frontiers | Research Topics

Advances in transcriptomics research and their applications in human diseases

Edited by
Hua Zhong, Duo Liu, Shuai Liu and Yuan Liu

Published in
Frontiers in Genetics

ISSN 1664-8714
ISBN 978-2-8325-4418-1
DOI 10.3389/978-2-8325-4418-1

About Frontiers

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

Frontiers journal series

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

Dedication to quality

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

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the *Frontiers journals series*: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area.

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers editorial office: frontiersin.org/about/contact

Advances in transcriptomics research and their applications in human diseases

Topic editors

Hua Zhong — University of Hawaii at Manoa, United States
Duo Liu — Harbin Medical University Cancer Hospital, China
Shuai Liu — University of Hawaii at Manoa, United States
Yuan Liu — Shanghai Jiao Tong University, China

Citation

Zhong, H., Liu, D., Liu, S., Liu, Y., eds. (2024). *Advances in transcriptomics research and their applications in human diseases*. Lausanne: Frontiers Media SA.
doi: 10.3389/978-2-8325-4418-1

Table of contents

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 13 September 2022
DOI 10.3389/fgene.2022.996310

OPEN ACCESS

EDITED BY
Shuai Liu,
University of Hawaii at Manoa,
United States

REVIEWED BY
Dirk Geerts,
University of Amsterdam, Netherlands
Jin Zhang,
I.M. Sechenov First Moscow State Medical University, Russia
Wenjun Wang,
Stanford University, United States

*CORRESPONDENCE
Yonghong Lin,
linyhcd2011@163.com
Zhenglin Yang,
zliny@yahoo.com

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

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

RECEIVED 17 July 2022
ACCEPTED 18 August 2022
PUBLISHED 13 September 2022

CITATION
Yu X, He L, Chen Y, Lin W, Liu H, Yang X, Ye Y, Zheng X, Yang Z and Lin Y (2022), Construction of a focal adhesion signaling pathway-related ceRNA network in pelvic organ prolapse by transcriptome analysis.
Front. Genet. 13:996310.
doi: 10.3389/fgene.2022.996310

Construction of a focal adhesion signaling pathway-related ceRNA network in pelvic organ prolapse by transcriptome analysis

Xia Yu[1†], Li He[2†], Ying Chen[2], Wenyi Lin[3], Hong Liu[4], Xiu Yang[4], Ying Ye[4], Xuemei Zheng[2], Zhenglin Yang[5*] and Yonghong Lin[2*]

[1]Department of Clinical Laboratory, Chengdu Women's and Children's Central Hospital, Sichuan Provincial People's Hospital, School of Medicine, University of Electronic Science and Technology of China, Chengdu, Sichuan, China, [2]Department of Obstetrics and Gynecology, Chengdu Women's and Children's Central Hospital, School of Medicine, University of Electronic Science and Technology of China, Chengdu, Sichuan, China, [3]Department of Medical Pathology, Chengdu Women's and Children's Central Hospital, School of Medicine, University of Electronic Science and Technology of China, Chengdu, Sichuan, China, [4]Department of Surgical, Chengdu Women's and Children's Central Hospital, School of Medicine, University of Electronic Science and Technology of China, Chengdu, Sichuan, China, [5]Sichuan Provincial Key Laboratory for Human Disease Gene Study and Institute of Laboratory Medicine, Sichuan Provincial People's Hospital, University of Electronic Science and Technology of China, Chengdu, China

Objective: Pelvic organ prolapse (POP) affects a large proportion of adult women, but the pathogenesis of POP remains unclear. The increase in global population aging will impose a substantial medical burden. Herein, we aimed to explore the related RNAs regulating the occurrence of POP and provide potential therapeutic targets.

Method: Tissue biopsies were collected from the anterior vaginal wall of six women with POP and six matched subjects without POP. The profiles of mRNAs, circRNAs, lncRNAs, and miRNAs were obtained by whole transcriptome RNA sequencing.

Result: The findings revealed that 71 circRNAs, 76 known lncRNAs, 84 miRNAs, and 931 mRNAs were significantly altered ($p < 0.05$ and $|log2FC| > 1$). GO and KEGG enrichment analyses indicated that the differentially expressed genes (DEGs) were mainly enriched in the focal adhesion signaling pathway. *FLT, ITGA9, VEGFD, PPP1R12B*, and *ROCK2* were identified as focal adhesion signaling pathway-related hub genes by protein–protein interaction network analysis. Based on the relationships between the DEGs and miRNA, lncRNA and circRNA targets, we constructed a focal adhesion signaling pathway-related ceRNA network. The ceRNA network includes hsa_circ_0002190/hsa_circ_0046843/lnc-CARMN -miR-23a-3p - ROCK2 and hsa_circ_0001326/hsa_circ_0007733/lnc-AC107959/lnc-TPM1-AS - miR-205-5p - ROCK2/PPP1R12B/VEGFD. Moreover, abnormalities in the cytoskeleton in fibroblasts from individuals with POP were observed.

Conclusion: In this study, a focal adhesion signaling pathway-related ceRNA network was constructed, and this network may serve as a target for finding suitable drugs for the treatment of POP.

KEYWORDS

pelvic organ prolapse, whole transcriptome RNA sequencing, competing endogenous RNA, PPI, focal adhesion signaling pathway

Introduction

Pelvic organ prolapse (POP) is a weakening of the pelvic floor support structure that causes the pelvic internal organs, including the bladder, rectum, and uterus, to localize into or outside the vagina, which results in a decline in quality of life and serious psychosocial problems (Collins et al., 2021; Harvey et al., 2021). According to epidemiological data, the prevalence of POP in women over the age of 60 years is as high as 50%, and the number of POP cases with clinical symptoms is expected to increase by more than 46% by 2050 (Wu et al., 2011). Anterior vaginal prolapse (AVP) is the most common form of POP. Several studies have found that vaginal wall weakness can be considered a possible cause of prolapse (Li et al., 2021). Surgery is the main treatment. However, because the pathogenesis of POP has not been fully elucidated, a possibility of reprolapse after surgery remains (Nussler et al., 2022), which results in certain difficulties in the treatment of POP.

In recent years, transcriptomics has become an important tool for exploring the pathogenesis of diseases (Wu et al., 2021). Song et al. performed RNA-seq analysis to identify the POP signatures of 81 genes in uterosacral ligament (USL) samples as well as a number of extracellular matrix (ECM)-related genes (Xie et al., 2016); but the study only examined differentially expressed (DE) genes (DEGs). In addition, many studies have examined the role of microRNAs (miRNAs) in the pathogenesis of POP. The overexpression of miRNA-30d and miRNA-181a regulates the expression of HOXA11 and collagen (Lin et al., 2020), and miRNA-92 expression may be associated with reduced estrogen receptor β1 mRNA levels in the USL of women with POP (He et al., 2016). Moreover, miR-19-3p targets IGF-1 to promote autophagy and apoptosis through the AKT/mTOR/p70s6k pathway in vaginal fibroblasts in POP (Yin et al., 2021). However, the roles of long noncoding RNAs (lncRNAs) and circular RNAs (circRNAs) in the pathogenesis of POP have not been reported. Increasing evidence shows that competing endogenous RNA (ceRNA) networks play important roles in many disease processes (Zhao et al., 2021; Liu et al., 2022). The potential roles of the circRNA/lncRNA–miRNA–mRNA ceRNA network in the pathogenesis of POP remain unclear and have not been characterized.

Prolonged stretching and mechanical stress cause progressive deterioration of the pelvic organ support (Deng et al., 2021). Four mechanoresponsive genes are responsible for the regulation of actin cytoskeleton remodeling in fibroblasts under stretching, and investigations of the cytoskeleton phenotype in POP samples revealed an abnormal F-actin configuration (Ewies et al., 2008; Wang et al., 2015). Focal adhesion is the cell-extracellular matrix contact point that bundles actin filaments and is linked to transmembrane receptors of the integrin family *via* a multimolecular complex of junctional plaque proteins. Some focal adhesion constituents are structurally involved in the link between membrane receptors and the actin cytoskeleton, and others are signaling molecules, such as various protein kinases and phosphatases, their substrates, and adaptor proteins. These signaling events result in actin cytoskeleton reorganization. As a result, the development of a focal adhesion signaling-related ceRNA network is critical for understanding the mechanism of cytoskeletal remodeling in POP.

In this study, the schematic workflow is shown in Figure 1. We designed a whole transcriptome RNA sequencing study to uncover the profiles of DE mRNAs, DEGs and ncRNAs between a POP group and a control group. Subsequently, the DEGs were evaluated by Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses to reveal the associated functions. Furthermore, a focal adhesion signaling pathway-related ceRNA network was constructed to explore the underlying molecular mechanisms of the ncRNAs. This study may serve as a target for finding suitable drugs to treat POP.

Materials and methods

Tissue sample collection

The study included 12 patients who underwent complete hysterectomy at Chengdu Women's and Children's Central Hospital's Department of Obstetrics and Gynecology. The patients were screened using the following inclusion criteria: diagnosis of POP at stage 3 or higher according to the pelvic organ prolapse quantification (POP-Q) exam; patients with benign diseases (uterus myoma, adenomyosis) who did not have POP. A patient was excluded from the study if any of the following criteria applied: collagen metabolic disease or a history of hormone therapy (including vaginal estrogen therapy). Six of these patients were enrolled in the POP group, and six patients were enrolled in the control group. There were no significant differences in clinical characteristics such as age, BMI, menopausal status, gravidity, or parity between the two groups (Supplementary Table S1). The Ethics Committee of Chengdu Women's and Children's Central Hospital reviewed and approved this study (serial number: 2021-65), and all subjects signed informed consent forms.

After hysterectomy, 1.0 × 1.0-cm full-thickness vaginal wall tissue biopsies were obtained from the pericervical region of the

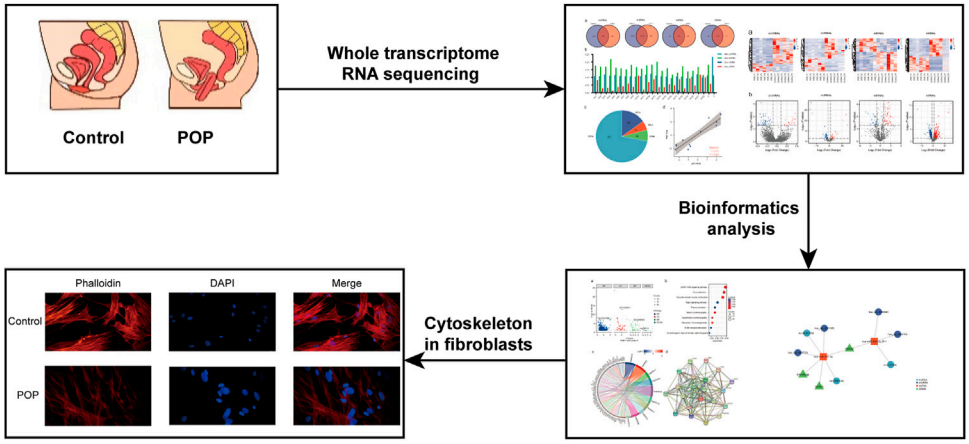


FIGURE 1
Research strategy employed in the present study.

anterior vaginal cuff in the controls, and from the prolapsed vaginal wall in the POP samples. Fresh tissue samples were collected aseptically, immediately frozen in liquid nitrogen, and stored at−80°C until analysis.

Total RNA extraction

Total RNA was extracted using TRIzol reagent (Invitrogen, CA, United States) according to the manufacturer's instructions. The RNA quantity and purity were determined using a NanoDrop ND-1000 (NanoDrop, Wilmington, DE, United States) and Bioanalyzer 2100 (Agilent, CA, United States). A small RNA library was prepared using approximately 1 μg of total RNA according to the TruSeq Small RNA Sample Prep Kit protocol (Illumina, San Diego, CA, United States). We then performed single-end sequencing (36 or 50 bp) on an Illumina HiSeq 2500 system (Illumina, San Diego, CA, United States) at LC-Bio (Hangzhou, China).

Ribosomal RNA was removed from approximately 2 μg of total RNA using the Epicenter Ribo-Zero Gold Kit (Illumina, San Diego, CA, United States) according to the manufacturer's instructions. The purified ribosomal RNA was fragmented into small pieces using divalent cations at elevated temperatures. The cleaved RNA fragments were then reverse-transcribed to generate the final cDNA library in accordance with the mRNA-Seq sample preparation kit protocol (Invitrogen, CA, United States). The final cDNA library had an average insert size of 300 ± 50 bp. We then performed 2 × 150-bp paired-end sequencing (PE150) using an Illumina NovaSeq™ 6000 system (LC-Bio Technology Co., Ltd., Hangzhou, China).

Bioinformatics analysis

As shown in Figure 2A, the raw data were cleaned and verified by fastp (Chen et al., 2018), and the remaining reads were mapped to the *Homo sapiens* GRCh38 genome using Bowtie2 (Langmead and Salzberg, 2012) and Tophat2 (Kim et al., 2013) and assembled using StringTie (Pertea et al., 2015). Additionally, StringTie was used to estimate the expression levels of all transcripts. The transcripts were then annotated with known mRNAs, known lncRNAs and transcripts shorter than 200 nt were discarded, and transcripts with coding potential were predicted using CPC (Kong et al., 2007) and CNCI (Sun et al., 2013). FPKM (FPKM = [total_exon_fragments/mapped_reads (millions) × exon_length (kB)]) values showed the expression levelsof mRNAs and lncRNAs, |log2FC|>1 and $p < 0.05$ obtained using a nonpaired test comparing nested linear models were identified as DE mRNAs and lncRNAs. TopHat fusion (Kim and Salzberg, 2011) and CIRCExplorer (Zhang et al., 2014; Zhang et al., 2016) were used to identify the circRNAs. The expression of circRNAs was calculated by SRPBM = (number of back-spliced junction reads)/(number of mapped reads) × 1,000,000,000. $p < 0.05$ obtained using the R package DESeq was regarded as indicating differential expression (Anders and Huber, 2010).

The process of miRNA identification is shown in Figure 2B. A BLAST search was performed to map unique sequences of 18–26 nucleotides in length to specific species precursors in miRBase 22.0 and therefore identify known miRNAs. RNAfold software (http://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAfold.cgi) was used to predict the secondary structures of all the obtained miRNAs. A nonpaired analysis was then conducted to calculate the

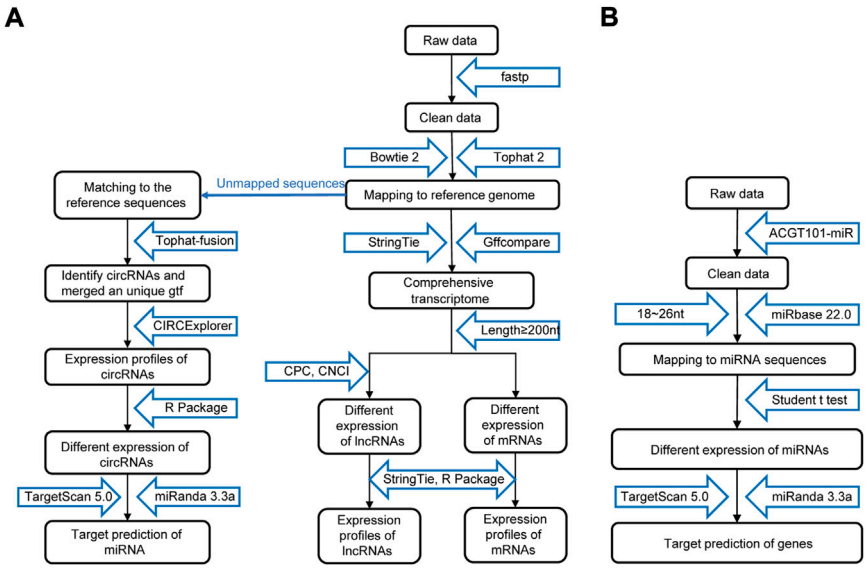


FIGURE 2
Procedure for preparing and analyzing an RNA library. **(A)** Process of preparing and analyzing circRNA, lncRNA, and mRNA libraries. **(B)** Process of preparing and analyzing miRNA libraries. CPC, coding potential calculator; CNCI, coding-noncoding index.

differential expression of miRNAs based on normalized deep-sequencing counts, and a significance threshold of 0.05 was chosen.

RNA functional enrichment analysis

The functions of DEGs were determined through GO and KEGG analyses using the R package (v3.6.3), and $p < 0.05$ indicated a significant difference.

Protein-protein interaction network

STRING (https://www.string-db.org) was used to build the PPI network based on the DEGs involved in the selected KEGG pathways between the POP and control groups. The extraction cutoff was set to an interaction score >0.4.

CeRNA network

TargetScan and miRanda were used to examine the targeting relationships between DEGs, DE miRNAs (DEMs), DE lncRNAs (DELs), and DE circRNAs (DECs). The circRNA/lncRNA–miRNA-mRNA ceRNA network was visualized using R package v3.6.3 and hub prognostic genes as the core. In the ceRNA network, this network displayed competitive binding.

Quantitative real-time PCR analysis

qRT–PCR analyses were performed using gene-specific primers (Supplementary Table S2). Total RNA was extracted using TRIzol reagent (TaKaRa, Kusatsu, Japan), and qRT–PCR was performed with an Applied Biosystems StepOnePlus Real-Time PCR system (Applied Biosystems, Foster City, CA, United States) using TB Green™ Premix EX Taq™ II (TaKaRa, Kusatsu, Japan) according to the manufacturer's instructions. Each reaction was conducted in triplicate, and the relative fold change values were measured in terms of the threshold cycle (Ct) and calculated using the formula $2^{-\Delta\Delta Ct}$ (Pfaffl, 2001).

Cell culture and identification

Fibroblasts were isolated from the anterior vaginal wall. Tissues removed during hysterectomy were washed with phosphate-buffered saline (PBS) containing 1% antibiotics, cut into slices with a diameter of less than 0.1 cm, and evenly distributed in the bottom of a 25-cm bottle. To avoid floating tissue blocks, upside-down containers were used in this study, and the bottle was carefully turned over after 4–6 h. The medium contained 15% FBS (Gibco, Waltham, MA, United States), 1% penicillin/streptomycin (Gibco, Waltham, MA, United States), and DMEM (Gibco, Waltham, MA, United States). The cells were passaged

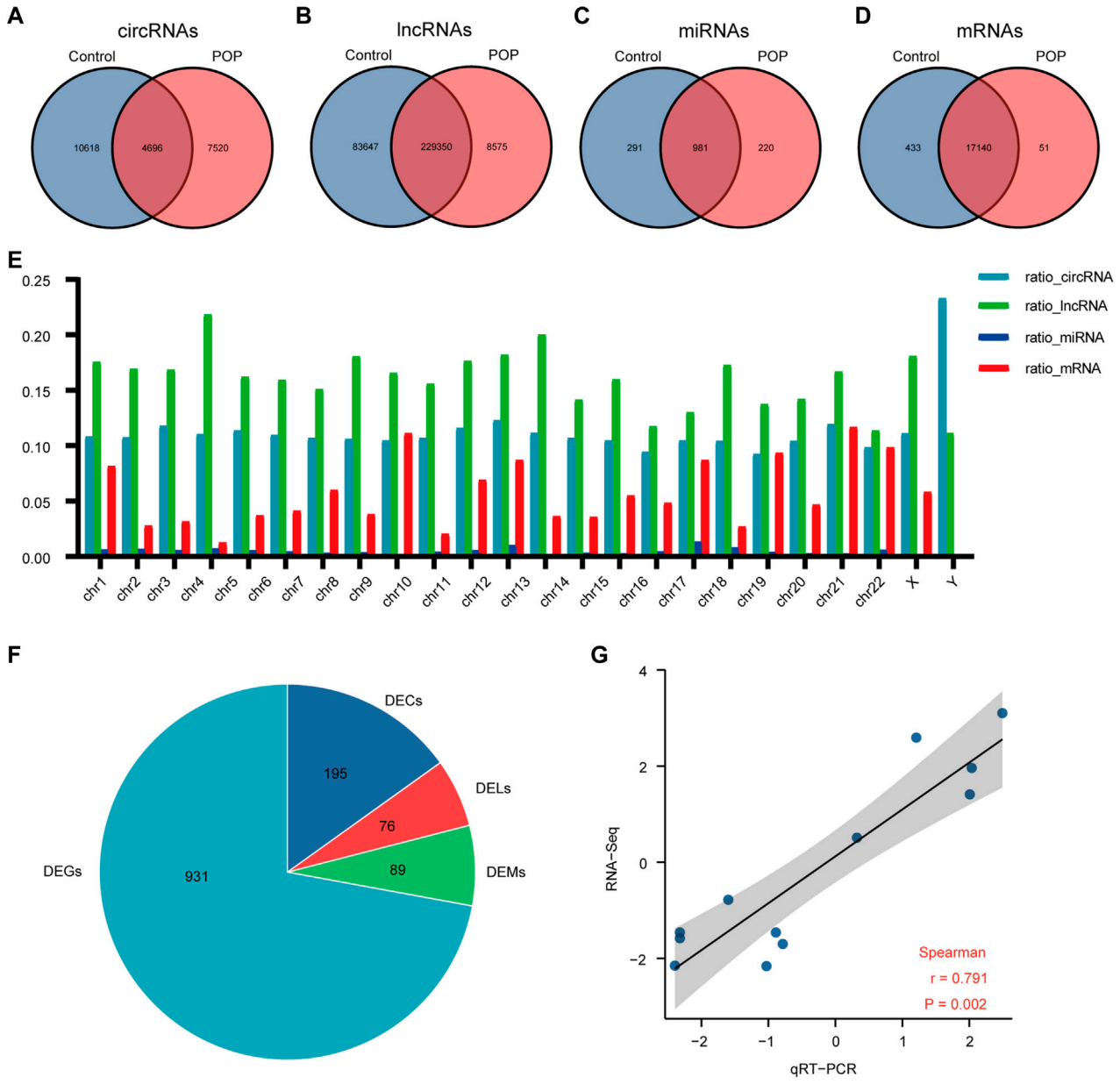


FIGURE 3
Whole transcriptome RNA sequencing results comparing the POP and control groups. **(A–D)** Venn diagrams depicting the profiles of circRNAs, lncRNAs, miRNAs, and mRNAs. **(E)** RNA expression ratio in a chromosome (number of RNAs in a related chromosome divided by the total number of RNAs). **(F)** Differentially expressed circRNAs, lncRNAs, miRNAs and mRNAs between the POP and control groups. **(G)** qRT–PCR validation of the RNA sequencing results using 3 DECs, 3 DELs, 3 DEMs, and 3 DEGs chosen at random. The log2-fold change from the q-RT–PCR analysis is plotted on the X-axis, and the log2-fold change from the RNA sequencing analysis is plotted on the Y-axis. POP, pelvic organ prolapse; DECs, differentially expressed circRNAs; DELs, differentially expressed lncRNAs; DEMs, differentially expressed miRNAs; DEGs, differentially expressed genes.

after approximately 15 days. The derived cells were identified by positive immunofluorescence staining for vimentin (1: 200; Bioss, Shanghai, China) and negative immunofluorescence staining for pancytokeratin (1:200; Bioss, Shanghai, China) (Supplementary Figure S1).

Fluorescence staining of F-actin

Microscopy was used to image the fluorescently labeled F-actin network. Cells were grown on a glass coverslip in 500 μl of cell culture media to a density of 3.0×10^4 cells/ml.

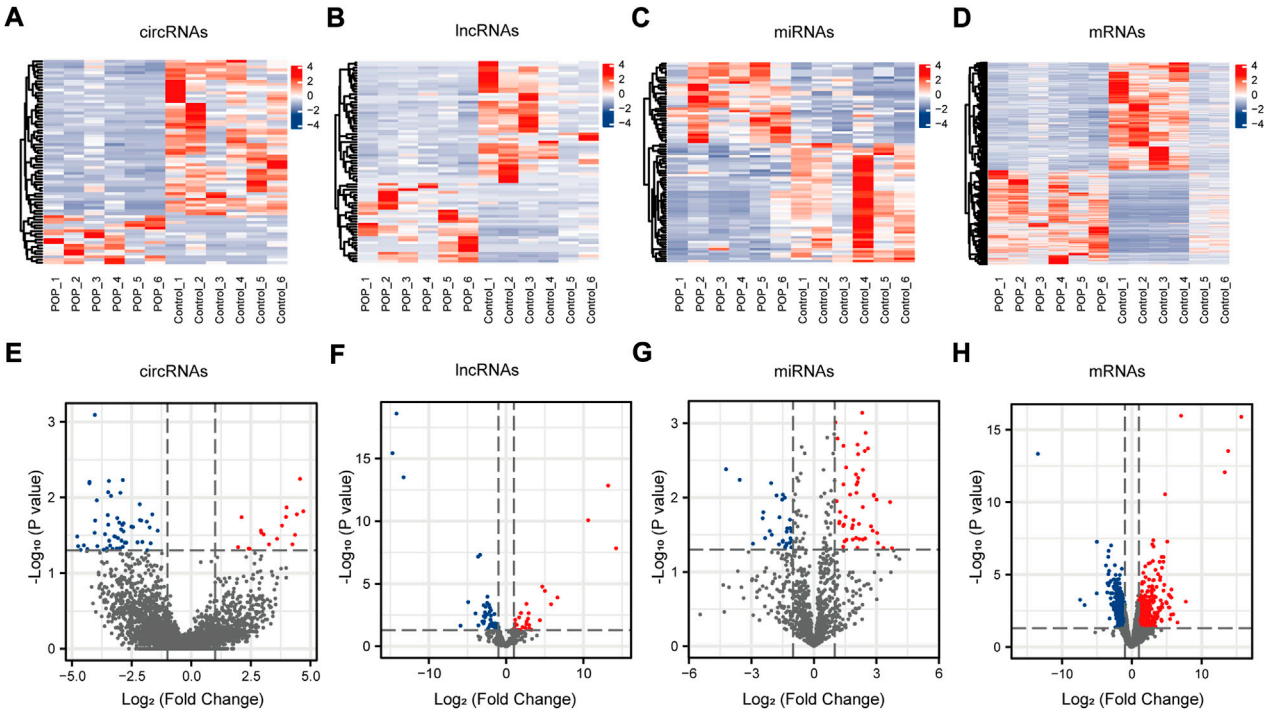


FIGURE 4
circRNA, lncRNA, miRNA, and mRNA expression profiles. **(A–D)** Heatmaps of DECs, DELs, DEMs, and DEGs in the anterior vaginal wall in the POP and control groups. **(E–H)** Volcano plots of DECs, DELs, DEMs, and DEGs in the anterior vaginal wall in the POP and control groups. POP, pelvic organ prolapse; circRNA, circular RNA; lncRNAs, long noncoding RNAs; miRNA, microRNA; DECs, differentially expressed circRNAs; DELs, differentially expressed lncRNAs; DEMs, differentially expressed miRNAs; DEGs, differentially expressed genes.

The cells were incubated at 37°C for 48 h and then stained with TRITC phalloidin to reveal the actin cytoskeleton. The cells were fixed with 4% paraformaldehyde for 10 min, washed three times with PBS, and incubated with 0.5% Triton-X for 5 min. The cells were blocked for 30 min at room temperature in 1% BSA and then incubated for 30 min with TRITC phalloidin (1:250; Solarbio, Beijing, China) in 3% BSA. The nuclei were sealed and counterstained with DAPI. All photographs were taken with a fluorescence microscope (Shunyu, Ningbo, China) at an excitation wavelength of 540–546 nm.

Results

Differences in transcription files

The expression profiles of circRNAs, lncRNAs, and mRNAs were determined using an Illumina NovaSeq™ 6000 system, and miRNAs were identified using an Illumina HiSeq 2500 system. After the removal of low-quality reads, the RNA-seq results revealed 22,834 circRNAs, 321,572 lncRNAs, 1,492 miRNAs, and 17,624 mRNAs (Figures 3A–D). Figure 3E depicts the chromosomal locations. Compared with the control group, 71 circRNAs, 76 known lncRNAs, 84 miRNAs, and 931 mRNAs were significantly altered in the POP group ($p < 0.05$ and $|\log2FC| > 1$) (Figure 3F). Upregulated expression of 17 circRNAs, 30 known lncRNAs, 53 miRNAs, and 433 mRNAs and downregulated expression of 54 circRNAs, 46 known lncRNAs, 34 miRNAs, and 498 mRNAs were observed (Figures 4A–H). To validate the accuracy of our whole transcriptome data, we chose three DECs, three DELs, three DEMs, and three DEGs at random for qRT–PCR analysis (Figure 3G). The correlation between the whole transcriptomic data and qRT–PCR data was 0.791, indicating that our sequencing results were repeatable and dependable. The raw data can be found at https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE208271. The transcription levels in the anterior vaginal wall tissue differed between the POP and control groups.

Focal adhesion is the main enrichment pathway

GO and KEGG enrichment analyses were performed to annotate the functional and pathway roles of the DEGs. The results showed that the GO terms “‘cell-substrate adhesion”, “collagen-containing extracellular matrix”, and “cell adhesion

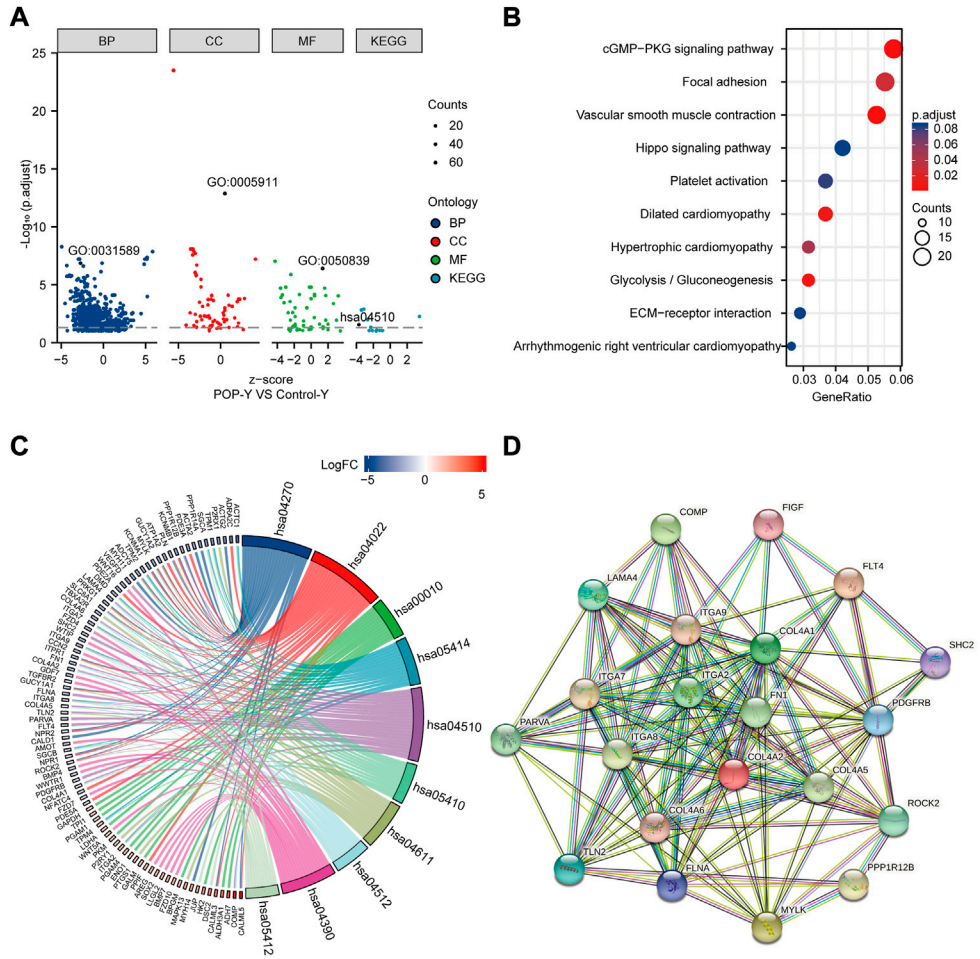


FIGURE 5
Functional enrichment analysis of DEGs using GO and KEGG and the PPI network. **(A)** Analysis of DEGs using GO and KEGG. **(B)** Top ten KEGG pathways. **(C)** Top ten KEGG pathways related to DEGs. **(D)** PPI network of focal adhesion pathway hub genes in the KEGG pathway.

molecule binding" were significantly enriched in the POP group compared with the control group (Figure 5A). Figure 5B depicts the top ten enriched KEGG pathways, which included the cGMP-PKG signaling pathway, focal adhesin, vascular smooth muscle contraction and hippo signaling pathway. Furthermore, to better analyze these pathways, Figure 5C depicts the DEGs involved in the pathways. Focal adhesion is an important structure for the transmission of mechanical signals from outside to inside cells, and the KEGG enrichment analysis showed that the DEGs were significantly enriched in focal adhesion. Thus, we mainly focus on focal adhesion in the subsequent analysis.

Hub genes that are closely related to focal adhesion signaling pathway

A PPI network was built using STRING online tools based on the DEGs associated with the focal adhesion signaling pathway (Figure 5D). Table 1 lists the enriched genes in this PPI network. We ultimately selected ITGA7, LAMA4, ROCK2, PPP1R12B, ITGA9, COL4A6, FLT4, COL4A5, COL4A2, ITGA2, VEGFD, FN1, ITGA8, and COL4A1 as hub genes because their interaction scores were >0.9.

Establishment of ROCK2-, PPP1R12B-, and VEGFD-related ceRNA networks

To investigate the potential regulatory involvement of DE ncRNAs, including circRNAs and lncRNAs, in targeting miRNAs in the focal adhesion signaling pathway in POP, we screened coexpressed genes from hub genes and then constructed and visualized the circRNA/lncRNA–miRNA–mRNA ceRNA network using R package v3.6.3. We selected miRNAs and circRNAs with TargetScan scores >95 to obtain more reliable results. The core miRNAs were determined to be miR-23a-3p

TABLE 1 The enriched genes in the KEGG pathway of focal adhesion signaling pathway.

KEGG pathway	Enriched genes
Focal adhesion	COMP, PPP1R12B, LAMA4, ITGA9, MYLK, COL4A6, FLT4, TLN2, ITGA7, COL4A5, COL4A2, SHC2, PARVA, ROCK2, ITGA2, VEGFD, FN1, FLNA, PDGFRB, ITGA8, COL4A1

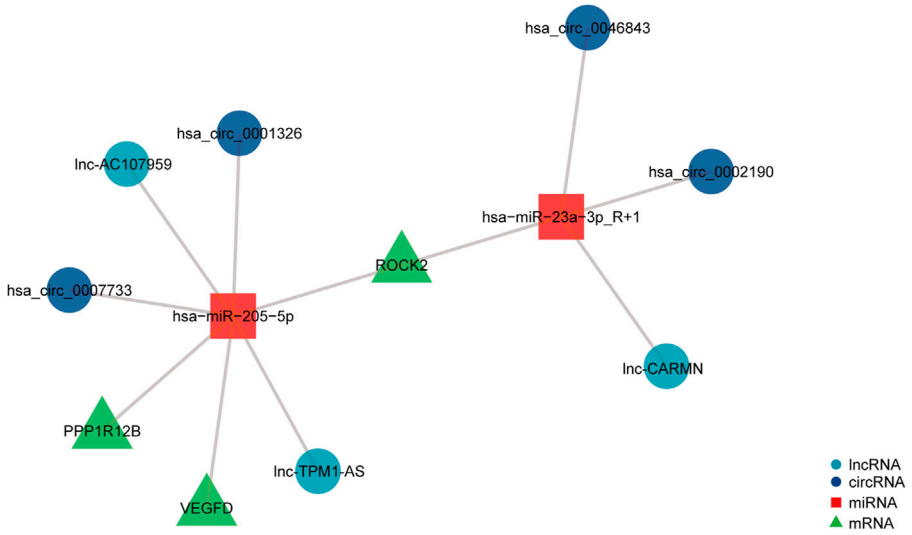


FIGURE 6
ceRNA network of differentially expressed circRNAs, lncRNAs, miRNAs, and mRNAs. The navy-blue circular nodes represent circRNAs, the light blue circular nodes represent lncRNAs, the square nodes represent miRNAs, and the triangular nodes represent mRNAs.

and miR-205-5p. Furthermore, hsa_circ_0002190, hsa_circ_0046843, and lnc-CARMN sponge miR-23a-3p with ROCK2 as the target gene. In addition, hsa_circ_0001326, hsa_circ_0007733, lnc-AC107959, and lnc-TPM1-AS regulate miR-205-5p, and hub genes such as ROCK2, PPP1R12B, and VEGFD were identified (Figure 6). Therefore, we have established ROCK2-, PPP1R12B-, and VEGFD-related ceRNA networks.

Abnormal cytoskeleton of vaginal fibroblasts in patients with pelvic organ prolapse

The action of focal adhesion on the cytoskeleton is the main mechanism for the transmission of a mechanical stimulation signal from the cell membrane to the intracellular space. Therefore, to analyze the changes in the cytoskeleton of fibroblasts in patients with POP, F-actin was stained with phalloidin. The results revealed that actin fibroblast filaments from the anterior vaginal wall were disrupted (Figure 7). An abnormal actin phenotype was observed in the POP group, and this phenotype involved decomposition of the stretched component of the filament. Thus, the cytoskeleton of vaginal fibroblasts in patients with POP is abnormal.

Discussion

POP is primarily characterized by the weakening of support tissue and can be caused by pelvic floor connective tissue extracellular matrix remodeling, activation of oxidative stress, genetic susceptibility, denervation of pelvic floor innervation, and reduced estrogen infiltration (Deng et al., 2021), however, the underlying mechanism, including the vital molecules and signaling pathways, is unknown. Furthermore, the role of regulatory ncRNAs in the pathogenesis of POP remains unclear. In this study, we performed whole transcriptome RNA sequencing to examine circRNAs, lncRNAs, miRNAs, and mRNAs in the anterior vaginal wall and identified 71 DECs, 76 DELs, 84 DEMs, and 931 DEGs. To test the accuracy of the whole transcriptomic data in this experiment, we chose 3 DECs, 3 DELs, 3 DEMs, and 3 DEGs at random for qRT–PCR verification. The results showed the same trend

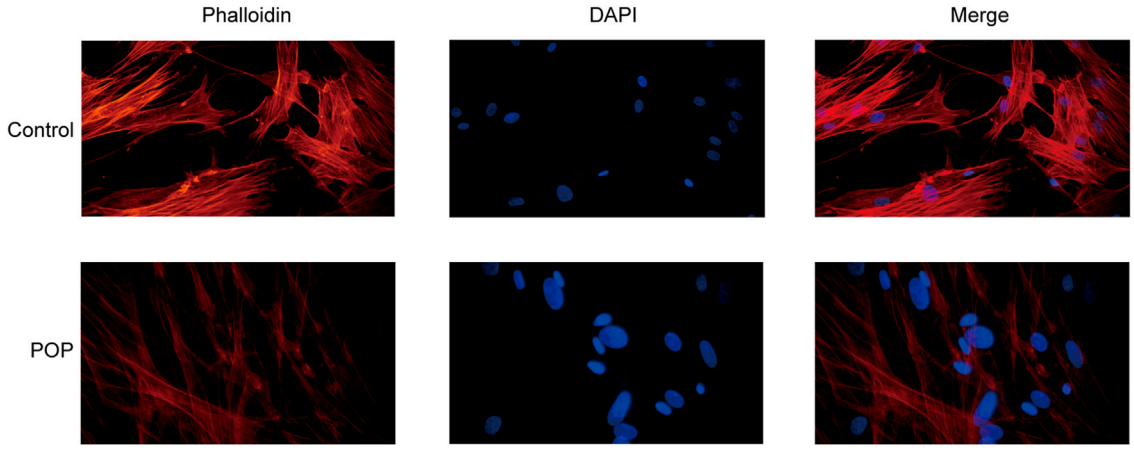


FIGURE 7
Cytoskeleton of fibroblasts from the POP and control groups. (400×).

between the two analyses and supported the high reliability of our data (R^2 = 0.791).

GO and KEGG enrichment analyses were then performed to analyze the function of DEGs and demonstrated that the pathways were mainly associated with focal adhesion. *In vivo*, cells are in contact with the extracellular matrix, bind to ECM proteins through integrins on the membrane and recruit force-sensitive proteins, such as vinculin, talin, paxillin, and focal adhesion kinase, to form a focal adhesion, which further mediates the connection between the cell membrane and the cytoskeleton, is then converted into biochemical signals, and thereby ultimately affect gene expression (Hoffmann et al., 2019). Therefore, changes in the extracellular matrix play an important role in the regulation of cell adhesion, proliferation, migration, and apoptosis through mechanical signal transduction (Fiorentino et al., 2022). It is well known that remodeling of the ECM in the pelvic floor connective tissue is one of the reasons for the occurrence of POP (Deng et al., 2021), particularly the proportions of collagen Ⅰ and Ⅲ (Deprest et al., 2022). Ruiz-Zapata et al. (2016) demonstrated that matrices increase stiffness in postmenopausal women with POP compared with controls. Therefore, the mechanical microenvironment of fibroblasts surrounded by extracellular matrix is also altered. The "focal adhesion - cytoskeleton system" is considered the main mechanism through which force stimulation signals are transmitted through the cell membrane to the intracellular cell. In our study, we found that patients with POP have a disorder of the cytoskeleton structure compared with the control group. Therefore, the focal adhesion signaling pathway was the focus of this study. Kufaishi et al. (2016). revealed that the transcript levels of integrins (ITGA1, ITGA4, ITGAV, and ITGB1) are downregulated in control vaginal fibroblasts during mechanical stretching, whereas ITGA2, ITGA4, ITGA6, ITGB1, contactin (CNTN1), catenins (A1 and B1), and laminins (A3 and C1) are significantly upregulated in vaginal fibroblasts from POP patients. Cecati et al. (2018) demonstrated a significant upregulation of extracellular matrix protein 1 (ECM1) and integrin beta 3 (ITGB3) and downregulation of FBLN5 in the POP group by real-time PCR and PCR array. The upregulation of ECM1 avoids collagen degradation and extracellular matrix remodeling by inhibiting matrix metalloproteinase-9 (MMP-9) activity (Fujimoto et al., 2006; Rapisarda et al., 2017). Furthermore, ITGB3 could cause a significant increase in transforming growth factor beta 1 (TGF-β1) activity, which could stimulate fibroblasts to increase FBLN-5 expression (Asano et al., 2005). Therefore, the reduced expression of FBLN5 may be the starting point. The pelvic tissue then increases the expression of both ITGB3 and ECM1 to compensate for a lack of FBLN5. In our study, integrins (ITGA2, ITGA7, ITGA8, and ITGA9) also showed differences in the anterior vaginal wall between the control and POP groups.

We chose 14 DEGs as hub genes and studied the interactions of ncRNAs, including circRNAs, lncRNAs, and miRNAs, using the PPI network. This analysis showed that hsa_circ_0002190, hsa_circ_0046843, and lnc-CARMN sponge miR-23a-3p with ROCK2 as the target gene. Furthermore, hsa_circ_0001326, hsa_circ_0007733, lnc-AC107959, and lnc-TPM1-AS regulate miR-205-5p, and hub genes such as ROCK2, PPP1R12B, and VEGFD were identified. A previous study showed that hsa_circ_0002190 expression is decreased in gastric cancer and that hsa_circ_0002190 accumulates preferentially in the cytoplasm (Dong et al., 2021). By regulating KIT and LAMC3, lnc-AC107959 may play a role in the mechanism of non-muscle invasive bladder cancer (He et al., 2018). Our study found that miR-23a-3p regulates ROCK2 expression. miR-23a-3p plays a

role in tumor pathology by regulating proliferation, invasion, and glycolysis (Shi et al., 2021). However, the biological role of miR-23a-3p in POP has not been investigated. We found that miR-23a-3p plays a regulatory role by targeting ROCK2. ROCK2 is a serine/threonine kinase that regulates actin-mediated cytoskeleton contractility and is a member of the AGC family of serine/threonine kinases (Weber and Herskowitz, 2021). ROCK2 is also targeted by miR-205-5p in our ceRNA network. miR-205-5p, similar to miR-23a-3p, is involved in modulating the proliferation and invasion of gastric cancer cells as part of the mechanism through which the lncRNA AFAP1-AS1 regulates AFAP1 (Dang et al., 2021). Furthermore, our findings indicate that PPP1R12B and VEGFD are involved in the regulation of miR-205-5p. Liang et al. (Tao et al., 2022) showed that inhibition of the PI3K/Akt signaling pathway can upregulate the expression of miR-205-5p and that miR-205-5p can inhibit the production of VEGF-A in breast cancer cells as well as tumor angiogenesis and metastasis. The present study found that miR-205-5p targets VEGFD to participate in the focal adhesion signaling pathway. We hypothesize that miR-205-5p inhibits fibroblast proliferation *via* the focal adhesion signaling pathway by targeting VEGFD because cell adhesion molecules also regulate angiogenesis *via* their involvement in cell proliferation, migration, and survival. More research is needed to confirm this theory. Furthermore, hsa_circ_0001326, hsa_circ_0007733, lnc-AC107959, and lnc-TPM1-AS sponge miR-205-5p. In preeclampsia, hsa_circ_0001326 increases IL16 expression to regulate proliferation, migration, invasion, and EMT (Liu et al., 2021). However, no information is available regarding the biological roles of hsa_circ_0007733, lnc-AC107959, and lnc-TPM1-AS. Even though many previous studies have investigated proliferation, apoptosis, and cytoskeleton morphology and confirmed that many molecules are associated with the pathological process in POP, they did not focus on the focal adhesion signaling pathway and did not construct a ceRNA network.

The current study provided a transcription database of POP and constructed a focal adhesion signaling pathway-related ceRNA network to uncover the pathogenesis of POP. Furthermore, this dataset can be used as a useful public resource for the identification of new biomarkers and may also provide new insights into elucidating the pathology of POP. However, this study has some limitations. First, the sample sizes of the control and POP groups were insufficient and need to be increased in future studies. Second, our findings are based on bioinformatics analysis, and more samples will be needed for confirmation via qRT–PCR, Western blot, and other methods. Furthermore, the different grades of POP may have an impact on whole transcriptome RNA sequencing. The whole transcriptome RNA sequencing results provide only a preliminary screening study result. Multiple experimental studies are needed to validate the regulatory mechanism of circRNA/lncRNA/miRNA/mRNA in POP.

In summary, we built a ceRNA network for the focal adhesion signaling pathway in POP pathogenesis, which included hsa_circ_0002190/hsa_circ_0046843/lnc-CARMN - miR-23a-3p - ROCK2 and hsa_circ_0001326/hsa_circ_0007733/lnc-AC107959/lnc-TPM1-AS - miR-205-5p - ROCK2/PPP1R12B/VEGFD. Furthermore, we discovered abnormal changes in the cytoskeleton, which are the final target of focal adhesion signaling events. This study not only contributes to the understanding of POP pathogenesis, but the findings can also be used to target molecular interventions for POP using drug-available gene databases.

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 studies involving human participants were reviewed and approved by The Ethics Committee of Chengdu Women's and Children's Central Hospital . The patients/participants provided their written informed consent to participate in this study.

Author contributions

ZY and YL: Conceptualization, Supervision. XY: Writing-Original draft preparation, Data curation. LH, YC: Writing - Reviewing and Editing. XY, LH, YC, WL, HL, XY, YY: Methodology. WL: Software, Validation. XZ: Data curation. All authors have read and approved the content of the manuscript.

Funding

This project was supported by Natural Science Foundation of Sichuan Province (grant number 2022NSFSC0815), the Yingcai Scheme, Chengdu Women's and Children's Central Hospital (grant number YC2021004), Youth Innovation Foundation of Sichuan Provincial Medical (grant number Q21060), Chengdu High-level Key Clinical Specialty Construction Project.

Acknowledgments

We thank Hu Li and Qingsong Liu for helpful cell culture. And we would like to thank all the participants, as well as all volunteers who helped us collect samples. We thank AJE Academic Services (https://www.aje.cn) for English-language editing and review services.

Conflict of interest

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

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.996310/full#supplementary-material

References

Anders, S., and Huber, W. (2010). Differential expression analysis for sequence count data. *Genome Biol.* 11, R106. doi:10.1186/gb-2010-11-10-r106

Asano, Y., Ihn, H., Yamane, K., Jinnin, M., Mimura, Y., and Tamaki, K. (2005). Increased expression of integrin alpha(v)beta3 contributes to the establishment of autocrine TGF-beta signaling in scleroderma fibroblasts. *J. Immunol.* 175, 7708–7718. doi:10.4049/jimmunol.175.11.7708

Cecati, M., Corradetti, A., Sartini, D., Pozzi, V., Giannubilo, S. R., Saccucci, F., et al. (2018). Expression of extracellular matrix and adhesion proteins in pelvic organ prolapse. *Cell. Mol. Biol.* 64, 142–148. doi:10.14715/cmb/2018.64.5.24

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

Collins, S. A., O'Shea, M., Dykes, N., Ramm, O., Edenfield, A., Shek, K. L., et al. (2021). International urogynecological consultation: Clinical definition of pelvic organ prolapse. *Int. Urogynecol. J.* 32, 2011–2019. doi:10.1007/s00192-021-04875-y

Dang, Y., Ouyang, X., Ren, W., Wang, L., and Huang, Q. (2021). LncRNA AFAP1-AS1 modulates the proliferation and invasion of gastric cancer cells by regulating AFAP1 via miR-205-5p. *Cancer Manag. Res.* 13, 5163–5175. doi:10.2147/CMAR.S307424

Deng, Z. M., Dai, F. F., Yuan, M. Q., Yang, D. Y., Zheng, Y. J., and Cheng, Y. X. (2021). Advances in molecular mechanisms of pelvic organ prolapse (Review). *Exp. Ther. Med.* 22, 1009. doi:10.3892/etm.2021.10442

Deprest, J. A., Cartwright, R., Dietz, H. P., Brito, L. G. O., Koch, M., Allen-Brady, K., et al. (2022). International urogynecological consultation (IUC): Pathophysiology of pelvic organ prolapse (POP). *Int. Urogynecol. J.* 33, 1699–1710. doi:10.1007/s00192-022-05081-0

Dong, Z., Liu, Z., Liang, M., Pan, J., Lin, M., Lin, H., et al. (2021). Identification of circRNA-miRNA-mRNA networks contributes to explore underlying pathogenesis and therapy strategy of gastric cancer. *J. Transl. Med.* 19, 226. doi:10.1186/s12967-021-02903-5

Ewies, A. A., Elshafie, M., Li, J., Stanley, A., Thompson, J., Styles, J., et al. (2008). Changes in transcription profile and cytoskeleton morphology in pelvic ligament fibroblasts in response to stretch: The effects of estradiol and levormeloxifene. *Mol. Hum. Reprod.* 14, 127–135. doi:10.1093/molehr/gam090

Fiorentino, G., Cimadomo, D., Innocenti, F., Soscia, D., Vaiarelli, A., Ubaldi, F. M., et al. (2022). Biomechanical forces and signals operating in the ovary during folliculogenesis and their dysregulation: implications for fertility. *Hum. Reprod. Update*. Advance online publication. doi:10.1093/humupd/dmac031

Fujimoto, N., Terlizzi, J., Aho, S., Brittingham, R., Fertala, A., Oyama, N., et al. (2006). Extracellular matrix protein 1 inhibits the activity of matrix metalloproteinase 9 through high-affinity protein/protein interactions. *Exp. Dermatol.* 15, 300–307. doi:10.1111/j.0906-6705.2006.00409.x

Harvey, M. A., Chih, H. J., Geoffrion, R., Amir, B., Bhide, A., Miotla, P., et al. (2021). International urogynecology consultation chapter 1 committee 5: relationship of pelvic organ prolapse to associated pelvic floor dysfunction symptoms: lower urinary tract, bowel, sexual dysfunction and abdominopelvic pain. *Int. Urogynecol. J.* 32, 2575–2594. doi:10.1007/s00192-021-04941-5

He, K., Niu, G., Gao, J., Liu, J. X., and Qu, H. (2016). MicroRNA-92 expression may be associated with reduced estrogen receptor β1 mRNA levels in cervical portion of uterosacral ligaments in women with pelvic organ prolapse. *Eur. J. Obstet. Gynecol. Reprod. Biol.* 198, 94–99. doi:10.1016/j.ejogrb.2016.01.007

He, Y. H., Deng, Y. S., Peng, P. X., Wang, N., Wang, J. F., Ding, Z. S., et al. (2018). A novel messenger RNA and long noncoding RNA signature associated with the progression of nonmuscle invasive bladder cancer. *J. Cell. Biochem.* 120 (5), 8101–8109. Advance online publication. doi:10.1002/jcb.28089

Hoffmann, G. A., Wong, J. Y., and Smith, M. L. (2019). On force and form: Mechano-biochemical regulation of extracellular matrix. *Biochemistry* 58, 4710–4720. doi:10.1021/acs.biochem.9b00219

Kim, D., and Salzberg, S. L. (2011). TopHat-fusion: An algorithm for discovery of novel fusion transcripts. *Genome Biol.* 12, R72. doi:10.1186/gb-2011-12-8-r72

Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., and Salzberg, S. L. (2013). TopHat2: Accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol.* 14, R36. doi:10.1186/gb-2013-14-4-r36

Kong, L., Zhang, Y., Ye, Z. Q., Liu, X. Q., Zhao, S. Q., Wei, L., et al. (2007). CPC: Assess the protein-coding potential of transcripts using sequence features and support vector machine. *Nucleic Acids Res.* 35, W345–W349. doi:10.1093/nar/gkm391

Kufaishi, H., Alarab, M., Drutz, H., Lye, S., and Shynlova, O. (2016). Static mechanical loading influences the expression of extracellular matrix and cell adhesion proteins in vaginal cells derived from premenopausal women with severe pelvic organ prolapse. *Reprod. Sci.* 23, 978–992. doi:10.1177/1933719115625844

Langmead, B., and Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat. Methods* 9, 357–359. doi:10.1038/nmeth.1923

Li, Y., Zhang, Q. Y., Sun, B. F., Ma, Y., Zhang, Y., Wang, M., et al. (2021). Single-cell transcriptome profiling of the vaginal wall in women with severe anterior vaginal prolapse. *Nat. Commun.* 12, 87. doi:10.1038/s41467-020-20358-y

Lin, W., Lin, L., Dong, B., Chen, L., Lei, H., Gao, Y., et al. (2020). The role of Obstetric factors, miRNA-30d and miRNA-181a in postpartum women with pelvic organ prolapse. *Risk Manag. Healthc. Policy* 13, 2309–2316. doi:10.2147/RMHP.S268235

Liu, Y., Ma, X., and Liu, Y. (2021). Hsa_circ_0001326 regulates proliferation, migration, invasion, and EMT of HTR-8/SVneo cells via increasing IL16 expression. *Am. J. Reprod. Immunol.* 86, e13484. doi:10.1111/aji.13484

Liu, Y., Khan, S., Li, L., Ten Hagen, T. L. M., and Falahati, M. (2022). Molecular mechanisms of thyroid cancer: A competing endogenous RNA (ceRNA) point of view. *Biomed. Pharmacother.* 146, 112251. doi:10.1016/j.biopha.2021.112251

Nussler, E., Granasen, G., Bixo, M., and Lofgren, M. (2022). Long-term outcome after routine surgery for pelvic organ prolapse-A national register-based cohort study. *Int. Urogynecol. J.* 33, 1863–1873. doi:10.1007/s00192-022-05156-y

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

Pfaffl, M. W. (2001). A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res.* 29, e45. doi:10.1093/nar/29.9.e45

Rapisarda, V., Borghesan, M., Miguela, V., Encheva, V., Snijders, A. P., Lujambio, A., et al. (2017). Integrin beta 3 regulates cellular senescence by activating the TGF-beta pathway. *Cell. Rep.* 18, 2480–2493. doi:10.1016/j.celrep.2017.02.012

Ruiz-Zapata, A. M., Kerkhof, M. H., Ghazanfari, S., Zandieh-Doulabi, B., Stoop, R., Smit, T. H., et al. (2016). Vaginal fibroblastic cells from women with pelvic organ prolapse produce matrices with increased stiffness and collagen content. *Sci. Rep.* 6, 22971. doi:10.1038/srep22971

Shi, Z., Li, G., Li, Z., Liu, J., and Tang, Y. (2021). TMEM161B-AS1 suppresses proliferation, invasion and glycolysis by targeting miR-23a-3p/HIF1AN signal axis in oesophageal squamous cell carcinoma. *J. Cell. Mol. Med.* 25, 6535–6549. doi:10.1111/jcmm.16652

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

Tao, Q., Qi, Y., Gu, J., Yu, D., Lu, Y., Liu, J., et al. (2022). Breast cancer cells-derived Von Willebrand Factor promotes VEGF-A-related angiogenesis through PI3K/Akt-miR-205-5p signaling pathway. *Toxicol. Appl. Pharmacol.* 440, 115927. doi:10.1016/j.taap.2022.115927

Wang, S., Zhang, Z., Lu, D., and Xu, Q. (2015). Effects of mechanical stretching on the morphology and cytoskeleton of vaginal fibroblasts from women with pelvic organ prolapse. *Int. J. Mol. Sci.* 16, 9406–9419. doi:10.3390/ijms16059406

Weber, A. J., and Herskowitz, J. H. (2021). Perspectives on ROCK2 as a therapeutic target for alzheimer's disease. *Front. Cell. Neurosci.* 15, 636017. doi:10.3389/fncel.2021.636017

Wu, J. M., Kawasaki, A., Hundley, A. F., Dieter, A. A., Myers, E. R., and Sung, V. W. (2011). Predicting the number of women who will undergo incontinence and prolapse surgery, 2010 to 2050. *Am. J. Obstet. Gynecol.* 205, 230 e1–e5. doi:10.1016/j.ajog.2011.03.046

Wu, J., Fang, Z., Liu, T., Hu, W., Wu, Y., and Li, S. (2021). Maximizing the utility of transcriptomics data in inflammatory skin diseases. *Front. Immunol.* 12, 761890. doi:10.3389/fimmu.2021.761890

Xie, R., Xu, Y., Fan, S., and Song, Y. (2016). Identification of differentially expressed genes in pelvic organ prolapse by RNA-seq. *Med. Sci. Monit.* 22, 4218–4225. doi:10.12659/msm.900224

Yin, Y., Qin, M., Luan, M., and Xia, Z. (2021). miR-19-3p promotes autophagy and apoptosis in pelvic organ prolapse through the AKT/mTOR/p70S6K pathway: Function of miR-19-3p on vaginal fibroblasts by targeting IGF-1. *Female Pelvic Med. Reconstr. Surg.* 27, e630–e638. doi:10.1097/SPV.0000000000001034

Zhang, X. O., Wang, H. B., Zhang, Y., Lu, X., Chen, L. L., and Yang, L. (2014). Complementary sequence-mediated exon circularization. *Cell.* 159, 134–147. doi:10.1016/j.cell.2014.09.001

Zhang, X. O., Dong, R., Zhang, Y., Zhang, J. L., Luo, Z., Zhang, J., et al. (2016). Diverse alternative back-splicing and alternative splicing landscape of circular RNAs. *Genome Res.* 26, 1277–1287. doi:10.1101/gr.202895.115

Zhao, B., Li, Z., Qin, C., Li, T., Wang, Y., Cao, H., et al. (2021). Mobius strip in pancreatic cancer: Biogenesis, function and clinical significance of circular RNAs. *Cell. Mol. Life Sci.* 78, 6201–6213. doi:10.1007/s00018-021-03908-5

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 28 September 2022
DOI 10.3389/fgene.2022.1005271

OPEN ACCESS

EDITED BY
Yuan Liu,
Shanghai Jiao Tong University, China

REVIEWED BY
Lin Wang,
University of California, San Francisco, United States
Cong Li,
Zhongshan School of Medicine, Sun Yat-sen University, China
Zao Zhang,
Queen's Medical Center, United States

*CORRESPONDENCE
Guohua Yang,
seawander@hotmail.com
Shunfang Liu,
liushunfang28@163.com

†These authors have contributed equally to this work

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

RECEIVED 28 July 2022
ACCEPTED 30 August 2022
PUBLISHED 28 September 2022

CITATION
Xiong R, Wang H, Li Y, Zheng J, Cheng Y, Liu S and Yang G (2022), Machine learning-based transcriptome analysis of lipid metabolism biomarkers for the survival prediction in hepatocellular carcinoma.
Front. Genet. 13:1005271.
doi: 10.3389/fgene.2022.1005271

Machine learning-based transcriptome analysis of lipid metabolism biomarkers for the survival prediction in hepatocellular carcinoma

Ronghong Xiong[1†], Hui Wang[2†], Ying Li[2], Jingpeng Zheng[2], Yating Cheng[2], Shunfang Liu[3*] and Guohua Yang[2*]

[1]Second Clinical College of Wuhan University, Zhongnan Hospital of Wuhan University, Wuhan, China, [2]Department of Medical Genetics, School of Basic Medical Science, Demonstration Center for Experimental Basic Medicine Education, Wuhan University, Wuhan, China, [3]Department of Oncology, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China

Hepatocellular carcinoma (HCC) is the most common primary malignancy of the liver with a very high fatality rate. Our goal in this study is to find a reliable lipid metabolism-related signature associated with prognostic significance for HCC. In this study, HCC lipid metabolism-related molecular subtype analysis was conducted based on the 243 lipid metabolism genes collected from the Molecular Signatures Database. Several significant disparities in prognosis, clinicopathological characteristics, and immune and ferroptosis-related status were found across the three subtypes, especially between C1 and C3 subgroups. Differential expression analysis yielded 57 differentially expressed genes (DEGs) between C1 and C3 subtypes. GO and KEGG analysis was employed for functional annotation. Three of 21 prognostic DEGs (CXCL8, SLC10A1, and ADH4) were finally selected through machine-learning-based discovery and validation strategy. The risk score = (0.103) × expression value of CXCL8 + (−0.0333) × expression value of SLC10A1 + (−0.0812) × expression value of ADH4. We used these three to construct a HCC prognostic risk model, which stratified the patients of the validation cohort into two risk subtypes with significantly different overall survival. Our work provides possible significance of the lipid metabolism-associated model in stratifying patient prognosis and its feasibility to guide therapeutic selection.

KEYWORDS

lipid metabolism, hepatocellular carcinoma, machine learning, prognostic risk model, biomarkers

Introduction

As a leading cause of cancer-related death worldwide, hepatocellular carcinoma (HCC) is the most prevalent type of primary liver malignancy (Balogh et al., 2016). Patients with HCC have a wide range in overall survival rates from region to region (Sung et al., 2021), with a 5-years survival rate of only 18% in the United States (Jemal et al., 2017). The most major risk factors for the development of HCC are chronic liver disease and cirrhosis, with viral hepatitis and excessive alcohol consumption being the primary well-known etiologies (Bruix et al., 2004). Therefore, it is vital and urgent to identify the prognostic value of novel markers that can aid in selecting patients who will benefit from patient-specific strategies.

The tumor microenvironment (TME) facilitates tumor metastasis, proliferation, and survival, which leads to abnormal metabolisms for tumor cells and those adjacent stromal cells. The TME in HCC might indicate a variety of metabolic disturbances, with lipid metabolic anomaly being a fresh subject that has sparked a lot of interest in recent years (Beloribi-Djefaflia et al., 2016). Lipid metabolic disturbance, particularly for fatty acid (FA) metabolism, is associated with altered lipid-metabolizing enzyme expression and activity due to aberrantly activated oncogenic signaling pathways (Hu et al., 2020). Lipid metabolism has been increasingly recognized as a critical phenomenon of metabolic rewiring within immune cells and cancer cells, which may be involved in the development of HCC. Furthermore, evidence from various solid tumor research suggests that tumor immune-metabolic reprogramming is significant, and it has been designated as a new critical subject for future HCC studies (Zhang et al., 2018). According to prior research, immune cells play an important role in the TME of HCC, and aberrant lipid metabolism may have a major impact on their activities and recruitment (Gajewski et al., 2013). Although growing studies have explored the genetic, cellular, and environmental mechanisms involved in the development of tumors (Chen et al., 2020; Zhang et al., 2020; Cao et al., 2021a; Zhong et al., 2021), clinicians currently have few choices for slowing HCC progression and extending patients' lives. Therefore, integrated lipid metabolism and liver cancer progression to build an effective prediction model is needed and is the focus of this investigation.

In this study, our goal is to identify a robust lipid metabolism-related signature associated with the HCC microenvironment to improve the prognostic prediction of HCC patients. Genes related to energy metabolism were collected from the Molecular Signatures Database. Gene expression data from The Cancer Genome Atlas (TCGA) were used in constructing HCC molecular subtypes based on genes related to energy metabolism. The relationship between molecular subtypes and prognosis was further evaluated. After differential expression analysis and machine-learning-based selection, three lipid metabolism-driven signatures were chosen from the 576 differentially expressed genes (DEGs) for establishing a prognostic risk model. Then, we validated the risk model, which may be used to assess the prognosis of HCC patients. Overall, this 3-signature prognostic risk model (CXCL8, SERPINC1, and ADH4) we built can be used as an independent prognostic evaluation index for HCC patients.

Materials and methods

Data collection and preprocessing

RNA-sequencing expression (level 3) profiles and corresponding clinical information for 371 HCC as well as 50 healthy subjects were derived from the TCGA dataset (https://portal.gdc.com). The raw data were preprocessed with the criteria which have been described elsewhere (Cao et al., 2020; Mao et al., 2020; Cao et al., 2021b; Mao et al., 2021). A total of 243 lipid metabolism-associated genes were gathered based on the Molecular Signatures Database v7.5.1 (c2: curated gene sets), including Fatty acid metabolism M699, Glycerophospholipid metabolism M9131, Glycerolipid metabolism M15902, Sphingolipid metabolism M15955, Ether lipid metabolism M2130, Glycosphingolipid biosynthesis-ganglio series M8535, Biosynthesis of unsaturated FAs M11673, Glycosphingolipid biosynthesis-globo series M12899, Glycosphingolipid biosynthesis-lacto, and neolacto series M17377 (http://www.broadinstitute.org/gsea/msigdb/index.jsp).

Identification and validation of the lipid-related subtypes

Consensus clustering was applied to identify a robust cluster of HCC patients based on the expression profile of The Cancer Genome Atlas Liver Hepatocellular Carcinoma (TCGA-LIHC) data (Cao et al., 2020). The 1,000 bootstraps with 80% item resampling and a range of K from 2 to 10 were selected for clustering analysis. Partition around the medoids classifier was trained in the discovery cohort. By calculating the in-group proportion and Euclidean correlation in the centroid of gene module scores, we quantitatively acquired and verified the consistency of immune subtypes among populations. The expression value for lipid metabolism-associated gene within each subtype was used for Principal component analysis (PCA) by the "prcomp" function in R.

Immune status and ferroptosis-related estimation

We used the CIBERSORT method (Chen et al., 2018) to assess the immune composition of a tumor biopsy and get

reliable results of immune score evaluation. The relative abundance enrichment score of 22 Tumor infiltrating leukocytes (TILs) was measured and standardized from 0 to 1. Moreover, potential immune checkpoint blockade (ICB) response was predicted with Tumor Immune Dysfunction and Exclusion (TIDE) algorithm (Jiang et al., 2018). Twenty-four Ferroptosis-related genes were collected from the previous study (Liu et al., 2020). The expression distribution of ferroptosis-related mRNA in tumor and normal tissues was implemented by the R program v4.0.3.

DEG identification and functional enrichment analysis

The differentially expressed genes (DEGs) of the subtypes were identified using the "limma" algorithm for subsequent analyses (FDR adjusted p-value < 0.05 and $|\log_2 FC| > 3$; FC, fold change, FDR, False Discovery Rate) (Ritchie et al., 2015). Afterward, Gene Ontology (GO) terms and Kyoto Encyclopedia of Genes and Genomes (KEGG) functional enrichment analyses were conducted based on the DEGs (Wu et al., 2021). The GO terms and KEGG pathways with a p-value of < 0.05 were considered significantly enriched function annotations.

Development and validation of the prognostic signature

To identify overall survival (OS)-related genes from DEGs and detect lipid metabolism-driven prognostic signature (LMSig), we randomly divided the mRNA expression profile of 371 HCC patients into two parts as the discovery (186 samples) and validation data (185 samples). Then, machine-learning-based variable selection was carried out on the discovery data using likelihood-based boosting in the Cox model as implemented in the R package "CoxBoost" [30]. The machine-learning algorithm has been as described previously in detail (Hou et al., 2021). For the CoxBoost model, the number of boosting iterations was then optimized through cross-validation after the optimal penalty had been determined through ten cross-validations using the R package "CoxBoost" (De Bin, 2016). The CoxBoost algorithm was used to automatically estimate the optimal number of LMSig.

In addition, we used the validation data to further confirm the relationship between LMSigs and clinical/prognostic features of HCC. The p-values and hazard ratio (HR) with 95% confidence interval (CI) were generated by log-rank tests and univariate cox proportional hazards regression in Kaplan–Meier curves analysis using the R package "survival" (Therneau and Lumley, 2015; Zhou et al., 2019). The concordance statistic (C-statistic index) was used to measure the goodness of fit of the prognostic model. The time-dependent receiver operating characteristic (ROC) curve was used to appraise the prognostic performance of the risk model for survival prediction, and the area under the ROC curve (AUC) values were calculated with the R package "timeROC" (Blanche and Blanche, 2019).

Statistical analysis

The statistical difference between the two groups was compared through the Wilcox test, the significant difference between the three groups was tested with the Kruskal–Wallis test. All statistical tests were two-sided, p-value < 0.05 was considered statistically significant. All the data were processed and analyzed by R program v4.0.3.

Results

Lipid metabolic molecular subtypes of HCC identification

An expression profile of 243 common lipid-related genes in 371 HCC patients from the TCGA database was used to implement the consensus clustering. The analysis clustered the patients with HCC into three subgroups C1, C2, and C3 (Figure 1A). These three conceivable subclusters were respectively distinguished *via* first and second principal components (PCs) (Figure 1B). As shown in Figure 1C, the expression levels of lipid metabolism-related genes seem to differ among the three subtypes. Prognosis signature among them was further analyzed. The Kaplan–Meier method was used to investigate the overall (OS) of the three subgroups, and we observed that the patients in the C3 subtype had the worst prognosis, while the C2 subgroup had significantly best OS (Figure 1D, p-value = 0.0043).

Clinicopathological and immune infiltration characteristics in three subgroups

The clinicopathological characteristics of the three subtypes were then compared. Tumor T stage, Gleason grade, and type of treatment among subtypes reached statistical significance (Figure 2A). The results manifested that those patients diagnosed with differential T stage and grade were clustered unevenly. T3 stage and G3 accounted for the major proportion of C3 while T1 and G2 were in the majority in C1 (Figure 2A).

The results of CIBERSORT showed significant differences in infiltrating immune cell types between the three subgroups. There were more abundant proportions of T cell CD8+and T cell regulatory (Tregs) in the C1 subgroup than in the C3 subgroup. On the other hand, Macrophage M0 was

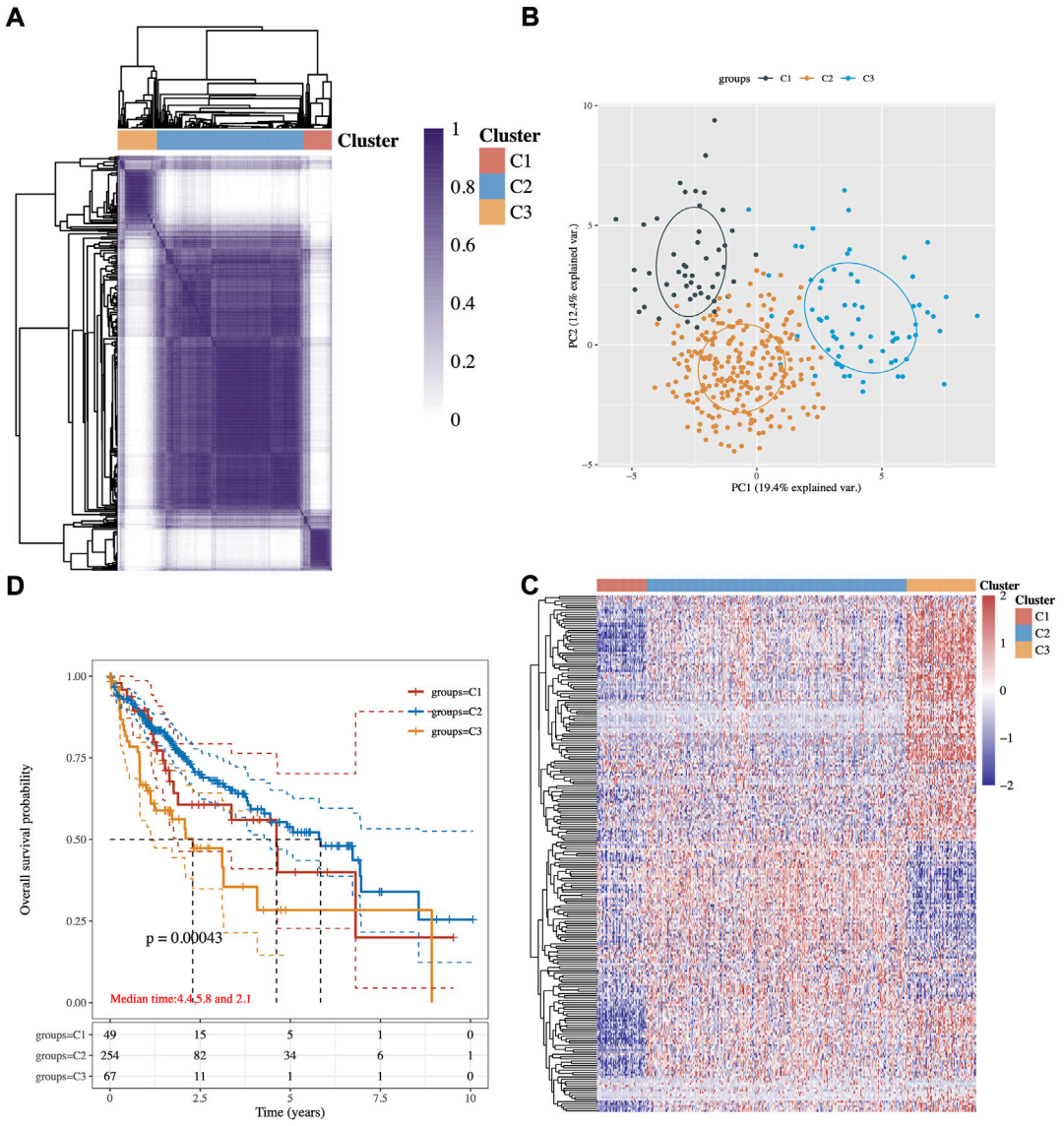


FIGURE 1
Identification of potential lipid metabolism-related subtypes of HCC. **(A)** Consistency of clustering results in the heatmap (k = 3), rows and columns represent samples, the different colors represent different types. **(B)** PCA analysis of different subgroups with PC1 and PC2. **(C)** The expression heatmap of lipid metabolism-related genes in three subgroups, red represents high expression, and blue represents low expression. **(D)** Kaplan–Meier survival analysis of the different groups of samples from TCGA dataset, comparison among different groups was made by log-rank test. HR (95% CI), the median survival time for different groups.

significantly enriched in C3 compared with C1 (p-value < 0.05; Figure 2B). Moreover, eight ICPs-associated genes were differentially expressed among C1-3. C3 showed significant upregulation of CD274, CTLA4, HAVCR2, PDCD1, PDCD1LG2, TIGIT, and SIGLEC15, while these genes were down-expressed separately in the C1 tumors (Figure 2C).

To predict the ICB response of identified HCC subtypes, the TIDE score was calculated. The findings showed that the TIDE score was significantly lower in the C1 subtype than in the C3 subtype (Kruskal–Wallis test, p-value = 4.9×10^{-11}; Figure 2D). These discoveries suggested that patients of the C1 subtype may be more sensitive to ICB therapy as judged by the TIDE score.

Ferroptosis-related estimation among three distinct subgroups

Ferroptosis is known as an iron-dependent form of regulated cell death (RCD) triggered by lipid peroxidation

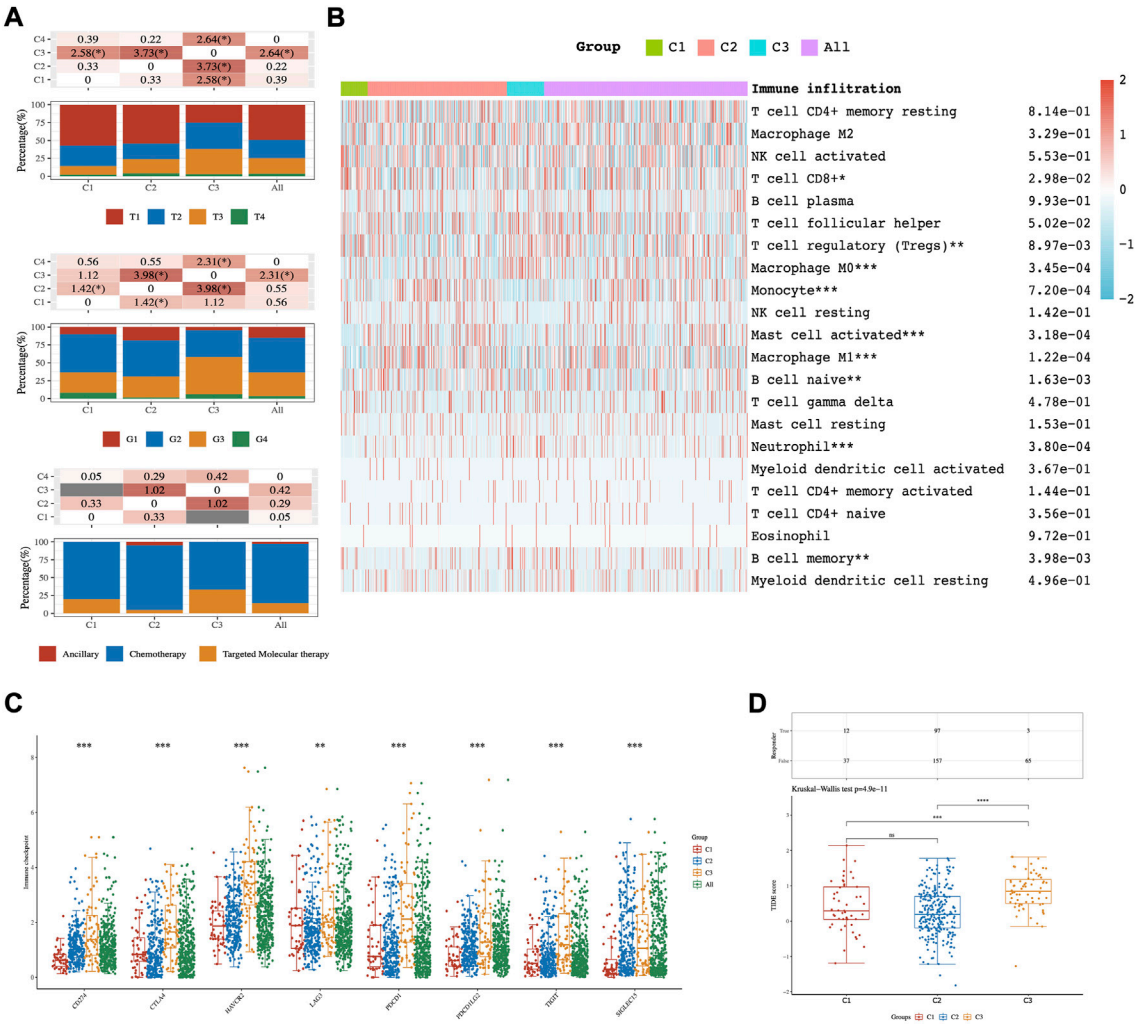


FIGURE 2
The distribution of clinical and immune characteristics in the samples from different groups. **(A)** Distribution of clinical characteristics across C1-C3 HCC patients. Association between three different subtypes and CIBERSORT immune infiltration **(B)**, ICPs **(C)**, and TIDE score **(D)**. Asterisks (*) stand for significance levels. * represents $p < 0.05$, ** represents $p < 0.01$,*** represents $p < 0.001$.

accumulation (Aguilera et al., 2021). Recently, triggering ferroptosis has emerged as a promising therapeutic option for inducing cancer cell death, particularly for malignancies that are resistant to traditional therapies (Liang et al., 2019). In our study, the comprehensive landscapes of ferroptosis-related gene interactions, connections, and their prognostic significance for HCC patients in three subgroups were depicted, respectively (Figure 3A). We found that ferroptosis-related genes in distinct subgroups presented a remarkably different correlation in expression. The expression of each ferroptosis-related gene also differed insignificantly among the three lipid subtypes. As shown in Figure 3B, the expression levels of CDKN1A, HSPA5, EMC2, SLC7A11, NFE2L2, FANCD2, SLC1A5, CS, and CARS1 were highly accumulated in C3 compared to C1 subgroups.

DEG identification and functional analysis

Transcriptome differential expression was performed between C1 and C3 subgroups of HCC patients according to the above difference among them. Fifty-seven genes were identified as DEGs at FDR < 0.05 and $\log_2$FC > 3, of which 27 DEGs were up-regulated, and 30 DEGs were down-regulated (Figures 4A,B). Then, GO functional and KEGG pathway enrichment analyses were performed. The results of KEGG analysis demonstrated that up-regulated DEGs majorly

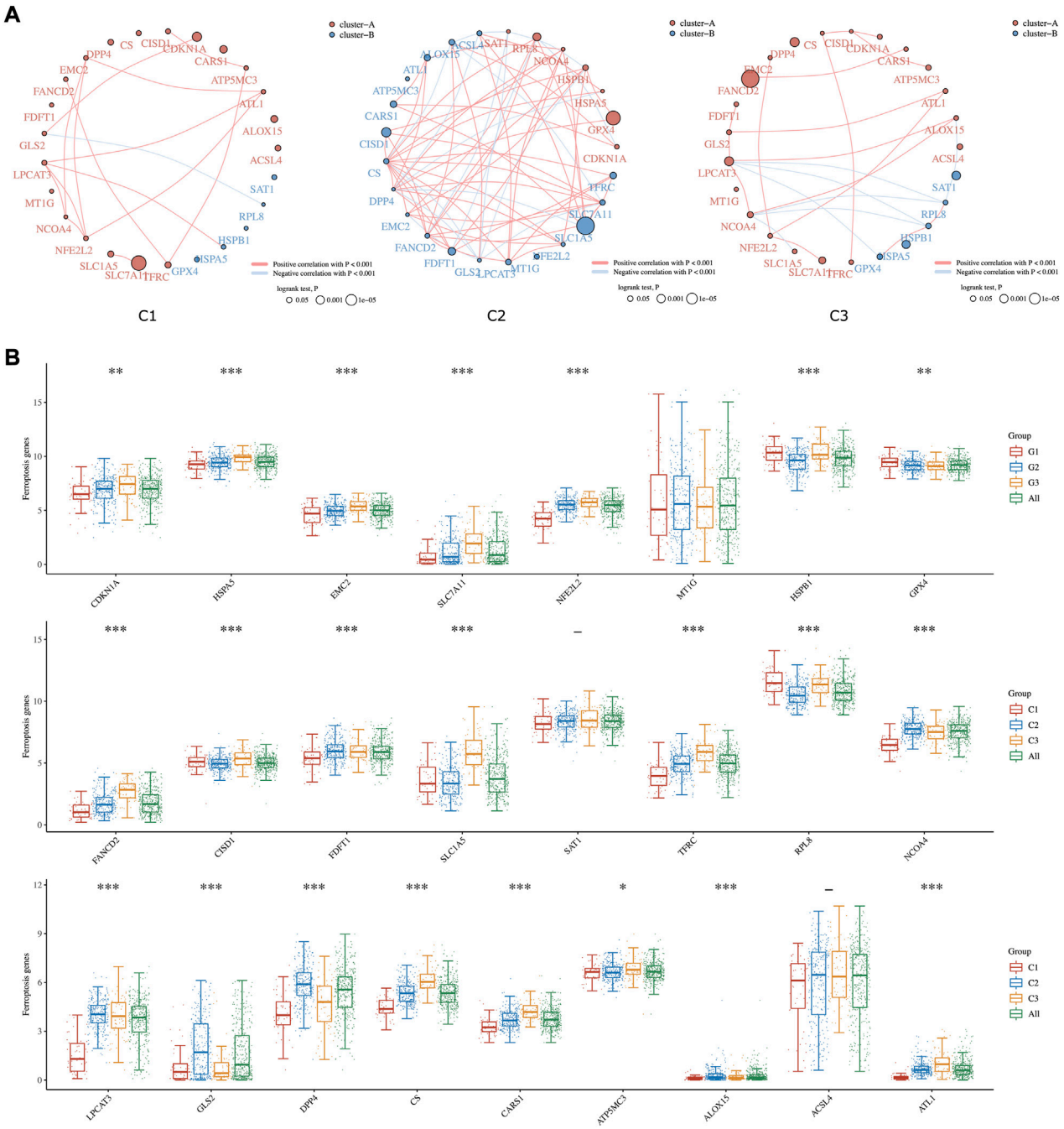


FIGURE 3
Ferroptosis-related Estimation in three subgroups. **(A)** The circles represent the ferroptosis-related mRNA, and the line represents the relationship between genes. Red represents positive correlation whereas blue represents negative correlation. The thicker the line, the higher the correlation coefficient. The larger the circle the smaller the log-rank *p* value. Different colors of circles represent different types of clusters. **(B)** The expression distribution of ferroptosis-related mRNA in tumor tissues and normal tissues. Asterisks (*) stand for significance levels. * represents $p < 0.05$, ** represents $p < 0.01$, and *** represents $p < 0.001$.

participated in complement and coagulation cascades, cholesterol metabolism, fatty acid degradation, and bile secretion (Figure 4C). Meanwhile, down-regulated DEGs were involved in Hepatitis B, proteoglycans in cancer, and the PI3K–Akt signaling pathway. GO analysis revealed that these DEGs were mostly enriched in steroid metabolic

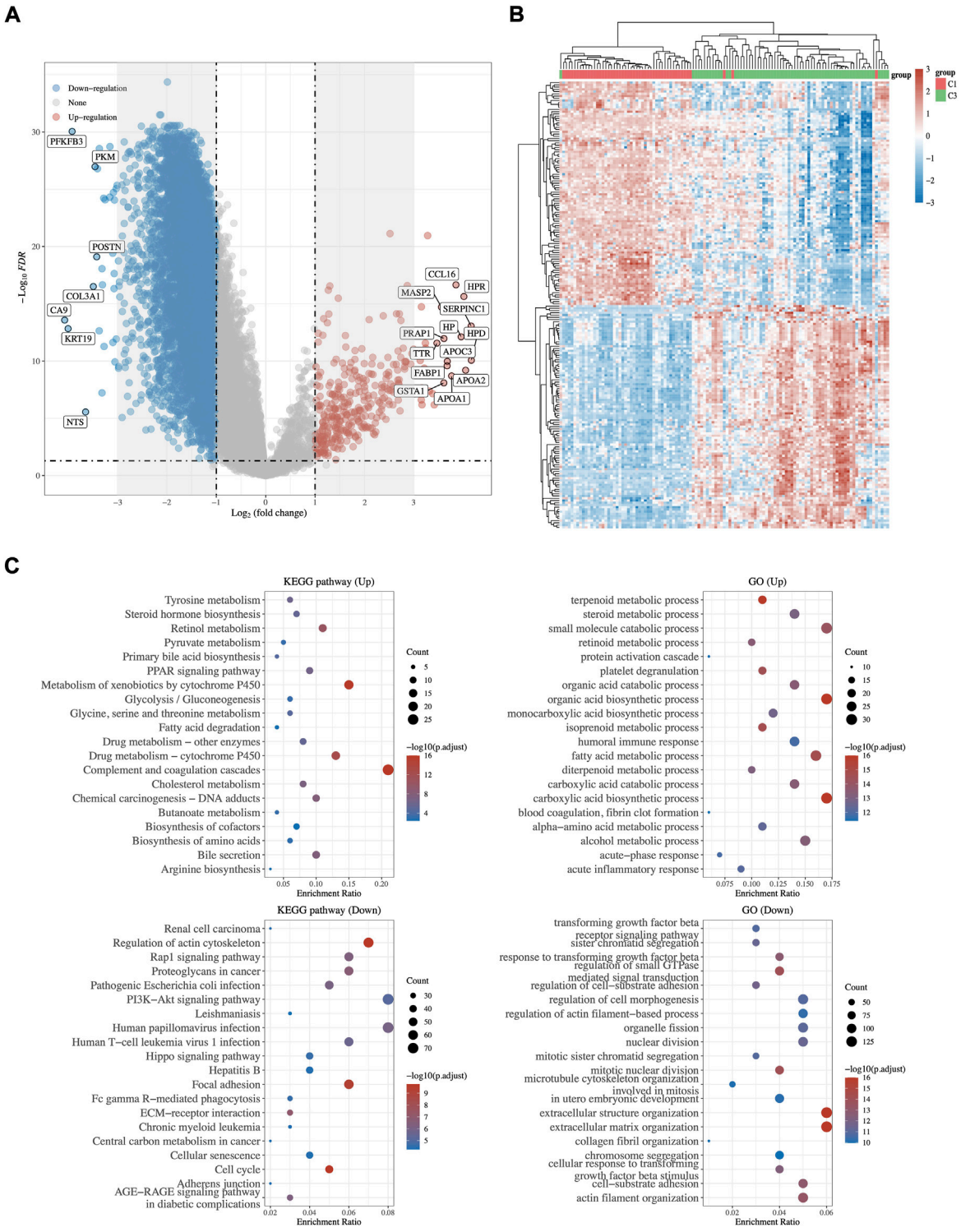


FIGURE 4
DEGs Identification and functional analysis between C1 and C3. **(A)** The volcano plot and heatmap **(B)** of differentially expressed gene analysis between C1 and C3. Red represents up regulation whereas blue represents down regulation. **(C)** GO and KEGG enrichment analysis of up- and down-regulated DEGs.

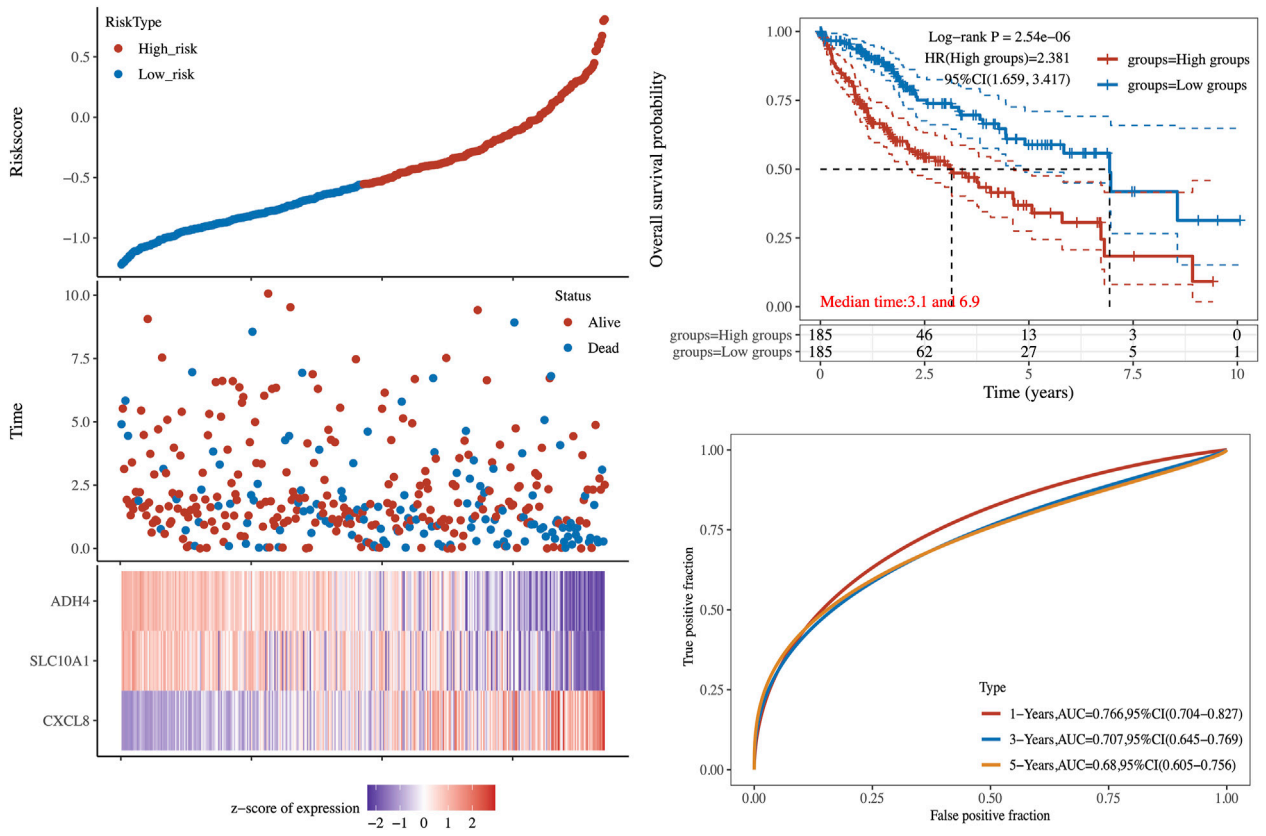


FIGURE 5
The validation of 3LMSig-associated prognostic risk model.

process, terpenoid metabolic process (up-regulated DEGs), and regulation of small GTPase (down-regulated DEGs), which were related to lipid metabolism (Figure 4C).

Establishing a prognostic risk model

Integrated the differential gene expression and patient survival data from the TCGA cohort, we screened DEGs to discover the feasibility and reliability of a prognostic signature for HCC. Among all 57 DEGs, 21 DEGs were associated with the OS of HCC patients (p-value < 0.05). Subsequently, a boosting machine learning algorithm was performed for signature selection from these 21 DEGs and identified three lipid metabolism-driven signatures (3LMSig). Then, the 3LMSig was transformed into a risk scoring model by linear combination of the expression of the 3LMSig as follows: risk score of 3LMSig = (0.103) × expression value of CXCL8 + (−0.0333) × expression value of SLC10A1 + (−0.0812) × expression value of ADH4. Among them, the expression value of CXCL8 was linked positively to HCC risk score, while the expression value of SERPINC1 and ADH4 showed a negative relationship with HCC risk score.

According to the risk score, the patients in validation cohort were divided into the high-risk group and the low-risk group. Low-risk patients had statistically significantly better OS than those in the high-risk group (Log-rank p-value = 2.54×10^{-6}, Figure 5). To compare the sensitivity and specificity of survival prediction, a time-dependent ROC curve analysis of this 3LMSig-based risk score model was performed. The area under the curves (AUCs) of the nomogram model in 1-, 3-, and 5-years were 0.766, 0.707, and 0.68, respectively (Figure 5), which suggested the good performance of the risk score signature.

Discussion

Prognostic prediction of Hepatocellular carcinoma (HCC) patients has been challenging due to the complicated etiologic variables and high-level heterogeneity of HCC (Liang et al., 2020). Therefore, there is an additional need for the development of novel prognostic models, considering the

limited treatment strategies for HCC. Emerging data suggest that changes in tumor lipid metabolism, including metabolite abundance and lipid metabolic product accumulation, contribute to tumor formation and local immunosuppression in the TME (Hao et al., 2019). As a result, we focused on learning more about the link between tumor lipid metabolic genes and prognosis in HCC. We seek to develop a panel of prognostic markers using molecular markers derived from tumor metabolic genes.

In this study, the public gene expression data from the TCGA-LIHC database were utilized to classify HCC patients into three molecular subtypes C1-3 based on 243 lipid metabolism-related genes. Several significant disparities in prognosis, clinicopathological characteristics, and immune and ferroptosis-related status were found across the three subtypes, especially between C1 and C3 subgroups. For example, a recently discovered cell death mechanism called ferroptosis may serve as a therapeutic biomarker for HCC. We observed C3 and C1 cluster can be classified as ferroptosis-high and ferroptosis-low groups according to the expression levels of ferroptosis-related genes. Previous research inferred the ferroptosis-high group have a worse prognosis and higher immune score (Deng et al., 2021), in line with our findings.

Then, 57 DEGs between these two subgroups were identified. GO and KEGG enrichment analysis displayed that these DEGs were closely associated functionally with lipid metabolism and tumorigenesis. Selected by machine-learning- based feature selection afterward, a prognostic risk model including 3LMSig was established. The risk model consisting of CXCL8, SLC10A1, and ADH4 was effective in predicting the prognosis of HCC patients. Moreover, the risk score calculated from the established risk model divided patients into high-risk and low-risk groups. The risk model showed that high CXCL8 expression level was associated with a bad prognosis, and high expression of SERPINC1 and ADH4 was related to better overall survival. The ROC curve analysis confirmed the moderate discriminatory accuracy of the model. According to findings, the lipid metabolism-related signature has prognostic significance for HCC. Previous research reported that SLC10A1 (solute carrier family 10 member 1) can inhibit the Warburg effect to suppress HCC tumor growth (Lu et al., 2020). ADH4 (Alcohol dehydrogenase 4), a member of the ADH family, metabolizes a wide variety of substrates including ethanol and retinol (Wei et al., 2012). CXCL8 is a promising prospective prognostic and tumor TME-related cluster (Zhu et al., 2020).

The advantage of this study is that we have identified a prognostic feature by 3LMSig that predicts 1-, 3-, and 5- year survival with relatively high AUC. However, there are limitations to this initial work. Some of the findings from this study could not be explained satisfactorily given our current limited knowledge of cancer biology. Moreover, independent studies are warranted to replicate our findings.

In summary, our study divided HCC patients into three lipid metabolism-related molecular subtypes with different prognoses and other molecular features. Then, a risk model with a good performance in prognostic prediction was built using the TCGA dataset. This model can be used as an independent prognostic evaluation index for HCC patients. Our work shed lighter on the possible significance of the lipid metabolism-associated model in stratifying patient prognosis and its feasibility to guide therapeutic selection.

Data availability statement

Publicly available datasets were analyzed in this study. The names of the repository/repositories and accession number(s) can be found in the article/supplementary material.

Author contributions

GY and SL designed this work. RX, HW, and YL integrated and analyzed the data. JZ and YC wrote this manuscript. RX, HW, GY, and SL edited and revised the manuscript. All authors approved this manuscript.

Funding

The present study was supported by the National Natural Science Foundation of China (project no. 81602626), Scientific Research Foundation for Returned Scholars of Tongji Hospital (project no. 2022hgry021), “School of Basic Medical Science and Zhongnan Hospital” Joint Development Fundamental Research Funds of Wuhan University. Chen Xiao-Ping Foundation for The Development of Science and Technology of Hubei Province (CXPJJH12000001-2020308).

Conflict of interest

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

Publisher’s note

References

Aguilera, A., Berdun, F., Bartoli, C., Steelheart, C., Alegre, M., Bayir, H., et al. (2021). C-ferroptosis is an iron-dependent form of regulated cell death in cyanobacteria. *J. Cell Biol.*, 221. Rockefeller University Press, e201911005. doi:10.1083/jcb.201911005

Balogh, J., Victor, D., III, Asham, E. H., Burroughs, S. G., Boktour, M., and Saharia, A. (2016). Hepatocellular carcinoma: A review. *J. Hepatocell. carcinoma*, 3. Dove Press, 41. doi:10.2147/jhc.s61146

Beloribi-Djefaflia, S., Vasseur, S., and Guillaumond, F. (2016). Lipid metabolic reprogramming in cancer cells. Oncogenesis. *Nat. Publ. Group* 5, e189. doi:10.1038/oncsis.2015.49

Blanche, P., and Blanche, M. P. (2019). *Package 'timeROC.*

Bruix, J., Boix, L., Sala, M., and Llovet, J. M. (2004). Focus on hepatocellular carcinoma. *Cancer Cell* 5, 215–219. doi:10.1016/s1535-6108(04)00058-3

Cao, F., Fan, Y., Yu, Y., Yang, G., and Zhong, H. (2021). Dissecting prognosis modules and biomarkers in glioblastoma based on weighted gene Co-expression network analysis. *Cancer Manag. Res.*, 13. . Dove Press, 5477. doi:10.2147/cmar.s310346

Cao, F., Guo, Y., Zhang, Q., Fan, Y., Liu, Q., Song, J., et al. (2020). Integration of transcriptome resequencing and quantitative proteomics analyses of collagenase vii-induced intracerebral hemorrhage in mice. *Front. Genet.* 11, 551065. doi:10.3389/fgene.2020.551065

Cao, F., Wang, C., Long, D., Deng, Y., Mao, K., and Zhong, H. (2021). Network-based integrated analysis of transcriptomic studies in. *Dissecting Gene Signatures LPS-Induced Acute Lung Inj.*, 44. Inflammation. Springer, 2486–2498. doi:10.1007/s10753-021-01518-8

Chen, B., Khodadoust, M. S., Liu, C. L., Newman, A. M., and Alizadeh, A. A. (2018). Profiling tumor infiltrating immune cells with CIBERSORT. *Methods Mol. Biol.* 1711, 243–259. doi:10.1007/978-1-4939-7493-1_12

Chen, Q., Li, F., Gao, Y., Xu, G., Liang, L., and Xu, J. (2020). Identification of energy metabolism genes for the prediction of survival in hepatocellular carcinoma. *Front. Oncol.* 10, 1210. doi:10.3389/fonc.2020.01210

De Bin, R. (2016). Boosting in cox regression: A comparison between the likelihood-based and the model-based approaches with focus on the R-packages CoxBoost and mboost. *Comput. Stat.* 31, 513–531. doi:10.1007/s00180-015-0642-2

Deng, T., Hu, B., Jin, C., Tong, Y., Zhao, J., Shi, Z., et al. (2021). A novel ferroptosis phenotype-related clinical-molecular prognostic signature for hepatocellular carcinoma. *J. Cell. Mol. Med.* 25, 6618–6633. doi:10.1111/jcmm.16666

Gajewski, T. F., Schreiber, H., and Fu, Y-X. (2013). Innate and adaptive immune cells in the tumor microenvironment. *Nat. Immunol.*, 14. Nature Publishing Group, 1014–1022. doi:10.1038/ni.2703

Hao, Y., Li, D., Xu, Y., Ouyang, J., Wang, Y., Zhang, Y., et al. (2019). Investigation of lipid metabolism dysregulation and the effects on immune microenvironments in pan-cancer using multiple omics data. *BMC Bioinforma.* 20, 195–239. doi:10.1186/s12859-019-2734-4

Hou, P., Bao, S., Fan, D., Yan, C., Su, J., Qu, J., et al. (2021). Machine learning-based integrative analysis of methylome and transcriptome identifies novel prognostic DNA methylation signature in uveal melanoma. *Brief. Bioinform*, 22. Oxford University Press, bbaa371. doi:10.1093/bib/bbaa371

Hu, B., Lin, J., Yang, X., and Sang, X. (2020). Aberrant lipid metabolism in hepatocellular carcinoma cells as well as immune microenvironment: A review. *Cell Prolif.*, 53. Wiley Online Library, e12772. doi:10.1111/cpr.12772

Jemal, A., Ward, E. M., Johnson, C. J., Cronin, K. A., Ma, J., Ryerson, A. B., et al. (2017). Annual report to the nation on the status of cancer, 1975–2014, featuring survival. *J. Natl. Cancer Inst.* 109, djx030. doi:10.1093/jnci/djx030

Jiang, P., Gu, S., Pan, D., Fu, J., Sahu, A., Hu, X., et al. (2018). Signatures of T cell dysfunction and exclusion predict cancer immunotherapy response. *Nat. Med.*, 24. Nature Publishing Group, 1550–1558. doi:10.1038/s41591-018-0136-1

Liang, C., Zhang, X., Yang, M., and Dong, X. (2019). Recent progress in ferroptosis inducers for cancer therapy. *Adv. Mater*, 31. Wiley Online Library, 1904197. doi:10.1002/adma.201904197

Liang, J., Wang, D., Lin, H., Chen, X., Yang, H., Zheng, Y., et al. (2020). A novel ferroptosis-related gene signature for overall survival prediction in patients with hepatocellular carcinoma. *Int. J. Biol. Sci.* 16, 2430–2441. doi:10.7150/ijbs.45050

Liu, Z., Zhao, Q., Zuo, Z-X., Yuan, S-Q., Yu, K., Zhang, Q., et al. (2020). Systematic analysis of the aberrances and functional implications of ferroptosis in cancer. *iScience* 23, 101302. doi:10.1016/j.isci.2020.101302

Lu, C., Fang, S., Weng, Q., Lv, X., Meng, M., Zhu, J., et al. (2020). Integrated analysis reveals critical glycolytic regulators in hepatocellular carcinoma. *Cell Commun. Signal.* 18, 97–14. doi:10.1186/s12964-020-00539-4

Mao, K., Geng, W., Liao, Y., Luo, P., Zhong, H., Ma, P., et al. (2020). Identification of robust genetic signatures associated with lipopolysaccharide-induced acute lung injury onset. *Aging (Albany NY).* 12. Impact Journals, LLC, 18716. doi:10.18632/aging.104042

Mao, K., Luo, P., Geng, W., Xu, J., Liao, Y., Zhong, H., et al. (2021). An integrative transcriptomic and metabolomic study revealed that melatonin plays a protective role in chronic lung inflammation by reducing necroptosis. *Front. Immunol.* 12, 1469. doi:10.3389/fimmu.2021.668002

Ritchie, M. E., Phipson, B., Wu, D. I., Hu, Y., Law, C. W., Shi, W., et al. (2015). Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* 43, e47. doi:10.1093/nar/gkv007

Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., et al. (2021).Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.*, 71. Wiley Online Library, 209–249. doi:10.3322/caac.21660

Therneau, T. M., and Lumley, T. (2015). Package 'survival. ' *R. Top. Doc.* 128, 28–33.

Wei, R-R., Zhang, M-Y., Rao, H-L., Pu, H-Y., Zhang, H-Z., and Wang, H-Y. (2012). Identification of ADH4 as a novel and potential prognostic marker in hepatocellular carcinoma. *Med. Oncol.*, 29. Springer, 2737–2743. doi:10.1007/s12032-011-0126-3

Wu, T., Hu, E., Xu, S., Chen, M., Guo, P., Dai, Z., et al. (2021). clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *Innovation.* 2, 100141. doi:10.1016/j.xinn.2021.100141

Zhang, Q., Zhong, H., Fan, Y., Liu, Q., Song, J., Yao, S., et al. (2020)., 8. Frontiers, 592. doi:10.3389/fbioe.2020.00592Immune and clinical features of CD96 expression in glioma by *in silico* analysis*Front. Bioeng. Biotechnol.*

Zhang, Q. I., Lou, Y. U., Bai, X-L., and Liang, T-B. (2018). Immunometabolism: A novel perspective of liver cancer microenvironment and its influence on tumor progression. *World J. Gastroenterol.* 24, 3500–3512. doi:10.3748/wjg.v24.i31.3500

Zhong, H., Liu, S., Cao, F., Zhao, Y., Zhou, J., Tang, F., et al. (2021). Dissecting tumor antigens and immune subtypes of glioma to develop mRNA vaccine. *Front. Immunol.* 12, 709986. doi:10.3389/fimmu.2021.709986

Zhou, J-G., Zhong, H., Zhang, J., Jin, S-H., Roudi, R., and Ma, H. (2019). Development and validation of a prognostic signature for malignant pleural mesothelioma. *Front. Oncol.* 9, 78. doi:10.3389/fonc.2019.00078

Zhu, J., Zhou, Y., Wang, L., Hao, J., Chen, R., Liu, L., et al. (2020). CXCL5/CXCL8 is a promising potential prognostic and tumor microenvironment-related cluster in hepatocellular carcinoma. *J. Gastrointest. Oncol.* 11, 1364–1380. doi:10.21037/jgo-20-556

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 13 October 2022
DOI 10.3389/fgene.2022.1014031

OPEN ACCESS

EDITED BY
Shuai Liu,
University of Hawaii at Manoa,
United States

REVIEWED BY
Liangqun Lu,
University of Memphis, United States
Binsen Li,
UCLA Health System, United States
Yuanyuan Leng,
National Cancer Institute (NIH),
United States

*CORRESPONDENCE
Xiaobo Xia,
xbxia21@csu.edu.cn

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

RECEIVED 08 August 2022
ACCEPTED 28 September 2022
PUBLISHED 13 October 2022

CITATION
Li Y, Lu Y, Du K, Yin Y, Hu T, Fu Q, Zhang Y, Wen D, Wu X and Xia X (2022), RNA-sequencing analysis reveals the long noncoding RNA profile in the mouse myopic retina.
Front. Genet. 13:1014031.
doi: 10.3389/fgene.2022.1014031

RNA-sequencing analysis reveals the long noncoding RNA profile in the mouse myopic retina

Yuanjun Li[1,2,3], Ying Lu[1,2,3], Kaixuan Du[1,2,3], Yewei Yin[1,2,3], Tu Hu[1,2,3], Qiuman Fu[1,2,3], Yanni Zhang[1,2,3], Dan Wen[1,2,3], Xiaoying Wu[1,2,3] and Xiaobo Xia[1,2,3]*

[1]Eye Center of Xiangya Hospital, Central South University, Changsha, China, [2]Hunan Key Laboratory of Ophthalmology, Xiangya Hospital, Central South University, Changsha, China, [3]National Clinical Research Center for Geriatric Disorders, Xiangya Hospital, Central South University, Changsha, China

Aim: Myopia is a prevalent public health problem. The long noncoding RNA (lncRNA) mechanisms for dysregulated retinal signaling in the myopic eye have remained elusive. The aim of this study was to analyze the expression profiles and possible pathogenic roles of lncRNAs in mouse form-deprived myopia (FDM) retinas.

Methods: A mouse FDM model was induced and retinas from the FDM right eyes and the contralateral eyes were collected for RNA sequencing. Gene Ontology (GO), Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment, and lncRNA-mRNA coexpression network analyses were conducted to explore the biological functions of the differentially expressed lncRNAs. In addition, the levels of differentially expressed lncRNAs in the myopic retinas were validated by quantitative real-time PCR (qRT–PCR). Fluorescence *in situ* hybridization (FISH) was used to detect the localization of lncRNAs in mouse retinas.

Results: FDM eyes exhibited reduced refraction and increased ocular axial length compared to control fellow eyes. RNA sequencing revealed that there were 655 differentially expressed lncRNAs between the FDM and control retinas. Functional enrichment analysis indicated that the differentially expressed RNAs were mostly enriched in cellular processes, cytokine-cytokine receptor interactions, retinol metabolism, and rhythmic processes. Differentially expressed lncRNAs were validated by qRT–PCR. Additionally, RNA FISH showed that XR_384718.4 (Gm35369) localized in the ganglion cell (GCL) and inner nuclear layers (INL).

Conclusion: This study identified the differential expression profiles of lncRNAs in myopic mouse retinas. Our results provide scientific evidence for investigations of myopia and the development of putative interventions in the future.

KEYWORDS

Long noncoding RNA, messenger RNA, RNA sequencing, myopia, retina

Introduction

Myopia is the most prevalent refractive error and a leading cause of visual impairment worldwide (Dolgin, 2015). In recent decades, there has been a pandemic increase in myopia prevalence, and uncorrected refractive error has become a major public health concern, affecting a large proportion of the world population (Lou et al., 2016; Cheng et al., 2020). Myopia is characterized by excessive elongation in ocular axial length (AL) accompanied by scleral thinning and stretching of other ocular tissues. High myopia (−6.00 D or worse) can lead to severe visual impairments caused by complications such as posterior staphyloma, glaucoma, choroidal neovascularization, myopic retinal degeneration, and detachment (Grossniklaus and Green, 1992; Wu et al., 2000). Multiple factors, including genetic anomalies, intensive near work, insufficient outdoor activities, etc., are involved in the development of myopia (Zhao et al., 2020). Although previous studies have implicated dopamine (Huang et al., 2020), nitric oxide (NO) (Carr and Stell, 2016), retinoic acid (RA) (Wang et al., 2014), glutamate (Guoping et al., 2017), the extracellular matrix (Liu et al., 2017), and, recently, scleral hypoxia (Wu et al., 2018) in the etiology of myopia, the mechanisms and pathogenesis of myopia still require further investigation.

Long noncoding RNAs (lncRNAs) are a class of transcripts greater than 200 nt in length that have little or no protein-coding potential (Carninci et al., 2005; Guttman et al., 2013). LncRNAs have been found to play important roles in a variety of biological processes, including chromatin organization, transcriptional/translational regulation, stem cell maintenance, differentiation, and cell fate reprogramming (Brockdorff et al., 1991; Geisler and Coller, 2013; Flynn and Chang, 2014).

Previous studies have shown that lncRNAs are associated with diverse ocular diseases, including diabetic retinopathy, retinal neovascularization, glaucoma (Zheng M. et al., 2020), cataracts (Tu et al., 2020), proliferative vitreoretinopathy (Ni et al., 2021) and retinoblastoma (Wang H. et al., 2021). RNA sequencing (RNA-seq) of a guinea pig form-deprived myopia (FDM) model and a lens-induced myopia (LIM) model has also suggested that there is differential lncRNA expression in the ocular posterior pole (Geng et al., 2020). lncRNA-associated extracellular matrix (ECM), ECM-receptor interaction, kinase activity, metabolism and multiple functional pathways are involved in myopia pathogenesis (Geng et al., 2020). lncRNAs can affect gene expression by functioning as competitive endogenous RNAs (ceRNAs) with microRNAs (miRNAs), competing with mRNAs for miRNA binding (Cesana et al., 2011). Since miRNA profiling in LIM mice, LIM guinea pigs, and highly myopic patients has suggested the existence of differentially regulated miRNA patterns (Tanaka et al., 2019; Guo et al., 2020; Zhu et al., 2020), lncRNAs might modulate gene expression through RNA interactions and thus regulate myopia.

The retina is a thin layer of complex neural tissue that receives light-stimuli and processes visual signal, and transmitted signal to the sclera. Numerous studies have suggested the retina playing important roles in the pathology in myopia, such as circuiting electrical and chemical synapses (Zhi et al., 2021). Moreover, lncRNAs play pathogenic roles in several retinal diseases, such as the lncRNA XIST and nuclear paraspeckle assembly transcript 1, which play roles in diabetic retinopathy (Li, 2018; Dong et al., 2020). Nevertheless, the detailed expression profiles and pathogenic mechanisms of lncRNAs in myopic retina remain largely elusive. As there were several sequencing studies focusing on the changes in sclera of myopia, while the role of retinal structure in the pathology of myopia remains unclear and complex, the sequencing analysis here hope to lay a foundation for the future study about the mechanisms (especially for the non-coding RNAs) in myopic retina.

Previous myopia studies have built well-established experimental myopic animal model, including FDM and lens-induced myopia (LIM). The two models differ from each other in the methods and behind mechanisms: FDM is induced by deprivation of form vision, while LIM by wearing concave lens to form image behind the retina and to induce excessive accommodation and extension of axial length (Xiao et al., 2014). Study with chicks indicated that the dopaminergic mechanisms mediating the protective effects of brief periods of unrestricted vision might differ for FDM vs. LIM, implying that the two might be different in the growth control mechanisms (Nickla and Totonelly, 2011). Form-deprivation has been well-developed and effective to induce myopia, and extensively used in research into the mechanisms, pathology, sequencing analysis of myopia, thus the current study applied FDM model to explore the lncRNA and mRNA expression pattern in myopic retina. Early study of experimental myopia has tested two mouse strains, C57BL/6 and DBA/2, and concluded that DBA/2J were unaffected by occlusion for 7 or 14 days; prolonged occlusion produces a significant myopic shift in C57BL/6 mice, but not in DBA/2J (Schaeffel et al., 2004). Thus, recent myopic mice studies applied the C57BL/6 strains in the model building. Our process of building the FDM model was almost the same to the procedure of Wu's report with the male C57BL/6, which was started at the age of 3 weeks postnatal, and deprived for 4 weeks by wearing monocular occlusion in the right eyes (Wu et al., 2015).

In the present study, a mouse FDM model was established and characterized. RNA sequencing was applied to compare lncRNA and mRNA expression patterns in retinal tissue between FDM and control mice. Next, the differentially expressed lncRNAs and mRNAs were utilized to conduct pathway enrichment and coexpression network analyses by bioinformatics methods. Furthermore, lncRNA expression was validated by qRT–PCR and localized by RNA FISH. This study aimed to provide experimental evidence of lncRNA profiles in the retina in the context of myopia, which might enable further investigation and the development of a therapy for this ocular disease.

Materials and methods

Animals

Male C57BL/6J mice (3 weeks of age, weight 10–15 g) from the Animal Unit of Central South University were used in this study. Mice were treated under the rules of the Association for Research in Vision and Ophthalmology Statement for the Use of Animals in Ophthalmic and Visual Research. They were housed in an indoor environment with a 12 h light/12 h dark cycle, a temperature of 24 ± 2°C, a luminance of approximately 100–200 lux, and free access to food and water. The animal procedures were approved by the Institutional Animal Care and Use Committee of Central South University (Approval No. 2020sydw0077).

Induction of mouse FDM

Induction of FDM in mice was performed following the procedures described in previous studies (Schaeffel et al., 2004; Wu et al., 2015) with minor modifications. Briefly, on the day of the experiment (Postnatal Day 21–24, weight 10–15 g), male C57BL/6J mice were anesthetized by an intraperitoneal injection of ketamine (90 mg/kg) and xylazine (10 mg/kg), and diffuser eye patches were attached to the skin surrounding the right eye. The diffuser eye patch was made in the laboratory from a plastic tube bottom (diameter: 7.5 mm) mounted on a matching soft latex ring. The eye diffuser was first glued to the periorbital skin around the right eye and then fixed with six to eight stitches (Prolene suture; size 4–0). TobraDex ophthalmic ointment (Alcon, United States) was applied to the eye to protect the cornea from drying. Collars made from plastic foils (outer diameter: 5.5–6.5 cm, inner diameter: 1.0 cm) were fitted around the neck to prevent the mice from removing their diffusers. Food pellets were placed on the floor of the cage to make eating easier. Mice wearing the diffusers were housed in groups of five to six in transparent plastic cages under 12:12 h light-dark conditions (approximately 200 lux illuminance) for 28 days. They were checked every day to ensure the attachment of the diffuser to the eye. A dropped or loose diffuser was reattached. Mice with cataracts or corneal opacity were excluded from the experiments.

Assessment of refraction and axial length

The diffusers were removed after 28 days of FDM treatment, and both eyes were refracted within the same day. The mice were intraperitoneally anesthetized as previously described. One drop of compound tropicamide solution (Santen Pharmaceutical Co., Ltd., JP) was instilled into each eye to ensure a pupil diameter of 1.5 mm. Full pupil dilatation took several minutes. To avoid cataract formation during anesthetization, the mice were refracted immediately (within minutes). The mice were examined using cycloplegic streak retinoscopy by an experienced optometrist. An interocular refractive difference greater than 5 diopters (D) was considered an indicator of successful induction of FDM, and successful models were used in the subsequent experiments.

The axial length of the mice was measured with spectral domain-optical coherence tomography (SD-OCT) under light anesthesia (Banerjee et al., 2020). The anesthetized mouse was placed in front of the light source (Visante OCT 1000, Carl Zeiss Meditec Inc., Dublin, California, United States). The cornea was hydrated with normal saline. The reference arm and focus dial were adjusted simultaneously to a point at which all structures of the eye were in focus. Alignment was confirmed by viewing the radial image of the surface of the eye and adjusting the light source for the central reflection along the horizontal and vertical optical meridians (Figure 1A). Each scan contained an average of 5 images. To measure AL, calipers were placed from the cornea to the retinal pigment epithelium (RPE) border by ImageJ software.

RNA extraction and sequencing

High-throughput sequencing was performed on the mouse retinas (Majorbio Bio-Pham Technology Co., Shanghai, China). There were 12 samples (6 FDM and 6 fellow eye controls), and each contained 3 retinas (total of 18 mice) to ensure that enough RNA was collected (Supplementary Figure.S1). Retinas were collected after 4 weeks of FDM induction. Total RNA was extracted from the retinas using TRIzol® Reagent (Invitrogen, Carlsbad, CA, United States), and genomic DNA was removed by DNase I RNase-free (Takara). The contamination or degradation of RNA was examined by agarose gel (1%) electrophoresis, and the concentration was measured using a NanoDrop-2000 (Thermo Scientific, Wilmington, DE, United States). RNA integrity was then assessed using a 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, United States). Only high-quality RNA samples (OD260/280 = 1.8~2.2, OD260/230 ≥ 2.0, RIN≥7, 28S: 18S ≥ 1.0, >5 μg) were used to construct a sequencing library. Ribosomal RNA depletion was performed using a Ribo-Zero Magnetic Kit (Epicentre Biotechnologies, Madison, WI, United States). A stranded RNA-seq transcriptome library was prepared with a TruSeq™ Stranded Total RNA Kit (Illumina, San Diego, CA, United States). In addition, 3 μg of total RNA was ligated with sequencing adapters with a TruSeq™ Small RNA Sample Prep Kit (Illumina, San Diego, CA, United States). Subsequently, cDNA was synthesized by reverse transcription and amplified with 12 PCR cycles to produce the library. After quantification, the RNA-seq library was sequenced with the HiSeq X Ten (Illumina, San Diego, CA, United States).

Analysis of sequencing data

The raw paired-end reads were trimmed and quality-controlled with SeqPrep (https://github.com/jstjohn/SeqPrep)

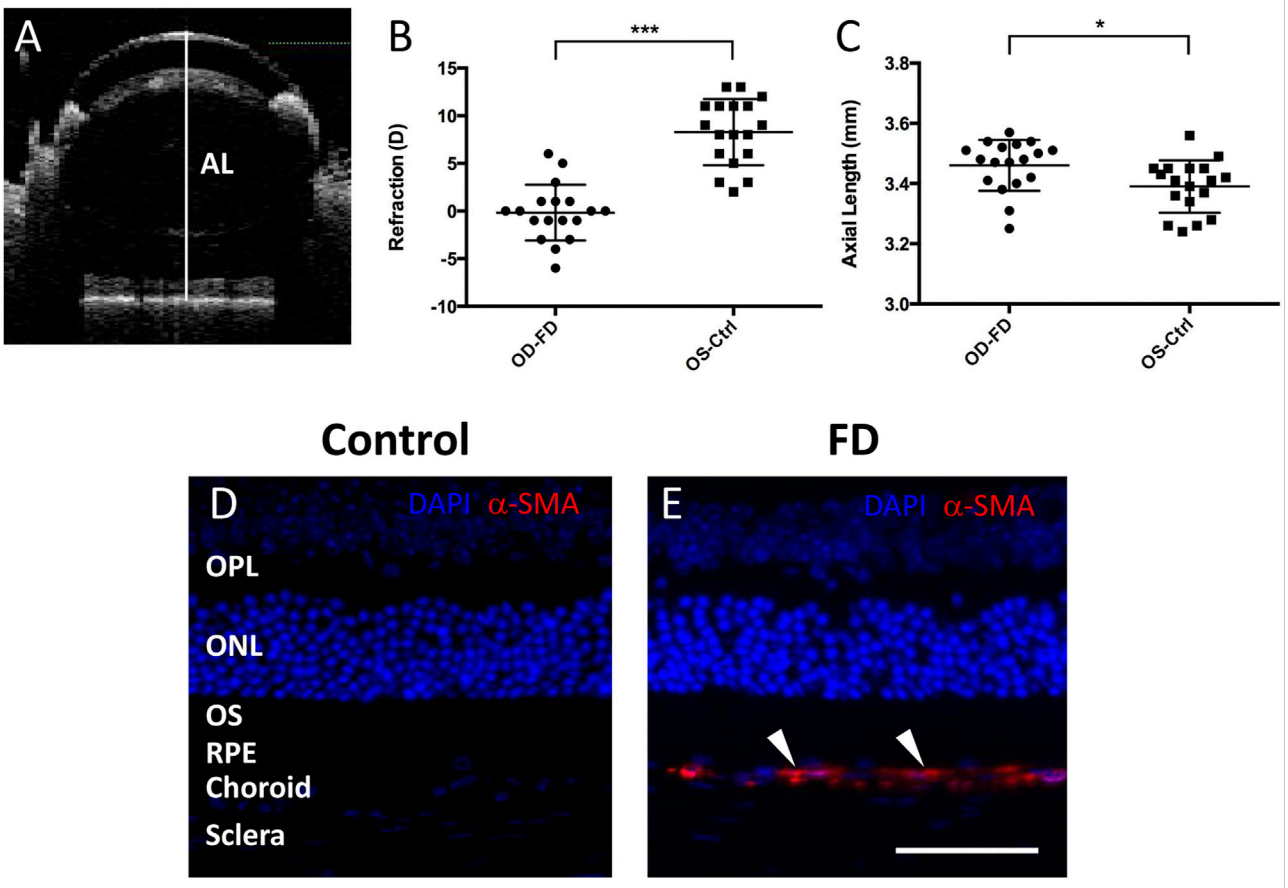


FIGURE 1
Assessment of FDM model mice. **(A)** Representative SD-OCT images of the axial length (AL) in mice. The white line indicates the ocular AL. After 4 weeks of form deprivation, FDM eyes (OD-FD) display significant myopic refraction **(B)** and significant increases in AL **(C)** compared to fellow controls (OS-Ctrl). Immunostaining for α-SMA (*red, arrowhead*) showed greater expression in the scleral and choroidal layers of FDM mouse eyes **(E)** than in those of control eyes **(D)** on Day 28 of form deprivation induction. DAPI (*blue*).

and Sickle (https://github.com/najoshi/sickle). The clean reads were aligned to a mouse reference genome (GRCm38.p6) using HISAT2 (V2.1.0) and using bowtie2 (V2.2.9). The mapped reads of each sample were assembled by StringTie (V1.3.3b) in a reference-based approach. Finally, assembled transcripts were annotated by Cuffcompare program from the Cufflinks (V2.2.1).

Identification of lncRNAs

Known lncRNAs were identified by alignment of the transcripts to the existing reference genome and reported lncRNA sequences in lncRNA-related databases, including NONCODE, Ensembl, NCBI, UCSC, LncRNAdb, GENCODE, GREENC, and LncRNA Disease. Novel lncRNAs were selected step-by-step with criteria. According to the definition and features of lncRNAs, the exclusion criteria for the transcripts were 1) overlapping with known protein-coding genes on the same strand, 2) a fragment count ≤3, 3) a length shorter than 200 nt, 4) an open reading frame (ORF) longer than 300 nt, and 5) an exon number less than 2. Next, the Coding Potential Calculator (CPC), Coding-Non-Coding index (CNCI), Coding Potential Assessment Tool (CPAT), and Pfam Scan were used to filter transcripts with coding potential. The remaining transcripts were considered reliably expressed lncRNAs. Using Cuffcompare in Cufflinks, lncRNAs were classified into intergenic, intronic, and antisense lncRNAs.

Expression analysis of lncRNAs and mRNAs

The quantitative expression of both lncRNAs and mRNAs in each sample was calculated in transcripts per kilobase of exon model per million mapped reads (TPM). lncRNAs with $|\log2(FDM/ctrl)| > 1$ and FDR (Q value) < 0.05 as determined

by EdgeR were considered significantly differentially expressed (DE) transcripts. Volcano plots and hierarchical clustering were used to analyze the DE lncRNAs and mRNAs identified between FDM and fellow control retinas. The predicted potential target genes whose loci were within a 10-kb window upstream or downstream of the given aberrantly expressed lncRNA were considered cis-regulated genes. Other genes in the co-expression network were identified as trans-regulated according to complementary base pairing by LncTar. Also the intaRNA (V2.3.1), RNAplex, RIblast (V1.1.3) were used to predict the target genes. The expressed correlation was calculated between lncRNAs and target genes, and a Pearson correlation coefficient >0.9 identified the target genes.

Quantitative real-time polymerase chain reaction (qRT–PCR)

The expression of ten lncRNAs in the retinas of four mice (control fellow retinas, $n = 4$; FDM retinas, $n = 4$) was assessed using qRT–PCR to verify the accuracy of the high-throughput sequencing results. Eyes were enucleated, and the retinas were immediately dissected. Total RNA was extracted from samples by using TRIzol® Reagent (Invitrogen, Carlsbad, CA, United States), and cDNA was synthesized by using a miScript II RT Kit (Qiagen, Hilden, Germany). Real-time PCR was performed with a miScript SYBR® Green PCR Kit (Qiagen, Hilden, Germany) using a 7500 FAST real-time PCR system (Applied Biosystems, Foster City, CA, United States). The expression of lncRNAs was calculated by the $2^{-\Delta\Delta Ct}$ method. A two-tailed Student's t test was used to compare lncRNA expression between samples from the fellow eyes and those from the FDM eyes in 3 experimental replicates. The forward and reverse primers for lncRNAs are shown in Supplementary Table S1.

Immunofluorescence

Eyes were enucleated and fixed in FAS eyeball fixative solution (G1109-100ML, Servicebio, Wuhan, China) at 4°C for 24 h. The tissues were cryoprotected in 20% sucrose in PBS and embedded in optimal cutting temperature (O.C.T.) compound (Tissue-Tek; Sakura Finetek, Torrance, CA, United States). Twenty-micrometer cryosections were first blocked with serum and immunolabeled with a primary rabbit IgG anti-SMA mAb (1:200; Abcam, Temecula, CA, United States), a primary rabbit IgG anti-RBMPS mAb (1:100; GTX118619 GeneTex, CA, United States), or a primary rabbit IgG anti-calbindin mAb (1:100; Bioworld, Nanjing, China) at 4°C overnight. Then, the sections were reacted with the corresponding fluorescein isothiocyanate-conjugated secondary antibody and finally evaluated by fluorescence microscopy. The slides were stained with DAPI (G1012, Servicebio) for mounting.

RNA fluorescence *in situ* hybridization (FISH)

After 4 weeks of FDM induction, mice were killed, and the eyes were enucleated. The eyes were fixed in FAS eyeball fixative solution (G1109-100ML, Servicebio, Wuhan, China) for more than 24 h. The eyes were dehydrated and embedded in paraffin. The paraffin-embedded eyes were sectioned at 5-μm thickness and baked on microscope slides in a hybridization oven at 62°C for 2 h before *in situ* hybridization. The probes labeled with DIG for lncRNA XR_384718.4 (Gm35369) are shown in Supplementary Table S1. The sections were washed with PBS and blocked with rabbit serum blocking buffer after prehybridization and hybridization. Next, the sections were incubated with mouse anti-DIG-HRP (Jackson ImmunoResearch Labs Inc., United States) for 40 min. After two washes with PBS for 5 min, fresh tyramide signal amplification chromogenic reagent (G3025, Servicebio) was used for the chromogenic reaction for 5 min. The sections were stained with DAPI (G1012, Servicebio) for 8 min and mounted. Photographs were obtained with a fluorescence microscope (Nikon Eclipse CI, Japan).

Statistical analysis

The data are reported as the mean ± standard error of the mean (SEM). Graphs were constructed using GraphPad Prism 9.0 software. Statistical Program for the Social Sciences 20.0 software (IBM SPSS Inc., New York, NY) was used for statistical analysis of the biometric parameters and qRT–PCR results. The ocular biometric parameters in the myopia-induced right eyes (OD) and the untreated contralateral left fellows (OS) were compared using the paired t test. The sample sizes are reported in the results. Overall comparisons of these indices in the right eyes among the experimental groups were performed with two-way analysis of variance (ANOVA) or one-way ANOVA, and pairwise comparisons were performed with Tukey's post hoc test. p values <0.05 were considered to indicate significance.

Results

Establishment and analysis of FDM mice

Visual form deprivation in C57BL/6J mice was started at approximately 3 weeks (P21, weight 12.82 ± 2.22 g, $n = 18$). After 28 days of form deprivation in photopic conditions, a myopic shift in ocular measurements was observed compared to that in the left control eyes (Figure 1B). The refraction (in diopters, D) in deprived eyes (−0.17 ± 2.94 D) was shifted toward myopia compared to that in the controls (+8.28 ± 3.48 D); the

interocular differences in refraction between the right and left eyes (OD-OS, −8.44 ± 4.93 D, $p < 0.001$) were statistically significant. The axial length (mm) was also elongated in the deprived eyes (3.46 ± 0.09 mm) relative to that in the controls (3.39 ± 0.09 mm), with a significant interocular difference (OD-OS, 0.07 ± 0.09 mm, $p < 0.05$) (Figure 1C). To further confirm the characteristics of FDM eyes, the expression of alpha-smooth muscle actin (α-SMA), a myofibroblast marker, was examined with immunostaining in the retina and sclera. In accordance with previous studies (Wu et al., 2015; Yuan et al., 2018), α-SMA was more highly expressed in the FDM sclera and choroid areas than in these areas in the control fellows after 4 weeks of induction (Figures 1D,E). These results suggested the development of significant FDM in the goggled mouse eyes.

Sequencing data summary

Libraries were constructed from retinal tissue samples from FDM eyes ($n = 6$; each sample consisted of three retinas from three FDM eyes to ensure that the RNA amount was sufficient) and control fellows ($n = 6$; each sample consisted of three retinas from the control fellow eyes of the FDM eyes) and subjected to sequencing analysis. RNA-seq yielded 501,329,752 and 547,167,536 raw reads from the FDM and control groups, respectively. Low-quality reads were filtered from the raw reads, and high-quality clean reads and clean bases were obtained. In total, 498,008,508 and 543,470,042 clean reads were retained for the FDM and control groups, respectively. The Q20 and Q30 quality scores of the clean data were higher than 90%, indicating the reliability of the RNA sequencing results. The clean reads were mapped to a mouse reference genome (GRCm38. p6, Ensembl) with a total mapping percentage ranging from 96.29 to 97.15%. Detailed data on the quality results are shown in Supplementary Table S2.

Identification and classification of lncRNAs in the retinas of mice

According to the mouse reference genome and related databases (NONCODE, Ensembl, and NCBI), 19,443 known lncRNAs were identified. Filtering and overlapping analyses in four programs (PfamScan, CPC, CPAT, and CNCI) identified a total of 561 novel lncRNAs (Supplementary Figure S2). Mapping of the reads to genomic regions with RSeQC-2.3.6 revealed the distributions of the lncRNAs from both FDM and control eyes in five areas: the 5′UTR (0.81%), intergenic regions (1.97%), the 3′UTR (14.06%), introns (24.57%), and the coding sequence (CDS, 58.58%) (Figure 2A). Based on the relative chromosomal position of the coding gene, the novel lncRNAs were classified into five categories: 5 were sense intronic overlapping lncRNAs (0.6%), 234 were intergenic lncRNAs (27.0%), 160 were antisense lncRNAs (18.5%), 416 were sense exonic overlapping lncRNAs (48.0%), and 51 were bidirectional lncRNAs (5.9%) (Figure 2B). Chromosomal distribution analysis of the lncRNAs showed that chromosomes 12, 11, 2 and 9 contained relatively higher amounts of lncRNAs than the other chromosomes (Figure 2C).

Differential expression patterns of lncRNAs and mRNAs in FDM

Among the 20,309 lncRNAs (19,443 from reference databases, 866 of novel) obtained from high-throughput sequencing, hierarchical clustering analysis showed that there were 655 differentially expressed lncRNAs between the FDM and control retinas, of which 296 were upregulated and 359 were downregulated. The top 20 differentially upregulated (Table 1) and downregulated lncRNAs (Table 2) between the FDM and control retinas, such as XR_003956022.1, NR_045075.1, and Oip5os1, are listed according to the statistical significance (p value) and log2FC. Among the 19,137 mRNAs (18,580 from reference databases, 557 of novel) obtained from RNA-seq, there were 478 differentially expressed mRNAs (206 upregulated and 272 downregulated) between the FDM and control. The top 20 differentially up- and downregulated mRNAs with the largest fold changes are also displayed (Tables 3, 4). Heatmaps and volcano plots were used to visualize the differentially expressed lncRNAs and mRNAs between the two groups (Figure 3).

Gene ontology and kyoto encyclopedia of genes and genomes analysis

Target genes of the differential lncRNAs were predicted by bioinformatics approaches, and the prediction results are illustrated in Supplementary Table S3. The differentially expressed mRNAs underwent GO and KEGG enrichment analyses. GO enrichment analysis examined the gene functions in three categories: the cellular component (CC), biological process (BP), and molecular function (MF) categories. The top 20 enriched GO terms of the significantly upregulated mRNAs are presented and included the cellular process (ontology: BP, GO: 0009987), cellular anatomical entity (ontology: CC, GO: 0110165), and binding (ontology: MF, GO: 0005488) terms (Figure 4A). The downregulated mRNAs were related to some different terms, such as rhythmic process (ontology: BP, GO: 0048511) and structural molecule activity (ontology: MF, GO: 0005198) (Figure 4B).

GO enrichment analyses were performed on the differentially expressed mRNAs. The top 20 GO and KEGG enrichment of the mRNAs are shown in Figures 5A,B, respectively. The terms “sensory perception of chemical stimulus” and “G-protein coupled receptor signaling pathway” were among the top

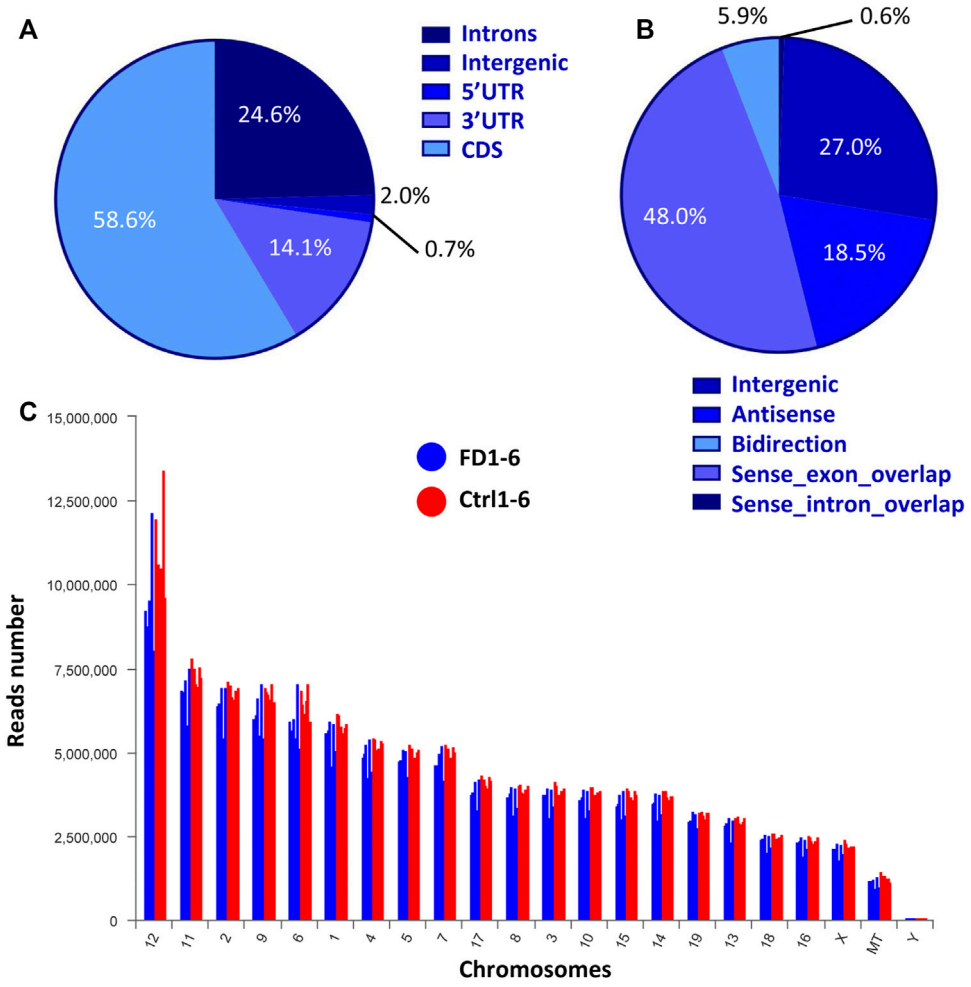


FIGURE 2
Expression characteristics of lncRNAs from both FDM and control eyes. **(A)** Percentages of reads mapped to genomic regions. **(B)** Percentages of lncRNAs classified into different groups. **(C)** Chromosomal distribution of lncRNA numbers.

enriched in the GO enrichment analysis (Figure 5A). Multiple pathways, such as cytokine-cytokine receptor interactions, retinol metabolism, olfactory transduction, metabolism of xenobiotics by cytochrome P450, chemical carcinogenesis, tyrosine metabolism, and proteasome, are likely involved in FDM (Figure 5B). The expression levels of altered genes and their related enriched KEGG pathways are illustrated in the KEGG chord plot (Supplementary Figure S3). For example, Ccl21d, Epo, Ccl22, Tnfrsf17, Il22ra1, Pf4, Ccl27a, Cd70, and Il21r were associated with cytokine-cytokine receptor interaction (pathway ID: map04060), while Bco1, Cyp3a13, Rdh9, Cyp2a5, and Adh7 were associated with retinol metabolism (pathway ID: map00860). These results indicate that the differentially expressed retinal lncRNAs participate in a variety of biological mechanisms and are intrinsically associated with form-deprived myopia.

LncRNA-mRNA coexpression network

To uncover the possible interactions between lncRNAs and mRNAs in the FDM retina, the lncRNA-mRNA coexpression relationship was identified based on the top differential lncRNAs and mRNAs. After screening (correlation coefficient, Corr >0.95, p value <0.05), a lncRNA-mRNA coexpression network, which consisted of 117 nodes (lncRNAs and mRNAs) and 755 edges connecting the nodes, was constructed (Figure 6). The interactive mRNAs included Vcan, Cmip, Trem2, Dmtf1, Cd59b, Shcpb1, Vmn2r89, and others, which might play regulatory roles in myopic biological processes. In particular, the downregulated lncRNA XR_869563.3 and the upregulated lncRNA NR_045075.1 were connected by a large number of mRNAs, which suggests that these dysregulated lncRNAs might be involved in additional functional pathways and mechanisms in the myopic retina.

TABLE 1 Top 20 Up-regulated LncRNAs ($p < 0.05$) Between FDM and Fellow Groups. Log2FC, Log2FC of (FDM retina/fellow control retina).

Transcript ID	Gene ID	Gene Name	Gene description	Log2FC	*p* value
XR_003956022.1	ENSMUSG00000100783	2310047D07Rik	RIKEN cDNA 2310047D07 gene	7.793	9.66E-03
NR_045075.1	ENSMUSG00000086363	A330102I10Rik	RIKEN cDNA A330102I10 gene	7.379	1.76E-03
chr4:21834991-21837868	ENSMUSG00000040455	Usp45	ubiquitin specific petidase 45	7.182	2.96E-03
ENSMUST00000147425	ENSMUSG00000085438	Oip5os1	Opa interacting protein 5, opposite strand 1	6.852	6.60E-03
chr16:32388756-32392334	ENSMUSG00000053774	Ubxn7	UBX domain protein 7	6.666	1.29E-02
chr6:92167828-92169115	ENSMUSG00000005893	Nr2c2	nuclear receptor subfamily 2, group C, member 2	6.249	6.69E-03
XR_871884.3	ENSMUSG00000112110	Gm15608	predicted gene 15608	6.235	1.26E-03
ENSMUST00000181960	ENSMUSG00000097290	1300002E11Rik	RIKEN cDNA 1300002E11 gene	6.172	4.21E-03
XR_004934313.1	ENSMUSG00000109233	Gm44866	predicted gene 44866	6.145	1.48E-02
chr7:101793411-101795506	ENSMUSG00000001829	Clpb	ClpB caseinolytic peptidase B	6.026	1.58E-02
chr13:114155322-114157022	ENSMUSG00000042348	Arl15	ADP-ribosylation factor-like 15	5.933	3.65E-03
chr17:88487362-88490321	ENSMUSG00000034998	Foxn2	forkhead box N2	5.837	1.58E-02
XR_004941494.1	ENSMUSG00000104178	Gm9916	predicted gene 9916	5.645	4.91E-03
XR_004940873.1	ENSMUSG00000086405	9330198N18Rik	RIKEN cDNA 9330198N18 gene	5.637	4.93E-03
chr1:34446795-34449308	ENSMUSG00000026127	Imp4	IMP4, U3 small nucleolar ribonucleoprotein	5.481	1.82E-02
ENSMUST00000128131	ENSMUSG00000086290	Snhg12	small nucleolar RNA host gene 12	5.464	9.49E-03
XR_381591.4	ENSMUSG00000112412	Gm35239	predicted gene, 35239	5.380	1.93E-02
ENSMUST00000152024	ENSMUSG00000086587	Gm11837	predicted gene 11837	5.352	2.69E-03
XR_001783522.3	ENSMUSG00000085317	Gssos2	glutathione synthase, opposite strand 2	5.169	5.48E-03
XR_880469.2	ENSMUSG00000090006	Gm16227	predicted gene 16227	5.166	2.14E-02

TABLE 2 Top 20 down-regulated LncRNAs ($p < 0.05$) between FDM and fellow groups.

Transcript ID	Gene ID	Gene Name	Gene description	Log2FC	*p* value
chr6:13086757-13089260	ENSMUSG00000029571	Tmem106b	transmembrane protein 106B	−6.240	1.57E-02
XR_004935657.1	ENSMUSG00000103640	Gm31406	predicted gene, 31406	−6.161	7.98E-03
chr9:110981867-110984062	ENSMUSG00000032495	Lrrc2	leucine rich repeat containing 2	−6.031	4.86E-03
XR_003955073.1	ENSMUSG00000087366	Junos	jun proto-oncogene, opposite strand	−5.990	2.54E-04
XR_386631.3	ENSMUSG00000117692	Gm50114	predicted gene, 50114	−5.928	1.73E-02
ENSMUST00000126380	ENSMUSG00000086290	Snhg12	small nucleolar RNA host gene 12	−5.873	2.33E-03
NR_040262.1	ENSMUSG00000044471	Lncpint	Trp53 induced transcript	−5.825	8.89E-03
XR_881968.1	ENSMUSG00000108711	Gm38991	predicted gene, 38991	−5.806	1.40E-02
XR_871885.3	ENSMUSG00000112110	Gm15608	predicted gene 15608	−5.731	4.93E-03
XR_003948926.1	ENSMUSG00000112110	Gm15608	predicted gene 15608	−5.723	1.15E-03
ENSMUST00000179924	ENSMUSG00000079179	Rab10os	RAB10, RAS oncogene family, opposite strand	−5.656	4.90E-03
XR_004934966.1	ENSMUSG00000110559	Gm26843	predicted gene, 26843	−5.612	1.04E-02
chr8:61504964-61506681	ENSMUSG00000031641	Cbr4	carbonyl reductase 4	−5.567	4.08E-03
XR_004942277.1	ENSMUSG00000086953	Aknaos	AT-hook transcription factor, opposite strand	−5.489	1.98E-02
chr7:113928073-113932081	ENSMUSG00000038156	Spon1	spondin 1, (f-spondin) extracellular matrix protein	−5.482	2.03E-02
ENSMUST00000238598	ENSMUSG00000097129	4930507D05Rik	RIKEN cDNA 4930507D05 gene	−5.468	6.20E-03
ENSMUST00000238778	ENSMUSG00000026736	4930426L09Rik	RIKEN cDNA 4930426L09 gene	−5.462	1.07E-02
chr1:15716879-15719068	ENSMUSG00000092083	Kcnb2	potassium voltage gated channel, Shab member 2	−5.370	1.19E-02
chr18:36299103-36299691	ENSMUSG00000110185	Igip	IgA inducing protein	−5.364	1.44E-02
XR_004941495.1	ENSMUSG00000104178	Gm9916	predicted gene 9916	−5.295	6.54E-03

TABLE 3 Top 20 upregulated mRNA ($p < 0.05$) between FDM and fellow groups.

Transcript ID	Gene ID	Gene Name	Gene description	Log2FC	*p* value
ENSMUST00000021662	ENSMUSG00000021236	Entpd5	ectonucleoside triphosphate diphosphohydrolase 5	9.540	4.56E-03
ENSMUST00000211820	ENSMUSG00000037270	4932438A13Rik	RIKEN cDNA 4932438A13 gene	9.436	1.39E-02
ENSMUST00000101375	ENSMUSG00000057113	Npm1	nucleophosmin 1	9.386	9.60E-04
ENSMUST00000095172	ENSMUSG00000034390	Cmip	c-Maf inducing protein	9.337	3.42E-03
ENSMUST00000097785	ENSMUSG00000026131	Dst	dystonin	9.328	2.92E-03
ENSMUST00000228412	ENSMUSG00000002496	Tsc2	TSC complex subunit 2	9.061	2.79E-02
ENSMUST00000132158	ENSMUSG00000026696	Vamp4	vesicle-associated membrane protein 4	8.938	8.63E-03
ENSMUST00000082170	ENSMUSG00000074505	Fat3	FAT atypical cadherin 3	8.701	1.89E-03
ENSMUST00000117805	ENSMUSG00000048240	Gng7	guanine nucleotide binding protein, gamma 7	8.681	2.95E-02
ENSMUST00000068367	ENSMUSG00000032396	Dis3l	DIS3 like exosome 3′–5′ exoribonuclease	8.422	3.03E-02
ENSMUST00000227200	ENSMUSG00000048038	Ccdc187	coiled-coil domain containing 187	8.308	3.35E-03
ENSMUST00000204198	ENSMUSG00000001632	Brpf1	bromodomain and PHD finger containing, 1	8.063	9.12E-03
ENSMUST00000238849	ENSMUSG00000068876	Cgn	cingulin	7.970	3.14E-02
ENSMUST00000061970	ENSMUSG00000031337	Mtm1	X-linked myotubular myopathy gene 1	7.904	1.85E-02
ENSMUST00000125774	ENSMUSG00000026426	Arl8a	ADP-ribosylation factor-like 8A	7.899	4.95E-03
ENSMUST00000233357	ENSMUSG00000117098	Gm49909	predicted gene, 49909	7.841	1.19E-02
ENSMUST00000111372	ENSMUSG00000040687	Madd	MAP-kinase activating death domain	7.786	1.98E-02
ENSMUST00000163854	ENSMUSG00000026074	Map4k4	mitogen-activated protein 4 kinase 4	7.753	1.08E-02
ENSMUST00000144936	ENSMUSG00000079020	Slc45a4	solute carrier family 45, member 4	7.732	1.86E-02
ENSMUST00000166592	ENSMUSG00000031691	Tnpo2	transportin 2 (importin 3, karyopherin beta 2b)	7.729	1.01E-02

TABLE 4 Top 20 downregulated mRNA ($p < 0.05$) between FDM and fellow groups.

Transcript ID	Gene ID	Gene Name	Gene description	Log2FC	*p* value
ENSMUST00000137823	ENSMUSG00000056342	Usp34	ubiquitin specific peptidase 34	−9.792	3.21E-03
ENSMUST00000085044	ENSMUSG00000006676	Usp19	ubiquitin specific peptidase 19	−9.352	5.12E-03
ENSMUST00000003191	ENSMUSG00000024070	Prkd3	protein kinase D3	−9.124	1.86E-02
ENSMUST00000234851	ENSMUSG00000061130	Ppm1b	protein phosphatase 1B, beta isoform	−8.760	1.08E-02
ENSMUST00000107417	ENSMUSG00000042626	Shc1	src homology 2 transforming protein C1	−8.704	3.18E-03
ENSMUST00000019246	ENSMUSG00000019102	Aldh3a1	aldehyde dehydrogenase family 3, A1	−8.484	9.23E-03
ENSMUST00000181981	ENSMUSG00000045659	Plekha7	pleckstrin homology family A member 7	−8.459	8.64E-04
ENSMUST00000084301	ENSMUSG00000028649	Macf1	microtubule-actin crosslinking factor 1	−8.457	3.38E-03
ENSMUST00000238066	ENSMUSG00000052387	Trpm3	transient receptor potential channel M3	−8.380	9.90E-03
ENSMUST00000212478	ENSMUSG00000036180	Gatad2a	GATA zinc finger domain containing 2A	−8.264	5.04E-03
ENSMUST00000099149	ENSMUSG00000025453	Nnt	nicotinamide nucleotide transhydrogenase	−8.218	2.02E-02
ENSMUST00000112990	ENSMUSG00000000266	Mid2	midline 2	−8.145	2.00E-02
ENSMUST00000076140	ENSMUSG00000033577	Myo6	myosin VI	−8.088	2.07E-02
ENSMUST00000095012	ENSMUSG00000028883	Sema3a	semaphorin 3A	−7.839	2.08E-02
ENSMUST00000169854	ENSMUSG00000056917	Sipa1	signal-induced proliferation associated 1 Symbol; Acc:MGI:107576	−7.783	2.13E-02
ENSMUST00000224209	ENSMUSG00000037824	Tspan14	tetraspanin 14	−7.778	1.71E-03
ENSMUST00000113530	ENSMUSG00000030087	Klf15	Kruppel-like factor 15	−7.753	1.14E-03
ENSMUST00000001043	ENSMUSG00000001017	Chtop	chromatin target of PRMT1	−7.696	3.60E-02
ENSMUST00000114617	ENSMUSG00000031337	Mtm1	X-linked myotubular myopathy gene 1	−7.648	1.21E-02
ENSMUST00000164039	ENSMUSG00000043531	Sorcs1	sortilin-related VPS10 domain containing receptor 1	−7.645	3.60E-02

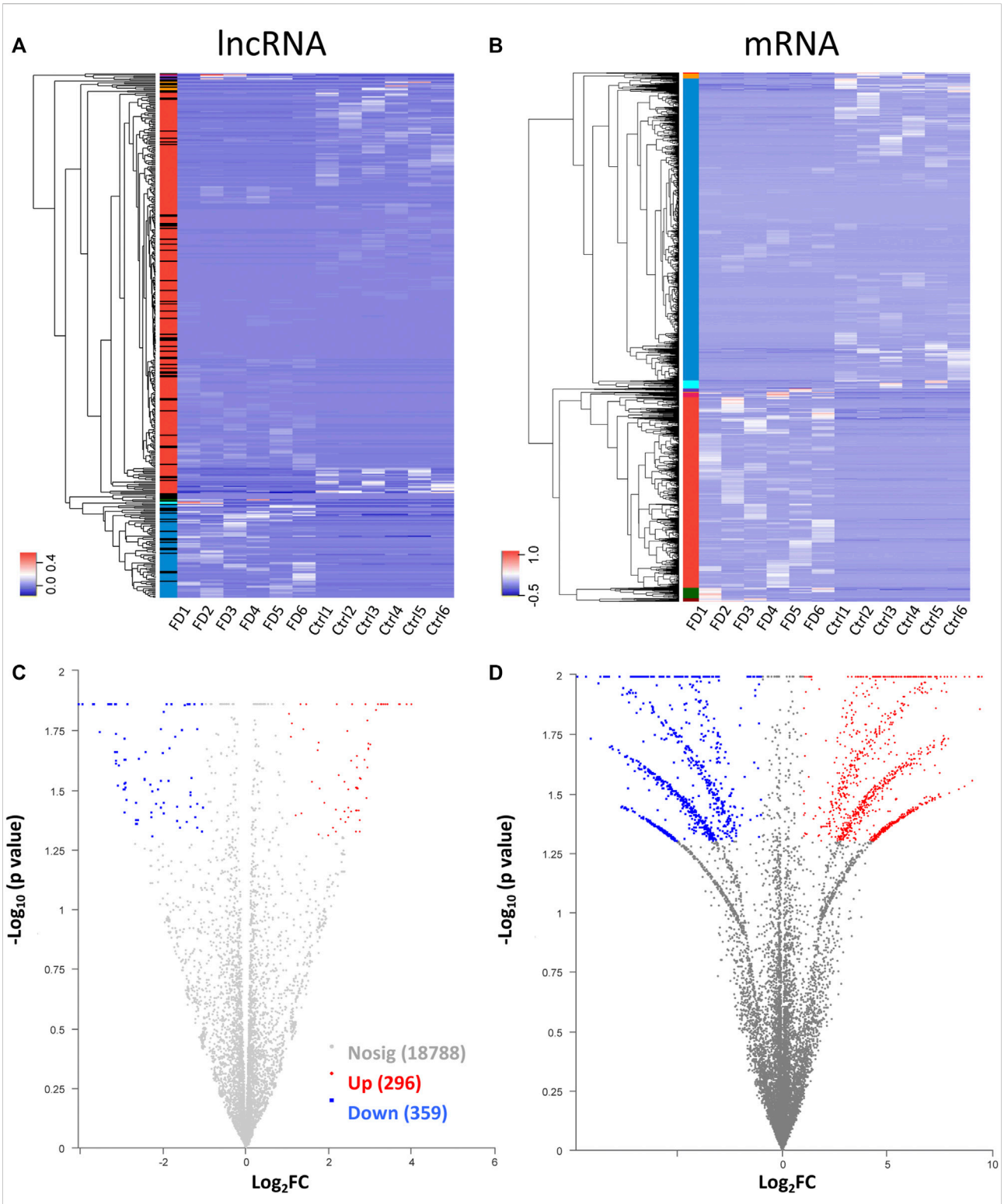


FIGURE 3
Heatmap for hierarchical clustering of lncRNAs **(A)** and mRNAs **(B)** from 12 samples (six for FDM and 6 for control). The colors in the panel represent the relative expression levels: blue and red represent low and high expression levels, respectively. Volcano plot of lncRNAs **(C)** and mRNAs **(D)**. Red/blue dots represent significantly up//downregulated RNAs (FC ≥ 2.0, $p < 0.05$). Gray indicates no differential expression.

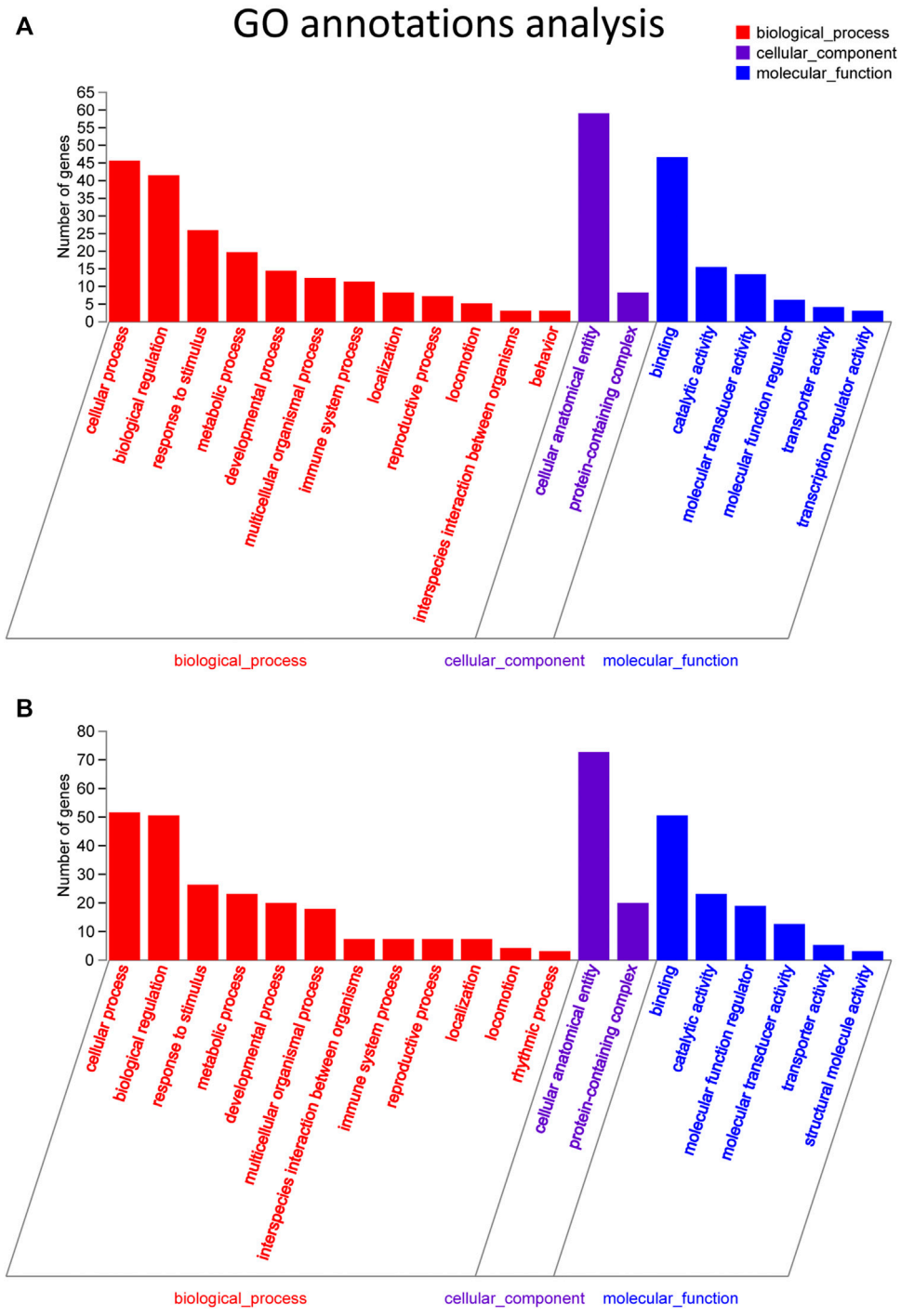


FIGURE 4
GO pathway analysis in FDM retinas. **(A)** GO analysis of significantly upregulated mRNAs. The top 20 GO terms in the biological process (BP), cellular component (CC) and molecular function (MF) categories are shown for the upregulated mRNAs. **(B)** GO analysis of significantly downregulated mRNAs. Y-axis, number of genes included in a single annotation; X-axis, GO pathway terms. The circle size represents the gene number. The FDR value is indicated by the color gradient. FDR <0.05 indicates significant enrichment of the functional pathway.

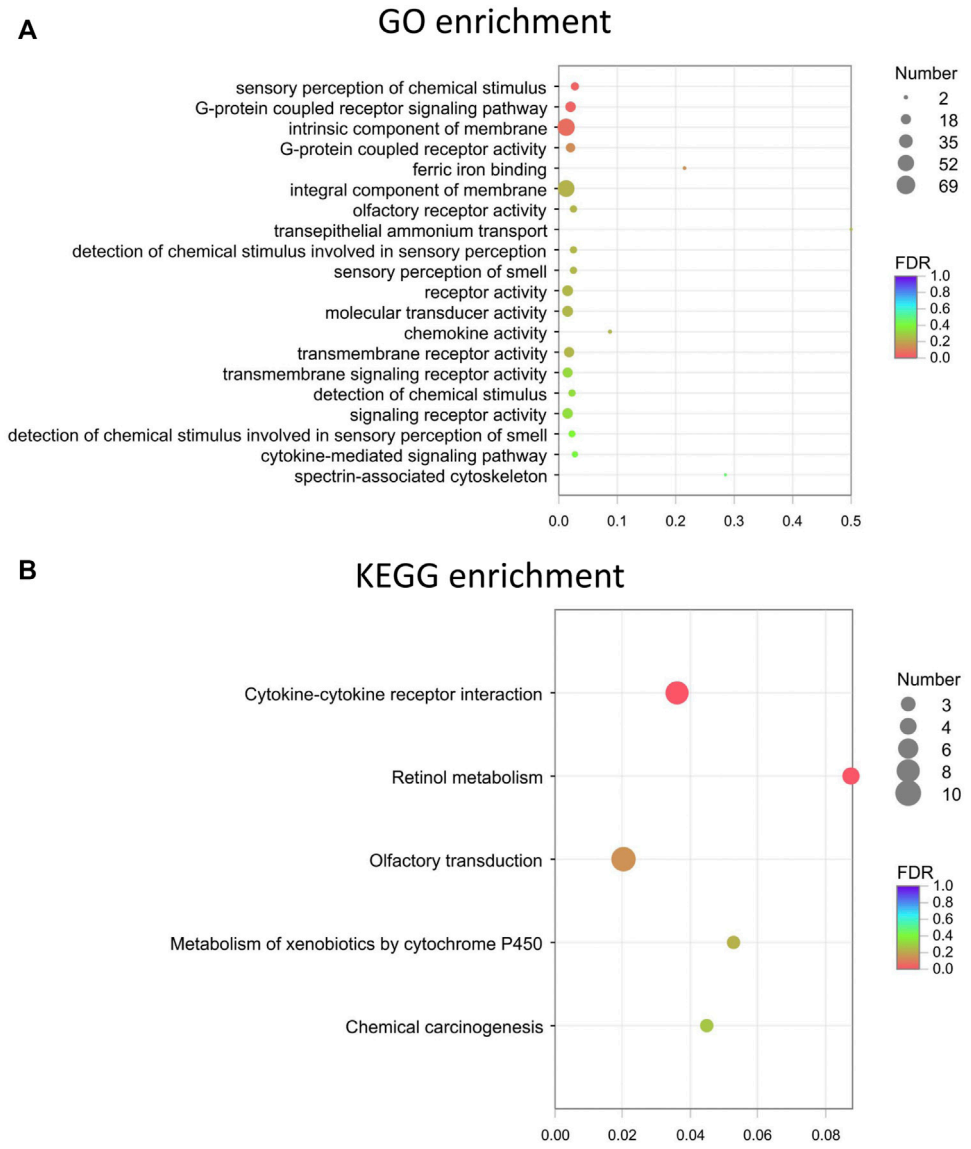


FIGURE 5
The top 20 GO and KEGG enrichment of the mRNAs. **(A)** Top 20 GO terms for the differentially expressed mRNAs. **(B)** Top KEGG enrichment for the differentially expressed mRNAs. Dot color: towards blue, FDR towards 1.0; towards red, FDR toward 0.0. Dot size: the number of core genes within the pathway.

Validation of the expression levels of lncRNAs by qRT–PCR

To verify the expression levels of lncRNAs in myopic retinas, we selected four lncRNAs for qRT–PCR based on the following criteria: A p value <0.05 with top-ranked fold change (FC). The LncRNAs were chose for qRT-PCR validation not only based on their expression fold changes, but also according to possible specific interesting correlations, and the lncRNAs records in NCBI databases. For instance, coexpression analysis of XR_377255.2 (log2FC = 5.028) showed correlation with Vcan (Versican), a critical extracellular matrix regulator of immunity and inflammation (Figure 6); while CMIP, C-Maf-inducing protein, was correlated with downregulation of the lncRNA XR_866459.4 (log2FC = 3.430). The XR_003955073.1 (log2FC = −5.990) was in the top 20 down-regulated list. The XR_384718.4 was a *Mus musculus* predicted gene (ncRNA), 35369 (Gm35369), transcript variant X2 by NCBI database. It is not in the top down-regulated list but still shows

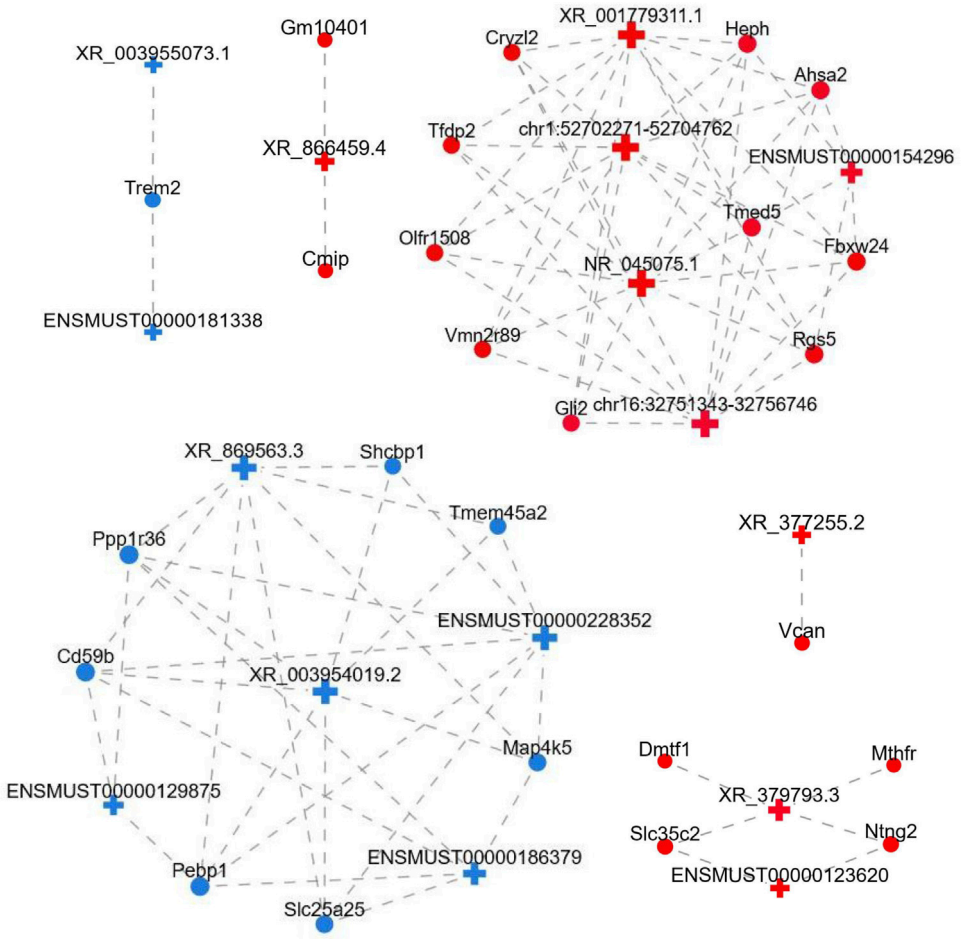


FIGURE 6
LncRNA-mRNA coexpression network. LncRNAs and mRNAs with Corr >0.95 and *p* value <0.05 were selected to construct the network. The network shows the interactions among the lncRNAs and their potential regulated coding genes. Red, upregulated; blue, downregulated; cross, lncRNA; circle, mRNA; gray dotted line, correlation relationship.

a log2FC of −2.316 (downregulated in myopic retinas as compared to the control). We also hope to test the DE lncRNA with moderate dysregulated expression level. Among the four lncRNAs, XR_377255.2, XR_866459.4 were upregulated, while XR_003955073.1, and XR_384718.4 were downregulated in FDM retinas according to the RNA sequencing data analysis (Supplementary Figure S4; Supplementary Table S4). Increased expression levels of XR_377255.2 (gene name: Gm15411) (p = 0.0149) and reduced expression of XR_003955073.1 (gene name: Junos), XR_384718.4 (gene name: Gm35369) and XR_866459.4 (gene name: Gm39857) were observed in FDM mouse retinas (p = 0.0451, p = 0.0217, and p = 0.0423, respectively; Figure 7). Among them, the changes in three (except for XR_866459.4) were consistent with the sequencing results. Thus, we validated the expression changes of XR_377255.2, XR_003955073.1, and XR_384718.4 by qRT-PCR.

Localization of differentially expressed lncRNAs in the retina by RNA FISH

LncRNAs act in different ways to interfere with cellular physiology, depending on their subcellular locations. For the retina, it is also important to identify the layers and cell types in which the targets are located. To further investigate the newly discovered targets in the retina, we conducted a preliminary localization experiment to identify the retinal layers of lncRNAs and their subcellular expression. FISH assay of the lncRNA XR_384718.4 (gene name: Gm35369) in myopic mouse eyecups showed that Gm35369 preferentially localized mostly in the GCL and INL (Figure 8A). Although qRT–PCR showed that it was downregulated in the myopic retina, Gm35369 localized in similar patterns in the retinas of both groups, as the difference between the myopia and control groups was too subtle to be observed by *in situ* hybridization (data not

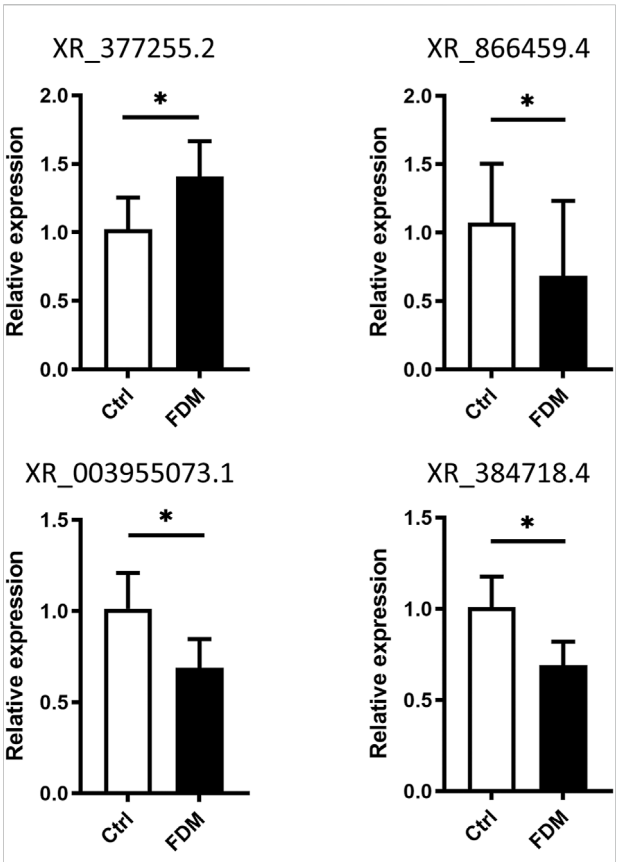


FIGURE 7
Validation of differential lncRNA expression by qRT–PCR. The relative expression of the lncRNAs XR_377255.2, XR_866459.4, XR_003955073.1, and XR_384718.4 in the retinas from FDM eyes and fellow control eyes is shown. For each group, $n = 6$. *, $p < 0.05$, Student t test.

shown). Further colocalization study revealed the overlap between the signal of Gm35369 and the retinal ganglion cell (RGC) marker RBPMS (Figure 8B) as well as the horizontal cell marker calbindin (CaBP) (Figure 8C). The signal was apparent in both the nucleus and cytoplasm. These data indicated that the lncRNA Gm35369 was located mainly in RGCs and horizontal cells.

Discussion

In this study, we analyzed the differential expression patterns of lncRNAs and mRNAs in FDM mouse retinas. The results from RNA FISH localized the target lncRNAs in specific retinal layers and cell types. We hope that this evidence might lay a foundation for future research on myopia. Previously, miRNA profiling of the whole eyes, retinas, and sclerae of mice (Tkatchenko et al., 2016; Mei et al., 2017) and lncRNA-mRNA sequencing of the ocular posterior poles of guinea pigs with experimental myopia have been reported (Geng et al., 2020). Our study focused on the retina rather than the whole eyeball or posterior poles, which reduced the possible heterogeneity from different tissues.

A mouse FDM model was induced for the sequencing analysis in this study, similar to the previously reported experimental myopia model. The Wu's results showed a myopic shift in the deprived eye with the refractive difference (OD 1.341 ± 0.298 D–OS 6.440 ± 0.292 D) of −5.099 ± 0.239 D ($p < 0.001$), which is slightly smaller than ours (−8.44 ± 4.93 D, $p < 0.001$). In another mice study with shorter deprivation of 10 days, the deprived eyes were induced to myopia of −6.93 ± 2.44 D ($p < 0.000001$) compared to the contralateral control eyes (Tkatchenko et al., 2016); and in a study of 2 weeks deprivation, the refraction difference was about −7 D (Wu et al., 2018). Taken together, these observations suggest a significant myopic shift in the refractive error, even with some variations in each report (possibly due to the animals, measuring method, or equipment), would be induced after monocular visual form deprivation ranging from 10 days to 4 weeks.

In sequencing analysis of eye, a sample size of three replicates or more than three replicates could be justifiable, as the differential gene expression from RNA-seq was successfully validated by qRT-PCT. Moreover, due to the small volume and low RNA content of ocular tissue, some studies pooled samples to get enough RNA for sequencing. For instance, Tkatchenko et al. analyzed the microRNA expression profiling in the retina and sclera of FDM mice by microarray, with the small sample size of only three replicates (3 eyes pooled together per replicate) in parallel (Tkatchenko et al., 2016). Similarly, four sclerae were pooled to form one sample in the bulk transcriptome sequencing (Zhao et al., 2021). Vocale *et al.* identified ligand-gated chloride efflux channels as a major pathway contributing to chick FDM using RNA-seq and gene set enrichment analysis, with four replicates (4 chicks) per condition (Vocale et al., 2021); and sample size of three or four replicates (3 or 4 guinea pig eyes) in the RNA-seq study to investigate the gene expression and pathways in sclera was also reported (Srinivasalu et al., 2018; Zeng et al., 2021). The current study included six replicates per experimental condition (FDM or contralateral control, with 3 retinas pooled per replicate), much more than the previous reported RNA-seq datasets, hoping to increase the confidence of our RNA-seq data. Besides, the validation of expression level of lncRNAs by qRT-PCR and sequencing showed a similar direction between the two techniques.

After the induction of FDM, the refractive error was measured by cycloplegic streak retinoscopy in the current study, which would possibly produce measurement error. Refraction in previous study was assessed by various methods, including automated eccentric infrared photorefractor (Schaeffel et al., 2004; Wu et al., 2018) and streak retinoscopy by optometrist (Huang et al., 2019; Zeng et al., 2021). In our study, the optometrist and the researchers who bled and built the model were different, and optometrist was masked to the

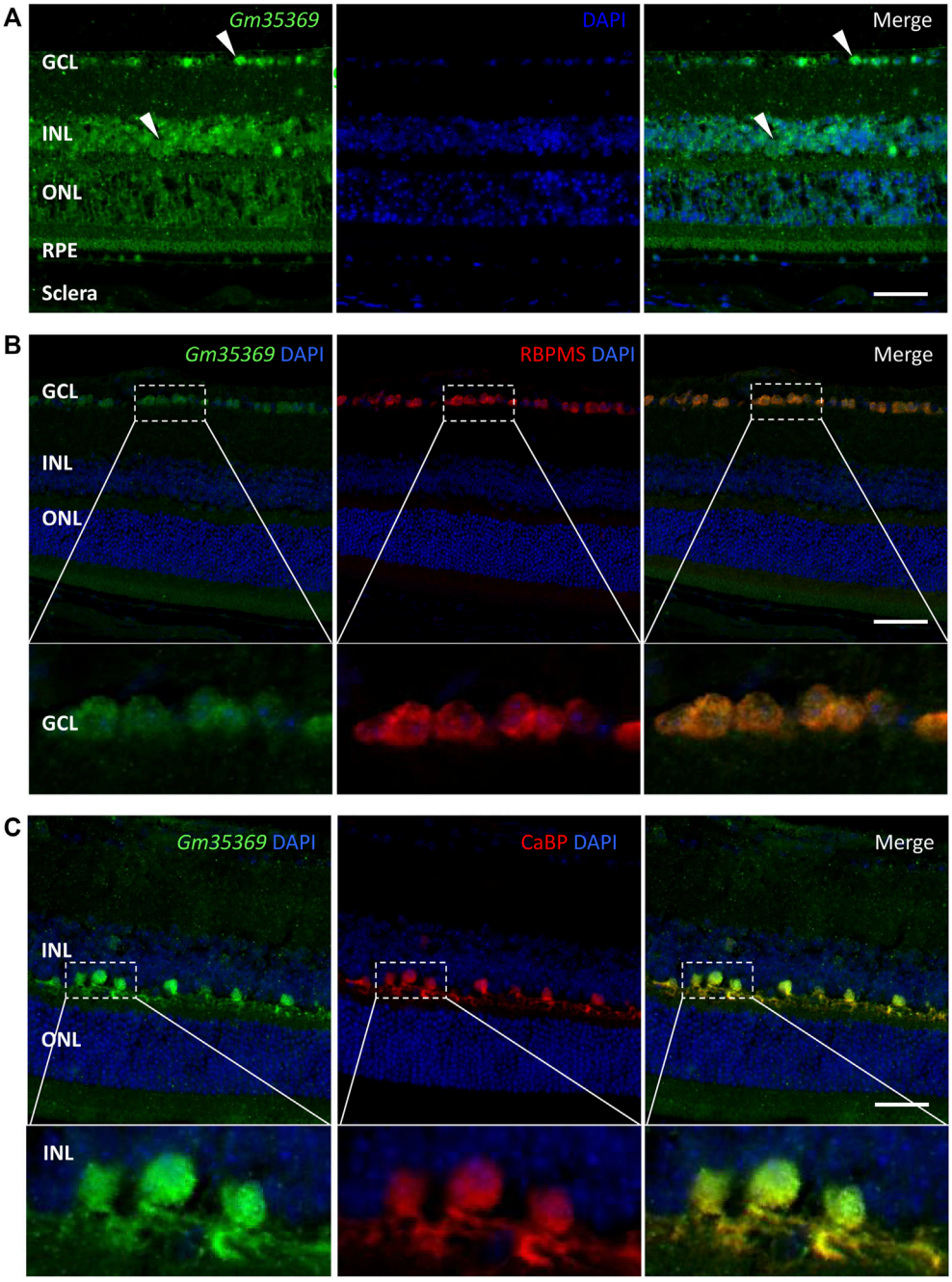


FIGURE 8
RNA FISH showing localization of the lncRNA Gm35369 in the retina. **(A)** RNA fluorescence *in situ* hybridization showing the localization of Gm35369 in the retina, mostly in the GCL and INL. **(B)** Colocalization of the retinal ganglion cell markers RBPMS (*red*) and Gm35369 (*green*) in the GCL. **(C)** Colocalization of the horizontal cell markers CaBP (*red*) and Gm35369 (*green*). Signals were apparent in both the nucleus and cytoplasm. DAPI (*blue*).

treatment status of the eyes during measurements of refractive errors. Moreover, an interocular refractive difference greater than 5 D was considered as successful induction of FDM, which adds to the confidence of the subsequent data. Axial length was measured by SD-OCT (Wu et al., 2015; Wu et al., 2018), MRI (Tkatchenko et al., 2010; Tkatchenko et al., 2013), or even digital caliper and video imaging micrometer in the early days (Barathi et al., 2008). The current study applied OCT to measure the mouse axial length as previous described. Another limitation of the generation of FDM model was that the ocular parameters

were only recorded in the end of the deprivation. Future experiments could assess the parameters at different time points (including before the deprivation) during the generation of FDM, to describe the development of myopia in a detailed way. Additionally, the age of initiation of the FDM and its correlation with the outcome of myopia as well as the gene expression profile would be an interesting direction to investigate, since recent study suggest that the mouse retina develops postnatally with dramatic changes from P10 to P120 by using the Claudin5 (a vascular endothelium-specific gene) eGFP transgenic mice and three-dimensional architecture analysis (Rust et al., 2019).

Among the differentially expressed lncRNAs identified (Tables 1, 2), Oip5os1 (OIP5-AS1, transcript ID: ENSMUST00000147425, FC = 115.55), UBXN7 (transcript ID: chr16: 32388756-32392334, FC = 101.58), and Snhg12 (transcript ID: ENSMUST00000128131, FC = 44.14) exhibited significantly increased expression levels, while Junos (transcript ID: XR_003955073.1, FC = 0.016) and another transcript of Snhg12 (transcript ID: ENSMUST00000126380, FC = 0.017) exhibited downregulated expression in FDM retinal tissue samples. The lncRNA OIP5-AS1 is involved in the pathogenesis of a variety of diseases, including colon cancer (Wang Y. et al., 2021), prostate cancer (Zhang Y. et al., 2021), papillary thyroid cancer (Zhang X. et al., 2021), and osteosarcoma (Li et al., 2021), as well as heart failure (Zhuang et al., 2021). More importantly, previous evidence has suggested that OIP5-AS1 might participate in the mechanisms of primary open angle glaucoma (POAG) and cataracts (Zhou et al., 2020) (Jing et al., 2020). Jing *et al* showed that by regulating apoptosis of lens epithelial cells and aggravating lens opacity, OIP5-AS1 can lead to the formation and development of cataracts (Jing et al., 2020). In POAG, OIP5-AS1, as well as three other lncRNAs, have been found to constitute a hub in a lncRNA-miRNA-mRNA competing endogenous RNA (ceRNA) network (Zhou et al., 2020). UBXN7 was upregulated in human epicardial adipose tissue samples from patients with heart failure (Zheng M. L. et al., 2020). Dysregulation of the lncRNA Snhg12 is involved in a variety of pathogeneses, such as those of LDL-induced endothelial cell injury in atherosclerosis (Mao et al., 2021) and endometrial (Cai et al., 2021), gastric (Zhang T. et al., 2021), and hepatocellular cancer (Zhang Q. et al., 2021). Junos is the opposite strand of the proto-oncogene c-Jun, which is the most extensively studied component of AP-1 and plays important roles in cellular physiology, including proliferation, apoptosis, and tumorigenesis (Meng and Xia, 2011). Therefore, the abnormally expressed lncRNAs in FDM retinas might affect the formation and development of myopia.

Moreover, GO analysis of mRNAs revealed that the differentially expressed mRNAs in the FDM retinas were associated with the cellular process, biological regulation, response to stimulus, metabolic process, developmental process, multicellular organismal process, immune system process, localization, locomotion, behavior, rhythmic process, cellular anatomical entity, binding, and catalytic activity terms. Among the top 20 enriched GO terms, the rhythmic process and the structural molecule activity were shown in the down-regulated but not in the up-regulated pattern. Consistent with recent studies, the downregulation of rhythmic processes or circadian processes might play an important role in the disruption of retinal functions and thus lead to myopia formation (Stone et al., 2013). For instance, the circadian rhythmic control of rod-cone electrical coupling switches the receipt of light signals during day and night (Ribelayga et al., 2008), and diurnal rhythms affect eye growth and refractive error development (Chakraborty et al., 2018). Recent ocular studies focusing on circadian clock genes have also identified the close relationship between myopia formation and circadian dysregulation. Retinal-specific knockout of the clock gene *Bmal1* in mice can induce a myopia shift and elongation of the vitreous chamber (Stone et al., 2019). Furthermore, both FDM and LIM in chicks significantly alter the expression of intrinsic circadian clock genes in the retina/RPE/choroid (Stone et al., 2020). In addition, the synchronization of the local circadian rhythm in the retina with the environmental light cycle requires an orphan opsin, OPN5, which has been found to be involved in emmetropization (Jiang et al., 2021). The term of structural molecule activity belongs to the MF (molecular function) and represents the action of a molecule that contributes to the structural integrity of a complex or its assembly within or outside a cell (Mouse Genome Database, www.informatics.jax.org). Annotated terms under the structural molecule activity include extracellular matrix structural constituent (GO:0005201), structural constituent of cytoskeleton (GO:0005200), structural constituent of eye lens (GO:0005212), and relevant terms. Alterations of extracellular matrix (ECM) remodeling and the cytoskeleton constitution in retina and sclera participate in the development of myopia, and it is generally accepted that the vision guided ocular growth *via* a cascade that firstly from chemical signals initiated in the retina and ultimately change the scleral remodeling (Boote et al., 2020). Early genome-wide association study on myopia in Europeans has revealed that LAMA2, a gene involved in the extracellular matrix, associated with the mechanism behind the development of myopia (Kiefer et al., 2013). Recent transcriptomics analysis of retinas from wavelength-induced myopic guinea pigs also suggest that the differentially expressed genes were primarily enriched in the extracellular matrix, and metabolism, receptor activity, and ion binding processes (Wen et al., 2022). It is worthy to investigate the mechanism of extracellular matrix in retina in myopia as well as the visual development by experimental methods. Our GO results also agree with previous evidence from chicks and guinea pigs. Riddell *et al* showed that fatty acid, sphingolipid, citrate, and mitochondrial metabolism pathways were strongly altered (bidirectionally; up- or downregulated) in retina/RPE/choroid samples from chicks with lens-induced myopia (Riddell et al., 2016). In addition, the levels of retinoic acid and retinaldehyde dehydrogenase-2 from retinoid acid metabolism are changed in the retinas of guinea pigs with lens-induced myopia, with the latter especially altered in the outer plexiform layer (Mao et al., 2012). These signaling pathways exhibiting strong differential expression in

the process of myopia might serve key functions in myopic retinas and merit further study.

The G-protein coupled receptor (GPCR) family member muscarinic acetylcholine receptor plays roles in mediating the development of myopia. Early evidence from FDM Syrian hamsters suggested that the muscarinic receptor M(3) might play important roles in the pathogenesis of myopia (Lin et al., 2012). A recent study involving experimental myopia with a mammalian model demonstrated that inhibition of myopia by muscarinic antagonists involved mainly M(1) and M(4) muscarinic receptor signaling (Arumugam and McBrien, 2012). However, Carr *et al* showed that muscarinic antagonist-mediated blockade of human α2A-adrenergic receptor signaling seemed to be able to inhibit chick FDM, but antagonists of the M(4) subtype did not (Carr et al., 2018). In addition, α2A-adrenoceptor agonists have been shown to be effective in inhibiting chick FDM, suggesting that adrenergic receptors are involved in myopia and visual processes (Carr et al., 2019). The KEGG enrichment pathway analysis showed that the interacting genes were enriched in chemokine signaling, GPCRs, intrinsic component of membrane, sensory perception, and catalytic activity-related pathways (Figure 5), indicating the intrinsic and complicated roles of GPCRs in myopia.

Coexpression analysis of the four qRT-PCR-validated lncRNAs showed that XR_377255.2 was correlated with Vcan (Versican), a critical extracellular matrix regulator of immunity and inflammation (Wight et al., 2020). Accumulating studies have suggested the important role of Vcan in cancer growth and metastasis in cancers such as ovarian, breast, and pancreatic cancers (Salem et al., 2018; Zhang et al., 2019; Gao et al., 2020). Although the samples in this study were retinas and previous reports of extracellular matrix-related genes, such as matrix metallopeptidase-2 and inhibitor of metalloproteinase 2, were from sclerae (Geng et al., 2020), it can be speculated that there might be a close relationship between the alteration of Vcan in the retina and the abnormal regulation of the scleral extracellular matrix. The coexpression analysis also revealed that the lncRNA XR_003955073.1 was correlated with downregulation of the mRNA TREM2, triggering receptor expressed in myeloid cells/microglia-2, which is a transmembrane-spanning sensor receptor critical for Aβ42-peptide clearance. In a study by Bhattacharjee *et al.*, TREM2 deficits in the retina and in oxidatively stressed microglia promoted the pathogenesis of amyloidogenesis in age-related macular degeneration (AMD) (Bhattacharjee et al., 2016). However, the pathogenic roles of TREM2 in myopic retinas remain unknown. In addition, CMIP, C-Maf-inducing protein, was correlated with downregulation of the lncRNA XR_866459.4 in the context of myopia. A previous study has suggested that CMIP is expressed in the nervous system and interacts with NF-κB, which is dysregulated in myopia (Lin et al., 2016; Ollero and Sahali, 2021).

Noncoding RNAs can play essential regulatory roles in many biological processes by acting as competing endogenous RNAs (ceRNAs) to suppress miRNAs by preventing them from interacting with target mRNAs (Grull and Masse, 2019). Several ceRNA pairs have been discovered and studied in the contexts of ocular diseases or pathogenesis. A previous report has shown that the lncRNA-MALAT1/miRNA-204-5p ceRNA mechanism is involved in the regulation of epithelial-mesenchymal transition of lens epithelial cells (Peng et al., 2021). The lncRNA MIR7-3HG can modulate miR-27a-3p/PEG10 and promote retinoblastoma progression (Ding et al., 2020). An integrative analysis of the lncRNA ceRNA network in human trabecular meshwork cells under oxidative stress revealed that 70 lncRNAs and 558 mRNAs were significantly dysregulated in HTMCs under oxidative stress compared to the control conditions (Yao et al., 2020). Moreover, ceRNA crosstalk between the lncRNA TUG1 and miRNA-145 has been found to be involved in the suppression of retinal microvascular endothelial cells under high-glucose conditions (Shi et al., 2021). In profiling retinal lncRNAs during myopia progression, we found dysregulation of lncRNAs and their cellular localization, which lays a foundation for further study of possible ceRNA crosstalk in the myopic eye.

Prior studies have suggested that the neurons of the inner retina play an important role in that process (Chen et al., 2006). RGCs and horizontal cells have been studied under the conditions of emmetropization and myopia progression in animal models. Altered cell-cell coupling by the gap junction protein connexin 36 in horizontal cells in the inner plexiform layer (IPL) has been found to play important roles in emmetropization and FDM in guinea pigs, as the uncoupling agent 18-β-GA induces myopic shifts and FDM decreases total connexin 36 levels and phosphorylation (Zhi et al., 2021). In addition, a recent study showed that stimulation of RGCs expressing neuropsin (OPN5) with violet light prevented experimental myopia in mice (Jiang et al., 2021). We found here that the downregulated lncRNA Gm35369 was located mainly in RGCs and horizontal cells. These results indicate that dysregulation of lncRNAs in specific cellular backgrounds is involved in myopia progression. However, the role of Gm35369 in these cell types remains unclear. Further knockdown or overexpression studies with RGCs or horizontal cell lines would help elucidate the mechanisms.

As with all transcriptomic profiling analyses, there were limitations to the present study. First, among the large numbers of differentially expressed lncRNAs and mRNAs, only 4 lncRNAs were verified by qRT–PCR, while the expression levels of the others remained uncertain. The predicted correlated mRNAs were not verified experimentally, so it is worth testing the levels of these mRNAs in the future. Second, some lncRNA levels verified by qRT-PCR were not consistent with the RNA sequencing data. For instance, the fold change (FC = FDM/ctrl) of the lncRNA XR_866459.4 in qRT-PCR was 0.45 (suggesting decreased expression in the myopic retina), while the FC from the sequencing data was 10.78 (data not shown). Thus, the RNA sequencing results can only be regarded as a reference dataset, and further experiments are needed to confirm the lncRNA targets of

interest. Second, although the lncRNA alterations dynamically changed over the course of visual development, we tested the changes in lncRNA levels at only one time point (after FDM induction for 4 weeks). Third, *in vitro* analysis of specific cultured retinal cell lines is required to identify specific pathways and targets for possible gene-based approaches or drugs to modulate the pathogenesis of myopia. Finally, it is necessary to identify the levels of these lncRNAs in patients with myopia.

Conclusion

Overall, this study analyzed the aberrant expression profiles of lncRNAs and mRNAs in the retinas of FDM mouse models with high-throughput sequencing. The potential roles of the significantly differentially expressed lncRNAs might be related to sensory perception of chemical stimuli, the G-protein coupled receptor signaling pathway, cytokine-cytokine receptor interactions, retinol metabolism, olfactory transduction, metabolism of xenobiotics by cytochrome P450, chemical carcinogenesis, tyrosine metabolism, and the proteasome, which might contribute to retinal myopic pathogenesis. We have preliminarily shown that the lncRNA Gm35369 is mainly located in RGCs and horizontal cells. These findings expand our understanding of lncRNAs in the myopic retina. By revealing a number of candidate target genes and the localization of lncRNAs in specific cell types, this study provides valuable evidence and will support future *in vitro/vivo* studies to investigate the potential mechanisms in myopia.

Data availability statement

The dataset presented in this study can be found in online repository. The name of the repository and accession number can be found below: NCBI SRA BioProject, PRJNA832969 (released upon publication).

Ethics statement

The animal study was reviewed and approved by The Institutional Animal Care and Use Committee of Central South University (Approval No. 2020sydw0077).

Author contributions

YLi and XX conceived the project. DW and XW supervised the project. YLi, YLu, and YY designed the experiments. KD, TH, and YZ provided materials and samples. YLi, YLu, QF, and KD performed the experiments. YLi and YLu analyzed experimental data. YLi and XX prepared the manuscript. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by National Key Research and Development Program of China (No. 2020YFC2008205), National Natural Science Foundation of China (Nos. 81974134, 82171058, and 82000937), Key R&D Plan of Hunan Province of China (No. 2020SK2076), and Hunan Natural Science Foundation (No. 2021JJ41003). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Conflict of interest

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

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.1014031/full#supplementary-Material

References

Arumugam, B., and McBrien, N. A. (2012). Muscarinic antagonist control of myopia: Evidence for M4 and M1 receptor-based pathways in the inhibition of experimentally-induced axial myopia in the tree shrew. *Invest. Ophthalmol. Vis. Sci.* 53 (9), 5827–5837. doi:10.1167/iovs.12-9943

Banerjee, S., Wang, Q., Zhao, F., Tang, G., So, C., Tse, D., et al. (2020). Increased Connexin36 phosphorylation in AII amacrine cell coupling of the mouse myopic retina. *Front. Cell. Neurosci.* 14, 124. doi:10.3389/fncel.2020.00124

Barathi, V. A., Boopathi, V. G., Yap, E. P., and Beuerman, R. W. (2008). Two models of experimental myopia in the mouse. *Vis. Res.* 48 (7), 904–916. doi:10.1016/j.visres.2008.01.004

Bhattacharjee, S., Zhao, Y., Dua, P., Rogaev, E. I., and Lukiw, W. J. (2016). microRNA-34a-Mediated down-regulation of the microglial-enriched triggering receptor and phagocytosis-sensor TREM2 in age-related macular degeneration. *PLoS One* 11 (3), e0150211. doi:10.1371/journal.pone.0150211

Boote, C., Sigal, I. A., Grytz, R., Hua, Y., Nguyen, T. D., and Girard, M. J. A. (2020). Scleral structure and biomechanics. *Prog. Retin. Eye Res.* 74, 100773. doi:10.1016/j.preteyeres.2019.100773

Brockdorff, N., Ashworth, A., Kay, G. F., Cooper, P., Smith, S., McCabe, V. M., et al. (1991). Conservation of position and exclusive expression of mouse Xist from the inactive X chromosome. *Nature* 351 (6324), 329–331. doi:10.1038/351329a0

Cai, P., Li, G., Wu, M., Zhang, B., and Bai, H. (2021). ZIC2 upregulates lncRNA SNHG12 expression to promote endometrial cancer cell proliferation and migration by activating the Notch signaling pathway. *Mol. Med. Rep.* 24 (3), 632. doi:10.3892/mmr.2021.12271

Carninci, P., Kasukawa, T., Katayama, S., Gough, J., Frith, M. C., Maeda, N., et al. (2005). The transcriptional landscape of the mammalian genome. *Science* 309 (5740), 1559–1563. doi:10.1126/science.1112014

Carr, B. J., Mihara, K., Ramachandran, R., Saifeddine, M., Nathanson, N. M., Stell, W. K., et al. (2018). Myopia-inhibiting concentrations of muscarinic receptor antagonists block Activation of Alpha2A-adrenoceptors *in vitro*. *Invest. Ophthalmol. Vis. Sci.* 59 (7), 2778–2791. doi:10.1167/iovs.17-22562

Carr, B. J., Nguyen, C. T., and Stell, W. K. (2019). Alpha2 -adrenoceptor agonists inhibit form-deprivation myopia in the chick. *Clin. Exp. Optom.* 102 (4), 418–425. doi:10.1111/cxo.12871

Carr, B. J., and Stell, W. K. (2016). Nitric oxide (NO) mediates the inhibition of form-deprivation myopia by atropine in chicks. *Sci. Rep.* 6 (1), 9. doi:10.1038/s41598-016-0002-7

Cesana, M., Cacchiarelli, D., Legnini, I., Santini, T., Sthandier, O., Chinappi, M., et al. (2011). A long noncoding RNA controls muscle differentiation by functioning as a competing endogenous RNA. *Cell* 147 (2), 358–369. doi:10.1016/j.cell.2011.09.028

Chakraborty, R., Ostrin, L. A., Nickla, D. L., Iuvone, P. M., Pardue, M. T., and Stone, R. A. (2018). Circadian rhythms, refractive development, and myopia. *Ophthalmic Physiol. Opt.* 38 (3), 217–245. doi:10.1111/opo.12453

Chen, J. C., Brown, B., and Schmid, K. L. (2006). Evaluation of inner retinal function in myopia using oscillatory potentials of the multifocal electroretinogram. *Vis. Res.* 46 (24), 4096–4103. doi:10.1016/j.visres.2006.07.033

Cheng, C. Y., Wang, N., Wong, T. Y., Congdon, N., He, M., Wang, Y. X., et al. (2020). Prevalence and causes of vision loss in east asia in 2015: Magnitude, temporal trends and projections. *Br. J. Ophthalmol.* 104 (5), 616–622. doi:10.1136/bjophthalmol-2018-313308

Ding, F., Jiang, K., Sheng, Y., Li, C., and Zhu, H. (2020). LncRNA MIR7-3HG executes a positive role in retinoblastoma progression via modulating miR-27a-3p/PEG10 axis. *Exp. Eye Res.* 193, 107960. doi:10.1016/j.exer.2020.107960

Dolgin, E. (2015). The myopia boom. *Nature* 519 (7543), 276–278. doi:10.1038/519276a

Dong, Y., Wan, G., Peng, G., Yan, P., Qian, C., and Li, F. (2020). Long non-coding RNA XIST regulates hyperglycemia-associated apoptosis and migration in human retinal pigment epithelial cells. *Biomed. Pharmacother.* 125, 109959. doi:10.1016/j.biopha.2020.109959

Flynn, R. A., and Chang, H. Y. (2014). Long noncoding RNAs in cell-fate programming and reprogramming. *Cell Stem Cell* 14 (6), 752–761. doi:10.1016/j.stem.2014.05.014

Gao, H., Cheng, Y., Chen, Y., Luo, F., Shao, Y., Sun, Z., et al. (2020). The expression of versican and its role in pancreatic neuroendocrine tumors. *Pancreatology* 20 (1), 142–147. doi:10.1016/j.pan.2019.11.009

Geisler, S., and Coller, J. (2013). RNA in unexpected places: Long non-coding RNA functions in diverse cellular contexts. *Nat. Rev. Mol. Cell Biol.* 14 (11), 699–712. doi:10.1038/nrm3679

Geng, C., Li, Y., Guo, F., Wang, J., Yue, Y., Zhou, K., et al. (2020). RNA sequencing analysis of long non-coding RNA expression in ocular posterior poles of Guinea pig myopia models. *Mol. Vis.* 26, 117–134.

Grossniklaus, H. E., and Green, W. R. (1992). Pathologic findings in pathologic myopia. *Retina* 12 (2), 127–133. doi:10.1097/00006982-199212020-00009

Grull, M. P., and Masse, E. (2019). Mimicry, deception and competition: The life of competing endogenous RNAs. *Wiley Interdiscip. Rev. RNA* 10 (3), e1525. doi:10.1002/wrna.1525

Guo, D., Ding, M., Song, X., Sun, Y., Li, G., Li, Z., et al. (2020). Regulatory roles of differentially expressed MicroRNAs in metabolic processes in negative Lens-induced myopia Guinea pigs. *BMC Genomics* 21 (1), 13. doi:10.1186/s12864-020-6447-x

Guoping, L., Xiang, Y., Jianfeng, W., Dadong, G., Jie, H., Wenjun, J., et al. (2017). Alterations of glutamate and gamma-aminobutyric acid expressions in normal and myopic eye development in Guinea pigs. *Invest. Ophthalmol. Vis. Sci.* 58 (2), 1256–1265. doi:10.1167/iovs.16-21130

Guttman, M., Russell, P., Ingolia, N. T., Weissman, J. S., and Lander, E. S. (2013). Ribosome profiling provides evidence that large noncoding RNAs do not encode proteins. *Cell* 154 (1), 240–251. doi:10.1016/j.cell.2013.06.009

Huang, F., Wang, Q., Yan, T., Tang, J., Hou, X., Shu, Z., et al. (2020). The role of the dopamine D2 receptor in form-deprivation myopia in mice: Studies with full and partial D2 receptor agonists and knockouts. *Invest. Ophthalmol. Vis. Sci.* 61 (6), 47. doi:10.1167/iovs.61.6.47

Huang, Y., Kee, C. S., Hocking, P. M., Williams, C., Yip, S. P., Guggenheim, J. A., et al. (2019). A genome-wide association study for susceptibility to visual experience-induced myopia. *Invest. Ophthalmol. Vis. Sci.* 60 (2), 559–569. doi:10.1167/iovs.18-25597

Jiang, X., Pardue, M. T., Mori, K., Ikeda, S. I., Torii, H., D'Souza, S., et al. (2021). Violet light suppresses lens-induced myopia via neuropsin (OPN5) in mice. *Proc. Natl. Acad. Sci. U. S. A.* 118 (22), e2018840118. doi:10.1073/pnas.2018840118

Jing, R., Ma, B., Qi, T., Hu, C., Liao, C., Wen, C., et al. (2020). Long noncoding RNA OIP5-AS1 inhibits cell apoptosis and cataract formation by blocking POLG expression under oxidative stress. *Invest. Ophthalmol. Vis. Sci.* 61 (12), 3. doi:10.1167/iovs.61.12.3

Kiefer, A. K., Tung, J. Y., Do, C. B., Hinds, D. A., Mountain, J. L., Francke, U., et al. (2013). Genome-wide analysis points to roles for extracellular matrix remodeling, the visual cycle, and neuronal development in myopia. *PLoS Genet.* 9 (2), e1003299. doi:10.1371/journal.pgen.1003299

Li, X. J. (2018). Long non-coding RNA nuclear paraspeckle assembly transcript 1 inhibits the apoptosis of retina Muller cells after diabetic retinopathy through regulating miR-497/brain-derived neurotrophic factor axis. *Diab. Vasc. Dis. Res.* 15 (3), 204–213. doi:10.1177/1479164117749382

Li, Y., Lin, S., Xie, X., Zhu, H., Fan, T., and Wang, S. (2021). Highly enriched exosomal lncRNA OIP5-AS1 regulates osteosarcoma tumor angiogenesis and autophagy through miR-153 and ATG5. *Am. J. Transl. Res.* 13 (5), 4211–4223.

Lin, H. J., Wan, L., Chen, W. C., Lin, J. M., Lin, C. J., and Tsai, F. J. (2012). Muscarinic acetylcholine receptor 3 is dominant in myopia progression. *Invest. Ophthalmol. Vis. Sci.* 53 (10), 6519–6525. doi:10.1167/iovs.11-9031

Lin, H. J., Wei, C. C., Chang, C. Y., Chen, T. H., Hsu, Y. A., Hsieh, Y. C., et al. (2016). Role of chronic inflammation in myopia progression: Clinical evidence and experimental validation. *EBioMedicine* 10, 269–281. doi:10.1016/j.ebiom.2016.07.021

Liu, H. H., Kenning, M. S., Jobling, A. I., McBrien, N. A., and Gentle, A. (2017). Reduced scleral TIMP-2 expression is associated with myopia development: TIMP-2 supplementation stabilizes scleral biomarkers of myopia and limits myopia development. *Invest. Ophthalmol. Vis. Sci.* 58 (4), 1971–1981. doi:10.1167/iovs.16-21181

Lou, L., Yao, C., Jin, Y., Perez, V., and Ye, J. (2016). Global patterns in health burden of uncorrected refractive error. *Invest. Ophthalmol. Vis. Sci.* 57 (14), 6271–6277. doi:10.1167/iovs.16-20242

Mao, J. F., Liu, S. Z., and Dou, X. Q. (2012). Retinoic acid metabolic change in retina and choroid of the Guinea pig with lens-induced myopia. *Int. J. Ophthalmol.* 5 (6), 670–674. doi:10.3980/j.issn.2222-3959.2012.06.04

Mao, P., Liu, X., Wen, Y., Tang, L., and Tang, Y. (2021). LncRNA SNHG12 regulates ox-LDL-induced endothelial cell injury by the miR-218-5p/IGF2 axis in atherosclerosis. *Cell Cycle* 20 (16), 1561–1577. doi:10.1080/15384101.2021.1953755

Mei, F., Wang, J., Chen, Z., and Yuan, Z. (2017). Potentially important MicroRNAs in form-deprivation myopia revealed by bioinformatics analysis of MicroRNA profiling. *Ophthalmic Res.* 57 (3), 186–193. doi:10.1159/000452421

Meng, Q., and Xia, Y. (2011). c-Jun, at the crossroad of the signaling network. *Protein Cell* 2 (11), 889–898. doi:10.1007/s13238-011-1113-3

Ni, Y., Liu, F., Hu, X., Qin, Y., and Zhang, Z. (2021). Coding and non-coding RNA interactions reveal immune-related pathways in peripheral blood mononuclear cells derived from patients with proliferative vitreoretinopathy. *BMC Med. Genomics* 14 (1), 30. doi:10.1186/s12920-021-00875-5

Nickla, D. L., and Totonelly, K. (2011). Dopamine antagonists and brief vision distinguish lens-induced- and form-deprivation-induced myopia. *Exp. Eye Res.* 93 (5), 782–785. doi:10.1016/j.exer.2011.08.001

Ollero, M., and Sahali, D. (2021). The enigmatic emerging role of the C-maf inducing protein in cancer. *Diagn. (Basel)* 11 (4), 666. doi:10.3390/diagnostics11040666

Peng, C., Wang, Y., Ji, L., Kuang, L., Yu, Z., Li, H., et al. (2021). LncRNA-MALAT1/miRNA-204-5p/Smad4 Axis regulates epithelial-mesenchymal transition, proliferation and migration of lens epithelial cells. *Curr. Eye Res.* 46 (8), 1137–1147. doi:10.1080/02713683.2020.1857778

Ribelayga, C., Cao, Y., and Mangel, S. C. (2008). The circadian clock in the retina controls rod-cone coupling. *Neuron* 59 (5), 790–801. doi:10.1016/j.neuron.2008.07.017

Riddell, N., Giummarra, L., Hall, N. E., and Crewther, S. G. (2016). Bidirectional expression of metabolic, structural, and immune pathways in early myopia and hyperopia. *Front. Neurosci.* 10, 390. doi:10.3389/fnins.2016.00390

Rust, R., Gronnert, L., Dogancay, B., and Schwab, M. E. (2019). A revised view on growth and remodeling in the retinal vasculature. *Sci. Rep.* 9 (1), 3263. doi:10.1038/s41598-019-40135-2

Salem, M., O'Brien, J. A., Bernaudo, S., Shawer, H., Ye, G., Brkic, J., et al. (2018). miR-590-3p promotes ovarian cancer growth and metastasis via a novel FOXA2-versican pathway. *Cancer Res.* 78 (15), 4175–4190. doi:10.1158/0008-5472.CAN-17-3014

Schaeffel, F., Burkhardt, E., Howland, H. C., and Williams, R. W. (2004). Measurement of refractive state and deprivation myopia in two strains of mice. *Optom. Vis. Sci.* 81 (2), 99–110. doi:10.1097/00006324-200402000-00008

Shi, Q., Tang, J., Wang, M., Xu, L., and Shi, L. (2021). Knockdown of long non-coding RNA TUG1 suppresses migration and tube formation in high glucose-stimulated human retinal microvascular endothelial cells by sponging miRNA-145. *Mol. Biotechnol.* 64, 171–177. doi:10.1007/s12033-021-00398-5

Srinivasalu, N., McFadden, S. A., Medcalf, C., Fuchs, L., Chung, J., Philip, G., et al. (2018). Gene expression and pathways underlying form deprivation myopia in the Guinea pig sclera. *Invest. Ophthalmol. Vis. Sci.* 59 (3), 1425–1434. doi:10.1167/iovs.16-21278

Stone, R. A., McGlinn, A. M., Chakraborty, R., Lee, D. C., Yang, V., Elmasri, A., et al. (2019). Altered ocular parameters from circadian clock gene disruptions. *PLoS One* 14 (6), e0217111. doi:10.1371/journal.pone.0217111

Stone, R. A., Pardue, M. T., Iuvone, P. M., and Khurana, T. S. (2013). Pharmacology of myopia and potential role for intrinsic retinal circadian rhythms. *Exp. Eye Res.* 114, 35–47. doi:10.1016/j.exer.2013.01.001

Stone, R. A., Wei, W., Sarfare, S., McGeehan, B., Engelhart, K. C., Khurana, T. S., et al. (2020). Visual image quality impacts circadian rhythm-related gene expression in retina and in choroid: A potential mechanism for ametropias. *Invest. Ophthalmol. Vis. Sci.* 61 (5), 13. doi:10.1167/iovs.61.5.13

Tanaka, Y., Kurihara, T., Hagiwara, Y., Ikeda, S. I., Mori, K., Jiang, X., et al. (2019). Ocular-component-specific miRNA expression in a murine model of lens-induced myopia. *Int. J. Mol. Sci.* 20 (15), E3629. doi:10.3390/ijms20153629

Tkatchenko, A. V., Luo, X., Tkatchenko, T. V., Vaz, C., Tanavde, V. M., Maurer-Stroh, S., et al. (2016). Large-scale microRNA expression profiling identifies putative retinal miRNA-mRNA signaling pathways underlying form-deprivation myopia in mice. *PLoS One* 11 (9), e0162541. doi:10.1371/journal.pone.0162541

Tkatchenko, T. V., Shen, Y., Braun, R. D., Bawa, G., Kumar, P., Avrutsky, I., et al. (2013). Photopic visual input is necessary for emmetropization in mice. *Exp. Eye Res.* 115, 87–95. doi:10.1016/j.exer.2013.06.025

Tkatchenko, T. V., Shen, Y., and Tkatchenko, A. V. (2010). Analysis of postnatal eye development in the mouse with high-resolution small animal magnetic resonance imaging. *Invest. Ophthalmol. Vis. Sci.* 51 (1), 21–27. doi:10.1167/iovs.08-2767

Tu, Y., Xie, L., Chen, L., Yuan, Y., Qin, B., Wang, K., et al. (2020). Long non-coding RNA MEG3 promotes cataractogenesis by upregulating TP53INP1 expression in age-related cataract. *Exp. Eye Res.* 199, 108185. doi:10.1016/j.exer.2020.108185

Vocale, L. G., Crewther, S., Riddell, N., Hall, N. E., Murphy, M., and Crewther, D. (2021). RNA-seq and GSEA identifies suppression of ligand-gated chloride efflux channels as the major gene pathway contributing to form deprivation myopia. *Sci. Rep.* 11 (1), 5280. doi:10.1038/s41598-021-84338-y

Wang, H., Zhang, Z., Zhang, Y., Liu, S., and Li, L. (2021a). Long non-coding RNA TP53TG1 upregulates SHCBP1 to promote retinoblastoma progression by sponging miR-33b. *Cell Transpl.* 30, 9636897211025223. doi:10.1177/09636897211025223

Wang, S., Liu, S., Mao, J., and Wen, D. (2014). Effect of retinoic acid on the tight junctions of the retinal pigment epithelium-choroid complex of Guinea pigs with lens-induced myopia *in vivo*. *Int. J. Mol. Med.* 33 (4), 825–832. doi:10.3892/ijmm.2014.1651

Wang, Y., Lin, C., and Liu, Y. (2021b). Molecular mechanism of miR-34b-5p and RNA binding protein HuR binding to lncRNA OIP5-AS1 in colon cancer cells. *Cancer Gene Ther.* 29, 612–624. doi:10.1038/s41417-021-00342-4

Wen, Y., Dai, B., Zhang, X., Zhu, H., Xie, C., Xia, J., et al. (2022). Retinal transcriptomics analysis reveals the underlying mechanism of disturbed emmetropization induced by wavelength defocus. *Curr. Eye Res.* 47 (6), 908–917. doi:10.1080/02713683.2022.2048395

Wight, T. N., Kang, I., Evanko, S. P., Harten, I. A., Chang, M. Y., Pearce, O. M. T., et al. (2020). Versican-A critical extracellular matrix regulator of immunity and inflammation. *Front. Immunol.* 11, 512. doi:10.3389/fimmu.2020.00512

Wu, H., Chen, W., Zhao, F., Zhou, Q., Reinach, P. S., Deng, L., et al. (2018). Scleral hypoxia is a target for myopia control. *Proc. Natl. Acad. Sci. U. S. A.* 115 (30), E7091–E7100. doi:10.1073/pnas.1721443115

Wu, P. C., Tsai, C. L., Gordon, G. M., Jeong, S., Itakura, T., Patel, N., et al. (2015). Chondrogenesis in scleral stem/progenitor cells and its association with form-deprived myopia in mice. *Mol. Vis.* 21, 138–147.

Wu, S. Y., Nemesure, B., and Leske, M. C. (2000). Glaucoma and myopia. *Ophthalmology* 107 (6), 1026–1027. doi:10.1016/s0161-6420(00)00051-8

Xiao, H., Fan, Z. Y., Tian, X. D., and Xu, Y. C. (2014). Comparison of form-deprived myopia and lens-induced myopia in Guinea pigs. *Int. J. Ophthalmol.* 7 (2), 245–250. doi:10.3980/j.issn.2222-3959.2014.02.10

Yao, K., Yu, Y., Li, F., Jin, P., Deng, C., and Zhang, H. (2020). Integrative analysis of an lncRNAassociated competing endogenous RNA network in human trabecular meshwork cells under oxidative stress. *Mol. Med. Rep.* 21 (3), 1606–1614. doi:10.3892/mmr.2020.10955

Yuan, Y., Li, M., To, C. H., Lam, T. C., Wang, P., Yu, Y., et al. (2018). The role of the RhoA/ROCK signaling pathway in mechanical strain-induced scleral myofibroblast differentiation. *Invest. Ophthalmol. Vis. Sci.* 59 (8), 3619–3629. doi:10.1167/iovs.17-23580

Zeng, L., Li, X., Liu, J., Liu, H., Xu, H., and Yang, Z. (2021). RNA-seq analysis reveals an essential role of the tyrosine metabolic pathway and inflammation in myopia-induced retinal degeneration in Guinea pigs. *Int. J. Mol. Sci.* 22 (22), 12598. doi:10.3390/ijms222212598

Zhang, Q., Cheng, M., Fan, Z., Jin, Q., Cao, P., and Zhou, G. (2021a). Identification of cancer cell stemness-associated long noncoding RNAs for predicting prognosis of patients with hepatocellular carcinoma. *DNA Cell Biol.* 40 (8), 1087–1100. doi:10.1089/dna.2021.0282

Zhang, T., Beeharry, M. K., Zheng, Y., Wang, Z., Li, J., Zhu, Z., et al. (2021b). Long noncoding RNA SNHG12 promotes gastric cancer proliferation by binding to HuR and stabilizing YWHAZ expression through the AKT/GSK-3β pathway. *Front. Oncol.* 11, 645832. doi:10.3389/fonc.2021.645832

Zhang, X., Li, D., Jia, C., Cai, H., Lv, Z., and Wu, B. (2021c). METTL14 promotes tumorigenesis by regulating lncRNA OIP5-AS1/miR-98/ADAMTS8 signaling in papillary thyroid cancer. *Cell Death Dis.* 12 (6), 617. doi:10.1038/s41419-021-03891-6

Zhang, Y., Guo, S., Wang, S., Li, X., Hou, D., Li, H., et al. (2021d). LncRNA OIP5-AS1 inhibits ferroptosis in prostate cancer with long-term cadmium exposure through miR-128-3p/SLC7A11 signaling. *Ecotoxicol. Environ. Saf.* 220, 112376. doi:10.1016/j.ecoenv.2021.112376

Zhang, Y., Zou, X., Qian, W., Weng, X., Zhang, L., Zhang, L., et al. (2019). Enhanced PAPSS2/VCAN sulfation axis is essential for Snail-mediated breast cancer cell migration and metastasis. *Cell Death Differ.* 26 (3), 565–579. doi:10.1038/s41418-018-0147-y

Zhao, F., Zhang, D., Zhou, Q., Zhao, F., He, M., Yang, Z., et al. (2020). Scleral HIF-1α is a prominent regulatory candidate for genetic and environmental interactions in human myopia pathogenesis. *EBioMedicine* 57, 102878. doi:10.1016/j.ebiom.2020.102878

Zhao, F., Zhou, H., Chen, W., Zhao, C., Zheng, Y., Tao, Y., et al. (2021). Declines in PDE4B activity promote myopia progression through downregulation of scleral collagen expression. *Exp. Eye Res.* 212, 108758. doi:10.1016/j.exer.2021.108758

Zheng, M. L., Du, X. P., Zhao, L., and Yang, X. C. (2020b). Expression profile of circular RNAs in epicardial adipose tissue in heart failure. *Chin. Med. J.* 133 (21), 2565–2572. doi:10.1097/CM9.0000000000001056

Zheng, M., Zheng, Y., Gao, M., Ma, H., Zhang, X., Li, Y., et al. (2020a). Expression and clinical value of lncRNA MALAT1 and lncRNA ANRIL in glaucoma patients. *Exp. Ther. Med.* 19 (2), 1329–1335. doi:10.3892/etm.2019.8345

Zhi, Z., Xiang, J., Fu, Q., Pei, X., Zhou, D., Cao, Y., et al. (2021). The role of retinal connexins Cx36 and horizontal cell coupling in emmetropization in Guinea pigs. *Invest. Ophthalmol. Vis. Sci.* 62 (9), 27. doi:10.1167/iovs.62.9.27

Zhou, M., Lu, B., Tan, W., and Fu, M. (2020). Identification of lncRNA-miRNA-mRNA regulatory network associated with primary open angle glaucoma. *BMC Ophthalmol.* 20 (1), 104. doi:10.1186/s12886-020-01365-5

Zhu, Y., Li, W., Zhu, D., and Zhou, J. (2020). microRNA profiling in the aqueous humor of highly myopic eyes using next generation sequencing. *Exp. Eye Res.* 195, 108034. doi:10.1016/j.exer.2020.108034

Zhuang, A., Calkin, A. C., Lau, S., Kiriazis, H., Donner, D. G., Liu, Y., et al. (2021). Loss of the long non-coding RNA OIP5-AS1 exacerbates heart failure in a sex-specific manner. *iScience* 24 (6), 102537. doi:10.1016/j.isci.2021.102537

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 19 October 2022
DOI 10.3389/fgene.2022.1030510

OPEN ACCESS

EDITED BY
Duo Liu,
Harbin Medical University Cancer Hospital, China

REVIEWED BY
Yu'E Liu,
Tongji University, China
Lingxiao Pang,
Fourth Affiliated Hospital of China Medical University, China

*CORRESPONDENCE
Laijing Du,
369593128@qq.com
Hengliang Zhang,
Hengliang.zhang@med.uni-muenchen.de

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

RECEIVED 29 August 2022
ACCEPTED 30 September 2022
PUBLISHED 19 October 2022

CITATION
Zhang H, Merkus D, Zhang P, Zhang H, Wang Y, Du L and Kottu L (2022), Predicting protective gene biomarker of acute coronary syndrome by the circRNA-associated competitive endogenous RNA regulatory network. *Front. Genet.* 13:1030510. doi: 10.3389/fgene.2022.1030510

Predicting protective gene biomarker of acute coronary syndrome by the circRNA-associated competitive endogenous RNA regulatory network

Hengliang Zhang[1,2]*, Daphne Merkus[2,3], Pei Zhang[1], Huifeng Zhang[1], Yanyu Wang[1], Laijing Du[1]* and Lakshme Kottu[3]

[1]The First Affiliated Hospital, and College of Clinical Medicine of Henan University of Science and Technology, Luoyang, China, [2]Walter-Brendel-Centre of Experimental Medicine, University Hospital, Ludwig-Maximilians-University München, Munich, Germany, [3]Department of Experimental Cardiology, Erasmus University Medical Center, Rotterdam, Netherlands

Background: The mortality and disability rates of acute coronary syndrome (ACS) are quite high. Circular RNA (circRNA) is a competitive endogenous RNA (ceRNA) that plays an important role in the pathophysiology of ACS. Our goal is to screen circRNA-associated ceRNA networks for biomarker genes that are conducive to the diagnosis or exclusion of ACS, and better understand the pathology of the disease through the analysis of immune cells.

Materials and methods: RNA expression profiles for circRNAs (GSE197137), miRNAs (GSE31568), and mRNAs (GSE95368) were obtained from the GEO database, and differentially expressed RNAs (DEcircRNAs, DEmiRNAs, and DEmRNAs) were identified. The circRNA-miRNA and miRNA-mRNA regulatory links were retrieved from the CircInteractome database and TargetScan databases, respectively. As a final step, a regulatory network has been designed for ceRNA. On the basis of the ceRNA network, hub mRNAs were verified by quantitative RT-PCR. Hub genes were validated using a third independent mRNA database GSE60993, and ROC curves were used to evaluate their diagnostic values. The correlation between hub genes and immune cells associated with ACS was then analyzed using single sample gene set enrichment analysis (ssGSEA).

Results: A total of 17 DEcircRNAs, 229 DEmiRNAs, and 27 DEmRNAs were found, as well as 52 circRNA-miRNA pairings and 10 miRNA-mRNA pairings predicted. The ceRNA regulatory network (circRNA-miRNA-mRNA) was constructed, which included 2 circRNA (hsa_circ_0082319 and hsa_circ_0005654), 4 miRNA (hsa-miR-583, hsa-miR-661, hsa-miR-671-5p, hsa-miR-578), and 5 mRNA (*XPNPEP1, UCHL1, DBNL, GPC6,* and *RAD51*). The qRT-PCR analysis result showed that the *XPNPEP1, UCHL1, GPC6* and *RAD51* genes had a significantly decreased expression in ACS patients. Based on ROC curve analysis, we found that *XPNPEP1* has important significance in preventing ACS occurrence and excluding ACS diagnosis. ACS immune infiltration analysis

revealed significant correlations between the other 3 hub genes (UCHL1, GPC6, RAD51) and the immune cells (Eosinophils, T folliculars, Type 2 T helper cells, and Imumature dendritic cells).

Conclusion: Our study constructed a circRNA-related ceRNA network in ACS. The *XPNPEP1* gene could be a protective gene biomarker for ACS. The *UCHL1, GPC6 and RAD51* genes were significantly correlated with immune cells in ACS.

KEYWORDS

acute coronary syndrome, circular RNA, competitive endogenous RNA, protective gene biomarkers, immune cells, XPNPEP1 gene

Introduction

One of the most serious, urgent, and lethal disorders in the clinic is acute coronary syndrome (ACS), which includes unstable angina pectoris (UA) and acute myocardial infarction (AMI). In high-income nations, the incidence of ACS is 200–250 cases per 100,000 person-years, and ACS is responsible for one-third of all mortality (Sanchis-Gomar et al., 2016; Ibanez et al., 2018; Collet et al., 2021). The exact pathogenic mechanism of ACS has not yet been extensively investigated. An early diagnosis, early intervention, and early prevention are essential tools in reducing ACS's harm and improving its prognosis. Researchers are investigating ACS' pathogenic process at the protein, molecular, and gene levels simultaneously. (Sposito, 2022; van den Broek and Ten Berg, 2022). Finding the genes that promote or protect ACS is critical since it can fundamentally prevent disease and give a theoretical foundation for the development of targeted medicine and precision therapy.

As a competitive endogenous RNA (ceRNA), circular RNA (circRNA) plays a significant role in the pathogenesis of ACS. CircRNA affects gene transcription and regulation by interacting with miRNA, mRNA, or protein *via* the ceRNA mechanism (Hansen et al., 2013; Zhao et al., 2018). In recent years, more evidence has emerged demonstrating that circRNA expression dysregulation plays a critical role in the ceRNA regulatory network and is a fundamental etiology of atherosclerotic diseases such as acute cerebral infarction and acute coronary syndrome (Zhao et al., 2020). In order to further clarify how circRNA affects the occurrence and development of ACS, we set up a circRNA-associated ceRNA network to investigate the pathophysiology of ACS at the molecular level. Through the analysis of the hub genes in the ceRNA network, we will try to screen the biomarker genes that are conducive to the diagnosis or exclusion of ACS, and explore its pathological mechanism through immune cell analysis.

Materials and methods

Data collection

The Gene Expression Omnibus database (GEO, https://www.ncbi.nlm.nih.gov/geo/) plays an important role in many fields, including comparative genomic analysis, proteomics, non-coding RNA, single nucleotide polymorphism genome and gene methylation status analysis (Sayers et al., 2022). A search was conducted from the GEO dataset for microarrays that met the following requirements: using "acute coronary syndrome" OR "unstable angina" OR "acute myocardial infarction" AND "circRNA".

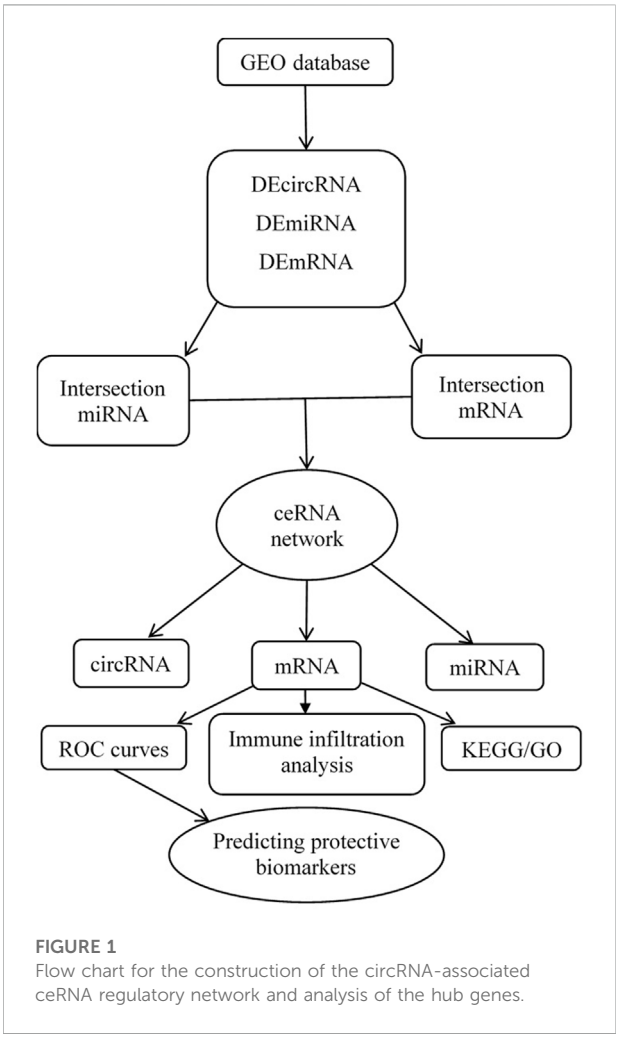


FIGURE 1
Flow chart for the construction of the circRNA-associated ceRNA regulatory network and analysis of the hub genes.

Three datasets were downloaded from the GEO database after a selection process. One circRNA expression profiling dataset [GSE197137 (GPL21825 platform)], consisted of 3 ACS patients and 6 subjects without ACS as controls. One miRNA expression profiling dataset [GSE31568 (GPL9040 platform)] included 20 ACS patients and 70 healthy controls, and one mRNA expression profiling dataset [GSE95368 (GPL23119 platform)] included 12 ACS patients and 6 healthy controls. Our research process is shown in Figure 1.

Data processing

The genes used for constructing the ceRNA network must be differentially expressed (DE) in ACS patient samples compared to non-ACS subject samples. Perl software version 5.30.0.1 was used to convert the probe matrix into an RNA matrix. DEcircRNA, DEmiRNA, and DEmRNA were obtained using the "limma" package in R version 4.0.3. The following screening criteria were considered statistically significant: $|\log_2$ fold chang $| > 2$ and adjusted p-value < 0.05. The pheatmap package, ggpubr package and reshape2 package in R are used for the visual analysis of differentially expressed RNA.

Construction of the competitive endogenous RNA network

According to the theory that circRNAs act as miRNA sponges in mammalian cells (Salmena et al., 2011), we constructed a circRNA-miRNA-mRNA regulatory network. The CircInteractome database (https://circinteractome.nia.nih.gov/index.html) was used to predict miRNA binding sites (MREs) (Dudekula et al., 2016). The circRNAs' chromosomal position, as well as the chromosomal location and length of the RNA necessary for the investigation, were all provided by CircBase. TargetScan databases were used to anticipate interactions between intersection miRNAs and target mRNAs (Xiong et al., 2018). Finally, the data for the regulatory network (circRNA-miRNA-mRNA) was processed through Perl software, and then a visual ceRNA regulatory network was established by using Cytoscape 3.8.0 software.

KEGG and GO pathway enrichment analysis

GO analysis (http://geneontology.org/) covers three areas, cellular components (CC), molecular functions (MF) and biological processes (BP). Each category explains the biological function of genes at different levels (Zhang et al., 2019). The Kyoto Encyclopedia of Genes and Genomes (KEGG) database (https://www.kegg.jp/) is a popular public resource for learning about the degree of enrichment of differential genes in pathway terms (Kanehisa and Goto, 2000). We performed GO function annotation and KEGG pathway analysis to further investigate the pathway and mechanism of DEmRNA in the ceRNA network affecting ACS by using the "cluster profiler" package in the R software.

Quantitative RT-PCR verification

We obtained peripheral blood from the patients with ACS who visited the emergency department within 12 h of the onset of chest pain. Total RNA was isolated from blood using Trizol and cDNA was synthesized using reverse transcription kits (Takara, Beijing, China) according to the manufacturer's instructions. Quantitative reverse transcription PCR was performed using SYBR Green Mix on an ABI7900HI (Thermo Fisher Scientific). The program was set to be a two-step method, 95°C for 5 s, 60°C for 30 s, and 40 cycles. The gene expression results were analyzed using the $2^{-\Delta\Delta CT}$ method, and GAPDH was used as an endogenous control for mRNA expression. The primer sequence information of the qRT-PCR experiment is shown in Supplementary Table S1.

Predicting the protective value of characteristic biomarkers in acute coronary syndrome

In order to test the protective value of identified biomarkers, receiver operating characteristic (ROC) curves were generated using the mRNA expression data from the GSE95368 and another independent mRNA database GSE60993. The predicting protective values of the identified hub genes were evaluated using the area under the ROC curve (AUC), which was between 0.5 and 1. The closer the AUC is to 1, the better the predictability of the protective effect. We obtained patient follow-up data from the GSE95368 dataset and used logistic regression analysis to investigate the relationship between the hub gene and major adverse cardiovascular events (MACE).

Expression of XPNPEP1 in cardiomyocytes by immunofluorescence staining

We obtained human induced pluripotent stem cells (iPSC) from skin fibroblasts of healthy human donor and ACS patients. We referred to Chen's scheme to differentiate the hiPSC to cardiomyocytes (Chen et al., 2011). Cardiomyocytes derived from human induced pluripotent stem cells (hiPSC-CM) were fixed with 4% paraformaldehyde for 1 h. After permeabilization (1% Triton X-100, 1 h), staining was performed with the primary

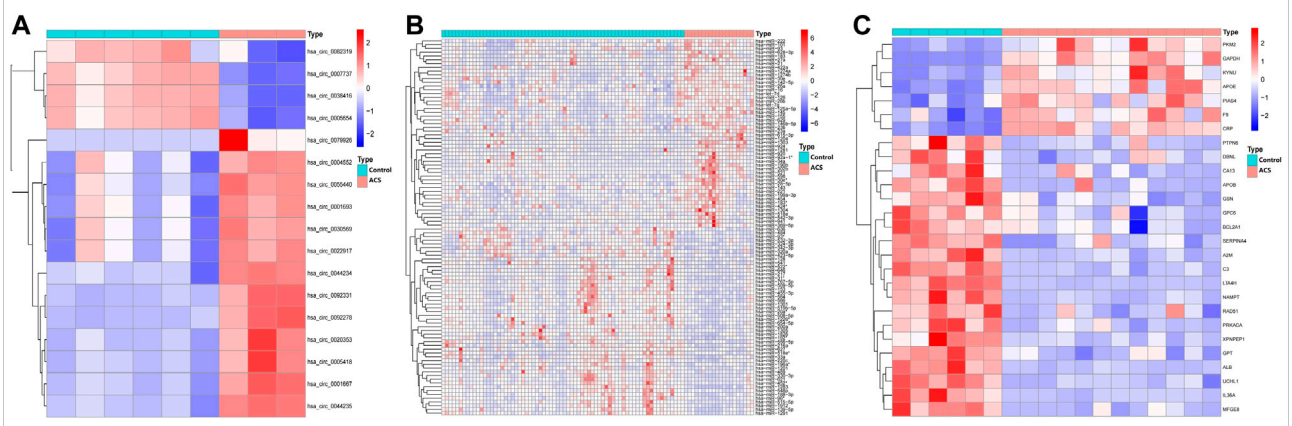


FIGURE 2
Heatmap of differentially expressed circRNA, miRNA, and mRNA. Red represents upregulated expression, and blue means downregulated expression. **(A)** 17 differentially expressed circRNAs in the GSE197137 dataset. **(B)** Top 50 of 229 differentially expressed miRNAs in the GSE31568 dataset. **(C)** 27 differentially expressed mRNAs in the GSE95368 dataset (|$\log_2$ fold change|>2, adjusted p-value<0.05).

antibody (anti-rabbit XPNPEP1 antibody, 1:50, Abcam), appropriate secondary antibody (goat anti-rabbit IgG Alexa Fluor 594, 1:100, Invitrogen) and DAPI (Invitrogen). The stained cytoskeleton structure was observed under a laser confocal microscope (LSM800; Carl Zeiss Meditec, Jena, Germany).

Correlation analysis between acute coronary syndrome hub genes and immune infiltrating cells

The single sample gene set enrichment analysis (ssGSEA) method was then used to evaluate the abundance of 23 immune cells in ACS to further investigate the correlation. Then we analyzed the correlation between hub genes and ACS related immune cells. The heat map and the vioplpt were constructed by the “ggplot2” package and the “ggpubr” package to visualize the features.

Results

Extraction of DEcircRNA, DEmiRNA, and DEmRNA

Human-derived RNA datasets (GSE197137, GSE31568, and GSE95368) containing patients with ACS and subjects without ACS were selected. The differential expression of the data set was analysed with the “limma” package. In total, 17 DEcircRNAs were obtained, of which 4 were downregulated and 13 were

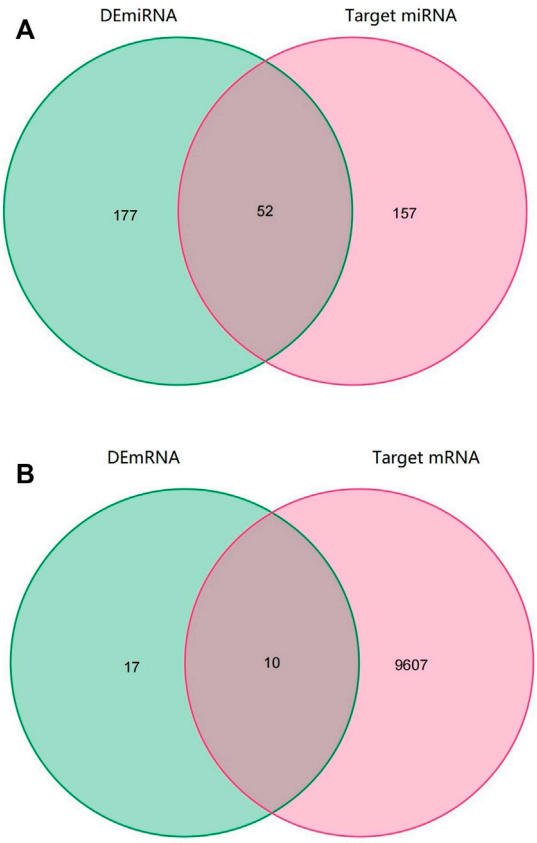


FIGURE 3
Venn diagram of RNAs involved in the ceRNA network. **(A)** The predicted target miRNA and DEmiRNA intersected **(B)** The target mRNA and DEmRNA intersected.

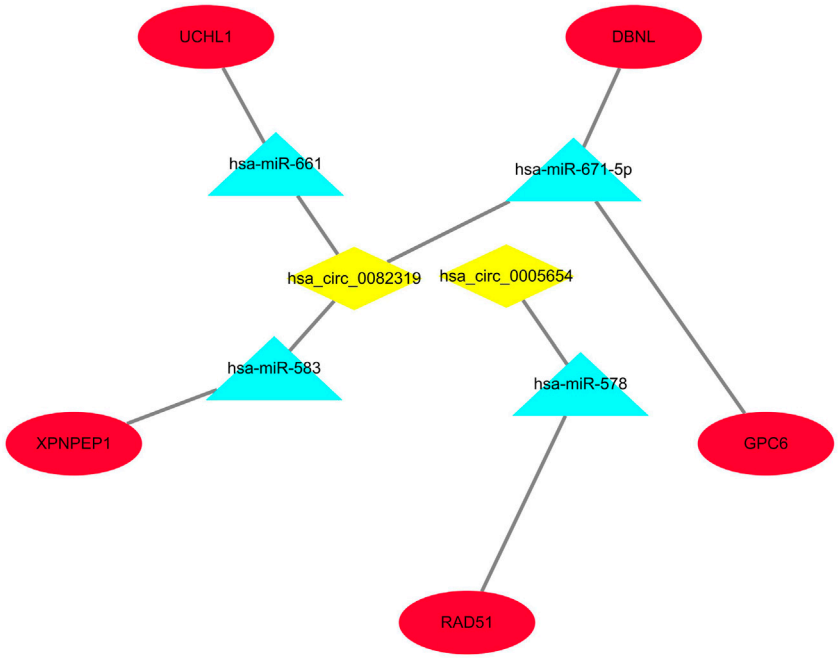


FIGURE 4
The visualized ceRNA regulatory network was constructed by Cytoscape software. The diamond represents 2 downregulated circRNAs, the triangle represents 4 upregulated miRNAs and the ellipse represents 5 downregulated mRNAs.

upregulated. Of the 229 DEmiRNAs, 103 were downregulated and 126 were upregulated. There were 27 DEmRNAs, 20 were downregulated and 7 were upregulated. As shown in Figure 2, we chose the top 50 downregulated and upregulated DERNAs for heat map analysis.

Identification of target miRNA and mRNA of circular RNA

Through the CircInteractome database, we predicted a total of 395 target miRNAs that bind to 15 DEcircRNAs. Following that, using the Venn diagram method, the predicted target miRNA and DEmiRNA were intersected, and 52 intersection miRNAs were obtained, as shown in Figure 3A. Next, genes identified in the TargetScan database were selected as potential target mRNAs. Then, the above-mentioned intersection miRNAs were predicted to obtain target 9,607 mRNAs. By intersecting the target mRNA and DEmRNA, 10 intersection mRNAs were obtained, as shown in Figure 3B.

Construction of the regulatory network

Only genes that meet the following criteria will be selected for inclusion in the ceRNA network: 1) All genes must be differentially expressed; 2) circRNAs and mRNAs have a binding relationship with miRNAs at the same time; 3) RNAs (circRNA, mRNA) and miRNAs that meet the above binding relationship must be negatively regulated. Then, through Cytoscape software, a ceRNA regulatory network (circRNA-miRNA-mRNA) was constructed (Figure 4), including 2 circRNAs (hsa_circ_0082319, hsa_circ_0005654), 4 miRNAs (hsa-miR-583, hsa-miR-661, hsa-miR-671-5p, hsa-miR-578), and 5 mRNAs (*XPNPEP1, UCHL1, DBNL, GPC6,* and *RAD51*). Meanwhile, the "ggpubr" package and "reshape2" package in R were used to visualize the circRNA, miRNA, and mRNA in the ceRNA regulatory network (Figures 5A–C). The structure pattern diagrams and basic characteristics of the 2 circRNAs, are shown in Supplementary Table S1; Supplementary Figure S1.

GO function annotation and KEGG pathway analysis

In the ceRNA regulatory network, GO analysis of 5 mRNAs indicated a major enrichment in: cortical actin cytoskeleton, protein C-terminus binding and adrenergic receptor binding. The histogram is displayed in Figures 6A–C. The results of the KEGG pathway analysis indicated that the main enrichment occurred in homologous recombination. (Figure 6D).

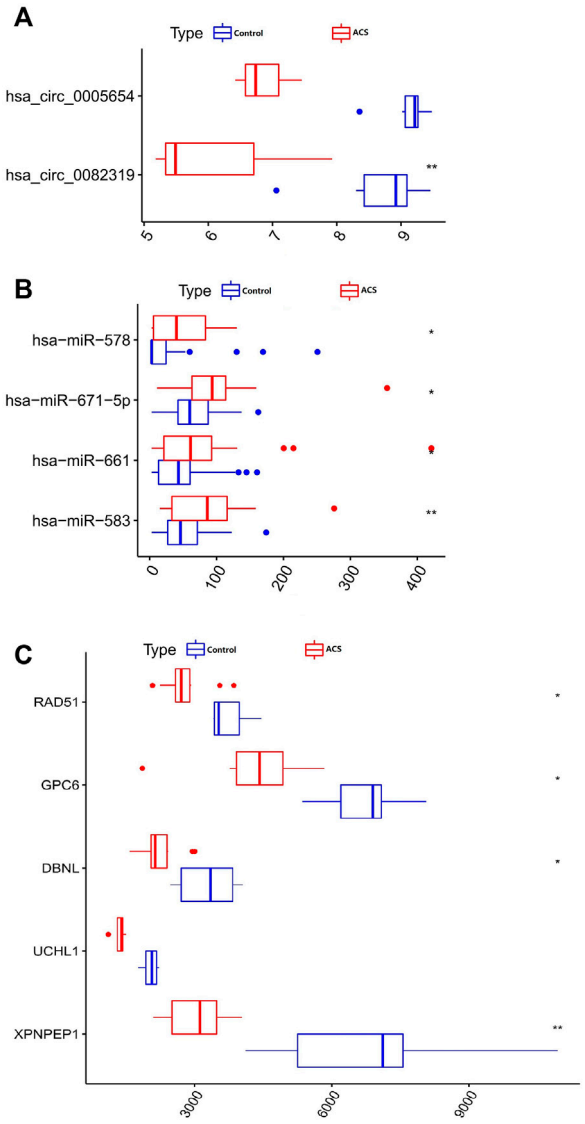


FIGURE 5
Boxplot of circRNA, miRNA, and mRNA in the ceRNA regulatory network. **(A)** 2 circRNAs. **(B)** 4 miRNAs. **(C)** 5 mRNAs. *$p < 0.05$, **$p < 0.01$.

Quantitative RT-PCR verification of hub genes

The information of datasets and verification patient were shown in the Supplementary Table S2. In order to keep consistency with patients in the GEO datasets, our research objects included STEMI, NSTEMI, and UA patients, and the collection time of samples was limited to 12 h after the onset of symptoms. The expression of the five hub genes (*XPNPEP1, UCHL1, DBNL, GPC6,* and *RAD51*) in the ceRNA network was analyzed by quantitative RT-PCR. Between the control and ACS groups, we found that the *XPNPEP1, UCHL1, GPC6,* and *RAD51* genes had significantly decreased expression in ACS patients. The *DBNL* gene did not show any significant difference (Figure 7).

Predicting the protective value of characteristic biomarkers in acute coronary syndrome

We compare the different expressions of *XPNPEP1, UCHL1, GPC6,* and *RAD51* in the GSE95368 and GSE60993 datasets (Figure 8; Supplementary Figure S2). Furthermore, we used the ROC curve to analyze the 4 mRNA from the GSE95368 and GSE60993 datasets for predicting and verifying characteristic biomarkers in ACS respectively. In the GSE95368 dataset, ROC curve analysis discovered that the *XPNPEP1, UCHL1, GPC6,* and *RAD51* genes have important clinical significance in preventing the occurrence and excluding diagnosis of ACS (Figures 9A–D). The AUC values of *XPNPEP1* in the GSE95368 dataset and the GSE60993 dataset were 1.000 (95% CI: 1.000–1.000, $p < 0.05$) and 0.777 (95% CI: 0.577–0.934, $p < 0.05$) dataset, respectively (Figure 9D; Supplementary Figure S3D). Logistic regression analysis showed that the *XPNPEP1* gene was significantly related to the occurrence of MACE in ACS patients [OR = 7.408 (95% CI: 0.762–72.009, $p < 0.05$)].

Expression of XPNPEP1 in cardiomyocytes

To explore the relationship between XPNPEP1 and ACS, we performed immunofluorescence staining on hiPSC-CM from patients with ACS and healthy human respectively. The results of immunofluorescence staining indicated that XPNPEP1 (red fluorescence) was expressed low in the cardiomyocytes of ACS patients (Figure 10).

Correlation analysis between acute coronary syndrome hub genes and immune infiltrating cells

Considering that ACS has been shown to be infiltrated with a large number of immune cells in previous studies (Kyaw et al., 2021). We analyzed the infiltrating immune cells between the different groups' mRNA in ACS patients (Supplementary Figures S4, S5). Moreover, we analyzed the relationship between the expression of the four hub genes and immune cell infiltration in ACS patients using the ssGSEA method (Figure 11). The results were as follows: *GPC6* was positively correlated with Eosinophil. RAD51 was negatively correlated with T follicular and positively correlated with Type 2 T helper cell. *UCHL1* was positively correlated with Imumature dendritic cell and Type 2 T helper cell. *XPNPEP1* levels were not statistically significantly correlated with immune cell infiltration.

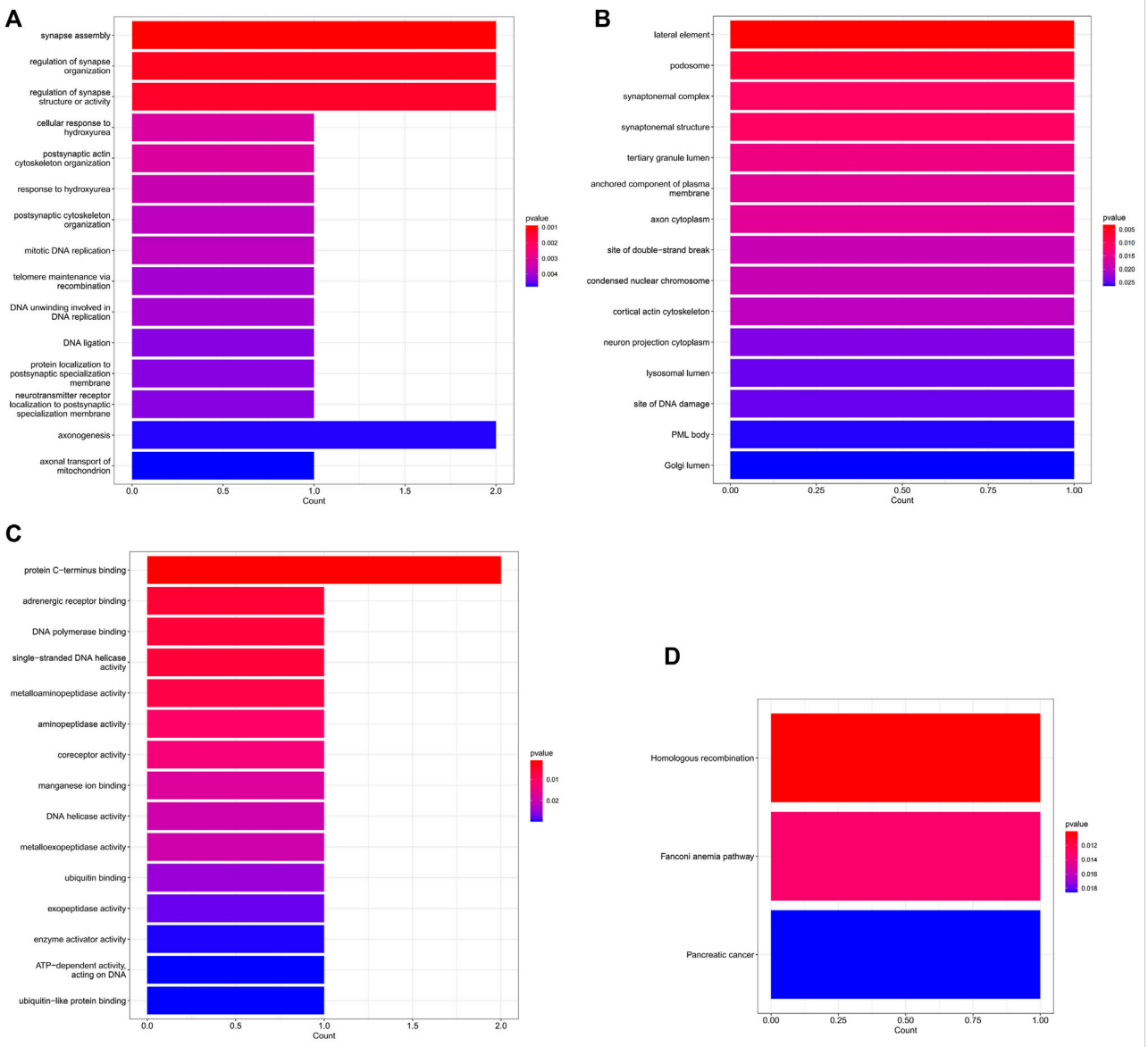


FIGURE 6
The GO and KEGG enrichment analysis of mRNA in the ceRNA network. **(A)** The top 15 GO terms of biological process. **(B)** The top 15 GO terms of cellular component. **(C)** The top 15 GO terms of molecular function. **(D)** The 3 KEGG pathways.

Discussion

In community, the incidence of missed ACS cases is 3.8 per 1,000 person-years among people aged over 55 years (de Torbal et al., 2006). Electrocardiogram and cardiac troponin I can provide effective help for rapid diagnosis, but their time window, specificity, and sensitivity have certain limitations (Lindahl, 2001; Chang et al., 2008). Therefore, it is necessary to study new biomarkers for early diagnosis or exclude ACS. The Gene chip microarray system has been widely employed in heart disease research since the development of gene chip technology (Chen et al., 2019). CircRNAs have long been thought to be a critical regulator in the pathogenesis of ACS, and their aberrant expression has been shown to have a major impact on disease progression (Cai et al., 2019; Si et al., 2020; Zhang et al., 2020).

CircRNA affects gene transcription and regulation by interacting with miRNA, mRNA, or protein *via* the ceRNA mechanism (Lasda and Parker, 2014). In our ceRNA networks, it was found that the downregulation of circRNA (hsa_circ_0082319 and hsa_circ_0005654) and downregulation of mRNA (*XPNPEP1, UCHL1, DBNL, GPC6,* and *RAD51*) formed a competitive relationship, which jointly led to the occurrence of ACS.

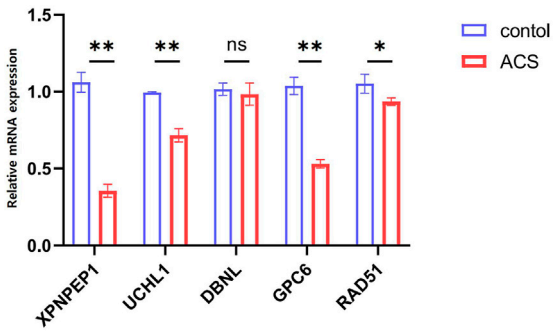


FIGURE 7
The expression levels of five hub genes in the ceRNA network analyzed by quantitative RT-PCR.*$p < 0.05$, **$p < 0.01$. Control group: $n = 5$, ACS group: $n = 8$.

It indicates that the expression of *UCHL1*, *DBNL*, *GPC6*, *RAD51*, and *XPNPEP1* can protect against ACS.

Every physician knows that the typical symptom of ACS is chest pain or pressure radiating across the chest and down the left arm, but some ACS can also present as dyspnoea, isolated jaw or arm pain, bilateral arm pain and back pain, or nausea and vomiting without any pain (Brieger et al., 2004; Wu et al., 2018; Chang et al., 2019; Kwok et al., 2021), which can lead both patients and physicians to fail to recognize ACS. Because of these factors, objective detection is critical for confirming or excluding ACS. Our study found that the *XPNPEP1* gene is the protective gene of ACS. By further plotting and analyzing the ROC curve, it can be utilized as a gene to exclude diagnosis. Outpatient follow-up of ACS patients found that the *XPNPEP1* gene was related to the occurrence of MACE in patients. Immunofluorescence staining showed that *XPNPEP1* was low expressed in the control group compared with ACS patients. The possible mechanism is that inhibition of *XPNPEP1* expression leads to the activation of CARD8 (caspase activation and recruitment domain 8) (Rao et al., 2022), which regulates the expression of cytokines and chemokines in endothelial cells and atherosclerotic lesions (Paramel et al., 2020). We speculate that the *XPNPEP1* gene indirectly affects the occurrence and development of atherosclerosis by participating in the function and metabolic mechanism of vascular endothelial cells. It is an

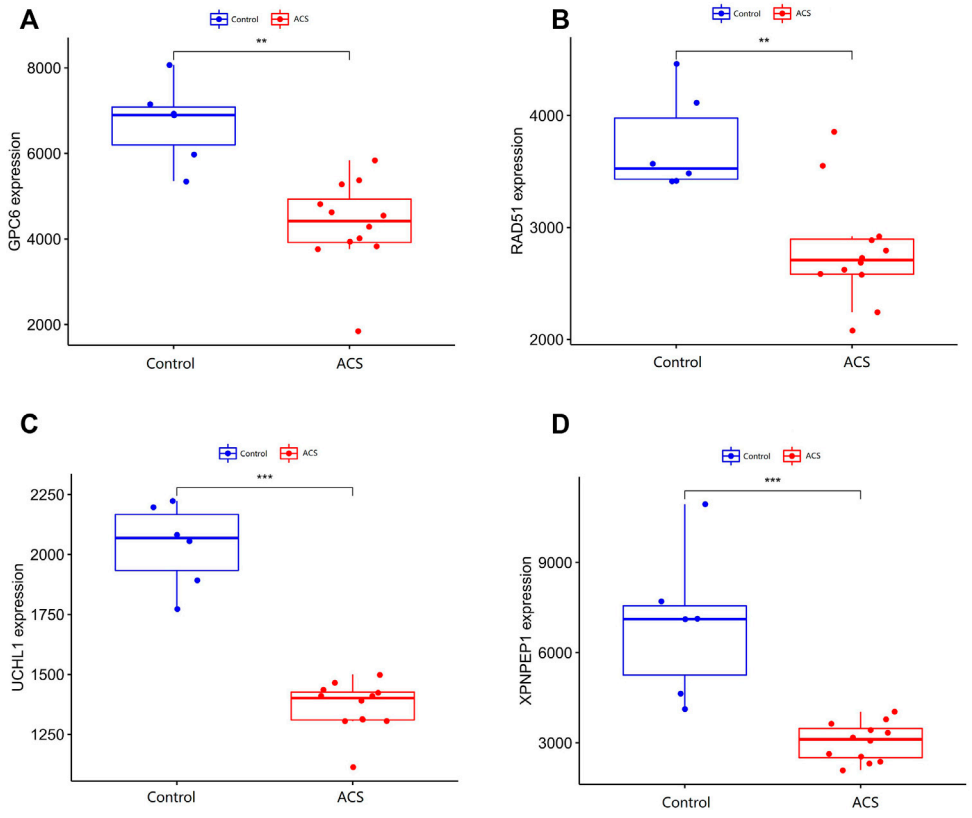


FIGURE 8
Validation of differentially expressed genes in ACS. **(A–D)** shows the expression of differentially expressed genes in the GSE95368 data set in ACS and non-ACS patients. The red box represents gene expression in the ACS group, and the bule box represents gene expression in the healthy control group.*$p < 0.05$, **$p < 0.01$, ***$p < 0.01$.

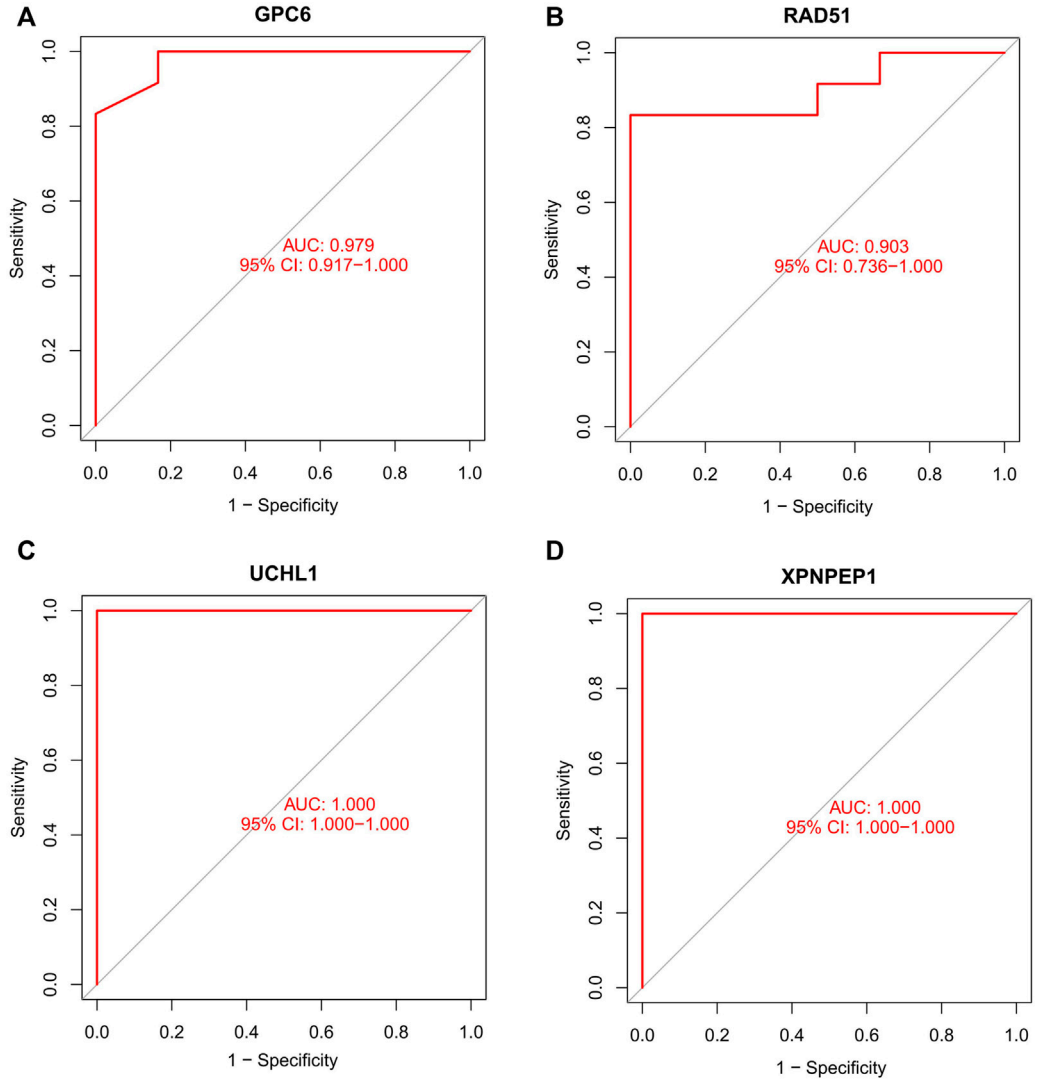


FIGURE 9
Receiver operating characteristic (ROC) curve of differentially gene's ability to exclude ACS diagnosis in the GSE95368 dataset.

important clue to clarify the regulatory mechanism of *XPNPEP1* at gene and protein levels in the pathogenesis of ACS.

It is known that *UCHL1* (Ubiquitin carboxyl-terminal hydrolase L1) is highly expressed and plays an important role in neurons, and it is usually used as a neuronal marker (Day and Thompson, 2010; Matuszczak et al., 2020). In recent years, there have also been some studies on the role of *UCHL1* in the cardiovascular system. Geng et al. (2022) compared the ischemic heart injury group with the control group and found the overexpression of *UCHL1* has a protective effect on myocardial injury after myocardial infarction. Upregulation of *UCHL1* can prevent cardiac remodeling and dysfunction after myocardial infarction by supporting autophagy flow and protein homeostasis (Wu et al., 2022). For the *GPC6* gene, a previous study found that compared to patients with heart failure after myocardial infarction, patients without heart failure had significantly lower glypican-6, suggesting that *GPC6* may be a protective gene for heart failure after myocardial infarction (Ozturk et al., 2021). Previous studies have found that RAD51 is a DNA damage repair molecule and it is involved in the pathogenesis of atherosclerosis (Davies et al., 2001; Singh et al., 2020). The conclusions of above studies are consistent with our results: *UCHL1, GPC6,* and *RAD*51 indirectly inhibit the occurrence of atherosclerosis or myocardial injury through various molecular biological level regulatory mechanisms.

ACS is not only an obstructive vascular disease, but also involves chronic vascular inflammation that causes atherosclerosis (Zhao and Mallat, 2019). For example,

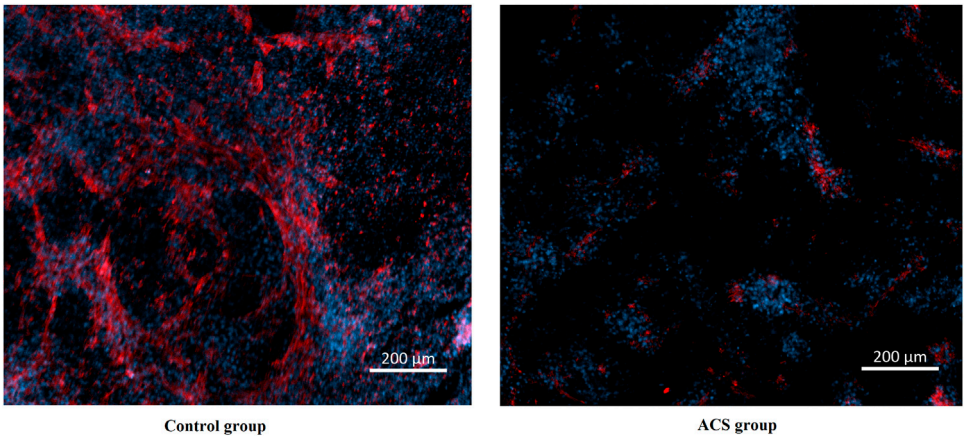


FIGURE 10
Immunofluorescence stain images of cardiomyocytes derived from human induced pluripotent stem cells. XPNPEP1: red. DNA: blue.

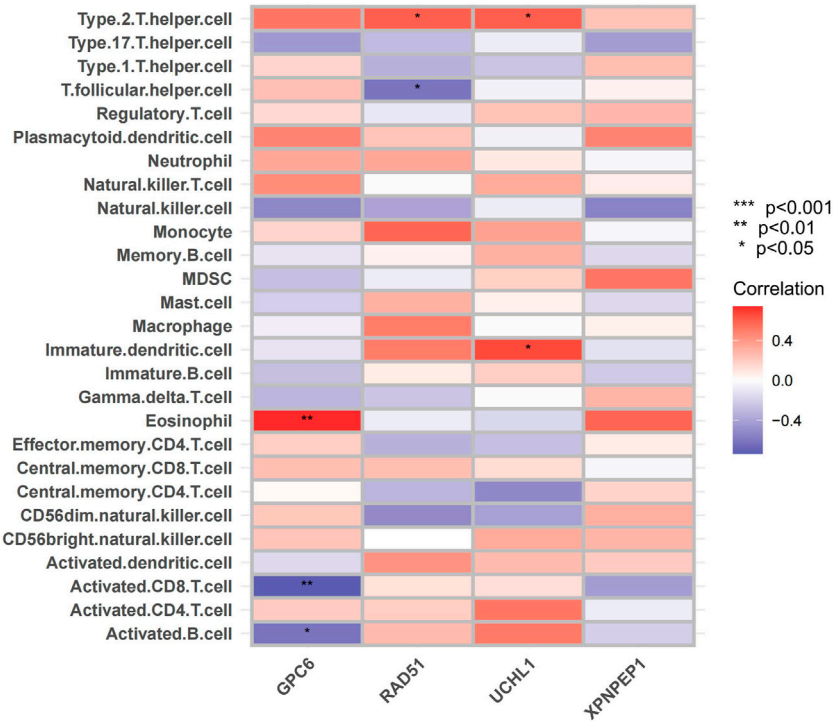


FIGURE 11
Immune cell infiltration analysis. The correlation analysis between immune cells: The red square represents positively correlated with upregulated immune cells. The bule square represents negatively correlated with decreased immune cells.

monocytes and macrophages play a pivotal role in the initiation, progression and instability of atherosclerotic plaque through multiple mechanisms such as necrosis and subsequent release of proinflammatory factors (Wei et al., 2018). It is not clear whether the hub genes in our study influence in the pathogenesis of ACS by immune cell infiltration. Therefore, we determined the relationship between the expression of three hub genes and immune cell infiltration by the ssGSEA method. The hub genes were highly enriched in immune-related or inflammation-related responses and pathways, which means that immune infiltration may be closely involved in the regulation of ACS pathogenesis.

The GO functional annotation revealed that hub mRNAs implicated in cortical actin cytoskeleton, protein C-terminus binding, and adrenergic receptor binding were overrepresented. Previous study has shown that the cortical actin cytoskeleton was involved in the pathogenesis of ACS through inflammatory and oxidative stress mechanisms (Hill et al., 2022). It is unknown how C-terminal protein plays a role in the pathophysiology of ACS. We speculate that the C-terminal protein may affect the pathogenesis of ACS as an inflammatory mediator (Hansson, 2005). According to a previous study, the adrenergic receptor blocker metoprolol has a special effect on neutrophils when inflammation intensifies, which provides cardiac protection (Clemente-Moragon et al., 2020). Adrenergic receptors are one of the targets for promoting the proliferation of adult cardiomyocytes and cardiac regeneration, which opens up a possible avenue for myocardial repair after myocardial infarction (Du et al., 2022). Through GO analysis of key mRNA, we found that the inflammatory mechanism is the basic pathological mechanism of ACS, and also explained that the adrenergic receptors are involved in the mechanism of myocardial damage repair after ACS.

The novelty of our study is that the results are based on the circRNA-associated ceRNA network in ACS patients. At the same time, the hub genes in the network were validated and analyzed by ROC curve and immune cells infiltration. At the same time, our research also has some deficiencies. On the one hand, because the sample size of the circRNA dataset is relatively small, no upregulated circRNA and mRNA were found in the ceRNA network of our study. On the other hand, detailed data for long-term follow-up to check for prognosis-predicting mRNAs is lacking. However, our research indirectly proves that the key RNAs may predict protective gene biomarkers of ACS and lays a solid framework for future research.

Conclusion

Timely diagnosis and treatment of ACS can help to improve the prognosis of patients and reduce the economic and social damage. Our study constructed a circRNA-related ceRNA network in ACS. We found that hsa_circ_0082319 and hsa_circ_0005654 were involved in the regulation of *XPNPEP1, UCHL1, DBNL, GPC6,* and *RAD51* expression or activity through sponging hsa-miR-583, hsa-miR-661, hsa-miR-671-5p and hsa-miR-578. The *XPNPEP1* gene could be a protective gene biomarker for ACS. Three other hub genes (*UCHL1, GPC6,* and *RAD51*) were significantly correlated with immune cells (Eosinophil, T follicular, Type 2 T helper cell and Imumature dendritic cell) in ACS.

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 studies involving human participants were reviewed and approved by the First Affiliated Hospital of Henan University of Science and Technology. The patients/participants provided their written informed consent to participate in this study.

Author contributions

HZ and DM conceived and designed the study, conducted the experiments, analyzed the data, interpreted the results, and drafted the manuscript. PZ and HZ prepared the charts. YW and LD help HZ analyzed the data. LK modify picture and language. All authors contributed to the article and approved the submitted version.

Funding

This research was funding by the China Scholarship Council (CSC, No. 202108410141). This research was funding by the cooperative project of the Henan Medical Science and Technology Research Program (No. LHGJ20200568). This research was funding by Luoyang Science and Technology Foundation, Medical and Health Project (No. 2030002A).

Conflict of interest

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

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.1030510/full#supplementary-material

References

Brieger, D., Eagle, K. A., Goodman, S. G., Steg, P. G., Budaj, A., White, K., et al. (2004). Acute coronary syndromes without chest pain, an underdiagnosed and undertreated high-risk group: Insights from the global registry of acute coronary events. *Chest* 126 (2), 461–469. doi:10.1378/chest.126.2.461

Cai, L., Qi, B., Wu, X., Peng, S., Zhou, G., Wei, Y., et al. (2019). Circular RNA Ttc3 regulates cardiac function after myocardial infarction by sponging miR-15b. *J. Mol. Cell. Cardiol.* 130, 10–22. doi:10.1016/j.yjmcc.2019.03.007

Chang, M. P., Vigen, R., Sharma, S., and Diercks, D. B. (2019). Possible missed acute coronary syndrome rate in north Texas: Is there room to improve? *Crit. Pathw. Cardiol.* 18 (3), 121–124. doi:10.1097/hpc.0000000000000188

Chang, S. A., Choi, S. I., Choi, E. K., Kim, H. K., Jung, J. W., Chun, E. J., et al. (2008). Usefulness of 64-slice multidetector computed tomography as an initial diagnostic approach in patients with acute chest pain. *Am. Heart J.* 156 (2), 375–383. doi:10.1016/j.ahj.2008.03.016

Chen, D. Q., Kong, X. S., Shen, X. B., Huang, M. Z., Zheng, J. P., Sun, J., et al. (2019). Identification of differentially expressed genes and signaling pathways in acute myocardial infarction based on integrated bioinformatics analysis. *Cardiovasc. Ther.* 2019, 8490707. doi:10.1155/2019/8490707

Chen, G., Gulbranson, D. R., Hou, Z., Bolin, J. M., Ruotti, V., Probasco, M. D., et al. (2011). Chemically defined conditions for human iPSC derivation and culture. *Nat. Methods* 8 (5), 424–429. doi:10.1038/nmeth.1593

Clemente-Moragón, A., Gómez, M., Villena-Gutiérrez, R., Lalama, D. V., García-Prieto, J., Martínez, F., et al. (2020). Metoprolol exerts a non-class effect against ischaemia-reperfusion injury by abrogating exacerbated inflammation. *Eur. Heart J.* 41 (46), 4425–4440. doi:10.1093/eurheartj/ehaa733

Collet, J. P., Thiele, H., Barbato, E., Barthélémy, O., Bauersachs, J., Bhatt, D. L., et al. (2021). 2020 ESC Guidelines for the management of acute coronary syndromes in patients presenting without persistent ST-segment elevation. *Eur. Heart J.* 42 (14), 1289–1367. doi:10.1093/eurheartj/ehaa575

Davies, A. A., Masson, J. Y., McIlwraith, M. J., Stasiak, A. Z., Stasiak, A., Venkitaraman, A. R., et al. (2001). Role of BRCA2 in control of the RAD51 recombination and DNA repair protein. *Mol. Cell* 7 (2), 273–282. doi:10.1016/s1097-2765(01)00175-7

Day, I. N., and Thompson, R. J. (2010). UCHL1 (pgp 9.5): Neuronal biomarker and ubiquitin system protein. *Prog. Neurobiol.* 90 (3), 327–362. doi:10.1016/j.pneurobio.2009.10.020

de Torbal, A., Boersma, E., Kors, J. A., van Herpen, G., Deckers, J. W., van der Kuip, D. A., et al. (2006). Incidence of recognized and unrecognized myocardial infarction in men and women aged 55 and older: The rotterdam study. *Eur. Heart J.* 27 (6), 729–736. doi:10.1093/eurheartj/ehi707

Du, J., Zheng, L., Gao, P., Yang, H., Yang, W. J., Guo, F., et al. (2022). A small-molecule cocktail promotes mammalian cardiomyocyte proliferation and heart regeneration. *Cell Stem Cell* 29 (4), 545–558.e13. e513. doi:10.1016/j.stem.2022.03.009

Dudekula, D. B., Panda, A. C., Grammatikakis, I., De, S., Abdelmohsen, K., and Gorospe, M. (2016). CircInteractome: A web tool for exploring circular RNAs and their interacting proteins and microRNAs. *RNA Biol.* 13 (1), 34–42. doi:10.1080/15476286.2015.1128065

Geng, B., Wang, X., Park, K. H., Lee, K. E., Kim, J., Chen, P., et al. (2022). UCHL1 protects against ischemic heart injury via activating HIF-1α signal pathway. *Redox Biol.* 52, 102295. doi:10.1016/j.redox.2022.102295

Hansen, T. B., Jensen, T. I., Clausen, B. H., Bramsen, J. B., Finsen, B., Damgaard, C. K., et al. (2013). Natural RNA circles function as efficient microRNA sponges. *Nature* 495 (7441), 384–388. doi:10.1038/nature11993

Hansson, G. K. (2005). Inflammation, atherosclerosis, and coronary artery disease. *N. Engl. J. Med.* 352 (16), 1685–1695. doi:10.1056/NEJMra043430

Hill, M. A., Jaisser, F., and Sowers, J. R. (2022). Role of the vascular endothelial sodium channel activation in the Genesis of pathologically increased cardiovascular stiffness. *Cardiovasc. Res.* 118 (1), 130–140. doi:10.1093/cvr/cvaa326

Ibanez, B., James, S., Agewall, S., Antunes, M. J., Bucciarelli-Ducci, C., Bueno, H., et al. (2018). 2017 ESC Guidelines for the management of acute myocardial infarction in patients presenting with ST-segment elevation: The Task Force for the management of acute myocardial infarction in patients presenting with ST-segment elevation of the European Society of Cardiology (ESC). *Eur. Heart J.* 39 (2), 119–177. doi:10.1093/eurheartj/ehx393

Kanehisa, M., and Goto, S. (2000). Kegg: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* 28 (1), 27–30. doi:10.1093/nar/28.1.27

Kwok, C. S., Bennett, S., Azam, Z., Welsh, V., Potluri, R., Loke, Y. K., et al. (2021). Misdiagnosis of acute myocardial infarction: A systematic review of the literature. *Crit. Pathw. Cardiol.* 20 (3), 155–162. doi:10.1097/hpc.0000000000000256

Kyaw, T., Loveland, P., Kanellakis, P., Cao, A., Kallies, A., Huang, A. L., et al. (2021). Alarmin-activated B cells accelerate murine atherosclerosis after myocardial infarction via plasma cell-immunoglobulin-dependent mechanisms. *Eur. Heart J.* 42 (9), 938–947. doi:10.1093/eurheartj/ehaa995

Lasda, E., and Parker, R. (2014). Circular RNAs: Diversity of form and function. *Rna* 20 (12), 1829–1842. doi:10.1261/rna.047126.114

Lindahl, B. (2001). Detection of myocardial damage - are the troponins the ultimate solution? *Scand. Cardiovasc. J.* 35 (4), 229–232. doi:10.1080/14017430152581323

Matuszczak, E., Tylicka, M., Komarowska, M. D., Debek, W., and Hermanowicz, A. (2020). Ubiquitin carboxy-terminal hydrolase L1 - physiology and pathology. *Cell biochem. Funct.* 38 (5), 533–540. doi:10.1002/cbf.3527

Ozturk, F., Atici, A., and Barman, H. A. (2021). Can glypican-6 level predict ejection fraction decline after myocardial infarction? *Angiology* 72 (6), 582–588. doi:10.1177/0003319720968376

Paramel, G. V., Karadimou, G., Eremo, A. G., Ljungberg, L. U., Hedin, U., Olofsson, P. S., et al. (2020). Expression of CARD8 in human atherosclerosis and its regulation of inflammatory proteins in human endothelial cells. *Sci. Rep.* 10 (1), 19108. doi:10.1038/s41598-020-73600-4

Rao, S. D., Chen, Q., Wang, Q., Orth-He, E. L., Saoi, M., Griswold, A. R., et al. (2022). M24B aminopeptidase inhibitors selectively activate the CARD8 inflammasome. *Nat. Chem. Biol.* 18 (5), 565–574. doi:10.1038/s41589-021-00964-7

Salmena, L., Poliseno, L., Tay, Y., Kats, L., and Pandolfi, P. P. (2011). A ceRNA hypothesis: The rosetta stone of a hidden RNA language? *Cell* 146 (3), 353–358. doi:10.1016/j.cell.2011.07.014

Sanchis-Gomar, F., Perez-Quilis, C., Leischik, R., and Lucia, A. (2016). Epidemiology of coronary heart disease and acute coronary syndrome. *Ann. Transl. Med.* 4 (13), 256. doi:10.21037/atm.2016.06.33

Sayers, E. W., Bolton, E. E., Brister, J. R., Canese, K., Chan, J., Comeau, D. C., et al. (2022). Database resources of the national center for biotechnology information. *Nucleic Acids Res.* 50 (D1), D20–d26. doi:10.1093/nar/gkab1112

Si, X., Zheng, H., Wei, G., Li, M., Li, W., Wang, H., et al. (2020). circRNA Hipk3 induces cardiac regeneration after myocardial infarction in mice by binding to Notch1 and miR-133a. *Mol. Ther. Nucleic Acids* 21, 636–655. doi:10.1016/j.omtn.2020.06.024

Singh, S., Nguyen, H., Michels, D., Bazinet, H., Matkar, P. N., Liu, Z., et al. (2020). BReast CAncer susceptibility gene 2 deficiency exacerbates oxidized LDL-induced DNA damage and endothelial apoptosis. *Physiol. Rep.* 8 (13), e14481. doi:10.14814/phy2.14481

Sposito, A. C. (2022). Soluble LOX-1 levels during acute coronary syndrome: A potent and multifaceted warning sign for cardiovascular risk. *Eur. Heart J.* 43 (19), 1861–1863. doi:10.1093/eurheartj/ehac173

van den Broek, W. W. A., and Ten Berg, J. M. (2022). Is a genotype-guided therapy the optimal strategy to personalize anti-thrombotic management in patients with acute coronary syndrome? *Eur. Heart J.* 43 (30), 2762–2763. doi:10.1093/eurheartj/ehac219

Wei, Y., Zhu, M., and Schober, A. (2018). Macrophage MicroRNAs as therapeutic targets for atherosclerosis, metabolic syndrome, and cancer. *Int. J. Mol. Sci.* 19 (6), E1756. doi:10.3390/ijms19061756

Wu, J., Gale, C. P., Hall, M., Dondo, T. B., Metcalfe, E., Oliver, G., et al. (2018). Editor's choice - impact of initial hospital diagnosis on mortality for acute myocardial infarction: A national cohort study. *Eur. Heart J. Acute Cardiovasc. Care* 7 (2), 139–148. doi:10.1177/2048872616661693

Wu, P., Li, Y., Cai, M., Ye, B., Geng, B., Li, F., et al. (2022). Ubiquitin carboxyl-terminal hydrolase L1 of cardiomyocytes promotes macroautophagy

and proteostasis and protects against post-myocardial infarction cardiac remodeling and heart failure. *Front. Cardiovasc. Med.* 9, 866901. doi:10.3389/fcvm.2022.866901

Xiong, D. D., Dang, Y. W., Lin, P., Wen, D. Y., He, R. Q., Luo, D. Z., et al. (2018). A circRNA-miRNA-mRNA network identification for exploring underlying pathogenesis and therapy strategy of hepatocellular carcinoma. *J. Transl. Med.* 16 (1), 220. doi:10.1186/s12967-018-1593-5

Zhang, R. J., Li, Y., Liu, Q., Gao, Y. J., Du, J., Ma, J., et al. (2019). Differential expression profiles and functional prediction of circular RNAs and long non-coding RNAs in the Hippocampus of nrf2-knockout mice. *Front. Mol. Neurosci.* 12, 196. doi:10.3389/fnmol.2019.00196

Zhang, S., Wang, W., Wu, X., and Zhou, X. (2020). Regulatory roles of circular RNAs in coronary artery disease. *Mol. Ther. Nucleic Acids* 21, 172–179. doi:10.1016/j.omtn.2020.05.024

Zhao, C., Liu, J., Ge, W., Li, Z., Lv, M., Feng, Y., et al. (2020). Identification of regulatory circRNAs involved in the pathogenesis of acute myocardial infarction. *Front. Genet.* 11, 626492. doi:10.3389/fgene.2020.626492

Zhao, T. X., and Mallat, Z. (2019). Targeting the immune system in atherosclerosis: JACC state-of-the-art review. *J. Am. Coll. Cardiol.* 73 (13), 1691–1706. doi:10.1016/j.jacc.2018.12.083

Zhao, Z., Wang, K., Wu, F., Wang, W., Zhang, K., Hu, H., et al. (2018). circRNA disease: a manually curated database of experimentally supported circRNA-disease associations. *Cell Death Dis.* 9 (5), 475. doi:10.1038/s41419-018-0503-3

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 08 November 2022
DOI 10.3389/fgene.2022.1002344

OPEN ACCESS

EDITED BY
William C. Cho,
QEH, Hong Kong SAR, China

REVIEWED BY
Ting Gong,
University of Hawaii at Manoa, United States
He Song,
University of California, Los Angeles, United States

*CORRESPONDENCE
Zhong Wang,
wangzhong523@vip.163.com

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

RECEIVED 25 July 2022
ACCEPTED 26 October 2022
PUBLISHED 08 November 2022

CITATION
Liao H, Chai Y, Sun Y, Guo Z, Wang X, Wang Z, Wang Z and Wang Z (2022), Hsa_circ_0074158 regulates the endothelial barrier function in sepsis and its potential value as a biomarker. *Front. Genet.* 13:1002344. doi: 10.3389/fgene.2022.1002344

Hsa_circ_0074158 regulates the endothelial barrier function in sepsis and its potential value as a biomarker

Haiyan Liao[1], Yan Chai[1], Yuming Sun[2], Zhe Guo[1,3], Xuesong Wang[1], Ziyi Wang[1], Ziwen Wang[1] and Zhong Wang[1,3]*

[1]School of Clinical Medicine, Tsinghua University, Beijing, China, [2]Clinical Laboratory, Beijing Tsinghua Changgung Hospital, School of Clinical Medicine, Tsinghua University, Beijing, China, [3]General Department, Beijing Tsinghua Changgung Hospital, School of Clinical Medicine, Tsinghua University, Beijing, China

Background: Sepsis is one of the main causes of death in critically ill patients with high morbidity and mortality. Circular RNAs (CircRNAs) are aberrantly expressed, and play significant regulatory roles in many diseases. However, the expression profiles and functions of circRNAs in sepsis have not yet been fully clarified.

Methods: Our present study performed an RNA sequencing (RNA-seq) analysis to assess the expression profiles of circRNAs *in vitro*. We applied the quantitative real-time polymerase chain reaction (RT-qPCR) to verify the RNA-seq results. The analyses of Gene Ontology (GO), Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway, the competitive endogenous RNA (ceRNA) regulatory networks, were performed to explore the potential mechanism in sepsis. And then, significantly up-regulated differentially expressed (DE) circRNA, hsa_circ_0074158, was selected for further study. Hsa_circ_0074158 was silenced to investigate its regulatory function in sepsis, and the barrier function was also examined *in vitro*. Endothelial cell junctions were valued using Vascular endothelial cadherin (VE-cadherin), which was detected by immunofluorescence staining. We measured endothelial permeability by transendothelial electrical resistance (TEER) and fluorescein isothiocyanate (FITC)-dextran extravasation.

Results: In total, 203 significantly DE circRNAs, including 77 up-regulated and 126 down-regulated, were identified. *In vitro,* the RT-qPCR assay showed that the expression pattern of hsa_circ_0074158, hsa_circ_RSBN1L_11059, hsa_circ_0004188, and hsa_circ_0005564 were consistent with the results from RNA-seq analysis. The expression of hsa_circ_0074158 detected by RT-

Abbreviations: CircRNAs, Circular RNAs; RT-qPCR, Quantitative real-time polymerase chain reaction; GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; CeRNA, Competitive endogenous RNA; DE, Differentially expressed; VE-cadherin, Vascular endothelial cadherin; TEER, Transendothelial electrical resistance; FITC, Fluorescein isothiocyanate; LPS, Lipopolysaccharide; HUVECs, Human umbilical vein endothelial cells; DMEM, Dulbecco's modified Eagle's Medium; FBS, Fetal bovine serum; PBS, Phosphate buffer saline; RNA-seq, RNA-sequencing; RNase, Ribonuclease; BP, Biological process; CC, Cellular component; MF, Molecular function.

qPCR *in vivo* was also consistent with the RNA-seq results. The ceRNA networks, GO enrichment, and the KEGG pathway analyses revealed that circRNAs may be related to the barrier function in sepsis. The immunofluorescence assay showed that the suppression of hsa_circ_0074158 expression significantly enhanced the expression of VE-cadherin, which was suppressed in lipopolysaccharide (LPS)-induced sepsis. Additionally, hsa_circ_0074158 knockdown could partially reverse the LPS-induced TEER reduction and FITC-dextran extravasation elevation in sepsis.

Conclusion: In conclusion, we have found DE circRNAs could serve as potential biomarkers and therapeutic targets for sepsis. Hsa_circ_0074158 plays a vital role in sepsis and is related to the disruption of the endothelial barrier.

KEYWORDS

barrier function, RNA-seq, CircRNAs, adherens junction, sepsis

1 Introduction

1.1 Background

Sepsis is defined as "life-threatening organ dysfunction caused by a dysregulated host response to infection" (Singer et al., 2016). Sepsis is also considered one of the leading causes of death in critically ill patients with high morbidity and mortality (Vincent et al., 2014; Machado et al., 2017). Multiple factors are involved in the complex pathophysiological process of sepsis, including impaired vascular endothelial barrier, enhanced vascular permeability, inflammatory response, immune response, and coagulation dysfunction. Sepsis is a serious threat to human health and the prevalence of sepsis is considered a significant burden for health systems. Research from 1979 to 2015 in seven high-income countries showed that morbidity was 288 hospital-treated sepsis cases and 148 hospital-treated severe sepsis cases per 100,000 person-years, while morbidity was 437 sepsis and 270 severe sepsis cases per 100,000 person-years in the last decade, with hospital mortality 17% for sepsis and 26% for severe sepsis (Fleischmann et al., 2016).

The vascular endothelium is composed of a monolayer of endothelial cells, a basement membrane, an extracellular matrix, and endothelial glycocalyx, which lines the luminal surface of the inner blood of vessels with extensive homeostasis functions (Rajendran et al., 2013). Vascular endothelial cadherin (VE-cadherin) is concentrated in adherent junctions and found exclusively in the vascular endothelium in vertebrates, additionally, it also plays an important role in stabilizing the barrier function of the endothelium (Brasch et al., 2011). As a dynamic and heterogeneous organ, the endothelium is involved in many biological functions, such as the secretory, metabolic, and immunologic functions (Rajendran et al., 2013; Godo and Shimokawa, 2017). Through the synthesizing and release of various relaxing factors, such as vasodilator prostaglandins, nitric oxide, hyperpolarization factors, and contracting factors, the endothelium is tightly involved in the regulation of vascular tone (Nava and Llorens, 2019; Shimokawa and Godo, 2020; Godo et al., 2021). Many pathological processes, including microvascular tone dysfunction, vascular permeability, inflammatory response, platelet adhesion and aggregation, coagulation and fibrinolysis, and immunological response, are significantly associated with endothelial dysfunction. The endothelium is considered a critical physical barrier in microcirculation, and in the physiopathological process of sepsis, endothelial dysfunction is a central event (Lee and Slutsky, 2010; Martin-Fernandez et al., 2021).

Circular RNAs (CircRNAs) were discovered for more than 40 years and also observed in the cytoplasmic fractions (Sanger et al., 1976; Hsu and Coca-Prados, 1979). CircRNAs are endogenous noncoding RNA in eukaryotic cell lines that are mainly derived from the pre-mRNAs by back-splicing of exons and are initially considered an aberrant splicing event during the gene transcription (Nigro et al., 1991; Cocquerelle et al., 1992; Cocquerelle et al., 1993; Pasman et al., 1996). Due to lacking polyadenylation [poly(A)] and capping, the 5′ and 3′ splice sites of circRNAs are covalently closed, and this structure determines its highly conservative, specific, and stable properties (Chen, 2016; Kristensen et al., 2019; Chen, 2020). CircRNAs are aberrantly expressed in many diseases and also act as miRNA and protein sponges to be involved in the regulation of different processes in sepsis (Ashwal-Fluss et al., 2014; Thomson and Dinger, 2016; Abdelmohsen et al., 2017; Panda, 2018; Kristensen et al., 2019; Beltran-Garcia et al., 2020; Chen, 2020). An increasing number of studies revealed that circRNAs have been significantly associated with the regulation of many physiological and pathological processes, such as innate immunity, cell proliferation and transformation, and neuronal function (Chen, 2020; Chen et al., 2021). Dysregulation of circRNAs was also observed in various human diseases, including neurogenesis, myogenesis, osteogenesis, cancer, cardiovascular disease, metabolic diseases, and other metabolism-associated diseases (Huang et al., 2021; Verduci

et al., 2021). Recent studies also demonstrated the potential values of circRNAs as diagnostic biomarkers and/or therapeutic targets in many diseases (Jiang et al., 2022; Wang et al., 2022; Zhang et al., 2022). Additionally, circRNAs have been reported to be used as new diagnostic biomarkers and molecular therapeutic targets (Tian et al., 2021). However, whether circRNAs are involved in the regulation of the endothelial barrier in sepsis is not fully explored.

Many RNA sequencing (RNA-seq) studies have found that differentially expressed (DE) circRNAs were involved in cellular functions (Chen, 2020; Chen et al., 2021; Huang et al., 2021; Verduci et al., 2021). In this study, we detected the expression profiles of circRNAs in lipopolysaccharide (LPS)-induced sepsis. Then we explored their value as biomarkers in sepsis and further explored the effect of hsa_circ_0074158 on the endothelial barrier.

2 Materials and methods

2.1 Samples and patients

In order to investigate DE circRNAs in LPS-induced sepsis *in vitro,* the human umbilical vein endothelial cells (HUVECs) were divided into two groups: 1) control group, treated with nothing; 2) LPS group, treated with 1 μg/ml LPS for 12 h. Each sample was collected for three independent biological replicates. The HUVECs came from the Institute of Immunology, Tsinghua University, and were grown in Dulbecco's modified Eagle's Medium (DMEM; Solarbio, China) supplemented with 10% fetal bovine serum (FBS; TianHang, China) at 37°C in a humidified atmosphere of 5% CO_2. LPS was purchased from Sigma-Aldrich (Cat No. L4516). After treatment, cells were washed once with Phosphate buffer saline (PBS; Solarbio, China). Subsequently, the samples were centrifuged at 1,500 rpm for 5 min and stored in a freezer at −80°C.

According to Sepsis −3, 44 patients with sepsis and 48 healthy individuals were enrolled at Beijing Tsinghua Changgung Hospital (Beijing, China) from April 2022 to June 2022. Whole blood samples were collected and the quantitative real-time polymerase chain reaction (RT-qPCR) was performed. The total RNA was immediately extracted using Hipure Blood RNA Kits (Magen, Guangzhou, China). All patients had no history of autoimmune disorders, neoplastic diseases, or oral immunosuppressants. This study was approved by the Ethics Committee (NCT05095324).

2.2 Construction of transcriptome libraries

The total RNA of HUVECs was extracted using the total RNA kit (TaKaRa, Japan), followed by the manufacturer's instructions. The concentration and purity of the extracted RNA were determined using Nanodrop One (Thermo Fisher, United States). The integrity and contamination of RNA were evaluated by agarose gel electrophoresis (Bio-RAD, United States). And the integrity of the RNA was further verified using an Agilent 2100 bioanalyzer (Agilent 2100, United States). The preparation of specific transcriptome libraries was performed by removing ribosomal RNA. RNA was degraded using Ribonuclease (RNase) H (TaKaRa, Japan). The first single-stranded cDNA was synthesized with reverse transcription of RNA, followed by the second double-stranded cDNA synthesized using dNTP (dUTP, dATP, dGTP, and dCTP). Subsequently, RT-qPCR was performed to amplify the sequences and then the preparation of the total RNA library was completed. The RT-qPCR and the Agilent 2100 bioanalyzer were used to control the quality and quantity of the library. Transcriptome sequencing was performed with Illumina PE150.

2.3 Sequence alignments

The alignment of valid sequencing data (clean reads) with the genome or transcriptome is the basis for subsequent analysis. We used Hisat2 software (version 2.0.1-beta) to align the transcriptome RNA-seq dataset. Hisat2 software is an upgrade of Tophat2 software, with a high mapping rate and high accuracy in finding Junction reads (Kim et al., 2015).

2.4 Differential expression analysis

The input data for the differential expression of circRNAs was the read counts data set obtained from the analysis of circRNA expression. For samples with biological duplications, the DESeq2 package was used for the analysis of miRNAs and circRNAs between the LPS and control groups (Love et al., 2014). For samples without biological duplications, DEGseq provides a TMM algorithm to standardize the read count data and then performs a difference analysis.

2.5 Enrichment analysis of differentially expressed circRNAs

After obtaining the DE circRNAs between the LPS and control groups, the host genes in each group were analyzed by enrichment of the Gene Ontology (GO) (version 1.1) and Kyoto Encyclopedia of Genes and Genomes (KEGG) (version 94.0) pathway, respectively. The GO is a major bioinformatics initiative to unify the representation of gene and gene product attributes across all species (Consortium, 2019). KEGG is a collection of databases dealing with genomes, biological pathways, diseases, drugs, and chemical substances (Kanehisa and Goto, 2000). The KEGG PATHWAY database is the core of the KEGG resource. The database is a collection of pathway maps that integrate many entities, including genes, proteins, RNAs,

chemical compounds, glycans, and chemical reactions, as well as disease genes and drug targets, which are stored as individual entries in the other KEGG databases. GO enrichment analysis consists of biological process (BP), cellular component (CC), and molecular function (MF), performed to construct gene annotations. The enrichment of the KEGG pathway was carried out to reveal the function and interactions among the DE genes, providing annotation information on signal transduction and disease pathways.

2.6 Construction of a regulatory network for the competitive endogenous RNA

CircRNA is known as a miRNA sponge. Through binding to target miRNA, circRNA can inhibit miRNA expression and subsequently affect the ceRNA network. Thus, analysis and identification of the binding sites of circRNAs and their target miRNAs are necessary. In this study, Starbase software (version 2.0) and miRanda software (version 3.3a) were used to predict target miRNAs and potential binding sites with DE circRNAs, and the target genes of predicted miRNAs were identified using miRWalk (version 2.0) (Li et al., 2014; Agarwal et al., 2015; Sticht et al., 2018).

2.7 Real-time polymerase chain reaction

The validation of the results obtained from the RNA-seq and the quantification of candidate circRNA expression were performed using RT-qPCR. Specific convergent primers that span the circRNA back-splice junction site were designed by the circPrimer (version 2.0) and Oligo 7 (version 7.37) software (Zhong et al., 2018). If there were no suitable convergent primers, divergent primers were designed similarly. The cDNA was synthesized using the PrimeScript™ RT reagent kit (TaKaRa, Beijing, China). The RT-qPCR was performed using TB Green Premix Ex Taq™ II (TaKaRa, Beijing, China). GAPDH was employed as an internal control. RT-qPCR was performed by three independent biological and technical duplications. The relative expression of the circRNAs was analyzed by the $2^{-\Delta\Delta CT}$ method.

2.8 Immunofluorescence staining

The HUVECs were seeded in 24-well plates and cultured for 24 h. After the indicated treatment, cells were washed with cold PBS and fixed in 4% paraformaldehyde. The cells were then blocked with 5% BSA in PBS for 1 h and incubated with specific primary antibodies (1:100; Abcam) at 4°C overnight. The cells were then incubated with a fluorochrome-labeled anti-rabbit secondary antibody (1:500; Beyotime) for 1 h at room temperature with protection from light. The cells were then washed with cold PBS three times and stained with DAPI (ready-to-use, Solarbio) for nucleus staining. The expression of VE-cadherin was observed by fluorescence microscopy (Olympus).

2.9 Transendothelial electrical resistance and fluorescein isothiocyanate -dextran assays

Cultivation of cell monolayers on the upper side of the Transwell insert (Corning, United States) at 37°C, 5% CO_2. In the upper and lower chambers, 200 μL and 600 μL culture medium were distributed, respectively. After treatment, the Volt-Ohm Meter (Millipore, United States) was used to measure TEER following the protocol. The values (Ω cm^2) of TEER were expressed by subtracting the resistance of the blank insert and correcting for the surface area.

Culture HUVECs as previously described. After treatment, FITC-dextran (1 mg/ml, Sigma-Aldrich, United States) was added to the upper chambers and cultured for 45 min 100 μL samples were collected from the upper and lower chambers for fluorescence. The dextran permeability coefficient (Pd) was used to evaluate the permeability of the endothelial monolayer.

$$Pd = \frac{[A]}{t} \times \frac{1}{A} \times \frac{V}{[L]}$$

$[A]$ is the dextran concentration of the lower chamber, t is time, A is the membrane area; V is the volume of the lower chamber $[L]$ is the dextran concentration of the upper chamber (Wu et al., 2020).

2.10 Statistical analysis

SPSS software (version 26.0) was used to analyze all the data collected from this study and the figures were produced using GraphPad Prism (version 8.0.2). Comparisons between two groups were analyzed using the student's t-test and comparisons between multiple groups were performed by one-way ANOVA. All data are presented as mean ± standard deviation (S.D.). The chi-square test or Fisher's exact test was applied to analyze the categorical variables. The p-value <0.05 was considered statistically significant.

3 Results

3.1 RNA-seq identified features of circRNAs and differentially expressed profiles

A total of 8766 circRNAs were identified by RNA-seq. The distribution of all cricRNAs on human chromosomes was analyzed and the results were shown in Figure 1A. Meanwhile, the length of

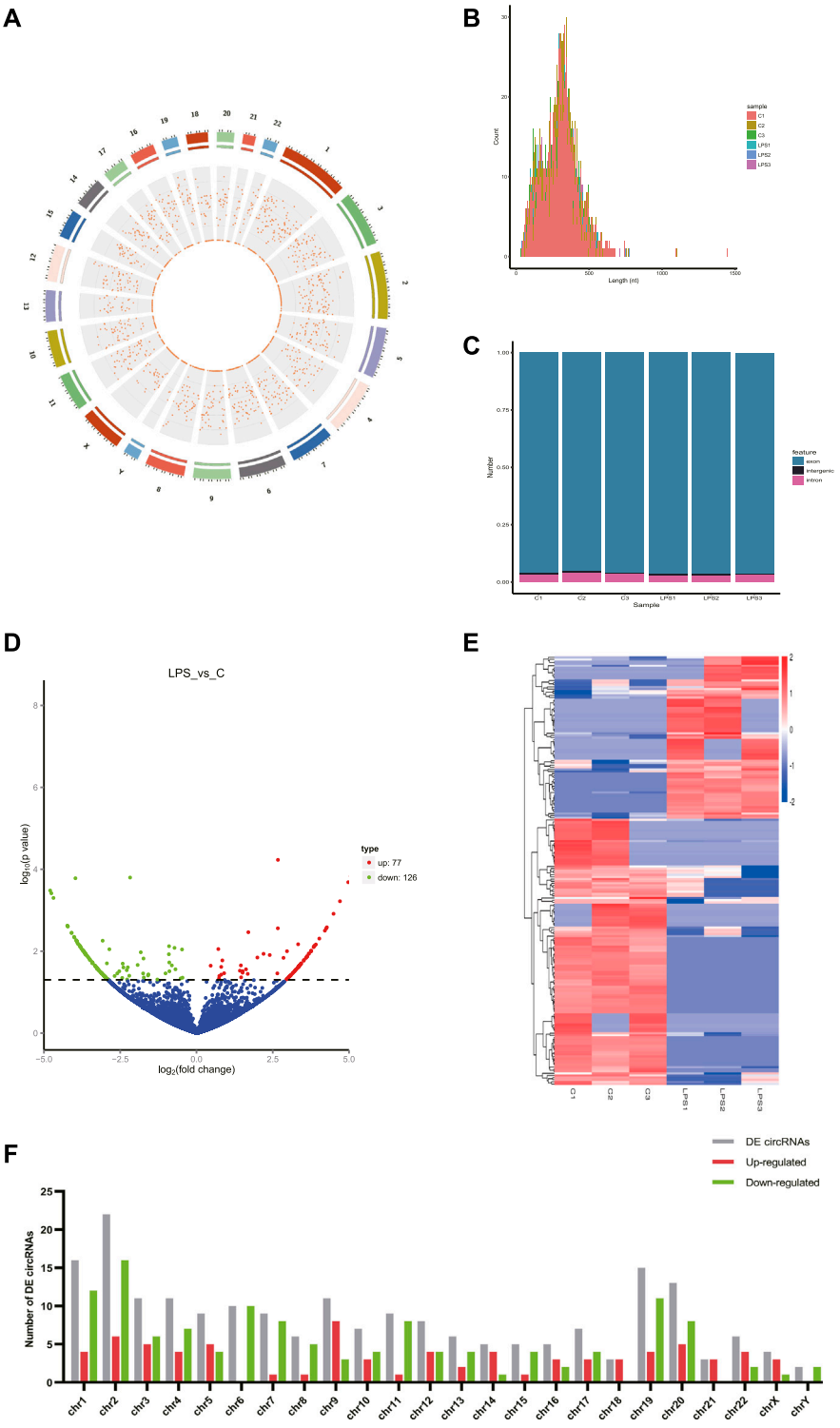


FIGURE 1
CircRNAs in LPS-induced sepsis. **(A)** The distribution of all cricRNAs in sepsis. The outermost layer was a chromosome map of the human genome. The red dots in the inner circles represented the circRNAs. **(B)** The length of the circRNA of all samples and the source of the circRNAs were represented. **(C)** The source of all circRNAs. **(D)** Representative volcano plot showing the up- and down-regulated circRNAs in the cells treated with or without LPS. Red dots represent up-regulated transcripts, green dots represent down-regulated transcripts, and blue dots represent genes without a significant change. **(E)** Cluster analysis of all differentially expressed transcripts. Up- and down-regulated circRNAs are colored red and blue, respectively. **(F)** Distribution of DE circRNAs on chromosomes. **(C)** Control groups. LPS: LPS groups.

TABLE 1 The 20 DE circRNAs of the RNA-seq dataset.

CircRNA ID	Chromosome	Feature	Gene symbol	Regulation	log2FoldChange	*p*-value
hsa_circ_0002360	21	exon	RUNX1	Up	5.0417	0.0001
hsa_circ_0069338	4	exon	SEPSECS	Up	4.9572	0.0002
hsa_circ_0007816	9	exon	UHRF2	Up	4.6863	0.0006
hsa_circ_0074158	5	exon	CTNNA1	Up	4.4764	0.0012
hsa_circ_TRAK1_8343	3	exon	TRAK1	Up	4.2562	0.0026
hsa_circ_RSBN1L_11059	7	exon	RSBN1L	Up	4.2336	0.0028
hsa_circ_0051427	19	exon	RELB	Up	4.1916	0.0032
hsa_circ_0004188	18	intron	PTPRM	Up	3.6088	0.0135
hsa_circ_ACTN1_2453	14	exon	ACTN1	Up	3.602	0.0137
hsa_circ_0063534	22	exon	RANGAP1	Up	1.6825	0.0035
hsa_circ_0001860	9	exon	ZCCHC7	Down	−6.7445	0.0000
hsa_circ_0004539	5	exon	LHFPL2	Down	−5.6548	0.0000
hsa_circ_0008585	2	exon	BIRC6	Down	−5.3167	0.0000
hsa_circ_0004161	1	exon	DPYD	Down	−4.8136	0.0003
hsa_circ_0002127	11	exon	HIPK3	Down	−4.7746	0.0004
hsa_circ_MAD1L1_10741	7	exon	AC069288.1; MAD1L1	Down	−4.7053	0.0005
hsa_circ_0005395	2	exon	NBEAL1	Down	−4.2523	0.0024
hsa_circ_0001491	5	exon	IPO11	Down	−4.2405	0.0025
hsa_circ_XPO7_11370	8	exon	XPO7	Down	−4.2344	0.0025
hsa_circ_0005564	8	exon	FGFR1	Down	−2.413	0.0250

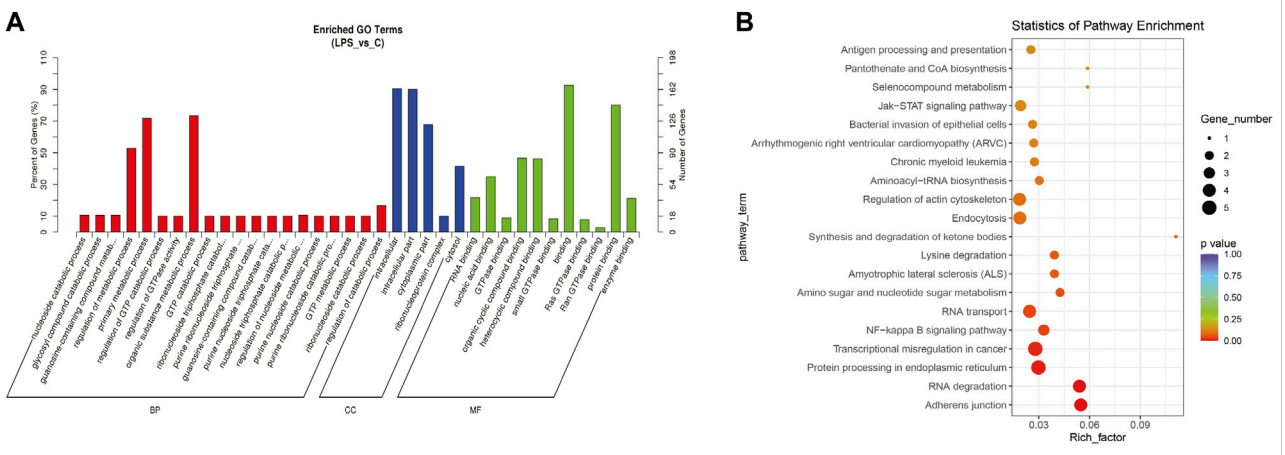


FIGURE 2
GO and KEGG pathway enrichment analyses for DE circRNAs. **(A)** The top significantly enriched GO terms for DE circRNAs. The horizontal axis represents the name of the GO term, and the vertical axis represents the number of genes. **(B)** The top 20 significantly enriched KEGG pathways. The size of the dots represents the number of DE circRNAs enriched in the pathway, and the color of the dots represents the *p*-value. The horizontal axis shows the rich factor (the degree of enrichment of DE circRNAs), and the vertical axis shows the pathway names.

the circRNAs of all samples and the source of these circRNAs were shown in Figures 1B,C, respectively. According to Figure 1B, we found that the length of most circRNAs in HUVECs is less than 500 nt. Figure 1C revealed that circRNAs can be derived from exons, introns, and intergenic, but most of the circRNAs in HUVECs are derived from the exon.

According to the filtering conditions of $|\log_2(\text{FoldChange})| > 1$ and the p-value <0.05, 203 circRNAs were identified as significantly DE circRNAs (77 circRNAs were up-regulated and 126 were down-regulated). Based on the top dysregulated circRNAs, 20 circRNAs were selected from those identified DE circRNAs (Table 1). To verify the

TABLE 2 CircRNAs related to the adherens junction pathway.

CircRNA_ID	Chromosome	Feature	Gene symbol	Regulation	log2FoldChange	*p*-value	DE circRNA
hsa_circ_0074158	5	exon	CTNNA1	Up	4.4764	0.0012	Yes
hsa_circ_0004188	18	intron	PTPRM	Up	3.6088	0.0135	Yes
hsa_circ_ACTN1_2453	14	exon	ACTN1	Up	3.602	0.0137	Yes
hsa_circ_0002872	18	exon	PTPRM	Up	1.8256	0.2126	No
hsa_circ_0074171	5	exon	CTNNA1	Up	1.5502	0.2802	No
hsa_circ_0046813	18	exon	PTPRM	Up	1.0413	0.4253	No
hsa_circ_0002913	14	exon	ACTN1	Up	0.85031	0.4615	No
hsa_circ_0032321	14	exon	ACTN1	Up	0.43822	0.7609	No
hsa_circ_0008196	5	exon	CTNNA1	Up	0.28868	0.8438	No
hsa_circ_0007440	5	exon	CTNNA1	Up	0.031718	0.9676	No
hsa_circ_0006114	18	exon	PTPRM	Down	−0.22938	0.7611	No
hsa_circ_0008016	8	exon	FGFR1	Down	−0.97928	0.1099	No
hsa_circ_PTPRM_4378	18	exon	PTPRM	Down	−1.4552	0.3111	No
hsa_circ_ACTN1_2450	14	exon	ACTN1	Down	−1.4768	0.3018	No
hsa_circ_0007644	14	exon	ACTN1	Down	−1.7358	0.2436	No
hsa_circ_0005564	8	exon	FGFR1	Down	−2.413	0.0250	Yes

results of the RNA-Seq analysis, the expression of the selected 20 DE circRNAs was verified using RT-qPCR. The significant differences in DE circRNAs between LPS and control groups were shown with a volcano plot (Figure 1D) and cluster analysis (Figure 1E). As shown in Figure 1F, sepsis-associated circRNAs were distributed on each chromosome, while more DE circRNAs were located on chromosome 2.

3.2 Enrichment analysis of the GO and KEGG pathways for differentially expressed circRNAs

The top significantly enriched GO terms of DE circRNAs were shown in Figure 2A, the results showed that the primary metabolic process and the organic substance metabolic process were more enriched for BP, and the intracellular and intracellular part had more enriched for CC, while binding was more enriched for MF. According to the analysis of KEGG pathway enrichment (Figure 2B), DE circRNAs were mainly enriched at the adherens junction, suggesting the significant associations of DE circRNAs with the endothelial barrier in LPS-induced sepsis, RNA degradation, and protein processing in the endoplasmic reticulum, *etc.* The adherens junction pathway was identified for our further study (Supplementary Figure 1S). The circRNAs associated with the adherens junction pathway were listed in Table 2, of which hsa_circ_0074158, hsa_circ_0004188, hsa_circ_0005564, and hsa_circ_ACTN1_2453 belonged to 203 DE circRNAs.

3.3 Analysis of the ceRNA regulatory networks

We constructed the ceRNA regulatory networks analysis to elucidate the relationship in sepsis using circRNA-miRNA-mRNA. 138 target miRNAs for 20 DE circRNAs and 21307 target mRNAs for the miRNAs were predicted. We selected top target mRNAs of the 20 DE circRNAs to conduct a circRNA-miRNA-mRNA network (Figure 3A). The network showed that CDH5 (VE-cadherin) is the target of hsa-miR-515-5p and hsa-miR-515-5p is the target of hsa_circ_0074158. To select hsa_circ_0074158 for further study, we also conducted a circRNA-miRNA-mRNA network for hsa_circ_0074158 (Figure 3B).

3.4 Validation of the RNA-seq results and demographic information of hsa_circ_0074158

RT-qPCR was performed to verify the 20 DE circRNAs identified by RNA-seq analysis above. Combined with $\log_2$FoldChange in Table 1, the results showed that only the expression patterns of hsa_circ_0074158, hsa_circ_RSBN1L_11059, hsa_circ_0004188, and hsa_circ_0005564 were highly consistent with the RNA-seq results (Figure 4A). The hsa_circ_0074158 was selected for our further study according to the expression level and KEGG pathway enrichment analysis. To verify the splice site of hsa_circ_0074158, the RT-qPCR amplification and Sanger sequencing were performed and the results confirmed the back splice junction site (Figure 4B). The linear RNAs were

removed using an RNase R treatment experiment (Epicentre Technologies, United States) (Figure 4C). The RT-qPCR results of 92 patients showed that the expression pattern of hsa_circ_0074158 was highly consistent with the results of the RNA-seq (Figure 4D). We also collected demographic information on sepsis patients. Based on the median expression of hsa_circ_0074158, we classified the 44 patients into a low expression group ($n = 22$) and a high expression group ($n = 22$). We found that the expression level of hsa_circ_0074158 was significantly correlated with chronic

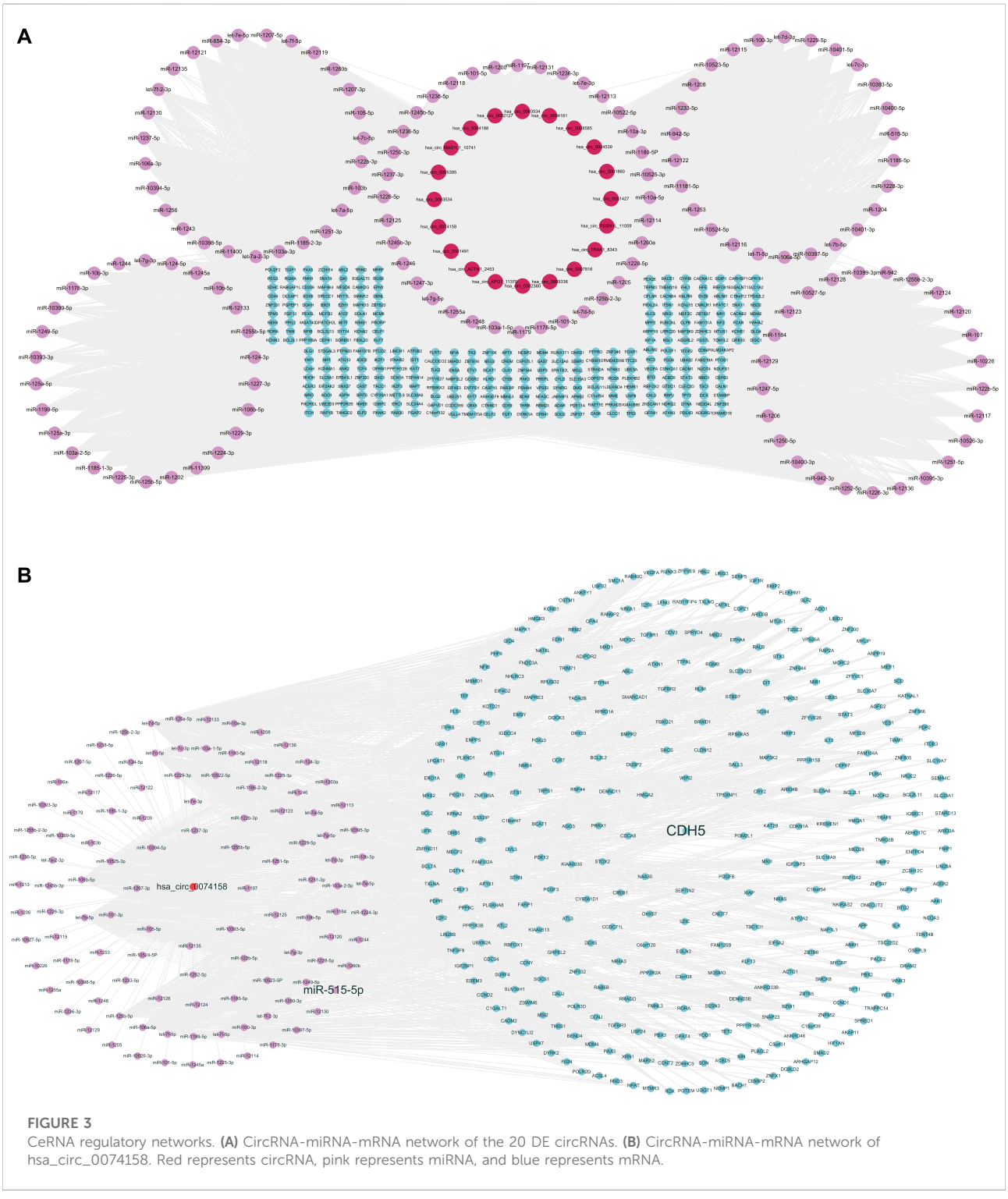


FIGURE 3
CeRNA regulatory networks. **(A)** CircRNA-miRNA-mRNA network of the 20 DE circRNAs. **(B)** CircRNA-miRNA-mRNA network of hsa_circ_0074158. Red represents circRNA, pink represents miRNA, and blue represents mRNA.

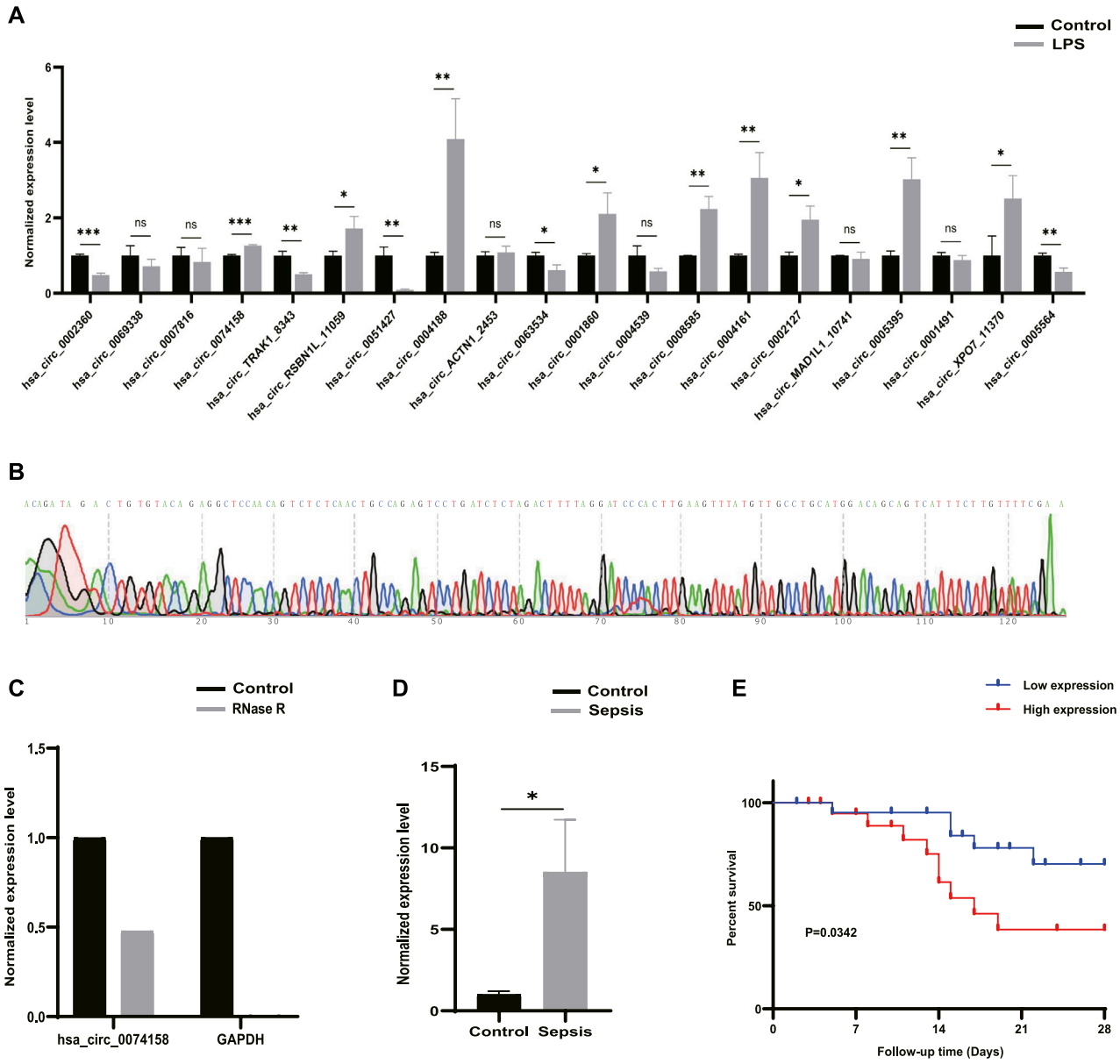


FIGURE 4
Candidate circRNAs. **(A)** Validation of 20 DE circRNAs expression by RT-qPCR. **(B)** Sanger sequencing confirmed the back splice junction site of hsa_circ_0074158. Green stands for A, red stands for T, black stands for G, and blue stands for C .**(C)** RNase R treatment experiments. **(D)** Validation of the hsa_circ_0074158 expression in the whole blood of 44 patients with sepsis and 48 healthy individuals. **(E)** Kaplan-Meier survival curve of patients with sepsis. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$. Control: Control groups. LPS: LPS groups. RNase R: RNase R treatment groups. Sepsis: Sepsis groups.

comorbidities (Table 3). Additionally, survival analyses were also performed based on the expression levels of hsa_circ_0074158. The Kaplan-Meier method with the log-rank test was used to assess the survival rate of sepsis patients (Figure 4E). And the analysis showed that the patients of the hsa_circ_0074158 high expression group exhibited poor overall survival ($p = 0.0342$) compared to those of the hsa_circ_0074158 low expression group.

3.5 Role of hsa_circ_0074158 in sepsis-induced endothelial barrier disruption

To illustrate the effects of hsa_circ_0074158 on the endothelial barrier function in sepsis, siRNA was used to confirm the role. Previous research has shown that LPS can cause VE-cadherin disruption and lead to adherens junction disruption (Chan et al., 2019). In this study, immunofluorescence

TABLE 3 Demographic information for sepsis patients according to hsa_circ_0074158 expression.

Parameters	Low expression	High expression	*p*-Value
Gender			0.540
Male	12	14	
Female	10	8	
Age (years)			0.093
≥ 60	17	21	
< 60	5	1	
Infection souce			0.793
Lung	13	10	
Urinary tract	5	5	
Skin or soft tissues	1	1	
Other	3	6	
Comorbidities			0.015
Yes	13	20	
No	9	2	

staining demonstrated that hsa_circ_0074158 knockdown can significantly enhance the expression of VE-cadherin, which was down-regulated in LPS-induced sepsis (Figure 5A). Experiments also showed that knockdown of hsa_circ_0074158 could increase TEER (Figure 5B) and decrease the extravasation of FITC-dextran (Figure 5C). In summary, these results fully confirmed that si-circ_0074158 reversed LPS-induced suppression of VE-cadherin and endothelial hyperpermeability, mimicking barrier protection.

4 Discussion

With the development of bioinformatic techniques and approaches, the associations of circRNA with diseases, such as cancer, cardiovascular disease, nervous system disease, immune-related disorders, and infections, were identified and verified (Chen et al., 2018; Wang et al., 2018; He et al., 2019; Hosaka et al., 2019; Wang et al., 2019). The differential expression of circRNA plays an important role in the diagnosis, treatment, and prognosis of the disease. Sepsis is characterized by complex pathogenesis, atypical clinical manifestations, and a difficult diagnosis. Despite the significant advances that have been made in treatment technologies, sepsis remains a serious clinical syndrome with substantial morbidity and mortality. To further clarify the pathogenesis and provide a novel approach for the diagnosis and treatment of sepsis, we explored the transcriptome characteristics of LPS-induced sepsis by RNA-seq.

Herein, 203 significantly DE circRNAs, including 77 up-regulated and 126 down-regulated, were identified. Then 20 DE circRNAs were selected for RT-qPCR validation. According to the results from the RNA-seq analysis and filtering conditions, only the expression of hsa_circ_0074158, hsa_circ_RSBN1L_11059, hsa_circ_0004188, and hsa_circ_0005564 were highly consistent with the results of the RNA-seq, which indicated that these circRNAs were potential biomarkers and drug targets for sepsis treatment. However, more research is needed to validate the results.

We performed GO and KEGG pathway enrichment analyses to explore the potential pathophysiological mechanism in sepsis. The results of the GO enrichment analysis showed that the most significantly enriched GO term in the BF was organic substance metabolic process, the CC was intracellular, and in MF, binding. The GO enrichment analysis indicated that circRNA functions were highly associated with the metabolic process and GTPase, which have been shown to have a relationship with sepsis and endothelial barrier function in previous studies, respectively (Mira et al., 2017; Aslam et al., 2019; Sanchez and Raja, 2022). The enrichment analysis of the KEGG pathway revealed that circRNAs were mainly involved in adherens junction, NF-kappa B signaling pathway, endocytosis, regulation of actin cytoskeleton, bacterial invasion of epithelial cells, antigen processing and presentation, and which also predicted that circRNAs might be related to endothelial barrier function in sepsis (Garcia et al., 2018; McRae et al., 2018; Radeva and Waschke, 2018; Leonard et al., 2019; Strauss and Gourdie, 2020).

We also constructed ceRNA regulatory networks to reveal the regulatory mechanism in sepsis. According to the ceRNA mechanism, there are 138 target miRNAs for 20 DE circRNAs and 21307 target mRNAs for the miRNAs were identified. In particular, CDH5 (VE-cadherin) is the target gene of hsa-miR-515-5p and hsa-miR-515-5p is the target of hsa_circ_0074158. The KEGG pathway analysis showed that hsa_circ_0074158 was related to the adherens junction. As we know, adherens junctions are specialized forms of adhesive contacts based on VE-cadherin, and adherens junctions are mainly composed of VE-cadherin (Dejana et al., 2008). Tight junctions, adherens junctions, and gaps are more closely correlated with the function of the endothelial barrier. The endothelial barrier function plays a key role in the development of sepsis. Therefore, we hypothesized that hsa_circ_0074158 regulated VE-cadherin and endothelial barrier function in sepsis through hsa-miR-515-5p. We also evaluated the adherens function of the endothelial barrier using VE-cadherin expression.

The hsa_circ_0074158 in the KEGG pathway of the adherens junction was selected for further study. High consistency of the *in vivo* RT-qPCR results with the RNA-seq analysis was observed. We used a combination of RNase R treatment and Sanger sequencing to confirm the cyclization feature of hsa_circ_0074158. Immunofluorescence assay and endothelial permeability demonstrated that LPS-induced sepsis can disrupt VE-cadherin expression and enhance endothelial permeability, while knockdown of hsa_circ_0074158 can suppress the sepsis-

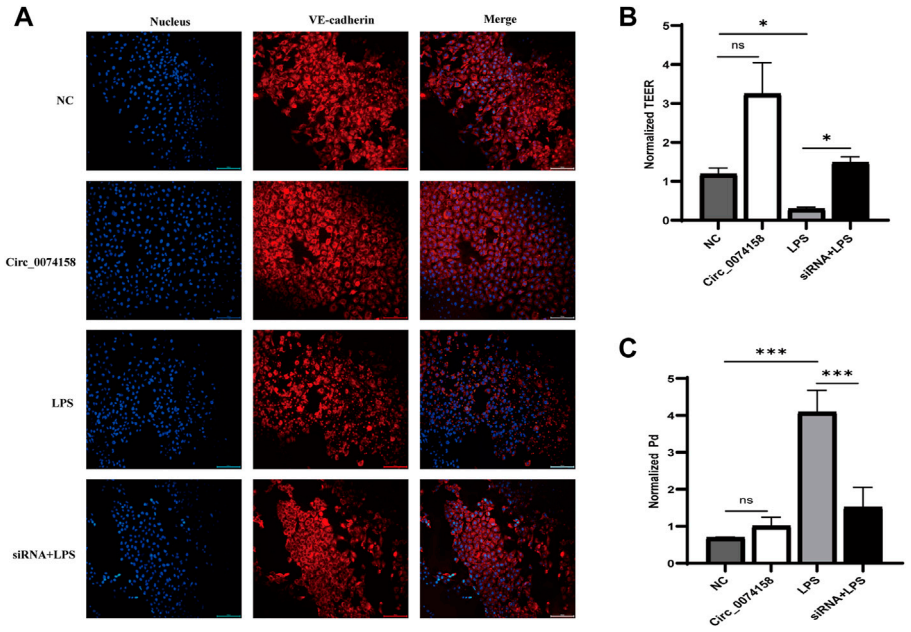


FIGURE 5
Hsa_circ_0074158 regulates the endothelial barrier function in sepsis. **(A)** The silencing of hsa_circ_0074158 enhanced the expression of VE-cadherin, which was down-regulated in LPS-induced sepsis. Bar = 100 μm. **(B)** The endothelial permeability was measured by TEER assay and showed. **(C)** The endothelial permeability was determined using the FITC-dextran assay and showed. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$. NC: Negative control groups. LPS: LPS groups. Circ_0074158: Circ_0074158 groups. siRNA + LPS: si-circ_0074158 + LPS groups.

caused aberrant regulation. Hence, we hypothesized that hsa_circ_0074158 was potentially involved in endothelial barrier dysfunction. However, these results are preliminary and further confirmation with an expanded study is needed.

Recently, an increasing number of studies have investigated the function of circRNAs. It is well known that circRNAs can act as miRNA sponges, thus regulating the expression of miRNAs (Thomson and Dinger, 2016; Panda, 2018). Many studies also demonstrated that circRNAs might be used as potential biomarkers and therapeutic targets for the treatment and therapy of sepsis. A clinical study demonstrated that circ-PRKCI was an independent factor to predict the mortality risk of sepsis patients (Wei and Yu, 2020). Another clinical study that included 25 patients with sepsis observed the differential expression of hsa_circRNA_104484 and hsa_circRNA_104670 in serum exosomes, indicating the potential of hsa_circRNA_104484 and hsa_circRNA_104670 as a biomarker and potential therapeutic target (Wei and Yu, 2020). A recent review analyzed the biological characteristics of circRNAs in sepsis based on the Medline database (Wei et al., 2022). In our study, survival analyses showed that people with high hsa_circ_0074158 expression have higher mortality, suggesting that hsa_circ_0074158 could be used as a biomarker. To date, more and more attention has been paid to the biomarkers of sepsis, which can help to diagnose, evaluate the disease, and judge the prognosis. However, the ideal biomarker for sepsis has not yet been determined. To prevent, identify, diagnose, and treat sepsis early, the combination of known biomarkers for sepsis may be more effective. If there are biomarkers that can be used to predict sepsis early, mortality and the use of unnecessary antibiotics can be reduced. The identification of the biomarkers for the prediction of sepsis has become one of the hotspots of research. In the future, when searching for sepsis biomarkers, we should pay more attention to early prevention of sepsis, which can reduce the morbidity and mortality of sepsis more efficiently. Furthermore, our research also showed that circRNA was potentially associated with endothelial barrier function in sepsis, which may provide more possibilities for the prevention of sepsis in advance.

5 Conclusion

In our present study, we clarified the differential expression and potential functions of circRNAs in LPS-induced sepsis. CirculaRNAs, including hsa_circ_0074158, hsa_circ_RSBN1L_11059, hsa_circ_0004188, and hsa_circ_0005564, could serve as potential biomarkers and therapeutic targets for sepsis treatment and therapy. Meanwhile, we observed significant associations of hsa_circ_0074158 with the endothelial barrier function in sepsis. The results also showed that the knockdown of hsa_circ_0074158 could significantly improved VE-cadherin expression, reduced endothelial permeability, and subsequently

protect the adherens junction in sepsis. Based on ceRNA regulatory network, we hypothesized that hsa_circ_0074158 regulated the endothelial barrier disruption in sepsis through hsa-miR-515-5p. Collectively, our data demonstrated that dysregulation of circRNAs was significantly associated with sepsis and circRNAs played a potential role in sepsis, suggesting the role of circRNA as a promising biomarker and the significant clinical values for sepsis treatment and therapy.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: https://www.ncbi.nlm.nih.gov/geo/, GSE204776.

Ethics statement

The studies involving human participants were reviewed and approved by Beijing Tsinghua Changgung Hospital. The patients/participants provided their written informed consent to participate in this study.

Author contributions

Conception: HL and ZW. Design, analysis, interpretation, and writing of the original manuscript: HL. Reviewing and supervision: HL, YC, YS, ZG, XW, ZyW, ZwW, and ZW. Editing: ZW. Supervision: HL, YC, YS, and ZW. All authors have read and agreed to the published version of the manuscript. All authors have agreed to be personally accountable for the author contributions and answer any questions related to final manuscript.

Conflict of interest

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

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.1002344/full#supplementary-material.

References

Abdelmohsen, K., Panda, A. C., Munk, R., Grammatikakis, I., Dudekula, D. B., De, S., et al. (2017). Identification of HuR target circular RNAs uncovers suppression of PABPN1 translation by CircPABPN1. *RNA Biol.* 14 (3), 361–369. doi:10.1080/15476286.2017.1279788

Agarwal, V., Bell, G. W., Nam, J. W., et al. (2015). *Predicting effective microRNA target sites in mammalian mRNAs*[J]. Cambridgeshire, United Kingdom: eLife Sciences Publications Ltd. 4.

Ashwal-Fluss, R., Meyer, M., Pamudurti, N. R., Ivanov, A., Bartok, O., Hanan, M., et al. (2014). circRNA biogenesis competes with pre-mRNA splicing. *Mol. Cell* 56 (1), 55–66. doi:10.1016/j.molcel.2014.08.019

Aslam, M., Troidl, C., Tanislav, C., Rohrbach, S., Gunduz, D., and Hamm, C. W. (2019). Inhibition of protein prenylation of GTPases alters endothelial barrier function. *Int. J. Mol. Sci* 21 (1), E2. doi:10.3390/ijms21010002

Beltran-Garcia, J., Osca-Verdegal, R., Nacher-Sendra, E., Pallardo, F. V., and Garcia-Gimenez, J. L. (2020). Circular RNAs in sepsis: Biogenesis, function, and clinical significance. *Cells* 9 (6), E1544. doi:10.3390/cells9061544

Brasch, J., Harrison, O. J., Ahlsen, G., Carnally, S. M., Henderson, R. M., Honig, B., et al. (2011). Structure and binding mechanism of vascular endothelial cadherin: A divergent classical cadherin. *J. Mol. Biol.* 408 (1), 57–73. doi:10.1016/j.jmb.2011.01.031

Chan, Y. H., Harith, H. H., Israf, D. A., and Tham, C. L. (2019). Differential regulation of LPS-mediated VE-cadherin disruption in human endothelial cells and the underlying signaling pathways: A mini review. *Front. Cell Dev. Biol.* 7, 280. doi:10.3389/fcell.2019.00280

Chen, B., Wei, W., Huang, X., Xie, X., Kong, Y., Dai, D., et al. (2018). circEPSTI1 as a prognostic marker and mediator of triple-negative breast cancer progression. *Theranostics* 8 (14), 4003–4015. doi:10.7150/thno.24106

Chen, L., Huang, C., and Shan, G. (2021). Circular RNAs in physiology and non-immunological diseases. *Trends biochem. Sci.* 47, 250–264. doi:10.1016/j.tibs.2021.11.004

Chen, L. L. (2016). The biogenesis and emerging roles of circular RNAs. *Nat. Rev. Mol. Cell Biol.* 17 (4), 205–211. doi:10.1038/nrm.2015.32

Chen, L. (2020). The expanding regulatory mechanisms and cellular functions of circular RNAs. *Nat. Rev. Mol. Cell Biol.* 21 (8), 475–490. doi:10.1038/s41580-020-0243-y

Cocquerelle, C., Daubersies, P., Majerus, M. A., Kerckaert, J. P., and Bailleul, B. (1992). Splicing with inverted order of exons occurs proximal to large introns. *EMBO J.* 11 (3), 1095–1098. doi:10.1002/j.1460-2075.1992.tb05148.x

Cocquerelle, C., Mascrez, B., Hetuin, D., and Bailleul, B. (1993). Mis-splicing yields circular RNA molecules. *FASEB J.* 7 (1), 155–160. doi:10.1096/fasebj.7.1.7678559

Consortium, T. G. O. (2019). The gene Ontology resource: 20 years and still GOing strong. *Nucleic Acids Res.* 47 (D1), D330–D338. doi:10.1093/nar/gky1055

Dejana, E., Orsenigo, F., and Lampugnani, M. G. (2008). The role of adherens junctions and VE-cadherin in the control of vascular permeability. *J. Cell Sci.* 121 (13), 2115–2122. doi:10.1242/jcs.017897

Fleischmann, C., Scherag, A., Adhikari, N. K., Hartog, C. S., Tsaganos, T., Schlattmann, P., et al. (2016). Assessment of global incidence and mortality of

hospital-treated sepsis. Current estimates and limitations. *Am. J. Respir. Crit. Care Med.* 193 (3), 259–272. doi:10.1164/rccm.201504-0781OC

Garcia, M. A., Nelson, W. J., and Chavez, N. (2018). Cell-cell junctions organize structural and signaling networks. *Cold Spring Harb. Perspect. Biol.* 10 (4), a029181. doi:10.1101/cshperspect.a029181

Godo, S., and Shimokawa, H. (2017). Endothelial functions. *Arterioscler. Thromb. Vasc. Biol.* 37 (9), e108–e114. doi:10.1161/ATVBAHA.117.309813

Godo, S., Takahashi, J., Yasuda, S., and Shimokawa, H. (2021). Endothelium in coronary macrovascular and microvascular diseases. *J. Cardiovasc. Pharmacol.* 78 (6), S19–S29. doi:10.1097/FJC.0000000000001089

He, J., Ren, M., Li, H., Yang, L., Wang, X., and Yang, Q. (2019). Exosomal circular RNA as a biomarker platform for the early diagnosis of immune-mediated demyelinating disease. *Front. Genet.* 10, 860. doi:10.3389/fgene.2019.00860

Hosaka, T., Yamashita, T., Tamaoka, A., and Kwak, S. (2019). Extracellular RNAs as biomarkers of sporadic amyotrophic lateral sclerosis and other neurodegenerative diseases. *Int. J. Mol. Sci.* 20 (13), E3148. doi:10.3390/ijms20133148

Hsu, M. T., and Coca-Prados, M. (1979). Electron microscopic evidence for the circular form of RNA in the cytoplasm of eukaryotic cells. *Nature* 280 (5720), 339–340. doi:10.1038/280339a0

Huang, Y., Zhang, C., Xiong, J., and Ren, H. (2021). Emerging important roles of circRNAs in human cancer and other diseases. *Genes Dis.* 8 (4), 412–423. doi:10.1016/j.gendis.2020.07.012

Jiang, B., Zhang, J., Sun, X., Yang, C., Cheng, G., Xu, M., et al. (2022). Circulating exosomal hsa_circRNA_0039480 is highly expressed in gestational diabetes mellitus and may be served as a biomarker for early diagnosis of GDM. *J. Transl. Med.* 20 (1), 5. doi:10.1186/s12967-021-03195-5

Kanehisa, M., and Goto, S. (2000). Kegg: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* 28 (1), 27–30. doi:10.1093/nar/28.1.27

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

Kristensen, L. S., Andersen, M. S., Stagsted, L., Ebbesen, K. K., Hansen, T. B., and Kjems, J. (2019). The biogenesis, biology and characterization of circular RNAs. *Nat. Rev. Genet.* 20 (11), 675–691. doi:10.1038/s41576-019-0158-7

Lee, W. L., and Slutsky, A. S. (2010). Sepsis and endothelial permeability. *N. Engl. J. Med.* 363 (7), 689–691. doi:10.1056/NEJMcibr1007320

Leonard, A., Millar, M. W., Slavin, S. A., Bijli, K. M., Dionisio Santos, D. A., Dean, D. A., et al. (2019). Critical role of autophagy regulator Beclin1 in endothelial cell inflammation and barrier disruption. *Cell. Signal.* 61, 120–129. doi:10.1016/j.cellsig.2019.04.013

Li, J. H., Liu, S., Zhou, H., Qu, L. H., and Yang, J. H. (2014). starBase v2.0: decoding miRNA-ceRNA, miRNA-ncRNA and protein-RNA interaction networks from large-scale CLIP-Seq data. *Nucleic Acids Res.* 42 D92–D97. doi:10.1093/nar/gkt1248

Love, M. I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15 (12), 550. doi:10.1186/s13059-014-0550-8

Machado, F. R., Cavalcanti, A. B., Bozza, F. A., Ferreira, E. M., Angotti Carrara, F. S., Sousa, J. L., et al. (2017). The epidemiology of sepsis in Brazilian intensive care units (the sepsis PREvalence assessment database, SPREAD): An observational study. *Lancet. Infect. Dis.* 17 (11), 1180–1189. doi:10.1016/S1473-3099(17)30322-5

Martin-Fernandez, M., Tamayo-Velasco, A., Aller, R., Gonzalo-Benito, H., and Martinez-Paz, P. (2021). Endothelial dysfunction and neutrophil degranulation as central events in sepsis physiopathology. *Int. J. Mol. Sci.* 22 (12), 6272. doi:10.3390/ijms22126272

McRae, M., LaFratta, L. M., Nguyen, B. M., Paris, J. J., Hauser, K. F., and Conway, D. E. (2018). Characterization of cell-cell junction changes associated with the formation of a strong endothelial barrier. *Tissue Barriers* 6 (1), e1405774. doi:10.1080/21688370.2017.1405774

Mira, J. C., Gentile, L. F., Mathias, B. J., Efron, P. A., Brakenridge, S. C., Mohr, A. M., et al. (2017). Sepsis pathophysiology, chronic critical illness, and persistent inflammation-immunosuppression and catabolism syndrome. *Crit. Care Med.* 45 (2), 253–262. doi:10.1097/CCM.0000000000002074

Nava, E., and Llorens, S. (2019). The local regulation of vascular function: From an inside-outside to an outside-inside model. *Front. Physiol.* 10, 729. doi:10.3389/fphys.2019.00729

Nigro, J. M., Cho, K. R., Fearon, E. R., Kern, S. E., Ruppert, J. M., Oliner, J. D., et al. (1991). Scrambled exons. *Cell* 64 (3), 607–613. doi:10.1016/0092-8674(91)90244-s

Panda, A. C. (2018). Circular RNAs act as miRNA sponges. *Adv. Exp. Med. Biol.* 1087, 67–79. doi:10.1007/978-981-13-1426-1_6

Pasman, Z., Been, M. D., and Garcia-Blanco, M. A. (1996). Exon circularization in mammalian nuclear extracts. *RNA* 2 (6), 603–610.

Radeva, M. Y., and Waschke, J. (2018). Mind the gap: Mechanisms regulating the endothelial barrier. *Acta Physiol.* 222 (1), e12860. doi:10.1111/apha.12860

Rajendran, P., Rengarajan, T., Thangavel, J., Nishigaki, Y., Sakthisekaran, D., Sethi, G., et al. (2013). The vascular endothelium and human diseases. *Int. J. Biol. Sci.* 9 (10), 1057–1069. doi:10.7150/ijbs.7502

Sanchez, L. D. N. A., and Raja, A. (2022). , *Metabolism physiology*,National Library of Medicine Bethesda, MD, USA

Sanger, H. L., Klotz, G., Riesner, D., Gross, H. J., and Kleinschmidt, A. K. (1976). Viroids are single-stranded covalently closed circular RNA molecules existing as highly base-paired rod-like structures. *Proc. Natl. Acad. Sci. U. S. A.* 73 (11), 3852–3856. doi:10.1073/pnas.73.11.3852

Shimokawa, H., and Godo, S. (2020). Nitric oxide and endothelium-dependent hyperpolarization mediated by hydrogen peroxide in health and disease. *Basic Clin. Pharmacol. Toxicol.* 127 (2), 92–101. doi:10.1111/bcpt.13377

Singer, M., Deutschman, C. S., Seymour, C. W., Shankar-Hari, M., Annane, D., Bauer, M., et al. (2016). The third international consensus definitions for sepsis and septic shock (Sepsis-3). *JAMA* 315 (8), 801–810. doi:10.1001/jama.2016.0287

Sticht, C., De La Torre, C., Parveen, A., and Gretz, N. (2018). miRWalk: An online resource for prediction of microRNA binding sites. *PLoS One* 13 (10), e0206239. doi:10.1371/journal.pone.0206239

Strauss, R. E., and Gourdie, R. G. (2020). Cx43 and the actin cytoskeleton: Novel roles and implications for cell-cell junction-based barrier function regulation. *Biomolecules* 10 (12), E1656. doi:10.3390/biom10121656

Thomson, D. W., and Dinger, M. E. (2016). Endogenous microRNA sponges: Evidence and controversy. *Nat. Rev. Genet.* 17 (5), 272–283. doi:10.1038/nrg.2016.20

Tian, C., Liu, J., Di, X., Cong, S., Zhao, M., and Wang, K. (2021). Exosomal hsa_circRNA_104484 and hsa_circRNA_104670 may serve as potential novel biomarkers and therapeutic targets for sepsis. *Sci. Rep.* 11 (1), 14141. doi:10.1038/s41598-021-93246-0

Verduci, L., Tarcitano, E., Strano, S., Yarden, Y., and Blandino, G. (2021). CircRNAs: Role in human diseases and potential use as biomarkers. *Cell Death Dis.* 12 (5), 468. doi:10.1038/s41419-021-03743-3

Vincent, J. L., Marshall, J. C., Namendys-Silva, S. A., Francois, B., Martin-Loeches, I., Lipman, J., et al. (2014). Assessment of the worldwide burden of critical illness: The intensive care over nations (ICON) audit. *Lancet. Respir. Med.* 2 (5), 380–386. doi:10.1016/S2213-2600(14)70061-X

Wang, Q., Wang, J., Xin, Y., He, Z., Zhou, X., Liu, X., et al. (2022). Hsa_circ_0005729 enhances the accuracy in diagnosing parathyroid carcinoma [J]. *Endocr Connect*,Bethesda, MD, USA 11(2)doi:10.1530/EC-21-0605

Wang, R., Zhang, S., Chen, X., Li, N., Li, J., Jia, R., et al. (2018). CircNT5E acts as a sponge of miR-422a to promote glioblastoma tumorigenesis. *Cancer Res.* 78 (17), 4812–4825. doi:10.1158/0008-5472.CAN-18-0532

Wang, W., Wang, Y., Piao, H., Li, B., Huang, M., Zhu, Z., et al. (2019). Circular RNAs as potential biomarkers and therapeutics for cardiovascular disease. *PeerJ* 7, e6831. doi:10.7717/peerj.6831

Wei, B., and Yu, L. (2020). Circular RNA PRKCI and microRNA-545 relate to sepsis risk, disease severity and 28-day mortality. *Scand. J. Clin. Lab. Invest.* 80 (8), 659–666. doi:10.1080/00365513.2020.1827291

Wei, L., Yang, Y., Wang, W., and Xu, R. (2022). Circular RNAs in the pathogenesis of sepsis and their clinical implications: A narrative review. *Ann. Acad. Med. Singap.* 51 (4), 221–227. doi:10.47102/annals-acadmedsg.2021405

Wu, J., Deng, Z., Sun, M., Zhang, W., Yang, Y., Zeng, Z., et al. (2020). Polydatin protects against lipopolysaccharide-induced endothelial barrier disruption via SIRT3 activation. *Lab. Invest.* 100 (4), 643–656. doi:10.1038/s41374-019-0332-8

Zhang, P., Sun, H., Wen, P., Wang, Y., Cui, Y., and Wu, J. (2022). circRNA circMED27 acts as a prognostic factor and mediator to promote lenvatinib resistance of hepatocellular carcinoma. *Mol. Ther. Nucleic Acids* 27, 293–303. doi:10.1016/j.omtn.2021.12.001

Zhong, S., Wang, J., Zhang, Q., Xu, H., and Feng, J. (2018). CircPrimer: A software for annotating circRNAs and determining the specificity of circRNA primers. *BMC Bioinforma.* 19 (1), 292. doi:10.1186/s12859-018-2304-1

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 16 November 2022
DOI 10.3389/fgene.2022.1047004

OPEN ACCESS

EDITED BY
Hua Zhong,
University of Hawaii at Manoa,
United States

REVIEWED BY
Tingxin Zhang,
University of Texas Southwestern Medical Center, United States
Baowen Yuan,
National Cancer Center, Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, China
Hehai Pan,
University of Pennsylvania, United States

*CORRESPONDENCE
Zhengjun Kang,
kzj7153@sina.com
Chaojie Xu,
1045310454@qq.com

†These authors have contributed equally to this work

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

RECEIVED 17 September 2022
ACCEPTED 31 October 2022
PUBLISHED 16 November 2022

CITATION
Yu Y, Huang Y, Li C, Ou S, Xu C and Kang Z (2022), Clinical value of M1 macrophage-related genes identification in bladder urothelial carcinoma and *in vitro* validation. *Front. Genet.* 13:1047004. doi: 10.3389/fgene.2022.1047004

Clinical value of M1 macrophage-related genes identification in bladder urothelial carcinoma and *in vitro* validation

Yang Yu[1†], Yuexi Huang[2†], Chen Li[3†], Santao Ou[2], Chaojie Xu[1*] and Zhengjun Kang[1*]

[1]The Fifth Affiliated Hospital of Zhengzhou University, Zhengzhou University, Zhengzhou, China, [2]Department of Nephrology, The Affiliated Hospital of Southwest Medical University, Luzhou, China, [3]Department of Biology, Chemistry, Pharmacy, Free University of Berlin, Berlin, Germany

Background: Tumor microenvironment (TME) takes a non-negligible role in the progression and metastasis of bladder urothelial carcinoma (BLCA) and tumor development could be inhibited by macrophage M1 in TME. The role of macrophage M1-related genes in BLCA adjuvant therapy has not been studied well.

Methods: CIBERSOR algorithm was applied for identification tumor-infiltrating immune cells (TICs) subtypes of subjects from The Cancer Genome Atlas (TCGA) and Gene Expression Omnibus (GEO) data sets. We identified potential modules of M1 macrophages by weighted gene co-expression network analysis (WGCNA). Nomogram was determined by one-way Cox regression and lasso regression analysis for M1 macrophage genes. The data from GEO are taken to verify the models externally. Kaplan-Meier and receiver operating characteristic (ROC) curves validated prognostic value of M1 macrophage genes. Finally, we divided patients into the low-risk group (LRG) and the high-risk group (HRG) based on the median risk score (RS), and the predictive value of RS in patients with BLCA immunotherapy and chemotherapy was investigated. Bladder cancer (T24, 5637, and BIU-87) and bladder uroepithelial cell line (SV-HUC-1) were used for *in vitro* validation. Reverse transcription-quantitative polymerase chain reaction (RT-qPCR) was employed to validate the associated genes mRNA level.

Results: 111 macrophage M1-related genes were identified using WGCNA. RS model containing three prognostically significant M1 macrophage-associated genes (FBXO6, OAS1, and TMEM229B) was formed by multiple Cox analysis, and a polygenic risk model and a comprehensive prognostic line plot was developed. The calibration curve clarified RS was a good predictor of prognosis. Patients in the LRG were more suitable for programmed cell death protein 1 (PD1) and cytotoxic T lymphocyte associate protein-4 (CTLA4) combination immunotherapy. Finally, chemotherapeutic drug models showed patients in the LRG were more sensitive to gemcitabine and

mitomycin. RT-qPCR result elucidated the upregulation of FBXO6, TMEM229B, and downregulation of OAS1 in BLCA cell lines.

Conclusion: A predictive model based on M1 macrophage-related genes can help guide us in the treatment of BLCA.

KEYWORDS

M1 macrophage, immunotherapy, chemotherapy, resistance, WGCNA

Introduction

Bladder urothelial carcinoma (BLCA) is a common malignant tumor in the bladder system, which was listed as one of the nine most common cancers by the World Health Organization (WHO). Based on histopathology, BLCA can be classified into two categories, muscle-invasive bladder cancer (MIBC) takes most of the BLCA, which is prone to recurrence, another type that accounts for 30% is non-muscle-invasive bladder cancer (NMIBC) with rapid metastasize and low survival rate (Xu et al., 2022a; Kubrak et al., 2022). Therefore, we urgently need a prognostic risk model to provide guidance for the treatment of BLCA.

Tumor microenvironment (TME) is constructed by interacting closely with the extracellular matrix (ECM) and stromal cells to evade detection and eradication by host immune surveillance (Ge and Ding, 2020). Mounting studies have indicated the composition of the TME correlates strongly with immune response and chemotherapy, and regulations in various immune cells from the TME effect the clinical outcome of malignancies (Erbani et al., 2022; Friedrich et al., 2022; Pang et al., 2022; Wang et al., 2022). Tumor-associated macrophages (TAM) drive tumor progression, metastasis, and therapeutic resistance (Larroquette et al., 2022)but heterogeneity is a considerable factor in macrophages (Wu et al., 2020). M1 macrophages inhibit solid tumorigenesis, progression, metastasis, and drug resistance, while M2 macrophages act the opposite. M1 macrophages phagocytose tumor cells *via* cell-mediated cytotoxicity. M1 macrophages exert anti-tumor effects through the production of pro-inflammatory factors (TNF-α, I L-1β, and iNOS), chemokines (CXCL10, CXCL11, and CCL2), antigen-presenting molecules (MHCII), co-stimulatory molecules (CD86, CD80), and antigen-processing peptidases (Jia et al., 2021). In addition, Zeng et al. found M1 infiltration to be a reliable biomarker for predicting prognosis of tumor patients and surpassed biomarkers such as CD8 T cells (Zeng et al., 2020). However, the biological role of M1 macrophages in the prognosis of BLCA has not been well studied (Wu et al., 2022).

Currently, transurethral resection of bladder tumors remains the first-line treatment for patients and is combined with chemotherapeutic agents, but with poor efficacy (Kubrak et al., 2022). Immune checkpoint inhibitors have been proven as a proper option for surgical treatments, however, they merely work in a few types of tumors (Garris et al., 2021; Groeneveld et al., 2021; Barone et al., 2022). Programmed death ligand 1 (PD-L1) expression is the main predictive biomarker for immune checkpoint inhibitor (ICI). Furthermore, tumor mutational burden (TMB) reflects an overall neoantigen load and has the potential to be a predictive biomarker for ICI (Cao et al., 2021a). In summary, we planned to evaluate the sensitivity of different adjuvant treatment modalities for bladder cancer treatment.

Therefore, this study tried to establish an M1 macrophage-based risk score (RS) to comprehensively investigate the sensitivity of the tumor to clinical treatments.

Materials and methods

Information collection

We downloaded The Cancer Genome Atlas-Bladder Cancer (TCGA-BLCA) (samples of bladder cancer, n = 414; normal tissue, n = 19) and GSE31684 (n = 93) genetic expression data to obtain sequencing profiles. Clinical data for TCGA-BLCA can also be found in The Cancer Genome Atlas (TCGA) database. A total of 407 tumor samples were uploaded after the missing clinical signs samples were extracted from TCGA database. To further analyze the variations in copy number of BLCA patients, we also obtained somatic mutation data from the TCGA database for BLCA patients to further analyze copy number variation (CNV).

Distribution of TICs

We obtained the abundance of 22 infiltrating immune cells (TICs) in the TME using the CIBERSORT algorithm on the TCGA-BLCA data of the samples (Zhong et al., 2021).

Weighted gene co-expression network analysis

The purpose of weighted gene co-expression network analysis (WGCNA) was to find co-expressed gene modules (Cao et al., 2021b). The expression of 16,394 genes in the TCGA-BLCA queue will be used as data, and the result of

CIBERSORT will be used as an explanation. To construct an approximate matrix, the soft optimal power (β) value is selected by using the pickSoftThreshold function, and the power level with a soft threshold of 1–20 is selected as a candidate. To obtain different gene modules, the tom matrix obtained from genetic expression is used to regroup genes, set the minimum number of modular genes, and cut the outcomes of gene synthesis. By using the "dynamic tree cutting" algorithm, similar genes are introduced into the same module at the same time. Our study targeted "M1 macrophages", so the modules of most significant relevance to M1 macrophages were selected.

Construction of prognostic signature in macrophage M1

Using the most important gene module to study the prognosis of M1 macrophage related genes, a prognostic risk signature was constructed in bladder cancer. In the first step, univariate regression analysis was conducted to identify genes that may be important for overall survival (OS). A multivariate Cox regression model was used to determine the final genes to be included after the lasso algorithm. Finally, three M1 macrophage-associated genes were developed, and RS were calculated according to the following equation:

$$riskscore = \sum_{i=1}^{n} (coef i^{*} Xi)$$

Here, *coef* was the regression coefficient in the multivariate Cox regression analysis as described previously. *X* indicated the expression of candidate genes. Where *i* indicated M1 macrophage-associated gene.

Validation of M1 macrophage-related prognostic features

Each BLCA sample was given a corresponding RS according to the previous risk formula. The cut-off point was set at the median RS. All samples were split into subgroups with low-risk group (LRG) and high-risk group (HRG). First, Kaplan-Meier curves were made to look for differences in prognosis. Moreover, the predictive value was verified by analyzing the time dependence of the receiver operating characteristic (ROC) curves (Xu et al., 2022a).

Creation and validation of nomogram

To predict overall survival at 1-, 3-, and 5-year, we developed a nomogram combining RS and other clinicopathological features. Finally, we plotted calibration curves capable of showing the prognostic validity of the nomogram.

Gene Set Enrichment Analysis

The GSEA software was used to look into the function annotation of the c2. cp.kegg.v7.4. symbols and c5. go.v7.4. symbols collections (Cao et al., 2021c). Results with p-value < 0.05 were being considered statistically significant. A graph was created based on the first eight outcomes.

Relationship between the TMB and RS

From the TCGA-BLCA cohort, we got information about somatic mutations. The number of somatic non-synonymous point mutations in each sample was found using the "maftools" R package.

Correlation of RS with TME

To determine if there was a correlation between RS and TICs, we measured immune cell infiltration in TME using seven different methods including XCELL, TIMER, QUANTISEQ, MCPcounter, EPIC, CIBERSORT, and CIBERSORT-ABS to evaluate the immune infiltrating situation (Xu et al., 2022b). Based on gene expression data, the ESTIMATE algorithm determines how many stromal cells and immune cells are present in a tumor sample using the stromal score and immune score. When you add up the two scores, you get the ESTIMATE score, which can be used to estimate how pure a tumor is. Spearman correlation analysis was used to find a link between RS and TICs.

Gene set variation analysis

We used the GSVA to estimate the activity of pathways so that we could compare the activity of pathways in different samples, which are listed in the MSigDB database (Xu et al., 2022c).

Prediction of patient response to immunotherapy

Immune checkpoints have been identified as key places where immune cells can be stopped from working (Zhang et al., 2020). In this study, we looked at how many 47 genes involved in blocking immune checkpoints were expressed in HRG and LRG. Immunophenoscore (IPS) predicts how a tumor will respond to immune checkpoint inhibitor treatment based on how immune-friendly it is. Each

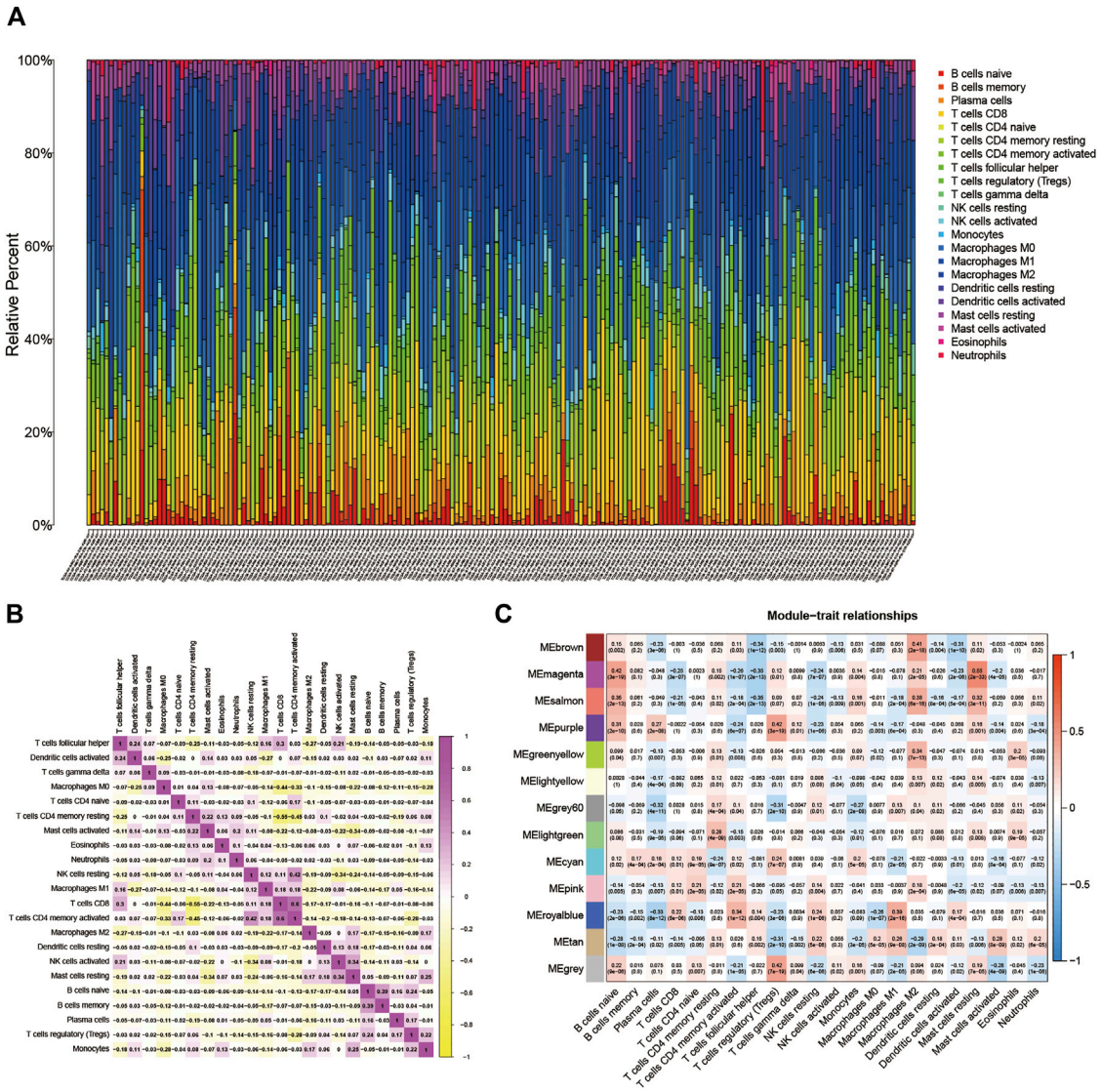


FIGURE 1
(A) Subpopulation of 22 immune cells. **(B)** The intrinsic correlation between 22 immune cells. **(C)** Heatmap of correlation between modules and TICs. Within every square, the number on the top refers to the coefficient between the cell infiltrating level and the corresponding module, and the bottom is the *p*-value.

immunophenotype (antigen-presenting, effector, suppressor, and checkpoint) is scored by IPS.

Prediction of the effects of chemotherapy

We made a ridge regression model using the Genomics of Drug Sensitivity in Cancer (GDSC) cell lines and the TCGA gene expression profiles to determine how drug sensitivity is different between HRG and LRG. The half-maximal inhibitory concentrations (IC50) of four chemotherapeutic agents in BLCA patients were calculated using the pRRophetic algorithm.

Experimental validation

The National Infrastructure of Cells was used to obtain SV-HUC-1 (a line of human bladder epithelial cells) and three lines of human bladder cancer cells (BIU-87, 5637, and T24). F-12K medium was used to grow the SV-HUC-1 cell line. In RPMI-1640 medium, 3 cell lines from people with bladder cancer were

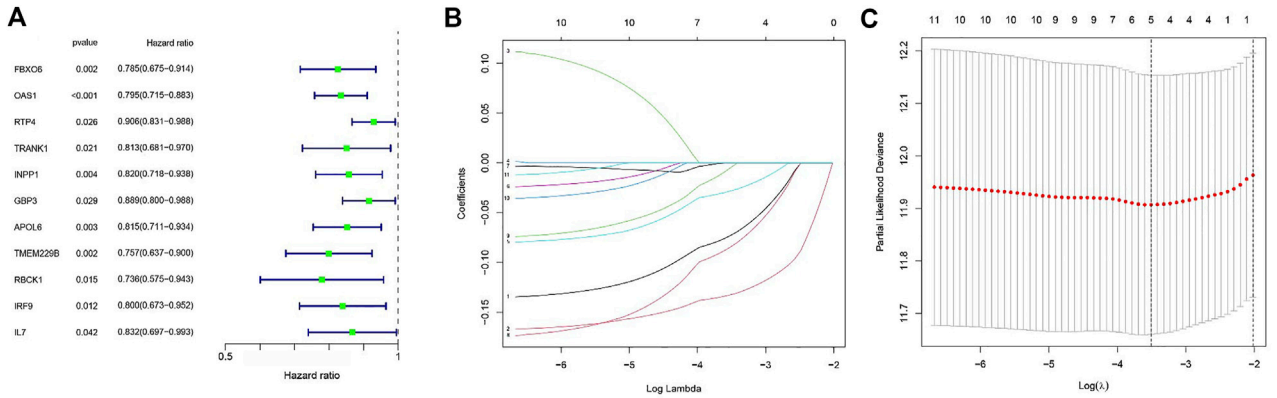


FIGURE 2
(A) Outcomes of univariate cox regression analysis. **(B)** LASSO coefficient profiles of 11 candidate genes. A vertical line is drawn at the value chosen by 10-fold cross-validation. **(C)** The setting parameters selected in the regression process were cross checked ten times. The vertical line represents the final three genes according to the best data.

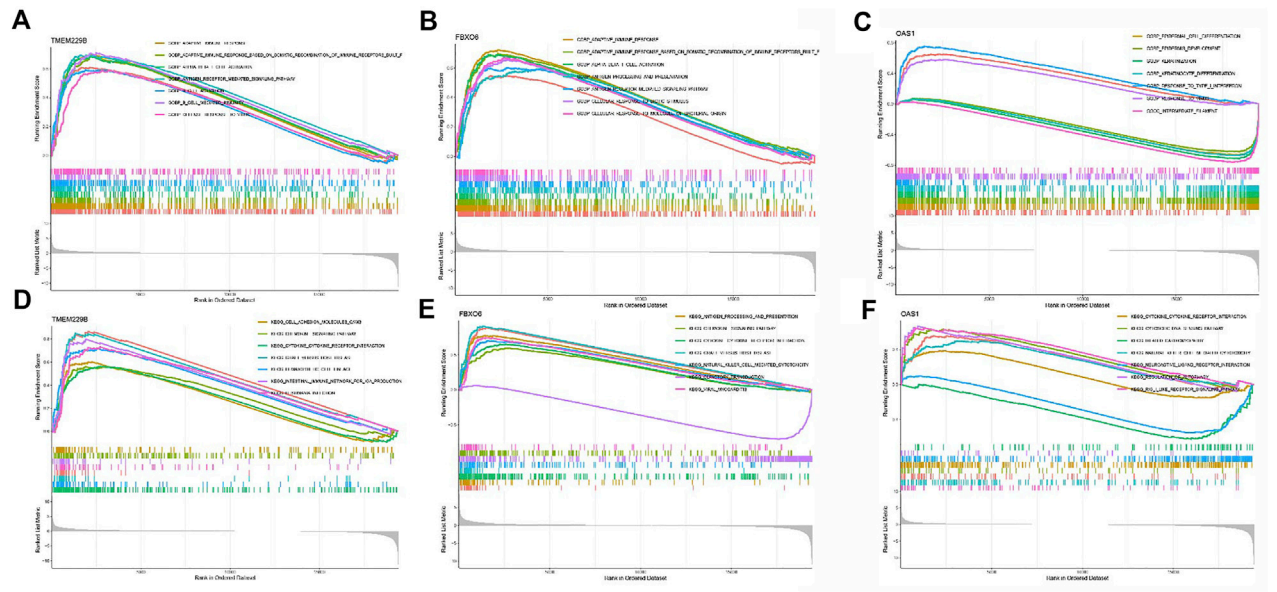


FIGURE 3
(A) The enriched gene sets in GO collection by the high TMEM229 expression sample. **(B)** The enriched gene sets in GO collection by the high FBXO6 expression sample. **(C)** The enriched gene sets in GO collection by the high OAS1 expression sample. **(D)** The enriched gene sets in KEGG collection by the high TMEM229 expression sample. **(E)** The enriched gene sets in KEGG collection by the high FBXO6 expression sample. **(F)** The enriched gene sets in KEGG collection by the high OAS1 expression sample.

grown. All cell lines were kept in an incubator that was set to 37°C and 5% CO_2. 10% fetal bovine serum and 1% double antibodies were added to all media. Four cell lines were put through reverse transcription-quantitative polymerase chain reaction (RT-qPCR). Three times, the experiment was done the same way. Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) levels were used as an endogenous control. We calculated the relative expression levels of genes FBXO6, OAS1, and TMEM229B using the $2^{-\Delta\Delta Ct}$ method. The sequences of the primers are as follows : FBXO6, 5′-CCC TACCGAGCTGTTGTCCA-3′ (forward) and 5′- GTTGAA CCGGGGCAGGAGTC-3′ (reverse); OAS1, 5′-AGACAC GTGTTTCCGCATGC-3′ (forward) and 5′-GAGCCACCC TTTACCACCTT-3′ (reverse); TMEM229B, 5′-GGAGAATGAGAGGAAGAA -3′ (forward) and 5′-AGAACCAGAACTGATACC -3′ (reverse); and GADPH,

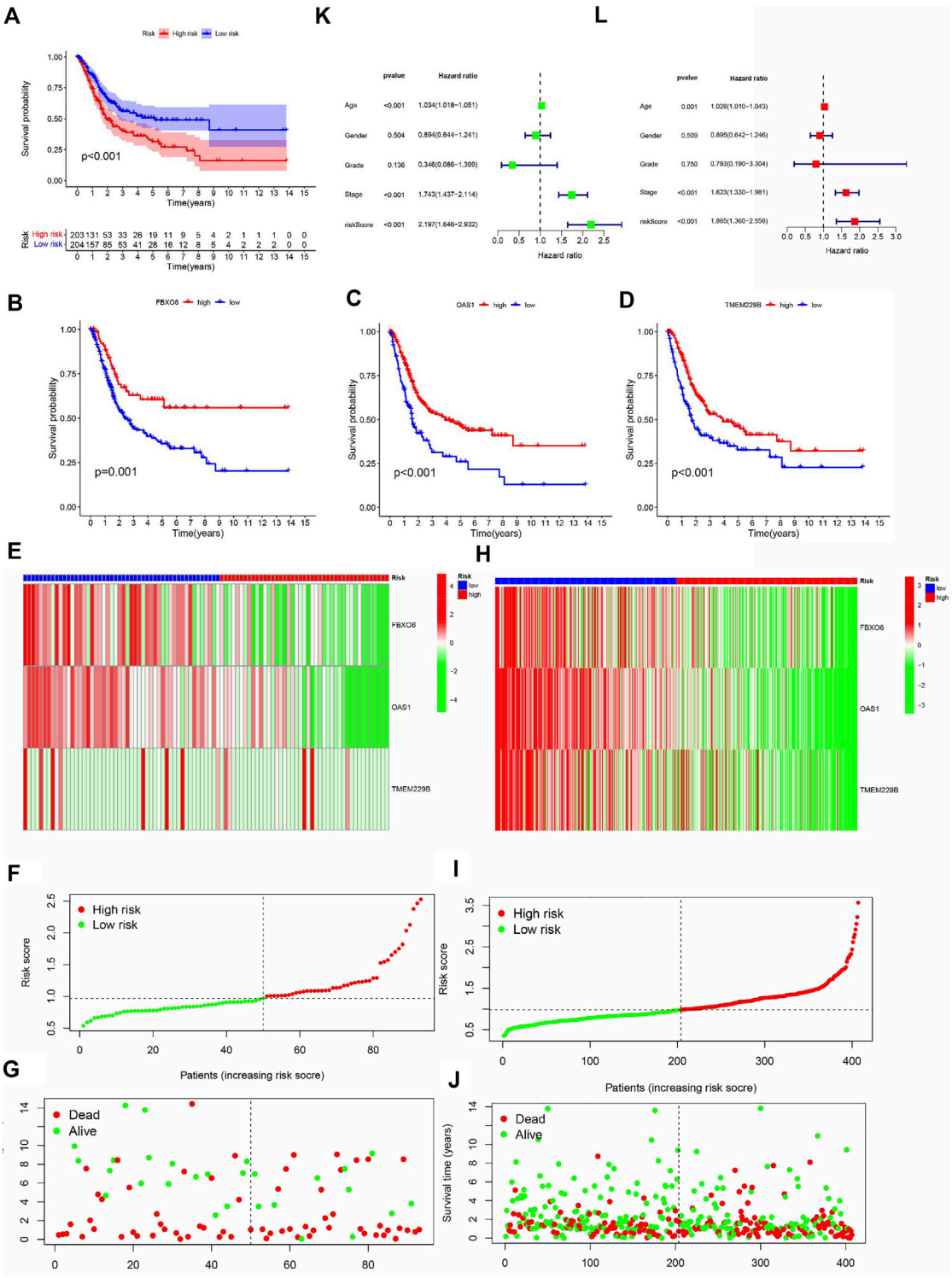


FIGURE 4
(A) The analysis of Kaplan-Meier curve shows that there is a difference in the overall survival rate between high-risk and low-risk groups. Kaplan-Meier curve analysis showed that there were differences in the overall survival rates of FBXO6 **(B)**, OAS1 **(C)**, and tem229b **(D)** in high/low-risk groups. In the GEO **(E)** and TCGA **(H)** cohorts, heat maps of FBXO6, OAS1, and TME229 gene ratios were drawn for each BLCA sample. In the GEO cohort**(F)** as well as in the TCGA cohort **(I)**, the distribution of multi-genes model risk score. In the GEO cohort**(G)** as well as in the TCGA cohort**(J)**, the survival status and duration of BLAC patients.

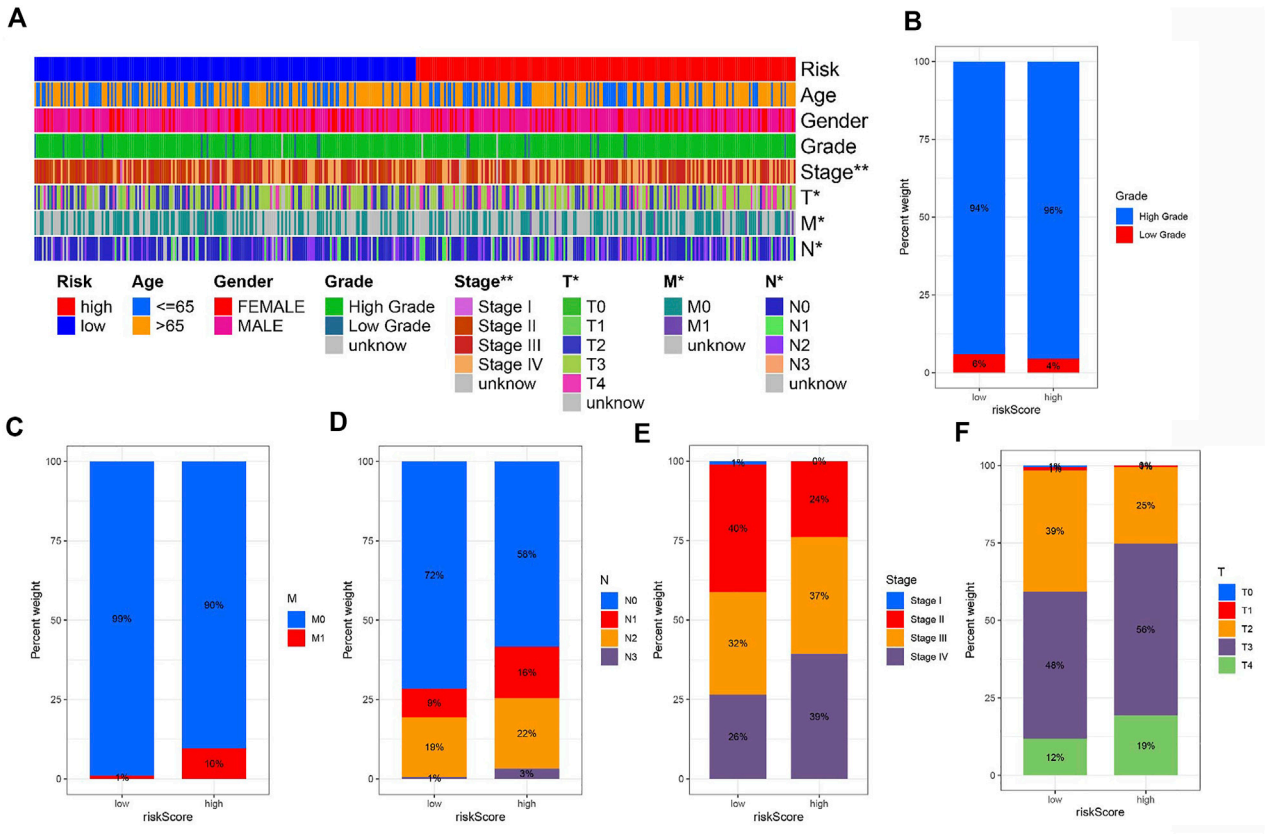


FIGURE 5
(A) The heat map shows the distribution of clinical characteristics and the corresponding risk score proportion in each sample. Rate of clinical variables subtypes in high or low risk score groups. **(B)** Grade, **(C)** M status, **(D)** N status, **(E)** Stage and **(F)** T status.

5′-CCTTCCGTGTTCCTACCC-3′ (forward) and 5′-CAA CCTGGTCCTCAGTGTAG-3′ (reverse).

Statistical analysis

Two groups were compared using Wilcoxon test, and more than two groups were compared using Kruskal Wallis test. Survival curves were analyzed by the Kaplan-Meier log rank test. Spearman analysis was used to determine the correlation coefficient between RS subgroup and somatic mutation frequency. A two-way p less than 0.05 was statistically significant. R software was used for all statistical analyses.

Results

The TME of BLCA

BLCA's TME situation is outlined using the CIBERSORT algorithm (Supplementary file: Supplementary Table S1). The 22 TICs in 407 samples are shown in Figure 1A. Supplementary Figure S1A depicts the relationship between the 22 TICs and the clinical phenotype. The potential connections between TICs and the associated relationships are further elucidated in Figure 1B.

WGCNA network establishment

16,394 permeating gene and TICs sequence documents were analyzed to establish the WGCNA network. First, the optimal soft threshold power (β) is set to 9 (Supplementary file: Supplementary Figure S1B). We set the module size to 60, and then inject genes with similar mappings into the same module to construct a hierarchical clustering tree of classes (Supplementary file: Supplementary Figure S1C). Based on the established criteria, the 13 gene modules were grouped and analyzed (using weighting and correlation). In Figure 1C, the horizontal coordinates are shown. There are 22 TIC types, and there are 13 modules in the vertical coordinate. There is a high correlation between the Meroyalblue module and M1 macrophages (cor = 0.39, p = 2e-16). In the present

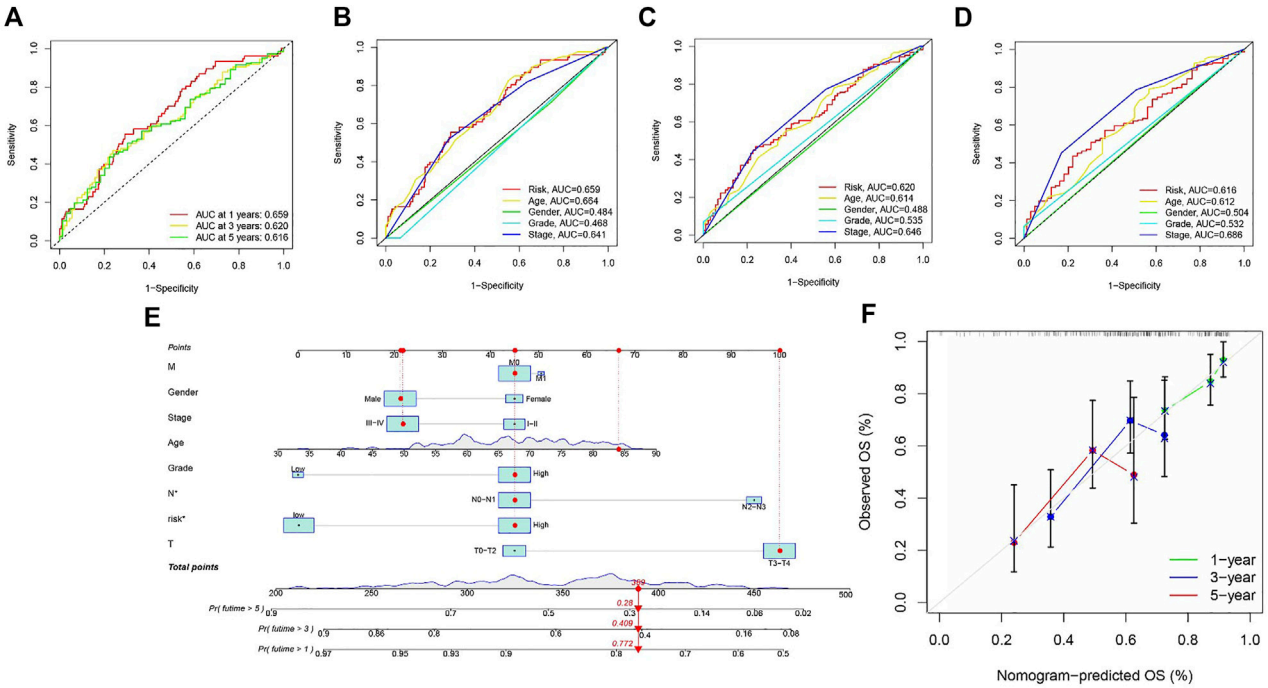


FIGURE 6
(A) ROC analysis is utilized to estimate the prediction characteristics of the prediction. **(B–D)** The prediction area under the risk assessment curve has other clinical characteristics of 1, 3, and 5-year total survival time. **(E)** Nomograms composed of clinical variables and risk characteristics were used to predict the survival rate of BLCA patients. **(F)** Calibration curves of 1-, 3-, and 5-year nomograms.

study, we focused on M1 macrophages, so we selected the Meroyalblue module (Supplementary file: Supplementary Table S2) for the follow-up study.

Establishment of risk signature

We extracted clinical information related to bladder cancer from the TCGA-BLCA. Eleven M1 macrophage-associated genes with prognostic value were identified based on univariate Cox analysis (p less than 0.05, Supplementary file: Supplementary Table S3), and gene-specific risk ratios are shown in Figure 2A. To prevent overfitting, we performed lasso regression analysis on the screened genes as well as cross-validation to find the optimal values of the penalty parameters. (Figures 2B,C). We then identified three M1 macrophage-associated genes (FBXO6, OAS1, and TMEM229B) as hub genes (all HRs <1, Supplementary file: Supplementary Table S4) by performing multivariate cox regression analysis.

We included three hub genes in the risk profile of BLCA patients and calculated risk scores (RS). Risk score = (−0.1255* expression value of FBXO6) + (−0.1619* expression value of OAS1) + (−0.1507* expression value of TMEM229B). A high-risk group and a low-risk group were determined based on the median cut-off value (0.983).

Subsequently, GSEA was used to determine the functional enrichment of the hub gene based on the median expression of the hub gene in all samples. Outcomes showed that TMEM229B expression was primarily influenced by *GRAFT* VS. *HOST DISEASE, CELL ADHESION MOLECULES CAMS, CHEMOKINE SIGNALING PATHWAY, HEMATOPOIETIC CELL LINEAGE, etc.* (Figures 3A,D). The elevated expression of FBXO6 was mainly associated with *ANTIGEN PROCESSING AND PRESENTATION, CHEMOKINE SIGNALING PATHWAY,* and so on (Figure 3B,E). And the elevated expression of OAS1 was mainly associated with *CYTOSOLIC DNA SENSING PATHWAY, REGULATION OF AUTOPHAGY*, and so on (Figures 3C,F).

Validation of risk signature

Kaplan-Meier curves showed that OS was lower in the HRG than in the LRG ($p < 0.001$) (Figure 4A), and OS was lower in the FBXO6, OAS1, and TMEM229B low-expression groups than in the high-expression group ($p = 0.001$, $p < 0.001$, $p < 0.001$) (Figures 4B–D). In the TCGA cohort, the expression of FBXO6, OAS1, and TMEM229B gradually decreased with increasing RS. And the point distribution of RS and survival status indicated that BLCA patients in the low-risk group had a longer OS than the HRG, also in the GEO

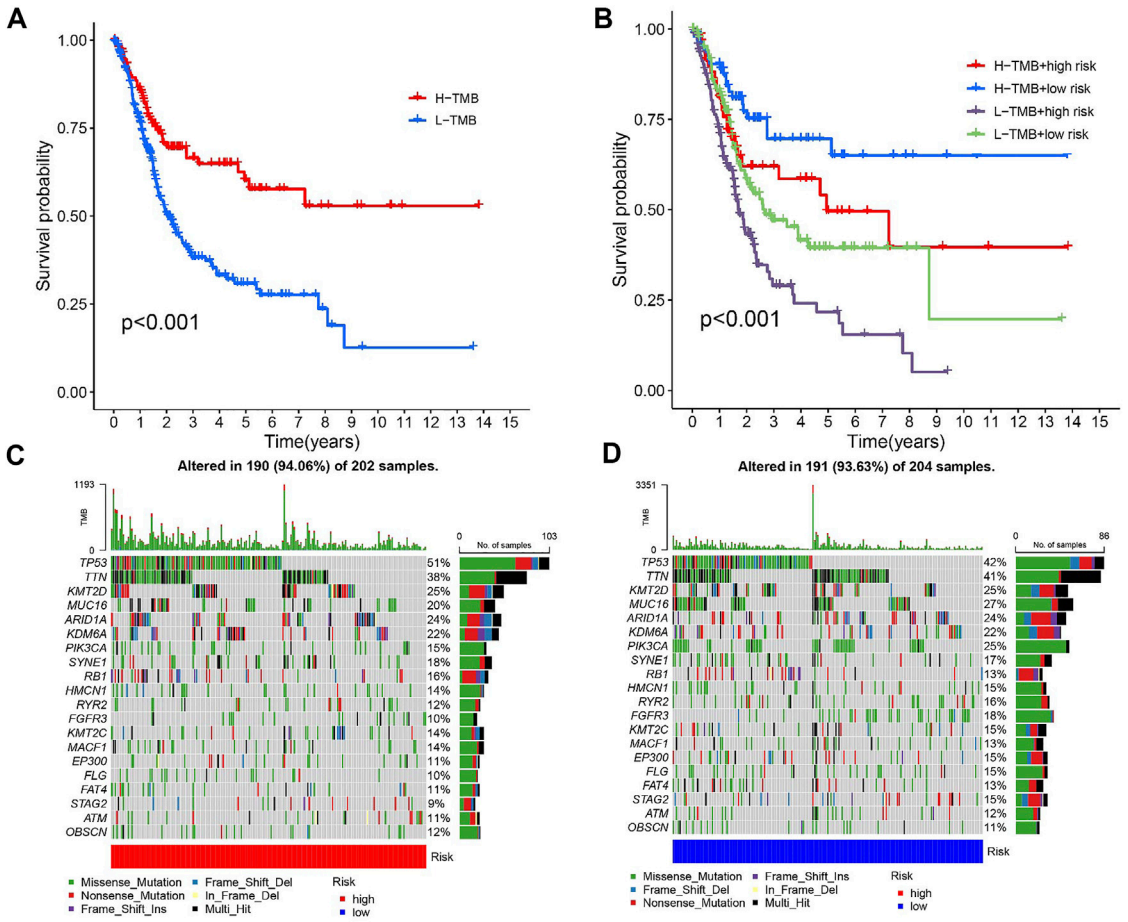


FIGURE 7
(A) Kaplan-Meier curve of high-level TMB group and low-level TMB group. **(B)** Kaplan-Meier curve is used to classify the risk level according to TMB. **(C)** The oncoPrint was constructed using a high risk score. **(D)** The oncoPrint was constructed using a low risk score.

cohort (Figures 4E–J). Univariate Cox analysis showed that the RS hazard ratios (HR) was 2.197 (95% of confidence interval 1.642–2.932) (Figure 4K), and the multivariate Cox regression analysis showed that the RS was 1.865 (95% of confidence interval 1.360–2.558) (Figure 4L).

To objectively evaluate the performance of our novel M1 macrophage signature, our studied signature was compared with traditional M1 macrophage markers (CD80, TNF-α, and iNOS) (Supplementary Figures S2B, S2F). Further, we combined the novel and traditional M1 macrophage signatures and found that the integrated signature had comparable performance to the novel signature we studied (Supplementary Figures S2A, S2F). T cells in the tumor microenvironment are essential bladder cancer immunotherapy-associated cells. We analyzed the correlation of the gene expression of T cell marker signature (CD4, CD8A, CCR4) and inhibitory molecule signature (CD279, CTLA4, HAVCR2) in HRG/LRG, and the results were shown in Supplementary Figure S2H. Our novel M1 macrophage marker signature was also compared with T cell marker signature and inhibitory molecule signature with ROC curves and C index at 1, 3 and 5 years as shown in Supplementary Figure S2C-E and Supplementary Figure S2G. In conclusion, RS can be used as an indicator to assess the prognosis of BLCA.

Risk signature and clinicopathological variables

To explore the correlation of risk and clinicopathological variables, we visualized a plot based on clinicopathological features. Figure 5A displays the distribution of clinical variables in HRG/LRG. There was different in grade, stage, T stage, N stage, and M stage (Figures 5B–F). The results were consistent with clinical practice. These findings, combined with results of univariable and multivariable regression analysis, emphasized that our risk score was indeed good prognostic predictive indicator independent from other clinical parameters.

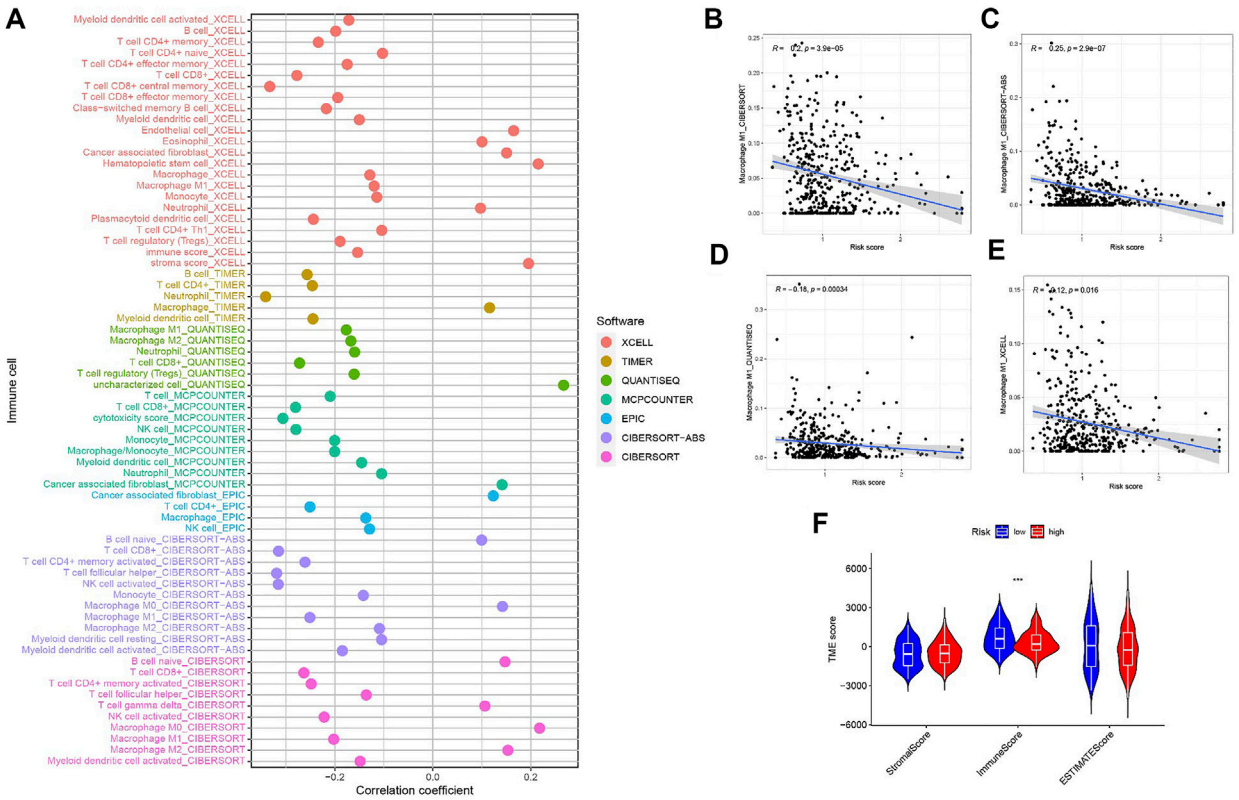


FIGURE 8
(A) Patients in the high-risk group were more positively associated with tumor-infiltrating immune cells, as shown by Spearman correlation analysis. **(B–E)** Correlation analysis of M1 macrophages and risk scores. **(F)** Correlation between TME score with risk scores.

Creation and validation of nomogram

Area under the Curve (AUC) of ROC of BLCA patients is 0.659, 0.620, and 0.616, respectively, which indicates high traceability for RS (Figure 6A). Next, we combined RS, age, gender, tumor grade, and clinical stage were then combined to analyze AUC for 1, 3, and 5 years. RS's AUC value was higher (Figures 6B–D), which further proved that RS had a better prognostic value. There is a nomogram consisting of RS, age, gender, tumor grade, clinical stage, T status, M category, and N category for quantitative prediction (Figure 6E). A good prediction performance is indicated by Figure 6F, indicating that the prediction nomogram was accurate.

TMB and prognosis

Previous studies have shown that patients with high TMB respond significantly better to immunotherapy (Samstein et al., 2019). Due to this, TMB has become one of the most important biological references for predicting tumor behavior and immunotherapy response. In this study, we found that the overall survival time was significantly lower for high TMB values ($p < 0.001$, Figure 7A). TMB and RS data were used to divide the patients into four subgroups. Patients with low TMB/high RS had the worst prognosis according to the survival analysis. (p less than 0.001, Figure 7B). The above outcomes demonstrate that TMB has an impact on the prognosis of BLCA. In addition, we further described the distribution of gene somatic mutations in the HRG/LRG (Figures 7C,D).

Risk signature in tumor immune microenvironment context of BLCA

Since M1 macrophages-based risk score and infiltration immune cells had intrinsic and intimate connection, we further explored the potential contribution of risk score in complexity and diversity of tumor immune microenvironment. Infiltration of immune cells and RS are correlated, as is shown in Figure 8A. Outcomes can be found in Supplementary file: Supplementary Table S5. Figures 8B–E reveals the correlation between M1 macrophages and risk score, suggesting a negative correlation between M1 macrophages and RS. Figure 8F shows that the immune score tends to be significantly higher in LRG.

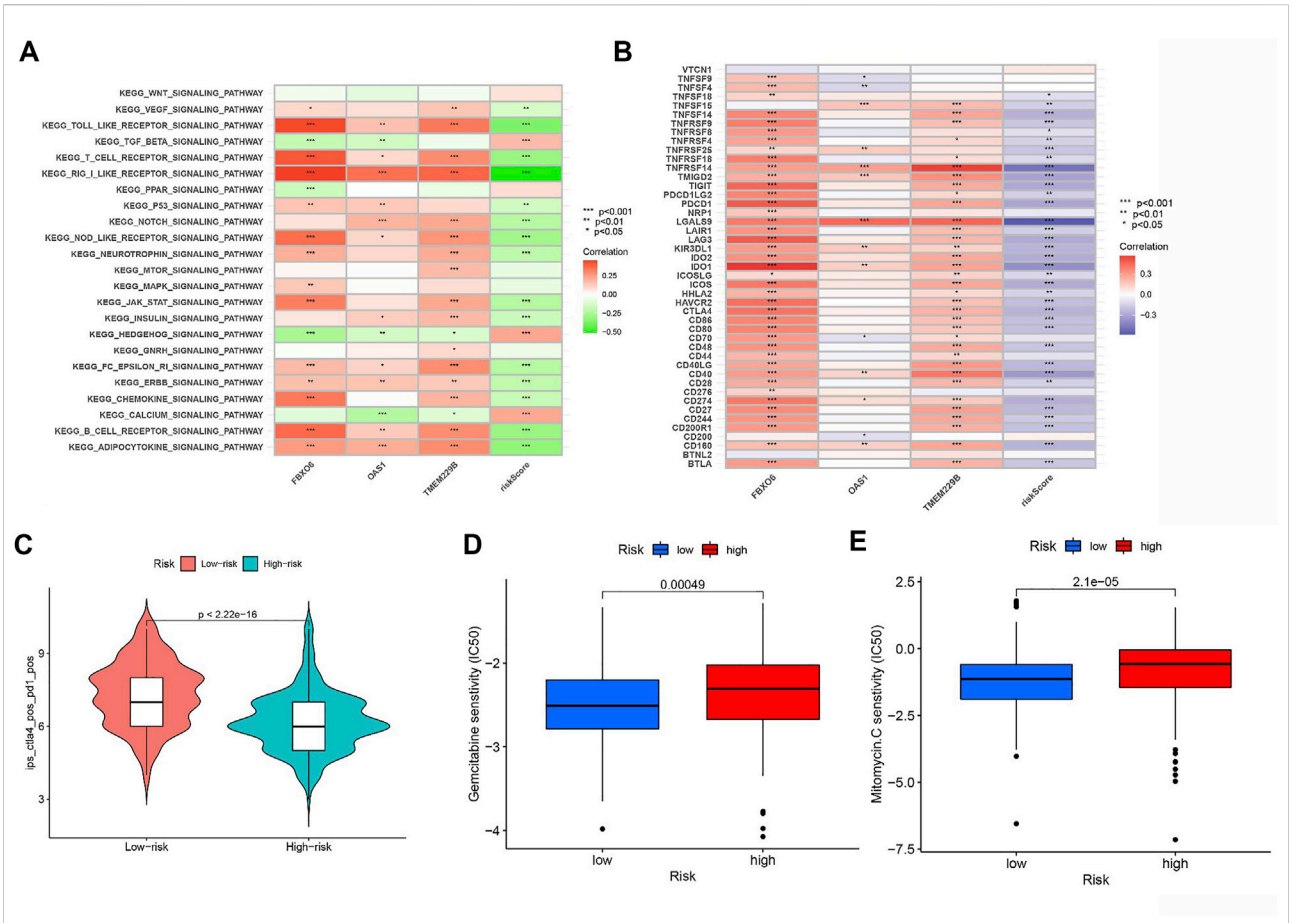


FIGURE 9
(A) Heatmap showing the correlation of representative pathway terms of KEGG with risk score. **(B)** Correlation of expression level of immune checkpoint blockade genes with risk score. **(C)** IPS score distribution plot. Estimation of risk score of chemotherapeutic effect. **(D)** Sensitivity analysis of Gemcitabine in patients with high and low risk score. **(E)** Sensitivity analysis of Mitomycin C in patients with high and low risk score.

Signaling pathways in two different risk groups

We further revealed the biological roles of signaling pathways in tumorigenesis and development in different risk groups by performing GSVA. Figure 9A shows that *TGF BETA SIGNALING PATHWAY, HEDGEHOG SIGNALING PATHWAY, and CALCIUM SIGNALING PATHWAY* activities were enhanced in the HRG.

Predicting clinical outcomes of immunotherapy in BLCA patients

Most genes associated with immune checkpoint blockade were found to be significantly negatively associated with risk scores (Figure 9B). According to the risk assessment system, LRG has a high IPS score (PD1-positive, CTLA4-positive, Figure 9C). Thus, PD1 and CTLA4 combined immunotherapy is suitable for LRG patients. While patients with HRG were more suitable for novel immunotherapies. The above outcomes suggest that RS correlates with immunotherapy response and that RS helps predict prognosis.

Prediction of chemotherapy response

We estimated the IC50 of two chemotherapeutic agents (gemcitabine and mitomycin C) in patients with BLCA according to the pRRophetic algorithm, which exhibited a higher IC50 in patients with HRG (both $p < 0.05$; Figures 9D,E). The results revealed that patients with LRG were more sensitive to chemotherapeutic agents.

Detection of mRNA levels of hub genes by RT-qPCR

To test the study hypothesis, in human bladder epithelial cells, we detected the expression of FBXO6, OAS1, and

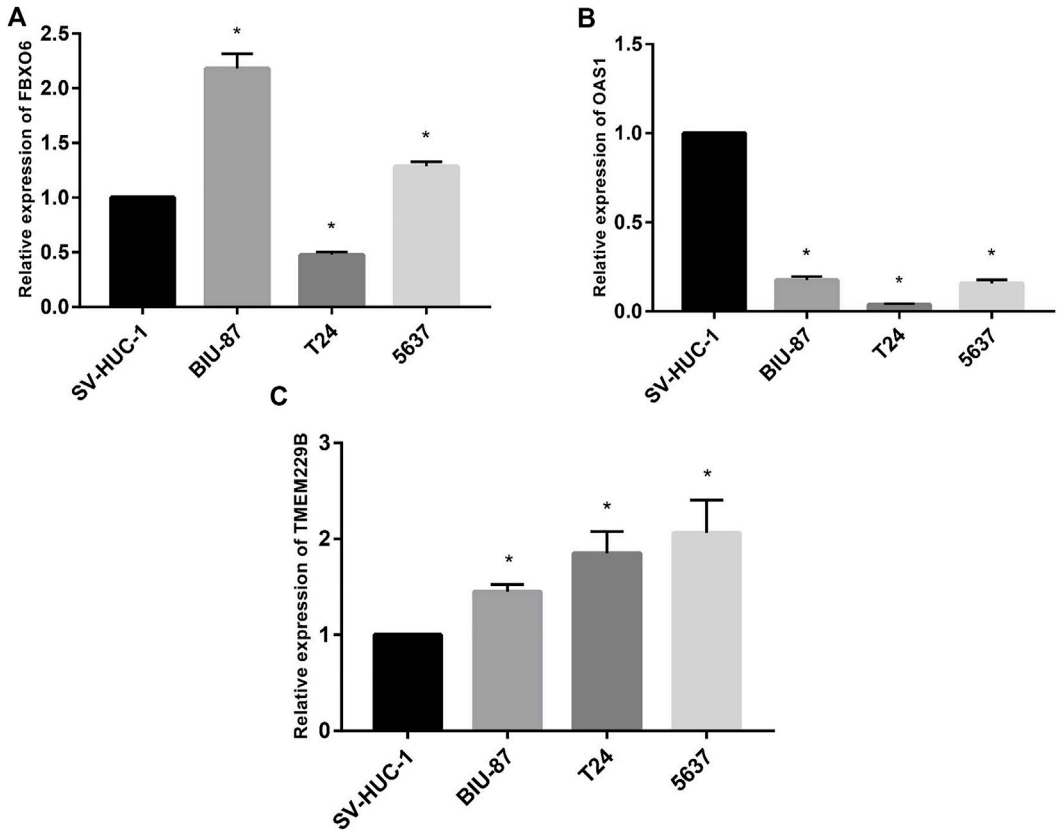


FIGURE 10
(A) The mRNA level of FBXO6 in normal urothelial cell line (SV-HUC-1) and three BLCA cell lines (BIU-87, T24, and 5637) was analyzed by RT-qPCR. **(B)** The mRNA level of OAS1 in normal urothelial cell line (SV-HUC-1) and three BLCA cell lines (BIU-87, T24, and 5637) was analyzed by RT-qPCR. **(C)** The mRNA level of TMEM229B in normal urothelial cell line (SV-HUC-1) and three BLCA cell lines (BIU-87, T24, and 5637) was analyzed by RT-qPCR.

TMEM229B genes, as well as three different bladder cancer cell lines by using RT-qPCR technique. OAS1 values were significantly lower in normal bladder epithelial cells than in bladder cancer cells (Figure 10B). Contrary to bladder epithelial cells, bladder cancer cells expressed significantly higher levels of FBXO6 and TMEM229B genes (Figures Figure10A,C). RT-qPCR outcomes well supported our findings.

Discussion

BLCA is one of the most common and most aggressive malignancies, the worried is that little progress has been made in the treatment of BLCA in the last few decades (Kimura et al., 2020; Lv et al., 2020). It is worth noticing that immunotherapy has shown great potential to treat MIBC and metastatic bladder cancer in the clinic (Nair et al., 2020; Pfail et al., 2021; Zhang et al., 2022). Previous studies have found that macrophages exert different effects on immunotherapeutic responses in advanced cancers (Petty and Yang, 2017; DeNardo and Ruffell, 2019; Zeng et al., 2020; Leblond et al., 2021). Sun et al. (Sun et al., 2022) found that the infiltration and polarization status of TAMs can predict outcomes of survival and chemotherapy benefits, as well as immunotherapy sensitivity in MIBC. Another study reported the value of M1 macrophages as a predictive biomarker for ICI treatment in patients with metastatic uroepithelial carcinoma (Zeng et al., 2020). These suggest that M1 macrophages take an irreplaceable role in tumor development.

In the current work, we obtained a total of 407 BLCA samples and 16,394 genes from the TCGA-BLCA and GSE31684 datasets to further study. The abundances of 22 TICs were obtained from the CIBERSORT algorithm, and the Meroyalblue module highly associated with M1 macrophages was constructed by WGCNA. Subsequently, Cox regression analysis was performed to identify three hub genes and took three hub genes in the risk profile of BLCA patients to calculate the RS. We also proved that RS can predict independently by Kaplan-Meier analysis and regression analysis. The nomogram was built and clinicopathological variables were investigated to strengthen the model prognostic value.

Our prognostic model was built based on three novel M1 macrophage-associated genes (FBXO6, OAS1, and TMEM229B). FBXO6 is an important member of the F-box protein family containing the FBA structural domain, which targets the DNA damage checkpoint kinase Chk1 to destroy S-phase arrest cells, phosphorylates during mitosis and dephosphorylates cells upon entry into the G1 phase. OAS1 is an interferon-induced protein that synthesizes adenosine oligomers from ATP to prevent tumor growth and cell differentiation (Piran et al., 2021). The TMEM is a family of proteins that span biological membranes. Many studies showed that the TMEM family can be described as tumor suppressors or oncogenes, and TMEM229B was reported can be a potential antigen for esophageal squamous cell carcinoma mRNA vaccines (Lu et al., 2022).

GSEA functional enrichment indicated that high expression of FBXO6 is related to immune response and chemokine signaling pathway, the elevated expression of OAS1 was associated with regulation of autophagy, and the elevated expression of TMEM229B was associated with humoral immune response. These outcomes suggest that the three hub genes are widely involved in tumor immunity, which provides a basis for our subsequent assessment of the efficacy of immunotherapy for RS.

TMB can predict survival after immunotherapy in types of cancer, especially when PD-1/PD-L1 is blocked. Moreover, TMB is a promising biomarker of the immune response. TMB may be an independent prognostic factor for multiple cancer immune responses. Therefore, we selected TMB as a prognostic indicator in this study and found patients with a high incidence rate of bladder cancer fared better than before. Subsequent stratified curves of survival showed that patients with low TMB/high RS had the worst prognosis, as well as that risk scores had independent prognostic predictive power.

Macrophages are important members of the innate immune response, and the polarization of macrophages allows them to have diverse functions. Their polarization status depends on environmental changes (Locati et al., 2020; Boutilier and SJIjoms, 2021). Direct metabolism of metabolites or cytokines drives the plasticity and heterogeneity of the tumor microenvironment (Mehla and Singh, 2019). We found that risk score was negatively associated with immune activating cell subpopulations such as CD8 T cells, and M1 macrophages were also negatively associated with RS. And immune scores were relatively low in the HRG. In summary, we speculated that the LRG group may be in immune activation state, which suppresses tumor progression and improves prognosis. In addition, TGFβ signaling pathway, *HEDGEHOG SIGNALING PATHWAY*, and *CALCIUM SIGNALING PATHWAY* were also activated in high-risk groups, suggesting that the high-risk group has different molecular mechanisms.

Finally, we explored the predictive value of RS in BLCA immunotherapy and chemotherapy. The results show that ICB-related genes had a significantly negative risk score, which indicates a higher IPS score for LRG. This suggests that LRG patients are suitable for PD1 and CTLA4 combination with immune therapy, whereas HRG patients prefer to novel immunotherapies.

There are still some limitations in our study, the functions of M1 macrophage-related genes need to be further explored in animal experiments. More importantly, we need to further validate the prediction model in different cohorts collected by multiple centers.

Conclusion

In conclusion, the RS model based on M1 macrophages can be utilized to predict clinical outcomes, treatment outcomes, as well as prognosis of BLCA patients and provide reference for the treatment of BLCA somehow.

Data availability statement

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

Author Contributions

CX designed this work. YH, YY, and SO analyzed the data. YH, CL, and CX wrote this manuscript. ZK edited and revised the manuscript. All authors approved this manuscript.

Funding

This study was funded by Henan Provincial Health Commission (LHGJ20190422).

Conflict of interest

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

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those

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

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.1047004/full#supplementary-material

References

Barone, B., Calogero, A., Scafuri, L., Ferro, M., Lucarelli, G., Di Zazzo, E., et al. (2022). Immune checkpoint inhibitors as a neoadjuvant/adjuvant treatment of muscle-invasive bladder cancer: A systematic review. *Cancers* 14 (10), 2545. doi:10.3390/cancers14102545

Boutilier, A., and Sjijoms, E. (2021). Macrophage polarization states in the tumor microenvironment. *Int. J. Mol. Sci.* 22 (13), 6995. doi:10.3390/ijms22136995

Cao, R., Yuan, L., Ma, B., Wang, G., and Tian, Y. J. C. (2021). Immunotherapy : CII. Tumour microenvironment (TME) characterization identified prognosis and immunotherapy response in muscle-invasive bladder cancer (MIBC). *Cancer Immunol. Immunother.* 70 (1), 1–18. doi:10.1007/s00262-020-02649-x

Cao, F., Fan, Y., Yu, Y., Yang, G., and Zhong, H. (2021). Dissecting prognosis modules and biomarkers in glioblastoma based on weighted gene Co-expression network analysis. *Cancer Manag. Res.* 13, 5477–5489. doi:10.2147/cmar.S310346

Cao, F., Wang, C., Long, D., Deng, Y., Mao, K., and Zhong, H. (2021). Network-based integrated analysis of transcriptomic studies in dissecting gene signatures for LPS-induced acute lung injury. *Inflammation* 44 (6), 2486–2498. doi:10.1007/s10753-021-01518-8

DeNardo, D., and Ruffell, B. J. NrI. (2019). Macrophages as regulators of tumour immunity and immunotherapy. *Nat. Rev. Immunol.* 19 (6), 369–382. doi:10.1038/s41577-019-0127-6

Erbani, J., Boon, M., and Ljsicb, A. (2022). Therapy-induced shaping of the glioblastoma microenvironment: Macrophages at play. *Semin. Cancer Biol.* 86, 41–56. doi:10.1016/j.semcancer.2022.05.003

Friedrich, M., Hahn, M., Michel, J., Sankowski, R., Kilian, M., Kehl, N., et al. (2022). Dysfunctional dendritic cells limit antigen-specific T cell response in glioma. *Neuro. Oncol.* 138, noac138. doi:10.1093/neuonc/noac138

Garris, C., Wong, J., Ravetch, J., and Djstm, K. (2021). Dendritic cell targeting with Fc-enhanced CD40 antibody agonists induces durable antitumor immunity in humanized mouse models of bladder cancer. *Sci. Transl. Med.* 13 (594), eabd1346. doi:10.1126/scitranslmed.abd1346

Ge, Z., and Ding, S. J. F. (2020). The crosstalk between tumor-associated macrophages (TAMs) and tumor cells and the corresponding targeted therapy. *Front. Oncol.* 10, 590941. doi:10.3389/fonc.2020.590941

Groeneveld, C., Fontugne, J., Cabel, L., Bernard-Pierrot, I., Radvanyi, F., Allory, Y., et al. (2021). Tertiary lymphoid structures marker CXCL13 is associated with better survival for patients with advanced-stage bladder cancer treated with immunotherapy. *Eur. J. Cancer* 148, 181–189. doi:10.1016/j.ejca.2021.01.036

Jia, W., Luo, S., Lai, G., Li, S., Huo, S., Li, M., et al. (2021). Homogeneous polyporus polysaccharide inhibits bladder cancer by polarizing macrophages to M1 subtype in tumor microenvironment. *BMC Complement. Med. Ther.* 21 (1), 150. doi:10.1186/s12906-021-03318-x

Kimura, T., Ishikawa, H., Kojima, T., Kandori, S., Kawahara, T., Sekino, Y., et al. (2020). Bladder preservation therapy for muscle invasive bladder cancer: The past, present and future. *Jpn. J. Clin. Oncol.* 50 (10), 1097–1107. doi:10.1093/jjco/hyaa155

Kubrak, T., Karakuła, M., Czop, M., Kawczyk-Krupka, A., and Aebisher, D. J. M. (2022). Advances in management of bladder cancer-the role of photodynamic therapy. *Molecules* 27 (3), 731. doi:10.3390/molecules27030731

Larroquette, M., Guegan, J., Besse, B., Cousin, S., Brunet, M., Le Moulec, S., et al. (2022). Spatial transcriptomics of macrophage infiltration in non-small cell lung cancer reveals determinants of sensitivity and resistance to anti-PD1/PD-L1 antibodies. *J. Immunother. Cancer* 10 (5), e003890. doi:10.1136/jitc-2021-003890

Leblond, M., Zdimerova, H., Desponds, E., and Verdeil, G. J. C. (2021). Tumor-associated macrophages in bladder cancer: Biological role, impact on therapeutic response and perspectives for immunotherapy. *Cancers* 13 (18), 4712. doi:10.3390/cancers13184712

Locati, M., Curtale, G., and Ajarop, M. (2020). Diversity, mechanisms, and significance of macrophage plasticity. *Annu. Rev. Pathol.* 15, 123–147. doi:10.1146/annurev-pathmechdis-012418-012718

Lu, T., Xu, R., Wang, C. H., Zhao, J. Y., Peng, B., Wang, J., et al. (2022). Identification of tumor antigens and immune subtypes of esophageal squamous cell carcinoma for mRNA vaccine development. *Front. Genet.* 13, 853113. doi:10.3389/fgene.2022.853113

Lv, J., Zhu, Y., Ji, A., Zhang, Q., and Liao, G. J. B. (2020). Mining TCGA database for tumor mutation burden and their clinical significance in bladder cancer. *Biosci. Rep.* 40 (4), BSR20194337. doi:10.1042/bsr20194337

Mehla, K., and Singh, P. J. T. (2019). Metabolic regulation of macrophage polarization in cancer. *Trends Cancer* 5 (12), 822–834. doi:10.1016/j.trecan.2019.10.007

Nair, S., Weil, R., Dovey, Z., Davis, A., and AjtucoNA, T. (2020). The tumor microenvironment and immunotherapy in prostate and bladder cancer. *Urol. Clin. North Am.* 47, e17–e54. doi:10.1016/j.ucl.2020.10.005

Pang, L., Khan, F., Heimberger, A., and Chen, P. J. T. (2022). Mechanism and therapeutic potential of tumor-immune symbiosis in glioblastoma. *Trends Cancer* 8 (10), 839–854. doi:10.1016/j.trecan.2022.04.010

Petty, A., and Yang, Y. J. I. (2017). Tumor-associated macrophages: Implications in cancer immunotherapy. *Immunotherapy* 9 (3), 289–302. doi:10.2217/imt-2016-0135

Pfail, J., Katims, A., Alerasool, P., and Jjwjou, S. (2021). Immunotherapy in non-muscle-invasive bladder cancer: Current status and future directions. *World J. Urol.* 39 (5), 1319–1329. doi:10.1007/s00345-020-03474-8

Piran, M., Sepahi, N., Moattari, A., Rahimi, A., and Ghanbariasad, A. (2021). Systems biomedicine of primary and metastatic colorectal cancer reveals potential therapeutic targets. *Front. Oncol.* 11, 597536. doi:10.3389/fonc.2021.597536

Samstein, R. M., Lee, C. H., Shoushtari, A. N., Hellmann, M. D., Shen, R., Janjigian, Y. Y., et al. (2019). Tumor mutational load predicts survival after immunotherapy across multiple cancer types. *Nat. Genet.* 51 (2), 202–206. doi:10.1038/s41588-018-0312-8

Sun, M., Zeng, H., Jin, K., Liu, Z., Hu, B., Liu, C., et al. (2022). Infiltration and polarization of tumor-associated macrophages predict prognosis and therapeutic benefit in muscle-invasive bladder cancer. *Cancer Immunol. Immunother.* 71 (6), 1497–1506. doi:10.1007/s00262-021-03098-w

Wang, X., Miao, J., Wang, S., Shen, R., Zhang, S., Tian, Y., et al. (2022). Single-cell RNA-seq reveals the Genesis and heterogeneity of tumor microenvironment in pancreatic undifferentiated carcinoma with osteoclast-like giant-cells. *Mol. Cancer* 21 (1), 133. doi:10.1186/s12943-022-01596-8

Wu, D., Liu, X., Mu, J., Yang, J., Wu, F., and Zhou, H. J. B. (2022). Therapeutic approaches targeting proteins in tumor-associated macrophages and their applications in cancers. *Biomolecules* 12 (3), 392. doi:10.3390/biom12030392

Wu, K., Lin, K., Li, X., Yuan, X., Xu, P., Ni, P., et al. (2020). Redefining tumor-associated macrophage subpopulations and functions in the tumor microenvironment. *Front. Immunol.* 11, 1731. doi:10.3389/fimmu.2020.01731

Xu, C., Pei, D., Liu, Y., Yu, Y., Guo, J., Liu, N., et al. (2022). Identification of a novel tumor microenvironment prognostic signature for bladder urothelial carcinoma. *Front. Oncol.* 12, 818860. doi:10.3389/fonc.2022.818860

Xu, C., Song, L., Peng, H., Yang, Y., Liu, Y., Pei, D., et al. (2022). Clinical eosinophil-associated genes can serve as a reliable predictor of bladder urothelial cancer. *Front. Mol. Biosci.* 9, 963455. doi:10.3389/fmolb.2022.963455

Xu, C., Song, L., Yang, Y., Liu, Y., Pei, D., Liu, J., et al. (2022). Clinical M2 macrophage-related genes can serve as a reliable predictor of lung adenocarcinoma. *Front. Oncol.* 12, 919899. doi:10.3389/fonc.2022.919899

Zeng, D., Ye, Z., Wu, J., Zhou, R., Fan, X., Wang, G., et al. (2020). Macrophage correlates with immunophenotype and predicts anti-PD-L1 response of urothelial cancer. *Theranostics* 10 (15), 7002–7014. doi:10.7150/thno.46176

Zhang, Q., Zhong, H., Fan, Y., Liu, Q., Song, J., Yao, S., et al. (2020). Immune and clinical features of CD96 expression in glioma by *in silico* analysis. *Front. Bioeng. Biotechnol.* 8, 592. doi:10.3389/fbioe.2020.00592

Zhang, Y., Huo, F., Cao, Q., Jia, R., Huang, Q., Wang, Z., et al. (2022). FimH confers mannose-targeting ability to Bacillus Calmette-Guerin for improved immunotherapy in bladder cancer. *J. Immunother. Cancer* 10 (3), e003939. doi:10.1136/jitc-2021-003939

Zhong, H., Liu, S., Cao, F., Zhao, Y., Zhou, J., Tang, F., et al. (2021). Dissecting tumor antigens and immune subtypes of glioma to develop mRNA vaccine. *Front. Immunol.* 12, 709986. doi:10.3389/fimmu.2021.709986

Glossary

TME tumor microenvironment

BLCA bladder urothelial carcinoma

TICs tumor-infiltrating immune cells

TCGA Cancer Genome Atlas

GEO Gene Expression Omnibus

WGCNA weighted gene co-expression network analysis

ROC the receiver operating characteristic

LRG low-risk group

HRG high-risk group

RS risk score

RT-qPCR reverse transcription-quantitative polymerase chain reaction

PD1 programmed cell death protein 1

CTLA4 cytotoxic T lymphocyte associate protein-4

WHO World Health Organization

MIBC muscle-invasive bladder cancer

NMIBC non-muscle invasive bladder cancer

PD-L1 programmed death ligand 1

ICIs immune checkpoint inhibitors

TMB tumor mutational burden

ECM extracellular matrix

TAM tumor-associated macrophages

TCGA-BLCA The Cancer Genome Atlas-Bladder Cancer

CNV copy number variation

OS overall survival

GSEA Gene Set Enrichment Analysis

GSVA Gene set variation analysis

IPS Immunophenoscore

GDSC Genomics of Drug Sensitivity in Cancer

IC50 half-maximal inhibitory concentrations

GAPDH Glyceraldehyde-3-phosphate dehydrogenase

HR hazard ratios

AUC Area under the Curve

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 01 December 2022
DOI 10.3389/fgene.2022.1061382

OPEN ACCESS

EDITED BY
Duo Liu,
Harbin Medical University Cancer Hospital, China

REVIEWED BY
Fei Yuan,
Baylor College of Medicine, United States
Hehai Pan,
University of Pennsylvania, United States

*CORRESPONDENCE
Guanghe Fei,
guanghefei@hotmail.com
Sijing Zhou,
zhousijing@yeah.net
Ran Wang,
wangran@ahmu.edu.cn

†These authors have contributed equally to this work

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

RECEIVED 04 October 2022
ACCEPTED 10 November 2022
PUBLISHED 01 December 2022

CITATION
Sun P, Xu H, Zhu K, Li M, Han R, Shen J, Xia X, Chen X, Fei G, Zhou S and Wang R (2022), The cuproptosis related genes signature predicts the prognosis and correlates with the immune status of clear cell renal cell carcinoma. *Front. Genet.* 13:1061382. doi: 10.3389/fgene.2022.1061382

The cuproptosis related genes signature predicts the prognosis and correlates with the immune status of clear cell renal cell carcinoma

Peng Sun[1†], Hua Xu[2†], Ke Zhu[1†], Min Li[3], Rui Han[1], Jiran Shen[1], Xingyuan Xia[1], Xiaojuan Chen[4], Guanghe Fei[1*], Sijing Zhou[5*] and Ran Wang[1*]

[1]Department of Respiratory and Critical Care Medicine, The First Affiliated Hospital of Anhui Medical University, Hefei, China, [2]Department of Nursing, Hefei Second People's Hospital, Hefei, China, [3]Department of Oncology, The First Affiliated Hospital of Anhui Medical University, Hefei, China, [4]Department of Infectious Diseases, Hefei Second People's Hospital, Hefei, China, [5]Department of Occupational Disease, Hefei Third Clinical College of Anhui Medical University, Hefei, China

Background: Clear cell renal cell carcinoma (CCRCC) has a high incidence and poor prognosis. Cuproptosis, an independent pattern of cell death associated with copper, plays an important role in cancer proliferation and metastasis. The role of cuproptosis-related genes (CRGs) in CCRCC is unclear.

Methods: Transcriptome and clinical information for CCRCC were downloaded from The Cancer Genome Atlas (TCGA) database. After dividing the training and testing cohort, a 4-CRGs risk signature (*FDX1*, *DLD*, *DLAT*, *CDKN2A*) was identified in the training cohort using Least absolute shrinkage and selection operator (LASSO) and Cox regression analysis. The effect of the 4-CRGs risk signature on prognosis was assessed using Kaplan-Meier (KM) curves and time-dependent receiver operating characteristic (ROC) curves and verified using the testing cohort. For different risk groups, the immune statue was assessed using the CIBERSORT algorithm, the ssGSEA method and immune checkpoint expression data. Finally, a competitive endogenous RNA (ceRNA) network was constructed using miRTarbase and starBase databases to identify molecules that may have a regulatory relationship with CRCCC.

Results: There were significant changes in the overall survival (OS), immune microenvironment, immune function, and checkpoint gene expression among the different risk groups. A ceRNA network consisting of one mRNA, two miRNAs, and 12 lncRNAs was constructed.

Conclusion: The 4-CRGs risk signature provides a new method to predict the prognosis of patients with CCRCC and the effect of immunotherapy. We propose a new cuproptosis-associated ceRNA network that can help to further explore the molecular mechanisms of CCRCC.

KEYWORDS

cuproptosis, clear cell renal cell carcinoma, immune, ceRNAs network, TCGA

1 Introduction

Renal cell carcinoma (RCC) is one of the top ten most common cancers in the world, ranking sixth and eighth in new cases in men and women, respectively (Siegel et al., 2020). Clear cell renal cell carcinoma (CCRCC), is the most common type of RCC, accounting for about 75–80% of RCCs (Nabi et al., 2018). Early diagnosis of RCC is difficult because only 6–10% of patients present with typical symptoms, such as hematuria, back pain, or abdominal mass (Patard et al., 2003). Furthermore, the effect of chemotherapy and radiation therapy in patients with CCRCC is not ideal, and tumor removal is the best treatment option (Sonpavde et al., 2012). Immunotherapy is an emerging and promising therapeutic option, and some immune checkpoint inhibitors (ICIs) have been approved for metastatic CCRCC (nivolumab) after failed targeted therapy or in combination with targeted drugs (pembrolizumab + axitinib/avelumab + axitinib) as first-line therapy (Motzer et al., 2015; Motzer et al., 2019; Rini et al., 2019). However, in actual clinical practice, there are still problems regarding which treatment methods can be used for individual patients, especially advanced patients. Therefore, it is necessary to establish a reliable predictive model to predict patient survival and guide the choice of different treatment options. We aimed to identify potential targets with prognostic implications for CCRCC from the perspective of cuproptosis, a newly discovered pattern of cell death.

Copper accumulates within cells and can induce cell death when a certain concentration is reached (Tsvetkov et al., 2022). In tumors, copper is involved in cell proliferation, epithelial-mesenchymal transition (EMT), angiogenesis, immunity, inflammation, and metastasis of tumors (De Luca et al., 2019; da Silva et al., 2022). Copper chelation may inhibit these processes to exert anti-tumor and anti-metastatic effects (Denoyer et al., 2015; Shanbhag et al., 2021). Copper ionophores have recently been shown to induce a novel mechanism of cell death (cuproptosis) (Tsvetkov et al., 2022). There have been studies on copper ion carriers that play a role in anti-tumor activity. For example, disulfiram has a significant tumor growth inhibition effect in patients with prostate (Safi et al., 2014) and breast cancer (Zhang et al., 2010; Allensworth et al., 2015). Disulfiram can improve the survival rate of patients treated with cisplatin and vinorelbine for non-small cell lung cancer (Nechushtan et al., 2015). Cuproptosis-related genes may serve as new targets for cancer treatment, but there are few studies on copper ionophores and cuproptosis.

Here, we downloaded the TCGA-KIRC dataset to identify differential genes associated with cuproptosis in tumor tissue and normal samples and validated them using the GSE53757 dataset. A risk score model containing four cuproptosis-related genes was constructed using LASSO and Cox regression, and the correlation between the risk score model and immune function, immune infiltrates, immune escape, and cancer treatment drugs were analyzed. Finally, a possible ceRNA network was constructed by searching for miRNAs and lncRNAs associated with cuproptosis-related genes using TCGA, miRTarbase, and starBase databases.

2 Materials and methods

2.1 Data collection

The RNA-sequencing TPM data and corresponding clinical data of KIRC were retrieved from the TCGA database (https://portal.gdc.cancer.gov/), including 541 KIRC samples with complete survival data and 72 normal samples. The RNA-sequencing TPM data and corresponding clinical data of LGG were retrieved from the TCGA database (https://portal.gdc.cancer.gov/), including 479 LGG samples. The GSE22541 dataset was downloaded from GEO (http://www.ncbi.nlm.nih.gov/geo) and included 68 CCRCC samples with survival data. The GSE53757 dataset was downloaded from GEO (http://www.ncbi.nlm.nih.gov/geo) and included 72 CCRCC samples and 72 normal samples (von Roemeling et al., 2014). *FDX1*, *LIPT1*, *LIAS*, *DLD*, *DLAT*, *PDHA1*, *PDHB*, *MTF1*, *GLS*, and *CDKN2A* are thought to be cuproptosis-related genes and are involved in two structurally distinct copper-loaded ionophores (Tsvetkov et al., 2022).

2.2 Expression patterns of cuproptosis-related genes in clear cell renal cell carcinoma

RNA-sequencing TPM data from the TCGA database were used to compare the expression of cuproptosis-related genes in CCRCC specimens and normal specimens using the Wilcoxon rank-sum test. Statistical significance was set at $p < 0.05$. The GSE53757 dataset was used for further validation.

2.3 Correlation analysis and GO and KEGG analysis

For significant prognosis-related cuproptosis-related genes, we performed gene co-expression analysis in TCGA CCRCC patients and set the absolute value of the correlation coefficient to greater than 0.4 with a p-value less than 0.001 to obtain the co-expression genes. To further

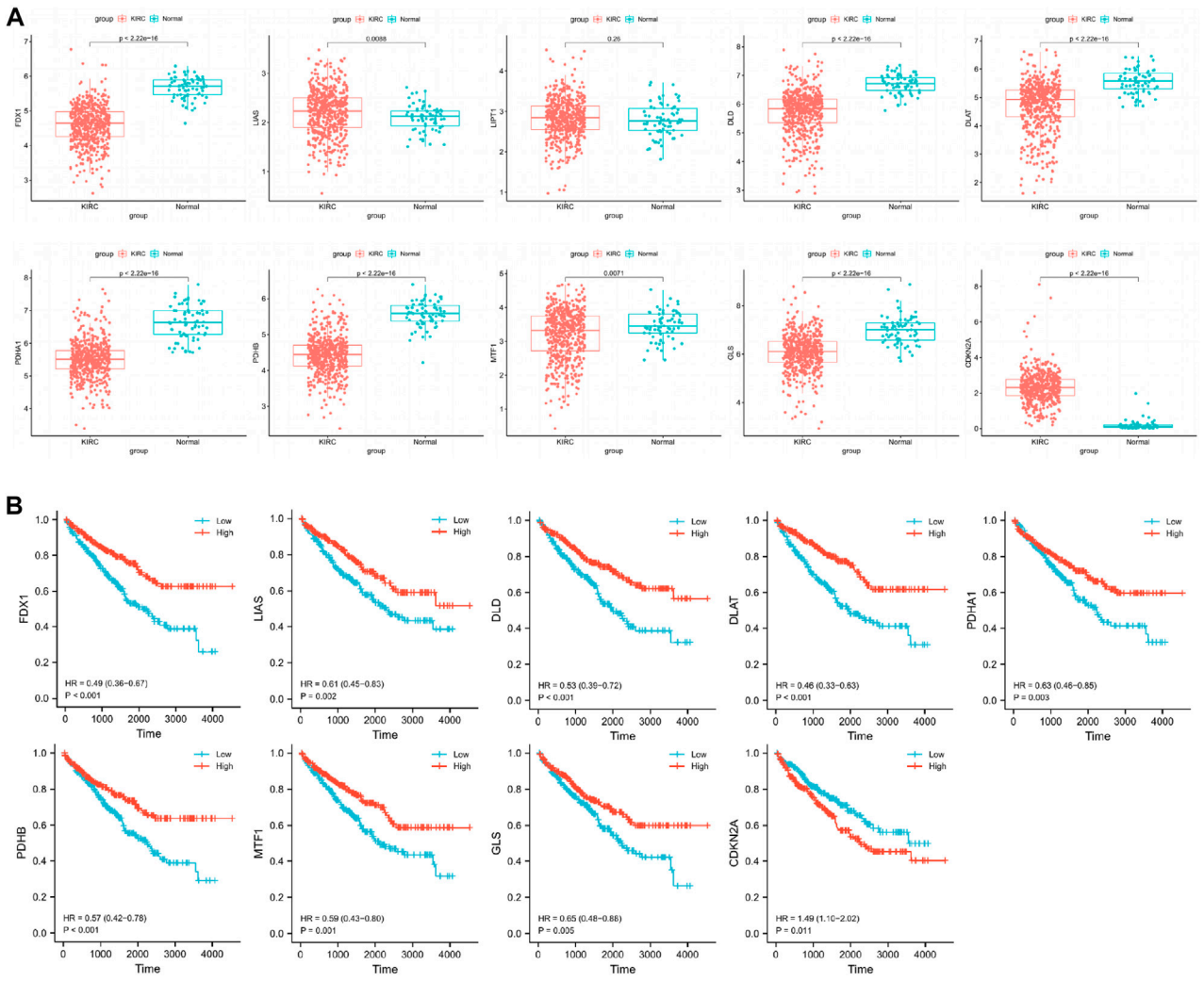


FIGURE 1
(A) Identification of differentially cuproptosis-related expressed genes in TCGA cohort. **(B)** Identification of cuproptosis-related genes with prognostic value.

understand the potential role of cuproptosis-related genes in CCRCC, GO and KEGG analyses were performed on co-expressed copper death-related genes.

2.4 Construction and validation of the 4-CRGs risk signature

We divided the training and testing cohorts into a ratio of 7:3 for patients with CCRCC. Clinical statistical analysis of the training and testing groups was performed using the chi-square test. In the training cohort, a univariate Cox regression analysis of cuproptosis-related genes was performed to identify the significant prognostically related genes. For significant prognosis-related cuproptosis-related genes, we used LASSO regression analysis to obtain independent prognostic genes in the training set. LASSO regression improves the accuracy and interpretability of the model and reduces the risk of overfitting (Tibshirani, 1997). Multivariate Cox regression analysis was conducted to obtain regression coefficients for independent prognostic genes. Finally, a 4-CRGs risk signature was established based on the multivariate Cox regression coefficient beta value, and the formula was as follows: risk score = EXPgene1∗ β1 + EXPgene2∗β2 + EXPgene3∗β3 + . . . + EXPgenen∗βn, where EXP is the expression level and β represents the regression coefficient from the multivariate Cox. In both cohort, by calculating the risk score for each sample, patients were divided into low- and high-risk groups using the median cut-off value. Furthermore, the KM curve was used to compare

TABLE 1 Characteristics of training, testing, and total cohort.

Clinical features	Type	Total (N = 518)	Testing cohort (N = 154)	Training cohort (N = 364)	*p*-value
Age	≤65	329 (63.51%)	99 (64.29%)	230 (63.19%)	0.8905
Age	>65	189 (36.49%)	55 (35.71%)	134 (36.81%)	
Stage	Stage I	261 (50.39%)	81 (52.6%)	180 (49.45%)	0.439
Stage	Stage II	58 (11.2%)	20 (12.99%)	38 (10.44%)	
Stage	Stage III	116 (22.39%)	34 (22.08%)	82 (22.53%)	
Stage	Stage IV	83 (16.02%)	19 (12.34%)	64 (17.58%)	
Gender	Female	175 (33.78%)	45 (29.22%)	130 (35.71%)	0.1846
Gender	Male	343 (66.22%)	109 (70.78%)	234 (64.29%)	
T	T1	267 (51.54%)	82 (53.25%)	185 (50.82%)	0.6336
T	T2	70 (13.51%)	24 (15.58%)	46 (12.64%)	
T	T3	170 (32.82%)	45 (29.22%)	125 (34.34%)	
T	T4	11 (2.12%)	3 (1.95%)	8 (2.2%)	
N	N0	230 (44.4%)	69 (44.81%)	161 (44.23%)	1
N	N1	15 (2.9%)	4 (2.6%)	11 (3.02%)	
N	Unknown	273 (52.7%)	81 (52.6%)	192 (52.75%)	
M	M0	414 (79.92%)	131 (85.06%)	283 (77.75%)	0.1085
M	M1	78 (15.06%)	17 (11.04%)	61 (16.76%)	
M	Unknown	26 (5.02%)	6 (3.9%)	20 (5.49%)	

the overall survival (OS) between the two groups using the log-rank test. A time-dependent ROC curve analysis was used to assess the predictive power of the 4-CRGs risk signature. Finally, we perform external validation with the external validation cohort GSE22451 and TCGA-LGG. In each independent external validation cohort, based on the risk score, patients were classified into two groups. Maximally selected rank statistics was applied by using an R package "survival", and "survminer" to identify the optimal cutting point to divide patients.

2.5 Construction of nomogram

We screened for prognostic predictive factors including clinical characteristics and risk scores. Specifically, the univariate Cox proportional hazard model was employed to analyze the correlation between the risk score and OS, and multivariate Cox regression analysis was used to evaluate whether the established risk score could serve as an independent prognostic predictor. Further, to comprehensively assess patient survival, we constructed a nomogram integrating distinct clinicopathological information, including age, stage, and risk score, using the "rms" package. Additionally, the decision curve analysis (DCA) of 1, 3, and 5 years was calculated to evaluate whether the synthetic nomogram we established was suitable for clinical application.

2.6 Immune function, immune infiltrates, immunomodulatory, and drugs

We used the CIBERSORT algorithm to assess the degree of infiltration of 22 immune cells in different CCRCC samples (Newman et al., 2015). Single-sample gene set enrichment analysis (ssGSEA) was applied to explore the different infiltration degrees of immune-related functions in different CCRCC samples of the TCGA database using the R package "GSVA".

TIMER is a website that can systematically analyze immune infiltration in various malignancies (https://timer.cistrome.org/) (Li et al., 2020). We investigated the relationship between gene expression and gene markers of TILs in CCRCC.

TISIDB (http://cis.hku.hk/TISIDB/index.php) was used to investigate the association of genes with immunostimulators in CCRCC.

The "pRRophetic" R package was used to predict the half-maximal inhibitory concentration (IC50) of some drugs in each sample regarding tumor treatment.

2.7 ceRNA network

We used RNAseq data from the TCGA database and miRNAseq data, including 541 KIRC samples and 72 normal samples. The difference analysis was performed using the DESeq2 package, and |logFC|>1 and adj. $p < 0.05$ were set as

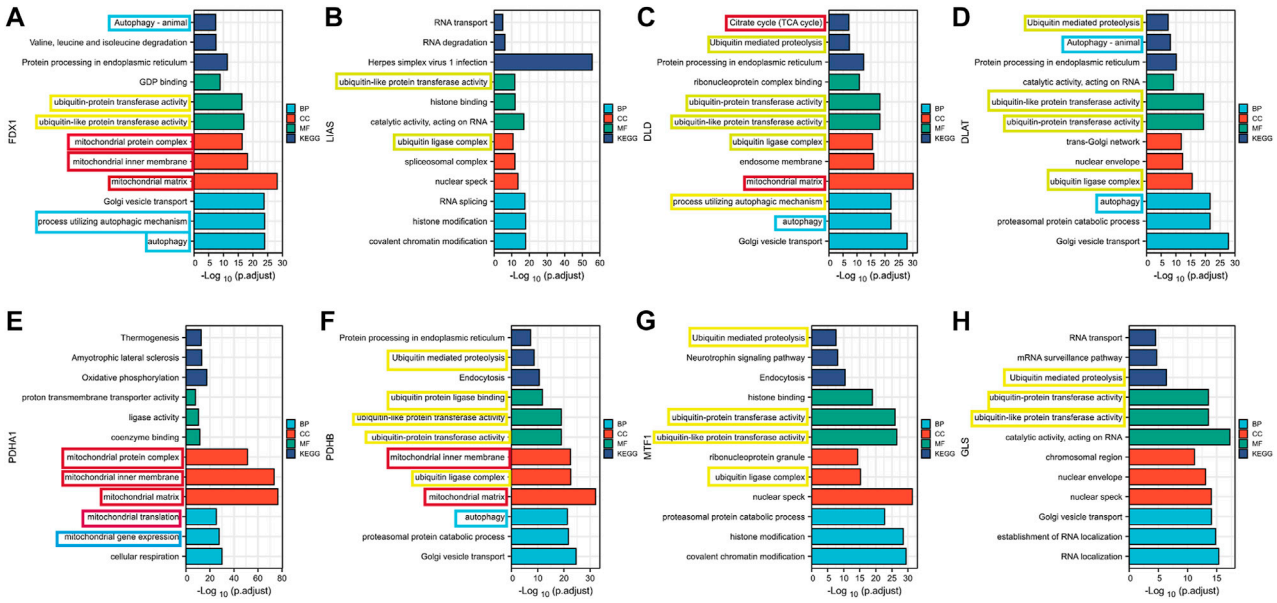


FIGURE 2
Results of GO and KEGG enrichment analysis of cuproptosis co-expression genes **(A)** FDX1; **(B)** LIAS **(C)** DLD; **(D)** DLAT; **(E)** PDHA1; **(F)** PDHB; **(G)** MTF1; **(H)** GLS.

thresholds to obtain the differential expression of lncRNAs (DElncRNAs) and miRNAs (DEmiRNAs) between CCRCC patients and normal patients. Subsequently, cuproptose-related miRNAs (CRMs) were predicted using the miRTarBase database (Huang et al., 2020). DEmiRNAs and CRMs were intersected to obtain cuproptosis-related DEmiRNAs (CRDEMs). Cuproptosis-related lncRNAs (CRLs) were predicted using the starBase database (Li et al., 2014). CRLs and DElncRNAs were intersected to obtain the cuproptosis-related DElncRNAs (CRDELs). Subsequently, we integrated the interactions between CRDEMs, CRDELs, and cuproptosis-related genes to construct a ceRNA regulatory network. Finally, Cytoscape (version 3.8.0) software was used to visualize the ceRNA regulatory network.

2.8 Statistical analysis

All statistical analyses were conducted using R version 4.1.2 software (https://www.r-project.org/). Univariate Cox hazard regression analyses were performed to identify the independent prognostic cuproptosis-related genes. Survival analysis was conducted by the Kaplan-Meier (K-M) method with the log-rank test. We also compared the expression of cuproptosis-related genes at different clinical stages by using the Wilcoxon rank-sum test.

3 Results

3.1 Identification of differentially expressed cuproptosis-related genes in normal and tumor samples

We compared the expression of cuproptosis-related genes between 541 CCRCC samples and 72 normal samples using the Wilcoxon rank-sum test in the TCGA cohort. We found that *LIAS* and *CDKN2A* were significantly upregulated in tumor samples, and *FDX1*, *DLD*, *DLAT*, *PDHA1*, *PDHB*, *MTF1*, and *GLS* were significantly downregulated in tumor samples (Figure 1A).

3.2 Identification of cuproptosis-related genes with prognostic value

To obtain reliable survival results for CCRCC, we first excluded samples with a survival time of less than 30 days. In total, 518 samples were obtained (Table 1). Nine differentially expressed cuproptosis-related genes (*FDX1*, *DLD*, *DLAT*, *PDHA1*, *PDHB*, *MTF1*, *GLS*, *LIAS*, and *CDKN2A*) were identified. Through the KM curve, we found that 9 selected genes all had an impact on the prognosis of CCRCC, including *FDX1* (hazard ratio, HR = 0.49; 95% confidence interval, 95% CI = 0.36–0.67; $p < 0.001$), *LIAS* (HR = 0.61; 95%CI =

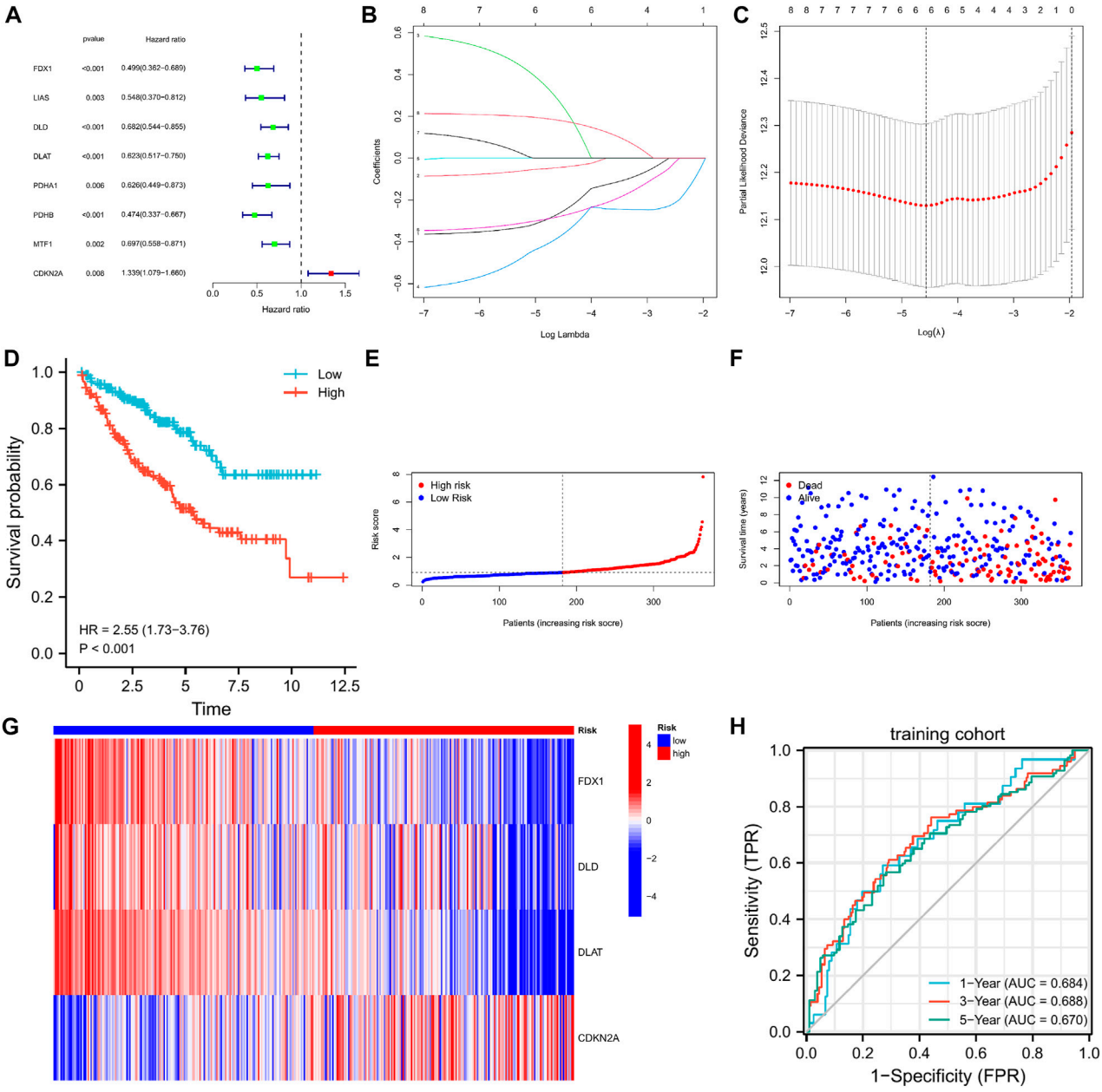


FIGURE 3
The 4-CRGs risk signature in the training cohort. **(A)** The CRGs with prognostic values were assessed by the univariate Cox proportional hazards regression model in the training cohort. **(B,C)** The selection of CRGs for risk signature by LASSO analysis in the training cohort. **(D)** K-M curves for OS in the training cohort. **(E–G)** The risk score, survival status, and heatmap of 4 CRGs in the training cohort. **(H)** Time-dependent ROC curves for OS in the training cohort.

0.45–0.83; $p = 0.002$), *DLD* (HR = 0.53; 95%CI = 0.39–0.72; $p < 0.001$), *DLAT* (HR = 0.46; 95%CI = 0.33–0.63; $p < 0.001$), *PDHA1* (HR = 0.63; 95%CI = 0.46–0.85; $p = 0.003$), *PDHB* (HR = 0.57; 95%CI = 0.42–0.78; $p < 0.001$), *MTF1* (HR = 0.59; 95%CI = 0.43–0.80; $p = 0.001$), *GLS* (HR = 0.65; 95%CI = 0.48–0.88; $p = 0.005$), and *CDKN2A* (HR = 1.49; 95%CI = 1.10–2.02; $p = 0.011$) (Figure 1B).

3.3 Clinicopathological features

We compared the expression of cuproptosis-related genes at different clinical stages using the Wilcoxon rank-sum test. *FDX1*, *LIAS*, *DLD*, *DLAT*, *PDHA1*, *PDHB*, *MTF1*, and *GLS* were highly expressed in T1 compared with T3, and *CDKN2A* was expressed at lower levels in T1 than in T3 (Supplementary Figure S1).

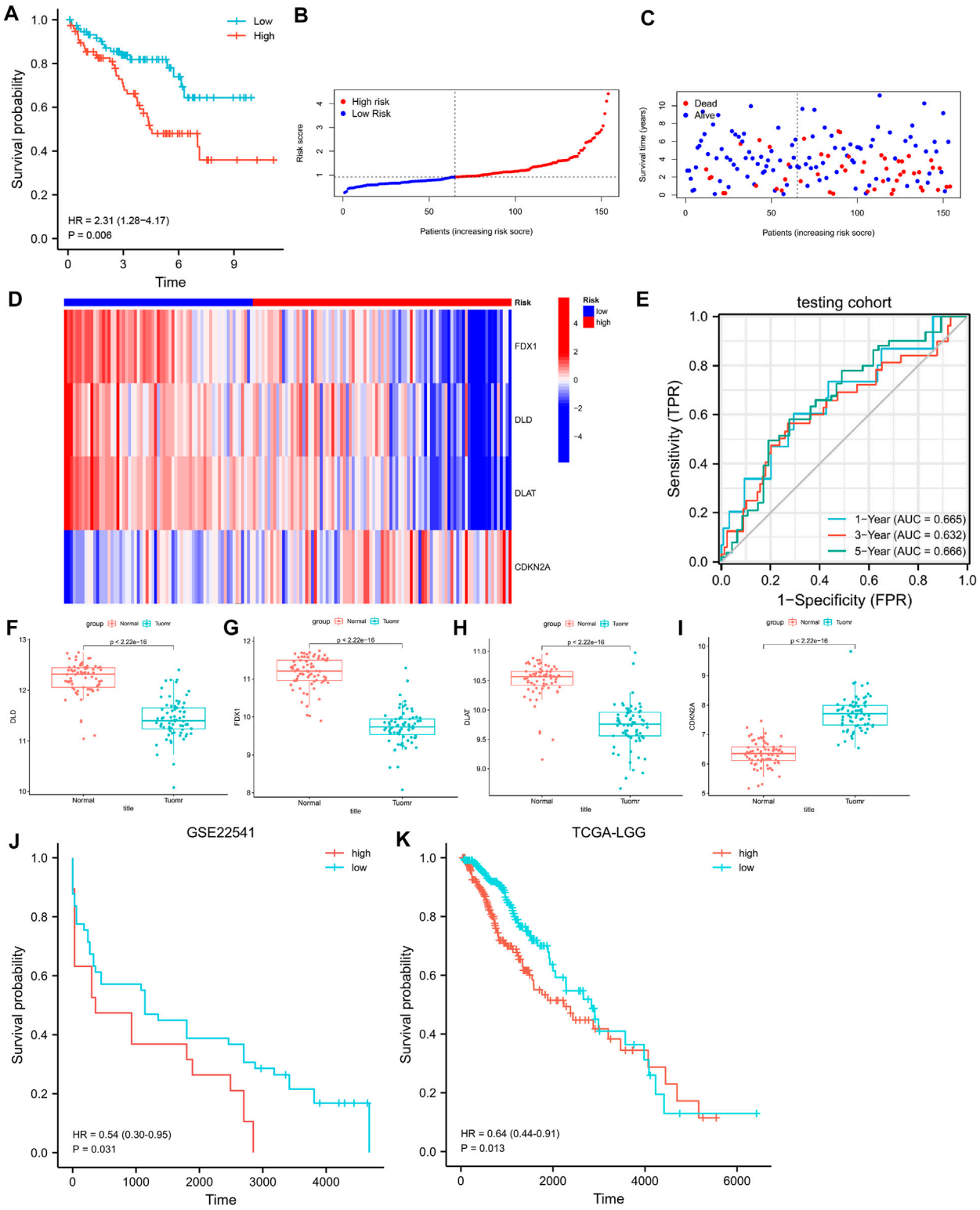


FIGURE 4
(A) K-M curves for OS in the testing cohort. **(B–D)** The risk score, survival status, and heatmap of 4 CRGs in the testing cohort. **(E)** Time-dependent ROC curves for OS in the testing cohort. **(F–I)** Differential expression of *FDX1*, *DLD*, *DLAT*, and *CDKN2A* in the GSE53757 dataset. **(J)** K-M curves for PFS in the GSE22541. **(K)** K-M curves for OS in the LGG.

Compared to N1, *FDX1* and *LIAS* were highly expressed in N0 (Supplementary Figure S2). *FDX1, LIAS, DLD, DLAT, PDHA1, PDHB, MTF1*, and *GLS* were highly expressed in M0 compared with M1, and *CDKN2A* was expressed at lower levels in M0 than in M1 (Supplementary Figure S3). *FDX1, LIAS, DLD, DLAT, PDHA1, PDHB*, and *MTF1* were highly expressed in Stage 1 compared to Stage 3 and 4, and *CDKN2A* had lower expression in Stage 1 compared to Stage 3 and 4 (Supplementary Figure S4).

3.4 GO and KEGG

For significant prognosis-related cuproptosis-related genes, we performed gene co-expression analysis in TCGA tumor patients and set the absolute value of the correlation coefficient to greater than 0.4 with a p-value less than 0.001 to obtain the co-expression genes. For co-expression genes, we performed GO and KEGG enrichment analyses and sorted them by p values (Figure 2). We found that co-expression genes were significantly enriched in the mitochondria during cell localization. The TCA cycle is thought to be associated with cancer progression, the site of biological processes in the mitochondria. This is consistent with Tsvetkov et al. (2022)'s view that Cu causes cell death by influencing the TCA cycle. The enrichment analysis results showed that co-expression of genes is correlated with autophagy and ubiquitin-mediated proteolysis, which provides a research direction for further exploration of the mechanism of cuproptosis.

3.5 Construction of the 4-CRGs risk signature

We found no significant differences between the training and testing cohorts in terms of age, sex, or TNM staging (Table 1). In the training cohort, univariate Cox regression analysis yielded eight cuproptosis-related genes that were significantly associated with prognosis (Figure 3A). Using lasso regression method, six optimal variables were obtained from the above 8 cuproptosis-prognostic-related gene (Figures 3B,C). By Cox regression analysis, the signature was finally established: risk score = EXP FDX1 * −0.499501220246694 + EXP DLD * −0.59322127824406 + EXP DLAT * −0.659153532219121 + EXP CDKN2A * 0.199116740963518. The KM curve showed that the prognosis of the high-risk group was worse than that of the low-risk group (Figure 3D, log-rank $p < 0.001$; HR = 2.55, 95%CI = 1.73–3.76). ROC curves were used to assess the accuracy of the established models in predicting overall survival (OS) in patients with CCRCC. As shown in Figure 3H, the AUC values at 1, 3, and 5 years were 0.684, 0.688, and 0.670, respectively, indicating the robustness and accuracy of the model in predicting patient prognosis.

3.6 Validation of the 4-CRGs risk signature and validation of differential expression of *FDX1*, *DLD*, *DLAT*, and *CDKN2A* in CCRCC

In the testing cohort, the KM curve showed that the prognosis of the high-risk group was worse than that of the low-risk group (Figure 4A, log-rank $p = 0.006$; HR = 2.31, 95% CI = 1.28–4.17). ROC curves were used to assess the accuracy of the established models in predicting OS in patients with CCRCC. As shown in Figure 4E, the AUC values at 1, 3, and 5 years were 0.665, 0.632, and 0.666, respectively, indicating the robustness and accuracy of the model in predicting patient prognosis.

We validated the differences in the expression of *FDX1*, *DLD*, *DLAT*, and *CDKN2A* between CCRCC and normal samples using GSE53757. The results showed that *FDX1*, *DLD*, and *DLAT* exhibited low expression in CCRCC, whereas *CDKN2A* was highly expressed in CCRCC (Figures 4F–I). This is consistent with the results obtained from the TCGA dataset. We further verify the above prediction method in external data cohorts" GSE22541" and "TCGA-LGG". In GSE22541 validation cohort, we divided the CCRCC patients into high-risk and low-risk groups based on the risk score. Survival comparison showed that low-risk group had significantly better prognosis outcome than high-risk group (Figure 4J). In addition, In TCGA-LGG validation cohort, we also found that based on the high and low-risk groups divided by risk score. Different groups have significantly different prognostic outcomes (Figure 4K). This demonstrates the generalization power of cuproptosis-related signature and has some value for the prediction of other cancers.

3.7 Nomogram and decision curve analysis

Univariate and multivariate Cox regression analyses showed that risk score, stage, and age were prognostic predictors of TGGA-KIRC (Figures 5A,B). Nomograms are widely used for the prognostic assessment of tumors. Various clinical features have prognostic value in clinical practice. Therefore, we established a nomogram containing multiple clinicopathological characteristics and risk scores. The scores for each variable were calculated and combined to predict the prognosis of patients with CCRCC (Figure 5C). DCA (Figure 5E) also proved that the nomogram combined with various clinical features had a better clinical application value.

3.8 Immune function and immune infiltrates

We used the CIBERSORT algorithm to estimate differences between 22 tumor-infiltrating immune cells between the low- and high-risk groups. Figure 6D shows that plasma cells, T cells CD8, T cells, regulatory T cells (Tregs), and activated NK cells

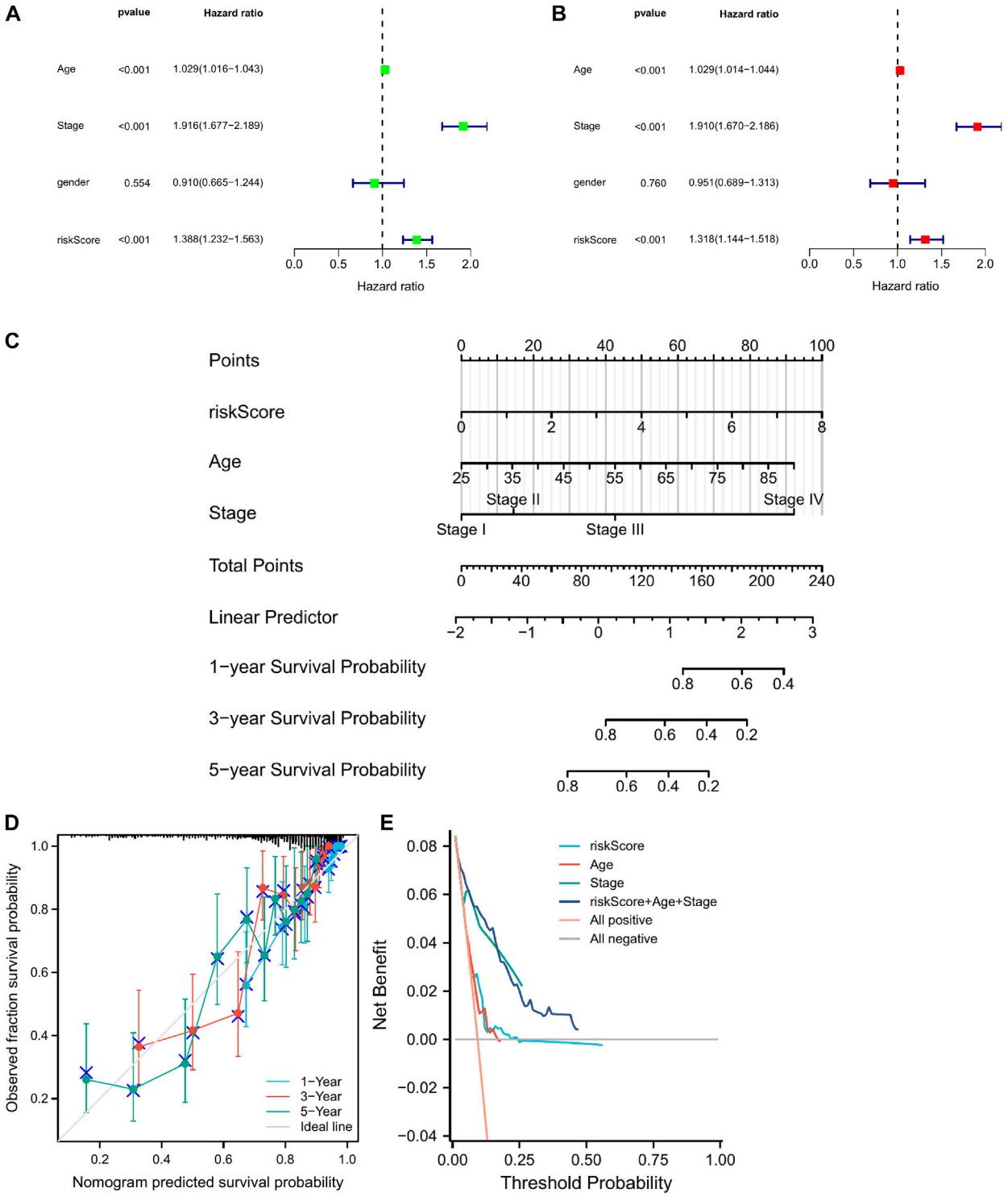


FIGURE 5
The construction of a nomogram for predicting survival. **(A,B)** Screening the independent predictors for OS in KIRC by univariate and multivariate Cox proportional hazards regression model. **(D)** A nomogram including risk score and clinicopathological features was constructed to predict 1/3/5-year OS. **(C)** The calibration plots for predicting 1/3/5-year OS are based on the CRGs nomogram. **(E)** Decision curve analysis (DCA) for the evaluation of the net benefits of CRGs and nomogram.

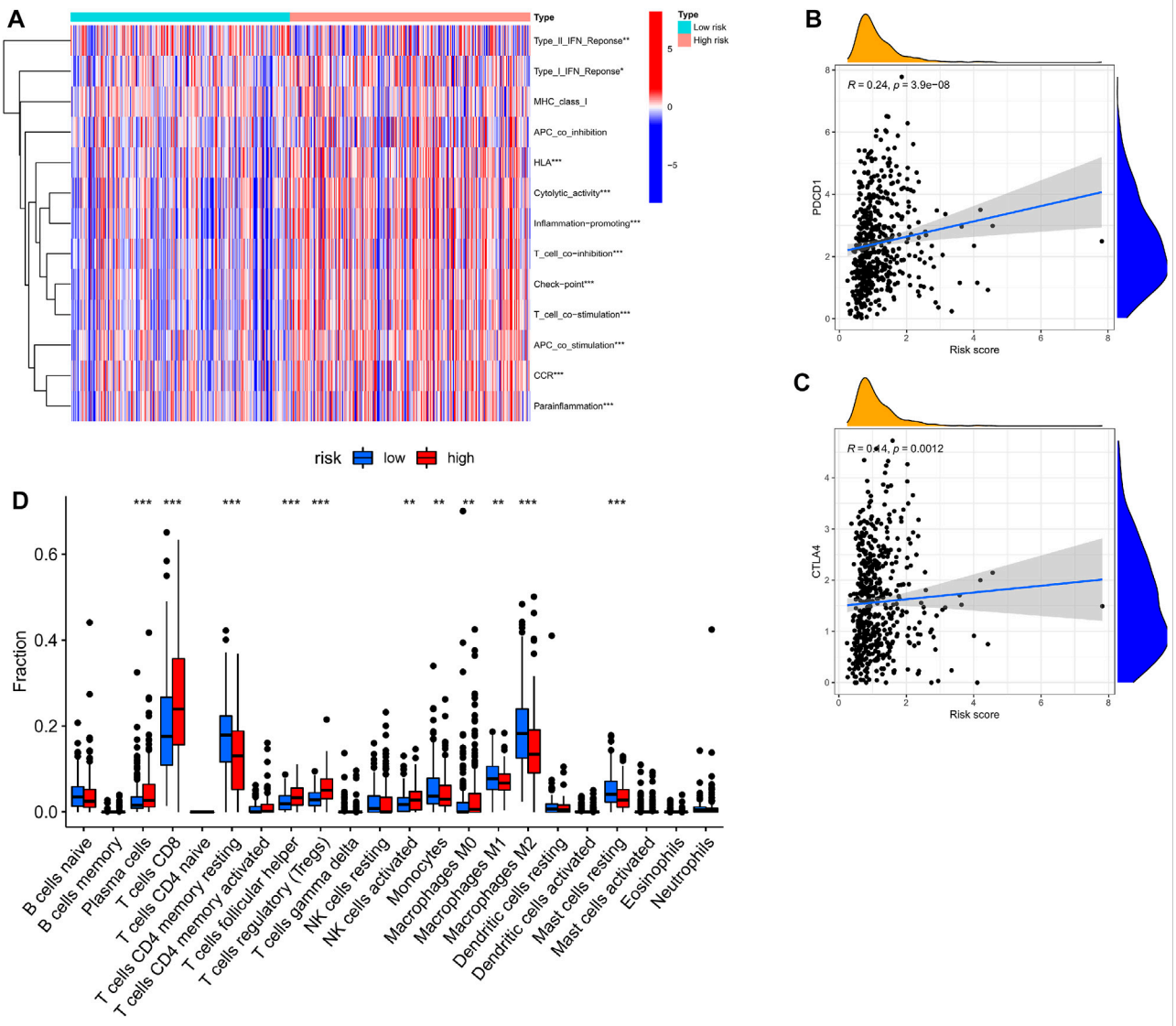


FIGURE 6
Correlation of CRGs signature with immunity. **(A)** Heatmap of the scores of immune-related functions between different risk groups. **(B,C)** The association between risk score and expression of PDCD1 or CTLA4 **(D)** Boxplots comparing the scores of immune cells between different risk groups. *$p < 0.05$; **$p < 0.01$, ***$p < 0.001$.

were more enriched in high-risk groups, while naive B cells, T cells, CD4 memory monocytes, macrophages M0, macrophages M1, and macrophages M2 were more enriched in the low-risk group. This indicates that there are differences in immune cell infiltration in different risk groups, suggesting that cuproptosis -related genes are closely related to immune cell infiltration.

The ssGSEA method was applied to KIRC patients in the high- and low-risk groups to assess the differences in immune function between the high- and low-risk groups. Figure 7A shows that Type-I-IFN-Response, HLA, Cytolytic activity, Inflammation-promoting, T-cell-co-inhibition, Checkpoint, T-cell-co-stimulation, APC-co-stimulation, CCR, and parainflammation were upregulated in the high-risk group, suggesting that cuproptosis-related genes are involved in immune regulation.

We further investigated the links between *FDX1*, *DLD*, *DLAT*, and *CDKN2A* and the TIL gene markers in the TIMER database (Supplementary Material S1). *DLD* was strongly correlated with STAT3 (rho = 0.419), STAT5B (rho = 0.47), and CD4 (rho = −0.401) (Supplementary Figure S5). There was a strong correlation between *DLAT* and TGFBR2 (rho = 0.49), STAT3 (rho = 0.501), AHR (rho = 0.449), STAT5B (rho = 0.571), MRC1 (rho = 0.437), CD7

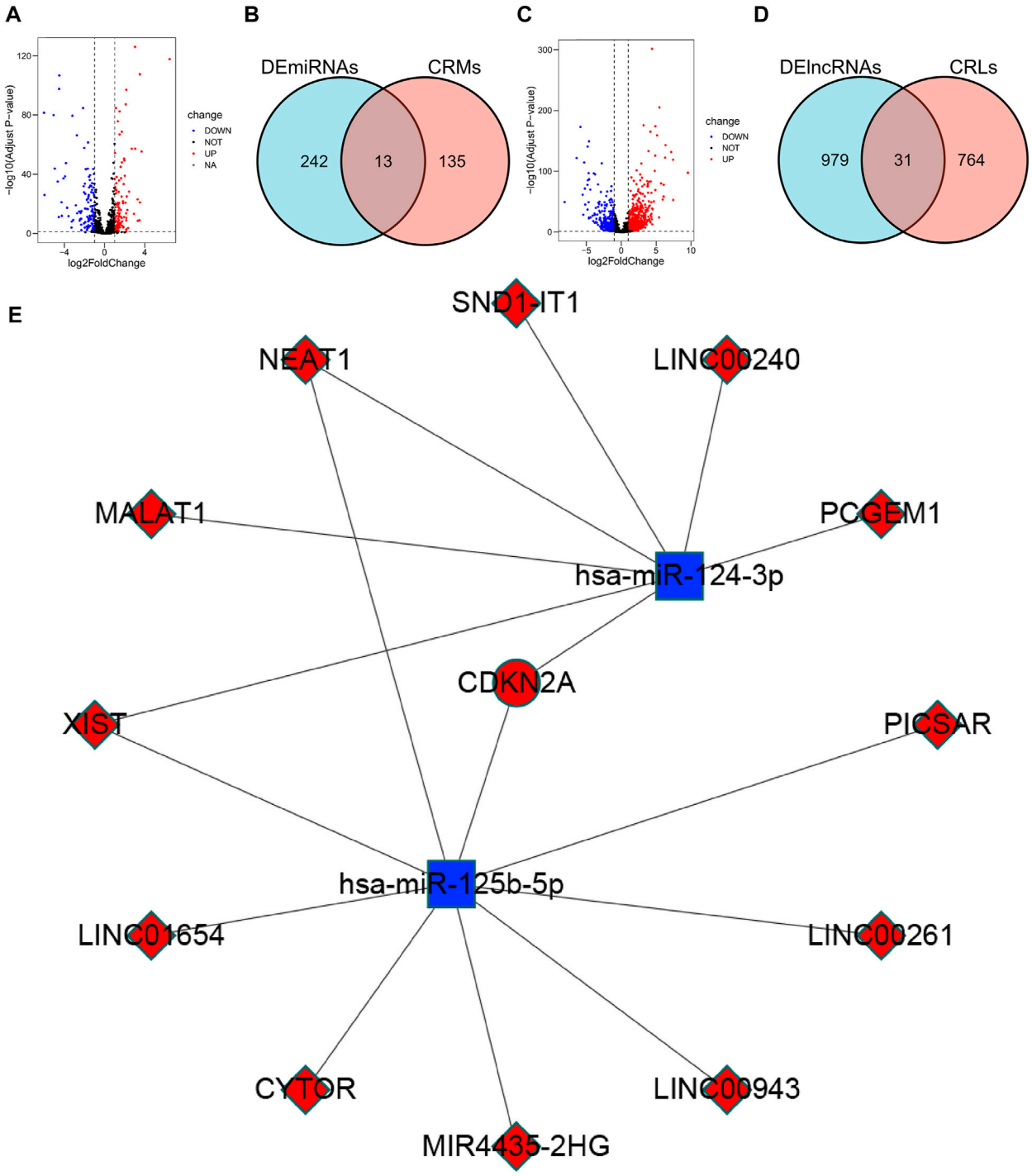


FIGURE 7
Construction of the ceRNA network. **(A)** Volcano plot of 255 differentially expressed miRNAs. **(B)** Venn diagram of the intersection of DEmiRNAs and CRMs. **(C)** Volcano plot of 1010 differentially expressed lncRNAs. **(D)** Venn diagram of the intersection of DElncRNAs and CRLs. **(E)** ceRNA network associated with cuproptosis: red indicates up-regulation in CCRCC, blue indicates down-regulation in CCRCC.

(rho = −0.417), and TGFBR2 (rho = 0.49) (Supplementary Figure S5). STAT3, STAT5B, and CD7 simultaneously show a strong correlation with *DLD* and *DLAT*, suggesting that STAT3, STAT5B, and CD7 may be associated with cuproptosis genes in an important link with levels of immune infiltration.

3.9 Immunomodulators and screening drugs

We explored the relationship between cuproptosis-related genes and immunomodulators associated with model building and found that *FDX1*, *DLD*, *DLAT*, *CDKN2A*, and immunostimulants were significantly associated. Specifically, high expression of *FDX1* was significantly associated with TNFRSF8 (Rho = −0.402), and high expression of *DLD* was significantly correlated with TNFRSF4 (Rho = −0.648), TNFRSF18 (Rho = −0.435), and TNFRSF25 (Rho = −0.6). *DLAT* was significantly correlated with TNFRSF4 (Rho = −0.522), TNFRSF8 (Rho = −0.436), TNFRSF18 (Rho = −0.557), TNFRSF25 (Rho = −0.642), CDKN2A, and TNFRSF18 (Rho = 0.398) (Supplementary Figure S6). Hence, *FDX1*, *DLD*, *DLAT*, and *CDKN2A* may play important roles in immune interactions and may be associated with tumor immune evasion.

We found that PDCD1 and CTLA4 (immune checkpoints) were elevated in patients with high-risk scores (Figures 6B,C). These discoveries suggested that patients of high-risk scores may be more sensitive to ICB therapy.

Finally, we predicted some drugs for tumor treatment using the pRRophetic R package and obtained some drugs that may show different sensitivities in patients in the high-risk and low-risk groups. Specifically, the low-risk group was more sensitive to AKT. inhibitor.VIII, AP.24534, AS601245, AUY922, axitinib, and AZ628, and the high-risk group were more sensitive to A.443654, ABT.888, AG.014699, AICAR, and AMG.706 (Supplementary Figure S7).

3.10 ceRNA

We analyzed DElncRNAs and DEmiRNAs between 541 KIRC samples and 72 normal samples and obtained a total of 255 DEmiRNAs, of which 129 were upregulated, 126 were downregulated; and 1010 DElncRNAs, of these, 777 were upregulated and 233 were downregulated. Using the miRTarBase database, 148 CRMs were identified. Thirteen CRDEMs were obtained by intersecting the CRMs with the DEmiRNAs. Subsequently, 795 CRLs were predicted using the starBase database, and 31 CRDELs were obtained by intersecting CRLs with DElncRNAs. Based on the existing theory that lncRNA inhibits the degradation of mRNA by miRNA through competitive binding of miRNA, we constructed a ceRNA network containing one mRNA, two miRNAs, and 12 lncRNAs (Figure 7E).

4 Discussion

We screened 10 copper ion carrier genes that are thought to be associated with cuproptosis. Nine out of 10 genes had differences in expression in CCRCC patients and non-tumor patients, and all 9 genes were valuable in assessing the prognosis of patients with CCRCC. We then looked for genes that were co-expressed with cuproptosis-related genes and found that cuproptosis-related genes may be associated with autophagy and ubiquitin-mediated proteolysis. Autophagy is associated with the survival of tumor cells but can either promote or inhibit apoptosis in different cellular contexts (Levy et al., 2017). Such context-dependent effects of autophagy are poorly understood; therefore, studying the relationship between apoptosis and autophagy may be a new research direction. Ubiquitin-mediated proteolysis is closely associated with cell proliferation. Studies have shown that the driving force of the cell cycle is the activation of cyclin-dependent kinases (CDKs), the activities of which are controlled by ubiquitin-mediated proteolysis of key regulators such as cyclins and CDK inhibitors (Nakayama and Nakayama, 2006). However, the link between cuproptosis and ubiquitin-mediated proteolysis needs to be experimentally confirmed. Using LASSO and multivariate Cox regression, we included four genes and constructed cuproptosis gene-related signatures containing *FDX1*, *DLD*, *DLAT*, and *CDKN2A*, of which *FDX1*, *DLAT*, and *CDKN2A* are correlated with CCRCC prognosis (Bian et al., 2022). Because *FDX1*, *DLD*, *DLAT*, and copper ion carriers are positively correlated, *CDKN2A* and copper ionophores are negatively correlated (Tsvetkov et al., 2022), which is consistent with the predictions. *FDX1* is a reductase that reduces cu^{2+} to cu^{1+} to promote cuproptosis (Tsvetkov et al., 2022). *FDX1* may modulate TP73 tumor suppressor through IRP2 to regulate tumor suppression (Zhang et al., 2020a), and *FDX1* may be a gene related to KIRC (Khouja et al., 2022). *DLD* is a homodimeric flavin-dependent enzyme that catalyzes NAD^+-dependent oxidation of dihydrolipoamide and participates in the TCA cycle to convert pyruvate to acetyl-CoA (Fleminger and Dayan, 2021). *DLD* may destroy cancer cells by producing ROS and by chelation with DNA (Dayan et al., 2019). *DLAT* is the subunit E2 of the PDC complex in the TCA cycle (Goh et al., 2015). *DLAT* may promote apoptosis by influencing energy production. *CDKN2A* encodes the tumor suppressors p15 INK4b and p16 INK4a to inhibit CDK4 and CDK6, which prevent pRB phosphorylation and block cell cycle progression (Hannou et al., 2015). *CDKN2A* mutations may play a role in renal cancer metastasis by influencing the expression of p16/p14 (Sun et al., 2021). In summary, cuproptosis-related genes may play important roles in CCRCC.

CCRCC is an immunogenic tumor whose tumor immune microenvironment has many different immune cells infiltrates with various immunomodulatory molecules, which may have a significant impact on the prognosis of patients, as well as the effect of immunotherapy (Díaz-Montero et al., 2020). The cytolytic activity index (CYT) in CCRCC is the highest among

18 human cancers (Rooney et al., 2015), and spontaneous regression in 1% of patients are considered immune-mediated (Janiszewska et al., 2013). $CD8^+$ T cells play an important role in tumor immunity, and their anti-tumor activity is the basis of ICI therapy (Şenbabaoğlu et al., 2016). Activated $CD8^+$ T cells have a significant positive effect on the prognosis of some tumor patients, such as those with early colon cancer (Pagès et al., 2005; Galon et al., 2006). However, for CCRCC, infiltration of $CD8^+$ T cells is associated with a high tumor grade and poor prognosis (Díaz-Montero et al., 2020). Our study showed that the high-risk group divided by the cuproptosis-related signature had a significant increase in the infiltration of $CD8^+$ T cells compared with the low-risk group, which is consistent with most studies. TILs that do not mediate anti-tumor function may be associated with Tregs (Díaz-Montero et al., 2020; Hah and Koo, 2021). We found that the Treg infiltration levels were significantly higher in the high-risk group than in the low-risk group. STAT5B, a marker gene for Tregs, is highly correlated with the *DLD* and *DLAT* genes involved in signature construction and may be the key gene mediating elevated Treg infiltration levels in the high-risk group. In addition, immune checkpoint molecules are also important factors that block $CD8^+$ T cells from exerting anti-tumor effects (Díaz-Montero et al., 2020). We found that both the most important immune checkpoint molecules, PDCD1 and CTLA-4, were positively correlated with the risk scores. In addition, we also found that the high-risk group had a higher TIDE value, which also suggested that the high-risk group was more likely to develop immune escape.

To make the cuproptosis-related signature more clinically relevant, we screened some of the drugs with different sensitivities in the high-risk and low-risk groups. Patients in the low-risk group had greater sensitivity to axitinib, an anti-VEGF-targeted drug used for the treatment of metastatic RCC (Hsieh et al., 2017).

The ceRNA hypothesis proposes that lncRNAs, as competing endogenous RNAs, regulate mRNA expression by competing for shared miRNAs (Karreth and Pandolfi, 2013). Specifically, upregulated lncRNA can competitively bind to miRNA, causing miRNA expression to be downregulated to inhibit the degradation of mRNA by miRNA and promote mRNA expression. We constructed a ceRNA network containing one upregulated mRNA, two down-regulated miRNAs, and 12 up-regulated lncRNAs. In the ceRNA network, we constructed CDKN2A, an upregulated mRNA whose high expression is thought to be associated with poor prognosis in CCRCC. Hsa-mir-124-3p, a downregulated miRNA predicted to bind to CDKN2A, is considered a key miRNA in CCRCC, inhibiting tumor migration, invasion, and proliferation (Butz et al., 2015). XIST, MALAT1, NEAT1, and LINC00240 up-regulated lncRNA were predicted to bind to hsa-mir-124-3p and promoted the proliferation and metastasis of other cancers by modulating hsa-mir-124-3p, but there are no related studies in CCRCC (Feng et al., 2016; Liu et al., 2018a; Xiao et al., 2019; Zhang et al., 2020b). In cervical cancer, Hsa-mir-125b-5p expression was downregulated, and CDKN2A expression was upregulated, suggesting that hsa-miR-125a-5p-CDKN2A is a possible ceRNA network (Wang et al., 2021). Hsa-mir-125b-5p was also found to be downregulated in bladder (Canturk et al., 2014) and prostate cancer (Lin et al., 2020). XIST, a lncRNA, has been shown to promote the progression of various cancers through its high expression (Liu et al., 2018b; Liu et al., 2019; Ning et al., 2021; Zheng et al., 2021). We found that the upregulated lncRNA XIST targets both hsa-mir-124-3p and hsa-miR-125b-5p and is positively correlated with CDKN2A. A XIST-hsa-mir-124-3p/hsa-miR-125b-5p-CDKN2A ceRNA network may exist in CCRCC and play an important role in its prognosis and development.

However, this study has some limitations and deficiencies. First, our study was retrospective, and prospective studies are needed to confirm these findings. Second, our conclusions were all obtained by data analysis and need to be further confirmed by experiments.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: https://www.ncbi.nlm.nih.gov/geo/, GSE53757.

Author contributions

PS, XJC, YX, and KZ were involved in conceptualization; PS, ZJS, and XYX were involved in data curation; PS, ML, and XJC analyzed the data; RW was involved in funding acquisition; GHF and were XJC involved in methodology; SJZ administrated the project; PS, XJC, and YX wrote the original draft; RW reviewed and edited the writing. All authors read and approved the final manuscript.

Funding

This research was supported by the fund for Natural Science Foundation of China (No. 81970051), Excellent Top Talent Cultivation Project of Anhui Higher Education Institutions (gxgwfx2021014), A sub-project of the Anhui Medical University National first-class undergraduate specialty construction program (clinical medicine), scientific research fund from Anhui medical university (2020xkj257), and the Applied Medical Research Project of

Hefei Health Commission (Hwk 2021zd008, Hwk 2022zd013).

Acknowledgments

We acknowledge the TCGA database for providing their platforms and contributors for uploading their meaningful datasets.

Conflict of interest

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

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.1061382/full#supplementary-material

References

Allensworth, J. L., Evans, M. K., Bertucci, F., Aldrich, A. J., Festa, R. A., Finetti, P., et al. (2015). Disulfiram (DSF) acts as a copper ionophore to induce copper-dependent oxidative stress and mediate anti-tumor efficacy in inflammatory breast cancer. *Mol. Oncol.* 9 (6), 1155–1168. doi:10.1016/j.molonc.2015.02.007

Bian, Z., Fan, R., and Xie, L. (2022). A novel cuproptosis-related prognostic gene signature and validation of differential expression in clear cell renal cell carcinoma. *Genes. (Basel)* 13 (5), 851. doi:10.3390/genes13050851

Butz, H., Szabo, P. M., Khella, H. W. Z., Nofech-Mozes, R., Patocs, A., and Yousef, G. M. (2015). miRNA-target network reveals miR-124as a key miRNA contributing to clear cell renal cell carcinoma aggressive behaviour by targeting CAV1 and FLOT1. *Oncotarget* 6 (14), 12543–12557. doi:10.18632/oncotarget.3815

Canturk, K. M., Ozdemir, M., Can, C., Öner, S., Emre, R., Aslan, H., et al. (2014). Investigation of key miRNAs and target genes in bladder cancer using miRNA profiling and bioinformatic tools. *Mol. Biol. Rep.* 41 (12), 8127–8135. doi:10.1007/s11033-014-3713-5

da Silva, D. A., De Luca, A., Squitti, R., Rongioletti, M., Rossi, L., Machado, C. M. L., et al. (2022). Copper in tumors and the use of copper-based compounds in cancer treatment. *J. Inorg. Biochem.* 226, 111634. doi:10.1016/j.jinorgbio.2021.111634

Dayan, A., Fleminger, G., and Ashur-Fabian, O. (2019). Targeting the Achilles' heel of cancer cells via integrin-mediated delivery of ROS-generating dihydrolipoamide dehydrogenase. *Oncogene* 38 (25), 5050–5061. doi:10.1038/s41388-019-0775-9

De Luca, A., Barile, A., Arciello, M., and Rossi, L. (2019). Copper homeostasis as target of both consolidated and innovative strategies of anti-tumor therapy. *J. Trace Elem. Med. Biol.* 55, 204–213. doi:10.1016/j.jtemb.2019.06.008

Denoyer, D., Masaldan, S., La Fontaine, S., and Cater, M. A. (2015). Targeting copper in cancer therapy: 'Copper that cancer. *Metallomics* 7 (11), 1459–1476. doi:10.1039/c5mt00149h

Díaz-Montero, C. M., Rini, B. I., and Finke, J. H. (2020). The immunology of renal cell carcinoma. *Nat. Rev. Nephrol.* 16 (12), 721–735. doi:10.1038/s41581-020-0316-3

Feng, T., Shao, F., Wu, Q., Zhang, X., Xu, D., Qian, K., et al. (2016). miR-124 downregulation leads to breast cancer progression via LncRNA-MALAT1 regulation and CDK4/E2F1 signal activation. *Oncotarget* 7 (13), 16205–16216. doi:10.18632/oncotarget.7578

Fleminger, G., and Dayan, A. (2021). The moonlighting activities of dihydrolipoamide dehydrogenase: Biotechnological and biomedical applications. *J. Mol. Recognit.* 34 (11), e2924. doi:10.1002/jmr.2924

Galon, J., Costes, A., Sanchez-Cabo, F., Kirilovsky, A., Mlecnik, B., Lagorce-Pages, C., et al. (2006). Type, density, and location of immune cells within human colorectal tumors predict clinical outcome. *Science* 313 (5795), 1960–1964. doi:10.1126/science.1129139

Goh, W. Q., Ow, G. S., Kuznetsov, V. A., Chong, S., and Lim, Y. P. (2015). DLAT subunit of the pyruvate dehydrogenase complex is upregulated in gastric cancer-implications in cancer therapy. *Am. J. Transl. Res.* 7 (6), 1140–1151.

Hah, Y. S., and Koo, K. C. (2021). Immunology and immunotherapeutic approaches for advanced renal cell carcinoma: A comprehensive review. *Int. J. Mol. Sci.* 22 (9), 4452. doi:10.3390/ijms22094452

Hannou, S. A., Wouters, K., Paumelle, R., and Staels, B. (2015). Functional genomics of the CDKN2A/B locus in cardiovascular and metabolic disease: What have we learned from GWASs? *Trends Endocrinol. Metab.* 26 (4), 176–184. doi:10.1016/j.tem.2015.01.008

Hsieh, J. J., Purdue, M. P., Signoretti, S., Swanton, C., Albiges, L., Schmidinger, M., et al. (2017). Renal cell carcinoma. *Nat. Rev. Dis. Prim.* 3, 17009. doi:10.1038/nrdp.2017.9

Huang, H. Y., miRTarBaseLi, J., Huang, K. Y., Shrestha, S., Hong, H. C., et al. (2020). miRTarBase 2020: updates to the experimentally validated microRNA-target interaction database. *Nucleic Acids Res.* 48 (1), D148–D154. doi:10.1093/nar/gkz896

Janiszewska, A. D., Poletajew, S., and Wasiutyński, A. (2013). Spontaneous regression of renal cell carcinoma. *Contemp. Oncol.* 17 (2), 123–127. doi:10.5114/wo.2013.34613

Karreth, F. A., and Pandolfi, P. P. (2013). ceRNA cross-talk in cancer: when ce-bling rivalries go awry. *Cancer Discov.* 3 (10), 1113–1121. doi:10.1158/2159-8290.CD-13-0202

Khouja, H. I., Ashankyty, I. M., Bajrai, L. H., Kumar, P. K. P., Kamal, M. A., Firoz, A., et al. (2022). Multi-staged gene expression profiling reveals potential genes and the critical pathways in kidney cancer. *Sci. Rep.* 12 (1), 7240. doi:10.1038/s41598-022-11143-6

Levy, J. M. M., Towers, C. G., and Thorburn, A. (2017). Targeting autophagy in cancer. *Nat. Rev. Cancer* 17 (9), 528–542. doi:10.1038/nrc.2017.53

Li, J. H., Liu, S., Zhou, H., Qu, L. H., and Yang, J. H. (2014). starBase v2.0: decoding miRNA-ceRNA, miRNA-ncRNA and protein-RNA interaction networks from large-scale CLIP-Seq data. *Nucleic Acids Res.* 42, D92–D97. doi:10.1093/nar/gkt1248

Li, T., Fu, J., Zeng, Z., Cohen, D., Li, J., Chen, Q., et al. (2020). TIMER2.0 for analysis of tumor-infiltrating immune cells. *Nucleic Acids Res.* 48 (1), W509–W514. doi:10.1093/nar/gkaa407

Lin, Y., Miao, Z., Zhang, X., Wei, X., Hou, J., Huang, Y., et al. (2020). Identification of key MicroRNAs and mechanisms in prostate cancer evolution based on biomarker prioritization model and carcinogenic survey. *Front. Genet.* 11, 596826. doi:10.3389/fgene.2020.596826

Liu, H., Deng, H., Zhao, Y., and Liang, Y. (2018). LncRNA XIST/miR-34a axis modulates the cell proliferation and tumor growth of thyroid cancer through MET-PI3K-AKT signaling. *J. Exp. Clin. Cancer Res.* 37 (1), 279. doi:10.1186/s13046-018-0950-9

Liu, J., Yao, L., Zhang, M., Jiang, J., Yang, M., and Wang, Y. (2019). Downregulation of LncRNA-XIST inhibited development of non-small cell lung cancer by activating miR-335/SOD2/ROS signal pathway mediated pyroptotic cell death. *Aging (Albany NY)* 11 (18), 7830–7846. doi:10.18632/aging.102291

Liu, X., Liang, Y., Song, R., Yang, G., Han, J., Lan, Y., et al. (2018). Long non-coding RNA NEAT1-modulated abnormal lipolysis via ATGL drives hepatocellular carcinoma proliferation. *Mol. Cancer* 17 (1), 90. doi:10.1186/s12943-018-0838-5

Motzer, R. J., Escudier, B., McDermott, D. F., George, S., Hammers, H. J., Srinivas, S., et al. (2015). Nivolumab versus everolimus in advanced renal-cell carcinoma. *N. Engl. J. Med.* 373 (19), 1803–1813. doi:10.1056/NEJMoa1510665

Motzer, R. J., Penkov, K., Haanen, J., Rini, B., Albiges, L., Campbell, M. T., et al. (2019). Avelumab plus axitinib versus sunitinib for advanced renal-cell carcinoma. *N. Engl. J. Med.* 380 (12), 1103–1115. doi:10.1056/NEJMoa1816047

Nabi, S., Kessler, E. R., Bernard, B., Flaig, T. W., and Lam, E. T. (2018). Renal cell carcinoma: A review of biology and pathophysiology. *F1000Res.* 7, 307. doi:10.12688/f1000research.13179.1

Nakayama, K. I., and Nakayama, K. (2006). Ubiquitin ligases: Cell-cycle control and cancer. *Nat. Rev. Cancer* 6 (5), 369–381. doi:10.1038/nrc1881

Nechushtan, H., Hamamreh, Y., Nidal, S., Gotfried, M., Baron, A., Shalev, Y. I., et al. (2015). A phase IIb trial assessing the addition of disulfiram to chemotherapy for the treatment of metastatic non-small cell lung cancer. *Oncologist* 20 (4), 366–367. doi:10.1634/theoncologist.2014-0424

Newman, A. M., Liu, C. L., Green, M. R., Gentles, A. J., Feng, W., Xu, Y., et al. (2015). Robust enumeration of cell subsets from tissue expression profiles. *Nat. Methods* 12 (5), 453–457. doi:10.1038/nmeth.3337

Ning, D., Chen, J., Du, P., Liu, Q., Cheng, Q., Li, X., et al. (2021). The crosstalk network of XIST/miR-424-5p/OGT mediates RAF1 glycosylation and participates in the progression of liver cancer. *Liver Int.* 41 (8), 1933–1944. doi:10.1111/liv.14904

Pagès, F., Berger, A., Camus, M., Sanchez-Cabo, F., Costes, A., Molidor, R., et al. (2005). Effector memory T cells, early metastasis, and survival in colorectal cancer. *N. Engl. J. Med.* 353 (25), 2654–2666. doi:10.1056/NEJMoa051424

Patard, J.-J., Leray, E., Rodriguez, A., Rioux-Leclercq, N., Guille, F., and Lobel, B. (2003). Correlation between symptom graduation, tumor characteristics and survival in renal cell carcinoma. *Eur. Urol.* 44 (2), 226–232. doi:10.1016/s0302-2838(03)00216-1

Rini, B. I., Plimack, E. R., Stus, V., Gafanov, R., Hawkins, R., Nosov, D., et al. (2019). Pembrolizumab plus axitinib versus sunitinib for advanced renal-cell carcinoma. *N. Engl. J. Med.* 380 (12), 1116–1127. doi:10.1056/NEJMoa1816714

Rooney, M. S., Shukla, S. A., Wu, C. J., Getz, G., and Hacohen, N. (2015). Molecular and genetic properties of tumors associated with local immune cytolytic activity. *Cell.* 160 (1-2), 48–61. doi:10.1016/j.cell.2014.12.033

Safi, R., Nelson, E. R., Chitneni, S. K., Franz, K. J., George, D. J., Zalutsky, M. R., et al. (2014). Copper signaling axis as a target for prostate cancer therapeutics. *Cancer Res.* 74 (20), 5819–5831. doi:10.1158/0008-5472.CAN-13-3527

Şenbabaoğlu, Y., Gejman, R. S., Winer, A. G., Liu, M., Van Allen, E. M., de Velasco, G., et al. (2016). Erratum to: Tumor immune microenvironment characterization in clear cell renal cell carcinoma identifies prognostic and immunotherapeutically relevant messenger RNA signatures. *Genome Biol.* 17 (1), 46. doi:10.1186/s13059-017-1180-8

Shanbhag, V. C., Gudekar, N., Jasmer, K., Papageorgiou, C., Singh, K., and Petris, M. J. (2021). Copper metabolism as a unique vulnerability in cancer. *Biochim. Biophys. Acta. Mol. Cell. Res.* 1868 (2), 118893. doi:10.1016/j.bbamcr.2020.118893

Siegel, R. L., Miller, K. D., and Jemal, A. (2020). Cancer statistics. *Ca. Cancer J. Clin.* 70 (1), 7–30. doi:10.3322/caac.21590

Sonpavde, G., Choueiri, T. K., Escudier, B., Ficarra, V., Hutson, T. E., Mulders, P. F., et al. (2012). Sequencing of agents for metastatic renal cell carcinoma: Can we customize therapy? *Eur. Urol.* 61 (2), 307–316. doi:10.1016/j.eururo.2011.10.032

Sun, Q., Chen, S., Hou, Y., Wen, X., Teng, X., Zhang, H., et al. (2021). Mutant CDKN2A regulates P16/p14 expression by alternative splicing in renal cell carcinoma metastasis. *Pathol. Res. Pract.* 223, 153453. doi:10.1016/j.prp.2021.153453

Tibshirani, R. (1997). The lasso method for variable selection in the Cox model. *Stat. Med.* 16 (4), 385–395. doi:10.1002/(sici)1097-0258(19970228)16:4<385::aid-sim380>3.0.co;2-3

Tsvetkov, P., Coy, S., Petrova, B., Dreishpoon, M., Verma, A., Abdusamad, M., et al. (2022). Copper induces cell death by targeting lipoylated TCA cycle proteins. *Science* 375 (6586), 1254–1261. doi:10.1126/science.abf0529

von Roemeling, C. A., Radisky, D. C., Marlow, L. A., Cooper, S. J., Grebe, S. K., Anastasiadis, P. Z., et al. (2014). Neuronal pentraxin 2 supports clear cell renal cell carcinoma by activating the AMPA-selective glutamate receptor-4. *Cancer Res.* 74 (17), 4796–4810. doi:10.1158/0008-5472.CAN-14-0210

Wang, T., Zhang, X. D., and Hua, K. Q. (2021). A ceRNA network of BBOX1-AS1-hsa-miR-125b-5p/hsa-miR-125a-5p-CDKN2A shows prognostic value in cervical cancer. *Taiwan. J. Obstet. Gynecol.* 60 (2), 253–261. doi:10.1016/j.tjog.2020.12.006

Xiao, D., Cui, X., and Wang, X. (2019). Long noncoding RNA XIST increases the aggressiveness of laryngeal squamous cell carcinoma by regulating miR-124-3p/EZH2. *Exp. Cell. Res.* 381 (2), 172–178. doi:10.1016/j.yexcr.2019.04.034

Zhang, H., Chen, D., Ringler, J., Chen, W., Cui, Q. C., Ethier, S. P., et al. (2010). Disulfiram treatment facilitates phosphoinositide 3-kinase inhibition in human breast cancer cells *in vitro* and *in vivo*. *Cancer Res.* 70 (10), 3996–4004. doi:10.1158/0008-5472.CAN-09-3752

Zhang, J., Kong, X., Zhang, Y., Sun, W., Wang, J., Chen, M., et al. (2020). FDXR regulates TP73 tumor suppressor via IRP2 to modulate aging and tumor suppression. *J. Pathol.* 251 (3), 284–296. doi:10.1002/path.5451

Zhang, Y., Li, X., Zhang, J., and Liang, H. (2020). Natural killer T cell cytotoxic activity in cervical cancer is facilitated by the LINC00240/microRNA-124-3p/STAT3/MICA axis. *Cancer Lett.* 474, 63–73. doi:10.1016/j.canlet.2019.12.038

Zheng, H., Zhang, M., Ke, X., Deng, X., Li, D., Wang, Q., et al. (2021). LncRNA XIST/miR-137 axis strengthens chemo-resistance and glycolysis of colorectal cancer cells by hindering transformation from PKM2 to PKM1. *Cancer Biomark.* 30 (4), 395–406. doi:10.3233/CBM-201740

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 07 December 2022
DOI 10.3389/fgene.2022.1088943

OPEN ACCESS

EDITED BY
Duo Liu,
Harbin Medical University Cancer Hospital, China

REVIEWED BY
Yijun Shen,
Fudan University, China
Yu'e Liu,
Tongji University, China
Liangqian Huang,
Nanjing University, China

*CORRESPONDENCE
Zi Ye,
yeziclover@163.com
Zhaohui Li,
zhaohuili202104@163.com

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

RECEIVED 03 November 2022
ACCEPTED 25 November 2022
PUBLISHED 07 December 2022

CITATION
Li H, Gao L, Du J, Ma T, Ye Z and Li Z (2022), Differentially expressed gene profiles and associated ceRNA network in ATG7-Deficient lens epithelial cells under oxidative stress.
Front. Genet. 13:1088943.
doi: 10.3389/fgene.2022.1088943

Differentially expressed gene profiles and associated ceRNA network in ATG7-Deficient lens epithelial cells under oxidative stress

Hongyu Li[1,2], Lixiong Gao[2], Jinlin Du[2], Tianju Ma[2], Zi Ye[2]* and Zhaohui Li[1,2]*

[1]Medical School of Chinese PLA, Beijing, China, [2]Senior Department of Ophthalmology, The Third Medical Center of PLA General Hospital, Beijing, China

Oxidation is an essential factor during cataract development. Autophagy, usually a cytoprotective process, is always found elevated in lens epithelial cells under oxidation, yet its roles and associated molecular mechanisms under such circumstances are rarely elucidated. Herein, we extracted and re-analyzed the RNA sequencing data of the GSE161701 dataset from the Gene Expression Omnibus database to identify the differentially expressed mRNAs and lncRNAs by using the R package "DESeq2". Further analyses of gene ontology and KEGG enrichment were implemented *via* the packages "clusterProfiler" and "enrichplot". We found that after the knockout of *ATG7*, differentially expressed genes were more associated with hemopoiesis, vasculature development, axonogenesis, and hypoxia regulation. When stimulated with H_2O_2, LECs displayed a gene expression profile correlating with apoptotic and proliferative pathways, such as the MAPK signaling pathway and FoxO signaling pathway. The differentially expressed gene profiles of the two types of LECs (wild type and ATG7 deficient) under oxidation were distinct to a large extent. Furthermore, 1,341 up-regulated and 1912 down-regulated differential mRNAs and 263 up-regulated and 336 down-regulated differential lncRNAs between these two types of LECs subjected to H_2O_2 were detected, among which 292 mRNAs and 24 lncRNAs possibly interacted with ten cataract-related miRNAs. A competing endogenous lncRNA-miRNA-mRNA network based on such interactions was finally constructed.

Abbreviations: ARC, age-related cataract; ceRNA, competing endogenouse RNA; DEGs, differentially expressed genes; DE-lncRNAs, differentially expressed lncRNAs; DE-mRNAs, differentially expressed mRNAs; GEO, Gene Expression Omnibus; GGH, glutathione; GO, gene ontology; GSEA, gene set enrichment analysis; HLE-B3, human lens epithelial B3; H2O2, hydrogen peroxide; KEGG, Kyoto Encyclopedia of Genes and Genomes; KO, knockout; LECs, lens epithelial cells; lncRNAs, long non-coding RNAs; MREs, microRNA response elements; ncRNAs, non-coding RNAs; PE, phosphatidylethanolamine; PPI, protein-protein interaction; ROS, reactive oxygen species; rRNAs, ribosomal RNAs; tRNAs, transfer RNAs; WT, wild type.

KEYWORDS

cataract, oxidative stress, autophagy, lens epithelial cells, differentially expressed genes (DEG), competing endogenous RNAs (ceRNA)

Background

Cataract, a common cause of vision loss, results from the gradual opacification of the lens, mostly at an elderly age. It has been recognized that risk factors such as aging, diabetes mellitus, ultraviolet B exposure, long-term use of corticosteroids, and smoking can contribute to the development of cataract. (Truscott, 2005). Physiologically, the lens epithelial cells (LECs) residing under the anterior capsule of the lens continuously proliferate and differentiate into the elongated fiber cells to form the compacted nuclear mass of the lens throughout its lifespan. These normal activities of LECs guarantee the homeostasis and transparency of the lens. And any disturbance to the LECs activities can result in cataractogenesis. A marked reduction of the lens epithelial cell density is observed in advanced senile cataract. (Tseng et al., 1994). With the aging of the lens, the LECs demonstrate a remarked increase of apoptosis in a time-dependent manner, and the experimentally induced apoptosis of LECs can give rise to the development of cataracts. (Li et al., 1995; Zhang et al., 2010). Moreover, there is a relatively high level of apoptosis of LECs in diabetes-induced cataracts. (Takamura et al., 2003; Xie et al., 2022). However, the molecular mechanisms of cataract development are not fully elucidated.

Oxidative stress is a prominent and critical factor for cataractogenesis. Generally, the metabolically produced reactive oxygen species (ROS), such as hydrogen peroxide (H_2O_2) and hydroxyl radical, can be reduced by the antioxidants like reduced glutathione (GGH), thus maintaining a stable and balanced redox environment in LECs. (Giblin, 2000). Any disturbance to this balance, whether due to aging, ultraviolet exposure, or some other factors, will result in redundancy of free radicals, which in turn contribute to cataract development. (Yildirim et al., 2009; Pescosolido et al., 2016; Hsueh and Chen, 2022). The lens proteins, such as crystallins, are found oxidized and aggregated in the nucleus of the lens and therefore scatter the light during the development of cataract. (Cobb and Petrash, 2002; Vetter et al., 2020; Hanafy and Cave, 2021). Of note, the most metabolically-active LECs are more susceptible to oxidative radicals. Excessive ROS can incur the apoptosis of LECs by targeting the DNA, membrane proteins, and many other constitutive components (Long et al., 2004; Cui et al., 2012), while antioxidants administration can ameliorate the H_2O_2-induced apoptosis of LECs(Zhou et al., 2016; Bai et al., 2017). Nevertheless, the exact and comprehensive molecular mechanisms of LECs apoptosis and the ensuing cataract under oxidative stress are, to our knowledge, not completely elaborated.

Autophagy (here referred specifically to macroautophagy) is an evolutionally conserved, catabolic process across a variety of species, which can break down the dysfunctional or unneeded macromolecules and membrane-coated organelles in the cytoplasm to recycle the necessary building substrates. It entails the formation of autophagosomes, their fusion with lysosomes, and the degradation of the inside components, each phase involving certain critical proteins, such as BECN1, LC3B, and ATG7. Although it is controversial whether autophagy is involved in the formation of the organelle-free zone during lens development (Matsui et al., 2006; Nishida et al., 2009; Morishita et al., 2013; Tu et al., 2021; Gheyas et al., 2022), the development of cataract is closely associated with dysfunctional autophagy. (Morishita, Eguchi, Kimura, Sasaki, Sakamaki, Robinson, Sasaki and Mizushima, 2013; Ping et al., 2021). A recent study revealed that rapamycin-induced autophagy can alleviate the level of ROS in LECs cultured in high glucose. (Liu J. et al., 2020). Furthermore, an elevated autography accompanied reduced apoptosis in H_2O_2-treated LECs(Han et al., 2021), and the overexpression of *ATG4a* in LECs can mitigate the apoptosis of cells. (Yan et al., 2020). It may suggest a cytoprotective role of autophagy toward LECs under oxidative stress by regulating cell apoptosis, which is quite contradictory to the findings of Huang J et al., whose study revealed autophagy-facilitated apoptosis. (Huang et al., 2022). Therefore, the detailed molecular mechanisms concerning the roles of autophagy in oxidative stress need further explored.

Non-coding RNAs (ncRNAs) refer to a considerable amount of transcriptomes without protein-coding function, which can be simply grouped into long non-coding RNAs (lncRNAs), small RNAs like microRNAs (miRNAs), transfer RNAs (tRNAs) and ribosomal RNAs (rRNAs) based on their length and function. Over the past few years, ncRNAs, especially lncRNAs and miRNAs, are found to play a vital role in the development of a variety of diseases, including cataracts. Recent studies showed that the elevated miR-23b-3p and miR-34a expressions could promote the apoptosis of LECs under the oxidative stress state while overexpression of miR-124 reduced the apoptosis of H_2O_2-treated LECs. (Fan et al., 2017; Gu, 2018; Zhou et al., 2019; Zhang et al., 2020). Moreover, lncRNA-H19 was up-regulated in cataract tissue and its knockdown could accelerate apoptosis of LECs under oxidative stress, and lncRNA TUG1 can promote the apoptosis of H_2O_2-treated LECs *via* targeting the miR-196a-5p. (Liu et al., 2018; Shen and Zhou, 2021). Although there is growing evidence of ncRNAs involved in pathogenic processes during cataract development, the comprehensive profile of ncRNAs and the interaction network concerning cataractogenesis have not been reported.

Autophagy requires the elongation of the phagophore at the beginning, during which ATG7 is critical for the formation and activation of two important complexes, ATG12-ATG5 complex and LC3-PE complex (Rubinsztein et al., 2012). Ablation of ATG7 is thought to completely block the elongation of the phagophore, thus abolishing the autophagy activity. Herein, we selected the LECs knocked out of *ATG7* to abolish the autophagy process during oxidative stress, and conducted detailed analyses of RNA sequencing data from the Gene Expression Omnibus (GEO) database to explore the expression profiles of mRNAs and lncRNAs of LECs under the circumstance of oxidative stress and to try to construct the possible molecular network of ATG7-associated cataractogenic mechanisms, hoping to offer some clues to the studies regarding the prevention and/or treatment of cataracts.

Materials and methods

RNA sequencing

Total RNA of human lens epithelial B3 (HLE-B3) cells treated with or without H_2O_2 for 12 h was extracted by using TRIzol (Thermo, United States) according to the manufactory's instructions. After being constructed by using NEBNext® UltraTM RNA Library Prep Kit for Illumina® (NEB, United States) and assessed by the Agilent Bioanalyzer 2,100 system, cDNA libraries were loaded on the Illumina NovaSeq 6,000 platform (Illumina, United States). Each group was sequenced in triplicate.

Data collection

Raw counts data of genes were extracted from the RNA-seq dataset of GSE161701 in the NCBI's GEO database (*https://www.ncbi.nlm.nih.gov/geo/*). (Huang, Yu, He, He, Yang, Chen and Han, 2022) Specifically, cultured HLE-B3 cells with or without the knockout of *the ATG7* gene were respectively treated with or without 200 μM H_2O_2 for 12 h, thus making four groups of samples for subsequent RNA sequencing (i.e., wild type cells for 0 h and 12 h of stimulation (WT-0h and WT-12 h), and *ATG7* knockout cells for 0 h and 12 h of stimulation (KO-0h and KO-12 h), each group with three biological replicates).

GO and KEGG enrichment analyses

Pearson correlation analysis was first applied to calculate the correlation coefficient between samples and the result was visualized through R package "pheatmap". Gene expression differences between groups (KO-0h vs. WT-0h, WT-12 h vs. WT-0h, KO-12 h vs. KO-0h, KO-12 h vs. WT-12 h) were detected by using R package "DESeq2" (Love et al., 2014) and genes with adjusted *p*-value < 0.05 and absolute log2 (fold change) > 1 were considered significantly differentially expressed. R packages "ggplot2" (Wickham, 2009), "clusterProfiler" (Wu et al., 2021), "enrichplot", and "org.Hs.eg.db" were applied for volcano plotting, gene set enrichment analysis (GSEA), and enrichment analysis of gene ontology (GO) (Ashburner et al., 2000; Mi et al., 2019; Consortium, 2021) and Kyoto Encyclopedia of Genes and Genomes (KEGG) (Kanehisa and Goto, 2000). A cutoff of adjusted *p*-value = 0.05 was set for the significance of enrichment analysis.

Constructions of PPI network and ceRNA network

Protein-protein interaction (PPI) network among the differentially expressed genes was constructed by using the STING database (Szklarczyk et al., 2021) (*http://string-db.org/*) and Cytoscape software (Shannon et al., 2003). A densely connected region concerning 10 critical genes from the PPI network was detected *via* the MCODE plugin (Bader and Hogue, 2003) of Cytoscape (Degree cutoff = 2, Node score cutoff = 0.2, K-score = 2, Max. depth = 100). Ten cataract-related miRNAs were collected *via* literature review and the predicted mRNAs and lncRNAs interacted with these miRNAs were obtained by searching the databases of miRDB (Chen and Wang, 2020) (*http://www.mirdb.org/*) and Targetscan (McGeary and Lin, 2019) (*http://www.targetscan.org/*) and databases of starBase(Li et al., 2014) (*https://starbase.sysu.edu.cn/*) and miRnet (Chang et al., 2020) (*https://www.mirnet.ca/*), respectively. Then these predicated mRNAs or lncRNAs that were not differentially expressed between the groups (p-adjusted<0.05) were filtered out and a competing endogenouse RNA (ceRNA) network of lncRNA-miRNA-mRNA was plotted *via* Cytoscape software.

Results

Gene expression profile of HLE-B3 cells after knockout of *ATG7*

First, Pearson correlation analysis was applied to explore the correlations between groups and the result showed that samples treated with H_2O_2 were more mutually correlated (Figure 1B), suggesting that H_2O_2 treatment exerted more influence on gene expression than *ATG7* knockout. To ascertain the effects of *ATG7* knockout on gene expression, we first performed differential expression analysis of genes between KO-0h and WT-0h groups, and 1,189 up-regulated and 965 down-regulated genes in KO-0h group were found compared with

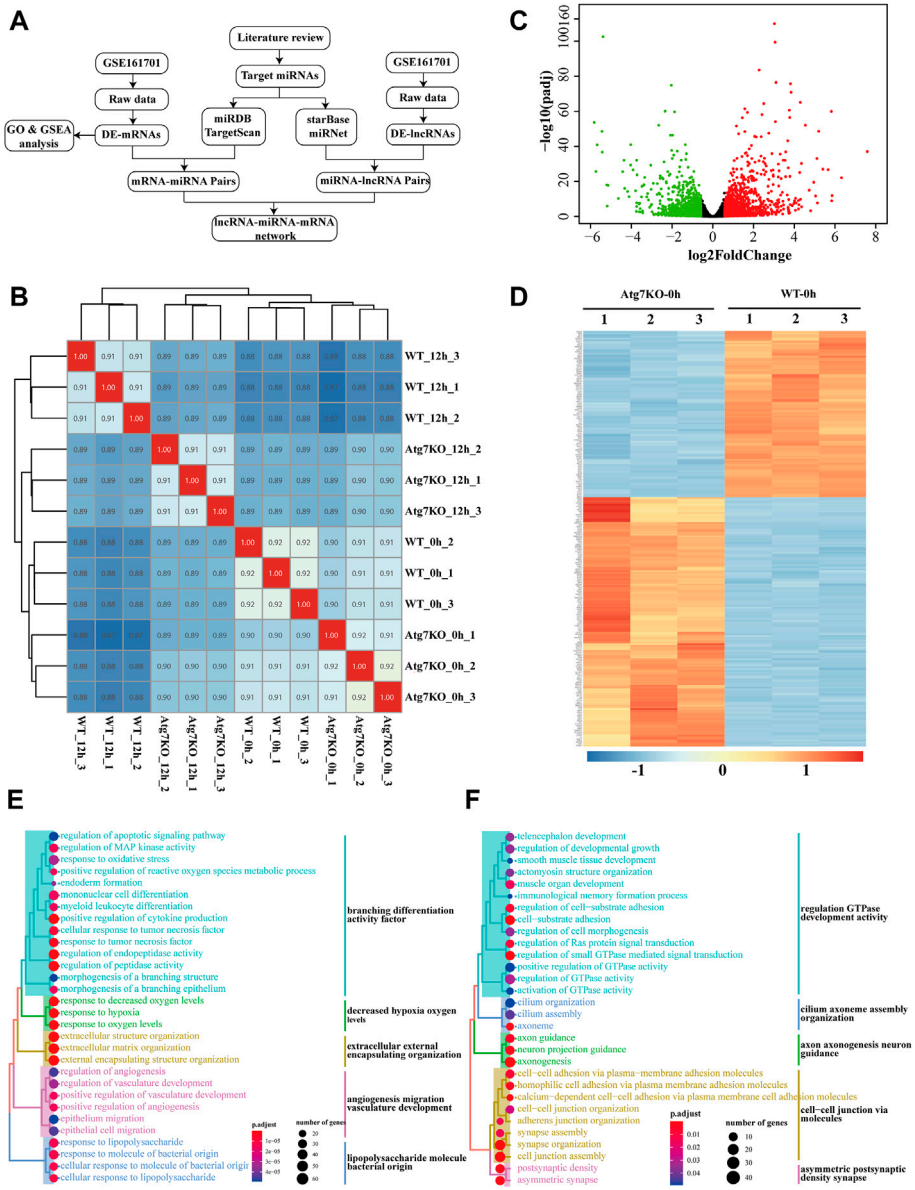


FIGURE 1
Gene expression profile of HLE-B3 cells after knockout of *ATG7*. **(A)** The flowchart showing the ways the study was carried out **(B)** Heatmap of the sample correlation results. Numbers in the cells denote the corresponding sample correlation coefficients and the blue gradient indicates the degree of coefficients. The darker the color is, the larger the coefficient is. Cells in red indicate the self-correlation. **(C)** Volcano plotting of differentially expressed mRNAs in KO-0h vs. WT-0h groups. Red dots and green dots signify the up-regulated and down-regulated mRNAs. Criteria of fold change >1.5 and adjusted *p*-value<0.05 were applied **(D)** Heatmap of clustered top 250 differentially expressed mRNAs in KO-0h vs. WT-0h groups. The color scale indicates the degrees of expression levels with the blue signifies the most down-regulated and the red the most up-regulated in KO-0h group. **(E and F)** GO enrichment analysis for the differentially up-regulated **(E)** and down-regulated **(F)** genes in KO-0h group. The size of the dots indicates the number of clustered genes and the color of the dots signifies the adjusted *p*-value of enrichment. The enriched GO terms were further clustered in different colors.

WT-0h group (p-adjusted<0.05 and absolute log2foldchange >1), as illustrated in the volcano plot (Figure 1C). The two group samples can be well clustered by the selected top 250 differentially expressed genes (shown in Figure 1D). To further explore the impacts of *ATG7* knockout, we performed GO analysis for the up-regulated and down-regulated genes, separately. The results showed that the top 30 enriched GO terms for up-regulated genes mainly resided in branching differentiation activity factor, decreased hypoxia oxygen levels, extracellular external encapsulating organization,

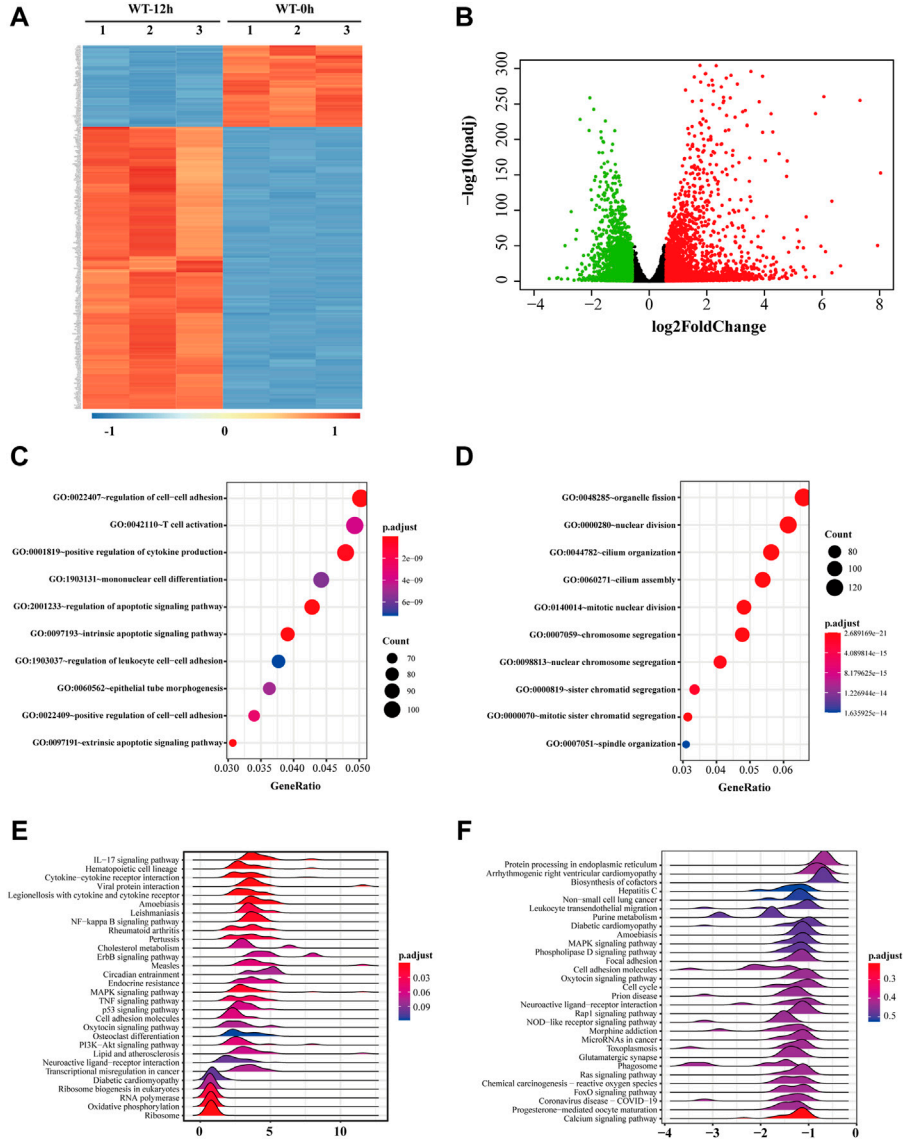


FIGURE 2
Gene expression profile of HLE-B3 cells after H_2O_2 treatment. **(A)** Top 250 differentially expressed mRNAs were clustered based on their expression levels in WT-12 h vs. WT-0h group. The color scale from blue to red is parallel with the expression level from most down-regulated to most up-regulated in WT-12 h group **(B)** Volcano plotting of the distribution of overall differentially expressed mRNAs in WT-12 h vs. WT-0h group. Red dots are for the up-regulated and green dots for the down-regulated. Criteria of fold change >1.5 and adjusted *p*-value<0.05 were applied. **(C and D)** Bubble charts of the top 10 enriched GO terms the up-regulated mRNAs **(C)** and the down-regulated mRNAs **(D)**. The size and color of the dot correspond to the clustered gene number and the related adjusted *p*-value of enrichment. **(E and F)** Ridgeline diagrams show the top 30 KEGG pathways of GSEA for the up-regulated **(E)** and the down-regulated **(F)** mRNAs. *X*-axis indicates the fold change of mRNA expressions. The color of each ridgeline represents the adjusted *p*-value for the enrichment significance of each pathway. GSEA, gene set enrichment analysis.

angiogenesis migration vasculature development and lipopolysaccharide molecule bacterial origin (Figure 1E), while the top 30 GO terms for down-regulated genes were more concerned with regulation GTPase development activity, cilium axoneme assembly organization, axon axonogenesis neuron guidance, cell-cell junction *via* molecules and asymmetric postsynaptic density synapse. (Figure 1F).

Effects of H_2O_2 stimulation on gene expression in HLE-B3 cells

H_2O_2 treatment was applied as a general approach to simulate oxidative stress commonly detected during cataract development. To investigate the biological changes of cultured cells under such oxidative stress, we first analyzed the differential expressed genes

TABLE 1 The top 10 up- and down-regulated mRNAs in WT 12 h vs. WT 0 h.

mRNA	Log_2FC	P_{adj}	Change
HSPA6	11.59	5.53E-56	UP
AREG	8.03	2.25E-153	UP
CSF3	7.93	4.34E-51	UP
GDF15	7.32	8.91E-256	UP
GREB1	6.64	2.17E-22	UP
LHX3	6.34	1.49E-12	UP
ANGPTL4	6.34	1.28E-113	UP
SH2D2A	6.12	2.16E-42	UP
BTG2	6.07	4.85E-261	UP
PLEKHA6	5.96	4.22E-50	UP
HLA-DRA	−3.47	2.59E-04	DOWN
DACH1	−3.21	5.44E-05	DOWN
CYBB	−3.17	7.82E-06	DOWN
C1QTNF3	−3.05	9.09E-04	DOWN
SIM1	−2.92	7.90E-51	DOWN
PDE7B	−2.86	1.70E-19	DOWN
SORBS2	−2.83	5.36E-04	DOWN
ASPM	−2.81	0.00E+00	DOWN
DIO2	−2.71	7.65E-99	DOWN
MUSK	−2.61	3.86E-02	DOWN

after H_2O_2 treatment of HLE-B3 cells. A total of 2,384 up-regulated and 2,219 down-regulated genes were finally determined (adjusted p-value < 0.05 and absolute fold change >1.5), as demonstrated in the volcano plot (Figure 2B). The top 10 of the up-regulated and down-regulated genes were shown in Table 1. The heatmap further illustrated that these differentially expressed genes can clearly distinguish the WT-12 h group from the WT-0h group (Figure 2A). To further ascertain the possible biological functions of these differentially expressed genes, we performed GO analysis for the up-regulated and down-regulated genes, separately. The results indicated that the top 10 GO terms for up-regulated genes were cell differentiation, regulation of cell-cell adhesion and cytokine production, reproductive associated development, regulation of apoptotic signaling pathway, and intrinsic and extrinsic apoptotic signaling pathway (Figure 2C), while the top10 terms for the down-regulated concerned organelle fission, nuclear division, chromosome segregation, cilium assembly and mitosis related organization (Figure 2D).

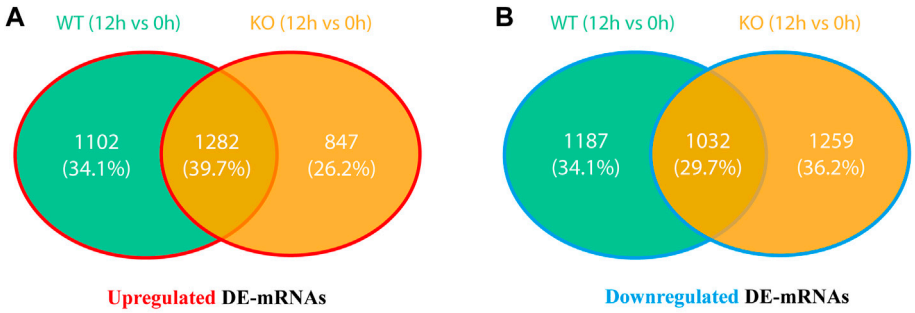
We also performed KEGG analysis *via* GSEA for the up-regulated and down-regulated genes separately to find out the potential pathways these genes involved, and ridgeline plots demonstrated separately the top 30 enriched pathways for the up-regulated and down-regulated (Figures 2E,F). Among them, pathways like hematopoietic cell lineage, NF-kappa B signaling pathway, ErbB signaling pathway, MAPK signaling pathway, TNF signaling pathway, and PI3K-Akt signaling pathway were enriched for the up-regulated while pathways like protein processing in the endoplasmic reticulum, MAPK signaling pathway, oxytocin signaling pathway, rap1 signaling pathway, Ras signaling pathway, and FoxO signaling pathway were enriched for the down-regulated, which suggested these genes under oxidative stress may be involved in cell proliferation and/or apoptosis activities.

Gene expression differences of two types of HLE-B3 cells under oxidative stress

To explore whether H_2O_2 treatment has the same effects on gene expressions of the two types of cells, we first analyzed the differentially expressed genes in KO-12 h vs. KO-0h groups. Among them, the up-regulated and down-regulated genes were intersected respectively with the up-regulated and the down-regulated ones in WT-12 h vs. WT-0h groups. Results showed that there were 1,282 and 1,032 differentially expressed mRNAs (DE-mRNAs) intersected respectively in the up-regulated and down-regulated groups. These shared genes may be regulated independent of *ATG7* expression under oxidative stress, whereas the genes exclusive to WT-12 h vs. WT-0h group (1,102 in up-regulated DE-mRNAs and 1,187 in down-regulated DE-mRNAs, denoted as green part) may contain the candidates that regulated by *ATG7* (Figures 3A,B). To better understand the functions of parted gene groups, we further performed GO analysis separately for these gene sets. Results indicated that the group-shared gene set of up-regulated DE-mRNAs were enriched mainly in the extrinsic and intrinsic apoptotic signaling pathway, regulation of cell-cell adhesion, and regulation of apoptotic signaling pathway, while the shared gene set of the down-regulated were more enriched in the structural organization involved in mitosis. The enriched GO terms for the up-regulated WT-specific group were mainly related to ribosome biogenesis and ncRNA processing, while the terms for the down-regulated WT-specific group involved cilium organization and assembly, endosomal transport, lysosomal transport, and membrane docking. The KO-specific gene set, however, enriched the GO terms of regulation of GTPase activity, axonogenesis, and regulation of cell morphogenesis for the up-regulated DE-genes and the GO terms concerning ATP metabolic process, small molecule catabolic process, and mitochondrial transport for the down-regulated genes. (Figure 3C).

Expressions and functions of *ATG7*-related genes in HLE-B3 cells challenged by H_2O_2

As a critical gene of autophagy and vacuole transport activity, *ATG7* has been reported to be involved in mitophagy and axonal homeostasis. To further investigate the roles of *ATG7* under the circumstances of H_2O_2-induced oxidative stress, we first investigated the differential expressed mRNAs between the



C

	Upregulated DE-mRNAs		Downregulated DE-mRNAs	
	GO-term	Padj	GO-term	Padj
WT KO Shared	extrinsic apoptotic signaling pathway	8.41E-14	mitotic nuclear division	1.91E-22
	intrinsic apoptotic signaling pathway	5.27E-12	chromosome segregation	2.62E-22
	regulation of cell-cell adhesion	1.39E-11	mitotic sister chromatid segregation	1.06E-21
	epithelial tube morphogenesis	1.41E-10	nuclear division	1.06E-21
	gland development	1.47E-10	organelle fission	3.20E-21
	T cell activation	1.63E-10	nuclear chromosome segregation	4.52E-19
	regulation of apoptotic signaling pathway	3.67E-10	sister chromatid segregation	7.80E-19
	positive regulation of cell-cell adhesion	3.67E-10	spindle organization	4.49E-14
	regulation of leukocyte cell-cell adhesion	1.09E-09	microtubule cytoskeleton organization involved in mitosis	4.88E-13
	positive regulation of cell adhesion	2.64E-09	regulation of mitotic nuclear division	1.63E-12
WT Specific	ribosome biogenesis	3.23E-08	cilium organization	9.17E-07
	rRNA metabolic process	9.88E-07	cilium assembly	1.46E-06
	rRNA processing	9.88E-07	ciliary basal body-plasma membrane docking	6.92E-03
	ribonucleoprotein complex biogenesis	2.77E-06	endosomal transport	6.92E-03
	ribosomal large subunit biogenesis	2.77E-06	organelle localization by membrane tethering	9.17E-03
	ncRNA metabolic process	2.20E-05	lysosomal transport	9.17E-03
	viral transcription	3.30E-05	vacuolar transport	9.17E-03
	ncRNA processing	3.30E-05	cell junction assembly	9.17E-03
	protein targeting to ER	1.15E-04	membrane docking	1.50E-02
	establishment of protein localization to ER	1.87E-04	centrosome cycle	1.54E-02
KO Specific	peptidyl-serine phosphorylation	3.98E-06	mitochondrial respiratory chain complex assembly	3.63E-05
	peptidyl-serine modification	1.33E-05	ATP metabolic process	3.63E-05
	regulation of GTPase activity	1.33E-05	electron transport chain	5.39E-05
	axonogenesis	6.94E-04	small molecule catabolic process	8.45E-05
	positive regulation of GTPase activity	1.41E-03	NADH dehydrogenase complex assembly	8.82E-05
	inositol phosphate-mediated signaling	4.90E-03	mitochondrial respiratory chain complex I assembly	8.82E-05
	regulation of cell shape	6.42E-03	mitochondrial transport	1.24E-04
	actin filament-based transport	6.50E-03	sterol biosynthetic process	1.95E-04
	regulation of cell morphogenesis	7.92E-03	cholesterol biosynthetic process	1.95E-04
	calcineurin-NFAT signaling cascade	1.26E-02	secondary alcohol biosynthetic process	2.38E-04

D

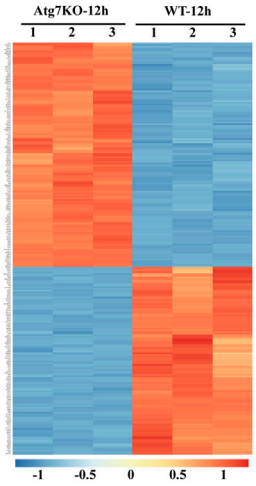


FIGURE 3
Gene expression differences between wild type and ATG7 knockout HLE-B3 cells under oxidative stress. **(A and B)** The Venn diagrams show the intersected, differentially up-regulated **(A)** and down-regulate **(B)** mRNAs in WT-12 h vs. WT-0h group and KO-12 h vs. KO-0h group, specifically. **(C)** The top 10 GO terms of enrichment analysis for the separate gene parts of the above intersections are summarized in table **(D)** The heatmap of top 250 differentially expressed genes between KO-12 h vs. WT-12 h group. The red and the blue indicate the high and low levels of gene expression, respectively.

KO-12 h group and the WT-12 h group. A total of 1,341 up-regulated and 1912 down-regulated mRNAs were detected based on the criteria of absolute fold change>1.5 and adjusted *p*-value < 0.05, and the top 250 DE-mRNAs can distinctly differentiate the two groups, as demonstrated by the heatmap (Figure 3D). Considering the possible two-way regulations of *ATG7* toward its downstream genes and to narrow down the scope of the downstream genes, the down-regulated genes in the KO-12 h vs. WT-12 h group and the up-regulated genes in the WT-12 h vs. WT-0h group were selected to mutually intersect while on the other hand, the up-regulated genes in the KO-12 h vs. WT-12 h group and the down-regulated genes in the WT-12 h vs. WT-0h group were also selected for another mutual intersection. A Venn diagrams showed that a total of 688 DE-mRNAs possibly positively regulated by *ATG7* and a total number of 419 DE-mRNAs possibly negatively regulated by *ATG7* were finally determined (Figures 4A,B). Next, GO enrichment analysis was performed separately on the above-selected DE-mRNAs. Results showed that enriched GO annotations for the 688 DE-mRNAs included but not limited to intrinsic apoptotic signaling pathway in response to endoplasmic reticulum stress, positive regulation of cytokine production, fat cell differentiation, and regulation of hemopoiesis; while the enriched GO terms for the 419 DE-mRNAs were more about GTPase regulator activity, cell leading edge, cell projection membrane, and organ morphogenesis. The top 10 enriched GO terms were shown in Figure 4C,D. The respective chord diagrams further illustrated the relationships between the five representative GO terms and the annotated input genes (Figures 4E,F).

Competing endogenous lncRNA-miRNA-mRNA network associated with *ATG7*

To further explore the functions of the differentially expressed genes potentially regulated by *ATG7*, we conducted the PPI network analysis *via* the STRING database by using the aforementioned 1,107 (688 + 419) differentially expressed genes

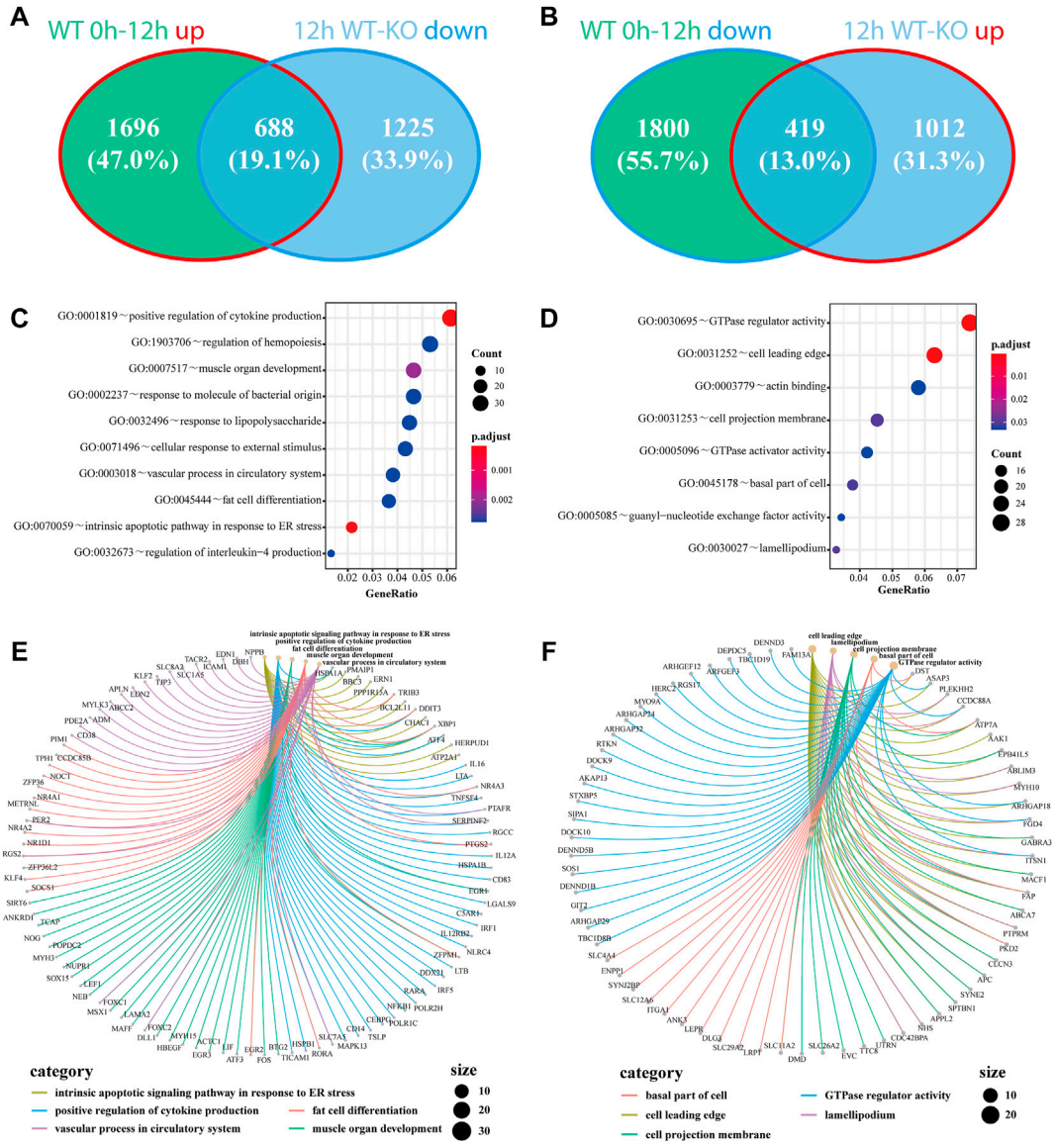


FIGURE 4
Genes regulated by the knockout of *ATG7* and their associated GO terms under oxidative stress. **(A and B)**The down-regulated genes in 12 h KO-WT group are intersected with the up-regulated genes in WT 12h−0 h group **(A)**, while the up-regulated genes in 12 h KO-WT group are intersected with the down-regulated genes in WT 12h−0 h group **(B)**. **(C and D)** The bubble charts show the top 10 enriched GO terms of the above overlapped parts of differentially expressed genes, separately. C and D correspond to A and B, respectively. Dot size denotes the number of involved genes and color scale represents the adjusted *p*-value of enrichment results **(E and F)** The chord diagrams separately demonstrate the correlations of five representative GO terms with their associated genes. Each color represents a single GO term and dot size signifies the involved gene number.

as an input list. The result visualized with Cytoscape showed that most of these genes mutually interacted (red dots denote genes from the 688 DE-mRNAs and blue dots represent genes from the 419 DE-mRNAs) (Figure 5A). To find out the downstream key gene module of *ATG7*, we next calculated the interactions within the genes and dug up two key gene modules, each consisting of seven DE-genes *via* the plugin of MCODE (Figures 5B,C). Since non-coding RNAs are emerging as important regulators during cataract development, we also detected the differentially expressed lncRNAs (DE-lncRNAs) in the KO-12 h vs. WT-12 h groups (shown in Figure 5D). There were 263 up-regulated and 336 down-regulated lncRNAs between the groups. The heatmap showed that the top 250 DE-lncRNAs were clustered well in a manner of expression levels between the groups (Figure 5E). The top 10 differentiated mRNAs and lncRNAs were shown in Table 2. Next, through literature

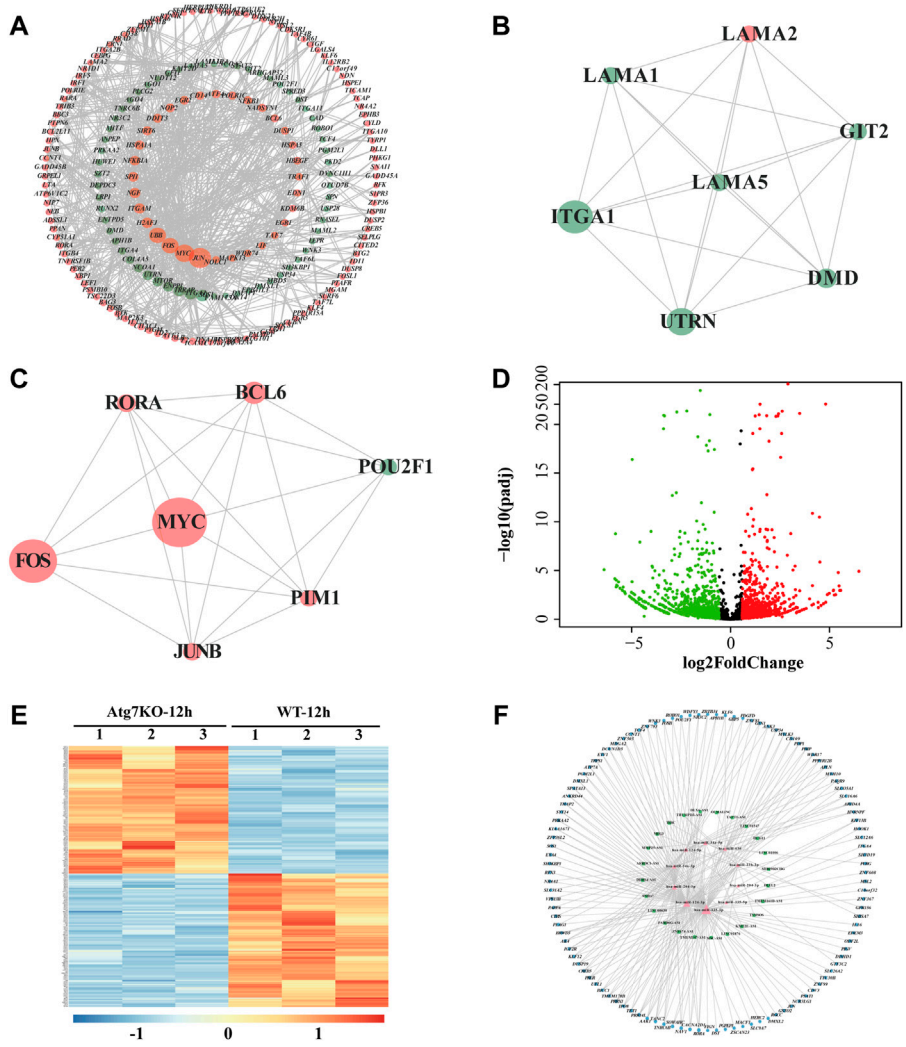


FIGURE 5
LncRNA-miRNA-mRNA network associated with ATG7. **(A)** PPI network of the intersected genes. Dots in orange represent the overlapped genes in Figure 4A and dots in green for the overlapped genes in Figure 4B. The dot size is in proportion to the number of correlations **(B and C)** Two critical gene modules are detected by the MCODE plugin of Cytoscape. **(D)** The volcano plot shows the differentially expressed lncRNAs in KO-12 h vs. WT-12 h group. Red dots for the up-regulated genes and green for the down-regulated. Criterion include fold change>1.5 and adjusted *p*-value > 0.05 **(E)** The heatmap shows the top 250 differentially expressed lncRNAs in KO-12 h vs. WT-12 h group. Color bar corresponds to the relative expression levels of lncRNAs **(F)** The lncRNA-miRNA-mRNA network constructed *via* Cytoscape. The inner pink triangles represent the selected 10 miRNAs reported involved in cataract development, the green rectangles for the possible competing lncRNAs and the outer blue circles for the targeted genes from the PPI network. PPI, protein-protein interaction.

review, 10 miRNAs having been reported participating in cataract development were selected and then input into the databases of miRDB and Targetscan and the databases of starBase and miRnet to predict the potential miRNA-mRNA interactions and miRNA-lncRNA interactions, respectively. Those predicted mRNAs and lncRNAs were then intersected with the aforementioned DE-mRNAs and DE-lncRNAs respectively to narrow their scope. Then the more interconnected mRNAs were further selected by the Cytoscape plugin with more strict criteria. Finally, 112 mRNAs and 24 lncRNAs were determined and these RNAs together with the 10 miRNAs were selected to construct the lncRNA-miRNA-mRNA network *via* Cytoscape (Figure 5F).

Discussion

Age-related cataract (ARC) is the most prominent type of cataracts and oxidative stress is one of the well-established culprit factors during cataract development, since there is a loss of

TABLE 2 The top 10 up- and down-regulated mRNAs and lncRNAs in KO 12 h vs. WT 12 h.

mRNA	Log_2FC	P_{adj}	lncRNA	Log_2FC	P_{adj}	Change
AKR1C3	6.59	1.74E-38	**AC006033.2**	6.49	1.20E-05	UP
CPNE8	6.52	2.38E-21	**AC005840.4**	5.57	1.09E-03	UP
LUM	4.93	1.14E-38	**CASC9**	5.52	1.11E-03	UP
KYNU	4.83	5.92E-54	**LINC02328**	5.48	3.83E-04	UP
PARP8	4.53	1.06E-20	**LINC01597**	5.44	1.62E-05	UP
EXOC3L2	4.13	5.14E-17	**AC010931.2**	5.31	2.98E-03	UP
F2RL1	3.96	1.89E-47	**AC006946.2**	5.25	3.25E-03	UP
SOX5	3.83	4.21E-07	**AC103808.3**	5.20	1.05E-03	UP
IL24	3.67	4.53E-06	**LINC01446**	5.19	1.14E-03	UP
TLR4	3.65	4.21E-61	**AC005838.2**	4.99	2.20E-03	UP
BEX1	−7.38	5.05E-145	**AL365181.3**	−6.39	8.12E-06	DOWN
SLC7A5	−7.00	1.84E-67	**AC092813.1**	−5.85	6.97E-05	DOWN
FBLN1	−6.22	1.48E-15	**AL360270.2**	−5.82	1.70E-09	DOWN
HSPA6	−5.25	3.63E-161	**LINC01315**	−5.78	4.78E-04	DOWN
OVGP1	−5.13	6.99E-10	**AC005921.4**	−5.78	9.76E-05	DOWN
GRB7	−4.95	3.71E-05	**AC138356.3**	−5.76	3.16E-04	DOWN
FBLN2	−4.89	1.54E-30	**AC009812.3**	−5.73	3.83E-04	DOWN
H2AFY2	−4.81	6.80E-25	**AC009108.3**	−5.66	1.49E-04	DOWN
NUP210	−4.59	2.76E-86	**TERC**	−5.59	6.50E-04	DOWN
DPYSL4	−4.59	8.00E-109	**LINC01165**	−5.48	1.02E-03	DOWN

antioxidants, such as reduced GSH, and increased protein oxidation in the nuclear region of the lens prior to the cataract development. (Giblin, 2000; Truscott, 2005; Vinson, 2006). Based on the location of opacity, ARC can be subdivided into three main types: nuclear, cortical, and posterior subcapsular cataracts. (Liu et al., 2017). Despite their association with aging, the underlying pathological changes and mechanisms seem somewhat different. Nuclear cataract, as the most common type, is more subject to oxidative damage compared with the other two types. (Truscott, 2005). In contrast, the destruction of cell structure and precipitation of soluble proteins are more likely observed in cortical cataract, although it is also associated with some degree of oxidation. As for the less common posterior subcapsular cataract, the failed elongation of swollen fiber cells may be the main cause. When considering the risk factors, there are also differences among the three types of cataracts. Cortical cataract is closely associated with high sunlight exposure and is more commonly seen in high-latitude regions, such as northern Japan and northern China. (Sasaki et al., 2002). Wearing sunglasses is verified as an effective way for its prevention. (Taylor et al., 1988). The posterior subcapsular cataract, however, is more associated with high myopia, diabetes, steroid administration, and ionizing radiation. (Beebe et al., 2010). Such discrepancies among age-related cataracts suggest intricate or even distinct mechanisms underlying cataractogenesis. Therefore, the cell line model stimulated by H_2O_2, although commonly used to simulate the factual oxidative stress, can only partly explain the possible mechanisms with caution.

In this study, we detected several pathways and differentially expressed genes that were associated with ATG7. Among them, the intrinsic apoptotic signaling pathway was one of the most enriched pathways for the differentially expressed genes between the groups. As one of the two types of apoptosis signaling, the intrinsic apoptotic signaling pathway is usually activated by internal stimuli such as hypoxia and free radical-induced oxidative stress. It has been found involved in tumor death due to chemotherapy and thus has been investigated as the therapeutic target for drug discovery. (Carneiro and El-Deiry, 2020; Kashyap et al., 2021). With the aging of the lens, lens epithelial cells also underwent certain apoptosis induced by oxidative stress from diabetes, ultraviolet exposure, or just senescence. (Su et al., 2017; Xie et al., 2022). In our study, we detected the intrinsic apoptosis signaling pathway as the potential autophagy-associated mechanism underlying the oxidative stress-induced injury toward LECs. The involved differentially expressed genes in this enriched signaling pathway, such as PMAIP1, BBC3, ENR1, and CHAC1, are thus worthy of further investigation as potential therapeutic targets for cataract development.

It is well recognized that autophagy can contribute to the degradation of damaged or wasted molecules and organelles, including oxidized proteins and lipids, which, if gradually accumulated, could result in cell apoptosis or even necrosis.

(Truscott, 2005). The overexpression of *ATG4* can activate autophagy and meanwhile inhibit apoptosis of the HLE-B3 cell line under H_2O_2-challenged circumstances. (Yan, Zhao, Qin, Zhao, Ji, and Zhang, 2020). And the rapamycin-induced autophagy can alleviate the ROS production in mice LECs cultured in high glucose. (Liu X. et al, 2020). However, some studies revealed the opposite results. The suppression of autophagy by EphA2 can attenuate the apoptosis of SRA01/04 cells induced by H_2O_2. (Han, Wang, Lv, Liu, Dong, Shi and Ji, 2021). Huang J et al. also found that autophagy facilitated the apoptosis of HLE-B3 cells under H_2O_2 stimulation. (Huang, Yu, He, He, Yang, Chen and Han, 2022). By mining the RNA sequencing data, our results showed that the differentially down-regulated genes in autophagy-deficient cells are mostly enriched in GO terms related to cell proliferation and differentiation, such as regulation of hemopoiesis, muscle organ development, and fat cell differentiation, which may suggest a pro-apoptotic function of autophagy. The controversial functions of autophagy toward cell apoptosis under oxidative stress may reside in the distinct cell lines and different concentrations and times of H_2O_2 stimulation. Further studies are warranted to clarify the relationship between autophagy and apoptosis in a detailed manner.

ATG7, as one of the autophagy-related proteins, is a ubiquitin-activating enzyme E1-like protein and participates, along with ATG3, in the conjugation of ATG8 family proteins to phosphatidylethanolamine (PE) during the phagophore expansion. (Dooley et al., 2014). Although ATG7 is an essential mediator in the canonical autophagy pathway, Atg7-independent autophagy was found in *Atg7*-modified mice, where knockout of *Atg7* did not affect the formation of autophagosome and the subsequent bulk degradation. (Nishida, Arakawa, Fujitani, Yamaguchi, Mizuta, Kanaseki, Komatsu, Otsu, Tsujimoto and Shimizu, 2009). The loss of *Atg7* in *Drosophila* did not prevent the occurrence of autophagy and the accompanied cell size reduction. (Chang et al., 2013). However, such ATG7-independent autophagy has not yet been reported in human cells. Furthermore, ATG components, including ATG7, have found to engage in non-autophagic activities, such as phagocytosis (Sanjuan et al., 2007), osteoclastic bone resorption (DeSelm et al., 2011) and antiviral activity of IFNγ (Hwang et al., 2012). In this study, we found that the DEGs shared by the two types of HLE-B3 cells (KO and WT) under oxidative stress are enriched in apoptotic signaling pathways (Figure 3C), which may suggest an *ATG7*-independent autophagy involved in cell apoptosis.

miRNAs, the highly conserved small ncRNAs across species, have long been found regulating substantial gene expressions by targeting the microRNA response elements (MREs) of mRNAs. (Thomas et al., 2010). Each miRNA can regulate many mRNAs and one mRNA can be regulated by a number of miRNAs. (Friedman et al., 2009). And lncRNA, a group of ncRNAs larger than 200 nucleotides in length, has emerged as an important regulator in a myriad of diseases, including cataract. (Liu, Liu, Shan, Zhang, Lu, Yan, and Luo, 2018; Liu X. et al., 2020; Tu et al., 2020; Shen and Zhou, 2021). On the one hand, they can directly target miRNAs to regulate certain mRNA expressions (Ye and Ma, 2020) while on the other, they can also be targeted by the miRNAs (Chi et al., 2009). Such complicated interactions within the ncRNAs and mRNAs suggest the mechanism of ceRNA. (Salmena et al., 2011). In this article, we constructed a predicated lncRNA-miRNA-mRNA ceRNA network based on the differentially expressed lncRNAs and mRNAs and the reported cataract-related miRNAs. Such a ceRNA network may propose an ATG7-associated potential regulatory mechanism underlying cataractogenesis, which warrants further verification.

Conclusion

In all, this study reveals the differentially expressed gene profiles of HLE-B3 cells with or without ATG7 knockout subjected to oxidative stress. And by comparing the gene expressions in these two types of cells under oxidative stimulation, we eventually detected differentially expressed 292 mRNAs and 24 lncRNAs that also interacted with the 10 cataract-associated miRNAs. Thus a competing endogenous lncRNA-miRNA-mRNA network was finally constructed based on such interactions, which 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 in the article/Supplementary Material.

Author contributions

HL carried out the project and drafted the manuscript. LG performed data analysis, JD complied the figures and the tables, and TM performed the language polishing. HL, ZY, and ZL conceived the study and its design as well as revised the manuscript. All authors read and approved the final manuscript.

Funding

This study was supported by the National Natural Science Foundation of China (82070937; 81870640) and the National Science Foundation for Young Scientists of China (82101097; 82000923).

Conflict of interest

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

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.1088943/full#supplementary-material

References

Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., et al. (2000). Gene ontology: Tool for the unification of biology. The gene ontology Consortium. *Nat. Genet.* 25 (1), 25–29. doi:10.1038/75556

Bader, G. D., and Hogue, C. W. (2003). An automated method for finding molecular complexes in large protein interaction networks. *BMC Bioinforma.* 4, 2. doi:10.1186/1471-2105-4-2

Bai, J., Yang, F., Dong, L., and Zheng, Y. (2017). Ghrelin protects human lens epithelial cells against oxidative stress-induced damage. *Oxid. Med. Cell. Longev.* 2017, 1910450. doi:10.1155/2017/1910450

Beebe, D. C., Holekamp, N. M., and Shui, Y. B. (2010). Oxidative damage and the prevention of age-related cataracts. *Ophthalmic Res.* 44 (3), 155–165. doi:10.1159/000316481

Carneiro, B. A., and El-Deiry, W. S. (2020). Targeting apoptosis in cancer therapy. *Nat. Rev. Clin. Oncol.* 17 (7), 395–417. doi:10.1038/s41571-020-0341-y

Chang, L., Zhou, G., Soufan, O., and Xia, J. (2020). miRNet 2.0: network-based visual analytics for miRNA functional analysis and systems biology. *Nucleic Acids Res.* 48 (W1), W244–w251. doi:10.1093/nar/gkaa467

Chang, T. K., Shravage, B. V., Hayes, S. D., Powers, C. M., Simin, R. T., Wade Harper, J., et al. (2013). Uba1 functions in Atg7- and atg3-independent autophagy. *Nat. Cell Biol.* 15 (9), 1067–1078. doi:10.1038/ncb2804

Chen, Y., and Wang, X. (2020). miRDB: an online database for prediction of functional microRNA targets. *Nucleic Acids Res.* 48 (D1), D127–d131. doi:10.1093/nar/gkz757

Chi, S. W., Zang, J. B., Mele, A., and Darnell, R. B. (2009). Argonaute HITS-CLIP decodes microRNA-mRNA interaction maps. *Nature* 460 (7254), 479–486. doi:10.1038/nature08170

Cobb, B. A., and Petrash, J. M. (2002). alpha-Crystallin chaperone-like activity and membrane binding in age-related cataracts. *Biochemistry* 41 (2), 483–490. doi:10.1021/bi0112457

Consortium, T. G. O. (2021). The gene ontology resource: Enriching a GOld mine. *Nucleic Acids Res.* 49 (D1), D325–d334. doi:10.1093/nar/gkaa1113

Cui, H., Kong, Y., and Zhang, H. (2012). Oxidative stress, mitochondrial dysfunction, and aging. *J. Signal Transduct.* 2012, 646354. doi:10.1155/2012/646354

DeSelm, C. J., Miller, B. C., Zou, W., Beatty, W. L., van Meel, E., Takahata, Y., et al. (2011). Autophagy proteins regulate the secretory component of osteoclastic bone resorption. *Dev. Cell* 21 (5), 966–974. doi:10.1016/j.devcel.2011.08.016

Dooley, H. C., Razi, M., Polson, H. E., Girardin, S. E., Wilson, M. I., and Tooze, S. A. (2014). WIPI2 links LC3 conjugation with PI3P, autophagosome formation, and pathogen clearance by recruiting Atg12-5-16L1. *Mol. Cell* 55 (2), 238–252. doi:10.1016/j.molcel.2014.05.021

Fan, F., Zhuang, J., Zhou, P., Liu, X., and Luo, Y. (2017). MicroRNA-34a promotes mitochondrial dysfunction-induced apoptosis in human lens epithelial cells by targeting Notch2. *Oncotarget* 8 (66), 110209–110220. doi:10.18632/oncotarget.22597

Friedman, R. C., Farh, K. K., Burge, C. B., and Bartel, D. P. (2009). Most mammalian mRNAs are conserved targets of microRNAs. *Genome Res.* 19 (1), 92–105. doi:10.1101/gr.082701.108

Gheyas, R., Ortega-Alvarez, R., Chauss, D., Kantorow, M., and Menko, A. S. (2022). Suppression of PI3K signaling is linked to autophagy activation and the spatiotemporal induction of the lens organelle free zone. *Exp. Cell Res.* 412 (2), 113043. doi:10.1016/j.yexcr.2022.113043

Giblin, F. J. (2000). Glutathione: A vital lens antioxidant. *J. Ocul. Pharmacol. Ther.* 16 (2), 121–135. doi:10.1089/jop.2000.16.121

Gu, X. L. (2018). MicroRNA-124 prevents H2O2-induced apoptosis and oxidative stress in human lens epithelial cells via inhibition of the NF-κB signaling pathway. *Pharmacology* 102 (3-4), 213–222. doi:10.1159/000491433

Han, J., Wang, L., Lv, H., Liu, J., Dong, Y., Shi, L., et al. (2021). EphA2 inhibits SRA01/04 cells apoptosis by suppressing autophagy via activating PI3K/Akt/mTOR pathway. *Arch. Biochem. Biophys.* 711, 109024. doi:10.1016/j.abb.2021.109024

Hanafy, B. I., Cave, G. W. V., Barnett, Y., and Pierscionek, B. K. (2021). Nanoceria prevents glucose-induced protein glycation in eye lens cells. *Nanomater. (Basel)* 11 (6), 1473. doi:10.3390/nano11061473

Hsueh, Y. J., Chen, Y. N., Tsao, Y. T., Cheng, C. M., and Wu, W. C. (2022). The pathomechanism, antioxidant biomarkers, and treatment of oxidative stress-related eye diseases. *Int. J. Mol. Sci.* 23 (3), 1255. doi:10.3390/ijms23031255

Huang, J., Yu, W., He, Q., He, X., Yang, M., Chen, W., et al. (2022). Autophagy facilitates age-related cell apoptosis-a new insight from senile cataract. *Cell Death Dis.* 13 (1), 37. doi:10.1038/s41419-021-04489-8

Hwang, S., Maloney, N. S., Bruinsma, M. W., Goel, G., Duan, E., Zhang, L., et al. (2012). Nondegradative role of Atg5-Atg12/Atg16L1 autophagy protein complex in antiviral activity of interferon gamma. *Cell Host Microbe* 11 (4), 397–409. doi:10.1016/j.chom.2012.03.002

Kanehisa, M., and Goto, S. (2000). Kegg: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* 28 (1), 27–30. doi:10.1093/nar/28.1.27

Kashyap, D., Garg, V. K., and Goel, N. (2021). Intrinsic and extrinsic pathways of apoptosis: Role in cancer development and prognosis. *Adv. Protein Chem. Struct. Biol.* 125, 73–120. doi:10.1016/bs.apcsb.2021.01.003

Li, J. H., Liu, S., Zhou, H., Qu, L. H., and Yang, J. H. (2014). starBase v2.0: decoding miRNA-ceRNA, miRNA-ncRNA and protein-RNA interaction networks from large-scale CLIP-Seq data. *Nucleic Acids Res.* 42, D92–D97. doi:10.1093/nar/gkt1248

Li, W. C., Kuszak, J. R., Dunn, K., Wang, R. R., Ma, W., Wang, G. M., et al. (1995). Lens epithelial cell apoptosis appears to be a common cellular basis for non-congenital cataract development in humans and animals. *J. Cell Biol.* 130 (1), 169–181. doi:10.1083/jcb.130.1.169

Liu, J., Dong, Y., Wen, Y., Shi, L., Zhu, Z., Ke, G., et al. (2020a). LncRNA KCNQ1OT1 knockdown inhibits viability, migration and epithelial-mesenchymal transition in human lens epithelial cells via miR-26a-5p/ITGAV/TGF-beta/Smad3 axis. *Exp. Eye Res.* 200, 108251. doi:10.1016/j.exer.2020.108251

Liu, X., Liu, C., Shan, K., Zhang, S., Lu, Y., Yan, B., et al. (2018). Long non-coding RNA H19 regulates human lens epithelial cells function. *Cell. Physiol. biochem.* 50 (1), 246–260. doi:10.1159/000494003

Liu, X., Zhao, X., Cheng, R., and Huang, Y. (2020b). Autophagy attenuates high glucose-induced oxidative injury to lens epithelial cells. *Biosci. Rep.* 40 (4), BSR20193006. doi:10.1042/bsr20193006

Liu, Y. C., Wilkins, M., Kim, T., Malyugin, B., and Mehta, J. S. (2017). Cataracts. *Lancet* 390 (10094), 600–612. doi:10.1016/s0140-6736(17)30544-5

Long, A. C., Colitz, C. M., and Bomser, J. A. (2004). Apoptotic and necrotic mechanisms of stress-induced human lens epithelial cell death. *Exp. Biol. Med. (Maywood)* 229 (10), 1072–1080. doi:10.1177/153537020422901012

Love, M. I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15 (12), 550. doi:10.1186/s13059-014-0550-8

Matsui, M., Yamamoto, A., Kuma, A., Ohsumi, Y., and Mizushima, N. (2006). Organelle degradation during the lens and erythroid differentiation is independent of autophagy. *Biochem. Biophys. Res. Commun.* 339 (2), 485–489. doi:10.1016/j.bbrc.2005.11.044

McGeary, S. E., Lin, K. S., Shi, C. Y., Pham, T. M., Bisaria, N., Kelley, G. M., et al. (2019). The biochemical basis of microRNA targeting efficacy. *Science* 366 (6472), eaav1741. doi:10.1126/science.aav1741

Mi, H., Muruganujan, A., Ebert, D., Huang, X., and Thomas, P. D. (2019). PANTHER version 14: More genomes, a new PANTHER GO-slim and improvements in enrichment analysis tools. *Nucleic Acids Res.* 47 (D1), D419–d426. doi:10.1093/nar/gky1038

Morishita, H., Eguchi, S., Kimura, H., Sasaki, J., Sakamaki, Y., Robinson, M. L., et al. (2013). Deletion of autophagy-related 5 (Atg5) and Pik3c3 genes in the lens causes cataract independent of programmed organelle degradation. *J. Biol. Chem.* 288 (16), 11436–11447. doi:10.1074/jbc.M112.437103

Nishida, Y., Arakawa, S., Fujitani, K., Yamaguchi, H., Mizuta, T., Kanaseki, T., et al. (2009). Discovery of Atg5/Atg7-independent alternative macroautophagy. *Nature* 461 (7264), 654–658. doi:10.1038/nature08455

Pescosolido, N., Barbato, A., Giannotti, R., Komaiha, C., and Lenarduzzi, F. (2016). Age-related changes in the kinetics of human lenses: Prevention of the cataract. *Int. J. Ophthalmol.* 9 (10), 1506–1517. doi:10.18240/ijo.2016.10.23

Ping, X., Liang, J., Shi, K., Bao, J., Wu, J., Yu, X., et al. (2021). Rapamycin relieves the cataract caused by ablation of Gja8b through stimulating autophagy in zebrafish. *Autophagy* 17 (11), 3323–3337. doi:10.1080/15548627.2021.1872188

Rubinsztein, D. C., Shpilka, T., and Elazar, Z. (2012). Mechanisms of autophagosome biogenesis. *Curr. Biol.* 22 (1), R29–R34. doi:10.1016/j.cub.2011.11.034

Salmena, L., Poliseno, L., Tay, Y., Kats, L., and Pandolfi, P. P. (2011). A ceRNA hypothesis: The rosetta stone of a hidden RNA language? *Cell* 146 (3), 353–358. doi:10.1016/j.cell.2011.07.014

Sanjuan, M. A., Dillon, C. P., Tait, S. W., Moshiach, S., Dorsey, F., Connell, S., et al. (2007). Toll-like receptor signalling in macrophages links the autophagy pathway to phagocytosis. *Nature* 450 (7173), 1253–1257. doi:10.1038/nature06421

Sasaki, H., Jonasson, F., Shui, Y. B., Kojima, M., Ono, M., Katoh, N., et al. (2002). High prevalence of nuclear cataract in the population of tropical and subtropical areas. *Dev. Ophthalmol.* 35, 60–69. doi:10.1159/000060806

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, Q., and Zhou, T. (2021). Knockdown of lncRNA TUG1 protects lens epithelial cells from oxidative stress-induced injury by regulating miR-196a-5p expression in age-related cataracts. *Exp. Ther. Med.* 22 (5), 1286. doi:10.3892/etm.2021.10721

Su, D., Hu, S., Guan, L., Wu, X., Shi, C., Yang, X., et al. (2017). Down-regulation of GJA3 is associated with lens epithelial cell apoptosis and age-related cataract. *Biochem. Biophys. Res. Commun.* 484 (1), 159–164. doi:10.1016/j.bbrc.2017.01.050

Szklarczyk, D., Gable, A. L., Nastou, K. C., Lyon, D., Kirsch, R., Pyysalo, S., et al. (2021). The STRING database in 2021: Customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets. *Nucleic Acids Res.* 49 (D1), D605–d612. doi:10.1093/nar/gkaa1074

Takamura, Y., Kubo, E., Tsuzuki, S., and Akagi, Y. (2003). Apoptotic cell death in the lens epithelium of rat sugar cataract. *Exp. Eye Res.* 77 (1), 51–57. doi:10.1016/s0014-4835(03)00083-6

Taylor, H. R., West, S. K., Rosenthal, F. S., Muñoz, B., Newland, H. S., Abbey, H., et al. (1988). Effect of ultraviolet radiation on cataract formation. *N. Engl. J. Med.* 319 (22), 1429–1433. doi:10.1056/nejm198812013192201

Thomas, M., Lieberman, J., and Lal, A. (2010). Desperately seeking microRNA targets. *Nat. Struct. Mol. Biol.* 17 (10), 1169–1174. doi:10.1038/nsmb.1921

Truscott, R. J. (2005). Age-related nuclear cataract-oxidation is the key. *Exp. Eye Res.* 80 (5), 709–725. doi:10.1016/j.exer.2004.12.007

Tseng, S. H., Yen, J. S., and Chien, H. L. (1994). Lens epithelium in senile cataract. *J. Formos. Med. Assoc.* 93 (2), 93–98.

Tu, C., Li, H., Liu, X., Wang, Y., Li, W., Meng, L., et al. (2021). TDRD7 participates in lens development and spermiogenesis by mediating autophagosome maturation. *Autophagy* 17 (11), 3848–3864. doi:10.1080/15548627.2021.1894058

Tu, Y., Xie, L., Chen, L., Yuan, Y., Qin, B., Wang, K., et al. (2020). Long non-coding RNA MEG3 promotes cataractogenesis by upregulating TP53INP1 expression in age-related cataract. *Exp. Eye Res.* 199, 108185. doi:10.1016/j.exer.2020.108185

Vetter, C. J., Thorn, D. C., Wheeler, S. G., Mundorff, C. C., Halverson, K. A., Wales, T. E., et al. (2020). Cumulative deamidations of the major lens protein γS-crystallin increase its aggregation during unfolding and oxidation. *Protein Sci.* 29 (9), 1945–1963. doi:10.1002/pro.3915

Vinson, J. A. (2006). Oxidative stress in cataracts. *Pathophysiology* 13 (3), 151–162. doi:10.1016/j.pathophys.2006.05.006

Wickham, H. (2009). ggplot2: Elegant graphics for data analysis. *J. (Issue)* VIII, 213. doi:10.1007/978-0-387-98141-3

Wu, T., Hu, E., Xu, S., Chen, M., Guo, P., Dai, Z., et al. (2021). clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *Innovation.* 2 (3), 100141. doi:10.1016/j.xinn.2021.100141

Xie, Q., Xue, L., Cao, X., Huang, L., and Song, Y. (2022). Apoptosis of lens epithelial cells and expression of NLRP3-related proteins in patients with diabetes and cataract. *Ocular Immunol. Inflamm.*, 1–8. doi:10.1080/09273948.2022.2079537

Yan, C., Zhao, J., Qin, Y., Zhao, F., Ji, L., and Zhang, J. (2020). Overexpression of ATG4a promotes autophagy and proliferation, and inhibits apoptosis in lens epithelial cells via the AMPK and Akt pathways. *Mol. Med. Rep.* 22 (2), 1295–1302. doi:10.3892/mmr.2020.11205

Ye, W., Ma, J., Wang, F., Wu, T., He, M., Li, J., et al. (2020). LncRNA MALAT1 regulates miR-144-3p to facilitate epithelial-mesenchymal transition of lens epithelial cells via the ROS/NRF2/Notch1/Snail pathway. *Oxid. Med. Cell. Longev.* 2020, 8184314. doi:10.1155/2020/8184314

Yildirim, Z., Yildirim, F., Ucgun, N. I., and Kilic, N. (2009). The evaluation of the oxidative stress parameters in nondiabetic and diabetic senile cataract patients. *Biol. Trace Elem. Res.* 128 (2), 135–143. doi:10.1007/s12011-008-8258-9

Zhang, G. B., Liu, Z. G., Wang, J., and Fan, W. (2020). MiR-34 promotes apoptosis of lens epithelial cells in cataract rats via the TGF-β/Smads signaling pathway. *Eur. Rev. Med. Pharmacol. Sci.* 24 (7), 3485–3491. doi:10.26355/eurrev_202004_20807

Zhang, L., Yan, Q., Liu, J. P., Zou, L. J., Liu, J., Sun, S., et al. (2010). Apoptosis: Its functions and control in the ocular lens. *Curr. Mol. Med.* 10 (9), 864–875. doi:10.2174/156652410793937741

Zhou, W., Xu, J., Wang, C., Shi, D., and Yan, Q. (2019). miR-23b-3p regulates apoptosis and autophagy via suppressing SIRT1 in lens epithelial cells. *J. Cell. Biochem.* 120 (12), 19635–19646. doi:10.1002/jcb.29270

Zhou, Y. F., Guo, B., Ye, M. J., Liao, R. F., and Li, S. L. (2016). Protective effect of rutin against H2O2-induced oxidative stress and apoptosis in human lens epithelial cells. *Curr. Eye Res.* 41 (7), 933–942. doi:10.3109/02713683.2015.1082186

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 10 January 2023
DOI 10.3389/fgene.2022.1110799

OPEN ACCESS

EDITED BY
Hua Zhong,
University of Hawaii at Manoa,
United States

REVIEWED BY
Shiqiang Jin,
Bristol Myers Squibb, United States
Guanghui Liu,
State University of New York at Oswego,
United States
Yao Zhu,
Cornell University, United States

*CORRESPONDENCE
Xiangming Mao,
mxm631221@126.com
Ke Wang,
wangke_225@126.com
Chen Fang,
cfang365@hotmail.com

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

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

RECEIVED 29 November 2022
ACCEPTED 23 December 2022
PUBLISHED 10 January 2023

CITATION
Zheng P, Long Z, Gao A, Lu J, Wang S, Zhong C, Lai H, Guo Y, Wang K, Fang C and Mao X (2023), A five-pseudouridylation-associated-LncRNA classifier for primary prostate cancer prognosis prediction. *Front. Genet.* 13:1110799. doi: 10.3389/fgene.2022.1110799

A five-pseudouridylation-associated-LncRNA classifier for primary prostate cancer prognosis prediction

Pengxiang Zheng[1,2†], Zining Long[1†], Anding Gao[3†], Jianming Lu[1], Shuo Wang[1], Chuanfan Zhong[1], Houhua Lai[1], Yufei Guo[1], Ke Wang[1,4]*, Chen Fang[2]* and Xiangming Mao[1]*

[1]Department of Urology, Zhujiang Hospital, Southern Medical University, Guangzhou, China, [2]Department of Urology, Fuqing City Hospital Affiliated to Fujian Medical University, Fuzhou, Fujian, China, [3]Department of Laboratory Medicine, Fuqing City Hospital Affiliated to Fujian Medical University, Fuzhou, Fujian, China, [4]Department of Urology, The Hospital of Trade-Business in Hunan Province, Changsha, China

Background: Prostate cancer (PCa) is one of the most common cancers in males around the globe, and about one-third of patients with localized PCa will experience biochemical recurrence (BCR) after radical prostatectomy or radiation therapy. Reportedly, a proportion of patients with BCR had a poor prognosis. Cumulative studies have shown that RNA modifications participate in the cancer-related transcriptome, but the role of pseudouridylation occurring in lncRNAs in PCa remains opaque.

Methods: Spearman correlation analysis and univariate Cox regression were utilized to determine pseudouridylation-related lncRNAs with prognostic value in PCa. Prognostic pseudouridylation-related lncRNAs were included in the LASSO (least absolute shrinkage and selection operator) regression algorithm to develop a predictive model. KM (Kaplan-Meier) survival analysis and ROC (receiver operating characteristic) curves were applied to validate the constructed model. A battery of biological cell assays was conducted to confirm the cancer-promoting effects of RP11-468E2.5 in the model.

Results: A classifier containing five pseudouridine-related lncRNAs was developed to stratify PCa patients on BCR and named the "ψ-lnc score." KM survival analysis showed patients in the high ψ-lnc score group experienced BCR more than those in the low ψ-lnc score group. ROC curves demonstrated that ψ-lnc score outperformed other clinical indicators in BCR prediction. An external dataset, GSE54460, was utilized to validate the predictive model's efficacy and authenticity. A ceRNA (competitive endogenous RNA) network was constructed to explore the model's potential molecular functions and was annotated through GO (Gene Ontology) and KEGG (Kyoto Encyclopedia of Genes and Genomes) pathway analyses. RP11-468E2.5 was picked for further investigation, including pan-cancer analysis and experimental validation. Preliminarily, RP11-468E2.5 was confirmed as a tumor promoter.

Conclusion: We provide some evidence that pseudouridylation in lncRNA played a role in the development of PCa and propose a novel prognostic classifier for clinical practice.

KEYWORDS

pseudouridylation (ψ), lncRNA, biochemical recurrence, prognostic model, prostate cancer

1 Introduction

According to the cancer statistics in the United States in 2022, prostate cancer alone will account for 27% (268,490/983,160) of cancer diagnoses in men, exceeding lung cancer to be the top one (Siegel et al., 2022). On the other hand, PCa was the second most diagnosed worldwide, only behind lung cancer (Siegel et al., 2021; Sung et al., 2021). Generally, localized PCa patients can yield a favorable prognosis after radical prostatectomy (RP) or radiation therapy (RT). However, biochemical recurrence (BCR), recognized as a detectable serum prostate-specific antigen (PSA) elevation within 10-year follow-ups, occurs in one-third of patients with RP or RT (Pound et al., 1999; Freedland et al., 2007; Boorjian et al., 2011; Van den Broeck et al., 2019). Furthermore, a long-term follow-up BCR study reported that about 24% of patients with BCR developed clinical progression, and the cancer-related mortality hit approximately 16% (Boorjian et al., 2011; Van den Broeck et al., 2019). Therefore, predicting the probability of developing BCR appears pivotal to the prognosis of PCa patients with the increasing incidence rate of PCa.

Thanks to the general application of next-generation sequencing to whole genomes and transcriptomes, numerous pieces of evidence show that less than 2% of the human genome encodes proteins while the rest is transcribed into non-coding RNAs (ncRNAs) (Djebali et al., 2012). Genetic mutations are primarily responsible for cancer, and most of the mutations reside inside the regions that transcribe ncRNAs (Huarte, 2015). In particular, more-than-200-nucleotide long non-coding RNAs (lncRNAs) take up a giant population of ncRNAs, and remarkably, they are gaining more and more attention in the cancer paradigm for exerting dual functions as both oncogenic and tumor-suppressive factors (Sánchez and Huarte, 2013). Given that lncRNAs are reportedly tissue-specific, it is likely that they may share some specific connections with certain cancer subtypes, shedding light on the development of novel biomarkers for the diagnosis, prognosis, or therapeutic targets of cancers (Ling et al., 2015). For instance, prostate cancer antigen 3 (PCA3), the first FDA-approved lncRNA, appears as a promising and pragmatic biomarker for supporting PCa diagnosis (Sartori and Chan, 2014; Sánchez-Salcedo et al., 2021).

RNA modifications are gradually coming into focus due to the development of novel modification detection methods and the realization that ncRNAs are no longer "junks" in the genome and their expression links to complex physiological and pathological processes (Ling et al., 2015; Barbieri and Kouzarides, 2020). Like DNA and proteins, RNAs can be subject to over 170 post-transcriptional modifications, catalyzed by highly conserved enzymes whose dysregulation leads to a broad spectrum of illnesses, including cancer (Jonkhout et al., 2017; Dinescu et al., 2019; Wiener and Schwartz, 2021). Among all these RNA modifications, seven kinds connect to cancer pathogenesis the strongest, such as 7-methylguanosine modification (m7G), N6-methyladenosine modification (m6A), N1-methyladenosine modification (m1A), 5-methylcytosine modification (m5C), pseudouridylation (ψ) and so forth but the underlying machinery of these modifications except m6A in the cancer field, has remained opaque (Barbieri and Kouzarides, 2020). Of the seven ones, pseudouridylation was the first discovered in the 1950s, once called the "fifth RNA nucleotide" (Davis and Allen, 1957) and the most abundant modification in total RNA of human cells (Penzo et al., 2017; Barbieri and Kouzarides, 2020). Pseudouridine used to be frequently detected and studied in tRNA, rRNA, and snRNA (small nuclear RNA); until recently, it was also discovered in mRNA and lncRNA, especially cancer-related lncRNA (Song and Yi, 2017; Dinescu et al., 2019). For example, ψ sites appeared in lncRNAs such as MALAT1 (metastasis-associated lung adenocarcinoma transcript one), XIST (X-inactive specific transcript), TERC (telomerase RNA component), SNHG1 (Small nucleolar RNA host gene one), ZFAS1 (Zinc finger antisense one), etc. Each of them is related to different malignant processes. Unfortunately, there is no established relationship between ψ modification and cancer events, and further studies are required to confirm this correlation. No previous study has revealed the value of ψ-related lncRNAs in PCa. As a result, in this study, we attempted to develop a ψ-related lncRNA predictive model to serve BCR-risk stratification in PCa patients, validate it internally and externally, and investigate its effects on cancer progression using preliminary experiments.

2 Materials and methods

2.1 Data processing

For bioinformatics analysis, TCGA (The Cancer Genome Atlas) dataset for prostate adenocarcinoma (PRAD) with patients' clinical data (n = 547; tumor tissue = 495, normal tissues = 52) was obtained from the TCGA website (https://portal.gdc.cancer.gov/), and only patients with biochemical recurrent time >1 month (n = 460) were included in the survival analyses. Additionally, the GSE54460 dataset (n = 100) was acquired from the Gene Expression Omnibus (GEO) database (http://www.ncbi.nlm.nih.gov/geo/). The baseline information for both datasets is deposited in Supplementary Table S1. And we processed the data following the instructions in one article (Li et al., 2021). FPKM (Fragments Per Kilobase Million) data was first transformed into TPM (Transcript Per Million) form and then normalized through log2 (TPM +1). We included R software (version: 4.1.0) and two website tools, "Sangerbox 3.0" (http://vip.sangerbox.com/) and "GEPIA2" (http://gepia2.cancer-pku.cn/), for analyses in the study.

2.2 Identification of pseudouridine-related lncRNAs

From literature mining (Rong et al., 2021), 13 pseudouridine-related genes were collected. Then, the expression data of these 13 genes and all lncRNAs from the TCGA-PRAD dataset was extracted. In addition, Spearman's correlation analysis (de Winter et al., 2016) was employed to look into the relationship between lncRNAs and the 13 Ψ-related genes (criteria: |Spearman R| > .4 and p < .05). Eventually, 265 lncRNAs were qualified (Supplementary Table S2). Next, univariate Cox regression analysis (Cox, 1972) was performed on these 265 lncRNAs to evaluate their prognostic values, and finally, 100 lncRNAs with p-value <.05 stood out (Supplementary Table S2).

2.3 Construction and validation of the Ψ-related predictive model

The LASSO (Least Absolute Shrinkage and Selection Operator) regression (Tibshirani, 1996) algorithm with ten-fold cross validation and penalty (R package "glmnet") was applied to narrow down the

number of genes for establishment of the predictive model. The algorithm constructed different models by including various numbers of Ψ-related lncRNAs (n = 100), and the minimum criteria chose the penalty parameter (λ). Ultimately, a five-gene model with the best performance was selected and named the "Ψ-lnc score". The Ψ-lnc score comes from the formula:

$$\Psi - \text{lnc score} = \Sigma_{i=1}^{N} \left(\text{Coefficient}_i \times \text{Expression level of lncRNA}_i\right)$$

Where "N" (N = 5) represents the total number of the lncRNAs in the predictive model, "Coefficient$_i$" denotes a specific lncRNA's coefficient, and "Expression level of lncRNA$_i$" refers to the relative expression level of a certain lncRNA.

The TCGA PCa patients were separated into two balanced subsets (the training subset and the testing subset, each number = 230) using the createDataPartition function in R, and the specific Ψ-lnc score for every patient was calculated using the formula above. Given the median scores in the subsets (.296 in the training subset and .288 in the testing subset), the low- and high- Ψ-lnc score subgroups were defined. The Kaplan–Meier (KM) survival analysis (Kaplan and Meier, 1958; Kim et al., 2018; Bichindaritz, 2021; Bichindaritz et al., 2021) in the "survminer" package depicted the BCR-free survival probability curves between the subgroups. The "survivalROC" package drew the 12-, 36-, and 60-month ROC (Receiver Operating Characteristic) curves (Mandrekar, 2010) to evaluate the predictive power of Ψ-lnc score, and the AUCs (Area Under the Curve) of Ψ-lnc score and typical clinicopathological traits were calculated to compare their clinical value. The GSE54460 dataset (N = 100) validated the predictive model externally.

2.4 Construction of ceRNA network and functional enrichment analysis

The "GDCRNAtools" package was introduced to help construct the potential competitive endogenous RNA (ceRNA) network (Salmena et al., 2011; Li et al., 2018), and the website tool, "Sangerbox 3.0" (http://vip.sangerbox.com/), conducted the functional enrichments of the mRNAs included in the ceRNA network mentioned above.

2.5 Cell culture, RNA extraction, and RT-qPCR assays

Two PCa cell lines, LNCaP and C4-2B, were acquired from the BeNa Culture Collection. Subsequently, both cell lines were cultured in RPMI-1640 media. In addition, 10% fetal bovine serum and 1% Penicillin-Streptomycin solution are combined to make the culture media. The cultivation temperature was 37°C, and the concentration of CO_2 was 5%. Total RNAs from LNCaP and C4-2B cells were extracted using Trizol reagent (15596018, Takara), and they were then reverse-transcribed into cDNA with the help of TransScript All-in-one First-Strand cDNA Synthesis SuperMix for qPCR (AT341-01, TransGen). RT-qPCR (Real-time quantitative PCR) assays were carried out using the PerfectStart Green (AQ601-02, TransGen) on an Applied Biosystems 7,500 Real-Time PCR System. Eventually, the relative expression of RP11-468E2.5 and other four lncRNAs (GAS1RR, RP11-400K9.4, RP11-400K9.3, and LINC02688) were calculated using glyceraldehyde 3-phosphate dehydrogenase (GAPDH) as the reference. All the experiments were equipped with three replicates. Supplementary Table S7 shows the primers for RP11-468E2.5, GAS1RR, RP11-400K9.4, RP11-400K9.3, and LINC02688.

2.6 Patient samples

Prostate cancer tissues (n = 10) and benign prostatic hyperplasia tissues (n = 10) were collected, respectively, from patients of Zhujiang Hospital, Southern Medical University. Fresh tissues were viewed and approved by two pathologists, frozen immediately in liquid nitrogen, and stored at −80°C.

2.7 RNA interference and loss of function assays

GenePharm Company synthesized siRNAs targeting RP11-468E2.5. RT-qPCR confirmed the transfection efficiency after the transfection of siRNAs along with siRNA-Mate (GenePharm) for 72 h. The CCK-8 (Cell Counting Kit-8, MA0218-5, Meilunbio) cell viability assay and colony formation assay inspected the proliferative ability of PCa cell lines after knocking down RP11-468E2.5. The transwell assay examined the change in the invasiveness of PCa cells with downregulation of RP11-468E2.5. Detailed procedures for the above assays are accessible in our previous study (Zhong et al., 2021). All experiments were performed in triplicates. siRNAs targeting sites in RP11-468E2.5 are in Supplementary Table S7.

2.8 Statistical analyses

All bioinformatics analyses were performed by R software version 4.1.0 (The R Project for Statistical Computing, Vienna, Austria). The Spearman's correlation analysis analyzed the correlation between the Ψ-related regulators and lncRNAs. The "survival" package carried out KM survival analysis, and the "survminer" package performed Cox regression analysis. GraphPad Prism 7.0 (GraphPad, La Jolla, CA, United States) analyzed the results of RT-qPCR and CCK-8 cell viability assays. We displayed all statistical results in mean ± SD (standard deviation) with a two-sided test and regarded the results with a p-value of less than .05 as statistically significant.

3 Results

3.1 The landscape of pseudouridylation-related modulators in PCa

The workflow diagram is displayed in Figure 1. Initially, a pseudouridylation-related gene list (PUS1, RPUSD3, TRUB1, PUS3, RPUSD4, RPUSD2, PUS10, PUS7, PUSL1, PUS7L, RPUSD1, DKC1, and TRUB2) was generated *via* literature mining, and then their expression profiling in the TCGA dataset for prostate adenocarcinoma (TCGA-PRAD) was investigated. As shown in Figure 2A, most of the pseudouridylation-related molecules (8 out of 13) were significantly upregulated in tumor samples (n = 492) compared to normal ones (n = 52). Then the CNV (copy number

variation) mutation data in these genes was examined (Figure 2B). Notably, CNV depletion exists in the majority of them (PUS1, RPUSD3, TRUB1, PUS3, RPUSD4, RPUSD2, PUS10, PUS7, PUSL1, and PUS7L), whereas CNV amplification is prevalent in three of them (RPUSD1, DKC1, and TRUB2). Moreover, Figure 2C depicted the locations of these genes with CNV mutations on chromosomes. In line with this, the somatic mutations of these molecules in PCa were determined using an R package called "maftools." As a result, only 8 (1.62%) of 495 samples experienced genetic mutations of these genes (Figure 2D). The missense mutation accounts for a giant proportion, followed by multi-hit mutation, in-frame deletion, frame-shift deletion, and splice-site mutation.

3.2 Establishment of the prognostic model with pseudouridylation-related LncRNAs and its association with clinical characteristics in PCa

The expression profile of all lncRNAs in TCGA-PRAD was extracted to ascertain the lncRNAs associated with pseudouridylation in PCa. Spearman's correlation analysis then defined the pseudouridylation-related lncRNAs as ones whose correlation coefficients exceed |.4| with a *p*-value less than .05. Consequently, we obtained 265 pseudouridylation-related lncRNAs (Supplementary Table S1). Next, univariate Cox proportional hazards regression was applied to figure out which lncRNAs presented prognostic value in PCa among these 266 lncRNAs. Consequently, 100 out of 265 lncRNAs appeared to be the prognostic ones (Supplementary Table S2). Subsequently, using the *createDataPartition* function in R, the TCGA-PRAD dataset with 460 samples were divided into two balanced subsets: one training subset and one testing subset, both of which contained 230 patients, respectively. In the training set, the LASSO regression with ten-fold cross validation and penalty was applied to determine the most appropriate prognostic model, using the 100 pseudouridylation-related lncRNAs above (Supplementary Figure S1A). And finally, a five-gene model was considered the most suitable one based on the LASSO results (Supplementary Figure S1B). Following that, the relationship between clinical characteristics and the expression of the five molecules was revealed in the form of a heatmap. Patients with high expression of RP11-468E2.5 (ENSG00000259321) tended to experience advanced T stage, high Gleason scores (GS), BCR, and lymph node metastasis (Supplementary Figure S1C). To further confirm our preliminary discovery, the samples were separated into several binary subgroups based on the GS (GS <= 7; GS > 7), N stage (N0; N1), T stage (T1/2; T3/4), etc. (Supplementary Figure S2). To begin with, patients with GS > 7 expressed more RP11-468E2.5 than those with GS <= 7 ($p < .001$); in contrast, patients with GS > 7 expressed the other four lncRNAs less (Supplementary Figure S2B). Aside from GS, patients in the N-stage and T-stage subgroups had the same expression patterns for RP11-468E2.5 ($p < .05$) and the other four lncRNAs (Supplementary Figures S2C,D). Next ten pairs of samples from local patients with PCa or benign prostatic hyperplasia (BPH) corroborated the difference in expression of these five lncRNAs between tumor ($n = 10$) and benign tissues ($n = 10$) (Supplementary Figure S2E). The expression disparity of four lncRNAs except for LINC02688 between tumor and benign prostate tissues was consistent with the findings above.

3.3 Performance and validation of the predictive model with the pseudouridylation-related LncRNAs

After generating the predictive model, Spearman's correlation analysis confirmed the association between the 13 pseudouridylation-related genes and the five pseudouridylation-related lncRNAs and it was presented in the form of a correlation heatmap; generally, a strong correlation showed up between these two subgroups of genes (Figure 3A). Given the LASSO results, a scoring formula based on the weighted expressions of the five chosen genes for scoring every PCa patient's prognosis was determined and named the "Ψ-lnc score." The weighted coefficients for each lncRNA were also displayed in a histogram (Figure 3B). In addition, univariate Cox regression analysis confirmed the prognostic value of these lncRNAs, and then the results were exhibited in a forest plot (Figure 3C). Notably, RP11-468E2.5 appeared to be the only risk factor with a hazard ratio (HR) of 2.36 (CI: 1.723–3.232), whereas the others were all protective variables. Subsequently, KM survival curve analysis were introduced to confirm the effects of their expression on PCa prognosis (Supplementary Figures S3A–E). Consistent with the results above, patients with high expression of RP11-468E2.5 had unfavorable BCR-free survival ($p < .001$); in contrast, patients with high expression of each of the other four lncRNAs experienced better BCR-free survival ($p < .05$).

Initially, every patient in the training subset was scored using the formula mentioned above; then, the median score served as the cutoff point to define the high-score and low-score groups inside the training subset. Consequently, Figure 3D depicts the distribution of the Ψ-lnc score, BCR status, expression of the five genes for two subgroups, and the survival analysis. Graphically, more patients who experienced BCR and showed highly-expressed RP11-468E2.5 were in the high-score group than those in the low-score group. In terms of survival, patients in the high-score group had a lower rate of BCR-free survival than those in the low-score group ($p < .0001$). Following that, ROC analysis was employed to draw the 1-year, 3-year, and 5-year ROC curves, calculating the corresponding AUCs to scrutinize the model's clinically predictive capability (Figure 3E). Remarkably, the five-gene predictive approach showed promise in predicting BCR prognosis in PCa patients (1-year AUC = .815; 3-year AUC = .804; 5-year AUC = .833). In parallel, a multivariate ROC analysis confirmed the feasibility of the model in clinical practice. Compared with some clinical traits like preoperative PSA, age at diagnosis, GS and T stage in BCR prognosis, the Ψ-lnc score outperformed them with its AUC ranking first (.744; AUC_{GS} = .671, AUC_{PSA} = .659, $AUC_{T\ stage}$ = .659, AUC_{Age} = .516). Additionally, two Cox regression models (the univariate and multivariate ones) were employed to investigate the clinical value of Ψ-lnc score and the aforementioned clinicopathological features (Supplementary Figure S4). Consequently, Ψ-lnc score surpassed all other features with the highest HR both in univariate and multivariate Cox regression analysis.

Likewise, the established model was then internally validated with the TCGA-PRAD testing subset. After separating the testing subset into the high-score and low-score groups based on the median Ψ-lnc score, the analyses above were repeated to verify the model's authenticity. Figure 4A displays the Ψ-lnc score distribution, BCR status, and gene expression profiles in the two groups. Figure 4B shows that patients in the low-score group yielded more favorable BCR-free survival outcomes than those in the high-score group ($p < .0001$), consistent with the previous results. In terms of predictive power, the

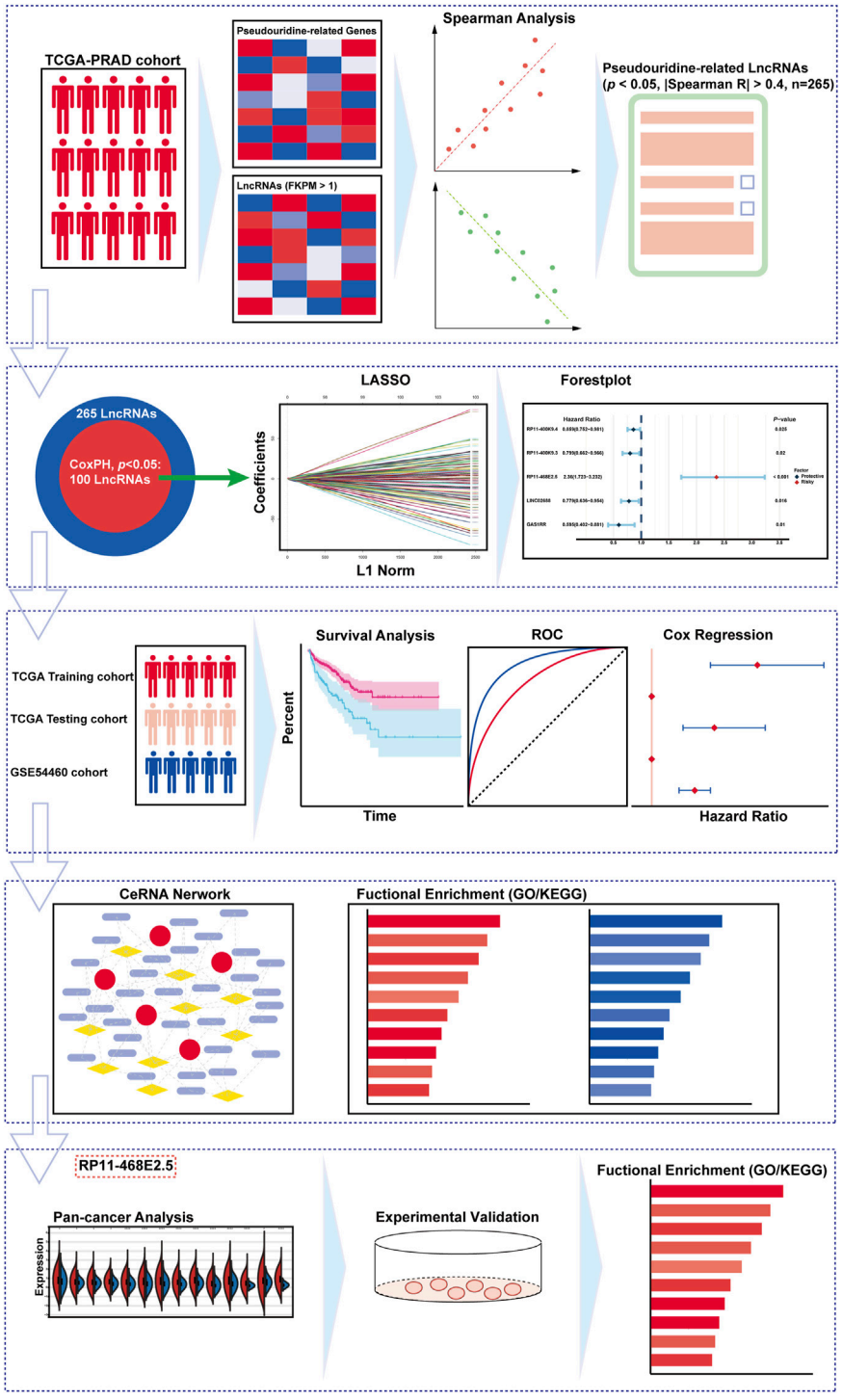


FIGURE 1
The overall design and the flowchart of the study.

model's 12-month, 36-month, and 60-month AUCs in the testing subset are .637, .715, and .775, respectively, harboring considerable outcomes (Figure 4C). Finally, the GSE54460 dataset was introduced to inspect the model's external validity (Figures 4D,E). Patients in the high-score and low-score groups showed a significant difference in BCR-free survival; high-score patients yielded worse outcomes than low-score ones.

3.4 Construction of the potential competing endogenous RNA network and functional enrichment analysis

Following a preliminary examination of the predictive model's performance, attention was drawn to the molecular functions that these genes may possess. It is well known that lncRNAs are likely

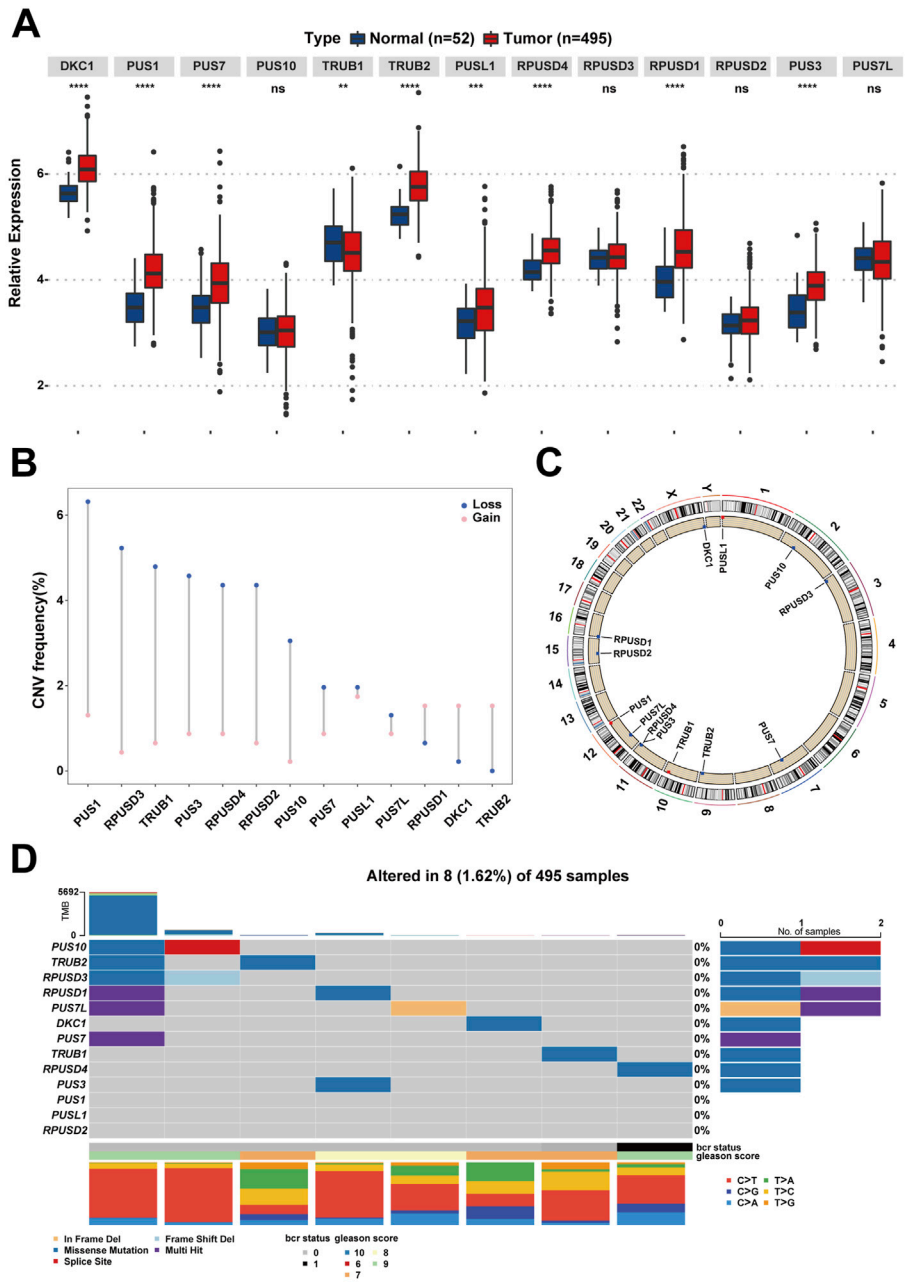


FIGURE 2
The landscape of Ψ-related regulators on expression, mutation, and chromosome location in PCa. **(A)** The differences of gene expression of the 13 Ψ-related regulators between tumor tissues and adjacent normal tissues in the TCGA-PRAD cohort. Ns, no significance; $^{**}p < .01$; $^{***}p < .001$; $^{****}p < .0001$. **(B)** The CNV frequency diagram of the 13 Ψ-related regulators. The two endpoints of each column correspond to two CNV values of a specific Ψ-related gene, with a blue point representing the depletion (Loss) frequency and a pink point denoting the amplification (Gain) frequency, respectively. **(C)** The exact mutation locations of the 13 Ψ-related regulators on chromosomes. **(D)** Eight of 495 (1.62%) PCa patients appeared genetic alterations in the 13 Ψ-related regulators, most of which were missense mutations. The percentages (0%) on the right indicate the mutation frequencies of each Ψ-related regulator, respectively. Each column represents an Ψ gene-mutated individual.

involved in the ceRNA network to exert their effects. Thus, the processed expression data from the TCGA-PRAD dataset was utilized to explore the potential ceRNA network with the help of an R package called "GDCRNAtools." Given the results, all the lncRNA-miRNA-mRNA pairs with their *p*-values and correlation coefficients were obtained. The pairs above were then filtered under the inclusive conditions ($p <$ .05 and |correlation coefficients| > .4) to form the ceRNA network. As a result, a ceRNA network of 754 molecules (5 lncRNAs, 121 microRNAs, and 628 mRNAs) was identified and then visualized using the software "Cytoscape" (Figure 5). Red circles indicate the five lncRNAs, yellow lozenges represent the 121 microRNAs, and blue rectangles represent the 628 mRNAs in the diagram. Detailed links among these three elements are available in Supplementary Table S3. Later, the 628 mRNAs were put into functional enrichment analysis to further investigate their potential

10.3389/fgene.2022.1110799

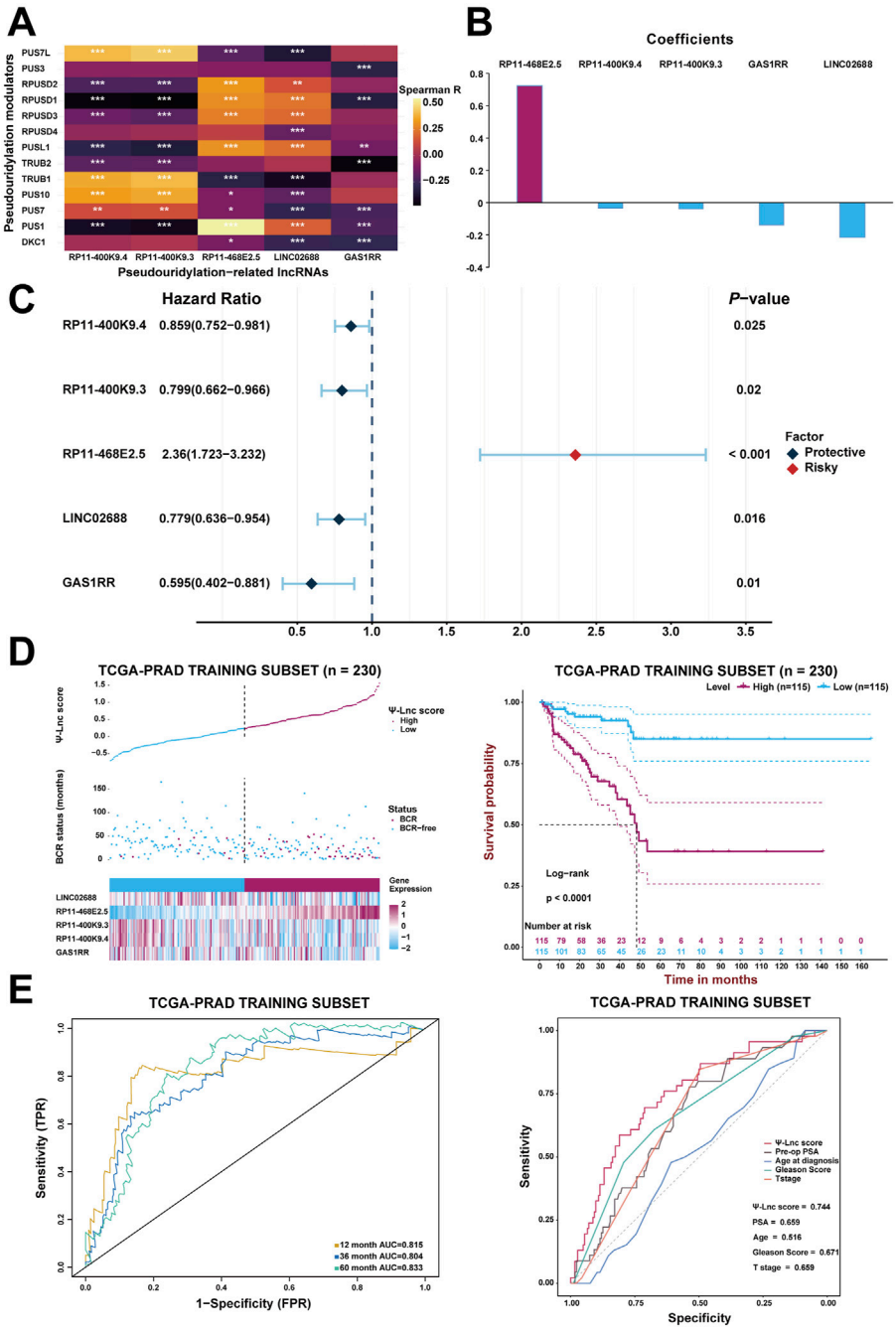


FIGURE 3
Construction of the prognostic model with Ψ-related lncRNAs. **(A)** The heatmap demonstrates the correlation between the 13 Ψ genes and the five lncRNAs included in the model. *$p < .05$; **$p < .01$; ***$p < .001$. **(B)** The coefficients of each selected lncRNA in the Ψ-lnc scoring formula. **(C)** The forest plot shows the univariate Cox regression results of the five lncRNAs. **(D)** The first diagram depicts PCa patients' profiles on three aspects, Ψ-lnc score, BCR status, and the five genes' expression levels, in the TCGA training subset. The second one shows the Kaplan-Meier BCR survival analysis between two Ψ-lnc score subgroups in the training subset. **(E)** The ROC curves show the accuracy of the Ψ-lnc score in predicting BCR-free survival, and the Ψ-lnc score outperforms other clinical indicators.

roles in biological processes. And the website tool called "Sangerbox 3.0" was applied to carry out the enrichment analyses, revealing the gene ontology (GO) terms and KEGG (Kyoto Encyclopedia of Genes and Genomes) pathways highly related to these genes. The GO terms with $p < .05$ and FDR (false discovery rate) < .25 were considered significant; the KEGG pathways with $p < .05$ were also considered meaningful. On the one hand, the top 10 GO terms from each of the three categories (BP, Biological Process; CC, Cellular Component; MF, Molecular Function) were chosen to exhibit in Figures 6A,B. In particular, attention was paid to the underlying biological processes. The top 10 GO terms in BP are regulation of alkaline phosphatase activity (GO:0010692), pigmentation (GO:0043473), positive regulation of alkaline phosphatase activity (GO:0010694), endosomal transport (GO:0016197), cell-substrate junction

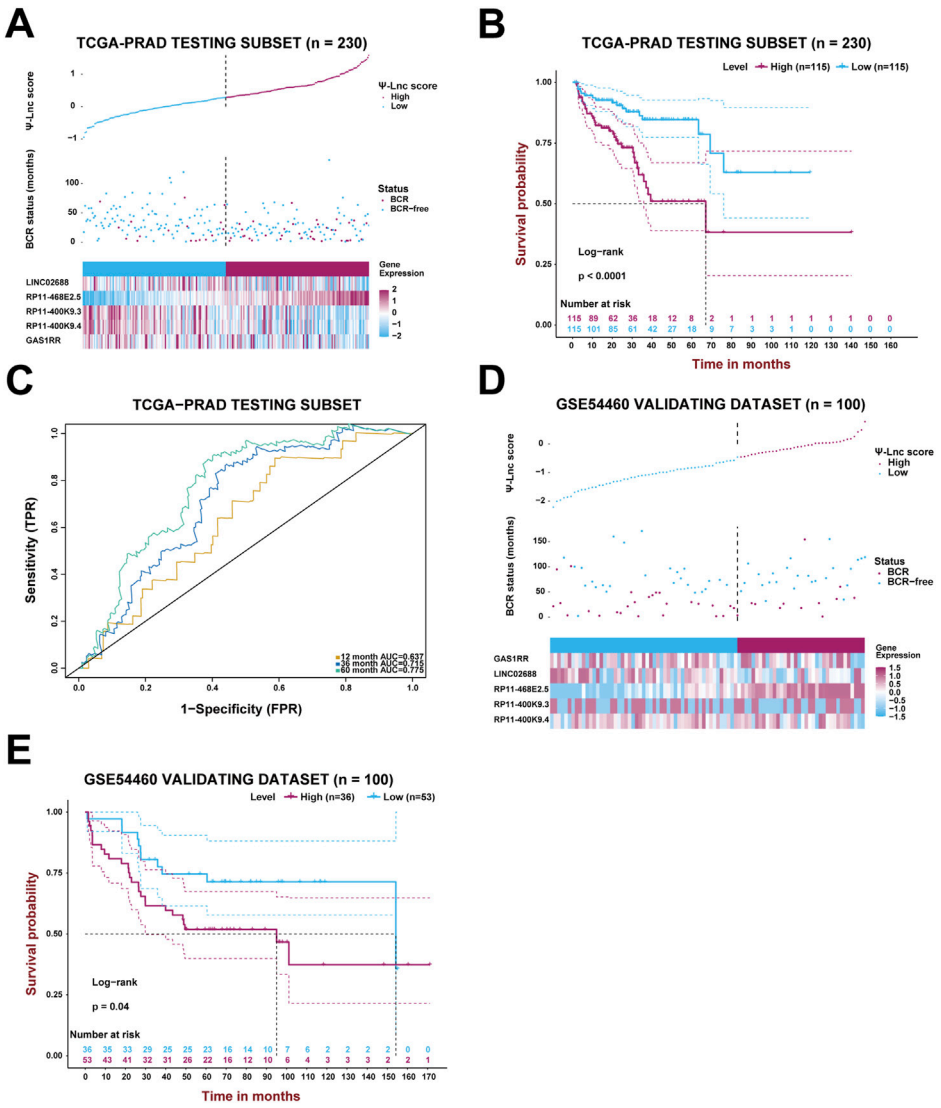


FIGURE 4
Validation of the prognostic model with Ψ-related lncRNAs. **(A)** The PCa patients' profiles on three aspects, Ψ-lnc score, BCR status, and the five genes' expression levels, in the TCGA testing subset. **(B)** The Kaplan-Meier BCR survival analysis between two Ψ-lnc score subgroups in the testing subset. **(C)** The ROC curves show the accuracy of the Ψ-lnc score in predicting BCR-free survival. **(D)** The PCa patients' profiles on three aspects, Ψ-lnc score, BCR status, and the five genes' expression levels, in the GSE54460 validating dataset. **(E)** The Kaplan-Meier BCR survival analysis between two Ψ-lnc score subgroups in the GSE54460 validating dataset.

assembly (GO:0007044), positive regulation of pseudopodium assembly (GO:0031274), response to cadmium ion (GO:0046686), atrial septum development (GO:0003283), regulation of pseudopodium assembly (GO:0031272), and adherens junction assembly (GO:0034333). On the other hand, the top 10 KEGG pathways were also displayed in the form of a ring plot as shown in Figure 6C, including axon guidance (hsa04360), dilated cardiomyopathy (DCM) (hsa05414), phosphonate and phosphinate metabolism (hsa00440), 2-oxocarboxylic acid metabolism (hsa01210), hypertrophic cardiomyopathy (HCM) (hsa05410), TGF-beta signaling pathway (hsa04350), necroptosis (hsa04217), regulation of actin cytoskeleton (hsa04810), sulfur relay system (hsa04122), and glutathione metabolism (hsa00480). The complete information about the GO and KEGG analyses is in Supplementary Table S4.

3.5 Pan-cancer analysis and experimental validation of RP11-468E2.5's effects on PCa

RP11-468E2.5 was the only risk factor with an HR of 1.86 in the established model (Figure 3C), indicating its cancer-promoting effects, so we decided to study its role in cancer, especially PCa, further. At the beginning, a pan-cancer analysis was performed to explore the relationship between its expression and tissue type (normal, tumor) and between its expression and cancer prognosis. Sangerbox 3.0 was used with TCGA data to determine the expression difference of RP11-468E2.5 between tumor-adjacent and tumor samples in each type of tumor, and unpaired Wilcoxon Rank Sum and Signed Rank Tests was implemented to analyze the significance of the difference. Consequently, RP11-468E2.5 is up-regulated significantly ($p < .05$)

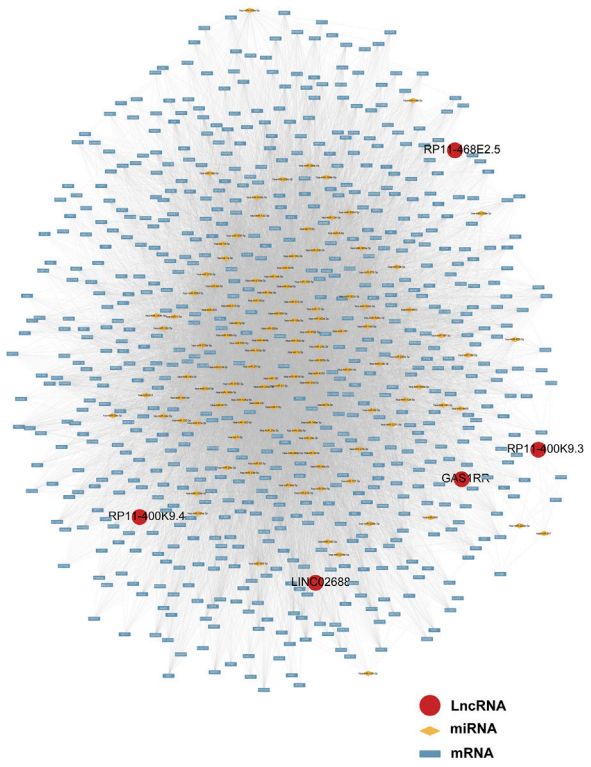


FIGURE 5
The ceRNA network of five lncRNAs with potential miRNAs and mRNAs.

in fourteen types of tumors such as PRAD, LUAD (Lung adenocarcinoma), COAD (Colon adenocarcinoma), COADREAD (Colon adenocarcinoma/Rectum adenocarcinoma), ESCA (Esophageal carcinoma), STES (Stomach and Esophageal carcinoma), KIRP (Kidney renal papillary cell carcinoma), KIRC (Kidney renal clear cell carcinoma), KIPAN (Pan-kidney cohort; KICH, Kidney Chromophobe; KIRC; KIRP), STAD (Stomach adenocarcinoma), HNSC(Head and Neck squamous cell carcinoma), LIHC (Liver hepatocellular carcinoma), BLCA (Bladder urothelial carcinoma), and CHOL (Cholangiocarcinoma), as shown in Supplementary Figure S5A. Next, the Cox proportional hazards regression model analyzed the relationship between RP11-468E2.5's expression and the prognosis of each tumor, one by one. Then the Log-rank test was run to obtain prognostic significance. Finally, the high expression of RP11-468E2.5 in the three types of tumors (PRAD; LUSC, Lung squamous cell carcinoma; ACC, Adrenocortical carcinoma) shows a poor prognosis while the low expression level of RP11-468E2.5 in another four types of tumors (PAAD, Pancreatic adenocarcinoma; SKCM, Skin Cutaneous Melanoma; BLCA; READ) indicates a poor prognosis (Supplementary Figure S5B). Therefore, RP11-468E2.5 is upregulated generally in tumors and its expression demonstrates dual effects on cancer patients' prognosis.

Then experiments were performed to confirm RP11-468E2.5's role in PCa. To begin with, detailed information on RP11-468E2.5 was scrutinized (Figure 7A). RP11-468E2.5 is a lncRNA of 1,000 bp, located on Chromosome 14: 24,139,445–24,140,444. The basal expression of RP11-468E2.5 was checked in six PCa cell lines and one normal prostate cell line (Figure 7B). As a result, RP11-468E2.5 is highly-expressed in four out of six PCa cell lines (LNCaP, C4-2, C4-2B, and 22Rv1) compared to the normal prostate cell line, BPH-1. Thus, two cell lines with the highest expression levels of RP11-468E2.5, LNCaP and C4-2B, were selected for further research. As shown in Figure 7C, three si-RNAs (si-62, si-122, and si-339) were designed to interrupt the expression of RP11-468E2.5 in LNCaP and C4-2B; however, only si-62 and si-122 silenced RP11-468E2.5 significantly, compared to the control group, si-NC. Fluorescence *in situ* hybridization (FISH) assays showed that RP11-468E2.5 mainly exists in the cytoplasmic part of LNCaP and C4-2B cell lines (Figure 7D). Furthermore, its subcellular localization was confirmed in tissues collected from patients with PCa or BPH (Figure 7E). Consistent with our previous findings, RP11-468E2.5 appears highly expressed in the tumor tissue compared to benign prostate tissue. Then the CCK-8 assay examined whether the two RP11-468E2.5-silenced cell lines' proliferative ability was attenuated. After the 5-day observation, silencing RP11-468E2.5 slowed PCa cells' proliferation significantly (Figure 7F). In another aspect, plate colony formation assay was performed to investigate the influence of knocking down RP11-468E2.5 on PCa cells' proliferation ability. Consequently, knock-down of RP11-468E2.5 imposed an attenuative effect on PCa cell viability, too (Figure 7G). The transwell assay demonstrated the decreased invasiveness of PCa cells after downregulating RP11-468E2.5 (Figure 7H). Silencing RP11-468E2.5 hindered PCa cells' invasive ability. Taken together, RP11-468E2.5 was preliminarily confirmed to act as a promoting factor in the development of PCa.

3.6 Functional enrichment analysis for RP11-468E2.5

In spite of RP11-468E2.5's cancer-promoting effects on PCa, the underlying mechanism remains opaque. Thus, RP11-468E2.5 and its top 1,000 similar genes (Supplementary Table S5) obtained from the website GEPIA2 were used to perform functional enrichment analysis. Likewise, the GO terms with $p < .05$ (FDR <.25) and the KEGG pathways with $p < .05$ were considered significant. Figure 8 exhibited the 10 GO terms (except for the MF category) and the top eight KEGG pathways. Specifically, the GO terms in the biological process category are as follow: RNA splicing (GO:0008380), mRNA processing (GO:0006397), RNA processing (GO:0006396), RNA splicing, *via* transesterification reactions with bulged adenosine as nucleophile (GO:0000377), mRNA splicing, *via* spliceosome (GO:0000398), RNA splicing, *via* transesterification reactions (GO:0000375), mRNA metabolic process (GO:0016071), cellular response to DNA damage stimulus (GO:0006974), mRNA export from nucleus (GO:0006406), and mRNA-containing ribonucleoprotein complex export from nucleus (GO:0071427). And the top eight KEGG pathways are as foloow: mRNA surveillance pathway (hsa03015), Spliceosome (hsa03040), Ether lipid metabolism (hsa00565), Fanconi anemia pathway (hsa03460), Base excision repair (hsa03410), Other glycan degradation (hsa00511), Glycerophospholipid metabolism (hsa00564), and alpha-Linolenic acid metabolism (hsa00592). These results may shed some light on the RP11-468E2.5's molecular functions. Detailed information about the functional enrichment results is in Supplementary Table S6.

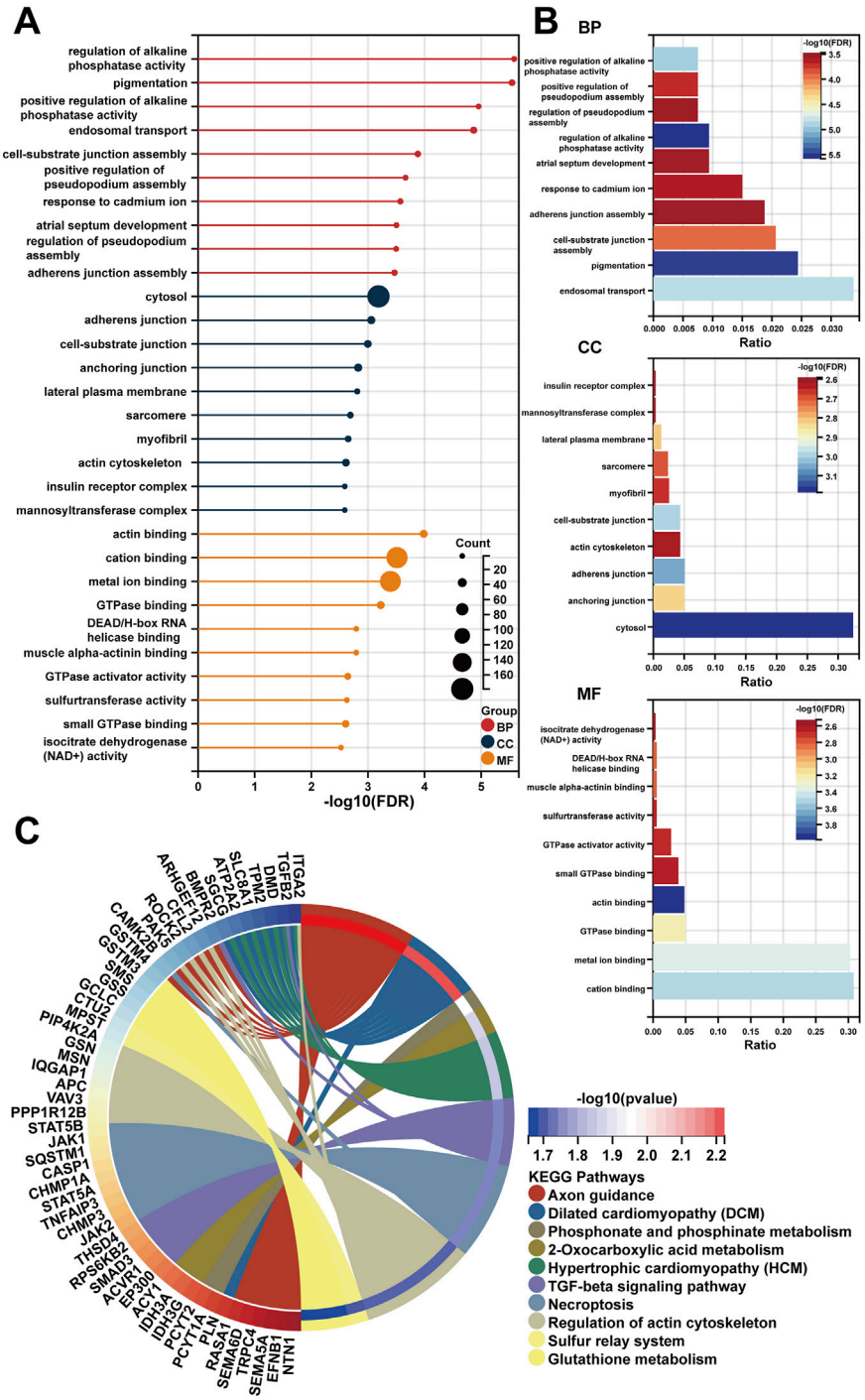


FIGURE 6
The functional enrichment analysis of 628 mRNAs in the ceRNA network. **(A,B)** The diagrams show the top 10 terms in three parts (BP, CC, and MF) of the GO analysis. **(C)** The top 10 pathways in the KEGG analysis.

4 Discussion

The "central dogma" has become the consensus in molecular biology for a drastically long period; The biological diversity all comes from the changes in the nucleotide sequences in DNA/RNA and the 64 codons together to determine the amino acid sequences (Boriack-Sjodin et al., 2018). With techniques for sequencing RNA and DNA pioneered by Fred Sanger in the 1960s and 1970s (Brownlee et al., 1967; Sanger et al., 1977), scientists have been gradually gaining access to the biological details inside these macromolecules. Simultaneously, the effects of chemical modifications to DNA and post-translational modifications to proteins on gene regulation and cancer biology have gained incredible attention in the research community (Esteller, 2007; Chen et al., 2017). Despite this, our understanding of an intermediate layer of regulation between DNA and proteins is still relatively limited. As numerous RNA modifications have come to light, they collectively

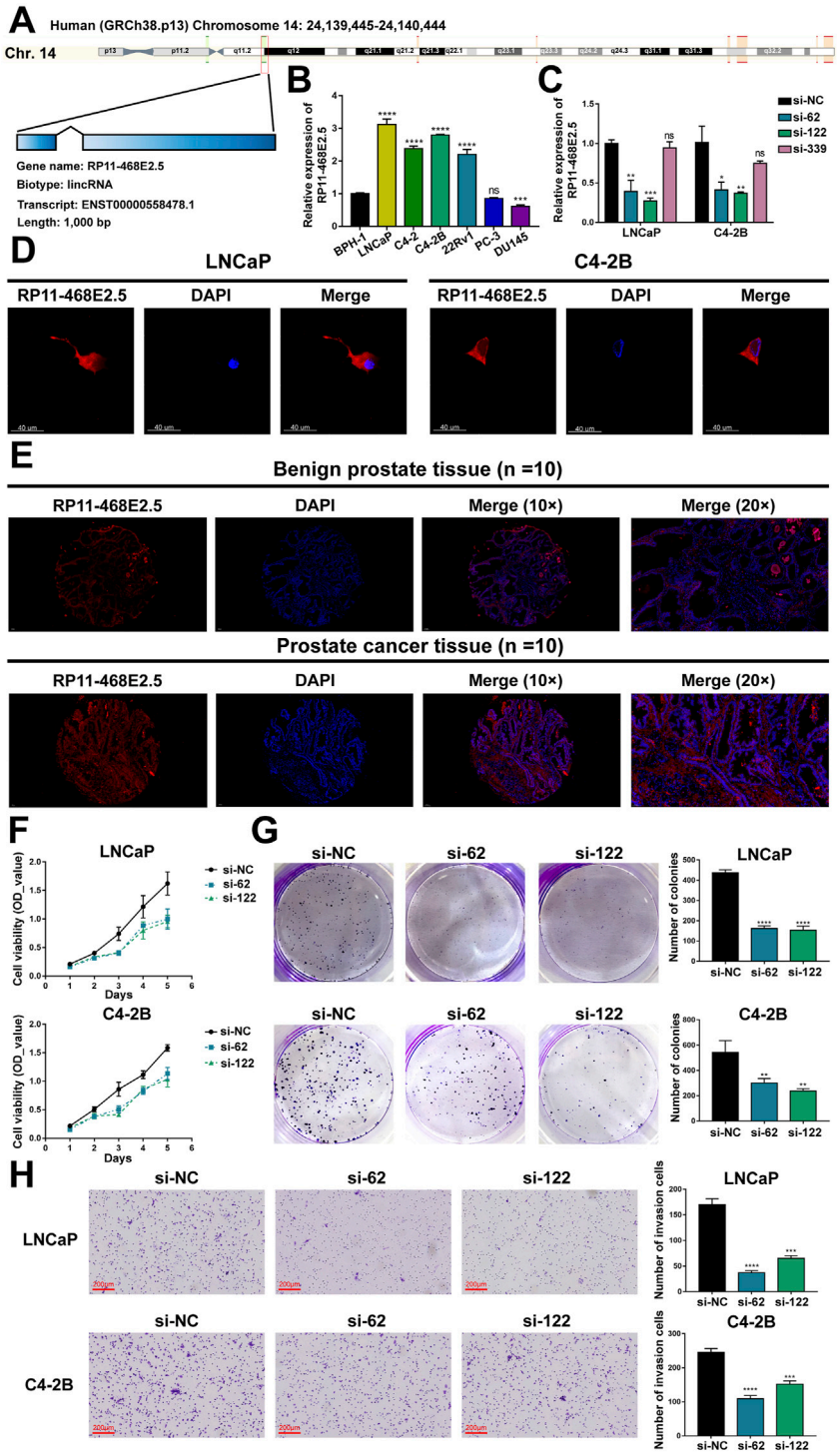


FIGURE 7
Experimental validation of RP11-468E2.5's cancer-promoting effects on PCa. **(A)** The gene information of RP11-468E2.5. **(B)** The basal expressions of RP11-468E2.5 in six PCa cell lines and one normal prostate cell line (BPH-1). **(C)** The gene-silencing efficiencies of three siRNAs in LNCaP and C4-2B PCa cell lines. **(D)** The fluorescence *in situ* hybridization (FISH) assays illustrated that RP11-468E2.5 mainly exists in the cytoplasmic part of LNCaP and C4-2B cell lines. **(E)** FISH assays confirmed that RP11-468E2.5 is highly expressed in tumor tissue compared to benign prostate tissue. **(F)** The proliferation (CCK-8) assays showed silencing RP11-468E2.5 compromised cell viability in LNCaP and C4-2B cell lines. **(G)** The plate colony formation assays demonstrated downregulating RP11-468E2.5 attenuated cell viability in LNCaP and C4-2B cell lines. **(H)** The transwell assay showed silencing RP11-468E2.5 hampered PCa cells' invasiveness.

constitute the concept of "epitranscriptome" (Saletore et al., 2012). These modifications regulate almost every aspect of RNA, such as splicing, nuclear export, translation, degradation, and so on (Gilbert et al., 2016; Peer et al., 2017). It is becoming clear that RNA functioning depends on RNA modifications greatly. And with the dysregulation of RNA epigenetic processes come common human

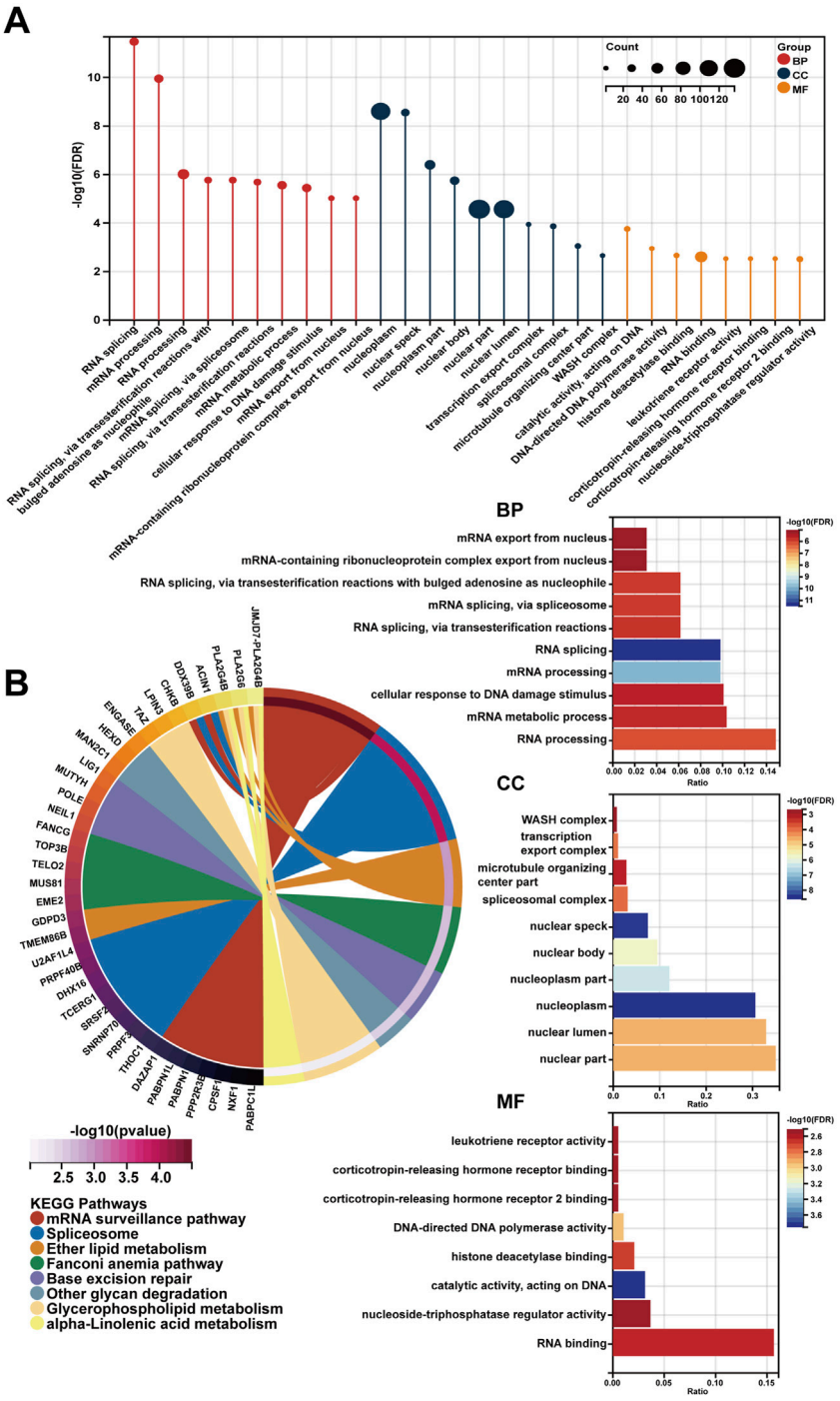


FIGURE 8
The functional enrichment analysis of a gene set including RP11-468E2.5 and its similar genes. **(A)** The diagrams show the top 10 terms in three parts (BP, CC, and MF) of the GO analysis. **(B)** The top eight pathways in the KEGG analysis.

diseases, including cancer (Esteller and Pandolfi, 2017; Barbieri and Kouzarides, 2020). Pseudouridylation is one kind of cancer-associated internal RNA modification but is still rarely investigated in the cancer field compared to two notable ones, m6A, and m5C chemical modifications (Esteller and Pandolfi, 2017; Barbieri and Kouzarides, 2020; Nombela et al., 2021). Pseudouridylation is reportedly the most abundant modification in ncRNAs, and previous studies confirmed its existence in tRNA, rRNA, and snoRNAs. But with the birth of various Ψ-Seq techniques, pseudouridine was also observed in lncRNAs such as XIST and MALAT1, and among ncRNAs, lncRNAs possess the highest abundance of pseudouridine (Li et al., 2015; Esteller and Pandolfi, 2017). How pseudouridylation impacts cancer through modulating lncRNA remains to be elucidated.

PCa is responsible for 7% of newly diagnosed malignancies in males worldwide (2021). According to the GLOBOCAN 2020 estimates, Asia accounted for 26.2% of the global PCa incidence rate and 32.1% of its

mortality rate in 2020 (Sung et al., 2021). PCa is becoming an unaffordable health issue and an economic burden for the public, even in low-incidence-rate areas like Asia. And indeed, a deeper understanding of PCa is urgent for improving prognosis prediction and offering therapeutic vision. From the academic standpoint, no research on pseudouridine-modified lncRNAs affecting PCa's carcinogenesis or progression has existed. Therefore, we aim to reveal some details about this novel topic with bioinformatics and preliminary experiments.

Initially, 13 pseudouridine-related modulators (DKC1, PUS1, PUS7, PUS10, TRUB1, TRUB2, PUSL1, RPUSD4, RPUSD3, RPUSD1, RPUSD2, PUS3, and PUS7L) were confirmed for further analysis (Penzo et al., 2017). Next, pseudouridine-related lncRNAs in PCa were identified by performing Spearman's correlation analysis between the Ψ-related genes and all lncRNAs in the TCGA-PRAD dataset. And a five-pseudouridine-related lncRNA scoring signature for predicting BCR survival in PCa, named "Ψ-lnc score", was developed by the LASSO approach (Tibshirani, 1996), given that LASSO is broadly introduced to the Cox proportional hazard regression model for survival analysis in the bioscience arena (Tibshirani, 1997; Zhang and Lu, 2007). The LASSO method generated a scoring formula based on the expression levels of the five selected genes, of which RP11-468E2.5 tends to be a risk factor, and the other four (GAS1RR, RP11-400K9.4, RP11-400K9.3, and LINC02688) serve as favorable ones.

RP11-468E2.5 is a lncRNA with a length of 1,000 nucleotides, and its influences on cancer are poorly understood. To date, only one study showed that RP11-468E2.5 could negatively target STAT5 and STAT6 to affect the JAK/STAT signaling pathway indirectly (Darnell et al., 1994; Leonard and O'Shea, 1998). Upregulating RP11-468E2.5 curtails the JAK/STAT signaling pathway by targeting two molecules, STAT5 and STAT6, and finally attenuates cell proliferation but boosts cell apoptosis in colorectal cancer (Jiang et al., 2019). However, how RP11-468E2.5 regulates STAT5 and STAT6 negatively remains to be elucidated. In contrast, LINC02688, one of the protective indicators in the constructed model, stays more poorly studied. Only one study unprecedentedly revealed that LINC02688 was expressed less in gastric cancer (GC) tissues compared to paired adjacent normal tissues, and its expression further decreased when GC developed into an advanced one (Fattahi et al., 2021). Additionally, it preliminarily showed considerable prognostic power in GC based on the AUC values of the ROC curve. Nevertheless, more rigorous studies with more clinical samples of different types of cancers and populations from different genetic backgrounds are necessary to explore the exact role of LINC02688 in cancer progression. Lastly, the other three novel lncRNAs haven't unveiled their roles in cancer yet.

After the model construction, the predictive accuracy of Ψ-lnc score was then inspected using KM survival analysis and uni-/multi-variate time-dependent ROC analysis (Heagerty et al., 2000). As a result, Ψ-lnc score appeared to be a satisfactory indicator with the highest AUC value, outperforming typical clinicopathological parameters such as PSA, GS, pathological T stage, and so forth. Subsequently, a dataset (GSE54460) was introduced for the model's external validation; the outcomes were consistent with the previous ones.

Increasing studies demonstrate that lncRNAs that harbor MREs (miRNA-response elements) come up as natural miRNA decoys (Karreth and Pandolfi, 2013). And they are bioinformatically presumed to be broad miRNA targets, suggesting their functioning as ceRNAs (competitive endogenous RNAs) (Griffiths-Jones et al., 2008; Paraskevopoulou et al., 2013). With the ceRNA hypothesis, we asked whether these five lncRNAs in the predictive model work as ceRNAs *via* the R package "GDCRNAtools" and consequently obtained an interactive ceRNA network. To better understand the ceRNA network's functions, functional annotation analysis (GO analysis and KEGG pathway analysis) was performed. As mentioned before, RP11-468E2.5 was the only risk factor with the highest coefficient in the scoring formula, suggesting its dominant role in the model. Given the pan-cancer analysis, RP11-468E2.5 is highly-expressed ($p < .05$) in fourteen types of tumors, including PCa, compared to their correspondent normal tissues. Additionally, its expression exerts tumor-suppressing or cancer-promoting effects on seven kinds of malignancies. Then *in vitro* experiments were implemented to validate its oncogenic role, and consistent results were found in cell proliferation assays in two PCa cell lines (C4-2B and LNCaP). Furthermore, its molecular functions were annotated bioinformatically; annotation analysis using RP11-468E2.5 and its 1,000 similar genes showed it might be involved in the RNA splicing process.

The current study has its limitations, too. Firstly, more public datasets are necessary for better external validation of the established model. Secondly, more advanced experimental validation of RP11-468E2.5 is meaningful for inspecting its molecular functions for the sake of novel pseudouridine-related biomarker development. In aggregate, the constructed model still has a long way to go before it comes into practice.

5 Conclusion

A predictive model containing pseudouridine-related lncRNAs was created to forecast BCR survival probabilities for PCa patients and validated internally and externally. Furthermore, preliminary experiments were performed to validate the cancer-promoting effects of the dominant lncRNA, RP11-468E2.5, in the model. This work sheds some insight into the influence of non-coding RNA modifications on PCa. Still, in-depth studies need to explore how the novel modification, pseudouridylation, functions in the cancer arena.

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

Author contributions

Conception and design: XM, KW, and CF. Collection and assembly of data: SW and JL. Data analysis and interpretation: PZ, ZL, HL, and YG. Experimental planning and execution: PZ, ZL, AG,

and CZ. Manuscript writing and final approval of manuscript: all authors.

Funding

This work was funded by China's National Natural Science Foundation [82173039, 81773277, and 82003271], Guangdong Province Basic and Applied Basic Research Fund Project [2021A1515010659], and Fujian Provincial Health Technology Project [2021QNA067].

Conflict of interest

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

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.1110799/full#supplementary-material

PRESENTATION 1
Contains Supplementary Figures S1-S5

DATA SHEET 1
Contains Supplementary Tables S1-S7

References

Barbieri, I., and Kouzarides, T. (2020). Role of RNA modifications in cancer. *Nat. Rev. Cancer* 20 (6), 303–322. doi:10.1038/s41568-020-0253-2

Bichindaritz, I., Liu, G., and Bartlett, C. (2021). Integrative survival analysis of breast cancer with gene expression and DNA methylation data. *Bioinforma. Oxf. Engl.* 37 (17), 2601–2608. doi:10.1093/bioinformatics/btab140

Bichindaritz, I. (2021). "Predicting with confidence: A case-based reasoning framework for predicting survival in breast cancer," in The International FLAIRS Conference Proceedings, North Miami Beach, FL, United States, May 17–19, 2021 (University of Florida George A Smathers Libraries).

Boorjian, S. A., Thompson, R. H., Tollefson, M. K., Rangel, L. J., Bergstralh, E. J., Blute, M. L., et al. (2011). Long-term risk of clinical progression after biochemical recurrence following radical prostatectomy: the impact of time from surgery to recurrence. *Eur. Urol.* 59 (6), 893–899. doi:10.1016/j.eururo.2011.02.026

Boriack-Sjodin, P. A., Ribich, S., and Copeland, R. A. (2018). RNA-modifying proteins as anticancer drug targets. *Nat. Rev. Drug Discov.* 17 (6), 435–453. doi:10.1038/nrd.2018.71

Brownlee, G. G., Sanger, F., and Barrell, B. G. (1967). Nucleotide sequence of 5S-ribosomal RNA from *Escherichia coli*. *Nature* 215 (5102), 735–736. doi:10.1038/215735a0

Chen, Y., Hong, T., Wang, S., Mo, J., Tian, T., and Zhou, X. (2017). Epigenetic modification of nucleic acids: From basic studies to medical applications. *Chem. Soc. Rev.* 46 (10), 2844–2872. doi:10.1039/c6cs00599c

Cox, D. R. (1972). Regression models and life-tables. *J. R. Stat. Soc. Ser. B Methodol.* 34 (2), 187–202. doi:10.1111/j.2517-6161.1972.tb00899.x

Darnell, J. E., Kerr, I. M., and Stark, G. R. (1994). Jak-STAT pathways and transcriptional activation in response to IFNs and other extracellular signaling proteins. *Sci. (New York, N.Y.)* 264 (5164), 1415–1421. doi:10.1126/science.8197455

Davis, F. F., and Allen, F. W. (1957). Ribonucleic acids from yeast which contain a fifth nucleotide. *J. Biol. Chem.* 227 (2), 907–915. doi:10.1016/s0021-9258(18)70770-9

de Winter, J. C. F., Gosling, S. D., and Potter, J. (2016). Comparing the pearson and spearman correlation coefficients across distributions and sample sizes: A tutorial using simulations and empirical data. *Psychol. Methods* 21 (3), 273–290. doi:10.1037/met0000079

Dinescu, S., Ignat, S., Lazar, A. D., Constantin, C., Neagu, M., and Costache, M. (2019). Epitranscriptomic signatures in lncRNAs and their possible roles in cancer. *Genes* 10 (1), 52. doi:10.3390/genes10010052

Djebali, S., Davis, C. A., Merkel, A., Dobin, A., Lassmann, T., Mortazavi, A., et al. (2012). Landscape of transcription in human cells. *Nature* 489 (7414), 101–108. doi:10.1038/nature11233

Esteller, M., and Pandolfi, P. P. (2017). The epitranscriptome of noncoding RNAs in cancer. *Cancer Discov.* 7 (4), 359–368. doi:10.1158/2159-8290.CD-16-1292

Esteller, M. (2007). Cancer epigenomics: DNA methylomes and histone-modification maps. *Nat. Rev. Genet.* 8 (4), 286–298. doi:10.1038/nrg2005

Fattahi, S., Nikbakhsh, N., Taheri, H., Ghadami, E., Ranaee, M., and Akhavan-Niaki, H. (2021). LINC02688 and PP7080 as novel biomarkers in early diagnosis of gastric cancer. *Non-coding RNA Res.* 6 (2), 86–91. doi:10.1016/j.ncrna.2021.04.002

Freedland, S. J., Humphreys, E. B., Mangold, L. A., Eisenberger, M., Dorey, F. J., Walsh, P. C., et al. (2007). Death in patients with recurrent prostate cancer after radical prostatectomy: Prostate-specific antigen doubling time subgroups and their associated contributions to all-cause mortality. *J. Clin. Oncol. Official J. Am. Soc. Clin. Oncol.* 25 (13), 1765–1771. doi:10.1200/JCO.2006.08.0572

Gilbert, W. V., Bell, T. A., and Schaening, C. (2016). Messenger RNA modifications: Form, distribution, and function. *Sci. (New York, N.Y.)* 352 (6292), 1408–1412. doi:10.1126/science.aad8711

Griffiths-Jones, S., Saini, H. K., van Dongen, S., and Enright, A. J. (2008). miRBase: tools for microRNA genomics. *Nucleic acids Res.* 36, D154–D158. doi:10.1093/nar/gkm952

Heagerty, P. J., Lumley, T., and Pepe, M. S. (2000). Time-dependent ROC curves for censored survival data and a diagnostic marker. *Biometrics* 56 (2), 337–344. doi:10.1111/j.0006-341x.2000.00337.x

Huarte, M. (2015). The emerging role of lncRNAs in cancer. *Nat. Med.* 21 (11), 1253–1261. doi:10.1038/nm.3981

Jiang, L., Zhao, X.-H., Mao, Y.-L., Wang, J.-F., Zheng, H.-J., and You, Q.-S. (2019). Long non-coding RNA RP11-468E2.5 curtails colorectal cancer cell proliferation and stimulates apoptosis via the JAK/STAT signaling pathway by targeting STAT5 and STAT6. *J. Exp. Clin. cancer Res. CR* 38 (1), 465. doi:10.1186/s13046-019-1428-0

Jonkhout, N., Tran, J., Smith, M. A., Schonrock, N., Mattick, J. S., and Novoa, E. M. (2017). The RNA modification landscape in human disease. *RNA (New York, N.Y.)* 23 (12), 1754–1769. doi:10.1261/rna.063503.117

Kaplan, E. L., and Meier, P. (1958). Nonparametric estimation from incomplete observations. *J. Am. Stat. Assoc.* 53 (282), 457–481. doi:10.1080/01621459.1958.10501452

Karreth, F. A., and Pandolfi, P. P. (2013). ceRNA cross-talk in cancer: when ce-bling rivalries go awry. *Cancer Discov.* 3 (10), 1113–1121. doi:10.1158/2159-8290.CD-13-0202

Kim, S. Y., Kim, T. R., Jeong, H.-H., and Sohn, K.-A. (2018). Integrative pathway-based survival prediction utilizing the interaction between gene expression and DNA methylation in breast cancer. *BMC Med. Genomics* 11, 68. doi:10.1186/s12920-018-0389-z

Leonard, W. J., and O'Shea, J. J. (1998). Jaks and STATs: biological implications. *Annu. Rev. Immunol.* 16, 293–322. doi:10.1146/annurev.immunol.16.1.293

Li, X., Zhu, P., Ma, S., Song, J., Bai, J., Sun, F., et al. (2015). Chemical pulldown reveals dynamic pseudouridylation of the mammalian transcriptome. *Nat. Chem. Biol.* 11 (8), 592–597. doi:10.1038/nchembio.1836

Li, R., Qu, H., Wang, S., Wei, J., Zhang, L., Ma, R., et al. (2018). GDCRNATools: an R/bioconductor package for integrative analysis of lncRNA, miRNA and mRNA data in GDC. *Bioinforma. Oxf. Engl.* 34 (14), 2515–2517. doi:10.1093/bioinformatics/bty124

Li, R., Wang, S., Cui, Y., Qu, H., Chater, J. M., Zhang, L., et al. (2021). Extended application of genomic selection to screen multiomics data for prognostic signatures of prostate cancer. *Briefings Bioinforma.* 22 (3), bbaa197. doi:10.1093/bib/bbaa197

Ling, H., Vincent, K., Pichler, M., Fodde, R., Berindan-Neagoe, I., Slack, F. J., et al. (2015). Junk DNA and the long non-coding RNA twist in cancer genetics. *Oncogene* 34 (39), 5003–5011. doi:10.1038/onc.2014.456

Mandrekar, J. N. (2010). Receiver operating characteristic curve in diagnostic test assessment. *J. Thorac. Oncol. Off. Publ. Int. Assoc. For Study Lung Cancer* 5 (9), 1315–1316. doi:10.1097/JTO.0b013e3181ec173d

Nombela, P., Miguel-López, B., and Blanco, S. (2021). The role of m^6A, m^5C and Ψ RNA modifications in cancer: Novel therapeutic opportunities. *Mol. cancer* 20 (1), 18. doi:10.1186/s12943-020-01263-w

Paraskevopoulou, M. D., Georgakilas, G., Kostoulas, N., Reczko, M., Maragkakis, M., Dalamagas, T. M., et al. (2013). DIANA-LncBase: experimentally verified and computationally predicted microRNA targets on long non-coding RNAs. *Nucleic acids Res.* 41, D239–D245. doi:10.1093/nar/gks1246

Peer, E., Rechavi, G., and Dominissini, D. (2017). Epitranscriptomics: regulation of mRNA metabolism through modifications. *Curr. Opin. Chem. Biol.* 41, 93–98. doi:10.1016/j.cbpa.2017.10.008

Penzo, M., Guerrieri, A. N., Zacchini, F., Treré, D., and Montanaro, L. (2017). RNA pseudouridylation in physiology and medicine: For better and for worse. *Genes* 8 (11), 301. doi:10.3390/genes8110301

Pound, C. R., Partin, A. W., Eisenberger, M. A., Chan, D. W., Pearson, J. D., and Walsh, P. C. (1999). Natural history of progression after PSA elevation following radical prostatectomy. *JAMA* 281 (17), 1591–1597. doi:10.1001/jama.281.17.1591

Prostate cancer (2021). Prostate cancer. *Nat. Rev. Dis. Prim.* 7 (1), 8. doi:10.1038/s41572-021-00249-2

Rong, D., Sun, G., Wu, F., Cheng, Y., Sun, G., Jiang, W., et al. (2021). Epigenetics: Roles and therapeutic implications of non-coding RNA modifications in human cancers. *Mol. Ther. Nucleic acids* 25, 67–82. doi:10.1016/j.omtn.2021.04.021

Saletore, Y., Meyer, K., Korlach, J., Vilfan, I. D., Jaffrey, S., and Mason, C. E. (2012). The birth of the epitranscriptome: deciphering the function of RNA modifications. *Genome Biol.* 13 (10), 175. doi:10.1186/gb-2012-13-10-175

Salmena, L., Poliseno, L., Tay, Y., Kats, L., and Pandolfi, P. P. (2011). A ceRNA hypothesis: the rosetta stone of a hidden RNA language? *Cell* 146 (3), 353–358. doi:10.1016/j.cell.2011.07.014

Sánchez, Y., and Huarte, M. (2013). Long non-coding RNAs: challenges for diagnosis and therapies. *Nucleic acid. Ther.* 23 (1), 15–20. doi:10.1089/nat.2012.0414

Sánchez-Salcedo, R., Miranda-Castro, R., de-Los-Santos-Álvarez, N., and Lobo-Castañón, M. J. (2021). Dual electrochemical genosensor for early diagnosis of prostate cancer through lncRNAs detection. *Biosens. Bioelectron.* 192, 113520. doi:10.1016/j.bios.2021.113520

Sanger, F., Air, G. M., Barrell, B. G., Brown, N. L., Coulson, A. R., Fiddes, C. A., et al. (1977). Nucleotide sequence of bacteriophage phi X174 DNA. *Nature* 265 (5596), 687–695. doi:10.1038/265687a0

Sartori, D. A., and Chan, D. W. (2014). Biomarkers in prostate cancer: what's new? *Curr. Opin. Oncol.* 26 (3), 259–264. doi:10.1097/CCO.0000000000000065

Siegel, R. L., Miller, K. D., Fuchs, H. E., and Jemal, A. (2021). Cancer statistics, 2021. *CA Cancer J. Clin.* 71 (1), 7–33. doi:10.3322/caac.21654

Siegel, R. L., Miller, K. D., Fuchs, H. E., and Jemal, A. (2022). Cancer statistics, 2022. *CA a Cancer J. For Clin.* 72 (1), 7–33. doi:10.3322/caac.21708

Song, J., and Yi, C. (2017). Chemical modifications to RNA: A new layer of gene expression regulation. *ACS Chem. Biol.* 12 (2), 316–325. doi:10.1021/acschembio.6b00960

Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., et al. (2021). Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* 71 (3), 209–249. doi:10.3322/caac.21660

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *J. R. Stat. Soc. Ser. B Methodol.* 58 (1), 267–288. doi:10.1111/j.2517-6161.1996.tb02080.x

Tibshirani, R. (1997). The lasso method for variable selection in the Cox model. *Statistics Med.* 16 (4), 385–395. doi:10.1002/(sici)1097-0258(19970228)16:4<385:aid-sim380>3.0.co;2-3

Van den Broeck, T., van den Bergh, R. C. N., Arfi, N., Gross, T., Moris, L., Briers, E., et al. (2019). Prognostic value of biochemical recurrence following treatment with curative intent for prostate cancer: A systematic review. *Eur. Urol.* 75 (6), 967–987. doi:10.1016/j.eururo.2018.10.011

Wiener, D., and Schwartz, S. (2021). The epitranscriptome beyond m6A. *Nat. Rev. Genet.* 22 (2), 119–131. doi:10.1038/s41576-020-00295-8

Zhang, H. H., and Lu, W. (2007). Adaptive Lasso for Cox's proportional hazards model. *Biometrika* 94 (3), 691–703. doi:10.1093/biomet/asm037

Zhong, C., Wu, K., Wang, S., Long, Z., Yang, T., Zhong, W., et al. (2021). Autophagy-related circRNA evaluation reveals hsa_circ_0001747 as a potential favorable prognostic factor for biochemical recurrence in patients with prostate cancer. *Cell Death Dis.* 12 (8), 726. doi:10.1038/s41419-021-04015-w

Glossary

PCa prostate cancer
RP radical prostatectomy
RT radiation therapy
BCR biochemical recurrence
PSA prostate-specific antigen
ncRNA non-coding RNA
lncRNA long non-coding RNA
m7G 7-methylguanosine
m6A N6-methyladenosine
m1A N1-methyladenosine
m5C 5-methylcytosine
ψ pseudouridylation/pseudouridine
MALAT1 metastasis-associated lung adenocarcinoma transcript 1
XIST X-inactive specific transcript
TERC telomerase RNA component
SNHG1 small nucleolar RNA host gene 1
ZFAS1 zinc finger anti-sense 1
TCGA the cancer genome atlas project
TCGA-PRAD TCGA dataset for prostate adenocarcinoma (TCGA-PRAD)
GEO gene expression omnibus
FPKM fragments per kilobase million
TPM transcript per million
LASSO least absolute shrinkage and selection operator
KM Kaplan-Meier
ROC receiver operating characteristic
AUC area under the curve
ceRNA competitive endogenous RNA
GAPDH glyceraldehyde 3-phosphate dehydrogenase
CCK-8 cell counting kit-8
CNV copy number variation
GS Gleason score
BPH benign prostatic hyperplasia
HR hazard ratio
GO gene ontology
KEGG Kyoto encyclopedia of genes and genomes
BP biological process
CC cellular component
MF molecular function
FISH fluorescence *in situ* hybridization
GC gastric cancer
MRE MiRNA-response element

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 20 January 2023
DOI 10.3389/fgene.2023.1106927

OPEN ACCESS

EDITED BY
Shuai Liu,
University of Hawaii at Manoa,
United States

REVIEWED BY
Huiliang Wen,
Nanchang University, China
Fulong Yu,
Broad Institute, United States
Shuxiong Wang,
University of California, Irvine,
United States

*CORRESPONDENCE
Zhongyu Jian,
1178590069@qq.com

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

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

RECEIVED 24 November 2022
ACCEPTED 10 January 2023
PUBLISHED 20 January 2023

CITATION
Di X, Xiang L and Jian Z (2023), YAP-mediated mechanotransduction in urinary bladder remodeling: Based on RNA-seq and CUT&Tag.
Front. Genet. 14:1106927.
doi: 10.3389/fgene.2023.1106927

YAP-mediated mechanotransduction in urinary bladder remodeling: Based on RNA-seq and CUT&Tag

Xingpeng Di[1†], Liyuan Xiang[1,2†] and Zhongyu Jian[1*]

[1]Department of Urology, Institute of Urology (Laboratory of Reconstructive Urology), West China Hospital, Sichuan University, Chengdu, Sichuan, China, [2]Department of Clinical Research Management, West China Hospital, Sichuan University, Chengdu, Sichuan, China

Yes-associated protein (YAP) is an important transcriptional coactivator binding to transcriptional factors that engage in many downstream gene transcription. Partial bladder outlet obstruction (pBOO) causes a massive burden to patients and finally leads to bladder fibrosis. Several cell types engage in the pBOO pathological process, including urothelial cells, smooth muscle cells, and fibroblasts. To clarify the function of YAP in bladder fibrosis, we performed the RNA-seq and CUT&Tag of the bladder smooth muscle cell to analyze the YAP ablation of human bladder smooth muscle cells (hBdSMCs) and immunoprecipitation of YAP. 141 differentially expressed genes (DEGs) were identified through RNA-seq between YAP-knockdown and nature control. After matching with the results of CUT&Tag, 36 genes were regulated directly by YAP. Then we identified the hub genes in the DEGs, including CDCA5, CENPA, DTL, NCAPH, and NEIL3, that contribute to cell proliferation. Thus, our study provides a regulatory network of YAP in smooth muscle proliferation. The possible effects of YAP on hBdSMC might be a vital target for pBOO-associated bladder fibrosis.

KEYWORDS

RNA-seq, CUT&Tag, YAP, bladder remodeling, mechanotransduction, fibrosis

Introduction

Fibrosis are characterized by excessive proliferation and transformation of fibroblast and extracellular matrix (ECM) deposition. The annual incidence of fibrotic diseases is 4,968 per 100,000 person-years, resulting in a severe burden on patients (Zhao et al., 2020). Fibrotic diseases include multiple organs, such as the liver, kidney, heart, lung, and urinary bladder (Henderson et al., 2020). Moreover, the abnormal activation of myofibroblast triggered by transforming growth factor, platelet-derived growth factor, and fibroblast growth factor is identified as a major alteration in fibrosis (Zhao et al., 2022). Hence, a large number of patients suffer from fibrotic diseases that need more effective therapies indeed.

The fibrotic process is initiated by tissue injury. Moderate injury often leads to a tissue repair process. In contrast, severe or long-term wound-healing strategies can cause fibrotic changes to tissue and organ. The etiologies vary in different situations, including hypertension, myocardial infarction, acute or chronic infection, diabetes, alcohol damage, radiation, and others (Rockey et al., 2015). In the urological system, fibrotic processes are triggered in the kidney, ureter, and urinary bladder. Bladder fibrosis is often caused by partial bladder outlet obstruction (pBOO), cystitis, and radiation (Di et al., 2022). We have investigated that mechanical cues in pBOO, including hydrostatic pressure, fluid shear stress, stretching force, and ECM stiffness, activate the bladder fibrosis process.

Yes-associated protein (YAP) has long been recognized as an intracellular mechanical transducer (Dupont et al., 2011). As an important transcriptional co-activator, YAP is also a crucial downstream effector of the Hippo signaling pathway (Moya and Halder, 2019). YAP can respond to cell geometry, density, and substrate adhesion to promote the progression of fibrotic diseases. YAP has been commonly identified in cell renewal, cell differentiation, epithelial-to-mesenchymal transition, and fibrosis (Tang et al., 2013; Ohgushi et al., 2015; Lin et al., 2018). In addition, YAP has been investigated in many mechanical cues-induced organ remodeling, such as atherosclerosis (Wang et al., 2016), orthodontic tooth movement (Deng et al., 2021), nerve regeneration (Li et al., 2021), and others.

YAP can sense the change of stretch and ECM stiffness in the urinary bladder to promote downstream gene expression. Besides ECM deposition, bladder smooth muscle proliferation is an important pathological change in the compensatory stage of bladder fibrosis (Lai et al., 2019; Chen et al., 2020). Therefore, an RNA transcriptome sequencing (RNA-seq) of YAP-knockdown (YAP-KD) and YAP CUT&Tag of human bladder smooth muscle cell (hBdSMC) were performed to investigate the key genes and regulatory network of smooth muscle proliferation in bladder fibrosis. Chromatin immunoprecipitation sequencing (ChIP-seq) is used to study transcriptional factors and target DNA. The prediction through DNA segments indicates the interaction between transcriptional factors and downstream molecules. The CUT&Tag is an improved ChIP-seq with higher quality and can be conducted with only 10^5 cells (Kaya-Okur et al., 2019). In fibrosis research, we aim to provide novel insights into YAP-associated smooth muscle proliferation.

Materials and methods

Cell line and cell culture

hBdSMCs cell line was purchased from ScienCell, United States (Cat No.4310), which was cultured with SMCM medium (ScienCell, United States, Cat No.1102) with fetal bovine serum (10%), streptomycin (100 μg/mL), penicillin (100 U/mL), and growth factor.

YAP adeno-associated virus (AAV) infection

For YAP knocking down, YAP-AAV with short hairpin RNAs (shRNAs) was obtained from *GeneChem* (Shanghai). Cells were infected by AAV and vector at a multiplicity of infection of 100 for 8 hours with minimal toxicity.

CUT&Tag sequencing

CUT&Tag sequencing was performed to analyze target genes of YAP. The sequencing technology was supported by Jiayin Biomedical Technology, Shanghai. The Raw Reads were sheltered by Trimmomatic software for Clean Reads (Bolger et al., 2014). Q20, Q30, and GC content parameters were used to assess the data quality (Supplementary Table S1).

RNA sequencing

Transcriptome sequencing and analysis between the control and YAP-KD group was carried out with the assistance of *Bioprofile*, Shanghai. Total RNA was extracted with a standard protocol of RaPure Total RNA Kit (*Magen*, Guangzhou, Cat No. R4011-02). Q20, Q30, and GC content parameters were used to assess the data quality (Supplementary Table S1).

Data analysis

The differentially expressed genes (DEGs) were clarified with the standard of adjusted $p < 0.05$, $|\log2FoldChange| \geq 1$. We intersected CUT&Tag sequencing with RNA-seq to obtain the common DEGs in both two sequencing for further analysis. To identify the function of DEGs, GO enrichment analysis and KEGG pathway enrichment analysis were performed with R package *clusterProfiler* (The R foundation; http://r-project.org; version, 4.2, United States). The protein-protein interactive (PPI) analysis of all DEGs was conducted with STRING (https://cn.string-db.org). The motif matching was performed with *HOMOR* software. *Cytoscape* software (version 3.9.1, JAVA version 11.0.6) was used to select hub genes.

Results

RNA-seq analysis of total DEGs for YAP-KD hBdSMC

To investigate whether YAP can regulate hBdSMC proliferation in bladder fibrosis, RNA-seq on the YAP-KD and nature control (NC) group with three duplicates. The results of RNA-seq indicated 141 DEGs in total, including 68 up-regulated DEGs and 73 down-regulated DEGs, which were presented in the heatmap and a volcano plot (Figures 1A, B). The top5 up-regulated and down-regulated DEGs were listed separately in Table1. To identify the function of DEGs, KEGG enrichment analysis was performed in Figure 1C showed that YAP regulated cell cycle and base excision repair pathways in hBdSMC. GO enrichment analysis was performed using all 141 DEGs. The results demonstrated that YAP engaged in DNA structure binding, meiotic cell cycle, recombination, and other proliferation-related cellular functions (Figure 1D). According to the DEGs acquired from RNA-seq, PPI analysis showed the interactions between proteins (Figure 1E).

The down-regulated DEGs regulatory network and gene functions

Apart from the analysis of total DEGs, we further analyzed up-regulated and down-regulated DEGs separately. Unfortunatey, we did not enrich enough function and signaling pathways in up-regulated genes *via* GO and KEGG analyses. Hence, we focused on down-regulated genes along with YAP-KD functions most in cell biology. GO results showed that YAP promoted the cell cycle, DNA binding, and TGFβ binding process (Figure 2A). KEGG results indicated that the DEGs enriched in base excision repair and cell

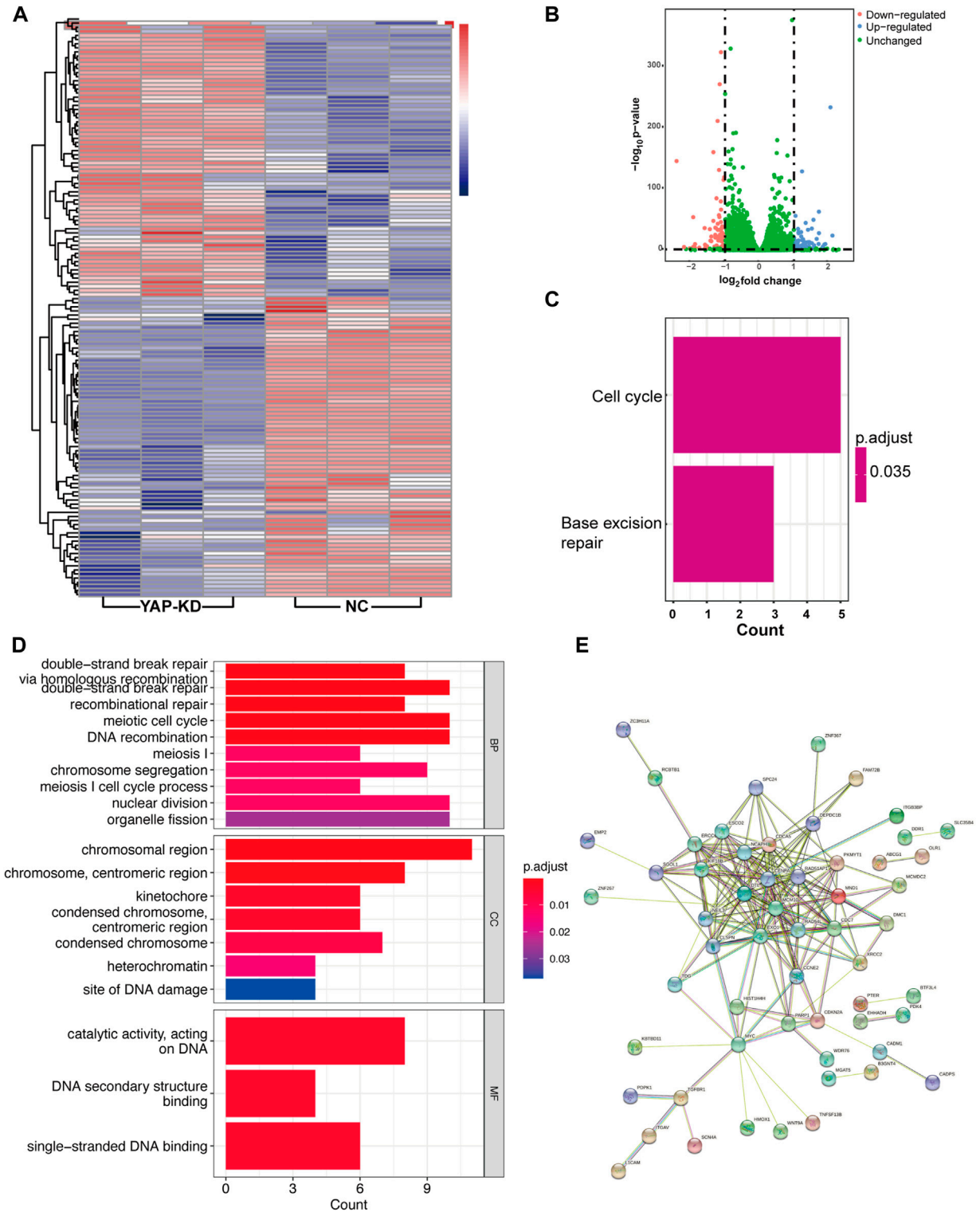


FIGURE 1
RNA-seq analysis of total DEGs for YAP-KD hBdSMC. **(A)**. The heatmap of YAP-KD and NC hBdSMCs. **(B)**. The volcano plot of DEGs. **(C)**. KEGG analysis of DEGs. **(D)**. GO analysis of DEGs. **(E)**. The PPI network of DEGs. hBdSMC: human bladder smooth muscle cell; KD: knock-down, NC: nature control, DEG: differentially expressed genes, PPI: protein-protein interaction.

process pathway (Figure 2B). The PPI network identified the interactions between down-regulated proteins (Figure 2C). The Maximum Climate Centrality (MCC) method was used to shelter the hub genes (Chin et al., 2014); the top five hub genes were RAD51AP1, CDCA5, EXO1, MCM10, and NCAPH, shown in Figure 2D; Supplementary Table S2.

TABLE 1 List of top 5 up-regulated and down-regulated DEGs of RNA-seq.

Gene_id	Gene_name	Expression_KD	Expression_NC	log2FC(KD/NC)	Adjusted P	Regulate
ENSG00000100292	HMOX1	314.9083048	75.52983456	2.059668088	2.35E-102	up
ENSG00000170801	HTRA3	16.46288202	6.991509256	1.234355134	8.50E-57	up
ENSG00000037749	MFAP3	1.594140556	0.475530677	1.724166201	4.07E-28	up
ENSG00000134717	BTF3L4	11.20840461	5.428414978	1.044609376	3.49E-25	up
ENSG00000173391	OLR1	1.924046244	0.64330886	1.565785182	2.42E-22	up
ENSG00000138448	ITGAV	13.09954398	28.48548158	−1.120115672	1.70E-141	down
ENSG00000143799	PARP1	13.70361237	30.55700841	−1.156366713	8.54E-119	down
ENSG00000122376	SHLD2	7.049022041	16.39803251	−1.216862024	1.18E-92	down
ENSG00000120253	NUP43	4.053020293	10.27518947	−1.339943896	1.79E-70	down
ENSG00000148841	NUP43	1.787120312	9.528147362	−2.408022075	3.95E-64	down

DEG, differentially expressed genes; KD, knock down; NC, nature control; FC, foldchange.

CUT&Tag sequencing analysis

As YAP is identified as a transcriptional co-activator, it plays an important role in the transcription regulatory process. To illustrate the function of YAP, immunoprecipitation for YAP CUT&Tag sequencing was performed. The Reads enrichment region is known as Peak. The heatmap of the Peak central region showed that the signals conversed near the enrichment location, indicating the data's satisfying quality (Supplementary Figure S1A). In addition, the distribution of Peaks in functional gene regions is shown in Supplementary Figure S1B. The predicted motifs were matched with HOMOR. The top 10 known sequences of predicted motifs are listed in Figure 3A. The GO annotation analysis showed that the genes enriched in positive regulation of the cellular process, metabolism, protein binding, signal transduction process, and others (Figures 3B–D). In KEGG enrichment analysis, the genes function in focal adhesion, MAPK signaling pathway, and PI3K/Akt signaling pathway that is crucial in cell development and proliferation processes (Figure 3E).

The cross-analysis of RNA-seq and CUT&Tag sequencing to identify the potential mechanism of the YAP target gene regulatory network

Not all the DEGs in RNA-seq are YAP directly interacted with. The DNA fragments that might directly interact with YAP were analyzed to identify the regulatory mechanisms of YAP. 9,875 predicted genes are binding to YAP. After combining RNA-seq and CUT&Tag sequencing, 36 DEGs were identified to bind to YAP (Figure 4A). The interactive genes between RNA-seq and CUT&Tag are listed in Supplementary Table S3. Among these genes, 10 DEGs were up-regulated, and 26 were down-regulated. The GO annotation indicated that the DEGs enriched the DNA activity process, binding process, and chromosomal region (Figure 4B). The KEGG annotation revealed that the DEGs enriched in the base excision repair process (Figure 4C). Then, a PPI analysis was conducted to clarify the interactions between the proteins (Figure 4D). The network was further analyzed in *Cytohubba, and the* MCC method selected the top five hub genes (Figure 4E). The five hub genes were identified in Table 2.

Discussion

YAP engages in many biological processes, such as cell proliferation, differentiation, apoptosis, and metabolism. In fibrotic diseases, YAP engages in fibroblast activation (Liu et al., 2015), wound healing process (Dey et al., 2020), and ECM remodeling (Zhang et al., 2022). Importantly, YAP can respond to mechanical cues to promote atherosclerosis, fibrosis, cardiac hypertrophy, muscular dystrophy, and cancer (Panciera et al., 2017). For example, the CRISPR/Cas9-mediated endothelial YAP knockdown attenuated the vascular plaque formation in mice, which indicates the association between YAP and atherosclerosis (Wang et al., 2016). The ECM stiffness triggered TGFβ-YAP signaling to promote kidney fibrosis (Szeto et al., 2016). In addition, ECM remodeling activated cardiac fibroblast and cardiac hypertrophy (Perestrelo et al., 2021). Furthermore, YAP also plays a pivotal role in mechanically sensing the urinary bladder. However, inflammatory responses, epithelial cell responses, smooth muscle cell hypertrophy, and ECM remodeling are all included in bladder fibrosis. Whether YAP functions in these processes are unclarified. Hence, the purpose of our study is to elaborate on the function of YAP of smooth muscle in bladder fibrosis progression.

We first infected the hBdSMCs with AAV in the current study to knock down YAP gene expression. Then RNA-seq was performed to illustrate YAP-regulated gene functions and pathways. The results revealed that YAP might engage in cell cycle regulation. With the highly development of high-throughput sequencing technologies, there comes great attention on multiple omics analyses to clarify human diseases (Yu et al., 2021). Hence, a CUT&Tag sequencing was conducted to identify the genes that directly bind to YAP. In the top 10 predicted genes, Fra-1-regulated transcription was reported to promote the proliferation process of breast cancer (Zhao et al., 2014). Studies demonstrated that the BATF-Jun family interacted with interferon regulatory factor 4 to promote lymphoid development

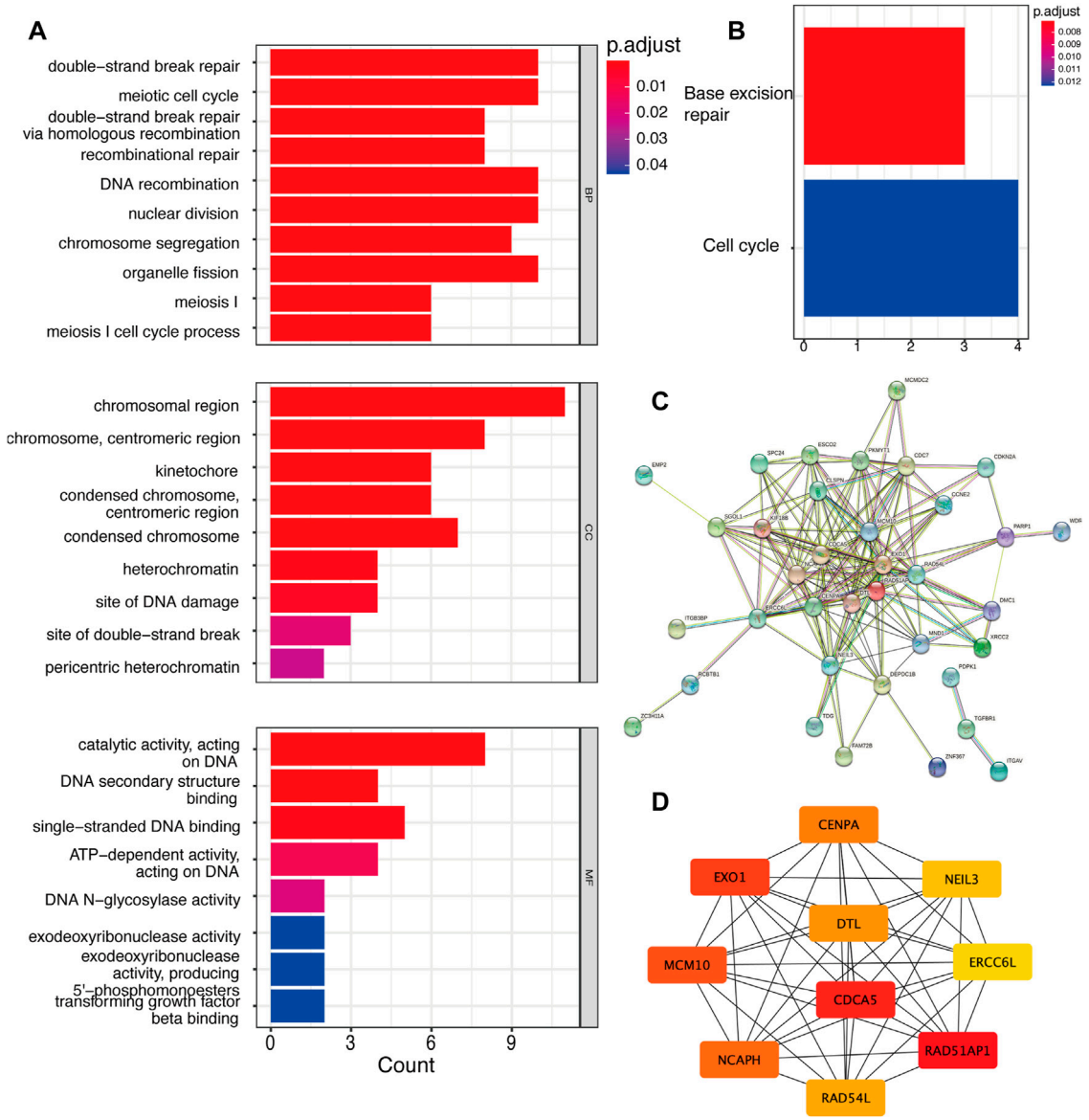


FIGURE 2
The down-regulated DEGs regulatory network and gene functions. **(A)**. GO analysis of down-regulated DEGs. **(B)**. KEGG analysis of DEGs. **(C)**. The PPI network of DEGs. **(D)**. Top 10 hub genes of down-regulated genes. DEG: differentially expressed genes, PPI: protein-protein interaction.

(Li et al., 2012). In the activator protein-1 (AP-1) transcription factor family, JunB functions dually in the cell cycle. JunB was initially recognized as a cell proliferation inhibitor. However, JunB also promotes cell division (Piechaczyk and Farràs, 2008), i,e., JunB is identified as pivotal in angiogenesis (Yoshitomi et al., 2021). Like JunB, c-Jun also belongs to the AP-1 family, which involves many biological processes, such as cell proliferation, survival, apoptosis, and tissue morphogenesis (Meng and Xia, 2011). In kidney fibrosis, G2/M-arrested proximal tubular cells facilitated c-Jun signaling to promote the production of fibrotic cytokines (Yang et al., 2010). YAP often binds to the TEA-binding domain (TEAD) family of transcription enhancers to promote downstream gene expression (Kaan et al., 2017). A recent study demonstrated that TEAD1 promoted vascular smooth muscle cells (VSMCs) through solute carrier family member 5 (SLC1A5), thereby activating mTORC1 signaling to facilitate endothelium formation (Osman et al., 2019).

After interactive matching, 36 DEGs were selected. The sheltered hub genes enriched the cell cycle process, which indicated that YAP promoted hBdSMCs proliferation in response to mechanical stimuli. The PPI network stated the top five hub genes through rankings. Coincidentally, the five hub genes are all down-regulated. Cell division cycle-associated 5 (CDCA5) has been widely studied in human cancer progression. In general, CDCA5 is identified as an oncogene and has a poor prognosis for cancers (Chang et al., 2015). For instance, CDCA5-knockdown inhibited cell proliferation, migration, and clone formation in breast cancer (Hu et al., 2022). The degradation of CDCA5 also inhibits prostate cancer progression (Luo et al., 2021). CDCA5 regulates cell proliferation through various signaling.

10.3389/fgene.2023.1106927

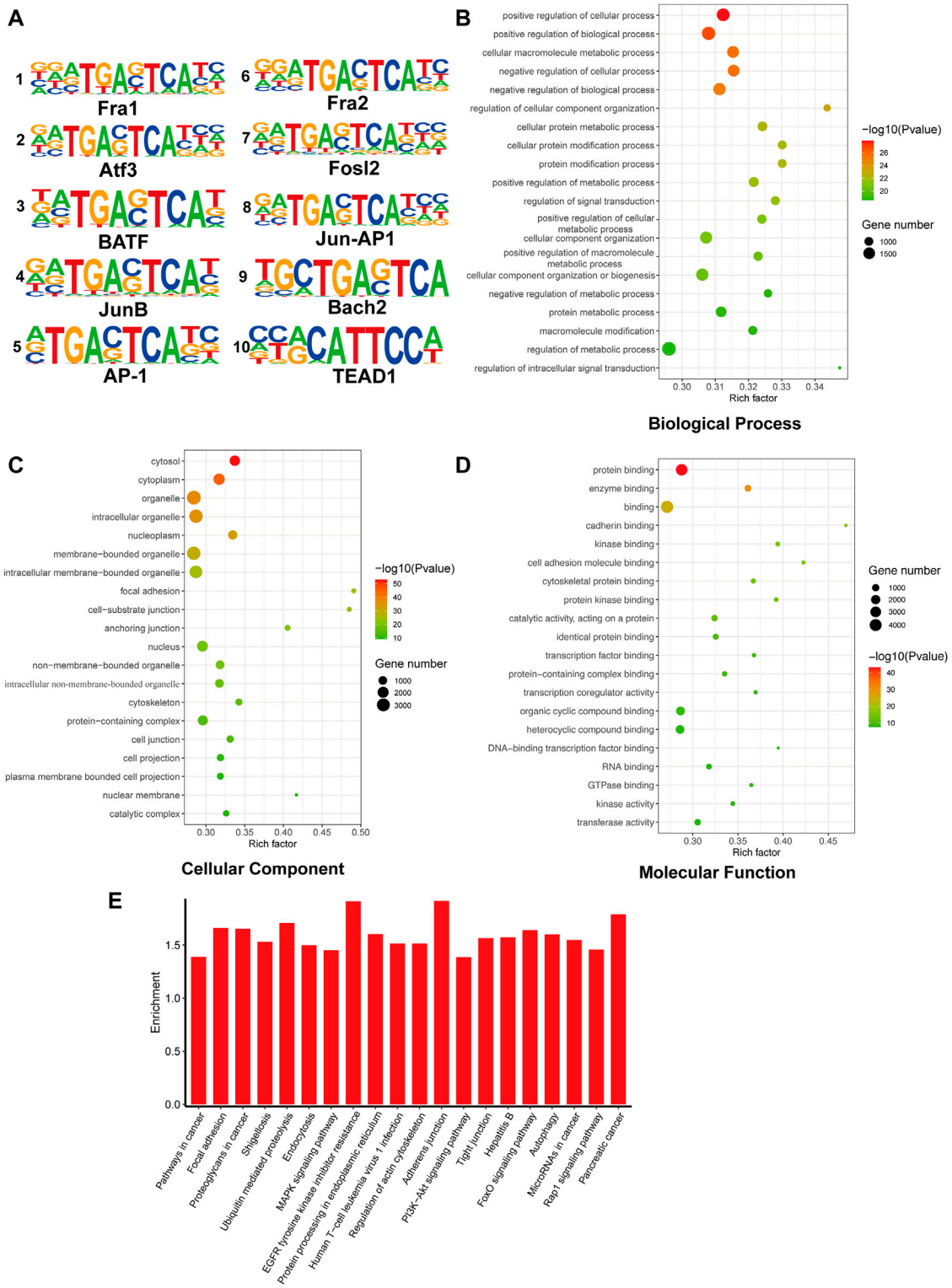


FIGURE 3
CUT&Tag seq analysis. **(A)**. The top 10 transcriptional factors binding to YAP. **(B–D)**. GO analysis of genes related to peaks. **(E)**. KEGG analysis of genes related to peaks.

Recent research illustrated that CDCA5 activated prostate cancer and colorectal cancer cell proliferation *via* ERK signaling pathway (Shen et al., 2019; Ji et al., 2021). In addition, the ablation of CDCA5 inhibited gastric cancer cell proliferation *via* downregulating Cyclin E1 expression (Zhang et al., 2018).

Centromere Protein A (CENPA) is highly correlated with cell proliferation. The centromere is a chromatin structure that provides an assembly site for cell machinery, which is essential in cell proliferation and survival (Fukagawa and Earnshaw, 2014). Mechanistically, CENPA is assembled into the centromeric

10.3389/fgene.2023.1106927

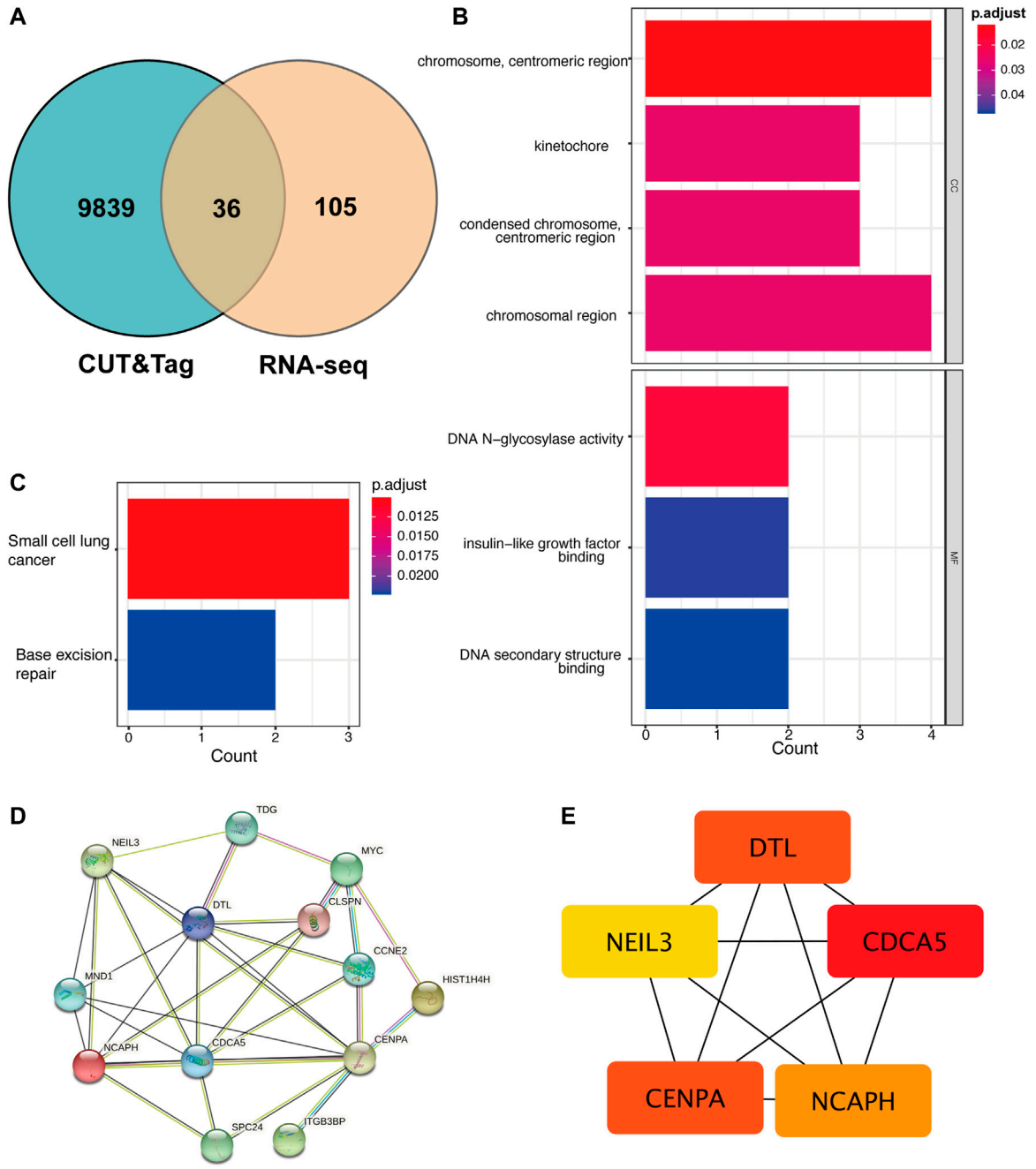


FIGURE 4
The cross-analysis of RNA-seq and CUT&Tag to identify the potential YAP target gene regulatory network mechanism. **(A)**. The Venn diagram of matching DEGs between RNA-seq and CUT&Tag seq. **(B)**. GO analysis of matching DEGs. **(C)**. KEGG analysis of matching DEGs. **(D)**. The PPI network of matching DEGs. **(E)**. Top five hub genes of matching genes. DEG: differentially expressed genes, PPI: protein-protein interaction.

chromatin in the cell cycle to the following cell cycle and generation (Wong et al., 2020). Studies demonstrated that the expression of CENPA in cardiac progenitor cells (CPCs) decreased along with aging. The expression level of CENPA is relatively higher in the early stage of life, thereby sustaining cell proliferation, inhibiting senescence, and triggering CPCs differentiation (McGregor et al., 2014). The ablation of CENPA inhibited cell proliferation in ovarian cancer (Han et al., 2021). Interestingly, circular RNAs (circRNAs), defined as crucial cancer regulators, decreased the expression of FOXM1 and promoted the expression of CENPA and CENPB to facilitate cell cycle progression (Cheng et al., 2019).

DTL is identified as CUL4-DDB1 associated factors (DCAFs), engaged in many tumorigenesis processes. Studies revealed that DTL enhanced the proliferation and migration of cancer cells in nude mice (Cui et al., 2019). In addition, the increase of DTL indicated a poor prognosis in malignant manners of bladder cancer through the mTOR/Akt signaling cascades (Luo et al., 2022). Similarly, the non-structural maintenance of chromosomes condensing I complex subunit H (NCAPH) also facilitated cell proliferation, migration, invasion, and

TABLE 2 Top 5 in PPI network of interactive genes ranked by Maximum Climate Centrality (MCC) method.

Rank	Name	Score	Regulate
1	CDCA5	138	down
2	CENPA	134	down
2	DTL	134	down
4	NCAPH	132	down
5	NEIL3	122	down

epithelial-to-mesenchymal transition (EMT) of cancer (Kim et al., 2020; Wang et al., 2020). And some microRNAs also target NCAPH to promote the degradation of β-catenin to reduce cancer stem cell maintenance (Wang et al., 2020). The fifth hub gene, the base excision repair enzyme NEIL3, plays a vital role in many biological processes, including fibrosis, lipid metabolism, tumorigenesis, and neurogenesis. For instance, in NEIL3$^{-/-}$ heart ruptured mice, the fibroblasts and myofibroblasts increased significantly, indicating that NEIL3-dependent regulation of DNA methylation affected the fibroblast proliferation and the ECM modulation (Olsen et al., 2017). Interestingly, we have found that epigenetics alterations in cell-free DNA genome were widely distributed in multiple diseases, which might be critical in early diagnosis of fibrosis and cancer disease (Yu et al., 2020). In the smooth muscle remodeling process, the depletion of DNA glycosylase NEIL3 promoted differentiation of aortic VSMCs through the Akt signaling pathway (Quiles-Jiménez et al., 2021). NEIL3 has also been reported to mediate the lipid metabolism and macrophage function in myocardial infarction (Skarpengland et al., 2016).

Collectively, the hub genes are enriched in the processes of cell proliferation and cell survival. The question is how smooth muscle proliferation behaves in bladder fibrosis. pBOO was caused by bladder wall inflammation, hypertrophy, and fibrosis (Siregar et al., 2022). Studies demonstrated that pBOO-induced bladder fibrosis was attributed to fibrosis, smooth muscle cell proliferation and hypertrophy, and urothelium proliferation (Qiao et al., 2018). In female pBOO rats, the bladder smooth muscle progresses in hematoxylin and eosin staining and is confirmed by increased bladder mass and thickness increase (Metcalfe et al., 2010). The bladder wall thickening functions in compensatory mechanisms against pathological mechanical forces in the urinary bladder, which ultimately results in a fibrotic bladder with low capacity and high pressure. Therefore, early intervention of pBOO is necessary, and knowing how to regulate these hub genes' functions might provide novel insights into preventing pBOO-induced fibrosis. In addition, a recent study demonstrated that computational framework for analyzing multi-omics profiles provided a novel direction for clinical diagnosis of many diseases (Yu et al., 2019). Combined analyses based on well-established tools may validate the results further.

Our study inevitably has some limitations. Although the CUT&Tag sequencing is more reliable than the common ChIP-seq, the low repeatability is still a common claw that cannot be solved to date. Hence, only two samples (Experiment and IgG) were applied for CUT&Tag sequencing. Furthermore, since bladder fibrosis is associated with several cell types, including urothelial cells, smooth muscle cells, and fibroblasts, the results from bladder smooth muscle cells can partly reflect the possible therapeutic targets. Further research on other cell types is needed.

Conclusion

The current study identified the potential mechanisms for YAP and its interactive hub genes in urinary bladder remodeling. The role of YAP in pBOO-induced bladder fibrosis is unclear. Our research confirmed that YAP is important in bladder smooth muscle proliferation and hypertrophy. Although pBOO-induced fibrosis is complex, the therapies targeting YAP might be a potential treatment for pBOO.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: https://www.ncbi.nlm.nih.gov/genbank/, PRJNA856421.

Author contributions

Conceptualization and Methodology: XD and ZJ; Data curation and Project administration:LX and XD; Manuscript Writing—Original draft: XD; Manuscript editing and manuscript review: ZJ and LX.

Funding

This work was supported by the Foundation of Science and Technology Department of Sichuan Province, Grant/Award Number: 2022NSFSC1571; and the Post-Doctor Research Project, West China Hospital, Sichuan University, Grant/Award Number: 2020HXBH016.

Conflict of interest

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

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2023.1106927/full#supplementary-material

References

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

Chang, I. W., Lin, V. C-H., He, H-L., Hsu, C-T., Li, C-C., Wu, W-J., et al. (2015). CDCA5 overexpression is an indicator of poor prognosis in patients with urothelial carcinomas of the upper urinary tract and urinary bladder. *Am. J. Transl. Res.* 7 (4), 710–722. doi:10.1007/s13277-015-3210-z

Chen, G., Jin, X., Luo, D., Ai, J., Xiao, K., Lai, J., et al. (2020). β-Adrenoceptor regulates contraction and inflammatory cytokine expression of human bladder smooth muscle cells via autophagy under pathological hydrostatic pressure. *Neurourol. Urodyn.* 39 (8), 2128–2138. doi:10.1002/nau.24517

Cheng, Z., Yu, C., Cui, S., Wang, H., Jin, H., Wang, C., et al. (2019). circTP63 functions as a ceRNA to promote lung squamous cell carcinoma progression by upregulating FOXM1. *Nat. Commun.* 10 (1), 3200. doi:10.1038/s41467-019-11162-4

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

Cui, H., Wang, Q., Lei, Z., Feng, M., Zhao, Z., Wang, Y., et al. (2019). DTL promotes cancer progression by PDCD4 ubiquitin-dependent degradation. *J. Exp. Clin. Cancer Res.* 38 (1), 350. doi:10.1186/s13046-019-1358-x

Deng, L., Chen, Y., Guo, J., Han, X., and Guo, Y. (2021). Roles and mechanisms of YAP/TAZ in orthodontic tooth movement. *J. Cell Physiol.* 236 (11), 7792–7800. doi:10.1002/jcp.30388

Dey, A., Varelas, X., and Guan, K-L. (2020). Targeting the Hippo pathway in cancer, fibrosis, wound healing and regenerative medicine. *Nat. Rev. Drug Discov.* 19 (7), 480–494. doi:10.1038/s41573-020-0070-z

Di, X., Jin, X., Xiang, L., Gao, X., Peng, L., Wang, W., et al. (2022). Survivin (BIRC5) regulates bladder fibrosis in a rat model of partial bladder outlet obstruction. *Chin. Med. J. (Engl).* Epub ahead of print. doi:10.1097/CM9.0000000000002022

Dupont, S., Morsut, L., Aragona, M., Enzo, E., Giulitti, S., Cordenonsi, M., et al. (2011). Role of YAP/TAZ in mechanotransduction. *Nature* 474 (7350), 179–183. doi:10.1038/nature10137

Fukagawa, T., and Earnshaw, W. C. (2014). The centromere: Chromatin foundation for the kinetochore machinery. *Dev. Cell* 30 (5), 496–508. doi:10.1016/j.devcel.2014.08.016

Han, J., Xie, R., Yang, Y., Chen, D., Liu, L., Wu, J., et al. (2021). CENPA is one of the potential key genes associated with the proliferation and prognosis of ovarian cancer based on integrated bioinformatics analysis and regulated by MYBL2. *Transl. Cancer Res.* 10 (9), 4076–4086. doi:10.21037/tcr-21-175

Henderson, N. C., Rieder, F., and Wynn, T. A. (2020). Fibrosis: From mechanisms to medicines. *Nature* 587 (7835), 555–566. doi:10.1038/s41586-020-2938-9

Hu, H., Xiang, Y., Zhang, X-Y., Deng, Y., Wan, F-J., Huang, Y., et al. (2022). CDCA5 promotes the progression of breast cancer and serves as a potential prognostic biomarker. *Oncol. Rep.* 48 (4), 172. doi:10.3892/or.2022.8387

Ji, J., Shen, T., Li, Y., Liu, Y., Shang, Z., and Niu, Y. (2021). CDCA5 promotes the progression of prostate cancer by affecting the ERK signalling pathway. *Oncol. Rep.* 45 (3), 921–932. doi:10.3892/or.2021.7920

Kaan, H. Y. K., Sim, A. Y. L., Tan, S. K. J., Verma, C., and Song, H. (2017). Targeting YAP/TAZ-TEAD protein-protein interactions using fragment-based and computational modeling approaches. *PLoS One* 12 (6), e0178381. doi:10.1371/journal.pone.0178381

Kaya-Okur, H. S., Wu, S. J., Codomo, C. A., Pledger, E. S., Bryson, T. D., Henikoff, J. G., et al. (2019). CUT&Tag for efficient epigenomic profiling of small samples and single cells. *Nat. Commun.* 10 (1), 1930. doi:10.1038/s41467-019-09982-5

Kim, B., Kim, S. W., Lim, J-Y., and Park, S-J. (2020). NCAPH is required for proliferation, migration and invasion of non-small-cell lung cancer cells. *Anticancer Res.* 40 (6), 3239–3246. doi:10.21873/anticanres.14305

Lai, J., Ai, J., Luo, D., Jin, T., Liao, B., Zhou, L., et al. (2019). β-Adrenoceptor signaling regulates proliferation and contraction of human bladder smooth muscle cells under pathological hydrostatic pressure. *J. Cell Biochem.* 120 (10), 17872–17886. doi:10.1002/jcb.29056

Li, H-X., Zhang, Z-C., and Peng, J. (2021). Low-intensity extracorporeal shock wave therapy promotes recovery of sciatic nerve injury and the role of mechanical sensitive YAP/TAZ signaling pathway for nerve regeneration. *Chin. Med. J. Engl.* 134 (22), 2710–2720. doi:10.1097/CM9.0000000000001431

Li, P., Spolski, R., Liao, W., Wang, L., Murphy, T. L., Murphy, K. M., et al. (2012). BATF-JUN is critical for IRF4-mediated transcription in T cells. *Nature* 490 (7421), 543–546. doi:10.1038/nature11530

Lin, M., Yuan, W., Su, Z., Lin, C., Huang, T., Chen, Y., et al. (2018). Yes-associated protein mediates angiotensin II-induced vascular smooth muscle cell phenotypic modulation and hypertensive vascular remodelling. *Cell Prolif.* 51 (6), e12517. doi:10.1111/cpr.12517

Liu, F., Lagares, D., Choi, K. M., Stopfer, L., Marinković, A., Vrbanac, V., et al. (2015). Mechanosignaling through YAP and TAZ drives fibroblast activation and fibrosis. *Am. J. Physiol. Lung Cell Mol. Physiol.* 308 (4), L344–L357. doi:10.1152/ajplung.00300.2014

Luo, Y., He, Z., Liu, W., Zhou, F., Liu, T., and Wang, G. (2022). DTL is a prognostic biomarker and promotes bladder cancer progression through regulating the AKT/mTOR axis. *Oxid. Med. Cell Longev.* 2022, 3369858. doi:10.1155/2022/3369858

Luo, Z., Wang, J., Zhu, Y., Sun, X., He, C., Cai, M., et al. (2021). SPOP promotes CDCA5 degradation to regulate prostate cancer progression via the AKT pathway. *Neoplasia* 23 (10), 1037–1047. doi:10.1016/j.neo.2021.08.002

McGregor, M., Hariharan, N., Joyo, A. Y., Margolis, R. L., and Sussman, M. A. (2014). CENP-A is essential for cardiac progenitor cell proliferation. *Cell Cycle* 13 (5), 739–748. doi:10.4161/cc.27549

Meng, Q., and Xia, Y. (2011). c-Jun, at the crossroad of the signaling network. *Protein Cell* 2 (11), 889–898. doi:10.1007/s13238-011-1113-3

Metcalfe, P. D., Wang, J., Jiao, H., Huang, Y., Hori, K., Moore, R. B., et al. (2010). Bladder outlet obstruction: Progression from inflammation to fibrosis. *BJU Int.* 106 (11), 1686–1694. doi:10.1111/j.1464-410X.2010.09445.x

Moya, I. M., and Halder, G. (2019). Hippo-YAP/TAZ signalling in organ regeneration and regenerative medicine. *Nat. Rev. Mol. Cell Biol.* 20 (4), 211–226. doi:10.1038/s41580-018-0086-y

Ohgushi, M., Minaguchi, M., and Sasai, Y. (2015). Rho-signaling-directed YAP/TAZ activity underlies the long-term survival and expansion of human embryonic stem cells. *Cell Stem Cell* 17 (4), 448–461. doi:10.1016/j.stem.2015.07.009

Olsen, M. B., Hildrestrand, G. A., Scheffler, K., Vinge, L. E., Alfsnes, K., Palibrk, V., et al. (2017). NEIL3-Dependent regulation of cardiac fibroblast proliferation prevents myocardial rupture. *Cell Rep.* 18 (1), 82–92. doi:10.1016/j.celrep.2016.12.009

Osman, I., He, X., Liu, J., Dong, K., Wen, T., Zhang, F., et al. (2019). TEAD1 (TEA domain transcription factor 1) promotes smooth muscle cell proliferation through upregulating SLC1A5 (solute carrier family 1 member 5)-mediated glutamine uptake. *Circ. Res.* 124 (9), 1309–1322. doi:10.1161/CIRCRESAHA.118.314187

Panciera, T., Azzolin, L., Cordenonsi, M., and Piccolo, S. (2017). Mechanobiology of YAP and TAZ in physiology and disease. *Nat. Rev. Mol. Cell Biol.* 18 (12), 758–770. doi:10.1038/nrm.2017.87

Perestrelo, A. R., Silva, A. C., Oliver-De La Cruz, J., Martino, F., Horváth, V., Caluori, G., et al. (2021). Multiscale analysis of extracellular matrix remodeling in the failing heart. *Circ. Res.* 128 (1), 24–38. doi:10.1161/CIRCRESAHA.120.317685

Piechaczyk, M., and Farràs, R. (2008). Regulation and function of JunB in cell proliferation. *Biochem. Soc. Trans.* 36 (5), 864–867. doi:10.1042/BST0360864

Qiao, L-Y., Xia, C., Shen, S., Lee, S. H., Ratz, P. H., Fraser, M. O., et al. (2018). Urinary bladder organ hypertrophy is partially regulated by Akt1-mediated protein synthesis pathway. *Life Sci.* 201, 63–71. doi:10.1016/j.lfs.2018.03.041

Quiles-Jiménez, A., Gregersen, I., Segers, F. M., Skarpengland, T., Kroustallaki, P., Yang, K., et al. (2021). DNA glycosylase Neil3 regulates vascular smooth muscle cell biology during atherosclerosis development. *Atherosclerosis* 324, 123–132. doi:10.1016/j.atherosclerosis.2021.02.023

Rockey, D. C., Bell, P. D., and Hill, J. A. (2015). Fibrosis--a common pathway to organ injury and failure. *N. Engl. J. Med.* 372 (12), 1138–1149. doi:10.1056/NEJMra1300575

Shen, A., Liu, L., Chen, H., Qi, F., Huang, Y., Lin, J., et al. (2019). Cell division cycle associated 5 promotes colorectal cancer progression by activating the ERK signaling pathway. *Oncogenesis* 8 (3), 19. doi:10.1038/s41389-019-0123-5

Siregar, S., Herlambang, M. S., Reza, M., Mustafa, A., and Stefanus, D. (2022). Role of human adipose-derived stem cells (hADSC) on TGF-β1, type I collagen, and fibrosis degree in bladder obstruction model of Wistar rats. *BMC Urol.* 22 (1), 69. doi:10.1186/s12894-022-01019-2

Skarpengland, T., Holm, S., Scheffler, K., Gregersen, I., Dahl, T. B., Suganthan, R., et al. (2016). Neil3-dependent base excision repair regulates lipid metabolism and prevents atherosclerosis in Apoe-deficient mice. *Sci. Rep.* 6, 28337. doi:10.1038/srep28337

Szeto, S. G., Narimatsu, M., Lu, M., He, X., Sidiqi, A. M., Tolosa, M. F., et al. (2016). YAP/TAZ are mechanoregulators of TGF--smad signaling and renal fibrogenesis. *J. Am. Soc. Nephrol.* 27 (10), 3117–3128. doi:10.1681/ASN.2015050499

Tang, Y., Rowe, R. G., Botvinick, E. L., Kurup, A., Putnam, A. J., Seiki, M., et al. (2013). MT1-MMP-dependent control of skeletal stem cell commitment via a β1-integrin/YAP/TAZ signaling axis. *Dev. Cell* 25 (4), 402–416. doi:10.1016/j.devcel.2013.04.011

Wang, L., Luo, J-Y., Li, B., Tian, X. Y., Chen, L-J., Huang, Y., et al. (2016). Integrin-YAP/TAZ-JNK cascade mediates atheroprotective effect of unidirectional shear flow. *Nature* 540 (7634), 579–582. doi:10.1038/nature20602

Wang, M., Qiao, X., Cooper, T., Pan, W., Liu, L., Hayball, J., et al. (2020). HPV E7-mediated NCAPH ectopic expression regulates the carcinogenesis of cervical carcinoma via PI3K/AKT/SGK pathway. *Cell Death Dis.* 11 (12), 1049. doi:10.1038/s41419-020-03244-9

Wong, C. Y. Y., Lee, B. C. H., and Yuen, K. W. Y. (2020). Epigenetic regulation of centromere function. *Cell Mol. Life Sci.* 77 (15), 2899–2917. doi:10.1007/s00018-020-03460-8

Yang, L., Besschetnova, T. Y., Brooks, C. R., Shah, J. V., and Bonventre, J. V. (2010). Epithelial cell cycle arrest in G2/M mediates kidney fibrosis after injury. *Nat. Med.* 16 (5), 535–543. doi:10.1038/nm.2144

Yoshitomi, Y., Ikeda, T., Saito-Takatsuji, H., and Yonekura, H. (2021). Emerging role of AP-1 transcription factor JunB in angiogenesis and vascular development. *Int. J. Mol. Sci.* 22 (6), 2804. doi:10.3390/ijms22062804

Yu, F., Li, K., Li, S., Liu, J., Zhang, Y., Zhou, M., et al. (2020). Cfea: A cell-free epigenome atlas in human diseases. *Nucleic Acids Res.* 48 (D1), D40–D44. doi:10.1093/nar/gkz715

Yu, F., Quan, F., Xu, J., Zhang, Y., Xie, Y., Zhang, J., et al. (2019). Breast cancer prognosis signature: Linking risk stratification to disease subtypes. *Brief. Bioinform* 20 (6), 2130–2140. doi:10.1093/bib/bby073

Yu, F., Sankaran, V. G., and Yuan, G-C. (2021). CUT&RUNTools 2.0: A pipeline for single-cell and bulk-level CUT&RUN and CUT&Tag data analysis. *Bioinformatics* 38 (1), 252–254. doi:10.1093/bioinformatics/btab507

Zhang, T., He, X., Caldwell, L., Goru, S. K., Ulloa Severino, L., Tolosa, M. F., et al. (2022). NUAK1 promotes organ fibrosis via YAP and TGF-β/SMAD signaling. *Sci. Transl. Med.* 14 (637), eaaz4028. doi:10.1126/scitranslmed.aaz4028

Zhang, Z., Shen, M., and Zhou, G. (2018). Upregulation of CDCA5 promotes gastric cancer malignant progression via influencing cyclin E1. *Biochem. Biophys. Res. Commun.* 496 (2), 482–489. doi:10.1016/j.bbrc.2018.01.046

Zhao, C., Qiao, Y., Jonsson, P., Wang, J., Xu, L., Rouhi, P., et al. (2014). Genome-wide profiling of AP-1-regulated transcription provides insights into the invasiveness of triple-negative breast cancer. *Cancer Res.* 74 (14), 3983–3994. doi:10.1158/0008-5472.CAN-13-3396

Zhao, M., Wang, L., Wang, M., Zhou, S., Lu, Y., Cui, H., et al. (2022). Targeting fibrosis, mechanisms and cilinical trials. *Signal Transduct. Target Ther.* 7 (1), 206. doi:10.1038/s41392-022-01070-3

Zhao, X., Kwan, J. Y. Y., Yip, K., Liu, P. P., and Liu, F-F. (2020). Targeting metabolic dysregulation for fibrosis therapy. *Nat. Rev. Drug Discov.* 19 (1), 57–75. doi:10.1038/s41573-019-0040-5

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 09 February 2023
DOI 10.3389/fgene.2023.1117713

OPEN ACCESS

EDITED BY
Shuai Liu,
University of Hawaii at Manoa,
United States

REVIEWED BY
Maochun Wang,
The Affiliated Hospital of Nanjing University Medical School, China
Yu Wang,
Zhejiang University School of Medicine, China

*CORRESPONDENCE
Pengyu Kong,
hrbmukpy@163.com
Jinglong Yan,
yanjinglong2020@126.com

†These authors have contributed equally to this work

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

RECEIVED 06 December 2022
ACCEPTED 01 February 2023
PUBLISHED 09 February 2023

CITATION
Xu W, Wang X, Liu D, Lin X, Wang B, Xi C, Kong P and Yan J (2023), Identification and validation of hub genes and potential drugs involved in osteoarthritis through bioinformatics analysis.
Front. Genet. 14:1117713.
doi: 10.3389/fgene.2023.1117713

Identification and validation of hub genes and potential drugs involved in osteoarthritis through bioinformatics analysis

Wenbo Xu[1†], Xuyao Wang[2†], Donghui Liu[3], Xin Lin[1], Bo Wang[1], Chunyang Xi[1], Pengyu Kong[1*] and Jinglong Yan[1*]

[1]The Second Affiliated Hospital of Harbin Medical University, Harbin, China, [2]Department of Pharmacy, Harbin Second Hospital, Harbin, China, [3]Department of Oncology, Heilongjiang Provincial Hospital, Harbin, China

Purpose: Osteoarthritis (OA) is a common degenerative disease, which still lacks specific therapeutic drugs. Synovitis is one of the most important pathological process in OA. Therefore, we aim to identify and analyze the hub genes and their related networks of OA synovium with bioinformatics tools to provide theoretical basis for potential drugs.

Materials and methods: Two datasets were obtained from GEO. DEGs and hub genes of OA synovial tissue were screened through Gene Ontology (GO) annotation, Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment as well as protein—protein interaction (PPI) network analysis. Subsequently, the correlation between expression of hub genes and ferroptosis or pyroptosis was analyzed. CeRNA regulatory network was constructed after predicting the upstream miRNAs and lncRNAs. The validation of hub genes was undertook through RT-qPCR and ELISA. Finally, potential drugs targeting pathways and hub genes were identified, followed by the validation of the effect of two potential drugs on OA.

Results: A total of 161 commom DEGs were obtained, of which 8 genes were finally identified as hub genes through GO and KEGG enrichment analysis as well as PPI network analysis. Eight genes related to ferroptosis and pyroptosis respectively were significantly correlated to the expression of hub genes. 24 miRNAs and 69 lncRNAs were identified to construct the ceRNA regulatory network. The validation of EGR1, JUN, MYC, FOSL1, and FOSL2 met the trend of bioinformatics analysis. Etanercept and Iguratimod reduced the secretion of MMP-13 and ADAMTS5 of fibroblast-like synoviocyte.

Conclusion: EGR1, JUN, MYC, FOSL1, and FOSL2 were identified as hub genes in the development of OA after series of bioinformatics analysis and validation. Etanercept and Iguratimod seemed to have opportunities to be novel drugs for OA.

KEYWORDS

osteoarthritis, hub genes, bioinformatics analysis, GEO, ferroptosis, pyroptosis

1 Background

Osteoarthritis (OA) is a common degenerative disease in middle-aged and elderly people all over the world, of which the prevalence has gradually increased due to the aging population and the trend of the overweight (Loeser et al., 2016; Roemer et al., 2022). According to the data from World Health Organization (WHO) in 2019, about 250 million people worldwide suffered from OA (Hunter and Bierma-Zeinstra, 2019). Synovitis is one of the most important pathological

manifestations in the occurrence and development of OA, running through the whole process (Sharma, 2021). Studies have shown that synovial lesions usually occur earlier than cartilage, which can be detected by MRI in several patients with small joint injury on X-ray. In addition, low-grade synovitis contributes to radiographic and pain progression (Sanchez-Lopez et al., 2022). The pathological changes of synovitis are complex and diverse, leading to the lack of specific treatment for OA. Therefore, it is necessary to explore the pathogenesis and diagnostic markers of OA from the perspective of synovium to find the therapeutic targets of OA, alleviate symptoms and promote the prognosis.

Ferroptosis is a programmed cell death mode caused by abnormal oxidation, which is regulated by glutathione peroxidase 4 (GPX4), with characteristics of iron-dependent accumulation of lipid peroxide. Kennish et al. (2014) showed that the iron level in serum of OA patients was positively correlated with the severity of OA, suggesting the existence of abnormal iron homeostasis, but its effect in synovium has not been reported. Pyroptosis is another mode of programmed cell death, which occurs when pattern recognition receptors (PRRS) induce the activation of cystine aspartic protein 1 (caspase-1) or caspase-11, which is characterized by the destruction of cell membrane and the release of cytokines. Pyroptosis has been shown to be involved in synovits, that is, Interleukin-1β (IL-1β) in the inflammatory environment derives from synovium rather than cartilage (Borgonio et al., 2014).

Bioinformatics tools have been widely used to process microarray data to determine differentially expressed genes (DEGs) and conduct various analysis. In this study, combining with bioinformatics analysis and verification, we explored and screened the hub genes of OA synovial tissue, and discussed the correlation between them and ferroptosis or pyroptosisy, then constructed the upstream ceRNA regulatory network to evaluate the reliability of these genes as the prediction or treatment targets of OA. Finally, we identified and validated potential drugs targeting pathways and hub genes through Comparative Toxicogenomics Database (CTD), Drugbank and Drugs-Genes interaction (DGI) database.

2 Materials and methods

2.1 Data collection

Gene Expression Omnibus (GEO) (https://www.ncbi.nlm.nih.gov/geo/) was used as our data source (Barrett et al., 2013), where "Osteoarthritis" was entered as the keyword in the search box for detection. The selected data set included the gene expression array of osteoarthritis and normal synovial tissue of human samples. Finally, GSE55235 (GPL96 platform), GSE55457 (GPL96 platform) were determined as the data sets of this study.

2.2 Data preprocessing and identification of DEGs

The data of GSE55235, GSE55457 were downloaded with the format of MINiML. The mRNA expression data of OA and normal synovial tissue in each data set were analyzed by limma software package of R software (Yu et al., 2012). *P*-value was analyzed in GEO to correct false-positive results, which <0.05 and |log2 (fold change)| > 1 was defined as the threshold. DEGs were then obtained, and were visualized by volcano map and heat map. common DEGs were identified after the intersection of the two datasets.

2.3 Enrichment analysis of common DEGs and PPI network construction

Gene Ontology (GO) annotation of the common DEGs was performed by the Database for Annotation, Visualization and Integrated Discovery (DAVID, version 2021) (Sherman et al., 2022), including biological process (BP), cellular component (CC), and molecular function (MF). $p < 0.05$ was determined as a significant margin for all analysis. Column chart was plotted by "http://www.bioinformatics.com.cn," a free online platform for data analysis and visualization. The Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis of the common DEGs was carried out by R package "clusterprofiler" (Version 4). Hypergeometric test was performed to evaluate the significance of pathway–pathway association: phyper (k-1, M, N-M, n, lower. tail = F). Metascape (http://metascape.org) was used for protein - protein interaction (PPI) network analysis of common DEGs (Zhou et al., 2019). PPI analysis was performed using the following databases: STRING, BioGrid (Kondo et al., 2019), OmniPath and InWeb_IM (Li et al., 2017). Physical score >0.132 was selected to be the standard of gene screening. In addition, molecular complexity detection (MCODE) was used to identify densely connected network components and obtain the hub genes of OA (Bader and Hogue, 2003). The network was visualized using Cytoscape.

2.4 Correlation between OA hub genes and ferroptosis and pyroptosis related genes

19 genes related to ferroptosis and 21 genes related to pyroptosis contained in microarray data were selected for correlation analysis with the hub genes (Liu et al., 2020a; Wu et al., 2021). Due to the small sample size of the three groups of data related to ferroptosis, the samples were combined and normalized (Zhang et al., 2019). The data was standardized using "normalize.quantiles" function in the preprocessCore package of R, and was evaluated through the box diagram. The batch effect of data was evaluated by comparing the visual PCA diagram before and after removal. The polygenic correlation map was displayed by pheatmap package of R. Spearman's correlation analysis was used to describe the correlation between quantitative variables without normal distribution. $p < 0.05$ was considered statistically significant.

2.5 Analysis of ceRNA regulatory network of hub genes

ENCORI (Li et al., 2014) (https://starbase.sysu.edu.cn) and TargetScan (McGeary et al., 2019) (https://www. targetscan. org/vert_80/) were used to predict miRNAs which regulated hub genes. The first three reliable miRNAs were selected after the intersection of the prediction. LncBase database v3.0 (Paraskevopoulou et al., 2016) (https://diana.e-ce.uth.gr/lncbasev3/interactions) was used to predict lncRNAs which regulated miRNAs above. The first

TABLE 1 Primers used for RT-qPCR amplification.

Gene	Primers	Sequence
ATF3	Forward	CCTCTGCGCTGGAATCAGTC
	Reverse	TTCTTTCTCGTCGCCTCTTTTT
EGR1	Forward	GGTCAGTGGCCTAGTGAGC
	Reverse	GTGCCGCTGAGTAAATGGGA
FOSB	Forward	GCTGCAAGATCCCCTACGAAG
	Reverse	ACGAAGAAGTGTACGAAGGGTT
FOSL1	Forward	CAGGCGGAGACTGACAAACTG
	Reverse	TCCTTCCGGGATTTTGCAGAT
FOSL2	Forward	CAGAAATTCCGGGTAGATATGCC
	Reverse	GGTATGGGTTGGACATGGAGG
JUN	Forward	TCCAAGTGCCGAAAAAGGAAG
	Reverse	CGAGTTCTGAGCTTTCAAGGT
JUNB	Forward	ACGACTCATACACAGCTACGG
	Reverse	GCTCGGTTTCAGGAGTTTGTAGT
MYC	Forward	GGCTCCTGGCAAAAGGTCA
	Reverse	CTGCGTAGTTGTGCTGATGT
GAPDH	Forward	CACTCAGACCCCCACCACAC
	Reverse	GATACATGACAAGGTGCGGCT

three reliable lncRNAs were selected for each miRNA. Finally, the ceRNA regulatory network of hub genes was constructed.

2.6 Extraction of human fibroblast-like synoviocyte (FLS)

The study was designed according to the Declaration of Helsinki, and was approved by Ethic Committee of the Second Affiliated Hospital of Harbin Medical University (KY 2021-256). Informed consent was obtained from each donor. Synovium of 3 OA patients (age 54–70 years, Kellgren-Lawrence grade 4) that underwent total joint arthroplasty (TKA) and 3 patients (age 56–68 years) that underwent meniscectomy without OA were obtained at the time after surgery. All patients were confirmed without Rheumatoid Arthritis, acute trauma, tumor or infection of knee joint. Briefly, synovium was cut into pieces at the final size about 0.5 mm*0.5 mm, and put into 0.1% type I collagenase (Biosharp, China, BS163). α-MEM medium (Cytiva, United States, SH30265.01) was added with 10% foetal bovine serum (ExCell Bio, China, FSD500) after 2 h. Primary cells could be seen climbing out after about 3–5 days. FLS of passage 6-8 (P6-8) were used in this study. 10 ng/ml of IL-1β (PEPROTECH, United States, 200-01B) were used to stimulate FLS of OA groups for 48 h in order to imitate the environment of OA, while complete medium was added into the FLS of control group. 10 μg/ml of Etanercept and Iguratimod (MedChem Express, China, HY-108847, HY-17009) were added along with IL-1β to FLS for the following test.

2.7 Screening of potential drugs for OA

CTD database (version 16766M) was used to search drugs targeting the KEGG pathways above. Drugs targeting hub genes were searched and compared using Drugbank (version 5.1.8) database and DGI database (version 4.2.0).

2.8 RT q-PCR

Trizol (Beyotime, China, R0016) was used to extract total RNA from FLS. After determining the concentration, 2 μg of total RNA was used to synthesize cDNA through cDNA synthesis kit. SYBR Green (ES Science, China, QP002) was used for qRT-PCR according to the instructions. The primer sequence of 8 hub genes and GAPDH were listed in Table 1. The mRNA level of a specific gene was calculated as $2^{-\Delta\Delta}$ Ct and normalized to GAPDH.

2.9 Elisa analysis

The supernatant was collected and centrifuged at 1,000 × g at 4°C for 10 min, then was added to 96-well plates coverd by the antibody of each protein (Elabscience, United States). After incubated for 90 min at 37°C, Biotinylated Detection Ab was added to each well. After 60 min, the plates were washed for 3 times. HRP Conjugate was then added to each well and incubated for 30 min at 37°C in the dark. 90 μL of Substrate Reagent and 50 μL Stop Solution were then added. The optical density (OD value) of each well were determined at once by microplate reader at 450 nm.

2.10 Detection of intracellular ROS

FLSs of each group were seeded at a density about 2000 cells per well in 96-well plates. The medium was replaced by H2DCFDA (Biosharp, China, BL714A) working solution with the concentration of 10 μM. FLSs were then incubated at 37°C for 2 h in the dark and washed twice to fully remove the probes outside the cells. OD value of each well was determined at once by microplate reader at 525 nm.

2.11 Detection of lipid peroxidation

FLSs of each group were broken by ultrasonic cell crusher and centrifuged at 3,500 xg for 10 min at 4°C. MDA standard solution (Solarbio, China, BC0025) and the supernatant to be tested with chromogenic agent were boiled, centrifuged, and added to a 96-well plate respectively. The absorbance was measured at 450 nm, 520 nm, and 600 nm for detection by microplate reader. Then the concentration of MDA was calculated according to the instruction.

2.12 Cell viability assay

FLSs were seeded in 96-well plates at a density at about 2000 cells per well. After adding 10 ul of CCK-8 reagent (Beyotime, China, C0037) at 24 h or 48 h respectively, the plates were put into the incubator for 30 min away from light. The absorbance of each well was measured at 450 nm by microplate reader.

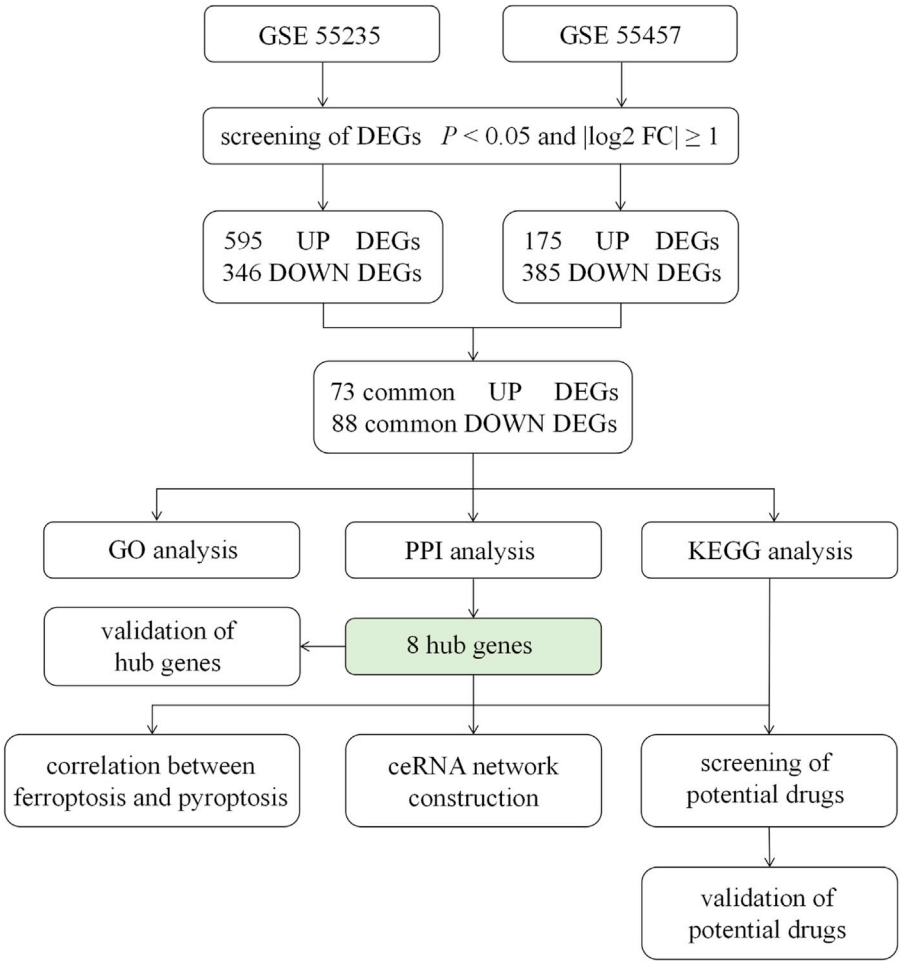


FIGURE 1
Flow chart of the study design. GSE55235, GSE55457 were determined as the data sets of this study, from where we screened DEGs by the standard of $p < 0.05$ and log2|FC| ≥ 1. After the intersection, 71 up-regulated DEGs and 88 down-regulated DEGs were obtained, followed by GO analysis, PPI network construction and KEGG pathway analysis. The correlation between hub genes and ferroptosis or pyroptosis and the construction of ceRNA were undertaken. Moreover, the expression of 8 hub genes was validated. Finally, potential drugs targeting pathways and hub genes were identified and validated. DEGs = differentially expressed genes; FC = fold change; GO = gene ontology; PPI = protein - protein interaction; KEGG = Kyoto Encyclopedia of Genes and Genomes; ceRNA = competing endogenous RNA.

2.13 Statistical analysis

All the data of validation was presented as the means ± standard deviation (SD). Statistical analysis was performed using GraphPad Prism version 6.02. Differences in numerical data between two groups were determined by *t*-test, while four groups were determined by One-way ANOVA followed by a Bonferroni post-hoc test. $p < 0.05$ was defined as statistically significant.

3 Results

3.1 Data collection and identification of DEGs

The flow chart of the study design was showed in Figure 1. A total of 20 OA synovium samples and 20 normal synovium samples were obtained from two GEO database, including 10 OA samples and 10 normal samples from GSE55235 dataset and 10 OA samples and 10 normal samples from GSE55457 dataset. Relevant clinical data were not provided in the two datasets above. After screening the differential expression of mRNA between OA and normal synovium samples, 595 up-regulated DEGs and 346 down-regulated DEGs from GSE55235 as well as 175 up-regulated DEGs and 385 down-regulated DEGs from GSE55457 were obtained (Figures 2A, B). After the intersection, a total of 161 common DEGs were obtained (Figure 2C; Table 2).

3.2 Enrichment analysis and protein-protein interaction analysis of DEGs in OA

In order to further explore the function of DEGs in OA, GO annotation and KEGG enrichment analysis were performed subsequently. As the results, 156 GO terms with $p < 0.05$ were screened out (S1). Top 5 BP, 5 CC and 4 MF terms were showed in

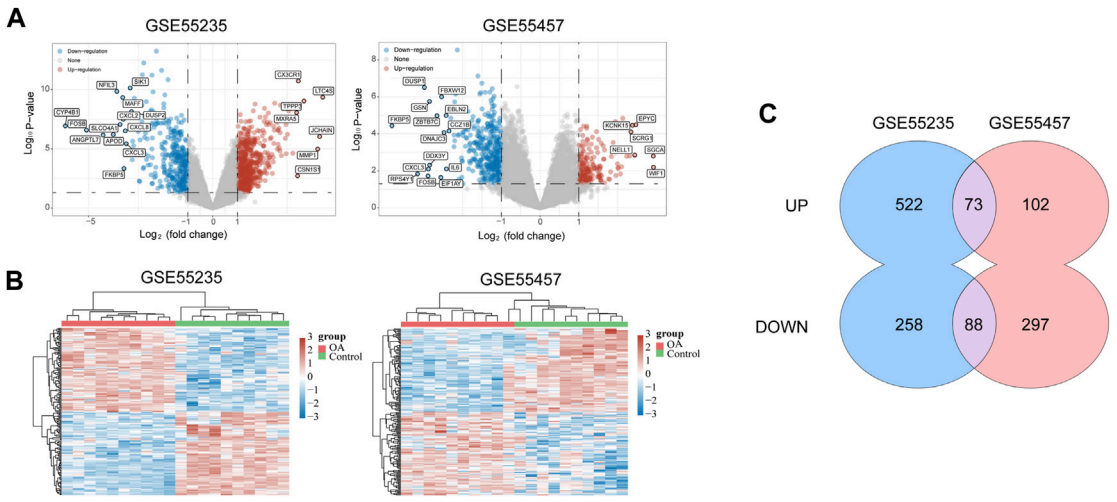


FIGURE 2
Data collection and identification of DEGs. **(A)** Volcano plot of DEGs in two datasets ($p < 0.05$ and log2|FC| ≥ 1). The red nodes represent up-regulated genes, and the blue nodes represent down-regulated genes. **(B)** Heat map of DEGs in two datasets, in which different colors represent different expression trend. Due to the large number of DEGs, the top 50 up-regulated and down-regulated genes with the largest FC were shown. **(C)** Venn diagram of two datasets which showed up-regulated and down-regulated genes respectively. DEGs = differentially expressed genes; FC = fold change.

TABLE 2 161 common DEGs of GSE 55235 and GSE 55457.

Regulation	161 common DEGs									
Up (73)	CX3CR1	LTC4S	WNT5B	ANOS1	TNFSF11	LRRC17	SCRG1	EPYC	MGAT4C	TLR7
	C1QTNF3	COPZ2	TREM2	NAP1L3	LTA4H	TRIL	SLC5A3	SLC18A2	GUCA1A	LRRC15
	TDO2	OGN	RTN1	LIPC	PNMA8A	ANKH	CAPG	RTP4	PTN	MTUS2
	PTHLH	CBR3	PTGS1	POU2AF1	NELL1	GPR1	DPT	HTR2B	DPYS	TMEM106C
	ZNF668	GPM6A	RPE65	TAC1	RGS13	NUDT11	FANCF	LCK	ZKSCAN4	MYOM2
	FGGY	GPR88	NDUFA4L2	ST8SIA1	RRAS	TNFRSF11B	MSTN	NUDT1	SIL1	TNIP3
	GLRB	STMN2	GSTZ1	CACNA2D3	ERMAP	WIF1	CLIC3	ERAP2	HSD11B2	MS4A1
	PDZRN4	APOC4	ZIC1							
Down (88)	SLC19A2	FOSL2	SIK1	NFIL3	KLF4	MAFF	GADD45B	MYC	TIPARP	APOLD1
	ZFP36	PPP1R15A	DUSP2	CDKN1A	ATF3	SOD2	BTG2	CXCL2	ADAMTS1	CCNL1
	SLC16A7	TNFAIP3	CCN1	NEDD9	ETS2	NAMPT	FOSB	SLC2A3	GRB10	NPAS2
	VEGFA	DUSP5	NR4A2	PTGS2	ZFP36L2	CXCL3	CRISPLD2	EDNRB	KLF6	STC1
	JUN	RND1	EGR1	IL1R1	JUNB	INHBB	SPRY1	HAS1	SLC2A14	FOXO3B
	IRAK3	DUSP1	SOCS3	LAMA3	NR4A1	IL6	RGS16	FKBP5	VEGFD	PFKFB3
	PHACTR1	LRCH1	ZMYM2	RGS1	ARNTL	FOXC2	SLC7A5	EFNB2	NR4A3	MCF2L
	RPS4Y1	SNAI1	CCL20	LGALSL	DDX3Y	ACACB	THBD	CCL25	DUSP4	KDM5D
	SELE	NAA15	FOSL1	USP9Y	SBNO2	KLF13	ANKRD11	NLGN4Y		

Figure 3A. It showed that the most significant BP of 161 commom DEGs was "positive regulation of transcription from RNA polymerase II promoter", the most significant CC was "extracellular space", and the most significant MF was "transcriptional activator activity, RNA polymerase II transcription regulatory region sequence-specific binding". As the result of KEGG enrichment analysis, 38 pathways with $p < 0.05$ were found (S2), and top 15 pathways were showed in Figure 3B. It showed that these genes were involved in pathways including tumor necrosis factor (TNF) signaling pathway, osteoclast differentiation, IL-17 signaling pathway, nuclear factor kappa-B (NF-

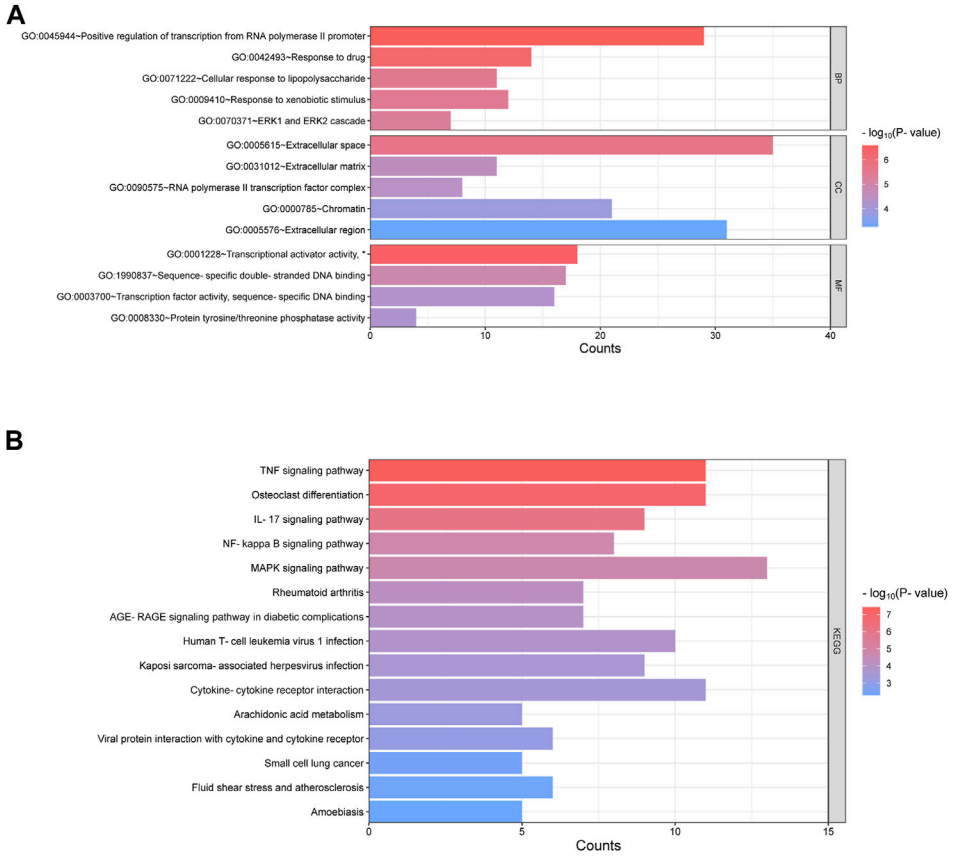


FIGURE 3
Enrichment of GO and KEGG of 161 common DEGs. **(A)** Go annotation categories of 161 common DEGs including top 5 of BP, top 5 of CC and top 4 of MF. The horizontal axis represents counts of enriched genes. **(B)** Top 15 KEGG pathway categories after enrichment analysis of 161 common DEGs. The horizontal axis represents counts of enriched genes. GO = gene ontology; BP = biological processes; CC = cellular components; MF = molecular function; KEGG = Kyoto Encyclopedia of Genes and Genomes. (*represents "RNA polymerase II transcription regulatory region sequence-specific binding", which was hided due to the space).

κB) signaling pathway and so on. In addition, as the result of PPI analysis combined with MCODE component algorithm, top 3 modules with highest MCODE score were showed in Figures 4A, B. Finally, 8 hub genes in the biggest module were identified as hub genes, which were ATF3, EGR1, FOSB, FOSL1, FOSL2, JUN, JUNB, and MYC (Table 3).

3.3 Correlation analysis between OA hub genes and ferroptosis related genes

We combined and normalized the data of 2 groups as described above, then evaluated the data standardization and batch effect through the box diagram (Figure 5A) and visual PCA diagram (Figure 5B), respectively. 19 ferroptosis related genes and 21 pyroptosis related genes were selected for correlation analysis and displayed by heat map. It showed that in the analysis of ferroptosis, ALOX15, CISD1, SAT1, and TFRC were positively correlated to hub genes, while ATP5MC3, GPX4, HSPB1, and MT1G were negatively correlated (Figure 5C). In the analysis of pyroptosis, Caspase-6, ELANE, GSDMB, IL-6 and NLRP1 were positively correlated to hub genes, while GPX4, NOD1 and PYCARD were negatively correlated (Figure 5D).

3.4 Analysis of ceRNA regulatory network of OA hub genes

A total of 24 reliable miRNAs that regulated mRNA of 8 hub genes were obtained, after the intersection of the prediction, using ENCORI and TargetScan database. Then, a total of 69 reliable lncRNAs were obtained after the analysis under LncBase v3.0 database and selection, which regulated miRNAs above. Finally, we constructed the ceRNA regulatory network of hub genes (Figure 6).

3.5 Validation of hub genes

qRT-PCR and ELISA were used to evaluate the mRNA and protein expression level of hub genes in FLS of control group and OA group. The results showed that the mRNA expression levels of JUN, MYC, EGR1, FOSL1, and FOSL2 of FLS in OA group was significantly lower than that in normal FLS ($p < 0.001$, $p < 0.01$, $p < 0.01$, $p < 0.05$, $p < 0.05$, respectively), meeting the trend of bioinformatics analysis (Figures 7A–E). The expression of JUNB and FOSB had no significant difference with control group (Figures 7F, G), while the expression of ATF3 was higher than that in normal FLS ($p < 0.001$, Figure 7H), which was opposite to bioinformatics analysis. The result of ELISA showed that the expression of

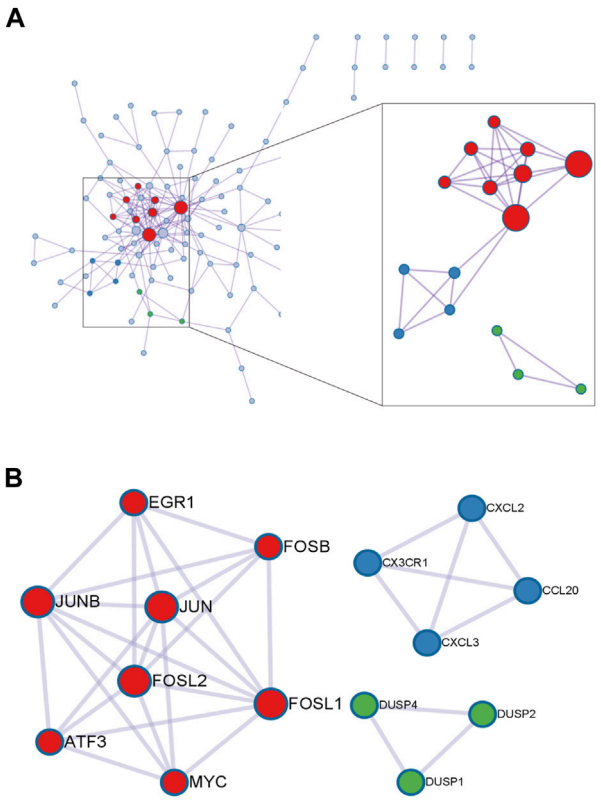


FIGURE 4
PPI analysis of 161 DEGs and screening of key modules. **(A)** Overview of the PPI network of 161 DEGs through metascape. The larger size of the points, the higher degree of the genes. **(B)** Genes in three key modules with the highest MCODE value.

EGR1, JUN, MYC were all significantly lower in the supernatant of FLS in OA group, same to the result of qRT-PCR ($p < 0.0001$, Figure 7I).

3.6 Identification of potential drugs for OA

Drugs targeting 8 hub genes and top 4 KEGG pathways were identified by CTD, DrugBank and DGI database. As the result, 85 drugs targeting 4 of hub genes were obtained, among which Quinapril targeted MYC and JUN at the same time (Figure 8A). 285 drugs were found targeting KEGG pathways, among which Elsubrutinib, Lenalidomide, Lenercept, Mifamurtide, Opinercept, Rebimastat, Tasonermin, Thalidomide and Tibulizumab targeted two pathways. Etanercept and Iguratimod targeted three pathways, which were selected for following validation (Table 4; Figure 8B).

3.7 Etanercept and Iguratimod protect FLS of OA from inflammation and cartilage degeneration

To verify the effect of Etanercept and Iguratimod on FLS, we co-treated FLS with IL-1β. By the test of cell viability, we found that Etanercept can significantly inhibit the growth rate of FLS compared with OA group ($p < 0.01$, Figure 9A). The results of intracellular ROS and MDA detection showed that the levels of these two species in Etanercept and Iguratimod groups were significantly lower than those of OA group ($p < 0.01$, $p < 0.05$, Figures 9B, C). At the same time, the mRNA expression of EGR1, JUN and MYC in Etanercept and Iguratimod group were significantly higher than those in OA group ($p < 0.001$, $p < 0.01$ for EGR1, $p < 0.001$, $p < 0.05$ for JUN and MYC, Figures 9D–F). Moreover, the level of MMP-13 and ADAMTS5 in the

TABLE 3 Features and functions of 8 hub genes in OA screened from DEGs.

Gene	Full name	Function	Regulation in OA synovium
ATF3	Activating transcription factor 3	ATF3 suppresses cyclin D1 expression in chondrocytes James et al. (2006)	Down
EGR1	Early growth response 1	EGR1 in chondrocytes could accelerate chondrocyte hypertrophy, prevent COl2A1	Down
		expression, and promote the release of inflammatory factors Sun et al. (2019)	
FOSB	FBJ murine osteosarcoma viral	FOSB promotes cell proliferation and inhibit apoptosis as a kind of proto	Down
	oncogene homolog B	oncogene Papoudou-Bai et al. (2017)	
FOSL1	FOS-like antigen 1	FOSL1 plays an oncogenic role by modulating various cellular processes Dhillon and Tulchinsky (2015)	Down
FOSL2	FOS-like antigen 2	The leucine chain encoded by FOSL2 dimerizes with the protein encoded by JUN	Down
		family to synthesize into AP1 Guo et al. (2019)	
JUN	Jun proto-oncogene	JUN family includes v-Jun, c-Jun, Jun B and Jun D.c-Jun transactivates Puma gene	Down
		expression to promote OA Lu et al. (2014)	
JUNB	Jun B proto-oncogene	Jun B positively regulates the expression of IL-4 in Th2 cells and regulates MMP13	Down
		expression Licht et al. (2006); Tanel et al. (2009)	
MYC	MYC proto-oncogene	MYC encodes transcription factors to regulate transcriptional activity and cell	Down
		proliferation, growth, and apoptosis O'Donnell et al. (2005)	

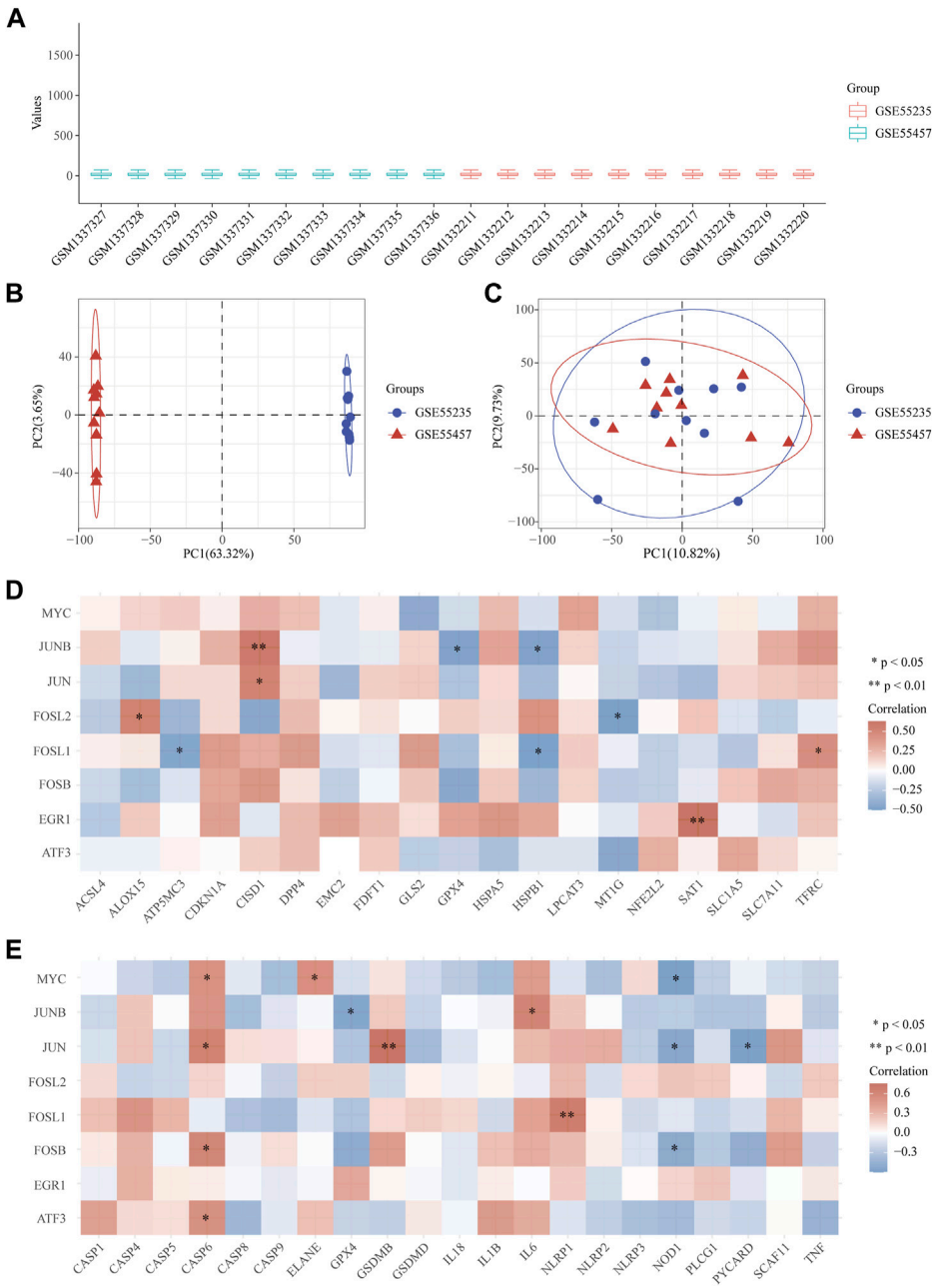


FIGURE 5
Correlation between OA hub genes and ferroptosis or pyroptosis related genes. **(A)** The data standardization evaluated by box diagram after the combination and normalization of 2 groups. **(B, C)** The visual PCA diagram before **(B)** and after **(C)** removal of the batch. **(D, E)** The heat map of the correlation between hub genes and genes related to ferroptosis and pyroptosis. The horizontal and vertical coordinates represent genes. Different colors represent the correlation coefficient (red represents positive correlation, blue represents negative correlation). PC = Principal Component.

supernatant of FLS was significantly lower in Etanercept and Iguratimod group than those in OA group ($p < 0.01$ for MMP-13, $p < 0.001$, $p < 0.01$ for ADAMTS5, Figures 9G, H).

4 Discussion

As one of the major chronic diseases endangering the middle-aged and elderly, the affected population of OA is also showing a trend of younger age (Mahmoudian et al., 2021). The compliance of non-drug treatment such as kinesitherapy and physiotherapy is usually inexact. The operations for KOA mainly includes arthroscopic debridement, high tibial osteotomy (HTO), unicondylar knee arthroplasty (UKA) and TKA (Zhang et al., 2020), which still have the problems of high risk and cost. The gathered evidence suggests that mononuclear infiltration and over expression of inflammatory mediators in synovium are seen in early OA and predate radiographic damage in OA (Sokolove and Lepus, 2013). Synovium and synovial fluid are the main contributors to inflammation that can secrete key cytokines, most of which are the main regulators of matrix metalloproteinases

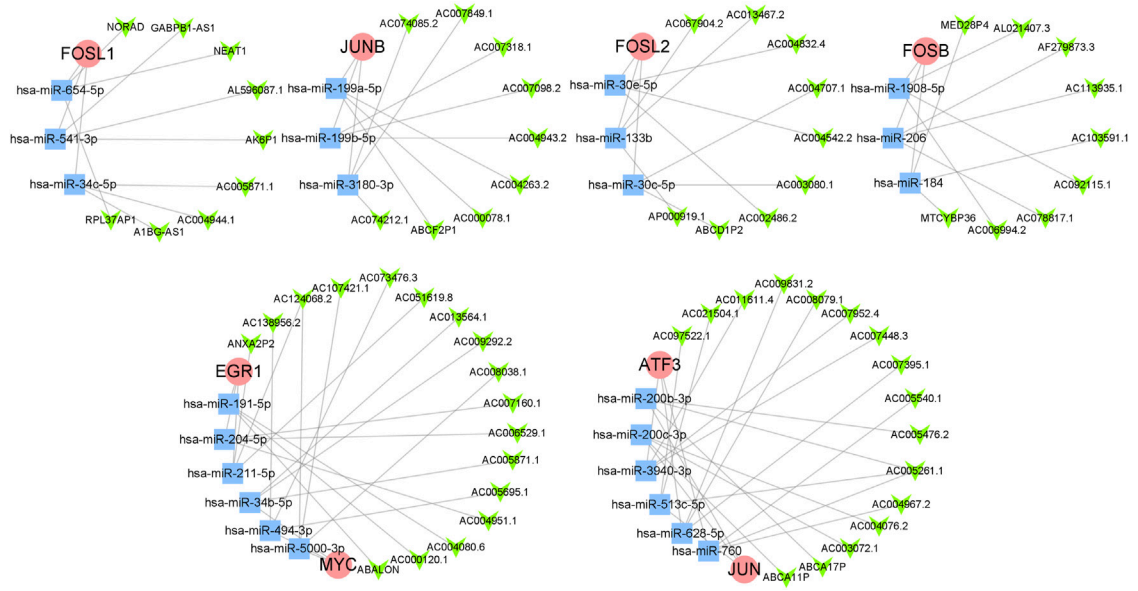


FIGURE 6
The ceRNA regulatory network of hub genes. The red circles represent hub genes, the blue squares represent miRNAs that targeted hub genes, and the green V-shapes represent lncRNAs that targeted miRNAs.

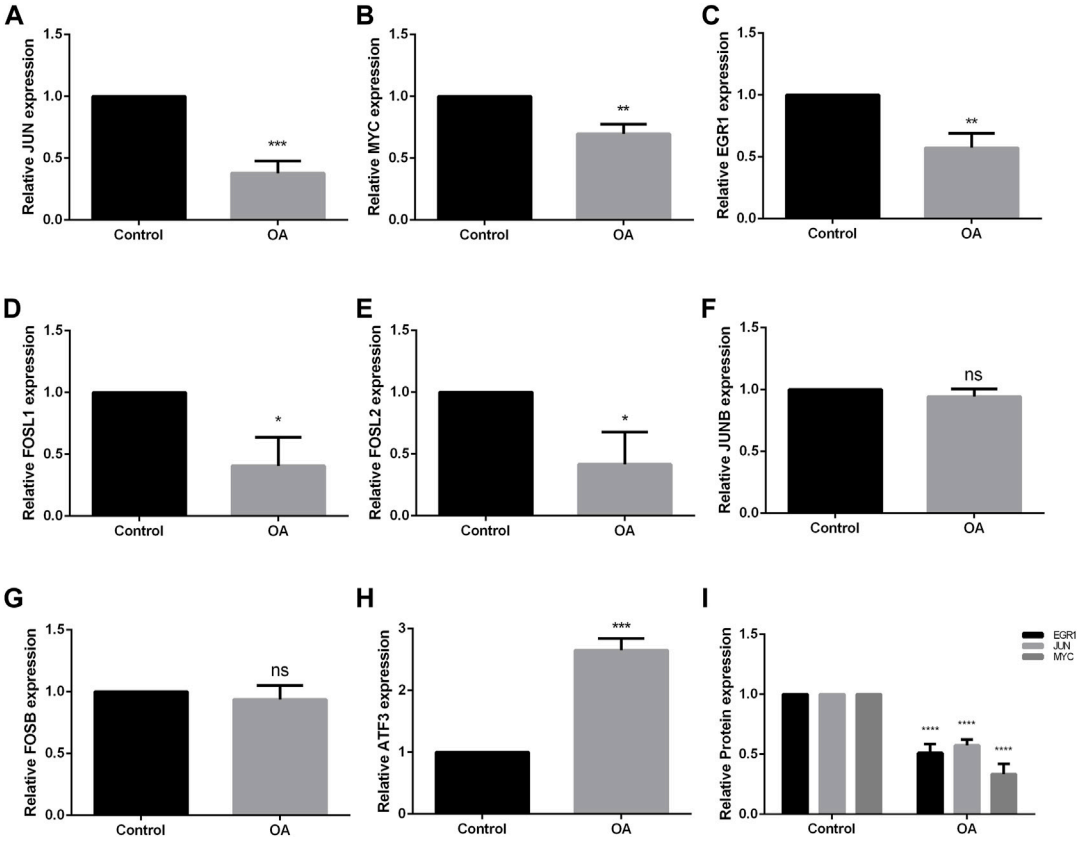


FIGURE 7
Validation of hub genes through qRT-PCR and ELISA of FLS. **(A–H)** mRNA expression of JUN, MYC, EGR1, FOSL2, FOSL1, JUNB, FOSB and ATF3. JUN, MYC, EGR1, FOSL2 and FOSL1 showed the same trend as the results of bioinformatics analysis. JUNB and FOSB showed no significant difference between two groups, while MYC showed the opposite trend. **(I)** The result of ELISA analysis of EGR1, JUN and MYC in the supernatant of FLS. OA = osteoarthritis. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, ****$p < 0.0001$. Error bars represent SD.

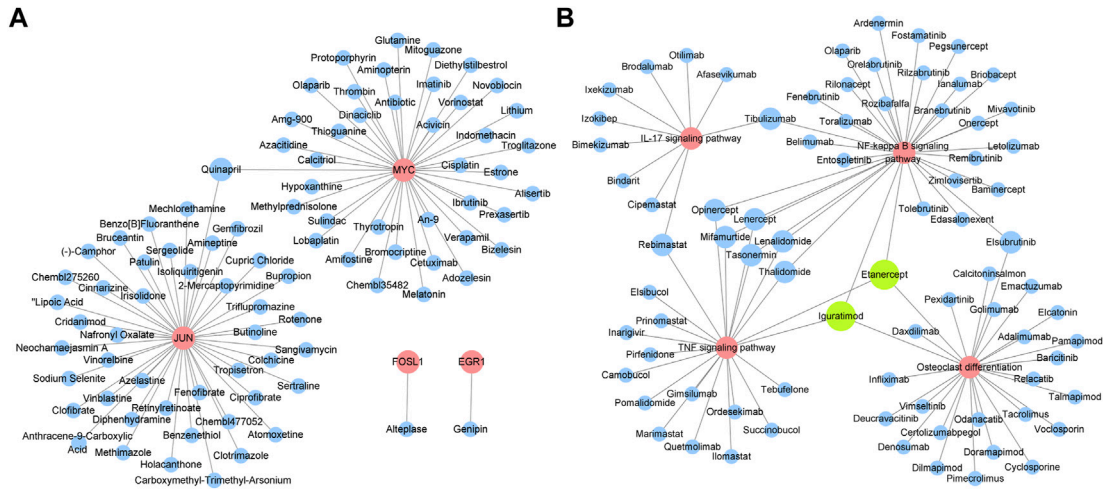


FIGURE 8
Identification of potential drugs for OA. **(A)** Drugs targeting hub genes. Red circles represent genes, blue circles represent drugs. **(B)** Drugs targeting top four KEGG pathways. Red circles represent pathways, blue circles represent drugs. Green circles represent drugs which have three targets.

TABLE 4 Drugs targeting KEGG pathways ≥2.

Drug	Target pathways	Functions or applications
Etanercept	Osteoclast differentiation	Fusion protein that binds TNF-α
	NF-kappa B signaling pathway	Treatment for severe rheumatoid arthritis
	TNF signaling pathway	Treatment for moderate to severe plaque psoriasis
Iguratimod	Osteoclast differentiation	Under investigation in rheumatoid arthritis
	NF-kappa B signaling pathway	
	TNF signaling pathway	
Elsubrutinib	Osteoclast differentiation	Bruton's tyrosine kinase inhibitor
	NF-kappa B signaling pathway	
Lenalidomide	NF-kappa B signaling pathway	Anti multiple myeloma and anemia
	TNF signaling pathway	
Lenercept	NF-kappa B signaling pathway	TNF receptor fusion protein
	TNF signaling pathway	Under investigation in multiple sclerosis
Mifamurtide	NF-kappa B signaling pathway	Antineoplasitc
	TNF signaling pathway	
Opinercept	NF-kappa B signaling pathway	TNF-α inhibitor
	TNF signaling pathway	Under investigation in rheumatoid arthritis
Rebimastat	IL-17 signaling pathway	Under investigation treatment in lung cancer and prostate cancer
	TNF signaling pathway	
Tasonermin	NF-kappa B signaling pathway	Recombinant soluble form of TNF-α
	TNF signaling pathway	An adjunt to surgery to remove soft tissue sarcomas of the limbs
Thalidomide	NF-kappa B signaling pathway	Antineoplasitc
	TNF signaling pathway	
Tibulizumab	IL-17 signaling pathway	Bispecific antibody targeting BAFF and IL-17A
	NF-kappa B signaling pathway	

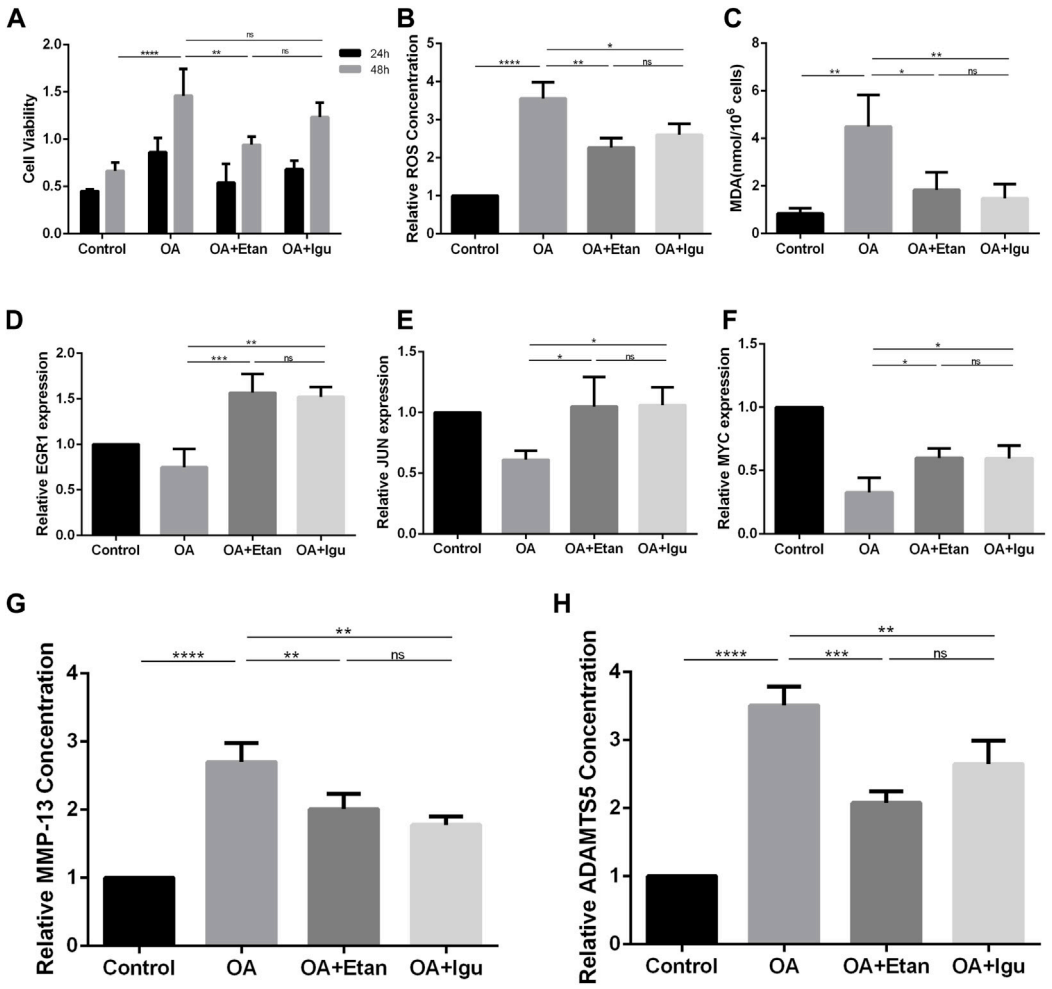


FIGURE 9
Validation of the effect of Etanercept and Iguratimod on FLS. **(A)** Cell viability of FLS at 24 and 48 h. **(B)** Relative Intracellular ROS concentration of FLS. **(C)** MDA concentration of FLS of different groups. **(D–F)** mRNA expression of EGR1, JUN and MYC of the four groups. **(G, H)** The level of MMP-13 and ADAMTS5 in the supernatant of FLS in different groups. OA = osteoarthritis; Etan = Etanercept; Igu = Iguratimod; ROS = reactive oxygen species; MDA = Malonic dialdehyde; MMP-13 = matrix metalloproteinase 13; ADAMTS5 = a disintegrin and metalloproteinase with thrombospondin motifs-5. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, ****$p < 0.0001$. Error bars represent SD.

(MMP), resulting in the loss of cartilage (Kulkarni et al., 2021). Therefore, in-depth study of the molecular mechanism of synovitis is of great significance to the prevention and treatment for OA.

In this study, we screened out 8 hub genes in synovim that may be involved in the progression of OA through a series of bioinformatics methods. EGR1 is potentially involved in postnatal bone biology and implicated in the regulation of osteoblastic cell growth and differentiation. Studies have shown that when EGR1 was inhibited, the bone mass of the limbs decreased in mice, with lower bone volume fraction and mineral density, same to our study (Reumann et al., 2011). However, over-expression of EGR1 regulated the expression of KLF5 and β-catenin signaling pathway, leading to the acceleration of cartilage hypertrophy and degeneration (Sun et al., 2019). MYC encodes transcription factors to regulate transcriptional activity and cell proliferation, growth, and had also been found to regulate OA process through multiple pathways. Wu et al. (O'Donnell et al., 2005) found that the effects of miR-24 on OA chondrocytes may be achieved by targeting MYC and further regulating the MAPK signaling pathway. The silence of c-MYC could promote proliferation of rat chondrocyte (Zou et al., 2018). ATF3 deficiency in chondrocytes had been reported to alleviates OA development (Iezaki et al., 2016). This effect may be achieved by ATF3 directly affecting the expression of MMP-13 thus reducing cartilage loss (Chan et al., 2017). However, There is little study on the mechanism of ATF3 in OA synoviocyte, except for the bioinformatics analysis.

The rest five hub genes, which were FOSB, FOSL1, FOSL2, JUN and JUNB, belonged to the AP-1 family. AP-1 was thought to be critically involved in the pathogenesis of arthritis due to the binding activity for its cognate recognition sites in the promoters of inflammatory cytokines and matrix-degrading enzymes (Huber et al., 2019). Study showed that mitochondrial dysfunction triggered a catabolic response in chondrocytes *via* activation of the JNK/AP1 pathway (Ansari et al., 2020). On the other hand, JUNB-FBXO21-ERK axis promoted cartilage degeneration in OA by inhibiting autophagy (Lin et al., 2021).

As independent forms of programmed death, ferroptosis and pyroptosis have been shown to play essential roles in the pathological processes of tumor, Alzheimer's disease, cerebral

hemorrhage, ischemia-reperfusion injury, OA and so on (Kenny et al., 2019; Liang et al., 2019; Liu et al., 2020b). In this study, we selected several ferroptosis and pyroptosis related genes. Although these genes were not differentially expressed in the two datasets, analyzing their relationship with hub genes will be still beneficial to explore the mechanism of FLS participating in OA. We found that FOSL1 and JUNB were associated with 3 of the 19 ferroptosis related genes. Overexpression of c-JUN inhibits ferroptosis induced by erastin in Schwann cells to promotes the rehabilitation of facial nerve function (Gao et al., 2022). Vertically, the strongest two genes related to hub genes were CISD1 and HSPB1. CISD1, an iron-containing outer mitochondrial membrane protein, inhibits ferroptosis by protection against mitochondrial lipid peroxidation (Yuan et al., 2016). The phosphorylation of HSPB1 at the Ser-15 site induced by Erastin is the key of the protective response to ferroptosis stress (Wang et al., 2022). In addition, over-expression of HSPB1 attenuates ferroptosis in rats through promoting G6PD expression (Dai and Hu, 2022). Therefore, we believe that CISD1 and HSPB1 are worth studying in the future to explore the mechanism of ferroptosis in OA synovium. In addition, we found that JUN was related to the most pyroptosis related genes. Study had showed that the Jun N-terminal kinases (JNK) pathway is largely upstream of the NLRP3 inflammasome, which exerts a crucial regulatory impact on microglia pyroptosis and inflammatory responses (He et al., 2021). However, the relationship between JUN and pyroptosis of FLS still needs further study. As a key regulator of innate immunity and inflammasome activation, caspase-6 promotes the activation of pyroptosis (Zheng et al., 2020).

Studies have found the significance of miRNAs and lncRNAs on OA. Xu et al. (2021) showed that SNHG7 ameliorated the development of OA by suppressing apoptosis through miR-214-5p-PPARGC1B-PPARγ axis. Wang et al. (2019) found that lncRNA FOXD2 Adjacent Opposite Strand RNA 1 (FOXD2-AS1) served as the protector for OA patients by inducing chondrocyte proliferation. The downregulation of lncRNA LOC101928134, which acts as a promoter of OA, can block the process of OA (Yang et al., 2019). Here, we built a ceRNA network after predicting the upstream miRNAs and lncRNAs, in order to provide more possible early biomarkers for diagnosis, or targeting drugs for OA patients.

Among the 11 drugs targeting two or more than two KEGG pathways, 10 acted on NF-κB signaling pathway, 9 acted on TNF signaling pathway, 3 acted on Osteoclast differentiation pathway, and 2 acted on IL-17 signaling pathway, suggesting that NF-κB was still the key point for OA treatment. Moreover, Etanercept and Iguratimod related to three pathways. Etanercept, a soluble fusion protein that binds TNF-α, has been proved to be an effective choice targeting several inflammatory diseases, especially rheumatoid arthritis (Graudal et al., 2015). Recent study showed that Etanercept lead to the repairment of cartilage with a special scaffold (Campos et al., 2022), and the reduction of MMP-3 (Kroon et al., 2020), suggesting the potential effect on OA. The effect of Iguratimod on OA still needs further study. Our study proved that both Etanercept and Iguratimod could reduce level of intracellular ROS and MDA of FLS as well as level of MMP-13 and ADAMTS5 in the supernatant. In addition, Etanercept could significantly alleviate the abnormal proliferation of FLS under OA environment. Drugs targeted to the hub genes can be divided into antineoplastic, antipsychotic, anti-inflammatory, anti-hyperglycemia, anti-hyperlipidemia, etc. Their roles in OA still needs to be explored.

Our study still had some limitations. First, we obtained gene arrays in synovial tissues of OA patients from GEO database. However, due to the lack of information of patients, it was hard to correlate the DEGs and hub genes obtained in this study with the severity or grade of OA. Second, the sample size of each group was not so large, which was needed for further research.

In summary, our study aimed to identify key genes involved in the pathophysiology of OA. 161 common DEGs and 8 hub genes were screened through GO analysis, KEGG pathway analysis, PPI network construction as well as MCODE, which may become potential targeting clinical diagnosis and treatment of OA. Furthermore, 5 of 8 genes met the similar expression trend with our result through the validation. Subsequently, we discussed the relationship between hub genes and key genes of ferroptosis and pyroptosis. Moreover, miRNAs and lncRNAs were identified to construct the ceRNA regulatory network of hub genes. Finally, we found that Etanercept and Iguratimod, as top two of potential drugs targeting KEGG pathways, had the protective effect on FLS in the OA environment.

5 Conclusion

EGR1, JUN, MYC, FOSL1, and FOSL2 were identified and validated as hub genes in the development of OA after series of bioinformatics analysis. They may have effect on OA development through different kinds of pathways, including the process of ferroptosis and pyroptosis. Etanercept and Iguratimod seemed to have stronger opportunities to be novel drugs for OA.

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 studies involving human participants were reviewed and approved by Ethic Committee of the Second Affiliated Hospital of Harbin Medical University (KY2021-256). The patients/participants provided their written informed consent to participate in this study.

Author contributions

DL, PK, and JY: conception and design. WX, XL, and XW: data collection. DL, CX, and WX: data analysis and interpretation. BW, WX, and DL: manuscript writing and validation. All authors have approved the final manuscript.

Funding

This work was supported by the Research & Practice Innovation Program of Harbin Medical University (No. 2020-KYYWF-1447) and National Natural Science Foundation of China (No. 82072472).

Conflict of interest

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

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2023.1117713/full#supplementary-material

References

Ansari, M. Y., Ahmad, N., Voleti, S., Wase, S. J., Novak, K., Haqqi, T. M., et al. (2020). Mitochondrial dysfunction triggers a catabolic response in chondrocytes via ROS-mediated activation of the JNK/AP1 pathway. *J. Of Cell Sci.* 133 (22), jcs247353. doi:10.1242/jcs.247353

Bader, G. D., and Hogue, C. V. (2003). An automated method for finding molecular complexes in large protein interaction networks. *BMC Bioinforma.* 4, 2. doi:10.1186/1471-2105-4-2

Barrett, T., Wilhite, S. E., Ledoux, P., Evangelista, C., Kim, I. F., Tomashevsky, M., et al. (2013). NCBI GEO: Archive for functional genomics data sets-update. *Nucleic Acids Res.* 41, D991–D995. doi:10.1093/nar/gks1193

Borgonio, C., González-Huerta, N. C., Romero-Córdoba, S., Hidalgo-Miranda, A., and Miranda-Duarte, A. (2014). Altered expression of circulating microRNA in plasma of patients with primary osteoarthritis and *in silico* analysis of their pathways. *PloS one* 9 (6), e97690. doi:10.1371/journal.pone.0097690

Campos, Y., Fuentes, G., Almirall, A., Que, I., Schomann, T., Chung, C. K., et al. (2022). The incorporation of Etanercept into a porous tri-layer scaffold for restoring and repairing cartilage tissue. *Pharmaceutics* 14 (2), 282. doi:10.3390/pharmaceutics14020282

Chan, C., Macdonald, C., Litherland, G., Wilkinson, D. J., Skelton, A., Europe-Finner, G. N., et al. (2017). Cytokine-induced MMP13 expression in human chondrocytes is dependent on activating transcription factor 3 (ATF3) regulation. *J. Of Biol. Chem.* 292 (5), 1625–1636. doi:10.1074/jbc.M116.756601

Dai, Y., and Hu, L. (2022). HSPB1 overexpression improves hypoxic-ischemic brain damage by attenuating ferroptosis in rats through promoting G6PD expression. *J. Of Neurophysiology* 128 (6), 1507–1517. doi:10.1152/jn.00306.2022

Dhillon, A. S., and Tulchinsky, E. (2015). FRA-1 as a driver of tumour heterogeneity: A nexus between oncogenes and embryonic signalling pathways in cancer. *Oncogene* 34, 4421–4428. doi:10.1038/onc.2014.374

Gao, D., Huang, Y., Sun, X., Yang, J., Chen, J., and He, J. (2022). Overexpression of c-Jun inhibits erastin-induced ferroptosis in Schwann cells and promotes repair of facial nerve function. *J. Of Cell. And Mol. Med.* 26 (8), 2191–2204. doi:10.1111/jcmm.17241

Graudal, N., Jürgens, G., and Combe, B. (2015). Etanercept tapering in rheumatoid arthritis. *N. Engl. J. Of Med.* 372 (5), 489–490. doi:10.1056/NEJMc1414787

Guo, J., Shen, S., Liu, X., Ruan, X., Zheng, J., Liu, Y., et al. (2019). Role of linc00174/miR-138-5p (miR-150-5p)/*FOSL2* feedback loop on regulating the blood-tumor barrier permeability. *Mol. Ther. Nucleic Acids.* 18, 1072–1090. doi:10.1016/j.omtn.2019.10.031

He, J., Liu, T., Li, Y., Mi, X., Han, D., Yang, N., et al. (2021). JNK inhibition alleviates delayed neurocognitive recovery after surgery by limiting microglia pyroptosis. *Int. Immunopharmacol.* 99, 107962. doi:10.1016/j.intimp.2021.107962

Huber, R., Kirsten, H., Näkki, A., Pohlers, D., Thude, H., Eidner, T., et al. (2019). Association of human FOS promoter variants with the occurrence of knee-osteoarthritis in a case control association study. *Int. J. Mol. Sci.* 20 (6), 1382. doi:10.3390/ijms20061382

Hunter, D. J., and Bierma-Zeinstra, S. (2019). *Osteoarthr. Lancet* 393 (10182), 1745–1759. doi:10.1016/S0140-6736(19)30417-9

Iezaki, T., Ozaki, K., Fukasawa, K., Inoue, M., Kitajima, S., Muneta, T., et al. (2016). ATF3 deficiency in chondrocytes alleviates osteoarthritis development. *J. Of Pathology* 239 (4), 426–437. doi:10.1002/path.4739

James, C. G., Woods, A., Underhill, T. M., and Beier, F. (2006). The transcription factor ATF3 is upregulated during chondrocyte differentiation and represses cyclin D1 and A gene transcription. *BMC Mol. Biol.* 7, 30. doi:10.1186/1471-2199-7-30

Kennish, L., Attur, M., Oh, C., Krasnokutsky, S., Samuels, J., Greenberg, J. D., et al. (2014). Age-dependent ferritin elevations and HFE C282Y mutation as risk factors for symptomatic knee osteoarthritis in males: A longitudinal cohort study. *BMC Musculoskelet. Disord.* 15, 8. doi:10.1186/1471-2474-15-8

Kenny, E. M., Fidan, E., Yang, Q., Anthonymuthu, T. S., New, L. A., Meyer, E. A., et al. (2019). Ferroptosis contributes to neuronal death and functional outcome after traumatic brain injury. *Crit. Care Med.* 47 (3), 410–418. doi:10.1097/CCM.0000000000003555

Kondo, K., Ohigashi, I., and Takahama, Y. (2019). Thymus machinery for T-cell selection. *Int. Immunol.* 31 (3), 119–125. doi:10.1093/intimm/dxy081

Kroon, F., Bay-Jensen, A. C., Wittoek, R., VerbruGGen, G., Smolen, J. S., Kloppenburg, M., et al. (2020). Etanercept therapy leads to reductions in matrix metalloproteinase-3 in patients with erosive hand osteoarthritis. *Scand. J. Of Rheumatology* 49 (2), 167–168. doi:10.1080/03009742.2019.1657493

Kulkarni, P., Martson, A., Vidya, R., Chitnavis, S., and Harsulkar, A. (2021). Pathophysiological landscape of osteoarthritis. *Adv. Clin. Chem.* 100, 37–90. doi:10.1016/bs.acc.2020.04.002

Li, J. H., Liu, S., Zhou, H., Qu, L. H., and Yang, J. H. (2014). starBase v2.0: decoding miRNA-ceRNA, miRNA-ncRNA and protein-RNA interaction networks from large-scale CLIP-Seq data. *Nucleic Acids Res.* 42, D92–D97. doi:10.1093/nar/gkt1248

Li, T., Wernersson, R., Hansen, R. B., Horn, H., Mercer, J., Slodkowicz, G., et al. (2017). A scored human protein-protein interaction network to catalyze genomic interpretation. *Nat. METHODS* 14 (1), 61–64. doi:10.1038/nmeth.4083

Liang, C., Zhang, X., Yang, M., and Dong, X. (2019). Recent progress in ferroptosis inducers for cancer therapy. *Adv. Mater.* 31 (51), e1904197. doi:10.1002/adma.201904197

Licht, A. H., Pein, O. T., Florin, L., Hartenstein, B., Reuter, H., Arnold, B., et al. (2006). JunB is required for endothelial cell morphogenesis by regulating core-binding factor beta. *J. Of Cell Biol.* 175 (6), 981–991. doi:10.1083/jcb.200605149

Lin, Z., Miao, J., Zhang, T., He, M., Wang, Z., Feng, X., et al. (2021). JUNB-FBXO21-ERK axis promotes cartilage degeneration in osteoarthritis by inhibiting autophagy. *Aging Cell* 20 (2), e13306. doi:10.1111/acel.13306

Liu, P., Feng, Y., Li, H., Chen, X., Wang, G., Xu, S., et al. (2020). Ferrostatin-1 alleviates lipopolysaccharide-induced acute lung injury via inhibiting ferroptosis. *Cell Mol. Biol. Lett.* 25 (10), 10. doi:10.1186/s11658-020-00205-0

Liu, Z., Zhao, Q., Zuo, Z. X., Yuan, S. Q., Yu, K., Zhang, Q., et al. (2020). Systematic analysis of the aberrances and functional implications of ferroptosis in cancer. *iScience* 23 (7), 101302. doi:10.1016/j.isci.2020.101302

Loeser, R. F., Collins, J. A., and Diekman, B. O. (2016). Ageing and the pathogenesis of osteoarthritis. *Nat. Rev. Rheumatol.* 12 (7), 412–420. doi:10.1038/nrrheum.2016.65

Lu, H., Hou, G., Zhang, Y., Dai, Y., and Zhao, H. (2014). c-Jun transactivates Puma gene expression to promote osteoarthritis. *Mol. Med. Rep.* 9 (5), 1606–1612. doi:10.3892/mmr.2014.1981

Mahmoudian, A., Lohmander, L. S., Mobasheri, A., Englund, M., and Luyten, F. P. (2021). Early-stage symptomatic osteoarthritis of the knee - time for action. *Nat. Rev. Rheumatol.* 17 (10), 621–632. doi:10.1038/s41584-021-00673-4

McGeary, S. E., Lin, K. S., Shi, C. Y., Pham, T. M., Bisaria, N., Kelley, G. M., et al. (2019). The biochemical basis of microRNA targeting efficacy. *Science* 366 (6472), eaav1741. doi:10.1126/science.aav1741

O'Donnell, K. A., Wentzel, E. A., Zeller, K. I., Dang, C. V., and Mendell, J. T. (2005). c-Myc-regulated microRNAs modulate E2F1 expression. *Nature* 435 (7043), 839–843. doi:10.1038/nature03677

Papoudou-Bai, A., Hatzimichael, E., Barbouti, A., and Kanavaros, P. (2017). Expression patterns of the activator protein-1 (AP-1) family members in lymphoid neoplasms. *Clin. And Exp. Med.* 17 (3), 291–304. doi:10.1007/s10238-016-0436-z

Paraskevopoulou, M. D., Vlachos, I. S., Karagkouni, D., Georgakilas, G., Kanellos, I., Vergoulis, T., et al. (2016). DIANA-LncBase v2: Indexing microRNA targets on non-coding transcripts. *Nucleic Acids Res.* 44 (D1), D231–D238. doi:10.1093/nar/gkv1270

Reumann, M. K., Strachna, O., Yagerman, S., Torrecilla, D., Kim, J., Doty, S. B., et al. (2011). Loss of transcription factor early growth response gene 1 results in impaired endochondral bone repair. *BONE* 49 (4), 743–752. doi:10.1016/j.bone.2011.06.023

Roemer, F. W., Guermazi, A., Hannon, M. J., Fujii, T., Omoumi, P., Hunter, D. J., et al. (2022). Presence of magnetic resonance imaging- defined inflammation particularly in overweight and obese women increases risk of radiographic knee osteoarthritis: The POMA study. *Arthritis Care & Res.* 74 (8), 1391–1398. doi:10.1002/acr.24568

Sanchez-Lopez, E., Coras, R., Torres, A., Lane, N. E., and Guma, M. (2022). Synovial inflammation in osteoarthritis progression. *Nat. Rev. Rheumatol.* 18 (5), 258–275. doi:10.1038/s41584-022-00749-9

Sharma, L. (2021). Osteoarthritis of the knee. *N. Engl. J. Of Med.* 384 (1), 51–59. doi:10.1056/NEJMcp1903768

Sherman, B. T., Hao, M., Qiu, J., Jiao, X., Baseler, M. W., Lane, H. C., et al. (2022). David: A web server for functional enrichment analysis and functional annotation of gene lists (2021 update). *Nucleic Acids Res.* 50, W216–W221. doi:10.1093/nar/gkac194

Sokolove, J., and Lepus, C. M. (2013). Role of inflammation in the pathogenesis of osteoarthritis: Latest findings and interpretations. *Ther. Adv. Musculoskelet. Dis.* 5 (2), 77–94. doi:10.1177/1759720X12467868

Sun, X., Huang, H., Pan, X., Li, S., Xie, Z., Ma, Y., et al. (2019). EGR1 promotes the cartilage degeneration and hypertrophy by activating the Krüppel-like factor 5 and β-catenin signaling. *Biochimica Biophysica Acta-Molecular Basis Of Dis.* 1865 (9), 2490–2503. doi:10.1016/j.bbadis.2019.06.010

Tanel, A., Fonseca, S. G., Yassine-Diab, B., Bordi, R., Zeidan, J., Shi, Y., et al. (2009). Cellular and molecular mechanisms of memory T-cell survival. *Expert Rev. Vaccines* 8 (3), 299–312. doi:10.1586/14760584.8.3.299

Wang, L., Wu, S., He, H., Ai, K., Xu, R., Zhang, L., et al. (2022). CircRNA-ST6GALNAC6 increases the sensitivity of bladder cancer cells to erastin-induced ferroptosis by regulating the HSPB1/P38 axis. *Lab. Investig.* 102 (12), 1323–1334. doi:10.1038/s41374-022-00826-3

Wang, Y., Cao, L., Wang, Q., Huang, J., and Xu, S. (2019). LncRNA FOXD2-AS1 induces chondrocyte proliferation through sponging miR-27a-3p in osteoarthritis. *Artif. Cells Nanomedicine Biotechnol.* 47 (1), 1241–1247. doi:10.1080/21691401.2019.1596940

Wu, J. B., Zhu, Y., Luo, M. M., and Li, L. (2021). Comprehensive analysis of pyroptosis-related genes and tumor microenvironment infiltration characterization in breast cancer. *Front. Immunol.* 12, 748221. doi:10.3389/fimmu.2021.748221

Xu, J., Pei, Y., Lu, J., Liang, X., Li, Y., Wang, J., et al. (2021). LncRNA SNHG7 alleviates IL-1β-induced osteoarthritis by inhibiting miR-214-5p-mediated PPARGC1B signaling pathways. *Int. Immunopharmacol.* 90, 107150. doi:10.1016/j.intimp.2020.107150

Yang, D. W., Zhang, X., Qian, G. B., Jiang, M. J., Wang, P., Wang, K. Z., et al. (2019). Downregulation of long noncoding RNA LOC101928134 inhibits the synovial hyperplasia and cartilage destruction of osteoarthritis rats through the activation of the Janus kinase/signal transducers and activators of transcription signaling pathway by upregulating IFNA1. *J. Of Cell. Physiology* 234 (7), 10523–10534. doi:10.1002/jcp.27730

Yu, G., Wang, L., Han, Y., and He, Q. Y. (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. *Omics-A J. Of Integr. Biol.* 16 (5), 284–287. doi:10.1089/omi.2011.0118

Yuan, H., Li, X., Zhang, X., Kang, R., and Tang, D. (2016). CISD1 inhibits ferroptosis by protection against mitochondrial lipid peroxidation. *Biochem. And Biophysical Res. Commun.* 478 (2), 838–844. doi:10.1016/j.bbrc.2016.08.034

Zhang, X., Zhang, W., Jiang, Y., Liu, K., Ran, L., and Song, F. (2019). Identification of functional lncRNAs in gastric cancer by integrative analysis of GEO and TCGA data. *J. Of Cell. Biochem.* 120 (10), 17898–17911. doi:10.1002/jcb.29058

Zhang, Z., Huang, C., Jiang, Q., Zheng, Y., Liu, Y., Liu, S., et al. (2020). Guidelines for the diagnosis and treatment of osteoarthritis in China (2019 edition). *Ann. Transl. Med.* 8 (19), 1213. doi:10.21037/atm-20-4665

Zheng, M., Karki, R., Vogel, P., and Kanneganti, T. D. (2020). Caspase-6 is a key regulator of innate immunity, inflammasome activation, and host defense. *Cell* 181 (3), 674–687. doi:10.1016/j.cell.2020.03.040

Zhou, Y., Zhou, B., Pache, L., Chang, M., Khodabakhshi, A. H., Tanaseichuk, O., et al. (2019). Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat. Commun.* 10 (1), 1523. doi:10.1038/s41467-019-09234-6

Zou, J., Li, X., Shi, Z., and Xue, J. F. (2018). Effects of C-myc gene silencing on interleukin-1β-induced rat chondrocyte cell proliferation, apoptosis and cytokine expression. *J. Of Bone And Mineral Metabolism* 36 (3), 286–296. doi:10.1007/s00774-017-0845-4

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 13 February 2023
DOI 10.3389/fgene.2023.1083615

OPEN ACCESS

EDITED BY
Shuai Liu,
University of Hawaii at Manoa,
United States

REVIEWED BY
Shangxue Yan,
Anhui Medical University, China
Félicie Costantino,
Université de Versailles Saint-Quentin-en-Yvelines, France

*CORRESPONDENCE
Lindi Jiang,
zsh-rheum@hotmail.com

†These authors have contributed equally to this work

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

RECEIVED 29 October 2022
ACCEPTED 26 January 2023
PUBLISHED 13 February 2023

CITATION
Huang H, Dong X, Mao K, Pan W, Nie B and Jiang L (2023), Identification of key candidate genes and pathways in rheumatoid arthritis and osteoarthritis by integrated bioinformatical analysis. *Front. Genet.* 14:1083615. doi: 10.3389/fgene.2023.1083615

Identification of key candidate genes and pathways in rheumatoid arthritis and osteoarthritis by integrated bioinformatical analysis

Huijing Huang[1†], Xinyi Dong[2†], Kaimin Mao[3†], Wanwan Pan[4], Bin'en Nie[5] and Lindi Jiang[1*]

[1]Department of Rheumatology, Zhongshan Hospital, Fudan University, Shanghai, China, [2]Shanghai Jiaotong University School of Medicine, Shanghai, China, [3]Department of Critical Care Medicine, Renji Hospital, School of Medicine, Shanghai Jiaotong University, Shanghai, China, [4]Yankuang New Journey General Hospital, Jingning, Shandong, China, [5]Department of Bone and Joint Surgery, Renji Hospital, School of Medicine, Shanghai Jiaotong University, Shanghai, China

Rheumatoid arthritis (RA) and osteoarthritis (OA) are the most common joint disorders. Although they have shown analogous clinical manifestations, the pathogenesis of RA and OA are different. In this study, we used the online Gene Expression Omnibus (GEO) microarray expression profiling dataset GSE153015 to identify gene signatures between RA and OA joints. The relevant data on 8 subjects obtained from large joints of RA patients (RA-LJ), 8 subjects obtained from small joints of RA patients (RA-SJ), and 4 subjects with OA were investigated. Differentially expressed genes (DEGs) were screened. Functional enrichment analysis of DEGs including the Gene Ontology terms and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways were identified, which were mainly associated with T cell activation or chemokine activity. Besides, protein-protein interaction (PPI) network analysis was performed, and key modules were identified. Hub genes of RA-LJ and OA groups were screened, they were *CD8A*, *GZMB*, *CCL5*, *CD2,* and *CXCL9*, whereas *CD8A*, *CD2*, *IL7R*, *CD27,* and *GZMB* were hub genes of RA-SJ and OA group. The novel DEGs and functional pathways between RA and OA identified in this study may provide new insight into the underlying molecular mechanisms and therapeutic strategies of RA and OA.

KEYWORDS

rheumatoid arthritis, osteoarthritis, microarray expression profiling dataset, differentially expressed genes, protein-protein interaction network

Introduction

According to recent studies, rheumatoid arthritis (RA) and osteoarthritis (OA) are considered to be the most prevalent rheumatic diseases, affecting 1% and 10% world's population, respectively (van der Woude and van der Helm-van Mil, 2018). RA is an autoimmune disorder presented with chronic aggressive multiple arthritis and systemic manifestation. Metacarpophalangeal joints and wrists are the most commonly involved joints, followed by large joints like the knee (Smolen et al., 2016). The pathogenesis of RA has not yet been completely understood. However, it is generally acknowledged that genetic and environmental factors play important roles in this disease (McInnes and Schett, 2011). The majority of patients with RA suffer damage to their small joints in the inchoate stage (Nakajima et al., 2016). Recent treatment-to-target strategies using classical or biological disease-modifying anti-rheumatic drugs (bDMARDs) have allowed patients to achieve remission and have delayed progressive damage to their small joints (Smolen and Aletaha, 2015). A

TABLE 1 Characteristics of the samples in this study.

Sample id	Diagnosis	Joint size	Age (year)	Gender	ACPA/RF positive	DAS28-CRP	CRP (mg/L)
GSM4633117	RA	LJ	36	F	Y	5.99	77
GSM4633119	RA	LJ	40	F	Y	6.22	44
GSM4633121	RA	LJ	60	F	Y	3.05	5
GSM4633123	RA	LJ	62	F	Y	4.54	1
GSM4633125	RA	LJ	78	F	Y	4.49	4
GSM4633127	RA	LJ	38	F	N	5.69	17
GSM4633129	RA	LJ	66	F	Y	4.27	14
GSM4633134	RA	LJ	30	F	Y	5.57	10
GSM4633116	RA	SJ	36	F	Y	5.99	77
GSM4633118	RA	SJ	40	F	Y	6.22	44
GSM4633120	RA	SJ	60	F	Y	3.05	5
GSM4633122	RA	SJ	62	F	Y	4.54	1
GSM4633124	RA	SJ	78	F	Y	4.49	4
GSM4633126	RA	SJ	38	F	N	5.69	17
GSM4633128	RA	SJ	66	F	Y	4.27	14
GSM4633132	RA	SJ	55	F	N	5.55	38
GSM4633136	OA	LJ	71	F	NA	NA	NA
GSM4633137	OA	LJ	77	F	NA	NA	NA
GSM4633138	OA	LJ	69	M	NA	NA	NA
GSM4633139	OA	LJ	66	F	NA	NA	NA

large joint (shoulder, elbow, hip, and knee) is usually destroyed during an advanced stage of RA. A patient with RA is not routinely monitored for progressive damage to large joints, even though damage to large joints has a considerably greater impact on functional ability than damage to smaller joints (Nakajima et al., 2016).

As the most common arthritis throughout the world, osteoarthritis is the leading cause of disability in the elderly (Abramoff and Caldera, 2020). OA destroys articular cartilage and stimulates the hyperplasia of the margin of bones as well as lesions in synovium and tissue of joints, which can lead to a series of biochemical and morphological changes in overused or weight-bearing joints unilaterally or bilaterally (Xia et al., 2014; Abramoff and Caldera, 2020). The pathogenesis of OA remains unclear. It is known that factors promoting inflammation, especially IL-1β and TNF-α play key roles in the development of OA. Several factors such as genetic predisposition, ageing, obesity, and joint misalignment have been implicated as contributing to the development of OA. A recent study by Butterfield et al. identified 14 genes involved in osteoarthritis pathogenesis, including *Pitx1*, and functionally characterize 6 candidate human osteoarthritis genes (*Unk, Josd1, Gsdme, Arhgap30, Ccdc6,* and *Col4a2*) in mouse models (Butterfield et al., 2021), revealing the genetic basis for osteoarthritis.

Although RA and OA share some clinical manifestations, they are two distinct arthritic disorders with RA being an autoimmune disease and OA being a degenerative disease (Vina and Kwoh, 2018). For identifying biomarkers for the diagnosis and prognosis of diseases, microarrays have become a promising and efficient tool for exploring significant genetic or epigenetic changes in disease (Cao et al., 2021b). RA patients and normal individuals, OA patients, and normal individuals as well as RA patients and OA patients, were found to have differentially expressed genes (DEGs) utilizing bioinformatic analysis (Lin et al., 2020; Triaille et al., 2020; Liu and Chen, 2021). Therefore, in this study, we compared gene expression profiles in synovial tissue between large and small joints of RA and joints of OA *via* bioinformatic analysis of an online dataset, seeking to identify possible key genes that involve the pathogenesis of joints of RA from joints of OA.

Materials and methods

Microarray data

The gene expression profile dataset GSE153015, deposited by Triaille C et al., was obtained from the Gene Expression Omnibus (GEO, https://www.ncbi.nlm.nih.gov/geo/) and was based on the platform of GPL570 Affymetrix Human Genome U133 Plus 2. 0 Array [HGU133_Plus_2]. We collected samples from 24 subjects, including 10 small joints (metacarpophalangeal joints or wrists) from 10 RA patients, 10 large joints (knees) from the same 10 RA patients,

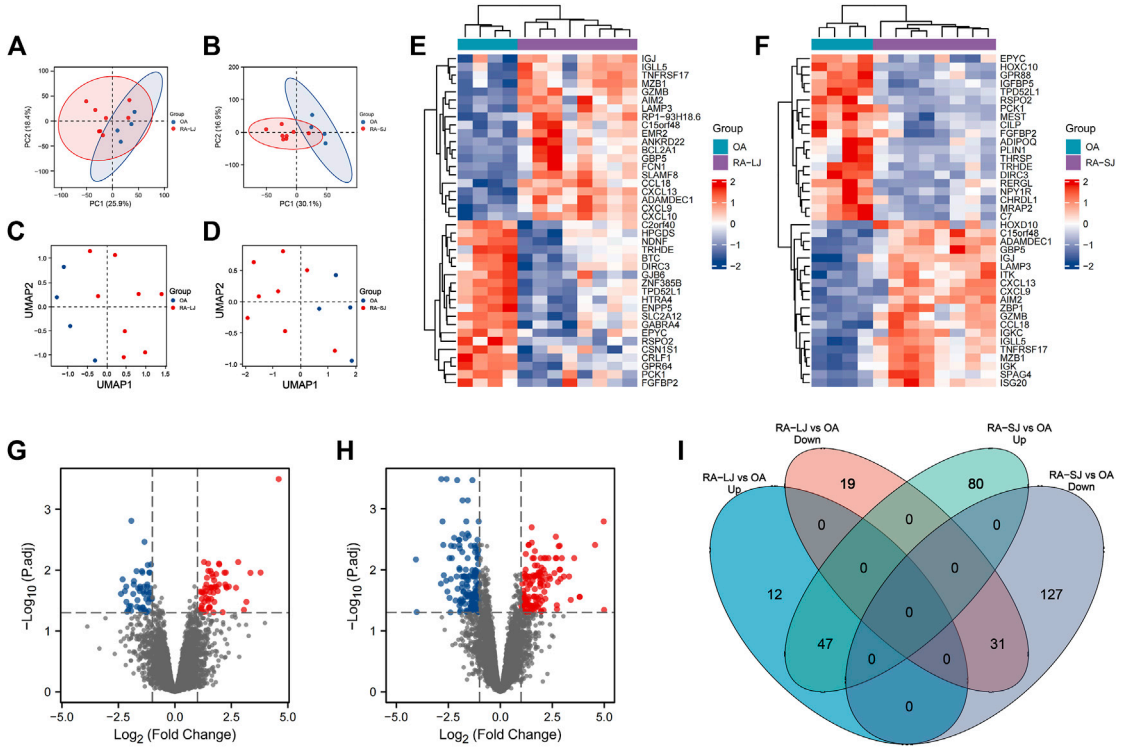


FIGURE 1
DEGs were identified from the online GEO dataset. **(A)–(D)** RA-LJ and RA-SJ samples are separated from OA samples in PCA and UMAP analyses. **(E) (F)** Heatmaps showed the top 20 upregulated and downregulated genes between RA and OA. **(G) (H)** Genes in RA samples and OA samples were significantly separated in the volcano plot. Red represents upregulated genes, and blue represents downregulated genes **(I)** Venn diagram of differentially expressed genes between RA-LJ and RA-SJ samples compared with OA samples.

and 4 OA subjects. The type of biological material used for gene expression quantification can be found in the study of Triaille C et al. (Triaille et al., 2020) Before the analysis, we did a principal component analysis (PCA) and found that two samples of large joints (LJ) of RA (GSM4633131, GSM4633133) and two samples of small joints (SJ) of RA group (GSM4633130, GSM4633135) were mixed with OA samples, respectively. As a result, we excluded the 4 samples to improve the quality of the samples. We obtained from the GEO dataset the age, gender, ACPA/RF status, DAS 28-CRP scores, and C reaction protein (CRP) levels of the individuals, as well as the annotation file for GPL570 and displayed in Table 1.

Differential expression analysis

To screen for DEGs, we compared expression profiles of large and small joints of RA patients and OA patients using the online analysis tool GEO2R. *p*-values and adjusted *p*-values were calculated *via* t-tests and statistically significant DEGs were defined with the criteria of 1) a $|\log2\ (\text{fold-change})| > 1$ and 2) an adjusted $p < 0.05$. Patients were divided according to their joint size. RA patients with large joints were compared with RA patients with small joints. R software was used to draw the volcano plot and PCA, HTML software was used to create the heatmap for the DEGs, and Venn diagrams of DEGs were drawn using Venny 2.1 (https://bioinfogp.cnb.csic.es/tools/venny/index.html), and uniform manifold approximation and projection (UMAP) were performed *via* R's "umap" package (Mao et al., 2020).

Functional enrichment analysis of DEGs

Using Enrichr (https://amp.pharm.mssm.edu/Enrichr/), functional enrichment analyses of DEGs were conducted, including Gene Ontology (GO) terms and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways. A GO analysis consists of three biological processes (BP), a cellular component (CC), and a molecular function (MF), which provide a framework for describing the functions of gene products in all organisms. We assigned DEGs to specific pathways using KEGG pathway analysis (15). The threshold for significance was set at Benjamini-adjusted $p < 0.05$ and an enriched gene count of at least five genes enriched in the pathway.

Protein-protein interaction (PPI) network construction

The PPI network analysis was carried out using Search Tool for the Retrieval of Interacting Genes (STRING, https://string-db.org/), a web-based database dedicated to predicting protein-protein interactions consisting of both physical and functional relations. We used Cytoscape v3.8.2 software to visualize and build the PPI network simultaneously while mapping the DEGs onto the PPI network with a medium confidence score of 0.4. Using ClusterOne from Cytoscape's software suite, the gene network clustering analysis was conducted to refine the key PPI network modules. An R software package with the "ggplot2" package was used to analyze the expression levels of key PPI

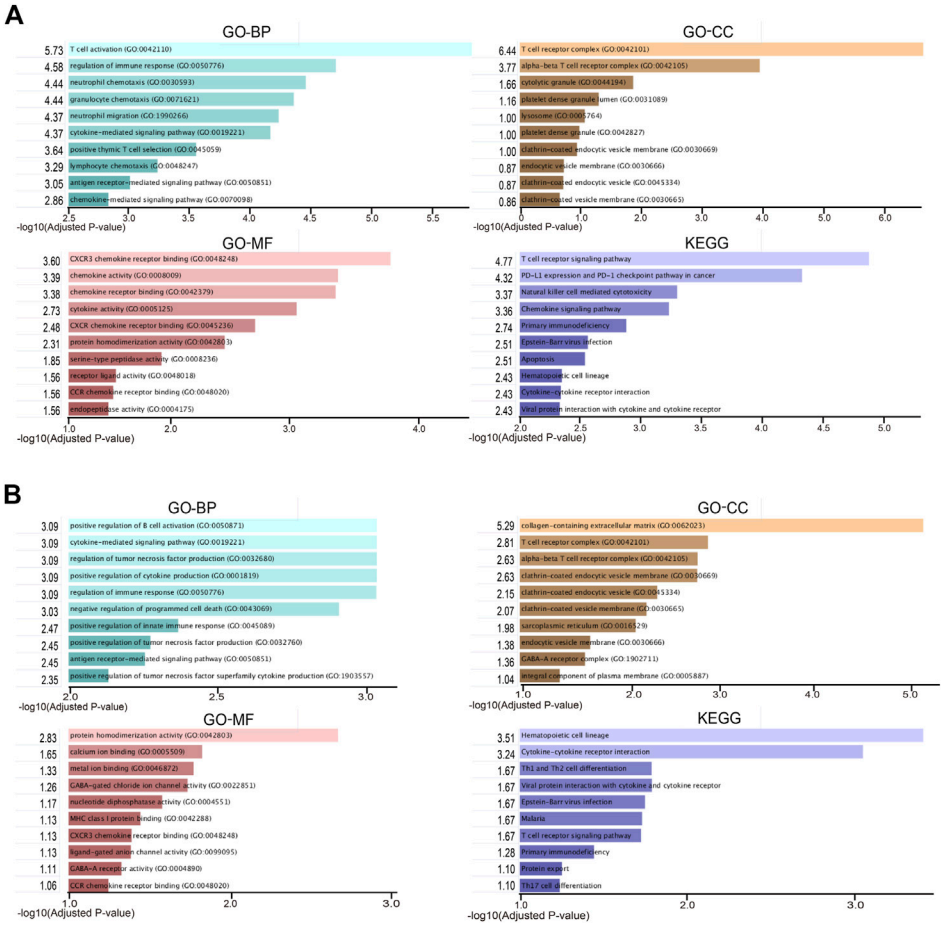


FIGURE 2
Representative enriched functional terms of RA samples compared with OA samples. **(A)** Most significantly enriched terms of GO-BP, MF, CC categories, and KEGG pathways in RA-LJ samples *versus* OA samples **(B)** Significantly enriched terms of GO-BP, MF, CC categories, and KEGG pathways in RA-SJ samples compared with OA samples. The *x*-axis represents the value of [-log (Adjusted *p*-value)].

network modules (Cao et al., 2021a). The degree topological algorithm of Cytohubba was used to determine which nodes had the greatest number of interactions with neighboring nodes as hub genes.

Immunohistochemistry

Paraffin joint tissue sections were cut into five mm-thick sections in the department of bone and joint surgery, at Renji hospital. Following deparaffinization and hydration, sections were treated with primary anti-CD8A antibodies (PA5-114369, Invitrogen), anti-CD2 antibodies (PA5-32312, Invitrogen), anti-GZMB antibodies (ab255598, Abcam), anti-CCL5 antibodies (ab52562, Abcam), anti-CXCL9 antibodies (ab290643, Abcam), anti-IL7R antibodies (ab259806, Abcam), and anti-CD27 antibodies (ab131254, Abcam). Incubation at room temperature for 30 min was followed by diluted biotinylated secondary antibodies. After incubation in Street Avidin-Biotin Complex for 20 min, sections were treated with diaminobenzidine (DAB) substrate solution until the desired color intensity was achieved (He et al., 2021; Mao et al., 2021). Images were taken under a light microscope (Nikon).

Statistical analysis

To correct the *p*-value, the Benjamini–Hochberg FDR (false discovery rate) was used. Hypergeometric tests were used to distinguish significantly enriched GO terms and KEGG pathways. Statistical analysis of significant differences was conducted using one-way ANOVA by Prism 7 software. A *p*-value less than 0.05 was considered statistically significant.

Results

DEGs between RA and OA joints

The characteristics of the samples were listed in Table 1. The microarray expression dataset GSE153015 was downloaded from the GEO database, and DEGs were obtained between RA-LJ and OA (Supplementary Tables 1, 2). Based on the established criteria, we found that 59 genes were upregulated and 50 genes were downregulated between RA-LJ and OA, and 127 genes were upregulated and 158 genes were downregulated between RA-SJ and OA. According to the PCA and UMAP analyses in Figures 1A–D,

the clusters of the two comparison groups were found in relatively independent quadrants, indicating that there was a significant difference in the RA-LJ and OA samples, as well as the RA-SJ and OA samples; Figures 1E, F show heatmaps for DEGs that show the top 20 upregulated genes in RA-LJ *versus* OA and RA-SJ *versus* OA, respectively; Figures 1G, H show volcano plots of compared groups. Our Venn diagram of the DEGs (Figure 1I) made it easier for the audience to understand the intersection and independence of the two DEG groups. According to the Venn diagram, 47 upregulated genes and 31 downregulated genes are shared by RA's large and small joints. The overlapped genes made up 79.67% of upregulated genes in the RA-LJ group *versus* the OA group and 62% of downregulated genes in the RA-LJ group *versus* the OA group.

Functional and pathway enrichment of DEGs

To take a further step in the investigation of the biological functions of DEGs, a functional enrichment analysis was conducted and the results are presented in Figure 2. As a result, the DEGs in RA-LJ *versus* OA groups are mainly T cell receptor complexes in the CC category, chemokine receptor binding in the MF category, and T cell activation in the BP category. The T cell receptor signaling pathway is the most enriched KEGG pathway. The DEGs in RA-SJ compared to OA groups include collagen-containing extracellular matrix components, protein homodimerization activity in MF, and positive regulation of B cell activation and cytokine-mediated signaling pathways in BP. Pathways associated with hematopoietic cells were identified as the most significantly altered pathways in RA-SJ compared to OA.

PPI network analysis of DEGs

In accordance with the STRING web-based database, two PPI networks were built with an interaction score >0.4 (Figure 3A). In order to identify the key PPI network modules, we performed network gene clustering in Cytoscape using ClusterOne, which corresponds to gene nodes and edges. In Figures 3B–E, a comparison between RA-LJ and OA reveals two key modules with 14 upregulated genes (*CST7, CD8A, GEMK, GZMH, GZMA, GZMB, PRF1, CD3D, CXCL9, NKG7, CD2, CD247, ITGAL, CCL5*) and three downregulated genes (*SOX8, ZIC1, and POU3F3*). While in Figures 3F–R, we identified 13 key modules with 33 upregulated genes (ITGAL, IL-7, CD27, CD38, GZMK, GZMA, GZMB, NKG7, KLRB1, CD247, CD2, CCL5, CD3D, CD3G, CXCL9, CXCL13, CD8A, IL7R, CCR2, ITK,

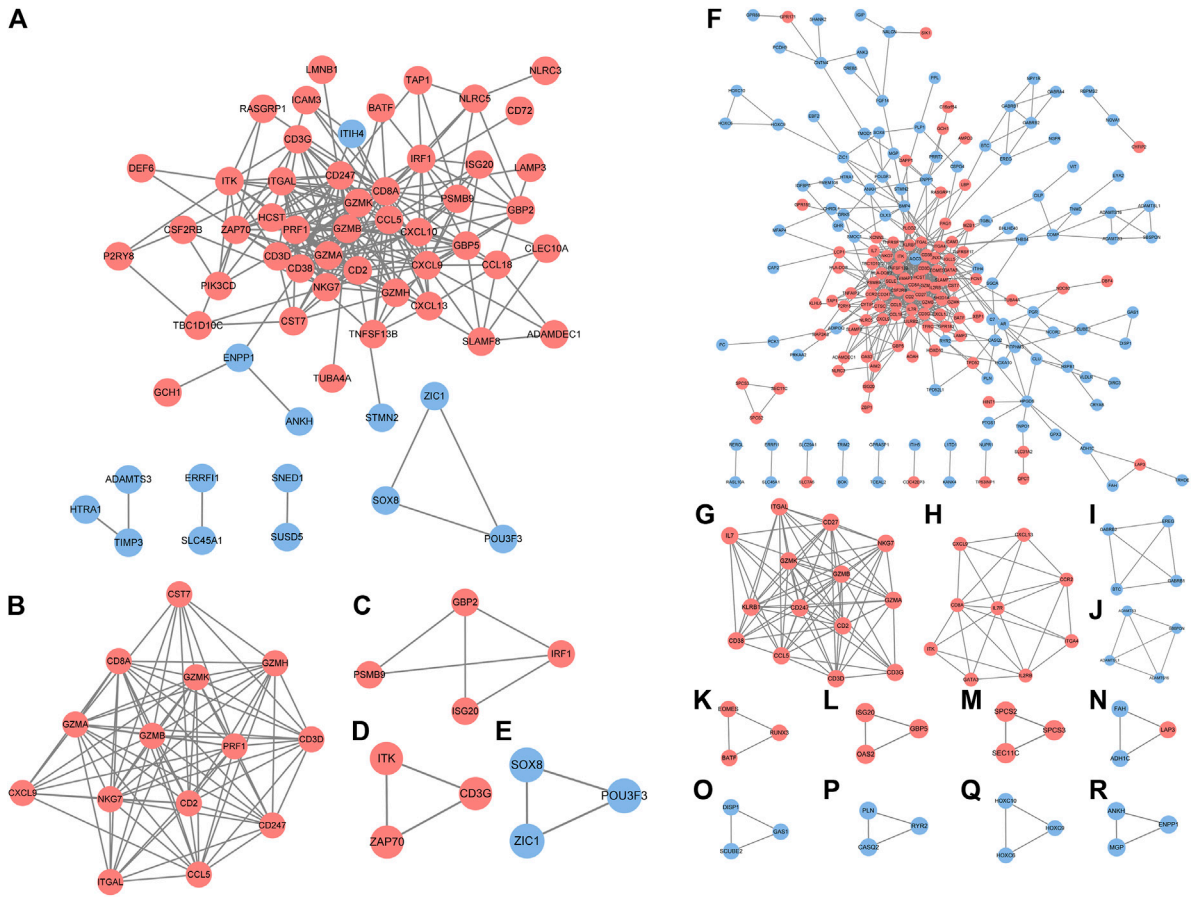


FIGURE 3
PPI network analysis of DEGs between RA-LJ/RA-SJ samples and OA samples. **(A)** Cytoscape network visualization of RA-LJ and OA samples that were obtained with interaction scores >0.4 according to the STRING online database **(B–E)** Four key modules were identified by ClusterOne. **(F)** Cytoscape network visualization of RA-SJ and OA samples **(G)–(R)** 12 key modules were identified by ClusterOne. The nodes represent genes, and the edges represent links between genes. Red represents upregulated genes, and blue represents downregulated genes.

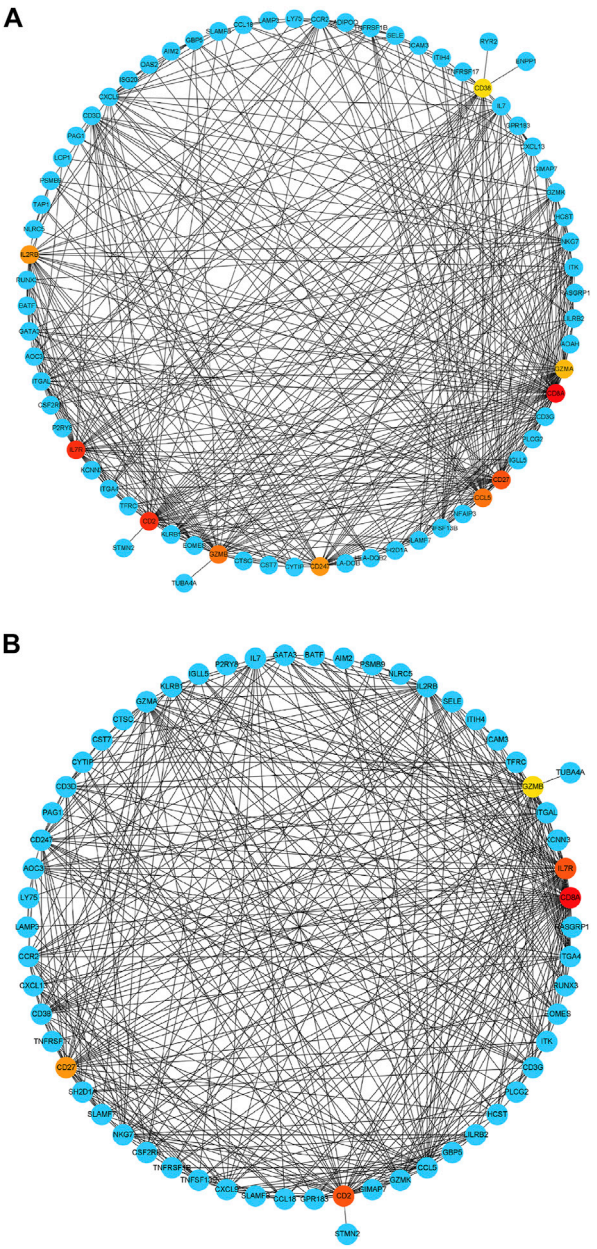


FIGURE 4
Detection of hub genes from the PPIs network of DEGs between RA and OA according to their degree value. **(A)** The highlighted five genes are *CD8A*, *GZMB*, *CCL5*, *CD2*, and *CXCL9* in RA-LJ samples compared with OA samples **(B)** The highlighted five genes are *CD8A*, *CD2*, *IL7R*, *CD27*, and *GZMB* in RA-SJ samples compared with OA samples. Red, orange and yellow modules represent hub genes, and the degree values of hub genes decrease from red to yellow modules. Blue modules refer to the genes which can interact with hub genes.

GATA3, IL2RB, ITGA4, EOMES, BATF, RUNX3, ISG20, OAS2, GBP5, LAP3, SPCS2, SCPS3, SEC11C) and 22 downregulated genes (ADAMTSL1, ADAMTS3, ADAMTS16, SBSPON, GABRB1, GABRB2, BTC, EREG, HOXC6, HOXC9, HOXC10, DISP1, GAS1, SCUBE2, FAH, ADH1C, PLN, CASQ2, RYR2, ANKH, MGP, ENPP1) in RA-SJ *versus* OA.

Based on their degree value, which represents the number of interactions among the genes, the hub genes were sorted. Figure 4A shows that the hub genes of the RA-LJ and OA groups are *CD8A, GZMB, CCL5, CD2,* and *CXCL9*, whereas Figure 4B shows the hub genes of RA-SJ and OA groups as *CD8A, CD2, IL7R, CD27,* and GZMB. Among them, *CD8A, GZMB,* and *CD2* are hub genes in both groups. In Figure 5A, the expression levels of these seven genes (*CD8A, CD2, GZMB, CCL5, CXCL9, IL7R, CD27*) were shown. In RA joints, *CD8A, CD2, GZMB, CCL5, CXCL9, IL7R,* and *CD27* levels were significantly higher than in OA joints. Further, IHC was used to determine the expression levels of seven different hub genes in large and small joints of RA and OA joints. As shown in Figures 5B, C, we found the protein expressions of CD8A, GZMB, CCL5, CD2, and CXCL9 were significantly higher in RA-LJ than that in OA, while the protein levels of CD8A, CD2, IL7R, CD27, and GZMB in RA-SJ were significantly higher than in OA joint (Figures 5D–F).

Discussion

Rheumatoid arthritis and osteoarthritis are two of the most common of the more than 100 different types of arthritis. As a result of chronic inflammation and autoimmunity associated with RA disease, synovial fluid is produced by the membrane lining, causing synovitis, joint pain, and eventually chronic and progressive erosion (Imas et al., 2020). As a whole-joint disease, OA causes increased remodeling of cartilage, subchondral bone, bone marrow, and synovium, as well as the joint (Hügle and Geurts, 2017). For proper treatment, it is necessary to understand the different mechanisms of the two diseases. As part of this study, we compared gene expression profiles in synovial tissue between large and small RA joints and OA joints using bioinformatic analysis. By using the online tool GEO2R, differentially expressed genes were screened. In addition, we identified 5 DEGs that have the most interactions in the two PPI networks as hub genes of two comparison groups. In RA-LJ *versus* OA groups, *CD8A, GZMB, CCL5, CD2,* and *CXCL9* are the hub genes, while *CD8A, CD2, IL7R, CD27,* and *GZMB* are the hub genes in RA-SJ *versus* OA groups. Lastly, immunohistochemistry was used to validate the protein expression levels of the 7 hub genes in RA and OA joint tissues.

Based on our analysis, we found that GO enrichments were mainly related to T cell activation, B cell activation, chemokine activity, and collagen-containing extracellular matrix in large and small joints of RA. In RA joints, T cell receptor signaling pathways and cytokine-cytokine receptor interactions were the main signaling pathways. It is easy to understand from these results for that rheumatoid arthritis is an autoimmune and inflammatory disease, which also suggests different treatment strategies for RA and OA. Rituximab, a kind of anti-CD20 chimeric monoclonal antibody that can inhibit the proliferation of stimulated human B cells, has long been a choice for refractory rheumatoid arthritis (Tavakolpour et al., 2019). Moreover, bDMARDS are now widely used to treat RA. The majority of bDMARDS are anti-cytokine or anti-chemokine molecules, but some patients do not respond to existing bDMARDs. Therefore, new bDMARDs for RA need to be developed.

Hub genes *CCL5*, also known as *RANTES*, its protein is a chemoattractant for blood monocytes, memory T helper cells, and eosinophils. It causes the release of histamine from basophils and activates eosinophils and plays an active role in the chemotactic activity of T cells in RA by facilitating leukocyte infiltration (Luterek-Puszyńska et al., 2017). Solomon A. et al. have demonstrated that CCL5 induced a positive inflammatory response in RA by activating synovial fibroblasts, thereby facilitating matrix

10.3389/fgene.2023.1083615

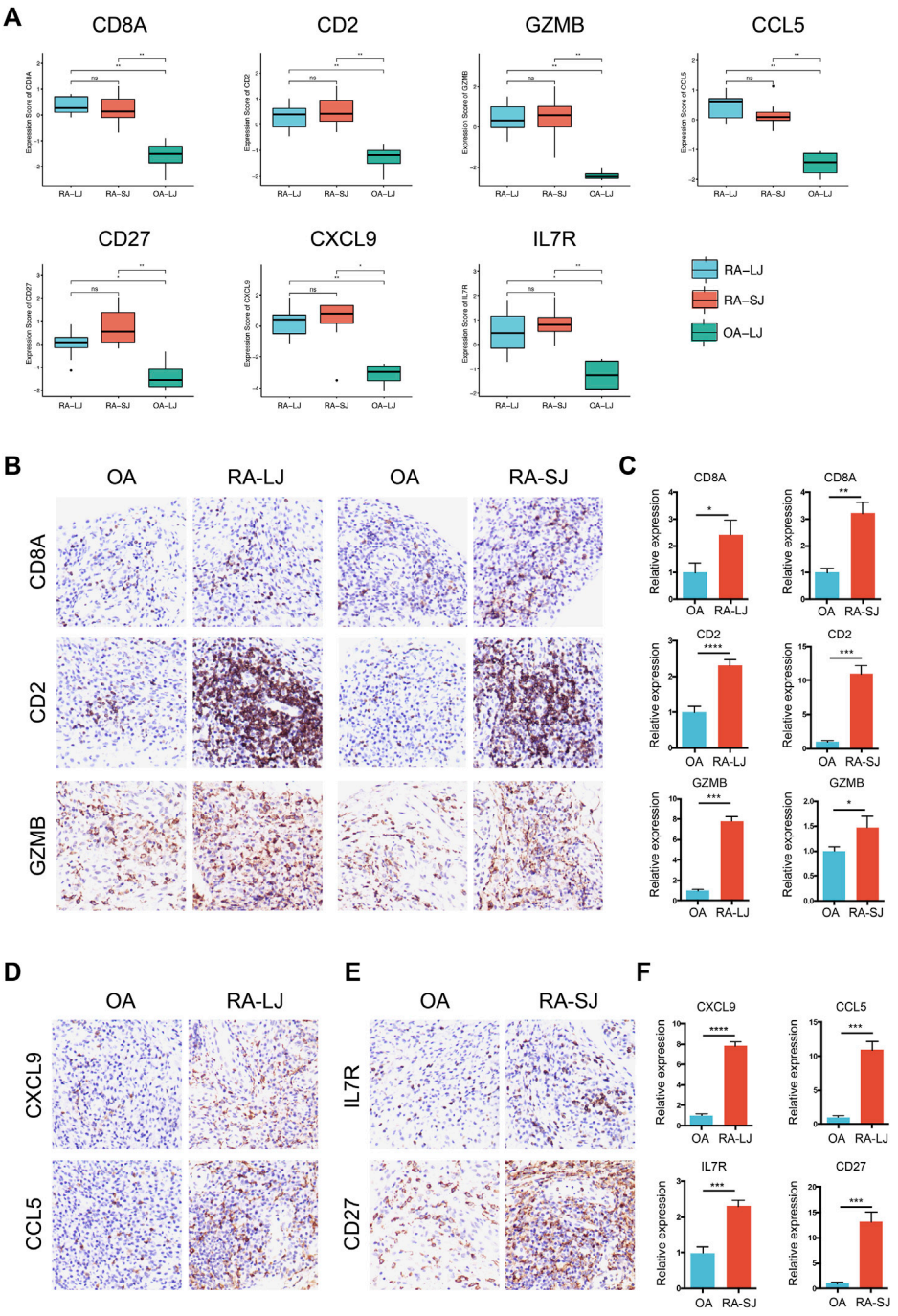


FIGURE 5
Expressions of hub genes in different groups in the microarray datasets and validations by immunohistochemical staining using joint biopsy sections. **(A)** Expressions of hub gene *CD8A*, *CD2*, *GZMB*, *CCL5*, *CXCL9*, *IL7R*, and *CD27* in RA-LJ, RA-SJ, and OA samples **(B)**–**(F)** The protein expressions of the 7 hub genes were determined by immunohistochemical staining using joint samples from patients. **p*-value <0.05.***p*-value <0.01. ****p*-value <0.001. *****p*-value <0.0001. Ns, no significance.

metalloproteinase-1 (MMP-1) and MMP-13-mediated destruction of the extracellular matrix (Agere et al., 2017). The pathological process of arthritis is also mediated by another gene, *CXCL9*, whose protein is detected in sera, synovial fluid, and synovial tissue (Lee et al., 2011; Yoshida et al., 2012; Pandya et al., 2017). By binding to CXCR3 on synovial tissue, inflammatory chemokines attract Th1 cells and macrophages, thereby contributing to arthritis accumulation. These findings suggest potential bDMARDS targets for treating RA in the future.

T cells participate in multiple pathways driving the disease process in rheumatoid arthritis. Naive $CD4^+$ T cells from patients with RA transition into highly proliferative, tissue-invasive and proinflammatory effector cells (Weyand and Goronzy, 2017). Equipped with tissue-invasive features, RA $CD4^+$ T cells rapidly induce synovitis in a human

synovium mouse chimera model (Weyand and Goronzy, 2021). It was found that synovial T cells, together with synovial macrophages, are the cellular origin of TNF (Zhang et al., 2019). To date, however, anti-T cell therapy has not been one of the major success stories of RA. Because there is still a lack of a highly selected target for anti-T cell therapy, completely blocking T cell activity is not feasible. F Zhang et al. defined distinct subsets of $CD8^+$ T cells characterized by GZMK+, GZMB+, and GNLY + phenotypes in RA joint synovial tissues by integrating single-cell transcriptomics and mass cytometry (Zhang et al., 2019), indicating an important role of hub gene *GZMB* in the classification of $CD8^+$ T cells in RA. It has already been found for decades that there was a remarkable increase in GAMB concentrations in plasma and synovial fluid of patients with established rheumatoid arthritis compared with disease controls (Tak et al., 1999). Besides, IL-7/IL-7R physiologically promotes T cell proliferation and prolonged survival as well as pathologically influencing Th1/Th17 cell differentiation, potentiated glycolysis, and expansion of osteoclast maturation, which contributes to the neovascularization in RA synovial tissue (Meyer et al., 2022). In addition, CD27, also known as TNFRSF7, is constitutively expressed on most T cells, and the interaction with its ligand CD70, can provide signals to T cells to control their accumulation and reactivity (Croft and Siegel, 2017). It has been found for a long time that in the synovial fluid of patients with RA, soluble CD27 levels and $CD27^+$ T cell numbers are elevated and correlate with the levels of rheumatoid factor, supporting a role for CD27 in human RA (Tak et al., 1996). Blocking $CD27^-CD70$ interactions with anti-CD70 antibody reduces bone and cartilage erosion and inflammatory infiltrates in the joints of mice with collagen-induced arthritis, moreover, decreases collagen-specific antibody production (Oflazoglu et al., 2009). In summary, molecules like GZMB, IL-7R, and CD27 offer the possibility of highly targeted and sophisticated therapies for RA.

Conclusion

In conclusion, we screened differentially expressed genes between large and small joints of RA between OA joints in this study. Several functional and pathway enrichments were also found between RA and OA joints, primarily related to T cell activation or chemokine activity. It was determined whether RA and OA joint samples expressed hub genes such as *CD8A, CD2, GZMB, CCL5, CXCL9, IL7R,* and *CD27*. As a result of our study, we may be able to identify new diagnostic markers and therapeutic targets for RA and OA.

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 studies involving human participants were reviewed and approved by the Ethics Review Board at Renji Hospital, Shanghai Jiaotong University. The patients/participants provided their written informed consent to participate in this study.

Author contributions

HH and KM conducted the formal analysis; XD wrote the original draft; WP participated in software and conducted data curation; BN collected joint samples and did the pathological staining; LJ designed the study and edited the manuscript. All authors contributed to the article and approved the submitted version.

Funding

This study is supported by the National Natural Science Foundation of China (grant number 82201976), the "Clinic Plus" Outstanding Project (grant number 2021ZYA010) from Shanghai Key Laboratory for Nucleic Acid Chemistry and Nanomedicine, Clinical Research Project of Zhongshan Hospital (grant number 2020ZSLC14) and Clinical Cultivation Project of Shanghai Shenkang Science and Technology Development Center (grant number SHDC1209X05).

Conflict of interest

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

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2023.1083615/full#supplementary-material

SUPPLEMENTARY FIGURE S1
The correlation between hub genes and CRP, DAS28-CRP, and ACPA/RFpositivity.**(A)-(D)** The level of IL-7R was negatively correlated with the concentrationof CRP in RA-LJ. No hub gene has a significant correlation with the DAS28-CRP score. **(E)** The expression of CD2 in the ACPA/RF positive group was significantly higherthan that in ACPA/RF negative group.

SUPPLEMENTARY TABLE S1
DEGs of RA-LJ and OA.

SUPPLEMENTARY TABLE S2
DEGs of RA-SJ and OA.

References

Abramoff, B., and Caldera, F. E. (2020). Osteoarthritis: pathology, diagnosis, and treatment options. *Med. Clin. North Am.* 104 (2), 293–311. doi:10.1016/j.mcna.2019.10.007

Agere, S. A., Akhtar, N., Watson, J. M., and Ahmed, S. (2017). RANTES/CCL5 induces collagen degradation by activating MMP-1 and MMP-13 expression in human rheumatoid arthritis synovial fibroblasts. *Front. Immunol.* 8, 1341. doi:10.3389/fimmu.2017.01341

Butterfield, N. C., Curry, K. F., Steinberg, J., Dewhurst, H., Komla-Ebri, D., Mannan, N. S., et al. (2021). Accelerating functional gene discovery in osteoarthritis. *Nat. Commun.* 12 (1), 467. doi:10.1038/s41467-020-20761-5

Cao, F., Fan, Y., Yu, Y., Yang, G., and Zhong, H. (2021a). Dissecting prognosis modules and biomarkers in glioblastoma based on weighted gene Co-expression network analysis. *Cancer Manag. Res.* 13, 5477–5489. doi:10.2147/cmar.s310346

Cao, F., Wang, C., Long, D., Deng, Y., Mao, K., and Zhong, H. (2021b). Network-based integrated analysis of transcriptomic studies in dissecting gene signatures for LPS-induced acute lung injury. *Inflammation* 44 (6), 2486–2498. doi:10.1007/s10753-021-01518-8

Croft, M., and Siegel, R. M. (2017). Beyond TNF: TNF superfamily cytokines as targets for the treatment of rheumatic diseases. *Nat. Rev. Rheumatol.* 13 (4), 217–233. doi:10.1038/nrrheum.2017.22

He, Y., Li, H., Yao, J., Zhong, H., Kuang, Y., Li, X., et al. (2021). HO-1 knockdown upregulates the expression of VCAM-1 to induce neutrophil recruitment during renal ischemia-reperfusion injury. *Int. J. Mol. Med.* 48 (4), 185. doi:10.3892/ijmm.2021.5018

Hügle, T., and Geurts, J. (2017). What drives osteoarthritis?-synovial versus subchondral bone pathology. *Rheumatol. Oxf.* 56 (9), 1461–1471. doi:10.1093/rheumatology/kew389

Imas, J. J., Ruiz Zamarreño, C., Zubiate, P., Sanchez-Martín, L., Campión, J., and Matías, I. R. (2020). Optical biosensors for the detection of rheumatoid arthritis (RA) biomarkers: A comprehensive review. *Sensors (Basel)* 20 (21), 6289. doi:10.3390/s20216289

Lee, E. Y., Seo, M., Juhnn, Y. S., Kim, J. Y., Hong, Y. J., Lee, Y. J., et al. (2011). Potential role and mechanism of IFN-gamma inducible protein-10 on receptor activator of nuclear factor kappa-B ligand (RANKL) expression in rheumatoid arthritis. *Arthritis Res. Ther.* 13 (3), R104. doi:10.1186/ar3385

Lin, X., Li, L., Liu, X., Tian, J., Zheng, W., Li, J., et al. (2020). Genome-wide analysis of aberrant methylation of enhancer DNA in human osteoarthritis. *BMC Med. Genomics* 13 (1), 1. doi:10.1186/s12920-019-0646-9

Liu, J., and Chen, N. (2021). A 9 mRNAs-based diagnostic signature for rheumatoid arthritis by integrating bioinformatic analysis and machine-learning. *J. Orthop. Surg. Res.* 16 (1), 44. doi:10.1186/s13018-020-02180-w

Luterek-Puszyńska, K., Malinowski, D., Paradowska-Gorycka, A., Safranow, K., and Pawlik, A. (2017). CD28, CTLA-4 and CCL5 gene polymorphisms in patients with rheumatoid arthritis. *Clin. Rheumatol.* 36 (5), 1129–1135. doi:10.1007/s10067-016-3496-2

Mao, K., Geng, W., Liao, Y., Luo, P., Zhong, H., Ma, P., et al. (2020). Identification of robust genetic signatures associated with lipopolysaccharide-induced acute lung injury onset and astaxanthin therapeutic effects by integrative analysis of RNA sequencing data and GEO datasets. *Aging (Albany NY)* 12 (18), 18716–18740. doi:10.18632/aging.104042

Mao, K., Luo, P., Geng, W., Xu, J., Liao, Y., Zhong, H., et al. (2021). An integrative transcriptomic and metabolomic study revealed that melatonin plays a protective role in chronic lung inflammation by reducing necroptosis. *Front. Immunol.* 12, 668002. doi:10.3389/fimmu.2021.668002

McInnes, I. B., and Schett, G. (2011). The pathogenesis of rheumatoid arthritis. *N. Engl. J. Med.* 365 (23), 2205–2219. doi:10.1056/NEJMra1004965

Meyer, A., Parmar, P. J., and Shahrara, S. (2022). Significance of IL-7 and IL-7R in RA and autoimmunity. *Autoimmun. Rev.* 21 (7), 103120. doi:10.1016/j.autrev.2022.103120

Nakajima, A., Aoki, Y., Sonobe, M., Takahashi, H., Saito, M., Terayama, K., et al. (2016). Radiographic progression of large joint damage in patients with rheumatoid arthritis treated with biological disease-modifying anti-rheumatic drugs. *Mod. Rheumatol.* 26 (4), 517–521. doi:10.3109/14397595.2015.1109785

Oflazoglu, E., Boursalian, T. E., Zeng, W., Edwards, A. C., Duniho, S., McEarchern, J. A., et al. (2009). Blocking of CD27-CD70 pathway by anti-CD70 antibody ameliorates joint disease in murine collagen-induced arthritis. *J. Immunol.* 183 (6), 3770–3777. doi:10.4049/jimmunol.0901637

Pandya, J. M., Lundell, A. C., Andersson, K., Nordström, I., Theander, E., and Rudin, A. (2017). Blood chemokine profile in untreated early rheumatoid arthritis: CXCL10 as a disease activity marker. *Arthritis Res. Ther.* 19 (1), 20. doi:10.1186/s13075-017-1224-1

Smolen, J. S., and Aletaha, D. (2015). Rheumatoid arthritis therapy reappraisal: strategies, opportunities and challenges. *Nat. Rev. Rheumatol.* 11 (5), 276–289. doi:10.1038/nrrheum.2015.8

Smolen, J. S., Aletaha, D., and McInnes, I. B. (2016). Rheumatoid arthritis. *Lancet* 388 (10055), 2023–2038. doi:10.1016/s0140-6736(16)30173-8

Tak, P. P., Hintzen, R. Q., Teunissen, J. J., Smeets, T. J., Daha, M. R., van Lier, R. A., et al. (1996). Expression of the activation antigen CD27 in rheumatoid arthritis. *Clin. Immunol. Immunopathol.* 80 (2), 129–138. doi:10.1006/clin.1996.0106

Tak, P. P., Spaeny-Dekking, L., Kraan, M. C., Breedveld, F. C., Froelich, C. J., and Hack, C. E. (1999). The levels of soluble granzyme A and B are elevated in plasma and synovial fluid of patients with rheumatoid arthritis (RA). *Clin. Exp. Immunol.* 116 (2), 366–370. doi:10.1046/j.1365-2249.1999.00881.x

Tavakolpour, S., Alesaeidi, S., Darvishi, M., GhasemiAdl, M., Darabi-Monadi, S., Akhlaghdoust, M., et al. (2019). A comprehensive review of rituximab therapy in rheumatoid arthritis patients. *Clin. Rheumatol.* 38 (11), 2977–2994. doi:10.1007/s10067-019-04699-8

Triaille, C., Vansteenkiste, L., Constant, M., Ambroise, J., Méric de Bellefon, L., Nzeusseu Toukap, A., et al. (2020). Paired rheumatoid arthritis synovial biopsies from small and large joints show similar global transcriptomic patterns with enrichment of private specificity TCRB and TCR signaling pathways. *Front. Immunol.* 11, 593083. doi:10.3389/fimmu.2020.593083

van der Woude, D., and van der Helm-van Mil, A. H. M. (2018). Update on the epidemiology, risk factors, and disease outcomes of rheumatoid arthritis. *Best. Pract. Res. Clin. Rheumatol.* 32 (2), 174–187. doi:10.1016/j.berh.2018.10.005

Vina, E. R., and Kwoh, C. K. (2018). Epidemiology of osteoarthritis: literature update. *Curr. Opin. Rheumatol.* 30 (2), 160–167. doi:10.1097/bor.0000000000000479

Weyand, C. M., and Goronzy, J. J. (2017). Immunometabolism in early and late stages of rheumatoid arthritis. *Nat. Rev. Rheumatol.* 13 (5), 291–301. doi:10.1038/nrrheum.2017.49

Weyand, C. M., and Goronzy, J. J. (2021). The immunology of rheumatoid arthritis. *Nat. Immunol.* 22 (1), 10–18. doi:10.1038/s41590-020-00816-x

Xia, B., Di, C., Zhang, J., Hu, S., Jin, H., and Tong, P. (2014). Osteoarthritis pathogenesis: a review of molecular mechanisms. *Calcif. Tissue Int.* 95 (6), 495–505. doi:10.1007/s00223-014-9917-9

Yoshida, S., Arakawa, F., Higuchi, F., Ishibashi, Y., Goto, M., Sugita, Y., et al. (2012). Gene expression analysis of rheumatoid arthritis synovial lining regions by cDNA microarray combined with laser microdissection: up-regulation of inflammation-associated STAT1, IRF1, CXCL9, CXCL10, and CCL5. *Scand. J. Rheumatol.* 41 (3), 170–179. doi:10.3109/03009742.2011.623137

Zhang, F., Wei, K., Slowikowski, K., Fonseka, C. Y., Rao, D. A., Kelly, S., et al. (2019). Defining inflammatory cell states in rheumatoid arthritis joint synovial tissues by integrating single-cell transcriptomics and mass cytometry. *Nat. Immunol.* 20 (7), 928–942. doi:10.1038/s41590-019-0378-1

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 23 February 2023
DOI 10.3389/fgene.2023.1132361

OPEN ACCESS

EDITED BY
Yuan Liu,
Shanghai Jiao Tong University, China

REVIEWED BY
Hanghui Ye,
University of Texas MD Anderson Cancer Center, United States
Zhengzheng Yan,
Southern Medical University, China
Mengmeng Zhai,
Alliance Pharma, United States

*CORRESPONDENCE
Na Jiao,
jiaona@zju.edu.cn
Qiang Shu,
shuqiang@zju.edu.cn

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

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

RECEIVED 27 December 2022
ACCEPTED 13 February 2023
PUBLISHED 23 February 2023

CITATION
Zhao S, Zhu K, Li X, Zhong X, Zhao Y, Le Z, Liu Z, Xiao Y, Lai D, Jiao N and Shu Q (2023), Co-expression and interaction network analysis reveals dysregulated neutrophil and T-cell activation as the core mechanism associated with septic shock.
Front. Genet. 14:1132361.
doi: 10.3389/fgene.2023.1132361

Co-expression and interaction network analysis reveals dysregulated neutrophil and T-cell activation as the core mechanism associated with septic shock

Shaobo Zhao[1†], Kun Zhu[2†], Xiaoyi Li[3†], Xiaohui Zhong[1], Yanan Zhao[1], Zhenkai Le[1], Zhicong Liu[1], Yi Xiao[3], Dengming Lai[1,3], Na Jiao[3]* and Qiang Shu[1]*

[1]Department of Pediatric Surgery, The Children's Hospital, Zhejiang University School of Medicine, Hangzhou, Zhejiang, China, [2]Department of Pathology, The Children's Hospital, Zhejiang University School of Medicine, Hangzhou, Zhejiang, China, [3]National Clinical Research Center for Child Health, The Children's Hospital, Zhejiang University School of Medicine, Hangzhou, Zhejiang, China

Septic shock as a subset of sepsis, has a much higher mortality, while the mechanism is still elusive. This study was aimed at identifying core mechanisms associated with septic shock and its high mortality by investigating transcriptome data. We screened 72 septic-shock-associated genes (SSAGs) with differential expression between septic shock and sepsis in the discovery dataset. Further gene set enrichment analysis identified upregulated neutrophil activation and impaired T-cell activation in septic shock. Co-expression analysis revealed nine co-expressed gene modules. In addition, we determined twenty-one prognostic SSAGs using cox regression analysis in an independent dataset. Moreover, protein–protein interaction (PPI) network revealed two clusters. Among these neutrophil activation was enriched in the most positively-related modules and the cluster2 PPI network, while T-cell activation was enriched in both the most negatively-related module and one of the most positively-related modules as well as the cluster1 PPI network. ELANE, LCN2 and IFI44 were identified as hub genes with CytoHubba methods and semantic similarity analysis. Notably, ELANE was the only prognostic gene and was further validated in an external dataset. Blood neutrophil count was demonstrated to increase in septic shock and be a risky factor of prognosis based on clinical data. In conclusions, septic shock is associated with upregulated neutrophil activation and dysregulated T-cell activation. Three hub genes might have potentials as sensitive markers for the further translational research and ELANE could be a robust prognostic biomarker and effective therapeutic target.

KEYWORDS

bioinformatics, neutrophil, ELANE, T cell, septic shock

1 Introduction

Sepsis is known as life-threatening organ dysfunction due to a dysregulated host response to infection (Singer et al., 2016), which has been a major global health concern because of high mortality and unacceptable hospital costs (Reinhart et al., 2017). A recent study reported a total of 11.0 million sepsis-related deaths in an estimated 48.9 million incident cases of sepsis worldwide in 2017 (Rudd et al., 2020). More importantly, the incidence of sepsis has still steadily increased over the past several decades (Esposito et al., 2017). Meanwhile, sepsis has been the most expensive condition for hospital stays in the United States, and the costs continue to increase (Liang et al., 2006; Torio and Andrews, 2006; Torio and Moore, 2006). In particular, septic shock, as a subset of sepsis with underlying circulatory and cellular/metabolic abnormalities (Singer et al., 2016), has a much higher mortality approaching 40%–60% than 10% of sepsis (Cecconi et al., 2018; Napolitano, 2018).

However, the significant biological and clinical heterogeneity of sepsis remains a major challenge, which has led to the failure of clinical sepsis trials of immunotherapy (Rubio et al., 2019). The understanding of sepsis and septic shock is still limited and keeps evolving over time. The Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3) was developed in 2016 (Singer et al., 2016), reflecting improved knowledge on the pathophysiology of sepsis and septic shock. Importantly, Sepsis-3 led to the new definition of septic shock by a more restrictive and unambiguous criteria that the criteria of sepsis and vasopressor therapy needed to elevate mean arterial pressure ≥65 mmHg and lactate >2 mmol/L (18 mg/dL) despite adequate fluid resuscitation (Shankar-Hari et al., 2016; Singer et al., 2016; Napolitano, 2018), which means distinguishing septic shock from sepsis more clearly than ever before. In this context, it could be necessary to determine core mechanisms under the new definition for a more accurate interpretation of septic shock.

Previous studies of septic shock based on different definitions have shown several important mechanisms. Tissue hypoxia has been discussed as an important pathophysiological mechanism under the action of microbial endotoxins during septic shock (Pavez et al., 2020). From an immunological perspective, the activation of monocytes, macrophages and neutrophils was considered to participate in the intimate mechanism of septic shock (Gorecki et al., 2021). In particular, polymorphonuclear neutrophils (PMNs) have been shown to lose their direct antimicrobial functions and acquire an immunosuppressive action and participate in the generation of disseminated intravascular coagulation (DIC) when septic shock develops (Stiel et al., 2018). However, few studies have focused on the difference between septic shock and sepsis without shock syndromes. The mechanism of septic shock is not yet fully understood, and the identification of the core mechanism is still needed.

In this study, we analyzed the gene expression profiles of patients between septic shock and sepsis from public databases to identify core mechanisms associated with septic shock and its high mortality. Weighted gene co-expression network analysis (WGCNA) was conducted to identify septic-shock-associated gene modules. Prognostic genes among septic-shock-associated genes (SSAGs) were identified to explain the higher mortality at the molecular level. Combining the protein–protein interaction (PPI) network and semantic similarity network based on gene annotation, hub genes were identified with the most connectivity among SSAGs. The main goal of the present study was to better understand the molecular changes and screen core mechanisms responsible for the development from sepsis to septic shock under the new Sepsis-3 definition. For more accurate interpretation, the "sepsis" declared after in this study refers specifically to sepsis without shock diagnosis.

2 Materials and methods

2.1 Data source

The included transcriptome data were downloaded from gene expression omnibus (GEO) databases (http://www.ncbi.nlm.nih.gov/geo/) (Barrett et al., 2013). Only peripheral blood samples collected within 24 h of diagnosis or ICU admission were included. The RNA sequencing data of 91 adult samples (including 19 septic shock, 20 sepsis, 12 uncomplicated infection and 40 healthy controls) in the GSE154918 dataset, which were pre-processed using the DESeq2 package by the contributors (Love et al., 2014; Herwanto et al., 2021), were used as discovery dataset to explore genes, modules and mechanisms associated with septic shock. Additionally, the array data and survival information of 479 adult sepsis samples with a 28-day cumulative death rate about 23.80% in the GSE65682 dataset were read in R language to determine the prognostic significance of interested genes in sepsis patients. The gene expression profiles of GSE65682 were background-subtracted and normalized by a robust multi-array average algorithm using the affy package. The row count matrix of 345 adult sepsis samples including 52 dead and 293 survival samples in the GSE185263 dataset was downloaded to validate survival significance of the hub gene.

Clinical blood laboratory examinations data of sepsis and septic shock patients were extracted from the MIMIC-IV (version 2.0) database in the physionet (https://physionet.org/content/mimiciv/2.0/) for the further validation (Goldberger et al., 2000; Johnson et al., 2022). One of the authors who has finished the required Collaborative Institutional Training Initiative examination (Certification number 53459610 for Zhao) can access the database. The adult ICU stay samples meeting the sepsis-3 definition at the first day of ICU admission were included (Singer et al., 2016). The patients' parameters including absolute neutrophil count, absolute CD3 count (i.e., T cell count), absolute CD4 count and absolute CD8 count from blood specimens and survival data were extracted for further analysis. Specifically, we extracted the max values of neutrophil counts of each ICU stay within 6 h before ICU admission and 24 h after; while the chart time requirements of the other three items were limited to 6 h before ICU admission and 48 h after, concerning their more time costs waiting for the reports. In our study, the data about neutrophil counts of 8250 ICU stays containing 40.5% septic shock samples and with a 28-day cumulative mortality rate (CMR) about 22.3% were extracted. However, among them only 69 had the time-limited data about CD3 counts and 68 had desirable CD4 counts and CD8 counts due to their less clinical applications. More details

were shown in Supplementary Table S1. The code used for data extraction can be available on GitHub (https://github.com/MIT-LCP/mimic-iv).

2.2 Differential gene expression analysis

Differential expression analysis was conducted using moderated *t*-test by the limma R package (Ritchie et al., 2015). The differential expression cutoff values were set to |log2 fold change (logFC)|≥ 1 and adjusted *p*-value (adj.P) < 0.05. *p* values were adjusted by the Benjamini–Hochberg (BH) method.

2.3 Functional enrichment analysis

Functional enrichment analysis was conducted using the clusterProfiler R package (Yu et al., 2012). Gene set enrichment analysis (GSEA) based on the rankings of logFC of all genes and over-representation analysis was utilized to determine enriched biological process (BP) GO terms and KEGG pathways. The cutoff of the adjusted *p*-value by the BH method was set to 0.05.

2.4 Weighted gene co-expression network analysis (WGCNA)

The weighted co-expression network was constructed using the WGCNA package (Zhang and Horvath, 2005; Langfelder and Horvath, 2008). The minimum module size was set to 30, the dendrogram cut height for module merging was set to 0.2 and the desired minimum scale free topology fitting index R^2 was set to 0.8 to screen optimal soft-thresholding power. Module eigengene (ME) was defined as the first principal component of the gene expression matrix of the corresponding module. The relationships between model eigengenes and phenotypes were assessed using the Spearman correlation.

2.5 Survival analysis

A univariate Cox proportional hazard regression model was conducted by the survival R package to screen prognostic factors from septic-shock-associated DEGs in an independent dataset (GSE65682), and *p* values were corrected by the BH method.

2.6 Protein-protein interaction (PPI) network analysis

The PPI network were constructed based on the DEGs of septic shock online in the STRING database (http://string-db.org/) (version 11.5) (Szklarczyk et al., 2021). Specifically, the gene list was input into the multiple protein mode with default parameters. The credibility was set to 0.40. Disconnected nodes were hidden. Then, the output table was input into Cytoscape software. CytoHubba, a Cytoscape plugin, was used to screen potential hub genes by providing 12 topological analysis algorithms (i.e., MCC, DMNC, MNC, Degree, EPC, Bottleneck, Eccentricity, Closeness, Radiality, Betweenness, Stress, and Clustering Coefficient) (Chin et al., 2014). In this research, genes appearing at least 5 times in the top 10 results of each algorithm were considered as potential hub genes.

2.7 Semantic similarity analysis

Semantic similarities were measured using the GOSemSim package (Yu et al., 2010; Yu, 2020). The pairwise semantic similarities were calculated by Wang's measure algorithm (Wang et al., 2007) from three aspects, including biological processes (BP), molecular function (MF) and cellular component (CC). The final adjacency matrix of semantic similarities between genes was identified as the geometric means of the similarities from these three aspects. The candidate hub genes were screened according to the decreasing order of average semantic similarities of each gene.

2.8 Immune cell correlation estimation

To estimate the immune cell fractions, CIBERSORTx, a suite of machine learning tools (https://cibersortx.stanford.edu/), was used to perform a deconvolution algorithm based on bulk expression profiles (Newman et al., 2019). The correlation between hub genes and cell fractions in the GSE65682 dataset was estimated using the Spearman rank correlation coefficient.

2.9 Clinical investigation of neutrophil and T-cell counts

To further validate the associations of neutrophils and T-cells with septic shock, the differences of the neutrophil counts, CD3 counts, CD4 counts and CD8 counts between septic shock and sepsis were accessed using two-sample Wilcoxon rank sum test, where the criteria of the statistical significance was set to $p < 0.05$ (two-sided). The prognostic associations were accessed using the univariate Cox proportional hazard regression model by the survival R package. Kaplan-Meier (KM) curves were further performed by the survminer R package to evaluate prognostic association in different subgroups of sepsis. The optimized cut-off values of each group were identified using X-tile tool respectively (Camp et al., 2004).

2.10 Software and version

R (version x64 3.6.2) and Cytoscape (version 3.8.2) were used through the analysis. The artworks were created by Adobe Illustrator CC (version 64-bit 22.1).

3 Results

3.1 Identification of differentially expressed genes

As shown in the flowchart in Figure 1, we first analyzed genes differentially expressed between septic shock (n = 19) and sepsis (n =

10.3389/fgene.2023.1132361

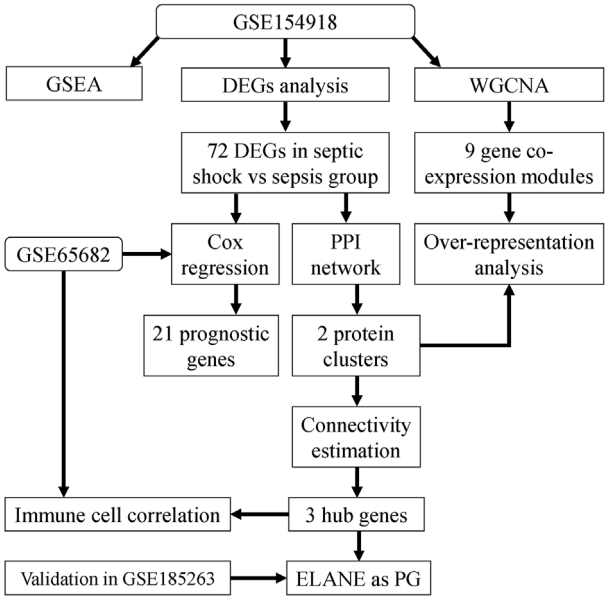


FIGURE 1
Study workflow. Bioinformatics analysis to screen core molecular mechanisms associated with septic shock and its high mortality. PPI, protein-protein interaction networks; GSEA, gene set enrichment analysis; WGCNA, weighted gene co-expression network analysis; DEG, Differential expressed gene; PG, prognostic gene.

20) groups in the discovery (GSE154918) dataset. A total of 72 septic-shock-associated DEGs were identified as septic-shock-associated genes (SSAGs) with |logFC>1| and adj.p < 0.05 (Figure 2A), among which 47 genes were upregulated and 25 genes were downregulated in septic shock. The heatmap showed the overall trend among the course from healthy to septic shock of the top 25 upregulated and 25 downregulated genes (Figure 2B). Of note, most of the DEGs showed significant increases from healthy to infection group, reflecting their possible participation in the infection-driven mechanisms.

3.2 Septic shock showed excessive neutrophil activation and impaired T Cell activation

To explore the mechanism of the development of septic shock, gene set enrichment analysis (GSEA) was used to provide global insight to assess the gene expression patterns of septic shock (Figures 2C, D).

Compared with sepsis group, we found that in septic shock, significant upregulation of the biological processed related to neutrophil activation, and those related to T cell activation were significantly downregulated (Figure 2C). Pathways related to energy metabolism were significantly upregulated, while the pathways including antigen presentation, T cell receptor (TCR) signaling pathways and NK cell mediated cytotoxity were downregulated (Figure 2D). Excessive neutrophil activation and impaired T cell activation could be the major characteristics of septic shock.

3.3 Identification of septic-shock-associated co-expression gene module

After excluding two outliers and setting soft-thresholding power to 14 (Figures 3A, B), a total of nine co-expression modules were identified based on the expression profiles of 5,000 genes with most median absolute deviation (Figure 3C). Correlational analysis between modules and phenotypes revealed positive correlations with septic shock of M4, M5, M6 and M7 and negative correlations of M1, M2 and M3. Moreover, M4 and M6 were shown to be the most positively related module to septic shock, while M2 showed the most negative correlation (Figure 3D).

Further over-representation analysis showed M4 and M6 enriched in processes and pathways about neutrophil activation while M2 and M6 enriched in those about T cell activation, suggesting the activation of neutrophil and T cell activation as key mechanism of septic shock (Figures 3E, F). In addition, M5 and M7, which showed nearly the highest positive correlations, were enriched in cell-division-related processes and RNA-metabolism-related processes, respectively.

Of note, we found modules correlated with neutrophil and T cell activation in part showed different trend among the step course from healthy to septic shock (Figure 3D), especially M2 related to T cell activation showed the most negative correlation with septic shock and infection as well as the most positive correlation with sepsis, meanwhile M6 both related to neutrophil and T cell activation as one of the most positively correlated modules with septic shock and infection showed the most negative correlation with sepsis. M4 module, which was related to neutrophil activation especially neutrophil extracellular trap (NET) formation, showed positive correlation with infection and septic shock while unsignificant correlation with sepsis. These findings further validate the specific transcriptomic changes from sepsis to septic shock.

However, T-cell-related modules were observed more perplexing associations that M6 was positively related and M2 was negatively related to septic shock, meanwhile M6 was also related to neutrophil activation. To further understand the functions of M6 in T cell activation, the gene-concept networks were constructed based on the enrichment analysis of M6 genes and their gene significances for septic shock (Supplementary Figures S1, S2). We found neutrophil-related genes showing highly consistent up-regulations while T-cell-related genes showed correlations in different directions and no obvious distribution tendency was observed. Interestingly, several HLA (Human leukocyte antigen) Class-II molecules, specifically HLA-DPA1 (Major Histocompatibility Complex, Class II, DP Alpha 1), HLA-DQA1 (Major Histocompatibility Complex, Class II, DQ Alpha 1), HLA-DRB1 (Major Histocompatibility Complex, Class II, DR Beta 1), HLA-DMB1 (Major Histocompatibility Complex, Class II, DM Beta 1) and HLA-DPB1 (Major Histocompatibility Complex, Class II, DP Beta 1), were all downregulated. On the other hand, the downregulated TCR signalling pathway were inferred based on the results of GSEA and the enrichment analysis for M2 genes (Figures 2C, D; Figure 3F). Combining these two findings, the suppression of the interaction of TCR and HLA-II might be an important part of the mechanism. Beyond that, the mechanism of dysregulation of T cell activation was still seemed more complicated.

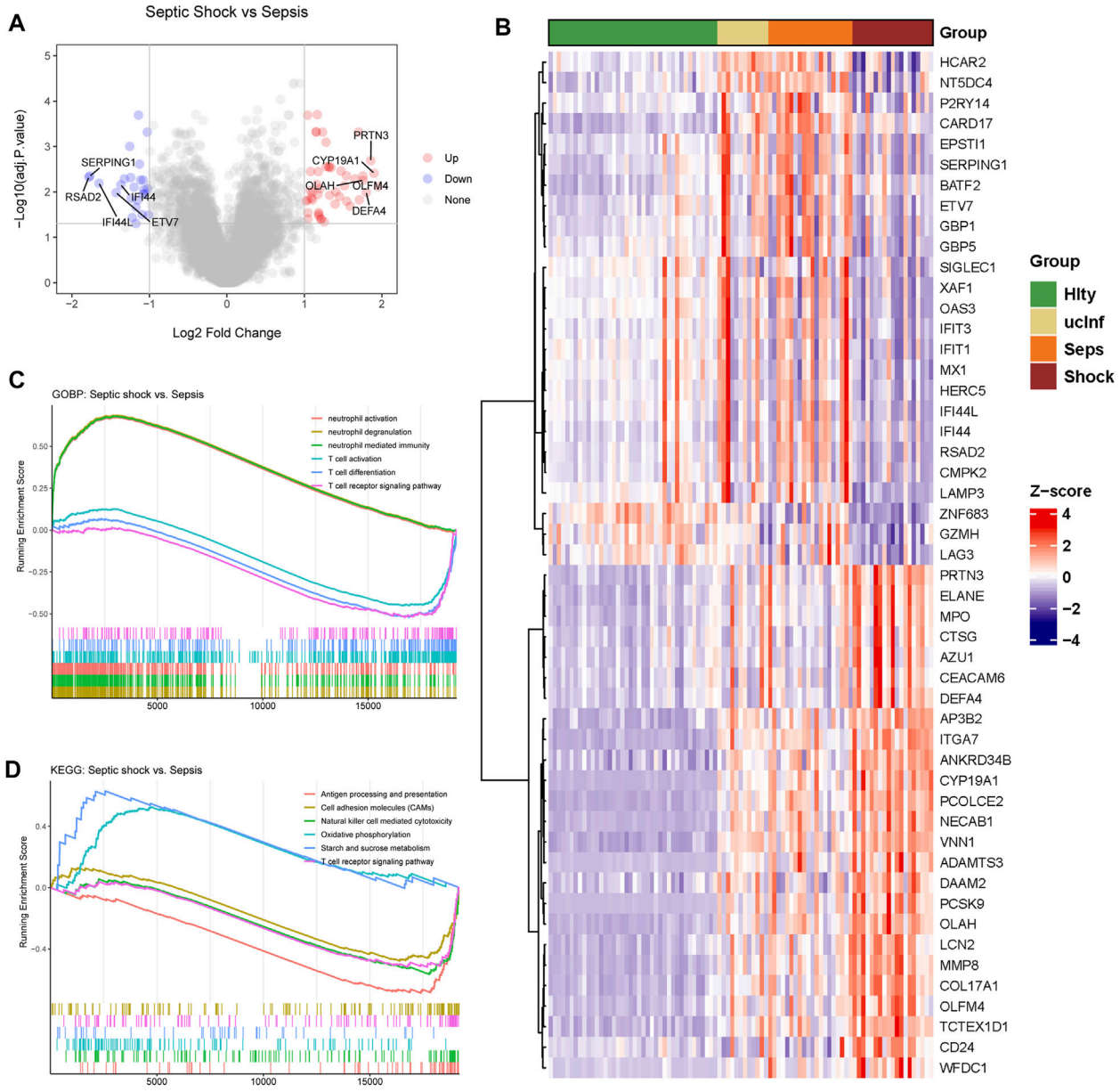


FIGURE 2
Differential gene expression between septic shock and sepsis patients. **(A)** Differently expressed genes in septic shock (n = 19) vs. sepsis (n = 20) group. **(B)** The expression heatmap of 25 top upregulated DEGs and 25 top downregulated DEGs. Hlty, healthy control; uclnf, uncomplicated infection; Seps, sepsis; Shock, septic shock. **(C)** Biological processes enriched in septic shock vs sepsis groups. **(D)** KEGG pathways enriched in septic shock vs. sepsis groups.

3.4 Identification of septic-shock-associated prognostic genes

To identify important genes associated with high mortality of septic shock. Septic-shock-associated DEGs with adjusted *p*-value <0.05 using univariate Cox analysis were further screened as septic-shock-associated prognostic genes. As a result, 21 genes were screened with significant correlations with 28-day cumulative death (Figure 4). Among them, 18 genes were identified as risky factors (log2HR > 0) and three genes were identified as protective factors. Notably, most of the risky factors belonged to M2, M4 and M6, which were identified as modules mainly related to neutrophil and T cell activation. Moreover, M2 showed the most negative correlation, and M4/M6 were two of the most positively correlated modules, reflecting their major association with septic shock and its high mortality.

3.5 Construction of PPI network

The PPI network based on 72 DEGs was distinctly separated into two clusters, and each of them showed high consistency of the trend

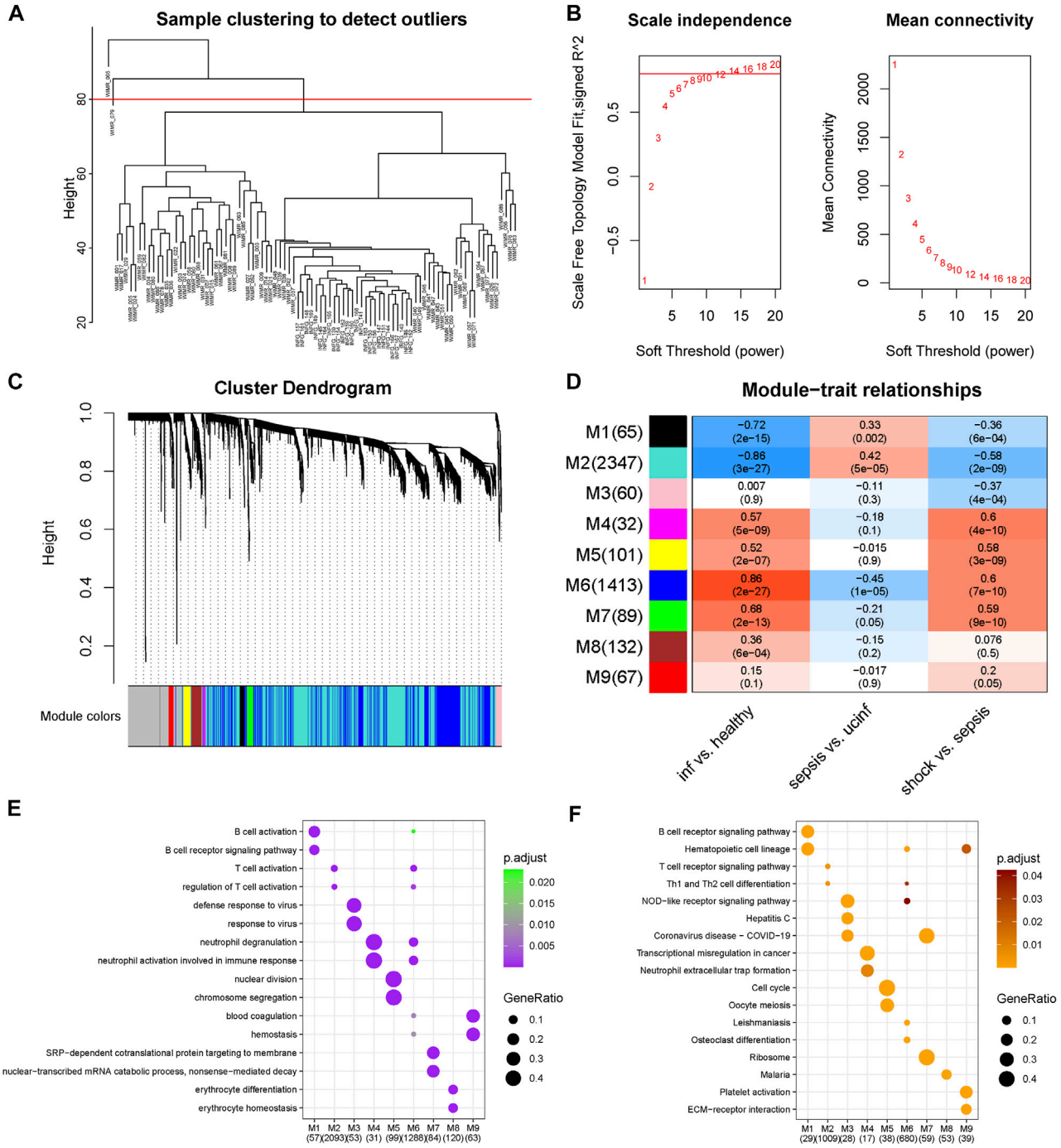


FIGURE 3
Co-expression network construction. **(A)** Hierarchical clustering dendrogram of adopted samples to detect outliers. **(B)** Determination of optimal soft-thresholding power. **(C)** Hierarchical clustering dendrogram and corresponding modules of involved genes. **(D)** Correlation heatmap of module eigengenes and phenotypes. The color of the squares gradually from blue to red represents the Spearman correlation coefficients. **(E)** Biological processes enriched in each module, the color and the size of the dots corresponded to the significance of enrichment and the ratio of enriched gene numbers in the corresponding terms respectively. **(F)** KEGG pathways enriched in each module, the color and the size of the dots corresponded to the significance of enrichment and the ratio of enriched gene numbers in the corresponding terms respectively.

of expression differences (Figure 5). All of the cluster1 genes were downregulated in septic shock, while almost all of the cluster2 genes were upregulated. Interestingly, most of the prognostic genes were concentrated in cluster2. We also found that most cluster1 genes did not show significant difference compared with healthy group. Functional over-representation analysis revealed the significant enrichment of response to virus and NOD-like receptor signaling pathway for the cluster1 genes, while

	Module	Cindex	Adj.pvalue	Log2(Hazard ratio)
CHDH	out	0.621	0.002	2.854(1.494~4.215)
INHBA	M6	0.567	0.049	0.648(0.129~1.168)
DAAM2	M6	0.602	0.004	0.502(0.215~0.789)
TCTEX1D1	M6	0.556	0.030	0.400(0.107~0.693)
WFDC1	M6	0.599	0.002	0.361(0.188~0.535)
CEACAM6	M4	0.611	0.002	0.359(0.168~0.550)
SERPINB10	M6	0.615	0.029	0.358(0.102~0.615)
PRTN3	M4	0.620	0.002	0.356(0.178~0.533)
IL1R2	M6	0.578	0.029	0.334(0.093~0.576)
ARG1	M6	0.582	0.029	0.305(0.083~0.526)
RNASE3	M4	0.581	0.008	0.277(0.106~0.447)
MPO	M4	0.606	0.003	0.264(0.119~0.409)
CD24	M4	0.594	0.005	0.264(0.111~0.416)
AZU1	M4	0.573	0.029	0.254(0.073~0.435)
ELANE	M4	0.613	0.002	0.245(0.121~0.369)
DEFA4	M4	0.602	0.002	0.228(0.111~0.344)
MS4A3	M4	0.605	0.002	0.225(0.110~0.340)
CTSG	M4	0.600	0.002	0.221(0.104~0.337)
HCAR2	out	0.581	0.033	−0.181(−0.316~−0.046)
GZMH	M2	0.598	0.005	−0.244(−0.387~−0.101)
ZNF683	M2	0.574	0.048	−1.149(−2.060~−0.238)

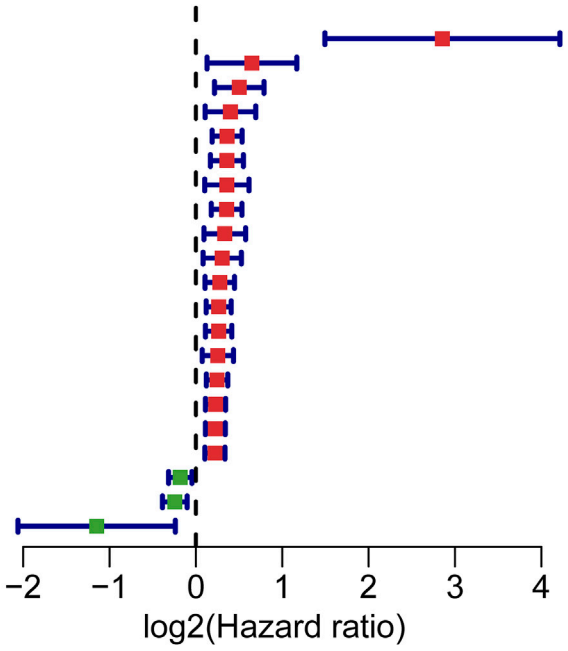


FIGURE 4
Prognostic SSAGs. The forest plots of prognostic SSAGs with adjusted *p*-value less than 0.05 by univariate cox analysis in the GSE65682 dataset. Red forest plots represent risky factors and green forest plots represent protective factors.

neutrophil-related processes and pathways such as neutrophil degranulation process and neutrophil extracellular trap (NET) formation pathway enriched in the cluster2 genes.

3.6 Identification of septic-shock-associated hub genes

Combining the results of 12 algorithms of cytoHubba app, a total of nine hub genes, including DEFA4 (Defensin Alpha 4), IFIT1 (Interferon Induced Protein With Tetratricopeptide Repeats 1), MMP8 (Matrix Metallopeptidase 8), MPO (Myeloperoxidase), MX1 (MX Dynamin Like GTPase 1), RSAD2 (Radical S-Adenosyl Methionine Domain Containing 2), ELANE (Elastase, Neutrophil Expressed), IFI44 (Interferon Induced Protein 44) and LCN2 (Lipocalin 2), were identified with at least five appearances in the top 10 results of each algorithm (Figure 6A).

On the other hand, based on the functional similarity among SSAGs calculated by the GOSemSim method, the hub genes including ELANE, IFI44, LCN2, S100A12 (S100 Calcium Binding Protein A12), CTSG (Cathepsin G), PRTN3 (Proteinase 3), OAS3 (2′-5′-Oligoadenylate Synthetase 3), IFIT3 (Interferon Induced Protein With Tetratricopeptide Repeats 3), AZU1 (Azurocidin 1) and GBP1 (Guanylate Binding Protein 1) with top 10 highest average semantic similarities were screened (Figure 6A).

We then considered the intersection of the two results above, specifically ELANE, IFI44 and LCN2, as hub genes with higher credibility (Figure 6A). ELANE and LCN2 were involved in neutrophil activation and IFI44 was involved in the response to the virus (Figure 5). Besides, ELANE and LCN2 showed persistent increase except for the period from uncomplicated infection to sepsis and were associated with increased classical monocyte as well as decreased neutrophils and memory T cells, while IFI44 was only downregulated in septic shock and showed roughly the opposite correlations with immune cell fractions (Figures 6B, C). We noticed ELANE was the only prognostic gene therein, and further validate its correlation with worse prognosis in an external dataset (GSE185263) (Figure 6D). According to the results of GSEA, upregulated neutrophil-related processes were enriched in ELANE-high group (Figure 6E). Meanwhile, downregulated TCR signaling pathway, NK cell mediated cytotoxicity and TLR signaling pathway were enriched (Figures 6E, F).

3.7 Neutrophil count was associated with septic shock and prognosis

A significant difference of neutrophil counts between septic shock and sepsis patients was validated (Figure 7A). Neutrophil counts were higher in septic shock than sepsis (13.70 vs. 10.80, $p < 0.001$). Further univariate Cox analysis demonstrated the prognostic significance of neutrophil count (Figure 7B). Moreover, KM curves showed the prognostic associations of neutrophil counts as a risky factor were not only significant for overall sepsis, but also significant for septic shock and non-shock sepsis (Figures 7C–E). However, we did not obtain any statistically different distribution or prognostic association of CD3 count, CD4 count and CD8 count (Figures 7A, B), which could be explained by the complicated mechanism of T cell activation or might be affected by the smaller sample sizes.

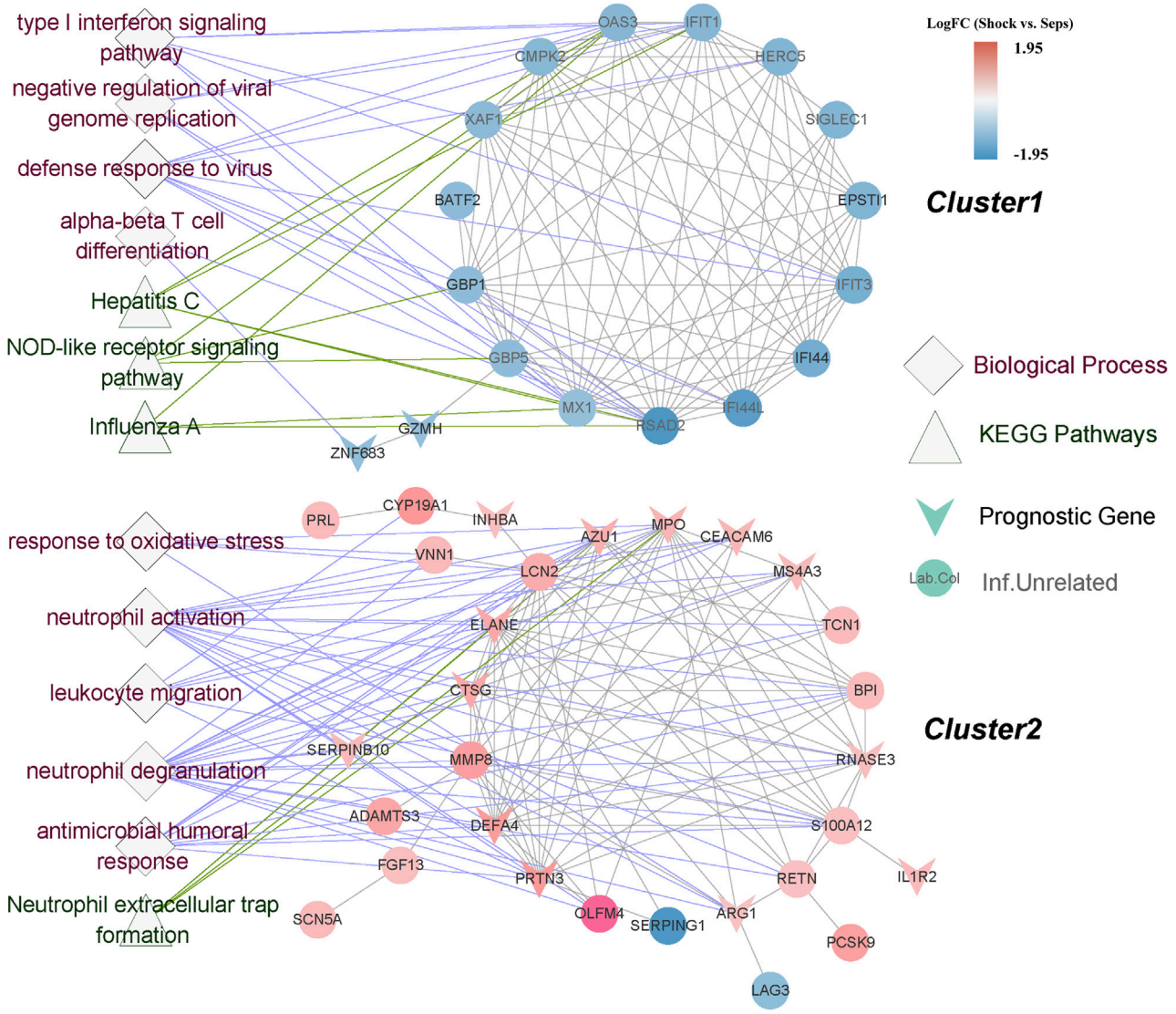


FIGURE 5
PPI network construction and over-representation analysis of cluster genes. PPI network constructed by SSAGs from STRING database, were clearly divided into 2 clusters. The color of gene nodes indicated log2 fold change. V-shape represented the prognostic genes. The diamond shaped nodes and the triangle shaped nodes represented enriched biological processes and KEGG pathways correspondingly.

4 Discussion

Sepsis causes life-threatening organ dysfunction (Singer et al., 2016), which places a great burden on human society (Fleischmann et al., 2016). Septic shock, as a subtype of sepsis, has a much higher mortality approaching 40%–60% than 10% of sepsis alone (Singer et al., 2016; Cecconi et al., 2018; Napolitano, 2018). It remains a big challenge to improve early and effective detection and management as well as the understanding of the mechanisms of septic shock. As Sepsis-3 revised in 2016 prompted new interests in sepsis immunobiology (Bermejo-Martin et al., 2016; Singer et al., 2016), further exploration of related molecular changes and underlying mechanisms could be helpful for better understanding and targeted therapy of septic shock. Most of the previous peripheral blood studies of sepsis were focused on identifying potential biomarkers, signatures, or endotypes (Scicluna et al., 2017; Baghela et al., 2022). However, since septic shock is characterized by circulatory and cellular metabolism abnormalities (Singer et al., 2016), peripheral leukocytes could be responsible for the development of septic shock.

Therefore, in the present study, we analyzed differences at the transcriptome level of peripheral blood between septic shock and sepsis. We found a significantly upregulated neutrophil activation and a dysregulated T cell activation at septic shock. The former was more likely to be associated with hyperinflammation and the latter could partially be related to suppressed interaction process of TCR and HLA-II. Interestingly, it could be inferred from the trends of module eigengenes among the step course from healthy to septic shock that neutrophils were activated when initially stimulated by infection, partially suppressed when sepsis (organ dysfunction)

10.3389/fgene.2023.1132361

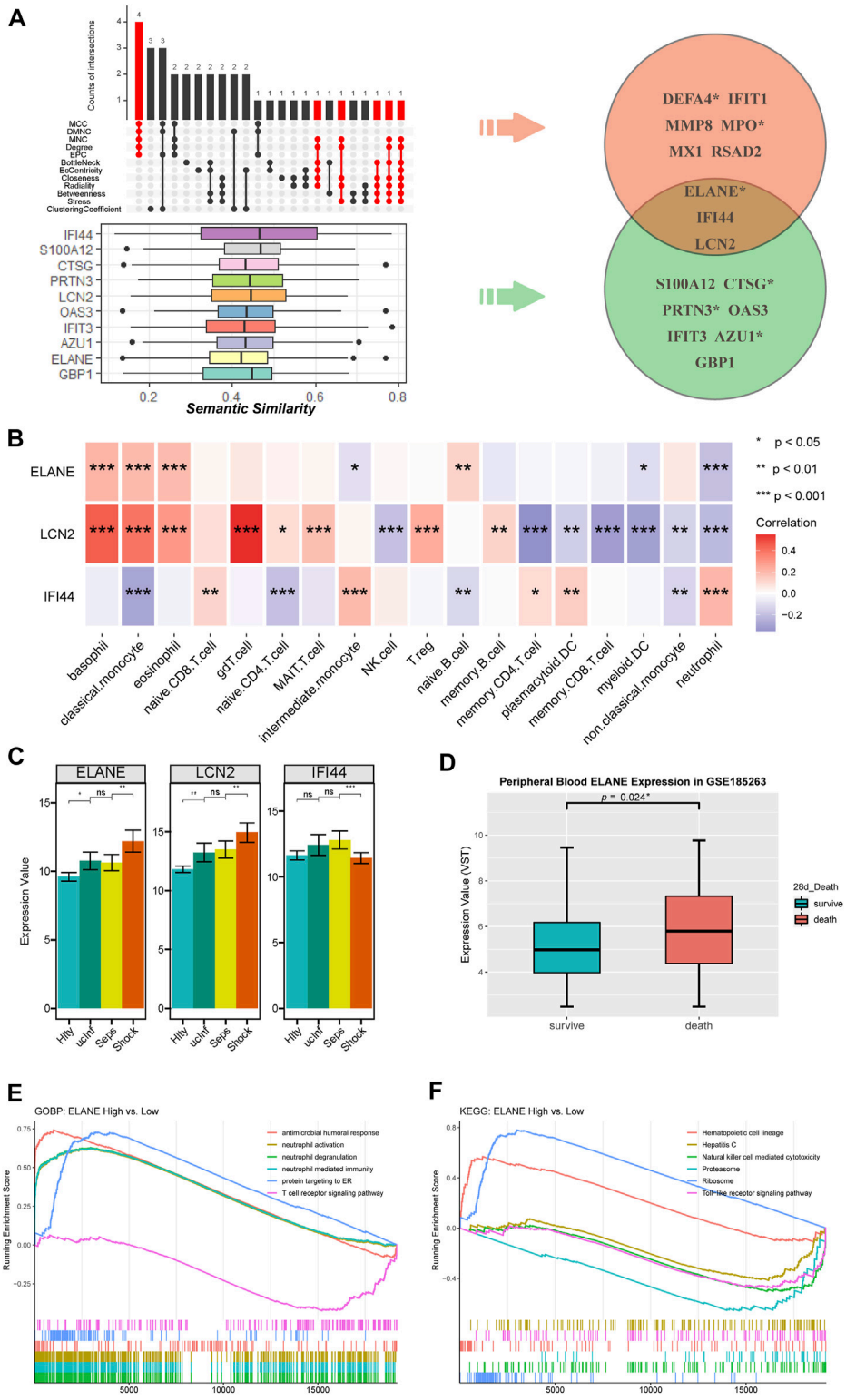


FIGURE 6
The identification of hub genes in SSAGs. **(A)** The genes appearing at least 5 times in top 10 results of each algorithm using cytoHubba were extracted (marked in red) and shown in the upset plot, below the SSAGs with top 10 functional similarities estimated by Semantic similarity analysis were shown in box plots. Three genes were identified as hub genes by overlapping the two results. The gene names marked with an asterisk indicated them as prognostic genes. **(B)** Correlations between hub genes and immune cell fractions. **(C)** Expression values of hub genes among the course from healthy to septic shock. **(D)** ELANE was significantly upregulated in death (n = 52) vs. survival (n = 293) group of the GSE185263 dataset. **(E)** Biological processes enriched in ELANE-high group. **(F)** KEGG pathways enriched in ELANE-high group.

10.3389/fgene.2023.1132361

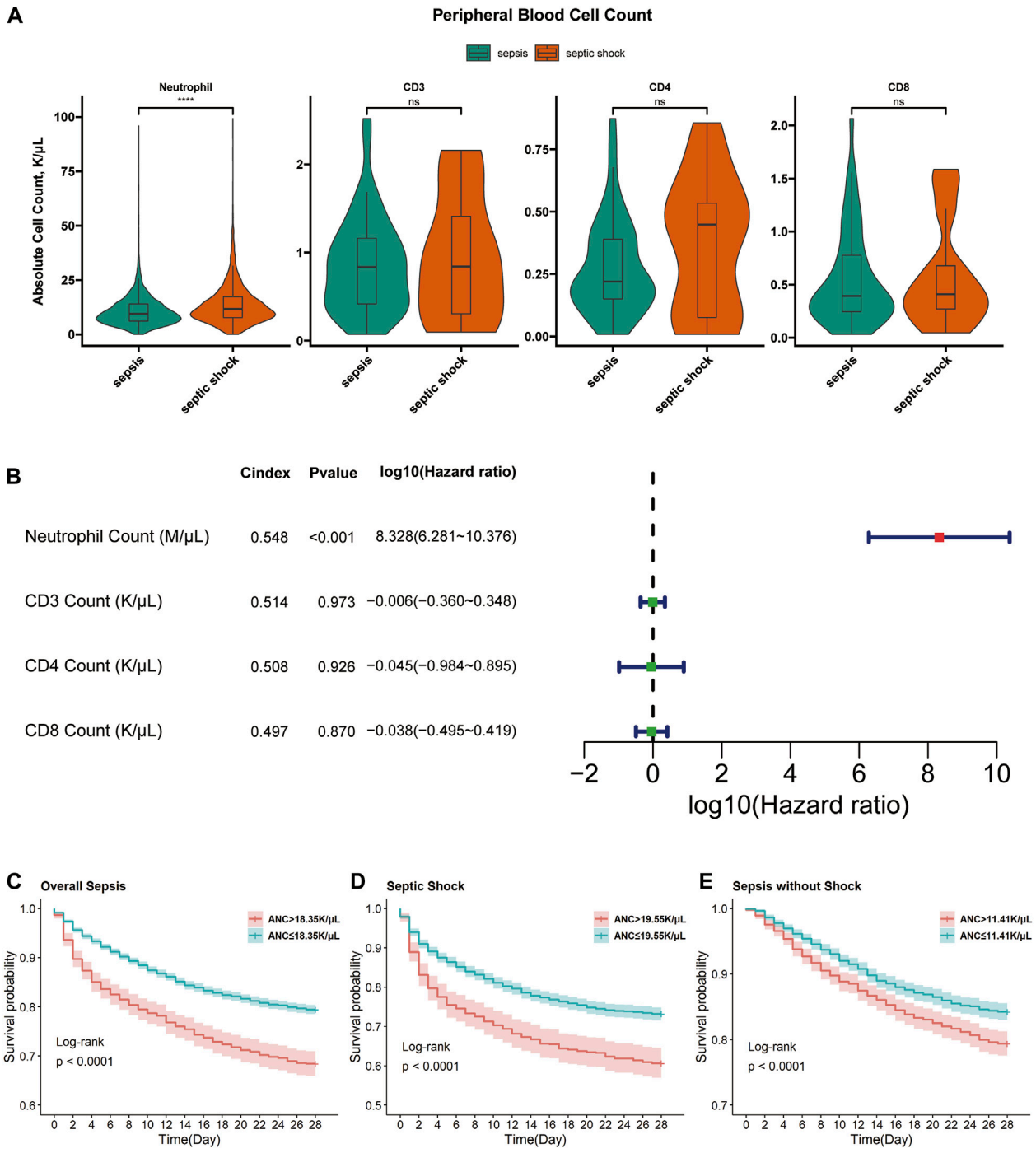


FIGURE 7
Associations of neutrophils and T cells with septic shock and prognosis of overall sepsis. **(A)** Comparisons using two-sample Wilcoxon rank sum test of neutrophil counts between sepsis (n = 4,905) and septic shock (n = 3,345) samples, CD3 count between sepsis (n = 46) and septic shock (n = 23) samples, CD4 and CD8 counts between sepsis (n = 46) and septic shock (n = 22) samples. ns, non-sense; ****, $p < 0.0001$. **(B)** The forrest plots of 4 cell counts for overall sepsis samples (n = 8,250 for neutrophil count, n = 69 for CD3 count and n = 68 for CD4 and CD8 count). Red forest plots represent risky factors and green forest plots represent protective factors. **(C)** The KM curves accessing the prognostic association of absolute neutrophil count (ANC) in overall sepsis samples. Samples with ANC>18.35 K/uL (n = 1,301) had higher 28-day cumulative mortality than samples with ANC≤18.35 K/μL (n = 6,949) (31.59% vs. 20.56%, $p < 0.0001$). **(D)** The KM curves accessing the prognostic association of ANC in septic shock samples. Samples with ANC>19.55 K/μL (n = 640) had higher 28-day cumulative mortality than samples with ANC≤19.55 K/μL (n = 2,705) (39.38% vs. 26.84%, $p < 0.0001$). **(E)** The KM curves accessing the prognostic association of ANC in non-shock sepsis samples. Samples with ANC>11.41 K/μL (n = 1827) had higher 28-day cumulative mortality than samples with ANC≤11.41 K/μL (n = 3,078) (20.63% vs 15.76%, $p < 0.0001$).

develops and finally abnormally reactivated involving the NETs formation under the situation of septic shock, while T cell activation showed more complicated changes involving modules with different trends. These could explain the phenotypic change patterns of these 2 cells at the different status from healthy to septic shock. A retrospective cross-sectional study has identified the neutrophil lymphocyte ratio (NLR) as a predictor of mortality and antibiotic responsiveness in ICU patients with septic shock and sepsis (Sari et al., 2019), suggesting both disorders as an important part of the mechanism of septic shock.

Neutrophils has been considered to play important and central roles during the early development of septic shock (Stiel et al., 2018). Neutrophils are known to acquire an immunosuppressive action during septic shock and participate in the generation of DIC where NETs exceed the regulatory and take an essential part (McDonald et al., 2017; Stiel et al., 2018). On the other hand, neutrophil activation was also significantly enriched in upregulated DEGs of septic shock compared with non-septic shock (Martinez-Paz et al., 2021), indicating the specific participation of excessive neutrophil activation in septic shock. We further demonstrated its most correlation with septic shock at the transcriptome level, and we found most of the prognostic SSAGs were concentrated in the neutrophil-related modules and cluster, which further revealed the major association of neutrophil with the high mortality of septic shock. Moreover, we validated the association of neutrophil count with the development and prognosis of septic shock based on the clinical data, indicating the great potential of neutrophil to help recognizing high-risk patients, and the prospect as an important line of the further target therapy research.

Septic shock is a time-dependent disease (Peltan et al., 2017). Early recognition of septic shock and effective targeted therapy in time could make sense to the practice of precision medicine thus is helpful to decrease the mortality of septic shock patients. Therefore, we identified ELANE, IFI44 and LCN2 as hub genes with the most connectivity, which have the potential to be more sensitive biomarkers for the detection of septic shock. ELANE and LCN2 were enriched in neutrophil related processes, while IFI44 was involved in adaptive-immune-response-related PPI cluster. As one of the prognostic genes, we noticed that ELANE was included in M4, enriched in the NET formation pathway and significantly correlated with neutrophil fractions, which could be the core part responsible for the high mortality. It encodes neutrophil elastase (NE), which is a serine protease and plays a critical role in innate host defense such as microbial killing (Horwitz et al., 1999; Voynow and Shinbashi, 2021). Under pathological conditions, NE, as one of the components of NETs, is released out of control during septic shock and has been proven to participate in multiple important mechanisms, such as chromatin decondensation and fibrinogenesis promotion (Massberg et al., 2010; Papayannopoulos et al., 2010). Besides, ELANE has been discussed to be involved specifically in the pyroptosis of neutrophil through mediating the cleavage and activation of GSDMD (Kambara et al., 2018), consistent with our findings about the correlation of upregulated ELANE with increased neutrophil activation and decreased neutrophil fraction. It has been proven that inhibition of NE synthesis can inhibit NET formation, reduce lipopolysaccharide (LPS)-induced acute lung injury in rats (Hagiwara et al., 2008; Okeke et al., 2020) and can significantly improve the survival rate of post-CLP septic rats (Kitamura et al., 1994), suggesting the great translational potential of ELANE as an important therapeutic target of septic shock. Moreover, previous transcriptomic studies have been published about identifying ELANE as an important signature related to the severity (SOFA score) (Baghela et al., 2022), and prognosis (Ding et al., 2022; Zhang et al., 2022), of sepsis patients. We further demonstrated the correlation of ELANE with septic shock and its vital participation in the core mechanism of septic shock. As for the other two, LCN2 encodes a secreted protein called neutrophil gelatinase-associated lipocalin (NGAL). It can be stimulated by Toll-like receptors and is pivotal in the innate immune response to bacterial infection through binding bacterial siderophores (Flo et al., 2004). LCN2 has been proven to differentially expressed between septic shock and sepsis in surgical patients (Martin-Fernandez et al., 2020), and has been reported as a potential biomarker of septic-shock-associated acute kidney injury (Tang et al., 2021). IFI44 is an interferon-alpha inducible protein associated with infection by several viruses (Power et al., 2015). The downregulation of IFI44 could in a way represent the suppressed adaptive immune response in septic shock. The specific roles of IFI44 in septic shock have not been defined yet. In summary, ELANE and LCN2 were enriched in neutrophil activation and correlated with infection and septic shock, especially ELANE as the only prognostic gene could participate through NETs formation and pyroptosis pathways. IFI44 was associated with adaptive immune response and specifically downregulated in septic shock. All three hub genes did not show any significant change between sepsis and uncomplicated infection.

Our study is the first in our knowledge to focus on the mechanisms about the contribution of peripheral leukocyte to the development of septic shock under the new Sepsis-3 definition since 2016. The roles of neutrophil activation and NETs in septic shock have been reported in previous studies. We further demonstrated their most correlation with septic shock combining the WGCNA and PPI network analysis. We highlighted septic shock as a subset of sepsis with much higher mortality showing different expression profiles in peripheral blood. The distinct distribution bias that most of the septic-shock-associated prognostic genes were concentrated in neutrophil-related modules and PPI cluster may be worthy of note. However, there are still some limitations of the study. The relatively small sample size of GSE154918 could partially limit the universal implication of our findings, however the consistency between the results of co-expression network and PPI network analysis could up to a point improve the credibility and represent an important endotype of septic shock. More importantly, the specific phenotypic changes of neutrophil and T cells, for example, whether ELANE was actually or only upregulated in neutrophil, and whether or how ELANE-mediated pyroptosis of neutrophil take part in were still unclear. The underlying cellular heterogenicity might partially reduce the credibility of our findings and the further well-designed research with directed focus is still needed.

Overall, septic shock showed an excessive neutrophil activation and a dysregulated T cell activation, of which the former was associated with hyperinflammation and the latter

could partially be related to suppressed interaction process of TCR and HLA-II. Neutrophil activation may play a core role during septic shock. ELANE, LCN2 and IFI44 were identified as hub genes during septic shock, among which ELANE as a neutrophil-related gene might have the greatest potential to be a clinical biomarker and therapeutic target. This study highlighted an important perspective about septic shock under the new definition and would help in designing further translational research to improve diagnosis and treatment.

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

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

Author contributions

SZ: study concept, data curation, formal analysis and visualization, original draft preparation. KZ: data curation, formal analysis and visualization. XL: review and editing. XZ: study concept, data curation. YZ: data curation. ZeL: validation. ZiL: validation. YX: study concept. DL: review and editing. NJ: review and editing, supervision. QS: review and editing, supervision. All authors have read and agreed to the published version of the manuscript.

Funding

This work was supported by the National Natural Science Foundation of China (No. 81901989) the Natural Science Foundation of Zhejiang Province (LY21H150005); and the Top-Notch Youth Talent Cultivation Project of Independent Design Project of National Clinical Research Center for Child Health (Q21B0007).

Acknowledgments

We would like to appreciate all of the data contributors of this study and all of the patients and volunteers who participated in this study.

Conflict of interest

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

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2023.1132361/full#supplementary-material

References

Baghela, A., Pena, O. M., Lee, A. H., Baquir, B., Falsafi, R., An, A., et al. (2022). Predicting sepsis severity at first clinical presentation: The role of endotypes and mechanistic signatures. *EBioMedicine* 75, 103776. doi:10.1016/j.ebiom.2021.103776

Barrett, T., Wilhite, S. E., Ledoux, P., Evangelista, C., Kim, I. F., Tomashevsky, M., et al. (2013). NCBI GEO: Archive for functional genomics data sets--update. *Nucleic Acids Res.* 41, D991–D995. doi:10.1093/nar/gks1193

Bermejo-Martin, J. F., Andaluz-Ojeda, D., Almansa, R., Gandia, F., Gomez-Herreras, J. I., Gomez-Sanchez, E., et al. (2016). Defining immunological dysfunction in sepsis: A requisite tool for precision medicine. *J. Infect.* 72, 525–536. doi:10.1016/j.jinf.2016.01.010

Camp, R. L., Dolled-Filhart, M., and Rimm, D. L. (2004). X-Tile: A new bio-informatics tool for biomarker assessment and outcome-based cut-point optimization. *Clin. Cancer Res.* 10, 7252–7259. doi:10.1158/1078-0432.CCR-04-0713

Cecconi, M., Evans, L., Levy, M., and Rhodes, A. (2018). Sepsis and septic shock. *Lancet* 392, 75–87. doi:10.1016/S0140-6736(18)30696-2

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

Ding, N., Xu, X., Wang, Y., Li, H., Cao, Y., and Zheng, L. (2022). Contribution of prognostic ferroptosis-related subtypes classification and hub genes of sepsis. *Transpl. Immunol.* 74, 101660. doi:10.1016/j.trim.2022.101660

Esposito, S., De Simone, G., Boccia, G., De Caro, F., and Pagliano, P. (2017). Sepsis and septic shock: New definitions, new diagnostic and therapeutic approaches. *J. Glob. Antimicrob. Resist* 10, 204–212. doi:10.1016/j.jgar.2017.06.013

Fleischmann, C., Scherag, A., Adhikari, N. K., Hartog, C. S., Tsaganos, T., Schlattmann, P., et al. (2016). Assessment of global incidence and mortality of hospital-treated sepsis. Current estimates and limitations. *Am. J. Respir. Crit. Care Med.* 193, 259–272. doi:10.1164/rccm.201504-0781OC

Flo, T. H., Smith, K. D., Sato, S., Rodriguez, D. J., Holmes, M. A., Strong, R. K., et al. (2004). Lipocalin 2 mediates an innate immune response to bacterial infection by sequestrating iron. *Nature* 432, 917–921. doi:10.1038/nature03104

Goldberger, A. L., Amaral, L. A., Glass, L., Hausdorff, J. M., Ivanov, P. C., Mark, R. G., et al. (2000). PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. *Circulation* 101, E215–E220. doi:10.1161/01.cir.101.23.e215

Gorecki, G., Cochior, D., Moldovan, C., and Rusu, E. (2021). Molecular mechanisms in septic shock (Review). *Exp. Ther. Med.* 22, 1161. doi:10.3892/etm.2021.10595

Hagiwara, S., Iwasaka, H., Togo, K., and Noguchi, T. (2008). A neutrophil elastase inhibitor, sivelestat, reduces lung injury following endotoxin-induced shock in rats by inhibiting HMGB1. *Inflammation* 31, 227–234. doi:10.1007/s10753-008-9069-z

Herwanto, V., Tang, B., Wang, Y., Shojaei, M., Nalos, M., Shetty, A., et al. (2021). Blood transcriptome analysis of patients with uncomplicated bacterial infection and sepsis. *BMC Res. Notes* 14, 76. doi:10.1186/s13104-021-05488-w

Horwitz, M., Benson, K. F., Person, R. E., Aprikyan, A. G., and Dale, D. C. (1999). Mutations in ELA2, encoding neutrophil elastase, define a 21-day biological clock in cyclic haematopoiesis. *Nat. Genet.* 23, 433–436. doi:10.1038/70544

Johnson, A., Bulgarelli, L., Pollard, T., Horng, S., Celi, L. A., and Mark, R. 2022 MIMIC-IV (version 2.0). doi:10.13026/7vcr-e114

Kambara, H., Liu, F., Zhang, X., Liu, P., Bajrami, B., Teng, Y., et al. (2018). Gasdermin D exerts anti-inflammatory effects by promoting neutrophil death. *Cell Rep.* 22, 2924–2936. doi:10.1016/j.celrep.2018.02.067

Kitamura, A., Takahashi, K., Okajima, A., and Kitamura, N. (1994). Induction of the human gene for p44, a hepatitis-C-associated microtubular aggregate protein, by interferon-alpha/beta. *Eur. J. Biochem.* 224, 877–883. doi:10.1111/j.1432-1033.1994.00877.x

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

Liang, L., Moore, B., and Soni, A. (2006). "National inpatient hospital costs: The most expensive conditions by payer, 2017: Statistical brief #261," in *Healthcare cost and utilization Project (HCUP) statistical briefs* (Rockville (MD): Agency for Healthcare Research and Quality US).

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

Martin-Fernandez, M., Vaquero-Roncero, L. M., Almansa, R., Gomez-Sanchez, E., Martin, S., Tamayo, E., et al. (2020). Endothelial dysfunction is an early indicator of sepsis and neutrophil degranulation of septic shock in surgical patients. *BJS Open* 4, 524–534. doi:10.1002/bjs5.50265

Martinez-Paz, P., Aragon-Camino, M., Gomez-Sanchez, E., Lorenzo-Lopez, M., Gomez-Pesquera, E., Fadrique-Fuentes, A., et al. (2021). Distinguishing septic shock from non-septic shock in postsurgical patients using gene expression. *J. Infect.* 83, 147–155. doi:10.1016/j.jinf.2021.05.039

Massberg, S., Grahl, L., Von Bruehl, M. L., Manukyan, D., Pfeiler, S., Goosmann, C., et al. (2010). Reciprocal coupling of coagulation and innate immunity via neutrophil serine proteases. *Nat. Med.* 16, 887–896. doi:10.1038/nm.2184

Mcdonald, B., Davis, R. P., Kim, S. J., Tse, M., Esmon, C. T., Kolaczkowska, E., et al. (2017). Platelets and neutrophil extracellular traps collaborate to promote intravascular coagulation during sepsis in mice. *Blood* 129, 1357–1367. doi:10.1182/blood-2016-09-741298

Napolitano, L. M. (2018). Sepsis 2018: Definitions and guideline changes. *Surg. Infect. (Larchmt)* 19, 117–125. doi:10.1089/sur.2017.278

Newman, A. M., Steen, C. B., Liu, C. L., Gentles, A. J., Chaudhuri, A. A., Scherer, F., et al. (2019). Determining cell type abundance and expression from bulk tissues with digital cytometry. *Nat. Biotechnol.* 37, 773–782. doi:10.1038/s41587-019-0114-2

Okeke, E. B., Louttit, C., Fry, C., Najafabadi, A. H., Han, K., Nemzek, J., et al. (2020). Inhibition of neutrophil elastase prevents neutrophil extracellular trap formation and rescues mice from endotoxic shock. *Biomaterials* 238, 119836. doi:10.1016/j.biomaterials.2020.119836

Papayannopoulos, V., Metzler, K. D., Hakkim, A., and Zychlinsky, A. (2010). Neutrophil elastase and myeloperoxidase regulate the formation of neutrophil extracellular traps. *J. Cell Biol.* 191, 677–691. doi:10.1083/jcb.201006052

Pavez, N., Kattan, E., Vera, M., Ferri, G., Valenzuela, E. D., Alegria, L., et al. (2020). Hypoxia-related parameters during septic shock resuscitation: Pathophysiological determinants and potential clinical implications. *Ann. Transl. Med.* 8, 784. doi:10.21037/atm-20-2048

Peltan, I. D., Mitchell, K. H., Rudd, K. E., Mann, B. A., Carlbom, D. J., Hough, C. L., et al. (2017). Physician variation in time to antimicrobial treatment for septic patients presenting to the emergency department. *Crit. Care Med.* 45, 1011–1018. doi:10.1097/CCM.0000000000002436

Power, D., Santoso, N., Dieringer, M., Yu, J., Huang, H., Simpson, S., et al. (2015). IFI44 suppresses HIV-1 LTR promoter activity and facilitates its latency. *Virology* 481, 142–150. doi:10.1016/j.virol.2015.02.046

Reinhart, K., Daniels, R., Kissoon, N., Machado, F. R., Schachter, R. D., and Finfer, S. (2017). Recognizing sepsis as a global health priority - a WHO resolution. *N. Engl. J. Med.* 377, 414–417. doi:10.1056/NEJMp1707170

Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., et al. (2015). Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* 43, e47. doi:10.1093/nar/gkv007

Rubio, I., Osuchowski, M. F., Shankar-Hari, M., Skirecki, T., Winkler, M. S., Lachmann, G., et al. (2019). Current gaps in sepsis immunology: New opportunities for translational research. *Lancet Infect. Dis.* 19, e422–e436. doi:10.1016/S1473-3099(19)30567-5

Rudd, K. E., Johnson, S. C., Agesa, K. M., Shackelford, K. A., Tsoi, D., Kievlan, D. R., et al. (2020). Global, regional, and national sepsis incidence and mortality, 1990-2017: Analysis for the global burden of disease study. *Lancet* 395, 200–211. doi:10.1016/S0140-6736(19)32989-7

Sari, R., Karakurt, Z., Ay, M., Celik, M. E., Yalaz Tekan, U., Ciyiltepe, F., et al. (2019). Neutrophil to lymphocyte ratio as a predictor of treatment response and mortality in septic shock patients in the intensive care unit. *Turk J. Med. Sci.* 49, 1336–1349. doi:10.3906/sag-1901-105

Scicluna, B. P., Van Vught, L. A., Zwinderman, A. H., Wiewel, M. A., Davenport, E. E., Burnham, K. L., et al. (2017). Classification of patients with sepsis according to blood genomic endotype: A prospective cohort study. *Lancet Respir. Med.* 5, 816–826. doi:10.1016/S2213-2600(17)30294-1

Shankar-Hari, M., Phillips, G. S., Levy, M. L., Seymour, C. W., Liu, V. X., Deutschman, C. S., et al. (2016). Developing a new definition and assessing new clinical criteria for septic shock: For the third international Consensus definitions for sepsis and septic shock (Sepsis-3). *JAMA* 315, 775–787. doi:10.1001/jama.2016.0289

Singer, M., Deutschman, C. S., Seymour, C. W., Shankar-Hari, M., Annane, D., Bauer, M., et al. (2016). The third international Consensus definitions for sepsis and septic shock (Sepsis-3). *JAMA* 315, 801–810. doi:10.1001/jama.2016.0287

Stiel, L., Meziani, F., and Helms, J. (2018). Neutrophil activation during septic shock. *Shock* 49, 371–384. doi:10.1097/SHK.0000000000000980

Szklarczyk, D., Gable, A. L., Nastou, K. C., Lyon, D., Kirsch, R., Pyysalo, S., et al. (2021). The STRING database in 2021: Customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets. *Nucleic Acids Res.* 49, D605–D612. doi:10.1093/nar/gkaa1074

Tang, Y., Yang, X., Shu, H., Yu, Y., Pan, S., Xu, J., et al. (2021). Bioinformatic analysis identifies potential biomarkers and therapeutic targets of septic-shock-associated acute kidney injury. *Hereditas* 158, 13. doi:10.1186/s41065-021-00176-y

Torio, C. M., and Andrews, R. M. (2006). "National inpatient hospital costs: The most expensive conditions by payer, 2011: Statistical brief #160," in *Healthcare cost and utilization Project (HCUP) statistical briefs* (Rockville (MD).

Torio, C. M., and Moore, B. J. (2006). "National inpatient hospital costs: The most expensive conditions by payer, 2013: Statistical brief #204," in *Healthcare cost and utilization Project (HCUP) statistical briefs* (Rockville (MD).

Voynow, J. A., and Shinbashi, M. (2021). Neutrophil elastase and chronic lung disease. *Biomolecules* 11, 1065. doi:10.3390/biom11081065

Wang, J. Z., Du, Z., Payattakool, R., Yu, P. S., and Chen, C. F. (2007). A new method to measure the semantic similarity of GO terms. *Bioinformatics* 23, 1274–1281. doi:10.1093/bioinformatics/btm087

Yu, G. (2020). Gene ontology semantic similarity analysis using GOSemSim. *Methods Mol. Biol.* 2117, 207–215. doi:10.1007/978-1-0716-0301-7_11

Yu, G., Li, F., Qin, Y., Bo, X., Wu, Y., and Wang, S. (2010). GOSemSim: an R package for measuring semantic similarity among GO terms and gene products. *Bioinformatics* 26, 976–978. doi:10.1093/bioinformatics/btq064

Yu, G., Wang, L. G., Han, Y., and He, Q. Y. (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* 16, 284–287. doi:10.1089/omi.2011.0118

Zhang, B., and Horvath, S. (2005). A general framework for weighted gene co-expression network analysis. *Stat. Appl. Genet. Mol. Biol.* 4, Article17. doi:10.2202/1544-6115.1128

Zhang, Q., Hu, Y., Wei, P., Shi, L., Shi, L., Li, J., et al. (2022). Identification of hub genes for adult patients with sepsis via RNA sequencing. *Sci. Rep.* 12, 5128. doi:10.1038/s41598-022-09175-z

Glossary

DEG Differentially expressed gene
SSAG Septic-shock-associated gene
BH Benjamini–Hochberg method
WGCNA weighted gene co-expression network analysis
PPI protein–protein interaction network
CMR cumulative mortality rate
adj.P adjusted *p*-value
BH Benjamini–Hochberg method
GSEA gene set enrichment analysis
GEO gene expression omnibus
logFC log2 fold change
BP biological process
TCR T cell receptor
HLA Human leukocyte antigen
HLA-DPA1 Major Histocompatibility Complex, Class II, DP Alpha 1
HLA-DQA1 Major Histocompatibility Complex, Class II, DQ Alpha 1)
HLA-DRB1 Major Histocompatibility Complex, Class II, DR Beta 1
HLA-DMB1 Major Histocompatibility Complex, Class II, DM Beta 1
HLA-DPB1 Major Histocompatibility Complex, Class II, DP Beta
ELANE Elastase, Neutrophil Expressed
IFI44 Interferon Induced Protein 44
LCN2 Lipocalin 2
DEFA4 Defensin Alpha 4
IFIT1 Interferon Induced Protein With Tetratricopeptide Repeats 1
MMP8 Matrix Metallopeptidase 8
MPO Myeloperoxidase
MX1 MX Dynamin Like GTPase 1
RSAD2 Radical S-Adenosyl Methionine Domain Containing 2
S100A12 S100 Calcium Binding Protein A12
CTSG Cathepsin G
PRTN3 Proteinase 3
OAS3 2′-5′-Oligoadenylate Synthetase 3
IFIT3 Interferon Induced Protein With Tetratricopeptide Repeats 3
AZU1 Azurocidin 1
GBP1 Guanylate Binding Protein 1
DIC disseminated intravascular coagulation
NET neutrophil extracellular trap
NE neutrophil elastase.

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 24 February 2023
DOI 10.3389/fgene.2023.1108805

OPEN ACCESS

EDITED BY
Shuai Liu,
University of Hawaii at Manoa,
United States

REVIEWED BY
Yunting Wang,
University of Houston, United States
Chen Li,
Free University of Berlin, Germany

*CORRESPONDENCE
Huijing Huang,
fangfeijin90@163.com
Yun Sun,
syun163@163.com

†These authors have contributed equally to this work

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

RECEIVED 26 November 2022
ACCEPTED 15 February 2023
PUBLISHED 24 February 2023

CITATION
Dong X, Zhou M, Li X, Huang H and Sun Y (2023), Gene profiling reveals the role of inflammation, abnormal uterine muscle contraction and vascularity in recurrent implantation failure.
Front. Genet. 14:1108805.
doi: 10.3389/fgene.2023.1108805

Gene profiling reveals the role of inflammation, abnormal uterine muscle contraction and vascularity in recurrent implantation failure

Xinyi Dong[1,2†], Mi Zhou[3†], Xinyu Li[1,2†], Huijing Huang[4*] and Yun Sun[1,2*]

[1]Center for Reproductive Medicine, Renji Hospital, School of Medicine, Shanghai Jiao Tong University, Shanghai, China, [2]Shanghai Key Laboratory for Assisted Reproduction and Reproductive Genetics, Shanghai, China, [3]Department of Rheumatology, Renji Hospital, School of Medicine, Shanghai Jiaotong University, Shanghai, China, [4]Department of Rheumatology, Zhongshan Hospital, Fudan University, Shanghai, China

Objective: Recurrent implantation failure (RIF) is now disturbing numerous infertile couples accepting assisted reproductive technology (ART). And the endometrial factors are crucial causes of recurrent implantation failure. However, its mechanism is still unclear. Thus, the aim of this study is to identify altered biologic processes in endometrium that may contribute to recurrent implantation failure.

Methods: We recruited two microarray datasets (GSE103465, GSE111974) from Gene Expression Omnibus database (GEO), which contain endometrium from RIF and normal women during implantation period. Using the online tools GEO2R and Venny, we identified Differentially Expressed Genes (DEGs) of selected datasets, and obtained common DEGs. Gene Ontology (GO) terms, Kyoto Encyclopedia of Genes and Genomes (KEGG) and BioCatar pathway enrichment were conducted with Enrichr platform, "ssgsea" and "ggplot2" package of RStudio. PPI networks and hub gene related TF-gene interaction and TF-miRNA co-regulation networks were built *via* online tools STRING and NetworkAnalyst. Immune infiltration analysis was performed by CIBERSORT platform. Recurrent implantation failure subgroup identification was achieved through "ConsensusClusterPlus," "tsne," "ssgsea", and "ggpubr" package in RStudio. Diagnostic characteristic ROC curves were constructed *via* "pROC" and "ggplot2" package of RStudio. Enrichr platform was utilized to find drugs targeting hub genes.

Results: 26 common DEGs were confirmed. Gene Ontology and Kyoto Encyclopedia of Genes and Genomes/BioCarta analysis determined common DEGs were mainly enriched in inflammation associated pathways including TNF, NF-κB, IL-4, IL-10, IL-6, and TGF-β signaling pathways. Five hub genes (*PTGS2*, *VCAM1*, *EDNRB*, *ACTA2*, and *LIF*) and related TF-gene and TF-miRNA interactions were identified. Immune infiltration analysis indicated the importance of macrophage M2 in recurrent implantation failure patients. Importantly, subgroup identification analysis highlighted that recurrent implantation failure patients can be divided into two subgroups with different phenotypes. Moreover, the ROC curves and drugs may provide new diagnostic and therapeutic thought for recurrent implantation failure.

KEYWORDS
recurrent implantation failure, inflammation, contraction, vascularity, expression profiling

Introduction

Nowadays, infertility depresses 8–12% of couples in reproductive age worldwide, and the boom of assisted reproductive technology (ART) has allowed numerous infertile couples to achieve feasible pregnancy (INHORN AND PATRIZIO, 2015). However, a challenging problem arising in this domain is recurrent implantation failure (RIF) (Bashiri et al., 2018) As far as we know, there is still lacking a world-wide acknowledged formal definition of RIF, but a relatively recognized definition is that RIF is failure in three *in vitro* fertilization-embryo transfer (IVF-ET) cycles after transferring good quality embryos (Orvieto et al., 2015; Bashiri et al., 2018). Among patients under infertility treatment, 15% suffer from RIF (Busnelli et al., 2020; Mrozikiewicz et al., 2021).

Implantation is a complex process requiring precise embryo-uterine cross-talk, which is still not well understood (Mrozikiewicz et al., 2021). The window of implantation (WOI) is a strict time span when blastocyst is overlain on the receptive state of the endometrium. Abnormality of each link in implantation can lead to RIF (Mrozikiewicz et al., 2021).

Risk factors of RIF include maternal age, smoking, stress and so on (Orvieto et al., 2015; Bashiri et al., 2018). Immunological factors including peripheral and uterine natural killer cells, Th1/Th2 ratio, tumor necrosis factor alpha (TNF-α) levels, auto-antibodies, antiphospholipid syndrome, hereditary thrombophilia as well as infection are considered to participated in the pathogenesis of RIF (Bashiri et al., 2018). Endometrium is the place of embryo to locate, adhere, penetrate and develop in. Abnormal status of endometrium, such as chronic endometritis, embryo-endometrial asynchrony, endometrial injuries (e.g., pipelle catheter, hysteroscopy and saline infusion) are factors to explain the origin of RIF (Bellver and Simón, 2018). Previous studies have reported that two-thirds of the RIF are caused by the abnormality of endometrial receptivity, so it is of great importance to focus on the role of endometrium in RIF (Margalioth et al., 2006). Various therapies are now being explored to treat RIF including different types and methods of embryo transfer, ovulation induction protocol, progesterone support, antithrombotic agents, immunotherapy, anti-infection, anatomical intervention and so on (Bashiri et al., 2018). However, the problem isn't fully resolved by the above-mentioned therapies, and we still dont have an ideal method to detect the causes of every RIF individual. Therefore, it is urgent to conduct bioinformatic analysis aiming to find potential mechanism and effective treatment of RIF.

Accumulating evidence has demonstrated that many genes have been proposed as potential receptivity markers, however, considering heterogeneity among those independent experiments as an outcome of variations in specimens or tissue and different data processing methods, the identification of those Differentially Expressed Genes (DEGs) is inconsistent. Therefore, in this research, we attempt to find more effective and reliable biological pathways and potential biomarkers contributing to the pathogenesis and development of RIF *via* integrating different studies. We downloaded two microarray datasets GSE103465 and GSE111974 from Gene Expression Omnibus database (GEO), which contain gene expression profiling from endometrial tissues of women with RIF and fertile women during WOI. We then performed further bioinformatic analysis, including common DEGs identification, gene ontology (GO), Kyoto Encyclopedia of Genes and Genomes (KEGG)/BioCarta pathway enrichment, protein-protein interaction (PPI) network analysis, TF-genes and TF-miRNA interaction analysis, subgroup identification, immune infiltration analysis, characteristic (ROC) curve analysis and drug searching The workflow of our analysis is displayed in Figure 1.

Materials and methods

Original data collection

We input "recurrent implantation failure" and "expression profile" as two keywords to the GEO database, then two datasets GSE103465 and GSE111974 were selected for analysis, and GSE26787 was chosen for validation. Both GSE103465 and GSE111974 contain expression profiles of endometrial tissue obtained from RIF and control women during WOI. GSE103465, in GPL16043 platform, contains whole-genome expression profiles of endometrial tissue from three women divided to the control group and RIF group (Guo et al., 2018). In GSE103465, RIF is defined as no pregnancy after ≥3 embryo transfers including a total of ≥4 good-quality embryos, and inclusion criteria of control group is infertile women with tubal factors who achieved a clinical pregnancy after the first embryo transfer (Guo et al., 2018). GSE111974, in GPL17077 platform, consists of 24 individuals with RIF and 24 fertile control patients, in which RIF is determined as failure of pregnancy in three consecutive IVF cycles with at least one transfer of good quality embryo in each cycle, while the fertile control refers to patients who had a history of at least one live birth with no related comorbidities (Bastu et al., 2019). The platform and series matrix files were all downloaded.

Analysis for DEGs

Using the online analysis tool GEO2R (https://www.ncbi.nlm.nih.gov/geo/geo2r/), the expression profiles of endometrium from RIF patients and fertile women were compared to screen the DEGs of the two datasets, independently. p values were calculated through t-tests and genes with the criteria of a $|\log2\ (\text{fold-change})| > 1$ and $p\text{-value} < 0.05$ were considered as DEGs. The volcano plot and box diagram were both created *via* the "ggplot2" package of RStudio software, and the heatmap for the DEGs was drawn using the "ComplexHeatmap" package of RStudio software (Gu et al., 2016; Wickham, 2016). Overlapping DEGs from two databases were defined as common DEGs and were displayed with Venn diagrams, which was drawn by utilizing the online platform Venny 2.1 (http://bioinfogp.cnb.csic.es/tools/venny/index.html).

Gene set enrichment analysis for GO terms, KEGG and BIOCARTA pathway finding

The GO terms of common DEGs were conducted with online tool Enrichr (https://amp.pharm.mssm.edu/Enrichr/). Significantly enriched function annotations were defined as GO terms and KEGG pathways with p values of <0.05 (Zhou et al., 2019). The GO analysis, including biological process (BP), cellular component (CC) and molecular function (MF), provides a set of concepts for describing molecular activity and the location where the genes execute their functions (Ashburner et al., 2000). The bubble plot of KEGG pathways used to understand specific metabolic pathways of common DEGs, were visualized *via* "ggplot2" package of RStudio software (Kanehisa and Goto, 2000; Wickham, 2016). To take a further step on gene enrichment in individual samples, we performed a single sample version of gene set enrichment analysis (ssGSEA) by "GSVA" package of RStudio software, which rules an enrichment score as the degree of absolute enrichment of a gene dataset in each sample with a designated database called BioCarta (Hänzelmann et al., 2013). The differences of enrichment scores were identified *via* wilcoxon test, and were visualized by "ggpubr" package in RStudio software (Kassambara, 2020).

Construction of PPI networks and identification of hub genes

Common DEGs are inserted into an online database called Search Tool for the Retrieval of Interacting Genes (STRING) (https://string-db.org/) to generate Protein-Protein Interaction (PPI) network. Those with a high level of confidence were regarded as valid interactions, and we set a convincing confidence score as 0.25 (Cao et al., 2021b). The obtained PPI network was then analyzed by Cytoscape 3.8.2 for a better visualization. The app Molecular Complex Detection (MCODE) on Cytoscape was applied to conduct the gene network clustering analysis to refine key modules, with a $p <$ 0.05. The app Cytohubba on Cytoscape was used to compute the degrees of nodes in PPI work, proteins with high degree might have key physiological regulatory functions, so the ones having the most interactions were considered as hub genes (Cao et al., 2021a). Additionally, we used a web-based tool GeneMANIA (http://genemania.org) to further visualize the interactions and roles of hub genes.

Analysis of TF-gene interactions and TF-miRNA co-regulation

TF-gene interactions with the identified hub genes point out the outcome of TF on functional pathways and expression levels of the genes (Ye et al., 2019). NetworkAnalyst (https://www.networkanalyst.ca/), a comprehensive online platform for analyzing gene expression, was used to identify TF-gene interaction and TF-miRNA co-regulation of identified hub genes. Then the two networks were mapped on Cytoscaope 3.8.2 for beautifying.

Immune infiltration analysis

With the help of Cell-type Identification By Estimating Relative Subsets Of RNA Transcripts (CIBERSORT) (https://cibersort.stanford.edu/), a web-based stool able to describe cell composition of complex tissues *via* their gene expression levels, we deconvoluted and compared the cellular composition of the two groups in our analysis. Then the box plots, bar charts, heat maps and scatter diagrams revealing the association of input datasets and immune pathways and cells were all drawn through "ggpubr" packages of RStudio software (Kassambara, 2020).

Identification of RIF subgroups

Consensus matrix of RIF subgroups identification in GSE111974 was obtained through "ConsensusClusterPlus" package of RStudio software, aiming to figure out if the five hub gene can distinguish the different subtypes of RIF(Wilkerson and Hayes, 2010). Optimal number of clusters was calculated *via* k-medoids clustering, indicating that k = 2. In order to verify the sample clustering condition of the 2 clusters we discerned, we conducted a diminished reduction analysis *via* "tsne" package of RStudio software (Donaldson, 2022). After clustering, we tried to find out the functional differences of the clusters. We used "ggpubr" package in RStudio software for visualizing the expression of hub genes in the 2 clusters, and subsequently utilized BioCarta database and "GSVA" package of RStudio software to identify the concerned pathways in the 2 clusters (Hänzelmann et al., 2013; Kassambara, 2020).

ROC curve analysis

For the purpose of identifying the role of the five hub genes in prediction of RIF, we conducted the characteristic ROC curves of the diagnostic model in the GSE111974, *via* "pROC" and "ggplot2" package of RStudio software (Robin et al., 2011; Wickham, 2016). To validate the results, we repeated the process in GSE26787, another dataset including RIF and control group.

Identification of potential drugs

Drug molecule identification is a pivotal component of the present study. We input five hub genes into the Drug Signatures database (DSigDB) on Enrichr (https://amp.pharm.mssm.edu/Enrichr/), thus obtained the candidate drugs interacting with hub genes, which may contribute to the treatment of RIF.

Results

Identification of DEGs using integrated bioinformatic analysis

The particulars of selected three datasets were presented in Table 1, and we finally chose GSE103465 and GSE111974 for

TABLE 1 Characteristics of the three GEO datasets selected.

GSE	GPL	Experiment type	Citation	Samples	Character	
103,465	16,043	Expression profiling by array	Guo et al. (2018)	Group: RIF (n = 3), fertile controls (n = 3). Sampling time: WOI (LH+7)	Women recruited in the RIF group had a history of implantation failure from at least three consecutive IVF attempts (including a total of ≥4 good-quality embryos)	The two datasets both contain the expression profile of endometrial tissue from RIF after IVF and control patients during WOI. Therefore, we chose them for analysis
111,974	17,077	Expression profiling by array	Bastu et al. (2019)	Group: RIF (n = 24), fertile controls (n = 24). Sampling time: WOI (LH+7–10)	RIF was determined as failure of pregnancy in ≥3 consecutive IVF cycles with ≥1 transfer(s) of good quality embryo in each cycle	
26,787	570	Expression profiling by array	Lédée et al. (2011)	Group: RIF (n = 5), fertile controls (n = 5), recurrent miscarriage (n = 5). Sampling time: 7–9 days after ovulation	It is originally aimed to identify the difference of endometrium from RIF and recurrent miscarriages, and did not provide the specific list of individuals accepting IVF or ICSI. Considering above mentioned reasons, we used it for the verification of ROC curves in our context	

analysis. As GSE26787 is originally aimed to identify the difference of endometrium from RIF and recurrent miscarriages, and didn't provide the specific list of individuals accepting IVF or ICSI, we used it for validation instead of analysis. 1,406 DEGs were obtained including 373 upregulated and 1,033 downregulated genes in GSE103465, while in GSE111974, 553 DEGs were collected, among which 326 genes were elevated and 227 were suppressed. The box plots shown in Figure 2 revealed the satisfying standardization of the samples. The expression of the top 20 DEGs for both two datasets were visualized on heatmaps (Figure 2). The volcano plots in Figure 2 highlighted the DEGs of the two datasets.

GO terms and KEGG/BioCarta pathway enrichment of common DEGs

Subsequently, venn diagrams were presented in Figure 3 to illustrate the overlap of DEGs from the two datasets. As presented in Figure 3, we finally identified 26 common DEGs containing 12 upregulated and 14 downregulated genes. The details of the common DEGs were displayed in Table 2. We visualized the GO terms and KEGG/BioCarta pathways of common DEGs in Figure 3, for the further understanding of biological functions. The biological processes analysis suggested that common DEGs mainly participated in vascular associated smooth muscle contraction and vasoconstriction (Figure 3C). Molecular function subsection indicated that common DEGs were associated with potassium channel regulator activity (Figure 3D). Predominate cellular components consisting of products by common DEGs were filopodium and caveola (Figure 3E). According to KEGG pathway database in Figure 3F, common DEGs mainly took a part in TNF and nuclear factor kappa B (NF-κB) signaling pathway. As pointed out by BioCarta pathway enrichment in Figure 3G, the inflammation associated pathways including interleukin (IL)-4, inerferon (IFN)-γ, IL-2 receptor β chain (IL2RB), IL-2, tumor growth factor (TGF) -β, tumor necrosis factor receptor (TNFR)-1, TNFR-2, and IL-12 pathways were dramatically downregulated in RIF group. Although the levels of IL-6 and IL-10 pathways didn't have a significant difference in two groups, we could still find a decreasing trend in RIF group. The specific data of pathway enrichment was displayed in Supplementary Table S1.

PPI network and hub genes analysis

The PPI networks of common DEGs built by STRING database and Cytoscape software, which contained 15 nodes and 32 edges, as picturized in Figure 4A. In Figure 4B, we displayed the key PPI network *via* network gene clustering analysis. It was exhibited in Figure 4C that hub genes we identified were *Prostaglandin-endoperoxide synthase (PTGS) 2*, *Vascular cell adhesion molecule 1 (VCAM1)*, *Endothelin receptor type B (EDNRB)*, *Actin alpha 2 (ACTA2)*, and *Leukaemia inhibitory factor (LIF)*. The networks of hub genes and their relative genes from GeneMANIA (Figure 4D) indicated that those hub genes had a strong relationship with *Leukemia inhibitory factor receptor (LIFR), endothelin (EDN) 3, integrin subunit alpha (ITGA) 9, EDN2, prostacyclin synthase (PTGIS), solute carrier family nine isoform 3 (SLC9A3), Thromboxane synthase (TBXAS1), prostaglandin D2 synthase (PTGDS), PTGS1, oncostatin M (OSM), GC, ITGA4, interleukin six cytokine family signal transducer (IL6ST), EDN1, myosin light chain 12A (MYL12A), Serotonin receptor 1B (HTR1B), myocardin (MYOCD), integrin alpha D (ITGAD), Myosin heavy chain 11 (MYH11)* and *mesoporous silica nanoparticle (MSN)*. The specific interaction patterns could be obtained from Supplementary Table S2.

TF-gene interaction and TF-miRNA coregulatory network

TF-gene interaction and TF-miRNA coregulatory network of common DEGs were identified with the aid of NetworkAnalyst, and were processed *via* Cytoscape. As shown in Figure 4E, the TF-gene interaction network consisted of 25 nodes and 24 edges. LIF was regulated by 19 TF-genes, while ACTA2 was regulated by four TF-genes, and the two hub genes shared a TF-gene called *AT-rich interaction domain (ARID4B)* (Figure 4E; Supplementary Table S3). TF-miRNA coregulatory network, as displayed in Figure 4F, contained 131 nodes and 151 edges, with 74 miRNAs and 52 TF-

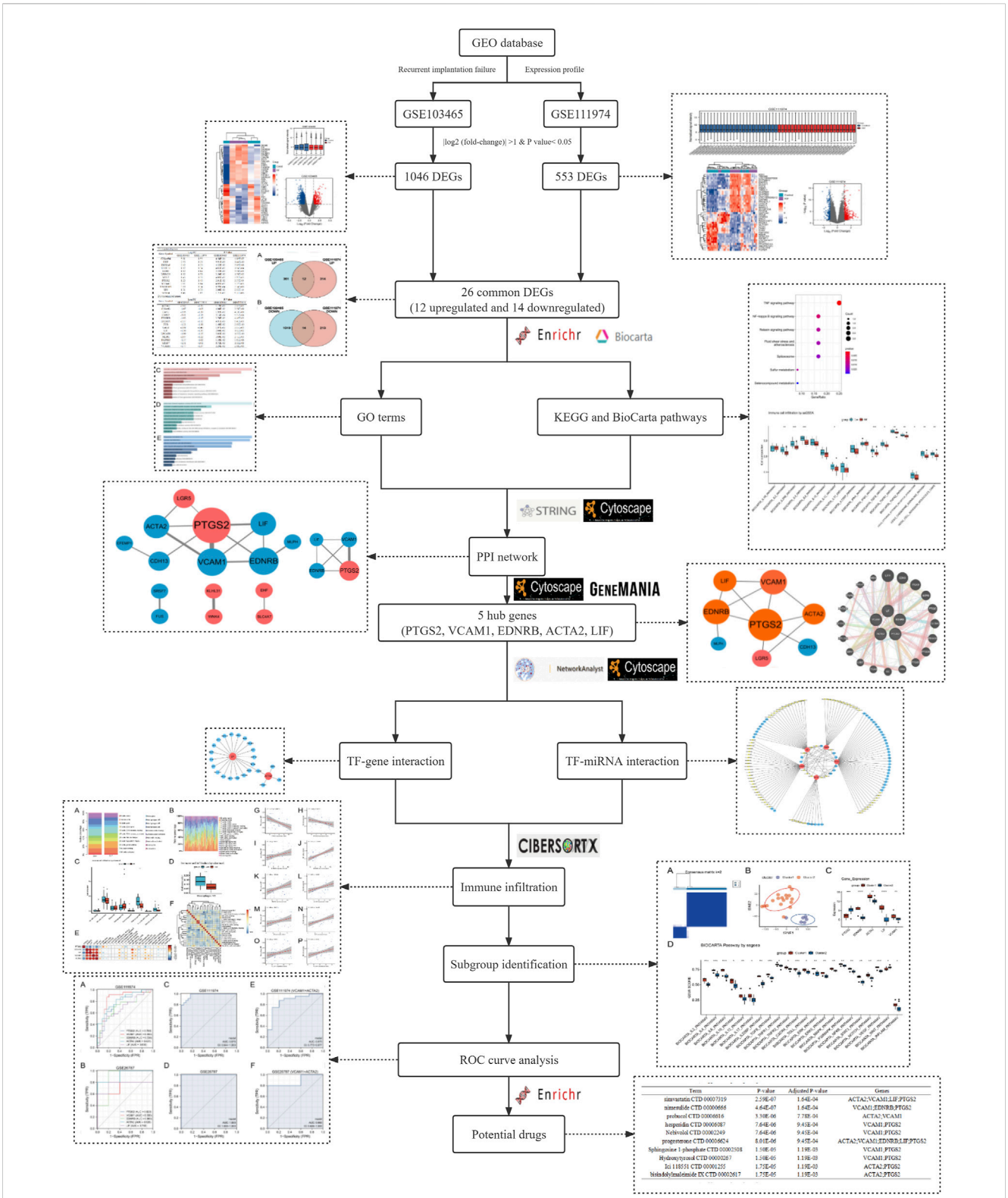


FIGURE 1
Methodical workflow for the current investigation.

genes being collected. Among the most interacted TF-genes in TF-miRNA coregulatory network, *NF-κB1* had the highest degree value of 4, *CCCTC-binding factor (CTCF)*, *E26 transformation specific-1 (ETS1)*, *RelA* and *NF-κB2* had eminent degree values of 3, *specificity protein 1 (SP1)*, *Jun* and *CCAAT/enhancer-binding protein beta (CEBPB)* have relatively higher degree values of 2 (Figure 4F;

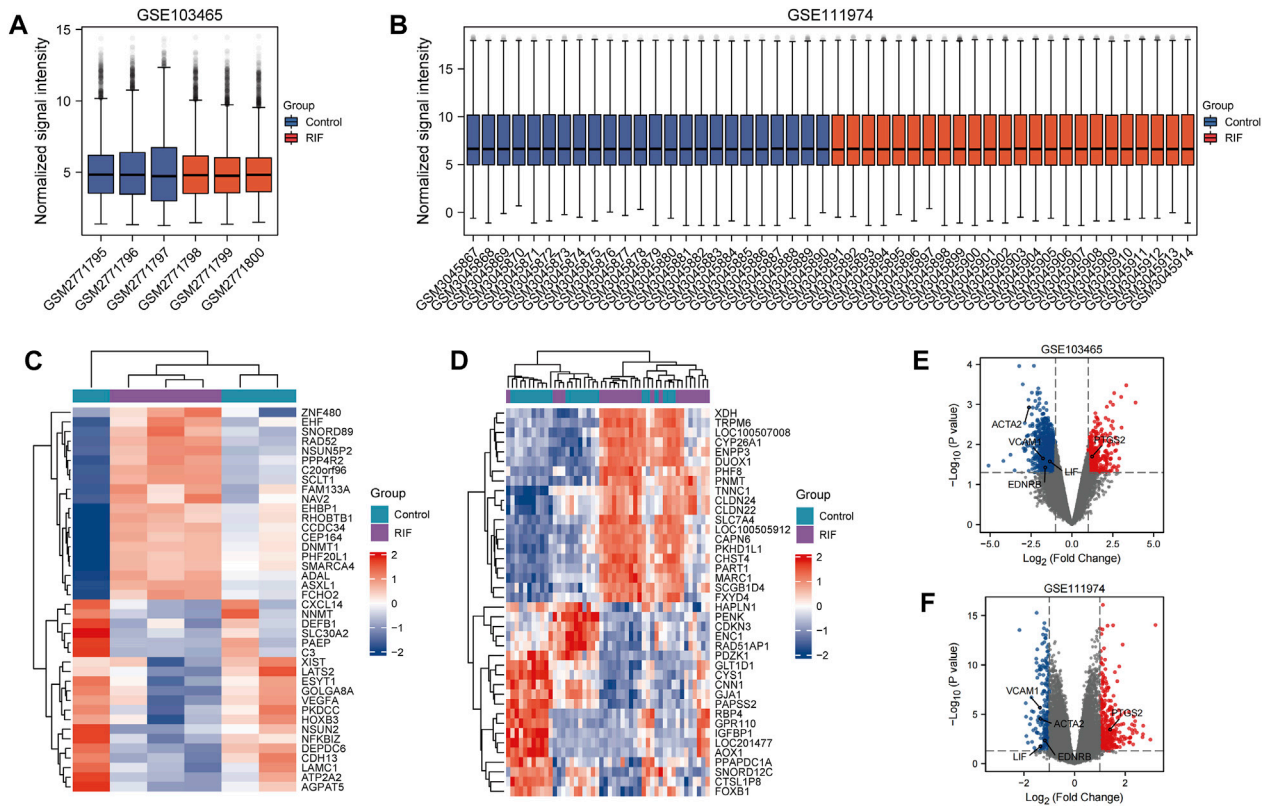


FIGURE 2
Box plots, Heatmaps and Volcano plots of DEGs between RIF and control group. **(A, B)** Box plots of GSE103465 **(A)** and GSE111974 **(B)**. Red refers to RIF group, while blue represents control group. **(C, D)** Heatmaps of 10 DEGs who have the highest log FC and 10 DEGs with the lowest log FC from GSE103465 **(C)** and GSE11194 **(D)**. Red refers to the elevated genes, and blue directs to the downregulated genes. **(E, F)** Volcano plots of DEGs in GSE103465 **(E)** and GSE111974 **(F)**. Red refers to the increased DEGs, and blue points to the reduced DEGs.

Supplementary Table S4). Among the most interacted TF-miRNAs in TF-miRNA coregulatory network, hsa-miR-181b and hsa-miR-181d had relatively higher degrees values of 3, and the degree values of hsa-miR-181a, hsa-miR-340, hsa-miR-590-3p, hsa-miR-26a, hsa-miR-29a, hsa-miR-29b and hsa-miR-29c were 2 (Figure 4F; Supplementary Table S4). Considering the role of aforementioned TF-genes and TF-miRNAs, especially *NF-κB*, *Jun*, *CEBPB*, hsa-miR-181 and miR-miR-29 in inflammation, we believed that the inflammation could affect RIF also in a transcription level.

Immune infiltration analysis

As presented in Figure 5, the percentage of macrophage M2, γδT cell and dendritic cells activated in RIF group were less than that in control group. The percentage of macrophage M2 in two groups was individually shown in Figure 5D. Although the levels of NK cells activated didn't have significant differences in two groups, we could still find a decreasing trend of NK cells activated in RIF group, and larger scale of samples are required (Figure 5C). The result might be significant if there was a larger sample size. Subsequently, we executed the relationships between hub genes and hub genes, hub genes and infiltrated immune cells, infiltrated immune cells and immune cells (Figure 5). The results showed that hub genes have a tight relationship with the amount of macrophage M2 and NK cell activated. Therefore, we then excavated the association of hub genes expression and above mentioned two immune cells. As exhibited in Figure 5, the expression of PTGS2 had a negative correlation with macrophage M2 and NK cell activated ($p < 0.05$), while EDNRB, ACTA2 and LIF were positively correlated with macrophage M2 and NK cell activated ($p < 0.05$). VCAM1 had a positive relationship with NK cells activated ($p < 0.05$), although it did not have a significant relationship with macrophage M2 ($p = 0.094$), we could still observe a positive trend.

Identification of RIF subgroups

The RIF individuals in GSE111974 could be grouped into 2 clusters *via* hub genes, which was displayed in Figure 6A, with a relatively larger cluster two and smaller cluster 1. Then the sample clustering condition in Figure 6B showed that the 2 clusters could be divided to two separate sections in the quadrant, which meant that there were remarkable differences between the two clusters. As shown in Figure 6C, the expression of PTGS2 was significantly higher in cluster 2, while the expression levels of *EDNRB*, *ACTA2*, *LIF*, and *VCAM1* were markedly decreased in cluster 2. The pathways which exhibited anti-inflammatory effect, such as IL-2,

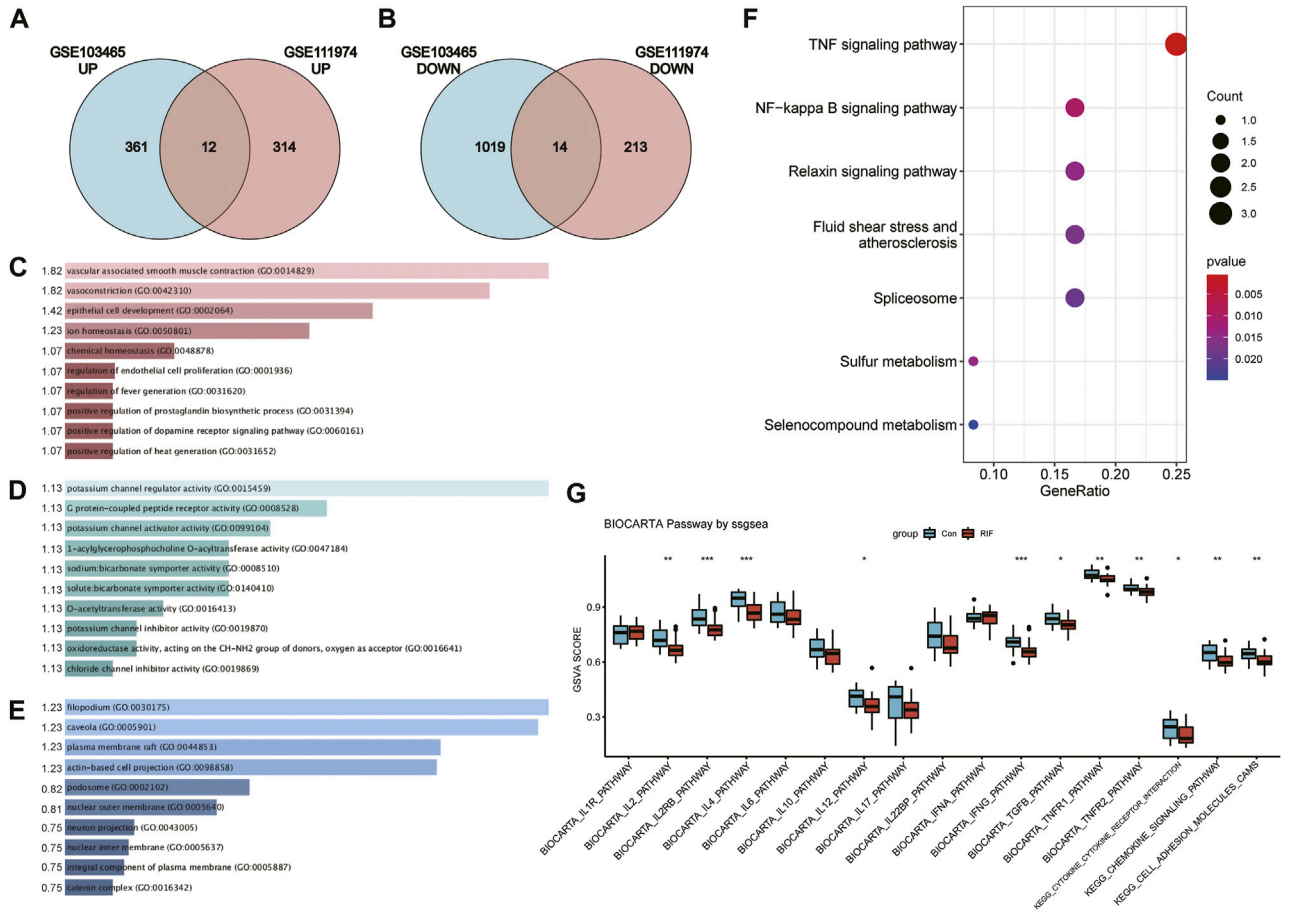


FIGURE 3
(A, B) Venn diagram of upregulated DEGs **(A)** and downregulated DEGs **(B)** from two datasets. **(C–E)** Bar graphs reflecting GO terms identification including Biological process **(C)**, molecular function **(D)** and cellular component **(E)** of common DEGs. The width represent the values of [−log (Adjusted *p*-value)], which are listed beside the bars. **(F)** Bubble chart visualizing KEGG pathway analysis of common DEGs. **(G)** Box plot of BioCarta and KEGG pathway analysis *via* ssGESA.

IL-4, IL-6, IL-10, TGF-β, TNFR2, and extracellular signal-regulated kinase 5 (ERK5) pathways were dramatically downregulated in cluster 2 (Figure 6D). Besides, Wnt pathway, mainly exerted pro-inflammatory effect, was significantly activated in cluster two compared with cluster 1 (Figure 6D). And other pathways like IL-12, IL-22 binding protein (IL-22 BP), Toll, extracellular signal-regulated kinase (ERK), IL-17, TNFR1, chemokine C-X-C motif ligand 4 (CXCR4), mitogen-activated protein kinases (MAPK), p38 MAPK, NF-κB, signal transducer and activator of transcription 3 (STAT3), Notch, vascular endothelial growth factor (VEGF) and Inflam pathways also had remarkable downregulations in cluster 2 (Figure 6D). Therefore, it was clear that the five hub genes could separate RIF patients into 2 clusters, with one presented a typical inflammatory environment in uterus (cluster 2) and another had atypical inflammatory responses (cluster 1) (Figure 6). Inflammatory alteration may contribute to RIF, but this cant explain all the patients, especially individuals in cluster 1, whose pathogenesis has been rarely discussed up to now. In future work, RIF individuals can be grouped by the five hub genes, and we can investigate specific etiology and accurate treatment methods for each group.

ROC curve analysis

In GSE111974, we found that the five hub genes showed important values in the diagnosis of RIF independently, especially VCAM1 and ACTA2, which had an AUC of 0.852 and 0.821, respectively (Figure 7A). The results were verified in GSE26787, with a high AUC (AUC>0.7) in each hub genes (Figure 7B). Combined the five hub genes, the AUC reached 0.976 when combined the five hub genes in GSE111974 (Figure 7C), while in GSE26787, the AUC was 1.0 (Figure 7D). Combined VCAM1 with ACTA2, the AUC achieved 0.875 in GSE11974 (Figure 7E) and 0.88 in GSE26787 (Figure 7F). To sum up, the hub genes we identified the five hub genes had good diagnostic values in RIF.

Identification of potential drugs

From DSigDB database, we selected 10 drugs who had the minimal adjust *p*-value, they were: Simvastatin CTD 00007319, nimesulide CTD 00000666, probucol CTD 00006616, hesperidin CTD 00006087, Nebivolol CTD 00002249, progesterone CTD

TABLE 2 Log FC and *p*-value of common DEGs.

Upregulated genes				
Gene.Symbol	Log FC		*p*.Value	
	GSE103465	GSE111974	GSE103465	GSE111974
C20orf96	3.01	1.35	6.50E-04	4.16E-07
EHF	2.53	1.76	9.91E-03	1.44E-10
HOXA6	2.42	1.20	4.52E-02	5.96E-08
KLHL31	1.07	1.34	4.63E-02	1.84E-04
LGR5	1.62	1.14	2.22E-02	3.51E-05
LRRC26	1.02	1.58	3.50E-02	1.51E-02
MTL5	1.43	1.08	4.85E-02	2.50E-03
PTGS2	1.23	1.40	2.01E-02	3.57E-04
SLC4A7	2.11	1.14	4.85E-02	3.35E-03
SNORD89	3.89	1.04	8.99E-04	2.74E-06
SP9	1.31	1.00	2.96E-02	2.02E-14
WNK4	1.46	2.15	1.21E-02	5.76E-06
Downregulated genes				
Gene.Symbol	Log FC		*p*.Value	
	GSE103465	GSE111974	GSE103465	GSE111974
ACTA2	−2.62	−1.35	1.19E-03	2.98E-05
C20orf3	−1.67	−1.05	4.96E-02	3.51E-09
CAV2	−1.95	−1.20	1.92E-02	4.21E-08
CDH13	−3.01	−1.13	3.14E-04	5.57E-04
EDNRB	−1.63	−1.17	3.76E-02	4.46E-03
EFEMP1	−2.17	−1.12	4.91E-03	5.40E-03
FUS	−1.23	−1.38	1.48E-02	2.69E-10
GAS1	−2.08	−1.46	1.07E-03	3.87E-05
LIF	−1.36	−1.35	2.64E-02	1.82E-02
LPCAT1	−1.99	−1.25	5.85E-03	1.07E-10
MLPH	−2.07	−1.22	3.68E-02	2.72E-03
PAPSS2	−1.57	−1.62	1.58E-02	5.62E-06
SRSF7	−1.51	−1.01	9.72E-03	1.06E-08
VCAM1	−1.77	−1.37	2.24E-02	2.11E-06

00006624, Sphingosine 1-phosphate CTD 00002508, Hydroxytyrosol CTD 00000267, Ici 118,551 CTD 00001255, bisindolylmaleimide IX CTD 00002617 (Table 3). Among the 10 candidate drugs, progesterone CTD 00006624 could interact with all the five hub genes, and simvastatin CTD 00007319 had an interaction with the five hub genes except EDNRB.

Discussion

In the current study, we tried to find biological changes contributing to the pathogenesis of RIF *via* gene profiling. Benefited from the combination of two microarray datasets GSE103465 and GSE111974, our results are more effective and reliable. We found that the disturbed inflammation regulation plays a key role in the pathogenesis of RIF, and abnormal uterine muscle contraction and vascularity also contribute to RIF. Although the five hub genes we identified have been discussed in previous studies, the molecular mechanism isn't fully understood. We first revealed the interaction existed among the five hub genes, which associated with the disturbed inflammation regulation, uterine muscle contraction and vascularity in RIF. Of importance, the subgroup identification revealed that a small number of patients have atypical phenotypes (cluster 1), which may be the reason for the poor prognosis of RIF. It hints us that RIF individuals can be grouped by the five hub genes, and we can investigate specific etiology and accurate treatment methods for each group in future work. Additionally, the drug prediction revealed the potential drug molecules, which sheds new light on the treatment of RIF.

Inflammation is essential in various pathophysiological processes and diseases including cancer, allergic diseases,

10.3389/fgene.2023.1108805

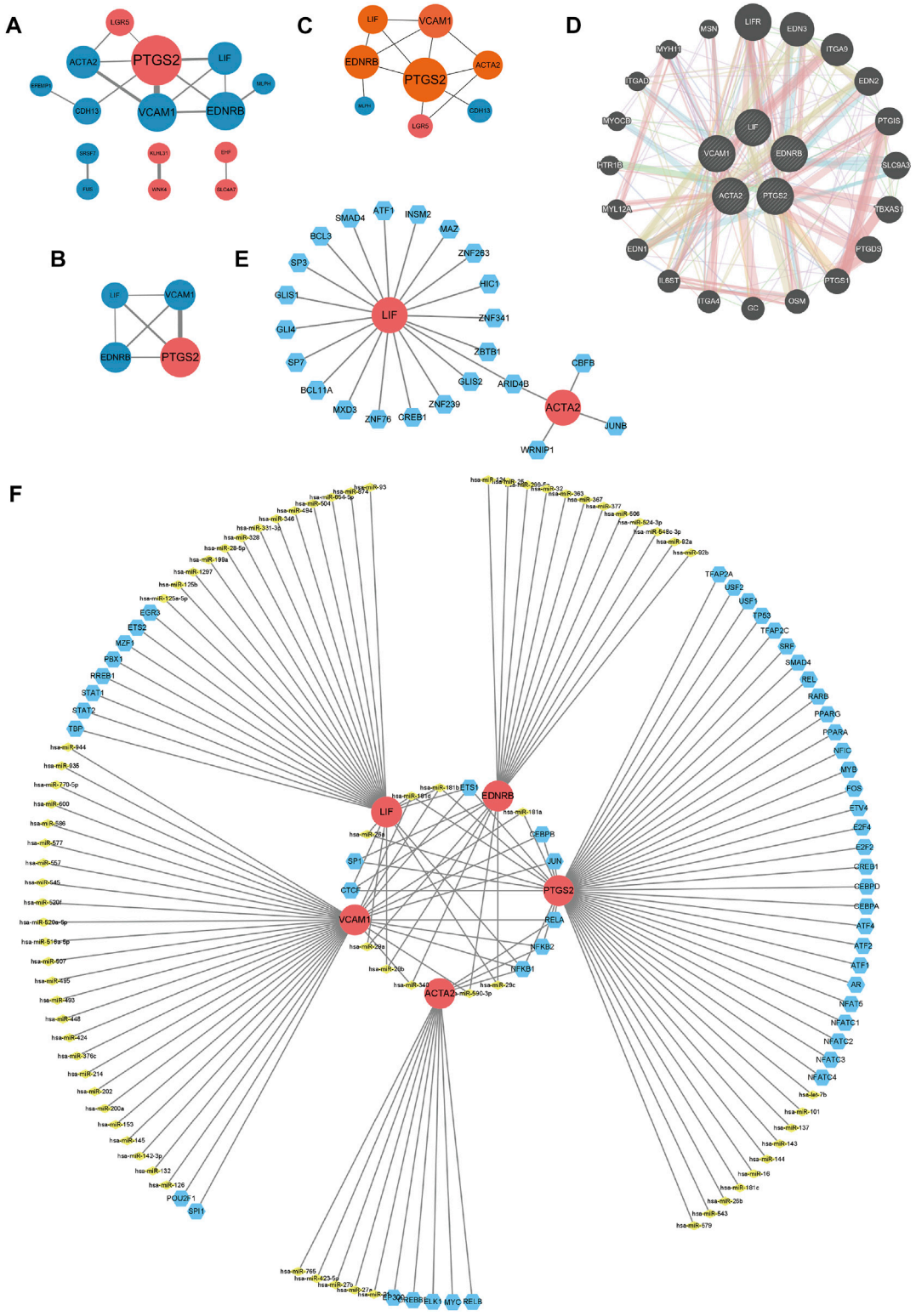


FIGURE 4
(A–D) PPI networks. **(A)** PPI network of 26 common DEGs. **(B)** Key modules of common DEGs. Larger nodes means higher degree scores. Thicker lines represent higher combined scores. Red nodes refer to upregulated genes, while blue ones are downregulated genes. **(C)** Display of hub genes, which are in orange nodes. **(D)** Interactions of hub genes and related genes. Pink lines represent physical interactions, purple lines refer to co-expression, orange lines direct to predicted interactions, blue lines mean co-localization, and green lines point to genetic interaction, cyan lines illustrate pathways, and yellow lines are shared protein domains. **(E)** Network for TF-gene interaction with hub genes. The highlighted red color node represents the hub genes and other blue nodes represent TF-genes. The TF-miRNA coregulatory network. The nodes red color are the hub genes, a yellow node represents miRNA and other blue nodes indicate TF-genes.

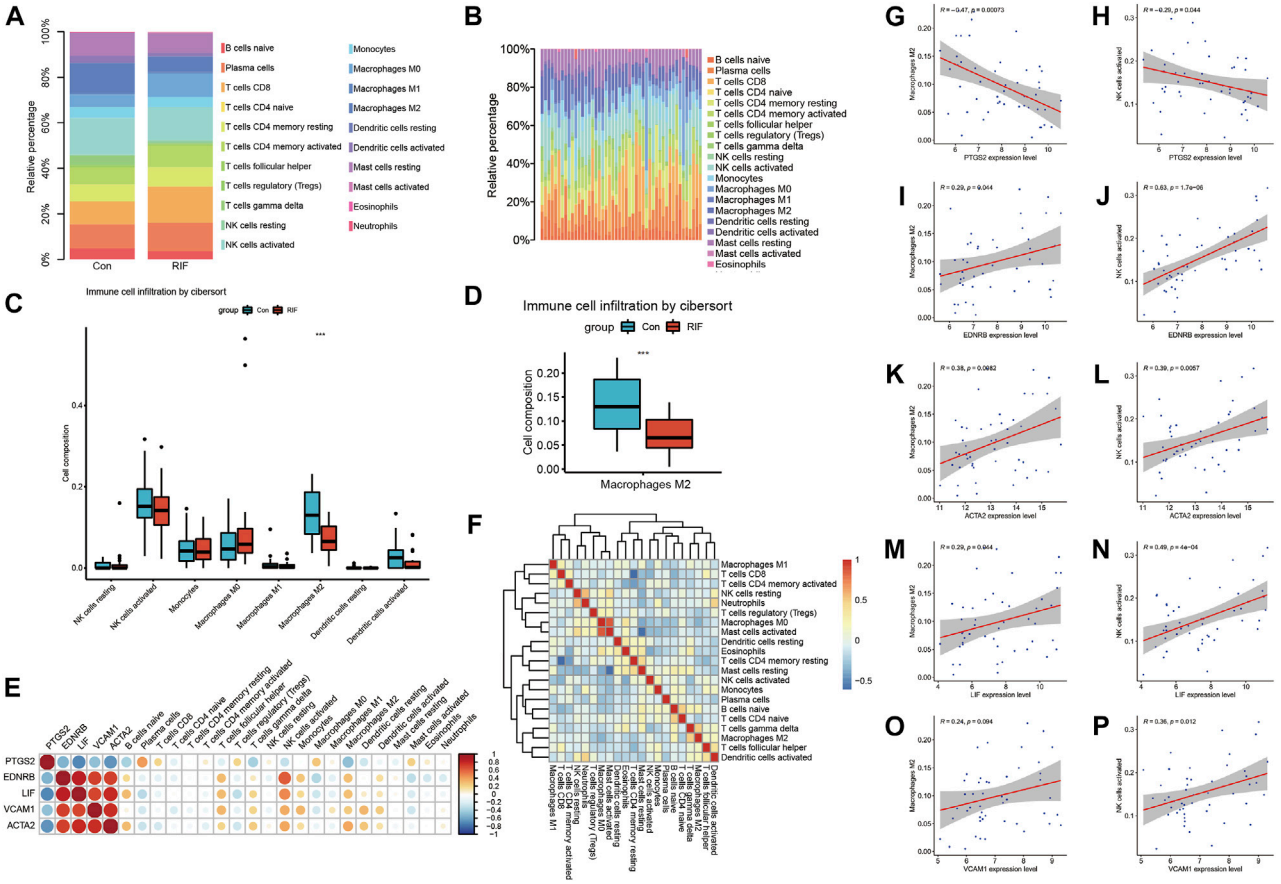


FIGURE 5
The immune infiltration analysis. **(A, B)** Immune infiltration analysis for 21 types of immune cell composition of endometrial tissue from RIF versues control group. **(C)** Box plot reveals the composition of eight types of immune cells in RIF and control group. **(D)** Box plot refers to the amount of M2 macrophages in RIF and control group. **(E)** Heat map exhibit the association of hub genes between hub genes and hub genes between immune cells. **(F)** Heat map reveals the relation of immune cells enriched. **(G,I,K,M,O)** Correlation scatter plots between expression of PTGS2/EDNRB/ACTA2/LIF/VCAM1 and Macrophages M2. **(H,J,L,N,P)** Correlation scatter plots between expression of PTGS2/EDNRB/ACTA2/LIF/VCAM1 and NK cells activated.

congenital diseases and so on (Gao et al., 2021). The role of inflammation in pregnancy has been argued for decades, of which importance is beyond doubt (Mor et al., 2011). Although a specific local pro-inflammatory environment is necessary for embryo implantation, excessive inflammatory response can also be harmful (Mor et al., 2011; Mekinian et al., 2016). Anti-inflammatory cytokines and cells perform multiple functions during normal pregnancies, such as promoting placental formation and angiogenesis, and modulating trophoblast differentiation and invasion (Chatterjee et al., 2014). Combining with results of current study, we believe that the ruined balance of pro-inflammatory and anti-inflammatory factors can ultimately lead to implantation failure.

KEGG pathway enrichment of common DEGs indicates that TNF and NF-κB signaling pathways play a major part in RIF. Canonical NF-κB pathway activation responds to a diversity of external stimuli involved in inflammatory and immune response, *via* inducing the expression of pro-inflammatory cytokines including TNF-α (Yu et al., 2020). Disturbed expression of NF-κB has been reported in women suffering from infertility, and gene polymorphism of NF-κB is noted to be related to RIF (Luo et al., 2016). Detectable TNF-α and NF-κB from feto-maternal surface are vital characters of successful implantation (Mor et al., 2011; Ersahin et al., 2016), but the overexpression of TNF-α and NF-κB also has adverse effect on implantation and lead to RIF (Ersahin et al., 2016; Mekinian et al., 2016). Moreover, Duan and colleagues found that TNF was the cytokine having the strongest correlation with all hub genes in their research on RIF (Duan et al., 2022). The aforementioned evidences suggest that NF-κB and TNF signaling pathways are essential in implantation, and should be cautiously controlled in an appropriate range. KEGG pathway calculated by ssGESA suggests that cytokine-cytokine receptor interaction pathway has a significant downregulation in RIF group. KEGG pathway calculated by ssGESA suggests that IL-2, IL-2 receptor beta (IL-2RB), IL-4, IL-12, IFN-γ, TGF-β, TNFR1, TNFR2, cytokine-cytokine receptor interaction, chemokine signaling pathway and cell adhesion molecules cams had significant downregulation in RIF group. Among the aforementioned pathways, IL-2RB, IL-4 and IFN-γ pathways showed the most remarkable decline ($p < 0.001$). IL-2 RB can induces growth potential for endometrial glandular epithelial cells, and its hypermethylation and downregulation has been found in ovarian endometriosis (Kusakabe et al., 2009; Zhang et al., 2022).

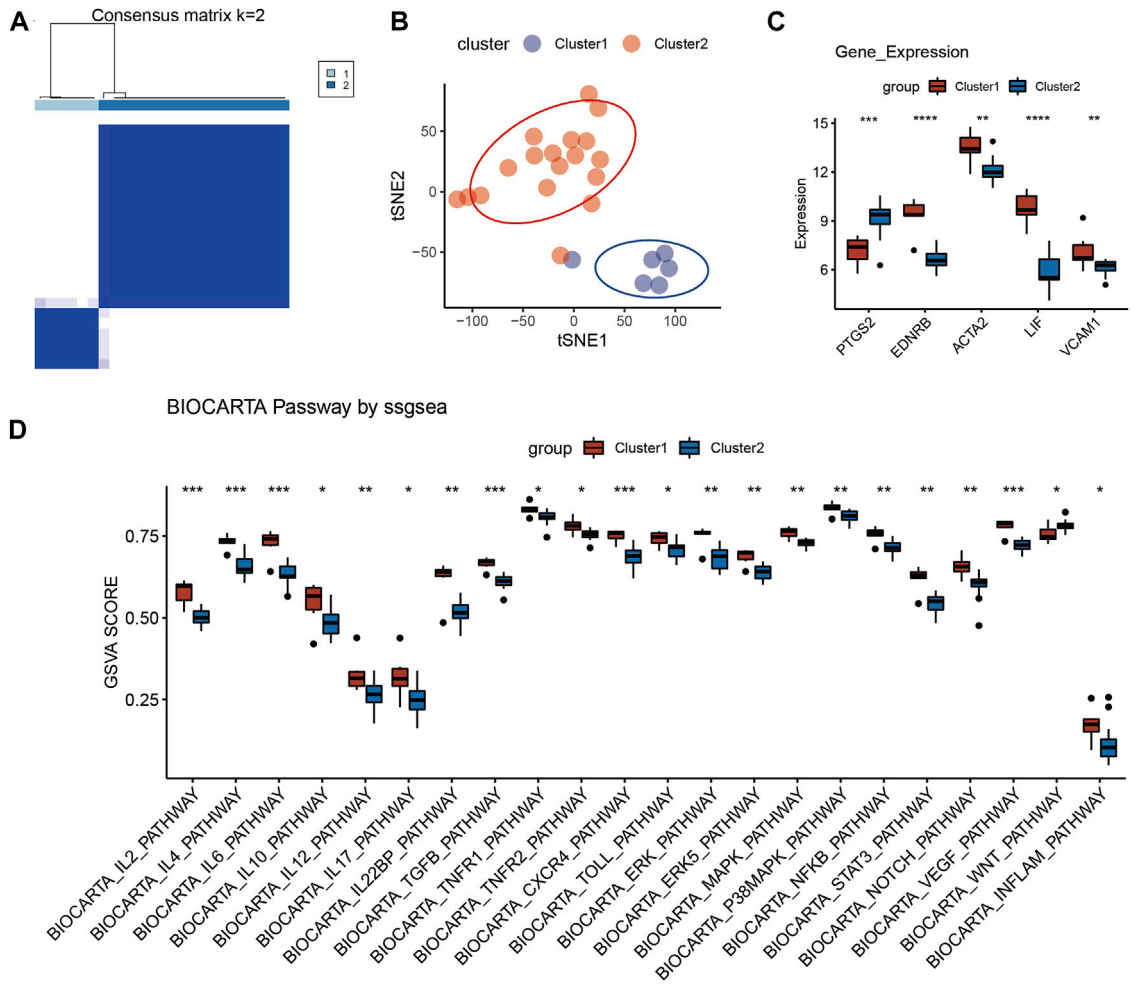


FIGURE 6
Identification of RIF subgroup in GSE111974. **(A)** Reordered consensus matrix on RIF compendium, by applying k-medoids with k = 2. **(B)** t-distributed Stochastic Neighbor Embedding (t-SNE) reduces the dimensions of a multivariate dataset. **(C)** Box plot reveals the expression of five hub genes in the two clusters. **(D)** Box plot of BioCarta pathway analysis in the two clusters *via* ssGESA.

IL-4 is always considered as a cytokine involved in anti-inflammatory effect, and is also a vital mediator of fetal tolerance in successful implantation and pregnancy (Feghali and Wright, 1997; Liang et al., 2015; Mekinian et al., 2016). IFN-γ, as a pro-inflammatory cytokine, has been reported to participate in uterine vascular modification and successful implantation (Feghali and Wright, 1997). The high ratios of IFN-γ/IL-4, IFN-γ/IL-10 and IFN-γ/TGF-β have been observed in the RIF and associated with adverse outcome of implantation (Liang et al., 2015). As pointed out by BioCarta pathway enrichment, the cytokine associated pathways, such as IL-2, IL-4, IL-6, CXCR4, and VEGF pathways, are dramatically downregulated ($p < 0.001$) in cluster 2, which contains the majority of RIF patients. Therefore, appropriate inflammatory activation is the crux of successful implantation.

Cytokines participating in implantation can be secreted by the endometrial cells and immune cells recruited to the position of implantation, 65%–70% of these cells are uterine-specific NK cells, and 10%–20% are macrophages (Mor et al., 2011). Depletion of these immune cells has deleterious effects on implantation, deciduation and placental development (Mor et al., 2011). In current immune infiltration analysis of cell composition, M2 macrophages and NK cells are memorably lower in RIF group than that in control group. Uterine NK cells have long been acknowledged to be essential in deciduation for its role in endometrial vascularity (Mor et al., 2011). M2 macrophages can be polarized by anti-inflammatory cytokines such as IL-4, IL-10, and TGF-β, conversely, they can produce large quantities of IL-10 and TGF-β to suppress the inflammation, for the purpose of tissue repair, remodeling and vasculogenesis (Shapouri-Moghaddam et al., 2018). The depletion of M2 macrophages in mice has been demonstrated to be a cause of implantation failure (Ono et al., 2020).

Consistent with previous studies, we find that the imbalance of pro-inflammatory and protective factors leads to a disordered immune environment in uterus, results in abnormal vascularity and muscle contraction, and finally bring about RIF. But our analysis revealed a more comprehensive inflammatory pathway spectrum in RIF. The function of identified hub genes (*PTGS2*, *VCAM1*, *EDNRB*,

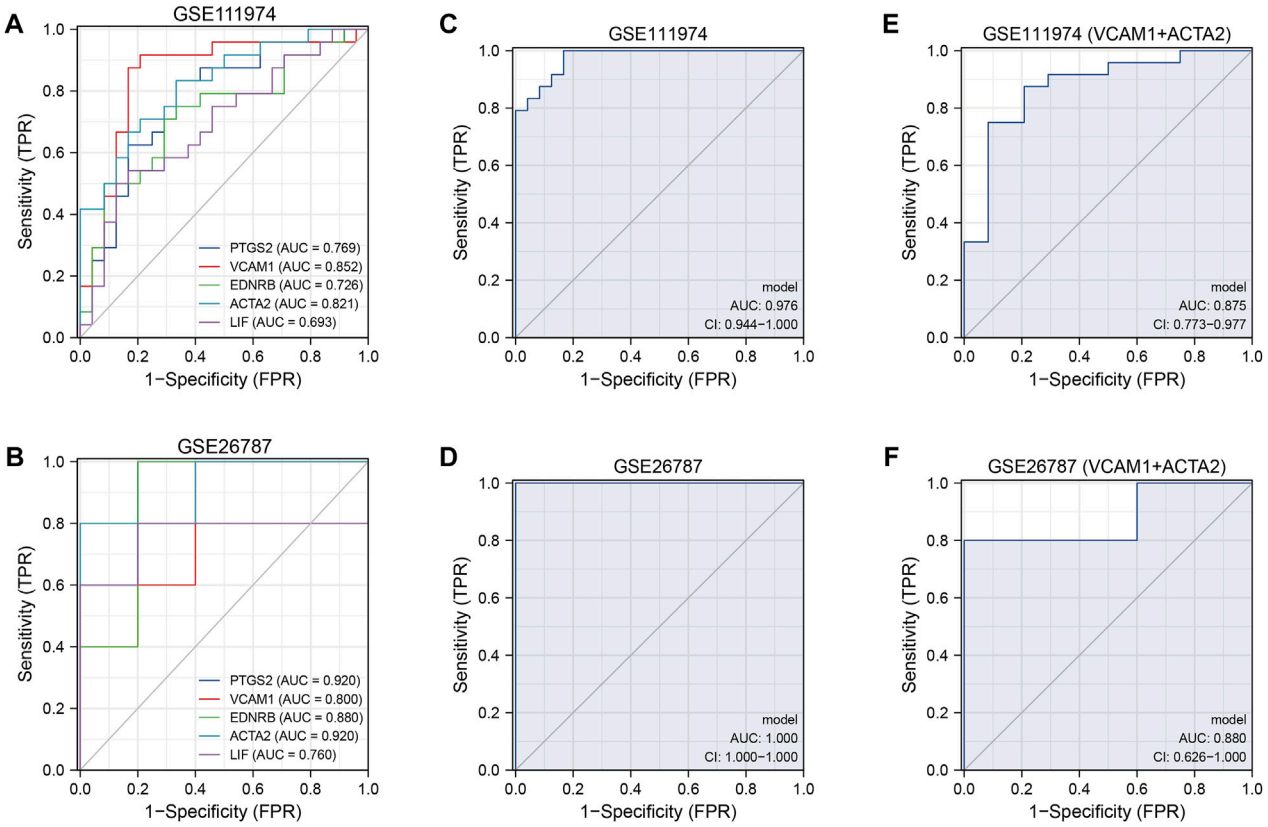


FIGURE 7
The ROC curves of the diagnostic model. **(A, B)** The diagnostic ROC curves of independent five hub genes in GSE111974 **(A)** and GSE26787 **(B)**. **(C, D)** The diagnostic ROC curves of combined five hub genes in GSE111974 **(C)** and GSE26787 **(D)**. **(E, F)** The diagnostic ROC curves of combined VCAM1 and ACTA2 in GSE111974 **(E)** and GSE26787 **(F)**.

TABLE 3 Suggested top drug compounds for RIF.

Term	*p*.Value	Adjusted p.Value	Genes
simvastatin CTD 00007319	2.59E-07	1.64E-04	ACTA2; VCAM1; LIF; PTGS2
nimesulide CTD 00000666	4.64E-07	1.64E-04	VCAM1; EDNRB; PTGS2
probucol CTD 00006616	3.30E-06	7.78E-04	ACTA2; VCAM1
hesperidin CTD 00006087	7.64E-06	9.45E-04	VCAM1; PTGS2
Nebivolol CTD 00002249	7.64E-06	9.45E-04	VCAM1; PTGS2
progesterone CTD 00006624	8.01E-06	9.45E-04	ACTA2; VCAM1; EDNRB; LIF; PTGS2
Sphingosine 1-phosphate CTD 00002508	1.50E-05	1.19E-03	VCAM1; PTGS2
Hydroxytyrosol CTD 00000267	1.50E-05	1.19E-03	VCAM1; PTGS2
Ici 118,551 CTD 00001255	1.75E-05	1.19E-03	ACTA2; PTGS2
bisindolylmaleimide IX CTD 00002617	1.75E-05	1.19E-03	ACTA2; PTGS2

ACTA2, and *LIF*) also highlights the importance of inflammation, uterine muscle contraction and vascularity in RIF.

PTGS2 is synthesized at very low levels under normal conditions, but can be stimulated by specific events and is responsible for the prostanoid biosynthesis under inflammation (Vane et al., 1998). Most of the stimuli that induce PTGS2 are those associated with inflammation, such as TNF-α, while IL-4 presented an inhibitory impact on PTGS2 (Vane et al., 1998). For decades, PTGS2 has been widely considered to be an indispensable molecule in female reproductive process including

ovulation, fertilization, implantation and embryo development (Anamthathmakula and Winuthayanon, 2021). However, our analysis showed that there was a significant increase of PTGS2 in RIF, suggesting that the over expression of PTGS2 may also be unfavorable for implantation. Rodent and simian models have revealed that premature uterine contraction may be associated with the activation of prostaglandin signaling, since PTGS2 inhibitors can dampen cytokine induced uterine contractility (Sadowsky et al., 2000; Mackler et al., 2003; Cirillo et al., 2007; Orsi and Tribe, 2008). Additionally, high-frequency uterine contraction at the time of embryo transfer had an adverse impact on implantation rates in IVF-ET (Fanchin et al., 1998). Our analysis provided a new sight that the overexpression of PTGS2 may also impair the endometrial receptivity during WOI, *via* leading to activated inflammatory cascade and uterine muscle contraction. As a consequence, the level of PTGS2 in endometrium needs a more precise regulation.

VCAM1 is a downregulated hub gene in our analysis. *VCAM1* is a NF-κB target gene induced by TNF-TNFR1 signaling pathway and IL-4 (Kong et al., 2018). As an adhesion molecule, VCAM1 is associated with epithelial cells activation, neutrophil recruitment and aggravated creatine kinase (He et al., 2021; Mao et al., 2021). In reproductive system, VCAM1 is known to appear in endometrial side of decidual stromal cells (Bai et al., 2014). Significantly lower expression of VCAM1 in endometrium at the peri-implantation stage is associated with unexplained infertility and implantation failure after IVF(Konac et al., 2009). Uterine VCAM1 expression is essential for conceptus-uterine endometrium adhesion and early placental development, so mutations or deficiencies of the *VCAM1* may contribute to a series of human placental insufficiencies. (Gurtner et al., 1995; Bai et al., 2014). In Shang and colleague's research, *VCAM1* was also regarded as a key gene in miRNA-mRNA interaction network of RIF (Shang et al., 2022). Result of our analysis is tied well with the previous studies, so we believe that VCAM1 is a vital molecule in implantation and has the potential to be a biomarker in RIF.

EDNRB widely locates in vascular endothelium of many human tissues including placenta (Gram et al., 2017). The expression of EDNRB was increased in the perivascular and vascular cells of branching vessels during the late secretory phase (Keator et al., 2011). And EDNRB was constantly expressed during pregnancy including peri-implantation phase (Gram et al., 2017). What intrigues us is that during prepartum luteolysis, elevated expression of the EDN receptors in placenta strongly resembles the placental localization of PGs family members (e.g., PTGS2) in dogs (Kowalewski et al., 2010; Gram et al., 2014). In addition, the elevation of EDN1 during normal prepartum luteolysis and antigestagen-induced parturition/abortion is associated with increased PGs output in dogs (Kowalewski et al., 2010; Gram et al., 2014). In brief, EDNRB is strongly associated with all stages of pregnancy including implantation and involved in the signaling cascade of leukocyte recruitment and PGs synthesis, loss of EDNRB has the probability to lead to implantation failure.

Actin alpha 2 (ACTA2) is a smooth muscle actin predominately participated in vascular contractility and blood pressure homeostasis (Maglott et al., 2011). Expressed in uterine myocytes, ACTA2 is associated with uterine muscle contraction and uterine remodeling in pregnancy (Cooper and Brown, 2017). Women with ACTA2 mutations may be more likely to suffer from uterine muscle dysfunction and hemorrhage, according to a case report by Kylie and colleagues (Cooper and Brown, 2017). Our study revealed the potential and indispensable role of ACTA2 in implantation *via* impacting the function of uterine muscle.

As a member of IL-6 cytokine family, LIF has been reported to exhibit both pro-inflammatory and anti-inflammatory effects (Gadient and Patterson, 1999). Like PTGS2, LIF is also strongly elevated by TNF and downregulated by IL-4 (Gadient and Patterson, 1999). Conversely, LIF induces the production of pro-inflammatory cytokines (Gadient and Patterson, 1999). LIF is viewed as a potential predictor of fertility after IVF, considering that high expression levels of LIF in endometrium during the mid-luteal phase are relevant to a higher rate of pregnancy success in women underwent IVF (Serafini et al., 2008). The mechanisms of LIF in modulating implantation have been well discussed. LIF influences endometrial receptivity through inducing decidualization, elevating IL-6 and IL-15 levels in decidual cells, and recruiting leukocytes during WOI (Kimber, 2005; Shuya et al., 2011). LIF also has the ability to accelerate the transformation of endometrial macrophages into an anti-inflammatory phenotype *via* LIF-signal transducer and activator of transcription (STAT) pathways (Brinsden et al., 2009; Dallagi et al., 2015).

TF-genes are reactors for the gene expression regulation, through binding with targeted genes and miRNAs (Zhang et al., 2015). Among the regulators, ARID4B has a significant interaction, which has been demonstrated to be a regulator of male fertility (Wu et al., 2013). Regulatory biomolecules work as potential biomarkers in plenty of complex diseases. As shown in TF-genes and TF-miRNA coregulatory network, NF-κB1, NF-κB2, RelA, CEBPB, miR-181, and miR-29 are important in RIF. NF-κB, RelA (a subunit of NF-κB) and CEBPB are both key regulators in inflammation (Zahid et al., 2020; Zhao et al., 2020). MiRNAs can regulate various target genes in numerous biological processes and diseases, such as endometrial cancer (Ni et al., 2022). It has been reported that miR-181 family present a central role in vascular inflammation (Sun et al., 2014). MiR-181b inhibits expression of VCAM1, and serves as an inhibitor of downstream NF-κB signaling pathway (Sun et al., 2012; Sun et al., 2014). MiR-181d has the responsibility to regulate the acute stress response in thymocytes *via* targeting LIF (Sun et al., 2014). Combing mRNA microarray GSE111974 with miRNA microarray GSE71332, Ahmadi and colleagues constructed a circRNA-miRNA-mRNA network in RIF, and found that miR-29c might be a crucial miRNA, which is consistent with our analysis (Ahmadi et al., 2022). In addition, overexpression of miR-29 has been reported to impair endometrial receptivity by inhibiting the differentiation of endometrial stromal cells and regulating decidualization (Zhou et al., 2021).

We then used cluster analysis, ROC analysis and DSigDB database, in order to the identify the clinical diagnostic and therapeutic value of the five hub genes in RIF. Considering the results of RIF subgroup identification, the patients with RIF can be divided into two subgroups. Most patients (cluster2) have similar characteristics with RIF we previously believed, but there are still

some patients have atypical phenotypes (cluster 1), which may be one of the reasons for the poor therapeutic effect of RIF. Therefore, our future research should investigate specific etiology and accurate treatment methods for each group. The ROC curve in GSE111974 shows that the five genes do make sense in diagnosis of RIF, and the result was furtherly verified in RIF group with another GEO datasets GSE26787. According to DSigDB database, the current study highlights the progesterone CTD 00006624 and simvastatin CTD 00007319 as two drug molecules that most hub genes interacted with. Progesterone is a necessary hormone for pregnancy, and its supplementation can prevent recurrent miscarriage and reduce implantation failure in IVT cycles (Nardo and Sallam, 2006). Simvastatin, a lipid regulating agent, also shows impact on polycystic ovary syndrome (PCOS), endometriosis and uterine fibrosis (Banaszewska et al., 2011; Taylor et al., 2017; Ali et al., 2018). Simvastatin has been reported to reduce the expression of neopterin in endometriosis, which is a marker of inflammation and immune system activation (Taylor et al., 2017). Simvastatin can also inhibit C-reactive protein in PCOS and suppress fibroid proliferation and extracellular matrix production in uterine fibrosis (Banaszewska et al., 2011; Ali et al., 2018).

However, this study also has limitations. The sample size of our study is still insufficient and our research lacks of experimental validation. Non-etheless, we believe that the inflammatory regulation network as well as uterine muscle contraction and vascularity play an essential role in RIF.

Conclusion

In present context, we accomplished DEGs analysis of endometrial tissue between RIF and control women in two datasets and identified 26 common DEGs. KEGG/BioCarta pathway and immune infiltration analysis determined inflammation associated pathways and cells might contribute to RIF. GO terms, five hub genes (*PTGS2*, *VCAM1*, *EDNRB*, *ACTA2*, and *LIF*) and related TF-gene and TF-miRNA interactions were identified, suggesting that inflammation, uterine muscle contraction and vascularity were key pathophysiological changes in RIF. Of interest, subgroup identification revealed that the patients with RIF can be divided into two subgroups, and a small number of patients have atypical phenotypes (cluster 1), which may be the reason for the poor prognosis of RIF. ROC curves and drugs affirmed the diagnostic and therapeutic values of hub genes. Those results may help us expand the understanding of RIF and may provide evidences for the treatment of RIF. Further researches should consider the underlying mechanisms of the inflammatory regulation as well as uterine muscle contraction and vascularity in RIF, and find meaningful diagnostic and treatment methods.

Data availability statement

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

Author contributions

XD and MZ conducted the formal analysis and wrote the original draft; HH and XL revised the original draft; HH and YS provided funding acquisition. All authors contributed to the article and approved the submitted version.

Funding

This study was supported by the National Natural Science Foundation of China (No. 82130046), National Natural Science Foundation of China (No. 82101708), National Natural Science Foundation of China (No. 82201976), National Key R&D Program of China (2019YFA0802604), Shanghai leading talent program, Innovative research team of high-level local universities in Shanghai (No. SHSMU-ZLCX20210201, No. SSMU-ZLCX20180401), Shanghai Jiaotong University School of Medicine Affiliated Renji Hospital Clinical Research Innovation Cultivation Fund Program (RJPY-DZX-003), Shanghai Municipal Education Commission-Gaofeng Clinical Medicine Grant Support (No. 20161413).

Conflict of interest

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

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2023.1108805/full#supplementary-material

References

Ahmadi, M., Pashangzadeh, S., Moraghebi, M., Sabetian, S., Shekari, M., Eini, F., et al. (2022). Construction of circRNA-miRNA-mRNA network in the pathogenesis of recurrent implantation failure using integrated bioinformatics study. *J. Cell Mol. Med.* 26 (6), 1853–1864. doi:10.1111/jcmm.16586

Ali, M., Jackson-Bey, T., and Al-Hendy, A. (2018). Simvastatin and uterine fibroids: Opportunity for a novel therapeutic option. *Fertil. Steril.* 110 (7), 1272–1273. doi:10.1016/j.fertnstert.2018.08.043

Anamthathmakula, P., and Winuthayanon, W. (2021). Prostaglandin-endoperoxide synthase 2 (PTGS2) in the oviduct: Roles in fertilization and early embryo development. *Endocrinology* 162 (4), bqab025. doi:10.1210/endocr/bqab025

Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., et al. (2000). Gene ontology: Tool for the unification of biology. The gene ontology consortium. *Nat. Genet.* 25 (1), 25–29. doi:10.1038/75556

Bai, R., Bai, H., Kuse, M., Ideta, A., Aoyagi, Y., Fujiwara, H., et al. (2014). Involvement of VCAM1 in the bovine conceptus adhesion to the uterine endometrium. *Reproduction* 148 (2), 119–127. doi:10.1530/rep-13-0655

Banaszewska, B., Pawelczyk, L., Spaczynski, R. Z., and Duleba, A. J. (2011). Effects of simvastatin and metformin on polycystic ovary syndrome after six months of treatment. *J. Clin. Endocrinol. Metab.* 96 (11), 3493–3501. doi:10.1210/jc.2011-0501

Bashiri, A., Halper, K. I., and Orvieto, R. (2018). Recurrent Implantation Failure-update overview on etiology, diagnosis, treatment and future directions. *Reprod. Biol. Endocrinol.* 16 (1), 121. doi:10.1186/s12958-018-0414-2

Bastu, E., Demiral, I., Gunel, T., Ulgen, E., Gumusoglu, E., Hosseini, M. K., et al. (2019). Potential marker pathways in the endometrium that may cause recurrent implantation failure. *Reprod. Sci.* 26 (7), 879–890. doi:10.1177/1933719118792104

Bellver, J., and Simón, C. (2018). Implantation failure of endometrial origin: What is new? *Curr. Opin. Obstet. Gynecol.* 30 (4), 229–236. doi:10.1097/gco.0000000000000468

Brinsden, P. R., Alam, V., de Moustier, B., and Engrand, P. (2009). Recombinant human leukemia inhibitory factor does not improve implantation and pregnancy outcomes after assisted reproductive techniques in women with recurrent unexplained implantation failure. *Fertil. Steril.* 91 (4), 1445–1447. doi:10.1016/j.fertnstert.2008.06.047

Busnelli, A., Reschini, M., Cardellicchio, L., Vegetti, W., Somigliana, E., and Vercellini, P. (2020). How common is real repeated implantation failure? An indirect estimate of the prevalence. *Reprod. Biomed. Online* 40 (1), 91–97. doi:10.1016/j.rbmo.2019.10.014

Cao, F., Fan, Y., Yu, Y., Yang, G., and Zhong, H. (2021a). Dissecting prognosis modules and biomarkers in glioblastoma based on weighted gene Co-expression network analysis. *Cancer Manag. Res.* 13, 5477–5489. doi:10.2147/cmar.S310346

Cao, F., Wang, C., Long, D., Deng, Y., Mao, K., and Zhong, H. (2021b). Network-based integrated analysis of transcriptomic studies in dissecting gene Signatures for LPS-induced acute lung injury. *Inflammation* 44 (6), 2486–2498. doi:10.1007/s10753-021-01518-8

Chatterjee, P., Chiasson, V. L., Bounds, K. R., and Mitchell, B. M. (2014). Regulation of the anti-inflammatory cytokines interleukin-4 and interleukin-10 during pregnancy. *Front. Immunol.* 5, 253. doi:10.3389/fimmu.2014.00253

Cirillo, R., Tos, E. G., Page, P., Missotten, M., Quattropani, A., Scheer, A., et al. (2007). Arrest of preterm labor in rat and mouse by an oral and selective nonprostanoid antagonist of the prostaglandin F2alpha receptor (FP). *Am. J. Obstet. Gynecol.* 197 (1), 54.e1–54.e9. doi:10.1016/j.ajog.2007.02.010

Cooper, K., and Brown, S. (2017). ACTA2 mutation and postpartum hemorrhage: A case report. *BMC Med. Genet.* 18 (1), 143. doi:10.1186/s12881-017-0505-5

Dallagi, A., Girouard, J., Hamelin-Morrissette, J., Dadzie, R., Laurent, L., Vaillancourt, C., et al. (2015). The activating effect of IFN-γ on monocytes/macrophages is regulated by the LIF-trophoblast-IL-10 axis via Stat1 inhibition and Stat3 activation. *Cell Mol. Immunol.* 12 (3), 326–341. doi:10.1038/cmi.2014.50

Donaldson, J. (2022). *Tsne: T-distributed stochastic neighbor embedding for R (t-SNE)*.

Duan, Y., Liu, Y., Xu, Y., and Zhou, C. (2022). Bioinformatics analysis identifies key genes in recurrent implantation failure based on immune infiltration. *Reprod. Sci.* doi:10.1007/s43032-022-01060-4

Ersahin, A., Acet, M., Acet, T., and Yavuz, Y. (2016). Disturbed endometrial NF-κB expression in women with recurrent implantation failure. *Eur. Rev. Med. Pharmacol. Sci.* 20 (24), 5037–5040.

Fanchin, R., Righini, C., Olivennes, F., Taylor, S., de Ziegler, D., and Frydman, R. (1998). Uterine contractions at the time of embryo transfer alter pregnancy rates after *in-vitro* fertilization. *Hum. Reprod.* 13 (7), 1968–1974. doi:10.1093/humrep/13.7.1968

Feghali, C. A., and Wright, T. M. (1997). Cytokines in acute and chronic inflammation. *Front. Biosci.* 2, d12–d26. doi:10.2741/a171

Gadient, R. A., and Patterson, P. H. (1999). Leukemia inhibitory factor, interleukin 6, and other cytokines using the GP130 transducing receptor: Roles in inflammation and injury. *Stem Cells* 17 (3), 127–137. doi:10.1002/stem.170127

Gao, G., Li, C., Fan, W., Zhang, M., Li, X., Chen, W., et al. (2021). Brilliant glycans and glycosylation: Seq and ye shall find. *Int. J. Biol. Macromol.* 189, 279–291. doi:10.1016/j.ijbiomac.2021.08.054

Gram, A., Boos, A., and Kowalewski, M. P. (2017). Cellular localization, expression and functional implications of the utero-placental endothelin system during maintenance and termination of canine gestation. *J. Reprod. Dev.* 63 (3), 235–245. doi:10.1262/jrd.2016-165

Gram, A., Fox, B., Büchler, U., Boos, A., Hoffmann, B., and Kowalewski, M. P. (2014). Canine placental prostaglandin E2 synthase: Expression, localization, and biological functions in providing substrates for prepartum PGF2alpha synthesis. *Biol. Reprod.* 91 (6), 154. doi:10.1095/biolreprod.114.122929

Gu, Z., Eils, R., and Schlesner, M. (2016). Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* 32 (18), 2847–2849. doi:10.1093/bioinformatics/btw313

Guo, F., Si, C., Zhou, M., Wang, J., Zhang, D., Leung, P. C. K., et al. (2018). Decreased PECAM1-mediated TGF-β1 expression in the mid-secretory endometrium in women with recurrent implantation failure. *Hum. Reprod.* 33 (5), 832–843. doi:10.1093/humrep/dey022

Gurtner, G. C., Davis, V., Li, H., McCoy, M. J., Sharpe, A., and Cybulsky, M. I. (1995). Targeted disruption of the murine VCAM1 gene: Essential role of VCAM-1 in chorioallantoic fusion and placentation. *Genes Dev.* 9 (1), 1–14. doi:10.1101/gad.9.1.1

Hänzelmann, S., Castelo, R., and Guinney, J. (2013). Gsva: Gene set variation analysis for microarray and RNA-seq data. *BMC Bioinforma.* 14, 7. doi:10.1186/1471-2105-14-7

He, Y., Li, H., Yao, J., Zhong, H., Kuang, Y., Li, X., et al. (2021). HO-1 knockdown upregulates the expression of VCAM-1 to induce neutrophil recruitment during renal ischemia-reperfusion injury. *Int. J. Mol. Med.* 48 (4), 185. doi:10.3892/ijmm.2021.5018

Inhorn, M. C., and Patrizio, P. (2015). Infertility around the globe: New thinking on gender, reproductive technologies and global movements in the 21st century. *Hum. Reprod. Update* 21 (4), 411–426. doi:10.1093/humupd/dmv016

Kanehisa, M., and Goto, S. (2000). Kegg: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* 28 (1), 27–30. doi:10.1093/nar/28.1.27

Kassambara, A. (2020). *ggpubr: 'ggplot2*. Based Publication Ready Plots.

Keator, C. S., Mah, K., Ohm, L., and Slayden, O. D. (2011). Estrogen and progesterone regulate expression of the endothelins in the rhesus macaque endometrium. *Hum. Reprod.* 26 (7), 1715–1728. doi:10.1093/humrep/der115

Kimber, S. J. (2005). Leukaemia inhibitory factor in implantation and uterine biology. *Reproduction* 130 (2), 131–145. doi:10.1530/rep.1.00304

Konac, E., Alp, E., Onen, H. I., Korucuoglu, U., Biri, A. A., and Menevse, S. (2009). Endometrial mRNA expression of matrix metalloproteinases, their tissue inhibitors and cell adhesion molecules in unexplained infertility and implantation failure patients. *Reprod. Biomed. Online* 19 (3), 391–397. doi:10.1016/s1472-6483(10)60174-5

Kong, D. H., Kim, Y. K., Kim, M. R., Jang, J. H., and Lee, S. (2018). Emerging roles of vascular cell adhesion molecule-1 (VCAM-1) in immunological disorders and cancer. *Int. J. Mol. Sci.* 19 (4), 1057. doi:10.3390/ijms19041057

Kowalewski, M. P., Beceriklisoy, H. B., Pfarrer, C., Aslan, S., Kindahl, H., Kücükaslan, I., et al. (2010). Canine placenta: A source of prepartal prostaglandins during normal and antiprogestin-induced parturition. *Reproduction* 139 (3), 655–664. doi:10.1530/rep-09-0140

Kusakabe, K., Kiso, Y., Hondo, E., Takeshita, A., Kato, K., Okada, T., et al. (2009). Spontaneous endometrial hyperplasia in the uteri of IL-2 receptor beta-chain transgenic mice. *J. Reprod. Dev.* 55 (3), 273–277. doi:10.1262/jrd.20171

Lédée, N., Munaut, C., Aubert, J., Sérazin, V., Rahmati, M., Chaouat, G., et al. (2011). Specific and extensive endometrial deregulation is present before conception in IVF/ICSI repeated implantation failures (IF) or recurrent miscarriages. *J. Pathol.* 225 (4), 554–564. doi:10.1002/path.2948

Liang, P. Y., Diao, L. H., Huang, C. Y., Lian, R. C., Chen, X., Li, G. G., et al. (2015). The pro-inflammatory and anti-inflammatory cytokine profile in peripheral blood of women with recurrent implantation failure. *Reprod. Biomed. Online* 31 (6), 823–826. doi:10.1016/j.rbmo.2015.08.009

Luo, L., Li, D. H., Li, X. P., Zhang, S. C., Yan, C. F., Wu, J. F., et al. (2016). Polymorphisms in the nuclear factor kappa B gene association with recurrent embryo implantation failure. *Genet. Mol. Res.* 15 (2). doi:10.4238/gmr.15027759

Mackler, A. M., Ducsay, T. C., Ducsay, C. A., and Yellon, S. M. (2003). Effects of endotoxin and macrophage-related cytokines on the contractile activity of the gravid murine uterus. *Biol. Reprod.* 69 (4), 1165–1169. doi:10.1095/biolreprod.103.015586

Maglott, D., Ostell, J., Pruitt, K. D., and Tatusova, T. (2011). Entrez gene: Gene-centered information at NCBI. *Nucleic Acids Res.* 39, D52–D57. Database issue). doi:10.1093/nar/gkq1237

Mao, K., Tan, Q., Ma, Y., Wang, S., Zhong, H., Liao, Y., et al. (2021). Proteomics of extracellular vesicles in plasma reveals the characteristics and residual traces of COVID-19 patients without underlying diseases after 3 months of recovery. *Cell Death Dis.* 12 (6), 541. doi:10.1038/s41419-021-03816-3

Margalioth, E. J., Ben-Chetrit, A., Gal, M., and Eldar-Geva, T. (2006). Investigation and treatment of repeated implantation failure following IVF-ET. *Hum. Reprod.* 21 (12), 3036–3043. doi:10.1093/humrep/del305

Mekinian, A., Cohen, J., Alijotas-Reig, J., Carbillon, L., Nicaise-Roland, P., Kayem, G., et al. (2016). Unexplained recurrent miscarriage and recurrent implantation failure: Is there a place for immunomodulation? *Am. J. Reprod. Immunol.* 76 (1), 8–28. doi:10.1111/aji.12493

Mor, G., Cardenas, I., Abrahams, V., and Guller, S. (2011). Inflammation and pregnancy: The role of the immune system at the implantation site. *Ann. N. Y. Acad. Sci.* 1221 (1), 80–87. doi:10.1111/j.1749-6632.2010.05938.x

Mrozikiewicz, A. E., Ożarowski, M., and Jędrzejczak, P. (2021). Biomolecular markers of recurrent implantation failure-A review. *Int. J. Mol. Sci.* 22 (18), 10082. doi:10.3390/ijms221810082

Nardo, L. G., and Sallam, H. N. (2006). Progesterone supplementation to prevent recurrent miscarriage and to reduce implantation failure in assisted reproduction cycles. *Reprod. Biomed. Online* 13 (1), 47–57. doi:10.1016/s1472-6483(10)62015-9

Ni, L., Tang, C., Wang, Y., Wan, J., Charles, M. G., Zhang, Z., et al. (2022). Construction of a miRNA-based nomogram model to predict the prognosis of endometrial cancer. *J. Pers. Med.* 12 (7), 1154. doi:10.3390/jpm12071154

Ono, Y., Yoshino, O., Hiraoka, T., Sato, E., Fukui, Y., Ushijima, A., et al. (2020). CD206+ M2-like macrophages are essential for successful implantation. *Front. Immunol.* 11, 557184. doi:10.3389/fimmu.2020.557184

Orsi, N. M., and Tribe, R. M. (2008). Cytokine networks and the regulation of uterine function in pregnancy and parturition. *J. Neuroendocrinol.* 20 (4), 462–469. doi:10.1111/j.1365-2826.2008.01668.x

Orvieto, R., Brengauz, M., and Feldman, B. (2015). A novel approach to normal responder patient with repeated implantation failures--a case report. *Gynecol. Endocrinol.* 31 (6), 435–437. doi:10.3109/09513590.2015.1005595

Robin, X., Turck, N., Hainard, A., Tiberti, N., Lisacek, F., Sanchez, J. C., et al. (2011). pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinforma.* 12, 77. doi:10.1186/1471-2105-12-77

Sadowsky, D. W., Haluska, G. J., Gravett, M. G., Witkin, S. S., and Novy, M. J. (2000). Indomethacin blocks interleukin 1beta-induced myometrial contractions in pregnant rhesus monkeys. *Am. J. Obstet. Gynecol.* 183 (1), 173–180. doi:10.1067/mob.2000.105968

Serafini, P., Rocha, A. M., Osório, C. T., da Silva, I., Motta, E. L., and Baracat, E. C. (2008). Endometrial leukemia inhibitory factor as a predictor of pregnancy after *in vitro* fertilization. *Int. J. Gynaecol. Obstet.* 102 (1), 23–27. doi:10.1016/j.ijgo.2007.12.005

Shang, J., Cheng, Y. F., Li, M., Wang, H., Zhang, J. N., Guo, X. M., et al. (2022). Identification of key endometrial MicroRNAs and their target genes associated with pathogenesis of recurrent implantation failure by integrated bioinformatics analysis. *Front. Genet.* 13, 919301. doi:10.3389/fgene.2022.919301

Shapouri-Moghaddam, A., Mohammadian, S., Vazini, H., Taghadosi, M., Esmaeili, S. A., Mardani, F., et al. (2018). Macrophage plasticity, polarization, and function in health and disease. *J. Cell Physiol.* 233 (9), 6425–6440. doi:10.1002/jcp.26429

Shuya, L. L., Menkhorst, E. M., Yap, J., Li, P., Lane, N., and Dimitriadis, E. (2011). Leukemia inhibitory factor enhances endometrial stromal cell decidualization in humans and mice. *PLoS One* 6 (9), e25288. doi:10.1371/journal.pone.0025288

Sun, X., Icli, B., Wara, A. K., Belkin, N., He, S., Kobzik, L., et al. (2012). MicroRNA-181b regulates NF-κB-mediated vascular inflammation. *J. Clin. Invest.* 122 (6), 1973–1990. doi:10.1172/jci61495

Sun, X., Sit, A., and Feinberg, M. W. (2014). Role of miR-181 family in regulating vascular inflammation and immunity. *Trends Cardiovasc Med.* 24 (3), 105–112. doi:10.1016/j.tcm.2013.09.002

Taylor, H. S., Alderman, M., Iii, D'Hooghe, T. M., Fazleabas, A. T., and Duleba, A. J. (2017). Effect of simvastatin on baboon endometriosis. *Biol. Reprod.* 97 (1), 32–38. doi:10.1093/biolre/iox058

Vane, J. R., Bakhle, Y. S., and Botting, R. M. (1998). Cyclooxygenases 1 and 2. *Annu. Rev. Pharmacol. Toxicol.* 38, 97–120. doi:10.1146/annurev.pharmtox.38.1.97

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. New York: Springer-Verlag.

Wilkerson, M. D., and Hayes, D. N. (2010). ConsensusClusterPlus: A class discovery tool with confidence assessments and item tracking. *Bioinformatics* 26 (12), 1572–1573. doi:10.1093/bioinformatics/btq170

Wu, R. C., Jiang, M., Beaudet, A. L., and Wu, M. Y. (2013). ARID4A and ARID4B regulate male fertility, a functional link to the AR and RB pathways. *Proc. Natl. Acad. Sci. U. S. A.* 110 (12), 4616–4621. doi:10.1073/pnas.1218318110

Ye, Z., Wang, F., Yan, F., Wang, L., Li, B., Liu, T., et al. (2019). Bioinformatic identification of candidate biomarkers and related transcription factors in nasopharyngeal carcinoma. *World J. Surg. Oncol.* 17 (1), 60. doi:10.1186/s12957-019-1605-9

Yu, H., Lin, L., Zhang, Z., Zhang, H., and Hu, H. (2020). Targeting NF-κB pathway for the therapy of diseases: Mechanism and clinical study. *Signal Transduct. Target Ther.* 5 (1), 209. doi:10.1038/s41392-020-00312-6

Zahid, M. D. K., Rogowski, M., Ponce, C., Choudhury, M., Moustaid-Moussa, N., and Rahman, S. M. (2020). CCAAT/enhancer-binding protein beta (C/EBPβ) knockdown reduces inflammation, ER stress, and apoptosis, and promotes autophagy in oxLDL-treated RAW264.7 macrophage cells. *Mol. Cell Biochem.* 463 (1-2), 211–223. doi:10.1007/s11010-019-03642-4

Zhang, H. M., Kuang, S., Xiong, X., Gao, T., Liu, C., and Guo, A. Y. (2015). Transcription factor and microRNA co-regulatory loops: Important regulatory motifs in biological processes and diseases. *Brief. Bioinform* 16 (1), 45–58. doi:10.1093/bib/bbt085

Zhang, H., Wu, J., Li, Y., Jin, G., Tian, Y., and Kang, S. (2022). Identification of key differentially methylated/expressed genes and pathways for ovarian endometriosis by bioinformatics analysis. *Reprod. Sci.* 29 (5), 1630–1643. doi:10.1007/s43032-021-00751-8

Zhao, J., Tian, M., Zhang, S., Delfarah, A., Gao, R., Rao, Y., et al. (2020). Deamidation shunts RelA from mediating inflammation to aerobic glycolysis. *Cell Metab.* 31 (5), 937–955.e7. doi:10.1016/j.cmet.2020.04.006

Zhou, J. G., Zhong, H., Zhang, J., Jin, S. H., Roudi, R., and Ma, H. (2019). Development and validation of a prognostic signature for malignant pleural mesothelioma. *Front. Oncol.* 9, 78. doi:10.3389/fonc.2019.00078

Zhou, T., Ni, T., Li, Y., Zhang, Q., Yan, J., and Chen, Z. J. (2021). circFAM120A participates in repeated implantation failure by regulating decidualization via the miR-29/ABHD5 axis. *Faseb J.* 35 (9), e21872. doi:10.1096/fj.202002298RR

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 27 February 2023
DOI 10.3389/fgene.2023.1118273

OPEN ACCESS

EDITED BY
Shuai Liu,
University of Hawaii at Manoa,
United States

REVIEWED BY
Chen Li,
Free University of Berlin, Germany
Shiqiang Jin,
Bristol Myers Squibb, United States
Chengxuan Chen,
Texas A&M University, United States

*CORRESPONDENCE
Yuan Gao,
rj_gaoyuan@163.com
Huijing Huang,
fangfeijin90@163.com

†These authors have contributed equally to this work

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

RECEIVED 07 December 2022
ACCEPTED 15 February 2023
PUBLISHED 27 February 2023

CITATION
Mao K, Tang R, Wu Y, Zhang Z, Gao Y and Huang H (2023), Prognostic markers of ferroptosis-related long non-coding RNA in lung adenocarcinomas.
Front. Genet. 14:1118273.
doi: 10.3389/fgene.2023.1118273

Prognostic markers of ferroptosis-related long non-coding RNA in lung adenocarcinomas

Kaimin Mao[1†], Ri Tang[1†], Yali Wu[2], Zhiyun Zhang[1], Yuan Gao[1*] and Huijing Huang[3*]

[1]Department of Critical Care Medicine, Renji Hospital, School of Medicine, Shanghai Jiaotong University, Shanghai, China, [2]Department of Respiratory and Critical Care Medicine, NHC Key Laboratory of Pulmonary Diseases, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China, [3]Department of Rheumatology, Zhongshan Hospital, Fudan University, Shanghai, China

Ferroptosis is a recently established type of iron-dependent programmed cell death. Growing studies have focused on the function of ferroptosis in cancers, including lung adenocarcinoma (LUAD). However, the factors involved in the regulation of ferroptosis-related genes are not fully understood. In this study, we collected data from lung adenocarcinoma datasets of the Cancer Genome Atlas (TCGA-LUAD). The expression profiles of 60 ferroptosis-related genes were screened, and two differentially expressed ferroptosis subtypes were identified. We found the two ferroptosis subtypes can predict clinical outcomes and therapeutic responses in LUAD patients. Furthermore, key long non-coding RNAs (lncRNAs) were screened by single factor Cox and least absolute shrinkage and selection operator (LASSO) based on which co-expressed with the 60 ferroptosis-related genes. We then established a risk score model which included 13 LUAD ferroptosis-related lncRNAs with a multi-factor Cox regression. The risk score model showed a good performance in evaluating the outcome of LUAD. What's more, we divided TCGA-LUAD tumor samples into two groups with high- and low-risk scores and further explored the differences in clinical characteristics, tumor mutation burden, and tumor immune cell infiltration among different LUAD tumor risk score groups and evaluate the predictive ability of risk score for immunotherapy benefit. Our findings provide good support for immunotherapy in LUAD in the future.

KEYWORDS

ferroptosis, lncRNA, lung adenocarcinoma, risk scores model, immunotherapy

Introduction

Lung cancer is one of the most common malignant tumors and the leading cause of cancer-related deaths worldwide. Despite the continuous emergence of new treatments, the prognosis of lung cancer is still very poor (Siegel et al., 2020). Non-small-cell lung cancer (NSCLC) is the main histologic subtype of lung cancer, it can be classified as lung adenocarcinoma (LUAD), lung squamous cell carcinoma (LUSC), and large-cell carcinoma, of which LUAD is the most common subtype (Relli et al., 2019). It is important to identify effective biomarkers for the prognosis of LUAD because, even

though there are a variety of treatment plans for this cancer, the average 5-year survival rate is only about 15% (Spella and Stathopoulos, 2021).

Ferroptosis is a new type of iron-dependent programmed cell death that differs from apoptosis, necrosis, and autophagy. It induces cell injury or death *via* the iron-dependent lipid peroxidation process (Latunde-Dada, 2017; Xu et al., 2023). Ferroptosis is characterized by increased mitochondrial membrane density and cell volume contraction, which is different from other morphological, biochemical, and genetically regulated cell deaths (Hassannia et al., 2019; Li et al., 2020). Studies have shown that ferroptosis inhibits tumor growth, kills tumor cells, and prevents tumor migration (Mou et al., 2019). Accumulating evidence has suggested that ferroptosis is associated with several biological processes in LUAD. For example, CAMP-responsive element binding protein 1 (CREB) can directly bind to the promoter region of glutathione peroxidase 4 (GPX4) to promote its expression, thereby inhibiting potential ferroptosis and promoting the growth of LUAD (Wang Z. et al., 2021). Besides, the novel 15-gene signature of ferroptosis provides a basis for an accurate prediction of the prognosis of LUAD, allowing for the development of new therapies and personalized outcome prediction in this population (Zhang A. et al., 2021). Therefore, it is necessary to find new treatment strategies to improve the prognosis of LUAD by regulating ferroptosis.

Recent advances in sequencing technologies have shown that 90% of RNAs do not encode proteins, which are called non-coding RNA (ncRNA) (Matsui and Corey, 2017). Long ncRNA (LncRNA) is a type of ncRNA. It has a length of more than 200 nucleotides and is mainly involved in regulating gene promoters and enhancers as well as RNA splicing (Ali and Grote, 2020). Several studies have indicated that RNA plays an important role in the development of cancer, its metastatic and genital development, and so it is now an important candidate for cancer treatment (Li et al., 2016; Liu S. J. et al., 2021). What's more, lncRNAs are increasingly recognized as crucial mediators in the regulation of ferroptosis (Gibb et al., 2011). For example, Chao Mao et al. demonstrated that the cytosolic lncRNA P53RRA promotes ferroptosis and apoptosis in lung cancer *via* nuclear sequestration of p53 (Jiang et al., 2015). In addition, it was demonstrated that lncRNA LINC00336, which is associated with ferroptosis, is highly expressed in lung cancer, and acts as a competitive endogenous RNA to function as an oncogene (Wang et al., 2020). However, the full role of ferroptosis-related lncRNAs in LUAD is still not completely understood. For new therapeutic strategies for patients with LUAD, ferroptosis-related lncRNAs must be identified to predict their outcome.

Anti-tumor immune response has long been a fundamental strategy in cancer immunotherapy (Liang et al., 2021). While ferroptosis plays a key role in tumor immunity. Therefore, it is important to explore biomarkers associated with tumor immunity and ferroptosis for immunotherapy of lung cancer. In this study, a ferroptosis-related lncRNA signature associated with LUAD prognosis is being explored based on the LUAD dataset of TCGA. To predict the survival of LUAD patients, a ferroptosis-related lncRNA risk score model was established by univariate and multivariate Cox regression analysis. In addition, the acting mechanism of ferroptosis-related lncRNAs in tumor progression was further mined by functional analysis and immune infiltration analysis to provide new insights into the prognosis and immunotherapy of LUAD. Our study provides insights into the mechanisms underlying ferroptosis in the treatment of LUAD, which may improve individualized therapy and the assessment of prognosis for LUAD.

Materials and methods

Acquisition of gene expression and clinical data

The process flow of this study is shown in Figure 1. Briefly, the LUAD expression profiles and clinical follow-up information were downloaded from the TCGA database (https://portal.gdc.cancer.gov/). The RNA-Seq data of TCGA-LUAD was processed in the following steps. Samples without clinical follow-up information and survival time were removed. We also excluded patients who survived less than 30 days and with no survival status. We converted probes to Gene Symbol, with one probe corresponding to multiple genes. Besides, we used the median value for the expression of multiple Gene Symbols. Finally, 489 tumor samples were included from the pre-processed TCGA-LUAD, as shown in Supplementary Table S1.

Consensus clustering of tumor ferroptosis-related gene expression

Ferroptosis is a new type of programmed cell death that differs from apoptosis, necrosis, and autophagy. As a result of divalent iron or ester oxygenase action, it causes unsaturated fatty acids highly expressed on the cell membrane to undergo lipid peroxidation, thus leading to cell death. Aside from this, it also acts as an antioxidant system (glutathione), which reduces the GPX4 enzyme. To ensure the stability of the classification, we used the ConsensuClusterPlus package in R and the Pam method based on Euclid and Ward linkages.

Differentially expressed genes among tumor ferroptosis subtypes (Fer_DEGs)

Two groups of samples of Fer-1 and Fer-2 were acquired based on the expression of tumor ferroptosis-related genes and consistent clustering results. The screening threshold was set as adjusted. $P < 0.05$ and $|\log2 (\text{Fold Change})| > 1$. A differentially expressed gene was analyzed between two subtypes using the "limma" package in R software. In addition, the Ensemble display was used to extract lncRNAs from differentially expressed genes.

Gene ontology and kyoto encyclopedia of genes and genomes pathway enrichment analyses

The co-expression genes of differential ferroptosis-related genes between high- and low-risk LUAD patients were chosen to perform

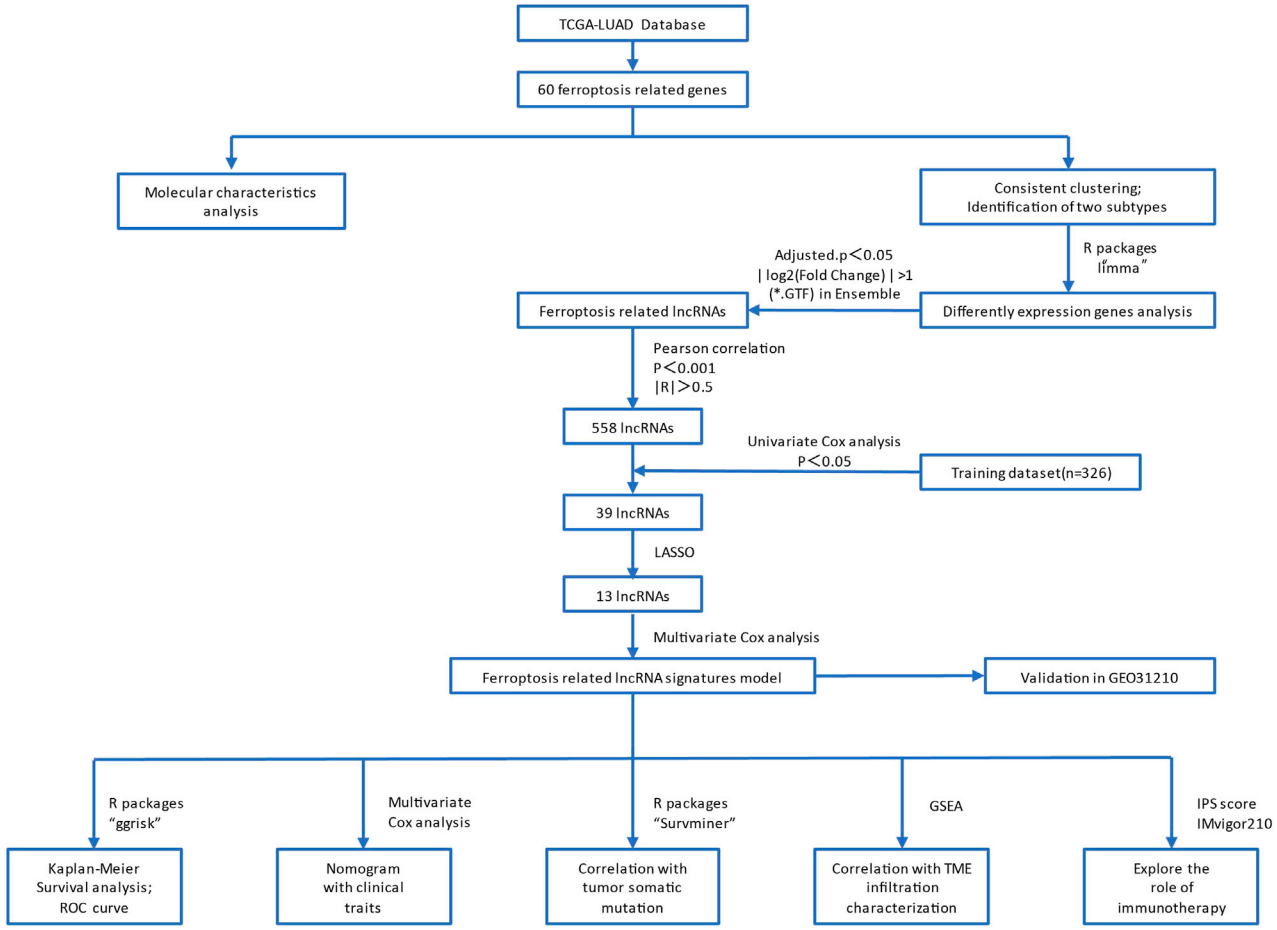


FIGURE 1
The flow chart of this study.

Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analyses, which was conducted by using the clusterProfiifiler package. Enrichment significance thresholds were set at $p < 0.05$ and false discovery rate (FDR) < 0.05 (Guo X. H. et al., 2021; Cao et al., 2021b). GO analysis was used to map all DEGs to GO terms in the GO database (http://www.geneontology.org/) to analyze the main functions of the DEGs. The KEGG pathway database (http://www.geneontology.org/) is a synthetic database, which was used to analyze the biochemical pathways of the DEGs of interest (Zhong H. et al., 2021).

Construction of ferroptosis-related lncRNA risk score model

To calculate the risk score for LUAD, we constructed a model based on the lncRNAs associated with ferroptosis subtypes. To reduce noise or redundant genes, a univariate Cox algorithm was applied to narrow the lncRNA set associated with immune cell infiltration subtypes. The best prognostic signature was identified by using the Lasso method [Least absolute shrinkage and selection operator, Tibshirani (1996)] A multi-factor Cox regression analysis contributed to the development of a risk score model for tumor immune cell infiltration. The formula was as follows:

$$\text{Risk\_}scores = \sum Coef(i) * Exp(i)$$

Gene set enrichment analysis (GSEA)

GSEA was published in 2005 based on gene set enrichment analysis. Genome-wide expression profiles can be interpreted using this knowledge-based approach. Using MSigDB (gene matrix transposition file format *.gmt) we selected one or more functional gene sets to analyze gene expression data (Guo Y. et al., 2021). We then sorted the gene expression data by correlation degree of phenotype (also known as a change in expression amount). To evaluate the influence of synergistic changes in genes on phenotypic changes, we sorted by phenotypic relevance the genes enriched in the upper and lower parts of the gene list.

Independent prognostic factors analysis of risk score and construction of a nomogram prediction model

After the extraction of clinical information (including age, gender, smoking, and TNM stage) of LUAD patients in the TCGA, univariate and multivariate prognostic analyses were used to demonstrate whether the risk score could be an independent prognostic factor. Based on the multivariate Cox regression analysis for risk score and other clinicopathological factors by the rms R package, a clinically adaptable nomogram prediction model was established to predict the survival probability of 489 LUAD individuals in 1-, 3-, and 5- years from the TCGA group. Then, the calibration analysis and time-dependent ROC curve were used to evaluate the prognostic value of the nomogram for LUAD patients (Sun et al., 2022).

Analysis of the tumor mutation burden in the high- and low-tumor risk score groups

Tumor mutational burden (TMB) is broadly defined as the number of somatic mutations per megabase of interrogated genomic sequence (Bravaccini et al., 2021). To inquire about the association between the TMB and tumor risk score, we compared the tumor mutation status between the low- and high-risk score groups. The somatic mutation file *.maf of TCGA-LUAD was downloaded from the GDC Data Portal (https://portal.gdc.cancer.gov) to calculate the TMB values. Significantly mutated genes ($p < 0.05$) between the low- and high-risk groups and the interaction effect of gene mutations were analyzed by maftools; only genes mutating more than 50 times in at least one group will be considered. The statistical significance test for the proportion of mutation was evaluated by Pearson correlation coefficient, student *t* test, Chi-square test, and survival analysis.

Relationship between tumor risk score and tumor microenvironment

Based on the LM22 signature and 1,000 permutations, the mutations of 22 different immune cells in TCGA-LUAD (B.cells.naive, B.cells.memory, Plasma.cells, T.cells.CD8, T.cells.CD4.naive, T.cells.CD4.memory.resting, T.cells.CD4.memory.activated, T.cells.follicular.helper, T.cells.regulatory.Tregs, T.cells.gamma.delta, NK.cells.resting, NK.cells.activated, Monocytes, Macrophages.M0, Macrophages.M1, Macrophages.M2, Dendritic.cells.resting, Dendritic.cells.activated, Mast.cells.resting, Mast.cells.activated, Eosinophils, Neutrophils) infiltration levels were quantified by using the CIBERSORT package in R. Besides, differences in the degree of immune cell infiltration between high- and low-risk groups were compared.

Correlation analyses between tumor risk score and immunotherapy response

The correlation between tumor risk score and immunotherapy response can evaluate the effect of the tumor risk score in predicting the benefit of immunotherapy in treating LUAD patients. In this study, we compared the immunotherapy response between the high- and low-risk groups based on expression profile data and clinical information in the IMvigor210 cohort (http://research-pub.gene.com/IMvigor210CoreBiologies/).

Reverse transcription-quantitative PCR (RT-qPCR)

Five paired LUAD tissues and corresponding adjacent non-tumorous tissues were obtained from patients who underwent radical resection of lung cancer in Renji hospital, Total RNA was extracted with TRIzol™ Reagent (Invitrogen). Reverse transcription of RNA was performed using PrimeScript™ RT Master Mix (Takara). In this study, Takara's TB Green™ Premix EX Taq™ II was used to perform the qPCR. GAPDH was used as an internal control (Cao et al., 2021a; Fei et al., 2021). The primer sequence of the tested genes is shown in Supplementary Table S6. The relative lncRNA expression level was quantified using the $2^{-\Delta\Delta Ct}$ method.

Statistical analysis and hypothesis testing

All statistical comparisons involved in this study, as well as hypothesis testing of the significance of differences between groups, were based on the statistical analysis method in R 3.6.

Results

Molecular characteristics of ferroptosis-related genes in LUAD

The flow chart of this study was shown in Figure 1. Based on the expression values of 60 ferroptosis-related genes in each sample of the TCGA-LUAD dataset, the genes were divided into a high-expression group and a low-expression group according to the optimal density algorithm. The high expressions of GLS2, PHKG2, ACACA, GPX4, DPP4, NCOA4, ACO1, PEBP1, NOX1, ZEB1, ALOX15, ALOX5, CRYAB, SAT1, and ACSF2 are significantly associated with better OS prognosis. While the low expressions of GCLM, GCLC, EMC2, SQLE, IREB2, FANCD2, AKR1C3, AKR1C2, TFRC, PGD, G6PD, ACSL4, CISD1, SLC7A11, ACSL3, and GOT1 have great significance with better OS prognosis (Figure 2).

Subsequently, the statistics of gene mutations in the TCGA-LUAD showed that 88.95% of tumor samples had gene mutations, including 47% of TP53 mutations, 41% of TTN mutations, 40% of MUC16 mutations, and 34% of RYR2 mutations (Supplementary Figure S1).

Furthermore, we conducted a hypothesis test on whether TP53 and TTN affect the expression of 60 ferroptosis-related genes. We found that the mutation of the TP53 gene was significantly associated with the high expression of CBS, GCLM, FANCD2, GSS, HSPB1, MT1G, TFRC, SQLE, FADS2, and NFS1 genes, while it has a remarkable correlation with the low expression of PEBP1, TP53, FDFT1, SLC7A11, CRYAB, NCOA4,

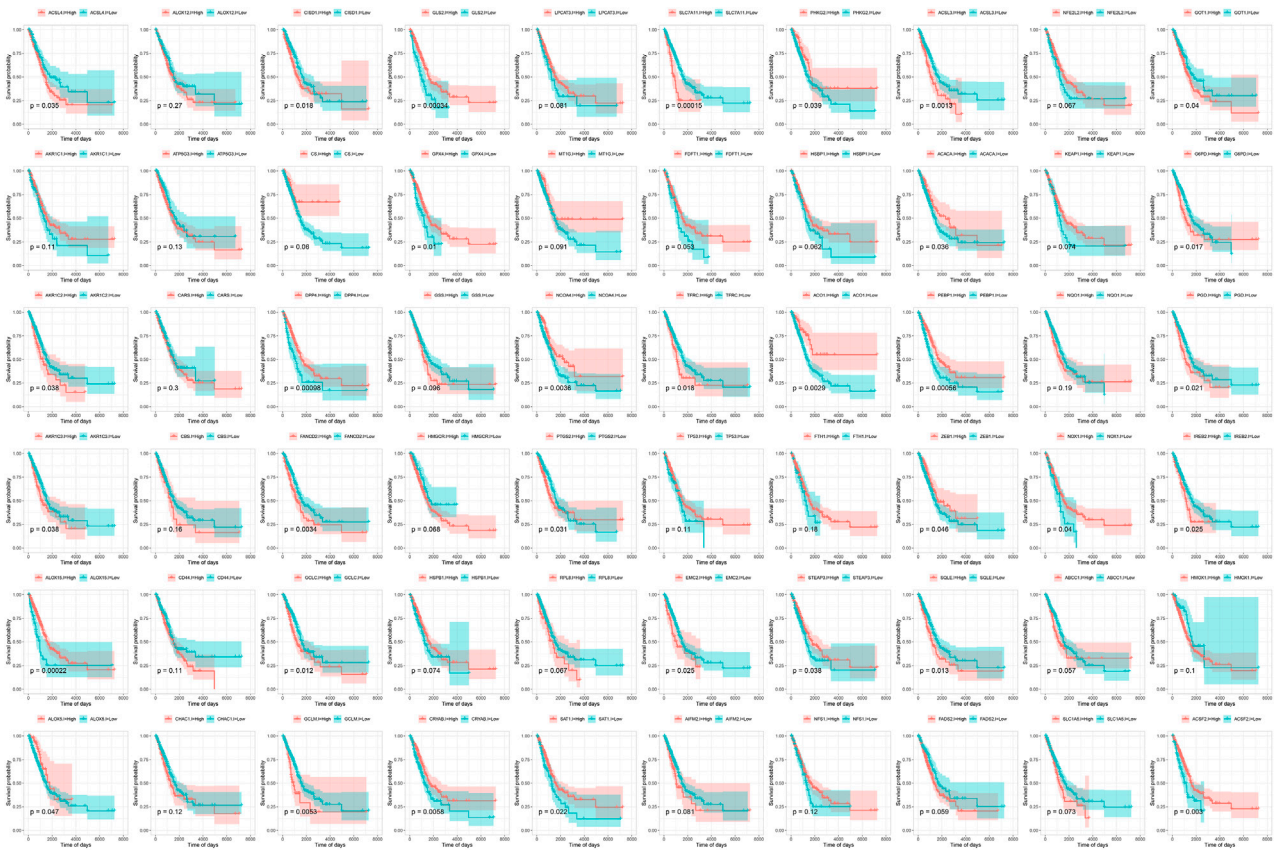


FIGURE 2
Survival curve of 60 ferroptosis-related genes and overall survival in the TCGA-LUAD data set.

SAT1, GLS2, AKR1C1, and AKR1C3 (Supplementary Figure S2). Among the mutation groups with TTN, ATP5G3, CARS, CBS, GPX4, GCLM, GCLC, FANCD2, CS, CISD1, CHAC1, GSS, HSPB1, RPL8, ACO1, EMC2, TFRC, NFS1, ZEB1, SQLE, FADS2, IREB2, PGD, and SLC1A5 were significantly highly expressed, while ALOX5, CD44, CRYAB, and SAT1 showed a significantly low expression status (Supplementary Figure S3). At the same time, we observed that most of the expressions of 60 ferroptosis-related genes were mutually promoting, as shown in Supplementary Figure S4.

Identification of ferroptosis subtypes and differentially expressed genes in LUAD

Consensus clustering was performed based on the expression of 60 ferroptosis-related genes in the TCGA-LUAD, and we determined two independent ferroptosis subtypes with a significant difference in survival. Among the two ferroptosis subtypes, Fer-1 has a significantly better prognosis than Fer-2, with a median survival time of 898 days. While Fer-2 indicated a worse disease prognosis, with a median survival time of 685 days (Figure 3).

In order to reveal the potential biological characteristics of different ferroptosis states, we used the "limma" package of R software to analyze differentially expressed genes between the subtypes. 882 genes were identified with an adjusted $p < 0.05$ and $|\log2\ (\text{Fold Change})| > 1$ (Supplementary Table S2). Among them, 511 genes were highly expressed in Fer-1 subtypes, while 371 genes were upregulated in Fer-2 (Figure 4A). Subsequently, we performed the Gene Ontology (GO) functional enrichment analysis on highly expressed genes. The first 10 pathways enriched in the three functional categories (BP, CC, and MF) were displayed with bubble diagrams (Figures 4B, C). Most of the pathways in Fer-1 were correlated with biological processes such as response to xenobiotic stimulus, hormone metabolic process, and antibiotic metabolic process. While in Fer-2, most of the enrichments were related to viral entry into the host cell, leukotriene metabolic process, and fluid transport.

Then, we performed Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis on the DEGs, and the first 12 enriched pathways were determined. As shown in Figure 4D, they were allograft rejection, graft *versus* host disease, asthma, intestinal immune network for iga production, hematopoietic cell lineage, metabolism of xenobiotics by cytochrome p450, ascorbate and aldarate metabolism, pentose and glucuronate interconversions, folate biosynthesis, phenylalanine metabolism, glutathione metabolism, porphyrin metabolism, and porphyrin metabolism, porphyrin metabolism. To further explore the relationship between tumor ferroptosis subtypes and tumor immune cells,

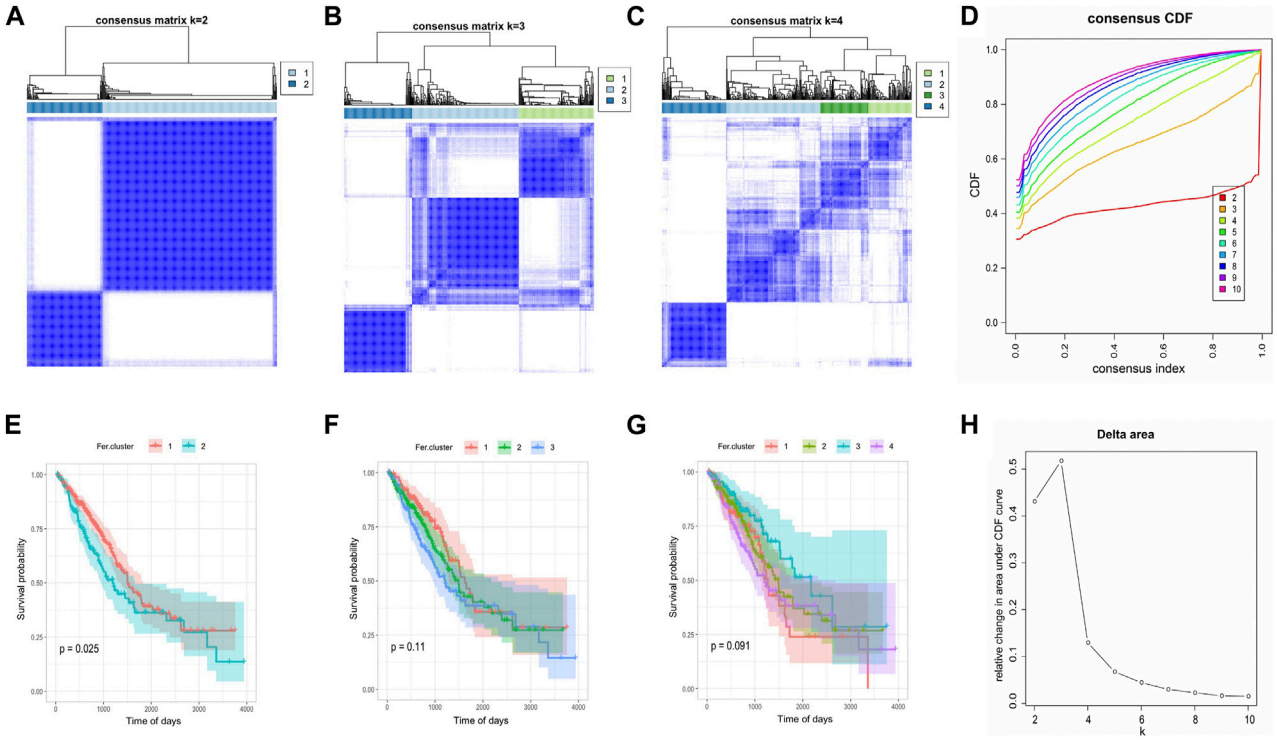


FIGURE 3
Consensus clustering of tumor ferroptosis-related genes expression profiles. **(A–C)** The clustering results when consensus matrix k = 2, k = 3, and k = 4. **(D)** Distribution of CDF curve of consensus clustering. **(E–G)** Survival curve when consensus matrix k = 2, k = 3, and k = 4, respectively. **(H)** Distribution of area under the CDF curve of consensus clustering.

firstly, we used principal component analysis (PCA) algorithm to visualize the expression profiles related to ferroptosis subtypes. As shown in Figure 4E, it is found that the samples in the first dimension and the second dimension have a good aggregation form, which indicates that the classification method of ferroptosis subtypes is reasonable. Secondly, as shown in Figure 4F, by comparing the immune cells infiltrating the difference between ferroptosis subtypes, it was found that mast cells, immature B cells, eosinophil, activated B cells, activated dendritic cells, and immature dendritic cells were significantly infiltrated at a high level in Fer-1 compared with Fer-2. In summary, the expression profile of ferroptosis-related genes in LUAD is consistent with the prognosis profile, indicating that it was a viable method to classify ferroptosis subtypes.

The construction of LUAD ferroptosis-related lncRNA risk score model

To explore the expression of ferroptosis-related lncRNAs and their role in the evaluation of OS of LUAD, we used the Pearson correlation coefficient to identify lncRNAs that co-expressed with ferroptosis-related genes (*P*-value <0.001 and |R| > 0.5). As a result, 558 lncRNAs were screened which have a significant co-expression relationship with at least one ferroptosis gene (Supplementary Table S3). In this study, we constructed a risk score model of tumor immune cell infiltration based on the ferroptosis-related lncRNAs. Firstly, according to an approximate 2:1 ratio, the TCGA-LUAD overall set (n = 489) was divided into a training set (n = 326) and a test set (n = 163). In the training set, we displayed univariate Cox analysis to analyze 558 candidate lncRNAs. As shown in Figure 5A, 39 lncRNAs were retained with a meaningful threshold of p.value < 0.05 (Supplementary Table S4). For the convenience of clinical application, 13 lncRNAs were identified by LASSO regression (Figures 5B, C). Multivariate Cox regression was used to construct the lncRNA risk score model based on the 13 lncRNAs, The final 13-lncRNA gene signature formula is as follows:

$$\begin{aligned}\text{Risk score} = {} & (-0.041) \times \text{AC008278.2} + (-0.098) \times \text{AC093911.1} \\ & + (-0.132) \times \text{ADPGK} - \text{AS1} + (-0.060) \times \text{APTR} \\ & + (-0.074) \times \text{CBR3} - \text{AS1} + (-0.122) \times \text{CRNDE} \\ & + (-0.072) \times \text{LINC00324} + (-0.088) \times \text{LINC00526} \\ & + (-0.041) \times \text{LINC00892} + (-0.109) \times \text{LINC01352} \\ & + (0.454) \times \text{OGFRP1} + (-0.021) \times \text{PAN3} - \text{AS1} \\ & + (-0.088) \times \text{ZNF674} - \text{AS1}\end{aligned}$$

An R package called “ggrisk” was used to evaluate the power of the risk score model in predicting OS. Based on the optimal density gradient algorithm, patients were divided into high-risk and low-risk groups. The high-risk group had a higher mortality rate, as shown in Figure 5D. Kaplan-Meier survival analysis showed that the high-risk

10.3389/fgene.2023.1118273

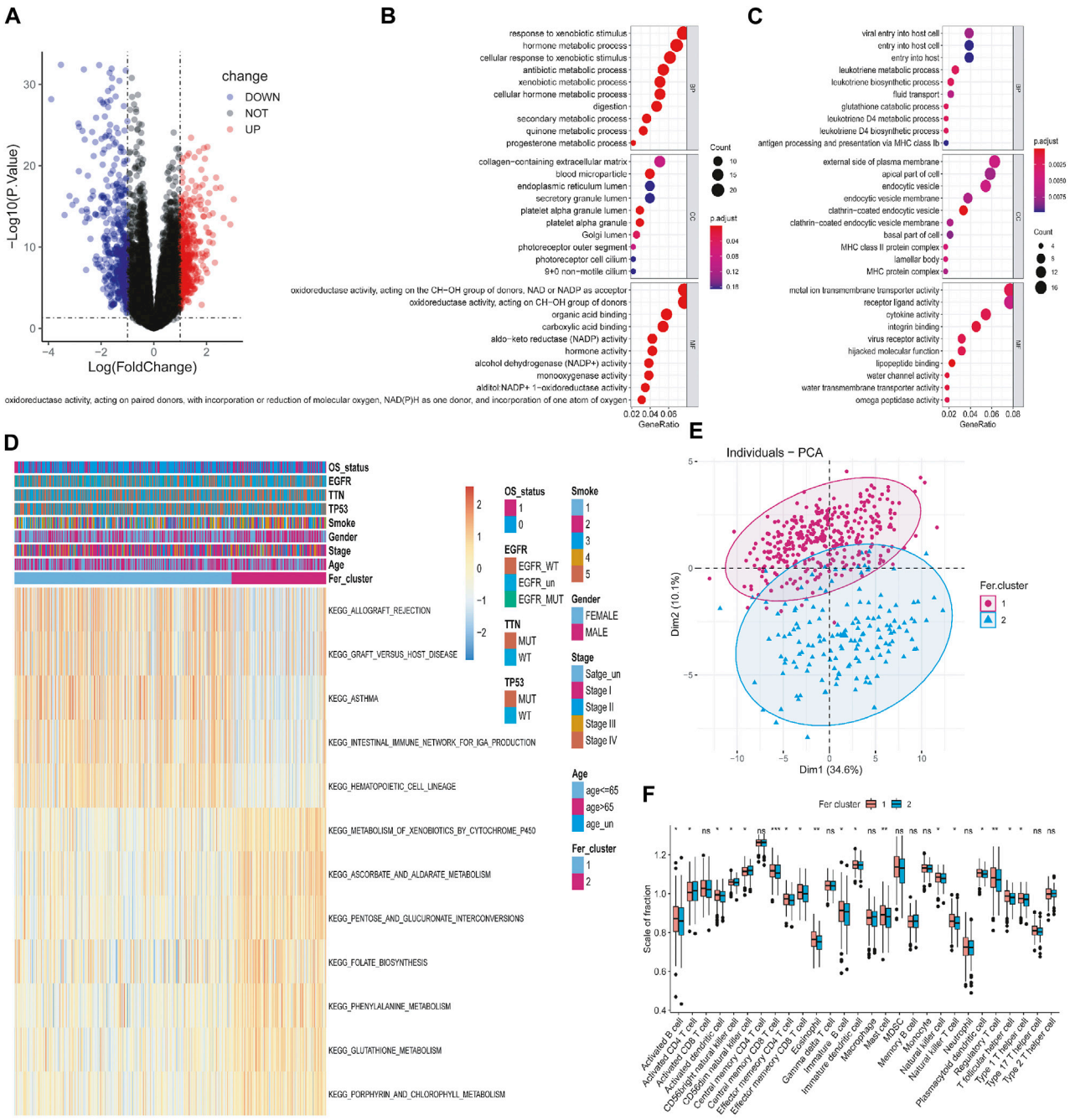


FIGURE 4
Identification and functional analysis of differentially expressed genes among different tumor ferroptosis subtypes. **(A)** Volcano map of differential expressed genes. **(B,C)** Bubble chart of GO enrichment analysis of upregulated and downregulated genes. **(D)** KEGG enrichment analysis of gene set. **(E)** PCA analysis of expression profile. **(F)** Tumor immune cell infiltration analysis of the gene dataset.

group has a significantly lower OS than the low-risk group (Figure 5E). The receiver operating characteristic curve (ROC) curves in Figure 5F indicated that the area under the curve (AUC) at TCGA-LUAD data sets was 0.7566, 0.7128, 0.7028 at 1-, 3-, and 5- years, respectively, indicating that the risk score is capable of predicting overall survival.

Subsequently, we used the test set and the overall set of TCGA-LUAD to access the predictive ability of risk score on OS. Based on the optimal density gradient algorithm, we assigned the patients to high-risk groups and low-risk groups. As shown in Figures 6A, D, the proportion of death samples in the high-risk group was relatively high. As Kaplan-Meier analyzed, the high-risk group has a significantly lower OS than the low-risk group (Figures 6B, E), suggesting that in the test set, the risk score model has a good predictive value. Its 1-, 3-, and 5- year AUC reached 0.6908, 0.6858, and 0.8546, respectively (Figure 6C). Similarly, in the overall dataset

10.3389/fgene.2023.1118273

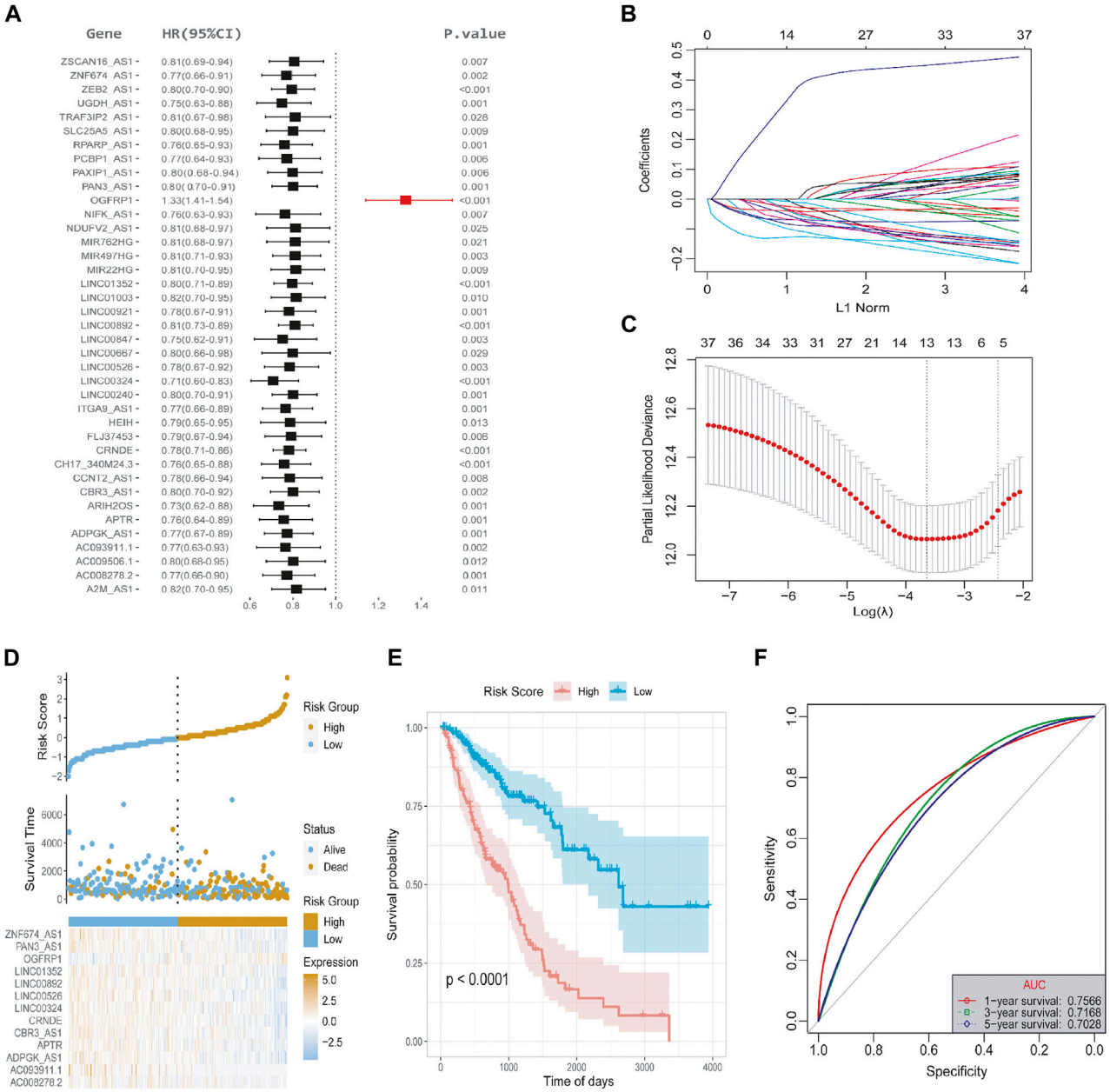


FIGURE 5
Screening of lncRNAs and the construction of risk score model. **(A)** Univariate Cox regression analysis was used to identify 558 candidate lncRNAs in the training set, and 39 lncRNAs were retained. The meaningful threshold was set as p-value < 0.05. **(B)** The changing trajectory of each independent variable. The horizontal axis represents the log value of the independent variable lambda, and the vertical axis represents the coefficient of the independent variable. **(C)** The confidence interval under each lambda. **(D)** The risk score distribution diagram; **(E)** Survival curve of LUAD patients with high- and low-risk scores. **(F)** ROC curve at 1-, 3-, and 5- years.

of TCGA-LUAD, the risk score model also has a good predictive value, with the 1-, 3-, and 5- year's AUC of 0.7400, 0.7125, and 0.7115, respectively (Figure 6F).

To evaluate the robustness of the risk score model in predicting OS of LUAD, the risk score model was validated by the external dataset GSE31210. By using the ggrisk software package in R, the samples were divided into high-risk and low-risk groups based on the optimal density gradient algorithm. We found that the proportion of death in the high-risk group was higher compared with the low-risk group (Figure 7A). In addition, Kaplan-Meier analysis showed that the OS of patients in the high-risk group was significantly lower than that in the low-risk group (Figure 7B). Therefore, the risk score model was also robust in predicting OS in the GSE31210 dataset (Figure 7C). The 1-, 3-, and 5- year's AUC was 0.7381, 0.7071, and 0.7296, respectively.

To better estimated the above bioinformatics results obtained from the public databases, we detected the levels of 13 key lncRNAs by using 5 paired LUAD tissues and corresponding adjacent non-

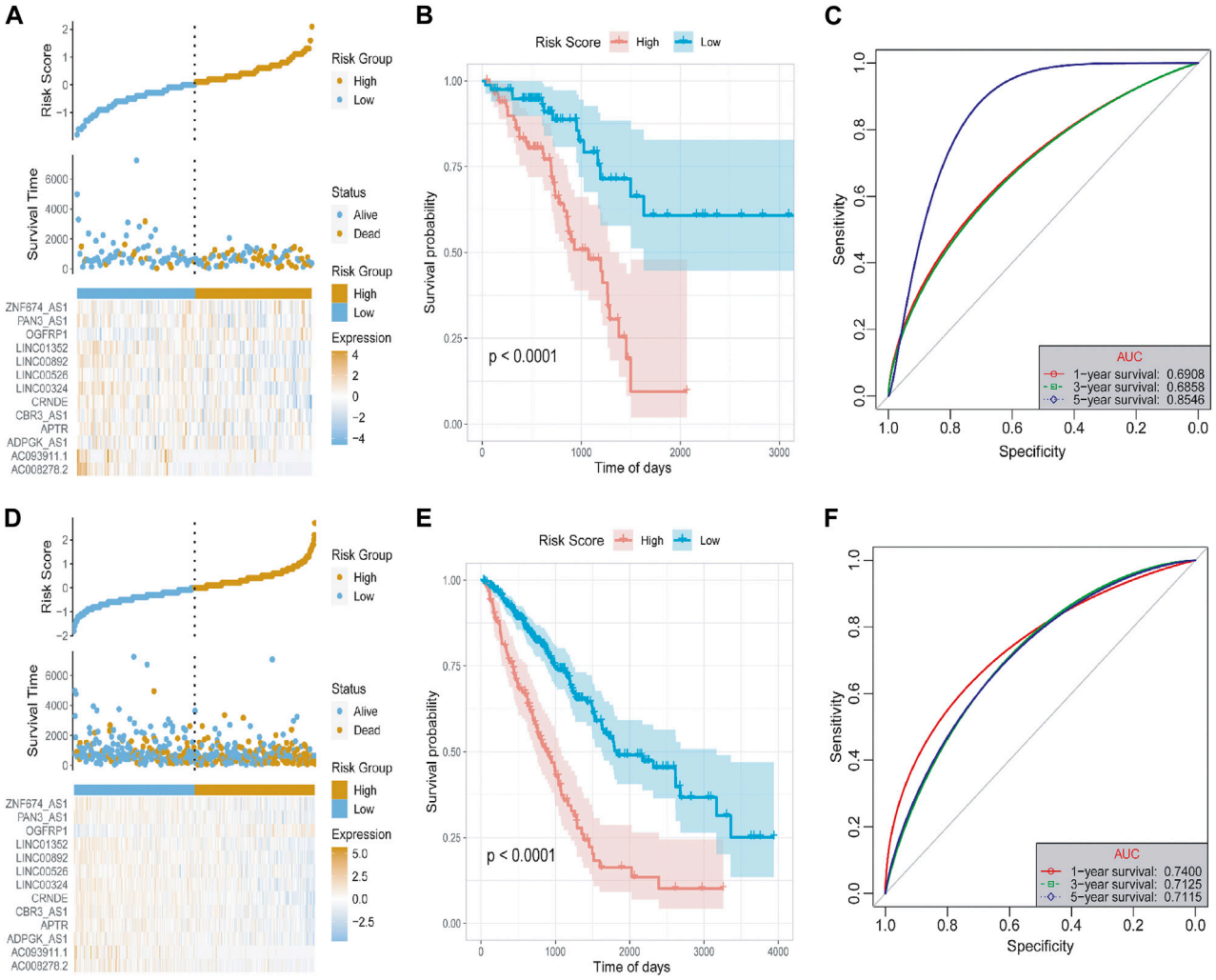


FIGURE 6
Test set and overall set to verify the risk model. **(A)** The distribution diagram of the risk score of the test set. **(B)** The survival curve of the test set. **(C)** The 1-, 3-, and 5- year's ROC curves of the test set. **(D)** Distribution chart of the risk score of the overall set. **(E)** Survival curve of the overall set. **(F)** The 1-, 3-, and 5- year's ROC curve of the overall set.

tumorous tissues. The quantitative RT-qPCR array in LUAD tissues shows enhanced expression of upregulated lncRNAs including APTR, CRNDE, LINC00324, OGFRP1, and LINC00526, as shown in Figure 7D. In contrast, LINC00892, LINC01352, PAN3-AS1, ZNF674-AS1, and ADPGK-AS1 have significantly diminished in non-tumorous tissues. Because of limited samples, we did not observe a significant difference in the expression of AC008278.2 and AC093911.1 in LUAD and non-tumorous tissues.

The relationship between risk score and clinical characteristics

It is necessary to clarify the relationship between tumor risk score and clinical characteristics, including age, smoke, and tumor grade. Firstly, multivariate Cox analysis determined that the lncRNA risk score was independent of other prognostic factors, such as age, gender, smoke and tumor stage, M-sage, N-stage, and T-stage (Figure 8A). Next, for the convenience of clinical evaluation, we construct a nomogram by using the risk score, T-stage, and N-stage (Figure 8B). The calibration curves of the nomogram 1-, 3-, and 5- years showed good stability. Notably, the ROC curve suggested that the predictive ability of the nomogram was higher than other factors (Figures 8C, D), with the AUC values reaching a high level above 0.75 (Figures 8E–G). Therefore, the lncRNA-based risk score was a relatively independent prognostic indicator in LUAD.

The relationship between lncRNA risk score and tumor mutation burden

Growing evidence suggests that tumor mutation burden (TMB) may determine the individual response to cancer immunotherapy (Bravaccini et al., 2021). It is important to explore the relationship between TMB and risk score to clarify the genetic characteristics of each ferroptosis subgroup. Correlation analysis (Figure 9A) showed that risk score was positively associated with TMB (R = 0.22, p =

10.3389/fgene.2023.1118273

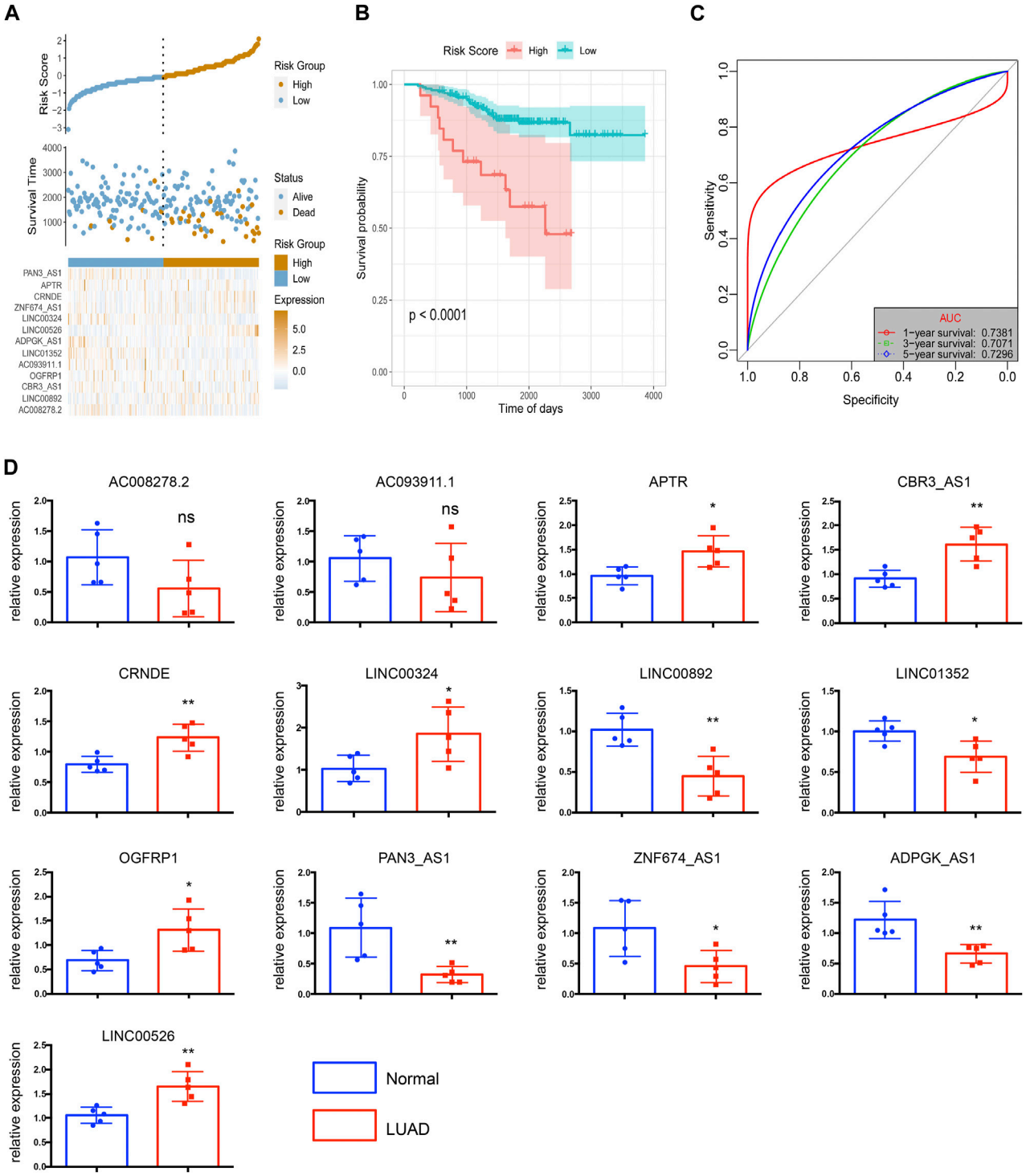


FIGURE 7
The risk score model was validated by the external dataset GSE31210. **(A)** The distribution of risk scores. **(B)** The survival probability was higher in the high-risk group compared with the low-risk group. **(C)** The 1-, 3-, and 5-year ROC curves of the external dataset. **(D)** Relative expressions of 13 key lncRNAs in LUAD tissues (LUAD) and corresponding adjacent non-tumorous tissues (normal). $N = 5$ in each group. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$, ns = no significance.

7.2×10^{-7}). By comparing the TMB of patients in subgroups (Figures 9B, C), we found that TMB in the high-risk score group was higher than in the low-risk score group. Furtherly, we used the Survminer package in R to calculate the optimal density gradient threshold associated with TMB and survival, and divided tumor samples in TCGA-LUAD into two groups with high- and low- TMB scores. As a result, we found a remarkable difference in survival between the two groups, as shown in Figure 9D.

10.3389/fgene.2023.1118273

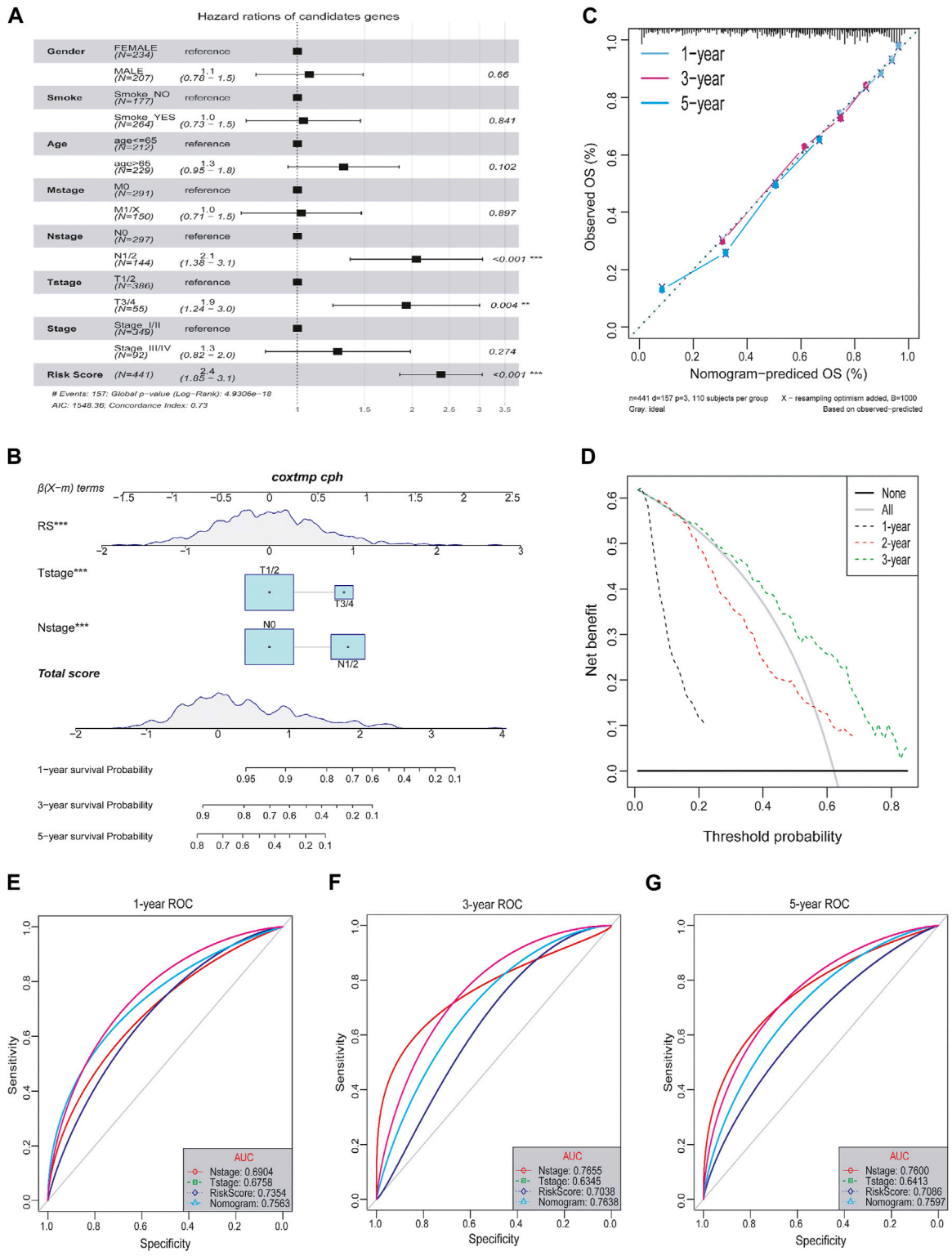


FIGURE 8
Relationship between tumor risk score and clinical characteristics. **(A)** Multivariate cox analysis of clinical characteristics and risk score. **(B)** Nomograms of clinical characteristics and risk score. **(C)** Calibration charts of nomograms in 1-, 3-, and 5-year. **(D)** DCA distribution map of nomograms in 1-, 3-, and 5-year. **(E–G)** ROC curves in 1-, 3-, and 5-year.

In addition, we quantified the distribution of somatic variation in LUAD driver genes between low-risk and high-risk score groups, meanwhile, the top 30 driver genes with the highest mutation frequency were compared (Figures 9E, F). By analyzing the mutation annotation files of the TCGA-LUAD cohort, we found that there were noteworthy differences in mutation profiles between

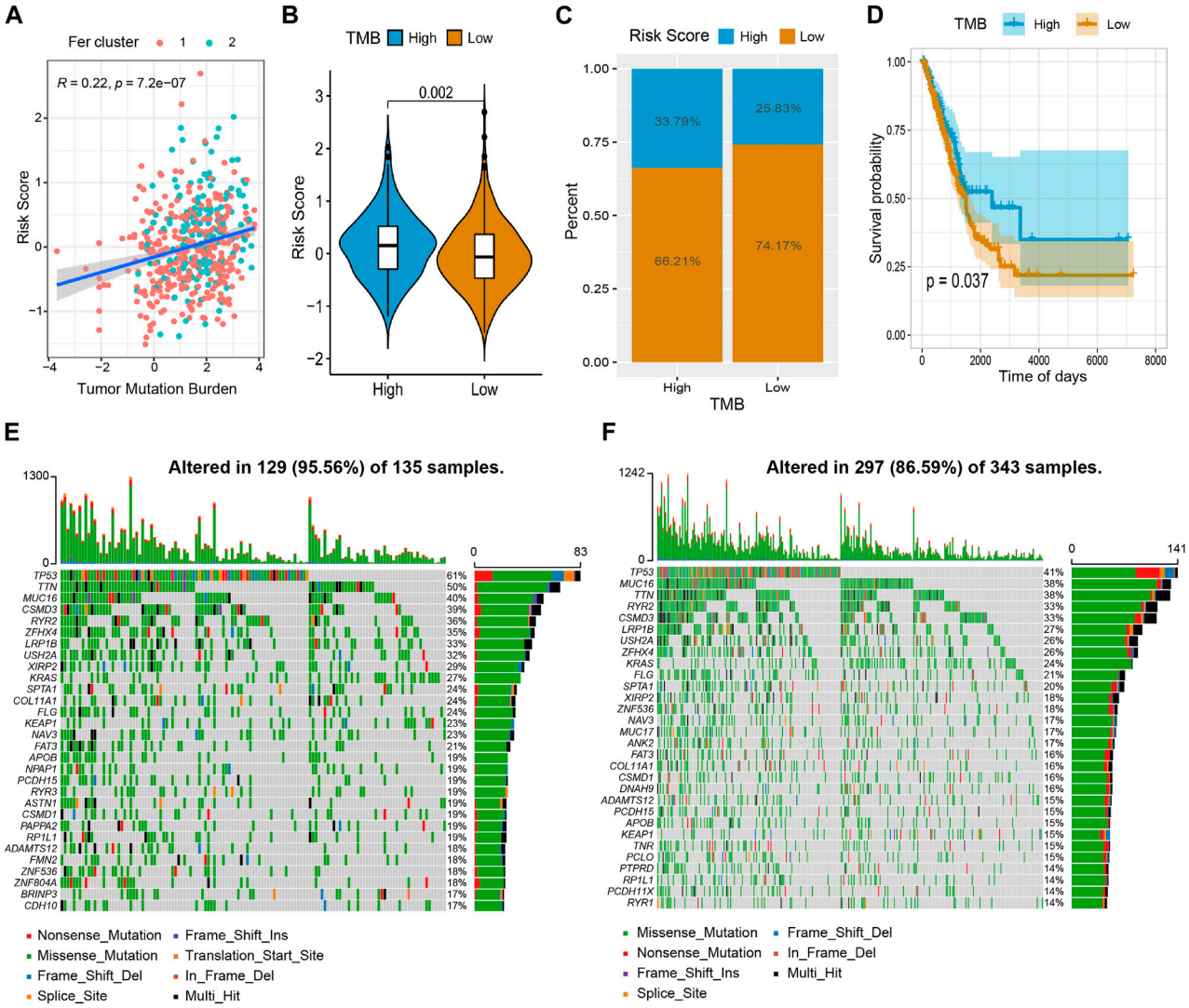


FIGURE 9
The relationship between tumor risk score and tumor mutation burden. **(A)** The risk score was positively correlated to TMB. **(B,C)** Violin chart and proportional distribution bar chart showed that TMB was higher in the high-risk score group than in the low-risk score group. **(D)** The survival curve showed patients in the high TMB have a better survival probability than the low TMB group. **(E)** High-risk group gene mutation waterfall chart. **(F)** Low-risk group gene mutation waterfall chart.

the low- and high-risk subgroups. These results may provide insight into understanding the mechanisms of LUAD ferroptosis status and gene mutations in immune checkpoints.

LncRNA risk score and immune cell infiltration (ICI)

To investigate the relationship between risk score and tumor immune microenvironment, we used GSEA to assess the state of infiltration of 28 different immune cells from the TCGA-LUAD dataset (Supplementary Table S5). As a whole, LUAD patients had a high infiltration ratio of $CD56^+$ dim natural killer cells, central memory $CD4^+$ T cells, central memory $CD8^+$ T cells, immature dendritic cells, myeloid-derived suppressor cell (MDSC), monocytes, natural killer cells, plasmacytoid dendritic cells, and regulatory T cells. LUAD tissues were less infiltrated by neutrophils, eosinophils, and type 17 T helpers (Figure 10A).

According to our hypotheses test, the infiltration level of active $CD4^+$ T cells was significantly higher in the group with high-risk score than in the group with low-risk score. In contrast, the infiltration of activated B cell, activated $CD8^+$ T cell, central memory $CD4^+$ T cell, eosinophil, γδ-T cell, immature B cell, immature dendritic cell, mast cell, monocyte, natural killer cells, T follicular helper cells, and type 2 T helper cells in high-risk score group were significantly lower than in the low-risk score group (Figure 10B).

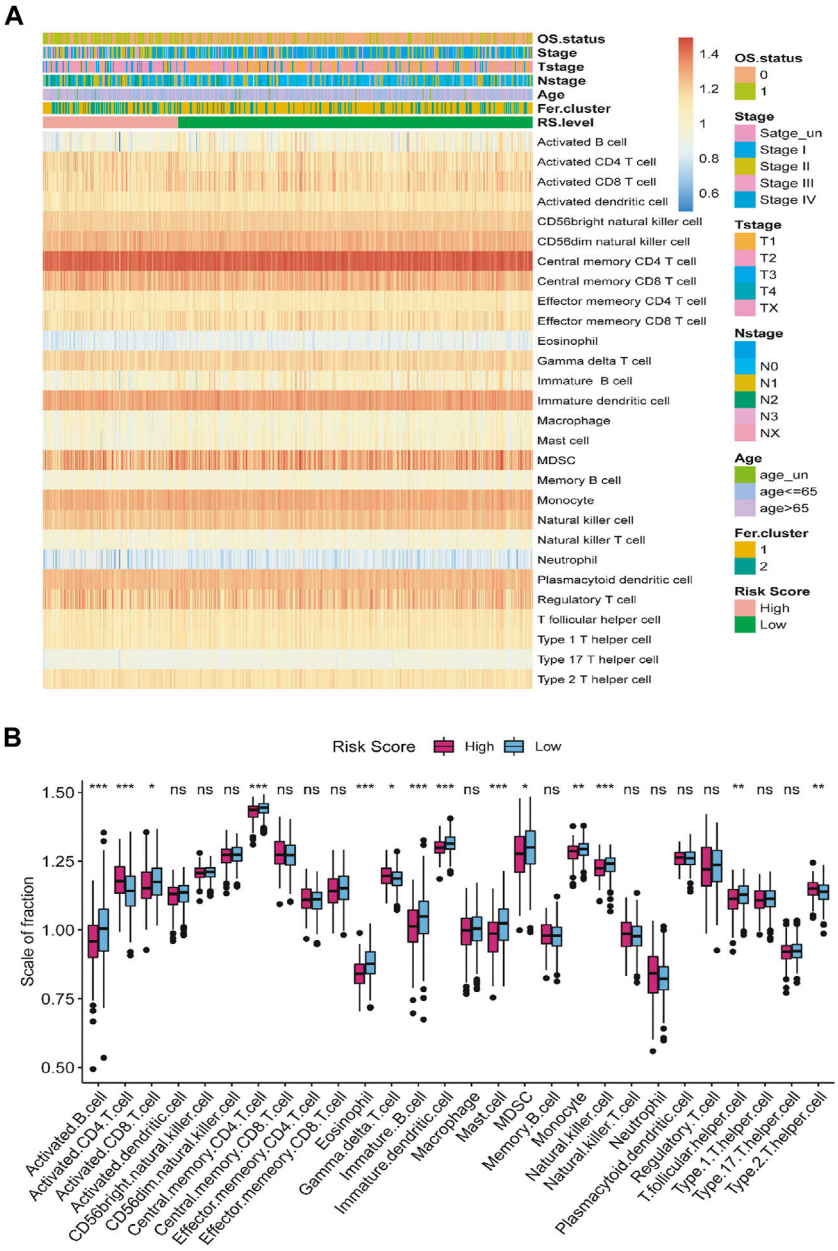


FIGURE 10
The relationship between tumor risk score and immune cells infiltration. **(A)** Heat map of the distribution of the immune cells infiltration. **(B)** Box plot of the difference in immune cells infiltration between high- and low-risk score group. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, ns = no significance.

The lncRNA risk score had a good predictive ability in evaluating the response of immunotherapy

To explore the predictive ability of risk score in predicting the benefit of immunotherapy, we analyzed the immunophenoscore (IPS) of samples from the TCIA database and the IMvigor210 cohort of immunotherapy patients (http://researchpub.gene.com/IMvigor210CoreBiologies). Multiple tumors can be predicted to respond to immunotherapy based on IPS, which can determine whether they are immunogenic. In Figures 11A–D, we found four types of low-risk score, namely, ips_ctla4_neg_pd1_neg, ips_ctla4_pos_pd1_neg, ips_ctla4_neg_pd1_pos, and ips_ctla4_pos_pd1_pos. IPS scores of patients in the low-risk group were significantly higher than those in the high-risk group, suggesting that immunotherapy was more likely to be effective. Patients who received anti-PD-L1 immunotherapy in the IMvigor210 cohort were divided into high- and low-risk groups. As a result, the group with low-risk scores showed a higher objective response to anti-PD-L1 therapy (Figure 11E). Moreover, patients with low-risk scores lived significantly longer than patients with high-risk scores (Figure 11F), and the increased risk in the IMvigor210 cohort correlated with the higher objective response rate (Figure 11G). In summary, these results suggest that

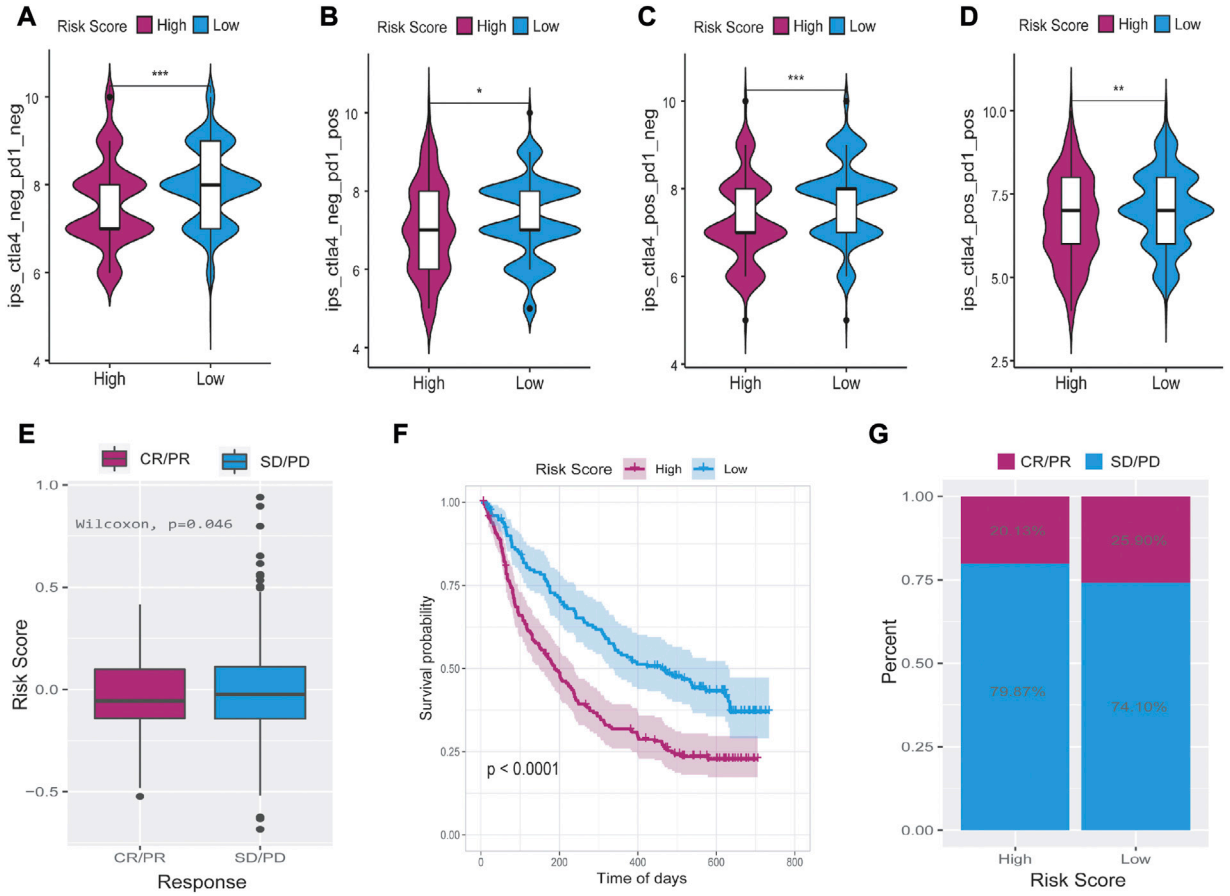


FIGURE 11
The relationship between tumor risk score and the response of immunotherapy. **(A–D)** Immunophenoscore was significantly higher in the low-risk group than in the high-risk score group. **(E–G)** Low-risk score patients who received anti-PD-L1 treatment had better responses to immunotherapy and significantly longer survival time than high-risk score patients. CR = complete response, PR = partial response, PD = progressive disease, SD = stable disease. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

the ferroptosis-related lncRNAs-based risk score may indicate the response to immunotherapy in LUAD.

Discussion

As the most common histological type of lung cancer, LUAD accounts for 40%–50% of all lung cancer cases (Bray et al., 2018). It severely affects human health and possesses both extremely high morbidity and mortality (Cheng et al., 2021). Despite great efforts having been made in developing novel treatments, however, LUAD still received a poor prognosis (Hirsch et al., 2017). In recent years, studies have demonstrated that ferroptosis is an important regulatory mechanism for tumor growth and is important for chemoradiotherapy and immunotherapy of tumors (Chen et al., 2021). In addition, lncRNAs have been a major focus of research into ferroptosis. However, the underlying relationship between ferroptosis-associated lncRNAs and the prognosis of LUAD patients remains quite limited. In this study, the expression profiles of ferroptosis-related genes in TCGA-LUAD dataset showed individual heterogeneity. Moreover, the expression profiles were correlated with the overall survival (OS) of LUAD patients. We also found that gene mutations could affect the expression of ferroptosis-related genes. Our results were then used to construct the risk score model with 13 ferroptosis-related lncRNAs. In univariate and multivariate Cox regression analysis, the risk score model was found to be a relatively independent prognostic indicator of the clinical features of LUAD patients. In addition, this study indicated the risk score model can well evaluate the benefit of LUAD patients receiving immunotherapy.

Liu et al. established the ferroptosis potential index (FPI) to reveal the functional roles of ferroptosis and found high FPI predicted poor prognosis in several tumors, highlighting the potential value of cancer classification based on ferroptosis-related genes expression (Liu et al., 2020). As a result of the expression of tumor ferroptosis-related genes and consensus clustering, we divided the samples into Fer-1 and Fer-2 groups. It was interesting to note that patients in the Fer-1 group had a median survival time of 898 days, significantly longer than Fer-2 group patients, who had a median survival time of 685 days. The differences in survival time between the two ferroptosis subtypes were probably determined by differences in biological functions and

signaling pathways as well as differences in immune cell infiltration. There seems to be a close relationship between Fer-1-enriched pathways and biological processes related to xenobiotic stimulus, hormone metabolism, and antibiotic metabolism. While Fer-2 was mostly enriched in viral entry into the host cell, leukotriene metabolism, and fluid transport. In addition, we discovered samples from Fer-1 were significantly more infiltrated with mast cells, immature B cells, eosinophils, activated B cells, activated dendritic cells, and immature dendritic cells than samples from Fer-2. In early-stage LUAD patients, mast cell abundance was associated with prolonged survival (Bao et al., 2020). Also, Han et al. found that upregulated glucose-6-phosphate isomerase (GPI) was associated with poorer survival, clinical stage, N stage, and primary therapy outcomes in LUAD. While GPI expression was negatively correlated with infiltrating levels of $CD8^+$ T cells, central memory T cells, dendritic cells, macrophages, mast cells, and eosinophils (Han et al., 2021), which is consistent with our study findings. Thus, this result showed the value of the classification of Fer-1 and Fer-2 in predicting the survival of LUAD patients.

A total of 13 ferroptosis-related key lncRNAs were identified by LASSO regression. What's more, a risk score model associated with tumor immune cell invasion was constructed based on these 13 lncRNAs. Interestingly, the risk score not only showed the ability to predict the overall survival of LUAD patients but was also associated with tumor mutation burden and evaluating the response of immunotherapy. Among the 13 key lncRNAs, LINC01352 is an important prognostic risk assessment factor for LUAD (Lu et al., 2021). By down-regulating miR-423-3p and inducing tumor suppressor protein p21, ZNF674-AS1 inhibits NSCLC growth. As a result, the low survival rate of NSCLC patients is significantly correlated with ZNF674-AS1 downregulation (Liu Y. et al., 2021). Linc00324 is over-expressed in a variety of cancer cell lines and tumoral tissues. Some researchers believe LINC00324 can be regarded as a promising candidate for the development of diagnostic and prognostic panels, what's more, can be used as a therapeutic target for a wide range of cancers (Ghafouri-Fard et al., 2022). A study suggested that Linc00324 overexpression accelerated the proliferation, migration, and invasion of LUAD cells by activating miR-615-5p/AKT1 axis (Zhang L. et al., 2021). CRNDE is a long non-coding RNA that has been demonstrated to be involved in multiple biological processes of different cancers as well as a potential diagnostic biomarker and prognostic predictor (Lu et al., 2020). Among the downstream targets of CRNDE, miR-641, CDK6, and miR-338-3p promote lung cancer cell proliferation and inhibit cell apoptosis (Fan et al., 2019; Jing et al., 2019). There have been reports that plncRNA-1, also known as CBR3-AS1, has different effects on different kinds of tumors. As an example, CBR3-AS1 modulates JNK1/MEK4 and enhances MAPK signaling by binding miR-25-3p competitively, suggesting it is a breast cancer prognosis marker (Zhang M. et al., 2021). Further, CBR3-AS1 is a poor prognostic molecule for osteosarcomas and colorectal cancer. Accordingly, high levels of CBR3-AS1 inhibit colorectal cancer metastasis by targeting the PI3K/Akt pathway (Zhang et al., 2018). Min Hou et al. found that CBR3-AS1 is associated with the prognoses of LUAD by activating the signal from the Wnt/β-catenin. (Hou et al., 2021). Despite its antisense lncRNA gene status, little is known about the role of ADPGK-AS1 in lung cancer. However, it has been reported to contribute to cervical, gastric, and colorectal cancer (Nagasaki et al., 2012; Jiang and Wang, 2020; Zhong Q. et al., 2021). ADPGK-AS1 has been shown to inhibit miR-205-5p downregulation in pancreatic cancer, which is negatively correlated with cancer cell proliferation, migration, and invasion, and positively correlated with apoptosis rates. The EMT process can thus be strongly induced *in vivo* by it (Song et al., 2018). Liu et al. demonstrated that downregulation of OGFRP1 inhibited the progression of NSCLC through miR-4640-5p/eIF5A axis (Liu X. et al., 2021). Furthermore, it has been reported that OGFRP1 is highly expressed in NSCLC tissues and significantly correlated with the prognosis of LUAD patients (Cui et al., 2021). As another core ferroptosis-related lncRNA noted in this study, APTR has been shown to reduce miR-132-3p and enhance YAP1 expression, which in turn promotes osteosarcoma progression (Guan et al., 2019). However, no study to date had demonstrated the relationship between APTR and lung cancer. It showed that AC008278.2 was a protective lncRNA was one of 19 genomic instability-related lncRNAs that correlated with somatic mutation pattern, immune microenvironment infiltration, immunotherapeutic response, drug sensitivity, and survival of NSCLC patients (Zhang et al., 2022). While as for PAN3-AS1 and AC093911.1, little has been studied in current diseases or molecular mechanisms. Further excavation is required to understand the role of these lncRNAs in lung cancer development.

TMB has emerged as a promising novel biomarker in predicting the prognosis and immune response in cancers, although the effect and the prognostic role of the TMB on outcomes varied dramatically across cancer types (Hellmann et al., 2018; Wang Z. M. et al., 2021). There are researches showed that higher TMB tends to form more new antigens, making tumors more immunogenic, improving clinical response to immunotherapy, and prolonging the overall survival (Lv et al., 2020; Wu et al., 2020). This is consistent with that patient in the high TMB scores group has better OS in our study. However, there are also studies showing the opposite. A study by Wang et al. found high TMB had a significantly poor prognosis in thymic epithelial tumors patients (Wang Z. M. et al., 2021). While Gao et al. discovered that higher TMB had a negative correlation with the prognosis of pancreatic ductal adenocarcinoma (Gao et al., 2020). The results of this study showed risk score had a modest positive correlation with TMB score, however, the risk score was negatively correlated with patients' OS, indicating an independent role of the risk score in predicting the response to immunotherapy in LUAD patients.

Harnessing an anti-tumor immune response has long been a fundamental strategy in cancer immunotherapy. According to the previously proposed tumor immunoediting hypothesis, tumor cells entering the immune escape phase can create an immunosuppressive state within the tumor microenvironment by subverting the same mechanisms that under normal conditions help regulate the immune response and prevent damage to healthy tissue (Carbone et al., 2015). In the last decade, higher objective response rates have been observed by targeting the PD-L1/PD-1 immune checkpoint pathway. This stems from distinct mechanisms of action that restore tumor-induced immunity deficiency selectively in a tumor microenvironment (TME) (Sanmamed and Chen, 2018). The therapeutic efficacy of these anti-PD1 therapies relies on endogenous tumor-antigen-specific T cells that are functionally held in check in the TME due to PD-L1 inhibitory signaling through PD-1. Anti-PD therapy results in the adaptive increase of functional T cells, which translates into tumor regression (Herbst et al., 2022). Until now, immunotherapy has shown considerable clinical success in the treatment response of many LUAD patients. Using T cells, monoclonal antibodies, or immune checkpoint inhibitors, immunotherapy stimulates the immune system to attack tumor cells

(Forde et al., 2018; Passiglia et al., 2018). What's more, growing studies have reported that the immune-related features of cancers such as the intensity of $CD4^+$ T cells and $CD8^+$ T cell infiltrates, macrophages, and natural killer (NK) cells, different B cell sub-populations were correlated with immunotherapeutic responsiveness in lung cancer (Stankovic et al., 2018). In this present study, we found a functional enrichment analysis that suggested that ferroptosis-related lncRNAs were mainly involved in immune pathways. Besides, immature dendritic cells, myeloid suppressive cells, monocytes, and regulatory T cells displayed a high level of LUAD. However, neutrophils, eosinophils, and type 17 T helper cells were the major low-level infiltrating cells. Additionally, our results revealed the relationship between immune cell infiltration (ICI) and the survival of LUAD patients. Based on these findings, these ferroptosis-related lncRNAs provide potential targets for combined treatments with immune checkpoint inhibitors.

There are some limitations of our study. Firstly, only data obtained from TCGA was used to construct a ferroptosis-related lncRNA prognostic model and to evaluate its validity. Secondly, the number of lung samples used on detecting the expression levels of the identified 13 key ferroptosis-associated lncRNAs was limited. Therefore, more work is needed to fully elucidate the mechanisms underlying the effects of ferroptosis-related lncRNAs on LUAD.

Conclusion

In conclusion, our study identified two ferroptosis subtypes to predict clinical outcomes and therapeutic responses in LUAD patients. The construction of a new risk score model with 13 ferroptosis-associated lncRNAs provides a candidate model for the evaluation of the LUAD prognosis. Our results demonstrate that LUAD patients in the high-risk score group presented worse OS, higher TMB, and lower immune activity. This study might contribute to the optimization of risk stratification for survival and personalized management of LUAD patients.

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 studies involving human participants were reviewed and approved by Ethics Review Board at Ren Ji Hospital, Shanghai Jiaotong University. The patients/participants provided their written informed consent to participate in this study.

Author contributions

KM conducted the formal analysis and wrote the original draft; RT performed the RT-qPCR experiments; YG contributed to writing, reviewing, and editing the article; YW participated in the software; ZZ conducted data curation; HH conceived and designed the study. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by the "Clinic Plus" Outstanding Project (No.2021ZYA010) from Shanghai Key Laboratory for Nucleic Acid Chemistry and Nanomedicine.

Conflict of interest

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

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2023.1118273/full#supplementary-material

SUPPLEMENTARY FIGURE S1
Waterfall chart of gene mutation in TCGA-LUAD dataset.

SUPPLEMENTARY FIGURE S2
The different expressions of 60 ferroptosis-related genes between TP53 mutation and wildtype in the TCGA-LUAD dataset.

SUPPLEMENTARY FIGURE S3
The different expressions of 60 ferroptosis-related genes between TTN mutation and wildtype in the TCGA-LUAD dataset.

SUPPLEMENTARY FIGURE S4
The correlation between the 60 ferroptosis-related genes in the TCGA-LUAD dataset.

SUPPLEMENTARY TABLE S1
Clinical information statistics of TCGA-LUAD data set.

SUPPLEMENTARY TABLE S2
Differentially expressed genes between the ferroptosis subtypes.

SUPPLEMENTARY TABLE S3
LncRNAs are co-expressed with ferroptosis-related genes.

SUPPLEMENTARY TABLE S4
Thirty-nine lncRNAs retained in the training set after the screening.

SUPPLEMENTARY TABLE S5
Infiltration state of 28 types of immune cells in the TCGA-LUAD dataset.

SUPPLEMENTARY TABLE S6
The primer sequences of the tested lncRNAs.

References

Ali, T., and Grote, P. (2020). Beyond the RNA-dependent function of LncRNA genes. *Elife* 9, e60583. doi:10.7554/eLife.60583

Bao, X., Shi, R., Zhao, T., and Wang, Y. (2020). Mast cell-based molecular subtypes and signature associated with clinical outcome in early-stage lung adenocarcinoma. *Mol. Oncol.* 14 (5), 917–932. doi:10.1002/1878-0261.12670

Bravaccini, S., Bronte, G., and Ulivi, P. (2021). Tmb in NSCLC: A broken dream? *Int. J. Mol. Sci.* 22 (12), 6536. doi:10.3390/ijms22126536

Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R. L., Torre, L. A., and Jemal, A. (2018). Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* 68 (6), 394–424. doi:10.3322/caac.21492

Cao, F., Fan, Y., Yu, Y., Yang, G., and Zhong, H. (2021a). Dissecting prognosis modules and biomarkers in glioblastoma based on weighted gene Co-expression network analysis. *Cancer Manag. Res.* 13, 5477–5489. doi:10.2147/cmar.s310346

Cao, F., Wang, C., Long, D., Deng, Y., Mao, K., and Zhong, H. (2021b). Network-based integrated analysis of transcriptomic studies in dissecting gene signatures for LPS-induced acute lung injury. *Inflammation* 44 (6), 2486–2498. doi:10.1007/s10753-021-01518-8

Carbone, D. P., Gandara, D. R., Antonia, S. J., Zielinski, C., and Paz-Ares, L. (2015). Non-small-cell lung cancer: Role of the immune system and potential for immunotherapy. *J. Thorac. Oncol.* 10 (7), 974–984. doi:10.1097/jto.0000000000000551

Chen, X., Kang, R., Kroemer, G., and Tang, D. (2021). Broadening horizons: The role of ferroptosis in cancer. *Nat. Rev. Clin. Oncol.* 18 (5), 280–296. doi:10.1038/s41571-020-00462-0

Cheng, Y., Hou, K., Wang, Y., Chen, Y., Zheng, X., Qi, J., et al. (2021). Identification of prognostic signature and gliclazide as candidate drugs in lung adenocarcinoma. *Front. Oncol.* 11, 665276. doi:10.3389/fonc.2021.665276

Cui, Y., Cui, Y., Gu, R., Liu, Y., Wang, X., Bi, L., et al. (2021). Identification of differentially expressed and prognostic lncRNAs for the construction of ceRNA networks in lung adenocarcinoma. *J. Oncol.* 2021, 2659550. doi:10.1155/2021/2659550

Fan, Y. F., Yu, Z. P., and Cui, X. Y. (2019). lncRNA colorectal neoplasia differentially expressed (CRNDE) promotes proliferation and inhibits apoptosis in non-small cell lung cancer cells by regulating the miR-641/CDK6 Axis. *Med. Sci. Monit.* 25, 2745–2755. doi:10.12659/msm.913420

Fei, X., Hu, C., Wang, X., Lu, C., Chen, H., Sun, B., et al. (2021). Construction of a ferroptosis-related long non-coding RNA prognostic signature and competing endogenous RNA network in lung adenocarcinoma. *Front. Cell Dev. Biol.* 9, 751490. doi:10.3389/fcell.2021.751490

Forde, P. M., Chaft, J. E., Smith, K. N., Anagnostou, V., Cottrell, T. R., Hellmann, M. D., et al. (2018). Neoadjuvant PD-1 blockade in resectable lung cancer. *N. Engl. J. Med.* 378 (21), 1976–1986. doi:10.1056/NEJMoa1716078

Gao, Y., Chen, S., Vafaei, S., and Zhong, X. (2020). Tumor-infiltrating immune cell signature predicts the prognosis and chemosensitivity of patients with pancreatic ductal adenocarcinoma. *Front. Oncol.* 10, 557638. doi:10.3389/fonc.2020.557638

Ghafouri-Fard, S., Safarzadeh, A., Hussen, B. M., Taheri, M., and Rashnoo, F. (2022). A concise review on the role of LINC00324 in different cancers. *Pathol. Res. Pract.* 240, 154192. doi:10.1016/j.prp.2022.154192

Gibb, E. A., Brown, C. J., and Lam, W. L. (2011). The functional role of long non-coding RNA in human carcinomas. *Mol. Cancer* 10, 38. doi:10.1186/1476-4598-10-38

Guan, H., Shang, G., Cui, Y., Liu, J., Sun, X., Cao, W., et al. (2019). Long noncoding RNA APTR contributes to osteosarcoma progression through repression of miR-132-3p and upregulation of yes-associated protein 1. *J. Cell Physiol.* 234 (6), 8998–9007. doi:10.1002/jcp.27572

Guo, X. H., Jiang, S. S., Zhang, L. L., Hu, J., Edelbek, D., Feng, Y. Q., et al. (2021a). Berberine exerts its antineoplastic effects by reversing the Warburg effect via downregulation of the Akt/mTOR/GLUT1 signaling pathway. *Oncol. Rep.* 46 (6), 253. doi:10.3892/or.2021.8204

Guo, Y., Qu, Z., Li, D., Bai, F., Xing, J., Ding, Q., et al. (2021b). Identification of a prognostic ferroptosis-related lncRNA signature in the tumor microenvironment of lung adenocarcinoma. *Cell Death Discov.* 7 (1), 190. doi:10.1038/s41420-021-00576-z

Han, J., Deng, X., Sun, R., Luo, M., Liang, M., Gu, B., et al. (2021). GPI is a prognostic biomarker and correlates with immune infiltrates in lung adenocarcinoma. *Front. Oncol.* 11, 752642. doi:10.3389/fonc.2021.752642

Hassannia, B., Vandenabeele, P., and Vanden Berghe, T. (2019). Targeting ferroptosis to iron out cancer. *Cancer Cell* 35 (6), 830–849. doi:10.1016/j.ccell.2019.04.002

Hellmann, M. D., Ciuleanu, T. E., Pluzanski, A., Lee, J. S., Otterson, G. A., Audigier-Valette, C., et al. (2018). Nivolumab plus ipilimumab in lung cancer with a high tumor mutational burden. *N. Engl. J. Med.* 378 (22), 2093–2104. doi:10.1056/NEJMoa1801946

Herbst, R. S., Wang, M., and Chen, L. (2022). When immunotherapy meets surgery in non-small cell lung cancer. *Cancer Cell* 40 (6), 603–605. doi:10.1016/j.ccell.2022.05.010

Hirsch, F. R., Scagliotti, G. V., Mulshine, J. L., Kwon, R., Curran, W. J., Jr., Wu, Y. L., et al. (2017). Lung cancer: Current therapies and new targeted treatments. *Lancet* 389 (10066), 299–311. doi:10.1016/s0140-6736(16)30958-8

Hou, M., Wu, N., and Yao, L. (2021). LncRNA CBR3-AS1 potentiates Wnt/β-catenin signaling to regulate lung adenocarcinoma cells proliferation, migration and invasion. *Cancer Cell Int.* 21 (1), 36. doi:10.1186/s12935-020-01685-y

Jiang, H. Y., and Wang, Z. J. (2020). ADPGK-AS1 promotes the progression of colorectal cancer via sponging miR-525 to upregulate FUT1. *Eur. Rev. Med. Pharmacol. Sci.* 24 (5), 2380–2386. doi:10.26355/eurrev_202003_20505

Jiang, L., Kon, N., Li, T., Wang, S. J., Su, T., Hibshoosh, H., et al. (2015). Ferroptosis as a p53-mediated activity during tumour suppression. *Nature* 520 (7545), 57–62. doi:10.1038/nature14344

Jing, H., Xia, H., Qian, M., and Lv, X. (2019). Long noncoding RNA CRNDE promotes non-small cell lung cancer progression via sponging microRNA-338-3p. *Biomed. Pharmacother.* 110, 825–833. doi:10.1016/j.biopha.2018.12.024

Latunde-Dada, G. O. (2017). Ferroptosis: Role of lipid peroxidation, iron and ferritinophagy. *Biochim. Biophys. Acta Gen. Subj.* 1861 (8), 1893–1900. doi:10.1016/j.bbagen.2017.05.019

Li, J., Cao, F., Yin, H. L., Huang, Z. J., Lin, Z. T., Mao, N., et al. (2020). Ferroptosis: Past, present and future. *Cell Death Dis.* 11 (2), 88. doi:10.1038/s41419-020-2298-2

Li, J., Meng, H., Bai, Y., and Wang, K. (2016). Regulation of lncRNA and its role in cancer metastasis. *Oncol. Res.* 23 (5), 205–217. doi:10.3727/096504016x14549667334007

Liang, R., Li, X., Li, W., Zhu, X., and Li, C. (2021). DNA methylation in lung cancer patients: Opening a "window of life" under precision medicine. *Biomed. Pharmacother.* 144, 112202. doi:10.1016/j.biopha.2021.112202

Liu, S. J., Dang, H. X., Lim, D. A., Feng, F. Y., and Maher, C. A. (2021a). Long noncoding RNAs in cancer metastasis. *Nat. Rev. Cancer* 21 (7), 446–460. doi:10.1038/s41568-021-00353-1

Liu, X., Niu, N., Li, P., Zhai, L., Xiao, K., Chen, W., et al. (2021b). LncRNA OGFRP1 acts as an oncogene in NSCLC via miR-4640-5p/eIF5A axis. *Cancer Cell Int.* 21 (1), 425. doi:10.1186/s12935-021-02115-3

Liu, Y., Huang, R., Xie, D., Lin, X., and Zheng, L. (2021c). ZNF674-AS1 antagonizes miR-423-3p to induce G0/G1 cell cycle arrest in non-small cell lung cancer cells. *Cell Mol. Biol. Lett.* 26 (1), 6. doi:10.1186/s11658-021-00247-y

Liu, Z., Zhao, Q., Zuo, Z. X., Yuan, S. Q., Yu, K., Zhang, Q., et al. (2020). Systematic analysis of the aberrances and functional implications of ferroptosis in cancer. *iScience* 23 (7), 101302. doi:10.1016/j.isci.2020.101302

Lu, L., Liu, L. P., Zhao, Q. Q., Gui, R., and Zhao, Q. Y. (2021). Identification of a ferroptosis-related LncRNA signature as a novel prognosis model for lung adenocarcinoma. *Front. Oncol.* 11, 675545. doi:10.3389/fonc.2021.675545

Lu, Y., Sha, H., Sun, X., Zhang, Y., Wu, Y., Zhang, J., et al. (2020). Crnde: An oncogenic long non-coding RNA in cancers. *Cancer Cell Int.* 20, 162. doi:10.1186/s12935-020-01246-3

Lv, J., Zhu, Y., Ji, A., Zhang, Q., and Liao, G. (2020). Mining TCGA database for tumor mutation burden and their clinical significance in bladder cancer. *Biosci. Rep.* 40 (4). doi:10.1042/bsr20194337

Matsui, M., and Corey, D. R. (2017). Non-coding RNAs as drug targets. *Nat. Rev. Drug Discov.* 16 (3), 167–179. doi:10.1038/nrd.2016.117

Mou, Y., Wang, J., Wu, J., He, D., Zhang, C., Duan, C., et al. (2019). Ferroptosis, a new form of cell death: Opportunities and challenges in cancer. *J. Hematol. Oncol.* 12 (1), 34. doi:10.1186/s13045-019-0720-y

Nagasaki, S., Nakamura, Y., Maekawa, T., Akahira, J., Miki, Y., Suzuki, T., et al. (2012). Immunohistochemical analysis of gastrin-releasing peptide receptor (GRPR) and possible regulation by estrogen receptor βcx in human prostate carcinoma. *Neoplasma* 59 (2), 224–232. doi:10.4149/neo_2012_029

Passiglia, F., Galvano, A., Rizzo, S., Incorvaia, L., Listì, A., Bazan, V., et al. (2018). Looking for the best immune-checkpoint inhibitor in pre-treated NSCLC patients: An indirect comparison between nivolumab, pembrolizumab and atezolizumab. *Int. J. Cancer* 142 (6), 1277–1284. doi:10.1002/ijc.31136

Relli, V., Trerotola, M., Guerra, E., and Alberti, S. (2019). Abandoning the notion of non-small cell lung cancer. *Trends Mol. Med.* 25 (7), 585–594. doi:10.1016/j.molmed.2019.04.012

Sanmamed, M. F., and Chen, L. (2018). A paradigm shift in cancer immunotherapy: From enhancement to normalization. *Cell* 175 (2), 313–326. doi:10.1016/j.cell.2018.09.035

Siegel, R. L., Miller, K. D., and Jemal, A. (2020). Cancer statistics, 2020. *CA Cancer J. Clin.* 70 (1), 7–30. doi:10.3322/caac.21590

Song, S., Yu, W., Lin, S., Zhang, M., Wang, T., Guo, S., et al. (2018). LncRNA ADPGK-AS1 promotes pancreatic cancer progression through activating ZEB1-mediated epithelial-mesenchymal transition. *Cancer Biol. Ther.* 19 (7), 573–583. doi:10.1080/15384047.2018.1423912

Spella, M., and Stathopoulos, G. T. (2021). Immune resistance in lung adenocarcinoma. *Cancers (Basel)* 13 (3), 384. doi:10.3390/cancers13030384

Stankovic, B., Bjørhovde, H. A. K., Skarshaug, R., Aamodt, H., Frafjord, A., Müller, E., et al. (2018). Immune cell composition in human non-small cell lung cancer. *Front. Immunol.* 9, 3101. doi:10.3389/fimmu.2018.03101

Sun, Z., Zeng, Y., Yuan, T., Chen, X., Wang, H., and Ma, X. (2022). Comprehensive analysis and reinforcement learning of hypoxic genes based on four machine learning algorithms for estimating the immune landscape, clinical outcomes, and therapeutic implications in patients with lung adenocarcinoma. *Front. Immunol.* 13, 906889. doi:10.3389/fimmu.2022.906889

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *J. Roy. Stat. Soc. B* 58, 267–288. doi:10.1111/j.2517-6161.1996.tb02080.x

Wang, M., Mao, C., Ouyang, L., Liu, Y., Lai, W., Liu, N., et al. (2020). Correction to: Long noncoding RNA LINC00336 inhibits ferroptosis in lung cancer by functioning as a competing endogenous RNA. *Cell Death Differ.* 27 (4), 1447. doi:10.1038/s41418-019-0394-6

Wang, Z. M., Xu, Q. R., Kaul, D., Ismail, M., and Badakhshi, H. (2021b). Significance of tumor mutation burden and immune infiltration in thymic epithelial tumors. *Thorac. Cancer* 12 (13), 1995–2006. doi:10.1111/1759-7714.14002

Wang, Z., Zhang, X., Tian, X., Yang, Y., Ma, L., Wang, J., et al. (2021a). CREB stimulates GPX4 transcription to inhibit ferroptosis in lung adenocarcinoma. *Oncol. Rep.* 45 (6), 88. doi:10.3892/or.2021.8039

Wu, Z., Wang, M., Liu, Q., Liu, Y., Zhu, K., Chen, L., et al. (2020). Identification of gene expression profiles and immune cell infiltration signatures between low and high tumor mutation burden groups in bladder cancer. *Int. J. Med. Sci.* 17 (1), 89–96. doi:10.7150/ijms.39056

Xu, L., Liu, Y., Chen, X., Zhong, H., and Wang, Y. (2023). Ferroptosis in life: To be or not to be. *Biomed. Pharmacother.* 159, 114241. doi:10.1016/j.biopha.2023.114241

Zhang, A., Yang, J., Ma, C., Li, F., and Luo, H. (2021a). Development and validation of a robust ferroptosis-related prognostic signature in lung adenocarcinoma. *Front. Cell Dev. Biol.* 9, 616271. doi:10.3389/fcell.2021.616271

Zhang, L., Zhang, K., Liu, S., Zhang, R., Yang, Y., Wang, Q., et al. (2021b). Identification of a ceRNA network in lung adenocarcinoma based on integration analysis of tumor-associated macrophage signature genes. *Front. Cell Dev. Biol.* 9, 629941. doi:10.3389/fcell.2021.629941

Zhang, M., Wang, Y., Jiang, L., Song, X., Zheng, A., Gao, H., et al. (2021c). LncRNA CBR3-AS1 regulates of breast cancer drug sensitivity as a competing endogenous RNA through the JNK1/MEK4-mediated MAPK signal pathway. *J. Exp. Clin. Cancer Res.* 40 (1), 41. doi:10.1186/s13046-021-01844-7

Zhang, Q., Liu, X., Chen, Z., and Zhang, S. (2022). Novel GIRlncRNA signature for predicting the clinical outcome and therapeutic response in NSCLC. *Front. Pharmacol.* 13, 937531. doi:10.3389/fphar.2022.937531

Zhang, Y., Meng, W., and Cui, H. (2018). LncRNA CBR3-AS1 predicts unfavorable prognosis and promotes tumorigenesis in osteosarcoma. *Biomed. Pharmacother.* 102, 169–174. doi:10.1016/j.biopha.2018.02.081

Zhong, H., Liu, S., Cao, F., Zhao, Y., Zhou, J., Tang, F., et al. (2021a). Dissecting tumor antigens and immune subtypes of glioma to develop mRNA vaccine. *Front. Immunol.* 12, 709986. doi:10.3389/fimmu.2021.709986

Zhong, Q., Lu, M., Yuan, W., Cui, Y., Ouyang, H., Fan, Y., et al. (2021b). Eight-lncRNA signature of cervical cancer were identified by integrating DNA methylation, copy number variation and transcriptome data. *J. Transl. Med.* 19 (1), 58. doi:10.1186/s12967-021-02705-9

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 09 May 2023
DOI 10.3389/fgene.2023.1169190

OPEN ACCESS

EDITED BY
Duo Liu,
Harbin Medical University Cancer Hospital, China

REVIEWED BY
Chen Li,
Free University of Berlin, Germany
Jin Zhang,
University of Mississippi Medical Center, United States

*CORRESPONDENCE
Chuanyong Liu,
liuwei199845@163.com
Fang Cao,
caof@zmu.edu.cn

†These authors have contributed equally to this work

RECEIVED 19 February 2023
ACCEPTED 04 April 2023
PUBLISHED 09 May 2023

CITATION
Shi G, He D, Xiao H, Liu Y, Liu C and Cao F (2023), Identification of the lncRNA–miRNA–mRNA regulatory network for middle cerebral artery occlusion-induced ischemic stroke. *Front. Genet.* 14:1169190. doi: 10.3389/fgene.2023.1169190

Identification of the lncRNA–miRNA–mRNA regulatory network for middle cerebral artery occlusion-induced ischemic stroke

Guixin Shi[1†], Dong He[1†], Hua Xiao[1], Yu'e Liu[2], Chuanyong Liu[3*] and Fang Cao[1*]

[1]Department of Neurosurgery, Affiliated Hospital of Zunyi Medical University, Zunyi, China, [2]Tongji University Cancer Center, Shanghai Tenth People's Hospital of Tongji University, School of Medicine, Tongji University, Shanghai, China, [3]Dingtao District Hospital of Traditional Chinese Medicine, Heze, China

Stroke known as a neurological disease has significant rates of disability and mortality. Middle cerebral artery occlusion (MCAO) models in rodents is crucial in stroke research to mimic human stroke. Building the mRNA and non-conding RNA network is essential for preventing MCAO-induced ischemic stroke occurrence. Herein, genome-wide mRNA, miRNA, and lncRNA expression profiles among the MCAO group at 3 h, 6 h, and 12 h after surgery and controls using high-throughput RNA sequencing. We detected differentially expressed mRNAs (DE-mRNAs), miRNAs (DE-miRNAs), and lncRNAs (DE-lncRNAs) between the MCAO and control groups. In addition, biological functional analyses were conducted, including GO/KEGG enrichment analysis, and protein-protein interaction analysis (PPI). GO analysis indicated that the DE-mRNAs were mainly enriched in several important biological processes as lipopolysaccharide, inflammatory response, and response to biotic stimulus. The PPI network analysis revealed that the 12 DE-mRNA target proteins showed more than 30˚ with other proteins, and the top three proteins with the highest node degree were Alb, IL-6, and TNF. In the DE-mRNAs, we found the mRNA of Gp6 and Elane interacting with two miRNAs (novel_miR_879 and novel_miR_528) and two lncRNAs (MSTRG.348134.3 and MSTRG.258402.19). As a result of this study, a new perspective can be gained into the molecular pathophysiology leading to the formation of MCAO. The mRNA-miRNA–lncRNA regulatory networks play an important role in MCAO-induced ischemic stroke pathogenesis and could be applied to the treatment and prevention of ischemic stroke in the future.

KEYWORDS

MCAO, mRNA, miRNA, lncRNA, ischemic stroke

1 Introduction

As an acute cerebral vascular disease, stroke is a leading cause of hospitalization for neurologic disease and long-term disability, with the majority (>80%) being ischemic stroke (an abrupt blockage of an artery) (Stegner et al., 2019; Shekhar et al., 2021; Fang et al., 2022). During an ischemic stroke, blood flow to the brain is decreased, triggering cascading events that eventually result in cell death (Mehta et al., 2007; Lai et al., 2011; Chen et al., 2022). Since

more than 80% of ischemic strokes occur in the territory of middle cerebral artery (MCA), experimental focal cerebral ischemia models including middle cerebral artery occlusion (MCAO) modelsare crucial in stroke research to mimic human stroke, which causes focal cerebral hypoperfusion and leads to ischemic stroke. The main therapeutic approach for MCAO-induced ischemic stroke is to restore or supply enough fresh blood flow to the brain by reperfusion. However, due to excitotoxicity and oxidative damage, side effects such as ischemic/reperfusion (I/R) may happen after cerebral ischemia and cerebral I/R, leading to brain injury (Lakhan et al., 2009; Enzmann et al., 2018). Therefore, there is still a substantial need for the development of therapeutic agents and the elucidation of the pathogenesis and molecular mechanisms of ischemic stroke to improve the functional outcome and prevent recurrence.

Previous research has demonstrated that gene mutation, DNA damage and oxidative stress contribute to the MCAO-induced ischemic stroke (Poudel et al., 2020; Liu et al., 2021; Liu et al., 2022). Recently, an increasing number of studies have indicated that non-coding RNAs (ncRNAs) cause functional alterations in ischemic stroke (Wang et al., 2018; Duan et al., 2019) Non-coding RNAs mainly including miRNAs, long non-coding RNAs (lncRNAs), and circular RNAs (circRNAs) participate in transcriptional regulation at different levels and play vital roles in various physiological and pathological processes(Panni et al., 2020). MiRNAs are small RNA molecules composed of 18–24 nucleotides, functioning in regulation of the expression of the target messenger RNAs (mRNAs), they bind to a short complementary sequence located at the 3' UTR region of the mRNA and lead to the target mRNA degradation(Ambros, 2004; Lu and Rothenberg, 2018). In an MCAO mice model, it is found that electroacupuncture (EA, a kind of stroke therapy) can reduce neuroinflammation and act in a neuroprotective manner by blocking the miR-223/NLRP3 pathway (Sha et al., 2019). In addition, lncRNAs are defined as transcripts of more than 200 nucleotides which cannot be translated to proteins, they play crucial roles in gene transcription(Ulitsky and Bartel, 2013; Kopp and Mendell, 2018). LncRNA MEG3 promotes cerebral I/R injury through increasing pyroptosis by targeting miR-485/AIM2 axis(Liang et al., 2020) (Chen et al., 2018). The RNA-RNA interplay emerges as a rising star in medical research now, the competing endogenous RNA (ceRNA) are the endogenous RNA transcripts that share the mutual miRNA response elements by competing for the same miRNA pools(Xiao et al., 2020). The ceRNA links different RNA species including mRNA, miRNA, and lncRNA together and enriching our understanding of molecular mechanism of various disease including ischemic stroke.

Thus, we hypothesized that the endogenous RNA regulatory network may be crucial in the emergence of MCAO-induced ischemic stroke. However, there is a lack of integrative analysis of mRNA-miRNA-lncRNA regulatory network in MCAO-induced Ischemic Stroke. To exploit the regulatory mechanism of ischemic stroke formation in this study, we conducted a high-throughput mRNA, miRNA, and lncRNA sequencing among the MCAO group at 3 h, 6h, and 12 h after surgery and the control group. We next explored differentially expressed lncRNA, miRNA and mRNA between these MCAO and control groups. Afterward, the mRNA-miRNA-lncRNA networks were developed to further understand the pathophysiology and underlying molecular mechanisms of MCAO-induced ischemic stroke.

2 Materials and methods

2.1 Animal experiments

A barrier system houses 10- to 12-week-old C57BL/6J mice with free access to food and water. Zunyi Medical University's Animal Experimentation Ethics Committee approved all procedures. Based on our previous research (Yao et al., 2016), we occluded the MCA in mice to perform transient MCAO. Anesthesia was induced in mice with sodium pentobarbital (40 mg/kg). There was exposure of the left common artery and the left external carotid artery. Through the right internal carotid artery, a 3 cm long MCAO suture (0.23/0.02 mm head/0.0104 mm body, RWD Life Science, MSMC23B104PK50) was inserted into the middle cerebral artery. Reperfusion was performed after 90 min of occlusion and removal of MCAO sutures. Similar operations were performed on sham control animals to expose the carotid arteries without occluding the middle cerebral artery.

2.2 Library preparation for ceRNA sequencing

Each sample contained 1.5 g RNA for removal of rRNA using the Ribo-Zero rRNA Removal Kit (Epicentre, Madison, WI, USA). In order to attribute sequences to each sample, NEBNextR UltraTM Directional RNA Library Prep Kit for IlluminaR (NEB, USA) was used to prepare sequencing libraries. NEBNext First Strand Synthesis Reaction Buffer (5X) was used to carry out fragmentation using divalent cations under elevated temperature. With the assistance of random hexamer primers and reverse transcriptase, first strand cDNA was synthesized (Mao et al., 2021a). DNA Polymerase I and RNase H were then used to synthesize second-strand cDNA. Exonuclease/polymerase activity was used to clear remaining overhangs. For hybridization, NEBNext Adaptor with hairpin loop structure was ligated after adenylation of 3′ ends of DNA fragments. By using AMPure XP beads (Beckman Coulter, Beverly, USA), library fragments were screened for preferred insert fragments that were 150–200 base pairs in length (Cao et al., 2020). Following that, 3 mL of USER Enzyme (NEB, USA) was used with cDNA that was adaptor ligated and size-selected before PCR. After that, PCR was performed with Phusion High-Fidelity DNA polymerase, Universal PCR primers, and Index(X) primers. Finally, the Agilent Bioanalyzer 2,100 and qPCR were used to assess library quality and purify PCR products (AMPure XP system).

2.3 Quality control

In-house Perl scripts were used to process raw reads in fastq format. We obtained clean data (clean reads) by removing adapter, ploy-N, and low quality reads from raw data. Sequences longer than 35 nt or smaller than 15 nt were removed from reads before

trimming and cleaning. Also, Q20, Q30, and GC-content of the clean data were calculated. By using Cutadapt software (v1.9.1), low-quality reads for base quality under 20 were eliminated. All the downstream analyses were based on clean data with high quality.

2.4 mRNA identification

Ensembl database was used to obtain gene annotations and reference genome files. HISAT2 software (v2.0.1) aligned the clean data with the reference genome (Kim et al., 2015). Expression analysis was performed by annotating and indexing the transcripts. RSEM software (v1.2.15) was used to estimate the expression of related genes using an annotated file as a reference gene set.

2.5 lncRNA identification

StringTie was used to assemble the transcriptome using reads mapped to the reference genome (Pertea et al., 2015). We annotated the assembled transcripts using the gffcompare program. We differentiated known lncRNAs from assembled transcripts if the sequencing species has lncRNA annotations. Putative lncRNAs were sought for using the remaining unknown transcripts. CPC/CNCI/Pfam/CPAT were combined to sort non-protein codingRNA candidates from putative protein-codingRNAs in the unknown transcripts. Aminimum lengths and exon number thresholds were used to filter out potential protein-coding RNAs. lncRNA candidates with a length of at least 200 nt and more than two exons were selected and further screened using the CPC/CNCI/Pfam/CPAT that differentiate protein-coding and non-coding genes. Additionally, lncRNAs were selected based on their types, such as lincRNA, intronic lncRNA, anti-sense lncRNA, and sense lncRNA.

2.6 miRNA identification

Utilizing Bowtie tools software, the Clean Reads database and Rfam database for sequence alignments, filtering out ribosomal RNAs (rRNAs), transfer RNAs (tRNAs), small nuclear RNAs (snRNAs), small nucleolar RNAs (snoRNAs) and repeats, and converting these to ncRNAs. By comparing the reads with known miRNAs from miRBase, we detected known miRNA and novel miRNA predictions(Ni et al., 2022). A novel miRNA secondary structure prediction was performed using Randfold tools software.

2.7 DE-ceRNA analysis

We used the R package 'DESeq2' to analyze differential expression of ceRNA as well as mRNA, miRNA, and lncRNA(Zhu X. et al., 2022). The ceRNA with p-value < 0.05 and $\log_2|\text{fold change}| \geq 2$ were identified as DE-ceRNAs. Volcano plots and heatmaps were constructed using the pheatmap package and ggplot2 R package(Mao et al., 2021b).

2.8 Functional analysis of DE-mRNA

DE-mRNA functions were revealed through Gene Ontology (GO) analysis, including biological processes (BP), cellular components (CC), and molecular functions (MF), along with KEGG pathway enrichment analysis. The R package 'clusterProfiler' was used to analyze GO enrichment and KEGG pathway enrichment analyses (Zhang et al., 2020). Those GO categories with adjust. p-value<0.05 and KEGG pathways with p-value<0.05 were considered as significantly enriched (Mao et al., 2020).

2.9 Protein–Protein interaction (PPI) analysis

DE-mRNA PPI network analysis was carried out based on STRING v11.5 database (https://string-db.org/) (Cao et al., 2021b). Direct interactions between DE-mRNAs were selected. In the PPI network, proteins associated with similar biological processes were grouped according to their attributes (event, betweenness, and degree) (Cao et al., 2021a).

2.10 Construction of lncRNA–miRNA–mRNA regulatory networks

TargetScan (http://www.targetscan.org/vert_72/) and miRanda algorithm were used to establish the lncRNA-miRNA-mRNA networks.

3 Results

3.1 DE-mRNA screening

We identified 31,011 mRNAs via high-throughput sequencing in at least one treatment (at 3h, 6h, and 12 h after MCAO surgery and control) (Supplementary Table S1). A total of 606, 685, and 799 DE-mRNAs were deemed to be significantly differentially expressed in 3 h vs. control, 6 h vs. control, and 12 h vs. control comparisons, respectively. Differences in mRNA expression in each treatment were evaluated by volcano plot analysis (Figure 1A). Among them, 188 common DE-mRNAs were overlapped in all three treatments which included 86 known mRNAs and 102 novel mRNAs detected in this study (Figure 1B). The common DE-mRNA expression patterns were visualized in a heatmap using hierarchical clustering analysis. (Figure 1C).

3.2 Functional analysis of DE-mRNA

GO and KEGG analyses were used to reveal the underlying roles of 188 common DE-mRNAs. The top enriched biological process (BP) GO terms included cellular response to lipopolysaccharide (GO: 0071222), inflammatory response (GO:0006954), and response to biotic stimulus (GO:0009607) (Figures 2A, B). The main enriched CCs included extracellular space (GO:0005615) and endomembrane system (GO:0012505) (Supplementary Figure S1). The main enriched

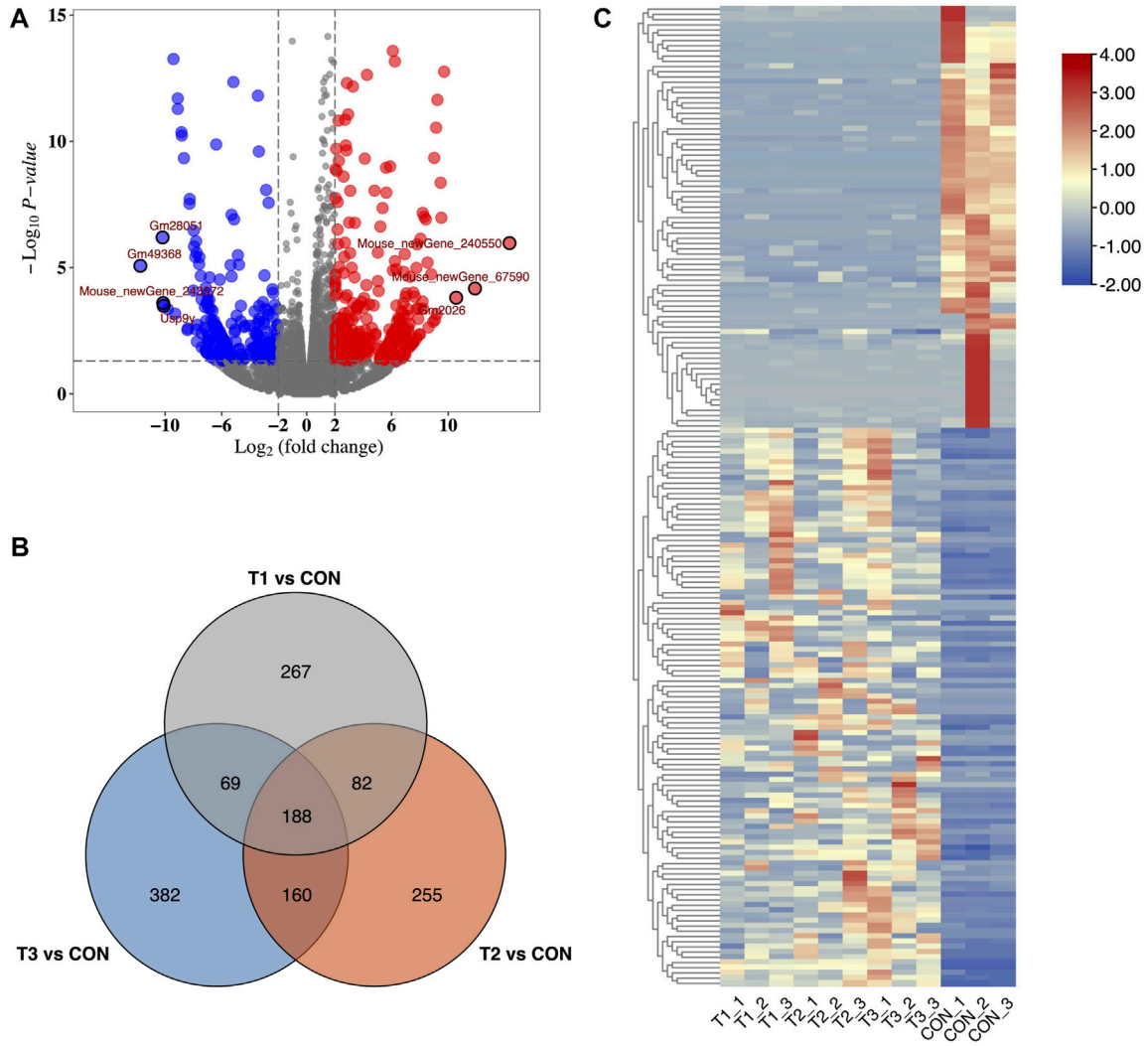


FIGURE 1
Identification of DE-mRNA in MCAO. **(A)** Volcano plots of DE-mRNAs for 3 h after MCAO treatments. Red and blue dots showed up- and down-regulated DE-mRNAs. **(B)** Heatmap of DE-mRNAs in MCAO and control groups. **(C)** Venn plot of potential DE-mRNA candidates with T1 vs. CON, T2 vs. CON, and T3 vs. CON. T1: 3 h after MCAO, T2: 6 h after MCAO, T3:12 h after MCAO, and CON: control.

MFs included cytokine activity (GO:0005125), chemokine activity (GO: 0008009), and endopeptidase inhibitor activity (GO:0004866) (Supplementary Figure S2). The top 25 KEGG pathways of common DE-mRNAs were shown in Figure 2C. The DE-mRNAs were primarily enriched in the following pathways: cytokine–cytokine receptor interaction, tumour necrosis factor (TNF) signaling pathway, JAK-STAT signaling pathway, and complement and coagulation cascades. The main KEGG annotation of common DE-mRNAs were enriched in metabolism global and overview maps, signal transduction, and immune system (Figure 2D).

3.3 PPI network analysis

The PPI network of DE-mRNA target proteins was established. As shown in Figure 3, this interaction network revealed that the 12 DE-mRNA target proteins showed more than 30° with other proteins. Among them, Albumin had the highest node degree (n = 55) in the network followed by IL-6 and TNF (with node degree of 42 and 41, separately).

3.4 Identification of DE-lncRNA and DE-miRNA

To explore the potential MCAO-related ncRNAs, we sequenced and detected lncRNA and miRNA expression profiles of MCAO using high-throughput sequencing (Figures 4A, B). We detected 4,509 lncRNAs in at least one treatment (at 3h, 6h, and 12 h after MCAO surgery and control) (Supplementary Table S2). A total of 803, 847, and 893 DE-lncRNAs were identified to be significantly differentially expressed with |log2(fold change) |≥2 and $p < 0.05$ in 3 h vs. control, 6 h vs. control, and 12 h vs. control comparisons, respectively. Among them, 186 common DE-lncRNAs were overlapped in all three treatments which included 26 known and 160 novel lncRNA (Figure 4C). Meanwhile,

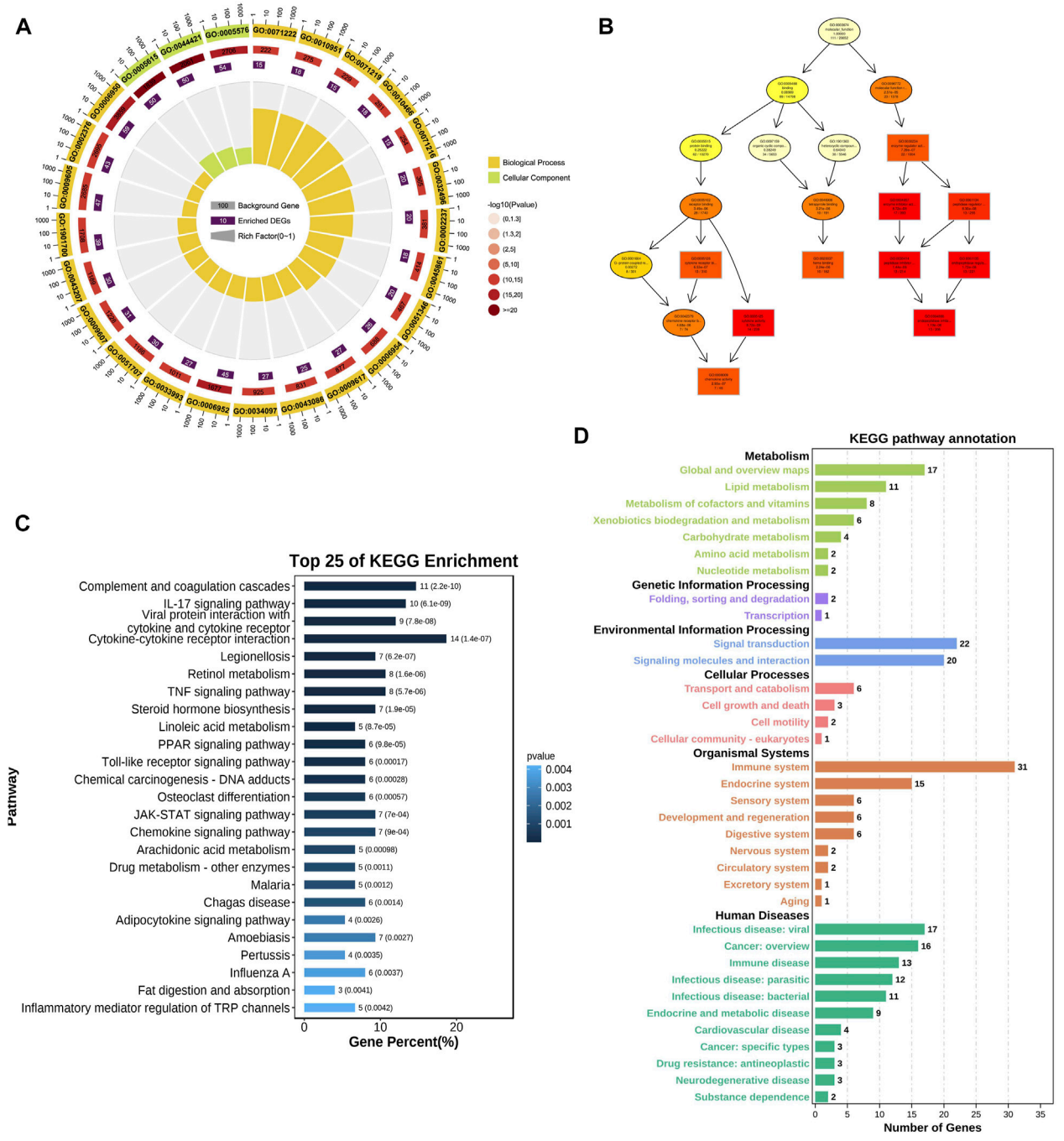


FIGURE 2
GO and KEGG enrichment analysis of DE-mRNA. **(A)** Top 25 BP GO enrichment analysis results of DE-mRNA. **(B)** GOplot of enriched GO terms **(C)** Top 25 KEGG pathway analysis results of DE-mRNA. **(D)** KEGG pathway annotation results of DE-mRNA. The color of bars represented different KEGG categories.

3,751 miRNAs were identified in at least one treatment (at 3h, 6h, and 12 h after MCAO surgery and control) (Supplementary Table S3). A total of 56, 142, and 155 DE-miRNAs were identified to be prominently differentially expressed in 3 h vs. control, 6 h vs. control, and 12 h vs. control comparisons, respectively. Twenty common DE-miRNA were overlapped in all three treatments which included 2 known miRNAs (mmu-miR-466m-3p and mmu-let-7j) and 18 novel miRNAs detected in this study (Figure 4D).

3.5 Construction of the lncRNA–miRNA–mRNA regulatory network

The lncRNA–miRNA–mRNA regulatory network was analyzed by miRanda algorithm and visualized in the network. We identified three DE-ceRNA networks consisting of DE-lncRNA–DE-miRNA–DE-mRNA. Among them, one network included one new Gene (Mouese_newGene_161645) and six novel miRNAs identified in

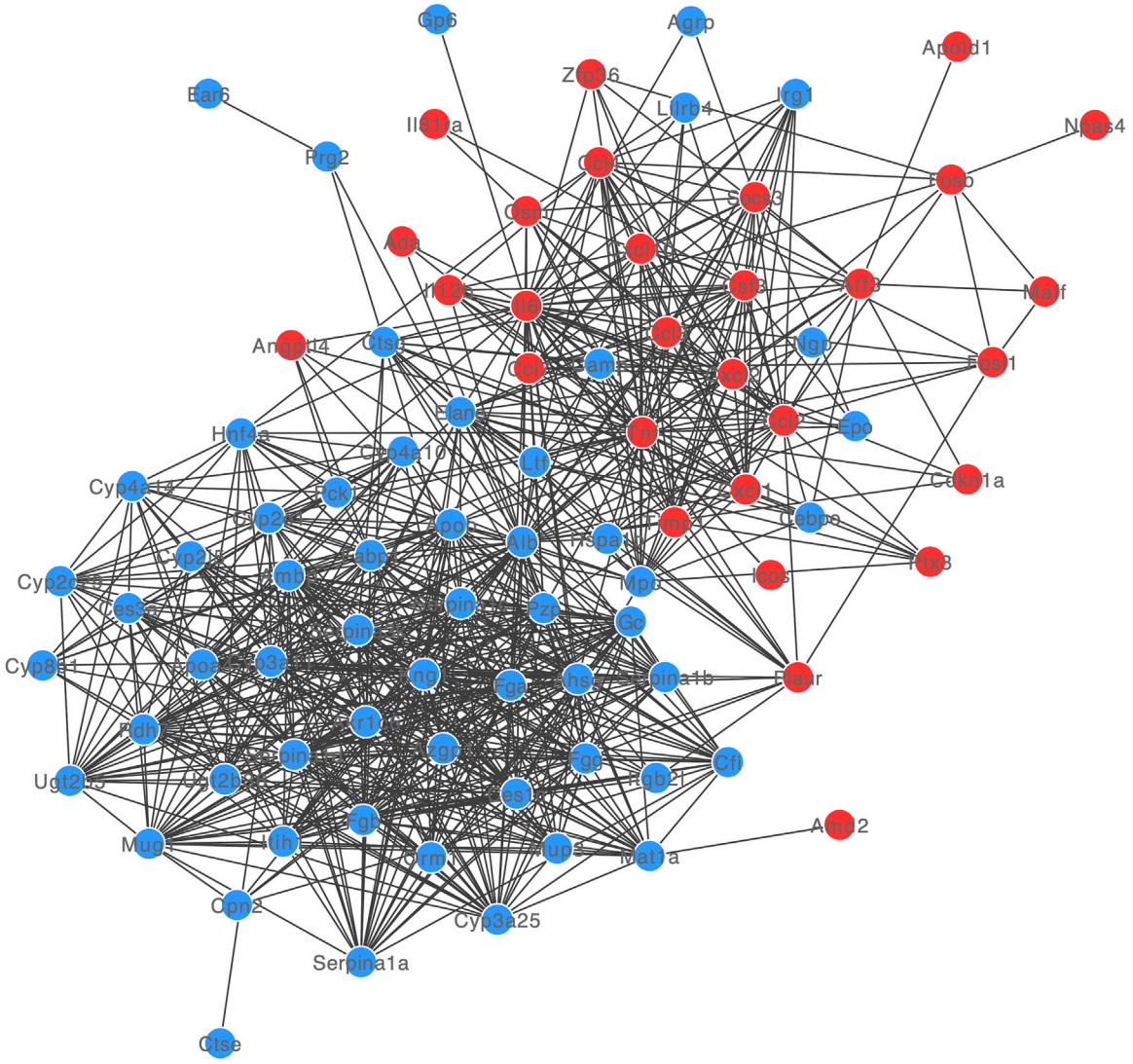


FIGURE 3
PPI network of DE-mRNAs. Red and blue dots showed up- and down-regulated DE-mRNAs.

this sutdy. Moreover, two networks contained known DE-ceRNA as Gp6 (mRNA)-novel_miR_879 (miRNA)-MSTRG.258402.19 (lncRNA) and Elane (mRNA)-novel_miR_528 (miRNA)-MSTRG.348134.31 (lncRNA) (Figure 5).

4 Discussion

An ischemic stroke occurs when a blood vessel in the neck or brain is blocked, which is currently a serious threat to human health and life. It is mainly characterized by the blockade of blood flow by an occlusive thrombus. The pathogenesis and molecular mechanism of MCAO remain elusive. Past studies have focused on the key mRNAs and ncRNAs which are essential in the development and curing of MCAO(Hou and Cheng, 2018; Duan et al., 2019). Hereby, we parsed out the mechanisms from another perspective, which focuses on the interactions among different RNAs.

We found DE-mRNAs and built their PPI network, among which Alb, IL-6, and Tnf have the highest node degree. Alb is the abbreviation of albumin, which is the main content of human blood. Several rodent models of ischemic stroke have demonstrated remarkable efficacy with human serum albumin (Huang and Xiao, 2021), it is a potential predictor of pneumonia after an acute ischemic stroke(Bath, 2013). Inflammation is a hallmark of stroke pathology. The interleukin-6 (Il6 or IL-6) is one of the interleukins and a major cytokine, produced by microglia under stroke (Zhu H. et al., 2022). Ischemic stroke pathogenesis is heavily influenced by IL-6 and has been demonstrated as the early marker of acute ischemic stroke(Li et al., 2022; Papadopoulos et al., 2022). Tumour necrosis factor (TNF or Tnf) is best known as a proinflammatory cytokine. After MCAO occurred, microglia produce IL-6 and TNF-α and the microglia-derived TNF-α mediate endothelial necroptosis aggravating blood brain-barrier disruption (Chen et al., 2019).

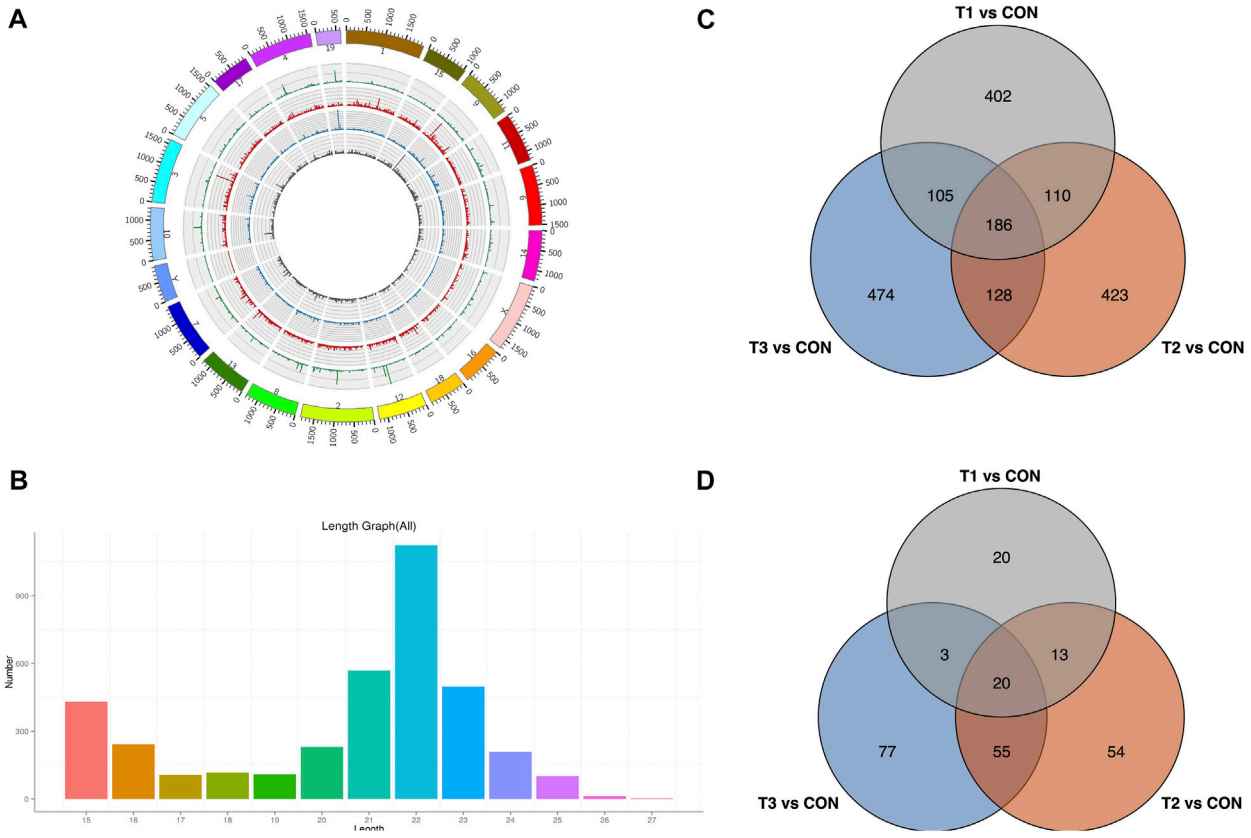


FIGURE 4
Identification of DE-lncRNA and DE-miRNA. **(A)** Chromosomal distribution of sequenced lncRNA. **(B)** Venn plot of potential DE-lncRNA candidates with T1 vs. CON, T2 vs. CON, and T3 vs. CON. **(C)** Length distribution of sequenced miRNA. **(D)** Venn plot of potential DE-miRNA candidates with T1 vs. CON, T2 vs. CON, and T3 vs. CON. T1: 3 h after MCAO, T2: 6 h after MCAO, T3:12 h after MCAO, and CON: control.

The top enriched BP GO analysis includes cellular response to lipopolysaccharide (GO:0071222) and inflammatory response (GO: 0006954). It has been verified that bacterial lipopolysaccharide is associated with stroke (Hakoupian et al., 2021). Poststroke cognitive impairment is common among stroke patients, and gut microbiota can contribute to it by influencing lipopolysaccharide levels (Wang et al., 2022). In recent years, the post-stroke immune response has emerged as a new breakthrough target in the treatment strategy for ischemic stroke pathobiology and outcome(Xu et al., 2020).

In this study, two ceRNA networks were identified, which contained known DE-ceRNA as Gp6(mRNA)-novel_miR_879(miRNA)-MSTRG.258402.19 (lncRNA) and Elane(mRNA)-novel_miR_528(miRNA)-MSTRG.348134.31 (lncRNA). Gp6, the membrane glycoprotein 6, is a platelet-specific collagen receptor exclusively expressed in the megakaryocytic lineage. Gp6 stimulates platelet activation and adhesion by interacting with collagen which is essential for thrombus formation, causing ischemic stroke (Jung and Moroi, 2008; Gao et al., 2021). A recent study has demonstrated that Gp6 contributes to atherosclerotic cerebral ischemic stroke development by activating the FYN-PKA-pPTK2/FAK1 signaling pathway, indicating the critical roles of Gp6 in ischemic stroke (Gu et al., 2021). Gp6 deficiency or inhibition suppresses thrombus formation and may still not cause a significant bleeding tendency, antibodies such as Abciximab, Glenzobimab and small molecule inhibitors as anti-thrombotic agents have been used for curing stroke clinically (Akkerhuis et al., 2001; Matsumoto et al., 2006; Wichaiyo et al., 2022). Therefore, Gp6 has been verified as a potential target for MCAO-induced ischemic stroke.

In addition, Elane is a neutrophil-expressed gene encoding Elastase. It is well known that neutrophils fight infection by phagocytosis and degranulation. Neutrophils release catalytically active Elane to kill cancer cells instead of non-cancer cells (Cui et al., 2021). Neutrophil extracellular traps (NETs) have been demonstrated to promote thrombus formation (Martinod and Wagner, 2014) and the elevated plasma NET biomarkers correlated with worse stroke outcomes (Denorme et al., 2022). Nets also impair revascularization and vascular remodeling after stroke (Kang et al., 2020). In acute ischemic stroke mice, the Elane inhibitor agaphelin reduces thrombosis, inflammation, and damage to the blood-brain barrier, indicating that Elane is a promising target for ischemic stroke (Leinweber et al., 2021).

In summary, we identified DE-mRNA, DE-miRNAs, and DE-lncRNAs in MCAO using high-throughput sequencing and then established an interplaying regulatory network. We found they work together to regulate MCAO-induced ischemic stroke development. Moreover, our study provided valuable and high-quality ceRNA sequencing data for ischemic stroke research as a reference. On the other hand, there were also

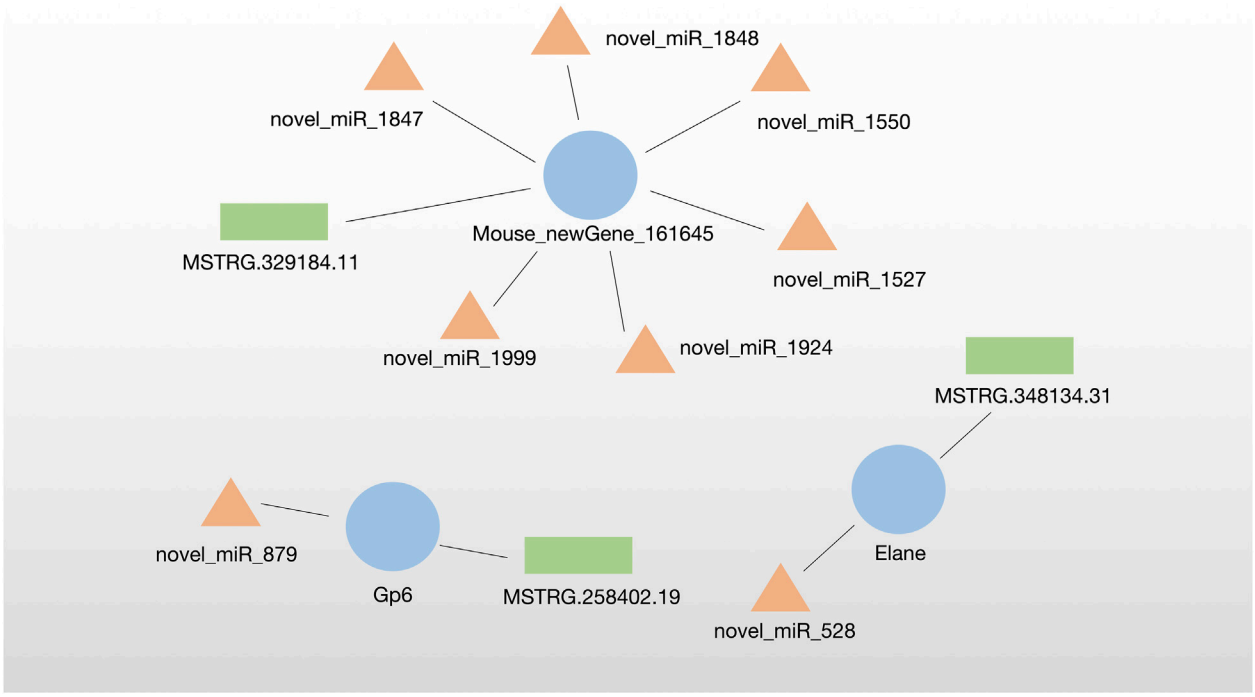


FIGURE 5
Identified three DE-ceRNA (lncRNA–miRNA–mRNA) regulatory networks.

several limitations. Firstly, the sample size was limited, which may affect the statistical power of subsequent analyses. Thus, we will further obtain a large number of sample data to validate our results. Secondly, the biological roles of the identified ceRNAs especially novel ceRNAs were not comprehensively explored. In the future, we will verify the regulatory functions through more bioinformatics analysis and experiments to facilitate elucidation of the underlying molecular mechanisms in ischemic stroke occurrence and development.

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/sra. Accession number: PRJNA957982.

Ethics statement

The animal study was reviewed and approved by Zunyi Medical University's Animal Experimentation Ethics Committee.

Author contributions

FC and CL designed the project. GS and YL wrote the manuscript. GS, DH, and HX completed the bioinformatic anlaysis. All authors approved the submitted manuscript.

Funding

The project was supported by Science and Technology Bureau of Zunyi (2018), 102, (2019)65. Young talents Found of Zunyi medical University (18zy-005). Guizhou Provincial Health Commission (gzwkj2022-089 and gzwkj2023-365).

Conflict of interest

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

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2023.1169190/full#supplementary-material

References

Akkerhuis, K. M., Deckers, J. W., Lincoff, A. M., Tcheng, J. E., Boersma, E., Anderson, K., et al. (2001). Risk of stroke associated with abciximab among patients undergoing percutaneous coronary intervention. *Jama-J Am. Med. Assoc.* 286 (1), 78–82. doi:10.1001/jama.286.1.78

Ambros, V. (2004). The functions of animal microRNAs. *Nature* 431 (7006), 350–355. doi:10.1038/nature02871

Bath, P. M. W. (2013). Albumin for hyperacute stroke: Another failed neuroprotectant. *Lancet Neurol.* 12 (11), 1036–1037. doi:10.1016/S1474-4422(13)70212-6

Cao, F., Fan, Y., Yu, Y., Yang, G., and Zhong, H. (2021a). Dissecting prognosis modules and biomarkers in glioblastoma based on weighted gene Co-expression network analysis. *Cancer Manag. Res.* 13, 5477–5489. doi:10.2147/CMAR.S310346

Cao, F., Guo, Y., Zhang, Q., Fan, Y., Liu, Q., Song, J., et al. (2020). Integration of transcriptome resequencing and quantitative proteomics analyses of collagenase VII-induced intracerebral hemorrhage in mice. *Front. Genet.* 11, 551065. doi:10.3389/fgene.2020.551065

Cao, F., Wang, C., Long, D., Deng, Y., Mao, K., and Zhong, H. (2021b). Network-based integrated analysis of transcriptomic studies in dissecting gene signatures for LPS-induced acute lung injury. *Inflammation* 44 (6), 2486–2498. doi:10.1007/s10753-021-01518-8

Chen, A. Q., Fang, Z., Chen, X. L., Yang, S., Zhou, Y. F., Mao, L., et al. (2019). Microglia-derived TNF-α mediates endothelial necroptosis aggravating blood brain-barrier disruption after ischemic stroke. *Cell Death Dis.* 10 (7), 487. doi:10.1038/s41419-019-1716-9

Chen, F. H., Zhang, L. X., Wang, E. W., Zhang, C. F., and Li, X. T. (2018). LncRNA GAS5 regulates ischemic stroke as a competing endogenous RNA for miR-137 to regulate the Notch1 signaling pathway. *Biochem. Bioph Res. Co.* 496 (1), 184–190. doi:10.1016/j.bbrc.2018.01.022

Chen, W., Li, C., Liang, W., Li, Y., Zou, Z., Xie, Y., et al. (2022). The roles of optogenetics and Technology in neurobiology: A review. *Front. Aging Neurosci.* 14, 867863. doi:10.3389/fnagi.2022.867863

Cui, C., Chakraborty, K., Tang, X. A., Zhou, G., Schoenfelt, K. Q., Becker, K. M., et al. (2021). Neutrophil elastase selectively kills cancer cells and attenuates tumorigenesis. *Cell* 184 (12), 3163–3177.e21. doi:10.1016/j.cell.2021.04.016

Denorme, F., Portier, I., Rustad, J. L., Cody, M. J., de Araujo, C. V., Hoki, C., et al. (2022). Neutrophil extracellular traps regulate ischemic stroke brain injury. *J. Clin. Invest.* 132 (10), e154225. doi:10.1172/JCI154225

Duan, X., Han, L., Peng, D., Peng, C., Xiao, L., Bao, Q., et al. (2019). Bioinformatics analysis of a long non-coding RNA and mRNA regulation network in rats with middle cerebral artery occlusion based on RNA sequencing. *Mol. Med. Rep.* 20 (1), 417–432. doi:10.3892/mmr.2019.10300

Enzmann, G., Kargaran, S., and Engelhardt, B. (2018). Ischemia-reperfusion injury in stroke: Impact of the brain barriers and brain immune privilege on neutrophil function. *Ther. Adv. Neurol. Disord.* 11, 1756286418794184. doi:10.1177/1756286418794184

Fang, X., Zhang, J., Roman, R. J., and Fan, F. (2022). From 1901 to 2022, how far are we from truly understanding the pathogenesis of age-related dementia? *Geroscience* 44 (3), 1879–1883. doi:10.1007/s11357-022-00591-7

Gao, G., Li, C., Fan, W., Zhang, M., Li, X., Chen, W., et al. (2021). Brilliant glycans and glycosylation: Seq and ye shall find. *Int. J. Biol. Macromol.* 189, 279–291. doi:10.1016/j.ijbiomac.2021.08.054

Gu, Y., Wu, Y., and Chen, L. (2021). GP6 promotes the development of cerebral ischemic stroke induced by atherosclerosis via the FYN-PKA-pPTK2/FAK1 signaling pathway. *Adv. Clin. Exp. Med.* 30 (8), 823–829. doi:10.17219/acem/135510

Hakoupian, M., Ferino, E., Jickling, G. C., Amini, H., Stamova, B., Ander, B. P., et al. (2021). Bacterial lipopolysaccharide is associated with stroke. *Sci. Rep-Uk* 11 (1), 6570. ARTN 6570. doi:10.1038/s41598-021-86083-8

Hou, X. X., and Cheng, H. (2018). Long non-coding RNA RMST silencing protects against middle cerebral artery occlusion (MCAO)-induced ischemic stroke. *Biochem. Biophys. Res. Commun.* 495 (4), 2602–2608. doi:10.1016/j.bbrc.2017.12.087

Huang, Y., and Xiao, Z. (2021). Albumin therapy for acute ischemic stroke: A meta-analysis. *Neurol. Sci.* 42 (7), 2713–2719. doi:10.1007/s10072-021-05244-9

Jung, S. M., and Moroi, M. (2008). Platelet glycoprotein VI. *Adv. Exp. Med. Biol.* 640, 53–63. doi:10.1007/978-0-387-09789-3_5

Kang, L., Yu, H., Yang, X., Zhu, Y., Bai, X., Wang, R., et al. (2020). Neutrophil extracellular traps released by neutrophils impair revascularization and vascular remodeling after stroke. *Nat. Commun.* 11 (1), 2488. doi:10.1038/s41467-020-16191-y

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

Kopp, F., and Mendell, J. T. (2018). Functional classification and experimental dissection of long noncoding RNAs. *Cell* 172 (3), 393–407. doi:10.1016/j.cell.2018.01.011

Lai, T. W., Shyu, W. C., and Wang, Y. T. (2011). Stroke intervention pathways: NMDA receptors and beyond. *Trends Mol. Med.* 17 (5), 266–275. doi:10.1016/j.molmed.2010.12.008

Lakhan, S. E., Kirchgessner, A., and Hofer, M. (2009). Inflammatory mechanisms in ischemic stroke: Therapeutic approaches. *J. Transl. Med.* 7 (97), 97. doi:10.1186/1479-5876-7-97

Leinweber, J., Mizurini, D. M., Francischetti, I. M. B., Fleischer, M., Hermann, D. M., Kleinschnitz, C., et al. (2021). Elastase inhibitor agaphelin protects from acute ischemic stroke in mice by reducing thrombosis, blood-brain barrier damage, and inflammation. *Brain Behav. Immun.* 93, 288–298. doi:10.1016/j.bbi.2020.12.027

Li, J., Lin, J., Pan, Y., Wang, M., Meng, X., Li, H., et al. (2022). Interleukin-6 and YKL-40 predicted recurrent stroke after ischemic stroke or TIA: Analysis of 6 inflammation biomarkers in a prospective cohort study. *J. Neuroinflammation* 19 (1), 131. doi:10.1186/s12974-022-02467-1

Liang, J., Wang, Q., Li, J. Q., Guo, T., and Yu, D. (2020). Long non-coding RNA MEG3 promotes cerebral ischemia-reperfusion injury through increasing pyroptosis by targeting miR-485/AIM2 axis. *Exp. Neurol.* 325, 113139. ARTN 113139. doi:10.1016/j.expneurol.2019.113139

Liu, H., Sun, S., and Liu, B. (2021). Smurf2 exerts neuroprotective effects on cerebral ischemic injury. *J. Biol. Chem.* 297 (2), 100537. doi:10.1016/j.jbc.2021.100537

Liu, Y., Chen, C., Wang, X., Sun, Y., Zhang, J., Chen, J., et al. (2022). An epigenetic role of mitochondria in cancer. *Cells* 11 (16), 2518. doi:10.3390/cells11162518

Lu, T. X., and Rothenberg, M. E. (2018). MicroRNA. *J. Allergy Clin. Immunol.* 141 (4), 1202–1207. doi:10.1016/j.jaci.2017.08.034

Mao, K., Geng, W., Liao, Y., Luo, P., Zhong, H., Ma, P., et al. (2020). Identification of robust genetic signatures associated with lipopolysaccharide-induced acute lung injury onset and astaxanthin therapeutic effects by integrative analysis of RNA sequencing data and GEO datasets. *Aging (Albany NY)* 12 (18), 18716–18740. doi:10.18632/aging.104042

Mao, K., Luo, P., Geng, W., Xu, J., Liao, Y., Zhong, H., et al. (2021a). An integrative transcriptomic and metabolomic study revealed that melatonin plays a protective role in chronic lung inflammation by reducing necroptosis. *Front. Immunol.* 12, 668002. doi:10.3389/fimmu.2021.668002

Mao, K., Tan, Q., Ma, Y., Wang, S., Zhong, H., Liao, Y., et al. (2021b). Proteomics of extracellular vesicles in plasma reveals the characteristics and residual traces of COVID-19 patients without underlying diseases after 3 months of recovery. *Cell Death Dis.* 12 (6), 541. doi:10.1038/s41419-021-03816-3

Martinod, K., and Wagner, D. D. (2014). Thrombosis: Tangled up in NETs. *Blood* 123 (18), 2768–2776. doi:10.1182/blood-2013-10-463646

Matsumoto, Y., Takizawa, H., Nakama, K., Gong, X., Yamada, Y., Tandon, N. N., et al. (2006). *Ex vivo* evaluation of anti-GPVI antibody in cynomolgus monkeys: Dissociation between anti-platelet aggregatory effect and bleeding time. *Thromb. Haemost.* 96 (2), 167–175. doi:10.1160/Th06-05-0266

Mehta, S. L., Manhas, N., and Raghubir, R. (2007). Molecular targets in cerebral ischemia for developing novel therapeutics. *Brain Res. Rev.* 54 (1), 34–66. doi:10.1016/j.brainresrev.2006.11.003

Ni, L., Tang, C., Wang, Y., Wan, J., Charles, M. G., Zhang, Z., et al. (2022). Construction of a miRNA-based nomogram model to predict the prognosis of endometrial cancer. *J. Pers. Med.* 12 (7), 1154. doi:10.3390/jpm12071154

Panni, S., Lovering, R. C., Porras, P., and Orchard, S. (2020). Non-coding RNA regulatory networks. *Biochim. Biophys. Acta Gene Regul. Mech.* 1863 (6), 194417. doi:10.1016/j.bbagrm.2019.194417

Papadopoulos, A., Palaiopanos, K., Bjorkbacka, H., Peters, A., de Lemos, J. A., Seshadri, S., et al. (2022). Circulating interleukin-6 levels and incident ischemic stroke: A systematic review and meta-analysis of prospective studies. *Neurology* 98 (10), e1002–e1012. doi:10.1212/WNL.0000000000013274

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

Poudel, S., Zeb, M., Kondapaneni, V., Gutlapalli, S. D., Choudhari, J., Sodiya, O. T., et al. (2020). Association of G20210A prothrombin gene mutation and cerebral ischemic stroke in young patients. *Cureus* 12 (12), e11984. doi:10.7759/cureus.11984

Sha, R., Zhang, B., Han, X., Peng, J., Zheng, C., Zhang, F., et al. (2019). Electroacupuncture alleviates ischemic brain injury by inhibiting the miR-223/NLRP3 pathway. *Med. Sci. Monit.* 25, 4723–4733. doi:10.12659/MSM.917213

Shekhar, S., Liu, Y., Wang, S., Zhang, H., Fang, X., Zhang, J., et al. (2021). Novel mechanistic insights and potential therapeutic impact of TRPC6 in neurovascular coupling and ischemic stroke. *Int. J. Mol. Sci.* 22 (4), 2074. doi:10.3390/ijms22042074

Stegner, D., Klaus, V., and Nieswandt, B. (2019). Platelets as modulators of cerebral ischemia/reperfusion injury. *Front. Immunol.* 10, 2505. ARTN 2505. doi:10.3389/fimmu.2019.02505

Ulitsky, I., and Bartel, D. P. (2013). lincRNAs: genomics, evolution, and mechanisms. *Cell* 154 (1), 26–46. doi:10.1016/j.cell.2013.06.020

Wang, H. D., Zhang, M. S., Li, J., Liang, J. H., Yang, M. J., Xia, G. H., et al. (2022). Gut microbiota is causally associated with poststroke cognitive impairment through lipopolysaccharide and butyrate. *J. Neuroinflamm* 19 (1), 76. ARTN 76. doi:10.1186/s12974-022-02435-9

Wang, S. W., Liu, Z., and Shi, Z. S. (2018). Non-coding RNA in acute ischemic stroke: Mechanisms, biomarkers and therapeutic targets. *Cell Transpl.* 27 (12), 1763–1777. doi:10.1177/0963689718806818

Wichaiyo, S., Parichatikanond, W., and Rattanavipanon, W. (2022). Glenzocimab: A gpVI (glycoprotein VI)-Targeted potential antiplatelet agent for the treatment of acute ischemic stroke. *Stroke* 53 (11), 3506–3513. doi:10.1161/STROKEAHA.122.039790

Xiao, H., Liang, S., and Wang, L. (2020). Competing endogenous RNA regulation in hematologic malignancies. *Clin. Chim. Acta* 509, 108–116. doi:10.1016/j.cca.2020.05.045

Xu, S., Lu, J., Shao, A., Zhang, J. H., and Zhang, J. (2020). Glial cells: Role of the immune response in ischemic stroke. *Front. Immunol.* 11, 294. doi:10.3389/fimmu.2020.00294

Yao, S., Tang, B., Li, G., Fan, R., and Cao, F. (2016). miR-455 inhibits neuronal cell death by targeting TRAF3 in cerebral ischemic stroke. *Neuropsychiatr. Dis. Treat.* 12, 3083–3092. doi:10.2147/NDT.S121183

Zhang, Q., Zhong, H., Fan, Y., Liu, Q., Song, J., Yao, S., et al. (2020). Immune and clinical features of CD96 expression in glioma by *in silico* analysis. *Front. Bioeng. Biotechnol.* 8, 592. doi:10.3389/fbioe.2020.00592

Zhu, H., Hu, S., Li, Y., Sun, Y., Xiong, X., Hu, X., et al. (2022a). Interleukins and ischemic stroke. *Front. Immunol.* 13, 828447. doi:10.3389/fimmu.2022.828447

Zhu, X., Tang, L., Mao, J., Hameed, Y., Zhang, J., Li, N., et al. (2022b). Decoding the mechanism behind the pathogenesis of the focal segmental glomerulosclerosis. *Comput. Math. Methods Med.* 2022, 1941038. doi:10.1155/2022/1941038

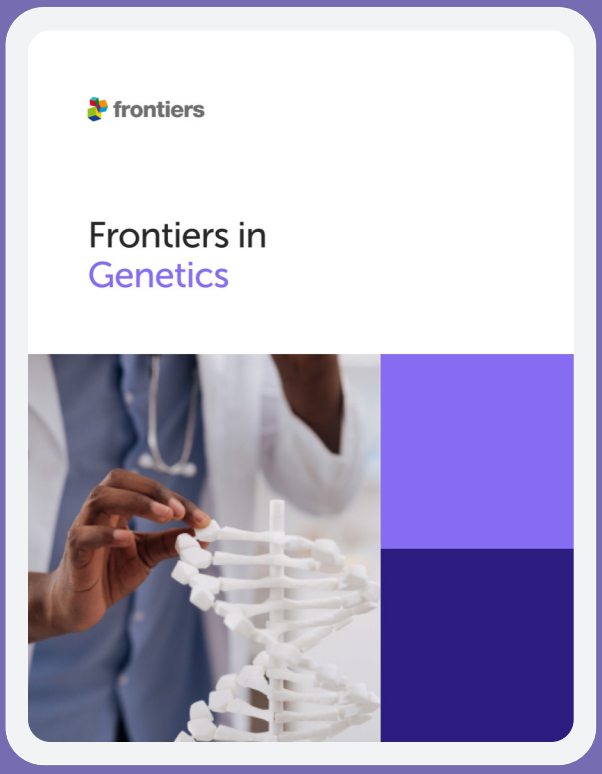
Frontiers in Genetics

Highlights genetic and genomic inquiry relating to all domains of life

The most cited genetics and heredity journal, which advances our understanding of genes from humans to plants and other model organisms. It highlights developments in the function and variability of the genome, and the use of genomic tools.

Discover the latest Research Topics

See more →



Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact