# GENE REGULATION EXPLORED BY SYSTEMS BIOLOGY IN LIVESTOCK SCIENCE

EDITED BY: Andressa Oliveira De Lima, Priyanka Banerjee and Aline Silva Mello Cesar

PUBLISHED IN: Frontiers in Genetics

frontiers Research Topics

ISSN 1664-8714
ISBN 978-2-88976-017-6
DOI 10.3389/978-2-88976-017-6

## About Frontiers

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

## Frontiers Journal Series

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

## Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.
Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

## What are Frontiers Research Topics?

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

# GENE REGULATION EXPLORED BY SYSTEMS BIOLOGY IN LIVESTOCK SCIENCE

Topic Editors:
**Andressa Oliveira De Lima,** University of Washington, United States
**Priyanka Banerjee,** Auburn University, United States
**Aline Silva Mello Cesar,** University of São Paulo, Brazil

**Citation:** De Lima, A. O., Banerjee, P., Cesar, A. S. M., eds. (2022). Gene Regulation Explored by Systems Biology in Livestock Science. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-88976-017-6

# Table of Contents

frontiers | Frontiers in Genetics

EDITORIAL
published: 08 April 2022
doi: 10.3389/fgene.2022.859061

# Editorial: Gene Regulation Explored by Systems Biology in Livestock Science

*Priyanka Banerjee*[1]*, *Aline Silva Mello Cesar*[2] *and Andressa Oliveira De Lima*[3]

[1]Department of Animal Sciences, College of Agriculture, Auburn University, Auburn, AL, United States, [2]Department of Agri-Food Industry, Food and Nutrition, University of São Paulo, Piracicaba, Brazil, [3]Department of Genome Sciences, University of Washington, Seattle, WA, United States

**Keywords: lncRNA, miRNA, transcriptomics, proteomics, methylation**

## Editorial on the Research Topic

### Gene Regulation Explored by Systems Biology in Livestock Science

In the livestock production system, comprehending animal health and production is increasingly necessary to meet the global food demands. The production of animals with desired traits is, therefore, a prerequisite. To meet future food requirements, the strategy targets vertical growth (increasing productivity) rather than expanding the livestock population. Improving productivity requires a better understanding of genes and genomes, and consequently, their influence on the trait of interest (i.e., production, reproduction, health and disease, or any other relevant traits in livestock), so that appropriate selection and breeding decisions can be implemented to improve livestock performance.

Over the past decade, much biological research has focused on gene expression and regulation by non-coding RNAs (ncRNAs), variants, cis- or distal regulatory elements (transcription factors, enhancers and silencers), histone modifications, and DNA methylation. Advancement in high-throughput technologies and cost reduction has enabled multi-omics studies to be performed on a large scale. Therefore, the crosstalk between multiple molecular layers is assessed by a systems biology approach that provides a systematic view of the regulatory mechanisms underlying complex traits. This approach proves beneficial for production systems (e.g., optimize animal nutrition, meat quality, or animal management) by selecting the desirable animals and integrating accurate breeding programs or innovative management systems.

For this research topic, we sought high-quality research papers describing novel insights into genetic and environmental factors that impact the mechanism and expression of production, reproduction and disease traits in livestock research. Research topics included the novel studies on the expression of genes, micro RNAs (miRNAs), long non-coding RNAs (lncRNAs), circular RNAs (circRNAs), proteins and methylome for the trait of interest in livestock. The final research topic has 12 articles covering the aforementioned aspects in bovine, ovine, caprine, porcine and poultry populations.

The quest for ideal therapeutic targets and biomarkers in disease traits in bovines has been riddled with many obstacles stemming from the molecular complexity of the disease and co-morbidities. Advances in omics technologies and the resulting amount of available data encompassing transcriptomics have created an opportunity to integrate omics datasets to identify molecular changes and represent data through networks in disease traits such as bovine respiratory disease (BRD) and mastitis. Hasankhani et al., identified key modules and potential hub genes from transcriptome data using co-expression networks that underlie BRD. With two different approaches of module–trait relationships and module preservation analysis, the authors identified eight candidate modules and 307 hub-genes involved in the immune response and BRD pathogenesis (Hasankhani et al.). Using a similar approach, Ghahramani et al., identified candidate genes and modules associated with mastitis in dairy cattle. Furthermore, they identified 360 meta genes within two modules by integrating microarray and RNA-Seq data. Additionally, the authors used attribute weighting and machine-learning methods to optimize predictive models using hub genes that were informative in *Escherichia coli* mastitis (Ghahramani et al.). These studies utilized a

OPEN ACCESS

***Edited and reviewed by:***
*Duy Ngoc Do,*
*Dalhousie University, Canada*

***Correspondence:**
*Priyanka Banerjee*
*pzb0035@auburn.edu*

***Specialty section:***
*This article was submitted to Livestock Genomics, a section of the journal Frontiers in Genetics*

***Received:*** *20 January 2022*
***Accepted:*** *18 March 2022*
***Published:*** *08 April 2022*

***Citation:***
*Banerjee P, Cesar ASM and Lima AOD (2022) Editorial: Gene Regulation Explored by Systems Biology in Livestock Science. Front. Genet. 13:859061. doi: 10.3389/fgene.2022.859061*

system's level approach to provide a complete understanding of complex biological systems of BRD and mastitis beyond the molecular level in cattle.

Even though omics technologies reveal global changes in RNA modifications under various conditions, researchers also focused on diverse functions of specific RNAs, particularly ncRNAs and circRNAs, due to their potential in regulating gene expression. The ncRNAs are classified into several sub-classes, such as small non-coding RNAs, including miRNAs, small interfering RNAs, and lncRNAs, to name a few. Although lncRNAs biochemically resemble messenger RNAs (mRNAs), they do not template protein synthesis. Advances in computational biology and the evolution of sensitive RNA sequencing have facilitated the discovery of numerous lncRNAs and encouraged the study of their roles in livestock. An interesting study reported within this research topic involved the regulatory role of lncRNA in the transcriptome of abomasal lymph node tissue samples from adult Spanish Churra sheep after experimental infection with the gastrointestinal nematode *Teladorsagia circumcincta* (Chitneedi et al.). The authors identified ten differentially expressed lncRNAs between samples from animals that differ in infection resistance associated with signaling pathways like cellular growth, proliferation and development, cellular stress and injury, intracellular and second messenger signaling and apoptosis (Chitneedi et al.). In a separate study on sheep hair follicle development and morphogenesis, Xu et al., identified the *STAT3* gene partially inhibiting cell proliferation via direct negative regulation of *FST* gene expression.

miRNAs are small and highly conserved non-coding RNA molecules that orchestrate various biological processes through post-transcriptional expression regulation. A study by Kusama et al., aimed to characterize proteins and exosomal miRNAs in the uterine flushing of pregnant and non-pregnant cows after artificial insemination. The authors identified 336 proteins, of which 260 were more than two-fold higher in pregnant cows. The authors identified *SUGT1* as the best predictor for the presence of embryos in the uterus that altered protein composition and exosomal miRNA contents in the uterine fluid (Kusama et al.). In an independent miRNA study in cattle, bta-miR-150 was found to have a negative regulatory effect on the differentiation of bovine adipocytes and promoted proliferation, inhibited adipocyte differentiation, and reduced lipid droplet formation (Chen et al.). The research clarified the relationship between bta-miR-150 and adipocyte differentiation in cattle. Differentiation of intramuscular preadipocytes was also inhibited by overexpression of *KLF4* by targeting C/EBPβ in goats (Xu et al.). Since the adipose tissue in meat impacts the economic value of animals, these studies focusing on the molecular mechanisms underlying adipose tissue generation were primarily crucial for the beef production industry. As beef producers continue to improve efficiency, another crucial aspect of skeletal muscle mass impacts consumer acceptance of meat products. Sheng et al., used the proteomics approach to unravel the mechanism of myostatin in regulating cattle skeletal development.

Another study involved co-expression analysis focusing on miRNA-mRNA interaction networks (Yang et al.). An integrative analysis of the miRNA-mRNA expression profiles in the lungs of high- and low-altitude pigs (Tibetan pigs and Landrace pigs, respectively) identified molecular pathways and networks involved in the genetic adaptation of Tibetan pigs to hypoxic conditions (Yang et al.).

A separate group of lncRNA—circRNA—influences cellular physiology through various molecular mechanisms by modulating gene expression. A study by Li et al. identified circTAF8 regulating myoblast development and associated carcass traits in chicken.

One of the advantages of livestock research is that it proves to be an excellent physiological model for studies related to human health or disease. Studies reported by Chen et al., revealed the similarities and diversity of sperm methylation patterns among three commercial pig breeds and between humans and pigs. These findings elucidate the mechanism of male fertility and the changes in commercial traits that undergo strong selection. In a separate study by Xu et al., methylation changes in *SLC4A11* and *MFSD3* genes were analyzed after Deoxynivalenol toxicity in porcine cells that may contribute to the detection of biomarkers and drug targets.

Considering all the above studies, we believe that a holistic approach combining statistics, bioinformatics, and mathematical modeling to integrate and analyze large amounts of data generated and targeting some of the key regulators has shed light on mechanisms regulating the economically important traits in livestock. We hope that the reader will find this research topic a helpful reference for the state-of-the-art in the emerging field of livestock research.

## AUTHOR CONTRIBUTIONS

PB wrote the first draft. AC and AL provided critical comments and editorial suggestions for revisions. All the authors agreed on the submitted version.

## ACKNOWLEDGMENTS

The guest editors wish to thank all the authors and reviewers for their valuable contributions to this Research Topic. We hope that this collection of articles will interest the livestock science and genetics community.

**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: 03 February 2021
doi: 10.3389/fgene.2021.636550

# Bovine Pre-adipocyte Adipogenesis Is Regulated by bta-miR-150 Through mTOR Signaling

*Xingyi Chen*[1], *Sayed Haidar Abbas Raza*[1], *Xinhao Ma*[1], *Jiangfang Wang*[1], *Xiaohui Wang*[1], *Chengcheng Liang*[1], *Xinran Yang*[1], *Chugang Mei*[1], *Syed Muhammad Suhail*[2] *and Linsen Zan*[1,3]*

[1] College of Animal Science and Technology, Northwest A&F University, Xianyang, China, [2] Department of Livestock Management, Breeding and Genetics, The University of Agriculture, Peshawar, Pakistan, [3] National Beef Cattle Improvement Center, Northwest A&F University, Xianyang, China

Micro RNA (miR) are recognized for their important roles in biological processes, particularly in regulatory componentization. Among the miR, miR-150 has been the focus of intense scrutiny, mostly due to its role in malignant tumors. A comparison between steer and bull adipose tissues identified bta-miR-150 as one of the nine downregulated miRNAs, although its function remains unknown (GEO:GSE75063). The present study aimed to further characterize the role of bta-miR-150 in cattle. bta-miR-150 has a negative regulatory effect on the differentiation of bovine adipocytes and promotes proliferation. Overexpression of bta-miR-150 can promote mRNA and protein expression of the marker genes *CDK1*, *CDK2*, and *PCNA*, increase the number of EdU-stained cells, promote adipocyte proliferation, inhibit adipocyte differentiation, and reduce lipid droplet formation. Results of RNA-seq and WGCNA analyses showed that the mammalian target of the rapamycin signaling pathway, which plays a major regulatory role, is dysregulated by the overexpression and inhibition of miR-150. We found that the target gene of bta-miR-150 is *AKT1* and that bta-miR-150 affects *AKT1* phosphorylation levels. These results showed that bta-miR-150 plays a role in adipogenic differentiation and might therefore have applications in the beef industry.

**Keywords: bta-miR-150, mTOR, RNA-seq, adipocyte differentiation, WGCNA**

OPEN ACCESS

***Edited by:***
*Andressa Oliveira De Lima, São Paulo State University, Brazil*

***Reviewed by:***
*Juliana Afonso, University of São Paulo, Brazil*
*Rajwali Khan, University of Agriculture, Pakistan*

***Correspondence:**
*Linsen Zan*
*zanlinsen@163.com*

***Specialty section:***
*This article was submitted to Livestock Genomics, a section of the journal Frontiers in Genetics*

***Received:*** *01 December 2020*
***Accepted:*** *04 January 2021*
***Published:*** *03 February 2021*

***Citation:***
*Chen X, Raza SHA, Ma X, Wang J, Wang X, Liang C, Yang X, Mei C, Suhail SM and Zan L (2021) Bovine Pre-adipocyte Adipogenesis Is Regulated by bta-miR-150 Through mTOR Signaling. Front. Genet. 12:636550. doi: 10.3389/fgene.2021.636550*

## INTRODUCTION

Adipose tissue that differentiates from mesenchymal stem cells during the embryonic period is responsible for many functions, including energy storage, lipid metabolism, and hormone secretion (Ambele et al., 2020; Poklukar et al., 2020). Since the amount of adipose tissue in meat impacts the economic value of animals, understanding the molecular mechanisms underlying adipose tissue generation in animals will ensure that production is maximized within an appropriate range. MicroRNA are non-coding RNA with a maturity length of 18–25 nt; they are ubiquitous in the animal kingdom and are produced by RNA polymerase II transcription. MicroRNA are involved in many different physiological processes in animals (Raza et al., 2020), including inhibition of the translation of target genes, affecting mRNA degradation, and regulating the transcription of target genes (Twayana et al., 2013). Differentiation of adipocytes is regulated by many miRNA, which are the focus of the current investigation. Advancements to high-throughput sequencing technology

has allowed deeper exploration of cellular regulatory processes (Katz et al., 2010). Previous studies have already revealed the regulatory effects of human miRNA, including miR-204 (Zhang Z. et al., 2020), miR-27a (Kim et al., 2010), miR-210 (Ren et al., 2020), miR-24 (Liu et al., 2020), miR-378 (Duarte et al., 2020), miR-149-5p (Khan et al., 2020), miR-143 (Zhang L. et al., 2020), and miR-145 (Wang et al., 2020), on adipogenesis. These classic mechanisms also occur in the adipose tissue, where they have a large impact on the surrounding tissue (Arcidiacono et al., 2020; Raza et al., 2020). For example, miR-143 promotes the differentiation of adipocytes by directly acting on its target gene, *MAP2K5*. Many biologically significant miRNA have been discovered, with high-throughput sequencing technology continuously improving in parallel with a reduction in cost (An et al., 2016).

Because of its high expression level in immune-related cells, hsa-miR-150 was one of the first miRNAs to be studied in humans. Human (hsa)-miR-150 plays an important role in the differentiation of hematopoietic cell lines; thus, investigations of hsa-miR-150 have mainly focused on malignant tumors (Yugawa et al., 2020). In animal, miR-150 has been implicated in the mTOR pathway, where it regulates the expression of leptin in adipocytes (Scrutinio et al., 2017) and the production of lipids such as triglycerides (TGs) and free fatty acids (FFAs). The master metabolic factor PGC-1a mediates the reciprocating cycle between TG and FFA to achieve physiologically stable energy consumption (Kang et al., 2018). In pigs, ssc-miR-150 is differentially regulated in the adipose tissue miRNA of lean, rather than fatty pig breeds. Ssc-miR-150 is thought to act directly on the 3′UTR of *CYP3A4* to promote FFA-induced adipose tissue degeneration (Ma et al., 2020). Since miRNA is highly conserved in the animal kingdom, bta-miR-150 might also function in the production of adipose tissue in cattle. Importantly, bta-miR-150 is one of the nine downregulated miRNA among differential expression profiles of steer and bull fat tissues (Liang et al., 2019).

With the global increase in human population and improvements in living standards, beef fat is increasingly being valued by breeders for the flavor it imparts, as well as its high energy value as a food source (Lillehammer et al., 2011). The 480-amino-acid enzyme, RAC-alpha serine/threonine-protein kinase (AKT1), located downstream of the mammalian target of the rapamycin signaling (mTOR) pathway regulates the growth of mammalian cells and is an important protein for regulating fat deposition (Zhang et al., 2017; Song et al., 2018). In mice, miR-150 inhibits *AKT1*. Bovine chromosome 21 encodes *AKT1*, which is an important member of the PI3K/AKT/mTOR pathway. The activity of *AKT* requires phosphorylation at Thr308 and is further enhanced by phosphorylation at Ser473. When Ser473 is simultaneously phosphorylated by *PDK1* and *mTORc2*, *AKT1* further regulates cell proliferation through downstream FOXO transcription factors and p53 regulatory factors (Cai et al., 2019; Sanchez-Gurmaches et al., 2019; Du et al., 2020) The present study aimed to determine the relationship between bta-miR-150 and *AKT1* by examining the effects and mechanisms of bta-miR-150 in regulating the proliferation and differentiation of Qinchuan beef cattle pre-adipocytes, thus providing a basis for targeted breeding and genetic improvement based on bta-miR-150. In beef cattle, the process of breeding to increase the intramuscular fat content would be supported by an understanding of the genetic basis of adipogenesis. Our research further clarifies the relationship between bta-miR-150 and adipocyte differentiation in cattle. This information paves the way for further research to expand our understanding of the complex process of adipocyte differentiation and adipose tissue generation.

# MATERIALS AND METHODS

## Isolation of Primary Bovine Pre-adipocytes and Cell Culture

The tissue and cell samples used in the experiment in this study were collected from three 1-day-old healthy Qinchuan cattle bulls with consistent growth bred from the National Beef Cattle Improvement Center of Northwest Agriculture and Forestry University (Yangling, China). Bacteria surgical instruments collected tissue samples such as heart, liver, spleen, lung, and muscle. The sample was placed in a sterile, DNase- and RNase-free cell cryotube. Then the sample was immediately put in liquid nitrogen for freezing, and finally stored in a refrigerator at −80°C for later use.

We used 1-day-old healthy bulls to isolate the original bovine pre-adipocytes, removed the adipose tissue from different parts under aseptic conditions, washed with PBS supplemented with 10% antibiotics (penicillin/streptomycin) 3 times, and then added I Collagenase digestion for 1–2 h in a 37°C water bath shaker. After digestion, the cells were filtered with a cell sieve and the supernatant was discarded after centrifugation. The red blood cell lysate was added for lysis, and then an appropriate amount of DMEM-F12 (Gibco, Grand Island, NY) medium containing 10% fetal bovine serum (FBS, Invitrogen) was added to resuspend the cells for seeding for subsequent experiments. The cells used in this experiment were all the 3rd generation cells after passage.

## Construction and Transfection of Plasmid and RNA Oligonucleotides

The cells were inoculated in DMEM-F12 supplemented with 10% FBS and 1% antibiotics and kept continuously at 37°C and 5% $CO_2$. According to the manufacturer's instructions, the miR-150 mimic (50 nM), miR-150 mimic NC (50 nm), miR-150 inhibitor (100 nM), miR-150 inhibitor NC (100 nM), si-*AKT1* (100 nM), and si-NC were transfected into the cells using Lipofectamine$^{TM}$ 3000 (Invitrogen, San Diego, GA, LUSA). We induced differentiation of adipocytes 48 h after transfection. After 2 days of induction with DMI (0.5 mm IBMX, 1 μm DXMS, and 2 μm insulin) induction solution, we changed to a maintenance medium containing 5 μg/mL insulin. After differentiation induction, adipocytes were harvested at d0, d2, d4, d6, d8, and d10. (We recorded the day of adding DMI as day 0 of induction of differentiation). Two restriction enzymes (Takara, Beijing, China), XhoI and NotI, were used to construct the vector psiCHECK-2(Laboratory retention). The detailed sequence is shown in **Table 1**. The Bta-miR-150 mimic, mimic NC, inhibitor, and inhibitor NC used in this stage were purchased from

**TABLE 1 |** The sequence of RNA oligonucleotides and plasmid.

| Name | Sequence (5′-3′) |
|---|---|
| miR-150 mimic | sense: UCUCCCAACCCUUGUACCAGUGU<br>antisense: ACACUGGTACAAGGGUUGGGAGA |
| mimic NC | UUGUACUACACAAAAGUACUG |
| miR-150 inhibitor | ACACUGGUACAAGGGUUGGGAGA |
| inhibitor NC | CAGUACUUUUGUGUAGUACAA |
| si-AKT1 | sense: GCCAUGAAGAUCCUAAAGATT<br>antisense: UCUUUAGGAUCUUCAUGGCTT |
| si-NC | sense: UUCUCCGAACGUGCACGUTT<br>antisense: ACGUGACACGUUCGGAGAATT |
| psiCHECK-2 | 5′… GAGCAGTAATTCTAGGCGATCGCTCGAGCCCG GGAATTCGTTTAAACCTAGAGCGGCCGCTGGCCGC AATAAAATA… 3′ |

RiboBio, Guangzhou, China; the si-*AKT1* and si-NC used in this stage were purchased from GenePharma, Shanghai, China. The detailed sequence in is shown **Table 1**.

## Differential Expression and Pathway Enrichment Analyses

Transfected miR-150 mimic, mimic NC, inhibitor, inhibitor NC, and bovine adipocytes were used to induce differentiation, and RNA on day 2 and day 6 were collected. A control was set for each treatment, with three technical replicates per group and a total of eight groups (24 samples). After RNA extraction, the HiSeq-PE150 platform (Illumina, Sandiego, CA, USA) was used for RNA sequence analysis. After filtering the original data, the reads were matched to the bovine genome using HiSat2 (Bos-taurusUMD3.1, release94, Ensembl database), and feature counts were used to count gene expression. The DESeq2 R software package was used to screen DEGs, with $Log_2$ fold change ($log_2FC$) $> 1$ and false discovery rate (FDR) $< 0.05$ as the screening criteria. ClusterProfiler was used to realize KEGG pathway analysis and FDR $< 0.05$ was considered significant.

## Extraction of RNA, Quantitative Real-Time PCR (qRT-PCR)

After 48 h of transfection, we used the RNAiso Plus Kit (Trizol, Takara, Beijing, China) to extract RNA. The specific extraction method can refer to the reference (Junjvlieke et al., 2020). After we removed the culture medium from the cells, we washed the cells with PBS 2∼3 times in each well, then added 1 mL Trizol, placed the mixture at 4° for 4 min, and transferred the mixture to a 1.5 mL centrifuge tube. Two hundred microliters of chloroform was added to the centrifuge tube, shaken vigorously for 30 s, left to stand for 5 min, and centrifuged for 15 min. The supernatant was then taken (3 layers in total), isopropanol was added, mixed upside down, left to stand on ice for 10 min, and then centrifuged for 10 min. After centrifugation, add 1 mL of 75% cold ethanol was slowly added to a small amount of precipitation at the bottom, centrifuged at 4°C for 5 min, and then the ethanol was discarded. The cells were dried for 2–5 min at room temperature, then an appropriate amount of RNase-free DEPC water was added to dissolve the precipitate. NanoQuant plateTM (TECAN, Infinite M200 PRO) was used to determine the concentration and purity of RNA, and it was stored at −80°C for later use. The method of extracting total RNA in tissues was to take out about 0.5 g tissue sample in a refrigerator at −80°C, add it to liquid nitrogen for pre-cooling, and then place it in a mortar cooled with liquid nitrogen for grinding. After grinding, the sample is transferred to a 1.5 mL enzyme-free centrifuge tube with 1 mL Trizol, and then placed on ice for 15 min. The next step is the same as that of extracting cellular RNA.

Reverse transcription of total RNA was performed following Prime Script ™ RT reagent Kit with gDNA Eraser (Takara, China), TB Green® Premix Ex TaqTM II(TAKARA, Beijing, China) instructions. The reverse transcription and qPCR of miRNA were performed with miRcute Plus miRNA First-Strand cDNA Kit and miRNA Plus miRNA qPCR Kit (SYBR Green) (Tiangen, Beijing, China). U6 and β – actin was used as the internal control for miRNA quantification and mRNA quantification, respectively. The sequence of all PCR primers is shown in **Table 2**. All experiments were performed with three biological replicates and three technical replicates. The 2-ΔΔCt method was used to analyze the relative expression levels of different qPT-PCR data.

## Western Blot Analysis

The methods of extraction and Western blot analysis of total protein have been introduced in references (Khan et al., 2019; Wang et al., 2019). The protein bands added with Millipore chemiluminescence solution were placed on the Gel DocTMXR+ System (Bio-Rad, Hercules, CA) for exposure imaging. The antibody in **Table 3** is the antibody we used in the experiment.

## EdU Staining and CCK-8 Assay

The Qinchuan cattle pre-adipocytes were seeded in a 96-well plate (NEST, Jiang Su, China) and cultured to a density of 40–60% confluence, which was reached at a density of $5.0 \times 10^4$ cells for transfection. The specific processing steps are detailed in reference (Chu et al., 2019). We added 100 μL PBS to each well of the last stained cells and washed them 1 to 3 times, and then took photographs under a microscope. The instrument used for taking pictures was an Olympus 1 × 71 microscope (Olympus, Tokyo, Japan). The treatment method of CCK-8 cell cycle detection was the same as that of EdU. The specific operation steps were carried out according to the reference (Khan et al., 2020). The kit used was TransDetect Cell Counting Kit (Transgen biotechnology, Beijing, China), and the detection instrument was an Infinite® 200PRO (Tecan Trading AG, Switzerland).

## Bodipy, Oil Red O Staining, and Triacylglycerol Assay

The differentiated cells were stained for different days. The cells on the second day were stained with Bodipy, and the cells on the fourth and sixth days were stained with oil red O. In short, both Bodipy (1:1,000) and DAPI (1:1,000) were diluted in PBS. The cells to be stained were washed with PBS 3 times, fixed with 4% paraformaldehyde for 30 min, and then stained with Bodipy for 30 min. After staining, the cells were washed with PBS 3 times

**TABLE 2 |** mRNA and miRNA Real-time quantitative PCR primer sequences.

| Genes | Primer Sequence (5′-3′) | Annealing temperature |
|---|---|---|
| β-actin | F: CATCGGCAATGAGCGGTTCC<br>R: CCGTGTTGGCGTAGAGGTC | 60°C |
| PCNA | F: CCTTGGTGCAGCTAACCCTT<br>R: TTGGACATGCTGGTGAGGTT | 60°C |
| CDK1 | F:AGTGGAAACCAGGAAGCTTAG<br>R:ATTCGTTTGGCAGGATCATAGA | 60°C |
| CDK2 | F:GGGTCCCTGTTCGTACTTATAC<br>R:CCACTGCTGTGGAGTAGTATTT | 60°C |
| AKT-1 | F:AAGGAGATCATGCAGCACCGATTC<br>R: GTCTTGGTCAGGTGGCGTAATGG | 60°C |
| CEBPα | F: ATCTGCGAACACGAGACG<br>R: CCAGGAACTCGTCGTTGAA | 60°C |
| PPARγ | F: GACGACAGACAAATCACCGT<br>R: CTTCCACGGAGCGAAACTGA | 60°C |
| FABP4 | F:TGAGATTTCCTTCAAATTGGG<br>R: CTTGTACCAGAGCACCTTCATC | 60°C |
| RRAGD | F: TAAGAAGGAGCGGCAAGT<br>R: GGAAGTCCCAAATCTGAAA | 60°C |
| RPS6KA3 | F:ACCTAGCAACATTCTTTATGTGGAT<br>R: GCATCATAGCCTTGTCGT | 60°C |
| RPS6KB1 | F: ATCACCAAGGTCACGTCAAAC<br>R: TGCTCCCAAACTCCACCAAT | 60°C |
| LPIN1 | F: AGTGAATCTTCAGATGCGTTTA<br>R: CAGGTGTCGGCTTCGTTT | 60°C |
| NFKB1 | F: GAGAACTTCGAGCCTCTGTAC<br>R: CTCATAGGGTTTCCCATTTA | 60°C |
| EIF4E | F: TCTAATCAGGAGGTTGCT<br>R: CCCACATAGGCTCAATAC | 60°C |
| PIK3CA | F:TGGGTTTCTCGGTCCTAA<br>R:CAGTCCGTCCAGTCATCC | 60°C |
| TNFRSF1B | F: CCAGCACAGCTCCAAGCA<br>R:CAACTATCAGAAGCAGACCCAAT | 60°C |
| INSR | F: GGAAGGCGAGAAGACCAT<br>R: TGACACCAGGGCATAGGA | 60°C |
| EIF4EBP2 | F: GTTCCTGATGGAGTGTCGGA<br>R: AACTGTGACTCTTCACCGCCT | 60°C |
| PIK3R1 | F: GAGCGGGAAGAGGACATTGA<br>R: TCTCCCCTGTCGTCTCGTTA | 60°C |
| mTOR | F: TGATGCAGAAGGTCGAGGTG<br>R: GATGGGTGTCTATCGCCCAG | 60°C |
| PIK3R1 | F: CTATCTCCTGGACTTACCG<br>R: GGCGACCTAATGAGTTTC | 60°C |
| U6 | F:TAGCCACCCTCAAGTATGTTCG<br>R:CGAGGTAGGAGGACAGGAGT | 60°C |
| miR-150 | UCUCCCAACCCUUGUACCAGUGU | 60°C |

**TABLE 3 |** Antibody information.

| Name of antibody | Immune features | The company | Dilution ratio |
|---|---|---|---|
| Anti-AKT1 antibody | Rabbit polyclonal | CST | 1:1,000 |
| Phospho-Akt (Ser473) antibody | Rabbit polyclonal | CST | 1:1,000 |
| Anti-CDK1 antibody | Rabbit monoclonal | Abcam | 1:1,000 |
| Anti-CDK2 antibody | Rabbit monoclonal | Abcam | 1:1,000 |
| Anti-PCNA antibody | Rabbit monoclonal | Abcam | 1:5,000 |
| Anti-PPARγ antibody | Rabbit polyclonal | Boster | 1:1,000 |
| Anti-FABP4 antibody | Rabbit monoclonal | Abcam | 1:1,000 |
| Anti-CEBPα antibody | Rabbit monoclonal | Abcam | 1:1,000 |
| Anti-β-actin antibody | Rabbit polyclonal | Abcam | 1:10,000 |

and then stained with DAPI. Finally, the cells were photographed under a microscope. The whole process needs to be done away from light.

The oil red O staining procedure was as follows: the cells were washed in the six-well plate with PBS 3 times, fixed with 4% paraformaldehyde for 30 min, dyed with 60% oil red O (solvent: isopropanol, 0.15 g oil red powder/100 mL) for 30 min, then washed with PBS and observed under a microscope. We measured the amount of intracellular TAG using the cell/tissue triglyceride assay kit (Applygen Technologies, Beijing, China) according to the protocol recommended by the manufacturer (Ma et al., 2018).

## Luciferase Reporter Assay

To verify whether bta-miR-150 targets the 3′-UTR of AKT1, we obtained the mature sequence of bta-miR-150 from the miRBase (http://www.mirbase.org/). The reporter of luciferase was produced by ShengGong (Sangon Biotech Co., Ltd, Shanghai, China) through synthesizing 3′-UTR of AKT1 wild-type (WT) and mutation-type (MUT) of AKT1 sequence. Xhol and Notl cutting sites of Wt-3′-UTR containing miR-150 target were amplified. AKT1-3′UTR WT or MUT were cloned into psiCHECK-2 vector (Laboratory retention) by Xhol and Notl restriction sites (Promega, Madison, WI, USA). After culturing HEK293A cells for 2 days, they were then co-transfected into bovine adipocytes with about 0.16 μg of MUT vector/AKT1-3′ UTR WT and 5 pmol miR-150 mimic/negative control. At 48 h after transfection, the luciferase activity was measured by Dual-Luciferase® Reporter Assay System (Promega, Madison, WI, USA). Luciferase assay was carried out in three repeat wells, and three experiments were carried out.

## Weighted Gene Co-expression Network Analysis (WGCNA)

WGCNA, an R package, was used to construct the weighted gene co-expression network. With the standardized gene expression matrix as input, the variation degree of each gene expression level among samples was calculated, and the top 50% genes with the largest variation were selected for WGCNA. After the threshold screening, the scaleless adjacency matrix was obtained by power processing with $\beta = 12$. In order to better evaluate the correlation of gene expression patterns, the adjacency matrix was further transformed into a topological overlap matrix (ToM), and the dynamic cutting algorithm was used to cluster and partition genes by using the topological difference matrix (distom = 1-tom).

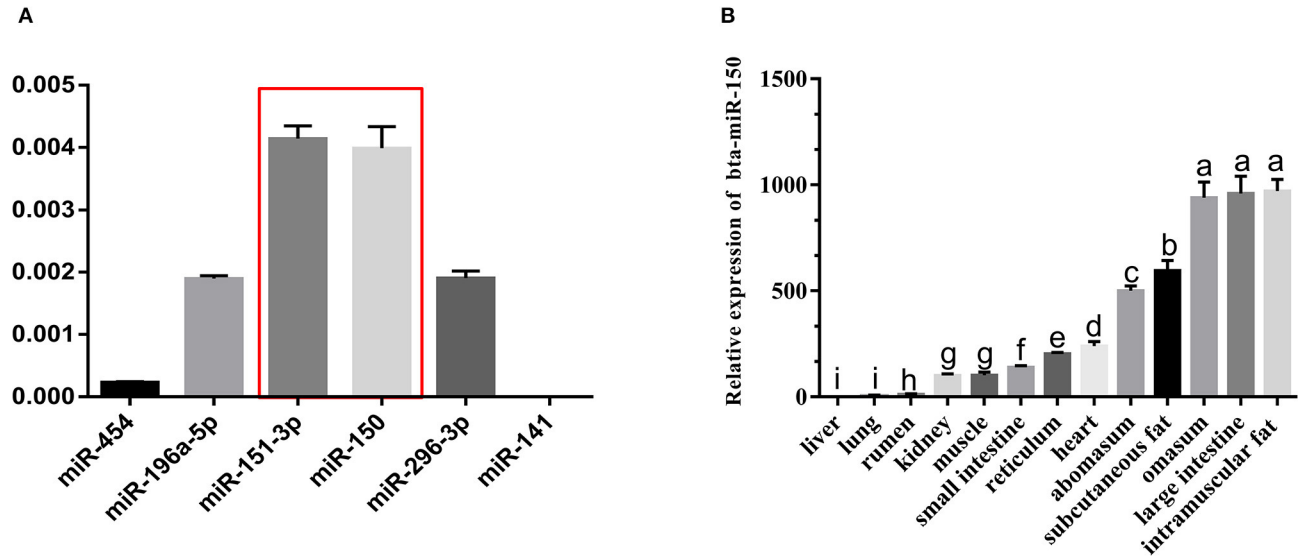


**FIGURE 1 |** miR-150 is expressed at high levels in bovine adipocytes. **(A)** Background expression of miR-150. **(B)** Tissue expression profile of miR-150. Values with superscript letters indicate mean significant difference ($p < 0.05$).

## Annotation of Gene Modules and KEGG Analysis

In order to explore the biological function of gene modules, David and KOBAS online tools were used to analyze the gene go and KEGG enrichment to describe the function of the modules and identify the relationship between these modules. The *p*-value was adjusted by the Benjamin Hochberg method.

## Statistical Analysis

Data are expressed as mean ± SD. Statistical analysis was performed using SPSS 19.0 software (SPSS, Chicago, IL, USA) and GraphPad Prism 6. One-way ANOVA and Dunnett's multiple comparison test were used to analyze *P*-values. Student's *t*-test was used for comparative analysis of the two groups, and $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$ were considered statistically significant.

# RESULTS

## miR-150 Is Up-Regulated in the Adipose Tissue From Qinchuan Cattle

To initially screen for candidate miRNA that regulate the development of bovine adipose tissue, we selected 52 differentially expressed (DE) miRNA ($\log_2$(FC) > 1, FDR < 0.05) based on our previous sequencing of transcriptomes from steers and bulls (GEO:GSE75063). Among these, 16 DE miRNA with sequence homology to human miRNA and their target DE genes (DEG) were assessed using Ingenuity Pathway Analysis (IPA) to construct a gene interaction network. In total, nine of these DE miRNA and 42 target DEG were mapped to a network that is involved in the differentiation of fat cells and fat tissue metabolism (Zhang et al., 2017). We tested background levels of miRNA with unknown or incomplete functional annotation by qPCR and found that background levels of miR-150 and miR-151-3p were higher than those of other miRNA in fatty tissues (**Figure 1A**), which was consistent with previous (Zhang et al., 2017). We therefore speculated that miR-150 is involved in the development and regulation of bovine adipose tissue (**Figure 1B**).

## Bovine Pre-adipocyte Proliferation Is Promoted by miR-150

We transfected bovine pre-adipocytes with a miR-150 mimic, mimic NC, miR-150 inhibitor, and inhibitor NC to further explore the effects of miR-150 on the proliferation of bovine pre-adipocytes. We initially stained cells with EdU. At 48 h after transfection, the rate of EdU-positive cells was significantly higher among pre-adipocytes incubated with the bta-miR-150 mimic than with mimic NC ($p < 0.05$), and the rate of EdU-positive cells was significantly lower in cells transfected with bta-miR-150 inhibitor than with inhibitor NC ($p < 0.05$) (**Figures 2A,B**). The results of CCK-8 assays were consistent with these findings (**Figure 2C**).

We then assessed changes in expression levels of proliferation marker genes at the mRNA and protein levels. Bovine pre-adipocytes were incubated for 48 h with mimic and inhibitor miR-150, and then, total RNA was extracted. The expression levels of genes associated with the cell cycle, such as *CDK1* (cyclin dependent kinase1), *CDK2* (cyclin dependent kinase2), and *PCNA* (proliferating cell nuclear antigen) significantly increased in cells incubated with the mimic and significantly decreased in those incubated with the inhibitor (**Figures 2D1–D3**) compared with that in the negative control. Western blotting showed

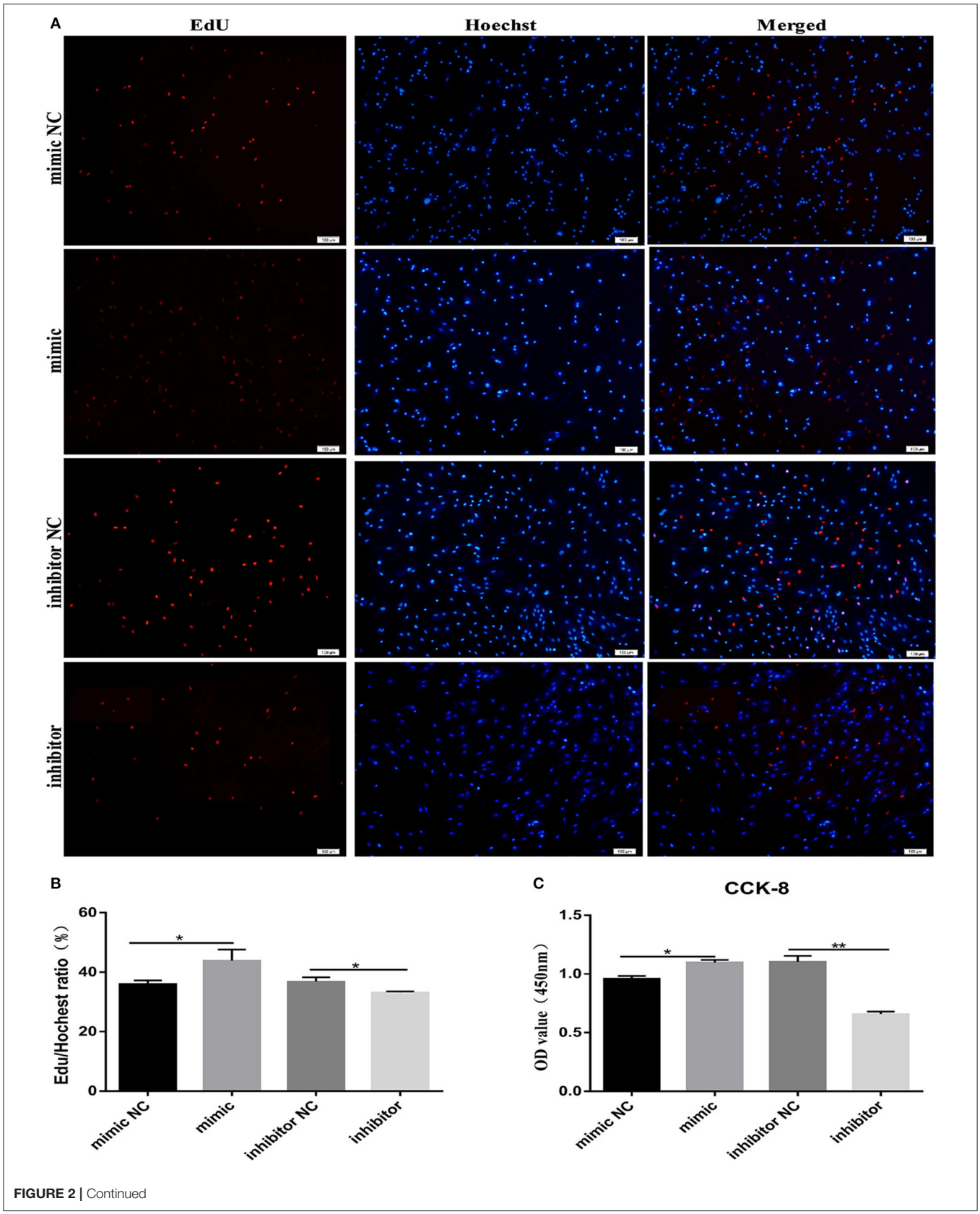


**FIGURE 2 |** Continued

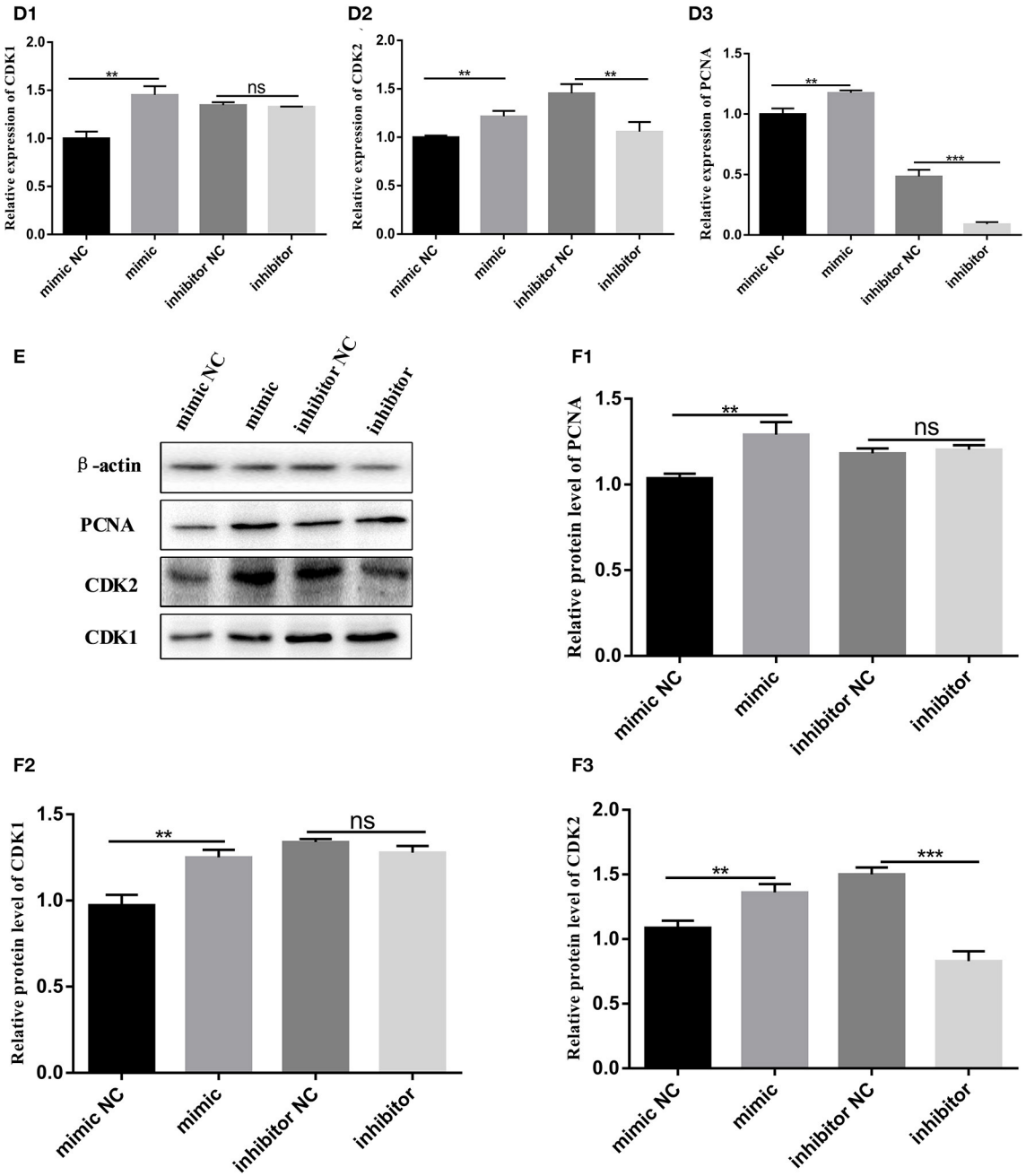


**FIGURE 2 |** Proliferation of Qinchuan bovine pre-adipocytes is promoted by bta-miR-150. **(A,B)** Bovine adipocytes transfected with mimic, mimic NC, inhibitor, and inhibitor NC. Cells harvested at 48 h after transfection were stained red with EdU (to indicate DNA replication) and blue with Hoechst (cell nuclei) **(C)** CCK-8 assays of bovine adipocytes. After incubating with 10% CCK-8 for 4 h, absorbance was measured at a wavelength of 450 nm. **(D1–D3)** Changes in mRNA levels of proliferation-related target genes according to qPCR findings. **(E, F1–F3)** Proliferation-related protein quantitation according to western blot findings. Error bars represent mean ± SD. $N = 3$ replicates. *Significantly different; $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$, Student *t*-tests.

that proliferating pre-adipocytes incubated with mimic miR-150 had significantly increased protein levels of CDK1, CDK2, and PCNA, while those incubated with miR-150 inhibitor had significantly reduced protein levels than the negative control (**Figures 2E,F1–F3**). These findings suggest that miR-150 promotes bovine pre-adipocyte proliferation.

## Inhibition of miR150 Promotes Adipogenesis

Total RNA extracted from bovine normal adipocyte cells at 0, 2, 4, 6, 8, and 10 days of induced differentiation *in vitro*, was reverse transcribed into cDNA. Quantitative PCR findings showed that miR-150 mRNA reached the highest level on day 2 of induced

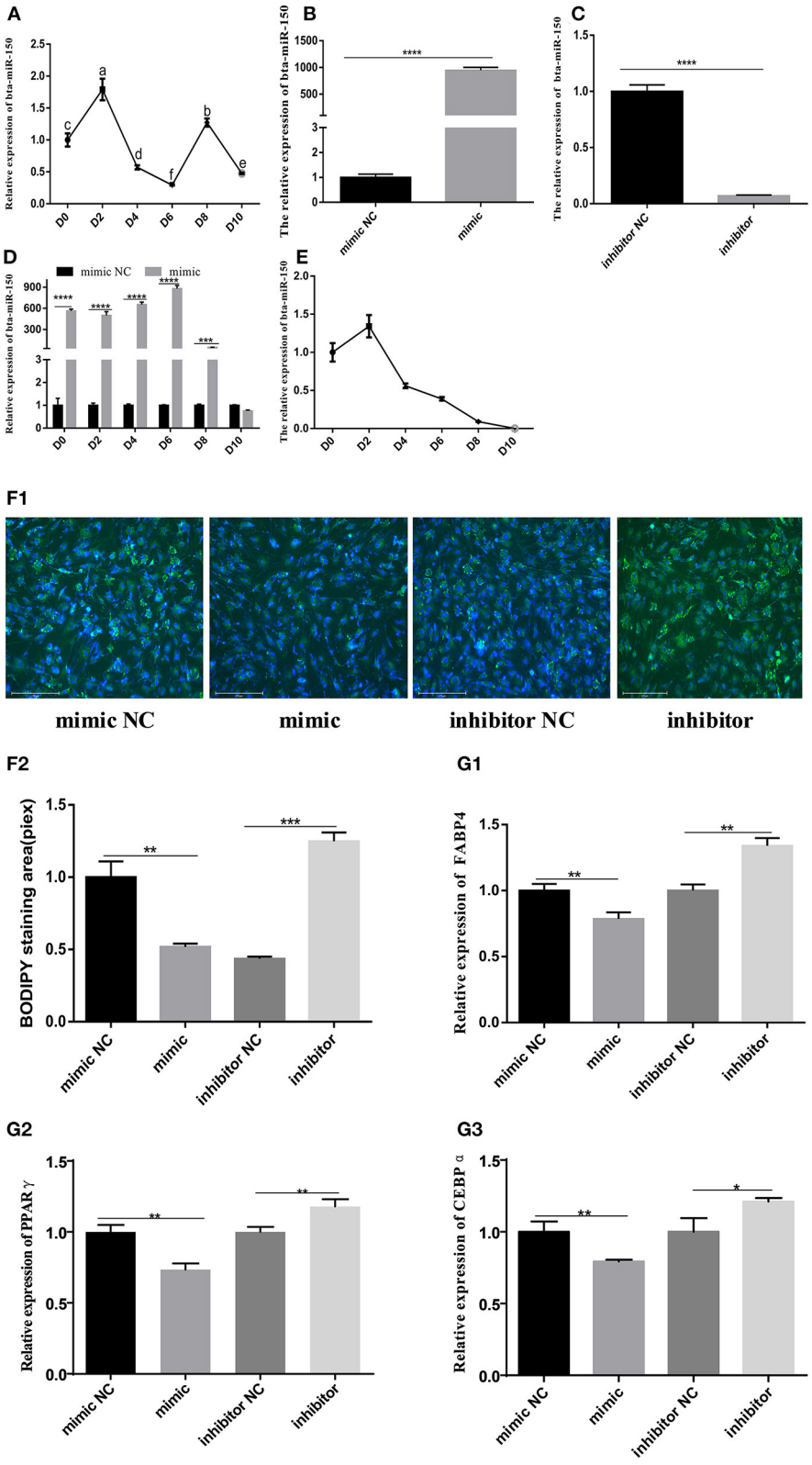


**FIGURE 3 |** Continued

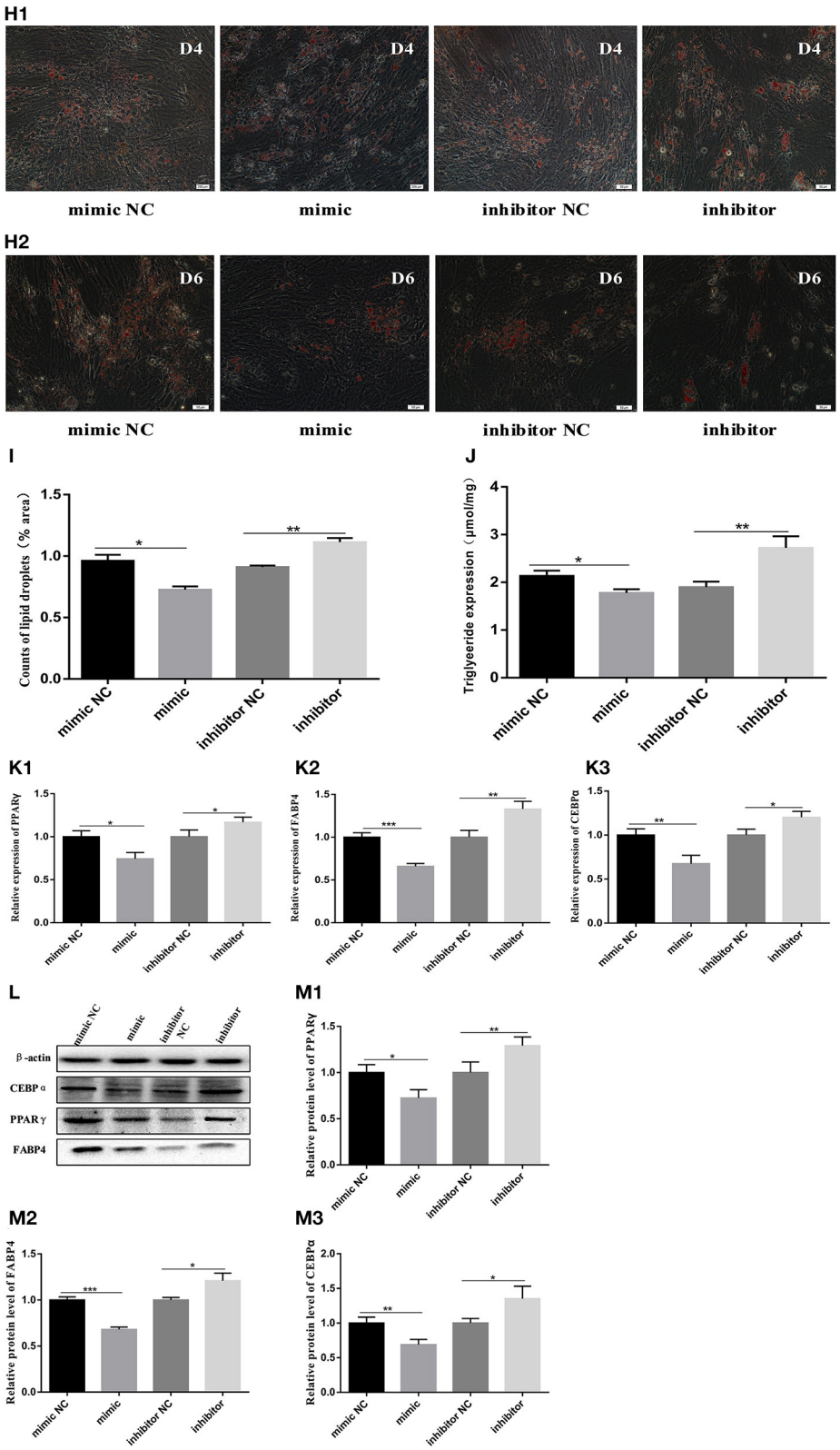


**FIGURE 3 |** Differentiation of pre-adipocytes from Qinchuan cattle. **(A)** Background expression of bta-miR-150 on different days after inducting differentiation in Qinchuan cattle adipocytes. **(B)** bta-miR-150 mimic transfection efficiency. **(C)** bta-miR-150 inhibitor transfection efficiency. **(D)** Overexpression of bta-miR-150

*(Continued)*

**FIGURE 3 |** treated with miR-150 mimic in 10 days. **(E)** Changes of bta-miR-150 content in Qinchuan cattle adipocytes treated with miR-150 mimics. **(F1,F2)** After transfection of miR-150 mimic, miR-150 inhibitor, or negative control, the adipocytes were stained with Bodipy and quantified on the second day of differentiation; photos were taken with an Evos-fl-auto2 automatic living cell fluorescence microscopy imaging system (Thermo Fisher, USA). **(G1–G3)** Changes in mRNA levels of adipose differentiation marker genes on the second day of differentiation (FABP4, PPARγ, and CEBPα). **(H1)** Oil red O staining of mimics, inhibitors, and their respective internal controls on day 4; the photos were taken with an Olympus Image IX71 microscope (Olympus, Japan). **(H2)** bta-miR-150 mimic treatment on day 6, Oil red O staining of substances, inhibitors, and their respective internal controls; the photos were taken with an Olympus Image IX71 microscope (Olympus, Japan). **(I)** Statistics of lipid droplet content on day 4 of differentiation, using Image J software. **(J)** Triglyceride content on day 4 of differentiation. **(K1–K3)** Changes in mRNA levels of fat differentiation marker genes (FABP4, PPARγ, CEBPα) on the 4th day. **(L)** Protein map of fat differentiation marker genes (FABP4, PPARγ, CEBPα). **(M1–M3)** Quantitative processing of fat differentiation marker gene protein results (FABP4, PPARγ, CEBPα); protein levels were quantified and analyzed by Image J software. Error bars represent mean ± SD. $n = 3$ replicates. *denotes significance according to Student's *t*-test, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$, and $^{****}p < 0.0001$.

differentiation *in vitro*, decreased to the lowest level on day 6, started to increase briefly on day 8, and then began to decline again (**Figure 3A**). Cells at ∼80% density were transfected, and transfection efficiency was measured 48 h later (**Figures 3B,C**). The effect of the bta-miR-150 mimic was prevalent until day 8 after transfection, but the bta-miR-150 content briefly reached a maximum on day 2 and then continuously fell (**Figures 3D,E**). We directly assessed the effects of miR-150 on adipose differentiation by staining lipid droplets using BODIPY and oil red O (**Figures 3F1,F2,H1,H2**). Compared with the control group, the number of lipid droplets in bta-miR-150 mimic-treated adipocytes decreased, while the number of lipid droplets significantly increased in adipocytes incubated with the inhibitor compared with the control. Overexpressed bta-miR-150 significantly inhibited the TG production, which increased when bta-miR-150 was inhibited (**Figure 3J**). The mRNA and protein levels of peroxisome proliferator-activated receptor gamma (*PPARγ*), fatty acid binding protein 4 (*FABP4)*, and CCAAT enhancer binding protein alpha (*CEBPα*) were higher with miR-150 inhibition compared with the levels with miR-150 overexpression NC (**Figures 3G1–G3,K1–K3**). Inhibition of bta-miR-150 therefore promotes adipogenesis.

## GO and KEGG Analysis of Bovine Adipocyte DEG

We further clarified the regulatory effect of miR-150 on the differentiation of bovine pre-adipocytes using RNA-seq. The miR-150 mimic, mimic NC, inhibitor, and inhibitor NC were transfected into bovine pre-adipocytes as described above, and cells were collected on days 2 and 6 after differentiation. The cells were then incubated daily with miR-150 mimic, mimic NC, inhibitor, and inhibitor NC. Transcriptomes from eight groups ($n = 3$ replicates per group) were sequenced (Beijing Novo Gene, Beijing, China) (**Figure 4A**). The RNA-seq results of cells transfected with miR-150 mimic, mimic NC, inhibitor, and inhibitor NC of D2, in which a total of 7,348 DEGs were detected, showed there were 1,234 common DEGs in four groups (**Figure 4B**). We identified the DEG in each group, then assessed them with the KEGG pathway and GO function enrichment analyses and WGCNA analysis (**Figures 4C–F**).

The soft threshold in the WGCNA analysis was set to 12, and 23 gene modules were detected (**Figure 4G**). By comparing different groups within the same modules, we identified five modules (green-yellow, light yellow, brown, blue, midnight blue). To better understand the biological function of specific gene modules related to the bovine adipocyte differentiation, we searched for DEG in these five modules using DAVID and KOBAS software using the key words, "fat fractionation" and "metabolism." We found enrichment for type II diabetes mellitus (ID: bta04930), the mTOR signaling pathway (bta04150), fatty acid degradation (ID: bta00071), and fatty acid metabolism (ID: bta001212) (**Figures 4H,I**). These pathways are all involved in the biological process of adipogenesis. We combined the RNA-seq data with the WGCNA analysis of KEGG findings and comprehensively integrated *P*-values, numbers of genes changed in the pathway, and the mTOR pathway as main research channels ($P < 0.05$).

## The mTOR Signaling Pathway Is Regulated by bta-miR-150 Through AKT1

We used targetscan7.1 to predict conserved target genes of miR-150 and the biological function of miR-150 during the development of bovine pre-adipocytes. We focused on the mTOR pathway based on the findings of the transcriptomic analysis (**Figure 5A**). Over 80% of the genes in the mTOR pathway were changed at the mRNA level by the miR-150 mimic or inhibitor (**Figures 5B,C**). Based on this and published information, we selected the adipocyte differentiation-related gene *AKT1* as a candidate target of bta-miR-150. We then constructed two luciferase reporter vectors for wild-type *AKT1* 3′UTR and mutant *AKT1* 3′UTR (**Figure 5D1**). Dual luciferase reporter assays revealed that *AKT1* is a target gene of bta-miR-150 (**Figure 5D2**). To further confirm the relationship between bta-miR-150 and *AKT1*, we measured *AKT1* expression during adipocyte differentiation at the mRNA and protein levels. The data indicated that bta-miR-150 promotes adipocyte differentiation by directly targeting *AKT1*. Western blot results showed that *AKT1* phosphorylation increased throughout differentiation up to a maximum of 1.0-fold after 48 h, compared with that at 0 h (**Figures 5E,F**). Taken together, these results implicate *AKT1* in fat differentiation in cattle and that *AKT1* is a target of bta-miR-150.

## si-AKT1 Can Inhibit Adipocyte Differentiation

We used si-*AKT1* interference to determine the role of *AKT1* in pre-adipocytes from Qinchuan beef cattle. We determined the optimal siRNA concentration required to ensure high knockout efficiency before transfection (**Figures 6A-C**). We then showed that silencing the *AKT1* gene with 100 nM si-*AKT1* reduced

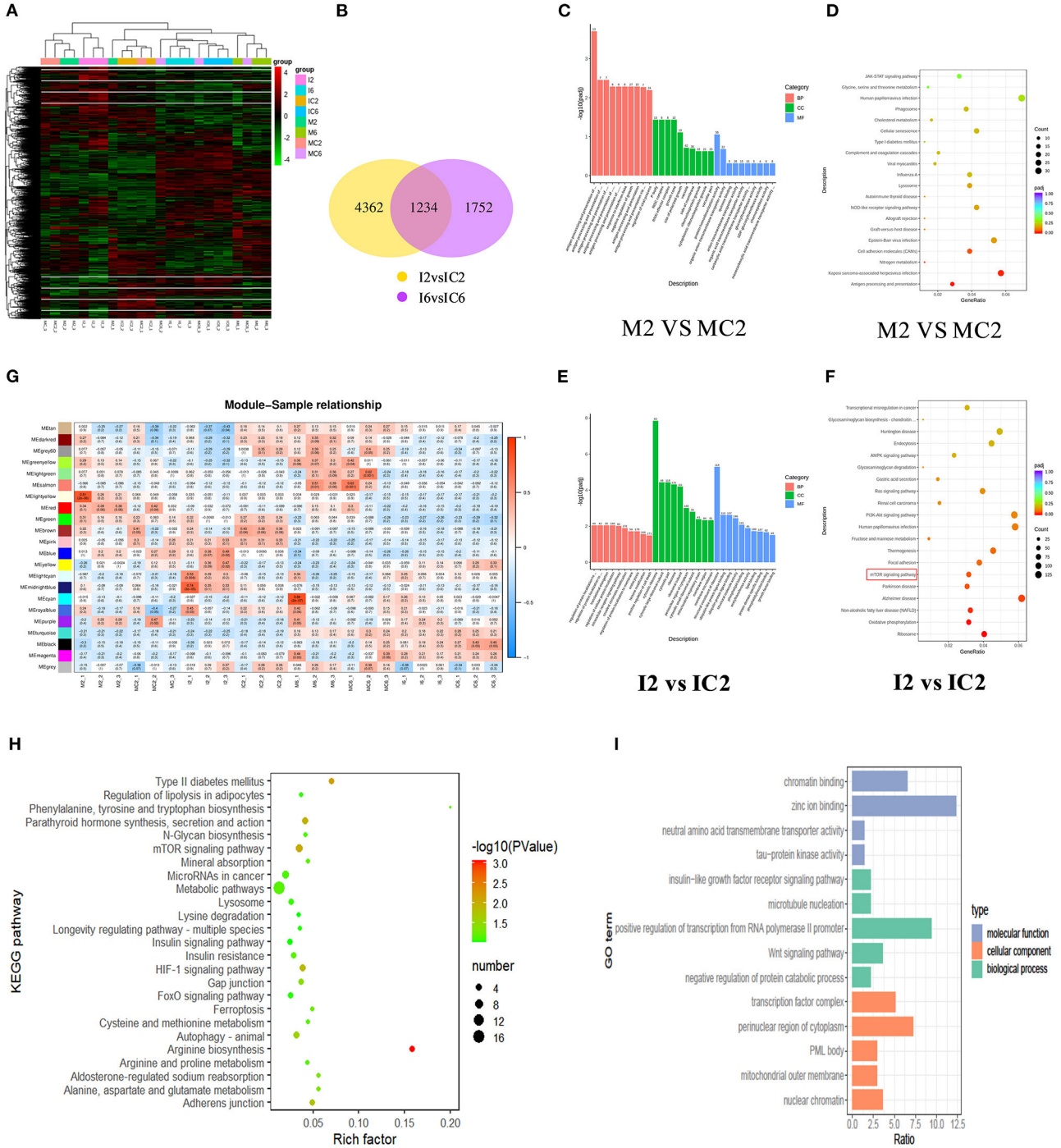


**FIGURE 4 |** GO and KEGG analysis of bovine adipocytes DEG and WGCNA analysis. **(A)** Differentially expressed gene clustering heat map. The abscissa is the sample name, and the ordinate is the normalized value of the differential gene FPKM. Red color indicates higher expression levels, while green represents lower expression levels; **(B)** The DEGs represented by a Venn diagram. Venn diagram showing the overlapping portion by sharing the expression pattern of the DEGs; **(C)** GO enrichment analysis of mimic treatment group and mimic NC group on the second day; **(D)** KEGG pathway enrichment analysis of mimic treatment group and mimic NC group on the second day;**(E)** GO enrichment analysis of inhibitor treatment group and inhibitor NC group on the second day; **(F)** KEGG pathway enrichment analysis of inhibitor treatment group and inhibitor NC group on the second day; **(G)** Correlation thermogram between samples and modules; **(H)** KEGG enrichment analysis of Midnightblue modules; **(I)** GO enrichment analysis of Midnightblue modules.

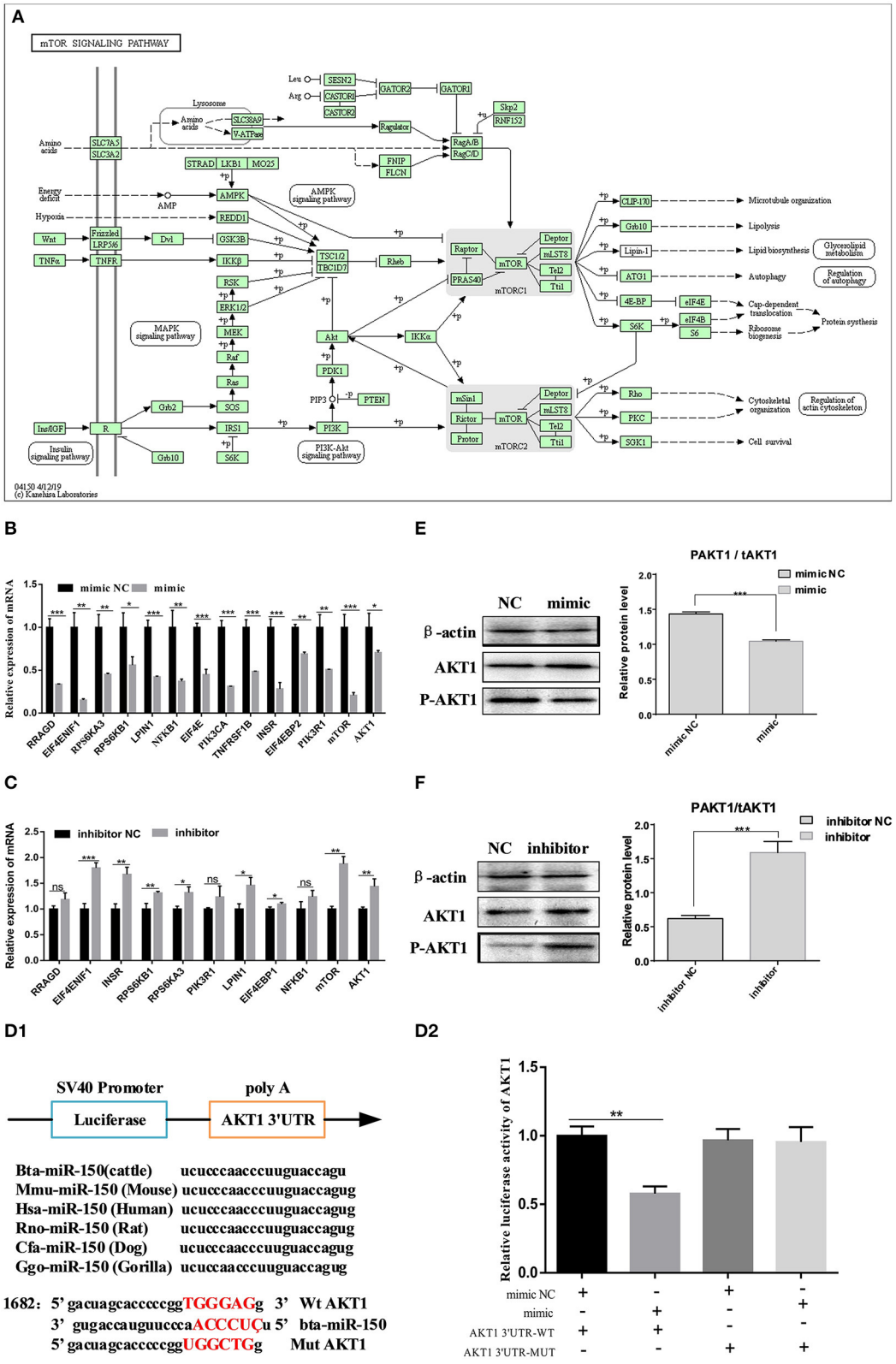


**FIGURE 5 |** bta-miR-150 directly targets AKT1 and regulates adipocyte differentiation through mTOR pathway. **(A)** mTOR signaling pathway. **(B,C)** Changes in mRNA levels in the mTOR pathway on day 2 after incubation with the mimic, inhibitor, and NC. **(D1)** Construction of a double luciferase reporter vector; **(D2)** Changes in luciferase activity after co-transfection. **(E,F)** Western blots of Akt1 phosphorylation on day 2 of differentiation. Error bars, means ± SD. $n = 3$ replicates. *denotes significance according to Student's *t*-test, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$, and $^{****}p < 0.0001$.

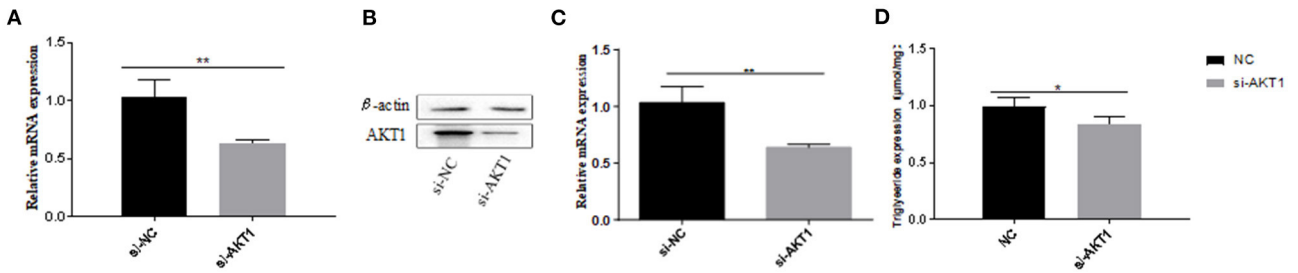


**FIGURE 6 |** Effects of AKT1 on the metabolism of adipocytes in Qinchuan cattle. **(A)** The efficiency of si-AKT1 transfection; **(B,C)** The expression levels of AKT1 mRNA and protein in pre-adipocytes transfected with si-AKT1 and negative control cells were compared; **(D)** After si-AKT1 treatment, the triglyceride level in pre-adipocytes decreased. Error bars represent mean ± SD. $n = 3$ replicates. *denotes significance according to Student's *t*-test, $^{*}p < 0.05$, $^{**}p < 0.01$, and $^{***}p < 0.001$.

triacylglycerol (TAG) levels in adipocyte precursors (**Figure 6D**), indicating that *AKT1* is a key transcription factor in pre-adipocytes from Qinchuan beef cattle.

## DISCUSSION

We stained induced adipocytes with BODIPY on day 2 after differentiation and with Oil Red O on days 4 and 6 after differentiation. The number of lipid droplets was lower in the cells incubated with bta-miR-150 mimic than with mimic NC and higher in those incubated with miR-150 inhibitor than with inhibitor NC. These phenotypes remained consistent on days 2, 4, and 6 of the induction of bovine adipocyte differentiation. Further experiments were based on this phenotype. The results confirmed that miR-150 functions in bovine fat formation, indicating that it is a candidate for the regulation of adipogenesis in cattle. At present, the precise role of miR-150 in the regulation of adipocyte differentiation remains unclear. Many previous studies used low-throughput methods, which cannot sufficiently explain the molecular regulatory mechanisms through which miR-150 affects the formation of pre-adipocytes. Therefore, we used high-throughput RNA-seq to identify DEG in adipocytes with overexpressed or inhibited miR-150. The results of KEGG enrichment analysis of the DEG in adipocytes incubated with inhibitor and inhibitor NC on day 2 implicated the mTOR pathway ($p = 0.0008$), because most genes in the pathway were changed due to miR-150 inhibition, and most of these DEG (such as *INSR* (Cignarelli et al., 2019), *IGF1R* (Hancock et al., 2019), *RPS6KA1* (Song and Richard, 2015), and *NFKB1* (Akbar et al., 2015) were related to "fat metabolism." To increase the richness of data analysis methods, we added the new WGCNA analysis (Langfelder and Horvath, 2008), which confirmed that among the multiple pathways enriched by DEG of five modules, the mTOR pathway was the most significant ($p = 0.0092$). Therefore, we verified that miR-150 mediates adipogenesis via the mTOR pathway. *AKT1* is part of the mTOR pathway and was implicated as a target of miR-150. Western blotting and qPCR findings confirmed that *AKT1* functions during adipogenesis.

Protein phosphorylation is a common and important post-translational modification (Humphrey et al., 2015). Approximately a third of cellular proteins are thought to be phosphorylated, and mTOR is a typical phosphorylation pathway (Rausch et al., 2019). Mammalian TOR plays an essential role in the regulation of cell growth and metabolism, such as responding to cell nutrients and growth signals, and it is associated with diabetes (Saxton and Sabatini, 2017). The seed sequence of miR-150 will be combined with the 3′ region sequence of *AKT1*. Qualcomm quantitative sequencing results have shown that when miR-150 is inhibited, insulin receptor substrates (IRS) are up-regulated, and receptor activation leads to the phosphorylation of key tyrosine residues on IRS proteins (Deng et al., 2020). After binding, the receptor activates p85 and recruits p10, which in turn catalyzes the generation of PI3P from PIP2 on the inner surface of the membrane. As the second messenger, PI3P further activates AKT and changes the phosphorylation level of *AKT* (Choi et al., 2010). *Akt* plays an important role in adipocyte differentiation (Xu and Liao, 2004). The "molecular switch" of BSTA can promote *AKT1* phosphorylation and promote adipocyte differentiation (Yao et al., 2013), whereas knocking out *Akt1* affects insulin-stimulated adipogenesis (Fischer-Posovszky et al., 2012). In addition, *Akt1*/*Akt2*-DKO mice have dyslipogenesis (Peng et al., 2003), and *AKT* deficiency in adipocytes can lead to severe lipodystrophy (Shearin et al., 2016). We found that miR-150 seems to inhibit bovine pre-adipocyte differentiation, which might inhibit the mTOR signaling pathway.

## CONCLUSIONS

Understanding the genetic basis of fat formation in beef cattle will support the breeding process to increase the intramuscular fat content of cattle. Our findings further validate and clarify the relationship between bta-miR-150 and bovine adipocyte differentiation; thus, overexpression of bta-miR-150 can inhibit adipocyte differentiation. Through RNA-seq sequencing data and WGCNA analysis methods, it was determined that the mTOR pathway, the signal pathway that plays a major role in bta-miR-150, affects bovine adipocytes' differentiation. This information

paves the way for further research to elucidate the complex adipocyte differentiation and fat tissue production process.

## DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: NCBI SRA (accession: PRJNA684031).

## ETHICS STATEMENT

All animal procedures for experiments were approved by Committee of Experimental Animal Management (EAMC) at Northwest A&F University, China (protocol number: NWAFUCAST2018–168). Moreover, the use of experimental animals was carried out in accordance to the rules and guidelines of the organization and government. Written informed consent was obtained from the owners for the participation of their animals in this study.

## AUTHOR CONTRIBUTIONS

XC, XM, SR, JW, XW, CL, and XY: conceptualization and data curation. XC, XM, SR, JW, CL, and XY: formal analysis. LZ: funding acquisition, project administration and supervision. XC, SR, CM, and XW: investigation. CL: methodology. JW and LZ: resources. XM, SR, and XY: software. XC, XM, CM, and XW: validation. XC: writing – original draft. SR: writing – review & editing. SS revised critically for content and grammar. All authors contributed to the article and approved the submitted version.

## FUNDING

This research was funded by the National Key Research and Development Program of China (2018YFD0501700), National Natural Science Foundation of China (31972994), Key Research and Development Program of Ningxia Province (2019BEF02004), the National Beef and Yak Industrial Technology System (CARS-37), and the China Postdoctoral Science Foundation (2020T130540).

## REFERENCES

Akbar, H., Grala, T. M., Riboni, M. V., Cardoso, F. C., Verkerk, G., McGowan, J., et al. (2015). Body condition score at calving affects systemic and hepatic transcriptome indicators of inflammation and nutrient metabolism in grazing dairy cows. *J. Dairy Sci.* 98, 1019–1032. doi: 10.3168/jds.2014-8584

Ambele, M. A., Dhanraj, P., Giles, R., and Pepper, M. S. (2020). Adipogenesis: a complex interplay of multiple molecular determinants and pathways. *Int. J. Mol. Sci.* 21:4283. doi: 10.3390/ijms21124283

An, X., Ma, K., Zhang, Z., Zhao, T., Zhang, X., Tang, B., et al. (2016). miR-17, miR-21, and miR-143 enhance adipogenic differentiation from porcine bone marrow-derived mesenchymal stem cells. *DNA Cell Biol.* 35, 410–416. doi: 10.1089/dna.2015.3182

Arcidiacono, B., Chiefari, E., Foryst-Ludwig, A., Curro, G., Navarra, G., Brunetti, F. S., et al. (2020). Obesity-related hypoxia via miR-128 decreases insulin-receptor expression in human and mouse adipose tissue promoting systemic insulin resistance. *EBioMedicine* 59:102912. doi: 10.1016/j.ebiom.2020.102912

Cai, R., Tang, G., Zhang, Q., Yong, W., Zhang, W., Xiao, J., et al. (2019). A novel lnc-RNA, named lnc-ORA, is identified by RNA-seq analysis, and its knockdown inhibits adipogenesis by regulating the PI3K/AKT/mTOR signaling pathway. *Cells* 8:477. doi: 10.3390/cells8050477

Choi, S. M., Tucker, D. F., Gross, D. N., Easton, R. M., DiPilato, L. M., Dean, A. S., et al. (2010). Insulin regulates adipocyte lipolysis via an Akt-independent signaling pathway. *Mol. Cell. Biol.* 30, 5009–5020. doi: 10.1128/MCB.00797-10

Chu, G., Zhou, X., Hu, Y., Shi, S., and Yang, G. (2019). Rev-erbα inhibits proliferation and promotes apoptosis of preadipocytes through the agonist GSK4112. *Int. J. Mol. Sci.* 20:4524. doi: 10.3390/ijms20184524

Cignarelli, A., Genchi, V. A., Perrini, S., Natalicchio, A., Laviola, L., and Giorgino, F. (2019). Insulin and insulin receptors in adipose tissue development. *Int. J. Mol. Sci.* 20:759. doi: 10.3390/ijms20030759

Deng, N., Guo, R., Zheng, B., Li, T., and Liu, R. H. (2020). IRS-1/PI3K/Akt pathway and miRNAs are involved in whole grain highland barley (*Hordeum vulgare* L.) ameliorating hyperglycemia of db/db mice. *Food Funct.* 11, 9535–9546. doi: 10.1039/D0FO01990A

Du, X., Hu, N., Yu, H., Hong, L., Ran, F., Huang, D., et al. (2020). miR-150 regulates endothelial progenitor cell differentiation via Akt and promotes thrombus resolution. *Stem Cell Res. Ther.* 11:354. doi: 10.1186/s13287-020-01871-9

Duarte, L., Garcia-Diaz, D. F., and Perez-Bravo, F. (2020). Body fat composition and miR-378 expression profiling in patients with type 1 diabetes. *Ann. Pediatr. Endocrinol. Metab.* 25, 118–125. doi: 10.6065/apem.1938088.044

Fischer-Posovszky, P., Tews, D., Horenburg, S., Debatin, K. M., and Wabitsch, M. (2012). Differential function of Akt1 and Akt2 in human adipocytes. *Mol. Cell. Endocrinol.* 358, 135–143. doi: 10.1016/j.mce.2012.03.018

Hancock, M. L., Meyer, R. C., Mistry, M., Khetani, R. S., Wagschal, A., Shin, T., et al. (2019). Insulin receptor associates with promoters genome-wide and regulates gene expression. *Cell* 177, 722–736. doi: 10.1016/j.cell.2019.02.030

Humphrey, S. J., James, D. E., and Mann, M. (2015). Protein phosphorylation: a major switch mechanism for metabolic regulation. *Trends Endocrinol. Metab.* 26, 676–687. doi: 10.1016/j.tem.2015.09.013

Junjvlieke, Z., Khan, R., Mei, C., Cheng, G., Wang, S., Raza, S. H. A., et al. (2020). Effect of ELOVL6 on the lipid metabolism of bovine adipocytes. *Genomics* 112, 2282–2290. doi: 10.1016/j.ygeno.2019.12.024

Kang, M., Liu, X., Fu, Y., and Timothy Garvey, W. (2018). Improved systemic metabolism and adipocyte biology in miR-150 knockout mice. *Metabolism* 83, 139–148. doi: 10.1016/j.metabol.2017.12.018

Katz, Y., Wang, E. T., Airoldi, E. M., and Burge, C. B. (2010). Analysis and design of RNA sequencing experiments for identifying isoform regulation. *Nat. Methods* 7, 1009–1015. doi: 10.1038/nmeth.1528

Khan, R., Raza, S. H. A., Junjvlieke, Z., Wang, X., Wang, H., Cheng, G., et al. (2020). Bta-miR-149-5p inhibits proliferation and differentiation of bovine adipocytes through targeting CRTCs at both transcriptional and posttranscriptional levels. *J. Cell. Physiol.* 235, 5796–5810. doi: 10.1002/jcp.29513

Khan, R., Raza, S. H. A., Junjvlieke, Z., Xiaoyu, W., Garcia, M., Elnour, I. E., et al. (2019). Function and transcriptional regulation of bovine TORC2 gene in adipocytes: roles of C/EBPγ, XBP1, INSM1 and ZNF263. *Int. J. Mol. Sci.* 20:4338. doi: 10.3390/ijms20184338

Kim, S. Y., Kim, A. Y., Lee, H. W., Son, Y. H., Lee, G. Y., Lee, J. W., et al. (2010). miR-27a is a negative regulator of adipocyte differentiation via suppressing PPARγ expression. *Biochem. Biophys. Res. Commun.* 392, 323–328. doi: 10.1016/j.bbrc.2010.01.012

Langfelder, P., and Horvath, S. (2008). WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* 9:559. doi: 10.1186/1471-2105-9-559

Liang, Y., Wang, Y., Ma, L., Zhong, Z., Yang, X., Tao, X., et al. (2019). Comparison of micro RNA s in adipose and muscle tissue from seven indigenous Chinese breeds and Yorkshire pigs. *Anim. Genet.* 50, 439–448. doi: 10.1111/age.12826

Lillehammer, M., Meuwissen, T. H., and Sonesson, A. K. (2011). A comparison of dairy cattle breeding designs that use genomic selection. *J. Dairy Sci.* 94, 493–500. doi: 10.3168/jds.2010-3518

Liu, Z., Deng, Y., Li, T., Zhu, F., Zhou, X., and He, Y. (2020). The opposite functions of miR-24 in the osteogenesis and adipogenesis of adipose-derived mesenchymal stem cells are mediated by the HOXB7/beta-catenin complex. *FASEB J.* 34, 9034–9050. doi: 10.1096/fj.202000006RR

Ma, L., Wang, R., Wang, H., Zhang, Y., and Zhao, Z. (2020). Long-term caloric restriction activates the myocardial SIRT1/AMPK/PGC-1α pathway in C57BL/6J male mice. *Food Nutr. Res.* 64. doi: 10.29219/fnr.v64.3668

Ma, X., Wei, D., Cheng, G., Li, S., Wang, L., Wang, Y., et al. (2018). Bta-miR-130a/b regulates preadipocyte differentiation by targeting PPARG and CYP2U1 in beef cattle. *Mol. Cell. Probes* 42, 10–17. doi: 10.1016/j.mcp.2018.10.002

Peng, X. D., Xu, P. Z., Chen, M. L., Hahn-Windgassen, A., Skeen, J., Jacobs, J., et al. (2003). Dwarfism, impaired skin development, skeletal muscle atrophy, delayed bone development, and impeded adipogenesis in mice lacking Akt1 and Akt2. *Genes Dev.* 17, 1352–1365. doi: 10.1101/gad.1089403

Poklukar, K., Candek-Potokar, M., Batorek Lukac, N., Tomazin, U., and Skrlep, M. (2020). Lipid deposition and metabolism in local and modern pig breeds: a review. *Animal* 10:424. doi: 10.3390/ani10030424

Rausch, S., Schollenberger, D., Hennenlotter, J., Stühler, V., Kruck, S., Stenzl, A., et al. (2019). mTOR and mTOR phosphorylation status in primary and metastatic renal cell carcinoma tissue: differential expression and clinical relevance. *J. Cancer Res. Clin. Oncol.* 145, 153–163. doi: 10.1007/s00432-018-2775-5

Raza, S. H. A., Kaster, N., Khan, R., Abdelnour, S. A., El-Hack, M. E. A., Khafaga, A. F., et al. (2020). The role of microRNAs in muscle tissue development in beef cattle. *Genes* 11:295. doi: 10.3390/genes11030295

Ren, L., Li, Q., Hu, X., Yang, Q., Du, M., Xing, Y., et al. (2020). A novel mechanism of bta-miR-210 in bovine early intramuscular adipogenesis. *Genes* 11:601. doi: 10.3390/genes11060601

Sanchez-Gurmaches, J., Martinez Calejman, C., Jung, S. M., Li, H., and Guertin, D. A. (2019). Brown fat organogenesis and maintenance requires AKT1 and AKT2. *Mol. Metab.* 23, 60–74. doi: 10.1016/j.molmet.2019.02.004

Saxton, R. A., and Sabatini, D. M. (2017). mTOR signaling in growth, metabolism, and disease. *Cell* 168, 960–976. doi: 10.1016/j.cell.2017.02.004

Scrutinio, D., Conserva, F., Passantino, A., Iacoviello, M., Lagioia, R., and Gesualdo, L. (2017). Circulating microRNA-150-5p as a novel biomarker for advanced heart failure: a genome-wide prospective study. *J. Hear. Lung Transpl.* 36, 616–624. doi: 10.1016/j.healun.2017.02.008

Shearin, A. L., Monks, B. R., Seale, P., and Birnbaum, M. J. (2016). Lack of AKT in adipocytes causes severe lipodystrophy. *Mol. Metab.* 5, 472–479. doi: 10.1016/j.molmet.2016.05.006

Song, J., and Richard, S. (2015). Sam68 regulates S6K1 alternative splicing during adipogenesis. *Mol. Cell. Biol.* 35, 1926–1939. doi: 10.1128/MCB.01488-14

Song, N. J., Chang, S. H., Kim, S., Panic, V., Jang, B. H., Yun, U. J., et al. (2018). PI3Ka-Akt1-mediated Prdm4 induction in adipose tissue increases energy expenditure, inhibits weight gain, and improves insulin resistance in diet-induced obese mice. *Cell Death Dis.* 9:876. doi: 10.1038/s41419-018-0904-3

Twayana, S., Legnini, I., Cesana, M., Cacchiarelli, D., Morlando, M., and Bozzoni, I. (2013). Biogenesis and function of non-coding RNAs in muscle differentiation and in Duchenne muscular dystrophy. *Biochem. Soc. Trans.* 41, 844–849. doi: 10.1042/BST20120353

Wang, L., Zhang, S., Cheng, G., Mei, C., Li, S., Zhang, W., et al. (2020). MiR-145 reduces the activity of PI3K/Akt and MAPK signaling pathways and inhibits adipogenesis in bovine preadipocytes. *Genomics* 112, 2688–2694. doi: 10.1016/j.ygeno.2020.02.020

Wang, Y., Mei, C., Su, X., Wang, H., and Zan, L. (2019). MEF2A regulates the MEG3–IO3 miRNA mega cluster-targeted PP2A signaling in bovine skeletal myoblast differentiation. *Int. J. Mol. Sci.* 20:2748. doi: 10.3390/ijms20112748

Xu, J., and Liao, K. (2004). Protein kinase B/AKT 1 plays a pivotal role in insulin-like growth factor-1 receptor signaling induced 3T3-L1 adipocyte differentiation. *J. Biol. Chem.* 279, 35914–35922. doi: 10.1074/jbc.M402297200

Yao, Y., Suraokar, M., Darnay, B. G., Hollier, B. G., Shaiken, T. E., Asano, T., et al. (2013). BSTA promotes mTORC2-mediated phosphorylation of Akt1 to suppress expression of FoxC2 and stimulate adipocyte differentiation. *Sci. Signal.* 6:ra2. doi: 10.1126/scisignal.2003295

Yugawa, K., Yoshizumi, T., Mano, Y., Itoh, S., Harada, N., Ikegami, T., et al. (2020). Cancer-associated fibroblasts promote hepatocellular carcinoma progression through downregulation of exosomal miR-150-3p. *Eur. J. Surg. Oncol.* doi: 10.1016/j.ejso.2020.08.002

Zhang, L., Wu, Z. Q., Wang, Y. J., Wang, M., and Yang, W. C. (2020). MiR-143 regulates milk fat synthesis by targeting smad3 in bovine mammary epithelial cells. *Animal* 10:1453. doi: 10.3390/ani10091453

Zhang, Y. Y., Wang, H. B., Wang, Y. N., Wang, H. C., Zhang, S., Hong, J. Y., et al. (2017). Transcriptome analysis of mRNA and microRNAs in intramuscular fat tissues of castrated and intact male Chinese Qinchuan cattle. *PLoS ONE* 12:e0185961. doi: 10.1371/journal.pone.0185961

Zhang, Z., Wang, W., Liu, J. B., Wang, Y., Hao, J. D., Huang, Y. J., et al. (2020). ssc-miR-204 regulates porcine preadipocyte differentiation and apoptosis by targeting TGFBR1 and TGFBR2. *J. Cell Biochem.* 121, 609–620. doi: 10.1002/jcb.29306

**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.

frontiers in Genetics

ORIGINAL RESEARCH
published: 14 June 2021
doi: 10.3389/fgene.2021.685341

# Identification of Regulatory Functions of LncRNAs Associated With *T. circumcincta* Infection in Adult Sheep

*Praveen Krishna Chitneedi[1], Rosemarie Weikard[1], Juan J. Arranz[2], María Martínez-Valladares[3,4], Christa Kuehn[1,5] and Beatriz Gutiérrez-Gil[2]**

*[1] Institute of Genome Biology, Leibniz Institute for Farm Animal Biology (FBN), Dummerstorf, Germany, [2] Departamento de Producción Animal, Facultad de Veterinaria, Universidad de León, León, Spain, [3] Departamento de Sanidad Animal, Facultad de Veterinaria, Universidad de León, León, Spain, [4] Instituto de Ganadería de Montaña, CSIC-Universidad de León, León, Spain, [5] Faculty of Agricultural and Environmental Sciences, University of Rostock, Rostock, Germany*

OPEN ACCESS

**Edited by:**
*Aline Silva Mello Cesar, University of São Paulo, Brazil*

**Reviewed by:**
*Rodrigo Baptista, University of Georgia, United States*
*Robert W. Li, United States Department of Agriculture, United States*

***Correspondence:**
*Beatriz Gutiérrez-Gil*
*beatriz.gutierrez@unileon.es*

**Specialty section:**
*This article was submitted to Livestock Genomics, a section of the journal Frontiers in Genetics*

**Received:** *25 March 2021*
**Accepted:** *20 May 2021*
**Published:** *14 June 2021*

**Citation:**
*Chitneedi PK, Weikard R, Arranz JJ, Martínez-Valladares M, Kuehn C and Gutiérrez-Gil B (2021) Identification of Regulatory Functions of LncRNAs Associated With T. circumcincta Infection in Adult Sheep. Front. Genet. 12:685341. doi: 10.3389/fgene.2021.685341*

Several recent studies have demonstrated the role of long non-coding RNAs (lncRNAs) in regulating the defense mechanism against parasite infections, but no studies are available that investigated their relevance for immune response to nematode infection in sheep. Thus, the aim of the current study was to (i) detect putative lncRNAs that are expressed in the abomasal lymph node of adult sheep after an experimental infection with the gastrointestinal nematode (GIN) *Teladorsagia circumcincta* and (ii) to elucidate their potential functional role associated with the differential host immune response. We hypothesized that putative lncRNAs differentially expressed (DE) between samples from animals that differ in resistance to infection may play a significant regulatory role in response to nematode infection in adult sheep. To obtain further support for our hypothesis, we performed co-expression and functional gene enrichment analyses with the differentially expressed lncRNAs (DE lncRNAs). In a conservative approach, we included for this predictive analysis only those lncRNAs that are confirmed and supported by documentation of expression in gastrointestinal tissues in the current sheep gene atlas. We identified 9,105 putative lncRNA transcripts corresponding to 7,124 gene loci. Of these, 457 were differentially expressed lncRNA loci (DELs) with 683 lncRNA transcripts. Based on a gene co-expression analysis via weighted gene co-expression network analysis, 12 gene network modules (GNMs) were found significantly correlated with at least one of 10 selected target DE lncRNAs. Based on the principle of "guilt-by-association," the DE genes from each of the three most significantly correlated GNMs were subjected to a gene enrichment analysis. The significant pathways associated with DE lncRNAs included ERK5 Signaling, SAPK/JNK Signaling, RhoGDI Signaling, EIF2 Signaling, Regulation of eIF4 and p70S6K Signaling and Oxidative Phosphorylation pathways. They belong to signaling pathway categories like Cellular

Growth, Proliferation and Development, Cellular Stress and Injury, Intracellular and Second Messenger Signaling and Apoptosis. Overall, this lncRNA study conducted in adult sheep after GIN infection provided first insights into the potential functional role of lncRNAs in the differential host response to nematode infection.

**Keywords: adult sheep, gastrointestinal infection, nematode, abomasal lymph node, long non-coding RNA, functional annotation, gene co-expression, pathways**

## INTRODUCTION

The central dogma of molecular biology states that DNA is transcribed into mRNA and mRNA is translated into protein products (Crick, 1970). There is an increasing interest, however, also in the non-protein coding RNAs (ncRNAs), which are estimated to comprise up to 80% of the entire transcriptome in mammals (The ENCODE Project Consortium, 2012; Hon et al., 2017; Srijyothi et al., 2018). The ncRNAs are classified into several sub-classes, such as small non-coding RNAs, including microRNAs and small interfering RNAs, processed small RNAs (Wilusz et al., 2009), transcription start site–associated RNAs (Seila et al., 2008), promoter associated RNAs (PARs) (Taft et al., 2009), enhancer RNAs (eRNAs) (Kim et al., 2015) and long non-coding RNAs (lncRNAs) (Wilusz et al., 2009). LncRNAs are defined as non-coding RNA transcripts with a length greater than 200 nucleotides. Although lncRNAs are biochemically resembling mRNAs, they generally do not encode protein products (Ponting et al., 2009). LncRNAs are divided into several subgroups according to the positional genomic relationship between lncRNAs and their neighboring protein-coding genes including overlapping, antisense, intronic and intergenic lncRNAs. Advances in computational biology and evolution of sensitive RNA sequencing and epigenomic technologies have facilitated the discovery of numerous lncRNAs and encouraged the study of their functional roles.

Based on their potential function, lncRNAs can be broadly classified into three main categories: (1) lncRNA transcripts, which are non-functional; (2) lncRNAs as regulators of transcription, which act as *cis*- and *trans*-active modulators of protein-coding gene expression (Kopp and Mendell, 2018), and (3) lncRNA transcripts involved in post-transcriptional regulation, which includes alternative mRNA splicing regulation, translational control and competing with regulatory endogenous RNAs (Riquelme et al., 2016; Dykes and Emanueli, 2017). It has been found that lncRNAs exhibit splice junctions and introns (Hiller et al., 2009) and cell- and tissue- specific expression patterns. The misexpression of lncRNAs has been shown to contribute to neurological disorders, cancer (Qureshi et al., 2010), susceptibility to infection, metabolic disorders such as diabetes and obesity (Moran et al., 2012; Zhao and Lin, 2015), and other diseases (Wapinski and Chang, 2011). In addition to their identification and cataloging, the functional annotation of the discovered lncRNAs is also challenging, as they are sparsely and cell type-specifically expressed, which might limit their function to few biological states (Derrien et al., 2012). Furthermore, at sequence level lncRNAs are poorly conserved across species (Ulitsky and Bartel, 2013; Hezroni et al., 2015).

The functional annotation of lncRNAs aims to identify or predict the possible biological process with which an lncRNA transcript can be interrelated, its putative mechanism of action, potential interacting partners and putative functional elements within the RNA locus (Cao et al., 2018). The annotation catalog of lncRNAs is far from complete even in humans and mice, and much work remains to be done in the field of lncRNAs annotation in the genomes of livestock species (Weikard et al., 2017). Currently, several lncRNA databases are available (Maracaja-Coutinho et al., 2019), which provide annotations of lncRNAs, mostly referring to the human genome. Among them, the NON-CODE database (Zhao et al., 2016) currently offers a comprehensive collection of lncRNA transcripts that have been experimentally confirmed in humans and 15 animal species (including the farm animals cow, pig, and chicken, but not sheep). With regard to the sheep genome, although there are many predicted protein coding genes known, only 30% of them are carrying an Ensembl identifier in the OAR v3.1 reference genome [Ensembl annotation based on adult Texel sheep (Clark et al., 2017)]. Most of the detected and characterized lncRNAs of the sheep genome have been annotated through the sheep atlas project based on the Texel breed (Clark et al., 2017).

Recently, another sheep reference genome Oar rambouillet v1.0 (GCA_002742125) has become available, providing a highly contiguous sheep genome with an annotation and mapping of transcription start sites (Salavati et al., 2020). However, this new reference genome contains a similar number of coding and non-coding genes compared to the OAR v3.1 reference genome. Currently, the Ensembl database contains 1,858 and 2,236 putative lncRNAs in the OAR v3.1 and the OAR rambouillet v1.0 genome annotation, respectively[1] (Oar rambouillet v1.0[2]).

LncRNAs are known to act as key regulators of the immune response by a variety of mechanisms (Chen et al., 2017; Menard et al., 2019). Some recent lncRNA studies conducted in humans (Rochet et al., 2019) and mice (Menard et al., 2018) have demonstrated the role of lncRNAs in regulating the defense mechanisms against parasite infections. However, up to now no sheep-specific studies are available that investigated the importance of lncRNAs for the immune response to nematode or parasite infection.

Gastrointestinal parasite infections in sheep are a major challenge to sheep husbandry due to great economic losses, impairment of animal welfare and difficulties in appropriate treatment and prophylaxis under the dominating extensive pasture system in sheep production (Roeber et al., 2013;

[1]http://www.ensembl.org/Ovis_aries_rambouillet/Info/Strains?db=core
[2]http://www.ensembl.org/Ovis_aries_rambouillet/Info/Index

Ruano et al., 2019). The aim of the current study was therefore to detect lncRNAs in the transcriptome of abomasal lymph node (ALN) tissue samples extracted from adult Spanish Churra sheep after an experimental infection with the gastrointestinal nematode (GIN) *Teladorsagia circumcincta* and to elucidate their potential functional role associated with the host immune response against GIN infection in adult sheep. The current study is based on the OAR v3.1 genome annotation, as it was the most widely used annotation and was also used as reference annotation in our previous RNA-sequencing (RNA-seq) studies (Chitneedi et al., 2018, 2020), which were conducted with the same samples as the current study.

We hypothesized that putative lncRNAs differentially expressed between samples from animals that differ in infection resistance may play a significant regulatory role in response to nematode infection in adult sheep. To validate our hypothesis, we performed co-expression and functional gene enrichment analyses with selected lncRNAs that (i) were differentially expressed in our study in response to GIN infection, (ii) are included in the current sheep gene atlas, and (iii) showed expression in gastrointestinal and lymph node tissue (Clark et al., 2017; Bush et al., 2018). Thus, this work will contribute to the aim of the global network for Functional Annotation of Animal Genomes (FAANG) to provide high quality functional annotations of farm animal genomes (Andersson et al., 2015).

# MATERIALS AND METHODS

## Experimental Design and RNA Sequencing

The experiment and the criteria for animal selection have been described earlier in detail in Chitneedi et al. (2018). The whole experiment was carried out according to the current National Spanish legislation on the protection of animals used in experimentation (Royal Decree 53/2013) and the approval of the competent body of the regional government, Junta de Castilla y León (ULE_024_2015). Initially, a commercial flock of the Churra dairy sheep breed including 119 dairy ewes, raised under a commercial, semi-intensive management system representing natural GIN infection conditions was sampled to measure the fecal egg count (FEC). Based on the distribution of FEC values in the selected flock [phenotypic values: average = 25.59 eggs per gram (epg); *SD* = 60.31 epg], we preselected a group of 10 animals showing extremely low FEC values (range 0–15 epg) and a group of eight animals showing very high FEC values (range 90–225 epg). These animals were transferred to the experimental farm of the Mountain Livestock Institute (IGM, León, Spain), where they were exposed to a first standardized experimental infection (EI1) with *T. circumcincta* third stage larvae (L3) after an antihelmintic treatment (oral dose of ivermectin). *T. circumcincta* is one of the most common GIN parasites that infects different sheep breeds, in particular with a high incidence in Churra sheep (Diez-Banos et al., 1992; Castilla-Gomez De Aguero et al., 2020). Based on the accumulative FEC at days 14–31 after the first experimental infection (EI1), six ewes were classified as 'susceptible' (SUS, range: 2,310–9,666 epg; average: 5,594 ± 2,661 epg) and six ewes as 'resistant' (RES, range: 0–915 epg; average: 308 ± 338 epg). After a second experimental infection (EI2), these 12 animals were sacrificed at day 7 after infection, and abomasal lymph node (ALN) tissue samples were collected. Total RNA extracted from these samples was used for library preparation using the KAPA Stranded mRNA-Sequencing Kit that starts with an oligo dT selection step (Roche). RNA sequencing (RNA-seq) using an Illumina Hi-Seq 2000 platform generated 'paired-end' reads of 75 bp, with a sequencing depth of 30 M fragments per sample. The FASTQ format data from the 12 ALN samples were subjected to the subsequent bioinformatics workflow summarized in **Supplementary Figure 1**.

## Alignment and Transcript Assembly

The script including the specific parameters used for different tools included in the bioinformatic workflow is provided in **Supplementary File 1**. Initially, potential adapter sequences were removed from RNA reads with *cutadapt_v1.18* (Martin, 2011) followed by trimming of poor-quality sequences from reads using *quality trim v1.6.0* (Robinson, 2015) to improve subsequent read alignment. After trimming, the reads were aligned against the *Ovis aries* reference genome Oar_v3.1 (GCA_000298735.1) with Ensembl annotation release 95[3] using the alignment tool *HISAT2 v2.1.0* (Pertea et al., 2015). The aligned unsorted reads were sorted using *samtools v1.9* (Li, 2011). The aligned and sorted read data were assembled into transcripts using *StringTie v1.3.5* (Pertea et al., 2015) and the *O. aries* reference genome Oar_v3.1 (GCA_000298735.1) with Ensembl annotation release 95 (see Text Footnote 3) for a reference-guided transcriptome assembly approach (Pertea et al., 2015). This concept of reference-guided transcriptome assembly, which depends on a genome assembly of the target organism, takes benefit from existing information about transcribed genome elements, but enables detection of previously unknown transcripts. This is of particular importance for livestock lncRNA investigation, because the catalog of lncRNAs is by far not complete up to now. *StringTie* starts by building clusters from reads aligned to the genome and then creates a splice graph to identify transcripts (Pertea et al., 2015). In addition to tagging the unannotated genes, *StringTie* also tags those genes provided in the reference genome annotation. After assembling the reads from all 12 samples individually, the full set of transcript assemblies was passed to *StringTie*'s merge function, which merges the gene structures found in any one of the samples. As in some samples only partially covered transcripts were assembled in the initial *StringTie* run, the merging step created a set of transcripts that is consistent across all samples, so that the transcripts can be compared in the subsequent steps (Pertea et al., 2015). Additionally, the merged annotation file was compared with the reference annotation Ensembl release 95 (see Text Footnote 3) of the OAR_v3.1 genome using *gffcompare v-0.10.6* (Pertea and Pertea, 2020) to identify novel transcripts and to determine how many assembled transcripts were fully or partially consistent with annotated genes.

[3]http://ftp.ensembl.org/pub/release-95/gtf/ovis_aries/

## LncRNA Detection

We used the lncRNA prediction tool *FEELnc* (Wucher et al., 2017) to indicate lncRNAs in the newly established merged annotation file of the ALN transcriptome in our sheep samples. The selection of this software was based on a previous study of our research group comparing different bioinformatic tools for lncRNA prediction (Weikard et al., 2018). The *FEELnc* analysis includes three steps: (i) filtering of candidate lncRNA transcripts; (ii) exploring coding potential and nucleotide composition of candidate lncRNA transcripts; and (iii) classification of the finally predicted lncRNA transcripts. In the first step, protein-coding genes, monoexonic transcripts (except for those in antisense direction to a coding gene) and transcripts with a size less than 200 nt were excluded, and a file was generated comprising potential candidate lncRNA transcripts. In the second step, the coding potential score was determined for each of the candidate lncRNA transcripts by considering absence of an open reading frame and *k*-mer nucleotide composition using the shuffle mode option of *FEELnc*. Finally, those putative lncRNAs with evidence for non-coding characteristics were classified with regard to their localization and their direction of transcription compared to adjacent RNA transcripts (protein-coding and non-coding loci). This classification of putative lncRNAs with respect to associated mRNAs or protein-coding genes (or other ncRNAs) might provide initial indication on potential lncRNA functions. In addition, to evaluate, if the putative lncRNA transcripts detected in the ALN sheep transcriptome were novel compared to the OAR_v3.1-r 95 annotation, the predicted lncRNAs were classified using the *gffcompare -r* option, and class code "u" assignment was taken as indication of a novel lncRNA.

## Read Count and Differential Gene Expression Analysis

The '*featureCounts*' option of the package *subread v 1.6.3* (Liao et al., 2013) was used with the newly generated merged annotation file to calculate read count matrices for all samples and all annotated loci. Prior to differential gene expression (DGE) analysis of the samples, we plotted the accumulated phenotypic FEC data of the animals and also performed an exploratory Principal Component Analysis (PCA) of the sample transcriptomes for the initial six resistant and six susceptible animals (**Supplementary Figure 2**). The phenotypic information available for sample ALN_14 (**Supplementary Figure 2A**) showed that the accumulated FEC value of the respective animal after EI1 was the lowest among the susceptible samples. Considering the first two principle components (PC1 and PC2) of the transcriptome expression data, after excluding the ALN_14 this sample originally allocated into the susceptible group could not be clearly assigned to this group (**Supplementary Figure 2B**). Thus, in order to perform the subsequent analysis on two groups with clearly distinct phenotypes (**Figures 1A,B**), sample ALN_14 was considered as an outlier and was excluded from the subsequent DGE analysis. This DGE analysis was performed with the *DESeq2* R package (Love et al., 2014) to identify differentially expressed (DE) loci (coding and non-coding) between the two animal groups differing in susceptibility to GIN infection. Loci were considered significantly DE at a FDR $< 0.05$ after Benjamini–Hochberg correction for multiple testing.

## Gene Co-expression Analysis

For a first indication on the potential functional role of lncRNA highlighted as DE in the differential expression analysis, we performed co-expression analyses. To this end, we applied the R package for Weighted gene co-expression network analysis (*WGCNA, v1.64*) (Langfelder and Horvath, 2008). This method was implemented using the parameters mentioned in Weikard et al. (2018). Briefly, WGCNA was performed initially with all the protein-coding and non-coding genes expressed in the 11 sheep ALN samples serving as expression data input except for those genes serving as target variables. We selected 10 lncRNAs as target variables for the WGCNA according to the following criteria: (i) they displayed significantly differential expression between the RES and SUS sheep in response to GIN infection, (ii) they had been reported previously as expressed in gastrointestinal and lymph node tissue in the sheep gene atlas study (Clark et al., 2017) and had been predicted as lncRNAs in the sheep lncRNA study linked to the sheep gene atlas study (Bush et al., 2018), (iii) finally, they displayed a clear exon-intron structure as verified by visual inspection of reads aligned to the selected DE lncRNA transcripts using the Interactive Genomics Viewer (IGV) (Thorvaldsdottir et al., 2013).

The co-expressed genes within the WGCNA were identified as gene network modules (GNMs) marked with different colors. We assumed that highly interconnected genes within a GNM are co-regulated and might be involved in similar biological pathways. Those GNMs that were highly significantly correlated with at least one DE lncRNA transcript ($r > |0.75|$ and $p < 0.01$) were selected for pathway enrichment analysis of the included DE genes to obtain indication on a potential functional role of the correlated DE lncRNA.

## Enrichment and Pathway Analysis

To predict the potential biological function of the DE lncRNAs and associated biological pathways, we performed the gene set enrichment pathway analysis using the Qiagen Ingenuity analysis package (IPA[4]) with the DE genes from those GNM that were correlated with a specific lncRNA. Only annotated protein-coding genes in significantly correlated GNMs were included in the IPA analysis, whereas functionally unannotated and non-protein-coding genes were excluded.

# RESULTS

## Alignment and Transcriptome Assembly

After trimming of adapters, primers and poor-quality sequences, the data from the 12 samples initially included in the transcriptomic study comprised 30 to 50 M fragments per sample, which were subjected to further analysis. An average of 89.8% of the reads were aligned against the OAR_v3.1

[4]https://digitalinsights.qiagen.com/products-overview/discovery-insights-portfolio/analysis-and-visualization/qiagen-ipa/

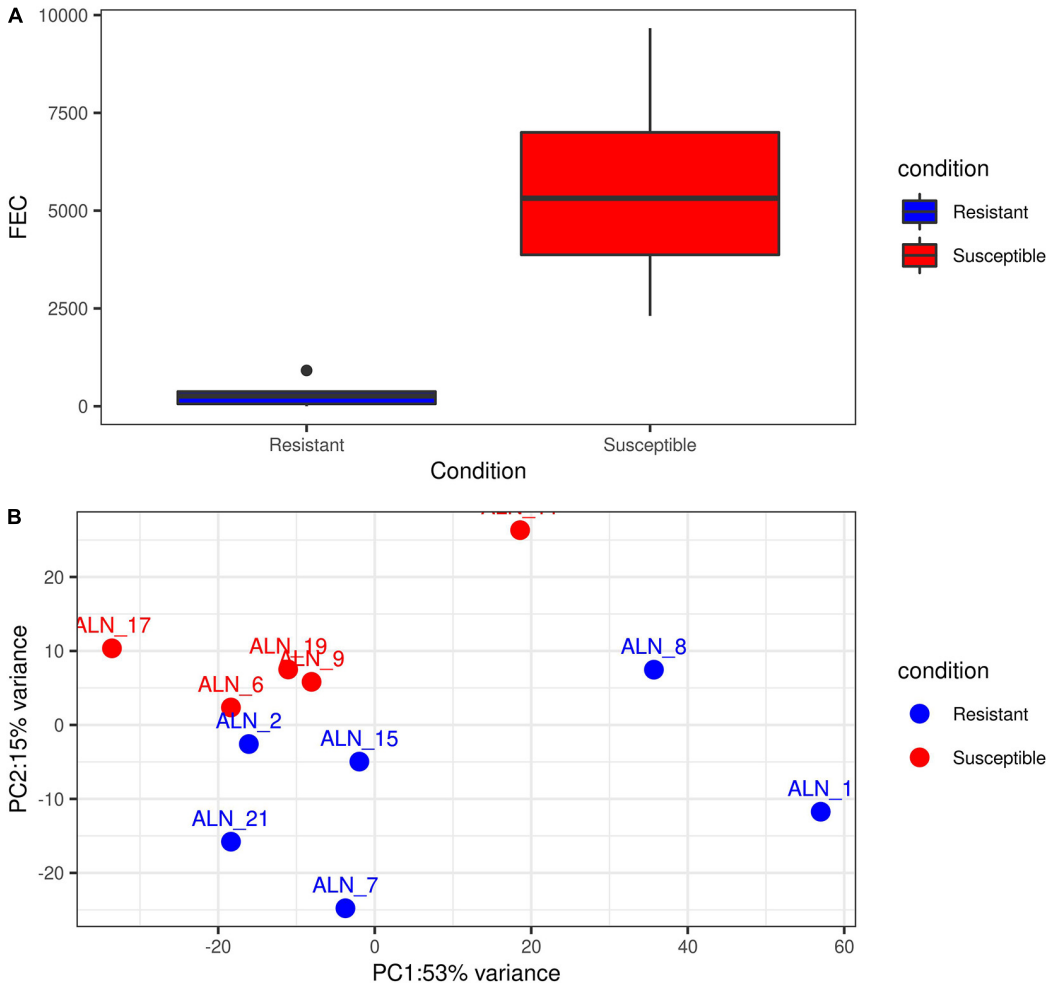


**FIGURE 1 |** Exploratory plots of GIN infected resistant and susceptible sheep. **(A)** Distribution of accumulated fecal egg count (FEC) in resistant and susceptible sheep after a second experimental infection with GIN *Teladorsagia circumcincta.* Blue box represents the group mean of the resistant group and red box represents the group mean of the susceptible group. **(B)** PCA performed with the read count matrix of the 11 ALN RNA-seq datasets finally included in the differential gene expression analysis specifying the sample condition [resistant (6) and susceptible (5)].

reference genome. Finally, after performing the reference-guided transcriptome assembly, the final new annotation file after merging information across all 12 samples comprised 77,039 transcripts. These transcripts corresponded to a total of 44,203 transcribed gene loci, which is an average of 1.74 isoform transcripts per gene locus.

## LncRNA Detection and Classification

After filtering out protein-coding transcripts, based on their coding potential score and nucleotide composition, the *FEELnc* program identified 9,105 lncRNA transcripts corresponding to 7,124 putative lncRNA loci. This corresponds to 1.28 lncRNA transcripts per lncRNA locus, which is lower than the average number of transcript isoforms across all gene loci. We observed an equal strand distribution of these lncRNA transcripts with 50.2% (3,580 lncRNA loci) identified on the plus strand and 49.7% (3,544 lncRNA loci) on the minus strand. The average number of exons per lncRNA locus was 3.3 (median 2) with a mean exon length of 432.2 nt (median149 nt). The average number of exons per non-lncRNA locus was 14.3 (median 2) with a mean exon length of 289.8 nt (median 130 nt).

To classify the identified lncRNA transcripts, we used the sliding window size of 10,000 to 100,000 nt to check for the possible overlap with the nearest transcript from the reference annotation. In total, we identified a total of 19,162 potential predicted spatial lncRNA interactions between 8,676 lncRNA transcripts (originating from 6,854 lncRNA loci) and other loci in the merged annotation. Out of these, 11,250 were intragenic interactions, called as 'genic' by the *FEELnc* classifier, and 7,912 were intergenic. No potential spatial interaction was predicted for the remaining 429 lncRNA transcripts (270 lncRNA loci). Some of the lncRNA transcripts interacted with both genic and intergenic regions of different partner transcripts. In total, 3,004 lncRNA transcripts were in sense direction with potentially interacting genic regions and 5,216 lncRNA transcripts in antisense direction. In case of intergenic regions, 3,143 lncRNA

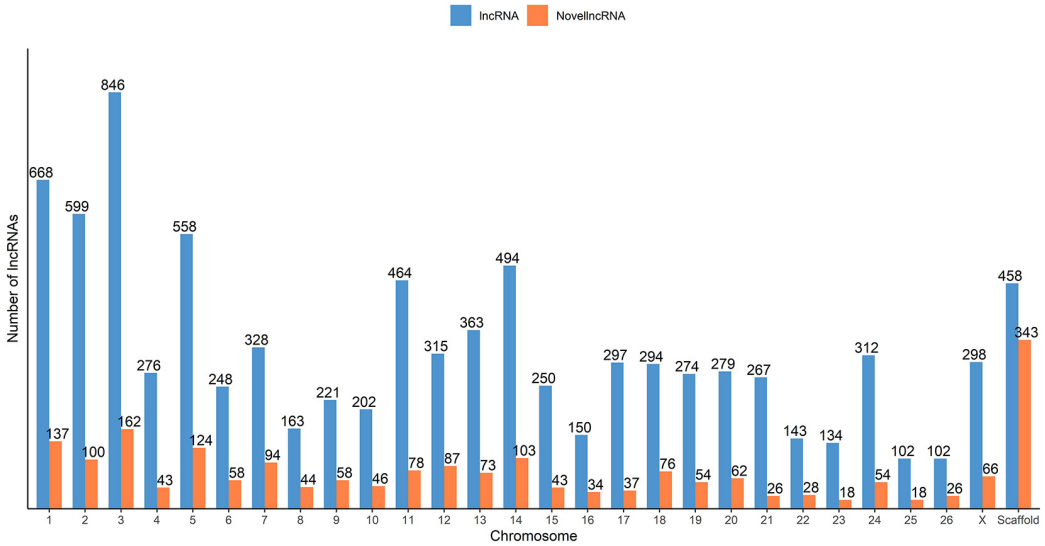


**FIGURE 2 |** Distribution of the predicted putative lncRNAs across the sheep genome from the merged ALN RNA-seq datasets of the 12 ALN sheep transcriptome samples under study.

transcripts were predicted in sense direction and 2,758 in antisense direction. The mean expression of the 7,124 lncRNA loci was 60.6 FPKM, with a median of 9.13 FPKM and in case of the rest 37,079 non-lncRNA gene loci, the mean expression was 12.9 FPKM (median 0.88). Of the total of 9,105 lncRNA transcripts found in our analysis (7,124 lncRNA loci), 2,092 lncRNA transcripts (1,393 lncRNA loci) were categorized as novel by gffcompare. The distribution of novel and known lncRNA transcripts across the ovine chromosomes is shown in **Figure 2**.

## Differential Gene Expression Analysis

After excluding the outlier sample (ALN_14) detected through the PCA and evaluation of phenotypic FEC values (**Figures 1A,B**), a total of 3,148 differentially expressed loci (DELs, protein-coding and non-coding) were detected at a significance threshold of $q < 0.05$ between the ALN transcriptomes of the RES and SUS sheep groups in response to GIN infection. Of these, 1,635 DELs were higher and 1,513 DELs were lower expressed after GIN infection in RES sheep compared to the SUS group. The list of DELs is shown in **Supplementary Table 1**. When investigating the 3,148 DELs, we found 683 lncRNA transcripts in 457 DELs. Of these, 263 lncRNA transcripts associated with 153 DELs were considered novel. The list of DE lncRNAs including the novel DE lncRNAs is provided in **Supplementary Table 1**.

## Gene Co-expression and Enrichment Analysis

The list of the 10 DE lncRNA loci selected for this analysis is shown in **Table 1** and **Figure 3** depicts results of the DE analysis for those 10 target DE lncRNA loci. The cluster dendrogram plot of the ALN transcriptome data from the RES and SUS sheep and the corresponding heatmap (**Figure 4**) show that the expression level of the 10 selected lncRNAs displayed a different pattern between the groups of differential resistance to GIN infection (**Figure 4**). Based on the module co-expression analysis with the 10 target lncRNAs, we found 88 GNMs (**Supplementary Figure 3**). Of them, 12 GNMs were significantly correlated ($p < 0.01$) with at least one of the 10 target lncRNAs included in the analysis (see **Figure 4**). Furthermore, we found that most lncRNAs were highly co-expressed with several GNMs at $r > |0.75|$ and $p < 0.01$.

For canonical biological pathway analysis, out of the 12 GNMs significantly correlated with at least one target lncRNA, we selected the DE genes from the three most highly significantly correlated GNMs: green, turquoise and yellow. Although other modules were also significantly co-expressed with the 10 selected lncRNAs, the genes from the other significantly correlated GNMs did not satisfy the criteria that were used to select the genes for enrichment pathway analysis. The blue and darkseagreen4 GNMs did not comprise DE genes and the other modules contained only a single or no protein-coding DE gene(s), which precluded biological pathway analysis. As shown in **Table 2**, Ingenuity pathway analysis revealed that the canonical pathways *ERK5 (Extracellular signal-related kinase) Signaling*, *SAPK/JNK (stress-activated protein kinases) Signaling*, *RhoGDI Signaling*, *EIF2 (eukaryotic translation initiation factors) Signaling*, *Regulation of eIF4 (Eukaryotic translation initiation factor 4) and p70S6K Signaling* and *Oxidative Phosphorylation* were significantly enriched by genes from GNMs green, turquoise and yellow. **Table 2** also contains the list of DE genes within the enriched pathways for the three GNMs significantly correlated to at least one target lncRNA.

The co-expression and IPA analyses revealed that the *ERK5 Signaling*, *SAPK/JNK Signaling* and *RhoGDI Signaling* canonical pathways were enriched by genes included in the GNMs 'green' and 'yellow' (**Table 2**), which were both negatively correlated

**TABLE 1 |** Differentially expressed (DE) lncRNAs selected for gene co-expression analysis.

| LncRNA | Chromosomal location (chr:start-end:strand, sense) | Overlapping gene | Classification of the identified lncRNA transcripts expressed in sheep abomasal lymph node relative to the overlapping gene[1] | Overlapping region from previous studies (chr:start-end:strand, sense) | Expression of overlapping gene regions from previous sheep studies[2] |
|---|---|---|---|---|---|
| MSTRG.2313 | 1:188445304–188449367:+ | *ENSOARG00000020259 (RPL35A)* | Intronic, antisense, class code - *S* | 1:188445366–188448530:– | Prescapular lymph node in TxBF adult[3] |
| MSTRG.9049 | 14:50414613–50422485: + | *RPS19* | Intronic, antisense, class code - *S* | 14:50414657–50422495: – | Prescapular lymph node in TxBF adult[3] |
| MSTRG.32373 | 9:36111002–36112522: + | *RPS20* | Intronic, antisense, class code - *S* | 9:36111034–36112627: – | Prescapular lymph node in TxBF adult[3] |
| MSTRG.1579 | 1:109258493–109270297: + | *ENSOARG00000025559 (lincRNA located near IGSF9 and TAGLN2)* | Complete, exact intronic match, class code -= | 1:109251911–109274557: + | Not calculated[4] |
| | 1:109264963–109270297: + | | Multi-exonic with at least one junction match, class code – *j* | 1:109267606–109271242: – | Abomasum in Texel lamb/6–10 months[3] |
| MSTRG.2319 | 1:188856696–188859141: – | *Located between MUC20 and TNK2* | Novel (unknown, intergenic), class code – *u* | 1:188856349–188859135: + | Hippocampus in TxBF adult[4] |
| | 1:188856697–188859104: – | | Novel (unknown, intergenic), class code – *u* | | |
| | 1:188856697–188859111: – | | Novel (unknown, intergenic), class code – *u* | | |
| | 1:188856697–188859118: – | | Novel (unknown, intergenic), class code – *u* | | |
| | 1:188856697–188859141: – | | Novel (unknown, intergenic), class code – *u* | | |
| | 1:188856701–188859137: – | | Novel (unknown, intergenic), class code – *u* | | |
| MSTRG.5010 | 11:39016413–39020308: + | *ENSOARG00000010029 (RPL23)* | Intronic, antisense, class code - *S* | 11:39016458–39020292: – | Omasum in TxBF lamb/new born[3] |
| MSTRG.4185 | 11:14061262–14065789: – | *ENSOARG00000004831 (CCL14)* | Intronic, antisense, class code - *S* | 11:14061328–14065335: + | Omentum in Texel adult[3] |
| MSTRG.22699 | 3:10979130–10983453: – | *ENSOARG00000013275 (RPL35)* | Intronic, antisense, class code - *S* | 3:10979613–10983240: + | Rumen in TxBF lamb/new born[3] |
| MSTRG.24044 | 3:133375792–133383230: – | *KRT8* | Intronic, antisense, class code - *S* | 3:133375727–133383019: + | Rectum in Texel adult[3] |
| MSTRG.32557 | 9:57536379–57541099: + | *FABP4* | Intronic, antisense, class code - *S* | 9:57536525–57541042: – | Omentum in Texel adult[3] |

[1] *The classification of transcripts is based on the annotation tool Gffcompare -r option.*
[2] *TxBF adult refers to adult Texel × Scottish Blackface sheep.*
[3] *Clark et al. (2017)*
[4] *Bush et al. (2018).*

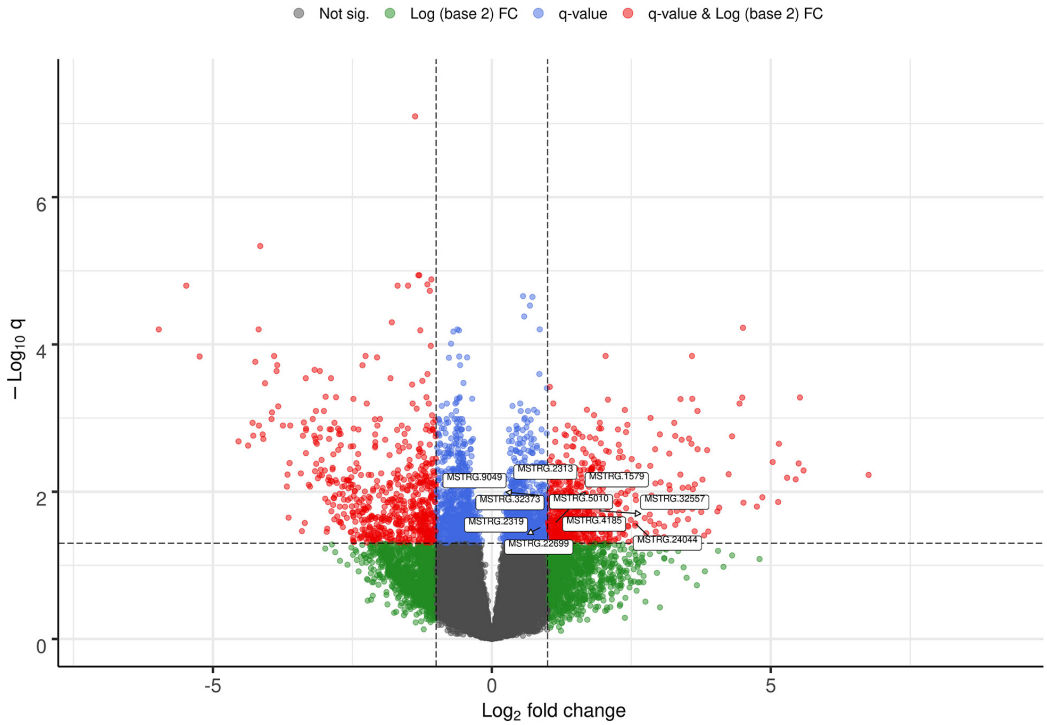


**FIGURE 3 |** Volcano plot showing the differentially expressed (DE) genes with the selected target lncRNA loci for co-expression analysis.

with expression levels of several DE lncRNAs (**Table 2** and **Figure 5**). In addition, genes from GNM 'turquoise' that showed a positive correlation with DE lncRNA expression levels (**Table 2** and **Figure 5**) were enriched in the canonical pathways *EIF2 Signaling* and *Regulation of eIF4 and p70S6K Signaling* (**Table 2**). It is striking that the same lncRNAs (MSTRG.1579, MSTRG.2313, MSTRG.2319, MSTRG.32373, MSTRG.32557, MSTRG.5010, MSTRG.9049, and MSTRG.4185) correlated with the GNMs 'green,' 'yellow,' and 'turquoise' indicating that they might be involved in the regulation of specific processes within the respective interconnected biological pathways, which display variation regarding resistance against GIN infection. An intersection between these pathways might be the mitogen-activated protein kinase (MAPK) pathway, which comprises the ERK, JNK and p38 mediated kinase cascades. It represents a conserved host-defense repertoire, and dysfunctions are known to result in hypersensitivity toward infection and stress (Johnson and Lapadat, 2002).

## DISCUSSION

LncRNAs are known to play multiple biological functions and their expression also varies with the developmental stage of cells and tissues and under different disease states and environmental challenges (Ma et al., 2012). Thus, by inferring the functional role of lncRNAs we can attempt to better understand the possible mechanisms of complex biological processes related to various metabolic disorders, disease conditions, divergent phenotypes and response to environmental challenges. However, compared to regular protein-coding genes, the sequences and secondary structures of lncRNA transcripts are usually not conserved (Mercer et al., 2009; Pang et al., 2009). This makes it difficult to investigate the function of lncRNAs directly based on their physical nucleotide structure and features. The availability of different bioinformatics tools such as *CNCI* (Sun et al., 2013), *PLEK* (Li et al., 2014), *PLAR* (Hezroni et al., 2015), and *FEELnc* (Wucher et al., 2017) has enabled to predict lncRNAs from unknown transcripts, but the detection of their biological functions is still challenging. It has been reported that lncRNAs exert their functions by regulating or interacting with other molecules like RNA, DNA and proteins (Ma et al., 2012; Schmitz et al., 2016). Thus, one possible approach to predict the function of lncRNAs is to explore the relationship between lncRNAs and other molecular interacting partners. One such approach is to identify the protein-coding genes that are co-expressed with the interrogated lncRNA. According to the principle of "guilt by association," the putative lncRNA function can be assigned via correlation to a group of co-expressed genes (in a module) with the biological pathway enriched by the group of co-expressed genes (Wolfe et al., 2005). This approach has been applied for a variety of lncRNA studies in other species (Petri et al., 2015; Li et al., 2016; Zhang et al., 2017; Weikard et al., 2018; Zheng et al., 2018; Ling et al., 2019; Nolte et al., 2019; Thiel et al., 2019; Wang et al., 2020).

In this context, the current study was performed to predict ovine lncRNA transcripts in the ALN transcriptome and to investigate the functional role of DE lncRNAs in respective samples obtained from adult sheep after an experimental GIN infection. This was carried out in several steps: In the

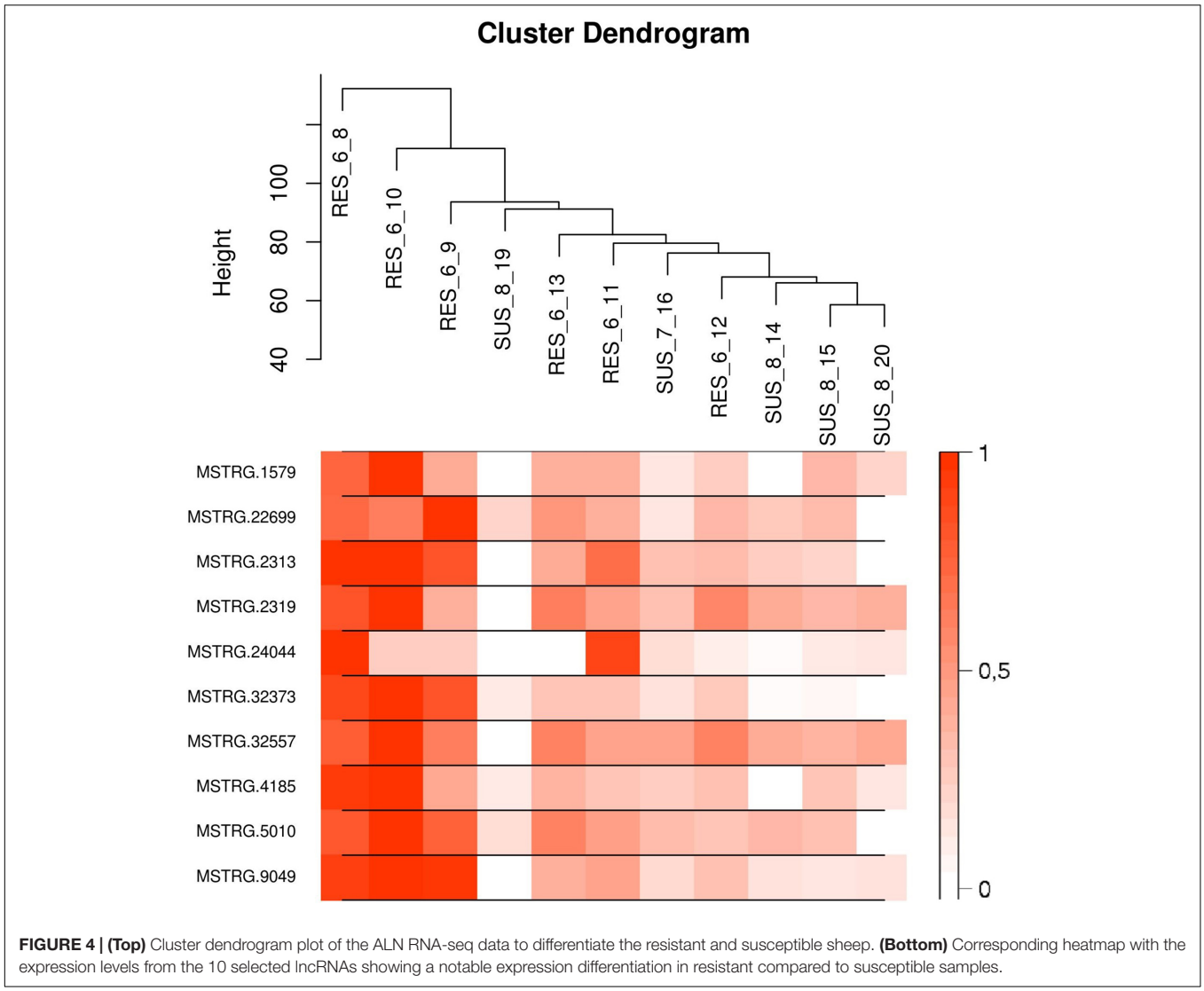


**FIGURE 4 | (Top)** Cluster dendrogram plot of the ALN RNA-seq data to differentiate the resistant and susceptible sheep. **(Bottom)** Corresponding heatmap with the expression levels from the 10 selected lncRNAs showing a notable expression differentiation in resistant compared to susceptible samples.

first step, we adopted a pipeline using bioinformatics tools for the detection and classification of lncRNAs expressed in the transcriptome of the ovine ALN. In the second step, differential gene expression and subsequent gene co-expression analyses were performed to establish co-expression networks of transcripts in gene expression modules, GNMs, followed by a correlation analysis between the expression level of 10 selected DE lncRNA loci and GNMs. In the final step, the potential functional role of these 10 target lncRNAs was predicted indirectly by performing gene enrichment and canonical biological pathway analyses with the DE genes of the significantly correlated GNMs.

## Co-expression of LncRNAs With Genes From ERK5 and SAPK/JNK Signaling Pathways

The canonical pathway *ERK5 (Extracellular signal-related kinase) Signaling* ($z$-score = −2.5) (**Supplementary Figure 4**) was on top of the enriched pathways in the GNM 'green.' The involved signal transduction molecules are known to elicit several biological responses and to regulate cell functions such as tissue morphogenesis, cell proliferation, differentiation, migration and survival, apoptosis, cytoskeletal rearrangements, immune response and adaptation or stress response as reviewed by different authors (Dong et al., 2002; Drew et al., 2012; Arthur and Ley, 2013; Stecca and Rovida, 2019). The genes from the GNM 'green,' belonging to the *ERK5 Signaling* pathway, were *ELK4* (encoding a transcription factor involved in both transcriptional activation and repression) and *GNA13* (associated with PKA Signaling and Rho-related Signaling). They had a higher expression level in the ALN transcriptome of RES compared to SUS sheep in response to GIN challenge and were also found to be significantly higher expressed in susceptible goat (Bhuiyan et al., 2017) and in susceptible selection lines of Perendale sheep (Diez-Tascon et al., 2005), respectively, in response to GIN infection (Bhuiyan et al., 2017).

**TABLE 2 |** Significant gene modules and canonical pathways associated with differentially expressed **(**DE) lncRNAs.

| GNM | Gene number[1] | Co-expressed DE lncRNAs | Correlation coefficient (lncRNA- GNM)[2] | Top canonical pathways associated with GNM | −log(*p*-value) | *z*-score | Genes involved |
|---|---|---|---|---|---|---|---|
| Green | 303 | MSTRG.1579 | −0.76 | ERK5 signaling | 4.55 | −2.45 | *RAP2A, PTPN11, GNAQ, KRAS, GNA13, ELK4* |
| | | MSTRG.2313 | −0.85 | SAPK/JNK signaling | 3.44 | −2.45 | *RAP2A, PTPN11, MAP3K13, MAP3K7, KRAS, GNA13* |
| | | MSTRG.2319 | −0.84 | 3-phosphoinositide degradation | 3.4 | −1.89 | *SYNJ2, PTPN11, NUDT9, PIKFYVE, PPP1CA, RNGTT, PPFIA4* |
| | | MSTRG.32373 | −0.8 | | | | |
| | | MSTRG.32557 | −0.88 | | | | |
| | | MSTRG.5010 | −0.93 | | | | |
| | | MSTRG.9049 | −0.84 | | | | |
| Yellow | 505 | MSTRG.1579 | −0.92 | RhoGDI signaling | 3.96 | 2.33 | *ROCK1, ROCK2, ITGA3, PAK1, CFL1, CDH6, ARHGDIA, ARHGEF10, DLC1, GNG12* |
| | | MSTRG.2313 | −0.82 | | | | |
| | | MSTRG.2319 | −0.89 | Axonal guidance signaling | 2.77 | NaN | *GLI2, CFL1, ADAM15, BCAR1, SEMA6C, ROCK1, ROCK2, ITGA3, PAK1, PLCB4, MMP11, PLXNB2, SEMA3C, PLCL1, GNG12* |
| | | MSTRG.32373 | −0.86 | | | | |
| | | MSTRG.32557 | −0.85 | | | | |
| | | MSTRG.4185 | −0.97 | | | | |
| | | MSTRG.5010 | −0.8 | | | | |
| | | MSTRG.9049 | −0.83 | | | | |
| Turquoise | 1517 | MSTRG.1579 | 0.92 | EIF2 signaling | 13.6 | 1.21 | *RPL11, RPLP1, SOS2, PDPK1, RPS11, BCL2, RPS20, UBA52, RPS13, EIF3A, IGF1R, RPS3, RPS5, RPL32, RPS19, NRAS, RALB, RPL29, RPS10, AGO2, RPS21, RPL23, RPL28, FAU, RPL10A, EIF3G, RPL8, RPS16, SREBF1, RPS27L, EIF4A1, EIF3I, RPS27A, RPL10, RPS25, RPS14, RPSA, EIF3K* |
| | | MSTRG.2313 | 0.9 | | | | |
| | | MSTRG.2319 | 0.94 | | | | |
| | | MSTRG.32373 | 0.88 | | | | |
| | | MSTRG.32557 | 0.93 | | | | |
| | | MSTRG.4185 | 0.92 | | | | |
| | | MSTRG.5010 | 0.93 | Regulation of eIF4 and p70S6K signaling | 8.83 | −1.34 | *SOS2, PDPK1, RPS11, RPS20, RPS13, EIF3A, RPS3, RPS5, ITGB1, RPS19, NRAS, RALB, RPS10, AGO2, RPS21, FAU, EIF3G, RPS16, RPS27L, EIF4A1, EIF3I, RPS27A, RPS25, RPSA, RPS14, EIF3K* |
| | | MSTRG.9049 | 0.83 | | | | |
| | | | | Oxidative phosphorylation | 5.01 | 4.0 | *SDHA, NDUFV1, COX4I2, COX6A1, ATP5F1D, NDUFB8, MT-CO2, NDUFA2, ATP5F1B, UQCR10, NDUFB7, CYC1, COX5A, CYB5A, UQCRC1, UQCRQ* |
| | | | | Mitochondrial dysfunction | 7.37 | NaN | *HSD17B10, SDHA, NDUFV1, COX4I2, ATP5F1D, COX6A1, PRDX5, TRAK1, BACE1, NDUFB8, MT-CO2, DHODH, GPX7, NDUFA2, BCL2, ATP5F1B, TXN2, UQCR10, CYC1,NDUFB7, COX5A, CYB5A, UQCRC1, ACO1, UQCRQ* |

[1] *Number of genes (in the GNM) with known function.*
[2] $p < 0.01$, $r > \pm 0.75$.

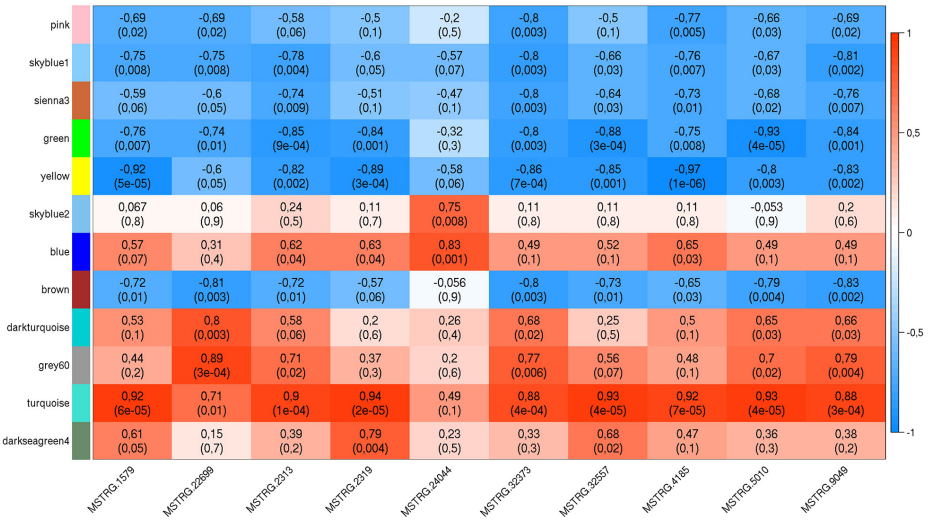


**FIGURE 5 |** Weighted gene co-expression plot with co-expression blocks corresponding to the significantly co-expressed ($p < 0.01$) gene network modules (GNMs) (on the **left**) and the 10 selected DE lncRNAs (on the **bottom**). The left axis consists of color codes of each GNM, the right axis is the correlation coefficient ($r$) reference scale with integer value and color intensity (red to blue) adopted for each lncRNA and GNM block. Each block consists of correlation coefficient (on the **top**) and p-value (in brackets) corresponding to each GNM-lncRNA pair.

Pathway enrichment analysis of the GNM 'green' showed co-expression of lncRNAs with genes included in the *SAPK/JNK Signaling* pathway ($z$-score = −2.5). The stress-activated protein kinase/c-Jun N-terminal kinases (SAPK/JNK) are known to be affected by many types of cellular stresses and extracellular signals, such as UV irradiation, inflammatory cytokines and growth factors. They participate in numerous different intracellular signaling pathways that control cellular processes, including cell proliferation, differentiation, transformation and migration, cytoskeletal integrity and DNA repair (Nishina et al., 2004). Finally, the modulation of the SAPK/JNK pathway by stress stimuli results in transcriptional regulation of stress-related genes. In our study, the genes *MAP3K13* and *MAP3K7* of the MAPK family were included in the *SAPK/JNK Signaling* pathway (**Table 2**), and both were lower expressed in the ALN transcriptome of RES compared to SUS sheep. MAP3K7 is a key signaling component of nuclear factor-κB (NF-κB) and MAPK signaling pathways and acts as an essential regulator of innate immune signaling and apoptosis and of the proinflammatory signaling pathway. It also plays a central role in adaptive immunity in response to physical and chemical stresses [reviewed by Dai et al. (2012)]. MAP3K13 is able to activate JNK (Ikeda et al., 2001), is implicated in NF-κB activation (Ikeda et al., 2001; Masaki et al., 2003) and a role in a variety of developmental, stress-sensing, and disease contexts is assumed (Jin and Zheng, 2019).

## Co-expression of LncRNAs With Genes Acting in RhoGDI and Axonal Guidance Signaling Pathways

Stress signals are known to be delivered to the SAPK/JNK signaling cascade by small GTPases of the Rho family. Our analysis revealed that lncRNAs mentioned above were also co-expressed with genes involved in pathways, which are linked to GTPases of the Rho family: *RhoGDI (Rho-specific guanine nucleotide dissociation inhibitor) Signaling* ($z$-score = 2.33) and 4 *Actin-based Motility by Rho* ($z$-score = −1.0) (**Supplementary Table 2** and **Figure 6**). These pathways all were enriched in the GNM 'yellow.' Rho GDP dissociation inhibitors (RhoGDIs) play important roles in various cellular processes, including cell migration, adhesion and proliferation, differentiation, cytoskeletal reorganization and membrane trafficking by regulating the functions of the Rho GTPase family members. Dissociation of Rho GTPases from RhoGDIs is necessary for their spatiotemporal activation (Dransart et al., 2005; Garcia-Mata et al., 2011; Cho et al., 2019). Several DE genes in the RhoGDI Signaling pathway, such as *ARHGEF10, DLC1* and *ARHGDIA*, were co-expressed with lncRNAs (**Supplementary Table 2**). The respective protein, ARHGEF10 is a guanine nucleotide exchange factor regulating activation of Rho GTPases, DLC1 belongs to the GTPase-activating proteins involved in the inactivation of Rho GTPases, whereas ARHGDIA is a RhoGDI, responsible for maintaining a stable pool of inactive Rho-GTPases (Stradal and Schelhaas, 2018). Co-expression of lncRNAs with the *ARHGDIA* gene that is higher expressed in the ALN transcriptome of RES compared to SUS sheep in our study might indicate a regulatory role of those lncRNAs for *ARHGDIA* expression. These lncRNAs were also co-expressed with ROCK kinase genes (*ROCK1* and *ROCK2*), which are included in all enriched Rho-associated pathways mentioned above but also in several *Ephrin Signaling* -related and *Axonal Guidance Signaling* pathways (**Supplementary Table 2**). In our study, both *ROCK* kinase isoform transcripts were lower expressed in the ALN transcriptome of RES compared to SUS sheep. Multiple functions

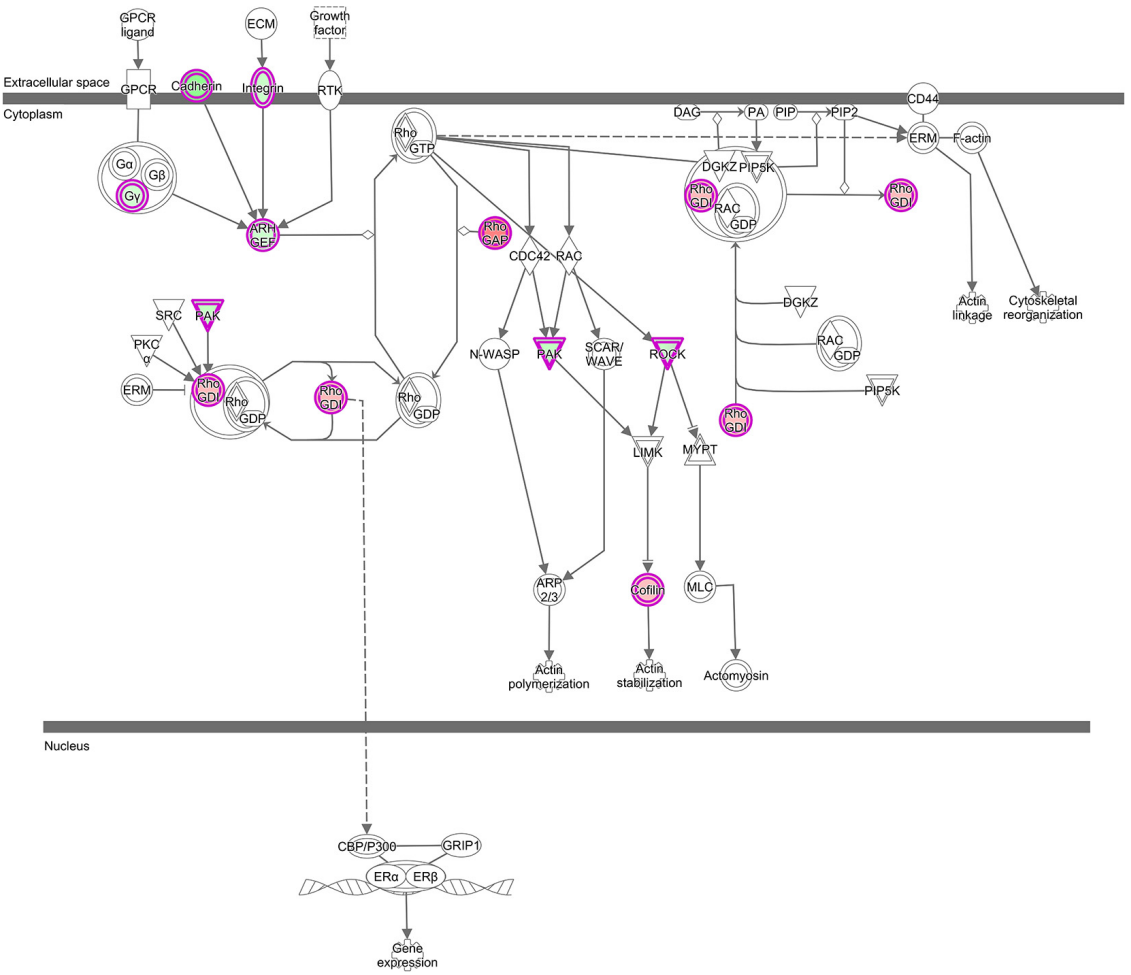


**FIGURE 6 |** "RhoGDI Signaling canonical pathway" from IPA, was the top enriched pathway with genes from GNM 'yellow' with *z*-score = 2.33. This figure is adapted from Ingenuity Pathway Analysis®.

of ROCK kinases have been detected in biological processes including cell contraction, migration, apoptosis, survival, and proliferation (Julian and Olson, 2014). The proteins are key regulators of actin organization, which link them to the actin-depolymerizing factor CFL1 that showed higher expression in the ALN transcriptome of RES versus SUS sheep in our study. CFL1 acts in Rho-induced reorganization of the actin cytoskeleton (actin depolymerisation/filament stabilization) and is known as key player in controlling the temporal and spatial pattern of actin dynamics, which is crucial for mediating host–pathogen interactions (Zheng et al., 2016). All intracellular and even some extracellular pathogens affect the host cell cytoskeleton to promote their own survival, replication, and dissemination (Stradal and Schelhaas, 2018). In summary, the RhoGDI canonical pathway also in conjunction with other Rho-related and Ephrin- related signaling cascades as well as with the *Axonal Guidance Signaling* pathway might be involved in controlling the differential resistance to nematode infection in sheep and possibly, all of these processes may be modulated by co-expressed lncRNAs.

## Co-expression of LncRNAs With Genes Involved in Protein Biosynthesis, Apoptosis and Mitochondrial Function Pathways

Furthermore, the co-expression analysis revealed that these lncRNAs were also connected with genes present in the GNM 'turquoise' suggesting a functional role of the canonical pathways *EIF2 Signaling* (*z*-score = 1.21) in conjunction with *Regulation of eIF4 and p70S6K Signaling* and *mTOR Signaling* (*z*-score = −1.34 for both) in sheep characterized by differential resistance to nematode challenge. The respective signaling pathways are engaged in the control of protein biosynthesis by regulating the translational machinery at different cascade steps (e.g., **Supplementary Figures 5**, **6**). As translation is a tightly regulated process in response to various stimuli, including extracellular and intracellular signals and environmental stress conditions, translational control plays a major role in host stress responses, including pathogenic infection and defense by enabling rapid responses to the challenge (Mohr and Sonenberg, 2012). In

our study, the participation of these biological pathways in the modulation of differential resistance to GIN challenge is supported by the upregulated expression of numerous ribosomal protein genes, and translation initiation factors, which are known to act in the ribosome biogenesis and the mRNA translational machinery. These genes from the GNM 'turquoise' showed a higher expression in the ALN transcriptome of sheep with a superior resistance against GIN. A crucial signaling mediator is the *PDPK1* gene (*3-phosphoinositide dependent protein kinase 1*) that is involved in all three canonical pathways *EIF2 Signaling*, *Regulation of eIF4 and p70S6K Signaling* and *mTOR Signaling*. The encoding protein PDK1 is a master kinase, able to control numerous physiological and pathological processes (Di Blasio et al., 2017). In our study, *PDPK1* had a lower expression level in the ALN transcriptome of SUS sheep compared to the RES group. But in goats, *PDPK1* was significantly higher expressed in susceptible animals after GIN infection (Bhuiyan et al., 2017).

GNM 'turquoise' also contained the antiapoptotic *BCL2* gene in the gene set associated with EIF2 signaling. The encoded BCL2 protein is known to be tightly connected with cellular survival. It suppresses apoptosis in many cell systems and regulates cell death by controlling the mitochondrial membrane permeability (Youle and Strasser, 2008; Siddiqui et al., 2015). Apoptosis, a mechanism by which cells infected with pathogens can be eliminated without triggering an unwanted inflammatory response, seemed to be upregulated in the RES group in response to GIN challenge in our study, which is supported by the higher *BCL2* expression in the RES compared to the SUS sheep. In the study of Ingham et al. (2008), the *BCL2* gene was quantified in the mucosa of the GIN tract of lambs of resistant and susceptible lines subjected to of *Haemonchus contortus* and *Trichostrongylus colubriformis* challenge, and it was identified as one of the candidate genes for maintaining the epithelial integrity of the gut in response to GIN. Furthermore, the *BCL2* gene was higher expressed in the duodenum of in Naïve Perendale sheep genetically susceptible to GIN compared to resistant animals (Keane et al., 2006), which is in line with our observations.

A regulatory link of these pathways is suggested due to the simultaneous co-expression with several lncRNAs that were also upregulated in the ALN transcriptome of those sheep characterized by a better resistance to nematode challenge. The functional interplay between protein biosynthesis cascades and PI3K/AKT- and MAPK- related pathways is clearly illustrated by the intertwining of the GNM 'turquoise' with the GNMs 'green' and 'yellow' and the associated lncRNAs. A closer look at the genomic location of the correlated lncRNAs revealed that some of them (MSTRG.9049, MSTRG.32373, MSTRG.2313, and MSTRG.5010) are located in genomic regions near ribosomal protein genes, which would indicate a regulatory potential for them to control the expression of the respective neighbored genes via *cis* -regulation. Considering that host cells respond to stress and imbalances by modifying gene expression at epigenetic, transcriptional and translational levels for recovering from the pathogen attacks, Knap et al. (2017) reviewed the state of knowledge on whether host cells would deploy lncRNAs to rapidly control host translation, the most energy-consuming process in cells, to counteract infection. The authors noted that the cause-effect relationship between the expression of lncRNAs and the activation of signaling pathways that control translation is currently unclear, but it is tempting to speculate that host cells could use this class of ncRNAs to fine-tune translation and cope with the imbalances triggered by pathogens.

Interestingly, numerous genes contained in the GNM 'turquoise' indicated an association of the correlated lncRNAs with the intertwined canonical pathways *Oxidative Phosphorylation* ($z$ – score = 2.45) and *Mitochondrial Dysfunction*. These genes encode for various protein components of all complexes forming the mitochondrial electron transport chain, which drives oxidative phosphorylation and appeared to be upregulated in the ALN transcriptome of the sheep group with a better resistance to GIN challenge (**Supplementary Figure 7**). Mitochondria are the key organelles of energy production and coordinate essential metabolic processes in the cells. In addition to their central role in metabolism, mitochondria also regulate cellular processes such as cell cycle, innate immunity and apoptosis [summarized by Mills et al. (2017) and Mohanty et al. (2019)]. Recent findings highlight the emerging role of mitochondria as important intracellular signaling platform that regulates innate immune and inflammatory responses (Jin et al., 2017). LncRNAs have been reported to be involved in regulating mitochondrial processes such as mitochondrial respiration, reactive oxygen production and apoptosis (De Paepe et al., 2018) and are hypothesized to coordinate functions between mitochondria and the nucleus (Dong et al., 2017). To maintain homeostasis of the cells, an intense cross-talk between mitochondria and the nucleus, mediated by lncRNAs is conceivable. Thus, lncRNAs that were highly correlated to genes of the GNM 'turquoise' might be associated with the regulation of mitochondrial bioenergetics and biosynthesis. Taken together, of the 10 DE lncRNAs selected for co-expression analysis, eight were found to be potentially involved in the regulation of these biological pathways with relevance to GIN infection, namely MSTRG.1579, MSTRG.2313, MSTRG.2319, MSTRG.32373, MSTRG.32557, MSTRG.5010, MSTRG.9049, and MSTRG.4185 (**Table 2**). Thus, these lncRNAs may contribute to the divergent resistance to gastrointestinal parasites.

## Potential Functional Genomic Interaction Partners of Target LncRNAs Potentially Involved in Regulating Divergent Resistance to Gastrointestinal Parasites in Sheep

LncRNAs might activate or repress the transcription of nearby genes (*cis*-regulation) present on the same (Engreitz et al., 2016) or opposite strand (Villegas and Zaphiropoulos, 2015; Tan-Wong et al., 2019).

As already mentioned, MSTRG.9049, MSTRG.32373, MSTRG.2313, and MSTRG.5010 are located in genomic regions antisense to DE ribosomal protein genes, *RPS19, RPS20,* and *RPL35A, RPL23* indicating a potential functional regulatory link to translation-associated processes.

MSTRG.1579 is overlapping with a known lincRNA (*ENSOARG00000025559*), and the *TAGLN2* gene region.

The *TAGLN2* is an actin binding protein and regulates the T cell activation in mammals by stabilizing the actin cytoskeleton (Na and Jun, 2015).

MSTRG.2319 represents a novel, not yet annotated intergenic lncRNA localized between the genes *MUC20* and *TNK2*, which however, were not DE between the different GIN-resistance groups. *MUC20* encodes a member of the mucin glycoprotein family implicated in protection of all mucosal surfaces. Membrane bound mucins have been suggested to play a functional role in cell signaling linked to health and disease (Corfield et al., 2001; Linden et al., 2008). A *MUC20* -lncRNA has been reported to bind ROCK1 and to be functionally involved in tumor suppression (Dai et al., 2020) indicating *trans*-regulation. *TNK2* encodes a tyrosine kinase that binds to the Rho family member Cdc42Hs and inhibits both the intrinsic and GTPase-activating protein-stimulated GTPase activity (Manser et al., 1993). Thus, MSTRG.2319 could be primarily involved in the pathway Signaling by Rho Family GTPases.

The lncRNA MSTRG.32557 is located near the *FABP4* gene region, which encodes a member of the fatty acid binding protein family that plays a role in lipid metabolism by binding and intracellular transport of long-chain fatty acids. *FABP4* showed a higher expression level in the ALN transcriptome of RES sheep and was found to be DE between RES and SUS sheep groups in our previous transcriptome study performed with the same tissue and conditions (Chitneedi et al., 2018). Other studies also imply roles of FABP family proteins in cell signaling, inhibition of cell growth and cellular differentiation. Furthermore, FABP also modulates tumor cell growth, metabolism, migration, differentiation and development involving the PI3K/AKT signaling and PPAR-associated pathways (Amiri et al., 2018). Specifically, FABP4 has been found to function as an adipokine that is involved in regulating macrophage and adipocyte interactions during inflammation (Thumser et al., 2014). A GWAS study showed that FABP4 contributes to resistance to fleece rot in Australian merino sheep (Smith et al., 2010). This gene was also found to be downregulated in response to GIN *Cooperia oncophora* infection in cattle (Li and Schroeder, 2012).

The lncRNA MSTRG.4185 is located close to the chemokine (C-C motif) ligand 14 *(CCL14)* gene region. The cytokine encoded by this gene induces changes in intracellular calcium concentration and enzyme release in monocytes and showed strong correlation with tumor immune cells infiltration (Gu et al., 2020). The expression of *CCL14* along with other chemokine ligands was reported to be significantly higher in a resistance sheep flock compared to a susceptible flock after *Haemonchus contortus* infection (Zhang et al., 2019). As the *CCL14* gene was not DE between RES and SUS sheep groups in our study, but MSTRG.4185 was highly correlated with GNMs (most negatively to the 'yellow' GNM and also positively correlated to the 'turquoise' GNM, **Figure 4**), a putative *trans*-regulation of genes included in correlated biological pathways could be assumed.

The lncRNAs, which are co-expressed with the GNMs 'green,' 'yellow,' and 'turquoise,' could possibly be involved in the regulation of the expression of genes included in the respective network modules. These modules, in turn are involved in pathways associated with different physiological and environmental conditions as well as with divergent phenotypic characteristics, such as the modulation of a differential resistance or susceptibility of adult sheep to parasite infection. Results obtained from this study only provide first hints on the genes and pathways that are primarily targeted by individual lncRNAs under interrogation. As this is an initial study, further research in sheep in response to GIN infection will be required to establish the functional role of the detected lncRNA transcripts.

Overall, this preliminary lncRNA study, conducted in adult sheep after GIN infection gave first insights into the potential functional role of selected lncRNAs by investigating their putative functions via co-expressed genes based on the principle of "guilt-by-association." Future multi-omics studies including DNA, RNA and metabolites will help to gain a better understanding of the general and specific roles of the selected lncRNAs and other lncRNAs significantly involved in the regulation of key physiological pathways associated with resistance against GIN infection in sheep.

## DATA AVAILABILITY STATEMENT

The 12 abomasal lymph node raw RNA-Seq datasets analyzed in this study have been submitted to the ENA Read Archive at EMBL-EBI following FAANG DCC guidelines. They are available at https://www.ebi.ac.uk/ena/browser/view/PRJEB33476 (Bioproject PRJEB33476). Data on the meta data of the samples are submitted as https://www.ebi.ac.uk/biosamples/samples/SAMEG4750853.

## ETHICS STATEMENT

The animal study was reviewed and approved by the subcommittee for experimentation and animal welfare (OBEA) of the University of León and the competent body of the Junta de Castilla y León regional government (Spain) (Ref. ULE_024_2015).

## AUTHOR CONTRIBUTIONS

BG-G and JA conceived and designed the study, and performed the collection and processing of tissue samples. MM-V designed and performed the experimental infections. BG-G, JA, and MM-V performed sampling in the commercial flock and selected the animals to be experimentally infected. CK and RW optimized the bioinformatics workflow. PC performed the bioinformatics analyses under the supervision of CK. PC and RW wrote the manuscript. BG-G, CK, and JA critically reviewed the manuscript. All authors read and approved the final manuscript.

## FUNDING

Financial support for the experimental work of this project was received from the LE248U14 project of the Junta de Castilla and

León regional government, whereas the storage and processing of the generated sequencing datasets have been funded by the RTI2018-093535-B-I00 project from the Spanish Ministry of Science and Innovation (MICINN). PC was funded by a Short-Term Scientific Missions of the Functional Annotation of Animal Genomes-European network (FAANG-Europe) COST Action CA15112 to do a short research stay at Leibniz Institute for Farm Animal Biology (FBN). MM-V was also funded by the "Ramón y Cajal" Programme (RYC-2015-18368) from MICINN.

## ACKNOWLEDGMENTS

The authors acknowledge the FAANG-Europe Cost Action (CA15112) for funding the Short Term Scientific Mission that settled the basis for this collaborative study. Authors thank to the National Breeders Association of Churra sheep for collaboration in the sampling of commercial flocks at the beginning of this project.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2021.685341/full#supplementary-material

**Supplementary Figure 1 |** Bioinformatics workflow showing the different steps performed to detect long non-coding RNAs (lncRNAs) in the abomasal lymph node transcriptome from resistant or susceptible sheep to GIN infection and to predict their potential functional role by gene co-expression and pathway enrichment analysis.

**Supplementary Figure 2 | (A)** Distribution of accumulated fecal egg count (FEC) in resistant and susceptible sheep after a second experimental infection with GIN *T. circumcincta*. Blue bars represent the FEC values in each resistant sample and yellow bars represent the FEC values in each susceptible sample. **(B)** PCA performed with the read count matrix of ALN RNA-seq datasets using R function "varianceStabilizingTransformation" by specifying the sample condition (resistant and susceptible) with the initially considered 6 resistant and 6 susceptible samples.

**Supplementary Figure 3 |** Weighted gene co-expression network matrix plot. Each block corresponds to the co-expression result of each gene network module (GNM, **left**) and a lncRNA (bottom). The left axis consists of color codes for each GNM and the right axis is the correlation coefficient (*r*) reference scale with integer values and color intensity (red to blue) adopted for each lncRNA and GNM block. Each block consists of the correlation coefficient **(top)** and *p*-value **(bottom)** corresponding to each GNM-lncRNA pair.

**Supplementary Figure 4 |** "EIF2 Signaling" from IPA, was the top enriched pathway with genes from the GNM 'turquoise' with *z*-score = 1.21. This figure is adapted from Ingenuity Pathway Analysis®.

**Supplementary Figure 5 |** "ERK5 Signaling" from IPA, was the top enriched pathway with genes from the GNM 'green' with *z*-score = −2.5. This figure is adapted from Ingenuity Pathway Analysis®.

**Supplementary Figure 6 |** "Regulation of eIF4 and p70S6K Signaling" from IPA, was the second top enriched pathway with genes from the GNM 'turquoise' with *z*-score = −1.34. This figure is adapted from Ingenuity Pathway Analysis®.

**Supplementary Figure 7 |** "Oxidative Phosphorylation" from IPA, was one of the top enriched pathways with genes from the GNM 'turquoise' with *z*-score = 4.00. This figure is adapted from Ingenuity Pathway Analysis®.

**Supplementary Table 1 |** Differentially expressed gene loci sorted based on p-value and the list of associated long non-coding RNAs (lncRNAs) along with the information of novel lncRNAs.

**Supplementary Table 2 |** List of top 20 pathways based on −log (*p*-value) enriched by the genes from the three GNMs green, yellow and turquoise according to the Ingenuity Pathway Analysis®.

**Supplementary File 1 |** Bash script with all the parameters adapted for different tools used in the bioinformatics workflow of lncRNA detection.

## REFERENCES

Amiri, M., Yousefnia, S., Seyed Forootan, F., Peymani, M., Ghaedi, K., and Nasr Esfahani, M. H. (2018). Diverse roles of fatty acid binding proteins (FABPs) in development and pathogenesis of cancers. *Gene* 676, 171–183. doi: 10.1016/j.gene.2018.07.035

Andersson, L., Archibald, A. L., Bottema, C. D., Brauning, R., Burgess, S. C., Burt, D. W., et al. (2015). Coordinated international action to accelerate genome-to-phenome with FAANG, the functional annotation of animal genomes project. *Genome Biol.* 16:57. doi: 10.1186/s13059-015-0622-4

Arthur, J. S., and Ley, S. C. (2013). Mitogen-activated protein kinases in innate immunity. *Nat. Rev. Immunol.* 13, 679–692. doi: 10.1038/nri3495

Bhuiyan, A. A., Li, J., Wu, Z., Ni, P., Adetula, A. A., Wang, H., et al. (2017). Exploring the genetic resistance to gastrointestinal nematodes infection in goat using RNA-sequencing. *Int. J. Mol. Sci.* 18:e040751. doi: 10.3390/ijms18040751

Bush, S. J., Muriuki, C., Mcculloch, M. E. B., Farquhar, I. L., Clark, E. L., and Hume, D. A. (2018). Cross-species inference of long non-coding RNAs greatly expands the ruminant transcriptome. *Genet. Sel. Evol.* 50:20. doi: 10.1186/s12711-018-0391-0

Cao, H., Wahlestedt, C., and Kapranov, P. (2018). Strategies to annotate and characterize long noncoding RNAs: advantages and pitfalls. *Trends Genet.* 34, 704–721. doi: 10.1016/j.tig.2018.06.002

Castilla-Gomez De Aguero, V., Gonzalez, J. F., Hernandez, J. N., Valderas-Garcia, E., Rojo Vazquez, F. A., Arranz, J. J., et al. (2020). Differences within Churra breed sheep in the early immune response to the infection by *Teladorsagia circumcincta*. *Parasitol. Res.* 120, 1115–1120. doi: 10.1007/s00436-020-06953-4

Chen, Y. G., Satpathy, A. T., and Chang, H. Y. (2017). Gene regulation in the immune system by long noncoding RNAs. *Nat. Immunol.* 18, 962–972. doi: 10.1038/ni.3771

Chitneedi, P. K., Arranz, J. J., Suarez-Vega, A., Martinez-Valladares, M., and Gutierrez-Gil, B. (2020). Identification of potential functional variants underlying ovine resistance to gastrointestinal nematode infection by using RNA-Seq. *Anim. Genet.* 51, 266–277. doi: 10.1111/age.12894

Chitneedi, P. K., Suarez-Vega, A., Martinez-Valladares, M., Arranz, J. J., and Gutierrez-Gil, B. (2018). Exploring the mechanisms of resistance to *Teladorsagia circumcincta* infection in sheep through transcriptome analysis of abomasal mucosa and abomasal lymph nodes. *Vet. Res.* 49:39. doi: 10.1186/s13567-018-0534-x

Cho, H. J., Kim, J. T., Baek, K. E., Kim, B. Y., and Lee, H. G. (2019). Regulation of Rho GTPases by RhoGDIs in human cancers. *Cells* 8:1037. doi: 10.3390/cells8091037

Clark, E. L., Bush, S. J., Mcculloch, M. E. B., Farquhar, I. L., Young, R., Lefevre, L., et al. (2017). A high resolution atlas of gene expression in the domestic sheep (*Ovis aries*). *PLoS Genet.* 13:e1006997. doi: 10.1371/journal.pgen.1006997

The ENCODE Project Consortium (2012). An integrated encyclopedia of DNA elements in the human genome. *Nature* 489, 57–74. doi: 10.1038/nature11247

Corfield, A. P., Carroll, D., Myerscough, N., and Probert, C. S. (2001). Mucins in the gastrointestinal tract in health and disease. *Front. Biosci.* 6, D1321–D1357. doi: 10.2741/corfield

Crick, F. (1970). Central dogma of molecular biology. *Nature* 227, 561–563. doi: 10.1038/227561a0

Dai, L., Aye Thu, C., Liu, X. Y., Xi, J., and Cheung, P. C. (2012). TAK1, more than just innate immunity. *IUBMB Life* 64, 825–834. doi: 10.1002/iub.1078

Dai, R., Zhou, Y., Chen, Z., Zou, Z., Pan, Z., Liu, P., et al. (2020). Lnc-MUC20-9 binds to ROCK1 and functions as a tumor suppressor in bladder cancer. *J. Cell Biochem.* 121, 4214–4225. doi: 10.1002/jcb.29626

De Paepe, B., Lefever, S., and Mestdagh, P. (2018). How long Noncoding RNAs enforce their will on mitochondrial activity: regulation of mitochondrial respiration, reactive oxygen species production, apoptosis, and metabolic reprogramming in cancer. *Curr. Genet.* 64, 163–172. doi: 10.1007/s00294-017-0744-1

Derrien, T., Johnson, R., Bussotti, G., Tanzer, A., Djebali, S., Tilgner, H., et al. (2012). The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. *Genome Res.* 22, 1775–1789. doi: 10.1101/gr.132159.111

Di Blasio, L., Gagliardi, P. A., Puliafito, A., and Primo, L. (2017). Serine/Threonine Kinase 3-phosphoinositide-dependent protein Kinase-1 (PDK1) as a key regulator of cell migration and cancer dissemination. *Cancers* 9:25. doi: 10.3390/cancers9030025

Diez-Banos, N., Cabaret, J., and Diez-Banos, P. (1992). Interspecific interactions in naturally acquired nematode communities from sheep abomasum in relation to age of host and season in four areas of Leon (Spain). *Int. J. Parasitol.* 22, 327–334. doi: 10.1016/s0020-7519(05)80010-4

Diez-Tascon, C., Keane, O. M., Wilson, T., Zadissa, A., Hyndman, D. L., Baird, D. B., et al. (2005). Microarray analysis of selection lines from outbred populations to identify genes involved with nematode parasite resistance in sheep. *Physiol. Genom.* 21, 59–69. doi: 10.1152/physiolgenomics.00257.2004

Dong, C., Davis, R. J., and Flavell, R. A. (2002). MAP kinases in the immune response. *Annu. Rev. Immunol.* 20, 55–72. doi: 10.1146/annurev.immunol.20.091301.131133

Dong, Y., Yoshitomi, T., Hu, J. F., and Cui, J. (2017). Long noncoding RNAs coordinate functions between mitochondria and the nucleus. *Epigenet. Chromat.* 10:41. doi: 10.1186/s13072-017-0149-x

Dransart, E., Olofsson, B., and Cherfils, J. (2005). RhoGDIs revisited: novel roles in Rho regulation. *Traffic* 6, 957–966. doi: 10.1111/j.1600-0854.2005.00335.x

Drew, B. A., Burow, M. E., and Beckman, B. S. (2012). MEK5/ERK5 pathway: the first fifteen years. *Biochim. Biophys. Acta* 1825, 37–48. doi: 10.1016/j.bbcan.2011.10.002

Dykes, I. M., and Emanueli, C. (2017). Transcriptional and post-transcriptional gene regulation by long non-coding RNA. *Genom. Proteom. Bioinform.* 15, 177–186. doi: 10.1016/j.gpb.2016.12.005

Engreitz, J. M., Haines, J. E., Perez, E. M., Munson, G., Chen, J., Kane, M., et al. (2016). Local regulation of gene expression by lncRNA promoters, transcription and splicing. *Nature* 539, 452–455. doi: 10.1038/nature20149

Garcia-Mata, R., Boulter, E., and Burridge, K. (2011). The 'invisible hand': regulation of RHO GTPases by RHOGDIs. *Nat. Rev. Mol. Cell Biol.* 12, 493–504. doi: 10.1038/nrm3153

Gu, Y., Li, X., Bi, Y., Zheng, Y., Wang, J., Li, X., et al. (2020). CCL14 is a prognostic biomarker and correlates with immune infiltrates in Hepatocellular carcinoma. *Aging* 12, 784–807. doi: 10.18632/aging.102656

Hezroni, H., Koppstein, D., Schwartz, M. G., Avrutin, A., Bartel, D. P., and Ulitsky, I. (2015). Principles of long noncoding RNA evolution derived from direct comparison of transcriptomes in 17 species. *Cell Rep.* 11, 1110–1122. doi: 10.1016/j.celrep.2015.04.023

Hiller, M., Findeiss, S., Lein, S., Marz, M., Nickel, C., Rose, D., et al. (2009). Conserved introns reveal novel transcripts in *Drosophila melanogaster*. *Genome Res.* 19, 1289–1300. doi: 10.1101/gr.090050.108

Hon, C. C., Ramilowski, J. A., Harshbarger, J., Bertin, N., Rackham, O. J., Gough, J., et al. (2017). An atlas of human long non-coding RNAs with accurate 5' ends. *Nature* 543, 199–204. doi: 10.1038/nature21374

Ikeda, A., Masaki, M., Kozutsumi, Y., Oka, S., and Kawasaki, T. (2001). Identification and characterization of functional domains in a mixed lineage kinase LZK. *FEBS Lett.* 488, 190–195. doi: 10.1016/s0014-5793(00)02432-7

Ingham, A., Reverter, A., Windon, R., Hunt, P. and Menzies, M. (2008). Gastrointestinal nematode challenge induces some conserved gene expression changes in the gut mucosa of genetically resistant sheep. *Int. J. Parasitol* 38, 431–442. doi: 10.1016/j.ijpara.2007.07.012

Jin, H. S., Suh, H. W., Kim, S. J., and Jo, E. K. (2017). Mitochondrial control of innate immunity and inflammation. *Immune Netw.* 17, 77–88. doi: 10.4110/in.2017.17.2.77

Jin, Y., and Zheng, B. (2019). Multitasking: dual leucine zipper-bearing kinases in neuronal development and stress management. *Annu. Rev. Cell Dev. Biol.* 35, 501–521. doi: 10.1146/annurev-cellbio-100617-062644

Johnson, G. L., and Lapadat, R. (2002). Mitogen-activated protein kinase pathways mediated by ERK, JNK, and p38 protein Kinases. *Science* 298, 1911–1912. doi: 10.1126/science.1072682

Julian, L., and Olson, M. F. (2014). Rho-associated coiled-coil containing kinases (ROCK): structure, regulation, and functions. *Small GTPases* 5:e29846. doi: 10.4161/sgtp.29846

Keane, O. M., Zadissa, A., Wilson, T., Hyndman, D. L., Greer, G. J., Baird, D. B., et al. (2006). Gene expression profiling of naive sheep genetically resistant and susceptible to gastrointestinal nematodes. *BMC Genom.* 7:42. doi: 10.1186/1471-2164-7-42

Kim, T. K., Hemberg, M., and Gray, J. M. (2015). Enhancer RNAs: a class of long noncoding RNAs synthesized at enhancers. *Cold Spring Harb. Perspect. Biol.* 7:a018622. doi: 10.1101/cshperspect.a018622

Knap, P., Tebaldi, T., Di Leva, F., Biagioli, M., Dalla Serra, M., and Viero, G. (2017). The unexpected tuners: are LncRNAs regulating host translation during infections? *Toxins* 9:357. doi: 10.3390/toxins9110357

Kopp, F., and Mendell, J. T. (2018). Functional classification and experimental dissection of long noncoding RNAs. *Cell* 172, 393–407. doi: 10.1016/j.cell.2018.01.011

Langfelder, P., and Horvath, S. (2008). WGCNA: an R package for weighted correlation network analysis. *BMC Bioinform.* 9:559. doi: 10.1186/1471-2105-9-559

Li, A., 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 Bioinform.* 15:311. doi: 10.1186/1471-2105-15-311

Li, H. (2011). A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics* 27, 2987–2993. doi: 10.1093/bioinformatics/btr509

Li, J., Gao, Z., Wang, X., Liu, H., Zhang, Y., and Liu, Z. (2016). Identification and functional analysis of long intergenic noncoding RNA genes in porcine pre-implantation embryonic development. *Sci. Rep.* 6:38333. doi: 10.1038/srep38333

Li, R. W., and Schroeder, S. G. (2012). Cytoskeleton remodeling and alterations in smooth muscle contractility in the bovine jejunum during nematode infection. *Funct. Integr. Genom.* 12, 35–44. doi: 10.1007/s10142-011-0259-7

Liao, Y., Smyth, G. K., and Shi, W. (2013). The Subread aligner: fast, accurate and scalable read mapping by seed-and-vote. *Nucleic Acids Res.* 41:e108. doi: 10.1093/nar/gkt214

Linden, S. K., Sutton, P., Karlsson, N. G., Korolik, V., and Mcguckin, M. A. (2008). Mucins in the mucosal barrier to infection. *Mucosal Immunol.* 1, 183–197. doi: 10.1038/mi.2008.5

Ling, Y., Zheng, Q., Sui, M., Zhu, L., Xu, L., Zhang, Y., et al. (2019). Comprehensive analysis of LncRNA reveals the temporal-specific module of goat skeletal muscle development. *Int. J. Mol. Sci.* 20:3950. doi: 10.3390/ijms20163950

Love, M. I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15:550. doi: 10.1186/s13059-014-0550-8

Ma, H., Hao, Y., Dong, X., Gong, Q., Chen, J., Zhang, J., et al. (2012). Molecular mechanisms and function prediction of long noncoding RNA. *Sci. World J.* 2012:541786. doi: 10.1100/2012/541786

Manser, E., Leung, T., Salihuddin, H., Tan, L., and Lim, L. (1993). A non-receptor tyrosine kinase that inhibits the GTPase activity of p21cdc42. *Nature* 363, 364–367. doi: 10.1038/363364a0

Maracaja-Coutinho, V., Paschoal, A. R., Caris-Maldonado, J. C., Borges, P. V., Ferreira, A. J., and Durham, A. M. (2019). Noncoding RNAs databases: current status and trends. *Methods Mol. Biol.* 1912, 251–285. doi: 10.1007/978-1-4939-8982-9_10

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J.* 17, 10–12. doi: 10.14806/ej.17.1.200

Masaki, M., Ikeda, A., Shiraki, E., Oka, S., and Kawasaki, T. (2003). Mixed lineage kinase LZK and antioxidant protein-1 activate NF-kappaB synergistically. *Eur. J. Biochem.* 270, 76–83. doi: 10.1046/j.1432-1033.2003.03363.x

Menard, K. L., Haskins, B. E., Colombo, A. P., and Denkers, E. Y. (2018). *Toxoplasma gondii* manipulates expression of host long Noncoding RNA

during intracellular infection. *Sci. Rep.* 8:15017. doi: 10.1038/s41598-018-33274-5

Menard, K. L., Haskins, B. E., and Denkers, E. Y. (2019). Impact of *Toxoplasma gondii* infection on host non-coding RNA responses. *Front. Cell Infect. Microbiol.* 9:132. doi: 10.3389/fcimb.2019.00132

Mercer, T. R., Dinger, M. E., and Mattick, J. S. (2009). Long non-coding RNAs: insights into functions. *Nat. Rev. Genet.* 10, 155–159. doi: 10.1038/nrg2521

Mills, E. L., Kelly, B., and O'neill, L. A. J. (2017). Mitochondria are the powerhouses of immunity. *Nat. Immunol.* 18, 488–498. doi: 10.1038/ni.3704

Mohanty, A., Tiwari-Pandey, R., and Pandey, N. R. (2019). Mitochondria: the indispensable players in innate immunity and guardians of the inflammatory response. *J. Cell Commun. Signal.* 13, 303–318. doi: 10.1007/s12079-019-00507-9

Mohr, I., and Sonenberg, N. (2012). Host translation at the nexus of infection and immunity. *Cell Host Microb.* 12, 470–483. doi: 10.1016/j.chom.2012.09.006

Moran, I., Akerman, I., Van De Bunt, M., Xie, R., Benazra, M., Nammo, T., et al. (2012). Human beta cell transcriptome analysis uncovers lncRNAs that are tissue-specific, dynamically regulated, and abnormally expressed in type 2 diabetes. *Cell Metab.* 16, 435–448. doi: 10.1016/j.cmet.2012.08.010

Na, B. R., and Jun, C. D. (2015). TAGLN2-mediated actin stabilization at the immunological synapse: implication for Cytotoxic T cell control of target cells. *BMB Rep.* 48, 369–370. doi: 10.5483/bmbrep.2015.48.7.132

Nishina, H., Wada, T., and Katada, T. (2004). Physiological roles of SAPK/JNK signaling pathway. *J. Biochem.* 136, 123–126. doi: 10.1093/jb/mvh117

Nolte, W., Weikard, R., Brunner, R. M., Albrecht, E., Hammon, H. M., Reverter, A., et al. (2019). Biological network approach for the identification of regulatory long Non-Coding RNAs associated with metabolic efficiency in cattle. *Front. Genet.* 10:1130. doi: 10.3389/fgene.2019.01130

Pang, K. C., Dinger, M. E., Mercer, T. R., Malquori, L., Grimmond, S. M., Chen, W., et al. (2009). Genome-wide identification of long noncoding RNAs in CD8+ T cells. *J. Immunol.* 182, 7738–7748. doi: 10.4049/jimmunol.0900603

Pertea, G., and Pertea, M. (2020). GFF Utilities: GffRead and GffCompare. *F1000Research* 9:304. doi: 10.12688/f1000research.23297.2

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, 290–295. doi: 10.1038/nbt.3122

Petri, A., Dybkaer, K., Bogsted, M., Thrue, C. A., Hagedorn, P. H., Schmitz, A., et al. (2015). Long Noncoding RNA expression during human B-Cell development. *PLoS One* 10:e0138236. doi: 10.1371/journal.pone.0138236

Ponting, C. P., Oliver, P. L., and Reik, W. (2009). Evolution and functions of long noncoding RNAs. *Cell* 136, 629–641. doi: 10.1016/j.cell.2009.02.006

Qureshi, I. A., Mattick, J. S., and Mehler, M. F. (2010). Long non-coding RNAs in nervous system function and disease. *Brain Res.* 1338, 20–35. doi: 10.1016/j.brainres.2010.03.110

Riquelme, I., Ili, C., Roa, J. C., and Brebi, P. (2016). Long non-coding RNAs in gastric cancer: mechanisms and potential applications. *Oncotarget* 5:48.

Robinson, A. (2015). *qualitytrim* [Online]. Available online at: https://bitbucket.org/arobinson/qualitytrim (accessed May 28, 2021).

Rochet, E., Appukuttan, B., Ma, Y., Ashander, L. M., and Smith, J. R. (2019). Expression of Long Non-coding RNAs by human retinal muller glial cells infected with clonal and exotic virulent *Toxoplasma gondii*. *Noncoding RNA* 5:48. doi: 10.3390/ncrna5040048

Roeber, F., Jex, A. R., and Gasser, R. B. (2013). Impact of gastrointestinal parasitic nematodes of sheep, and the role of advanced molecular tools for exploring epidemiology and drug resistance - an Australian perspective. *Parasit Vect.* 6:153. doi: 10.1186/1756-3305-6-153

Ruano, Z. M., Cortinhas, A., Carolino, N., Gomes, J., Costa, M., and Mateus, T. L. (2019). Gastrointestinal parasites as a possible threat to an endangered autochthonous portuguese sheep breed. *J. Helminthol.* 94:e103. doi: 10.1017/S0022149X19000968

Salavati, M., Caulton, A., Clark, R., Gazova, I., Smith, T. P. L., Worley, K. C., et al. (2020). Global analysis of transcription start sites in the new ovine reference genome (*oar rambouillet v1.0*). *Front. Genet.* 11:580580. doi: 10.3389/fgene.2020.580580

Schmitz, S. U., Grote, P., and Herrmann, B. G. (2016). Mechanisms of long noncoding RNA function in development and disease. *Cell Mol. Life Sci.* 73, 2491–2509. doi: 10.1007/s00018-016-2174-5

Seila, A. C., Calabrese, J. M., Levine, S. S., Yeo, G. W., Rahl, P. B., Flynn, R. A., et al. (2008). Divergent transcription from active promoters. *Science* 322, 1849–1851. doi: 10.1126/science.1162253

Siddiqui, W. A., Ahad, A., and Ahsan, H. (2015). The mystery of BCL2 family: Bcl-2 proteins and apoptosis: an update. *Arch. Toxicol.* 89, 289–317. doi: 10.1007/s00204-014-1448-7

Smith, W. J., Li, Y., Ingham, A., Collis, E., Mcwilliam, S. M., Dixon, T. J., et al. (2010). A genomics-informed, SNP association study reveals FBLN1 and FABP4 as contributing to resistance to fleece rot in Australian Merino sheep. *BMC Vet. Res.* 6:27. doi: 10.1186/1746-6148-6-27

Srijyothi, L., Ponne, S., Prathama, T., Cheemala, A., and Baluchamy, S. (2018). *Roles of Non-Coding RNAs in Transcriptional Regulation*. London, UK: IntechOpen. 50.

Stecca, B., and Rovida, E. (2019). Impact of ERK5 on the hallmarks of cancer. *Int. J. Mol. Sci.* 20:1426. doi: 10.3390/ijms20061426

Stradal, T. E. B., and Schelhaas, M. (2018). Actin dynamics in host-pathogen interaction. *FEBS Lett.* 592, 3658–3669. doi: 10.1002/1873-3468.13173

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:e166. doi: 10.1093/nar/gkt646

Taft, R. J., Kaplan, C. D., Simons, C., and Mattick, J. S. (2009). Evolution, biogenesis and function of promoter-associated RNAs. *Cell Cycle* 8, 2332–2338. doi: 10.4161/cc.8.15.9154

Tan-Wong, S. M., Dhir, S., and Proudfoot, N. J. (2019). R-Loops promote antisense transcription across the mammalian genome. *Mol. Cell* 76, 600–616.e606. doi: 10.1016/j.molcel.2019.10.002

Thiel, D., Conrad, N. D., Ntini, E., Peschutter, R. X., Siebert, H., and Marsico, A. (2019). Identifying lncRNA-mediated regulatory modules via ChIA-PET network analysis. *BMC Bioinform.* 20:292. doi: 10.1186/s12859-019-2900-8

Thorvaldsdottir, H., Robinson, J. T., and Mesirov, J. P. (2013). Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief Bioinform.* 14, 178–192. doi: 10.1093/bib/bbs017

Thumser, A. E., Moore, J. B., and Plant, N. J. (2014). Fatty acid binding proteins: tissue-specific functions in health and disease. *Curr. Opin. Clin. Nutr. Metab. Care* 17, 124–129. doi: 10.1097/MCO.0000000000000031

Ulitsky, I., and Bartel, D. P. (2013). lincRNAs: genomics, evolution, and mechanisms. *Cell* 154, 26–46. doi: 10.1016/j.cell.2013.06.020

Villegas, V. E., and Zaphiropoulos, P. G. (2015). Neighboring gene regulation by antisense long non-coding RNAs. *Int. J. Mol. Sci.* 16, 3251–3266. doi: 10.3390/ijms16023251

Wang, J., Chai, Z., Deng, L., Wang, J., Wang, H., Tang, Y., et al. (2020). Detection and integrated analysis of lncRNA and mRNA relevant to plateau adaptation of Yak. *Reprod. Domest. Anim.* 55, 1461–1469. doi: 10.1111/rda.13767

Wapinski, O., and Chang, H. Y. (2011). Long noncoding RNAs and human disease. *Trends Cell Biol.* 21, 354–361. doi: 10.1016/j.tcb.2011.04.001

Weikard, R., Demasius, W., and Kuehn, C. (2017). Mining long noncoding RNA in livestock. *Anim. Genet.* 48, 3–18. doi: 10.1111/age.12493

Weikard, R., Hadlich, F., Hammon, H. M., Frieten, D., Gerbert, C., Koch, C., et al. (2018). Long noncoding RNAs are associated with metabolic and cellular processes in the jejunum mucosa of pre-weaning calves in response to different diets. *Oncotarget* 9, 21052–21069. doi: 10.18632/oncotarget.24898

Wilusz, J. E., Sunwoo, H., and Spector, D. L. (2009). Long noncoding RNAs: functional surprises from the RNA world. *Genes Dev.* 23, 1494–1504. doi: 10.1101/gad.1800909

Wolfe, C. J., Kohane, I. S., and Butte, A. J. (2005). Systematic survey reveals general applicability of "guilt-by-association" within gene coexpression networks. *BMC Bioinform.* 6:227. doi: 10.1186/1471-2105-6-227

Wucher, V., Legeai, F., Hedan, B., Rizk, G., Lagoutte, L., Leeb, T., et al. (2017). FEELnc: a tool for long non-coding RNA annotation and its application to the dog transcriptome. *Nucleic Acids Res.* 45:e57. doi: 10.1093/nar/gkw1306

Youle, R. J., and Strasser, A. (2008). The BCL-2 protein family: opposing activities that mediate cell death. *Nat. Rev. Mol. Cell Biol.* 9, 47–59. doi: 10.1038/nrm2308

Zhang, R., Liu, F., Hunt, P., Li, C., Zhang, L., Ingham, A., et al. (2019). Transcriptome analysis unraveled potential mechanisms of resistance to *Haemonchus contortus* infection in Merino sheep populations bred for parasite resistance. *Vet. Res.* 50:7. doi: 10.1186/s13567-019-0622-6

Zhang, T., Zhang, X., Han, K., Zhang, G., Wang, J., Xie, K., et al. (2017). Analysis of long noncoding RNA and mRNA using RNA sequencing during

the differentiation of intramuscular preadipocytes in chicken. *PLoS One* 12:e0172389. doi: 10.1371/journal.pone.0172389

Zhao, X. Y., and Lin, J. D. (2015). Long Noncoding RNAs: a new regulatory code in metabolic control. *Trends Biochem. Sci.* 40, 586–596. doi: 10.1016/j.tibs.2015.08.002

Zhao, Y., Li, H., Fang, S., Kang, Y., Wu, W., Hao, Y., et al. (2016). NONCODE 2016: an informative and valuable data source of long non-coding RNAs. *Nucleic Acids Res.* 44, D203–D208. doi: 10.1093/nar/gkv1252

Zheng, K., Kitazato, K., Wang, Y., and He, Z. (2016). Pathogenic microbes manipulate cofilin activity to subvert actin cytoskeleton. *Crit. Rev. Microbiol.* 42, 677–695. doi: 10.3109/1040841X.2015.1010139

Zheng, X., Ning, C., Zhao, P., Feng, W., Jin, Y., Zhou, L., et al. (2018). Integrated analysis of long noncoding RNA and mRNA expression profiles reveals the potential role of long noncoding RNA in different bovine lactation stages. *J. Dairy Sci.* 101, 11061–11073. doi: 10.3168/jds.2018-14900

**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.

frontiers in Genetics

ORIGINAL RESEARCH
published: 22 June 2021
doi: 10.3389/fgene.2021.678667

# STAT3 Partly Inhibits Cell Proliferation *via* Direct Negative Regulation of *FST* Gene Expression

*Haidong Xu[1†], Guangwei Ma[1,2†], Fang Mu[1], Bolin Ning[1], Hui Li[1] and Ning Wang[1*]*

[1] College of Animal Science and Technology, Northeast Agricultural University, Harbin, China, [2] Ministry of Education Key Laboratory for Ecology of Tropical Islands, Key Laboratory of Tropical Animal and Plant Ecology of Hainan Province, College of Life Sciences, Hainan Normal University, Haikou, China

Follistatin (FST) is a secretory glycoprotein and belongs to the TGF-β superfamily. Previously, we found that two single nucleotide polymorphisms (SNPs) of sheep *FST* gene were significantly associated with wool quality traits in Chinese Merino sheep (Junken type), indicating that FST is involved in the regulation of hair follicle development and hair trait formation. The transcription regulation of human and mouse *FST* genes has been widely investigated, and many transcription factors have been identified to regulate *FST* gene. However, to date, the transcriptional regulation of sheep *FST* is largely unknown. In the present study, genome walking was used to close the genomic gap upstream of the sheep genomic *FST* gene and to obtain the *FST* gene promoter sequence. Transcription factor binding site analysis showed sheep *FST* promoter region contained a conserved putative binding site for signal transducer and activator of transcription 3 (STAT3), located at nucleotides −423 to −416 relative to the first nucleotide (A, +1) of the initiation codon (ATG) of sheep *FST* gene. The dual-luciferase reporter assay demonstrated that STAT3 inhibited the *FST* promoter activity and that the mutation of the putative STAT3 binding site attenuated the inhibitory effect of STAT3 on the *FST* promoter activity. Additionally, chromatin immunoprecipitation assay (ChIP) exhibited that STAT3 is directly bound to the *FST* promoter. Cell proliferation assay displayed that FST and STAT3 played opposite roles in cell proliferation. Overexpression of sheep *FST* significantly promoted the proliferation of sheep fetal fibroblasts (SFFs) and human keratinocyte (HaCaT) cells, and overexpression of sheep *STAT3* displayed opposite results, which was accompanied by a significantly reduced expression of *FST* gene ($P < 0.05$). Taken together, STAT3 directly negatively regulates sheep *FST* gene and depresses cell proliferation. Our findings may contribute to understanding molecular mechanisms that underlie hair follicle development and morphogenesis.

**Keywords: STAT3, FST, transcriptional regulation, cell proliferation, sheep**

OPEN ACCESS

***Edited by:***
*Priyanka Banerjee, North Dakota State University, United States*

***Reviewed by:***
*Doreen Becker, Leibniz Institute for Farm Animal Biology (FBN), Germany*
*Pablo Fonseca, University of Guelph, Canada*

***Correspondence:**
*Ning Wang*
*wangning@neau.edu.cn*

*[†]These authors have contributed equally to this work*

***Specialty section:***
*This article was submitted to Livestock Genomics, a section of the journal Frontiers in Genetics*

***Received:*** *10 March 2021*
***Accepted:*** *27 May 2021*
***Published:*** *22 June 2021*

***Citation:***
*Xu H, Ma G, Mu F, Ning B, Li H and Wang N (2021) STAT3 Partly Inhibits Cell Proliferation via Direct Negative Regulation of FST Gene Expression. Front. Genet. 12:678667. doi: 10.3389/fgene.2021.678667*

## INTRODUCTION

The hair follicle is a skin appendage with a complex structure composed of the dermal papilla, hair bulbs, outer root sheaths (ORS), inner root sheaths (IRS), and the hair matrix (Schneider et al., 2009; Plowman and Harland, 2018). Hair follicle morphogenesis and development involve proliferation, differentiation, and apoptosis of hair follicle stem cells (Alonso and Fuchs, 2006;

Plowman and Harland, 2018). The hair follicle undergoes life-long cyclic transformations exhibiting anagen (growth), catagen (regression), and telogen (rest) phases, respectively (Stenn and Paus, 2001; Schneider et al., 2009). A variety of growth factors and cytokines have been shown to tightly regulate the hair follicle morphogenesis and development through acting on the epithelial-mesenchymal interaction (Müller-Röver et al., 2001; Wang et al., 2012), such as fibroblast growth factor (FGF) (Pispa and Thesleff, 2003), tumor necrosis factor (TNF) (Liu et al., 2018) and transforming growth factor-β (TGF-β) (Li et al., 2016). As an antagonist of the TGF-β superfamily, follistatin (FST) is highly expressed in the matrix of hair follicles, which consist of cells with a strong proliferation ability (Ma et al., 2017). *FST* transgenic mice exhibited shinier and more irregular hair (Guo et al., 1998; Wankell et al., 2001). *FST* knockout mice died within hours after birth and displayed curlier whiskers (Matzuk et al., 1995; Jhaveri et al., 1998; Nakamura et al., 2003). Our previous association analysis demonstrated that two single nucleotide polymorphisms (SNPs) in sheep *FST* gene were associated with wool quality traits in Chinese Merino sheep (Junken Type) (Ma et al., 2017). Collectively, these data indicated that FST is involved in the regulation of hair follicle development and hair trait formation.

The transcriptional regulation of human and mouse *FST* genes has been widely investigated, and many transcription factors have been identified to regulate *FST* gene. For example, erythroid 2 related factor (Nrf2) directly regulates *FST* gene and inhibits the apoptosis of human lung epithelial cells and A549 cells (Lin et al., 2016). It has been shown that transcription of *FST* gene is directly regulated by β-catenin/transcription factor 4 (TCF4) transcription factor complex to promote the murid myogenic differentiation and myoblast fusion *in vitro* and *in vivo* (Han et al., 2014). *FST* gene, involving in skeletal muscle development of *L. crocea*, is suppressed by MyoD and Sox8 (Yang et al., 2016). Myogenin promotes the satellite cell differentiation of adult mouse myogenesis in an *FST*-dependent manner (Jones et al., 2015). Estrogen-related receptor β (ERRβ) inhibits epithelial to mesenchymal transition in breast cancer through directly boosting *FST* expression (Sengupta et al., 2014). Moreover, transcription factor SP1 induces *FST* transcription in intestinal epithelial cells and kidney mesangial cells (Necela et al., 2008; Mehta et al., 2019). *FST* gene expression is predominantly up-regulated by GLI family zinc finger 2 (*GLI2*) in human keratinocytes (Eichberger et al., 2008). However, to date, the transcriptional regulation of sheep *FST* is largely unknown.

In this study, we investigated the transcription regulation of sheep *FST* gene by a transcription factor, signal transducer and activator of transcription 3 (STAT3), and our results demonstrated that *FST* gene is a target of STAT3 and that STAT3 inhibits cell proliferation at least partly *via* direct negative regulation *FST* gene expression.

# MATERIALS AND METHODS

## Ethics Statement

All animal experiments were conducted according to the guidance for the care and use of experimental animals established by the Ministry of Science and Technology of the People's Republic of China (Approval number: 2006-398) and approved by the Laboratory Animal Management Committee of Northeast Agricultural University.

## Cell Culture

HEK293T and HaCaT cells, purchased from the China Center for Type Culture Collection, were cultured in Dulbecco's Modified Eagle's Medium (DMEM) (Gibco, United States). Sheep fetal fibroblasts (SFFs), gifted from Dr. Tiezhu An, Northeast Forestry University, Harbin, were cultured in DMEM nutrient mixture F12 (DMEM-F12, Gibco, United States). All cells were cultured in the medium supplemented with 10% fetal bovine serum (FBS) (Biological Industries, Germany) plus 1% penicillin/streptomycin (Gibco, United States) at 37°C in 5% $CO_2$.

## Genome Walking

There is a genomic gap immediately upstream of the sheep *FST* gene according to *Ovis aries* reference genome (ISGC Oar_v3.1/oviAri3). To close the genomic gap upstream of the sheep FST gene, genome walking was performed as previously described (Shapter and Waters, 2014). Briefly, genomic DNA was isolated from sheep skin samples, previously collected and preserved (Ma et al., 2017), using the phenol-chloroform method (Green and Sambrook, 2014). Three *FST* gene-specific reverse primers: FST-SP1, FST-SP2, and FST-SP3, were designed and synthesized. Their sequences and location are presented in **Supplementary Table 1** and **Supplementary Figure 1**, respectively. Three forward primers: ZFP2, ZSP1, and ZSP2 were provided by the KX Genome Walking Kit (Zomanbio, China). Three rounds of polymerase chain reaction (PCR) were performed to amplify the genomic gap region with these primers. The 3′ end of ZFP2 is a random sequence and its 5′ end is a specific sequence, which can be matched by primers ZSP1 and ZSP2 in second- and third-round PCR reactions, respectively. The first PCR (primers: ZFP2 and FST-SP1) was performed using genomic DNA as a template. Second PCR (primers: ZSP1 and FST-SP2) and the third PCR (primers: ZSP2 and FST-SP3) were performed using the product (1 μL) from the first and second rounds of PCR as a template, respectively. The first PCR was performed in a reaction volume of 50 μL including 200 ng genomic DNA, 10 μL dNTPs (2.5 mM), 25 μL 2 × Kx PCR Buffer (with $Mg^{2+}$), 1 μL Kx Pfu DNA Polymerase (1 U/μL), 7.5 μL ZFP2primers (10 pmol/μL), and 1.5 μL *FST*-SP1 (10 pmol/μL). The first PCR conditions were as follows: initial denaturation at 94°C for 2 min, followed by 2 cycles (98°C for 10 s, 60°C for 30 s, 68°C for 2 min), 98°C for 10 s, 25°C for 2 min, 25 to 68°C for 0.2°C/s, 68°C for 2 min, 6 cycles (98°C for 10 s, 60°C for 30 s, 68°C for 2 min, 98°C for 10 s, 60°C for 30 s, 68°C for 2 min, 98°C for 10 s, 44°C for 30 s, 68°C for 2 min), with a final extension at 68°C for 5 min. The second and third PCRs were also conducted in a 50 μL reaction volume including 1 μL template, 10 μL dNTPs (2.5 mM), 25 μL 2 × Kx PCR Buffer (with $Mg^{2+}$), 1 μL Kx Pfu DNA Polymerase (1U/μL), and 1.5 μL primers (ZSP1 and *FST*-SP2 for the second PCR, ZSP2, and FST-SP3 for the third PCR, 10 pmol/μL), and run with a thermal protocol of 94°C for 2 min, followed by 30 cycles (98°C for 10 s,

60°C for 30 s, 68°C for 2 min), with a final extension at 68°C for 5 min. The third PCR product was resolved on a 1.2% agarose gel, recovered, and cloned into pEASY-T1 Simple Cloning Vector (TransGen Biotech, China) for sequencing in both directions.

## Bioinformatics Analysis

In this study, the first nucleotide (A) of the initiation codon (ATG) of *FST* was assigned position + 1. The *FST* promoter sequences of different animal species were obtained from the UCSC website[1]. The conserved transcription factor binding sites were predicted by using the Mulan website tool[2] with the option "optimized for function" in matrix similarity and "vertebrates" in biological species (Ovcharenko et al., 2005).

## Plasmid Construction and Transient Transfection

For the construction of *FST* and *STAT3* expression vectors, based on *FST* (NM_001257093.1) and *STAT3* (XM_015098787) sequences, two pairs of primers (FST-V and STAT3-V, **Supplementary Table 1**) were designed to amplify the full-length coding regions (CDSs) of sheep *FST* and *STAT3* genes, respectively. The full-length CDSs of *FST* and *STAT3* were amplified by reverse transcription PCR (RT-PCR) from the pooled total RNA of sheep skin ($n$ = 3) using the primer pairs FST-V and STAT3-V, respectively. The *FST* and *STAT3* PCR products were individually ligated into the pCMV-Myc (Clontech, United States), and the resulting plasmids were named pCMV-Myc-FST and pCMV-Myc-STAT3, respectively.

To construct *FST* promoter luciferase reporter vectors, the highly conserved region (−980/−340) of sheep *FST* promoter, which harbored the conserved putative STAT3 binding site (from −423 to −416), was PCR amplified with two primer pairs FST-P(+) and FST-P(−) (**Supplementary Table 1**) using sheep genomic DNA as the template. Subsequently, the two amplified *FST* promoter fragments were inserted into the *Kpn*I and *Hin*dIII sites of pGL3-basic (Promega, United States) to yield two *FST* promoter reporters. The reporter with the *FST* promoter fragment in the right direction was named pGL3-FST(−980/−340) and the other one with the *FST* promoter fragment in opposite direction was named pGL3-FST(−340/−980).

There was only one putative STAT3 binding site "CGATTCCCC" in sheep FST promoter (locating from −423 to −416). The mutation of this putative STAT3 binding site was expected to prevent STAT3 from binding to the *FST* promoter (Shackleford et al., 2011). This site mutation has not been reported to be associated with the wool trait. To test whether STAT3 directly regulates sheep *FST* gene, this putative STAT3 binding site was mutated to CGAGGTACC in the reporter pGL3-FST(−980/−340) using the Fast Mutagenesis System (TransGen, China) and the primer pairs FST-M according to the manufacturer's recommendation. The resulting reporter construct was named pGL3-FST(−980/−340)-mutSTAT3. All primers were synthesized by Invitrogen (Shanghai, China) and all constructs were confirmed by Sanger sequencing (Invitrogen, Shanghai, China).

[1]http://genome.ucsc.edu/

[2]https://mulan.dcode.org/

## Dual-Luciferase Reporter Assay

Briefly, the HEK293T cells were seeded in a 24-well plate ($2 \times 10^5$ cells/well) and cultured in the DMEM medium supplemented with 10% FBS and 1% penicillin/streptomycin. After the HEK293T cell reached 70–80% confluence, HEK293T cells were co-transfected with either pGL3-basic, pGL3-FST(980/−340), pGL3-FST(−340/−980) or pGL3-FST(−980/−340)-mutSTAT3 and either pCMV-Myc or pCMV-Myc-STAT3 using Lipofectamine 2000 (Invitrogen, United States) according to the manufacturer's instructions. Dual-luciferase reporter assays were performed 48 h post-transfection using the dual-luciferase reporter assay system (Promega, United States) according to the manufacturer's instructions. The firefly luciferase (*Fluc*) signal was normalized to that of Renilla luciferase (*Rluc*).

## Western Blot Analysis

Western blotting was performed to identify the two expression vectors (pCMV-Myc-FST and pCMV-Myc-STAT3). Briefly, HEK293T cells were transfected with pCMV-Myc-FST or pCMV-Myc-STAT3, 48 h after transfection, the cells were harvested in RIPA lysis buffer (Beyotime, China) containing 1% phenyl methane sulfonyl fluoride (Beyotime, China). After incubation on ice for 30 min, the supernatant was collected by centrifugation at 10, 000 × g for 5 min at 4°C. Equal amounts of protein from the cell lysates were separated by 12% sulfate dodecyl sodium-polyacrylamide gel electrophoresis (SDS-PAGE) and transferred to nitrocellulose membranes (Millipore, United States). After blocking with 5% (w/v) dry milk and 0.1% Tween 20 for 2 h, the membranes were incubated with Myc-tag mouse monoclonal antibody (Abcam, 1:1,000) at room temperature for 2 h. Subsequently, the membranes were incubated with horseradish peroxidase-conjugated secondary rabbit anti-mouse IgG (H&L) antibody (Abcam, 1:5,000) at room temperature for 1 h, followed by visualization using an ECL Plus detection kit (Beyotime, China).

## Chromatin Immunoprecipitation Assay

Chromatin immunoprecipitation was accomplished using a ChIP assay kit (Beyotime, China) as previously described (Deng et al., 2012). Briefly, HEK293T cells were co-transfected with pGL3-FST(−980/−340) and either pCMV-Myc-STAT3 or pCMV-Myc, at 48 h post-transfection, the cells were fixed with 1% formaldehyde at room temperature for 10 min. The Chromatin was digested with 0.5 μL micrococcal nuclease into 100–900 bp DNA/protein fragments, following immunoprecipitated with 5 μg of anti-Myc antibody (Abcam, United States) and mouse IgG (negative control) (Beyotime, China), respectively. The purified DNA fragments were measured by quantitative PCR (qPCR) using the FST-C primer pairs (**Supplementary Table 1**), which was performed on the 7,500 Fast Real-Time PCR System (Applied Biosystems, United States) with SYBR Green PCR Master Mix (Roche Molecular Systems, United States). The qPCR reaction volume was 20 μL including 1 μL of cDNA,

10 μL of 2 × SYBR Green PCR MasterMix (Roche Molecular Systems, United States), 0.5 μL each of the forward and reverse primers (10 μM), and 8 μL double-distilled water. The qPCR conditions were as follows: 95°C for 10 min; 40 cycles at 95°C for 30 s, 60°C for 30 s. Non-immunoprecipitated DNA (2%) was used as input control. Two additional negative controls, mouse IgG (A) and anti-Myc antibody (B), were prepared by the co-transfection of HEK293T cells with pGL3-FST(−980/−340) and pCMV-Myc. The qPCR data were normalized to input chromatin DNA and presented as fold enrichment over the input control using ΔCt equation (Tatler et al., 2016), which signal relative to input = $0.2 \times 2^{-\Delta\ Ct}$, $\Delta Ct = Ct_{[IPsample]} - Ct_{[Inputsample]}$ (Patrik, 2018). The amplification product of the qPCR was analyzed by agarose gel electrophoresis with 1.5% consistency (g/mL).

## Cell Counting Kit-8 Assay

The Cell Counting Kit-8 (CCK-8) assay was used to assay cell proliferation. Briefly, the SFFs and HaCaT cells were seeded in a 96-well plate ($2 \times 10^4$ cells/well) and cultured in the DMEM-F12 and DMEM medium, respectively, supplemented with 10% FBS and 1% penicillin/streptomycin. The cells were individually transfected with pCMV-Myc, pCMV-Myc-FST or pCMV-Myc-STAT3 for 24, 48, and 72 h, every well was added with 10 μL CCK-8 solution (TransGen, China) and incubated at 37°C for 2 h, following the absorbance was measured at 450 nm using a Model 680 Microplate Reader (Bio-Rad, United States). The cells transfected with pCMV-Myc were used as the negative control.

## RNA Isolation, Reverse Transcription, and qPCR

The SFFs and HaCaT cells were transfected with pCMV-Myc, pCMV-Myc-FST, or pCMV-Myc-STAT3, at 48 h after transfection, total RNAs were isolated using Trizol reagent (Invitrogen, United States) according to the standard procedures, and RNA quality was assessed by denaturing formaldehyde agarose gel electrophoresis. Reverse transcription was performed with ImProm-II Reverse Transcriptase (Promega, United States) according to the manufacturer's protocols.

The expression of proliferation marker genes, *Ki67* and proliferating cell nuclear antigen (*PCNA*), were detected by quantitative real-time PCR (qPCR). The qPCR reaction volume was 20 μL containing 1 μL of cDNA, 10 μL of 2 × SYBR Green PCR MasterMix (Roche Molecular Systems, United States), 0.5 μL each of the forward and reverse primers (10 μM), and 8 μL double-distilled water. Thermal cycling consisted of an initial step at 95°C for 10 min followed by 40 cycles at 95°C for 30 s and 60°C for 30 s. The primers used for qPCR are listed in **Supplementary Table 1**. The target gene expression was normalized to the glyceraldehyde 3-phosphate dehydrogenase (*GAPDH*) gene using the $2^{-\Delta\ Ct}$ method, where $\Delta Ct = Ct_{[targetgene]}\text{-}Ct_{[GAPDH]}$. The negative control was the cells transfected with pCMV-Myc.

## Statistical Analysis

Results were expressed as the mean ± standard error of the mean (SEM), and all experiments were performed in triplicate. The equity of variance was tested using the Hartley F test and the result showed that the variance was homogeneous. In our present study, every group had three samples, which was not enough for normality testing. Because of the continuity and regularity of gene expression, we considered that the data were normally distributed. The statistical significance of differences was evaluated with the student's *t*-test using SAS 9.1.3 (SAS Institute lnc., NC). Statistical significance was indicated by $^*P < 0.05$, $^{**}P < 0.01$, or different letters above error bars indicating a statistical significance ($P < 0.05$).

# RESULTS

## Sheep FST Promoter Contains a Conserved Putative STAT3 Binding Site

There is a genomic gap immediately upstream of the sheep *FST* gene according to *Ovis aries* reference genome (ISGC Oar_v3.1/oviAri3). To obtain the promoter sequence of the sheep *FST* gene, we first closed the genomic gap by genome walking. The results showed that the gap sequence was 775 bp in length and we submitted the sequence to GenBank (Accession No. MT917184). The complete genomic sequence immediately upstream of the initiation codon (ATG) of the sheep *FST* gene was obtained by sequence assembly using the acquired genomic gap sequence and the genomic sequence from the UCSC Genome Browser database (see text footnote 1). Then we performed sequence alignment of *FST* promoters (3-kb genomic sequences immediately upstream of the initiation codon (ATG) of *FST* genes) from various animal species, which were obtained from the UCSC Genome Browser database (see text footnote 1), including sheep, cow, pig, human, mouse, and rat. The result showed that there was a conserved region, which was located at the −1,900/−1 region of sheep *FST* gene promoter. Using Mulan website tool (see text footnote 2), several putative binding sites for transcription factors, such as homeobox A4 (HOXA4), E2F transcription factor 2 (E2F2), hepatocyte nuclear factor 4 (HNF4), and STAT3, were predicted within the conserved region of the sheep *FST* gene promoter. As shown in **Figure 1A**, sequence alignment showed that a putative STAT3 binding site was conserved among various animal species. Interestingly, STAT3 interested us. STAT3 belongs to the signal transduction and transcription activation factor family in cell signal activation and transduction (Kong et al., 2019). It has been shown that STAT3 plays a vital role in the hair follicle and morphogenesis and development (Sano et al., 2008; Yu et al., 2014; Nelson et al., 2015; Gong et al., 2018, 2020). Whether STAT3 regulates *FST* is not clear.

## STAT3 Inhibits the FST Promoter Activity

To test the hypothesis that STAT3 directly regulates *FST* gene expression, firstly, we constructed and verified the *STAT3* expression vector, pCMV-Myc-STAT3 by western blotting (**Figure 1B**). Subsequently, dual-luciferase reporter assays were performed. The promoter reporter gene assay showed that, as expected, both pGL3-basic and pGL3-FST(−340/−980), as a negative control, had very lower luciferase activity, and no

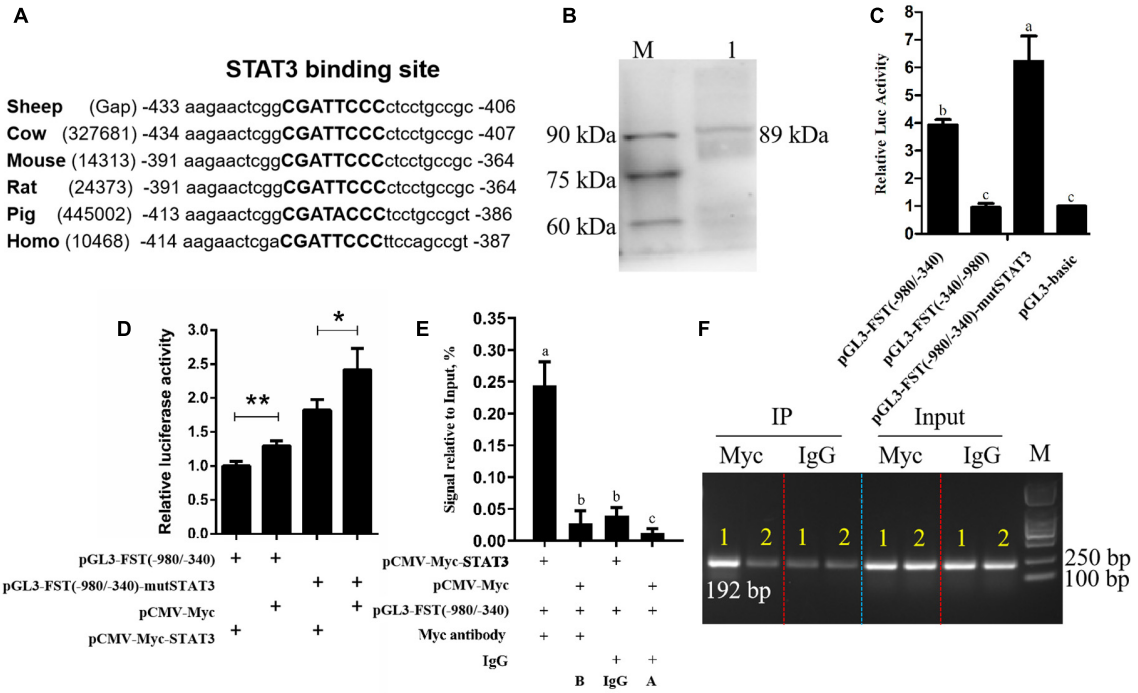


**FIGURE 1 |** Effects of STAT3 overexpression on *FST* gene promoter activity. **(A)** Conservation analysis of STAT3 binding sites in *FST* promoter among various species, STAT3 binding sites are announced by capital letters and their locations are relative to the first nucleotide (A, + 1) of the initiation codon (ATG) of the *FST* gene. **(B)** Western blot recognition of the STAT3 expression vector (pCMV-Myc-STAT3). Lane 1: the lysate of the cells transfected with pCMV-Myc-STAT3 (89 kDa); M: Protein markers (25 kDa-90 kDa). **(C)** The effect of the mutation of STAT3 binding site on *FST* promoter activity. **(D)** Effects of STAT3 on *FST* gene promoter activity. **(E)** ChIP-qPCR analysis of the binding of STAT3 to the *FST* promoter. **(F)** The agarose gel electrophoresis analysis of ChIP-qPCR products. Lane 1: The qPCR products generated from the immunoprecipitated DNA which was isolated from the cells co-transfected with pCMV-Myc-STAT3 and pGL3-FST(–980/–340); Lane 2: The qPCR products generated from the immunoprecipitated DNA which was isolated from the cells co-transfected with pCMV-Myc and pGL3-FST(–980/–340). All data are representative of three independent experiments and display as the mean ± SEM. For each figure layer, statistical significance was indicated by $^{*}P < 0.05$, $^{**}P < 0.01$, and different letters above error bars indicated a statistical significance ($P < 0.05$).

difference in luciferase activity was observed between them ($P > 0.05$, **Figure 1C**). The luciferase activities of pGL3-FST(−980/−340) and pGL3-FST(−980/−340)-mutSTAT3 were 3.39- and 6.23-fold, respectively, higher than that of pGL3-basic ($P < 0.05$, **Figure 1C**). Moreover, the luciferase activity of pGL3-FST(−980/−340)-mutSTAT3 was significantly higher than that of pGL3-FST(−980/−340) ($P < 0.05$, **Figure 1C**). These data suggest that the −980/−340 region has promoter activity and that STAT3 inhibits sheep FST promoter activity.

Further co-transfection analysis showed that the luciferase activity of pGL3-FST(−980/−340) was significantly reduced by 22.83% in the cells co-transfected with pCMV-Myc-STAT3, as compared with the cells co-transfected with pCMV-Myc ($P < 0.05$, **Figure 1D**). Consistent with the above mutation analysis result (**Figure 1C**), this result also supports that STAT3 inhibits sheep FST promoter activity.

Furthermore, to test whether STAT3 directly regulates sheep *FST* promoter, the pGL3-FST(−980/−340) and either pCMV-Myc-STAT3 or pCMV-Myc were co-transfected into HEK293T cells, and chromatin immunoprecipitation (ChIP) assay was employed with anti-Myc antibody or mouse IgG (negative control). The ChIP-qPCR results exhibited that the *FST* promoter fragment (−547/−356) was significantly enriched (6.16, 20.55, and 8.89-fold, respectively) in the DNA immunoprecipitated by the anti-Myc antibody related to negative controls (mouse IgG, A and B) ($P < 0.05$, **Figure 1E**). Consistent with the ChIP-qPCR results, agarose gel electrophoresis analysis showed that, compared with negative controls (mouse IgG, A, and B), more PCR products (−547/−356 region of *FST* promoter) were obtained from the DNA fragments immunoprecipitated by the anti-Myc antibody (**Figure 1F**). In summary, these data indicated that STAT3 directly binds to and negatively regulates the *FST* promoter.

## STAT3 and FST Have Opposite Effects on Cell Proliferation

To test whether FST mediates the roles of STAT3 in cell proliferation, we constructed and confirmed the *FST* expression vector (pCMV-Myc-FST) by western blotting (**Figure 2A**), and investigated the effects of overexpression of *STAT3* and *FST* on cell proliferation using the CCK-8 assay. The results showed that the absorbance of both the SFFs and HaCaT cells transfected with pCMV-Myc-FST was significantly higher than those transfected with pCMV-Myc at 96 h of transfection ($P < 0.01$, **Figures 2B,C**), suggesting that *FST* promotes the

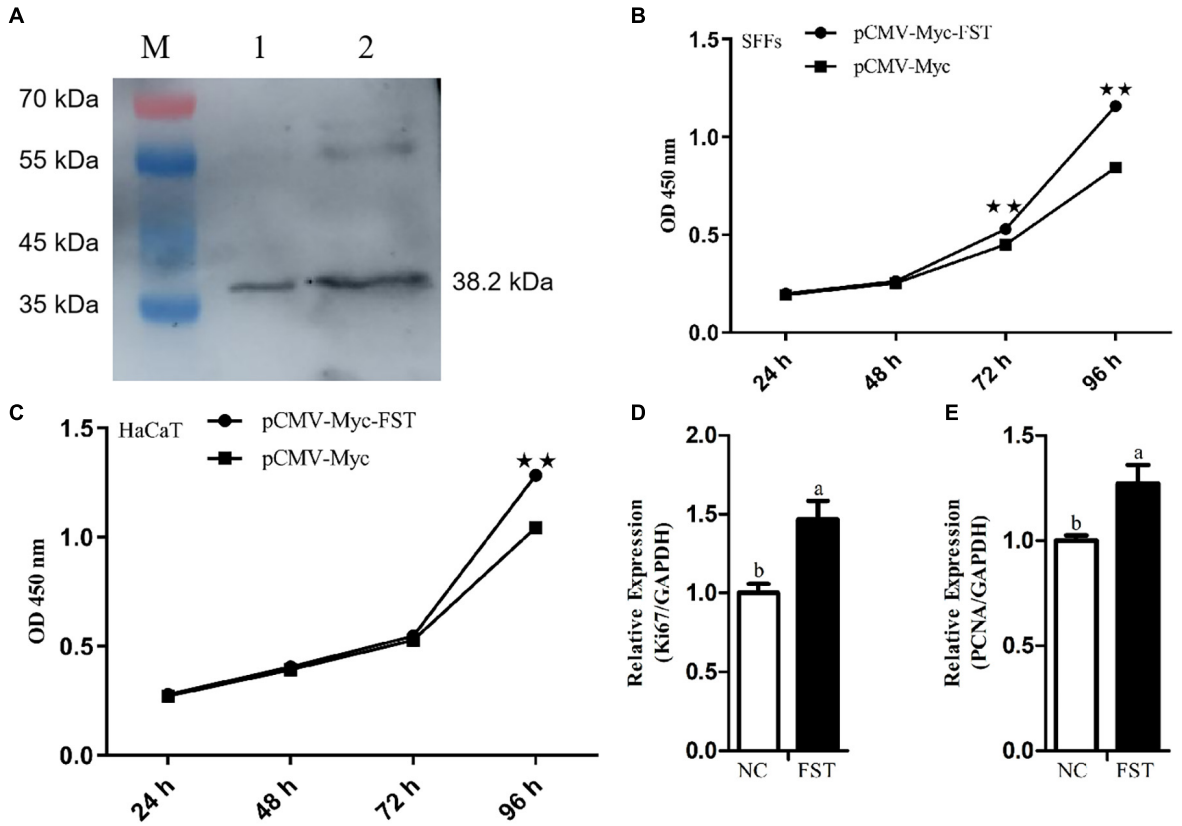


**FIGURE 2 |** Effects of *FST* overexpression on cell proliferation. **(A)** Western blot recognition of the *FST* expression vector (pCMV-Myc-FST). Lanes 1 and 2: the lysate of the cells transfected with pCMV-Myc-FST (38.2 kDa). **(B,C)** Effects of *FST* overexpression on the proliferation of SFFs and HaCaT cells. **(D,E)** Expression of *Ki67* and *PCNA* in the SFFs transfected with pCMV-Myc-FST. Fold change was calculated referring to the expression of the SFFs transfected with pCMV-Myc at 48 h. All data are representative of three independent experiments and display as the mean ± SEM. For each figure layer, statistical significance was indicated by $^{*}P < 0.05$, $^{**}P < 0.01$, and different letters above error bars indicated a statistical significance ($P < 0.05$).

proliferation of SFFs and HaCaT cells. In contrast, the absorbance of both the SFFs and HaCaT cells transfected with pCMV-Myc-STAT3 was significantly lower than those transfected with pCMV-Myc at 48 h and 72 h ($P < 0.01$, **Figures 3A,B**), suggesting that STAT3 represses the proliferation of SFFs and HaCaT cells. Consistently, *FST* overexpression significantly promoted *Ki67* and *PCNA* expression in the SFFs ($P < 0.05$, **Figures 2D,E**), while *STAT3* overexpression significantly inhibited *Ki67* and *PCNA* expression in the SFFs, compared with the cells transfected with pCMV-Myc at 48h ($P < 0.05$, **Figures 3C,D**). Further gene expression analysis showed *STAT3* overexpression significantly reduced the endogenous *FST* expression in both SFFs and HaCaT cells by 76.39 and 71.36%, respectively, compared with the cells transfected with pCMV-Myc at 48 h ($P < 0.05$, **Figures 3E,F**).

## DISCUSSION

In this study, we revealed that STAT3 directly negatively regulates sheep *FST* gene. Our evidence is as follows: (1) Bioinformatics analysis showed that *FST* promoter harbored a conserved putative STAT3 binding site (**Figure 1A**). (2) The luciferase reporter assay showed that mutation of STAT3 binding site led to an increase in the *FST* promoter activity and that STAT3 inhibited the *FST* promoter activity (**Figures 1C,D**). (3) The ChIP-qPCR assay showed that STAT3 directly bound to the *FST* promoter (**Figures 1E,F**). (4) Further functional analysis showed that *FST* and *STAT3* overexpression had opposite effects on the proliferation of SFFs and HaCaT cells (**Figures 2**, **3**) and that *STAT3* overexpression inhibited the endogenous *FST* expression in SFFs and HaCaT cells (**Figures 3E,F**). Moreover, Many target genes of STAT3 have been identified, such as forkhead box L2 (*FOXL2*) (Han et al., 2017), interleukin 17A (*IL-17A*) (Kunkl et al., 2019), interferon regulatory factor 4 (IR*F-4*), and B cell lymphoma 6 (*Bcl-6*) (Chen et al., 2019). To our knowledge, for the first time, we demonstrated that *FST* is a bona fide target gene of STAT3 and that STAT3 directly negatively regulates the *FST* gene and inhibits cell proliferation.

In the present study, the bioinformatics analysis showed that besides STAT3, several transcription factors had their binding sites in sheep *FST* gene promoter, such as HOXA4, E2F2, and HNF4. Previous studies have demonstrated that HOXA4 and E2F2 were involved in the development of epidermis and dermis, as well as hair follicles (Stelnicki et al., 1998; Lorz et al., 2010).

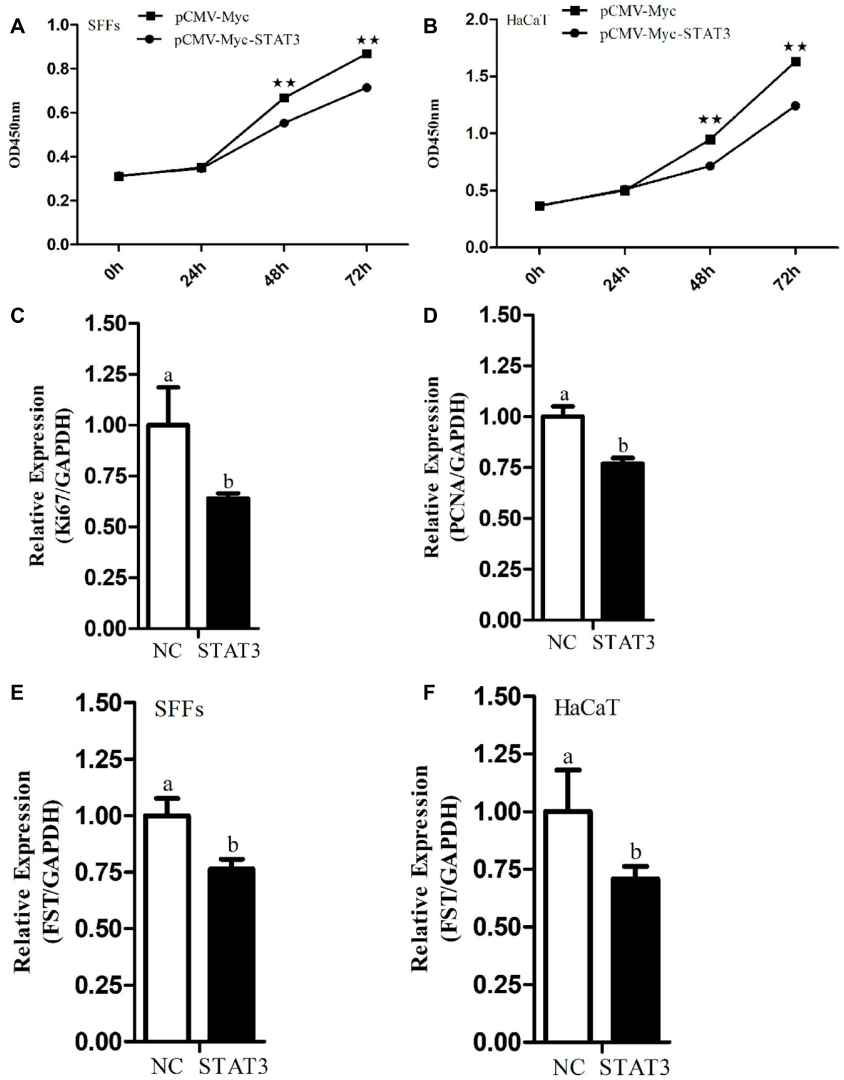


**FIGURE 3 |** Effects of *STAT3* gene overexpression on cell proliferation and endogenous *FST* expression. **(A,B)** Effect of *STAT3* overexpression on the proliferation of SFFs and HaCaT cells. **(C,D)** Expression levels of *Ki67* and *PCNA* in the SFFs transfected with pCMV-Myc STAT3. **(E,F)** Effects of STAT3 overexpression on the endogenous expression of *FST* gene in SFFs and HaCaT cells. Fold change was calculated referring to the expression of the SFFs transfected with pCMV-Myc at 48 h. All data are representative of three independent experiments and display as the mean ± SEM. For each figure layer, statistical significance was indicated by $^{*}P < 0.05$, $^{**}P < 0.01$, and different letters above error bars indicated a statistical significance ($P < 0.05$).

To better understand the transcriptional regulation of the *FST* gene in sheep hair follicles, it is worth investigating the regulation of the sheep *FST* gene by these predicted transcription factors as well.

In the present study, we found that the STAT3 negatively regulated *FST* gene and inhibited cell proliferation (**Figures 1–3**). Considering that transcription factors have numerous target genes, we cannot eliminate the probability that STAT3 inhibits cell proliferation partly by regulating the expression of its other target genes. Interestingly, a partial inhibitory repercussion of STAT3 on the promoter activity of pGL3-FST(−980/−340)-mutSTAT3 was observed, as compared with the cells co-transfected with pCMV-Myc ($P < 0.05$, **Figure 1D**). This may be dual for several reasons. Firstly, STAT3 may bind to its non-canonical binding sites in sheep *FST* promoter and inhibit *FST* promoter activity. Secondly, STAT3 may indirectly regulate *FST* promoter activity through regulation of the expression of the transcription factors which have binding sites in the *FST* promoter. Lastly, STAT3 may indirectly regulate *FST* promoter activity by interaction with some transcription factors, which have binding sites in the *FST* promoter. Further study is required to determine the precise mechanism underlying the partial inhibitory effect of STAT3 on the reporter pGL3-FST(−980/−340)-mutSTAT3 in the future.

In the present study, we demonstrated that sheep *FST* overexpression promoted SFFs and HaCaT cell proliferation (**Figures 2B–E**). In agreement with our results, it has been shown that FST promotes the proliferation of duck primary

myoblasts (Li et al., 2014). Moreover, *FST* overexpression promoted satellite cell proliferation and stimulated muscle fiber hypertrophy in mice (Gilson et al., 2009) and duck (Liu et al., 2012). The knock-down of *FST* significantly reduced the proliferation of the immortalized ovarian surface epithelial and human ovarian carcinoma cell line SKOV3 (Karve et al., 2012). Previous studies showed *STAT3* overexpression inhibited the proliferation of mouse leukocyte and hepatocyte *via* inhibiting *cyclin D* expression (Lee et al., 2002; Matsui et al., 2002), as well as chondrogenic cell line ATDC5 (Suemoto et al., 2007). In agreement, our results showed that *STAT3* overexpression inhibited the proliferation of SFFs and HaCaT cells (**Figures 3A–D**). However, it has been shown that *STAT3* overexpression has been shown to promote human breast cancer (Bromberg et al., 1999; Ma et al., 2020) and thyroid carcinoma (Kong et al., 2019) cell proliferation. The different effects of STAT3 overexpression on cell proliferation suggest that STAT3 may play different roles in cell proliferation, depending on cell type, cellular context, and species.

Accumulating evidence has demonstrated that STAT3 and FST function in hair follicle morphogenesis and development. STAT3 activation is a prerequisite for the early anagen of hair follicles (Sano et al., 2008) and keratinocytes-specific STAT3 knockout mice exhibited impaired hair cycle (Sano et al., 1999). Additionally, STAT3 can maintain keratinocyte stem/progenitor cell homeostasis *via* facilitating the maturation of the bulge region in mouse hair follicle development (Sano et al., 2000, 2008; Rao et al., 2015; Shibata et al., 2015; Nelson et al., 2016). FST promotes hair follicle development *via* binding activins and preventing the activation of activin receptors (Mcdowall et al., 2008). FST knockout mice displayed thin and curlier vibrissae (Matzuk et al., 1995; Nakamura et al., 2003), and FST transgenic mice exhibited smaller hair follicles and rough and irregular pelage (Wankell et al., 2001). Our previous study showed that sheep *FST* gene polymorphisms were associated with wool quality traits (Ma et al., 2017). Given these previous reports and our previous and present results, we hypothesize that STAT3 controls sheep hair follicle development at least in part *via* direct negative regulation of *FST* expression. Considering the STAT3 binding site in *FST* promoter are conserved across different species, we presume that our results may not be limited to sheep.

There are several limitations in the present study. First, SFFs and HaCaT cell lines were used for transcriptional regulation and function of sheep STAT3 and FST in hair follicle morphogenesis and development. These two cell lines do not originate from hair follicles and the HaCaT cell line is a non-sheep cell line. Both of these two cell lines may not be the best *in vitro* model for our study. However, hair follicles consist of mesenchymal cells and epithelial cells. SFFs are a type of mesenchymal cells (Martin, 1997; Haniffa et al., 2009), and HaCaT cells, a spontaneously immortalized, human keratinocyte line, represent epithelial cells (Deyrieux and Wilson, 2007; Wilson, 2014). These two types of cell lines may reflect some extent the *in vivo* situation of hair follicles. Additionally, these cell lines are widely used to study hair follicle morphogenesis and development (Inui et al., 2000; Ahmed et al., 2011; Luanpitpong et al., 2011; Nakamura et al., 2014; Koobatian et al., 2015). Second, only a 3-kb promoter fragment upstream of the sheep *FST* gene was used for promoter analysis, and the distal promoter region of the sheep *FST* gene was not investigated. Third, only the *in vitro* study was performed in our study. An *In vivo* study needs to be carried out to investigate the regulation of FST by STAT3 in sheep hair follicle development and wool trait formation. Nevertheless, even though there are several limitations in our present study, our results suggests that STAT3 regulates *FST* gene in sheep.

## CONCLUSION

In summary, in the present study, we closed the genomic gap upstream of sheep genomic *FST* gene (Accession No. MT917184) and demonstrated that STAT3 inhibits the proliferation of SFFs and HaCaT cells at least in part *via* direct negative regulation of *FST* gene expression. Our findings will contribute to an understanding of the *FST* transcriptional regulation and the molecular mechanisms underlying hair follicle development. To gain a better understanding of the mechanisms underlying sheep hair follicle development and morphogenesis, *in vivo* studies will be needed to validate the regulatory relationship between STAT3 and FST in sheep hair follicle development.

## 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 all animal works were conducted according to the guidance for the care and use of experimental animals established by the Ministry of Science and Technology of the People's Republic of China (Approval number: 2006-398) and approved by the Laboratory Animal Management Committee of Northeast Agricultural University.

## AUTHOR CONTRIBUTIONS

NW designed the study and provided funding support. HX and GM carried out the experiments, analyzed data, and wrote the first draft of the manuscript. FM and BN contributed to the subject discussion. HL and NW critically revised the manuscript. All authors reviewed and approved the final version of the manuscript.

## FUNDING

This work was supported by Domain-Specific projects for transgenic biological breeding (2014ZX08009-002 and 2009ZX08009-160B).

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2021.678667/full#supplementary-material

**Supplementary Figure 1 |** Determination of the genomic gap upstream of sheep *FST* gene by genome walking. The three reverse primers, *FST*-SP1, *FST*-SP2, and *FST*-SP3 were designed according to the published sequences (GCF_002742125.1) on the NCBI website, and three forward primers, ZFP2, ZSP1, and ZSP2 were provided by KX Genome Walking Kit (Zoman Biotechnology, China). The primer pairs ZFP2/*FST*-SP1, ZSP1/*FST*-SP2, and ZSP2/*FST*-P3 were used to conduct the first-, second-, and third-round PCRs for closing the genomic gap upstream of sheep *FST* gene according to the manufacturer's directions (more detailed information was shown in "MATERIALS AND METHODS").

**Supplementary Table 1 |** Primers used in this study.

**Supplementary Table 2 |** Original data.

## REFERENCES


Ahmed, M. I., Mardaryev, A. N., Lewis, C. J., Sharov, A. A., and Botchkareva, N. V. (2011). MicroRNA-21 is an important downstream component of BMP signalling in epidermal keratinocytes. *J. Cell Sci.* 124, 3399–3404. doi: 10.1242/jcs.086710

Alonso, L., and Fuchs, E. (2006). The hair cycle. *J. Cell Sci.* 119, 391–393. doi: 10.1242/jcs02793

Bromberg, J., Wrzeszczynska, M., Devgan, G., Zhao, Y., Pestell, R., Albanese, C., et al. (1999). STAT3 as an oncogene. *Cell* 98, 295–303. doi: 10.1016/s0092-8674(00)81959-5

Chen, W., Nyuydzefe, M. S., Weiss, J. M., Zhang, J. Y., Waksal, S. D., and Zanin-Zhorov, A. (2019). ROCK2, but not ROCK1 interacts with phosphorylated STAT3 and co-occupies TH17/TFH gene promoters in TH17-activated human T cells. *Sci. Rep.* 8:16636. doi: 10.1038/s41598-018-35109-9

Deng, X., Zhang, W., O-Sullivan, I., Williams, J. B., Dong, Q., Park, E. A., et al. (2012). FoxO1 inhibits sterol regulatory element-binding protein-1c (SREBP-1c) gene expression via transcription factors Sp1 and SREBP-1c. *J. Biol. Chem.* 287, 20132–20143. doi: 10.1074/jbc.M112.347211

Deyrieux, A. F., and Wilson, V. G. (2007). In vitro culture conditions to study keratinocyte differentiation using the HaCaT cell line. *Cytotechnology* 54, 77–83. doi: 10.1007/s10616-007-9076-1

Eichberger, T., Kaser, A., Pixner, C., Schmid, C., Klingler, S., Winklmayr, M., et al. (2008). GLI2-specific transcriptional activation of the bone morphogenetic protein/activin antagonist follistatin in human epidermal cells. *J. Biol. Chem.* 283, 12426–12437. doi: 10.1074/jbc.M707117200

Gilson, H., Schakman, O., Kalista, S., Lause, P., Tsuchida, K., and Thissen, J. P. (2009). Follistatin induces muscle hypertrophy through satellite cell proliferation and inhibition of both myostatin and activin. *Am. J. Physiol. Endocrinol. Metab.* 297, E157–E164. doi: 10.1152/ajpendo.00193.2009

Gong, L., Xiao, J., Li, X., Li, Y., Gao, X., and Xu, X. (2020). IL-36α promoted wound induced hair follicle neogenesis via hair follicle stem/progenitor cell proliferation. *Front. Cell Dev. Biol.* 8:627. doi: 10.3389/fcell.2020.00627

Gong, L., Xu, X. G., and Li, Y. H. (2018). Embryonic-like regenerative phenomenon: wound-induced hair follicle neogenesis. *Regen. Med.* 13, 729–739. doi: 10.2217/rme-2018-0028

Green, M. R., and Sambrook, J. (2014). *Molecular Cloning: A Laboratory Manual*, 4th Edn. New York: Cold Spring Harbor Laboratory Press.

Guo, Q., Kumar, T. R., Woodruff, T., Hadsell, L. A., DeMayo, F. J., and Matzuk, M. M. (1998). Overexpression of mouse follistatin causes reproductive defects in transgenic mice. *Mol. Endocrinol.* 12, 96–106. doi: 10.1210/mend.12.1.0053

Han, X., Jin, Y., Tan, L., Kosciuk, T., Lee, J. S., and Yoon, J. K. (2014). Regulation of the follistatin gene by RSPO-LGR4 signaling via activation of the WNT/β-catenin pathway in skeletal myogenesis. *Mol. Cell. Biol.* 34, 752–764. doi: 10.1128/MCB.01285-13

Han, Y., Wang, T., Sun, S., Zhai, Z., and Tang, S. (2017). Cloning of the promoter region of a human gene, FOXL2, and its regulation by STAT3. *Mol. Med. Rep.* 16, 2856–2862. doi: 10.3892/mmr.2017.6914

Haniffa, M. A., Collin, M. P., Bukley, C. D., and Dazzi, F. (2009). Mesenchymal stem cells: the fibroblasts' new clothes? *Haematologica* 94, 258–263. doi: 10.3324/haematol.13699

Inui, S., Itami, S., Pan, H. J., and Chang, C. (2000). Lack of androgen receptor transcriptional activity in human keratinocytes. *J. Dermatol. Sci.* 23, 87–92. doi: 10.1016/s0923-1811(99)00091-2

Jhaveri, S., Erzurumlu, R. S., Chiaia, N., Kumar, T. R., and Matzuk, M. M. (1998). Defective whisker follicles and altered brainstem patterns in activin and follistatin knockout mice. *Mol. Cell. Neurosci.* 12, 206–219. doi: 10.1006/mcne.1998.0710

Jones, A. E., Price, F. D., Le Grand, F., Soleimani, V. D., Dick, S. A., Megeney, L. A., et al. (2015). Wnt/β-catenin controls follistatin signalling to regulate satellite cell myogenic potential. *Skelet. Muscle* 5:14. doi: 10.1186/s13395-015-0038-6

Karve, T. M., Anju, P., Sneed, R., Salamanca, C., Li, X., Xu, J. W., et al. (2012). BRCA1 regulates follistatin function in ovarian cancer and human ovarian surface epithelial cells. *PLoS One* 7:e37697. doi: 10.1371/journal.pone.0037697

Kong, D., Li, A., Liu, Y., Cui, Q., Wang, K., Zhang, D., et al. (2019). SIX1 activates STAT3 signaling to promote the proliferation of thyroid carcinoma via EYA1. *Front. Oncol.* 9:1450. doi: 10.3389/fonc.2019.01450

Koobatian, M. T., Liang, M. S., Swartz, D. D., and Andreadis, S. T. (2015). Differential effects of culture senescence and mechanical stimulation on the proliferation and leiomyogenic differentiation of MSC from different sources: implications for engineering vascular grafts. *Tissue Eng. Part A* 21, 1364–1375. doi: 10.1089/ten.TEA.2014.0535

Kunkl, M., Mastrogiovanni, M., Porciello, N., Caristi, S., Monteleone, E., Arcieri, S., et al. (2019). CD28 individual signaling up-regulates human IL-17A expression by promoting the recruitment of RelA/NF-κB and STAT3 transcription factors on the proximal promoter. *Front. Immunol.* 10:864. doi: 10.3389/fimmu.2019.00864

Lee, C. K., Raz, R., Gimeno, R., Gertner, R., Wistinghausen, B., Takeshita, K., et al. (2002). STAT3 is a negative regulator of granulopoiesis but is not required for G-CSF-dependent differentiation. *Immunity* 17, 63–72. doi: 10.1016/s1074-7613(02)00336-9

Li, X., Liu, H., Wang, H., Sun, L., Ding, F., Sun, W., et al. (2014). Follistatin could promote the proliferation of duck primary myoblasts by activating PI3K/Akt/mTOR signalling. *Biosci. Rep.* 34:e00143. doi: 10.1042/BSR20140085

Li, Z., Ryu, S. W., Lee, J., Choi, K., Kim, S., and Choi, C. (2016). Protopanaxatirol type ginsenoside Re promotes cyclic growth of hair follicles via inhibiting transforming growth factor β signaling cascades. *Biochem. Biophys. Res. Commun.* 470, 924–929. doi: 10.1016/j.bbrc.2016.01.148

Lin, C., Zhao, X., Sun, D., Zhang, L., Fang, W., Zhu, T., et al. (2016). Transcriptional activation of follistatin by Nrf2 protects pulmonary epithelial cells against silica nanoparticle-induced oxidative stress. *Sci. Rep.* 6:21133. doi: 10.1038/srep21133

Liu, H., Wang, J., Yu, H. Y., Zhang, R., Chen, X., Jin, H., et al. (2012). Injection of duck recombinant follistatin fusion protein into duck muscle tissues stimulates satellite cell proliferation and muscle fiber hypertrophy. *Appl. Microbiol. Biotechnol.* 94, 1255–1263. doi: 10.1007/s00253-012-3908-4

Liu, J. J., Xu, Y. X., Wu, Q. F., Ding, Q., and Fan, W. X. (2018). Sirtuin-1 protects hair follicle stem cells from TNFα-mediated inflammatory stress via activating the MAPK-ERK-Mfn2 pathway. *Life Sci.* 212, 213–224. doi: 10.1016/j.lfs.2018.10.003

Lorz, C., García-Escudero, R., Segrelles, C., Garín, M. I., Ariza, J. M., Santos, M., et al. (2010). A functional role of RB-dependent pathway in the control of quiescence in adult epidermal stem cells revealed by genomic profiling. *Stem Cell Rev. Rep.* 6, 162–177. doi: 10.1007/s12015-010-9139-0

Luanpitpong, S., Nimmannit, U., Chanvorachote, P., Leonard, S. S., Pongrakhananon, V., Wang, L., et al. (2011). Hydroxyl radical mediates cisplatin-induced apoptosis in human hair follicle dermal papilla cells and keratinocytes through Bcl-2-dependent mechanism. *Apoptosis* 16, 769–782. doi: 10.1007/s10495-011-0609-x

Ma, G., Chu, Y., Zhang, W., Qin, F., Xu, S., Yang, H., et al. (2017). Polymorphisms of *FST* gene and their association with wool quality traits in Chinese Merino sheep. *PLoS One* 12:e0174868. doi: 10.1371/journal.pone.0174868

Ma, J., Qin, L., and Li, X. (2020). Role of STAT3 signaling pathway in breast cancer. *Cell Commun. Signal.* 18:33. doi: 10.1186/s12964-020-0527-z

Martin, P. (1997). Wound healing-aiming for perfect skin regeneration. *Science* 276, 75–81. doi: 10.1126/science.276.5309.75

Matsui, T., Kinoshita, T., Hirano, T., Yokota, T., and Miyajima, A. (2002). STAT3 down-regulates the expression of cyclin D during liver development. *J. Biol. Chem.* 277, 36167–36173. doi: 10.1074/jbc.M203184200

Matzuk, M. M., Lu, N., Vogel, H., Sellheyer, K., Roop, D. R., and Bradley, A. (1995). Multiple defects and perinatal death in mice deficient in follistatin. *Nature* 374, 360–363. doi: 10.1038/374360a0

Mcdowall, M., Edwards, N. M., Jahoda, C. A. B., and Hynd, P. I. (2008). The role of activins and follistatins in skin and hair follicle development and function. *Cytokine Growth Factor Rev.* 19, 415–426. doi: 10.1016/j.cytogfr.2008.08.005

Mehta, N., Zhang, D., Li, R. Z., Wang, T., Gava, A., Parthasarathy, P., et al. (2019). Caveolin-1 regulation of Sp1 controls production of the antifibrotic protein follistatin in kidney mesangial cells. *Cell Commun. Signal.* 17:37. doi: 10.1186/s12964-019-0351-5

Müller-Röver, S., Handjiski, B., van der Veen, C., Eichmüller, S., Foitzik, K., McKay, I. A., et al. (2001). A comprehensive guide for the accurate classification of murine hair follicles in distinct hair cycle stages. *J. Invest. Dermatol.* 117, 3–15. doi: 10.1046/j.0022-202x.2001.01377.x

Nakamura, M., Matzuk, M. M., Gerstmayer, B., Bosio, A., Lauster, R., Miyachi, Y., et al. (2003). Control of pelage hair follicle development and cycling by complex interactions between follistatin and activin. *FASEB J.* 17, 497–499. doi: 10.1096/fj.02-0247fje

Nakamura, T., Yoshitomi, Y., Sakai, K., Patel, V., Fukumoto, S., and Yamada, Y. (2014). Epiprofin orchestrates epidermal keratinocyte proliferation and differentiation. *J. Cell Sci.* 127, 5261–5272. doi: 10.1242/jcs.156778

Necela, B. M., Su, W. D., and Thompson, E. A. (2008). Peroxisome proliferator-activated receptor gamma down-regulates follistatin in intestinal epithelial cells through SP1. *J. Biol. Chem.* 283, 29784–29794. doi: 10.1074/jbc.M804481200

Nelson, A. M., Katseff, A. S., Resnik, S. R., Ratliff, T. S., Zhu, A. S., and Garza, L. A. (2016). Interleukin-6 null mice paradoxically display increased STAT3 activity and wound-induced hair neogenesis. *J. Invest. Dermatol.* 136, 1051–1053. doi: 10.1016/j.jid.2015.12.043

Nelson, A. M., Reddy, S. K., Ratliff, T. S., Hossain, M. Z., Katseff, A. S., Zhu, A. S., et al. (2015). dsRNA released by tissue damage activates TLR3 to drive skin regeneration. *Cell Stem Cell* 17, 139–151. doi: 10.1016/j.stem.2015.07.008

Ovcharenko, I., Loots, G. G., Giardine, B. M., Hou, M. M., Ma, J., Hardison, R. C., et al. (2005). Mulan: multiple-sequence local alignment and visualization for studying function and evolution. *Genome Res.* 15, 184–194. doi: 10.1101/gr.3007205

Patrik, A. (2018). How to combine ChIP with qPCR. *Methods Mol. Biol.* 1689, 29–42. doi: 10.1007/978-1-4939-7380-4_3

Pispa, J., and Thesleff, I. (2003). Mechanisms of ectodermal organogenesis. *Dev. Biol.* 262, 195–205. doi: 10.1016/s0012-1606(03)00325-7

Plowman, J. E., and Harland, D. P. (2018). The follicle cycle in brief. *Adv. Exp. Med. Biol.* 1054, 15–17. doi: 10.1007/978-981-10-8195-8_2

Rao, D., Macias, E., Carbajal, S., Kiguchi, K., and DiGiovanni, J. (2015). Constitutive STAT3 activation alters behavior of hair follicle stem and progenitor cell populations. *Mol. Carcinog.* 54, 121–133. doi: 10.1002/mc.22080

Sano, S., Chan, K. S., and Digiovanni, J. (2008). Impact of STAT3 activation upon skin biology: a dichotomy of its role between homeostasis and diseases. *J. Dermatol. Sci.* 50, 1–14. doi: 10.1016/j.jdermsci.2007.05.016

Sano, S., Itami, S., Takeda, K., Tarutani, M., Yamaguchi, Y., Miura, H., et al. (1999). Keratinocyte-specific ablation of STAT3 exhibits impaired skin remodeling, but does not affect skin morphogenesis. *EMBO J.* 18, 4657–4668. doi: 10.1093/emboj/18.17.4657

Sano, S., Kira, M., Takagi, S., Yoshikawa, K., Takeda, J., and Itami, S. (2000). Two distinct signaling pathways in hair cycle induction: STAT3-dependent and -independent pathways. *Proc. Natl. Acad. Sci. U.S.A.* 97, 13824–13829. doi: 10.1073/pnas.240303097

Schneider, M. R., Schmidt-Ullrich, R., and Paus, R. (2009). The hair follicle as a dynamic miniorgan. *Curr. Biol.* 19, R132–R142. doi: 10.1016/j.cub.2008.12.005

Sengupta, D., Bhargava, D. K., Dixit, A., Sahoo, B. S., Biswas, S., Biswas, G., et al. (2014). ERR-βsignalling through FST and BCAS2 inhibits cellular proliferation in breast cancer cells. *Br. J. Cancer* 110, 2144–2158. doi: 10.1038/bjc.2014.53

Shackleford, T. J., Zhang, Q., Tian, L., Vu, T. T., Korapati, A. L., Baumgartner, A. M., et al. (2011). Stat3 and CCAAT/enhancer binding protein beta (C/EBP-beta) regulate Jab1/CSN5 expression in mammary carcinoma cells. *Breast Cancer Res.* 13:R65. doi: 10.1186/bcr2902

Shapter, F. M., and Waters, D. L. E. (2014). Genome walking. *Methods Mol. Biol.* 1099, 133–146. doi: 10.1007/978-1-62703-715-0_12

Shibata, A., Tanahashi, K., Sugiura, K., and Akiyama, M. (2015). TRPS1 haploinsufficiency results in increased STAT3 and SOX9 mRNA expression in hair follicles in trichorhinophalangeal syndrome. *Acta Derm. Venereol.* 95, 620–621. doi: 10.2340/00015555-1948

Stelnicki, E. J., Kömüves, L. G., Kwong, A. O., Holmes, D., Klein, P., Rozenfeld, S., et al. (1998). HOX homeobox genes exhibit spatial and temporal changes in expression during human skin development. *J. Invest. Dermatol.* 110, 110–115. doi: 10.1046/j.1523-1747.1998.00092.x

Stenn, K. S., and Paus, R. (2001). Controls of hair follicle cycling. *Physiol. Rev.* 81, 449–494. doi: 10.1152/physrev.2001.81.1.449

Suemoto, H., Muragaki, Y., Nishioka, K., Stao, M., Ooshima, A., Itoh, S., et al. (2007). TRPS1 regualtes proliferation and apotosis of chondrocytes through STAT3 signaling. *Dev. Biol.* 312, 572–581. doi: 10.1016/j.ydbio.2007.10.001

Tatler, A. L., Habgood, A., Porte, J., John, A. E., Stavrou, A., Hodge, E., et al. (2016). Reduced ets domain-containing protein Elk1 promotes pulmonary fibrosis via increased integrin αvβ6 expression. *J. Biol. Chem.* 291, 9540–9553. doi: 10.1074/jbc.M115.692368

Wang, X. S., Tredget, E. E., and Wu, Y. (2012). Dynamic signals for hair follicle development and regeneration. *Stem Cells Dev.* 21, 7–18. doi: 10.1089/scd.2011.0230

Wankell, M., Munz, B., Hübner, G., Hans, W., Wolf, E., Goppelt, A., et al. (2001). Impaired wound healing in transgenic mice overexpressing the activin antagonist follistatin in the epidermis. *EMBO J.* 20, 5361–5372. doi: 10.1093/emboj/20.19.5361

Wilson, V. G. (2014). Growth and differentiation of HaCaT keratinocytes. *Methods Mol. Biol.* 1195, 33–41. doi: 10.1007/7651_2013_42

Yang, H., Zhang, Z., and Xue, L. (2016). Structural characterization and functional analysis of the follistatin promoter of *Larimichthys crocea*. *DNA Cell Biol.* 35, 471–479. doi: 10.1089/dna.2015.3178

Yu, H., Lee, H., Herrmann, A., Buettner, R., and Jove, R. (2014). Revisiting STAT3 signalling in cancer: new and unexpected biological functions. *Nat. Rev. Cancer* 14, 736–746. doi: 10.1038/nrc3818

**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.

frontiers in Genetics

ORIGINAL RESEARCH
published: 28 June 2021
doi: 10.3389/fgene.2021.676791

# Day 7 Embryos Change the Proteomics and Exosomal Micro-RNAs Content of Bovine Uterine Fluid: Involvement of Innate Immune Functions

*Kazuya Kusama[1], Mohammad B. Rashid[2,3], Rasoul Kowsar[2,4], Mohamed A. Marey[2,5], Anup K. Talukder[2,6], Kentaro Nagaoka[7], Masayuki Shimada[8], Hasan Khatib[9], Kazuhiko Imakawa[10] and Akio Miyamoto[2]**

[1]Department of Endocrine Pharmacology, Tokyo University of Pharmacy and Life Sciences, Tokyo, Japan, [2]Global Agromedicine Research Center (GAMRC), Obihiro University of Agriculture and Veterinary Medicine, Obihiro, Japan, [3]Department of Physiology and Pharmacology, Hajee Mohammad Danesh Science and Technology University, Dinajpur, Bangladesh, [4]Department of Animal Sciences, College of Agriculture, Isfahan University of Technology, Isfahan, Iran, [5]Department of Theriogenology, Faculty of Veterinary Medicine, Damanhour University, Damanhour, Egypt, [6]Department of Gynecology, Obstetrics and Reproductive Health, Bangabandhu Sheikh Mujibur Rahman Agricultural University, Gazipur, Bangladesh, [7]Laboratory of Veterinary Physiology, Cooperative Department of Veterinary Medicine, Faculty of Agriculture, Tokyo University of Agriculture and Technology, Tokyo, Japan, [8]Graduate School of Integrated Sciences for Life, Hiroshima University, Higashi-Hiroshima, Japan, [9]Department of Animal and Dairy Sciences, University of Wisconsin-Madison, Madison, WI, United States, [10]Research Institute of Agriculture, Tokai University, Kumamoto, Japan

OPEN ACCESS

**Edited by:**
*Aline Silva Mello Cesar, University of São Paulo, Brazil*

**Reviewed by:**
*Guilherme Pugliesi, University of São Paulo, Brazil*
*Sayed Haidar Abbas Raza, Northwest A & F University, China*

***Correspondence:**
*Akio Miyamoto*
*akiomiya@obihiro.ac.jp*

**Specialty section:**
*This article was submitted to Livestock Genomics, a section of the journal Frontiers in Genetics*

**Received:** *06 March 2021*
**Accepted:** *28 May 2021*
**Published:** *28 June 2021*

**Citation:**
*Kusama K, Rashid MB, Kowsar R, Marey MA, Talukder AK, Nagaoka K, Shimada M, Khatib H, Imakawa K and Miyamoto A (2021) Day 7 Embryos Change the Proteomics and Exosomal Micro-RNAs Content of Bovine Uterine Fluid: Involvement of Innate Immune Functions. Front. Genet. 12:676791. doi: 10.3389/fgene.2021.676791*

This study aimed to characterize proteins and exosomal microRNAs (miRNAs) in the uterine flushings (UF) of cows associated with Day 7 (D7) pregnancy using the embryo donor cows of the embryo transfer program. Superovulated cows either were inseminated (AI cows) or remained non-inseminated (Ctrl cows). UF was collected on D7 in the presence of multiple embryos (AI cows) or without embryos (Ctrl cows) and subjected to isobaric tags for relative and absolute quantification protein analysis. A total of 336 proteins were identified, of which 260 proteins were more than 2-fold higher in AI cows than Ctrl cows. Gene ontology analysis revealed that many differentially expressed proteins were involved in "neutrophil-related" and "extracellular vesicular exosome-related" terms. In silico analysis of proteins with higher concentrations in the UF of AI identified 18 uniquely expressed proteins. Exosomes were isolated from the UF, from which RNA was subjected to miRNA-seq, identifying 37 miRNAs. Of these, three miRNAs were lower, and six miRNAs were higher in the UF of AI cows than those of Ctrl ones. The principal component analysis displayed a close association in miRNA and protein between bta-miR-29a, bta-miR-199b, SUGT1, and PPID. In addition, the receiver operating characteristic curve analysis showed that SUGT1 was the best predictor for the presence of embryos in the uterus. These findings suggest that the presence of multiple D7 embryos in the uterus can lead to significant changes in the protein composition and exosomal miRNA contents of UF, which could mediate innate immunological interactions between the pre-hatching embryo and the uterus in cows.

**Keywords: Day 7 embryo, uterine flushing, protein, exosome, miRNA, immunity, cow**

## INTRODUCTION

Up to 50% of potential bovine pregnancies are lost in the first week following insemination (Diskin and Morris, 2008; Wiltbank et al., 2016). Thus, proper interactions between the early embryo and mother are crucial for successful pregnancy in cows. Pregnancy is established by complex interactions of the factors/molecules derived from both the developing embryo and the cow's endometrium (Talukder et al., 2020). The pre-hatching bovine blastocyst can synthesize and secrete interferon-tau (IFNT), the pregnancy recognition signal in ruminants, on Day 7 (D7) of pregnancy. Subsequently, IFNT regulates the maternal immune response toward an anti-inflammatory (Th2) action to tolerate the semi-allogenic embryo (Sponchiado et al., 2017; Talukder et al., 2017; Passaro et al., 2018; Rashid et al., 2018; Sponchiado et al., 2019). In addition to IFNT, other factors from both the pre-hatching embryo or the endometrium may modulate maternal immunity. However, the nature of these interactions during the pre-hatching period of pregnancy has not yet been well characterized.

Several global analyses of bovine uterine flushing (UF) revealed some changes in intrauterine protein levels during the peri-implantation period (Forde et al., 2013, 2014a,b; Kusama et al., 2016, 2018b). Furthermore, extracellular vesicles (EVs), including exosomes, have been identified in UF (Burns et al., 2014; Nakamura et al., 2016; Kusama et al., 2018b) and found to be involved in the conceptus-endometrial interactions during the implantation period (Burns et al., 2016; Nakamura et al., 2016; Kusama et al., 2018b). In sheep, endometrial exosomes regulate IFNT production by the conceptus (Ruiz-González et al., 2015). Several studies investigated the effects of intrauterine exosomes on immunological interactions between the conceptus and endometrium during the peri-implantation period (Burns et al., 2014, 2016; Nakamura et al., 2016; Kusama et al., 2018b), but not in earlier stages, particularly during the pre-hatching period in cows.

MicroRNAs (miRNAs) are small, non-coding RNAs that regulate gene expression in various cell types (Bi et al., 2009). The pre-hatching D7 bovine blastocysts can secrete miRNAs into the extracellular environment through EVs/exosomes (Gross et al., 2017b). Given that both the early embryo and the maternal tract produce miRNAs, these miRNAs may serve as communication signals between the pre-hatching embryo and mother during early pregnancy (Gross et al., 2017b). miRNAs were found to be involved in gametogenesis, early embryo development, implantation, and placental formation (Bidarimath et al., 2014; Ioannidis and Donadeu, 2016; Gross et al., 2017a). Recent studies also indicate that miRNAs contribute to the immune tolerance at conception through regulatory effects of seminal fluid in generating tolerogenic dendritic cells and T-reg cells (Bidarimath et al., 2014; Schjenken et al., 2016).

We hypothesized that the presence of D7 embryos could induce global changes in the protein composition and exosomal contents of UF during the pre-hatching period, regulating the local uterine immunity on D7 post-insemination for the embryo tolerance, thereby ensuring the establishment of pregnancy in cows. Therefore, the present study aimed to investigate the embryo- and endometrium-derived factors involved in regulating the local immune environment of the uterus at D7 of pregnancy in cows. Additionally, bioinformatics analysis was used to identify significant proteins and miRNAs and their interrelationship associated with D7 pregnancy in cows. In the present investigation, we used a superovulation model for multiple embryo production in embryo transfer donor cows. This model could amplify embryo-derived signals in the uterus that can not be easily detected using the single-embryo model. Indeed, we previously found that interferon-stimulated genes were upregulated in circulating immune cells in the superovulation cow model (Talukder et al., 2019). Furthermore, we used the Japanese Black cows that were prepared exclusively for embryo donors and have only one-time experience of pregnancy and lactation, so their responsiveness to hormone treatment was expected to be highly homogeneous.

## MATERIALS AND METHODS

### Ethics Statement

All animal experiments were conducted in a commercial herd (Nobel's Co. Ltd., Hokkaido, Japan) under the approval of the Animal Experiments Ethics Committee, Obihiro University of Agriculture and Veterinary Medicine, Japan (Permit number 25–101).

### Experimental Design

To investigate the effect of multiple D7 embryos on uterine proteomic and exosomal miRNA contents, donor Japanese Black cows (3–7 years old) were treated for superovulation 7 days before estrus (D7). UF (20–25 ml) was collected on D7 from inseminated cows and remained non-inseminated cows on D0 (estrus). Pregnancy was confirmed by the presence of multiple embryos in the UF. The UF content was analyzed using isobaric tags for relative and absolute quantification (iTRAQ) analysis, followed by GO and enriched signaling pathway analyses. Exosomes were isolated from the UF, followed by RNA extraction and miRNA sequencing. Proteins and miRNAs were subjected to network analysis to identify associations with D7 pregnancy outcomes in cows (**Figure 1**).

### Generation and Collection of D7 Bovine Uterine Flushing

Twenty-one cows from a commercial herd (Nobel's Co. Ltd., Hokkaido, Japan) were assigned randomly to the experiment. All cows were synchronized for estrus using the superovulation regimen as described previously (Rashid et al., 2018). Based on the behavioral estrus on D0, cows were either inseminated ($n$ = 15) or remained non-inseminated ($n$ = 6). On D7, embryos were non-surgically flushed out from the uterus using Ringer solution (SOLULACT, Terumo Co., Tokyo, Japan, 500 ml/uterine horn × 2 horns) as described in Rashid et al. (2018). Briefly, a balloon catheter (Fujihira Industries Co., Tokyo, Japan) was first inserted up to the tip of the uterine horn ipsilateral to the side existing a higher number of corpora lutea in the ovary, and an inflatable cuff on the catheter was filled with air to hold the catheter in place. Flushing solution

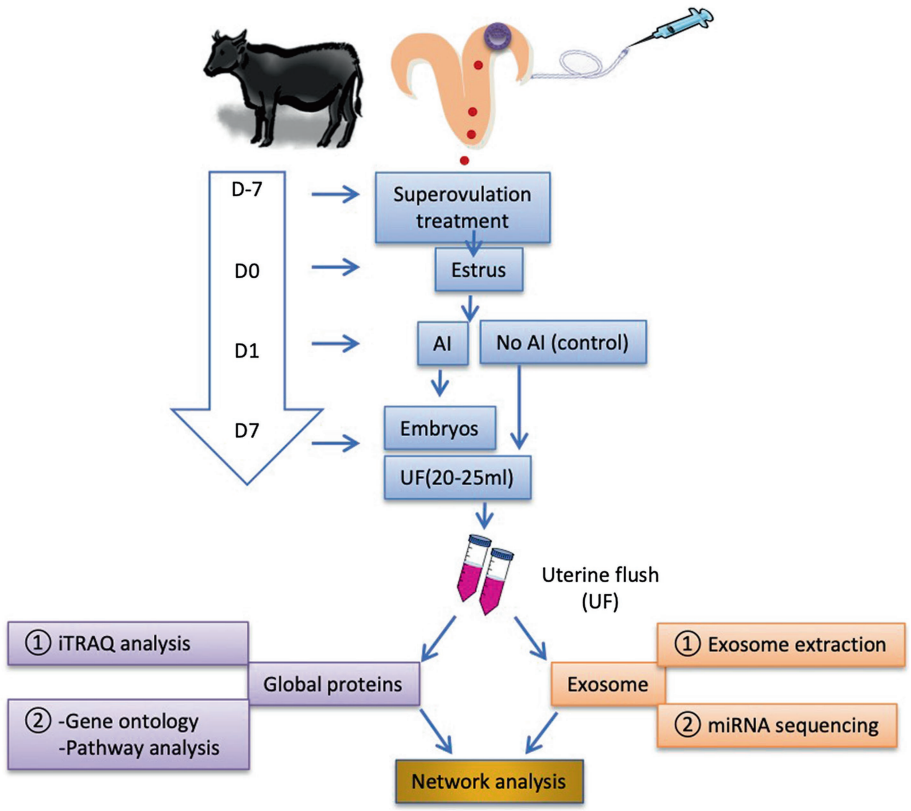


**FIGURE 1 |** Schematic illustration of the experimental design. Cows were superovulated 7 days before the estrus (D7). In estrus (D0), the cows were either inseminated (AI cows) or non-inseminated (Ctrl cows) on D1. The embryos were flushed out on D7, and the first uterine flush (UF) was collected (20–25 ml). To reduce individual variations in the animal responses and the cost for analyses, the equal volume of UF samples was pooled together (UF from each three AI cows and each two Ctrl cows were pooled as a single sample, respectively; thus, three pooled AI UF from nine cows and three pooled Ctrl UF from six cows) and stored at −80°C until use. The number of recovered embryos with good quality per AI cow was 9.5 ± 1.4 (mean ± SEM). Global UF protein analysis was performed using isobaric tags for relative and absolute quantification (iTRAQ) analysis followed by GO and enriched signaling pathway analyses. Exosomes were isolated from D7 UF followed by RNA extraction and microRNA (miRNA) sequencing for the identification of exosome-derived miRNAs contents. The identified proteins and miRNAs were mapped to network analysis. AI, artificial insemination; UF, uterine flush; and D, day.

(20–25 ml) was infused inside the uterus and rapidly aspirated to avoid fluid overflow inside the lumen of the uterus and prevent injury of the soft uterine horn. The flushing process was repeated several times using small amounts of Ringer solution (totally 500 ml/uterine horn × 2 horns). Embryos were classified and graded according to the guidelines of the International Embryo Transfer Society (Bó and Mapletoft, 2014). Only inseminated cows, referred to as AI cows ($n$ = 9), with relatively large numbers of good quality embryos (Grades 1 and 2) were used in this experiment. The first 20–25 ml of UF collected from the uterine horn ipsilateral to the side existing a higher number of corpora lutea in the ovary was kept for further analysis. The Ctrl cows were exposed to a single uterine flushing process using 20–25 ml of UF. To minimize individual variation in the animal responses, three pools of UF samples from AI cows and three pools from Ctrl cows were constructed. Each pool comprised an equal volume (4 ml) of UF. Pooled samples were centrifuged at 1,000 × g for 15 min, and the supernatant was collected and stored at −80°C until use.

## iTRAQ Analysis

Global protein analysis was performed using iTRAQ analysis as described previously (Kusama et al., 2018b). Briefly, total protein (100 μg) from D7 UF samples was subjected to trypsin digestion and then reacted with appropriate iTRAQ reagent according to the manufacturer's instructions. Sample fractionation was performed with an Agilent 3100 OFFGEL Fractionator (Agilent Technologies, Santa Clara, CA, United States). Mass spectrometry analysis was performed with a Thermo Scientific LTQ Orbitrap XL mass spectrometer (Thermo Fisher Scientific, Waltham, MA, United States). Mascot software was used to identify and quantify proteins simultaneously.

## Isolation of Exosomes From the Uterine Flush

Exosomes were isolated from D7 UF using exosome precipitation solution (Exo-Quick-TC, System Biosciences, Mountain View, CA, United States) according to the manufacturer's instructions. The UF with Exo-Quick-TC was incubated overnight at 4°C

and then centrifuged at 1,500 × g for 30 min at 4°C to pellet exosomes as previously described (Nakamura et al., 2016; Kusama et al., 2018b).

## Western Blotting

Exosomes lysed with M-PER (10 μg) were separated through SDS-PAGE, then transferred onto polyvinylidene difluoride membranes (Bio-Rad, Hercules, CA, United States). After blocking with bullet blocking one (Nacalai Tesque, Kyoto, Japan), the membranes were incubated with rabbit polyclonal anti-AKR1B1 (1:500, sc-33219, Santa Cruz Biotechnology, Dallas, TX, United States), goat polyclonal anti-CAPG (1:100, sc-33084, Santa Cruz Biotechnology), rabbit polyclonal anti-HSP70 antibody (1:2,000, EXOAB-HSP70A-1, System Biosciences), rabbit polyclonal anti-CD63 (1:1,000, EXOAB-CD63A-1, System Biosciences), or rabbit polyclonal anti-RAB5 (ab13253, 1:1,000, Abcam, Tokyo, Japan). Immunoreactive bands were detected using enhanced chemiluminescence (EMD Millipore, Temecula, CA, United States) after incubation with horseradish peroxidase-labeled goat anti-rabbit IgG or horse anti-goat IgG (1:5,000, Vector Laboratories, Burlingame, CA, United States). Signals were detected using C-DiGit Blot Scanner (LI-COR; Kusama et al., 2018a). Total proteins were stained with colloidal gold total protein stain solution according to the manufacturer's instructions (Bio-Rad Laboratories, Hercules, CA, United States).

## Nanoparticle Tracking Analysis

Nanoparticle tracking analysis of exosomes, isolated from D7 UF and suspended in PBS (2–6 × $10^8$ particles/ml) was performed using a NanoSight NS300 (NanoSight Ltd., Amesbury, United Kingdom) instrument with 488 nm laser and a complementary metal oxide-semiconductor camera (Andor Technology, Belfast, United Kingdom) and NanoSight NTA 3.2 software calibrated with 100 nm polystyrene beads (Thermo Fisher Scientific; Kusama et al., 2017).

## RNA Extraction and miRNA Sequencing

For miRNA-seq analysis, RNA was extracted from exosomes using the SeraMir Exosome RNA Amplification Kit (System Biosciences) according to the manufacturer's instructions. The miRNA libraries were constructed using the SMARTer smRNA-seq kit (Clontech, Tokyo, Japan) according to the manufacturer's instructions. RNA sequencing was performed on an Illumina Hiseq 2500 platform (Macrogen, Tokyo, Japan), and 50-bp, single-end reads were generated. Trimmed reads of 18–50 nt were mapped to a reference sequence by Bowtie. Mapped fragments were annotated with a known miRNA database (miRBase22.1) using feature counts. Annotated data were normalized using edgeR to the total reads of each sample as the standardized to count per million. The primary data were deposited to the (DNA Data Bank of Japan) Sequence Read Archive[1] (accession number DRA010067). Data analysis was performed as described previously (Nakamura et al., 2020).

## Statistical Analysis

Data are expressed as the mean ± SEM. Significance was assessed using ANOVA and the Dunnet comparisons test. A value of $p < 0.05$ was considered statistically significant. Gene ontology (GO) and enriched signaling pathway analyses were performed with the Enrichr tool.[2] The multivariate method, the principal component analysis (PCA), was used to estimate the association of the inter-correlated variables (i.e., an association between proteins and the presence of embryos in the UF or an association between proteins and miRNAs). Network analysis with the Pearson similarity index was performed to identify central nodes (i.e., proteins or miRNAs). The Fruchterman-Reingold algorithm was used as a force-directed layout algorithm. For PCA and network mapping, the PAST program (accessible at http://folk.uio.no/ohammer/past) was used. Using the easyROC web-tool,[3] a receiver operating characteristic (ROC) curve analysis was performed to identify predictive factors for the presence of embryos. The optimal cutoff for the presence of embryos was determined by maximizing the Youden index. Finally, the area under the curve (AUC) was determined by the asymptotic statistic (a parametric unbiased likelihood estimator) to detect UF-containing embryos.

# RESULTS

## Quantitative Changes in the Protein Abundance of D7 UF

The number of recovered embryos with good quality (Grade 1 + Grade 2) per AI cows was 12 ± 0.95. Moreover, no significant differences were detected in the numbers of corpora lutea (Ctrl cows, 11.8 ± 0.9 and AI cows, 14.7 ± 1.6) or circulating P4 concentration (Ctrl cows, 12.2 ± 1.5 ng/ml and AI cows, 13.8 ± 0.8 ng/ml). The iTRAQ analysis revealed a significant difference in protein content and level in the D7 UF from AI cows with multiple embryos vs. those of Ctrl cows without embryos. A total of 336 proteins were identified, of which 260 proteins were more than 2-fold higher (**Figure 2A**) in AI cows compared to Ctrl cows. The PCA of the 260 proteins identified 18 unique proteins in AI cows, which significantly differed from those of Ctrl cows. Additionally, the presence of embryos in the UF was positively associated with SUGT1, PPID, PDLIM1, DBI, and NMRAL1 (vectors<45°, yellow circle; **Figure 2B**). In contrast, the presence of embryos in the UF showed a negative association with PNP, EFHD1, and ENO1 (vectors approaching 180°, red circle; **Figure 2B**). The ROC curve analysis was performed to identify UF proteins that can predict the presence of embryos in the uterus on D7. The area under the ROC curve analysis (AUC) determined that SUGT1 (AUC: 0.999), PPID (AUC: 0.998), NMRAL1 (AUC:

[1]https://www.ddbj.nig.ac.jp/dra/index-e.html

[2]http://amp.pharm.mssm.edu/Enrichr/

[3]http://www.biosoft.hacettepe.edu.tr/easyROC/

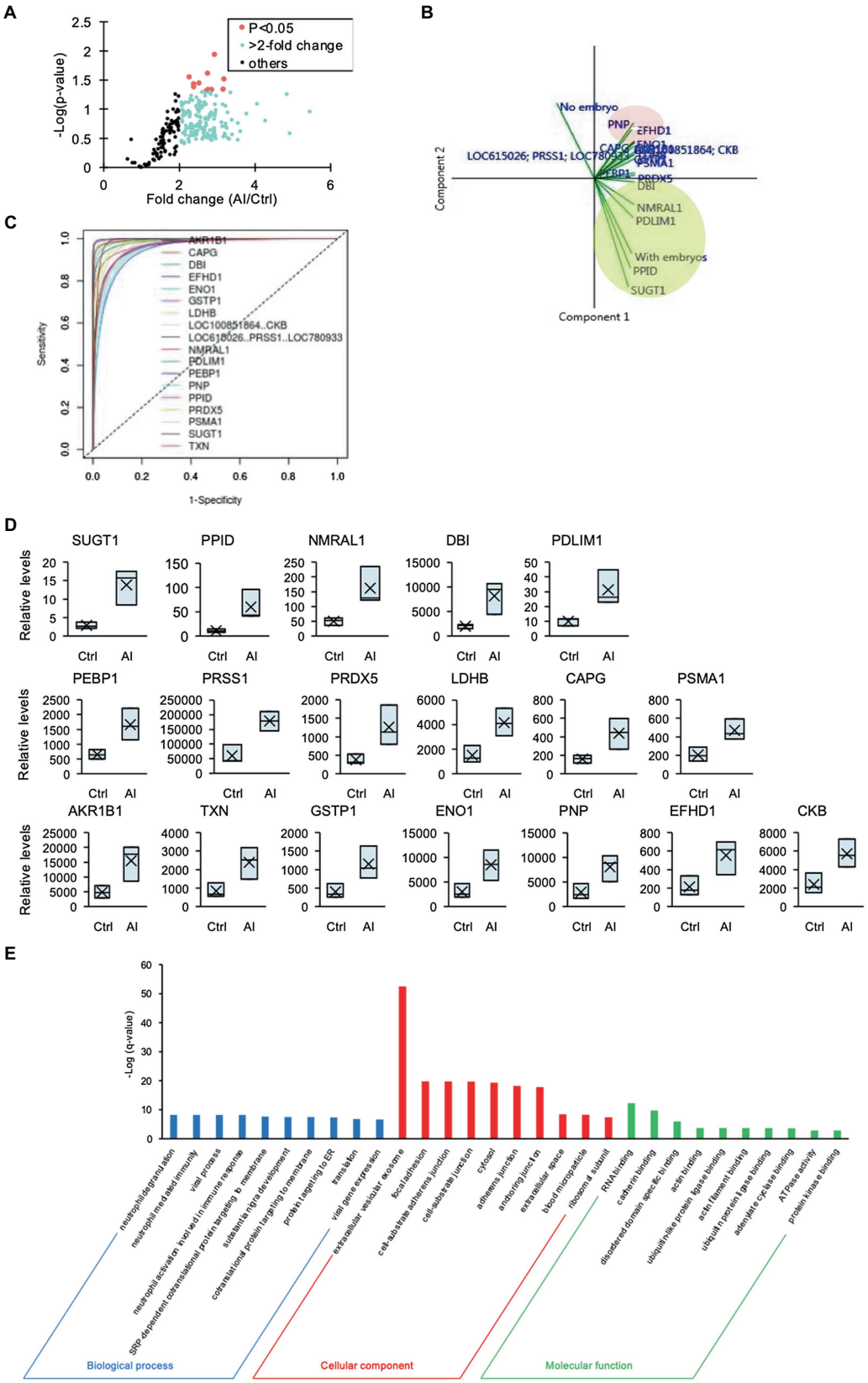


**FIGURE 2 |** The iTRAQ analysis of uterine flush (UF) collected on D7. **(A)** Volcano plot showing the protein levels found by the iTRAQ analysis. The proteins

*(Continued)*

**FIGURE 2 |** highlighted in light blue are more than 2-fold higher, and those proteins highlighted in red are significantly higher ($p < 0.05$) in UF-containing embryos obtained from inseminated (AI) cows. **(B)** Biplot produced by the principal component analysis (PCA) of factors, including higher-concentration proteins in UF of cows. Close-angles vectors (<45°) indicate a strong correlation, perpendicular vectors indicate no correlation, and vectors in opposite directions (approaching 180°) indicate a negative correlation. **(C)** Receiver operating characteristic (ROC) curve analysis was performed to assess the predictive power of variables and to measure the optimum cutoff point for the presence of embryos in UF. **(D)** Changes in identified proteins by the comparison between Ctrl and AI cows were determined by iTRAQ analysis. Upper box plots were AUC > 0.99 by ROC. **(E)** Higher-concentration proteins in UF of AI cows were functionally classified in GO analysis by the biological process (BP), cellular component (CC), and molecular function (MF) terms.

**TABLE 1 |** The list of parameters of factors identified by ROC analysis.

| Gene ID | AUC | Optimal cutoff | Sensitivity | Specificity | PPV | NPV | PLR | NLR |
|---|---|---|---|---|---|---|---|---|
| SUGT1 | 0.999 | 8.4 | 100 | 100 | 100 | 100 | INF | 0 |
| PPID | 0.998 | 41.5 | 100 | 100 | 100 | 100 | INF | 0 |
| NMRAL1 | 0.995 | 122.2 | 100 | 100 | 100 | 100 | INF | 0 |
| DBI | 0.994 | 4419.7 | 100 | 100 | 100 | 100 | INF | 0 |
| PDLIM1 | 0.992 | 22.9 | 100 | 100 | 100 | 100 | INF | 0 |
| PEBP1 | 0.987 | 1147.9 | 100 | 100 | 100 | 100 | INF | 0 |
| LOC615026; PRSS1; LOC780933 | 0.986 | 145090.9 | 100 | 100 | 100 | 100 | INF | 0 |
| PRDX5 | 0.985 | 801 | 100 | 100 | 100 | 100 | INF | 0 |
| LDHB | 0.979 | 3096.6 | 100 | 100 | 100 | 100 | INF | 0 |
| CAPG | 0.975 | 265.2 | 100 | 100 | 100 | 100 | INF | 0 |
| PSMA1 | 0.974 | 378.1 | 100 | 100 | 100 | 100 | INF | 0 |
| AKR1B1 | 0.965 | 8649.6 | 100 | 100 | 100 | 100 | INF | 0 |
| TXN | 0.965 | 1483.8 | 100 | 100 | 100 | 100 | INF | 0 |
| GSTP1 | 0.965 | 774.9 | 100 | 100 | 100 | 100 | INF | 0 |
| ENO1 | 0.96 | 5348.7 | 100 | 100 | 100 | 100 | INF | 0 |
| PNP | 0.952 | 5055.7 | 100 | 100 | 100 | 100 | INF | 0 |
| EFHD1 | 0.947 | 346.2 | 100 | 100 | 100 | 100 | INF | 0 |
| LOC100851864; CKB | 0.91 | 4293.1 | 100 | 100 | 100 | 100 | INF | 0 |

0.995), DBI (AUC: 0.994), and PDLIM1 (AUC: 0.992) had a strong predictive power for the presence of embryos in the uterus (**Figures 2C,D; Table 1**).

Furthermore, GO analysis of the 260 proteins revealed significantly enriched biological process (BP), cellular component (CC), and molecular function (MF) terms, from which several neutrophil-related terms were identified in BP terms, and the most enriched CC term was "extracellular vesicular exosome" (**Figure 2E**).

## Proteins Identified in D7 UF Were Involved Mainly in Neutrophil-Mediated Immune Responses

Since the GO analysis identified "neutrophil-related terms," the iTRAQ-identified proteins were compared with the immune-related gene database, from which a heat map was generated (**Figure 3A**). The matched proteins were re-evaluated with GO BP and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analyses. The GO BP analysis identified enriched the neutrophil degranulation term, neutrophil-mediated immunity term, and neutrophil activation involved in the immune response term (**Figure 3B**). KEGG pathway analysis revealed enriched complement and coagulation cascade terms (**Figure 3C**).

## Exosomes Were Enriched in the UF-Containing Embryos on D7 Post-insemination

GO analysis identified "extracellular vesicular exosome" terms. Similar to neutrophil-related terms, we compared iTRAQ-identified proteins with the database of exosome-expressed genes and generated a heat map (**Figure 4A**). Using Western blot analysis, we further examined exosome markers HSP70, CD63, RAB5, AKR1B1, and CAPG. The Western blot analysis revealed exosomal markers present in UF, and the levels of these markers in UF-containing embryos were higher than those in UF from Ctrl cows (**Figure 4B**). These results indicate that exosomes were secreted into the uterine lumen at 7 days post-insemination. The exosomes isolated from UF in AI cows were subjected to nanoparticle tracking analysis, which showed an average size of 122.5 nm ± 35.2 nm (mean ± SEM; **Figure 4C**).

## Exosomal miRNAs Were Differentially Expressed in D7 UF-Containing Embryos

It was reported that exosomal miRNAs participate in dynamic changes in uterine gene expression patterns (Liang et al., 2017). To elucidate functional roles of miRNAs in exosomes, exosomal RNA isolated from D7 UF was subjected to miRNA-seq analysis, then compared with a database of miRNAs, followed by the generation of a heat map. The analysis detected 37 miRNAs in the intrauterine exosomes from Ctrl to AI cows (**Figure 4D**), of which three miRNAs (miR-2332, -199b, and -29a) were lowly expressed and six miRNAs (miR-1248, -12034, -2890, -11,972, -2892, and -423-5p) were highly expressed in AI vs. Ctrl cows (**Figures 4E,F**).

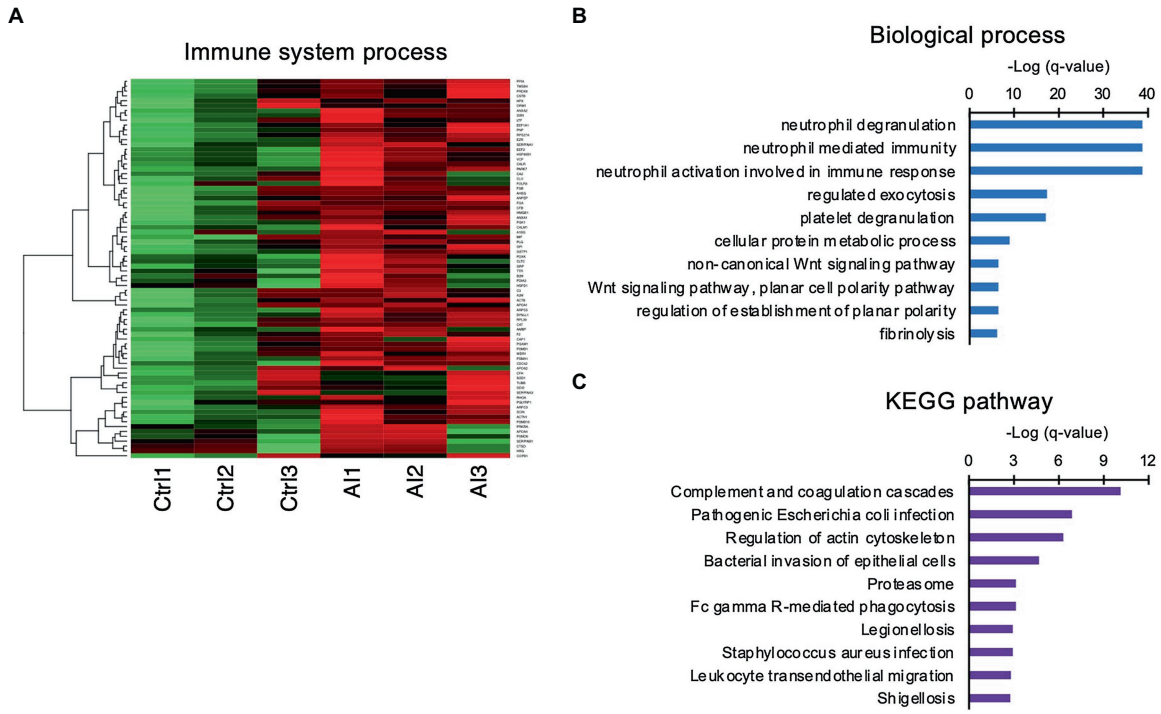


**FIGURE 3 |** Involvement of innate immune functions in the UF obtained on D7. **(A)** Heat-map study of innate immune system-related genes in UF of inseminated (AI) and non-inseminated (Ctrl) cows. High-concentration proteins are shown in red, and low-concentration proteins are shown in green. **(B,C)** Identified immune system-related genes were functionally classified in biological GO terms **(B)** and enriched pathway analyses **(C)**.

## Interaction of Exosomal miRNAs and Proteins in D7 UF-Containing Embryos

A network and PCA analyses were carried out to elucidate the relationship between miRNAs and proteins found in the UF of AI cows. To this aim, 18 proteins that significantly differed between AI cows and Ctrl cows and nine differentially expressed miRNAs were analyzed. The network analysis showed that three miRNAs, including bta-miR-29a, bta-miR-199b, and bta-miR-423-5p, were associated with many uterine proteins (**Figure 5A**). More precisely, the PCA showed that bta-miR-29a, bta-miR-199b, SUGT1, and PPID exhibited a strong and positive association with each other (vectors<45°, red ovals; **Figure 5B**). Other miRNAs, such as bta-miR-12034 and bta-miR-1246, were negatively associated with UF proteins, PDLIM1, NMRAL1, SUGT1, and PPID (vectors approaching 180°, blue oval; **Figure 5B**).

## DISCUSSION

This study characterized the proteomics and exosome-derived miRNA content in UF collected from superovulated donor cows on D7 post-insemination. The proteomic analysis revealed a remarkable difference in protein content of UF in AI cows vs. Ctrl cows, and most of the identified proteins were associated with neutrophil-mediated immune responses. The expression of exosome markers was higher in UF from AI cows, and miRNA-seq identified differential expression of exosomal miRNAs in UF of AI cows vs. Ctrl cows. These findings demonstrate that the presence of D7 bovine embryos induces significant changes in uterine luminal proteins and exosome-derived miRNAs and suggest that as early as D7, a proper embryo-maternal interaction may be required to adjust the uterine environment toward the establishment of pregnancy in cattle.

One of the key immunological events that characterize the pre-hatching period of pregnancy in ruminants is the ability of the developing embryo to express antigenic MHC molecule-I on D7 post-insemination in cows (Templeton et al., 1987; Low et al., 1990). Therefore, the maternal immune system must be in place on D7 of pregnancy to prevent rejection of the semi-allogenic embryo. Indeed, recent *in vivo* investigations showed that D7 bovine embryos could modify the endometrial transcriptome and the biochemical composition of the uterine luminal fluid in the most cranial portion of the uterine horn ipsilateral to the corpus luteum (Sponchiado et al., 2017, 2019). Recently, we reported that UF collected from superovulated donor cows on D7 post-insemination induced anti-inflammatory responses and upregulation of interferon-stimulated gene expression in peripheral blood mononuclear cells *in vitro* (Rashid et al., 2018). In the present study, 336 proteins in the D7 UF, of which 260 proteins showed over a 2-fold increase in AI cows vs. Ctrl cows. Of these, many proteins were associated with "neutrophil-related" terms and "extracellular vesicular exosome" terms, suggesting that D7 bovine embryos exist in the intrauterine innate immune microenvironment.

In particular, many proteins identified in the D7 UF were related to "neutrophil degradation," "neutrophil-mediated immunity,"

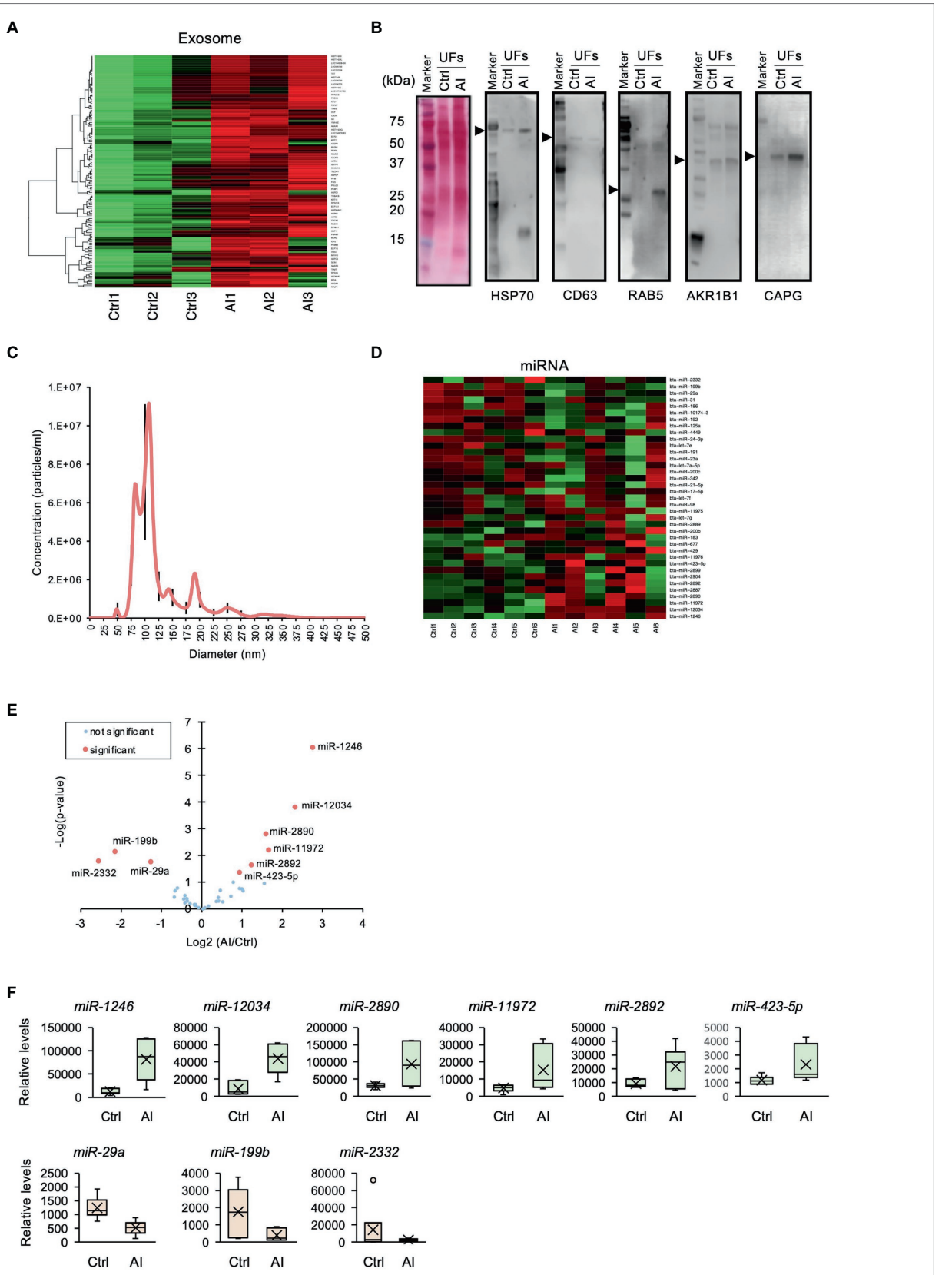


**FIGURE 4 |** Profile of exosomal miRNAs in uterine flush (UF) collected on D7 after the estrus. **(A)** Heat-map analysis of exosome-related genes found in the UF of

*(Continued)*

**FIGURE 4 |** inseminated (AI) cows and non-inseminated (Ctrl) cows. High-concentration proteins are shown in red, and low-concentration proteins are shown in green. **(B)** The UF of Ctrl and AI cows was subjected to Western blotting, which revealed the presence of exosomal markers. The arrow heads show each target band. The polyvinylidene difluoride membrane was stained with colloidal gold. Total protein stain solution confirms total loading of protein. **(C)** Nanoparticle tracking analysis of the UF of AI cows revealed an exosomal size range of 50–150 nm. Black lines are SEM. **(D)** Heat-map analysis detected 37 exosomal miRNAs in the UF of AI and Ctrl cows. Upregulated miRNAs are shown in red, and downregulated miRNAs are shown in green. **(E)** Volcano plot showing the miRNA levels identified by miRNA-seq. The nine miRNAs highlighted in red are differentially expressed ($p < 0.05$). **(F)** Changes in identified miRNAs by the comparison between Ctrl and AI cows were determined by miRNA-seq analysis.

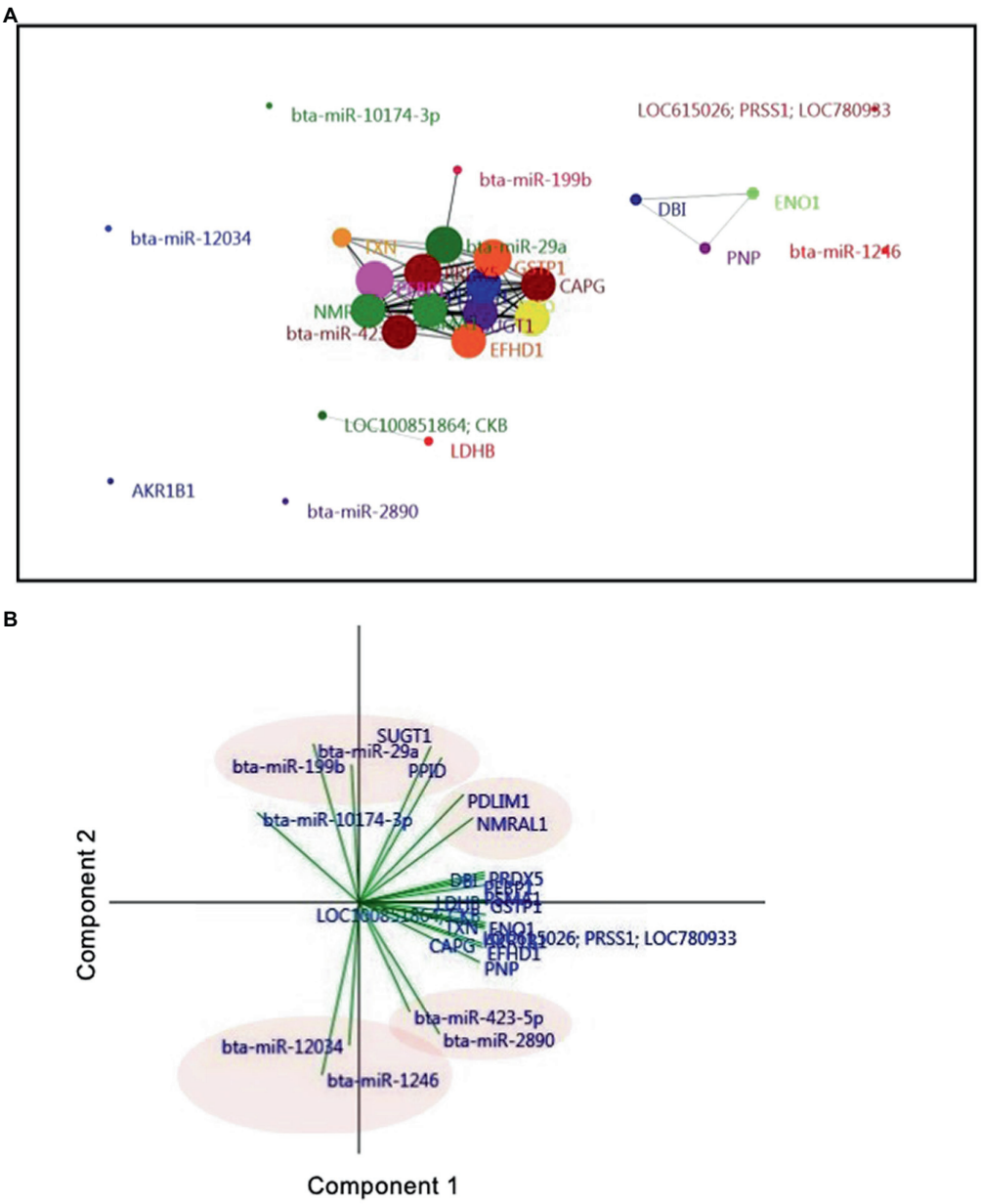


**FIGURE 5 |** Interaction of exosomal miRNAs and proteins in D7 UF-containing embryos. **(A)** Correlation-based network analysis of exosomal miRNAs and proteins in UF. Eighteen unique proteins and nine differentially expressed miRNAs in UF of cows were mapped to a network analysis. Network analysis and visualization were carried out using the PAST and Fruchterman-Reingold algorithm as a force-directed layout algorithm. The Pearson correlation thresholds of 99% were chosen to assess the relationships between edges and nodes. Nodes are proteins and miRNAs. Edges are the interactions of all variables. The size of the nodes and edges refers to the coefficient of clustering and the coefficient of correlation, respectively. Small nodes and thin edges refer to small values. **(B)** Biplot generated from the PCA of factors, including 18 unique proteins and nine differentially expressed miRNAs found in the UF of cows. Vectors with close angles (<45°) indicate a strong correlation, perpendicular vectors indicate no correlation, and vectors in opposite directions (approaching 180°) indicate a negative correlation.

and "neutrophil activation involved in immune response." Neutrophils are the first line of the defense mechanism in the innate immune system (Tapper, 1996); thus, they might immediately respond to the presence of embryo(s) in the uterus and play key roles in modulation of the local immune cascade for acceptance or rejection of the embryo. Neutrophils could also transfer the local IFNT-signal from the uterus to the peripheral blood neutrophils. This hypothesis was supported by our previous study, demonstrating that the peripheral blood neutrophils could respond to the IFNT-signal much earlier (D5) than peripheral blood mononuclear cells (D8; Shirasuna et al., 2012). Additionally, we reported that superovulated D7 embryos generated an anti-inflammatory immune response in peripheral blood neutrophils through the upregulation of transcripts for anti-inflammatory cytokines (TGFB1 and IL10; Talukder et al., 2019). Moreover, it was found that insemination induces rapid and transient infiltrations of neutrophils into the uterine lumen for removal of bacteria, excess/dead sperm, and tissue debris, which enhances uterine clearance and subsequent embryo receptivity (Katila, 2012). The data from the current study suggest that the pre-hatching D7 bovine embryos begin to communicate with the uterus locally through the innate immune system, which might pave the way for a tolerance of the semi-allogenic embryo toward successful pregnancy in the cow. However, the definite role of neutrophils, as a component of innate immune functions of the uterus, in pregnancy establishment in cows warrants further investigations, particularly during the pre-hatching period.

This study provides the first evidence for the existence of exosomes in the lumen of the bovine uterus on D7 of pregnancy. The intrauterine exosomes are secreted from both embryo and endometrium (Burns et al., 2016), and the number of intrauterine exosomes was shown to be markedly higher in the presence of a D17 embryo in ruminants (Nakamura et al., 2016). Although the presence of intrauterine exosomes was validated in this study, the origin of exosomes (trophectoderm or endometrium) was not determined. We speculate that the higher concentration of exosomes in the UF of AI cows could be released from the endometrium in response to the zona-pellucida encapsulated embryos.

Nucleic acids, especially miRNAs, are enriched in exosomes, suggesting that exosomes can serve as means of genetic information transfer from one cell to another. It has been reported that several miRNAs in the embryo- and endometrium-derived exosomes affect the expression of adhesion- and migration-related genes in the endometrium (Kurian and Modi, 2019). In addition, enrichment analysis of the exosome-derived miRNAs indicated their involvement in several pathways necessary for embryo-endometrial crosstalk at implantation, inflammation, cell remodeling, proliferation, and angiogenesis (Ng et al., 2013; Bidarimath et al., 2017). In this study, miRNA-seq of exosomal RNA from UF revealed 37 miRNAs, of which nine miRNAs were differently expressed between AI and Ctrl cows. Of these, miR-1246, upregulated in exosomes from AI cows, targets GSK3ß and AXIN2, and inhibitors of WNT/ß-catenin signaling, which is crucial for embryo development and placentation (Chai et al., 2016; Muralimanoharan et al., 2018; Yang et al., 2019). Indeed, the enrichment analysis identified several WNT signaling-related terms in this study. It has been reported that several WNTs and WNT-related molecules are produced and secreted in the oviduct and endometrium, which regulate embryo development during early pregnancy, such as D7 (Tríbulo et al., 2018). These findings suggest that miR-1246-containing exosomes could accelerate WNT-induced early embryonic development through inhibition of GSK3ß and AXIN2. The functional impact of other deferentially expressed miRNAs on embryo-maternal interaction during the pre-hatching period in cows is unknown and requires further investigations.

The PCA analysis identified SUGT1 as a unique protein associated with the presence of embryos in the UF. Moreover, the AUC showed that SUGT1 is an indicator of embryo presence in the UF. It has been reported that SUGT1 may be involved in innate immunity through the control of inflammasome and NF-kappaB (NF-kB) activities in mammals (Mayor et al., 2007). It has also been shown that the presence of D8 bovine embryos decreased NF-kB contents in UF, adding to the thoughts on immune privilege during early embryonic development (Muñoz et al., 2012). Of note, the network analysis and PCA revealed a strong association between bta-miR-29a, bta-miR-199b, SUGT1, and PPID. It has been reported that miR-29 decreases the chemotaxis of human neutrophils through CDK2 suppression (Hsu et al., 2019), and miR-199b negatively regulates innate and adaptive immunity by suppression of IFN gamma (Ma et al., 2011). These findings suggest that D7 embryos can regulate local innate immunity in the uterine microenvironment *via* regulation of miR-29, miR-199b, and SUGT1. However, further studies are required to confirm this hypothesis.

One could argue that the superovulation model for multiple embryo production used in this study may not represent the physiological condition normally seen with a single embryo. It was reported that the maintenance of higher circulating P4 concentration by P4 supplementation during early pregnancy leads to subtle changes in a large number of genes in conceptus and endometrium, resulting in enhancing histotroph composition and contributing to advanced conceptus elongation (Forde et al., 2009; Carter et al., 2010; Forde et al., 2010). Thus, the present model used may advance the endometrial response to the embryos, i.e., the observed changes on D7 of the present study could represent the physiological changes induced by a single embryo normally seen several days later. We previously reported that the presence of multiple D7 embryos amplified the embryo-derived signals in the uterus while a single embryo did not (Talukder et al., 2017; Passaro et al., 2018). As expected, in the present study, Japanese Black cows' responsiveness to hormone treatment was highly homogeneous, of which the number of recovered embryos with good quality per AI cows was 12 ± 0.95. Therefore, it is likely that using the model with multiple embryos in our study may have amplified the response in gene expression, rather than causing unphysiological condition or response, in the intrauterine environment.

## CONCLUSION

In conclusion, this study provides evidence that the presence of D7 multiple pre-hatching blastocysts induces significant changes in the relative and absolute protein and miRNA contents

of UF. The comparison of global protein analysis of UF between AI and Ctrl cows revealed the involvement of innate immunity and the increase of exosomes, including several miRNAs, in the uterine lumen on D7 in the presence of multiple embryos. These findings suggest that a network of proteins and miRNA components in exosomes could contribute to the uterine microenvironment for embryo development and successful pregnancy. The impact of these proteins and miRNAs on the immunological regulation of embryo-maternal communication and the establishment of pregnancy in cattle warrants further investigations.

## DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found at https://www.ddbj.nig.ac.jp/, DRA010067.

## ETHICS STATEMENT

The animal study was reviewed and approved by the Animal Experiments Ethics Committee, Obihiro University of Agriculture and Veterinary Medicine, Japan.

## AUTHOR CONTRIBUTIONS

Conceptualization: MS, KI, and AM. Methodology: KK, MR, RK, AT, and KN. Software, data curation, and investigation: KK. Validation: MM and AM. Formal analysis: KK and KN. Resources, supervision, project administration, and funding acquisition: AM. Writing: KK, MM, HK, and AM. Visualization: KI. All authors contributed to the article and approved the submitted version.

## FUNDING

This work was supported by the Grant-in-Aid for Scientific Research (16H05013, 17F17407, and 20H03122) from the Japan Society for the Promotion of Science (JSPS), Japan Association for Livestock New Technology, and Livestock Promotional Funds of Japan Racing Association (JRA).

## ACKNOWLEDGMENTS

Computations were partially performed on the NIG supercomputer at the ROIS National Institute of Genetics. The authors would like to thank Dr. Toshiro Takedomi (Taketomi Reproduction Clinic, Obihiro, Japan) for his kind help in collecting the uterine flushing.

## REFERENCES

Bi, Y., Liu, G., and Yang, R. (2009). MicroRNAs: novel regulators during the immune response. *J. Cell. Physiol.* 218, 467–472. doi: 10.1002/jcp.21639

Bidarimath, M., Khalaj, K., Kridli, R. T., Kan, F. W., Koti, M., and Tayade, C. (2017). Extracellular vesicle mediated intercellular communication at the porcine maternal-fetal interface: a new paradigm for conceptus-endometrial cross-talk. *Sci. Rep.* 7:40476. doi: 10.1038/srep40476

Bidarimath, M., Khalaj, K., Wessels, J. M., and Tayade, C. (2014). MicroRNAs, immune cells and pregnancy. *Cell. Mol. Immunol.* 11, 538–547. doi: 10.1038/cmi.2014.45

Bó, G. A., and Mapletoft, R. J. (2014). Historical perspectives and recent research on superovulation in cattle. *Theriogenology* 81, 38–48. doi: 10.1016/j.theriogenology.2013.09.020

Burns, G., Brooks, K., Wildung, M., Navakanitworakul, R., Christenson, L. K., and Spencer, T. E. (2014). Extracellular vesicles in luminal fluid of the ovine uterus. *PLoS One* 9:e90913. doi: 10.1371/journal.pone.0090913

Burns, G. W., Brooks, K. E., and Spencer, T. E. (2016). Extracellular vesicles originate from the conceptus and uterus during early pregnancy in sheep. *Biol. Reprod.* 94:56. doi: 10.1095/biolreprod.115.134973

Carter, F., Rings, F., Mamo, S., Holker, M., Kuzmany, A., Besenfelder, U., et al. (2010). Effect of elevated circulating progesterone concentration on bovine blastocyst development and global transcriptome following endoscopic transfer of in vitro produced embryos to the bovine oviduct. *Biol. Reprod.* 83, 707–719. doi: 10.1095/biolreprod.109.082354

Chai, S., Ng, K. Y., Tong, M., Lau, E. Y., Lee, T. K., Chan, K. W., et al. (2016). Octamer 4/microRNA-1246 signaling axis drives Wnt/β-catenin activation in liver cancer stem cells. *Hepatology* 64, 2062–2076. doi: 10.1002/hep.28821

Diskin, M. G., and Morris, D. G. (2008). Embryonic and early foetal losses in cattle and other ruminants. *Reprod. Domest. Anim.* 43(Suppl. 2), 260–267. doi: 10.1111/j.1439-0531.2008.01171.x

Forde, N., Carter, F., Fair, T., Crowe, M. A., Evans, A. C., Spencer, T. E., et al. (2009). Progesterone-regulated changes in endometrial gene expression contribute to advanced conceptus development in cattle. *Biol. Reprod.* 81, 784–794. doi: 10.1095/biolreprod.108.074336

Forde, N., McGettigan, P. A., Mehta, J. P., O'Hara, L., Mamo, S., Bazer, F. W., et al. (2014a). Proteomic analysis of uterine fluid during the pre-implantation period of pregnancy in cattle. *Reproduction* 147, 575–587. doi: 10.1530/rep-13-0010

Forde, N., Mehta, J. P., McGettigan, P. A., Mamo, S., Bazer, F. W., Spencer, T. E., et al. (2013). Alterations in expression of endometrial genes coding for proteins secreted into the uterine lumen during conceptus elongation in cattle. *BMC Genomics* 14:321. doi: 10.1186/1471-2164-14-321

Forde, N., Simintiras, C. A., Sturmey, R., Mamo, S., Kelly, A. K., Spencer, T. E., et al. (2014b). Amino acids in the uterine luminal fluid reflects the temporal changes in transporter expression in the endometrium and conceptus during early pregnancy in cattle. *PLoS One* 9:e100010. doi: 10.1371/journal.pone.0100010

Forde, N., Spencer, T. E., Bazer, F. W., Song, G., Roche, J. F., and Lonergan, P. (2010). Effect of pregnancy and progesterone concentration on expression of genes encoding for transporters or secreted proteins in the bovine endometrium. *Physiol. Genomics* 41, 53–62. doi: 10.1152/physiolgenomics.00162.2009

Gross, N., Kropp, J., and Khatib, H. (2017a). MicroRNA signaling in embryo development. *Biology* 6:34. doi: 10.3390/biology6030034

Gross, N., Kropp, J., and Khatib, H. (2017b). Sexual dimorphism of miRNAs secreted by bovine in vitro-produced embryos. *Front. Genet.* 8:39. doi: 10.3389/fgene.2017.00039

Hsu, A. Y., Wang, D., Liu, S., Lu, J., Syahirah, R., Bennin, D. A., et al. (2019). Phenotypical microRNA screen reveals a noncanonical role of CDK2 in regulating neutrophil migration. *Proc. Natl. Acad. Sci. U. S. A.* 116, 18561–18570. doi: 10.1073/pnas.1905221116

Ioannidis, J., and Donadeu, F. X. (2016). Circulating miRNA signatures of early pregnancy in cattle. *BMC Genomics* 17:184. doi: 10.1186/s12864-016-2529-1

Katila, T. (2012). Post-mating inflammatory responses of the uterus. *Reprod. Domest. Anim.* 47(Suppl. 5), 31–41. doi: 10.1111/j.1439-0531.2012.02120.x

Kurian, N. K., and Modi, D. (2019). Extracellular vesicle mediated embryo-endometrial cross talk during implantation and in pregnancy. *J. Assist. Reprod. Genet.* 36, 189–198. doi: 10.1007/s10815-018-1343-x

Kusama, K., Bai, R., Ideta, A., Aoyagi, Y., Okuda, K., and Imakawa, K. (2016). Regulation of epithelial to mesenchymal transition in bovine conceptuses through the interaction between follistatin and activin A. *Mol. Cell. Endocrinol.* 434, 81–92. doi: 10.1016/j.mce.2016.06.017

Kusama, K., Bai, R., and Imakawa, K. (2018a). Regulation of human trophoblast cell syncytialization by transcription factors STAT5B and NR4A3. *J. Cell. Biochem.* 119, 4918–4927. doi: 10.1002/jcb.26721

Kusama, K., Bai, R., Nakamura, K., Okada, S., Yasuda, J., and Imakawa, K. (2017). Endometrial factors similarly induced by IFNT2 and IFNTc1 through transcription factor FOXS1. *PLoS One* 12:e0171858. doi: 10.1371/journal.pone.0171858

Kusama, K., Nakamura, K., Bai, R., Nagaoka, K., Sakurai, T., and Imakawa, K. (2018b). Intrauterine exosomes are required for bovine conceptus implantation. *Biochem. Biophys. Res. Commun.* 495, 1370–1375. doi: 10.1016/j.bbrc.2017.11.176

Liang, J., Wang, S., and Wang, Z. (2017). Role of microRNAs in embryo implantation. *Reprod. Biol. Endocrinol.* 15:90. doi: 10.1186/s12958-017-0309-7

Low, B. G., Hansen, P. J., Drost, M., and Gogolin-Ewens, K. J. (1990). Expression of major histocompatibility complex antigens on the bovine placenta. *J. Reprod. Fertil.* 90, 235–243. doi: 10.1530/jrf.0.0900235

Ma, F., Xu, S., Liu, X., Zhang, Q., Xu, X., Liu, M., et al. (2011). The microRNA miR-29 controls innate and adaptive immune responses to intracellular bacterial infection by targeting interferon-γ. *Nat. Immunol.* 12, 861–869. doi: 10.1038/ni.2073

Mayor, A., Martinon, F., De Smedt, T., Pétrilli, V., and Tschopp, J. (2007). A crucial function of SGT1 and HSP90 in inflammasome activity links mammalian and plant innate immune responses. *Nat. Immunol.* 8, 497–503. doi: 10.1038/ni1459

Muñoz, M., Corrales, F. J., Caamaño, J. N., Díez, C., Trigal, B., Mora, M. I., et al. (2012). Proteome of the early embryo-maternal dialogue in the cattle uterus. *J. Proteome Res.* 11, 751–766. doi: 10.1021/pr200969a

Muralimanoharan, S., Kwak, Y. T., and Mendelson, C. R. (2018). Redox-sensitive transcription factor NRF2 enhances trophoblast differentiation via induction of miR-1246 and aromatase. *Endocrinology* 159, 2022–2033. doi: 10.1210/en.2017-03024

Nakamura, K., Kusama, K., Bai, R., Sakurai, T., Isuzugawa, K., Godkin, J. D., et al. (2016). Induction of IFNT-stimulated genes by conceptus-derived exosomes during the attachment period. *PLoS One* 11:e0158278. doi: 10.1371/journal.pone.0158278

Nakamura, K., Kusama, K., Ideta, A., Imakawa, K., and Hori, M. (2020). IFNT-independent effects of intrauterine extracellular vesicles (EVs) in cattle. *Reproduction* 159, 503–511. doi: 10.1530/REP-19-0314

Ng, Y. H., Rome, S., Jalabert, A., Forterre, A., Singh, H., Hincks, C. L., et al. (2013). Endometrial exosomes/microvesicles in the uterine microenvironment: a new paradigm for embryo-endometrial cross talk at implantation. *PLoS One* 8:e58502. doi: 10.1371/journal.pone.0058502

Passaro, C., Tutt, D., Mathew, D. J., Sanchez, J. M., Browne, J. A., Boe-Hansen, G. B., et al. (2018). Blastocyst-induced changes in the bovine endometrial transcriptome. *Reproduction* 156, 219–229. doi: 10.1530/REP-18-0188

Rashid, M. B., Talukder, A. K., Kusama, K., Haneda, S., Takedomi, T., Yoshino, H., et al. (2018). Evidence that interferon-tau secreted from Day-7 embryo in vivo generates anti-inflammatory immune response in the bovine uterus. *Biochem. Biophys. Res. Commun.* 500, 879–884. doi: 10.1016/j.bbrc.2018.04.178

Ruiz-González, I., Xu, J., Wang, X., Burghardt, R. C., Dunlap, K. A., and Bazer, F. W. (2015). Exosomes, endogenous retroviruses and toll-like receptors: pregnancy recognition in ewes. *Reproduction* 149, 281–291. doi: 10.1530/REP-14-0538

Schjenken, J. E., Zhang, B., Chan, H. Y., Sharkey, D. J., Fullston, T., and Robertson, S. A. (2016). miRNA regulation of immune tolerance in early pregnancy. *Am. J. Reprod. Immunol.* 75, 272–280. doi: 10.1111/aji.12490

Shirasuna, K., Matsumoto, H., Kobayashi, E., Nitta, A., Haneda, S., Matsui, M., et al. (2012). Upregulation of interferon-stimulated genes and interleukin-10 in peripheral blood immune cells during early pregnancy in dairy cows. *J. Reprod. Dev.* 58, 84–90. doi: 10.1262/jrd.11-094K

Sponchiado, M., Gomes, N. S., Fontes, P. K., Martins, T., Del Collado, M., Pastore, A. A., et al. (2017). Pre-hatching embryo-dependent and -independent programming of endometrial function in cattle. *PLoS One* 12:e0175954. doi: 10.1371/journal.pone.0175954

Sponchiado, M., Gonella-Diaza, A. M., Rocha, C. C., Turco, E. G. L., Pugliesi, G., Leroy, J., et al. (2019). The pre-hatching bovine embryo transforms the uterine luminal metabolite composition in vivo. *Sci. Rep.* 9:8354. doi: 10.1038/s41598-019-44590-9

Talukder, A. K., Marey, M. A., Shirasuna, K., Kusama, K., Shimada, M., Imakawa, K., et al. (2020). Roadmap to pregnancy in the first 7 days post-insemination in the cow: immune crosstalk in the corpus luteum, oviduct, and uterus. *Theriogenology* 150, 313–320. doi: 10.1016/j.theriogenology.2020.01.071

Talukder, A. K., Rashid, M. B., Takedomi, T., Moriyasu, S., Imakawa, K., and Miyamoto, A. (2019). Day-7 embryos generate an anti-inflammatory immune response in peripheral blood immune cells in superovulated cows. *Am. J. Reprod. Immunol.* 81:e13069. doi: 10.1111/aji.13069

Talukder, A. K., Yousef, M. S., Rashid, M. B., Awai, K., Acosta, T. J., Shimizu, T., et al. (2017). Bovine embryo induces an anti-inflammatory response in uterine epithelial cells and immune cells in vitro: possible involvement of interferon tau as an intermediator. *J. Reprod. Dev.* 63, 425–434. doi: 10.1262/jrd.2017-056

Tapper, H. (1996). The secretion of preformed granules by macrophages and neutrophils. *J. Leukoc. Biol.* 59, 613–622. doi: 10.1002/jlb.59.5.613

Templeton, J. W., Tipton, R. C., Garber, T., Bondioli, K., and Kraemer, D. C. (1987). Expression and genetic segregation of parental BoLA serotypes in bovine embryos. *Anim. Genet.* 18, 317–322. doi: 10.1111/j.1365-2052.1987.tb00775.x

Tríbulo, P., Siqueira, L. G. B., Oliveira, L. J., Scheffler, T., and Hansen, P. J. (2018). Identification of potential embryokines in the bovine reproductive tract. *J. Dairy Sci.* 101, 690–704. doi: 10.3168/jds.2017-13221

Wiltbank, M. C., Baez, G. M., Garcia-Guerra, A., Toledo, M. Z., Monteiro, P. L., Melo, L. F., et al. (2016). Pivotal periods for pregnancy loss during the first trimester of gestation in lactating dairy cows. *Theriogenology* 86, 239–253. doi: 10.1016/j.theriogenology.2016.04.037

Yang, F., Xiong, H., Duan, L., Li, Q., Li, X., and Zhou, Y. (2019). MiR-1246 promotes metastasis and invasion of A549 cells by targeting GSK-3β–mediated Wnt/β-catenin pathway. *Cancer Res. Treat.* 51, 1420–1429. doi: 10.4143/crt.2018.638

**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.

frontiers in Genetics

ORIGINAL RESEARCH
published: 21 July 2021
doi: 10.3389/fgene.2021.697883

# *SLC4A11* and *MFSD3* Gene Expression Changes in Deoxynivalenol Treated IPEC-J2 Cells

*Yafei Xu[1], Xiaolei Chen[1], Luchen Yu[1], Yi Wang[1], Haifei Wang[1], Zhengchang Wu[1], Shenglong Wu[1,2] and Wenbin Bao[1,2]**

[1] Key Laboratory for Animal Genetic, Breeding, Reproduction and Molecular Design of Jiangsu Province, College of Animal Science and Technology, Yangzhou University, Yangzhou, China, [2] Joint International Research Laboratory of Agriculture and Agri-Product Safety, Yangzhou University, Yangzhou, China

Deoxynivalenol (DON) caused serious cytotoxicity for animal cells. However, genes involved in regulating DON toxicity and the underlying molecular mechanisms remain largely unknown. This study explored the role of *SLC4A11* and *MFSD3* in alleviating DON toxicity and analyzed the DNA methylation changes of these two genes. Viability and cell cycle analysis showed that DON exposure decreased the IPEC-J2 viability ($P < 0.01$), blocked the cell cycle in the G2/M phase ($P < 0.01$), and increased the rate of apoptosis ($P < 0.05$). Expression of the *SLC4A11* and *MFSD3* genes was significantly downregulated upon DON exposure ($P < 0.01$). Overexpression of *SLC4A11* and *MFSD3* can enhance the cell viability ($P < 0.01$). DNA methylation assays indicated that promoter methylation of *SLC4A11* (mC-1 and mC-23) and *MFSD3* (mC-1 and mC-12) were significantly higher compared with those in the controls and correlated negatively with mRNA expression ($P < 0.05$). Further analysis showed that mC-1 of *SLC4A11* and *MFSD3* was located in transcription factor binding sites for NF-1 and Sp1. Our findings revealed the novel biological functions of porcine *SLC4A11* and *MFSD3* genes in regulating the cytotoxic effects induced by DON, and may contribute to the detection of biomarkers and drug targets for predicting and eliminating the potential toxicity of DON.

**Keywords: pig, DON, *SLC4A11*, *MFSD3*, methylation regulation**

OPEN ACCESS

***Edited by:***
*Aline Silva Mello Cesar, University of São Paulo, Brazil*

***Reviewed by:***
*Doreen Becker, Leibniz Institute for Farm Animal Biology (FBN), Germany*
*Carlotta Giromini, University of Milan, Italy*

***Correspondence:**
*Wenbin Bao*
*wbbao@yzu.edu.cn*

***Specialty section:***
*This article was submitted to Livestock Genomics, a section of the journal Frontiers in Genetics*

***Received:*** *20 April 2021*
***Accepted:*** *21 June 2021*
***Published:*** *21 July 2021*

***Citation:***
*Xu Y, Chen X, Yu L, Wang Y, Wang H, Wu Z, Wu S and Bao W (2021) SLC4A11 and MFSD3 Gene Expression Changes in Deoxynivalenol Treated IPEC-J2 Cells. Front. Genet. 12:697883. doi: 10.3389/fgene.2021.697883*

## INTRODUCTION

Deoxynivalenol (DON), belonging to the family of trichothecenes, is one of the metabolites of Fusarium, induces emesis in swine, leading to it being known as "vomitoxin" (Pestka and Smolinski, 2005; Pestka, 2010; Borutova et al., 2012). DON is extremely harmful to pigs, and its main toxic effects are antifeeding, vomiting, decreased immunity, necrosis of the digestive tract, decreased reproductive performance, and finally, death (Eriksen and Pettersson, 2004). Early research suggested that DON induced the apoptosis of piglet mesenteric lymph node cells, stimulated immune cells to secrete cytokines, and resulted in intestinal inflammation by destroying the intestinal barrier and inducing inflammation (Pestka, 2010; Gerez et al., 2015). Once the

intestinal barrier is damaged, in addition to affecting the digestion and absorption of pigs, it is likely to cause other intestinal diseases, which will seriously affect the growth and development of pigs, resulting in huge economic losses to the pig industry. Therefore, it is necessary to determine the toxic effects and mechanism of DON, which will contribute to discovering new methods of prevention and treatment of diseases induced by DON.

Epigenetic modification is a type of gene regulation that is ubiquitous in all organisms, and is essential to maintain the normal life activities of mammals (Shi and Wu, 2009). Epigenetic modification includes DNA methylation, histone modification, and non-coding RNA regulation, among which DNA methylation has attracted much attention from epigenetic modification researchers (Dahl and Guldberg, 2003). DNA methylation regulates gene expression by affecting chromatin structure, DNA conformation, chromosome stability, and DNA-protein interaction. In eukaryotes, it plays an important role in cell differentiation, embryonic development, environmental adaptation, and the development of diseases (Deng, 2014; Dai et al., 2019). After DON exposure, the levels of methyltransferase H3K4me2 increased significantly on *EZH2*, encoding enhancer of Zeste 2 polycomb repressive complex 2 Subunit, a key regulator in the early development of oocytes. Thus, DON can cause the interruption of oocyte maturation by changing epigenetic modifications (Han et al., 2016). Studies have also reported increased DNA methylation levels and the expression of H3K9me3 and H4K20me3 in the oocytes of mice fed a diet containing DON, and that the expression levels of H3K27me3 and H4K20me2 were reduced, revealing that epigenetic modifications might be one of the reasons for the decrease of the level of oocyte development (Zhu et al., 2014). These studies showed that epigenetic modification plays an important regulatory role in the process of DON induction.

Our previous transcriptome analysis in IPEC-J2 cells found significant decreases in the expression levels of the *SLC4A11* (encoding solute carrier family 4 member 11) and *MFSD3* (encoding major facilitator superfamily domain containing 3) genes in response to DON (Wang et al., 2019). *SLC4A11* and *MFSD3* are both membrane-bound solute carriers (SLCs), which maintain nutrient uptake, ion transport, and waste removal associated with physiological functions (Perland et al., 2017). *SLC4A11* is an electrogenic Na/borate cotransporter that stimulates cell growth and proliferation by increasing intracellular borate levels and activating the mitogen activated protein kinase (MAPK) pathway (Jiao et al., 2007; Lopez et al., 2009). *MFSD3* is a kind of membrane-bound solute carrier that belongs to the major facilitator superfamily (MFS), which is the largest phylogenetic group of SLCs in humans (Nicoletti et al., 2019). Studies reported that the expression of *MFSD3* was associated with nutrient intake and adipose tissue homeostasis (Hoglund et al., 2011). Therefore, *SLC4A11* and *MFSD3* may play an important role in DON-induced cell damage, we further explored the expression regulation mechanism of *SLC4A11* and *MFSD3* genes associated with the activity of IPEC-J2 cells induced by DON. We examined the effects of DON on the viability, cell cycle, and apoptosis of IPEC-J2 cells, as well as the regulation of *SLC4A11* and *MFSD3* expression levels in IPEC-J2 cells induced by DON, including a comprehensive analysis of the degree of methylation and expression changes of these two genes. The present study explored the regulatory role of *SLC4A11* and *MFSD3* in resisting DON-induced cytotoxicity. Better understanding of DON pathogenesis and identification of the responsive genes provided the theoretical basis for further study of the molecular regulation mechanism of DNA methylation modification in DON-induced cytotoxicity, and may contribute to the identification of biomarkers and drug targets for DON contamination.

# MATERIALS AND METHODS

## Ethics Statement

The animal study proposal was approved by the Institutional Animal Care and Use Committee (IACUC) of the Yangzhou University Animal Experiments Ethics Committee [permit number: SYXK (Su) IACUC 2012-0029]. All experimental methods were conducted in accordance with the related guidelines and regulations.

## Cell Culture

The IPEC-J2 cells were preserved in our laboratory, and cultured in Dulbecco's modified Eagle's medium (DMEM) containing 10% fetal bovine serum (FBS) and 1% penicillin streptomycin (1 mg/mL) at 37°C with 5% CO2.

## Cell Viability

IPEC-J2 cells were cultivated in 96-well plates at a density of $2 \times 10^3$ cells/well and cultured for 24 h. Based on a previous study, (Wang et al., 2019) we could see that treatment with a DON (Sigma, Germany) concentration of 1 μg/ml for 48 h induces cytotoxicity in IPEC-J2 cells. When the cells reached 70–80% confluence, they were incubated with DON (1 μg/mL) for 24, 48, and 72 h. Cell viability was assessed using a Cell Counting Kit-8 (MedChemExpress, Monmouth Junction, NJ, United States) according to the manufacturer's protocol. The absorbance was measured on a Tecan Infinite Pro (Sunrise, Tecan, Switzerland) at 450 nm.

## Cell Apoptosis Assay

IPEC-J2 cells were seeded into six-well plates at a density of $2 \times 10^5$ cells/well and randomly assigned into a control group and a DON treated group. When the cells reached 70–80% confluence, they were incubated with DON (1 μg/mL) for 48 h in the DON treated group. Subsequently, cells were collected, and stained with Annexin V-FITC according to the instructions of the Apoptosis Detection kit (Solarbio, Beijing, China). Finally, apoptosis was analyzed using a Flow Cytometer (FAC Scan, Becton Dickinson, Franklin Lakes, NJ, United States) within 1 h.

## Cell Cycle Analysis

First, IPEC-J2 cells were cultured in a six-well plates and incubated at 37°C with 5% $CO_2$ overnight and divided into a

control group and a DON treated group. Then, digestion was performed with trypsin without EDTA followed by washing twice in pre-cooled phosphate-buffered saline. Cell cycle analysis was performed according to a Cell Cycle and Apoptosis Analysis Kit (Beyotime Institute of Biotechnology, Jiangsu, China) and the percentage of the cell population at a particular phase was estimated using ModFit LT for Windows V3.1 (Verity Software House, Topsham, ME, United States).

## RNA Extraction and Quantitative Real Time Reverse Transcription PCR

Total RNA was extracted from IPEC-J2 cells using Trizol. cDNA was reverse transcribed from the mRNA was reverse transcribed using a HiScript II Q RT SuperMix kit (Vazyme Biotech Co., Ltd., Nanjing, China). The cDNA was used as the template for the quantitative real-time PCR (qPCR) step using a fast real-time PCR system (ABI Step One Plus; Applied Biosystems, Foster City, CA, United States). *GAPDH* (encoding glyceraldehyde-3-phosphate dehydrogenase) was used as the control gene. The relative expression levels were determined using the $2^{-\Delta\Delta CT}$ method (Livak and Schmittgen, 2001) and overexpression efficiency was also calculated using the $2^{-\Delta\Delta CT}$ method. The primers used are shown in **Table 1**. All primers were synthesized by Sangon (Shanghai, China).

## Construction of Overexpression Cell Lines

Primers with *Sac*I and *Pme*I restriction enzyme sites were designed according to the coding sequences (CDS) of porcine *SLC4A11* (XM_021077564.1) and *MFSD3* (XM_021090521.1) gene published in the GenBank database[1] (**Table 1**). The primers were used to amplify the *SLC4A11* and *MFSD3* genes, which were ligated into the florescence expression vector, pEGFP–C1, after double digestion. The sequences of the *SLC4A11* and *MFSD3* overexpression vectors were confirmed by Sangon sequencing. Cells were incubated with 20 nM *SLC4A11* and *MFSD3* overexpression vector or negative control using Jet PRIME (Polyplus, Illkirch, France). After transfection, stable cells lines were established by incubation with geneticin (800 μg/mL) (Invivogen, Toulouse, France) for 15 days.

**TABLE 1 |** The primer sequence of genes for quantitative real-time PCR (qPCR) and their sequences.

| Gene | Primer sequence | Product length (bp) |
|---|---|---|
| *SLC4A11-q1* | F: AGTAGTAGGGAGCAGGGTGG | 234 |
| | R: AAGCAAGCAGAGAGTGAGCC | |
| *MFSD3-q2* | F: CGCCTCAGCCATCAGAACCCCGCCG | 224 |
| | R: TGAGCCCACAATGGAACAGA | |
| *GAPDH-q3* | F: GGTCGGAGTGAACGGATTT | 245 |
| | R: ATTTGATGTTGGCGGGAT | |
| *SLC4A11-c1* | F: CGAGCTCATGTCACAGAGTGGATACCC | *2631* |
| | R: AGCTTTGTTTAAACTCAATGATTCGGGGCAGCAG | |
| *MFSD3-c2* | F: CGAGCTCATGCACGGGAAGCTGCTGGT | *1635* |
| | R: AGCTTTGTTTAAACTCAGCCAAGGCTCCGGCCAG | |
| *SLC4A11-p1* | F: GTTATTTTGGGAAGGAAAAAGTT | *345* |
| | R: CCCTAATATCAACTTTAAATATATAACC | |
| *MFSD3-p2* | F: ATTCTGCCCCCCA | *260* |
| | R: GCCTCCAAGCAGCCAA | |

*Note q1–q3 stands for qPCR primers. c1–c2 stands for overexpression primers. The underlined sequences represent restriction sites. p1–p2 stands for BSP primers, respectively.*

## Bioinformatic Analysis and Primer Design

Analysis and identification of the promoter region, CpG islands and transcription factor binding sites in the 5′ flanking regions of *SLC4A11* and *MFSD3* were performed using the online tool MethPrimer and Alibaba (Li and Dahiya, 2002). Based on this analysis and the location of the predicted CpG islands, bisulfite-sequencing PCR (BSP) primers were designed using MethPrimer (**Table 1**).

## Genomic DNA Preparation and Promoter Methylation Analysis

Genomic DNA was extracted from IPEC-J2 cells using a Universal DNA purification kit (Tiangen, Beijing, China), following the manufacturer's instructions. The genomic DNA was treated with bisulfite and used as the template for PCR (ZymoTaq PreMix; Zymo Research, Irvine, CA, United States). Purified PCR products were ligated into vector PMD-19T overnight (Bao Biological Engineering Co., Dalian, China). The recombinant clones were transformed into competent *Escherichia coli* DH5α cells (Tiangen, Beijing, China). 200 μL of the bacteria solution were applied to Luria-Bertani (LB) agar plates containing 100 ng/mL ampicillin, and maintained at 37°C overnight. The positive recombinant clones were selected on LB-ampicillin and confirmed by sequencing (Sangon Shanghai, China). The raw *SLC4A11* and *MFSD3* mRNA sequences were aligned using QUMA online software to analyze the degree of methylation at each CpG site. The methylation rate for each sample was calculated as the number of CpG methylated loci/the number of CpG loci.

## Statistical Analysis

One-way analysis of variance and Student's *t*-test were used for comparisons between the control and DON treated groups. The association between an individual CG methylation and the expression levels of genes after DON-induced cytotoxicity in IPEC-J2 cells was analyzed using Pearson's correlation analysis. All statistical analyses were performed using the SPSS 23.0 (IBM Corp., Armonk, NY, United States) or GraphPad Prism version 8.0 for Windows (GraphPad Software Inc., San Diego, CA, United States). $P < 0.05$ was considered statistically significant and all experimental samples had three replicates.

[1] https://www.ncbi.nlm.nih.gov/gene/100157069

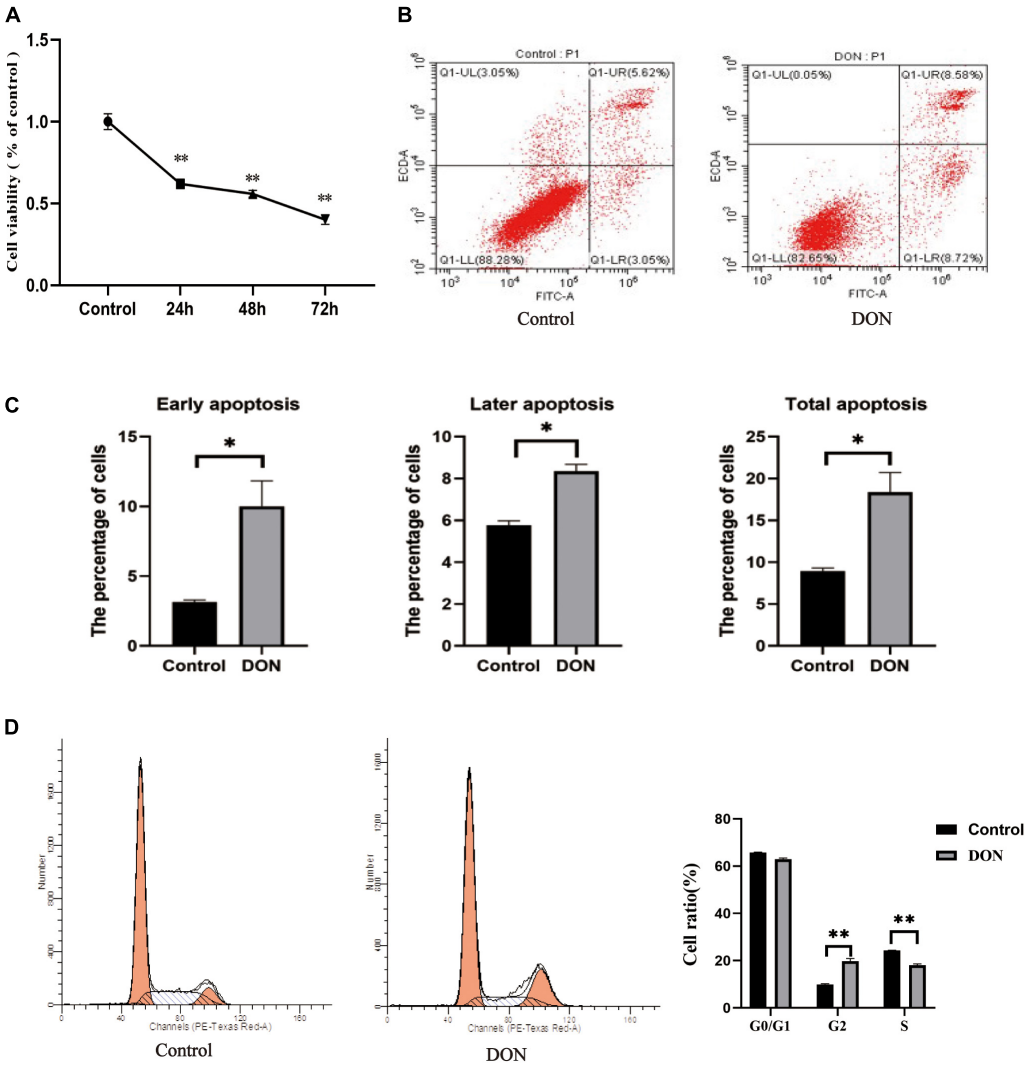


**FIGURE 1 |** The viability, cell cycle, and apoptosis of IPEC-J2 cells in the control group and the DON treated group. **(A)** The viability of IPEC-J2 cells in the control group and the DON treated group. **(B)** Apoptosis of IPEC-J2 cells in the control group and the Deoxynivalenol (DON) treatment group was evaluated by measurement of Annexin V by flow cytometry. Apoptotic cells were Annexin V-positive and PI-negative. Q1-UL stands for nuclear debris, Q1-LL stands for living cells, Q1-LR stands for early apoptotic cells, and Q1-UR stands for later apoptotic cells **(C)** Quantification of panel **(B)**. $^{*}P < 0.05$. **(D)** The cell cycle distribution in the control group and the DON treated group. $^{**}P < 0.01$.

# RESULTS

## DON-Induced Cytotoxicity in IPEC-J2 Cells

To investigate the effect of DON on the viability of IPEC-J2 cells, the cells were incubated with DON (1 μg/mL) for 24, 48, and 72 h. Cell viability was significantly lower compared with that of the control group at 24, 48, and 72 h ($P < 0.01$), indicating that DON induced increased cell death with time (**Figure 1A**). Flow Cytometry analysis also showed that DON increased the percentage of early, late, and total apoptotic cells significantly compared with that in the control group following 48 h of incubation (**Figures 1B,C**). In subsequent experiments, the cell samples were treated with DON for 48 h. Flow cytometry analysis of the cell cycle showed that after 48 h of DON treatment the cell cycle was arrested at the G2/M phase compared with the controls ($P < 0.01$), and the percentage of cells in the S phase decreased ($P < 0.01$) (**Figure 1D**).

## Effects of DON Induction on the Expression Levels of *SLC4A11* and *MFSD3* in IPEC-J2 Cells

The mRNA expression levels of *SLC4A11* and *MFSD3* were assessed using qRT-PCR in DON-treated IPEC-J2 cells. In the DON treated group, the mRNA expression levels of *SLC4A11* ($P < 0.01$) (**Figure 2A**) and *MFSD3* ($P < 0.01$) were significantly lower compared with those in the controls (**Figure 2B**).

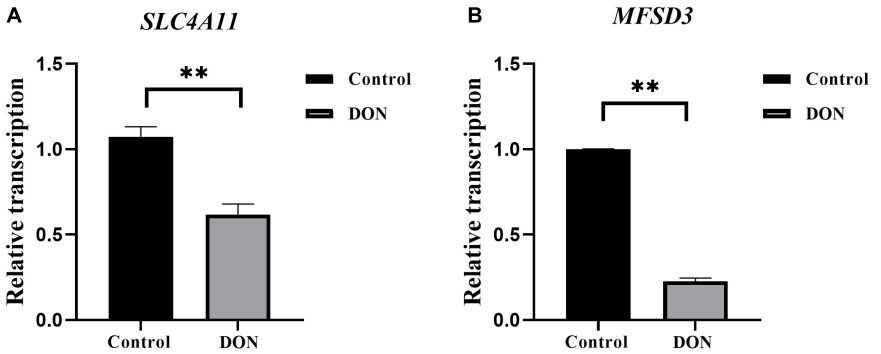


**FIGURE 2 |** *SLC4A11* and *MFSD3* mRNA expression levels in DON-induced IPEC-J2 cells. **(A)** *SLC4A11* mRNA expression levels induced by DON. $^{*}P < 0.05$. **(B)** *MFSD3* mRNA expression levels induced by DON. $^{**}P < 0.01$.

## Construction of Overexpression Cell Line *SLC4A11* and *MFSD3*

The recombinant plasmids pEGFP-C1-SLC4A11 and pEGFP-C1-MFSD3 were verified by DNA sequencing (**Supplementary Figure 1**). The gene sequences of the overexpression vectors were consistent with those of the pEGFP vector, and the *SLC4A11* and *MFSD3* CDS, which indicated that plasmids were successfully constructed. After 15 days of screening with geneticin, green fluorescent cells could still be observed under the fluorescence microscope, which indicated that the recombinant fusion proteins of pEGFP-C1-SLC4A11 and pEGFP-C1-MFSD3 were stably expressed in IPEC-J2 cells (**Supplementary Figure 2**). The relative mRNA expression levels of *SLC4A11* in the overexpression group represent a 585.09-fold increase compared with that in the control group ($P < 0.01$) (**Figure 3A**). The relative expression levels of the *MFSD3* mRNA in the overexpression group representing a 6.68-fold increase compared with that in the control group ($P < 0.01$) (**Figure 3B**).

## Effects of Overexpression of *SLC4A11* and *MFSD3* on the Viability of IPEC-J2 Cells Induced by DON

*SLC4A11* and *MFSD3* overexpressing cells were incubated with DON (1 μg/mL) for 24, 48, and 72 h to detect cell viability. The results showed that the viability of the *SLC4A11* overexpression group under DON treatment was significantly increased compared to the control group at 24, 48, and 72 h ($P < 0.01$) (**Figure 4A**). For the *MFSD3* overexpression group under DON treatment, the viability was significantly lower that of the control group at 24 and 48 h ($P < 0.05$). There was an increasing trend at 72 h, but the changes were not significant ($P > 0.05$) (**Figure 4B**).

## Prediction of *SLC4A11* and *MFSD3* CpG Islands and Methylation Analysis

The CpG islands in the porcine *SLC4A11* and *MFSD3* 5′ flanking region were predicted by MethPrimer, respectively. The analysis indicated that the *SLC4A11* 5′ flanking region contained two CpG islands, and primers were designed to amplify a 345 bp fragment (**Figure 5A**). The *MFSD3* 5′ flanking region also contained two CpG islands, and primers were designed to amplify a 260 bp fragment (**Figure 5B**). After determining the CpG island regions of the porcine *SLC4A11* and *MFSD3* genes, we analyzed the rest of the gene sequences and found that the *SLC4A11* gene contained 28 CpG sites (**Figure 6A**) and the *MFSD3* gene contained 19 CpG sites (**Figure 6D**). Methylation levels of both genes in the DON group (48 h of treatment) and the control group were very high. In addition, there was no significant difference in the overall methylation degree of *SLC4A11* and *MFSD3* between the DON treated group and the control group (**Figures 6B,E**). Interestingly, the methylation levels of the mC-1 and mC-23 sites of *SLC4A11* and the mC-1 and mC-12 sites of *MFSD3* were significantly higher compared with those in the controls ($P < 0.05$) (**Figures 6C,F**).

## Correlation Between the Methylation Level of *SLC4A11* and *MFSD3* Gene Amplified Fragment and mRNA Expression

Pearson analysis was performed to examine the correlation between the methylation degree of CpG islands in the promoter regions of *SLC4A11* and *MFSD3* genes and the mRNA expression level. The methylation level of 19 (out of 28) CpG sites in *SLC4A11* CpG islands correlated negatively with the mRNA expression level, in which the methylation levels of mC-1 and mC-23 correlated significantly and negatively with the mRNA expression level ($P < 0.05$) (**Figure 7A**). The methylation level of 11 (out of 19) CpG sites in *MFSD3* correlated negatively with the mRNA expression level (**Figure 7B**), in which the methylation level of mC-1 and mC-12 correlated significantly and negatively with the mRNA expression level ($P < 0.05$).

## Prediction of Transcription Factor Binding Sites

The potential transcription factor binding sites in the regions containing the CpG islands in the *SLC4A11* and *MFSD3* genes were predicted. The results showed that mC-1 of *SLC4A11* was located in the NF-1 and Antp binding region, while mC-23 was

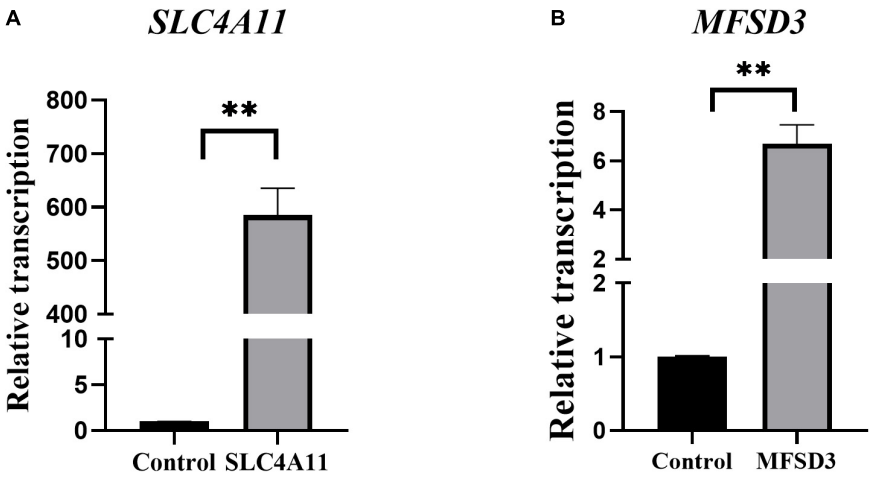


**FIGURE 3 |** Construction of recombinant plasmids and cell lines overexpressing pEGFP-C1-SLC4A11, pEGFP-C1-MFSD3. **(A)** pEGFP-C1-SLC4A11 plasmid expression efficiency. $^{**}P < 0.01$. **(B)** pEGFP-C1-MFSD3 plasmid expression efficiency. $^{**}P < 0.01$.

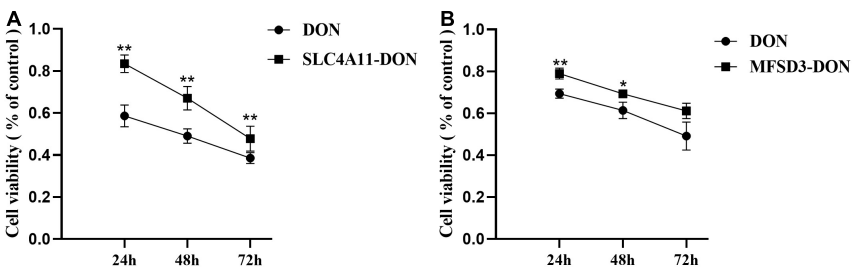


**FIGURE 4 |** Effects of DON induction on the viability of *SLC4A11* and *MFSD3* overexpressing IPEC-J2 cell lines after 24, 48, and 72 h. **(A)** The cell viability of the *SLC4A11* cell line and DON treated group induced by DON. $^{**}P < 0.01$. **(B)** The cell viability of the *MFSD3* cell line DON treated group. $^{*}P < 0.05$. $^{**}P < 0.01$.

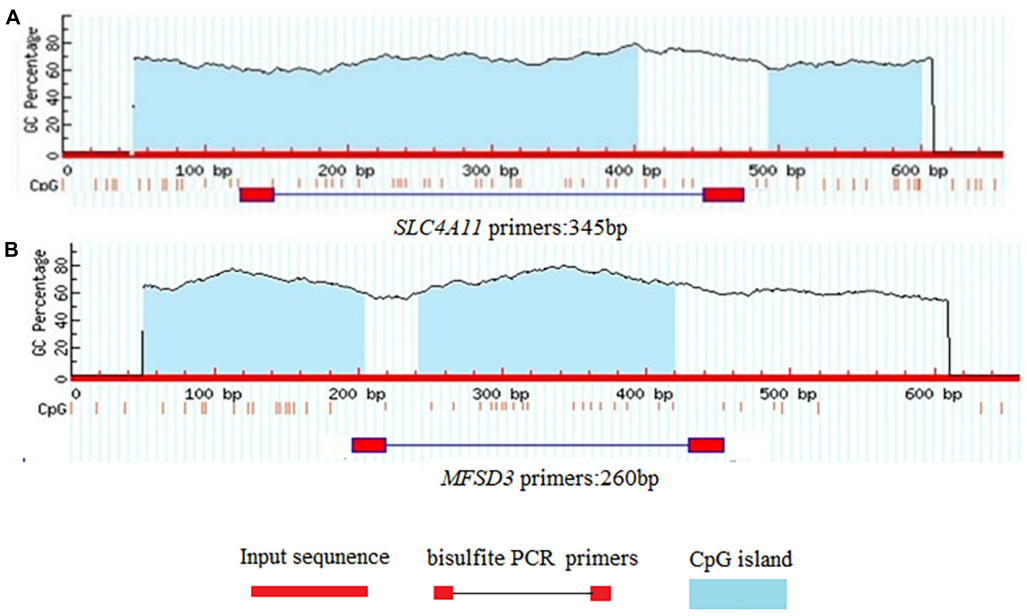


**FIGURE 5 |** Prediction of CpG islands in gene promoter regions. **(A)** CpG islands in the *SLC4A11* gene promoter region. **(B)** CpG islands in the *MFSD3* gene promoter region.

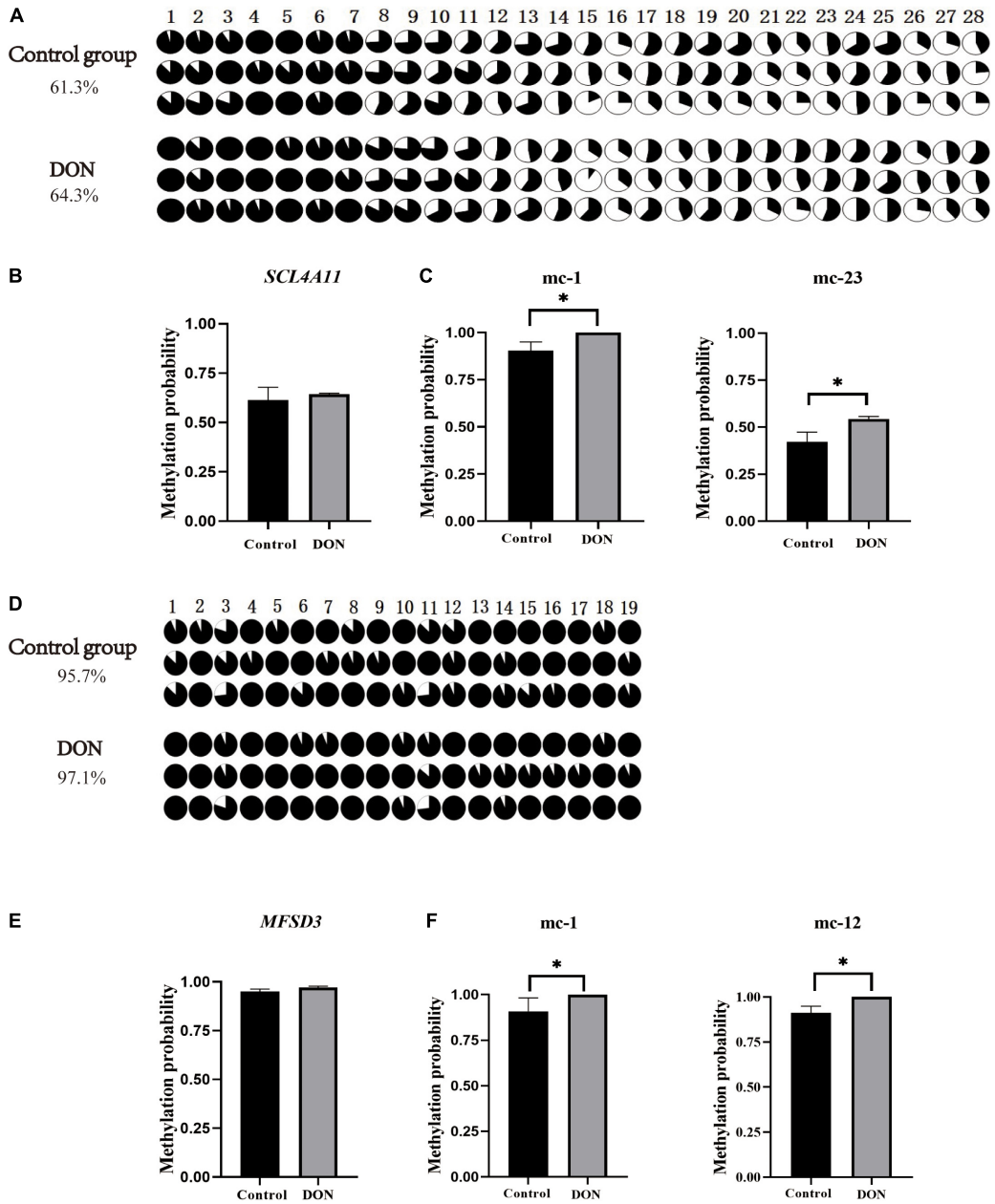


**FIGURE 6 |** CpG island methylation analysis of the *SLC4A11* and *MFSD3* gene promoter regions. **(A)** The methylation level of CpG site of the *SLC4A11* gene. **(B)** Degree of methylation of the *SLC4A11* gene between the DON treated and control groups. **(C)** the methylation levels of the mC-1 and mC-23 sites of *SLC4A11* between the DON treated and control groups. *$P < 0.05$. **(D)** The methylation level of CpG site of the *MFSD3* gene. **(E)** Degree of methylation of the *MFSD3* gene between the two groups. **(F)** the methylation levels of the mC-1 and mC-12 sites of *MFSD3* between the DON treated and control group. *$P < 0.05$.

not located in a transcription binding site (**Figure 8A**). The mC-1 site of *MFSD3* was located in the Sp1 binding region, while mC-12 was not located in a transcription binding site (**Figure 8B**).

## DISCUSSION

The growth and function of intestinal epithelial cells might be affected by DON through various pathological mechanisms, including the activation of cell signaling and ribosomal stress (Diesing et al., 2012). In pigs, the intestines are the main organ that absorbs DON, and most of it is absorbed in the jejunum (Diesing et al., 2011). The porcine intestinal epithelial IPEC-J2 cell line provides an ideal *in vitro* model system to study porcine-specific pathogenesis, which can show basic features similar to those *in vivo* (Goyarts and Dänicke, 2006). In this study, we examined the effects of DON on the viability, cell cycle, and apoptosis of IPEC-J2 cells. We found

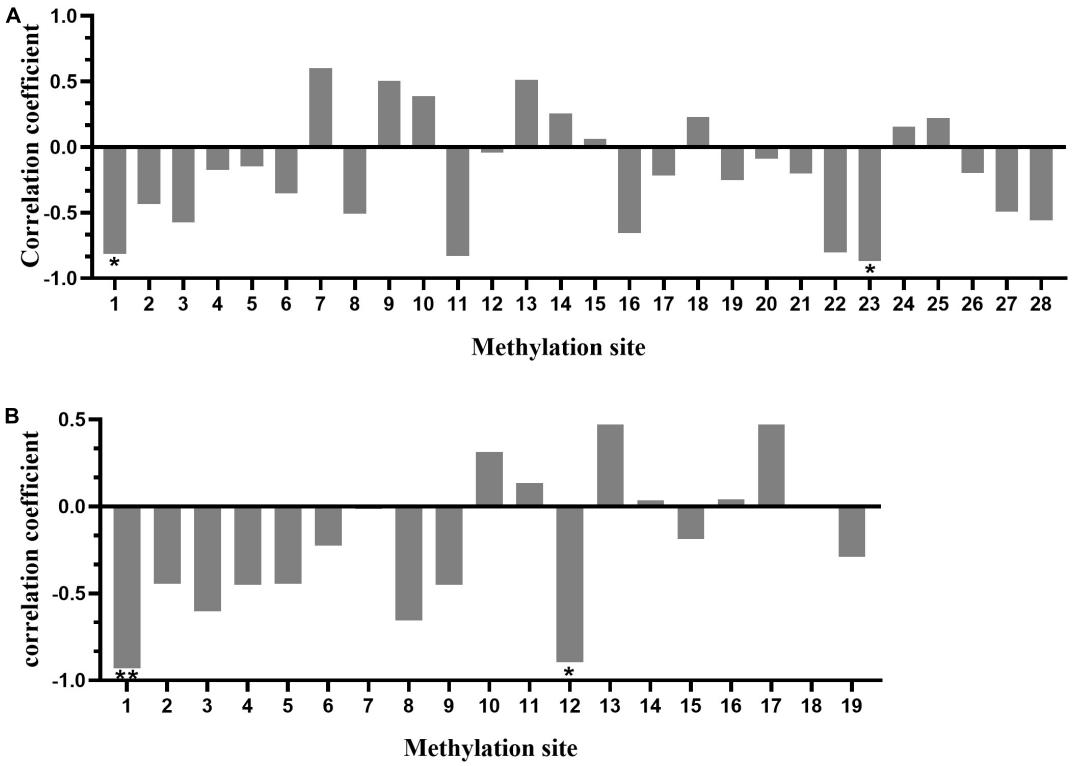


**FIGURE 7 |** Correlation between the degree of methylation of CpG islands in the gene promoter region and mRNA expression. **(A)** Correlation between the CpG island methylation degree in the promoter region of *SLC4A11* gene and mRNA expression. $^{*}P < 0.05$. **(B)** Correlation between CpG island methylation degree in the promoter region of the *MFSD3* gene and mRNA expression. $^{*}P < 0.05$; $^{**}P < 0.01$.

that DON significantly decreased cell viability, induced cell cycle arrest in the G2/M phase, and increased the apoptosis rate. Previously, it was shown that DON affects cell growth significantly by inducing IPEC-J2 cell apoptosis and arresting the cell cycle in the G2/M phase (Schierack et al., 2006), which is consistent with our results. Wang et al. (2019) showed that DON significantly increased the expression of the *BCL10* and *AEN* genes, whose protein products function in the induction and enhancement of apoptosis. Previous research also showed that DON decreases the activity of porcine endometrial cells, resulting in apoptotic phenotypes, such as mitochondrial swelling, membrane rupture, and cytoplasmic vacuolization, which hinders the synthesis of DNA, arrest the cell cycle at the G0 and G1 phases, and reduces the expression of PCNA, a key protein in the cell cycle, thus inhibiting cell proliferation (Tiemann et al., 2003; Diesing et al., 2011). Other studies have indicated the toxicological effects of DON on cell growth of different cell types via activation of the apoptosis and cell cycle arrest pathway. Our previous transcriptome analysis found that the expression levels of *SLC4A11* and *MFSD3* decreased significantly in the DON treated group. DON activates MAPKs related to differentiation and apoptosis, thereby disrupting normal cell functions. Previous research showed DON induces apoptosis and disrupts cellular homeostasis through MAPK signaling pathways in bovine mammary epithelial cells (Lee et al., 2019) and *SLC4A11* activated the MAPK pathway to stimulate cell growth and proliferation (Lopez et al., 2009). As MAPKs are important molecules in regulation of cell cycle and growth (Vithana et al., 2006), we speculate that *SLC4A11* and *MFSD3* may play a central role in DON-induced cell damage. In the present study, the expression levels of *SLC4A11* and *MFSD3* in the DON treated group were significantly lower than those in the control group. The cell viability of *SLC4A11* and *MFSD3* IPEC-J2 overexpressing cell lines under DON induction enhanced compared with the control group. *SLC4A11* and *MFSD3* can maintain nutrient physiological functions, and overexpression of *SLC4A11* and *MFSD3* may help resist the toxicity of DON. A previous study showed that relative to *SLC4A11* wild-type cells, *SLC4A11* knockout caused obvious oxidative damage to the corneal endothelial cells and depressed glutamine (Gln) catabolism, suggesting that *SLC4A11* can protect cells from Gln-induced toxicity (Bonanno and Ogando, 2018). A recent study found that *SLC4A11* plays a key role in the oxidative stress response in human corneal endothelial cells (HCEnC) mediated by the *NRF2* gene, and overexpression of *SLC4A11* in HEK 293 cells resulted in a significant increase in cell viability and reduced reactive oxygen species (ROS) in these cells compared with those in the control cells (Guha et al., 2017). In addition, DON exerts a direct toxic effect on the IPEC-J2 cells by enhancing ROS accumulation, activating NF-kB and apoptotic signaling pathways (Ruifen et al., 2019). Therefore, we hypothesized that the downregulation of

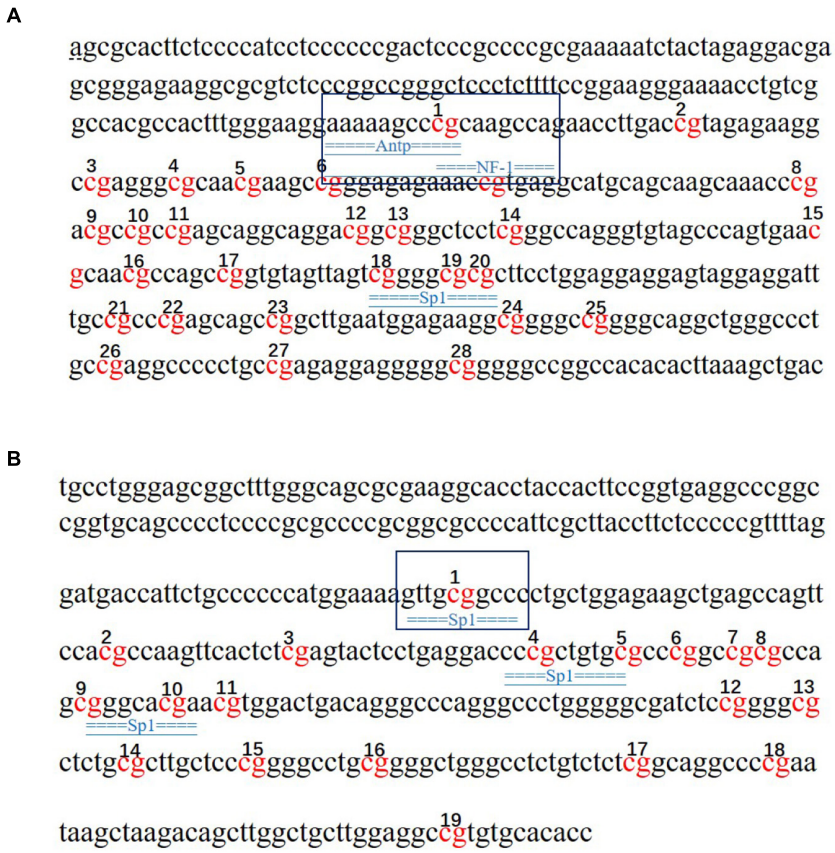


**FIGURE 8 |** Prediction of transcription factors that bind to the *SLC4A11* and *MFSD3* genes. **(A)** Prediction of transcription factor binding sites in the *SLC4A11* gene promoter region. **(B)** Prediction of transcription factor binding sites in the *MFSD3* gene promoter region. The underlined regions represent the transcription factor binding sites, and the red bases are the CpG sites.

*SLC4A11* and *MFSD3* expression is closely related to growth inhibition induced by DON.

DNA methylation is a significant epigenetic modification, which can regulate gene expression by promoting or inhibiting the ability of transcription factors to bind to DNA (Keshet et al., 1985; Lister et al., 2009). Liu et al. (2020) showed that DON affected the expression of growth related-genes in liver through DNA methylation. This suggested that epigenetic modifications play an important regulatory role in the mechanism of DON's effects. In view of the role of *SLC4A11* and *MFSD3* gene expression on DON-induced injury, the present study explored the regulatory mechanism of promoter methylation on DON-induced toxicity. The results showed that the methylation levels of *SLC4A11* gene methylation sites (mC-1 and mC-23) and *MFSD3* methylation sites (mC-1 and mC-12) correlated significantly and negatively with the mRNA expression level. Previous research showed that promoter region methylation was associated with inhibition of gene expression, which occurs mainly through interference, in which the binding of transcription factors to the target gene or histone deacetylase (HDACs) is hindered (Keshet et al., 1985; Nan et al., 1998). Our study found that the mC-1 site of *SLC4A11* is located in an NF1 transcription factor binding domain and the mC-1 site of *MFSD3* is located in an Sp1 transcription factor binding domain. Sp1-like transcription regulators participate in the regulation of cell function, including cell proliferation, apoptosis, differentiation, and tumor transformation, by regulating the expression of many genes with CG-rich promoters (Black et al., 2001). NF-1 is considered as a common transcription factor (Inoue et al., 1990). However, the interaction with specific NF-1 subtypes in different cell types might contribute to the selective transcriptional activation or silencing of target genes (Chaudhry et al., 1997). ANTP belongs to Transcription factors of the homeodomain family (Bernd et al., 2008) NF-1, ANTP and Sp1 transcription factors fulfill important roles in cell proliferation and growth. Various components in different CpG islands play their respective roles in gene regulation. Some specific CpG sites have a crucial impact on the function of CpG islands and are decisive factors for the methylation of CpG islands (Mikeska et al., 2007; Barrera and Peinado, 2012). A single methylation site in the *CCR3* (C–C motif chemokine receptor 3) region is a candidate region for causing narcolepsy (Mihoko et al., 2018). Wang et al. (2020) found that changes in the methylation degree of mC-6 site would significantly change the expression of gonadotropin releasing hormone 1 (*GNRH*), indicating that the methylation level of specific sites would affect the expression level of the gene. Therefore, it can be speculated that in the regulatory regions of *SLC4A11* and *MFSD3*, hypermethylation of

mC-1 in *SLC4A11* and *MFSD3* induced by DON in IPEC-J2 cells might hinder the binding of transcription factors to their target sequences, which would inhibit gene expression.

In the present study, we found that overexpression of *SLC4A11* and *MFSD3* can enhance the cell viability and alleviate the cytotoxicity of DON. The methylation levels of the mC-1 and mC-23 sites of *SLC4A11*, and the mC-1 and mC-12 sites of *MFSD3* showed significantly negative correlations with mRNA expression. These findings indicated that the *SLC4A11* and *MFSD3* genes may play important roles in regulating the growth of IPEC-J2 cells and the DON-induced cell damages. This study provided novel insights into the biological functions of *SLC4A11* and *MFSD3* genes in regulating the cytotoxic effects induced by DON, which lays important foundations for future studies on the identification of functional gene and toxic mechanisms associated with DON. Further studies are needed to use chip-assay to confirm transcription factor binding, to reveal the detailed relationship between *SLC4A11* and *MFSD3* promoter methylation and the specific mechanism of DON-induced cytotoxicity.

## DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

## AUTHOR CONTRIBUTIONS

WB and SW designed the study. YX, LY, and YW acquired and interpreted the data. YX analyzed the data and was a major contributor in writing the manuscript. YX and XC prepared figures and tables. ZW, HW, and WB prepared the manuscript and supervised the study. All authors read and approved the final manuscript.

## FUNDING

This research was funded by Key Research and Development Project (Modern Agriculture) of Jiangsu Province (BE2019341) and College Students' Innovation and Entrepreneurship Training Program of Yangzhou University (X20200649) and the Priority Academic Program Development of Jiangsu Higher Education Institutions.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2021.697883/full#supplementary-material

## REFERENCES

Barrera, V., and Peinado, M. A. (2012). Evaluation of single CpG sites as proxies of CpG island methylation states at the genome scale. *Nucleic Acids Res.* 40, 11490–11498. doi: 10.1093/nar/gks928

Bernd, S., Kai, K., Mansi, S., Daniel, R., Rosengarten, R. D., Dellaporta, S. L., et al. (2008). The early antp gene repertoire: insights from the placozoan genome. *PLoS One* 3:e2457. doi: 10.1371/journal.pone.0002457

Black, A. R., Black, J. D., and Cliffod, J. A. (2001). Sp1 and krüppel-like factor family of transcription factors in cell growth regulation and cancer. *J. Cell. Physiol.* 188, 143–160. doi: 10.1002/jcp.1111

Bonanno, J., and Ogando, D. (2018). SLC4A11 provides NH3 sensitive mitocho ndri- al uncoupling and ROS prevention that facilitates glutamine catabolism. *Free Radic. Biol. Med.* 128:S80.

Borutova, R., Aragon, Y. A., NäHrer, K., and Berthiller, F. (2012). Co-occurrence and statistical correlations between mycotoxins in feedstuffs collected in the Asia–Oceania in 2010. *Anim. Feed Sci. Tech.* 178, 190–197. doi: 10.1016/j.anifeedsci.2012.09.015

Chaudhry, A. Z., Lyons, G. E., and Gronostajski, R. M. (1997). Expression patterns of the four nuclear factor I genes during mouse embryogenesis indicate a potential role in development. *Dev. Dyn.* 208, 313–325.

Dahl, C., and Guldberg, P. (2003). DNA methylation analysis techniques. *Biogerontology.* 4, 233–250. doi: 10.1023/A:1025103319328

Dai, C., Yang, L., Jin, J., Wang, H., and Bao, W. (2019). Regulation and molecular mechanism of TLR5 on resistance to *Escherichia coli* F18 in Weaned Piglets. *Animals* 9:735. doi: 10.3390/ani9100735

Deng, D. J. (2014). DNA methylation and demethylation: current status and future perspective. *Heredity* 36, 403–410. doi: 10.3724/SP.J.1005.2014.0403

Diesing, A. K., Nossol, C., Panther, P., Walk, N., Post, A., Kluess, J., et al. (2011). Mycotoxin deoxynivalenol (DON) mediates biphasic cellular response in intestinal porcine epithelial cell lines IPEC-1 and IPEC-J2. *Toxicol. Lett.* 200, 8–18. doi: 10.1016/j.toxlet.2010.10.006

Diesing, A. K., Nossol, S., Ponsuksili, S., Wimmers, K., Kluess, J., Walk, N., et al. (2012). Gene regulation of intestinal porcine epithelial cells IPEC-J2 is dependent on the site of deoxynivalenol toxicological action. *PLoS One* 7:e34136. doi: 10.1371/journal.pone.0034136

Eriksen, G. S., and Pettersson, H. (2004). Toxicological evaluation of trichothecenes in animal feed. *Anim. Feed Sci. Tech.* 114, 205–239.

Gerez, J. R., Pinton, P., Callu, P., Grosjean, F., Oswald, I. P., and Bracarense, A. P. F. I. (2015). Deoxynivalenol alone or in combination with nivalenol and zear alenone induce systemic histological changes in pigs. *Exp. Toxicol. Pathol.* 67, 89–98. doi: 10.1016/j.etp.2014.10.001

Goyarts, T., and Dänicke, S. (2006). Bioavailability of the Fusarium toxin deoxynival- enol (DON) from naturally contaminated wheat for the pig. *Toxicol. Lett.* 163, 171–182. doi: 10.1016/j.toxlet.2005.10.007

Guha, S., Chaurasia, S., Ramachandran, C., and Roy, S. (2017). SLC4A11 depletion impairs NRF2 mediated antioxidant signaling and increases reactive oxygen species in human corneal endothelial cells during oxidative stress. *Sci. Rep.* 7:4074. doi: 10.1038/s41598-017-03654-4

Han, J., Wang, Q. C., Zhu, C. C., Liu, J., Zhang, Y., Cui, X. S., et al. (2016). Deoxynivalenol exposure induces autophagy/apoptosis and epigenetic modification changes during porcine oocyte maturation. *Toxicol. Appl. Pharmacol.* 300, 70–67. doi: 10.1016/j.taap.2016.03.006

Hoglund, P. J., Nordstrom, K. J., Schioth, H. B., and Fredriksson, R. (2011). The solute carrier families have a remarkably long evolutionary history with the majority of the human families present before divergence of Bilaterian species. *Mol. Biol. Evol.* 28, 153–141. doi: 10.1093/molbev/msq350

Inoue, T., Tamura, T., Furuichi, T., and Mikoshiba, K. (1990). Isolation of complementary DNAs encoding a cerebellum-enriched nuclear factor I family that activates transcription from the mouse myelin basic protein promoter. *J. Biol. Chem.* 265, 19065–19070. doi: 10.1016/0005-2728(90)90052-6

Jiao, X., Sultana, A., Garg, P., Ramamurthy, B., Vemuganti, G. K., Gangopadhyay, N., et al. (2007). Autosomal recessive corneal endothelial dystrophy (CHED2) is associated with mutations in SLC4A11. *J. Med. Genet.* 44:407. doi: 10.1136/jmg.2006.044644

Keshet, I., Yisraeli, J., and Cedar, H. (1985). Effect of regional DNA methylation on gene expression. *Proc. Natl. Acad. Sci. U. S. A.* 82, 2560–2564. doi: 10.2307/25215

Lee, J. Y., Lim, W., Park, S., Kim, J., You, S., and Song, G. (2019). Deoxynivalenol induces apoptosis and disrupts cellular homeostasis through MAPK signaling pathways in bovine mammary epi- thelial cells. *Environ. Pollut.* 252, 879–887.

Li, L. C., and Dahiya, R. (2002). Methprimer: designing primers for methylation PCRs. *Bioinformatics* 18, 1427–1431. doi: 10.1093/bioinformatics/18.11.1427

Lister, R., Pelizzola, M., Dowen, R. H., Hawkins, R. D., Hon, G., Tonti-Filippini, J., et al. (2009). Human DNA methylomes at base resolution show wide spread epigenomic differences. *Nature* 462, 315–322. doi: 10.1038/nature08514

Liu, A., Hu, S., Zhu, X., Ihsan, A., Wu, Q., Sun, L., et al. (2020). Dietary exposure of deoxynivalenol affected cytochrome P450 and growth related-gene expression via DNA methylation in piglet liver. *Research square. Preprint.* doi: 10.21203/rs.3.rs-34338/v1

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

Lopez, I. A., Rosenblatt, M. I., Kim, C., Galbraith, G. C., Jones, S. M., Kao, L., et al. (2009). SLC4A11 gene disruption in mice. *J. Biol. Chem.* 284, 26882–26896. doi: 10.1074/jbc.M109.008102

Mihoko, S., Taku, M., Hiromi, T., Katsushi, T., and Makoto, H. (2018). Epigeno mew- ide association study of DNA methylation in narcolepsy: an integrated genetic an- d epigenetic approach. *Sleep* 41, 32–42. doi: 10.1093/sleep/zsy019

Mikeska, T., Bock, C., El-Maarri, O., Hübner, A., Ehrentrau, D., Schramm, J., et al. (2007). Optimization of quantitative MGMT promoter methylation analysis using pyrosequencing and combined bisulfite restriction analysis. *J. Mol. Diagn.* 9, 368–381. doi: 10.2353/jmoldx.2007.060167

Nan, X., Ng, H. H., Johnson, C. A., Laherty, C. D., Turner, B. M., Eisenman, R. N., et al. (1998). Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. *Nature* 393, 386–389. doi: 10.1038/30764

Nicoletti, C. F., Pinhel, M. A. S., Noronha, N. Y., Jacomo, A., Crujeiras, A. B., and Nonino, C. B. (2019). Association of MFSD3 promoter methylation level and weight regain after gastric bypass: assessment for three years after surgery. *Nutrition* 70:110499. doi: 10.1016/j.nut.2019.04.010

Perland, E., Hellsten, S. V., Lekholm, E., Eriksson, M. M., Arapi, V., and Fredriksson, R. (2017). The novel membrane-bound proteins MFSD1 and MFSD3 are putative SLC transporters affected by altered nutrient intake. *J. Mol. Neurosci.* 61, 1–16. doi: 10.1007/s12031-016-0867-8

Pestka, J. J. (2010). Deoxynivalenol-induced proinflammatory gene expression: mechanisms and pathological sequelae. *Toxins* 2, 1300–1317.

Pestka, J. J., and Smolinski, A. T. (2005). Deoxynivalenol: toxicology and potential effects on humans. *J. Toxicol. Environ. Health B* 8, 39–69. doi: 10.1080/10937400590889458

Ruifen, K., Ruonan, L., Pengyuan, D., Zhaojian, L., Yansen, L., Chunmei, L., et al. (2019). Deoxynivalenol induced apoptosis and inflammation of IPEC-J2 cells by promoting ROS production. *Environ. Pollut.* 251, 689–698. doi: 10.1016/j.envpol.2019.05.026

Schierack, P., Nordhoff, M., Pollmann, M., Weyrauch, K. D., and Amasheh, S. (2006). Characterization of a porcine intestinal epithelial cell line for in vitro studies of microbial pathogenesis in swine. *Histochem. Cell Biol.* 125, 293–305. doi: 10.1007/s00418-005-0067-z

Shi, L., and Wu, J. (2009). Epigenetic regulation in mammalian preimplantation embryo development. *Reprod. Biol. Endocrinol.* 7:59. doi: 10.118-6/1477-7827-7-59

Tiemann, U., Viergutz, T., Jonas, L. F., and Schneider, F. (2003). Influence of the my- cotoxins α- and β zearalenol and deoxynivalenol on the cell cycle of cultured por- cine endometrial cells. *Reprod. Toxicol.* 17, 209–218.

Vithana, E. N., Morgan, P., Sundaresan, P., Ebenezer, N. D., Tan, D. T. H., Mohamed, M. D., et al. (2006). Mutations in sodium-borate cotransporter SLC4A11 cause recessive congenital hereditary endothelial dystro- phy (CHED2). *Nat. Genet.* 38, 755–757. doi: 10.1038/ng1824

Wang, H. F., Zong, Q. F., Wang, S. Q., Zhao, C. X., Wu, S. L., and Bao, W. B. (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, X., Li, X. J., Chen, J., Xiao, J., Li, K., and Zhou, Y. (2020). The effect of miR-29 on the methylation of specific sites in the promoter region of GnRH gene. *J. Biol.* 37, 2095–1736. (In Chinese)

Zhu, C. C., Hou, Y. J., Han, J., Liu, H. L., Cui, X. S., Kim, N. H., et al. (2014). Effect of mycotoxin-containing diets on epigenetic modifications of mouse oocytes by fluorescence microscopy analysis. *Microsc. Microanal.* 20, 1158–1166. doi: 10.1017/S1431927614000919

**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.

frontiers in Genetics

ORIGINAL RESEARCH
published: 04 August 2021
doi: 10.3389/fgene.2021.663759

# KLF4 Inhibits the Differentiation of Goat Intramuscular Preadipocytes Through Targeting C/EBPβ Directly

*Qing Xu[1,2,3†], Yanyan Li[1,3†], Sen Lin[1,2], Yong Wang[1,2,3]*, Jiangjiang Zhu[1,2] and Yaqiu Lin[1,2,3]**

[1] Key Laboratory of Qinghai-Tibetan Plateau Animal Genetic Resource Reservation and Utilization, Ministry of Education, Southwest Minzu University, Chengdu, China, [2] Key Laboratory of Sichuan Province for Qinghai-Tibetan Plateau Animal Genetic Resource Reservation and Exploitation, Southwest Minzu University, Chengdu, China, [3] College of Animal Science and Veterinary Medicine, Southwest Minzu University, Chengdu, China

Intramuscular fat (IMF) deposition is a complicated process, and most of the underlying regulators of this biological process are unknown. Here, we cloned the intact CDS of *KLF4* gene, investigated the role of KLF4 by gaining or losing function *in vitro* and further explored the pathways of KLF4 regulating differentiation of intramuscular preadipocytes in goat. Our results show that goat *KLF4* gene consists of 1,536 bp encoding a protein of 486 amino acids. The expression of KLF4 is higher in the lung while lower in the heart and muscle in goat. Knockdown of KLF4 mediated by siRNA technique significantly promotes intramuscular preadipocyte lipid accumulation and upregulates mRNA expression of adipogenic related genes including C/EBPα, C/EBPβ, and PPARγ *in vivo* cultured cells. Consistently, overexpression of KLF4 inhibits intramuscular adipocyte lipid accumulation and significantly downregulation gene expression of C/EBPβ, PPARγ, *aP2*, and Pref-1. Further, we found that other members of KLFs were upregulated or downregulated after interference or overexpression of KLF4, including KLF2 and KLF5–7. We also found that C/EBPβ was a potential target of KLF4, because it had an opposite expression pattern with KLF4 during the differentiation of intramuscular preadipocytes and had putative binding sites of KLF4. The dual-luciferase reporter assay indicated that overexpression of KLF4 inhibited the transcriptional activity of C/EBPβ. These results demonstrate that KLF4 inhibits the differentiation of intramuscular preadipocytes in goat by targeting C/EBPβ.

**Keywords: goat, KLF4, intramuscular preadipocytes, differentiation, C/EBPβ**

OPEN ACCESS

***Edited by:***
*Aline Silva Mello Cesar, University of São Paulo, Brazil*

***Reviewed by:***
*Chunping Zhao, Northwest A&F University, China*
*Jian Xu, Chinese Academy of Fishery Sciences (CAFS), China*
*Fernanda Cristina da Veiga, University of São Paulo, Brazil*

***Correspondence:**
*Yaqiu Lin*
*linyq1999@163.com*
*Yong Wang*
*wangyong010101@hotmail.com*

*† These authors have contributed equally to this work*

***Specialty section:***
*This article was submitted to Livestock Genomics, a section of the journal Frontiers in Genetics*

***Received:*** *03 February 2021*
***Accepted:*** *06 April 2021*
***Published:*** *04 August 2021*

***Citation:***
*Xu Q, Li Y, Lin S, Wang Y, Zhu J and Lin Y (2021) KLF4 Inhibits the Differentiation of Goat Intramuscular Preadipocytes Through Targeting C/EBPβ Directly. Front. Genet. 12:663759. doi: 10.3389/fgene.2021.663759*

## INTRODUCTION

Intramuscular fat (IMF) content is a crucial indicator for meat sensory quality evaluation and has been positively correlated with meat color, marbling, moisture content, and flavor (Cheng et al., 2015). Optimizing the number of intramuscular adipocytes can enhance IMF content. Alongside that, the balance between triglyceride (TG) anabolism and catabolism determines the volume of lipid droplets in intramuscular preadipocytes, which misguidedly cause IMF deposition (Gao and Zhao, 2009). IMF deposition is a complicated process that is controlled by triacylglycerol catabolic or synthesis genes and a complex network of transcription factors, including CCAAT enhancer binding protein family, peroxisome proliferator-activated receptor gamma (PPARγ),

sterol regulatory element binding protein isoform 1c (SREBP1c), Krüppel-like factor (KLF) family, lipoprotein lipase (LPL), and fatty acid binding protein (FABP4) (Lowe et al., 2011). It is believed that a complicated regulatory network exists in these molecular regulators and that many genes were involved in this biological process (Tontonoz et al., 1994; Tzameli et al., 2004). However, little information is known about the molecular regulatory mechanisms understanding the IMF deposition.

Krüppel-like factors are important transcription factors including 18 members, and they have important roles in gluconeogenesis, obesity, and the differentiation of preadipocytes (Gray et al., 2002; White and Stephens, 2010; Wang et al., 2017). *KLF4*, the core of transcription cascade of preadipocyte differentiation, is a target gene of multiple microRNAs and regulates numerous downstream genes (Rivero et al., 2012; Shen et al., 2018). The Evidence that KLF4 regulates adipogenesis is increasing. However, the discrepancies of animal species and the source (endogenous or exogenous) result in different functional significance (Cervantes-Camacho et al., 2015; Park et al., 2017). For example, KLF4 knockdown in 3T3-L1 cells downregulates the expression level of C/EBPβ and inhibits adipogenesis, suggesting that KLF4 can function as a regulator of adipogenesis and preadipocyte differentiation (Birsoy et al., 2008). On the contrary, deletion of KLF4 in white preadipocytes and brown preadipocytes had little effect on adipogenesis, indicating that endogenous KLF4 is dispensable for adipogenesis in mice (Park et al., 2017). Thus, the role of KLF4 in the differentiation of intramuscular preadipocytes in goat needs to be further investigated in view of the different cells and species.

To explore the functional effects of KLF4 on regulating the differentiation of intramuscular preadipocytes in goats, *KLF4* gene was cloned for the first time, and its expression was monitored in various tissues, and adipogenic differentiation process was detected in goats. Intramuscular preadipocytes were separated and cultured as we described previously (Xiong et al., 2018). siRNAs and adenovirus were used for transfection or infection in intramuscular preadipocytes to reveal the role of KLF4 in intramuscular preadipocyte differentiation. Further, we explored the regulation between KLF4 and other members of KLF family. Finally, we predicted the putative binding site of KLF4 and identified its target genes. Our results show that KLF4 inhibits the adipogenic differentiation of goat intramuscular preadipocytes through targeting C/EBPβ directly.

# MATERIALS AND METHODS

## Animals, Tissue Collection, and Cell Culture

The 7-day-old ($n$ = 3) and 1-year-old ($n$ = 6) Jianzhou Daer male goats were used as experimental models, and the goats were purchased from Sichuan Jianyang Dageda Animal Husbandry Co., Ltd (Sichuan, China). Animal studies were approved by the Institutional Animal Care and Use Committee, Southwest Minzu University (Chengdu, China). One-year-old goats were slaughtered by using carotid bleeding. The heart, liver, spleen, lung, kidney, subcutaneous fat, longissimus dorsi, biceps femoris, and triceps brachii were harvested and stored at liquid nitrogen. Seven-day-old goats were slaughtered on an empty stomach and washed twice with 10 mg/L of benzalkonium bromide and 75% ethyl alcohol. The longissimus dorsi was isolated under sterile conditions, washed thrice in phosphate-buffered saline (PBS) (HyClone, Logan, UT, United States) supplemented with 3% penicillin/streptomycin and minced. Triploid collagenase type II (Sigma-Aldrich Corp., St. Louis, MO, United States) was added to the minced tissues at 37°C for 1 h with gentle agitation to sufficiently digest. The same volume of DMEM/F12 (HyClone) containing 10% fetal bovine serum (FBS) was used to terminate enzymatic digestion. The cell suspension was filtrated through 75-μm mesh and centrifuged at 2,000 r/min for 5 min, and then the suspension was disposed with red blood cell (RBC) lysed solution. After centrifugation at 2,000 r/min for 5 min, the intramuscular preadipocytes were re-suspended in DMEM/F12 supplemented with 10% FBS and seed in cell culture flask. These cells were cultured in an incubator with 5% $CO_2$ at 37°C.

## Cloning and Sequence Analysis of *KLF4* Gene

Based on the predicted sequence of goat *KLF4* gene at GenBank with accession number: XM_005684390.1, clonal primers were designed with Primer Premier 5.0 software, which are KLF4-KS: 5′-TACCCCTTCTGCTTCGGA-3′ and KLF4-KA: 5′-TGTGGGTCACATCCACTGTT-3′. A 2 × GC-rich PCR MasterMix (TIANGEN, Beijing, China) and AG 22331 Hambury PCR (Thermo Fisher Scientific, Waltham, MA, United States) were used to perform polymerase chain reaction (PCR). The right separated DNA was purified by 1% agarose gel and connected with pMD-19T vector. These recombinants were converted into *Escherichia coli* DH5α (Tiangen, China). Finally, certified recombinants were submitted to TSINGKE (Chengdu, Sichuan, China) for sequencing in two directions. The tools for sequence analysis, including ORF Finder, DNAMAN, Version 5.2.10) and the Conserved Domains (CD) Search Service as described (Qing et al., 2018b), were used to analyze the open reading frame (ORF), the homology of derived amino acid sequences, and protein domain, respectively. The protein interaction network and neighbor-joining (NJ) phylogenetic tree were constructed by using STRING and MEGA7.0, respectively.

## Construction of the KLF4 Adenoviruses Vector

The coding sequence (CDS) of KLF4 was cloned into the pHBAD-EF1-MCS-3flag-CMV-GFP vector, named as pHBAD-KLF4. The SnapGene program was used for designing the primers, KLF4-Eco/Eco-Fg: gtgaccggcgcctacgccaccATGAGGC AGCCACCTGGC, KLF4-Eco/Eco-Rg: ggatcccgcccggggAAAGT GCCTTTTCATATGT (lowercase letters represent the sequence of vector). The pHBAD-KLF4 was transformed into DH5α and screened by PCR amplification. pHBAD-KLF4 and pHBAD-BHG were co-transfected in HEK293A cells by Lipofectamine$^{TM}$ 3000 (Invitrogen, Carlsbad, CA, United States) to package adenovirus. Following multiplication in 293A cells, high titer of adenovirus was harvested for KLF4 expression. The adenovirus

titer was measured by $TCID_{50}$ method. An adenovirus expression green fluorescent protein was used as the control (vector) and was stored in our laboratory.

## Chemical Synthesis of siRNA

Three gene-specific siRNAs for KLF4 were designed and synthesized by Invitrogen according to the sequence of goat KLF4 (KU041754.1), named as KLF4 siRNA-1 (5′-GACCUGGACUUUAUCCUCUCCAACUdTdT-3′), KLF4 si RNA-2 (5′-CCUACACGAAGAGUUCUCAUCUCAAdTdT-3′), and KLF4 siRNA-3 (5′-ACCACCUCGCCUUACAUAUGAAdT dT-3′). The sequence of negative control was 5′-UUCUCC GAACGUGUCACGUdTdT-3′.

## Cell Induction, Transfection, and Infection

The goat intramuscular preadipocytes whose confluence reached 80% were adipogenic inducted by DMEM/F12 supplemented with 10% FBS and 100 μM of oleic acid (Sigma) as described (Shang et al., 2014). For KLF4 knockdown, intramuscular preadipocytes at 80% confluence were transfected by siRNAs *via* Lipofectamine® RNAIMAX Reagent (Invitrogen). Ad-GFP [negative control, (NC)] or Ad-KLF4 was used to infect cells to perform the experiment to overexpress KLF4. The cells were collected and monitored at day 1 with an analysis by qPCR, Oil Red O staining, or Bodipy staining after adipogenic differentiation.

## Oil Red O Staining and Bodipy Staining

The cells were washed twice with PBS and fixed with 500 μl of 10% formaldehyde for 30 min. Then, the cells were washed and stained using the Oil Red O or Bodipy working solutions for 20 min. Cells were observed and photographed with an Olympus TH4-200 microscope after staining and washing. Finally, Oil Red O dye was extracted with 1 ml of isopropanol and the Oil Red signal was quantified by measuring the absorbance at 490 nm (OD 490) to determine the extent of differentiation.

## Total RNA Extraction and Quantitative Real-Time PCR

Total RNA was extracted from 1-year-old goats using TRIzol reagent (TaKaRa, Dalian, China) according to the manufacturer. The integrality and concentration of total RNA from cultured cell samples or tissues were detected by 1% agarose gel electrophoresis and ultraviolet spectrophotometer. Total RNA of 1 μg for each sample were reverse-transcribed by RevertAid First Strand cDNA Synthesis Kit (Thermo) according to the instructions. Peptidylprolyl isomerase A (*PPIA*) was selected to normalize the expression levels. qPCR was by using TB Green™ Premix EX Taq™ (Tli RNase H Plus) (Takara) and CFX96 (Bio-Rad, Hercules, CA, United States). Relative mRNA expressions were normalized by *UXT*. The primers' information for qPCR is listed in **Table 1**. The $2^{-\Delta\Delta Ct}$ method was used to analyze the relative expression level of each gene (Livak and Schmittgen, 2001).

**TABLE 1 |** Primer information for quantitative real-time PCR (qPCR).

| Gene (GeneBank number) | Forward sequence (5′–3′) | Reverse sequence (5′–3′) |
|---|---|---|
| *PPIA* (XM_005679322.2) | ACAAAGTCCCG AAGACAGCAG | AAGTCACCACC CTGGCACAT |
| C/EBPα (XM_018062278) | CCGTGGACAAGA ACAGCAAC | AGGCGGTCATT GTCACTGGT |
| C/EBPβ (XM_018058020.1) | CAAGAAGACGGT GGACAAGC | AACAAGTTCC GCAGGGTG |
| PPARγ (NM_001285658) | AAGCGTCAGGG TTCCACTATG | GAACCTGATGGCG TTATGAGAC |
| *aP2* (NM_001285623.1) | TGAAGTCACTCC AGATGACAGG | TGACACATTCC AGCACCAGC |
| *SREBP1* (NM_001285755) | AAGTGGTGG GCCTCTCTGA | GCAGGGGTTT CTCGGACT |
| *Pref1* (KP686197.1) | CCGGCTTCATG GATAAGACCT | GCCTCGCACTT GTTGAGGAA |
| *KLF2* (KU041748) | GCGGCAAGACC TACACCAA | TGTGCTTGCGGTAG TGGC |
| *KLF4* (KU041754.1) | GTCGGTCATCA GTGTTAGCAAAGG | ACGGTGCACGAGGA GACAGTCT |
| *KLF5* (KU041751) | CACCTCCATCCTA TGCTGCTAC | CAGCCTGGGTAATC GCAGTAGT |
| *KLF6* (KU041749) | GCCTCTGAGATC AAATTCGACA | GGAGGACTCGCTG CTCACAT |
| *KLF7* (KU041750) | TTCGGTGAGGACTT GGACTGTT | TGTCCCGAGAG AGCAGAATGTC |
| *UXT* (XM_005700842.2) | GCAAGTGGATTT GGGCTGTAAC | ATGGAGTCC TTGGTGAGGTTGT |

## Total Protein Extraction and Western Blot Analysis

The cells were washed twice with PBS. Total protein was extracted from cells using 150 μl of radioimmunoprecipitation assay (RIPA) (Biosharp, Hefei City, Anhui, China) solution supplemented with 1% phenylmethylsulfonyl fluoride (PMSF) for each well. The protein concentrations were measured by bicinchoninic acid (BCA) protein quantitation assay (KeyGen Biotech, Nanjing, China) according to the instructions. Total proteins of 40 μg were separated by sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE) and transferred to polyvinylidene difluoride (PVDF) (Roche Diagnostics, Basel, Switzerland) membrane. The membrane was blocked with 5% fat-free milk and then incubated for 2 h with antibodies to C/EBPβ (WanleiBio, Shenyang, China; WL01710)/β-actin (Abcam, Cambridge, United Kingdom; ab176323), which were diluted to 1:500. The membrane was incubated with horseradish peroxidase (HRP)-conjugated secondary antibody (diluted to 1:1,000) for 1 h. Subsequently, an immunodetection was performed using Clarity™ Western ECL Substrate (Bio-Rad) and analyzed by FluorChem R System (ProteinSimple, San Jose, CA, United States).

## Dual-Luciferase Reporter Assay

The promoter of goat C/EBPβ containing KLF4 target sites was cloned from goat genome DNA using primers tagged

with *Xho*I and *Hin*dIII restriction sites (sense primer: 5′-CCGCTCGAGCACAATCGGCCATCCCAGG-3′ and antisense primer: 5′-CCCAAGCTTTAACTGAAGGCGGGAATGGG-3′). The wild-type promoter fragment was cloned into pGL3-basic vector, named as pGL3-C/EBPβ. The constructed plasmids of (pGL3-C/EBPβ)/pGL3-basic and pHBAD-KLF4/vector were co-transfected into goat intramuscular preadipocytes using Lipofectamine® 3000 (Invitrogen). Then the cells were induced to adipogenic differentiation for 48 h; after that, the cells were harvested, and the luciferase activity was detected by an automated microplate reader.

## Phylogenetic Tree Construction

Phylogenetic tree constructed based on the deduced KLF4 amino acid sequences with NJ method in MEGA v7.0.14 (Kumar et al., 2016). The GeneBank accession number: *Ovis aries* (ALI16866.1), *Bos taurus* (NP_001098855.1), *Sus scrofa* (NP_001026952.2), *Equus caballus* (XP_005605741.1), *Homo sapiens* (ABG25917.1), *Macaca mulatta* (NP_001136265.1), *Rattus norvegicus* (NP_446165.1) *Mus musculus* (NP_034767.2) *Canis lupus familiaris* (XP_005627053.1), *Felis catus* (NP_001166915.1), *Oryctolagus cuniculus* (XP_017202748.1), *Gallus gallus* (XP_004949426.1), and *Danio rerio* (NP_001106955.1).

## Statistical Analysis

All data were presented as "means ± SD." The variance of data was analyzed by SPSS; to evaluate the significance of the differences between two groups, the means were compared using Student's *t* test. Multiple-group comparisons were performed by Duncan's multiple comparisons test. $P < 0.05$ was considered to be significant difference. All experiments in our study were carried out three times at least.

# RESULTS

## Cloning and Sequence Analysis of the *KLF4* Gene in Goats

Previous studies reported that the function discrepancies of KLF4 were observed in different animal species and cell models (Birsoy et al., 2008; Park et al., 2017). To further explore its role in goat intramuscular adipogenesis, the subcutaneous fat tissue from 1-year-old Jianzhou Daer male goats was harvested. By means

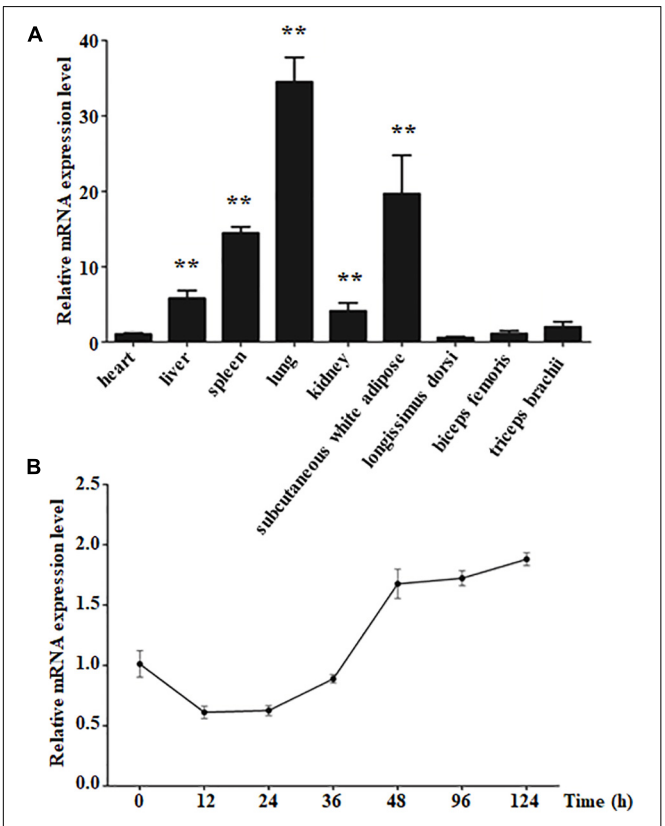


**FIGURE 1 |** The expression level of KLF4 in different tissues and during intramuscular preadipocyte differentiation. **(A)** The mRNA level of KLF4 in the heart, liver, spleen, lung, kidney, subcutaneous white adipose tissue, longissimus dorsi, biceps femoris, and triceps brachii of goats ($n = 6$). **(B)** The mRNA level of KLF4 at 0, 12, 24, 36, 48, 96, and 124 h after intramuscular preadipocytes adipogenic differentiation. The data are presented as the mean values ± SD. Each experiment was performed at least in triplicate, producing consistent results. $^{*}p < 0.05$, $^{**}p < 0.01$.

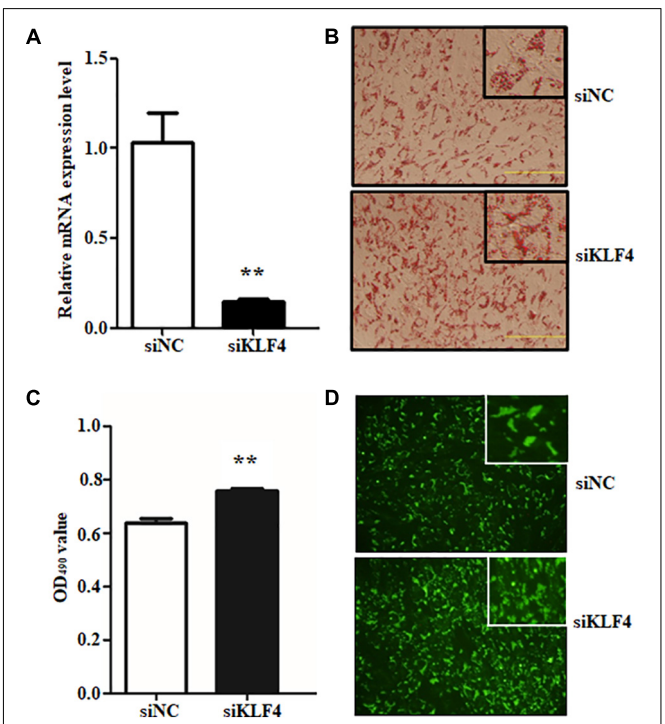


**FIGURE 2 |** Knockdown of KLF4 expression promotes goat intramuscular adipocyte adipogenic differentiation. **(A)** Knockdown efficiency detection of KLF4 at mRNA level mediated by siRNA. **(B–D)** Representative images (×400) of Bidipy staining and Oil Red O staining and extracting between control and siRNA-treatment intramuscular adipocytes. $^{**}p < 0.01$ versus negative control (siNC) groups. The Data are presented as mean values ± SD. Each experiment was performed at least in triplicate, producing consistent results. $^{*}p < 0.05$, $^{**}p < 0.01$.

of PCR amplification, the complete CDS of KLF4 (GenBank No. KU041754.1) was obtained, which contains 1,461 bp and codes 311 amino acids (**Supplementary Figure 1A**). The deduced protein contains three zinc finger structures, which were the representative structure of KLFs (**Supplementary Figure 1B**). The amino acid sequence of goat KLF4 shows a similarity of 97.94, 97.94, 92.61, and 92.39% with *O. aries*, *B. taurus*, *H. sapiens*, and *M. musculus*, respectively (**Supplementary Figure 2**). According to the corresponding amino acid sequences, an NJ phylogenetic tree was constructed, which suggested that the KLF4 protein in goat has higher genetic relationship with *O. aries* and *B. taurus* (**Supplementary Figure 2**).

## The Expression Pattern of KLF4 in Various Goat Tissues and During Intramuscular Preadipocyte Differentiation

To explore the characteristics of KLF4 expression, qPCR was performed to detect its expression in various goat tissues. The result showed that KLF4 expression was the highest in the lung, whereas it was of middle expression level in the spleen and subcutaneous white adipose tissue with ∼13.7- and 18.6-fold change to that of the heart, respectively (**Figure 1A**). Alongside that, KLF4 presents a similar mRNA level in the heart, longissimus dorsi, biceps femoris, and triceps brachii, where the expression level was the lowest (**Figure 1A**). In view of subcutaneous white adipose tissue will be fractionated into mature adipocytes and stromal vascular fraction and the skeletal muscle tissues containing minimal intramuscular adipocytes, we isolated intramuscular preadipocytes from longissimus dorsi and induced them to adipogenic differentiation. The expression of KLF4 during differentiation showed that KLF4 mRNA abundance ratio was downregulated for the previous 36 h during differentiation and exhibited an increasing trend from 36 to 124 h (**Figure 1B**). All the above suggest that KLF4 might regulate the adipogenic differentiation of goat intramuscular preadipocytes.

## Loss of Function of KLF4 Promotes Intramuscular Preadipocyte Differentiation

To reveal the function of KLF4 on intramuscular preadipocyte differentiation, three independent siRNAs were transfected into preadipocytes to knockdown of KLF4 expression. The interference efficiency assay suggested that the siRNAs dramatically decreased the expression of KLF4 (**Supplementary Figure 3**). Interference efficiency reached 71.62, 85.11, and 61.50% compared with negative control (siNC) for siRNA1, siRNA2, and siRNA3, respectively (**Supplementary Figure 3**). Because of the prominent effect of siRNA2, the subsequent KLF4 knockdown experiment was performed by siRNA2; we named it siKLF4 (**Figure 2A**). At morphological observation, which

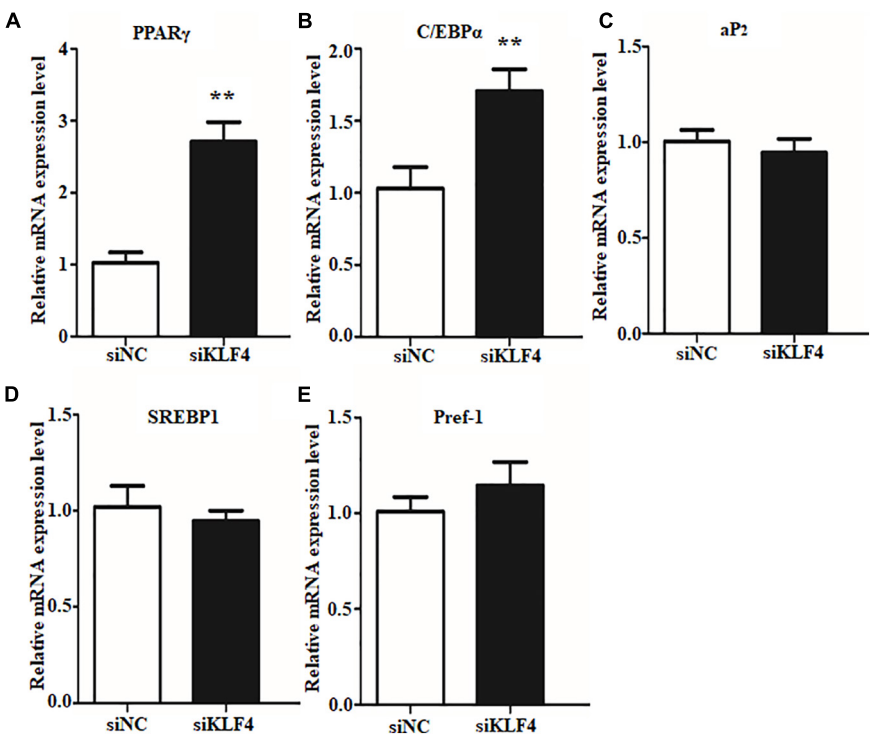


**FIGURE 3 |** Knockdown of KLF4 expression upregulates the mRNA expression of C/EBPα and PPARγ. The mRNA levels of PPARγ **(A)**, C/EBPα **(B)**, *aP2* **(C)**, *SREBP1* **(D)**, and Pref1 **(E)** in control and KLF4 siRNA-treated intramuscular adipocytes. *$p < 0.05$, **$p < 0.01$ versus negative control (siNC) groups. The data are presented as the mean values ± SD. Each experiment was performed at least in triplicate, producing consistent results. *$p < 0.05$, **$p < 0.01$.

was performed by Oil Red O staining and Bodipy staining, we found that not only lipid accumulation was augmented but also lipid droplets were increased, both of which were caused by knockdown of KLF4 expression (**Figures 2B–D**). Consistently, the Oil Red O signal was significantly increased in siKLF4 group compared with that of siNC as shown in **Figure 2C**. These data suggest that knockdown of KLF4 expression can increase intramuscular preadipocyte lipid accumulation.

Preadipocyte differentiation required the expression of numerous lipogenic genes, and then the cells present lipid accumulation and cellular morphology change (Ali et al., 2013; Cohen and Spiegelman, 2016). To explore the enhancive lipid content in KLF4 knockdown intramuscular preadipocytes whether caused by lipolysis inhibiting or promoting differentiation into adipocytes, we further detected the expression changes of several adipogenic and lipolysis genes, including C/EBPα, PPARγ, *SREBP1*, *aP2*, and *Pref1*. We found that the interference of *KLF4* upregulated the mRNA level of PPARγ, which increased by ∼2.6-fold in siRNA2-treated cells (**Figure 3A**). Moreover, the mRNA level of C/EBPα, cooperatively active with PPARγ, was also upregulated in KLF4 knockdown group as shown in **Figure 3B**. However, the mRNA levels of *aP2*, *SREBP1*, and *Pref1* were comparable with those of control cells (**Figures 3C–E**). These evidences suggest that loss function of KLF4 promotes adipogenic genes expression of PPARγ and C/EBPα and may promote goat intramuscular preadipocyte differentiation.

## Overexpression of KLF4 Inhibits Goat Intramuscular Preadipocyte Differentiation

To further illustrate the role of KLF4 in the differentiation of goat intramuscular preadipocytes, stable KLF4 overexpression (KLF4) and the negative control (vector) cells were established by adenovirus-mediated technique. Overexpression gave rise to ∼71.3-fold increase of KLF4 mRNA level when compared with vector (**Figure 4A**). In addition, overexpression of KLF4 substantially inhibited lipid contents based on the evaluation of Oil Red O staining and Bodipy staining (**Figures 4B–D**). Consistently, Oil Red O signal measurement showed that the OD value at 490 nm was decreased in overexpression group (**Figure 4C**). These data demonstrated that gain of function KLF4 can inhibit intramuscular preadipocyte differentiation.

The efficient evidences between gain and loss function of KLF4 in differentiated intramuscular preadipocytes confirmed our initial speculation. And the speculation was further confirmed by detecting the expression of adipogenic and lipolysis genes as described above. Overexpression of KLF4 significantly downregulated the mRNA level of PPARγ, as shown in **Figure 5A**. Surprisingly, *aP2* and *Pref1*, two key genes in TG synthesis and differentiation, were also inhibited by KLF4 overexpression (**Figures 5C,E**). However, no obvious change was observed about mRNA level of C/EBPα and *SREBP1* when compared with the control group (**Figures 5B,D**). All the above suggest that KLF4 inhibits intramuscular preadipocyte differentiation and gene expression of PPARγ, *aP2*, and *Pref1*.

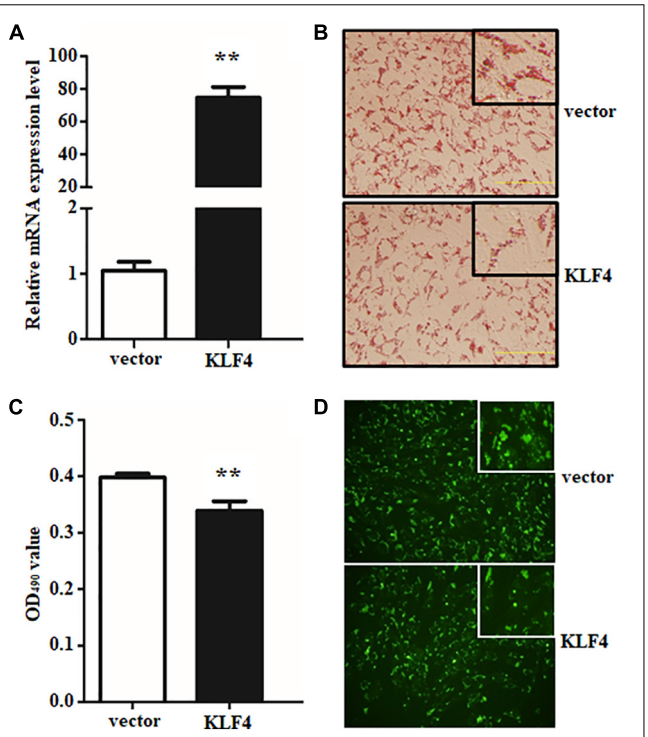


**FIGURE 4 |** Overexpression of KLF4 inhibits goat intramuscular adipocyte adipogenic differentiation. **(A)** KLF4 Overexpression efficiency detection at mRNA level mediated by adenovirus. **(B–D)** Representative images (×400). The Bodipy staining and Oil Red O staining and extracting between control and adenovirus treatment intramuscular adipocytes. $^{**}p < 0.01$ versus the negative control (vector) groups. Data are presented as mean values ± SD. Each experiment was performed at least in triplicate, producing consistent results. $^{*}p < 0.05$, $^{**}p < 0.01$.

## KLF4 Regulates mRNA Expression of Other Members of KLF Family

The internal regulation of KLF family members has been reported to be a well-orchestrated multistep process (Eaton et al., 2008; Guo et al., 2018). To determine whether the expression of the other members of KLF family will be affected by KLF4 in goat intramuscular preadipocytes, the mRNA levels of KLF family members were measured by qPCR technique. We first noticed the expression of KLF1, KLF2, and KLF17 belonging to the same subfamily with KLF4. Knockdown of KLF4 significantly enhanced the mRNA level of KLF2 as shown in **Figure 6A**. Similarly, overexpression of KLF4 suppressed the mRNA level of KLF2 by ∼0.31-fold compared with control group (**Figure 6B**). However, knockdown or overexpression of KLF4 did not affect the expression of KLF1 and KLF17 (data was not shown). KLF5, KLF6, and KLF7, belonging to the same subfamily of KLF, exhibited a dramatically higher mRNA level in KLF4 knockdown preadipocytes and a markedly lower mRNA level in KLF4 overexpression cells (**Figures 6A,B**). No significant change was recorded of KLF3, KLF4, KLF8, KLF9, KLF10, KLF11, KLF12, KLF13, KLF14, KLF15, and KLF16 (data not shown). On the whole, these data imply that KLF4

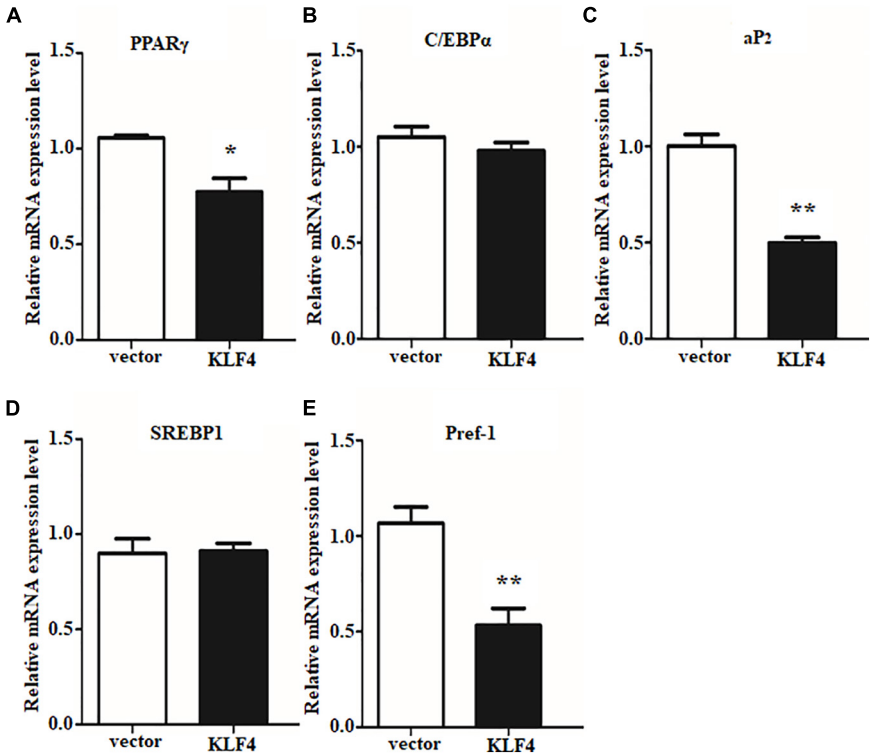


**FIGURE 5 |** Overexpression of KLF4 downregulates PPARγ, aP2, and Pref1 expression. The mRNA levels of PPARγ **(A)**, C/EBPα **(B)**, aP2 **(C)**, SREBP1 **(D)**, and Pref1 **(E)** in control and adenovirus-treated intramuscular adipocytes. $^{*}p < 0.05$, $^{**}p < 0.01$ versus negative control (vector) groups. The data are presented as the mean values ± SD. Each experiment was performed at least in triplicate, producing consistent results. $^{*}p < 0.05$, $^{**}p < 0.01$.

inhibits the expression of KLF2, KLF5, KLF6, and KLF7 in goat intramuscular preadipocytes.

## KLF4 Regulates Intramuscular Preadipocyte Differentiation Through Targeting C/EBPβ Directly

Based on the fact that KLF4 inhibits goat preadipocyte differentiation and regulates the expression of several adipogenic genes, the potential downstream targets of KLF4 should be identified. Interestingly, we found that C/EBPβ expression was opposite with KLF4 during goat intramuscular preadipocyte differentiation (Xiong et al., 2018). Combining that C/EBPβ is a very important gene in regulating preadipocyte differentiation and bioinformatic analysis showed that C/EBPβ may be a target of KLF4 stimulated us to further investigate the regulation between C/EBPβ and *KLF4* expression. We found that both mRNA and protein level of C/EBPβ were significantly upregulated by KLF4 interference when compared with the control groups (**Figures 7A,B**). Consistently, KLF4 overexpression visibly downregulated both the mRNA and the protein level of C/EBPβ (**Figures 7A,B**). Thus, we speculated that C/EBPβ might be a potential target gene of KLF4. To confirm this hypothesis, we first analyzed the transcriptional binding DNA motif of KLF4 using JASPAR software[1]. The result showed that KLF4 binding sequence is TGGGTGGGGC as shown in **Figure 7C**. Then we analyzed the promoter region of C/EBPβ and found 77 potential binding sites of KLF4, and the top 12 is shown in **Figure 7D**. Further, through dual-luciferase reporter assay, we showed that the transcriptional activity of C/EBPβ in cells was inhibited by KLF4 by about 73.18% when compared with the control group (**Figure 7E**). These data jointly suggest that KLF4 inhibits goat intramuscular preadipocyte differentiation through targeting C/EBPβ directly.

[1]http://jaspar.genereg.net/

# DISCUSSION

Intramuscular fat content is an important indicator for meat quality, and adipose lipid deposition is related to numerous genes.

In this study, we revealed that goat KLF4 has three zinc finger motifs similar to those of other species (Pearson et al., 2008; Wu and Wang, 2013), which were located at the C-terminus of the protein and recognized as GC- and CACCC-boxes of DNA. Based on the STRING analysis, KLF4 protein might interact with several proteins, such as FGF2, FGF4, FGF5, BMP4, MYC, and FOXA2 (data not shown). A previous study reported that KLF4/STAT3 pathway plays a critical role in podocyte injury and regulates aberrant glomerular epithelial cell (GEC) proliferation (Estrada et al., 2018). KLF4 and MYC drive the differentiation

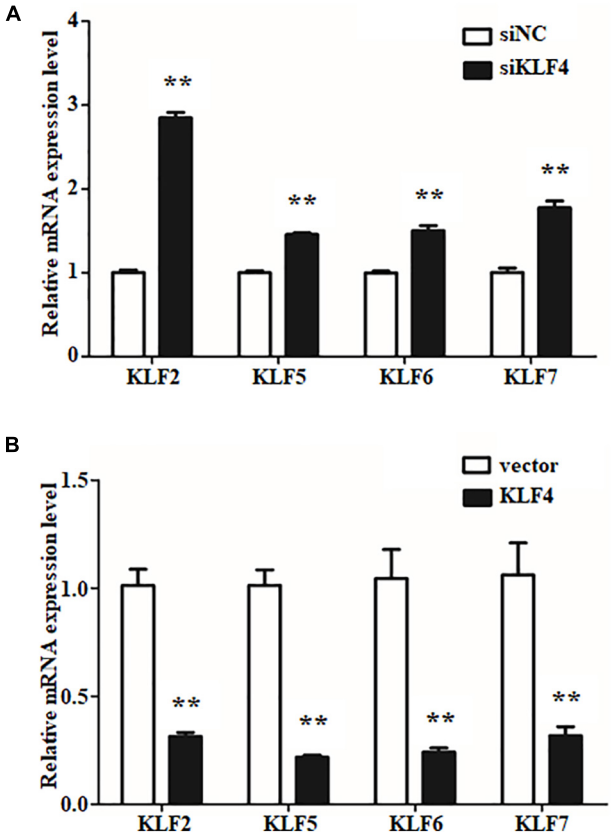


**FIGURE 6 |** The expressions of other KLF members in gain or loss function of KLF4 intramuscular adipocytes. The mRNA levels of *KLF2*, *KLF5*, *KLF6*, and *KLF7* in knockdown **(A)** and overexpression **(B)** KLF4 intramuscular adipocytes. $^{*}p < 0.05$, $^{**}p < 0.01$ versus negative control groups. The data are presented as the mean values ± SD. Each experiment was performed at least in triplicate, producing consistent results. $^{*}p < 0.05$, $^{**}p < 0.01$.

of cardiac mesenchymal progenitors into adipocytes (Kami et al., 2016). Furthermore, KLF4 was upregulated by BMP4 to promote a change of esophageal squamous epithelium in phenotype (Yan et al., 2016). KLF4, as a novel regulator of lipid signal, enhances the FGF2 induction of SK1 expression in endothelial (de Assuncao et al., 2014), and our previous research showed that the interaction between KLFs and FGFs might exist in differentiation of goat intramuscular preadipocytes (Qing et al., 2018a). Altogether, KLF4 is an important transcription factor involved in a variety of biological functions. Thus, understanding that KLF4 regulates intramuscular adipogenesis might be important to regulate the IMF content for improving the meat quality of goat.

The expression characteristic of genes is indispensable for understanding specific functional mechanism. Researches have shown that KLF4 is highly expressed in the late stage of embryonic development and many types of epithelial cells (Garrett-Sinha et al., 1996; Foster et al., 2000). Nevertheless, the expression pattern of KLF4 in goats is unclear. We examined the expression pattern of KLF4 in multiple tissues and found that the highest level of KLF4 in lung, which was supported by KLF4, inhibits lung carcinoma growth through downregulating hTERT expression and telomerase activity in mice (Hu et al., 2016). More KLF enrichment to SWAT pointed that it might regulate fat deposition in goats on account that KLF4 plays a critical role in differentiation of preadipocytes (Park et al., 2017). The reason of the heart having the lower relative expression level of KLF4 is a vital role of KLF4 in cardiac hypertrophy and atherosclerotic plaque (Yoshida et al., 2014; Shankman et al., 2015). However, the lowest mRNA level of KLF4 in muscle, including LD, BF, and TB, could be explained by these tissues containing a small proportion of intramuscular preadipocytes (Hausman et al., 2014). To better uncover the role of KLF4 in goat intramuscular adipocytes, we investigated its temporal expression during adipogenic differentiation. Surprisingly, the mRNA level of KLF4 was lower at early stage of differentiation in accordance with our previous study based on RNA -seq analysis (data not shown), whereas it reached to the peak at 2 h during differentiation in 3T3-L1 cells (Birsoy et al., 2008; Park et al., 2017). It can be supported by the discrepancies in selected differentiation stages and experimental conditions of experimental animals (Jiang et al., 2015). However, Birsoy and colleagues found that KLF4, as an early regulator, promotes adipogenesis through inducing the expression of C/EBPβ in 3T3-L1 cells (Birsoy et al., 2008). As mentioned above, whether KLF4 acts as a positive or negative regulator in goat intramuscular adipocytes should be further elucidated.

In this study, knockdown of KLF4 in intramuscular preadipocytes dramatically increased lipid accumulation and mRNA level of several adipogenic genes. The number and size of lipid droplets were reduced by KLF4 knockdown; moreover, this effect was partially attributable to the higher expression of C/EBPα and PPARγ, but not to downregulate a negative regulator for adipogenesis genes such as *Pref1*. The C/EBPα and PPARγ were the major regulatory genes for adipogenesis, and C/EBPα promotes adipocytes into terminal differentiation to form mature adipocytes (Satoh et al., 2011; Mota de Sá et al., 2017). Further, we demonstrated that KLF4 is indispensable for differentiation by infecting KLF4 adenovirus into intramuscular preadipocytes in goats. The intramuscular adipogenesis was significantly enhanced in adenovirus-treated cells. Interestingly, the mRNA level of PPARγ was in proportion to lipid content in both interference and overexpression groups. This result might be explained by the nature ligands of PPARγ, which contain some fatty acids and derivatives (Tzameli et al., 2004). Moreover, the lower expression of *aP2*, an important TS synthesis gene, revealed that the lipid synthesis activity is weak in overexpressing KLF4, compared with NC groups. The Pref1 was also inhibited; however, the mechanisms underlying the inhibition of Pref1 in adipogenic differentiation remain unclear. Taken together, KLF4 antagonizing PPARγ inhibits the differentiation of intramuscular preadipocytes in goats. On the contrary, KLF4, as a positive regulator, promotes differentiation of preadipocytes in 3T3-L1 cells (Birsoy et al., 2008). We speculate that this might be related to the specificity of cells and species. In our previous studies, we found that many members of KLF family play differential roles in

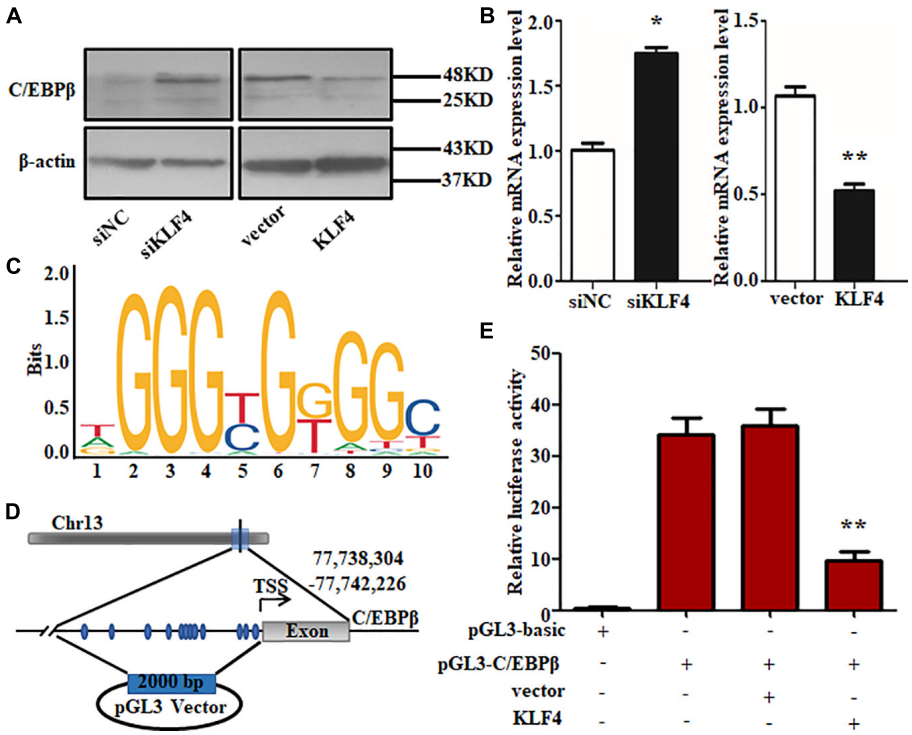


**FIGURE 7 |** KLF4 effects on intramuscular adipocyte differentiation through targeting C/EBPβ directly. The protein **(A)** and mRNA **(B)** levels of C/EBPβ in KLF4 siRNA and adenovirus-treated cells. **(C)** KLF4 binding DNA motif. **(D)** The predicted of KLF4 binding sites at C/EBPβ promoters. **(E)** Luciferase assay of transfected with pGL3-C/EBPβ P in goat intramuscular adipocytes. The data are presented as the mean values ± SD. Each experiment was performed at least in triplicate, producing consistent results. $^{*}p < 0.05$, $^{**}p < 0.01$.

subcutaneous preadipocyte and intramuscular preadipocytes differentiation in goats.

Recently, transcription factor cross-talk in KLF family attracts much attention in adipogenesis. For instance, KLF15 triggers adipogenesis through promoting transcription of KLF3 (Mori et al., 2005; Guo et al., 2018). As expected, our data show that KLF4 inhibits the expression of KLF2 and KLF5–7. In chicken, it was demonstrated that KLF2 inhibits adipogenesis partly through downregulating the expression of PPARγ and C/EBPα (Zhang et al., 2014), but the reason why KLF4 inhibits KLF2 and the regulatory relationships among them are unknown during differentiation of intramuscular preadipocytes in goats. In addition, we also reveal that KLF4 has an effect on the expression of KLF5-7 during adipogenic differentiation. According to KLF member clusters in the phylogenetic analyses, *KLF5*, *KLF6*, and *KLF7* are divided into the same subgroups (Chen et al., 2010). *KLF5* cooperates with C/EBPβ and C/EBPδ to activate the transcription of PPARγ (Oishi et al., 2011). The expression of PPARγ was controlled by *KLF6* to promote TG accumulation (Escalona-Nande et al., 2015). And *KLF7* promotes the differentiation through *aP2*, C/EBPα, C/EBPβ, and *SREBP1c* in our previous research (data not shown). Thus, KLF4 antagonizes with *KLF5*, *KLF6*, and *KLF7* during differentiation of intramuscular preadipocytes in goats. As mentioned above, the network of KLF4 inhibition role has been described in **Supplementary Figure 4**.

On the other hand, we found that the expression pattern of KLF4 and C/EBPβ is interesting, and the transcriptional binding site of KLF4 is found in the promoter of C/EBPβ. Meanwhile, the mRNA and protein levels of C/EBPβ are blocked by KLF4 during adipogenic differentiation. Based on the dual-luciferase reporter assay, we reveal that, in intramuscular adipocytes of goat, the pGL3-C/EBPβ transcriptional activity is decreased by KLF4 overexpression. These results indicate that C/EBPβ is a marker gene for differentiation with putative KLF4 binding sites in promoter region, which is similar to those reported in a previous study where KLF4 binds to a 1.45- to 1.1-kb region of C/EBPβ promoter to take part in the differentiation (Birsoy et al., 2008). Also, we speculate that the decreased expression level of PPARγ may be a secondary effect caused by C/EBPβ. Meanwhile, KLF2 and KLF5-7 might interact with C/EBPβ or PPARγ as mentioned above. Therefore, KLF4 directly downregulates the transcriptional expression of C/EBPβ and inhibits expressions of KLF2 and KLF5-7 and further results in the decrease of PPARγ expression and lipid accumulation at the early stage of differentiation in goat intramuscular preadipocytes. However, the expression of KLF4 at terminal stage is higher than that of early stage. Whether the regulatory effects of KLF4 are different between early and terminal stages or a feedback loop exists in C/EBPβ and KLF4 (Birsoy et al., 2008) in intramuscular preadipocytes of goats remain to be further explored.

## CONCLUSION

In summary, we showed that KLF4 inhibits the differentiation of intramuscular preadipocytes in goat through targeting C/EBPβ directly and regulates the expression of *KLF2* and *KLF5–7*. These results suggest that *KLF4* is a new candidate gene of regulating IMF deposition in goat.

## DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

## ETHICS STATEMENT

The animal study was reviewed and approved by the Institutional Animal Care and Use Committee, Southwest Minzu University.

## AUTHOR CONTRIBUTIONS

QX, YYL, SL, YW, JZ, and YL: concept and design. QX, YYL, YW, and YL: development of methodology. QX and YL: acquisition of data and analysis and interpretation of data. QX, YYL, YW, and YL: writing, review, and/or revision of the manuscript. YYL, SL, YW, and JZ: administrative, technical, or material support, YW and YL: study supervision.

## FUNDING

This study was supported by grants from the National Natural Science Foundation of China (31672395 and 32072723) and by the Fundamental Research Funds for the Central Universities, Southwest Minzu University (2019NQN51).

## ACKNOWLEDGMENTS

We are grateful to Wenlin Bai for proofreading this article and modifying the statement critically.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2021.663759/full#supplementary-material

## REFERENCES

Ali, A. T., Hochfeld, W. E., Myburgh, R., and Pepper, M. S. (2013). Adipocyte and adipogenesis. *Eur. J. Cell Biol.* 92, 229–236.

Birsoy, K., Chen, Z., and Friedman, J. (2008). Transcriptional regulation of adipogenesis by KLF4. *Cell Metab.* 7, 339–347. doi: 10.1016/j.cmet.2008.02.001

Cervantes-Camacho, C., Beltrán-Langarica, A., Ochoa-Uribe, A. K., Marsch-Moreno, M., Ayala-Sumuano, J. T., Velez-delValle, C., et al. (2015). The transient expression of Klf4 and Klf5 during adipogenesis depends on GSK3β activity. *Adipocyte* 4, 248–255. doi: 10.1080/21623945.2015.1007823

Chen, Z., Lei, T., Chen, X., Zhang, J., Yu, A., Long, Q., et al. (2010). Porcine KLF gene family: structure, mapping, and phylogenetic analysis. *Genomics* 95, 111–119. doi: 10.1016/j.ygeno.2009.11.001

Cheng, W. W., Cheng, J. H., Sun, D. W., and Pu, H. B. (2015). Marbling analysis for evaluating meat quality: methods and techniques. *Comprehens. Rev. Food Sci. Food Safety* 14, 523–535. doi: 10.1111/1541-4337.12149

Cohen, P., and Spiegelman, B. M. (2016). Cell biology of fat storage. *Mol. Biol. Cell* 27, 2523–2527. doi: 10.1091/mbc.e15-10-0749

de Assuncao, T. M., Lomberk, G., Cao, S., Yaqoob, U., Mathison, A., Simonetto, D. A., et al. (2014). New role for Kruppel-like factor 14 as a transcriptional activator involved in the generation of signaling lipids. *J. Biol. Chem.* 289, 15798–15809. doi: 10.1074/jbc.m113.544346

Eaton, S. A., Funnell, A. P., Sue, N., Nicholas, H., Pearson, R. C., and Crossley, M. (2008). A network of Kruppel-like Factors (Klfs). Klf8 is repressed by Klf3 and activated by Klf1 in vivo. *J. Biol. Chem.* 283, 26937–26947. doi: 10.1074/jbc.m804831200

Escalona-Nande, I., Guerrero-Escalera, D., Estanes-Hernandez, A., Ortíz-Ortega, V., Tovar, A. R., and Pérez-Monter, C. (2015). The activation of peroxisome proliferator-activated receptor gamma is regulated by Kruppel-like transcription factors 6 & 9 under steatotic conditions. *Biochem. Biophys. Res. Commun.* 458, 751–756. doi: 10.1016/j.bbrc.2015.01.145

Estrada, C. C., Paladugu, P., Guo, Y., Pace, J., Revelo, M. P., Salant, D. J., et al. (2018). Krüppel-like factor 4 is a negative regulator of STAT3-induced glomerular epithelial cell proliferation. *JCI Insight* 3:e98214.

Foster, K. W., Frost, A. R., McKie-Bell, P., Lin, C. Y., Engler, J. A., Grizzle, W. E., et al. (2000). Increase of GKLF messenger RNA and protein expression during progression of breast cancer. *Cancer Res.* 60, 6488–6495.

Gao, S. Z., and Zhao, S. M. (2009). Physiology, affecting factors and strategies for control of pig meat intramuscular fat. *Recent Patents Food Nutritions Agricul.* 1, 59–74. doi: 10.2174/1876142910901010059

Garrett-Sinha, L. A., Eberspaecher, H., Seldin, M. F., and de Crombrugghe, B. (1996). A gene for a novel zinc-finger protein expressed in differentiated epithelial cells and transiently in certain mesenchymal cells. *J. Biol. Chem.* 271, 31384–31390. doi: 10.1074/jbc.271.49.31384

Gray, S., Feinberg, M. W., Hull, S., Kuo, C. T., Watanabe, M., Sen-Banerjee, S., et al. (2002). The Kruppel-like factor KLF15 regulates the insulin-sensitive glucose transporter GLUT4. *J. Biol. Chem.* 277, 34322–34328. doi: 10.1074/jbc.m201304200

Guo, H., Khan, R., Raza, S. H. A., Ning, Y., Wei, D., Wu, S., et al. (2018). KLF15 promotes transcription of KLF3 gene in bovine adipocytes. *Gene* 659, 77–83. doi: 10.1016/j.gene.2018.03.049

Hausman, G. J., Basu, U., Du, M., Fernyhough-Culver, M., and Dodson, M. V. (2014). Intermuscular and intramuscular adipose tissues: bad vs. good adipose tissues. *Adipocyte* 3, 242–255. doi: 10.4161/adip.28546

Hu, W., Jia, Y., Xiao, X., Lv, K., Chen, Y., Wang, L., et al. (2016). KLF4 downregulates hTERT expression and telomerase activity to inhibit lung carcinoma growth. *Oncotarget* 7, 52870–52887. doi: 10.18632/oncotarget.9141

Jiang, S., Wei, H., Song, T., Yang, Y., Zhang, F., Zhou, Y., et al. (2015). KLF13 promotes porcine adipocyte differentiation through PPARγ activation. *Cell Biosci.* 5:28.

Kami, D., Kitani, T., Kawasaki, T., and Gojo, S. (2016). Cardiac mesenchymal progenitors differentiate into adipocytes via Klf4 and c-Myc. *Cell Death Dis.* 7:e2190. doi: 10.1038/cddis.2016.31

Kumar, S., Stecher, G., and Tamura, K. (2016). MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* 33, 1870–1874. doi: 10.1093/molbev/msw054

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

Lowe, C. E., O'Rahilly, S., and Rochford, J. J. (2011). Adipogenesis at a glance. *J. Cell Sci.* 124, 2681–2686. doi: 10.1242/jcs.079699

Mori, T., Sakaue, H., Iguchi, H., Gomi, H., Okada, Y., Takashima, Y., et al. (2005). Role of Krüppel-like factor 15 (KLF15) in transcriptional regulation of adipogenesis. *J. Biol. Chem.* 280, 12867–12875. doi: 10.1074/jbc.m410515200

Mota de Sá, P., Richard, A. J., Hang, H., and Stephens, J. M. (2017). Transcriptional regulation of adipogenesis. *Comprehens. Physiol.* 7, 635–674. doi: 10.1002/cphy.c160022

Oishi, Y., Manabe, I., and Nagai, R. (2011). Krüppel-like family of transcription factor 5 (KLF5). KLF5 is a key regulator of adipocyte differentiation. *Nihon Rinsho* 69, 264–268.

Park, Y. K., Wang, L., Giampietro, A., Lai, B., Lee, J. E., and Ge, K. (2017). Distinct roles of transcription factors KLF4, Krox20, and peroxisome proliferator-activated receptor γ in adipogenesis. *Mol. Cell. Biol.* 37:e00554-16.

Pearson, R., Fleetwood, J., Eaton, S., Crossley, M., and Bao, S. (2008). Krüppel-like transcription factors: a functional family. *Int. J. Biochem. Cell Biol.* 40, 1996–2001.

Qing, X., Sen, L., Yong, W., Jiangjiang, Z., and Yiaqiu, L. (2018a). Fibroblast growth factor 10 (FGF10) promotes the adipogenesis of intramuscular preadipocytes in goat. *Mol. Biol. Rep.* 45, 1881–1888. doi: 10.1007/s11033-018-4334-1

Qing, X., Yong, W., Jiangjiang, Z., Yanying, Z., and Yaqiu, L. (2018b). Molecular characterization of GTP binding protein overexpressed in skeletal muscle (GEM) and its role in promoting adipogenesis in goat intramuscular preadipocytes. *Anim. Biotechnol.* 20, 1–8. doi: 10.1080/10495398.2020.1800484

Rivero, S., Díaz-Guerra, M. J., Monsalve, E. M., Laborda, J., and García-Ramírez, J. J. (2012). DLK2 is a transcriptional target of KLF4 in the early stages of adipogenesis. *J. Mol. Biol.* 417, 36–50. doi: 10.1016/j.jmb.2012.01.035

Satoh, A., Stein, L., and Imai, S. (2011). The role of mammalian sirtuins in the regulation of metabolism, aging, and longevity. *Handb. Exp. Pharmacol.* 206, 125–162. doi: 10.1007/978-3-642-21631-2_7

Shang, Z., Guo, L., Wang, N., Shi, H., Wang, Y., and Li, H. (2014). Oleate promotes differentiation of chicken primary preadipocytes in vitro. *Biosci. Rep.* 34:e00093.

Shankman, L. S., Gomez, D., Cherepanova, O. A., Salmon, M., Alencar, G. F., Haskins, R. M., et al. (2015). KLF4-dependent phenotypic modulation of smooth muscle cells has a key role in atherosclerotic plaque pathogenesis. *Nat. Med.* 21, 628–637. doi: 10.1038/nm.3866

Shen, L., Gan, M., Li, Q., Wang, J., Li, X., Zhang, S., et al. (2018). MicroRNA-200b regulates preadipocyte proliferation and differentiation by targeting KLF4. *Biomed. Pharmacotherapy* 103, 1538–1544. doi: 10.1016/j.biopha.2018.04.170

Tontonoz, P., Hu, E., and Spiegelman, B. M. (1994). Stimulation of adipogenesis in fibroblasts by PPAR gamma 2, a lipid-activated transcription factor. *Cell* 79, 1147–1156. doi: 10.1016/0092-8674(94)90006-x

Tzameli, I., Fang, H., Ollero, M., Shi, H., Hamm, J. K., Kievit, P., et al. (2004). Regulated production of a peroxisome proliferator-activated receptor-g ligand during an early phase of adipocyte differentiation in 3T3-L1 adipocytes. *J. Biol. Chem.* 279, 36093–36102. doi: 10.1074/jbc.m405346200

Wang, C., Ha, X., Li, W., Xu, P., Gu, Y., Wang, T., et al. (2017). Correlation of TLR4 and KLF7 in inflammation induced by obesity. *Inflammation* 40, 42–51. doi: 10.1007/s10753-016-0450-z

White, U. A., and Stephens, J. M. (2010). Transcriptional factors that promote formation of white adipose tissue. *Mol. Cell. Endocrinol.* 318, 10–14. doi: 10.1016/j.mce.2009.08.023

Wu, Z., and Wang, S. (2013). Role of kruppel-like transcription factors in adipogenesis. *Dev. Biol.* 373, 235–243. doi: 10.1016/j.ydbio.2012.10.031

Xiong, Y., Xu, Q., Lin, S., Wang, Y., Lin, Y., and Zhu, J. (2018). Knockdown of LXRα inhibits goat intramuscular preadipocyte differentiation. *Int. J. Mol. Sci.* 19:E3037.

Yan, W., Zhang, H., Li, J., Shen, C., Xia, Y., Wang, P., et al. (2016). BMP4 promotes a phenotype change of an esophageal squamous epithelium via up-regulation of KLF4. *Exp. Mol. Pathol.* 101, 259–266. doi: 10.1016/j.yexmp.2016.09.007

Yoshida, T., Yamashita, M., Horimai, C., and Hayashi, M. (2014). Kruppel-like factor 4 protein regulates isoproterenol-induced cardiac hypertrophy by modulating myocardin expression and activity. *J. Biol. Chem.* 289, 26107–26118. doi: 10.1074/jbc.m114.582809

Zhang, Z. W., Rong, E. G., Shi, M. X., Wu, C. Y., Sun, B., Wang, Y. X., et al. (2014). Expression and functional analysis of Krüppel-like factor 2 in chicken adipose tissue. *J. Anim. Sci.* 92, 4797–4805.

**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: 06 October 2021
doi: 10.3389/fgene.2021.740036

# Comparative Analyses of Sperm DNA Methylomes Among Three Commercial Pig Breeds Reveal Vital Hypomethylated Regions Associated With Spermatogenesis and Embryonic Development

*Siqian Chen, Shuli Liu, Siyuan Mi, Wenlong Li, Shengli Zhang, Xiangdong Ding* and Ying Yu**

*Key Laboratory of Animal Genetics, Breeding and Reproduction, Ministry of Agriculture and National Engineering Laboratory for Animal Breeding, College of Animal Science and Technology, China Agricultural University, Beijing, China*

OPEN ACCESS

**Edited by:**
*Aline Silva Mello Cesar, University of SÃ£o Paulo, Brazil*

**Reviewed by:**
*Ran LI, Northwest A and F University, China*
*Guillermo Giovambattista, CONICET Institute of Veterinary Genetics (IGEVET), Argentina*

***Correspondence:**
*Xiangdong Ding*
*xding@cau.edu.cn*
*Ying Yu*
*yuying@cau.edu.cn*

**Specialty section:**
*This article was submitted to Livestock Genomics, a section of the journal Frontiers in Genetics*

**Received:** *12 July 2021*
**Accepted:** *08 September 2021*
**Published:** *06 October 2021*

**Citation:**
*Chen S, Liu S, Mi S, Li W, Zhang S, Ding X and Yu Y (2021) Comparative Analyses of Sperm DNA Methylomes Among Three Commercial Pig Breeds Reveal Vital Hypomethylated Regions Associated With Spermatogenesis and Embryonic Development. Front. Genet. 12:740036. doi: 10.3389/fgene.2021.740036*

Identifying epigenetic changes is essential for an in-depth understanding of phenotypic diversity and pigs as the human medical model for anatomizing complex diseases. Abnormal sperm DNA methylation can lead to male infertility, fetal development failure, and affect the phenotypic traits of offspring. However, the whole genome epigenome map in pig sperm is lacking to date. In this study, we profiled methylation levels of cytosine in three commercial pig breeds, Landrace, Duroc, and Large White using whole-genome bisulfite sequencing (WGBS). The results showed that the correlation of methylation levels between Landrace and Large White pigs was higher. We found that 1,040–1,666 breed-specific hypomethylated regions (HMRs) were associated with embryonic developmental and economically complex traits for each breed. By integrating reduced representation bisulfite sequencing (RRBS) public data of pig testis, 1743 conservated HMRs between sperm and testis were defined, which may play a role in spermatogenesis. In addition, we found that the DNA methylation patterns of human and pig sperm showed high similarity by integrating public data from WGBS and chromatin immunoprecipitation sequencing (ChIP-seq) in other mammals, such as human and mouse. We identified 2,733 conserved HMRs between human and pig involved in organ development and brain-related traits, such as *NLGN1* (neuroligin 1) containing a conserved-HMR between human and pig. Our results revealed the similarities and diversity of sperm methylation patterns among three commercial pig breeds and between human and pig. These findings are beneficial for elucidating the mechanism of male fertility, and the changes in commercial traits that undergo strong selection.

**Keywords: pig, sperm, DNA methylation, hypomethylated region, comparative epigenomics**

## INTRODUCTION

Pigs are an important source of fats, proteins, and human biomedical models (Swindle et al., 2012). After a long artificial and natural selection period, considerable phenotypic differences have emerged between pig and human in morphology, physiological structure, and behavior. Identifying epigenetic markers subject to evolution in different pig breeds will contribute to elucidating the epigenetic mechanism of important economic trait changes and supporting pigs as human biomedical models.

DNA methylation is the most stable and commonly studied epigenetic marker (Curradi et al., 2002), which participates in genome imprinting, silencing of the transposon element, and transcription inhibition (Barlow, 1993; Zamudio et al., 2015). Previous studies have shown that DNA methylation plays an important role in the growth, fertility, and health of pigs (Wang et al., 2017). In addition, by comparing epigenome-wide skeletal muscle DNA methylation profiles in distinct metabolic types of pig breeds, Ponsuksili et al. (2019) demonstrated that breed-specific methylated genes are linked to muscle metabolism and trigger extensive compensatory processes.

Sperm is an important heritable lineage. The comparison of sperm DNA methylomes in the whole genome across different pig breeds is unknown. Thus, to fill this gap, we conducted WGBS of sperm DNA samples from different pig breeds to explore the potential genetic mechanisms of phenotypic diversity and male fertility. Duroc, Landrace, and Large White are three common commercial pig breeds known for their excellent production performance under long-term artificial selection. Comparing breed-specific epigenomic markers in three commercial pig breeds helps us to understand how epigenetic regulation leads to phenotypic changes during evolution.

This study aimed to: 1. Investigate the sperm DNA methylomes using WGBS and analyze DNA methylation variation in three pig breeds—Landrace, Duroc, and Yorkshire; 2. Excavate breed-specific hypomethylated regions and identify genes that are enriched by lineage-specific hypomethylated regions around promoters; 3. Further identify the epigenetic biomarkers by integrating public reduced representation bisulfite sequencing (RRBS) data in pig testis; 4. Combine public data from WGBS and chromatin immunoprecipitation sequencing (ChIP-seq) in other mammals, the changes in sperm DNA methylation across mammalian evolution, and the epigenetic mechanism of pigs as human medical models (**Figure 1**).

## MATERIALS AND METHODS

### Sample Collection and Sequence Library Preparation

Six sperm samples were collected from three pig breeds in Landrace, Duroc, and Large White (3 breeds × 2 individuals). The pigs were raised using the same feed type on the same farm. Semen samples were collected by professional artificial

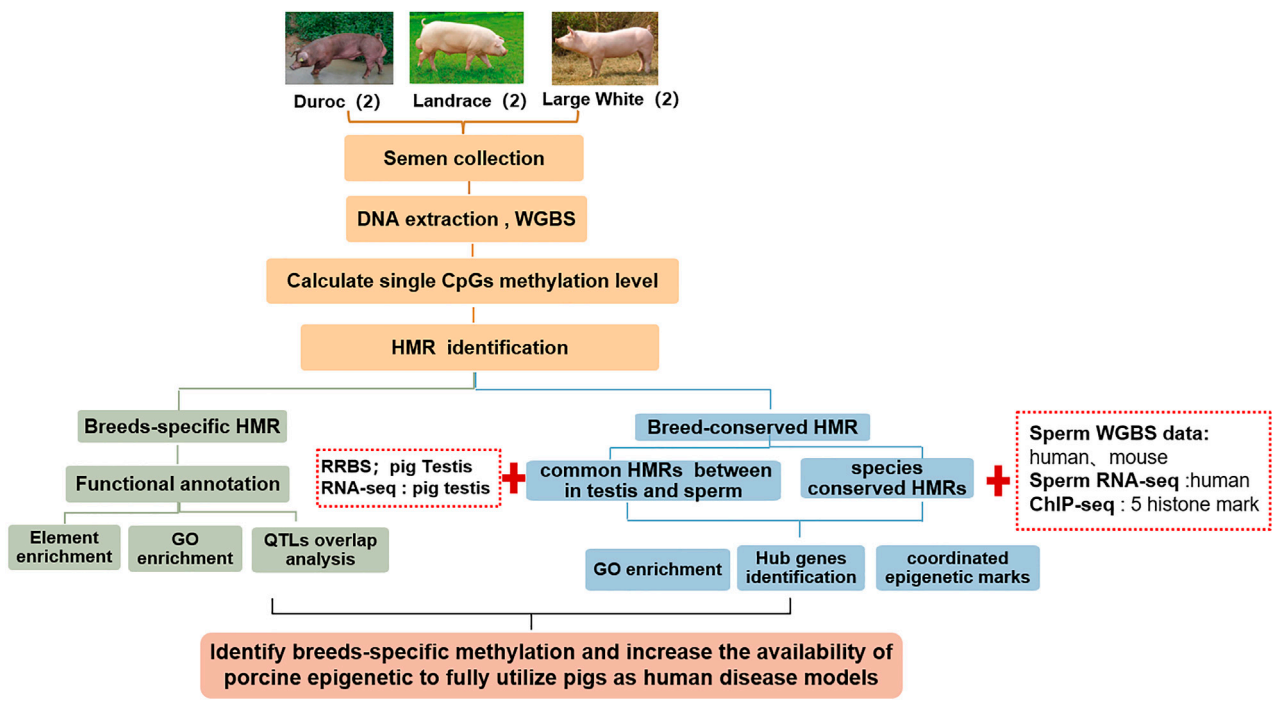


**FIGURE 1 |** Schematic overview of this study. We identified breed-specific/breeds-conserved HMRs of pigs using whole-genome bisulfite sequencing of six sperm samples from three pig breeds in Landrace, Duroc, and Yorkshire. Thereafter, we annotated them by integrating GO/KEGG, traditional quantitative trait loci (QTL) and transcriptional factor binding sites. Finally, we analyzed changes of HMRs in mammalian evolution by integrating WGBS public data on humans and mice.

insemination personnel according to a standardized procedure with artificial vaginas. Genomic DNA was extracted using a salt-fractionation protocol. DNA quality was assessed using a 2,100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, United States) and a spectrophotometer (NanoDrop Technologies, Rockland, DE). The qualified genomic DNA were spiked with unmethylated lambda DNA and fragmented into 200–300 bp, followed by terminal repair, the addition of 3′ A and adapter ligation. The DNA fragments were treated twice with bisulfite using an EZ DNA Methylation-Gold™ kit (Zymo Research, Irvine, CA, United States), under the manufacturer's instructions. Then the DNA fragments were amplified by PCR to screen the qualified library and sequenced using a paired-end 150 bp flow cell on an Illumina HiSeq X Ten machine (PE-150bp FC; Illumina, San Diego, CA, United States).

## WGBS Data Processing and Hypomethylated Region Identification

The public WGBS data used in this study included human sperm (GSE30340 and GSE57097) (Hammoud et al., 2014; Molaro et al., 2011) and mouse sperm (GSE49623) (Molaro et al., 2011). Three RRBS data of Landrace testis tissues from the NCBI GEO database (GSE129385) (Wang and Kadarmideen, 2019). FastQC v 0.11.2 (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/) and Trim Galore v 0.4.0 (https://www.bioinformatics. babraham. ac. uk/projects/trim _galore/) was used to assess the quality of the sequence data and filter low-quality reads ($q < 30$), respectively. The cleaned data were then mapped to the respective reference genomes: sscrofa11.1 (pig), hg38 (human), and mm10 (mouse), using bowtie2 under the Bismark software (0.14.5) with default parameters (Langmead et al., 2009; Krueger and Andrews, 2011). Furthermore, bismark_methylation_extractor was employed to obtain methyl CpG site information.

CpG sites with coverage greater than five were used for the HMR analysis. We used tools from the MethPipe package to identify HMRs, as described previously (Song et al., 2013). HMRs were inferred according to the methylation level and coverage at individual cytosines, using the HMM algorithm, which is trained using the Baum-Welch algorithm and posterior decoding (Rabiner, 1989). We only kept HMRs with an average regional methylation level of less than 20% and at least five CpG sites for further analysis.

## Identification of Conserved/Specific HMRs

The conservation and variability of HMR was analyzed using BEDTools (version 2.26.0) (Quinlan and Hall, 2010). This study required conserved HMRs overlapping more than 25% of the base pair unless otherwise stated. Specific and conserved HMRs were identified according to the number of overlapping HMRs across breeds, tissues, or species.

## RNA-Seq Data and Gene Expression Quantification

The transcriptomes of pig testes (PRJEB33381) and human sperm (PRJNA573604) were collected from a public database. Thereafter, the RNA-seq reads were mapped to the respective reference genomes of pigs (sscrofa11.1) and human (Hg38) using the HISAT2 software (Kim et al., 2015). Finally, we obtained FPKM using StringTie software to quantify gene expression levels for downstream analyses (Pertea et al., 2015).

## DNA Motif Enrichment

DNA motif enrichment analysis of breed-specific HMRs was conducted in three pig breeds. Moreover, we identified overrepresented DNA sequences using MEME suite 5.0.5 with the following parameters:' -dna -minw 8 -maxw 20 -mod zoops -objfun de -nmotifs 10' (Bailey et al., 2009). In addition, 2000 HMRs were randomly selected as the control region. We compared motifs against known motifs using the Tomtom of MEME suit software to identify enriched motifs. The databases of known motifs consist of jolma2013 motifs (Jolma et al., 2013), HOCOMOCOv11 full HUMAN mono meme format motifs (Kulakovskiy et al., 2018) and JASPAR2018 CORE vertebrate non-redundant motifs (Khan et al., 2018). Significant motifs were identified at $p < 0.05$.

## ChIP-Seq Data and Epigenetic Features Analysis

Histone data were downloaded from the SRA dataset, including five histone modifications: H3K27me3, H3K4me1, H3K4me3, H3K36me3, and H3K27ac (PRJNA173071 and PRJNA281061). First, we obtained clean reads using the Trim Galore software. The clean reads were mapped to hg38 (human) using the Bowtie2 software. Then, the Picard tool (http://broadinstitute.github.io/picard/) was used to remove duplicated reads with parameters 'REMOVE_DUPLICATES = true'. Histone peaks were obtained using the MACS2 software. The overlap between the HMR datasets and five histone modifications was identified using BEDTools, as previously described.

## Other Downstream Bioinformatics Analysis

The proportion of breed-specific HMRs that fell completely within the QTLs were calculated. The QTLs were downloaded from the Pig QTL database (https://www.animalgenome.org/cgi-bin/QTLdb/SS/index). Genes with promoter HMRs were annotated using the online software DAVID (http://metascape.org/). Significant GO terms and pathways were identified based on $p < 0.05$. GO terms were visualized using the R package GOplot (version 1.0.2) (Walter et al., 2015). Thereafter, the protein-protein interaction network (PPI) encoded by important genes was investigated using the STRING Genomics database (https://www.string-db.org/) and identified hub genes with a higher degree of connectivity using CytoHubba under the Cytoscape software (Shannon et al., 2003).

# RESULTS

## General Characteristics of Sperm DNA Methylome

In this study, we conducted WGBS of individual sperm DNA samples from three commercial pig breeds—Landrace, Duroc,

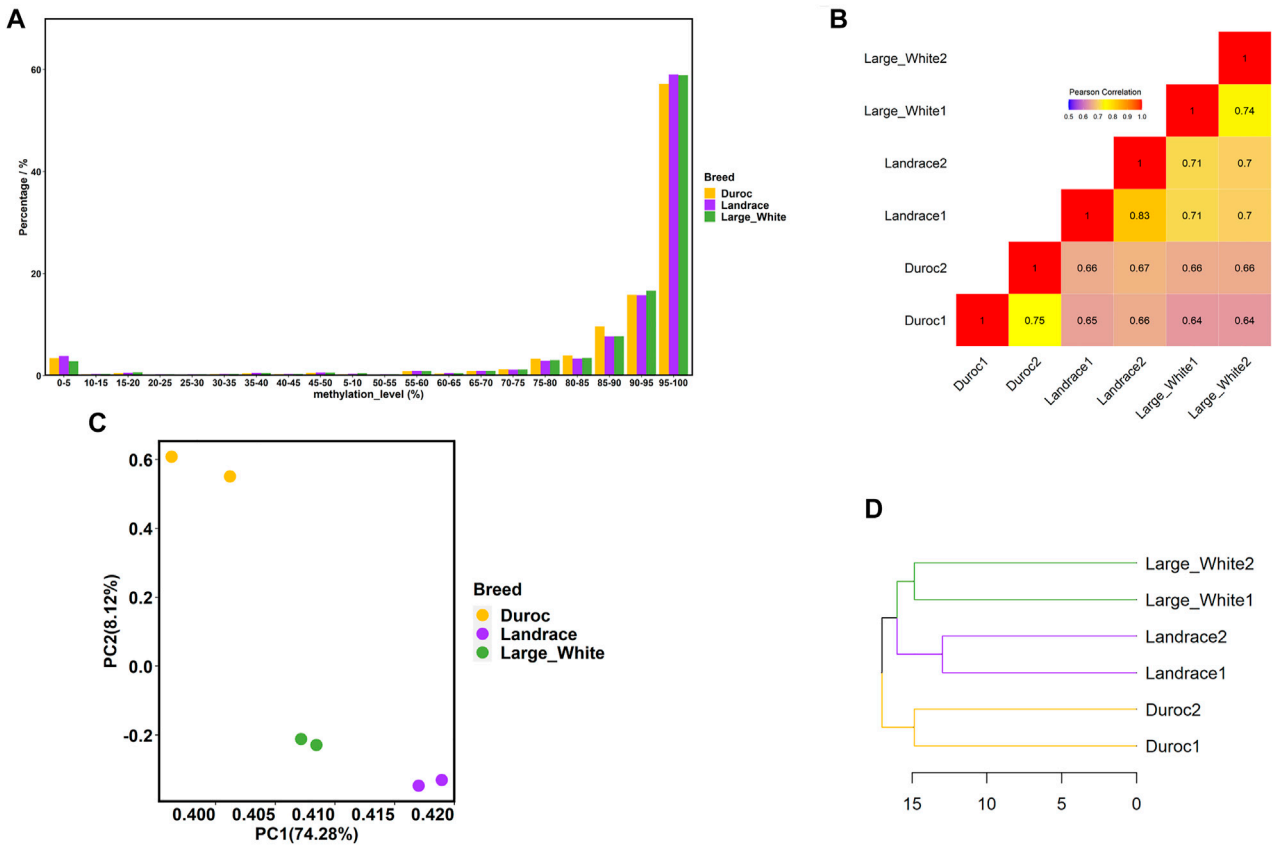


**FIGURE 2 |** Porcine sperm DNA methylome characteristics. **(A)** Distribution of methylation levels of CpG sites (only CpG sites covered by at least five reads were included). **(B)** Correlation analysis between each sample using common CpG sites. **(C)** PCA of methylation levels in six samples using common CpG sites. **(D)** Hierarchical clustering of three pig breeds based on the methylation level of promoter regions.

and Large White. We obtained 204 to 307 million unique mapped reads with an average coverage from 12.21 to 18.43× (**Supplementary Table S1**). An overall methylation level of 86.20–87.80% was observed for all CpG sites in the three pig breeds, and non-CpG methylation levels (CHG and CHH) of Landrace (5.90%) were higher than those of Duroc (1.40%) and Large White (1.55%). Moreover, the bisulfite conversion rates of all samples were greater than 99%; therefore, we faithfully captured the patterns of porcine sperm genomic DNA methylation.

In general, we observed that approximately 86% of the analyzed CpG sites were extremely hypermethylated (methylation level ≥80%) for all samples (**Figure 2A**), which was consistent with the pattern of sperm methylation in other mammals (Molaro et al., 2011; Song et al., 2013; Fang et al., 2019; Liu et al., 2019). Then, we compared the global methylation levels between pairs of samples at the common CpG sites within and among breeds, with a minimum coverage of five sequencing reads. As expected, DNA methylation varied more across breeds than within breeds (**Figure 2B**). The correlations were higher within breeds, ranging from 0.74 to 0.83. The correlations of DNA methylation among the three pig breeds were lower, with the correlations between Duroc and Large White being the lowest ($r$ = 0.64). Principal component analysis (PCA) also demonstrated this phenomenon (**Figure 2C**). PC1 successfully divided samples into three clusters (Landrace, Duroc, and Large White) according to the variations among the three breeds, which explained 74.28% of the variance.

To study the evolution of DNA methylation among the three porcine breeds, we further investigated the promoter methylation levels of different breeds. A promoter region was defined as the segment 1.5 kb upstream and 0.5 kb downstream of transcription start sites (TSSs) (de Vooght et al., 2009; Shen et al., 2012; Zhou et al., 2021). The hierarchy in the sperm DNA methylation cluster (**Figure 2D**) was consistent with the known genetic relationship between the three porcine breeds (Cai et al., 2020). **Figure 2D** shows that, compared with Duroc, the similarity between Large White and Landrace was higher.

## Inter-Breed Variation and Conservation in Sperm DNA Methylome

Hypomethylated regions (HMRs) are associated with gene activation and coincided with gene regulatory elements, including gene promoters and enhancers (Jones, 2012; Wagner et al., 2014); therefore, we identified HMRs (methylation level ≤20%) using

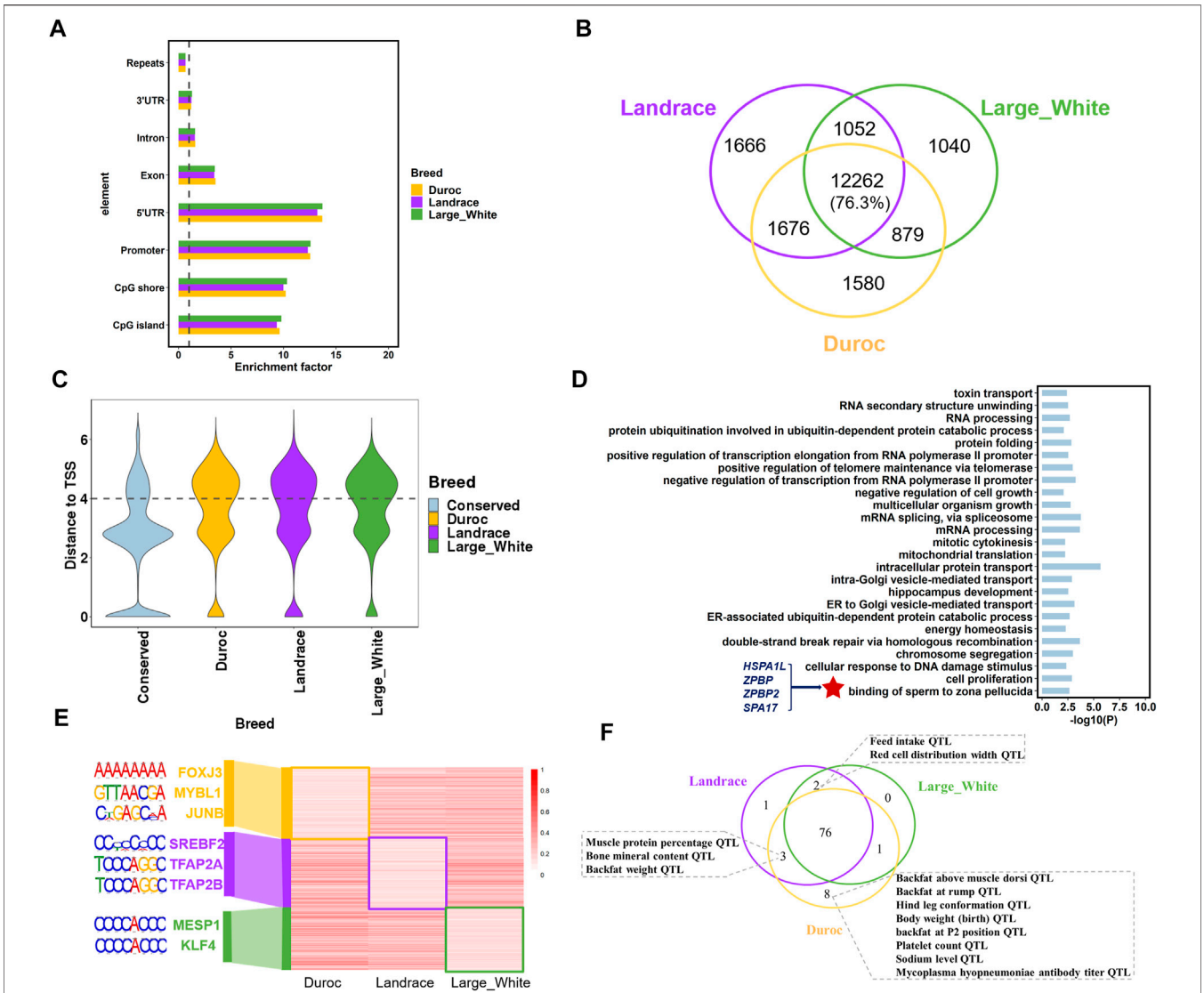


**FIGURE 3 |** Porcine sperm hypomethylated regions (HMRs) characteristics. **(A)** The enrichment of HMRs across genomic elements. **(B)** The number of conserved HMRs of the three porcine breeds. **(C)** The distribution of distances from the different types of HMRs to closet TSS. **(D)** Go term ($p < 0.05$) analysis of genes associated with breeds-conserved hypomethylated (<20%) promoters (hypo-genes) in pigs. **(E)** The methylation level and the significant DNA motifs of breed-specific HMRs. **(F)** The overlapped pig QTLs and breed-specific HMRs in the three porcine breeds.

the 2-state hidden Markov model with beta-binomial emission distributions as previously described (Song et al., 2013). The algorithm detected 15,233 to 16,656 HMRs with an average size range of 1,109–1,146 bp. Although HMRs were distributed throughout the genome, they were significantly enriched in the 5' UTR, promoters, CpG islands, and CpG shores ($p < 0.001$; 1,000 times of permutation test) (**Figure 3A**). The distribution of HMRs in these genomic regions further indicated that HMRs play an important role in regulating transcriptional initiation and gene expression.

To analyze the difference in DNA methylation across pig breeds, we defined two categories of regions from HMRs: 1) breeds-conserved HMRs, containing a region that is hypomethylated in all breeds, and 2) breed-specific HMRs, which are regions that are hypomethylated in one breed compared to the other two breeds. As expected, most HMRs were conserved in the three porcine breeds, with an average of 76.3% in each breed (**Figure 3B**). Moreover, the genomic distribution of breeds-conserved HMRs and breed-specific HMRs was obviously different. A total of 49.62% of breed-specific HMRs were located in the distal regions (more than 10 kb away from TSS), contrary to 22.08% of breeds-conserved HMRs (**Figure 3C**). These results were consistent with high conservation at promoters, as reported in previous studies (Qu et al., 2018; Villar et al., 2015). In addition, this result also indicates that we need to pay attention to the HMRs located at the distal genome in future studies of breeds divergence. These breeds-specific regions may be located at distal regulatory

element regions, thus affecting gene expression differences in breeds. Functional enrichment analysis revealed that the genes ($n$ = 1800) with breeds-conserved HMRs in the promoters were engaged in basic cell function and fertilization, including energy homeostasis, cell proliferation, multicellular organism growth, and binding of sperm to the zona pellucida (**Figure 3D**, **Supplementary Table S2**). In the binding of sperm to the zona pellucida GO term, we found several genes associated with spermatogenesis and fertilization, namely, *HSPA1L* (heat shock protein family A member 1 like), *SPA17* (sperm autoantigenic protein 17), *ZPBP* (zona pellucida binding protein), and *ZPBP2* (zona pellucida binding protein *2*). *HSPA1L* has also been shown to be related to spermiogenesis in cattle, human, and mouse (Liu et al., 2019; Wang et al., 2020). This implies that DNA methylation patterns of key genes associated with spermatogenesis are highly conserved across species (**Figure 3D**, **Supplementary Table S2**).

We identified 1,580 Duroc-specific, 1,666 Landrace-specific, and 1,040 Large White-specific HMRs (**Figure 3B**). The genes with breed-specific hypomethylation in the gene promoter revealed a strong enrichment for transcription and development (**Supplementary Table S3**), such as post-embryonic development, kidney development, and regulation of transcription from RNA polymerase II promoter. Genes were also enriched in GO terms related to complex traits, such as high-density lipoprotein particle remodeling and prostate gland growth. Moreover, we discovered significant DNA motifs in breed-specific HMRs (**Figure 3E**, **Supplementary Table S4**). Landrace and large white HMRs were enriched for embryonic development and cholesterol synthesis. In the Duroc-specific HMRs, we identified significant motifs associated with skeletal muscle regeneration and male meiosis, such as *FOXJ3* and *MYBL1*.

To further investigate the relationship between breed-specific HMRs and complex traits, we examined the QTL regions of five categories (meat and carcass, health, exterior, production, and reproduction) of traits ($n$ = 691) from the Pig QTL database (https://www.animalgenome.org/cgi-bin/QTLdb/SS/index). We observed that Duroc-specific HMR had higher overlapping QTL signals of production traits, meat, carcass traits such as body weight (birth) QTL, and backfat above muscle dorsi QTL. In addition, Duroc and Landrace had specific HMRs overlapping with muscle protein percentage QTL and backfat weight QTL (**Figure 3F**). These results were consistent with their varietal characteristics.

## Comparison of HMRs Between Porcine Sperm and Testis

Testis is the organ which produces sperm, the male reproductive cell, and androgens, the male hormones (Schumacher, 2012). Moreover, spermatogenic activity co-varies with testis mass and sperm quality across (Pintus et al., 2015). To better understand the epigenetic mechanism that regulates male fertility and semen quality, we investigated the conservation and divergence of breeds-conserved HMRs in sperm and HMRs in Landrace testis. We found that 14.23% of HMRs ($n$ = 1743) in sperm were common to the testis (**Figure 4A**). This result may imply that HMRs have higher tissue specificity. Functional annotation revealed that for genes with conserved HMRs at the promoters between sperm and testis, GO terms of mRNA splicing, cell proliferation, regulation of double-strand break repair via homologous recombination, and regulation of cell cycle during spermatogenesis (**Supplementary Table S5**).

To further investigate the important genes associated with the spermatogenesis process, we conducted PPI analyses and identified hub genes from genes with conserved HMRs at the promoter between sperm and testis. We found that among 1743 conserved HMRs, 30 genes with a higher degree of connectivity were hub genes (lower panel, **Figure 4A**). Several genes with hypomethylated promoter regions were highly expressed in the testis (**Figure 4B**). In particular, *EXOSC10* (exosome component 10) and *HSPA9* (heat shock protein family member 9) have been suggested to play important roles in male germ cell development (Jamin et al., 2017; Jeong et al., 2017).

## The Conservation of Breeds-Conserved HMRs Across Species

HMRs around promoters can reflect the evolution of the mammalian epigenome (Qu et al., 2018); therefore, we compared HMRs in the sperms of three species, *i.e.,* pig, human and mouse, to explore the evolution of germline DNA methylation and epigenetic mechanisms underlying remarkable changes. Consequently, our results suggested that the proportion of conserved HMRs in three species only constituted 31.66–32.96% (**Figure 5A**), implying that the epigenome exhibits evolutionary changes and strong lineage-specific aspects across millions of years of evolution in the three mammals. Functional enrichment analysis showed that genes associated with three species-conserved HMRs in the promoters were enriched in transcription processing and embryonic development, such as transcription from RNA polymerase II promoter, nervous system development, and embryonic digit morphogenesis (**Figure 5B**, **Supplementary Table S6**). This suggests that genes involved in basic physiological functions and development processes tend to have conserved DNA methylation levels during mammalian evolution.

Pairwise comparison between human with pig and mouse revealed that an average of 61.14% ($n$ = 8,553) of HMRs were conserved between humans and the other two species (**Figure 5A**). Interestingly, the percentage of conserved HMRs between pig and human (56.24%) was higher than that between mouse and human (55.61%); therefore, the use of pigs as an animal model for human diseases research is reasonable.

To investigate the similarities between human and pig sperm DNA methylation patterns and identify stable epigenetic markers, we compared the conservation of HMRs across the three pig breeds and humans. A total of 32.08% ($n$ = 2,733) of the conserved HMRs across the three pig breeds were conserved in human (**Figure 5C**). Then, we obtained hub genes (top 40) from genes with conserved HMRs between pig and human at the promoters using Cytoscape (**Figure 5C**). We identified several important genes with hypomethylated regions at promoters

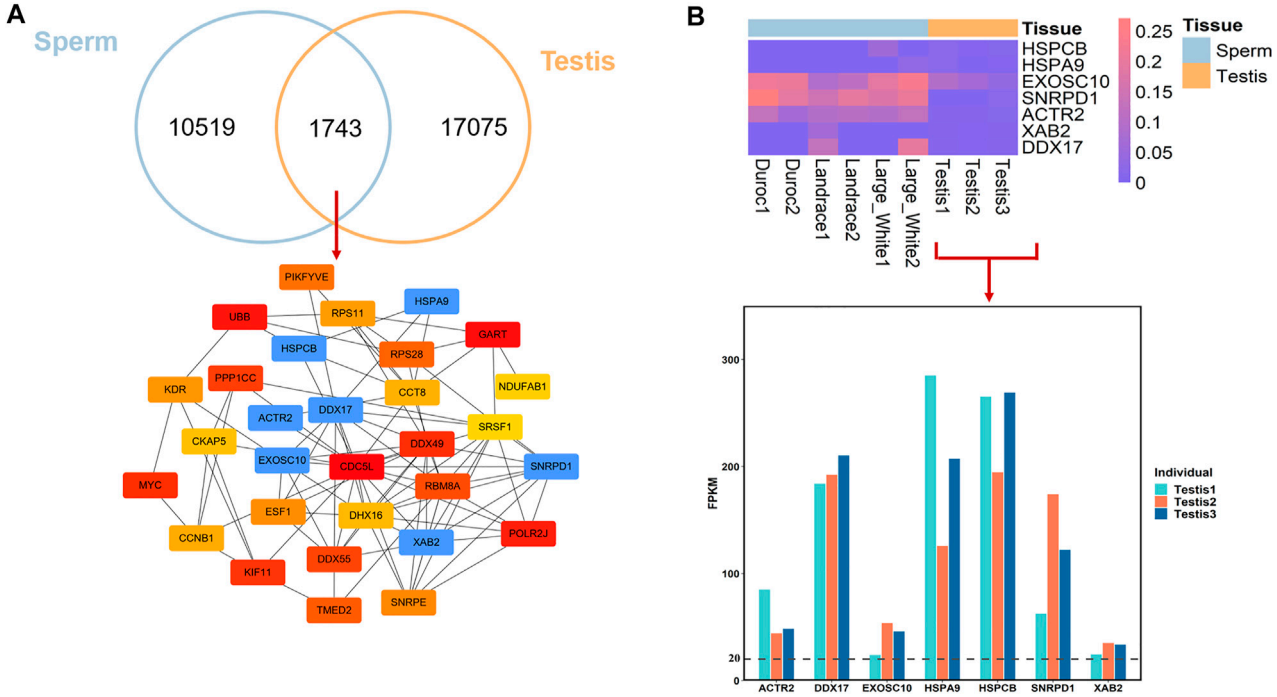


**FIGURE 4 |** DNA methylation conservation and divergences between sperm and testis in pig. **(A)** Overlap of HMRs and hub genes identification. **(B)** DNA methylation level in promoter regions between sperm and testis and gene expression of hub genes in testis.

highly expressed in human sperm (**Figure 5D**). These findings further suggest that these genes are related to mitosis progression and regulation of neurotransmitter release.

## Coordinated Epigenetic Marks at Conserved HMRs Between Human and Pig

In addition to DNA methylation, chromatin states also play an important role in spermiogenesis, fertilization, and development. Then, we analyzed the chromatin states of the conserved HMRs between human and pigs using five histone marks (H3K27me3, H3K4me1, H3K4me3, H3K36me3, and H3K27ac, downloaded from the SAR database). We found that 75.01% ($n$ = 2050) of conserved HMRs between human and pigs overlapped with H3K4me3 marks, which is in accordance with the antagonistic relationship between DNA methylation and H3K4me3. In contrast, less conserved HMRs were enriched for H3K4me1 ($n$ = 90) and H3K36me3 ($n$ = 17) (**Figure 6A**). In addition, no conserved HMRs were enriched with H3K27ac ($n$ = 0).

In addition, we found that some conserved HMRs (24.48%, $n$ = 669) were simultaneously marked by H3K4me3 and H3K27me3 (**Figure 6B**), which are termed bivalent domains. Mounting evidence suggests that genes with bivalent domains are found in embryonic stem cells and germ cells and play an important role in initiating sexual differentiation, embryonic development, and completion of meiosis (Sin et al., 2015; Vastenhouw and Schier, 2012). We further found that genes containing HMRs and bivalent domains at promoters were significantly enriched in development and behavior terms, such as nervous system development and locomotory behavior (**Figure 6C**, **Supplementary Table S7**), as previously described (Vastenhouw and Schier, 2012). It implies that genes associated with organ specification are poised in germ cells, and that DNA methylation levels are highly conserved between species. Among them, *NLGN1* plays an important role in synaptic signal transmission and in neuropsychiatric disorders (Nakanishi et al., 2017). *NLGN1* was further found to have a conserved-HMR between human and pig, and contain a bivalent domain at the promoter in human sperm (**Figure 6D**). This result further suggested that *NLGN1* was regulated by the conserved epigenomic markers and may play an important role in brain-related traits between human and pig. Overall, These findings further support pig as an ideal animal model for human complex traits, particularly for brain-related diseases.

## DISCUSSION

This study analyzed DNA methylation patterns in three commercial pig breeds (Durocs-Duroc, Landrace, and Large White) at a high resolution and investigated the relationship between pig and human with DNA methylation and histone modifications. In addition, the evolutionary properties of HMR in three commercial pig breeds, which may play an important role in embryonic developmental traits and adaptive traits, were analyzed. Our results also showed that most genes with conserved HMRs at promoters between human and pig are in the poised chromatin state (H3K4me3 and H3K27me3 bivalent), which are involved in brain-related traits.

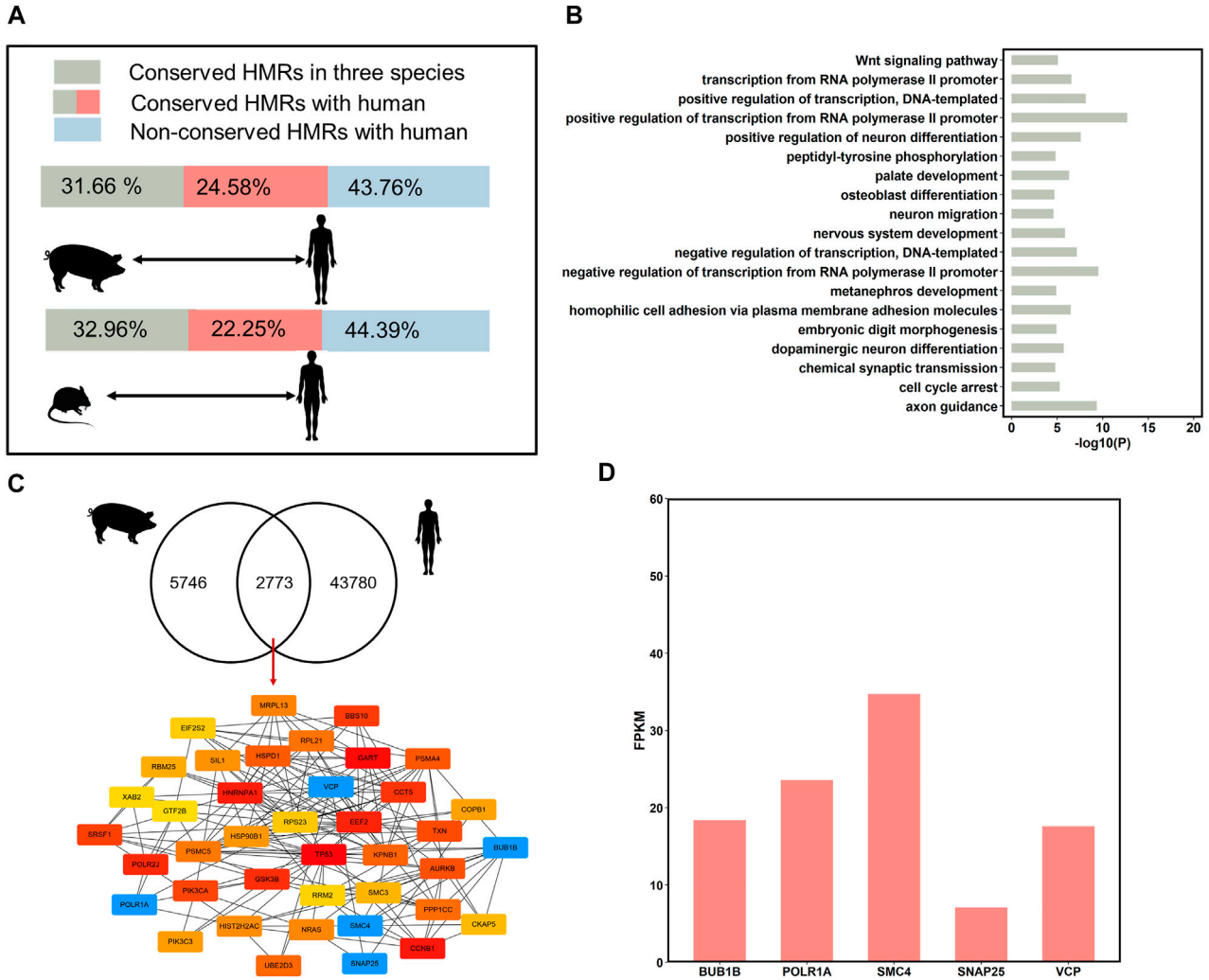


**FIGURE 5 |** DNA methylation conservation and divergences of species. **(A)** Comparison HMRs in human to pig, and mouse. The green bars correspond to the conservative HMR percentage of three species in the pig and mouse. The green bars plus red bars correspond to the percentage of conserved HMRs in pig and mouse, compared to the human. The blue bars correspond to non-conservative HMR percentage in the pig and mouse, compared to the human. **(B)** Gene Ontology analysis of genes associated with conserved hypomethylated (<20%) promoters in three species (top 20, $p < 0.05$). **(C)** The hub genes with conserved HMRs between humans and the three pigs at the promoters (top 40). **(D)** The expression level of the important hub genes in human sperm.

Based on the comparison of sperm DNA methylome variation across three commercial pigs, our results showed that the divergence of sperm DNA methylomes fully recapitulates phylogenetic relationships, as previously reported for DNA sequences (Cai et al., 2020). Principal component analyses and hierarchical clustering demonstrated that breeds explain the largest amount of variation in sperm DNA methylomes, which PC1 could explained most (74.28%) of the variances. Taken together, these results suggested the importance of DNA methylation during breeds divergencies. Interestingly, we also noted that there were some differences in DNA methylation levels within breeds. Through comparing the intra-breed difference and conservatism of HMRs in three commercial pigs, we found that only the small proportion of HMRs are variable in intra-breed, which constituted 10.96–14.43% (**Supplementary Figure S1**). The methylation variations within breeds need to be analyzed through expanding the sample size for subsequent research.

In primate, the majority of species-specific DMRs locate outside promoters and have similar transcriptional potential as promoter DMRs (Mendizabal et al., 2016). In this study, we also found that many of the breed-specific HMRs were located in distal regions of genes. This was consistent with more rapid evolution of enhancers and slower changes at promoters in mammals (Villar et al., 2015). Moreover, we also observed that the important QTLs for carcass and production traits were obviously overlapped with breed-specific HMRs. For example, we found that 9 Duroc-specific HMRs were located with an important muscle protein percentage QTL for carcass and meat traits on chromosome 15 (BTA15, roughly located between 127.9 and 135.9 Mb) by comparing crossing porcine from the Pietrain and Duroc breeds (Choi et al., 2011).

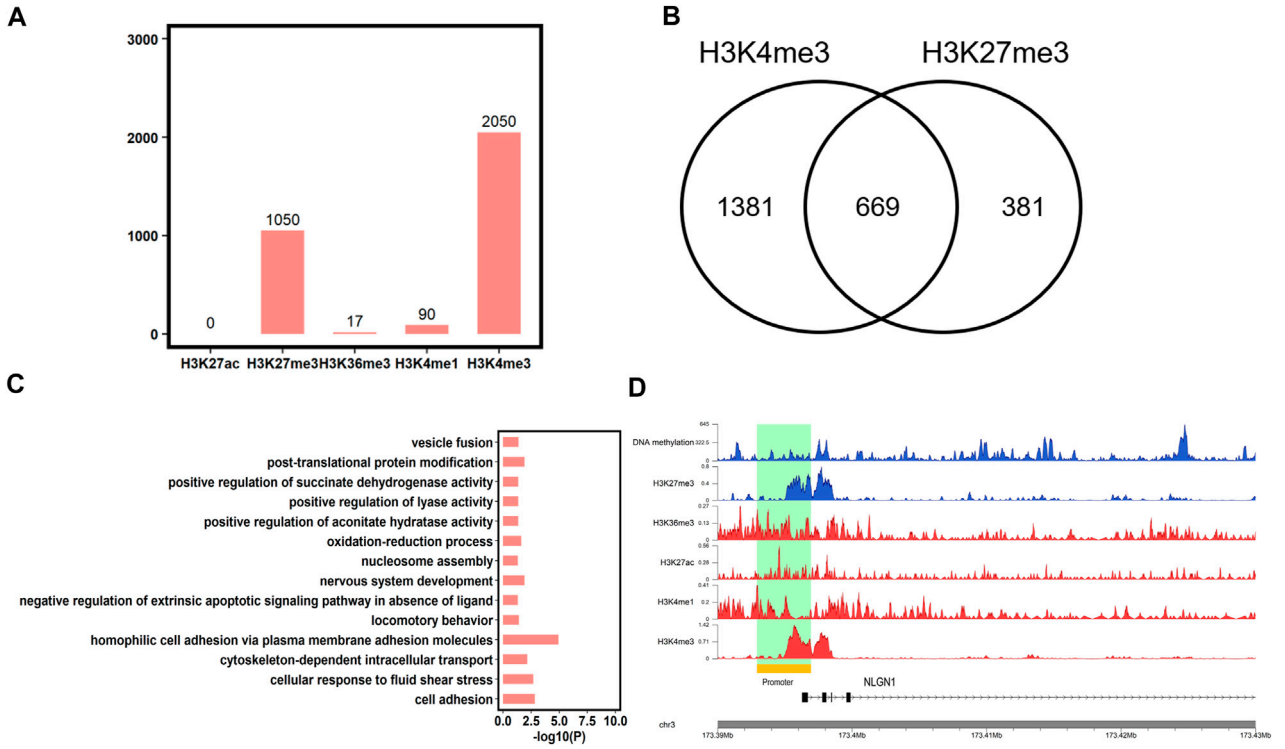


**FIGURE 6 |** Epigenetic marks at conserved HMRs between human and pig. **(A)** The number of conserved HMRs between human and pig showed overlap with H3K4me1, H3K4me3, H3K36me3, H3K27ac, and H3K27me3. **(B)** The number of conserved HMRs simultaneously marked by H3K4me3 and H3K27me3. **(C)** Gene Ontology analysis of genes associated with conserved HMRs and bivalent domains at the promoters ($p < 0.05$). **(D)** Gene tracks showing DNA methylation, H3K4me1, H3K4me3, H3K36me3, H3K27ac, and H3K27me3 signal in human sperm at one representative poised gene, *NLGN1*.

Six Duroc-specific HMRs were also located at the body weight (birth) QTL for production traits on chromosome 15 (BTA12, roughly located between 54.4 and 58.5 Mb) (Rothammer et al., 2014). These results imply that complex traits may be regulated by genetics and epigenetics, and breed-specific HMRs may be the candidate epigenetic markers that influence production and reproductive traits.

Through pig breeds-conserved HMRs in sperms compared with testis, we observed several conserved HMRs between pig sperm and testis close to or located in the promoter regions of important genes involved in spermatogenesis, such as *EXOSC10* and *HSPA9*. *EXOSC10* is a well-known target of autoantibodies in patients with systemic sclerosis (scleroderma) (Staals and Pruijn, 2010). Recent studies have also reported that *EXOSC10* is essential for normal growth-to-maturation transition in mouse oocytes and male germ cell proliferation and development (Jamin et al., 2017; Wu and Dean, 2020). This study also found that *EXOSC10* exhibited hypomethylated levels at the promoters of sperm and testis and high expression in the testis. Consistently, a previous study reported that the loss of *EXOSC10* in spermatogonia could lead to abnormal testicular development and a strongly decreased size (Jamin et al., 2017). *HSPA9* has also shown low methylation levels at the promoter and is expressed at a high level in the testis, which plays an important role in prophase I of spermatogenesis by binding to testis-specific *MAGEG2* (Melanoma Antigen Family G2) (Jeong et al., 2017). These results show that these conserved HMRs between the testis and sperm could be important candidate regions for the study of spermatogenesis and male infertility.

By analyzing the divergence of sperm DNA methylomes in three mammalian species: human, mouse and pig, we found that DNA methylomes were more highly conserved in pig and human than in human and mouse. Moreover, conserved HMRs between human and pig are also associated with GO terms related to development functions, particularly brain-related traits. This may be explained by brain-associated genes under strong selective constraints during the evolution of species (Cardoso-Moreira et al., 2019; Sarropoulos et al., 2019). For instance, SNAP25 (synaptosome-associated protein 25) is the core component of the soluble N-ethylmaleimide fusion protein attachment protein receptor (SNARE) and is associated with brain-related diseases, such as autism and schizophrenia (Braida et al., 2015; Ramos-Miguel et al., 2019; Washbourne et al., 2002). Moreover, we found that the conserved HMRs between human and pig were highly co-located with bivalent domains and were significantly enriched for GO terms of nervous system development and locomotor behavior. This was consistent with previous studies showing that developmental genes remain poised for later activation and are simultaneously marked by H3K4me3 and H3K27me3 in the male germline, such as the homeobox family of genes and the SRY-related HMG-box family of genes (Sin et al., 2015). In summary, these observations indicate that brain-associated genes may have highly conserved DNA methylation patterns across species evolution and are poised for activation at specific developmental stages.

## CONCLUSION

This study is the first to report a genome-wide comparative DNA methylation map of adult pig sperm in three commercial pig breeds using WGBS technology. Our results showed that sperm HMRs were highly conserved in the three commercial pig breeds. Moreover, our results indicated that breed-specific HMRs are related to phenotypic changes and economically complex traits for each breed. The conserved HMRs between pig sperm and testis close to or located in the promoter regions of important genes are mainly involved in DNA repair and spermatogenesis. Additionally, we found considerable similarities in DNA methylomes between pig and human sperms. The conserved HMRs between human and pig are related to brain-associated genes. In summary, our study of sperm methylomes in three commercial pig breeds contributes to understanding the mechanism of complex traits (particularly male fertility and brain-related traits) undergoing selection, further supporting pigs as human medical models from an epigenetic standpoint.

## DATA AVAILABILITY STATEMENT

The original contributions presented in the study are publicly available. This data can be found here: https://www.ncbi.nlm.nih.gov/geo/, GSE180099.

## ETHICS STATEMENT

The animal study was reviewed and approved by Animal handling and sample collection were conducted according to protocols approved by the Institutional Animal Care and Use Committee (IACUC) at China Agricultural University.

## AUTHOR CONTRIBUTIONS

SC analyzed the data and wrote manuscript. SL participated in samples collection and WGBS data analyses. SM and ML performed RRBS data analyses. YY and SZ participated in the result interpretation and paper revision. YY and XD conceived and designed the study and revised the manuscript. Both authors read and approved the final manuscript.

## FUNDING

This research was financially supported by the National Natural Science Foundation of China (32072718), the National Key Research and Development Project (2019YFE0106800) and China Agriculture Research System of MOF and MARA (CARS-35).

## ACKNOWLEDGMENTS

We thank all co-authors for their contributions and Qinglei Xu for assistance in the collection and processing of samples.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2021.740036/full#supplementary-material

## REFERENCES

Bailey, T. L., Boden, M., Buske, F. A., Frith, M., Grant, C. E., Clementi, L., et al. (2009). MEME SUITE: Tools for Motif Discovery and Searching. *Nucleic Acids Res.* 37, W202–W208. doi:10.1093/nar/gkp335

Barlow, D. (1993). Methylation and Imprinting: from Host Defense to Gene Regulation?. *Science* 260, 309–310. doi:10.1126/science.8469984

Braida, D., Guerini, F. R., Ponzoni, L., Corradini, I., De Astis, S., Pattini, L., et al. (2015). Association Between SNAP-25 Gene Polymorphisms and Cognition in Autism: Functional Consequences and Potential Therapeutic Strategies. *Transl. Psychiatry* 5 (1), e500. doi:10.1038/tp.2014.136

Cai, Z., Sarup, P., Ostersen, T., Nielsen, B., Fredholm, M., Karlskov-Mortensen, P., et al. (2020). Genomic Diversity Revealed by Whole-Genome Sequencing in Three Danish Commercial Pig Breeds. *J. Anim. Sci.* 98. doi:10.1093/jas/skaa229

Cardoso-Moreira, M., Halbert, J., Valloton, D., Velten, B., Chen, C., Shao, Y., et al. (2019). Gene Expression across Mammalian Organ Development. *Nature* 571, 505–509. doi:10.1038/s41586-019-1338-5

Cho, C., Jeong, J., Jin, S., Choi, H., Kwon, J., Kim, J., et al. (2017). Characterization of MAGEG2 with Testis-specific Expression in Mice. *Asian J. Androl.* 19, 659–665. doi:10.4103/1008-682X.192033

Choi, I., Steibel, J. P., Bates, R. O., Raney, N. E., Rumph, J. M., and Ernst, C. W. (2011). Identification of Carcass and Meat Quality QTL in an F2 Duroc × Pietrain Pig Resource Population Using Different Least-Squares Analysis Models. *Front. Gene.* 2, 18. doi:10.3389/fgene.2011.00018

Curradi, M., Izzo, A., Badaracco, G., and Landsberger, N. (2002). Molecular Mechanisms of Gene Silencing Mediated by DNA Methylation. *Mol. Cel. Biol.* 22, 3157–3173. doi:10.1128/MCB.22.9.3157-3173.2002

de Vooght, K. M. K., van Wijk, R., and van Solinge, W. W. (2009). Management of Gene Promoter Mutations in Molecular Diagnostics. *Clin. Chem.* 55 (4), 698–708. doi:10.1373/clinchem.2008.120931

Fang, L., Zhou, Y., Liu, S., Jiang, J., Bickhart, D. M., Null, D. J., et al. (2019). Comparative Analyses of Sperm DNA Methylomes Among Human, Mouse and Cattle Provide Insights into Epigenomic Evolution and Complex Traits. *Epigenetics* 14, 260–276. doi:10.1080/15592294.2019.1582217

Hammoud, S. S., Low, D. H. P., Yi, C., Carrell, D. T., Guccione, E., and Cairns, B. R. (2014). Chromatin and Transcription Transitions of Mammalian Adult Germline Stem Cells and Spermatogenesis. *Cell Stem Cell* 15, 239–253. doi:10.1016/j.stem.2014.04.006

Jamin, S. P., Petit, F. G., Kervarrec, C., Smagulova, F., Illner, D., Scherthan, H., et al. (2017). EXOSC10/Rrp6 Is post-translationally Regulated in Male Germ Cells and Controls the Onset of Spermatogenesis. *Sci. Rep.* 7, 15065. doi:10.1038/s41598-017-14643-y

Jolma, A., Yan, J., Whitington, T., Toivonen, J., Nitta, K. R., Rastas, P., et al. (2013). DNA-binding Specificities of Human Transcription Factors. *Cell* 152, 327–339. doi:10.1016/j.cell.2012.12.009

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

Khan, A., Fornes, O., Stigliani, A., Gheorghe, M., Castro-Mondragon, J. A., van der Lee, R., et al. (2018). JASPAR 2018: Update of the Open-Access

Database of Transcription Factor Binding Profiles and its Web Framework. *Nucleic Acids Res.* 46, D260–D266. doi:10.1093/nar/gkx1126

Kim, D., Langmead, B., and Salzberg, S. L. (2015). HISAT: a Fast Spliced Aligner with Low Memory Requirements. *Nat. Methods* 12, 357–360. doi:10.1038/nmeth.3317

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

Kulakovskiy, I. V., Vorontsov, I. E., Yevshin, I. S., Sharipov, R. N., Fedorova, A. D., Rumynskiy, E. I., et al. (2018). HOCOMOCO: towards a Complete Collection of Transcription Factor Binding Models for Human and Mouse via Large-Scale ChIP-Seq Analysis. *Nucleic Acids Res.* 46, D252–D259. doi:10.1093/nar/gkx1106

Langmead, B., Trapnell, C., Pop, M., and Salzberg, S. L. (2009). Ultrafast and Memory-Efficient Alignment of Short DNA Sequences to the Human Genome. *Genome Biol.* 10, R25. doi:10.1186/gb-2009-10-3-r25

Liu, S., Chen, S., Cai, W., Yin, H., Liu, A., Li, Y., et al. (2019). Divergence Analyses of Sperm DNA Methylomes between Monozygotic Twin AI Bulls. *Epigenomes* 3, 21. doi:10.3390/epigenomes3040021

Mendizabal, I., Shi, L., Keller, T. E., Konopka, G., Preuss, T. M., Hsieh, T.-F., et al. (2016). Comparative Methylome Analyses Identify Epigenetic Regulatory Loci of Human Brain Evolution. *Mol. Biol. Evol.* 33 (11), 2947–2959. doi:10.1093/molbev/msw176

Molaro, A., Hodges, E., Fang, F., Song, Q., McCombie, W. R., Hannon, G. J., et al. (2011). Sperm Methylation Profiles Reveal Features of Epigenetic Inheritance and Evolution in Primates. *Cell* 146, 1029–1041. doi:10.1016/j.cell.2011.08.016

Nakanishi, M., Nomura, J., Ji, X., Tamada, K., Arai, T., Takahashi, E., et al. (2017). Functional Significance of Rare Neuroligin 1 Variants Found in Autism. *PLOS Genet.* 13, e1006940. doi:10.1371/journal.pgen.1006940

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, 290–295. doi:10.1038/nbt.3122

Pintus, E., Ros-Santaella, J. L., and Garde, J. J. (2015). Beyond Testis Size: Links between Spermatogenesis and Sperm Traits in a Seasonal Breeding Mammal. *PLoS One* 10 (10), e0139240. doi:10.1371/journal.pone.0139240

Ponsuksili, S., Trakooljul, N., Basavaraj, S., Hadlich, F., Murani, E., and Wimmers, K. (2019). Epigenome-wide Skeletal Muscle DNA Methylation Profiles at the Background of Distinct Metabolic Types and Ryanodine Receptor Variation in Pigs. *BMC Genom.* 20, 492. doi:10.1186/s12864-019-5880-1

Qu, J., Hodges, E., Molaro, A., Gagneux, P., Dean, M. D., Hannon, G. J., et al. (2018). Evolutionary Expansion of DNA Hypomethylation in the Mammalian Germline Genome. *Genome Res.* 28, 145–158. doi:10.1101/gr.225896.117

Quinlan, A. R., and Hall, I. M. (2010). BEDTools: a Flexible Suite of Utilities for Comparing Genomic Features. *Bioinformatics* 26, 841–842. doi:10.1093/bioinformatics/btq033

Rabiner, L. R. (1989). A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition. *Proc. IEEE* 77, 257–286. doi:10.1109/5.18626

Ramos-Miguel, A., Barakauskas, V., Alamri, J., Miyauchi, M., Barr, A. M., Beasley, C. L., et al. (2019). The SNAP25 Interactome in Ventromedial Caudate in Schizophrenia Includes the Mitochondrial Protein ARF1. *Neuroscience* 420, 97–111. doi:10.1016/j.neuroscience.2018.12.045

Rothammer, S., Kremer, P. V., Bernau, M., Fernandez-Figares, I., Pfister-Schär, J., Medugorac, I., et al. (2014). Genome-wide QTL Mapping of Nine Body Composition and Bone mineral Density Traits in Pigs. *Genet. Sel. Evol.* 46, 68. doi:10.1186/s12711-014-0068-2

Sarropoulos, I., Marin, R., Cardoso-Moreira, M., and Kaessmann, H. (2019). Developmental Dynamics of lncRNAs across Mammalian Organs and Species. *Nature* 571, 510–514. doi:10.1038/s41586-019-1341-x

Schumacher, J. (2012). "Testis," in *Equine Surgery*. Editors J. A. Auer and J. A. Stick. Fourth Edition (Saint Louis: W.B. Saunders), 804–840. doi:10.1016/b978-1-4377-0867-7.00059-4

Shannon, P., Markiel, A., Ozier, 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

Shen, Y., Yue, F., McCleary, D. F., Ye, Z., Edsall, L., Kuan, S., et al. (2012). A Map of the Cis-Regulatory Sequences in the Mouse Genome. *Nature* 488 (7409), 116–120. doi:10.1038/nature11243

Sin, H.-S., Kartashov, A. V., Hasegawa, K., Barski, A., and Namekawa, S. H. (2015). Poised Chromatin and Bivalent Domains Facilitate the Mitosis-To-Meiosis Transition in the Male Germline. *BMC Biol.* 13, 53. doi:10.1186/s12915-015-0159-8

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

Staals, R. H. J., and Pruijn, G. J. M. (2010). "The Human Exosome and Disease," in *RNA Exosome*. Editor T. H. Jensen (New York, NY: Springer US), 132–142. doi:10.1007/978-1-4419-7841-7_11

Swindle, M. M., Makin, A., Herron, A. J., Clubb, F. J., and Frazier, K. S. (2012). Swine as Models in Biomedical Research and Toxicology Testing. *Vet. Pathol.* 49, 344–356. doi:10.1177/0300985811402846

Vastenhouw, N. L., and Schier, A. F. (2012). Bivalent Histone Modifications in Early Embryogenesis. *Curr. Opin. Cel Biol.* 24, 374–386. doi:10.1016/j.ceb.2012.03.009

Villar, D., Berthelot, C., Aldridge, S., Rayner, T. F., Lukk, M., Pignatelli, M., et al. (2015). Enhancer Evolution across 20 Mammalian Species. *Cell* 160, 554–566. doi:10.1016/j.cell.2015.01.006

Wagner, J. R., Busche, S., Ge, B., Kwan, T., Pastinen, T., and Blanchette, M. (2014). The Relationship between DNA Methylation, Genetic and Expression Inter-individual Variation in Untransformed Human Fibroblasts. *Genome Biol.* 15, R37. doi:10.1186/gb-2014-15-2-r37

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, X., and Kadarmideen, H. N. (2019). An Epigenome-wide DNA Methylation Map of Testis in Pigs for Study of Complex Traits. *Front. Genet.* 10, 405. doi:10.3389/fgene.2019.00405

Walter, W., Sánchez-Cabo, F., and Ricote, M. (2015). GOplot: An R Package for Visually Combining Expression Data with Functional Analysis. *Bioinformatics* 31 (17), 2912–2914. doi:10.1093/bioinformatics/btv300

Wang, X., Xie, W., Yao, Y., Zhu, Y., Zhou, J., Cui, Y., et al. (2020). The Heat Shock Protein Family Gene Hspa1l in Male Mice Is Dispensable for Fertility. *PeerJ* 8, e8702. doi:10.7717/peerj.8702

Washbourne, P., Thompson, P. M., Carta, M., Costa, E. T., Mathews, J. R., Lopez-Benditó, G., et al. (2002). Genetic Ablation of the t-SNARE SNAP-25 Distinguishes Mechanisms of Neuroexocytosis. *Nat. Neurosci.* 5 (1), 19–26. doi:10.1038/nn783

Wu, D., and Dean, J. (2020). EXOSC10 Sculpts the Transcriptome during the Growth-To-Maturation Transition in Mouse Oocytes. *Nucleic Acids Res.* 48, 5349–5365. doi:10.1093/nar/gkaa249

Zamudio, N., Barau, J., Teissandier, A., Walter, M., Borsos, M., Servant, N., et al. (2015). DNA Methylation Restrains Transposons from Adopting a Chromatin Signature Permissive for Meiotic Recombination. *Genes Dev.* 29, 1256–1270. doi:10.1101/gad.257840.114

Zhou, Z., Zhu, Y., Zhang, Z., Jiang, T., Ling, Z., Yang, B., et al. (2021). Comparative Analysis of Promoters and Enhancers in the Pituitary Glands of the Bama Xiang and Large White Pigs. *Front. Genet.* 12, 697994. doi:10.3389/fgene.2021.697994

**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: 07 October 2021
doi: 10.3389/fgene.2021.691592

# The Expression Regulatory Network in the Lung Tissue of Tibetan Pigs Provides Insight Into Hypoxia-Sensitive Pathways in High-Altitude Hypoxia

*Yanan Yang[1], Haonan Yuan[1], Tianliang Yang[1], Yongqing Li[2], Caixia Gao[3], Ting Jiao[1,4], Yuan Cai[1] and Shengguo Zhao[1]**

[1] College of Animal Science and Technology, Gansu Agricultural University, Lanzhou, China, [2] Research on Quality Standard of Animal Husbandry, Xinjiang Academy of Animal Sciences, Xinjiang, China, [3] State Key Laboratory of Veterinary Biotechnology, Harbin Veterinary Research Institute, Chinese Academy of Agricultural Sciences, Harbin, China, [4] College of Grassland Science, Gansu Agricultural University, Lanzhou, China

To adapt to a low-oxygen environment, Tibetan pigs have developed a series of unique characteristics and can transport oxygen more effectively; however, the regulation of the associated processes in high-altitude animals remains elusive. We performed mRNA-seq and miRNA-seq, and we constructed coexpression regulatory networks of the lung tissues of Tibetan and Landrace pigs. *HBB*, *AGT*, *COL1A2*, and *EPHX1* were identified as major regulators of hypoxia-induced genes that regulate blood pressure and circulation, and they were enriched in pathways related to signal transduction and angiogenesis, such as *HIF-1*, *PI3K-Akt*, *mTOR*, and *AMPK*. *HBB* may promote the combination of hemoglobin and oxygen as well as angiogenesis for high-altitude adaptation in Tibetan pigs. The expression of *MMP2* showed a similar tendency of alveolar septum thickness among the four groups. These results indicated that *MMP2* activity may lead to widening of the alveolar wall and septum, alveolar structure damage, and collapse of alveolar space with remarkable fibrosis. These findings provide a perspective on hypoxia-adaptive genes in the lungs in addition to insights into potential candidate genes in Tibetan pigs for further research in the field of high-altitude adaptation.

**Keywords: hypoxia, Tibetan pigs, PI3K-Akt pathway, MiRNA-mRNA network, lung tissue**

OPEN ACCESS

***Edited by:***
*Aline Silva Mello Cesar, University of São Paulo, Brazil*

***Reviewed by:***
*Bárbara Silva-Vignato, University of São Paulo, Brazil*
*Wellison J. S. Diniz, Auburn University, United States*

****Correspondence:***
*Shengguo Zhao*
*zhaoshengguo0628@hotmail.com*

***Specialty section:***
*This article was submitted to Livestock Genomics, a section of the journal Frontiers in Genetics*

***Received:*** *06 April 2021*
***Accepted:*** *31 August 2021*
***Published:*** *07 October 2021*

***Citation:***
*Yang Y, Yuan H, Yang T, Li Y, Gao C, Jiao T, Cai Y and Zhao S (2021) The Expression Regulatory Network in the Lung Tissue of Tibetan Pigs Provides Insight Into Hypoxia-Sensitive Pathways in High-Altitude Hypoxia. Front. Genet. 12:691592. doi: 10.3389/fgene.2021.691592*

## INTRODUCTION

Tibetans are a unique and geographically isolated pig breed that inhabits the Qinghai-Tibet Plateau, which has an extreme environment with high altitudes (Wang et al., 2018; Ma et al., 2019). This unique ecological condition is characterized by low air pressure, reduced oxygen content, and high ultraviolet radiation, imposing extreme physiological challenges on domestic animals, and failure to adapt will lead to altitude illness or even death (Cao et al., 2017; Lancuo et al., 2019; Qi et al., 2019). Native high-altitude species have been selected through evolutionary processes to evolve adaptive mechanisms to cope with this harsh environment (Liu et al., 2019). Special lung properties of the Tibetan pig, yak, and Tibetan sheep living in the plateau, such as larger lungs, thicker alveolar septa, and more developed capillaries, have been previously reported by Qi and Yang

(Yang et al., 2014; Qi et al., 2019). Tibetan pigs exhibit heritable adaptations to high-altitude environments as a result of natural selection. Exposure to hypoxia changes the gene profiles in various cell types and is associated with adaptation to high altitudes (Zhang T. et al., 2019). mRNAs and miRNAs are involved in many biological processes in animals, and not surprisingly, transcriptional analyses have revealed the differential expression of hypoxia regulators that enable adaptation to a hypoxic environment (Ni and Leng, 2016). The hypoxia-inducible factor-1 (HIF-1), vascular endothelial growth factor (VEGF), and mitogen-activated protein kinase (MAPK) signaling pathways are typical hypoxia-associated pathways (Lee et al., 2016; Zhang et al., 2018; Nicolas et al., 2019), and some mRNAs (*PHD2*, *VHL*, and *FIH*-1) and miRNAs (miR-363, miR-421, and miR-204) have been implicated in the regulation of the HIF-1 signaling pathway (Semenza, 2007; Ge et al., 2016; Wang et al., 2016; Xie et al., 2016).

Studies of the molecular mechanisms of livestock adaptation to high altitude have focused on miRNA-mRNA interaction networks. Here, we performed an integrative analysis of the miRNA-mRNA expression profiles in the lungs of high- and low-altitude pigs (Tibetan pigs and Landrace pigs, respectively) to identify molecular pathways and networks involved in the genetic adaptation of Tibetan pigs to hypoxic conditions.

# MATERIALS AND METHODS

## Ethics Statement

All animal experiments were conducted according to the guidelines for the care and use of experimental animals established by the Ministry of Science and Technology of the People's Republic of China (Approval number: 2006–398). The procedures for animal care were approved by the Gansu Agricultural University Animal Care and Use Committee of Gansu Agricultural University, and all experiments were conducted in accordance with approved relevant guidelines and regulations.

## Sample Collection

In total, 18 Tibetan male piglets from the highlands (TH group; Gannan Tibetan Autonomous Prefecture, Gansu, representing an altitude of 3,000 m) and 18 Landrace male piglets from the lowlands (LL group; Jingchuan, Gansu, representing an altitude of 1,000 m) with similar weights and non-genetic relationships were selected, and nine piglets from each group migrated to low altitude (TL group; Tibetan pigs at low altitude) or high altitude (LH group; Landrace pigs at high altitude) from their original rearing facility at the age of 1 month. We randomly selected six pigs from each group to collect the left lower lobes of the lung from indigenous and imported adult male pigs at the age of 6 months. These animals ($n = 6$ in each group) were feed restricted for 12 h and slaughtered in their feeding place. Six samples from each group were immediately stored in stationary liquid for hematoxylin and eosin (H&E) staining, and three of the six samples were randomly selected and collected within 1 h after the pigs were harvested and stored immediately in liquid nitrogen for subsequent RNA extraction.

**Abbreviations:** TH, Tibetan male piglets from the highlands; LL, Landrace male piglets from the lowlands; TL, Tibetan male piglets migrated to the lowlands; LH, Landrace male piglets migrated to the highlands.

## Hematoxylin and Eosin Staining

Sections from the left lower lobes of the lung were stained with H&E (Ban et al., 2018; Zhang et al., 2019b), observed under a microscope (Sunny Optical Technology Co. Ltd, Ningbo, China), and then photographed using Image View (Sunny Optical Technology Co. Ltd).

## RNA Extraction

Total RNA from the lungs was extracted using a TRIzol reagent kit (Invitrogen, Carlsbad, CA, USA) according to the manufacturer's protocol, and eukaryotic mRNA was enriched by oligo (dT) beads (Epicenter, Madison, WI, USA). RNA quality was assessed on an Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA, USA) and verified by 1% gel electrophoresis. All samples presented an RNA integrity number (RIN) $> 7.5$.

## Library Construction and Sequencing for mRNA

After total RNA was extracted, eukaryotic mRNA was enriched by oligo (dT) beads (Epicenter) and reverse-transcribed into cDNA using random primers. mRNA was ligated with proper 5′ and 3′ adapters. The ligation products were reverse-transcribed by PCR amplification to generate a cDNA library, which was sequenced using an Illumina HiSeq™ 2500 by Gene Denovo Biotechnology Co. (Guangzhou, China).

## Library Construction and miRNA Sequencing

After total RNA was extracted for miRNA sequencing, 18–30 nt RNA molecules were enriched by polyacrylamide gel electrophoresis (PAGE). A 3′ adapter was added to enrich the 36–44 nt RNAs, and the 5′ adapter was then connected to the RNA. PCR products of 140–160 bp were amplified by reverse transcription. A cDNA library was generated and sequenced using Illumina HiSeq™ 2500 sequencing (Illumina Inc., San Diego, CA, USA) by Gene Denovo Biotechnology Co., Ltd.

## Expression Analysis of mRNAs

High-quality clean raw data were screened by removing low-quality data with fastp (Chen et al., 2018). The short-read alignment tool, Bowtie 2 (Langmead and Salzberg, 2012) was used to map reads to the ribosome RNA (rRNA) database. An index of the reference genome was built, and paired-end clean reads were mapped to *Sus scrofa* RefSeq (*Sus scrofa* 11.1) using HISAT 2 (Kim et al., 2015). The mapped reads of each sample were assembled using StringTie v1.3.1 (Pertea et al., 2015, 2016) in a reference-based approach. For each transcription region, a fragment per kilobase of transcript per million mapped reads (FPKM) value was calculated to quantify its expression abundance and variations using RSEM software. RNA differential expression analysis was performed with DESeq 2 (Love et al.,

2014) software between the two groups. The raw mRNA-seq data (accession number PRJNA687172) were submitted to the Sequence Read Archive (SRA) database of NCBI.

## Expression Analysis of miRNAs

Clean reads were obtained by filtering raw reads, and all of them were aligned with small RNAs in the GenBank database (Benson et al., 2013). All the clean reads were aligned with small RNAs in the Rfam database (Griffiths-Jones et al., 2003) to identify and remove rRNAs, scRNAs, snoRNAs, snRNAs, and tRNAs. All the clean reads were also aligned with the reference genome and were searched against the miRbase database (Griffiths-Jones et al., 2006) to identify known (*Sus scrofa*) miRNAs. All the unannotated reads were aligned with the reference genome by HISAT2. 2.4. Novel miRNA candidates were identified according to their genome positions and hairpin structures predicted by mirdeep2 software. The miRNA expression levels were calculated and normalized to transcripts per million (TPM). The raw miRNA-seq data (accession number PRJNA687649) were submitted to the NCBI Sequence Read Archive (SRA) database.

## Functional Annotation of DEmRNAs

DEmRNAs were analyzed using Kyoto Encyclopedia of Genes and Genomes (KEGG) and Gene Ontology (GO) analyses using the online tool Database for Annotation, Visualization and Integrated Discovery (DAVID) (Huang et al., 2009) to explore their roles, functions, and enrichment in different biological pathways. Gene Ontology (GO) terms and pathways with $q < 0.05$ were considered significantly enriched by DEmRNAs. The hypoxic DEmRNAs were filtered based on the intersection of our results and published hypoxia-related genes in the HIF-1 signaling pathway. The hypoxia-related genes and target genes of miRNAs were also mapped to GO terms in the GO database and pathways in the KEGG (Kyoto Encyclopedia of Genes and Genomes) database to further elucidate their functions.

## Target Prediction and Integrative Analysis of the Hypoxia-Related miRNA–mRNA Regulatory Network

We identified mRNAs with a fold change ≥2 and a false discovery rate (FDR) <0.05 as DEmRNAs. To explore more DEmiRNAs, we identified miRNAs with fold change ≥ 2 and $p < 0.05$ as DEmiRNAs. The potential target genes of DEmiRNAs were predicted using RNAhybrid 89 (version 2.1.2) + svm_light (version 6.01), miRanda (version 3.3a), and TargetScan (version 7.0), and the genes at the intersection of the results from the three software packages were selected as predicted miRNA target genes.

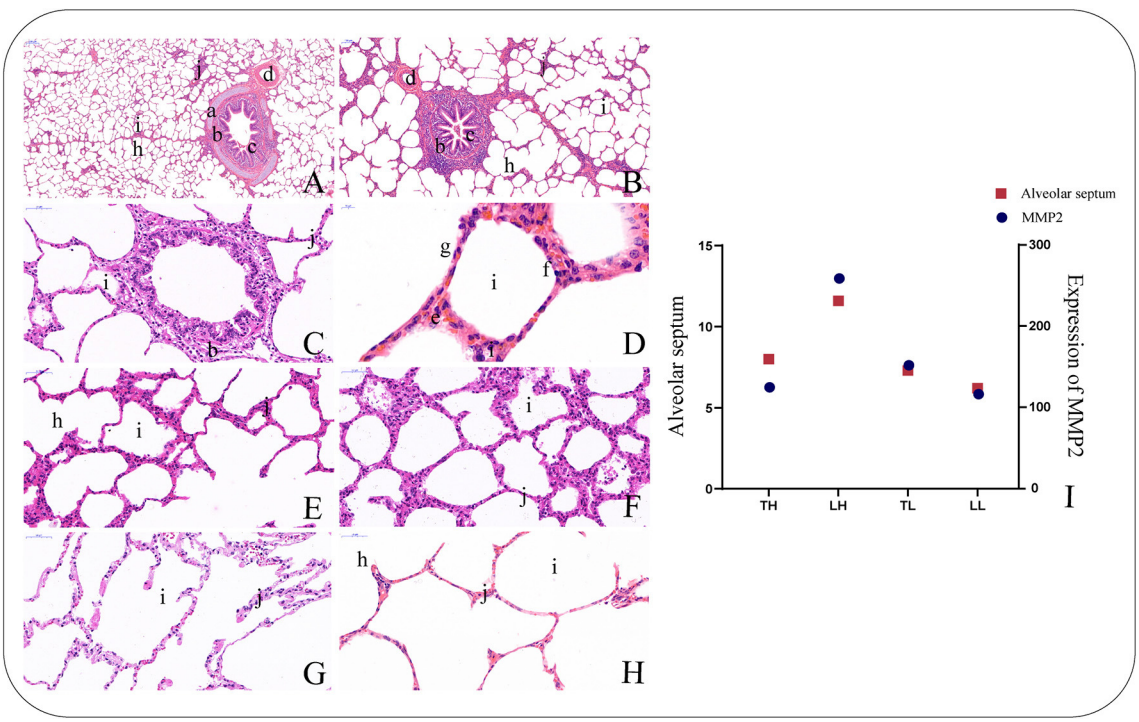


**FIGURE 1 |** Morphological characteristics of the lungs in Tibetan pigs by H and E staining. (40×). **(A)** Morphological characteristics of small bronchiole. **(B)** Morphological characteristics of bronchiole. **(C)** Morphological characteristics of terminal bronchiole. **(D)** Morphological characteristics of alveolar cells. **(E)** Morphological characteristics of alveolar septa in the Tibetan pigs raised in highland (TH) group. **(F)** Morphological characteristics of alveolar septa in the Landrace pigs raised in highland (LH) group. **(G)** Morphological characteristics of alveolar septa in the Tibetan pigs raised in lowland (TL) group. **(H)** Morphological characteristics of alveolar septa in the Landrace pigs raised in lowland (LL) group. **(I)** The histogram shows the expression levels of *MMP2* in the lungs of the four types of pigs and the relationship with alveolar septa. **a.** Piece of cartilage. **b.** Smooth muscle. **c.** Plica. **d.** Arteriole. **f.** Alveolar epithelial type II cells. **g.** Alveolar epithelial type I cells. **h.** Alveolar duct. **i.** Pulmonary alveoli. **j.** Alveolar septa.

Because mRNAs and miRNAs have potential negative regulatory relationships, we assessed the expression correlation between a miRNA and its predicted target by the Pearson correlation coefficient (PCC). Subsequently, the negatively coexpressed miRNA–mRNA pairs with PCC $< -0.7$ and $p < 0.05$ were screened to construct miRNA–mRNA networks.

The coexpression network diagram of DEmRNAs and DEmiRNAs was generated using the PCC, and only the relationship pair network diagram of the top 300 is shown. The coexpression network diagram of the 273 hypoxic DEmRNAs is displayed, and the correlation between miRNA and mRNA was required to account for the top 5% of the total correlation. The potential regulatory network was constructed by Cytoscape (Szklarczyk et al., 2015).

## Quantitative Real-Time PCR Validation

Total RNA from pulmonary tissues was extracted with a TRIzol reagent kit and reverse-transcribed into cDNA using a FastQuant cDNA first-strand synthesis kit (TianGen, China). SYBR® Premix Ex Taq™ II (TaKaRa, China) was used for real-time fluorescence quantitative analysis. In total, eight DEmRNAs and eight DEmiRNAs were randomly selected to determine sequencing accuracy. The primers used here were designed using Primer 5.0 software and are listed in **Supplementary Tables 1**, **2** (**Supplementary Material 1**).

The experimental data were analyzed with the $2^{-\Delta\Delta CT}$ method (Livak and Schmittgen, 2001). Statistical analyses were performed using GraphPad Prism 8.0 (GraphPad Software, San Diego, CA, USA) and SPSS 20.0 (SPSS, Chicago, IL, USA). The comparisons were conducted by one-way analysis of variance (ANOVA), and $p < 0.05$ was considered statistically significant.

# RESULTS

## Morphological Structure

H&E staining showed that the lung sections exhibited the following connective tissues with epithelia: pulmonary alveoli, smooth muscle, blood capillaries, bronchial tubes, and alveolar septa (**Figures 1A–H**). The sections from the TH group were characterized by smooth muscle hyperplasia and larger alveoli, while those from the LH group were characterized by a thicker alveolar septum. In addition, the analysis showed that *MMP*2 expression had a similar tendency to the alveolar septum thickness among the four groups (**Figure 1I**).

## Identification of DEmRNAs in the Lung

In total, 12 cDNA libraries, which included six Tibetan pigs and six Landrace pigs at high and low altitudes, were sequenced from lung tissues (**Supplementary Material 2**). After quality filtering, 51,193,662–69,112,222 clean paired reads were obtained with 99.70–99.79% of clean reads mapped to the porcine reference genome (**Table 1**). A total of 471 DEmRNAs (247 up- and 224 downregulated) were identified in the TH group compared to the TL group (**Figure 2**, **Supplementary Table 3** in **Supplementary Materials 1**, **3**). Furthermore, 809 novel genes were identified in the sequencing data. Eight mRNAs were randomly selected and detected using qRT-PCR to validate the

**TABLE 1** | Overview of the reads and quality control of the 12 libraries of the mRNA sequencing from swine lung tissue.

| Sample | Raw data | Clean data(bp) | Clean reads | Q20 (%) | Q30 (%) | GC (%) | Total mapped (%) | Unique mapped (%) |
|---|---|---|---|---|---|---|---|---|
| LL-1 | 62,519,034 | 9,305,579,147 | 62,348,446 | 9,060,896,883 (97.37%) | 8,641,334,794 (92.86%) | 5,271,030,180 (56.64%) | 59,539,505 (95.92%) | 56,671,208 (91.30%) |
| LL-2 | 65,884,878 | 9,804,022,124 | 65,723,768 | 9,553,412,100 (97.44%) | 9,118,107,584 (93.00%) | 5,492,767,034 (56.03%) | 61,941,659 (95.83%) | 58,584,049 (90.63%) |
| LL-3 | 62,450,712 | 9,299,635,729 | 62,295,524 | 9,062,302,058 (97.45%) | 8,648,230,615 (93.00%) | 5,167,314,374 (55.56%) | 58,667,319 (95.66%) | 56,224,782 (91.68%) |
| LH-1 | 62,323,778 | 9,266,150,211 | 62,180,224 | 9,048,646,333 (97.65%) | 8,658,401,070 (93.44%) | 5,409,933,897 (58.38%) | 59,678,317 (96.35%) | 57,179,752 (92.31%) |
| LH-2 | 69,319,172 | 10,309,201,883 | 69,112,222 | 10,032,974,281 (97.32%) | 9,562,740,889 (92.76%) | 5,793,441,899 (56.20%) | 65,082,781 (95.76%) | 62,361,180 (91.76%) |
| LH-3 | 60,877,578 | 9,057,244,483 | 60,694,770 | 8,807,812,613 (97.25%) | 8,387,856,985 (92.61%) | 5,060,253,346 (55.87%) | 56,200,627 (95.66%) | 53,771,349 (91.53%) |
| TH-1 | 68,421,860 | 10,180,216,302 | 68,240,484 | 9,907,699,993 (97.32%) | 9,443,251,897 (92.76%) | 5,793,799,001 (56.91%) | 64,529,051 (95.18%) | 61,876,279 (91.27%) |
| TH-2 | 68,109,198 | 10,130,609,161 | 67,917,554 | 9,862,307,770 (97.35%) | 9,401,034,430 (92.80%) | 5,643,159,581 (55.70%) | 64,353,888 (95.25%) | 61,919,184 (91.65%) |
| TH-3 | 51,303,388 | 7,628,671,504 | 51,193,662 | 7,442,138,700 (97.55%) | 7,108,738,685 (93.18%) | 4,248,139,426 (55.69%) | 48,728,948 (95.77%) | 46,580,671 (91.54%) |
| TL-1 | 53,728,490 | 7,983,823,434 | 53,609,850 | 7,782,090,571 (97.47%) | 7,429,812,850 (93.06%) | 4,508,645,123 (56.47%) | 48,895,354 (91.49%) | 44,125,432 (82.56%) |
| TL-2 | 48,511,592 | 7,230,580,616 | 48,384,742 | 7,041,331,495 (97.38%) | 6,713,834,168 (92.85%) | 4,035,618,474 (55.81%) | 45,033,230 (93.36%) | 42,700,246 (88.53%) |
| TL-3 | 53,850,276 | 8,020,766,562 | 53,729,084 | 7,827,485,412 (97.59%) | 7,484,806,819 (93.32%) | 4,541,433,778 (56.62%) | 49,703,779 (92.99%) | 46,478,435 (86.96%) |

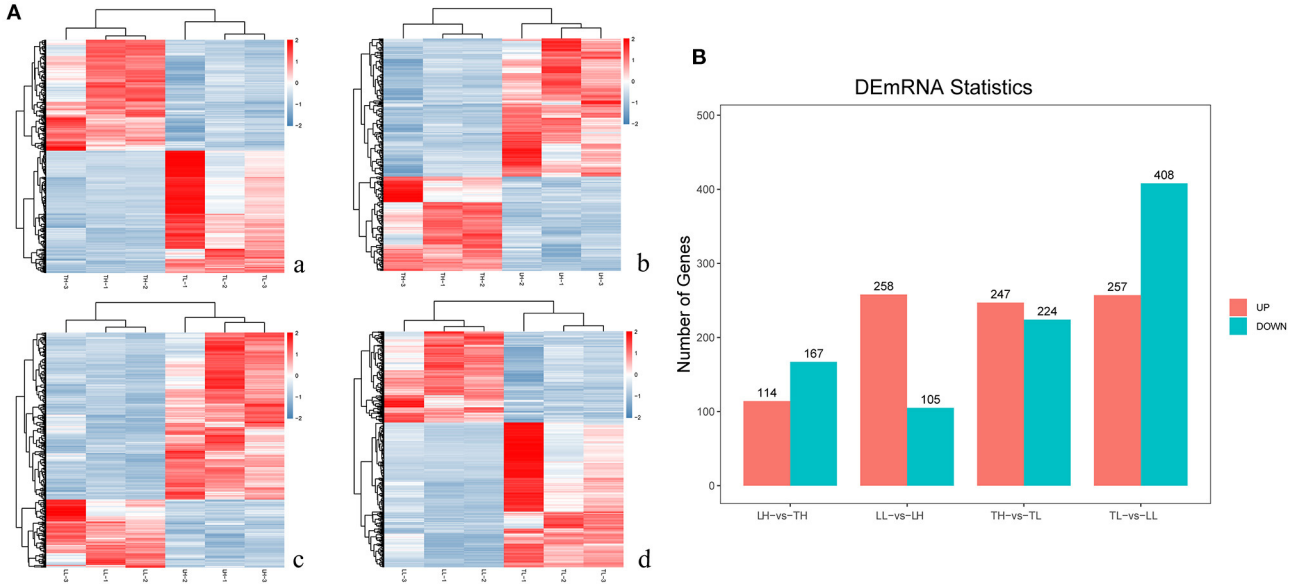


**FIGURE 2 |** mRNA expression results among the four groups by RNA-seq. **(A)** The heatmap shows the relative expression patterns of DEmRNAs among the four groups. Each column represents a sample, and each row represents the expression levels of a single mRNA in various samples. The color scale of the heat map ranges from blue (low expression) to red (high expression). **a.** Heatmap of mRNA for Tibetan pigs raised in highland (TH) and Tibetan pigs raised in lowlands (TL). **b.** Heatmap of mRNA for TH and Landrace pigs raised in highlands (LH). **c.** Heatmap of mRNA for LH and Landrace pigs raised in lowlands (LL). **d.** Heatmap of mRNA for TL and LL. **(B)** The histogram shows the number of DEmRNAs identified among the four groups.

accuracy of the sequencing data. Our verification test indicated that the qRT-PCR results were consistent with the mRNA-seq data (**Figure 4A**).

## Identification of DEmiRNAs in the Lung

A total of 12 cDNA libraries were sequenced from lung tissues. In the miRNA-seq data, 10,810,538–14,920,316 clean reads were obtained by removing low-quality data and data with sequences shorter than 18 nt and longer than 30 nt, and 94.380–97.30% clean reads were obtained and mapped (**Table 2**). A total of 464 DEmiRNAs (324 up- and 140-downregulated) were identified in the TH group compared to the TL group (**Figure 3**, **Supplementary Table 4** in **Supplementary Materials 1**, **4**). Eight miRNAs were randomly selected and detected using qRT-PCR to validate the accuracy of the sequencing data. Our verification test indicated that the qRT-PCR results were consistent with the miRNA-seq data (**Figure 4B**).

## Functional Analysis of DEmRNAs

GO and KEGG enrichment analyses showed that most DEmRNAs were involved in cellular processes and pathways related to cytokine-cytokine receptor interaction, the PI3K-Akt signaling pathway, and pathways in cancer (**Figure 5**). Interestingly, a number of genes were mainly enriched in "response to stimulus (GO: 0050896)" of biological process among the four groups. GO: 0001071 is associated with nucleic acid binding transcription factor activity and was significantly enriched between the TH and LH groups. The top 20 pathways with the most significant enrichment were obtained. KEGG enrichment results revealed that most of these genes were significantly enriched in cancer pathways among Landrace pigs (LH and LL) (breast cancer and transcriptional misregulation in cancers) or high-altitude groups (LH and TH) (proteoglycans in cancer, pathways in cancer, breast cancer). A number of genes were significantly enriched in cytokine–cytokine receptor interaction, hematopoietic cell lineage, and African trypanosomiasis among Tibetan pigs and Landrace pigs in the high- or low-altitude groups. Six pathways were significantly enriched among the high- (TH and LH) or low-altitude (TL and LL) groups, and 15 pathways were significantly enriched between the Tibetan pig (TH and TL) and Landrace pig (LH and LL) groups.

## Identification and Prediction Targets of DEmiRNAs

A total of 59,636 target DEmRNAs of 1,630 DEmiRNAs (365 functionally annotated miRNAs, 989 known miRNAs and 276 novel miRNAs) were analyzed (**Supplementary Materials 5**). In addition, multiple pathways and GO terms were associated with hypoxia traits. The analysis revealed KEGG pathways that were significantly related to genes targeted by DEmiRNAs, and the Wnt signaling pathway, metabolic pathway and hepatocellular carcinoma were the most significantly related (**Figure 6**). Interestingly, the results showed that the targets were primarily enriched in terms related to hypoxia adaptation. ssc-miR-210, ssc-miR-101, ssc-miR-7136-5p, ssc-miR-10b, ssc-miR-206, ssc-miR-1343, ssc-miR-142-5p, ssc-miR-421-5p, and ssc-miR-4331 were identified as key miRNAs. Functional assessment showed that 100, 56, and 104 putative targets were mainly enriched in the HIF, PI3K-Akt, and MAPK signaling pathways, respectively.

**TABLE 2 |** Overview of the reads and quality control of the 12 libraries of the miRNA sequencing from swine lung tissue.

| Sample | Clean_reads | High_quality | Smaller_than_18nt | Clean reads | Match | Ratio (%) |
|---|---|---|---|---|---|---|
| LL-1 | 11,823,173 (100%) | 11,800,040 (99.8043%) | 117,090 (0.9923%) | 147,719,777 | 122,350,552 | 82.83 |
| LL-2 | 11,353,289 (100%) | 11,339,875 (99.8818%) | 90,118 (0.7947%) | 11,422,900 | 9,303,536 | 81.45 |
| LL-3 | 13,427,720 (100%) | 13,404,468 (99.8268%) | 107,640 (0.8030%) | 10,937,366 | 9,065,662 | 82.89 |
| LH-1 | 13,721,166 (100%) | 13,702,350 (99.8629%) | 303,894 (2.2178%) | 13,042,101 | 10,696,421 | 82.01 |
| LH-2 | 13,752,929 (100%) | 13,569,709 (98.6678%) | 261,550 (1.9275%) | 13,042,192 | 11,012,185 | 84.44 |
| LH-3 | 13,530,906 (100%) | 13,507,736 (99.8288%) | 153,784 (1.1385%) | 12,902,983 | 10,703,895 | 82.96 |
| TH-1 | 12,421,634 (100%) | 12,401,306 (99.8364%) | 165,479 (1.3344%) | 13,093,151 | 10,784,226 | 82.37 |
| TH-2 | 12,972,521 (100%) | 12,954,856 (99.8638%) | 247,892 (1.9135%) | 11,909,246 | 9,793,972 | 82.24 |
| TH-3 | 13,503,266 (100%) | 13,476,246 (99.7999%) | 161,856 (1.2010%) | 12,330,439 | 10,255,345 | 83.17 |
| TL-1 | 10,810,538 (100%) | 10,788,450 (99.7957%) | 132,232 (1.2257%) | 12,998,511 | 10,664,603 | 82.04 |
| TL-2 | 14,920,316 (100%) | 14,892,778 (99.8154%) | 388,770 (2.6105%) | 10,283,765 | 8,634,859 | 83.97 |
| TL-3 | 12,392,682 (100%) | 12,373,187 (99.8427%) | 226083 (1.8272%) | 14,056,433 | 11,724,322 | 83.41 |

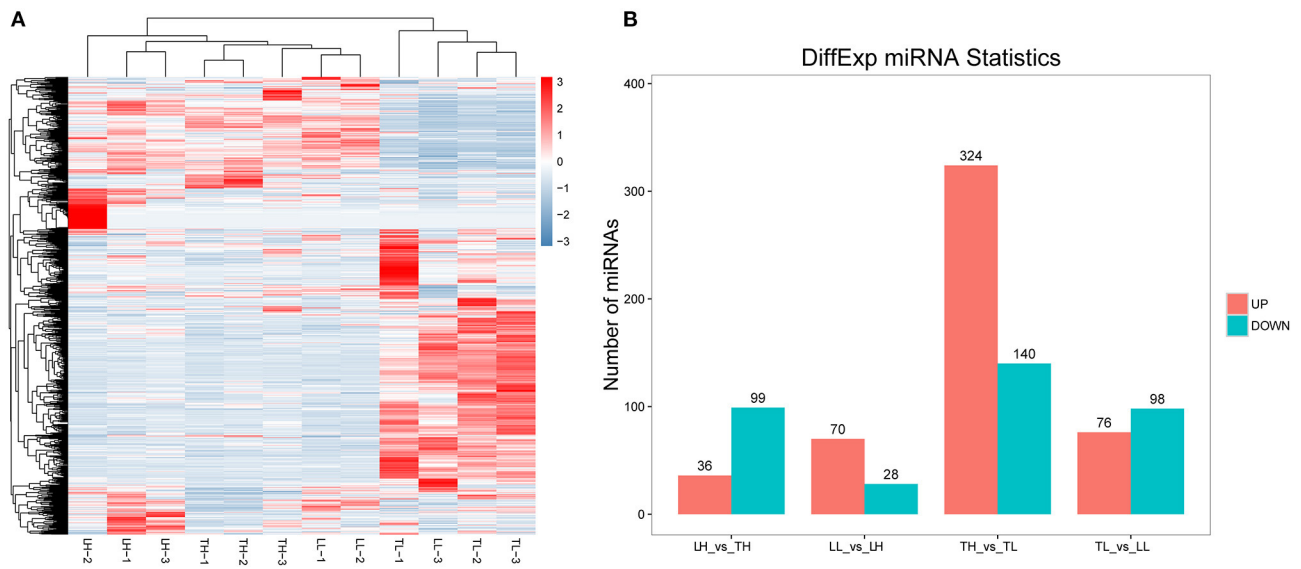


**FIGURE 3 |** miRNA expression results among the four groups by RNA-seq. **(A)** The heatmap shows the relative expression patterns of DEmiRNAs among the four groups. Each column represents a sample, and each row represents the expression levels of a single miRNA in various samples. The color scale of the heat map ranges from blue (low expression) to red (high expression). **(B)** The histogram shows the number of DEmiRNAs identified among the four groups.

## Screening of Differentially Expressed Hypoxia-Related mRNA Target miRNAs and Their Functional Enrichment Analysis

Functional analysis was conducted to understand the pathways and molecular interactions of DEmRNAs and DEmiRNAs. The DEmRNAs were enriched in a number of important pathways related to hypoxia, and we identified 273 significant DEmRNAs involved in hypoxia adaptation among the four groups (**Supplementary Material 6**). We predicted potential target miRNAs of mRNAs according to the negative regulatory effects of miRNAs on mRNAs, which were further considered veritable miRNA–mRNA pairs. To further reveal the regulatory relationship of node mRNAs and non-coding miRNAs, the resulting potential regulatory networks of miRNA-target genes associated with hypoxia-genes were constructed (**Figure 7**). The target DEmRNAs of DEmiRNAs were assessed using KEGG and GO analyses. The results indicated that 71.09, 17.00, and 11.90% (total of 273) of the genes were enriched in the biological process (BP), cell component (CC), and molecular function (MF) categories, respectively, in the TH-vs.-TL comparison ($p < 0.05$). A number of genes were targeted by hub miRNAs, such as novel-m0237-5p, novel-m0173-3p, and novel-m0142-5p, which had 45, 19, and 14 target mRNAs among the four groups, respectively. Furthermore, miR-2465-x targeted HIF-1α, while novel-m0087-3p and novel-m0237-5p targeted *HIF*-3α.

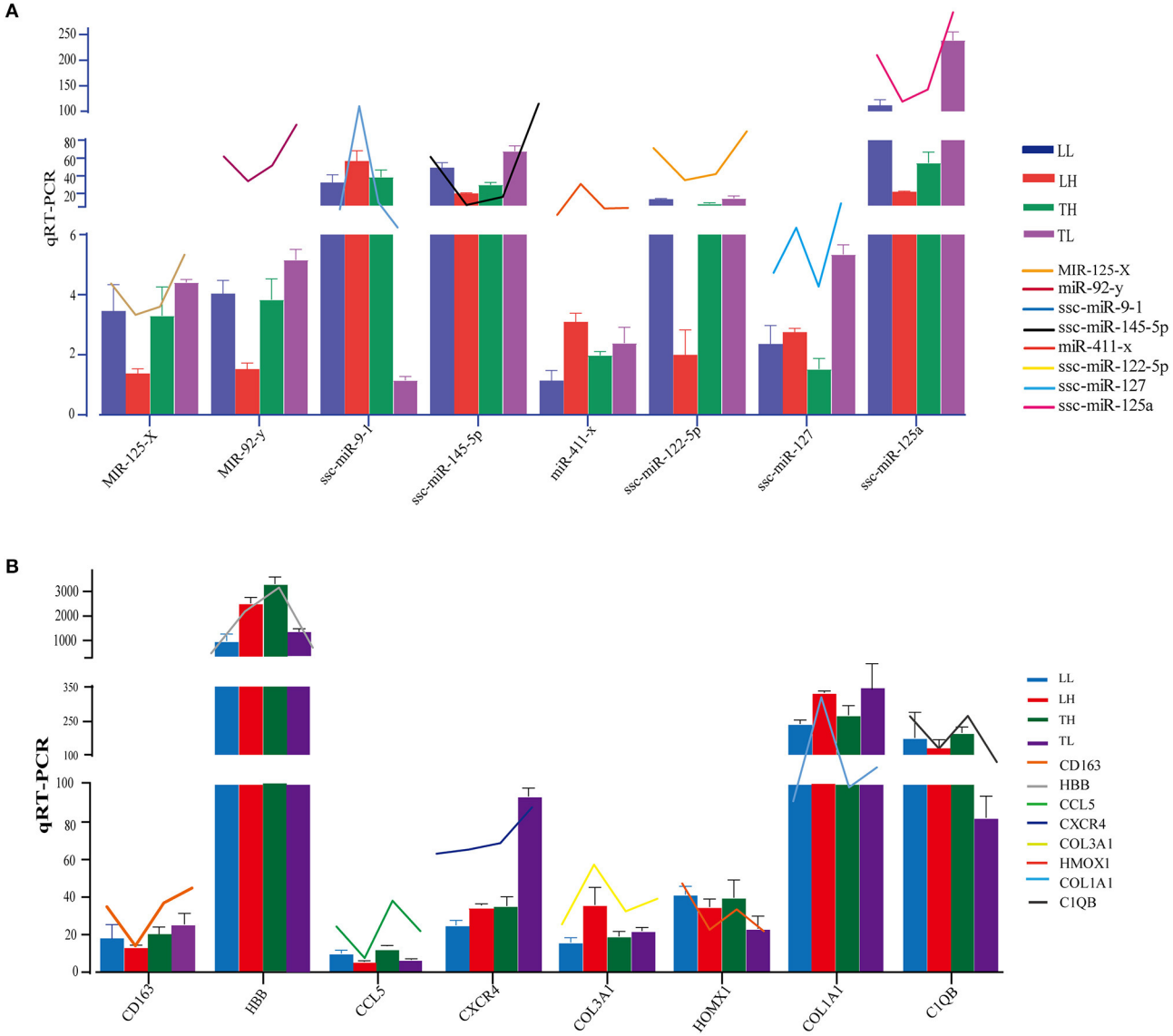


**FIGURE 4 |** Expression patterns of randomly selected DEmRNAs and DEmiRNAs. **(A)** Eight mRNA expression levels were confirmed by qRT-PCR in comparison to corresponding data detected in mRNA-Seq. GAPDH was used as control. **(B)** Eight miRNA expression levels were confirmed by qRT-PCR in comparison to corresponding data detected in mRNA-Seq. U6 was used as control. The broken line indicates the change in transcript level according to the FPKM value of mRNA-seq and miRNA-seq. Three biological replicates with three technical replicates each were used. The values represent the mean ± SE ($n = 3$).

## Construction of the Coexpression Network Between DEmRNAs and DEmiRNAs in Response to Hypoxia

To explore the relationship between miRNAs and mRNAs in a hypoxic environment, a coexpression network of DEmRNAs and DEmiRNAs was constructed, and the top 300 relationship pair network diagrams are listed (**Figure 8A**, **Supplementary Material 7.1**). The intersection of differentially expressed hypoxia mRNAs and miRNAs identified from the four group comparisons represented their differential expression in pig lungs with increasing altitude. *TAR1-A*, *GPD*1, *ST8SIA*5, and *LENG*8 were selected as the most affected mRNAs, and there were strong correlations with a number of miRNAs. Furthermore, a coexpression network of 273 hypoxic DEmRNAs and DEmiRNAs was constructed (**Figure 8B**, **Supplementary Material 7.2**). *MEF*2*C*, AKAP6, *NTRK*2, *MAPT*, and *GPR*146 were selected as the most affected mRNAs, and there were strong correlations with a number of miRNAs.

## DISCUSSION

A high-altitude environment plays an important role in the adaptation of native species, and it may modify gene

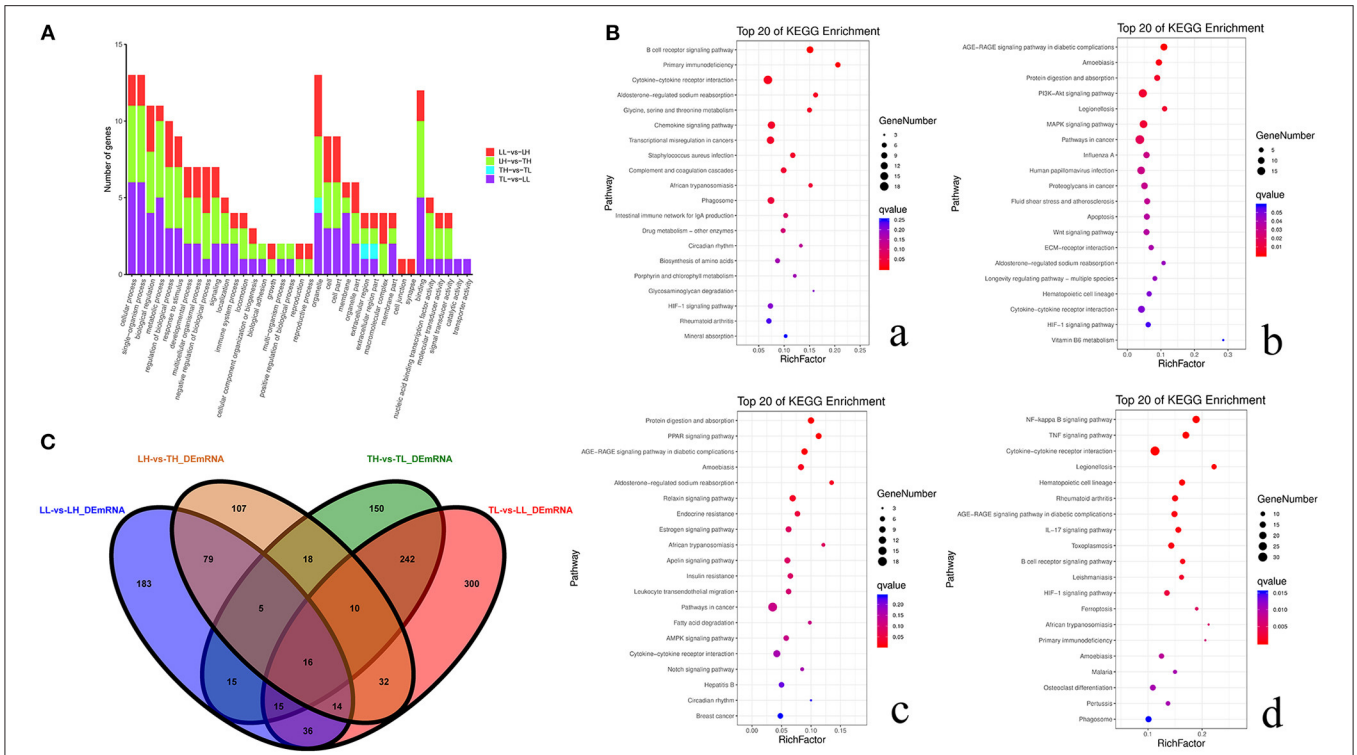


**FIGURE 5 |** Functional annotation analysis of DEmRNAs in porcine lungs among the four groups. **(A)** Histogram of GO annotation results of DEmRNAs. The abscissa is the second level GO term, and the ordinate is the number of DEmRNAs in the term. Red indicates the number of DEmRNAs between Landrace pigs raised in lowlands (LL) and Landrace pigs raised in highlands (LH) groups, green indicates the number of DEmRNAs between LH and Tibetan pigs raised in highlands (TH) groups, blue indicates the number of DEmRNAs between TH and Tibetan pigs raised in lowlands (TL) groups, and purple indicates the number of DEmRNAs between TL and LL groups. **(B)** Top 20 KEGG enrichment pathways of DEmRNAs. The ordinate is the pathway, and the abscissa is the enrichment factor. Darker colors indicate smaller q-values. **a.** Pathway enrichment analysis of DEmRNAs between TH and TL. **b.** Pathway enrichment analysis of DEmRNAs between TH and LH. **c.** Pathway enrichment analysis of DEmRNAs between LH and LL. **d.** Pathway enrichment analysis of DEmRNAs between TL and LL. **(C)** Venn diagram of mRNA interactions based on the overlapping mRNAs among the four groups.

transcription and may irreversibly affect specific phenotypes (Zhang et al., 2015; Ni et al., 2019). We used a complete migrant design to evaluate genes interacting with the environment and selected Tibetan pigs and Landrace pigs in both their native altitude environments and as migrants in a non-native environment. Our previous research identified that Tibetan pigs have heavier and wider lungs, thicker alveolar septa, and a denser vascular network than Landrace pigs. The hemoglobin (HGB) and mean corpuscular hemoglobin concentration (MCHC) of high-altitude pigs (Tibetan and Landrace pigs) were significantly higher than those of low-altitude pigs (Yang et al., 2021). We next investigated whether there are gene expression changes specific to Tibetan pigs that are responsible for hypoxic adaptation. Sequencing of multiple pigs from different breeds revealed that certain genomic regions, including genes involved in the hypoxia response, were under selection in Tibetan pigs (Zhang B. et al., 2016; Zhang et al., 2017, 2019a). We screened for key genes related to hypoxic adaptation through genotype and environment interaction effects *via* RNA-seq analyses. Several pathways were enriched in DEmRNAs among Tibetan pigs and Landrace pigs at different altitudes, including the VEGF signaling pathway, PI3K-AKT signaling pathway, and mTOR signaling pathway (Ai et al., 2015; Zhang et al., 2017). Moreover, GO enrichment analysis revealed that these DEmRNAs were associated with vascular regulation, regulatory region DNA binding, or extracellular region. The identified hypoxia-related signaling pathways may form a complex cascade of responses that occur in hypoxic conditions in Tibetan pigs to reduce the risk of pulmonary damage.

Hypoxia-regulated miRNAs play vital roles in cell survival and have been implicated in the regulation of both upstream and downstream HIF-1 signaling pathways under hypoxic conditions. For example, miR-199a, miR-17-92, and miR-20b induce HIFs (Dai et al., 2015; Chen et al., 2016; Danza et al., 2016). *HIF*-1 regulates the expression of various genes to protect cells from hypoxic injury through cell apoptosis, glucose metabolism, and mitochondrial function (Bhattarai et al., 2018; Yu et al., 2018, 2020). *HIF*-1α is a potential therapeutic proangiogenic molecule that regulates the levels of *VEGF* to elevate interstitial pressure (Zhi et al., 2018; Lin et al., 2019). Several putative target genes (*FOXO*3, *RASGRF*1, and *CX3CR*1) that are regulated by ssc-miR-214, ssc-miR-320, and ssc-miR-101 have been found to be involved in the HIF-1 related signaling pathway. miR-210 is located on human chromosome

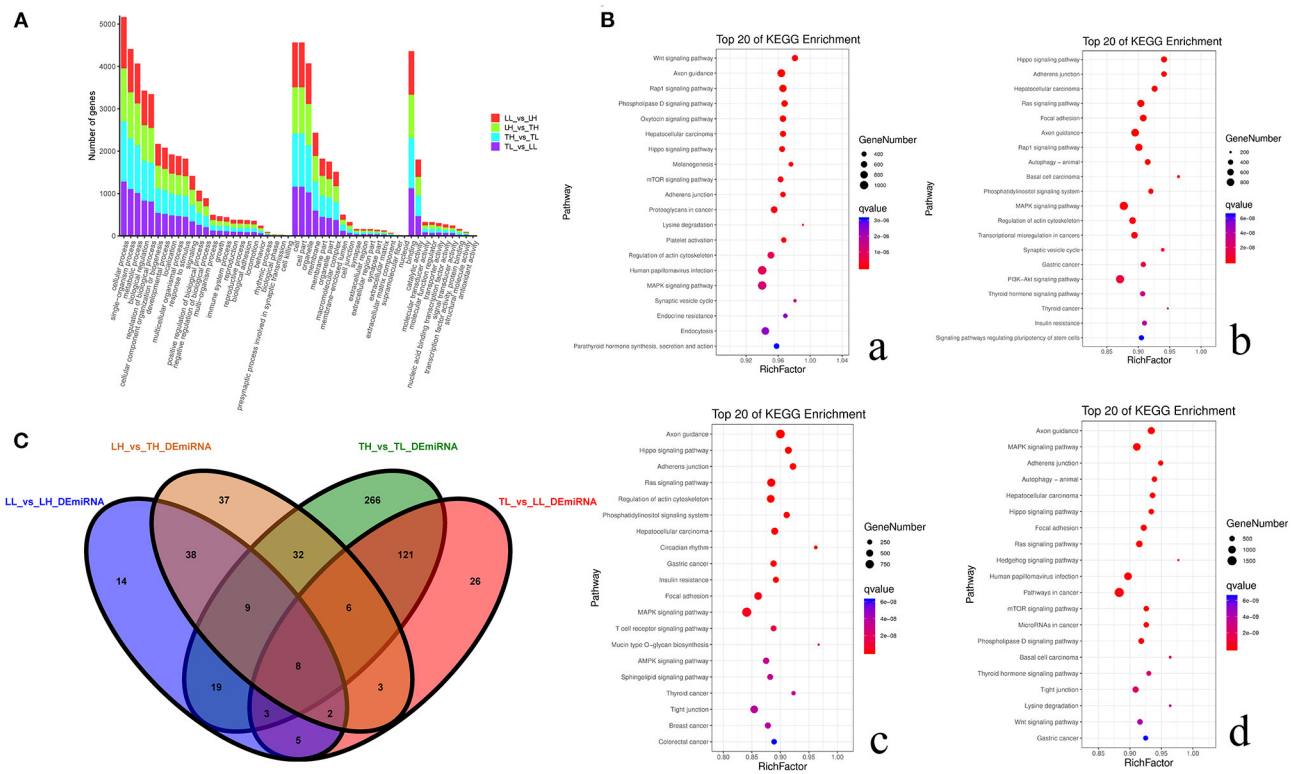


**FIGURE 6 |** Functional annotation analysis of DEmiRNA-target genes in porcine lungs among the four groups. **(A)** Histogram of GO annotation results of DEmiRNA-target genes. The abscissa is the second-level GO term, and the ordinate is the number of DEmiRNA-target genes in the term. Red indicates the number of DEmiRNA-target genes between Landrace pigs raised in lowlands (LL) and Landrace pigs raised in highland (LH) groups, green indicates the number of DEmiRNA-target genes between LH and Tibetan pigs raised in highland (TH) groups, blue indicates the number of DEmiRNA-target genes between TH and Tibetan pigs raised in lowland (TL) groups, and purple indicates the number of DEmiRNA-target genes between TL and LL groups. **(B)** Top 20 KEGG enrichment pathways of DEmiRNA-target genes. The ordinate is the pathway, and the abscissa is the enrichment factor. Darker colors indicate smaller q-values. **a.** Pathway enrichment analysis of DEmiRNA-target genes between TH and TL. **b.** Pathway enrichment analysis of DEmiRNA-target genes between TH and LH. **c.** Pathway enrichment analysis of DEmiRNA-target genes between LH and LL. **d.** Pathway enrichment analysis of DEmiRNA-target genes between TL and LL. **(C)** Venn diagram of miRNA interactions based on the overlapping miRNAs among the four groups.

11p15.5 and correlates with angiogenesis and *VEGF* regulation in breast cancer patients (Forkens et al., 2008; Dai et al., 2015; Tang et al., 2018; Zhang H. et al., 2019). In the present study, the expression of miR-210-x and miR-210-z was significantly lower in TH than in TL but not significantly different in the other groups, which may play vital roles in the expression of proteins in homology-dependent repair pathways and nucleotide excision repair pathways to reverse cellular DNA damage in the lungs of Tibetan pigs during hypoxia (Crosby et al., 2009; Hui et al., 2019). *HBB* is involved in the malaria reference pathway and downregulates *IL*-6, which is a key gene in the HIF-1 pathway. Comparison of *HBB* expression between Tibetan pigs and Landrace pigs showed that among the beta globin amino acid substitutions at positions 58, 75, 119 and 137, the replacement of alanine at position 137 with valine and the locus mutation improved the affinity of HGB and $O_2$ (Zhang B. et al., 2016). The expression of the HBB gene in Tibetan pigs (TH and TL) was significantly higher than that in Landrace pigs (LH and LL), agreeing with a similar trend previously reported by other authors, and there was similar variation in the HGB concentration in Tibetan pigs (Taliercio et al., 2013; Zhang B. et al., 2016; Yang et al., 2021), indicating that hypoxia transcriptionally upregulates HBB to increase HGB in the blood to ensure the transport of blood and nutrients. These findings may (Jang et al., 2014; Zhang G. et al., 2016; Cai et al., 2018) explain why Tibetan pigs have better adaption than Landrace pigs in hypoxic environments regardless of altitude.

The PI3K/Akt pathway is an intracellular signaling pathway that is promoted by several biological molecules, including calmodulin, insulin-like growth factor (IGF), and multiple EGF-like domains 6 (MEGF6) (Pompura and Dominguez-Villar, 2018; Ellis and Ma, 2019; Revathidevi and Munirajan, 2019; Zhang et al., 2020). *IGF*2 is the target gene of miR-506-y and ssc-miR-181d-3p. The expression levels of *IGF*2 and *MEGF*6 were significantly upregulated in LH compared to LL, but no differences were found in the Tibetan pigs (TH and TL). We hypothesized that these genes may induce the growth, proliferation, and differentiation of tumor cells in the lungs of Landrace pigs living in a hypoxic environment (Mohlin et al., 2013). Activated Akt induces various biological

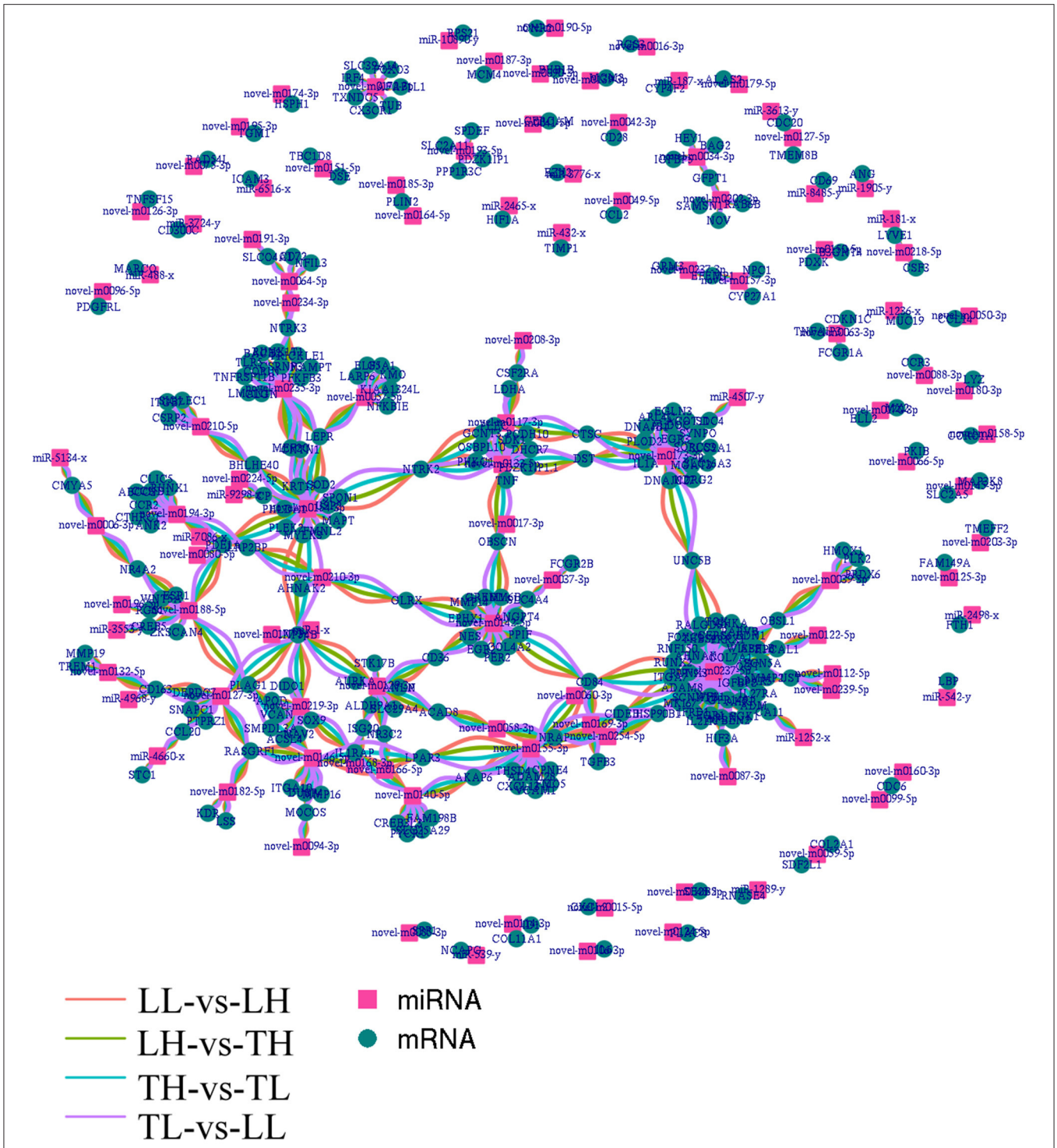


**FIGURE 7 |** Integrated miRNA-target negative correlation regulatory network. The square nodes represent miRNAs, and the circle nodes represent target genes. The differentially expressed miRNAs/target genes are highlighted as follows: red indicated differential expression between Landrace pigs raised in lowlands (LL) and Landrace pigs raised in highland (LH), green indicates differential expression between LH and Tibetan pigs raised in highland (TH), blue indicates differential expression between TH and Tibetan pigs raised in lowland (TL), and purple indicates differential expression between TL and LL.

processes, including activating mTOR, localizing FOXO to the cytoplasm, and activating cAMP-response element binding protein (CREB) (Zhang et al., 2011; Gaecía-Morales et al., 2017; Marquard and Jücker, 2020). The FOXO signaling pathway was also enriched in a comparison of pigs living at different altitudes. It has been shown that alcohol suppresses P450

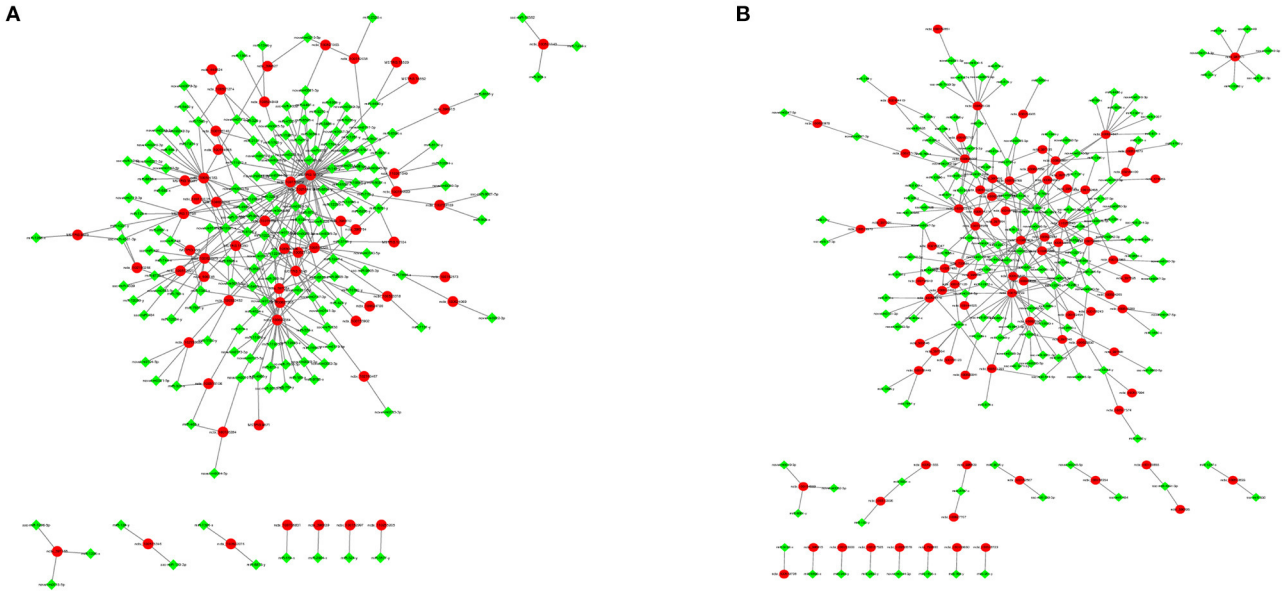


**FIGURE 8 |** Gene coexpression network analyses. Red nodes indicate hypoxic DEmRNAs, and green nodes indicate DEmiRNAs. **(A)** Gene coexpression network analyses of DEmRNAs and DEmiRNAs. **(B)** Gene coexpression network analyses of hypoxic DEmRNAs and DEmiRNAs.

oxidoreductase (POR) and glutathione reductase (GSR) gene expression by upregulating miR-214, which induces oxidative stress and plays a crucial role in adaptation to hypoxia (Zhou et al., 2013; Dong et al., 2014; Stefanetti et al., 2018; Li et al., 2019). *FOXO3* is a targeted gene of ssc-miR-214-3p, and the expression of *FOXO3* in the TH group was significantly higher than that in the TL group, but not significantly different between the LH and LL groups. Changes in ssc-miR-214-3p expression may inhibit the cell cycle and promote apoptosis, thereby inhibiting cell proliferation through *FOXO3* prolyl hydroxylation in hypoxic conditions. The regulation of the expression levels of *IGF2*, *MEGF6*, and *FOXO3* through miRNAs may lead to the better adaption of Tibetan pigs in hypoxic environments compared to the Landrace breed.

Collagens, such as *COL1A1*, *COL1A2*, and *COL3A1*, are widely represented in ECM–receptor interactions and focal adhesion pathways (Gelse et al., 2003), and their expression was significantly higher in the LH group than in the LL group but not significantly different between the TH and TL groups. All of these genes function as mechanoreceptors and may provide a force-transmitting physical link between the EMC and cytoskeleton, indicating that enhanced expression of *COL1A1*, *COL1A2*, and *COL3A1* may be another reason for the superior adaption to hypoxic conditions of TH. Our study revealed that high expression of fibroblast growth factors (such as *FGF1*, *FGF2*, *FGF9*) was higher in the native groups (TH and LL) than in the migrated groups (TL and LH), which was alleviated by activating *AKT3* (Pompura and Dominguez-Villar, 2018; Revathidevi and Munirajan, 2019). These findings indicated that Tibetan pigs may increase the expression of *FGF1* and the cross-sectional area of a blood vessel to increase blood flow in response to hypoxia (Karar and Maity, 2011; Kir et al., 2018; Sajib et al., 2018).

*GPR146* may be upregulated by a number of miRNAs (such as miR-8903, miR-11972 and miR-466-x) under hypoxic stimulation and has been suggested to be an important hypoxia-inducible gene in recent years. C-peptide inhibits low $O_2$-induced ATP release in erythrocytes as a putative ligand of *GPR146*, which was consistent with our results (Richards et al., 2014). Ncbi_397391 (*MMP2*), ncbi_102159047 (*FOXC1*), ncbi_100738910 (*PRRX1*), and ncbi_100520318 (*TUB*) are potentially regulated by novel-m0237-5p. In the present study, *MMP2* expression was significantly higher in the LH group than in the LL group, but no significant differences were found between the TH and TL groups. The expression of *MMP2* showed a similar tendency to the results of alveolar septum thickness among the four groups, indicating that *MMP2* activities may lead to the widening of the alveolar wall and septum as well as alveolar structure damage and collapse of the alveolar space with remarkable fibrosis in Landrace pigs (Tan et al., 2006).

## CONCLUSION

The comparisons between Tibetan pigs and Landrace pigs from high or low altitudes revealed genes and regulatory pathways with possible adaptive changes in response to high-altitude hypoxia. We identified several molecular pathways and hypoxia genes showing adaptive changes in the lung, including increased blood circulation and regulation of blood pressure and circulation

as well as regulation of HGB concentration and angiogenesis. Integrated analysis of mRNAs and miRNAs demonstrated that a number of hypoxia genes may be regulated by miRNAs and participate in the hypoxic regulation of the lung. For example, novel-m0237-5p may potentially upregulate the expression levels of *MMP*2, resulting in widened alveolar septum and alveolar structure damage. These results provide a better understanding of the molecular mechanisms regulating the hypoxia response in the lungs of Tibetan pigs and will help to prevent damage to the lungs caused by hypoxia.

## 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 Ministry of Science and Technology of the People's Republic of China (Approval number: 2006-398). Written informed consent was obtained from the owners for the participation of their animals in this study.

## AUTHOR CONTRIBUTIONS

SZ was the overall project leader who provided financial support and experimental conception. YY was involved in data analyses, statistical analyses, language revisions, journal selection, and manuscript submissions and revisions. HY, TJ, and TY contributed to the experimental design and implementation. CG contributed to the supervision and assistance of students in managing animals and collecting and analyzing samples. YL was responsible for the trial implementation, supervision of students collecting and analyzing samples, and manuscript preparation. YC contributed to supervision of sample collection and analysis and manuscript editing. All authors contributed to the article and approved the submitted version.

## FUNDING

This study was supported by the National Natural Science Foundation of China (31760644, 32060730).

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2021.691592/full#supplementary-material

## REFERENCES

Ai, H., Fang, X., Yang, B., Huang, Z., Chen, H., Mao, L., et al. (2015). Adaptation and possible ancient interspecies introgression in pigs identified by whole-genome sequencing. *Nat. Genet.* 47, 217–225. doi: 10.1038/ng.3199

Ban, S., Min, E., Ahn, Y., Popescu, G., and Jung, W. (2018). Effect of tissue staining in quantitative phase imaging. *J. Biophoton.* 11:e201700402. doi: 10.1002/jbio.201700402

Benson, D. A., Cavanaugh, M., Clark, K., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J., et al. (2013). GenBank. *Nucleic Acids Res.* 41, D36–D42. doi: 10.1093/nar/gks1195

Bhattarai, D., Xu, X., and Cee, K. (2018). Hypoxia-inducible factor-1 (HIF-1) inhibitors from the last decade (2007 to 2016): A "structure-activity relationship" perspective. *Med. Res. Rev.* 38, 1404–1442. doi: 10.1002/med.21477

Cai, L., Bai, H., Mahairaki, V., Gao, Y., He, C., Wen, Y., et al. (2018). A universal approach to correct various HBB gene mutations in human stem cells forgene therapy of beta-thalassemia and sickle cell disease. *Stem Cells Transl. Med.* 7, 87–97. doi: 10.1002/sctm.17-0066

Cao, X. F., Bai, Z. Z., Ma, L., Ma, S., and Ge, R. L. (2017). Metabolic alterations of Qinghai-Tibet Plateau pikas in adaptation to high altitude. *High Alt. Med. Biol.* 18, 219–225. doi: 10.1089/ham.2016.0147

Chen, S., Zhou, Y., Chen, Y., and Gu, J. (2018). fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* 34, i884–i890. doi: 10.1093/bioinformatics/bty560

Chen, T., Zhou, Q., Tang, H., Bozkanat, M., Yuan, J. X., Raj, J. U., et al. (2016). MiR-17 / 20 controls prolyl hydroxylase 2 (PHD2) / hypoxia-inducible factor 1 (HIF1) to regulate pulmonary artery smooth muscle cell proliferation. *J. Am. Heart Assoc.* 5:4510. doi: 10.1161/JAHA.116.004510

Crosby, M. E., Kulshreshtha, R., Ivan, M., and Glazer, P. M. (2009). MicroRNA regulation of DNA repair gene expression in hypoxic stress. *Cancer Res.* 69, 1221–1229. doi: 10.1158/0008-5472.CAN-08-2516

Dai, L., Lou, W., Zhu, J., Zhou, X., and Di, W. (2015). MiR-199a inhibits the angiogenic potential of endometrial stromal cells under hypoxia by targeting HIF-1α / VEGF pathway. *Int. J. Clin. Exp. Pathol.* 8, 4735–4744.

Danza, K., Silvestris, N., Simone, G., Signorile, M., Saragoni, L., Brunetti, O., et al. (2016). Role of miR-27a, miR-181a and miR-20b in gastric cancer hypoxia-induced chemoresistance. *Cancer Biol. Ther.* 17, 400–406. doi: 10.1080/15384047.2016.1139244

Dong, X., Liu, H., Chen, F., Li, D., and Zhao, Y. (2014). MiR-214 promotes the alcohol-induced oxidative stress *via* down-regulation of glutathione reductase and cytochrome P450 oxidoreductase in liver cells. *Alcohol Clin. Exp. Res.* 38, 68–77. doi: 10.1111/acer.12209

Ellis, H., and Ma, C. X. (2019). PI3K inhibitors in breast cancer therapy. *Curr. Oncol. Rep.* 21:110. doi: 10.1007/s11912-019-0846-7

Forkens, J. A., Sieuwerts, A. M., Smid, M., Look, M. P., De Weerd, V., Boersma, A. W., et al. (2008). Four miRNAs associated with aggressiveness of lymph node-negative, estrogen receptor-positive human breast cancer. *Proc. Natl. Acad. Sci. USA* 105, 13021–13026. doi: 10.1073/pnas.0803304105

Gaecía-Morales, V., Luaces-Regueira, M., and Campos-Toimil, M. (2017). The cAMP effectors PKA and Epac activate endothelial NO synthase through PI3K / Akt pathway in human endothelial cells. *Biochem. Pharmacol.* 145, 94–101. doi: 10.1016/j.bcp.2017.09.004

Ge, X., Liu, X., Lin, F., Li, P., Liu, K., Geng, R., et al. (2016). MicroRNA-421 regulated by HIF-1α promotes metastasis, inhibits apoptosis, and induces cisplatin resistance by targeting E-cadherin and caspase-3 in gastric cancer. *Oncotarget* 7, 24466–24482. doi: 10.18632/oncotarget.8228

Gelse, K., Pöschl, E., and Aigner, T. (2003). Collagens–structure, function, and biosynthesis. *Adv. Drug. Deliv. Rev.* 55, 1531–1546. doi: 10.1016/j.addr.2003.08.002

Griffiths-Jones, S., Bateman, A., Marshall, M., Khanna, A., and Eddy, S. R. (2003). Rfam: an RNA family database. *Nucleic Acids Res.* 31, 439–441. doi: 10.1093/nar/gkg006

Griffiths-Jones, S., Grocock, R. J., Van, D. S., Bateman, A., and Enright, A. J. (2006). miRBase: microRNA sequences, targets and gene nomenclature. *Nucleic Acids Res.* 34, D140–D144. doi: 10.1093/nar/gkj112

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

Hui, X., Al-Ward, H., Shaher, F., Liu, C. Y., and Liu, N. (2019). The role of miR-210 in the biological system: a current overview. *Hum. Hered.* 84, 233–239. doi: 10.1159/000509280

Jang, J. H., Seo, J. Y., Jang, J., Jung, C. W., Lee, K. O., Kim, S. H., et al. (2014). Hereditary gene mutations in korean patients with isolated erythrocytosis. *Ann. Hematol.* 93, 931–935. doi: 10.1007/s00277-014-2006-3

Karar, J., and Maity, A. (2011). PI3K / AKT / mTOR pathway in angiogenesis. *Front. Mol. Neurosci.* 4:51. doi: 10.3389/fnmol.2011.00051

Kim, D., Langmead, B., and Salzberg, S. L. (2015). HISAT: a fast spliced aligner with low memory requirements. *Nat. Methods* 12, 357–360. doi: 10.1038/nmeth.3317

Kir, D., Schnettler, E., Modi, S., and Ramakrishnan, S. (2018). Regulation of angiogenesis by microRNAs in cardiovascular diseases. *Angiogenesis* 21, 699–710. doi: 10.1007/s10456-018-9632-7

Lancuo, Z., Hou, G., Xu, C., Liu, Y., Zhu, Y., Wang, W., et al. (2019). Simulating the route of the tang-tibet ancient road for one branch of the silk road across the qinghai-tibet plateau. *PLoS ONE* 14:e0226970. doi: 10.1371/journal.pone.0226970

Langmead, B., and Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat Methods*. 9, 357–359. doi: 10.1038/nmeth.1923

Lee, Y., Kim, Y. J., Kim, M. H., and Kwak, J. M. (2016). MAPK Cascades in guard cell signal transduction. *Front. Plant Sci.* 7:80. doi: 10.3389/fpls.2016.00080

Li, L., Kang, H., Zhang, Q., D'agati, V. D., Al-Awoqti, Q., and Lin, F. (2019). FOXO3 activation in hypoxic tubules prevents chronic kidney disease. *J. Clin. Invest.* 129, 2374–2389. doi: 10.1172/JCI122256

Lin, C. J., Lan, Y. M., Ou, M. Q., Ji, L. Q., and Lin, S. D. (2019). Expression of miR-217 and HIF-1α / VEGF pathway in patients with diabetic foot ulcer and its effect on angiogenesis of diabetic foot ulcer rats. *J. Endocrinol. Invest.* 42, 1307–1317. doi: 10.1007/s40618-019-01053-2

Liu, X., Zhang, Y., Li, Y., Pan, J., Wang, D., Chen, W., et al. (2019). EPAS1 gain-of-function mutation contributes to high-altitude adaptation in Tibetan horses. *Mol. Biol. Evol.* 36, 2591–2603. doi: 10.1093/molbev/msz158

Livak, K. J., and Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. *Methods* 25, 402–408. doi: 10.1006/meth.2001.1262

Love, M. I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15, 550. doi: 10.1186/s13059-014-0550-8

Ma, Y. F., Han, X. M., Huang, C. P., Zhong, L., Adeola, A. C., Irwin, D. M., et al. (2019). Population genomics analysis revealed origin and high-altitude adaptation of Tibetan pigs. *Sci. Rep.* 9:11463. doi: 10.1038/s41598-019-47711-6

Marquard, F. E., and Jücker, M. (2020). PI3K / AKT / mTOR signaling as a molecular target in head and neck cancer. *Biochem. Pharmacol.* 172:113729. doi: 10.1016/j.bcp.2019.113729

Mohlin, S., Hamidian, A., and Påhlman, S. (2013). HIF2A and IGF2 expression correlates in human neuroblastoma cells and normal immature sympathetic neuroblasts. *Neoplasia* 15, 328–334. doi: 10.1593/neo.121706

Ni, L., Song, C., Wu, X., Zhao, X., Wang, X., Li, B., et al. (2019). RNA-seq transcriptome profiling of porcine lung from two pig breeds in response to mycoplasma hyopneumoniae infection. *Peer J.* 7:e7900. doi: 10.7717/peerj.7900

Ni, W. J., and Leng, X. M. (2016). MiRNA-dependent activation of mRNA translation. *Microrna* 5, 83–86. doi: 10.2174/2211536605666160825151201

Nicolas, S., Abdellatef, S., Haddad, M. A., Fakhoury, I., and El-Sabai, M. (2019). Hypoxia and EGF stimulation regulate VEGF expression in human glioblastoma multiforme (GBM) cells by differential regulation of the PI3K / Rho-GTPase and MAPK pathways. *Cells* 8:1397. doi: 10.3390/cells8111397

Pertea, M., Kim, D., Pertea, G. M., Leek, J. T., and Salzberg, S. L. (2016). Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. *Nat. Protoc.* 11, 1650–1667. doi: 10.1038/nprot.2016.095

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

Pompura, S. L., and Dominguez-Villar, M. (2018). The PI3K/AKT signaling pathway in regulatory T-cell development, stability, and function. *J. Leukoc Biol.* 22:349. doi: 10.1002/JLB.2MIR0817-349R

Qi, X., Zhang, Q., He, Y., Yang, L., Zhang, X., Shi, P., et al. (2019). The transcriptomic landscape of yaks reveals molecularpathways for high altitude adaptation. *Genome Biol. Evol.* 11, 72–85. doi: 10.1093/gbe/evy264

Revathidevi, S., and Munirajan, A. K. (2019). Akt in cancer: mediator and more. *Semin. Cancer Biol.* 59, 80–91. doi: 10.1016/j.semcancer.2019.06.002

Richards, J., Yosten, G. L., Kolar, G. R., Jones, C. W., Stephenson, A. H., Ellsworth, M. L., et al. (2014). Low O2-induced ATP release from erythrocytes of humans with type 2 diabetes is restored by physiological ratios of C-peptide and insulin. *Am. J. Physiol. Regul. Integr. Comp. Physiol.* 307, R862–R868. doi: 10.1152/ajpregu.00206.2014

Sajib, S., Zahra, F. T., Lionakis, M. S., German, N. A., and Mikelis, C. M. (2018). Mechanisms of angiogenesis in microbe-regulated inflammatory and neoplastic conditions. *Angiogenesis* 21, 1–14. doi: 10.1007/s10456-017-9583-4

Semenza, G. L. (2007). Hypoxia-inducible factor 1 (HIF-1) pathway. *Sci STKE* 407:cm8. doi: 10.1126/stke.4072007cm8

Stefanetti, R. J., Voisin, S., Russell, A., and Lamon, S. (2018). Recent advances in understanding the role of FOXO3. *Front. Genet.* 7:F1000. doi: 10.12688/f1000research.15258.1

Szklarczyk, D., Franceschini, A., Wyder, S., Forslund, K., Heller, D., Huerta-Cepas, J., et al. (2015). STRING v10: protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Res.* 43, D447–D452. doi: 10.1093/nar/gku1003

Taliercio, R. M., Ashton, R. W., Horwitz, L., Swanson, K. C., Wendt, P. C., Hoyer, J. D., et al. (2013). Hb Grove City [β38(C4)Thr→Ser, ACC>AGC; HBB: c.116C>G]: a new low oxygen affinity β chain variant. *Hemoglobin* 37, 396–403. doi: 10.3109/03630269.2013.789794

Tan, S. Z., Liu, C. H., Zhang, W., Lu, X., Ye, W. C., Cai, Z. Z., et al. (2006). Feature changes of MMP-2/9 activities and TIMP-1/2 protein expressions during the progression of pulmonary fibrosis in rats. *Zhong Xi Yi Jie He Xue Bao* 4, 402–407. doi: 10.3736/jcim20060417

Tang, T., Yang, Z., Zhu, Q., Wu, Y., Sun, K., Alahdal, M., et al. (2018). Up-regulation of miR-210 induced by a hypoxic microenvironment promotes breast cancer stem cells metastasis, proliferation, and self-renewal by targeting E-cadherin. *Faseb J.* 6:fj201801013R. doi: 10.1096/fj.201801013R

Wang, M., Wang, Y., Baloch, A. R., Pan, Y., Tian, L., Xu, F., et al. (2018). Detection and genetic characterization of porcine deltacoronavirus in Tibetan pigs surrounding the Qinghai-Tibet Plateau of China. *Transbound Emerg. Dis.* 65, 363–369. doi: 10.1111/tbed.12819

Wang, X., Li, J., Wu, D., Bu, X., and Qiao, Y. (2016). Hypoxia promotes apoptosis of neuronal cells through hypoxia-inducible factor-1α-microRNA-204-B-cell lymphoma-2 pathway. *Exp. Biol. Med.* 241, 177–183. doi: 10.1177/1535370215600548

Xie, Y., Li, W., Feng, J., Wu, T., and Li, J. (2016). MicroRNA-363 and GATA-1 are regulated by HIF-1α in K562 cells under hypoxia. *Mol. Med. Rep.* 14, 2503–2510. doi: 10.3892/mmr.2016.5578

Yang, Y., Gao, C., Yang, T., Sha, Y., Wang, X., Yang, Q., et al. (2021). Characteristics of Tibetan pig lung tissue in response to a hypoxic environment on the Qinghai–Tibet Plateau. *Arch. Anim. Breed.* 64, 283–292. doi: 10.5194/aab-64-283-2021

Yang, Z., Zhao, T. Z., Zou, Y. J., Zhang, J. H., and Feng, H. (2014). Hypoxia induces autophagic cell death through hypoxia-inducible factor 1α in microglia. *PLoS ONE* 9:e96509. doi: 10.1371/journal.pone.0096509

Yu, J., Liang, F., Huang, H., Pirttiniemi, P., and Yu, D. (2018). Effects of loading on chondrocyte hypoxia, HIF-1α and VEGF in the mandibular condylar cartilage of young rats. *Orthod. Craniofac. Res.* 21, 41–47. doi: 10.1111/ocr.12212

Yu, Y., Ma, L., Zhang, H., Sun, W., Zheng, L., Liu, C., et al. (2020). EPO could be regulated by HIF-1 and promote osteogenesis and accelerate bone repair. *Artif. Cells Nanomed. Biotechnol.* 48, 206–217. doi: 10.1080/21691401.2019.1699827

Zhang, B., Ban, D., Gou, X., Zhang, Y., Yang, L., Chamba, Y., et al. (2019a). Genome-wide DNA methylation profiles in Tibetan and Yorkshire

pigs under high-altitude hypoxia. *J. Anim. Sci. Biotechnol.* 10, 25–29. doi: 10.1186/s40104-019-0316-y

Zhang, B., Ban, D., Xiao, G., Zhang, Y., Yang, L., Yangzom, C. (2016). *Identification of plateau adaptation genes and pathways in Tibetan pigs by multiple expression omics of myocardial tissue*. Ph.D., China Agricultural University.

Zhang, B., Chamba, Y., Shang, P., Wang, Z., Ma, J., Wang, L., et al. (2017). Comparative transcriptomic and proteomic analyses provide insights into the key genes involved in high-altitude adaptation in the Tibetan pig. *Sci. Rep.* 7:3654. doi: 10.1038/s41598-017-03976-3

Zhang, B., Qiangba, Y., Shang, P., Wang, Z., Ma, J., Wang, L., et al. (2015). A Comprehensive microRNA expression profile related to hypoxia adaptation in the Tibetan Pig. *PLoS ONE* 10:e0143260. doi: 10.1371/journal.pone.0143260

Zhang, B., Zhang, X., Xu, H., Gao, X., Zhang, G., Zhang, H., et al. (2019b). Dynamic variation of RAS on silicotic fibrosis pathogenesis in rats. *Curr. Med. Sci.* 39, 551–559. doi: 10.1007/s11596-019-2073-8

Zhang, G., Yin, S., Mao, J., Liang, F., Zhao, C., Li, P., et al. (2016). Integrated analysis of mRNA-seq and miRNA-seq in the liver of pelteobagrus vachelli in response to hypoxia. *Sci. Rep.* 6:22907. doi: 10.1038/srep22907

Zhang, L., Li, Y., Wang, Q., Chen, Z., Li, X., Wu, Z., et al. (2020). The PI3K subunits, P110α and P110β are potential targets for overcoming P-gp and BCRP-mediated MDR in cancer. *Mol Cancer*, 19. doi: 10.1186/s12943-019-1112-1

Zhang, T., Suo, C., Zheng, C., and Zhang, H. (2019). Hypoxia and metabolism in metastasis. *Adv. Exp. Med. Biol.* 1136, 87–95. doi: 10.1007/978-3-030-12734-3_6

Zhang, X., Tang, N., Hadden, T. J., and Rishi, A. K. (2011). Akt, FoxO and regulation of apoptosis. *Biochim. Biophys. Acta* 1813, 1978–1986. doi: 10.1016/j.bbamcr.2011.03.010

Zhang, Z., Yao, L., Yang, J., Wang, Z., and Du, G. (2018). PI3K/Akt and HIF-1 signaling pathway in hypoxia-ischemia. *Mol. Med. Rep.* 18, 3547–3554. doi: 10.3892/mmr.2018.9375

Zhang, H., Wu, J., Wu, J., Fan, Q., Zhou, J., Wu, J., et al. (2019). Exosome-mediated targeted delivery of miR-210 for angiogenic therapy after cerebral ischemia in mice. *J. Nanobiotechnol.* 17:29. doi: 10.1186/s12951-019-0461-7

Zhi, Z., Yang, W., Liu, L., Jiang, X., and Pang, L. (2018). Early missed abortion is associated with villous angiogenesis *via* the HIF-1α/VEGF signaling pathway. *Arch. Gynecol. Obstet.*298, 537–543. doi: 10.1007/s00404-018-4802-9

Zhou, Y., Tang, X., Song, Q., Ji, Y., Wang, H., Wang, H., et al. (2013). Identification and characterization of pig embryo microRNAs by solexa sequencing. *Reprod. Domest. Anim.* 48, 112–120. doi: 10.1111/j.1439-0531.2012.02040.x

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

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Genetics

ORIGINAL RESEARCH
published: 08 October 2021
doi: 10.3389/fgene.2021.712306

# Integrative Systems Biology Analysis Elucidates Mastitis Disease Underlying Functional Modules in Dairy Cattle

*Nooshin Ghahramani*[1]*, *Jalil Shodja*[1], *Seyed Abbas Rafat*[1], *Bahman Panahi*[2] and *Karim Hasanpur*[1]

[1] Department of Animal Science, Faculty of Agriculture, University of Tabriz, Tabriz, Iran, [2] Department of Genomics, Branch for Northwest & West Region, Agricultural Biotechnology Research Institute of Iran (ABRII), Agricultural Research, Education and Extension Organization (AREEO), Tabriz, Iran

**Background:** Mastitis is the most prevalent disease in dairy cattle and one of the most significant bovine pathologies affecting milk production, animal health, and reproduction. In addition, mastitis is the most common, expensive, and contagious infection in the dairy industry.

**Methods:** A meta-analysis of microarray and RNA-seq data was conducted to identify candidate genes and functional modules associated with mastitis disease. The results were then applied to systems biology analysis *via* weighted gene coexpression network analysis (WGCNA), Gene Ontology, enrichment analysis for the Kyoto Encyclopedia of Genes and Genomes (KEGG), and modeling using machine-learning algorithms.

**Results:** Microarray and RNA-seq datasets were generated for 2,089 and 2,794 meta-genes, respectively. Between microarray and RNA-seq datasets, a total of 360 meta-genes were found that were significantly enriched as "peroxisome," "NOD-like receptor signaling pathway," "IL-17 signaling pathway," and "TNF signaling pathway" KEGG pathways. The turquoise module ($n = 214$ genes) and the brown module ($n = 57$ genes) were identified as critical functional modules associated with mastitis through WGCNA. *PRDX5*, *RAB5C*, *ACTN4*, *SLC25A16*, *MAPK6*, *CD53*, *NCKAP1L*, *ARHGEF2*, *COL9A1*, and *PTPRC* genes were detected as hub genes in identified functional modules. Finally, using attribute weighting and machine-learning methods, hub genes that are sufficiently informative in *Escherichia coli* mastitis were used to optimize predictive models. The constructed model proposed the optimal approach for the meta-genes and validated several high-ranked genes as biomarkers for *E. coli* mastitis using the decision tree (DT) method.

**Conclusion:** The candidate genes and pathways proposed in this study may shed new light on the underlying molecular mechanisms of mastitis disease and suggest new approaches for diagnosing and treating *E. coli* mastitis in dairy cattle.

**Keywords: attribute weighting, *E. coli*, machine learning, mastitis, meta-analysis, WGCNA**

OPEN ACCESS

***Edited by:***
*Aline Silva Mello Cesar, University of São Paulo, Brazil*

***Reviewed by:***
*Claas Heuer, ST Genetics, United States*
*Andressa Oliveira De Lima, University of Washington, United States*

***Correspondence:**
*Nooshin Ghahramani*
*gh.nooshin69@gmail.com*

***Specialty section:***
*This article was submitted to Livestock Genomics, a section of the journal Frontiers in Genetics*

***Received:*** *20 May 2021*
***Accepted:*** *30 August 2021*
***Published:*** *08 October 2021*

***Citation:***
*Ghahramani N, Shodja J, Rafat SA, Panahi B and Hasanpur K (2021) Integrative Systems Biology Analysis Elucidates Mastitis Disease Underlying Functional Modules in Dairy Cattle. Front. Genet. 12:712306. doi: 10.3389/fgene.2021.712306*

## BACKGROUND

Mastitis is a severe disease characterized as an inflammatory condition affecting the mammary glands (Gelasakis et al., 2015). *Escherichia coli*, *Staphylococcus aureus*, and *Streptococcus uberis* are the three major bacterial pathogens associated with mastitis disease (Vasudevan et al., 2003), with *E. coli* causing severe inflammation in dairy cattle (Vangroenweghe et al., 2005). The focus of current research has shifted to elucidating the underlying mechanisms and developing effective treatment strategies for mastitis disease (Takeshima et al., 2008; Compton et al., 2009). *E. coli* typically infects the mammary glands during the dry period, and inflammation progresses during the early stages of lactation (Burvenich et al., 2003). Recent research indicates that *E. coli's* pathogenicity is entirely dependent on a protein called FecA (Blum et al., 2018).

Recent advancements in high-throughput transcriptome profiling technologies, such as microarray and RNA sequencing (RNA-seq), have enabled opportunities for precision critical care medicine to understand better the molecular mechanisms underlying diverse biological functions (Bansal and Di Bernardo, 2007; Farhadian et al., 2020; Panahi et al., 2020). On the other hand, identifying disease biomarkers can aid breeders in optimizing their genetic programs for dairy animals (Kulkarni and Kaliwal, 2013; Duarte et al., 2015; Lai et al., 2017). Previous research identified *TNF*- and *SAA3* (Swanson et al., 2009), *STAT3*, *MAPK14*, *TNF* (Gorji et al., 2019), *IL8RB*, *CXCL6*, *MMP9* (Li et al., 2019), *IRF9*, *CCL* (Buitenhuis et al., 2011), *S100A12*, *MT2A*, *SOD2* (Mitterhuemer et al., 2010), *CXCL8*, *CXCL2*, S100A9 (Sharifi et al., 2018), *PSMA6*, *HCK*, and *STAT1* (Bakhtiarizadeh et al., 2020) as potential biomarkers for mastitis disease.

Meta-analysis is a systematic and quantitative study methodology used to evaluate prior research and reach a conclusion (Haidich, 2010). On the other hand, independent studies have limitations in sample size, statistical power, and the reliability of the results (Panahi and Hejazi, 2021). Meta-analysis has demonstrated that combining *p* values resolves several issues (Rhodes et al., 2002; Tseng et al., 2012; Panahi et al., 2019a). When combining *p* values using Fisher's technique, the null hypothesis is that the actual effect is zero in each of the combined datasets (Evangelou and Ioannidis, 2013), suggesting that the techniques should be sensitive even when only a subset of the combined datasets has an impact magnitude more significant than zero. Fisher's approach outperformed other methods for establishing associations. In addition, the *p* value combination method shows considerable promise for identifying novel markers or differentially expressed genes (DEGs) (Evangelou and Ioannidis, 2013). Moreover, connectivity analysis of known meta-genes has been presented as a promising approach for disentangling the complicated method (Panahi et al., 2020).

Weighted gene coexpression network analysis (WGCNA) has been proposed as a versatile tool for gene coexpression analysis, which provide valuable information about gene connectivity based on gene expression levels (Ebrahimie et al., 2014; Farhadian et al., 2021). A combination of machine-learning algorithms and microarray meta-analysis was used to identify mastitis genes in dairy cattle (Sharifi et al., 2018), However, they did not include RNA-seq data in their analysis and instead focused on the expression patterns of meta-genes.

The present study is the first that the authors are aware of that integrates meta-analysis of microarray and RNA-seq datasets, connectivity analysis, and model optimization in mastitis disease. Thus, in this integrative study, we identified master genes associated with mastitis disease using a combination of meta-analysis, WGCNA, and machine-learning algorithms.

## MATERIALS AND METHODS

### Data Collection

The National Center for Biotechnology Information's Gene Expression Omnibus (GEO) repository (https://www.ncbi.nlm.nih.gov/gds/) was explored for datasets related to dairy cattle mastitis. This database was searched for RNA-seq and microarray studies using the keywords "*Bos taurus*," "mastitis," and "*Escherichia coli*." For this research, six microarrays and two RNA-seq datasets were chosen. **Table 1** lists the platform, accession number, species, and references for each dataset. All healthy and mastitis animal samples were from the *Bos taurus* species, which have a high sensitivity to *E. coli*. Fifteen healthy German Holstein Frisian cows in midlactation (3–6 months postpartum) were included in dataset GSE15025. The animals were inoculated with *E. coli* in one-quarter and died after 6 h ($n = 5$) or 24 h ($n = 5$) in two distinct infection scenarios. Five heifers were used as controls; they were not given any medication and died after 24 h. At 4 to 6 weeks following parturition, 16 healthy primiparous Danish Holstein-Friesian cows were tested with *E. coli* for the GSE24217 dataset. The overall udder health of 24 dairy cows was assessed before the disease challenge. Control quarters were selected based on bacteriological tests performed before *E. coli* inoculation and the quarter foremilk SCC at 24 and 192 h. From the total German Holstein population, 11 heifers at day 42 postpartum, either with high or low sensitivity to mastitis, were chosen for the GSE24560 dataset. Heat-inactivated *E. coli* plus *S. aureus* was used to challenge the cells, as a control. The cells were collected after 1, 6, and 24 h, and mRNA expression was compared. Four first-lactation Holstein cows in the fourth month of lactation were also experimentally inoculated with the mastitis-causing *E. coli* pathogen strain 1303 for the GSE25413 dataset. The transcriptomes of the treated and untreated cells were examined at 1, 3, 6, and 24 h. In the GSE32186 dataset, four first-lactation Holstein cows were given primary MEC (pbMEC) cultures for 6 h, and some cultures were stimulated. *E. coli* particles were collected from the udders of three healthy, pregnant (day 130 of gestation) cows in the middle of their first lactation 12 or 42 h later. Six Holstein Friesian cows were challenged with *E. coli* mastitis for the GSE50685 investigations. Every 6 h after infection, blood and milk samples were taken. At successive milkings, the treatment was repeated (12, 24, and 36 h postchallenge). At 24 h ($n = 3$) and 48 h ($n = 3$) following infection, the cows were sacrificed for tissue collection. GSE75379 and GSE159286 were two datasets related to RNA-seq. Sixteen healthy primiparous Holstein cows at 4–6 weeks of lactation were included in the GSE75379 dataset. Biopsy specimens of healthy and diseased udder tissue were taken at

**TABLE 1 |** Selected microarray and RNA-seq datasets for systems biology analysis of mastitis disease.

| Accession no. | Species | Bacteria | Platform | Samples* (C:T) | References |
|---|---|---|---|---|---|
| **Microarray datasets** | | | | | |
| GSE15025 | *Bos taurus* | *Escherichia coli* | Affymetrix | 15:15 | Mitterhuemer et al., 2010 |
| GSE24217 | *B. taurus* | *E. coli* | Affymetrix | 23:26 | Buitenhuis et al., 2011 |
| GSE24560 | *B. taurus* | *E. coli* | Affymetrix | 27:31 | Brand et al., 2011 |
| GSE25413 | *B. taurus* | *E. coli* | Affymetrix | 6:24 | Günther et al., 2011 |
| GSE32186 | *B. taurus* | *E. coli* | Affymetrix | 12:12 | Günther et al., 2012 |
| GSE50685 | *B. taurus* | *E. coli* | Affymetrix | 5:15 | Sipka et al., 2014 |
| **RNA-seq datasets** | | | | | |
| GSE75379 | *B. taurus* | *E. coli* | Illumina | 6:12 | Moyes et al., 2016 |
| GSE159286 | *B. taurus* | *E. coli* | Illumina | 53:52 | Cebron et al., 2020 |

**Number of healthy and infected samples.*

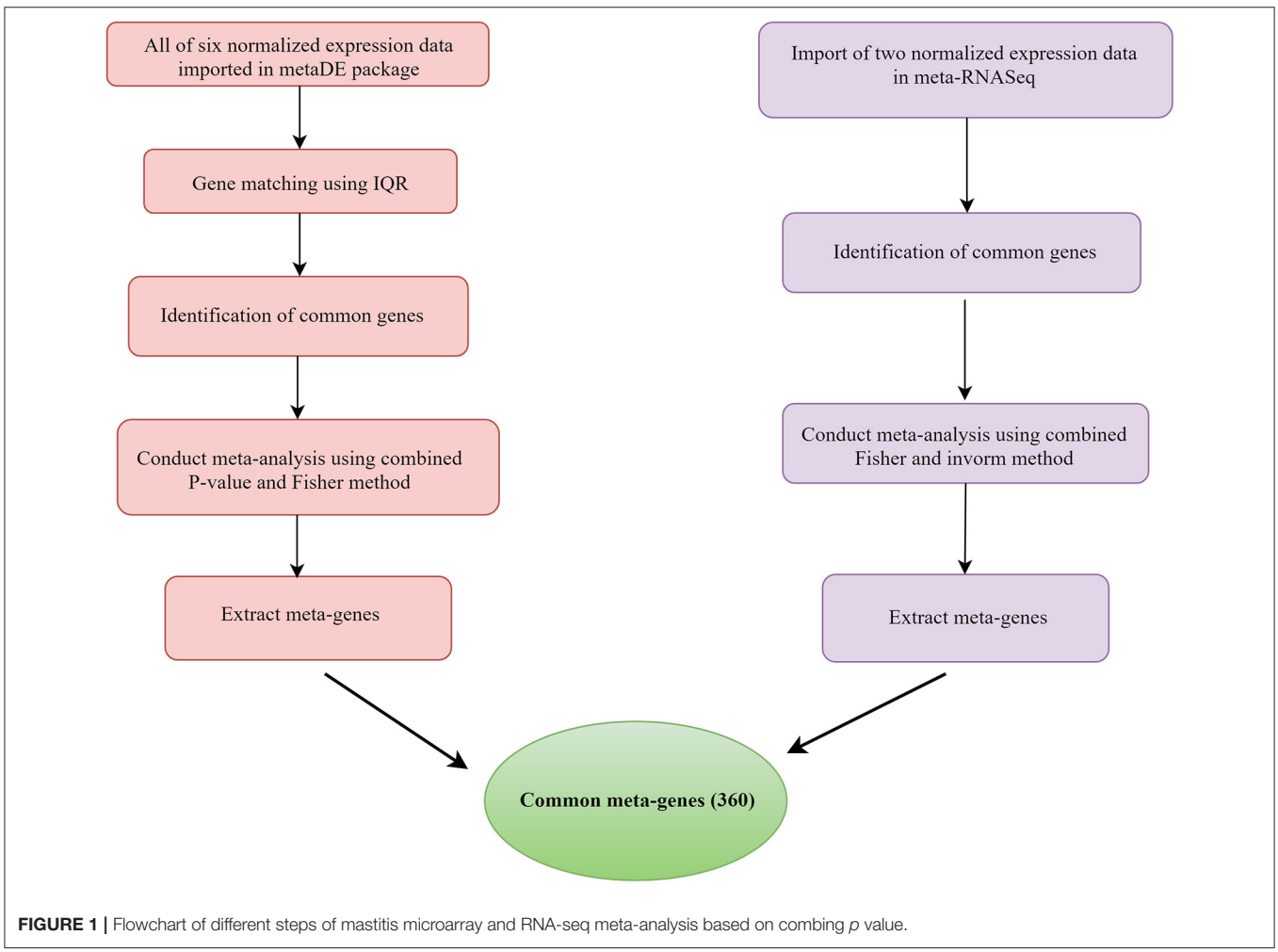


**FIGURE 1 |** Flowchart of different steps of mastitis microarray and RNA-seq meta-analysis based on combing *p* value.

$T = 24$ h and $T = 192$ h after infection. In total, 12 heifers were intramuscularly vaccinated with heat-killed *E. coli* for the GSE159286 investigations. Half of the heifers (IM group, n = 6) received a booster injection 2 months later. Others (IMM group, n = 6) received 50 g of protein concentrate produced from *E. coli* in two quarters. Cows were then challenged with an *E. coli* P4 bacterial suspension infusion at 49 days in milk inside one healthy quarter (10e3 bacteria). Before the trial, blood was taken 7 days after the first and second injections (immunization phase) and then at 0, 12, and 40 h following infection (challenge phase).

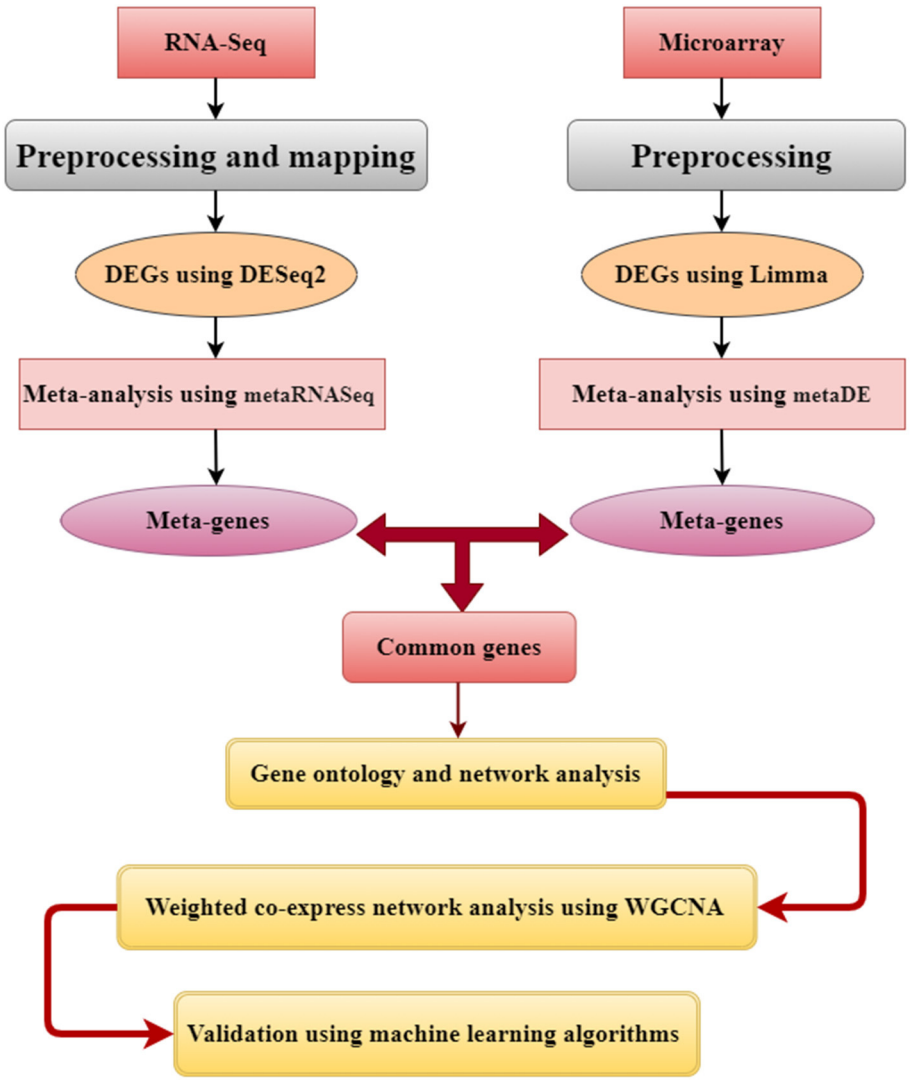


**FIGURE 2 |** Flowchart of applied systems biology approach in this study.

## Preprocessing and Analysis of Microarray Datasets

The GEO database was used to collect all microarray expression raw data and associated annotations for each study, and microarray data were preprocessed to obtain reliable findings. Nonbiological data variances were then removed, and appropriate scales were used for further analysis (Bolstad et al., 2003). The quantile normalization method and batch effects reduction were used to conduct effective gene expression analysis and eliminate variability between studies. The Limma software (2.16.0) (Smyth et al., 2005) was used to calculate DEGs among each control and treatment group after preprocessing the raw data. DEGs were deemed significant when the false discovery rate (FDR) using the Benjamini–Hochberg method was $p < 0.05$ and the logarithm of fold change $> \pm 0.5$ (Benjamini and Hochberg, 1995).

## Preprocessing and Analysis of Individual RNA-seq Datasets

The data generated by RNA-seq can be skewed due to biases introduced during library preparation, polymerase chain reaction, and sequencing. The technique of trimmed mean of m-values was used to eliminate the effect of known nonbiological features on the RNA-seq data (Robinson and Oshlack, 2010). Each sample was inspected for quality using the FastQC tool version 0.11.5 (Trapnell et al., 2012), and low-quality reads were trimmed using the Trimmomatic (v 0.32) software (Goldman et al., 2006). Bowtie2 (2.2.4) software was used to index reference genomes, and clean reads were then mapped to the *B. taurus* reference genome (ARS-UCD1.2 version) employing Tophat2 (2.0.10) software (Kim et al., 2013; Love et al., 2014) with default settings. The sample mapping rates are listed in **Supplementary Material 1**, **Table 1**. The htseq-count package

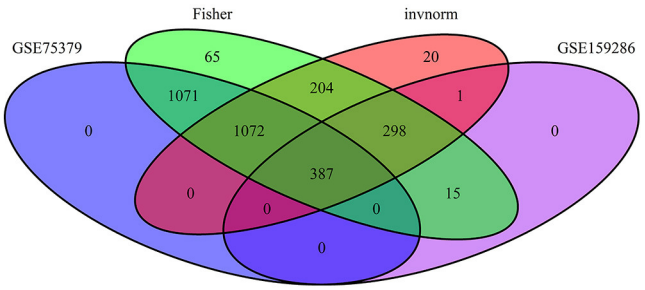


**FIGURE 3 |** Results of meta-analysis of RNA-seq data using Fisher and invorm method.

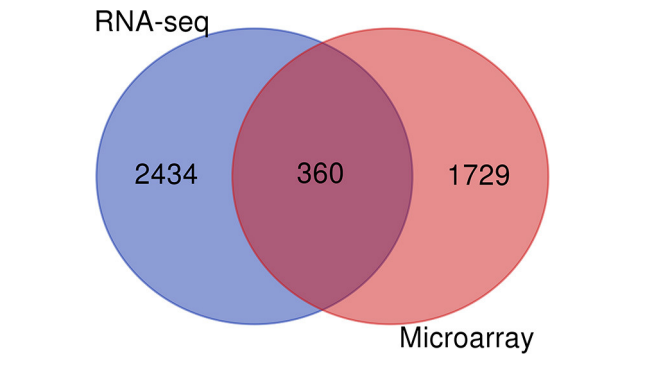


**FIGURE 4 |** Results of common meta-genes using Venn diagram.

(2.7.3) (Anders et al., 2013) was used to calculate the expression count matrix. The Bioconductor DESeq2 software (1.10.1) was used to determine the differential gene expressions in each research (Love et al., 2014). In terms of normalization and batch effect correction, the methods outlined in the studies (Farhadian et al., 2021; Panahi and Hejazi, 2021) were followed. In summary, genes with a CV <10% were initially removed. The group was then used as a covariate, using DESeq2 library size, size factor normalization factors. The variance-stabilizing transformations (VSTs) function was used to reduce sample disparities. The VST function does not typically remove variations related to the batch or other variables. As a result, the "removeBatchEffect" function was used to remove batch variations. The blind = false option was selected as re-estimation of the dispersion values was not required. This process leveled the library size and other normalization variables. Each study's samples were normalized jointly, implying that each dataset was normalized individually (Love et al., 2014).

## Meta-Analysis of Microarray and RNA-seq Datasets

In microarray studies, the MetaDE package (1.0.5) was used to identify meta-genes (Wang et al., 2012). The meta-analysis included the following steps: after quantile normalization, labeling samples as "Infected" or "Healthy"; the "Gene Symbol" was matched to multiple probe IDs using the interquartile range for probe selection (Wang et al., 2012). Merging genes is an approach used to determine which genes should be studied further (Wang et al., 2021). Because the number of genes in the research varied, the multiple gene expression datasets may not have been adequately matched by genes. In this study, the direct merging method was used to obtain common genes across different investigations. The Fisher technique (Marot et al., 2020) was used to identify meta-gens in RNA-seq data using the metaRNASeq software (1.0.5). Initially, the DEGs for each study were defined using the DESeq2 package (1.30.1), and the corresponding $p$ value was extracted. Then, the fishcomb function included in the metaRNASeq package was used to combine the $p$ values. For downstream analysis, the genes shared between meta-analyses of microarray and RNA-seq data were extracted (**Figure 1**).

## Weighted Gene Coexpression Network Analysis

Common genes were used to construct coexpression networks using the WGCNA Bioconductor R package (version 3.5.1) to understand the correlation patterns among genes and identify significant modules associated with mastitis disease (Langfelder and Horvath, 2008). Initially, networks were constructed using Pearson correlations across all common genes (Botía et al., 2017). A soft threshold was used to evaluate the correlation and noise-filtering power to fulfill the scale-free topology requirement. The weighted gene coexpression network design promotes strong correlations at the cost of low correlations by increasing the absolute magnitude of the correlation to a soft threshold (Langfelder and Horvath, 2008). The topological overlap matrix (TOM) and its corresponding dissimilarity matrix (1–TOM) were used to visualize the network, which resulted in a network diagram for module detection. Module eigengenes and module membership were used to identify the hub genes for each significant coexpressed module (Langfelder and Horvath, 2008). The following parameters were used to construct the modules: cut height of 0.975, minimum module size of 30 genes, "hybrid" method, and deepSplit = 2.

## Functional Enrichment Analysis

The STRING (Szklarczyk et al., 2015) database was used to conduct enrichment analysis on the Kyoto Encyclopedia of Genes and Genomes (KEGG) and Gene Ontology (GO) (Dennis et al., 2003). The FDR (<0.05) correction was used to determine the statistical significance of GO and KEGG terms.

## Protein–Protein Interaction Networks of Common Genes

Gene network analysis of protein–protein interaction between common genes was performed using Cytoscape software to visualize gene networks and identify hub genes. Hub genes are defined as those with the highest degree of connectivity and those with a greater biological significance than other gene members (Shannon et al., 2003).

**TABLE 2 |** Significant common meta-genes in mastitis disease.

| Gene symbol | Full name | Chromosome | Compartment |
|---|---|---|---|
| *ELMO2* | Engulfment and cell motility 2 | 20 | Cytosol |
| *ORM1* | Orosomucoid 1 (alpha-1-acid glycoprotein) | 9 | Extracellular |
| *ABCB7* | ATP-binding cassette subfamily B member 7 | X | Mitochondrion |
| *LRRC41* | Leucine-rich repeat containing 41 | 1 | Nucleus |
| *CXCL3* | C-X-C motif chemokine ligand 3 | 4 | Extracellular |
| *SOD2* | Superoxide dismutase 2 | 6 | Extracellular |
| *ZFYVE1* | Zinc finger FYVE-type containing 1 | 14 | Mitochondrion |
| *SORL1* | Sortilin-related receptor 1 | 11 | Plasma membrane |
| *PTCD3* | Pentatricopeptide repeat domain 3 | 2 | Mitochondrion |
| *RPS6KA5* | Ribosomal porotein S6 kinase A5 | 14 | Nucleus |
| *LHFPL2* | LHFPL tetraspan subfamily member 2 | 5 | Plasma membrane |
| *TRIQK* | Triple QxxK/R motif containing | 8 | Endoplasmic reticulum |
| *MAOA* | Monoamine oxidase A | X | Mitochondrion |
| *CORO2A* | Coronin 2A | 9 | Cytosol |
| *TPM3* | Tropomyosin 3 | 1 | Extracellular |
| *PTPRC* | Protein tyrosine phosphatase receptor type C | 1 | Plasma membrane |
| *LOC407171* | Fc gamma 2 receptor | 18 | Extracellular |

## Supervised Machine-Learning Models

The common meta-genes identified were utilized to select features using 10 different weighting algorithms, including information gain, information gain ratio, $\chi^2$, deviation, rule, support vector machine, Gini index, uncertainty, relief, and PCA to validate the hub genes' efficacy in distinguishing different genes involved in mastitis disease (Farhadian et al., 2018a, 2021). The Rapid Miner software (Rapid Miner 5.0.001, Dortmund, Germany) was used for attribute weighting (Ebrahimi et al., 2011; Farhadian et al., 2018a; Panahi et al., 2019a; Nami et al., 2021). The primary objective of attribute weighting algorithms was to extract a subset of input features (genes) by excluding those that contained little or no information (Panahi et al., 2019b). The decision trees (DTs) were constructed using features with weighting values greater than 0.5. The DTs were constructed using the following methods: information gain, information gain ratio, Gini index, and accuracy criteria. **Figure 2** depicts the flowchart of an analytical strategy for microarray and RNA-seq.

# RESULTS

## Meta-Analysis

We conducted a meta-analysis of DEGs using data from microarray and RNA-seq experiments. Six raw microarray datasets containing 211 samples and two RNA-seq datasets containing 123 independent dairy cattle experiments were chosen separately for the meta-analysis. Finally, a total of 2,089 and 2,794 meta-genes in response to *E. coli* mastitis in microarray and RNA-seq data, respectively, were observed using the Fisher method in the metaDE and metaRNASeq packages. **Supplementary Material 2** and **Figure 3** contain the results of the meta-analysis of RNA-seq data.

## Identification of Common Genes by Meta-Gene Comparison

A total of 360 genes were identified as common meta-genes in meta-analysis of microarray and RNA-seq data (**Figure 4**). **Table 2** and **Supplementary Material 3** contain additional information about the significant common meta-genes.

## Functional Enrichment Analysis of Common Genes

The STRING database was used to conduct GO analyses on 360 common meta-genes to ascertain their biological process (BP), molecular function (MF), and cellular component (CC) roles in mastitis disease. The results found 170, 33, and 36 GO terms for BPs, MFs, and CCs, respectively. The terms "cellular process," "response to stimulus," "biological regulation," "regulation of a biological process," and "regulation of a cellular process" were used to denote the most critical process in the BP category. "Binding," "ion binding," "actin binding," "cation binding," and "metal ion binding" were all significantly overrepresented in the MF category. In terms of CC, the terms "intracellular," "cell," "cytoplasm," "intracellular organelle," and "organelle" were significantly enriched. Additional information is available in **Table 3** and **Supplementary Material 4**.

This analysis identified a total of nine significant KEGG pathways. In addition, the results indicated that the "peroxisome," "NOD-like receptor signaling pathway," "IL-17 signaling pathway," and "TNF signaling pathway" were significantly overrepresented. **Table 4** contains additional information about KEGG pathways.

Cytoscape demonstrates the involvement of DEGs in protein–protein interaction. **Figure 5** illustrates the gene network

**TABLE 3 |** Significant GO terms of common genes.

| Term ID | Description | GO terms | *p* value |
|---|---|---|---|
| **GO:0009987** | Cellular process | BP | 1.52E-10 |
| **GO:0050896** | Response to stimulus | BP | 1.96E-10 |
| **GO:0065007** | Biological regulation | BP | 8.53E-10 |
| **GO:0050789** | regulation of BP | BP | 3.66E-09 |
| **GO:0050794** | Regulation of cellular process | BP | 3.66E-09 |
| **GO:0051716** | Cellular response to stimulus | BP | 6.09E-08 |
| **GO:0019222** | Regulation of metabolic process | BP | 2.26E-07 |
| **GO:0051171** | Regulation of nitrogen compound metabolic process | BP | 2.26E-07 |
| **GO:0080090** | Regulation of primary metabolic process | BP | 2.26E-07 |
| **GO:0031323** | Regulation of cellular metabolic process | BP | 2.63E-07 |
| **GO:0005488** | Binding | MF | 3.90E-10 |
| **GO:0043167** | Ion binding | MF | 2.55E-06 |
| **GO:0005515** | Protein binding | MF | 1.30E-05 |
| **GO:0003779** | Actin binding | MF | 0.00011 |
| **GO:0043169** | Cation binding | MF | 0.00024 |
| **GO:0046872** | Metal ion binding | MF | 0.00029 |
| **GO:0008092** | Cytoskeletal protein binding | MF | 0.00034 |
| **GO:1901363** | Heterocyclic compound binding | MF | 0.0014 |
| **GO:0097159** | Organic cyclic compound binding | MF | 0.0019 |
| **GO:0036094** | Small molecule binding | MF | 0.002 |
| **GO:0005622** | Intracellular | CC | 7.27E-13 |
| **GO:0005623** | Cell | CC | 8.74E-13 |
| **GO:0005737** | Cytoplasm | CC | 8.74E-13 |
| **GO:0043229** | Intracellular organelle | CC | 1.39E-09 |
| **GO:0043226** | Organelle | CC | 1.74E-09 |
| **GO:0043227** | Membrane-bound organelle | CC | 5.37E-09 |
| **GO:0043231** | Intracellular membrane-bound organelle | CC | 7.48E-09 |
| **GO:0005829** | Cytosol | CC | 2.03E-07 |
| **GO:0070013** | Intracellular organelle lumen | CC | 1.28E-05 |
| **GO:0005634** | Nucleus | CC | 3.30E-05 |

*Only the significantly enriched ($p < 0.05$) GO terms are presented.*

**TABLE 4 |** The significant KEGG metabolic pathways associated with the common genes.

| Pathway name | *p* value | Total genes in pathway | Strength |
|---|---|---|---|
| Peroxisome | 0.0012 | 10 | 0.84 |
| NOD-like receptor signaling pathway | 0.0058 | 12 | 0.61 |
| IL-17 signaling pathway | 0.0058 | 9 | 0.76 |
| TNF signaling pathway | 0.008 | 9 | 0.68 |
| Salmonella infection | 0.008 | 8 | 0.76 |
| Viral carcinogenesis | 0.008 | 13 | 0.55 |
| Human papillomavirus infection | 0.0089 | 16 | 0.46 |
| Necroptosis | 0.0194 | 10 | 0.56 |
| Autophagy—animal | 0.0216 | 9 | 0.59 |

*TNF, tumor necrosis factor.*

visualization of common meta-genes. The top genes were STAT1, RTPRC, SOD2, and VCP (**Supplementary Material 5**).

## Weighted Gene Coexpression Network Construction

A WGCNA was performed to identify genes with a high correlation and classified the common genes into four modules. The turquoise module ($n = 214$ genes) and the brown module ($n = 57$ genes) were identified as critical functional modules associated with mastitis through WGCNA analysis (**Figure 6A**). The remaining modules, such as the blue module ($n = 84$ genes) plus the gray module ($n = 5$ genes), were not notable. **Figure 6B** illustrates the hierarchical clustering of common genes.

The correlation coefficient and $p$ value for the significant modules in the mastitis and healthy groups were $r = 0.28$, $p = 0.002$, and $r = 0.36$, $p = 4e-05$, respectively, for the turquoise and brown modules (**Figure 7**).

Turquoise had a negative correlation with mastitis disease, whereas brown had a positive correlation. **Table 5** lists the top five hub genes in brown and turquoise modules. **Supplementary Material 6** contains a list of the more significant modules identified.

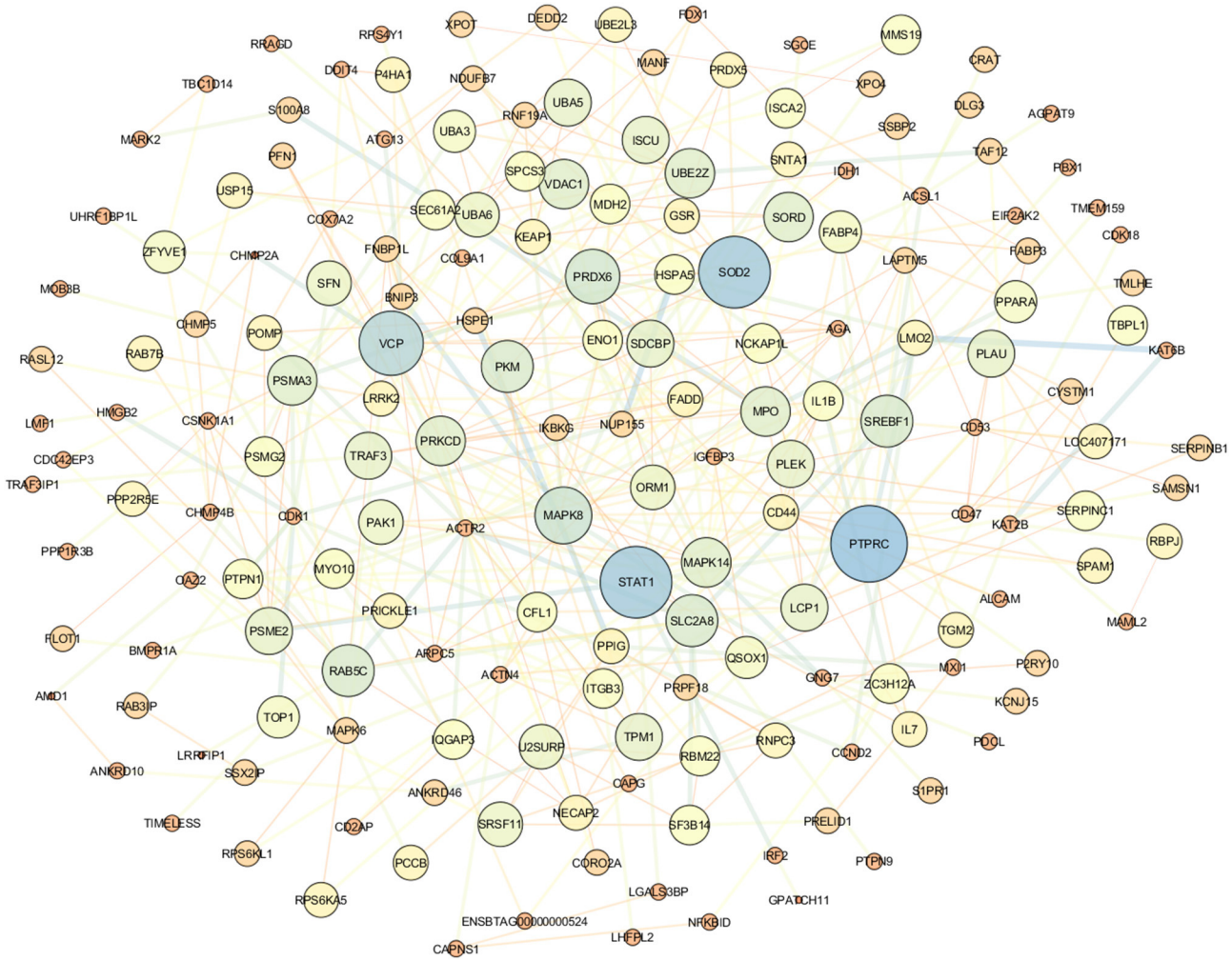


**FIGURE 5 |** Protein–protein interaction network for the common genes using Cytoscape.

"Peroxisome," "viral carcinogenesis," and "arginine and proline metabolism" were determined as the most significantly enriched pathways based on the enriched functional analysis in these modules that were potentially associated with mastitis development. These modules enriched for genes involved in "negative regulation of peptidyl-serine phosphorylation," "response to stimulus," "cell process regulation," "protein hydroxylation," "actin filament-based process," and "cellular process." These modules carried out critical MFs such as "actin binding," "binding," "transcription factor binding," and "peroxidase activity." "Intracellular," "organelle," "cytoplasm," "cell," "cortical actin cytoskeleton," and "microvillus" were identified as CCs.

## Attribute Weighting

The data-cleaning process was used to eliminate redundant and highly correlated (>95%) attributes. Finally, modeling was performed on the 360 genes. If an attribute was assigned a weight >0.5 by a specific attribute weighting algorithm, it was considered essential. **Supplementary Material 7** contains the results of 10 different attribute weighting algorithm applications. **Table 6** summarizes the number of attribute weighting algorithms that supported the selected DEGs.

## Validation Hub Genes in Coexpressed Modules

The DT technique was used to validate the identified hub genes. Thus, the accuracy of various models was calculated and presented in **Supplementary Material 8** using four different criteria, namely, information gain ratio, information gain, Gini index, and accuracy. According to the results, the DT with the gain ratio criterion achieved the highest accuracy (75%) (**Table 7**). The DT validated the role of the top-ranked genes in mastitis classification using the expression values of common meta-genes.

As illustrated in **Figure 8**, because the *LOC407171* gene is located at the root of the constructed tree, it can be considered a biomarker for mastitis. When the *LOC407171* gene value

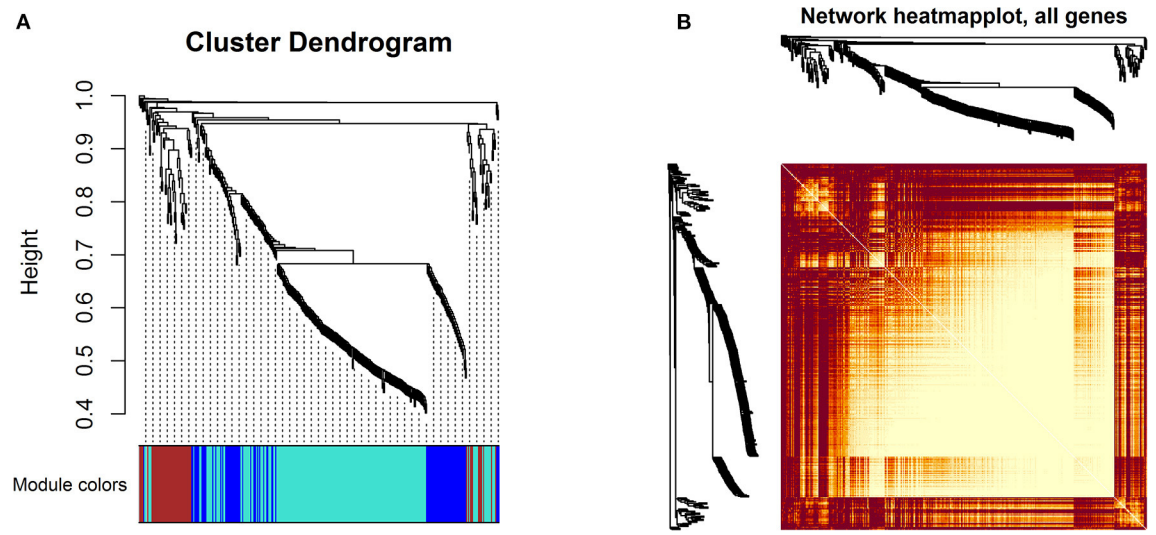


**FIGURE 6 |** WGCNA: **(A)** Cluster dendrogram of the common genes. The branches and color bands demonstrate the specific module. **(B)** TOM plot; light color symbolizes low overlap, and progressively darker red color symbolizes higher overlap between common genes. Blocks of darker colors along the diagonal correspond to modules.

exceeded 8.119, and the *SFN* gene value exceeded 5.291, the samples were classified as having mastitis. When the *LOC407171* gene value is equal to or <8.119, the sample is considered healthy. When *LOC407171* exceeded 8.119, SFN was equal to or <5.291, and *PTPRC* was equal to or <14.390, the sample was classified as healthy. In addition, if the last feature exceeded 14.390 and the expression of *IDH1* was present, *PTPRC* would be classified as having mastitis.

The significance of the *LOC407171*, *PTPRC*, *ABCG2*, and *IDH1* genes in the turquoise module was confirmed using DT models and attribute weighting, highlighting their critical roles in mastitis.

## DISCUSSION

Mastitis is a significant disease involving multiple genes that may interact to enrich specific signaling pathways. We performed a meta-analysis of RNA-seq and microarray transcriptome data to gain a comprehensive understanding of the master/key genes during mastitis disease that may play a significant role in response to *E. coli* mastitis. As individual studies have limitations in statistical power and reproducibility, several small impact genes remain unknown. Meta-analysis has been suggested as a practical approach for resolving this issue (Farhadian et al., 2018b; Sharifi et al., 2018). The BP, biological regulation, and reaction to a stimulus, the study's primary enriched GO terms, have been described as BPs in mastitis disease (Asselstine et al., 2019). These words include various activities, including cell proliferation, cell growth, biochemical processes, and signaling pathways (Long et al., 2001; Arnellos, 2018). The terminology used to describe MF in this research, such as binding, ion binding, protein binding, actin binding, cation binding, and catalytic activity, has been previously described in immune response and protein transport (Swanson et al., 2009; Asselstine et al., 2019).

Enrichment analysis of metabolic KEGG pathways was used to identify metabolic pathways that were significantly overrepresented among 360 common genes. Several significant pathways were enriched, including peroxisomes and three subcategories of signaling pathways [interleukin 17 (IL-17) signaling pathway, nucleotide-binding and oligomerization domain (NOD)–like receptor signaling pathway, TNF signaling pathway]. Peroxisomes are required to oxidize specific biomolecules and the inflammatory response to environmental stress (Trindade Da Rosa, 2016; Su et al., 2019). Mammary epithelial cells have been shown to have immune activity, activating signaling pathways during mastitis (Song et al., 2014). TNF plays a role in various pathological processes, including immune cell regulation and immune response modulation (Shah et al., 2012; Gao et al., 2015). The NOD-like receptor regulates the immune and inflammatory responses in mammals' innate immune systems (Saxena and Yeretssian, 2014). IL-17 expression in milk peaked 24 to 48 h after pathogen challenge. These findings indicated that IL-17 was a significant cytokine in the development of dairy goat mastitis and played a critical role in mastitis development (Jing et al., 2012). Previously published research indicated that mastitis involves the NOD-like receptor, IL-17, and TNF signaling pathways (Asselstine et al., 2019). As a result of their function, we can deduce that these pathways are involved in immune system responses to mastitis disease.

The PPI networks constructed using Cytoscape revealed that the hub genes are *PTPRC*, SOD2, and *STAT1*. As a result, these hub genes may affect mastitis and thus warrant further validation. The *PTPRC* gene is required to signal T- and B-cell antigen receptors (Miterski et al., 2002; Porcu et al., 2012). *PTPRC* is a highly connected gene in PPI networks and is

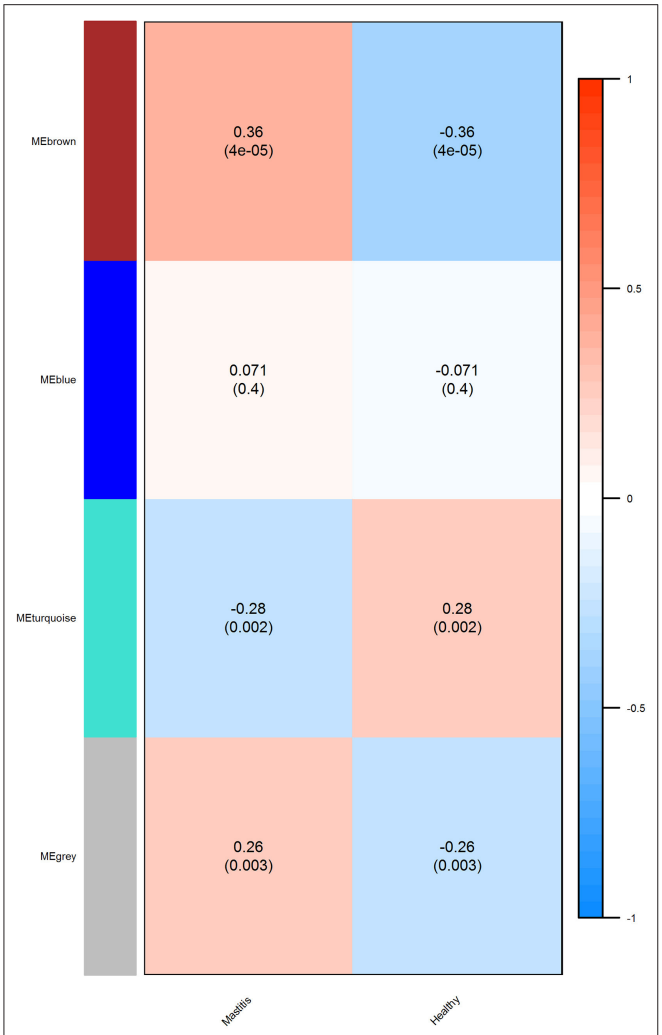


**FIGURE 7 |** The module trait relationship (*p* value) for identified modules (y axis) in relation with traits (x axis). The relationship was colored according to the correlation between the module and traits (turquoise, strong negative correlation; brown, strong positive correlation).

**TABLE 5 |** Top hub genes in significant modules in mastitis disease.

| **Mastitis disease** | |
|---|---|
| **Brown** | **Turquoise** |
| *PRDX5* | *CD53* |
| *RAB5C* | *NCKAP1L* |
| *ACTN4* | *ARHGEF2* |
| *SLC25A16* | *COL9A1* |
| *MAPK6* | *PTPRC* |

**TABLE 6 |** Results of different attribute weighting algorithms confirmed the most important genes.

| **Attribute** | **No. of weighting models** |
|---|---|
| *LOC407171* | 5 |
| *MT2A* | 4 |
| *PTPRC* | 4 |
| *LPCAT2* | 4 |
| *SAMSN1* | 4 |
| *IL1B* | 4 |
| *SELPLG* | 4 |
| *CD53* | 4 |
| *PLEK* | 4 |
| *SFN* | 3 |
| *KCNJ15* | 3 |
| *SPCS3* | 3 |
| *SOD2* | 3 |
| *IDH1* | 3 |
| *SYNGR1* | 3 |
| *TANC2* | 3 |
| *CXCL3* | 3 |

**TABLE 7 |** Accuracy comparison of constructed DT models by different criteria.

| **Criteria** | **Accuracy (%)** |
|---|---|
| Gain ratio | 75 |
| Information gain | 63.89 |
| Gini index | 58.33 |
| Accuracy | 63.89 |

involved in the development of mastitis (Bakhtiarizadeh et al., 2020). *SOD2* and *IDH1* genes have been up-regulated in ewes' mammary glands using functional enrichment analysis (Gao et al., 2018, 2019). *SOD2* gene expression increased in mammary tissue of cows and ewes with mastitis caused by *S. aureus* and *E. coli* (Mitterhuemer et al., 2010; Jensen et al., 2013). Also, in addition, *STAT1* regulates genes involved in milk protein synthesis, fat metabolism, and immune cell activation (Cobanoglu et al., 2006). The analysis of common meta-gene coexpression networks identified four modules, two of which were significant. These modules were the most significant in the current study based on the enriched functional terms related to mastitis development. The brown module's most essential genes included *PRDX5*, *RAB5C*, *ACTN4*, and *MAPK6*. The *PRDX5* gene is expressed ubiquitously in tissues and protects cells from oxidative stress by detoxifying peroxides (Knoops et al., 2011). *PRDX5* has been shown to play a critical role in inflammation in mice by protecting cells from oxidative stress (Argyropoulou et al., 2016). In addition, the *PRDX5* gene expression is increased in mastitis sheep milk (Pisanu et al., 2015). *RAB5C* and *MAPK6* genes were identified as candidate genes for mastitis in dairy cattle following intramammary infection with *E. coli* or *S. uberis* using a combination of GWAS and DEG data analyses (Chen et al., 2015). The ACTN4 gene was identified as the DEG in mastitis vs. healthy samples of sheep milk by transcriptomic analysis (Bonnefont et al., 2011). Furthermore, ACTN4 was identified as a hub gene in mastitis-related modules (Bakhtiarizadeh et al., 2020). On the other hand, the turquoise module's master genes were the *CD53*,

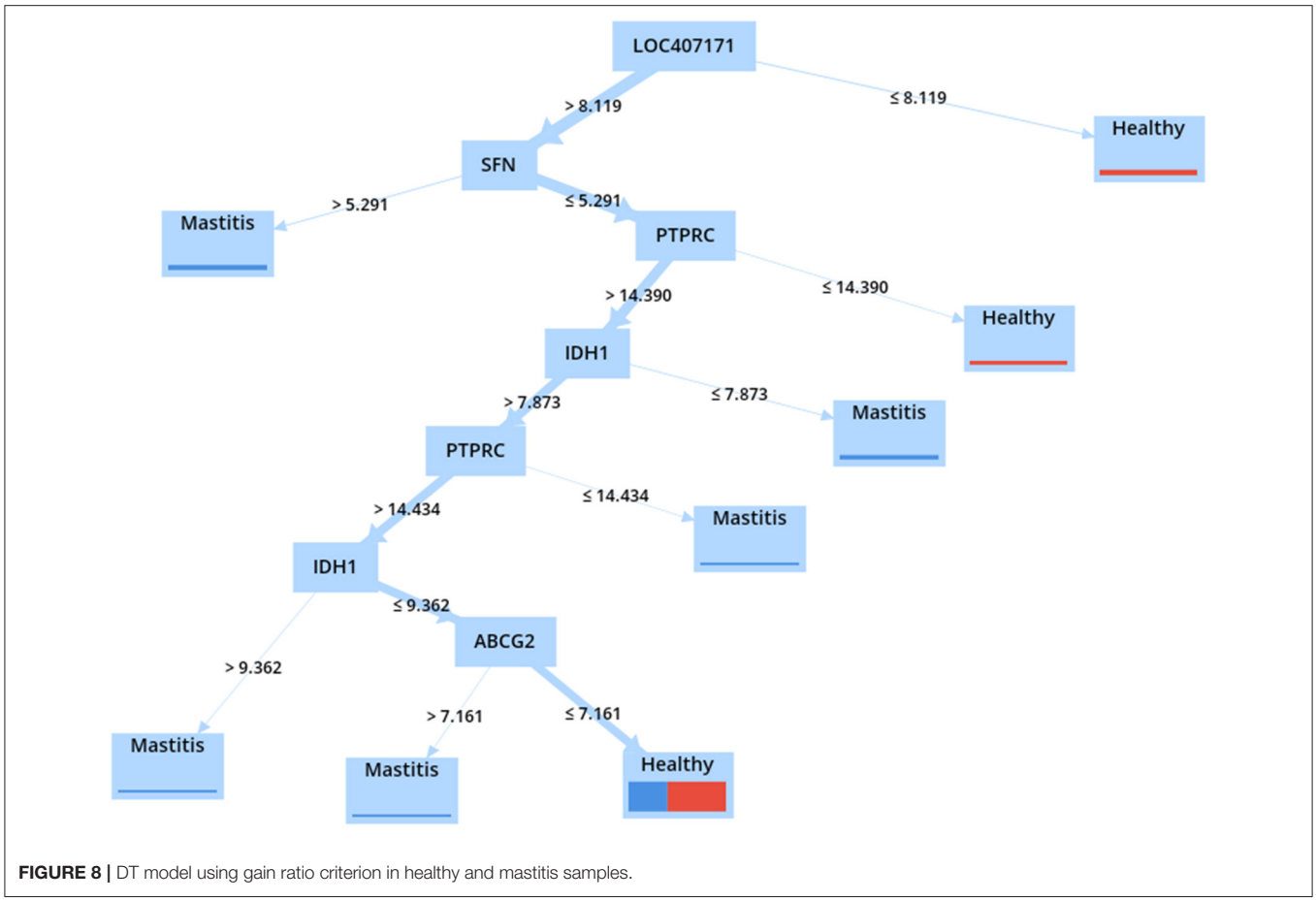


**FIGURE 8 |** DT model using gain ratio criterion in healthy and mastitis samples.

*ARHGEF2*, and *COL9A1* genes. *CD53* regulates cell development, and its function has been implicated in mastitis disease (Rinaldi et al., 2010). The results of a high-throughput analysis on infected bovine mammary glands with *E. coli* indicated the *ARHGEF2* gene's importance (Bagnicka et al., 2021). The *COL9A1* gene has been implicated in research involving identifying genomic regions and expression analysis of mastitis (Lu et al., 2020).

Several genes, including *LOC407171*, *MT2A*, *LPCAT2*, *CXCL3*, *SFN*, *IDH1*, and *ABCG2*, were confirmed as essential genes based on the outcome of the attribute weighting algorithm. The *LOC407171* gene is associated with the innate immune response in beef cattle and has been identified as an up-regulated gene in a dairy cow with *E. coli* mastitis (Li et al., 2019). *MT2A* plays a role in stimulus response in the pathogenesis of bovine *E. coli* in early lactation cows (Cheng et al., 2021). *LPCAT2* regulates the glycerophospholipid metabolism in periparturient dairy cattle (Bakhtiarizadeh et al., 2020). *CXCL3* is recognized as a proinflammatory cytokine in dairy cows with experimentally induced *S. aureus* clinical mastitis (Peralta et al., 2020). *SFN* was reported to regulate cell cycle progression in bovine mastitis *via* genome-wide association (Miles et al., 2021). *IDH1* was identified as a candidate gene in the milk transcriptome of dairy cattle implicated in innate immunity by pathway and network analysis (Banos et al., 2017). *ABCG2* gene, which is regulated by the mammary gland, responsible for the active secretion of several compounds into milk (Otero et al., 2015).

The DT model identified the *LOC407171* gene as a critical player in mastitis disease in this study. *LOC407171* has been validated using an attribute weighting algorithm and a machine-learning algorithm. In addition, *SFN* and *IDH1* were identified using attribute weighting and machine-learning techniques, with *IDH1*, validated using WGCNA. Furthermore, *ABCG2* is recognized using weighted attributes, machine learning, and WGCNA. In addition, machine learning, attribute weighting, PPI network, and WGCNA were used to confirm *PTPRC*.

We examined possible changes in gene expression and connectivity during mastitis, and it was concluded that genes involved in the development, proliferation, and differentiation of cells in the mammary gland, as well as genes involved in immune system improvement, were primarily altered in their expression.

## CONCLUSION

Because of the complexity of mastitis disease in dairy animals, far more relevant research is required to identify biomarkers associated with mastitis. The current study's findings from meta-analysis, WGCNA, and machine-learning approach allow us to represent the primary contribution to our understanding

of the most valuable genes for *E. coli* mastitis, which may provide a more robust biosignature and thus serve as reliable biomarker candidates in future studies. Our study suggests that all identified genes affect mastitis disease *via* their immune system–related functions.

## DATA AVAILABILITY STATEMENT

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

## AUTHOR CONTRIBUTIONS

NG: research concept and design, data analysis and interpretation, wrote the article, and final approval of the article. JS, SR, and KH: wrote the article. BP: data analysis, interpretation, wrote the article, and final approval of the article. All authors contributed to the article and approved the submitted version.

## ACKNOWLEDGMENTS

We are grateful to Dr. Mohammad Farhadian and Dr. Sana Farhadi for their kindly help.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2021.712306/full#supplementary-material

## REFERENCES


Anders, S., Mccarthy, D. J., Chen, Y., Okoniewski, M., Smyth, G. K., Huber, W., et al. (2013). Count-based differential expression analysis of RNA sequencing data using R and Bioconductor. *Nat. Protoc.* 8:1765. doi: 10.1038/nprot.2013.099

Argyropoulou, V., Goemaere, J., Clippe, A., Lefort, C., Tissir, F., Schakman, O., et al. (2016). Peroxiredoxin-5 as a novel actor in inflammation and tumor suppression. *Free Radic. Biol. Med.* 100:S92. doi: 10.1016/j.freeradbiomed.2016.10.229

Arnellos, A. (2018). From organizations of processes to organisms and other biological individuals. *Everything Flows: Toward. Process. Philo. Biol.* 3, 199–221. doi: 10.1093/oso/9780198779636.003.0010

Asselstine, V., Miglior, F., Suárez-Vega, A., Fonseca, P., Mallard, B., Karrow, N., et al. (2019). Genetic mechanisms regulating the host response during mastitis. *J. Dairy Sci.* 102, 9043–9059. doi: 10.3168/jds.2019-16504

Bagnicka, E., Kawecka-Grochocka, E., Pawlina-Tyszko, K., Zalewska, M., Kapusta, A., Kościuczuk, E., et al. (2021). MicroRNA expression profile in bovine mammary gland parenchyma infected by coagulase-positive or coagulase-negative staphylococci. *Vet. Res.* 52, 1–20. doi: 10.1186/s13567-021-00912-2

Bakhtiarizadeh, M. R., Mirzaei, S., Norouzi, M., Sheybani, N., and Vafaei Sadi, M. S. (2020). Identification of gene modules and hub genes involved in mastitis development using a systems biology approach. *Front. Genet.* 11:722. doi: 10.3389/fgene.2020.00722

Banos, G., Bramis, G., Bush, S., Clark, E., Mcculloch, M., Smith, J., et al. (2017). The genomic architecture of mastitis resistance in dairy sheep. *BMC Genom.* 18, 1–18. doi: 10.1186/s12864-017-3982-1

Bansal, M., and Di Bernardo, D. (2007). Inference of gene networks from temporal gene expression profiles. *IET Syst. Biol.* 1, 306–312. doi: 10.1049/iet-syb:20060079

Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. Roy. Statistic. Soc. Series B* 57, 289–300. doi: 10.1111/j.2517-6161.1995.tb02031.x

Blum, S. E., Goldstone, R. J., Connolly, J. P., Répérant-Ferter, M., Germon, P., Inglis, N. F., et al. (2018). Postgenomics characterization of an essential genetic determinant of mammary pathogenic Escherichia coli. *MBio* 9:18. doi: 10.1128/mBio.00423-18

Bolstad, B. M., Irizarry, R. A., Åstrand, M., and Speed, T.P. (2003). A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics* 19, 185–193. doi: 10.1093/bioinformatics/19.2.185

Bonnefont, C. M., Toufeer, M., Caubet, C., Foulon, E., Tasca, C., Aurel, M.-R., et al. (2011). Transcriptomic analysis of milk somatic cells in mastitis resistant and susceptible sheep upon challenge with Staphylococcus epidermidis and Staphylococcus aureus. *BMC Genom.* 12, 1–16. doi: 10.1186/1471-2164-12-208

Botía, J. A., Vandrovcova, J., Forabosco, P., Guelfi, S., D'sa, K., Hardy, J., et al. (2017). An additional k-means clustering step improves the biological features of WGCNA gene co-expression networks. *BMC Syst. Biol.* 11, 1–16. doi: 10.1186/s12918-017-0420-6

Brand, B., Hartmann, A., Repsilber, D., Griesbeck-Zilch, B., Wellnitz, O., Kühn, C., et al. (2011). Comparative expression profiling of E. coli and S. aureus inoculated primary mammary gland cells sampled from cows with different genetic predispositions for somatic cell score. *Gene. Select. Evol.* 43, 1–17. doi: 10.1186/1297-9686-43-24

Buitenhuis, B., Røntved, C. M., Edwards, S. M., Ingvartsen, K. L., and Sørensen, P. (2011). In depth analysis of genes and pathways of the mammary gland involved in the pathogenesis of bovine Escherichia coli-mastitis. *BMC Genom.* 12, 1–10. doi: 10.1186/1471-2164-12-130

Burvenich, C., Van Merris, V., Mehrzad, J., Diez-Fraile, A., and Duchateau, L. (2003). Severity of *E. coli* mastitis is mainly determined by cow factors. *Veterinary Res.* 34, 521–564. doi: 10.1051/vetres:2003023

Cebron, N., Maman, S., Walachowski, S., Gausserès, B., Cunha, P., Rainard, P., et al. (2020). Th17-related mammary immunity, but not a high systemic Th1 immune response is associated with protection against E. *coli mastitis. NPJ vaccines* 5, 1–13. doi: 10.1038/s41541-020-00258-4

Chen, X., Cheng, Z., Zhang, S., Werling, D., and Wathes, D. C. (2015). Combining genome wide association studies and differential gene expression data analyses identifies candidate genes affecting mastitis caused by two different pathogens in the dairy cow. *Open J. Anim. Sci.* 5, 358–393. doi: 10.4236/ojas.2015.54040

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*. doi: 10.21203/rs.3.rs-204708/v1

Cobanoglu, O., Zaitoun, I., Chang, Y., Shook, G., and Khatib, H. (2006). Effects of the signal transducer and activator of transcription 1 (STAT1) gene on milk production traits in Holstein dairy cattle. *J. Dairy Sci.* 89, 4433–4437. doi: 10.3168/jds.S0022-0302(06)72491-2

Compton, C. W., Cursons, R. T., Barnett, C. M., and Mcdougall, S. (2009). Expression of innate resistance factors in mammary secretion from periparturient dairy heifers and their association with subsequent infection status. *Vet. Immunol. Immunopathol.* 127, 357–364. doi: 10.1016/j.vetimm.2008.10.331

Dennis, G., Sherman, B. T., Hosack, D. A., Yang, J., Gao, W., Lane, H. C., et al. (2003). DAVID: database for annotation, visualization, and integrated discovery. *Genome Biol.* 4, 1–11. doi: 10.1186/gb-2003-4-9-r60

Duarte, C. M., Freitas, P. P., and Bexiga, R. (2015). Technological advances in bovine mastitis diagnosis: an overview. *J. Vet. Diagn. Investig.* 27, 665–672. doi: 10.1177/1040638715603087

Ebrahimi, M., Lakizadeh, A., Agha-Golzadeh, P., Ebrahimie, E., and Ebrahimi, M. (2011). Prediction of thermostability from amino acid attributes by combination of clustering with attribute weighting: a new vista in engineering enzymes. *PLoS ONE* 6:e23146. doi: 10.1371/journal.pone.0023146

Ebrahimie, M., Esmaeili, F., Cheraghi, S., Houshmand, F., Shabani, L., and Ebrahimie, E. (2014). Efficient and simple production of insulin-producing cells from embryonal carcinoma stem cells using mouse neonate pancreas extract, as a natural inducer. *PLoS ONE* 9:e90885. doi: 10.1371/journal.pone.0090885

Evangelou, E., and Ioannidis, J. P. (2013). Meta-analysis methods for genome-wide association studies and beyond. *Nat. Rev. Genet.* 14, 379–389. doi: 10.1038/nrg3472

Farhadian, M., Rafat, S. A., Hasanpur, K., Ebrahimi, M., and Ebrahimie, E. (2018a). Cross-species meta-analysis of transcriptomic data in combination with supervised machine learning models identifies the common gene signature of lactation process. *Front. Genet.* 9:235. doi: 10.3389/fgene.2018.00235

Farhadian, M., Rafat, S. A., Hasanpur, K., and Ebrahimie, E. (2018b). Transcriptome signature of the lactation process, identified by meta-analysis of microarray and RNA-Seq data. BioTechnologia. *J. Biotechnol. Comput. Biol. Bionanotechnol.* 99:75659. doi: 10.5114/bta.2018.75659

Farhadian, M., Rafat, S. A., Panahi, B., and Ebrahimie, E. (2020). Transcriptome signature of two lactation stages in Ghezel sheep identifies using RNA-Sequencing. *Anim. Biotechnol.* 20, 1–11. doi: 10.1080/10495398.2020.1784185

Farhadian, M., Rafat, S. A., Panahi, B., and Mayack, C. (2021). Weighted gene co-expression network analysis identifies modules and functionally enriched pathways in the lactation process. *Sci. Rep.* 11, 1–15. doi: 10.1038/s41598-021-97893-1

Gao, J., Li, T., Lu, Z., Wang, X., Zhao, X., and Ma, Y. (2019). Proteomic analyses of mammary glands provide insight into the immunity and metabolism pathways associated with clinical mastitis in meat sheep. *Animals* 9:309. doi: 10.3390/ani9060309

Gao, Y., Theng, S. S., Mah, W.-C., and Lee, C. G. (2015). Silibinin down-regulates FAT10 and modulate TNF-α/IFN-γ-induced chromosomal instability and apoptosis sensitivity. *Biol. Open* 4, 961–969. doi: 10.1242/bio.011189

Gao, J., Ma, Y., Li, T., Lu, Z., Chen, C., and Zhao, X. (2018). Expression and localization of SOD2 gene in breast tissue of clinical mastitis sheep (Ovis aries). *J. Agric. Biotechnol.* 26, 246–252. https://pubmed.ncbi.nlm.nih.gov/12154050/

Gelasakis, A., Mavrogianni, V., Petridis, I., Vasileiou, N., and Fthenakis, G. (2015). Mastitis in sheep–The last 10 years and the future of research. *Vet. Microbiol.* 181, 136–146. doi: 10.1016/j.vetmic.2015.07.009

Goldman, B., Nierman, W., Kaiser, D., Slater, S., Durkin, A. S., Eisen, J. A., et al. (2006). Evolution of sensory complexity recorded in a myxobacterial genome. *Proc. Nat. Acad. Sci.* 103, 15200–15205. doi: 10.1073/pnas.0607335103

Gorji, A. E., Roudbari, Z., Sadeghi, B., Javadmanesh, A., and Sadkowski, T. (2019). Transcriptomic analysis on the promoter regions discover gene networks involving mastitis in cattle. *Microb. Pathog.* 137:103801. doi: 10.1016/j.micpath.2019.103801

Günther, J., Esch, K., Poschadel, N., Petzl, W., Zerbe, H., Mitterhuemer, S., et al. (2011). Comparative kinetics of Escherichia coli-and Staphylococcus aureus-specific activation of key immune pathways in mammary epithelial cells demonstrates that S. aureus elicits a delayed response dominated by interleukin-6 (IL-6) but not by IL-1A or tumor necrosis factor alpha. *Infect. Immun.* 79, 695–707. doi: 10.1128/IAI.01071-10

Günther, J., Petzl, W., Zerbe, H., Schuberth, H.-J., Koczan, D., Goetze, L., et al. (2012). Lipopolysaccharide priming enhances expression of effectors of immune defence while decreasing expression of pro-inflammatory cytokines in mammary epithelia cells from cows. *BMC Genom.* 13, 1–13. doi: 10.1186/1471-2164-13-17

Haidich, A.-B. (2010). Meta-analysis in medical research. *Hippokratia* 14, 29.

Jensen, K., Günther, J., Talbot, R., Petzl, W., Zerbe, H., Schuberth, H.-J., et al. (2013). Escherichia coli-and Staphylococcus aureus-induced mastitis differentially modulate transcriptional responses in neighbouring uninfected bovine mammary gland quarters. *BMC Genom.* 14, 1–19. doi: 10.1186/1471-2164-14-36

Jing, X., Zhao, Y., Shang, C., Yao, Y., Tian, T., Li, J., et al. (2012). Dynamics of cytokines associated with IL-17 producing cells in serum and milk in mastitis of experimental challenging with Staphylococcus aureus and *Escherichia coli* in dairy goats. *J. Anim. Veterin. Adv.* 11, 475–479. doi: 10.3923/javaa.2012.475.479

Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., and Salzberg, S. L. (2013). TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol.* 14, 1–13. doi: 10.1186/gb-2013-14-4-r36

Knoops, B., Goemaere, J., Van Der Eecken, V., and Declercq, J.-P. (2011). Peroxiredoxin 5: structure, mechanism, and function of the mammalian atypical 2-Cys peroxiredoxin. *Antioxid. Redox Signal.* 15, 817–829. doi: 10.1089/ars.2010.3584

Kulkarni, A. G., and Kaliwal, B. (2013). Bovine mastitis: a review. *Int. J. Recent Sci. Res.* 4, 543–548. http://recentscientific.com/bovine-mastitis-review

Lai, Y.-C., Fujikawa, T., Maemura, T., Ando, T., Kitahara, G., Endo, Y., et al. (2017). Inflammation-related microRNA expression level in the bovine milk is affected by mastitis. *PLoS ONE* 12:e0177182. doi: 10.1371/journal.pone.0177182

Langfelder, P., and Horvath, S. (2008). WGCNA: an R package for weighted correlation network analysis. *BMC Bioinform.* 9, 1–13. doi: 10.1186/1471-2105-9-559

Li, L., Chen, X., and Chen, Z. (2019). Identification of key candidate genes in dairy cow in response to escherichia coli mastitis by bioinformatical analysis. *Front. Genet.* 10:1251. doi: 10.3389/fgene.2019.01251

Long, E., Capuco, A., Wood, D., Sonstegard, T., Tomita, G., Paape, M., et al. (2001). Escherichia coli induces apoptosis and proliferation of mammary cells. *Cell Death Differ.* 8, 808–816. doi: 10.1038/sj.cdd.4400878

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, 1–21. doi: 10.1186/s13059-014-0550-8

Lu, X., Duan, A., Liang, S., Ma, X., and Deng, T. (2020). Genomic identification, evolution, and expression analysis of collagen genes family in water buffalo during lactation. *Genes* 11:515. doi: 10.3390/genes11050515

Marot, G., Jaffrézic, F., and Rau, A. (2020). metaRNASeq: Differential meta-analysis of RNA-seq data. *dim (param)* 1:3. https://cran.r-project.org/web/packages/metaRNASeq/vignettes/metaRNASeq.pdf

Miles, A. M., Posbergh, C. J., and Huson, H. J. (2021). Direct phenotyping and principal component analysis of type traits implicate novel QTL in bovine mastitis through genome-wide association. *Animals* 11:1147. doi: 10.3390/ani11041147

Miterski, B., Sindern, E., Haupts, M., Schimrigk, S., and Epplen, J. T. (2002). PTPRC (CD45) is not associated with multiple sclerosis in a large cohort of German patients. *BMC Med. Genet.* 3, 1–3. doi: 10.1186/1471-2350-3-3

Mitterhuemer, S., Petzl, W., Krebs, S., Mehne, D., Klanner, A., Wolf, E., et al. (2010). Escherichia coli infection induces distinct local and systemic transcriptome responses in the mammary gland. *BMC Genom.* 11, 1–16. doi: 10.1186/1471-2164-11-138

Moyes, K., Sørensen, P., and Bionaz, M. (2016). The impact of intramammary Escherichia coli challenge on liver and mammary transcriptome and cross-talk in dairy cows during early lactation using RNAseq. *PLoS ONE* 11:e0157480. doi: 10.1371/journal.pone.0157480

Nami, Y., Imeni, N., and Panahi, B. (2021). Application of machine learning in bacteriophage research. *BMC Microbiol.* 21, 1–8. doi: 10.1186/s12866-021-02256-5

Otero, J., Barrera, B., De La Fuente, A., Prieto, J., Marqués, M., Álvarez, A., et al. (2015). The gain-of-function Y581S polymorphism of the ABCG2 transporter increases secretion into milk of danofloxacin at the therapeutic dose for mastitis treatment. *J. Dairy Sci.* 98, 312–317. doi: 10.3168/jds.2014-8288

Panahi, B., Farhadian, M., and Hejazi, M. A. (2020). Systems biology approach identifies functional modules and regulatory hubs related to secondary metabolites accumulation after transition from autotrophic to heterotrophic growth condition in microalgae. *PLoS ONE* 15:e0225677. doi: 10.1371/journal.pone.0225677

Panahi, B., Frahadian, M., Dums, J. T., and Hejazi, M. A. (2019a). Integration of cross species RNA-Seq meta-analysis and machine-learning models identifies the most important salt stress–responsive pathways in microalga Dunaliella. *Front. Genet.* 10:752. doi: 10.3389/fgene.2019.00752

Panahi, B., and Hejazi, M. A. (2021). Weighted gene co-expression network analysis of the salt-responsive transcriptomes reveals novel hub genes in green halophytic microalgae Dunaliella salina. *Sci. Rep.* 11, 1–11. doi: 10.1038/s41598-020-80945-3

Panahi, B., Mohammadi, S. A., and Doulati-Baneh, H. (2019b). Characterization of Iranian grapevine cultivars using machine learning models. *Proceed. Nat. Acad. Sci. India Sect. B: Biol. Sci.* 19, 1–7. doi: 10.1007/s40011-019-01131-8

Peralta, O., Carrasco, C., Vieytes, C., Tamayo, M., Muñoz, I., Sepulveda, S., et al. (2020). Safety and efficacy of a mesenchymal stem cell intramammary therapy in dairy cows with experimentally induced Staphylococcus aureus clinical mastitis. *Sci. Rep.* 10, 1–12. doi: 10.1038/s41598-020-59724-7

Pisanu, S., Cubeddu, T., Pagnozzi, D., Rocca, S., Cacciotto, C., Alberti, A., et al. (2015). Neutrophil extracellular traps in sheep mastitis. *Vet. Res.* 46, 1–14. doi: 10.1186/s13567-015-0196-x

Porcu, M., Kleppe, M., Gianfelici, V., Geerdens, E., De Keersmaecker, K., Tartaglia, M., et al. (2012). Mutation of the receptor tyrosine phosphatase PTPRC (CD45) in T-cell acute lymphoblastic leukemia. *Blood* 119, 4476–4479. doi: 10.1182/blood-2011-09-379958

Rhodes, D. R., Barrette, T. R., Rubin, M. A., Ghosh, D., and Chinnaiyan, A. M. (2002). Meta-analysis of microarrays: interstudy validation of gene expression profiles reveals pathway dysregulation in prostate cancer. *Cancer Res.* 62, 4427–4433. https://pubmed.ncbi.nlm.nih.gov/12154050/

Rinaldi, M., Li, R. W., Bannerman, D. D., Daniels, K. M., Evock-Clover, C., Silva, M. V., et al. (2010). A sentinel function for teat tissues in dairy cows: dominant innate immune response elements define early response to E. *coli* mastitis. *Funct. Integrat. Genom.* 10, 21–38. doi: 10.1007/s10142-009-0133-z

Robinson, M. D., and Oshlack, A. (2010). A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol.* 11, 1–9. doi: 10.1186/gb-2010-11-3-r25

Saxena, M., and Yeretssian, G. (2014). NOD-like receptors: master regulators of inflammation and cancer. *Front. Immunol.* 5:327. doi: 10.3389/fimmu.2014.00327

Shah, N., Dhar, D., Mohammed, F. E. Z., Habtesion, A., Davies, N. A., Jover-Cobos, M., et al. (2012). Prevention of acute kidney injury in a rodent model of cirrhosis following selective gut decontamination is associated with reduced renal TLR4 expression. *J. Hepatol.* 56, 1047–1053. doi: 10.1016/j.jhep.2011.11.024

Shannon, P., Markiel, A., Ozier, 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, 2498–2504. doi: 10.1101/gr.1239303

Sharifi, S., Pakdel, A., Ebrahimi, M., Reecy, J. M., Fazeli Farsani, S., and Ebrahimie, E. (2018). Integration of machine learning and meta-analysis identifies the transcriptomic bio-signature of mastitis disease in cattle. *PLoS ONE* 13:e0191227. doi: 10.1371/journal.pone.0191227

Sipka, A., Klaessig, S., Duhamel, G. E., Swinkels, J., Rainard, P., and Schukken, Y. (2014). Impact of intramammary treatment on gene expression profiles in bovine *Escherichia coli* mastitis. *PLoS ONE* 9:e85579. doi: 10.1371/journal.pone.0085579

Smyth, G. K., Ritchie, M., Thorne, N., and Wettenhall, J. (2005). LIMMA: linear models for microarray data. in bioinformatics and computational biology solutions using R and bioconductor. *Stat. Biol. Health* 62:23. doi: 10.1007/0-387-29362-0_23

Song, X., Guo, M., Wang, T., Wang, W., Cao, Y., and Zhang, N. (2014). Geniposide inhibited lipopolysaccharide-induced apoptosis by modulating TLR4 and apoptosis-related factors in mouse mammary glands. *Life Sci.* 119, 9–17. doi: 10.1016/j.lfs.2014.10.006

Su, T., Li, W., Wang, P., and Ma, C. (2019). Dynamics of peroxisome homeostasis and its role in stress response and signaling in plants. *Front. Plant Sci.* 10:705. doi: 10.3389/fpls.2019.00705

Swanson, K., Stelwagen, K., Dobson, J., Henderson, H., Davis, S., Farr, V., et al. (2009). Transcriptome profiling of Streptococcus uberis-induced mastitis reveals fundamental differences between immune gene expression in the mammary gland and in a primary cell culture model. *J. Dairy Sci.* 92, 117–129. doi: 10.3168/jds.2008-1382

Szklarczyk, D., Franceschini, A., Wyder, S., Forslund, K., Heller, D., Huerta-Cepas, J., et al. (2015). STRING v10: protein–protein interaction networks, integrated over the tree of life. *Nucleic Acids Res.* 43, D447–D452. doi: 10.1093/nar/gku1003

Takeshima, S., Matsumoto, Y., Chen, J., Yoshida, T., Mukoyama, H., and Aida, Y. (2008). Evidence for cattle major histocompatibility complex (BoLA) class II DQA1 gene heterozygote advantage against clinical mastitis caused by Streptococci and Escherichia species. *Tissue Antigens* 72, 525–531. doi: 10.1111/j.1399-0039.2008.01140.x

Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D. R., et al. (2012). Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat. Protoc.* 7, 562. doi: 10.1038/nprot.2012.016

Trindade Da Rosa, F. (2016). Role of peroxisome proliferator-activated receptor gamma on prevention/cure of mastitis. *Review* 54(8):2460–2470. doi: 10.2337/diabetes.54.8.2460

Tseng, G. C., Ghosh, D., and Feingold, E. (2012). Comprehensive literature review and statistical considerations for microarray meta-analysis. *Nucleic Acids Res.* 40, 3785–3799. doi: 10.1093/nar/gkr1265

Vangroenweghe, F., Duchateau, L., Boutet, P., Lekeux, P., Rainard, P., Paape, M., et al. (2005). Effect of carprofen treatment following experimentally induced *Escherichia coli* mastitis in primiparous cows. *J. Dairy Sci.* 88, 2361–2376. doi: 10.3168/jds.S0022-0302(05)72914-3

Vasudevan, P., Nair, M. K. M., Annamalai, T., and Venkitanarayanan, K. S. (2003). Phenotypic and genotypic characterization of bovine mastitis isolates of Staphylococcus aureus for biofilm formation. *Vet. Microbiol.* 92, 179–185. doi: 10.1016/S0378-1135(02)00360-7

Wang, G., Muschelli, J., and Lindquist, M. A. (2021). Moderated t-tests for group-level fMRI analysis. *Neuroimage* 237:118141. doi: 10.1016/j.neuroimage.2021.118141

Wang, X., Kang, D. D., Shen, K., Song, C., Lu, S., Chang, L.-C., et al. (2012). An R package suite for microarray meta-analysis in quality control, differentially expressed gene analysis and pathway enrichment detection. *Bioinformatics* 28, 2534–2536. doi: 10.1093/bioinformatics/bts485

**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: 18 October 2021
doi: 10.3389/fgene.2021.753839

# Integrated Network Analysis to Identify Key Modules and Potential Hub Genes Involved in Bovine Respiratory Disease: A Systems Biology Approach

*Aliakbar Hasankhani[1]*, Abolfazl Bahrami[1,2]*, Negin Sheybani[3], Farhang Fatehi[1], Roxana Abadeh[4], Hamid Ghaem Maghami Farahani[1], Mohammad Reza Bahreini Behzadi[5], Ghazaleh Javanmard[1], Sadegh Isapour[1], Hosein Khadem[6] and Herman W. Barkema[7]*

[1]Department of Animal Science, College of Agriculture and Natural Resources, University of Tehran, Karaj, Iran, [2]Biomedical Center for Systems Biology Science Munich, Ludwig-Maximilians-University, Munich, Germany, [3]Department of Animal and Poultry Science, College of Aburaihan, University of Tehran, Tehran, Iran, [4]Department of Animal Science, Science and Research Branch, Islamic Azad University, Tehran, Iran, [5]Department of Animal Science, Faculty of Agriculture, Yasouj University, Yasouj, Iran, [6]Department of Agronomy and Plant Breeding, University of Tehran, Karaj, Iran, [7]Department of Production Animal Health, Faculty of Veterinary Medicine, University of Calgary, Calgary, AB, Canada

**Background:** Bovine respiratory disease (BRD) is the most common disease in the beef and dairy cattle industry. BRD is a multifactorial disease resulting from the interaction between environmental stressors and infectious agents. However, the molecular mechanisms underlying BRD are not fully understood yet. Therefore, this study aimed to use a systems biology approach to systematically evaluate this disorder to better understand the molecular mechanisms responsible for BRD.

**Methods:** Previously published RNA-seq data from whole blood of 18 healthy and 25 BRD samples were downloaded from the Gene Expression Omnibus (GEO) and then analyzed. Next, two distinct methods of weighted gene coexpression network analysis (WGCNA), i.e., module–trait relationships (MTRs) and module preservation (MP) analysis were used to identify significant highly correlated modules with clinical traits of BRD and non-preserved modules between healthy and BRD samples, respectively. After identifying respective modules by the two mentioned methods of WGCNA, functional enrichment analysis was performed to extract the modules that are biologically related to BRD. Gene coexpression networks based on the hub genes from the candidate modules were then integrated with protein–protein interaction (PPI) networks to identify hub–hub genes and potential transcription factors (TFs).

**Abbreviations:** BCV, bovine coronavirus; BHV-1, bovine herpes virus type 1; BPIV-3, bovine parainfluenza type 3 virus; BRD, bovine respiratory disease; BRSV, bovine respiratory syncytial virus; BVDV, bovine viral diarrhea virus; GEO, Gene Expression Omnibus; GO, Gene Ontology; GS, gene significance; KEGG, Kyoto Encyclopedia of Genes and Genomes; ME, module eigengene; MM, module membership; MTRs, module–trait relationships; MP, module preservation; PPI, protein–protein interaction; RSV, respiratory syncytial virus; RNA-seq, RNA-sequencing; TF, transcription factor; TOM, topological overlap matrix; WGCNA, weighted gene coexpression network analysis.

OPEN ACCESS

**Edited by:**
*Andressa Oliveira De Lima, University of Washington, United States*

**Reviewed by:**
*Bárbara Silva-Vignato, University of São Paulo, Brazil*
*Hugo Oswaldo Toledo-Alvarado, National Autonomous University of Mexico, Mexico*

***Correspondence:**
*Aliakbar Hasankhani*
*A.hasankhani74@ut.ac.ir*
*Abolfazl Bahrami*
*A.Bahrami@ut.ac.ir*

**Specialty section:**
*This article was submitted to Livestock Genomics, a section of the journal Frontiers in Genetics*

**Received:** *05 August 2021*
**Accepted:** *28 September 2021*
**Published:** *18 October 2021*

**Citation:**
*Hasankhani A, Bahrami A, Sheybani N, Fatehi F, Abadeh R, Ghaem Maghami Farahani H, Bahreini Behzadi MR, Javanmard G, Isapour S, Khadem H and Barkema HW (2021) Integrated Network Analysis to Identify Key Modules and Potential Hub Genes Involved in Bovine Respiratory Disease: A Systems Biology Approach. Front. Genet. 12:753839. doi: 10.3389/fgene.2021.753839*

**Results:** Four significant highly correlated modules with clinical traits of BRD as well as 29 non-preserved modules were identified by MTRs and MP methods, respectively. Among them, two significant highly correlated modules (identified by MTRs) and six nonpreserved modules (identified by MP) were biologically associated with immune response, pulmonary inflammation, and pathogenesis of BRD. After aggregation of gene coexpression networks based on the hub genes with PPI networks, a total of 307 hub–hub genes were identified in the eight candidate modules. Interestingly, most of these hub–hub genes were reported to play an important role in the immune response and BRD pathogenesis. Among the eight candidate modules, the turquoise (identified by MTRs) and purple (identified by MP) modules were highly biologically enriched in BRD. Moreover, *STAT1*, *STAT2*, *STAT3*, *IRF7*, and *IRF9* TFs were suggested to play an important role in the immune system during BRD by regulating the coexpressed genes of these modules. Additionally, a gene set containing several hub–hub genes was identified in the eight candidate modules, such as *TLR2*, *TLR4*, *IL10*, *SOCS3*, *GZMB*, *ANXA1*, *ANXA5*, *PTEN*, *SGK1*, *IFI6*, *ISG15*, *MX1*, *MX2*, *OAS2*, *IFIH1*, *DDX58*, *DHX58*, *RSAD2*, *IFI44*, *IFI44L*, *EIF2AK2*, *ISG20*, *IFIT5*, *IFITM3*, *OAS1Y*, *HERC5*, and *PRF1*, which are potentially critical during infection with agents of bovine respiratory disease complex (BRDC).

**Conclusion:** This study not only helps us to better understand the molecular mechanisms responsible for BRD but also suggested eight candidate modules along with several promising hub–hub genes as diagnosis biomarkers and therapeutic targets for BRD.

**Keywords: bovine respiratory disease, RNA-seq, weighted gene co-expression network, protein-protein interaction, hub-hub genes**

## INTRODUCTION

Bovine respiratory disease (BRD) is the most common and costly infectious disease in the beef and dairy cattle industry. It causes 70–80% morbidity and 40–50% mortality in feedlot cattle in the United States (Edwards, 2010; Tizioto et al., 2015). BRD is a multifactorial disease, and its onset is usually associated with stress factors (nutritional or environmental risk factors) and the presence of infectious agents (Gagea et al., 2006; Grissett et al., 2015). Stress factors such as weaning, shipping distance, and commingling that negatively affect the immune system, can predispose cattle to a primary infection (Snowder et al., 2006; Timsit et al., 2016b). Infection is commonly caused by bovine respiratory disease complex (BRDC) including the viral and bacterial pathogens, which can affect the upper and lower respiratory system (Ellis, 2001; Caswell, 2014; Kirchhoff et al., 2014). Clinical diagnosis of BRD is made by visual observations and is usually based on clinical signs such as high rectal temperature, depression/lethargy, nasal or ocular discharge, increased respiration rate, reduced feed intake, and reduced average daily gain (Amrine et al., 2013; Behura et al., 2017). However, this method has low detection sensitivity and specificity, and the diagnosis is often made without identifying the cause of the disease (White and Renter, 2009; Timsit et al., 2016a). On the other hand, among the animals that are vaccinated against BRD, approximately 75% of them are protected (Hodgins et al., 2002), and the animals that are diagnosed based on clinical signs are treated with antimicrobials (Wilkinson, 2009). Moreover, excessive use of antimicrobial therapies for BRD facilitates the antibiotic resistance of microbes (Schaefer et al., 2007; Portis et al., 2012). In addition to vaccination and antimicrobials, other intervention methods such as nutritional manipulation and processing procedures have a limited effect on reducing morbidity and mortality rates, and despite extensive studies, BRD is still an issue (Taylor et al., 2010). Although the predisposing factors, viral and bacterial agents that cause BRD are relatively well known, the pathogenic mechanisms of BRD, the molecular immune response of the host to infection, and their association with the disease outcomes are not fully understood yet (Taylor et al., 2010; Johnston et al., 2019). Also, due to insufficient knowledge of the disease mechanisms, it is not possible to develop an effective method to identify animals with BRD (Sun et al., 2020). Therefore, understanding the infection dynamics and identification of new candidate biomarkers involved in BRD can help to better understanding the molecular mechanisms of BRD.

Functional genomic methods such as RNA-sequencing-based transcriptomics can provide a global gene expression profile, and their use in BRD studies can accelerate the understanding of disease mechanisms and the development of diagnosis (Rai et al., 2015). Several transcriptome studies have been performed on BRD in various tissues, such as the lung, bronchial lymph nodes (Behura et al., 2017; Johnston et al., 2019), and whole blood (Jiminez et al., 2021; Scott et al., 2021). For example, Sun et al.

(2020) reported differentially expressed genes (DEGs) in the blood tissue as biomarkers to recognize BRD cattle at entry, such as *MX1*, *IFIT3*, *ISG15*, *OAS2*, and *IFI6* involved in the interferon signaling pathway. Moreover, Tizioto et al. (2015) identified 142 DEGs that were located in quantitative trait locus regions associated with BRD risk. The most common application of differential expression analysis is to identify DEGs between different conditions. On the other hand, it is known that the differential expression analysis focuses only on the individual effect of genes. However, genes and gene products do not work individually but interact in complex gene networks (Liu et al., 2020). Therefore, individual evaluation of gene expression may not explain the cause of complex diseases such as BRD.

Systems biology is one of the suitable methods to better understand the mechanism of diseases (Darzi et al., 2021) and other complex traits (Salleh et al., 2018). In systems biology, there are various computational methods based on the network approach, and one of the fundamental aspects of the network approach in systems biology is the construction of gene coexpression networks using high-throughput gene expression data (Kadarmideen and Watson-Haigh, 2012). In this regard, one of the most widely used methods for building gene coexpression networks is weighted gene coexpression network analysis (WGCNA) (Langfelder and Horvath, 2008). The WGCNA method identifies clusters of coexpressed genes (also called a module) based on correlation patterns between expression profiles of genes across samples (Langfelder and Horvath, 2008). Furthermore, the WGCNA identifies highly connected genes (hub genes) by calculating intramodular gene connectivity, which are centrally in their modules and can be involved in important roles during the disease (Bakhtiarizadeh et al., 2018). The WGCNA approach has been used successfully in different disease studies in humans (Wang Y. et al., 2020), cattle (Yan et al., 2020), swine (Wilkinson et al., 2016), mice (Rangaraju et al., 2018), chickens (Liu and Cai, 2017), and sheep (Kadarmideen et al., 2011). One of the most widely used methods of WGCNA is module–trait relationship analysis. In this method, after identifying the modules across samples, module–trait relationships are calculated according to the correlation between module eigengenes and traits of interest, and finally, significant modules are identified (Kommadath et al., 2014; Sabino et al., 2018). Moreover, WGCNA provides another unique network-based method called module preservation analysis. This method focuses on determining network topology changes across different conditions. For example, it can be checked whether the network density and topological pattern of the modules identified in normal samples (as a reference) are preserved in the disease samples (as a test) (Langfelder et al., 2011). In this regard, the differences in the topology of these two networks indicate a significant perturbation of the coexpression patterns by the disease. Thus, the nonpreserved modules between these conditions, as well as their hub genes may exert crucial roles in the pathological processes of the disease (Mukund and Subramaniam, 2015; Riquelme Medina and Lubovac-Pilav, 2016; Bakhtiarizadeh et al., 2020).

In the present study, we used previously published RNA-seq data (Jiminez et al., 2021) and the WGCNA method for constructing weighted gene coexpression networks to better understand the molecular regulatory mechanisms responsible for the immune response to BRD. It should be noted that in the current study we used two distinct WGCNA methods to identify key modules, their hub genes, hub–hub genes, and regulatory factors involved in BRD: 1) module–trait relationships analysis across all samples to identify significant highly correlated modules with clinical traits of BRD, and 2) module preservation analysis between healthy samples (as reference set) and BRD samples (as test set) to identify nonpreserved modules between these conditions. The main hypothesis was that significant highly correlated modules with clinical traits of BRD (identified by the first method) and nonpreserved modules between healthy and BRD samples (identified by the second method) may contain potential functionally related genes and identifies biological regulatory systems involved in pathological processes of BRD, and it is also expected that these two methods confirm each other's results.

# MATERIALS AND METHODS

## Datasets

RNA sequencing data from feedlot cattle with and without BRD were obtained from the Gene Expression Omnibus (GEO) database at the National Center for Biotechnology Information (NCBI) under the accession number of GSE162156. Moreover, clinical traits of BRD were obtained from the supplementary material section of the original paper (Jiminez et al., 2021) and then filtered for useful measurements. The data included samples from the whole blood of 25 and 18 mixed-breed beef heifers with and without BRD, respectively. An Illumina HiSeq 4000 platform was used to generate 43 paired-end (2 × 100 bp) libraries that included 18 healthy and 25 BRD samples. More information about the data can be found in the original paper (Jiminez et al., 2021).

## RNA-Seq Data Analysis and Preprocessing

Quality control of the raw sequencing data was performed using FastQC[1] (version 0.11.9). Raw reads were then trimmed by removing low-quality bases and adaptor sequences using the Trimmomatic software (version 0.39) (Bolger et al., 2014) with the following options: ILLUMINACLIP: Adapter. fa:2:30:10, LEADING:20, TRAILING:20, and MINLEN:60. To confirm quality improvements, the clean reads were checked again using FastQC. Then the Hisat2 (version 2.2.1) (Kim et al., 2015) software was used for aligning the clean reads to the latest bovine reference genome (ARS-UCD1.2) with default parameters. To calculate counts of uniquely mapped reads to annotated genes based on the bovine GTF file (release 104), the python script HTSeq-count (version 0.13.5) (Anders et al., 2014) was applied using intersection-strict mode. All count files were then merged and finally, a raw gene expression matrix was created containing read counts information of all genes for all samples.

## Weighted Gene Coexpression Network Analysis

Raw gene expression matrix obtained from the previous steps was normalized to log-counts per million (log-cpm) using the "voom" function of the limma package (version 3.46.0) (Smyth, 2005). This normalization method opens access for the gene expression data generated by the RNA-seq analysis, to various computational methods, such as WGCNA (Law et al., 2014). Because the low-expressed or low-variance genes usually represent sampling noise and correlations based on these genes are not significant, genes with <1 CPM (counts per million) in at least five samples were removed. In addition, genes with a standard deviation >0.25 were selected for further analysis. Weighted gene coexpression network analysis was performed based on the functions of the WGCNA R package (version 1.70) (Langfelder and Horvath, 2008).

### Module–Trait Relationships Analysis

To identify significant highly correlated modules with clinical traits of BRD, all 43 samples (18 healthy and 25 BRD) were used for module–trait relationships (MTRs) analysis. Because the gene coexpression analysis is very sensitive to outliers, the distance-based adjacency metrics of samples was calculated and samples with a standardized connectivity < −2.5 were removed, considered as an outlier. In addition, samples and genes with >50% missing entries and genes with zero variance were identified and excluded from the WGCNA analysis. In this study, a signed weighted coexpression network was constructed in which correlation values between 0 and 1 are considered and values <0.50 are considered as negative correlation, and values >0.50 are considered as positive correlation (van Dam et al., 2017). Signed networks considers only positively correlated genes, and especially, network construction based on this method leads to more significantly enriched modules (Mason et al., 2009). Furthermore, bi-weight mid-correlation coefficient was used for the coexpression network construction since it is more robust to outliers in comparison to the Pearson correlation (Zheng et al., 2014). Briefly, a correlation matrix of expression values was constructed using pairwise bi-weight mid-correlation coefficients between all pairs of genes across the selected samples. Then, the correlation matrix at $\beta = 10$ as a soft threshold (power) was transformed into weighted adjacency matrix. Subsequently, the weighted adjacency matrix was transformed into topological overlap matrix (TOM), which considers each pairs of genes concerning all other genes by comparing their connections with all other genes in the network (interconnectedness). In other words, the genes in a module share strong interconnectedness (Zhang and Horvath, 2005; Li and Horvath, 2006; Yip and Horvath, 2007). Finally, average linkage hierarchical clustering analysis was performed by the topological overlap-based dissimilarity matrix (1-TOM) as input, and modules were identified by dynamic hybrid tree cutting algorithm. Then the modules with the highly correlated eigengenes were merged. The above steps were performed using automatic, one-step network construction and module detection function "blockwiseModules" of the WGCNA R package with the following parameters: power = 10, corType = "bicor," maxBlockSize = 12,000, networkType = "signed," TOMType = "signed," minModuleSize = 30, reassignThreshold = 0, and mergeCutHeight = 0.25. Next, in order to identify the BRD-related modules, the correlation between the clinical traits of BRD and module eigengenes (the first principal component of the expression matrix for a given module) was taken using Pearson correlation coefficient. The cutoff of significant moderately or highly correlated modules with clinical traits of BRD was defined as $p$-value < 0.05 and $0.30 < |R| < 0.50$, and $p$-value < 0.05 and $|R| > 0.50$, respectively. Moreover, gene significance (GS), which is a criterion for biological association of a gene with an interest trait was calculated for each gene through the correlation between gene expression profile and clinical traits of BRD.

### Module Preservation Analysis

In this method, based on the assumption that BRD may cause a topological change in the coexpression patterns of the healthy samples, and that nonpreserved modules between the healthy and disease samples may be biologically related to BRD, the healthy samples ($n = 18$) were selected as a reference set for construction the coexpression network and modules detection. So, after outlier detection, removing them, and set $\beta = 13$ as a soft threshold, automatic module detection function "blockwiseModules" of the WGCNA was used for a signed network construction, as well as identification of modules in healthy samples with following parameters: networkType = "signed," TOMType = "signed," corType = "bicor," mergeCutHeight = 0.25, power = 13, maxBlockSize = 12,000, minModuleSize = 30, and reassignThreshold = 0. After identifying the modules, module preservation analysis was performed using the "module Preservation" function of WGCNA R package to investigate whether the network density and connectivity patterns of the modules were preserved between the healthy and BRD samples. For this purpose, two composite preservation statistics were investigated using a permutation test (based on 200 random permutations). The first preservation composite statistic was Zsummary that was calculated from a combination of several preservation statistics, which investigated whether the mean connection strength among all genes in a module (known as network density) identified in the healthy samples remain highly connected in the disease samples and it also evaluates whether the sum of the connection strengths for a gene with other network genes (known as connectivity) in the healthy samples are similar in the disease samples (Langfelder et al., 2011). A higher value of Zsummary indicates strong preservation between conditions (healthy vs. BRD). However, Zsummary increases with increasing module size, Therefore, it is strongly dependent on the module size (Langfelder et al., 2011). The second preservation composite statistic used is medianRank, which is a module size-independent statistic; this rank-based measure relies on observed preservation statistics. Unlike Zsummary, modules with low medianRank values are highly preserved between conditions. In this study, modules with Zsummary > 10 and medianRank

< 8, 5 < Zsummary ≤ 10 and medianRank < 8, and Zsummary ≤ 5 or medianRank ≥ 8 were considered as highly-preserved, semipreserved, and nonpreserved, respectively.

## Functional Enrichment Analysis and Transcription Factors Prediction

To determine which modules are biologically related to BRD, all genes in each module were analyzed for Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways using the Enrichr web tool (Chen et al., 2013). The cutoffs of significant terms were defined as adjusted $p$-value <0.05 (correction by the Benjamini–Hochberg method). Moreover, to identify potential regulatory factors in the modules, the genes of each module were aligned to bovine transcription factors (TFs) set from the AnimalTFDB3.0 database (Hu et al., 2018).

## Hub Genes Identification and Protein–Protein Interaction Network Construction

Highly connected genes (hub genes) in a coexpression module are suitable candidates for explaining behavior and biological function of that module. In other words, highly connected intramodular hub genes have the highest degree of connection in a module and are central to modules in a network, and compared with other genes, they have more biological relevance to the module functions (van Dam et al., 2017). In this regard, module membership (MM) also known as eigengene-based connectivity $k_{ME}$ for each gene was calculated by the WGCNA R package through the correlation between the gene expression profile and the module eigengenes. Next, this criterion was used to identify hub genes in the significant highly correlated modules with clinical traits of BRD that were identified by the MTRs method and nonpreserved modules that were identified by the MP method. In fact, the MM assesses how well the genes of a module correlate with the characteristics of that module. Genes with $k_{ME} \geq 0.7$ were considered as highly connected hub genes in the respective modules. Furthermore, to investigate the connections of proteins encoded by the hub genes, Search Tool for the Retrieval of Interacting Genes (STRING) database (version 11.0) (Szklarczyk et al., 2018) was used and the protein–protein interaction (PPI) network of the hub genes was attained for further analysis.

## Hub–Hub Gene Detection and Network Visualization

For detection of the highly connected and central genes in the PPI network based on the hub genes (hub–hub genes), cytoHubba application (version 0.1) was used (Chin et al., 2014). This application is a cytoscape plugin and explores important genes and subnetworks in a given biological network such as the PPI network by several topological analysis methods including local-based and global-based methods. Local-based methods only considers the direct neighborhood of a gene, including degree (Deg), maximum neighborhood component (MNC), density of maximum neighborhood component (DMNC), clustering coefficient (CC), and maximal clique centrality (MCC) methods. Global-based methods focus on the shortest paths, such as closeness (Clo), eccentricity (EcC), radiality (Rad), bottleneck (BN), stress (Str), and betweenness (BC) methods (Chin et al., 2014). These 12 topological analysis methods were used separately to rank 60 important genes in each PPI network that were derived from the hub genes. Next, for rank aggregation of important genes lists, RankAggreg R package (version 0.6.6) was used based on cross-entropy (CE) algorithm and genetic algorithm (GA) (Pihur et al., 2009). Finally, the common genes between these two methods were considered as hub–hub genes. Significant highly correlated modules (identified by MTRs) and nonpreserved modules (identified by MP) that were biologically associated with BRD were visualized using Cytoscape (version 3.7.1) (Cline et al., 2007).

# RESULTS

## RNA-Seq Data Analysis

A summary for the RNA-seq data analysis pipeline and the steps for constructing the weighted gene coexpression network is presented in **Figure 1**. RNA-seq data included 43 samples (18 healthy and 25 BRD), and there was a mean of 31.781 million paired-end (2 × 100 bp) reads per sample. A total of 1.366 billion reads were analyzed and after trimming, a total of 1.352 billion clean reads were obtained (approximately 31.469 million clean reads per sample). On average, 95% of all clean reads were mapped to the bovine reference genome (ranging from 92 to 97%). Moreover, a mean of 82% of all clean reads were uniquely mapped to the bovine reference genome. The details about RNA-seq data, trimming, and mapping summary of all samples is provided in **Supplementary Table S1**. Finally, after applying various parameters to filter the low-expressed and low-variance genes, a total of 10,099 genes were further used in the WGCNA analysis. A list of filtered genes along with the normalized values of their expression is provided in **Supplementary Table S2**.

## Module–Trait Relationship Analysis

To prevent the negative effects of outlier samples on gene coexpression network analysis, after identifying the outliers, two samples (GSM4943645 and GSM4943656) with a standardized connectivity score < −2.5 were removed (**Figure 2A**). The weighted adjacency matrix was constructed at $\beta$ = 10 whose scale-free topology fitting index ($R^2$) was ≥0.80 (**Figure 2B**). After network construction, 12 coexpression modules (excluding grey module with 690 uncorrelated genes) were identified through hierarchical clustering and dynamic hybrid tree cutting with an average size of 784 genes. The turquoise and tan modules were the largest and smallest module with 2,592 and 72 genes each, respectively (**Supplementary Table S3**). In **Figure 3A**, a clustering dendrogram is presented in which the branches represent the modules that are labeled with a specific color by the WGCNA R package. Clinical traits related to BRD that were used in MTRs included clinical signs measurements of

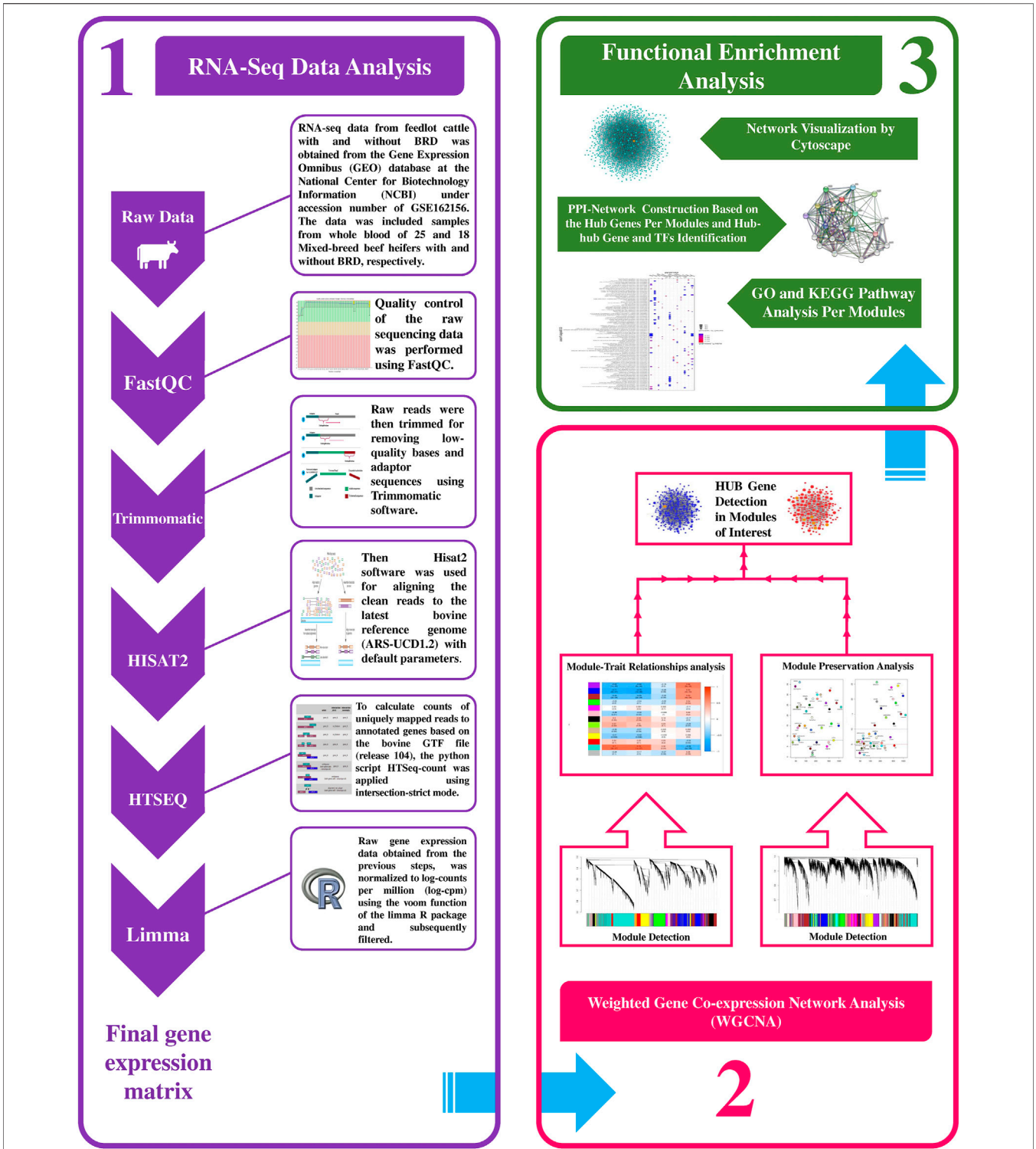


**FIGURE 1 |** Schematic pipeline for RNA-seq data analysis and weighted gene coexpression network construction in this study.

BRD such as rectal temperature (°C), haptoglobin level (g/L), respiratory rate (per min), and average daily gain. The sample dendrogram and trait heatmap of clinical traits related to BRD across all samples are presented in **Figure 3B**. The results of MTRs indicate that the rectal temperature, haptoglobin level, average daily gain, and respiratory rate have eight, eight, six, and two significant modules, respectively. Among the significant modules are as follows: MEpurple ($R$ = −0.63, $p$ = 1e−05), MEblue ($R$ =

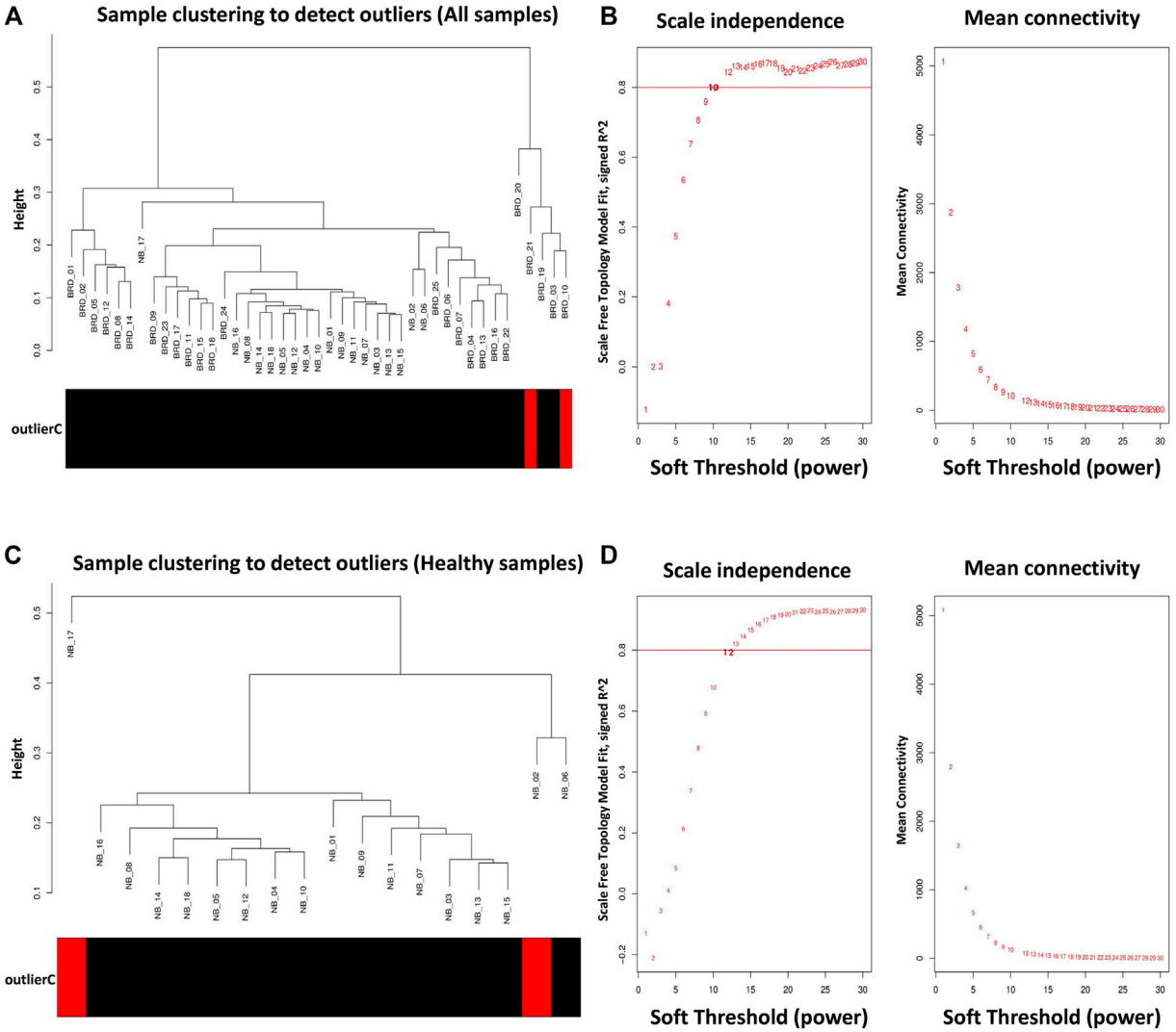


**FIGURE 2 |** Sample clustering to detect outliers and network topology analysis. **(A)** All samples except GSM4943645 and GSM4943656 were clustered and then selected for module–trait relationships analysis. **(B)** Scale-free topology fitting index (left) and mean connectivity (right) for different soft-threshold powers ($\beta$). For module–trait relationships analysis, coexpression network was constructed at $\beta = 10$ whose scale-free topology fitting index ($R^2$) was ≥0.80. **(C)** All samples except GSM4943661 and GSM4943667 were clustered and then selected for module preservation analysis. **(D)** Scale-free topology fitting index (left) and mean connectivity (right) for different soft-threshold powers ($\beta$). For module preservation analysis, co-expression network was constructed at $\beta = 13$ whose scale-free topology fitting index ($R^2$) was >0.80.

−0.71, $p$ = 2e−07), MEbrown ($R$ = −0.55, $p$ = 2e−04), and MEturquoise ($R$ = 0.7, $p$ = 3e−07) modules were significantly highly correlated and MEyellow ($R$ = −0.41, $p$ = 0.008), MEtan ($R$ = −0.32, $p$ = 0.04), MEgreenyellow ($R$ = 0.37, $p$ = 0.02), and MEpink ($R$ = −0.38, $p$ = 0.01) modules were significantly moderately correlated with rectal temperature, respectively (**Figure 3C**). Also, MEpurple ($R$ = −0.64, $p$ = 8e−06), MEblue ($R$ = −0.75, $p$ = 1e−08), MEbrown ($R$ = −0.55, $p$ = 2e−04), and MEturquoise ($R$ = 0.72, $p$ = 1e−07) modules were significantly highly-correlated and MEyellow ($R$ = −0.42, $p$ = 0.006), MEtan ($R$ = −0.31, $p$ = 0.04), MEblack ($R$ = 0.33, $p$ = 0.04), and MEpink ($R$ = −0.34, $p$ = 0.03) modules were significantly moderately correlated with haptoglobin level, respectively (**Figure 3C**). Moreover, MEpurple ($R$ = 0.64, $p$ = 8e−06), MEblue ($R$ = 0.53, $p$ = 4e−04), MEbrown ($R$ = 0.53, $p$ = 4e−04), and MEturquoise ($R$ = −0.59, $p$ = 6e−05) modules were significantly highly correlated, and MEred ($R$ = −0.34, $p$ = 0.03) and MEgreen ($R$ = 0.34, $p$ = 0.03) modules were significantly moderately correlated with average daily gain, respectively (**Figure 3C**). The MEturquoise ($R$ = 0.32, $p$ = 0.04) and MEbrown ($R$ = −0.32, $p$ = 0.04) modules were significantly moderately correlated with respiratory rate, respectively (**Figure 3C**), but no significant highly correlated modules were found for this trait. Then, the significant highly correlated modules were selected for downstream analysis. Briefly, the turquoise, blue, brown, and purple modules with module sizes of 2,592, 1,691,

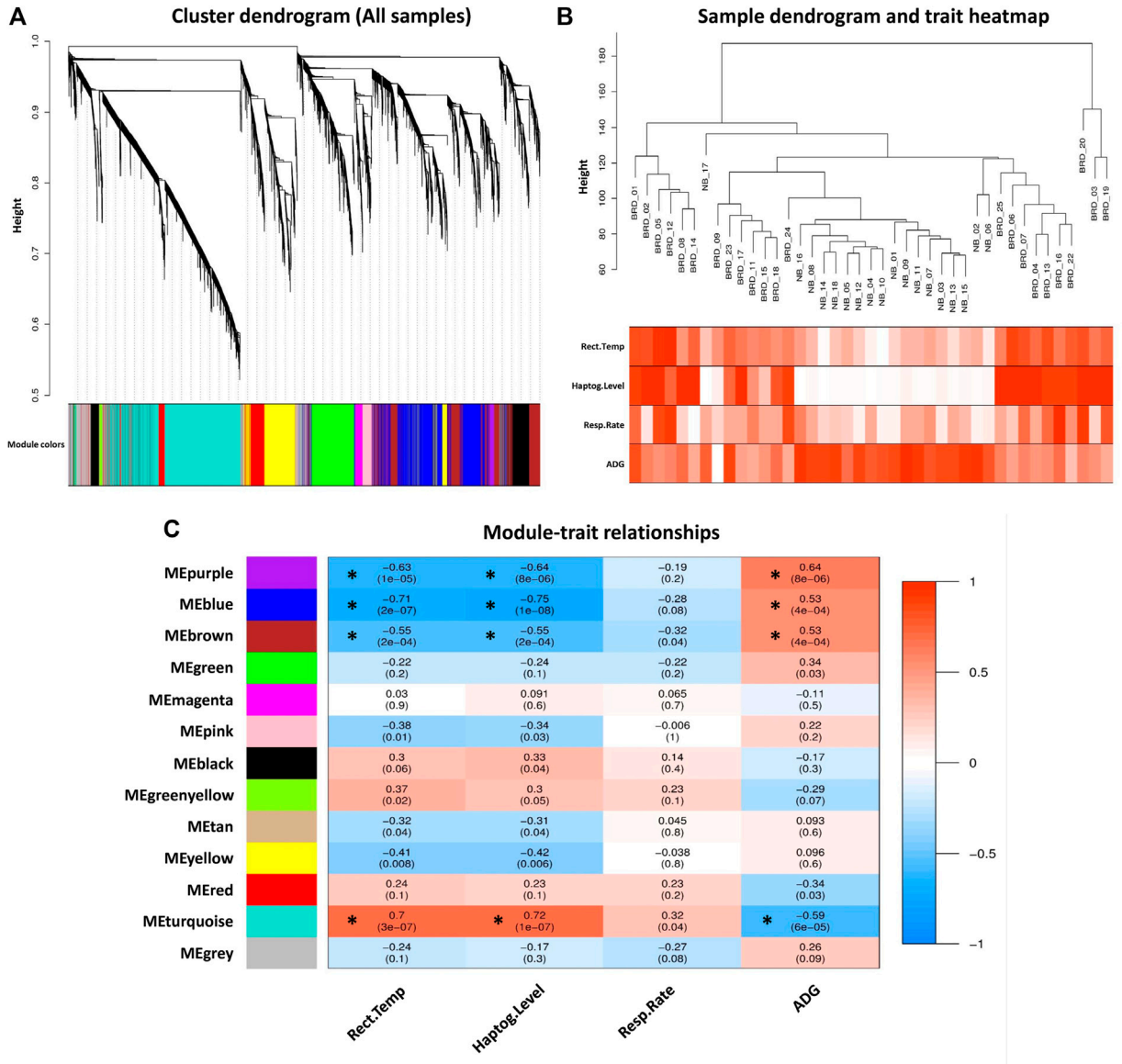


**FIGURE 3 |** Module–trait relationships analysis. **(A)** Gene hierarchical clustering dendrogram of 12 detected modules based on a dissimilarity (1-TOM) measure across all samples, the y-axis represents the coexpression distance and the x-axis represents the genes. The branches indicate the modules, and each module is marked with a separate color, the gray module encompass genes that are not assigned to any of the modules. **(B)** Sample dendrogram and trait heatmap of clinical traits related to BRD across all samples. The gradient from white to red indicates the low to high level of the respective traits in the samples. **(C)** Module–trait relationships between detected modules and clinical traits of BRD. Module–trait relationships are obtained by calculating the correlation between the traits and the module eigengenes. The red and blue colors indicate strong positive correlation and strong negative correlation, respectively. Rows represent module eigengene (ME) and columns indicate clinical traits of BRD. Rect. Temp, rectal temperature (°C); Haptog. Level, haptoglobin level (g/L); Resp. Rate, respiratory rate (per min), ADG, average daily gain. Asterisks corresponds to significant highly correlated values.

1,214, and 141 genes, respectively, were identified as significantly highly correlated modules with rectal temperature, haptoglobin level, and average daily gain (**Figure 3C**).

## Functional Enrichment Analysis of Highly Correlated Modules

In order to understand the biological performance of the significant highly correlated modules with clinical traits of BRD, functional enrichment analysis was performed and a total of 356 biological process and 129 KEGG pathways were significantly enriched in the respective modules. The turquoise module had the highest number of enriched terms and pathways, including 305 biological processes and 116 KEGG pathways. The most significant GO term and KEGG pathway in the turquoise module were "neutrophil-mediated immunity" (GO:0002446, adjusted *p*-value = 7.45E−55) and "Lysosome" (adjusted *p*-value = 5.16E−13), respectively. On the other hand, 19 biological processes and 13 KEGG pathways were

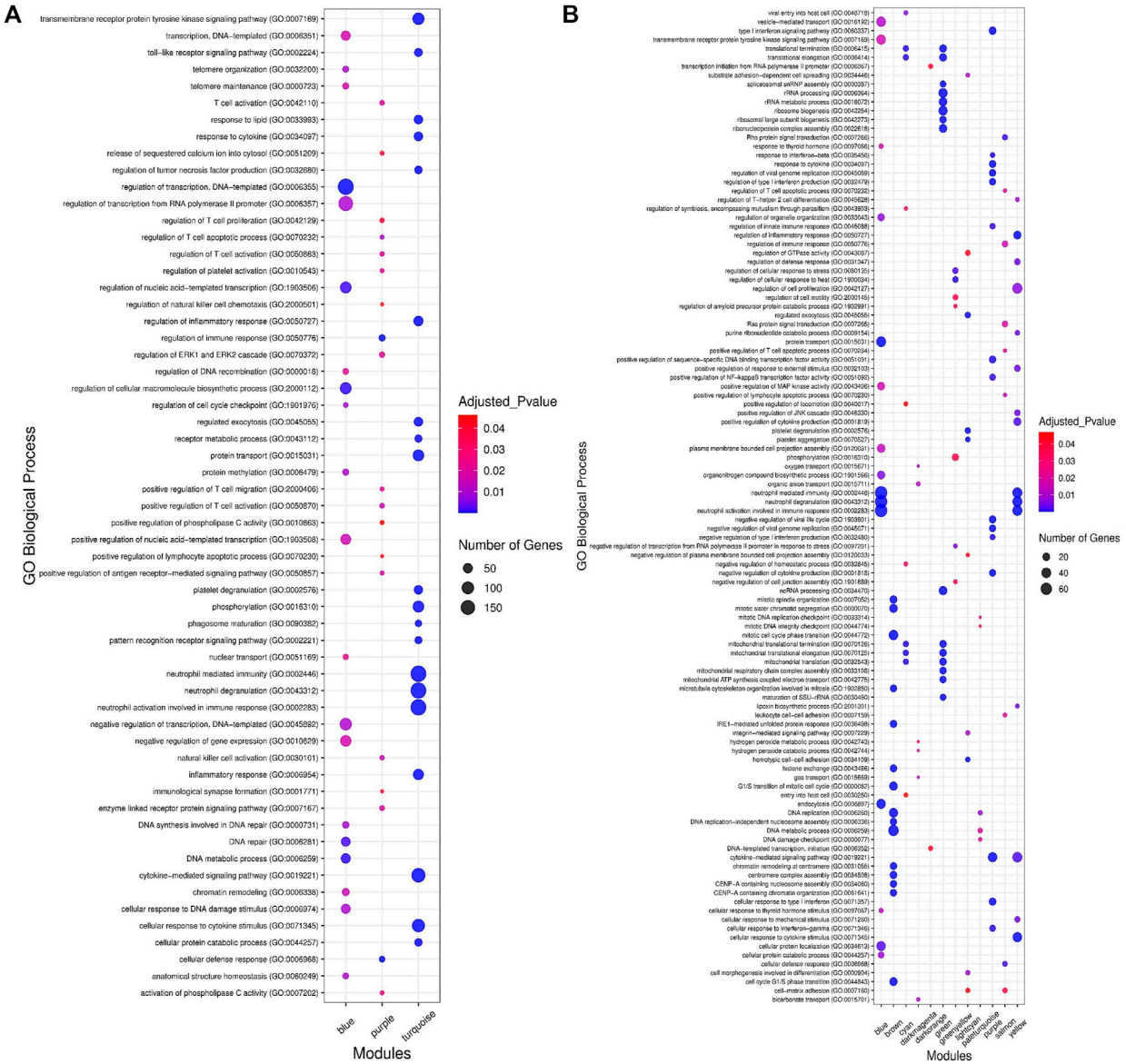


**FIGURE 4 |** Functional enrichment analysis results. **(A)** The top 20 significant biological processes for significant highly correlated and biologically related modules to bovine respiratory disease (BRD). Color and size and each point represent –log2(FDR) and number of genes for each term, respectively. **(B)** The top 15 significant biological processes for significant enriched nonpreserved modules. Color and size and each point represent –log2(FDR) and number of genes for each term, respectively.

significantly enriched in the purple module. The most significant GO term and KEGG pathway in the purple module were "cellular defense response" (GO:0006968, adjusted *p*-value = 2.27E–09) and "natural killer cell-mediated cytotoxicity" (adjusted *p*-value = 2.09E–06), respectively. Only 32 biological processes were enriched in the blue module, and no biological process or KEGG pathway was significantly enriched in the brown module. The top 20 significant biological process terms for turquoise, blue, and purple modules are presented in **Figure 4A**. Moreover, the complete information of the functional enrichment analysis for the significant highly correlated modules with clinical traits of BRD is provided in **Supplementary Table S4**. Based on the functional enrichment analysis, among the significant highly correlated modules with clinical traits of BRD, turquoise and purple modules were associated with BRD mechanisms and host immune response. To identify potential TFs that may control transcription of coexpressed genes in the modules, a total number of 100 and 11 TFs were found in the turquoise and purple modules, respectively (**Supplementary Table S5**).

## Hub and Hub–Hub Gene Detection in Highly Correlated Modules

In this study, coexpressed genes in both turquoise and purple modules (as significant highly correlated as well as biologically

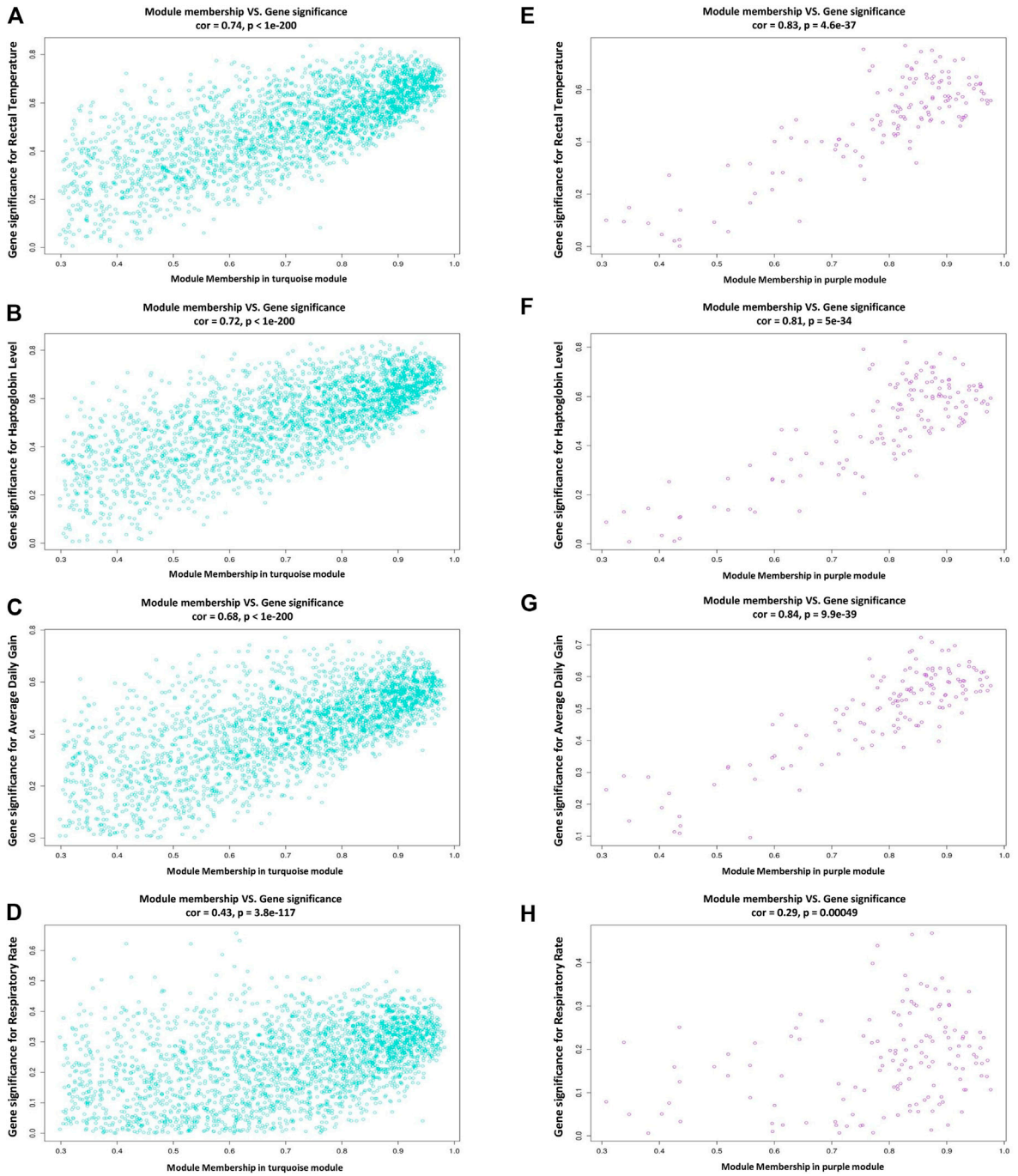


**FIGURE 5 |** Scatterplots of module membership (MM) versus gene significance (GS) plots. **(A–D)** module membership versus gene significance for rectal temperature, haptoglobin level, average daily gain, and respiratory rate in the turquoise module, respectively. **(E–H)** module membership versus gene significance for rectal temperature, haptoglobin level, average daily gain, and respiratory rate in the purple module, respectively.

related modules with BRD) were further evaluated. The MM versus GS plots of these modules are presented in **Figure 5**, which shows a strong correlation between GS and MM. In other words, the most significant genes with clinical traits of BRD are often the central genes in the respective modules. More information about GS for clinical traits of BRD can be found in **Supplementary Table S6**. Highly connected intramodular hub genes often have high levels of MM and may play an important role in BRD. So, in this regard, a total of 1,476 and 114 hub genes were identified in turquoise and purple modules, respectively (**Supplementary Table S7**). Then, the connection of proteins encoded by these hub genes in each module was examined and the PPI network related to the hub genes in

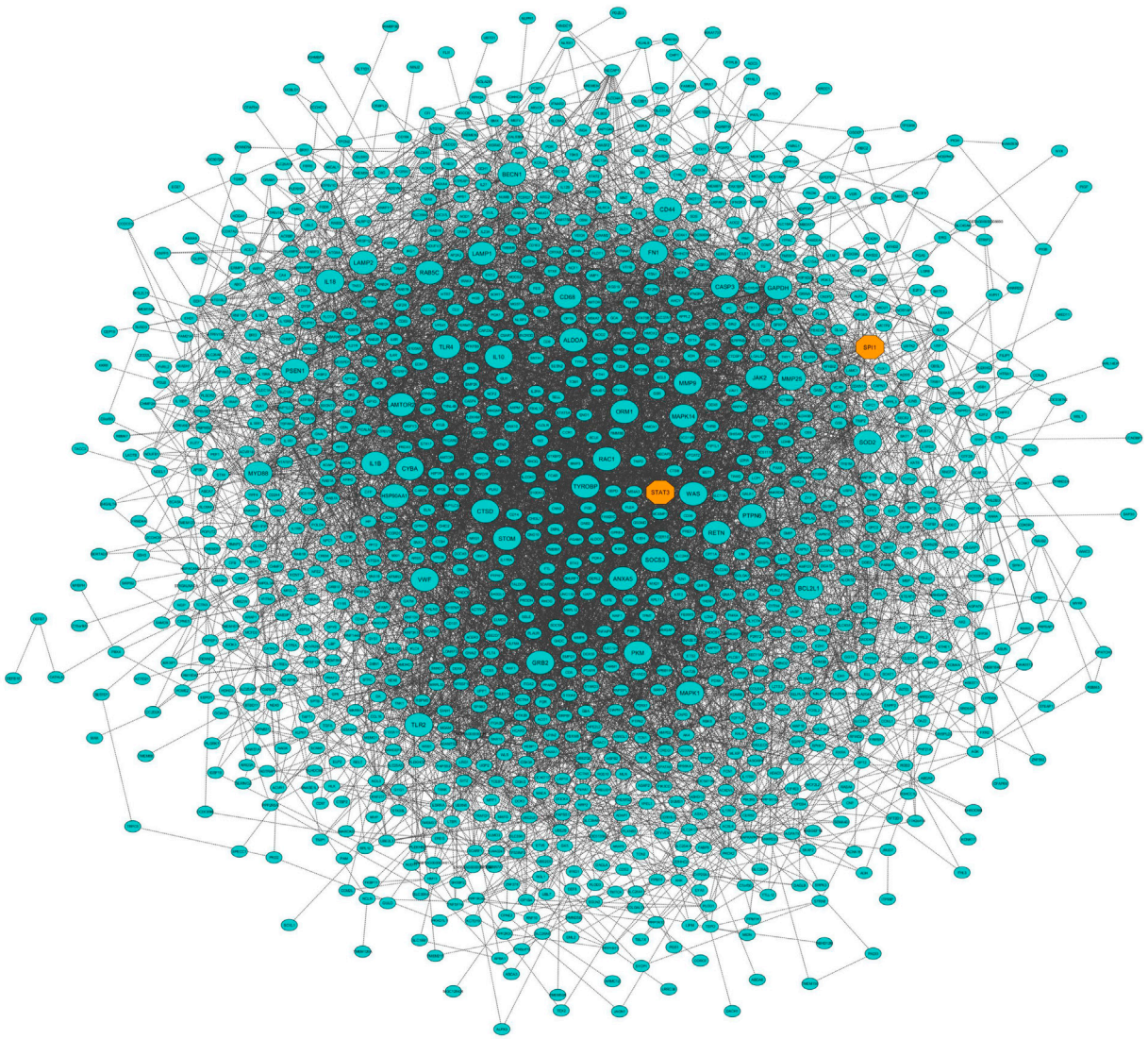


**FIGURE 6 |** Protein–protein interaction (PPI) network based on the hub genes of the turquoise module (identified by module–trait relationships method). Larger nodes and orange octagons represent hub–hub genes and transcription factors, respectively.

turquoise and purple modules was obtained. Interestingly, PPI networks based on the hub genes in the turquoise and purple modules were of high density based on the STRING database information. The PPI network based on the hub genes in the turquoise module is presented in **Figure 6**. Hub–hub genes, which were highly connected in the respective coexpression modules and also are central genes in the hub genes-based PPI networks, can be considered as prognostic and therapeutic targets in BRD development. Here, a total of 42 and 23 hub–hub genes were found in turquoise and purple modules, respectively (**Table 1**).

## Module Preservation Analysis

For module preservation analysis, healthy samples were used as a reference set to construct the weighted gene coexpression network. Two samples, GSM4943661 and GSM4943667 were identified as outliers based on the distance adjacency metrics of samples and then removed (**Figure 2C**). For network construction, the soft threshold power was set to 13 which showed high scale-free topology ($R^2$ > 0.80; **Figure 2D**). Using WGCNA, a total of 36 modules were identified in the healthy samples. The modules had different sizes ranging from 40 in the yellowgreen module to 1,163 in the turquoise module. In addition, the grey module contained 247 genes that were not assigned to any of the other modules (**Supplementary Table S8**). The identified modules in the healthy samples with different colors as branches of the hierarchical clustering dendrogram and the relationship between them are presented in **Figures 7A,B**, respectively. Next, for the module preservation analysis, we used the BRD samples as a test set to investigate whether the network density and connectivity pattern of the modules identified in the healthy samples were preserved in the BRD samples. The results of the preservation analysis showed that based on the thresholds set in the *Materials and methods* section, six modules, including lightgreen

**TABLE 1 |** List of the hub–hub genes in the turquoise and purple modules as significant highly correlated and biologically related modules with bovine respiratory disease (BRD) along with their module memberships (MM) and gene significance (GS) for rectal temperature (identified by module–trait relationships analysis).

| Module | | | | | |
|---|---|---|---|---|---|
| Turquoise | | | Purple | | |
| Genes | MM | GS | Genes | MM | GS |
| GAPDH | 0.85 | 0.56 | PRF1 | 0.97 | −0.54 |
| IL10 | 0.93 | 0.68 | KLRK1 | 0.92 | −0.71 |
| STAT3 | 0.89 | 0.65 | IL2RB | 0.95 | −0.62 |
| MAPK1 | 0.81 | 0.66 | LCK | 0.85 | −0.55 |
| CASP3 | 0.81 | 0.51 | ITK | 0.82 | −0.60 |
| LAMP1 | 0.86 | 0.56 | EOMES | 0.82 | −0.64 |
| FN1 | 0.72 | 0.45 | KLRD1 | 0.88 | −0.52 |
| MAPK14 | 0.91 | 0.64 | CD40LG | 0.81 | −0.52 |
| TLR4 | 0.93 | 0.67 | NCR1 | 0.91 | −0.58 |
| TLR2 | 0.89 | 0.50 | CCL5 | 0.92 | −0.49 |
| CD68 | 0.90 | 0.65 | LOC618565 | 0.74 | −0.36 |
| RAC1 | 0.76 | 0.37 | TBX21 | 0.93 | −0.64 |
| CD44 | 0.77 | 0.49 | CD8A | 0.88 | −0.62 |
| JAK2 | 0.87 | 0.55 | RUNX3 | 0.83 | −0.68 |
| IL1B | 0.86 | 0.50 | XCL2 | 0.74 | −0.30 |
| CTSD | 0.85 | 0.65 | CCR8 | 0.95 | −0.64 |
| MMP9 | 0.93 | 0.62 | CX3CR1 | 0.90 | −0.57 |
| GRB2 | 0.82 | 0.42 | SH2D1A | 0.86 | −0.65 |
| ANXA5 | 0.81 | 0.46 | GPR55 | 0.82 | −0.76 |
| TYROBP | 0.96 | 0.66 | CTSW | 0.92 | −0.57 |
| PTPN6 | 0.79 | 0.51 | KIR2DS1 | 0.72 | −0.34 |
| RAB5C | 0.90 | 0.57 | NKG7 | 0.93 | −0.54 |
| SOD2 | 0.93 | 0.77 | CD96 | 0.95 | −0.61 |
| BECN1 | 0.81 | 0.41 | — | — | — |
| WAS | 0.94 | 0.68 | — | — | — |
| LAMTOR2 | 0.87 | 0.75 | — | — | — |
| CYBA | 0.86 | 0.56 | — | — | — |
| SOCS3 | 0.94 | 0.66 | — | — | — |
| ALDOA | 0.93 | 0.67 | — | — | — |
| SPI1 | 0.96 | 0.61 | — | — | — |
| RETN | 0.78 | 0.47 | — | — | — |
| MYD88 | 0.93 | 0.68 | — | — | — |
| VWF | 0.77 | 0.63 | — | — | — |
| PKM | 0.97 | 0.66 | — | — | — |
| ORM1 | 0.80 | 0.62 | — | — | — |
| STOM | 0.91 | 0.52 | — | — | — |
| BCL2L1 | 0.81 | 0.70 | — | — | — |
| HSP90AA1 | 0.89 | 0.72 | — | — | — |
| MMP25 | 0.96 | 0.67 | — | — | — |
| LAMP2 | 0.91 | 0.57 | — | — | — |
| PSEN1 | 0.82 | 0.47 | — | — | — |
| IL18 | 0.81 | 0.58 | — | — | — |

*Note that the rectal temperature is one of the most important and widely used clinical signs of BRD.*

(Zsummary = 26, medianRank = 1), magenta (Zsummary = 24, medianRank = 5), pink (Zsummary = 20, medianRank = 7), grey60 (Zsummary = 19, medianRank = 2), midnightblue (Zsummary = 17, medianRank = 7), and darkred (Zsummary = 11, medianRank = 4), were identified as highly preserved modules, and one module, yellowgreen (Zsummary = 5.8, medianRank = 7) was identified as a semipreserved module (**Figure 7C**). Interestingly, in agreement with our hypothesis, of the 36 modules identified, 29 modules were non-preserved between healthy and BRD samples, indicating that their connectivity pattern and topological structure have been affected and changed by the BRD (**Figure 7C**). Skyblue (Zsummary = 1.5, medianRank = 36), darkmagenta (Zsummary = 2.1, medianRank = 34), and darkolivegreen (Zsummary = 3.6, medianRank = 24) modules were identified as the most non-preserved modules between healthy and BRD samples, respectively (**Supplementary Table S9**).

## Functional Enrichment Analysis of Highly Preserved, Semipreserved, and Nonpreserved Modules

To understand the biological significance and the functional difference between highly preserved, semipreserved, and nonpreserved modules, all the 36 modules were subjected to GO terms and KEGG pathway analysis, and a total of 521 biological processes and 158 KEGG pathways were significantly enriched. Enrichment analysis of four highly preserved modules that included magenta, pink, grey60, and midnightblue revealed a total of 108 and 27 biological processes and KEGG pathways, respectively. Among them, the magenta module with 37 biological processes and 17 KEGG pathways was identified as the most significant enriched module. The other two highly preserved modules, including lightgreen and darkred, showed no biological process or KEGG pathway enrichment. Moreover, one biological process was enriched in the yellowgreen module as the only semipreserved module. The complete information of the functional enrichment analysis for the highly preserved and semipreserved modules is provided in **Supplementary Table S10**. On the other hand, among the nonpreserved module, 13 modules including blue, brown, cyan, darkmagenta, darkorange, green, greenyellow, lightcyan, paleturquoise, purple, red, salmon, and yellow were enriched in at least one biological process or KEGG pathway (significantly enriched nonpreserved modules). Furthermore, functional enrichment analysis of the significantly enriched nonpreserved modules identified a total of 412 biological processes and 131 KEGG pathways. The top 15 significant biological process terms for significantly enriched nonpreserved modules are presented in **Figure 4B**. In addition, the complete information of the functional enrichment analysis for the nonpreserved modules is provided in **Supplementary Table S11**. Based on the enrichment results, among the significantly enriched nonpreserved modules, six nonpreserved modules including blue, greenyellow, purple, red, salmon, and yellow were associated with pathogenic mechanisms of BRD and immune response. Also, based on the bovine transcription factors set from the AnimalTFDB3.0 database, a total number of 37, 18, 25, 55, 26, and 39 TFs were identified in blue, greenyellow, purple, red, salmon, and yellow modules, respectively (**Supplementary Table S12**).

## Hub and Hub–Hub Gene Identification in Nonpreserved Modules

Six nonpreserved modules were identified as potential candidate modules and biologically related to BRD based on the module preservation and functional enrichment analysis, respectively. Then, MM measures were used to identify intramodular hub genes as important and central genes in these modules. A total number of 326, 251, 216, 145, 131, and 99 hub genes were

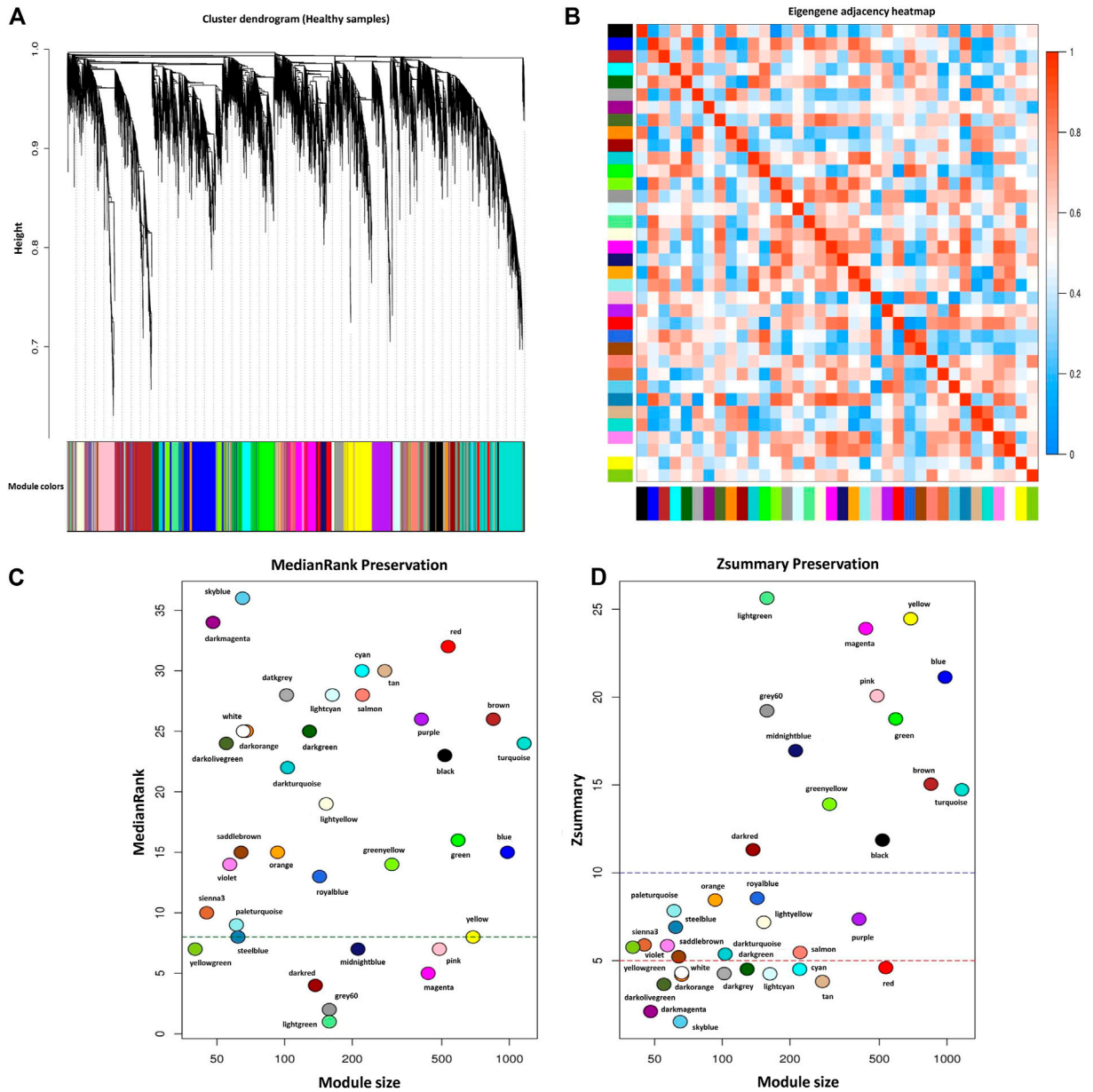


**FIGURE 7 |** Module preservation analysis. **(A)** Gene hierarchical clustering dendrogram of 36 detected modules based on a dissimilarity (1-TOM) measure across healthy samples as reference set, the y-axis represents the coexpression distance and the x-axis represents the genes. The branches indicate the modules, and each module is marked with a separate color, the gray module encompass genes that are not assigned to any of the modules. **(B)** Eigengene adjacency heatmap indicate relationship among all the modules. **(C)** The medianRank preservation statistics of the modules. The y-axis and the x-axis represent medianRank values and module size, respectively. Each point indicates a module labeled by a respective color. The green dashed line represents the medianRank threshold (medianRank ≥8). **(D)** The Zsummary preservation statistics of the modules. The y-axis and the x-axis represent Zsummary values and module size, respectively. Each point indicates a module labeled by a respective color. The red dashed line represents the Zsummary threshold (Zsummary ≤5). Modules with Zsummary ≤5 or medianRank ≥8 were considered as nonpreserved between healthy and BRD conditions.

detected in blue, yellow, red, purple, greenyellow, and salmon modules, respectively. The complete list of the hub genes in the nonpreserved modules can be found in **Supplementary Table S13**. Hub genes-based PPI networks extracted from the STRING database identified 303 nodes (proteins) and 2,942 edges (interactions) for the blue module, 204 nodes and 1,085 edges for the yellow module, 183 nodes and 982 edges for the red module, 131 nodes, and 1,905 edges for the purple module, 121 nodes and 716 edges for the greenyellow module, and 91 nodes and 345 edges for the salmon module indicating high connection density of proteins encoded by the genes of these modules. **Figure 8** represents the PPI network of the purple module as the nonpreserved and potential biologically BRD-related module. Additionally, based on the PPI networks obtained by hub genes, a total number of 48, 51, 48, 49, 27, and

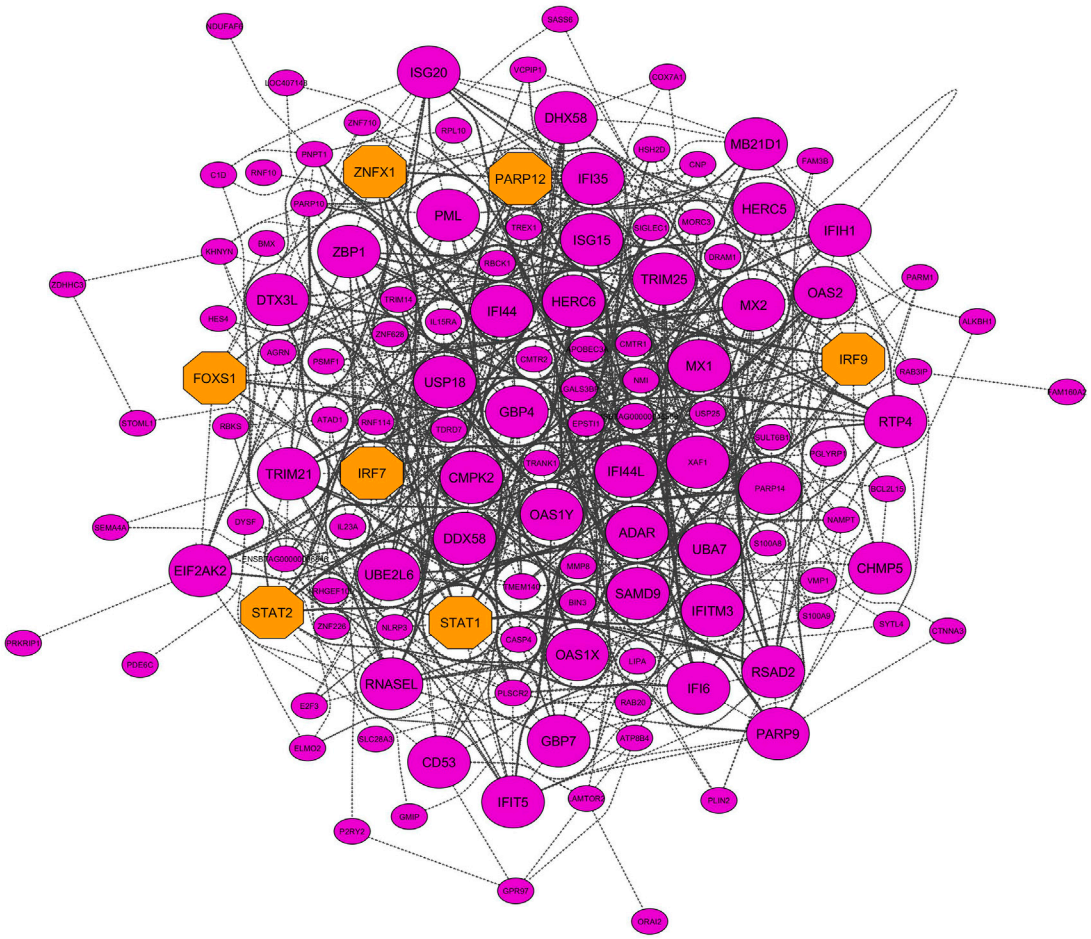


**FIGURE 8 |** PPI network based on the hub genes of the purple module (identified by module preservation method). Larger nodes and orange octagons represent hub–hub genes and transcription factors, respectively.

19 hub–hub genes were found in the blue, yellow, red, purple, greenyellow, and salmon modules, which can be potential biomarkers and candidate disease genes in the etiology and the diagnostics of BRD (**Table 2**). PPI networks of significant highly correlated modules (identified by MTRs) and nonpreserved modules (identified by MP) that were biologically associated with BRD are available in **Supplementary Figure S1**.

## DISCUSSION

BRD is a multifactorial disease that results from the interaction of environmental stressors and infectious agents of BRDC (Gagea et al., 2006). BRDC includes the viral pathogens such as bovine respiratory syncytial virus (BRSV), bovine parainfluenza type 3 virus (BPIV-3), bovine viral diarrhea virus (BVDV), bovine coronavirus (BCV), and bovine herpes virus type 1 (BHV-1) that affect upper respiratory system and also contain the bacterial pathogens *Trueperella pyogenes*, *Mycoplasma bovis*, *Pasteurella multocida*, *Histophilus somni*, *Bibersteinia trehalosi*, and *Mannheimia haemolytica*, which can affect the lower respiratory system (Caswell, 2014; Kirchhoff et al., 2014). Despite numerous studies, BRD is still the most common disease and the leading cause of morbidity and mortality in the cattle industry (Taylor et al., 2010). Understanding the molecular mechanisms involved in the bovine immune response to BRD is necessary given the persistence of the disease in recent years. Combining high-throughput technologies with various computational methods based on the network approach, can provide an exceptional opportunity to better understand the pathological processes of diseases and the molecular mechanisms of the host immune responses (Kadarmideen and Watson-Haigh, 2012). In this study, we combined gene expression matrix obtained by RNA-seq data analysis with two co-expression network-based methods of WGCNA, module–trait relationships, and module preservation analysis, to identify potential gene modules and candidate genes involved in molecular processes induced by BRD.

### Module–Trait Relationships Analysis

Module–trait relationship analysis identified four significant highly correlated modules including turquoise, purple, blue,

**TABLE 2 |** List of the hub–hub genes of the nonpreserved and biologically BRD-related modules that were identified by module preservation analysis.

| Modules | | | | | |
|---|---|---|---|---|---|
| **Blue** | **Greenyellow** | **Purple** | **Red** | **Salmon** | **Yellow** |
| LAMP1 | STAG2 | ISG15 | RAB11FIP2 | IL2RB | C5AR1 |
| LAMP2 | PIK3CA | STAT1 | TNF | KLRK1 | IL1B |
| TLR4 | SMURF2 | IFIH1 | MAU2 | PRF1 | UBA52 |
| PSAP | PTEN | RTP4 | UNKL | ITGAL | PTAFR |
| ANXA5 | ADAM10 | USP18 | TAF1 | CCL5 | IL15 |
| C3 | TSPAN13 | DDX58 | NR2C2 | CCR5 | GRB2 |
| PSEN1 | SNX27 | IRF7 | LCK | GZMA | CD68 |
| RAP1B | KLHL11 | MX1 | CD2 | NCR1 | PLEK |
| CTSB | PIP5K1B | PARP9 | ARIH2 | RUNX3 | SOCS3 |
| STOM | SOS2 | DHX58 | ZFYVE20 | S1PR5 | NFKBIA |
| ACTR2 | SNX13 | IFI35 | TSC1 | CX3CR1 | MYD88 |
| BECN1 | MGAT4A | RSAD2 | ITK | NKG7 | ANXA1 |
| CAT | YES1 | IFI44 | SMC5 | IL12RB2 | FCAR |
| NPC2 | GSK3B | IFI44L | LCP2 | GZMB | CYTH4 |
| CD86 | SGK3 | UBA7 | PHF8 | CCR8 | CRKL |
| ATG7 | UBE2R2 | UBE2L6 | JAK3 | PDCD1 | HECW2 |
| VPS35 | TMEM30A | EIF2AK2 | RIC1 | TRPM2 | RAF1 |
| PRCP | HECTD1 | IRF9 | SETDB1 | TMEM63A | IL18RAP |
| RAB1A | TSPAN33 | ISG20 | FMR1 | GZMH | GADD45B |
| RAB5A | RNF217 | MX2 | PLCG1 | — | FAS |
| GAA | ROCK1 | XAF1 | CHD3 | — | CNR2 |
| MGST1 | TAB2 | PARP14 | FBXO4 | — | CCR1 |
| CCR2 | PDP1 | PARP12 | FBXL20 | — | HCAR3 |
| CD59 | WAPAL | IFIT5 | PIK3CD | — | SNX18 |
| CTSC | WAC | STAT2 | FBXO21 | — | SGK1 |
| ACTR1A | PI4K2B | TRIM21 | TIA1 | — | IFNAR2 |
| TLR7 | SERPINE2 | OAS1X | TNRC6A | — | VNN2 |
| LYZ | — | HERC6 | TLR3 | — | DDIT3 |
| GM2A | — | CMPK2 | POGZ | — | WDFY3 |
| CST3 | — | ZBP1 | PDPR | — | IFNGR2 |
| TNFRSF1B | — | DTX3L | CRAMP1L | — | PPP2R5A |
| CAPZA1 | — | ZNFX1 | RERE | — | SELL |
| RAB6A | — | IFITM3 | RAB11FIP4 | — | KDM4B |
| ARF1 | — | RNASEL | CD6 | — | NCF1 |
| FTL | — | SAMD9 | ZC3H11A | — | LOC407171 |
| APLP2 | — | GBP4 | TNFRSF10D | — | ARRDC4 |
| PECAM1 | — | TRIM25 | UBP1 | — | TARM1 |
| ATF6 | — | MB21D1 | SLC37A3 | — | VASP |
| CSF2RB | — | OAS1Y | IKBKE | — | SLC2A3 |
| RAB18 | — | HERC5 | STAT5A | — | BST1 |
| SPTLC1 | — | ADAR | NAA16 | — | MCL1 |
| KTN1 | — | GBP7 | ZBTB43 | — | NFAM1 |
| SHISA5 | — | OAS2 | CAMSAP1 | — | TNIP1 |
| MOSPD2 | — | IFI6 | TWF1 | — | GADD45A |
| CD163 | — | FOXS1 | ABI2 | — | NUDT3 |
| ATP6V1A | — | PML | VAMP5 | — | MEFV |
| IFNAR1 | — | CD53 | CDC7 | — | MXD1 |
| CAPN2 | — | CHMP5 | UPB1 | — | LONRF3 |
| — | — | GPR97 | — | — | KDM6B |
| — | — | — | — | — | ICAM3 |
| — | — | — | — | — | TCN1 |

and brown with clinical traits of BRD. Functional enrichment analysis indicated that three of the four (75%) identified modules including blue, purple, and turquoise modules had significant enriched terms and pathways. This shows the power and high accuracy of signed networks in separating the modules from each other and identifying more significantly enriched terms or pathways in the modules. The significant enriched terms in the blue module were mostly related to basic cellular activities including "tRNA transport," "transcription DNA-templated," "DNA metabolism," "macromolecule biosynthetic," and "cell cycle." On the other hand, the turquoise and purple modules had many biological processes and KEGG pathways closely related to the BRD mechanisms. So, our focus on the turquoise and purple modules (identified by MTRs method) as significant highly correlated and key biologically related modules to BRD.

Coregulated genes in the turquoise module were closely related to the mechanisms of the innate immune system as the first defense line against BRD and this shows the great importance of this module during BRD development. We also found a number of enriched terms related to the adaptive immune system in this module. After infection with various viral or bacterial pathogens, extensive complex interactions begin between the host and the pathogen (Kumar et al., 2011). Pathogens produce various molecules known as pathogen-associated molecular pattern (PAMPs) to continue their life cycle and pathogenic activity (Janeway and Medzhitov, 2002; Tang et al., 2012). These PAMPs are recognized by pattern recognition receptors (PRRs) in the host, which are proteins expressed by key innate immune system cells such as neutrophils, macrophages, monocytes, dendritic cells, and epithelial cells (Medzhitov, 2007; Yuki and Koutsogiannaki, 2021). Upon recognition of PAMPs by PRRs, a cascade of signaling pathways is induced, leading to an inflammatory response and subsequent rapid response of the innate immune system to eliminate the pathogens (Lee and Kim, 2007; Kawai and Akira, 2010). In this regard, the turquoise module had important KEGG signaling pathways as well as their downstream biological processes associated with PRRs and inflammatory response including "Toll-like receptor signaling pathway," "MyD88-dependent Toll-like receptor signaling pathway," "TRIF-dependent Toll-like receptor signaling pathway," "C-type lectin receptor signaling pathway," "NOD-like receptor signaling pathway," "NF-kappa B signaling pathway," and "MAPK signaling pathway."

Toll-like receptors (TLRs) are the most important signaling maker PRRs and act as the primary sensors of pathogens (Iwasaki and Medzhitov, 2004; Akira et al., 2006). The Toll-like receptor signaling pathway is activated through the recognition of PAMPs by membrane and cytoplasmic TLRs (Kumar et al., 2009; Blasius and Beutler, 2010; Heidarzadeh et al., 2020). The TLR signaling pathway is divided into two distinct pathways, including the MyD88-dependent and the TRIF-dependent signaling pathway, depending on the type of TLR sensitized and subsequently equipped adapters (Takeuchi and Akira, 2010). All TLRs family (*TLR1* to *10*), except *TLR3*, activate the MyD88-dependent Toll-like receptor signaling pathway, which activates the NF-kappa B signaling pathway and MAPK signaling pathway to produce proinflammatory cytokines and chemokines (Takeda and Akira, 2005; Lin et al., 2010; Kawasaki and Kawai, 2014). On the other hand, *TLR3* as well as *TLR4* activate the TRIF-dependent Toll-like receptor signaling pathway, which leads to induce production of proinflammatory cytokines and type I interferons by activating *NF-κB* and *IRF3/IRF7* transcription factors (Takeda and Akira, 2005; Lin et al., 2010; Kawasaki and Kawai, 2014). Various studies in cattle diseases such as bovine tuberculosis (Nalpas et al., 2015), Johne's disease (Ferwerda et al., 2007; Wang et al., 2019), endometritis (Turner et al., 2014), and mastitis (Luoreng et al., 2018; Islam et al., 2020) have reported the TLR-signaling pathway being induced during these diseases. A previous study also reported that the Toll-like receptor signaling pathway was activated during different challenges with a group of BRDC, including BoHV-1, BRSV, BVDV, *Mannheimia haemolytica*, and *Pasteurella multocida* (Tizioto et al., 2015).

C-type lectin receptor signaling pathway is another pattern recognition receptor related pathway that is activated by CLRs membrane receptors by identifying different carbohydrates such as mannose, glucan, and fucose in viruses, bacteria, and fungi and activates MAP kinases, the transcription factor *NF-AT*, and *NF-κB* that eventually induces the production of proinflammatory cytokines (Geijtenbeek and Gringhuis, 2009; Takeuchi and Akira, 2010). Recent studies have examined the importance of C-type lectin receptors in human infectious diseases (Lugo-Villarino et al., 2018; Zhao et al., 2019). NOD-like receptors are cytosolic receptors that can detect a wide range of bacteria, viruses, and other pathogens that enter the cytoplasm (Franchi et al., 2009). These receptors, like Toll-like receptors activate the NF-kappa B and MAPK signaling pathways, regulate the production of inflammatory cytokines such as *IL-1β*, and can also induce apoptosis (Chen et al., 2009; Banse et al., 2013).

The NF-kappa B signaling pathway is a key pathway that acts as a major mediator in inflammatory responses. Activation of the *NF-κB* transcription factor induces the transcription of many genes that encode proinflammatory cytokines and chemokines such as *IL-1β*, *IL-6*, *TNF-α*, *IL-12p40*, and *cyclooxygenase-2* (Oeckinghaus and Ghosh, 2009; Liu et al., 2017). In addition, a study investigated the effect of SH protein expressed by the BRSV genome on the lack of *NF-κB* phosphorylation in the host, which reduces the production of proinflammatory cytokines and thus modulates the immune system (Pollock et al., 2017). The MAPK signaling pathway is another key mediator pathway during inflammation that regulates cytokine production by phosphorylating and activating certain kinases (Salojin and Oravecz, 2007). Several transcriptomics and proteomics studies have reported MAPK signaling pathway activity during BRD (Tizioto et al., 2015; Behura et al., 2017; Liyang et al., 2021).

Furthermore, other important immune-related terms identified in the turquoise module include "JAK-STAT signaling pathway," "TNF signaling pathway," "PI3K-Akt signaling pathway," "regulation of interferon-gamma production," and "positive regulation of interleukin-8 secretion." The JAK-STAT signaling pathway is one of the most important intracellular signaling pathways that regulates communication between cytokine transmembrane receptors and the nucleus and is involved in many biological processes in the body, such as immune regulation, cell differentiation/proliferation, apoptosis, and keep homeostasis in inflammatory conditions (O'Shea et al., 2012; O'Shea et al., 2015; Xin et al., 2020). This pathway, mediates cytokine responses through binding of cytokines such as *IL-6/12/17/23*, as well as type I (alpha and beta), and II (gamma) interferons to their respective receptors at the cell surface and activation of STATs transcription factors to regulate their target genes in the nucleus (Charles Jay and Eric, 2009). Interferons by activating this pathway, can cause antiviral conditions (Horvath, 2004b; a). BoHV-1, as an infectious agent of BRDC, has the ability to suppress the host immune system by expressing the *UL41* gene and subsequently increase its viral replication. The *UL41* blocks the JAK-STAT signaling pathway by suppressing *STAT1* expression. Thus, this

virus has the ability to deal with the antiviral condition resulting from the JAK-STAT signaling pathway (Ma et al., 2019). On the other hand, Ma et al. (2020) reported that *Bta-miR-2890*, by directly targeting BoHV-1 *UL41*, increases *STAT1* and *JAK1* expression and, thus, opens the JAK-STAT signaling pathway as well as prevent viral replication. This suggests the importance of JAK-STAT signaling pathway during viral infections.

The PI3K–Akt signaling pathway is a key pathway in all mammalian cells that is involved in several processes, such as cell growth, migration, proliferation, and metabolism as well as the role of this pathway in regulatory T-cell development and memory CD8 T-cell differentiation, has been reported (Kim and Suresh, 2013; Pompura and Dominguez-Villar, 2018). It has also been suggested that some genes in the PI3K–Akt signaling pathway may play an important role in eliciting downstream cascades in lung lesions during BRD (Behura et al., 2017). Tumor necrosis factor (TNF) is a proinflammatory cytokine that is mainly secreted by macrophages and alerts other cells during the inflammatory response. TNF is also known to be a major regulator of the production of proinflammatory cytokines and has been considered as a therapeutic target for the treatment of some inflammatory diseases such as rheumatoid arthritis and inflammatory disease (Liu, 2005; Parameswaran and Patial, 2010). The TNF signaling pathway plays an important role in the control of inflammation, immunity, and cell survival (Rana et al., 2019). It has been shown that one of the strategies for bacterial survival and immunosuppression by *Mycoplasma bovis* is to inhibit *TNF-α* and interferon-gamma production (Mulongo et al., 2014).

Interleukin-8 is a chemokine, that is expressed in various cells especially in macrophages. *IL8* is responsible for the induction of chemotaxis and causes guided migration of neutrophils to the site of infection (Remick, 2005). An increase in *IL8* expression has been observed in challenges with BRSV and *Mannheimia haemolytica* as well as a significant association between increased expressions of *IL8* with lung lesions. Therefore, it has been suggested that *IL8* antagonist drugs be used to prevent inflammatory lung lesions (Caswell et al., 1998; Malazdrewich et al., 2001; Singh et al., 2011; Redondo et al., 2014).

In response to signaling proinflammatory cytokines and chemokine such as *IL8*, neutrophils are the first cells to migrate from the blood to the infection site (Kolaczkowska and Kubes, 2013). These cells play an important role in killing extracellular pathogens through phagocytosis (Nordenfelt and Tapper, 2011; DeLeo and Allen, 2020). Neutrophil-related biological processes and KEGG pathways in the turquoise module included "neutrophil mediated immunity," "neutrophil degranulation," "neutrophil activation involved in immune response," and "neutrophil extracellular trap formation." However, neutrophils play an important role in the pathogenesis of BRD by destruction and damaging lung tissue during infection (McGill and Sacco, 2020). Moreover, neutrophils release their nuclear DNA and related proteins in the extracellular environment through NETosis, a unique form of cell death that leads to the formation of neutrophil extracellular traps (McGill and Sacco, 2020). NETs trap and kill bacteria, fungi, viruses, and parasites. In addition to their antimicrobial role, NETs can also play a role in the pathogenesis of inflammatory diseases (Papayannopoulos, 2018). Furthermore, evidence suggests that NETs play a role in the host defense in response to *Mannheimia haemolytica* and *Histophilus somni* infection (Aulik et al., 2010; Hellenbrand et al., 2013). The findings also indicate that *Mycoplasma bovis*, through its endonucleases, is able to digest NETs and escape the immune system (Gondaira et al., 2017). However, Cortjens et al. (2016) showed that NETs has the ability to trap respiratory syncytial virus (RSV), but their excessive accumulation leads to airway obstruction, which can contribute to RSV pathogenesis (Cortjens et al., 2016).

We also identified several microbicidal mechanisms, including "Phagocytosis," "Fc gamma R-mediated phagocytosis," "Phagosome," "Lysosome," and "phagosome maturation" in the turquoise module. In the immune system of multicellular organisms, phagocytosis is a cellular process for the elimination of pathogens and dead cell debris, and is an essential for maintaining tissue homeostasis (Flannagan et al., 2012). Fc gamma R (Fc$\gamma$R) receptors are glycoproteins found on the surface of immune cells such as neutrophils, macrophages, and natural killer cells that stimulate phagocytosis through antigen-binding IgG antibodies (Rosales, 2017). Phagocytosis involves several stages. After ingestion of pathogens or foreign particles by immune cells, the ingested particles become specialized vacuoles and distinct organelles called phagosomes. During the stage called phagosome maturation, the lysosome fuses with the phagosome membrane, and its contents are poured into the phagosome, and an organelle called phagolysosome is formed. This newly formed organelle contains enzymes that can break down the digested particles (Canton, 2014; Levin et al., 2016; Uribe-Querol and Rosales, 2020). Because of the importance of phagocytosis pathway in inducing an innate and adaptive immune response, pathogens use specific strategies to control and suppress phagocytes (Srikumaran et al., 2007). *Mannheimia haemolytica* and *Pasteurella multocida* use their toxins and extracellular components to kill phagocytes, thus preventing phagocytosis and subsequently releasing the reactive oxygen metabolite contents of the phagocytes, which exacerbates pulmonary inflammation (CUSACK et al., 2003). Research also shows that *Mycoplasma bovis* has the ability to survive long-term in necrotic lung lesions and phagocytic cells by overcoming phagocytosis (Kleinschmidt et al., 2013). Moreover, virulent isolates of *Histophilus somni* have the ability to survive in phagocytic cells by interfering with phagosome–lysosome maturation (Pan et al., 2018). In addition to the bacterial agents of the BRDC, viruses such as BVDV, PIV3, and BRSV have the ability to inhibit phagocytosis by macrophages (Bell et al., 2021).

The turquoise module also identified some of the major pathways associated with programmed cell death, including "Apoptosis" and "Necroptosis." Apoptosis is the first type of programmed cell death that clears cells infected with pathogens (especially viruses) to prevent them from replication and spread (Amarante-Mendes et al., 2018). For example, a previous study showed that apoptosis of BRSV-infected epithelial bronchial cells is an effective way to clear the virus (Viuff et al., 2002) as well as it has been suggested that apoptosis may play a role in modulating

airway inflammation during BRSV infection (Cristina et al., 2001). On the other hand, the BHV-1 uses the products of the *LR* gene to prevent apoptosis of infected cells and can therefore proliferate sufficiently in infected cells (Geiser and Jones, 2005; Srikumaran et al., 2007). Moreover, BHV-1 inhibits the efficient immune response by infecting $CD41^+$ T-cells and inducing apoptosis in them (Jones and Chowdhury, 2010). Bacterial pathogens also have the ability to survive and manipulate intracellular mechanisms of host cells to escape the immune system. A recent study showed that *Mycoplasma bovis* has the ability to inhibit the apoptosis of infected bovine alveolar macrophages by increasing the expression of several anti-apoptotic genes (Maina et al., 2019). Necroptosis is a programmed form of inflammatory cell death (necrosis) that actively causes cell death of infected cells or the propagation of danger signals to stimulate the immune system. However, uncontrolled necroptosis may lead to the pathogenesis of inflammatory diseases. For instance, the findings show that RSV, through the activation of necroptosis, lead to neutrophilic cell death (Muraro et al., 2018). Bedient et al. (2020) showed that the RSV causes cell death of alveolar macrophages through various cell death mechanisms such as necroptosis, pyroptosis, and apoptosis, which can contribute to the pathogenesis of the RSV and exacerbate inflammation in the lungs (Bedient et al., 2020).

Interestingly, we identified the terms "regulation of nitric oxide biosynthetic process," "positive regulation of nitric oxide metabolic process," and "positive regulation of nitric oxide biosynthetic process" in the turquoise module. Nitric oxide (NO) is a natural molecule with antimicrobial properties that is produced by most mammalian cells (Crepieux et al., 2016). Several studies have shown the antibacterial and antiviral properties of NO during BRD, and it has also been demonstrated that NO therapy, as a non-antibiotic-based treatment can be a safe and effective method to control BRD (Regev-Shoshani et al., 2013; Regev-Shoshani et al., 2014). Another study demonstrated that NO reduced levels of proinflammatory cytokines, such as *IL-1β* and *TNF*, thereby limiting inflammation during BRD. On the other hand, NO increases the ability of the host to detect pathogens by increasing the expression of *TLR4* gene (Sheridan et al., 2016).

We also identified some pathways in the turquoise module associated with the adaptive immune system such as "B-cell receptor signaling pathway", "positive regulation of activated T-cell proliferation," "Th1/Th2 cell differentiation," and "Th17 cell differentiation." B cells are vital cells for the humoral response, and research shows that the B-cell receptor signaling pathway is one of the most important pathways in the immune system of animals to respond to infection with viral agents of BRDC (Tizioto et al., 2015). T cells are essential cells for cell-mediated immunity and include several subtypes that play a variety of roles, including killing virus-infected cells, secreting interferon-gamma, and other cytokines (Platt et al., 2006). The importance and application of helper and cytotoxic T cells in viral clearance in the first and second challenges with BVDV have been investigated. The results show that these T cells in the second challenge with BVDV have the ability to quickly clear the virus (Silflow et al., 2005). Furthermore, T cells have been reported to be critical for the response to BRSV infection (Gaddum et al., 2003).

Intramodular hub genes (especially hub–hub genes) are highly correlated with the biological function of the module. In this regard, hub–hub genes identified in the turquoise module, such as *FN1*, *MAPK14*, *PSEN1* (Neupane et al., 2018), *IL-1β*, *IL-18* (Taylor et al., 2014), *MYD88* (Dubbert et al., 2013), *SOD2* (Hofstetter and Sacco, 2020), *CTSD* (Gray et al., 2019), *JAK2* (Amat et al., 2019; Chao et al., 2019), *CD68* (Lee et al., 2009), and *TLR2* (Mariotti et al., 2009; Tizioto et al., 2015) have been reported as important genes in previous BRD studies. *SOCS3* hub–hub gene was another key gene identified in the turquoise module. This gene blocks both the production and signal transduction of type I and II interferons by disrupting the JAK/STAT signaling pathway (Akhtar and Benveniste, 2011). Several studies have shown that viral agents of BRDC disrupt interferon-dependent antiviral responses in the host by inducing the expression of *SOCS1* and *SOCS3* and subsequently provide a suitable condition for their survival and proliferation (Akhtar and Benveniste, 2011; Ye et al., 2015; Zheng et al., 2015; Alkheraif et al., 2017; Salem et al., 2019). These findings indicate the importance of the *SOCS3* gene as a therapeutic target for infections caused by BRDC viral agents.

Moreover, other important hub–hub genes in the turquoise module included *IL10* and *TLR4*. Interleukin-10 (*IL10*) is an anti-inflammatory cytokine that inhibits inflammatory responses initiated by proinflammatory cytokines and subsequently regulates inflammation (Pestka et al., 2004). Therefore, some studies show that increasing the expression of *IL10* regulates the inflammatory response during BRD (Molina et al., 2014; Gondaira et al., 2015; Rodríguez et al., 2015). On the other hand, Risalde et al. (2011) showed that BHV-1 secondary infection in BVDV-infected cows suppressed *IL10* expression, which leads to exacerbate the inflammatory response and more severe clinical lesions (Risalde et al., 2011). Furthermore, failure in upregulation of *IL10* expression level due to weaning and transport is associated with a doubling of BRD mortality (Hodgson et al., 2005). *TLR4* is one of the most important cell surface PRRs that induces an inflammatory response in response to lipopolysaccharide (LPS) derived from Gram-negative bacteria by activation of different signaling pathways (Ciesielska et al., 2021). However, overexpression and abnormal activation of *TLR4* leads to chronic and acute inflammatory disorders such as endotoxemia and sepsis in human and equine (Werners and Bryant, 2012; Kuzmich et al., 2017). Also, due to the key role that *TLR4* plays in activating the signaling pathways that lead to the secretion of proinflammatory cytokine, this gene has been suggested as a very attractive therapeutic target for inflammatory diseases, such as sepsis in human and equine (Werners and Bryant, 2012; Kuzmich et al., 2017). Moreover, a significant correlation has been reported between increased *TLR4* expression level and increased mortality during BRD (Hodgson et al., 2012). Therefore, these results indicate a key role for *IL10* and *TLR4* during BRD that can be further explored as important targets.

Among the detected TFs, *STAT3* is one of the most important hub–hub TFs identified in the turquoise module. Signal

transducer and activator of transcription 3 (*STAT3*) is a transcription factor belonging to the STAT protein family that regulates various biological activities such as apoptosis, angiogenesis, differentiation, cell proliferation, inflammation, and the immune response (Furtek et al., 2016). In different studies, *STAT3* has been suggested as a key gene involved in various bovine diseases (Jaslow et al., 2018; Villaseñor et al., 2019; Bakhtiarizadeh et al., 2020). Moreover, several transcriptomic studies have identified *STAT3* among the top DEGs in response to infection with agents of BRDC, indicating its important role during BRD (Wu et al., 2017; Chao et al., 2019).

Several functional terms identified in the purple module included "Natural killer cell-mediated cytotoxicity," "T-cell receptor signaling pathway," "cellular defense response," "Chemokine signaling pathway," and "T-cell activation." Interestingly, in addition to the antiviral activity of T cells that have been discussed above, the T-cell receptor signaling pathway has also been observed in challenges with *Mannheimia haemolytica*, which indicates the importance and participation of T cells in responding to different types of pathogens (Tizioto et al., 2015). In agreement with the previous studies, *PRF1* (Johnston et al., 2021a), *LCK* (Smirnova et al., 2009), *NCR1* (Osman and Griebel, 2017), *CCL5* (N'jai et al., 2013), *CD8A* (Knapek et al., 2020; Lebedev et al., 2021), *CCR8* (Lopez et al., 2020), *CX3CR1* (Salem et al., 2019), and *TBX21* hub–hub TF (Johnston et al., 2019) were identified as highly connected genes in the purple module and have been reported as immune-related genes during BRD. For instance, the *PRF1* hub–hub gene is an important gene that encodes the perforin-1, which is present in cytotoxic T-lymphocytes (CTLs) and natural killer cells (NK cells) and is involved in cytolysis and the regulation of the immune system. This gene has been suggested by Johnston et al. (2021a) as one of the biomarkers for diagnosis of subclinical BRD from blood samples.

## Module Preservation Analysis

Module preservation analysis and functional enrichment showed that six modules (blue, greenyellow, purple, red, salmon, and yellow) first changed their connectivity pattern and network density due to BRD and second, were biologically related to the BRD development. As expected, highly preserved and semipreserved modules were more active in basic and general cellular activities such as "Ribosome," "rRNA processing," "DNA packaging," "peptide biosynthetic process," and "translation." Therefore, based on highly preserved and semipreserved modules, it is not possible to explain the molecular mechanisms involved in BRD. On the other hand, the six mentioned nonpreserved modules were closely related to the immune system and the pathogenesis of BRD.

Functional terms in the blue module were mostly related to the innate immune system, including "lysosome," "phagosome," "Fc gamma R-mediated phagocytosis," "leukocyte transendothelial migration," "Toll-like receptor signaling pathway," "Chemokine signaling pathway," "neutrophil activation involved in immune response," "activation of MAPK activity," and "positive regulation of NF-kappaB import into nucleus." The leukocyte transendothelial migration is one of the important steps of the innate immune system in triggering the inflammatory immune response and the migration of the first immune response cells such as neutrophils to the sites of infection (Getter et al., 2019; Bakhtiarizadeh et al., 2020) which that chemokines control cellular responses at inflammatory sites through this pathway (Krishnan et al., 2004). In agreement with similar BRD studies (Tizioto et al., 2015; Behura et al., 2017; Johnston et al., 2021c; Lebedev et al., 2021), the leukocyte transendothelial migration was identified as one of the important pathways of the immune system in the blue module. Additionally, some of the hub–hub genes identified in the blue module have also been reported in the previous BRD studies, including *TLR4* (Scott et al., 2021), *ANXA5* (Mitchell et al., 2008), *C3*, *PSEN1* (Neupane et al., 2018), *CTSB*, *CD59*, *FTL* (Nilson et al., 2020), *CAT* (Joshi et al., 2018), *TLR7*, *CD86* (Palomares et al., 2014), *ATG7* (Lipkin et al., 2016), and *IFNAR1* (Amat et al., 2019). For example, stresses from weaning, transport, and commingling reduced the expression level of the *ANXA5* (hub–hub gene) and, thus, lead to an increase in apoptotic cells in the lungs or epithelial lining fluid, which can increase the susceptibility of cattle to a primary infection (Mitchell et al., 2008).

Coexpressed genes in the greenyellow module were enriched in KEGG pathways and biological process such as "T-cell receptor signaling pathway," "focal adhesion," "leukocyte transendothelial migration," and "regulation of cell motility." These pathways have also been observed in previous transcriptomic studies in response to infection with the agents of BRDC (Tizioto et al., 2015; Behura et al., 2017). Among the hub–hub genes identified in the greenyellow module, the *PTEN* gene plays an important role in the pathogenesis of BRD. Research has shown that *MicroRNA-26b* induces the NF-kB signaling pathway by directly targeting *PTEN*, which exacerbates inflammation in the lungs during infection with Gram-negative bacteria (Zhang et al., 2015). Moreover, other hub–hub genes in the greenyellow module including *ADAM10* (Neupane et al., 2018), *MGAT4A* (Johnston et al., 2021b), and *GSK3B* (Chao et al., 2019) were identified as important genes involved in BRD.

Functional enrichment analysis revealed that the purple module was enriched in the "NOD-like receptor signaling pathway," "C-type lectin receptor signaling pathway," "RIG-I-like receptor signaling pathway," and "necroptosis" KEGG pathways as well as "inflammatory response," "type I interferon signaling pathway," "positive regulation of type I interferon production," "interferon-gamma-mediated signaling pathway," and "cellular response to interferon-gamma" biological processes. RIG-I-like receptors are important cytoplasmic PRRs that detect intracellular viruses through their genomic RNA (Takeuchi and Akira, 2010). Recognition of PAMPs by RIG-I-like receptors initiates the RIG-I-like receptor signaling pathway, which activates *NF-κB* and *IRF3/IRF7* transcription factors which subsequently lead to the production of proinflammatory cytokines and type I interferons (Kumar et al., 2011). Type I (IFN-α and IFN-β) and II (IFN-γ) interferons are cytokines that are the first line of defense against viral infections that play a key role in the development of antiviral states during the immune response (Platanias, 2005). Type I interferons are polypeptides that are secreted from virus-infected cells and activate

antimicrobial states in infected cells and healthy neighboring cells to prevent the proliferation and growth of infectious agents, especially viruses. They also promote antigen presentation, increase the activity of natural killer cells, and activate the adaptive immune system (Ivashkiv and Donlin, 2014). Furthermore, type II interferon, whose major producers include natural killer cells, T cells, macrophages, dendritic cells, and B cells play a key role in regulating many protective functions such as inducing antiviral states, enhancing antimicrobial functions, increasing leukocyte trafficking, effect on cell proliferation, and apoptosis (Kak et al., 2018).

Several recent studies have shown that the type I interferon signaling pathway has been identified as a key pathway in cows with BRD atfeedlot entry, which indicates that animals show antiviral responses at the entry stage (Sun et al., 2020; Johnston et al., 2021a; Scott et al., 2021). Interestingly, some of the important hub–hub genes, which are involved in anti-viral interferon response were identified in the purple module, such as *IFI6*, *ISG15*, *MX1*, *OAS2*, *IFIH1*, *DDX58*, *DHX58*, *RSAD2*, *IFI44*, *IFI44L*, *EIF2AK2*, *ISG20*, *MX2*, *IFIT5*, *IFITM3*, *OAS1Y*, and *HERC5*, have been suggested by these studies as potential biomarkers for diagnostic and prediction of subclinical BRD at early stage of infection (Sun et al., 2020; Johnston et al., 2021a; Scott et al., 2021). Additionally, some of the key hub–hub TFs that regulate the expression of coexpressed genes in this module, and play critical roles in interferon antiviral responses during BRD, were included *IRF9*, *STAT1*, *STAT2* (Sun et al., 2020), and IRF7 (Johnston et al., 2019). Type I interferons activate the IFN-stimulated gene factor 3 (*ISGF3*) complex during JAK/STAT signaling pathway. This complex consists of three transcription factors *STAT1*, *STAT2*, and IFN-regulatory factor 9 (*IRF9*), which induce the expression of antiviral genes (Ivashkiv and Donlin, 2014). In addition to type I interferons (IFN-α and IFN-β), type II interferons (IFN-$\gamma$) also cause the formation of STAT1-STAT1 homodimers, which, following phosphorylation and after being transferred to the nucleus, induce the expression of antiviral genes (Platanias, 2005; Kak et al., 2018). Besides, several studies have demonstrated that cytoplasmic localized infected cell protein 0 (*bICP0*) encoded by the BHV-1 (a viral agent of BRDC), through interaction with *IRF7*, disrupts the activity of the IFN-β promoter (Saira et al., 2007; da Silva et al., 2011; Jones, 2019). Moreover, one study demonstrated that BPIV-3 had a negative effect on the JAK/STAT signaling pathway by reducing phosphorylation of *STAT1*, thereby inhibiting the production of antiviral molecules (Eberle et al., 2016). These explain why the *STAT1*, *STAT2*, *IRF9*, and *IRF7* transcription factors are highlighted in this module as very important regulators. Other members of the purple module, including *USP18*, *OAS1X*, *CMPK2*, *GBP4* (Nilson et al., 2020), *IFI35* (Sun et al., 2020), *PARP14* (Oguejiofor et al., 2015), *RTP4* (Johnston et al., 2019), and *TRIM21* (Quick et al., 2020), have also been observed in different BRD studies that may play an important role in the immune system in response to BRDC agents.

In agreement with the previous studies, functional terms including "VEGF signaling pathway," "T-cell receptor signaling pathway" (Tizioto et al., 2015), and "C-type lectin receptor signaling pathway" as well as hub–hub genes including *LCK* (Smirnova et al., 2009), *TLR3* (Marin et al., 2016), *TIA1* (Johnston et al., 2021b), *TNF* (El-Deeb et al., 2020), and *STAT5A* hub–hub TF (Lin et al., 2015) demonstrate the importance of the red module during BRD. Furthermore, the salmon module was mainly enriched in "cellular defense response," "regulation of immune response," "leukocyte cell–cell adhesion," and "natural killer cell-mediated cytotoxicity." Recent studies have demonstrated that some of the hub–hub genes in the salmon module, such as *CCR5* (Salem et al., 2019), *CCR8* (Amat et al., 2019; Lopez et al., 2020), *CX3CR1* (Scott et al., 2021), *ITGAL*, *IL12RB2* (Neupane et al., 2018), *NCR1* (Osman and Griebel, 2017), *CCL5* (N'jai et al., 2013), and *PRF1* (Johnston et al., 2021a) tend to participate in BRD. Additionally, *GZMB* gene, which plays a major role in stimulating cytotoxic T-cell responses and limiting virus replication in the host (Jiminez et al., 2021), was found among hub–hub genes in the salmon module. Granzyme B-protein, which is encoded by this gene, is an important serine protease that is expressed in cytotoxic T-lymphocytes (CTL) and natural killer (NK) cells and kills viral infected cells through apoptosis (Xu et al., 2018). *GZMB* was the most highly upregulated gene in the bronchial lymph node in response to BRSV infection, indicating a close relationship between this gene and the response to viral infection (Johnston et al., 2019). This gene has also been found to be among the DEGs with the highest expression in animals that are resistant to BRD (Scott et al., 2020). The *GZMB* gene is likely to play an important role in the host defense against BRSV infection and maybe other BRDC viral agents, and changes in this gene are critical for BRD resistance and susceptibility (Johnston et al., 2019), as shown, the presence of polymorphisms in the *GZMB* gene in mice caused the cytotoxic T cells to lose their ability to kill viral infected cells (Andoniou et al., 2014).

The yellow module showed that its genes were enriched in some important biological processes and KEGG pathways associated with BRD such as "neutrophil activation involved in immune response," "regulation of inflammatory response," "positive regulation of T-cell proliferation," "MAPK signaling pathway," "NF-kappa B signaling pathway," "apoptosis," "TNF signaling pathway," "JAK-STAT signaling pathway," and "Cytokine-cytokine receptor interaction." Furthermore, examination of the relationship between the hub–hub genes of this module and BRD showed that many of these genes, such as *IL1B* (Behura et al., 2017), *IL15* (Leach et al., 2012; Amat et al., 2019), *CD68* (Buchenau et al., 2010; Hermeyer et al., 2011), *SOCS3* (Ye et al., 2015; Zheng et al., 2015), *NFKBIA*, *RAF1*, *SGK1* (Chao et al., 2019), *ANXA1* (Mitchell et al., 2008), *MYD88* (Dubbert et al., 2013), *FAS* (Xu et al., 2012), *CCR1* (Lindholm-Perry et al., 2018), *IFNAR2* (Schlender et al., 2000), *NFAM1*, *NUDT3* (Johnston et al., 2021b), and *SLC2A3* (N'jai et al., 2013) as well as *DDIT3* hub–hub TF (Wang S. et al., 2020) have important effects on the interaction between the host and the pathogen. For example, stressful stimuli directly increase the expression level of the *SGK1* (hub–hub gene), and

consequently an upregulation in the expression of this gene leads to stimulate BoHV-1 and HSV-1 replication (Kook and Jones, 2016; Zhu et al., 2017). Therefore, the use of SGK inhibitors may be a suitable strategy to reduce BoHV-1 and HSV-1 replication (Kook and Jones, 2016).

These findings demonstrate the relevance of the mentioned modules as well as their genes, especially hub–hub genes and TFs as important candidates in the development of BRD, helping us to better understand the molecular mechanisms responsible for the immune response to BRD. Further researches are needed to more closely examine the biological behavior and functions of these modules and their genes during BRD.

## CONCLUSION

Given that BRD is the main cause of morbidity and mortality in beef and dairy cattle and has a potential impact on economic losses in the livestock industry, a systems biology approach was used to further investigate the molecular mechanisms of BRD as well as to identify diagnosis biomarkers and therapeutic targets for BRD. In this study by using WGCNA distinct methods (MTRs and MP) and functional enrichment analysis, we identified eight candidate modules that are involved in the immune response and BRD pathogenesis. It is noteworthy that both WGCNA methods showed a similar ability to identify candidate modules during BRD, confirming each other results. Integrated coexpressed hub genes of eight candidate modules with PPI networks, allowed us to identify hub–hub genes that act as central genes in both coexpression and PPI networks and may be important candidates during BRD development. In total, we identified 307 hub–hub genes in eight candidate modules, most of which were potentially involved in BRD. These genes along with other members of the eight candidate modules could be important targets in the pathogenesis of BRD for future researches. Therefore, more research is needed to validate the hub–hub genes reported in this study, especially those whose role in the immune system in response to BRD is still unclear.

## 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

AH conceived the ideas. AH designed the study. AH and NS analyzed the data. AH, AB, and FF interpreted the data. AH wrote the main manuscript. GJ, MB, HK, and SI helped in writing the manuscript. AB and RA reviewed and edited the manuscript. HM, AB, and HB reviewed the manuscript. All authors read and approved the final version of manuscript.

## ACKNOWLEDGMENTS

We would like to thank the University of Tehran for supporting this research.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2021.753839/full#supplementary-material

## REFERENCES

Akhtar, L. N., and Benveniste, E. N. (2011). Viral Exploitation of Host SOCS Protein Functions. *J. Virol.* 85 (5), 1912–1921. doi:10.1128/JVI.01857-10

Akira, S., Uematsu, S., and Takeuchi, O. (2006). Pathogen Recognition and Innate Immunity. *Cell* 124 (4), 783–801. doi:10.1016/j.cell.2006.02.015

Alkheraif, A. A., Topliff, C. L., Reddy, J., Massilamany, C., Donis, R. O., Meyers, G., et al. (2017). Type 2 BVDV Npro Suppresses IFN-1 Pathway Signaling in Bovine Cells and Augments BRSV Replication. *Virology* 507, 123–134. doi:10.1016/j.virol.2017.04.015

Amarante-Mendes, G. P., Adjemian, S., Branco, L. M., Zanetti, L. C., Weinlich, R., and Bortoluci, K. R. (2018). Pattern Recognition Receptors and the Host Cell Death Molecular Machinery. *Front. Immunol.* 9, 2379. doi:10.3389/fimmu.2018.02379

Amat, S., Timsit, E., Baines, D., Yanke, J., Alexander, T. W., and Liu, S.-J. (2019). Development of Bacterial Therapeutics against the Bovine Respiratory Pathogen Mannheimia Haemolytica. *Appl. Environ. Microbiol.* 85 (21), e01359–01319. doi:10.1128/AEM.01359-19

Amrine, D. E., White, B. J., Larson, R., Anderson, D. E., Mosier, D. A., and Cernicchiaro, N. (2013). Precision and Accuracy of Clinical Illness Scores, Compared with Pulmonary Consolidation Scores, in Holstein Calves with Experimentally inducedMycoplasma Bovispneumonia. *Am. J. Vet. Res.* 74 (2), 310–315. doi:10.2460/ajvr.74.2.310

Anders, S., Pyl, P. T., and Huber, W. (2014). HTSeq--a Python Framework to Work with High-Throughput Sequencing Data. *Bioinformatics* 31 (2), 166–169. doi:10.1093/bioinformatics/btu638

Andoniou, C. E., Sutton, V. R., Wikstrom, M. E., Fleming, P., Thia, K. Y. T., Matthews, A. Y., et al. (2014). A Natural Genetic Variant of Granzyme B Confers Lethality to a Common Viral Infection. *Plos Pathog.* 10 (12), e1004526. doi:10.1371/journal.ppat.1004526

Aulik, N. A., Hellenbrand, K. M., Klos, H., and Czuprynski, C. J. (2010). Mannheimia Haemolytica and its Leukotoxin Cause Neutrophil Extracellular Trap Formation by Bovine Neutrophils. *Infect. Immun.* 78 (11), 4454–4466. doi:10.1128/IAI.00840-10

Bakhtiarizadeh, M. R., Hosseinpour, B., Shahhoseini, M., Korte, A., and Gifani, P. (2018). Weighted Gene Co-expression Network Analysis of Endometriosis and Identification of Functional Modules Associated with its Main Hallmarks. *Front. Genet.* 9, 453. doi:10.3389/fgene.2018.00453

Bakhtiarizadeh, M. R., Mirzaei, S., Norouzi, M., Sheybani, N., and Vafaei Sadi, M. S. (2020). Identification of Gene Modules and Hub Genes Involved in Mastitis Development Using a Systems Biology Approach. *Front. Genet.* 11, 722. doi:10.3389/fgene.2020.00722

Banse, H., Wollums, A., and Step, D. L. (2013). A Review of Host Pulmonary Defenses with Reference to Cattle. *Bovine Pract.* 48 (1), 13–24. doi:10.21423/bovine-vol48no1p13-24

Bedient, L., Pokharel, S. M., Chiok, K. R., Mohanty, I., Beach, S. S., Miura, T. A., et al. (2020). Lytic Cell Death Mechanisms in Human Respiratory Syncytial

Virus-Infected Macrophages: Roles of Pyroptosis and Necroptosis. *Viruses* 12 (9), 932. doi:10.3390/v12090932

Behura, S. K., Tizioto, P. C., Kim, J., Grupioni, N. V., Seabury, C. M., Schnabel, R. D., et al. (2017). Tissue Tropism in Host Transcriptional Response to Members of the Bovine Respiratory Disease Complex. *Sci. Rep.* 7 (1), 17938. doi:10.1038/s41598-017-18205-0

Bell, R. L., Turkington, H. L., and Cosby, S. L. (2021). The Bacterial and Viral Agents of BRDC: Immune Evasion and Vaccine Developments. *Vaccines* 9 (4), 337. doi:10.3390/vaccines9040337

Blasius, A. L., and Beutler, B. (2010). Intracellular Toll-like Receptors. *Immunity* 32 (3), 305–315. doi:10.1016/j.immuni.2010.03.012

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

Buchenau, I., Poumarat, F., Grand, D. L., Linkner, H., Rosengarten, R., and Hewicker-Trautwein, M. (2010). Expression of Mycoplasma Bovis Variable Surface Membrane Proteins in the Respiratory Tract of Calves after Experimental Infection with a Clonal Variant of Mycoplasma Bovis Type Strain PG45. *Res. Vet. Sci.* 89 (2), 223–229. doi:10.1016/j.rvsc.2010.03.014

Canton, J. (2014). Phagosome Maturation in Polarized Macrophages. *J. Leukoc. Biol.* 96 (5), 729–738. doi:10.1189/jlb.1MR0114-021R

Caswell, J. L. (2014). Failure of Respiratory Defenses in the Pathogenesis of Bacterial Pneumonia of Cattle. *Vet. Pathol.* 51 (2), 393–409. doi:10.1177/0300985813502821

Caswell, J. L., Middleton, D. M., Sorden, S. D., and Gordon, J. R. (1998). Expression of the Neutrophil Chemoattractant Interleukin-8 in the Lesions of Bovine Pneumonic Pasteurellosis. *Vet. Pathol.* 35 (2), 124–131. doi:10.1177/030098589803500206

Chao, J., Han, X., Liu, K., Li, Q., Peng, Q., Lu, S., et al. (2019). Calves Infected with Virulent and Attenuated Mycoplasma Bovis Strains Have Upregulated Th17 Inflammatory and Th1 Protective Responses, Respectively. *Genes* 10 (9), 656. doi:10.3390/genes10090656

Chen, E. Y., Tan, C. M., Kou, Y., Duan, Q., Wang, Z., Meirelles, G. V., et al. (2013). Enrichr: Interactive and Collaborative HTML5 Gene List Enrichment Analysis Tool. *BMC Bioinformatics* 14 (1), 128. doi:10.1186/1471-2105-14-128

Chen, G., Shaw, M. H., Kim, Y.-G., and Nuñez, G. (2009). NOD-like Receptors: Role in Innate Immunity and Inflammatory Disease. *Annu. Rev. Pathol. Mech. Dis.* 4 (1), 365–398. doi:10.1146/annurev.pathol.4.110807.092239

Chin, C.-H., Chen, S.-H., Wu, H.-H., Ho, C.-W., Ko, M.-T., and Lin, C.-Y. (2014). cytoHubba: Identifying Hub Objects and Sub-networks from Complex Interactome. *BMC Syst. Biol.* 8 (4), S11. doi:10.1186/1752-0509-8-S4-S11

Ciesielska, A., Matyjek, M., and Kwiatkowska, K. (2021). TLR4 and CD14 Trafficking and its Influence on LPS-Induced Pro-inflammatory Signaling. *Cell. Mol. Life Sci.* 78 (4), 1233–1261. doi:10.1007/s00018-020-03656-y

Cline, M. S., Smoot, M., Cerami, E., Kuchinsky, A., Landys, N., Workman, C., et al. (2007). Integration of Biological Networks and Gene Expression Data Using Cytoscape. *Nat. Protoc.* 2 (10), 2366–2382. doi:10.1038/nprot.2007.324

Cortjens, B., de Boer, O. J., de Jong, R., Antonis, A. F., Sabogal Piñeros, Y. S., Lutter, R., et al. (2016). Neutrophil Extracellular Traps Cause Airway Obstruction during Respiratory Syncytial Virus Disease. *J. Pathol.* 238 (3), 401–411. doi:10.1002/path.4660

Crepieux, T., Miller, C., Regev-Shoshani, G., Schaefer, A., Dorin, C., Alexander, T., et al. (2016). Randomized, Non-inferiority Trial Comparing a Nitric Oxide Releasing Solution with a Macrolide Antibiotic for Control of Bovine Respiratory Disease in Beef Feedlot Calves at High-Risk of Developing Respiratory Tract Disease. *Res. Vet. Sci.* 105, 216–221. doi:10.1016/j.rvsc.2016.02.020

Cristina, J., Yunus, A. S., Rockemann, D. D., and Samal, S. K. (2001). Bovine Respiratory Syncytial Virus Can Induce Apoptosis in MDBK Cultured Cells. *Vet. Microbiol.* 83 (4), 317–320. doi:10.1016/S0378-1135(01)00435-7

Cusack, P., McMENIMAN, N., and Lean, I. (2003). The Medicine and Epidemiology of Bovine Respiratory Disease in Feedlots. *Aust. Vet J* 81 (8), 480–487. doi:10.1111/j.1751-0813.2003.tb13367.x

da Silva, L. F., Gaudreault, N., and Jones, C. (2011). Cytoplasmic Localized Infected Cell Protein 0 (bICP0) Encoded by Bovine Herpesvirus 1 Inhibits β Interferon Promoter Activity and Reduces IRF3 (Interferon Response Factor 3) Protein Levels. *Virus. Res.* 160 (1), 143–149. doi:10.1016/j.virusres.2011.06.003

Darzi, M., Gorgin, S., Majidzadeh-A, K., and Esmaeili, R. (2021). Gene Co-expression Network Analysis Reveals Immune Cell Infiltration as a Favorable Prognostic Marker in Non-uterine Leiomyosarcoma. *Sci. Rep.* 11 (1), 2339. doi:10.1038/s41598-021-81952-8

DeLeo, F. R., and Allen, L.-A. H. (2020). Phagocytosis and Neutrophil Extracellular Traps. *Fac. Rev.* 9, 25. doi:10.12703/r/9-25

Dubbert, J., Bowers, A., Su, Y., and McClenahan, D. (2013). Effect of TRIF on Permeability and Apoptosis in Bovine Microvascular Endothelial Cells Exposed to Lipopolysaccharide. *Vet. J.* 198 (2), 419–423. doi:10.1016/j.tvjl.2013.08.025

Eberle, K. C., McGill, J. L., Reinhardt, T. A., Sacco, R. E., and Perlman, S. (2016). Parainfluenza Virus 3 Blocks Antiviral Mediators Downstream of the Interferon Lambda Receptor by Modulating Stat1 Phosphorylation. *J. Virol.* 90 (6), 2948–2958. doi:10.1128/JVI.02502-15

Edwards, T. A. (2010). Control Methods for Bovine Respiratory Disease for Feedlot Cattle. *Vet. Clin. North America: Food Anim. Pract.* 26 (2), 273–284. doi:10.1016/j.cvfa.2010.03.005

El-Deeb, W., Elsohaby, I., Fayez, M., Mkrtchyan, H. V., El-Etriby, D., and ElGioushy, M. (2020). Use of Procalcitonin, Neopterin, Haptoglobin, Serum Amyloid A and Proinflammatory Cytokines in Diagnosis and Prognosis of Bovine Respiratory Disease in Feedlot Calves under Field Conditions. *Acta Tropica* 204, 105336. doi:10.1016/j.actatropica.2020.105336

Ellis, J. A. (2001). The Immunology of the Bovine Respiratory Disease Complex. *Vet. Clin. North America: Food Anim. Pract.* 17 (3), 535–550. doi:10.1016/S0749-0720(15)30005-0

Ferwerda, G., Kullberg, B. J., de Jong, D. J., Girardin, S. E., Langenberg, D. M. L., van Crevel, R., et al. (2007). Mycobacterium Paratuberculosisis Recognized by Toll-like Receptors and NOD2. *J. Leukoc. Biol.* 82 (4), 1011–1018. doi:10.1189/jlb.0307147

Flannagan, R. S., Jaumouillé, V., and Grinstein, S. (2012). The Cell Biology of Phagocytosis. *Annu. Rev. Pathol. Mech. Dis.* 7 (1), 61–98. doi:10.1146/annurev-pathol-011811-132445

Franchi, L., Warner, N., Viani, K., and Nuñez, G. (2009). Function of Nod-like Receptors in Microbial Recognition and Host Defense. *Immunological Rev.* 227 (1), 106–128. doi:10.1111/j.1600-065X.2008.00734.x

Furtek, S. L., Backos, D. S., Matheson, C. J., and Reigan, P. (2016). Strategies and Approaches of Targeting STAT3 for Cancer Treatment. *ACS Chem. Biol.* 11 (2), 308–318. doi:10.1021/acschembio.5b00945

Gaddum, R. M., Cook, R. S., Furze, J. M., Ellis, S. A., and Taylor, G. (2003). Recognition of Bovine Respiratory Syncytial Virus Proteins by Bovine CD8+ T Lymphocytes. *Immunology* 108 (2), 220–229. doi:10.1046/j.1365-2567.2003.01566.x

Gagea, M. I., Bateman, K. G., van Dreumel, T., McEwen, B. J., Carman, S., Archambault, M., et al. (2006). Diseases and Pathogens Associated with Mortality in Ontario Beef Feedlots. *J. VET. Diagn. Invest.* 18 (1), 18–28. doi:10.1177/104063870601800104

Geijtenbeek, T. B. H., and Gringhuis, S. I. (2009). Signalling through C-type Lectin Receptors: Shaping Immune Responses. *Nat. Rev. Immunol.* 9 (7), 465–479. doi:10.1038/nri2569

Geiser, V., and Jones, C. (2005). Localization of Sequences within the Latency-Related Gene of Bovine Herpesvirus 1 that Inhibit Mammalian Cell Growth. *J. Neurovirol.* 11 (6), 563–570. doi:10.1080/13550280500385286

Getter, T., Margalit, R., Kahremany, S., Levy, L., Blum, E., Khazanov, N., et al. (2019). Novel Inhibitors of Leukocyte Transendothelial Migration. *Bioorg. Chem.* 92, 103250. doi:10.1016/j.bioorg.2019.103250

Gondaira, S., Higuchi, H., Iwano, H., Nakajima, K., Kawai, K., Hashiguchi, S., et al. (2015). Cytokine mRNA Profiling and the Proliferative Response of Bovine Peripheral Blood Mononuclear Cells to Mycoplasma Bovis. *Vet. Immunol. Immunopathology* 165 (1), 45–53. doi:10.1016/j.vetimm.2015.03.002

Gondaira, S., Higuchi, H., Nishi, K., Iwano, H., and Nagahata, H. (2017). Mycoplasma Bovis Escapes Bovine Neutrophil Extracellular Traps. *Vet. Microbiol.* 199, 68–73. doi:10.1016/j.vetmic.2016.12.022

Gray, D., Ellis, J. A., Gow, S. P., Lacoste, S. R., Allan, G. M., and Mooney, M. H. (2019). Profiling of Local Disease-Sparing Responses to Bovine Respiratory Syncytial Virus in Intranasally Vaccinated and Challenged Calves. *J. Proteomics* 204, 103397. doi:10.1016/j.jprot.2019.103397

Grissett, G. P., White, B. J., and Larson, R. L. (2015). Structured Literature Review of Responses of Cattle to Viral and Bacterial Pathogens Causing Bovine Respiratory Disease Complex. *J. Vet. Intern. Med.* 29 (3), 770–780. doi:10.1111/jvim.12597

Heidarzadeh, M., Roodbari, F., Hassanpour, M., Ahmadi, M., Saberianpour, S., and Rahbarghazi, R. (2020). Toll-like Receptor Bioactivity in Endothelial Progenitor Cells. *Cell Tissue Res* 379 (2), 223–230. doi:10.1007/s00441-019-03119-2

Hellenbrand, K. M., Forsythe, K. M., Rivera-Rivas, J. J., Czuprynski, C. J., and Aulik, N. A. (2013). Histophilus Somni Causes Extracellular Trap Formation by Bovine Neutrophils and Macrophages. *Microb. Pathogenesis* 54, 67–75. doi:10.1016/j.micpath.2012.09.007

Hermeyer, K., Jacobsen, B., Spergser, J., Rosengarten, R., and Hewicker-Trautwein, M. (2011). Detection of Mycoplasma Bovis by In-Situ Hybridization and Expression of Inducible Nitric Oxide Synthase, Nitrotyrosine and Manganese Superoxide Dismutase in the Lungs of Experimentally-Infected Calves. *J. Comp. Pathol.* 145 (2), 240–250. doi:10.1016/j.jcpa.2010.12.005

Hodgins, D. C., Conlon, J. A., and Shewen, P. E. (2002). "Respiratory Viruses and Bacteria in Cattle," in *Polymicrobial Diseases* (Washington (DC): ASM Press), 213–229.

Hodgson, P. D., Aich, P., Manuja, A., Hokamp, K., Roche, F. M., Brinkman, F. S. L., et al. (2005). Effect of Stress on Viral-Bacterial Synergy in Bovine Respiratory Disease: Novel Mechanisms to Regulate Inflammation. *Comp. Funct. Genomics* 6, 244–250. doi:10.1002/cfg.474

Hodgson, P. D., Aich, P., Stookey, J., Popowych, Y., Potter, A., Babiuk, L., et al. (2012). Stress Significantly Increases Mortality Following a Secondary Bacterial Respiratory Infection. *Vet. Res.* 43 (1), 21. doi:10.1186/1297-9716-43-21

Hofstetter, A. R., and Sacco, R. E. (2020). Oxidative Stress Pathway Gene Transcription after Bovine Respiratory Syncytial Virus Infection *In Vitro* and *Ex Vivo*. *Vet. Immunol. Immunopathology* 219, 109956. doi:10.1016/j.vetimm.2019.109956

Horvath, C. M. (2004b). The Jak-STAT Pathway Stimulated by Interferon. *Sci. Signaling* 2004 (260), tr8. doi:10.1126/stke.2602004tr8

Horvath, C. M. (2004a). The Jak-STAT Pathway Stimulated by Interferon or Interferon. *Sci. Signaling* 2004 (260), tr10. doi:10.1126/stke.2602004tr10

Hu, H., Miao, Y.-R., Jia, L.-H., Yu, Q.-Y., Zhang, Q., and Guo, A.-Y. (2018). AnimalTFDB 3.0: a Comprehensive Resource for Annotation and Prediction of Animal Transcription Factors. *Nucleic Acids Res.* 47 (D1), D33–D38. doi:10.1093/nar/gky822

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), 200. doi:10.3390/pathogens9030200

Ivashkiv, L. B., and Donlin, L. T. (2014). Regulation of Type I Interferon Responses. *Nat. Rev. Immunol.* 14 (1), 36–49. doi:10.1038/nri3581

Iwasaki, A., and Medzhitov, R. (2004). Toll-like Receptor Control of the Adaptive Immune Responses. *Nat. Immunol.* 5 (10), 987–995. doi:10.1038/ni1112

Janeway, C. A., and Medzhitov, R. (2002). Innateimmunerecognition. *Annu. Rev. Immunol.* 20 (1), 197–216. doi:10.1146/annurev.immunol.20.083001.084359

Jaslow, S. L., Gibbs, K. D., Fricke, W. F., Wang, L., Pittman, K. J., Mammel, M. K., et al. (2018). Salmonella Activation of STAT3 Signaling by SarA Effector Promotes Intracellular Replication and Production of IL-10. *Cel Rep.* 23 (12), 3525–3536. doi:10.1016/j.celrep.2018.05.072

Jiminez, J., Timsit, E., Orsel, K., van der Meer, F., Guan, L. L., and Plastow, G. (2021). Whole-Blood Transcriptome Analysis of Feedlot Cattle with and without Bovine Respiratory Disease. *Front. Genet.* 12, 257. doi:10.3389/fgene.2021.627623

Johnston, D., Earley, B., McCabe, M. S., Kim, J., Taylor, J. F., Lemon, K., et al. (2021a). Messenger RNA Biomarkers of Bovine Respiratory Syncytial Virus Infection in the Whole Blood of Dairy Calves. *Sci. Rep.* 11 (1), 9392. doi:10.1038/s41598-021-88878-1

Johnston, D., Earley, B., McCabe, M. S., Kim, J., Taylor, J. F., Lemon, K., et al. (2021b). Elucidation of the Host Bronchial Lymph Node miRNA Transcriptome Response to Bovine Respiratory Syncytial Virus. *Front. Genet.* 12, 526. doi:10.3389/fgene.2021.633125

Johnston, D., Earley, B., McCabe, M. S., Lemon, K., Duffy, C., McMenamy, M., et al. (2019). Experimental challenge with Bovine Respiratory Syncytial Virus in Dairy Calves: Bronchial Lymph Node Transcriptome Response. *Sci. Rep.* 9 (1), 14736. doi:10.1038/s41598-019-51094-z

Johnston, D., Kim, J., Taylor, J. F., Earley, B., McCabe, M. S., Lemon, K., et al. (2021c). ATAC-seq Identifies Regions of Open Chromatin in the Bronchial Lymph Nodes of Dairy Calves Experimentally Challenged with Bovine Respiratory Syncytial Virus. *BMC Genomics* 22 (1), 14. doi:10.1186/s12864-020-07268-5

Jones, C. (2019). Bovine Herpesvirus 1 Counteracts Immune Responses and Immune-Surveillance to Enhance Pathogenesis and Virus Transmission. *Front. Immunol.* 10, 1008. doi:10.3389/fimmu.2019.01008

Jones, C., and Chowdhury, S. (2010). Bovine Herpesvirus Type 1 (BHV-1) Is an Important Cofactor in the Bovine Respiratory Disease Complex. *Vet. Clin. North America: Food Anim. Pract.* 26 (2), 303–321. doi:10.1016/j.cvfa.2010.04.007

Joshi, V., Gupta, V. K., Bhanuprakash, A. G., Mandal, R. S. K., Dimri, U., and Ajith, Y. (2018). Haptoglobin and Serum Amyloid A as Putative Biomarker Candidates of Naturally Occurring Bovine Respiratory Disease in Dairy Calves. *Microb. Pathogenesis* 116, 33–37. doi:10.1016/j.micpath.2018.01.001

Kadarmideen, H. N., Watson-Haigh, N. S., and Andronicos, N. M. (2011). Systems Biology of Ovine Intestinal Parasite Resistance: Disease Gene Modules and Biomarkers. *Mol. Biosyst.* 7 (1), 235–246. doi:10.1039/c0mb00190b

Kadarmideen, H. N., and Watson-Haigh, N. S. (2012). Building Gene Co-expression Networks Using Transcriptomics Data for Systems Biology Investigations: Comparison of Methods Using Microarray Data. *Bioinformation* 8 (18), 855–861. doi:10.6026/97320630008855

Kak, G., Raza, M., and Tiwari, B. K. (2018). Interferon-gamma (IFN-$\gamma$): Exploring its Implications in Infectious Diseases. *Biomol. Concepts* 9 (1), 64–79. doi:10.1515/bmc-2018-0007

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

Kim, D., Langmead, B., and Salzberg, S. L. (2015). HISAT: a Fast Spliced Aligner with Low Memory Requirements. *Nat. Methods* 12 (4), 357–360. doi:10.1038/nmeth.3317

Kim, E. H., and Suresh, M. (2013). Role of PI3K/Akt Signaling in Memory CD8 T Cell Differentiation. *Front. Immun.* 4, 20. doi:10.3389/fimmu.2013.00020

Kirchhoff, J., Uhlenbruck, S., Goris, K., Keil, G. M., and Herrler, G. (2014). Three Viruses of the Bovine Respiratory Disease Complex Apply Different Strategies to Initiate Infection. *Vet. Res.* 45 (1), 20. doi:10.1186/1297-9716-45-20

Kleinschmidt, S., Spergser, J., Rosengarten, R., and Hewicker-Trautwein, M. (2013). Long-term Survival of Mycoplasma Bovis in Necrotic Lesions and in Phagocytic Cells as Demonstrated by Transmission and Immunogold Electron Microscopy in Lung Tissue from Experimentally Infected Calves. *Vet. Microbiol.* 162 (2), 949–953. doi:10.1016/j.vetmic.2012.11.039

Knapek, K. J., Georges, H. M., Van Campen, H., Bishop, J. V., Bielefeldt-Ohmann, H., Smirnova, N. P., et al. (2020). Fetal Lymphoid Organ Immune Responses to Transient and Persistent Infection with Bovine Viral Diarrhea Virus. *Viruses* 12 (8), 816. doi:10.3390/v12080816

Kolaczkowska, E., and Kubes, P. (2013). Neutrophil Recruitment and Function in Health and Inflammation. *Nat. Rev. Immunol.* 13 (3), 159–175. doi:10.1038/nri3399

Kommadath, A., Bao, H., Arantes, A. S., Plastow, G. S., Tuggle, C. K., Bearson, S. M., et al. (2014). Gene Co-expression Network Analysis Identifies Porcine Genes Associated with Variation in Salmonella Shedding. *BMC Genomics* 15 (1), 452. doi:10.1186/1471-2164-15-452

Kook, I., and Jones, C. (2016). The Serum and Glucocorticoid-Regulated Protein Kinases (SGK) Stimulate Bovine Herpesvirus 1 and Herpes Simplex Virus 1 Productive Infection. *Virus. Res.* 222, 106–112. doi:10.1016/j.virusres.2016.06.007

Krishnan, S., Halonen, M., and Welliver, R. C. (2004). Innate Immune Responses in Respiratory Syncytial Virus Infections. *Viral Immunol.* 17 (2), 220–233. doi:10.1089/0882824041310612

Kumar, H., Kawai, T., and Akira, S. (2011). Pathogen Recognition by the Innate Immune System. *Int. Rev. Immunol.* 30 (1), 16–34. doi:10.3109/08830185.2010.529976

Kumar, H., Kawai, T., and Akira, S. (2009). Toll-like Receptors and Innate Immunity. *Biochem. Biophysical Res. Commun.* 388 (4), 621–625. doi:10.1016/j.bbrc.2009.08.062

Kuzmich, N., Sivak, K., Chubarev, V., Porozov, Y., Savateeva-Lyubimova, T., and Peri, F. (2017). TLR4 Signaling Pathway Modulators as Potential Therapeutics in Inflammation and Sepsis. *Vaccines* 5 (4), 34. doi:10.3390/vaccines5040034

Langfelder, P., and Horvath, S. (2008). WGCNA: an R Package for Weighted Correlation Network Analysis. *BMC Bioinformatics* 9 (1), 559. doi:10.1186/1471-2105-9-559

Langfelder, P., Luo, R., Oldham, M. C., and Horvath, S. (2011). Is My Network Module Preserved and Reproducible. *Plos Comput. Biol.* 7 (1), e1001057. doi:10.1371/journal.pcbi.1001057

Law, C. W., Chen, Y., Shi, W., and Smyth, G. K. (2014). Voom: Precision Weights Unlock Linear Model Analysis Tools for RNA-Seq Read Counts. *Genome Biol.* 15 (2), R29. doi:10.1186/gb-2014-15-2-r29

Leach, R. J., O'Neill, R. G., Fitzpatrick, J. L., Williams, J. L., and Glass, E. J. (2012). Quantitative Trait Loci Associated with the Immune Response to a Bovine Respiratory Syncytial Virus Vaccine. *PLOS ONE* 7 (3), e33526. doi:10.1371/journal.pone.0033526

Lebedev, M., McEligot, H. A., Mutua, V. N., Walsh, P., Carvallo Chaigneau, F. R., and Gershwin, L. J. (2021). Analysis of Lung Transcriptome in Calves Infected with Bovine Respiratory Syncytial Virus and Treated with Antiviral And/or Cyclooxygenase Inhibitor. *PLOS ONE* 16 (2), e0246695. doi:10.1371/journal.pone.0246695

Lee, M. S., and Kim, Y.-J. (2007). Signaling Pathways Downstream of Pattern-Recognition Receptors and Their Cross Talk. *Annu. Rev. Biochem.* 76 (1), 447–480. doi:10.1146/annurev.biochem.76.060605.122847

Lee, S.-R., Nanduri, B., Pharr, G. T., Stokes, J. V., and Pinchuk, L. M. (2009). Bovine Viral Diarrhea Virus Infection Affects the Expression of Proteins Related to Professional Antigen Presentation in Bovine Monocytes. *Biochim. Biophys. Acta (Bba) - Proteins Proteomics* 1794 (1), 14–22. doi:10.1016/j.bbapap.2008.09.005

Levin, R., Grinstein, S., and Canton, J. (2016). The Life Cycle of Phagosomes: Formation, Maturation, and Resolution. *Immunol. Rev.* 273 (1), 156–179. doi:10.1111/imr.12439

Li, A., and Horvath, S. (2006). Network Neighborhood Analysis with the Multi-Node Topological Overlap Measure. *Bioinformatics* 23 (2), 222–231. doi:10.1093/bioinformatics/btl581

Li, L., Li, P., Chen, A., Li, H., Liu, Z., Yu, L., et al. (2021). Quantitative Proteomic Analysis Shows Involvement of the P38 MAPK Pathway in Bovine Parainfluenza Virus Type 3 Replication. *Res. Square*. doi:10.21203/rs.3.rs-253558/v1

Lin, J., Zhao, D., Wang, J., Wang, Y., Li, H., Yin, X., et al. (2015). Transcriptome Changes upon *In Vitro* challenge with Mycobacterium Bovis in Monocyte-Derived Macrophages from Bovine Tuberculosis-Infected and Healthy Cows. *Vet. Immunol. Immunopathology* 163 (3), 146–156. doi:10.1016/j.vetimm.2014.12.001

Lin, S.-C., Lo, Y.-C., and Wu, H. (2010). Helical Assembly in the MyD88-IRAK4-IRAK2 Complex in TLR/IL-1R Signalling. *Nature* 465 (7300), 885–890. doi:10.1038/nature09121

Lindholm-Perry, A. K., Kuehn, L. A., McDaneld, T. G., Miles, J. R., Workman, A. M., Chitko-McKown, C. G., et al. (2018). Complete Blood Count Data and Leukocyte Expression of Cytokine Genes and Cytokine Receptor Genes Associated with Bovine Respiratory Disease in Calves. *BMC Res. Notes* 11 (1), 786. doi:10.1186/s13104-018-3900-x

Lipkin, E., Strillacci, M. G., Eitam, H., Yishay, M., Schiavini, F., Soller, M., et al. (2016). The Use of Kosher Phenotyping for Mapping QTL Affecting Susceptibility to Bovine Respiratory Disease. *PLoS One* 11 (4), e0153423. doi:10.1371/journal.pone.0153423

Liu, B.-H., and Cai, J.-P. (2017). Identification of Transcriptional Modules and Key Genes in Chickens Infected with *Salmonella enterica* Serovar Pullorum Using Integrated Coexpression Analyses. *Biomed. Res. Int.* 2017, 1–12. doi:10.1155/2017/8347085

Liu, L., Amorín, R., Moriel, P., DiLorenzo, N., Lancaster, P. A., and Peñagaricano, F. (2020). Differential Network Analysis of Bovine Muscle Reveals Changes in Gene Coexpression Patterns in Response to Changes in Maternal Nutrition. *BMC Genomics* 21 (1), 684. doi:10.1186/s12864-020-07068-x

Liu, T., Zhang, L., Joo, D., and Sun, S.-C. (2017). NF-κB Signaling in Inflammation. *Sig Transduct Target. Ther.* 2 (1), 17023. doi:10.1038/sigtrans.2017.23

Liu, Z.-g. (2005). Molecular Mechanism of TNF Signaling and beyond. *Cell Res* 15 (1), 24–27. doi:10.1038/sj.cr.7290259

Lopez, B. I., Santiago, K. G., Lee, D., Ha, S., and Seo, K. (2020). RNA Sequencing (RNA-Seq) Based Transcriptome Analysis in Immune Response of Holstein Cattle to Killed Vaccine against Bovine Viral Diarrhea Virus Type I. *Animals* 10 (2), 344. doi:10.3390/ani10020344

Lugo-Villarino, G., Troegeler, A., Balboa, L., Lastrucci, C., Duval, C., Mercier, I., et al. (2018). The C-type Lectin Receptor DC-SIGN Has an Anti-inflammatory Role in Human M(IL-4) Macrophages in Response to *Mycobacterium tuberculosis*. *Front. Immunol.* 9, 1123. doi:10.3389/fimmu.2018.01123

Luoreng, Z.-M., Wang, X.-P., Mei, C.-G., and Zan, L.-S. (2018). Expression Profiling of Peripheral Blood miRNA Using RNAseq Technology in Dairy Cows with Escherichia Coli-Induced Mastitis. *Sci. Rep.* 8 (1), 12693. doi:10.1038/s41598-018-30518-2

Ma, W., Wang, H., and He, H. (2019). Bovine Herpesvirus 1 Tegument Protein UL41 Suppresses Antiviral Innate Immune Response via Directly Targeting STAT1. *Vet. Microbiol.* 239, 108494. doi:10.1016/j.vetmic.2019.108494

Ma, W., Wang, H., and He, H. (2020). Bta-miR-2890 Up-Regulates JAK-STAT Pathway to Inhibit BoHV-1 Replication by Targeting Viral Gene UL41. *Vet. Microbiol.* 245, 108709. doi:10.1016/j.vetmic.2020.108709

Maina, T., Prysliak, T., and Perez-Casal, J. (2019). Mycoplasma Bovis Delay in Apoptosis of Macrophages Is Accompanied by Increased Expression of Anti-apoptotic Genes, Reduced Cytochrome C Translocation and Inhibition of DNA Fragmentation. *Vet. Immunol. Immunopathology* 208, 16–24. doi:10.1016/j.vetimm.2018.12.004

Malazdrewich, C., Ames, T. R., Abrahamsen, M. S., and Maheswaran, S. K. (2001). Pulmonary Expression of Tumor Necrosis Factor Alpha, Interleukin-1 Beta, and Interleukin-8 in the Acute Phase of Bovine Pneumonic Pasteurellosis. *Vet. Pathol.* 38 (3), 297–310. doi:10.1354/vp.38-3-297

Malemud, C., and Pearlman, E. (2009). Targeting JAK/STAT Signaling Pathway in Inflammatory Diseases. *Cst* 4 (3), 201–221. doi:10.2174/157436209789057467

Marin, M. S., Quintana, S., Leunda, M. R., Odeón, A. C., and Pérez, S. E. (2016). Distribution of Bovine Alpha-Herpesviruses and Expression of Toll-like Receptors in the Respiratory System of Experimentally Infected Calves. *Res. Vet. Sci.* 105, 53–55. doi:10.1016/j.rvsc.2016.01.011

Mariotti, M., Williams, J., Dunner, S., Valentini, A., and Pariset, L. (2009). Polymorphisms within the Toll-like Receptor (TLR)-2, -4, and -6 Genes in Cattle. *Diversity* 1 (1), 7–18. doi:10.3390/d1010007

Mason, M. J., Fan, G., Plath, K., Zhou, Q., and Horvath, S. (2009). Signed Weighted Gene Co-expression Network Analysis of Transcriptional Regulation in Murine Embryonic Stem Cells. *BMC Genomics* 10 (1), 327. doi:10.1186/1471-2164-10-327

McGill, J. L., and Sacco, R. E. (2020). The Immunology of Bovine Respiratory Disease. *Vet. Clin. North America: Food Anim. Pract.* 36 (2), 333–348. doi:10.1016/j.cvfa.2020.03.002

Medzhitov, R. (2007). Recognition of Microorganisms and Activation of the Immune Response. *Nature* 449 (7164), 819–826. doi:10.1038/nature06246

Mitchell, G. B., Clark, M. E., Siwicky, M., and Caswell, J. L. (2008). Stress Alters the Cellular and Proteomic Compartments of Bovine Bronchoalveolar Lavage Fluid. *Vet. Immunol. Immunopathology* 125 (1), 111–125. doi:10.1016/j.vetimm.2008.05.005

Molina, V., Risalde, M. A., Sánchez-Cordón, P. J., Romero-Palomo, F., Pedrera, M., Garfia, B., et al. (2014). Cell-Mediated Immune Response during Experimental Acute Infection with Bovine Viral Diarrhoea Virus: Evaluation of Blood Parameters. *Transbound Emerg. Dis.* 61 (1), 44–59. doi:10.1111/tbed.12002

Mukund, K., and Subramaniam, S. (2015). Dysregulated Mechanisms Underlying Duchenne Muscular Dystrophy from Co-expression Network Preservation Analysis. *BMC Res. Notes* 8 (1), 182. doi:10.1186/s13104-015-1141-9

Mulongo, M., Prysliak, T., Scruten, E., Napper, S., Perez-Casal, J., and Blanke, S. R. (2014). In VitroInfection of Bovine Monocytes with Mycoplasma Bovis Delays Apoptosis and Suppresses Production of Gamma Interferon and Tumor Necrosis Factor Alpha but Not Interleukin-10. *Infect. Immun.* 82 (1), 62–71. doi:10.1128/IAI.00961-13

Muraro, S. P., De Souza, G. F., Gallo, S. W., Da Silva, B. K., De Oliveira, S. D., Vinolo, M. A. R., et al. (2018). Respiratory Syncytial Virus Induces the Classical ROS-dependent NETosis through PAD-4 and Necroptosis Pathways Activation. *Sci. Rep.* 8 (1), 14166. doi:10.1038/s41598-018-32576-y

Nalpas, N. C., Magee, D. A., Conlon, K. M., Browne, J. A., Healy, C., McLoughlin, K. E., et al. (2015). RNA Sequencing Provides Exquisite Insight into the Manipulation of the Alveolar Macrophage by Tubercle Bacilli. *Sci. Rep.* 5 (1), 13629. doi:10.1038/srep13629

Neupane, M., Kiser, J. N., Neibergs, H. L., and Neibergs, H. L. (2018). Gene Set Enrichment Analysis of SNP Data in Dairy and Beef Cattle with Bovine Respiratory Disease. *Anim. Genet.* 49 (6), 527–538. doi:10.1111/age.12718

Nilson, S. M., Workman, A. M., Sjeklocha, D., Brodersen, B., Groteluesch en, D. M., and Petersen, J. L. (2020). Upregulation of the Type I Interferon Pathway in Feedlot Cattle Persistently Infected with Bovine Viral Diarrhea Virus. *Virus. Res.* 278, 197862. doi:10.1016/j.virusres.2020.197862

N'jai, A. U., Rivera, J., Atapattu, D. N., Owusu-Ofori, K., and Czuprynski, C. J. (2013). Gene Expression Profiling of Bovine Bronchial Epithelial Cells Exposed *In Vitro* to Bovine Herpesvirus 1 and Mannheimia Haemolytica. *Vet. Immunol. Immunopathology* 155 (3), 182–189. doi:10.1016/j.vetimm.2013.06.012

Nordenfelt, P., and Tapper, H. (2011). Phagosome Dynamics during Phagocytosis by Neutrophils. *J. Leukoc. Biol.* 90 (2), 271–284. doi:10.1189/jlb.0810457

O'Shea, J. J., Plenge, R., and Plenge, R. (2012). JAK and STAT Signaling Molecules in Immunoregulation and Immune-Mediated Disease. *Immunity* 36 (4), 542–550. doi:10.1016/j.immuni.2012.03.014

O'Shea, J. J., Schwartz, D. M., Villarino, A. V., Gadina, M., McInnes, I. B., and Laurence, A. (2015). The JAK-STAT Pathway: Impact on Human Disease and Therapeutic Intervention. *Annu. Rev. Med.* 66 (1), 311–328. doi:10.1146/annurev-med-051113-024537

Oeckinghaus, A., and Ghosh, S. (2009). The NF- B Family of Transcription Factors and its Regulation. *Cold Spring Harbor Perspect. Biol.* 1 (4), a000034. doi:10.1101/cshperspect.a000034

Oguejiofor, C. F., Cheng, Z., Abudureyimu, A., Anstaett, O. L., Brownlie, J., Fouladi-Nashta, A. A., et al. (2015). Global Transcriptomic Profiling of Bovine Endometrial Immune Response *In Vitro*. II. Effect of Bovine Viral Diarrhea Virus on the Endometrial Response to Lipopolysaccharide1. *Biol. Reprod.* 93, 4. doi:10.1095/biolreprod.115.128876

Osman, R. A., and Griebel, P. J. (2017). CD335 (NKp46)+ T-Cell Recruitment to the Bovine Upper Respiratory Tract during a Primary Bovine Herpesvirus-1 Infection. *Front. Immunol.* 8, 1393. doi:10.3389/fimmu.2017.01393

Palomares, R. A., Parrish, J., Woolums, A. R., Brock, K. V., and Hurley, D. J. (2014). Expression of Toll-like Receptors and Co-stimulatory Molecules in Lymphoid Tissue during Experimental Infection of Beef Calves with Bovine Viral Diarrhea Virus of Low and High Virulence. *Vet. Res. Commun.* 38 (4), 329–335. doi:10.1007/s11259-014-9613-2

Pan, Y., Tagawa, Y., Champion, A., Sandal, I., Inzana, T. J., and Palmer, G. H. (2018). Histophilus Somni Survives in Bovine Macrophages by Interfering with Phagosome-Lysosome Fusion but Requires IbpA for Optimal Serum Resistance. *Infect. Immun.* 86 (12), e00365–00318. doi:10.1128/IAI.00365-18

Papayannopoulos, V. (2018). Neutrophil Extracellular Traps in Immunity and Disease. *Nat. Rev. Immunol.* 18 (2), 134–147. doi:10.1038/nri.2017.105

Parameswaran, N., and Patial, S. (2010). Tumor Necrosis Factor-α Signaling in Macrophages. *Crit. Rev. Eukaryot. Gene Expr.* 20 (2), 87–103. doi:10.1615/CritRevEukarGeneExpr.v20.i2.10

Pestka, S., Krause, C. D., Sarkar, D., Walter, M. R., Shi, Y., and Fisher, P. B. (2004). Interleukin-10andRelatedCytokines andReceptors. *Annu. Rev. Immunol.* 22 (1), 929–979. doi:10.1146/annurev.immunol.22.012703.104622

Pihur, V., Datta, S., and Datta, S. (2009). RankAggreg, an R Package for Weighted Rank Aggregation. *BMC Bioinformatics* 10 (1), 62. doi:10.1186/1471-2105-10-62

Platanias, L. C. (2005). Mechanisms of Type-I- and Type-II-Interferon-Mediated Signalling. *Nat. Rev. Immunol.* 5 (5), 375–386. doi:10.1038/nri1604

Platt, R., Burdett, W., and Roth, J. A. (2006). Induction of Antigen-specific T-Cell Subset Activation to Bovine Respiratory Disease Viruses by a Modified-Live Virus Vaccine. *Am. J. Vet. Res.* 67 (7), 1179–1184. doi:10.2460/ajvr.67.7.1179

Pollock, N., Taylor, G., Jobe, F., and Guzman, E. (2017). Modulation of the Transcription Factor NF-Kb in Antigen-Presenting Cells by Bovine Respiratory Syncytial Virus Small Hydrophobic Protein. *J. Gen. Virol.* 98 (7), 1587–1599. doi:10.1099/jgv.0.000855

Pompura, S. L., and Dominguez-Villar, M. (2018). The PI3K/AKT Signaling Pathway in Regulatory T-Cell Development, Stability, and Function. *J. Leukoc. Biol.* 103 (6), 1065–1076. doi:10.1002/JLB.2MIR0817-349R

Portis, E., Lindeman, C., Johansen, L., and Stoltman, G. (2012). A Ten-Year (2000-2009) Study of Antimicrobial Susceptibility of Bacteria that Cause Bovine Respiratory Disease Complex-Mannheimia Haemolytica, Pasteurella Multocida, and Histophilus Somni-In the United States and Canada. *J. VET. Diagn. Invest.* 24 (5), 932–944. doi:10.1177/1040638712457559

Quick, A. E., Ollivett, T. L., Kirkpatrick, B. W., and Weigel, K. A. (2020). Genomic Analysis of Bovine Respiratory Disease and Lung Consolidation in Preweaned Holstein Calves Using Clinical Scoring and Lung Ultrasound. *J. Dairy Sci.* 103 (2), 1632–1641. doi:10.3168/jds.2019-16531

Rai, A. N., Epperson, W. B., and Nanduri, B. (2015). Application of Functional Genomics for Bovine Respiratory Disease Diagnostics. *Bioinform Biol. Insights* 9s2, BBI.S30525. doi:10.4137/bbi.S30525

Rana, H. K., Akhtar, M. R., Islam, M. B., Ahmed, M. B., Liò, P., Quinn, J. M. W., et al. (2019). Genetic Effects of Welding Fumes on the Development of Respiratory System Diseases. *Comput. Biol. Med.* 108, 142–149. doi:10.1016/j.compbiomed.2019.04.004

Rangaraju, S., Dammer, E. B., Raza, S. A., Rathakrishnan, P., Xiao, H., Gao, T., et al. (2018). Identification and Therapeutic Modulation of a Pro-inflammatory Subset of Disease-Associated-Microglia in Alzheimer's Disease. *Mol. Neurodegeneration* 13 (1), 24. doi:10.1186/s13024-018-0254-8

Redondo, E., Gázquez, A., Vadillo, S., García, A., Franco, A., and Masot, A. J. (2014). Induction of Interleukin-8 and Interleukin-12 in Neonatal Ovine Lung Following Experimental Inoculation of Bovine Respiratory Syncytial Virus. *J. Comp. Pathol.* 150 (4), 434–448. doi:10.1016/j.jcpa.2013.08.002

Regev-Shoshani, G., Church, J. S., Cook, N. J., Schaefer, A. L., and Miller, C. (2013). Prophylactic Nitric Oxide Treatment Reduces Incidence of Bovine Respiratory Disease Complex in Beef Cattle Arriving at a Feedlot. *Res. Vet. Sci.* 95 (2), 606–611. doi:10.1016/j.rvsc.2013.06.016

Regev-Shoshani, G., Vimalanathan, S., Prema, D., Church, J. S., Reudink, M. W., Nation, N., et al. (2014). Safety, Bioavailability and Mechanism of Action of Nitric Oxide to Control Bovine Respiratory Disease Complex in Calves Entering a Feedlot. *Res. Vet. Sci.* 96 (2), 328–337. doi:10.1016/j.rvsc.2013.12.012

Remick, D. G. (2005). Interleukin-8. *Crit. Care Med.* 33 (12), S466–S467. doi:10.1097/01.Ccm.0000186783.34908.18

Riquelme Medina, I., and Lubovac-Pilav, Z. (2016). Gene Co-expression Network Analysis for Identifying Modules and Functionally Enriched Pathways in Type 1 Diabetes. *PLOS ONE* 11 (6), e0156006. doi:10.1371/journal.pone.0156006

Risalde, M. A., Molina, V., Sánchez-Cordón, P. J., Pedrera, M., Panadero, R., Romero-Palomo, F., et al. (2011). Response of Proinflammatory and Anti-inflammatory Cytokines in Calves with Subclinical Bovine Viral Diarrhea Challenged with Bovine Herpesvirus-1. *Vet. Immunol. Immunopathology* 144 (1), 135–143. doi:10.1016/j.vetimm.2011.07.022

Rodríguez, F., González, J. F., Arbelo, M., Zucca, D., and Fernández, A. (2015). Cytokine Expression in Lungs of Calves Spontaneously Infected with Mycoplasma Bovis. *Vet. Res. Commun.* 39 (1), 69–72. doi:10.1007/s11259-014-9620-3

Rosales, C. (2017). Fc*γ* Receptor Heterogeneity in Leukocyte Functional Responses. *Front. Immunol.* 8, 280. doi:10.3389/fimmu.2017.00280

Sabino, M., Carmelo, V. A. O., Mazzoni, G., Cappelli, K., Capomaccio, S., Ajmone-Marsan, P., et al. (2018). Gene Co-expression Networks in Liver and Muscle Transcriptome Reveal Sex-specific Gene Expression in Lambs Fed with a Mix of Essential Oils. *BMC Genomics* 19 (1), 236. doi:10.1186/s12864-018-4632-y

Saira, K., Zhou, Y., and Jones, C. (2007). The Infected Cell Protein 0 Encoded by Bovine Herpesvirus 1 (bICP0) Induces Degradation of Interferon Response Factor 3 and, Consequently, Inhibits Beta Interferon Promoter Activity. *J. Virol.* 81 (7), 3077–3086. doi:10.1128/JVI.02064-06

Salem, E., Hägglund, S., Cassard, H., Corre, T., Näslund, K., Foret, C., et al. (2019). Pathogenesis, Host Innate Immune Response, and Aerosol Transmission of Influenza D Virus in Cattle. *J. Virol.* 93 (7), e01853–01818. doi:10.1128/JVI.01853-18

Salleh, S. M., Mazzoni, G., Løvendahl, P., and Kadarmideen, H. N. (2018). Gene Co-expression Networks from RNA Sequencing of Dairy Cattle Identifies Genes and Pathways Affecting Feed Efficiency. *BMC Bioinformatics* 19 (1), 513. doi:10.1186/s12859-018-2553-z

Salojin, K., and Oravecz, T. (2007). Regulation of Innate Immunity by MAPK Dual-Specificity Phosphatases: Knockout Models Reveal New Tricks of Old Genes. *J. Leukoc. Biol.* 81 (4), 860–869. doi:10.1189/jlb.1006639

Schaefer, A. L., Cook, N. J., Church, J. S., Basarab, J., Perry, B., Miller, C., et al. (2007). The Use of Infrared Thermography as an Early Indicator of Bovine Respiratory Disease Complex in Calves. *Res. Vet. Sci.* 83 (3), 376–384. doi:10.1016/j.rvsc.2007.01.008

Schlender, J., Bossert, B., Buchholz, U., and Conzelmann, K.-K. (2000). Bovine Respiratory Syncytial Virus Nonstructural Proteins NS1 and NS2 Cooperatively Antagonize Alpha/Beta Interferon-Induced Antiviral Response. *J. Virol.* 74 (18), 8234–8242. doi:10.1128/JVI.74.18.8234-8242.2000

Scott, M. A., Woolums, A. R., Swiderski, C. E., Perkins, A. D., Nanduri, B., Smith, D. R., et al. (2021). Comprehensive At-Arrival Transcriptomic Analysis of post-weaned Beef Cattle Uncovers Type I Interferon and Antiviral Mechanisms Associated with Bovine Respiratory Disease Mortality. *PLOS ONE* 16 (4), e0250758. doi:10.1371/journal.pone.0250758

Scott, M. A., Woolums, A. R., Swiderski, C. E., Perkins, A. D., Nanduri, B., Smith, D. R., et al. (2020). Whole Blood Transcriptomic Analysis of Beef Cattle at Arrival Identifies Potential Predictive Molecules and Mechanisms that Indicate

Animals that Naturally Resist Bovine Respiratory Disease. *PLOS ONE* 15 (1), e0227507. doi:10.1371/journal.pone.0227507

Sheridan, M. P., Regev-Shoshani, G., Martins, J., Vimalanathan, S., and Miller, C. (2016). Nitric Oxide Modulates the Immunological Response of Bovine PBMCs in an *In Vitro* BRDc Infection Model. *Res. Vet. Sci.* 109, 21–28. doi:10.1016/j.rvsc.2016.09.004

Silflow, R. M., Degel, P. M., and Harmsen, A. G. (2005). Bronchoalveolar Immune Defense in Cattle Exposed to Primary and Secondary challenge with Bovine Viral Diarrhea Virus. *Vet. Immunol. Immunopathology* 103 (1), 129–139. doi:10.1016/j.vetimm.2004.09.008

Singh, K., Ritchey, J. W., and Confer, A. W. (2011). Mannheimia Haemolytica. *Vet. Pathol.* 48 (2), 338–348. doi:10.1177/0300985810377182

Smirnova, N. P., Ptitsyn, A. A., Austin, K. J., Bielefeldt-Ohmann, H., Van Campen, H., Han, H., et al. (2009). Persistent Fetal Infection with Bovine Viral Diarrhea Virus Differentially Affects Maternal Blood Cell Signal Transduction Pathways. *Physiol. Genomics* 36 (3), 129–139. doi:10.1152/physiolgenomics.90276.2008

Smyth, G. K. (2005). "Limma: Linear Models for Microarray Data," in *Bioinformatics and Computational Biology Solutions Using R and Bioconductor*. Editors R. Gentleman, V. J. Carey, W. Huber, R. A. Irizarry, and S. Dudoit (New York, NY: Springer New York), 397–420.

Snowder, G. D., Van Vleck, L. D., Cundiff, L. V., and Bennett, G. L. (2006). Bovine Respiratory Disease in Feedlot Cattle: Environmental, Genetic, and Economic Factors. *J. Anim. Sci.* 84 (8), 1999–2008. doi:10.2527/jas.2006-046

Srikumaran, S., Kelling, C. L., and Ambagala, A. (2007). Immune Evasion by Pathogens of Bovine Respiratory Disease Complex. *Anim. Health Res. Rev.* 8 (2), 215–229. doi:10.1017/S1466252307001326

Sun, H.-Z., Srithayakumar, V., Jiminez, J., Jin, W., Hosseini, A., Raszek, M., et al. (2020). Longitudinal Blood Transcriptomic Analysis to Identify Molecular Regulatory Patterns of Bovine Respiratory Disease in Beef Cattle. *Genomics* 112 (6), 3968–3977. doi:10.1016/j.ygeno.2020.07.014

Szklarczyk, D., Gable, A. L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., et al. (2018). STRING V11: Protein-Protein Association Networks with Increased Coverage, Supporting Functional Discovery in Genome-wide Experimental Datasets. *Nucleic Acids Res.* 47 (D1), D607–D613. doi:10.1093/nar/gky1131

Takeda, K., and Akira, S. (2005). Toll-like Receptors in Innate Immunity. *Int. Immunol.* 17 (1), 1–14. doi:10.1093/intimm/dxh186

Takeuchi, O., and Akira, S. (2010). Pattern Recognition Receptors and Inflammation. *Cell* 140 (6), 805–820. doi:10.1016/j.cell.2010.01.022

Tang, D., Kang, R., Coyne, C. B., Zeh, H. J., and Lotze, M. T. (2012). PAMPs and DAMPs: Signal 0s that spur Autophagy and Immunity. *Immunol. Rev.* 249 (1), 158–175. doi:10.1111/j.1600-065X.2012.01146.x

Taylor, G., Wyld, S., Valarcher, J.-F., Guzman, E., Thom, M., Widdison, S., et al. (2014). Recombinant Bovine Respiratory Syncytial Virus with Deletion of the SH Gene Induces Increased Apoptosis and Pro-inflammatory Cytokines *In Vitro*, and Is Attenuated and Induces Protective Immunity in Calves. *J. Gen. Virol.* 95 (6), 1244–1254. doi:10.1099/vir.0.064931-0

Taylor, J. D., Fulton, R. W., Lehenbauer, T. W., Step, D. L., and Confer, A. W. (2010). The Epidemiology of Bovine Respiratory Disease: what Is the Evidence for Preventive Measures. *Can. Vet. J.* 51 (12), 1351–1359. doi:10.1128/CMR.16.1.79

Timsit, E., Dendukuri, N., Schiller, I., and Buczinski, S. (2016a). Diagnostic Accuracy of Clinical Illness for Bovine Respiratory Disease (BRD) Diagnosis in Beef Cattle Placed in Feedlots: A Systematic Literature Review and Hierarchical Bayesian Latent-Class Meta-Analysis. *Prev. Vet. Med.* 135, 67–73. doi:10.1016/j.prevetmed.2016.11.006

Timsit, E., Holman, D. B., Hallewell, J., and Alexander, T. W. (2016b). The Nasopharyngeal Microbiota in Feedlot Cattle and its Role in Respiratory Health. *Anim. Front.* 6 (2), 44–50. doi:10.2527/af.2016-0022

Tizioto, P. C., Kim, J., Seabury, C. M., Schnabel, R. D., Gershwin, L. J., Van Eenennaam, A. L., et al. (2015). Immunological Response to Single Pathogen Challenge with Agents of the Bovine Respiratory Disease Complex: An RNA-Sequence Analysis of the Bronchial Lymph Node Transcriptome. *PLOS ONE* 10 (6), e0131459. doi:10.1371/journal.pone.0131459

Turner, M. L., Cronin, J. G., Healey, G. D., and Sheldon, I. M. (2014). Epithelial and Stromal Cells of Bovine Endometrium Have Roles in Innate Immunity and Initiate Inflammatory Responses to Bacterial Lipopeptides *In Vitro* via Toll-like Receptors TLR2, TLR1, and TLR6. *Endocrinology* 155 (4), 1453–1465. doi:10.1210/en.2013-1822

Uribe-Querol, E., and Rosales, C. (2020). Phagocytosis: Our Current Understanding of a Universal Biological Process. *Front. Immunol.* 11, 1066. doi:10.3389/fimmu.2020.01066

van Dam, S., Võsa, U., van der Graaf, A., Franke, L., and de Magalhães, J. P. (2017). Gene Co-expression Analysis for Functional Classification and Gene-Disease Predictions. *Brief Bioinform* 19 (4), bbw139–592. doi:10.1093/bib/bbw139

Villaseñor, T., Madrid-Paulino, E., Maldonado-Bravo, R., Pérez-Martínez, L., and Pedraza-Alva, G. (2019). Mycobacterium Bovis BCG Promotes IL-10 Expression by Establishing a SYK/PKCα/β Positive Autoregulatory Loop that Sustains STAT3 Activation. *Pathog. Dis.* 77 (3), ftz032. doi:10.1093/femspd/ftz032

Viuff, B., Tjørnehøj, K., Larsen, L. E., Røntved, C. M., Uttenthal, A., Rønsholt, L., et al. (2002). Replication and Clearance of Respiratory Syncytial Virus. *Am. J. Pathol.* 161 (6), 2195–2207. doi:10.1016/S0002-9440(10)64496-3

Wang, S., Ma, X., Wang, H., and He, H. (2020a). Induction of the Unfolded Protein Response during Bovine Alphaherpesvirus 1 Infection. *Viruses* 12 (9), 974. doi:10.3390/v12090974

Wang, Y., Miao, L., Tao, L., Chen, J.-H., Zhu, C.-M., Li, Y., et al. (2020b). Weighted Gene Coexpression Network Analysis Identifies the Key Role Associated with Acute Coronary Syndrome. *Aging* 12 (19), 19440–19454. doi:10.18632/aging.103859

Wang, Z., Kong, L. C., Jia, B. Y., Chen, J. R., Dong, Y., Jiang, X. Y., et al. (2019). Analysis of the microRNA Expression Profile of Bovine Monocyte-Derived Macrophages Infected with *Mycobacterium avium* Subsp. Paratuberculosis Reveals that miR-150 Suppresses Cell Apoptosis by Targeting PDCD4. *Ijms* 20 (11), 2708. doi:10.3390/ijms20112708

Werners, A. H., and Bryant, C. E. (2012). Pattern Recognition Receptors in Equine Endotoxaemia and Sepsis. *Equine Vet. J.* 44 (4), 490–498. doi:10.1111/j.2042-3306.2012.00574.x

White, B. J., and Renter, D. G. (2009). Bayesian Estimation of the Performance of Using Clinical Observations and Harvest Lung Lesions for Diagnosing Bovine Respiratory Disease in Post-weaned Beef Calves. *J. VET. Diagn. Invest.* 21 (4), 446–453. doi:10.1177/104063870902100405

Wilkinson, A. (2009). Future of BRD Research: an Animal Health Industry Perspective. *Anim. Health Res. Rev.* 10 (2), 163–164. doi:10.1017/S1466252309990235

Wilkinson, J. M., Bao, H., Ladinig, A., Hong, L., Stothard, P., Lunney, J. K., et al. (2016). Genome-wide Analysis of the Transcriptional Response to Porcine Reproductive and Respiratory Syndrome Virus Infection at the Maternal/fetal Interface and in the Fetus. *BMC Genomics* 17 (1), 383. doi:10.1186/s12864-016-2720-4

Wu, C., Qin, X., Li, P., Pan, T., Ren, W., Li, N., et al. (2017). Transcriptomic Analysis on Responses of Murine Lungs to Pasteurella Multocida Infection. *Front. Cel. Infect. Microbiol.* 7, 251. doi:10.3389/fcimb.2017.00251

Xin, P., Xu, X., Deng, C., Liu, S., Wang, Y., Zhou, X., et al. (2020). The Role of JAK/STAT Signaling Pathway and its Inhibitors in Diseases. *Int. Immunopharmacology* 80, 106210. doi:10.1016/j.intimp.2020.106210

Xu, M., Liu, Y., Liu, Y., Li, X., Chen, G., Dong, W., et al. (2018). Genetic Polymorphisms of GZMB and Vitiligo: A Genetic Association Study Based on Chinese Han Population. *Sci. Rep.* 8 (1), 13001. doi:10.1038/s41598-018-31233-8

Xu, X., Zhang, K., Huang, Y., Ding, L., Chen, G., Zhang, H., et al. (2012). Bovine Herpes Virus Type 1 Induces Apoptosis through Fas-dependent and Mitochondria-Controlled Manner in Madin-Darby Bovine Kidney Cells. *Virol. J.* 9 (1), 202. doi:10.1186/1743-422X-9-202

Yan, Z., Huang, H., Freebern, E., Santos, D. J. A., Dai, D., Si, J., et al. (2020). Integrating RNA-Seq with GWAS Reveals Novel Insights into the Molecular Mechanism Underpinning Ketosis in Cattle. *BMC Genomics* 21 (1), 489. doi:10.1186/s12864-020-06909-z

Ye, S., Lowther, S., Stambas, J., and Sandri-Goldin, R. M. (2015). Inhibition of Reactive Oxygen Species Production Ameliorates Inflammation Induced by Influenza A Viruses via Upregulation of SOCS1 and SOCS3. *J. Virol.* 89 (5), 2672–2683. doi:10.1128/JVI.03529-14

Yip, A. M., and Horvath, S. (2007). Gene Network Interconnectedness and the Generalized Topological Overlap Measure. *BMC Bioinformatics* 8 (1), 22. doi:10.1186/1471-2105-8-22

Yuki, K., and Koutsogiannaki, S. (2021). Pattern Recognition Receptors as Therapeutic Targets for Bacterial, Viral and Fungal Sepsis. *Int. Immunopharmacology* 98, 107909. doi:10.1016/j.intimp.2021.107909

Zhang, B., and Horvath, S. (2005). A General Framework for Weighted Gene Co-expression Network Analysis. *Stat. Appl. Genet. Mol. Biol.* 4, 17. doi:10.2202/1544-6115.1128

Zhang, L., Huang, C., Guo, Y., Gou, X., Hinsdale, M., Lloyd, P., et al. (2015). MicroRNA-26b Modulates the NF- B Pathway in Alveolar Macrophages by Regulating PTEN. *J. Immunol.* 195 (11), 5404–5414. doi:10.4049/jimmunol.1402933

Zhao, X., Chu, H., Wong, B. H.-Y., Chiu, M. C., Wang, D., Li, C., et al. (2019). Activation of C-type Lectin Receptor and (RIG)-I-Like Receptors Contributes to Proinflammatory Response in Middle East Respiratory Syndrome Coronavirus-Infected Macrophages. *J. Infect. Dis.* 221 (4), 647–659. doi:10.1093/infdis/jiz483

Zheng, C.-H., Yuan, L., Sha, W., and Sun, Z.-L. (2014). Gene Differential Coexpression Analysis Based on Biweight Correlation and Maximum Clique. *BMC Bioinformatics* 15 (15), S3. doi:10.1186/1471-2105-15-S15-S3

Zheng, J., Yang, P., Tang, Y., Pan, Z., and Zhao, D. (2015). Respiratory Syncytial Virus Nonstructural Proteins Upregulate SOCS1 and SOCS3 in the Different Manner from Endogenous IFN Signaling. *J. Immunol. Res.* 2015, 1–11. doi:10.1155/2015/738547

Zhu, L., Thompson, J., Ma, F., Eudy, J., and Jones, C. (2017). Effects of the Synthetic Corticosteroid Dexamethasone on Bovine Herpesvirus 1 Productive Infection. *Virology* 505, 71–79. doi:10.1016/j.virol.2017.02.012

**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.

frontiers in Genetics

ORIGINAL RESEARCH
published: 16 November 2021
doi: 10.3389/fgene.2021.752129

# Proteomic Studies on the Mechanism of Myostatin Regulating Cattle Skeletal Muscle Development

*Hui Sheng[1], Yiwen Guo[1], Linlin Zhang[1], Junxing Zhang[1], Manning Miao[1], Haoyun Tan[1], Debao Hu[1], Xin Li[1], Xiangbin Ding[1], Guangpeng Li[2]* and Hong Guo[1]**

[1]Tianjin Key Laboratory of Agricultural Animal Breeding and Healthy Husbandry, College of Animal Science and Veterinary Medicine, Tianjin Agricultural University, Tianjin, China, [2]The Key Laboratory of Mammalian Reproductive Biology and Biotechnology of the Ministry of Education, Inner Mongolia University, Hohhot, China

Myostatin (MSTN) is an important negative regulator of muscle growth and development. In this study, we performed comparatively the proteomics analyses of gluteus tissues from MSTN$^{+/-}$ Mongolian cattle (MG.MSTN$^{+/-}$) and wild type Mongolian cattle (MG.WT) using a shotgun-based tandem mass tag (TMT) 6-plex labeling method to investigate the regulation mechanism of MSTN on the growth and development of bovine skeletal muscle. A total of 1,950 proteins were identified in MG.MSTN$^{+/-}$ and MG.WT. Compared with MG.WT cattle, a total of 320 differentially expressed proteins were identified in MG.MSTN cattle, including 245 up-regulated differentially expressed proteins and 75 down-regulated differentially expressed proteins. Bioinformatics analysis showed that knockdown of the MSTN gene increased the expression of extracellular matrix and ribosome-related proteins, induced activation of focal adhesion, PI3K-AKT, and Ribosomal pathways. The results of proteomic analysis were verified by muscle tissue Western blot test and *in vitro* MSTN gene knockdown test, and it was found that knockdown MSTN gene expression could promote the proliferation and myogenic differentiation of bovine skeletal muscle satellite cells (BSMSCs). At the same time, Co-Immunoprecipitation (CO-IP) assay showed that MSTN gene interacted with extracellular matrix related protein type I collagen α 1 (COL1A1), and knocking down the expression of COL1A1 could inhibit the activity of adhesion, PI3K-AKT and ribosome pathway, thus inhibit BSMSCs proliferation. These results suggest that the MSTN gene regulates focal adhesion, PI3K-AKT, and Ribosomal pathway through the COL1A1 gene. In general, this study provides new insights into the regulatory mechanism of MSTN involved in muscle growth and development.

**Keywords: myostatin, proteomics, extracellular matrix, ribosome, focal adhesion**

OPEN ACCESS

***Edited by:***
*Andressa Oliveira De Lima, University of Washington, United States*

***Reviewed by:***
*Loan To Nguyen, University of Queensland, Australia*
*Kerong Shi, Shandong Agricultural University, China*

****Correspondence:***
*Hong Guo*
*guohong64@163.com*
*Guangpeng Li*
*gpeng@imu.edu.cn*

***Specialty section:***
*This article was submitted to Livestock Genomics, a section of the journal Frontiers in Genetics*

***Received:*** *02 August 2021*
***Accepted:*** *28 October 2021*
***Published:*** *16 November 2021*

***Citation:***
*Sheng H, Guo Y, Zhang L, Zhang J, Miao M, Tan H, Hu D, Li X, Ding X, Li G and Guo H (2021) Proteomic Studies on the Mechanism of Myostatin Regulating Cattle Skeletal Muscle Development. Front. Genet. 12:752129. doi: 10.3389/fgene.2021.752129*

## INTRODUCTION

Myostatin (MSTN), also known as growth differentiation factor-8 (GDF-8), is a highly conservative member of the transforming growth factor β (TGF-β) superfamily (McPherron et al., 1997). In previous studies, MSTN has been confirmed to be a secreted growth factor expressed predominantly in skeletal muscle (McPherron, 1997; Kollias and McDermott, 2008) and plays a key role in the negative regulation of muscle development (Tsuchida, 2008). *In vitro* and *in vivo* studies have shown

that the Mstn signal is mediated by binding to activin receptor type-IIB (ActRIIB). Mstn has also been shown to bind to ActRIIB, although the functional correlation of this interaction has not been fully determined (Lee and McPherron, 2001). MSTN deficiency resulting from genetic ablation of both copies ($MSTN^{-/-}$) or single copy ($MSTN^{+/-}$) of the germline allele or loss-of-function mutations usually leads to a "double muscle" phenotype, mainly characterized by a significant increase in muscle mass (McPherron, 1997; McPherron et al., 1997). Similar phenotypes were also demonstrated in natural mutations of MSTN genes in cattle (Grobet et al., 1997; McPherron, 1997), sheep (Clop et al., 2006), dogs (Mosher et al., 2007), and humans (Schuelke et al., 2004). On the contrary, overexpression of MSTN or systemic administration can lead to muscle atrophy (Zimmers et al., 2002; Durieux et al., 2007). All these effects are mainly achieved by regulating the proliferation and differentiation of myoblasts (Thomas et al., 2000; Rios et al., 2001; Langley et al., 2002; Rios et al., 2002). MSTN has been shown to interfere with proliferation and protein synthesis and protein decomposition of adult muscle fibers (Sartori et al., 2009; Trendelenburg et al., 2009). Many studies have shown that MSTN can regulate IGF-I signal pathway (Morissette et al., 2009), WNT4/β-catenin signal pathway (Steelman et al., 2006), Erk1/2, c-Jun N-terminal kinase (JNK) signal pathway, p38 mitogen-activated protein (MAP)K (Allendorph et al., 2006; Huang et al., 2007; Joulia-Ekaza and Cabello, 2007; Elkasrawy and Hamrick, 2010), and PI3K/Akt signal pathway (Lipina et al., 2010) through transforming growth factor-β pathway (Kollias and McDermott, 2008; Elkasrawy and Hamrick, 2010). Besides, muscle somatostatin mediates CKD-induced muscle catabolism by coordinating the activation of autophagy and the ubiquitin-proteasome system and may limit cell proliferation by activating miRNAs (Wang et al., 2015; Huang et al., 2019). In addition to its effects on muscle development, MSTN also plays an important role in fat and glucose metabolism. The results of previous studies in our laboratory showed that knockout MSTN gene expression enhanced glycolysis and fat β-oxidation in bovine muscle tissue (Yang et al., 2018; Xin et al., 2020). High levels of MSTN have been shown to cause muscle atrophy and are associated with a variety of diseases, such as cancer (Aversa et al., 2012), chronic obstructive pulmonary disease (Hayot et al., 2011), chronic heart failure (Lenk et al., 2009), acquired immune deficiency syndrome (Gonzalez-Cadavid et al., 1998), obesity, insulin resistance, and type 2 diabetes (McPherron and Lee, 2002; Zhang et al., 2011; Zhang et al., 2012). Therefore, MSTN dysfunction has been considered a promising strategy for animal breeding or for fighting muscle atrophy in different diseases, including neuromuscular diseases (Mariot et al., 2017). Although MSTN signal cascade plays a central role in regulating muscle weight, the mechanism of this signal cascade is still unclear (Elkina et al., 2011).

Proteins play an important role in many types of molecular networks and perform most of the biochemical functions of organisms. Label-free liquid chromatography-mass spectrometry (LC-MS/MS) can be used to quantify and identify thousands of proteins in multiple samples in one operation, which provides an unprecedented opportunity to study the proteomic changes of biological components or organisms (Wang et al., 2019). Protein de/phosphorylation is a ubiquitous post-translational modification, which plays a regulatory role in protein structure, function, cell signal transduction, and enzyme activity regulation. Phosphorylation usually occurs on serine, threonine, and tyrosine residues and is catalyzed by upstream protein kinases by adding covalently bound phosphate groups (Chen et al., 2019). Some studies have shown that reversible protein phosphorylation plays an important role in the transformation of muscle to meat by regulating the development of meat quality by regulating proteins involved in glycolysis and muscle contraction (Huang et al., 2011; Huang et al., 2012; Li et al., 2012; Huang et al., 2014; Li et al., 2015a; Li et al., 2015b; Chen et al., 2016). Protein phosphorylation of various muscle samples has been reported, for example, myosin and myosin regulatory light chain 2 have been identified (Obermann et al., 1997; Muroya et al., 2007). Because there are many proteins in the pathway affected by MSTN, the single-target study can not fully analyze its function and mechanism, so proteomics and detection of the phosphorylation level of key proteins in signal pathway can be used to comprehensively and effectively analyze its mechanism or regulatory proteins.

In this study, in order to explore the mechanism of MSTN gene regulating signal pathway during muscle development, we used 6-plex TMT labeling method to determine the global protein abundance in the gluteal muscle tissue of $MG.MSTN^{+/-}$ cattle with artificially knocked down expression of MSTN gene and MG.WT cattle with normal expression of MSTN gene. These proteomic analyses enable us to compare quantitative changes in overall protein abundance. After statistical and bioinformatics analysis, based on the results of protein abundance changes related to MSTN gene deletion, we carried out Western blot tests of muscle tissue samples and *in vitro* experiments and confirmed that the data sets from global proteomics are reliable.

# RESULTS

## Statistical Analysis of Mass Spectrometry Results

$MSTN^{+/-}$ associated changes in abundance of any identified proteins were determined based on the TMT 6-plex reporter ion ratios. The results of mass spectrometry analysis showed that a total of 1,950 quantitative proteins were identified in the gluteal muscle tissues of $MG.MSTN^{+/-}$ cattle and MG.WT cattle, and each protein contained at least two unique peptides (Supplement Table 3A). After scatter plotting analysis (**Supplementary Figure S1**) used for determining the internal error of the biological replicates and student t-test analysis of the data set (Supplement Table 3B), the fold change in values more than 1.3 were determined based on the value of the log2 TMT ratio (log2 1.3 = 0.38) at which 95% of all proteins had no deviation (Martin et al., 2016). Thus, the fold-change (≥1.3) and *p*-value (≤0.05) from the *t*-test were applied to rank and filter the quantitative data. The proteins with fold change ≥1.30 or ≤0.77 in relative abundance and a $p$-value ≤ 0.05 were identified as differentially expressed proteins (Supplement Table 3B).

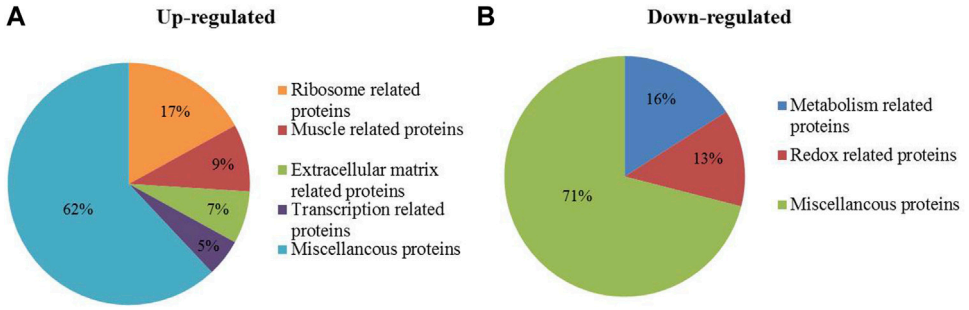


**FIGURE 1 |** The biological functions of 320 differentially expressed proteins identified in MG.MSTN+/– vs MG.WT data. **(A, B)** show the proportions of biological functions among 245 up-regulated proteins and 75 down-regulated proteins.

## Identification of Differentially Expressed Proteins Between MG.MSTN$^{+/-}$ and MG.WT Cattle

Compared with wild MG.WT, 320 differentially expressed proteins were identified in MG.MSTN$^{+/-}$ group, including 245 proteins with increased abundance and 75 proteins with decreased abundance (**Supplementary Table S3**), which were used in subsequent bioinformatics analysis and verification experiments. GO analysis showed that the functions of 245 up-regulated proteins were divided into ribosome-related proteins (17%), muscle-related proteins (9%), extracellular matrix-related proteins (7%), and transcription-related proteins (5%) (**Figure 1A**). While the 75 down-regulated proteins were classified as metabolism-related proteins (16%) and redox-related proteins (13%) (**Figure 1B**). Interestingly, we found that the expression abundance of a large number of extracellular matrix related proteins (such as collagen alpha-1(I) chain, laminin subunit beta 1) and ribosomal related proteins (such as 40S ribosomal protein S6, 60S ribosomal protein L26) increased in MG.MSTN$^{+/-}$ (**Table 1**). This result is expected to activate focal adhesion, PI3K-AKT, and ribosomal pathways, enhance actin-binding and protein synthesis, and promote muscle contraction and growth.

## Functional Analysis of Differentially Expressed Proteins Between MG.MSTN$^{+/-}$ and MG.WT Cattle

The results of the GO analysis are shown in **Figure 2A** with $p \leq 0.01$ as a significant threshold. In the GO annotation and KEGG analysis of 320 differentially expressed proteins, the translation, focal adhesion, and structural constituent of ribosome were significantly enriched as the primary categories of biological process (BP), cellular component (CC), and molecular functional (MF), respectively. At the same time, ribosome, ECM receptor interaction, focal adhesion, and other pathways have also been greatly enriched in the current KEGG database. Using the STRING database to further analyze the possible protein-protein interaction networks of these changed proteins, it was found that there were three major protein-protein interaction networks among the 320 differentially expressed proteins (**Figure 2B**). The largest network highlighted with an pink background in **Figure 2B** includes more than 60 proteins known to be associated with translation (such as eukaryotic translation initiation factor 5A-1 and 40S ribosomal protein S21). The second network with a violet background contains about 30 proteins including myosin light polypeptide 6 and is mainly involved in muscle contraction. The third network with a green background includes about 20 extracellular matrix related proteins (such as collagen alpha-1(I) chain).

## Tissue Samples Western Blot to Verify the Accuracy of Data Analysis

By comparing the proteomic data of the gluteal muscle of MG.MSTN$^{+/-}$ and MG.WT, it was found that knocking down the expression of MSTN increased the expression of a large number of extracellular matrix related proteins and ribosomal related proteins. And these proteins are important components of focal adhesion, PI3K-AKT, and ribosomal pathway. Therefore, we speculate that knocking down the expression of the MSTN gene may activate focal adhesion, PI3K-AKT, and ribosomal pathway. To verify the TMT-based quantitative proteomics results, we used classical Western blot analyses to validate the accuracy of data analysis. Considering the availability of some antibodies, we performed Western blot analysis of MSTN, one differentially expressed protein (MYL6), and alpha-tubulin (for internal control). As shown in **Figure 3**, the expression trend of Western blot results is consistent with that of data analysis (supplement Table 3), indicating that quantitative proteomics data have reasonable accuracy. At the same time, the expressions of key proteins pFAK, pAKT, and pRPS6 in focal adhesion, PI3K-AKT, and ribosomal pathway were verified by Western blot. The results showed that the expression of pFAK, pAKT, and pRPS6 in the gluteal muscle of MG.MSTN$^{+/-}$ was significantly higher than that of WT Mongolian cattle (**Figure 3**), which was consistent with our conjecture.

## Knockdown of Myostatin Expression Promotes Proliferation and Myogenic Differentiation of Bovine Skeletal Muscle Satellite Cells

To explore the mechanism of the effect of the MSTN gene on the development of bovine skeletal muscle, this experiment studied the effect of MSTN knockdown on the proliferation

**TABLE 1 |** A partial list of the differentially expressed proteins involving extracellular matrix related proteins and ribosomal related in MG.MSTN$^{+/-}$ vs MG.WT

| Protein accession | Protein description | Protein name | Fold-change (MG.MSTN$^{+/-}$ vs MG.WT) | Regulated |
|---|---|---|---|---|
| **Extracellular matrix related proteins** | | | | |
| 77404252 | collagen alpha-1(I) chain precursor | COL1A1 | 2.32 | Up |
| 27806257 | collagen alpha-2(I) chain precursor | COL1A2 | 2.95 | Up |
| 528938209 | PREDICTED: collagen alpha-2(VI) chain isoform X1 | COL6A2 | 1.81 | Up |
| 528945453 | PREDICTED: collagen alpha-3(VI) chain isoform X2 | COL6A3 | 2.26 | Up |
| 982945253 | PREDICTED: collagen alpha-1 (XIV) chain isoform X1 | COL14A1 | 1.66 | Up |
| 116003881 | collagen alpha-1(III) chain precursor | COL3A1 | 1.52 | Up |
| 300796391 | collagen alpha-1 (XV) chain precursor | COL15A1 | 1.43 | Up |
| 982987742 | PREDICTED: laminin subunit alpha-2, partial | LAMA2 | 2.03 | Up |
| 329664360 | laminin subunit alpha-4 precursor | LAMA4 | 1.93 | Up |
| 330688474 | laminin subunit beta-1 precursor | LAMB1 | 2.06 | Up |
| 332205887 | laminin subunit gamma-1 precursor | LAMC1 | 1.62 | Up |
| 982993396 | PREDICTED: LOW QUALITY PROTEIN: laminin subunit alpha-5 isoform X2 | LAMA5 | 0.58 | Down |
| 78045497 | vitronectin precursor | VTN | 0.70 | Down |
| **Ribosomal related proteins** | | | | |
| 66792868 | 40S ribosomal protein S15 | RPS15 | 1.85 | Up |
| 528974050 | PREDICTED: 40S ribosomal protein S21 isoform X1 | RPS21 | 1.68 | Up |
| 77797830 | 40S ribosomal protein S10 | RPS10 | 1.66 | Up |
| 70778778 | 40S ribosomal protein S13 | RPS13 | 1.57 | Up |
| 529010831 | PREDICTED: 40S ribosomal protein S24 isoform X1 | RPS24 | 1.56 | Up |
| 75812924 | 40S ribosomal protein S18 | RPS18 | 1.53 | Up |
| 82697365 | 40S ribosomal protein S19 | RPS19 | 1.53 | Up |
| 149642623 | 40S ribosomal protein S7 | RPS7 | 1.47 | Up |
| 62752040 | 40S ribosomal protein S6 | RPS6 | 1.38 | Up |
| 149642675 | 40S ribosomal protein S17 | RPS17 | 1.36 | Up |
| 62461611 | 40S ribosomal protein S12 | RPS12 | 1.36 | Up |
| 155372029 | 40S ribosomal protein S9 | RPS9 | 1.34 | Up |
| 70778964 | 40S ribosomal protein S25 | RPS25 | 1.32 | Up |
| 70778960 | 40S ribosomal protein S28 | RPS28 | 1.31 | Up |
| 741965251 | PREDICTED: 28S ribosomal protein S27, mitochondrial isoform X1 | MRPS27 | 1.71 | Up |
| 982996641 | PREDICTED: 60S ribosomal protein L22 isoform X1 | RPL22 | 1.64 | Up |
| 62751887 | 60S ribosomal protein L26 | RPL26 | 1.58 | Up |
| 77404275 | 60S ribosomal protein L27 | RPL27 | 1.53 | Up |
| 70778766 | 60S ribosomal protein L31 | RPL31 | 1.48 | Up |
| 77735941 | 60S ribosomal protein L35 | RPL35 | 1.48 | Up |
| 27806129 | 60S ribosomal protein L24 | RPL24 | 1.45 | Up |
| 114051890 | 60S ribosomal protein L23a | RPL23A | 1.45 | Up |
| 528936538 | PREDICTED: 60S ribosomal protein L35a isoform X1 | RPL35A | 1.44 | Up |
| 27807523 | 60S acidic ribosomal protein P2 | RPLP2 | 1.42 | Up |
| 529001374 | PREDICTED: 60S ribosomal protein L29 isoform X1 | RPL29 | 1.40 | Up |
| 78042478 | 60S ribosomal protein L37a | RPL37A | 1.38 | Up |
| 528993469 | PREDICTED: 60S ribosomal protein L28 isoform X1 | RPL28 | 1.37 | Up |
| 528988589 | PREDICTED: 60S ribosomal protein L6 isoform X1 | RPL6 | 1.35 | Up |
| 62751646 | 60S ribosomal protein L13 | RPL13 | 1.33 | Up |
| 62460552 | 60S ribosomal protein L7 | RPL7 | 1.32 | Up |
| 94966839 | 60S ribosomal protein L7a | RPL7A | 1.32 | Up |

*Note: Corrected* p*-value ≤ 0.05, Fold change ≥1.30 or ≤0.77.*

and myogenic differentiation of BSMSCs. SiRNA (si-MSTN) for MSTN knockdown was designed and synthesized which resulted in a significant decrease in the expression of MSTN in cells GM and DM3 respectively (**Figure 4A**). Meanwhile, proliferation-related markers Pax7 and MyoD were detected to evaluate whether cell proliferation was affected. The results showed that the expression of MSTN was silenced, the mRNA expression levels of Pax7 did not change significantly, the mRNA expression levels of MyoD was significantly up-regulated (**Figure 4B**); the protein expression levels of Pax7 did not change significantly, and the protein expression levels of MyoD was significantly up-regulated (**Figure 4C**). The results of the 5-ethynyl-2′-deoxyuridine (EdU) cell proliferation assay showed that when MSTN was knocked down, the number of EDU positive cells (**Figure 4D**) and EDU labeling index (**Figure 4E**) were all up-regulated. Meanwhile, the differentiation process of satellite cells was observed under the light microscope. Compared with the wild group (WT) and control group (si-NC), when MSTN was knocked down, thick myotubes were formed in DM3 (**Figure 5A**). The expression levels of differentiation-related markers MyoG and MyHC were detected. The results showed

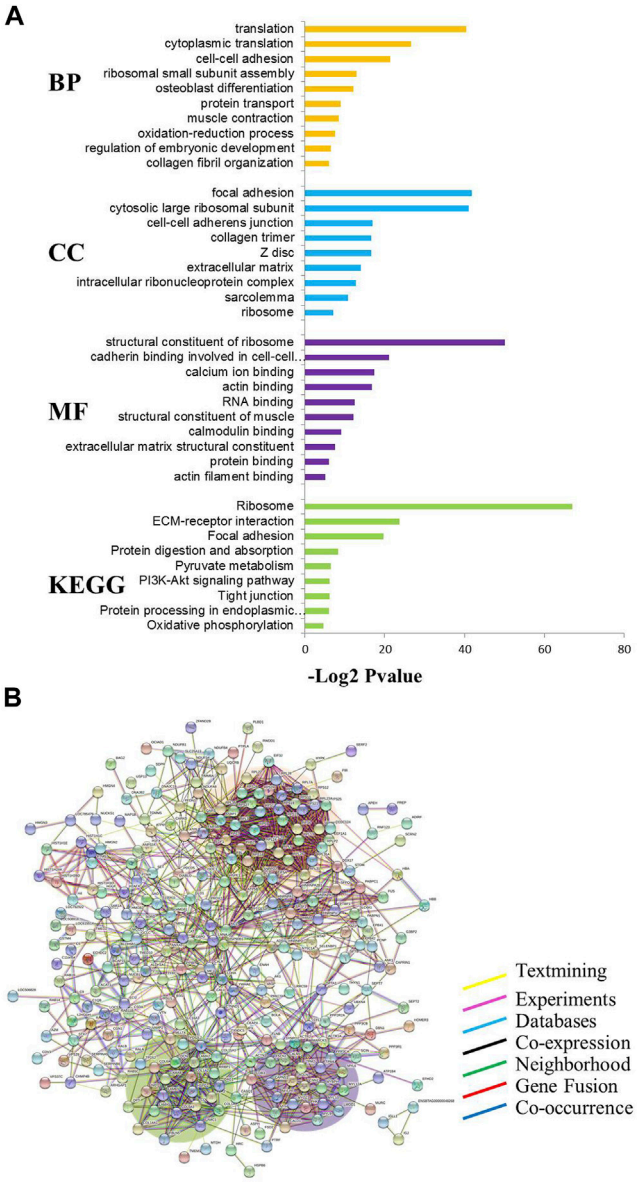


**FIGURE 2 |** Functional classification of differentially expressed proteins in MG.MSTN+/– vs MG.WT. **(A)** GO and KEGG analyses. BP: Biological process, CC: Cellular component, MF: Molecular function, KEGG: Kyoto encyclopedia of genes and genomes. The values given in each of the enriched terms or pathways are corrected at $p \leq 0.01$. **(B)** Protein-protein interaction analysis.

that the expression of MSTN was silenced, the mRNA expression level of MyoG did not change significantly, but the mRNA expression level of MyHC was significantly up-regulated (**Figure 5B**); the protein expression levels of MyoG and MyHC were all significantly up-regulated (**Figures 5C,D**). To sum up, our experiments showed that knocking down the expression of the MSTN gene promotes the proliferation and myogenic differentiation of BSMSCs.

## Knockdown of Myostatin Expression Activates Focal Adhesion, PI3K-AKT, and Ribosomal Pathway

To further verify the accuracy of the data analysis results, to explore the effects of the MSTN gene on focal adhesion, PI3K-AKT, and ribosomal pathway. The RNA and protein were extracted from GM and DM3 cells transfected with si-MSTN and were used to detect the differentially expressed proteins involved in focal adhesion, PI3K-AKT, and ribosomal pathways in the gluteal muscles of MG.MSTN$^{+/-}$ and MG.WT, as well as the expression levels of key genes FAK, AKT, RPS6, pFAK, pAKT, and pRPS6 in these pathways. The results showed that the mRNA expression level and protein expression level of these differentially expressed proteins, as well as FAK, AKT, and RPS6 were mostly up-regulated in GM and DM3 cells silenced by MSTN. Meanwhile, it was found that the protein expression levels of p-FAK, p-AKT, and p-RPS6 were also significantly increased (**Figure 6**, **Figure 7**). The test results are consistent with speculation, which proves that knocking down the MSTN gene expression can activate focal adhesion, PI3K-AKT, and ribosomal pathways.

## Verification of Interaction Between Myostatin and COL1A1

To explore the mechanism of MSTN acting on downstream target proteins to regulate Focal adhesion, PI3K-AKT, and Ribosomal pathways. By comparing proteomics analysis, we found that compared with MG.WT, the expression of a large number of extracellular matrix-related proteins in the gluteal muscle tissues of MG.MSTN$^{+/-}$ was up-regulated, among which the expression of COL1A1 The levels were raised by 2.3 times (see **Supplementary Table S3B**). At the same time, in GM and DM3 BSMSCs transfected with si-MSTN, the mRNA expression level of the COL1A1 gene did not change significantly (**Figure 8A**), but the protein expression level was significantly increased (**Figures 8B,C**). To investigate the interaction between MSTN and COL1A1, wild-type GM BSMSCs were collected. CO-IP was performed against MSTN using an anti-MSTN antibody. COL1A1 protein was detected in the eluent of the immunocomplex by western blot using an anti-COL1A1 antibody (**Figure 8D**), indicating that MSTN interacted with COL1A1.

## Knockdown of COL1A1 Expression Promotes Proliferation of Bovine Skeletal Muscle Satellite Cells

After confirming the interaction between MSTN and COL1A1, we tested the effect of knocking down COL1A1 expression on the proliferation of BSMSCs. Designed and synthesized siRNA (si-COL1A1) for COL1A1 knockdown, which resulted in a very significant decrease in the mRNA expression level and protein expression level of COL1A1 in GM BSMSCs (**Figure 9A**). The expression levels of Pax7 and MyoD, which are related to proliferation, were detected. The results showed that after silencing COL1A1 expression, the mRNA expression level of Pax7 was significantly decreased, the mRNA expression level of MyoD was significantly increased (**Figure 9B**), and the protein expression levels

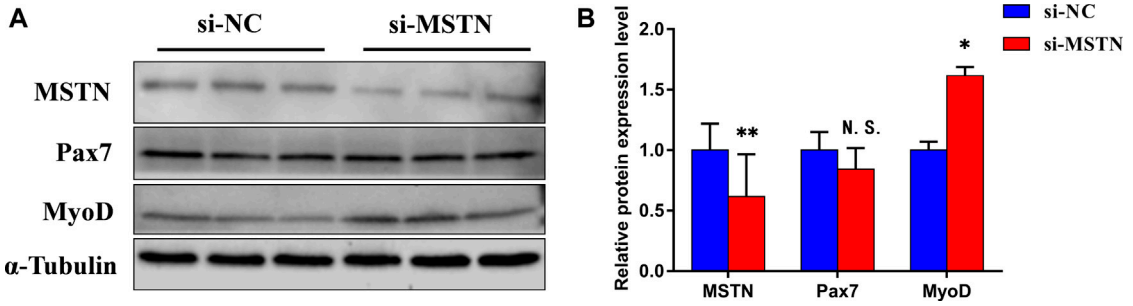


**FIGURE 3 |** Comparison and verification of quantitative results of phosphorylation levels of differentially expressed proteins and key proteins in the pathway from Western blotting analysis. **(A)** Western blotting results of MSTN,MYL6, pFAK (Tyr-473), pAKT1 (Ser-473) and pRPS6 (Ser-235/236) proteins. **(B)** Results of muscle tissue protein Western blot quantification. These imprints are cropped and the original image is shown in **Supplementary Figure S2**.

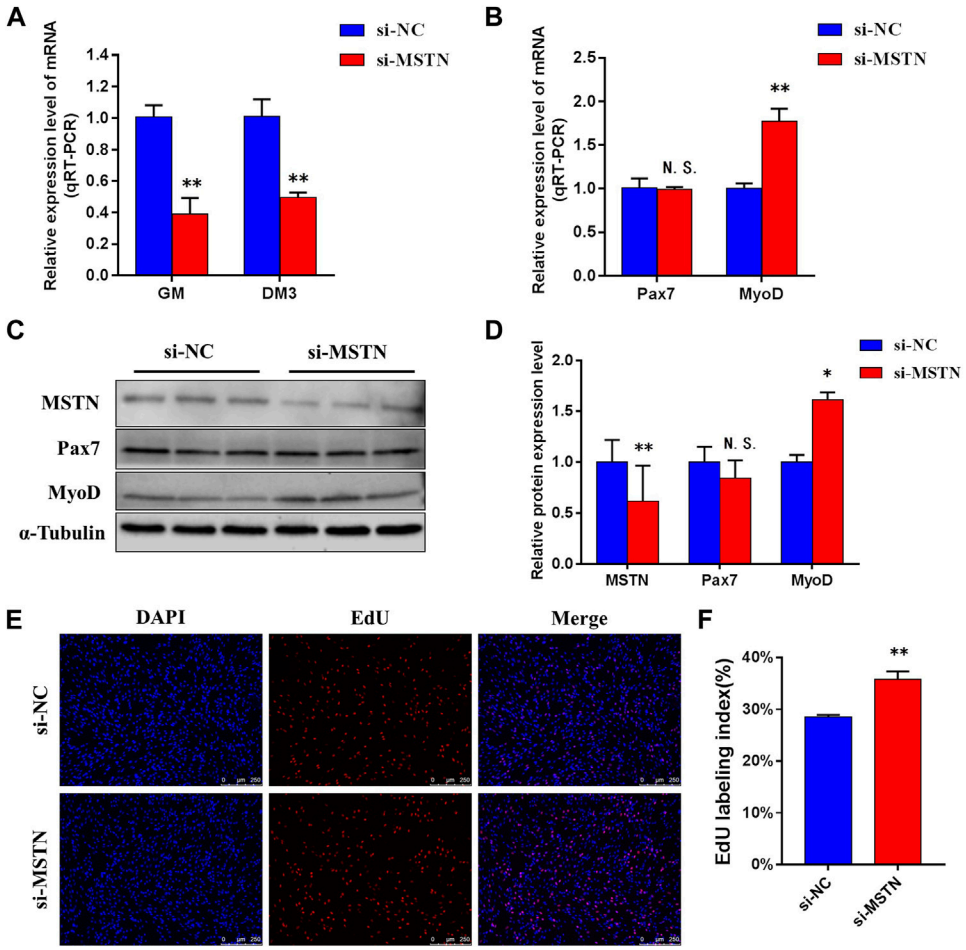


**FIGURE 4 |** Knockdown of MSTN expression promotes the maintenance of bovine skeletal muscle satellite cells (BSMSCs) proliferation. **(A)** Knockdown MSTN significantly decreased MSTN expression level in proliferative phase (GM) and the third day of differentiation (DM3). **(B)** resulting in no change in the expression level of mRNA of Pax7, but significantly increased the expression level of mRNA of MyoD. **(C–D)** no significant change of protein expression of Pax7, and significant increase of protein expression of MyoD. (E–F) EdU assay was performed at 24 h after transfection. The number of 5-ethynyl-2′-deoxyuridine (EdU)-positive cells (E, ×200; scale bars 250 μm) and EdU labeling index increased with MSTN knockdown. These blots were cropped and original images were shown in **Supplementary Figure S3**.

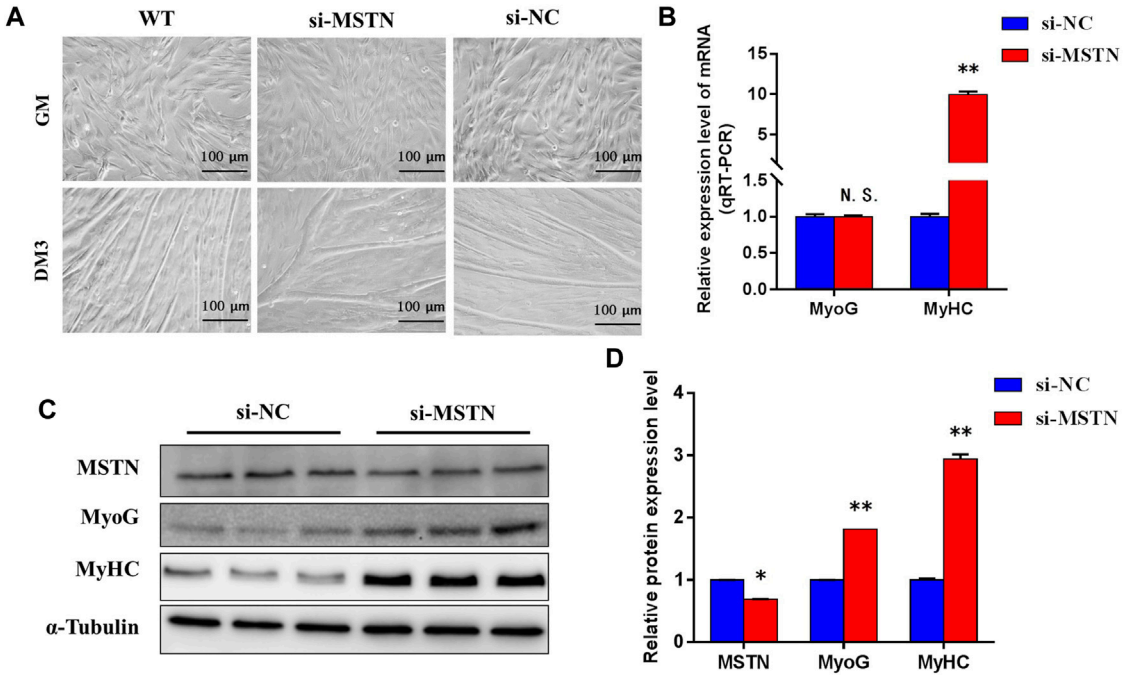


**FIGURE 5 |** MSTN functions in the differentiation of bovine myoblasts. **(A)** Knockdown of MSTN expression promoted the differentiation process of bovine myoblasts (×200; scale bars 100 μm). **(B)** Knockdown experiments resulted in no change in the mRNA expression level of MyoG, but increased the expression level of mRNA of MyHC. **(C–D)** The protein expression levels of MyoG and MHC were significantly increased. These blots were cropped and original images were shown in **Supplementary Figure S4**.

of Pax7 and MyoD were significantly decreased (**Figures 9C,D**). The results of the 5-ethynyl-2′-deoxyuridine (EdU) cell proliferation assay showed that when COL1A1 was knocked down, the number of EDU positive cells (**Figure 9E**) and EDU labeling index (**Figure 9F**) were all down-regulated. In summary, our experiments show that COL1A1 is essential for maintaining the proliferation of BSMSCs.

## Knockdown of COL1A1 Expression Inhibits Focal Adhesion, PI3K-AKT, and Ribosomal Pathways

The effects of knocking down COL1A1 on Focal adhesion, PI3K-AKT, and Ribosomal pathway were detected. The results showed that in BSMSCs transfected with si-COL1A1, the mRNA expression levels of majority differentially expressed proteins involved in focal adhesion, PI3K-AKT, and ribosomal pathways did not change significantly (**Figure 10A**), but the protein expression levels decreased significantly, especially the protein expression levels of key genes p-FAK, p-AKT1 and p-RPS6 in the pathway (**Figures 10B,C**). It shows that knockdown of COL1A1 expression inhibits Focal adhesion, PI3K-AKT, and Ribosomal pathways.

# DISCUSSION

Our group has a long-standing research interest in understanding the molecular mechanisms of MSTN knockout in transgenic cattle. The main purpose of these studies is to improve livestock breeds to obtain the best economic benefits. We demonstrated that MSTN knockout promoted skeletal muscle growth in Luxi-MSTN$^{-/-}$ cattle, which showed a higher percentage of lean meat (Yang et al., 2018). It was found that MSTN knockout promoted fatty acid β oxidation and glycolysis in MSTN$^{-/-}$ Inner Mongolia black beef cattle (Xin et al., 2020). In order to further explore the global regulatory mechanism of MSTN related to skeletal muscle growth and development, in the present study, the global proteomics of gluteus muscle of MG.MSTN$^{+/-}$ and MG. WT were analyzed. We believe that the differentially expressed proteins identified in MG.MSTN$^{+/-}$ vs MG.WT group data are caused by MSTN gene knockdown. Through bioinformatics analysis, it was found that the expression abundance of extracellular matrix-related proteins and ribosome-related proteins in the gluteal muscle of MG.MSTN$^{+/-}$ was higher than that of MG.WT (**Supplementary Table S3B**). Extracellular matrix-related proteins and ribosome-related proteins are important components of focal adhesion, PI3K-AKT, and ribosomal pathways. Therefore, we speculate that MSTN may regulate focal adhesion, PI3K-AKT, and ribosomal pathway.

Extracellular matrix (ECM) is a non-cellular three-dimensional macromolecular network composed of collagens, elastin, fibronectin (FN), laminins, glycoproteins, proteoglycans (PGs), glycosaminoglycans (GAGs), and several other glycoproteins (Theocharis et al., 2016; Birch, 2018; Karamanos et al., 2018; Muncie and Weaver, 2018). The ECM not only

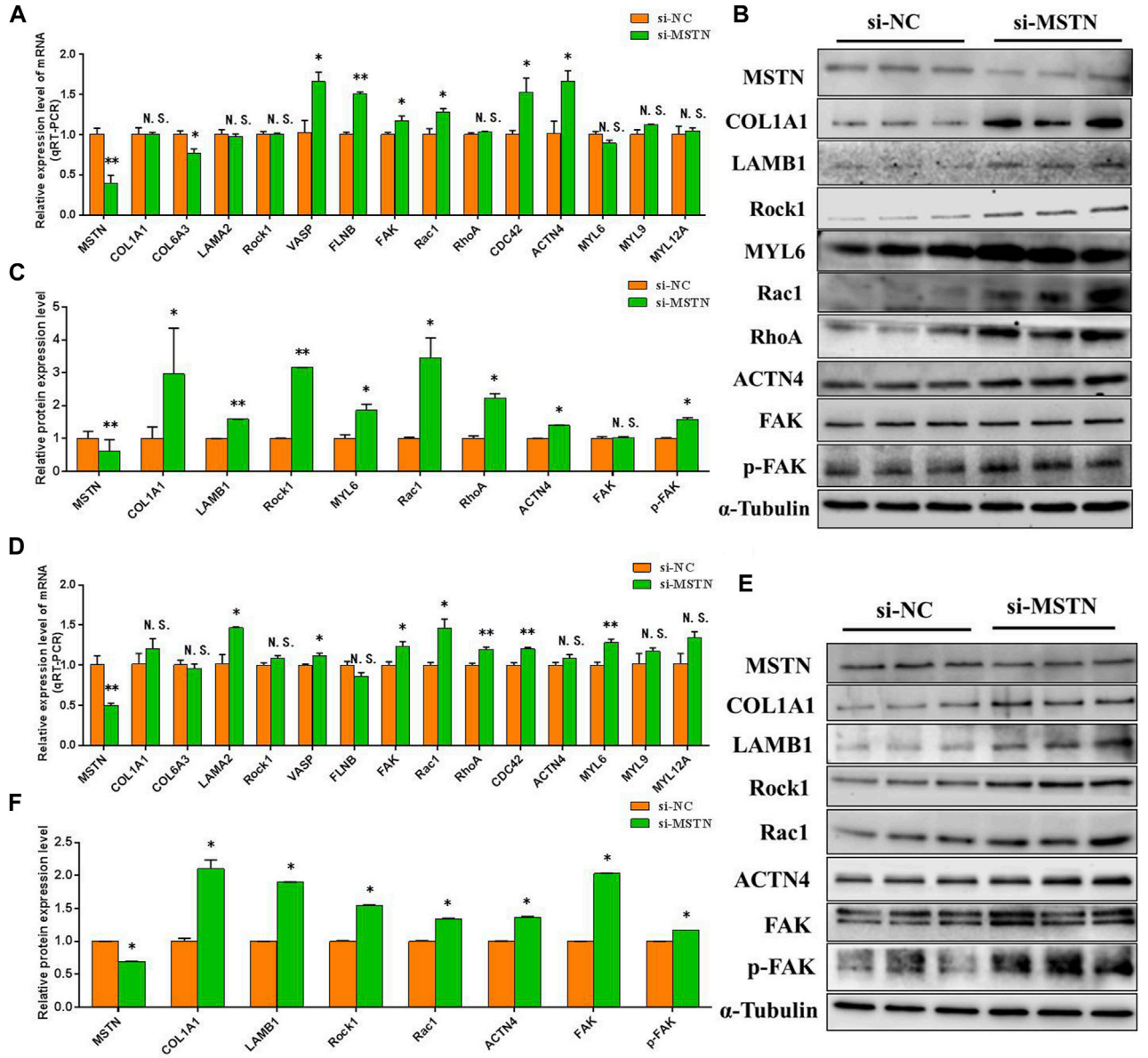


**FIGURE 6 |** The effect of MSTN on the focal adhesion pathway. **(A–C)** Knockdown of MSTN expression resulted in a significant increase in the mRNA expression level and protein expression level of the majority of differentially expressed proteins involved in the focal adhesion pathway in GM, as well as the expression level of the key gene FAK and p-FAK in the pathway. **(D–F)** The mRNA expression level and protein expression level of these proteins were also generally increased in DM3. These blots were cropped and original images were shown in **Supplementary Figure S5** and **Figure 6**.

provides physical scaffolds embedded by cells but also regulates many cellular processes, including survival, migration, growth, differentiation, homeostasis, and morphogenesis (Gumbiner, 1996; Frantz et al., 2010; Clause and Barker, 2013). The ECM is an important source of growth factors, providing the attachment and controlled release of many growth factors and signal molecules, thereby aiding cell-to-cell signaling (Birch, 2018). Transcriptome studies showed that in the rat model of overload, the expression abundance of the MSTN gene decreased, a large number of extracellular matrix-related proteins increased, and skeletal muscle increased (Mendias et al., 2017; Stantzou et al., 2021). Biological scaffolds composed of ECM have been shown to facilitate the functional reconstruction of several tissue types including the esophagus (Badylak et al., 2005; Desai et al., 2006), heart and vascular structures (Badylak et al., 2006; Quarti et al., 2011), lower urinary tract (Reddy et al., 2000; Boruch et al., 2010), and musculoskeletal tissues (Smith et al., 2004; Franklin et al., 2008; Turner et al., 2010; Valentin et al., 2010; Zhao and Bass, 2018), among others. Mutations in COL1A1 cause osteogenic insufficiency (OI), resulting in decreased muscle mass and function (Veilleux et al., 2014; Veilleux et al., 2015). The extracellular matrix is an important part of focal adhesions

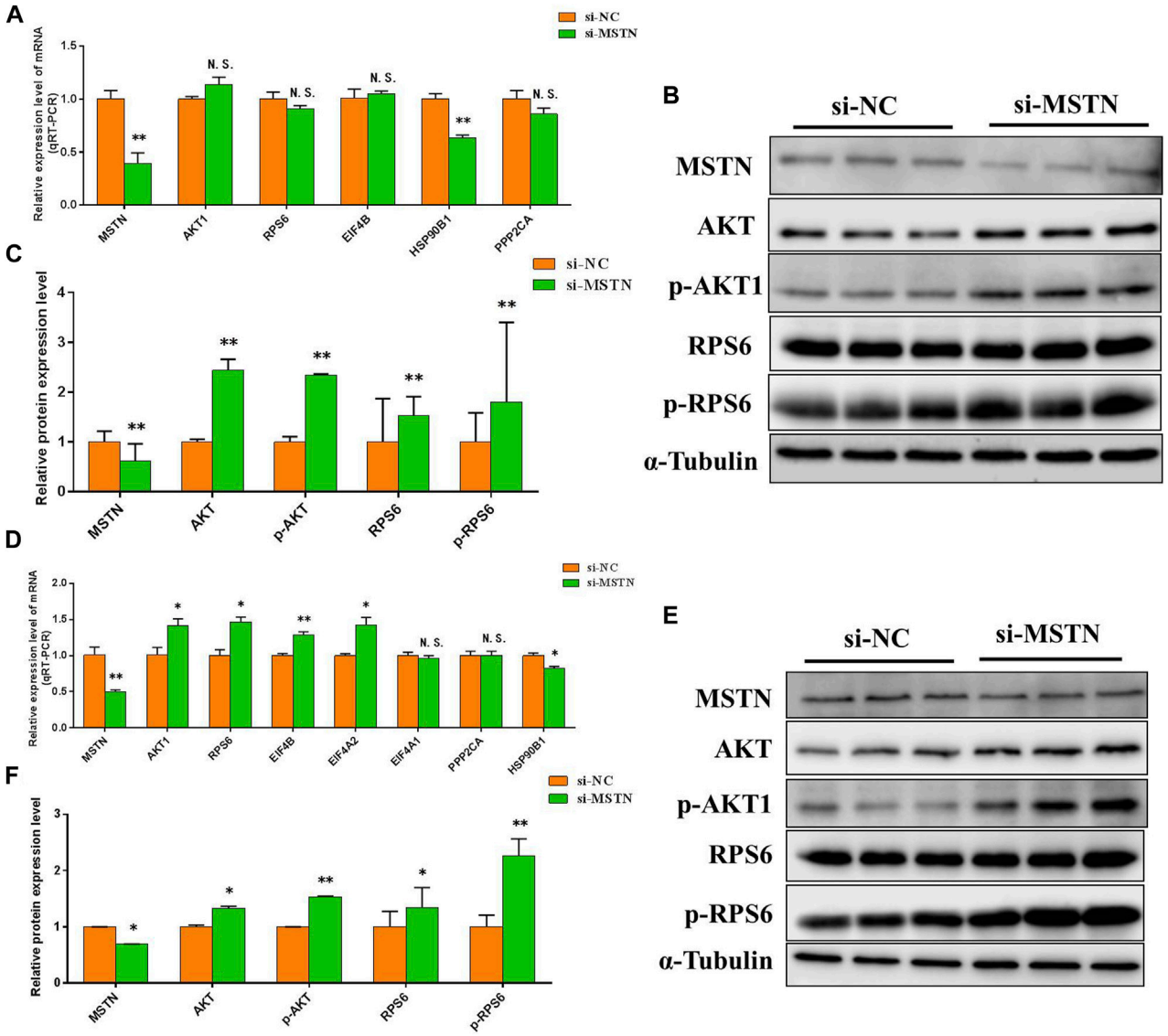


**FIGURE 7 |** The effect of MSTN on the PI3K-AKT and ribosomal pathway. **(A–C)** Knockdown of MSTN expression resulted in a significant increase in mRNA expression level and protein expression level of majority differentially expressed proteins involved in PI3K-AKT and ribosomal pathway in GM, as well as a significant increase in the expression of key genes AKT1, RPS6, p-AKT1and p-RPS6 in the pathway. **(D–F)** The mRNA expression level and protein expression level of these proteins were also generally increased in DM3. These blots were cropped and original images were shown in **Supplementary Figure S7** and **Figure 8**.

and the PI3K-AKT pathway. Focal adhesions are structural links between the extracellular matrix and actin cytoskeleton, which play critical roles in normal physiological events such as cellular adhesion, movement, cytoskeletal structure, and intracellular signaling pathways (Yam et al., 2009). Tyrosine 397 phosphorylation of FAK, a key gene in the adhesion pathway, is essential for normal myoblast differentiation (de Oliveira et al., 2009) and skeletal muscle hypertrophy (Gordon et al., 2001) after hindlimb suspension in rodents, and leads to the binding of FAK to the SH2 domain of the 85 kDa subunit of PI3K, which may lead to the increase of PI3K activity (Appeddu, 1996). The PI3K/AKT signaling pathway is an essential node in mammalian cells that controls cell growth, migration, proliferation, and metabolism (Pompura and Dominguez-Villar, 2018). Previous studies have shown that knockout of the expression of MSTN activates the PI3K-AKT pathway, which can promote muscle development in pigs (Li et al., 2020), which is consistent with the results of this study.

Muscle hypertrophy occurs when the rate of protein synthesis exceeds the rate of degradation. The main factor determining the rate of protein synthesis is ribosome abundance or translational capacity (Wen et al., 2016). In the last couple of decades, studies have revealed that the ribosome has an essential role in the regulation of cell proliferation and growth (Volarevic, 2000; Kirn-Safran et al., 2007) and homeostasis in mammalian organisms. For example, deletions and mutations in genes

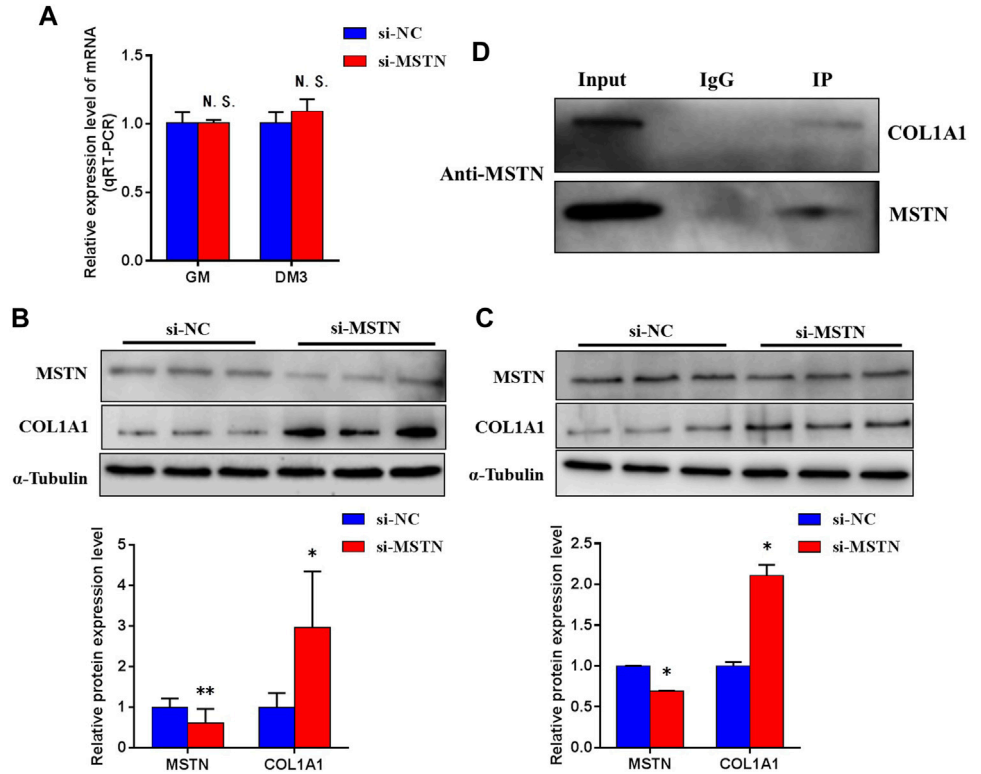


**FIGURE 8 |** The physical interaction between MSTN and COL1A1. **(A)** Knockdown of MSTN expression in BSMSCs resulted in no significant change in COL1A1 mRNA expression, **(B–C)** but protein expression significantly increased in GM and DM3. **(D)** Co-IP against the endogenous MSTN using anti-MSTN antibody was conducted with the lysate from BSMSCs and the eluent was immunoblotted against anti-COL1A1 antibody. A normal rabbit IgG was used as a negative control for immunoprecipitation. These blots were cropped and original images were shown in **Supplementary Figure S9**.

linked to ribosome biogenesis result in pathologies known as ribosomopathies, which are associated with malformation and growth retardation (Teng et al., 2013). A known role of ribosomal protein RPS6 phosphorylation is the regulation of ribosomal protein synthesis (Jefferies et al., 1994).

The expression levels of MYL6 in muscle tissue were detected by western blot to verify the accuracy of proteomics data. Our Western blotting with anti-phospho-FAK (Tyr-576), Akt (Ser-473), RPS6(Ser-235/236) confirmed the dramatically increased expression of pFAK, pAkt, pRPS6 in MG.MSTN$^{+/-}$ (**Figure 3**), thus activating the focal adhesion, PI3K-AKT, and ribosomal signaling pathways. Satellite cells are tissue-specific stem cells responsible for skeletal muscle growth and regeneration (Kanisicak et al., 2009). We constructed an MSTN siRNA model by using BSMSCs and found that knockdown of the expression of MSTN activated focal adhesion, PI3K-AKT, and ribosomal signal pathways, and promoted cell proliferation and myogenic differentiation. The interaction between MSTN and COL1A1 is verified by CO-IP technology. At the same time, the COL1A1 siRNA model was constructed, and it was found that knockdown of the expression of COL1A1 inhibited the adhesion plaque, PI3K-AKT, and ribosomal pathway, and inhibited cell proliferation.

Our study using TMT-based quantitative proteomics analyses on skeletal muscle from MG.MSTN$^{+/-}$ and MG.WT, and verified the results of proteomics at the cellular level for the first time. The quantitative proteomics data reveal that MSTN induces abundance changes of extracellular matrix and ribosome-related proteins. Using muscle tissue and cell model verification, we found that MSTN regulates focal adhesion, PI3K-AKT, and ribosomal pathways through its interaction with COL1A1 (**Figure 11**). By the combination of the results reported in this work and previous studies we conclude that MSTN plays a critical role in muscle development, and knock down the expression of MSTN can activate focal adhesion, PI3K-AKT, and ribosomal pathways, and elevated actin polymerization and protein synthesis.

## MATERIALS AND METHODS

### Experimental Design and Sample Collection

In this experiment, 3 MSTN$^{+/-}$ Mongolian cattle and 3 WT Mongolian cattle were selected as the research objects in the animal experimental base of Inner Mongolia University. The experimental cattle were 2-year-old female cattle, healthy and reared in the same environment. The muscle tissue of bovine leg and gluteal muscle was collected by *in vivo* sampling technique, and the individual information was recorded. Quantitative proteomics analysis based on TMT 6-plex were designed and

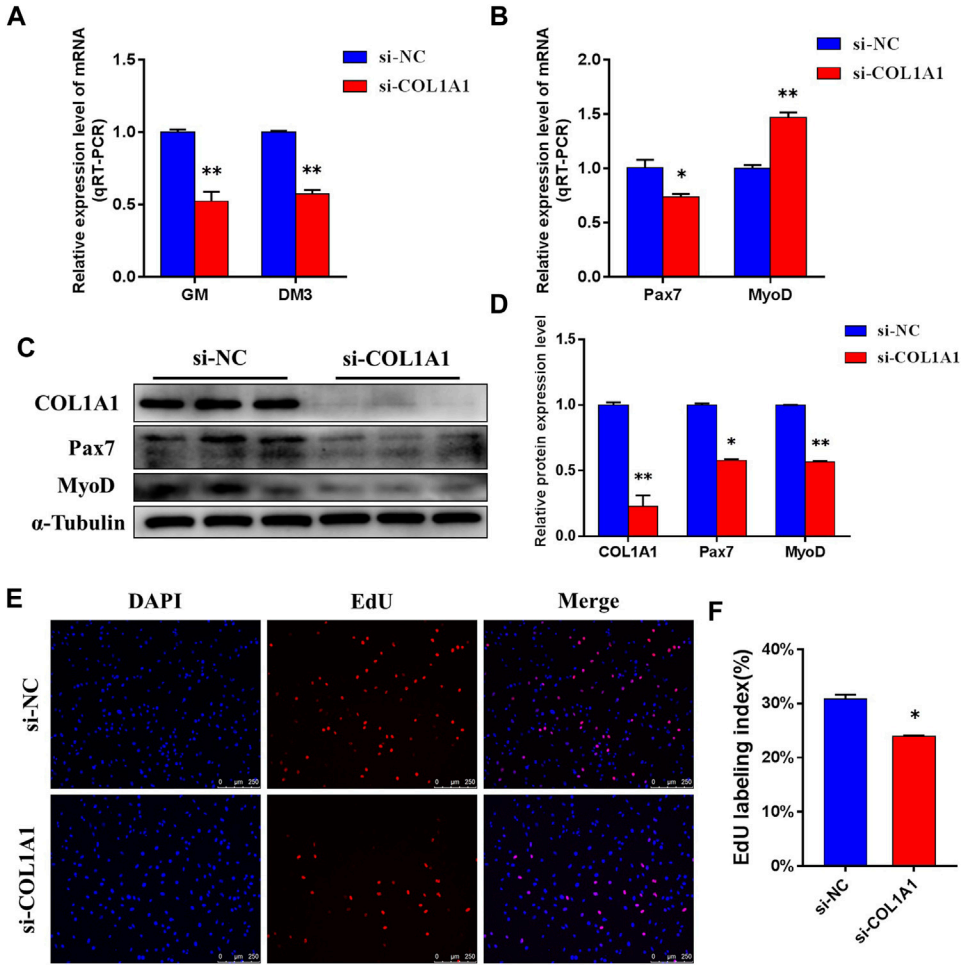


**FIGURE 9 |** COL1A1 maintains the proliferation of BSMSCs. **(A)** Knockdown COL1A1 significantly decreased COL1A1 expression level in GM and DM3, **(B)** resulting in a significant decrease in the expression level of mRNA in Pax7 and a significant increase in the expression level of mRNA in MyoD, **(C–D)** but the protein expression levels of Pax7 and MyoD were significantly decreased. **(E–F)** EdU assay was performed at 24 h after transfection. The number of EdU-positive cells (E, ×200; scale bars 250 μm) and EdU labeling index (F) were decreased with COL1A1 knockdown. These blots were cropped and original images were shown in **Supplementary Figure S10**.

carried out to study the regulation mechanism of MSTN in muscle development. A schematic diagram for experimental design and workflow is shown in **Figure 12**. All the experiments were conducted in strict accordance with the recommendations in the guidelines for Animal Protection and Utilization of Inner Mongolia University and approved by the Animal Welfare Committee of Inner Mongolia University.

## Protein Extraction and Digestion

Crush gluteal muscle tissue by conventional grinding method, and add RIPA lysis buffer (1% NP-40, 0.5% sodium deoxycholate, 1% SDS solution, 5 ml 1×PBS, 42 ml milli-Q water) containing protease inhibitor and phosphatase inhibitor (RIPA: protease inhibitor: phosphatase inhibitor:100:1:1) to mix well. The homogenized samples were centrifuged at 4°C for 20 min at 15,000 g after incubated for 30 min at room temperature. The supernatant was collected to obtain protein samples, and then the protein concentration was determined by BCA (CWbiotech Ltd., Beijing, China). The proteins were lyophilized and stored at −80°C for further analysis.

The lyophilized proteins were resuspended and denatured for 5 min sonication and 0.5 h of the vortex in 7 M urea, 2 M thiourea with a final concentration of 100 mM phosphate buffer (pH 7.8) containing 0.5 tablets PhosSTOP (phosphatase Inhibitor Cocktail Tablets from Roche) for 5 ml buffer. The protein concentration for each of the 6 samples was quantified by a gel-based analysis as described previously (Chen et al., 2013). Each protein solution was reduced in 10 mM dithiothreitol (DTT) for 1 h at 60°C before being alkylated in 40 mM freshly made iodoacetamide for 10min at room temperature. The reduced and alkylated protein solutions were transferred to 10 K ultrafiltration tubes before being centrifuged at 12,000 g for 20 min. After discarding the liquid in the button of the tube, the supernatant was transferred to a new ultrafiltration tube and subsequently mixed with trypsin (AB Sciex) (1:50 w/w) for overnight protein digestion at 37°C. After digestion, the samples were centrifuged at 12,000 g for

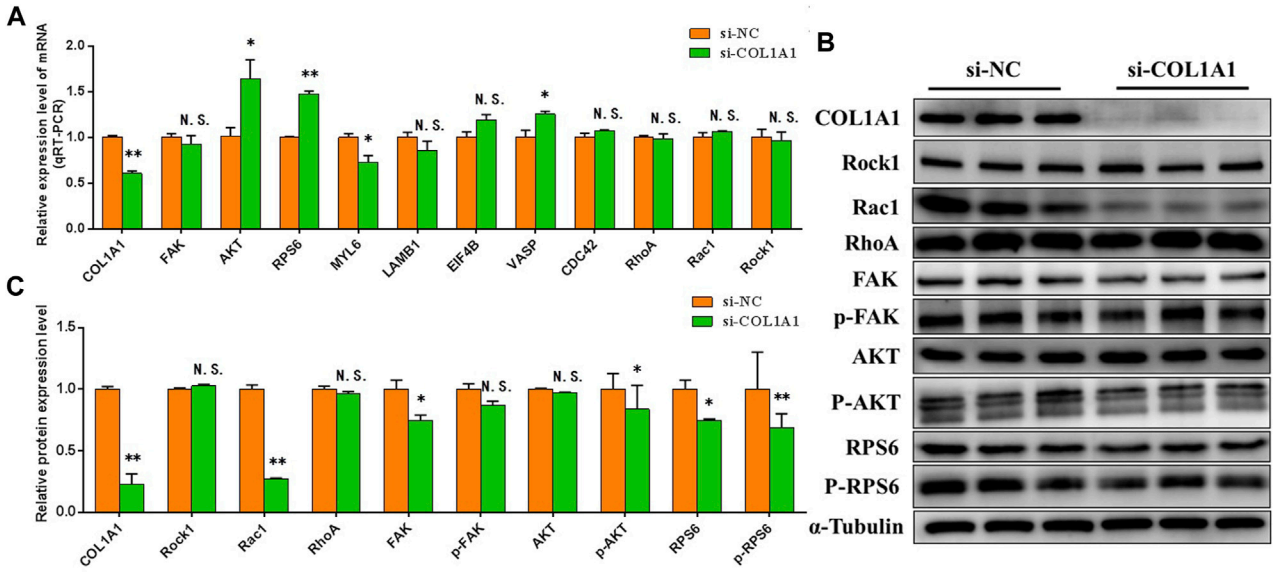


**FIGURE 10 |** COL1A1 functions in the proliferation of BSMSCs. **(A)** knockdown of COL1A1 expression resulted in focal adhesion, PI3K-AKT and ribosomal pathway-related differentially expressed protein mRNA expression level did not change significantly, **(B–C)** but the protein expression level decreased significantly. These blots were cropped and original images were shown in **Supplementary Figure S11**.

20 min, the sediment (peptides) was dried by vacuum centrifugation.

## Tandem Mass Tag 6-Plex Labeling and High pH Reverse Phase LC Fractionation

The 6-plex tandem mass tag (TMT) labeling was carried out according to Thermo Scientific's TMT Mass Tagging Kits protocol. Each of the TMT 6-plex label reagents was reconstituted in 45 μL of acetonitrile (ACN), and the digested peptides from each sample were incubated with a specific tag (tags 128N, 129N, 129C used for 3 MG.MSTN$^{+/-}$ samples and tags 130N, 130C, 131 used for 3 MG.WT samples, respectively) for 1 h in room temperature (**Figure 12**). And then the samples were pooled and dry-out to cation exchange chromatography. The hpRP chromatography was carried out as described previously (Yang et al., 2011). Forty-eight fractions were obtained at 1 min intervals and pooled into 6 fractions based on UV absorbance at 214 nm and with a multiple fraction concatenation strategy (Wang et al., 2011). The trypsin peptides labeled in each component are used for global proteomic analysis of nanoLC-MS/MS analysis.

## Nano-Scale Reverse Phase Chromatography and Tandem MS Analysis

NanoLC-MS/MS analysis was carried out using an Orbitrap Fusion mass spectrometer (Thermo-Fisher Scientific, San Jose, CA) with an UltiMate3000RSLCnano system (Thermo-Dionex, Sunnyvale, CA). 10 μL supernatant were loaded onto a reverse-phase trap column (Thermo Scientific Acclaim PepMap100, 75 μm*2 cm, nanoViper C18) which was connected to a C18 reversed-phase analytical column (Thermo Scientific Easy Column, 10 cm long, 75 μm inner diameter). The samples were loaded at 4 μL/min for 5 min, then run at 400 ml/min from 5 to 30% for 40 min in mobile phase B (98% acetonitrile, 0.1% formic acid), followed by 8 min linear gradient to 80%, then keep at 80% for 8 min in mobile phase B, and finally back to 5% in 2 min. MS data were acquired using a data-dependent top 20 method dynamically choosing the most abundant precursor ions from the survey scan (350–2000 m/z) for Higher Collision Dissociation (HCD) fragmentation. The automatic gain control (AGC) target was set to 3e6. The resolution was set to 70,000 for Survey scans and 17,500 for HCD spectra. Isolation width was 2 m/z. The normalized collision energy was 27 eV and the underfill ratio was 1%.

## Protein Identification and Data Analysis

The NanoLC-MS/MS data were searched against the Bos Taurus database from NCBI using Sequest HT software and converted into MGF files using the Proteome Discoverer 1.4 (PD1.4, Thermo) as previously described (Xin et al., 2017). The parameters of default search for quantitative processing and protein identification were also the same as our previous reports. For the assignment of proteins, at least two distinct peptides should be identified for one protein. The final ratios were subjected to statistical analysis with Perseus software (http://www.coxdocs.org/doku.php), and significance was assessed with t-tests.

## Bioinformatics Analysis

UniProt website (http://www.uniprot.org/) was used to find protein function. DAVID Bioinformatics DAVID Bioinformatics Resources 6.8 (https://david.ncifcrf.gov/) was

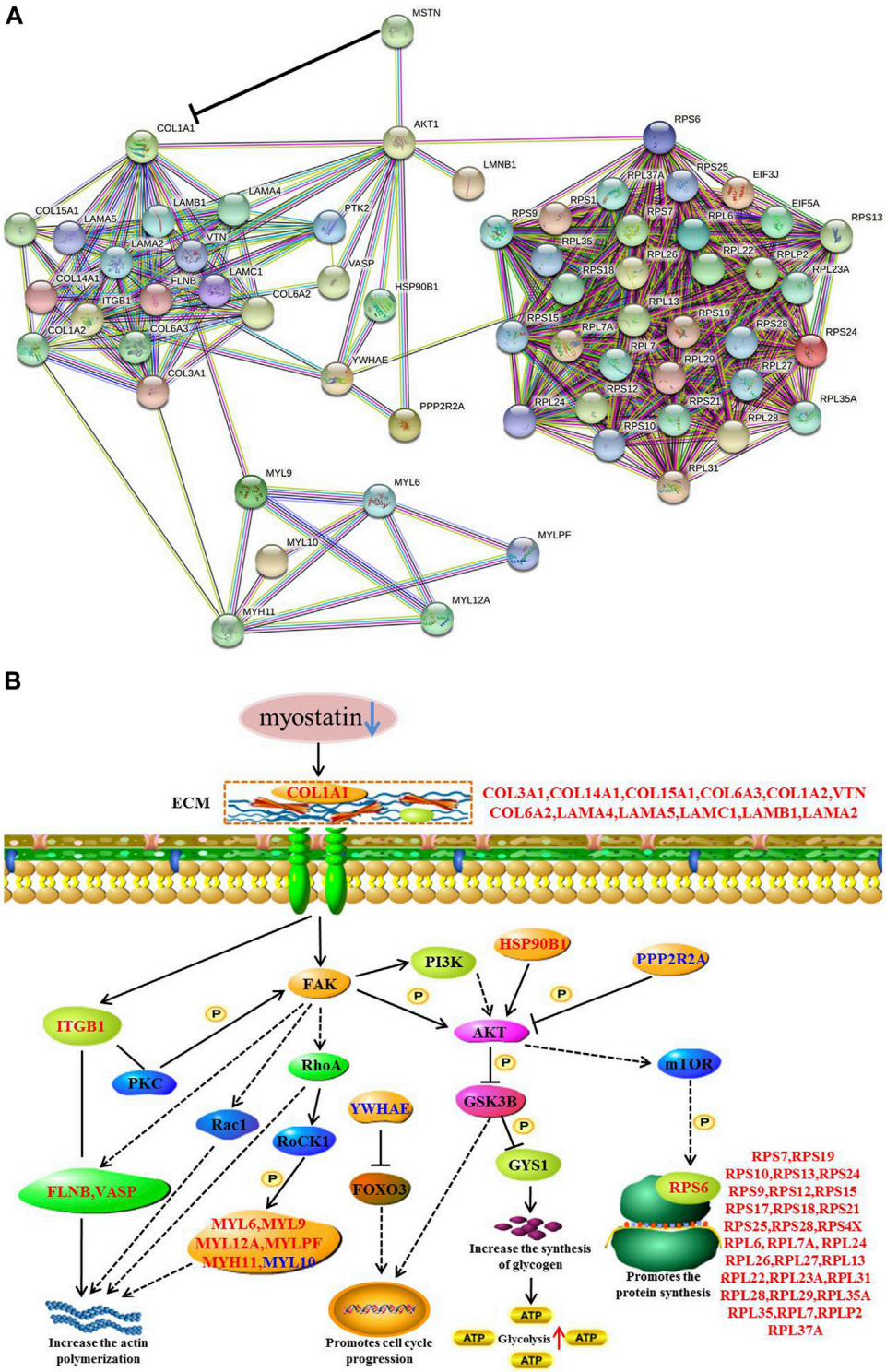


**FIGURE 11 |** Functional interaction network and a diagrammatic mode illustrate the regulatory mechanism of the MSTN gene on the focal adhesion, PI3K-AKT, and ribosomal pathways. **(A)** MSTN, FAK, AKT1, RPS6, and 55 Differentially expressed proteins were submitted to conduct blast searching against the existing databases in STRING 11 software. **(B)** "→" The activation of the process, " -| " the inhibition of the process, "– –" the presence of intermediate steps either unknown or omitted. Up-regulated differentially expressed proteins were marked with red font, while down-regulated differentially expressed proteins were marked with blue font.

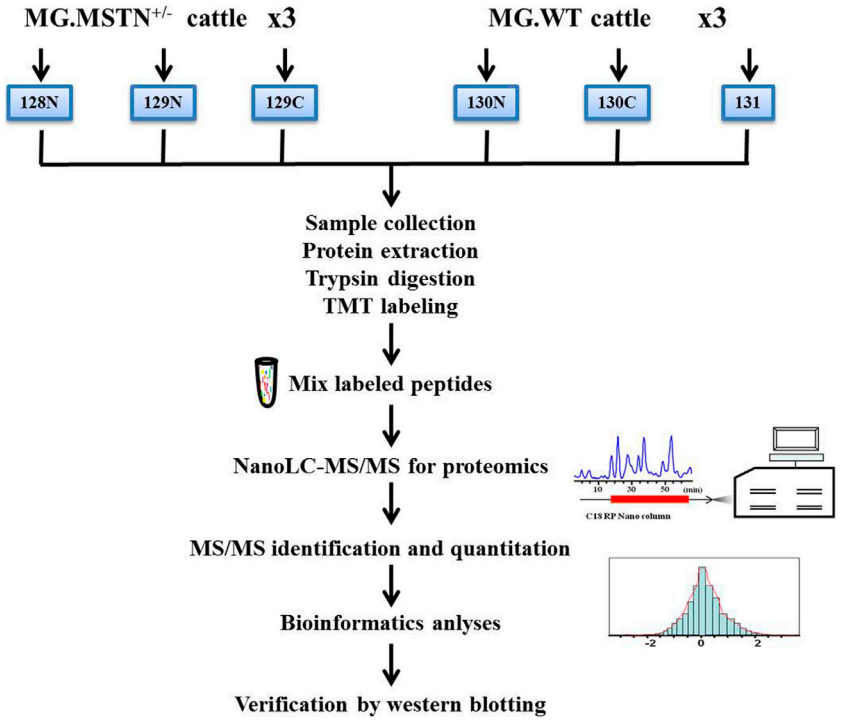


**FIGURE 12 |** Experimental design and schematic diagram of the proteomics workflow. A total of 3 MSTN$^{+/-}$ Mongolian cattle (MG.MSTN$^{+/-}$) and 3 wild type Mongolian cattle (MG.WT) were analyzed by TMT 6-plex based shotgun-based quantitative proteomics, using the high pH reverse phase label-free liquid chromatography-mass spectrometry (hpRP-nanoLC-MS/MS) workflow, and the discovery results by multiple methods including real-time fluorescence quantitative PCR (qRT-PCR), co-Immunoprecipitation (CO-IP), and western blotting approaches.

used to perform the GO annotation and KEGG pathway analysis of differentially expressed proteins. All interactions and construct networks were performed in the STRING 11.0 database (http://string-db.org/).

## Cell Isolation and Culture

Primary BSMSCs were isolated, cultured, and differentiated according to the previous study (Dai et al., 2016). To put this point simply, the muscle tissue of the hindlimb was isolated from 5-6-month-old fetal cattle and digested with type II collagenase (Gibco, Waltham, MA) and trypsin (Gibco). The isolated cells were cultured in a growth medium containing 80% Dulbecco's modified Eagle's medium (DMEM, Gibco) and 20% fetal bovine serum (FBS, Gibco) until the cell density reached about 70%, the differentiation medium containing 98% DMEM and 2% horse serum (HS, Gibco) was used to induce differentiation.

## SiRNA Synthesis

SiRNA for specific knockdown of MSTN and COL1A1 expression was designed and synthesized by RiboBio (Guangzhou, China). The sequences of si-MSTN and si-COL1A1 are shown in **Supplementary Table S1**.

## Cell Transfection

According to the manufacturer's protocol, when the cell density reaches 70–90%, Lipofectamine 3000 (Invitrogen, Carlsbad, CA) is used to transfect siRNA into the original BSMSCs. The final concentrations used for siRNA were 100 nM. After 24 h of transfection, the (GM) of proliferative cells was collected. The differentiation medium was changed and cultured for 72 h, and the cells on the third day of differentiation (DM3) were collected.

## RNA Isolation and Quantitative Real-Time PCR

The total RNA was isolated from BSMSCs using Trizol reagent (Invitrogen). The integrity and concentration of RNA samples were measured by 2.0% agarose gel electrophoresis and NanoDrop 2000c. Then, the PrimeScript II 1st Strand cDNA Synthesis Kit (Takara, Dalian, China) was employed to prepare the first-strand cDNA. Quantitative real-time PCR (qRT-PCR) was performed using All-in-One™ qRT-PCR Mix (Genocopoeia, Guangzhou, China) in a LightCycler® 96 Instrument (Roche, Germany) to detect the expression level of mRNAs. The gene relative expression level was calculated by the $2^{-\Delta\Delta Ct}$ method with TUBB mRNA as an endogenous control of the basal level. All primers used were listed in **Supplementary Table S2**.

## 5-Ethynyl-2′-deoxyuridine Assay

After 24 h of transfection, BSMSCs were incubated at 37°C for 2 h in 96-well plates with 50 μM EdU (RiboBio, Guangzhou, China). Then, the cells were fixed with 4% paraformaldehyde for 30 min and

neutralized with 2 mg/ml glycine solution. The Apollo® staining solution which contains EdU was added and incubated at room temperature for 30 min in the dark to label the DNA in the synthesis stage, and the nucleus was then counterstained with DAPI solution. Three images were randomly obtained by an inverted fluorescence microscope (Leica, Germany) at a magnification of ×200, and the number of EdU positive cells was calculated.

### Co-Immunoprecipitation

Co-immunoprecipitation was conducted using a Pierce™ Crosslink Magnetic IP/Co-IP Kit (Thermo-Fisher Scientific, San Jose, CA). The cells were lysed with a buffer containing 25 mM Tris, 150 mM NaCl, 1 mM EDTA, 1% NP40, 5% glycerol, and a protease inhibitor cocktail. The lysates were immunoprecipitated with either anti-MSTN antibody or anti-IgG antibody, and the eluents were loaded on a 10% SDS–PAGE gel. The blots were incubated overnight at 4°C with either anti-MSTN antibody or anti-COL1A1 antibody. After washing in TBS containing 0.05% Tween 20, then incubated with secondary antibody for 1 h before detection using ECL chemiluminescent substrate (Solarbio).

### Western Blot Analysis

GM phase and DM3 phase cells were collected and lysed with RIPA buffer with protease inhibitor and phosphatase inhibitor, and the total proteins were identified. The total protein concentration of the extract was determined by BCA method. Then, equal amounts of cells lysate were resolved by 10% SDS-PAGE and transferred onto polyvinylidene difluoride membranes (Millipore, Burlington, MA). The membranes were blocked with 5% BSA for 1 h, incubated with primary antibody at 4°C overnight, then incubated with secondary antibody for 1 h prior to detection using ECL chemiluminescent substrate (Solarbio). Several proteins were validated by western blot analysis using antibodies against the following proteins: Pax7 (1:100, DSHB, United States), MyoG (1:100, DSHB, United States), MyoD (1:100, DSHB, United States), MyHC (1:100, DSHB, United States), MSTN (1:100, Santa Cruz, United States), p-AKT1 (Ser-473) (1:200, Santa Cruz, Dallas, United States), COL1A1 (1:5,000, Abcam, United States), FAK (1:1,000, Abcam, United States), Rock1 (1: 5,000, Abcam, United States), AKT1 (1:5,000, Abcam, United States), RhoA (1:1,000, NewEast, China), Rac1 (1:1,000, NewEast, China), LAMB1 (1:1,500, Sangon Biotech, China), MYL6 (1:500, Sangon Biotech, China), ACTN4 (1:1,000, Sangon Biotech, China), RPS6 (1:500, Sangon Biotech, China), p-FAK (Tyr-473) (1:500, Sangon Biotech, China), p-RPS6 (Ser-235/236) (1:500, Sangon Biotech, China), and reference gene Alpha-Tubulin (1:5,000, Abcam, United States). The gray values of each protein band were calculated using ImageJ software.

## 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 Ethical and Welfare Committee of Inner Mongolia University.

## AUTHOR CONTRIBUTIONS

HS, GL, and HG contributed to conception and design of the study. HS, MM, and HT organized the database. HS performed the statistical analysis. HS wrote the first draft of the manuscript. YG, LZ, JZ, XL, DH and XD wrote sections of the manuscript. All authors contributed to manuscript revision, read, and approved the submitted version.

## FUNDING

This work was supported by the National Transgenic Animal Program (grant numbers: 2016ZX08007-002) and the Natural Science Foundation of Tianjin (grant numbers: 18JCYBJC29700).

## ACKNOWLEDGMENTS

We thanks to Research Center for Laboratory Animal Science, Inner Mongolia University for providing experimental materials and the proteomic data provided by Cornell University. We are grateful for the help from the Tianjin Key Laboratory of Agricultural Animal Breeding and Healthy Husbandry of Tianjin Agricultural University.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2021.752129/full#supplementary-material

## REFERENCES

Allendorph, G. P., Vale, W. W., and Choe, S. (2006). Structure of the Ternary Signaling Complex of a TGF-Beta Superfamily Member. *Proc. Natl. Acad. Sci.* 103 (20), 7643–7648. doi:10.1073/pnas.0602558103

Aversa, Z., Bonetto, A., Penna, F., Costelli, P., Di Rienzo, G., Lacitignola, A., et al. (2012). Changes in Myostatin Signaling in Non-weight-losing Cancer Patients. *Ann. Surg. Oncol.* 19 (4), 1350–1356. doi:10.1245/s10434-011-1720-5

Badylak, S. F., Kochupura, P. V., Cohen, I. S., Doronin, S. V., Saltman, A. E., Gilbert, T. W., et al. (2006). The Use of Extracellular Matrix as an Inductive Scaffold for the Partial Replacement of Functional Myocardium. *Cel Transpl.* 15, 29–40. doi:10.3727/000000006783982368

Badylak, S. F., Vorp, D. A., Spievack, A. R., Simmons-Byrd, A., Hanke, J., Freytes, D. O., et al. (2005). Esophageal Reconstruction with ECM and Muscle Tissue in a Dog Model. *J. Surg. Res.* 128 (1), 87–97. doi:10.1016/j.jss.2005.03.002

Birch, H. L. (2018). Extracellular Matrix and Ageing. *Subcell Biochem.* 90, 169–190. doi:10.1007/978-981-13-2835-0_7

Boruch, A. V., Nieponice, A., Qureshi, I. R., Gilbert, T. W., and Badylak, S. F. (2010). Constructive Remodeling of Biologic Scaffolds Is Dependent on Early Exposure to Physiologic Bladder Filling in a Canine Partial Cystectomy Model. *J. Surg. Res.* 161 (2), 217–225. doi:10.1016/j.jss.2009.02.014

Chen, H.-C., Appeddu, P. A., Isoda, H., and Guan, J.-L. (1996). Phosphorylation of Tyrosine 397 in Focal Adhesion Kinase Is Required for Binding Phosphatidylinositol 3-Kinase. *J. Biol. Chem.* 271 (42), 26329–26334. doi:10.1074/jbc.271.42.26329

Chen, J.-W., Scaria, J., Mao, C., Sobral, B., Zhang, S., Lawley, T., et al. (2013). Proteomic Comparison of Historic and Recently Emerged Hypervirulent Clostridium difficile Strains. *J. Proteome Res.* 12 (3), 1151–1161. doi:10.1021/pr3007528

Chen, L., Li, X., Ni, N., Liu, Y., Chen, L., Wang, Z., et al. (2016). Phosphorylation of Myofibrillar Proteins in post-mortem Ovine Muscle with Different Tenderness. *J. Sci. Food Agric.* 96 (5), 1474–1483. doi:10.1002/jsfa.7244

Chen, L., Li, Z., Everaert, N., Lametsch, R., and Zhang, D. (2019). Quantitative Phosphoproteomic Analysis of Ovine Muscle with Different Postmortem Glycolytic Rates. *Food Chem.* 280, 203–209. doi:10.1016/j.foodchem.2018.12.056

Clause, K. C., and Barker, T. H. (2013). Extracellular Matrix Signaling in Morphogenesis and Repair. *Curr. Opin. Biotechnol.* 24 (5), 830–833. doi:10.1016/j.copbio.2013.04.011

Clop, A., Marcq, F., Takeda, H., Pirottin, D., Tordoir, X., Bibé, B., et al. (2006). A Mutation Creating a Potential Illegitimate microRNA Target Site in the Myostatin Gene Affects Muscularity in Sheep. *Nat. Genet.* 38 (7), 813–818. doi:10.1038/ng1810

Dai, Y., Wang, Y. M., Zhang, W. R., Liu, X. F., Li, X., Ding, X. B., et al. (2016). The Role of microRNA-1 and microRNA-206 in the Proliferation and Differentiation of Bovine Skeletal Muscle Satellite Cells. In Vitro *Cell.Dev.Biol.-Animal* 52 (1), 27–34. doi:10.1007/s11626-015-9953-4

de Oliveira, M. V., Marin, T. M., Clemente, C. F., Costa, A. P. D., Judice, C. C., and Franchini, K. G. (2009). SHP-2 Regulates Myogenesis by Coupling to FAK Signaling Pathway. *Febs Lett.* 583 (18), 2975–2981. doi:10.1016/j.febslet.2009.08.022

Desai, K. M., Diaz, S., Dorward, I. G., Winslow, E. R., La Regina, M. C., Halpin, V., et al. (2006). Histologic Results 1 Year after Bioprosthetic Repair of Paraesophageal Hernia in a Canine Model. *Surg. Endosc.* 20 (11), 1693–1697. doi:10.1007/s00464-006-0680-5

Durieux, A.-C., Amirouche, A., Banzet, S., Koulmann, N., Bonnefoy, R., Pasdeloup, M., et al. (2007). Ectopic Expression of Myostatin Induces Atrophy of Adult Skeletal Muscle by Decreasing Muscle Gene Expression. *Endocrinology* 148 (7), 3140–3147. doi:10.1210/en.2006-1500

Elkasrawy, M. N., and Hamrick, M. W. (2010). Myostatin (GDF-8) as a Key Factor Linking Muscle Mass and Bone Structure. *J. Musculoskelet. Neuronal Interact* 10 (1), 56–63. doi:10.1007/s00232-010-9238-5

Elkina, Y., von Haehling, S., Anker, S. D., and Springer, J. (2011). The Role of Myostatin in Muscle Wasting: an Overview. *J. Cachexia Sarcopenia Muscle* 2 (3), 143–151. doi:10.1007/s13539-011-0035-5

Franklin, M. E., Treviño, J. M., Portillo, G., Vela, I., Glass, J. L., and González, J. J. (2008). The Use of Porcine Small Intestinal Submucosa as a Prosthetic Material for Laparoscopic Hernia Repair in Infected and Potentially Contaminated fields: Long-Term Follow-Up. *Surg. Endosc.* 22 (9), 1941–1946. doi:10.1007/s00464-008-0005-y

Frantz, C., Stewart, K. M., and Weaver, V. M. (2010). The Extracellular Matrix at a Glance. *J. Cel Sci.* 123 (24), 4195–4200. doi:10.1242/jcs.023820

Gonzalez-Cadavid, N. F., Taylor, W. E., Yarasheski, K., Sinha-Hikim, I., Ma, K., Ezzat, S., et al. (1998). Organization of the Human Myostatin Gene and Expression in Healthy Men and HIV-Infected Men with Muscle Wasting. *Proc. Natl. Acad. Sci.* 95 (25), 14938–14943. doi:10.1073/pnas.95.25.14938

Gordon, S. E., Flück, M., and Booth, F. W. (2001). Selected Contribution: Skeletal Muscle Focal Adhesion Kinase, Paxillin, and Serum Response Factor Are Loading Dependent. *J. Appl. Physiol.* 90 (3), 1174–1183. doi:10.1.1.337.90110.1152/jappl.2001.90.3.1174

Grobet, L., Royo Martin, L. J., Poncelet, D., Pirottin, D., Brouwers, B., Riquet, J., et al. (1997). A Deletion in the Bovine Myostatin Gene Causes the Double-Muscled Phenotype in Cattle. *Nat. Genet.* 17 (1), 71–74. doi:10.1038/ng0997-71

Gumbiner, B. M. (1996). Cell Adhesion: the Molecular Basis of Tissue Architecture and Morphogenesis. *Cell* 84 (3), 345–357. doi:10.1016/s0092-8674(00)81279-9

Hayot, M., Rodriguez, J., Vernus, B., Carnac, G., Jean, E., Allen, D., et al. (2011). Myostatin Up-Regulation Is Associated with the Skeletal Muscle Response to Hypoxic Stimuli. *Mol. Cell Endocrinol.* 332 (1-2), 38–47. doi:10.1016/j.mce.2010.09.008

Huang, H., Larsen, M. R., Karlsson, A. H., Pomponio, L., Costa, L. N., and Lametsch, R. (2011). Gel-based Phosphoproteomics Analysis of Sarcoplasmic Proteins in Postmortem Porcine Muscle with pH Decline Rate and Time Differences. *Proteomics* 11 (20), 4063–4076. doi:10.1002/pmic.201100173

Huang, H., Larsen, M. R., and Lametsch, R. (2012). Changes in Phosphorylation of Myofibrillar Proteins during Postmortem Development of Porcine Muscle. *Food Chem.* 134 (4), 1999–2006. doi:10.1016/j.foodchem.2012.03.132

Huang, H., Larsen, M. R., Palmisano, G., Dai, J., and Lametsch, R. (2014). Quantitative Phosphoproteomic Analysis of Porcine Muscle within 24h Postmortem. *J. Proteomics* 106, 125–139. doi:10.1016/j.jprot.2014.04.020

Huang, P., Pang, D., Wang, K., Xu, A., Yao, C., Li, M., et al. (2019). The Possible Role of Complete Loss of Myostatin in Limiting Excessive Proliferation of Muscle Cells (C2C12) via Activation of MicroRNAs. *Ijms* 20 (3), 643. doi:10.3390/ijms20030643

Huang, Z., Chen, D., Zhang, K., Yu, B., Chen, X., and Meng, J. (2007). Regulation of Myostatin Signaling by C-Jun N-Terminal Kinase in C2C12 Cells. *Cell Signal.* 19 (11), 2286–2295. doi:10.1016/j.cellsig.2007.07.002

Jefferies, H. B., Reinhard, C., Kozma, S. C., and Thomas, G. (1994). Rapamycin Selectively Represses Translation of the "polypyrimidine Tract" mRNA Family. *Proc. Natl. Acad. Sci.* 91 (10), 4441–4445. doi:10.1073/pnas.91.10.4441

Jouliaekaza, D., and Cabello, G. (2007). The Myostatin Gene: Physiology and Pharmacological Relevance. *Curr. Opin. Pharmacol.* 7 (3), 310–315. doi:10.1016/j.coph.2006.11.011

Kanisicak, O., Mendez, J. J., Yamamoto, S., Yamamoto, M., and Goldhamer, D. J. (2009). Progenitors of Skeletal Muscle Satellite Cells Express the Muscle Determination Gene, MyoD. *Dev. Biol.* 332 (1), 131–141. doi:10.1016/j.ydbio.2009.05.554

Karamanos, N. K., Piperigkou, Z., Theocharis, A. D., Watanabe, H., Franchi, M., Baud, S., et al. (2018). Proteoglycan Chemical Diversity Drives Multifunctional Cell Regulation and Therapeutics. *Chem. Rev.* 118 (18), 9152–9232. doi:10.1021/acs.chemrev.8b00354

Kirn-Safran, C. B., Oristian, D. S., Focht, R. J., Parker, S. G., Vivian, J. L., and Carson, D. D. (2007). Global Growth Deficiencies in Mice Lacking the Ribosomal Protein HIP/RPL29. *Dev. Dyn.* 236 (2), 447–460. doi:10.1002/dvdy.21046

Kollias, H. D., and McDermott, J. C. (2008). Transforming Growth Factor-β and Myostatin Signaling in Skeletal Muscle. *J. Appl. Physiol.* 104 (3), 579–587. doi:10.1152/japplphysiol.01091.2007

Langley, B., Thomas, M., Bishop, A., Sharma, M., Gilmour, S., and Kambadur, R. (2002). Myostatin Inhibits Myoblast Differentiation by Down-Regulating MyoD Expression. *J. Biol. Chem.* 277 (51), 49831–49840. doi:10.1074/jbc.M204291200

Lee, S.-J., and McPherron, A. C. (2001). Regulation of Myostatin Activity and Muscle Growth. *Proc. Natl. Acad. Sci.* 98 (16), 9306–9311. doi:10.1073/pnas.151270098

Lenk, K., Schur, R., Linke, A., Erbs, S., Matsumoto, Y., Adams, V., et al. (2009). Impact of Exercise Training on Myostatin Expression in the Myocardium and Skeletal Muscle in a Chronic Heart Failure Model. *Eur. J. Heart Fail.* 11 (4), 342–348. doi:10.1093/eurjhf/hfp020

Li, C. B., Li, J., Zhou, G. H., Lametsch, R., Ertbjerg, P., Brüggemann, D. A., et al. (2012). Electrical Stimulation Affects Metabolic Enzyme Phosphorylation, Protease Activation, and Meat Tenderization in Beef1. *J. Anim. Sci.* 90 (5), 1638–1649. doi:10.2527/jas.2011-4514

Li, C., Zhou, G., Xu, X., Lundström, K., Karlsson, A., and Lametsch, R. (2015a). Phosphoproteome Analysis of Sarcoplasmic and Myofibrillar Proteins in Bovine Longissimus Muscle in Response to Postmortem Electrical Stimulation. *Food Chem.* 175, 197–202. doi:10.1016/j.foodchem.2014.11.139

Li, R., Zeng, W., Ma, M., Wei, Z., Liu, H., Liu, X., et al. (2020). Precise Editing of Myostatin Signal Peptide by CRISPR/Cas9 Increases the Muscle Mass of Liang Guang Small Spotted Pigs. *Transgenic Res.* 29 (1), 149–163. doi:10.1007/s11248-020-00188-w

Li, X., Fang, T., Zong, M., Shi, X., Xu, X., Dai, C., et al. (2015b). Phosphorproteome Changes of Myofibrillar Proteins at Early Post-mortem Time in Relation to

Pork Quality as Affected by Season. *J. Agric. Food Chem.* 63 (47), 10287–10294. doi:10.1021/acs.jafc.5b03997

Lipina, C., Kendall, H., McPherron, A. C., Taylor, P. M., and Hundal, H. S. (2010). Mechanisms Involved in the Enhancement of Mammalian Target of Rapamycin Signalling and Hypertrophy in Skeletal Muscle of Myostatin-Deficient Mice. *FEBS Lett.* 584 (11), 2403–2408. doi:10.1016/j.febslet.2010.04.039

Mariot, V., Joubert, R., Hourdé, C., Féasson, L., Hanna, M., Muntoni, F., et al. (2017). Downregulation of Myostatin Pathway in Neuromuscular Diseases May Explain Challenges of Anti-myostatin Therapeutic Approaches. *Nat. Commun.* 8, 1859. doi:10.1038/s41467-017-01486-4

Martin, L. B. B., Sherwood, R. W., Nicklay, J. J., Yang, Y., Muratore-Schroeder, T. L., Anderson, E. T., et al. (2016). Application of Wide Selected-ion Monitoring Data-independent Acquisition to Identify Tomato Fruit Proteins Regulated by the CUTIN DEFICIENT2 Transcription Factor. *Proteomics* 16 (15-16), 2081–2094. doi:10.1002/pmic.201500450

McPherron, A. C., Lawler, A. M., and Lee, S.-J. (1997). Regulation of Skeletal Muscle Mass in Mice by a New TGF-P Superfamily Member. *Nature* 387 (6628), 83–90. doi:10.1038/387083a0

McPherron, A. C., and Lee, S.-J. (1997). Double Muscling in Cattle Due to Mutations in the Myostatin Gene. *Proc. Natl. Acad. Sci. U S A.* 94 (23), 12457–12461. doi:10.1073/pnas.94.23.12457

McPherron, A. C., and Lee, S.-J. (2002). Suppression of Body Fat Accumulation in Myostatin-Deficient Mice. *J. Clin. Invest.* 109 (5), 595–601. doi:10.1172/jci0213562

Mendias, C. L., Schwartz, A. J., Grekin, J. A., Gumucio, J. P., and Sugg, K. B. (2017). Changes in Muscle Fiber Contractility and Extracellular Matrix Production during Skeletal Muscle Hypertrophy. *J. Appl. Physiol.* 122 (3), 571–579. doi:10.1152/japplphysiol.00719.2016

Morissette, M. R., Cook, S. A., Buranasombati, C., Rosenberg, M. A., and Rosenzweig, A. (2009). Myostatin Inhibits IGF-I-Induced Myotube Hypertrophy through Akt. *Am. J. Physiology-Cell Physiol.* 297 (5), C1124–C1132. doi:10.1152/ajpcell.00043.2009

Mosher, D. S., Quignon, P., Bustamante, C. D., Sutter, N. B., Mellersh, C. S., Parker, H. G., et al. (2007). A Mutation in the Myostatin Gene Increases Muscle Mass and Enhances Racing Performance in Heterozygote Dogs. *Plos Genet.* 3 (5), e79. doi:10.1371/journal.pgen.0030079

Muncie, J. M., and Weaver, V. M. (2018). The Physical and Biochemical Properties of the Extracellular Matrix Regulate Cell Fate. *Curr. Top. Dev. Biol.* 130, 1–37. doi:10.1016/bs.ctdb.2018.02.002

Muroya, S., Ohnishi-Kameyama, M., Oe, M., Nakajima, I., Shibata, M., and Chikuni, K. (2007). Double Phosphorylation of the Myosin Regulatory Light Chain during Rigor Mortis of Bovine Longissimus Muscle. *J. Agric. Food Chem.* 55 (10), 3998–4004. doi:10.1021/jf063200o

Obermann, W. M., Gautel, M., Weber, K., and Furst, D. O. (1997). Molecular Structure of the Sarcomeric M Band: Mapping of Titin and Myosin Binding Domains in Myomesin and the Identification of a Potential Regulatory Phosphorylation Site in Myomesin. *EMBO J.* 16 (2), 211–220. doi:10.1093/emboj/16.2.211

Pompura, S. L., and Dominguez-Villar, M. (2018). The PI3K/AKT Signaling Pathway in Regulatory T-Cell Development, Stability, and Function. *J. Leukoc. Biol.* 103 (6), 1065–1076. doi:10.1002/Jlb.2mir0817-349r

Quarti, A., Nardone, S., Colaneri, M., Santoro, G., and Pozzi, M. (2011). Preliminary Experience in the Use of an Extracellular Matrix to Repair Congenital Heart Diseases. *Interactive Cardiovasc. Thorac. Surg.* 13 (6), 569–572. doi:10.1510/icvts.2011.280016

Reddy, P. P., Barrieras, D. J., Wilson, G., Bagli, D. J., McLorie, G. A., Khoury, A. E., et al. (2000). Regeneration of Functional Bladder Substitutes Using Large Segment Acellular Matrix Allografts in a Porcine Model. *J. Urol.* 164 (3), 936–941. doi:10.1016/S0022-5347(05)67221-710.1097/00005392-200009020-00005

Rios, R., Carneiro, I., Arce, V. M., and Devesa, J. (2002). Myostatin Is an Inhibitor of Myogenic Differentiation. *Am. J. Physiology-Cell Physiol.* 282 (5), C993–C999. doi:10.1152/ajpcell.00372.2001

Rios, R., Carneiro, I., Arce, V. M., and Devesa, J. (2001). Myostatin Regulates Cell Survival during C2C12 Myogenesis. *Biochem. Biophys. Res. Commun.* 280 (2), 561–566. doi:10.1006/bbrc.2000.4159

Sartori, R., Milan, G., Patron, M., Mammucari, C., Blaauw, B., Abraham, R., et al. (2009). Smad2 and 3 Transcription Factors Control Muscle Mass in Adulthood. *Am. J. Physiol. Cel Physiol* 296 (6), C1248–C1257. doi:10.1152/ajpcell.00104.2009

Schuelke, M., Wagner, K. R., Stolz, L. E., Hubner, C., Riebel, T., Komen, W., et al. (2004). Myostatin Mutation Associated with Gross Muscle Hypertrophy in a Child. *N. Engl. J. Med.* 350 (26), 2682–2688. doi:10.1056/NEJMoa040933

Smith, M. J., Paran, T. S., Quinn, F., and Corbally, M. T. (2004). The SIS Extracellular Matrix Scaffold - Preliminary Results of Use in Congenital Diaphragmatic Hernia (CDH) Repair. *Pediatr. Surg. Int.* 20 (11-12), 859–862. doi:10.1007/s00383-004-1298-0

Stantzou, A., Relizani, K., Morales-Gonzalez, S., Gallen, C., Grassin, A., Ferry, A., et al. (2021). Extracellular Matrix Remodelling Is Associated with Muscle Force Increase in Overloaded Mouse Plantaris Muscle. *Neuropathol. Appl. Neurobiol.* 47 (2), 218–235. doi:10.1111/nan.12655

Steelman, C. A., Recknor, J. C., Nettleton, D., and Reecy, J. M. (2006). Transcriptional Profiling of Myostatin-Knockout Mice Implicates Wnt Signaling in Postnatal Skeletal Muscle Growth and Hypertrophy. *FASEB J.* 20 (3), 580–582. doi:10.1096/fj.05-5125fje

Teng, T., Thomas, G., and Mercer, C. A. (2013). Growth Control and Ribosomopathies. *Curr. Opin. Genet. Dev.* 23 (1), 63–71. doi:10.1016/j.gde.2013.02.001

Theocharis, A. D., Skandalis, S. S., Gialeli, C., and Karamanos, N. K. (2016). Extracellular Matrix Structure. *Adv. Drug Deliv. Rev.* 97, 4–27. doi:10.1016/j.addr.2015.11.001

Thomas, M., Langley, B., Berry, C., Sharma, M., Kirk, S., Bass, J., et al. (2000). Myostatin, a Negative Regulator of Muscle Growth, Functions by Inhibiting Myoblast Proliferation. *J. Biol. Chem.* 275 (51), 40235–40243. doi:10.1074/jbc.M004356200

Trendelenburg, A. U., Meyer, A., Rohner, D., Boyle, J., Hatakeyama, S., and Glass, D. J. (2009). Myostatin Reduces Akt/TORC1/p70S6K Signaling, Inhibiting Myoblast Differentiation and Myotube Size. *Am. J. Physiol. Cel Physiol* 296 (6), C1258–C1270. doi:10.1152/ajpcell.00105.2009

Tsuchida, K. (2008). Targeting Myostatin for Therapies against Muscle-Wasting Disorders. *Curr. Opin. Drug Discov. Devel* 11 (4), 487–494.

Turner, N. J., Yates, A. J., Weber, D. J., Qureshi, I. R., Stolz, D. B., Gilbert, T. W., et al. (2010). Xenogeneic Extracellular Matrix as an Inductive Scaffold for Regeneration of a Functioning Musculotendinous Junction. *Tissue Eng. A* 16 (11), 3309–3317. doi:10.1089/ten.tea.2010.0169

Valentin, J. E., Turner, N. J., Gilbert, T. W., and Badylak, S. F. (2010). Functional Skeletal Muscle Formation with a Biologic Scaffold. *Biomaterials* 31 (29), 7475–7484. doi:10.1016/j.biomaterials.2010.06.039

Veilleux, L. N., Lemay, M., Pouliot-Laforte, A., Cheung, M. S., Glorieux, F. H., and Rauch, F. (2014). Muscle Anatomy and Dynamic Muscle Function in Osteogenesis Imperfecta Type I. *J. Clin. Endocrinol. Metab.* 99 (2), E356–E362. doi:10.1210/jc.2013-3209

Veilleux, L. N., Pouliot-Laforte, A., Lemay, M., Cheung, M. S., Glorieux, F. H., and Rauch, F. (2015). The Functional Muscle-Bone Unit in Patients with Osteogenesis Imperfecta Type I. *Bone* 79, 52–57. doi:10.1016/j.bone.2015.05.019

Volarevic, S. (2000). Proliferation, but Not Growth, Blocked by Conditional Deletion of 40S Ribosomal Protein S6. *Science* 288 (5473), 2045–2047. doi:10.1126/science.288.5473.2045

Wang, D. T., Yang, Y. J., Huang, R. H., Zhang, Z. H., and Lin, X. (2015). Myostatin Activates the Ubiquitin-Proteasome and Autophagy-Lysosome Systems Contributing to Muscle Wasting in Chronic Kidney Disease. *Oxid Med. Cel Longev* 2015, 684965. doi:10.1155/2015/684965

Wang, L., Huang, Y., Wang, X., and Chen, Y. (2019). Label-Free LC-MS/MS Proteomics Analyses Reveal Proteomic Changes Accompanying MSTN KO in C2C12 Cells. *Biomed. Res. Int.* 2019, 7052456. doi:10.1155/2019/7052456

Wang, Y. X., Yang, F., Gritsenko, M. A., Wang, Y. C., Clauss, T., Liu, T., et al. (2011). Reversed-phase Chromatography with Multiple Fraction Concatenation Strategy for Proteome Profiling of Human MCF10A Cells. *Proteomics* 11 (10), 2019–2026. doi:10.1002/pmic.201000722

Wen, Y., Alimov, A. P., and McCarthy, J. J. (2016). Ribosome Biogenesis Is Necessary for Skeletal Muscle Hypertrophy. *Exerc. Sport Sci. Rev.* 44 (3), 110–115. doi:10.1249/Jes.0000000000000082

Xin, X. B., Liu, X. F., Li, X., Ding, X. B., Yang, S. P., Jin, C. F., et al. (2017). Comparative Muscle Proteomics/phosphoproteomics Analysis Provides New Insight for the Biosafety Evaluation of Fat-1 Transgenic Cattle. *Transgenic Res.* 26 (5), 625–638. doi:10.1007/s11248-017-0032-3

Xin, X. B., Yang, S. P., Li, X., Liu, X. F., Zhang, L. L., Ding, X. B., et al. (2020). Proteomics Insights into the Effects of MSTN on Muscle Glucose and Lipid Metabolism in Genetically Edited Cattle. *Gen. Comp. Endocrinol.* 291, 113237. doi:10.1016/j.ygcen.2019.113237

Yam, J. W. P., Tse, E. Y. T., and Ng, I. O. L. (2009). Role and Significance of Focal Adhesion Proteins in Hepatocellular Carcinoma. *J. Gastroenterol. Hepatol.* 24 (4), 520–530. doi:10.1111/j.1440-1746.2009.05813.x

Yang, S., Xin, L., Liu, X., Ding, X., and Hong, G. J. O. (2018). Parallel Comparative Proteomics and Phosphoproteomics Reveal that Cattle Myostatin Regulates Phosphorylation of Key Enzymes in Glycogen Metabolism and Glycolysis Pathway. *Oncotarget* 9 (13), 11352–11370. doi:10.18632/oncotarget.24250

Yang, Y., Qiang, X., Owsiany, K., Zhang, S., Thannhauser, T. W., and Li, L. (2011). Evaluation of Different Multidimensional LC-MS/MS Pipelines for Isobaric Tags for Relative and Absolute Quantitation (iTRAQ)-Based Proteomic Analysis of Potato Tubers in Response to Cold Storage. *J. Proteome Res.* 10 (10), 4647–4660. doi:10.1021/pr200455s

Zhang, C., McFarlane, C., Lokireddy, S., Bonala, S., Ge, X., Masuda, S., et al. (2011). Myostatin-deficient Mice Exhibit Reduced Insulin Resistance through Activating the AMP-Activated Protein Kinase Signalling Pathway. *Diabetologia* 54 (6), 1491–1501. doi:10.1007/s00125-011-2079-7

Zhang, C., McFarlane, C., Lokireddy, S., Masuda, S., Ge, X., Gluckman, P. D., et al. (2012). Inhibition of Myostatin Protects against Diet-Induced Obesity by Enhancing Fatty Acid Oxidation and Promoting a Brown Adipose Phenotype in Mice. *Diabetologia* 55 (1), 183–193. doi:10.1007/s00125-011-2304-4

Zhao, J. Y., and Bass, K. D. (2018). Skeletal Muscle Regeneration by Extracellular Matrix Biological Scaffold: a Case Report. *J. Wound Care* 27 (9), S11–S14. doi:10.12968/jowc.2018.27.Sup9.S11

Zimmers, T. A., Davies, M. V., Koniaris, L. G., Haynes, P., Esquela, A. F., Tomkinson, K. N., et al. (2002). Induction of Cachexia in Mice by Systemically Administered Myostatin. *Science* 296 (5572), 1486–1488. doi:10.1126/science.1069525

**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: 04 January 2022
doi: 10.3389/fgene.2021.743757

# circTAF8 Regulates Myoblast Development and Associated Carcass Traits in Chicken

*Kan Li*[1,2], *Weichen Huang*[1,2], *Zhijun Wang*[1,2], *Yangfeng Chen*[1,2], *Danfeng Cai*[1,2] *and Qinghua Nie*[1,2]*

[1]Department of Animal Genetics, Breeding and Reproduction, College of Animal Science, South China Agricultural University, Guangzhou, China, [2]National-Local Joint Engineering Research Center for Livestock Breeding, Guangdong Provincial Key Lab of Agro-Animal Genomics and Molecular Breeding, and Key Laboratory of Chicken Genetics, Breeding and Reproduction, Ministry of Agriculture, Guangzhou, China

Recent studies have shown that circular RNAs (circRNAs) play important roles in skeletal muscle development. CircRNA biogenesis is dependent on the genetic context. Single-nucleotide polymorphisms in the introns flanking circRNAs may be intermediate-inducible factors between circRNA expression and phenotypic traits. Our previous study showed that circTAF8 is an abundantly and differentially expressed circRNA in leg muscle during chicken embryonic development. Here, we aimed to investigate circTAF8 function in muscle development and the association of the SNPs in the circTAF8 flanking introns with carcass traits. In this study, we observed that overexpression of circTAF8 could promote the proliferation of chicken primary myoblasts and inhibit their differentiation. In addition, the SNPs in the introns flanking the circTAF8 locus and those associated with chicken carcass traits were analyzed in 335 partridge chickens. A total of eight SNPs were found associated with carcass traits such as leg muscle weight, live weight, and half and full-bore weight. The association analysis results of haplotype combinations were consistent with the association analysis of a single SNP. These results suggest that circTAF8 plays a regulatory role in muscle development. These identified SNPs were found correlated with traits to muscle development and carcass muscle weight in chickens.

**Keywords: circTAF8, snps, flanking introns, non-coding RNA, muscle development, carcass traits, chicken**

OPEN ACCESS

***Edited by:***
*Andressa Oliveira De Lima, University of Washington, United States*

***Reviewed by:***
*Theros Ng, Western University of Health Sciences, United States*
*Juliana Afonso, Embrapa Pecuária Sudeste, Brazil*

***Correspondence:**
*Qinghua Nie*
*nqinghua@scau.edu.cn*

***Specialty section:***
*This article was submitted to Livestock Genomics, a section of the journal Frontiers in Genetics*

***Received:*** *19 July 2021*
***Accepted:*** *30 November 2021*
***Published:*** *04 January 2022*

***Citation:***
*Li K, Huang W, Wang Z, Chen Y, Cai D and Nie Q (2022) circTAF8 Regulates Myoblast Development and Associated Carcass Traits in Chicken. Front. Genet. 12:743757. doi: 10.3389/fgene.2021.743757*

## INTRODUCTION

Skeletal muscle development directly impacts carcass yield for meat consumption and is affected by heredity, nutrition, breed, sex, and environment (Fortin et al., 1987; Houba et al., 2004; Halevy et al., 2006). From a genetic point of view, muscle development is under the precise regulation of a series of specific genes and signals, mainly including the myogenic regulatory factor family and myocyte enhancer factor-2 family, the paired box transcription factors Pax3 and Pax7, and myostatin (Pas and Visscher, 1994; Grefte et al., 2007). In addition to these coding genes, a growing number of studies have found that noncoding RNAs also play important roles in muscle development (Luo et al., 2013; Cai et al., 2017; Simona et al., 2018).

Circular RNAs (circRNAs) are closed circular RNA molecules formed by back-splicing of a precursor mRNA, lacking a 3′ end poly-A tail and an 5′ end cap structure (Kristensen et al., 2019). CircRNAs are widely present in eukaryotic animals and participate in various biological processes. Studies have found that circRNAs are closely related to myogenesis, the transformation of muscle

fiber types, and skeletal muscle diseases (Kyei et al., 2020; Li et al., 2020; Chen et al., 2021). Our group has previously reported that circSVIL (Ouyang et al., 2018a), circHIPK3 (Chen et al., 2019), and circFGFR2 (Chen et al., 2018) promote the proliferation and differentiation of chicken myoblasts. Biological functions of circRNAs mainly include acting as miRNA sponges, interacting with various RNA-binding proteins (RBPs), and cap-independent translation themselves (Chen and Yang, 2015). Currently, studies on the mechanism whereby circRNAs regulate muscle development mainly focus on the interaction of circRNAs with miRNA (Ouyang et al., 2018a; Chen et al., 2019).

Single-nucleotide polymorphisms (SNPs) are variations of only one nucleotide variation and are widely used in studying animal and plant genetics. To date, numerous SNP markers have been identified to be associated with various important economic traits of chickens (Gorbach et al., 2010; Niknafs et al., 2014). The three SNPs of the growth hormone receptor have been found and genotyped in an F2 full-sib chicken population. G6631778A is related to body weight at various ages, dressed weight, subcutaneous fat thickness, and hatching weight in the roosters. G6631778A is only related to the 28-day-old body weight in the hens (Ouyang et al., 2008). Several SNPs of amylase alpha 1A were associated with leg muscle weight and daily gain (Zhang et al., 2021). Because of SNPs' wide distribution, high marker density, and high genetic stability, they have become an indispensable tool in chicken genetic breeding. With the in-depth study on circRNAs, studies have shown that SNPs could affect the formation of circRNAs and change their expression level (Paraboschi et al., 2018; Liu et al., 2019; Gao et al., 2021). The multiple-sclerosis-associated SNPs on the STAT3 gene affect the expression level of circRNA has_circ_0043813 (Paraboschi et al., 2018). The rs12196996 polymorphism in the introns flanking circFOXO3 can change circFOXO3 expression and increase the risk of coronary artery disease (Zhou et al., 2020). The biogenesis of circRNAs is influenced by *cis*-acting elements and *trans*-acting splicing factors, both of which require the participation of circRNA-flanking introns (Kristensen et al., 2019). SNPs in introns flanking circRNAs may modulate back-splicing of circRNA precursors, thereby affecting the production of circRNAs.

We have previously shown that circTAF8 is one of the top ten abundantly expressed circRNAs in chicken leg muscle at three different time points in embryonic development (GSE89355) (Ouyang et al., 2018b), indicating that circTAF8 functions in muscle development. In this study, we assessed the function of circTAF8 in the proliferation and differentiation of primary myoblasts and analyzed the association of the SNPs in the introns flanking circTAF8 with chicken carcass traits. Accordingly, this study aimed to assess whether phenotypic traits are associated with circRNAs and the SNPs in the circRNA-flanking introns.

# MATERIALS AND METHODS

All animal experiments and sampling procedures used in this study were strictly implemented in accordance with the regulations of the ethics committee of laboratory animals of South China Agricultural University (approval ID: SCAU#2020C030). All samples and carcass-trait data were collected in Guangzhou Kwangfeng Industrial Co., Ltd. Guangzhou Kwangfeng Industrial Co., Ltd., is the animal experimental unit operated under South China Agricultural University.

## Experimental Animal and Sample Collection

A total of 335 healthy partridge chickens of the M3 line (77 males and 258 females) were selected to screen for SNPs. Blood samples were collected in anticoagulant tubes containing 0.5 M EDTA, and E.Z.N.A.® Blood DNA Mini Kit (Omega Bio-tek, Norcross, GA, United States) was used to isolate the genomic DNA from blood. The concentration and quality of DNA were determined using NanoDrop One (Thermo Fisher Scientific, Seattle, WA, United States) and 1% agarose gel electrophoresis. The DNA samples were stored in an ultra-low temperature refrigerator at –80°C for later use.

The phenotypic traits mainly included live weight before slaughter (LWBS), carcass weight (CW), half-bore weight (HBW), full-bore weight (FBW), pectoral muscle weight (PMW), leg muscle weight (LMW), wing weight (WW), foot weight (FW), head weight (HW), heart weight (HW), liver weight (LW), stomach weight (SW), abdominal fat weight (AFW), shank length (SL), and shank diameter (SD). These obtained data were quantified and analyzed in Microsoft Excel 365 (Microsoft Corporation, Redmond, WA, USA).

In addition, four 7-week-old white Recessive Rock (WRR) chickens with similar weights were selected. The heart, liver, spleen, lung, kidney, pectoral, leg chicken, cerebellum, epencephalon, and abdominal fat were collected to characterize the tissue expression profiles of circTAF8 and TAF8.

## Genotyping SNPs *via* Polymerase Chain Reaction

Primers based on the TAF8 gene sequence provided by NCBI (accession number: NC_052,555.1) were designed using the Oligo 7 software version 7.56 (Molecular Biology Insights, Cascade, CO, USA) and synthesized by Beijing Tsingke Biotechnology Co., Ltd. (Beijing, China). The information about the primers used is shown in **Supplementary Table S1**. PCR was performed in a total volume of 30 μl including the following: 2 μl DNA template, 15 μl 2× Flash PCR Master Mix (CWBio, Beijing, China), 2 μl forward and reverse primers each, and 9 μl $ddH_2O$. The PCR program was as follows: pre-denaturation 95°C for 5 min; 35 cycles of denaturation at 94°C for 25 s, annealing at 60°C for 25 s, and extension at 72°C for 10 s; final extension at 72°C for 5 min. The PCR products were analyzed *via* 1% agarose gel electrophoresis.

## Isolation, Culture, Differentiation, and Transfection of Chicken Primary Myoblasts

Primary myoblasts were isolated from the leg muscle of chick embryos (eggs were from Zhuhai Yuhe Agriculture and Animal Husbandry Co., Ltd.) on day 11, as previously described (Cai et al., 2017). The isolated leg muscles were washed with phosphate-buffered saline (PBS; Gibco, Carlsbad, CA, USA)

containing 1% penicillin/streptomycin (Gibco, CA, USA), and visible skin and bone tissue were removed using sterile forceps. The minced muscle tissue was digested with 0.25% trypsin-EDTA (Gibco, CA, USA) at 37°C for 20 min. The cell suspension after digestion was filtered and centrifuged at 1,000 *g* for 5 min. The supernatant was discarded, and the precipitated cell pellet was suspended in the growth medium, consisting of RPMI-1640 medium (Gibco, CA, USA), 20% fetal bovine serum (Gibco, CA, USA), and 1% penicillin/streptomycin. A twice 40-min interval attachment was used to purify myoblast. Then, cells were counted and seeded in cell culture dishes and cultured in an incubator at 37°C in a 5% $CO_2$-humidified atmosphere. After the cell density reached 95%–100%, the medium was replaced with the differentiation medium (RPMI 1640 Medium, 2% horse serum [Solarbio, Beijing, China], and 1% penicillin/streptomycin).

All cells were transfected using Lipofectamine 3000 reagent (Invitrogen, CA, USA) by following the manufacturer's directions.

## RNA Extraction, cDNA Synthesis, and RT-PCR

Total RNA was extracted according to the SteadyPure Universal RNA Extraction Kit (Accurate Biology, Wuhan, China) instructions. The concentration and integrity of extracted RNA were tested for using NanoDrop One and 1% gel electrophoresis. Evo M-MLV RT Kit (Accurate Biology, Wuhan, China) was used for reverse transcription. The cDNA obtained by transcription was used for the next step of RT-PCR. The PCR amplification procedure was as follows: pre-denaturation at 95°C for 20 s; 40 cycles of denaturation at 95°C for 20 s, and annealing and extension at (58–60) °C for 30 s. The GAPDH gene was used as an internal reference gene. All relative RNA expression was calculated using $2^{-\Delta\Delta\ CT}$ method (Livak and Schmittgen, 2001). The primer information is shown in **Supplementary Table S1**.

## CircTAF8-Overexpressed Plasmid Construction

Forward and reverse primers based on the general rules of PCR-primer design were designed to amplify the linear sequence of circTAF8. After the design was completed, an EcoRI restriction site, a forward circularization-mediating sequence, and an AG acceptor were added in the 5′ end of the forward primer; a BamHI restriction site, a reverse circularization -mediating sequence, and a GT donor were added in the 5′ end of the reverse primer. The linear sequence of circTAF8 was amplified *via* PCR from the myoblast's total cDNA, which was cloned into the pCD25-ciR vector by using BamHI and EcoRI restriction sites.

The primer sequences were as follows:

Forward primer: CGGAATTCTAATACTTTCAGAGGT CGGGCAGCAAGCACAC.

Reverse former: CGGGATCCAGTTGTTCTTACTGGTGTC TTGATGTAGGTGT.

## Nuclear and Cytoplasmic RNA Extraction

The total number of cells required for this experiment was about $2 \times 10^7$. First, the cells were washed 3 times with precooled PBS; then, the cells were scraped off using a cell scraper, collected in a 50-ml centrifuge tube, and centrifuged at 1,000 *g* for 10 min. Subsequently, each cell pellet was resuspended in 1 ml cell-disruption buffer (2 M KCl, 1 M $MgCl_2$, 1 M pH 7.5 Tris–HCl, 0.5 mM DTT) and then incubated on ice for 10 min. Afterward, the cells were transferred to a pre-baked Dounce cell homogenizer (Solarbio, Beijing, China) and homogenized with 15 strokes using a pestle. The homogenates were transferred to a new microtube, gently mixed with Triton X-100 (Beyotime, Shanghai, China) at a final concentration of 0.1%, and centrifuged at 1,500 g for 15 min. The supernatants, corresponding to the cytoplasmic fraction, were transferred to a new microtube. The pellets correspond to the nuclear fraction. Nuclear and cytoplasmic RNA was extracted using the TRIzol reagent (Sigma-Aldrich, San Francisco, CA, USA).

## RNase R Digestion and Actinomycin D Treatment

Total RNA (2.5 μg) was treated with 10U RNase R (Geneseed, Guangzhou, China) for 30 min at 37°C. The RNA digestion efficiency was evaluated using electrophoresis.

Primary chicken myoblasts were cultured in the presence or absence of 5 μg/ml actinomycin D (MedChemExpress, Shanghai, China) for 6 h. Then, total RNA from various treatment groups was extracted using the SteadyPure Universal RNA Extraction Kit.

## Cell Counting Kit 8 Assay

Cells at the proliferation phase were seeded in a 96-well plate. When they reached 60% confluence, they were transfected with the indicated constructs. The CCK-8 reagent (Beyotime, Shanghai, China) was used to detect cell proliferation at 12, 24, 36, and 48 h of culture. The absorbance at 450-nm wavelength was measured using a microplate reader Model 680 (Bio-Rad, Berkeley, CA, USA).

## 5-Ethynyl-2′-Deoxyuridine Assay

As previously reported (Cai et al., 2021), Edu staining was performed using Cell Light EdU Apollo 488 *In Vitro* Kit (Ruibo, Guangzhou, China). Briefly, cells were incubated in 50 μM EdU solution for 2 h and then fixed for 30 min using 4% paraformaldehyde solution. Subsequently, 0.1% Triton X-100 solution was used for cell permeabilization. Finally, the cells were incubated with Hoechst reaction solution at room temperature in the dark for 30 min. A Leica DMi8 fluorescence microscope (Leica Microsystems, Wetzlar, Germany) and ImageJ software (U.S. National Institutes of Health, Bethesda, MD, United States) (Schneider et al., 2012) were used to acquire and analyze the image, respectively.

## Flow Cytometric of the Cell Cycle

After 48 h of transfection, cells were collected and fixed in 70% ethanol at –20°C for 24 h. Then, they were centrifuged at 1,000 *g* for 5 min, and the cell pellets were resuspended in precooled PBS. Subsequently, 0.5 ml *p*I/RNase staining buffer (BD Biosciences, Franklin Lakes, NJ, USA) was added to each sample, and the

sample was incubated at 37°C in the dark for 30 min. Flow cytometric analysis was performed on a BD FACSCalibur (BD Biosciences, NJ, USA), and the data were processed using FlowJo software (Treestar, Woodburn, OR, USA).

## Immunofluorescence

The cells after transfection were washed twice with PBS, fixed in 4% formaldehyde solution for 30 min, and permeated with 0.1% Triton X-100 at room temperature for 10 min. After blocking with goat serum (Boster, Wuhan, China) for 30 min, the cells were incubated overnight with MyHC antibody (1:50 dilution; DHSB, IA, USA) at 4°C. After washing 3 times with PBS, the goat anti-mouse IgM/FITC antibody (1:1,000 dilution; Bioss, Beijing, China) was added and incubated at room temperature for 1 h. Finally, the nucleus was stained with a DAPI staining solution (Beyotime, Shanghai, China). A fluorescence microscope and ImageJ software were used to acquire and analyze the image, respectively.

## Western Blot

rBriefly, cells were washed twice with precooled PBS, mixed with the RIPA buffer (Solarbio, Beijing, China) containing 1% PMSF (Solarbio, Beijing, China), and then incubated on ice for 30 min. The cell lysates were centrifuged at 12,000 *g* for 10 min. The supernatant was mixed with the 5× SDS-PAGE loading buffer and then incubated at 95°C for 5 min. The extracted proteins were separated *via* SDS-PAGE and transferred to a PVDF membrane (400 mA, 30 min). The membrane was blocked with 5% skimmed milk powder for 1 h, then incubated with a primary antibody solution overnight at 4°C. Afterward, the PVDF membrane was washed three times with TBST solution (Beyotime, Shanghai, China) for 5 min and then incubated with the secondary antibody solution at room temperature for 60 min. The protein immunoblot results were analyzed using the Odyssey Fc system (LI-COR, Lincoln, NE, USA). β-Tubulin was used as an internal reference. The relative protein levels normalized the β-tubulin protein content. The antibody information is as follows: rabbit anti-β-tubulin (1:1,000 dilution; Bioss, Beijing, China), rabbit anti-MyoD1 (1:1,000 dilution; Bioss, Beijing, China), MyHC antibody (1:1,000 dilution; DHSB, IA, USA), goat anti-rabbit IgG H&L antibody (1:3,000 dilution; Bioss, Beijing, China), and goat anti-mouse IgG H&L antibody (1:3,000 dilution; Bioss, Beijing, China).

## Statistical Analyses

The genotype frequency and gene frequency distribution of different alleles were evaluated using Microsoft Excel. The polymorphism information content (PIC) calculation of SNPs sites was based on the formula

$$PIC = 1 - \sum_{i=1}^{n} p_i^2 - \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} 2p_i^2 p_j^2$$

Pi and Pj are the frequencies of the "ith" and "jth" alleles, respectively, and n is the number of multiple alleles. PIC >0.5 indicates high polymorphism, PIC <0.25 indicates low polymorphism, and 0.25 < PIC <0.5 indicates moderate polymorphism.

The mixed linear model in SPSS 25.0 software (IBM, Armonk, NY, USA) was used to conduct an association analysis between different combinations of genotypes and haplotypes of polymorphic sites and chicken carcass traits. The results are shown as "mean ± standard error." The analysis model is as follows:

$$Y_{ijlm} = \mu + G_i + S_j + f_l + e_{ijlm}$$

*Yijlm* is the observed value, $\mu$ is the overall average, *Gi* is the fixed effect of the genotype, *Sj* is the fixed effect of sex, $fl$ is the random effect of the family, and $e_{ijlm}$ is the random error term. Bonferroni's test was performed to control for multiple comparisons. In addition, we used online SHesis (Yong and He, 2005) and EMBOSS Needle (Madeira et al., 2019) software to analyze the linkage disequilibrium of SNP sequence complementary information, respectively.

*A priori* power analyses were performed using G*Power software version 3.1.9.7 (Heinrich-Heine-Universität Düsseldorf, Düsseldorf, Germany) (Faul et al., 2009) based on the effect size and standard deviation of a previous publication (Chen et al., 2021) or preliminary data to achieve a significance level (α) of 0.05 and a power of 0.8. In this study, three biological replicates were included in the analyses unless stated otherwise. The statistical analyses and drawings were performed using GraphPad Prism 8.0 (GraphPad Software, CA, USA). The statistical significance of the difference between the two groups was calculated using the unpaired Student's t-test. One-way analysis of variance (ANOVA) was used to evaluate the differences among multiple groups, followed by Tukey's multiple-comparison test. The data were shown as mean ± standard error of the mean (SEM).

# RESULTS

## Identification of circTAF8 Molecular Characteristics

In this study, to verify that circTAF8 is not a technical artifact, specific divergent and convergent primers were designed based on the back-splicing sites of circTAF8. Primary myoblast cDNA and genomic DNA (gDNA) were used for PCR. The PCR products of divergent primer were sent to Beijing Tsingke Biotechnology Co., Ltd., for Sanger sequencing. The sequencing results confirmed that circTAF8 is the product of the head-to-tail cyclization of exons 2, 3, 4, and 5 of the protein-coding gene TAF8 (441 bp in total), located on chromosome 26 (**Figure 1A**). The electrophoresis results showed that both divergent and convergent primers could yield PCR products when cDNA was used as the template, whereas only the convergent primers could yield a product when gDNA was used as the template (**Figure 1B**).

CircRNAs showed resistance to digestion with exonuclease RNase R due to the lack of a 3′ end poly-A tail. Using RT-PCR, we quantified the expression levels of circTAF8 and its parent gene *TAF8* after RNase R digestion. Meanwhile, circHIPK3 was used as a positive control and *β-actin* as a negative control. The results

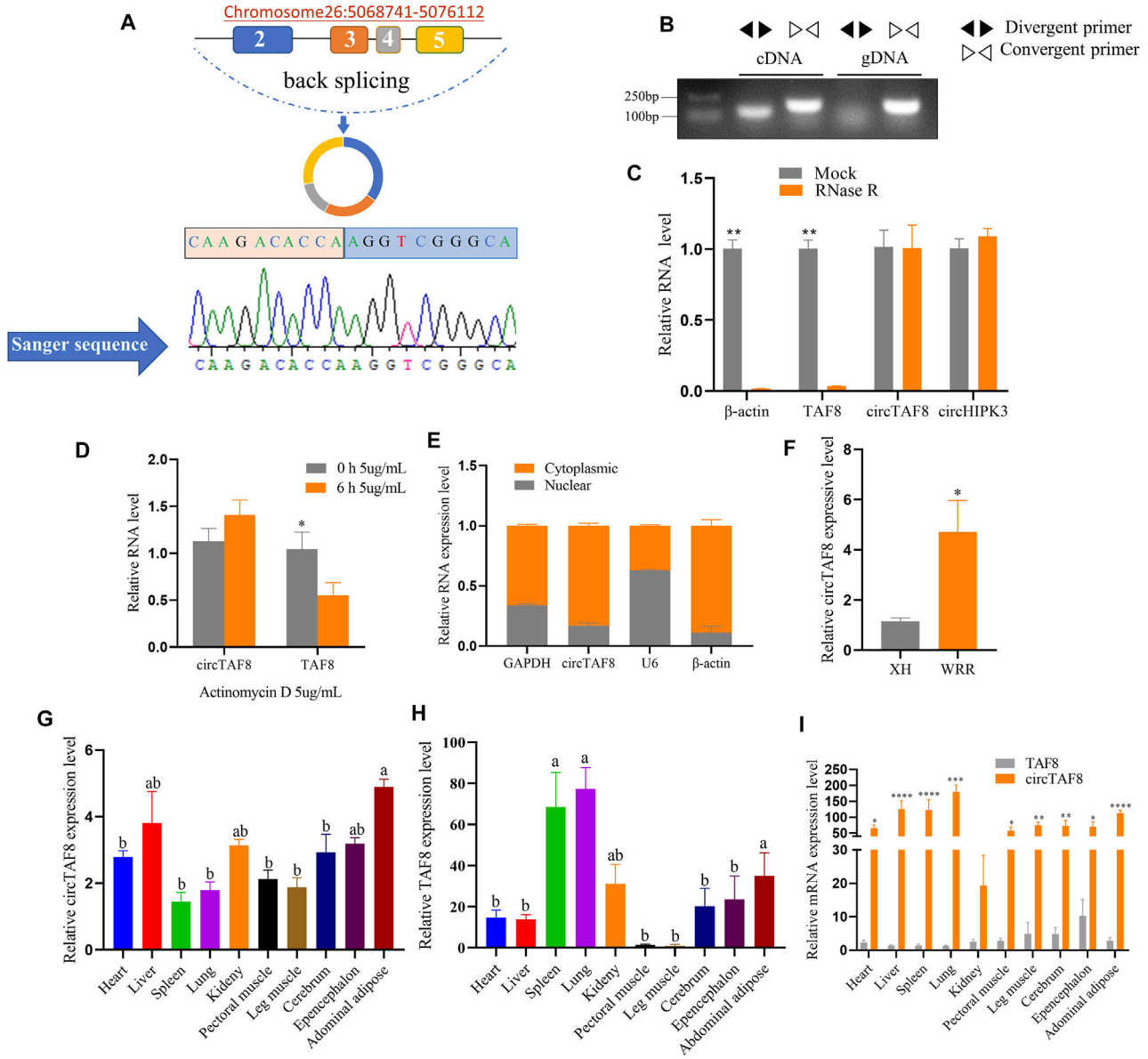


**FIGURE 1 |** Identification of circTAF8 in chicken. **(A)** Genomic location and looping model of circTAF8 and Sanger sequencing results. **(B)** Electrophoresis results of circTAF8 amplified from cDNA and gDNA with divergent and convergent primers, respectively. **(C)** Reverse transcription-polymerase chain reaction (RT-PCR) assays to detect β-actin, TAF8, circTAF8, and circHIPK3 RNA expression levels with and without RNase R treatment. **(D)** RT-PCR to detect circTAF8 and TAF8 RNA expression levels with and without actinomycin D treatment. **(E)** RT-PCR results showing the circTAF8 subcellular localization. **(F)** RT-PCR results showing the circTAF8 expression level in XH and WRR groups. **(G)** Relative RNA expression level of circTAF8 in diverse tissues. **(H)** Relative *TAF8* RNA levels in diverse tissues. **(I)** The relative levels of *TAF8* RNA and circTAF8 in each tissue. The results of are shown as mean ± S.E.M from at least three biological replicates. The statistical significance of the differences was assessed using the unpaired Student's *t*-test or one-way analysis of variance (ANOVA) (* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; **** $p < 0.0001$).

showed that the expression levels of circTAF8 and circHIPK3 did not change, and linear *TAF8* and *β-actin* were almost undetectable ($p < 0.01$; **Figure 1C**). Then, since circRNAs usually have a longer half-life than linear RNA, we performed actinomycin D (5 μg/ml) treatment to myoblasts for 6 h for the stability evaluation of circTAF8. RT-PCR results showed no significant difference in the expression level of circTAF8 after actinomycin D treatment. In contrast, the expression of the parent gene *TAF8* was significantly reduced ($p < 0.05$; **Figure 1D**). Nuclear and cytoplasmic RNA extraction assays were employed to determine the subcellular localization of circTAF8. *U6*, *β-actin*, and *GAPDH* are used as controls; *U6* is mainly distributed in the nucleus, and *β-actin* and *GAPDH* are mainly distributed in the cytoplasm. The results showed that circTAF8 is present in both the cytoplasm and the nucleus (**Figure 1E**). Next, the expression level of circTAF8 was detected in the Xinghua (XH) and WRR chicken groups

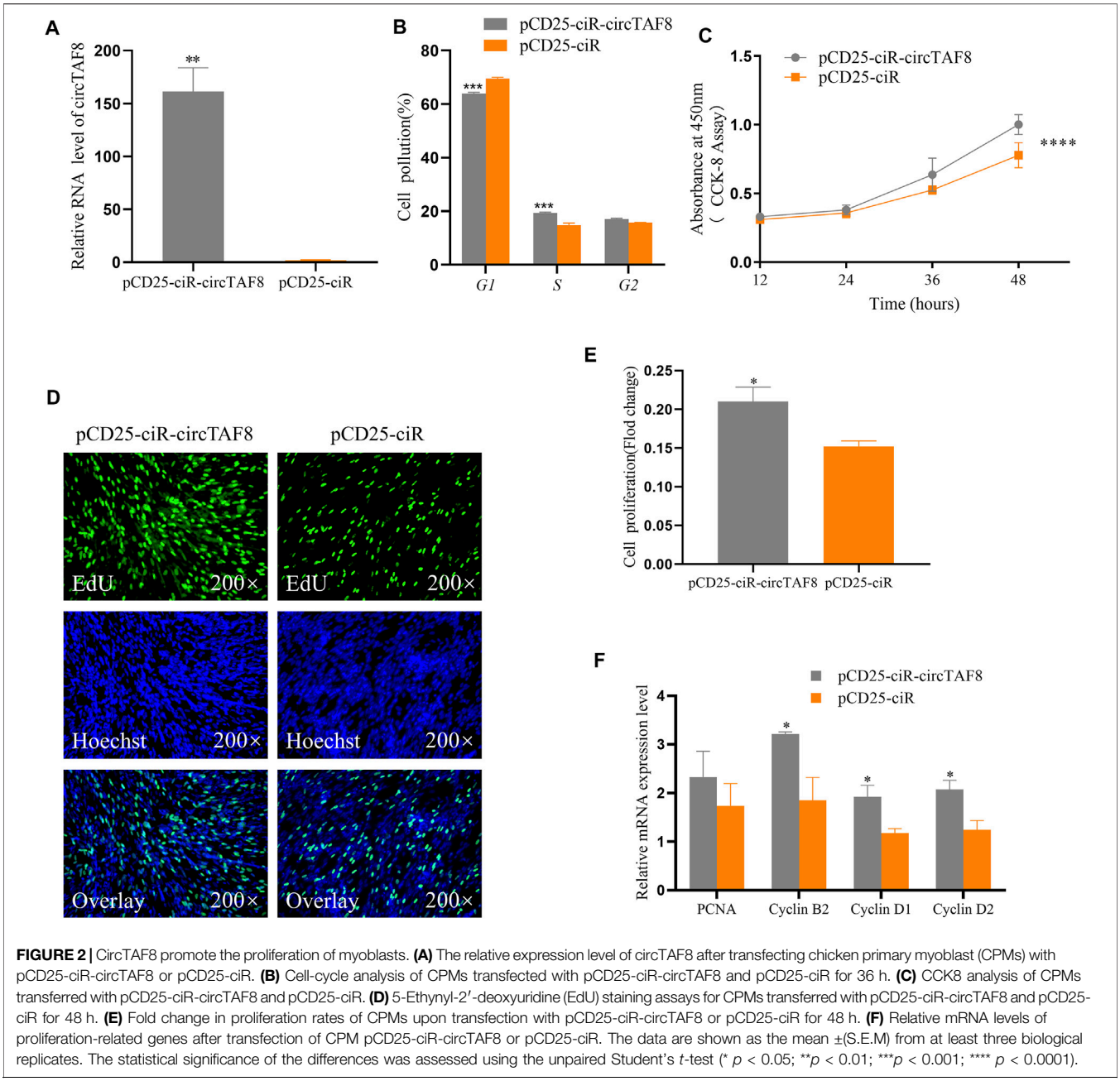


**FIGURE 2 |** CircTAF8 promote the proliferation of myoblasts. **(A)** The relative expression level of circTAF8 after transfecting chicken primary myoblast (CPMs) with pCD25-ciR-circTAF8 or pCD25-ciR. **(B)** Cell-cycle analysis of CPMs transfected with pCD25-ciR-circTAF8 and pCD25-ciR for 36 h. **(C)** CCK8 analysis of CPMs transferred with pCD25-ciR-circTAF8 and pCD25-ciR. **(D)** 5-Ethynyl-2′-deoxyuridine (EdU) staining assays for CPMs transferred with pCD25-ciR-circTAF8 and pCD25-ciR for 48 h. **(E)** Fold change in proliferation rates of CPMs upon transfection with pCD25-ciR-circTAF8 or pCD25-ciR for 48 h. **(F)** Relative mRNA levels of proliferation-related genes after transfection of CPM pCD25-ciR-circTAF8 or pCD25-ciR. The data are shown as the mean ±(S.E.M) from at least three biological replicates. The statistical significance of the differences was assessed using the unpaired Student's *t*-test (* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; **** $p < 0.0001$).

by RT-PCR, and the results showed that the expression of circTAF8 in WRR was significantly higher than in XH ($p < 0.05$; **Figure 1F**). XH and WRR are two breeds with different growth rates. The expression levels of circTAF8 and TAF8 in diverse tissues were quantified using qPCR. The expression profile data showed that the circTAF8 level was highest in the abdominal adipose (**Figure 1G**), and TAF8 was the highest in the lung (**Figure 1H**). Interestingly, compared with other tissues, the expressions of circTAF8 and TAF8 in muscle seems to be relatively low. In addition, the circTAF8 level in various tissues is significantly higher than that of TAF8 RNA (**Figure 1I**). In the lung and muscle, the circTAF8 level is approximately 200- and 50-fold the *TAF8* RNA level, respectively.

## CircTAF8 Promotes Myoblast Proliferation

To study the effect of circTAF8 on the proliferation of primary myoblasts, primary myoblasts were transfected with the circTAF8 overexpression plasmid. The expression level of circTAF8 was detected by RT-PCR 48 h after transfection. The results showed that compared with the control group, the overexpression effect of circTAF8 reached a significant level ($p < 0.01$; **Figure 2A**). Then, flow cytometric analysis was performed to evaluate the cell cycle in primary myoblasts, and the results showed that the

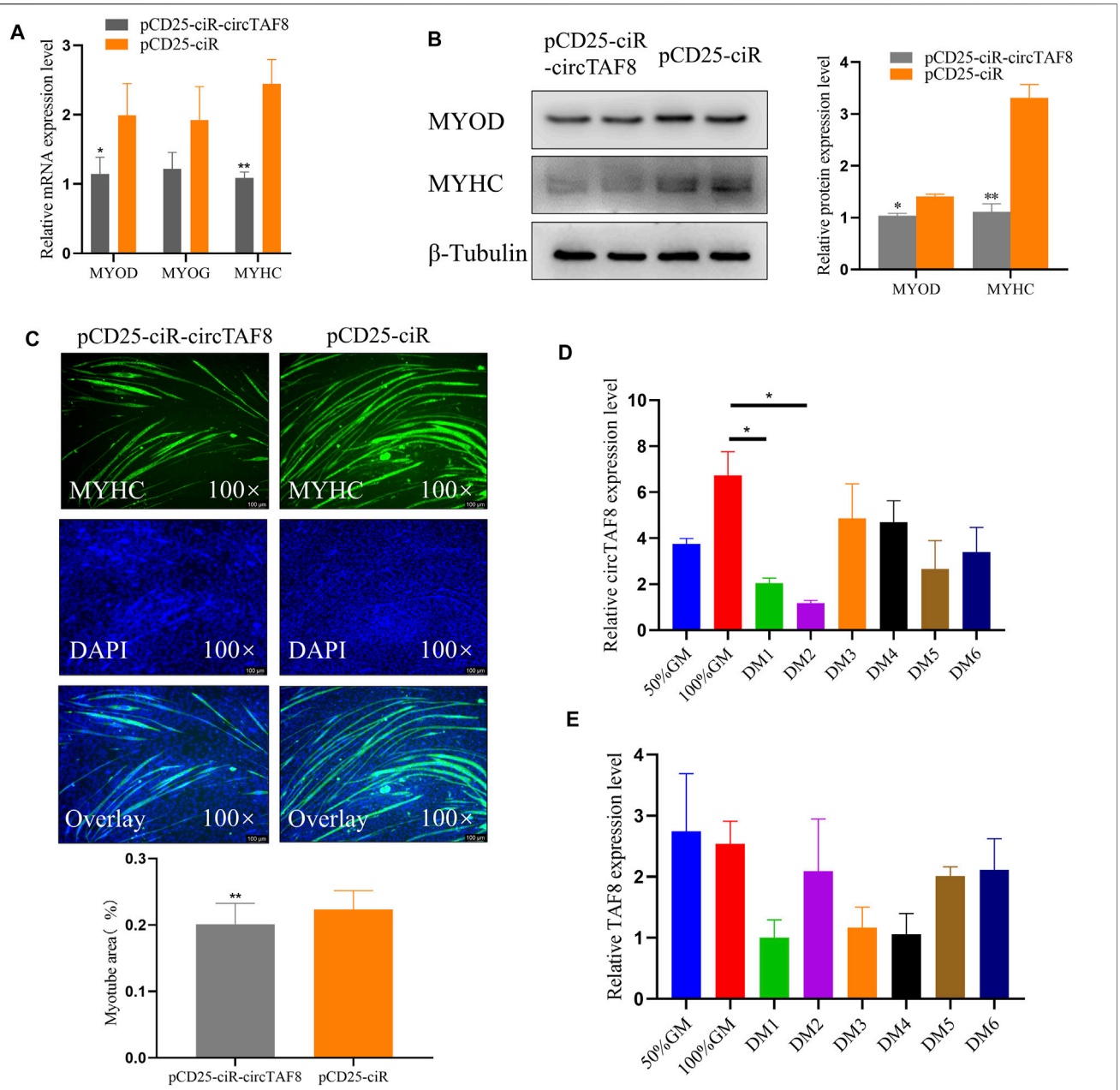


**FIGURE 3 |** CircTAF8 inhibits myoblast differentiation. **(A)** Reverse transcription–polymerase chain reaction results showing the mRNA level of the cell differentiation-related genes in CPMs upon transfection with pCD25-ciR-circTAF8 or pCD25-ciR. **(B)** Western blot results showing the MYOD and MYOG protein levels in CPMs upon transfection with pCD25-ciR-circTAF8 or pCD25-ciR. **(C)** MyHC Immunofluorescence analysis of the myotubes transfected with pCD25-ciR-circTAF8 or pCD25-ciR. **(D)** Relative circTAF8 RNA level at various timepoints during myoblast differentiation. **(E)** Relative *TAF8* RNA levels at various timepoints during myoblast differentiation. The results are shown as mean ± S.E.M from at least three biological replicates. The statistical significance of the differences was assessed using the unpaired Student's *t*-test or one-way analysis of variance (ANOVA) (* $p < 0.05$; ** $p < 0.01$).

overexpression of circTAF8 increased the number of cells in the S phase and decreased the number of cells in the G1/0 phase ($p < 0.001$; **Figure 2B**). Besides, the CCK8 experiment was used to determine the proliferation status of myoblasts at various time points after overexpression of circTAF8. The results showed that the proliferation activity of the overexpression circTAF8 group was higher than that of the control group ($p < 0.0001$; **Figure 2C**). Afterward, the proliferation of the primary myoblasts was detected by EdU incorporation assays. EdU staining results showed that the ectopic expression of circTAF8 could significantly promote the proliferation of primary myoblasts ($p < 0.05$; **Figures 2D,E**). RT-PCR was used to detect the expression of multiple proliferation-related marker genes, and the results showed that the expression levels of Cyclin D1, Cyclin

D2, and Cyclin B2 were significantly increased ($p < 0.05$; **Figure 2F**).

## CircTAF8 Inhibit Myoblast Differentiation

We transfected the circTAF8 overexpression plasmid into chicken myoblast cells to investigate the potential role of circTAF8 on primary myoblast differentiation. After cell density achieved 95%–100%, the growth medium was replaced by the differentiation medium. The total RNA and protein were extracted after 36 h of continuous differentiation. RT-PCR was used to detect the expression level of relevant differentiation marker genes, and the results showed that compared with the control group, the expressions of MyoD and MyHC were significantly or extremely significantly reduced ($p < 0.05$; $p < 0.01$; **Figure 3A**). In addition, the protein expressions of MYOD and MYHC were detected by Western blot, and the results were similar to the changes in mRNA level ($p < 0.05$; $p < 0.01$; **Figure 3B**). The immunofluorescence of MyHC showed that the overexpression of circTAF8 significantly reduced the myotube and the overall muscle fiber area ($p < 0.01$; **Figure 3C**). Afterward, we compared the expression levels of circTAF8 and TAF8 at different differentiation times of myoblasts (**Figures 3D,E**). The results showed that the expression level of circTAF8 during the proliferation phase was relatively high, and the expression level significantly decreased after differentiation began and then slowly increased. The expression level of TAF8 in different developmental stages of myoblasts did not reach a statistically significant difference.

## Association of the SNPs in the Introns Flanking circTAF8 With Carcass Traits

To investigate the distribution of SNPs in the introns flanking circTAF8, primers P1, P2, and P3 were used for PCR of a mixed pool of 50 DNA samples, and the PCR products were sent to Beijing Tsingke Biological Co., Ltd., for sequencing. The results show that 10 and seven SNP sites were detected in the PCR products obtained from P1 and P2, respectively, and no SNPs were detected in the PCR products of P3. The peak figures of the above sequencing results are shown in **Supplementary Figure S2**. All the detected SNP sites are located in the 5′ flanking region of the TAF8 gene. Subsequently, direct sequencing was applied to all 335 partridge chicken DNA PCR products amplified with P1 and P2, and the sequencing files were analyzed by SeqMan software (DNASTAR, Madison, WI, USA). The analysis results showed that all the above SNPs had detected three genotypes. The statistical results of the frequency of genotypes and alleles are shown in **Supplementary Table S2**. $X^2$ fitness test results showed that only 12 SNPs were in Hardy–Weinberg equilibrium ($p > 0.05$). The calculation results of PIC showed that except for g.-1576A > G, the remaining 16 SNPs were all moderate polymorphic, indicating that the genetic variation TAF8 was at a medium level and had great potential for selection.

The association analysis results showed that only eight SNPs were related to carcass traits (**Table 1**). g.-1771 G > C was significantly correlated with FBW and AFW ($p < 0.05$) and was extremely significantly correlated with SL ($p < 0.01$), and individuals of GG genotype had significantly higher FBW and SL. The locus g.-1576A > G was significantly correlated with HBW ($p < 0.05$). Both g.-1554 T > C and g.1480A > C were extremely significantly related to the LMW ($p < 0.01$), and TT and AA genotype individuals had significantly higher LMW ($p < 0.05$). The remaining four sites were associated with similar traits. g.-289 C > T, g.-288A > G, and g.-210 T > C were significantly associated with LWBS, HBW, and FBW, respectively ($p < 0.05$). g.-288A > G, g.-210 T > C, and g.-173A > G were significantly correlated with WW. g.-173A > G was also significantly related to FBW ($p < 0.05$).

## Linkage Disequilibrium and Haplotype Analysis of Eight SNPs circTAF8 Flanking Introns

The SHEsis online analysis software was used to analyze the linkage disequilibrium and haplotype of the above eight SNPs. The results showed that g.-1554 T > C and g.-1480A > C can form strong linkage disequilibrium. In addition, the remaining four loci g.-289 C > T, g.-288A > G, g.-210 T > C, and g.-173A > G were in a state of strong linkage disequilibrium (**Figure 4**). A total of seven haplotypes were found in the g.-1554 T > C and g.-1480A > C groups. Only four haplotypes were selected for further analysis, excluding the small individuals (**Table 2**). These four were CCAC (6), CCCC (191), TCAC (101), and TTAA (20). Ten haplotypes consist of other four SNPs. Haplotype combinations CCAATCAA (6), CCAATTAA (65), CCAGTCAG (16), CTAGTCAG 135), and TTGGCCGG 82) were used to analyze further the association with carcass traits, except for groups with fewer individuals.

The association results showed that haplotypes of g.-1554 T > C and g.-1480A > C significantly related to LWM ($p < 0.05$) and individuals of TTAA haplotypes had higher LMW (**Table 2**). The haplotypes composed of g.-289 C > T, g.-288A > G, g.-210 T > C, and g.-173A > G were significantly related to LWBS and FBW ($p < 0.05$) and were extremely significantly related to HBW and LMW ($p < 0.01$). Individuals of the CCAATCAA haplotype had higher LWBS, and individuals of the CCAGTCAG haplotype had higher LMW. Haplotype individuals of TTGGCCGG haplotype had higher FBW and HBW. The results of the association analysis of haplotype combinations were consistent with a single SNP.

## Complementary Pairing Analysis in Flank Sequence of circTAF8

We first used the online software EMBOSS Needle to analyze the sequence complementarity between the flanking introns of circTAF8 (**Figure 5A**). The comparison results showed multiple complementary sequences in the two introns. Then, the sequences of the short fragments upstream of the eight SNP sites were aligned with the sequences of the downstream flanking introns (**Figure 5B**). The results showed that all SNP sites have short complementary sequences.

# DISCUSSION

Muscle development is a complex process precisely regulated by specific genes and signaling pathways (Houba et al., 2004; Grefte

**TABLE 1 |** Association of eight SNPs in introns flanking circTAF8 with carcass traits.

| SNP | Trait | *p*-value | Least square mean ± standard error | | |
|---|---|---|---|---|---|
| g.-1771 G > C | | | GG (n = 6) | GC (n = 103) | CC (n = 214) |
| | FBW(g) | 0.037 | 1,128.33 ± 22.70$^{a}$ | 1,078.48 ± 6.30 | 1,071.99 ± 4.93$^{a}$ |
| | SL (mm) | 0.004 | 71.88 ± 1.58$^{A}$ | 66.44 ± 0.44$^{B}$ | 66.82 ± 0.35$^{AB}$ |
| | AFW(g) | 0.044 | 35.57 ± 4.32 | 26.52 ± 1.20 | 25.24 ± 0.95 |
| g.-1576A > G | | | AA (n = 253) | AG (n = 63) | GG (n = 7) |
| | HBW(g) | 0.030 | 1,290.27 ± 8.53$^{a}$ | 1,252.62 ± 15.22$^{a}$ | 1,333.36 ± 42.56 |
| g.-1554 T > C | | | TT (n = 21) | TC (n = 105) | CC(n = 197) |
| | LMW(g) | 0.001 | 253.78 ± 15.71$^{AB}$ | 192.21 ± 7.55$^{A}$ | 196.10 ± 5.75$^{B}$ |
| g.-1480A > C | | | AA (n = 21) | AC (n = 107) | CC(n = 195) |
| | LMW(g) | 0.001 | 254.39 ± 15.71$^{AB}$ | 192.36 ± 7.43$^{A}$ | 196.01 ± 5.87$^{B}$ |
| g.-289 C > T | | | CC(n = 88) | CT (n = 143) | TT (n = 83) |
| | LWBS(g) | 0.022 | 1,608.47 ± 9.92$^{a}$ | 1,599.25 ± 7.89 | 1,629.83 ± 9.97$^{a}$ |
| | HBW(g) | 0.045 | 1,278.47 ± 13.18 | 1,275.18 ± 10.16 | 1,312.68 ± 13.26 |
| | FBW(g) | 0.012 | 1,072.26 ± 6.63 | 1,068.22 ± 5.20$^{a}$ | 1,090.64 ± 6.66$^{a}$ |
| g.-288A > G | | | AA (n = 72) | AG (n = 157) | GG (n = 85) |
| | LWBS(g) | 0.012 | 1,615.45 ± 10.66 | 1,597.32 ± 7.77$^{a}$ | 1,628.30 ± 9.84$^{a}$ |
| | HBW(g) | 0.046 | 1,280.74 ± 14.32 | 1,274.13 ± 9.92$^{a}$ | 1,311.62 ± 13.06$^{a}$ |
| | FBW(g) | 0.007 | 1,076.01 ± 7.14 | 1,066.81 ± 5.09$^{A}$ | 1,090.05 ± 6.55$^{A}$ |
| | WW(g) | 0.034 | 64.23 ± 0.55 | 64.18 ± 0.41$^{a}$ | 65.55 ± 0.5$^{a}$ |
| g.-210 T > C | | | TT (n = 67) | TC (n = 158) | CC(n = 89) |
| | LWBS(g) | 0.017 | 1,612.00 ± 11.02 | 1,598.30 ± 7.82$^{a}$ | 1,628.94 ± 9.80$^{a}$ |
| | HBW(g) | 0.036 | 1,292.94 ± 14.88 | 1,270.77 ± 9.91$^{a}$ | 1,308.98 ± 12.91$^{a}$ |
| | FBW(g) | 0.002 | 1,075.76 ± 7.38 | 1,065.75 ± 5.11$^{A}$ | 1,092.10 ± 6.50$^{A}$ |
| | WW(g) | 0.048 | 64.39 ± 0.57 | 64.14 ± 0.41$^{a}$ | 65.48 ± 0.51$^{a}$ |
| g.-173A > G | | | AA (n = 66) | AG (n = 160) | GG (n = 88) |
| | FBW(g) | 0.004 | 1,075.28 ± 7.32 | 1,066.41 ± 5.09$^{A}$ | 1,090.77 ± 6.42$^{A}$ |
| | WW(g) | 0.047 | 64.53 ± 0.56 | 64.10 ± 0.41$^{a}$ | 65.43 ± 0.50$^{a}$ |

*The above values are "average values ± standard errors"; in each group of SNPs, unmarked letters in the same line indicate that the difference is not significant (*$p > 0.05$*). When the letters are the same, lowercase letters indicate significant differences (*$p < 0.05$*), and uppercase letters indicate significant differences (*$p < 0.01$*).*

et al., 2007). CircRNAs, as a new type of posttranscriptional regulator in skeletal muscle, have also been discovered in rhesus monkeys, mice, pigs, cattle, and sheep (Huang et al., 2018; Xu et al., 2018; Hao et al., 2020; Chen et al., 2021). Our previous studies have shown that circRNAs are abundantly and dynamically expressed during chicken muscle development (Ouyang et al., 2018b). According to the previous sequencing data, circTAF8 is highly and differentially expressed in the skeletal muscle at 11 embryo age, 16 embryo age, and 1-day post-hatch (Ouyang et al., 2018b), indicating that circTAF8 has a potential role in muscle development. To confirm this hypothesis, we first identified the molecular properties of circTAF8. PCR and Sanger sequencing results showed that circTAF8 was a back-splicing product derived from the two to five exons of protein-coding gene TAF8. In general, high cellular stability and longer half-life time were standard features of circRNAs due to their closed-ring structure (Kristensen et al., 2018; Kristensen et al., 2019). CircTAF8 showed resistance to digestion with exonuclease RNase R compared with the linear transcript in our results. In addition, the subcellular localization of circRNAs is closely related to their function (Wang et al., 2014). Our results showed that circTAF8 is mainly present in the cytoplasm. Multiple reports have found that most circRNAs that are related to skeletal muscle and function as miRNA sponges are located in the cytoplasm (Ouyang et al., 2018a; Chen et al., 2019; Chen et al., 2021). These results indicate that circTAF8 was reliable and stable in chicken muscle.

Breed type affects the rate of muscle development (Fortin et al., 1987). The circTAF8 expression level ranked second in circRNA sequencing data from the breast muscle of 7-week-old XH and WRR chickens (unpublished data), and the level was higher in WRR chicken than in XH chicken. XH chicken is a native slow-growing broiler; WRR is a typical fast-growing broiler (Ouyang et al., 2015). Our results presented here are consistent with sequencing data, suggesting that circTAF8 is a positive regulator of chicken muscle development. Cell confluence is essential for maintaining cell phenotype and regulating gene expression (Abo-Aziza and Zaki, 2017). Cell dynamics indicated that myoblasts would exit the exponential growth phase and begin to differentiate when they reach 100% confluence (Tanaka et al., 2011). In our study, the confluence of the chicken primary myoblasts for cell proliferation-related assays was about 60% when transfected, and the cells will fuse to 100% for 48 h after transfection. All proliferation-related assays were performed within 48 h. In comparison, the cell density for the differentiation-related study was about 70%–80% when transfected, and the myoblasts were continually differentiated using a differentiation medium for 36 h when cell density achieved 100% after transfection. Accordingly, the gain-of-function test of circTAF8 showed that circTAF8 promotes the proliferation of skeletal myoblasts and inhibits their differentiation in chickens. During myogenesis, the extent of myoblast proliferation largely determines the number of muscle fibers (te Pas, 2004). Commercial broilers usually have more muscle fibers than slow-growing chickens (Scheuermann et al., 2004; Al-Musawi et al., 2011). Therefore, the pro-proliferation effect of circTAF8 on myoblasts is essential for muscle development. However, the specific mechanism of circTAF8 in skeletal muscle regulation has not been elucidated yet, and further investigation is needed.

**A**

| D': | site2 | site3 | site4 | site5 | site6 | site7 | site8 |
|---|---|---|---|---|---|---|---|
| site1 | 0. 996 | 0. 999 | 0. 998 | 0. 599 | 0. 727 | 0. 901 | 0. 933 |
| site2 | – | 0. 977 | 0. 938 | 0. 818 | 0. 879 | 0. 882 | 0. 764 |
| site3 | – | – | 0. 955 | 0. 770 | 0. 703 | 0. 683 | 0. 660 |
| site4 | – | – | – | 0. 771 | 0. 726 | 0. 705 | 0. 682 |
| site5 | – | – | – | – | 0. 986 | 0. 985 | 0. 949 |
| site6 | – | – | – | – | – | 0. 993 | 0. 953 |
| site7 | – | – | – | – | – | – | 0. 955 |

**B**

| r²: | site2 | site3 | site4 | site5 | site6 | site7 | site8 |
|---|---|---|---|---|---|---|---|
| site1 | 0. 029 | 0. 064 | 0. 065 | 0. 084 | 0. 110 | 0. 160 | 0. 171 |
| site2 | – | 0. 038 | 0. 036 | 0. 089 | 0. 115 | 0. 123 | 0. 092 |
| site3 | – | – | 0. 897 | 0. 166 | 0. 155 | 0. 155 | 0. 145 |
| site4 | – | – | – | 0. 170 | 0. 168 | 0. 168 | 0. 158 |
| site5 | – | – | – | – | 0. 867 | 0. 817 | 0. 757 |
| site6 | – | – | – | – | – | 0. 932 | 0. 857 |
| site7 | – | – | – | – | – | – | 0. 911 |

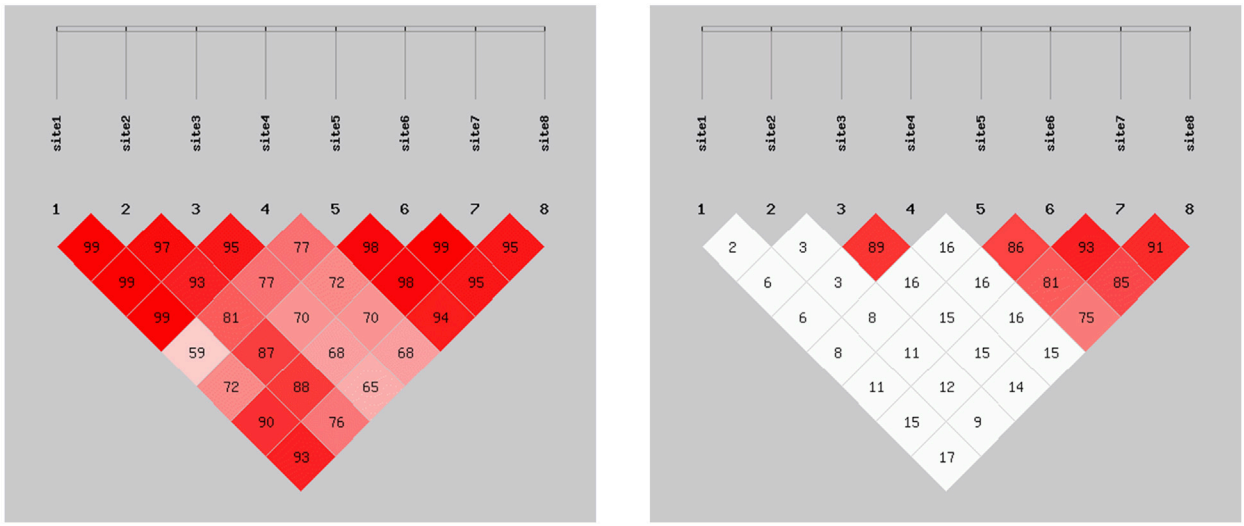


**FIGURE 4 |** LinKage disequilibrium analysis among 8 SNPs in the introns flanking of circTAF8. The linkage among SNPs was evaluated based on D′ and $r^2$. **(A)** D′ on the left. D′ = 1 indicates full linkage, D′ = 0 indicates no linkage or linkage equilibrium. **(B)** $r^2$ on the right. $r^2$ = 1 indicates full linkage, $r^2$ > 0.33 indicates strong linkage, $r^2$ = 0 indicates no linkage or linkage equilibrium.

**TABLE 2 |** Association of the combinations of TAF8 haplotypes with carcass traits.

| Group | Traits | Least square mean ± standard error | | | | | |
|---|---|---|---|---|---|---|---|
| 1 | | CCAC | CCCC | TCAC | TTAA | *p* value | |
| | LMW | 196.77 ± 29.13 | 195.92 ± 5.93$^{A}$ | 193.02 ± 7.72$^{B}$ | 257.328 ± 16.15$^{AB}$ | 0.024 | |
| 2 | | CCAATCAA | CCAATTAA | CCAGTCAG | CTAGTCAG | TTGGCCGG | *p* Value |
| | LWBS | 1,648.41 ± 32.97 | 1,612.11 ± 11.12 | 1,582.80 ± 20.66 | 1,598.61 ± 8.06$^{a}$ | 1,631.47 ± 10.08$^{a}$ | 0.02 |
| | HBW | 1,152.30 ± 46.55$^{aB}$ | 1,292.32 ± 15.19$^{a}$ | 1,271.24 ± 29.05 | 1,276.11 ± 10.61 | 1,314.13 ± 13.55$^{B}$ | 0.007 |
| | FBW | 1,073.51 ± 22.19 | 1,075.37 ± 7.48 | 1,057.68 ± 13.90 | 1,066.23 ± 5.43$^{A}$ | 1,092.37 ± 6.78$^{A}$ | 0.011 |
| | LMW | 198.65 ± 29.93 | 196.23 ± 9.68$^{A}$ | 271.25 ± 18.67$^{ABC}$ | 194.16 ± 6.68$^{B}$ | 199.32 ± 8.86$^{C}$ | 0.003 |

*[a]Group 1 indicates haplotype combination of g.-1554 T > C and g.-1480A > C.*

*[b]Group 2 indicates haplotype combination of g.-289 C > T, g.-288A > G, g.-210 T > C, and g.-173A > G.*

*The above values are "mean values ± standard errors"; in each group of SNPs, unmarked letters in the same line indicate that the difference is not significant (p > 0.05). When the letters are same, lowercase letters indicate significant differences (p < 0.05), and uppercase letters indicate significant differences (p < 0.01).*

Genome-wide association studies have identified millions of SNPs associated with complex growth traits in chickens (Muir et al., 2008). However, how these genetic variations are related to phenotypes is often unclear. Most variation sites are located in noncoding regions such as introns and intergenic regions (Zhao et al., 2003; Abdollahi-Arpanahi et al., 2016). Interestingly, there have been reports confirming that the inverted repeat elements, complementary sequences, and certain specific motifs in the flanking introns of circRNAs can regulate the circularization of the circRNAs (Zhang et al., 2014; Ivanov et al., 2015; Conn et al., 2015). In our study, by comparing the flanking intron sequences at both ends of circTAF8, we found multiple complementary sequences between the upstream and downstream introns, indicating that the complementary pairing of introns may directly or indirectly affect the formation of circTAF8. Recent studies have reported that the polymorphism of the introns flanking circRNAs might regulate the expression of circRNAs (Burd et al., 2010; Paraboschi et al., 2018; Zhou et al., 2020). CircANRIL is transcribed from the lncANRIL

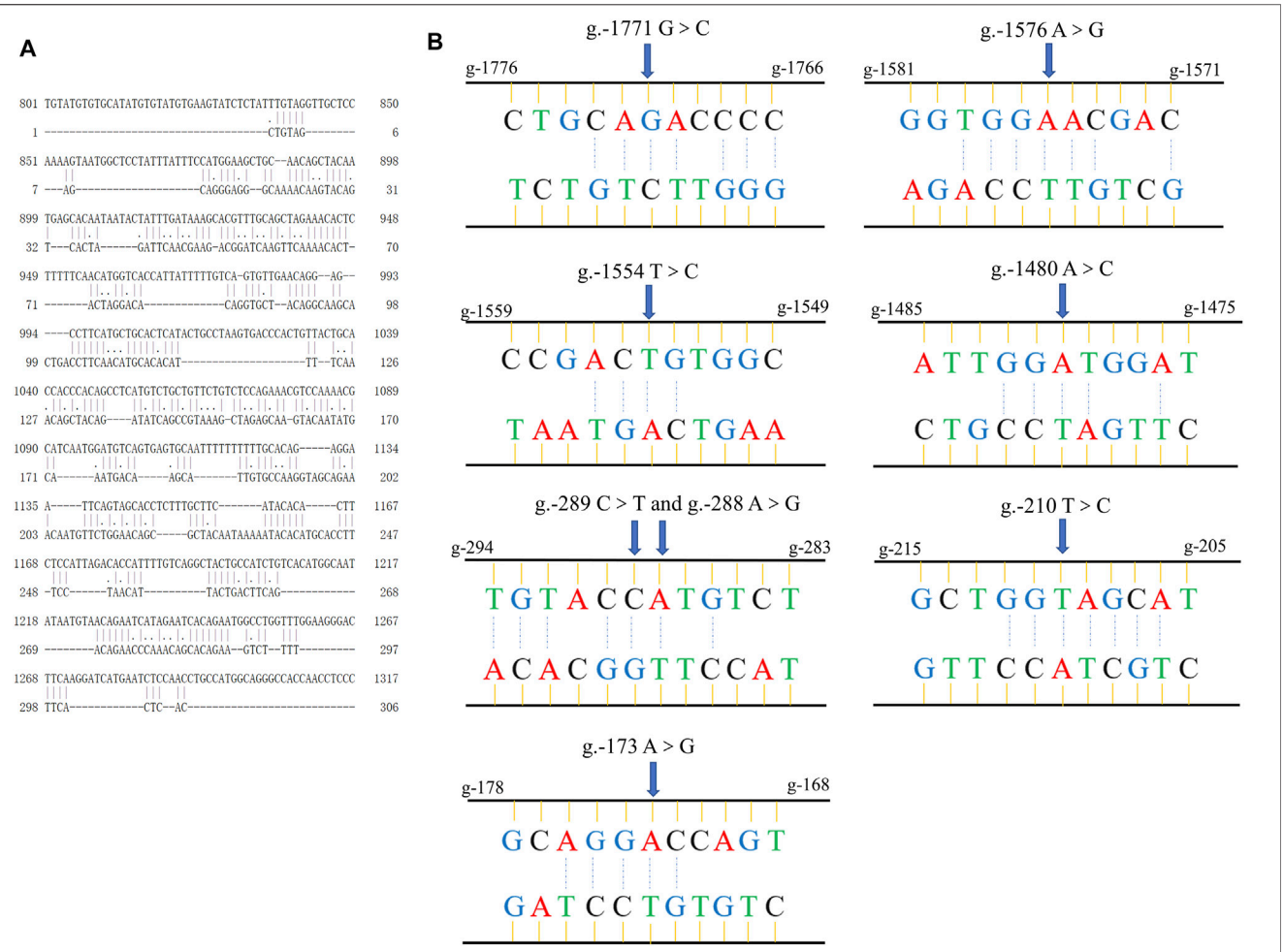


**FIGURE 5 |** Sequence alignment information. **(A)** The global alignment of introns flanking circTAF8. **(B)** The short reverse complementary sequences of SNPs in flanking introns of circTAF8.

gene, and the SNP within 200 bp of an ANRIL intron–exon boundary may contribute circANRIL expression, leading to risk of atherosclerotic vascular disease (Burd et al., 2010). The SNP site rs12196996 in the flanking intron of circFOXO3 is associated with the risk of coronary artery disease, due to its effect on the expression level of circFOXO3 (Zhou et al., 2020). At present, several research teams have provided databases on the correlation between circRNAs quantitative trait loci (circQTLs) and complex diseases (Ghosal et al., 2013; Liu et al., 2019; Gao et al., 2021), but research on the growth traits of animals has not been reported. As a layer of gene regulation network, the expression of circRNAs may be an intermediate phenotype that connects genetic variation and phenotypic change.

Given the importance of flanking introns in circRNA transcription level and phenotypic variation, we investigated the polymorphism of flanking introns of circTAF8. The association results showed that a total of eight SNPs were related to carcass traits, and four SNP sites (g.-289 C > T, g.-288A > G, g.-210 T > C, and g.-173A > G) are located within the 150-bp region of the 5′ upstream splice acceptor. These observations were consistent with previous studies, and most circQTL SNPs were close to the splicing sites (Ahmed et al., 2019; Liu et al., 2019). It is well known that skeletal muscle is the largest tissue organ in the body, accounting for about 40%–50% of the total body weight (Frontera and Ochala, 2015). Full-bore and leg muscle weight were closely related to muscle development. In this study, g.-289 C > T, g.-288A > G, g.-210 T > C, and g.-173A > G were significantly related to full-bore weight. Both g.-1480A > C and g.-1554 T > C were extremely significantly correlated with leg muscle weight. Haplotypes can usually provide more information than a single SNP site can because animal phenotypes can be affected by multiple mutations (Liu et al., 2008). The association of haplotype combinations was consistent with the results of the association analysis of a single SNP. It is worth noting that four SNPs close to the splice receptor are strongly linked to form a haplotype. This haplotype was significantly related to live weight, full-bore weight, half-bore, and leg muscle weight. In addition, we found that all eight SNP sites have short complementary sequences, indicating that the polymorphisms of these SNP sites may regulate the production of circTAF8. Based on the above results, TAF8 may serve a new SNP maker in chicken

genetics and breeding, and circTAF8 may contribute to understanding the potential regulatory mechanisms of this genetic trait.

Interestingly, the results of tissue expression profiling showed that the expression level of circTAF8 was significantly higher than that of the linear transcript in various tissues of chicken; circTAF8 seems to be the major transcript of pre-TAF8. Generally, the level of circRNAs is lower than those of corresponding parent genes (Guo et al., 2014). However, some circRNAs are expressed at high levels in some unique cell lines or tissues independently of their host genes (Salzman et al., 2013; Rybak-Wolf et al., 2015). TAF8 is a TBP-binding protein of the multi-subunit transcription factor TFIID. TFIID plays a crucial role in the binding of RNA polymerase II to transcription factors and core promoters (Trowitzsch et al., 2015). Research related to TAF8 mainly focuses on the TFIID assembly; relatively few studies focus on its function and polymorphism. Previous reports have shown that TAF8 can promote the differentiation of 3T3-L1 preadipocyte cells into adipocytes. However, it does not appear to play a role in the myogenesis of C2C12 cells (Guermah et al., 2003). Interestingly, our results showed that the expression level of the *TAF8* gene was the lowest in breast muscle and leg muscle. In addition, we found that the protein sequence of TAF8 was conserved entirely between chicken and mouse by the NCBI protein blast tool. CircRNAs are often closely related to the expression or function of their linear host genes. Various circRNAs and their host genes, such as circRBFOX2, circSVIL, circLMO7, circFGFR4, circFGFR2, and circTTN, all play a regulatory role in muscle development (Wei et al., 2017; Li et al., 2018a; Ouyang et al., 2018a; Ouyang et al., 2018b; Chen et al., 2018; Wang X. et al., 2019). However, it is worth noting that some parent genes of circRNAs related to skeletal muscle development have not been reported to be involved in muscle development, such as circTMTC1, circZNF609, circFUT10, and circHIPK3 (Li et al., 2018b; Wang Y. et al., 2019; Chen et al., 2019; Shen et al., 2019). These results suggested that pre-TAF8 is involved in regulating muscle development which might mainly be through this circular transcript. Although the function of TAF8 on chicken muscle development has not been explored, considering that circTAF8 is the main transcript, we speculate that the association effect of the flanking introns of circTAF8 with carcass traits might be mainly realized through the expression of circTAF8 in chicken. Whether circRNAs are a regulator between phenotypic traits and SNP needs to be confirmed *via* additional research.

## CONCLUSION

In summary, circTAF8 regulates skeletal muscle development by promoting myoblast proliferation and inhibiting myoblast differentiation. The SNPs in introns flanking circTAF8 are significantly correlated with multiple carcass traits, such as live weight, full and half-bore weight, and leg muscle weight. The association between SNPs and phenotypic traits may be achieved through the expression of circTAF8.

## 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 ethics committee of laboratory animals of South China Agricultural University.

## AUTHOR CONTRIBUTIONS

Conceptualization, methodology, KL, QN; software, WH; validation, KL, ZW, WH; formal analysis, ZW; investigation, resources, YC and DC; data curation, writing—original draft preparation, KL; writing—review and editing, ZW, QN; visualization, KL; supervision, QN; project administration, KL; funding acquisition, QN. All authors have read and agreed to the published version of the manuscript.

## FUNDING

This work was supported by the Natural Scientific Foundation of China (U1901206 and 31761143014), Local Innovative and Research Teams Project of Guangdong Province (2019BT02N630), Guangzhou Science and Technology Key Project (202103000084), and China Agriculture Research System (CARS-41-G03).

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2021.743757/full#supplementary-material

## REFERENCES

Abdollahi-Arpanahi, R., Morota, G., Valente, B. D., Kranis, A., Rosa, G. J. M., and Gianola, D. (2016). Differential Contribution of Genomic Regions to Marked Genetic Variation and Prediction of Quantitative Traits in Broiler Chickens. *Genet. Sel Evol.* 48 (1), 1–13. doi:10.1186/s12711-016-0187-z

Abo-Aziza, F. A. M., and A A, Z. (2017). The Impact of confluence on Bone Marrow Mesenchymal Stem (BMMSC) Proliferation and Osteogenic Differentiation. *Int. J. Hematol. Oncol. Stem Cell Res* 11 (2), 121–132. doi:10.1002/bem.20550

Ahmed, I., Karedath, T., Al-Dasim, F. M., and Malek, J. A. (2019). Identification of Human Genetic Variants Controlling Circular RNA Expression. *RNA* 25 (12), 1765–1778. doi:10.1261/rna.071654.119

Al-Musawi, S. L., Lock, F., Simbi, B. H., Bayol, S. A. M., and Stickland, N. C. (2011). Muscle Specific Differences in the Regulation of Myogenic Differentiation in Chickens Genetically Selected for Divergent Growth Rates. *Differentiation* 82 (3), 127–135. doi:10.1016/j.diff.2011.05.012

Burd, C. E., Jeck, W. R., Liu, Y., Sanoff, H. K., Wang, Z., Sharpless, N. E., et al. (2010). Expression of Linear and Novel Circular Forms of an INK4/ARF-Associated Non-coding RNA Correlates with Atherosclerosis Risk. *Plos Genet.* 6 (12), e1001233. doi:10.1371/journal.pgen.1001233

Cai, B., Li, Z., Ma, M., Wang, Z., Han, P., Abdalla, B. A., et al. (2017). LncRNA-Six1 Encodes a Micropeptide to Activate Six1 in Cis and Is Involved in Cell Proliferation and Muscle Growth. *Front. Physiol.* 8, 230. doi:10.3389/fphys.2017.00230

Cai, B., Li, Z., Ma, M., Zhang, J., Kong, S., Abdalla, B. A., et al. (2021). Long Noncoding RNA SMUL Suppresses SMURF2 Production-Mediated Muscle Atrophy via Nonsense-Mediated mRNA Decay. *Mol. Ther. - Nucleic Acids* 23, 512–526. doi:10.1016/j.omtn.2020.12.003

Chen, B., Yu, J., Guo, L., Byers, M., Wang, Z., Chen, X., et al. (2019). Circular RNA circHIPK3 Promotes the Proliferation and Differentiation of Chicken Myoblast Cells by Sponging miR-30a-3p. *Cells* 8 (2), 177. doi:10.3390/cells8020177

Chen, L.-L., and Yang, L. (2015). Regulation of circRNA Biogenesis. *Rna Biol.* 12 (4), 381–388. doi:10.1080/15476286.2015.1020271

Chen, M., Wei, X., Song, M., Jiang, R., Huang, K., Deng, Y., et al. (2021). Circular RNA circMYBPC1 Promotes Skeletal Muscle Differentiation by Targeting MyHC. *Mol. Ther. - Nucleic Acids* 24, 352–368. doi:10.1016/j.omtn.2021.03.004

Chen, X., Ouyang, H., Wang, Z., Chen, B., and Nie, Q. (2018). A Novel Circular RNA Generated by FGFR2 Gene Promotes Myoblast Proliferation and Differentiation by Sponging miR-133a-5p and miR-29b-1-5p. *Cells* 7 (11), 199. doi:10.3390/cells7110199

Conn, S. J., Pillman, K. A., Toubia, J., Conn, V. M., Salmanidis, M., Phillips, C. A., et al. (2015). The RNA Binding Protein Quaking Regulates Formation of circRNAs. *Cell* 160 (6), 1125–1134. doi:10.1016/j.cell.2015.02.014

Faul, F., Erdfelder, E., Buchner, A., and Lang, A.-G. (2009). Statistical Power Analyses Using G*Power 3.1: Tests for Correlation and Regression Analyses. *Behav. Res. Methods* 41 (4), 1149–1160. doi:10.3758/BRM.41.4.1149

Fortin, A., Wood, J. D., and Whelehan, O. P. (1987). Breed and Sex Effects on the Development, Distribution of Muscle, Fat and Bone, and the Partition of Fat in Pigs. *J. Agric. Sci.* 108 (1), 141–153. doi:10.1017/S0021859600064212

Frontera, W. R., and Ochala, J. (2015). Skeletal Muscle: a Brief Review of Structure and Function. *Calcif Tissue Int.* 96 (3), 183–195. doi:10.1007/s00223-014-9915-y

Gao, Y., Li, X., Shang, S., Guo, S., Wang, P., Sun, D., et al. (2021). LincSNP 3.0: an Updated Database for Linking Functional Variants to Human Long Non-coding RNAs, Circular RNAs and Their Regulatory Elements. *Nucleic Acids Res.* 49 (D1), D1244–D1250. doi:10.1093/nar/gkaa1037

Ghosal, S., Das, S., Sen, R., Basak, P., and Chakrabarti, J. (2013). Circ2Traits: a Comprehensive Database for Circular RNA Potentially Associated with Disease and Traits. *Front. Genet.* 4, 283. doi:10.3389/fgene.2013.00283

Gorbach, D. M., Fan, B., Onteru, S. K., Zhao, X., Du, Z.-Q., Garrick, D. J., et al. (2010). Genome-wide Association Studies for Important Economic Traits in Domestic Animals Using High Density SNP Genotyping. *Iowa State. Univ. Anim. Industry Rep.* 7 (1). doi:10.31274/ans_air-180814-980

Greco, S., Cardinali, B., Falcone, G. F., and Martelli, F. (2018). Circular RNAs in Muscle Function and Disease. *Ijms* 19, 3454. doi:10.3390/ijms19113454

Grefte, S., Kuijpers-Jagtman, A. M., Torensma, R., and Von den Hoff, J. W. (2007). Skeletal Muscle Development and Regeneration. *Stem Cell Dev.* 16 (5), 857–868. doi:10.1089/scd.2007.0058

Guermah, M., Ge, K., Chiang, C.-M., and Roeder, R. G. (2003). The TBN Protein, Which Is Essential for Early Embryonic Mouse Development, Is an Inducible TAFII Implicated in Adipogenesis. *Mol. Cell* 12 (4), 991–1001. doi:10.1016/S1097-2765(03)00396-4

Guo, J. U., Agarwal, V., Guo, H., and Bartel, D. P. (2014). Expanded Identification and Characterization of Mammalian Circular RNAs. *Genome Biol.* 15 (7), 1–14. doi:10.1186/s13059-014-0409-z

Halevy, O., Yahav, S., and Rozenboim, I. (2006). Enhancement of Meat Production by Environmental Manipulations in Embryo and Young Broilers. *World's Poult. Sci. J.* 62 (03), 485–497. doi:10.1017/S0043933906001103

Hao, Z., Zhou, H., Hickford, J. G. H., Gong, H., Wang, J., Hu, J., et al. (2020). Identification and Characterization of Circular RNA in Lactating Mammary Glands from Two Breeds of Sheep with Different Milk Production Profiles Using RNA-Seq. *Genomics* 112 (3), 2186–2193. doi:10.1016/j.ygeno.2019.12.014

Houba, P. H. J., Pas, M. F. W. T., and Pas, M. (2004). "The Muscle Regulatory Factors Gene Family in Relation to Meat Production," in *Muscle Development of Livestock Animals – Physiology, Genetics, and Meat Quality*. Editors M. F. W. te Pas, M. E. Everts, and H. P. Haagsman (Oxfordshire, UK: CABI Publishing), 201–223. doi:10.1079/9780851998114.0201

Huang, M., Shen, Y., Mao, H., Chen, L., Chen, J., Guo, X., et al. (2018). Circular RNA Expression Profiles in the Porcine Liver of Two Distinct Phenotype Pig Breeds. *Asian-australas J. Anim. Sci.* 31 (6), 812–819. doi:10.5713/ajas.17.0651

Ivanov, A., Memczak, S., Wyler, E., Torti, F., Porath, H. T., Orejuela, M. R., et al. (2015). Analysis of Intron Sequences Reveals Hallmarks of Circular RNA Biogenesis in Animals. *Cell Rep.* 10 (2), 170–177. doi:10.1016/j.celrep.2014.12.019

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

Kristensen, L. S., Hansen, T. B., Venø, M. T., and Kjems, J. (2018). Circular RNAs in Cancer: Opportunities and Challenges in the Field. *Oncogene* 37 (5), 555–565. doi:10.1038/onc.2017.361

Kyei, B., Li, L., Yang, L., Zhan, S., and Zhang, H. (2020). CDR1as/miRNAs-related Regulatory Mechanisms in Muscle Development and Diseases. *Gene* 730, 144315. doi:10.1016/j.gene.2019.144315

Li, B., Yin, D., Li, P., Zhang, Z., Zhang, X., Li, H., et al. (2020). Profiling and Functional Analysis of Circular RNAs in Porcine Fast and Slow Muscles. *Front Cell Dev Biol* 8, 322. doi:10.3389/fcell.2020.00322/

Li, H., Wei, X., Yang, J., Dong, D., Hao, D., Huang, Y., et al. (2018a). circFGFR4 Promotes Differentiation of Myoblasts via Binding miR-107 to Relieve its Inhibition of Wnt3a. *Mol. Ther. - Nucleic Acids* 11, 272–283. doi:10.1016/j.omtn.2018.02.012

Li, H., Yang, J., Wei, X., Song, C., Dong, D., Huang, Y., et al. (2018b). CircFUT10 Reduces Proliferation and Facilitates Differentiation of Myoblasts by Sponging miR-133a. *J. Cell Physiol* 233 (6), 4643–4651. doi:10.1002/jcp.26230

Liu, N., Zhang, K., and Zhao, H. (2008). Haplotype-Association Analysis. *Adv. Genet.* 60, 335–405. doi:10.1016/S0065-2660(07)00414-2

Liu, Z., Ran, Y., Tao, C., Li, S., Chen, J., and Yang, E. (2019). Detection of Circular RNA Expression and Related Quantitative Trait Loci in the Human Dorsolateral Prefrontal Cortex. *Genome Biol.* 20 (1), 1–16. doi:10.1186/s13059-019-1701-8

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 (4), 402–408. doi:10.1006/meth.2001.1262

Luo, W., Nie, Q., and Zhang, X. (2013). MicroRNAs Involved in Skeletal Muscle Differentiation. *J. Genet. Genomics* 40 (3), 107–116. doi:10.1016/j.jgg.2013.02.002

Madeira, F., Park, Y. m., Lee, J., Buso, N., Gur, T., Madhusoodanan, N., et al. (2019). The EMBL-EBI Search and Sequence Analysis Tools APIs in 2019. *Nucleic Acids Res.* 47 (W1), W636–W641. doi:10.1093/nar/gkz268

Muir, W. M., Wong, G. K.-S., Zhang, Y., Wang, J., Groenen, M. A. M., Crooijmans, R. P. M. A., et al. (2008). Genome-wide Assessment of Worldwide Chicken SNP Genetic Diversity Indicates Significant Absence of Rare Alleles in Commercial Breeds. *Proc. Natl. Acad. Sci.* 105 (45), 17312–17317. doi:10.1073/pnas.0806569105

Niknafs, S., Javaremi, A. N., and Sadeghi, M. (2014). Single Nucleotide Polymorphisms in BMPR-IB and STAT5B Genes and Their Association with Growth and Reproductive Traits in Chicken. *Songklanakarin J. Sci. Tech.* 36 (2), 137–142.

Ouyang, H., Chen, X., Li, W., Li, Z., Nie, Q., and Zhang, X. (2018a). Circular RNA circSVIL Promotes Myoblast Proliferation and Differentiation by Sponging miR-203 in Chicken. *Front. Genet.* 9, 172. doi:10.3389/fgene.2018.00172

Ouyang, H., Chen, X., Wang, Z., Yu, J., Jia, X., Li, Z., et al. (2018b). Circular RNAs Are Abundant and Dynamically Expressed during Embryonic Muscle Development in Chickens. *Dna Res.* 25 (1), 71–86. doi:10.1093/dnares/dsx039

Ouyang, H., He, X., Li, G., Xu, H., Jia, X., Nie, Q., et al. (2015). Deep Sequencing Analysis of miRNA Expression in Breast Muscle of Fast-Growing and Slow-Growing Broilers. *Ijms* 16 (7), 16242–16262. doi:10.3390/ijms160716242

Ouyang, J. H., Xie, L., Nie, Q., Luo, C., Liang, Y., Zeng, H., et al. (2008). Single Nucleotide Polymorphism (SNP) at theGHRgene and its Associations with Chicken Growth and Fat Deposition Traits. *Br. Poult. Sci.* 49 (2), 87–95. doi:10.1080/00071660801938817

Paraboschi, E. M., Cardamone, G., Soldà, G., Duga, S., and Asselta, R. (2018). Interpreting Non-coding Genetic Variation in Multiple Sclerosis Genome-wide Associated Regions. *Front. Genet.* 9, 647. doi:10.3389/fgene.2018.00647

Pas, M. F. W., and Visscher, A. H. (1994). Genetic Regulation of Meat Production by Embryonic Muscle Formation - a Review. *J. Anim. Breed. Genet.* 111 (1-6), 404–412. doi:10.1111/j.1439-0388.1994.tb00477.x

Rybak-Wolf, A., Stottmeister, C., Glažar, P., Jens, M., Pino, N., Giusti, S., et al. (2015). Circular RNAs in the Mammalian Brain Are Highly Abundant, Conserved, and Dynamically Expressed. *Mol. Cell* 58 (5), 870–885. doi:10.1016/j.molcel.2015.03.027

Salzman, J., Chen, R. E., Olsen, M. N., Wang, P. L., and Brown, P. O. (2013). Cell-type Specific Features of Circular RNA Expression. *Plos Genet.* 9 (9), e1003777. doi:10.1371/journal.pgen.1003777

Scheuermann, G. N., Bilgili, S. F., Tuzun, S., and Mulvaney, D. R. (2004). Comparison of Chicken Genotypes: Myofiber Number in Pectoralis Muscle and Myostatin Ontogeny. *Poult. Sci.* 83 (8), 1404–1412. doi:10.1093/ps/83.8.1404

Schneider, C. A., Rasband, W. S., and Eliceiri, K. W. (2012). NIH Image to ImageJ: 25 Years of Image Analysis. *Nat. Methods* 9 (7), 671–675. doi:10.1038/nmeth.2089

Shen, X., Liu, Z., Cao, X., 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

Tanaka, K., Sato, K., Yoshida, T., Fukuda, T., Hanamura, K., Kojima, N., et al. (2011). Evidence for Cell Density Affecting C2C12 Myogenesis: Possible Regulation of Myogenesis by Cell-Cell Communication. *Muscle Nerve* 44 (6), 968–977. doi:10.1002/mus.22224

te Pas, M. F. (2004). *Muscle Development of Livestock Animals Physiology Genetics and Meat Quality*. Oxfordshire, UK: CABI Publishing.

Trowitzsch, S., Viola, C., Scheer, E., Conic, S., Chavant, V., Fournier, M., et al. (2015). Cytoplasmic TAF2-TAF8-TAF10 Complex Provides Evidence for Nuclear Holo-TFIID Assembly from Preformed Submodules. *Nat. Commun.* 6 (1), 1–14. doi:10.1038/ncomms7011

Wang, P. L., Bao, Y., Yee, M.-C., Barrett, S. P., Hogan, G. J., Olsen, M. N., et al. (2014). Circular RNA Is Expressed across the Eukaryotic Tree of Life. *Plos One* 9 (3), e90859. doi:10.1371/journal.pone.0090859

Wang, X., Cao, X., Dong, D., Shen, X., Cheng, J., Jiang, R., et al. (2019a). Circular RNA TTN Acts as a miR-432 Sponge to Facilitate Proliferation and Differentiation of Myoblasts via the IGF2/PI3K/AKT Signaling Pathway. *Mol. Ther. - Nucleic Acids* 18, 966–980. doi:10.1016/j.omtn.2019.10.019

Wang, Y., Li, M., Wang, Y., Liu, J., Zhang, M., Fang, X., et al. (2019b). A Zfp609 Circular RNA Regulates Myoblast Differentiation by Sponging miR-194-5p. *Int. J. Biol. Macromolecules* 121, 1308–1313. doi:10.1016/j.ijbiomac.2018.09.039

Wei, X., Li, H., Yang, J., Hao, D., Dong, D., Huang, Y., et al. (2017). Circular RNA Profiling Reveals an Abundant circLMO7 that Regulates Myoblasts Differentiation and Survival by Sponging miR-378a-3p. *Cell Death Dis* 8 (10), e3153. doi:10.1038/cddis.2017.541

Xu, K., Chen, D., Wang, Z., Ma, J., Zhou, J., Chen, N., et al. (2018). Annotation and Functional Clustering of circRNA Expression in Rhesus Macaque Brain during Aging. *Cell Discov* 4 (1), 1–18. doi:10.1038/s41421-018-0050-1

Yong, Y., and He, L. (2005). SHEsis, a Powerful Software Platform for Analyses of Linkage Disequilibrium, Haplotype Construction, and Genetic Association at Polymorphism Loci. *Cell Res* 15 (2), 97–98. doi:10.1038/sj.cr.7290272

Zhang, X.-O., Wang, H.-B., Zhang, Y., Lu, X., Chen, L.-L., and Yang, L. (2014). Complementary Sequence-Mediated Exon Circularization. *Cell* 159 (1), 134–147. doi:10.1016/j.cell.2014.09.001

Zhang, Z., Zhong, H., Lin, S., Liang, L., Ye, S., Xu, Z., et al. (2021). Polymorphisms of AMY1A Gene and Their Association with Growth, Carcass Traits and Feed Intake Efficiency in Chickens. *Genomics* 113 (2), 583–594. doi:10.1016/j.ygeno.2020.10.041

Zhao, Z., Fu, Y.-X., Hewett-Emmett, D., and Boerwinkle, E. (2003). Investigating Single Nucleotide Polymorphism (SNP) Density in the Human Genome and its Implications for Molecular Evolution. *Gene* 312, 207–213. doi:10.1016/S0378-1119(03)00670-X

Zhou, Y.-L., Wu, W.-P., Cheng, J., Liang, L.-L., Cen, J.-M., Chen, C., et al. (2020). CircFOXO3 Rs12196996, a Polymorphism at the Gene Flanking Intron, Is Associated with circFOXO3 Levels and the Risk of Coronary Artery Disease. *Aging* 12 (13), 13076–13089. doi:10.18632/aging.103398

**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.

# Advantages of publishing in Frontiers

**OPEN ACCESS**
Articles are free to read for greatest visibility and readership

**FAST PUBLICATION**
Around 90 days from submission to decision

**HIGH QUALITY PEER-REVIEW**
Rigorous, collaborative, and constructive peer-review

**TRANSPARENT PEER-REVIEW**
Editors and reviewers acknowledged by name on published articles

**Frontiers**
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

**Visit us:** www.frontiersin.org
**Contact us:** frontiersin.org/about/contact

**REPRODUCIBILITY OF RESEARCH**
Support open data and methods to enhance research reproducibility

**DIGITAL PUBLISHING**
Articles designed for optimal readership across devices

**FOLLOW US**
@frontiersin

**IMPACT METRICS**
Advanced article metrics track visibility across digital media

**EXTENSIVE PROMOTION**
Marketing and promotion of impactful research

**LOOP RESEARCH NETWORK**
Our network increases your article's readership