First: 2020/01/04]. doi:10.1007/s00223-019-00654-6

Wang, S., Lv, W., Li, T., Zhang, S., Wang, H., Li, X., et al. (2022). Dynamic regulation and functions of mRNA m6A modification. *Cancer Cell. Int.* 22 (1), 48. [published Online First: 2022/01/31]. doi:10.1186/s12935-022-02452-x

Wei, A., Zhao, F., Hao, A., Liu, B., and Liu, Z. (2022). N-acetyl-seryl-aspartyl-lysyl-proline (AcSDKP) mitigates the liver fibrosis via WTAP/m(6)A/Ptch1 axis through Hedgehog pathway. *Gene* 813, 146125. [published Online First: 2021/12/19]. doi:10.1016/j.gene.2021.146125

Xia, X., Zhang, L., Chi, J., Liu, X., Li, H., Hu, T., et al. (2020). *Helicobacter pylori* infection impairs endothelial function through an exosome-mediated mechanism. *J. Am. Heart Assoc.* 9 (6), e014120. [published Online First: 2020/03/17]. doi:10.1161/JAHA.119.014120

Xiao, Y., Chen, J., Yang, S., Sun, H., Xie, L., Li, J., et al. (2022). Maternal mRNA deadenylation and allocation via Rbm14 condensates facilitate vertebrate blastula development. *EMBO J.* 42, e111364. [published Online First: 2022/12/09]. doi:10.15252/embj.2022111364

Xiao, Q., Lei, L., Ren, J., Peng, M., Jing, Y., Jiang, X., et al. (2022). Mutant NPM1-regulated FTO-mediated m(6)A demethylation promotes leukemic cell survival via PDGFRB/ERK signaling Axis. *Front. Oncol.* 12, 817584. [published Online First: 2022/02/26]. doi:10.3389/fonc.2022.817584

Xu, W., Huang, Y., Yang, Z., Hu, Y., Shu, X., Xie, C., et al. (2018). *Helicobacter pylori* promotes gastric epithelial cell survival through the PLK1/PI3K/Akt pathway. *Onco Targets Ther.* 11, 5703–5713. [published Online First: 2018/09/27]. doi:10.2147/OTT.S164749

Yang, Y., Ma, Q., Li, Z., Wang, H., Zhang, C., Liu, Y., et al. (2021). Harmine alleviates atherogenesis by inhibiting disturbed flow-mediated endothelial activation via protein tyrosine phosphatase PTPN14 and YAP. *Br. J. Pharmacol.* 178 (7), 1524–1540. [published Online First: 2021/01/22]. doi:10.1111/bph.15378

Yang, G., Yao, G., Xu, Z., Fan, H., Liu, X., He, J., et al. (2021). Expression level of ADAMTS1 in granulosa cells of PCOS patients is related to granulosa cell function, oocyte quality, and embryo development. *Front. Cell. Dev. Biol.* 9, 647522. [published Online First: 2021/04/30]. doi:10.3389/fcell.2021.647522

Yao, Y., Bi, Z., Wu, R., Zhao, Y., Liu, Y., Liu, Q., et al. (2019). METTL3 inhibits BMSC adipogenic differentiation by targeting the JAK1/STAT5/C/EBPβ pathway via an m6A-YTHDF2-dependent manner. *FASEB J.* 33 (6), 7529–7544. [published Online First: 2019/03/14]. doi:10.1096/fj.201802644R

Zhang, S. Y., Zhang, S. W., Fan, X. N., Meng, J., Chen, Y., Gao, S. J., et al. (2019). Global analysis of N6-methyladenosine functions and its disease association using deep learning and network-based methods. *PLoS Comput. Biol.* 15 (1), e1006663. [published Online First: 2019/01/03]. doi:10.1371/journal.pcbi.1006663

Zhang, M., Yang, C., Ruan, X., Liu, X., Wang, D., Liu, L., et al. (2022). CPEB2 m6A methylation regulates blood-tumor barrier permeability by regulating splicing factor SRSF5 stability. *Commun. Biol.* 5 (1), 908. [published Online First: 2022/09/07]. doi:10.1038/s42003-022-03878-9

Zhao, X., Yang, Y., Sun, B. F., Shi, Y., Xiao, W., Yang, X., et al. (2014). FTO-dependent demethylation of N6-methyladenosine regulates mRNA splicing and is required for adipogenesis. *Cell. Res.* 24 (12), 1403–1419. [published Online First: 2014/11/22]. 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 (1), 18–29. [published Online First: 2012/11/28]. doi:10.1016/j.molcel.2012.10.015

Zhou, R., Ni, W., Qin, C., Zhou, Y., Li, Y., Huo, J., et al. (2022). A functional loop between YTH domain family protein YTHDF3 mediated m(6)A modification and phosphofructokinase PFKL in glycolysis of hepatocellular carcinoma. *J. Exp. Clin. Cancer Res.* 41 (1), 334. [published Online First: 2022/12/07]. doi:10.1186/s13046-022-02538-4

Zuo, W., Yang, H., Li, N., Ouyang, Y., Xu, X., and Hong, J. (2022). *Helicobacter pylori* infection activates Wnt/β-catenin pathway to promote the occurrence of gastritis by upregulating ASCL1 and AQP5. *Cell. Death Discov.* 8 (1), 257. [published Online First: 2022/05/11]. doi:10.1038/s41420-022-01026-0

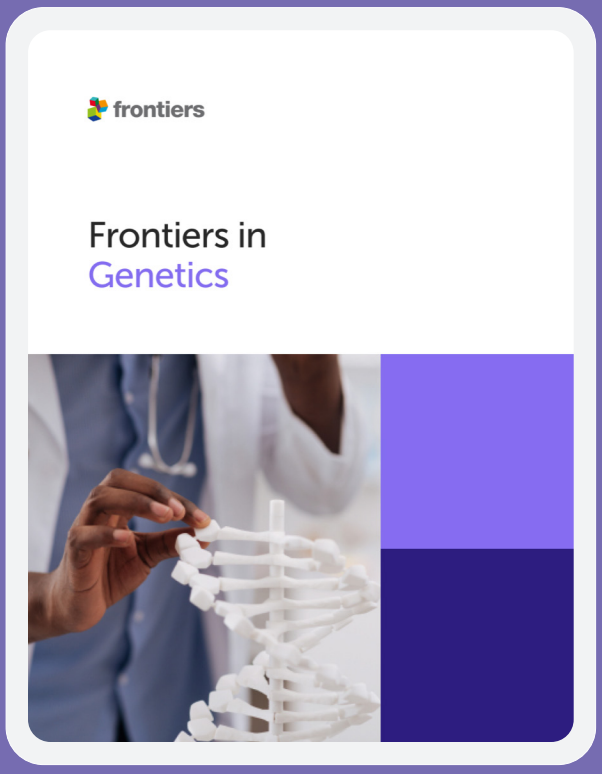
Frontiers in Genetics

Highlights genetic and genomic inquiry relating to all domains of life

The most cited genetics and heredity journal, which advances our understanding of genes from humans to plants and other model organisms. It highlights developments in the function and variability of the genome, and the use of genomic tools.

Discover the latest Research Topics

See more →



Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact