

Bioinformatics applied to neuroscience

Edited by

Thayne Kowalski, Mariana Recamonde-Mendoza,
Clévia Rosset and Jaqueline Bohrer Schuch

Published in

Frontiers in Genetics
Frontiers in Neuroscience



FRONTIERS EBOOK COPYRIGHT STATEMENT

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

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

Each article within this ebook, and the ebook itself, are published under the most recent version of the Creative Commons CC-BY licence. The version current at the date of publication of this ebook is CC-BY 4.0. If the CC-BY licence is updated, the licence granted by Frontiers is automatically updated to the new version.

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

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

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

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

ISSN 1664-8714
ISBN 978-2-8325-3639-1
DOI 10.3389/978-2-8325-3639-1

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

Bioinformatics applied to neuroscience

Topic editors

Thayne Kowalski — Federal University of Rio Grande do Sul, Brazil

Mariana Recamonde-Mendoza — Federal University of Rio Grande do Sul, Brazil

Clévia Rosset — Federal University of Rio Grande do Sul, Brazil

Jaqueline Bohrer Schuch — Federal University of Rio Grande do Sul, Brazil

Citation

Kowalski, T., Recamonde-Mendoza, M., Rosset, C., Schuch, J. B., eds. (2023).

Bioinformatics applied to neuroscience. Lausanne: Frontiers Media SA.

doi: 10.3389/978-2-8325-3639-1

Table of contents

- 05 **Editorial: Bioinformatics applied to neuroscience**
Clevia Rosset, Jaqueline Bohrer Schuch,
Mariana Recamonde-Mendoza and Thayne Woycinck Kowalski
- 08 **Association Between C-Reactive Protein and Risk of Amyotrophic Lateral Sclerosis: A Mendelian Randomization Study**
Yahui Zhu, Mao Li, Jinghong Zhang and Xusheng Huang
- 17 **Identification of Immune-Related Hub Genes in Parkinson's Disease**
Lin Chen, Yong Wang, Juan Huang, Binbin Hu and Wei Huang
- 28 **Identification of programmed cell death-related gene signature and associated regulatory axis in cerebral ischemia/reperfusion injury**
Jun Shu, Lu Yang, Wenshi Wei and Li Zhang
- 46 **A machine learning model for predicting patients with major depressive disorder: A study based on transcriptomic data**
Sitong Liu, Tong Lu, Qian Zhao, Bingbing Fu, Han Wang, Ginhong Li, Fan Yang, Juan Huang and Nan Lyu
- 57 **Development and validation of immune-based biomarkers and deep learning models for Alzheimer's disease**
Yijie He, Lin Cong, Qinfei He, Nianping Feng and Yun Wu
- 72 **Experimental verification and validation of immune biomarkers based on chromatin regulators in ischemic stroke**
Beibei Yu, Yunze Tian, Yongfeng Zhang, Boqiang Lv, Jianzhong Li and Shouping Gong
- 85 **Decoding competitive endogenous RNA regulatory network in postoperative cognitive dysfunction**
Wei Wang, Pengwei Huo, Lei Zhang, Gang Lv and Zhongyuan Xia
- 98 **Expression pattern and clinical value of Key RNA methylation modification regulators in ischemic stroke**
Xinyue Zhang, Yuanlin Wang, Beibei Dong, Yi Jiang, Dan Liu, Kelian Xie and Yonghao Yu
- 116 **Potential value of differentially expressed circular RNAs derived from circulating exosomes in the pathogenesis of rat spinal cord injury**
Chunfang Zan, Jianan Li, Fengsong Lin and Zengliang Wang
- 129 **Machine learning-based identification of the novel circRNAs circERBB2 and circCHST12 as potential biomarkers of intracerebral hemorrhage**
Congxia Bai, Xiaoyan Hao, Lei Zhou, Yingying Sun, Li Song, Fengjuan Wang, Liu Yang, Jiayun Liu and Jingzhou Chen

- 147 **Experimental verification and comprehensive analysis of m7G methylation regulators in the subcluster classification of ischemic stroke**
Yunze Tian, Beibei Yu, Boqiang Lv, Yongfeng Zhang, Longhui Fu, Shijie Yang, Jianzhong Li and Shouping Gong
- 159 **Multi-omics analysis revealed the role of CCT2 in the induction of autophagy in Alzheimer's disease**
Xueting Ma, Yuxin Feng, Xiangyu Quan, Bingyu Geng, Guodong Li, Xueqi Fu and Linlin Zeng
- 173 **Identification of immune signatures in Parkinson's disease based on co-expression networks**
Xiaolin Dong, Yanping Li, Qingyun Li, Wenhao Li and Gang Wu
- 184 **Exploring the bi-directional relationship and shared genes between depression and stroke *via* NHANES and bioinformatic analysis**
Zhanghuan Yang, Maokun He, Qian Zhang, Shifu Li, Hua Chen and Di Liao
- 202 **The use of multiple datasets to identify autophagy-related molecular mechanisms in intracerebral hemorrhage**
Yinggang Xiao, Yang Zhang, Cunjin Wang, Yali Ge, Ju Gao and Tianfeng Huang



OPEN ACCESS

EDITED AND REVIEWED BY

Sarah H. Elsea,
Baylor College of Medicine, United States

*CORRESPONDENCE

Mariana Recamonde-Mendoza
✉ mrmendoza@inf.ufrgs.br
Thayne Woycinck Kowalski
✉ tkowalski@hcupa.edu.br

[†]These authors have contributed equally to this work

RECEIVED 11 August 2023

ACCEPTED 08 September 2023

PUBLISHED 21 September 2023

CITATION

Rosset C, Schuch JB, Recamonde-Mendoza M and Kowalski TW (2023) Editorial: Bioinformatics applied to neuroscience.
Front. Neurosci. 17:1276346.
doi: 10.3389/fnins.2023.1276346

COPYRIGHT

© 2023 Rosset, Schuch, Recamonde-Mendoza and Kowalski. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](#). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Editorial: Bioinformatics applied to neuroscience

Clevia Rosset^{1,2,3†}, Jaqueline Bohrer Schuch^{4†},
Mariana Recamonde-Mendoza^{5,6*} and
Thayne Woycinck Kowalski^{2,6,7,8*}

¹Graduate Program in Medical Sciences (PPGCM), Universidade Federal do Rio Grande do Sul (UFRGS), Porto Alegre, Brazil, ²Laboratory of Genomic Medicine, Experimental Research Service, Hospital de Clínicas de Porto Alegre (HCPA), Porto Alegre, Brazil, ³Laboratory Research Unit, Experimental Research Service, Hospital de Clínicas de Porto Alegre (HCPA), Porto Alegre, Brazil, ⁴Graduate Program in Psychiatry and Behavioral Sciences, Department of Psychiatry, Universidade Federal do Rio Grande do Sul (UFRGS), Porto Alegre, Brazil, ⁵Graduate Program in Computer Science (PPGC), Institute of Informatics, Universidade Federal do Rio Grande do Sul (UFRGS), Porto Alegre, Brazil, ⁶Bioinformatics Core, Hospital de Clínicas de Porto Alegre (HCPA), Porto Alegre, Brazil, ⁷Graduate Program in Genetics and Molecular Biology (PPGBM), Department of Genetics, Universidade Federal do Rio Grande do Sul (UFRGS), Porto Alegre, Brazil, ⁸Laboratory Genetics Unit, Medical Genetics Service, Hospital de Clínicas de Porto Alegre (HCPA), Porto Alegre, Brazil

KEYWORDS

transcriptomics, systems biology, epigenetics, machine learning, neurodegenerative diseases, ischemia, autophagy, Mendelian randomization analysis

Editorial on the Research Topic

Bioinformatics applied to neuroscience

Genetics and molecular biology studies have revealed many new associations between the most diverse diseases, pointing to the involvement of a wide range of pathways. This increasing volume of data through interdisciplinary lenses and methods can provide insights into pathophysiologic mechanisms in several diseases. Therefore, new methods have arisen to analyze the data provided by these studies. Bioinformatics has emerged as a necessary discipline, revolutionizing this field and providing tools to jointly study DNA variants, gene expression, epigenetic marks, and biological networks. Biological public databases and computational tools are essential to develop data-driven approaches for understanding human diseases. This Research Topic contains 15 articles presenting current bioinformatics approaches applied to understanding the biological underpinnings of psychiatric and neurological conditions.

Dong et al. analyzed the association of immune and Parkinson's disease (PD)-related genes through protein-protein interaction networks, using co-expression data. Modular clustering analysis was also performed to identify central core genes, and the findings were validated by analyzing the expression of specific key genes. In summary, the authors were able to show strong correlations between immune- and PD-related genes, which could have the potential for diagnostic and therapeutic approaches. The integration between differential gene expression, weighted gene co-expression network analysis (WGCNA), and neurodegenerative disorders were also the topics studied by Chen et al. and He et al. The first study focused on PD, and their results also pointed to immune-related hub genes and the PD pathogenesis, which were presented as key modules in the WGCNA. He et al. focused on interconnected expression analysis and a machine learning model, through a random forest algorithm. They proposed seven hub genes that could confer the theoretical basis for studying biomarkers in Alzheimer's disease.

Other two studies used machine learning models to identify biomarkers in psychiatry and neurological conditions. [Liu et al.](#) proposed a diagnostic model of major depressive disorder (MDD) using machine learning. Data consisted of differentially expressed genes in MDD individuals and healthy controls. Most genes were involved in immune pathways and response to external stimuli. A robust diagnostic model was created through random forest and artificial neural network machine learning algorithms and potential driver genes were identified, as *C3AR1*, *BST2*, *TREM1*, *BTG3*, *LY6E*, and *IER5*. [Bai et al.](#) also used machine learning methods to identify potential biomarkers but in intracerebral hemorrhage (ICH). They examine the expression profiles of circRNAs in the peripheral blood and identify their potential functions using bioinformatic tools. Three circRNAs, named *hsa_circ_0005505*, *circERBB2* and *circCHST12*, were identified as promising biomarkers for ICH based on machine learning algorithms.

[Tian et al.](#) studied for the first time the epigenetic role of m7G-regulated genes in ischemic stroke. Based on previous studies on m7G, the expression of 34 m7G key regulatory genes was searched in the datasets from the Gene Expression Omnibus (GEO), including patients and controls. Two widely used machine learning algorithms, random forest (RF) and support vector machine (SVM) were subsequently used to identify eight key regulators of m7G. Five of the eight key regulators were significantly different in the middle cerebral artery occlusion model and quantitative polymerase chain reaction validations. In summary, their findings suggest that *EIF3D*, *CYFIP2*, *NCBP2*, *DCPS*, and *NUDT1* genes may serve as potential diagnostic biomarkers for ischemic stroke and could predict clinical risk. Another study by [Shu et al.](#) also used the middle cerebral artery occlusion model. They identified 15 differentially expressed genes related to the three types of programmed cell death (apoptosis, pyroptosis, and necroptosis) in transcriptome signatures of brain tissue samples from mice subjected to middle cerebral artery occlusion/reperfusion (MCAO/R). They conclude that these processes and the crosstalk among them might be involved in ischemic stroke and that the key nodes and regulatory axes identified in this study might play vital roles in regulating the above processes.

Ischemic stroke and epigenetics were also the object of study of other researchers. [Zhang et al.](#) evaluated the regulators of RNA methylation in ischemic stroke and suggested therapeutic targets by applying WGCNA followed by quantitative PCR analysis in an animal model; finally, the authors performed molecular docking to predict the interaction with the hub genes and drug molecules, which identified *GNMB* and chloroquine as potential targets. WGCNA was also applied in the study conducted by [Yu et al.](#) that evaluated chromatin regulators in ischemic stroke; the analysis pointed to four immune biomarkers (*DPF2*, *LMNB1*, *MLLT3*, and *JAK2*), which were validated through quantitative PCR and evaluated in regard to molecular docking. In a different approach, [Yang et al.](#) also evaluated immune biomarkers in ischemic stroke by studying the hub shared genes between the condition and major depressive disorder; differentially expressed genes were analyzed and it was found that innate immunity genes were upregulated whilst acquired immunity genes were downregulated.

[Zhu et al.](#) conducted a two-sample Mendelian randomization study to evaluate the association between C-Reactive Protein (CRP) levels and risk of Amyotrophic Lateral Sclerosis. Amyotrophic Lateral Sclerosis data were extracted from GWAS performed in people of European ancestry and included 20,806 cases and 59,804 controls. Six Mendelian randomization methods were selected, including the inverse variance weighted (IVW), weighted median, MR-Egger, MRPRESSO, simple mode, and weighted mode test. Fifty-seven independent SNPs were found to be associated with CRP. However, there was no significant causal relationship between genetically predicted CRP levels and disease risk ($OR = 1.123$, $95\% CI = 0.963-1.309$, $p = 0.139$), which is a relevant result since it excludes the possible association between CRP levels and Amyotrophic Lateral Sclerosis in the European population.

Autophagy was another important topic addressed. [Xiao et al.](#) investigated combinations of DGE with enrichment and systems biology analysis in intracerebral hemorrhage datasets. They suggested four autophagy-related genes, *IL1B*, *STAT3*, *NLRP3*, and *NOD2*, as key factors associated with intracerebral hemorrhage. In another study by [Ma et al.](#) autophagy was studied in the context of Alzheimer's disease through multi-omics analysis. They studied the novel ubiquitin-binding receptor, Chaperonin containing TCP1 subunit 2 (CCT2), which promotes aggrephagy, a process in which autophagy selectively degrades protein aggregates. All the datasets used in the study were obtained from the Gene Expression Omnibus database. The CCT2-high-associated genes screened by Pearson coefficients were enriched in protein folding, autophagy, and messenger RNA stability regulation pathways. The logistic prediction model screened in this study is a favorable candidate for predicting potential biological targets and small molecule inhibitors for Alzheimer's disease treatment.

Finally, two studies evaluated the role of circRNAs in neurologic conditions. The first study by [Wang et al.](#) evaluated competitive endogenous RNA regulatory networks in postoperative cognitive dysfunction (POCD). In their study, the authors extracted the transcriptomic signatures in the hippocampus of POCD mice derived from Gene Expression Omnibus (GEO) datasets in order to identify the circRNA, miRNA, and mRNA expression profiles of POCD mice compared with controls, respectively. A set of differentially expressed RNAs, including 119 circRNAs, 33 miRNAs, and 49 mRNAs were identified. Transcript validation by qPCR confirmed the enhanced expression of *circ_0001634*, *circ_0001345*, and *circ_0001493*. A regulatory network was constructed using circRNA-miRNA pairs and miRNA-mRNA pairs, resulting in a competing endogenous RNA regulatory network composed of three circRNAs, three miRNAs, and six mRNAs. The hub mRNAs in the network were further found to be involved in the hormone catabolic process and regulation of the canonical Wnt signaling pathway, revealing their crucial role in POCD.

The second circRNA study evaluated the potential value of differentially expressed circular RNAs derived from circulating exosomes in the pathogenesis of rat spinal cord injury (SCI). SCI remains a catastrophically injured condition for humans, thereby bringing severe social and economic burdens. The study by [Zan et al.](#) analyzed differentially expressed circRNAs derived from circulating exosomes in SCI rats in comparison with the control

rats. Subsequently, functional enrichment analyses including Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway and Gene Ontology (GO) were performed to evaluate the possible biological functions of upregulated as well as downregulated circRNAs involved in SCI. Five upregulated circulating circRNAs including and five downregulated circulating circRNAs were verified through reverse transcription-polymerase chain reaction. They also constructed a circRNA-miRNA-mRNA gene interaction network to predict the possible functionalities of circRNAs in SCI through anticipating specific interactive miRNAs. Their main findings suggest the possible involvement and functional significance of circRNAs in SCI.

The manuscripts included in this Research Topic show the diversity of bioinformatics tools that could be applied to investigate intriguing questions in the neuroscience field. They could direct more precise *in vitro* analyses and provide interesting findings to clarify the pathogenesis of complex diseases, as well as to suggest novel biomarkers and treatment options.

Author contributions

CR: Conceptualization, Supervision, Writing—original draft, Writing—review and editing. JS: Conceptualization, Supervision, Writing—original draft, Writing—review and editing. MR-M: Supervision, Writing—review and editing. TK: Conceptualization, Supervision, Writing—original draft, Writing—review and editing.

Funding

This work has been supported in part by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) - Finance Code 001, Fundação de Amparo à Pesquisa do Rio Grande do Sul (FAPERGS), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) grant no. 23/2551-0000115-2, and FIPE/HCPA grant no. 2022-0567. TK is the recipient of a CNPq scholarship (grant no. 150181/2023-0). MR-M is the recipient of a CNPq scholarship grant (grant no. 308075/2021-8).

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.



Association Between C-Reactive Protein and Risk of Amyotrophic Lateral Sclerosis: A Mendelian Randomization Study

Yahui Zhu^{1,2}, Mao Li^{1,2}, Jinghong Zhang^{1,2} and Xusheng Huang^{1,2*}

¹Medical School of Chinese PLA, Beijing, China, ²Department of Neurology, The First Medical Center, Chinese PLA General Hospital, Beijing, China

Background: Until now, the relationship between C-reactive protein (CRP) levels and amyotrophic lateral sclerosis (ALS) risk has not been fully established. It is necessary to assess whether there is a causal relationship between C-reactive protein levels and ALS risk.

Objective and Methods: We aimed to determine whether CRP has causal effects on risk of ALS. In this present study, summary-level data for ALS (20,806 cases and 59,804 controls) was obtained from large analyses of genome-wide association studies. For instrumental variables, 37 single nucleotide polymorphisms that had been previously identified to be related to CRP levels were used, including 4 SNPs of conservative CRP genetic variants and 33 SNPs of liberal CRP genetic variants. MR estimates were calculated using the inverse-variance weighted method, supplemented by MR-Egger, weighted median, and MR-PRESSO methods.

Results: There was no significant causal relationship between genetically predicted CRP levels and ALS risk (OR = 1.123, 95% CI = 0.963–1.309, $p = 0.139$) and results for the conservative CRP instruments were consistent (OR = 0.964, 95% CI = 0.830–1.119, $p = 0.628$). Pleiotropic bias was not observed in this study.

Conclusions: This study suggests that genetically predicted CRP levels may not be a causal risk factor for ALS.

Keywords: amyotrophic lateral sclerosis, mendelian randomization, C-reactive protein, single-nucleotide polymorphisms, causal relationship

INTRODUCTION

Amyotrophic lateral sclerosis (ALS) is a lethal neurodegenerative disease characterized by motor neuron axonal degeneration, with an average survival of 3–5 years after symptom onset (Turner et al., 2013). Serum C-reactive protein (CRP) is a biomarker of systemic inflammation (Koenig et al., 1999) and has also been considered a biomarker of neurodegeneration (Luan and Yao., 2018). The previous studies have found elevated levels of serum C-reactive protein in ALS patients (Ryberg et al., 2010) (Cui et al., 2020), but Huang et al. suggests no significant difference of CRP levels in ALS patients when compared to controls (Huang et al., 2020). Several studies have evaluated the role of CRP as a prognostic marker in ALS and the relationship between CRP and disease progression/

OPEN ACCESS

Edited by:

Clévia Rosset,
Clinical Hospital of Porto Alegre, Brazil

Reviewed by:

Renata Cupertino,
University of Vermont, United States
Sarah A Gagliano Taliun,
Université de Montréal, Canada

*Correspondence:

Xusheng Huang
lewish301@sina.com

Specialty section:

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

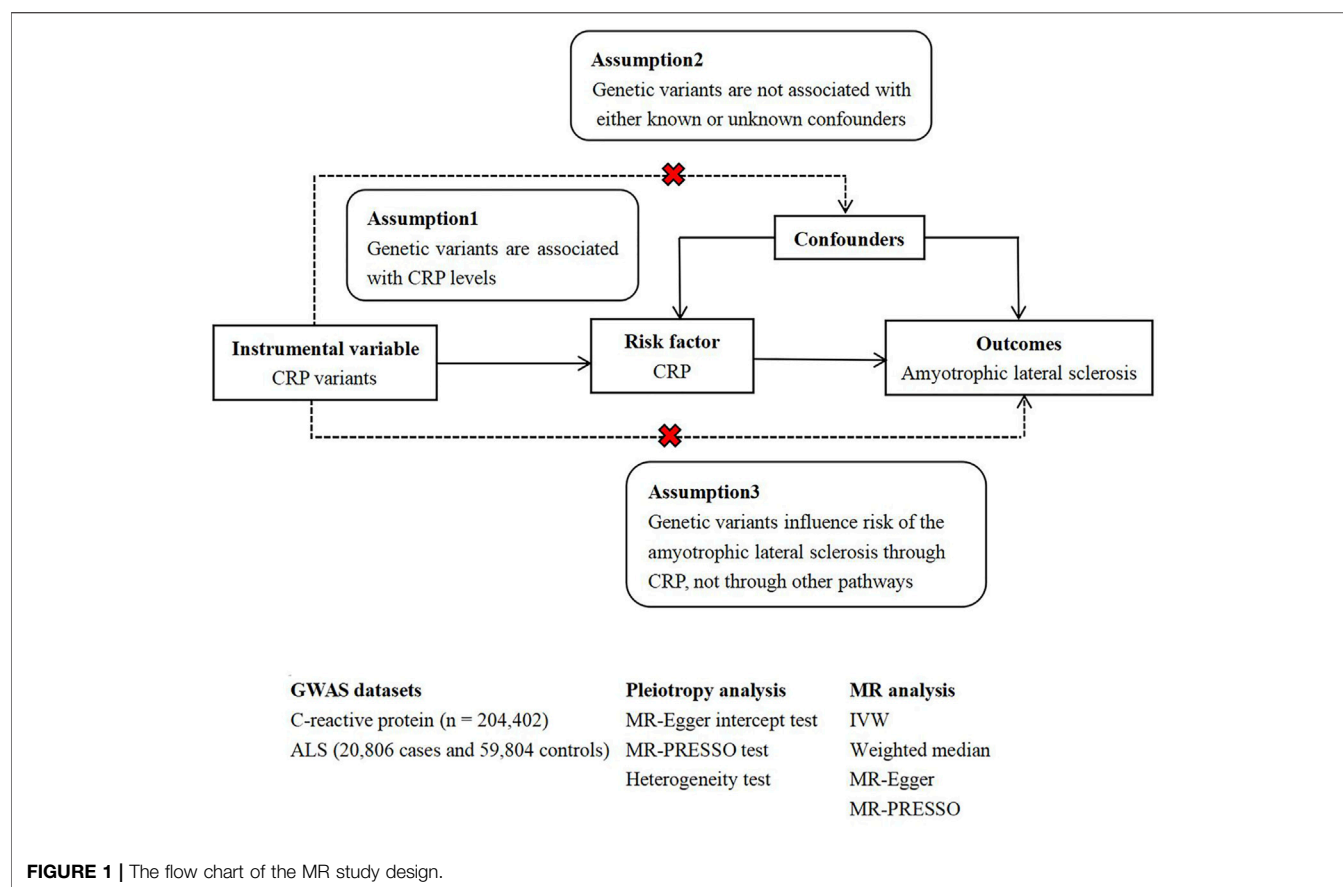
Received: 13 April 2022

Accepted: 04 May 2022

Published: 20 May 2022

Citation:

Zhu Y, Li M, Zhang J and Huang X
(2022) Association Between C-
Reactive Protein and Risk of
Amyotrophic Lateral Sclerosis: A
Mendelian Randomization Study.
Front. Genet. 13:919031.
doi: 10.3389/fgene.2022.919031



survival rate. One study suggests that ALS patients with elevated CRP levels have faster disease progression than those with lower CRP levels and that serum CRP may be a prognostic biomarker in ALS (Lunetta et al., 2017). However, the other study showed no association of serum CRP with survival rate (De Schaepdryver et al., 2020). ALS usually occurs in middle-aged people, and this age group is often accompanied by other complications, which may affect CRP level. Since data on the association of CRP levels with ALS are often derived from observational studies, which could be subject to potential confounding bias and reverse causes, such as chronic diseases, cardiovascular risk factors and so on, it is unclear whether CRP levels are a risk factor for ALS.

Mendelian randomization (MR) is a novel method to evaluate the causal relationship between risk factors and diseases using genetic variation in observational studies (Davies et al., 2018). Due to the random assignment of genes at conception, genetic variants predate disease development and are not influenced by environmental risk factors. Thus, MR overcomes the core deficiencies of observational studies, minimizing confounding bias and reverse causality, allowing the assessment of potential causality (Walker et al., 2017).

Here, we performed a two-sample Mendelian randomization to further understand the causal effects of C-reactive protein levels on the risk of ALS. In this study, single nucleotide polymorphisms (SNPs) related to CRP levels were used as instrumental variables.

MATERIALS AND METHODS

As all analyses were performed using publicly available genome-wide association study (GWAS) summary data, no additional ethical permission was required from institutional research ethics committees. This study followed the Strengthening the Reporting of Observational Studies in Epidemiology Using Mendelian Randomization (STROBE-MR) guide (Skrivankova et al., 2021).

Study Design

In general, MR studies must satisfy three principal assumptions, as shown in **Figure 1**, which is a flow chart of our research design. Both the second and third hypotheses are designed to ensure independence from pleiotropy, as described in some previous studies (He et al., 2020) (Zhang H. et al., 2020).

C-Reactive Protein Genetic Variants

We selected two sets of CRP genetic variants as instrumental variables, including conservative CRP genetic variants (Wensley et al., 2011) and liberal CRP genetic variants (Ligthart et al., 2018).

Wensley et al. (Wensley et al., 2011) used detailed information about the composition of the CRP gene to select a set of SNPs (rs3093077, rs1205, rs1130864, and rs1800947) that fully covers the common variations of CRP gene in populations of European descent (that is, minor allele frequency ≥ 0.05 and an r^2 threshold of ≥ 0.8). The 4 SNPs (rs3093077, rs1130864, rs1205 and

rs1800947) were at the CRP locus and used as conservative CRP instrumental variables.

Liberal CRP instrumental variables were extracted from a pooled analysis of GWAS by using data from 88 studies comprising 204,402 European individuals. The GWAS meta-analyses of CRP revealed 58 distinct genetic loci ($p < 5 \times 10^{-8}$). The lead variants at the distinct loci explained up to 7.0% of the variance in circulating amounts of CRP. In this study, serum CRP levels were measured using standard laboratory techniques. The authors excluded individuals with autoimmune disease, taking immunomodulators, or C-reactive protein levels four standard deviations or more from the mean. Analyses were adjusted for age, sex, population substructure, and correlation. Liberal CRP instrumental variants represented that SNPs were extracted from CRP GWAS dataset mentioned above and at the distinct genetic loci, such as C6orf173, FABP1, IL1R1 and so on.

Amyotrophic Lateral Sclerosis Genome-Wide Association Study Dataset

The current study was based on publicly available ALS GWAS summary statistics data, including 20,806 ALS cases and 59,804 controls in people of European ancestry (Nicolas et al., 2018). In the study, Nicolas et al. undertook a large-scale GWAS involving 12,663 patients diagnosed with ALS and 53,439 control subjects and incorporated into a meta-analysis with GWAS involving 12,577 ALS cases and 23,475 control subjects (van Rheenen et al., 2016). After imputation and quality-control measures, variants from 20,806 ALS cases and 59,804 control samples were available for association analysis. All ALS patients included in the case cohort were diagnosed by neurologists specializing in ALS according to the El Escorial criteria (Brooks, 1994).

We compared the sources of participants in the CRP GWAS dataset and ALS GWAS dataset. Since the participants were from different studies or consortiums, we thought that the probability of overlapping samples between CRP GWAS and ALS GWAS was small.

Mendelian Randomization Analysis

Conservative CRP instrumental variables were rs3093077, rs1130864, rs1205 and rs1800947. For liberal CRP instrumental variables, we first identified significant SNPs ($p < 5 \times 10^{-8}$) related to C-reactive protein from the summary analysis of GWAS. SNPs for CRP were clumped using standard parameters (clumping window of 10000 kb, r^2 cutoff value of 0.001) to discard variants in linkage disequilibrium (LD). Here, 57 independent SNPs were found to be associated with CRP. If SNPs were absent in the ALS GWAS dataset and for which proxies ($r^2 > 0.9$) were not available by searching the online website SNIIPA (<http://snipa.helmholtzmuench.de/snipa3/>), these unavailable SNPs would be excluded from downstream analysis. Subsequently, to satisfy the second assumption, we used the PhenoScanner tool (Kamat et al., 2019) to examine whether selected SNPs were associated with potential confounders affecting ALS. When using the PhenoScanner tool, the threshold for genome-wide significance was set at $p < 5 \times$

10^{-8} . In addition, we applied MR Steiger filtering (Hemani et al., 2017) to test the causal direction of the obtained SNPs on exposures and outcomes. We excluded SNPs with “FALSE” results because these SNPs mainly affected the outcomes, not exposures. Finally, we assessed the power of remaining SNPs using the F statistics ($F = \beta^2/se^2$) for each SNP. SNPs with less statistical power would be removed to avoid weak instrumental variables (F statistics < 10) (Chen L. et al., 2021).

Pleiotropy analyses were mainly based on three different statistical methods, including the MR-Egger intercept test (Verbanck et al., 2018), MR Pleiotropy RESidual Sum and Outlier (MR-PRESSO) global test (Verbanck et al., 2018), and the heterogeneity test using Cochran’s Q statistic (Greco et al., 2015). Statistically significant differences for the above analyses were set at p value < 0.05 . In addition, we depicted funnel plots to visualize any heterogeneity of effect estimates. Asymmetry about the vertical line is indicative of the heterogeneity.

Six MR analysis methods were selected including the inverse-variance weighted (IVW), weighted median, MR-Egger, MR-PRESSO, simple mode, and weighted mode test. The random-effects IVW method, the main method of the study, essentially assumed a zero intercept and performed a weighted regression of the SNP-exposure effects with the SNP-outcome effects. The MR Egger method provided more conservative causal estimates in the presence of pleiotropy and was less likely to produce exaggerated test statistics (Burgess and Thompson, 2017). Even when up to 50% of the information in the analysis came from invalid IVs, the weighted median method could provide valid estimates (Bowden et al., 2016). The MR-PRESSO method was used to detect outliers that might bias the results and to assess whether causal estimates change after removing outliers (Verbanck et al., 2018). In addition, we performed a leave-one-SNP-out analysis. In this analysis, we systematically removed one SNP at a time, assessing the impact of potentially pleiotropic SNPs on causal effects. Estimates were expressed as odds ratio (OR) and 95% confidence interval (CI) per unit increase in natural log-transformed genetically predicted CRP levels (mg/L). Statistical analysis was performed in version R4.1.2 (TwoSampleMR and MR-PRESSO packages). The significance threshold was p value < 0.05 .

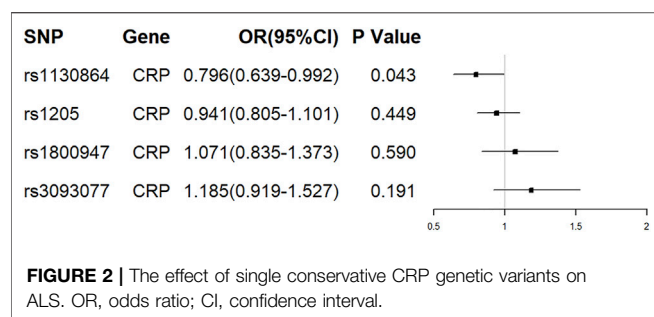
Power Analysis

The proportion of CRP variance was explained by each instrument SNP R^2 , which was calculated using the following formula: $R^2 = 2\beta^2 \text{MAF} (1 - \text{MAF})$ (Park et al., 2010). Where MAF represents the minor allele frequency of the instrument SNP, and β denotes the effect size for SNP. The statistical power was calculated using the web-based tool mRnd, where the two-sided type-I error rate α was 0.05 (Brion et al., 2013).

RESULTS

Conservative CRP Genetic Variants With Amyotrophic Lateral Sclerosis

Of the 4 SNPs, only 1 was shown to be significantly associated with the risk of ALS (OR = 0.796, 95% CI = 0.639–0.992, $p =$



0.043) and the other 3 SNPs were showed no association with the risk of ALS (**Figure 2**). Overall, there was no evidence to suggest a causal association between CRP levels and ALS risk in the analysis using IVW method with an OR of 0.964 (95% CI = 0.830–1.119, $p = 0.628$). The results were consistent in the analyses using weighted median and MR Egger methods (**Figure 3**). The detailed characteristics of the 4 SNPs was showed in **Table 1**.

Heterogeneity was not observed using Cochran Q statistic based on IVW ($p = 0.102$) and MR-Egger ($p = 0.519$). The MR Egger intercept (intercept = -0.092, SE = 0.042, $p = 0.158$) showed no horizontal pleiotropy. In PhenoScanner database, none of the 4 SNPs were observed to be associated with other traits, diseases or risk factors.

Liberal CRP Genetic Variants With Amyotrophic Lateral Sclerosis

A total of 57 independent SNPs were found to be associated with CRP. Rs644234 was excluded in downstream analysis because it was not available in ALS data. Since rs2794520 was at the CRP locus, in order to reduce horizontal pleiotropy, we would analyze the effect of rs2794520 and other 55 SNPs on ALS risk separately.

Rs2794520 explained 1.48% variance of CRP. For rs2794520, we did not observe a causal relationship between CRP levels and ALS risk (OR = 0.949, 95% CI = 0.813–1.108, $p = 0.505$).

For the 55 SNPs, harmonising CRP and ALS, the following SNPs for being palindromic with intermediate allele frequencies were excluded: rs10778215 and rs11108056. When using the PhenoScanner tool, we excluded 20 SNPs that were associated with confounders, which were proved to be causally associated with ALS such as low density lipoprotein (LDL), total cholesterol (Chen et al., 2018), type 2 diabetes (T2DM) (Chen H. et al., 2021), childhood body mass index (Zhang L. et al., 2020), neutrophil count, white blood cell count (Li et al., 2020), systolic blood pressure and diastolic blood pressure (Xia et al., 2022). The remaining SNPs were all with true causal direction identified by the MR Steiger filtering. The F statistics for each SNP were greater than the statistical threshold of 10, indicating sufficient validity for all SNPs. Thus, the MR analysis of ALS included 33 SNPs related to CRP. The detailed characteristics of the SNPs was showed in **Supplementary Table S1**. All these 33 genetic variants could explain 1.55% variance of CRP.

Cochran Q statistic based on IVW ($p = 0.491$) and MR-Egger ($p = 0.613$) showed no evidence of heterogeneity and the symmetry of the funnel plot supported that (**Figure 4**). The MR Egger intercept (intercept = 0.011, SE = 0.006, $p = 0.077$) suggested no horizontal pleiotropy for instrumental variables, and the MR-PRESSO global test ($p = 0.482$) supported that. The MR-PRESSO outlier test did not identify outlier SNPs. Therefore, the selected 33 SNPs associated with CRP were used as instrumental variables for the downstream MR analysis.

For the 33 SNPs, there was no evidence of a causal relationship between genetically predicted CRP levels and ALS risk, with p values >0.05 in the analysis using IVW approach with an OR of 1.123 (95% CI = 0.963–1.309, $p = 0.139$). The other analyses, MR Egger (OR = 0.838, 95% CI = 0.591–1.188, $p = 0.329$), weighted median (OR = 1.018, 95% CI

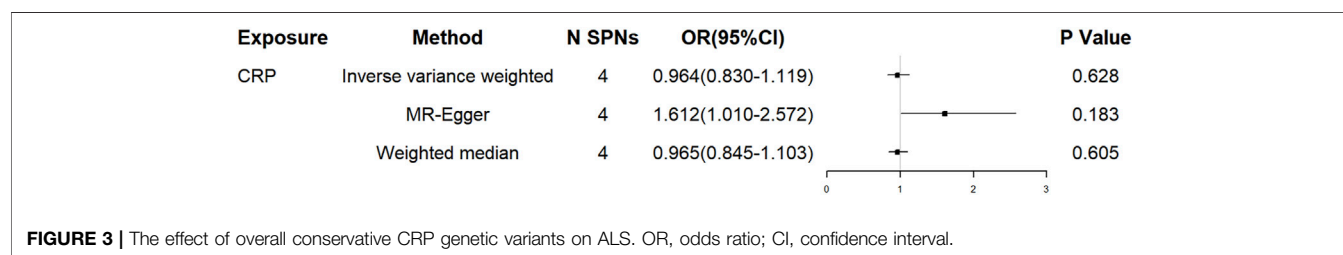


TABLE 1 | The characteristics of the selected conservative CRP instrumental variables.

| SNP | Effect Allele | Other Allele | Association with CRP | | | Amyotrophic Lateral Sclerosis | | |
|-----------|---------------|--------------|----------------------|---------|-------|-------------------------------|--------|---------|
| | | | EAF | β | SE | β | SE | P |
| rs3093077 | C | A | 0.0716 | 0.21 | 0.018 | 0.0356 | 0.0272 | 0.1913 |
| rs1205 | C | T | 0.6899 | 0.18 | 0.01 | -0.0109 | 0.0144 | 0.4501 |
| rs1130864 | A | G | 0.327 | 0.13 | 0.008 | -0.0296 | 0.0146 | 0.04302 |
| rs1800947 | C | G | 0.9429 | 0.26 | 0.015 | 0.0178 | 0.033 | 0.5905 |

SNPs, single nucleotide polymorphisms; EAF, effect allele frequency; SE, standard error.

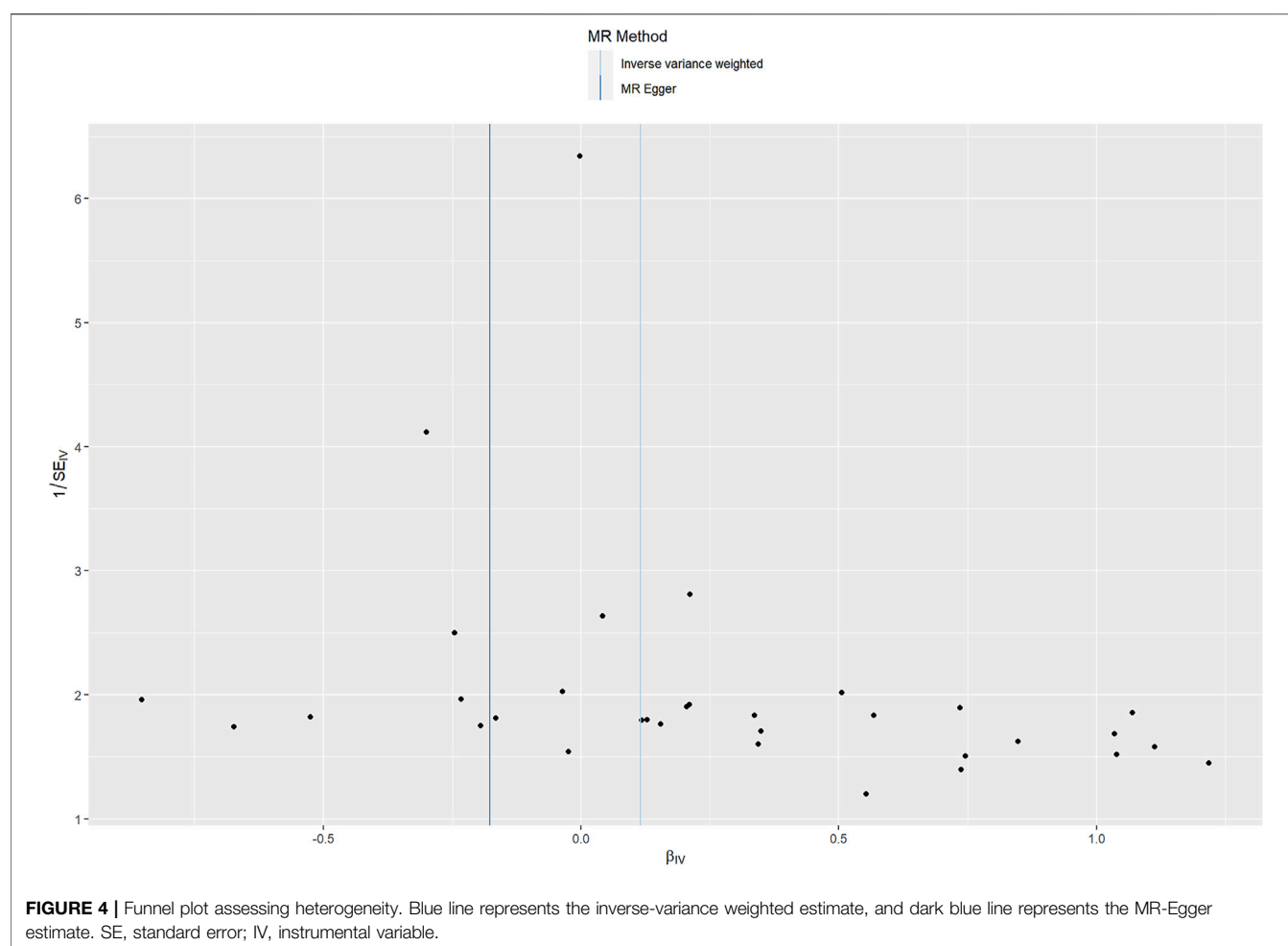


TABLE 2 | The causal association of CRP levels with ALS risk.

| Method | N SNPs | OR | 95%CI | p Value |
|----------------------------|--------|-------|-------------|---------|
| IWW | 33 | 1.123 | 0.963–1.309 | 0.139 |
| MR-Egger | 33 | 0.838 | 0.591–1.188 | 0.329 |
| Weighted median | 33 | 1.018 | 0.795–1.304 | 0.885 |
| MR-PRESSO (raw, 0 outlier) | 33 | 1.123 | 0.964–1.307 | 0.146 |
| simple mode | 33 | 1.133 | 0.715–1.795 | 0.598 |
| weighted mode | 33 | 1.002 | 0.770–1.303 | 0.990 |

SNPs, single nucleotide polymorphisms; OR, odds ratio; CI, confidence interval; IWW, inverse-variance weighted; MR-PRESSO, MR, Pleiotropy RESidual Sum and Outlier.

= 0.795–1.304, $p = 0.885$) and MR-PRESSO (OR = 1.123, 95% CI = 0.964–1.307, $p = 0.146$) methods supported these results (Table 2; Figure 5 and Figure 6). In the leave-one-out analysis, we did not observe a single SNP of CRP to have an influence on the association (Figure 7).

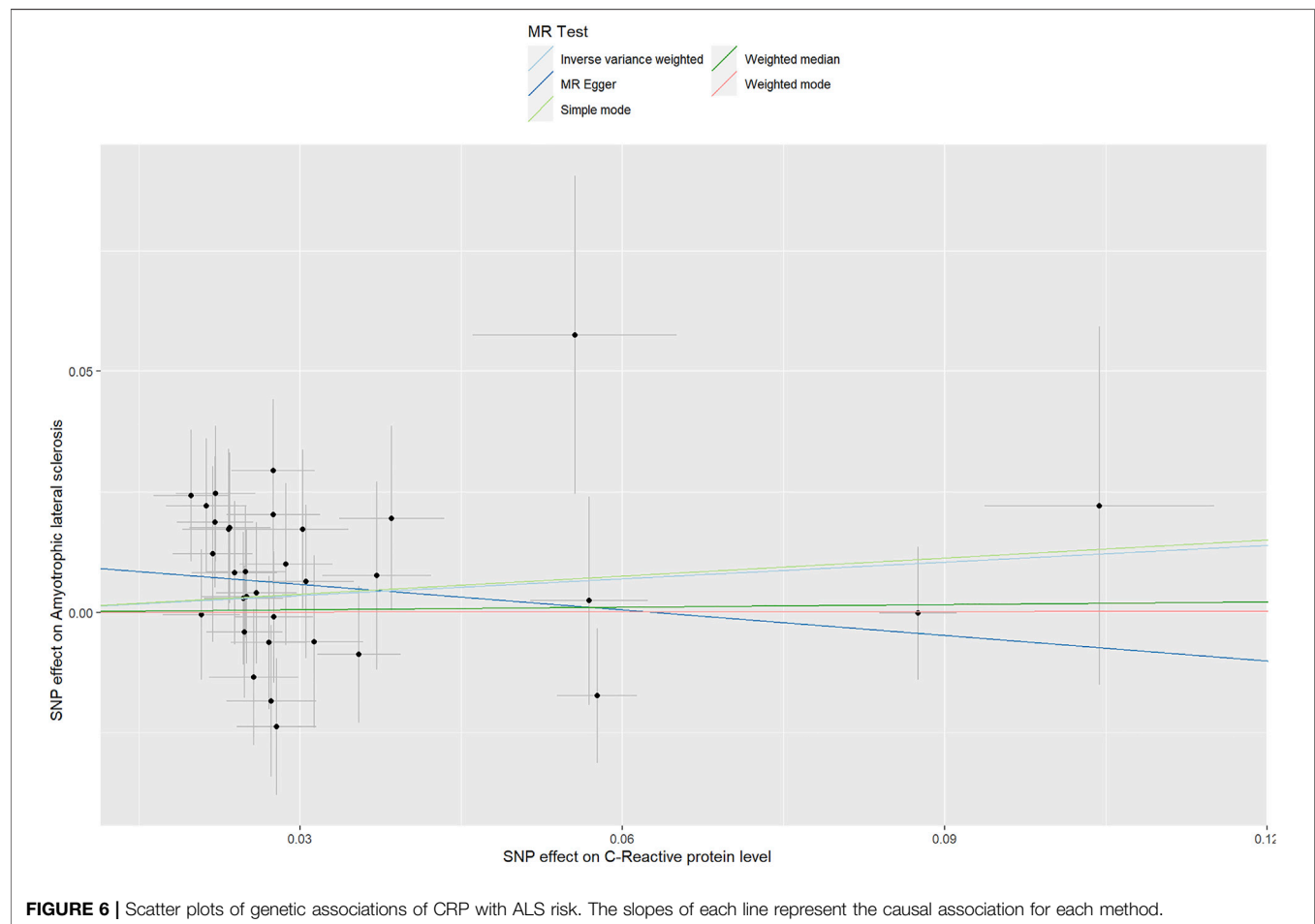
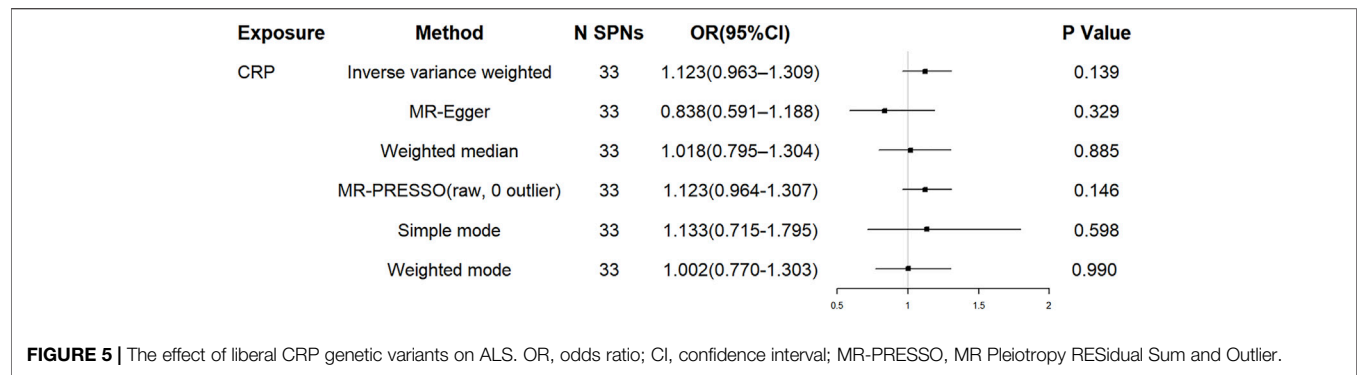
Power Analysis

For liberal CRP genetic variants with ALS, our MR study had 80% power to detect an OR of 1.191 or higher per SD increase in C-reactive protein levels for ALS with an alpha of 5%.

DISCUSSION

Until now, the relationship between C-reactive protein levels and ALS risk has not been fully established. Hence, it is necessary to clarify the causal relationship between CRP levels and ALS risk in order to develop effective therapeutic and preventive measures. In our study, we selected 4 SNPs and 33 SNPs as conservative CRP instrumental variables and liberal CRP instrumental variables and obtained their corresponding summary statistics in the ALS GWAS dataset. After that, we evaluated the causal link of CRP levels with ALS risk by MR analysis. Our study showed no causal effects between CRP levels and ALS risk. CRP levels might not increase or reduce the risk of ALS.

Inflammation is involved in the pathogenesis of central nervous system neurodegenerative diseases, including ALS (Stephenson et al., 2018). The neuroinflammation of ALS is predominantly characterized by activation of microglia and astrocytes innate immune sensing pathways to the central nervous system (McCauley and Baloh, 2019). CRP, as a biomarker of low-grade inflammation, has been proposed to play a role in the development of ALS. However, a recent study showed that compared to controls, patients with ALS had lower CRP levels before diagnosis until 1 year after



diagnosis. After that, ALS patients had higher CRP levels when compared to controls. This study suggested that C-reactive protein was involved in the disease course of ALS, but perhaps only in the later disease stages (Cui et al., 2020). Therefore, this study also supports our findings to a certain extent. That is, C-reactive protein may not be a risk factor for the development of ALS. Although persistent inflammation is thought to be a contributor in the development of ALS, inflammation may promote the development of ALS through inflammatory

mediators other than C-reactive protein. Hence, there is no clear causal relationship between C-reactive protein levels and ALS risk.

One of the advantages of our study is that the use of MR analysis reduces potential confounding factors and reverse causality, minimizing bias from the traditional observational studies. To our knowledge, this is the first study to investigate the causal relationship between CRP and ALS risk using genetic variation. In addition, the large sample size of this study and the inclusion of multiple SNPs as

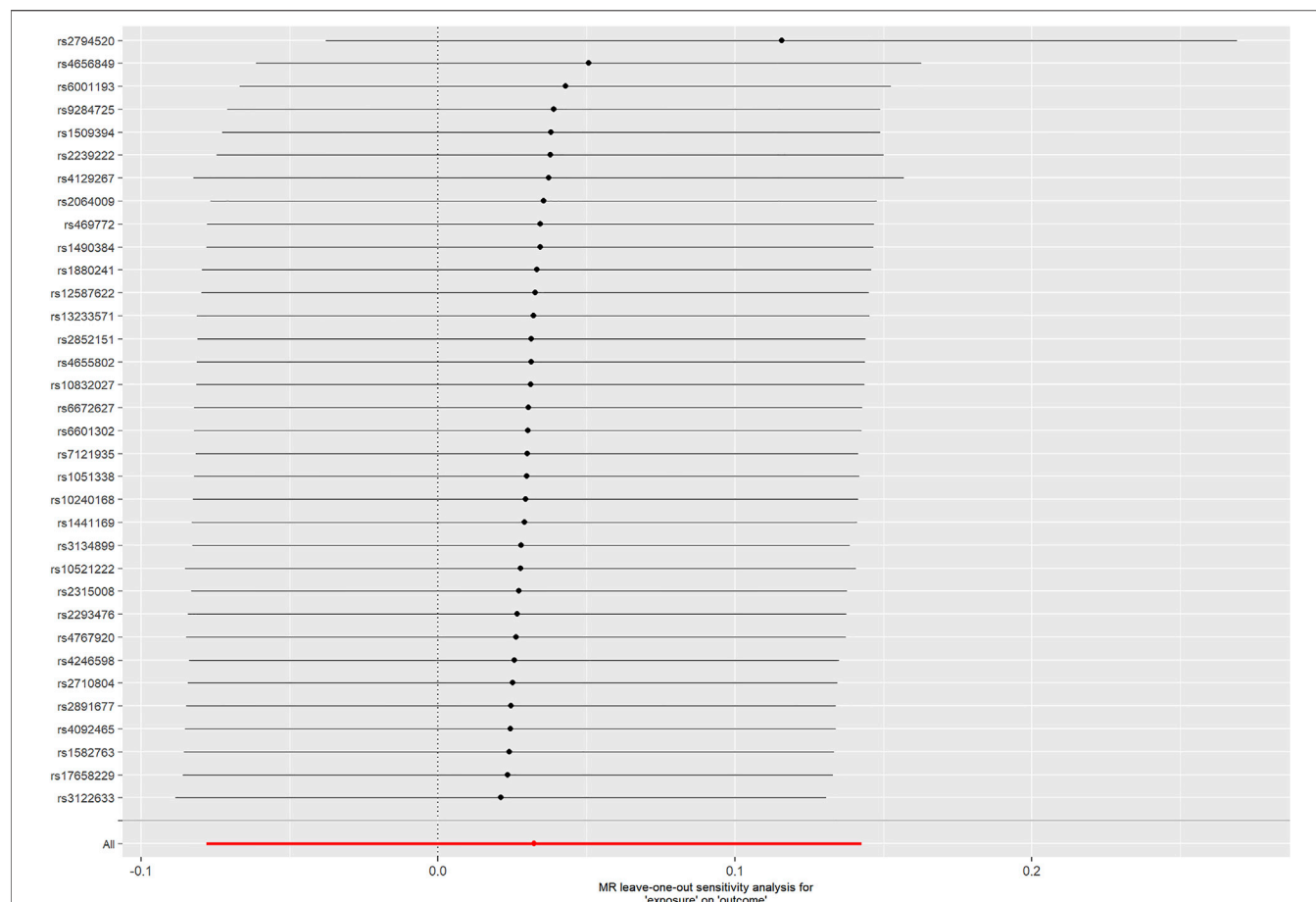


FIGURE 7 | MR leave-one-out sensitivity analysis for CRP on ALS.

instrumental variables improve statistical power, which increases the ability to identify weak associations. Finally, no pleiotropy is observed in this study, and these models produce similar conclusions, suggesting the robustness of the findings.

Our study also has some limitations. First, since this study used summary-level data, the possibility of a non-linear association between CRP levels and the risk of ALS cannot be completely ruled out. Second, as this study was based on pooled-level data, we could not conduct a more detailed subgroup analysis. Therefore, in further studies, the opportunity to obtain more detailed data on the individual-level, such as patient's age, gender, and so on, will help us further understand the causal relationship between CRP levels and the risk of ALS in each subgroup. We look forward to the opportunity for authors of the publicly available GWAS to share individual-level data for further research. Third, bi-directional MR analysis was not carried out. For ALS as exposure, we identified significant SNPs ($p < 5 \times 10^{-8}$) associated with ALS from the summary analysis of GWAS. Then, SNPs for ALS were clumped using standard parameters (clumping window of 10000 kb, r^2 cutoff value of 0.001) to discard variants in linkage disequilibrium (LD). Here, 6 independent SNPs associated with ALS were found. However, only one SNP (rs3849938) in the CRP (outcome) GWAS dataset was available. We considered that the number of instrumental

variables was too small for MR analysis, so bi-directional MR analysis was not carried out. In addition, since all participants (including CRP levels and ALS) are of European ancestry, this study may not be available in other ethnic groups, making it difficult to extrapolate the research results to other ancestries. Hence, the findings should be further replicated in other ancestries.

To conclude, we found no evidence to support a causal relationship between CRP levels and the risk of ALS. In the future, we expect a larger GWAS database and individual-level data to be available. Meanwhile, additional studies are also expected to further confirm our findings.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: <https://gwas.mrcieu.ac.uk/>.

AUTHOR CONTRIBUTIONS

XH contributed to conception and design of the study. YZ, ML, JZ organized the database and performed the statistical analysis. YZ

wrote the first draft of the manuscript. XH revised the manuscript. All authors contributed to interpretation of the results and critical revision of the manuscript for important intellectual content and approved the final version of the manuscript.

FUNDING

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

REFERENCES

- Bowden, J., Smith, G. D., Haycock, P. C., and Burgess, S. (2016). Consistent Estimation in Mendelian Randomization with Some Invalid Instruments Using a Weighted Median Estimator. *Genet. Epidemiol.* 40, 304–314. doi:10.1002/gepi.21965
- Brion, M.-J. A., Shakhbazov, K., and Visscher, P. M. (2013). Calculating Statistical Power in Mendelian Randomization Studies. *Int. J. Epidemiol.* 42 (5), 1497–1501. doi:10.1093/ije/dyt179
- Brooks, B. R. (1994). El Escorial World Federation of Neurology Criteria for the Diagnosis of Amyotrophic Lateral Sclerosis. *J. Neurol. Sci.* 124, 96–107. doi:10.1016/0022-510x(94)90191-0
- Burgess, S., and Thompson, S. G. (2017). Interpreting Findings from Mendelian Randomization Using the MR-Egger Method. *Eur. J. Epidemiol.* 32 (5), 377–389. doi:10.1007/s10654-017-0255-x
- Chen, H., Zhang, J., Wang, T., Zhang, S., Lai, Q., Huang, S., et al. (2021). Type 2 Diabetes Mellitus and Amyotrophic Lateral Sclerosis: Genetic Overlap, Causality, and Mediation. *J. Clin. Endocrinol. Metab.* 106 (11), e4497–e4508. doi:10.1210/clinem/dgab465
- Chen, L., Yang, H., Li, H., He, C., Yang, L., and Lv, G. (2021). Insights into Modifiable Risk Factors of Cholelithiasis: A Mendelian Randomization Study. *Hepatology* 75, 785–796. doi:10.1002/hep.32183
- Chen, X., Yazdani, S., Piehl, F., Magnusson, P. K. E., and Fang, F. (2018). Polygenic Link between Blood Lipids and Amyotrophic Lateral Sclerosis. *Neurobiol. Aging* 67, 202.e1–202.e6. doi:10.1016/j.neurobiolaging.2018.03.022
- Cui, C., Sun, J., Pawitan, Y., Piehl, F., Chen, H., Ingre, C., et al. (2020). Creatinine and C-Reactive Protein in Amyotrophic Lateral Sclerosis, Multiple Sclerosis and Parkinson's Disease. *Brain Commun.* 2 (2), fcaa152. doi:10.1093/braincomms/fcaa152
- Davies, N. M., Holmes, M. V., and Smith, G. D. (2018). Reading Mendelian Randomisation Studies: A Guide, Glossary, and Checklist for Clinicians. *Bmj* 362, k601. doi:10.1136/bmj.k601
- De Schaepdryver, M., Lunetta, C., Tarlarini, C., Mosca, L., Chio, A., Van Damme, P., et al. (2020). Neurofilament Light Chain and C Reactive Protein Explored as Predictors of Survival in Amyotrophic Lateral Sclerosis. *J. Neurol. Neurosurg. Psychiatry* 91 (4), 436–437. doi:10.1136/jnnp-2019-322309
- Greco M, F. D., Minelli, C., Sheehan, N. A., and Thompson, J. R. (2015). Detecting Pleiotropy in Mendelian Randomisation Studies with Summary Data and a Continuous Outcome. *Stat. Med.* 34 (21), 2926–2940. doi:10.1002/sim.6522
- He, Y., Zhang, H., Wang, T., Han, Z., Ni, Q.-b., Wang, K., et al. (2020). Impact of Serum Calcium Levels on Alzheimer's Disease: A Mendelian Randomization Study. *Jad* 76 (2), 713–724. doi:10.3233/jad-191249
- Hemani, G., Tilling, K., and Smith, G. D. (2017). Orienting the Causal Relationship between Imprecisely Measured Traits Using GWAS Summary Data. *PLoS Genet.* 13 (11), e1007081. doi:10.1371/journal.pgen.1007081
- Huang, F., Zhu, Y., Hsiao-Nakamoto, J., Tang, X., Dugas, J. C., Moscovitch-Lopatin, M., et al. (2020). Longitudinal Biomarkers in Amyotrophic Lateral Sclerosis. *Ann. Clin. Transl. Neurol.* 7 (7), 1103–1116. doi:10.1002/actn.3.51078
- Kamat, M. A., Blackshaw, J. A., Young, R., Surendran, P., Burgess, S., Danesh, J., et al. (2019). PhenoScanner V2: An Expanded Tool for Searching Human

ACKNOWLEDGMENTS

We thank all the investigators for making the summary data publicly available, and we are grateful for all the investigators and participants who contributed to those studies.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2022.919031/full#supplementary-material>

- Genotype-Phenotype Associations. *Bioinformatics* 35 (22), 4851–4853. doi:10.1093/bioinformatics/btz469
- Koenig, W., Sund, M., Fröhlich, M., Fischer, H.-G., Löwel, H., Döring, A., et al. (1999). C-reactive Protein, a Sensitive Marker of Inflammation, Predicts Future Risk of Coronary Heart Disease in Initially Healthy Middle-Aged Men. *Circulation* 99 (2), 237–242. doi:10.1161/01.cir.99.2.237
- Li, C., Yang, W., Wei, Q., and Shang, H. (2020). Causal Association of Leukocytes Count and Amyotrophic Lateral Sclerosis: A Mendelian Randomization Study. *Mol. Neurobiol.* 57 (11), 4622–4627. doi:10.1007/s12035-020-02053-7
- Ligthart, S., Vaez, A., Vösa, U., Stathopoulou, M. G., de Vries, P. S., Prins, B. P., et al. (2018). Genome Analyses of >200,000 Individuals Identify 58 Loci for Chronic Inflammation and Highlight Pathways that Link Inflammation and Complex Disorders. *Am. J. Hum. Genet.* 103 (5), 691–706. doi:10.1016/j.ajhg.2018.09.009
- Luan, Y.-y., and Yao, Y.-m. (2018). The Clinical Significance and Potential Role of C-Reactive Protein in Chronic Inflammatory and Neurodegenerative Diseases. *Front. Immunol.* 9, 1302. doi:10.3389/fimmu.2018.01302
- Lunetta, C., Lizio, A., Maestri, E., Sansone, V. A., Mora, G., Miller, R. G., et al. (2017). Serum C-Reactive Protein as a Prognostic Biomarker in Amyotrophic Lateral Sclerosis. *JAMA Neurol.* 74 (6), 660–667. doi:10.1001/jama.2016.6179
- McCauley, M. E., and Baloh, R. H. (2019). Inflammation in ALS/FTD Pathogenesis. *Acta Neuropathol.* 137 (5), 715–730. doi:10.1007/s00401-018-1933-9
- Nicolas, A., Kenna, K. P., Renton, A. E., Ticozzi, N., Faghri, F., Chia, R., et al. (2018). Genome-wide Analyses Identify KIF5A as a Novel ALS Gene. *Neuron* 97 (6), 1268–1283. e1266. doi:10.1016/j.neuron.2018.02.027
- Park, J.-H., Wacholder, S., Gail, M. H., Peters, U., Jacobs, K. B., Chanock, S. J., et al. (2010). Estimation of Effect Size Distribution from Genome-wide Association Studies and Implications for Future Discoveries. *Nat. Genet.* 42 (7), 570–575. doi:10.1038/ng.610
- Ryberg, H., An, J., Darko, S., Lustgarten, J. L., Jaffa, M., Gopalakrishnan, V., et al. (2010). Discovery and Verification of Amyotrophic Lateral Sclerosis Biomarkers by Proteomics. *Muscle Nerve* 42 (1), 104–111. doi:10.1002/mus.21683
- Skrivankova, V. W., Richmond, R. C., Woolf, B. A. R., Yarmolinsky, J., Davies, N. M., Swanson, S. A., et al. (2021). Strengthening the Reporting of Observational Studies in Epidemiology Using Mendelian Randomization. *JAMA* 326, 1614–1621. doi:10.1001/jama.2021.18236
- Stephenson, J., Nutma, E., van der Valk, P., and Amor, S. (2018). Inflammation in CNS Neurodegenerative Diseases. *Immunology* 154 (2), 204–219. doi:10.1111/imm.12922
- Turner, M. R., Hardiman, O., Benatar, M., Brooks, B. R., Chio, A., de Carvalho, M., et al. (2013). Controversies and Priorities in Amyotrophic Lateral Sclerosis. *Lancet Neurology* 12 (3), 310–322. doi:10.1016/s1474-4422(13)70036-x
- van Rheenen, W., Shatunov, A., Dekker, A. M., McLaughlin, R. L., Diekstra, F. P., Pulit, S. L., et al. (2016). Genome-wide Association Analyses Identify New Risk Variants and the Genetic Architecture of Amyotrophic Lateral Sclerosis. *Nat. Genet.* 48 (9), 1043–1048. doi:10.1038/ng.3622
- Verbanck, M., Chen, C.-Y., Neale, B., and Do, R. (2018). Detection of Widespread Horizontal Pleiotropy in Causal Relationships Inferred from Mendelian Randomization between Complex Traits and Diseases. *Nat. Genet.* 50 (5), 693–698. doi:10.1038/s41588-018-0099-7

- Walker, V. M., Davey Smith, G., Davies, N. M., and Martin, R. M. (2017). Mendelian Randomization: A Novel Approach for the Prediction of Adverse Drug Events and Drug Repurposing Opportunities. *Int. J. Epidemiol.* 46 (6), 2078–2089. doi:10.1093/ije/dyx207
- Wensley, F., Gao, P., Burgess, S., Kaptoge, S., Di Angelantonio, E., Shah, T., et al. (2011). Association between C Reactive Protein and Coronary Heart Disease: Mendelian Randomisation Analysis Based on Individual Participant Data. *Bmj* 342, d548. doi:10.1136/bmj.d548
- Xia, K., Zhang, L., Tang, L., Huang, T., and Fan, D. (2022). Assessing the Role of Blood Pressure in Amyotrophic Lateral Sclerosis: A Mendelian Randomization Study. *Orphanet J. Rare Dis.* 17 (1), 56. doi:10.1186/s13023-022-02212-0
- Zhang, H., Wang, T., Han, Z., and Liu, G. (2020). Mendelian Randomization Study to Evaluate the Effects of Interleukin-6 Signaling on Four Neurodegenerative Diseases. *Neurol. Sci.* 41 (10), 2875–2882. doi:10.1007/s10072-020-04381-x
- Zhang, L., Tang, L., Huang, T., and Fan, D. (2020). Life Course Adiposity and Amyotrophic Lateral Sclerosis: A Mendelian Randomization Study. *Ann. Neurol.* 87 (3), 434–441. doi:10.1002/ana.25671

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

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

Copyright © 2022 Zhu, Li, Zhang and Huang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Identification of Immune-Related Hub Genes in Parkinson's Disease

Lin Chen¹, Yong Wang², Juan Huang¹, Binbin Hu¹ and Wei Huang^{1*}

¹Department of Neurology, The Second Affiliated Hospital of Nanchang University, Nanchang, China, ²Department of Oncology, The First Affiliated Hospital of Nanchang University, Nanchang, China

Background: Parkinson's disease (PD) is a common, age-related, and progressive neurodegenerative disease. Growing evidence indicates that immune dysfunction plays an essential role in the pathogenic process of PD. The objective of this study was to explore potential immune-related hub genes and immune infiltration patterns of PD.

Method: The microarray expression data of human postmortem substantia nigra samples were downloaded from GSE7621, GSE20141, and GSE49036. Key module genes were screened via weighted gene coexpression network analysis, and immune-related genes were intersected to obtain immune-key genes. Functional enrichment analysis was performed on immune-key genes of PD. In addition to, immune infiltration analysis was applied by a single-sample gene set enrichment analysis algorithm to detect differential immune cell types in the substantia nigra between PD samples and control samples. Least absolute shrinkage and selection operator analysis was performed to further identify immune-related hub genes for PD. Receiver operating characteristic curve analysis of the immune-related hub genes was used to differentiate PD patients from healthy controls. Correlations between immune-related hub genes and differential immune cell types were assessed.

Result: Our findings identified four hub genes (*SLC18A2*, *L1CAM*, *S100A12*, and *CXCR4*) and seven immune cell types (neutrophils, T follicular helper cells, myeloid-derived suppressor cells, type 1 helper cells, immature B cells, immature dendritic cells, and CD56 bright natural killer cells). The area under the curve (AUC) value of the four-gene-combined model was 0.92. The AUC values of each immune-related hub gene (*SLC18A2*, *L1CAM*, *S100A12*, and *CXCR4*) were 0.81, 0.78, 0.78, and 0.76, respectively.

Conclusion: In conclusion, *SLC18A2*, *L1CAM*, *S100A12*, and *CXCR4* were identified as being associated with the pathogenesis of PD and should be further researched.

Keywords: Parkinson's disease, weighted gene coexpression network analyses (WGCNA), LASSO, immune cell, hub genes

OPEN ACCESS

Edited by:

Thayne Kowalski,
Centro Universitário Cesuca, Brazil

Reviewed by:

Gerda Cristal Villalba Silva,
Hospital de Clínicas de Porto Alegre,
Brazil

Paul C. Marcogliese,
University of Manitoba, Canada

*Correspondence:

Wei Huang
13677080198@163.com

Specialty section:

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

Received: 07 April 2022

Accepted: 02 June 2022

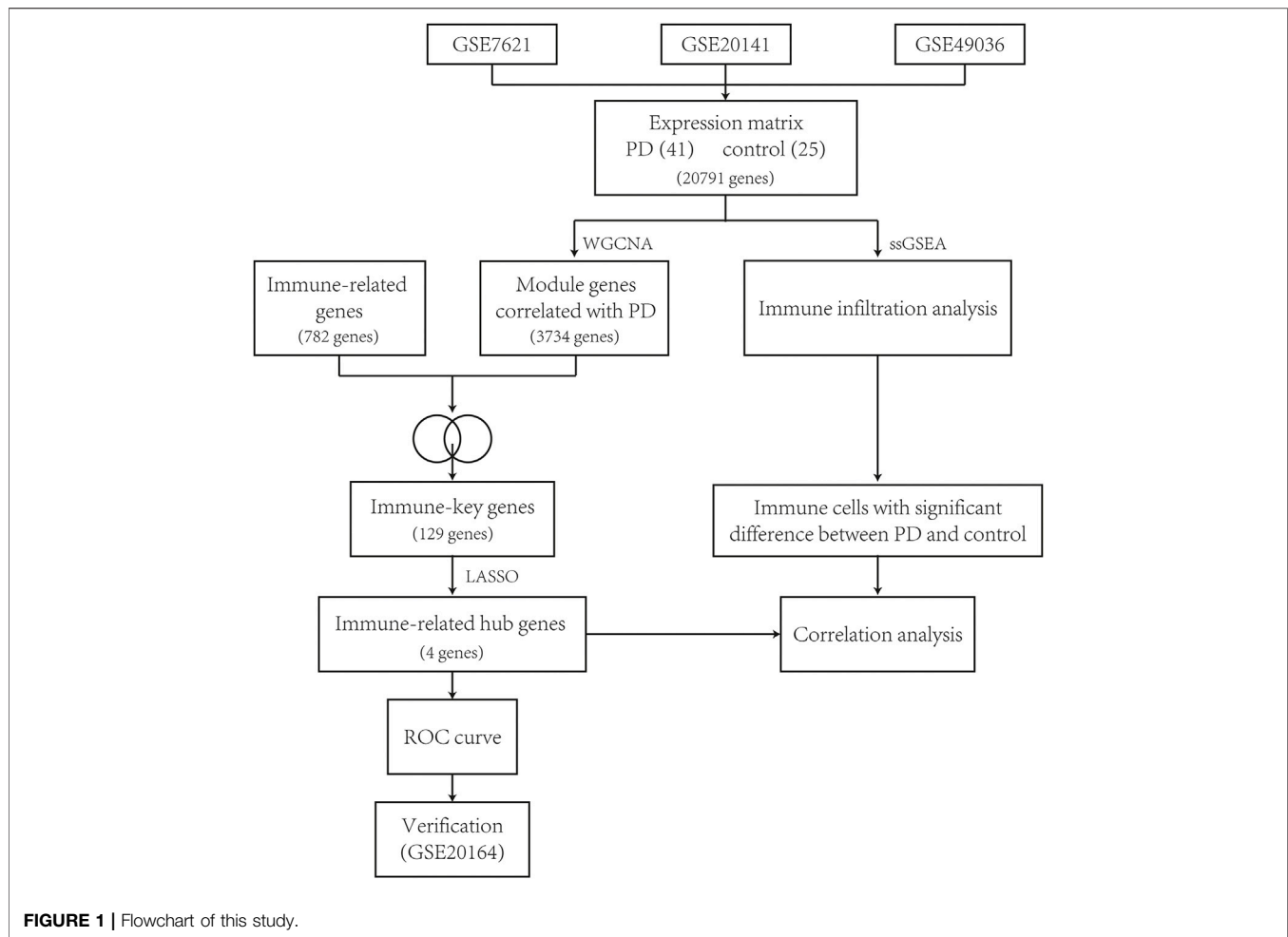
Published: 22 July 2022

Citation:

Chen L, Wang Y, Huang J, Hu B and
Huang W (2022) Identification of
Immune-Related Hub Genes in
Parkinson's Disease.
Front. Genet. 13:914645.
doi: 10.3389/fgene.2022.914645

INTRODUCTION

Parkinson's disease (PD) is a progressive neurodegenerative disease. The condition is characterized by motor symptoms, comprising bradykinesia, resting tremor, rigidity, and nonmotor symptoms, including olfactory loss, autonomic dysfunction, depression, cognitive impairment, and insomnia. In 2016, more than 6 million people suffered PD worldwide, and the number of PD patients is estimated



to increase with population age (GBD 2016 Parkinson's Disease Collaborators, 2018). The major pathological features of PD are progressive loss of dopaminergic neurons in the substantia nigra and aberrant α -synuclein aggregation called Lewy bodies (Kalia and Lang, 2015). The exact etiology of sporadic PD remains unknown.

Growing evidence supports that immune dysfunction plays an essential role in the pathogenic process of PD. A nationwide epidemiological study from Sweden involving 310,522 patients with autoimmune disorders showed that 932 patients developed subsequent PD after follow-up; patients with an autoimmune disease had a 33% overall excess risk of PD (Li et al., 2012). A nested case-control study published in 2022 demonstrated that people with rheumatoid arthritis using immunosuppressant treatments were associated with a potentially decreased risk of developing PD (Paakinaho et al., 2022). Imaging of neuroinflammation with [^{11}C] (R)-PK11195 PET showed that microglial activation is higher in PD patients than in healthy controls, confirming a link between neuroinflammation and the pathological process in PD (Gerhard et al., 2006). The levels of activated immune cells and proinflammatory cytokines are increased in the brain or cerebrospinal fluid of patients with PD (Mogi et al., 1996; Brodacki et al., 2008; Zhang et al., 2016;

Schröder et al., 2018; Lin et al., 2019) and in the substantia nigra of the MPTP-induced PD animal model (Hébert et al., 2003). Hence, regulation of immune function might provide a potential therapeutic strategy to improve the prognosis of PD.

We focused on identifying potential immune-related hub genes and immune infiltration patterns of PD. The immune-related hub genes and immune infiltration patterns of PD were identified by bioinformatics analysis in this article. Weighted gene coexpression network analysis (WGCNA) was performed to obtain key module genes. A dataset of immune-related genes was acquired from an article (<https://www.sciencedirect.com/science/article/pii/S2211124716317090>), which is representative of 28 peripheral immune cell types. The immune-related genes and key module genes were intersected to obtain immune-key genes, and then, functional enrichment analysis was applied to immune-key genes of PD. In addition to, immune infiltration analysis was performed between PD samples and control samples using the single-sample gene set enrichment analysis (ssGSEA) algorithm to explore the differential immune cell types in the substantia nigra. Least absolute shrinkage and selection operator (LASSO) analysis was used to discover the immune-related hub genes, and receiver operating characteristic (ROC) curve analysis of immune-related hub genes was used to differentiate PD

TABLE 1 | The number of samples for PD and controls in included datasets.

| Tissue | GSE | Platform | PD | Controls |
|------------------|----------|----------|----|----------|
| Substantia nigra | GSE7621 | GPL570 | 16 | 9 |
| | GSE20141 | | 10 | 8 |
| | GSE49036 | | 15 | 8 |

patients from healthy controls. The correlations between immune-related hub genes and differential immune cell types were analyzed. The expression levels of immune-related hub genes were verified in GSE20164. The immune-related hub genes and immune infiltration patterns of PD can be regarded as new therapeutic targets for PD.

MATERIALS AND METHODS

Data Processing

The workflow of this study is described as a flowchart in **Figure 1**.

The RNA microarray data of human postmortem substantia nigra samples were downloaded from the GSE7621, GSE20141, and GSE49036 datasets of the Gene Expression Omnibus (<https://www.ncbi.nlm.nih.gov/geo/>). These datasets were analyzed using Affymetrix Human Genome U133 Plus 2.0 Array on the platform of GPL570 (HG-U133_Plus_2). There were a total of 66 samples, including 41 samples from PD patients and 25 samples from healthy controls as a control group (**Table 1**). Using the ComBat function from the “sva” package of R software version 4.1.0, batch effects of the expression data were removed.

Weighted Gene Coexpression Network Analysis

All genes (20,791 genes) of the datasets were analyzed using the “WGCNA” package. First, the variances of all genes were calculated and ranked from high to low, and genes with variance beyond quartiles of all variances were chosen for further analysis (5,198 genes). In a subsequent manner, we performed hierarchical cluster analysis of all samples to remove the abnormal samples. The Pearson correlation coefficient between every two genes was calculated to construct a similarity matrix. Then, we chose the soft thresholding power value from 1 to 20 based on the pickSoftThreshold function to build scale-free topology. We checked the scale-free topology based on the connection degrees k and $p(k)$. Via dynamic tree cutting to construct coexpression modules, the highly correlated genes were classified into the same module, and each module contained at least 30 genes. The module eigengene (ME) value of each module was calculated. Furthermore, the correlation coefficient and p value between the ME value and phenotype of clinical traits (type of disease, PD vs. control) were calculated. When the p value was less than 0.05, the modules associated with PD were considered key modules. Module membership (MM) was the correlation between a gene and its module. Gene significance (GS) is the relationship between a gene and a clinical trait.

Immune Infiltration Analysis

ssGSEA was used to quantify the infiltration levels of 28 immune cells by converting each sample's gene expression profile into an immune gene set enrichment profile. Both the PD and control samples were fit into immune infiltration analysis via the ssGSEA algorithm using the “GSEA” R package to calculate the infiltration abundance of immune cells in the substantia nigra. The differential immune cells between the two groups were screened (p value < 0.05).

Identification of Immune-Key Genes and Functional Enrichment Analysis

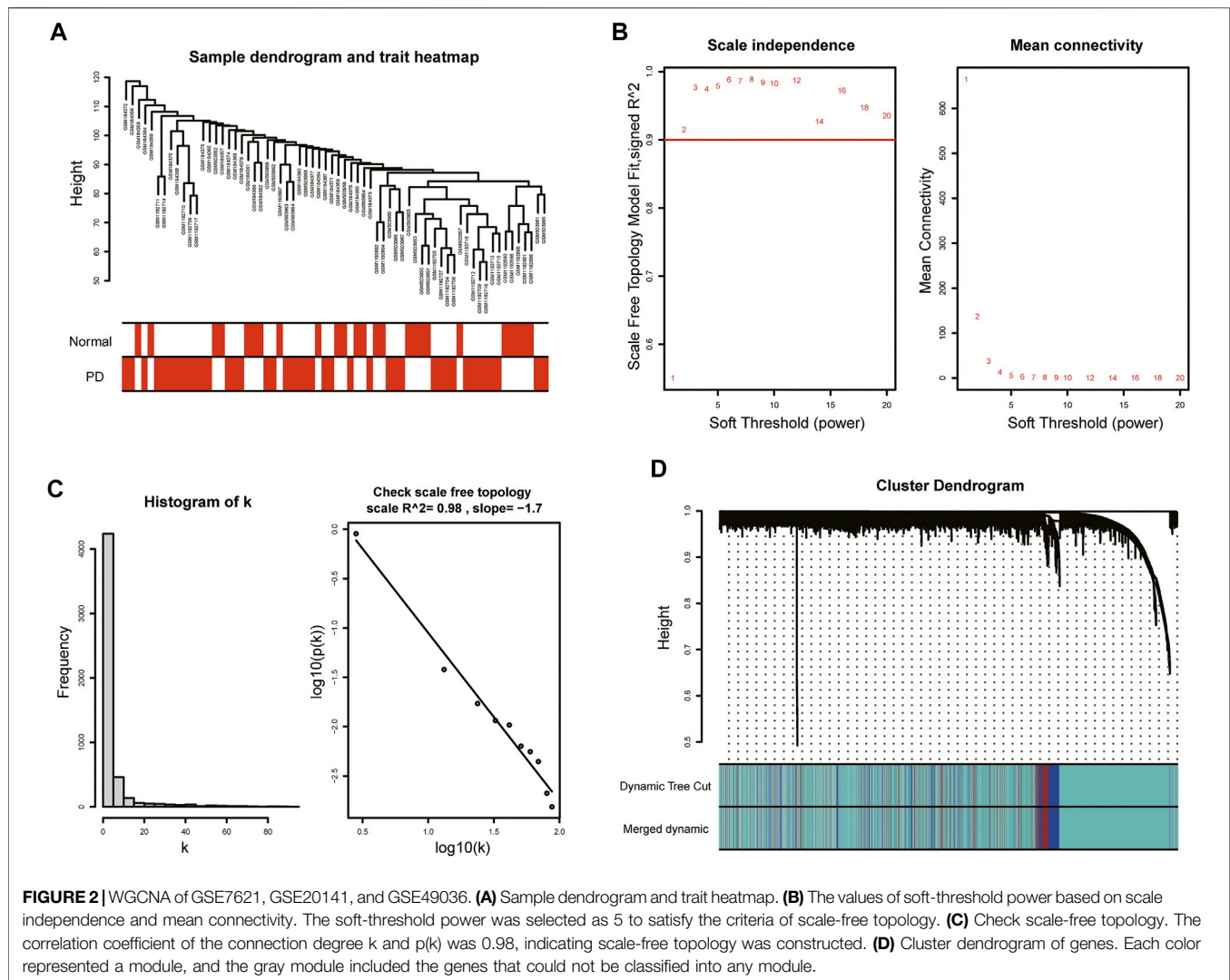
A widely recognized immune-related gene dataset of 28 peripheral immune cell types was used to investigate (Charoentong et al., 2017). This dataset included 782 immune-related genes. Immune-related genes and key module genes screened via WGCNA were intersected to obtain immune-key genes. Functional enrichment analysis was performed on immune-key genes to explore the potential biological implications. GO enrichment and KEGG pathway analysis were performed using the “clusterProfiler” and “GOplot” R packages. The BH method was used for p value adjustment in both GO enrichment and KEGG pathway analysis. The results of functional enrichment analysis were considered significantly enriched if the p adjust value < 0.05 . In addition to, the top 5 GO terms and top 10 KEGG pathways are shown visually in the bubble chart.

Identification of Immune-Related Hub Genes Using LASSO Logistic Regression

LASSO analysis was used to select the best features for high-dimensional data on account of its strong predictive value and low correlation. We used LASSO analysis to further identify immune-related hub genes for PD. The LASSO model was established using the “glmnet” R package, which could distinguish PD patients from controls. The expression levels of immune-key genes and clinical traits (type of disease, PD vs. control) were applied to build a LASSO logistic regression. ROC curve analysis of the immune-related hub genes was used to differentiate PD patients from healthy controls using the “pROC” R package. Furthermore, the correlations between immune-related hub genes and differential immune cell types were assessed.

Validation of Hub Genes

To verify the expression difference of the immune-related hub genes between PD and controls, the GSE20164 dataset was used for validation. There were six PD patients and five controls in this dataset. The expression difference of hub genes in GSE20164 is shown with a boxplot using the draw_boxplot function “tinyarray” package. The Kruskal–Wallis test was used to compare the expression levels of hub genes between the PD and control samples. Statistical significance was set at $p < 0.05$.



RESULTS

Identification of Gene Coexpression Modules

With the purpose of more precise following analysis, 20,791 genes were obtained after removal of batch effects using ComBat, an empirical Bayes method (**Supplementary Figure S1**). The top 5,198 most varied genes from a total of 66 samples were selected to construct the coexpression network. All samples passed the cut-off line with a height of 120 followed by hierarchical clustering, and the clinical characteristic heatmap was drawn (**Figure 2A**). The soft thresholding power was selected as 5 based on the criteria of scale-free topology, with a scale-free R^2 value of 0.98 and a slope value of -1.7 (**Figures 2B,C**). Modules with divergences of less than 25% and fewer than 30 genes were merged into larger modules. Finally, four coexpression modules were determined (**Figure 2D**). The gray module consisting of non-coexpressed genes was considered an invalid module, which was excluded from the following analysis. There were

3,734 genes in the turquoise module, 521 genes in the blue module, 343 genes in the brown module, and 600 genes in the gray module.

Calculation of Module-Trait Correlations

According to the ME values of the obtained modules, the correlations between these modules and clinical traits (PD vs. control) were performed. The turquoise module revealed the highest correlation with PD ($r = -0.31$, $p = 0.01$) and was selected as the key module for further analysis (**Figure 3A**). The relationship between MM and GS was evaluated in the key modules, for which the correlation coefficient was 0.61 ($p < 0.001$), as depicted in **Figure 3B**. The heatmap of the eigengene network suggested that the turquoise module is highly related to the clinical trait status (**Figure 3C**).

Immune Infiltration Analysis in PD

The infiltration abundance of 28 immune cell types in the substantia nigra between PD and control samples was calculated by the ssGSEA algorithm. The results revealed that

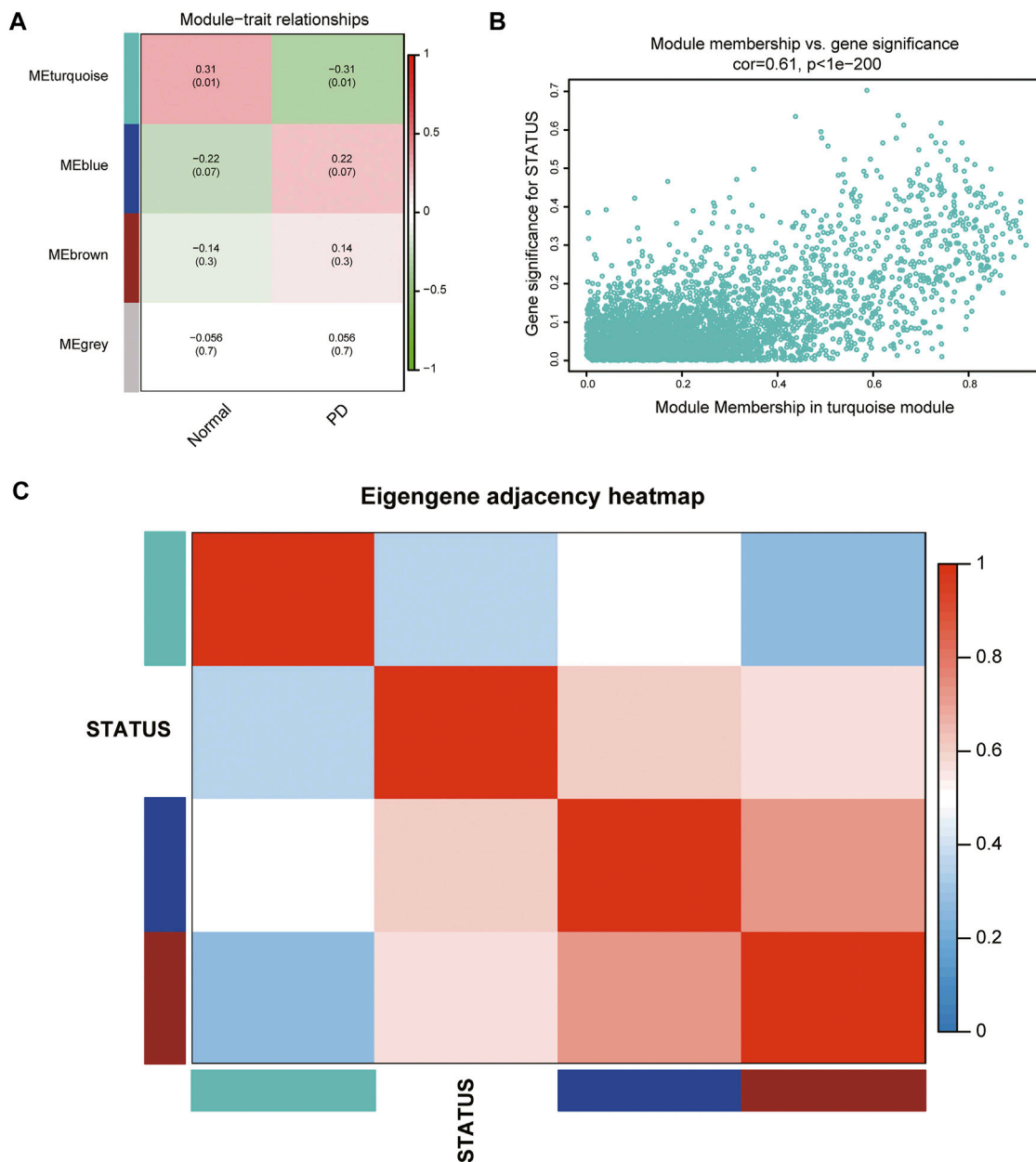


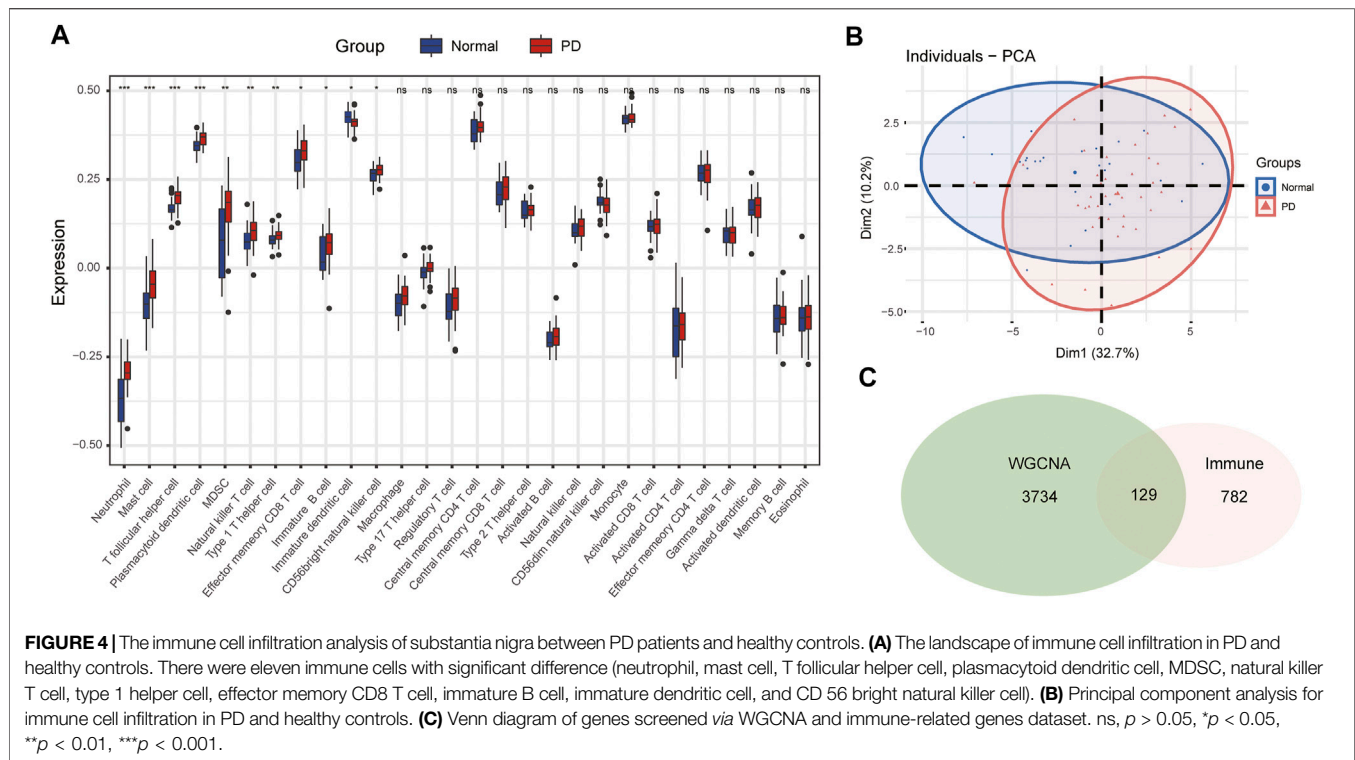
FIGURE 3 | Identification of key modules associated with PD. **(A)** Heatmap of correlations between MEs and phenotype of clinical traits (type of disease). Red represented positive correlation and green represented negative correlation, and the corresponding p value was indicated in brackets. **(B)** GS and MM in the turquoise module. **(C)** Heatmap of the eigengene network representing the relationships among the modules and the clinical trait status.

neutrophils, mast cells, T follicular helper (Tfh) cells, plasmacytoid dendritic cells (DCs), myeloid-derived suppressor cells (MDSCs), natural killer (NK) T cells, type 1 helper (Th1) cells, effector memory CD8 T cells, immature B cells, immature DCs, and CD56 bright NK cells were significantly different in PD samples compared with control samples (**Figure 4A**). The infiltration abundance of immature DCs decreased in PD samples among differential immune cell types, whereas others increased in samples of PD compared to control samples. Furthermore, we performed principal component analysis

(PCA) of immune cell infiltration among 66 substantia nigra samples. The PCA plot suggested that immune cell infiltration in the PD group was significantly different from that in the control group (**Figure 4B**).

Identification of Immune-Key Genes and Functional Enrichment Analysis

The immune-related gene dataset of 28 peripheral immune cell types was used for analysis, and 782 immune-related genes are



shown in Supplementary dataset 1. A total of 129 immune-key genes were acquired via the overlap of immune-related genes and the key module genes (turquoise module), as displayed in **Figure 4C**. Then, GO enrichment and KEGG pathway enrichment analyses were performed on immune-key genes. GO enrichment analysis revealed that these genes were largely enriched in biological functions related to immune activities, such as regulation of leukocyte cell–cell adhesion and T-cell activation and cytokine binding (**Figure 5A**). KEGG pathway enrichment analysis showed that these genes were mainly involved in immune pathways, such as cell adhesion molecules and Th17-, Th1-, and Th2-cell differentiation (**Figure 5B**).

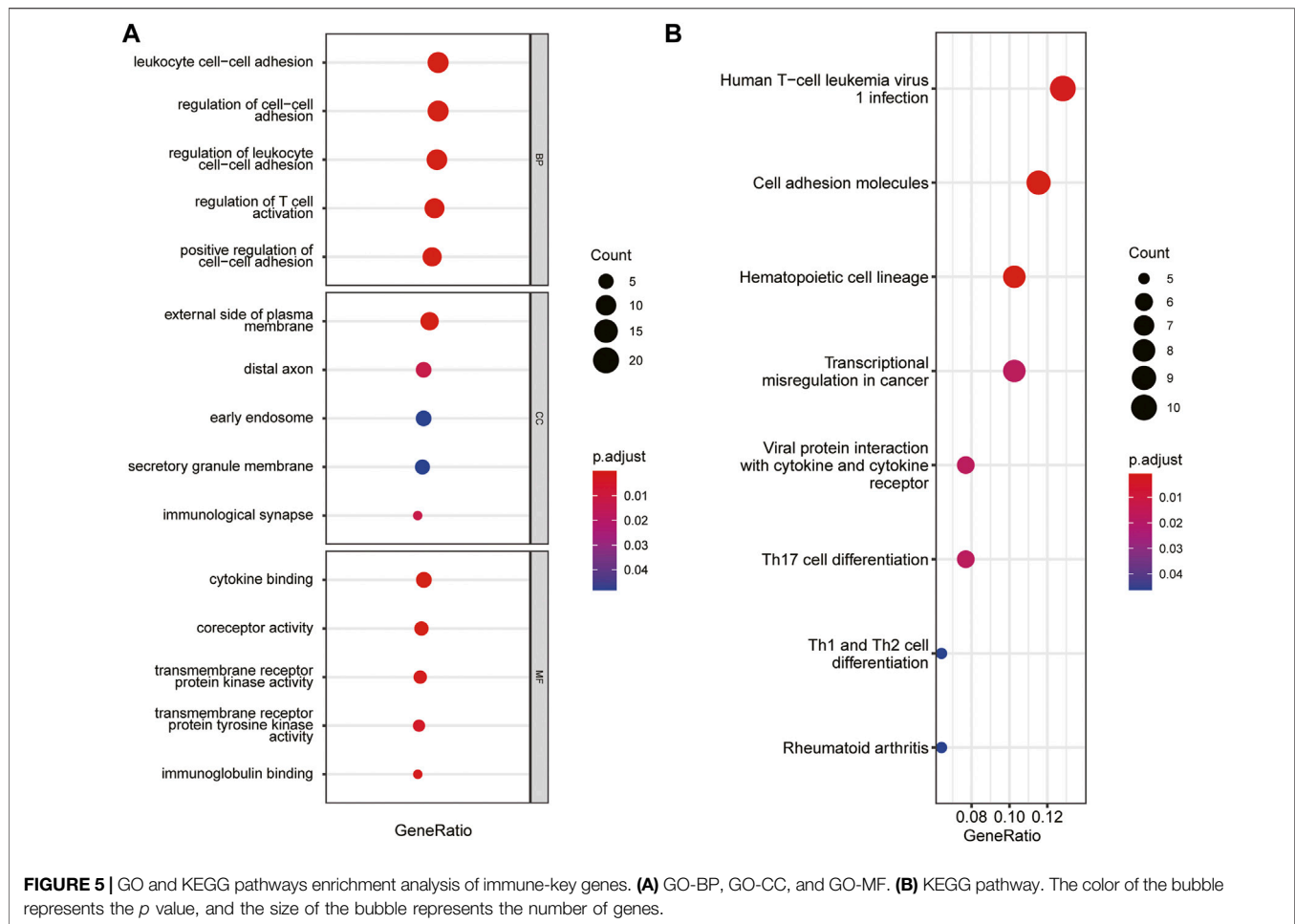
Immune-Related Hub Gene Screening and Correlation Analysis of Hub Genes and Immune Cells

To identify immune-related hub genes for PD, the expression levels of the immune-key genes and clinical traits (PD vs. control) in all samples were used to build the LASSO model (**Figure 6A**). A total of 129 immune-key genes were fit into LASSO logistic regression. According to the value of the lambda minimum criteria, four immune-related hub genes (*SLC18A2*, *LICAM*, *S100A12*, and *CXCR4*) were identified to have nonzero regression coefficients (**Figure 6B**). The relationship among immune-related hub genes showed that *SLC18A2* was positively correlated with *LICAM* and negatively correlated with *S100A12* and *CXCR4*; *LICAM* was negatively correlated with *S100A12* and *CXCR4*; and *S100A12* was positively

correlated with *CXCR4* (**Figure 6C**). The correlations between 4 immune-related hub genes and 11 differential immune cell types were assessed (**Figure 6D**). *SLC18A2* had a strong correlation with immature DCs (correlation coefficient = 0.82) and a moderate correlation with MDSCs, neutrophils, and Tfh cells (correlation coefficients were -0.35 , -0.39 , and -0.32 , respectively). *LICAM* had a moderate correlation with immature B cells, immature DCs, MDSCs, Tfh cells, and Th1 cells (correlation coefficients were -0.40 , 0.56 , -0.37 , -0.42 , and 0.34 , respectively). There was a strong correlation between *S100A12* and neutrophils (correlation coefficient = 0.72) and a moderate relationship with immature DCs, MDSCs, and Tfh cells (correlation coefficients were -0.31 , 0.35 , and 0.38 , respectively). *CXCR4* was moderately related to MDSCs (correlation coefficient = 0.40).

ROC Curve Analysis of Immune-Related Hub Genes

We compared the expression levels of immune-related hub genes in 66 samples, including 41 PD and 25 healthy control samples. The expression levels of *SLC18A2* and *LICAM* were significantly downregulated, and *S100A12* and *CXCR4* were significantly upregulated in PD samples compared with control samples (**Figure 7A**). ROC curve analysis of the immune-related hub genes was performed to differentiate PD patients from healthy controls. As shown in **Figure 7B**, the area under the curve (AUC) value of the four-gene-combined model was 0.92. The AUC values of each immune-related hub gene (*SLC18A2*, *LICAM*,



S100A12, and *CXCR4*) were 0.81, 0.78, 0.78, and 0.76, respectively (Figures 7C–F).

Validation of Immune-Related Hub Genes in GSE20164

The expression levels of immune-related hub genes were verified in GSE20164. In GSE20164, there were six PD patients and five healthy controls. Figure 8 shows that the expression of *SLC18A2* and *LICAM* was significantly downregulated and *CXCR4* was significantly upregulated in PD patients. There was no significant difference in *S100A12*.

DISCUSSION

PD was previously regarded as a movement disorder, but PD is now thought to be a multisystem disorder with notable neuroinflammation and immune dysfunction. Inflammatory manifestations that have been identified in PD patients include intestinal dysbiosis and inflammation, elevated circulating proinflammatory cytokine levels, innate and adaptive immune cell activation, blood–brain barrier breakdown allowing

peripheral immune cell infiltration of the central nervous system, and chronic neuroinflammation (Tansey et al., 2022). In this study, via bioinformatics analysis, we explored the potential immune-related hub genes and immune infiltration patterns in the substantia nigra of PD patients, which provided new potential therapeutic biomarkers for PD.

Three datasets, GSE7621, GSE20141, and GSE49036, were included in our research. WGCNA established four distinct coexpression modules, among which the turquoise module was significantly related to the pathogenesis of PD. There were 3,734 genes in the turquoise module. The overlap of 782 immune-related genes and the turquoise module genes included 129 immune-key genes. GO and KEGG enrichment analyses demonstrated that immune-key genes focused on immune activities and pathways, including regulation of leukocyte cell–cell adhesion, regulation of T-cell activation, and Th17-, Th1-, and Th2-cell differentiation.

The immune infiltration analysis revealed eleven differential infiltrative immune cell types in the substantia nigra between PD and control samples: neutrophils, mast cells, Tfh cells, plasmacytoid DCs, MDSCs, NK T cells, Th1 cells, effector memory CD8 T cells, immature B cells, immature DCs, and CD56 bright NK cells. Based on LASSO logistic regression, four

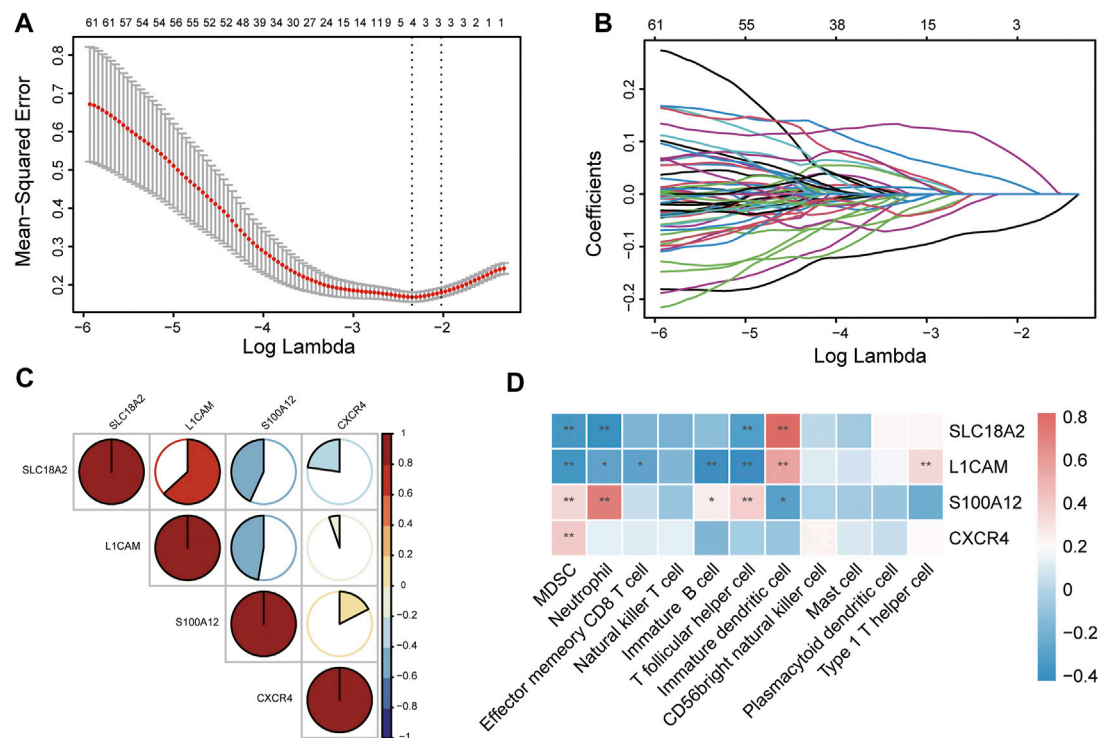


FIGURE 6 | Identification of immune-related hub genes via LASSO model. **(A)** Tuning parameter (lambda) selection in the LASSO regression model. The vertical lines were drawn at the optimal values by minimum criteria and 1-SE criteria, and we selected minimum criteria to construct the model. **(B)** The LASSO coefficient profiles. **(C)** The relationship among immune-related hub genes. **(D)** The relationship between immune-related hub genes and immune cells. * $p < 0.05$, ** $p < 0.01$.

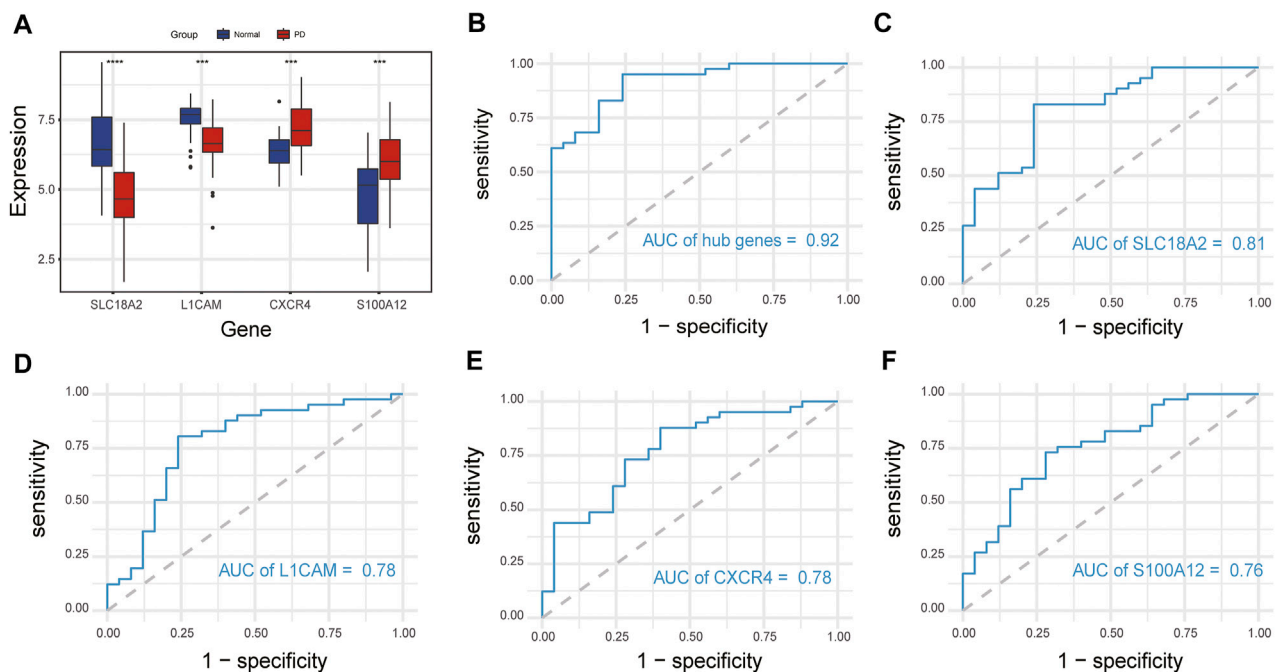
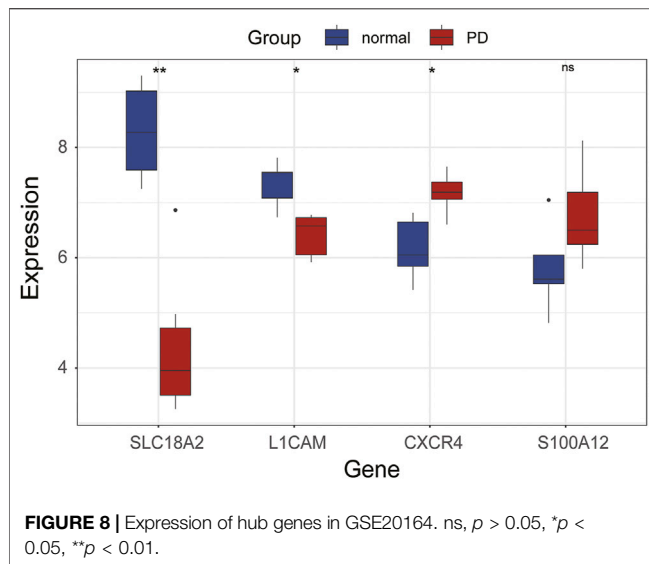


FIGURE 7 | Expression of immune-related hub genes and ROC curve of immune-related hub genes. **(A)** Expression of hub genes in 41 PD and 25 healthy control samples. **(B)** The ROC curve of four immune-related hub genes. **(C-F)** The ROC curve of each immune-related hub gene. ns, $p > 0.05$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$.



hub genes were finally screened: *SLC18A2*, *L1CAM*, *S100A12*, and *CXCR4*.

SLC18A2 (also known as VMAT2, vesicular monoamine transporter 2) is expressed in both the peripheral and central nervous systems. VMAT2 takes up dopamine into intracellular vesicles. It was reported that dysfunction of VMAT2 proteins can result in cytoplasmic dopamine accumulation and lead to dopaminergic neuron death. Furthermore, VMAT2 mRNA levels were significantly reduced in PD patients versus healthy controls (Sala et al., 2010). *L1CAM* (cell adhesion molecule L1) is a transmembrane protein expressed in the brain and peripheral nerves that plays an important role in the development of the nervous system, cell adhesion, and synaptic plasticity. A study showed that the expression levels of *L1CAM* mRNA decreased in α -syn-treated cells (Sugeno et al., 2016). *S100A12* (S100 calcium binding protein A12) is an inflammation-associated protein expressed in neutrophils, macrophages, and epithelial cells and is related to Alzheimer's disease (Shepherd et al., 2006), brain injury (Petrone et al., 2017), cancer, and so on. *CXCR4*, which is expressed in the central nervous system, is a chemokine receptor implicated in immune activity, microglia recruitment, and neurodevelopmental processes. Emerging evidence indicates that *CXCR4* is associated with neurodegenerative diseases, such as PD and progressive supranuclear palsy (Bonham et al., 2018). Postmortem brains showed that the expression levels of *CXCR4* in the substantia nigra and striatum of PD patients were higher than those in controls, accompanied by an increase in activated microglia (Shimoji et al., 2009). Moreover, *CXCR4* expression in the circulating mononuclear cells of PD patients is increased in comparison with that in controls (Bagheri et al., 2018).

In addition, we performed ROC curve analysis of immune-related hub genes to differentiate PD patients from healthy controls. The AUC value of the four-gene-combined model was 0.92. The AUC values of each immune-related hub gene were all above 0.75, indicating that these four hub genes might be signature genes of PD.

The correlations between 4 hub genes and 11 immune cells were evaluated. We discovered that neutrophils, Tfh cells, MDSCs, Th1 cells, immature B cells, immature DCs, and CD56 bright NK cells had moderate or strong correlations with hub genes, indicating that these immune cells are closely related to the pathogenesis of PD.

Oxidative injury is a characteristic feature of neurodegenerative diseases. Peripheral blood neutrophils are the predominant sources of reactive oxygen species. A study reported that oxidative stress levels in circulating neutrophils were higher in patients with PD than in controls. This research also revealed that mitochondrial mass and function were altered in neutrophils of patients with PD (Vitte et al., 2004). In another study, neutrophil counts were higher in PD patients than in healthy persons (Akil et al., 2015).

Tfh cells, which are a subtype of CD4⁺ T cells, are relevant to B-cell differentiation, germinal center formation and humoral immune responses. Tfh cells in PD patients were significantly higher than those in controls (Zhao et al., 2020). In addition to, IL-4 was identified as a cytokine of Tfh cells and may participate in the degeneration of dopamine neurons in patients with PD (Bok et al., 2018).

MDSCs play an important role in the pathogenesis of cancers and inflammatory and autoimmune diseases. According to research findings, MDSCs inhibit inflammatory responses by suppressing CD4⁺ T-cell activation (Ostanin and Bhattacharya, 2013). The circulating MDSCs of newly diagnosed PD patients were higher than those of healthy controls and exhibited a pro-neuroinflammatory effect in PD (Chen et al., 2017).

Th1 cells are a subtype of CD4⁺ T cells. Flow cytometric analysis revealed that Th1 and regulatory T cells were increased in the midbrain tissue of a PD mouse model (Williams et al., 2021). The Th1 cytokines IFN γ and TNF have been shown to be increased in PD patient blood (Sulzer et al., 2017; Kustrimovic et al., 2018). A study published in 2009 suggested that CD4⁺ T and CD8⁺ T cells infiltrated specifically in the substantia nigra in the brains of PD patients (Brochard et al., 2009).

Peripheral immature B cells are an important member of the immune system, producing natural antibodies and regulating CD4⁺ T-cell responses (Lee et al., 2015). Dysfunction of immature B cells may contribute to autoimmune diseases. Emerging evidence suggests that B cells contribute to the pathogenesis of PD (Sabatino et al., 2019). B cells might be reduced in patients with PD (Bas et al., 2001; Stevens et al., 2012).

DCs are professional antigen-presenting cells related to the pathogenesis of neuroinflammation. DCs are recruited to the brain across the blood-brain barrier. A study revealed that immature DCs adhered to activated endothelial cells more avidly than mature DCs (Arjmandi et al., 2009).

NK cells are primarily divided into two major subsets (CD56 dim and CD56 bright) in humans. In the central nervous system, the majority of NK cells in cerebrospinal fluid are CD56 bright cell subsets (Han et al., 2014). CD56 bright NK cells modulate immune responses through cytokine production (Poli et al., 2009). Circulating NK cell counts increased in PD patients in contrast with non-PD controls (Mihara et al., 2008). In

addition to, NK cells can internalize and degrade α -syn aggregates (Earls and Lee, 2020).

In this study, we discovered four immune-related hub genes (*SLC18A2*, *LICAM*, *S100A12*, and *CXCR4*) and seven peripheral immune cell types (neutrophils, Tfh cells, MDSCs, Th1 cells, peripheral immature B cells, DCs, and NK cells) that are closely related to the pathogenesis of PD. The immune-related hub genes were mainly associated with biological pathways, including the dopaminergic neurotransmitter release cycle, developmental biology (such as axon guidance), and immune and inflammatory activity. Mounting evidence indicates that immune dysfunction is involved in the pathogenesis of PD. The physiological function of the axon guidance pathway includes neuronal network formation during central nervous system development and the maintenance and plasticity of neural synapses (Bae et al., 2021). Abnormal axon-guidance-molecule signaling can lead to loss of connectivity and eventually trigger PD (Lin et al., 2009; Tomiyama, 2011). In addition, peripheral immune cells infiltrate the brain, possibly because of blood-brain barrier breakdown, which may affect neuroinflammation in the central nervous system by modulating microglial and astrocyte functions. Microglia and astrocytes are central to neuronal function and health. Activation of microglia can transform neuroprotective astrocytes to neurotoxic astrocytes, inducing loss of neuroprotective function, and overactivated microglia can result in cerebral inflammation and neuronal injury (Tan et al., 2020).

There were several limitations in our study that should be acknowledged. First, our findings require *in vitro* experiments to verify the results. Second, samples from early-stage or prodromal PD patients and animal models are required to explore immune-related hub genes, which can provide diagnostic biomarkers and timely drug intervention. In future investigations, we will focus on the mechanism of *SLC18A2*, *LICAM*, *S100A12*, and *CXCR4* and explore early diagnostic biomarkers in PD.

CONCLUSION

Using bioinformatics analysis, immune-related hub genes in the substantia nigra of PD patients were identified. *SLC18A2*,

LICAM, *S100A12*, and *CXCR4* were regarded as candidate genes for further investigation. Our study provides immune-related genes involved in the pathogenesis of PD and promising therapeutic targets for PD.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Materials**, and further inquiries can be directed to the corresponding author.

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

LC and YW designed the study. LC, YW, JH, and BH performed the data analysis and wrote the manuscript. All authors reviewed and approved the final manuscript for publication.

FUNDING

Clinical research project of the Second Affiliated Hospital of Nanchang University, 2021efyC03.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2022.914645/full#supplementary-material>

REFERENCES

- Akil, E., Bulut, A., Kaplan, İ., Özdemir, H. H., Arslan, D., and Aluçlu, M. U. (2015). The Increase of Carcinoembryonic Antigen (CEA), High-Sensitivity C-Reactive Protein, and Neutrophil/Lymphocyte Ratio in Parkinson's Disease. *Neurol. Sci.* 36 (3), 423–428. doi:10.1007/s10072-014-1976-1
- Arjmandi, A., Liu, K., and Dorovini-Zis, K. (2009). Dendritic Cell Adhesion to Cerebral Endothelium: Role of Endothelial Cell Adhesion Molecules and Their Ligands. *J. Neuropathol. Exp. Neurol.* 68 (3), 300–313. doi:10.1097/nen.0b013e31819a8dd1
- Bae, Y. J., Kim, J.-M., Sohn, C.-H., Choi, J.-H., Choi, B. S., Song, Y. S., et al. (2021). Imaging the Substantia Nigra in Parkinson Disease and Other Parkinsonian Syndromes. *Radiology* 300 (2), 260–278. doi:10.1148/radiol.2021203341
- Bagheri, V., Khorramdelazad, H., Hassanshahi, G., Moghadam-Ahmadi, A., and Vakilian, A. (2018). CXCL12 and CXCR4 in the Peripheral Blood of Patients with Parkinson's Disease. *Neuroimmunomodulation* 25 (4), 201–205. doi:10.1159/000494435
- Bas, J., Calopa, M., Mestre, M., Molleví, D. G., Cutillas, B., Ambrosio, S., et al. (2001). Lymphocyte Populations in Parkinson's Disease and in Rat Models of Parkinsonism. *J. Neuroimmunol.* 113 (1), 146–152. doi:10.1016/s0165-5728(00)00422-7
- Bok, E., Cho, E. J., Chung, E. S., Shin, W.-H., and Jin, B. K. (2018). Interleukin-4 Contributes to Degeneration of Dopamine Neurons in the Lipopolysaccharide-Treated Substantia Nigra *In Vivo*. *Exp. Neurol.* 27 (4), 309–319. doi:10.5607/en.2018.27.4.309
- Bonham, L. W., Karch, C. M., Karch, C. M., Fan, C. C., Tan, C., Geier, E. G., et al. (2018). CXCR4 Involvement in Neurodegenerative Diseases. *Transl. Psychiatry* 8 (1), 73. doi:10.1038/s41398-017-0049-7
- Brochard, V., Combadière, B., Prigent, A., Laouar, Y., Perrin, A., Beray-Berthet, V., et al. (2009). Infiltration of CD4+ Lymphocytes into the Brain Contributes to Neurodegeneration in a Mouse Model of Parkinson Disease. *J. Clin. Invest.* 119 (1), 182–192. doi:10.1172/JCI36470
- Brodacki, B., Staszewski, J., Toczyłowska, B., Kozłowska, E., Drela, N., Chalimoniuk, M., et al. (2008). Serum Interleukin (IL-2, IL-10, IL-6, IL-4), TNF α , and INF γ Concentrations Are Elevated in Patients with Atypical and

- Idiopathic Parkinsonism. *Neurosci. Lett.* 441 (2), 158–162. doi:10.1016/j.neulet.2008.06.040
- Charoentong, P., Finotello, F., Angelova, M., Mayer, C., Efremova, M., Rieder, D., et al. (2017). Pan-Cancer Immunogenomic Analyses Reveal Genotype-Immunophenotype Relationships and Predictors of Response to Checkpoint Blockade. *Cell Rep.* 18 (1), 248–262. doi:10.1016/j.celrep.2016.12.019
- Chen, S., Liu, Y., Niu, Y., Xu, Y., Zhou, Q., Xu, X., et al. (2017). Increased Abundance of Myeloid-Derived Suppressor Cells and Th17 Cells in Peripheral Blood of Newly-Diagnosed Parkinson's Disease Patients. *Neurosci. Lett.* 648, 21–25. doi:10.1016/j.neulet.2017.03.045
- Earls, R. H., and Lee, J.-K. (2020). The Role of Natural Killer Cells in Parkinson's Disease. *Exp. Mol. Med.* 52 (9), 1517–1525. doi:10.1038/s12276-020-00505-7
- GBD 2016 Parkinson's Disease Collaborators (2018). Global, Regional, and National Burden of Parkinson's Disease, 1990–2016: A Systematic Analysis for the Global Burden of Disease Study 2016. *Lancet Neurol.* 17 (11), 939–953. doi:10.1016/S1474-4422(18)30295-3
- Gerhard, A., Pavese, N., Hotton, G., Turkheimer, F., Es, M., Hammers, A., et al. (2006). *In Vivo* imaging of Microglial Activation with [¹¹C](R)-PK11195 PET in Idiopathic Parkinson's Disease. *Neurobiol. Dis.* 21 (2), 404–412. doi:10.1016/j.nbd.2005.08.002
- Han, S., Lin, Y. C., Wu, T., Salgado, A. D., Mexhitaj, I., Wuest, S. C., et al. (2014). Comprehensive Immunophenotyping of Cerebrospinal Fluid Cells in Patients with Neuroimmunological Diseases. *J. Immunol.* 192 (6), 2551–2563. doi:10.4049/jimmunol.1302884
- Hébert, G., Arsaut, J., Dantzer, R., and Demotes-Mainard, J. (2003). Time-Course of the Expression of Inflammatory Cytokines and Matrix Metalloproteinases in the Striatum and Mesencephalon of Mice Injected with 1-Methyl-4-Phenyl-1,2,3,6-Tetrahydropyridine, a Dopaminergic Neurotoxin. *Neurosci. Lett.* 349 (3), 191–195. doi:10.1016/s0304-3940(03)00832-2
- Kalia, L. V., and Lang, A. E. (2015). Parkinson's Disease. *Lancet* 386 (9996), 896–912. doi:10.1016/s0140-6736(14)61393-3
- Kustrimovic, N., Comi, C., Magistrelli, L., Rasini, E., Legnaro, M., Bombelli, R., et al. (2018). Parkinson's Disease Patients Have a Complex Phenotypic and Functional Th1 Bias: Cross-Sectional Studies of CD4+ Th1/Th2/T17 and Treg in Drug-Naïve and Drug-Treated Patients. *J. Neuroinflammation* 15 (1), 205. doi:10.1186/s12974-018-1248-8
- Lee, J., Sim, J.-H., and Kim, I.-J. (2015). Peripheral Immature B Cells: Modulators of Autoimmunity. *Int. J. Rheum. Dis.* 18 (2), 200–207. doi:10.1111/1756-185x.12432
- Li, X., Sundquist, J., and Sundquist, K. (2012). Subsequent Risks of Parkinson Disease in Patients with Autoimmune and Related Disorders: A Nationwide Epidemiological Study from Sweden. *Neurodegener. Dis.* 10 (1–4), 277–284. doi:10.1159/000333222
- Lin, C.-H., Chen, C.-C., Chiang, H.-L., Liou, J.-M., Chang, C.-M., Lu, T.-P., et al. (2019). Altered Gut Microbiota and Inflammatory Cytokine Responses in Patients with Parkinson's Disease. *J. Neuroinflammation* 16 (1), 129. doi:10.1186/s12974-019-1528-y
- Lin, L., Lesnick, T. G., Maragone, D. M., and Isacson, O. (2009). Axon Guidance and Synaptic Maintenance: Preclinical Markers for Neurodegenerative Disease and Therapeutics. *Trends Neurosci.* 32 (3), 142–149. doi:10.1016/j.tins.2008.11.006
- Mihara, T., Nakashima, M., Kuroiwa, A., Akitake, Y., Ono, K., Hosokawa, M., et al. (2008). Natural Killer Cells of Parkinson's Disease Patients are Set up for Activation: A Possible Role for Innate Immunity in the Pathogenesis of This Disease. *Park. Relat. Disord.* 14 (1), 46–51. doi:10.1016/j.parkreldis.2007.05.013
- Mogi, M., Harada, M., Narabayashi, H., Inagaki, H., Minami, M., and Nagatsu, T. (1996). Interleukin (IL)-1 β , IL-2, IL-4, IL-6 and Transforming Growth Factor- α Levels Are Elevated in Ventricular Cerebrospinal Fluid in Juvenile Parkinsonism and Parkinson's Disease. *Neurosci. Lett.* 211 (1), 13–16. doi:10.1016/0304-3940(96)12706-3
- Ostanin, D. V., and Bhattacharya, D. (2013). Myeloid-Derived Suppressor Cells in the Inflammatory Bowel Diseases. *Inflamm. Bowel Dis.* 19 (11), 2468–2477. doi:10.1097/mib.0b013e3182902b11
- Paakinaho, A., Koponen, M., Tiihonen, M., Kauppi, M., Hartikainen, S., and Tolppanen, A.-M. (2022). Disease-Modifying Antirheumatic Drugs and Risk of Parkinson Disease: Nested Case-Control Study of People with Rheumatoid Arthritis. *Neurology* 98 (12), e1273–e1281. doi:10.1212/wnl.0000000000013303
- Petrone, A. B., Gionis, V., Giersch, R., and Barr, T. L. (2017). Immune Biomarkers for the Diagnosis of Mild Traumatic Brain Injury. *NeuroRehabilitation* 40 (4), 501–508. doi:10.3233/nre-171437
- Poli, A., Michel, T., Thérèse, M., Andrès, E., Hentges, F., and Zimmer, J. (2009). CD56bright Natural Killer (NK) Cells: an Important NK Cell Subset. *Immunology* 126 (4), 458–465. doi:10.1111/j.1365-2567.2008.03027.x
- Sabatino, J. J., Jr, Pröbstel, A.-K., and Zamvil, S. S. (2019). B Cells in Autoimmune and Neurodegenerative Central Nervous System Diseases. *Nat. Rev. Neurosci.* 20 (12), 728–745. doi:10.1038/s41583-019-0233-2
- Sala, G., Brighina, L., Saracchi, E., Fermi, S., Riva, C., Carrozza, V., et al. (2010). Vesicular Monoamine Transporter 2 mRNA Levels are Reduced in Platelets from Patients with Parkinson's Disease. *J. Neural Transm.* 117 (9), 1093–1098. doi:10.1007/s00702-010-0446-z
- Schröder, J. B., Pawlowski, M., Meyer Zu Hörste, G., Gross, C. C., Wiendl, H., Meuth, S. G., et al. (2018). Immune Cell Activation in the Cerebrospinal Fluid of Patients with Parkinson's Disease. *Front. Neurol.* 9, 1081. doi:10.3389/fneur.2018.01081
- Shepherd, C. E., Goyette, J., Utter, V., Rahimi, F., Yang, Z., Geczy, C. L., et al. (2006). Inflammatory S100A9 and S100A12 Proteins in Alzheimer's Disease. *Neurobiol. Aging* 27 (11), 1554–1563. doi:10.1016/j.neurobiolaging.2005.09.033
- Shimoji, M., Pagan, F., Heaton, E. B., and Mocchetti, I. (2009). CXCR4 and CXCL12 Expression is Increased in the Nigro-Striatal System of Parkinson's Disease. *Neurotox. Res.* 16 (3), 318–328. doi:10.1007/s12640-009-9076-3
- Stevens, C. H., Rowe, D., Morel-Kopp, M. C., Orr, C., Russell, T., Ranola, M., et al. (2012). Reduced T Helper and B Lymphocytes in Parkinson's Disease. *J. Neuroimmunol.* 252 (1–2), 95–99. doi:10.1016/j.jneuroim.2012.07.015
- Sugeno, N., Jäckel, S., Voigt, A., Wassouf, Z., Schulze-Hentrich, J., and Kahle, P. J. (2016). α -Synuclein Enhances Histone H3 Lysine-9 Dimethylation and H3K9me2-Dependent Transcriptional Responses. *Sci. Rep.* 6, 36328. doi:10.1038/srep36328
- Sulzer, D., Alcalay, R. N., Garretti, F., Cote, L., Kanter, E., Agin-Lieb, J., et al. (2017). T Cells from Patients with Parkinson's Disease Recognize α -Synuclein Peptides. *Nature* 546 (7660), 656–661. doi:10.1038/nature22815
- Tan, E.-K., Chao, Y.-X., West, A., Chan, L.-L., Poewe, W., and Jankovic, J. (2020). Parkinson Disease and the Immune System - Associations, Mechanisms and Therapeutics. *Nat. Rev. Neurol.* 16 (6), 303–318. doi:10.1038/s41582-020-0344-4
- Tansey, M. G., Wallings, R. L., Houser, M. C., Herrick, M. K., Keating, C. E., and Joers, V. (2022). Inflammation and Immune Dysfunction in Parkinson Disease. *Nat. Rev. Immunol.*, 1–17. doi:10.1038/s41577-022-00684-6 [Online ahead of print].
- Tomiyama, H. (2011). A Commentary on Axon Guidance Pathway Genes and Parkinson's Disease. *J. Hum. Genet.* 56 (2), 102–103. doi:10.1038/jhg.2010.153
- Vitte, J., Michel, B. F., Bongrand, P., and Gastaut, J.-L. (2004). Oxidative Stress Level in Circulating Neutrophils is Linked to Neurodegenerative Diseases. *J. Clin. Immunol.* 24 (6), 683–692. doi:10.1007/s10875-004-6243-4
- Williams, G. P., Schonhoff, A. M., Jurkuvenaite, A., Gallups, N. J., Standaert, D. G., and Harms, A. S. (2021). CD4 T Cells Mediate Brain Inflammation and Neurodegeneration in a Mouse Model of Parkinson's Disease. *Brain* 144 (7), 2047–2059. doi:10.1093/brain/awab103
- Zhang, P., Shao, X.-Y., Qi, G.-J., Chen, Q., Bu, L.-L., Chen, L.-J., et al. (2016). Cdk5-Dependent Activation of Neuronal Inflammation in Parkinson's Disease. *Mov. Disord.* 31 (3), 366–376. doi:10.1002/mds.26488
- Zhao, X., Jin, T., Zheng, C., Ma, D., and Zhang, Y. (2020). Imbalance of Circulating Th1/Th2 Cells in Patients with Parkinson's Disease. *Front. Neurol.* 11, 572205. doi:10.3389/fneur.2020.572205

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

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

Copyright © 2022 Chen, Wang, Huang, Hu and Huang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



OPEN ACCESS

EDITED BY

Mariana Recamonde-Mendoza,
Federal University of Rio Grande do Sul,
Brazil

REVIEWED BY

Qihuan Yuan,
Shandong University, China
Deqin Geng,
The Affiliated Hospital of Xuzhou
Medical University, China

*CORRESPONDENCE

Wenshi Wei,
wenshiwei1999@163.com
Li Zhang,
lizhang_huadong@163.com

SPECIALTY SECTION

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

RECEIVED 02 May 2022

ACCEPTED 06 July 2022

PUBLISHED 04 August 2022

CITATION

Shu J, Yang L, Wei W and Zhang L (2022),
Identification of programmed cell
death-related gene signature and
associated regulatory axis in cerebral
ischemia/reperfusion injury.
Front. Genet. 13:934154.
doi: 10.3389/fgene.2022.934154

COPYRIGHT

© 2022 Shu, Yang, Wei and Zhang. This
is an open-access article distributed
under the terms of the [Creative
Commons Attribution License \(CC BY\)](#).
The use, distribution or reproduction in
other forums is permitted, provided the
original author(s) and the copyright
owner(s) are credited and that the
original publication in this journal is
cited, in accordance with accepted
academic practice. No use, distribution
or reproduction is permitted which does
not comply with these terms.

Identification of programmed cell death-related gene signature and associated regulatory axis in cerebral ischemia/reperfusion injury

Jun Shu, Lu Yang, Wenshi Wei* and Li Zhang*

Department of Neurology, Cognitive Disorders Center, Huadong Hospital Affiliated to Fudan University, Shanghai, China

Background: Numerous studies have suggested that programmed cell death (PCD) pathways play vital roles in cerebral ischemia/reperfusion (I/R) injury. However, the specific mechanisms underlying cell death during cerebral I/R injury have yet to be completely clarified. There is thus a need to identify the PCD-related gene signatures and the associated regulatory axes in cerebral I/R injury, which should provide novel therapeutic targets against cerebral I/R injury.

Methods: We analyzed transcriptome signatures of brain tissue samples from mice subjected to middle cerebral artery occlusion/reperfusion (MCAO/R) and matched controls, and identified differentially expressed genes related to the three types of PCD (apoptosis, pyroptosis, and necroptosis). We next performed functional enrichment analysis and constructed PCD-related competing endogenous RNA (ceRNA) regulatory networks. We also conducted hub gene analysis to identify hub nodes and key regulatory axes.

Results: Fifteen PCD-related genes were identified. Functional enrichment analysis showed that they were particularly associated with corresponding PCD-related biological processes, inflammatory response, and reactive oxygen species metabolic processes. The apoptosis-related ceRNA regulatory network was constructed, which included 24 long noncoding RNAs (lncRNAs), 41 microRNAs (miRNAs), and 4 messenger RNAs (mRNAs); the necroptosis-related ceRNA regulatory network included 16 lncRNAs, 20 miRNAs, and 6 mRNAs; and the pyroptosis-related ceRNA regulatory network included 15 lncRNAs, 18 miRNAs, and 6 mRNAs. Hub gene analysis identified hub nodes in each PCD-related ceRNA regulatory network and seven key regulatory axes in total, namely, lncRNA Malat1/miR-181a-5p/Mapt, lncRNA Malat1/miR-181b-5p/Mapt, lncRNA Neat1/miR-181a-5p/Mapt, and lncRNA Neat1/miR-181b-5p/Mapt for the apoptosis-related ceRNA regulatory network; lncRNA Neat1/miR-181a-5p/Tnf for the necroptosis-related ceRNA regulatory network; lncRNA Malat1/miR-181c-5p/Tnf for the pyroptosis-related ceRNA regulatory network; and lncRNA Malat1/miR-181a-5p for both necroptosis-related and pyroptosis-related ceRNA regulatory networks.

Conclusion: The results of this study supported the hypothesis that these PCD pathways (apoptosis, necroptosis, pyroptosis, and PANoptosis) and crosstalk among them might be involved in ischemic stroke and that the key nodes and regulatory axes identified in this study might play vital roles in regulating the above processes. This may offer new insights into the potential mechanisms underlying cell death during cerebral I/R injury and provide new therapeutic targets for neuroprotection.

KEYWORDS

apoptosis, pyroptosis, necroptosis, cerebral ischemia/reperfusion (I/R) injury, competing endogenous RNA (ceRNA) network

1 Introduction

Ischemic stroke is one of the leading causes of long-term severe disability and death worldwide, which is usually caused by a permanent or transient local reduction in blood supply to the brain (Campbell and Khatri, 2020; Mendelson and Prabhakaran, 2021). Currently, the most effective strategy for ischemic stroke patients is to restore cerebral blood flow in a timely manner through drugs and surgery (Herpich and Rincon, 2020). However, injury to brain tissue caused by ischemia and hypoxia is further aggravated following the short-term recovery of blood perfusion, which is known as cerebral ischemia/reperfusion (I/R) injury. The mechanism by which cerebral ischemia/reperfusion injury occurs has not been fully elucidated. Nonetheless, a growing body of evidence suggests that the overproduction of ROS and activation of inflammation and immune responses might be involved, which ultimately trigger cell death, including apoptosis, necroptosis, and pyroptosis (Eltzschig and Eckle, 2011; Jurcau and Simion, 2021). There is thus a need for a comprehensive understanding of the mechanisms underlying cell death during cerebral ischemia/reperfusion (I/R) injury to rescue injured cells, especially injured neurons in the brain, and seek new neuroprotective therapies.

Multiple cell death pathways are currently believed to be involved in cell death in ischemic stroke, among which apoptosis, pyroptosis, and necroptosis are three key programmed cell death (PCD) pathways (Tuo et al., 2022). Apoptosis can be triggered through the intrinsic and/or extrinsic pathway and may contribute to a significant proportion of neuron death following cerebral ischemia/reperfusion (Radak et al., 2017; Datta et al., 2020). Meanwhile, necroptosis is a newly discovered mechanism of cell death that is mainly regulated by receptor-interacting protein kinase 1 (RIPK1), receptor-interacting protein kinase 3 (RIPK3), and mixed-lineage kinase domain-like pseudokinase (MLKL) (Liao et al., 2020). Increasing studies have suggested that necroptosis participates in the pathogenesis of various diseases including ischemia stroke. Studies have also indicated that the inhibition of necroptosis can exert neuroprotective effects after cerebral I/R in mice by reducing cerebral infarct volume and improving motor and

cognitive function (Deng et al., 2019; Yao et al., 2021). Pyroptosis is a kind of inflammatory programmed cell death that is characterized by rapid plasma-membrane rupture and the release of proinflammatory intracellular contents as well as cytokines (Yu et al., 2021). Pyroptosis was reported to be triggered by certain inflammasomes and activating caspases and executed by gasdermin family members (Dong et al., 2018). Accumulating evidence has shown that these three PCD pathways participate in the pathogenesis of ischemic stroke and that their inhibition could attenuate ischemic brain injury (Tuo et al., 2022). Recently, further evidence has also shown significant crosstalk among the three PCD pathways (Banoth et al., 2020; Zheng et al., 2020; Karki et al., 2021). Against this background, the concept of PANoptosis was proposed, which is defined as a proinflammatory PCD pathway with key features of pyroptosis, apoptosis, and/or necroptosis that cannot be accounted for by any of these PCD pathways alone (Malireddi et al., 2020; Wang and Kanneganti, 2021). PANoptosis is regulated by the cytoplasmic multimeric protein complex called the PANoptosome, which can participate in the three PCD pathways in parallel (Samir et al., 2020). PANoptosis has been implicated in various conditions, including infection, sterile inflammation, and cancer (Karki et al., 2020; Zheng et al., 2020; Place et al., 2021). A recent study that collected, integrated, and analyzed reports on research on cerebral I/R indicated that PANoptosis is observed in ischemic brain injury (Yan et al., 2022). Despite efforts to reveal the role of PCD pathways in cerebral I/R injury, the mechanisms underlying the involvement of the three PCD pathways in cerebral I/R injury are extremely complicated and remain largely unknown.

In this study, we collected PCD (apoptosis, pyroptosis, and necroptosis)-related genes based on previous literature and related databases, and analyzed transcriptome signatures of brain tissue samples from mice subjected to middle cerebral artery occlusion/reperfusion (MCAO/R) and matched controls to identify differentially expressed genes related to the three types of PCD. We then performed functional enrichment analysis of these differentially expressed PCD-related genes and their potential regulatory axes to explore their potential biological functions and regulatory mechanisms. This bioinformatic

analysis might provide new insights into the potential mechanisms underlying cell death during cerebral I/R injury and new therapeutic targets for neuroprotection.

2 Materials and methods

2.1 Collection of datasets and programmed cell death-related genes

We searched the Gene Expression Omnibus (GEO) database (Barrett et al., 2013) (<https://www.ncbi.nlm.nih.gov/geo>) using the following terms: “cerebral ischemia–reperfusion OR cerebral ischemia OR ischemia stroke” AND “*Mus musculus*.” We included the gene expression profiling of adult mouse brain tissues after transient focal ischemia at 24 h of reperfusion and matched control samples. Then two datasets, GSE131193 and GSE58720, were downloaded for analysis. The dataset GSE131193 based on the GPL19057 platform is an mRNA high-throughput sequencing series that includes data on contralateral and ipsilateral brain tissues from mice subjected to transient middle cerebral artery occlusion (tMCAO) at different reperfusion timepoints (24 h and 7 days) and matched sham-operated mice. We selected a subset comprising three ipsilateral brain tissues after transient focal ischemia at 24 h of reperfusion and three matched sham-operated mice for analysis. The dataset GSE58720 based on the GPL10787 platform contains microarray gene expression data of brain tissue samples from three MCAO-operated mice at 24 h of reperfusion and three matched sham-operated mice.

For apoptosis-related genes (ARGs), 101 ARGs were downloaded from Reactome Pathway Database (<https://reactome.org/>) (Jassal et al., 2020) and two were extracted from the literature, thus 103 ARGs were collected (Supplementary Table S1); for necroptosis-related genes (NRGs), twenty-seven NRGs were downloaded from Reactome Pathway Database, eighty-two NRGs were extracted from the literature, after removing the overlapping genes, ninety-three NRGs were obtained (Supplementary Table S2); for pyroptosis-related genes (PRGs), twenty PRGs were downloaded from Reactome Pathway Database, sixty-seven PRGs were extracted from the literature, after removing the overlapping genes, seventy-eight PRGs were obtained for further study (Supplementary Table S3).

2.2 Screening strategy for differentially expressed programmed cell death-related genes

Differentially expressed genes (DEGs) of the microarray dataset GSE58720 were identified with NCBI's GEO2R tool

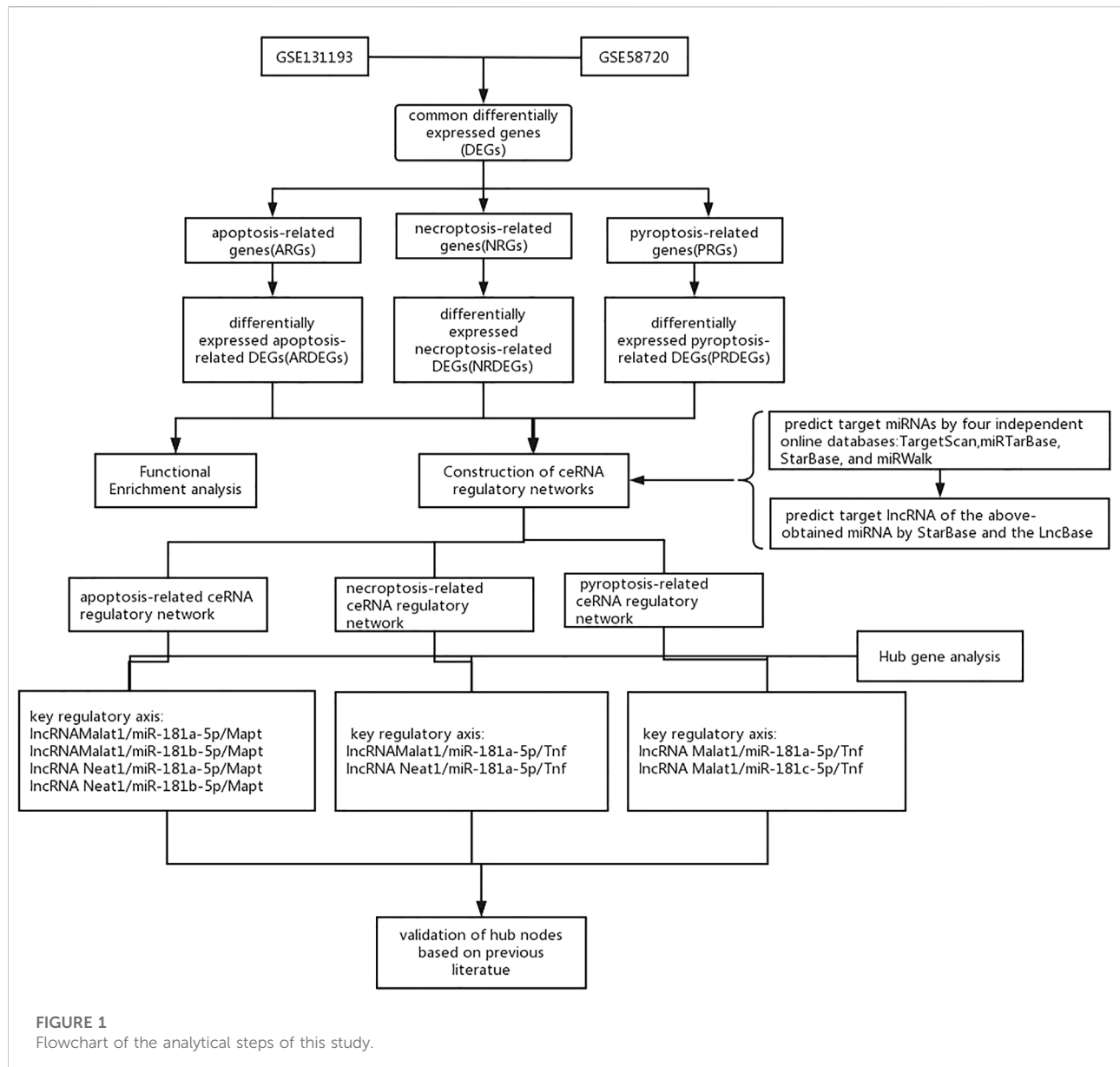
(<https://www.ncbi.nlm.nih.gov/geo/geo2r/>) using the Limma package. For the sequencing dataset GSE131193, the processed count matrix data was downloaded and differential analysis between tMCAO-operated mice and their matched control was conducted using the “lmFit” and “eBayes” functions in the Limma package (Ritchie et al., 2015). A p -value < 0.05 and $|\log_2$ fold change (FC)| > 1 were regarded as cut-off criteria for significant DEGs. The common DEGs in common between the GSE58720 dataset and the GSE131193 dataset were intersected with PCD (apoptosis, necroptosis, and pyroptosis)-related genes, respectively, to obtain apoptosis-related DEGs (ARDEGs), necroptosis-related DEGs (NRDEGs), and pyroptosis-related DEGs (PRDEGs). The above results were visualized using the online tool Jvenn (<http://jvenn.toulouse.inra.fr/app/index.html>) (Bardou et al., 2014).

2.3 Functional enrichment analysis

To obtain a better understanding of the biological mechanisms of the differentially expressed PCD-related genes, functional enrichment analysis, including Gene Ontology (GO) and pathway enrichment analysis, were performed using Metascape (<http://metascape.org>) (Zhou et al., 2019). The Kyoto Encyclopedia of Genes and Genomes (KEGG) (Kanehisa et al., 2017), Reactome (Jassal et al., 2020), and WikiPathways (Martens et al., 2021) databases were used for pathway annotations.

2.4 Construction of PCD-related ceRNA networks

To explore the potential regulatory mechanisms of these differentially expressed PCD-related genes, we constructed intricate competing endogenous RNA (ceRNA) networks. First, target microRNAs (miRNAs) of the obtained differentially expressed PCD-related genes were predicted by four independent online databases: TargetScan (Agarwal et al., 2015) (v7.2, http://www.targetscan.org/vert_72/), miRTarBase (Huang et al., 2020) (v8.0, <http://mirtarbase.mbc.nctu.edu.tw/php/index.php>), StarBase (Li et al., 2014) (<http://starbase.sysu.edu.cn/>), and miRWalk (Sticht et al., 2018) (<http://mirwalk.umm.uni-heidelberg.de/>). Only the miRNAs that were shared by any three or all four databases were regarded as eligible. Next, target long noncoding RNAs (lncRNAs) of the above-obtained miRNAs were predicted by StarBase and the LncBase module of the DIANA tool (<http://carolina.imis.athena-innovation.gr/>) (Karagkouni et al., 2020). Only the lncRNAs that were shared between the two databases were regarded as eligible. Finally, we selected lncRNA–mRNA interactions and miRNA–mRNA interactions that shared the same miRNAs to



construct the ceRNA network and visualized it using Cytoscape software (Shannon et al., 2003) (Version 3.8.0, <http://cytoscape.org>).

2.5 Hub gene analysis

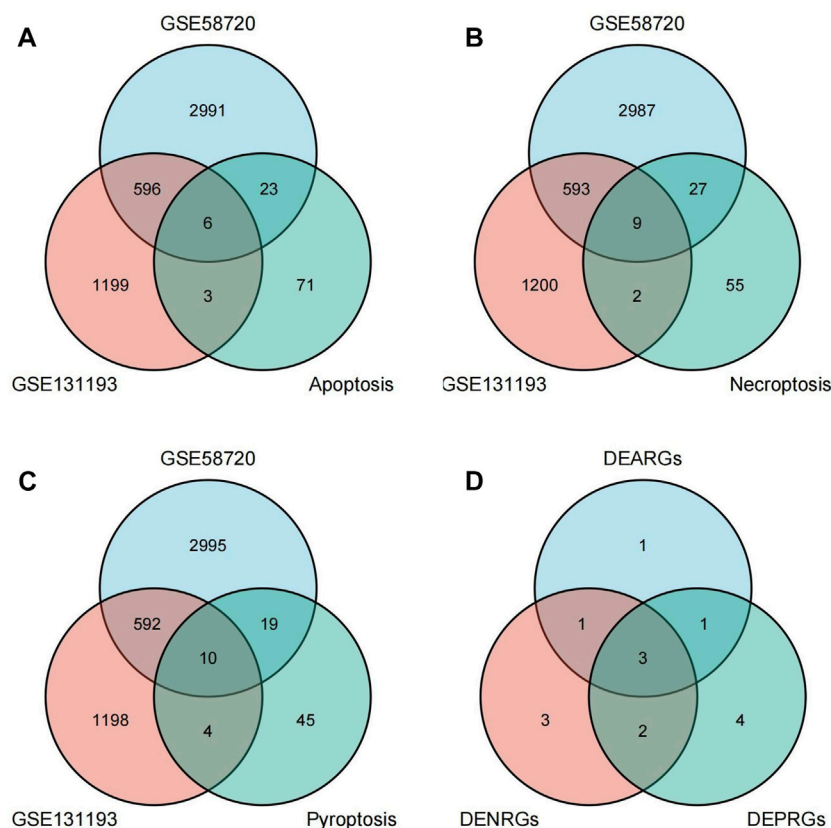
The cytoHubba plugin was applied to screen out the top ten genes of the above three ceRNA regulatory networks through seven different algorithms: MCC, Degree, Edge Percolated Component (EPC), EcCentricity, Closeness, Radiality, and Betweenness (Chin et al., 2014). UpSet R package was used to extract the overlapping genes obtained by the above seven

different algorithms and visualize them (Conway et al., 2017). These overlapping genes were confirmed as the hub nodes.

3 Results

3.1 Identification of differentially expressed PCD-related genes

A flow chart of this study is shown in Figure 1. We first analyzed the two datasets GSE58720 and GSE131193 to identify the common DEGs in MCAO/R-operated mice at 24 h of reperfusion compared with matched controls. Then, we

**FIGURE 2**

Identification of differentially expressed programmed cell death (PCD)-related genes. **(A)** Differentially expressed apoptosis-related genes (DEARGs) were identified by determining the overlap of datasets GSE58720 and GSE131193 with apoptosis-related genes. **(B)** Differentially expressed necroptosis-related genes (DENRGs) were identified by determining the overlap of datasets GSE58720 and GSE131193 with necroptosis-related genes. **(C)** Differentially expressed pyroptosis-related genes (DEPRGs) were identified by determining the overlap of datasets GSE58720 and GSE131193 with pyroptosis-related genes. **(D)** The overlapping genes were identified among DEARGs, DENRGs, and DEPRGs.

identified the common DEGs that overlapped with PCD (apoptosis, necroptosis, and pyroptosis)-related genes to obtain differentially expressed PCD-related genes (DEARGs, DENRGs, DEPRGs). A total of six DEARGs (Figure 2A), nine DENRGs (Figure 2B), and ten DEPRGs (Figure 2C) were identified. The six DEARGs included *Cd14*, *Zbp1*, *Tnfrsf10b*, *Bax*, *Mapt*, and *Pycard*, among which *Cd14*, *Zbp1*, *Tnfrsf10b*, *Bax*, and *Pycard* were upregulated in the dataset GSE58720 but downregulated in the dataset GSE131193, while *Mapt* showed the opposite pattern (Table 1). The nine DENRGs included *Cxcl1*, *Zbp1*, *Il1b*, *Tnf*, *Ripk3*, *Tnfrsf10b*, *Mkl1*, *Pycard*, and *Bax* (Table 1) and the ten DEPRGs included *Cd14*, *Zbp1*, *Il1b*, *Mefv*, *Tnf*, *Il1rn*, *Anxa2*, *Ccr5*, *Pycard*, and *Bax* (Table 1). All of these DENRGs and DEPRGs were upregulated in the dataset GSE58720 but downregulated in the dataset GSE131193. We also attempted to identify the common genes among the above three kinds of differentially expressed PCD-related genes and found that three genes, namely, *Zbp1*, *Bax*, and *Pycard*, overlapped among the three sets of differentially expressed PCD-related genes. *Cd14*

was in common between DEARGs and DEPRGs, *Tnfrsf10b* was in common between DEARGs and DENRGs, while *Il1b* and *Tnf* were in common between DENRGs and DEPRGs (Figure 2D and Table 1).

3.2 Functional enrichment analysis

To further explore the potential functions of DEARGs, DENRGs, and DEPRGs, functional enrichment analysis was performed using the online database Metascape. The results of GO analysis revealed that the DEARGs were particularly associated with the positive regulation of cell death, apoptotic signaling pathway, negative regulation of mitochondrial membrane potential, extrinsic apoptotic signaling pathway via death domain receptors, positive regulation of interleukin-8 production, membrane rafts, and left-handed Z-DNA binding (Figure 3A and Table 2). The DENRGs were mainly associated with programmed necrotic cell death, response to virus, positive

TABLE 1 Differentially expressed PCD-related genes.

| Gene symbol | Gene name | Expression in GSE58720 | Expression in GSE131193 | Belong to which kind of PCD related genes (DEARGs, DENRGs, DEPRGs) |
|-------------|--|------------------------|-------------------------|--|
| Bax | BCL2-associated X protein | upregulated | downregulated | DEARG, DENRG, DEPRG |
| Pycard | PYD and CARD domain containing | upregulated | downregulated | DEARG, DENRG, DEPRG |
| Zbp1 | Z-DNA binding protein 1 | upregulated | downregulated | DEARG, DENRG, DEPRG |
| Tnfrsf10b | tumor necrosis factor receptor superfamily, member 10b | upregulated | downregulated | DEARG, DENRG |
| Il1b | interleukin 1 beta | upregulated | downregulated | DENRG, DEPRG |
| Tnf | tumor necrosis factor | upregulated | downregulated | DENRG, DEPRG |
| Cd14 | CD14 antigen | upregulated | downregulated | DEARG, DEPRG |
| Mapt | microtubule-associated protein tau | downregulated | upregulated | DEARG |
| Cxcl1 | chemokine (C-X-C motif) ligand 1 | upregulated | downregulated | DENRG |
| Ripk3 | receptor-interacting serine-threonine kinase 3 | upregulated | downregulated | DENRG |
| Mlkl | mixed lineage kinase domain-like | upregulated | downregulated | DENRG |
| Mefv | Mediterranean fever | upregulated | downregulated | DEPRG |
| Il1rn | interleukin 1 receptor antagonist | upregulated | downregulated | DEPRG |
| Anxa2 | annexin A2 | upregulated | downregulated | DEPRG |
| Ccr5 | chemokine (C-C motif) receptor 5 | upregulated | downregulated | DEPRG |

Note: PCD, programmed cell death; DEARG, differentially expressed apoptosis-related gene. DENRG, differentially expressed necroptosis-related gene; DEPRG, differentially expressed pyroptosis-related gene.

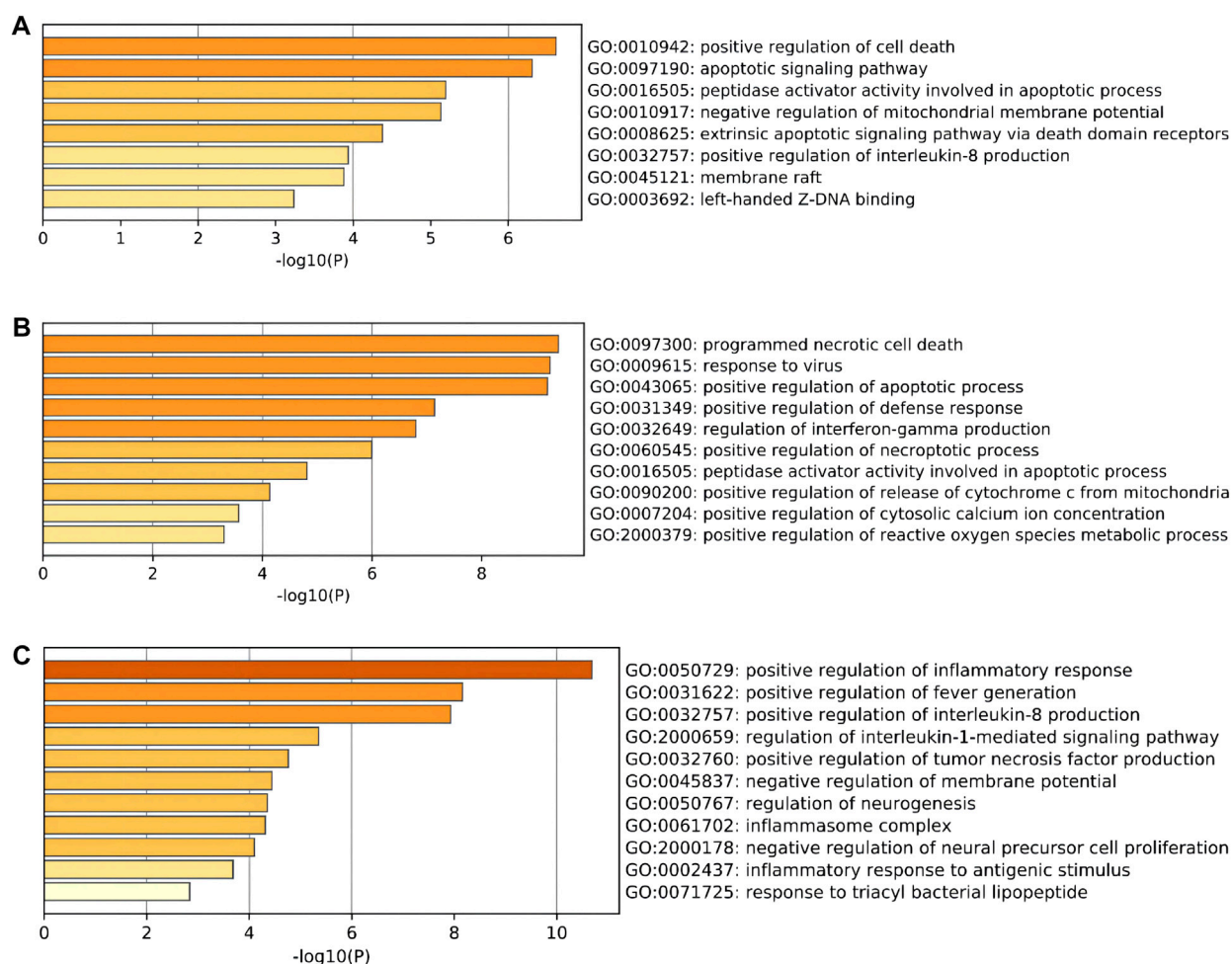
regulation of apoptotic process, necroptotic process, defense response, release of cytochrome c from mitochondria, cytosolic calcium ion concentration, reactive oxygen species metabolic process, and regulation of interferon-gamma production (Figure 3B and Table 3). The DEPRGs were particularly involved in the positive regulation of inflammatory response, tumor necrosis factor production, fever generation, interleukin-8 production, regulation of interleukin-1-mediated signaling pathway, regulation of neurogenesis, negative regulation of membrane potential and neural precursor cell proliferation, and inflammatory response to antigenic stimulus (Figure 3C and Table 4).

Moreover, regarding the results of pathway analysis, these revealed that the DEARGs were mainly associated with apoptosis, necroptosis, lipids and atherosclerosis, caspase activation via death receptors in the presence of ligand influenza A, legionellosis, activation, translocation and oligomerization of BAX, and the Mapk signaling pathway (Figure 4A and Table 2). The DENRGs were particularly involved in necroptosis, lipids and atherosclerosis, TNF signaling pathway, legionellosis, cytosolic DNA-sensing pathway, RIPK1-mediated regulated necrosis, TRAIL signaling, Kaposi sarcoma-associated herpesvirus infection, and the NLRP3 inflammasome (Figure 4B and Table 3). The DEPRGs were particularly associated with *Salmonella* infection, *Yersinia* infection, *tuberculosis*, CLEC7A/inflammasome pathway, viral protein interaction with cytokine and cytokine

receptor, transfer of LPS from LBP carrier to CD14, and neutrophil degranulation (Figure 4C and Table 4).

3.3 Construction of PCD-related ceRNA networks

To clarify the potential molecular regulatory mechanisms of these differentially expressed PCD-related genes, we then constructed PCD-related ceRNA regulatory networks of lncRNA-miRNA-mRNA. First, four independent online databases, namely, TargetScan, miRTarBase, StarBase, and miRWalk, were used to predict the interactions between miRNAs and mRNAs. Only the miRNAs that were shared by any three or all four of the databases were regarded as eligible. A total of 133 miRNA-mRNA interactions for apoptosis-related ceRNA regulatory networks, 91 miRNA-mRNA interactions for necroptosis-related ceRNA regulatory networks, and 70 miRNA-mRNA interactions for pyroptosis-related ceRNA regulatory networks were obtained based on the above methods. Next, target lncRNAs of the above-obtained miRNAs were predicted by StarBase and the LncBase module of the DIANA tool; only the lncRNAs that were shared between the two databases were regarded as eligible. A total of 107 lncRNA-miRNA pairs for apoptosis-related ceRNA regulatory networks, 58 lncRNA-miRNA pairs for necroptosis-related ceRNA regulatory networks, and 49 lncRNA-miRNA pairs for pyroptosis-related ceRNA

**FIGURE 3**

Gene Ontology (GO) enrichment analysis of these differentially expressed programmed cell death (PCD)-related genes. **(A)** Significantly enriched GO terms of differentially expressed apoptosis-related genes (DEARGs). **(B)** Significantly enriched GO terms of differentially expressed necroptosis-related genes (DENRGs). **(C)** Significantly enriched GO terms of differentially expressed pyroptosis-related genes (DEPRGs).

regulatory networks were identified. Then, the ceRNA networks were constructed using the miRNAs shared between the lncRNAs and mRNAs according to the ceRNA hypothesis. The apoptosis-related ceRNA regulatory network included 152 edges and 69 nodes (including 24 lncRNAs, 41 miRNAs, and 4 mRNAs) (Figure 5A). The necroptosis-related ceRNA regulatory network included 82 edges and 42 nodes (including 16 lncRNAs, 20 miRNAs, and 6 mRNAs) (Figure 5B). Finally, the pyroptosis-related ceRNA regulatory network included 69 edges and 39 nodes (including 15 lncRNAs, 18 miRNAs, and 6 mRNAs) (Figure 5C).

3.4 Hub gene analysis

The cytoHubba plugin was used to identify hub nodes of each of these PCD-related ceRNA regulatory networks based on the

above methods. In the hub gene analysis, five hub nodes, namely, the mRNA Mapt, miR-181a-5p, miR-181b-5p, and the lncRNAs Malat1 and Neat1, were identified in the apoptosis-related ceRNA regulatory network (Figure 6A) and these hub nodes formed four ceRNA regulatory pathways, namely, lncRNA Malat1/miR-181a-5p/Mapt, lncRNA Malat1/miR-181b-5p/Mapt, lncRNA Neat1/miR-181a-5p/Mapt, and lncRNA Neat1/miR-181b-5p/Mapt (Figure 6B). Hub nodes in the necroptosis-related ceRNA regulatory network included mRNA Tnf, miR-181a-5p, lncRNA Malat1, lncRNA Xist, and lncRNA Neat1 (Figure 6C). These hub nodes formed two ceRNA regulatory pathways, namely, lncRNA Malat1/miR-181a-5p/Tnf and lncRNA Neat1/miR-181a-5p/Tnf (Figure 6D). Hub nodes in the pyroptosis-related ceRNA regulatory network included the mRNA Tnf, miR-181a-5p, miR-181c-5p, lncRNA Malat1, and lncRNA Xist (Figure 6E); these hub nodes formed two ceRNA

TABLE 2 Functional enrichment analysis of differentially expressed apoptosis-related genes (DEARGs).

Significant enriched GO terms of DEARGs

| Category | Term | Description | p value | Gene symbols |
|---|--------------|--|---------|---------------------------------|
| GO Biological Processes | GO:0010942 | positive regulation of cell death | 0.0000 | Bax,Mapt, Tnfrsf10b,Zbp1,Pycard |
| GO Biological Processes | GO:0097190 | apoptotic signaling pathway | 0.0000 | Bax,Mapt, Tnfrsf10b,Pycard |
| GO Biological Processes | GO:0010917 | negative regulation of mitochondrial membrane potential | 0.0000 | Bax,Mapt |
| GO Biological Processes | GO:0008625 | extrinsic apoptotic signaling pathway via death domain receptors | 0.0000 | Bax,Tnfrsf10b |
| GO Biological Processes | GO:0032757 | positive regulation of interleukin-8 production | 0.0001 | Cd14,Pycard,Bax |
| GO Cellular Components | GO:0045121 | membrane raft | 0.0001 | Cd14,Mapt, Tnfrsf10b |
| GO Molecular Functions | GO:0003692 | left-handed Z-DNA binding | 0.0006 | Zbp1,Bax |
| Significant enriched pathways of DEARGs | | | | |
| Reactome Gene Sets | R-MMU-109581 | Apoptosis | 0.0000 | Bax,Cd14,Mapt, Tnfrsf10b |
| KEGG Pathway | mmu04217 | Necroptosis | 0.0000 | Bax,Tnfrsf10b,Zbp1,Pycard |
| KEGG Pathway | mmu05417 | Lipid and atherosclerosis | 0.0000 | Bax,Cd14,Tnfrsf10b,Pycard |
| Reactome Gene Sets | R-MMU-140534 | Caspase activation via Death Receptors in the presence of ligand | 0.0000 | Cd14,Tnfrsf10b |
| KEGG Pathway | mmu05164 | Influenza A | 0.0000 | Bax,Tnfrsf10b,Pycard |
| KEGG Pathway | mmu05134 | Legionellosis | 0.0001 | Cd14,Pycard,Zbp1 |
| Reactome Gene Sets | R-MMU-114294 | Activation, translocation and oligomerization of BAX | 0.0006 | Bax,Cd14,Mapt |
| WikiPathways | WP493 | Mapk signaling pathway | 0.0009 | Cd14,Mapt |

regulatory pathways, namely, lncRNA Malat1/miR-181a-5p/Tnf and lncRNA Malat1/miR-181c-5p/Tnf (Figure 6F).

3.5 Validation of hub nodes in the programmed cell death-related ceRNA regulatory networks

To validate the hub nodes in these PCD-related ceRNA regulatory networks, we searched the literature and found that they were abnormally expressed in ischemic stroke. Shi et al. (Shi et al., 2021) found that the levels of acetylated tau (ac-MAPT) and phosphorylated tau (p-MAPT) increased in rats subjected to MCAO/R compared with that in the sham group. The protein and mRNA levels of total-tau (T-MAPT) showed no significant differences between the sham and MCAO/R groups. Basurto-Islas et al. (Basurto-Islas et al., 2018) observed higher phosphorylation of tau and total tau in MCAO/R mice. Other studies also reported that the hyperphosphorylation of tau increases during MCAO/R in animal models (Dong et al., 2014; Fujii et al., 2017). Tnf was also reported to be significantly upregulated in MCAO/R animal models and OGD/R cell models (Li et al., 2019; Zhang et al., 2021a; Zhou et al., 2021). Moreover, it was reported that miR-181a-5p was highly expressed in serum of ischemic stroke patients, brain tissues of MCAO/R mice, and an oxygen-glucose-deprivation/reoxygenation (OGD/R) N2a cell model (Ouyang et al., 2012; Wu et al., 2017; Song et al., 2021). Studies also reported that miR-181b-5p and miR-181c-5p expression was significantly decreased

in cerebral ischemia *in vivo* and *in vitro* (Deng et al., 2016; Ma et al., 2016; Zhang et al., 2018; Meng et al., 2020). Accumulating evidence has also revealed that expression of the lncRNA Malat1 was upregulated after MCAO/R in rats and mice and OGD/R in different cells including primary neuronal cells, HT-22 cells, mouse astrocyte cells, and brain vascular endothelial cells (Xin and Jiang, 2017; Zhang et al., 2021b; Jia et al., 2021; Tan et al., 2021). Moreover, a recent study reported that the lncRNA Malat1 significantly increased in the blood of ischemic stroke patients compared with the level in normal controls (Tan et al., 2021). Furthermore, several studies reported that the lncRNA Neat1 was increased in an MCAO/R animal model, an OGD/R-induced cell model, and ischemic stroke patients (Ni et al., 2020; Zhang et al., 2021c; Jin et al., 2021). Another recent study reported that expression of the lncRNA Neat1 was significantly decreased in OGD/R-induced BV-2 and N2a cells compared with that in control cells (Zhou et al., 2022). The lncRNA XIST was also reported to be highly expressed in an MCAO/R-treated animal model and an OGD/R-treated cell model (Zhang et al., 2021d; Wang et al., 2021; Xiong et al., 2021). Wang et al. also reported that the lncRNA XIST was upregulated in brain tissues under MCAO/R treatment and in OGD/R-treated PC12 cells (Wang et al., 2021). Furthermore, Xiong et al. found that XIST was significantly highly expressed in SH-SY5Y cells after OGD/R treatment (Xiong et al., 2021). Finally, another study identified that XIST expression was upregulated in the brain tissues of an I/R mouse model and OGD/R-induced N2a cells (Zhang et al., 2021d). The findings of these previous studies are in accordance with our results, indicating the robustness of our analysis.

TABLE 3 Functional enrichment analysis of differentially expressed necroptosis-related genes (DENRGs).

Significant enriched GO terms of DENRGs

| Category | Term | Description | p value | Symbols |
|---|---------------|--|---------|--|
| GO Biological Processes | GO:0097300 | programmed necrotic cell death | 0.0000 | Bax,Tnf,Ripk3,Mkl,Il1b,Tnfrsf10b,Pycard, Cxcl1 |
| GO Biological Processes | GO:0009615 | response to virus | 0.0000 | Bax,Tnf,Ripk3,Zbp1,Pycard, Mkl |
| GO Biological Processes | GO:0043065 | positive regulation of apoptotic process | 0.0000 | Bax,Il1b,Tnf,Tnfrsf10b,Ripk3,Zbp1,Pycard |
| GO Biological Processes | GO:0031349 | positive regulation of defense response | 0.0000 | Cxcl1,Il1b,Tnf,Zbp1,Pycard, Ripk3,Bax |
| GO Biological Processes | GO:0032649 | regulation of interferon-gamma production | 0.0000 | Il1b,Tnf,Ripk3,Pycard,Zbp1,Bax,Cxcl1,Mkl |
| GO Biological Processes | GO:0060545 | positive regulation of necroptotic process | 0.0000 | Ripk3,Zbp1,Bax |
| GO Biological Processes | GO:0090200 | positive regulation of release of cytochrome c from mitochondria | 0.0001 | Bax,Pycard, Ripk3 |
| GO Biological Processes | GO:0007204 | positive regulation of cytosolic calcium ion concentration | 0.0003 | Bax,Cxcl1,Il1b |
| GO Biological Processes | GO:2000379 | positive regulation of reactive oxygen species metabolic process | 0.0005 | Cxcl1,Ripk3 |
| Significant enriched pathways of DENRGs | | | | |
| KEGG Pathway | mmu04217 | Necroptosis | 0.0000 | Bax,Il1b,Tnf,Tnfrsf10b,Ripk3,Zbp1,Pycard, Mkl |
| KEGG Pathway | mmu05417 | Lipid and atherosclerosis | 0.0000 | Bax,Cxcl1,Il1b,Tnf,Tnfrsf10b,Pycard |
| KEGG Pathway | mmu04668 | TNF signaling pathway | 0.0000 | Cxcl1,Il1b,Tnf,Ripk3,Mkl |
| KEGG Pathway | mmu05134 | Legionellosis | 0.0000 | Cxcl1,Il1b,Tnf,Pycard, Ripk3,Tnfrsf10b |
| KEGG Pathway | mmu04623 | Cytosolic DNA-sensing pathway | 0.0000 | Il1b,Ripk3,Zbp1,Pycard |
| Reactome Gene Sets | R-MMU-5213460 | RIPK1-mediated regulated necrosis | 0.0000 | Tnfrsf10b,Ripk3,Mkl,Tnf |
| Reactome Gene Sets | R-MMU-75158 | TRAIL signaling | 0.0022 | Tnfrsf10b |
| KEGG Pathway | mmu05167 | Kaposi sarcoma-associated herpesvirus infection | 0.0040 | Bax,Cxcl1 |
| Reactome Gene Sets | R-MMU-844456 | The NLRP3 inflammasome | 0.0044 | Pycard, Cxcl1 |

4 Discussion

Accumulating evidence supports the involvement of PCD pathways in the pathogenesis of ischemic stroke and highlights the importance of each form of cell death. However, the specific mechanisms underlying them remain incompletely clarified. There is also currently a lack of specific neuroprotective drugs in clinical practice. Nonetheless, increasing studies have indicated significant crosstalk among these PCD pathways. Therefore, we applied bioinformatic analysis to identify differentially expressed PCD-related genes during cerebral I/R injury and investigated their potential regulatory axes by constructing ceRNA networks. This may contribute to elucidating the molecular mechanisms behind these PCD pathways and provide a basis for developing novel therapeutic targets against cerebral I/R injury.

A total of six DEARGs, nine DENRGs, and ten DEPRGs were identified in this study. Among them, three genes, namely, Bax, Zbp1, and Pycard, overlapped among these three sets of genes,

indicating that they may play key roles in the crosstalk among these PCD pathways. The protein encoded by the *Bax* gene belongs to the BCL2 protein family and is regarded as the fundamental effector of the intrinsic apoptotic pathway (Spitz and Gavathiotis, 2022). Numerous studies have indicated that Bax-dependent initiation and activation of subsequent apoptotic pathways play critical roles in ischemic brain injury (Li et al., 2021; Tu and Hu, 2021). In addition, it has been suggested that inhibition of Bax function may provide a new strategy for neuroprotection and functional improvement against cerebral ischemia (Han et al., 2011). Recently, some studies also reported that Bax is a key regulator of caspase-independent necroptosis and pyroptosis (Cabon et al., 2012; Hu et al., 2020). However, whether Bax is involved in necroptosis and pyroptosis in cerebral ischemia–reperfusion injury has remained unclear, so further research on this issue is needed. The gene *Pycard* encodes the adaptor protein ASC, which comprises two protein–protein interaction domains: an N-terminal PYRIN-PAAD-DAPIN domain (PYD) and a C-terminal caspase-recruitment domain

TABLE 4 Functional enrichment analysis of differentially expressed pyroptosis-related genes (DEPRGs).

Significant enriched GO terms of DEPRGs

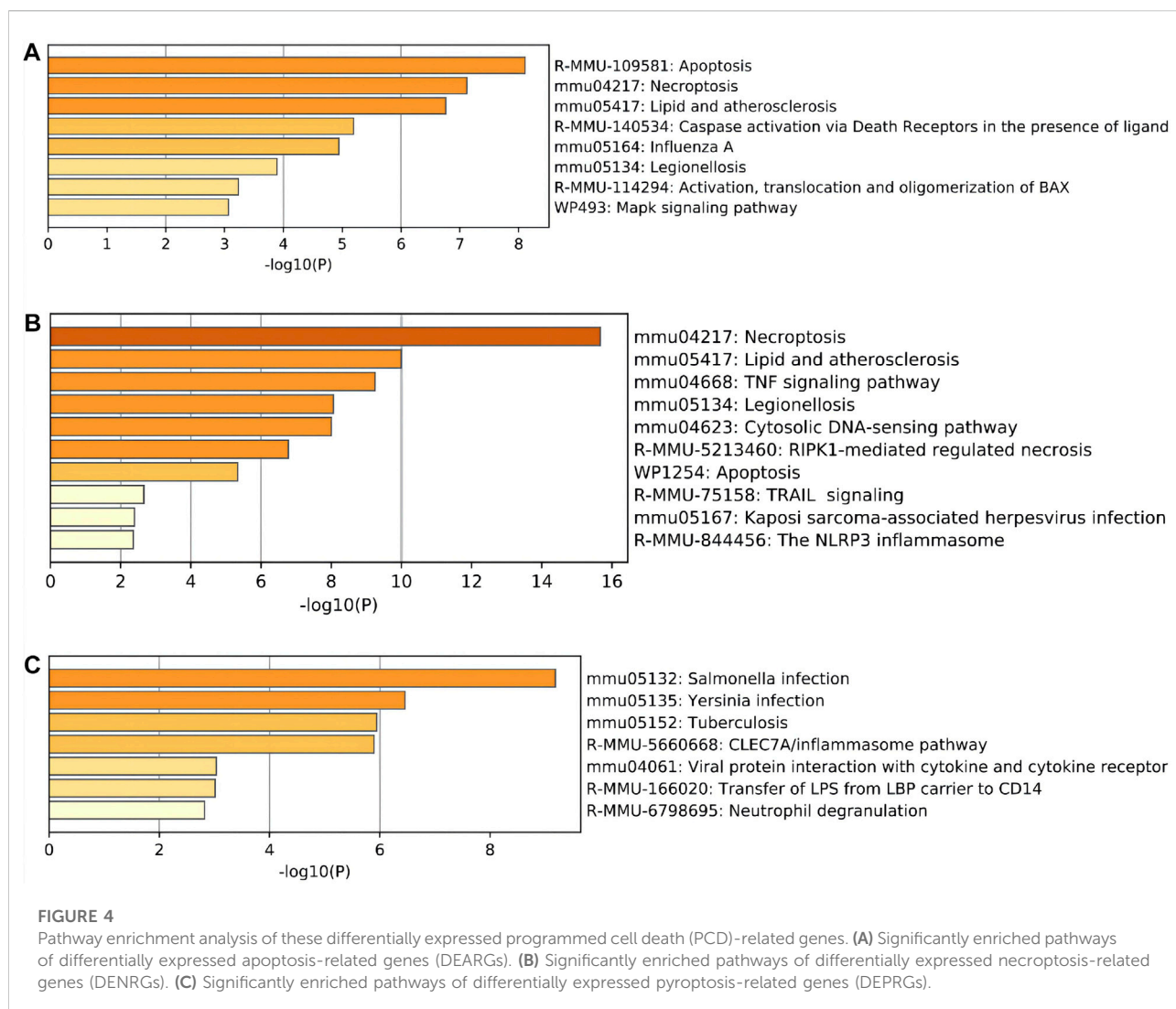
| Category | Term | Description | p value | Gene symbols |
|---|---------------|---|---------|--|
| GO Biological Processes | GO:0050729 | positive regulation of inflammatory response | 0.0000 | Ccr5,Ill1b,Tnf,Mefv,Zbp1,Pycard,Cd14,Ill1rn |
| GO Biological Processes | GO:0031622 | positive regulation of fever generation | 0.0000 | Ccr5,Ill1b,Tnf,Mefv, Pycard,Ill1rn,Bax,Anxa2,Zbp1 |
| GO Biological Processes | GO:0032757 | positive regulation of interleukin-8 production | 0.0000 | Cd14,Ill1b,Tnf,Pycard,Ccr5,Ill1rn,Anxa2,Bax |
| GO Biological Processes | GO:2000659 | regulation of interleukin-1-mediated signaling pathway | 0.0000 | Ill1rn,Zbp1,Bax,Tnf,Pycard,Ill1b |
| GO Biological Processes | GO:0032760 | positive regulation of tumor necrosis factor production | 0.0000 | Cd14,Ccr5,Pycard,Bax,Ill1b,Anxa2 |
| GO Biological Processes | GO:0045837 | negative regulation of membrane potential | 0.0000 | Bax,Ill1rn,Anxa2 |
| GO Biological Processes | GO:0050767 | regulation of neurogenesis | 0.0000 | Anxa2,Ccr5,Ill1b,Tnf,Cd14,Mefv |
| GO Cellular Components | GO:0061702 | inflammasome complex | 0.0000 | Mefv, Pycard |
| GO Biological Processes | GO:2000178 | negative regulation of neural precursor cell proliferation | 0.0001 | Ccr5,Ill1b,Bax,Tnf,Mefv |
| GO Biological Processes | GO:0002437 | inflammatory response to antigenic stimulus | 0.0002 | Ill1rn,Tnf,Ccr5,Ill1b,Anxa2 |
| Significant enriched pathways of DEPRGs | | | | |
| KEGG Pathway | mmu05132 | <i>Salmonella</i> infection | 0.0000 | Bax,Anxa2,Cd14,Ill1b,Tnf,Pycard,Zbp1,Ccr5 |
| KEGG Pathway | mmu05135 | <i>Yersinia</i> infection | 0.0000 | Ill1b,Tnf,Mefv, Pycard |
| KEGG Pathway | mmu05152 | Tuberculosis | 0.0000 | Bax,Cd14,Ill1b,Tnf,Ccr5 |
| Reactome Gene Sets | R-MMU-5660668 | CLEC7A/inflammasome pathway | 0.0000 | Ill1b,Pycard,Zbp1,Ccr5,Ill1rn,Tnf |
| KEGG Pathway | mmu04061 | Viral protein interaction with cytokine and cytokine receptor | 0.0009 | Ccr5,Tnf |
| Reactome Gene Sets | R-MMU-166020 | Transfer of LPS from LBP carrier to CD14 | 0.0010 | Cd14 |
| Reactome Gene Sets | R-MMU-6798695 | Neutrophil degranulation | 0.0015 | Anxa2,Cd14,Pycard |

(CARD) (Hoss et al., 2017). Previous studies demonstrated that ASC was upregulated in an ischemic stroke model and played a key role in cerebral ischemia–reperfusion injury by participating in the inflammatory response and cell death, including apoptosis, necroptosis, and pyroptosis (Meng et al., 2019; Liang et al., 2020; Xu et al., 2021a). The gene *Zbp1* encodes Z-DNA binding protein 1 with two Za domains, which is a critical innate immune sensor of not only viral RNA products but also endogenous nucleic acid ligands (Zhang et al., 2020). Previous studies showed that *Zbp1* plays a role in the innate immune response by binding to foreign DNA and inducing type I interferon production (Jiao et al., 2020). In addition, in response to influenza virus infection, it could induce cell death in the form of pyroptosis, apoptosis, and necroptosis, that is, PANoptosis (Zheng and Kanneganti, 2020). However, the role of *Zbp1* in cerebral I/R injury remains unknown. Recent studies demonstrated that *Zbp1* and ASC are components of the PANoptosome (Zheng and Kanneganti, 2020; Lee et al., 2021). Yan et al. (Yan et al., 2022) also indicated that PANoptosis is observed in ischemic brain injury. Based on previous related studies, we speculated that these three genes might be components of the PANoptosome and be involved in PANoptosis in cerebral I/R injury. Our study further confirmed previous findings supporting the hypothesis that these three

genes might be involved in PANoptosis and crosstalk among apoptosis, necroptosis, and pyroptosis in cerebral I/R injury, thus providing new targets for neuroprotection.

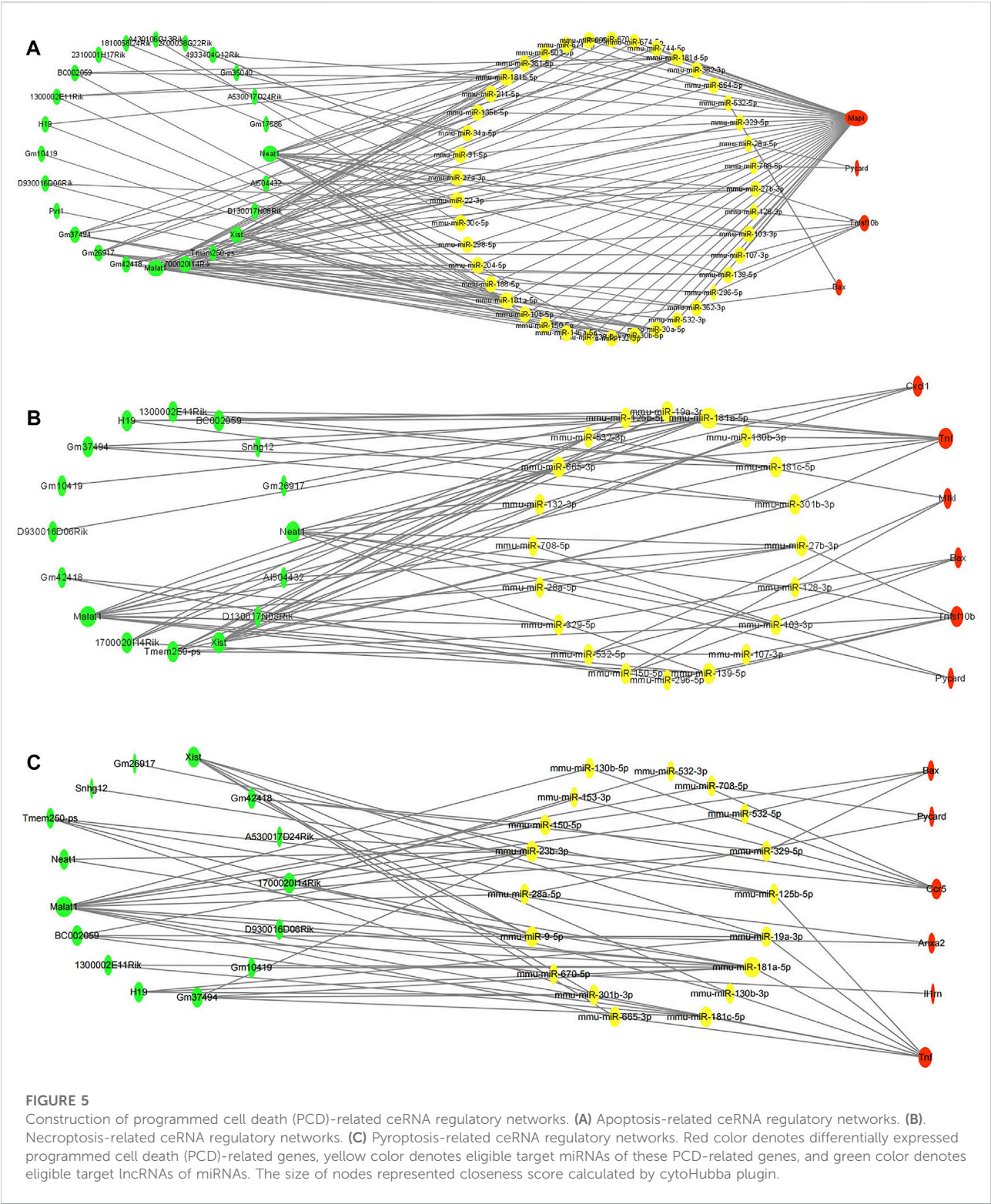
Functional enrichment analyses were performed to obtain a more in-depth understanding of the differentially expressed PCD-related genes. The results showed that these genes were not only particularly associated with corresponding PCD-related biological processes and pathways, but also involved in other biological processes and pathways, such as inflammatory response and reactive oxygen species metabolic process. This indicates that these genes have different functions under particular circumstances and that there might be crosstalk among these biological processes. These results are in line with previous studies that revealed significant crosstalk between PCD and inflammatory response (Jayaraj et al., 2019).

In recent years, increasing studies have suggested that the regulatory network composed of lncRNAs, miRNAs, and mRNAs plays a critical role in the mechanisms underlying cerebral ischemia–reperfusion injury (Xu et al., 2021b). To better understand the molecular regulatory mechanisms of these differentially expressed PCD-related genes, we constructed PCD-related ceRNA regulatory networks and



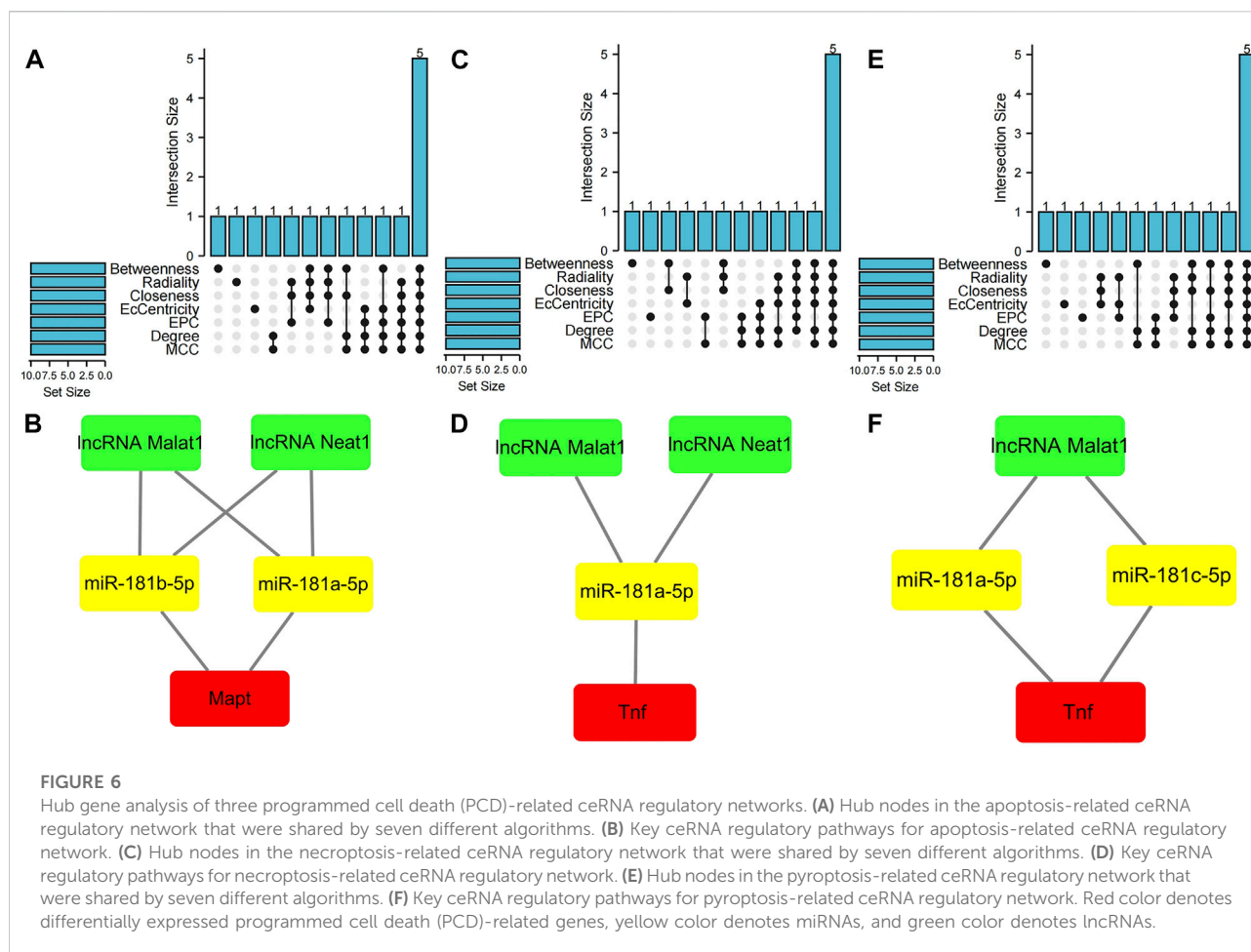
performed hub gene analysis to identify key nodes in these networks. The gene *Mapt* was found to be a hub node of the apoptosis-related ceRNA regulatory network. It can encode several isoforms of tau protein as a result of complex, regulated alternative splicing of its messenger RNA (Zhang et al., 2009). *Mapt* transcripts are differentially expressed in the nervous system, depending on the stage of neuronal maturation and neuron type. *Mapt* gene mutations have been shown to be associated with several neurodegenerative disorders (Michalicova et al., 2020). In recent years, increasing evidence has demonstrated that *Mapt* plays a role in ischemic stroke. In addition, Basurto-Islas et al. (Basurto-Islas et al., 2018) found that a large amount of hyperphosphorylated MAPT (Ser262/356) was colocalized with apoptotic cells in MCAO/R-treated mice. Moreover, Fujii et al. (Fujii et al., 2017) illustrated that the knockout of MAPT reduced infarct area and alleviated symptoms of neurological deficit. A recent study also showed

that astragaloside IV exerted neuroprotective effects in rats with cerebral ischemia/reperfusion (CIR) injury, probably through the Sirt1/Mapt pathway (Shi et al., 2021). In this study, the gene *Tnf* was identified as a hub node in both the necroptosis-related ceRNA regulatory network and the pyroptosis-related ceRNA regulatory network. This gene encodes the multifunctional proinflammatory cytokine TNF- α , which belongs to the tumor necrosis factor (TNF) superfamily (Watters and O'Connor, 2011). It can bind to its surface receptors and functions through their activation. Generally, TNF is a classical activator of necroptosis that binds to its receptor to recruit RIPK1, which interacts with RIPK3 to form necrosome and phosphorylate MLKL to mediate necroptosis in the absence of caspases-8 (Green, 2019). Studies have shown that the level of TNF- α was elevated in ischemic stroke and it has been implicated in cerebral I/R injury, exerting effects by regulating the inflammatory response and PCD pathways including



apoptosis, necroptosis, and pyroptosis (Hallenbeck, 2002; Maddahi et al., 2011), which is consistent with our results. Many studies have demonstrated that the inhibition of TNF

signaling pathways may have neuroprotective effects against cerebral I/R injury. For example, Zhang et al. showed that preconditioning with Carbonisatus significantly decreased the



levels of TNF- α and IL-6, reduced ischemic lesion volume, and improved neurological deficits in MCAO/R rats (Zhang et al., 2021a).

MicroRNAs (miRNAs) are conserved small regulatory noncoding RNAs of about 20–22 bp in length. They can regulate protein expression by binding to the 3' untranslated region (3'UTR) of their target genes, degrading or inhibiting their expression (Krol et al., 2010). Numerous studies have shown that miRNAs are involved in the regulation of PCD pathways in many diseases, including ischemic stroke (Ghafouri-Fard et al., 2020). The hub gene analysis in this study demonstrated that miR-181a-5p was a hub node in all of the above-mentioned three PCD-related ceRNA networks, and that miR-181b-5p was a hub node of the apoptosis-related ceRNA regulatory network and miR-181c-5p was a hub node of the pyroptosis-related ceRNA regulatory network. miR-181a-5p, miR-181b-5p, and miR-181c-5p all belong to the miR-181 family and their aberrant expression has been associated with various diseases including stroke, neurodegeneration, and cancer (Indrieri et al., 2020). Previous studies suggested that the miR-181 family

participates in the regulation of a range of biological processes including cell proliferation (Huo et al., 2016), apoptosis (Zhang et al., 2018), autophagy (Guo et al., 2019), and immune and inflammatory responses (Hutchison et al., 2013; Lu et al., 2019). Moreover, several studies have demonstrated that the inhibition of miR-181a-5p played a neuroprotective role in cerebral ischemic injury, as evidenced by reductions in cell apoptosis, pyroptosis, and cerebral infarction area (Moon et al., 2013; Stary et al., 2017; Yan et al., 2020; Song et al., 2021). However, the roles of miR-181b-5p and miR-181c-5p in cerebral ischemia have remained controversial. Peng et al. showed that downregulated miR-181b played a neuroprotective role against ischemic injury through negatively regulating HSPA5 and UCHL1 protein levels (Peng et al., 2013). In addition, Zhang et al. suggested that the downregulation of miRNA-181b protects against cerebral ischemic injury via the inhibition of NF- κ B-mediated inflammatory and apoptotic responses (Zhang et al., 2018). In contrast, another two reports demonstrated the possible neuroprotective effects of increased miR-181b in ischemia-caused neuronal cell apoptosis and mechanical repair of

brain tissue (Deng et al., 2016; Liu et al., 2016). Most studies supported the assertion that miR-181c-5p plays a positive role in brain injury caused by cerebral ischemia–reperfusion and that its overexpression can inhibit brain injury caused by ischemic stroke through regulating proliferation, inflammatory response, and apoptosis of neuronal cells (Zhang et al., 2019; Cao et al., 2020; Bu et al., 2021). However, in two other studies, the opposite conclusions were drawn. Specifically, Ma et al. (Ma et al., 2016) found a positive correlation between the NIHSS score and miR-181c level, and showed that plasma miR-181c concentration was positively correlated with the number of neutrophils and blood platelet count and negatively correlated with the number of lymphocytes. They also found that miR-181c promoted the apoptosis of BV2 and Neuro-2a cells and aggravated brain ischemia–reperfusion injury in a mouse model of stroke via the modulation of pro- and anti-apoptotic proteins. Moreover, a recent study showed that downregulated miR-181c ameliorated cerebral ischemic injury via increasing the expression of c-Fos and its downstream genes (Meng et al., 2020). Taken together, these findings indicated that miR-181a-5p, miR-181b-5p, and miR-181c-5p are all involved in the mechanism of cerebral I/R injury, but might play different roles depending on the specific target gene to which they bind.

lncRNAs are the most abundant noncoding RNAs (ncRNAs). They are greater than 200 bp in length, lack protein-coding function, and are associated with a variety of neurological diseases including ischemic stroke (Wu et al., 2013; Bao et al., 2018). Our hub analysis identified the lncRNA Malat1 as a hub node in all three PCD-related ceRNA regulatory networks, which is consistent with previous studies, indicating its critical role in regulating PCD pathways. Malat1 is known as a long intergenic noncoding RNA and is highly abundant in the nervous system. Accumulating evidence has indicated that this lncRNA plays vital roles in regulating various physiological processes, including apoptosis, autophagy, immune and inflammatory responses, and endothelial dysfunction of ischemic stroke (Wang et al., 2022). The expression of Malat1 was also found to be upregulated in ischemic stroke, while its downregulation was shown to improve the neurological deficit score and reduce neuronal apoptosis and the size of cerebral infarction by regulating miR-211-5p to in turn regulate the expression of COX-2 (Tan et al., 2021). Other studies also demonstrated that the inhibition of Malat1 expression could protect against cerebral I/R injury by alleviating neuronal apoptosis, endoplasmic reticulum stress, and inflammation (Shi et al., 2019; Cao et al., 2020; Jia et al., 2021). Moreover, it was reported that Malat1 was highly expressed in OGD/R-induced astrocyte injury models, and that its silencing protected against cerebral ischemia–reperfusion injury by downregulating AQP4 levels via miR-145 (Wang et al., 2020). In contrast, some studies

supported the neuroprotective role of Malat1 in cerebral ischemia–reperfusion injury. For example, Xin et al. (Xin and Jiang, 2017) found that Malat1 could protect human brain vascular endothelial cells from OGD/R-induced apoptosis via a PI3K-dependent mechanism. Another study showed that mice with lncRNA Malat1 KO presented larger brain infarct size and worse neurological scores, indicating that Malat1 plays critical protective roles in ischemic stroke via anti-apoptotic and anti-inflammatory effects in the brain microvasculature (Zhang et al., 2017). Accumulating evidence has also indicated that Malat1 is an important regulator of pyroptosis in various diseases (Song et al., 2019; Shu et al., 2021). However, the specific roles and mechanisms by which Malat1 regulates pyroptosis and necroptosis in cerebral ischemia–reperfusion injury have remained unclear. In addition, the lncRNAMalat1/miR-181a-5p/Mapt regulatory axis, lncRNAMalat1/miR-181b-5p/Mapt regulatory axis, and lncRNA Malat1/miR-181a-5p/Tnf regulatory axis were not previously reported to be involved in cerebral I/R injury, so they need further investigation. In this study, the lncRNA Neat1 was also identified as a hub node in both apoptosis-related and necroptosis-related ceRNA regulatory networks. Recently, increasing evidence has shown that this lncRNA plays an essential role in physiological and pathological responses in ischemic stroke (Ni et al., 2020; Jin et al., 2021). Li et al. (Li et al., 2020) found that the expression of Neat1 was elevated in patients with ischemic stroke compared with that in controls, and that lncRNA Neat1 expression positively correlated with NIHSS score and inflammatory factors and could predict an increased risk of recurrence/death. Ni et al. (Ni et al., 2020) also showed that Neat1 knockdown alleviated OGD/R-induced apoptosis and increased neuronal viability. Another study demonstrated that Gastrodin significantly alleviated cerebral I/R injury by regulating the lncRNA Neat1/miR-22-3p axis; it also showed that the overexpression of Neat1 promoted neuronal pyroptosis (Zhang et al., 2021c). Previous studies also reported that the downregulation of Neat1 could exert anti-inflammatory effects in cerebral I/R injury (Han and Zhou, 2019; Jin et al., 2021). Taken together, these findings indicate that the lncRNA Neat1 might play crucial roles in PCD pathways and inflammation in cerebral I/R injury and is a potential therapeutic target. Another lncRNA identified as a hub node in both necroptosis-related ceRNA and pyroptosis-related regulatory networks is Xist. Previous studies confirmed that Xist contributes to cerebral I/R injury through modulating cell apoptosis, ROS production, and inflammation. Wang et al. also demonstrated that the silencing of XIST protected against cerebral I/R injury by inhibiting neuronal deficit and inflammation via the miR-362/ROCK2 axis (Wang et al., 2021). In addition, Xiong et al. reported that XIST reduced cell viability and induced cell apoptosis via modulating miR-486-5p and GAB2, which promoted cerebral I/R injury (Xiong

et al., 2021). Another two studies also indicated that knockdown of XIST inhibited brain injury by suppressing apoptosis and ROS production (Zhang et al., 2021d; Weng et al., 2021). Moreover, a recent study illustrated that Xist was involved in the regulation of pyroptosis in MCAO/R-treated rats and OGD/R-treated rat brain microvascular endothelial cells (Guo et al., 2022). Nevertheless, the molecular roles and regulatory mechanisms of XIST in necroptosis and pyroptosis in cerebral I/R injury have not been fully elucidated and require further research.

In summary, we analyzed datasets GSE131193 and GSE58720 to identify PCD-related genes signature and potential regulatory axes in cerebral I/R injury and the results were validated through previous work. To our knowledge, this was the first study to focused on PCD (apoptosis, necroptosis, and pyroptosis)-related genes and potential regulatory axes in cerebral I/R injury, which might have profound significance for ischemia stroke. We identified hub nodes and seven key ceRNA regulatory axes that has never been reported before in ischemia stroke, which may contribute to elucidating the molecular mechanisms and provide a basis for developing novel therapeutic targets against cerebral I/R injury. Further *in vivo* and *in vitro* studies should be conducted to verify these regulatory axes.

There were several limitations to this study that should be acknowledged. First, the PCD (apoptosis, necroptosis, and pyroptosis)-related genes included in this study were mainly identified in previous studies, so some unreported related genes may have been ignored or excluded. Second, both lncRNAs and miRNAs were obtained by online database prediction because neither lncRNA nor miRNA datasets on adult mouse brain tissues after transient focal ischemia at 24 h of reperfusion and matched control samples were available. In future studies, if available, lncRNA and miRNA datasets should be analyzed simultaneously to increase the reliability of the results. Third, the selected datasets were performed in different laboratories and therefore, the differences in sample preparation, sample collection methods, and microarray platforms might influence the results. Finally, our hypothesized potential binding affinity among lncRNAs, miRNAs, and mRNAs should be subjected to further experimental investigation.

5 Conclusion

Taken together, our findings indicated that these PCD pathways (apoptosis, necroptosis, pyroptosis, and PANoptosis) and crosstalk among them might be involved in ischemic stroke. And the key nodes and regulatory axes identified in this study might play vital roles in regulating the above processes, which may offer new insights into the potential mechanisms underlying cell death during cerebral I/R injury and provide new therapeutic targets for neuroprotection.

Data availability statement

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

Author contributions

JS, LY, WW, and LZ conceived and designed the study, analyzed the data, wrote the manuscript, and contributed to the designs of the methods used in this study. All authors have read and agreed to publication of this version of the manuscript.

Funding

This work was funded by Shanghai Municipal Key Clinical Specialty, grant number shslczdzk02801, the Shanghai Municipal Health Commission, grant number 2020YJZX0109, the Shanghai Pujiang Program, grant number 2019PJD013, and the National Natural Science Foundation of China (NSFC), grant number 81471103.

Acknowledgments

We thank Liwen Bianji (Edanz) (www.liwenbianji.cn) for editing the language of a draft of this 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.2022.934154/full#supplementary-material>

References

- Agarwal, V., Bell, G. W., and Nam, J. W. (2015). *Predicting effective microRNA target sites in mammalian mRNAs* [J]. United Kingdom: Elife, 4.
- Banoth, B., Tuladhar, S., Karki, R., Sharma, B. R., Briard, B., Kesavardhana, S., et al. (2020). ZBP1 promotes fungi-induced inflammasome activation and pyroptosis, apoptosis, and necroptosis (PANoptosis). *J. Biol. Chem.* 295 (52), 18276–18283. doi:10.1074/jbc.RA120.015924
- Bao, M., Szeto, V., Yang, B. B., Zhu, S. Z., Sun, H. S., and Feng, Z. P. (2018). Long non-coding RNAs in ischemic stroke. *Cell Death Dis.* 9 (3), 281. doi:10.1038/s41419-018-0282-x
- Bardou, P., Mariette, J., Escudie, F., Djemiel, C., and Klopp, C. (2014). jvenn: an interactive Venn diagram viewer. *BMC Bioinforma.* 15, 293. doi:10.1186/1471-2105-15-293
- 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 (Database issue), D991–D995. doi:10.1093/nar/gks1193
- Basurto-Islas, G., Gu, J., Tung, Y. C., Liu, F., and Iqbal, K. (2018). Mechanism of tau hyperphosphorylation involving lysosomal enzyme asparagine endopeptidase in a mouse model of brain ischemia. *J. Alzheimer's Dis.* 63 (2), 821–833. doi:10.3233/jad-170715
- Bu, X., Zhao, Y., Chang, M., and Ge, X. (2021). Downregulation of lncRNA SNHG14 alleviates neurons injury by modulating the miR-181c-5p/BMF axis in ischemic stroke. *Brain Res. Bull.* 174, 379–388. doi:10.1016/j.brainresbull.2021.06.026
- Cabon, L., Galan-Malo, P., Bouharrou, A., Brunelle-Navas, M. N., Lorenzo, H. K., Gross, A., et al. (2012). BID regulates AIF-mediated caspase-independent necroptosis by promoting BAX activation. *Cell Death Differ.* 19 (2), 245–256. doi:10.1038/cdd.2011.91
- Campbell, B., and Khatri, P. (2020). Stroke. *Lancet* 396 (10244), 129–142. doi:10.1016/S0140-6736(20)31179-X
- Cao, D. W., Liu, M. M., Duan, R., Tao, Y. F., Zhou, J. S., Fang, W. R., et al. (2020). The lncRNA Malat1 functions as a ceRNA to contribute to berberine-mediated inhibition of HMGB1 by sponging miR-181c-5p in poststroke inflammation. *Acta Pharmacol. Sin.* 41 (1), 22–33. doi:10.1038/s41401-019-0284-y
- 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 (Suppl. 4), S11. doi:10.1186/1752-0509-8-S4-S11
- Conway, J. R., Lex, A., and Gehlenborg, N. (2017). UpSetR: an R package for the visualization of intersecting sets and their properties. *Bioinformatics* 33 (18), 2938–2940. doi:10.1093/bioinformatics/btx364
- Datta, A., Sarmah, D., Mounica, L., Kaur, H., Kesharwani, R., Verma, G., et al. (2020). Cell death pathways in ischemic stroke and targeted pharmacotherapy. *Transl. Stroke Res.* 11 (6), 1185–1202. doi:10.1007/s12975-020-00806-z
- Deng, B., Bai, F., Zhou, H., Zhou, D., Ma, Z., Xiong, L., et al. (2016). Electroacupuncture enhances rehabilitation through miR-181b targeting PirB after ischemic stroke. *Sci. Rep.* 6, 38997. doi:10.1038/srep38997
- Deng, X. X., Li, S. S., and Sun, F. Y. (2019). Necrostatin-1 prevents necroptosis in brains after ischemic stroke via inhibition of RIPK1-mediated RIPK3/MLKL signaling. *Aging Dis.* 10 (4), 807–817. doi:10.14336/AD.2018.0728
- Dong, D. W., Zhang, Y. S., Yang, W. Y., Wang-Qin, R. Q., Xu, A. D., and Ruan, Y. W. (2014). Hyperphosphorylation of tau protein in the ipsilateral thalamus after focal cortical infarction in rats. *Brain Res.* 1543, 280–289. doi:10.1016/j.brainres.2013.11.004
- Dong, Z., Pan, K., Pan, J., Peng, Q., and Wang, Y. (2018). The possibility and molecular mechanisms of cell pyroptosis after cerebral ischemia. *Neurosci. Bull.* 34 (6), 1131–1136. doi:10.1007/s12264-018-0294-7
- Eltzschig, H. K., and Eckle, T. (2011). Ischemia and reperfusion—from mechanism to translation. *Nat. Med.* 17 (11), 1391–1401. doi:10.1038/nm.2507
- Fujii, H., Takahashi, T., Mukai, T., Tanaka, S., Hosomi, N., Maruyama, H., et al. (2017). Modifications of tau protein after cerebral ischemia and reperfusion in rats are similar to those occurring in Alzheimer's disease - hyperphosphorylation and cleavage of 4- and 3-repeat tau. *J. Cereb. Blood Flow. Metab.* 37 (7), 2441–2457. doi:10.1177/0271678X16668889
- Ghafouri-Fard, S., Shoori, H., and Taheri, M. (2020). Non-coding RNAs participate in the ischemia-reperfusion injury. *Biomed. Pharmacother.* 129, 110419. doi:10.1016/j.biopha.2020.110419
- Green, D. R. (2019). The coming decade of cell death research: Five riddles. *Cell* 177 (5), 1094–1107. doi:10.1016/j.cell.2019.04.024
- Guo, J., Ma, Y., Peng, X., Jin, H., and Liu, J. (2019). LncRNA CCAT1 promotes autophagy via regulating ATG7 by sponging miR-181 in hepatocellular carcinoma. *J. Cell. Biochem.* 120 (10), 17975–17983. doi:10.1002/jcb.29064
- Guo, Y., Yang, J. H., He, Y., Zhou, H. F., Wang, Y., Ding, Z. S., et al. (2022). Protocatechuic aldehyde prevents ischemic injury by attenuating brain microvascular endothelial cell pyroptosis via lncRNA Xist. *Phytomedicine.* 94, 153849. doi:10.1016/j.phymed.2021.153849
- Hallenbeck, J. M. (2002). The many faces of tumor necrosis factor in stroke. *Nat. Med.* 8 (12), 1363–1368. doi:10.1038/nm1202-1363
- Han, B., Wang, Q., Cui, G., Shen, X., and Zhu, Z. (2011). Post-treatment of Bax-inhibiting peptide reduces neuronal death and behavioral deficits following global cerebral ischemia. *Neurochem. Int.* 58 (2), 224–233. doi:10.1016/j.neuint.2010.12.008
- Han, D., and Zhou, Y. (2019). YY1-induced upregulation of lncRNA NEAT1 contributes to OGD/R injury-induced inflammatory response in cerebral microglial cells via Wnt/ β -catenin signaling pathway. *Vitro Cell. Dev. Biol. Anim.* 55 (7), 501–511. doi:10.1007/s11626-019-00375-y
- Herpich, F., and Rincon, F. (2020). Management of acute ischemic stroke. *Crit. Care Med.* 48 (11), 1654–1663. doi:10.1097/CCM.0000000000004597
- Hoss, F., Rodriguez-Alcázar, J. F., and Latz, E. (2017). Assembly and regulation of ASC specks. *Cell. Mol. Life Sci.* 74 (7), 1211–1229. doi:10.1007/s00018-016-2396-6
- Hu, L., Chen, M., Chen, X., Zhao, C., Fang, Z., Wang, H., et al. (2020). Chemotherapy-induced pyroptosis is mediated by BAK/BAX-caspase-3-GSDME pathway and inhibited by 2-bromopalmitate. *Cell Death Dis.* 11 (4), 281. doi:10.1038/s41419-020-2476-2
- Huang, H. Y., Lin, Y. C., Li, 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 (D1), D148–D154. doi:10.1093/nar/gkz896
- Huo, X., Zhang, K., Yi, L., Mo, Y., Liang, Y., Zhao, J., et al. (2016). Decreased epithelial and plasma miR-181b-5p expression associates with airway eosinophilic inflammation in asthma. *Clin. Exp. Allergy* 46 (10), 1281–1290. doi:10.1111/cea.12754
- Hutchison, E. R., Kawamoto, E. M., Taub, D. D., Lal, A., Abdelmohsen, K., Zhang, Y., et al. (2013). Evidence for miR-181 involvement in neuroinflammatory responses of astrocytes. *Glia* 61 (7), 1018–1028. doi:10.1002/glia.22483
- Indrieri, A., Carrella, S., Carotenuto, P., Banfi, S., and Franco, B. (2020). The pervasive role of the miR-181 family in development, neurodegeneration, and cancer. *Int. J. Mol. Sci.* 21 (6), 2092. doi:10.3390/ijms21062092
- Jassal, B., Matthews, L., Viteri, G., Gong, C., Lorente, P., Fabregat, A., et al. (2020). The reactome pathway knowledgebase. *Nucleic Acids Res.* 48 (D1), D498–D503. doi:10.1093/nar/gkz1031
- Jayaraj, R. L., Azimullah, S., Beiram, R., Jalal, F. Y., and Rosenberg, G. A. (2019). Neuroinflammation: Friend and foe for ischemic stroke. *J. Neuroinflammation* 16 (1), 142. doi:10.1186/s12974-019-1516-2
- Jia, Y., Yi, L., Li, Q., Liu, T., and Yang, S. (2021). LncRNA MALAT1 aggravates oxygen-glucose deprivation/reoxygenation-induced neuronal endoplasmic reticulum stress and apoptosis via the miR-195a-5p/HMGA1 axis. *Biol. Res.* 54 (1), 8. doi:10.1186/s40659-021-00331-9
- Jiao, H., Wachsmuth, L., Kumari, S., Schwarzer, R., Lin, J., Eren, R. O., et al. (2020). Z-nucleic-acid sensing triggers ZBP1-dependent necroptosis and inflammation. *Nature* 580 (7803), 391–395. doi:10.1038/s41586-020-2129-8
- Jin, F., Ou, W., Wei, B., Fan, H., Wei, C., Fang, D., et al. (2021). Transcriptome-wide analysis to identify the inflammatory role of lncRNA *Neat1* in experimental ischemic stroke. *J. Inflamm. Res.* 14, 2667–2680. doi:10.2147/JIR.S315281
- Jurcau, A., and Simion, A. (2021). Neuroinflammation in cerebral ischemia and ischemia/reperfusion injuries: From pathophysiology to therapeutic strategies. *Int. J. Mol. Sci.* 23 (1), 14. doi:10.3390/ijms23010014
- Kanehisa, M., Furumichi, M., Tanabe, M., Sato, Y., and Morishima, K. (2017). Kegg: New perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res.* 45 (D1), D353–D361. doi:10.1093/nar/gkw1092
- Karagkouni, D., Paraskevopoulou, M. D., Tastsoglou, S., Skoufos, G., Karavangeli, A., Pierros, V., et al. (2020). DIANA-LncBase v3: Indexing experimentally supported miRNA targets on non-coding transcripts. *Nucleic Acids Res.* 48 (D1), D101–D110. doi:10.1093/nar/gkz1036
- Karki, R., Sharma, B. R., Lee, E., Banoth, B., Malireddi, R. K. S., Samir, P., et al. (2020). Interferon regulatory factor 1 regulates PANoptosis to prevent colorectal cancer. *JCI Insight* 5 (12), 136720. doi:10.1172/jci.insight.136720

- Karki, R., Sharma, B. R., Tuladhar, S., Williams, E. P., Zalduondo, L., Samir, P., et al. (2021). Synergism of TNF- α and IFN- γ triggers inflammatory cell death, tissue damage, and mortality in SARS-CoV-2 infection and cytokine shock syndromes. *Cell* 184 (1), 149–168.e17. doi:10.1016/j.cell.2020.11.025
- Krol, J., Loedige, I., and Filipowicz, W. (2010). The widespread regulation of microRNA biogenesis, function and decay. *Nat. Rev. Genet.* 11 (9), 597–610. doi:10.1038/nrg2843
- Lee, S., Karki, R., Wang, Y., Nguyen, L. N., Kalathur, R. C., and Kanneganti, T. D. (2021). AIM2 forms a complex with pyrin and ZBP1 to drive PANoptosis and host defence. *Nature* 597 (7876), 415–419. doi:10.1038/s41586-021-03875-8
- Li, J. H., Liu, S., Zhou, H., Qu, L. H., and Yang, J. H. (2014). starBase v2.0: decoding miRNA-cRNA, miRNA-ncRNA and protein-RNA interaction networks from large-scale CLIP-Seq data. *Nucleic Acids Res.* 42 (Database issue), D92–D97. doi:10.1093/nar/gkt1248
- Li, J., Zhang, J., Zhang, Y., Wang, Z., Song, Y., Wei, S., et al. (2019). TRAF2 protects against cerebral ischemia-induced brain injury by suppressing necroptosis. *Cell Death Dis.* 10 (5), 328. doi:10.1038/s41419-019-1558-5
- Li, P., Duan, S., and Fu, A. (2020). Long noncoding RNA NEAT1 correlates with higher disease risk, worse disease condition, decreased miR-124 and miR-125a and predicts poor recurrence-free survival of acute ischemic stroke. *J. Clin. Lab. Anal.* 34 (2), e23056. doi:10.1002/jcla.23056
- Li, Y., Xiang, L., Wang, C., Song, Y., Miao, J., and Miao, M. (2021). Protection against acute cerebral ischemia/reperfusion injury by Leonuri Herba Total Alkali via modulation of BDNF-TrkB-PI3K/Akt signaling pathway in rats. *Biomed. Pharmacother.* 133, 111021. doi:10.1016/j.biopha.2020.111021
- Liang, Y., Song, P., Zhu, Y., Xu, J., Zhu, P., Liu, R., et al. (2020). TREM-1-targeting LP17 attenuates cerebral ischemia-induced neuronal injury by inhibiting oxidative stress and pyroptosis. *Biochem. Biophys. Res. Commun.* 529 (3), 554–561. doi:10.1016/j.bbrc.2020.05.056
- Liao, S., Apaijai, N., Chattapakorn, N., and Chattapakorn, S. C. (2020). The possible roles of necroptosis during cerebral ischemia and ischemia/reperfusion injury. *Arch. Biochem. Biophys.* 695, 108629. doi:10.1016/j.abb.2020.108629
- Liu, X., Hou, L., Huang, W., Gao, Y., Lv, X., and Tang, J. (2016). The mechanism of long non-coding RNA MEG3 for neurons apoptosis caused by hypoxia: Mediated by miR-181b-12/15-LOX signaling pathway. *Front. Cell. Neurosci.* 10, 201. doi:10.3389/fncel.2016.00201
- Lu, Y., Xu, X., Dong, R., Sun, L., Chen, L., Zhang, Z., et al. (2019). MicroRNA-181b-5p attenuates early postoperative cognitive dysfunction by suppressing hippocampal neuroinflammation in mice. *Cytokine* 120, 41–53. doi:10.1016/j.cyt.2019.04.005
- Ma, Q., Zhao, H., Tao, Z., Wang, R., Liu, P., Han, Z., et al. (2016). MicroRNA-181c exacerbates brain injury in acute ischemic stroke. *Aging Dis.* 7 (6), 705–714. doi:10.14336/AD.2016.0320
- Maddahi, A., Kruse, L. S., Chen, Q. W., and Edvinsson, L. (2011). The role of tumor necrosis factor- α and TNF- α receptors in cerebral arteries following cerebral ischemia in rat. *J. Neuroinflammation* 8, 107. doi:10.1186/1742-2094-8-107
- Malireddi, R., Tweedell, R. E., and Kanneganti, T. D. (2020). PANoptosis components, regulation, and implications. *Aging (Albany NY)* 12 (12), 11163–11164. doi:10.18632/aging.103528
- Martens, M., Ammar, A., Riutta, A., Waagmeester, A., Slenter, D. N., Hanspers, K., et al. (2021). WikiPathways: Connecting communities. *Nucleic Acids Res.* 49 (D1), D613–D621. doi:10.1093/nar/gkaa1024
- Mendelson, S. J., and Prabhakaran, S. (2021). Diagnosis and management of transient ischemic attack and acute ischemic stroke: A review. *JAMA* 325 (11), 1088–1098. doi:10.1001/jama.2020.26867
- Meng, C., Zhang, J., Zhang, L., Wang, Y., Li, Z., and Zhao, J. (2019). Effects of NLRP6 in cerebral ischemia/reperfusion (I/R) injury in rats. *J. Mol. Neurosci.* 69 (3), 411–418. doi:10.1007/s12031-019-01370-4
- Meng, Q., Ye, C., and Lu, Y. (2020). miR-181c regulates ischemia/reperfusion injury-induced neuronal cell death by regulating c-Fos signaling. *Pharmazie* 75 (2), 90–93. doi:10.1691/ph.2020.9856
- Michalicova, A., Majerova, P., and Kovac, A. (2020). Tau protein and its role in blood–brain barrier dysfunction[J]. *Front. Mol. Neurosci.*, 13.
- Moon, J. M., Xu, L., and Giffard, R. G. (2013). Inhibition of microRNA-181 reduces forebrain ischemia-induced neuronal loss. *J. Cereb. Blood Flow. Metab.* 33 (12), 1976–1982. doi:10.1038/jcbfm.2013.157
- Ni, X., Su, Q., Xia, W., Zhang, Y., Jia, K., Su, Z., et al. (2020). Knockdown lncRNA NEAT1 regulates the activation of microglia and reduces AKT signaling and neuronal apoptosis after cerebral ischemic reperfusion. *Sci. Rep.* 10 (1), 19658. doi:10.1038/s41598-020-71411-1
- Ouyang, Y. B., Lu, Y., Yue, S., Xu, L. J., Xiong, X. X., White, R. E., et al. (2012). miR-181 regulates GRP78 and influences outcome from cerebral ischemia *in vitro* and *in vivo*. *Neurobiol. Dis.* 45 (1), 555–563. doi:10.1016/j.nbd.2011.09.012
- Peng, Z., Li, J., Li, Y., Yang, X., Feng, S., Han, S., et al. (2013). Downregulation of miR-181b in mouse brain following ischemic stroke induces neuroprotection against ischemic injury through targeting heat shock protein A5 and ubiquitin carboxyl-terminal hydrolase isozyme L1. *J. Neurosci. Res.* 91 (10), 1349–1362. doi:10.1002/jnr.23255
- Place, D. E., Lee, S., and Kanneganti, T. D. (2021). PANoptosis in microbial infection. *Curr. Opin. Microbiol.* 59, 42–49. doi:10.1016/j.mib.2020.07.012
- Radak, D., Katsiki, N., Resanovic, I., Jovanovic, A., Sudar-Milovanovic, E., Zafirovic, S., et al. (2017). Apoptosis and acute brain ischemia in ischemic stroke. *Curr. Vasc. Pharmacol.* 15 (2), 115–122. doi:10.2174/1570161115666161104095522
- 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 (7), e47. doi:10.1093/nar/gkv007
- Samir, P., Malireddi, R. K. S., and Kanneganti, T. (2020). The PANoptosome: A deadly protein complex driving pyroptosis, apoptosis, and necroptosis (PANoptosis). *Front. Cell. Infect. Microbiol.* 10, 238. doi:10.3389/fcimb.2020.00238
- 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
- Shi, Y. L., Wang, Q., and Wei, J. C. (2019). Influence of lncRNA-MALAT1 on neuronal apoptosis in rats with cerebral infarction through regulating the ERK/MAPK signaling pathway. *Eur. Rev. Med. Pharmacol. Sci.* 23 (18), 8039–8048. doi:10.26355/eurrev_201909_19020
- Shi, Y., Zhang, X., and Ying, P., (2021). Neuroprotective effect of astragaloside IV on cerebral ischemia/reperfusion injury rats through sirt1/mapt pathway[J]. *Front. Pharmacol.*, 12.
- Shu, B., Zhou, Y. X., Li, H., Zhang, R. Z., He, C., and Yang, X. (2021). The METTL3/MALAT1/PTBP1/USP8/TAK1 axis promotes pyroptosis and M1 polarization of macrophages and contributes to liver fibrosis. *Cell Death Discov.* 7 (1), 368. doi:10.1038/s41420-021-00756-x
- Song, X., Xue, Y., and Cai, H. (2021). Down-regulation of miR-181a-5p prevents cerebral ischemic injury by upregulating En2 and activating wnt/ β -catenin pathway. *J. Stroke Cerebrovasc. Dis.* 30 (3), 105485. doi:10.1016/j.jstrokecerebrovasdis.2020.105485
- Song, Y., Yang, L., Guo, R., Lu, N., Shi, Y., and Wang, X. (2019). Long noncoding RNA MALAT1 promotes high glucose-induced human endothelial cells pyroptosis by affecting NLRP3 expression through competitively binding miR-22. *Biochem. Biophys. Res. Commun.* 509 (2), 359–366. doi:10.1016/j.bbrc.2018.12.139
- Spitz, A. Z., and Gavathiotis, E. (2022). Physiological and pharmacological modulation of BAX. *Trends Pharmacol. Sci.* 43 (3), 206–220. doi:10.1016/j.tips.2021.11.001
- Stary, C. M., Xu, L., Li, L., Sun, X., Ouyang, Y. B., Xiong, X., et al. (2017). Inhibition of miR-181a protects female mice from transient focal cerebral ischemia by targeting astrocyte estrogen receptor- α . *Mol. Cell. Neurosci.* 82, 118–125. doi:10.1016/j.mcn.2017.05.004
- 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
- Tan, X., Guo, W., Peng, Z., Gu, C., Xiang, P., Tu, Y., et al. (2021). lncRNA-Malat1 down-regulates miR-211-5p expression to promote neuronal damage from cerebral ischemia reperfusion injury. *Biochem. Pharmacol.* 192, 114694. doi:10.1016/j.bcp.2021.114694
- Tu, Y., and Hu, Y. (2021). MiRNA-34c-5p protects against cerebral ischemia/reperfusion injury: Involvement of anti-apoptotic and anti-inflammatory activities. *Metab. Brain Dis.* 36 (6), 1341–1351. doi:10.1007/s11011-021-00724-5
- Tuo, Q. Z., Zhang, S. T., and Lei, P. (2022). Mechanisms of neuronal cell death in ischemic stroke and their therapeutic implications. *Med. Res. Rev.* 42 (1), 259–305. doi:10.1002/med.21817
- Wang, H., Zheng, X., Jin, J., Zheng, L., Guan, T., Huo, Y., et al. (2020). lncRNA MALAT1 silencing protects against cerebral ischemia-reperfusion injury through miR-145 to regulate AQP4. *J. Biomed. Sci.* 27 (1), 40. doi:10.1186/s12929-020-00635-0
- Wang, J., Fu, Z., Wang, M., Lu, J., Yang, H., and Lu, H. (2021). Knockdown of XIST attenuates cerebral ischemia/reperfusion injury through regulation of miR-362/ROCK2 Axis. *Neurochem. Res.* 46 (8), 2167–2180. doi:10.1007/s11064-021-03354-6

- Wang, L., Li, S., Stone, S. S., Liu, N., Gong, K., Ren, C., et al. (2022). The role of the lncRNA MALAT1 in neuroprotection against hypoxic/ischemic injury. *Biomolecules* 12 (1), 146. doi:10.3390/biom12010146
- Wang, Y., and Kanneganti, T. (2021). From pyroptosis, apoptosis and necroptosis to PANoptosis: A mechanistic compendium of programmed cell death pathways. *Comput. Struct. Biotechnol. J.* 19, 4641–4657. doi:10.1016/j.csbj.2021.07.038
- Watters, O., and O'Connor, J. J. (2011). A role for tumor necrosis factor- α in ischemia and ischemic preconditioning. *J. Neuroinflammation* 8, 87. doi:10.1186/1742-2094-8-87
- Weng, S., Wang, S., and Jiang, J. (2021). Long noncoding RNA X-inactive specific transcript regulates neuronal cell apoptosis in ischemic stroke through miR-98/BACH1 Axis. *DNA Cell Biol.* 40 (7), 979–987. doi:10.1089/dna.2020.6354
- Wu, J., Fan, C. L., Ma, L. J., Liu, T., Wang, C., Song, J. X., et al. (2017). Distinctive expression signatures of serum microRNAs in ischaemic stroke and transient ischaemic attack patients. *Thromb. Haemost.* 117 (5), 992–1001. doi:10.1160/TH16-08-0606
- Wu, P., Zuo, X., Deng, H., Liu, X., Liu, L., and Ji, A. (2013). Roles of long noncoding RNAs in brain development, functional diversification and neurodegenerative diseases. *Brain Res. Bull.* 97, 69–80. doi:10.1016/j.brainresbull.2013.06.001
- Xin, J. W., and Jiang, Y. G. (2017). Long noncoding RNA MALAT1 inhibits apoptosis induced by oxygen-glucose deprivation and reoxygenation in human brain microvascular endothelial cells. *Exp. Ther. Med.* 13 (4), 1225–1234. doi:10.3892/etm.2017.4095
- Xiong, F., Wei, W. P., Liu, Y. B., Wang, Y., Zhang, H. Y., and Liu, R. (2021). Long noncoding RNA XIST enhances cerebral ischemia-reperfusion injury by regulating miR-486-5p and GAB2. *Eur. Rev. Med. Pharmacol. Sci.* 25 (4), 2013–2020. doi:10.26355/eurrev_202102_25103
- Xu, Q., Guohui, M., Li, D., Bai, F., Fang, J., Zhang, G., et al. (2021). lncRNA C2dat2 facilitates autophagy and apoptosis via the miR-30d-5p/DDIT4/mTOR axis in cerebral ischemia-reperfusion injury. *Aging (Albany NY)* 13 (8), 11315–11335. doi:10.18632/aging.202824
- Xu, Q., Zhao, B., Ye, Y., Li, Y., Zhang, Y., Xiong, X., et al. (2021). Relevant mediators involved in and therapies targeting the inflammatory response induced by activation of the NLRP3 inflammasome in ischemic stroke. *J. Neuroinflammation* 18 (1), 123. doi:10.1186/s12974-021-02137-8
- Yan, W., Yang, Y., Hu, X., Ning, W. Y., Liao, L. S., Lu, S., et al. (2022). Do pyroptosis, apoptosis, and necroptosis (PANoptosis) exist in cerebral ischemia? Evidence from cell and rodent studies[J]. *Neural Regen. Res.* 17 (8), 1761–1768. doi:10.4103/1673-5374.331539
- Yan, Y., Chen, L., Zhou, J., and Xie, L. (2020). SNHG12 inhibits oxygen-glucose deprivation-induced neuronal apoptosis via the miR-181a-5p/NEGR1 axis. *Mol. Med. Rep.* 22 (5), 3886–3894. doi:10.3892/mmr.2020.11459
- Yao, D., Zhang, S., Hu, Z., Luo, H., Mao, C., Fan, Y., et al. (2021). CHIP ameliorates cerebral ischemia-reperfusion injury by attenuating necroptosis and inflammation. *Aging (Albany NY)* 13 (23), 25564–25577. doi:10.18632/aging.203774
- Yu, P., Zhang, X., Liu, N., Tang, L., Peng, C., and Chen, X. (2021). Pyroptosis: Mechanisms and diseases. *Signal Transduct. Target. Ther.* 6 (1), 128. doi:10.1038/s41392-021-00507-5
- Zhang, G., Wang, Q., and Su, D. (2021). Long non-coding RNAMALAT1 knockdown alleviates cerebral ischemia/reperfusion injury of rats through regulating the miR-375/pde4d Axis[J]. *Front. Neurology*, 11.
- Zhang, H., Ouyang, B., Ji, X., and Liu, M. F. (2021). Gastrodin alleviates cerebral ischaemia/reperfusion injury by inhibiting pyroptosis by regulating the lncRNA NEAT1/miR-22-3p Axis. *Neurochem. Res.* 46 (7), 1747–1758. doi:10.1007/s11064-021-03285-2
- Zhang, H., Xia, J., Hu, Q., Xu, L., Cao, H., Wang, X., et al. (2021). Long non-coding RNA XIST promotes cerebral ischemia/reperfusion injury by modulating miR-27a-3p/FOXO3 signaling. *Mol. Med. Rep.* 24 (2), 566. doi:10.3892/mmr.2021.12205
- Zhang, H., Zhong, K., Lu, M., Mei, Y., Tan, E., Sun, X., et al. (2018). Neuroprotective effects of isosteviol sodium through increasing CYLD by the downregulation of miRNA-181b. *Brain Res. Bull.* 140, 392–401. doi:10.1016/j.brainresbull.2018.05.015
- Zhang, T., Yin, C., Boyd, D. F., Quarato, G., Ingram, J. P., Shubina, M., et al. (2020). Influenza virus Z-RNAs induce ZBP1-mediated necroptosis. *Cell* 180 (6), 1115–1120. doi:10.1016/j.cell.2020.02.050
- Zhang, X., Liu, Z., Shu, Q., Yuan, S., Xing, Z., and Song, J. (2019). lncRNA SNHG6 functions as a ceRNA to regulate neuronal cell apoptosis by modulating miR-181c-5p/BIM signalling in ischaemic stroke. *J. Cell. Mol. Med.* 23 (9), 6120–6130. doi:10.1111/jcmm.14480
- Zhang, X., Tang, X., Liu, K., Hamblin, M. H., and Yin, K. J. (2017). Long noncoding RNA Malat1 regulates cerebrovascular pathologies in ischemic stroke. *J. Neurosci.* 37 (7), 1797–1806. doi:10.1523/JNEUROSCI.3389-16.2017
- Zhang, Y., Tian, Q., Zhang, Q., Zhou, X., Liu, S., and Wang, J. Z. (2009). Hyperphosphorylation of microtubule-associated tau protein plays dual role in neurodegeneration and neuroprotection. *Pathophysiology* 16 (4), 311–316. doi:10.1016/j.pathophys.2009.02.003
- Zhang, Y., Wang, S., Lu, F., Zhang, M., Kong, H., Cheng, J., et al. (2021). The neuroprotective effect of pretreatment with carbon dots from Crinis Carbonisatus (carbonized human hair) against cerebral ischemia reperfusion injury. *J. Nanobiotechnology* 19 (1), 257. doi:10.1186/s12951-021-00908-2
- Zheng, M., and Kanneganti, T. D. (2020). The regulation of the ZBP1-NLRP3 inflammasome and its implications in pyroptosis, apoptosis, and necroptosis (PANoptosis). *Immunol. Rev.* 297 (1), 26–38. doi:10.1111/imr.12909
- 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, F., Wang, Y. K., Zhang, C. G., and Wu, B. Y. (2021). miR-19a/b-3p promotes inflammation during cerebral ischemia/reperfusion injury via SIRT1/FoxO3/SPHK1 pathway. *J. Neuroinflammation* 18 (1), 122. doi:10.1186/s12974-021-02172-5
- 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
- Zhou, Z., Ren, X., Zheng, L., Li, A. P., and Zhou, W. S. (2022). lncRNA NEAT1 ameliorate ischemic stroke via promoting Mfn2 expression through binding to Nova and activates Sirt3. *Metab. Brain Dis.* 37 (3), 653–664. doi:10.1007/s11011-021-00895-1



OPEN ACCESS

EDITED BY

Jaqueline Bohrer Schuch,
Federal University of Rio Grande do
Sul, Brazil

REVIEWED BY

Thiago Henrique Roza,
Federal University of Rio Grande do
Sul, Brazil
Rammohan Shukla,
University of Toledo, United States
Yun-Ai Su,
Peking University Sixth Hospital, China

*CORRESPONDENCE

Nan Lyu
lvnangirl@163.com

†These authors have contributed
equally to this work

SPECIALTY SECTION

This article was submitted to
Neurogenomics,
a section of the journal
Frontiers in Neuroscience

RECEIVED 21 May 2022

ACCEPTED 25 July 2022

PUBLISHED 08 August 2022

CITATION

Liu S, Lu T, Zhao Q, Fu B, Wang H, Li G,
Yang F, Huang J and Lyu N (2022) A
machine learning model for predicting
patients with major depressive
disorder: A study based on
transcriptomic data.
Front. Neurosci. 16:949609.
doi: 10.3389/fnins.2022.949609

COPYRIGHT

© 2022 Liu, Lu, Zhao, Fu, Wang, Li,
Yang, Huang and Lyu. This is an
open-access article distributed under
the terms of the [Creative Commons
Attribution License \(CC BY\)](#). The use,
distribution or reproduction in other
forums is permitted, provided the
original author(s) and the copyright
owner(s) are credited and that the
original publication in this journal is
cited, in accordance with accepted
academic practice. No use, distribution
or reproduction is permitted which
does not comply with these terms.

A machine learning model for predicting patients with major depressive disorder: A study based on transcriptomic data

Sitong Liu^{1,2†}, Tong Lu^{3†}, Qian Zhao^{1,2}, Bingbing Fu^{1,2},
Han Wang^{1,2}, Ginhong Li^{1,2}, Fan Yang^{1,2}, Juan Huang^{1,2} and
Nan Lyu^{1,2*}

¹The National Clinical Research Center for Mental Disorders & Beijing Key Laboratory of Mental Disorders, Beijing Anding Hospital, Capital Medical University, Beijing, China, ²Advanced Innovation Center for Human Brain Protection, Capital Medical University, Beijing, China, ³Department of Thoracic Surgery, The Second Affiliated Hospital of Harbin Medical University, Harbin, China

Background: Identifying new biomarkers of major depressive disorder (MDD) would be of great significance for its early diagnosis and treatment. Herein, we constructed a diagnostic model of MDD using machine learning methods.

Methods: The GSE98793 and GSE19738 datasets were obtained from the Gene Expression Omnibus database, and the limma R package was used to analyze differentially expressed genes (DEGs) in MDD patients. Gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analyses were performed to identify potential molecular functions and pathways. A protein-protein interaction network (PPI) was constructed, and hub genes were predicted. Random forest (RF) and artificial neural network (ANN) machine-learning algorithms were used to select variables and construct a robust diagnostic model.

Results: A total of 721 DEGs were identified in peripheral blood samples of patients with MDD. GO and KEGG analyses revealed that the DEGs were mainly enriched in cytokines, defense responses to viruses, responses to biotic stimuli, immune effector processes, responses to external biotic stimuli, and immune systems. A PPI network was constructed, and CytoHubba plugins were used to screen hub genes. Furthermore, a robust diagnostic model was established using a RF and ANN algorithm with an area under the curve of 0.757 for the training model and 0.685 for the test cohort.

Conclusion: We analyzed potential driver genes in patients with MDD and built a potential diagnostic model as an adjunct tool to assist psychiatrists in the clinical diagnosis and treatment of MDD.

KEYWORDS

major depressive disorder, machine learning, random forest, artificial neural network, bioinformatics analysis

Introduction

Major depressive disorder (MDD) is a general chronic psychiatric disorder affecting people of all ages, which can ultimately lead to chronic disability, financial difficulties, and shortened life expectancy (Murray and Lopez, 1997; Zhdanova et al., 2021). In recent years, as increasing importance has been placed on the treatment of mental illness, the proportion of patients seeking treatment for MDD rose from 43.5% in 2007–2008 to 52.9% in 2015–2016 (Rhee et al., 2020). In China, MDD is the most common mood disorder, with a lifetime prevalence of 3.4% and a prevalence of 2.1% at 12 months (Huang et al., 2019). However, the misdiagnosis rate of MDD can reach as high as 78%, and misdiagnosis often leads to improper treatment (Fernández et al., 2010). The early identification of MDD is particularly important; therefore, it is of great significance to identify new and feasible biomarkers for the early diagnosis and treatment of MDD.

Obtaining peripheral blood biomarkers is a more convenient and practical method of diagnosis than brain imaging or biopsy. Recently, many studies have focused on the use of mRNA expression data from peripheral blood groups to investigate the differential characteristics between patients with MDD and healthy populations. For example, Woo et al. (2018) analyzed gene expression in peripheral blood of 38 patients with MDD and 14 healthy controls and identified seven differentially expressed genes (DEGs). Spijker et al. (2010) studied the peripheral blood of 21 patients with MDD and 21 healthy control participants and found significant differences in the expression levels of CAPRN1, CLEC4A, CKRT23, MLC1, PLSCR1, PROK2, and ZBTB16, indicating that this signature could distinguish patients with depression from healthy individuals. Therefore, a predictive diagnostic model based on mRNA expression could help us understand the potential pathophysiology of MDD and may further support clinical decisions.

Currently, machine learning is increasingly applied in the medical field and has come to play an important role in the diagnosis and prognosis in the fields of oncology, neurology, and cardiology. For example, Ciobanu et al. (2020) used a random forest (RF) approach to predict depression and suicide risk in 39 patients with depression and 87 healthy controls using blood methylation and transcriptomic data, with an accuracy of 87.3% in distinguishing between the two groups. Bill et al.

used a regularization gradient enhancement machine to classify microarray gene expression data in the blood of 1581 patients with MDD and 369 controls, with an average area under the curve (AUC) of 0.64 (Qi et al., 2021). An artificial neural network (ANN) model has further been applied in the diagnosis of many asymptomatic and early diseases (He et al., 2020) using a neural network classifier to the voice of patients with depression. The control group variables were analyzed, and the diagnostic accuracy rate was between 82.40 and 93.02% (Navarro et al., 2019). ANNs can accurately predict the positive or negative effects of Alzheimer's disease and *Mycobacterium tuberculosis*, with a total accuracy of 93.8 and 94% (Khan et al., 2019; Swietlik and Bialowas, 2019). However, no ANN-related diagnostic model has been applied to construct an auxiliary peripheral blood diagnostic model for patients with MDD.

In this study, we obtained the MDD-related datasets, GSE98793 and GSE19738, and analyzed DEGs in patients with MDD. Functional analyses, including gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG), were performed to investigate the enriched molecular functions and pathways. Machine learning algorithms, including RF and ANN, were used for variable selection and diagnostic model construction. Based on these results, we compared the discrimination and accuracy of single genes and diagnostic models for MDD.

Materials and methods

Data collection and data processing

The GSE98793 and GSE19738 datasets were searched from the Gene Expression Omnibus database¹ using the following keywords: “MDD, blood, normal” [All Fields] AND “Homo sapiens” AND “Expression profiling by array” [All Fields]. The screening standards for microarray datasets included the following: reference to profiles of gene expression with genome-wide whole blood; containing samples from patients with MDD and healthy controls; all included samples were not treated with drugs; the number of samples was greater than 40. Eventually, GSE98793 (Leday G. G. R. et al., 2018) and GSE19738 (Spijker et al., 2010) were screened for in-depth investigation. The GSE98793 dataset was provided by Kelly et al., who examined a total of 192 peripheral blood samples, including 128 from MDD patients and 64 from healthy volunteers. And Affymetrix Human Genome U133 Plus 2.0 Array was used to test. GSE19738 data set was provided by Spijker, using the chip Agilent-012391 Whole Human Genome Oligo Microarray G4112A. They detected 132 peripheral blood samples in total from 34 healthy volunteers and 33 MDD

TABLE 1 The information of datasets.

| Dataset | Platform | Organism | Tissue | Sample | |
|----------|----------|--------------|--------|--------|---------|
| | | | | Normal | Disease |
| GSE19738 | GPL6848 | Homo Sapiens | Blood | 34 | 33 |
| GSE98793 | GPL570 | Homo Sapiens | Blood | 64 | 128 |

¹ <https://www.ncbi.nlm.nih.gov/geo/>

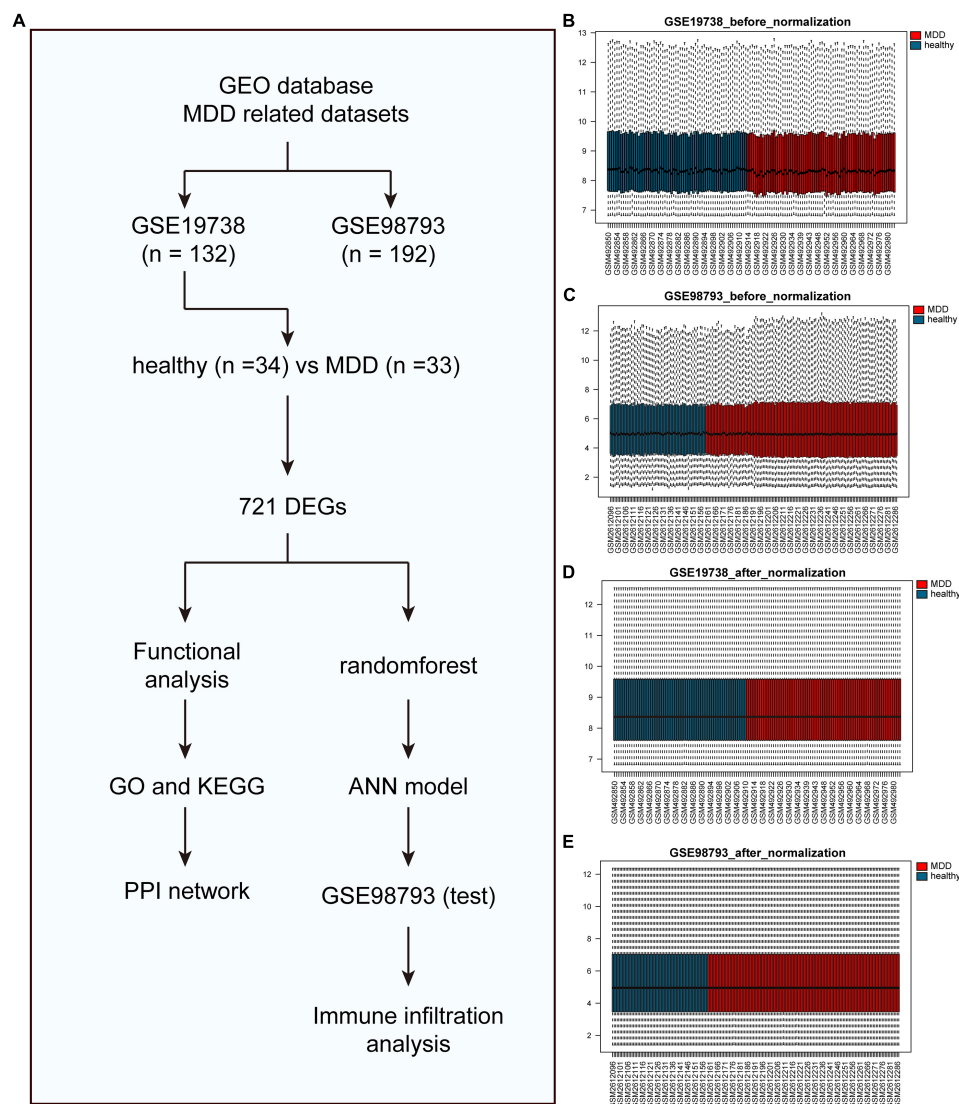


FIGURE 1

Before and after box diagram of standardization of GSE19738 and GSE98793 datasets. (A) The flow chart of this study. (B) Box diagram of the GSE19738 dataset before correction; (C) Box diagram of the GSE98793 dataset before correction; (D) Box diagram of the GSE19738 dataset after correction; and (E) Box diagram of the GSE98793 dataset after correction; Red represents the MDD samples, and blue represents the normal samples.

patients, respectively. The information for patients with MDD and healthy participants was provided in Table 1.

Differential expression analysis

The limma package in R software was used for standardized processing of the datasets to eliminate changes in gene expression caused by experimental techniques, and the normalized data were used for subsequent analysis. Differential analysis was carried out on the GSE19738 dataset, and the cutoff value was set to $|\log_2FC| > 1$, adj. $P < 0.05$.

Functional enrichment analysis and protein-protein interaction network

To explore the function and pathways of the identified DEGs, GO and KEGG pathway enrichment analyses were performed using the clusterProfiler R package. Statistical significance was set at $P < 0.05$. The protein-protein interaction (PPI) network between the DEGs was analyzed using the STRING database². The interaction score was set

² <https://string-db.org>

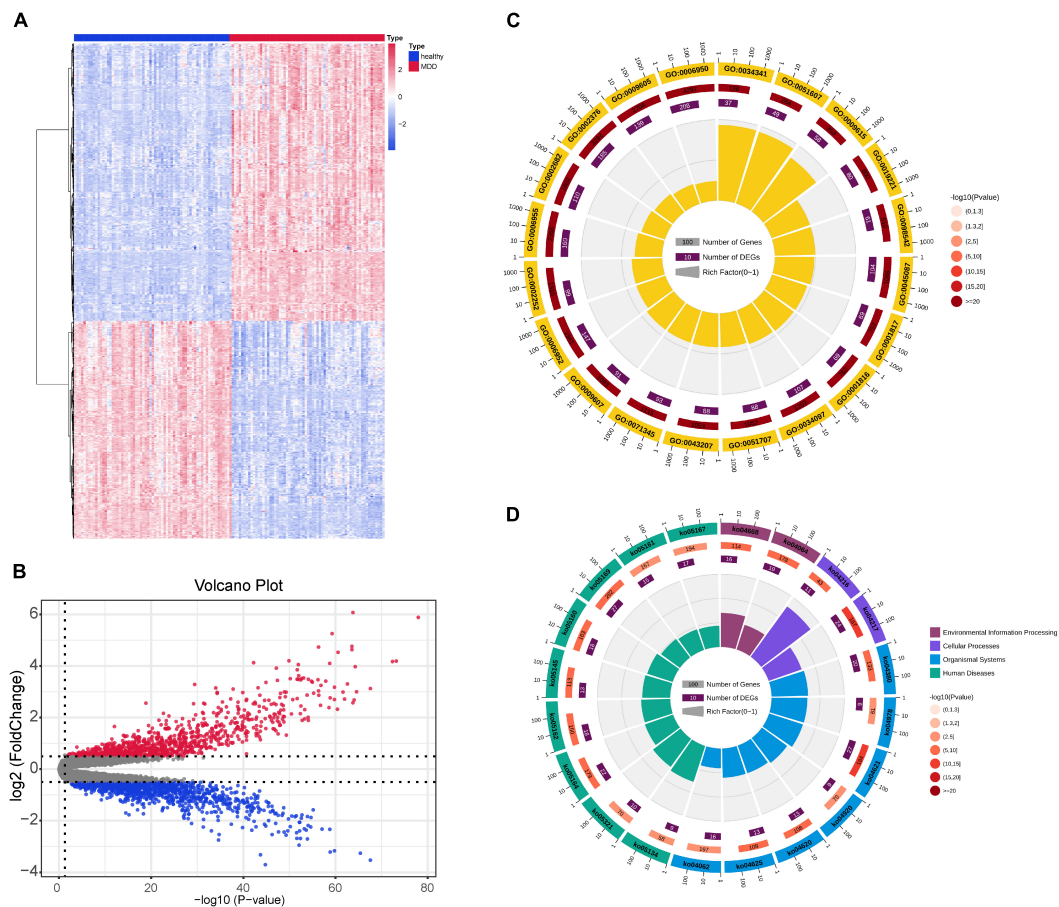


FIGURE 2

DEG identification of the GSE19738 dataset. (A) DEGs of peripheral blood samples from patients with MDD and healthy controls were obtained from the DEG heat map constructed from the GSE19738 dataset. Horizontal coordinate blue represents the control group, red represents the experimental group, blue indicates low expression, and red indicates high expression. (B) Volcano diagram, black indicates genes with no differential expression, blue indicates down-regulated genes, and red indicates up-regulated genes. (C) GO enrichment analysis. The outer circle represents the number of GO term, the outer circle number represents all genes in GO term, and the inner circle number represents the number of enriched genes. The inner circle pie chart represents the percentage of genes that are enriched. (D) KEGG pathways. The outer circle represents the KEGG ID, the outer circle number represents all genes in the KEGG pathway, and the inner circle number represents the number of genes enriched in the pathway. The inner circle pie chart represents the percentage of genes that are enriched. DEGs, differentially expressed genes; GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes.

to 0.900 (highest confidence), and the nodes in the network were randomly clustered using the *k*-means algorithm to reveal potential regulatory relations between nodes. Subsequently, the ten most significant hub genes were screened using cytoHubba.

Variable selection and diagnostic model construction

The RF algorithm is a classifier that contains multiple decision trees. It uses the RF package for analysis and sorts genes according to their importance. Genes with high importance were extracted from the list of different genes for visualization.

After the MDD characteristic genes were screened using the RF algorithm, information redundancy was removed by collinearity analysis. Taking the threshold of the Spearman Rho absolute value as >0.5 , the parameters with collinearity were removed, and the model was further constructed using an ANN.

The ANN consists of the following three layers: the input, hidden, and output layers of six, five, and two neurons, respectively. An ANN software simulator was used to solve the return of the mission, including the forecast revision numbers. The network's answer to each test case ranged from 0 to 1. The level of activation and inhibition of the output neurons is automatically selected by the stimulator of the ANN to minimize losses. The error function of the ANN was chosen as the sum of the square of the prior given value and the actual value of the output neuron. GSE19738 and GSE98793 were used as the

TABLE 2 GO enrichment analyses results.

| GO ID | Description | Number | P value | Q value |
|-------------|--|--------|----------|----------|
| GO: 0034341 | Response to interferon-gamma | 37 | 1.46E-24 | 4.15E-22 |
| GO: 0051607 | Defense response to virus | 49 | 3.74E-32 | 2.36E-29 |
| GO:0009615 | Response to virus | 58 | 5.70E-35 | 4.63E-32 |
| GO:0019221 | Cytokine-mediated signaling pathway | 80 | 2.27E-32 | 1.61E-29 |
| GO:0098542 | Defense response to other organism | 61 | 1.42E-24 | 4.15E-22 |
| GO:0045087 | Innate immune response | 104 | 8.81E-42 | 1.25E-38 |
| GO:0001817 | Regulation of cytokine production | 69 | 2.04E-25 | 6.84E-23 |
| GO:0001816 | Cytokine production | 69 | 3.01E-25 | 9.51E-23 |
| GO:0034097 | Response to cytokine | 107 | 2.77E-39 | 2.62E-36 |
| GO:0051707 | Response to other organism | 88 | 7.14E-30 | 3.11E-27 |
| GO:0043207 | Response to external biotic stimulus | 88 | 7.66E-30 | 3.11E-27 |
| GO:0071345 | Cellular response to cytokine stimulus | 93 | 1.43E-31 | 8.12E-29 |
| GO:0009607 | Response to biotic stimulus | 91 | 8.60E-31 | 4.45E-28 |
| GO:0006952 | Defense response | 147 | 2.56E-47 | 1.46E-43 |
| GO:0002252 | Immune effector process | 99 | 4.33E-30 | 2.05E-27 |
| GO:0006955 | Immune response | 160 | 1.07E-45 | 3.03E-42 |
| GO:0002682 | Regulation of immune system process | 110 | 8.62E-26 | 3.06E-23 |
| GO:0002376 | Immune system process | 185 | 1.02E-42 | 1.93E-39 |
| GO:0009605 | Response to external stimulus | 139 | 7.34E-26 | 2.78E-23 |
| GO:0006950 | Response to stress | 208 | 1.86E-39 | 2.12E-36 |

training and test groups, respectively. GSE19738 was used in the initial receiver operating characteristic (ROC) curve analysis, ANN model predictive value of the MDD model. The GSE98793 dataset was used to test the model. The code is provided in [Supplementary material 1](#) and [Supplementary material 2](#).

Immune cell infiltration analysis and correlation analysis

CIBERSORT³ and the LM22 characteristic gene matrix were used to predict the proportion of 22 immune cells in all samples of the dataset. The CIBERSORT package was used to assess the abundance of 22 immune cells in the GSE19738 dataset. Using the median prediction index of the ANN model as the cutoff value, the samples were divided into high-and low-score groups, and the differences in immune cell infiltration among the 22 groups were analyzed.

³ <http://CIBERSORT.stanford.edu/>

TABLE 3 KEGG enrichment analyses results.

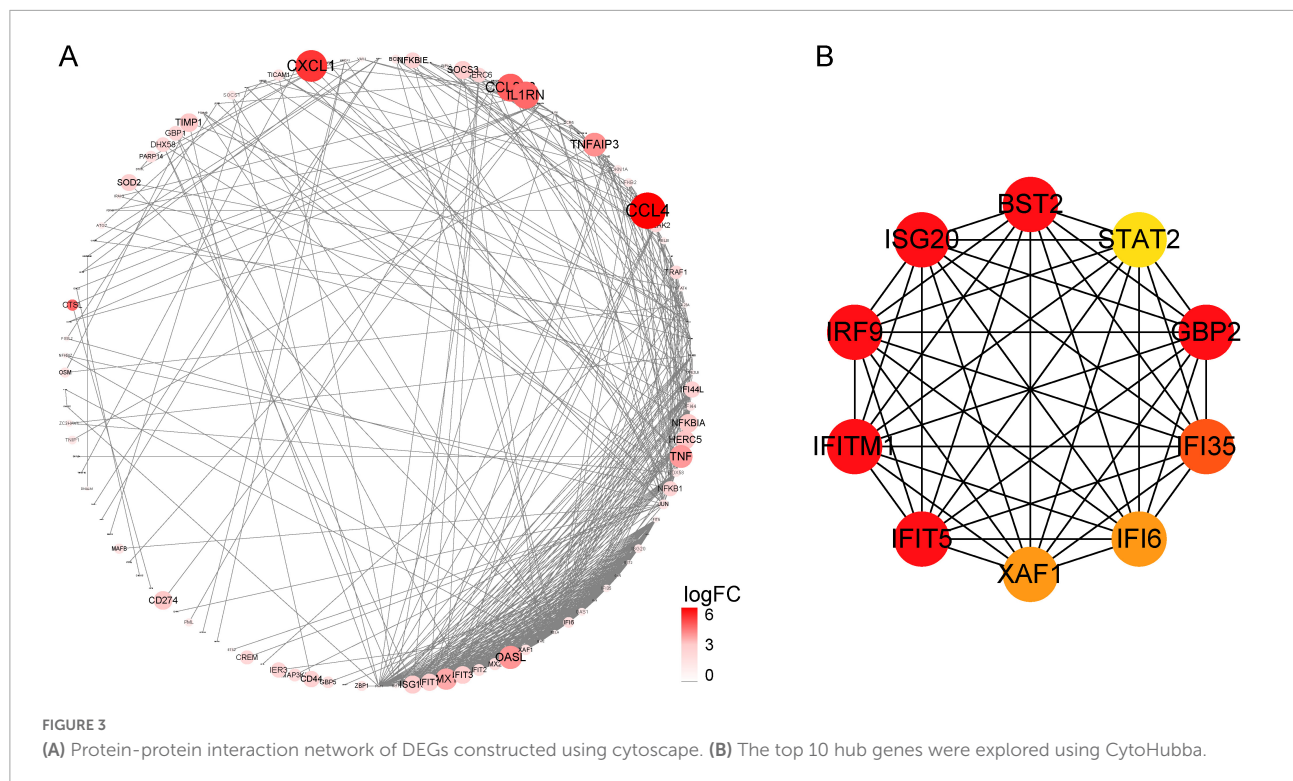
| KEGG ID | Description | Number | P value | Q value |
|---------|---|--------|----------|----------|
| ko04668 | TNF signaling pathway | 16 | 4.09E-08 | 1.39E-06 |
| ko04064 | NF-kappa B signaling pathway | 19 | 1.94E-07 | 5.11E-06 |
| ko04216 | Ferroptosis | 11 | 9.06E-09 | 3.58E-07 |
| ko04217 | Necroptosis | 24 | 8.24E-12 | 9.77E-10 |
| ko04621 | NOD-like receptor signaling pathway | 27 | 2.20E-13 | 5.22E-11 |
| ko04380 | Osteoclast differentiation | 20 | 2.17E-10 | 1.71E-08 |
| ko05164 | Influenza A | 22 | 1.40E-09 | 8.30E-08 |
| ko05169 | Epstein-Barr virus infection | 27 | 4.83E-09 | 2.29E-07 |
| ko05162 | Measles | 18 | 6.86E-08 | 2.03E-06 |
| ko05160 | Hepatitis C | 18 | 2.47E-07 | 5.85E-06 |
| ko04620 | Toll-like receptor signaling pathway | 13 | 4.66E-06 | 9.21E-05 |
| ko04625 | C-type lectin receptor signaling pathway | 13 | 4.66E-06 | 9.21E-05 |
| ko05145 | Toxoplasmosis | 13 | 7.74E-06 | 1.41E-04 |
| ko05321 | Inflammatory bowel disease (IBD) | 10 | 1.31E-05 | 2.07E-04 |
| ko05167 | Kaposi sarcoma-associated herpesvirus infection | 17 | 1.31E-05 | 2.07E-04 |
| ko05134 | Legionellosis | 9 | 1.80E-05 | 2.67E-04 |
| ko04978 | Mineral absorption | 9 | 2.74E-05 | 3.82E-04 |
| ko05161 | Hepatitis B | 15 | 3.21E-05 | 4.23E-04 |
| ko04062 | Chemokine signaling pathway | 16 | 6.01E-05 | 7.50E-04 |
| ko04920 | Adipocytokine signaling pathway | 9 | 8.36E-05 | 9.90E-04 |

Results

Identification of differentially expressed genes and enrichment analyses

A flowchart of this study is shown in [Figure 1A](#). [Figures 1B–E](#) shows data distribution before and after standardization of data sets GSE19738 and GSE98793. First, compared to healthy participants, 721 genes were differentially expressed in the peripheral blood samples of patients with MDD. Among them, 404 DEGs were upregulated, and 317 DEGs were downregulated. The corresponding volcano and heat maps are shown in [Figures 2A,B](#).

By analyzing the GO and KEGG pathway enrichment analyses, MDD peripheral blood raises the biological function of genes. The results of GO annotation revealed that the DEGs mainly comprised genes related to cytokines, defense responses to viruses, responses to biotic stimuli, immune



effector processes, and responses to external biotic stimuli (Figure 2C). Prior studies have suggested that abnormal cytokine homeostasis may be related to the pathogenesis of MDD and that cytokine profiles may be used to distinguish patients with MDD (Petrulia et al., 2020). KEGG signal pathway enrichment analysis results showed that the DEGs were mainly enriched in the immune system, cell growth and death, development, and infectious diseases (Figure 2D). MDD is associated with proinflammatory activation of the peripheral innate immune system, coupled with relative inactivation of the adaptive immune system (Leday G. et al., 2018). Several growth factors have been shown to play important roles in cell survival, growth, programmed death, and neuroplasticity and are associated with MDD (Li et al., 2021). The detailed data are presented in Tables 2, 3.

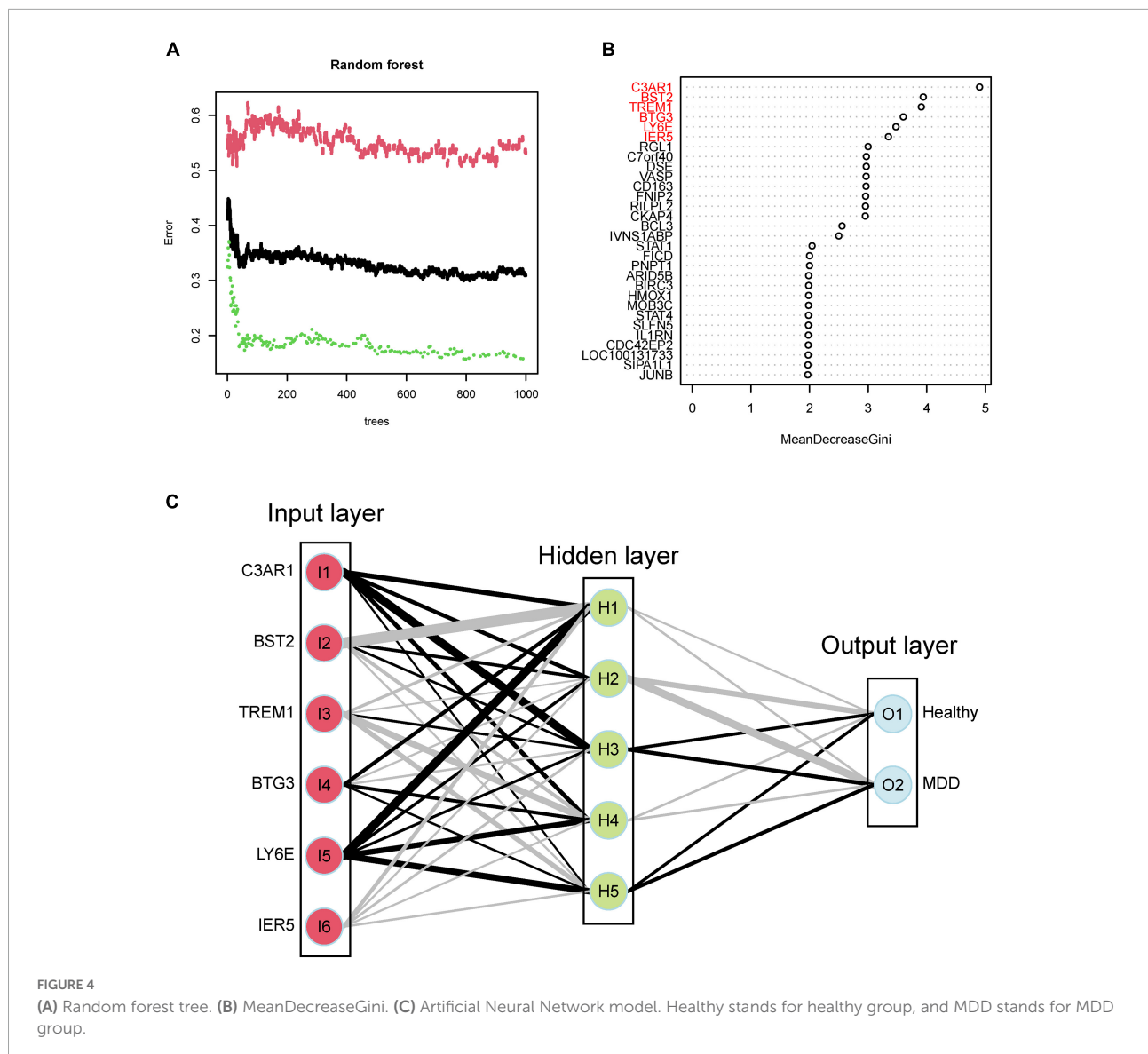
Construction of protein-protein interaction network

To further understand the relationship between DEGs at the protein level, we built a PPI network using the STRING database, which contains 487 nodes and 886 edges (Figure 3A). CytoHubba plugins were used to screen the top 10 hub genes, namely *BST2*, *STAT2*, *GBP2*, *IFI35*, *IFI6*, *XAF1*, *IFIT5*, *IFITM1*, *IRF9*, and *ISG20* (Figure 3B). Among them, dysregulation of the *GBP2* gene could indicate a relationship between cell surface receptors and intracellular effectors that can transmit

extracellular information into cells, as well as an intracellular signal transduction protein (Jiang et al., 2015). Furthermore, the proteins encoded by *IFI6* may play an important role in the regulation of apoptosis and restrict various viral infections by targeting different stages of the viral life cycle (Sajid et al., 2021).

Random forest screening of characteristic genes and major depressive disorder diagnostic model construction and validation

Using the identified DEGs, we further screened MDD-related characteristic genes using the RF algorithm. Six characteristic genes related to MDD were screened according to a gene score of >3 (Figures 4A,B). Also, we tested the correlation between these genes, and found there was no significant covariance between them (Supplementary Table 1). Furthermore, based on the features of the six MDD-related genes, we built an ANN model. The MDD diagnostic model constructed using the ANN includes the input, hidden, and output layers, as shown in Figure 4C. Among them, the dimension of the input vector is six, and the dimensions of the output vector of the control and disease. Based on the scoring values of the ANN model, ROC analysis was performed on the model to verify its accuracy. As shown in Figure 5, the AUC was 0.757 in the training set, indicating good accuracy. In addition, the accuracy of the model was further tested using the GSE98793



dataset, and the AUC of the test set was 0.685, indicating the high accuracy of the model.

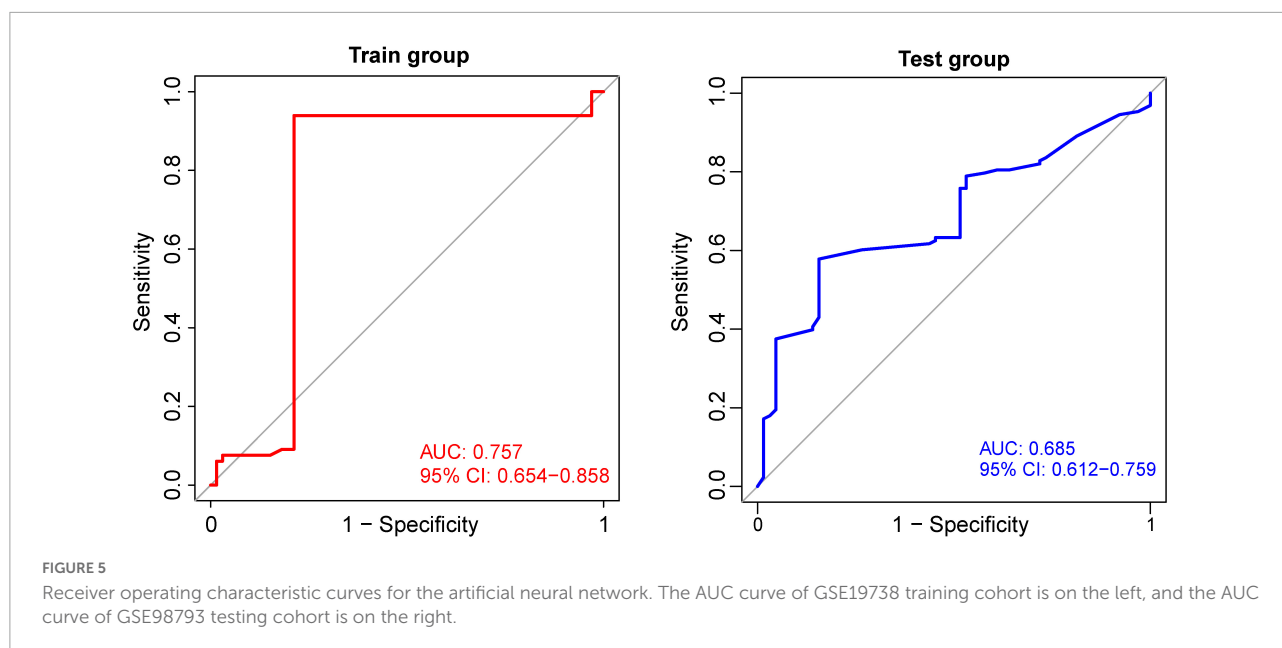
Immune infiltration and correlation analysis

To analyze the relationship between the ANN model and immune cell infiltration, the CIBERSORT algorithm was used to calculate the proportion of immune cell infiltration in the peripheral blood of the healthy and MDD groups. In the immune analysis, we drew the immune landscape of infiltrated immune cells of patients with MDD and normal volunteers (Figure 6A). We found that a variety of immune cells were significantly correlated in the disease group (Figures 6B,C). However, we did not find any significant differences in the

immune cells between the MDD and normal groups. This may be related to the sources of the specimens used in this study.

Discussion

Patients with depression have a high rate of misdiagnosis; as such, it is important to find new and feasible biomarkers to facilitate the early diagnosis and treatment of MDD. Previous studies have reported a variety of algorithms used to construct diagnostic models for patients with MDD. However, at present, there are no reports on the application of ANNs to construct auxiliary peripheral blood diagnostic models for patients with MDD. Here, we tried to find diagnostic markers related to MDD and applied ANN machine learning methods to construct



diagnostic models to explore better diagnostic models for patients with MDD that are suitable for different populations.

We identified 404 DEGs in the peripheral blood of patients with MDD compared to healthy controls. Through enrichment analysis of GO and KEGG pathways, we further identified that the biological functions of upregulated genes in the peripheral blood of patients with MDD mainly targeted cytokines, defense response to viruses, response to biotic stimulus, immune effector process, response to external biotic stimulus, response to external stimulus and immune system, cell growth and death, development, infectious diseases, and other pathways. We constructed a PPI network to screen out the top 10 hub genes, namely *BST2*, *STAT2*, *GBP2*, *IFI35*, *IFI6*, *XAF1*, *IFIT5*, *IFITM1*, *IRF9*, and *ISG20*. Moreover, we selected the following six characteristic genes in order of importance by RF algorithm: *C3AR1*, *BST2*, *TREM1*, *BTG3*, *LY6E*, and *IER5*. *C3AR1* is a critical regulator of central immune homeostasis in tau pathology whose signaling operates intracellularly in human CD4⁺ cells and participates in several T-cell functions (Arbore et al., 2016). Litvinchuk et al. (2018) showed that the expression of the C3a receptor (*C3AR1*) is positively correlated with cognitive decline and Braak staging in human Alzheimer's disease brains. *BST2* has been identified as a marker of immunomodulatory bone marrow mesenchymal stem cell cloning, as well as an effective inhibitor of enveloped virus release (James et al., 2015). *TREM-1* is an activating receptor expressed at high levels in neutrophils and monocytes that infiltrate human tissues infected with bacteria. Furthermore, it is upregulated in peritoneal neutrophils of patients with microbial sepsis and mice with experimental lipopolysaccharide-induced shock (Bouchon et al., 2001). Chronic stress contributes to the development of psychiatric disorders, including anxiety

and depression. Mouse hippocampal RNA sequences showed that stress increases the *TREM1* pathway associated with inflammation (DiSabato et al., 2021). Monocytes are in a pro-inflammatory state in patients with severe psychiatric disease. The expression of *TREM-1* is significantly increased in monocytes of patients with SCZ and BD and tends to be overexpressed in patients with major depression (Weigelt et al., 2011). Regulating the imbalance in *TREM1* expression ameliorates depression-like behaviors and impairment of learning and memory in rats (Fang et al., 2019). In addition, Rosie Owens suggested that neuroinflammatory conditions that alter the balance of *TREM1* expression may be important factors affecting microglial inflammation and homeostasis activity and may be associated with neuroinflammation and neurodegenerative disease (Owens et al., 2017). *BTG3* is a member of the anti-proliferative protein family. *IER5* may play an important role in mediating the cellular response to mitogenic signals (Williams et al., 1999). Savitz et al. found using genome-wide expression analysis of peripheral blood mononuclear cells that *IER5* was differentially expressed between 29 unmedicated depressed patients with a mood disorder (8 bipolar disorder and 21 MDD) compared to 24 healthy controls (Savitz et al., 2013). These differentially upregulated genes were closely related to MDD and mental and neurological diseases.

Based on the application of the above six characteristic genes using the ANN method, we successfully established an ANN model and further calculated the infiltration of two groups of immune cells in peripheral blood. In the training set, the AUC was 0.757, indicating good accuracy. The accuracy of the model was further tested, and the AUC of the test set was calculated to be 0.685. The reason of that the AUC is

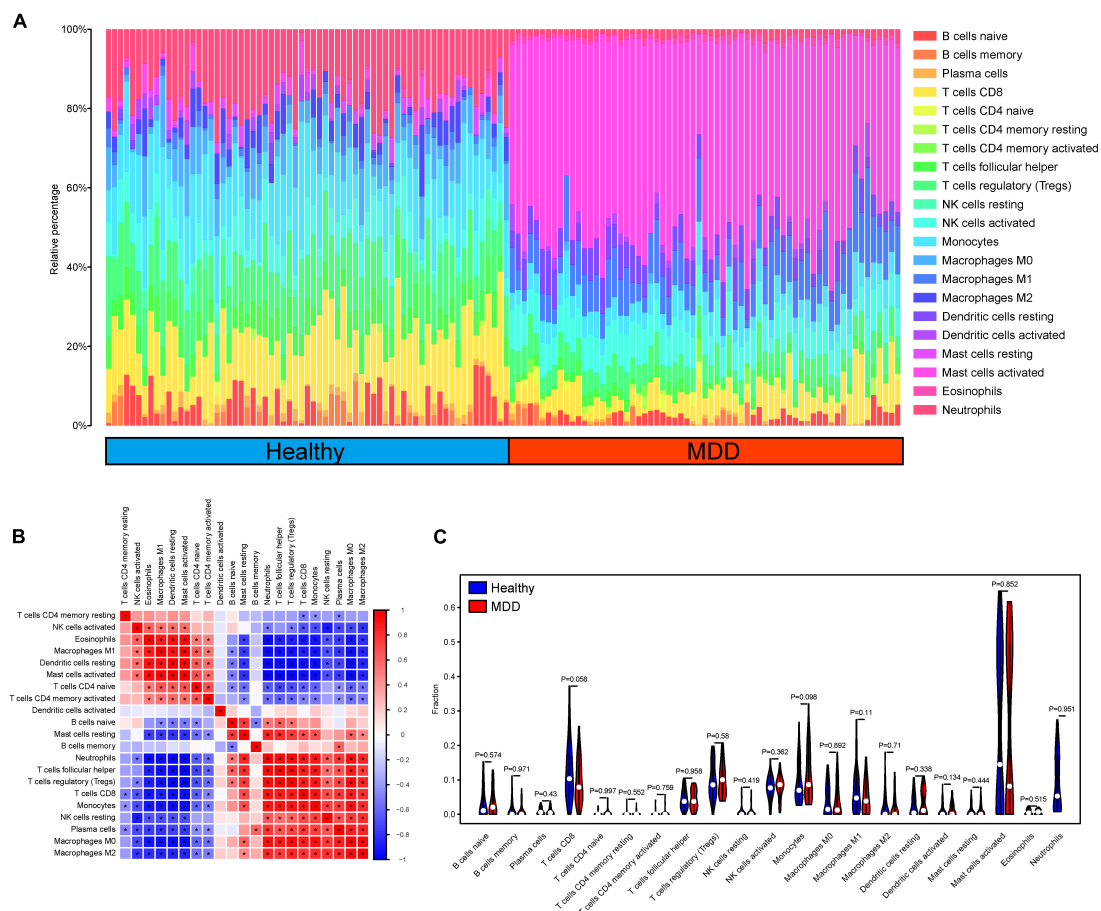


FIGURE 6

Evaluation and correlation analysis of immune cell infiltration. (A) Panoramic view of 22 immune cell infiltrates in peripheral blood samples; (B, C) High and low expression group of immune cell infiltration difference.

less than 0.7 may be due to the small sample size in the test set. At present, many studies have reported the application of RF, support vector machine, k -nearest neighbors (kNN), and naive bayesian (NB) algorithms to build diagnostic models for patients with MDD. These studies have reported high classification accuracies ranging from 70 to 100% (Yi et al., 2012; Yu et al., 2016; Bhak et al., 2019). In particular, Zhao et al. compared different machine learning approaches using the same data set which also used by us. It was found that compared with other methods, such as SVM, RF, kNN, NB, SVM could distinguish MDD from healthy controls more accurately (Zhao et al., 2021). Overall, compared with previous studies, our model evaluation provides new ideas for the application of peripheral blood in aiding diagnostic machine learning. ANNs are the most common form of neuromorphic computing, and breakthrough progress has been made in many areas. Neural networks are composed of multiple layers, each made up of a collection of cells called artificial neurons that are connected by artificial synapses (Subbulakshmi Radhakrishnan et al., 2021).

One difference in our study is that upregulated differential genes are used in the analysis of differences between patients with MDD and control groups as clinical indicators tend to focus more on increased indicators, and upregulated genes can be more effectively applied to the analysis of blood indicators in outpatient and inpatient patients. In previously reported studies, although cross-validation from the same dataset can also be used for model validation, compared with the external validation with completely independent data, the previous experiments do not reflect the universality or replicability of the model (Steyerberg and Harrell, 2016). In this study, we used previously unused datasets to calculate the AUC of the classifier.

In this study, an ANN was used for the first time to establish an auxiliary diagnosis model of MDD, which provides a new method for MDD diagnosis using machine learning and could help clinicians reduce the misdiagnosis rate of MDD. However, this study had several limitations. First, the data used in this study were obtained from public databases. They did not provide comprehensive clinical information, such as

age, sex, and BMI, which should be controlled. Meanwhile, transcriptome data are the main data in GEO database in the field of mental disorder, other types of data such as genomic and proteomic data, are lacking. Second, compared with traditional methods, such as logistic regression, the ANN method cannot obtain an image of the patient score for each variable. Third, the sample size of this study was relatively small; the more samples the machine learning model is subjected to, the more similar the sample sources are, and the more accurate the model is constructed. Fourth, due to the relatively small sample size of mental diseases in the public database, the potential transcriptome biomarkers for mental disorders was not well studied, so further comparison was not conducted between MDD and other mental disorders. In addition, the MDD samples we selected were not divided into subtypes and may have different clinical characteristics. As such, the representativeness and extensibility of the model will be limited to some extent. In future research, we will further explore the differential gene of MDD, expand the sample size, collect our own data, consider the influence of clinical features and subtypes of MDD. In order to find better diagnostic models for patients with MDD, we will compare more machine learning method such as SVM, RF, kNN, NB, etc. And continue to optimize and improve the model design to provide more reliable auxiliary tools for the diagnosis of MDD.

Conclusion

In conclusion, we adopted a popular machine learning algorithm, RFs, and ANNs to filter the characteristics of patients with MDD and construct a diagnostic model. This model was then verified in an external test set. This validation established that this model could clearly distinguish between patients with MDD and healthy controls. This model could serve as a potential adjunct tool to help psychiatrists make clinical diagnoses and treatment plans.

Data availability statement

Publicly available datasets were analyzed in this study. This data can be found from the Gene Expression Omnibus: <https://www.ncbi.nlm.nih.gov/geo/>, GSE19738 and GSE98793 datasets.

References

Arbore, G., West, E., Spolski, R., Robertson, A., Klos, A., Rheinheimer, C., et al. (2016). T helper 1 immunity requires complement-driven NLRP3 inflammasome activity in CD4+ T cells. *Science* 352:aad1210. doi: 10.1126/science.aad1210

Author contributions

SL and TL collected the data, conducted the analysis, and drafted the manuscript. QZ, BF, GL, FY, and JH conducted the data analysis and helped to interpret the results. HW and NL revised the manuscript. All authors contributed to the article and approved the submitted version.

Funding

The study was funded by National Natural Science Foundation of China (No. 82101600).

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/fnins.2022.949609/full#supplementary-material>

SUPPLEMENTARY TABLE 1

Correlation analysis between features genes.

SUPPLEMENTARY MATERIAL 1

The code of Artificial Neural Network (ANN) model.

SUPPLEMENTARY MATERIAL 2

The code of Random Forest (RF) model.

Bhak, Y., Jeong, H., Cho, Y., Jeon, S., Cho, J., Gim, J., et al. (2019). Depression and suicide risk prediction models using blood-derived multi-omics data. *Transl. Psychiatry* 9:262. doi: 10.1038/s41398-019-0595-2

- Bouchon, A., Facchetti, F., Weigand, M., and Colonna, M. (2001). TREM-1 amplifies inflammation and is a crucial mediator of septic shock. *Nature* 410, 1103–1107. doi: 10.1038/35074114
- Cioabanu, L. G., Sachdev, P. S., Trollor, J. N., Reppermund, S., Thalamuthu, A., Mather, K. A., et al. (2020). Downregulated transferrin receptor in the blood predicts recurrent MDD in the elderly cohort: a fuzzy forests approach. *J. Affect Disord.* 267, 42–48. doi: 10.1016/j.jad.2020.02.001
- DiSabato, D., Nemeth, D., Liu, X., Witcher, K., O'Neil, S., Oliver, B., et al. (2021). Interleukin-1 receptor on hippocampal neurons drives social withdrawal and cognitive deficits after chronic social stress. *Mol. Psychiatry* 26, 4770–4782. doi: 10.1038/s41380-020-0788-3
- Fang, K., Li, H., Chen, X., Gao, X., Huang, L., Du, A., et al. (2019). viaQuercetin Alleviates LPS-Induced Depression-Like Behavior in Rats Regulating BDNF-Related Imbalance of Copine 6 and TREM1/2 in the Hippocampus and PFC. *Front. Pharmacol.* 10:1544. doi: 10.3389/fphar.2019.01544
- Fernández, A., Pinto-Meza, A., Bellón, J., Roura-Poch, P., Haro, J., Autonell, J., et al. (2010). Is major depression adequately diagnosed and treated by general practitioners? Results from an epidemiological study. *Gen. Hosp. Psychiatry* 32, 201–209. doi: 10.1016/j.genhosppsy.2009.11.015
- He, R., Zhang, W., Chen, S., Liu, Y., Yang, W., and Li, J. (2020). Transcriptional profiling reveals the regulatory role of DNER in promoting pancreatic neuroendocrine neoplasms. *Front. Genet.* 11:587402. doi: 10.3389/fgene.2020.587402
- Huang, Y., Wang, Y., Wang, H., Liu, Z., Yu, X., Yan, J., et al. (2019). Prevalence of mental disorders in China: a cross-sectional epidemiological study. *Lancet Psychiatry* 6, 211–224. doi: 10.1016/s2215-0366(18)30511-x
- James, S., Fox, J., Afsari, F., Lee, J., Clough, S., Knight, C., et al. (2015). Multiparameter analysis of human bone marrow stromal cells identifies distinct immunomodulatory and differentiation-competent subtypes. *Stem Cell Rep.* 4, 1004–1015. doi: 10.1016/j.stemcr.2015.05.005
- Jiang, Y., Wu, R., Chen, C., You, Z., Luo, X., and Wang, X. (2015). Six novel rare non-synonymous mutations for migraine without aura identified by exome sequencing. *J. Neurogenet.* 29, 188–194. doi: 10.3109/01677063.2015.1122787
- Khan, M. T., Kaushik, A. C., Ji, L., Malik, S. I., Ali, S., and Wei, D. Q. (2019). Artificial neural networks for prediction of tuberculosis disease. *Front. Microbiol.* 10:395. doi: 10.3389/fmicb.2019.00395
- Leday, G., Vértés, P., Richardson, S., Greene, J., Regan, T., Khan, S., et al. (2018). Replicable and coupled changes in innate and adaptive immune gene expression in two case-control studies of blood microarrays in major depressive disorder. *Biol. Psychiatry* 83, 70–80.
- Leday, G. G. R., Vertes, P. E., Richardson, S., Greene, J. R., Regan, T., Khan, S., et al. (2018). Replicable and coupled changes in innate and adaptive immune gene expression in two case-control studies of blood microarrays in major depressive disorder. *Biol. Psychiatry* 83, 70–80. doi: 10.1016/j.biopsych.2017.01.021
- Li, Y., Jia, Y., Wang, D., Zhuang, X., Li, Y., Guo, C., et al. (2021). Programmed cell death 4 as an endogenous suppressor of BDNF translation is involved in stress-induced depression. *Mol. Psychiatry* 26, 2316–2333. doi: 10.1038/s41380-020-0692-x
- Litvinchuk, A., Wan, Y., Swartzlander, D., Chen, F., Cole, A., Propson, N., et al. (2018). Complement C3aR Inactivation attenuates tau pathology and reverses an immune network deregulated in tauopathy models and Alzheimer's Disease. *Neuron* 100, 1337–1353. doi: 10.1016/j.neuron.2018.10.031
- Murray, C., and Lopez, A. (1997). Alternative projections of mortality and disability by cause 1990–2020: Global Burden of Disease Study. *Lancet* 349, 1498–1504. doi: 10.1016/s0140-6736(96)07492-2
- Navarro, J., Fernandez Rosell, M., Castellanos, A., Del Moral, R., Lahoz-Beltra, R., and Marijuan, P. C. (2019). Plausibility of a neural network classifier-based neuroprosthesis for depression detection via laughter records. *Front. Neurosci.* 13:267. doi: 10.3389/fnins.2019.00267
- Owens, R., Grabert, K., Davies, C., Alfieri, A., Antel, J., Healy, L., et al. (2017). Divergent Neuroinflammatory Regulation of Microglial TREM Expression and Involvement of NF- κ B. *Front. Cell. Neurosci.* 11:56. doi: 10.3389/fncel.2017.00056
- Petralia, M., Mazzon, E., Fagone, P., Basile, M., Lenzo, V., Quattropiani, M., et al. (2020). The cytokine network in the pathogenesis of major depressive disorder. Close to translation? *Autoimmun. Rev.* 19:102504. doi: 10.1016/j.autrev.2020.102504
- Qi, B., Ramamurthy, J., Bennani, I., and Trakadis, Y. J. (2021). Machine learning and bioinformatic analysis of brain and blood mRNA profiles in major depressive disorder: A case-control study. *Am. J. Med. Genet. B. Neuropsychiatr. Genet.* 186, 101–112. doi: 10.1002/ajmg.b.32839
- Rhee, T. G., Wilkinson, S. T., Steffens, D. C., Rosenheck, R. A., and Olfson, M. (2020). Prevalence of treatment for depression among us adults who screen positive for depression, 2007–2016. *JAMA Psychiatry* 77, 1193–1195. doi: 10.1001/jamapsychiatry.2020.1818
- Sajid, M., Ullah, H., Yan, K., He, M., Feng, J., Shereen, M., et al. (2021). The functional and antiviral activity of interferon alpha-inducible ifi6 against hepatitis b virus replication and gene expression. *Front. Immunol.* 12:634937. doi: 10.3389/fimmu.2021.634937
- Savitz, J., Frank, M., Victor, T., Bekak, M., Marino, J., Bellgowan, P., et al. (2013). Inflammation and neurological disease-related genes are differentially expressed in depressed patients with mood disorders and correlate with morphometric and functional imaging abnormalities. *Brain Behav. Immun.* 31, 161–171. doi: 10.1016/j.bbi.2012.10.007
- Spijker, S., Van Zanten, J. S., De Jong, S., Penninx, B. W., van Dyck, R., Zitman, F. G., et al. (2010). Stimulated gene expression profiles as a blood marker of major depressive disorder. *Biol. Psychiatry* 68, 179–186. doi: 10.1016/j.biopsych.2010.03.017
- Steyerberg, E., and Harrell, F. (2016). Prediction models need appropriate internal, internal-external, and external validation. *J. Clin. Epidemiol.* 69, 245–247. doi: 10.1016/j.jclinepi.2015.04.005
- Subbulakshmi Radhakrishnan, S., Sebastian, A., Oberoi, A., Das, S., and Das, S. (2021). A biomimetic neural encoder for spiking neural network. *Nat. Commun.* 12:2143. doi: 10.1038/s41467-021-22332-8
- Swietlik, D., and Bialowas, J. (2019). Application of artificial neural networks to identify Alzheimer's Disease using cerebral perfusion SPECT data. *Int. J. Environ. Res. Public Health* 16:1303. doi: 10.3390/ijerph16071303
- Weigelt, K., Carvalho, L., Drexhage, R., Wijkhuijs, A., de Wit, H., van Beveren, N., et al. (2011). TREM-1 and DAP12 expression in monocytes of patients with severe psychiatric disorders. EGR3, ATF3 and PU.1 as important transcription factors. *Brain Behav. Immun.* 25, 1162–1169. doi: 10.1016/j.bbi.2011.03.006
- Williams, M., Lyu, M., Yang, Y., Lin, E., Dunbrack, R., Birren, B., et al. (1999). Ier5, a novel member of the slow-kinetics immediate-early genes. *Genomics* 55, 327–334. doi: 10.1006/geno.1998.5679
- Woo, H., Lim, S., Myung, W., Kim, D., and Lee, S. (2018). Differentially expressed genes related to major depressive disorder and antidepressant response: genome-wide gene expression analysis. *Exp. Mol. Med.* 50, 1–11. doi: 10.1038/s12276-018-0123-0
- Yi, Z., Li, Z., Yu, S., Yuan, C., Hong, W., Wang, Z., et al. (2012). Blood-based gene expression profiles models for classification of subsyndromal symptomatic depression and major depressive disorder. *PLoS One* 7:e31283. doi: 10.1371/journal.pone.0031283
- Yu, J., Xue, A., Redei, E., and Bagheri, N. (2016). A support vector machine model provides an accurate transcript-level-based diagnostic for major depressive disorder. *Transl. Psychiatry* 6:e931. doi: 10.1038/tp.2016.198
- Zhao, S., Bao, Z., Zhao, X., Xu, M., Li, M. D., and Yang, Z. (2021). Identification of diagnostic markers for major depressive disorder using machine learning methods. *Front. Neurosci.* 15:645998. doi: 10.3389/fnins.2021.645998
- Zhdanova, M., Pilon, D., Ghelerter, I., Chow, W., Joshi, K., Lefebvre, P., et al. (2021). The prevalence and national burden of treatment-resistant depression and major depressive disorder in the United States. *J. Clin. Psychiatry* 82:20m13699. doi: 10.4088/JCP.20m13699



OPEN ACCESS

EDITED BY

Thayne Kowalski,
Centro Universitário Cesuca, Brazil

REVIEWED BY

Nancy Lucero Martínez Rodríguez,
Federico Gómez Children's Hospital,
Mexico
Wei Kong,
Shanghai Maritime University, China

*CORRESPONDENCE

Yun Wu,
wuyun7770@163.com
Nianping Feng,
fnp2580@126.com

SPECIALTY SECTION

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

RECEIVED 14 June 2022

ACCEPTED 22 July 2022

PUBLISHED 22 August 2022

CITATION

He Y, Cong L, He Q, Feng N and Wu Y
(2022), Development and validation of
immune-based biomarkers and deep
learning models for Alzheimer's disease.
Front. Genet. 13:968598.
doi: 10.3389/fgene.2022.968598

COPYRIGHT

© 2022 He, Cong, He, Feng and Wu.
This is an open-access article
distributed under the terms of the
[Creative Commons Attribution License](#)
(CC BY). The use, distribution or
reproduction in other forums is
permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original
publication in this journal is cited, in
accordance with accepted academic
practice. No use, distribution or
reproduction is permitted which does
not comply with these terms.

Development and validation of immune-based biomarkers and deep learning models for Alzheimer's disease

Yijie He, Lin Cong, Qinfei He, Nianping Feng* and Yun Wu*

Department of Neurology, The Second Affiliated Hospital of Harbin Medical University, Harbin, China

Background: Alzheimer's disease (AD) is the most common form of dementia in old age and poses a severe threat to the health and life of the elderly. However, traditional diagnostic methods and the ATN diagnostic framework have limitations in clinical practice. Developing novel biomarkers and diagnostic models is necessary to complement existing diagnostic procedures.

Methods: The AD expression profile dataset GSE63060 was downloaded from the NCBI GEO public database for preprocessing. AD-related differentially expressed genes were screened using a weighted co-expression network and differential expression analysis, and functional enrichment analysis was performed. Subsequently, we screened hub genes by random forest, analyzed the correlation between hub genes and immune cells using ssGSEA, and finally built an AD diagnostic model using an artificial neural network and validated it.

Results: Based on the random forest algorithm, we screened a total of seven hub genes from AD-related DEGs, based on which we confirmed that hub genes play an essential role in the immune microenvironment and successfully established a novel diagnostic model for AD using artificial neural networks, and validated its effectiveness in the publicly available datasets GSE63060 and GSE97760.

Conclusion: Our study establishes a reliable model for screening and diagnosing AD that provides a theoretical basis for adding diagnostic biomarkers for the AD gene.

KEYWORDS

Alzheimer's disease, immune infiltration, biomarkers, random forest, artificial neural networks

Introduction

Relevant studies have shown that in 2018 the prevalence of dementia is about 50 million people worldwide and is expected to triple by 2050 (Scheltens et al., 2021). Alzheimer's disease (AD), the most common form of dementia, is a significant threat to the health and lives of older adults, with initial symptoms of memory loss, decreased

verbal skills, and impaired logical thinking (Sabayan and Sorond, 2017). The onset of AD is insidious, and some pathophysiological changes are thought to occur years or even decades before the clinical diagnosis of dementia (Morris, 2005). It was not until 2011 that the concept of the preclinical stage of AD was explicitly introduced in the NIA-AA diagnostic criteria for Alzheimer's disease (Sperling et al., 2011). The introduction of this concept is critical and suggests that interventions can be made in the preclinical stage of AD to ultimately delay the disease's progression.

The latest NIA-AA AD diagnostic framework-ATN framework (Jack et al., 2018), officially published in 2018, is considered promising for the early identification of disease development in the preclinical phase of AD. In this framework, the diagnosis of AD is determined by the biomarkers A β and tau. Still, the framework is currently only used for scientific research and is not widely used in the clinic. Therefore, how identifying and diagnosing early becomes an urgent problem for us.

Genetic factors are considered a significant risk for Alzheimer's disease, accounting for 60%–80% of the disease (Gatz et al., 2006). In addition to the well-known APOE4 risk alleles, there are many genes involved in AD that we do not recognize (Jansen et al., 2019). Second-generation sequencing technology has revealed the potential of some of these genes in the development of AD, such as SORL1 (Holstege et al., 2017), ABCA7 (Bossaerts et al., 2021), TREM2, and R47H (Cheng-Hathaway et al., 2018; Sudom et al., 2018). With the development of science and technology, bioinformatics analysis has been widely used in diseases. Weighted co-expression networks (WGCNA) have become the most prevalent gene screening tool. They have been validated in numerous conditions by constructing free-scale gene co-expression networks to explore the association between clinical features and genes with co-expression patterns. In addition, some machine learning algorithms have been gradually introduced into medical research. Random Forest (RF) algorithms have been applied in acute myeloid leukemia (Shi and Xu, 2019), ALS (Hothorn and Jung, 2014), and cardiovascular diseases (Yang et al., 2020). Artificial neural networks have also demonstrated their powerful functions in medical research applications. Some scholars have validated diagnostic models for ulcerative colitis and heart failure (Li et al., 2020; Tian et al., 2020). Using WGCNA combined with machine learning to analyze AD biological data to find AD susceptibility genes may be a breakthrough.

In this study, we identified datasets of AD serological sources in the GEO database, used WGCNA with differential expression analysis to screen out differentially expressed genes (DEGs) between AD and normal control samples from them, and applied the random forest algorithm to screen out hub genes, constructed an artificial neural network AD diagnostic model, and further analyzed the role of hub genes in the immune microenvironment to provide early identification and intervention of AD and a better understanding of the

molecular immune mechanisms provide new perspectives. The technical route is shown in Figure 1.

Materials and methods

Download and processing of expression spectrum data

In this study, the AD gene expression profile datasets GSE63060, GSE63061, and GSE97760 were downloaded from Gene Expression Omnibus (Table 1). The annotation information of the microarray probes was obtained through the soft annotation list of the corresponding platform. Multiple probes corresponding to the same gene symbol may be encountered during the annotation of probe data. We use the average probe expression as the gene expression level. The process is annotated through a Perl language script (<https://www.perl.org/>).

Principal component analysis

To analyze the distribution of samples in the expression profiling dataset GSE63060 and the differences in gene expression between samples. We first performed a z-score on the expression spectrum, further performed dimensionality reduction analysis using the prcomp function to obtain the reduced matrix, and visualized the results. The results of PCA (Figure 2) show that there is little variability between the AD and CTL groups, so constructing a diagnostic model is necessary. We have invoked the R package stats (version 3.6.0) for the above procedure.

Differential expression analysis

We performed differential expression analysis on the expression profiling dataset GSE63060 for differential expression genes between the AD and CTL groups with the R package limma (version 3.40.6) (Ritchie et al., 2015). The significance criteria for DEGs were set to FoldChang >1.2 and adjusted *p*-value < 0.05. Heatmaps of DEGs were implemented by the pheatmap package (version 1.0.12), and differentially expressed genes were represented by volcano maps constructed by the ggplot2 package (version 3.3.5).

Construction of weighted gene co-expression network analysis

First, the obtained expression profile matrix was read in. The MAD value, also known as median absolute deviation,

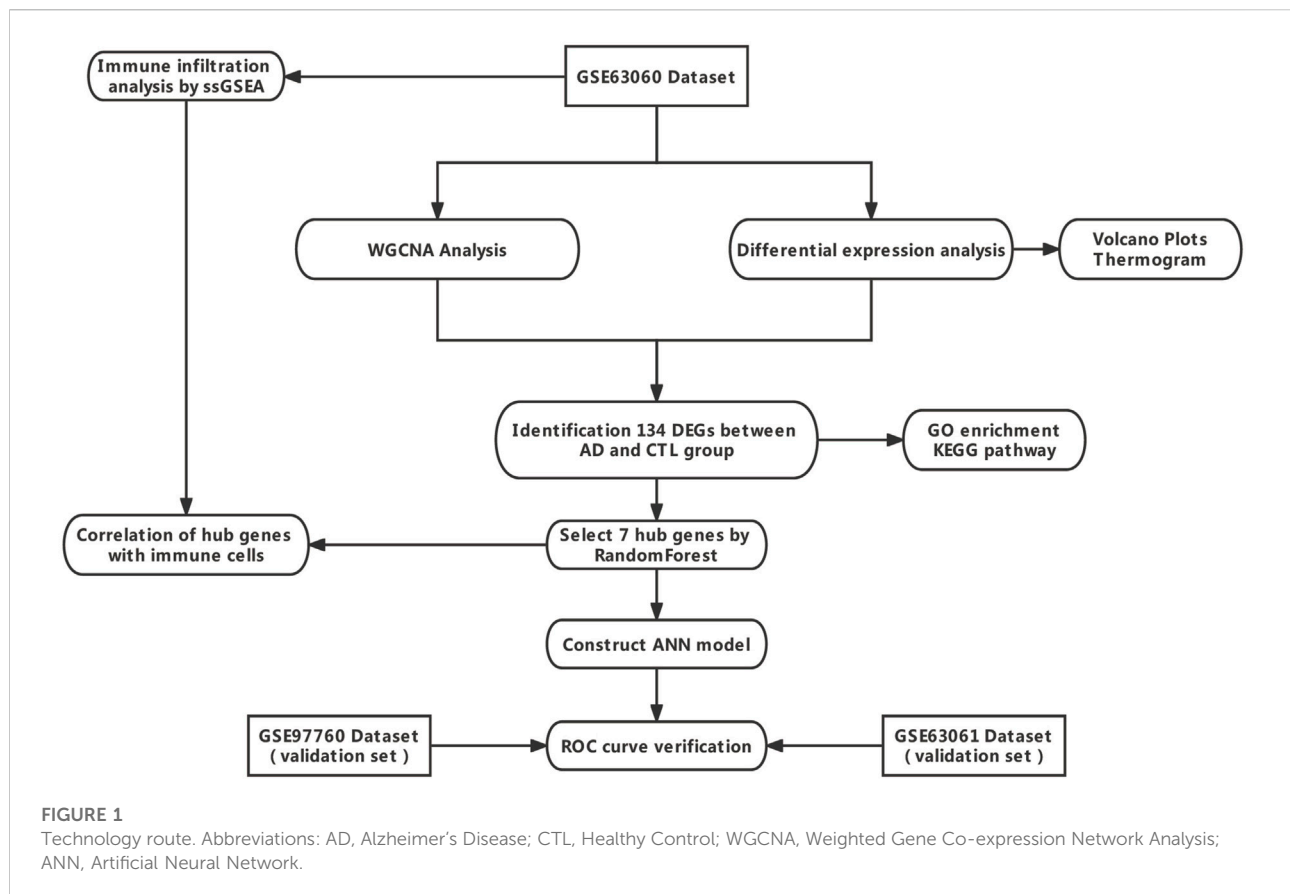


TABLE 1 Dataset information from the GEO database.

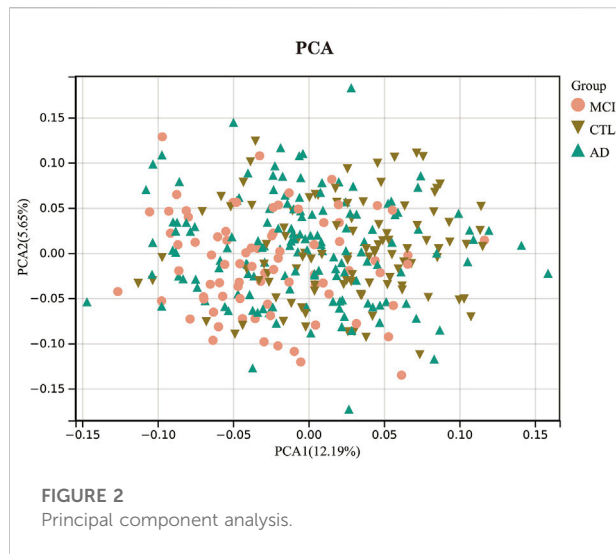
| Location | Dataset ID | Platform | Type | Number |
|----------|------------|----------|------------|------------------------|
| Blood | GSE63060 | GPL6947 | Microarray | 104 control vs. 145 AD |
| Blood | GSE63061 | GPL10558 | Microarray | 134 control vs. 139 AD |
| Blood | GSE97760 | GPL16699 | Microarray | 10 control vs. 9 AD |

was calculated separately for each gene. The first 50% of genes with the smallest MAD values were eliminated. The goodSampleGenes function of the R package WGCNA was used to eliminate outlier genes and samples, on which the scale-free co-expression network was further constructed. The genes were then hierarchically clustered to identify modules. Pearson's correlation analysis determined correlations between clinical phenotypes and the resulting modules. Among all the obtained modules, we selected the most correlated modules with the normal control group (CTL group) and AD group for further analysis. The genes in the key modules were those that met the

following criteria: gene significance (GS) > 0.1 and module membership (MM) > 0.8.

Overlapping weighted co-expression networks-related module genes with differential expression genes

The 200 Differential Expression genes obtained from the differential expression analysis overlapped with the WGCNA correlation module genes. Venn (Bardou et al., 2014) diagrams were used to visualize the results.

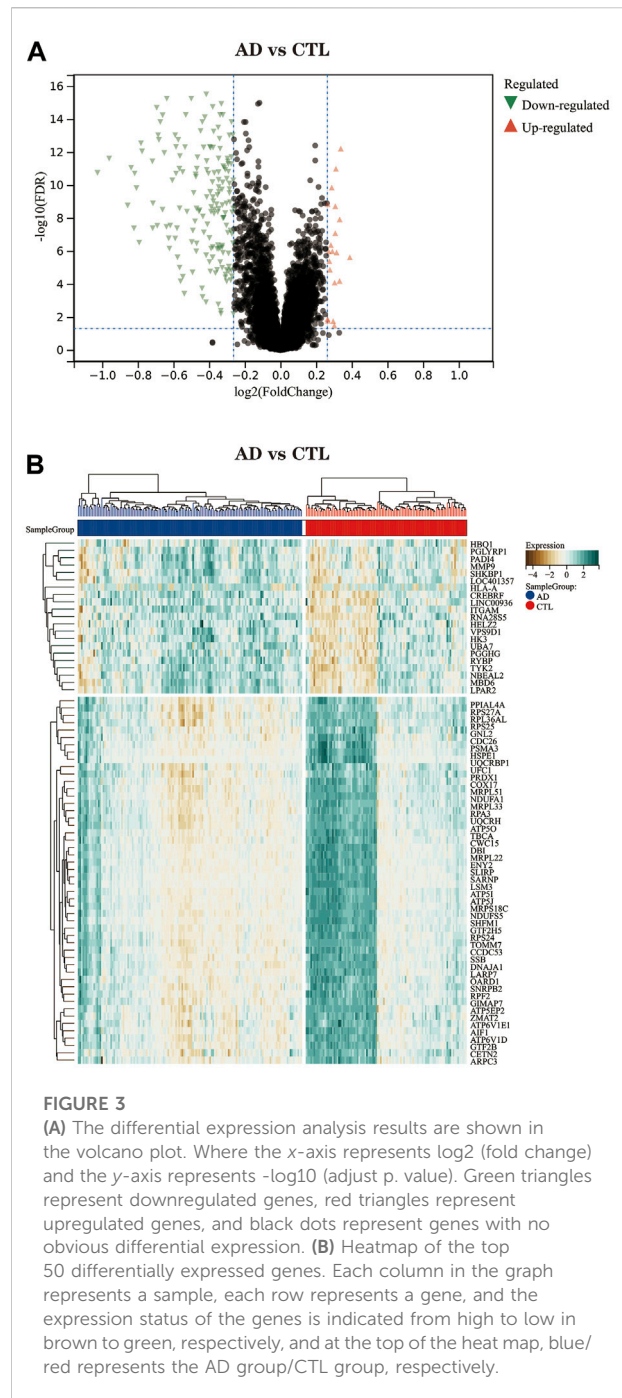


Screening for hub genes using random forest

Random forest models for differentially expressed genes were constructed using the randomForest package (version 4.6–14). First, the number of decision trees needed to achieve the highest model accuracy in cross-validation was calculated based on the expression matrix of differentially expressed genes. Second, the random forest model was constructed. The importance value scores of dimensions were obtained from the random forest model using the Gini coefficient method. Genes with an importance value greater than two were identified as hub genes for subsequent analysis. The hub genes were clustered, and heatmaps were drawn with the pheatmap package (version 1.0.12).

Correlation analysis between hub genes and immune characteristics

To determine the role of hub genes in the immune microenvironment, we analyzed the correlation between them and immune cell infiltration by applying the ssGSEA approach to analyze the proportion of 28 different immune cell distributions and infiltration scores in each sample of the dataset GSE63060. Pheatmap package (version 1.0.12) was used to map the immune cell distribution maps. Further, the voplot package (version 0.3.7) was used to present the differences in immune cell infiltration scores between the CTL and AD groups. Finally, we used the Spearman correlation test to assess the correlation between hub genes and immune cells and visualized the results using the ggplot2 package (version 3.3.5).



Construction and validation of artificial neural network models

The hub genes expression matrix obtained was extracted, and the data were first scaled by Min-Max processing. An artificial neural network model was constructed using the R package Neuralnet (version 1.44.2). The parameters were set

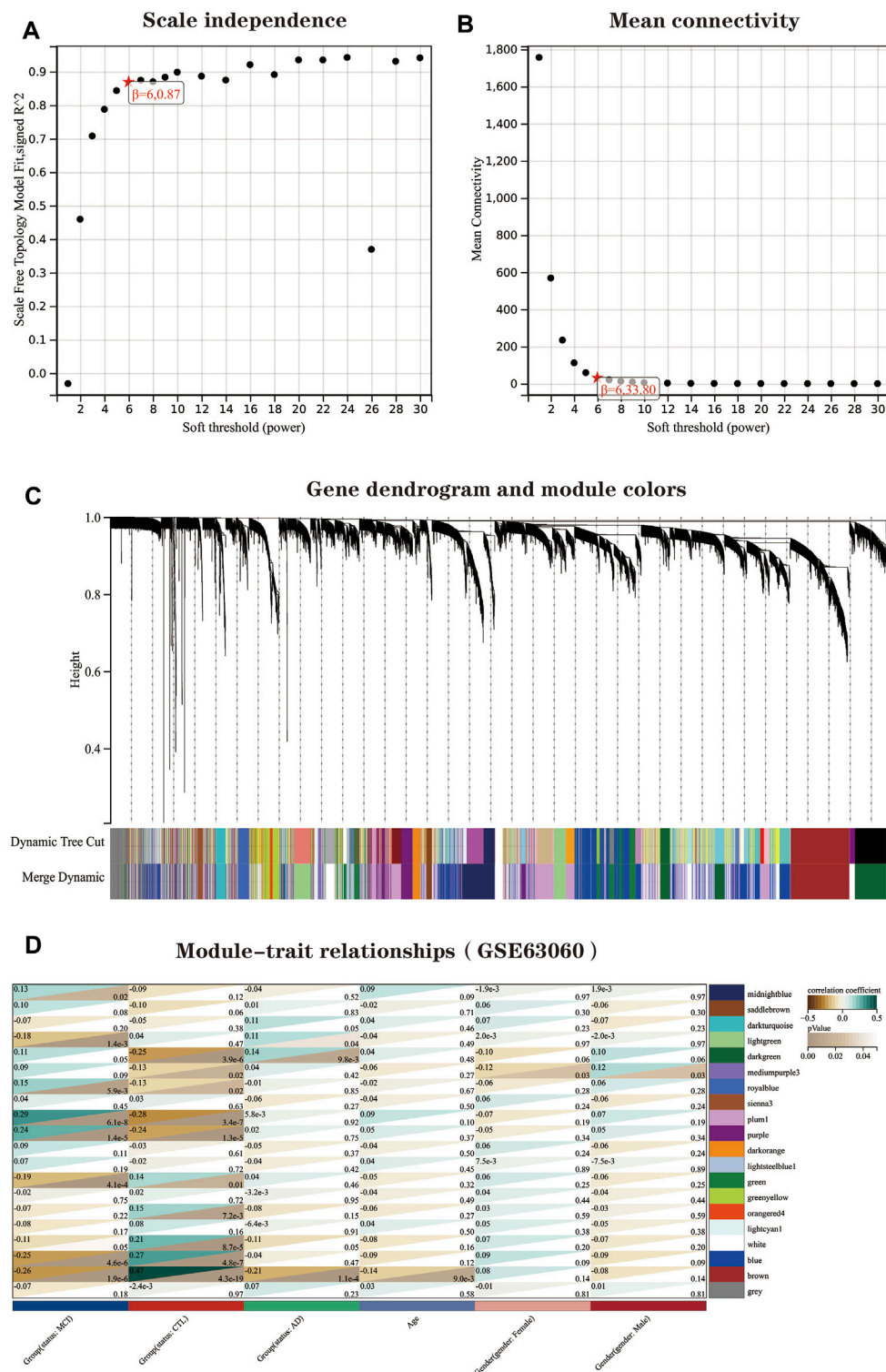


FIGURE 4

WGCNA of AD dataset GSE63060. (A) Scale-free index for analyzing the power of various soft thresholds. The horizontal coordinate represents the power of soft thresholds, and the best soft threshold is marked with an asterisk. (B) Average connectivity of various soft thresholds. (C) Identification of co-expressed gene modules. A dendrogram of all differentially expressed genes was clustered based on a measure of gene similarity. Cut lines of modules were identified, and a different color indicated each module. (D) Heat map of the correlation between modules and clinical phenotypes. Each row represents a module; each column represents a clinical trait. Each cell indicates the correlation between the (Continued)

FIGURE 4

module and the clinical phenotype. The corresponding cor value and *p*-value are labeled therein. The brown and dark green modules have the strongest correlation with AD, and the brown module and plum1 module have the strongest correlation with CTL. (E) Correlation between module membership (MM) and gene significance (GS) in the AD correlation module. *r* denotes the absolute correlation coefficient between GS and MM. (F) Correlation between module membership (MM) and gene significance (GS) in the CTL correlation module. *r* denotes the absolute correlation coefficient between GS and MM. Abbreviations: WGCNA (Weighted Gene Co-expression Network Analysis), AD (Alzheimer's Disease), CTL (Healthy Control).

to five hidden layers, the neural network algorithm obtained the gene weight information, and the disease classification score Neural AD was obtained using “Gene Expression” × “Gene_Weight.” The model was visualized using the software package NeuralNetTools (version 1.5.2). Finally, the classification performance of the artificial neural network model was evaluated using the pROC package (version 1.18.0) and the ggplot2 package (version 3.3.5) to calculate the AUC scores and plot the ROC curves. Two independent datasets, GSE63060 and GSE97760, were used to validate the accuracy of the artificial neural network model, while ROC curves were plotted and the area under the curve AUC was calculated.

**FIGURE 5**

A total of 134 DEGs were screened for further analysis. Red represents 335 genes in the module with the strongest correlation to the CTL group, blue represents 284 genes in the module with the strongest correlation to the AD group, and the green represents 200 genes obtained from differential expression analysis. Abbreviations: AD (Alzheimer's Disease), CTL (Healthy Control), DEGs (Differentially Expressed Genes).

Results

Differential expression analysis

The volcano plot (Figure 3A) shows that after differential expression analysis, we screened 200 differential expression genes (Supplementary Material S1), including 179 downregulated genes and 21 upregulated genes. The heat map (Figure 3B) shows the expression status of the top 50 differential expression genes.

Construction of weighted gene co-expression network analysis and identification of core modules

Before performing the analysis, we first processed all pairs of genes using the Pearson correlation matrix and the average linkage method. Then, a weighted adjacency matrix is constructed, which is built by a power function, and we usually use the formula.

$$A_{mn} = |C_{mn}|^{\beta}$$

Where C_{mn} is the Pearson correlation coefficient between gene *m* and gene *n*, and A_{mn} is the adjacency relationship between gene *m* and gene *n*. An important parameter in the construction of the weighted adjacency matrix is the soft threshold parameter β , which effectively emphasizes correlations between genes and, at the same time, penalizes weak correlations between genes. This study determines the soft threshold parameter as 6 (Figures 4A,B). Immediately after, we transformed this adjacency into a TOM matrix (topological overlap matrix), which better reflects the connectivity and adjacency between genes, and 1-TOM was defined as the difference between genes. To group genes with similar expression characteristics into the same module, we clustered genes in an average linkage hierarchy based on the dissimilarity measure of the TOM matrix. We set the minimum number of genes in the gene dendrogram to 30. We chose a cut line for the module dendrogram, calculated the similarity of module feature genes, and merged some similar modules to better delineate the modules (Figure 4C). After a series of calculations, we finally obtained 20 co-expression modules and visualized the correlation between modules and clinical

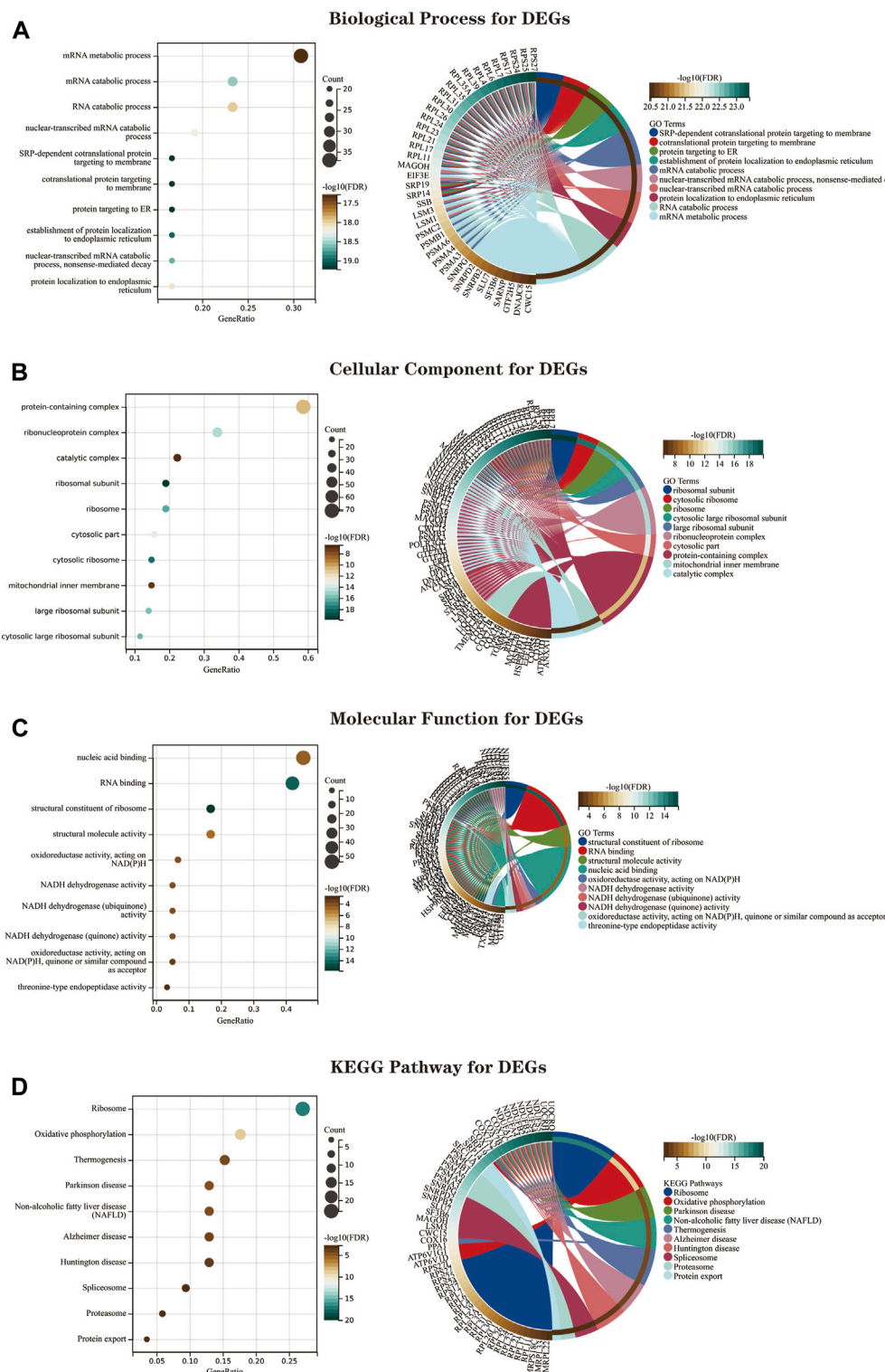


FIGURE 6

The GO and KEGG enrichment analysis results of 134 DEGs are shown as bubble and circle plots. (A) Shows the top 10 significantly enriched BP (biological process). (B) Shows the top 10 enriched CC (cellular component) considerably. (C) The top 10 enriched MF (molecular function) considerably. (D) The top 10 enriched KEGG pathways. In the bubble plot of GO and KEGG enrichment analysis, the x-axis represents the GeneRatio, the y-axis represents the $-\log_{10}(\text{FDR})$ value, the bubble size represents the number of genes, and the color shades represent the size of the FDR value. The linkage between the left and right sides indicates the correlation between DEGs and terms. Abbreviations: DEGs (Differential Expressed Genes), GO (Gene Ontology), KEGG (Kyoto Encyclopedia of Genes and Genomes).

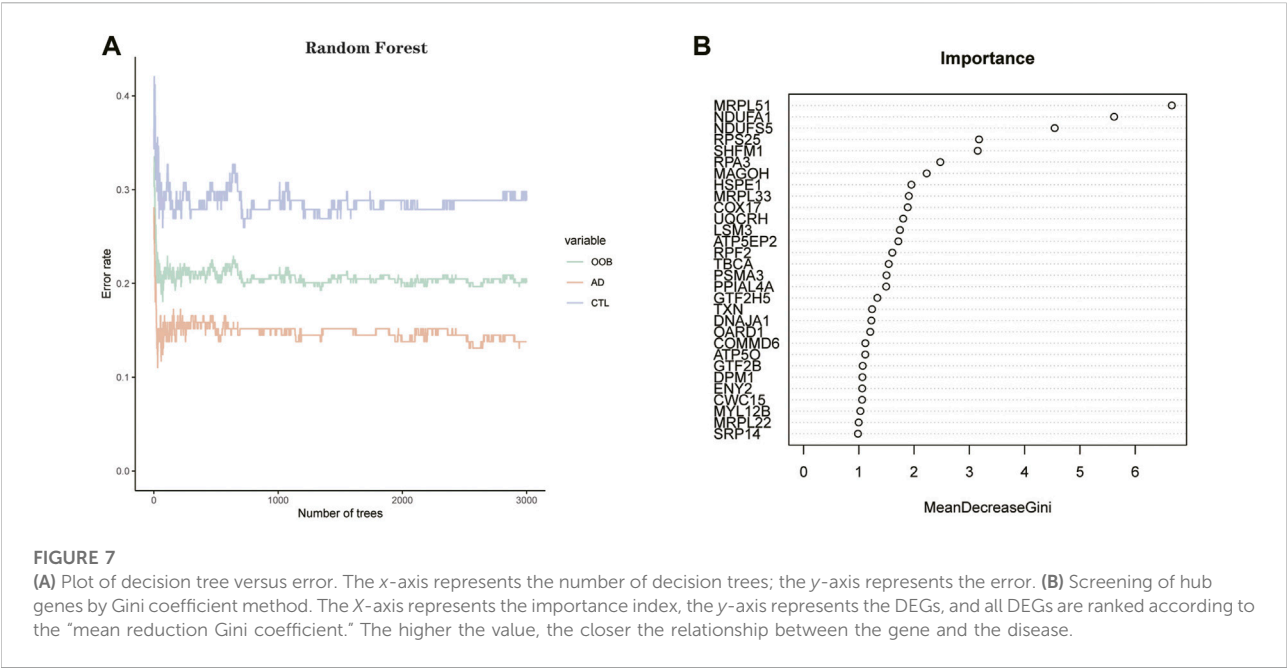


FIGURE 7 (A) Plot of decision tree versus error. The x-axis represents the number of decision trees; the y-axis represents the error. (B) Screening of hub genes by Gini coefficient method. The X-axis represents the importance index, the y-axis represents the DEGs, and all DEGs are ranked according to the "mean reduction Gini coefficient." The higher the value, the closer the relationship between the gene and the disease.

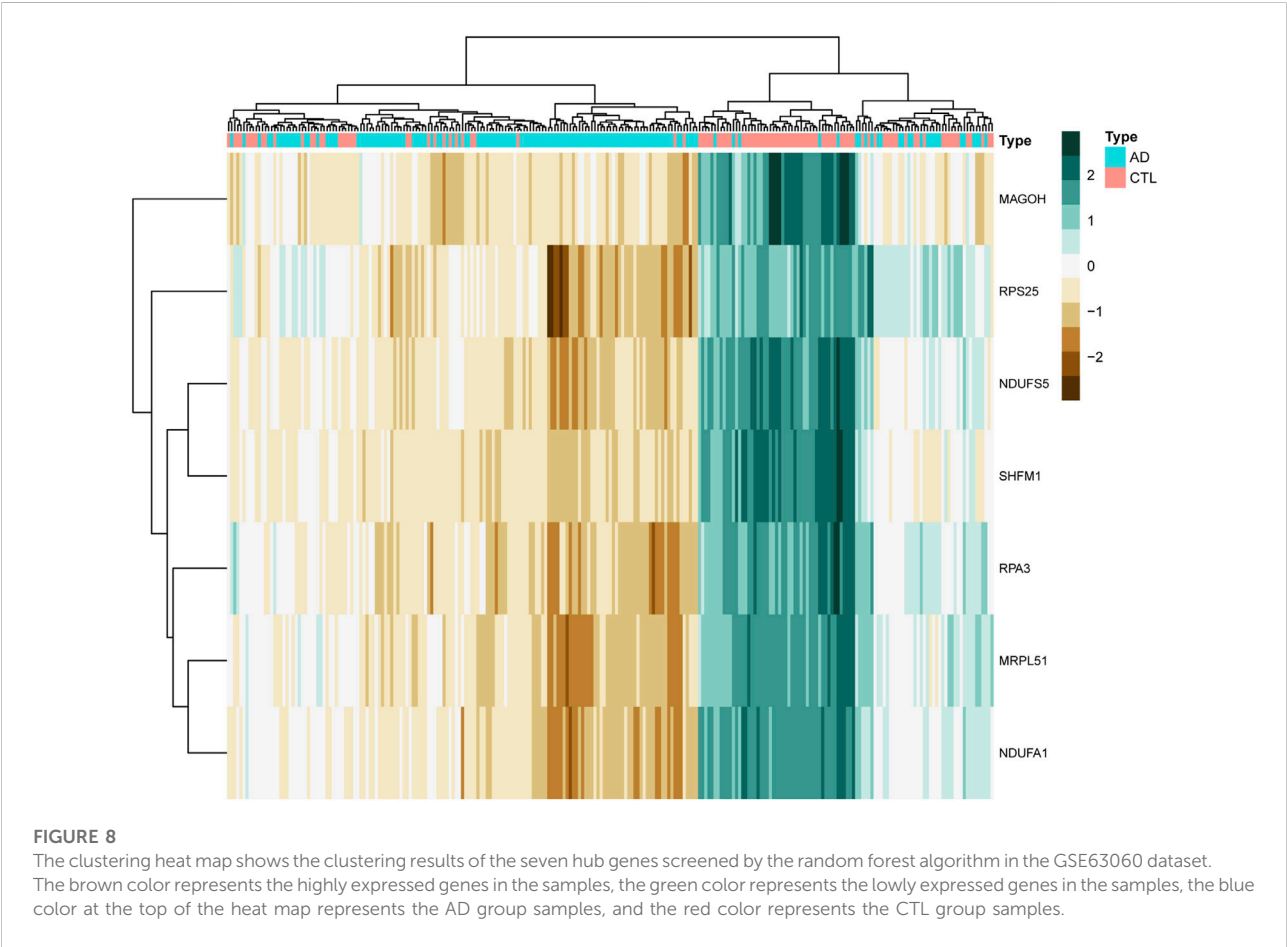


FIGURE 8 The clustering heat map shows the clustering results of the seven hub genes screened by the random forest algorithm in the GSE63060 dataset. The brown color represents the highly expressed genes in the samples, the green color represents the lowly expressed genes in the samples, the blue color at the top of the heat map represents the AD group samples, and the red color represents the CTL group samples.

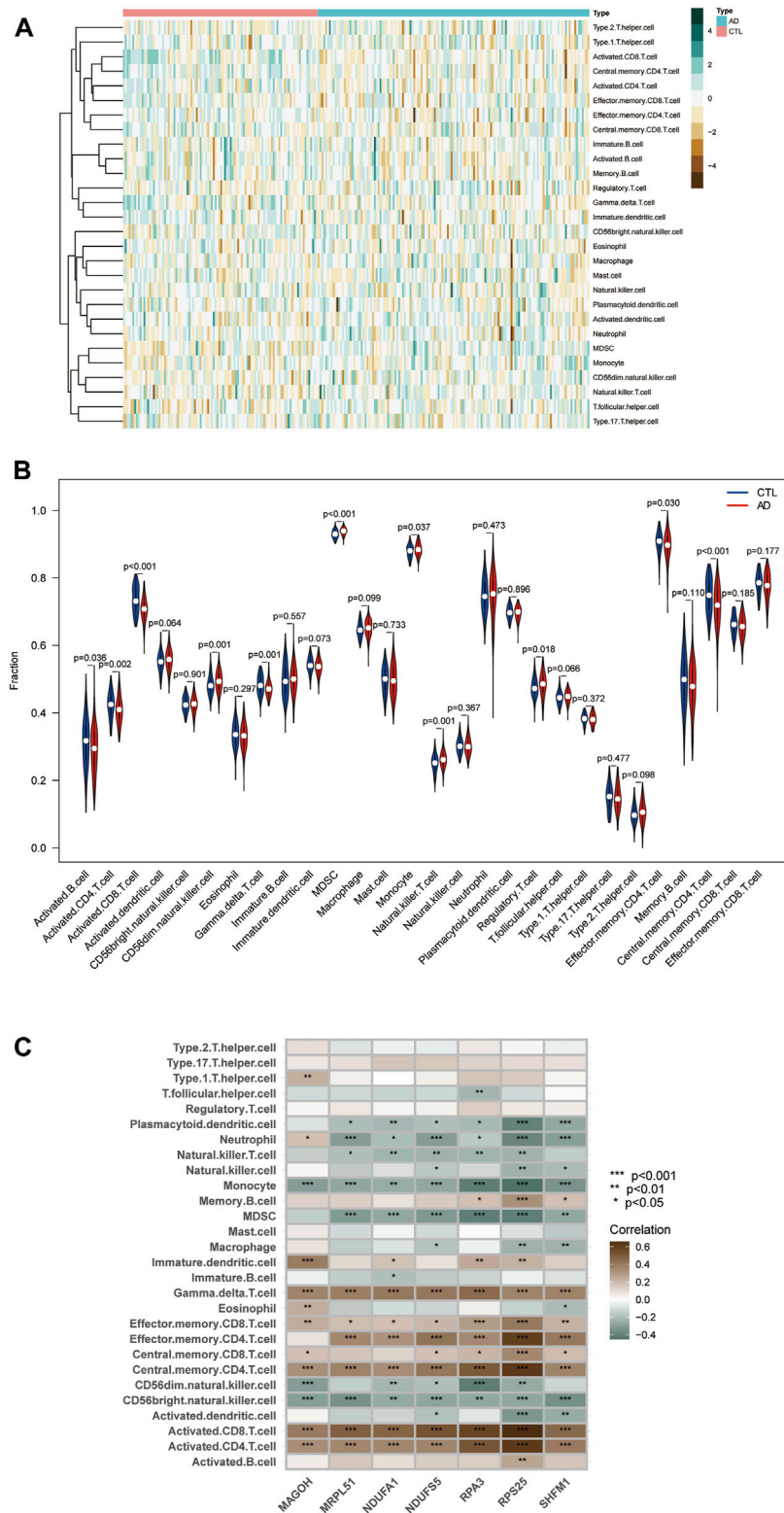


FIGURE 9 Immune infiltration landscape between AD and CTL obtained by ssGSEA analysis. **(A)** Heat map summarizing the scores of immune cell infiltration between AD patients and non-AD patients. **(B)** Violin plot showing the difference in immune cell infiltration between AD (red) and CTL (blue), $p < 0.05$, was considered statistically significant. **(C)** Shows the correlation between hub genes and immune cells. The colors from brown to green represent the change from positive to negative correlations, respectively. More asterisks and darker colors of the modules represent stronger correlations.

features using the form of a heat map (Figure 4D). Notably, the grey module was considered a set of genes that could not be assigned to any module. From the correlation heat map of clinical phenotypes and modules, we could learn that the brown module negatively correlated with age and AD groups. In contrast, the genes in the mediumpurple3 module were differentially expressed between genders. The correlation between modules and clinical features was used to estimate the association of modules with features. Two methods were used to identify the key modules of the network. In the first method, Pearson correlation coefficients were calculated between the ME of each module and each clinical trait, allowing the identification of modules significantly associated with traits ($p < 0.05$). In the second approach, the Pearson correlation coefficient [gene significance (GS)] between the expression level of each gene and each clinical trait was calculated; then, the mean absolute value GS of all genes in the module was calculated. The larger the mean total value, the stronger the correlation between the module and the clinical trait.

Also, we plotted the scatter plot of GS and MM correlations for each module (Figure 4E). Combining Figure 4D as well as Figures 4E,F, we can see that the brown module ($\text{cor} = -0.21$, $p = 1.1 \times 10^{-4}$) and the dark green module ($\text{cor} = 0.14$, $p = 9.8 \times 10^{-3}$) had the highest correlations with the AD group, and a total of 284 genes were extracted from these two modules (Supplementary Material S2). In contrast, the brown module ($\text{cor} = 0.47$, $p = 4.3 \times 10^{-19}$) and the plum1 module ($\text{cor} = -0.28$, $p = 3.4 \times 10^{-7}$) had the highest correlation with the CTL group, with a total of 335 genes extracted in the same way (Supplementary Material S3). The effect of clinical phenotype on module genes was also considered in the extraction of genes.

Overlapping weighted co-expression networks-related module genes with differential expression genes

We overlapped the genes derived from the genes obtained in the WGCNA analysis (335 genes in the CTL group-related module and 284 genes in the AD group-related module) and the 200 differentially expressed genes obtained in the differential expression analysis, and a total of 134 DEGs were screened for further analysis. The results were visualized by the Venn diagram (Figure 5) and recorded in the table (Supplementary Material S4).

Kyoto encyclopedia of genes and genomes and gene ontology enrichment analysis for overlapping differential genes

We performed GO and KEGG enrichment analysis on the screened 134 DEGs. The results of our GO enrichment analysis included BP (Figure 6A), CC (Figure 6B), and MF (Figure 6C).

GO-BP was mainly enriched in RNA metabolic pathways such as mRNA metabolic process, mRNA catabolic process, and RNA catabolic process; GO-MF enrichment results showed DEGs were primarily associated with nucleic acid binding, RNA binding, structural constituent of ribosome, structural molecule activity, oxidoreductase activity, acting on NAD (P) H, and GO-CC analysis showed that these genes were significantly enriched in ribosomal structures such as protein-containing complex, ribonucleoprotein complex, catalytic complex, etc. The results of GO enrichment (Supplementary Material S5) indicate that DEGs play a role in ribosome function and mitochondrial function.

The KEGG pathway shows (Figure 6D; Supplementary Material S6) that DEGs are mainly involved in the “ribosome,” “oxidative phosphorylation,” “thermogenesis,” “Parkinson’s disease,” “NAFLD,” and “Alzheimer’s disease” pathways.

Random forest screening for Alzheimer’s disease hub genes

We imported the expression profile files of 134 differentially expressed genes into the random forest model. Before calculation, firstly, we set the random seed to 123,456 and calculate the number of decision trees needed to achieve the highest accuracy of the model in cross-validation. The optimal number of decision trees is 72 by operation. Next, we construct the random forest model. Regarding parameter settings, “importance” is the parameter for judging the importance of variables, the “proximity” parameter is used to set the proximity matrix for calculating the model, and “ntree” is used to set the number of random forest decision trees. The importance score of each gene in the random forest model was calculated using the Gini coefficient method. Specifically, it is a method of decreasing accuracy. The relationship between the random forest model error and the number of decision trees is shown in the following figures (Figures 7A,B). The genes with a better importance score than two were selected as the hub genes for the subsequent analysis. MRPL51, NDUFA1, NDUFS5, RPS25, SHFM1, RPA3, and MAGOH, respectively. Among these genes, NDUFS5, SHFM1, RPA3, and MAGOH have never been mentioned or confirmed associated with AD development in studies.

We clustered the GSE63061 dataset and plotted a heat map (Figure 8), confirming that the seven hub genes mentioned above perform well in distinguishing between diseased and normal samples.

Correlation analysis between hub genes and immune characteristics

We quantified the immune infiltration scores of 28 immune cells in the samples using the method of ssGSEA. We plotted heat

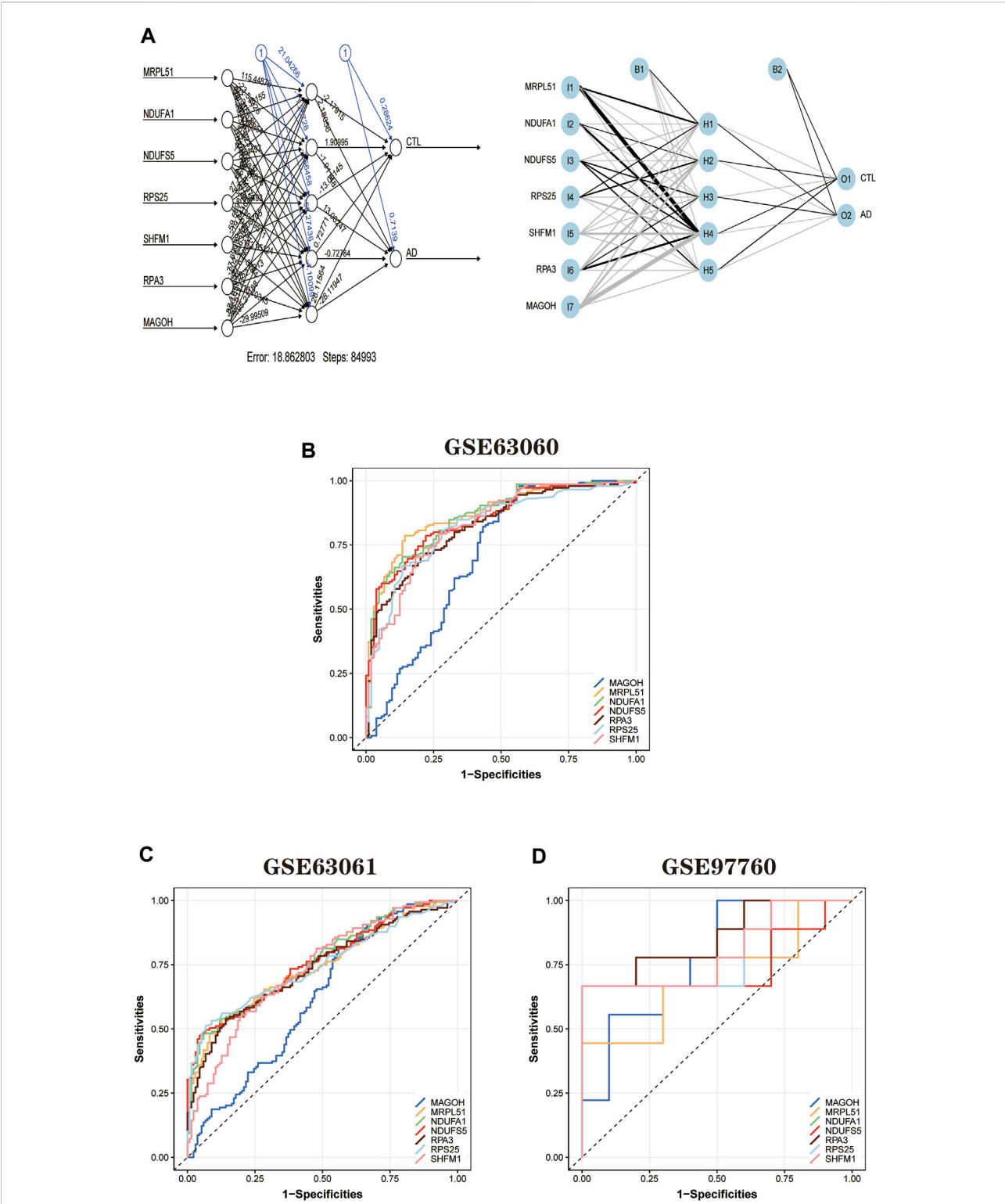


FIGURE 10
Artificial neural network model building and ROC curve validation. **(A)** Visualization of the artificial neural network model has undergone 84,993 training sessions and contains five hidden layers and two output layers. **(B)** ROC curves of the training set GSE63060 dataset. **(C)** The validation results of ROC curves in the validation set GSE63061 dataset. **(D)** The validation results of ROC curves in the validation set GSE97760 dataset. The different color lines represent different genes.

maps showing the distribution of immune cells in different samples and the infiltration scores (Figure 9A; Supplementary Material S7). The discrepancy in immune cell infiltration between the AD and CTL groups was then computed and visualized the results. The results showed (Figure 9B; Supplementary Material S8) that the proportion of CD56dim.natural.killer. Cells, MDSC, Monocyte, Natural.killer.T.cells, and Regulatory.T.cells were substantially higher in the AD group than in the CTL group. And many cells had lower fractions than normal patients, such as Activated.B.cells, Activated. CD4.T.cells, Activated. CD8.T.cells, Gamma.delta.T.cells, Effector. memory.CD4.T.cells, Central. CD4.T.cell.

We analyzed the correlation between immune cell infiltration scores and hub genes using Spearman correlation to explore the role played by hub genes in the immune microenvironment and the corresponding mechanisms. The results are shown (Figure 9C): the seven hub genes identified by the random forest algorithm were strongly correlated with the level of immune cell infiltration, suggesting that these genes may play a role in the development of AD by regulating the immune microenvironment.

Construction and validation of artificial neural network models

First, the expression profile data of the seven hub genes identified by the random forest algorithm were imported. Normalization of the input data was used to normalize the data. The input variables were normalized through the input nodes, and the normalized values fell between 0 and 1, or -1 to 1. We chose min-max (0,1) and performed the extrapolation. When choosing the parameters, we set the number of hidden layers to 5. There is no fixed rule for the number of hidden layers and the number of input neurons. The number of neurons is generally between two-thirds of the input layer size and one-third of the output layer size. In this study, the number of neurons is set to 7. The training and validation sets used to train the model are created randomly from the input data set. The purpose of the training set is to calculate the importance value score (gene weights) for each candidate gene. And the validation set is used to test the classification performance of the model scores using the expression of genes and gene weights. Finally, the formula was used.

$$\text{neural AD} = \text{Gene Expression} \times \text{Gene Weight}$$

The disease neural network classification score neural AD is obtained. The specific training process is as follows: ① The initial value of the network weights is set to 0, and the function of each node estimates the target variable value of the data. ② Compare the error between the actual and estimated values and readjust the bias of each weight according to the error value. Step ① is repeatedly executed until the error between the actual and

calculated values is minimized, at which point learning is stopped to obtain the best weights. The model's training process went through a total of 84,993 steps, and the termination condition (reaching the threshold) was the absolute partial derivative of the error function < 0.01. The output results of the artificial neural network model and the weight information of the candidate genes are shown in the table (Figure 10A; Supplementary Material S9). The accuracy of the artificial neural network model is reflected by the AUC values of the hub genes, and the larger the value, the higher the accuracy of the model is proved. We calculated the AUC values of the hub genes (Figure 10B): MRPL51 (0.87), NDUFA1 (0.86), NDUFS5 (0.85), RPS25 (0.82), SHFM1 (0.83) RPA3 (0.83), and MAGOH (0.81).

In addition, to further validate the accuracy of the ANN model, two independent datasets (GSE63061 and GSE97760) were selected for analysis. During the validation of the ANN model accuracy using the independent dataset GSE63061, we calculated the AUC values of seven hub genes using the same method (Figure 10C): MRPL51 (0.74), NDUFA1 (0.76), NDUFS5 (0.76), RPS25 (0.74), SHFM1 (0.73), RPA3 (0.73), and MAGOH (0.62). It is worth noting that the AUC values of the above genes remained significant when validated using the independent dataset GSE97760 (Figure 10D) for MRPL51 (0.69), NDUFA1 (0.77), NDUFS5 (0.74), RPS25 (0.79), SHFM1 (0.80), and RPA3 (0.86), and MAGOH (0.78). The validation results confirm that the ANN model has good classification performance for AD and normal samples.

Discussion

In this study, we used a combination of bioinformatics analysis and machine learning to obtain differential genes (DEGs) for AD serology and did GO and KEGG enrichment analysis. We got the following results from the enrichment analysis of the obtained DEGs. GO analysis showed that DEGs were significantly enriched in ribosomal and mitochondrial functions. The KEGG pathway leads that DEGs are mainly involved in "ribosome," "oxidative phosphorylation," "thermogenesis," "Parkinson's disease," and "non-alcoholic fatty liver disease (NAFLD)" and "Alzheimer's disease" pathways. Previous studies have confirmed the role of the ribosomal (Ding et al., 2005; Nyhus et al., 2019) and oxidative stress pathways (Simunkova et al., 2019; Butterfield and Mattson, 2020; Zhang et al., 2020) in AD development, which deserves further exploration.

Then seven hub genes MRPL51, NDUFA1, NDUFS5, RPS25, SHFM1, RPA3, and MAGOH were obtained by a random forest algorithm. Among these seven hub genes, NDUFA1 and MRPL51 were considered potential biomarkers of AD in previous bioinformatics analyses (Li et al., 2018; Liu et al., 2021). NDUFA1 is an essential component of the human

respiratory chain complex I. It is involved in mitochondrial function and oxidative phosphorylation and is a critical coding gene in the human body. Some studies have confirmed that partial deletion of respiratory chain function may impair ATP synthesis and chronic increase of oxidative stress (Carelli et al., 2002). A survey of optic neuropathy also suggested that reduced gene expression of mitochondrial proteins leads to neuronal degeneration (Qi et al., 2003). Both studies indicate that downregulation of the *NDUFA1* gene in the serum of AD patients is likely to lead to impaired oxidative phosphorylation and partial deficiency of mitochondrial function, which is ultimately involved in the disease development of AD. In other studies, it has been suggested that *RPS25* and *AIF1* may also play a role in AD development (Sanfilippo et al., 2020; Wang et al., 2021). In contrast, one study confirmed that *RPS25* is a therapeutic target for neurodegenerative diseases caused by nucleotide repeat amplification, which indirectly confirms the role of *RPS25* in the pathophysiological process of AD (Yamada et al., 2019). However, few studies have mentioned or established the association of *NDUFS5*, *SHFM1*, *RPA3*, and *MAGOH* genes with AD development. *NDUFS5* is a member of the iron-sulfur family of NADH dehydrogenases (ubiquinone) and encodes a subunit of the mitochondrial respiratory chain complex I (Wirth et al., 2016). Previous studies have highlighted the role of mitochondrial dysfunction in AD (Cai and Tammineni, 2017; Perez Ortiz and Swerdlow, 2019), leading us to speculate that *NDUFS5* may be involved in the pathogenesis of AD by affecting mitochondrial function and oxidative phosphorylation processes. *SHFM1* encodes the 26S proteasome subunit, one of the proteasome components. Earlier studies have confirmed the role of the proteasome in inhibiting neurodegeneration and that impaired proteasome function occurs in the early stages of AD (Keller et al., 2000; Cecarini et al., 2007). The differential expression of *SHFM1* in AD patients is likely to be a manifestation of proteasome dysfunction, offering the possibility of its use as a biomarker for early screening of AD. *RPA3* is a protein-coding gene mainly involved in DNA repair and DNA replication. It has been shown that disruption of DNA repair may lead to increased DNA damage in AD patients and increase the risk of AD, providing a theoretical basis for *RPA3* as a biomarker for AD. *MAGOH* is a protein-coding gene involved in the development of the nervous system. Little research has been done on this gene, and further studies are needed to elucidate its potential association with AD.

We further constructed a diagnostic model for AD using artificial neural networks based on the above seven hub genes. We validated the efficacy of the model in two publicly available datasets. The bioinformatics analysis combined with the machine learning approach is the innovation of this study, and good results were obtained. The random forest (RF) algorithm is an emerging and high precision machine learning algorithm that has been widely used in numerous

fields, and of course, its role in the medical field is also exact. RF algorithms have been used for clinical diseases, such as using random forests to identify biomarkers for glioblastoma to find potential targets for treatment (Li et al., 2021), building COPD risk prediction models (Perret et al., 2021), and detecting and predicting type 2 diabetes (Muneeb and Henschel, 2021), all with good results. An artificial neural network is a new type of algorithm derived from imitating the structure and function of the human brain, which has the characteristics of self-learning ability and high efficiency compared with the traditional machine learning algorithm. It has also found many applications in clinical settings. Studies have been using artificial neural network models to accurately predict the risk of liver failure after hepatectomy in patients with hepatocellular carcinoma who underwent hemihepatectomy (Mai et al., 2020). Artificial neural networks have also been used in AD for a long time. Some scholars have applied artificial neural networks to the diagnosis of AD based on the information contained in the digital images of SPECT cerebral blood flow assessment (Świetlik and Białowas, 2019). There is a precedent for combining two machine learning algorithms to diagnose and predict diseases (Mozafari et al., 2020; Xie et al., 2020). Still, it is noteworthy that no research has yet used this combination of the two in the field of AD (Feng et al., 2021). Therefore, combining random forests and artificial neural networks to build AD diagnosis models is a bold attempt and an excellent complement to the existing diagnosis methods. At the same time, our study revealed AD susceptibility genes that may be involved in the regulation of mitochondrial function and ribosomal pathways. We hope that their essential value will be reflected in future studies. Meanwhile, immunoassays showed that hub genes are closely related to immune cell infiltration, confirming that dysregulation of the immune microenvironment plays an essential role in the pathogenesis of AD.

However, there are limitations to this study. First, AD is highly heterogeneous, which affects our understanding and judgment of the disease (Jellinger, 1996). Second, some AD patients have other neurodegenerative lesions in combination, affecting the model's accuracy. And the data sample used in this study is still insufficient, and the sample size needs to be increased and further studied and optimized. Biological experiments for critical steps may be more revealing, but they cannot be completed at this time for objective reasons. In subsequent studies, we will continue to analyze these genes to further their upstream and downstream pathways to understand AD's biomolecular mechanisms better.

Conclusion

Using bioinformatics analysis and machine learning algorithm modeling, we uncovered potential biomarkers of AD based on

immune cell infiltration while constructing a random forest and artificial neural network AD diagnostic model. We confirmed its excellent classification performance in two independent datasets. This study nicely complements the existing tools for early screening and diagnosis of AD and reveals AD susceptibility genes that may be involved in the regulation of mitochondrial function and ribosomal function; and also provides new perspectives for a better understanding of molecular immune mechanisms and finding drug targets.

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

YH and LC completed the design and wrote the manuscript for this study, and these authors share first authorship. QH preprocessed the data used in this study and provided suggestions for data analysis. NF and YW provided constructive suggestions for this study. All authors contributed to the article and approved the submitted version.

Funding

This study was supported by the First-class disciplines and specialties project of The Second Affiliated Hospital of Harbin

References

- Bardou, P., Mariette, J., Escudié, F., Djemiel, C., and Klopp, C. (2014). Jvarkit: an interactive Venn diagram viewer. *BMC Bioinform.* 15, 293. doi:10.1186/1471-2105-15-293
- Bossaerts, L., Hens, E., Hanseeuw, B., Vandenberghe, R., Cras, P., De Deyn, P. P., et al. (2021). Premature termination codon mutations in ABCA7 contribute to Alzheimer's disease risk in belgian patients. *Neurobiol. Aging* 106, 307.e1–307307.e7. doi:10.1016/j.neurobiolaging.2021.04.023
- Butterfield, D., and Mattson, M. (2020). Apolipoprotein E and oxidative stress in brain with relevance to Alzheimer's disease. *Neurobiol. Dis.* 138, 104795. doi:10.1016/j.nbd.2020.104795
- Cai, Q., and Tammineni, P. (2017). Mitochondrial aspects of synaptic dysfunction in Alzheimer's disease. *J. Alzheimers Dis.* 57 (4), 1087–1103. doi:10.3233/JAD-160726
- Carrelli, V., Ross-Cisneros, F., and Sadun, A. (2002). Optic nerve degeneration and mitochondrial dysfunction: genetic and acquired optic neuropathies. *Neurochem. Int.* 40 (6), 573–584. doi:10.1016/s0197-0186(01)00129-2
- Cecarini, V., Ding, Q., and Keller, J. (2007). Oxidative inactivation of the proteasome in Alzheimer's disease. *Free Radic. Res.* 41 (6), 673–680. doi:10.1080/10715760701286159
- Cheng-Hathaway, P., Reed-Geaghan, E., Jay, T., Casali, B. T., Bemiller, S. M., Puntambekar, S. S., et al. (2018). The Trem2 R47H variant confers loss-of-function-like phenotypes in Alzheimer's disease. *Mol. Neurodegener.* 13 (1), 29. doi:10.1186/s13024-018-0262-8
- Ding, Q., Markesbery, W., Chen, Q., Li, F., and Keller, J. (2005). Ribosome dysfunction is an early event in Alzheimer's disease. *J. Neurosci.* 25 (40), 9171–9175. doi:10.1523/JNEUROSCI.3040-05.2005
- Feng, J., Chen, Y., Feng, Q., Ran, Z., and Shen, J. (2021). Novel gene signatures predicting primary non-response to infliximab in ulcerative colitis: development and validation combining random forest with artificial neural network. *Front. Med.* 8, 678424. doi:10.3389/fmed.2021.678424
- Gatz, M., Reynolds, C., Fratiglioni, L., Johansson, B., Mortimer, J. A., Berg, S., et al. (2006). Role of genes and environments for explaining Alzheimer disease. *Arch. Gen. Psychiatry* 63 (2), 168–174. doi:10.1001/archpsyc.63.2.168
- Holstege, H., van der Lee, S., Hulsman, M., Wong, T. H., van Rooij, J. G., Weiss, M., et al. (2017). Characterization of pathogenic SORL1 genetic variants for association with Alzheimer's disease: a clinical interpretation strategy. *Eur. J. Hum. Genet.* 25 (8), 973–981. doi:10.1038/ejhg.2017.87
- Hothorn, T., and Jung, H. (2014). RandomForest4Life: a random forest for predicting ALS disease progression. *Amyotroph. Lateral Scler. Front. Degener.* 15, 444–452. doi:10.3109/21678421.2014.893361
- Jack, C., Bennett, D., Blennow, K., Carrillo, M. C., Dunn, B., Haeberlein, S. B., et al. (2018). NIA-AA Research Framework: toward a biological definition of Alzheimer's disease. *Alzheimers Dement.* 14 (4), 535–562. doi:10.1016/j.jalz.2018.02.018
- Jansen, I., Savage, J., Watanabe, K., Bryois, J., Williams, D. M., Steinberg, S., et al. (2019). Genome-wide meta-analysis identifies new loci and functional pathways

Medical University (No. YLXK202007) and the Heilongjiang Province Postdoctoral Grant (No. LBH-Z16306).

Acknowledgments

We want to thank the GEO database for the data support.

Conflict of interest

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

Publisher's note

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

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2022.968598/full#supplementary-material>

- influencing Alzheimer's disease risk. *Nat. Genet.* 51 (3), 404–413. doi:10.1038/s41588-018-0311-9
- Jellinger, K. (1996). Recent update on the heterogeneity of the Alzheimer's disease spectrum. *J. Neural. Transmission*, 2021.
- Keller, J., Hanni, K., and Markesbery, W. (2000). Impaired proteasome function in Alzheimer's disease. *J. Neurochem.* 75 (1), 436–439. doi:10.1046/j.1471-4159.2000.0750436.x
- Li, H., Lai, L., and Shen, J. (2020). Development of a susceptibility gene based novel predictive model for the diagnosis of ulcerative colitis using random forest and artificial neural network. *Aging* 12 (20), 20471–20482. doi:10.18632/aging.103861
- Li, J., BiZhang, X., Cao, Y., Lv, K., and Jiang, L. (2021). Network pharmacology and inflammatory microenvironment strategy approach to finding the potential target of *Siraitia grosvenorii* (Luo han guo) for glioblastoma. *Front. Genet.* 12, 799799. doi:10.3389/fgene.2021.799799
- Li, X., Wang, H., Long, J., Pan, G., He, T., Anichtchik, O., et al. (2018). Systematic analysis and biomarker study for Alzheimer's disease. *Sci. Rep.* 8 (1), 17394. doi:10.1038/s41598-018-35789-3
- Liu, Z., Li, H., and Pan, S. (2021). Discovery and validation of key biomarkers based on immune infiltrates in Alzheimer's disease. *Front. Genet.* 12, 658323. doi:10.3389/fgene.2021.658323
- Mai, R., Lu, H., Bai, T., Liang, R., Lin, Y., Ma, L., et al. (2020). Artificial neural network model for preoperative prediction of severe liver failure after hemihepatectomy in patients with hepatocellular carcinoma. *Surgery* 168 (4), 643–652. doi:10.1016/j.surg.2020.06.031
- Morris, J. (2005). Early-stage and preclinical Alzheimer disease. *Alzheimer Dis. Assoc. Disord.* 19 (3), 163–165. doi:10.1097/01.wad.0000184005.22611.cc
- Mozafari, Z., Arab Chamjangali, M., Beglari, M., and Doosti, R. (2020). The efficiency of ligand-receptor interaction information alone as new descriptors in QSAR modeling via random forest artificial neural network. *Chem. Biol. Drug Des.* 96 (2), 812–824. doi:10.1111/cbdd.13690
- Muneeb, M., and Henschel, A. (2021). Eye-color and Type-2 diabetes phenotype prediction from genotype data using deep learning methods. *BMC Bioinforma.* 22 (1), 198. doi:10.1186/s12859-021-04077-9
- Nyhus, C., Pihl, M., Hyttel, P., and Hall, V. (2019). Evidence for nucleolar dysfunction in Alzheimer's disease. *Rev. Neurosci.* 30 (7), 685–700. doi:10.1515/revneuro-2018-0104
- Perez Ortiz, J., and Swerdlow, R. (2019). Mitochondrial dysfunction in Alzheimer's disease: role in pathogenesis and novel therapeutic opportunities. *Br. J. Pharmacol.* 176 (18), 3489–3507. doi:10.1111/bph.14585
- Perret, J., Vicendese, D., Simons, K., Jarvis, D. L., Lowe, A. J., Lodge, C. J., et al. (2021). Ten-year prediction model for post-bronchodilator airflow obstruction and early detection of COPD: development and validation in two middle-aged population-based cohorts. *BMJ Open Respir. Res.* 8 (1), e001138. doi:10.1136/bmjresp-2021-001138
- Qi, X., Lewin, A., Hauswirth, W., and Guy, J. (2003). Suppression of complex I gene expression induces optic neuropathy. *Ann. Neurol.* 53 (2), 198–205. doi:10.1002/ana.10426
- Ritchie, M., 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 (7), e47. doi:10.1093/nar/gkv007
- Sabayan, B., and Sorond, F. (2017). Reducing risk of dementia in older age. *JAMA* 317 (19), 2028. doi:10.1001/jama.2017.2247
- Sanfilippo, C., Castrogiovanni, P., Imbesi, R., and Di Rosa, M. (2020). CHI3L2 expression levels are correlated with AIF1, PECAM1, and CALB1 in the brains of Alzheimer's disease patients. *J. Mol. Neurosci.* 70 (10), 1598–1610. doi:10.1007/s12031-020-01667-9
- Scheltens, P., De Strooper, B., Kivipelto, M., Holstege, H., Chetelat, G., Teunissen, C. E., et al. (2021). Alzheimer's disease. *Lancet* 397 (10284), 1577–1590. doi:10.1016/S0140-6736(20)32205-4
- Shi, M., and Xu, G. (2019). Development and validation of GMI signature based random survival forest prognosis model to predict clinical outcome in acute myeloid leukemia. *BMC Med. Genomics* 12 (1), 90. doi:10.1186/s12920-019-0540-5
- Simunkova, M., Alwasel, S., Alhazza, I., Jomova, K., Kollar, V., Rusko, M., et al. (2011). Management of oxidative stress and other pathologies in Alzheimer's disease. *Arch. Toxicol.* 93 (9), 2491–2513. doi:10.1007/s00204-019-02538-y
- Sperling, R. A., Aisen, P. S., Beckett, L. A., Bennett, D. A., Craft, S., Fagan, A. M., et al. (2011). Toward defining the preclinical stages of Alzheimer's disease: recommendations from the national institute on aging-alzheimer's association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement.* 7 (3), 280–292. doi:10.1016/j.jalz.2011.03.003
- Sudom, A., Talreja, S., Danao, J., Bragg, E., Kegel, R., Min, X., et al. (2018). Molecular basis for the loss-of-function effects of the Alzheimer's disease-associated R47H variant of the immune receptor TREM2. *J. Biol. Chem.* 293 (32), 12634–12646. doi:10.1074/jbc.RA118.002352
- Świetlik, D., and Białowąs, J. (2019). Application of artificial neural networks to identify Alzheimer's disease using cerebral perfusion SPECT data. *Int. J. Environ. Res. Public Health* 16 (7), E1303. doi:10.3390/ijerph16071303
- Tian, Y., Yang, J., Lan, M., and Zou, T. (2020). Construction and analysis of a joint diagnosis model of random forest and artificial neural network for heart failure. *Aging* 12 (24), 26221–26235. doi:10.18632/aging.202405
- Wang, H., Han, X., and Gao, S. (2021). Identification of potential biomarkers for pathogenesis of Alzheimer's disease. *Hereditas* 158 (1), 23. doi:10.1186/s41065-021-00187-9
- Wirth, C., Brandt, U., Hunte, C., and Zickermann, V. (2016). Structure and function of mitochondrial complex I. *Biochim. Biophys. Acta* 1857 (7), 902–914. doi:10.1016/j.bbabo.2016.02.013
- Xie, N., Wang, F., Zhou, J., Liu, C., and Qu, F. (2020). Establishment and analysis of a combined diagnostic model of polycystic ovary syndrome with random forest and artificial neural network. *Biomed. Res. Int.* 2020, 2613091. doi:10.1155/2020/2613091
- Yamada, S., Gendron, T., Niccoli, T., Genuth, N. R., Grosely, R., Shi, Y., et al. (2019). RPS25 is required for efficient RAN translation of C9orf72 and other neurodegenerative disease-associated nucleotide repeats. *Nat. Neurosci.* 22 (9), 1383–1388. doi:10.1038/s41593-019-0455-7
- Yang, L., Wu, H., Jin, X., Zheng, P., Hu, S., Xu, X., et al. (2020). Study of cardiovascular disease prediction model based on random forest in eastern China. *Sci. Rep.* 10 (1), 5245. doi:10.1038/s41598-020-62133-5
- Zhang, L., Xu, M., Ren, Q., Liu, G., Meng, S., Xiahou, K., et al. (2020). Human induced pluripotent stem cell-derived neural cells from Alzheimer's disease patients exhibited different susceptibility to oxidative stress. *Stem Cells Dev.* 29 (22), 1444–1456. doi:10.1089/scd.2020.0103



OPEN ACCESS

EDITED BY

Thayne Kowalski,
Centro Universitário Cesuca, Brazil

REVIEWED BY

Patrycja Chylińska-Wrzos,
Lublin, Medical University of Lublin,
Poland
Fnu Shiridhar,
University of Central Florida,
United States

*CORRESPONDENCE

Jianzhong Li,
jianzhong-0520@163.com
Shouping Gong,
shpingg@126.com

[†]These authors have contributed equally
to this work

SPECIALTY SECTION

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

RECEIVED 13 July 2022

ACCEPTED 08 August 2022

PUBLISHED 29 August 2022

CITATION

Yu B, Tian Y, Zhang Y, Lv B, Li J and
Gong S (2022), Experimental verification
and validation of immune biomarkers
based on chromatin regulators in
ischemic stroke.
Front. Genet. 13:992847.
doi: 10.3389/fgene.2022.992847

COPYRIGHT

© 2022 Yu, Tian, Zhang, Lv, Li and Gong.
This is an open-access article
distributed under the terms of the
[Creative Commons Attribution License](#)
(CC BY). The use, distribution or
reproduction in other forums is
permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original
publication in this journal is cited, in
accordance with accepted academic
practice. No use, distribution or
reproduction is permitted which does
not comply with these terms.

Experimental verification and validation of immune biomarkers based on chromatin regulators in ischemic stroke

Beibei Yu^{1†}, Yunze Tian^{1†}, Yongfeng Zhang¹, Boqiang Lv¹,
Jianzhong Li^{2*} and Shouping Gong^{1*}

¹Department of Neurosurgery, the Second Affiliated Hospital of Xi'an Jiao Tong University, Xi'an, China,

²Department of Thoracic Surgery, the Second Affiliated Hospital of Xi'an Jiao Tong University, Xi'an, China

Ischemic stroke (IS) is a disease characterized by rapid progression and high mortality and disability rates. Its pathophysiological process is inseparable from immune dysfunction. Recently, chromatin regulators (CRs) have been described as a class of enzymes that can recognize, form, and maintain the epigenetic state of an organism, and are closely associated with immune regulation. Nevertheless, the role of CR-related genes in IS has not been fully elucidated. In this study, seven CR-related immune biomarkers in the GSE58294 and GSE22255 datasets were identified by combining differential gene expression analysis, weighted correlation network analysis, and single sample gene set enrichment analysis. After experimental validation using quantitative polymerase chain reaction, four genes (*DPF2*, *LMNB1*, *MLLT3*, and *JAK2*) were screened as candidate immune biomarkers. These four biomarkers demonstrated good predictive power in the clinical risk model (area under the curve, 0.775). Molecular docking simulations revealed that mevastatin, WP1066, cladribine, trichostatin A, mequitazine, and zuclophene may be potential immunomodulatory drugs for IS. Overall, the results of this study contribute to the identification of CR-related immune therapeutics target in IS and provide an important reference for further research.

KEYWORDS

ischemic stroke, immune infiltration, chromatin regulators, biomarker, bioinformatics

Introduction

Stroke is the second leading cause of death and disability, killing >5.5 million individuals annually (Lindsay et al., 2019). Among stroke types, ischemic stroke (IS) accounts for approximately 71–87% of cases and is caused mainly by blockage of blood flow in the brain (Strong et al., 2007; Campbell et al., 2019). The current treatment for IS focuses on achieving rapid reperfusion through intravenous thrombolysis and/or endovascular thrombectomy. Although these treatments can

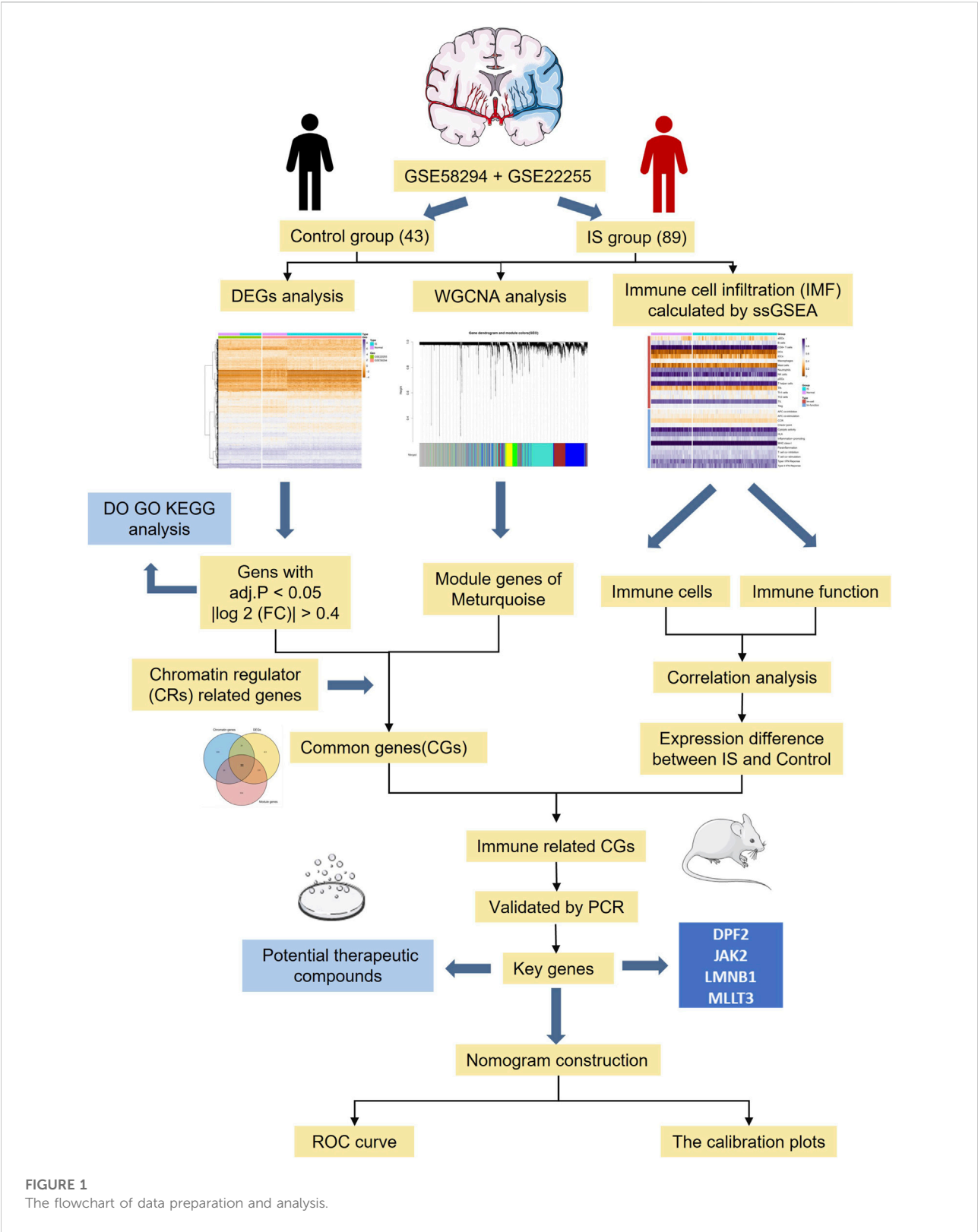


FIGURE 1
The flowchart of data preparation and analysis.

reduce disability, patients continue to experience serious economic and social burdens (Rochmah et al., 2021). Recent studies have reported that IS can induce immune-inflammatory responses, which in turn aggravate neurological deficits and increase patient mortality (Yu et al., 2020). As such, there is an urgent need to explore the influence of the immune system on IS and to identify potential therapeutic targets.

Chromatin regulators (CRs) are a class of enzymes with special structural functions that can recognize, form, and maintain the epigenetic state of the organism (Lu et al., 2018). Somatic alterations or abnormal expression of CRs may lead to fatal diseases including glioma, lung cancer, bladder cancer, and other tumors. At the same time, CRs are also considered to be important therapeutic targets in diseases such as colitis, and congenital heart disease (Li et al., 2022; Linglart and Bonnet, 2022). In addition, recent research has also found that modulation of these epigenetic genes could improve cellular immune responses in immunotherapy (Belk et al., 2022). Consequently, actively exploring the role of CR-related genes will help to better understand the regulatory role of the immune system in IS.

With the development of high-throughput sequencing and gene chip technologies, bioinformatics can be used to conduct extensive and in-depth analyses of messenger RNA (mRNA) expression profiles of IS (Cao et al., 2021a). In this study, we identified key CR-related immune biomarkers in the GSE58294 and GSE22255 datasets by combining differential gene expression analysis, weighted correlation network analysis (WGCNA), and single-sample gene set enrichment analysis (ssGSEA). After experimental validation of a middle cerebral artery occlusion (MCAO) model, we evaluated the accuracy of key immune therapeutics target in predicting the occurrence of IS. Results of this study may provide a theoretical molecular basis for the diagnosis and targeted treatment of IS. A flow-diagram illustrating our research process is presented in Figure 1.

Methods

Establishment of the MCAO model

Twelve specific pathogen-free male Sprague-Dawley rats (weight 280 g) were provided by the Medical Experimental Animal Center (Xi'an Jiaotong University). The Xian Jiaotong University Medical College and Xian Jiaotong University Experimental Animal Center approved all protocols involving animal models or specimens. The rat MCAO model was established based on the modified Zea-Longa model, which removed coil occlusion after 2 h (Longa et al., 1989). Thereafter, a total of 12 rats were randomly assigned to sham and the IS groups (n = 6 each). Two hours after MCAO,

the neurobehavioral scores of the rats in each group were evaluated according to the Longa scale (Longa et al., 1989). Animals that did not exhibit neurological deficits after surgery were excluded. After 3 days of reperfusion, three rats in each group were euthanized using sodium pentobarbital (30 mg/kg) injected through the tail vein. After the brains were removed and sliced, specimens were placed in 2% TTC (Solarbio Life Science, Beijing, China) and incubated at 37°C for 30 min.

GEO dataset screening

Two publicly available GEO datasets, GSE58294 (Stamova et al., 2014) and GSE22255 (Krug et al., 2012), were used to identify differentially expressed genes (DEGs) between those with IS and healthy individuals (Table 1). GSE58294 contains the whole blood mRNA expression profiles of 69 IS patients and 23 healthy individuals, while GSE22255 contains data from 20 IS patients and 20 healthy individuals.

Identification of DEGs and key module genes

Statistical analysis was performed using R version 4.1.0 (R Foundation for Statistical Computing, Vienna, Austria). To analyze microarray data, the gene expression matrix was quantile-normalized and log₂-transformed after merging it with the probe data (GPL570). Batch effects were removed from the gene expression profiles by merging these two matrices using the comBat algorithm in the sva package of R. The “LIMMA” package was used to identify DEGs, and the filter was |log₂ (fold-change) > 0.4 and adjusted *p*-value < 0.05.

To explore interactions between genes, the “WGCNA” package was used to perform WGCNA and identify the key module. Thereafter, the “pickSoftThreshold” function was used to obtain the optimal value of the adjacent function weighting parameters, which were used as a soft threshold for subsequent network construction. The related gene modules were then constructed based on hierarchical clustering of the dissimilarity measure. Finally, the correlation between each module and the sample trait (presence or absence of IS) was assessed. Modules with the highest correlation coefficients were selected for further analysis.

Screening of CR-related biomarkers

A total of 870 CR-related genes was obtained from the supplemental information published by Lu et al. (2018). To identify CRs-related biomarkers, Venn analysis was used to determine the intersection of the above three gene sets, DEGs,

TABLE 1 Detailed information of the gene expression matrixes and platform.

| GEO dataset | Platform | Country | Author | Stroke | Normal | Type |
|-------------|----------|----------|----------------|--------|--------|------|
| GSE58294 | GPL570 | USA | Stamova et al. | 69 | 23 | mRNA |
| GSE22255 | GPL570 | Portugal | Krug et al. | 20 | 20 | mRNA |

the key module genes derived from WGCNA, and CR-related genes. Gene Ontology (GO), Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway, and disease enrichment (DO) analyses were performed on DEGs using the “clusterProfiler” package to identify the underlying pathogenesis and biological pathways of IS.

Immune infiltration analysis

The degree of immune cell infiltration in each sample was calculated using the ssGSEA algorithm. After downloading the gene set data in “gmt” format with 29 immune-related score, the “GSVA” package was used to score each sample in GSE58294 and GSE22255. Subsequently, the differences in the expression of 16 immune cells and 13 immune-related functions between the IS and healthy samples were further distinguished. Finally, Spearman’s correlation analysis was performed to calculate the correlation coefficients between the CR-related biomarkers obtained in the previous step. Screening criteria for key immune biomarkers related to IS were as follows: > 1/5 of immune cells and function correlation; correlation coefficient, $r > 0.3$; and $p < 0.05$.

Construction and validation of IS risk models

Based on the median expression levels of polymerase chain reaction (PCR)-validated immune biomarkers in the samples, they were divided into high and low expression groups. Subsequently, the “rms” package was used to build nomogram model for predicting the incidence of IS. Finally, the “ROCR” package was used to plot a receiver operating characteristic (ROC) curve and a calibration curve to evaluate the effectiveness of the risk model (Cao et al., 2021b).

More importantly, these immune therapeutics target were uploaded to the DSigDB database (<http://dsigdb.tan-lab.org/DSigDBv1.0/>) to screen small-molecule compounds for IS (Yoo et al., 2015). The screening criteria included $p < 0.05$ and $n \geq 3$. The molecular docking software AutoDock Vina was used to simulate the docking between proteins and polymers.

TABLE 2 Specific primers used for quantitative real-time PCR.

| Primer | Sequence |
|-----------|---------------------------|
| GAPDH-F | GGTCGGTGTGAACGGATTT |
| GAPDH-R | TGAACTTGCCGTGGGTAGA |
| JAK2-F | ACACCTCTGATCCCTCAGC |
| JAK2-R | GCGAATGATAAACAGGCAGGATG |
| MLLT3-F | ACAACGAGGAGGAGTCTGATGAGG |
| MLLT3-R | CACTGTCAGTCCGTCACCTCAAG |
| LMNB1-F | AGCTCTCTCCAAGTCTTCTTCCC |
| LMNB1-R | CACTACTGCTCGCCTCTGATTCTTC |
| GLYATL1-F | GCAGTGAGAGGAGCCAACGATTTC |
| GLYATL1-R | ATCAGAGCCCAGGACACAGGAG |
| DPF2-F | TAAGCCAGACACGGACCAGACTC |
| DPF2-R | CAGTACGCAGCAGAGCCTCTAAAC |
| BRCA1-F | CAGATCGAGAGTTGTGGTAGCAGTG |
| BRCA1-R | TTGGCTCGTTCTTCTTGGCATCAG |

Quantitative real-time PCR

After 72 h of reperfusion, three rats from each group were euthanized under anesthesia. Tissue samples were collected from the ischemic penumbra of the rats and immediately stored in liquid nitrogen. Total RNA was extracted from each sample using TRIzol reagent (Invitrogen, United States). The extracted RNA was reverse transcribed into complementary DNA using PrimeScript™ RT Master Mix (TaKaRa, Japan). The primer sequences are listed in Table 2. Relative mRNA expression was calculated using the 2- $\Delta\Delta C_t$ method and compared to that of the control group (GAPDH mRNA expression). Statistical comparisons were performed using the student’s *t*-test; differences with $p < 0.05$ were considered to be statistically significant.

3 Results

Identification of DEGs using the LIMMA package

After the successful merging of GSE58294 and GSE22255, 987 genes were differentially expressed, of which 445 were upregulated and 542 were

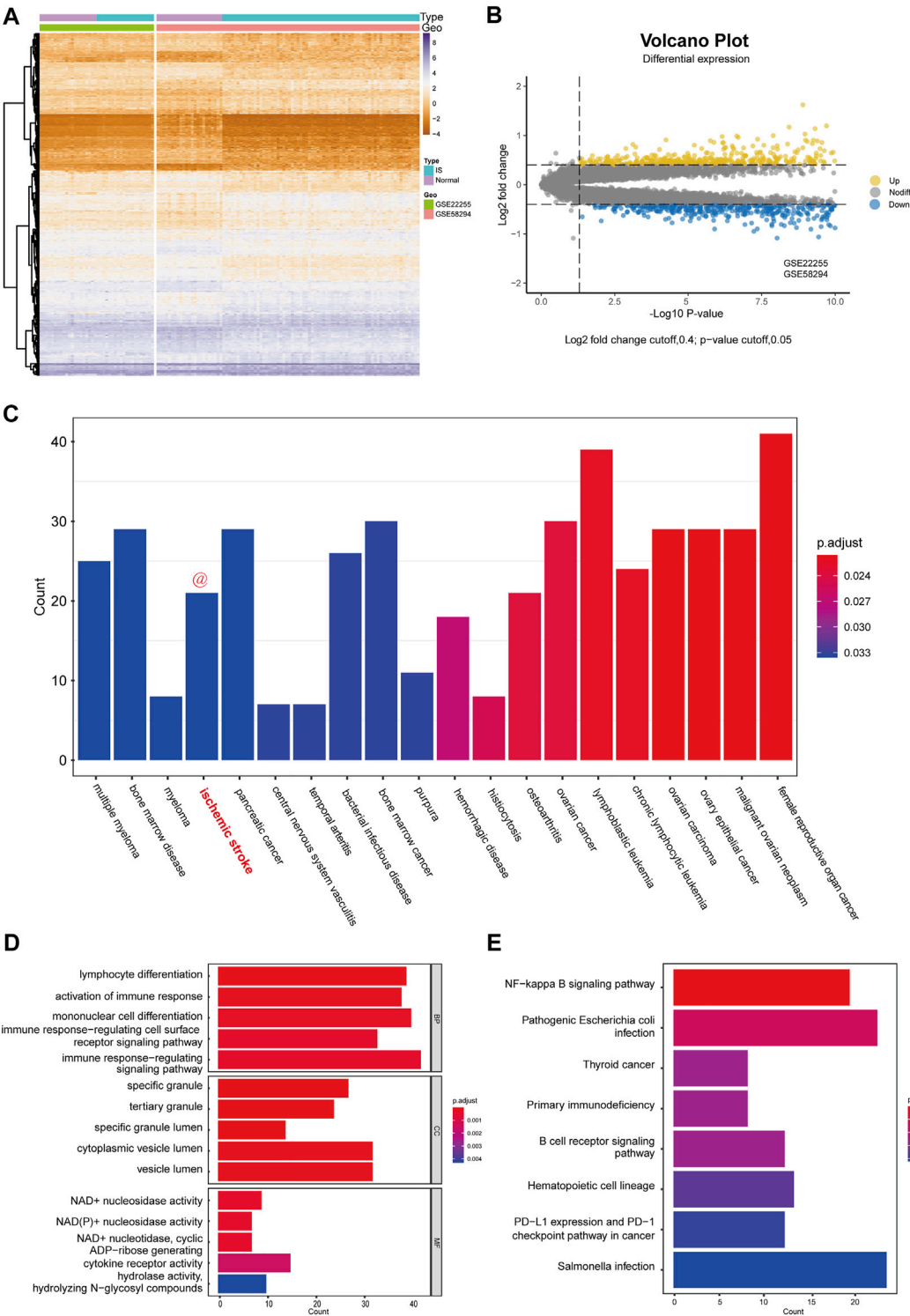


FIGURE 2
Differential expression analysis and functional enrichment analysis. **(A)** Cluster heatmap for DEmRNAs in GSE58294 and GSE22255 dataset. Yellow represents high gene expression and purple represents low expression. **(B)** Volcano plot for DEmRNAs in GSE58294 and GSE22255 dataset. **(C)** Disease enrichment analysis of DEGs. **(D)** GO functional enrichment analysis of DEGs. **(E)** KEGG pathway analysis of DEGs.

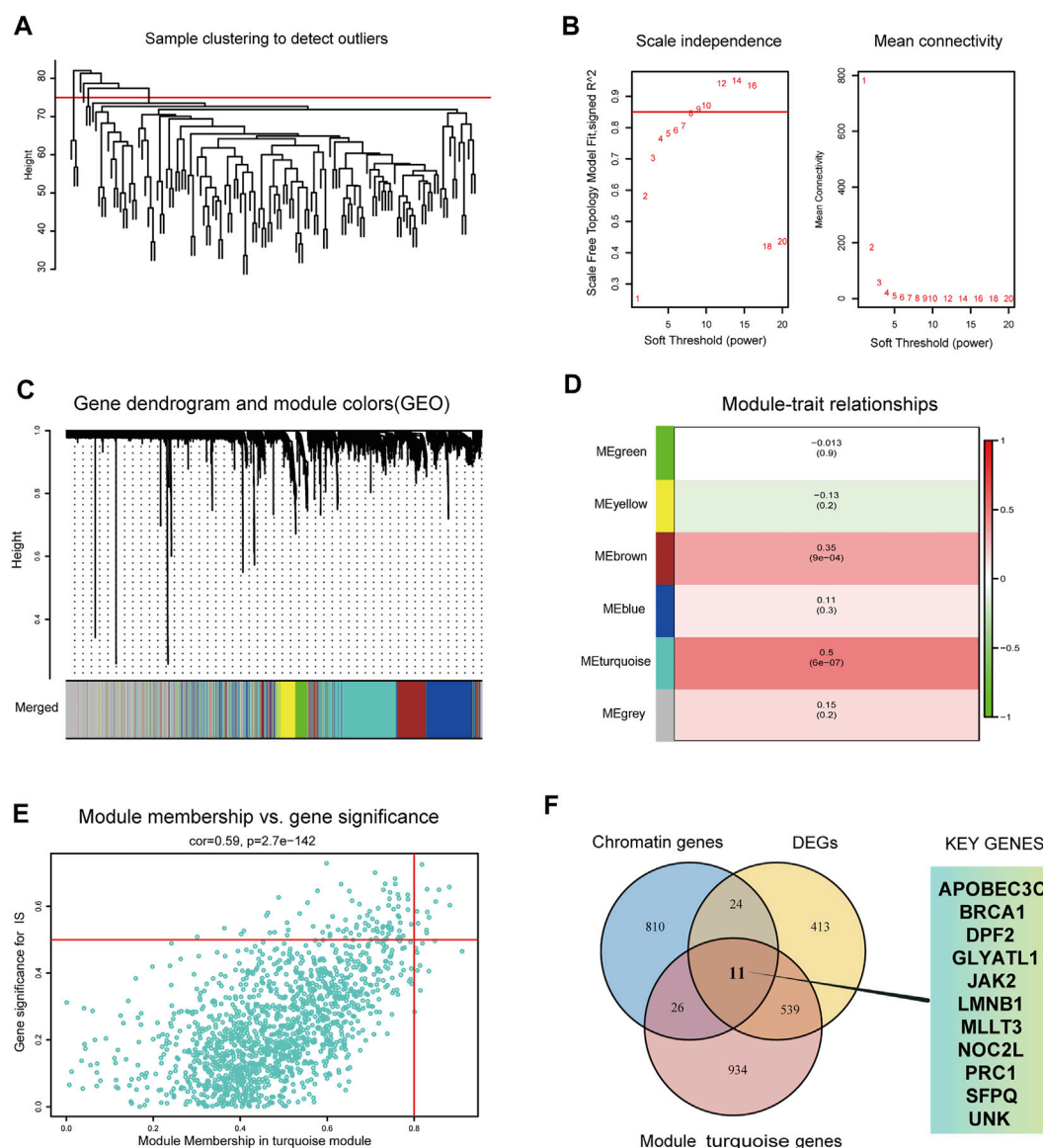


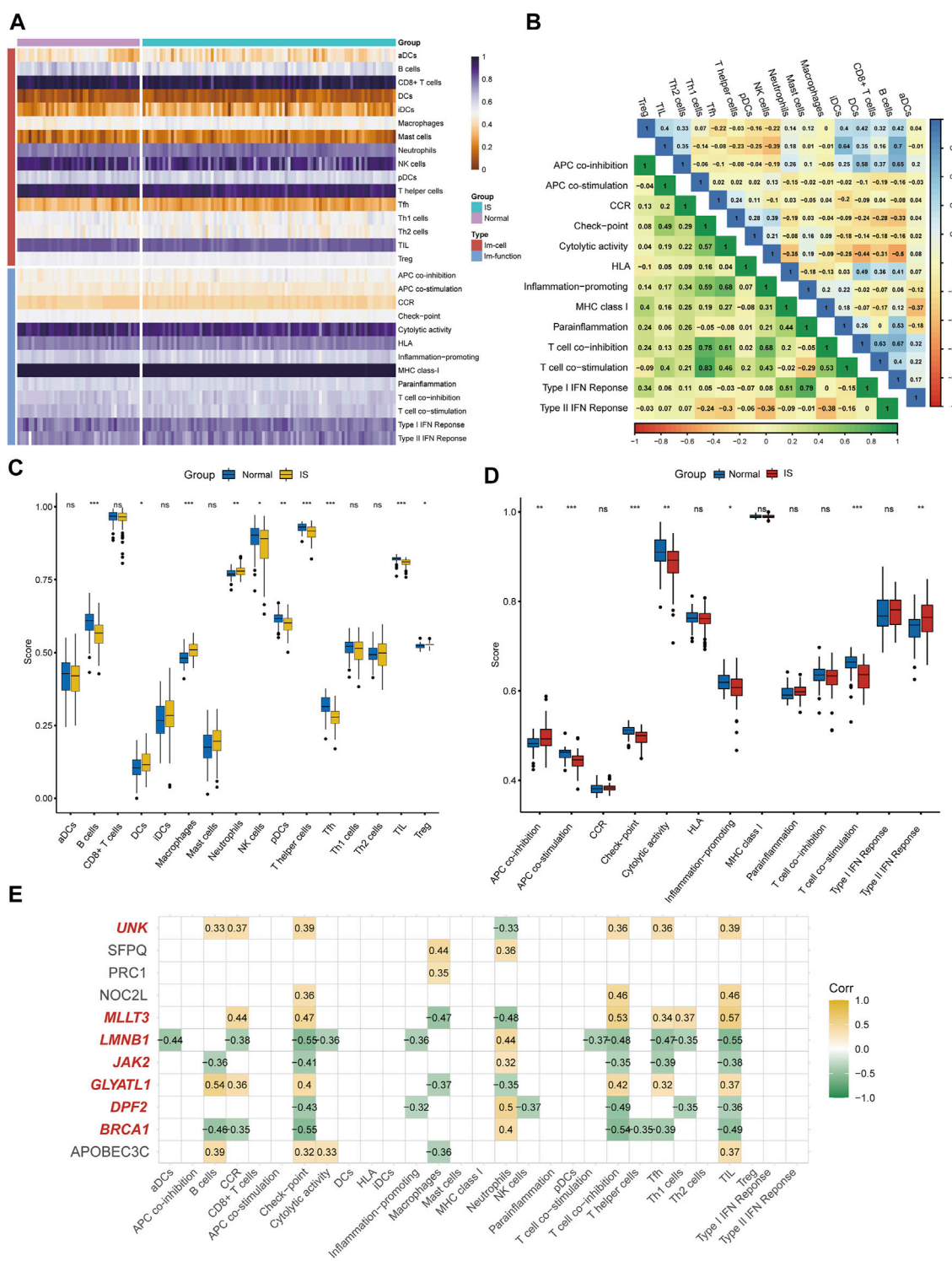
FIGURE 3

Identification of key modules genes and CRs related biomarkers (A) Outlier removal with “h > 75”. (B) Soft threshold setting according to $R^2 = 0.85$. (C) The gene set is divided into 6 different modules. (D) The correlation of modules with IS occurrence. (E) The correlation between turquoise module and IS gene significance. (F) The Venn plot of the three gene sets, DEGs, turquoise module genes and CRs-related genes.

downregulated. The DEGs results are presented in the form of a heatmap and volcano plot (Figure 2A,B). DO analysis revealed that the DEGs were specific to the occurrence of IS (Figure 2C).

To better assess the function of the DEGs, GO functional enrichment (Figure 2D) and KEGG pathway analyses (Figure 2E) were performed. Biological processes (BP) included “lymphocyte differentiation,” “activation of immune response,” “mononuclear cell differentiation,” “immune response-regulating cell surface

receptor signaling pathway,” “immune response-regulating signaling pathway,” which reflected that DEGs has a strong correlation with immune function. In addition, cellular components (CC) indicated that DEGs were involved in the release of specific granules of immune cells, such as neutrophils, while molecular functions (MF) indicated that DEGs were involved in the metabolism of $NAD^+/NADH$. In addition, KEGG pathway analysis indicated that the DEGs were mostly associated with three pathways: the NF-kappa B signaling



pathway; primary immunodeficiency; and B cell receptor signaling pathway.

Identification of turquoise genes and CRs-related biomarkers

First, WGCNA was used to select genes with a variance >25%, followed by sample cluster analysis to eliminate outlier samples with $h > 75$ (Figure 3A). The optimal soft threshold $\beta = 9$ was determined according to a scale-free fitting index (R2) of 0.85 (Figure 3B). As shown in Figures 3C,D, WGCNA identified six modules, of which the turquoise module was strongly correlated with IS (module feature correlation = 0.50). Moreover, correlation analysis revealed that the correlation between module membership in turquoise modules and gene significance for IS was 0.59, which was statistically significant (Figure 3E). The intersection of the three gene sets of DEGs, turquoise module genes, and CR-related genes is shown in Figure 3F. A total of 11 overlapping biomarkers related to CRs were identified for further analysis.

Immune infiltration analysis and immune biomarker screening

According to the ssGSEA algorithm, the immune infiltration scores of 16 immune cells and 13 immune-related functions in 132 samples were evaluated (Figure 4A). For immune cells, B cells had a strong positive correlation ($r = 0.64$) with follicular helper T cells, whereas neutrophils had a strong negative correlation ($r = -0.44$) with T-helper one cells (Figure 4B). For immune function, type-I interferon (IFN) response had a strong positive correlation ($r = 0.79$) with para-inflammation, and type-II IFN response had a strong negative correlation ($r = -0.38$) with T cell co-inhibition.

Differences in immune cell content and immune function between IS and healthy individuals were further explored (Figures 4C,D). It was clearly shown that the proportion of dendritic cells, macrophages, neutrophils, and regulatory T cells was higher in the IS group than in the control group, and the functions of antigen presenting cell (APC) co-inhibition and type-II IFN response were better than those of healthy individuals. Meanwhile, the proportion of B cells, NK cells, plasmacytoid DC cells, T-helper cells, follicular helper T cells, and tumor-infiltrating lymphocytes in the control group were higher than those in the IS group, and the functions of APC co-stimulation, checkpoint, cytolytic activity, inflammation-promoting, and T cell co-stimulation were greater than those of IS patients. Finally, seven of 11 DEGs (*BRCA1*, *DPF2*, *GLYATL1*, *LMNB1*, *MLLT3*, *JAK2*, and *UNK*) were identified as immune biomarkers based on correlation analysis (Figure 4E).

Validation of key immune biomarkers

As shown in Figures 5A,B, compared with the sham group, the IS group exhibited obvious cerebral infarction lesions and increased Longa scores, indicating that the MCAO model was successful. During PCR validation of the seven immune biomarkers, *UNK* was excluded due to its poor species conservation between humans and rats. In the GEO expression matrix, the mRNA expression levels of *BRCA1*, *DPF2*, *LMNB1*, and *JAK2* in the IS group were higher than those in the control group, and *MLLT3* and *GLYATL1* were lower than those in the control group (Figure 5C). Real-time-qPCR (Figure 5D) further demonstrated that the mRNA expression of *DPF2*, *LMNB1*, *MLLT3*, and *JAK2* was consistent with the prediction of bioinformatics analysis and was statistically significant ($p < 0.05$).

Construction of clinical risk models

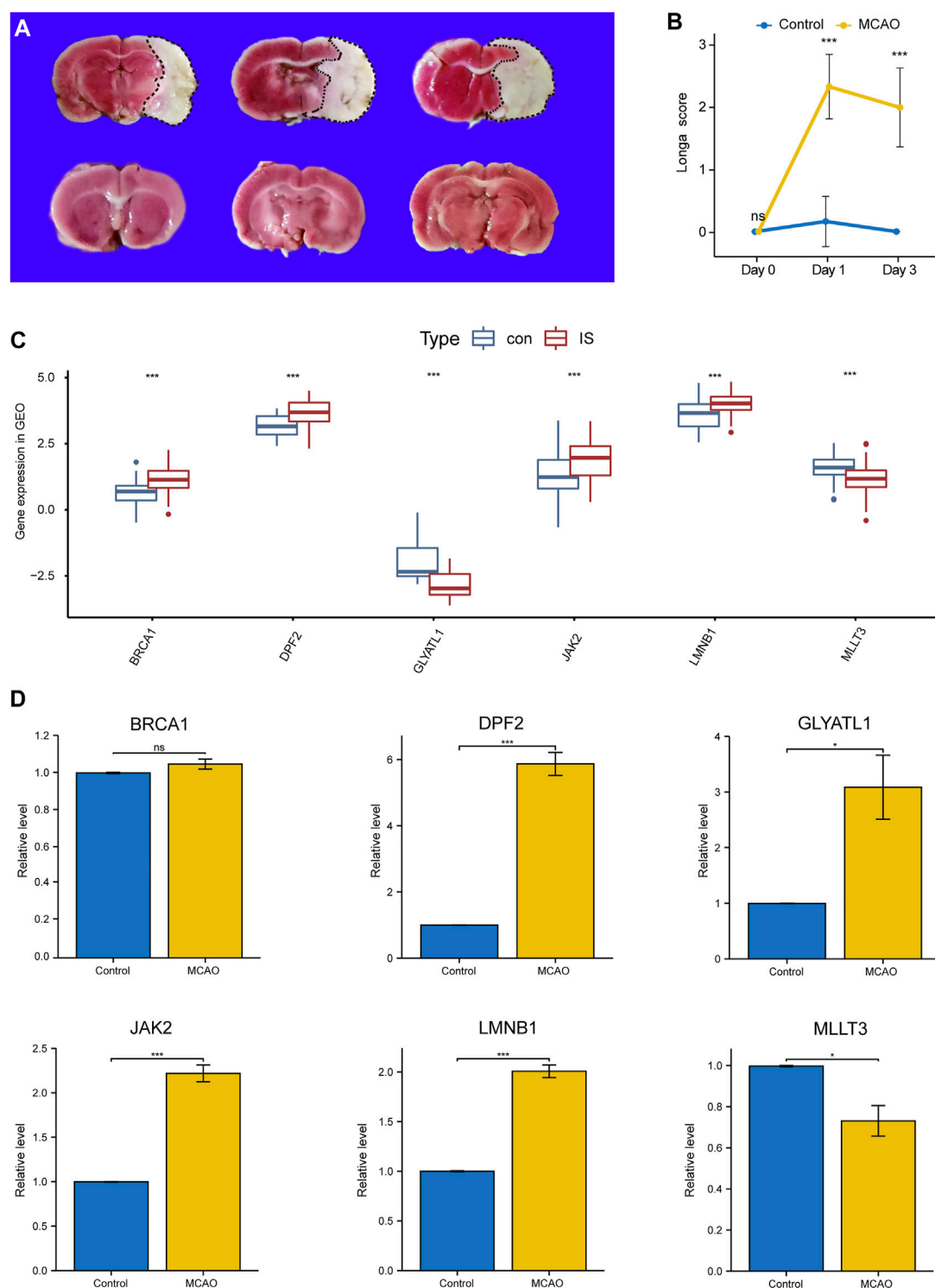
The median expression levels of *DPF2*, *LMNB1*, *MLLT3*, and *JAK2* were 3.58, 3.93, 1.27, and 1.74, respectively, which were the baseline levels in the multivariate logistic regression analysis. Subsequently, the “rms” package was used to construct a nomogram model for predicting the risk for IS, as shown in Figure 6A. A ROC curve was used to evaluate the internal validation results of the dataset, and its AUC was 0.775 (Figure 6B). Furthermore, good agreement between the estimated values and actual observations was found with the calibration curve.

Potential therapeutic compounds for IS

Using the DSigDB database, 34 small-molecule compounds that may bind to *LMNB1*, *MLLT3*, and *JAK2* were identified. Accordingly, the protein structures of *LMNB1*, *MLLT3*, and *JAK2* in the PDB database were investigated. Following the calculation of binding energies, the top 10% of the small-molecule compounds with the lowest binding energies were retained. Among them, *JAK2* may be combined with mevastatin and WP1066 (Figure 7A,B), *LMNB1* may be combined with cladribine and trichostatin A (Figure 7C,D), and *MLLT3* may be combined with mequitazine and zuclophene (Figure 7E,F).

Discussion

In the complex pathogenesis of IS, lymphoid organs are activated, followed by stroke-induced immunosuppression to reduce inflammatory damage (Lindsay et al., 2019; Westendorp et al., 2022). From an epigenetic perspective,

**FIGURE 5**

Validation of Key Immune biomarkers by MCAO models. **(A)** TTC staining in Sham group and IS group. **(B)** Longa scores in Sham group and IS group. **(C)** Expression of CRs-related immune biomarkers in GSE58294 and GSE22255 datasets. **(D)** qPCR verifies the expression of CRs-related immune biomarkers.

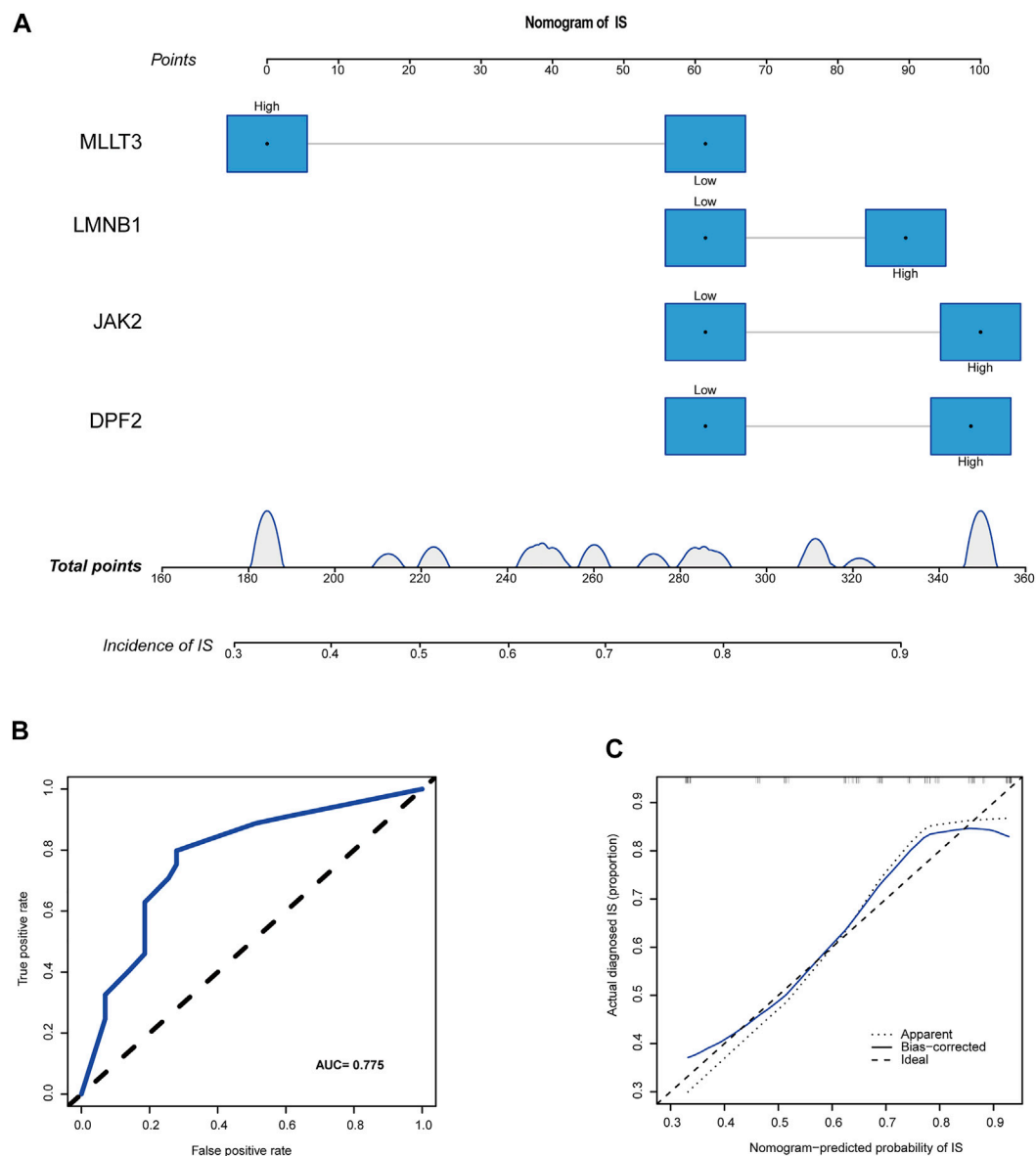


FIGURE 6

A novel nomogram for predicting IS. (A) CRs-related immune biomarkers predict the occurrence of IS. (B) ROC curve for nomogram. (C) The calibration curve for nomogram.

CR-related biomarkers could help shed light on the mechanisms of central nervous system and immune system dysregulation in IS. In this study, we explored four CR-related immune biomarkers associated with IS. To our knowledge, this was the first study to use WGCNA to screen and identify key immune therapeutics target based on CRs in patients who experienced IS. We verified our conclusions using PCR with the MCAO model. Our study provides a rationale and promising research recommendations for the possible epigenetic mechanism(s), treatment, and prognosis of this lethal disease.

Epidemiological data suggest that post-stroke infection is the leading cause of death and disability among patients, and is associated with higher rates of recurrence and readmission (Zhang et al., 2021). The down-regulation of immune system function after IS reduces inflammation on the one hand, but on the other, conversely increases the risk for infection (Qin et al., 2020). Animal experiments have also shown that immune cell therapy can produce significant beneficial effects by improving infarct size and neurological scores in animal models. Consistently, the GO and KEGG results of DEGs in our study indicated that IS was closely associated with the activation of

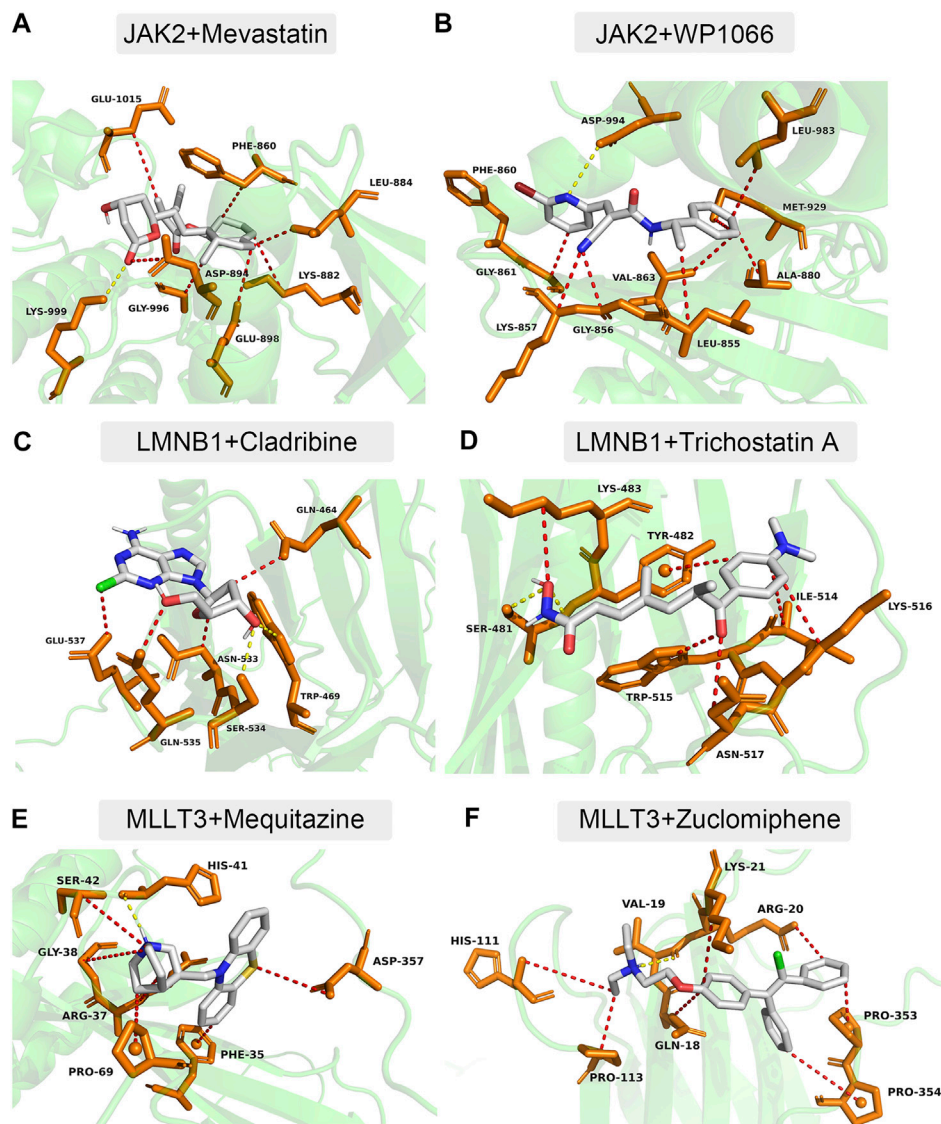


FIGURE 7

Docking simulation of proteins and small molecule compounds. (A) *JAK2* and Mevastatin. (B) *JAK2* and WP1066. (C) *LMNB1* and Cladribine. (D) *LMNB1* and Trichostatin A. (E) *MLLT3* and Mequitazine. (F) *MLLT3* and Zuclophene.

immune response, immune response-regulating signaling pathway, and primary immunodeficiency.

Given the close association between IS and the immune system, we further dissected immune infiltration of the disease using ssGSEA. According to analysis of the immune cell infiltration landscape, the first step of an inflammatory response following IS is the activation of resident microglia, followed by the infiltration of peripheral immune cells such as neutrophils and macrophages. These variations promote neuroinflammation and tissue repair after ischemia. Nathan (2006) showed that brain damage caused by IS was induced by neutrophil-mediated oxidative stress and the release of

proteolytic enzymes. In addition, our findings also revealed immunosuppression after IS, which is embodied in the downregulation of B cells, NK cells, plasmacytoid DC cells, T-helper cells, follicular helper T cells, and tumor-infiltrating lymphocytes. Extensive literature suggests that immunosuppression is associated with the post-stroke-induced activation of the sympathetic and parasympathetic nervous systems (Haspula and Clark, 2018; Zera and Buckwalter, 2020).

After validation by qPCR, we screened four biomarkers (*DPF2*, *LMNB1*, *MLLT3*, and *JAK2*) that were most relevant to immunity. These four biomarkers demonstrated good predictive power in the clinical risk model (AUC, 0.775), and

which may bring assistance for IS patients with negative MRI imaging. Using a rat model of cerebral ischemia, Wang et al. reported that the *Jak2* inhibitor AG490 could improve neurological deficits, cerebral infarction, edema, oxidative stress, and inflammation (Wang et al., 2021). A recent study also identified *Jak2* as an immune-related gene involved in IS pathophysiology (Wang et al., 2022). However, the roles of *DPF2*, *LMNB1*, and *MLLT3* in IS remain to be explored.

Another vital finding from our study was that several potential therapeutic compounds for the treatment of IS were identified. *JAK2*, *LMNB1*, *MLLT3* and can bind to mevastatin, WP1066, cladribine, trichostatin A, mequitazine, and zuclopenthixol. Clinical practice and animal experiments further validated our predictions of immunomodulatory drugs. Some studies have reported that mevastatin, a HMG-CoA reductase inhibitor, can reduce infarction damage (Amin-Hanjani et al., 2001). WP1066 also demonstrated potential to ameliorate ischemic brain injury in rats (Yu et al., 2020). A study by Lingling et al. reported that trichostatin A exerts neuroprotective effects by improving autophagy/lysosomal dysfunction in neurons (Lingling et al., 2022).

It is important to note, however, that the present study had some inherent limitations. First, it was based on bioinformatics analysis and animal experiments, and the results still need to be validated in rat whole blood and clinical IS patients. Second, the sample size of IS patients included in the study was limited. Although the risk model constructed in this study demonstrated good performance, it remains necessary to integrate more sample data to improve the stability of the model. This study also suffered from the inherent drawback of confounding bias in time and space, which include race, region, and time period of IS patients. Finally, the development of immunomodulatory drugs represents a feasible treatment method for IS; however, this needs to be verified in animal experiments, which in turn can be applied to clinical settings.

Collectively, we identified four candidate genes as potential therapeutics target for IS according to bioinformatics analysis and qPCR, and further explored immunomodulatory drugs that may bind to these immune genes related to CRs. Results of the present study contribute to the discovery of CR-related immune therapeutics target in IS and provide an important reference for further research.

References

- Amin-Hanjani, S., Stagliano, N. E., Yamada, M., Huang, P. L., Liao, J. K., and Moskowitz, M. A. (2001). Mevastatin, an HMG-CoA reductase inhibitor, reduces stroke damage and upregulates endothelial nitric oxide synthase in mice. *Stroke* 32 (4), 980–986. doi:10.1161/01.str.32.4.980
- Belk, J. A., Yao, W., Ly, N., Freitas, K. A., Chen, Y. T., Shi, Q., and, et al. (2022). Genome-wide CRISPR screens of T cell exhaustion identify chromatin remodeling factors that limit T cell persistence. *Cancer Cell*. 40, 768–786.e7. doi:10.1016/j.ccell.2022.06.001

Data availability statement

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

Ethics statement

The animal study was reviewed and approved by Xian Jiaotong University Medical College.

Author contributions

BY and YT: manuscript preparation, data analysis, and the research conception. SG and JL: manuscript revision. YZ and BL: data analysis.

Funding

This work was supported by the Construction and application of big data self-evolution remote diagnosis and treatment system in ICU for severe neurological diseases (Grant No.2022ZDLSF04-01) and the National Natural Science Foundation of China (Grant No. 81903268).

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.

- Campbell, B. C. V., De Silva, D. A., Macleod, M. R., Coutts, S. B., Schwamm, L. H., Davis, S. M., et al. (2019). Ischaemic stroke. *Nat. Rev. Dis. Prim.* 5 (1), 70. doi:10.1038/s41572-019-0118-8

- Cao, S., Li, J., Zhang, J., and Li, H. (2021b). Development and validation of a prognostic nomogram for predicting the overall survival of myxofibrosarcoma patients: A large population-based study. *Transl. Cancer Res.* 10 (2), 923–937. doi:10.21037/tcr-20-2588

- Cao, S., Liu, H., Fan, J., Yang, K., Yang, B., Wang, J., et al. (2021a). An oxidative stress-related gene pair (CCNB1/PKD1), competitive endogenous RNAs, and immune-infiltration patterns potentially regulate intervertebral disc degeneration development. *Front. Immunol.* 12, 765382. doi:10.3389/fimmu.2021.765382
- Haspula, D., and Clark, M. A. (2018). Neuroinflammation and sympathetic overactivity: Mechanisms and implications in hypertension. *Auton. Neurosci.* 210, 10–17. doi:10.1016/j.autneu.2018.01.002
- Krug, T., Gabriel, J. P., Taipa, R., Fonseca, B. V., Domingues-Montanari, S., Fernandez-Cadenas, I., and et al. (2012). TTC7B emerges as a novel risk factor for ischemic stroke through the convergence of several genome-wide approaches. *J. Cereb. Blood Flow. Metab.* 32 (6), 1061–1072. doi:10.1038/jcbfm.2012.24
- Li, P., Shi, D. P., Jin, T., Tang, D., Wang, W., and Wang, L. H. (2022). MTA1 aggravates experimental colitis in mice by promoting transcription factor HIF1A and up-regulating AQP4 expression. *Cell. Death Discov.* 8 (1), 298. doi:10.1038/s41420-022-01052-y
- Lindsay, M. P., Norrving, B., Sacco, R. L., Brainin, M., Hacke, W., Martins, S., et al. (2019). World stroke organization (WSO): Global stroke fact sheet 2019. *Int. J. Stroke* 14 (8), 806–817. doi:10.1177/1747493019881353
- Linglart, L., and Bonnet, D. (2022). Epigenetics and congenital heart diseases. *J. Cardiovasc. Dev. Dis.* 9 (6), 185. doi:10.3390/jcdd9060185
- Lingling, D., Miaomiao, Q., Yili, L., Hongyun, H., and Yihao, D. (2022). Attenuation of histone H4 lysine 16 acetylation (H4K16ac) elicits a neuroprotection against ischemic stroke by alleviating the autophagic/lysosomal dysfunction in neurons at the penumbra. *Brain Res. Bull.* 184, 24–33. doi:10.1016/j.brainresbull.2022.03.013
- Longa, E. Z., Weinstein, P. R., Carlson, S., and Cummins, R. (1989). Reversible middle cerebral artery occlusion without craniectomy in rats. *Stroke* 20 (1), 84–91. doi:10.1161/01.str.20.1.84
- Lu, J., Xu, J., Li, J., Pan, T., Bai, J., Wang, L., et al. (2018). FACER: Comprehensive molecular and functional characterization of epigenetic chromatin regulators. *Nucleic Acids Res.* 46 (19), 10019–10033. doi:10.1093/nar/gky679
- Nathan, C. (2006). Neutrophils and immunity: Challenges and opportunities. *Nat. Rev. Immunol.* 6 (3), 173–182. doi:10.1038/nri1785
- Qin, X., Akter, F., Qin, L., Cheng, J., Guo, M., Yao, S., and et al. (2020). Adaptive immunity regulation and cerebral ischemia. *Front. Immunol.* 11, 689. doi:10.3389/fimmu.2020.00689
- Rochmah, T. N., Rahmawati, I. T., Dahlui, M., Budiarto, W., and Bilqis, N. (2021). Economic burden of stroke disease: A systematic review. *Int. J. Environ. Res. Public Health* 18 (14), 7552. doi:10.3390/ijerph18147552
- Stamova, B., Jickling, G. C., Ander, B. P., Zhan, X., Liu, D., Turner, R., and et al. (2014). Gene expression in peripheral immune cells following cardioembolic stroke is sexually dimorphic. *PLoS One* 9 (7), e102550. doi:10.1371/journal.pone.0102550
- Strong, K., Mathers, C., and Bonita, R. (2007). Preventing stroke: Saving lives around the world. *Lancet. Neurol.* 6 (2), 182–187. doi:10.1016/s1474-4422(07)70031-5
- Wang, X., Wang, Q., Wang, K., Ni, Q., Li, H., Su, Z., et al. (2022). Is immune suppression involved in the ischemic stroke? A study based on computational biology. *Front. Aging Neurosci.* 14, 830494. doi:10.3389/fnagi.2022.830494
- Wang, Y. Y., Lin, S. Y., Chang, C. Y., Wu, C. C., Chen, W. Y., Liao, S. L., et al. (2021). Jak2 inhibitor AG490 improved poststroke central and peripheral inflammation and metabolic abnormalities in a rat model of ischemic stroke. *Antioxidants (Basel)* 10 (12), 1958. doi:10.3390/antiox10121958
- Westendorp, W. F., Dames, C., Nederkoorn, P. J., and Meisel, A. (2022). Immunodepression, infections, and functional outcome in ischemic stroke. *Stroke* 53 (5), 1438–1448. doi:10.1161/strokeaha.122.038867
- Yoo, M., Shin, J., Kim, J., Ryall, K. A., Lee, K., Lee, S., and et al. (2015). DSigDB: Drug signatures database for gene set analysis. *Bioinformatics* 31 (18), 3069–3071. doi:10.1093/bioinformatics/btv313
- Yu, L., Liu, Z., He, W., Chen, H., Lai, Z., Duan, Y., et al. (2020). Hydroxysafflor yellow A confers neuroprotection from focal cerebral ischemia by modulating the crosstalk between JAK2/STAT3 and SOCS3 signaling pathways. *Cell. Mol. Neurobiol.* 40 (8), 1271–1281. doi:10.1007/s10571-020-00812-7
- Zera, K. A., and Buckwalter, M. S. (2020). The local and peripheral immune responses to stroke: Implications for therapeutic development. *Neurotherapeutics* 17 (2), 414–435. doi:10.1007/s13311-020-00844-3
- Zhang, S. R., Phan, T. G., and Sobey, C. G. (2021). Targeting the immune system for ischemic stroke. *Trends Pharmacol. Sci.* 42 (2), 96–105. doi:10.1016/j.tips.2020.11.010



OPEN ACCESS

EDITED BY

Clévia Rosset,
Clinical Hospital of Porto Alegre, Brazil

REVIEWED BY

Jiaxin Zhou,
Sun Yat-sen University, China
Mariana Recamonde-Mendoza,
Federal University of Rio Grande do
Sul, Brazil

*CORRESPONDENCE

Gang Lv
cocolv2022@126.com
Zhongyuan Xia
xiazhongyuan2005@aliyun.com

†These authors have contributed
equally to this work

SPECIALTY SECTION

This article was submitted to
Neurogenetics,
a section of the journal
Frontiers in Neuroscience

RECEIVED 19 June 2022

ACCEPTED 31 August 2022

PUBLISHED 20 September 2022

CITATION

Wang W, Huo P, Zhang L, Lv G and
Xia Z (2022) Decoding
competitive endogenous RNA
regulatory network in postoperative
cognitive dysfunction.
Front. Neurosci. 16:972918.
doi: 10.3389/fnins.2022.972918

COPYRIGHT

© 2022 Wang, Huo, Zhang, Lv and Xia.
This is an open-access article
distributed under the terms of the
[Creative Commons Attribution License](#)
(CC BY). The use, distribution or
reproduction in other forums is
permitted, provided the original
author(s) and the copyright owner(s)
are credited and that the original
publication in this journal is cited, in
accordance with accepted academic
practice. No use, distribution or
reproduction is permitted which does
not comply with these terms.

Decoding competitive endogenous RNA regulatory network in postoperative cognitive dysfunction

Wei Wang^{1†}, Pengwei Huo^{2†}, Lei Zhang¹, Gang Lv^{1*} and
Zhongyuan Xia^{1*}

¹Department of Anesthesiology, Renmin Hospital of Wuhan University, Wuhan, China, ²Department of Anesthesiology, Yulin No.2 Hospital, Yulin, China

Postoperative cognitive dysfunction (POCD) is a common postoperative neurological complication in elderly patients. Circular RNAs (circRNAs) are abundant in the mammalian brain and can probably regulate cognitive function. However, the competitive endogenous RNA (ceRNA) regulatory network in POCD remains illiterate. Transcriptomic signatures in the hippocampus of POCD mice derived from the Gene Expression Omnibus (GEO) dataset GSE190880, GSE95070, and GSE115440 were used to identify the circRNA, miRNA, and mRNA expression profiles of POCD mice compared with controls, respectively. A set of differentially expressed RNAs, including 119 circRNAs, 33 miRNAs, and 49 mRNAs were identified. Transcript validation showed the enhanced expression of circ_0001634, circ_0001345, and circ_0001493. A ceRNA regulatory network composed of three circRNAs, three miRNAs, and six mRNAs was established. The hub mRNAs in the ceRNA network were further found to be involved in the hormone catabolic process and regulation of canonical Wnt signaling pathway, revealing their crucial role in POCD. Finally, three miRNAs and four mRNAs were verified by qRT-PCR. These results based on bioinformatics and PCR array suggest that circ_0001634/miR-490-5p/Rbm47, circ_0001634/miR-490-5p/Sostdc1, circ_0001634/miR-7001-5p/Sostdc1, circ_0001345/miR-7001-5p/Sostdc1, and circ_0001493/miR-7001-5p/Sostdc1 may be novel diagnostic biomarkers and therapeutic targets for POCD.

KEYWORDS

postoperative cognitive dysfunction, competitive endogenous RNA network, circRNA, Wnt signaling, bioinformatic analysis

Introduction

Postoperative cognitive dysfunction (POCD) is a neurocognitive disorder such as acute or persistent impairments in attention, learning, memory, and information processing that occurs predominantly in geriatric patients who undergo anesthesia and major surgery (Suwanabol et al., 2022). The long-term impact of POCD is associated with a worse overall quality of life, increased mortality, and a heavy burden on society and families (Boone et al., 2020; Tang et al., 2020). Recent evidence suggests that various elements, such as aberrant expression of apolipoprotein E4 genotype, blood-brain barrier (BBB) compromise, surgery-induced neuroinflammation, microglial activation, mitophagy impairment, and iron accumulation (McDonagh et al., 2010; Chen et al., 2020; Danielson et al., 2020; Wu et al., 2020; Yang et al., 2020), are intimately involved in all stages of postoperative cognitive decline. Despite major advances in comprehending pathogenesis and prevention methods, the overall curative effect remains inadequate. Consequently, it is imperative to screen promising therapeutic targets for POCD.

Circular RNA (circRNA), unlike conventional linear RNA, is a subtype of endogenous non-coding RNAs with covalently closed-loop structures that confer high resistance to degradation (Chen, 2020). CircRNAs have been proposed to modulate gene transcription, to act as microRNA (miRNA) or RNA-binding protein (RBP) decoys, and to function as protein scaffolds (Kristensen et al., 2019). However, the overwhelming majority of circRNAs are thought to act as competing endogenous RNAs (ceRNAs), which specify that circRNAs can compete with mRNAs for binding to the shared miRNAs and thereby indirectly modify gene expression at the post-transcriptional level (Chen, 2020). The abundance of circRNAs in the mammalian brain, particularly synaptoneurosomes, is highlighted by RNA sequencing data (Rybak-Wolf et al., 2015). Mounting evidence also suggests circRNAs play a vital role in cognitive impairments such as Alzheimer impisease (AD) (Dube et al., 2019). Downregulation of circCwc27 improved cognitive capacity in the AD mouse (Song et al., 2022). Knockdown of circTshz2-2 alleviated obesity-induced spatial memory decline via modulating BDNF/TrkB signaling pathway (Yoon et al., 2021). CircPtk2 contributed to sepsis-provoked cognitive dysfunction by serving as a sponge of miR-181c-5p to facilitate HMGB1 expression (Li et al., 2021). Niu et al. (2021) proposed that aerobic exercise mitigated vascular cognitive impairment by activating circRIMS2/miR-186/BDNF axis. Additionally, plasma circRNA-089763 is positively correlated with the occurrence of POCD (Wang et al., 2019; Zhou et al., 2020). A microarray analysis screened 210 differentially expressed circRNAs in POCD patients' serum, such as circCPNE1, circUBE3B, and circITSN1 (Gao et al., 2020). Another microarray profiling highlighted three

circRNAs (circ_22058, circ_44122, and circ_22673) as key elements in POCD (Wu et al., 2021). Circ_009789- and circ_004229-associated ceRNA networks were identified to elucidate the mechanism underlying susceptibility to POCD in aged mice (Zhang et al., 2022). CircShank3 may be involved in dexmedetomidine-mediated protection against POCD via targeting the p53 and NF- κ B signaling pathways (Cao et al., 2020). Nevertheless, the role of circRNAs in POCD remains elusive.

MiRNAs are small (18–22 nucleotide long) non-coding RNAs that trigger post-transcriptional repression of gene expression by directly binding to the 3'-untranslated region (UTR) of mRNAs, profoundly governing neurodevelopment and neurodegeneration (Singh and Yadav, 2020). Multiple aberrantly expressed miRNAs, such as miR-124, miR-146a, and miR-381, were implicated in the development of POCD (Chen et al., 2019a,b; Wang et al., 2021). However, the underlying mechanisms of miRNAs in the neuropathogenesis of POCD are still unclear.

In this study, we extracted the differentially expressed circRNAs, miRNAs, and mRNAs in POCD from Gene Expression Omnibus (GEO) databases. According to the flowchart diagram (Figure 1), circRNA-miRNA pairs and miRNA-mRNA pairs were successively identified for constructing a ceRNA regulatory network. Furthermore, functional enrichment analysis and transcript validation were implemented to identify the hub genes and interpret potential regulatory mechanisms in the development of POCD. The present findings may strengthen our understanding of POCD.

Materials and methods

Expression microarray data

The circRNA dataset (GSE190880) was retrieved from the GPL21826 platform and contained 3 pairs of hippocampus from POCD and control mice (Ran et al., 2022). The miRNA dataset (GSE95070) was retrieved from the GPL19117 platform and contained 5 pairs of hippocampus from POCD and control mice (Wei et al., 2017). Meanwhile, we also downloaded the mRNA expression profiles from GSE115440 (GPL11533) (Yang et al., 2019), including 3 pairs of hippocampus tissues from POCD and control mice, to further construct the ceRNA network about POCD. The profiles of these three microarray datasets (GSE190880, GSE95070, and GSE115440) regarding circRNAs, miRNAs, and mRNAs were listed in Table 1. In these three microarray datasets, aseptic tibia fracture in C57BL/6 male mice was applied to mimic POCD.

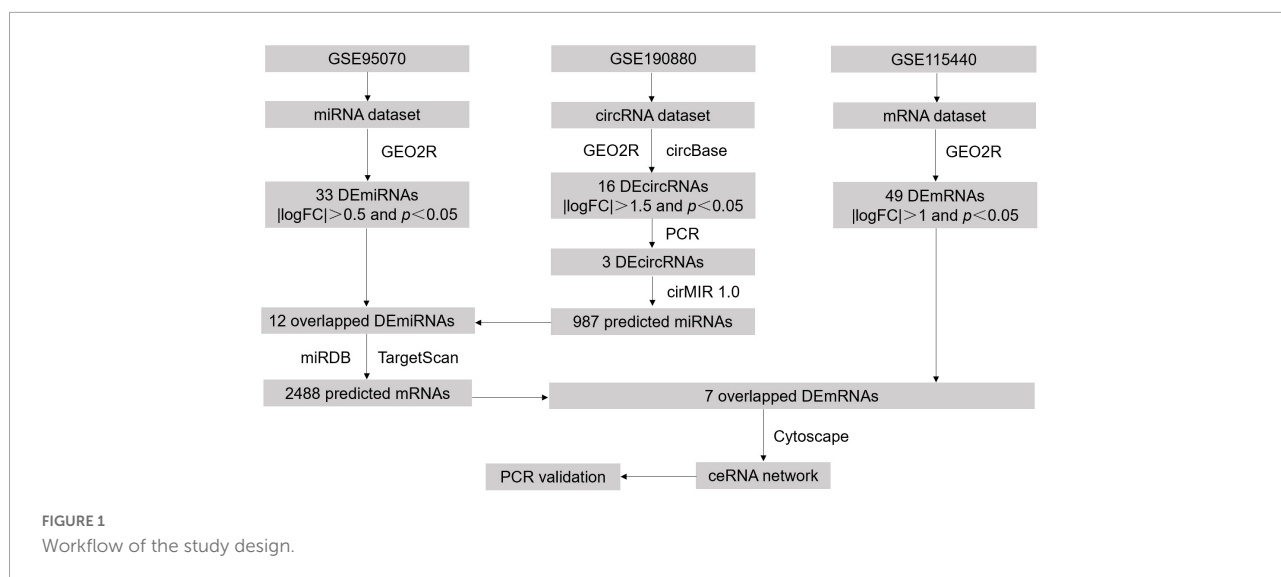


TABLE 1 The profiles of three microarray datasets from the GEO database.

| Data source | Series | Platform | Author | Year | Country | Sample (POCD/control) |
|-------------|-----------|----------|-----------|------|---------|-----------------------|
| circRNA | GSE190880 | GPL21826 | Ran | 2021 | China | 3/3 |
| miRNA | GSE95070 | GPL19117 | Wei | 2017 | China | 5/5 |
| mRNA | GSE115440 | GPL11533 | Mkrtchian | 2018 | Sweden | 3/3 |

Identification of the differentially expressed circular RNAs, microRNAs, and mRNAs

Differentially expressed (DE) analysis between POCD and control hippocampus samples were performed using GEO2R (Barrett et al., 2013), an online program based on R language that can analyze any GEO series to compare two groups of data. The cut-off standard of DE circRNAs was set to $p < 0.05$ and $|\log FC| > 1.5$, the cut-off standard of DE miRNAs was set to $p < 0.05$ and $|\log FC| > 0.5$, and the cut-off standard of DE mRNAs was set to $p < 0.05$ and $|\log FC| > 1$. Principal component analysis (PCA) and heatmap (linkage method: row clustering; distance measure: euclidean) were visualized using the Xiantao search tool.¹

Establishment of circular RNAs–microRNA–mRNA network

The circRNA identification (ID) of DE circRNAs were first converted into the circRNA ID in the circBase (Glažar et al., 2014),² an online database that integrates thousands

of public circRNAs across five species and their genomic profiles. Only genomic sequences in FASTA format of annotated circRNAs were further analyzed using the circMIR1.0 software, a forecasting instrument of circRNA–miRNA interactions based on miRanda and RNAhybrid that can visualize the binding sites of miRNA adsorption. Overlapping miRNAs between the target miRNAs and DE miRNAs screened in the GSE95070 were considered candidate miRNAs. Moreover, miRDB (Chen and Wang, 2020)³ and Targetscan (McGeary et al., 2019) (version 8.0)⁴ databases were used to predict the target mRNAs of the candidate miRNAs. The Venn diagram drawn in the Xiantao search tool was utilized to overlap the miRNA-forecasted mRNAs and DE mRNAs in the GSE115440. Finally, the ceRNA network was constructed using Cytoscape (Otasek et al., 2019) (version 3.9.1), by integrating circRNA–miRNA pairs and miRNA–mRNA pairs.

Functional classifications and pathway enrichment analysis

The Gene Ontology (GO) project is an ontological annotation resource that describes gene product function based on multiple databases (Ashburner et al., 2000). Kyoto

¹ <https://www.xiantao love/>

² <http://www.circbase.org>

³ <http://www.mirdb.org/>

⁴ <https://www.targetscan.org>

Encyclopedia of Genes and Genomes (KEGG)⁵ is an integrated tool that provides a new perspective on molecular-level functions, diseases and drugs (Kanehisa et al., 2016). The DAVID Knowledgebase⁶ is a bioinformatics data resource based on a single-linkage method for the agglomeration of millions of genes that can be entered into DAVID gene clusters (Huang et al., 2009a,b). GO annotation and KEGG pathway enrichment analysis were implemented using the Xiantao search tool. A *p*-value of <0.05 was set as the cut-off point.

Animals

Twelve-month-old C57BL/6 male mice weighing 25–35 g were purchased from Hubei Provincial Center for Disease Control and Prevention. Mice with *ad libitum* access to food and water were housed in a colony room with a temperature of 22–25°C, a humidity of 50%, and a 12-h light/dark cycle. The animals were acclimated to the environment for 7 days before experimental manipulation. All experimental procedures were approved by the Laboratory Animal Ethics Committee in Renmin Hospital of Wuhan University (No. WDRM-20210709).

Animal model of postoperative cognitive dysfunction

The animals were randomly assigned to two groups: the POCD group (*n* = 6) and control group (*n* = 6). Intramedullary fixation for open tibial fracture under isoflurane anesthesia has been widely used as a POCD model (Yang et al., 2019). Briefly, mice were exposed to 3% isoflurane for the anesthesia induction, followed by 1.5% isoflurane in 100% oxygen for the maintenance. After shaving and disinfecting the animal, a skin incision was made on the lateral tibia, and a 0.3-mm intramedullary fixation pin was inserted into the medullary cavity. Next, an osteotomy was performed at the middle and distal third of the bone, and the wound was closed with 5-0 Vicryl suture (Ethicon, Somerville, NJ, USA). The entire procedure from the induction to the end of surgery lasted 30 min. A warm pad was utilized to keep the body temperature around 37°C throughout the surgery. Then, 2% lidocaine solution and 1% tetracaine hydrochloride mucilage was applied locally for postoperative analgesia twice daily until 3 days after surgery. The control group received 100% oxygen without surgery or anesthesia.

Open field test

The open field test (OFT) was employed to assess the locomotor activity of the mice 3 days after the surgery. The apparatus was an opaque plastic cube box with a side length of 45 cm. Each mouse was gently placed in the central area and left to move freely for 5 min. The total distance and time spent in the center area were recorded using SuperMaze software (XinRuan Information Technology, Shanghai, China).

Morris water maze

Spatial learning and memory was assessed using Morris water maze test 4 days after the surgery. The apparatus was a circular plastic pool of 120-cm diameter and 50-cm height surrounded by four curtains. The pool was equipped with a 6-cm-diameter hidden platform placed in the center of the fourth quadrant. Prior to the experiment, the pool was filled with opaque water at around 25°C to a depth of about 30 cm, approximately 1 cm above the platform. Each mouse received four daily trials with a 30-min intertrial interval for four consecutive days to find the platform, and was randomly placed at one of the four quadrants. Mice that failed to find the platform within 90 s were manually directed to the platform. The escape latency (time taken to reach the platform) in the four quadrants on the same day was averaged. The platform was removed on the fifth day to allow for probe trial. During the 90-s session, the total time spent in target quadrant and the frequency of crossing over the target quadrant were recorded. The swimming trajectory of the animals were recorded automatically via a video tracking system (XinRuan Information Technology, Shanghai, China).

RNA extraction and quantitative real-time polymerase chain reaction

Firstly, five circRNAs (circ_0001634, circ_0001345, circ_0001493, circ_0000487, and circ_0001468) whose host genes may be related to cognitive function were selected for quantitative real-time polymerase chain reaction (qRT-PCR). To verify the reliability and accuracy of the predicted ceRNA network, three miRNAs (miR-6912-5p, miR-490-5p, and miR-7001-5p), and four mRNAs (Rbm47, Sostdc1, Cdh3, and Sfrp5) were selected for qRT-PCR. Total RNA was extracted from the hippocampus tissue of the control and POCD group using TRIpure reagent (ELK, Wuhan, China). For circRNAs and mRNAs, cDNA was generated from template RNA using EntiLink™ Reverse Transcriptase Kit (ELK, Wuhan, China), and analyzed with EnTurbo™ SYBR Green PCR SuperMix (ELK, Wuhan, China) using StepOne™ Real-Time PCR system (Life Technologies, Carlsbad, CA, USA). The reaction

⁵ <http://www.kegg.jp>

⁶ <https://david.ncifcrf.gov/>

procedure was as follows: 95°C for 1 min, followed by 40 cycles of 95°C for 15 s, 58°C for 20 s and 72°C for 45 s. For miRNAs, cDNA was generated using EntiLink™ 1st Strand cDNA Synthesis Kit (ELK, Wuhan, China), and analyzed with EnTurbo™ SYBR Green PCR SuperMix (ELK, Wuhan, China) using QuantStudio 6 Flex real-time PCR system (Life Technologies, Carlsbad, CA, USA). The reaction procedure was as follows: 95°C for 30 s, followed by 95°C for 10 s, 58°C for 30 s, and 72°C for 30 s. The relative expression levels of these RNAs were calculated via the $2^{-\Delta\Delta C_t}$ method. GAPDH was used as an internal control gene for circRNA and mRNA expression, while U6 was used as an internal control gene for miRNA expression. The primers for these RNAs are shown as Table 2.

Statistical analysis

Quantitative data were presented as the mean \pm standard deviation (SD) or median (range). GraphPad Prism version 9.0 (GraphPad Software Inc., San Diego, California, USA) was used for statistical analyses. The statistical significance between the two groups was determined using Student's *t*-test or non-parametric test. A *p*-value of <0.05 was considered statistically significant.

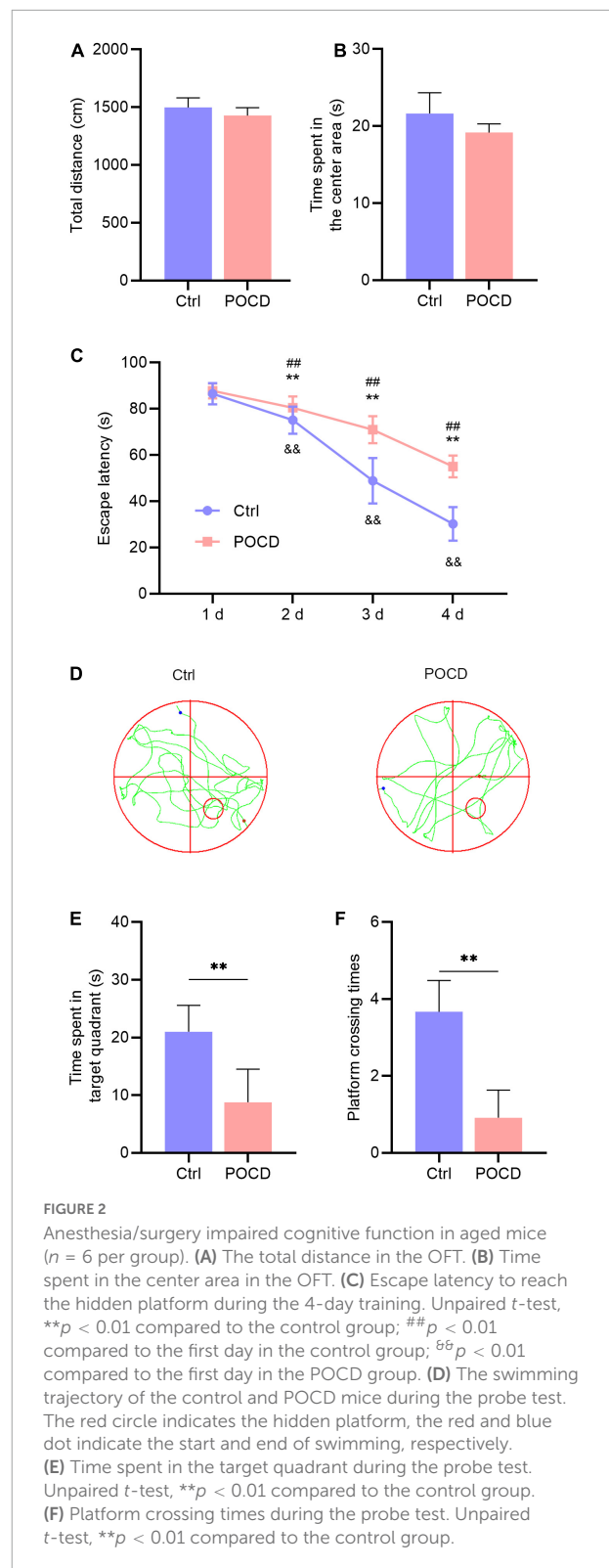
Results

Isoflurane plus orthopedic surgery caused postoperative cognitive dysfunction in mice

The open field test showed no statistical difference in the total distance and duration of the central area between the two groups, indicating that anesthesia/surgery did not affect the locomotor activity of the mice (Figures 2A,B). To address whether isoflurane/surgery impaired cognitive function in mice, we performed Morris water maze to assess learning and memory following orthopedic surgery. During hidden platform test, the mice undergoing anesthesia/surgery showed significantly longer escape latency on day 2, 3, and 4 than the untreated control (Figure 2C). During the probe test, time spent in target quadrant and platform crossing times in the mice undergoing anesthesia/surgery were less than those in the control group (Figures 2D–F). These results suggested that isoflurane plus orthopedic surgery caused POCD.

Data preprocessing

The accuracy and reliability of microarray data were evaluated using PCA. By comparing the distribution patterns of



all identified circRNAs (Figure 3A) and miRNAs (Figure 4A), we could completely separate POCD samples from control hippocampus.

TABLE 2 The primer sequences used for qRT-PCR.

| | Sense | Antisense |
|--------------|---------------------------|---------------------------|
| circ_0001634 | CATGAGCAGTTTCTCTCCAG | GGAGAGTGAGGTCAGTAAACAG |
| circ_0001345 | GACCAAGAGACTGGACGAGT | GGAGAGCTTATTGTCAGAGTGACA |
| circ_0001493 | CAGCAAGCAGACATACCAC | GTGTCGTAGTGGTCTGGAC |
| circ_0000487 | GTGTTCTGACAAAACACCTGAGG | CTGTAATGGTGTCCAGGCAGTAAC |
| circ_0001468 | GCTGACCTCAAACCAGAAAACAT | GTTTCTTCACACTACAGAAGGCA |
| miR-6912-5p | GGCTACAGGGAGGGTGCT | CTCAACTGGTGTCTGGAGTC |
| miR-490-5p | GGCCCCATGGATCTCCA | CTCAACTGGTGTCTGGAGTC |
| miR-7001-5p | GGAGGCAGGGTGTGAGC | CTCAACTGGTGTCTGGAGTC |
| Rbm47 | ACCCAGCTACGTGTACTCTGT | GTTTCATATCCTTTCTCTCTGCTG |
| Sostdc1 | CATTTCAGTAGCACTGGACTGG | GCTCCAGTACTTTGTTCATAGC |
| Cdh3 | ATCAGCTCAAATCTAATAAGGACAG | CCATAAAGCTCGTACTTGACAATCT |
| Sfrp5 | GGGACCGAAAGTTGATTGG | TGAATTTGACTGCAAACTTCATC |
| GAPDH | TGAAGGGTGGAGCCAAAAG | AGTCTTCTGGGTGGCAGTGAT |
| U6 | CTCGCTTCGGCAGCACAT | AACGCTTCACGAATTTCGCT |

Circular RNA expression profiles in postoperative cognitive dysfunction

In the GSE190880 dataset, a total of 119 DEcircRNAs with thresholds of $|\log FC| > 1.5$ and $p < 0.05$ were detected in the hippocampus of POCD mice, of which 103 were up-regulated and 16 were down-regulated. Additionally, 76.47% of DEcircRNAs originated from exonic regions, while 8.40% originated from intronic regions. Sense-overlapping and intergenic circRNAs accounted for 13.45 and 1.68%, respectively (Figure 3B). Chromosomal distribution demonstrated that DEcircRNAs are highly abundant in chr1-19 and chrX, but absent in chr20 and chrY (Figure 3C).

GO and KEGG pathway analyses were used to categorize and annotate the host genes of DEcircRNAs in order to further characterize these in POCD. The enriched GO terms were mainly associated with locomotory behavior (gene ratio = 8/105, $p = 2.06E-05$), vesicle-mediated transport in synapse (gene ratio = 8/105, $p = 1.64E-05$), regulation of synaptic transmission, glutamatergic (gene ratio = 5/105, $p = 2.63E-05$), intrinsic component of synaptic membrane (gene ratio = 9/106, $p = 1.69E-06$), integral component of synaptic membrane (gene ratio = 7/106, $p = 7.67E-05$), Schaffer collateral-CA1 synapse (gene ratio = 6/106, $p = 1.27E-05$), guanyl-nucleotide exchange factor activity (gene ratio = 7/106, $p = 3.92E-05$), calcium ion transmembrane transporter activity (gene ratio = 6/106, $p = 4.49E-05$), and cation:cation antiporter activity (gene ratio = 3/106, $p = 4.49E-05$) (Figure 3D and Supplementary Table 1). Moreover, KEGG pathway enrichment analysis concluded that DEcircRNAs were principally enriched in retrograde endocannabinoid signaling (gene ratio = 5/43, $p = 0.0007$), dilated cardiomyopathy (gene ratio = 4/43, $p = 0.0010$), and glutamatergic synapse (gene ratio = 4/43, $p = 0.0021$) (Figure 3D and Supplementary Table 1). Ulteriorly,

the DEcircRNAs were blasted by circBase and 16 annotated circRNAs were identified, of which 14 were up-regulated and 2 were down-regulated, as shown by the hierarchical clustering heatmap (Figure 3E and Table 3).

Validation of differentially expressed circRNAs

To validate the expression of DEcircRNAs, five circRNAs whose host genes are possibly related to cognitive function were selected for verification by qRT-PCR. The expression of mmu_circ_0001634 (circFam53b), mmu_circ_0001345 (circFbxl5), and circ_0001493 (circIqsec1) were found to be upregulated in the hippocampus of POCD mice, which was consistent with the sequencing results. There was no significant difference in the expression of mmu_circ_0000487 (circXrcc4) between the two groups. Contrary to the sequencing result, isoflurane/surgery obviously diminished mmu_circ_0001468 (circHpk2) expression (Figure 5). To unveil insights into the potential functional mechanisms of circFam53b, circFbxl5, and circIqsec1, we predicted 987 miRNAs interacting with circRNAs and the putative binding domains with the aid of circMIR1.0 software (Figure 4B).

MicroRNA expression profiles in postoperative cognitive dysfunction

A total of 33 significant DEMiRNAs with cut-off criteria of $|\log FC| > 0.5$ and $p < 0.05$ were screened in the hippocampus of POCD mice. We identified 12 overlapping DEMiRNAs by integrating GSE95070 data and predicted results, of which five were upregulated and seven were downregulated

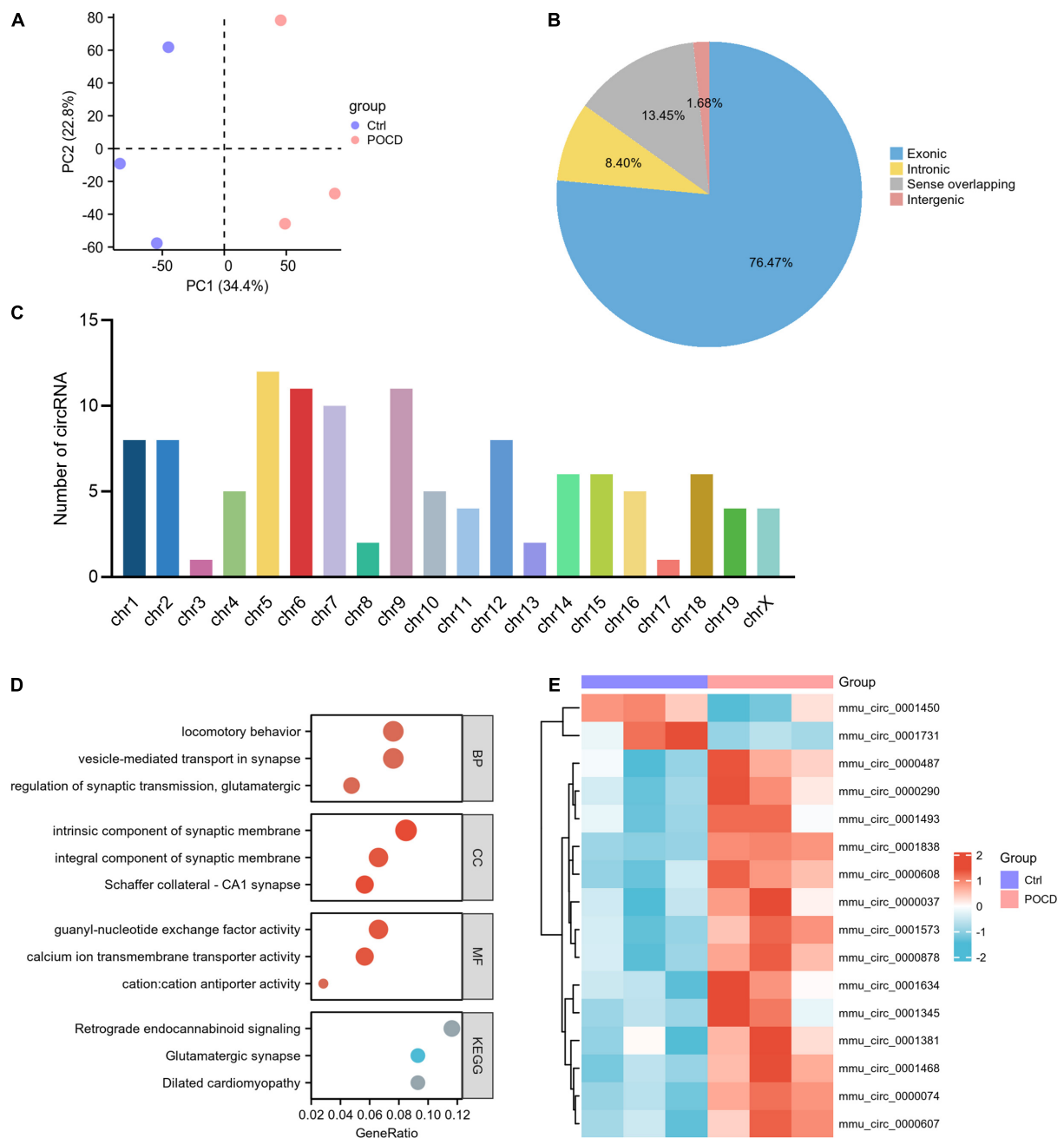


FIGURE 3

Characteristics of differentially expressed circRNAs (DEcircRNAs) in postoperative cognitive dysfunction (POCD). **(A)** Principal component analysis (PCA) of circRNA expression between POCD and control groups. **(B)** Genomic origin of the DEcircRNAs. **(C)** Chromosomal distribution of the DEcircRNAs. **(D)** Functional classifications and pathway enrichment analysis (GO and KEGG) of the host genes of DEcircRNAs. The horizontal axis denotes the proportion of the host genes in each cluster, and the vertical axis denotes biological process (BP), cellular component (CC), molecular function (MF), and KEGG pathway, respectively. **(E)** Heatmap plots of the 16 circRNAs annotated by circBase.

(Figures 4B,C and Supplementary Table 2). According to the ceRNA hypothesis that circRNAs could compete with mRNAs for the same miRNAs, we predicted the mRNAs downstream of 12 miRNAs. The results reflected that 2,488 target genes of the 12 DEMiRNAs were acquired by the databases of miRDB and

TargetScan. Furthermore, a total of 49 significant DEMiRNAs were extracted in GSE115440. The Venn diagram revealed that seven mRNAs were shared by the predicted genes of the above-mentioned circRNA-targeted miRNAs and the DEMiRNAs from GSE115440 (Figure 4D and Supplementary Table 3).

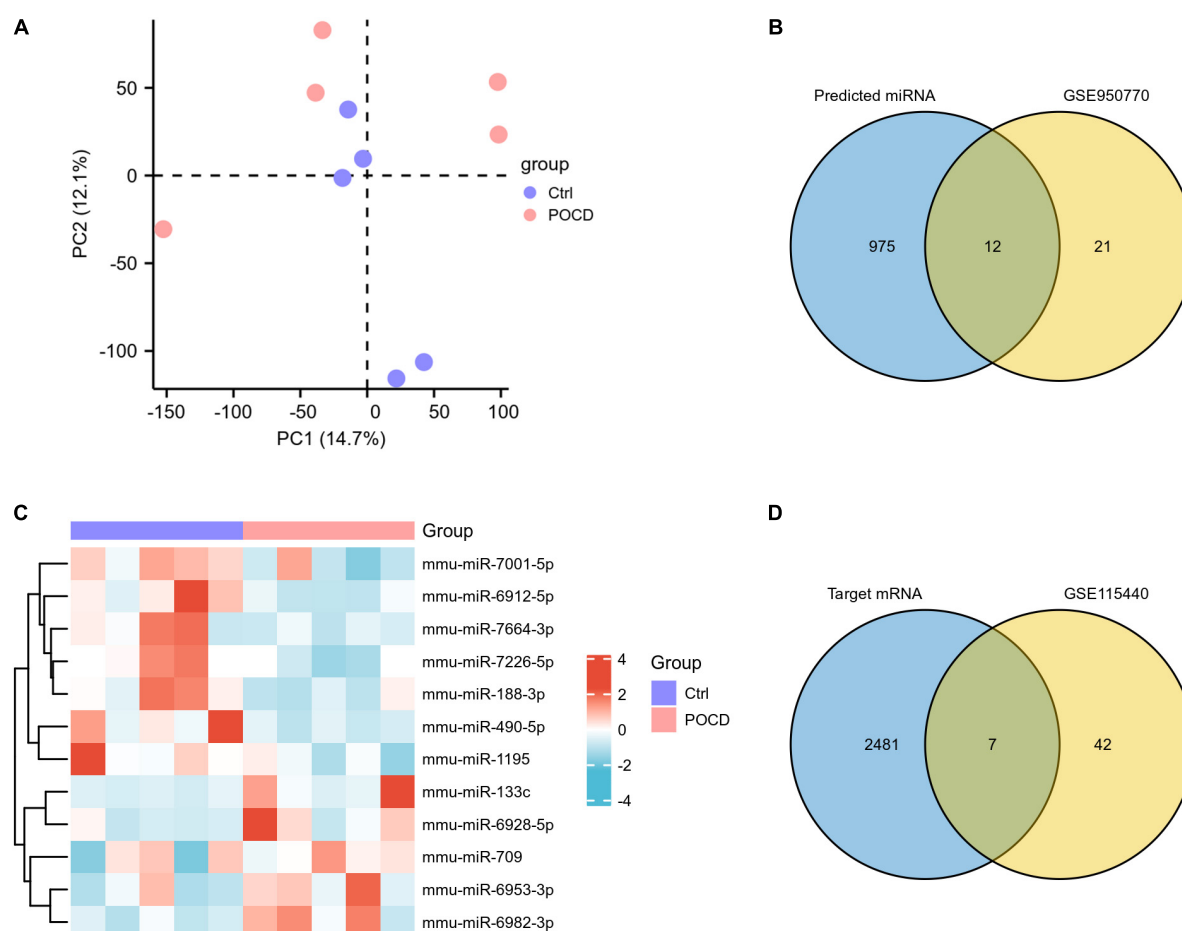


FIGURE 4

Characteristics of differentially expressed miRNAs (DEmiRNAs) in postoperative cognitive dysfunction (POCD). **(A)** Principal component analysis (PCA) of miRNA expression between POCD and control groups. **(B)** The 12 DEmiRNAs were obtained by overlapping the 987 target miRNAs binding to the three DEcircRNAs and the 33 DEmiRNAs identified in GSE950770. **(C)** The heatmap of the 12 overlapped DEmiRNAs. **(D)** The seven DEmRNAs were obtained by overlapping the 2,488 target mRNAs binding to the 12 DEmiRNAs and the 49 DEmRNAs identified in GSE950770.

Construction of the competitive endogenous RNA regulatory network and experimental verification

A ceRNA regulatory network based on 3 circRNAs, 12 miRNAs, and 7 mRNAs was created to further investigate the mechanisms by which circRNAs and miRNAs affect the occurrence and development of POCD. Eventually, 10 circRNA-miRNA pairs and 7 miRNA-mRNA pairs were identified, which were composed of three circRNAs, three miRNAs, and six mRNAs (Figure 6A). The enriched GO terms were chiefly associated with the hormone catabolic process, regulation of canonical Wnt signaling pathway, sperm midpiece, catenin complex, dipeptidyl-peptidase activity, and chloride ion binding. Moreover, the enriched KEGG pathway were renin-angiotensin system and renin secretion (Figure 6B and Supplementary Table 4). To substantiate the potential

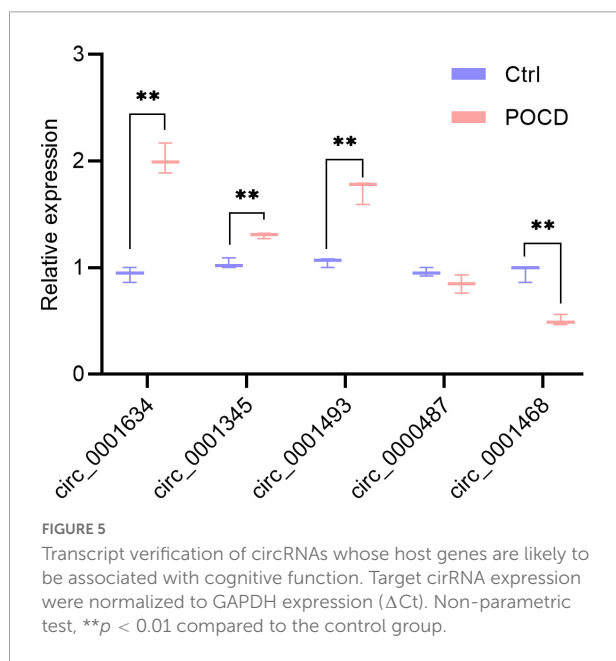
interaction in the ceRNA, these three miRNAs and Wnt-related genes were selected for verification by qRT-PCR. The enhanced expression of miR-6912-5p, Rbm47 and Sostdc1 were observed in the POCD group, while the expression of miR-490-5p, miR-7001-5p, Cdh3, and Sfrp5 were significantly decreased in the POCD group (Figures 6C,D).

Discussion

In our study, we outlined comprehensive transcriptome profiles of POCD. A total of 119 circRNAs, 33 miRNAs, and 49 mRNAs were identified as differentially expressed. The circRNA-miRNA-mRNA triple regulatory network consisted of three circRNAs, three miRNAs, and six mRNAs. Biological process enrichment analysis revealed that the bulk of DEmRNAs in the ceRNA network were involved in the

TABLE 3 Differentially expressed circRNAs annotated by circBase in POCD.

| circRNA_ID | circBase_ID | Chrom | Strand | Location | Type | Gene | LogFC | P-value |
|-------------|--------------|-------|--------|-------------------------|-------------------|---------|--------|---------|
| circ_011823 | circ_0000878 | chr18 | + | 63,755,035–63,760,785 | Exonic | Wdr7 | 1.633 | 0.009 |
| circ_009299 | circ_0000607 | chr15 | – | 68,165,752–68,170,223 | Exonic | Zfat | 1.543 | 0.009 |
| circ_007883 | circ_0001634 | chr7 | – | 132,759,388–132,779,385 | Exonic | Fam53b | 1.596 | 0.032 |
| circ_016800 | circ_0000608 | chr15 | + | 69,013,357–69,029,910 | exonic | Khdrbs3 | 1.526 | 0.005 |
| circ_011539 | circ_0000074 | chr1 | + | 127,791,604–127,799,553 | Exonic | Ccnt2 | 1.546 | 0.001 |
| circ_008286 | circ_0000487 | chr13 | – | 89,991,072–90,001,084 | Exonic | Xrcc4 | 1.506 | 0.028 |
| circ_009389 | circ_0001381 | chr5 | – | 106,619,539–106,666,845 | Exonic | Zfp644 | 1.639 | 0.047 |
| circ_009434 | circ_0000037 | chr1 | – | 52,708,163–52,709,755 | Exonic | Mfsd6 | 1.601 | 0.033 |
| circ_011555 | circ_0001345 | chr5 | – | 43,758,221–43,773,659 | Exonic | Fbxl5 | 1.576 | 0.039 |
| circ_009748 | circ_0001838 | chr9 | – | 106,952,311–106,978,774 | Exonic | Dock3 | 1.541 | <0.001 |
| circ_017841 | circ_0001493 | chr6 | – | 90,689,579–90,694,850 | Exonic | Iqsec1 | 1.506 | 0.039 |
| circ_016934 | circ_0001573 | chr7 | – | 59,479,061–59,481,464 | Sense overlapping | Gm22632 | 2.078 | 0.007 |
| circ_002179 | circ_0001468 | chr6 | – | 38,818,229–38,819,313 | Exonic | Hipk2 | 1.581 | 0.005 |
| circ_016597 | circ_0000290 | chr11 | + | 75,390,071–75,391,227 | Exonic | Smyd4 | 1.560 | 0.018 |
| circ_009489 | circ_0001450 | chr6 | + | 29,372,580–29,372,670 | Intronic | Calu | –1.594 | 0.047 |
| circ_011181 | circ_0001731 | chr8 | – | 122,908,667–122,916,045 | Exonic | Ankrd11 | –1.668 | 0.035 |



“hormone catabolic process” and “regulation of canonical Wnt signaling pathway.” Transcript validation suggested that circ_0001634/miR-490-5p/Rbm47, circ_0001634/miR-490-5p/Sostdc1, circ_0001634/miR-7001-5p/Sostdc1, circ_0001345/miR-7001-5p/Sostdc1, and circ_0001493/miR-7001-5p/Sostdc1 axis were likely to participate in the development of POCD.

The vast majority of circRNAs are derived from a single or multiple exons of known coding genes and thereby located in the cytoplasm (Kristensen et al., 2019). Our analysis showed that

up to 76.47% DEcircRNAs were from exons, and 14 out of 16 annotated circRNAs were of exon origin, which implied that circRNAs probably function as competing endogenous RNAs in the pathogenesis of POCD because the ceRNA network merely exists in the cytoplasm (Chen, 2020). Enrichment analysis showed the corresponding parental or host genes of DEcircRNAs were implicated in synaptic plasticity and mainly located in the Schaffer collateral-CA1 synapse, suggesting circRNAs are extensively involved in the regulation of cognitive function.

Circ_0001493, circ_0000487, circ_0001468, and circ_0001345 are spliced from IQ motif and sec7 domain-containing protein 1 (Iqsec1), x-ray repair cross complementing 4 (Xrcc4), homeodomain-interacting protein kinase 2 (Hipk2), and F-box/LRR repeat protein 5 (Fbxl5), respectively, which may be linked to neurocognitive disorders (Zhang et al., 2013; Gerez et al., 2019; Liang et al., 2020; Briševac et al., 2021). Moreover, circ_0001634 originates through back-splicing events from Fam53b, which has been reported to regulate Wnt signal transduction by altering β -catenin nuclear localization (Kizil et al., 2014). Wnt signaling has been implicated in the modulation of synaptogenesis, long-term potentiation (LTP), and dendrite arborization (Narvaes and Furini, 2022). Intranuclear accumulation of β -catenin marks activation of canonical Wnt signaling that sequentially leads to enhance the transcription of Wnt target genes via the interaction between β -catenin and T-cell factor/lymphoid enhancer factor (TCF/LEF) (Narvaes and Furini, 2022). Hu et al. (2016) found that prolonged exposure to 3.6% sevoflurane could disrupt BBB components via suppressing Wnt/ β -catenin/Annexin A1 pathway in brain microvascular endothelial cells, indicating that promoting β -catenin synthesis can alleviate POCD.

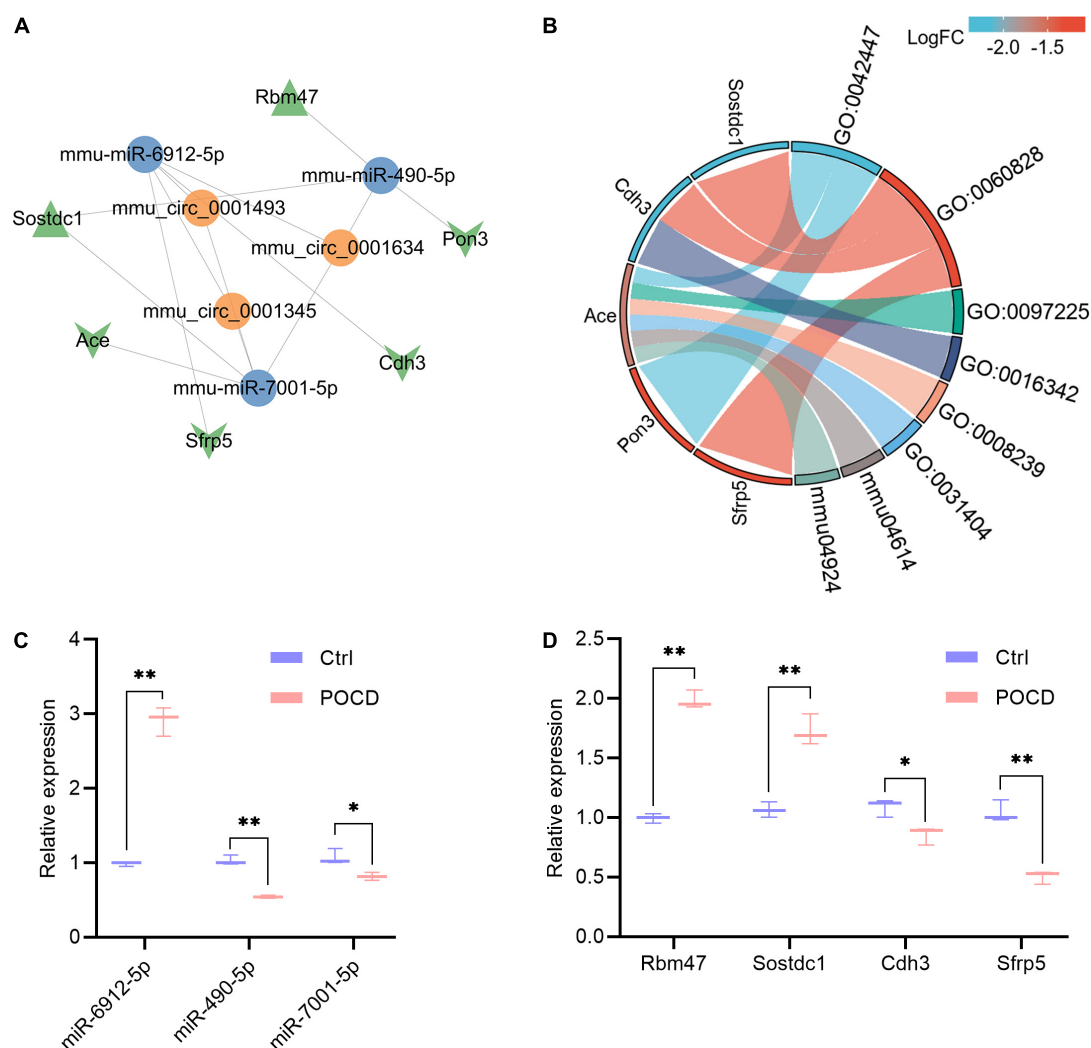


FIGURE 6

Potential competing endogenous RNA (ceRNA) regulatory network, enrichment analysis of DEmRNAs in the ceRNA network, and qRT-PCR validation of three miRNAs and four mRNAs. **(A)** The ceRNA regulatory network includes three circRNAs, three miRNAs, and six mRNAs. The orange color indicates circRNAs, the blue color indicates miRNAs, and the green color indicates mRNAs (Triangle and V denote upregulation and downregulation, respectively). **(B)** Functional classifications and pathway enrichment analysis (GO and KEGG) of DEmRNAs in the ceRNA network. The right half part indicates enriched biological process (BP), cellular component (CC), molecular function (MF), and KEGG pathway; the left half part indicates the genes involved in the corresponding BP, CC, MF, and pathways. GO:0042447, hormone catabolic process; GO:0060828, regulation of canonical Wnt signaling pathway; GO:0097225, sperm midpiece; GO:0016342, catenin complex; GO:0008239, dipeptidyl-peptidase activity; GO:0031404, chloride ion binding; mmu04614, renin-angiotensin system; mmu04924, renin secretion. **(C,D)** Quantitative RT-PCR validation of miR-6912-5p, miR-490-5p, miR-7001-5p, Rbm47, Sostdc1, Cdh3, and Sfrp5. Target miRNA expression were normalized to U6 expression, and mRNA expression were normalized to GAPDH expression (ΔCt). Non-parametric test, * $p < 0.05$, ** $p < 0.01$ compared to the control group.

Hence, these five circRNAs were selected for verification. Quantitative RT-PCR array revealed that the augmented levels of circ_0001634, mmu_circ_0001345, and circ_0001493 were consistent with the sequencing results.

Based on ceRNA hypothesis that a protein-coding RNA and a non-coding RNA compete for the same miRNA through the shared miRNA response elements (MREs), miRNA serves as a bridge for circRNA-induced translation and/or stabilization of the mRNAs (Chen, 2020). In our

study, miR-6912-5p, miR-490-5p, and miR-7001-5p were screened as hub nodes to compete with three circRNAs for governing the expression of six mRNAs. GO classification and enrichment analysis uncovered that cadherin 3 (Cdh3), secreted frizzled related protein 5 (Sfrp5), and sclerostin domain containing 1 (Sostdc1) in the ceRNA network participate in the regulation of the canonical Wnt signaling pathway. PCR analysis showed that anesthesia/surgery increased the expression of miR-6912-5p, RNA binding motif protein 47

(Rbm47), and Sostdc1 in the hippocampus, while decreased the expression of miR-490-5p, miR-7001-5p, Cdh3, and Sfrp5. Due to a positive expression correlation between genes and their corresponding circRNAs, circ_0001634/miR-490-5p/Rbm47, circ_0001634/miR-490-5p/Sostdc1, circ_0001634/miR-7001-5p/Sostdc1, circ_0001345/miR-7001-5p/Sostdc1, and circ_0001493/miR-7001-5p/Sostdc1 axis might be existent in the pathogenesis of POCD. Pei et al. (2022) found that CircFAM53B, highly homologous to mmu_circ_0001634, impeded glioma cell apoptosis through sponging miR-532-3p, suggesting a potential role in neurological disorders. Interestingly, circFbxl5 (homologous with circ_0001345) could function as a miR-146a sponge in mouse cardiomyocyte (Li et al., 2022b), and miR-146a has been proven to ameliorate surgery-induced cognitive decline (Chen et al., 2019a). Accordingly, it is speculated that depletion of circ_0001345 may be an emerging therapeutic perspective on POCD. Most notably, miR-490-5p has been shown to ameliorate stroke-induced neurological dysfunction by repressing cyclin-dependent kinases 1 (CDK1) (Skovira et al., 2016; Ding et al., 2021), revealing a cognition-protective role for miR-490-5p. Rbm47, a central mediator of mRNA alternative splicing and stability, has been elaborated to suppress Wnt activity in cancer cells via maintaining AXIN1 or DKK1 mRNA stability (Vanharanta et al., 2014; Shen et al., 2020). Moreover, Sostdc1 has been documented to negatively modulate Wnt signaling (Li et al., 2022a). These imply that inhibition of Rbm47 or Sostdc1 may mitigate postoperative neurocognitive disorder.

Nonetheless, some limitations of this study must be considered. First, these three datasets are retrieved from different platform and samples, which certainly reduces confidence in the regulator-target pairs outlined. Second, the ideal scenario would be the same logFC threshold and adjusted *p*-value. However, adjusted *p*-value is either absent (GSE190880) or greater than 0.9 (GSE95070 and GSE115440), which may be attributed to the relative small sample size and unknown distribution of their genomic data. There are many DEcircRNAs but few DEmiRNAs in the case of $p < 0.05$ and $|\logFC| > 1.0$. Hence, *p*-value and different logFC thresholds were employed for screening DERNAs, as previously described (Wang et al., 2022). Third, limited RNA ID information of other datasets available such as GSE165798 (Wu et al., 2021) hampers the integration of multiple datasets. Finally, the circRNA-miRNA and miRNA-mRNA interaction relationships in the ceRNA network were both based upon a prediction algorithm that required further experimental verification.

In summary, a ceRNA regulatory network in the hippocampus of POCD mice included three circRNAs, three miRNAs, and six mRNAs. Through comprehensive bioinformatic analysis, the present study broadens our horizon on the occurrence and progression of POCD.

Data availability statement

The original contributions presented in this study are included in the article/Supplementary material, further inquiries can be directed to the corresponding author/s.

Ethics statement

The animal study was reviewed and approved by the Laboratory Animal Ethics Committee in Renmin Hospital of Wuhan University.

Author contributions

WW, PH, LZ, GL, and ZX designed the study. WW and PH collected and analyzed the data. WW drafted the manuscript. GL and ZX undertook a critical revision of the manuscript. All authors contributed to the article and approved the submitted version.

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/fnins.2022.972918/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, 25–29. doi: 10.1038/75556
- 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
- Boone, M. D., Sites, B., Von Recklinghausen, F. M., Mueller, A., Taenzer, A. H., and Shaefi, S. (2020). Economic burden of postoperative neurocognitive disorders among US medicare patients. *JAMA Network Open* 3:e208931. doi: 10.1001/jamanetworkopen.2020.8931
- Briševac, D., Scholz, R., Du, D., Elagabani, M. N., Köhr, G., and Kornau, H. C. (2021). The small GTPase Arf6 is dysregulated in a mouse model for fragile X syndrome. *J. Neurochem.* 157, 666–683. doi: 10.1111/jnc.15230
- Cao, C., Deng, F., and Hu, Y. (2020). Dexmedetomidine alleviates postoperative cognitive dysfunction through circular RNA in aged rats. *3 Biotech* 10:176. doi: 10.1007/s13205-020-2163-0
- Chen, L., Dong, R., Lu, Y., Zhou, Y., Li, K., Zhang, Z., et al. (2019a). MicroRNA-146a protects against cognitive decline induced by surgical trauma by suppressing hippocampal neuroinflammation in mice. *Brain Behav. Immun.* 78, 188–201. doi: 10.1016/j.bbi.2019.01.020
- Chen, Y., Sun, J. X., Chen, W. K., Wu, G. C., Wang, Y. Q., Zhu, K. Y., et al. (2019b). miR-124/VAMP3 is a novel therapeutic target for mitigation of surgical trauma-induced microglial activation. *Signal Transduct. Target Ther.* 4:27. doi: 10.1038/s41392-019-0061-x
- Chen, L. L. (2020). The expanding regulatory mechanisms and cellular functions of circular RNAs. *Nat. Rev. Mol. Cell Biol.* 21, 475–490. doi: 10.1038/s41580-020-0243-y
- Chen, Y., and Wang, X. (2020). miRDB: an online database for prediction of functional microRNA targets. *Nucleic Acids Res.* 48, D127–D131. doi: 10.1093/nar/gkz757
- Chen, Y., Zhang, P., Lin, X., Zhang, H., Miao, J., Zhou, Y., et al. (2020). Mitophagy impairment is involved in sevoflurane-induced cognitive dysfunction in aged rats. *Aging* 12, 17235–17256. doi: 10.18632/aging.103673
- Danielson, M., Wiklund, A., Granath, F., Blennow, K., Mkrtychian, S., Nellgård, B., et al. (2020). Neuroinflammatory markers associate with cognitive decline after major surgery: findings of an explorative study. *Ann. Neurol.* 87, 370–382. doi: 10.1002/ana.25678
- Ding, L., Ning, J., Wang, Q., Lu, B., and Ke, H. (2021). Sevoflurane improves nerve regeneration and repair of neurological deficit in brain damage rats via microRNA-490-5p/CDK1 axis. *Life Sci.* 271:119111. doi: 10.1016/j.lfs.2021.119111
- Dube, U., Del-Aguila, J. L., Li, Z., Budde, J. P., Jiang, S., Hsu, S., et al. (2019). An atlas of cortical circular RNA expression in Alzheimer disease brains demonstrates clinical and pathological associations. *Nat. Neurosci.* 22, 1903–1912. doi: 10.1038/s41593-019-0501-5
- Gao, R., Chen, C., Zhao, Q., Li, M., Wang, Q., Zhou, L., et al. (2020). Identification of the Potential Key Circular RNAs in Elderly Patients With Postoperative Cognitive Dysfunction. *Front. Aging Neurosci.* 12:165. doi: 10.3389/fnagi.2020.00165
- Gerez, J. A., Prymaczek, N. C., Rockenstein, E., Herrmann, U. S., Schwarz, P., Adame, A., et al. (2019). A cullin-RING ubiquitin ligase targets exogenous α -synuclein and inhibits Lewy body-like pathology. *Sci. Transl. Med.* 11:eaa6722. doi: 10.1126/scitranslmed.aau6722
- Glažar, P., Papavasileiou, P., and Rajewsky, N. (2014). circBase: a database for circular RNAs. *RNA* 20, 1666–1670. doi: 10.1261/rna.043687.113
- Hu, N., Wang, C., Zheng, Y., Ao, J., Zhang, C., Xie, K., et al. (2016). The role of the Wnt/ β -catenin-Annexin A1 pathway in the process of sevoflurane-induced cognitive dysfunction. *J. Neurochem.* 137, 240–252. doi: 10.1111/jnc.13569
- Huang da, W., Sherman, B. T., and Lempicki, R. A. (2009a). Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res.* 37, 1–13. doi: 10.1093/nar/gkn923
- Huang da, W., Sherman, B. T., and Lempicki, R. A. (2009b). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* 4, 44–57. doi: 10.1038/nprot.2008.211
- Kanehisa, M., Sato, Y., and Morishima, K. (2016). BlastKOALA and GhostKOALA: KEGG tools for functional characterization of genome and metagenome sequences. *J. Mol. Biol.* 428, 726–731. doi: 10.1016/j.jmb.2015.11.006
- Kizil, C., Küchler, B., Yan, J. J., Özhan, G., Moro, E., Argenton, F., et al. (2014). Simplex/Fam53b is required for Wnt signal transduction by regulating β -catenin nuclear localization. *Development* 141, 3529–3539. doi: 10.1242/dev.108415
- Kristensen, L. S., Andersen, M. S., Stagsted, L. V. W., Ebbesen, K. K., Hansen, T. B., and Kjems, J. (2019). The biogenesis, biology and characterization of circular RNAs. *Nat. Rev. Genet.* 20, 675–691. doi: 10.1038/s41576-019-0158-7
- Li, D., You, J., Mao, C., Zhou, E., Han, Z., Zhang, J., et al. (2022b). Circular RNA Fbx15 Regulates Cardiomyocyte Apoptosis During Ischemia Reperfusion Injury via Sponging microRNA-146a. *J. Inflamm. Res.* 15, 2539–2550. doi: 10.2147/jir.S360129
- Li, C., Wang, M., Shi, Y., and Xin, H. (2022a). SOSTDC1 acts as a tumor inhibitor in acute myeloid leukemia by downregulating the Wnt/ β -catenin pathway. *Environ. Toxicol.* 37, 1934–1943. doi: 10.1002/tox.23540
- Li, M., Hu, J., Peng, Y., Li, J., and Ren, R. (2021). CircPTK2-miR-181c-5p-HMGB1: a new regulatory pathway for microglia activation and hippocampal neuronal apoptosis induced by sepsis. *Mol. Med.* 27:45. doi: 10.1186/s10020-021-00305-3
- Liang, L., Xie, R., Lu, R., Ma, R., Wang, X., Wang, F., et al. (2020). Involvement of homodomain interacting protein kinase 2-c-Jun N-terminal kinase/c-Jun cascade in the long-term synaptic toxicity and cognition impairment induced by neonatal Sevoflurane exposure. *J. Neurochem.* 154, 372–388. doi: 10.1111/jnc.14910
- McDonagh, D. L., Mathew, J. P., White, W. D., Phillips-Bute, B., Laskowitz, D. T., Podgoreanu, M. V., et al. (2010). Cognitive function after major noncardiac surgery, apolipoprotein E4 genotype, and biomarkers of brain injury. *Anesthesiology* 112, 852–859. doi: 10.1097/ALN.0b013e3181d31fd7
- 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:eaav1741. doi: 10.1126/science.aav1741
- Narvaez, R. F., and Furini, C. R. G. (2022). Role of Wnt signaling in synaptic plasticity and memory. *Neurobiol. Learn. Mem.* 187:107558. doi: 10.1016/j.nlm.2021.107558
- Niu, Y., Wan, C., Zhang, J., Zhang, S., Zhao, Z., Zhu, L., et al. (2021). Aerobic exercise improves VCI through circRIMS2/miR-186/BDNF-mediated neuronal apoptosis. *Mol. Med.* 27:4. doi: 10.1186/s10020-020-00258-z
- Otasek, D., Morris, J. H., Bouças, J., Pico, A. R., and Demchak, B. (2019). Cytoscape automation: empowering workflow-based network analysis. *Genome Biol.* 20:185. doi: 10.1186/s13059-019-1758-4
- Pei, J., Dou, H., and Deng, X. (2022). CircFAM53B promotes the proliferation and metastasis of glioma through activating the c-MET/PI3K/AKT pathway via sponging miR-532-3p. *Cell Cycle* 21, 462–476. doi: 10.1080/15384101.2021.2014738
- Ran, W., Liang, N., Yuan, R., Wang, Z., and Gao, J. (2022). Identification of potential key circRNAs in aged mice with postoperative delirium. *Front. Mol. Neurosci.* 15:836534. doi: 10.3389/fnmol.2022.836534
- Rybak-Wolf, A., Stottmeister, C., Glažar, P., Jens, M., Pino, N., Giusti, S., et al. (2015). Circular RNAs in the mammalian brain are highly abundant, conserved, and dynamically expressed. *Mol. Cell* 58, 870–885. doi: 10.1016/j.molcel.2015.03.027
- Shen, D. J., Jiang, Y. H., Li, J. Q., Xu, L. W., and Tao, K. Y. (2020). The RNA-binding protein RBM47 inhibits non-small cell lung carcinoma metastasis through modulation of AXIN1 mRNA stability and Wnt/ β -catenin signaling. *Surg. Oncol.* 34, 31–39. doi: 10.1016/j.suronc.2020.02.011
- Singh, T., and Yadav, S. (2020). Role of microRNAs in neurodegeneration induced by environmental neurotoxins and aging. *Ageing Res. Rev.* 60:101068. doi: 10.1016/j.arr.2020.101068
- Skovira, J. W., Wu, J., Matyas, J. J., Kumar, A., Hanscom, M., Kabadi, S. V., et al. (2016). Cell cycle inhibition reduces inflammatory responses, neuronal loss, and cognitive deficits induced by hypobaric exposure following traumatic brain injury. *J. Neuroinflammation* 13:299. doi: 10.1186/s12974-016-0769-2
- Song, C., Zhang, Y., Huang, W., Shi, J., Huang, Q., Jiang, M., et al. (2022). Circular RNA Cwc27 contributes to Alzheimer's disease pathogenesis by repressing Pur- α activity. *Cell Death Differ.* 29, 393–406. doi: 10.1038/s41418-021-00865-1
- Suwanabol, P. A., Li, Y., Abrahamse, P., De Roo, A. C., Vu, J. V., Silveira, M. J., et al. (2022). Functional and cognitive decline among older adults after high-risk surgery. *Ann. Surg.* 275, e132–e139. doi: 10.1097/sla.0000000000003950
- Tang, V. L., Jing, B., Boscardin, J., Ngo, S., Silvestrini, M., Finlayson, E., et al. (2020). Association of functional, cognitive, and psychological measures with 1-year mortality in patients undergoing major surgery. *JAMA Surg.* 155, 412–418. doi: 10.1001/jamasurg.2020.0091
- Vanharanta, S., Marney, C. B., Shu, W., Valiente, M., Zou, Y., Mele, A., et al. (2014). Loss of the multifunctional RNA-binding protein RBM47 as a source of

selectable metastatic traits in breast cancer. *eLife* 3:e02734. doi: 10.7554/eLife.02734

Wang, J., Niu, Y., Luo, L., Lu, Z., Chen, Q., Zhang, S., et al. (2022). Decoding ceRNA regulatory network in the pulmonary artery of hypoxia-induced pulmonary hypertension (HPH) rat model. *Cell Biosci.* 12:27. doi: 10.1186/s13578-022-00762-1

Wang, M., Su, P., Liu, Y., Zhang, X., Yan, J., An, X., et al. (2019). Abnormal expression of circRNA_089763 in the plasma exosomes of patients with postoperative cognitive dysfunction after coronary artery bypass grafting. *Mol. Med. Rep.* 20, 2549–2562. doi: 10.3892/mmr.2019.10521

Wang, Y. L., Zhang, Y., and Cai, D. S. (2021). Dexmedetomidine ameliorates postoperative cognitive dysfunction via the MicroRNA-381-Mediated EGR1/p53 Axis. *Mol. Neurobiol.* 58, 5052–5066. doi: 10.1007/s12035-021-02417-7

Wei, C., Luo, T., Zou, S., Zhou, X., Shen, W., Ji, X., et al. (2017). Differentially expressed lncRNAs and miRNAs with associated ceRNA networks in aged mice with postoperative cognitive dysfunction. *Oncotarget* 8, 55901–55914. doi: 10.18632/oncotarget.18362

Wu, J., Yang, J. J., Cao, Y., Li, H., Zhao, H., Yang, S., et al. (2020). Iron overload contributes to general anaesthesia-induced neurotoxicity and cognitive deficits. *J. Neuroinflammation* 17:110. doi: 10.1186/s12974-020-01777-6

Wu, Y. Q., Liu, Q., Wang, H. B., Chen, C., Huang, H., Sun, Y. M., et al. (2021). Microarray analysis identifies key differentially expressed circular RNAs in aged mice with postoperative cognitive dysfunction. *Front. Aging Neurosci.* 13:716383. doi: 10.3389/fnagi.2021.716383

Yang, T., Velagapudi, R., and Terrando, N. (2020). Neuroinflammation after surgery: from mechanisms to therapeutic targets. *Nat. Immunol.* 21, 1319–1326. doi: 10.1038/s41590-020-00812-1

Yang, T., Xu, G., Newton, P. T., Chagin, A. S., Mkrtchian, S., Carlström, M., et al. (2019). Maresin 1 attenuates neuroinflammation in a mouse model of perioperative neurocognitive disorders. *Br. J. Anaesth.* 122, 350–360. doi: 10.1016/j.bja.2018.10.062

Yoon, G., Lim, Y. H., Jo, D., Ryu, J., Song, J., and Kim, Y. K. (2021). Obesity-linked circular RNA circTshz2-2 regulates the neuronal cell cycle and spatial memory in the brain. *Mol. Psychiatry* 26, 6350–6364. doi: 10.1038/s41380-021-01303-x

Zhang, L. Y., Chen, L. S., Sun, R., Ji, S. J., Ding, Y. Y., Wu, J., et al. (2013). Effects of expression level of DNA repair-related genes involved in the NHEJ pathway on radiation-induced cognitive impairment. *J. Radiat. Res.* 54, 235–242. doi: 10.1093/jrr/rrs095

Zhang, M. X., Lin, J. R., Yang, S. T., Zou, J., Xue, Y., Feng, C. Z., et al. (2022). Characterization of circRNA-Associated-ceRNA networks involved in the pathogenesis of postoperative cognitive dysfunction in aging mice. *Front. Aging Neurosci.* 14:727805. doi: 10.3389/fnagi.2022.727805

Zhou, H., Li, F., Ye, W., Wang, M., Zhou, X., Feng, J., et al. (2020). Correlation between plasma CircRNA-089763 and postoperative cognitive dysfunction in elderly patients undergoing non-cardiac surgery. *Front. Behav. Neurosci.* 14:587715. doi: 10.3389/fnbeh.2020.587715



OPEN ACCESS

EDITED BY

Thayne Kowalski,
Centro Universitário Cesuca, Brazil

REVIEWED BY

Ba Nguyen,
University of Hawaii at Mānoa,
United States
Michael Sierk,
National Cancer Institute at Frederick
(NIH), United States

*CORRESPONDENCE

Yonghao Yu,
yyu@tmu.edu.cn

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

SPECIALTY SECTION

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

RECEIVED 01 August 2022

ACCEPTED 16 September 2022

PUBLISHED 03 October 2022

CITATION

Zhang X, Wang Y, Dong B, Jiang Y, Liu D,
Xie K and Yu Y (2022), Expression pattern
and clinical value of Key RNA
methylation modification regulators in
ischemic stroke.
Front. Genet. 13:1009145.
doi: 10.3389/fgene.2022.1009145

COPYRIGHT

© 2022 Zhang, Wang, Dong, Jiang, Liu,
Xie and Yu. This is an open-access
article distributed under the terms of the
[Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/)
(CC BY). The use, distribution or
reproduction in other forums is
permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original
publication in this journal is cited, in
accordance with accepted academic
practice. No use, distribution or
reproduction is permitted which does
not comply with these terms.

Expression pattern and clinical value of Key RNA methylation modification regulators in ischemic stroke

Xinyue Zhang^{1,2†}, Yuanlin Wang^{1,2}, Beibei Dong^{1,2}, Yi Jiang^{1,2},
Dan Liu³, Keliang Xie^{1,2,4} and Yonghao Yu^{1,2*}

¹Department of Anesthesiology, Tianjin Medical University General Hospital, Tianjin, China, ²Tianjin Institute of Anesthesiology, Tianjin, China, ³School of Medicine, Nankai University, Tianjin, China, ⁴Department of Critical Care Medicine, Tianjin Medical University General Hospital, Tianjin, China

Ischemic stroke (IS) is one of the major causes of death and disability worldwide, and effective diagnosis and treatment methods are lacking. RNA methylation, a common epigenetic modification, plays an important role in disease progression. However, little is known about the role of RNA methylation modification in the regulation of IS. The aim of this study was to investigate RNA methylation modification patterns and immune infiltration characteristics in IS through bioinformatics analysis. We downloaded gene expression profiles of control and IS model rat brain tissues from the Gene Expression Omnibus database. IS profiles were divided into two subtypes based on RNA methylation regulators, and functional enrichment analyses were conducted to determine the differentially expressed genes (DEGs) between the subtypes. Weighted gene co-expression network analysis was used to explore co-expression modules and genes based on DEGs. The IS clinical diagnosis model was successfully constructed and four IS characteristic genes (*GFAP*, *GPNMB*, *FKBP9*, and *CHMP5*) were identified, which were significantly upregulated in IS samples. Characteristic genes were verified by receiver operating characteristic curve and real-time quantitative PCR analyses. The correlation between characteristic genes and infiltrating immune cells was determined by correlation analysis. Furthermore, *GPNMB* was screened using the protein-protein interaction network, and its regulatory network and the potential therapeutic drug chloroquine were predicted. Our finding describes the expression pattern and clinical value of key RNA methylation modification regulators in IS and novel diagnostic and therapeutic targets of IS from a new perspective.

KEYWORDS

characteristic gene, epigenetics, immune infiltration, ischemic stroke, RNA methylation modification

Introduction

Ischemic stroke (IS) is a serious cerebrovascular disease characterized by a high disability rate and mortality, imposing a massive burden on society (Zhou et al., 2019). Current evidence shows that the focus of IS treatment is emergency intervention and long-term secondary prevention (Herpich and Rincon, 2020). However, owing to the narrow therapeutic window and hemorrhage-related complications, the clinical treatment options for IS are very limited and only a minority of patients benefit (Henderson et al., 2018). Therefore, effective diagnostic biomarkers and treatments are urgently needed to improve early diagnosis, reduce mortality, and improve prognosis of IS.

Several studies have been performed to improve the understanding of the molecular mechanisms of IS based on microarray and bioinformatics analysis. A previous study isolated 10 hub genes and five key miRNAs between IS and normal control groups by analyzing two datasets (GSE58294 and GSE16561) (Yang et al., 2022). Li et al. (2020) studied IS from the perspective of immune regulation and identified immune-related gene expression modules and hub genes in the peripheral blood of patients with IS, which might become important targets for immunotherapy of IS. However, these studies merely identified differentially expressed genes (DEGs), without exploring the detailed molecular mechanisms and potential drug molecules.

As a research hotspot in recent years, the post-transcriptional chemical modification of RNA is rapidly emerging as a pivotal player in regulating gene expression. To date, more than 170 types of RNA modifications have been identified that modify coding and noncoding RNAs, which account for more than 50% of methylations (Boccalletto et al., 2022). RNA methylation, an abundant and widely studied epigenetic modification, plays an important role in modulating multiple biological functions (Zhou et al., 2020). The occurrence of RNA methylation is reversible and dynamically regulated by groups of proteins called RNA-modifying proteins, including “writers” (methyltransferases), “erasers” (demethylases), and “readers” (methyl binding proteins) (Zaccara et al., 2019). N1-methyladenosine (m1A), N6-methyladenosine (m6A), and 5-methylcytosine (m5C) are common types of eukaryotic RNA methylation modifications (Xu et al., 2021), among which m6A RNA methylation has been reported to be highly enriched in the mammalian brain and closely associated with the pathological mechanism of IS (Yu et al., 2021). For instance, Xu S. et al. (2020) found that lnc-D63785 m6A methylation leads to the accumulation of miR-422a and neuronal death in an oxygen-glucose deprivation/reperfusion model. Moreover, as one of the m6A “readers,” YTHDC1 has been found to alleviate brain injury through the PTEN/Akt pathway and provide a potential therapeutic target for treating IS (Zhang Z. et al., 2020). However, as new types of RNA methylation, the relationship between m1A- and m5C-related regulators and IS has not been reported, and their mechanisms need to be further explored.

A growing body of research has confirmed that the immune microenvironment plays a vital role in IS (Zera and Buckwalter, 2020; Liu et al., 2021). Following IS, peripheral immune cells migrate through the broken blood-brain barrier to the damaged area and activate host immune cells, such as microglia (Chavda et al., 2021). Infiltrated inflammatory cells and the activated immune response lead to the dysfunction of the immune microenvironment, which dramatically hinders neurological functional recovery (Shi et al., 2022). Further evidence indicates that RNA methylation modifications are involved in immune regulation, especially in the tumor immune microenvironment. For example, m5C regulators have been shown to promote the expression and infiltration of CD8⁺ T cells and are associated with poor prognosis in patients with lung squamous cell carcinoma (Pan et al., 2021). The m6A-binding protein YTHDF1 facilitates tumor immune escape by impairing the cross-presentation of tumor neoantigens and cross-priming of CD8⁺ T cells (Han et al., 2019). However, the role of RNA methylation regulators in immune infiltration in IS have yet to be explored.

In this study, we first comprehensively analyzed the GSE97537 dataset to identify differentially expressed RNA methylation-related regulators (m1A, m6A, and m5C) and evaluate immunocyte infiltration in IS and control samples, and we identified IS-related subtypes. Gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses were performed to identify DEGs between subtypes. Weighted gene co-expression network analysis (WGCNA) was used to identify the co-expressed genes and modules. Then, we constructed a clinical diagnostic model of IS and identified characteristic genes. Protein-protein interaction (PPI) networks, transcription factor (TF) correlation, competing endogenous RNA (ceRNA) networks, and potential drug molecules for IS therapy were identified based on characteristic genes. Our study may provide insight into the role of RNA methylation in pathogenesis and immune infiltration in IS.

Materials and methods

Data acquisition and preprocessing

Gene expression profile data, focusing on ischemic reperfusion, were obtained from the Gene Expression Omnibus database (<https://www.ncbi.nlm.nih.gov/geo/>). In total, 22 samples from GSE97537 (7 IS and five control rat samples) (Wang et al., 2015) and GSE61616 (5 IS and five control rat samples) were selected. The same platform, GPL1355, was used for the two datasets. Detailed information from GSE97537 and GSE61616 is listed in [Supplementary Table S1](#). RNA methylation-related regulators from previous studies, including 11 m1A methylation regulators (Gao et al., 2021),

21 m5C methylation regulators (Chen et al., 2020), and 23 m6A methylation regulators (Zhao et al., 2017; He et al., 2019) were collected. The “Affy” package (Ritchie et al., 2015) was utilized to normalize gene expression values from the two datasets. Next, Log2 transformation was also carried out. Principal component analysis (PCA) was performed to detect the distribution of samples in the two groups. GSE97537 was considered the primary analysis dataset and training set. GSE61616 was chosen as the testing set to check the diagnostic ability of this diagnosis model.

Screen of RNA methylation regulators

The differential expression analysis and visualization of m1A, m5C, and m6A methylation regulators between IS samples and control samples were performed using the “limma” and “pheatmap” packages. The “RCircos” package (Zhang et al., 2013) in R, which can display the chromosomal location of DEGs, was used. Next, the correlation and interaction between DEGs of m1A, m5C, and m6A methylation regulators were calculated based on the Pearson algorithm. Gene interaction networks showing these factors were drawn using the “Corrplot” package.

Infiltration characteristics of the immune microenvironment in IS

CIBERSORTx (<https://cibersort.stanford.edu/>), an R tool for the deconvolution of expression matrices of immune cell subtypes, was designed by combining linear support vector regression and immune infiltration theory (Chen et al., 2018). We used the CIBERSORTx algorithm to profile the landscape of 22 types of immune cells in the immune microenvironments of the IS and normal groups according to gene expression levels in datasets. The correlation among immune cells was considered very influential to understand immune pathway and function. Therefore, the correlation coefficients between immune cells were calculated by Spearman analysis and visualized through heatmaps. Statistical differences in the proportion of infiltrating immunocytes between IS and control groups were calculated by Wilcoxon test using R software (v. 3.5.1).

Identification of IS-related molecular subtype

Consensus clustering was performed using the “ConsensusClusterPlus” package (Wilkerson and Hayes, 2010) to identify IS subgroups based on differentially expressed RNA methylation regulators. By combining consensus cumulative distribution function (CDF) plots, delta area plots, tracking plots, and clustering heatmaps, the optimal number of clusters

was identified. These clusters were defined as IS-related molecular subtypes.

GO and KEGG enrichment analyses

The “limma” package was used to screen DEGs between the subtypes. The cut-off criteria for statistical significance were adjusted p value (P_{adj}) < 0.05 and logFC > 1. GO analysis (Ashburner et al., 2000) is a major bioinformatics tool designed for complex functional enrichment analyses, composed of annotations of biological process (BP), molecular function (MF), and cellular component (CC). KEGG (Kanehisa and Goto, 2000), an integrated database resource, is used to understand high-level functions and utilities of biological systems from genomic and molecular-level information. GO annotation and KEGG pathway enrichment analyses of DEGs were performed using the “clusterProfiler” package (Yu et al., 2012). Results with a false discovery rate < 0.05 were considered statistically significant. The pathway with the highest enrichment of DEGs in KEGG analysis was visualized using the “Pathview” package (Luo and Brouwer, 2013).

Gene set enrichment analysis (GSEA) and gene set variation analysis (GSVA)

GSEA, an analytical method based on the entire gene expression matrix, was conducted to derive the significant differences in biological processes between the IS subtypes. Reference gene sets, “c2. all.v7.5.2. entrez.gmt,” were downloaded from the Molecular Signature Database (Liberzon et al., 2015). P_{adj} < 0.05 and |normalized enrichment score| > 1 were considered to indicate statistical significance. GSVA, a nonparametric unsupervised analysis method, was used to evaluate different pathways enriched in the different samples. In our study, GSVA was performed using the “GSVA” package (Hänzelmann et al., 2013).

WGCNA

WGCNA is a systems biology method that can be used to identify modules of highly correlated genes among different samples and identify candidate biomarkers or potential therapeutic targets based on the association of modules to one another and to phenotype (Yue et al., 2016). The top 1,000 genes in gene expression data of IS samples, which were ranked by median absolute deviation (MAD), were analyzed using the “WGCNA” package (Langfelder and Horvath, 2008). Then, we removed outliers and set an optimal soft threshold. The settings of minModuleSize = 25 and set height = 0.15 were used to obtain the final co-expression modules. Finally, genes in the most important modules were screened.

Construction of a diagnostic model

To identify characteristic genes associated with IS and analyze their diagnostic ability, least absolute shrinkage and selection operator (LASSO) regression was performed. A diagnostic model was constructed based on the training dataset, GSE97537. Further validation of this model was performed on the GSE61616 dataset. Receiver operating characteristic (ROC) curves drawn using the “ROCR” package (Sing et al., 2005) were used to illustrate the diagnostic ability of this model on the test set.

Correlation analysis between characteristic genes and immune cell infiltration

The expression levels of characteristic genes associated with IS and immune cell infiltration score were integrated. Spearman correlation analysis was used to determine the correlation between characteristic genes and immunocyte fractions. Detailed results were displayed as a lollipop plot.

Establishment of an animal model and RT-qPCR

All experimental procedures were approved by the Animal Experimental Ethics Committee of Tianjin Medical University General Hospital. Male C57BL/6J mice (aged 6–8 weeks, 20–25 g) were used to establish the middle cerebral artery occlusion (MCAO) animal model. Specific operations and evaluation methods are detailed in a previous study (Espinosa et al., 2020). Sham-operated mice underwent the same surgical procedures except for the occlusion of the middle cerebral artery. Twenty-4 hours after reperfusion, mice were sacrificed by euthanasia and the cerebral cortex of the lesioned side was removed from the mice. Total RNA was extracted from the cortex using TRIzol reagent (Invitrogen, Carlsbad, CA, United States) and used as a template for reverse transcription into cDNA using a cDNA synthesis kit (Thermo Scientific, Waltham, MA, United States). Then, RT-qPCR amplification was carried out. *GAPDH* was used for normalization. Primer sequences are shown in Supplementary Table S2.

Construction of protein-protein interaction networks and hub gene regulatory networks

PPI networks were constructed in the Search Tool for the Retrieval of Interacting Genes (STRING) online database ([http://](http://string-db.org)

string-db.org; v. 10.5). Visualization was performed in Cytoscape (v. 3.9.0) (Shannon et al., 2003). Maximal clique centrality (MCC) was calculated using CytoHubba (Chin et al., 2014), a Cytoscape plugin. Genes with the highest MCC value were selected as hub genes.

Prediction of ceRNA network

The ENCODE database (<https://www.encodeproject.org/>) (Davis et al., 2018) was used to screen possible TFs of the hub gene. To explore the potential relationship between the hub gene and various noncoding RNAs, we constructed a ceRNA network using Cytoscape. The interaction information between mRNA and miRNA and between miRNA and lncRNA were predicted using miRTarBase (Huang et al., 2022) and StarBase database (Li et al., 2014), respectively.

Construction of drug-gene network and molecular docking

We used the Comparative Toxicogenomics Database (Davis et al., 2021) (<http://ctdbase.org/>) to predict drug molecules that might be useful in the treatment of IS by targeting hub genes. Cytoscape was used to visualize the interaction network between hub genes and drug molecules. According to the targeting relationship and reference scores of these potential components, a potential therapeutic drug was identified. PubChem (<https://pubchem.ncbi.nlm.nih.gov/>) and PDB (<http://www.rcsb.org/>) databases (Burley et al., 2017), which contain three-dimensional structures of small molecules and large-sized proteins, were searched and detailed structures for drug and hub mRNA and proteins were obtained. Autodock (v. 4.2.6) and Pymol (v. 2.3.0) were used to calculate and visualize the results of docking for drugs and mRNA/proteins. To verify docking results, YASARA was utilized through another algorithm (Krieger and Vriend, 2014). I Mutant3.0 was also utilized to identify function of important binding location.

Statistical analysis

All data processing and analyses were completed in R software (v. 4.1.1). Statistical significance between non-normally distributed variables was analyzed using the Mann-Whitney U test (Wilcoxon rank sum test). Correlation coefficients among different genes were calculated using Pearson correlation analysis. Spearman correlation analysis was used to calculate the correlation coefficients between different immune cells and with genes. $p < 0.05$ was considered to indicate statistically significant results.

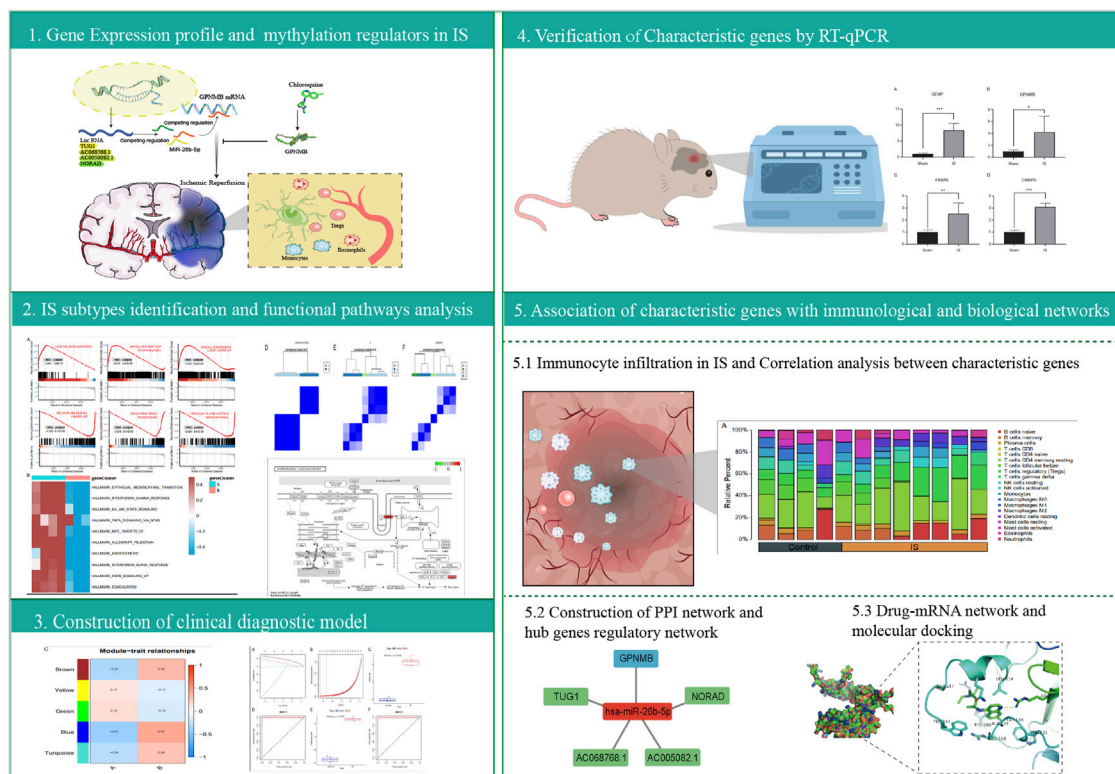


FIGURE 1
Protocol flowchart.

Results

Analysis flow chart and data preprocessing

The analysis flow chart of this study is shown in Figure 1. Gene expression data of the two datasets were normalized to eliminate the batch effect (Supplementary Figures S1A, B). According to the results of PCA, IS samples and control samples were well-classified (Supplementary Figures S1C, D).

Expression profile and chromosomal localization of m1A, m5C, and m6A regulators in IS.

DEGs of m1A, m5C, and m6A regulators (readers, writers, and erasers), with significant differences, were clearly separated in the heatmap according to groups (Supplementary Figure S2A). Detailed chromosomal locations of DEGs (6 m1A regulators, 15 m5C regulators, and 13 m6A regulators) were displayed using a chromosomal circle diagram (Supplementary Figure S2B–D). m1A-related genes were mainly located on chromosomes 1, 3, 8, 11, and 14; m5C-related genes were mainly located on

chromosomes 2, 3, 19, and 18; and m6A-related genes were mainly located on chromosomes X, 17, and 7. The correlation and interaction among DEGs are exhibited in Figures 2A–C. DEGs belonging to m1A, m5A, and m6A could be linked in each network.

Immunocyte infiltration and correlation in IS and control samples

The CIBERSORTx algorithm was used to calculate the abundance ratios of 22 types of immune cells, and cells with abundance ratios of 0 were removed in the subsequent analysis (Figure 3A). The correlation coefficients between immune cells were analyzed. Positive relationships were observed between M0 macrophages and follicular helper T cells and between gamma delta T cells and resting mast cells. Negative relationships were observed between regulatory T cells and follicular helper T cells, M1 macrophages and resting NK cells, and M2 macrophages and M0 macrophages (Figure 3B). Significant differences in the proportion of immune cells between IS and control samples were calculated. Differences in naïve B cells, memory B cells, follicular helper T cells, activated NK

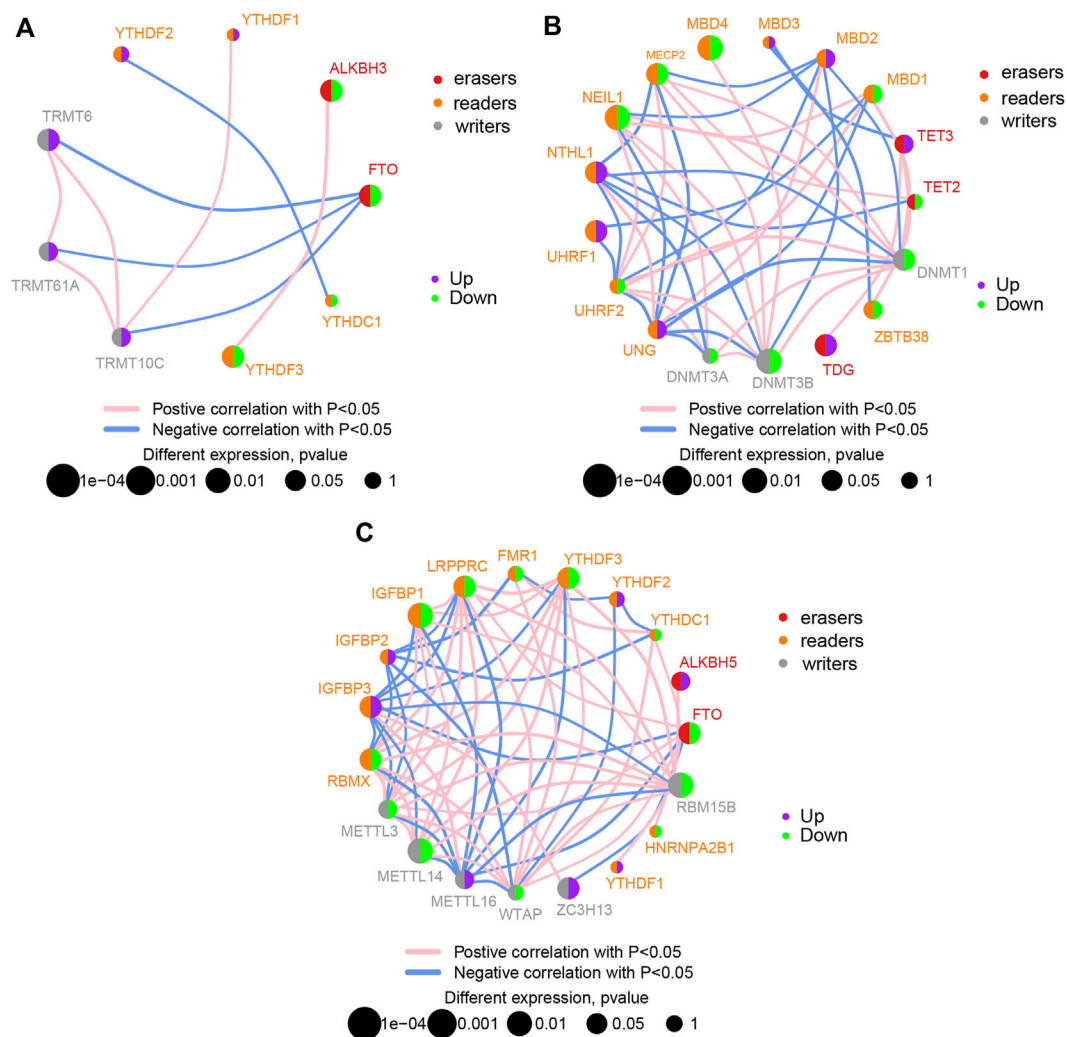


FIGURE 2

Correlation between differentially expressed genes. Correlation network diagram of differentially expressed m1A regulators (A), m5C regulators (B), and m6A regulators (C) in ischemic stroke samples.

cells, M0 macrophages, activated mast cells, neutrophils ($p < 0.05$), and T cells regulatory (Tregs) ($p < 0.01$) between IS and control samples were significant (Figure 3C).

Identification of IS subtypes based on the RNA methylation-related regulators

Gene expression profiles of 34 differentially expressed RNA methylation-related regulators were constructed to investigate the IS molecular subtypes. The CDF plots, delta area plots, and tracking plot (Supplementary Figure S3A–C) were used to assess the appearance of different k values. Three kinds of selection on k ($k = 2, 3, 4$) and the

probable separating subtypes in ConsensusClusterPlus are separately shown in Supplementary Figure S3D–F. The optimal division was reached when $k = 2$, thus, two IS subtypes were identified.

GO functional enrichment analysis and KEGG pathway analysis

GO functional enrichment analysis and KEGG pathway analysis were performed on 28 DEGs between the IS subtypes to obtain more detailed information on their potential functions and correlated pathways. According to the results of GO functional enrichment analysis, the DEGs were mainly

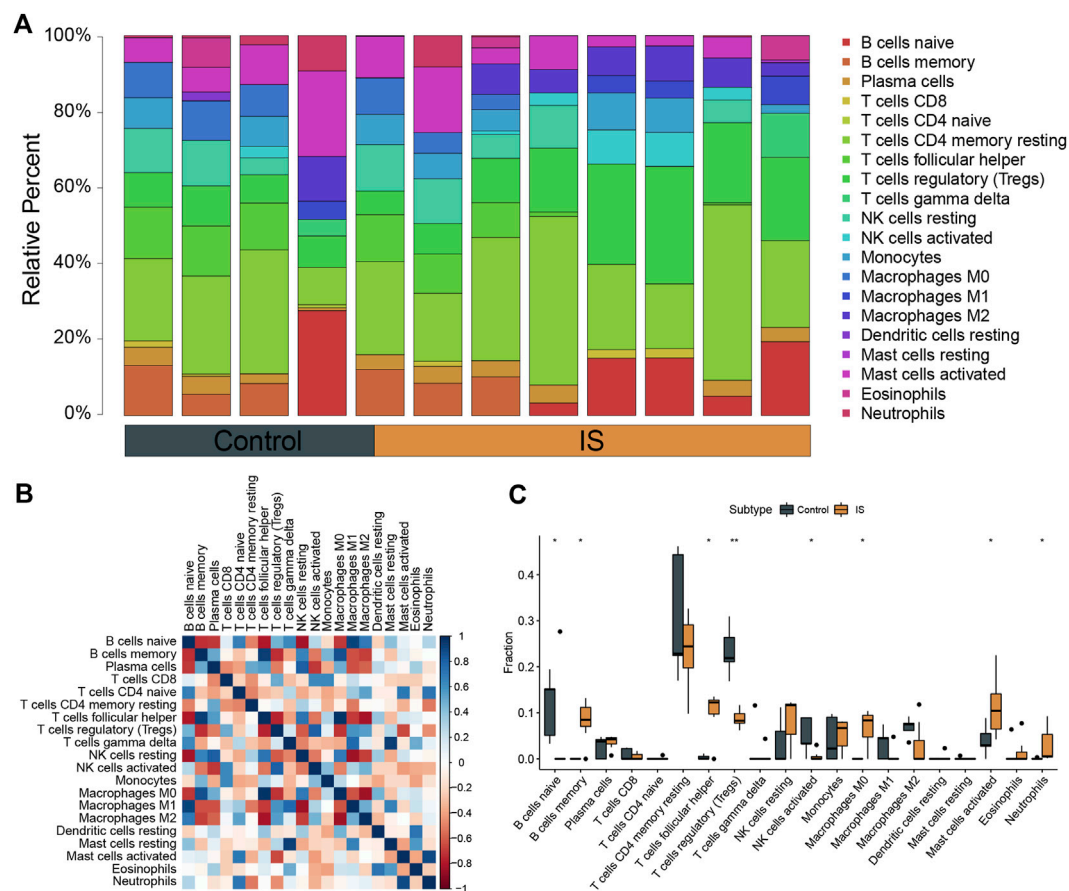


FIGURE 3

Analysis of immune infiltration in ischemic stroke (IS). (A) Overall expression of 20 infiltrating immune cell types in IS and control groups. (B) Correlation heat map of immune infiltrating cells. (C) Differential expression of infiltrating immune cells between the groups.

enriched in BP: regulation of neurotrophin TRK receptor signaling pathway, neurotrophin signaling pathway, transmission of nerve impulse, and regulation of ion transmembrane transport; in CC: postsynaptic membrane; and in MF: protein tyrosine kinase activity (Figure 4A). Hypertrophic cardiomyopathy, selected as the pathway with the highest score in KEGG analysis, is shown in Supplementary Figure S4. The overall results of GO and KEGG analysis are shown in Table 1, 2.

GSEA and GSVA

The top three pathways with the highest and lowest normalized enrichment score are shown in Figure 4B (highest: LEIN_NEURON_MARKERS, MIKKELSEN_MEF_HCP_WITH_H3K27ME3, and KIM_ALL_DISORDERS_CALB1_CORR_UP; lowest: MCLACHLAN_DENTAL_CARIES_UP, REACTOME_

RRNA_PROCESSING, and VERHAAK_GLIOMASTOMA_MESENCHYMAL). The overall results of GSEA are shown in Table 3. The top 10 enrichment results in GSVA with the highest MAD are shown in Figure 4C.

WGCNA

The “WGCNA” package in RStudio was used to identify co-expressed genes and modules. The clustering results based on characters showed good clustering, with no outlier samples were detected (Supplementary Figure S5A). In total, five modules were identified in WGCNA (Supplementary Figure S5B). By comparing the correlation between module genes and the two IS subtypes, blue modules with the largest correlation difference were identified as the most important modules (Figure 5A). DEGs in blue modules were subsequently analyzed.

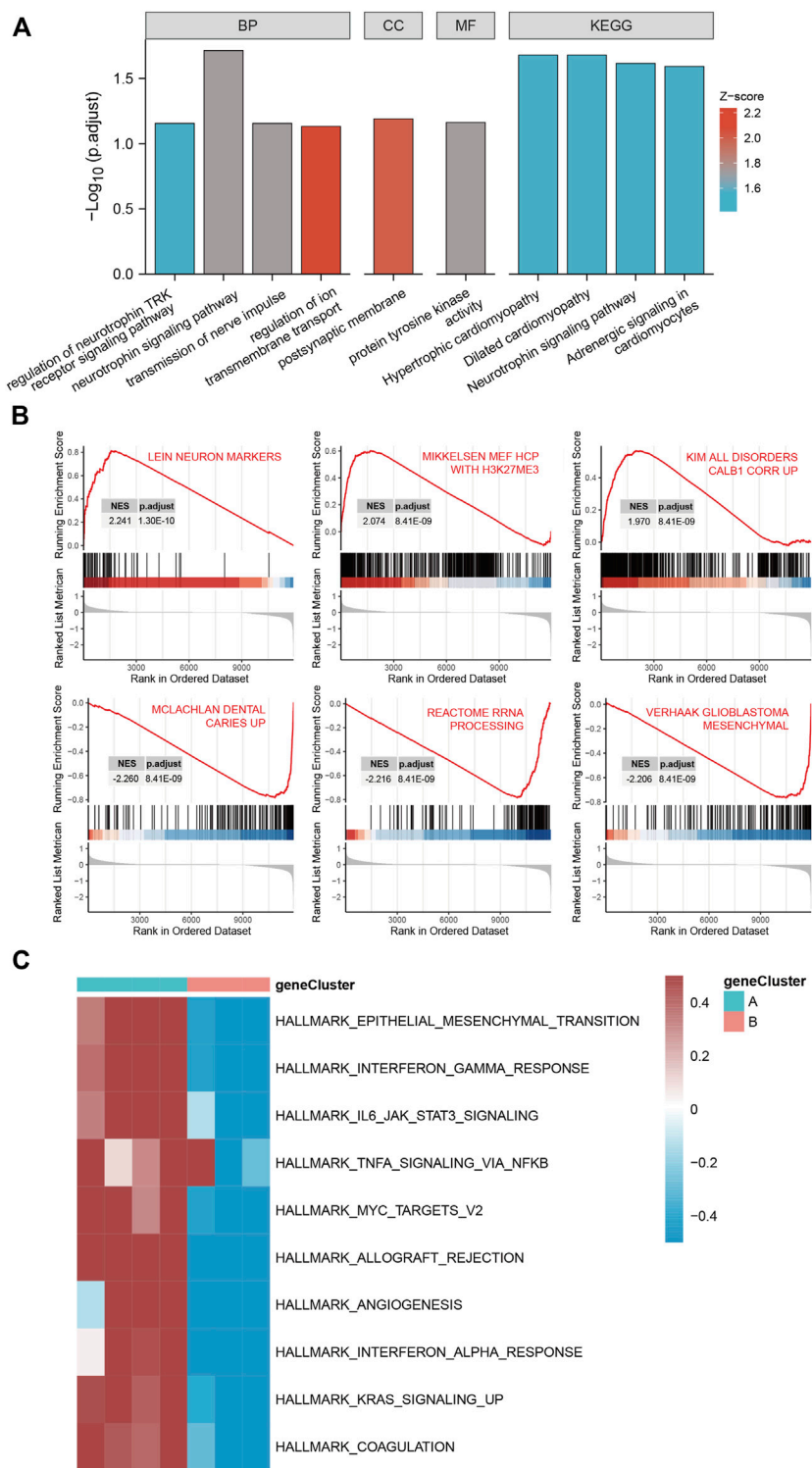


FIGURE 4 Gene ontology (GO), Kyoto Encyclopedia of Genes and Genomes (KEGG), gene set enrichment analysis (GSEA) and gene set variation analysis (GSVA) of DEGs between two IS subtypes. **(A)** Items with minimum Padj of GO enrichment analysis (biological processes, cellular components, and molecular functions) and KEGG pathway enrichment. **(B)** Clustering of the top three pathways with the highest and lowest normalized enrichment scores in GSEA. **(C)** Heat map of the top 10 enrichment results with the highest median absolute deviation of GSVA; red represents upregulation and blue represents downregulation.

TABLE 1 GO enrichment analysis.

GO Enrichment Results

| Category | ID | Description | BgRatio | pvalue | p.adjust | qvalue | geneID | Count |
|----------|------------|---|-----------|----------|----------|----------|---------------------|-------|
| BP | GO:0038179 | neurotrophin signaling pathway | 43/17859 | 0.000022 | 0.02064 | 0.015211 | Wasf1/Ntrk3/Agt | 3 |
| BP | GO:0051386 | regulation of neurotrophin TRK receptor signaling pathway | 16/17859 | 0.000188 | 0.076186 | 0.056146 | Wasf1/Agt | 2 |
| BP | GO:0019226 | transmission of nerve impulse | 96/17859 | 0.000247 | 0.076186 | 0.056146 | Cacng3/Ntrk3/Agt | 3 |
| BP | GO:0048011 | neurotrophin TRK receptor signaling pathway | 31/17859 | 0.000721 | 0.157632 | 0.116168 | Wasf1/Agt | 2 |
| CC | GO:0045211 | postsynaptic membrane | 322/18211 | 0.006574 | 0.158162 | 0.12973 | Cacng3/Ntrk3/Lzts1 | 3 |
| CC | GO:0043235 | receptor complex | 395/18211 | 0.011483 | 0.158162 | 0.12973 | Cacng3/Ntrk3/Tyro3 | 3 |
| CC | GO:0014069 | postsynaptic density | 425/18211 | 0.013980 | 0.158162 | 0.12973 | Cacng3/Rnf112/Lzts1 | 3 |
| CC | GO:0031209 | SCAR complex | 12/18211 | 0.014405 | 0.158162 | 0.12973 | Wasf1 | 1 |
| MF | GO:0004714 | transmembrane receptor protein tyrosine kinase activity | 106/16532 | 0.000361 | 0.020574 | 0.017646 | Ntrk3/Tyro3/Matk | 3 |
| MF | GO:0019199 | transmembrane receptor protein kinase activity | 123/16532 | 0.000558 | 0.020574 | 0.017646 | Ntrk3/Tyro3/Matk | 3 |
| MF | GO:0004713 | protein tyrosine kinase activity | 124/16532 | 0.000571 | 0.020574 | 0.017646 | Ntrk3/Tyro3/Matk | 3 |
| MF | GO:0016247 | channel regulator activity | 145/16532 | 0.000901 | 0.021156 | 0.018146 | Cacng3/Fxyd7/Agt | 3 |

TABLE 2 KEGG enrichment analysis.

KEGG Enrichment Results

| ID | Description | BgRatio | pvalue | p.adjust | qvalue | geneID | Count |
|----------|--|----------|-----------|----------|------------|------------|-------|
| rno05410 | Hypertrophic cardiomyopathy | 91/8947 | 0.0014946 | 0.023111 | 0.01509998 | Cacng3/Agt | 2 |
| rno05414 | Dilated cardiomyopathy | 94/8947 | 0.0015939 | 0.023111 | 0.01509998 | Cacng3/Agt | 2 |
| rno04722 | Neurotrophin signaling pathway | 120/8947 | 0.0025834 | 0.024973 | 0.01631621 | Ntrk3/Matk | 2 |
| rno04261 | Adrenergic signaling in cardiomyocytes | 148/8947 | 0.003903 | 0.028296 | 0.01848773 | Cacng3/Agt | 2 |

Construction of clinical diagnostic model

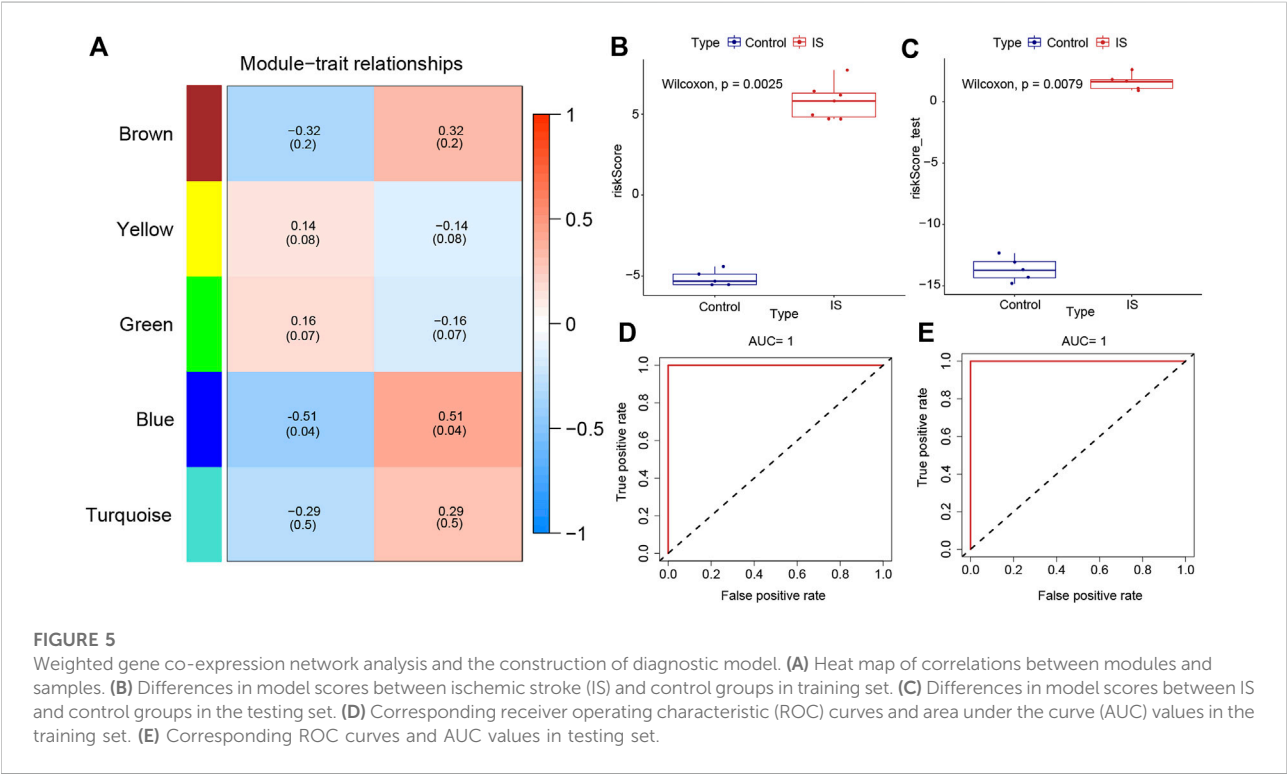
Datasets GSE97537 and GSE61616 were regarded as the training and testing sets, respectively. With the increase of parameter λ , the selected characteristic parameters decreased and absolute value of coefficients increased (Supplementary Figure S5C). After the simulation and selection of characteristic parameters, two models were obtained (optimal model and minimalist model) (Supplementary Figure S5D). We selected the minimalist model to

construct the diagnostic model and identified four genes as characteristic genes of IS, namely, *GFAP*, *GPNUMB*, *FKBP9*, and *CHMP5*. Then, model scores of IS and control groups in the two datasets were analyzed, and the results showed significant differences between two groups (Wilcoxon test, $p < 0.05$) (Figures 5B,C). The ROC curves of the training and testing sets were plotted to determine the area under the curve (AUC) value to verify the accuracy of the diagnostic model. The AUC values were 1 and 0.748, respectively (Figures 5D,E).

TABLE 3 GSEA analysis results.

GSEA Analysis Results

| ID | setSize | enrichmentScore | NES | pvalue | p.adjust | qvalues | rank |
|-----------------------------------|---------|-----------------|----------|----------|----------|----------|-------|
| LEIN_NEURON_MARKERS | 59 | 0.814076967 | 2.241206 | 1.30E-10 | 1.08E-08 | 7.07E-09 | 1,584 |
| MIKKELSEN_MEF_HCP_WITH_H3K27ME3 | 440 | 0.601882423 | 2.074171 | 1.00E-10 | 8.41E-09 | 5.53E-09 | 1,670 |
| KIM_ALL_DISORDERS_CALB1_CORR_UP | 468 | 0.567695316 | 1.970307 | 1.00E-10 | 8.41E-09 | 5.53E-09 | 1968 |
| WP_SYNAPTIC_VESICLE_PATHWAY | 47 | 0.744183724 | 1.96983 | 1.02E-05 | 0.000171 | 0.00011 | 1869 |
| REACTOME_ION_HOMEOSTASIS | 52 | 0.736104382 | 1.947812 | 5.60E-06 | 0.000108 | 7.07E-05 | 1727 |
| POOLA_INVASIVE_BREAST_CANCER_UP | 194 | -0.739888555 | -2.17464 | 1.00E-10 | 8.41E-09 | 5.53E-09 | 1744 |
| REACTOME_INTERLEUKIN_10_SIGNALING | 42 | -0.903693749 | -2.19738 | 1.00E-10 | 8.41E-09 | 5.53E-09 | 516 |
| VERHAAK_GLIOBLASTOMA_MESENCHYMAL | 169 | -0.76491492 | -2.20585 | 1.00E-10 | 8.41E-09 | 5.53E-09 | 1,481 |
| REACTOME_RRNA_PROCESSING | 135 | -0.785930466 | -2.21597 | 1.00E-10 | 8.41E-09 | 5.53E-09 | 1958 |
| MCLACHLAN_DENTAL_CARIES_UP | 169 | -0.783539634 | -2.25956 | 1.00E-10 | 8.41E-09 | 5.53E-09 | 1,083 |



Correlation analysis between characteristic genes and immune cell infiltration

The correlation of characteristic genes and immunocyte fractions was determined by Spearman correlation analysis and the results are displayed as a lollipop plot (Figures 6A–D).

Characteristic genes were verified by RT-qPCR

An MCAO mouse model was established to simulate IS, and RT-qPCR was performed to verify the expression of the four characteristic genes in the cerebral cortex of MCAO and Sham-operated mice. The expression of *GFAP*, *GPNMB*, *FKBP9*, and *CHMP5* was significantly higher in MCAO than in Sham-operated samples ($p < 0.05$) (Figure 7).

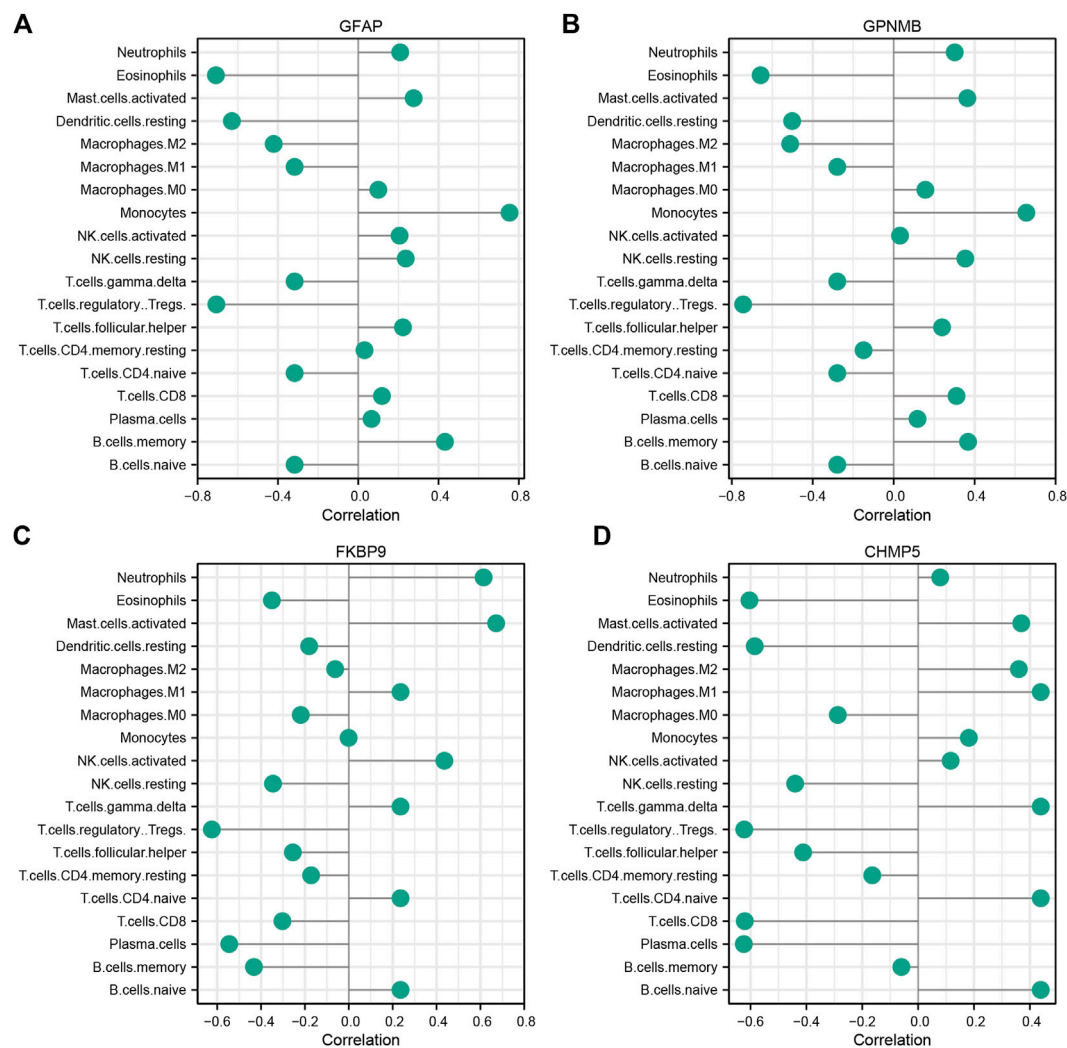


FIGURE 6

Correlation between characteristic ischemic stroke genes GFAP (A), GPNMB (B), FKBP9 (C), and CHMP5 (D) and infiltrating immune cells.

Construction of PPI network and hub gene regulatory network

The STRING database was used to construct the PPI network of characteristic genes associated with IS (Figure 8A). The interactions between genes were imported into Cytoscape, and the MCC value of each gene was calculated by CytoHubba. *GPNMB*, which had the highest MCC value, was identified as a hub gene (Figure 8B). The interaction network between *GPNMB* and TFs was obtained using the ENCODE database (Figure 8C). An miRNA, has-miR-26b-5p, was predicted to interact with *GPNMB*, and lncRNAs related to has-miR-26b-5p were further predicted. Subsequently, the ceRNA network of *GPNMB* was constructed based on these prediction results (Figure 8D).

Drug-mRNA network and molecular docking

In total, 241 drug molecules were predicted as potential drugs targeting *GPNMB* (Supplementary Figure S6). Putative *GPNMB*-drug interaction networks are shown in Figure 8E. The binding energy between chloroquine and *GPNMB* in most possible mode, of the lowest binding energy, was -5.6 kcal/mol, less than -5 kcal/mol, indicating small molecules were capable of binding with protein receptor to an extent. A hydrogen bond existed between alanine in *GPNMB* and small drug molecules with a distance of 3.2 Å. This specific bond was observed as basic interaction in the docking pocket area. According to results of YASARA software, the most possible docking mode was shown in (Figure 8F). Binding location

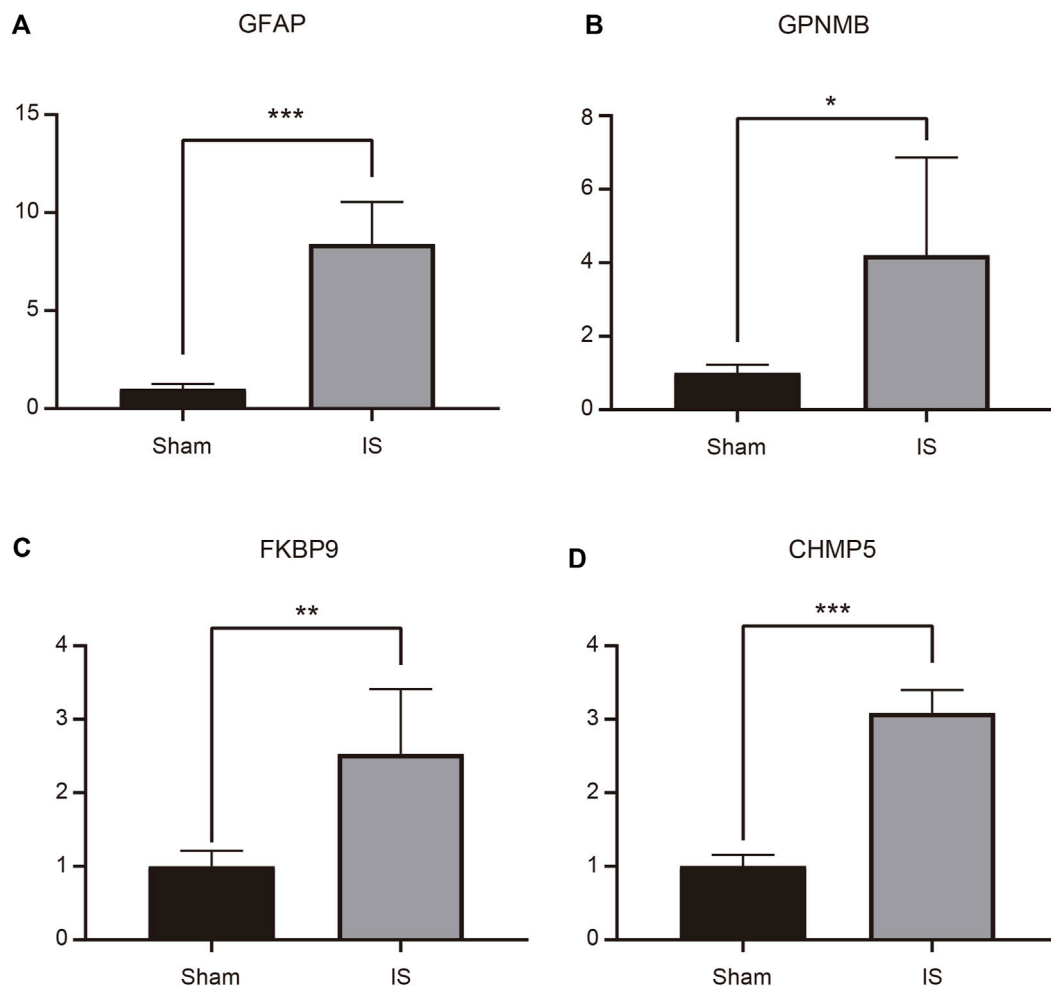


FIGURE 7

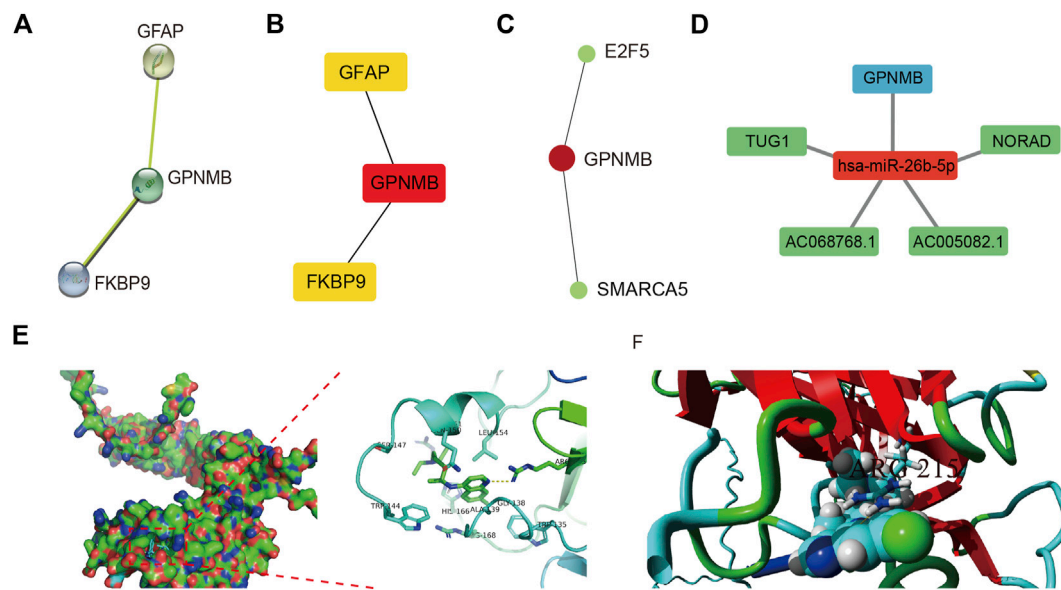
Expression of GFAP (A), GPNMB (B), FKBP9 (C), and CHMP5 (D) in ischemic stroke brain tissue determined by RT-qPCR. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

between two algorithms were the same, both of hydrogen bond with ARG215. I-Mutant 3.0 web server, assessing the influence of one amino acid dot mutation on GPNMB protein through free energy stability change (DDG) was utilized for ARG 215 (Fang et al., 2019). The average DDG was -0.73 kcal/mol. According to reference criteria, $DDG < -0.5$, indicating large decrease of protein structure stability (Lim et al., 2021). In our article, ARG 215 was hydrogen binding location, results of mutation revealing the important function of this location.

Discussion

Stroke, as a sudden disorder of cerebral blood circulation, has emerged as the second leading cause of death and disability worldwide owing to the lack of early diagnosis and effective treatment. IS the most prevalent form of stroke, accounting for

approximately 75–80% of cases (Krishnamurthi et al., 2018). Increasing evidence supports the involvement of epigenetic alterations in the pathogenesis of IS (Stanzione et al., 2020). However, as one of the most ubiquitous epigenetic modifications in mammalian cells, the role of RNA methylation modification in the regulation of IS remains unclear. To determine the role of RNA methylation-related regulators (m1A, m6A, and m5C) in the pathogenesis and immune microenvironment of IS, we determined the overall expression of RNA methylation regulators and immune infiltration in rat IS gene sets and divided IS into two molecular subtypes according to the expression levels of RNA methylation regulators. An IS clinical diagnosis model was successfully constructed and four upregulated characteristic genes were identified, which were significantly negatively correlated with the degree of Treg infiltration. Furthermore, we identified a hub gene by PPI network analysis and predicted its regulatory networks and the potential therapeutic drug, chloroquine.

**FIGURE 8**

Protein-protein interaction (PPI) network, transcription factor (TF) correlation network, competing endogenous RNA (ceRNA) network and drug-molecular docking of the hub gene. **(A)** PPI network; each node represents a different gene. **(B)** Maximal clique centrality (MCC) value of each gene was calculated. The deeper the color the higher, the MCC value of genes. **(C)** TF correlation network of the hub gene. Red represents genes and green represents TFs. **(D)** ceRNA network of the hub gene. Red represents miRNA, blue represents mRNA, and green represents lncRNA. **(E)** Docking results between GPNMB and chloroquine in Autodock. Hydrogen bond between two was marked in yellow dotted line. **(F)** Docking results of YASARA. Bonds marked in yellow represent hydrogen bond.

Two IS molecular subtypes were constructed based on the differentially expressed RNA methylation-related regulators between IS and normal samples, and 28 DEGs were identified between them. GO analysis revealed that these DEGs were engaged in BP of regulation of neurotrophin TRK receptor signaling pathway, neurotrophin signaling pathway, transmission of nerve impulse, and regulation of ion transmembrane transport, which have been referred to as the main mechanisms of IS (Li et al., 2021; Zhu et al., 2022). Additionally, the protein products of these DEGs are mainly distributed in the postsynaptic membrane and primarily involved in regulating protein tyrosine kinase activity. A previous study has revealed that protein tyrosine kinase activity, which is closely associated with various synaptic and cellular functions in the brain, is upregulated in IS (Takagi, 2014). Pathway enrichment assessments determined that the top four enriched pathways were hypertrophic cardiomyopathy, dilated cardiomyopathy, neurotrophin signaling pathway, and adrenergic signaling in cardiomyocytes. We speculated that DEGs may influence the occurrence and progression of IS through these potential pathways. Furthermore, GSEA and GSVA results showed that IS was influenced by the inflammation and immune regulation pathway, which is consistent with the current theory that inflammation and immune responses play key roles in the regulatory network of IS (Endres et al., 2022).

In this study, we constructed a diagnostic model of IS and identified four characteristic genes (*GFAP*, *GPNMB*, *FKBP9*, and *CHMP5*) with good diagnostic value. The results of RT-qPCR showed that these four characteristic genes were significantly upregulated in a mouse MCAO model, which verified our results. Glial fibrillary acidic protein (GFAP), an intermediate filament protein only produced by astrocytes, is a well-established marker of astrocyte activation in central nervous system (CNS) diseases (Sayad et al., 2022). Growing evidence suggests the potential clinical application value of blood GFAP levels in numerous neuroinflammatory and neurodegenerative diseases, as they can be used to detect even subtle injury to the CNS (Abdelhak et al., 2022; Heimfarth et al., 2022). A previous study has shown that IS can induce the transformation of astrocytes into a neurotoxic A1 phenotype and increase GFAP expression (Zhang et al., 2022). Amalia, 2021 reported that GFAP is highly expressed in cerebrospinal fluid and serum from patients with IS, demonstrating its potential as a reliable biological marker to help diagnose IS. Our results revealed a similar expression tendency of GFAP as these reports.

Glycoprotein nonmetastatic melanoma protein B (GPNMB) is a type-I transmembrane protein, also known as dendritic cell heparan sulfate proteoglycan integrin-dependent ligand, that has been demonstrated to be overexpressed in numerous cancers and is associated with a metastatic phenotype (Huang et al., 2021). It

has been reported that tumor endothelial cells can induce tumor-infiltrating CD8 T cell exhaustion and promote the escape of cancer cells from immune surveillance by upregulating the expression of GPNMB (Sakano et al., 2022). Furthermore, GPNMB plays an important role in various diseases in addition to cancer by regulating inflammation and immune responses. GPNMB can negatively regulate macrophage inflammatory capacity *via* the inhibition of NF- κ B signaling by binding to CD44 (Prabata et al., 2021). In a cellular amyotrophic lateral sclerosis model, it has been shown that GPNMB exerts neuroprotective effects by binding to Na/K-ATPase, an ion pump and receptor that modulates neuroinflammation (Ono et al., 2016). Consistent with our results, Nakano et al. (2014) have demonstrated that GPNMB is upregulated after ischemic reperfusion, and the overexpression of GPNMB has neuroprotective effects against IS, although the mechanism has not been fully characterized.

FK506-binding protein 9 (FKBP9), a member of the immunophilin family FKBP9, binds to the immunosuppressive drug tacrolimus (FK506) (Ghartey-Kwansah et al., 2018). FKBP9 is widely expressed in multiple human organs and tissues and involved in the regulation of various physiological processes. It has been reported that FKBP9 is associated with metastasis and poor prognosis in a variety of cancers (Annett et al., 2020). For example, Xu H. et al. (2020) demonstrated that FKBP9 is upregulated in human glioblastoma samples and promotes malignant phenotypes by regulating unfolded protein response signaling. Additionally, FKBP9 is closely related to physiological functions such as T cell activation and plays an important role in immune system regulation (Jiang et al., 2020).

Charged multivesicular body protein 5 (CHMP5), a component of the endosomal sorting complex required for transport-III, is responsible for the final conversion of late endosomal multivesicular body to lysosomes (Shim et al., 2006). CHMP5 is a multifunctional protein with potential roles in cellular signaling. It was previously reported that CHMP5 has antiapoptotic functions because silencing CHMP5 induces apoptosis by caspase cascade activation (Mo et al., 2018). Additionally, CHMP5 prevents Bcl-2, a widely recognized apoptosis suppressor gene that intrinsically regulates apoptosis, from deleterious oxidation by reactive oxygen species (ROS) formation (Adoro et al., 2017). Furthermore, it has been shown that CHMP5 has a key role in T-cell receptor signaling and its deficiency affects T-cell receptor expression on the cell surface (Wi et al., 2016). However, the role of CHMP5 in IS has not been studied to date.

In terms of the immune response, the infiltration of Tregs was significantly lower in IS samples than in control samples. This result is consistent with the previous findings of Noh et al. (2018), indicating that Tregs are closely related to the pathogenesis of IS. Tregs are an important subpopulation of T lymphocytes that are involved in resisting immune response

overactivity, maintaining immune homeostasis, and regulating inflammation (Wang et al., 2021). In Figure 3, immune infiltration levels were evaluated through Cibersort algorithm. Total eight types of immune cells were found to be of statistical significance, Treg cells were the most significant among them. In Figure 6, Several species of immune cells including Eosinophils, Treg and so on were of significant correlation relationship ($|\text{Correlation coefficient}| > 0.5$) with four characteristic genes. Treg, being of the most significant statistical meanings between IS and control, were closely negatively related with all characteristic genes. Activation of Treg has been verified to slow down the process of progress of IR through reducing IFN- γ in the IR microenvironment (Hu et al., 2013). GPNMB protein could bind to heparan sulphate-like structures, blocking the activation of T cells (Chung et al., 2013). These are consistent with our results.

Our study showed that the four IS RNA methylation-related characteristic genes were significantly negatively correlated with the degree of Tregs infiltration, indicating that these genes participate in the immune regulation of IS. M0 macrophage is another important cell type in immune infiltration, was found to secrete interleukin-1 β accompanying with the progress of early IR (Mabuchi et al., 2000). Thus, accumulation of interleukin-1 β can promote the development of follicular helper T cells (Kobayashi et al., 2017), been seen as a bridge linking M0 macrophages with follicular helper T cells in our study.

We built a PPI network and identified GPNMB as a hub gene. miRNAs are a class of small, single-stranded noncoding RNAs that regulate target gene expression on a post-transcriptional level (Correia de Sousa et al., 2019). In our study, miR-26b-5p was predicted to act on GPNMB, and the expression of GPNMB in IS samples was upregulated. Previous studies have shown that miR-26b-5p is associated with various disease states, such as tumors, inflammation, autoimmune disease, and IS. For example, a bioinformatics analysis reported that miR-26b-5p can be recognized as a potential biomarker for IS (Barrera-Vázquez et al., 2022). Additionally, Xiao et al. (2021) have observed that miR-26b-5p is downregulated in the brain of an MCAO rat model and that the overexpression of miR-26b-5p reduces apoptosis and the inflammatory response. Another study has shown that miR-26b-5p alleviates IS injury by negatively regulating the expression of Smad1, which promotes apoptosis and inflammation by increasing the level of ROS in cells (Shanguan et al., 2020). The findings of these reports are consistent with our results.

Finally, we predicted potential drug molecules that may bind to GPNMB, and the most prominent was chloroquine. Chloroquine, an established drug originally used for the treatment of malaria, has been reported to have anti-inflammatory and immunomodulatory properties (Silva et al., 2021). Recently, several studies have shown that chloroquine pretreatment can alleviate brain injury in IS through a variety of mechanisms, including the inhibition of the inflammatory

response by lowering myeloperoxidase activity and inflammatory cytokine gene expression (Cui et al., 2013; Zhang Y. P. et al., 2020) and alleviation of neuronal injury by restoring ganglioside homeostasis (Caughlin et al., 2019). Gabriel et al. (2014) reported that chloroquine can effectively increase *Gpnmb* transcription in mice as a lysosomal stress inducer. All these reports are mostly consistent with our analysis.

There were some limitations to this study. First, single microarray analysis may be associated with high false-positive rates, and it is necessary to integrate multiple individual datasets in future studies to improve the reliability of the results. Second, although the clinical diagnostic model constructed in this study showed high accuracy, the sample sizes of the training and validation sets were small, resulting in insufficient statistical efficacy. Performing cross-validation internally and increasing the sample size for external validation in future studies would be beneficial. Finally, our research was retrospective, and a large number of prospective studies are needed to validate the results.

In conclusion, this study is the first to comprehensively analyze the correlations between RNA methylation-related regulators and IS and immune infiltration. We identified two highly heterogeneous RNA methylation subtypes in IS, with significantly different BP and MF. An IS clinical diagnosis model was constructed and four characteristic genes with effective diagnostic value were identified using bioinformatics methodologies (such as WGCNA and LASSO regression). *GPNMB* was identified as a hub gene by PPI network analysis, and its regulatory networks and binding to the potential therapeutic drug chloroquine may provide guidance for clinical diagnosis and treatment. Overall, our study may provide insight into the potential molecular mechanisms underlying IS and a new basis for optimizing the clinical diagnosis and treatment of patients with IS.

Data availability statement

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

Ethics statement

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

Author contributions

XZ conceived and designed the research, conducted bioinformatic analyses, and drafted the manuscript. YW, BD, YJ, DL, and KX provided insightful guidance and revised the manuscript. YY administered the project, revised the manuscript, and provided support. All authors contributed to the article and read and approved the final version of the manuscript.

Funding

This work was supported by the National Natural Science Foundation of China (Grant 82072150) and Tianjin Natural Science Foundation (Grant 21JCQNJC01100).

Acknowledgments

We thank the Gene Expression Omnibus database for providing the platform and contributors for uploading meaningful datasets. We also acknowledge Editage (www.editage.cn) for professional linguistic editing 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

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.2022.1009145/full#supplementary-material>

References

- Abdelhak, A., Foschi, M., Abu-Rumeileh, S., Yue, J. K., D'Anna, L., Huss, A., et al. (2022). Blood GFAP as an emerging biomarker in brain and spinal cord disorders. *Nat. Rev. Neurol.* 18 (3), 158–172. doi:10.1038/s41582-021-00616-3
- Adoro, S., Park, K. H., Bettigole, S. E., Lis, R., Shin, H. R., Seo, H., et al. (2017). Post-translational control of T cell development by the ESCRT protein CHMP5. *Nat. Immunol.* 18 (7), 780–790. doi:10.1038/ni.3764
- Amalia, L. (2021). Glial fibrillary acidic protein (GFAP): Neuroinflammation biomarker in acute ischemic stroke. *J. Inflamm. Res.* 14, 7501–7506. doi:10.2147/jir.S342097
- Annett, S., Moore, G., and Robson, T. (2020). FK506 binding proteins and inflammation related signalling pathways; basic biology, current status and future prospects for pharmacological intervention. *Pharmacol. Ther.* 215, 107623. doi:10.1016/j.pharmthera.2020.107623
- 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
- Barrera-Vázquez, O. S., Gomez-Verjan, J. C., Ramírez-Aldana, R., Torre, P. G., and Rivero-Segura, N. A. (2022). Structural and pharmacological network analysis of miRNAs involved in acute ischemic stroke: A systematic review. *Int. J. Mol. Sci.* 23 (9), 4663. doi:10.3390/ijms23094663
- Boccalletto, P., Stefaniak, F., Ray, A., Cappannini, A., Mukherjee, S., Purta, E., et al. (2022). Modomics: A database of RNA modification pathways. 2021 update. *Nucleic Acids Res.* 50 (D1), D231–d235. doi:10.1093/nar/gkab1083
- Burley, S. K., Berman, H. M., Kleywegt, G. J., Markley, J. L., Nakamura, H., and Velankar, S. (2017). Protein data bank (PDB): The single global macromolecular structure archive. *Methods Mol. Biol.* 1607, 627–641. doi:10.1007/978-1-4939-7000-1_26
- Caughlin, S., Hepburn, J., Liu, Q., Wang, L., Yeung, K. K., Cechetto, D. F., et al. (2019). Chloroquine restores ganglioside homeostasis and improves pathological and behavioral outcomes post-stroke in the rat. *Mol. Neurobiol.* 56 (5), 3552–3562. doi:10.1007/s12035-018-1317-0
- Chavda, V., Madhwani, K., and Chaurasia, B. (2021). Stroke and immunotherapy: Potential mechanisms and its implications as immune-therapeutics. *Eur. J. Neurosci.* 54, 4338–4357. doi:10.1111/ejn.15224
- 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, Y. T., Shen, J. Y., Chen, D. P., Wu, C. F., Guo, R., Zhang, P. P., et al. (2020). Identification of cross-talk between m(6)A and 5mC regulators associated with onco-immunogenic features and prognosis across 33 cancer types. *J. Hematol. Oncol.* 13 (1), 22. doi:10.1186/s13045-020-00854-w
- 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, S11. doi:10.1186/1752-0509-8-s4-s11
- Chung, J. S., Tomihari, M., Tamura, K., Kojima, T., Cruz, P. D., Jr., and Ariizumi, K. (2013). The DC-HIL ligand syndecan-4 is a negative regulator of T-cell alloreactivity responsible for graft-versus-host disease. *Immunology* 138 (2), 173–182. doi:10.1111/imm.12027
- Correia de Sousa, M., Gjorgjieva, M., Dolicka, D., Sobolewski, C., and Foti, M. (2019). Deciphering miRNAs' action through miRNA editing. *Int. J. Mol. Sci.* 20 (24), E6249. doi:10.3390/ijms20246249
- Cui, G., Ye, X., Zuo, T., Zhao, H., Zhao, Q., Chen, W., et al. (2013). Chloroquine pretreatment inhibits toll-like receptor 3 signaling after stroke. *Neurosci. Lett.* 548, 101–104. doi:10.1016/j.neulet.2013.02.072
- Davis, A. P., Grondin, C. J., Johnson, R. J., Sciaky, D., Wieggers, J., Wieggers, T. C., et al. (2021). Comparative Toxicogenomics database (CTD): Update 2021. *Nucleic Acids Res.* 49 (D1), D1138–d1143. doi:10.1093/nar/gkaa891
- Davis, C. A., Hitz, B. C., Sloan, C. A., Chan, E. T., Davidson, J. M., Gabdank, I., et al. (2018). The encyclopedia of DNA elements (ENCODE): Data portal update. *Nucleic Acids Res.* 46 (D1), D794–d801. doi:10.1093/nar/gkx1081
- Endres, M., Moro, M. A., Nolte, C. H., Dames, C., Buckwalter, M. S., and Meisel, A. (2022). Immune pathways in etiology, acute phase, and chronic sequelae of ischemic stroke. *Circ. Res.* 130 (8), 1167–1186. doi:10.1161/circresaha.121.319994
- Espinosa, A., Meneses, G., Chavarria, A., Mancilla, R., Pedraza-Chaverri, J., Fleury, A., et al. (2020). Intranasal dexamethasone reduces mortality and brain damage in a mouse experimental ischemic stroke model. *Neurotherapeutics* 17 (4), 1907–1918. doi:10.1007/s13311-020-00884-9
- Fang, X., Huang, J., Zhang, R., Wang, F., Zhang, Q., Li, G., et al. (2019). Convolution neural network-based prediction of protein thermostability. *J. Chem. Inf. Model.* 59 (11), 4833–4843. doi:10.1021/acs.jcim.9b00220
- Gabriel, T. L., Tol, M. J., Ottenhof, R., van Roomen, C., Aten, J., Claessen, N., et al. (2014). Lysosomal stress in obese adipose tissue macrophages contributes to MITF-dependent Gpnmb induction. *Diabetes* 63 (10), 3310–3323. doi:10.2337/db13-1720
- Gao, Y., Wang, H., Li, H., Ye, X., Xia, Y., Yuan, S., et al. (2021). Integrated analyses of m(1)A regulator-mediated modification patterns in tumor microenvironment-infiltrating immune cells in colon cancer. *Oncoimmunology* 10 (1), 1936758. doi:10.1080/2162402x.2021.1936758
- Ghartey-Kwansah, G., Li, Z., Feng, R., Wang, L., Zhou, X., Chen, F. Z., et al. (2018). Comparative analysis of FKBP family protein: Evaluation, structure, and function in mammals and *Drosophila melanogaster*. *BMC Dev. Biol.* 18 (1), 7. doi:10.1186/s12861-018-0167-3
- Han, D., Liu, J., Chen, C., Dong, L., Liu, Y., Chang, R., et al. (2019). Anti-tumour immunity controlled through mRNA m(6)A methylation and YTHDF1 in dendritic cells. *Nature* 566 (7743), 270–274. doi:10.1038/s41586-019-0916-x
- 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, L., Li, H., Wu, A., Peng, Y., Shu, G., and Yin, G. (2019). Functions of N6-methyladenosine and its role in cancer. *Mol. Cancer* 18 (1), 176. doi:10.1186/s12943-019-1109-9
- Heimfarth, L., Passos, F. R. S., Monteiro, B. S., Araújo, A. A. S., Quintans Júnior, L. J., and Quintans, J. S. S. (2022). Serum glial fibrillary acidic protein is a body fluid biomarker: A valuable prognostic for neurological disease - a systematic review. *Int. Immunopharmacol.* 107, 108624. doi:10.1016/j.intimp.2022.108624
- Henderson, S. J., Weitz, J. I., and Kim, P. Y. (2018). Fibrinolysis: Strategies to enhance the treatment of acute ischemic stroke. *J. Thromb. Haemost.* 16 (10), 1932–1940. doi:10.1111/jth.14215
- Herpich, F., and Rincon, F. (2020). Management of acute ischemic stroke. *Crit. Care Med.* 48 (11), 1654–1663. doi:10.1097/ccm.0000000000004597
- Hu, J., Zhang, L., Wang, N., Ding, R., Cui, S., Zhu, F., et al. (2013). Mesenchymal stem cells attenuate ischemic acute kidney injury by inducing regulatory T cells through splenocyte interactions. *Kidney Int.* 84 (3), 521–531. doi:10.1038/ki.2013.114
- Huang, H. Y., Lin, Y. C., Cui, S., Huang, Y., Tang, Y., Xu, J., et al. (2022). miRTarBase update 2022: an informative resource for experimentally validated miRNA-target interactions. *Nucleic Acids Res.* 50 (D1), D222–d230. doi:10.1093/nar/gkab1079
- Huang, Y. H., Chu, P. Y., Chen, J. L., Huang, C. T., Huang, C. C., Tsai, Y. F., et al. (2021). Expression pattern and prognostic impact of glycoprotein non-metastatic B (GPNMB) in triple-negative breast cancer. *Sci. Rep.* 11 (1), 12171. doi:10.1038/s41598-021-91588-3
- Jiang, F. N., Dai, L. J., Yang, S. B., Wu, Y. D., Liang, Y. X., Yin, X. L., et al. (2020). Increasing of FKBP9 can predict poor prognosis in patients with prostate cancer. *Pathol. Res. Pract.* 216 (1), 152732. doi:10.1016/j.prp.2019.152732
- 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
- Kobayashi, T., Iijima, K., Dent, A. L., and Kita, H. (2017). Follicular helper T cells mediate IgE antibody response to airborne allergens. *J. Allergy Clin. Immunol.* 139 (1), 300–313. e307. doi:10.1016/j.jaci.2016.04.021
- Krieger, E., and Vriend, G. (2014). YASARA View - molecular graphics for all devices - from smartphones to workstations. *Bioinformatics* 30 (20), 2981–2982. doi:10.1093/bioinformatics/btu426
- Krishnamurthi, R. V., Barker-Collo, S., Parag, V., Parmar, P., Witt, E., Jones, A., et al. (2018). Stroke incidence by major pathological type and ischemic subtypes in the auckland regional community stroke studies: Changes between 2002 and 2011. *Stroke* 49 (1), 3–10. doi:10.1161/strokeaha.117.019358
- Langfelder, P., and Horvath, S. (2008). Wgcna: an R package for weighted correlation network analysis. *BMC Bioinforma.* 9, 559. doi:10.1186/1471-2105-9-559
- Li, 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. Database issue). doi:10.1093/nar/gkt1248
- Li, Z., Cui, Y., Feng, J., and Guo, Y. (2020). Identifying the pattern of immune related cells and genes in the peripheral blood of ischemic stroke. *J. Transl. Med.* 18 (1), 296. doi:10.1186/s12967-020-02463-0

- Li, Z., Wang, H., Xiao, G., Du, H., He, S., Feng, Y., et al. (2021). Recovery of post-stroke cognitive and motor deficiencies by Shu Xuening injection via regulating hippocampal BDNF-mediated Neurotrophin/Trk Signaling. *Biomed. Pharmacother.* 141, 111828. doi:10.1016/j.biopha.2021.111828
- Liberzon, A., Birger, C., Thorvaldsdóttir, H., Ghandi, M., Mesirov, J. P., and Tamayo, P. (2015). The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst.* 1 (6), 417–425. doi:10.1016/j.cels.2015.12.004
- Lim, S. W., Tan, K. J., Azuraiddi, O. M., Sathiy, M., Lim, E. C., Lai, K. S., et al. (2021). Functional and structural analysis of non-synonymous single nucleotide polymorphisms (nsSNPs) in the MYB oncoproteins associated with human cancer. *Sci. Rep.* 11 (1), 24206. doi:10.1038/s41598-021-03624-x
- Liu, R., Song, P., Gu, X., Liang, W., Sun, W., Hua, Q., et al. (2021). Comprehensive landscape of immune infiltration and aberrant pathway activation in ischemic stroke. *Front. Immunol.* 12, 766724. doi:10.3389/fimmu.2021.766724
- Luo, W., and Brouwer, C. (2013). Pathview: An R/bioconductor package for pathway-based data integration and visualization. *Bioinformatics* 29 (14), 1830–1831. doi:10.1093/bioinformatics/btt285
- Mabuchi, T., Kitagawa, K., Ohtsuki, T., Kuwabara, K., Yagita, Y., Yanagihara, T., et al. (2000). Contribution of microglia/macrophages to expansion of infarction and response of oligodendrocytes after focal cerebral ischemia in rats. *Stroke* 31 (7), 1735–1743. doi:10.1161/01.str.31.7.1735
- Mo, J. S., Han, S. H., Yun, K. J., and Chae, S. C. (2018). MicroRNA 429 regulates the expression of CHMP5 in the inflammatory colitis and colorectal cancer cells. *Inflamm. Res.* 67 (11–12), 985–996. doi:10.1007/s00011-018-1194-z
- Nakano, Y., Suzuki, Y., Takagi, T., Kitashoji, A., Ono, Y., Tsuruma, K., et al. (2014). Glycoprotein nonmetastatic melanoma protein B (GPNMB) as a novel neuroprotective factor in cerebral ischemia-reperfusion injury. *Neuroscience* 277, 123–131. doi:10.1016/j.neuroscience.2014.06.065
- Noh, M. Y., Lee, W. M., Lee, S. J., Kim, H. Y., Kim, S. H., and Kim, Y. S. (2018). Regulatory T cells increase after treatment with poly (ADP-ribose) polymerase-1 inhibitor in ischemic stroke patients. *Int. Immunopharmacol.* 60, 104–110. doi:10.1016/j.intimp.2018.04.043
- Ono, Y., Tsuruma, K., Takata, M., Shimazawa, M., and Hara, H. (2016). Glycoprotein nonmetastatic melanoma protein B extracellular fragment shows neuroprotective effects and activates the PI3K/Akt and MEK/ERK pathways via the Na⁺/K⁺-ATPase. *Sci. Rep.* 6, 23241. doi:10.1038/srep23241
- Pan, J., Huang, Z., and Xu, Y. (2021). m5C RNA methylation regulators predict prognosis and regulate the immune microenvironment in lung squamous cell carcinoma. *Front. Oncol.* 11, 657466. doi:10.3389/fonc.2021.657466
- Prabata, A., Ikeda, K., Rahardini, E. P., Hirata, K. I., and Emoto, N. (2021). GPNMB plays a protective role against obesity-related metabolic disorders by reducing macrophage inflammatory capacity. *J. Biol. Chem.* 297 (5), 101232. doi:10.1016/j.jbc.2021.101232
- 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 (7), e47. doi:10.1093/nar/gkv007
- Sakano, Y., Noda, T., Kobayashi, S., Sasaki, K., Iwagami, Y., Yamada, D., et al. (2022). Tumor endothelial cell-induced CD8(+) T-cell exhaustion via GPNMB in hepatocellular carcinoma. *Cancer Sci.* 113 (5), 1625–1638. doi:10.1111/cas.15331
- Sayad, A., Uddin, S. M., Yao, S., Wilson, H., Chan, J., Zhao, H., et al. (2022). A magnetoimpedance biosensor microfluidic platform for detection of glial fibrillary acidic protein in blood for acute stroke classification. *Biosens. Bioelectron.* 211, 114410. doi:10.1016/j.bios.2022.114410
- Shangguan, Y., Han, J., and Su, H. (2020). GAS5 knockdown ameliorates apoptosis and inflammatory response by modulating miR-26b-5p/Smad1 axis in cerebral ischaemia/reperfusion injury. *Behav. Brain Res.* 379, 112370. doi:10.1016/j.bbr.2019.112370
- 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
- Shi, J., Yang, Y., Yin, N., Liu, C., Zhao, Y., Cheng, H., et al. (2022). Engineering CXCL12 biomimetic decoy-integrated versatile immunosuppressive nanoparticle for ischemic stroke therapy with management of overactivated brain immune microenvironment. *Small Methods* 6 (1), e2101158. doi:10.1002/smt.202101158
- Shim, J. H., Xiao, C., Hayden, M. S., Lee, K. Y., Trombetta, E. S., Pypaert, M., et al. (2006). CHMP5 is essential for late endosome function and down-regulation of receptor signaling during mouse embryogenesis. *J. Cell Biol.* 172 (7), 1045–1056. doi:10.1083/jcb.200509041
- Silva, R., Tan, L., Rodrigues, D. A., Prestes, E. B., Gomes, C. P., Gama, A. M., et al. (2021). Chloroquine inhibits pro-inflammatory effects of heme on macrophages and in vivo. *Free Radic. Biol. Med.* 173, 104–116. doi:10.1016/j.freeradbiomed.2021.07.028
- Sing, T., Sander, O., Beerenwinkel, N., and Lengauer, T. (2005). ROCR: Visualizing classifier performance in R. *Bioinformatics* 21 (20), 3940–3941. doi:10.1093/bioinformatics/bti623
- Stanzione, R., Cotugno, M., Bianchi, F., Marchitti, S., Forte, M., Volpe, M., et al. (2020). Pathogenesis of ischemic stroke: Role of epigenetic mechanisms. *Genes (Basel)* 11 (1), E89. doi:10.3390/genes11010089
- Takagi, N. (2014). Protein tyrosine phosphorylation in the ischemic brain. *J. Pharmacol. Sci.* 125 (4), 333–339. doi:10.1254/jphs.14r04cp
- Wang, H., Wang, Z., Wu, Q., Yuan, Y., Cao, W., and Zhang, X. (2021). Regulatory T cells in ischemic stroke. *CNS Neurosci. Ther.* 27 (6), 643–651. doi:10.1111/cns.13611
- Wang, L., Yu, Y., Yang, J., Zhao, X., and Li, Z. (2015). Dissecting Xuesaitong's mechanisms on preventing stroke based on the microarray and connectivity map. *Mol. Biosyst.* 11 (11), 3033–3039. doi:10.1039/c5mb00379b
- Wi, S. M., Min, Y., and Lee, K. Y. (2016). Charged MVB protein 5 is involved in T-cell receptor signaling. *Exp. Mol. Med.* 48 (1), e206. doi:10.1038/emm.2015.102
- 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
- Xiao, Y., Zheng, S., Duan, N., Li, X., and Wen, J. (2021). MicroRNA-26b-5p alleviates cerebral ischemia-reperfusion injury in rats by inhibiting the N-myc/PTEN axis by downregulating KLF10 expression. *Hum. Exp. Toxicol.* 40 (8), 1250–1262. doi:10.1177/0960327121991899
- Xu, H., Liu, P., Yan, Y., Fang, K., Liang, D., Hou, X., et al. (2020a). FKBP9 promotes the malignant behavior of glioblastoma cells and confers resistance to endoplasmic reticulum stress inducers. *J. Exp. Clin. Cancer Res.* 39 (1), 44. doi:10.1186/s13046-020-1541-0
- Xu, S., Li, Y., Chen, J. P., Li, D. Z., Jiang, Q., Wu, T., et al. (2020b). Oxygen glucose deprivation/re-oxygenation-induced neuronal cell death is associated with Lnc-D63785 m6A methylation and miR-422a accumulation. *Cell Death Dis.* 11 (9), 816. doi:10.1038/s41419-020-03021-8
- Xu, Y., Zhang, M., Zhang, Q., Yu, X., Sun, Z., He, Y., et al. (2021). Role of main RNA methylation in hepatocellular carcinoma: N6-Methyladenosine, 5-methylcytosine, and N1-methyladenosine. *Front. Cell Dev. Biol.* 9, 767668. doi:10.3389/fcell.2021.767668
- Yang, X., Wang, P., Yan, S., and Wang, G. (2022). Study on potential differentially expressed genes in stroke by bioinformatics analysis. *Neurol. Sci.* 43 (2), 1155–1166. doi:10.1007/s10072-021-05470-1
- 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 (5), 284–287. doi:10.1089/omi.2011.0118
- Yu, J., She, Y., and Ji, S. J. (2021). m(6)A modification in mammalian nervous system development, functions, disorders, and injuries. *Front. Cell Dev. Biol.* 9, 679662. doi:10.3389/fcell.2021.679662
- Yue, H., Yang, B. O., Yang, F., Hu, X. L., and Kong, F. B. (2016). Co-expression network-based analysis of hippocampal expression data associated with Alzheimer's disease using a novel algorithm. *Exp. Ther. Med.* 11 (5), 1707–1715. doi:10.3892/etm.2016.3131
- Zaccara, S., Ries, R. J., and Jaffrey, S. R. (2019). Reading, writing and erasing mRNA methylation. *Nat. Rev. Mol. Cell Biol.* 20 (10), 608–624. doi:10.1038/s41580-019-0168-5
- Zera, K. A., and Buckwalter, M. S. (2020). The local and peripheral immune responses to stroke: Implications for therapeutic development. *Neurotherapeutics* 17 (2), 414–435. doi:10.1007/s13311-020-00844-3

- Zhang, H., Meltzer, P., and Davis, S. (2013). RCircos: an R package for Circos 2D track plots. *BMC Bioinforma.* 14, 244. doi:10.1186/1471-2105-14-244
- Zhang, Q., Liu, C., Shi, R., Zhou, S., Shan, H., Deng, L., et al. (2022). Blocking C3d(+)/GFAP(+) A1 astrocyte conversion with semaglutide attenuates blood-brain barrier disruption in mice after ischemic stroke. *Aging Dis.* 13 (3), 943–959. doi:10.14336/ad.2021.1029
- Zhang, Y. P., Cui, Q. Y., Zhang, T. M., Yi, Y., Nie, J. J., Xie, G. H., et al. (2020a). Chloroquine pretreatment attenuates ischemia-reperfusion injury in the brain of ob/ob diabetic mice as well as wildtype mice. *Brain Res.* 1726, 146518. doi:10.1016/j.brainres.2019.146518
- Zhang, Z., Wang, Q., Zhao, X., Shao, L., Liu, G., Zheng, X., et al. (2020b). YTHDC1 mitigates ischemic stroke by promoting Akt phosphorylation through destabilizing PTEN mRNA. *Cell Death Dis.* 11 (11), 977. doi:10.1038/s41419-020-03186-2
- Zhao, B. S., Roundtree, I. A., and He, C. (2017). Post-transcriptional gene regulation by mRNA modifications. *Nat. Rev. Mol. Cell Biol.* 18 (1), 31–42. doi:10.1038/nrm.2016.132
- Zhou, M., Wang, H., Zeng, X., Yin, P., Zhu, J., Chen, W., et al. (2019). Mortality, morbidity, and risk factors in China and its provinces, 1990–2017: A systematic analysis for the global burden of disease study 2017. *Lancet* 394 (10204), 1145–1158. doi:10.1016/s0140-6736(19)30427-1
- Zhou, Y., Kong, Y., Fan, W., Tao, T., Xiao, Q., Li, N., et al. (2020). Principles of RNA methylation and their implications for biology and medicine. *Biomed. Pharmacother.* 131, 110731. doi:10.1016/j.biopha.2020.110731
- Zhu, M., Sun, H., Cao, L., Wu, Z., Leng, B., and Bian, J. (2022). Role of Na(+)/K(+)-ATPase in ischemic stroke: In-depth perspectives from physiology to pharmacology. *J. Mol. Med.* 100 (3), 395–410. doi:10.1007/s00109-021-02143-6



OPEN ACCESS

EDITED BY

Clévia Rosset,
Clinical Hospital of Porto Alegre, Brazil

REVIEWED BY

Igor Araujo Vieira,
Federal University of Rio Grande do
Sul, Brazil
Jun Li,
Sichuan University, China

*CORRESPONDENCE

Chunfang Zan
zancf2710@hotmail.com;
chunfang.zan@med.uni-muenchen.de
Zengliang Wang
wangzengliang1978@163.com

†These authors have contributed
equally to this work

SPECIALTY SECTION

This article was submitted to
Neurogenomics,
a section of the journal
Frontiers in Neuroscience

RECEIVED 26 July 2022

ACCEPTED 10 October 2022

PUBLISHED 11 November 2022

CITATION

Zan C, Li J, Lin F and Wang Z (2022)
Potential value of differentially
expressed circular RNAs derived from
circulating exosomes
in the pathogenesis of rat spinal cord
injury.
Front. Neurosci. 16:1003628.
doi: 10.3389/fnins.2022.1003628

COPYRIGHT

© 2022 Zan, Li, Lin and Wang. This is
an open-access article distributed
under the terms of the [Creative
Commons Attribution License \(CC BY\)](#).
The use, distribution or reproduction in
other forums is permitted, provided
the original author(s) and the copyright
owner(s) are credited and that the
original publication in this journal is
cited, in accordance with accepted
academic practice. No use, distribution
or reproduction is permitted which
does not comply with these terms.

Potential value of differentially expressed circular RNAs derived from circulating exosomes in the pathogenesis of rat spinal cord injury

Chunfang Zan^{1†}, Jianan Li^{2†}, Fengsong Lin² and
Zengliang Wang^{2*}

¹Division of Vascular Biology, Institute for Stroke and Dementia Research (ISD), LMU Klinikum, Ludwig-Maximilian-University (LMU), Munich, Germany, ²Department of Orthopedics, Tianjin Hospital, Tianjin, China

Spinal cord injury (SCI) remains one kind of devastating neurological damage, and specific molecular mechanisms involved need to be understood deeply. Currently, circular RNAs (circRNAs), as a newly discovered type of non-coding RNAs (ncRNAs), have been under active investigation. Through functional interactions with disease-associated microRNAs (miRNAs), exosome-derived circRNAs have been extensively implicated in various organ pathogenesis. Nevertheless, the functional involvement of circulating circRNAs in SCI onset, progression as well as repair remains poorly explored until now. Of note, there still lacks clinical and experimental evidence in this regard. To obtain some relevant knowledge in this field, this study was originally designed to have a general overview of differentially expressed circRNAs derived from circulating exosomes in SCI rats in comparison with the control rats. It turned out that 709 types of downregulated circRNAs and 346 kinds of upregulated circRNAs were preliminarily screened out. Functional enrichment analyses including Kyoto encyclopedia of genes and genomes (KEGG) pathway and gene ontology (GO) were performed to evaluate the possible biological functions of upregulated as well as downregulated circRNAs involved in SCI. Furthermore, five types of upregulated circulating circRNAs including chr4:208359914–208362182+, chr15:20088296–20092102+, chr1:175098934–175134845–, chr1:175099657–175128203–, and chr1:175104454–175134845–, and plus five kinds of downregulated circulating circRNAs including chr11:74154652–74159524–, chr12:45412398–45412635–, chr7:137630261–137648924–, chr6:6280974–6281188+, and chr4:225251864–225254087+, were verified through reverse transcription-polymerase chain reaction (RT-PCR). At last, taking these differentially expressed circRNAs in the center, the circRNA-miRNA-mRNA gene interaction network was constructed to predict the possible functionalities of circRNAs in SCI through anticipating

specific interactive miRNAs, giving new insights into how circRNAs contribute to this pathological process. Taken together, these findings suggest the possible involvement and functional significance of circRNAs in SCI.

KEYWORDS

spinal cord injury, exosome, ncRNA, circRNA, miRNA, mRNA, ceRNA

Introduction

Spinal cord injury (SCI) remains a catastrophic injured condition for humans, and causes impaired mobility and even permanent neurological deficits, thereby bringing serious social and economic burdens (Ahuja et al., 2017; David et al., 2019; Zipser et al., 2022). Given that there is still a lack of effective therapies for SCI until now, fundamental research is under active investigation in order to provide some promising targets for controlling SCI progression, exacerbation as well as recovery (Yamazaki et al., 2020; Kopper and Gensel, 2021). In terms of mechanisms, the pathological changes of the injured spinal cord are characterized by traumatic injury, ischemia, chronic neuroinflammation, and redox imbalance (Anjum et al., 2020; Alcántar-Garibay et al., 2022). Recently, non-coding RNAs (ncRNAs), for example, microRNAs (miRNAs) and long non-coding RNAs (lncRNAs), have been identified as pivotal players in SCI (Kimura et al., 2021; Liu et al., 2022). Of note, the knockdown of certain lncRNAs such as TUG1, XIST, and ZFAS1 has shown a beneficial effect on SCI development through mouse as well as rat model studies (Chen Y. et al., 2021; Wu et al., 2021; Zhong et al., 2021), indicating the functional involvement and preclinical significance of these ncRNAs in SCI.

Distinct from traditional linear RNAs, circular RNAs (circRNAs), as a new class of ncRNAs, are more stable, not easy to degrade, and not affected by RNA exonuclease due to their unique closed annular structure (Saaoud et al., 2021; Liu and Chen, 2022). In the functional sense, circRNAs have been found to regulate transcription, splicing, cytoplasmic mRNA stability and translation, interference with signaling pathways, etc. Of interest, circRNAs contain plenty of miRNA binding sites and thereby act as miRNA sponges, further counteracting the inhibitory effects of miRNAs on their target genes and meanwhile upregulating their corresponding target gene expression (Liu and Chen, 2022; Misir et al., 2022). This regulatory process involves the competing endogenous RNA (ceRNA) mechanism. Especially, the translation of circRNAs can produce some novel isoforms of proteins, which further determines their clinical significance in human diseases (Wen et al., 2022).

On the one hand, exosomes, as mediators of cell-to-cell communication as well as vehicles of circRNAs and other

ncRNAs, are key regulators in specific molecular signaling (Isaac et al., 2021). On the other hand, circRNAs are one class of the most abundant components in exosomes (Li et al., 2021; Jafari et al., 2022). Therefore, exosomal circRNAs attract much attention from researchers in the context of SCI. In a similar vein, Han et al. (2022) gave a comprehensive review about separated/cooperative biological functions of circRNAs and exosomes, highlighting the potential of these exosomal circRNAs in disease states. So far, exosomal circRNAs have been found to participate in tumorigenesis, and cardiovascular inflammation as well as metabolic disorders through targeting the corresponding miRNA-mediated axis (Li et al., 2021; Lin et al., 2021). Even so, the role of exosomal circRNAs in the injured spinal cord has not been well elucidated, which needs to be explored further.

For this reason, we mainly focus on circulating exosome-derived circRNAs in this study, and aim to investigate their potential role and relevant functions in the context of SCI. To this end, we utilized rat SCI model and collected exosomes, procedures and results of which have been well described in our recently published paper (Li et al., 2023). Based on this, exosomal circRNAs were obtained, functional enrichment analyses including kyoto encyclopedia of genes and genomes (KEGG) pathway and gene ontology (GO) ontology were performed, and several differentially expressed circRNAs were confirmed through reverse transcription-polymerase chain reaction (RT-PCR). Furthermore, the interaction network of circRNA-miRNA-mRNA would better explain the predicted functions of exosomal circRNAs in the injured spinal cord based on the ceRNA mechanism.

Materials and methods

Rat spinal cord injury model establishment and exosome extraction and identification

The successful establishment of rat SCI model, the isolation of blood exosomes, and the identification of exosomes were detailedly described in our recently published paper (Li et al., 2023). In brief, six adult female rats were randomly divided

into two groups, the experimental group (EG) in which SCI surgery was performed on rats, and the control group (CG) in which the skin was injured on rats. Both cohorts of rats were maintained under the same feeding conditions before and after surgery. After 24 h, blood was collected and exosomes were extracted by using the density gradient centrifugation. Then, the classical characteristics of isolated exosomes were verified through Western blot, electron microscopy, and nanoparticle tracking analysis (NTA). Related data on this part would be referred to our previous publication (Li et al., 2023).

Exosomal RNA isolation

According to the manufacturer's protocol, Trizol reagent (Thermo Fisher Scientific, Waltham, MA, USA) was used to obtain high-quality RNA from circulating exosomes. Concentrations of different RNA samples were measured through using Nanodrop Spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA). Then the quantitative control as well as the integrity control were performed for isolated RNA prior to experiments in the next step.

In the process of quantitative control, the purity of RNA was evaluated by using OD260/OD280 values. The final results would be considered as "Pass", if these OD260/OD280 values are within the range between 1.8 and 2.1, suggesting isolated RNA has high quality and can be used further.

Regarding the integrity control, 1% agarose gel (Thermo Fisher Scientific, Waltham, MA, USA) was utilized to assess the integrity of isolated RNA in this study.

RNA library construction and RNA sequencing

Prior to RNA library construction, ribosomal RNA (rRNA) needs to be completely removed from total RNA. To this end, NEBNext rRNA Depletion Kit (New England Biolabs, Ipswich, MA, USA) was used according to the producer's protocol. Following this, the RNA library was then constructed through using NEBNext® Ultra™ II Directional RNA Library Prep Kit (New England Biolabs, Ipswich, MA, USA) according to the standard procedures. Furthermore, RNA library was quality-controlled and then quantified *via* a BioAnalyzer 2100 system (Agilent Technologies, Santa Clara, CA, USA). In the end, circRNA sequencing was conducted on an illumina HiSeq sequencer from Cloud-Seq Biotech (Shanghai, China). All raw data were uploaded to GEO repository as the GSE213561 study (GSM6589180; GSM6589181; GSM6589182; GSM6589183; GSM6589184; GSM6589185), and the link is as follows: <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE213561>.

RNA sequencing data analysis

As described in the above part, raw data which contain original paired-end reads were initially acquired from Illumina HiSeq 4000 systems (Cloud-Seq Biotech, Shanghai, China). Through using Cutadapt software,¹ 3' adaptor was modified, and low-quality reads were removed in order to keep high-quality reads for further analysis (Han et al., 2020). To display relative levels of these circulating circRNAs, these modified reads were analyzed. HISAT2 software² was used to make an alignment of these pre-selected reads to the human reference genome. Next, according to gene transfer format (GTF) files which were obtained from the Ensembl database,³ the expression levels of circRNAs, characterized by the fragments per kilobase of exon model per million mapped fragments (FPKM), were calculated *via* Cuffdiff software (Trapnell et al., 2010). Following this, specific fold changes and corresponding *P*-values were calculated based on the FPKM index, and differentially expressed circRNAs from circulating exosomes were eventually determined.

Moreover, in order to predict their probably involved pathways and functional processes, KEGG pathway analysis⁴ and GO analysis⁵ were performed for circulating upregulated as well as downregulated circRNAs, respectively (Kanehisa and Sato, 2020). The value $-\log_{10}(P\text{-value})$ was calculated and presented as the enrichment score. For both KEGG pathway and GO analysis, fold change > 2 and *P*-value < 0.05 were taken as the threshold of the differential expression of these circRNAs.

Determination of differentially expressed circular RNAs through reverse transcription-polymerase chain reaction

Next, we would like to separately verify the top ten dysregulated circulating circRNAs in SCI rats, which were revealed by RNA sequencing previously. To this end, RT-PCR was utilized here to show the relative levels of these 10 circulating exosomal circRNAs from both group rats. As mentioned in the above part, Trizol reagent was used to isolate RNA, and then cDNA was synthesized by utilizing SuperScript™ IV First-Strand synthesis kit (Thermo Fisher Scientific, Waltham, MA, USA). For RT-PCR process, SYBR Green master mix (CloudSeq, Shanghai,

¹ <https://cutadapt.readthedocs.io/>

² <http://daehwankimlab.github.io/hisat2/>

³ <http://www.ensembl.org/>

⁴ <http://www.genome.jp/kegg>

⁵ <http://www.geneontology.org>

China) was additionally employed. Primer 5.0 software (PREMIER Biosoft, Palo Alto, CA, USA) was used to design all primer sequences for circRNA targets as well as the housekeeping gene, i.e., glyceraldehyde-3-phosphate dehydrogenase (GAPDH). All primers were purchased from CloudSeq Biotechnology (Shanghai, China). The primer sequences are shown in Table 1. All raw data were acquired by QuantStudio 5 Real-Time PCR System (Thermo Fisher Scientific, Waltham, MA, USA), and further $2^{-\Delta\Delta Ct}$ method was applied to show the relative expression of these circulating circRNAs. Given that most of differentially expressed circRNA are novel, we used their chromosome locations to name them.

Construction of circRNA-miRNA-mRNA network

Considering the importance of functional interactions among circRNAs, mRNAs and miRNAs during the gene regulation, the upstream as well as the downstream on the genome of these differential expressed circRNAs were extended in order to find functional genes near these circRNAs. To this end, corresponding miRNAs were predicted through employing miRNA target

prediction software miRanda⁶ based on these differentially expressed circRNAs, and then potential mRNAs were further anticipated by applying Targetscan⁷ as well as miRDB⁸ (Sun et al., 2020). In the end, circRNA-miRNA-mRNA interaction network was established and visualized via Cytoscape⁹ (Zhao et al., 2021), where these differentially expressed circRNAs from circulating exosomes were centered.

Statistical analysis

GraphPad Prism 8 software was utilized to make statistical analyses for these experimental results. Data were shown as means \pm standard deviation (SD) from more than three statistical independent experiments. To compare the statistical significances among different groups, *t*-test of two independent samples was utilized, respectively. When *P*-value < 0.05, the difference was regarded as statistically significant. Especially, **P* < 0.05, ***P* < 0.01, ****P* < 0.001, *****P* < 0.0001.

6 http://cbio.mskcc.org/microna_data/miRanda-aug2010.tar.gz

7 https://www.targetscan.org/vert_80/

8 <http://www.mirdb.org/>

9 <https://cytoscape.org/>

TABLE 1 The primer sequences for circular RNA (circRNA) identification.

| Genes | | Sequences | |
|-------|---------------------------|-----------|-----------------------|
| 1 | chr11:74154652–74159524– | Forward | CATCTCCTACGCTTGCCTGA |
| | | Reverse | CCAGAGAAACAAAGTGGCACC |
| 2 | chr12:45412398–45412635– | Forward | GAGACGAACCAACCTGGTG |
| | | Reverse | CGTGCCCTCCAAAATTGTACC |
| 3 | chr4:208359914–208362182+ | Forward | ACTGGTGTGAATACTCGGCG |
| | | Reverse | CTGTATGGGCAATTCCGGT |
| 4 | chr7:137630261–137648924– | Forward | TCATGGGCAGTGGGATCTTG |
| | | Reverse | GCAGTGAATAGAGATGCCCGA |
| 5 | chr15:20088296–20092102+ | Forward | TGTGTGAGGCCTTGGTTTGA |
| | | Reverse | CCACAAGTCCGTATCTTTGGC |
| 6 | chr1:175098934–175134845– | Forward | ACCTGGGCAAGGAATTCACC |
| | | Reverse | GTGGTACTTGTGAGCCAGGG |
| 7 | chr1:175099657–175128203– | Forward | TCCACTGTGACAAGCTGCAT |
| | | Reverse | GAATTCCTTGCCAGGTGGT |
| 8 | chr1:175104454–175134845– | Forward | TTATGATGGGCCACCACCTG |
| | | Reverse | GTGGTACTTGTGAGCCAGGG |
| 9 | chr6:6280974–6281188+ | Forward | ACTCTGGAGAACTCGGGAT |
| | | Reverse | CCAGCTGCTACTTGCTCAGT |
| 10 | chr4:225251864–225254087+ | Forward | GGAGCTGGAGAAGGACTTGG |
| | | Reverse | CCTCTTGCCATTGTCCGTGA |
| 11 | GAPDH | Forward | GACATGCCGCTGGAGAAAC |
| | | Reverse | AGCCCAGGATGCCCTTTAGT |

Results

Expression profile analysis of circular RNAs derived from circulating exosomes

The experimental design of this study was shown in **Figure 1**. Part data of preliminary work including SCI rat model establishment and exosome identification were presented in our recently published paper (Li et al., 2023).

Through enrichment analyses, a total of 709 types of circRNAs were found to be downregulated whereas 346 kinds of circRNAs were shown to be upregulated in SCI rats in comparison with the control rats. In this experiment, three SCI rats and plus three control rats were included. Through the normalization as FPKM, the heat map of differentially expressed circRNAs from circulating exosomes including 15 kinds of downregulated circRNAs and 7 kinds of upregulated circRNAs was especially visualized (**Figure 2A**), suggesting that there exists a significantly different expression pattern of circulating exosomal circRNAs between SCI rats and the control rats ($P < 0.05$). In a similar vein, the volcano plots demonstrated consistent changes of differentially expressed circRNAs in SCI rats (**Figure 2B**). Taken together, these data give the first impression that some circulating circRNAs display differential expression in SCI rats.

Enrichment analyses of potentially upregulated circular RNAs

To further speculate their specific functional activities and involved signaling pathways in the context of SCI, both GO analysis and KEGG analysis were next carried out in this study. **Figures 3, 4** together demonstrated the top enriched GO terms and KEGG pathways for these differentially expressed circRNAs. Especially, upregulated circRNAs were analyzed in **Figure 3**, and downregulated circRNAs were analyzed in **Figure 4**. These results together provide an overview of these probably involved cellular components, biological processes, and molecular functions as well as signaling pathways for differentially expressed circRNAs in SCI rats (**Figures 3, 4**), which may offer some meaningful indications for future research.

The top five enriched biochemical pathways for upregulated circRNAs were African trypanosomiasis, Malaria, non-small lung cancer, small cell lung cancer, and purine metabolism, as demonstrated by KEGG pathway analysis based on located genes (**Figure 3A**). On the other hand, GO enrichment analysis showed that these upregulated circRNAs were enriched in plenty of biological processes associated with inflammation, cellular components, and molecular functions. The top five enriched terms of biological process were oxygen

transport, negative regulation of proteolysis involved in cellular protein catabolic process, negative regulation of proteolysis, negative regulation of protein catabolic process, and negative regulation of proteasomal ubiquitin-dependent protein catabolic process (**Figure 3B**). The top four enriched terms of cellular component were hemoglobin complex, cytosolic part, cytosol, and cytoplasmic part (**Figure 3C**). At last, oxygen transporter activity, oxygen binding, heme binding, tetrapyrrole binding, and transition metal ion binding were the top five enriched terms of molecular function (**Figure 3D**). Of special interest, the potential involvement of some circRNAs in tumors may indicate the potency of circRNA-elicited microenvironment changes in inflammatory processes.

Enrichment analyses of potentially downregulated circular RNAs

The top eight enriched biochemical pathways were GABAergic synapse, ErbB signaling pathway, mRNA surveillance pathway, T cell receptor signaling pathway, glutamatergic synapse, thyroid hormone signaling pathway, axon guidance and RNA transport, as demonstrated by KEGG pathway analysis based on located genes (**Figure 4A**). In addition, GO enrichment analysis showed that these downregulated circRNAs were enriched in plenty of biological processes associated with inflammation, cellular components, and molecular functions. The top five enriched terms of biological process were transcription from RNA polymerase II promoter, RNA metabolic process, regulation of transcription, DNA-templated, and positive regulation of T cell activation (**Figure 4B**). The top five enriched terms of cellular component were SWI/SNF superfamily-type complex, SWI/SNF complex, proteasome complex, pericentric heterochromatin, and npBAF complex (**Figure 4C**). At last, transcription factor binding transcription factor activity, transcription factor activity, RNA polymerase II activating transcription factor binding, protein binding transcription activity, and protein-lysine *N*-methyltransferase activity were the top five enriched terms of molecular function (**Figure 4D**). In combination with data for upregulated circRNAs, these bioinformatics data together suggest there are different pathways and functions involved for these dysregulated circRNAs.

Verification of differentially expressed circular RNAs through reverse transcription-polymerase chain reaction

As addressed in the above part, several significantly upregulated and downregulated circRNAs were preliminarily

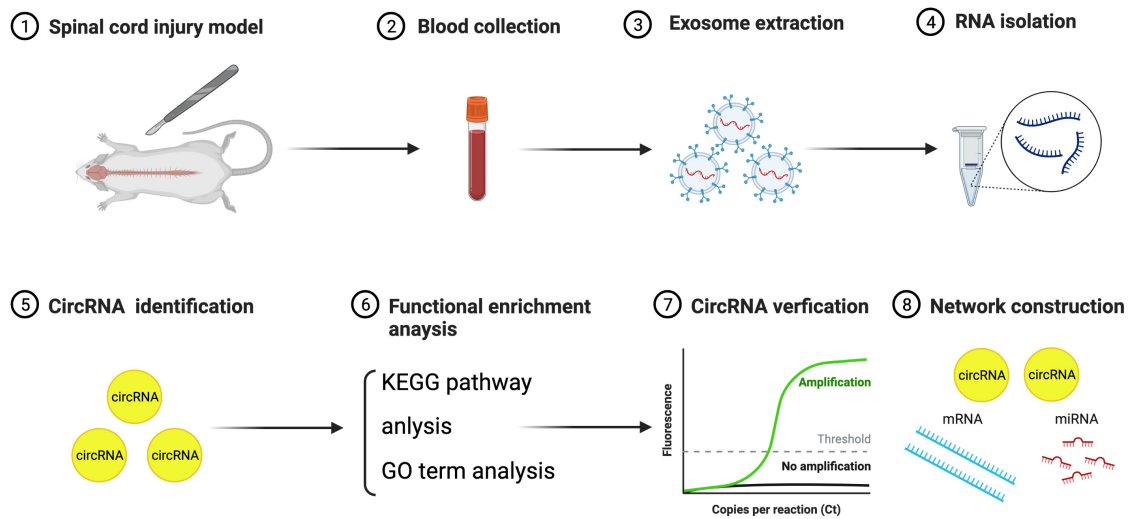


FIGURE 1
The flowchart of this study.

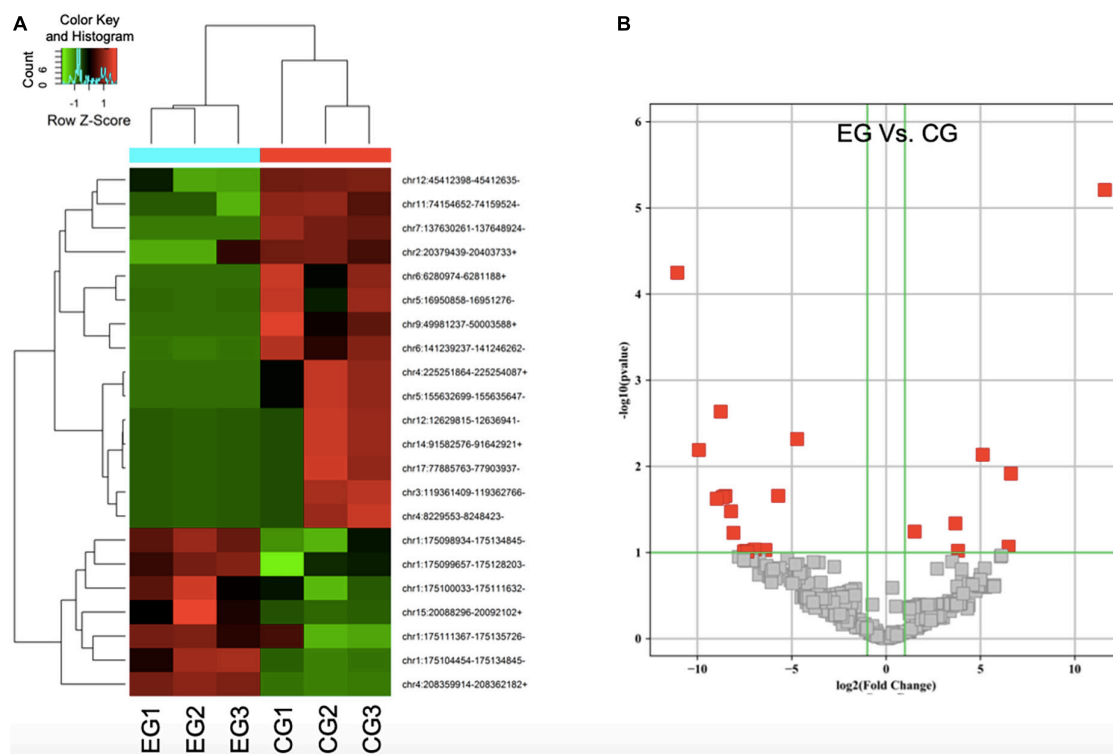


FIGURE 2
The expression profile of differentially expressed circular RNAs (circRNAs) in spinal cord injury (SCI) rats is identified. (A) Heat map of differentially expressed circRNAs from circulating exosomes. (B) Two-dimensional presentation for cluster analysis of differentially expressed circRNAs from circulating exosomes. Red represents relatively high expression, green represents relatively low expression, and black represents the average expression. $P < 0.05$, fold change > 2.0 .

selected through RNA sequencing (Figure 2). Later on, the potential roles of these circulating circRNAs in SCI rats were predicted *via* KEGG and GO analyses (Figures 3, 4).

Therefore, we would like to take advantage of qRT-PCR assay, to verify whether these candidate circRNAs are significantly dysregulated indeed. It turned out that there

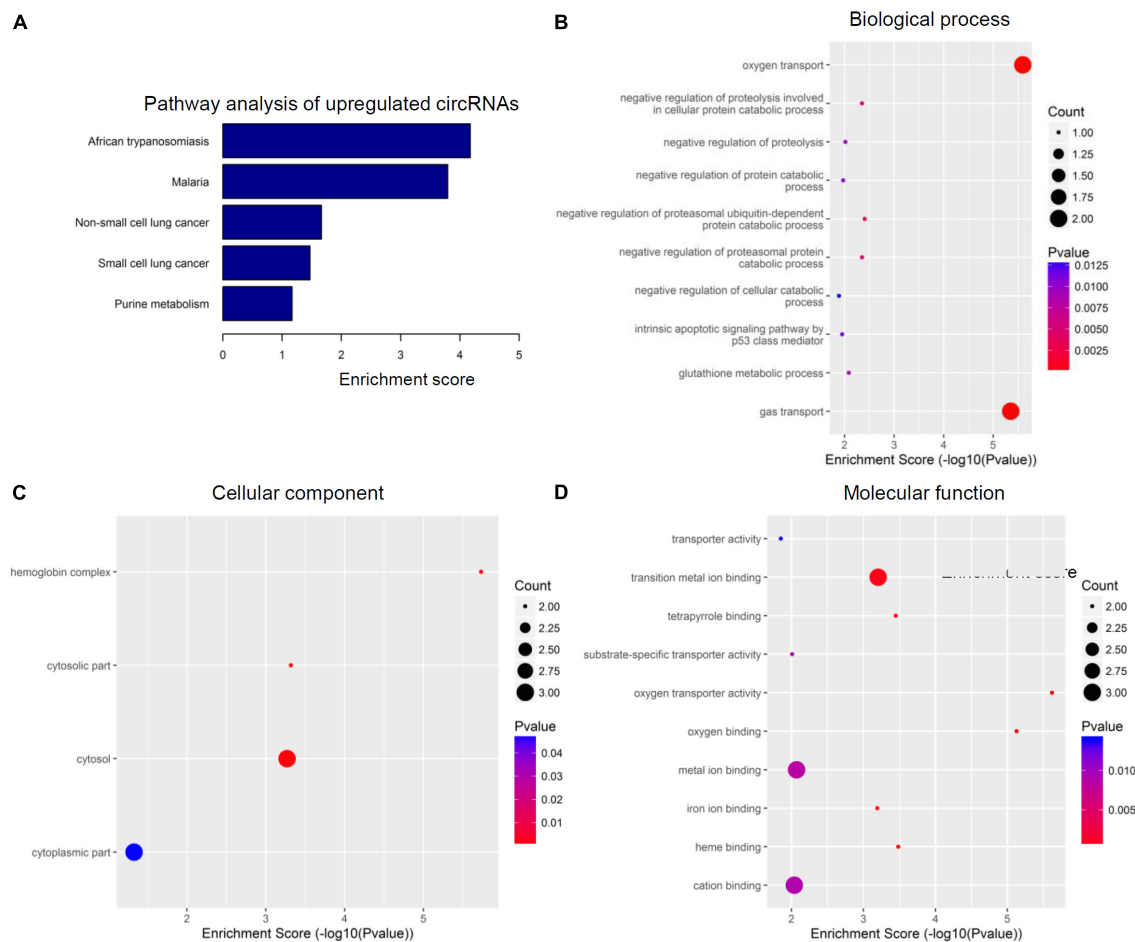


FIGURE 3

Kyoto encyclopedia of genes and genomes (KEGG) pathway analysis and gene ontology (GO) function analysis of upregulated circular RNAs (circRNAs) from circulating exosomes. (A) KEGG pathway analysis of upregulated circRNAs from circulating exosomes. (B–D) GO function analysis of upregulated circRNAs from circulating exosomes. (C) Cellular component analysis of upregulated circRNAs from circulating exosomes. (D) Molecular function analysis of upregulated circRNAs from circulating exosomes.

were a total of ten kinds of differentially expressed circRNAs, which were further verified by qRT-PCR in this study. Quantification results displayed that the relative expression of chr4:208359914–208362182+, chr15:20088296–20092102+, chr1:175098934–175134845–, chr1:175099657–175128203–, and chr1:175104454–175134845–, was significantly elevated in SCI rats in comparison with the control rats (Figure 5A). By contrast, the relative levels of chr11:74154652–74159524–, chr12:45412398–45412635–, chr7:137630261–137648924–, chr6:6280974–6281188+, and chr4:225251864–225254087+ were dramatically diminished in SCI rats compared to the controls (Figure 5B). Moreover, the detailed information of these ten candidate circRNAs is summarized in Table 2. Given that most of differentially expressed circRNA are novel and the gene symbol of three kinds of circRNA is the same, we named them according to their chromosome locations. Relevant predicted pathway information of each

circRNA can be checked through circBase¹⁰ (Table 2). These quantitative results not only provide some promising targets for future *in vivo* and *in vitro* research, but also validate the expression profile of circulating circRNAs indicated by RNA sequencing.

Construction of circRNA-miRNA-mRNA interaction network to predict functions of differentially expressed circular RNAs

Based on ceRNA mechanism, we would like to take advantage of circRNA-miRNA-mRNA interaction network

¹⁰ <http://www.circbase.org/>

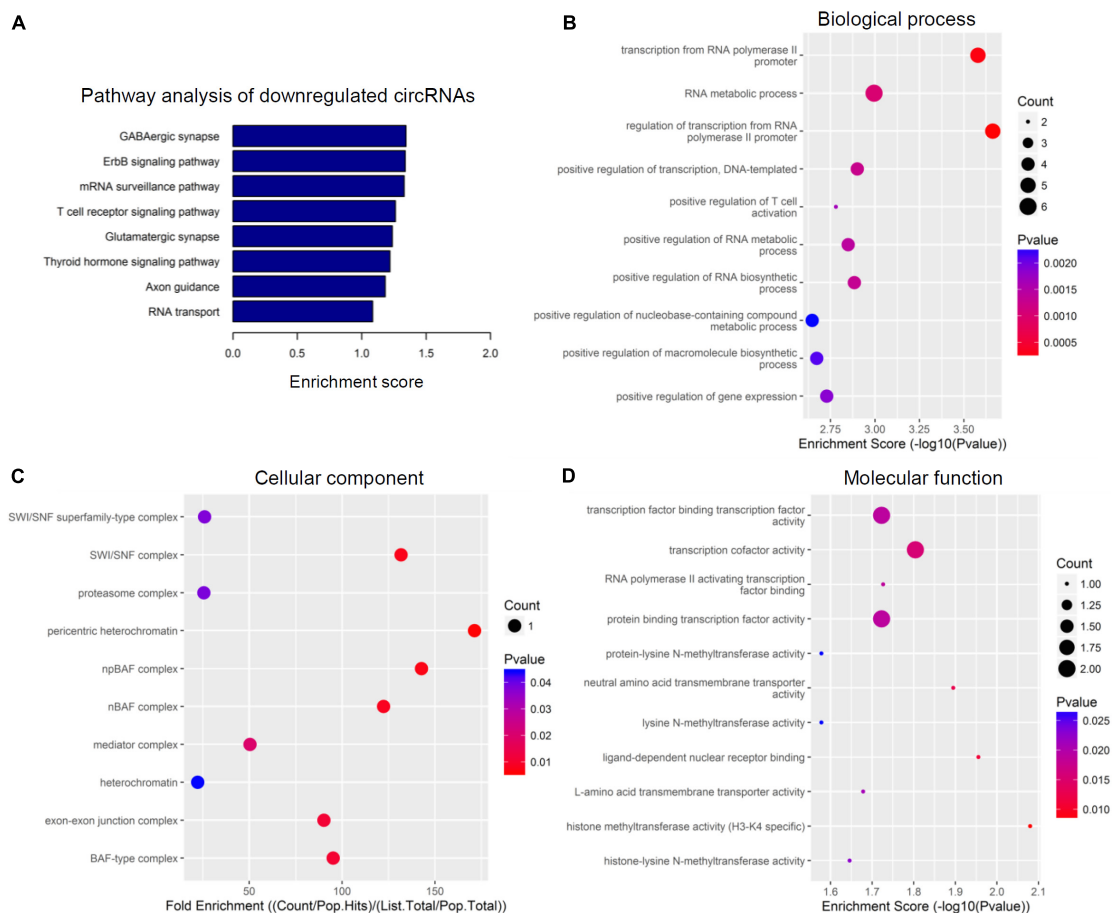


FIGURE 4

Kyoto encyclopedia of genes and genomes (KEGG) pathway analysis and gene ontology (GO) function analysis of downregulated circular RNAs (circRNAs) from circulating exosomes. (A) KEGG pathway analysis of downregulated circRNAs from circulating exosomes. (B–D) GO function analysis of downregulated circRNAs from circulating exosomes. (B) Biological process analysis of downregulated circRNAs from circulating exosomes. (C) Cellular component analysis of downregulated circRNAs from circulating exosomes. (D) Molecular function analysis of downregulated circRNAs from circulating exosomes.

to anticipate potential functions of circRNAs in the injured spinal cord. As shown in **Figure 6**, five kinds of upregulated circRNAs as well as five kinds of downregulated circRNAs verified by RT-PCR were chosen and centered in this network. These circRNAs could co-express and interact with 37 miRNAs which are indicated by red triangles and 145 mRNAs which are indicated by blue boxes (**Figure 6**). All these upregulated circRNAs had four to six target miRNAs, as exemplified by chr4:208359914–208362182+ which has five target miRNAs including rno-miR-26b-3p, rno-miR-17-1-3p, rno-miR-3068-5p, rno-miR-3564, and rno-miR-20b-3p (**Figure 6**). In the meanwhile, these downregulated circRNAs such as chr11:74154652–74159524–, chr12:45412398–45412635–, chr7:137630261–137648924–, chr6:6280974–6281188+, and chr4:225251864–225254087+ had three to six target miRNAs generally (**Figure 6**). Thus, these bioinformatics data together imply the potential important functions of the

above-mentioned circulating circRNAs in regulating other genes.

Discussion

In the current study, bioinformatics analyses of rat specimens revealed 709 downregulated circRNAs and 346 upregulated circRNAs from circulating exosomes in SCI rats in comparison with the control rats, and RT-PCR further confirmed their differential expression. These results together highlight the potential involvement of exosomal circRNAs in injured spinal cord.

In fact, there have been plenty of published studies revealing differentially expressed circRNAs during different stages of SCI (Qin et al., 2019; Wu et al., 2019; Liu et al., 2020; Peng et al., 2020). Liu et al. (2020) found that 1101 circRNAs were upregulated and 897 circRNAs were downregulated at

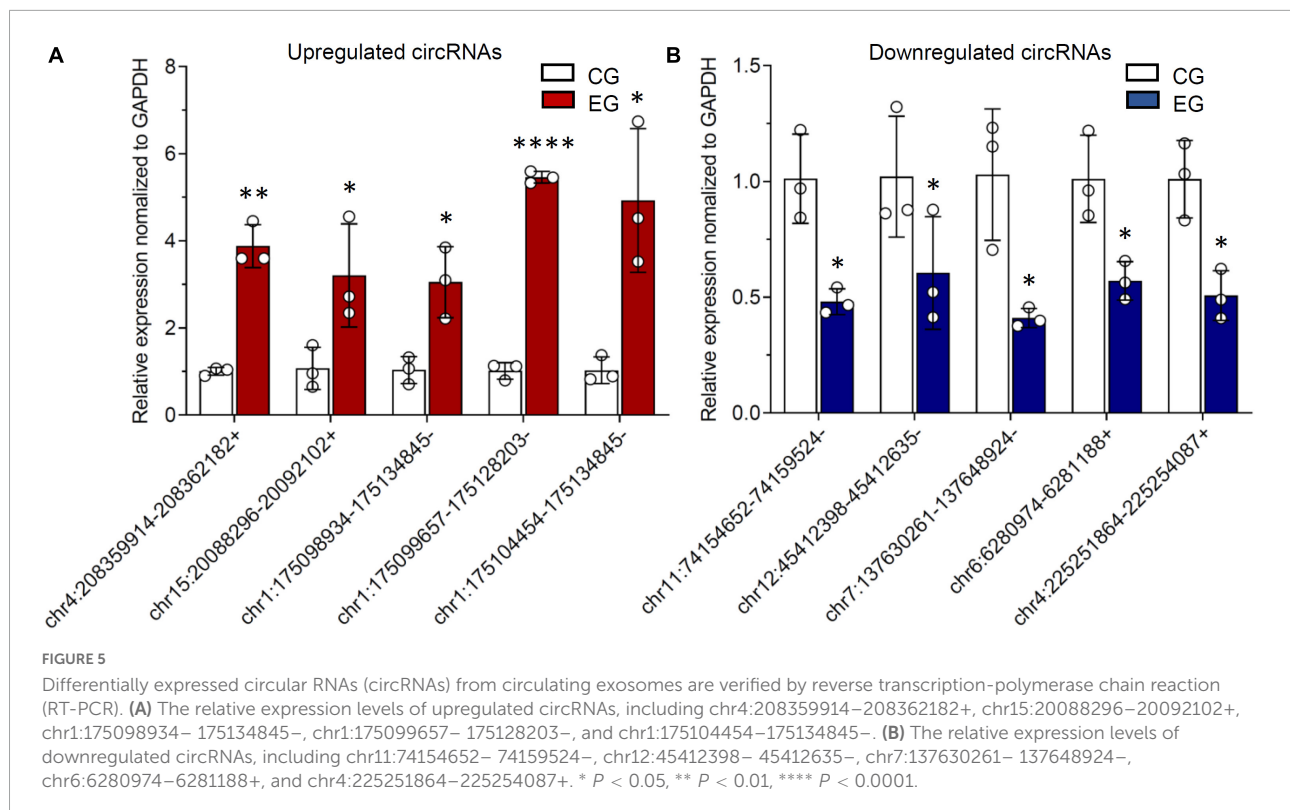


TABLE 2 Basic information of differentially expressed circular RNAs (circRNAs).

| Source | Position | Genomic length | Strand | Best transcript | Gene symbol | Catalog | Regulation |
|----------------|--------------------------|----------------|--------|-----------------|-------------|-------------------|------------|
| PMID: 25714049 | chr11:74154652–74159524 | 4,872 | – | XM_003751063 | Lrch3 | Exonic | Down |
| PMID: 25714049 | chr12:45412398–45412635 | 237 | – | XM_003751198 | Med13l | Exonic | Down |
| PMID: 25714049 | chr4:208359914–208362182 | 2,268 | + | NM_001106614 | Setd5 | Exonic | Up |
| PMID: 25714049 | chr7:137630261–137648924 | 18,663 | – | NM_138832 | Slc38al | Exonic | Down |
| Novel | chr15:20088296–20092102 | 3,806 | + | NM_021774 | Fhit | Intronic | Up |
| Novel | chr1:175098934–175134845 | 35,911 | – | NM_033234 | Hbb | Sense overlapping | Up |
| Novel | chr1:175099657–175128203 | 28,546 | – | NM_033234 | Hbb | Sense overlapping | Up |
| Novel | chr1:175104454–175134845 | 30,391 | – | NM_033234 | Hbb | Sense overlapping | Up |
| Novel | chr6:6280974–6281188 | 214 | + | XM_592846 | Loc103692 | Exonic | Down |
| Novel | chr4:225251864–225254087 | 2,223 | + | – | – | Intergenic | Down |

the immediate stage of SCI. Furthermore, they identified eight out of ten candidate circRNAs through RT-PCR, such as rno_circRNA_011494 and rno_circRNA_009608, which displayed different levels of dysregulated expression in SCI rats in comparison with the control group, suggesting several possible targets for early intervention of SCI (Liu et al.,

2020). In a similar vein, Wu et al. (2019) displayed a more broad expression profile of spinal cord-derived circRNAs at seven time points including immediate and chronic phases. Of note, they especially focused on exonic circRNA_01477, and demonstrated silencing of circRNA_01477 ameliorated astrocyte proliferation and migration through *in vitro* functional

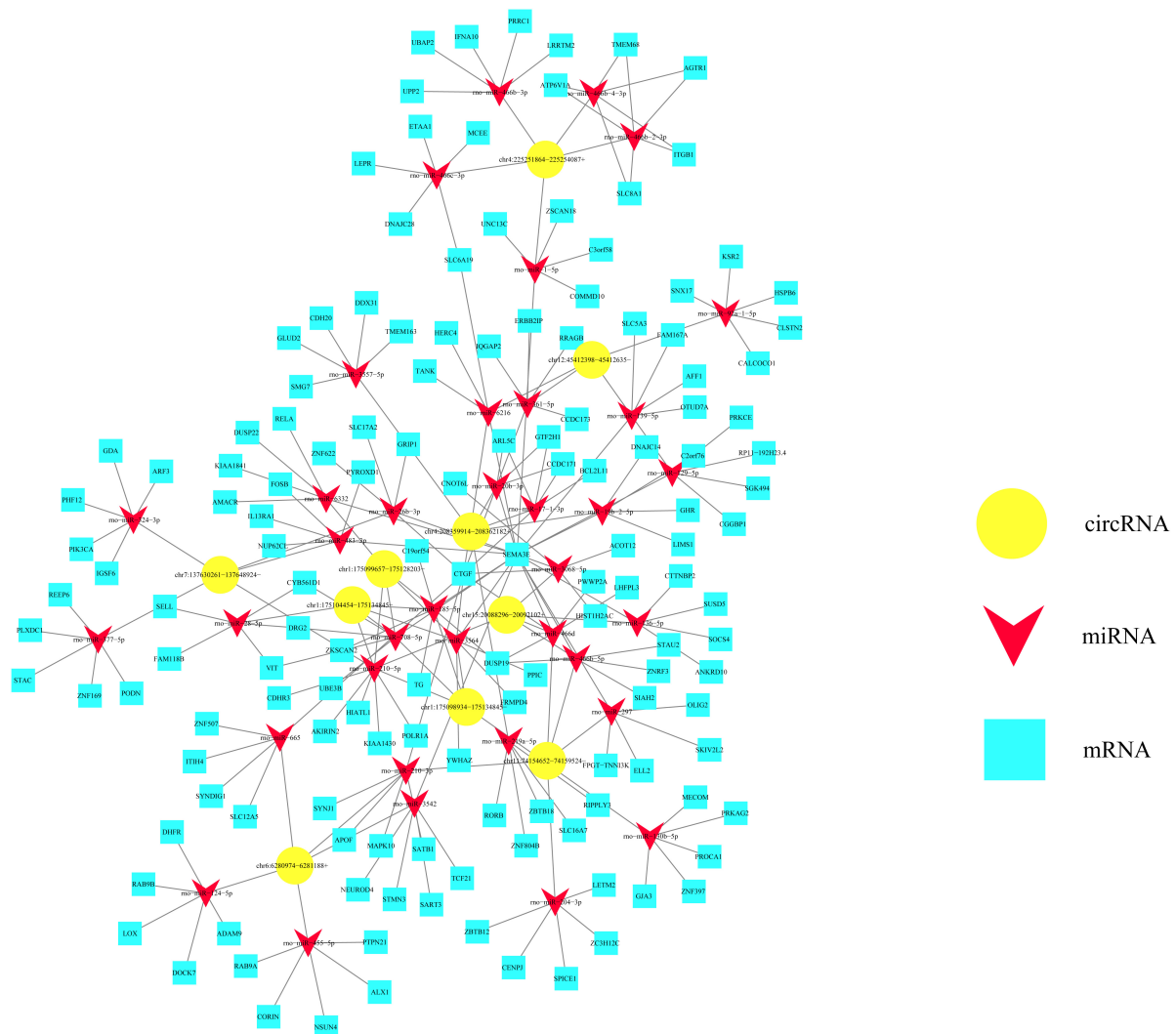


FIGURE 6

Potential value of differentially expressed circular RNAs (circRNAs) from circulating exosomes is predicted based on the circRNA-miRNA-mRNA interaction network. This interaction network is centered by 10 kinds of differentially expressed circRNAs, as indicated by yellow polygons. These circRNAs could co-express and interact with 37 miRNAs which are indicated by red triangles, and 145 mRNAs which are indicated by blue boxes.

assay. Mechanistically, circRNA_01477 depletion was associated with significantly reduced miRNA-423-5p expression (Wu et al., 2019), indicating that circRNA_01477/miRNA-423-5p may contribute to SCI progression. Taken together, the important role of circRNAs in SCI has been well established.

In the functional sense, circRNAs have been extensively implicated in SCI through multiple mechanisms, as systematically reviewed by Sámano et al. (2021) recently. Through SCI mouse as well as rat models, circular RNAs have been found to regulate proliferation and migration of vascular endothelial cells, angiogenesis, neuronal cell apoptosis, and neuroinflammation, based on ceRNA pattern (Zhao et al., 2020; Chen J. et al., 2021; Xie et al., 2021; Ye et al., 2021). Among these reported circRNAs, circRNA_014301 showed a

pro-inflammatory effect and a pro-apoptotic property upon lipopolysaccharide (LPS) stimulation, which could be reversed by its silencing in PC12 cells (Xie et al., 2021). It suggests the suppression of circRNA_014301 may be a promising strategy to control SCI progression. By contrast, circ-HIPK3 was found to be downregulated in SCI rats and further circ-HIPK3/miR-558/DPYSL5 axis protects against neuronal cell apoptosis in injured spinal cord (Zhao et al., 2020). Moreover, Chen Y. et al. (2021) found another novel circRNA-2960, which could aggravate inflammatory response and induce apoptosis through downregulating miRNA-124 at the lesion site. More recently, Qi et al. (2022) identified circ-Ctnnb1 as a potent regulator of neuronal injury in SCI through the Wnt/ β -catenin signaling pathway. Collectively, there have been plenty of well-studied

circRNAs in the context of SCI so far, and these circRNAs functionally contribute to SCI progression and even recovery.

However, the majority of current mechanistic studies utilized the spinal cord as the major specimen source and thereby investigated spinal cord-derived circRNAs after injury (Liu et al., 2020; Peng et al., 2020; Chen J. et al., 2021), as discussed in the above part. In addition to the spinal cord itself, circulating mediators from blood including inflammatory factors and bioactive lipids are also considered as important regulators for SCI development (Francos-Quijorna et al., 2017; David and López-Vales, 2021). On the one hand, sudden injury of spinal cord could induce inflammation, a complex response that contributes to secondary damage and even severe functional loss (Hellenbrand et al., 2021; Liu et al., 2021). On the other hand, exosomal circRNAs can participate in multiple inflammatory processes, such as NLRP3 inflammasome activation (Li M. et al., 2020), macrophage polarization and inflammation (Song et al., 2022), autoimmune response (Lodde et al., 2020). Given the importance of circRNAs in inflammation, we are curious about whether circulating circRNAs contribute to SCI. However, there are no relevant data yet. Therefore, our study for the first time addressed the potential value of circulating exosomal circRNAs in SCI progression. Newly identified differentially expressed circRNAs would be promising candidates for early intervention and treatment of SCI.

Of note, we applied enrichment analyses including KEGG pathway, GO, and interaction network prediction in this study, which are well established to anticipate potent genes involved in SCI (Li Z. et al., 2020; Chen Q. et al., 2021). Top enriched biochemical pathways and GO terms of dysregulated circRNAs could provide us some meaningful hints for further mechanistic study. More importantly, the circRNA-miRNA-mRNA interactive network sheds light on the functional involvement of circRNAs in SCI development (Peng et al., 2020; Tong et al., 2021; Wang et al., 2021). Our data demonstrated these differentially expressed circRNAs could co-express and interact with 37 miRNAs and 145 mRNAs. Previous studies have identified several circRNA-miRNA pairs in SCI, such as circAbca1/miR-135b, circRNA_01477/miR-423-5p, circ-HIPK3/miR-558 (Wu et al., 2019; Zhao et al., 2020; Wang et al., 2021). In this study, we also constructed the interaction network, which was centered by 10 kinds of differentially expressed circRNAs. These circRNAs could co-express and interact with 37 types of miRNAs, and 145 types of mRNAs. In combination with the previously published study, these bioinformatics results together kick off functional research of circulating exosomal circRNAs in this field.

In the meanwhile, we acknowledge that there exist several limitations in our study. Even though the expression profile of circulating circRNAs has been mapped in this study, we did not go deeper considering any kind of circRNA. Apparently, *in vitro* functional assays and pharmacological investigations are called for further mechanistic exploration. In addition,

concerning differences in circRNAs among species, the clinical significance of human circRNAs in patients should be clarified. Furthermore, solid experimental evidence for the interaction between circRNAs and corresponding miRNAs would bring new avenues for targeting SCI pathological process.

In conclusion, our data identify several circulating exosomal circRNAs in SCI rats, which show a differential expression pattern. These circRNAs are predicted to be involved in different pathways and functional processes, and thereby contribute to SCI development, providing some potential targets for future research on SCI.

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/>; GSE213561.

Ethics statement

This animal study was reviewed and approved by the Animal Care and Use Committee of Tijian Hospital.

Author contributions

JL, CZ, and ZW designed and conceptualized this study. JL and CZ wrote the manuscript. FL and ZW further polished the manuscript. All authors prepared and performed experiments, collected, analyzed, and interpreted data, read, and approved the 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.

References

- Ahuja, C. S., Wilson, J. R., Nori, S., Kotter, M., Druschel, C., Curt, A., et al. (2017). Traumatic spinal cord injury. *Nat. Rev. Dis. Primers* 3:17018. doi: 10.1038/nrdp.2017.18
- Alcántar-Garibay, O. V., Incontri-Abraham, D., and Ibarra, A. (2022). Spinal cord injury-induced cognitive impairment: A narrative review. *Neural Regen. Res.* 17, 2649–2654. doi: 10.4103/1673-5374.339475
- Anjum, A., Yazid, M. D. I., Fauzi Daud, M., Idris, J., Ng, A. M. H., Selvi Naicker, A., et al. (2020). Spinal cord injury: Pathophysiology, multimolecular interactions, and underlying recovery mechanisms. *Int. J. Mol. Sci.* 21:7533. doi: 10.3390/ijms21207533
- Chen, Y., Wei, Z., Liu, J., Xie, H., Wang, B., Wu, J., et al. (2021). Long noncoding RNA ZFAS1 aggravates spinal cord injury by binding with miR-1953 and regulating the PTEN/PI3K/AKT pathway. *Neurochem. Int.* 147:104977. doi: 10.1016/j.neuint.2021.104977
- Chen, J., Fu, B., Bao, J., Su, R., Zhao, H., and Liu, Z. (2021). Novel circular RNA 2960 contributes to secondary damage of spinal cord injury by sponging miRNA-124. *J. Comp. Neurol.* 529, 1456–1464. doi: 10.1002/cne.25030
- Chen, Q., Zhao, Z., Yin, G., Yang, C., Wang, D., Feng, Z., et al. (2021). Identification and analysis of spinal cord injury subtypes using weighted gene co-expression network analysis. *Ann. Transl. Med.* 9:466. doi: 10.21037/atm-21-340
- David, G., Mohammadi, S., Martin, A. R., Cohen-Adad, J., Weiskopf, N., Thompson, A., et al. (2019). Traumatic and nontraumatic spinal cord injury: Pathological insights from neuroimaging. *Nat. Rev. Neurol.* 15, 718–731. doi: 10.1038/s41582-019-0270-5
- David, S., and López-Vales, R. (2021). Bioactive lipid mediators in the initiation and resolution of inflammation after spinal cord injury. *Neuroscience* 466, 273–297. doi: 10.1016/j.neuroscience.2021.04.026
- Franco-Quijorna, I., Santos-Nogueira, E., Gronert, K., Sullivan, A. B., Kopp, M. A., Brommer, B., et al. (2017). Maresin 1 promotes inflammatory resolution, neuroprotection, and functional neurological recovery after spinal cord injury. *J. Neurosci.* 37, 11731–11743. doi: 10.1523/JNEUROSCI.1395-17.2017
- Han, F., Jiang, X., Li, Z.-M., Zhuang, X., Zhang, X., Ouyang, W.-M., et al. (2020). Epigenetic inactivation of SOX30 is associated with male infertility and offers a therapy target for non-obstructive azoospermia. *Mol. Ther. Nucleic Acids* 19, 72–83. doi: 10.1016/j.omtn.2019.10.038
- Han, Z., Chen, H., Guo, Z., Shen, J., Luo, W., Xie, F., et al. (2022). Circular RNAs and their role in exosomes. *Front. Oncol.* 12:848341. doi: 10.3389/fonc.2022.848341
- Hellenbrand, D. J., Quinn, C. M., Piper, Z. J., Morehouse, C. N., Fixel, J. A., and Hanna, A. S. (2021). Inflammation after spinal cord injury: A review of the critical timeline of signaling cues and cellular infiltration. *J. Neuroinflammation* 18:284. doi: 10.1186/s12974-021-02337-2
- Isaac, R., Reis, F. C. G., Ying, W., and Olefsky, J. M. (2021). Exosomes as mediators of intercellular crosstalk in metabolism. *Cell Metab.* 33, 1744–1762. doi: 10.1016/j.cmet.2021.08.006
- Jafari, N., Lleven, P., and Denis, G. V. (2022). Exosomes as novel biomarkers in metabolic disease and obesity-related cancers. *Nat. Rev. Endocrinol.* 18, 327–328. doi: 10.1038/s41574-022-00666-7
- Kanehisa, M., and Sato, Y. (2020). KEGG mapper for inferring cellular functions from protein sequences. *Protein Sci.* 29, 28–35. doi: 10.1002/pro.3711
- Kimura, T., Horikoshi, Y., Kuriyagawa, C., and Niiyama, Y. (2021). Rho/rock pathway and noncoding RNAs: Implications in ischemic stroke and spinal cord injury. *Int. J. Mol. Sci.* 22:11573. doi: 10.3390/ijms222111573
- Kopper, T. J., and Gensel, J. C. (2021). Continued development of azithromycin as a neuroprotective therapeutic for the treatment of spinal cord injury and other neurological conditions. *Neural Regen. Res.* 16, 508–509. doi: 10.4103/1673-5374.293146
- Li, C., Ni, Y.-Q., Xu, H., Xiang, Q.-Y., Zhao, Y., Zhan, J.-K., et al. (2021). Roles and mechanisms of exosomal non-coding RNAs in human health and diseases. *Signal Transduct. Target. Ther.* 6:383. doi: 10.1038/s41392-021-00779-x
- Li, J.-A., Shi, M.-P., Cong, L., Gu, M.-Y., Chen, Y.-H., Wang, S.-Y., et al. (2023). Circulating exosomal lncRNA-elicited microenvironment changes contribute to rat spinal cord injury development. *Neural Regen. Res.* 18, 889–894. doi: 10.4103/1673-5374.353504
- Li, M., Hua, Q., Shao, Y., Zeng, H., Liu, Y., Diao, Q., et al. (2020). Circular RNA circBbs9 promotes PM2. 5-induced lung inflammation in mice via NLRP3 inflammasome activation. *Environ. Int.* 143:105976. doi: 10.1016/j.envint.2020.105976
- Li, Z., Yu, F., Yu, X., and Wang, S. (2020). Potential molecular mechanism and biomarker investigation for spinal cord injury based on bioinformatics analysis. *J. Mol. Neurosci.* 70, 1345–1353. doi: 10.1007/s12031-020-01549-0
- Lin, H., Yu, J., Gu, X., Ge, S., and Fan, X. (2021). Novel insights into exosomal circular RNAs: Redefining intercellular communication in cancer biology. *Clin. Transl. Med.* 11:e636. doi: 10.1002/ctm2.636
- Liu, C.-X., and Chen, L.-L. (2022). Circular RNAs: Characterization, cellular roles, and applications. *Cell* 185, 2016–2034. doi: 10.1016/j.cell.2022.04.021
- Liu, W., Tao, J.-C., Zhu, S.-Z., Dai, C.-L., Wang, Y.-X., Yu, B., et al. (2022). Expression and regulatory network of long noncoding RNA in rats after spinal cord hemisection injury. *Neural Regen. Res.* 17, 2300–2304. doi: 10.4103/1673-5374.337052
- Liu, X., Zhang, Y., Wang, Y., and Qian, T. (2021). Inflammatory response to spinal cord injury and its treatment. *World Neurosurg.* 155, 19–31. doi: 10.1016/j.wneu.2021.07.148
- Liu, Y., Liu, J., and Liu, B. (2020). Identification of circular RNA expression profiles and their implication in spinal cord injury rats at the immediate phase. *J. Mol. Neurosci.* 70, 1894–1905. doi: 10.1007/s12031-020-01586-9
- Lodde, V., Murgia, G., Simula, E. R., Steri, M., Floris, M., and Idda, M. L. (2020). Long noncoding RNAs and circular RNAs in autoimmune diseases. *Biomolecules* 10:1044. doi: 10.3390/biom10071044
- Misir, S., Wu, N., and Yang, B. B. (2022). Specific expression and functions of circular RNAs. *Cell Death Differ.* 29, 481–491. doi: 10.1038/s41418-022-00948-7
- Peng, P., Zhang, B., Huang, J., Xing, C., Liu, W., Sun, C., et al. (2020). Identification of a circRNA-miRNA-mRNA network to explore the effects of circRNAs on pathogenesis and treatment of spinal cord injury. *Life Sci.* 257:118039. doi: 10.1016/j.lfs.2020.118039
- Qi, J., Wang, T., Zhang, Z., Yin, Z., Liu, Y., Ma, L., et al. (2022). Circctnbn1 regulates neuronal injury in spinal cord injury through the Wnt/ β -catenin signaling pathway. *Dev. Neurosci.* 44, 131–141. doi: 10.1159/000521172
- Qin, C., Liu, C.-B., Yang, D.-G., Gao, F., Zhang, X., Zhang, C., et al. (2019). Circular RNA expression alteration and bioinformatics analysis in rats after traumatic spinal cord injury. *Front. Mol. Neurosci.* 11:497. doi: 10.3389/fnmol.2018.00497
- Saaoud, F., Iv, C. D., Shao, Y., Sun, Y., Lu, Y., Xu, K., et al. (2021). Circular RNAs are a novel type of non-coding RNAs in ROS regulation, cardiovascular metabolic inflammations and cancers. *Pharmacol. Ther.* 220:107715. doi: 10.1016/j.pharmthera.2020.107715
- Sámano, C., Mladinic, M., and Mazzone, G. L. (2021). Circular RNAs: The novel actors in pathophysiology of spinal cord injury. *Front. Integr. Neurosci.* 15:758340. doi: 10.3389/fnint.2021.758340
- Song, H., Yang, Y., Sun, Y., Wei, G., Zheng, H., Chen, Y., et al. (2022). Circular RNA Cdy1 promotes abdominal aortic aneurysm formation by inducing M1 macrophage polarization and M1-type inflammation. *Mol. Ther.* 30, 915–931. doi: 10.1016/j.ymthe.2021.09.017
- Sun, Q., Li, X., Xu, M., Zhang, L., Zuo, H., Xin, Y., et al. (2020). Differential expression and bioinformatics analysis of circRNA in non-small cell lung cancer. *Front. Genet.* 11:586814. doi: 10.3389/fgene.2020.586814
- Tong, D., Zhao, Y., Tang, Y., Ma, J., Wang, Z., and Li, C. (2021). CircUsp10 promotes microglial activation and induces neuronal death by targeting miRNA-152-5p/CD84. *Bioengineered* 12, 10812–10822. doi: 10.1080/21655979.2021.2004362
- Trapnell, C., Williams, B. A., Pertea, G., Mortazavi, A., Kwan, G., Van Baren, M. J., et al. (2010). Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.* 28, 511–515. doi: 10.1038/nbt.1621
- Wang, W.-Z., Li, J., Liu, L., Zhang, Z.-D., Li, M.-X., Li, Q., et al. (2021). Role of circular RNA expression in the pathological progression after spinal cord injury. *Neural Regen. Res.* 16, 2048–2055. doi: 10.4103/1673-5374.308100
- Wen, S.-Y., Qadir, J., and Yang, B. B. (2022). Circular RNA translation: Novel protein isoforms and clinical significance. *Trends Mol. Med.* 28, 405–420. doi: 10.1016/j.molmed.2022.03.003
- Wu, H., Li, Y., Wang, X., Zhang, Z., and Huang, Y. (2021). Long non-coding RNA TUG1 knockdown prevents neurons from death to alleviate acute spinal cord injury via the microRNA-338/BIK axis. *Bioengineered* 12, 5566–5582. doi: 10.1080/21655979.2021.1966258

- Wu, R., Mao, S., Wang, Y., Zhou, S., Liu, Y., Liu, M., et al. (2019). Differential circular RNA expression profiles following spinal cord injury in rats: A temporal and experimental analysis. *Front. Neurosci.* 13:1303. doi: 10.3389/fnins.2019.01303
- Xie, X., Xiao, Y., and Xu, K. (2021). Mechanism underlying circularRNA_014301-mediated regulation of neuronal cell inflammation and apoptosis. *Exp. Ther. Med.* 22:1432. doi: 10.3892/etm.2021.10867
- Yamazaki, K., Kawabori, M., Seki, T., and Houkin, K. (2020). Clinical trials of stem cell treatment for spinal cord injury. *Int. J. Mol. Sci.* 21:3994. doi: 10.3390/ijms21113994
- Ye, X., Chen, Y., Wang, J., Chen, J., Yao, Y., Wang, L.-L., et al. (2021). Identification of circular RNAs related to vascular endothelial proliferation, migration, and angiogenesis after spinal cord injury using microarray analysis in female mice. *Front. Neurol.* 12:666750. doi: 10.3389/fneur.2021.666750
- Zhao, H., Chen, L., Shan, Y., Chen, G., Chu, Y., Dai, H., et al. (2021). Hsa_circ_0038383-mediated competitive endogenous RNA network in recurrent implantation failure. *Aging (Albany NY)* 13, 6076–6090. doi: 10.18632/aging.202590
- Zhao, J., Qi, X., Bai, J., Gao, X., and Cheng, L. (2020). A circRNA derived from linear HIPK3 relieves the neuronal cell apoptosis in spinal cord injury via ceRNA pattern. *Biochem. Biophys. Res. Commun.* 528, 359–367. doi: 10.1016/j.bbrc.2020.02.108
- Zhong, X., Bao, Y., Wu, Q., Xi, X., Zhu, W., Chen, S., et al. (2021). Long noncoding RNA XIST knockdown relieves the injury of microglia cells after spinal cord injury by sponging miR-219-5p. *Open Med.* 16, 1090–1100. doi: 10.1515/med-2021-0292
- Zipser, C. M., Cragg, J. J., Guest, J. D., Fehlings, M. G., Jutzeler, C. R., Anderson, A. J., et al. (2022). Cell-based and stem-cell-based treatments for spinal cord injury: Evidence from clinical trials. *Lancet Neurol.* 21, 659–670. doi: 10.1016/S1474-4422(21)00464-6



OPEN ACCESS

EDITED BY

Jaqueline Bohrer Schuch,
Federal University of Rio Grande do
Sul, Brazil

REVIEWED BY

Yanfang Liu,
Jinan University, China
Jiankun Zang,
Jinan University, China

*CORRESPONDENCE

Jiayun Liu
jiayun@fmmu.edu.cn
Jingzhou Chen
chendragon1976@aliyun.com

SPECIALTY SECTION

This article was submitted to
Neurogenetics,
a section of the journal
Frontiers in Neuroscience

RECEIVED 25 July 2022

ACCEPTED 14 November 2022

PUBLISHED 29 November 2022

CITATION

Bai C, Hao X, Zhou L, Sun Y, Song L,
Wang F, Yang L, Liu J and Chen J
(2022) Machine learning-based
identification of the novel circRNAs
circERBB2 and circCHST12 as
potential biomarkers of intracerebral
hemorrhage.
Front. Neurosci. 16:1002590.
doi: 10.3389/fnins.2022.1002590

COPYRIGHT

© 2022 Bai, Hao, Zhou, Sun, Song,
Wang, Yang, Liu and Chen. This is an
open-access article distributed under
the terms of the [Creative Commons
Attribution License \(CC BY\)](#). The use,
distribution or reproduction in other
forums is permitted, provided the
original author(s) and the copyright
owner(s) are credited and that the
original publication in this journal is
cited, in accordance with accepted
academic practice. No use, distribution
or reproduction is permitted which
does not comply with these terms.

Machine learning-based identification of the novel circRNAs circERBB2 and circCHST12 as potential biomarkers of intracerebral hemorrhage

Congxia Bai¹, Xiaoyan Hao¹, Lei Zhou¹, Yingying Sun²,
Li Song², Fengjuan Wang¹, Liu Yang¹, Jiayun Liu^{1*} and
Jingzhou Chen^{2,3*}

¹Department of Clinical Laboratory Medicine, Xijing Hospital, Fourth Military Medical University, Xi'an, China, ²State Key Laboratory of Cardiovascular Disease, Fuwai Hospital, National Center for Cardiovascular Diseases, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China, ³National Health Commission Key Laboratory of Cardiovascular Regenerative Medicine, Fuwai Central-China Hospital, Central-China Branch of National Center for Cardiovascular Diseases, Zhengzhou, China

Background: The roles and potential diagnostic value of circRNAs in intracerebral hemorrhage (ICH) remain elusive.

Methods: This study aims to investigate the expression profiles of circRNAs by RNA sequencing and RT-PCR in a discovery cohort and an independent validation cohort. Bioinformatics analysis was performed to identify the potential functions of circRNA host genes. Machine learning classification models were used to assess circRNAs as potential biomarkers of ICH.

Results: A total of 125 and 284 differentially expressed circRNAs (fold change > 1.5 and FDR < 0.05) were found between ICH patients and healthy controls in the discovery and validation cohorts, respectively. Nine circRNAs were consistently altered in ICH patients compared to healthy controls. The combination of the novel circERBB2 and circCHST12 in ICH patients and healthy controls showed an area under the curve of 0.917 (95% CI: 0.869–0.965), with a sensitivity of 87.5% and a specificity of 82%. In combination with ICH risk factors, circRNAs improved the performance in discriminating ICH patients from healthy controls. Together with hsa_circ_0005505, two novel circRNAs for differentiating between patients with ICH and healthy controls showed an AUC of 0.946 (95% CI: 0.910–0.982), with a sensitivity of 89.1% and a specificity of 86%.

Conclusion: We provided a transcriptome-wide overview of aberrantly expressed circRNAs in ICH patients and identified hsa_circ_0005505 and novel circERBB2 and circCHST12 as potential biomarkers for diagnosing ICH.

KEYWORDS

intracerebral hemorrhage, RNA sequencing, circular RNA, biomarkers, machine learning algorithms

Introduction

Stroke causes high levels of mortality and disability globally. Intracerebral hemorrhage (ICH) is a deadly stroke subtype with an estimated annual incidence of 16 per 100,000 persons worldwide (Wilkinson et al., 2018). ICH accounts for approximately 23.8% of stroke cases in China, compared with Western countries, where it accounts for 10–15% of stroke cases, causing a median fatality ratio of 40.4% per month (Qureshi et al., 2009; Benjamin et al., 2017). The diagnosis of stroke is often made with computed tomography (CT) or magnetic resonance imaging (MRI), and although most patients are hospitalized with typical neurological symptoms, it is difficult to distinguish ICH from ischemic stroke (IS) in the super acute period (Hankey, 2017). Thus, identifying potential biomarkers for the early prediction and diagnosis of ICH is important.

Non-coding RNAs (ncRNAs) have been extensively studied in the pathophysiology of cerebrovascular diseases (Weng et al., 2022). Changes in RNA levels during stroke have the potential to aid stroke diagnosis and provide insight into stroke diagnosis and management (Montaner et al., 2020). Emerging evidence has revealed that ncRNA expression profiles are altered in the peripheral blood of patients with ICH (Kim et al., 2019; Li et al., 2019; Cheng et al., 2020). CircRNAs are a novel class of ncRNAs that are produced in eukaryotic cells during posttranscriptional processes; these covalently closed RNAs lack a free 3' or 5' end and are resistant to exonuclease digestion (Kristensen et al., 2019). Thus, circRNAs are promising diagnostic and prognostic biomarkers for many human diseases because of their stability, specificity and abundance in human blood (Jeck and Sharpless, 2014; Zhang et al., 2018). Growing evidence has demonstrated that circRNAs are implicated in a variety of pathological conditions, including coronary artery disease (Cardona-Monzonis et al., 2020), acute ischemic stroke (Liu Y. et al., 2022) and cancers (Kristensen et al., 2022). Moreover, the expression of circRNAs was found to be significantly altered in IS (Tiedt et al., 2017; Dong et al., 2020; Li et al., 2020; Lu et al., 2020; Ostolaza et al., 2020; Zuo et al., 2020), and these studies implied that aberrantly expressed circRNAs may be novel biomarkers for IS diagnosis and prognosis. Our previous study revealed that circRNA profiles were significantly altered in hypertensive ICH patients compared to hypertensive subjects without

ICH and found that hsa_circ_0001240, hsa_circ_0001947 and hsa_circ_0001386 were potential biomarkers for predicting and diagnosing hypertensive ICH (Bai et al., 2021). In addition, circRNA expression is significantly altered in rat brain tissue after ICH (Dou et al., 2020; Zhong et al., 2020), indicating that circRNAs are novel clinical biomarkers for ICH. However, comprehensive circRNA expression profiles and their potential diagnostic value in the peripheral blood of ICH patients remain elusive.

Artificial intelligence techniques such as machine learning tools have been increasingly used in precision diagnosis (Chang et al., 2021). Machine learning algorithms are artificial intelligence techniques used to select the best model from a set of alternatives to fit a set of observations (Li, 2018). Machine learning has remained a fundamental and indispensable tool due to its efficacy and efficiency in both feature extraction of relevant biomarkers and the classification of samples as validation of the discovered biomarkers (Ledesma et al., 2021).

In this study, we investigated the expression profile of circRNAs in peripheral blood cells from patients with ICH, patients with IS and healthy controls by RNA sequencing in the discovery and validation cohorts. The significantly altered circRNA host genes were examined with Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analyses to characterize the potential functions. We further validated the altered circRNAs by quantitative reverse transcription-PCR (RT-PCR) analysis of all samples. Logistic regression models were performed to identify whether circRNAs were independent factors for ICH. Additionally, we performed Spearman's correlation analysis to investigate the correlation between ICH risk factors and candidate circRNAs. Furthermore, machine learning classification algorithms and ROC curves were used to assess circRNAs as potential biomarkers of ICH.

Materials and methods

Study design and sample collection

We recruited 64 patients with ICH, 59 patients with IS and 50 sex- and age-matched healthy controls between 2014 and 2019 from two individual cohorts for RNA sequencing. In the discovery cohort, 44 patients with ICH, 43 patients

with IS and 31 healthy controls were enrolled from Cangzhou Central Hospital between 2014 and 2017. In the validation cohort, 20 patients with ICH were enrolled from the Affiliated Hospital of Hebei University, 16 patients with IS were enrolled from General Hospital of Ningxia Medical University, and 19 healthy control subjects were enrolled from the Tsinghua University Hospital between 2017 and 2019. Patients with ICH were diagnosed by professional neurologists based on their histories and examinations, and ICH was confirmed by CT or MRI. Healthy controls without a history of stroke or cardiovascular events were selected. The demographic and clinical characteristics of the study population were obtained through a face-to-face survey and by checking hospital records or medical examination records. The exclusion criteria included autoimmune diseases, cardiac disease, liver diseases, renal diseases, cancer or a history of stroke and cerebral infarction with hemorrhagic transformation. This study was reviewed and approved by the Human Ethics Committee, Fuwai Hospital (Approval No. 2016-732), and conducted in accordance with the principles of Good Clinical Practice and the Declaration of Helsinki. Written informed consent was obtained from all participants or their legal proxies.

RNA isolation and sequencing

RNA was isolated from human peripheral blood and used to perform RNA sequencing by Annoroad

Gene Technology Company Ltd. (Beijing, China), as previously described (Bai et al., 2021). Total RNA from all samples was isolated with an RNeasy Mini kit (QIAGEN). An Agilent 2100 RNA Nano 6000 Assay Kit (Agilent Technologies, CA, USA) was used to measure RNA integrity. The libraries were constructed using an RNA integrity number ≥ 7.5 and a 28S:18S rRNA ratio ≥ 1.8 . Ribo-Zero™ Gold Kits (Illumina, San Diego, CA, USA) were utilized to eliminate all ribosomal RNAs from total RNA. RNase R (Epicenter, Madison, WI, USA) digestion was used to eliminate linear RNAs. The purified circRNAs were subjected to the NEB Next Ultra Directional RNA Library Prep Kit for Illumina (NEB, Ipswich, USA) according to the manufacturer's instructions. The obtained libraries were subjected to paired-end sequencing with 150 bp reads performed on the Illumina PE150 platform. The sequence depth was approximately 15G. The raw sequencing data were analyzed using Q30 statistics from FastQC, and clean reads were obtained by removing adaptor-polluted and low-quality reads. The RNA-seq data have been deposited into the Genome Sequence Archive (Chen T. et al., 2021) in the National Genomics Data Center (CNCB-NGDC Members and Partners, 2022), China National Center for Bioinformatics/Beijing Institute of Genomics, Chinese Academy of Sciences (GSA-Human: HRA001807), which are publicly accessible at <https://ngdc.cncb.ac.cn/gsa-human>.

TABLE 1 Demographics and characteristics of the discovery and validation cohorts.

| | Discovery cohort | | | | Validation cohort | | | |
|------------------------|---------------------|-----------------|----------------|---------|---------------------|-----------------|----------------|---------|
| | Control (n = 31) | ICH (n = 44) | IS (n = 43) | P-value | Control (n = 19) | ICH (n = 20) | IS (n = 16) | P-value |
| Age, y | 58.9 ± 5.3 | 55.9 ± 7.2 | 57.4 ± 5.5 | 0.09 | 57.2 ± 7.0 | 56.7 ± 7.1 | 57.2 ± 7.7 | 0.86 |
| Men, % | 17 (54.8) | 24 (54.5) | 21 (48.8) | 0.83 | 10 (52.6) | 10 (50) | 8 (50) | 0.98 |
| BMI, kg/m ² | 24.8 ± 2.9 | 26.1 ± 6.6 | 27.6 ± 6.9 | 0.09 | 24.9 ± 2.4 | 25.8 ± 6.8 | 25.0 ± 2.6 | 0.90 |
| SBP, mmHg | 125.7 ± 10.1 | 137.4 ± 17.6 | 138.6 ± 13.6 | < 0.001 | 120.3 ± 9.7 | 171.2 ± 25.7 | 150.6 ± 19.4 | < 0.001 |
| DBP, mmHg | 79.2 ± 4.3 | 87.9 ± 10.7 | 91.8 ± 16.6 | < 0.001 | 77.6 ± 9.1 | 103.7 ± 13.3 | 89.3 ± 13.9 | < 0.001 |
| HDL-C, mmol/L | 1.4 ± 0.3 | 1.1 ± 0.3 | 1.1 ± 0.2 | < 0.001 | 1.3 ± 0.3 | 0.9 ± 0.5 | 1.0 ± 0.3 | 0.007 |
| LDL-C, mmol/L | 2.9 ± 0.7 | 2.4 ± 0.8 | 2.3 ± 0.8 | < 0.001 | 2.9 ± 0.9 | 2.8 ± 0.8 | 2.7 ± 0.9 | 0.82 |
| TC, mmol/L | 5.5 ± 1.0 | 4.5 ± 1.0 | 4.5 ± 1.0 | < 0.001 | 4.5 ± 1.0 | 4.3 ± 0.9 | 4.9 ± 1.3 | 0.23 |
| TG, mmol/L | 1.4 ± 0.8 | 1.5 ± 0.9 | 1.6 ± 0.6 | 0.41 | 1.2 ± 0.5 | 1.4 ± 0.6 | 2.3 ± 1.5 | 0.004 |
| GLU, mmol/L | 6.0 ± 1.8 | 6.3 ± 1.6 | 5.9 ± 1.3 | 0.17 | 5.3 ± 0.6 | 5.5 ± 1.7 | 6.0 ± 1.1 | 0.21 |
| Smoking, % | | | | 0.95 | | | | 0.92 |
| Never | 19 (61.3) | 28 (63.7) | 26 (60.5) | | 13 (68.4) | 14 (70) | 11 (68.7) | |
| Former | 4 (12.9) | 5 (13.6) | 8 (18.6) | | 3 (15.8) | 2 (10) | 2 (12.5) | |
| Current | 8 (25.8) | 11 (22.7) | 9 (20.9) | | 3 (15.8) | 4 (20) | 3 (18.8) | |
| Drinking, % | | | | 0.98 | | | | 0.96 |
| Non-drinker | 20 (64.5) | 28 (63.6) | 27 (62.8) | | 11 (57.9) | 12 (60) | 10 (62.5) | |
| Drinker | 11 (35.5) | 16 (36.4) | 16 (37.2) | | 8 (42.1) | 8 (40) | 6 (37.5) | |

Data are expressed as the mean ± standard deviation or n (%).

BMI, Body mass index; SBP, Systolic blood pressure; DBP, Diastolic blood pressure; TC, Total cholesterol; TG, Triacylglycerol; HDL-C, High-density lipoprotein cholesterol; LDL-C, Low-density lipoprotein cholesterol; GLU, Glucose; ICH, Intracerebral hemorrhage; IS, ischemic stroke.

Statistical comparisons for percentages were performed using the chi-square test. Comparisons between means or medians were performed using one-way ANOVA.

Differential expression analysis

The differential expression circRNA analysis was performed as previously described (Bai et al., 2021). Briefly, CIRI2 (Gao et al., 2018) was used to detect paired chiasmic clipping signals according to the mapping of reads. The reads were mapped to the reference genome¹ using the BWA-MEM method. Back-spliced junction reads were integrated and measured by spliced reads per billion mapping to quantify circRNA. Differential expression analysis was performed using the DESeq2 R package (Wang et al., 2010) and edgeR (Robinson et al., 2010). Fold differences of each circRNA were calculated to identify differentially expressed circRNAs between ICH patients and healthy controls (or IS patients) by Student's *t*-test. A *P* value was assigned to each circRNA and adjusted by multiple testing using the Benjamini–Hochberg method for controlling the false discovery rate (FDR). The differentially expressed circRNAs were defined as those with a fold change ≥ 1.5 and $\text{FDR} < 0.05$.

Bioinformatics analysis

Volcano plots and hierarchical clustering using heatmaps were generated based on the normalized values of differentially expressed genes using the R package. Venn diagrams were used to present the consistently differentially expressed genes in the discovery and validation cohorts. GO enrichment and KEGG analyses were performed to determine the biological functions and pathways of differentially expressed circRNA host genes. *P* values were calculated using Fisher's exact test with the hypergeometric algorithm.

Quantitative real-time polymerase chain reaction validation

To validate the expression levels of differentially expressed circRNAs identified by RNA-seq, the candidate circRNAs were selected for further validation of expression levels by quantitative RT-PCR. Total RNA was incubated with RNase R or RNase-free water as a control at 37°C for 30 min to purify the circRNAs. After incubation, cDNA synthesis was completed using 1 μg of total RNA and a Transcriptor First Stand cDNA Synthesis Kit (Takara, Dalian, China), and Taq premix (Takara, Dalian, China) was added to start PCR according to the manufacturer's protocol. The products were used for Sanger sequencing. Quantitative RT-PCR was performed using SYBR Master Mix (Yeasen, Shanghai, China) on the ViiA 7 Real-time PCR System (Applied Biosystems) according to

the manufacturer's instructions. The circRNA primers were designed to overlap the back-spliced junction using the NCBI Primer-BLAST website.² The primers used in this study are listed in **Supplementary Table 7**. The relative expression of the corresponding genes was quantified and normalized to that of GAPDH.

Performance evaluation of candidate biomarkers with classification algorithms

To evaluate the applicable biomarkers for ICH, we used mutual information (MI) (Blok and Stambler, 2017) and random forest (RF) algorithms (Ambale-Venkatesh et al., 2017; Kawakami et al., 2019) to screen circRNA biomarker signatures according to the expression levels in all samples. To assess the diagnostic values of the specific circRNAs, we used six machine learning classification algorithms (Chang et al., 2021; Chen Y. et al., 2021; Liu D. et al., 2022), support vector machine (SVM), RF, K-nearest neighbor (KNN), logistic regression (LR), decision tree (DT) and Gaussian naive Bayes (GNB), to discriminate ICH patients from healthy controls or IS patients according to the expression levels of circRNAs by Python packages. To ensure the stability and accuracy of the classifiers, we used 10-fold cross-validation; 90% of the data were used for the training set, and 10% were used for the test set. We calculated five measurements, including sensitivity, specificity, accuracy, positive predictive value (PPV), and negative predictive value (NPV) (Shu et al., 2020). The ROC curve was illustrated based on sensitivity and 1-specificity scores. For each area under the curve (AUC) value, the 95% CI was computed with 1000 stratified bootstrap replicates.

Statistical analysis

Statistical analysis was performed using SPSS 21.0 (IBM Corp., NY, USA). The sample distribution was determined using the Kolmogorov–Smirnov normality test. For parametric data, the two-tailed unpaired Student's *t*-test was used to determine differences between two groups. The data are represented as the means \pm standard deviations or medians (interquartile range). Statistical comparisons for percentages were performed using chi-square statistical analysis. In the RNA sequencing analysis, differentially expressed RNAs were selected if there were significant differences (fold change > 1.5 and $\text{FDR} < 0.05$) between the ICH patients and healthy controls (or IS patients) using Student's *t*-test. Logistic regression models were used to evaluate

¹ <http://www.ensembl.org/index.html>

² <https://www.ncbi.nlm.nih.gov/tools/primer-blast/>

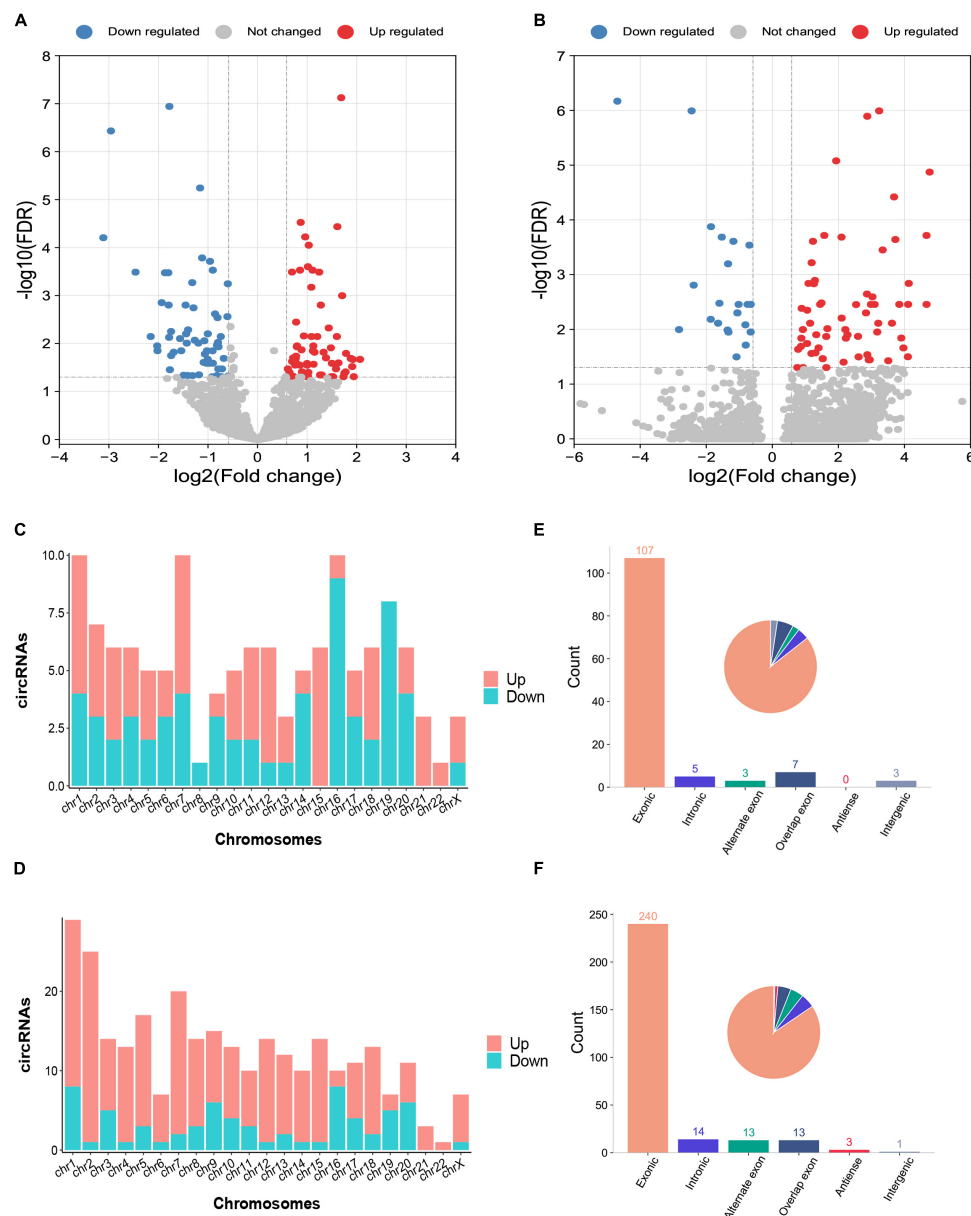


FIGURE 1

Differentially expressed circRNAs between intracerebral hemorrhage (ICH) patients and healthy controls in the discovery and validation cohorts. (A,B) The volcano plot of circRNA expression profiles in ICH patients and controls (fold change ≥ 1.5 and FDR < 0.05) in the discovery ($n = 44$ vs. 31) (A) and validation ($n = 20$ vs. 19) (B) cohorts. Red dots represent upregulated genes, and blue dots represent downregulated genes. (C) The bar diagram shows the circRNA distribution in the chromosomes between 44 ICH patients and 31 healthy controls in the discovery cohort. The red columns represent upregulated circRNAs, while blue columns represent downregulated circRNAs. (D) The bar diagram shows the circRNA distribution in the chromosomes between 20 ICH patients and 19 healthy controls in the validation cohort. The red columns represent upregulated circRNAs, while blue columns represent downregulated circRNAs. (E) The bar diagram and pie chart show the differentially expressed circRNA distribution in the chromosome region (exonic, intronic, intergenic, alternate exon, overlapping exon and antisense) in 44 ICH patients compared with 31 healthy controls in the discovery cohort. (F) The bar diagram and pie chart show the differentially expressed circRNA distribution in the chromosome region (exonic, intronic, intergenic, alternate exon, overlapping exon and antisense) in 20 ICH patients compared with 19 healthy controls in the validation cohort.

whether circRNAs were independent predictive factors for ICH. Spearman's correlation analysis was performed to investigate the correlation between ICH risk factors and circRNAs. The net reclassification index (NRI) and integrated

discrimination improvement (IDI) were calculated to evaluate the effect of the candidate biomarkers as previously described (Wu et al., 2020). $P < 0.05$ was considered indicative of statistical significance.

Results

CircRNA expression profiles were significantly altered in intracerebral hemorrhage patients

The characteristics and demographics of the cohorts of ICH patients, IS patients and healthy controls are shown in [Table 1](#). In RNA sequencing, the significantly differentially expressed circRNAs were determined by a fold change > 1.5 and FDR < 0.05 by DESeq2 methods. In total, 125 circRNAs were significantly altered between

patients with ICH and controls, including 63 upregulated circRNAs and 62 downregulated circRNAs in the discovery cohort ([Figure 1A](#) and [Supplementary Table 1](#)), and 284 circRNAs were significantly altered between patients with ICH and healthy controls in the validation cohort, including 218 upregulated circRNAs and 66 downregulated circRNAs ([Figure 1B](#) and [Supplementary Table 2](#)). Additionally, the circRNAs were distributed across all chromosomes in both cohorts ([Figures 1C,D](#)). There were 107 circRNAs produced by classic exon back-splicing, 3 alternate exons, 5 introns, 7 overlapping exons, and 3 intergenic circRNAs detected between ICH patients and controls in the discovery cohort

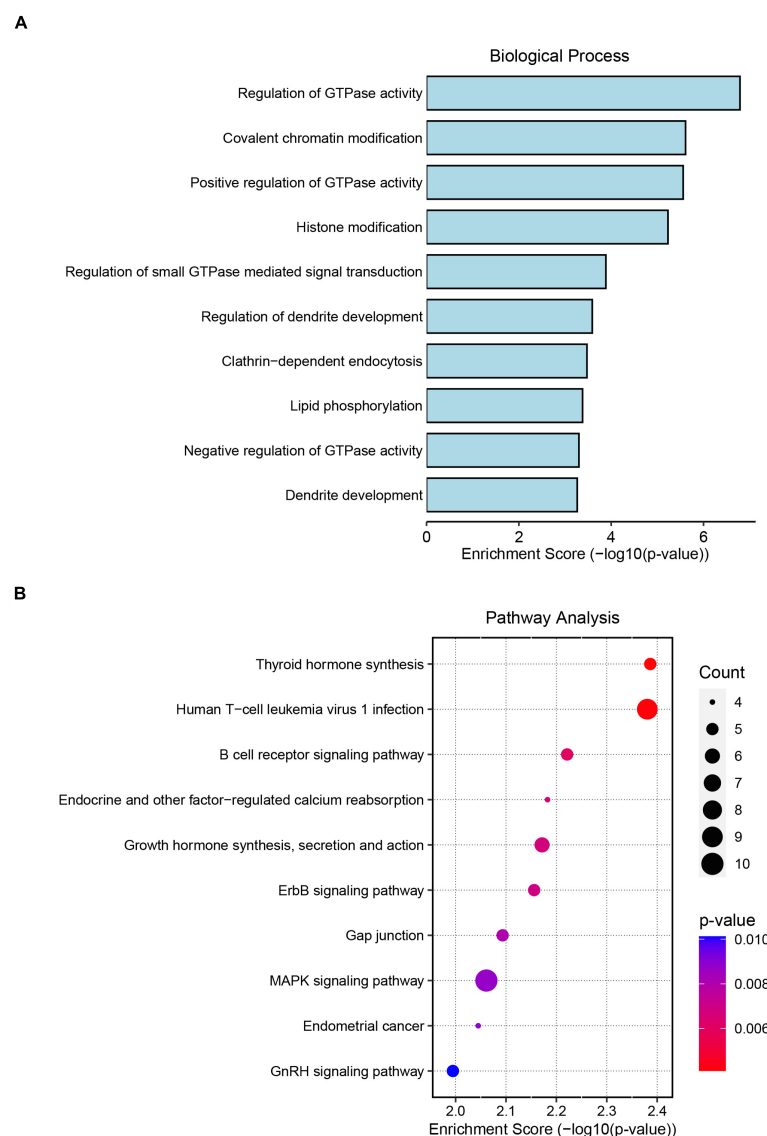


FIGURE 2

Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analyses of significantly altered circRNA host genes. (A) The top 10 biological process terms from GO enrichment analysis of differentially expressed circRNA host genes. (B) The top 10 KEGG pathway analyses of differentially expressed circRNA host genes.

(Figure 1E), and 240 circRNAs produced by classic exon back-splicing, 13 alternate exons, 14 introns, 13 overlapping exons, 3 antisense and 1 intergenic circRNA were detected

between ICH patients and controls in the validation cohort (Figure 1F). Moreover, we observed that 302 and 395 circRNAs were significantly altered between ICH and IS patients in the

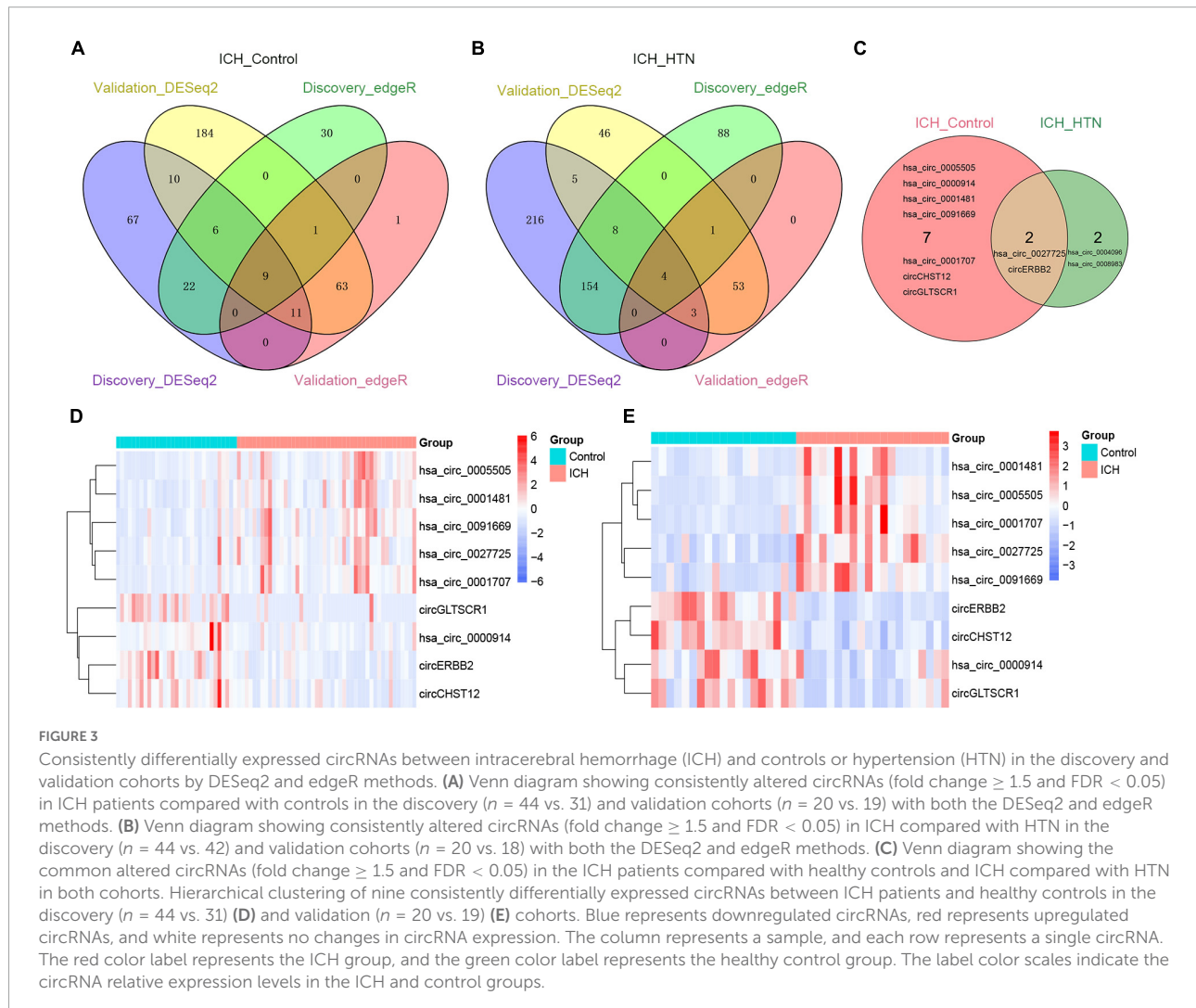


TABLE 2 The consistently altered circRNAs in intracerebral hemorrhage (ICH) patients compared with controls.

| Location | circRNA ID | Discovery | | Validation | | Up/Down | Host gene | Type |
|-----------------------------|------------------|-----------|---------|------------|---------|---------|-----------|----------------|
| | | FC | FDR | FC | FDR | | | |
| chr7:48541721-48542148: + | hsa_circ_0001707 | 3.038 | 2.4E-04 | 3.817 | 2.1E-09 | up | Null | Intronic |
| chrX:147733519-147744289: + | hsa_circ_0091669 | 1.827 | 0.005 | 2.138 | 1.4E-06 | up | AFF2 | Exonic |
| chr12:66597490-66622150: + | hsa_circ_0005505 | 2.903 | 0.007 | 7.329 | 2.8E-10 | up | IRAK3 | Exonic |
| chr5:49694940-49707217: - | hsa_circ_0001481 | 1.850 | 0.012 | 2.803 | 3.9E-06 | up | EMB | Exonic |
| chr12:94562928-94580249: + | hsa_circ_0027725 | 2.179 | 0.037 | 2.281 | 4.9E-07 | up | PLXNC1 | Exonic |
| chr7:2477438-2483381: + | circCHST12 | 0.334 | 0.007 | 0.395 | 5.5E-07 | down | CHST12 | Alternate exon |
| chr19:18648410-18649246: - | hsa_circ_0000914 | 0.510 | 0.010 | 0.595 | 4.6E-06 | down | FKBP8 | Exonic |
| chr17:37866065-37872192: + | circERBB2 | 0.278 | 0.020 | 0.184 | 1.7E-10 | down | ERBB2 | Exonic |
| chr19:48185232-48198731: + | circGLTSCR1 | 0.183 | 0.037 | 0.275 | 1.3E-05 | down | GLTSCR1 | Exonic |

FC, fold change; FDR, false discovery rate.

discovery and validation cohorts, respectively (Supplementary Figures 1A,B).

Gene ontology enrichment and kyoto encyclopedia of genes and genomes pathway analyses of circRNA host genes

To assess the potential regulatory mechanism of differentially expressed circRNAs in host gene transcription after ICH, we performed GO and KEGG pathway analyses of the host genes of the altered circRNAs in the two cohorts. The top GO terms in the biological process category indicated that the host genes were involved in the regulation of GTPase activity, covalent chromatin modification, histone modification, regulation of dendrite development and lipid phosphorylation (Figure 2A). KEGG pathway analysis showed that the host genes were mainly involved in the MAPK signaling network, B-cell receptor signaling, ERBB receptor signaling network, thyroid hormone synthesis and lysine degradation (Figure 2B).

Consistently altered circRNAs in the discovery and validation cohorts

To elucidate the underlying mechanism by which the circRNAs affected ICH more specifically, we screened the common circRNAs in the two cohorts by both DESeq2 and edgeR methods (Supplementary Tables 1–4) and found that 9 circRNAs overlapped between the ICH patients and controls (Figure 3A). Similarly, there were 4 consistent circRNAs between ICH and hypertension (HTN) in our previous study (Figure 3B) (Bai et al., 2021); 2 of them were consistently altered in the two comparison groups, including hsa_circ_0027725 and a novel circRNA (host gene *ERBB2*) we named circERBB2 (Figure 3C).

The nine consistently altered circRNAs included five upregulated circRNAs and four downregulated circRNAs. The five upregulated circRNAs in ICH were hsa_circ_0001707, hsa_circ_0091669, hsa_circ_0005505, hsa_circ_0001481 and hsa_circ_0027725; the 4 downregulated circRNAs in ICH were hsa_circ_0000914 and three novel circRNAs that we named according to their host genes, circCHST12 (host gene *CHST12*), circERBB2 and circGLTSCR1 (host gene *GLTSCR1*) (Table 2). The 9 circRNA expression variants are shown with hierarchical clustering heatmaps in the discovery and validation cohorts (Figures 3D,E), which indicated that the circRNA expression profiles in ICH patients were distinct from those in healthy control groups.

Likewise, we detected 20 consistent circRNAs between ICH and IS patients in the two cohorts by both DESeq2 and edgeR methods (Supplementary

Figure 1C). Notably, 3 circRNAs were in the intersection between ICH versus controls (9 consistent circRNAs) and ICH versus IS (20 consistent circRNAs), including circERBB2, circCHST12 and hsa_circ_0005505 (Supplementary Figure 1D).

Investigation of the nine circRNAs as independent predictors of intracerebral hemorrhage

To further explore the potential value of candidate circRNAs as ICH biomarkers, logistic regression models were performed to identify whether nine circRNAs could be predictors of ICH occurrence. As shown in Table 3, after adjusting for age, sex, body mass index (BMI), systolic blood pressure (SBP), diastolic blood pressure (DBP), total cholesterol (TC), triacylglycerol (TG), high-density lipoprotein cholesterol (HDL-C), low-density lipoprotein cholesterol (LDL-C), smoking and alcohol consumption, per unit of increase in hsa_circ_0001707, hsa_circ_0091669, hsa_circ_0005505, hsa_circ_0001481 and hsa_circ_0027725, the odds ratios for ICH occurrence were 2.23 (95% CI: 1.294–3.842; $P = 0.004$), 3.372 (95% CI: 1.665–6.867; $P = 0.001$), 2.216 (95% CI: 1.363–3.316; $P = 0.001$), 4.750 (95% CI: 2.054–10.985; $P < 0.001$) and 2.156 (95% CI: 1.170–3.974; $P = 0.014$), respectively. In addition, the adjusted ORs were 0.009 (95% CI: 0.001–0.097; $P < 0.001$), 0.160 (95% CI: 0.051–0.507; $P = 0.002$), 0.019 (95% CI: 0.002–0.157; $P < 0.001$) and 0.122 (95% CI: 0.037–0.410; $P = 0.001$) per unit increase in circCHST12, hsa_circ_0000914, circERBB2 and circGLTSCR1, respectively.

TABLE 3 Logistic regression analysis to identify circRNAs as independent predictive factors of intracerebral hemorrhage (ICH).

| circRNA ID | Adjusted risk factors | | | Up/Down Host gene | |
|------------------|-----------------------|--------------|---------|-------------------|---------|
| | OR | 95% CI | P-value | | |
| hsa_circ_0001707 | 2.230 | 1.294–3.842 | 0.004 | up | Null |
| hsa_circ_0091669 | 3.372 | 1.655–6.867 | 0.001 | up | AFF2 |
| hsa_circ_0005505 | 2.216 | 1.363–3.316 | 0.001 | up | IRAK3 |
| hsa_circ_0001481 | 4.750 | 2.054–10.985 | < 0.001 | up | EMB |
| hsa_circ_0027725 | 2.156 | 1.170–3.974 | 0.014 | up | PLXNC1 |
| circCHST12 | 0.009 | 0.001–0.097 | < 0.001 | down | CHST12 |
| hsa_circ_0000914 | 0.160 | 0.051–0.507 | 0.002 | down | FKBP8 |
| circERBB2 | 0.019 | 0.002–0.157 | < 0.001 | down | ERBB2 |
| circGLTSCR1 | 0.122 | 0.037–0.410 | 0.001 | down | GLTSCR1 |

Risk factors included SBP, systolic blood pressure; DBP, diastolic blood pressure; TG, triacylglycerol; TC, total cholesterol; LDL-C, low-density lipoprotein cholesterol; HDL-C, high-density lipoprotein cholesterol, smoking and alcohol consumption; ICH, intracerebral hemorrhage; OR, odds ratio; CI, confidence interval.

Validation of the differentially expressed circRNAs by quantitative real-time polymerase chain reaction

To verify the novel circRNAs circERBB2 and circCHST12 are really circular form, we first blasted the sequences and confirmed the back-splice junction sites and assayed them by RT-PCR with divergent primers. Next, Sanger sequencing was performed to illustrate the junction site. The results showed that circERBB2, located at chr17:37866065-37872192 (genomic length: 6127 bp, spliced sequence length: 939 bp), was derived from exons 9–16 of the *ERBB2* gene (Figure 4A). circCHST12, located at chr7:2477438-2483381 (genomic length: 5943 bp, spliced sequence length: 5943 bp), was derived from exon 1 and partial exon 2 of the *CHST12* gene (Figure 4B). RT-qPCR analysis of total RNA after RNase R or control treatment indicated that circERBB2 and circCHST12 were resistant, while

ERBB2, *CHST12* and *GAPDH* mRNA transcripts were degraded (Figures 4C,D). These data established that circERBB2 and circCHST12 are two *bona fide* circRNAs.

Next, to confirm the expression of circRNAs in the high-throughput results, we selected three upregulated circRNAs (hsa_circ_0001707, hsa_circ_0005505 and hsa_circ_0027725) and three downregulated circRNAs (hsa_circ_0000914, circERBB2 and circCHST12) of the above consistently altered circRNAs for further validation by RT-qPCR in all samples. The expression levels of these circRNAs were consistent with the RNA sequencing results, including three upregulated circRNAs and three downregulated circRNAs that were significantly altered in patients with ICH compared with control subjects (Figures 5A–F). Moreover, the expression levels of circERBB2, circCHST12 and hsa_circ_0005505 were also significantly altered between ICH and IS patients (Figures 5G–I). These results were consistent with the levels obtained by RNA sequencing, supporting the accuracy and reliability of the data.

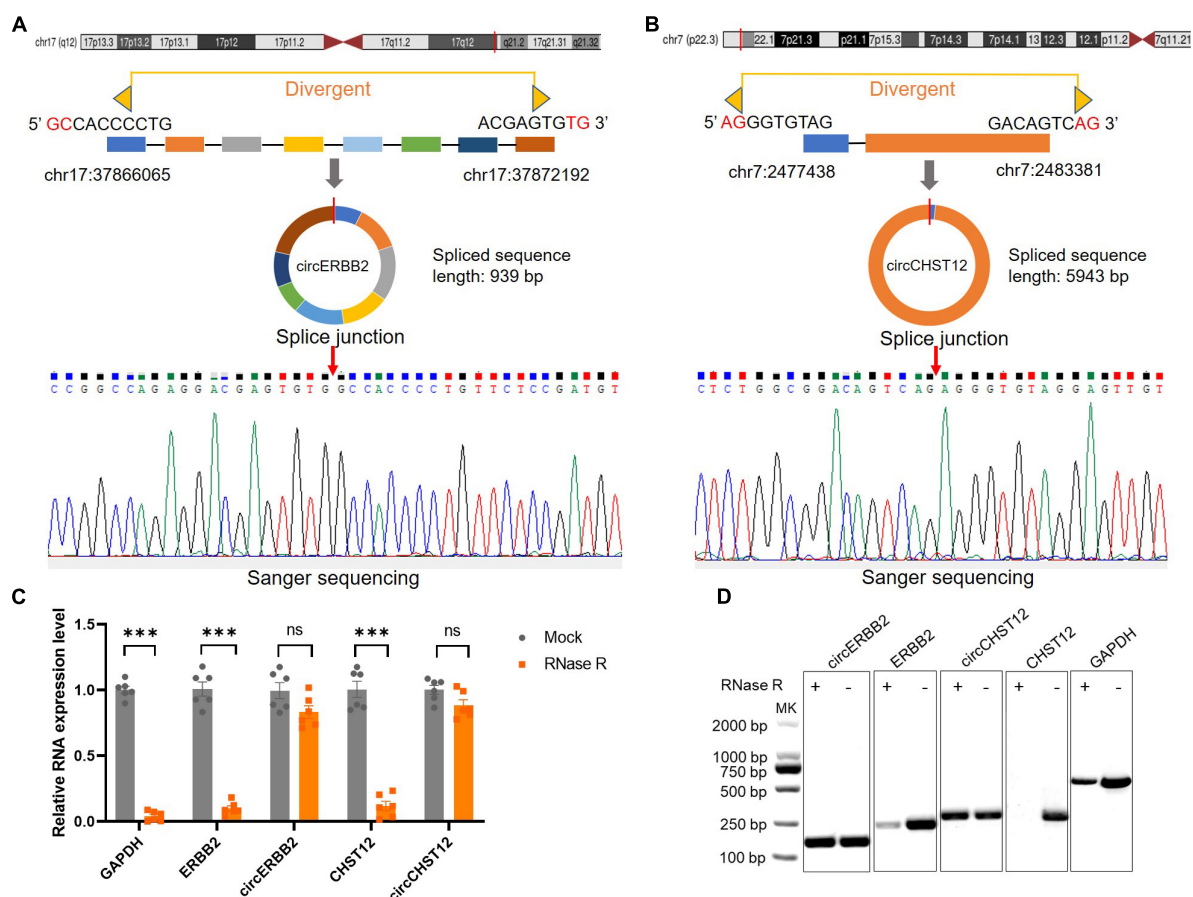


FIGURE 4

Identification of novel circular RNAs circERBB2 and circCHST12. (A,B) Schematic diagrams and Sanger sequencing illustrated the back-splice junction site of circERBB2 (A) and circCHST12 (B). (C) RT-qPCR showed the expression of *GAPDH*, *ERBB2*, *circERBB2*, *CHST12* and *circCHST12* administered RNase R or mock control ($n = 6$ per group). (D) Representative agarose gel pictures showing the relative expression of *GAPDH*, *ERBB2*, *circERBB2*, *CHST12*, and *circCHST12* administered RNase R or mock control. Data are presented as the mean \pm standard deviation. *** $p < 0.001$. ns: no significant. Statistical significance was assessed using unpaired two-tailed Student's t -test.

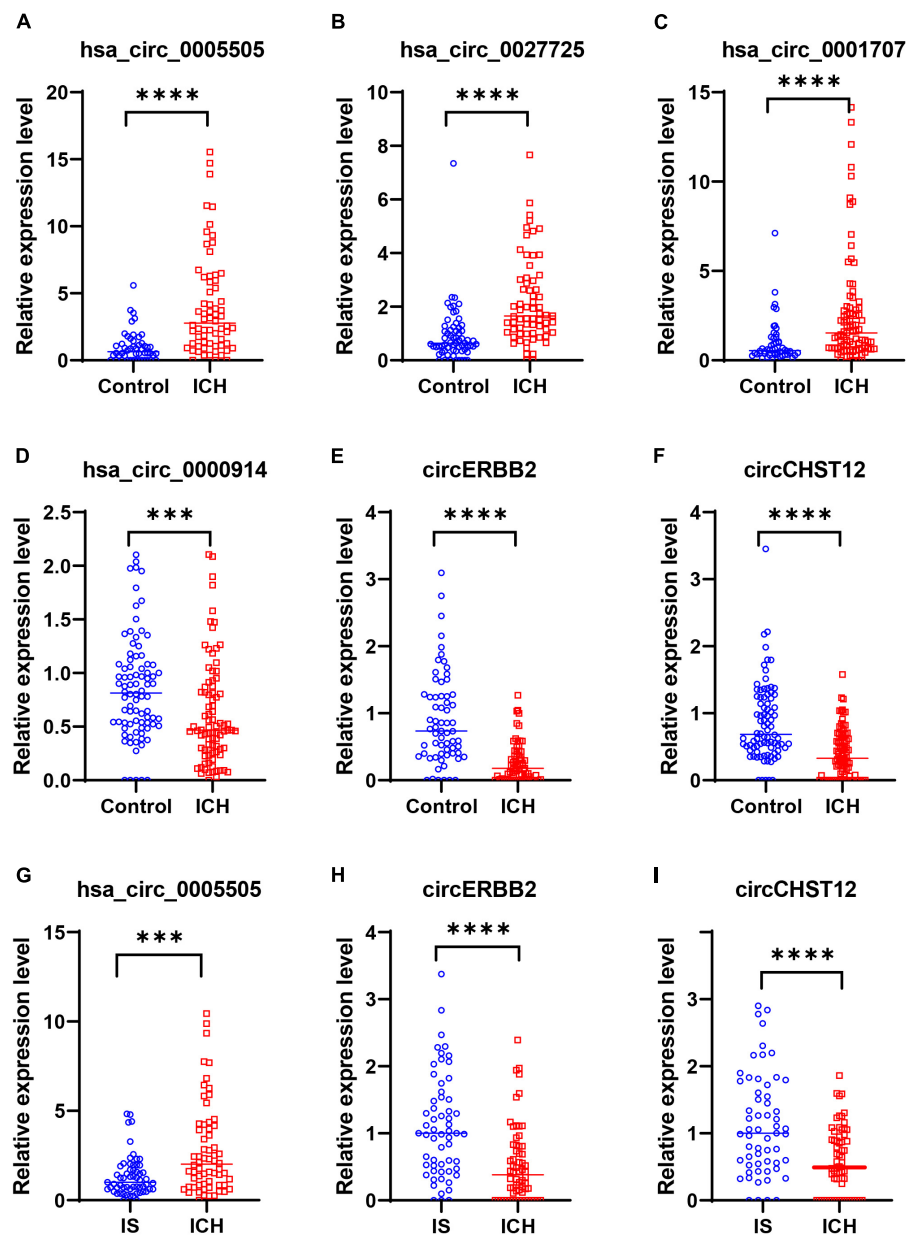


FIGURE 5

Validation of circRNA expression levels by quantitative real-time polymerase chain reaction (RT-qPCR). (A–F) RT-qPCR results validated the expression levels of candidate circRNAs in all samples between 64 intracerebral hemorrhage (ICH) patients and 50 healthy controls. (A) hsa_circ_0005505, (B) hsa_circ_0027725, (C) hsa_circ_0001707, (D) hsa_circ_0000914, (E) circERBB2 and (F) circCHST12. (G–I) RT-qPCR results validated the expression levels of hsa_circ_0005505 (G), circERBB2 (H) and circCHST12 (I) between 64 ICH patients and 59 ischemic stroke (IS) patients. The data are presented as the median (interquartile range). *** $p < 0.001$, **** $p < 0.0001$. Statistical significance was assessed using the Mann–Whitney U test.

Performance evaluation of the candidate circRNAs with classification algorithms

To evaluate applicable biomarkers for ICH, we used mutual information (MI) and random forest (RF) algorithms to screen circRNA marker signatures according to the expression levels in

all samples. We obtained the signature of the top 10 circRNAs in the two algorithms and found 4 circRNAs [hsa_circ_0005806, circERBB2, circCHST12, circFBRS (host gene *FBRS*)] in the intersection (Supplementary Table 5). However, there was no significant difference in hsa_circ_0005806 or circFBRS expression levels between the ICH patients and controls in the validation cohort (Supplementary Figure 2). Finally, we

focused on evaluating the diagnostic value of circERBB2 and circCHST12 as potential ICH biomarkers in further statistical analysis.

Furthermore, six different classifier algorithms were executed to assess the validity of the candidate circRNAs. By using 10-fold cross-validation, the average performance measurement values of the candidate circRNAs in ICH were computed and are summarized in **Table 4**. The six machine learning classifiers based on test accuracies and AUCs in the training set and validation set are presented in **Figure 6**. The RF provides greater accuracy values of 0.995 and 0.910 than the other five classifiers in the training and test sets between ICH and controls, respectively (**Figures 6A,B**). We also evaluated the performance of the circERBB2 and circCHST12 signatures for discriminating ICH from IS patients and observed that the RF had the highest value of 0.989 in the training set and the SVM had the highest value of 0.779 in the test set (**Figures 6C,D** and **Supplementary Table 6**). These results indicate that the combination of the circERBB2 and circCHST12 signatures is capable of identifying ICH with high accuracy according to expression levels.

Correlation of the circERBB2 and circCHST12 expression levels with clinical characteristics

Additionally, we performed Spearman's correlation analysis to test the correlation of the expression levels of circCHST12

and circERBB2 with ICH patient clinical characteristics. The results showed that the circERBB2 expression levels positively correlated with HDL-C and negatively correlated with SBP, DBP and alcohol consumption in ICH patients ($P < 0.05$); the circCHST12 expression levels positively correlated with LDL-C and negatively correlated with SBP, DBP, glucose, white blood cells and alcohol consumption ($P < 0.05$) (**Table 5**). These results indicated that circERBB2 and circCHST12 may be involved in the pathogenesis of ICH.

Evaluation of the diagnostic value of circERBB2 and circCHST12 in intracerebral hemorrhage patients

Receiver operating curve (ROC) analysis was performed to explore the potential diagnostic value of circERBB2 and circCHST12. The signatures of circERBB2 for differentiating between patients with ICH and healthy control subjects showed an AUC of 0.883 (95% CI: 0.811–0.937) with a sensitivity of 68.2% and a specificity of 92%; the signatures of circCHST12 showed an AUC of 0.838 (95% CI: 0.769–0.908) with a sensitivity of 93% and a specificity of 71.6% (**Figure 7A**). The combination of circERBB2 and circCHST12 for differentiating between patients with ICH and healthy controls showed an AUC of 0.917 (95% CI: 0.869–0.965), with a sensitivity of 87.5% and a specificity of 82% (**Figure 7A**). We next performed a multifactor risk logistic regression model, the combination of circERBB2 and circCHST12 together with the risk factors (age, sex, BMI,

TABLE 4 Classification performance for the two-circRNA signatures between intracerebral hemorrhage (ICH) patients.

| | Sensitivity (%) | Specificity (%) | Accuracy (%) | PPV (%) | NPV (%) | AUC |
|--------------|-----------------|-----------------|--------------|---------|---------|--------------------|
| RF | | | | | | |
| Training set | 100 | 100 | 100 | 100 | 100 | 0.995(0.983–1) |
| Test set | 80.83 | 74.81 | 78.48 | 80.10 | 77.48 | 0.910(0.857–0.963) |
| KNN | | | | | | |
| Training set | 91.71 | 81.32 | 87.26 | 86.66 | 88.33 | 0.938(0.894–0.982) |
| Test set | 80.56 | 69.63 | 76.89 | 78.08 | 72.62 | 0.827(0.753–0.901) |
| DT | | | | | | |
| Training set | 100 | 100 | 100 | 100 | 100 | 0.995(0.983–1) |
| Test set | 77.03 | 69.81 | 73.26 | 78.21 | 66.48 | 0.734(0.644–0.824) |
| LR | | | | | | |
| Training set | 86.35 | 76.82 | 82.28 | 83.16 | 80.99 | 0.906(0.852–0.960) |
| Test set | 84.03 | 72.64 | 80.38 | 80.62 | 76.33 | 0.883(0.822–0.944) |
| GNB | | | | | | |
| Training set | 90.90 | 65.11 | 79.79 | 77.48 | 84.41 | 0.897(0.840–0.954) |
| Test set | 90.14 | 63.98 | 79.62 | 77.46 | 82.17 | 0.882(0.821–0.943) |
| SVM | | | | | | |
| Training set | 93.25 | 64.25 | 80.75 | 77.51 | 87.90 | 0.902(0.846–0.957) |
| Test set | 88.89 | 63.98 | 79.62 | 76.78 | 85.83 | 0.885(0.825–0.945) |

ICH, intracerebral hemorrhage; RF, random forest; KNN, K-nearest neighbor; LR, logistic regression; DT, decision tree; GNB, Gaussian naive Bayes; SVM, support vector machine; PPV, positive predictive value; NPV, negative predictive value; AUC, area under the curve.

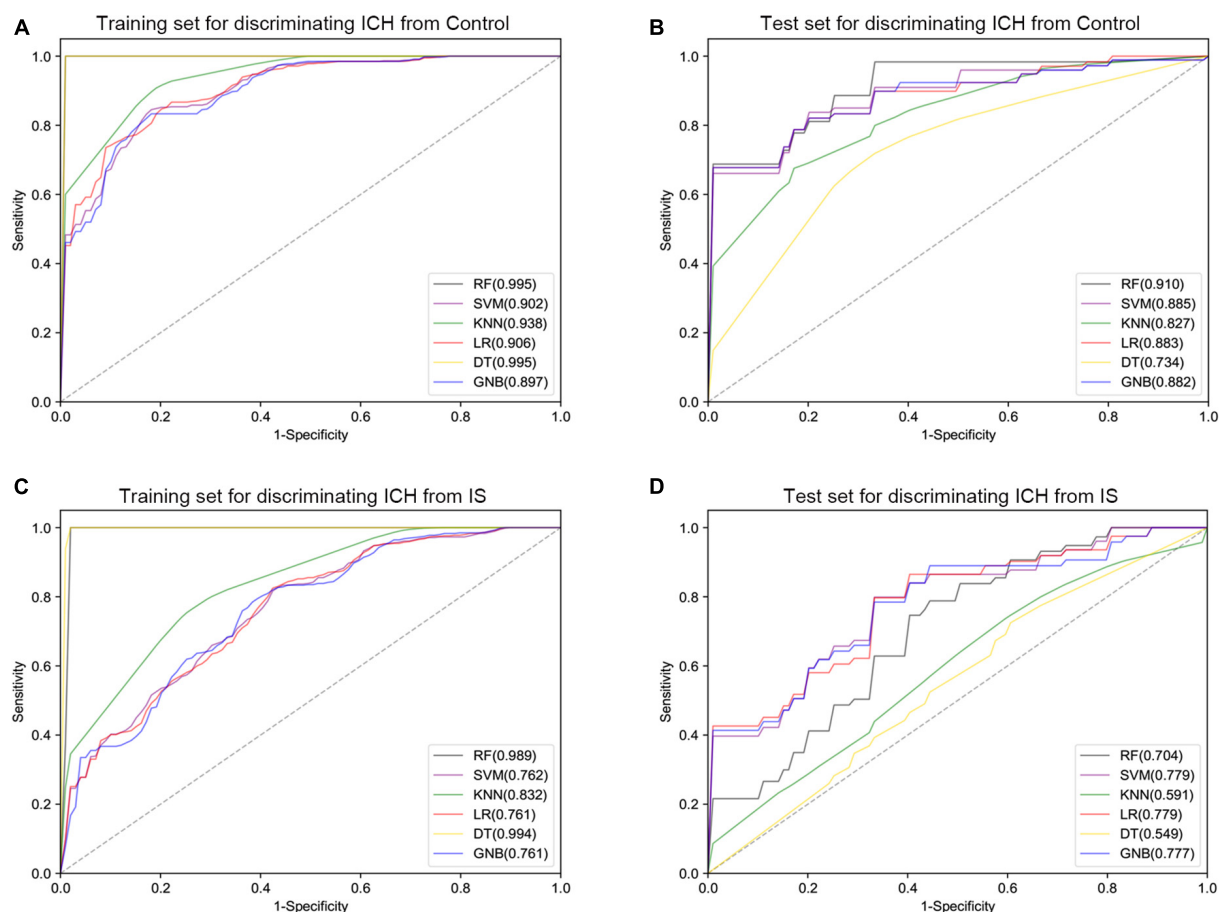


FIGURE 6

Receiver operating curve (ROC) plot of the six classifier performances based on AUC in the training set and test set. (A,B) ROC plot of the six classifier performances based on AUC in the training set (A) and test set (B) for discriminating intracerebral hemorrhage (ICH) from healthy controls. (C,D) ROC plot of the six classifier performances based on AUC in the training set (C) and test set (D) for discriminating ICH from ischemic stroke (IS) patients. SVM, support vector machine; RF, random forest; KNN, K-nearest neighbor; LR, logistic regression; DT, decision tree; GNB, Gaussian naive Bayes.

SBP, DBP, TC, TG, HDL-C, LDL-C, smoking and alcohol consumption) showed that the AUC was increased to 0.980 (95% CI: 0.959–1), the sensitivity was 93.8%, and the specificity was 96% (Figure 7B). The addition of circERBB2 and circCHST12 to the previously known risk factors improved the predictive ability, with an NRI of 20.3% and IDI of 23.7% ($P < 0.001$). The AUC of circERBB2 and circCHST12 for differentiating between ICH and IS patients was 0.765 (95% CI: 0.682–0.847); the sensitivity was 57.6%, and the specificity was 85.9% (Figure 7C).

hsa_circ_0005505 was upregulated in both ICH compared with controls and ICH compared IS patients. Furthermore, we evaluated the diagnostic values of the two novel circRNA combinations of hsa_circ_0005505 for identifying ICH. The combination of hsa_circ_0005505, circERBB2 and circCHST12 for differentiating between patients with ICH and healthy controls showed an AUC of 0.946 (95% CI: 0.910–0.982), with a sensitivity of 89.1% and a specificity of 86% (Figure 7A); the AUC was 0.799 (95% CI: 0.722–0.875), with a sensitivity

of 59.3% and a specificity of 89.5% for differentiating between patients with ICH and IS patients (Figure 7D). These results indicate that hsa_circ_0005505, novel circERBB2 and circCHST12, individually or combined, serve as potential diagnostic biomarkers for identifying ICH (Figure 8).

Discussion

In the present study, we first investigated the circRNA profiles in the peripheral blood of ICH patients and healthy controls by using RNA sequencing in two independent cohorts. Functional analysis indicated that the differentially expressed circRNAs are involved in many pathophysiologic processes of ICH. By using two independent analysis strategies, we obtained nine circRNAs that were consistently altered in both cohorts, including five upregulated circRNAs and four downregulated circRNAs. Furthermore, based on machine

TABLE 5 Correlation between baseline characteristic and circRNA levels in intracerebral hemorrhage (ICH) patients.

| Parameters | circERBB2 | | circCHST12 | |
|-------------------------|-------------|----------|-------------|---------|
| | Coefficient | P-value | Coefficient | P-value |
| Age, y | 0.143 | 0.128 | −0.018 | 0.850 |
| Sex (male) | 0.017 | 0.895 | 0.022 | 0.814 |
| BMI, kg/m ² | 0.044 | 0.646 | 0.103 | 0.274 |
| SBP, mmHg | −0.373 | < 0.001* | −0.240 | 0.010* |
| DBP, mmHg | −0.418 | < 0.001* | −0.309 | 0.001* |
| HDL-C, mmol/L | 0.190 | 0.043* | 0.153 | 0.104 |
| LDL-C, mmol/L | 0.157 | 0.096 | 0.224 | 0.016* |
| TC, mmol/L | 0.165 | 0.079 | 0.016 | 0.419 |
| TG, mmol/L | −0.085 | 0.367 | −0.164 | 0.182 |
| GLU, mmol/L | −0.06 | 0.525 | −0.273 | 0.003* |
| UA, μmol/L | 0.193 | 0.097 | 0.218 | 0.060 |
| TBIL, μmol/L | −0.023 | 0.846 | 0.001 | 0.992 |
| BUN, mmol/L | −0.027 | 0.817 | 0.094 | 0.442 |
| WBC, 10 ⁹ /L | −0.283 | 0.014* | −0.366 | 0.001* |
| Smoking | −0.063 | 0.504 | −0.153 | 0.104 |
| Alcohol consumption | −0.215 | 0.022* | −0.307 | 0.001* |

ICH, Intracerebral hemorrhage; BMI, Body mass index; SBP, Systolic blood pressure; DBP, Diastolic blood pressure; TC, Total cholesterol; TG, Triacylglycerol; HDL-C, High-density lipoprotein cholesterol; LDL-C, Low-density lipoprotein cholesterol; GLU, Glucose; UA, Uric acid; TBIL, Total bilirubin; BUN, Blood urea nitrogen; WBC, White blood cell. * $p < 0.05$.

learning classification, we screened two candidates, circERBB2 and circCHST12, to explore their diagnostic value as potential biomarkers in ICH patients. The AUC was 0.917 (95% CI: 0.869–0.965), with a sensitivity of 87.5% and a specificity of 82% for distinguishing between ICH patients and healthy controls. In combination with ICH risk factors, the AUC was 0.980 (95% CI: 0.959–1), sensitivity was 93.8% and specificity was 96% in ICH diagnosis. Moreover, logistic regression analysis and Spearman's correlation test demonstrated that downregulation of circERBB2 and circCHST12 may be independent risk factors for ICH. Additionally, the expression level of circERBB2 correlated with SBP and HDL-C; circCHST12 expression levels correlated with LDL-C, SBP, DBP and white blood cells, indicating that circERBB2 and circCHST12 might be heavily involved in the pathology of ICH. Our data show that circERBB2 and circCHST12 may be novel biomarkers for ICH diagnosis. Together with hsa_circ_0005505, circERBB2 and circCHST12 showed high accuracy for identifying ICH. A previous study revealed that hsa_circ_0005505 was upregulated in ruptured intracranial aneurysm tissues, promoted proliferation and migration and suppressed apoptosis of vascular smooth muscle cells *in vitro* (Chen X. et al., 2021), indicating that hsa_circ_0005505 may be associated with the pathological process of cerebrovascular diseases.

Intracerebral hemorrhage (ICH) is a multifactorial disease with high incidence and mortality that imposes a large

socioeconomic burden. Identifying novel potential biomarkers for the early diagnosis of ICH would be part of risk prediction. CircRNAs are produced by host gene back-splicing, and closed RNAs without a free 3' or 5' end are resistant to exonuclease digestion (Jeck and Sharpless, 2014), which makes them more stable and better biomarkers of human disease. Furthermore, circRNAs are highly expressed in many tissues, particularly the human brain, and in blood (Patop et al., 2019). There is growing evidence that the circRNA expression profile is altered in IS (Dong et al., 2020; Ostolaza et al., 2020; Zuo et al., 2020; Liu Y. et al., 2022), indicating that circRNAs have the potential to serve as biomarkers and therapeutic targets in IS. Moreover, the circRNA expression profiles were altered in rat brain tissues after ICH (Zhong et al., 2020; Bai et al., 2021). However, the changes in circRNA expression in the peripheral blood of ICH patients remain unclear. Our previous study demonstrated that hsa_circ_0001240, hsa_circ_0001947 and hsa_circ_0001386 were promising biomarkers for predicting and diagnosing hypertensive ICH (Bai et al., 2021). In this study, we first investigated whether circRNA profiles were significantly altered between ICH patients and healthy controls, which provides new insights into understanding the epigenomic mechanisms of ICH.

In this study, we found that circERBB2 may serve as a novel biomarker in ICH diagnosis. Previous studies have identified blood biomarkers, such as glial fibrillary acid protein (GFAP), retinol binding protein 4 and N-terminal pro B-type natriuretic peptide, that distinguish IS from ICH with moderate accuracy (Bustamante et al., 2021) and metabolic biomarkers for ICH diagnosis (Zhang et al., 2021). The AUCs of S100 and IL6 were 0.65 and 0.59 (Bhatia et al., 2020), respectively, and GFAP had a sensitivity of 78% and a specificity of 95% between ICH and IS (Kumar et al., 2020). ncRNAs have been identified as critical novel regulators of cardiovascular risk factors and cell functions and are thus important candidates to improve diagnostics and prognosis assessment (Poller et al., 2018). In the present study, we identified that the AUC of circERBB2 was 0.883 for distinguishing between ICH patients and healthy controls, with a sensitivity and specificity of 68.2% and 92%, respectively. The signatures of circCHST12 showed an AUC of 0.838 with a sensitivity of 93% and a specificity of 71.6%. The combination of circERBB2 and circCHST12 with ICH risk factors increased the predictive value for the identification of ICH. These findings were better than the diagnostic value of three previously identified circRNAs [hsa_circ_0001240 (AUC = 0.808), hsa_circ_0001947 (AUC = 0.798) and hsa_circ_0001386 (AUC = 0.806)] in ICH (Bai et al., 2021). Additionally, we observed that downregulation of circERBB2 was positively associated with HDL-C and negatively correlated with SBP and DBP. Lowering blood lipids was associated with an increased risk of ICH (Sun et al., 2019), and high blood pressure was found to be the most prevalent stroke risk factor (Feigin et al., 2016; Wang et al., 2017). Thus,

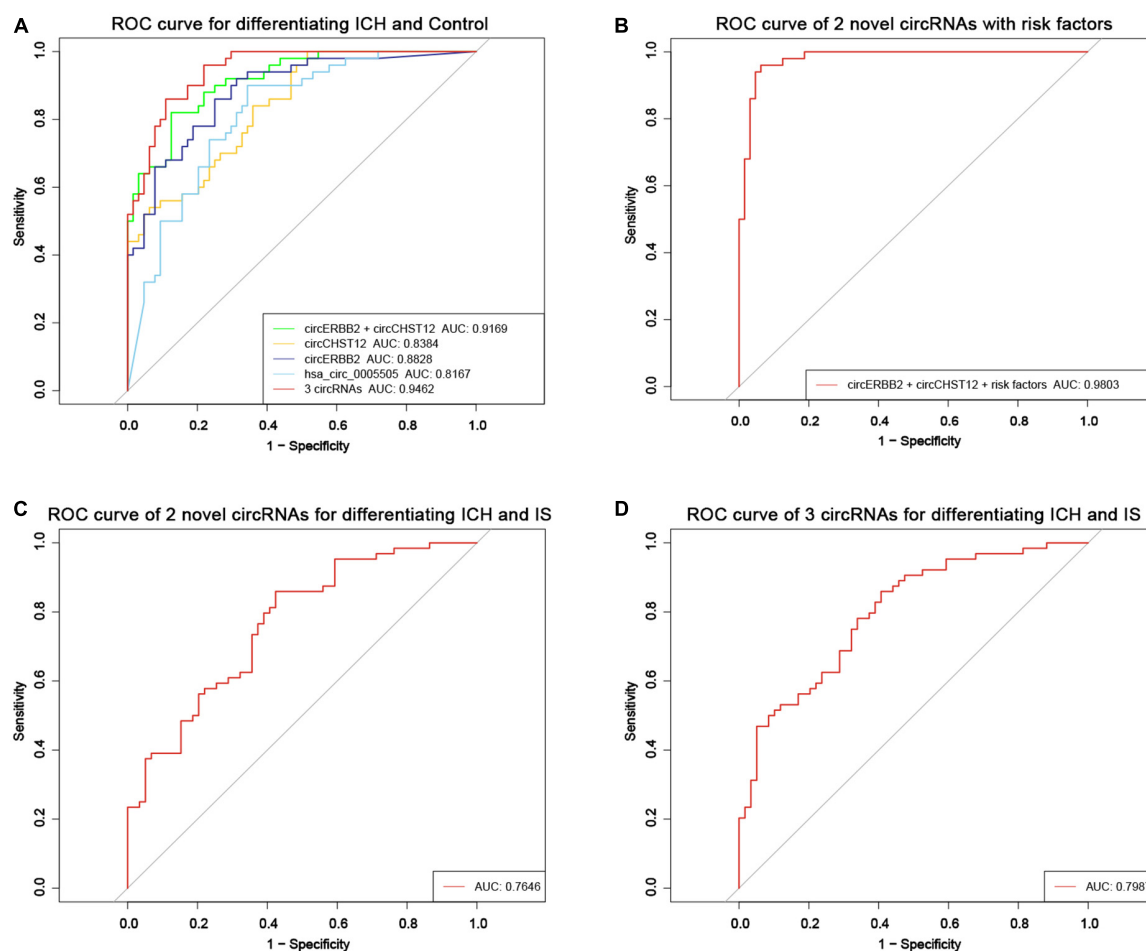


FIGURE 7

Evaluation of the circRNA diagnostic value in ICH patients. (A) Receiver operating characteristic (ROC) curves were calculated using the expression levels of circERBB2, circCHST12 and hsa_circ_0005505 for differentiating patients with intracerebral hemorrhage (ICH) and healthy controls ($n = 64$ vs. 50). (B) ROC curves of combining circERBB2 and circCHST12 with ICH risk factors to differentiate patients with ICH and healthy controls in all samples ($n = 64$ vs. 50). (C) ROC curves of combining circERBB2 and circCHST12 for differentiating patients with ICH and IS patients in all samples ($n = 64$ vs. 59). (D) ROC curves of two novel circRNAs, circERBB2 and circCHST12, combined with hsa_circ_0005505 for differentiating patients with ICH and IS patients in all samples ($n = 64$ vs. 59).

we speculate that a decrease in circERBB2 expression levels might correlate with an increased risk of ICH occurrence. These findings indicate that circERBB2 might play vital roles in the pathogenesis and pathology of ICH.

The protein ERBB2 is a member of a family of epidermal growth factor receptors that are involved in aberrant signaling and cell migration, growth, adhesion, and differentiation (Strickler et al., 2022). A previous study demonstrated that circERBB2 (chr17: 39,708,320–39,710,481; length: 676 bp) serves as an important regulator of cancer cell proliferation and has the potential to be a new therapeutic target for gallbladder cancer (Huang et al., 2019) and breast cancer (Huang Y. et al., 2021). Our study identified circERBB2 (chr17: 37,866,065–37,872,192; genomic length: 6127 bp, spliced sequence length: 939 bp), which is a novel back-splicing circRNA that has never been reported thus far, at a different chromosomal

position. Carbohydrate sulfotransferases (CHSTs) are a class of key enzymes that contribute to tissue remodeling. CHST12 is a significant member of the CHST family, and a previous study demonstrated that CHST12 may be a novel biomarker for glioblastoma; it regulates cell proliferation and mobility via the WNT/ β -catenin pathway (Wang et al., 2021). One study reported that hsa_circ_0134005 (chr7:2472197–2477555; genomic length: 5358 bp, spliced sequence length: 5358 bp) is derived from the CHST12 gene (Rybak-Wolf et al., 2015). This study identified circCHST12 (chr7:2477438–2483381; genomic length: 5943 bp, spliced sequence length: 5943 bp) derived from exon 1 and partial exon 2 of the CHST12 gene, which is a novel back-splicing circRNA that has never been reported thus far at a different chromosomal position.

CircRNAs are involved in the translational and transcriptional regulation of the pathological mechanisms

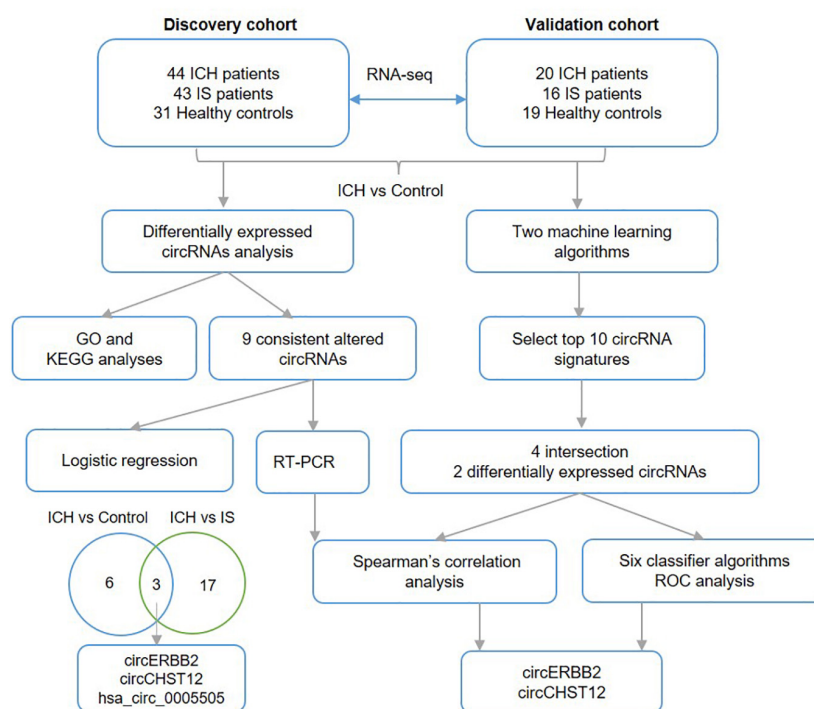


FIGURE 8

Work flow. The diagram of the data analysis process in this study.

of many disorders (Shan et al., 2017; Aufero et al., 2019). CircRNAs can act as miRNA sponges and are expected to influence downstream miRNA function, further regulating the expression levels of target mRNAs (Hansen et al., 2013). We performed GO and KEGG analyses to investigate the enrichment of differentially expressed circRNAs. Functional analysis demonstrated that the circRNA host genes were mainly involved in GTPase activity, covalent chromatin modification, histone modification, the MAPK signaling pathway and the ERBB signaling pathway. Activation of the MAPK signaling pathway is involved in the progression of injury following ICH (Ding et al., 2020; Guo et al., 2020). Recently, research identified that knockdown of circERBB2 suppressed the PDGF-BB-induced proliferation, migration, and inflammatory response of human airway smooth muscle cells via miR-98-5p/IGF1R signaling (Huang J. Q. et al., 2021). The phenotype of smooth muscle cells transforming from a contractile to a synthetic phenotype plays an essential role in the onset of brain vascular pathological progression (Bennett et al., 2016; Rho et al., 2017). In this study, we speculated that the downregulation of the novel circERBB2 in ICH patients might contribute to the pathogenesis of ICH via the phenotype of smooth muscle cell transformation.

Notably, there are some limitations of this study. First, we should perform a larger multicenter study with more participants to externally validate the candidate biomarkers.

Second, further studies should be performed to explore how hsa_circ_0005505, circERBB2 and circCHST12 contribute to the pathogenesis and development of ICH with cell- or animal-based experiments. Additionally, our study lacked follow-up information for ICH patients, and the prognostic value of these candidate circRNAs should be assessed in subsequent studies. We expect that hsa_circ_0005505, circERBB2 and circCHST12 will provide new insights for a better understanding of the pathogenesis of ICH and help to improve the diagnosis and prognostic assessment of ICH in clinical practice.

Conclusion

In this study, we provided a transcriptome-wide overview of aberrantly expressed circRNAs in the peripheral blood of ICH patients and identified hsa_circ_0005505 and novel circERBB2 and circCHST12 as promising biomarkers for diagnosing ICH based on machine learning algorithms.

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 Human Ethics Committee, Fuwai Hospital (Approval No. 2016-732). The patients/participants provided their written informed consent to participate in this study.

Author contributions

CB, YS, and LS: design and experiment. XH and FW: data analyses. CB and LZ: manuscript preparation. JL, LY, and JC: manuscript review. All authors contributed to the article and approved the submitted version.

Funding

This study was supported by the National Natural Science Foundation of China (91539113 and 82130013 to JC), the National Basic Research Program of China (2014CB541601 to JC), and the CAMS Innovation Fund for

Medical Sciences (2021-CXGC02-3CAMS-I2 M and 2021-1-I2 M-007 to JC).

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/fnins.2022.1002590/full#supplementary-material>

References

- Ambale-Venkatesh, B., Yang, X., Wu, C. O., Liu, K., Hundley, W. G., McClelland, R., et al. (2017). Cardiovascular event prediction by machine learning: The multi-ethnic study of atherosclerosis. *Circ. Res.* 121, 1092–1101. doi: 10.1161/CIRCRESAHA.117.311312
- Aufiero, S., Reckman, Y. J., Pinto, Y. M., and Creemers, E. E. (2019). Circular RNAs open a new chapter in cardiovascular biology. *Nat. Rev. Cardiol.* 16, 503–514. doi: 10.1038/s41569-019-0185-2
- Bai, C., Liu, T., Sun, Y., Li, H., Xiao, N., Zhang, M., et al. (2021). Identification of circular RNA expression profiles and potential biomarkers for intracerebral hemorrhage. *Epigenomics* 13, 379–395. doi: 10.2217/epi-2020-0432
- Benjamin, E. J., Blaha, M. J., Chiuve, S. E., Cushman, M., Das, S. R., Deo, R., et al. (2017). Heart disease and stroke statistics-2017 update: A report from the American heart association. *Circulation* 135, e146–e603. doi: 10.1161/CIR.0000000000000485
- Bennett, M. R., Sinha, S., and Owens, G. K. (2016). Vascular smooth muscle cells in atherosclerosis. *Circ. Res.* 118, 692–702. doi: 10.1161/CIRCRESAHA.115.306361
- Bhatia, R., Warrier, A. R., Sreenivas, V., Bali, P., Sisodia, P., Gupta, A., et al. (2020). Role of blood biomarkers in differentiating ischemic stroke and intracerebral hemorrhage. *Neurol. India* 68, 824–829. doi: 10.4103/0028-3886.293467
- Blokh, D., and Stambler, I. (2017). The application of information theory for the research of aging and aging-related diseases. *Prog. Neurobiol.* 157, 158–173. doi: 10.1016/j.pneurobio.2016.03.005
- Bustamante, A., Penalba, A., Orset, C., Azurmendi, L., Llobart, V., Simats, A., et al. (2021). Blood biomarkers to differentiate ischemic and hemorrhagic strokes. *Neurology* 96, e1928–e1939. doi: 10.1212/WNL.00000000000011742
- Cardona-Monzonis, A., Garcia-Gimenez, J. L., Mena-Molla, S., Pareja-Galeano, H., de la Guia-Galipienso, F., and Pallardo, F. V. (2020). Non-coding RNAs and coronary artery disease. *Adv. Exp. Med. Biol.* 1229, 273–285. doi: 10.1007/978-981-15-1671-9_16
- Chang, C. H., Lin, C. H., and Lane, H. Y. (2021). Machine learning and novel biomarkers for the diagnosis of Alzheimer's disease. *Int. J. Mol. Sci.* 22:2761. doi: 10.3390/ijms22052761
- Chen, T., Chen, X., Zhang, S., Zhu, J., Tang, B., Wang, A., et al. (2021). The genome sequence archive family: Toward explosive data growth and diverse data types. *Genom. Proteom. Bioinform.* 19, 578–583.
- Chen, X., Yang, S., Yang, J., Liu, Q., Li, M., Wu, J., et al. (2021). The potential role of hsa_circ_0005505 in the rupture of human intracranial aneurysm. *Front. Mol. Biosci.* 8:670691. doi: 10.3389/fmolb.2021.670691
- Chen, Y., Chen, B., Song, X., Kang, Q., Ye, X., and Zhang, B. (2021). A data-driven binary-classification framework for oil fingerprinting analysis. *Environ. Res.* 201:111454. doi: 10.1016/j.envres.2021.111454
- Cheng, X., Ander, B. P., Jickling, G. C., Zhan, X., Hull, H., Sharp, F. R., et al. (2020). MicroRNA and their target mRNAs change expression in whole blood of patients after intracerebral hemorrhage. *J. Cereb. Blood Flow Metab.* 40, 775–786. doi: 10.1177/0271678X19839501
- CNCB-NGDC Members and Partners (2022). Database resources of the national genomics data center, china national center for bioinformatics in 2022. *Nucleic Acids Res.* 50, D27–D38.
- Ding, Y., Flores, J., Klebe, D., Li, P., McBride, D. W., Tang, J., et al. (2020). Annexin A1 attenuates neuroinflammation through FPR2/p38/COX-2 pathway after intracerebral hemorrhage in male mice. *J. Neurosci. Res.* 98, 168–178. doi: 10.1002/jnr.24478
- Dong, Z., Deng, L., Peng, Q., Pan, J., and Wang, Y. (2020). CircRNA expression profiles and function prediction in peripheral blood mononuclear cells of patients with acute ischemic stroke. *J. Cell. Physiol.* 235, 2609–2618. doi: 10.1002/jcp.29165

- Dou, Z., Yu, Q., Wang, G., Wu, S., Reis, C., Ruan, W., et al. (2020). Circular RNA expression profiles alter significantly after intracerebral hemorrhage in rats. *Brain Res.* 1726:146490. doi: 10.1016/j.brainres.2019.146490
- Feigin, V. L., Roth, G. A., Naghavi, M., Parmar, P., Krishnamurthi, R., Chugh, S., et al. (2016). Global burden of stroke and risk factors in 188 countries, during 1990–2013: A systematic analysis for the global burden of disease study 2013. *Lancet Neurol.* 15, 913–924. doi: 10.1016/S1474-4422(16)30073-4
- Gao, Y., Zhang, J., and Zhao, F. (2018). Circular RNA identification based on multiple seed matching. *Brief. Bioinform.* 19, 803–810. doi: 10.1093/bib/bb1014
- Guo, F., Xu, D., Lin, Y., Wang, G., Wang, F., Gao, Q., et al. (2020). Chemokine CCL2 contributes to BBB disruption via the p38 MAPK signaling pathway following acute intracerebral hemorrhage. *FASEB J.* 34, 1872–1884. doi: 10.1096/fj.201902203RR
- Hankey, G. J. (2017). Stroke. *Lancet* 389, 641–654. doi: 10.1016/S0140-6736(16)30962-X
- 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, 384–388. doi: 10.1038/nature11993
- Huang, J. Q., Wang, F., Wang, L. T., Li, Y. M., Lu, J. L., and Chen, J. Y. (2021). Circular RNA ERBB2 promotes gallbladder cancer progression by regulating PA2G4-dependent rDNA transcription. *Mol. Cancer* 18:166. doi: 10.1186/s12943-019-1098-8
- Huang, Y., Zheng, S., Lin, Y., and Ke, L. (2021). Circular RNA circ-ERBB2 elevates the warburg effect and facilitates triple-negative breast cancer growth by the MicroRNA 136-5p/pyruvate dehydrogenase kinase 4 axis. *Mol. Cell. Biol.* 41:e0060920. doi: 10.1128/MCB.00609-20
- Jeck, W. R., and Sharpless, N. E. (2014). Detecting and characterizing circular RNAs. *Nat. Biotechnol.* 32, 453–461. doi: 10.1038/nbt.2890
- Kawakami, E., Tabata, J., Yanaihara, N., Ishikawa, T., Koseki, K., Iida, Y., et al. (2019). Application of artificial intelligence for preoperative diagnostic and prognostic prediction in epithelial ovarian cancer based on blood biomarkers. *Clin. Cancer Res.* 25, 3006–3015. doi: 10.1158/1078-0432.CCR-18-3378
- Kim, J. M., Moon, J., Yu, J. S., Park, D. K., Lee, S. T., Jung, K. H., et al. (2019). Altered long noncoding RNA profile after intracerebral hemorrhage. *Ann. Clin. Transl. Neurol.* 6, 2014–2025. doi: 10.1002/acn3.50894
- Kristensen, L. S., Andersen, M. S., Stagsted, L. V. W., Ebbesen, K. K., Hansen, T. B., and Kjems, J. (2019). The biogenesis, biology and characterization of circular RNAs. *Nat. Rev. Genet.* 20, 675–691. doi: 10.1038/s41576-019-0158-7
- Kristensen, L. S., Jakobsen, T., Hager, H., and Kjems, J. (2022). The emerging roles of circRNAs in cancer and oncology. *Nat. Rev. Clin. Oncol.* 19, 188–206. doi: 10.1038/s41571-021-00585-y
- Kumar, A., Misra, S., Yadav, A. K., Sagar, R., Verma, B., Grover, A., et al. (2020). Role of glial fibrillary acidic protein as a biomarker in differentiating intracerebral haemorrhage from ischaemic stroke and stroke mimics: A meta-analysis. *Biomarkers* 25, 1–8. doi: 10.1080/1354750X.2019.1691657
- Ledesma, D., Symes, S., and Richards, S. (2021). Advancements within modern machine learning methodology: Impacts and prospects in biomarker discovery. *Curr. Med. Chem.* 28, 6512–6531. doi: 10.2174/0929867328666210208111821
- Li, L., Wang, P., Zhao, H., and Luo, Y. (2019). Noncoding RNAs and intracerebral hemorrhage. *CNS Neurol. Disord. Drug Targets* 18, 205–211. doi: 10.2174/1871527318666190204102604
- Li, R. (2018). data mining and machine learning methods for dementia research. *Methods Mol. Biol.* 1750, 363–370. doi: 10.1007/978-1-4939-7704-8_25
- Li, S., Chen, L., Xu, C., Qu, X., Qin, Z., Gao, J., et al. (2020). Expression profile and bioinformatics analysis of circular RNAs in acute ischemic stroke in a South Chinese han population. *Sci. Rep.* 10:10138. doi: 10.1038/s41598-020-66990-y
- Liu, D., Zhao, L., Jiang, Y., Li, L., Guo, M., Mu, Y., et al. (2022). Integrated analysis of plasma and urine reveals unique metabolomic profiles in idiopathic inflammatory myopathies subtypes. *J. Cachexia Sarcopenia Muscle* 13, 2456–2472. doi: 10.1002/jcsm.13045
- Liu, Y., Li, Y., Zang, J., Zhang, T., Li, Y., Tan, Z., et al. (2022). CircOGDH Is a penumbra biomarker and therapeutic target in acute ischemic stroke. *Circ. Res.* 130, 907–924. doi: 10.1161/CIRCRESAHA.121.319412
- Lu, D., Ho, E. S., Mai, H., Zang, J., Liu, Y., Li, Y., et al. (2020). Identification of blood circular RNAs as potential biomarkers for acute ischemic stroke. *Front. Neurosci.* 14:81. doi: 10.3389/fnins.2020.00081
- Montaner, J., Ramiro, L., Simats, A., Tiedt, S., Makris, K., Jickling, G. C., et al. (2020). Multilevel omics for the discovery of biomarkers and therapeutic targets for stroke. *Nat. Rev. Neurol.* 16, 247–264. doi: 10.1038/s41582-020-0350-6
- Ostolaza, A., Blanco-Luquin, I., Urdanoz-Casado, A., Rubio, I., Labarga, A., Zandio, B., et al. (2020). Circular RNA expression profile in blood according to ischemic stroke etiology. *Cell Biosci.* 10:34. doi: 10.1186/s13578-020-00394-3
- Patop, I. L., Wust, S., and Kadener, S. (2019). Past, present, and future of circRNAs. *EMBO J.* 38:e100836. doi: 10.15252/embj.2018100836
- Poller, W., Dimmeler, S., Heymans, S., Zeller, T., Haas, J., Karakas, M., et al. (2018). Non-coding RNAs in cardiovascular diseases: Diagnostic and therapeutic perspectives. *Eur. Heart J.* 39, 2704–2716. doi: 10.1093/eurheartj/ehx165
- Qureshi, A. I., Mendelow, A. D., and Hanley, D. F. (2009). Intracerebral haemorrhage. *Lancet* 373, 1632–1644. doi: 10.1016/S0140-6736(09)60371-8
- Rho, S. S., Ando, K., and Fukuhara, S. (2017). Dynamic regulation of vascular permeability by vascular endothelial cadherin-mediated endothelial cell-cell junctions. *J. Nippon Med. Sch.* 84, 148–159. doi: 10.1272/jnms.84.148
- Robinson, M. D., McCarthy, D. J., and Smyth, G. K. (2010). edgeR: A bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26, 139–140. doi: 10.1093/bioinformatics/btp616
- Rybak-Wolf, A., Stottmeister, C., Glazar, P., Jens, M., Pino, N., Giusti, S., et al. (2015). Circular RNAs in the mammalian brain are highly abundant, conserved, and dynamically expressed. *Mol. Cell* 58, 870–885. doi: 10.1016/j.molcel.2015.03.027
- Shan, K., Liu, C., Liu, B. H., Chen, X., Dong, R., Liu, X., et al. (2017). Circular noncoding RNA HIPK3 mediates retinal vascular dysfunction in diabetes mellitus. *Circulation* 136, 1629–1642. doi: 10.1161/CIRCULATIONAHA.117.029004
- Shu, T., Ning, W., Wu, D., Xu, J., Han, Q., Huang, M., et al. (2020). Plasma proteomics identify biomarkers and pathogenesis of COVID-19. *Immunity* 53, 1108–1122e5. doi: 10.1016/j.immuni.2020.10.008
- Strickler, J. H., Yoshino, T., Graham, R. P., Siena, S., and Bekaii-Saab, T. (2022). Diagnosis and treatment of ERBB2-positive metastatic colorectal cancer: A review. *JAMA Oncol.* 8, 760–769. doi: 10.1001/jamaoncol.2021.8196
- Sun, L., Clarke, R., Bennett, D., Guo, Y., Walters, R. G., Hill, M., et al. (2019). Causal associations of blood lipids with risk of ischemic stroke and intracerebral hemorrhage in Chinese adults. *Nat. Med.* 25, 569–574. doi: 10.1038/s41591-019-0366-x
- Tiedt, S., Prestel, M., Malik, R., Schieferdecker, N., Duering, M., Kautzky, V., et al. (2017). RNA-seq identifies circulating miR-125a-5p, miR-125b-5p, and miR-143-3p as potential biomarkers for acute ischemic stroke. *Circ. Res.* 121, 970–980. doi: 10.1161/CIRCRESAHA.117.311572
- Wang, J., Xia, X., Tao, X., Zhao, P., and Deng, C. (2021). Knockdown of carbohydrate sulfotransferase 12 decreases the proliferation and mobility of glioblastoma cells via the WNT/beta-catenin pathway. *Bioengineered* 12, 3934–3946. doi: 10.1080/21655979.2021.1944455
- Wang, L., Feng, Z., Wang, X., Wang, X., and Zhang, X. (2010). DEGseq: An R package for identifying differentially expressed genes from RNA-seq data. *Bioinformatics* 26, 136–138. doi: 10.1093/bioinformatics/btp612
- Wang, W., Jiang, B., Sun, H., Ru, X., Sun, D., Wang, L., et al. (2017). Prevalence, incidence, and mortality of stroke in china: Results from a nationwide population-based survey of 480 687 adults. *Circulation* 135, 759–771. doi: 10.1161/CIRCULATIONAHA.116.025250
- Weng, R., Jiang, Z., and Gu, Y. (2022). Noncoding RNA as diagnostic and prognostic biomarkers in cerebrovascular disease. *Oxid. Med. Cell. Longev.* 2022:8149701. doi: 10.1155/2022/8149701
- Wilkinson, D. A., Pandey, A. S., Thompson, B. G., Keep, R. F., Hua, Y., and Xi, G. (2018). Injury mechanisms in acute intracerebral hemorrhage. *Neuropharmacology* 134(Pt B), 240–248. doi: 10.1016/j.neuropharm.2017.09.033
- Wu, J., Zhang, H., Li, L., Hu, M., Chen, L., Xu, B., et al. (2020). A nomogram for predicting overall survival in patients with low-grade endometrial stromal sarcoma: A population-based analysis. *Cancer Commun.* 40, 301–312. doi: 10.1002/cac2.12067

Zhang, J., Su, X., Qi, A., Liu, L., Zhang, L., Zhong, Y., et al. (2021). Metabolomic profiling of fatty acid biomarkers for intracerebral hemorrhage stroke. *Talanta* 222:121679. doi: 10.1016/j.talanta.2020.121679

Zhang, Z., Yang, T., and Xiao, J. (2018). Circular RNAs: Promising biomarkers for human diseases. *EBioMedicine* 34, 267–274. doi: 10.1016/j.ebiom.2018.07.036

Zhong, Y., Li, X., Li, C., Li, Y., He, Y., Li, F., et al. (2020). Intracerebral hemorrhage alters circular RNA expression profiles in the rat brain. *Am. J. Transl. Res.* 12, 4160–4174.

Zuo, L., Zhang, L., Zu, J., Wang, Z., Han, B., Chen, B., et al. (2020). Circulating circular RNAs as biomarkers for the diagnosis and prediction of outcomes in acute ischemic stroke. *Stroke* 51, 319–323. doi: 10.1161/STROKEAHA.119.027348



OPEN ACCESS

EDITED BY
Clévia Rosset,
Clinical Hospital of Porto Alegre, Brazil

REVIEWED BY
Ulrike Schumann,
Australian National University, Australia
Fan Zhang,
Harbin Medical University, China

*CORRESPONDENCE
Jianzhong Li,
✉ jianzhong-0520@163.com
Shouping Gong,
✉ shpingg@126.com

[†]These authors have contributed equally
to this work

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

RECEIVED 04 September 2022
ACCEPTED 09 December 2022
PUBLISHED 04 January 2023

CITATION
Tian Y, Yu B, Lv B, Zhang Y, Fu L, Yang S,
Li J and Gong S (2023), Experimental
verification and comprehensive analysis
of m7G methylation regulators in the
subcluster classification of
ischemic stroke.
Front. Genet. 13:1036345.
doi: 10.3389/fgene.2022.1036345

COPYRIGHT
© 2023 Tian, Yu, Lv, Zhang, Fu, Yang, Li
and Gong. This is an open-access article
distributed under the terms of the
[Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/)
(CC BY). The use, distribution or
reproduction in other forums is
permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original
publication in this journal is cited, in
accordance with accepted academic
practice. No use, distribution or
reproduction is permitted which does
not comply with these terms.

Experimental verification and comprehensive analysis of m7G methylation regulators in the subcluster classification of ischemic stroke

Yunze Tian^{1,2†}, Beibei Yu^{1†}, Boqiang Lv¹, Yongfeng Zhang¹,
Longhui Fu¹, Shijie Yang¹, Jianzhong Li^{2*} and Shouping Gong^{1*}

¹Department of Neurosurgery, The Second Affiliated Hospital of Xi'an Jiao Tong University, Xi'an, China, ²Department of Thoracic Surgery, The Second Affiliated Hospital of Xi'an Jiao Tong University, Xi'an, China

Background: Ischemic stroke (IS) is a fatal cerebrovascular disease involving several pathological mechanisms. Modification of 7-methylguanosine (m7G) has multiple regulatory functions. However, the expression pattern and mechanism of m7G in IS remain unknown. Herein, we aimed to explore the effect of m7G modification on IS.

Methods: We screened significantly different m7G-regulated genes in Gene Expression Omnibus datasets, GSE58294 and GSE22255. The random forest (RF) algorithm was selected to identify key m7G-regulated genes that were subsequently validated using the middle cerebral artery occlusion (MCAO) model and quantitative polymerase chain reaction (qPCR). A risk model was subsequently generated using key m7G-regulated genes. Then, "ConsensusClusterPlus" package was used to distinguish different m7G clusters of patients with IS. Simultaneously, between two m7G clusters, differentially expressed genes (DEGs) and immune infiltration differences were also explored. Finally, we investigated functional enrichment and the mRNA-miRNA-transcription factor network of DEGs.

Results: RF and qPCR confirmed that *EIF3D*, *CYFIP2*, *NCBP2*, *DCPS*, and *NUDT1* were key m7G-related genes in IS that could accurately predict clinical risk (area under the curve = 0.967). *NCBP2* was the most significantly associated gene with immune infiltration. Based on the expression profiles of these key m7G-related genes, the IS group could be divided into two clusters. According to the single-sample gene set enrichment analysis algorithm, four types of immune cells (immature dendritic cells, macrophages, natural killer T cells, and TH1 cells) were significantly different in the two m7G clusters. The functional enrichment of 282 DEGs between the two clusters was mainly concentrated in the "regulation of apoptotic signaling pathway," "cellular response to DNA damage stimulus," "adaptive immune system," and "pyroptosis." The miR-214-*LTF-FOXJ1* axis may be a key regulatory pathway for IS.

Conclusion: Our findings suggest that *EIF3D*, *CYFIP2*, *NCBP2*, *DCPS*, and *NUDT1* may serve as potential diagnostic biomarkers for IS and that the m7G clusters developed by these genes provide more evidence for the regulation of m7G in IS.

KEYWORDS

ischemic stroke, modification of 7-methylguanosine, immunity, consensus clustering, transcription factor

1 Introduction

Ischemic stroke (IS) is the most common cerebrovascular disease, with high mortality and morbidity. It affects approximately 15 million people worldwide, of which approximately 5 million die and 5 million are disabled for life (Maida et al., 2020). In recent years, with the aging of the population, the risk of IS has greatly increased, resulting in great pain and economic burden to patients (Matsuzono et al., 2021). Currently, studies in this area mainly focus on the regulation of pathological mechanisms, including apoptosis, inflammation, oxidative stress, and calcium overload (Feske, 2021). Multiple genes and regulatory methods are involved in IS, such as phosphorylation signal transduction and RNA methylation modification (Zhang et al., 2020). Identifying key genes and intervening in their regulation can improve the IS prognosis and provide newer ideas for its treatment.

Recently, the role of RNA modifications in gene regulation has received increasing attention. More than 150 RNA modification methods have been discovered, of which methylation modifications are the most abundant (Chen et al., 2019). Methylation modifications include 1-methyladenosine, 5-methyluridine, 5-methylcytidine (m5C), and G methylation of m1G, m2G, and m7G, 2'-O-ribonucleoside, and N6-methyladenosine (m6A) (Yang et al., 2021). Modification of 7-methylguanosine (m7G) is one of the most common base modifications in post-transcriptional regulation. It is widely distributed in the 5' cap region of tRNA, rRNA, and eukaryotic mRNA (Tomikawa, 2018). Zhao et al. (Zhao et al., 2021) found that m7G-regulated genes are differentially expressed and induce angiogenesis in other ischemic diseases. In addition, m7G-regulated genes play an irreplaceable role in many diseases, such as tumors and gastrointestinal diseases (Dai et al., 2021). However, the exact regulatory role of m7G-regulated genes in IS remains unclear.

To the best of our knowledge, this is the first study to explore the epigenetic role of m7G-regulated genes in IS. After screening using machine learning, we identified five m7G-regulated genes involved in IS using the middle cerebral artery occlusion (MCAO) animal model, which were clearly clustered IS patients into two m7G clusters, and the immune infiltration of each cluster was further analyzed. Through functional enrichment and the mRNA-miRNA-transcription

factor (TF) network, we further revealed the biological functions and regulation modes of different m7G clusters. This study provides a novel m7G cluster method that extensively participates in the regulation of IS occurrence and treatment.

2 Methods

2.1 Data collection

Two IS-related mRNA expression profiling datasets, GSE58294 and GSE22255, were downloaded from the Gene Expression Omnibus (GEO) database using the R package “GEOquery.” GSE58294 contains 92 samples, including 23 control samples and 69 IS samples, whereas GSE22255 contains 20 patients with IS and 20 healthy individuals. These samples were all detected by GPL570 probe (Affymetrix Human Genome U133 Plus 2.0 Array). The “normalizeBetween-Arrays” function of the “limma” package was used to normalize the expression matrix. The gene probes were annotated using official symbols. We calculated the mean values if multiple gene probes matched the same gene.

2.2 Establishment of the middle cerebral artery occlusion (MCAO) model

In total, 200–240 g Sprague–Dawley rats were purchased from the Animal Experiment Center of Xi'an Jiaotong University. Rat MCAO model was established, as previously developed and described (Longa et al., 1989). In brief, the external carotid artery of the rat was carefully isolated and an incision was made. A suture (RWD, Shenzhen, China) with a head diameter of approximately 0.34 ± 0.01 mm was inserted from the incision in the external carotid artery into the internal carotid artery up to the middle cerebral artery. Two hours later, the suture was removed and the wound was sutured. After 3 days, the rats were euthanized. The rat brain was snap-frozen, cut into 2-mm coronal slices, and immersed in 2, 3, 5-triphenyl tetrazolium chloride (TTC) solution in a 37°C water bath for 30 min. Images were taken using a digital camera after dyeing.

TABLE 1 Specific primers used for quantitative real-time PCR.

| Primer name | Sequence |
|-------------|----------------------------|
| GAPDH-F | TGCCACTCAGAAGACTGTGG |
| GAPDH-R | TTCAGCTCTGGGATGACCTT |
| NCBP2-F | AGCGTGTGGGTTCTGTTTCGTG |
| NCBP2-R | CATACTGCCTGCCCTCCTTAAAGC |
| CYFIP1-F | GATGGTGAGAGGATTGCAAGTTC |
| CYFIP1-R | CTGGCTAGGGACTGGTGGATGG |
| NUDT1-F | TACTACAGCCTCAGCGAGTTCTCC |
| NUDT1-R | TCCCTCTTAGCCCCATCCTCAATG |
| DCPS-F | AAGCAGCGGTTGGCAATGGTAC |
| DCPS-R | TCCCCAGAGTCTCATTACCTTC |
| NSUN2-F | CGCTGCTATCTGCTCGTCCATC |
| NSUN2-R | CTGTGAGTCTAGGAATGCTGGATGC |
| CYFIP2-F | CCACCACCAACTGAAGGACATCATC |
| CYFIP2-R | TCTATGAGGAGGCAGAACAGGATGG |
| EIF4E3-F | GAGTGTGCCTCGAACCTGAAGAAG |
| EIF4E3-R | TGGTCGCCTCTCTCCTCTCATTAAAG |
| EIF3D-F | CAACAAGCAGGTCATCCGAGTCTAC |
| EIF3D-R | CCTCCTCTCCTCCTCATCCTCTTC |

2.3 Machine learning screens 7-methylguanosine (m7G) key genes between healthy individuals and patients with ischemic stroke (IS)

Based on previous studies on m7G, 34 m7G key regulatory genes were included in this study as study objects, including *DCP2*, *AGO2*, *CYFIP1*, *CYFIP2*, *DCPS*, *EIF3D*, *EIF4A1*, *EIF4E*, *EIF4E1B*, *EIF4E2*, *EIF4E3*, *EIF4G3*, *GEMIN5*, *IFIT5*, *LARP1*, *LSM1*, *METTL1*, *NCBP1*, *NCBP2*, *NCBP2L*, *NCBP3*, *NSUN2*, *NUDT1*, *NUDT10*, *NUDT11*, *NUDT16*, *NUDT16L1*, *NUDT3*, *NUDT4*, *NUDT4B*, *NUDT5*, *NUDT7*, *SNUPN*, and *WDR4* (Tomikawa, 2018; Chen et al., 2022). Differences in the expression patterns of these genes between patients and controls were detected using the Wilcoxon test, with a selection criterion of $p < 0.05$. Spearman correlation analysis was performed on these differentially expressed genes (DEGs), and their chromosomal locations were marked. This study utilized two widely used machine learning algorithms, random forest (RF) and support vector machine (SVM), to identify key regulators of m7G between patients with IS and controls by the “randomForest” package. The algorithm with the smaller residual was considered to be a more precise algorithm and was used. The R package “pROC” was used to calculate the area

under the curve (AUC) and evaluate the accuracy of the two algorithms.

2.4 Quantitative real-time polymerase chain reaction

Total RNA was extracted from the ischemic penumbra of rats and from the same site in the control group using TRIzol (Invitrogen, USA). After reverse transcription, real-time PCR was performed on genes with significant m7G differences. The primer sequences for these genes are listed in Table 1. Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) was used as the internal reference gene. The results are expressed as relative mRNA expression at cycle thresholds and normalized by parallel amplification of the endogenous control GAPDH. The relative mRNA expression level (target mRNA/GAPDH value) of the control group was set as 100%, and the mRNA values of the other groups were converted into fold changes after comparison with the control group.

2.5 Establishment and validation of clinical prediction models

The expression of the five m7G-related genes was packed by the “datadist” function of the “rms” package, and subsequently, the model was fitted using the “lrm” function. The “nomogram” function was used to build a suitable model and draw a nomogram by these risk genes. The total score of the nomogram was the sum of the corresponding scores assigned to each differential gene, and the score corresponded to the corresponding disease risk. The higher the score, the higher the risk of gene-induced IS development. Internal validation using the “caret” package and Bootstrap self-sampling method to derive the consistency index (C-Index). The calibration, clinical decision analysis, and receiver operating characteristic (ROC) curves were used to further evaluate the accuracy of the risk model.

2.6 Cluster analysis of patients with IS by m7G-regulated genes

Cluster analysis was used to distinguish different IS patient classifications based on the regulation of key m7G genes. The R package “ConsensusClusterPlus” was used to classify patients with IS into different subgroups according to experimentally validated m7G key regulatory genes. In this study, the PAM algorithm and spearman distance were used as parameters, and the sampling was repeated 1,000 times for a more stable classification. The number of clusters was determined using a cumulative distribution function. The “Rtsne” package was used

to display the distribution of samples for different clusters. The expression of m7G key regulatory genes was compared between the two clusters using the Kruskal–Wallis test.

2.7 Predicting the immune properties of m7G key regulatory genes

The single-sample gene set enrichment analysis (ssGSEA) algorithm was used to assess the immune infiltration of samples and genes by the “gsva” package. This study analyzed 23 immune cell types using ssGSEA. These included activated B cells, activated CD4 T cells, activated CD8 T cells, activated dendritic cells, CD56 bright natural killer (NK) cells, CD56 dim NK cells, eosinophils, gamma delta T cells, immature B cells, immature dendritic cells, myeloid-derived suppressor cells (MDSCs), macrophages, mast cells, monocytes, NK T cells, NK cells, neutrophils, plasmacytoid dendritic cells, regulatory T cells, T follicular helper cells, type 1 T helper cells, type 17 T helper cells, and type 2 T helper cells. The infiltrating immune cell abundance scores in two different patient clusters were compared using the Kruskal–Wallis test. A heatmap was drawn by the “pheatmap” package to show the correlation between five m7G key regulatory genes and these immune cells and to select a key gene that best represents the cluster analysis.

2.8 Enrichment analysis

After cluster analysis, the DEGs between the two clusters were screened by the “limma” package, and the screening conditions were as follows: $|\log_2(\text{fold change})| > 0.5$, adjustment p -value < 0.05 . Metascape (<https://metascape.org/gp/index.html>) is an excellent tool for pathway and biological function enrichment analysis. These genes were functionally enriched using Metascape, with output options, including Gene Ontology (GO) biological processes, canonical pathways, Kyoto Encyclopedia of Genes and Genomes pathway, and Reactome gene sets.

2.9 Construction of the mRNA–miRNA–transcription factor (TF) network

The STRING database (<https://cn.string-db.org/>) can be used to assess protein–protein interactions (PPIs). The DEGs between the two m7G clusters were inputted into the STRING database to construct a PPI network. After forming the PPI network, we performed cluster analysis on the PPI network using MCODE of Cytoscape and explored the cluster with the highest MCODE score as the key genes network. The possible binding miRNAs of

the key genes were predicted using the TargetScan (<https://www.targetscan.org/>) and miRTarBase databases (<https://www.mirbase.org/>). Predicted transcription factors (TF) may bind to key genes in the Enrichr database (<https://maayanlab.cloud/Enrichr/>). Finally, Cytoscape 3.7.2 was used to construct the mRNA–miRNA–TF network.

2.10 Statistical analyses

R version 4.0.2 was applied for all statistical analyses. Between-group comparisons were made using the independent samples t -test and Mann–Whitney U test. All analyses were based on two-tailed tests, and statistical significance was set at $p < 0.05$.

3 Results

3.1 Expression patterns and differences of m7G-regulated genes in IS

We explored the differential expression of 34 m7G-regulated genes in IS and found that 11 genes were significantly differentially expressed. Among these, *CYFIP1*, *EIF4E2*, and *EIF4E3* were significantly upregulated in IS, whereas *CYFIP2*, *DCPS*, *EIF3D*, *GEMIN5*, *NCBP2*, *NSUN2*, *NUDT1*, and *SNUPN* were significantly downregulated (Figures 1A, B). To explore whether these m7G-regulated genes played a key role in IS, we assessed the correlation between these genes (Figure 1C). In IS, *DCPS* and *NUDT1* showed a high positive correlation ($r = 0.67$), and *EIF3D* and *EIF4E3* showed a high negative correlation ($r = -0.60$). This suggested that m7G-regulated genes play an important role in IS. We further marked the location of these genes on the chromosomes (Figure 1D).

3.2 Machine learning and m7G key gene screening

The machine learning algorithm was used to further screen for m7G key regulatory genes. We compared two machine learning algorithms and found that the residual of RF was significantly smaller than that of SVM (Figures 2A, B). In the ROC curve, the RF algorithm ($AUC = 1$) also showed better accuracy than SVM (Figure 2C). Therefore, the RF algorithm was selected as the machine learning algorithm in this experiment. When the number of trees was 93, the machine learning error of the RF algorithm was the smallest (Figure 2D). Finally, eight genes with an importance score greater than 3 were selected: *CYFIP1*, *CYFIP2*, *DCPS*, *EIF3D*, *EIF4E3*, *NCBP2*, *NSUN2*, and *NUDT1* (Figure 2E).

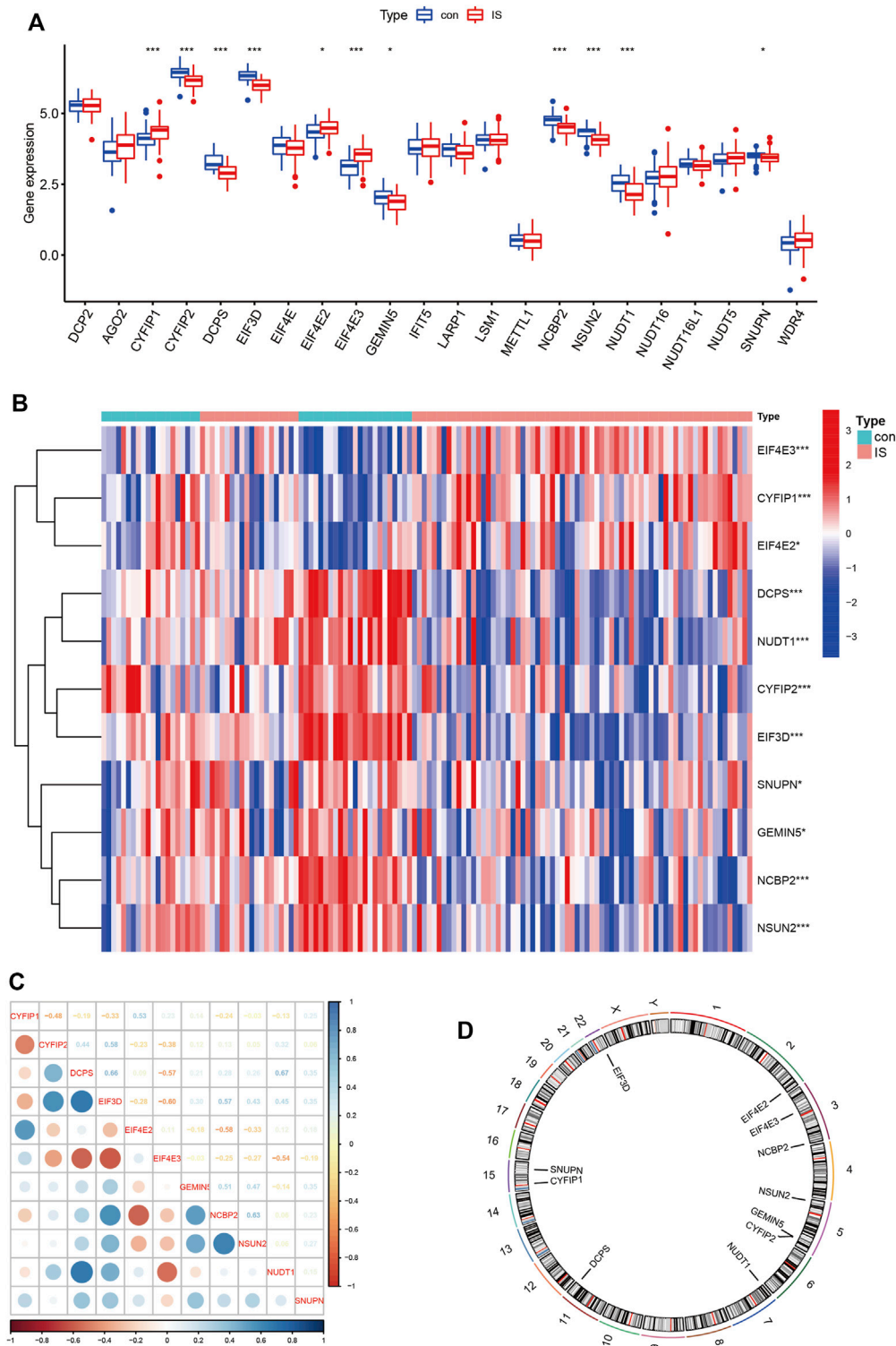
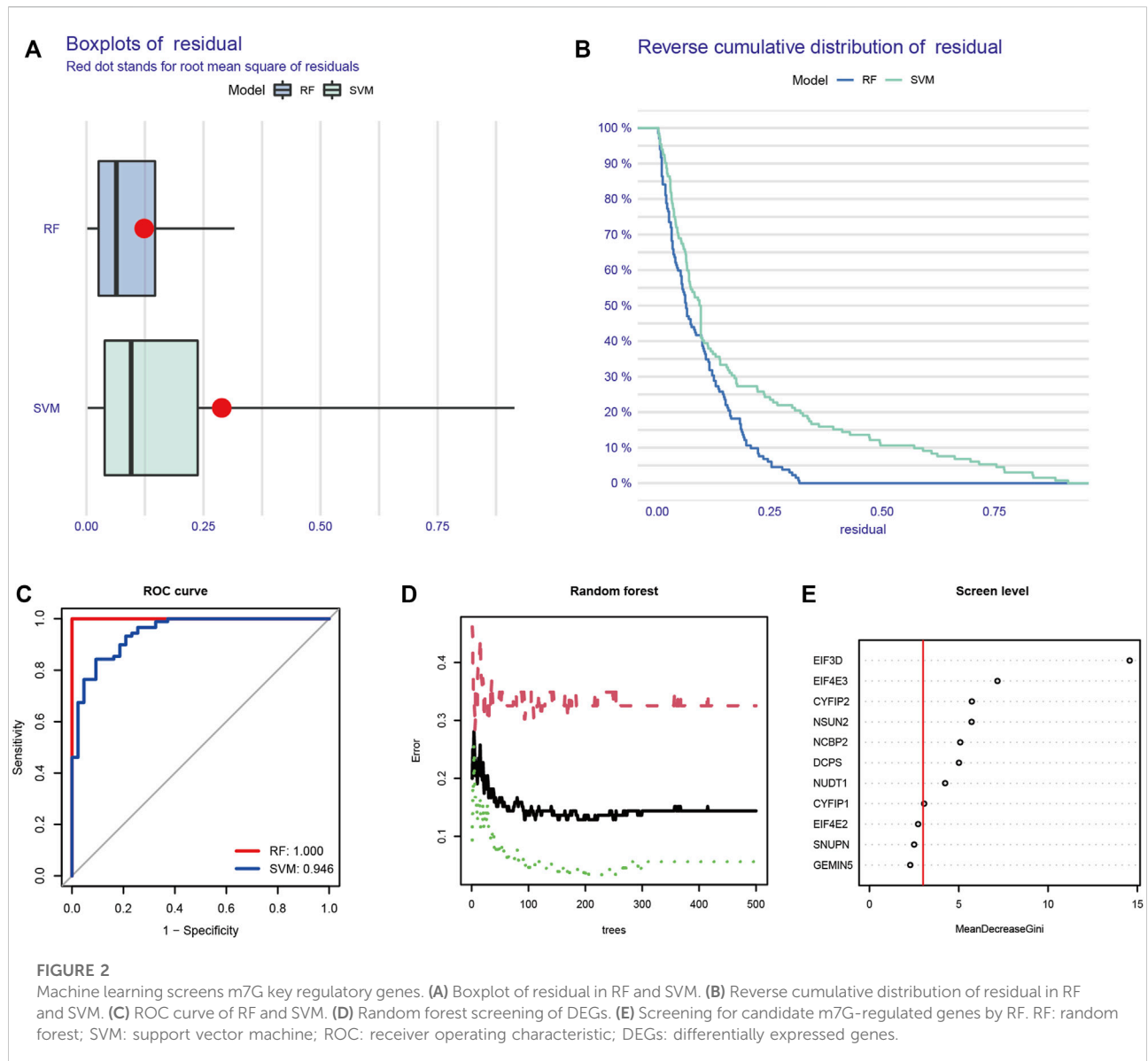


FIGURE 1 Expression patterns and differences of m7G-regulated genes in IS. **(A)** Boxplot of 34 m7G genes expression between control and IS. **(B)** Heatmap of 11 differentially expressed m7G genes between control and IS. Red represents high expression and blue represents low expression. **(C)** Correlations of m7G DEGs in IS. Blue represents positive correlation and red represents negative correlation. **(D)** Chromosomal positions of m7G DEGs. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. IS: ischemia stroke; con: control; DEGs: differentially expressed genes.



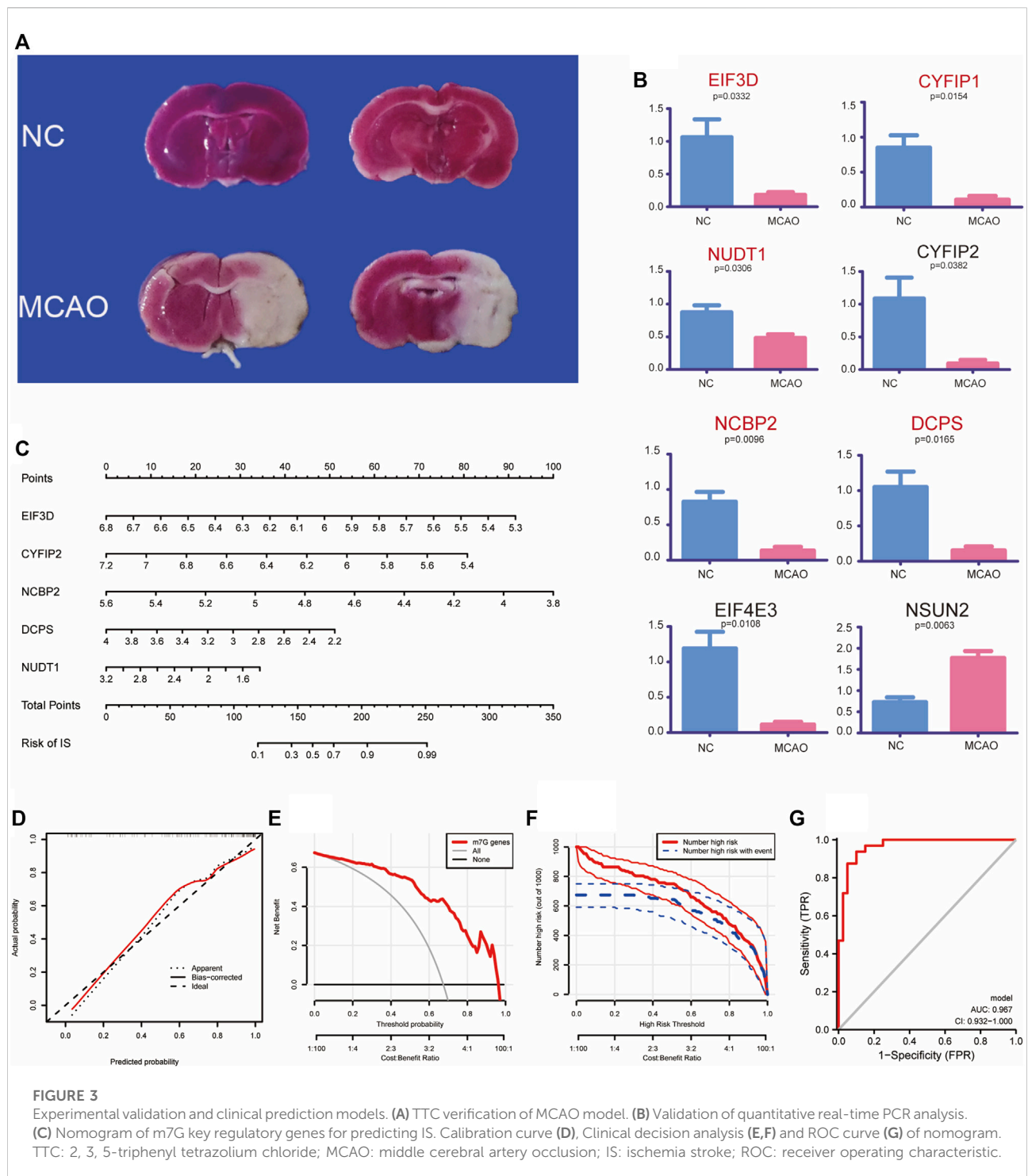
3.3 Expression profiles of m7G-regulated genes in the MCAO model

To explore the expression of m7G-regulated genes in IS, we constructed a MCAO model. We used rat brain tissue for TTC staining 3 days after modeling to verify the success of the modeling. A clear white infarct appeared in the left cerebral hemisphere of the model group, whereas the whole brain of the control group showed a red active state (Figure 3A). Eight screened m7G-regulated genes were verified using qPCR. The results showed that *EIF3D*, *CYFIP2*, *NCBP2*, *DCPS*, and *NUDT1* exhibited significant differences in the MCAO model, which was consistent with the differential analysis of the expression profile dataset (Figure 3B). These results

confirm that these m7G-regulated genes play a significant regulatory role in IS.

3.4 Establishment of a clinical prediction model

We established a clinical prediction model to evaluate the risk and correlation between five key m7G-regulated genes in IS. Our nomogram showed the risk of developing IS for each gene (Figure 3C). The internal validation of the model using Bootstrap self-sampling method with 1,000 samples yielded a model C-Index of 0.888. A calibration curve was used to further confirm the accuracy of the model, which showed that the



prediction model had a good accuracy (Figure 3D). The decision analysis curve also showed that m7G could predict the risk of disease more accurately (Figures 3E, F). The ROC curve (AUC = 0.967, 95% CI 0.932–1.000) further supported these results (Figure 3G). In conclusion, we used a clinical predictive model to accurately assess the risk of m7G-regulated genes in IS.

3.5 Cluster of patients with IS according to m7G key regulatory genes

Based on the five validated key regulatory genes of m7G, we performed cluster analysis on patients with IS. The tracking plot showed that it was prudent to divide the patients into two clusters

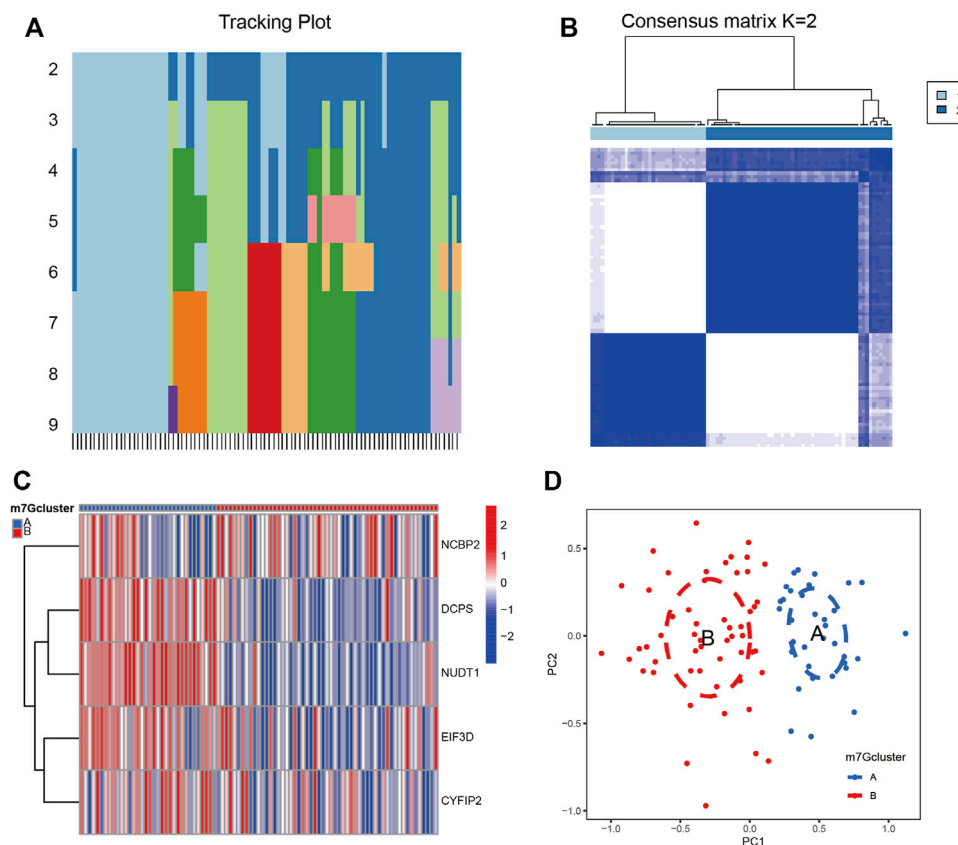


FIGURE 4

Cluster analysis of IS by m7G key regulatory genes. (A) Sample distribution for $k = 2-9$. (B) Consensus clustering matrix with $k = 2$. (C) Heatmap of m7G key regulatory genes between clusters. Red represents high expression and blue represents low expression. (D) PCA analysis between clusters. Group A and group B represent two clusters of IS patients divided according to the expression of m7G key regulatory genes. PCA: principal component analysis.

for accuracy (Figures 4A, B). We displayed the expression profiles of five m7G key regulatory genes according to these two clusters and found that their expression levels varied significantly in different clusters (Figure 4C). Principal component analysis (PCA) revealed that this clustering method could completely and accurately distinguish patients with IS (Figure 4D). Therefore, we accurately clustered patients with IS according to the expression patterns of m7G-regulated genes.

3.6 Immune infiltration signatures of m7G clusters

We used the ssGSEA algorithm to evaluate the level of immune cell infiltration between different clusters to explore the differences in their immune microenvironment characteristics. We found that the four types of immune cell infiltration were significantly different between the two clusters: immature dendritic cells, macrophages, NK T cells, and type1 T helper cells (Figure 5A). The correlation of

the five m7G-regulated genes experimentally identified with immune cell infiltration was also calculated (Figure 5B). Among them, the correlation of *NCBP2* was the most evident, with a maximum positive correlation coefficient of 0.58 and a maximum negative correlation coefficient of -0.69 . Therefore, *NCBP2* may play a critical role in immune cell infiltration. As shown in Figure 5C, among the cells with different *NCBP2* expression levels, there were more cell types with significant differences in immune infiltration, including activated B cells, activated CD4 T cells, activated CD8 T cells, activated dendritic cells, eosinophils, MDSC, macrophages, plasmacytoid dendritic cells, mast cells, NK cells, neutrophils, and type 2 T helper cells.

3.7 Enrichment among different m7G clusters

To explore the characteristics of the biological functions under different m7G gene expression patterns, we performed

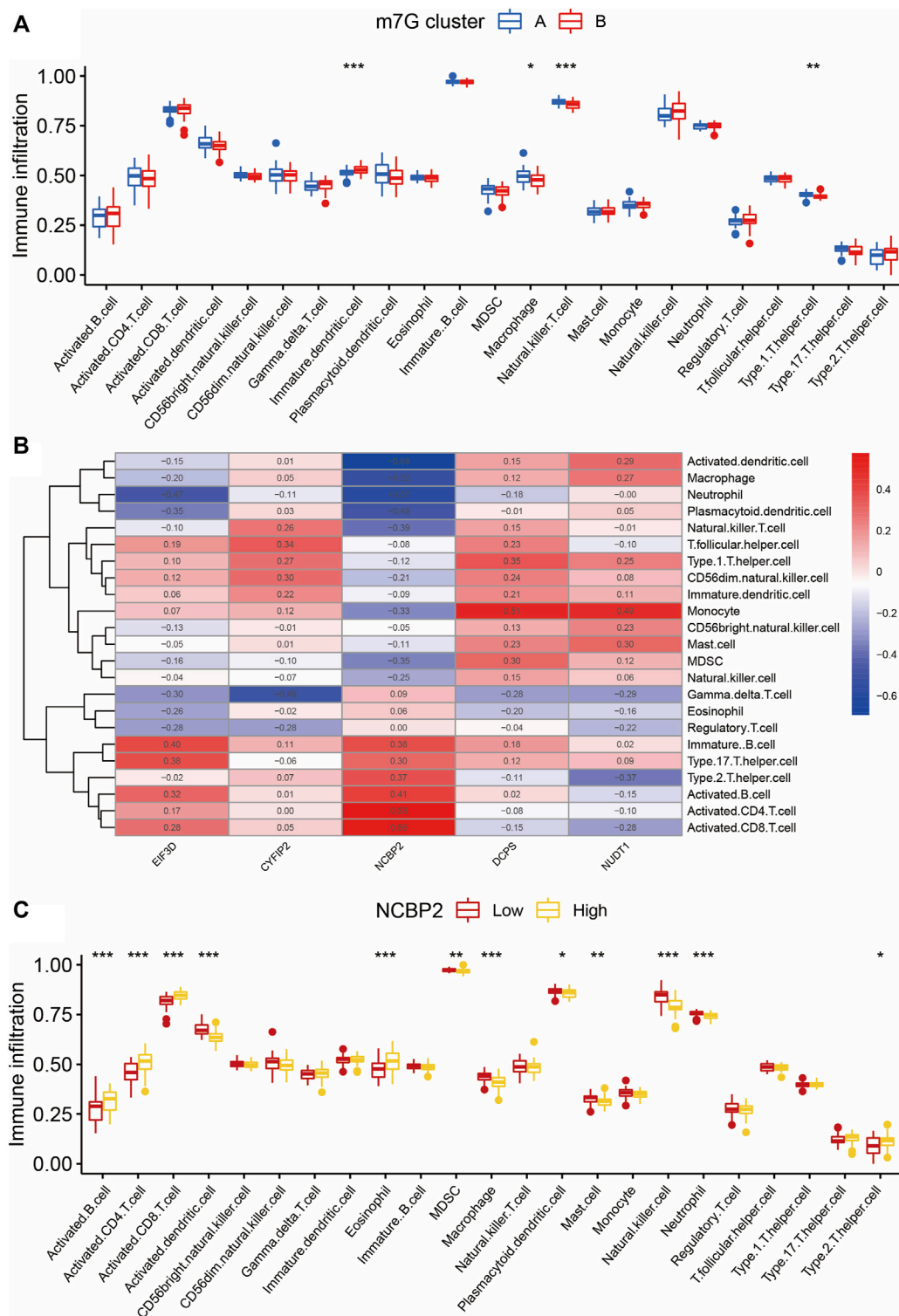


FIGURE 5 Immune infiltration analysis of two m7G clusters. (A) Differences in immune infiltration abundances between two m7G clusters. Group A and group B represent two clusters of IS patients divided according to the expression of m7G key regulatory genes. (B) Immune cell infiltration correlation heatmap of m7G key regulatory genes. Red represents positive correlation and blue represents negative correlation. (C) Immune infiltration analysis between clusters with different *NCBP2* expression levels. Group Low and group High represent cell clusters with low and high *NCBP2* expression, respectively. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. IS: ischemia stroke.

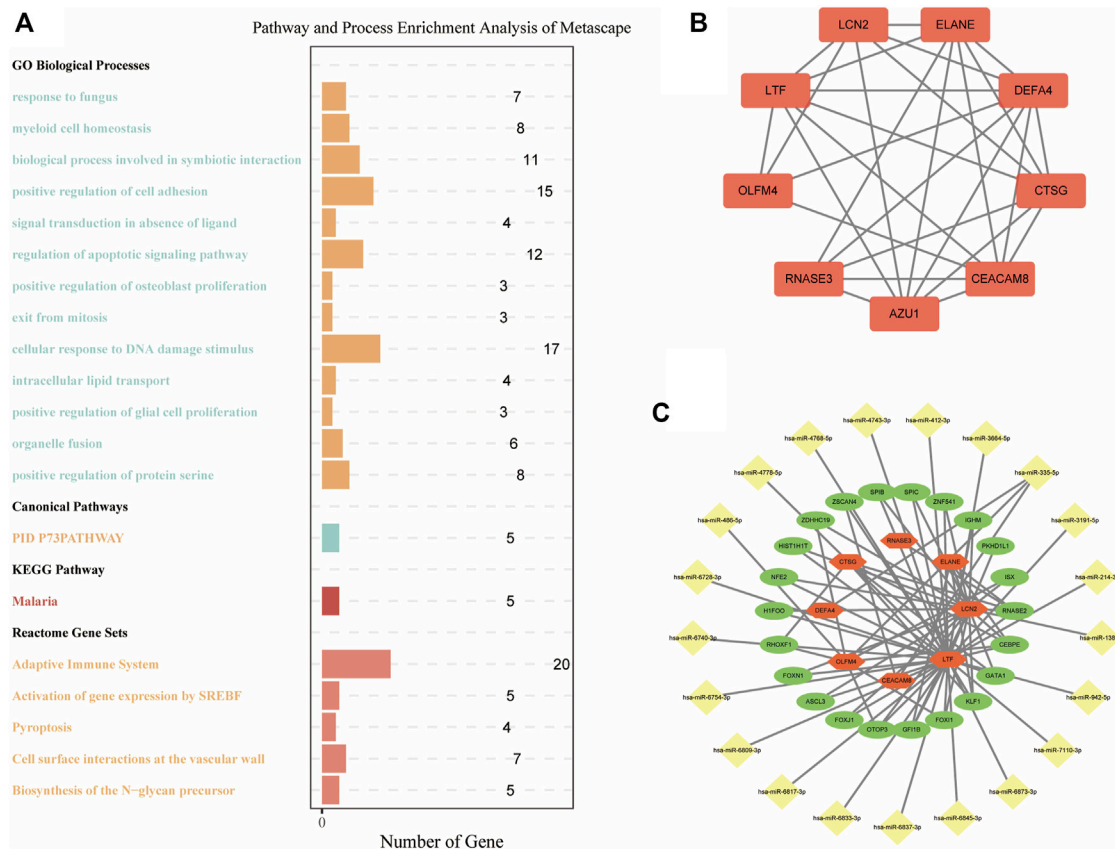


FIGURE 6

Functional enrichment and mRNA-miRNA-TF networks. (A) Pathway and process enrichment analysis of DEGs between m7G clusters. (B) PPI network of DEGs between m7G clusters. (C) mRNA-miRNA-TF networks of DEGs between m7G clusters. TF: transcription factor; PPI: protein-protein interactions; DEGs: differentially expressed genes.

biological functions and pathway enrichment analysis. Specifically, we screened 282 DEGs between the two m7G clusters (Supplementary Table S1). Metascape was used for the enrichment analysis (Figure 6A). The results showed that in the “GO Biological Processes” analysis, the “regulation of apoptotic signaling pathway” and “cellular response to DNA damage stimulus,” the mechanisms closely related to IS pathogenesis, were enriched in 12 and 17 genes, respectively. In the “Reactome Gene Set” analysis, the genes enriched in the two IS-related pathways of “adaptive immune system” and “pyroptosis” were 20 and 4 genes, respectively. This indicates that m7G-regulated genes are closely related to IS in terms of biological functions.

3.8 Construction of the mRNA-miRNA-TF network

DEGs between the two m7G clusters were used to build a PPI network to explore the interaction relationship between genes. There

were 254 nodes and 274 edges in this PPI network (Supplementary Figure S1). After further identification of key modules and hub genes using MCODE, a PPI network with 9 nodes and 26 edges was identified, which included *LTF*, *LCN2*, *ELANE*, *RNASE3*, *CTSG*, *DEFA4*, *OLFM4*, and *CEACAM8* (Figure 6B). We combined “TargetsCan” and “miRTarBase” databases to successfully predict 21 miRNAs that may bind to key regulatory genes of m7G. Next, we showed the mRNA-miRNA-TF networks (Figure 6C). Among them, *LTF* had the largest number of nodes and edges and may bind to 18 miRNAs and 19 TFs. This indicates that *LTF* plays an important regulatory role in IS.

4 Discussion

In this study, we discovered the epigenetic and immune microenvironmental regulatory mechanisms of m7G in IS. First, we screened out the differentially expressed m7G regulatory genes in IS. Second, we identified that *EIF3D*, *CYFIP2*, *NCBP2*, *DCPS*, and *NUDT1* were five key m7G-

regulated genes differentially expressed in IS according to the RF algorithm and qPCR of the MCAO model. The risk impact of these genes on developing IS was assessed separately, and patients with IS were divided into two clusters based on these genes. Finally, immune infiltration between the two clusters and the functional enrichment and regulatory network of differential genes were also revealed.

IS is a complex disease that involves multiple molecular mechanisms and methylation modifications. Chokkalla et al. (Chokkalla et al., 2019) found that regulation of m6A methylation is involved in IS development and can be considered an important marker of IS. Zhang et al. (Zhang et al., 2020) found that YTH domain-containing 1 acts as an m6A reader and alleviates IS by promoting the activation of the AKT signaling pathway. However, studies on m7G and IS are limited. Therefore, our study provides evidence for epigenetic studies on methylation and IS.

We screened key m7G-regulated genes in IS and further verified this using the rat MCAO model by qPCR, which greatly improved the accuracy of the screening. These genes included *EIF3D*, *CYFIP2*, *NCBP2*, *DCPS*, and *NUDT1*. *EIF3D* and *NUDT1* have been shown to play important regulatory roles in tumor and immune infiltration (Huang et al., 2019; Huang et al., 2022). *CYFIP2* has been shown to play vital regulatory role in the central nervous system (Schaks et al., 2020). *NCBP2* and *DCPS* are believed to be involved in neurogenesis, which may inextricably be associated with IS (Singh et al., 2020; Salamon et al., 2022). However, the specific mode of regulation between them and the IS has not yet been studied. Therefore, our study successfully confirmed their close correlation with IS using an animal model.

Machine learning and clinical predictions are excellent tools for bioinformatics analysis, enabling accurate assessment of disease regulatory mechanisms and risks. The algorithm we used, RF, has been used for long-term outcome prediction of mortality and morbidity in patients with stroke. Heo et al. (Heo et al., 2019) found that the RF algorithm can also predict the long-term prognosis of IS. Our study not only selected the key genes with the RF algorithm but also proved that RF was more suitable for our study, which provides evidence for the precise selection of the appropriate machine learning. Nomograms have been widely used in clinical prediction models of stroke. Yuan et al. (Yuan et al., 2020) used a nomogram to accurately predict the risk of stroke using multiple risk factors, including hypertension, diabetes, and smoking. Our study further refines the risk factors for genes, providing a more precise theoretical basis for the prevention and treatment of stroke through molecular mechanisms.

Our study makes the first attempt to cluster patients with IS into two defined clusters based on m7G key gene expression profiles, as well as presents novel methodologies for identifying different types of patients with IS and their precise treatment. In addition, we analyzed the differences in immune infiltration between the two clusters. Li et al. (Li et al., 2022) clustered patients with liver cancer by m7G-regulated gene expression patterns. In our study, the immune infiltrating cells with significant differences between the

different clusters were immature dendritic cells, macrophages, NK T cells, and type 1 T helper cells. Therefore, we confirmed that m7G-regulated genes have profound effects on immune cell infiltration and play different immune regulatory roles in various diseases. We also found that these differences in immune cell infiltration were closely related to *NCBP2*.

Differential m7G gene expression profiles between clusters were screened and functionally enriched and mRNA-miRNA-TF networks were established. Several reports have suggested that *LCN2* can mediate the phagocytosis of astrocytes to trigger demyelination, which exacerbates IS (Wan et al., 2022). The critical role of the miR-214-*LTF*-*FOXJ1* axis was also observed in our study. Although *LTF* is believed to mediate neuronal ferroptosis in hemorrhagic stroke, it has rarely been reported in IS (Zhao et al., 2018). MIR-214 attenuates neuronal apoptosis and ferroptosis in IS, and *FOXJ1* is believed to induce neurogenesis (Devaraju et al., 2013; Lu et al., 2020). Therefore, we hypothesized that the miR-214-*LTF*-*FOXJ1* axis may play an important regulatory role in IS, thus becoming an important molecular target for the prevention and treatment of IS. However, the specific role of this axis has not yet been verified, which may become the focus of our next study.

This study has its own limitations. First, only vivo experiments but no vitro cell experiments were performed. This may be improved in subsequent studies. Second, although we innovatively discovered the miR-214-*LTF*-*FOXJ1* axis, this could not be verified by basic experiments. Third, although we were able to establish detailed predictions on the mediation network of m7G, we did not further explore therapeutic drugs based on this, which is insufficient for clinical guidance. In addition, we still need to obtain more clinical data from patients as an analysis basis to augment the accuracy of assessment and prediction.

Data availability statement

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

Ethics statement

The animal study was reviewed and approved by Xi'an Jiaotong University.

Author contributions

YT and BY: Manuscript preparation, data analysis, and the research conception. SG and JL: Manuscript revision. YZ, SY, LF, and BL: Data analysis.

Funding

This work was supported by the Construction and application of big data self-evolution remote diagnosis and treatment system in ICU for severe neurological diseases (grant number NO.2022ZDLSF04-01) and the National Natural Science Foundation of China (Grant No. 81903268).

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.

References

- Chen, X., Sun, Y. Z., Liu, H., Zhang, L., Li, J. Q., and Meng, J. (2019). RNA methylation and diseases: Experimental results, databases, web servers and computational models. *Brief. Bioinform* 20 (3), 896–917. doi:10.1093/bib/bbx142
- Chen, Z., Zhang, Z., Ding, W., Zhang, J. H., Tan, Z. L., Mei, Y. R., et al. (2022). Expression and potential biomarkers of regulators for m7G RNA modification in gliomas. *Front. Neurol.* 13, 886246. doi:10.3389/fneur.2022.886246
- Chokkalla, A. K., Mehta, S. L., Kim, T., Chelluboina, B., Kim, J., and Vemuganti, R. (2019). Transient focal ischemia significantly alters the m(6)A epitranscriptomic tagging of RNAs in the brain. *Stroke* 50 (10), 2912–2921. doi:10.1161/strokeaha.119.026433
- Dai, Z., Liu, H., Liao, J., Huang, C., Ren, X., Zhu, W., et al. (2021). N(7)-Methylguanosine tRNA modification enhances oncogenic mRNA translation and promotes intrahepatic cholangiocarcinoma progression. *Mol. Cell* 81 (16), 3339–3355.e8. doi:10.1016/j.molcel.2021.07.003
- Devaraju, K., Barnabé-Heider, F., Kokaia, Z., and Lindvall, O. (2013). FoxJ1-expressing cells contribute to neurogenesis in forebrain of adult rats: Evidence from *in vivo* electroporation combined with piggyBac transposon. *Exp. Cell Res.* 319 (18), 2790–2800. doi:10.1016/j.yexcr.2013.08.028
- Feske, S. K. (2021). Ischemic stroke. *Am. J. Med.* 134 (12), 1457–1464. doi:10.1016/j.amjmed.2021.07.027
- Heo, J., Yoon, J. G., Park, H., Kim, Y. D., Nam, H. S., and Heo, J. H. (2019). Machine learning-based model for prediction of outcomes in acute stroke. *Stroke* 50 (5), 1263–1265. doi:10.1161/strokeaha.118.024293
- Huang, B., Lyu, Z., Qian, Q., Chen, Y., Zhang, J., Li, B., et al. (2022). NUDT1 promotes the accumulation and longevity of CD103(+) TRM cells in primary biliary cholangitis. *J. Hepatol.* 77, 1311–1324. doi:10.1016/j.jhep.2022.06.014
- Huang, H., Gao, Y., Liu, A., Yang, X., Huang, F., Xu, L., et al. (2019). EIF3D promotes sunitinib resistance of renal cell carcinoma by interacting with GRP78 and inhibiting its degradation. *EBioMedicine* 49, 189–201. doi:10.1016/j.ebiom.2019.10.030
- Li, X. Y., Zhao, Z. J., Wang, J. B., Shao, Y. H., Hui, L., You, J. X., et al. (2022). m7G methylation-related genes as biomarkers for predicting overall survival outcomes for hepatocellular carcinoma. *Front. Bioeng. Biotechnol.* 10, 849756. doi:10.3389/fbioe.2022.849756
- Longa, E. Z., Weinstein, P. R., Carlson, S., and Cummins, R. (1989). Reversible middle cerebral artery occlusion without craniectomy in rats. *Stroke* 20 (1), 84–91. doi:10.1161/01.str.20.1.84
- Lu, J., Xu, F., and Lu, H. (2020). LncRNA PVT1 regulates ferroptosis through miR-214-mediated TFR1 and p53. *Life Sci.* 260, 118305. doi:10.1016/j.lfs.2020.118305
- Maida, C. D., Norrito, R. L., Daidone, M., Tuttolomondo, A., and Pinto, A. (2020). Neuroinflammatory mechanisms in ischemic stroke: Focus on cardioembolic stroke, background, and therapeutic approaches. *Int. J. Mol. Sci.* 21 (18), 6454. doi:10.3390/ijms21186454
- Matsuzono, K., Mashiko, T., Ozawa, T., Miura, K., Suzuki, M., Furuya, K., et al. (2021). Characteristics of aged ischemic stroke patients indicative of cardioembolism. *J. Thromb. Thrombolysis* 51 (2), 522–529. doi:10.1007/s11239-020-02198-1
- Salamon, I., Palsule, G., Luo, X., Roque, A., Tucai, S., Khosla, I., et al. (2022). mRNA-decapping associated DcpS enzyme controls critical steps of neuronal development. *Cereb. Cortex* 32 (7), 1494–1507. doi:10.1093/cercor/bhab302
- Schaks, M., Reinke, M., Witke, W., and Rottner, K. (2020). Molecular dissection of neurodevelopmental disorder-causing mutations in CYFIP2. *Cells* 9 (6), 1355. doi:10.3390/cells9061355
- Singh, M. D., Jensen, M., Lasser, M., Huber, E., Yusuff, T., Pizzo, L., et al. (2020). NCBP2 modulates neurodevelopmental defects of the 3q29 deletion in *Drosophila* and *Xenopus laevis* models. *PLoS Genet.* 16 (2), e1008590. doi:10.1371/journal.pgen.1008590
- Tomikawa, C. (2018). 7-Methylguanosine modifications in transfer RNA (tRNA). *Int. J. Mol. Sci.* 19 (12), 4080. doi:10.3390/ijms19124080
- Wan, T., Zhu, W., Zhao, Y., Zhang, X., Ye, R., Zuo, M., et al. (2022). Astrocytic phagocytosis contributes to demyelination after focal cortical ischemia in mice. *Nat. Commun.* 13 (1), 1134. doi:10.1038/s41467-022-28777-9
- Yang, B., Wang, J. Q., Tan, Y., Yuan, R., Chen, Z. S., and Zou, C. (2021). RNA methylation and cancer treatment. *Pharmacol. Res.* 174, 105937. doi:10.1016/j.phrs.2021.105937
- Yuan, K., Chen, J., Xu, P., Zhang, X., Gong, X., Wu, M., et al. (2020). A nomogram for predicting stroke recurrence among young adults. *Stroke* 51 (6), 1865–1867. doi:10.1161/strokeaha.120.029740
- Zhang, Z., Wang, Q., Zhao, X., Shao, L., Liu, G., Zheng, X., et al. (2020). YTHDC1 mitigates ischemic stroke by promoting Akt phosphorylation through destabilizing PTEN mRNA. *Cell Death Dis.* 11 (11), 977. doi:10.1038/s41419-020-03186-2
- Zhao, X., Ting, S. M., Sun, G., Roy-O'Reilly, M., Mobley, A. S., Bautista Garrido, J., et al. (2018). Beneficial role of neutrophils through function of lactoferrin after intracerebral hemorrhage. *Stroke* 49 (5), 1241–1247. doi:10.1161/strokeaha.117.020544
- Zhao, Y., Kong, L., Pei, Z., Li, F., Li, C., Sun, X., et al. (2021). m7G methyltransferase METTL1 promotes post-ischemic angiogenesis via promoting VEGFA mRNA translation. *Front. Cell Dev. Biol.* 9, 642080. doi:10.3389/fcell.2021.642080

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.2022.1036345/full#supplementary-material>



OPEN ACCESS

EDITED BY

Clévia Rosset,
Clinical Hospital of Porto Alegre, Brazil

REVIEWED BY

Yizhuo Wang,
University of Texas MD Anderson
Cancer Center, United States
Xiaoxiao He,
Northeast Normal University, China

*CORRESPONDENCE

Linlin Zeng,
✉ zenglinlin@jlu.edu.cn

SPECIALTY SECTION

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

RECEIVED 13 June 2022

ACCEPTED 07 December 2022

PUBLISHED 10 January 2023

CITATION

Ma X, Feng Y, Quan X, Geng B, Li G, Fu X
and Zeng L (2023), Multi-omics analysis
revealed the role of CCT2 in the
induction of autophagy in
Alzheimer's disease.
Front. Genet. 13:967730.
doi: 10.3389/fgene.2022.967730

COPYRIGHT

© 2023 Ma, Feng, Quan, Geng, Li, Fu and
Zeng. This is an open-access article
distributed under the terms of the
[Creative Commons Attribution License](#)
(CC BY). The use, distribution or
reproduction in other forums is
permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original
publication in this journal is cited, in
accordance with accepted academic
practice. No use, distribution or
reproduction is permitted which does
not comply with these terms.

Multi-omics analysis revealed the role of CCT2 in the induction of autophagy in Alzheimer's disease

Xueting Ma¹, Yuxin Feng¹, Xiangyu Quan¹, Bingyu Geng¹,
Guodong Li², Xueqi Fu¹ and Linlin Zeng^{1*}

¹Edmond H. Fischer Signal Transduction laboratory, School of Life Sciences, Jilin University, Changchun, China, ²Department of General Surgery, The Second Hospital of Jilin University, Changchun, China

Chaperonin containing TCP1 subunit 2 (CCT2) is essential in various neurodegenerative diseases, albeit its role in the pathogenesis of Alzheimer's disease (AD) remains elusive. This study aimed to evaluate the role of *CCT2* in Alzheimer's disease. First, bioinformatics database analysis revealed that *CCT2* was significantly downregulated in patients with Alzheimer's disease and associated with autophagic clearance of β -amyloid. The 789 differentially expressed genes overlapped in AD-group and *CCT2*-low/high group, and the *CCT2*-high-associated genes screened by Pearson coefficients were enriched in protein folding, autophagy, and messenger RNA stability regulation pathways. These results suggest that *CCT2* is significantly and positively associated with multiple pathways linked to autophagy and negatively associated with neuronal death. The logistic prediction model with 13 key genes, such as *CCT2*, screened in this study better predicts Alzheimer's disease occurrence (AUC = 0.9671) and is a favorable candidate for predicting potential biological targets of Alzheimer's disease. Additionally, this study predicts reciprocal micro RNAs and small molecule drugs for hub genes. Our findings suggest that low *CCT2* expression may be responsible for the autophagy suppression in Alzheimer's disease, providing an accurate explanation for its pathogenesis and new targets and small molecule inhibitors for its treatment.

KEYWORDS

alzheimer's disease, autophagy, CCT2, microRNA, logistic model

Introduction

Alzheimer's disease (AD) is a neurodegenerative disease responsible for 60–80% of dementia cases, which is characterized by memory loss and reduced cognitive function (Liu et al., 2020). This report indicates that neuronal fibrous tangles caused by Tau hyperphosphorylation in neurons, sedimentation of amyloid beta (A β) plaques (Ma et al., 2022), apoptosis of numerous neurons, and loss of neural synapses all contribute to AD. Drugs approved by the FDA for AD are designed to improve the quality of life of patients with the disease albeit may not play an effective therapeutic role in the treatment of AD

(2020 AD facts and figures, 2020). AD-related therapeutic drugs based on A β starch spot and Tau protein have not made significant progress (Kopeikina et al., 2011); thus, the development of the specific pathogenesis of AD requires further research and exploration. It has been reported that the *chaperonin containing TCP1 subunit 2* (*CCT2*) is poorly expressed in AD; however, the relationship between *CCT2* and AD remains elusive (Yuan et al., 2019), implying that there is some link between *CCT2* gene expression and the occurrence of AD.

Aggrephagy, a process in which autophagy selectively degrades protein aggregates, is important for removing intracellular toxic protein aggregates and is a key target for the treatment of aggregate-related diseases such as neurodegenerative diseases. Several studies have reported that autophagy deficiency occurs in the early stages of AD (Vaillant-Beuchot L et al., 2021; Roca-Agujetas V et al., 2021). Autophagy is important in the production and metabolism of A β , and its dysfunction may contribute to the progression of AD (Li et al., 2017). Traditional ubiquitin-binding receptors (*P62*, *NBR1*, and *TAX1BP1*) can mediate aggrephagy and other types of ubiquitin-related selective autophagy (Zellner S et al., 2021). The novel ubiquitin-binding receptor, *CCT2*, promotes autophagic clearance of various toxic protein aggregates associated with neurodegenerative diseases (Zhang and Klionsky., 2022). Similar to the conventional ubiquitin-binding receptors, *CCT2* binds to LC3 and protein aggregates. *CCT2* binds protein aggregates in a ubiquitin-independent manner through its apical domain, laying the groundwork for *CCT2*-specific aggregate recognition. Research has indicated that conventional autophagy receptors degrade liquid aggregates whereas *CCT2* degrades solid aggregates (Ma et al., 2022). Consequently, *CCT2* is more likely than autophagy receptors to function and become an AD drug target in pathological states. *CCT2* mediates aggrephagy as a monomer, exposing the VLIR domain of the binding site to LC3. The presence of aggregates inhibits the formation of the Chaperonin complex, thus, releasing more *CCT2* monomers to promote aggregate clearance (Khaminets et al., 2016; Johansen and Lamark, 2020; Gatica et al., 2018). Full-length tau protein has been reported to preferentially be degraded by macrophage whereas caspase-cleaved tau, tau Δ C, which is more likely than natural proteins to aggregate and cause neurotoxicity, is preferentially degraded by autophagy and can turnover faster than the full-length tau. Thus, the autophagy degradation pathway is important in inhibiting the formation of pathological manifestations of AD and has the potential to be a novel target for its treatment (Zare-Shahabadi et al., 2015).

Therefore, this study aimed to investigate the changes in the expression level of *CCT2* in patients with AD and its possible pathway involved in autophagy and predict the possible micro RNA targets. Our study may help researchers investigate how *CCT2* affects AD via autophagy, contributing to the understanding of disease causes, mechanisms, and treatments.

Materials and methods

Data acquisition

All the datasets used in this study were obtained from the Gene Expression Omnibus database (<https://www.ncbi.nlm.nih.gov/geo/>) (Barrett et al., 2013). The AD transcriptome datasets screened from the database included brain tissue sequencing samples, GSE33000, GSE44768, GSE44770, and GSE44771, based on the GPL 4372 platform, peripheral blood samples of patients with AD, GSE140829, based on GPL5988 platform, and serum microRNA (miRNA) sequencing samples, GSE120584, based on the GPL21263 platform. GSE33000, which included 310 patients with AD and 157 controls, was used to explore the potential role of *CCT2* in AD. GSE44768, GSE44770, and GSE44771 were obtained from the cerebellum, frontal cortex, and visual cortex, respectively, and included 129 patients with AD and 101 controls. GSE140829 included 204 patients with AD and 249 controls to validate the model and explore *CCT2* expression in different tissues. GSE120584 included 1,021 patients with AD and 288 controls and was used to probe the possible messenger RNA (mRNA)-miRNA interaction networks. Component differences were observed using principal component analysis (PCA) plots drawn by the FactoMineR and factoextra packages. The data in GSE33000 was normalized using the normalizeBetweenArrays function in the Limma package (Ritchie et al., 2015), and the first group was retained for duplicated genes in the sequencing data.

Screening of differential genes (DEGs) and associated genes

LmFit, eBayes from the limma package, and the topTable function were used to identify differentially expressed genes DEGs between AD-con and *CCT2*-low/high expression groups. According to the false discovery rate (FDR), $p \leq 0.05$ was statistically significant, and log2fold change (FC) was used to comprehensively analyze the upregulated and downregulated genes.

For the AD-con group, we selected the first 30% genes with larger $|\log FC|$ under the $p \leq 0.05$ condition as the DEGs. Further, we divided all patients with AD into high- and low-expression groups based on the median of *CCT2*, and under the $p \leq 0.05$ condition, $|\log FC|$ the larger top 10% genes were selected as DEGs in the *CCT2*-low/high group. DEGs were intersected between AD-con and *CCT2*-low/high groups for further analysis.

Cor function was used for the raw data, and Pearson correlation analysis was performed between *CCT2* and other genes. If p -value was ≤ 0.05 and the gene was positively associated with *CCT2*, it was selected as the related gene.

Functional enrichment analysis

CCT2 and 460 genes with the strongest positive correlation with *CCT2* were uploaded to the online Fdatabase—Database for Annotation, Visualization, and Integrated Discovery, 2021 (Sherman et al., 2022; Huang et al., 2009)—for analysis. The official gene symbol was selected as the identifier, and the species was *Homo sapiens*. This was followed by Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis. The top eight pathway are displayed in ascending order of *p*-value ($p \leq 0.05$).

Gene set enrichment analysis (GSEA)

The differential expression analysis results of the limma package were analyzed by GSEA using the gseKEGG and gseGO functions of the clusterProfiler package in R (Wu et al., 2021), and biological process (BP) GO terms and KEGG pathways that may be related to AD and *CCT2* expression were explored. $p \leq 0.05$ and $|\text{NES}| > 1$ indicated significant differences.

Gene set variation analysis (GSVA)

Gene sets related to autophagy and protein folding were obtained from the GSEA website (<http://www.gsea-msigdb.org/gsea/index.jsp>) (Subramanian et al., 2005; Mootha et al., 2003). Before standardization, the GSVA package in R was used to calculate the functional enrichment scores of all AD groups in the GSE33000 dataset, and the parameters were set as default (Hänzelmann et al., 2013). Results were visualized by drawing heatmaps using the pheatmap package in R, and Pearson correlation analysis was used to determine the correlation between *CCT2* and autophagy and protein folding processes. Further, the top ten genes with the strongest positive and negative correlation with *CCT2* were drawn to exhibit their correlation with *CCT2* using data from the HADb database (<http://www.autophagy.lu/index.Autophagy-related> gene sets of html) (Moussay et al., 2011) and circo package (Krzywinski et al., 2009). Relevant gene sets from different stages of autophagy were selected for GSVA analysis; Pearson correlation analysis was used to calculate its correlation coefficient; the corrgram package was used to construct matrix plots.

Construction of the protein-protein interaction (PPI) network and identification of the hub genes

The DEGs from the AD-con groups intersected with the *CCT2* low-high groups, and 295 upregulated and 494 downregulated genes were removed as co-DEGs and uploaded to the online database

(STRING version 11.0, <https://cn.string-db.org/>) (Szkarczyk et al., 2021) to predict the PPI network, with the default parameters. The PPI interaction network was further drawn using Cytoscape, and 36 hub genes associated with *CCT2* were removed using the MCODE plugin.

Logistic model construction and receiver operating characteristic (ROC) curve analysis

The least absolute shrinkage and selection operator (LASSO) is a compression estimation method that has a strong factor screening ability (Tibshirani, 1997; Zou et al., 2019). The hub genes were intersected using Pearson's analysis results ($|r| \geq 0.65$, $p \leq 0.05$) to obtain 26 genes, and the expression profiles of these genes were used to construct the LASSO model, with 13 genes whose regression coefficient was not zero. These genes were used to construct a logistic regression model using the glmnet package. This model had the following formula: $\text{index} = \text{EXGene1} \times \text{Coef1} + \text{EXGene2} \times \text{Coef2} + \text{EXGene3} \times \text{Coef3} + \dots$ (Coef was the regression coefficient, derived from the logistic regression (Domínguez-Almendros et al., 2011); EXGene was the gene expression level).

Further, data from the GSE33000 dataset were randomly assigned to the test set (30%) and validated with those of the GSE44768, GSE44770, GSE44771, and GSE140829 datasets, and the ROC curve was drawn using the pROC package.

CCT2 expression and single-cell correlation analysis in different brain tissues

Using the online database, AlzDate (<http://www.alzdata.org/>) (Xu et al., 2018; Zhang et al., 2019) and the Single Cell Expression tool, *CCT2* expression in single cells was obtained. Using the Differential Expression tool, the differential expression of *CCT2* in multiple databases was obtained.

MicroRNA-mRNA interaction network analysis

MicroRNA is a type of single-stranded RNA molecule that is encoded by endogenous genes and binds to mRNA inside cells to inhibit protein translation. Exploring the interaction between miRNA and its target genes can provide a reference for investigating the disease causes and therapeutic methods. Databases for predicting gene-miRNA interactions include MiRDB, miRWalk, RNA22, and RNAInter (Chen and Wang, 2020; Dweep et al., 2011; Miranda et al., 2006; Lin et al., 2020). They were used to predict the miRNA interactions with hub genes, and the results were cross-checked to improve prediction

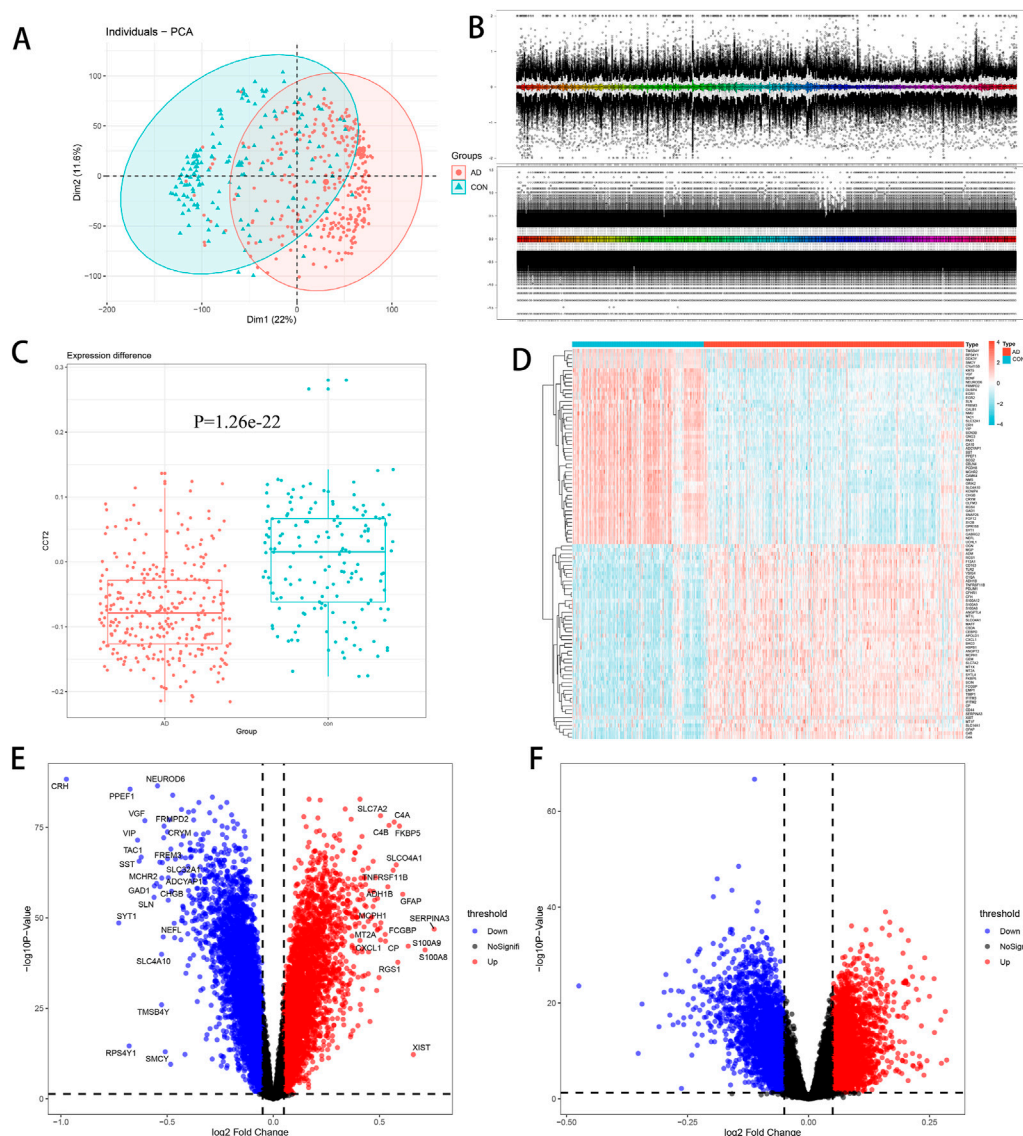


FIGURE 1

Differential expression analysis. (A) Principal component analysis (PCA) plot demonstrating differences among groups, with Alzheimer's disease (AD) in red and con in green. (B) Boxchart before and after standardization—the upper figure is before standardization; the following figure is after standardization. (C) Chaperonin containing TCP1 subunit 2 (CCT2) was downregulated in AD ($p = 1.26e-22$, $\log FC = -0.07632$). (D) Heatmap of the top 50 upregulated and downregulated genes between AD and control. (E) Volcano plot of AD-con, with upregulated genes in red and downregulated genes in blue. (F) Volcano plot of CCT2-low/high, with upregulated genes in red and downregulated genes in blue.

accuracy. Simultaneously, serum miRNA sequencing samples (GSE120584) were analyzed for differential expression using the limma package, and miRNAs with $p \leq 0.05$ were considered differentially expressed. Additionally, the differentially expressed miRNAs that interacted with hub genes were visualized using the Cytoscape software.

Related drug prediction

Drug development has always prioritized research on drugs for AD. Numerous effective drugs are ineffective in AD treatment as

they cannot cross the blood-brain barrier (BBB) whereas small-molecule drugs have natural advantages in crossing the BBB.

CCT2 has been reported to be used as a target of small-molecule drugs in the treatment of neurodegenerative diseases. Consequently, the prediction of CCT2-related DEGs serves as a reference for AD therapy. The Drug Signatures database (DSigDB) on the Enrichr website was used in this study to identify relevant targeted drugs for DEGs (Chen EY et al., 2013; Kuleshov et al., 2016; Xie Z et al., 2021). The results were reviewed and displayed (Kuleshov et al., 2016).

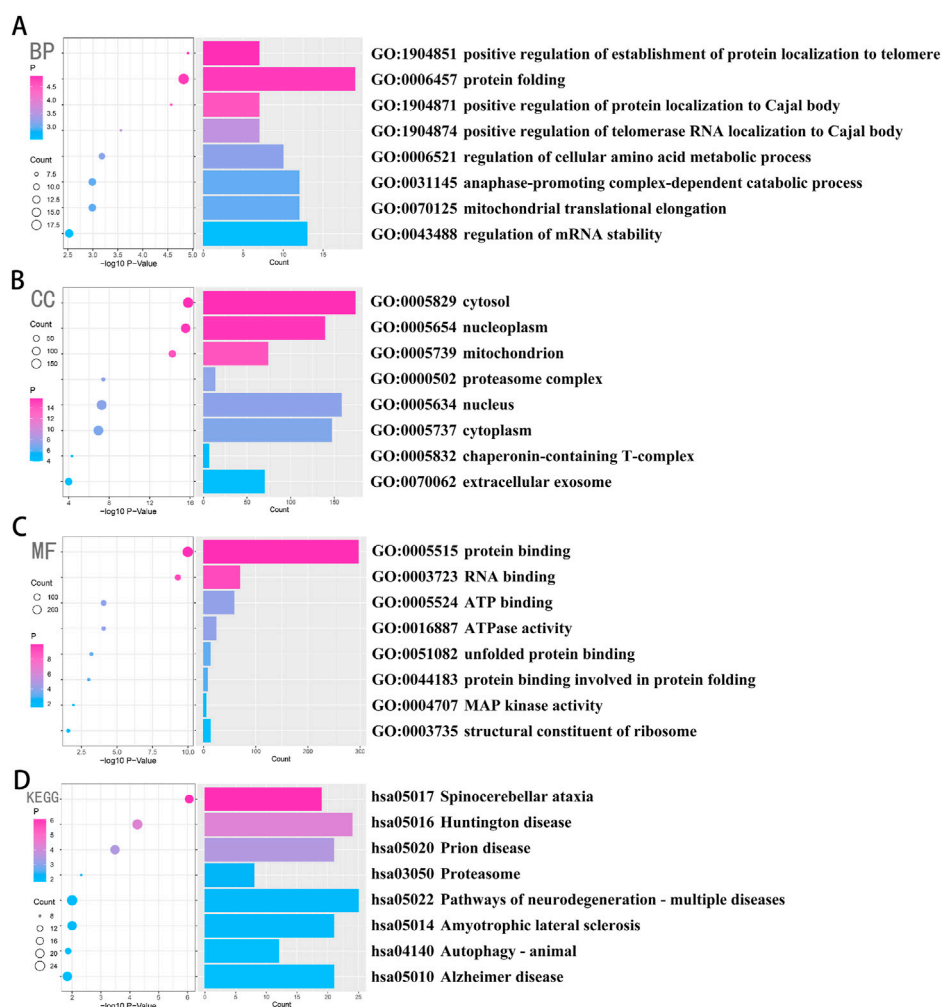


FIGURE 2

Chaperonin containing TCP1 subunit 2 (CCT2) was closely related to the process of β -amyloid formation and clearance in AD. (A–C) Biological processes (BP), cellular components (CC), and molecular functions (MF) were enriched in gene sets that were positively associated with CCT2 in Pearson's test, whose credibility gradually increases from blue to red, and the size of the circle exhibits the number of genes contained in the corresponding pathway. (D) The signaling pathways (Kyoto Encyclopedia of Genes and Genomes (KEGG)) were enriched in the gene set that was positively associated with CCT2 in Pearson's test, whose credibility gradually increased from blue to red, and the size of the circle exhibits the number of genes contained in the corresponding pathway.

Results

Identification of the DEGs in AD

To investigate the differences in gene transcriptome between AD and normal controls, we conducted the following analysis. First, the PCA chart demonstrated that there are significant differences between AD and con groups in GSE33000, allowing for subsequent analysis (Figure 1A). Boxchart displayed standardized data, eliminating intra-group differences (Figure 1B). Second, we observed that CCT2 was significantly downregulated in AD, $p = -1.26e-$

22 and $\log FC = -0.07632$ (Figure 1C), which suggested that the low CCT2 expression is associated with AD. There were 4,381 DEGs in AD, with 2,152 upregulated and 2,229 downregulated genes (Figure 1E), among which the top 50 upregulated and 50 downregulated genes are indicated in the heatmap (Figure 1D). However, the CCT2-low group had 1,273 DEGs compared with the CCT2-high group, of which 561 were upregulated and 712 were downregulated (Figure 1F). A total of 789 genes were either upregulated or downregulated in the AD-con and CCT2 low-high groups, which may be associated with both CCT2 expression and AD.

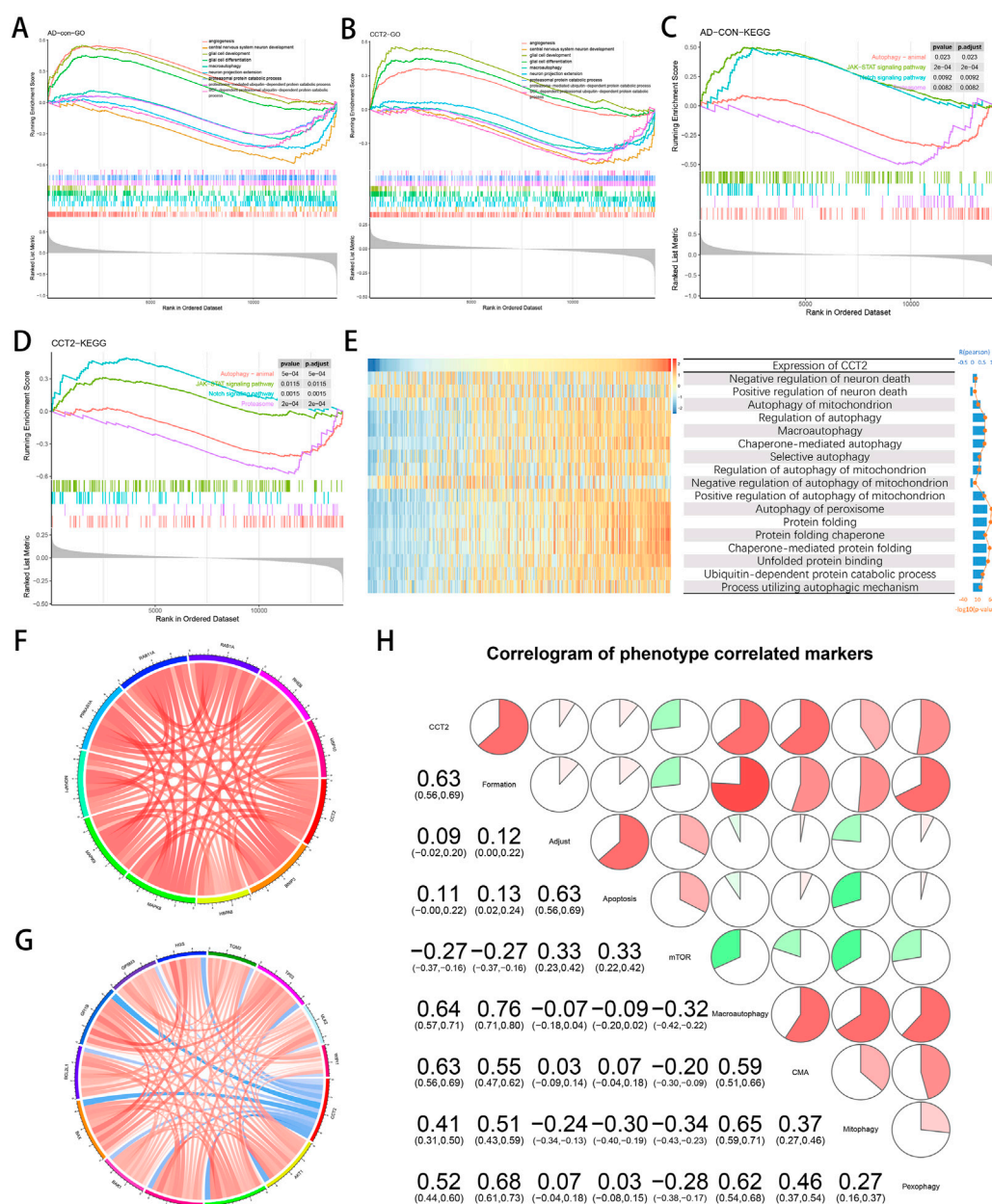


FIGURE 3

Correlation of Chaperonin containing TCP1 subunit 2 (*CCT2*) expression with autophagy and other gene sets. (A) Biological processes (BP) of Alzheimer's disease (AD) enrichment. (B) BP of *CCT2*-low enrichment. (C) The Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways enriched in AD. (D) The KEGG pathway enriched in *CCT2*-low. (E) Heatmaps display the enrichment scores for *CCT2* expression and related pathways in GSE33000; samples are arranged in *CCT2* ascending order, and bar and line plots on the right indicate the analyzed R and p-values. (F) Autophagy-related genes positively associated with *CCT2*, as indicated in red. (G) Autophagy-related genes negatively associated with *CCT2*, as indicated in blue. (H) Correlation of *CCT2* with autophagy-related gene sets, the correlation coefficient is presented in the lower left and simultaneously in the upper right, red for positive correlation and green for the negative correlation.

CCT2 downregulation was linked to protein misfolding and neurodegenerative diseases

Four hundred and sixty genes with the strongest positive correlation with *CCT2* were selected using Pearson correlation

analysis to explore the relevant biological functions of *CCT2*. GO and KEGG analyses were performed according to the above-mentioned gene sets. Genes associated with *CCT2* in biological processes (BP) are primarily enriched in protein folding pathways, regulation of telomere protein localization related to the Cajal body, and regulation of mRNA stability (Figure 2A).

Additionally, the most relevant cellular components (CC) of *CCT2* and its related genes included cell cytoplasm and T complex proteins (Figure 2B) and were related to exosomes (Figure 2C), whose molecular function (MF) was protein binding and folding, RNA binding, and ribosome composition. Moreover, the most related signaling pathway (KEGG) was mainly associated with various neurodegenerative diseases, including AD, and autophagy (Figure 2D). These findings indicated that downregulating *CCT2* in patients with AD may be significant for snRNP formation, mRNA splicing, protein folding, and clearance of misfolded proteins by autophagy. Thus, *CCT2* was associated with the production and clearance of amyloid proteins, and a possible cause of AD was *CCT2* downregulation.

CCT2 positively regulates the occurrence of multiple autophagy and reduces neuronal death

The transcriptomic data was analyzed using GSEA and GSVA. In GSEA, the Janus kinase (JAK)-signal transducer and activator of transcription (STAT) signaling pathway and the Notch signaling pathway were significantly enriched in the AD group compared to the con group; however, there was a contrasting observation in proteasome and animal autophagy (Figure 3A). Meanwhile, there was a similar trend in the *CCT2*-low/high group (Figure 3B). Additionally, compared with the con group, glial cell development and differentiation and angiogenesis-related pathways were significantly enriched in AD, opposing the observation in protein catabolism and neuronal development (Figure 3C), with a similar trend in the *CCT2*-low/high group (Figure 3D). This suggests that the downregulation of *CCT2* may be a cause of AD. It has been proven that the accumulation of amyloid protein can affect the production of angiogenic factors (Skaaraas et al., 2021).

In AD progression, *CCT2* may alleviate amyloid aggregation by promoting normal protein folding and autophagy. Thus, the method used by the study was GSVA for calculating the enrichment score of *CCT2* expression levels for several pathway and BP. These include neuronal death regulation, mitochondrial autophagy, chaperone-mediated autophagy, peroxisomal autophagy, and other related pathway. BP such as protein folding and its decomposition. The enrichment score indicated that *CCT2* was positively correlated with most autophagy-related BP and protein folding; however, it was reversed in the negative regulation of neuronal death and the negative regulation of mitochondrial autophagy. (Figure 3E). These results indicate that high *CCT2* expression can inhibit neuronal death while enhancing mitophagy, which is considered to be the target of AD treatment (Xie et al., 2022).

Since *CCT2* was observed to be involved in AD, we investigated its relationship with the expression levels of some important autophagy-related genes. The results indicated that *CCT2* expression was positively correlated with *MAPK8*, *HSPA8*, *NCKAP1*, *RAB11A*, and *RAB1A* (Figure 3F), and negatively associated with *BAX*, *MAPK3*, *ITGB4*, *ATG16L2*, and *ERBB2* (Figure 3G). Using the Pearson matrix diagram (Figure 3H), high *CCT2* expression revealed a significant and positive correlation with autophagy formation, macroautophagy, and autophagy mediated by molecular chaperones and a negative correlation with the mTOR pathway. This validates that *CCT2* downregulation affects normal autophagy for clearing Tau and A β , thus, causing AD.

The logistic model was constructed for AD prediction

Through ppi network analysis, we screened 36 hub genes from co-DEGs, including *CCT2*, *ACTR2*, *CLTA* et al. Using the results obtained from the PPI network analysis (Supplementary Figure one to two), we extracted the expression profiles of the hub genes to construct a predictive model. Using LASSO regression, 12 genes were selected with non-zero regression coefficients and value of lambda.min = 0.003690707 (Figures 4A,B). *CCT2* was further used to construct logistic regression prediction models as follows: risk score = (4.0041 \times ARAF- 1.9746 \times ACTR2- 5.8043 \times ATP5F1+ 15.7535 \times ATP6V1A+ 9.4168 \times ATP6V1C1- 15.9159 \times CA10-1.8964 \times GNG11 + 4.0073 \times NRXN1- 10.1124 \times PPP1A2+ 4.2734 \times PPP1R1B- 11.3482 \times PPP2CA- 4.1251 \times RAN+ 2.2023 \times *CCT2*). The heatmap indicated the relationship between prediction score and disease, age, and related genes (Figure 4C). The ROC curve indicates that the area under curve (AUC) is 0.9671 and 0.9700 (Figure 4D,E) in the training and validation sets, respectively. In the external validation set (GSE44768, GSE44770, GSE44771, and GSE140829), AUC values were 0.9681, 0.9724, 0.923, and 0.6342 for prefrontal samples, hippocampal samples, cerebellar samples, and whole blood samples, respectively (Figure 4F), indicating that the model has high accuracy in AD prediction.

Additionally, *CCT2* was expressed in all cell types in the human brain (Figure 4G) and significantly downregulated in various brain regions (Figure 4H), indicating that *CCT2* and its related genes are significantly correlated with AD and have broad prospects as a biomarker.

Transcriptome was combined with micro RNA omics analysis

There are several studies suggesting that miRNA acts on target genes through exosomes and thus affects neurodegenerative diseases (Lydie et al., 2013; Jiang et al.,

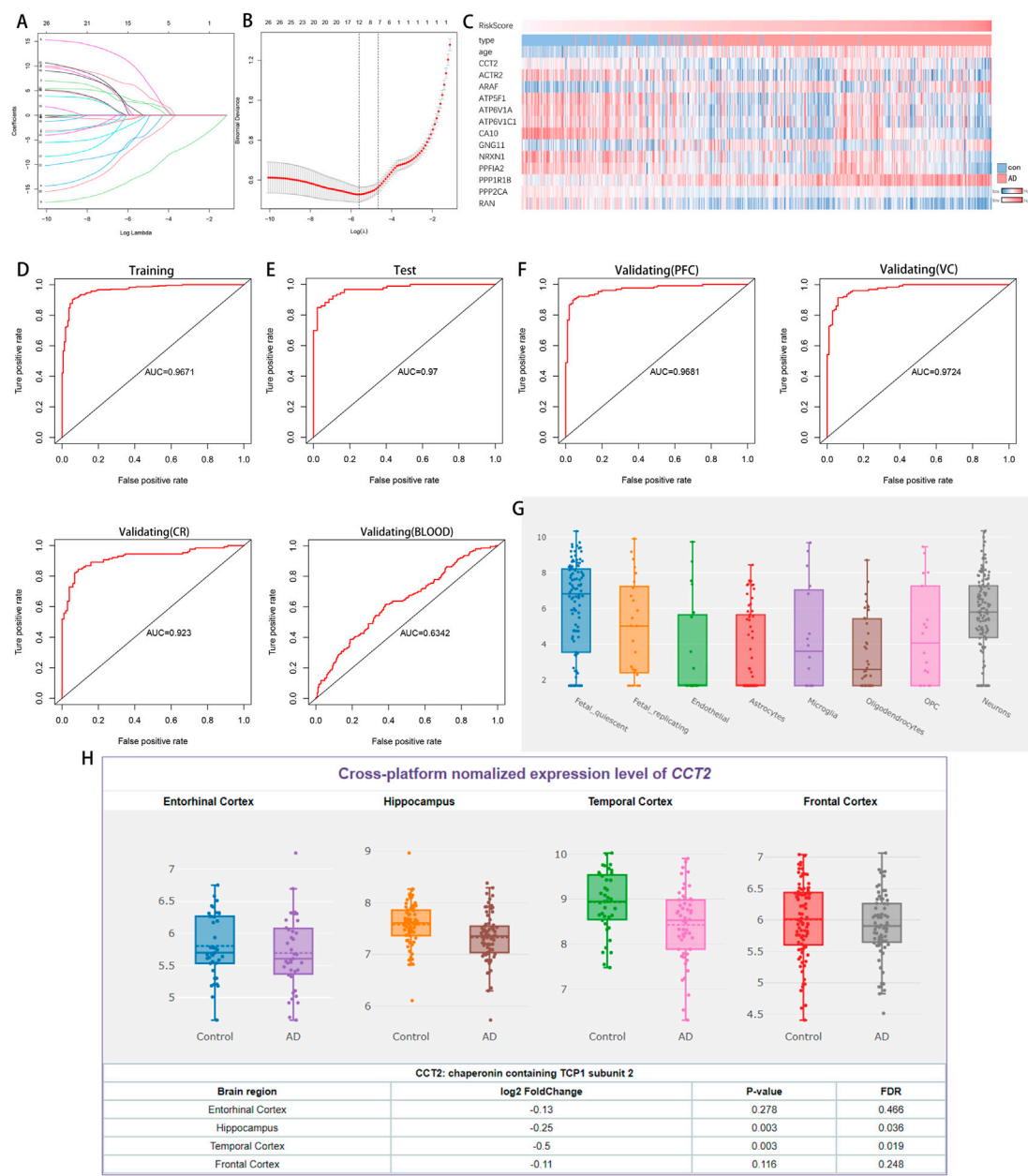
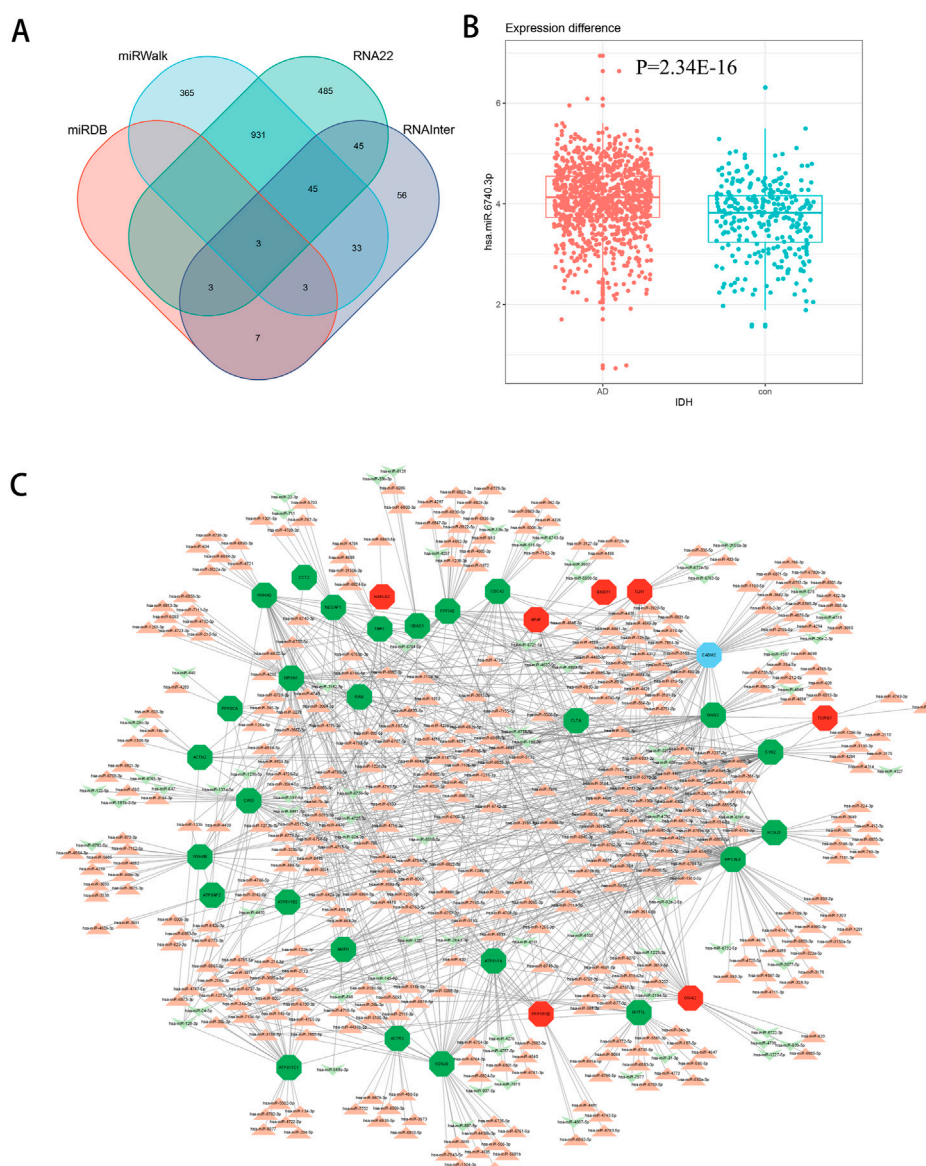


FIGURE 4 A logistic model for identifying Alzheimer's disease (AD) and the expression of *Chaperonin containing TCP1 subunit 2* (*CCT2*) in different tissues and cells. (A–B) Least absolute shrinkage and selection operator (LASSO) model. (C) Heatmap demonstrates changes in category, age, and gene expression as predictive scores rise. (D) Receiver operating characteristic (ROC) curve analysis of the training set (GSE33000). (E) ROC curve analysis of the validation set (GSE33000). (F) The ROC curve analysis of the external validation set (GSE44768, GSE44770, GSE44771, and GSE140829). (G) *CCT2* expression observed in various cells of the brain. (H) *CCT2* was significantly downregulated in various brain regions of patients with AD.

2019). We used four miRNA databases for joint prediction (Figure 5A), among which three were predicted for *CCT2*, including miR-196b-3p, miR-4778-3p, and miR-6740-3p, where miR-6740-3p was significantly different in blood samples (Figure 5B). Thus, miR-6740-3p may inhibit *CCT2*

translation by binding to its transcript, which may be a potential cause of AD. We analyzed all 36 hub genes in the same way, and the mRNA-miRNA interaction network revealed that the majority of the miRNAs corresponding to the downregulated genes in AD were upregulated, confirming

**FIGURE 5**

The mRNA-miRNA interaction network. **(A)** Four databases—miRDB, miRWalk, RNA22, and RNAInter—were combined for CCT2-targeting miRNA prediction. **(B)** The boxplot demonstrates the expression level of miR-6740-3p in the serum samples (GSE120584) ($\log_{2}FC = 0.38540845$, $p = 2.34E-16$). **(C)** The interaction network presents the hub genes and their corresponding miRNA, with octagon nodes representing genes and other shapes representing miRNA. Upregulation is indicated in red; downregulation is indicated in green.

that the interaction network had a good predictive value (Figure 5C).

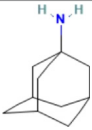
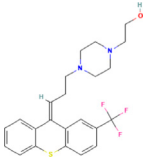
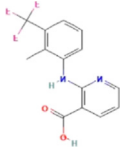
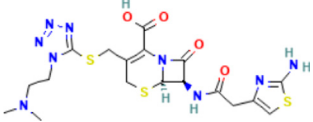
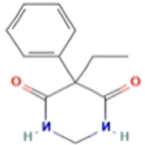
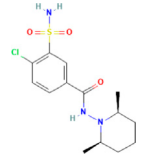
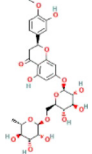
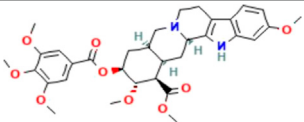
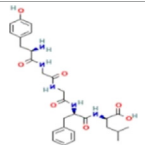
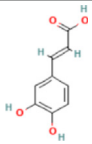
Small-molecule drug prediction based on the hub genes

CCT2 may be a novel target for AD therapy, thus, the protein-drug interactions must be predicted. Since

macromolecular drugs are difficult to cross the blood-brain barrier, we used the DSingDB database for model gene prediction to identify ten viable small molecule drugs. Additionally, they were sorted and displayed based on the p -value (Table 1).

Some studies have shown that *Amantadine* may have a new beneficial effect on axial symptoms of Parkinson's disease (PD) patients with subthalamic nucleus deep brain stimulation (Chan et al., 2013). Additionally, *Reserpine* is an antihypertensive agent

TABLE 1 List of drugs recommended for treating AD by targeting CCT2

| Name | <i>p</i> -value | Chemical formula | Structure |
|------------------|-----------------|--|---|
| Amantadine HL60 | 3.55E-05 | C ₁₀ H ₁₇ N |  |
| Flupentixol HL60 | 7.71E-05 | C ₂₃ H ₂₃ F ₃ N ₂ O ₂ S |  |
| Flunixin HL60 | 4.36E-04 | C ₁₄ H ₁₁ F ₃ N ₂ O ₂ |  |
| Cefotiam PC3 | 4.88E-04 | C ₁₈ H ₂₃ N ₉ O ₄ S ₃ |  |
| Primidone PC3 | 0.001139 | C ₁₂ H ₁₄ N ₂ O ₂ |  |
| Clopidamide HL60 | 0.00128 | C ₁₄ H ₂₀ ClN ₃ O ₃ S |  |
| Hesperidin PC3 | 0.001862 | C ₂₈ H ₃₄ O ₁₅ |  |
| Reserpine TTD | 0.014852 | C ₃₃ H ₄₀ N ₂ O ₉ |  |
| Enkephalin | 0.016133 | C ₂₈ H ₃₇ N ₅ O ₇ |  |
| Caffeic acid PC3 | 0.018692 | C ₉ H ₈ O ₄ |  |

whose action is attributed to its ability to inhibit the vesicle monoamine transporter *VMAT*, thereby reducing the level of bioamine neurotransmitters in synaptic vesicles. One study found that reserpine improves A β toxicity in *caenorhabditis elegans* model of AD (Arya et al., 2009). Similarly, several small-molecule drugs seem to be associated with AD. Elevated enkephalins cause neuronal and behavioral disorders in transgenic mouse models of AD (Meilandt et al., 2008). Caffeic acid slows the development of AD by increasing cognitive function, alleviating brain damage, and inhibiting the AD-induced increase in AChE activity and nitrite production (Wang et al., 2016).

Discussion

AD has gradually grown to be one of the most significant diseases of this century as a typical neurodegenerative condition affecting the elderly. Over 50 million people worldwide currently suffer from dementia, with AD accounting for 60–80% of all dementia cases (Porsteinsson et al., 2021). Acetylcholinesterase inhibitor (AChEI) is a common drug for the treatment of AD; however, it can only treat its symptoms and have some side effects (Lane et al., 2018). Studies have reported that AD is caused by abnormal folding of A β protein, and the neurofibrillary tangles are caused by excessive phosphorylation of Tau (Kozlov et al., 2017), which is accompanied by neuronal apoptosis, which is irreversible. Therefore, the onset of AD is usually irreversible (Ganzer, 2007).

Previous studies have demonstrated the eukaryotic chaperone tailless complex polypeptide one ring complex and its eight subunits can prevent the formation of protein aggregates (Behrends et al., 2006; Noormohammadi et al., 2016). A recently published paper has demonstrated that *CCT2* binds to protein aggregates, recruits autophagosomes to endocytose, and degrades Tau (Ma et al., 2022). Therefore, our study used the transcriptome sequencing dataset of AD prefrontal cells to explore the regulatory mechanism of *CCT2* in AD. We observed that *CCT2* was significantly downregulated in patients with AD, suggesting that the *CCT2* downregulation may be a contributing factor for AD. By considering the intersection of AD-con and *CCT2*-low/high, we identified the common DEGs. Further, using PPI network analysis, 36 hub genes of co-DEGs were screened out, including *CCT2*, *ACTR2*, and *TCIRG1*. Among them, *MYT1L* was observed to induce cell reprogramming into cholinergic neurons and provide a strategy for treating AD (Liang et al., 2018). *ATP6V1A* is involved in AD via the synaptic vesicle cycle, phagosome, and oxidative phosphorylation downregulation (Zhou et al., 2021). Additionally, *VDAC1* has been observed to impact AD occurrence by regulating mitochondrial function (Shoshan-Barmatz et al., 2018).

This indicates a strong correlation between hub genes and AD occurrence. Since *CCT2* is also closely related to these genes, it may play an important role in the regulation of hub genes, providing a reference for future research.

However, functional enrichment analysis revealed that the genes positively correlated with *CCT2*, selected by Pearson correlation coefficient, were associated with telomerase, Cajal body positioning, mRNA splicing, protein folding, and MAPK activity, indicating that the genes highly associated with *CCT2* are involved in the entire process of transcription and translation. This is also consistent with the mechanism of co-regulation of protein homeostasis by a molecular chaperone and aggregate autophagy reported in a study by Ma et al. (Ma et al., 2022). Meanwhile, the MAPK signaling pathway is believed to function during the early stages of AD (Johnson and Bailey, 2003), and the pathway enrichment analysis revealed that it is related to multiple neurodegenerative diseases and autophagy in animals whereas autophagic deficiency is a widely recognized cause of AD (Reddy and Oliver, 2019; Fang et al., 2019).

Meanwhile, GSEA analysis of the DEGs and GSVA analysis of the *CCT2*-related genes revealed that *CCT2* affects autophagy. GSEA analysis demonstrated that the JAK-STAT pathway, the Notch signaling pathway, angiogenesis, and development and differentiation of glial cells were enriched in AD-con. At the same time, *CCT2*-low/high groups whereas proteasome, animal autophagy, protein breakdown, and neuronal development were not enriched. Several pathways have been reported to correlate with AD occurrence. For example, the JAK-STAT signaling pathway has been reported to play a potential role in the molecular mechanism regulating cellular autophagy (Chen et al., 2021). The Notch pathway is linked to autophagy (Ko et al., 2020), and angiogenesis is also believed to promote AD (Vagnucci and Li, 2003). Additionally, there is evidence that the AD and the *CCT2*-low expression groups can clear A β and Tau by inhibiting proteasomes and autophagy (Bonet-Costa et al., 2016). In contrast, GSVA revealed that the expression level of *CCT2* was negatively associated with neuronal death, suggesting that high *CCT2* expression can inhibit neuronal death. Furthermore, the *CCT2* expression level was significantly and positively correlated with mitophagy, macroautophagy, pexophagy, protein folding, and protein metabolism; however, the R-value for mitophagy was relatively low, possibly since *CCT2* does not directly affect mitophagy. This is consistent with the first report of *CCT2*-mediated aggregate elimination and indicates that *CCT2* can regulate the levels of A β and Tau by regulating protein folding and promoting autophagy. Increasing evidence suggests that autophagy plays a role in scavenging abnormal proteins, thus, affecting the clearance of A β and Tau when autophagy activation is hampered (Dou et al., 2020). Hence, autophagy

dysfunction plays a crucial role in the pathological process of AD (Li et al., 2010).

We also analyzed the correlation between *CCT2* and autophagy-related genes, and top ten autophagy genes with the strongest positive correlation included *BNIP3*, *HSPA8*, and *MAPK8* etc. Whereas those with the strongest negative correlation included *ATG16L2*, *ATG9B*, and *BAX* etc. This indicates that *CCT2* may affect autophagy by interacting with genes that are highly associated with autophagy. The autophagy gene set correlation analysis revealed that *CCT2* is positively correlated with the initiation stage of autophagy, macroautophagy, and chaperone-mediated autophagy, which also suggested that *CCT2* may affect the occurrence of autophagy via some mechanism. We also observed the inhibition of the mammalian target of the rapamycin (mTOR) pathway. Additionally, high expression of mTOR-inhibiting autophagy has been demonstrated in neurodegenerative diseases (Zhu et al., 2019).

During the construction of the clinical prediction model, we used LASSO regression to screen for genes with regression coefficients greater than zero and then combined these genes with *CCT2* to build the logistic model. The model performed well with high AUC values in the prefrontal cortex, visual cortex, and hippocampus; however, it performed poorly in whole blood samples, which could be attributed to the fact that the brain tissue samples were used to construct the model. Some of the genes involved in the model have been observed to be involved in AD, of which *CCT2* is significantly downregulated in AD, *PPP1R1B* can regulate *cAMP response element-binding protein (CREB)* phosphorylation, and *CREB* dysfunction is one of the causes of AD (Cho et al., 2015) whereas *NRXN1* is involved in memory recovery in rats by affecting synaptic plasticity (Zhang et al., 2021). The ROC curve revealed that the model had high AUC values in the training, test, and external validation sets; thus, the expression of these genes can be used as a biomarker for AD. Online database analysis also demonstrated that *CCT2* is expressed in various brain cells, with neuronal cells having the highest levels of expression. *CCT2* expression, in contrast, was significantly reduced in various brain regions, including the entorhinal cortex, hippocampus, frontal cortex, and frontal cortex, providing further evidence that low *CCT2* expression is one of the mechanisms of AD pathogenesis.

The study also investigated the causes of low *CCT2* expression and observed that gene mutations, DNA methylation, and miRNA may cause changes in gene expression, causing AD (Qin et al., 2020; De Jager et al., 2014; Akhter and Bekris, 2019). Thus, we aimed to explore the miRNA interacting with hub genes, most of which had the opposite expression profile of their target genes in AD. For example, miR-6740-3p, which interacts with *CCT2*, is significantly upregulated in AD ($\log_{2}FC = 0.38540845$, $p = 2.34E-16$) and miR-661, which interacts with *HPCAL4*, *NECAP1*, *CLTA*, and *GNAI2*,

has been observed to be involved in AD via metabolic and stress pathways (Hojati et al., 2021). The miR-501-3p, which interacts with *CADM2*, may impact AD by regulating cell division (Hara et al., 2017); hsa-miR-107, which interacts with *ACTR2*, *AMPH*, and *RAN*, targets A β precursor protein (APP) and influences AD (Hébert et al., 2008). This study can assist researchers in screening for appropriate miRNA and validating their biological functions to obtain effective biological results. (The specific pathways by which most miRNAs affect AD are currently unknown, albeit the mRNA-miRNA interaction network serves as a reference for AD diagnosis and treatment. Similarly, the prediction of the last small-molecule drugs provides a point of reference for targeting *CCT2* to treat AD.

Conclusion

Using bioinformatic analysis, this study used multiple datasets and revealed that the low expression of *CCT2* in AD may be responsible for the inhibition of autophagy in AD. The PPI network was used to screen out potential AD biomarkers with diagnostic value, and the mRNA-miRNA interaction network was constructed to predict the potential miRNA. These findings contribute to our understanding of the pathogenesis of AD and provide new guidelines for the treatment and diagnosis of the disease.

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

Z-LL provided the experimental idea of this study and reviewed and modified the manuscript. M-XT and F-YX designed the experiments and collected the data; M-XT verified the experiments, analyzed the data, and visualized the results; G-BY and Q-XY edited pictures and tables; M-XT, F-YX, and Q-XY wrote the initial manuscript, with F-XQ and L-GD providing amendments and financial assistance. All authors contributed to this article and validated the final version to be submitted.

Funding

This work was supported by the National Key Research and Development Program of China [Grant Number 2021YFA1500400]. Science and Technology Department of

Jilin Province [Grant Number 20200201386JC]. Science and Technology Department of Jilin Province [Grant Number 20190701037 GH]. Education department of Jilin Province [Grant Number JJKH20200948KJ]. The funding bodies played no role in study design, in the collection, analysis, and interpretation of data, in the report's writing, and in the decision to submit the article for publication.

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.

References

- Alzheimer's Association (2020). 2020 Alzheimer's disease facts and figures. *Alzheimer's Dementia* 16 (3), 391–460. doi:10.1002/alz.12068
- Akhter, R., and Bekris, L. M. (2019). Potential role of miRNA-140 in Alzheimer's disease. *Aging (Albany, NY)* 11 (4), 1087–1088. doi:10.18632/aging.101827
- Arya, U., Dwivedi, H., and Subramaniam, J. R. (2009). Reserpine ameliorates Abeta toxicity in the Alzheimer's disease model in *Caenorhabditis elegans*. *Exp. Gerontol.* 44 (6–7), 462–466. doi:10.1016/j.exger.2009.02.010
- 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
- Behrends, C., Langer, C. A., Boteva, R., Böttcher, U. M., Stemp, M. J., Schaffar, G., et al. (2006). Chaperonin TRiC promotes the assembly of polyQ expansion proteins into nontoxic oligomers. *Mol. Cell.* 23 (6), 887–897. doi:10.1016/j.molcel.2006.08.017
- Bonet-Costa, V., Pomatto, L. C., and Davies, K. J. (2016). The proteasome and oxidative stress in Alzheimer's disease. *Antioxid. Redox Signal.* 25 (16), 886–901. doi:10.1089/ars.2016.6802
- Chan, H. F., Kukkle, P. L., Merello, M., Lim, S. Y., Poon, Y. Y., and Moro, E. (2013). Amantadine improves gait in PD patients with STN stimulation. *Park. Relat. Disord.* 19 (3), 316–319. doi:10.1016/j.parkreldis.2012.11.005
- Chen, D., Liu, Y., Chen, J., Lin, H., Guo, H., Wu, Y., et al. (2021). JAK/STAT pathway promotes the progression of diabetic kidney disease via autophagy in podocytes. *Eur. J. Pharmacol.* 902, 174121. doi:10.1016/j.ejphar.2021.174121
- Chen, E. Y., Tan, C. M., Kou, Y., Duan, Q., Wang, Z., Meirelles, G. V., et al. (2013). Enrichr: Interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinforma.* 14, 128. doi:10.1186/1471-2105-14-128
- 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
- Cho, K., Cho, M. H., Seo, J. H., Peak, J., Kong, K. H., Yoon, S. Y., et al. (2015). Calpain-mediated cleavage of DARPP-32 in Alzheimer's disease. *Aging Cell.* 14 (5), 878–886. doi:10.1111/acel.12374
- De Jager, P. L., Srivastava, G., Lunnon, K., Burgess, J., Schalkwyk, L. C., Yu, L., et al. (2014). Alzheimer's disease: Early alterations in brain DNA methylation at ANK1, BIN1, RHBDF2 and other loci. *Nat. Neurosci.* 17 (9), 1156–1163. doi:10.1038/nn.3786
- Dominguez-Almendros, S., Benítez-Parejo, N., and Gonzalez-Ramirez, A. R. (2011). Logistic regression models. *Allergol. Immunopathol.* 39 (5), 295–305. doi:10.1016/j.aller.2011.05.002
- Dou, J., Su, P., Xu, C., Wen, Z., Mao, Z., and Li, W. (2020). Targeting Hsc70-based autophagy to eliminate amyloid β oligomers. *Biochem. Biophys. Res. Commun.* 524 (4), 923–928. doi:10.1016/j.bbrc.2020.02.016
- Dweep, H., Sticht, C., Pandey, P., and Gretz, N. (2011). miRWalk-database: prediction of possible miRNA binding sites by "walking" the genes of three genomes. *J. Biomed. Inf.* 44 (5), 839–847. doi:10.1016/j.jbi.2011.05.002
- Fang, E. F., Hou, Y., Palikaras, K., Adriaanse, B. A., Kerr, J. S., Yang, B., et al. (2019). Mitophagy inhibits amyloid- β and tau pathology and reverses cognitive

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.2022.967730/full#supplementary-material>

deficits in models of Alzheimer's disease. *Nat. Neurosci.* 22 (3), 401–412. doi:10.1038/s41593-018-0332-9

Ganzer, C. A. (2007). Assessing Alzheimer's disease and dementia: Best practices in nursing care. *Geriatr. Nurs.* 28 (6), 358–365. doi:10.1016/j.gerinurse.2007.10.008

Gatica, D., Lahiri, V., and Klionsky, D. J. (2018). Cargo recognition and degradation by selective autophagy. *Nat. Cell. Biol.* 20 (3), 233–242. doi:10.1038/s41556-018-0037-z

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

Hara, N., Kikuchi, M., Miyashita, A., Hatsuta, H., Saito, Y., Kasuga, K., et al. (2017). Serum microRNA miR-501-3p as a potential biomarker related to the progression of Alzheimer's disease. *Acta Neuropathol. Commun.* 5 (1), 10. doi:10.1186/s40478-017-0414-z

Hébert, S. S., Horré, K., Nicolaï, L., Papadopoulou, A. S., Mandemakers, W., Silahatoglu, A. N., et al. (2008). Loss of microRNA cluster miR-29a/b-1 in sporadic Alzheimer's disease correlates with increased BACE1/beta-secretase expression. *Proc. Natl. Acad. Sci. U. S. A.* 105 (17), 6415–6420. doi:10.1073/pnas.0710263105

Hojati, Z., Omid, F., Dehbashi, M., and Mohammad Soltani, B. (2021). The highlighted roles of metabolic and cellular response to stress pathways engaged in circulating hsa-miR-494-3p and hsa-miR-661 in Alzheimer's disease. *Iran. Biomed. J.* 25 (1), 62–67. doi:10.29252/ibj.25.1.62

Huang, d., Sherman, B. T., and Lempicki, R. A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* 4 (1), 44–57. doi:10.1038/nprot.2008.211

Jiang, Y., Liu, J., Chen, L., Jin, Y., Zhang, G., Lin, Z., et al. (2019). Serum secreted miR-137-containing exosomes affects oxidative stress of neurons by regulating OXRI in Parkinson's disease. *Brain Res.* 1722, 146331. doi:10.1016/j.brainres.2019.146331

Johansen, T., and Lamark, T. (2020). Selective autophagy: ATG8 family proteins, LIR motifs and cargo receptors. *J. Mol. Biol.* 432 (1), 80–103. doi:10.1016/j.jmb.2019.07.016

Johnson, G. V., and Bailey, C. D. (2003). The p38 MAP kinase signaling pathway in Alzheimer's disease. *Exp. Neurol.* 183 (2), 263–268. doi:10.1016/s0014-4886(03)00268-1

Khaminets, A., Behl, C., and Dikic, I. (2016). Ubiquitin-Dependent and independent signals in selective autophagy. *Trends Cell. Biol.* 26 (1), 6–16. doi:10.1016/j.tcb.2015.08.010

Ko, S. H., Apple, E. C., Liu, Z., and Chen, L. (2020). Age-dependent autophagy induction after injury promotes axon regeneration by limiting NOTCH. *Autophagy* 16 (11), 2052–2068. doi:10.1080/15548627.2020.1713645

Kopeikina, K. J., Carlson, G. A., Pitstick, R., Ludvigson, A. E., Peters, A., Luebke, J. I., et al. (2011). Tau accumulation causes mitochondrial distribution deficits in neurons in a mouse model of tauopathy and in human Alzheimer's disease brain. *Am. J. Pathol.* 179 (4), 2071–2082. doi:10.1016/j.ajpath.2011.07.004

Kozlov, S., Afonin, A., Evsyukov, I., and Bondarenko, A. (2017). Alzheimer's disease: As it was in the beginning. *Rev. Neurosci.* 28 (8), 825–843. doi:10.1515/revneuro-2017-0006

- Krzywinski, M., Schein, J., Birol, I., Connors, J., Gascoyne, R., Horsman, D., et al. (2009). Circos: An information aesthetic for comparative genomics. *Genome Res.* 19 (9), 1639–1645. doi:10.1101/gr.092759.109
- Kuleshov, M. V., Jones, M. R., Rouillard, A. D., Fernandez, N. F., Duan, Q., Wang, Z., et al. (2016). Enrichr: A comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res.* 44 (W1), W90–W97. doi:10.1093/nar/gkw377
- Lane, C. A., Hardy, J., and Schott, J. M. (2018). Alzheimer's disease. *Eur. J. Neurol.* 25 (1), 59–70. doi:10.1111/ene.13439
- Li, L., Zhang, X., and Le, W. (2010). Autophagy dysfunction in Alzheimer's disease. *Neurodegener. Dis.* 7 (4), 265–271. doi:10.1159/000276710
- Li, Q., Liu, Y., and Sun, M. (2017). Autophagy and alzheimer's disease. *Cell. Mol. Neurobiol.* 37 (3), 377–388. doi:10.1007/s10571-016-0386-8
- Liang, X. G., Tan, C., Wang, C. K., Tao, R. R., Huang, Y. J., Ma, K. F., et al. (2018). Myt1l induced direct reprogramming of pericytes into cholinergic neurons. *CNS Neurosci. Ther.* 24 (9), 801–809. doi:10.1111/cns.12821
- Lin, Y., Liu, T., Cui, T., Wang, Z., Zhang, Y., Tan, P., et al. (2020). RNAInter in 2020: RNA interactome repository with increased coverage and annotation. *Nucleic Acids Res.* 48 (D1), D189–D197. doi:10.1093/nar/gkz804
- Liu, L., Wu, Q., Zhong, W., Chen, Y., Zhang, W., Ren, H., et al. (2020). Microarray analysis of differential gene expression in alzheimer's disease identifies potential biomarkers with diagnostic value. *Med. Sci. Monit. Int. Med. J. Exp. Clin. Res.* 26, e919249. doi:10.12659/MSM.919249
- Lydie, M., Melissa, R., Haruki, H., Seng, K. N., Christine, E., Svetlana, V., et al. (2013). Neuronal exosomal miRNA-dependent translational regulation of astroglial glutamate transporter GLT1. *J. Biol. Chem.* 288 (10), 7105–7116. doi:10.1074/jbc.M112.410944
- Ma, X., Lu, C., Chen, Y., Li, S., Ma, N., Tao, X., et al. (2022). CCT2 is an aggregophagy receptor for clearance of solid protein aggregates. *Cell.* 185 (8), 1325–1345. e22. doi:10.1016/j.cell.2022.03.005
- Meilandt, W. J., Yu, G. Q., Chin, J., Roberson, E. D., Palop, J. J., Wu, T., et al. (2008). Enkephalin elevations contribute to neuronal and behavioral impairments in a transgenic mouse model of Alzheimer's disease. *J. Neurosci.* 28 (19), 5007–5017. doi:10.1523/JNEUROSCI.0590-08.2008
- Miranda, K. C., Huynh, T., Tay, Y., Ang, Y. S., Tam, W. L., Thomson, A. M., et al. (2006). A pattern-based method for the identification of MicroRNA binding sites and their corresponding heteroduplexes. *Cell.* 126 (6), 1203–1217. doi:10.1016/j.cell.2006.07.031
- Mootha, V. K., Lindgren, C. M., Eriksson, K. F., Subramanian, A., Sihag, S., Lehar, J., et al. (2003). PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nat. Genet.* 34 (3), 267–273. doi:10.1038/ng1180
- Moussay, E., Kaoma, T., Baginska, J., Muller, A., Van Moer, K., Nicot, N., et al. (2011). The acquisition of resistance to TNF α in breast cancer cells is associated with constitutive activation of autophagy as revealed by a transcriptome analysis using a custom microarray. *Autophagy* 7 (7), 760–770. doi:10.4161/auto.7.7.15454
- Noormohammadi, A., Khodakarami, A., Gutierrez-Garcia, R., Lee, H. J., Koyuncu, S., König, T., et al. (2016). Somatic increase of CCT8 mimics proteostasis of human pluripotent stem cells and extends *C. elegans* lifespan. *Nat. Commun.* 7, 13649. doi:10.1038/ncomms13649
- Porsteinsson, A. P., Isaacson, R. S., Knox, S., Sabbagh, M. N., and Rubino, I. (2021). Diagnosis of early alzheimer's disease: Clinical practice in 2021. *J. Prev. Alzheimers Dis.* 8 (3), 371–386. doi:10.14283/jpad.2021.23
- Qin, Q., Yin, Y., Wang, Y., Lu, Y., Tang, Y., and Jia, J. (2020). Gene mutations associated with early onset familial Alzheimer's disease in China: An overview and current status. *Mol. Genet. Genomic Med.* 8 (10), e1443. doi:10.1002/mgg3.1443
- Reddy, P. H., and Oliver, D. M. (2019). Amyloid beta and phosphorylated tau-induced defective autophagy and mitophagy in alzheimer's disease. *Cells* 8 (5), 488. doi:10.3390/cells8050488
- 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 (7), e47. doi:10.1093/nar/gkv007
- Roca-Agujetas, V., de Dios, C., Abadin, X., and Colell, A. (2021). Upregulation of brain cholesterol levels inhibits mitophagy in Alzheimer disease. *Autophagy* 17 (6), 1555–1557. doi:10.1080/15548627.2021.1920814
- 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.*, 194. doi:10.1093/nar/gkac194
- Shoshan-Barmatz, V., Nahon-Crystal, E., Shtein-Kuzmine, A., and Gupta, R. (2018). VDAC1, mitochondrial dysfunction, and Alzheimer's disease. *Pharmacol. Res.* 131, 87–101. doi:10.1016/j.phrs.2018.03.010
- Skaaraas, G., Melbye, C., Puchades, M. A., Leung, D., Jacobsen, Ø., Rao, S. B., et al. (2021). Cerebral amyloid angiopathy in a mouse model of alzheimer's disease associates with upregulated angiopoietin and downregulated hypoxia-inducible factor. *J. Alzheimers Dis.* 83 (4), 1651–1663. doi:10.3233/JAD-210571
- Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., et al. (2005). Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U. S. A.* 102 (43), 15545–15550. doi:10.1073/pnas.0506580102
- 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
- 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
- Vagnucci, A. H., and Li, W. W. (2003). Alzheimer's disease and angiogenesis. *Lancet (London, Engl.)* 361 (9357), 605–608. doi:10.1016/S0140-6736(03)12521-4
- Vaillant-Beuchot, L., Mary, A., Pardossi-Piquard, R., Bourgeois, A., Lauritzen, I., Eysert, F., et al. (2021). Accumulation of amyloid precursor protein C-terminal fragments triggers mitochondrial structure, function, and mitophagy defects in Alzheimer's disease models and human brains. *Acta Neuropathol.* 141 (1), 39–65. doi:10.1007/s00401-020-02234-7
- Wang, Y., Wang, Y., Li, J., Hua, L., Han, B., Zhang, Y., et al. (2016). Effects of caffeic acid on learning deficits in a model of Alzheimer's disease. *Int. J. Mol. Med.* 38 (3), 869–875. doi:10.3892/ijmm.2016.2683
- 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, C., Zhuang, X. X., Niu, Z., Ai, R., Lautrup, S., Zheng, S., et al. (2022). Amelioration of Alzheimer's disease pathology by mitophagy inducers identified via machine learning and a cross-species workflow. *Nat. Biomed. Eng.* 6 (1), 76–93. doi:10.1038/s41551-021-00819-5
- Xie, Z., Bailey, A., Kuleshov, M. V., Clarke, D. J. B., Evangelista, J. E., Jenkins, S. L., et al. (2021). Gene set knowledge discovery with Enrichr. *Curr. Protoc.* 1 (3), e90. doi:10.1002/cpz1.90
- Xu, M., Zhang, D. F., Luo, R., Wu, Y., Zhou, H., Kong, L. L., et al. (2018). A systematic integrated analysis of brain expression profiles reveals YAP1 and other prioritized hub genes as important upstream regulators in Alzheimer's disease. *Alzheimers Dement.* 14 (2), 215–229. doi:10.1016/j.jalz.2017.08.012
- Yuan, O., Lin, C., Wagner, J., Archard, J. A., Deng, P., Halmaj, J., et al. (2019). Exosomes derived from human primed mesenchymal stem cells induce mitosis and potentiate growth factor secretion. *Stem Cells Dev.* 28 (6), 398–409. doi:10.1089/scd.2018.0200
- Zare-Shahabadi, A., Masliah, E., Johnson, G. V., and Rezaei, N. (2015). Autophagy in Alzheimer's disease. *Rev. Neurosci.* 26 (4), 385–395. doi:10.1515/revneuro-2014-0076
- Zellner, S., Schifferer, M., and Behrends, C. (2021). Systematically defining selective autophagy receptor-specific cargo using autophagosome content profiling. *Mol. Cell.* 81 (6), 1337–1354.e8. e8. doi:10.1016/j.molcel.2021.01.009
- Zhang, D. F., Fan, Y., Xu, M., Wang, G., Wang, D., Li, J., et al. (2019). Complement C7 is a novel risk gene for Alzheimer's disease in Han Chinese. *Natl. Sci. Rev.* 6, 257–274. doi:10.1093/nsr/nwy127
- Zhang, S., Wu, D., Xu, Q., You, L., Zhu, J., Wang, J., et al. (2021). The protective effect and potential mechanism of NRXN1 on learning and memory in ADHD rat models. *Exp. Neurol.* 344, 113806. doi:10.1016/j.expneurol.2021.113806
- Zhang, Z., and Klionsky, D. J. (2022). CCT2, a newly identified aggregophagy receptor in mammals, specifically mediates the autophagic clearance of solid protein aggregates. *Autophagy* 18 (7), 1483–1485. doi:10.1080/15548627.2022.2083305
- Zhou, Z., Bai, J., Zhong, S., Zhang, R., Kang, K., Zhang, X., et al. (2021). Downregulation of ATP6V1A involved in alzheimer's disease via synaptic vesicle cycle, phagosome, and oxidative phosphorylation. *Oxid. Med. Cell. Longev.* 2021, 5555634. doi:10.1155/2021/5555634
- Zhu, Z., Yang, C., Iyaswamy, A., Krishnamoorthi, S., Sreenivasamurthy, S. G., Liu, J., et al. (2019). Balancing mTOR signaling and autophagy in the treatment of Parkinson's disease. *Int. J. Mol. Sci.* 20 (3), 728. doi:10.3390/ijms20030728
- Zou, D., Li, R., Huang, X., Chen, G., Liu, Y., Meng, Y., et al. (2019). Identification of molecular correlations of RBM8A with autophagy in Alzheimer's disease. *Aging* 11 (23), 11673–11685. doi:10.18632/aging.102571



OPEN ACCESS

EDITED BY

Jaqueline Bohrer Schuch,
Federal University of Rio Grande do Sul,
Brazil

REVIEWED BY

Xinglong Yang,
The First Affiliated Hospital of Kunming
Medical University, China
Laralynne Przybyla,
University of California, San Francisco,
United States

*CORRESPONDENCE

Gang Wu,
✉ xd19840611@126.com

SPECIALTY SECTION

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

RECEIVED 05 November 2022

ACCEPTED 09 January 2023

PUBLISHED 17 January 2023

CITATION

Dong X, Li Y, Li Q, Li W and Wu G (2023),
Identification of immune signatures in
Parkinson's disease based on co-
expression networks.
Front. Genet. 14:1090382.
doi: 10.3389/fgene.2023.1090382

COPYRIGHT

© 2023 Dong, Li, Li, Li and Wu. This is an
open-access article distributed under the
terms of the [Creative Commons
Attribution License \(CC BY\)](#). The use,
distribution or reproduction in other
forums is permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original publication in
this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

Identification of immune signatures in Parkinson's disease based on co-expression networks

Xiaolin Dong, Yanping Li, Qingyun Li, Wenhao Li and Gang Wu*

Department of Neurology, The Affiliated Yan'An Hospital of Kunming Medical University, Kunming, Yunnan, China

Parkinson's disease (PD) is a common neurodegenerative disease in middle-aged and elderly people, and there is less research on the relationship between immunity and PD. In this study, the protein-protein interaction networks (PPI) data, 2747 human immune-related genes (HIRGs), 2078 PD-related genes (PDRGs), and PD-related datasets (GSE49036 and GSE20292) were downloaded from the Human Protein Reference Database (HPRD), Amigo 2, DisGeNET, and Gene Expression Omnibus (GEO) databases, respectively. An immune- or PD-directed neighbor co-expressed network construction (IOPDNC) was drawn based on the GSE49036 dataset and HPRD database. Furthermore, a PD-directed neighbor co-expressed network was constructed. Modular clustering analysis was performed on the genes of the gene interaction network obtained in the first step to obtain the central core genes using the GraphWeb online website. The modules with the top 5 functional scores and the number of core genes greater than six were selected as PD-related gene modules. The Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analyses of different module genes were performed. The single sample Gene Set Enrichment Analysis (ssGSEA) algorithm was used to calculate the immune cell infiltration of the PD and the normal samples. The quantitative Reverse Transcription Polymerase Chain Reaction (qRT-PCR) was performed to investigate the expression of module genes. An IOPDNC and PD-directed neighbor co-expressed network (PDNC network) were constructed. Furthermore, a total of 5 immune-PD modules were identified which could distinguish between PD and normal samples, and these module genes were strongly related to PD in protein interaction level or gene expression level. In addition, functional analysis indicated that module genes were involved in various neurodegenerative diseases, such as Alzheimer disease, Huntington disease, Parkinson disease, and Long-term depression. In addition, the genes of the 6 modules were significantly associated with these 4 differential immune cells (aDC cells, eosinophils, neutrophils, and Th2 cells). Finally, the result of qRT-PCR manifested that the expression of 6 module genes was significantly higher in normal samples than in PD samples. In our study, the immune-related genes were found to be strongly related to PD and might play key roles in PD.

KEYWORDS

Parkinson's disease, bioinformatics analysis, co-expression network, immunity, diagnostic markers

Introduction

PD is the second most common neurodegenerative disease after Alzheimer's disease (Del Rey et al., 2018). According to the report, the incidence rate is about 1%–2% (Tarsy, 2012) in the elderly over 60. The typical symptoms are static tremor, slow movement, increased muscle tone, abnormal postural gait, and some non-motor symptoms such as

insomnia and constipation. The main pathological manifestations of the disease are degeneration and loss of nigra dopaminergic neurons and abnormal accumulation of α -synuclein (α -syn) (Kalia and Lang, 2016). However, the occurrence of disease also involves the influence of the environment and epigenetics, so additional research on these underlying factors is required (Elsworth, 2020).

By analyzing the mRNA expression levels of inflammatory mediators, it was found that the intensity of inflammation in PD nigra was notably increased (Pajares et al., 2020). Neuroinflammatory markers include reactive CNS myeloid cells, T lymphocytes, and increased proinflammatory cytokines/chemokines in the blood, cerebrospinal fluid (CSF), and brain parenchyma of the patients (Marras et al., 2018). These inflammatory markers change with elevated levels of T cells and autoantibodies (anti- α -syn and anti-GM1-gangliosides) in peripheral blood and CSF of PD patients. The accumulation of α -syn triggers an immune response characterized by inflammation (Kline et al., 2021). In rat studies, overexpression of α -syn was found to cause microglial activation and release of inflammatory factors (IFN- γ and resolvin D1) (Krashia et al., 2019). Moreover, α -syn can trigger neuronal autoantigen presentation (Cebrián et al., 2014), which relies on MHC I and MHC II. There are a large number of drugs that have been proven to be effective in the treatment of PD. These drugs mainly include anti-melanin antibodies (Double et al., 2009), α -syn-related drugs (Yanamandra et al., 2011; Horvath et al., 2017; Huang et al., 2019), and GM1 ganglioside-related immune responders (Zappia et al., 2002). All these studies suggest that the pathogenesis and progression of PD may be related to the immune response.

In recent years, with the development of bioinformatics analysis, many significant advances have been made in a wide range of diseases. Several potentially therapeutic drugs (Sun et al., 2016) and key pathways (Zhang et al., 2012) have been identified by bioinformatics in PD. As the genes and the proteins they encode play key roles in physiological activities, it would be useful to study their networks in the disease. In our study, the association between immunity and PD was systematically analyzed using bioinformatics techniques based on the construction of co-expression network, providing a new perspective for the treatment and research in PD.

Materials and methods

Data source

The high-confidence protein-protein interaction (PPI) data with score >10000 were downloaded from the Human Protein Reference Database (HPRD, <http://www.hprd.org/>). The 2747 immune-related genes (HIRGs) were downloaded from the Amigo 2 database (<http://amigo.geneontology.org/amigo>) with immune as the key word. The 2078 PD-related genes (PDRGs) were downloaded from the DisGeNET database (<https://www.disgenet.org/search>). The FPKM expression profiles of GSE49036 and GSE20292 datasets were downloaded from the Gene Expression Omnibus (GEO) database. In the GSE49036 dataset, 8 normal and 14 PD samples were selected for data analysis, and 15 normal and 11 PD samples of GSE20292 dataset were selected for validation analysis. The clinical characteristics of GSE20292 and GSE49036 datasets were shown in Supplementary Table S1.

Construction of an immune- or PD-directed neighbor co-expressed network construction (IOPDNC)

The fragments per kilobase of transcript per million fragments mapped (FPKM) values of gene expression in the GSE49036 dataset were log2-transformed, and the Pearson correlation of the two genes was calculated using the R package psych (version 2.1.9). Then, according to a threshold of |Pearson coefficient value| > 0.7 and FDR < 0.05 to obtain the correlation among genes. Furthermore, based on the correlation between the filtered genes, mapped into the protein interaction network of the HPRD database, the common network was selected. The common network was drawn using Cytoscape software (version 3.8.2) (Shannon et al., 2003). According to the high connectivity score of genes in the common network, the number of four types of genes (PD, immune-PD, immune, and others) was counted. The Venn online tool (<http://www.bioinformatics.com.cn/static/others/jvonn/example.html>) was used to draw a Venn diagram of the correlation of protein-interacting genes, PDRGs, and immune genes. The R package ggplot (version 2.3.3.2) (Villanueva, 2019) was applied to draw a histogram of the four gene categories.

Dissecting PD and immune-associated gene features in network

Based on the above-mentioned high connectivity score of genes in the common network, the core genes of PD-related genes were extracted, including PD and immune-PD genes, and their connected genes. Next, only the PD genes and their direct-acting genes were extracted as core genes. Subsequently, the number of four types of genes (PD, immune-PD, immune, other) were counted, and visualized by ggplot (Version 2.3.3.2) package (Villanueva, 2019). Notably, the immune-PD genes were both immune-related genes and PD-related genes. In the interaction network where the core gene was only PD gene, the correlation of different types of genes was calculated. Finally, the expression data of all genes in the network (the core genes are only the PD genes) were extracted and compared with the core genes. Wilcoxon rank-sum test was applied to compare the coexpression correlation coefficients between different gene groups (immune, immune-PD, PD, and other genes). Subsequently, to investigate the level of interaction between different gene groups with neighbors, cumulative distribution function (CDF) was utilized to assess the degree of the expression correlation for each gene group. The Pearson correlations of genes and the genes that related to the corresponding core genes were calculated, and the R package pheatmap was used to draw correlation heatmaps.

Module cluster analysis and validation of its classification power

Modular clustering analysis was performed on the genes in the IOPDNC to obtain the central core genes using the GraphWeb online website (<https://biit.cs.ut.ee/graphweb/>). Subsequently, the number of core genes (i.e. PD-related genes) was adjusted to six in the GraphWeb database, with the rest set to default, and the

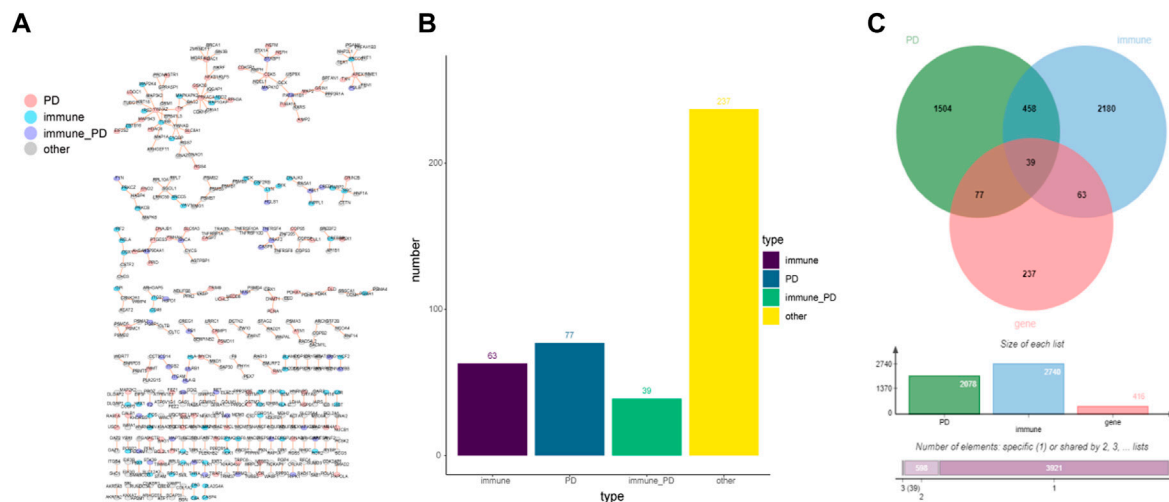


FIGURE 1

The immune- or PD-directed neighbor co-expressed network (IOPDNC network). (A) The global IOPDNC network was constructed to identify a common network by the GSE49036 dataset and HPRD database. (B) The histogram chart of the common genes in the IOPDNC network. (C) The Venn diagram showed the intersections of immune-related genes, PD-related genes, and common network genes.

modules with top five functional score values were selected as PD-related gene modules. Then, the R package ConsensusClusterPlus (version 1.54.0) (Wilkerson and Hayes, 2010) was used to perform consistent clustering analysis on the genes with the top 5 functional scores in the module, and the appropriate K value was selected based on the clustering results. In the external validation set GSE20292 dataset, the expression levels of modular genes were extracted in the same way and the accuracy of our screening of modular genes was validated against the same consistent clustering criteria.

Functional enrichment analysis and pathway enrichment analysis of modular genes

The Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses of different module genes were analyzed using the R package ClusterProfiler package (version 4.0.2) (Wu et al., 2021). According to the significance threshold $p < 0.05$ and count value, the enrichment analysis of each module was carried out separately, and the ggplot (version 2.3.3.2) was used for plotting. According to the website of Pathview (<https://pathview.uncc.edu/>), the immune-related pathway hsa04650 was selected to visualize the most immune-related pathways in the module.

Differential analysis of immune cell infiltration by modular genes

Based on the 24 immune cell sets, the single sample Gene Set Enrichment Analysis (ssGSEA) algorithm was used to calculate the immune cell infiltration of the PD and normal samples, and the rank sum test was used to analyze the immune differences between the PD and normal samples of cell infiltration.

Blood samples correlation

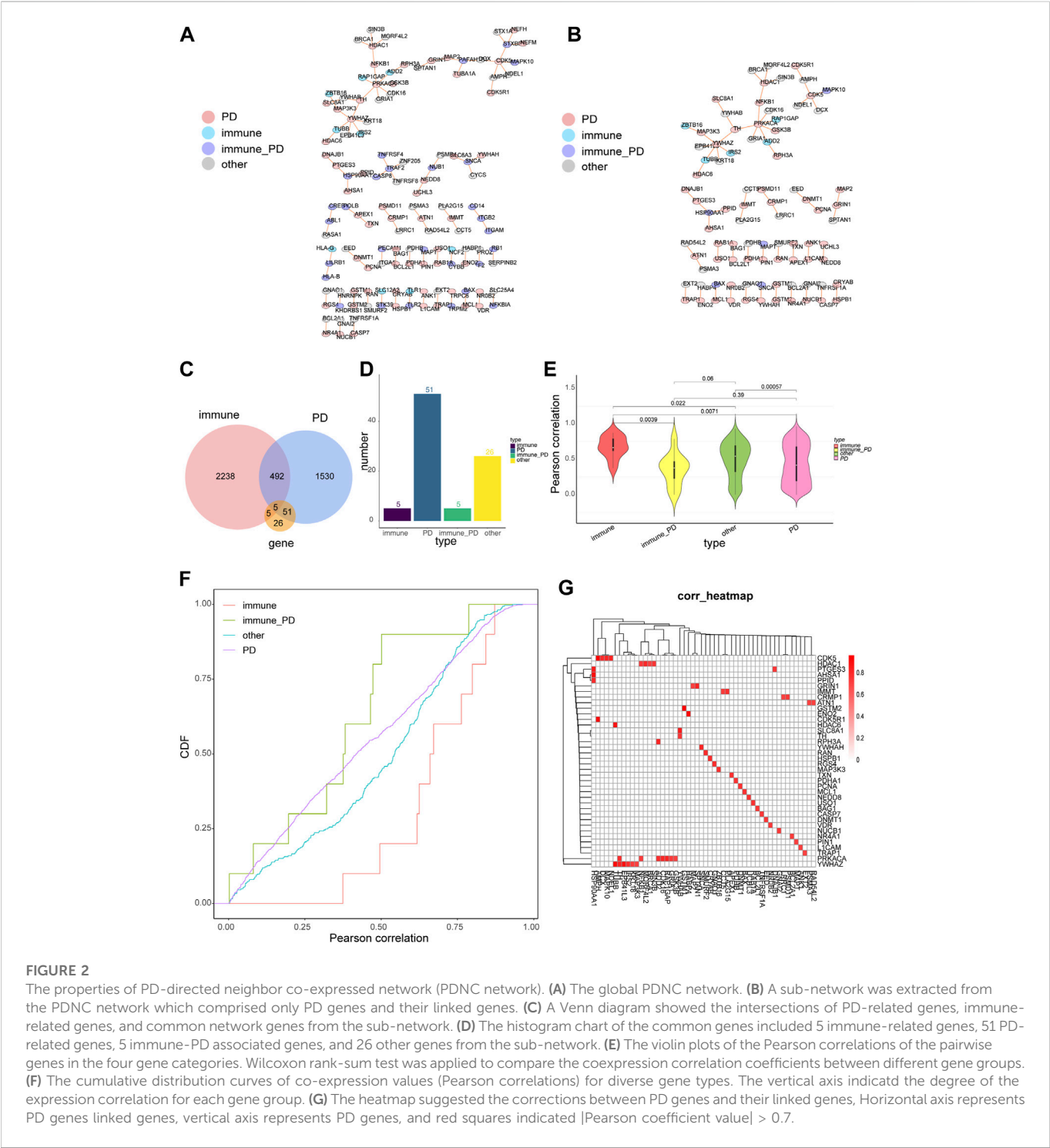
Peripheral blood mononuclear cells (PBMC) samples from eight normal samples and eight PD patients was collected using vacuum blood tubes containing EDTA anticoagulant in accordance with clinical blood collection techniques. Each PBMC sample was gently shaken repeatedly and loaded into a 4°C thermostat and transferred to the laboratory for subsequent manipulation according to biosafety requirements.

The quantitative reverse transcription polymerase chain reaction (qRT-PCR) analysis

The total RNA of 16 PBMC samples (8 normal samples and 8 PD samples) was extracted to verify the results of the bioinformatics analysis. The top 1 gene of each module (module 1, module 2, module 3, module 4, module 5, and module all) was selected for qRT-PCR experiments. The total RNA of 16 samples was extracted with TRIzol Reagent (Life Technologies-Invitrogen, Carlsbad, CA, United States). Then, these total RNA were reverse transcription into cDNA with the SureScript-First-strand-cDNA-synthesis-kit (Genecopoeia, Guangzhou, China) prior to qRT-PCR. The primers of these genes for qPCR were as follows:

PSMB7-For:CATGGGTTCTGGCTCCTTGG; PSMB7-Rev:CTGGTCCCCTTCTTGTGTGGG; GRIN1-For:CAAGAAGGAGTGAATGGGATG; GRIN1-Rev:GCTCGTTGTTATGGTTAGCGG; NME1-For:CAACCTGCAGACTCCAAGC; NME1-Rev:GGTGAAACCACAAGCCGATC; SIN3B-For:ACCCTGCCACCTACAACGG; SIN3B-Rev:TTGTCA GAGGCGAC TGTATGTTTA; HABP4-For:GAGGCAGGCAGACTTCACAG HABP4-Rev:CGAACTCCACATCCACCCAT; STX1A-For:CAATGTGGAACACGCGGTAG; STX1A-Rev: ACAGTGGAGGCGATGACGAT.

The expression was uniformized to the internal reference GAPDH and computed employing the $2^{-\Delta\Delta Ct}$ method.



Western blotting

RIPA Lysis Bufferb (Servucebui) containing a protease inhibitor (Servucebio) was utilized to obtain the protein from tissues, subsequently, immunoblotting was performed. The bicinchoninic acid (BCA) quantification kit was applied to determine protein concentration of the cell lysates. The protein samples were loaded and separated by SDS-PAGE and shifted to PVDF membranes

(Millpore, Sigma). Membranes were incubated with specific primary antibodies against PSMB7 (cst), GRIN1 (Affinity), NME1 (Affinity), SIN3B (Affinity), HABP4 (Proteintech), STX1A (BOSTER), and β -Actin (Proteintech) after blocked with 5% nonfat dry milk at 4°C. Furthermore, secondary antibodies (IgG) were incubated at room temperature for 60 min and visualized using an ECL system. The dilution factors of the primary and secondary antibodies were shown in [Supplementary Table S2](#).

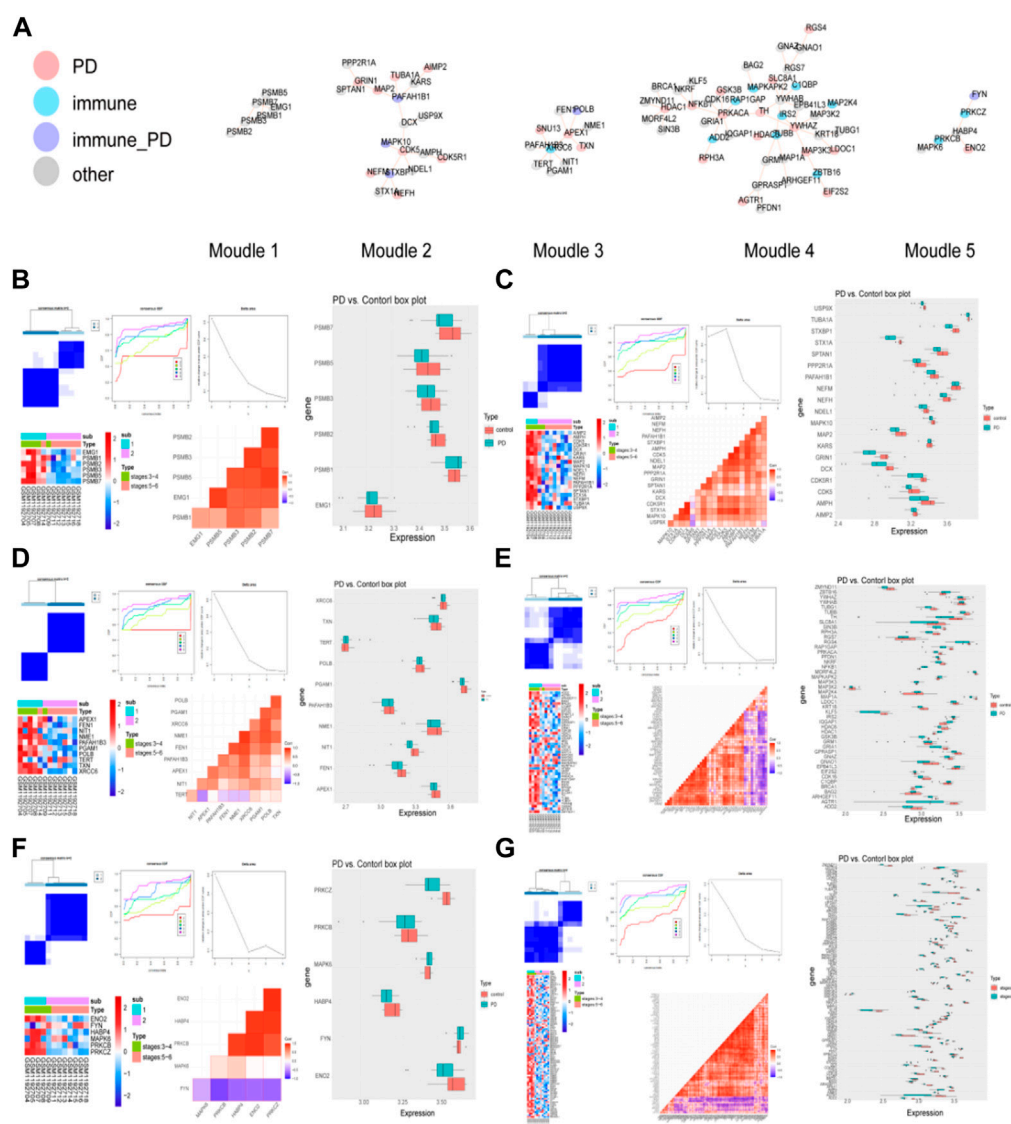


FIGURE 3

Detection of PD-related clusters and validation of their classification power. (A) Important clusters of modules were generated in the IOPDNC network. (B–G) The consensus cluster heatmap, cumulative distribution function (CDF) plot, delta area plot, gene expression heatmap, gene correlation heatmap, and gene expression box plot of modules including module 1 (B), module 2 (C), module 3 (D), module 4 (E), module 5 (F), and the common genes (G).

Results

The construction of IOPDNC

A total of 1022077 gene relationship pairs were selected from the GSE49036 dataset. Then, the common network with 416 nodes and 281 edges was obtained from the GSE49036 dataset and HPRD database (Figure 1A, Supplementary Table S3). Among the common network genes of Figure 1A, the number of immune genes was 63; the number of PD genes was 77; the number of immune-PD genes was 39; the number of other genes were 237 (Figure 1B). Genes with high protein interaction connectivity scores were selected, and a Venn diagram of protein-interacting genes associated with PD and immune genes was drawn. There were 39 intersection genes between 2078 PD genes, 2740 immune genes, and 416 common network genes (Figure 1C). The results

indicated that immune-related genes played a vital role in the IOPDNC network. Together, these results suggested that immune-related genes might be important contributors for PD.

Dissecting PD and immune-associated gene features in the network

The core genes of PD, immune-PD genes, and their linked genes were extracted from the network of 3.1. Then, a network of these genes was constructed with 130 nodes and 91 nodes (Figure 2A, Supplementary Table S4). The genes whose core genes only were PD and the genes that were directly affected by PD were extracted to construct a network that including 87 nodes and 60 edges (Figure 2B, Supplementary Table S5). Totally five common genes were detected between 2078 PD genes, 2740 immune genes, and 87 core genes

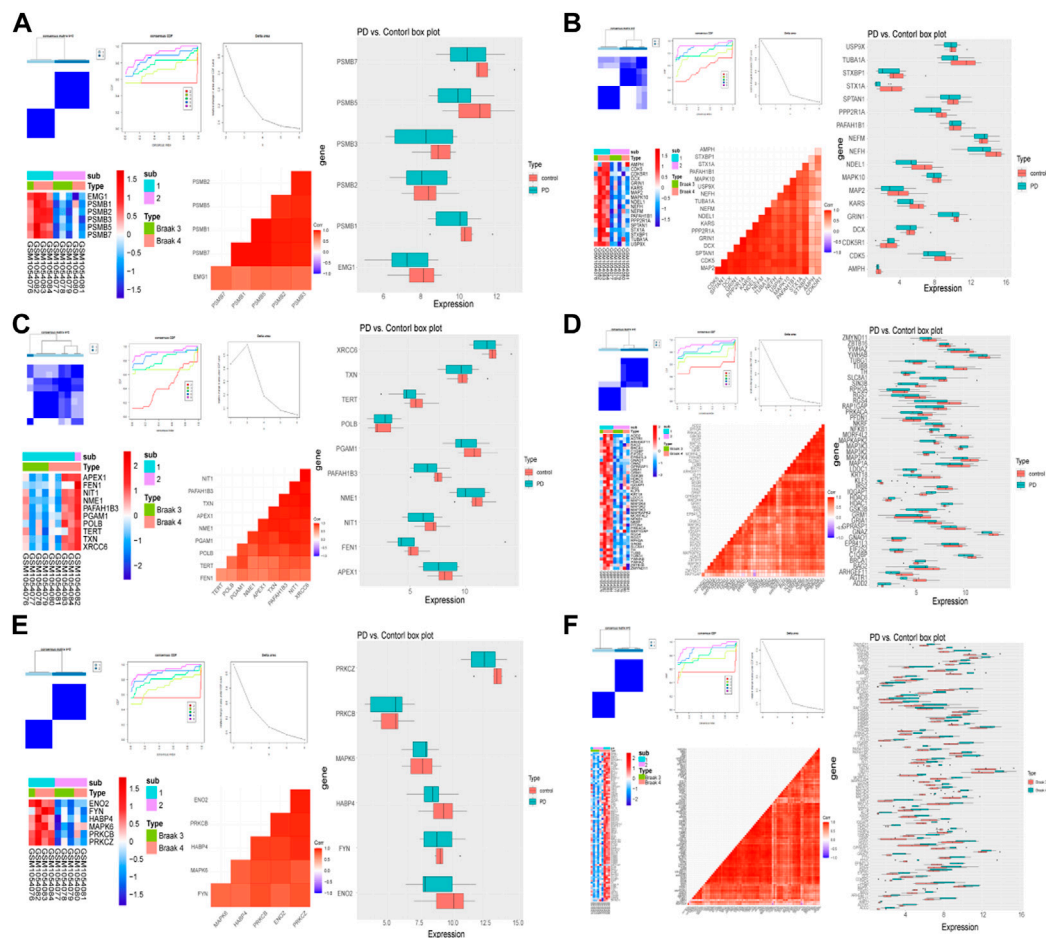


FIGURE 4

Validation of the classification power by GSE20292 data set. The consensus cluster heatmap, cumulative distribution function (CDF) plot, delta area plot, gene expression heatmap, gene correlation heatmap, and gene expression box plot of six modules including module 1 (A), module 2 (B), module 3 (C), module 4 (D), module 5 (E), and common genes (F).

(Figure 2C). Furthermore, the number of genes in the four gene categories (PD, immune-PD, immune, and other genes) that the core genes only were PD of the PD gene interaction network were accounted (Figure 2D). Among the protein-interaction network, the number of immune genes was 5; the number of PD genes was 51; the number of immune-PD genes was 5; the number of other genes was 26 (Figure 2D). Moreover, there were significant differently expressed correlations between different gene groups, except immune-PD and PD genes ($p = 0.39$) (Figure 2E). In the PD genes and its linked genes, the correlation of immune genes, immune-PD genes and other genes was significant (Figure 2F). The correlation between PD genes and its linked genes was significant with |Pearson coefficient value| > 0.7 and FDR < 0.05 (Figure 2G). Totally, the results showed that there were topological interactions and expression patterns among the correlations between PD- and immune-related genes.

Module cluster analysis and validation of its classification power

The modules with top 5 functional scores were selected (Figure 3A, Supplementary Table S6). In the all of 5 modules, 25 PD genes,

11 immune genes, 5 immune-PD genes, and 46 other genes were contained.

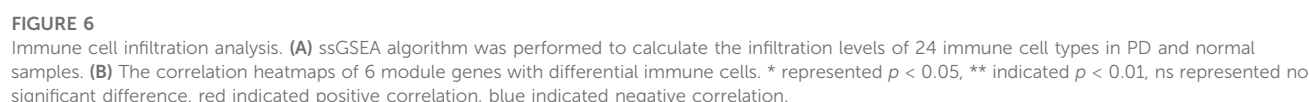
In these 6 modules, the genes were divided into 2 clusters when the $K = 4$, and the expression of these module genes were higher in cluster 1 (Figures 3B–G). In these modules, the expression levels of module genes except PSMB1, KARS, TERT, ZBTB16, NFKB1, MAPKAPK2, MAP3K3, MAP2K4, IRS2, IQGAP1, HDAC1, EPB41L3, C1QBP, BRCA1, BAG2, FYN, MAPK6, MAP3K2, and FNY in the PD group were higher than that in the control group.

The same method was used to validate the accuracy of the screened module genes in the external validation set GSE20292 data set (Figure 4). These genes have better representation and can screen out patients at different stages. These results demonstrated that the model genes could distinguish PD and control samples well.

Functional enrichment analysis of modular genes

Module 1 genes were enriched in 104 GO BPs (including 12 GO CCs, 7 GO MFs) and 8 KEGG pathways, and these GO terms and KEGG pathways were mainly related to various metabolic





the brain parenchyma of mouse models of PD (Earls et al., 2020). NK cells can reduce synuclein burden *in vitro*, and systemic depletion of NK cells in a preclinical mouse model of PD results in increased pathological α -syn burden in numerous brain regions, including the striatum, SNpc, and brainstem (Peng et al., 2019). In the natural killer (NK) cell-mediated cytotoxicity pathway, Fyn, Vav, and PKC were significantly enriched. Among them, Fyn is a tyrosine phosphotransferase of Src family non-receptor kinases, which is mainly related to immune regulation, cell proliferation, and brain development (Guglietti et al., 2021). In previous studies, Fyn was

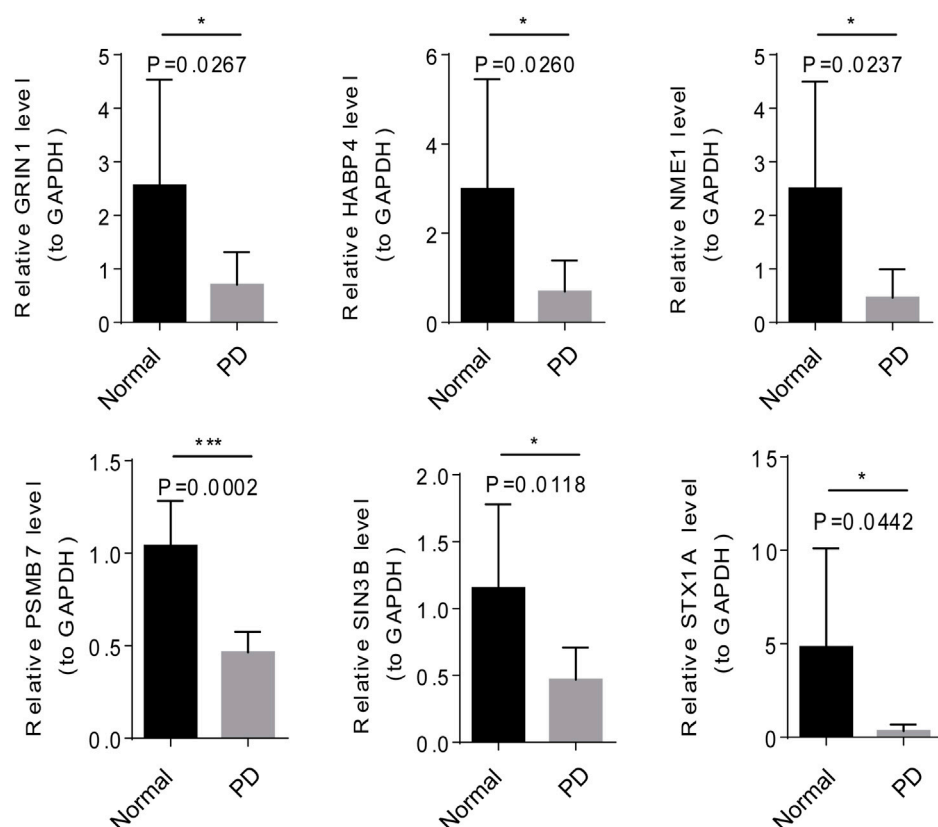


FIGURE 7
Verification of the mRNA expression of six modular genes by qRT-PCR and Western blot.

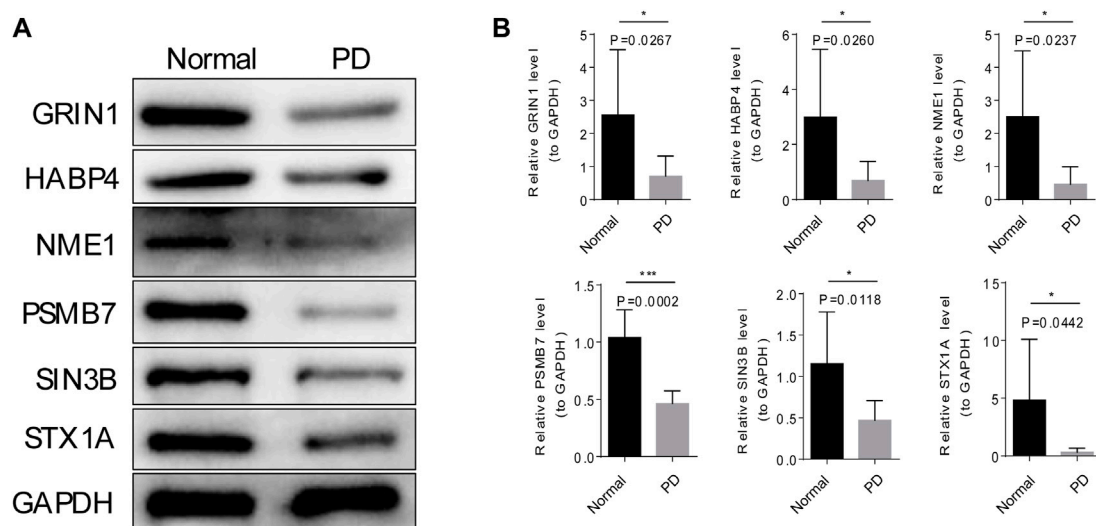


FIGURE 8
Verification of the protein expression of six modular genes by Western blot. (A) The grayscale values six of modular genes in PD and normal samples by Western blot. (B) The mages of six modular genes in PD and normal samples by Western blot.

confirmed to be a major upstream regulator of proinflammatory signaling pathway involving BDNF/TrkB, PKC δ , MAPK, AMPK, NF- κ B, Nrf2, and NMDAR axis. Fyn is also being used as a potential signaling node for the development of novel anti-

neuroinflammatory drug candidates for the treatment of PD and other related neurodegenerative diseases (Peng et al., 2019). For example, saracatinib, a non-selective Fyn inhibitor, has been tested in clinical trials to treat PD (Angelopoulou et al., 2021). The protein

kinase C (PKC) family is a phospholipid-dependent serine/threonine kinase (Kishimoto et al., 1980). The protein kinase family consists of more than 15 subgroups and 500 kinases whose expression affects the progression of various diseases, including neurodegenerative diseases (Zisopoulou et al., 2013; Jha et al., 2015; Crawley et al., 2017). PKC α inhibits the expression of peroxisome proliferator-activated receptor C coactivator 1 (PGC-1) by inducing miR-129-2 in neural tube defect (NTD) embryonic mouse models, and the overexpression of PGC-1 protects neurons from mitochondrial dysfunction under oxidative stress in PD (Mudò et al., 2012).

There is growing evidence linking the immune system to neuronal death and the pathogenesis of PD. Previous studies have shown that detection of immune cell components in the blood can identify the early stages of PD progression, leading to earlier detection and confirmation of PD (Farmen et al., 2021). Activated microglia (brain's resident immune cells) correlate directly with the clinical and pathological severity of PD (Lanskey et al., 2018). Through immune infiltration analysis executed by the ssGSEA algorithm, we discovered that aDC and Th2 cells were significantly decreased in PD samples, and eosinophils and neutrophils cells were significantly upregulated in PD samples. But there is still a gap in how these cells play a role in the progression of cup-like lesions in Parkinson's disease, however, this provided a basis and direction to further unravel the immune-related mechanisms of PD.

The top1 gene was selected from the 6 modules for qRT-PCR and Western blot validation. NME1 was a protein with serine/threonine specific protein kinase activity (Yu et al., 2021). NME1 has been shown to play an important role in neuronal growth by increasing mitochondrial respiration and preventing α -synuclein and LRRK2-induced degeneration. In PD treatment, NME1 can promote neurite growth in PD cell models and restore damaged mitochondrial respiration and cellular pathways (Anantha et al., 2022). GRIN1 (encoding NMDAR subunit n-methyl-D-aspartate 1) gene has been shown to be closely associated with neurodevelopmental disorders (Platzer et al., 1993), and its polymorphism has also been demonstrated as a potential biomarker for reducing the risk of PD in previous studies (Wu et al., 2010). The module genes such as the expression of NME1, SIN3B, HABP4, STX1A, SIN3B, HABP4, and STX1A could distinguish PD and normal samples, indicating these genes may become promising candidate genes for PD.

In conclusion, these results indicated strong correlations between immune- and PD-related genes not only in terms of network structures but also in expression patterns. According to the differential expression and functional enrichment analyses, some immune-related genes may have the potential as diagnostic and therapeutic biomarkers for PD. However, there still have two main limitations in this study. Firstly, this study was a retrospective study based on a public database with limited sample sizes. Second, the important genes and mechanisms in this study need further experimental studies to be validated. Altogether, we have revealed the association between immunity and PD through systematic network studies and bioinformatics approaches, providing a theoretical basis for further studies on the pathogenesis of PD and clinical therapeutic targets.

Conclusion

In summary, all the results presented here indicate a strong association between immune and PD-related genes not only in

network structure but also in expression patterns. After analyzing the expression patterns and functions of the genes in the five modules, we believe that these genes have potential as molecular diagnostic markers.

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 Ethical Committee of Yan'An Hospital of Kunming Medical University. The patients/participants provided their written informed consent to participate in this study.

Author contributions

GW conceived and designed the research. XD obtained and analyzed the data, and drafted the manuscript. YL, QL, and WL participated in the data analysis and determination of analytical method. All authors read, revised manuscript and approved the manuscript for publication.

Funding

This work was supported by the health research project of Kunming Health Commission (2022-03-07-001).

Conflict of interest

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

The reviewer XY declared a shared parent affiliation with the author(s) to the handling editor at the time of review.

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.2023.1090382/full#supplementary-material>

References

- Anantha, J., Goulding, S. R., Tuboly, E., O'Mahony, A. G., Moloney, G. M., Lomansey, G., et al. (2022). NME1 protects against neurotoxin- α -synuclein- and LRRK2-induced neurite degeneration in cell models of Parkinson's disease. *Mol. Neurobiol.* 59 (1), 61–76. doi:10.1007/s12035-021-02569-6
- Angelopoulou, E., Paudel, Y. N., Julian, T., Shaikh, M. F., and Piperi, C. (2021). Pivotal role of Fyn kinase in Parkinson's disease and levodopa-induced dyskinesia: A novel therapeutic target? *Mol. Neurobiol.* 58 (4), 1372–1391. doi:10.1007/s12035-020-02201-z
- Cebrián, C., Zucca, F. A., Mauri, P., Steinbeck, J. A., Studer, L., Scherzer, C. R., et al. (2014). MHC-I expression renders catecholaminergic neurons susceptible to T-cell-mediated degeneration. *Nat. Commun.* 5, 3633. doi:10.1038/ncomms4633
- Crawley, O., Giles, A. C., Desbois, M., Kashyap, S., Birnbaum, R., and Grill, B. (2017). A MIG-15/JNK-1 MAP kinase cascade opposes Rpm-1 signaling in synapse formation and learning. *PLoS Genet.* 13 (12), e1007095. doi:10.1371/journal.pgen.1007095
- Del Rey, N. L., Quiroga-Varela, A., Garbayo, E., Carballo-Carbajal, I., Fernandez-Santiago, R., Monje, M. H. G., et al. (2018). Advances in Parkinson's disease: 200 Years later. *Front. Neuroanat.* 12, 113. doi:10.3389/fnana.2018.00113
- Double, K. L., Rowe, D. B., Carew-Jones, F. M., Hayes, M. Chan, D. K. Y., Blackie, J., et al. (2009). Anti-melanin antibodies are increased in sera in Parkinson's disease. *Exp. Neurol.* 217 (2), 297–301. doi:10.1016/j.expneurol.2009.03.002
- Earls, R. H., Menees, K. B., Chung, J., Gutekunst, C. A., Lee, H. J., Hazim, M. G., et al. (2020). NK cells clear α -synuclein and the depletion of NK cells exacerbates synuclein pathology in a mouse model of α -synucleinopathy. *Proc. Natl. Acad. Sci. U. S. A.* 117, 1762–1771. doi:10.1073/pnas.1909110117
- Elsworth, J. D. (2020). Parkinson's disease treatment: Past, present, and future. *J. Neural Transm. (Vienna)* 127 (5), 785–791. doi:10.1007/s00702-020-02167-1
- Farmen, K., Nissen, S. K., Stokholm, M. G., Iranzo, A., Ostergaard, K., Serradell, M., et al. (2021). Monocyte markers correlate with immune and neuronal brain changes in REM sleep behavior disorder. *Proc. Natl. Acad. Sci.* 118 (10), e2020858118. doi:10.1073/pnas.2020858118
- Guan, Q., Liu, W., Mu, K., Hu, Q., Xie, J., Cheng, L., et al. (2022). Single-cell RNA sequencing of CSF reveals neuroprotective RAC1+ NK cells in Parkinson's disease. *Front. Immunol.* 13, 992505. doi:10.3389/fimmu.2022.992505
- Guglietti, B., Sivasankar, S., Mustafa, S., Corrigan, F., and Collins-Praino, L. E. (2021). Fyn kinase activity and its role in neurodegenerative disease pathology: A potential universal target? *Mol. Neurobiol.* 58 (11), 5986–6005. doi:10.1007/s12035-021-02518-3
- Horvath, I., Iashchishyn, I. A., Forsgren, L., and Morozova-Roche, L. A. (2017). Immunochemical detection of α -synuclein autoantibodies in Parkinson's disease: Correlation between plasma and cerebrospinal fluid levels. *ACS Chem. Neurosci.* 8 (6), 1170–1176. doi:10.1021/acschemneuro.7b00063
- Huang, Y. R., Xie, X. X., Ji, M., Yu, X. L., Zhu, J., Zhang, L. X., et al. (2019). Naturally occurring autoantibodies against α -synuclein rescues memory and motor deficits and attenuates α -synuclein pathology in mouse model of Parkinson's disease. *Neurobiol. Dis.* 124, 202–217. doi:10.1016/j.nbd.2018.11.024
- Jha, S. K., Jha, N. K., Kar, R., Ambasta, R. K., and Kumar, P. (2015). p38 MAPK and PI3K/AKT signalling cascades in Parkinson's disease. *Int. J. Mol. Cell Med.* 4 (2), 67–86.
- Kalia, L. V., and Lang, A. E. (2016). Parkinson disease in 2015: Evolving basic, pathological and clinical concepts in PD. *Nat. Rev. Neurol.* 12 (2), 65–66. doi:10.1038/nrneuro.2015.249
- Kishimoto, A., Takai, Y., Mori, T., Kikkawa, U., and Nishizuka, Y. (1980). Activation of calcium and phospholipid-dependent protein kinase by diacylglycerol, its possible relation to phosphatidylinositol turnover. *J. Biol. Chem.* 255 (6), 2273–2276. doi:10.1016/s0021-9258(19)85886-6
- Kline, E. M., Houser, M. C., Herrick, M. K., Seibler, P., Klein, C., West, A., et al. (2021). Genetic and environmental factors in Parkinson's disease converge on immune function and inflammation. *Mov. Disord.* 36 (1), 25–36. doi:10.1002/mds.28411
- Krashia, P., Cordella, A., Nobili, A., La Barbera, L., Federici, M., Leuti, A., et al. (2019). Author Correction: Blunting neuroinflammation with resolvin D1 prevents early pathology in a rat model of Parkinson's disease. *Nat. Commun.* 10 (1), 4725. doi:10.1038/s41467-019-12538-2
- Lanskey, J. H., McColgan, P., Schrag, A. E., Acosta-Cabrero, J., Rees, G., Morris, H. R., et al. (2018). Can neuroimaging predict dementia in Parkinson's disease? *Brain* 141 (9), 2545–2560. doi:10.1093/brain/awy211
- Marras, C., Beck, J. C., Bower, J. H., RobErt, E., Ritz, B., Ross, G. W., et al. (2018). Prevalence of Parkinson's disease across north America. *NPJ Park. Dis.* 4, 21. doi:10.1038/s41531-018-0058-0
- Mudò, G., Mäkelä, J., Di Liberto, V., Tselykh, T. V., Olivieri, M., Piepponen, P., et al. (2012). Transgenic expression and activation of PGC-1 α protect dopaminergic neurons in the MPTP mouse model of Parkinson's disease. *Cell Mol. Life Sci.* 69 (7), 1153–1165. doi:10.1007/s00018-011-0850-z
- Pajares, M., Manda, G., Bosca, L., and Cuadrado, A. (2020). Inflammation in Parkinson's disease: Mechanisms and therapeutic implications. *Cells* 9 (7), 1687. doi:10.3390/cells9071687
- Peng, T., Liu, X., Wang, J., Liu, Y., Fu, Z., Ma, X., et al. (2019). Long noncoding RNA HAGLROS regulates apoptosis and autophagy in Parkinson's disease via regulating miR-100/ATG10 axis and PI3K/Akt/mTOR pathway activation. *Artif. Cells Nanomed Biotechnol.* 47 (1), 2764–2774. doi:10.1080/21691401.2019.1636805
- Platzer, K., Lemke, J. R., Everman, D. B., Mirzaa, G. M., Pagon, R. G., and Wallace, S. E. (1993). Grin1-Related neurodevelopmental disorder [M]/Adam M P, Seattle. GeneReviews is a registered trademark of the University of Washington, Seattle. All rights reserved.
- 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
- Sun, A. G., Lin, A. Q., Huang, S. Y., Huo, D., and Cong, C. H. (2016). Identification of potential drugs for Parkinson's disease based on a sub-pathway method. *Int. J. Neurosci.* 126 (4), 318–325. doi:10.3109/00207454.2014.986673
- Tarsy, D. (2012). Treatment of Parkinson disease: A 64-year-old man with motor complications of advanced Parkinson disease. *Jama* 307 (21), 2305–2314. doi:10.1001/jama.2012.4829
- Villanueva, C. (2019). *ggplot2: Elegant graphics for data analysis*. 2nd ed. 17. Measurement: Interdisciplinary Research and Perspectives.
- 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, S. L., Wang, W. F., Shyu, H. Y., Ho, Y. J., Shieh, J. C., Fu, Y. P., et al. (2010). Association analysis of GRIN1 and GRIN2B polymorphisms and Parkinson's disease in a hospital-based case-control study. *Neurosci. Lett.* 478 (2), 61–65. doi:10.1016/j.neulet.2010.04.063
- 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. *Innov. (Camb)* 2 (3), 100141. doi:10.1016/j.xinn.2021.100141
- Yanamandra, K., Gruden, M. A., Casate, V., Meskys, R., Forsgren, L., and Morozova-Roche, L. A. (2011). α -synuclein reactive antibodies as diagnostic biomarkers in blood sera of Parkinson's disease patients. *PLoS One* 6 (4), e18513. doi:10.1371/journal.pone.0018513
- Yu, B. Y. K., Tossounian, M. A., Hristov, S. D., Lawrence, R., Arora, P., Tsuchiya, Y., et al. (2021). Regulation of metastasis suppressor NME1 by a key metabolic cofactor coenzyme A. *Redox Biol.* 44, 101978. doi:10.1016/j.redox.2021.101978
- Zappia, M., Crescibene, L., Bosco, D., Arabia, G., Nicoletti, G., BagAla, A., et al. (2002). Anti-GM1 ganglioside antibodies in Parkinson's disease. *Acta Neurol. Scand.* 106 (1), 54–57. doi:10.1034/j.1600-0404.2002.01240.x
- Zhang, B., Xia, C., Lin, Q., and Huang, J. (2012). Identification of key pathways and transcription factors related to Parkinson disease in genome wide. *Mol. Biol. Rep.* 39 (12), 10881–10887. doi:10.1007/s11033-012-1985-1
- Zisopoulou, S., Asimaki, O., Leondaritis, G., Vasilaki, A., Sakellaris, N., Pitsikas, N., et al. (2013). PKC-epsilon activation is required for recognition memory in the rat. *Behav. Brain Res.* 253, 280–289. doi:10.1016/j.bbr.2013.07.036



OPEN ACCESS

EDITED BY

Thayne Kowalski,
Centro Universitário Cesuca, Brazil

REVIEWED BY

Julia do Amaral Gomes,
Ministério da Saúde, Brazil
Giovanna Giudicelli,
Federal University of Rio Grande do Sul,
Brazil

*CORRESPONDENCE

Hua Chen,
✉ 936483206@qq.com
Di Liao,
✉ liaodi1118@csu.edu.cn

[†]These authors have contributed equally
to this work

SPECIALTY SECTION

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

RECEIVED 27 July 2022

ACCEPTED 13 March 2023

PUBLISHED 31 March 2023

CITATION

Yang Z, He M, Zhang Q, Li S, Chen H and
Liao D (2023), Exploring the bi-directional
relationship and shared genes between
depression and stroke via NHANES and
bioinformatic analysis.
Front. Genet. 14:1004457.
doi: 10.3389/fgene.2023.1004457

COPYRIGHT

© 2023 Yang, He, Zhang, Li, Chen and
Liao. This is an open-access article
distributed under the terms of the
[Creative Commons Attribution License](#)
(CC BY). The use, distribution or
reproduction in other forums is
permitted, provided the original author(s)
and the copyright owner(s) are credited
and that the original publication in this
journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

Exploring the bi-directional relationship and shared genes between depression and stroke via NHANES and bioinformatic analysis

Zhanghuan Yang^{1†}, Maokun He^{2†}, Qian Zhang^{3,4}, Shifu Li^{3,4},
Hua Chen^{5*} and Di Liao^{6*}

¹Department of Oncology, Xiangya Cancer Center, Xiangya Hospital, Central South University, Changsha, Hunan, China, ²Hainan Medical University, Haikou, China, ³Department of Neurosurgery, Xiangya Hospital, Central South University, Changsha, Hunan, China, ⁴National Clinical Research Center for Geriatric Disorders, Central South University, Changsha, Hunan, China, ⁵Department of Neurosurgery, The First people's Hospital of Changde, Changde, China, ⁶Department of Neurology, Xiangya Hospital, Central South University, Changsha, Hunan, China

Background: Stroke and depression are the two most common causes of disability worldwide. Growing evidence suggests a bi-directional relationship between stroke and depression, whereas the molecular mechanisms underlying stroke and depression are not well understood. The objectives of this study were to identify hub genes and biological pathways related to the pathogenesis of ischemic stroke (IS) and major depressive disorder (MDD) and to evaluate the infiltration of immune cells in both disorders.

Methods: Participants from the United States National Health and Nutritional Examination Survey (NHANES) 2005–2018 were included to evaluate the association between stroke and MDD. Two differentially expressed genes (DEGs) sets extracted from GSE98793 and GSE16561 datasets were intersected to generate common DEGs, which were further screened out in cytoHubba to identify hub genes. GO, KEGG, Metascape, GeneMANIA, NetworkAnalyst, and DGIdb were used for functional enrichment, pathway analysis, regulatory network analysis, and candidate drugs analysis. ssGSEA algorithm was used to analyze the immune infiltration.

Results: Among the 29706 participants from NHANES 2005–2018, stroke was significantly associated with MDD (OR = 2.79, 95% CI: 2.26–3.43, $p < 0.0001$). A total of 41 common upregulated genes and eight common downregulated genes were finally identified between IS and MDD. Enrichment analysis revealed that the shared genes were mainly involved in immune response and immune-related pathways. A protein-protein interaction (PPI) was constructed, from which ten (CD163, AEG1, IRAK3, S100A12, HP, PGLYRP1, CEACAM8, MPO, LCN2, and DEFA4) were screened. In addition, gene-miRNAs, transcription factor-gene interactions, and protein-drug interactions coregulatory networks with hub genes were also identified. Finally, we observed that the innate immunity was activated while acquired immunity was suppressed in both disorders.

Conclusion: We successfully identified the ten hub shared genes linking the IS and MDD and constructed the regulatory networks for them that could serve as novel targeted therapy for the comorbidities.

KEYWORDS

ischemic stroke, major depressive disorder, bioinformatics, shared genes, immune infiltration, NHANES

Introduction

Stroke is the leading cause of death and disability, leading to significant economic losses as a result of its functional impairments (Meschia et al., 2014). Every year more than 795,000 people in the United States (United States) have a stroke, of which 610,000 are first-time strokes, whereas 185,000 patients have previously had a stroke. The vast majority of stroke cases occur as a result of two specific types of stroke (hemorrhagic and ischemic strokes). United States strokes are dominated by ischemic strokes, which account for 87% of all strokes (Barthels and Das, 2020). The prevalence of depression is growing among the general population, typically characterized by anhedonia and the inability to experience pleasure. A variety of somatic symptoms (psychological disturbance, fatigue, and weight fluctuations) as well as cognitive symptoms (poor concentration and negative cognitions) can accompany depression (Bucciarelli et al., 2020). Depression is sub-categorized into major depressive disorder (MDD) and dysthymia. Epidemiologic data indicates that the average lifetime and 12-month prevalence estimates of MDD are 14.6% and 5.5% in high-income and 11.1% and 5.9% in the low-to middle-income countries (Bromet et al., 2011).

Stroke patients are more likely to suffer from depression than the general population. Growing evidence suggests a bi-directional relationship between stroke and depression: 1) depression is associated with an increased risk of stroke (Pan et al., 2011; Wassertheil-Smoller et al., 2018; Harshfield et al., 2020), and 2) depression is particularly prevalent among stroke survivors (Robinson and Jorge, 2016; Das and Rajanikant, 2018). The prevalence of post-stroke depression (PSD) is estimated to be 29% at any time point up to 5 years following a stroke (Hackett and Pickles, 2014). However, the mechanisms underlying the association between depression and stroke are poorly investigated. Biological factors such as alterations in ascending monoamine systems, neuroplasticity, and glutamate neurotransmission and an increasing of pro-inflammatory cytokines were proposed to explain the mechanisms of PSD (Robinson and Jorge, 2016). Moreover, multiple mechanisms may play roles in depression contributing to stroke. First, Smoking, obesity (Ho et al., 2008), and poor health behaviors (i.e., poor diet, physical inactivity, and smoking) (Strine et al., 2008) may increase stroke risk in depression patients. Second, other major comorbidities, such as diabetes (Wesołowska et al., 2018), atherosclerosis (Joynt et al., 2003), and hypertension (Patten et al., 2009), accompanied by depression, are major risk factors for stroke. Finally, the use of antidepressant medication may potentially contribute to the occurrence of stroke events.

Apart from the above-mentioned mechanisms, genetic factors are likely involved in the pathogenesis of depression and stroke. Increased risk for depression in first-degree relatives of depression probands was observed with an estimated odds ratio of 2.84 from a meta-analysis of the highest-quality family studies (Sullivan et al., 2000). The heritability of MDD has been found to be greater in women (42%) than in men (29%) in a Swedish national twin study

(Kendler et al., 2006). There are multiple risk factor genes that were thought to participate in the pathogenesis of depression with extremely complex, polygenic, and epistatic inheritance patterns (Zhao et al., 2019a). There is significant evidence that stroke has a hereditary component based on studies of twins, siblings, and families (Humphries and Morgan, 2004). Heritability for all IS is estimated to be 37.9% (Bevan et al., 2012). The heritability of stroke subtypes varies markedly, with 40.3% for large vessels and 32.6% for cardioembolics but lower for cardioembolic small vessels (16.1%). The genetic involvement in the pathogenesis of both stroke and depression as well as the comorbidity frequency is not yet fully established or whether common overlapping genes and biological mechanisms are subserving both disorders.

A common transcription feature may provide new insights into the pathogenesis of depression and stroke. This study aims to identify hub genes and biological pathways related to the pathogenesis of IS and MDD. Furthermore, as increasing evidence points to the involvement of an immune response in both disorders (Beurel et al., 2020; Iadecola et al., 2020), we evaluate the immune cell infiltration and identify the common immune cells.

Materials and methods

Dataset collection and processing

The data used in the present work was downloaded from the National Health and Nutrition Examination Survey (NHANES) (<https://www.cdc.gov/nchs/index.htm>) and the Gene Expression Omnibus (GEO) database (<https://www.ncbi.nlm.nih.gov/geo/>) based on a microarray or RNA-seq dataset of major depressive disorder (MDD) and ischemic stroke (IS). The NHANES is a research project aimed to assess the health and nutritional status of adults and children in the United States, combining interviews and physical examinations to provide vital and health statistics. The GSE98793 microarray profile included 128 MDD whole blood samples and 64 health samples at the platform of GPL570 Affymetrix U133_Plus2.0 Genechips. The effect of two batches in the GSE98793 dataset were removed by applying removed BatchEffect function of the limma package (Ritchie et al., 2015). The GSE76826 dataset is a microarray profile at the platform of GPL17077 Agilent-039494 SurePrint G3 Human GE v2 8 × 60 K Microarray 039381. The GSE16561 microarray profile contained whole blood from 39 IS patients and was compared with 24 healthy control subjects, measured using a GPL6883 Illumina HumanRef-8 v3.0 expression beadchip. The GSE122709 dataset (including 10 peripheral blood mononuclear cells samples of IS patients and five controls) is a RNA-sequencing dataset and measured at GPL20795 HiSeq X Ten. When multiple probes were matched with one gene, the probe with the highest expression values was annotated in the homologous gene symbol based on the annotation information on the platform.

NHANES

Data of 70190 participants were available in NHANES 2005–2018. Age, sex, race or ethnicity, education level, poverty, marital status, smoking status, stroke, body mass index (BMI), waist circumference and diabetes was included as variables in the analysis. Depression was measured using the Patient Health Questionnaire (PHQ-9). Participants with PHQ-9 total scores ≥ 10 were considered as having MDD. After excluding participants with missing data, 29706 participants were included in our analysis. Continuous variables are presented as the mean (standard deviation), and categorical variables are presented as the frequency (percentage). The chi-square test or Student's *t*-test were performed to evaluate the differences between the non-exposure and exposure condition on stroke and MDD. Logistic regression models were performed to calculate odds ratios (ORs) for stroke and MDD.

Identification of differentially expressed genes (DEGs)

After the data standardization and normalization of datasets using the `normalizeBetweenArrays` function in the “limma” R package, a principal component analysis (PCA) was conducted by using the “factoextra” R package. The DEGs between cases and healthy controls were analyzed by using the “limma” R package. The criteria of p -value < 0.05 and $|\log \text{fold change (FC)}| > 0.2$ were used to screen the DEGs of MDD and controls, and $|\log \text{FC}| > 0.5$ were regarded as cut-off criteria for significant DEGs for IS patients and controls. A volcano plot and a heat map plot were performed by using the R software `ggplot2` package (Ginestet, 2011) and “ComplexHeatmap” (Gu et al., 2016) to show significant DEGs, respectively.

Screening of communal DEGs of MDD and IS

After having separately identified the DEGs of MDD and IS, we intersected their DEGs to screen out the communal DEGs that may participate in the pathogenesis of the two diseases. Only when the DEGs had the same expression trends in both diseases were these common genes kept. The processes were conducted and visualized using the “ggVennDiagram” package (Gao et al., 2021). The overlapped genes were further shown in two disorders with a heat map from the perspective of $\log \text{FC}$ and p -value.

Function enrichment analysis

The “clusterProfiler” package (Yu et al., 2012) was used to enrich the biological processes (BP), cellular components (CC), and molecular function (MF) of Gene Ontology (GO) (Gene Ontology Consortium, 2015) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways (Kanehisa and Goto, 2000) of common DEGs.

Protein-protein interaction (PPI) network

To detect potential relationships among the DEG-encoded proteins common to both MDD and IS, a protein-protein

interaction (PPI) network was constructed using the Search Tool for the Retrieval of Interacting Genes database (STRING, www.string-db.org) (Szkarczyk et al., 2019). Low confidence of 0.15 was set to find more interactions between proteins. The other parameters were set to the default values (i.e., a full STRING network for network type; evidence for meaning of network edges; and all active interaction sources). The constructed network was imported into Cytoscape to be visualized and further analyzed.

Selection and analysis of hub genes

In this work, we used six common algorithms MCC (Maximal Clique Centrality), MNC (Maximum neighborhood component), DMNC (Density of Maximum Neighborhood Component), Degree, Closeness, and Betweenness) in the cytoHubba plugin of Cytoscape to evaluate and identify hub genes. The detailed information about the six algorithms were described in previous article (Chin et al., 2014). The relationships among genes were calculated using Pearson's correlation methods. The GSE76826 and GSE122709 datasets was applied to validate the expression levels of hub genes.

Subsequently, a co-expression network of these hub genes was constructed via GeneMANIA (<http://www.genemania.org/>) (Wardle-Farley et al., 2010), and their potential functional processes were enriched using the Metascape tool (<https://metascape.org/>) (Zhou et al., 2019a).

DEG-miRNA interaction analysis

NetworkAnalyst (<https://www.networkanalyst.ca/>) is an online platform that aimed to provide a wide-range for meta-analyzing gene expression data and constructing gene regulatory networks in a user-friendly manner (Zhou et al., 2019b). The miRTarBase database provided comprehensive information on experimentally validated miRNA-target interactions and was used to identify regulatory miRNAs that influenced DEGs at the post-transcriptional level in NetworkAnalyst.

Recognition of transcription factors

Transcription factors influence target genes at a transcriptional level. Using the Binding and Expression Target Analysis Minus algorithm, ENCODE targeted the transcription factor of genes derived from the ChIP-seq data. We adopted the ENCODE to predict regulatory TFs of our hub genes.

Prediction of potential drugs of hub genes

The Drug-Gene Interaction Database (DGIdb) (<http://www.dgiddb.org/>) is an online database for identifying drug-gene interaction by integrating the data from, for example the Drug Target Commons, DrugBank, TTD, PharmGKB, and ChEMBL and so on (Wagner et al., 2016). The common hub genes were imported into the database to search for potential drugs. The drug-gene interactions were visualized by the “ggalluvial” R package.

TABLE 1 Baseline characteristics and odds ratio of participants by stroke levels in NHANES (2005–2018).

| Variables | Stroke | | <i>p</i> -value1 | OR | 95% CI | <i>p</i> -value2 |
|--------------------------|----------------------|------------------------|------------------|------|------------------|------------------|
| | No | Yes | | | | |
| MDD | | | <0.0001 | | | |
| No | 26,216 (92.69) | 916 (81.98) | | ref | ref | ref |
| Yes | 2,366 (7.31) | 208 (18.02) | | 2.79 | 2.79 (2.26,3.43) | <0.0001 |
| Age (years) | 46.85 (46.37,47.32) | 63.46 (62.32,64.59) | <0.0001 | 1.07 | 1.07 (1.06,1.07) | <0.0001 |
| Poverty | 3.06 (3.00,3.13) | 2.34 (2.21,2.47) | <0.0001 | 0.76 | 0.76 (0.72,0.80) | <0.0001 |
| BMI(kg.m ²) | 28.99 (28.83,29.16) | 30.07 (29.51,30.64) | <0.001 | 1.02 | 1.02 (1.01,1.03) | <0.001 |
| Waist-circumference (cm) | 99.17 (98.73, 99.61) | 104.53 (103.12,105.94) | <0.0001 | 1.02 | 1.02 (1.01,1.02) | <0.0001 |
| Sex | | | 0.01 | | | |
| Male | 14,361 (49.61) | 557 (44.51) | | ref | ref | ref |
| Female | 14,221 (50.39) | 567 (55.49) | | 1.23 | 1.23 (1.05,1.43) | 0.01 |
| Race | | | <0.0001 | | | |
| Non-Hispanic White | 12,584 (69.21) | 566 (70.84) | | ref | ref | ref |
| Non-Hispanic Black | 5,969 (10.48) | 313 (14.75) | | 1.38 | 1.38 (1.19,1.59) | <0.0001 |
| Mexican American | 4,352 (8.03) | 103 (4.51) | | 0.55 | 0.55 (0.43,0.70) | <0.0001 |
| Other Hispanic | 2,641 (5.20) | 66 (2.83) | | 0.53 | 0.53 (0.38,0.74) | <0.001 |
| Other | 3,036 (7.08) | 76 (7.07) | | 0.98 | 0.98 (0.69,1.38) | 0.89 |
| Education level | | | <0.0001 | | | |
| Less than 9th grade | 2,629 (4.57) | 148 (8.44) | | ref | ref | ref |
| 9–11th grade | 3,891 (10.01) | 207 (15.23) | | 0.83 | 0.83 (0.63,1.09) | 0.17 |
| High school graduate | 6,569 (23.11) | 324 (31.87) | | 0.75 | 0.75 (0.60,0.93) | 0.01 |
| Some college | 8,649 (31.94) | 301 (27.00) | | 0.46 | 0.46 (0.36,0.58) | <0.0001 |
| College graduate | 6,844 (30.36) | 144 (17.46) | | 0.31 | 0.31 (0.24,0.41) | <0.0001 |
| Marital status | | | 0.2 | | | |
| Unmarried | 13,730 (44.15) | 582 (46.82) | | ref | ref | ref |
| Married | 14,852 (55.85) | 542 (53.18) | | 0.9 | 0.90 (0.76,1.06) | 0.20 |
| Smoke | | | <0.0001 | | | |
| Never | 15,711 (54.93) | 415 (38.18) | | ref | ref | ref |
| Former | 6,909 (24.78) | 417 (35.86) | | 2.08 | 2.08 (1.76,2.46) | <0.0001 |
| Now | 5,962 (20.28) | 292 (25.96) | | 1.84 | 1.84 (1.52,2.23) | <0.0001 |
| Diabetes | | | <0.0001 | | | |
| No | 23,538 (86.80) | 665 (63.27) | | ref | ref | ref |
| Yes | 5,044 (13.20) | 459 (36.73) | | 3.82 | 3.82 (3.26,4.48) | <0.0001 |

Immune infiltration analysis

The enrichment for 28 immune infiltrating cells (Bindea et al., 2013) in the MDD and IS was assessed using a single-sample gene set enrichment analysis (ssGSEA) by using the “GSVA” R package (Hänzelmann et al., 2013). The immune cells with the same enrichment trends for both diseases and significant differences between diseases and the healthy controls were identified as the potential immune cells involved in the pathogenesis. The relationships between hub DEGs and immune cells were also constructed.

Statistical analyses

R software (version R-4.1.0) performed all statistical analyses. The Wilcoxon test was used for statistical analysis between two groups. The relationships of genes with genes and genes with immune cells were constructed by using Pearson’s correlation method. A *p*-value less than 0.05 ($p < 0.05$) was considered to indicate statistical significance. The significance level is denoted as follows: * $p < 0.05$, ** $p < 0.01$, and *** $p < 0.001$.

TABLE 2 Baseline characteristics and odds ratio of participants by MDD levels in NHANES (2005–2018).

| Variables | MDD | | <i>p</i> -value1 | OR | 95% CI | <i>p</i> -value2 |
|--------------------------|----------------------|------------------------|------------------|------|------------------|------------------|
| | No | Yes | | | | |
| Stroke | | | <0.0001 | | | |
| No | 26,216 (97.53) | 2,366 (93.41) | | ref | ref | ref |
| Yes | 916 (2.47) | 208 (6.59) | | 2.79 | 2.79 (2.26,3.43) | <0.0001 |
| Age (years) | 47.37 (46.87,47.88) | 46.54 (45.71,47.37) | 0.08 | 1 | 1.00 (0.99,1.00) | 0.08 |
| Poverty | 3.12 (3.05,3.18) | 2.13 (2.02,2.24) | <0.0001 | 0.68 | 0.68 (0.65,0.71) | <0.0001 |
| BMI(kg.m ²) | 28.90 (28.73,29.06) | 30.54 (30.12,30.97) | <0.0001 | 1.03 | 1.03 (1.02,1.04) | <0.0001 |
| Waist-circumference (cm) | 99.06 (98.63, 99.49) | 102.41 (101.36,103.47) | <0.0001 | 1.01 | 1.01 (1.01,1.02) | <0.0001 |
| Sex | | | <0.0001 | | | |
| Male | 13,960 (50.56) | 958 (36.20) | | ref | ref | ref |
| Female | 13,172 (49.44) | 1,616 (63.80) | | 1.8 | 1.80 (1.62,2.01) | <0.0001 |
| Race | | | <0.0001 | | | |
| Non-Hispanic White | 12,019 (69.60) | 1,131 (65.15) | | ref | ref | ref |
| Non-Hispanic Black | 5,730 (10.41) | 552 (12.85) | | 1.32 | 1.32 (1.16,1.50) | <0.0001 |
| Mexican American | 4,075 (7.96) | 380 (7.60) | | 1.02 | 1.02 (0.85,1.22) | 0.83 |
| Other Hispanic | 2,390 (4.96) | 317 (7.19) | | 1.55 | 1.55 (1.27,1.89) | <0.0001 |
| Other | 2,918 (7.07) | 194 (7.21) | | 1.09 | 1.09 (0.89,1.33) | 0.41 |
| Education level | | | <0.0001 | | | |
| Less than 9th grade | 2,428 (4.43) | 349 (7.73) | | ref | ref | ref |
| 9–11th grade | 3,568 (9.61) | 530 (16.73) | | 0.96 | 0.96 (0.83,1.12) | 0.64 |
| High school graduate | 6,259 (23.05) | 634 (27.07) | | 0.65 | 0.65 (0.56,0.76) | <0.0001 |
| Some college | 8,167 (31.61) | 783 (34.22) | | 0.61 | 0.61 (0.52,0.72) | <0.0001 |
| College graduate | 6,710 (31.30) | 278 (14.25) | | 0.25 | 0.25 (0.20,0.32) | <0.0001 |
| Marital status | | | <0.0001 | | | |
| Unmarried | 12,644 (42.72) | 1,668 (62.43) | | ref | ref | ref |
| Married | 14,488 (57.28) | 906 (37.57) | | 0.45 | 0.45 (0.40,0.50) | <0.0001 |
| Smoke | | | <0.0001 | | | |
| never | 15,111 (55.83) | 1,015 (37.94) | | ref | ref | ref |
| former | 6,743 (25.30) | 583 (22.54) | | 1.31 | 1.31 (1.12,1.54) | 0.001 |
| now | 5,278 (18.87) | 976 (39.53) | | 3.08 | 3.08 (2.73,3.49) | <0.0001 |
| Diabetes | | | <0.0001 | | | |
| No | 22,299 (86.69) | 1904 (79.52) | | ref | ref | ref |
| Yes | 4,833 (13.31) | 670 (20.48) | | 1.68 | 1.68 (1.48,1.90) | <0.0001 |

Results

Association between stroke and MDD

Baseline characteristics and the results of logistic regression analysis for stroke and MDD were shown in

Table 1 and Table 2. The results indicated that MDD was significantly associated with an increased risk of stroke. Compared with non-exposure condition, the odds ratios (ORs) with 95% confidence intervals (CIs) for exposure condition between stroke and MDD was 2.79 (2.26,3.43), $p < 0.0001$.

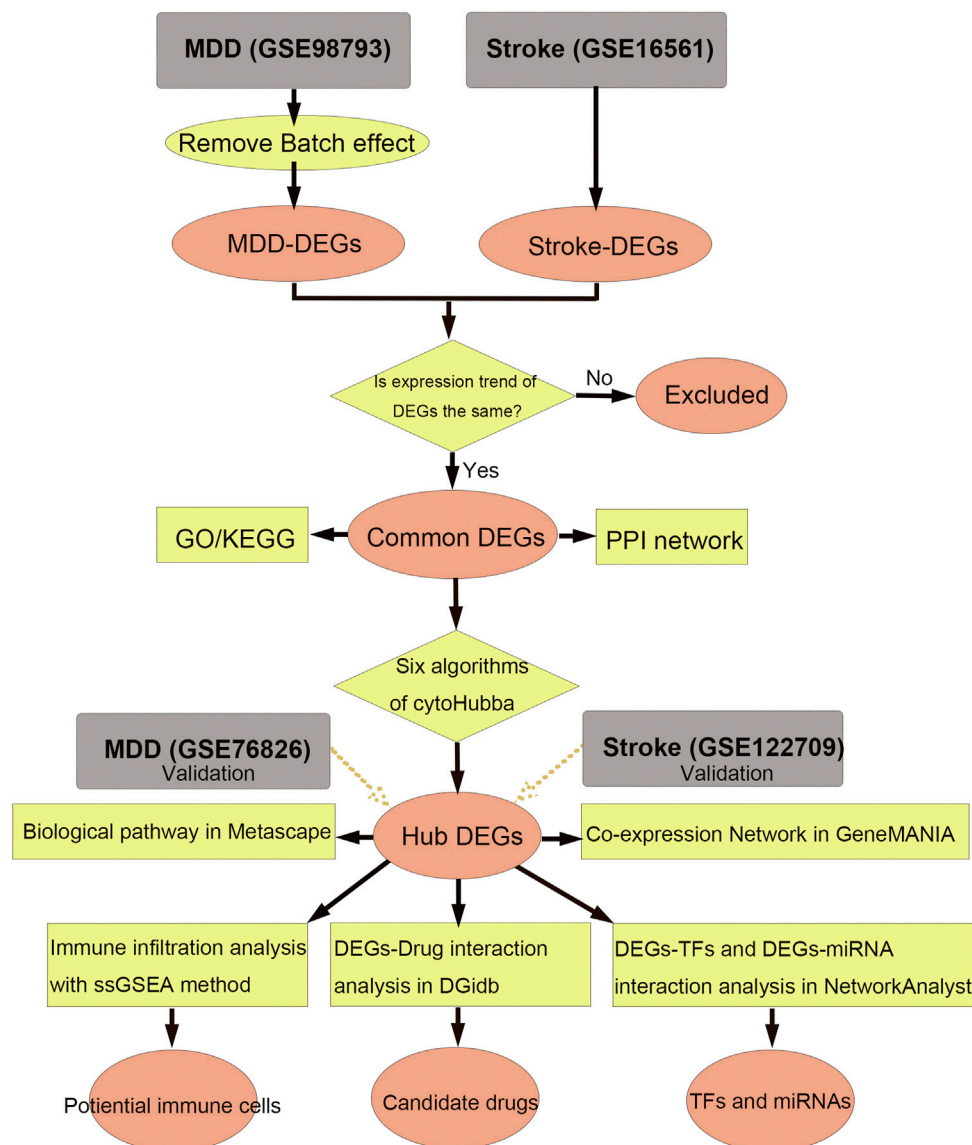


FIGURE 1

Workflow of data analysis in our present work. MDD, major depressive disorder; DEGs, differentially expressed genes; ssGSEA, single sample gene set enrichment analysis; TF, transcription factor; GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; PPI, protein-protein interaction.

Identification of DEGs and communal DEGs between MDD and IS

The flow diagram for this study is shown in Figure 1. To uncover the interrelationships of IS with MDD, we first analyzed the human gene expression datasets from the GEO database to identify the dysregulated genes that stimulate MDD and IS separate. PCA results showed that there were two distinctive batches in GSE989793 in Supplementary Figure S1A, and the batch effect was removed in Supplementary Figure S1B. A volcano plot showed that a total of 336 DEGs were identified based on the following criteria: $|\log_2FC| > 0.2$ and a p -value < 0.05 , including 194 that were upregulated and 142 that were downregulated between MDD patients and healthy controls in Figure 2A. These deregulated

genes are shown with a heat map in Figure 2B. The GSE16561 dataset identified 360 upregulated and 295 downregulated genes taken from IS patient peripheral blood with the cut-off of $|\log_2FC| > 0.5$ and a p -value < 0.05 . The deregulated genes were presented with a volcano and heat map plot in Figures 2C, D, respectively. The PCA result for IS patients and controls are shown in Supplementary Figure S2.

We further overlapped the deregulated genes of MDD and IS with the same expression trends. The Venn diagram showed that 41 common upregulated genes and eight common downregulated genes were finally identified in Figures 3A, B. The differential expression patterns in the two groups were presented with heat map plots from the perspective of \log_2FC and p -value in Figures 3C, D, respectively.

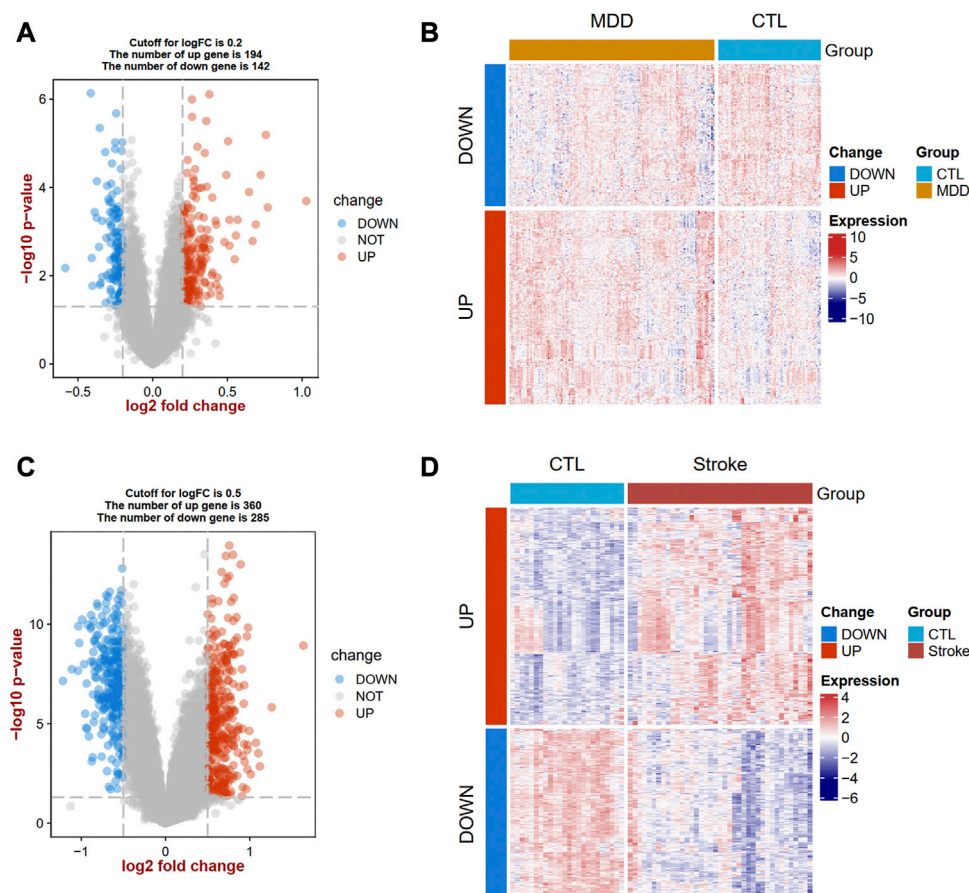


FIGURE 2

Landscapes of differentially expressed genes (DEGs) in MDD and IS. A volcano plot (A) and heat map (B) show the DEGs in MDD. A volcano plot (C) and heat map (D) show the DEGs in IS. MDD, major depressive disorder; IS, ischemic stroke.

Functional enrichment analysis

The biological processes (BP) results showed that the shared genes were mainly enriched in neutrophil activation involved in immune response, defense response to the bacterium, humoral immune response, and reactive oxygen species metabolic process (Figure 4A). In Figure 4B, we observed that these genes were involved in the vesicle lumen, specific granule, and tertiary granule cell component (CC). The molecular functions (MF) of these shared DEGs were enriched in protein heterodimerization activity and serine-related activity in Figure 4C. The KEGG result showed that the primary immunodeficiency, T cell receptor signaling pathway, and antigen processing and presentation pathways were enriched in Figure 4D.

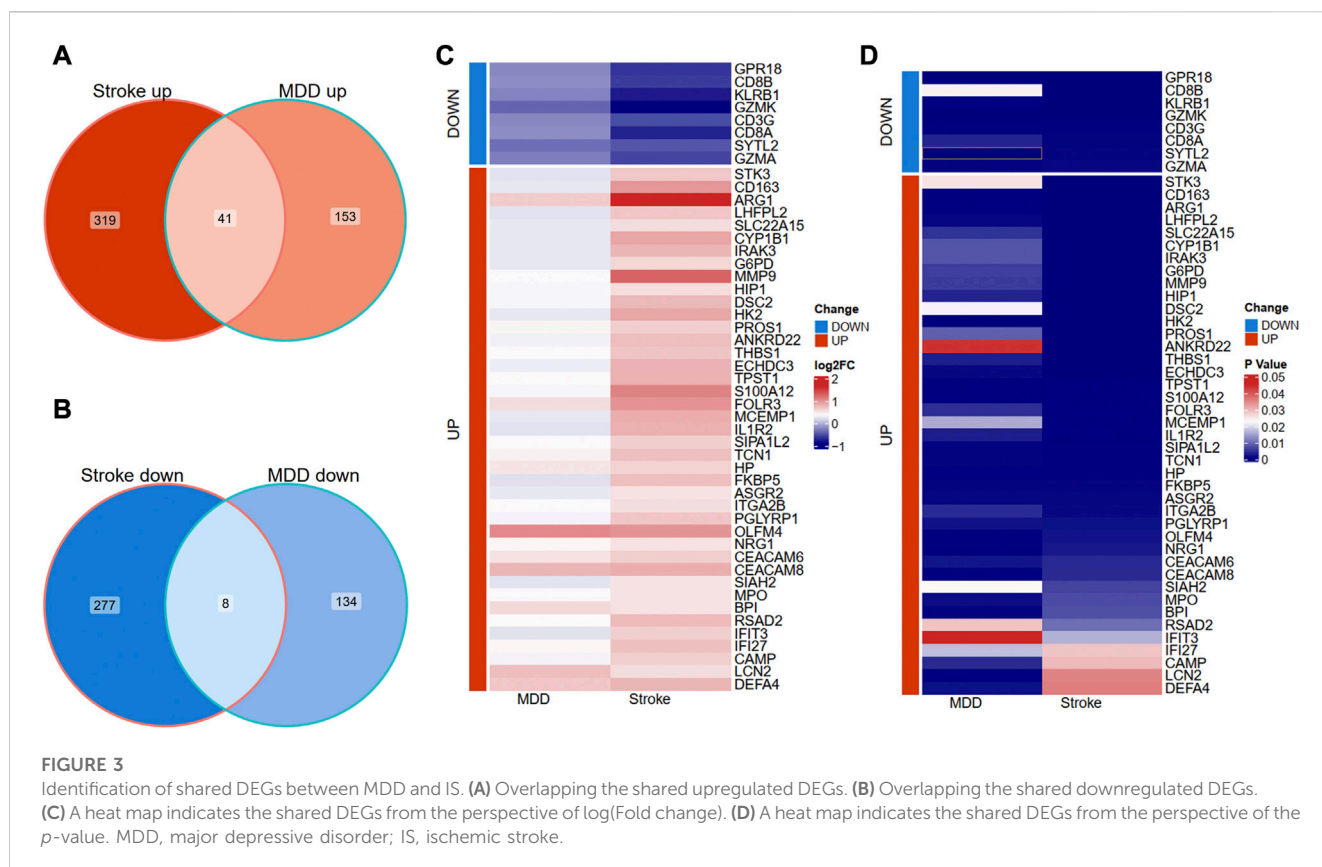
Identification and analysis of hub common DEGs

The shared genes were imported into the STRING database to construct a protein-protein network (Supplementary Figure S3). The network was further visualized in Cytoscape in Figure 5. The red

represents the upregulated DEGs, while the turquoise represents the downregulated DEGs. The size of the node shows the degree of interaction with other genes.

Next, we adopted six algorithms in the cytoHubb plugin of Cytoscape to identify the hub genes. The top 20 genes in each method were visualized in Supplementary Figure S4 and listed in Table 3. We then intersected the 20 genes for each method, and ten overlapped genes (CD163, AEG1, IRAK3, S100A12, HP, PGLYRP1, CEACAM8, MPO, LCN2, and DEFA4) denominated as hub communal DEGs were selected in Figure 6A. The detailed descriptions of the hub genes were listed in Table 4, and that of other DEGs were in Supplementary Table S1. The locations of the 10 genes in the corresponding chromosome are presented in Figure 6B. The violin plot showed that the hub shared genes were significantly expressed in MDD and IS with the same trend in Figures 6C, D. The relationships among genes show that most genes were significantly positively related to each other in MDD (Figure 6E) and stroke (Figure 6F).

The diagnostic ability of the hub genes in MDD (Figure 7A) and IS (Figure 7B) were visualized with receiver operating characteristic curves. The results shows that HP present the greatest diagnostic value with AUC = 0.671 in MDD, while CD163 display the greatest diagnostic value with AUC = 0.965 in IS.



We further validated the expression of the ten hub genes in other external datasets (Supplementary Figure S5). However, only S100A12 was validated as the common differentially expressed gene of the two disorders, which need to be verified by *in vivo* or *in vitro* experiments.

In Figure 8A, the co-expression network of hub genes was constructed using the GeneMANIA website. In the complex PPI network, the interaction of the co-expression accounted for 74.83%, physical interactions for 22.14%, and colocalization for 3.04%. In function analysis, these genes were involved in humoral immune response, secretory granule lumen, cell killing, and regulation of inflammatory response, which was almost consistent with the results from the Metascape analysis (Figure 8B; Supplementary Table S2). In Figure 8C and Supplementary Table S3, we also predicted the potential diseases that the hub genes may be involved in through gene-disease association information collected from the DisGeNET database (Piñero et al., 2020) in Metascape. The results showed that these genes participate in intravascular hemolysis, endotoxemia, and bacterial infections.

Gene regulatory network analysis of Genes-miRNAs and Genes-TFs

TarBase database was utilized to predict the miRNA of hub genes. All hub shared genes were predicted for their interacted miRNA, and a total of 28 miRNA were determined in Figure 9A. In the gene-miRNA interaction network, LCN2 interacted with the

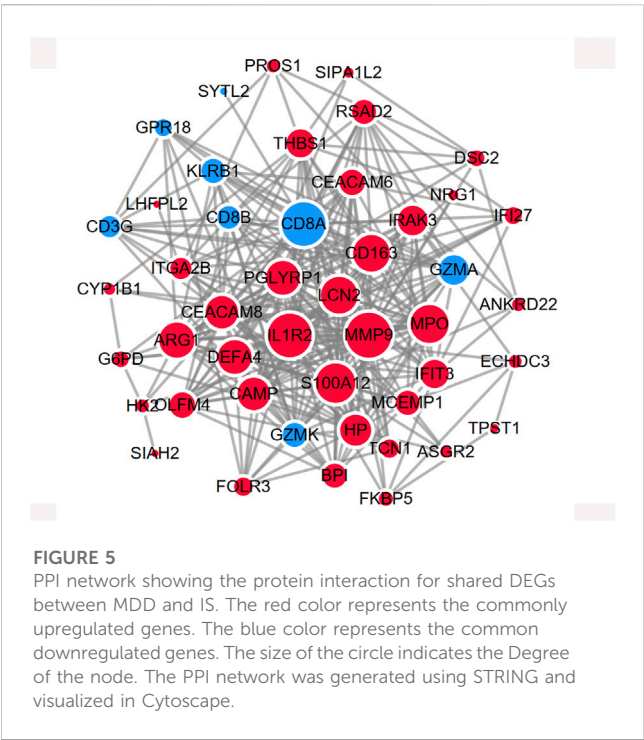
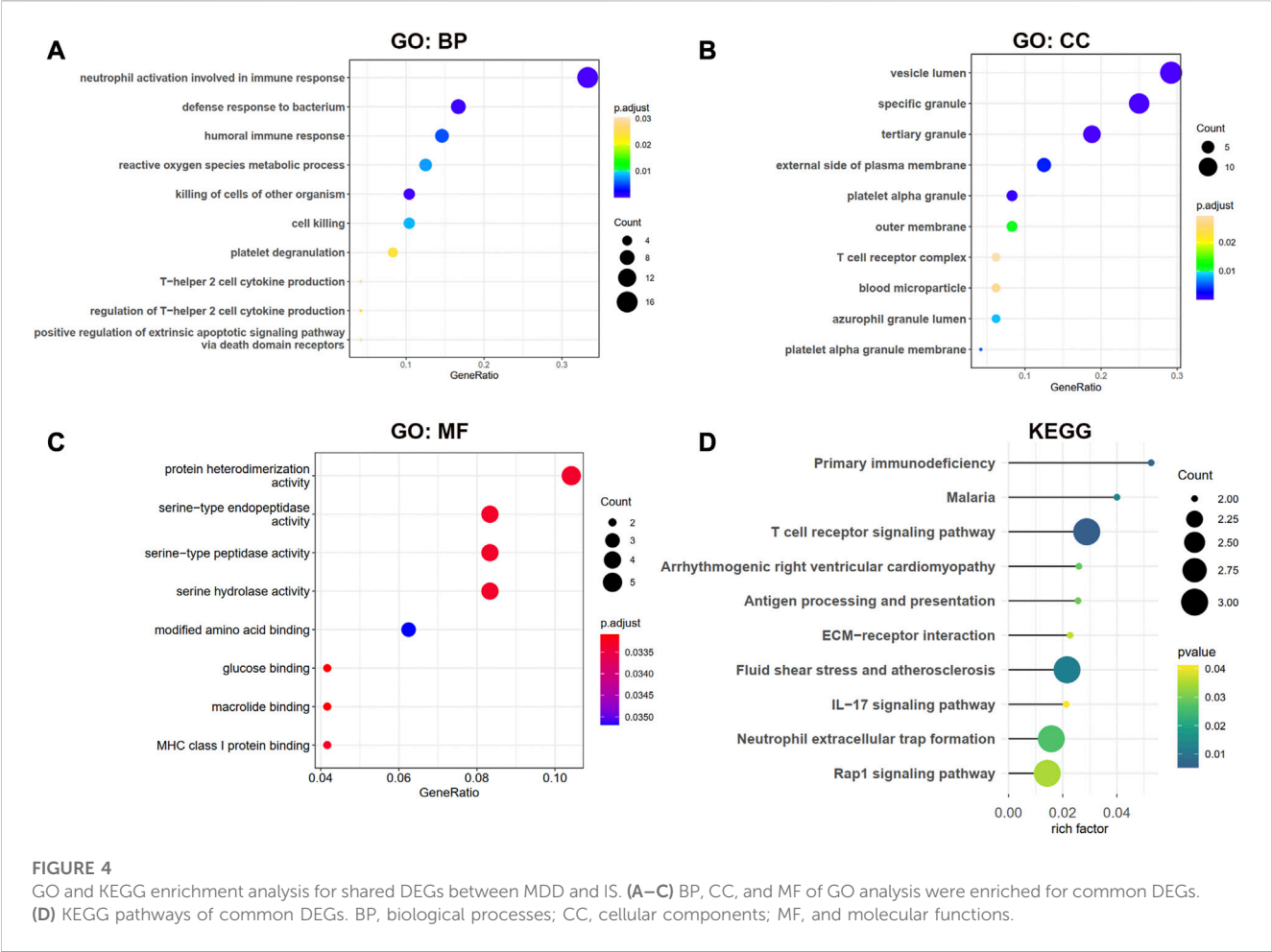
most miRNAs with 13 predicted, followed by HP and IRAK3 with 6 miRNAs. hsa-mir-27a-3p were located in a conspicuous place due to interacting with five hub genes.

We also predicted the experimentally validated TFs of hub genes using ENCODE database in Figure 9B. Only four hub genes predicted their regulatory TFs. A total of 61 TFs were identified, and LCN2 also had the most targeting nodes with 45 TFs, followed by HP with 20 TFs.

The Sankey diagram showed the potential drugs that targeted the hub genes from the DGIdb database (Figure 9C). A total of 30 drugs were predicted, and the detailed information were listed in Supplementary Table S4. Of these, 21 drugs targeted MPO; five drugs targeted S100A12, and 2 drugs each targeted HP and ARG1. No potential drugs could be identified for LCN2, DEFA4, PGLYRP1, CEACAM8, CS163, and IRAK3.

Immune cell infiltration analysis

Using the ssGSEA algorithm, we obtained the immune infiltration of 28 immune cells in the MDD group, IS group, and control group. The immune cells with significant differences between cases and the healthy control group and the same trends were regarded as the potential cells. A total of five immune cells among these 28 types of cells, including activated B cell, activated dendritic cell, effector memory CD8 T cell, macrophage, and natural killer cell were identified, among which active B cell and effector memory CD8 T cell were downregulated, whereas other cells were



upregulated, implying the innate immunity was activated while acquired immunity was suppressed in the two diseases (Figures 10A, B). Figures 10C, D also showed the strong relationships between the hub gene and immune cells.

Discussion

Depression is a global health problem with a high prevalence and the third leading cause of disability globally (Park and Zarate, 2019). The incidence of suicide associated with depression has been increasing and is the 10th leading cause of death in the United States. Similarly, stroke, a neurological disorder characterized by blockage of blood vessels, is a major cause of death and disability worldwide (Johnston et al., 2009). Early studies indicated that depression increased the risk of stroke. A prospective longitudinal study showed that a history of depression was associated with an increased risk of stroke by over twofold (Jackson and Mishra, 2013). Compared with participants with stable low/no depressive symptoms, the participants with the stable high and remitted depressive symptoms had a 2.14 and 1.66 elevated hazard risk of stroke, respectively (Gilsanz et al., 2015). In addition, post-stroke depression (PSD) is one of the common and serious sequelae of stroke. Folstein et al. (1977) first demonstrated that mood disorder is

TABLE 3 The top20 genes identified by six different methods.

| Rank | MNC | MCC | DMNC | Degree | Closeness | Betweenness |
|------|---------|---------|---------|---------|-----------|-------------|
| 1 | MMP9 | MMP9 | BPI | MMP9 | MMP9 | MMP9 |
| 2 | IL1R2 | S100A12 | CAMP | IL1R2 | IL1R2 | CD8A |
| 3 | CD8A | MPO | OLFM4 | CD8A | CD8A | IL1R2 |
| 4 | S100A12 | LCN2 | LCN2 | S100A12 | S100A12 | ITGA2B |
| 5 | MPO | CAMP | CEACAM8 | MPO | MPO | S100A12 |
| 6 | LCN2 | ARG1 | PGLYRP1 | LCN2 | LCN2 | HK2 |
| 7 | CD163 | CD8A | TCN1 | CD163 | CD163 | THBS1 |
| 8 | ARG1 | PGLYRP1 | IRAK3 | ARG1 | ARG1 | ARG1 |
| 9 | PGLYRP1 | HP | HP | PGLYRP1 | PGLYRP1 | DEFA4 |
| 10 | CEACAM8 | IL1R2 | MPO | DEFA4 | DEFA4 | IRAK3 |
| 11 | CAMP | CEACAM8 | CD163 | CEACAM8 | CEACAM8 | MPO |
| 12 | DEFA4 | CD163 | KLRB1 | CAMP | CAMP | FKBP5 |
| 13 | HP | BPI | S100A12 | HP | HP | LCN2 |
| 14 | GZMA | DEFA4 | ARG1 | IRAK3 | IRAK3 | CD163 |
| 15 | IRAK3 | IRAK3 | DEFA4 | GZMA | GZMA | HP |
| 16 | IFIT3 | OLFM4 | GZMA | IFIT3 | THBS1 | IFIT3 |
| 17 | THBS1 | THBS1 | IFI27 | THBS1 | IFIT3 | MCEMP1 |
| 18 | CEACAM6 | GZMA | GZMK | CEACAM6 | CEACAM6 | PGLYRP1 |
| 19 | OLFM4 | CEACAM6 | CD3G | OLFM4 | OLFM4 | CEACAM8 |
| 20 | RSAD2 | MCEMP1 | ITGA2B | RSAD2 | MCEMP1 | ECHDC3 |

a more specific complication of stroke. Disability, anxiety, stroke severity, depression pre-stroke, and cognitive impairment all play an important role in PSD, according to a meta-analysis (Ayerbe et al., 2013). Although the bi-directional relationship between stroke and depression is recognized, the underlying mechanism remains a provocative and unresolved question. Considering that stroke and depression have genetic roots, as well as their frequent comorbidity, we speculate shared genes and biological pathways for both stroke and depression.

With the rapid development of sequencing technology, for the first time, we explored the shared gene signatures and molecular mechanisms between MDD and IS from transcriptome data. In our study, we observe a total of 41 genes are simultaneously upregulated, and eight genes are downregulated in both MDD and IS. Biological enrichment analysis shows that these common genes are involved in the immune response, cell killing, and defense response to the bacterium. Moreover, the T cell receptor signaling pathway, primary immunodeficiency, malaria, IL-17 signaling pathway, and rap1 signaling pathway are enriched in the KEGG pathway analysis. Ten overlapped genes (CD163, AEG1, IRAK3, S100A12, HP, PGLYRP1, CEACAM8, MPO, LCN2, and DEFA4) denominated as hub communal DEGs are identified. We observe that the ten hub genes participate in the immune response and cell killing processes. The gene-diseases analysis reveals that intravascular hemolysis, endotoxemia, and bacterial infections are

correlated with these genes. Furthermore, we construct gene regulatory networks with gene-miRNA, gene-TF, and gene-drugs, which further provide targets for therapeutic interventions. Finally, we depict the immune landscapes for both disorders and found that five immune cells, including activated B cell, activated dendritic cell, effector memory CD8 T cell, macrophage, and natural killer cell, were significantly different in both diseases. Further analysis indicates innate immunity may be activated whereas acquired immunity may be suppressed.

There is consistently and robust evidence supporting the role of inflammation in depression. The inflammatory response in MDD patients was characterized by increased production of complement, chemoattractors, and pro-inflammatory cytokines in peripheral blood cerebrospinal fluid, and post-mortem brain samples (Miller and Raison, 2016). Cytokines, which mediate the innate immune response, including IL-1, tumor necrosis factor (TNF)-alpha, C-reactive protein (CRP), and IL-6, from peripheral blood are considered the most reliable biomarkers of inflammation in patients with depression (Miller et al., 2009). In addition, by inhibiting pro-inflammatory cytokines or their signaling pathways, depressed mood can be improved and conventional antidepressants better tolerated (Kenis and Maes, 2002; Bluthé et al., 2006). Furthermore, by producing anti-inflammatory cytokines (IL-2, IL-4, and IL-10) and/or activating T regulatory (Treg) cells, the effects of immune response were also counter-

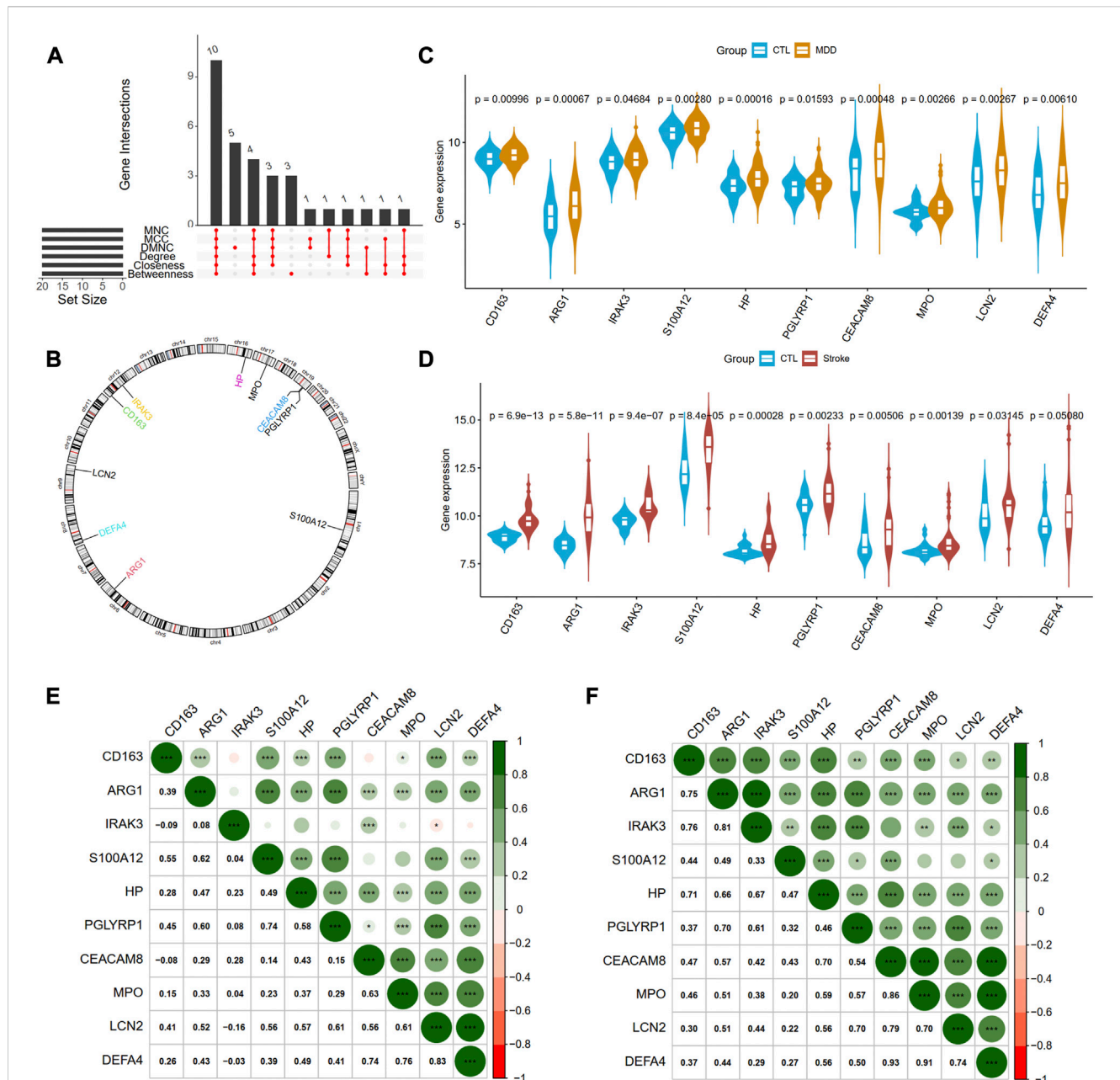


FIGURE 6

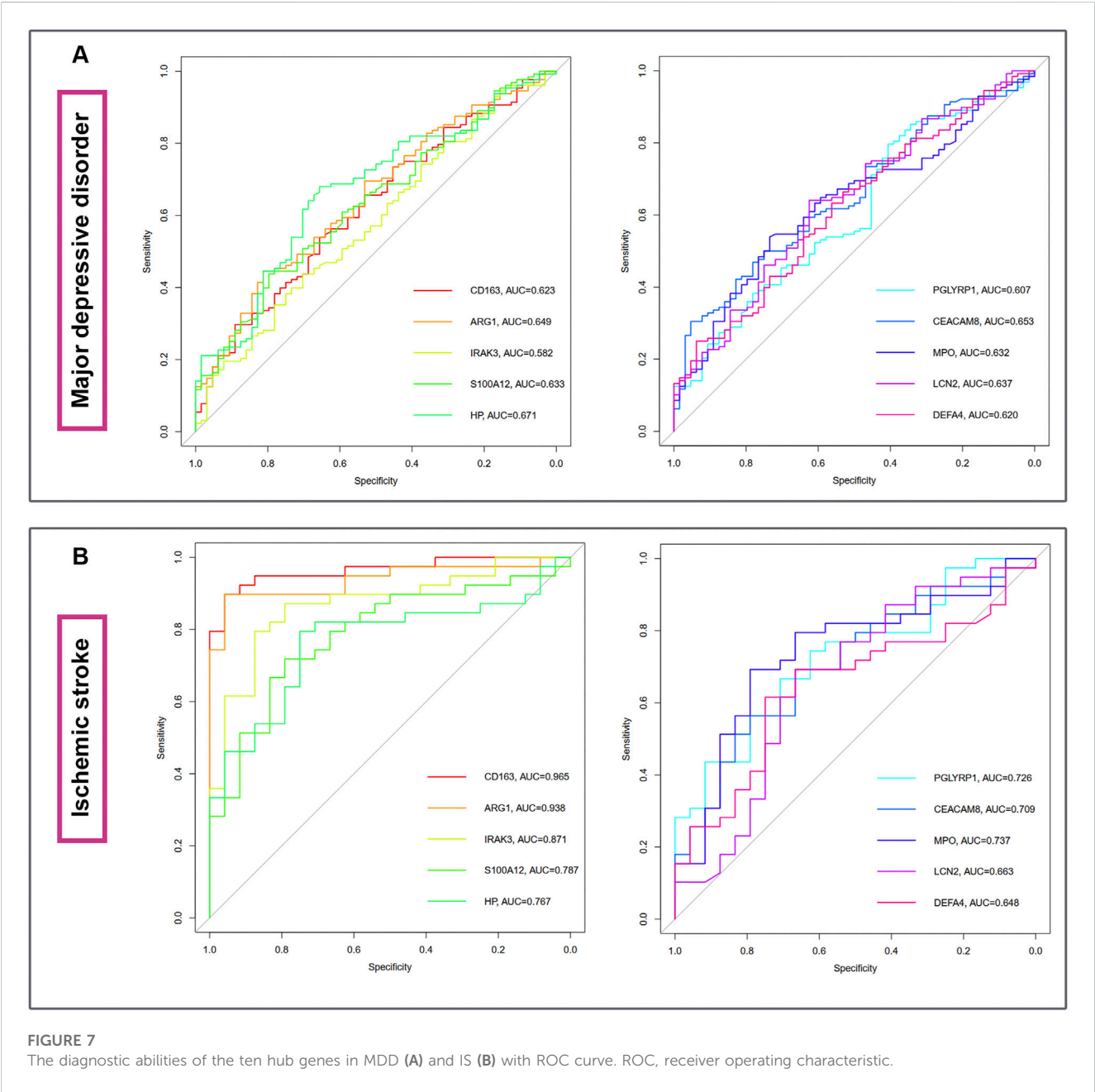
Identification of hub genes from PPI network. (A) A Venn diagram shows ten overlapped genes that were screened out by the six methods from the top 20 genes of six methods in Cytohubba plug in Cytoscape. (B) The gene locations of the ten hub genes. (C) and (D) shows the expression levels of the ten hub genes in MDD and IS, respectively. (E) and (F) depicted the correlations of the ten genes with each other in MDD and IS, respectively. Significance level was denoted by * p -value < 0.05, ** p -value < 0.01, *** p -value < 0.001.

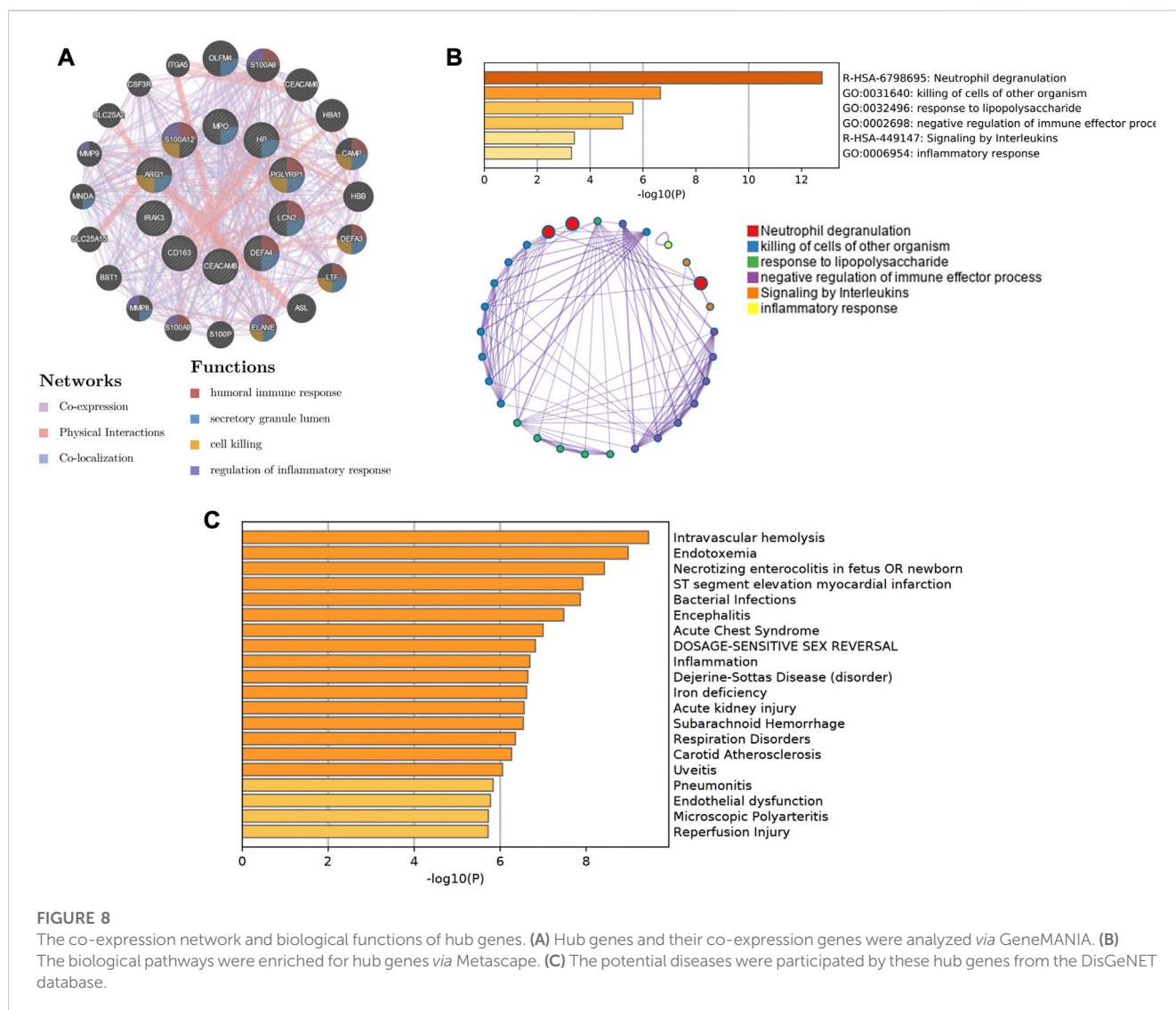
balanced or compensated in MDD patients (Dowlati et al., 2010). Growing evidence also revealed an intimate relationship between the immune system and all stages of the ischemic cascade, from the acute intravascular events induced by a blockage of the blood supply to the parenchymal process causing brain damage (Iadecola and Anrather, 2011; Endres et al., 2022). A recent review summarized that pro-inflammatory interleukins (IL-1b, IL-6, IL-8, IL-12, IL-15, IL-16, IL-20, IL-18, and IL-23/IL-17) and anti-inflammatory interleukins (IL-2, IL-4, IL-10, IL-13, IL-19, and IL-33) were involved in the pathogenesis of IS (Zhu et al., 2022). As

inflammation is common after stroke and depression, immunological processes were proposed as the underlying mechanism triggering PSD (Pascoe et al., 2011). Inflammatory markers such as CRP, ferritin, and neopterin have been linked to PSD development later in life (Becker, 2016). Our GO and KEGG analysis demonstrates that the immune response is enriched in the common DEGs and the ten hub genes of MDD and IS. Moreover, in our immune infiltration analysis, the abundance of active B cells and effector memory CD8 T cell decreases, while that of activated dendritic cell, macrophage, and natural killer cell increases in

TABLE 4 The detailed information and descriptions of hub genes.

| Gene name | Ensembl id | Gene description | Chromosome | Change |
|-----------|-----------------|--|------------|--------|
| CD163 | ENSG00000177575 | CD163 molecule | 12 | UP |
| ARG1 | ENSG00000118520 | arginase 1 | 6 | UP |
| IRAK3 | ENSG00000090376 | interleukin 1 receptor associated kinase 3 | 12 | UP |
| S100A12 | ENSG00000163221 | S100 calcium binding protein A12 | 1 | UP |
| HP | ENSG00000257017 | haptoglobin | 16 | UP |
| PGLYRP1 | ENSG00000008438 | peptidoglycan recognition protein 1 | 19 | UP |
| CEACAM8 | ENSG00000124469 | CEA cell adhesion molecule 8 | 19 | UP |
| MPO | ENSG00000005381 | myeloperoxidase | 17 | UP |
| LCN2 | ENSG00000148346 | lipocalin 2 | 9 | UP |
| DEFA4 | ENSG00000285318 | defensin alpha 4 | 8 | UP |



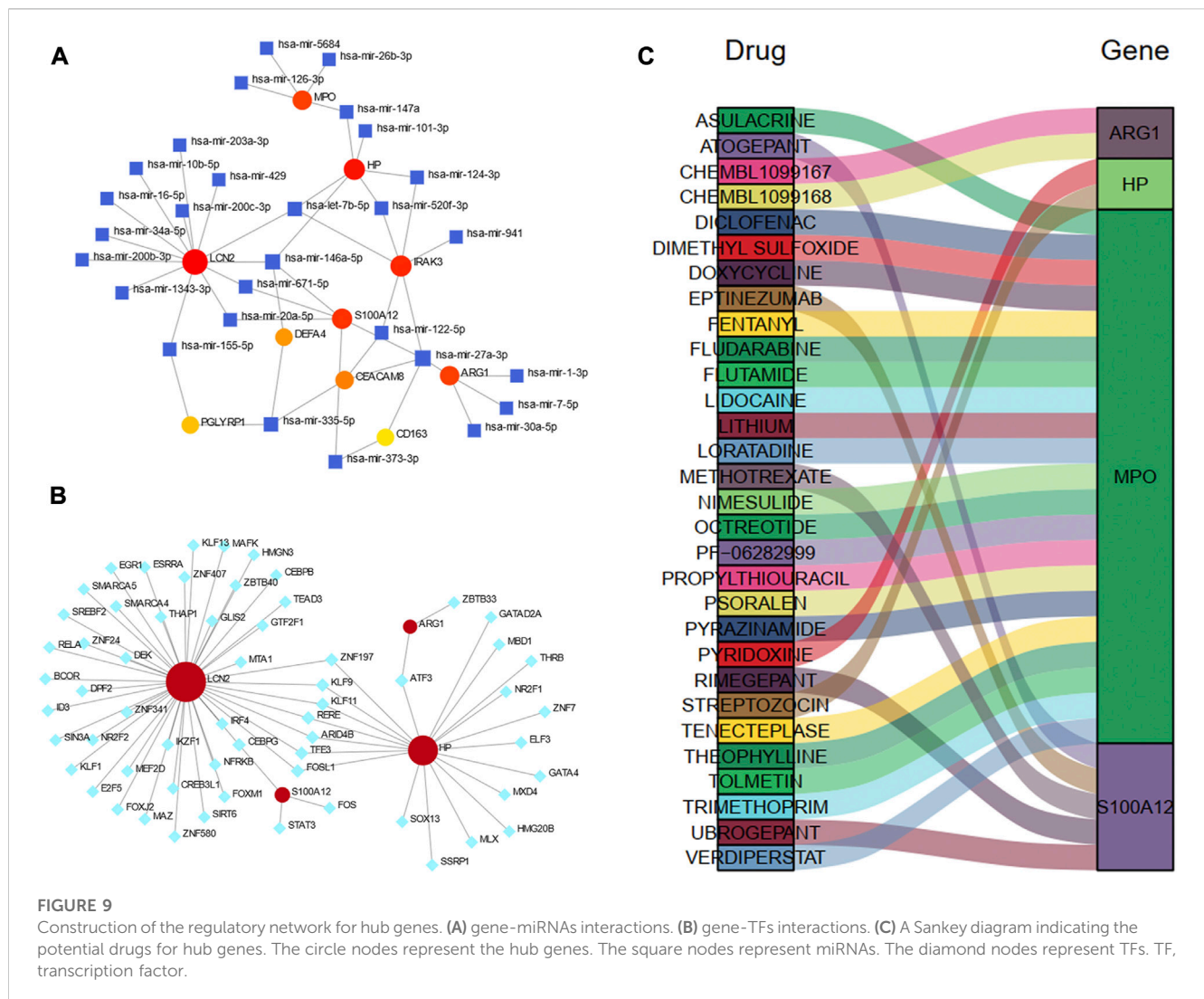


both disorders, which may provide new insight into the common pathogenesis and immunotherapy for both diseases. Specifically, the IL-17 signaling pathway was also observed in our biological function annotation of shared genes. The IL-17 family is an evolutionarily old cytokine family consisting of six members (IL-17A-F), dominantly produced by immune cells of the adaptive and innate lymphocyte lineages, including CD4⁺ Th17 cells, CD8⁺ Tc17 cells, $\gamma\delta$ T17 cells, MAIT cells, and innate lymphoid cells ILC3 (Majumder and McGeachy, 2021). It has been observed that IL-17 levels are high in the central nervous system (CNS) during inflammatory responses, including IS and MDD. Peripheral blood samples from patients with IS show an increased expression of IL-17 compared with healthy individuals (Kostulas et al., 1999). High plasma levels of IL-17 were also detected in MDD patients (Waisman et al., 2015). Combined with our results, we speculate that the IL-17 signaling pathway plays an important role in the shared mechanisms of MDD and IS.

A PPI network analysis was conducted among the proteins derived from shared DEGs to depict functional and physical interactions between IS and MDD. By integrating with six

algorithms (MCC, MNC, DMNC, Degree, Closeness, and Betweenness) in the cytoHubba plugin of Cytoscape, we identify ten hub communal DEGs (CD163, AEG1, IRAK3, S100A12, HP, PGLYRP1, CEACAM8, MPO, LCN2, and DEFA4), which may serve as potential interventional targets.

CD163, the hemoglobin scavenger receptor, is a macrophage-specific protein of the “alternative activation” phenotype and played a major role in dampening the inflammatory response (Moestrup and Møller, 2004). The upregulation of CD163 in monocytes was observed in MDD patients compared with healthy controls (Simon et al., 2021). However, in MDD *in vitro* experiments, sub-anesthetic doses of ketamine, an antidepressant (Murrough et al., 2013), program human monocytes into M2c-like macrophages (anti-inflammatory phenotype) by inducing high levels of CD163 and MERTK (Nowak et al., 2019). Compared with the CD14⁺ classical subtype, CD163 expression was more pronounced in CD16⁺ non-classical and intermediate monocytes after IS and may serve as a potential biomarker of monocyte activation (Greco et al., 2021). Moreover, the percentage of CD163+/CD16⁺ events 24 h after IS was positively associated with stroke severity and disability. In our



analysis, in comparison with the controls, the higher expression of CD163 in both MDD and IS is observed, which may act as a shared risk gene for IS and MDD.

The interleukin receptor-associated kinase (IRAK) family [including IRAK-1, IRAK-2, IRAK-M (IRAK-3), and IRAK-4] are involved in regulating Toll-like receptor (TLR) and interleukin-1 (IL-1) signaling pathways. Interleukin one receptor-associated kinase 3 (IRAK3) is a protein of 596 amino acids with a molecular mass of 68 kDa and is limited to monocytes and macrophages (Wesche et al., 1999). A recent genome-wide association study (GWAS) identified a genome-wide significant locus (rs11465988) in IRAK3 for esketamine efficacy of anti-depression (i.e., percentage change in symptom severity score compared with baseline). The potential roles of IRAK3 in IS have also been discovered recently. The expression levels of IRAK3 that may link natural killer cells to apoptosis were upregulated in IS through bioinformatics analysis (Feng et al., 2022). In experimental stroke mice, IRAK3 has neuroprotective effects, and its deletion can exacerbate neurovascular damages (Lyu et al., 2018). However, our results identify the enhanced expression in both IS and MDD.

Myeloperoxidase (MPO) is a member of the superfamily of heme peroxidases, that is, mainly found in neutrophils and monocytes. High levels of MPO have been detected in the serum of depressive patients in a twin study (Vaccarino et al., 2008). Inhibiting MPO activity and serotonin reuptake may be a potential new approach to MDD with inflammatory syndrome (Soubhye et al., 2014). Moreover, a significant increase in MPO mRNA expression was observed in peripheral blood cells from patients with recurrent depressive disorder (rDD) compared to controls (Galecki et al., 2012; Talarowska et al., 2015). The expression of MPO was also associated with the risk of IS (Wright et al., 2009). Concentrations of serum MPO are increased after IS and was associated with stroke severity (Palm et al., 2018; Orion et al., 2020). Inhibiting MPO activity increased cell proliferation and improved neurogenesis after IS (Kim et al., 2016; Kim et al., 2019). Although MPO contributes to both IS and MDD, few studies were conducted to explore the potential mechanism of IS complicated with MDD. Our result may provide a bridge linking the two disorders.

Lipocalin-2 (LCN2) is a member of the highly heterogeneous lipocalin family of secretory proteins. The roles of LCN2 in IS and depression have been proved recently (Zhao et al., 2019b; Vichaya et al., 2019). A study demonstrated that the relationship between

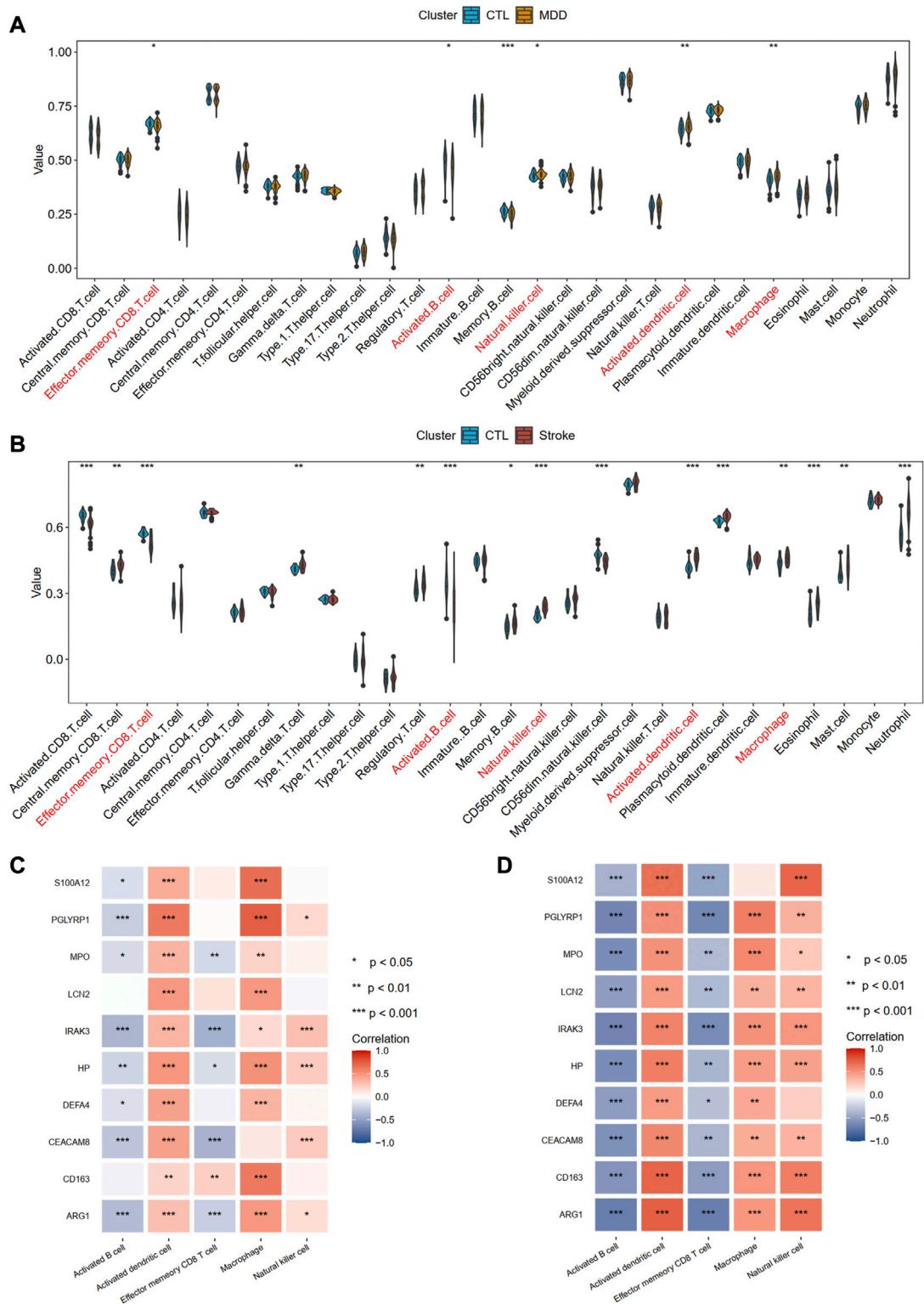


FIGURE 10 Identification of common immune cells between MDD and IS. **(A)** The abundance of the immune cell in MDD using the ssGSEA method. **(B)** The abundance of the immune cell in IS using the ssGSEA method. The immune cells with red color indicate the significantly common immune cells. **(C)** and **(D)** A heat map visualized the correlations between common immune cells and shared hub genes. Significance level was denoted by * p -value < 0.05, ** p -value < 0.01, *** p -value < 0.001.

LCN2 and the process of PSD may be mediated *via* the P38 MAPK pathway (Wei et al., 2021). Our study provides potential association for LCN2 and comorbidity between depression and IS.

Haptoglobin encoded by HP participates in the process of depression and stroke from the perspective of genetic and proteomic levels (Maes et al., 1993; Kiga et al., 2008; Ijäs et al., 2013). Considering few studies focusing on the intermediate role of haptoglobin between depression and stroke, this study provides new insight and reference for investigating its potential roles in the comorbidity, such as PSD.

Regulatory biomolecules may serve as potential interventional targets in multiple complex illnesses. TFs play a key role in regulating the ratio of transcription, and miRNAs handle gene regulation and RNA silencing at the post-transcription level. Given the crucial roles of the ten hub common genes, we also analyze the TFs–gene, miRNAs–gene, and drugs–gene interaction to find transcriptional, post-transcriptional, and therapeutic regulators. mir-27a-3p, mir-146a-5p, mir-335-5p, and let-7b-5p are identified to be interacting with at least three hub genes. Furthermore, we discovered that TFs (such as ZNF197, KLF9, KLF11, RERE, ARID4B, TFE3, and FOSL1) target LCN2 and HP simultaneously. Finally, 31 candidate drugs were predicted, among which 21 drugs target MPO. Combined with the above-mentioned roles of MPO in both depression and stroke, these drugs may serve as potential therapeutics to treat the comorbidities.

Some limitations should be noted in our work. First, although the gene expression profiling from both diseases are derived from the same tissues, there is inadequate information regarding the blood sample collection time for the studies. The disease course of depression and stroke are different. Second, in this study, all the results were acquired by bioinformatic analysis, and we have not conducted any *in vivo* or *in vitro* experiments to verify the different expression levels. Hence, the findings should be interpreted with caution.

Conclusion

We performed a bioinformatic analysis to identify overlapping DEGs subserving both MDD and IS. The communal DEGs participate in the immune response and cell killing processes. Furthermore, ten hub DEGs (CD163, AEG1, IRAK3, S100A12, HP, PGLYRP1, CEACAM8, MPO, LCN2, and DEFA4) were screened out based on six algorithms (MCC, MNC, DMNC, Degree, Closeness, and Betweenness). Immune infiltration analysis shows that the innate immunity was activated whereas acquired immunity was suppressed in both diseases. These findings increase our understating of the association of IS with depression at a transcriptional level. The final gene regulatory network may shed light on novel therapeutic targets for both disorders.

References

- Ayerbe, L., Ayis, S., Wolfe, C. D., and Rudd, A. G. (2013). Natural history, predictors and outcomes of depression after stroke: Systematic review and meta-analysis. *Br. J. psychiatry J. Ment. Sci.* 202, 14–21. doi:10.1192/bjp.bp.111.107664
- Barthels, D., and Das, H. (2020). Current advances in ischemic stroke research and therapies. *Biochimica biophysica acta Mol. basis Dis.* 1866, 165260. doi:10.1016/j.bbdis.2018.09.012
- Becker, K. J. (2016). Inflammation and the silent sequelae of stroke. *Neurother. J. Am. Soc. Exp. Neurother.* 13, 801–810. doi:10.1007/s13311-016-0451-5
- Beurel, E., Toups, M., and Nemeroff, C. B. (2020). The bidirectional relationship of depression and inflammation: Double trouble. *Neuron* 107, 234–256. doi:10.1016/j.neuron.2020.06.002
- Bevan, S., Traylor, M., Adib-Samii, P., Malik, R., Paul, N. L. M., Jackson, C., et al. (2012). Genetic heritability of ischemic stroke and the contribution of previously reported candidate gene and genome-wide associations. *Stroke* 43, 3161–3167. doi:10.1161/strokeaha.112.665760
- Bindea, G., Mlecnik, B., Tosolini, M., Kirilovsky, A., Waldner, M., Obenaus, A. C., et al. (2013). Spatiotemporal dynamics of intratumoral immune cells reveal the immune landscape in human cancer. *Immunity* 39, 782–795. doi:10.1016/j.immuni.2013.10.003

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

Conceptualization, ZY, SL, and MH; methodology, ZY, SL, and MH; software, ZY and MH; validation, QZ; formal analysis, QZ; investigation, QZ; writing-original draft preparation, ZY, SL, and MH; writing-review and editing, HC; visualization, DL; supervision, DL.

Acknowledgments

We acknowledge GEO databases for providing their platforms and contributors for uploading their meaningful datasets. Thanks to Jing Zhang (Shanghai Tongren Hospital) for his work in the NHANES database. His outstanding work, nhanesR package and webpage, makes it easier for us to explore the NHANES database.

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.2023.1004457/full#supplementary-material>

- Bluthé, R. M., Kelley, K. W., and Dantzer, R. (2006). Effects of insulin-like growth factor-I on cytokine-induced sickness behavior in mice. *Brain, Behav. Immun.* 20, 57–63. doi:10.1016/j.bbi.2005.02.003
- Bromet, E., Andrade, L. H., Hwang, I., Sampson, N. A., Alonso, J., de Girolamo, G., et al. (2011). Cross-national epidemiology of DSM-IV major depressive episode. *BMC Med.* 9, 90. doi:10.1186/1741-7015-9-90
- Bucciarelli, V., Caterino, A. L., Bianco, F., Caputi, C. G., Salerni, S., Sciomer, S., et al. (2020). Depression and cardiovascular disease: The deep blue sea of women's heart. *Trends Cardiovasc. Med.* 30, 170–176. doi:10.1016/j.tcm.2019.05.001
- 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, S11. doi:10.1186/1752-0509-8-s4-s11
- Das, J., and Rajanikant, G. K. (2018). Post stroke depression: The sequelae of cerebral stroke. *Neurosci. Biobehav. Rev.* 90, 104–114. doi:10.1016/j.neubiorev.2018.04.005
- Dowlati, Y., Herrmann, N., Swardfager, W., Liu, H., Sham, L., Reim, E. K., et al. (2010). A meta-analysis of cytokines in major depression. *Biol. psychiatry* 67, 446–457. doi:10.1016/j.biopsych.2009.09.033
- Endres, M., Moro, M. A., Nolte, C. H., Dames, C., Buckwalter, M. S., and Meisel, A. (2022). Immune pathways in etiology, acute phase, and chronic sequelae of ischemic stroke. *Circulation Res.* 130, 1167–1186. doi:10.1161/circresaha.121.319994
- Feng, L., Tian, R., Mu, X., Chen, C., Zhang, Y., Cui, J., et al. (2022). Identification of genes linking natural killer cells to apoptosis in acute myocardial infarction and ischemic stroke. *Front. Immunol.* 13, 817377. doi:10.3389/fimmu.2022.817377
- Folstein, M. F., Maiberger, R., and McHugh, P. R. (1977). Mood disorder as a specific complication of stroke. *J. Neurol. Neurosurg. Psychiatry* 40, 1018–1020. doi:10.1136/jnnp.40.10.1018
- Galecki, P., Galecka, E., Maes, M., Chamielec, M., Orzechowska, A., Bobińska, K., et al. (2012). The expression of genes encoding for COX-2, MPO, iNOS, and sPLA2-IIA in patients with recurrent depressive disorder. *J. Affect. Disord.* 138, 360–366. doi:10.1016/j.jad.2012.01.016
- Gao, C. H., Yu, G., and Cai, P. (2021). ggVennDiagram: An intuitive, easy-to-use, and highly customizable R package to generate Venn diagram. *Front. Genet.* 12, 706907. doi:10.3389/fgene.2021.706907
- Gene Ontology Consortium (2015). Gene Ontology Consortium: Going forward. *Nucleic acids Res.* 43, D1049–D1056. doi:10.1093/nar/gku1179
- Gilsanz, P., Walter, S., Tchetgen Tchetgen, E. J., Patton, K. K., Moon, J. R., Capistrant, B. D., et al. (2015). Changes in depressive symptoms and incidence of first stroke among middle-aged and older US adults. *J. Am. Heart Assoc.* 4, e001923. doi:10.1161/jaha.115.001923
- Ginestet, C. (2011). ggplot2: Elegant graphics for data analysis. *J. R. Stat. Soc. Ser. A-STATISTICS Soc.* 174, 245–246. doi:10.1111/j.1467-985X.2010.00676.9.x
- Greco, R., Demartini, C., Zanaboni, A. M., Tumelero, E., Persico, A., Candeloro, E., et al. (2021). CD163 as a potential biomarker of monocyte activation in ischemic stroke patients. *Int. J. Mol. Sci.* 22, 6712. doi:10.3390/ijms22136712
- Gu, Z., Eils, R., and Schlesner, M. (2016). Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinforma. Oxf. Engl.* 32, 2847–2849. doi:10.1093/bioinformatics/btw313
- Hackett, M. L., and Pickles, K. (2014). Part I: Frequency of depression after stroke: An updated systematic review and meta-analysis of observational studies. *Int. J. stroke official J. Int. Stroke Soc.* 9, 1017–1025. doi:10.1111/ijis.12357
- 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
- Harshfield, E. L., Pennells, L., Schwartz, J. E., Willeit, P., Kaptoge, S., Bell, S., et al. (2020). Association between depressive symptoms and incident cardiovascular diseases. *Jama* 324, 2396–2405. doi:10.1001/jama.2020.23068
- Ho, R. C., Niti, M., Kua, E. H., and Ng, T. P. (2008). Body mass index, waist circumference, waist-hip ratio and depressive symptoms in Chinese elderly: A population-based study. *Int. J. geriatric psychiatry* 23, 401–408. doi:10.1002/gps.1893
- Humphries, S. E., and Morgan, L. (2004). Genetic risk factors for stroke and carotid atherosclerosis: Insights into pathophysiology from candidate gene approaches. *Lancet Neurology* 3, 227–235. doi:10.1016/s1474-4422(04)00708-2
- Iadecola, C., and Anrather, J. (2011). The immunology of stroke: From mechanisms to translation. *Nat. Med.* 17, 796–808. doi:10.1038/nm.2399
- Iadecola, C., Buckwalter, M. S., and Anrather, J. (2020). Immune responses to stroke: Mechanisms, modulation, and therapeutic potential. *J. Clin. investigation* 130, 2777–2788. doi:10.1172/jci135530
- Ijäs, P., Saksi, J., Soenne, L., Tuimala, J., Jauhiainen, M., Jula, A., et al. (2013). Haptoglobin 2 allele associates with unstable carotid plaque and major cardiovascular events. *Atherosclerosis* 230, 228–234. doi:10.1016/j.atherosclerosis.2013.07.008
- Jackson, C. A., and Mishra, G. D. (2013). Depression and risk of stroke in midaged women: A prospective longitudinal study. *Stroke* 44, 1555–1560. doi:10.1161/strokeaha.113.001147
- Johnston, S. C., Mendis, S., and Mathers, C. D. (2009). Global variation in stroke burden and mortality: Estimates from monitoring, surveillance, and modelling. *Lancet Neurology* 8, 345–354. doi:10.1016/s1474-4422(09)70023-7
- Joynt, K. E., Whellan, D. J., and O'Connor, C. M. (2003). Depression and cardiovascular disease: Mechanisms of interaction. *Biol. psychiatry* 54, 248–261. doi:10.1016/s0006-3223(03)00568-7
- Kanehisa, M., and Goto, S. (2000). Kegg: Kyoto encyclopedia of genes and genomes. *Nucleic acids Res.* 28, 27–30. doi:10.1093/nar/28.1.27
- Kendler, K. S., Gatz, M., Gardner, C. O., and Pedersen, N. L. (2006). A Swedish national twin study of lifetime major depression. *Am. J. psychiatry* 163, 109–114. doi:10.1176/appi.ajp.163.1.109
- Kenis, G., and Maes, M. (2002). Effects of antidepressants on the production of cytokines. *Int. J. Neuropsychopharmacol.* 5, 401–412. doi:10.1017/s1461145702003164
- Kiga, C., Sakurai, H., Goto, H., Hayashi, K., Shimada, Y., and Saiki, I. (2008). Proteomic identification of haptoglobin as a stroke plasma biomarker in spontaneously hypertensive stroke-prone rats. *Life Sci.* 83, 625–631. doi:10.1016/j.lfs.2008.08.013
- Kim, H., Wei, Y., Lee, J. Y., Wu, Y., Zheng, Y., Moskowitz, M. A., et al. (2016). Myeloperoxidase inhibition increases neurogenesis after ischemic stroke. *J. Pharmacol. Exp. Ther.* 359, 262–272. doi:10.1124/jpet.116.235127
- Kim, H. J., Wei, Y., Wojtkiewicz, G. R., Lee, J. Y., Moskowitz, M. A., and Chen, J. W. (2019). Reducing myeloperoxidase activity decreases inflammation and increases cellular protection in ischemic stroke. *J. Cereb. blood flow metabolism official J. Int. Soc. Cereb. Blood Flow Metabolism* 39, 1864–1877. doi:10.1177/0271678x18771978
- Kostulas, N., Pelidou, S. H., Kivisaak, P., Kostulas, V., and Link, H. (1999). Increased IL-1beta, IL-8, and IL-17 mRNA expression in blood mononuclear cells observed in a prospective ischemic stroke study. *Stroke* 30, 2174–2179. doi:10.1161/01.str.30.12.2174
- Lyu, C., Zhang, Y., Gu, M., Huang, Y., Liu, G., Wang, C., et al. (2018). IRAK-M deficiency exacerbates ischemic neurovascular injuries in experimental stroke mice. *Front. Cell. Neurosci.* 12, 504. doi:10.3389/fncel.2018.00504
- Maes, M., Scharpé, S., Meltzer, H. Y., and Cosyns, P. (1993). Relationships between increased haptoglobin plasma levels and activation of cell-mediated immunity in depression. *Biol. psychiatry* 34, 690–701. doi:10.1016/0006-3223(93)90042-c
- Majumder, S., and McGeachy, M. J. (2021). IL-17 in the pathogenesis of disease: Good intentions gone awry. *Annu. Rev. Immunol.* 39, 537–556. doi:10.1146/annurev-immunol-101819-092536
- Meschia, J. F., Bushnell, C., Boden-Albala, B., Braun, L. T., Bravata, D. M., Chaturvedi, S., et al. (2014). Guidelines for the primary prevention of stroke: A statement for healthcare professionals from the American heart association/American stroke association. *Stroke* 45, 3754–3832. doi:10.1161/str.0000000000000046
- Miller, A. H., Maletic, V., and Raison, C. L. (2009). Inflammation and its discontents: The role of cytokines in the pathophysiology of major depression. *Biol. psychiatry* 65, 732–741. doi:10.1016/j.biopsych.2008.11.029
- Miller, A. H., and Raison, C. L. (2016). The role of inflammation in depression: From evolutionary imperative to modern treatment target. *Nat. Rev. Immunol.* 16, 22–34. doi:10.1038/nri.2015.5
- Moestrup, S. K., and Möller, H. J. (2004). CD163: A regulated hemoglobin scavenger receptor with a role in the anti-inflammatory response. *Ann. Med.* 36, 347–354. doi:10.1080/07853890410033171
- Murrough, J. W., Iosifescu, D. V., Chang, L. C., Al Jurdi, R. K., Green, C. E., Perez, A. M., et al. (2013). Antidepressant efficacy of ketamine in treatment-resistant major depression: A two-site randomized controlled trial. *Am. J. psychiatry* 170, 1134–1142. doi:10.1176/appi.ajp.2013.13030392
- Nowak, W., Grendas, L. N., Sanmarco, L. M., Estecho, I. G., Arena, A. R., Eberhardt, N., et al. (2019). Pro-inflammatory monocyte profile in patients with major depressive disorder and suicide behaviour and how ketamine induces anti-inflammatory M2 macrophages by NMDAR and mTOR. *EBioMedicine* 50, 290–305. doi:10.1016/j.ebiom.2019.10.063
- Orion, D., von Landenberg, P., Itsekson-Hayosh, Z., Schwammethal, Y., Tsaabari, R., Merzeliak, O., et al. (2020). Plasma myeloperoxidase levels in acute brain ischaemia and high grade carotid stenosis. *Eur. J. neurology* 27, 1604–1611. doi:10.1111/ene.14279
- Palm, F., Pussinen, P. J., Safer, A., Tervahartala, T., Sorsa, T., Urbanek, C., et al. (2018). Serum matrix metalloproteinase-8, tissue inhibitor of metalloproteinase and myeloperoxidase in ischemic stroke. *Atherosclerosis* 271, 9–14. doi:10.1016/j.atherosclerosis.2018.02.012
- Pan, A., Sun, Q., Okereke, O. I., Rexrode, K. M., and Hu, F. B. (2011). Depression and risk of stroke morbidity and mortality: A meta-analysis and systematic review. *Jama* 306, 1241–1249. doi:10.1001/jama.2011.1282
- Park, L. T., and Zarate, C. A., Jr. (2019). Depression in the primary care setting. *N. Engl. J. Med.* 380, 559–568. doi:10.1056/NEJMcp1712493
- Pascoe, M. C., Crewther, S. G., Carey, L. M., and Crewther, D. P. (2011). Inflammation and depression: Why poststroke depression may be the norm and not the exception. *Int. J. stroke official J. Int. Stroke Soc.* 6, 128–135. doi:10.1111/j.1747-4949.2010.00565.x
- Patten, S. B., Williams, J. V., Lavorato, D. H., Campbell, N. R. C., Eliasziw, M., and Campbell, T. S. (2009). Major depression as a risk factor for high blood pressure: Epidemiologic evidence from a national longitudinal study. *Psychosom. Med.* 71, 273–279. doi:10.1097/PSY.0b013e3181988e5f
- Piñero, J., Ramírez-Anguaita, J. M., Saich-Pitarch, J., Ronzano, F., Centeno, E., Sanz, F., et al. (2020). The DisGeNET knowledge platform for disease genomics: 2019 update. *Nucleic acids Res.* 48, D845–D855. doi:10.1093/nar/gkz1021

- 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
- Robinson, R. G., and Jorge, R. E. (2016). Post-stroke depression: A review. *Am. J. psychiatry* 173, 221–231. doi:10.1176/appi.ajp.2015.15030363
- Simon, M. S., Schiweck, C., Arteaga-Henriquez, G., Poletti, S., Haarman, B. C. M., Dik, W. A., et al. (2021). Monocyte mitochondrial dysfunction, inflamming, and inflammatory pyroptosis in major depression. *Prog. neuro-psychopharmacology Biol. psychiatry* 111, 110391. doi:10.1016/j.pnpbp.2021.110391
- Soubhye, J., Aldib, I., Prévost, M., Elfving, B., Gelbcke, M., Podrecca, M., et al. (2014). Hybrid molecules inhibiting myeloperoxidase activity and serotonin reuptake: A possible new approach of major depressive disorders with inflammatory syndrome. *J. Pharm. Pharmacol.* 66, 1122–1132. doi:10.1111/jphp.12236
- Strine, T. W., Mokdad, A. H., Dube, S. R., Balluz, L. S., Gonzalez, O., Berry, J. T., et al. (2008). The association of depression and anxiety with obesity and unhealthy behaviors among community-dwelling US adults. *General Hosp. psychiatry* 30, 127–137. doi:10.1016/j.genhosppsych.2007.12.008
- Sullivan, P. F., Neale, M. C., and Kendler, K. S. (2000). Genetic epidemiology of major depression: Review and meta-analysis. *Am. J. psychiatry* 157, 1552–1562. doi:10.1176/appi.ajp.157.10.1552
- Szklarczyk, D., Gable, A. L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., et al. (2019). STRING v11: Protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic acids Res.* 47, D607–D613. doi:10.1093/nar/gky1131
- Talarowska, M., Szemraj, J., and Galecki, P. (2015). Myeloperoxidase gene expression and cognitive functions in depression. *Adv. Med. Sci.* 60, 1–5. doi:10.1016/j.advms.2014.06.001
- Vaccarino, V., Brennan, M. L., Miller, A. H., Bremner, J. D., Ritchie, J. C., Lindau, F., et al. (2008). Association of major depressive disorder with serum myeloperoxidase and other markers of inflammation: A twin study. *Biol. psychiatry* 64, 476–483. doi:10.1016/j.biopsych.2008.04.023
- Vichaya, E. G., Gross, P. S., Estrada, D. J., Cole, S. W., Grossberg, A. J., Evans, S. E., et al. (2019). Lipocalin-2 is dispensable in inflammation-induced sickness and depression-like behavior. *Psychopharmacology* 236, 2975–2982. doi:10.1007/s00213-019-05190-7
- Wagner, A. H., Coffman, A. C., Ainscough, B. J., Spies, N. C., Skidmore, Z. L., Campbell, K. M., et al. (2016). DGIdb 2.0: Mining clinically relevant drug-gene interactions. *Nucleic acids Res.* 44, D1036–D1044. doi:10.1093/nar/gkv1165
- Waisman, A., Hauptmann, J., and Regen, T. (2015). The role of IL-17 in CNS diseases. *Acta neuropathol.* 129, 625–637. doi:10.1007/s00401-015-1402-7
- Warde-Farley, D., Donaldson, S. L., Comes, O., Zuberi, K., Badrawi, R., Chao, P., et al. (2010). The GeneMANIA prediction server: Biological network integration for gene prioritization and predicting gene function. *Nucleic acids Res.* 38, W214–W220. doi:10.1093/nar/gkq537
- Wassertheil-Smolter, S., Qi, Q., Dave, T., Mitchell, B. D., Jackson, R. D., Liu, S., et al. (2018). Polygenic risk for depression increases risk of ischemic stroke: From the stroke genetics network study. *Stroke* 49, 543–548. doi:10.1161/strokeaha.117.018857
- Wei, L., Du, Y., Xie, Y., Yu, X., Chen, H., and Qiu, Y. (2021). Lipocalin-2 regulates hippocampal microglial activation in poststroke depression. *Front. aging Neurosci.* 13, 798335. doi:10.3389/fnagi.2021.798335
- Wesche, H., Gao, X., Li, X., Kirschning, C. J., Stark, G. R., and Cao, Z. (1999). IRAK-M is a novel member of the Pelle/interleukin-1 receptor-associated kinase (IRAK) family. *J. Biol. Chem.* 274, 19403–19410. doi:10.1074/jbc.274.27.19403
- Wesołowska, K., Elovainio, M., Hintsa, T., Jokela, M., Pulkki-Raback, L., Lipsanen, J., et al. (2018). Is the association between depressive symptoms and glucose bidirectional? A population-based study. *Health Psychol. official J. Div. Health Psychol. Am. Psychol. Assoc.* 37, 603–612. doi:10.1037/hea0000612
- Wright, C. B., Moon, Y., Paik, M. C., Brown, T. R., Rabbani, L., Yoshita, M., et al. (2009). Inflammatory biomarkers of vascular risk as correlates of leukoariorosis. *Stroke* 40, 3466–3471. doi:10.1161/strokeaha.109.559567
- Yu, G., Wang, L. G., Han, Y., and He, Q. Y. (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. *Omics a J. Integr. Biol.* 16, 284–287. doi:10.1089/omi.2011.0118
- Zhao, F., Yue, Y., Jiang, H., and Yuan, Y. (2019a). Shared genetic risk factors for depression and stroke. *Prog. neuro-psychopharmacology Biol. psychiatry* 93, 55–70. doi:10.1016/j.pnpbp.2019.03.003
- Zhao, N., Xu, X., Jiang, Y., Gao, J., Wang, F., Xu, X., et al. (2019b). Lipocalin-2 may produce damaging effect after cerebral ischemia by inducing astrocytes classical activation. *J. neuroinflammation* 16, 168. doi:10.1186/s12974-019-1556-7
- Zhou, G., Soufan, O., Ewald, J., Hancock, R. E. W., Basu, N., and Xia, J. (2019). NetworkAnalyst 3.0: A visual analytics platform for comprehensive gene expression profiling and meta-analysis. *Nucleic acids Res.* 47, W234–W241. doi:10.1093/nar/gkz240
- Zhou, Y., Zhou, B., Pache, L., Chang, M., Khodabakhshi, A. H., Tanaseichuk, O., et al. (2019). Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat. Commun.* 10, 1523. doi:10.1038/s41467-019-09234-6
- Zhu, H., Hu, S., Li, Y., Sun, Y., Xiong, X., Hu, X., et al. (2022). Interleukins and ischemic stroke. *Front. Immunol.* 13, 828447. doi:10.3389/fimmu.2022.828447



OPEN ACCESS

EDITED BY

Thayne Kowalski,
Federal University of Rio Grande do Sul,
Brazil

REVIEWED BY

Mi Zhang,
Huazhong University of Science and
Technology, China
Xiang Cao,
Nanjing Drum Tower Hospital, China

*CORRESPONDENCE

Ju Gao,
✉ gaoju_003@163.com
Tianfeng Huang,
✉ 18051063400@yzu.edu.cn

SPECIALTY SECTION

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

RECEIVED 31 August 2022

ACCEPTED 23 March 2023

PUBLISHED 03 April 2023

CITATION

Xiao Y, Zhang Y, Wang C, Ge Y, Gao J and
Huang T (2023), The use of multiple
datasets to identify autophagy-related
molecular mechanisms in
intracerebral hemorrhage.
Front. Genet. 14:1032639.
doi: 10.3389/fgene.2023.1032639

COPYRIGHT

© 2023 Xiao, Zhang, Wang, Ge, Gao and
Huang. This is an open-access article
distributed under the terms of the
[Creative Commons Attribution License](#)
(CC BY). The use, distribution or
reproduction in other forums is
permitted, provided the original author(s)
and the copyright owner(s) are credited
and that the original publication in this
journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

The use of multiple datasets to identify autophagy-related molecular mechanisms in intracerebral hemorrhage

Yinggang Xiao^{1,2,3}, Yang Zhang^{1,2,3}, Cunjin Wang^{1,2,3}, Yali Ge^{1,2,3},
Ju Gao^{1,2,3*} and Tianfeng Huang^{1,2,3*}

¹Department of Anesthesiology, Clinical Medical College of Yangzhou University, Yangzhou, Jiangsu, China, ²Department of Anesthesiology, Yangzhou University Affiliated Northern Jiangsu People's Hospital, Yangzhou, Jiangsu, China, ³Yangzhou Key Laboratory of Anesthesiology, Yangzhou, Jiangsu, China

Background: Intracerebral hemorrhage (ICH) is a stroke syndrome with high mortality and disability rates, but autophagy's mechanism in ICH is still unclear. We identified key autophagy genes in ICH by bioinformatics methods and explored their mechanisms.

Methods: We downloaded ICH patient chip data from the Gene Expression Omnibus (GEO) database. Based on the GENE database, differentially expressed genes (DEGs) for autophagy were identified. We identified key genes through protein–protein interaction (PPI) network analysis and analyzed their associated pathways in Gene Ontology (GO) and the Kyoto Encyclopedia of Genes and Genomes (KEGG). Gene-motif rankings, miRWalk and ENCORI databases were used to analyze the key gene transcription factor (TF) regulatory network and ceRNA network. Finally, relevant target pathways were obtained by gene set enrichment analysis (GSEA).

Results: Eleven autophagy-related DEGs in ICH were obtained, and *IL-1B*, *STAT3*, *NLRP3* and *NOD2* were identified as key genes with clinical predictive value by PPI and receiver operating characteristic (ROC) curve analysis. The candidate gene expression level was significantly correlated with the immune infiltration level, and most of the key genes were positively correlated with the immune cell infiltration level. The key genes are mainly related to cytokine and receptor interactions, immune responses and other pathways. The ceRNA network predicted 8,654 interaction pairs (24 miRNAs and 2,952 lncRNAs).

Conclusion: We used multiple bioinformatics datasets to identify *IL-1B*, *STAT3*, *NLRP3* and *NOD2* as key genes that contribute to the development of ICH.

KEYWORDS

intracerebral hemorrhage, autophagy, immune infiltration, bioinformatics analysis, ceRNA network

1 Introduction

Intracerebral hemorrhage (ICH) is a common stroke syndrome, accounting for approximately 15% of strokes, and nearly 50% of stroke-related deaths worldwide are related to ICH (Feigin et al., 2009; Biffi et al., 2016). ICH is caused by the sudden rupture of blood vessels caused by pathological accumulation of blood in the brain parenchyma (Jia et al., 2020). ICH injury is divided into primary and secondary injuries, with the former being caused by direct mechanical action of the hematoma (Fu et al., 2022). Edema around the hematoma occurs within hours of ICH, disrupting the blood–brain barrier and adjacent tissues and leading to secondary damage (Li et al., 2020). Second, mitochondrial dysfunction, neurotransmitter disturbance, microglial activation, and the release of inflammatory mediators are also important mechanisms for aggravating brain injury (Kim-Han et al., 2006). The death of nerve cells after ICH is closely related to the sequelae of ICH and death from ICH. Programmed cell death (PCD) refers to the autonomous and orderly death of cells controlled by genes to maintain the stability of the internal environment. PCD is an active suicidal behavior of cells (Huysmans et al., 2018). PCD, including autophagy, apoptosis and pyroptosis, plays an important role in neuronal cell death after ICH (Bobinger et al., 2018). Autophagy, as an important category of PCD, has been identified in ICH, but its mechanism in intracerebral hemorrhage remains unclear.

Autophagy is one of the important subcellular events occurring from eukaryotic cells to mammals, and the process of autophagy is highly conserved. Autophagy refers to the process in which cells can wrap their intracellular contents under stress and integrate with lysosomes to degrade into these contents into biomacromolecules, which are reused by cells (Ohsumi, 2014). Recent studies have shown that autophagy is closely related to the occurrence of various neurological diseases (Moujalled et al., 2021). In recent years, autophagy has been found to be closely related to secondary brain tissue damage after ICH (Duan et al., 2016; Zhang et al., 2021). After ICH occurs, thrombin is produced in the blood coagulation process, while the hematoma gradually degrades, releasing degradation products such as hemoglobin, heme and iron that invade the surrounding brain tissue. When iron overload and abnormal thrombin expression occur in brain tissue, autophagy is activated and involved in the brain protection process to reduce injury, remove harmful substances and maintain intracellular environmental homeostasis. The protective role of autophagy in ICH has been demonstrated (Wang et al., 2020; Li et al., 2021). However, the overactivation of autophagy, which activates microglia to produce proinflammatory factors and damages neurons, leads to the aggravation of secondary injury after ICH (Shi et al., 2018; Zhang et al., 2021). In summary, autophagy is extremely important for the progression of ICH, but the key genes involved in this process are still not clearly known. The diagnosis of the severity of ICH on the basis of autophagy-related gene expression is also a clinical blind spot. The key autophagy-related genes in ICH need to be identified.

To explore and identify potential biomarkers and the key autophagy-related genes in ICH, we obtained microarray and gene

information from multiple databases and used the R statistical programming language for analysis. We selected DEGs in perihematomal tissue (PH) and contralateral normal tissue from intracerebral hemorrhage patients obtained from multiple sources as raw data. Then, four key genes were screened by analyzing the interactions and relationships of DEGs highly related to autophagy with the ROC curve method. Finally, we analyzed the impact of key genes on the immune microenvironment and the mechanisms by which these genes are regulated by transcription factors and non-coding RNAs. We innovatively used methods such as ceRNA network construction, motif-TF annotation and xCell to analyze autophagy after ICH. These results will contribute to the study of the mechanism of secondary injury following ICH and provide new ideas for the diagnosis and treatment of ICH in the clinic.

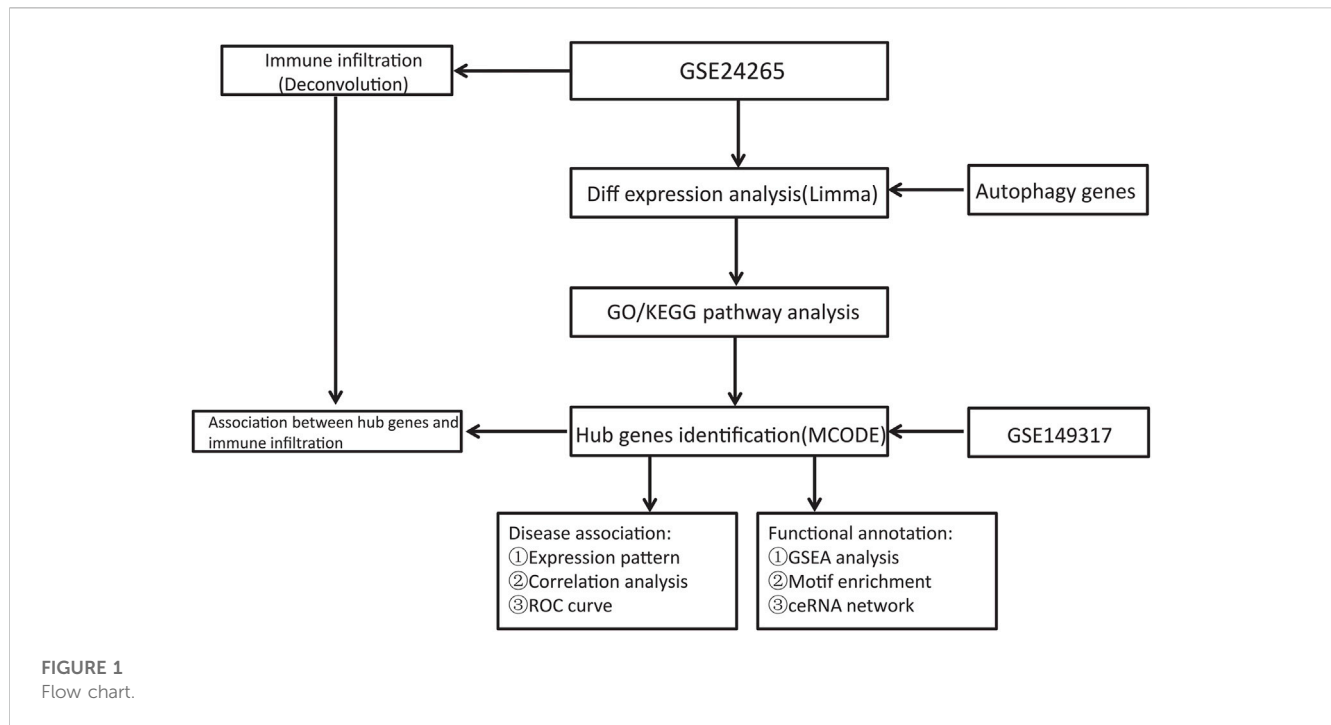
2 Materials and methods

2.1 Data download

The NCBI GEO Database (<http://www.ncbi.nlm.nih.gov/geo/>) is a repository of microarray, next-generation sequencing, and other high-throughput sequencing data (Edgar et al., 2002). The GSE24265 Series Matrix File was downloaded from the GEO public database, noted by the GPL570 annotation file, of which the expression profile data belonged to 11 samples, including the perihematomal areas, gray matters, and white matters of 7 patients in the healthy control group and 4 patients with ICH (Rosell et al., 2011). The GSE149317 Series Matrix File (only used to verify the expression level of key genes) was downloaded from the GEO public database, and the annotated File is GPL24688 (Yuan et al., 2020). The microarray data included 6 cases in the healthy control group and 6 patients in the ICH group. We used the R package limma to count the differentially expressed genes between ICH patient and healthy control samples (Ritchie et al., 2015). The screening conditions for differential genes were P -Value < 0.05 and $|\log FC| > 1$. Using the GeneCards database (<https://www.genecards.org/>) (Stelzer et al., 2016), 7236 autophagy-related genes were obtained. The relevance scores of 269 genes were greater than 3, and these genes were chosen for analysis as an autophagy gene set. Another 1139 ICH-related genes were also obtained from the database. The flow chart of this study is shown in Figure 1.

2.2 Functional annotation

The R package clusterProfiler was used to comprehensively explore the functional correlation of these differentially expressed genes (Yu et al., 2012). GO and KEGG were used for the evaluation of relevant functional categories. GO and KEGG enriched pathways with both p values and q -values less than 0.05 were considered significant pathways. To comprehensively explore the functional correlation of differentially expressed genes, we also used the Metascape database (www.metascape.org) for gene annotation (Zhou et al., 2019). GO and KEGG were used to analyze the potential pathways of the selected genes. Min overlap ≥ 3 and $p \leq 0.01$ were considered statistically significant.



2.3 Protein–protein interaction network analysis

The protein–protein interaction (PPI) information of genes was retrieved through the STRING database (Szklarczyk et al., 2021), and the confidence scores were set to ≥ 0.4 . Cytoscape software was used to visualize the results, and the gene coexpression network was obtained. The MCODE algorithm of Cytoscape identified densely connected sets of genes in the PPI network.

2.4 Analysis of immune cell infiltration

Developed by the Dviraran team in 2017, xCell is a widely used method to evaluate immune cell types in the microenvironment (Aran et al., 2017). This method integrates the strengths of gene enrichment analysis *via* deconvolution to assess 64 cell types that include multiple adaptive and innate immune cells, hematopoietic progenitor cells, epithelial cells, and extracellular stromal cells, including 48 tumor microenvironment-related cells. With the R package xCell, we analyzed the patient data to infer the relative proportion of infiltrating immune cells and performed Pearson correlation analysis on the level of immune cell infiltration. Pearson correlation analysis was used to evaluate the immune cell content and the expression level of some key genes.

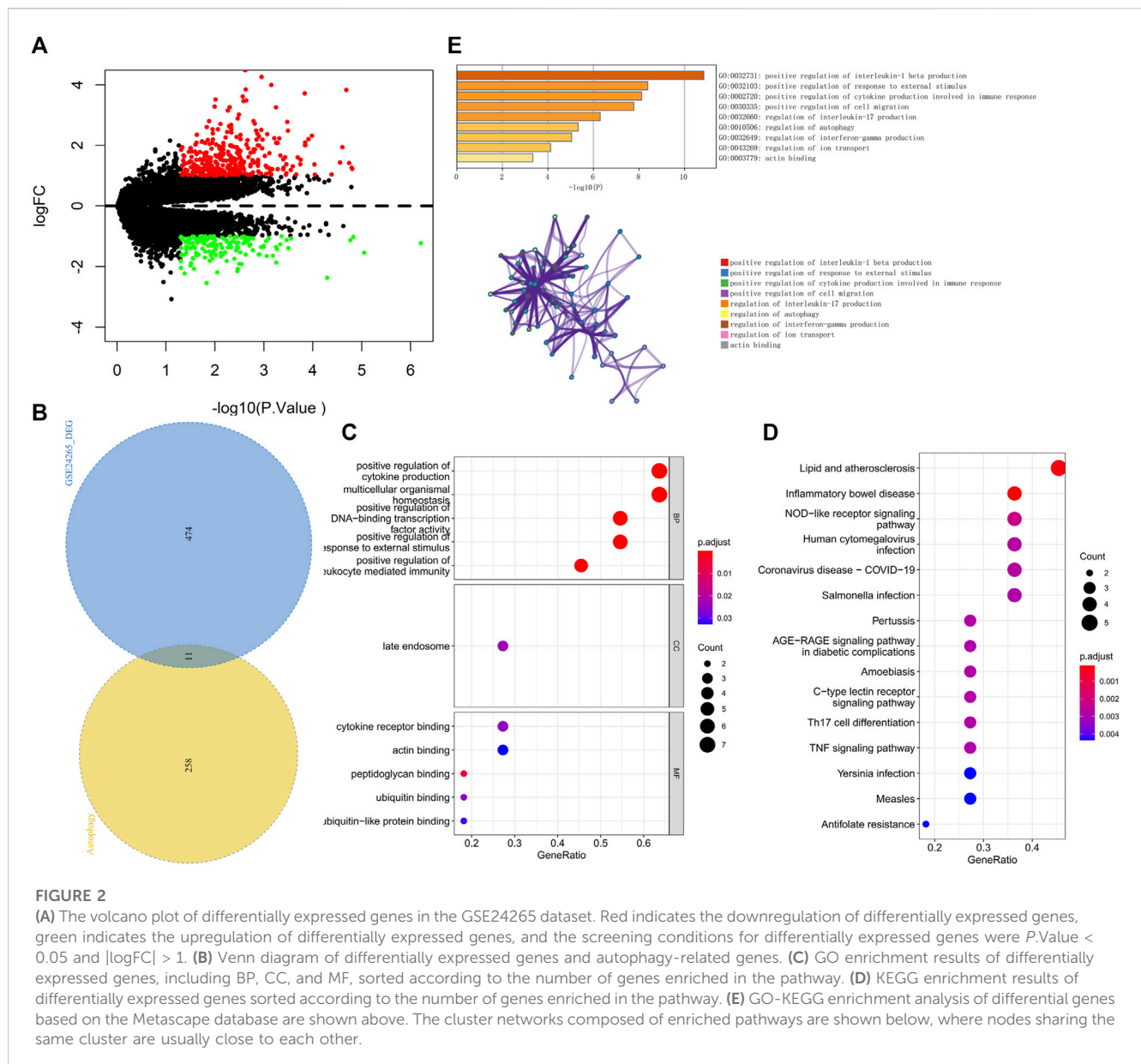
2.5 Transcription factor regulatory network analysis of key genes

The transcription initiation process of eukaryotes is very complex and often requires the assistance of various protein factors. TFs and RNA polymerase II form a transcription

initiation complex and participate in the process of transcription initiation together. TFs can be divided into two categories according to their function. The first category is universal transcription factors, which, when acting together with RNA polymerase II to form the transcription initiation complex, can start transcription at the correct position. Another category is cis-acting elements, which are sequences present in sequences flanking genes that can affect gene expression. Cis-acting elements include promoters, enhancers, regulatory sequences, and inducible elements that participate in the regulation of gene expression. The cis-acting element itself does not encode any protein but provides an action site to interact with the trans-acting factor. This analysis was mainly performed using the R package cisTarget (<https://resources.aertslab.org/cistarget/>), in which we used mm9-500bp-upstream-7species.mc9nr.feather version 1.6.0 for the Genemotif rankings database. The main TFs were predicted by the cisTarget function, when nesThreshold was 3, geneErnMethod was aprox, and geneErnMmaxRank was 5000.

2.6 Gene set enrichment analysis

According to a predefined set of genes, GSEA is a statistical procedure to rank genes according to their degree of differential expression in two types of samples and then test whether the predefined gene set is enriched at the top or bottom of the ranking list (Subramanian et al., 2005). In this study, GSEA was used to compare the discrepancies in signaling pathways between the high expression group and the low expression group and to explore the molecular mechanisms of the core genes of patients. The number of substitutions was 1000, and the substitution type was phenotype.



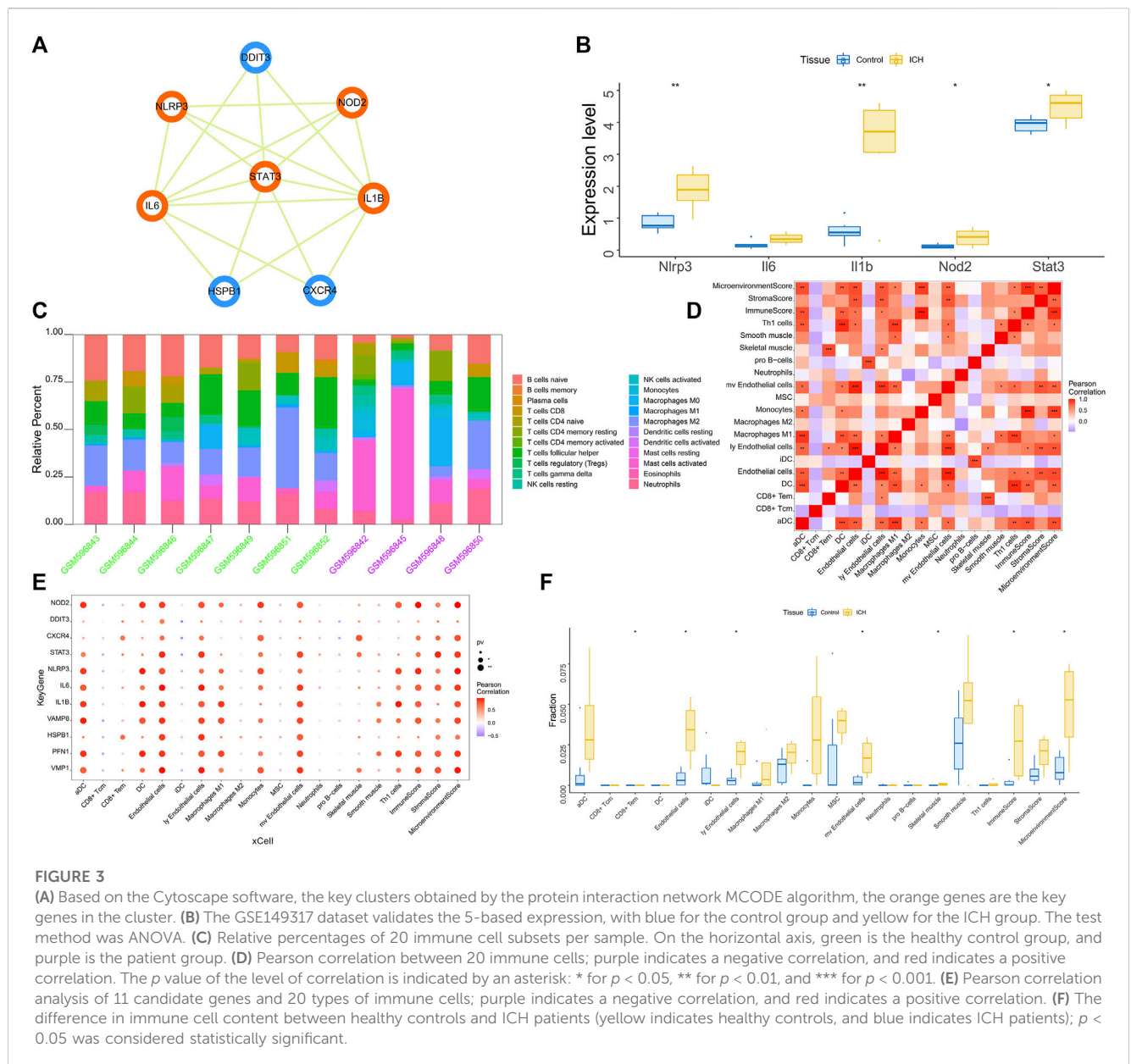
2.7 Analysis of the ceRNA network

Representing a new mode of gene expression regulation, ceRNA has attracted much attention in the academic community in recent years. Compared with the miRNA regulatory network, the ceRNA network is more elaborate and complex, involving more RNA molecules, including mRNAs, gene-coding pseudogenes, long non-coding RNAs, and miRNAs. In addition, we combined four databases, miRWalk, miRDB, TargetScan and ENCORI, to predict the interaction between key mRNAs and non-coding RNAs. Moreover, we selected coidentified targeted mRNAs for further analysis. Finally, ceRNA networks were established with the combination of mRNA-miRNA and miRNA-lncRNA interactions and visualized with Cytoscape.

3 Results

3.1 Identification of Hub genes

We downloaded the GSE24265 dataset from the NCBI GEO public database, which contained the data from a total of 11 individuals, including 7 in the healthy control group and 4 in the disease group. Through comparison with the healthy control group, we used the limma package to screen out a total of 341 upregulated genes and 144 downregulated genes in the patient samples (Figure 2A). Among them, 11 autophagy-related genes (all upregulated genes) were included (Figure 2B). Ultimately, we used these 11 autophagy-related differentially expressed genes as candidate gene sets for further analysis.

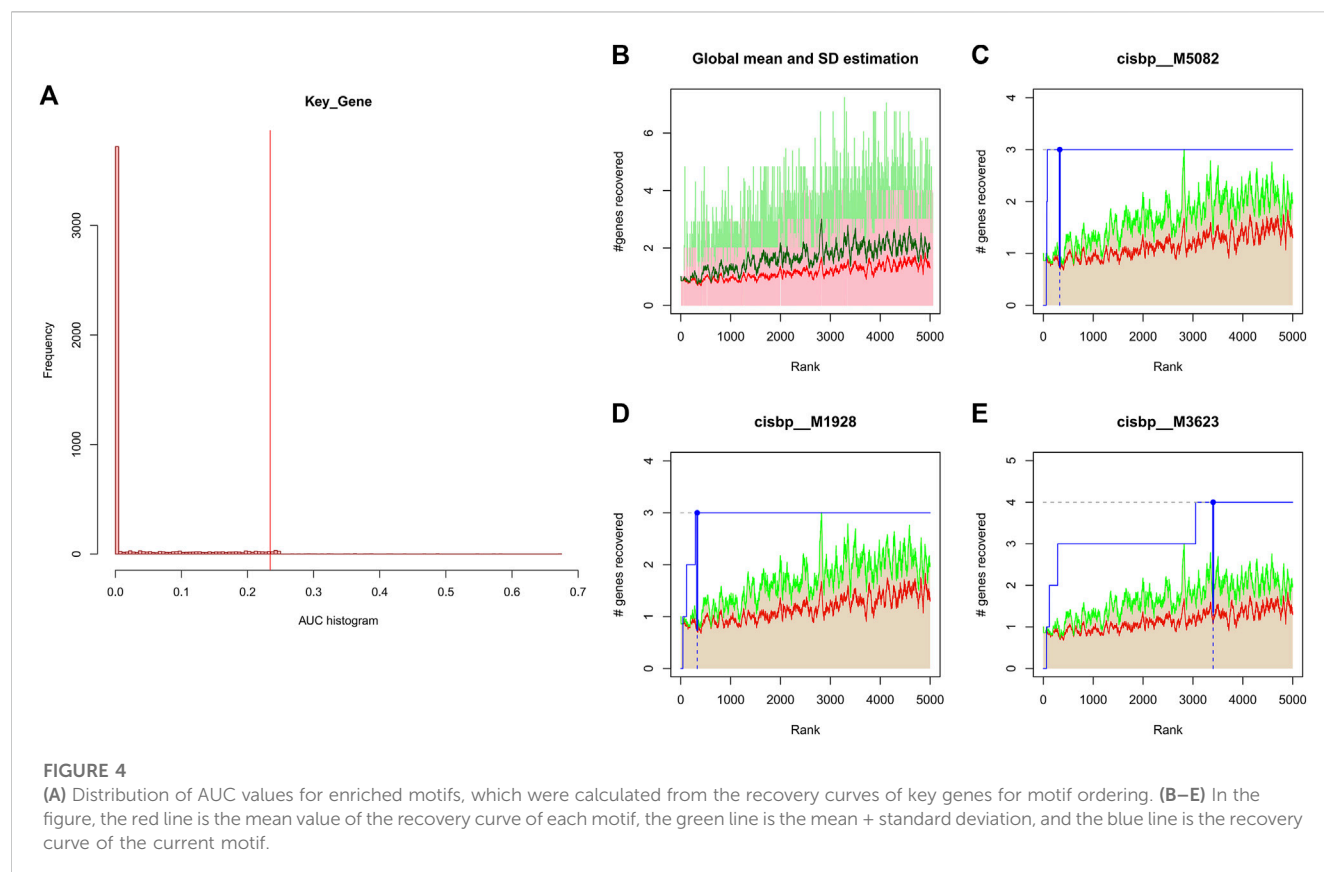


3.2 Functional enrichment analysis

We further performed pathway analysis on these 11 candidate genes. GO enrichment analysis showed that these candidate genes were mainly enriched in the positive regulation of cytokine production and cytokine receptor binding pathways (Figure 2C). KEGG enrichment analysis revealed that these candidate genes were mainly enriched in pathways such as lipid and atherosclerosis and the nucleotide-binding oligomerization domain (NOD)-like receptor signaling pathway (Figure 2D). The Metascape database was used for further pathway analysis of candidate genes. The results showed that these candidate genes were mainly enriched in positive regulation of interleukin-1 beta production, the regulation of interleukin-17 production and the regulation of autophagy pathways (Figure 2E).

3.3 Identification of key genes and ROC curve analysis

We found multiple protein interaction pairs among 11 candidate genes through the STRING online database. Moreover, five key genes, including *IL1B*, *STAT3*, *IL6*, *NOD2* and *NLRP3*, were obtained by MCODE analysis in Cytoscape (Figure 3A). Then, we analyzed the expression levels of these five key genes in the GSE149317 dataset and found that the expression levels of interleukin-1beta (*IL1B*), signal transducer and activator of transcription 3 (*STAT3*), nucleotide-binding oligomerization domain containing 2 (*NOD2*) and NOD-1-like receptor pyrin domain containing three (*NLRP3*) were significantly higher in the ICH group than in the healthy control group (Figure 3B). The area under the receiver operating characteristic curve (AUC) for the four key genes was no less than 0.75 (Supplementary Figure S1). Based on



the GSE24265 dataset, we once again analyzed the predictive power of these key genes for ICH. The results showed that the AUCs of *IL1B*, *STAT3*, *NOD2* and *NLRP3* were greater than 0.8 (Supplementary Figure S1).

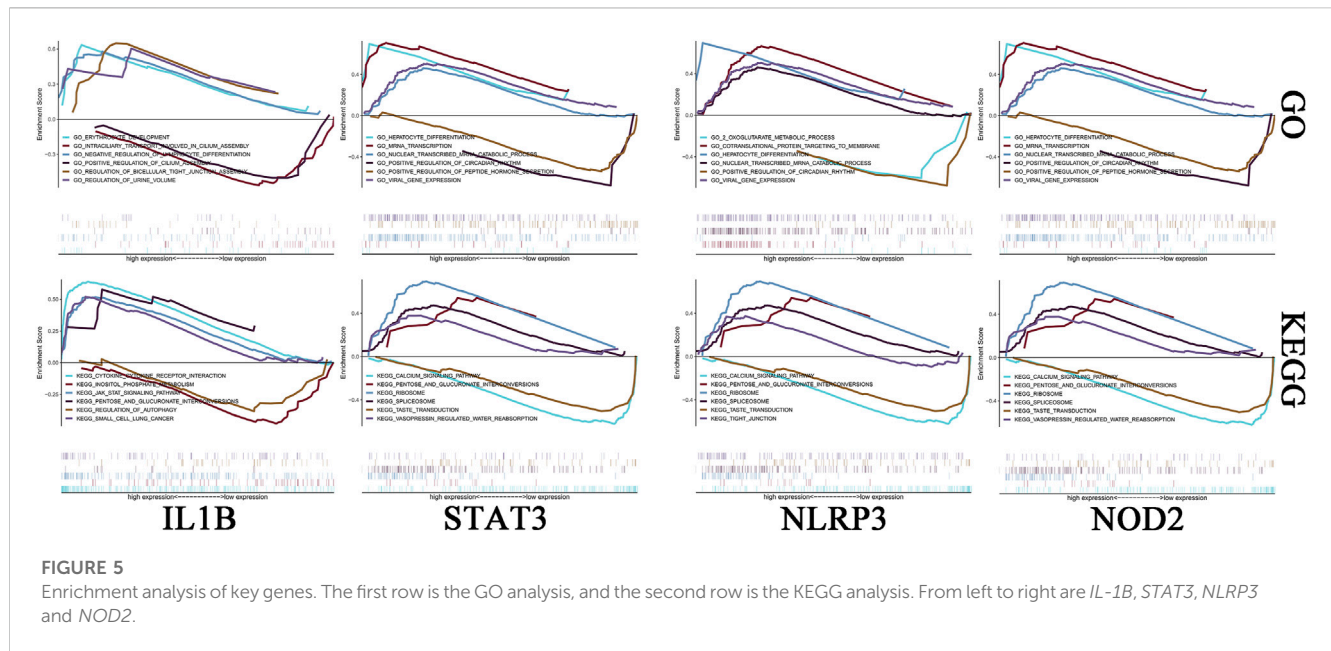
3.4 Analyses of the immune microenvironment

The immune microenvironment is mainly composed of immune-related fibroblasts, immune cells, extracellular matrix, various growth factors, inflammatory factors and special physicochemical characteristics. The immune microenvironment significantly affects the diagnosis, survival outcome and clinical severity of disease. Analyzing the relationship between core genes and immune infiltration in the GSE24265 dataset, we further explored the potential molecular mechanisms affecting disease progression. The 20 most significant immune factors in the Wilcoxon test were selected for analysis. The research results showing the proportion of immune cells and the correlation with immunity are shown in Figures 3C, D. There were multiple significant correlation pairs between the expression level of candidate genes and the level of immune infiltration (Figure 3E). In addition, the levels of endothelial cells and Ly endothelial cells in the ICH group were higher than those in the healthy controls (Figure 3F). We further explored the relationship between key genes and immune cells and found

that key genes were mostly positively correlated with immune cell infiltration levels. For example, Endothelial cells, MicroenvironmentScore, aDC, ly Endothelial cells and ImmuneScore were significantly positively correlated with 4 key genes, but CD8⁺ Tcm, iDC and pro B-cells were significantly negatively correlated with key genes (Supplementary Figure S2). We further obtained the correlations between these key genes and different immune factors from the TISIDB database, including immunomodulators, chemokines and cell receptors (Supplementary Figure S3). These data confirmed that these key genes are closely related to immune cell infiltration levels and play important roles in the immune microenvironment.

3.5 The correlation between key genes and ICH-related genes

We obtained 1,139 ICH-related pathogenic genes through the GeneCards database. Based on the GSE24265 dataset, we analyzed the expression levels of the 4 key genes and the top 20 genes in the Relevance score from GeneCards. Statistical analysis by ANOVA showed that the expression levels of these disease-related genes were significantly different between the healthy control group and the disease-related group. In addition, the expression levels of key genes were significantly correlated with the expression levels of multiple disease-related genes (Supplementary Figure S4).



3.6 Transcription factors of key genes

We applied these four key genes to the gene set for this analysis and found that they are regulated by a common mechanism including multiple transcription factors. Therefore, enrichment analysis (Figure 4), motif-TF annotation and the selection of important genes were performed for these transcription factors using accumulative recovery curves. The analysis results showed that the motif with the highest normalized enrichment score (NES: 7.70) was annotated as *cisbp_M5082*. Three genes were enriched in this motif, namely, *IL1B*, *NLRP3* and *NOD2*. We displayed all enriched motifs and corresponding transcription factors of core genes (Supplementary File S1)

3.7 GSEA of key genes

We investigated the specific signaling pathways enriched by the 4 key genes and explored the underlying molecular mechanisms by which the core genes affect the progression of ICH. Some of these highly significant pathways were selected to be displayed in detail (Figure 5). The *IL1B* gene GO enrichment pathways were ERYTHROCYTE DEVELOPMENT, INTRACILIARY TRANSPORT INVOLVED IN CILIUM ASSEMBLY, etc. The *IL1B* gene KEGG enrichment pathways were CYTOKINE-CYTOKINE RECEPTOR INTERACTION, JAK STAT SIGNALING PATHWAY, etc. The *NLRP3* gene GO enrichment pathways were COTRANSLATIONAL PROTEIN TARGETING TO MEMBRANE, 2 OXOGLUTARATE METABOLIC PROCESS, etc. The *NLRP3* gene KEGG enrichment pathways were CALCIUM SIGNALING PATHWAY, PENTOSE AND GLUCURONATE INTERCONVERSIONS. The *NOD2* gene GO enrichment pathways were CELLULAR METABOLIC COMPOUND SALVAGE, HISTONE H4 K16 ACETYLATION, etc. The *NOD2* gene KEGG enrichment pathways were

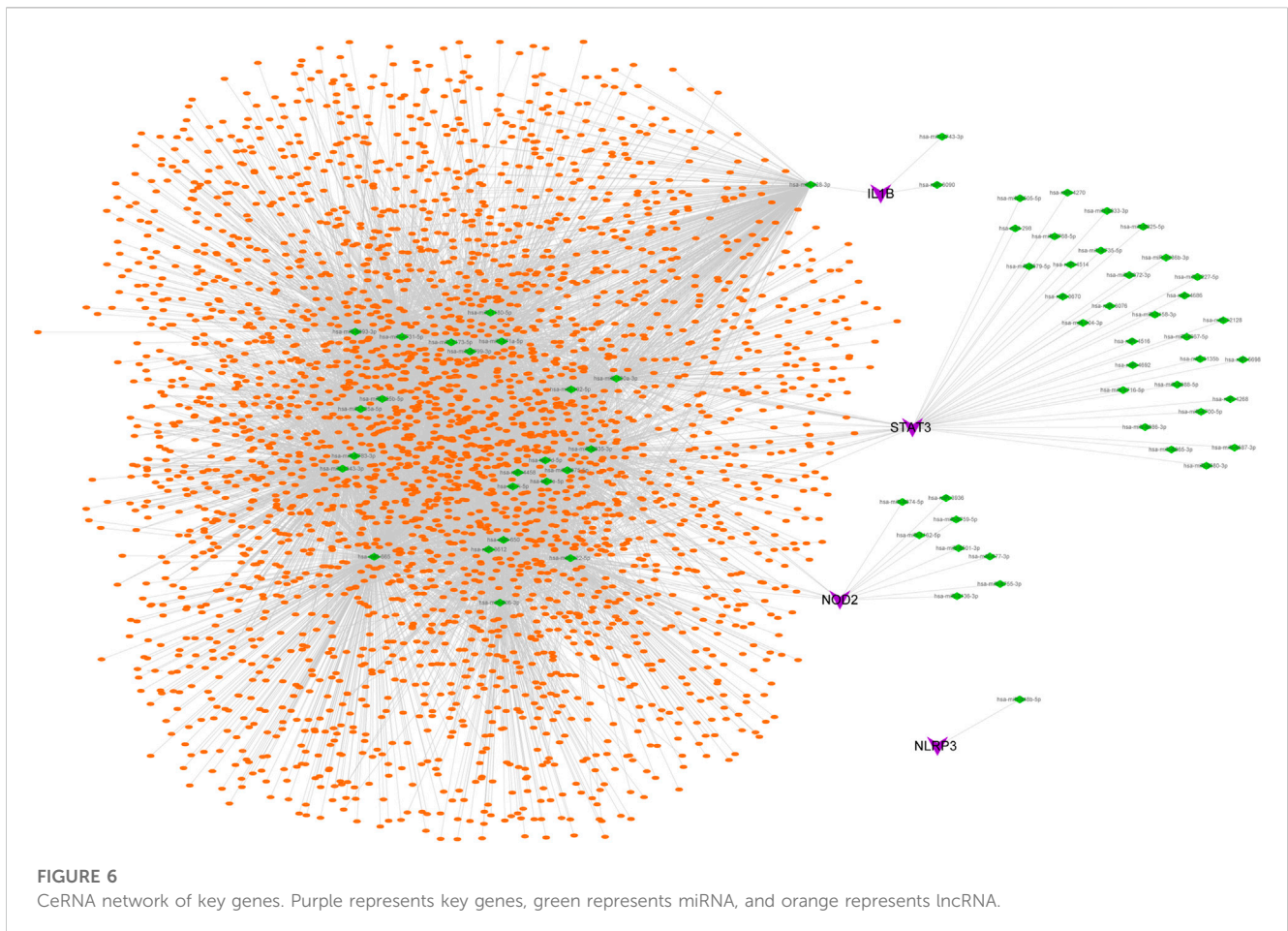
ANTIGEN PROCESSING AND PRESENTATION, AUTOIMMUNE THYROID DISEASE, etc. The *STAT3* gene GO enrichment pathways were HEPATOCYTE DIFFERENTIATION, MRNA TRANSCRIPTION, etc. The *STAT3* gene KEGG enrichment pathways were CALCIUM SIGNALING PATHWAY, PENTOSE AND GLUCURONATE INTERCONVERSIONS, etc.

3.8 Further ceRNA interaction and mining

The possible miRNAs and lncRNAs of the 4 key genes were obtained from the miRWalk database and ENCORI database, respectively. First, the four key mRNA-related mRNA-miRNA relationship pairs were extracted from the miRWalk database, but we retained only 67 mRNA-miRNA pairs (4 mRNAs and 66 miRNAs) that were validated in TargetScan or miRDB. Then, interacting lncRNAs were predicted based on these miRNAs, and a total of 8,654 pairs of interactions (24 miRNAs and 2,952 lncRNAs) were predicted. Finally, a ceRNA network was constructed by Cytoscape (V3.7) (Figure 6).

4 Discussion

Defined as a primary, non-traumatic intraparenchymal hemorrhage, ICH can lead to severe disability and is associated with a high fatality rate of 30%–50% within 6 months (Mayer and Rincon, 2005). The mortality rate of ICH within 30 days is 32%–50%, and only 28%–35% of patients who survive 3 months are able to live independently (Martini et al., 2012). As a subtype of stroke, the pathogenesis and treatment of ICH have been extensively studied, and there is still a lack of effective acute treatment. Autophagy, as an important regulatory mechanism of intracellular homeostasis, has been gradually recognized in ICH, but the regulatory effects of autophagy on intracellular homeostasis



and the immune microenvironment after ICH remain to be further explored.

Inflammation in secondary injury after ICH is mainly due to the activation of and increase in inflammatory cells and inflammatory factors (Wang, 2010). After ICH, components in the blood, including blood cells, cytokines and various immune cells, quickly cross the blood–brain barrier and enter and accumulate at the center of the injured site. This is followed by the activation of infiltrated immune cells and immune cells of the central nervous system, including the polarization of macrophages and microglia, the activation of leukocytes and astrocytes, and brain tissue damage and repair (Xue and Del Bigio, 2000; Sheth and Rosand, 2014). Immune cells include peripheral blood-derived leukocytes and macrophages, innate microglia, astrocytes, and mast cells. Many studies have shown that leukocytes, macrophages, activated microglia, and astrocytes are the main cellular mediators of secondary injury in ICH (Illanes et al., 2011). These immune cells can release cytokines, chemokines, prostaglandins, proteases, ferrous iron, and other immunologically active molecules (Hua et al., 2006). The R package X cell analysis indicated that four key genes can cause macrophages, neutrophils and CD8⁺ T cells to infiltrate the lesions during ICH and can also promote the increase in related immune factors and aggravate the inflammatory response.

Although the pathogenesis of ICH has been extensively studied at the transcriptional level, there are some limitations of these studies. Most of the

related research at the RNA level is on only the regulatory relationship between a single type of RNA, such as lncRNAs, miRNAs or mRNAs, and ICH, but little is known about the interaction of different RNAs in the development of ICH. The discovery of ceRNAs in recent years has solved this problem. ceRNA refers to RNA that has a miRNA binding site and can compete with mRNA to bind miRNA, thereby inhibiting the regulatory effect of miRNA on target genes (Ma et al., 2020). The ceRNA regulatory network rigorously integrates the mutual regulatory relationship between mRNA and non-coding RNA (ncRNA), providing significant help for the study of posttranscriptional mechanisms of diseases (Qi et al., 2015). Numerous studies have shown that the ceRNA regulatory network plays an important role in secondary injury following ICH (Liu et al., 2021; Wang et al., 2021; Yang et al., 2022). Based on multiple databases, the key gene-related ceRNA network described in this study shows miRNAs and related lncRNAs that play major regulatory roles.

This study identified *IL-1B*, *STAT3*, *NOD2* and *NLRP3* as key causative genes for secondary injury in ICH and demonstrated the critical role of autophagy in ICH. We combined two datasets, mainly using GO/KEGG analysis, immune infiltration analysis and ceRNA network construction, to screen key autophagy-related genes and analyze their mechanisms affecting ICH progression.

Recent studies have shown that cytokines, including proinflammatory cytokines and anti-inflammatory cytokines, play an important regulatory role in the course of various inflammatory-related diseases. *IL-1B*, a member of the interleukin 1 cytokine

family, is a key proinflammatory factor that plays an important role in the body's immune response and regulates inflammatory responses to brain injury (You et al., 2020). After inflammation occurs in the body, the secretion of IL-1B increases rapidly. In general, IL-1B has a proinflammatory effect in the local inflammatory response, causing vascular dilation and inducing the transfer of monocytes and neutrophils to the inflammatory site, resulting in a stress response and tissue damage (Schett et al., 2016). Our GSEA suggested that *IL-1B* was involved in the process of cytokine binding to its corresponding receptor, which also suggested that *IL-1B* plays an important role in the inflammatory response to ICH. In addition, KEGG analysis of *IL-1B* also showed enrichment of the Janus kinase-signal transducer and activator of transcription (JAK-STAT) pathway. Previously, researchers found that miRNAs/mRNAs changes in whole-blood samples for patients with ICH were important links with the JAK-STAT pathway (Cheng et al., 2020). The JAK-STAT pathway has also been associated with ICH progression in rat models (Ji et al., 2020). Our GSEA results also showed that *STAT3*, which is closely related to mRNA catabolism, is a key gene leading to ICH. The STAT protein family, which includes seven members, plays a key role in regulating cytokine-dependent inflammation and immunity. *STAT3* is considered to be the most conserved and can be activated by various factors and stimuli, such as cytokines and chemokines. *STAT3* is closely related to ischemic stroke and ischemia-reperfusion injury, and its high expression aggravates nerve damage (Zhu et al., 2021). Zhu H reported that *STAT3* activation can promote the occurrence and development of inflammation, leading to increased cerebral edema after ICH and damage to neurons around the hematoma, and *NLRP3* is a downstream molecule of *STAT* (Lee et al., 2006). In addition, the findings from mouse experiments suggest that *NLRP3* is the key to the aggravation of ICH injury caused by *STAT3* (Ji et al., 2022). Our results showed that *NLRP3* was significantly upregulated in the brain tissues of ICH patients, and the AUC of *NLRP3* was greater than 0.89, which indicates that *NLRP3* is a key gene for ICH and has strong predictive value for ICH. *NLRP3*, a member of the intracytoplasmic pattern recognition receptor NOD-like receptors (NLRs), is an important part of the innate immune system and plays an important regulatory role in the process of innate immune inflammation. *NLRP3* can sense tissue cell damage and is then activated by a variety of damage-associated molecular patterns (DAMPs) or pathogen-associated molecular patterns (PAMPs) (Mangan et al., 2018). Activated *NLRP3* protein can form the *NLRP3* inflammasome, which can cleave biologically inactive pro-IL-1B into IL-1B and exert its proinflammatory effect (Mangan et al., 2018). The last key gene identified in our analysis, *NOD2*, is also one of the main NLRs. As an important intracytoplasmic pattern recognition receptor, *NOD2* is widely involved in the recognition of immune cells and the induction of inflammatory responses (Huang et al., 2013). Activated *NOD2* receptors recruit the downstream signaling molecule receptor interacting protein 2 (RIP2), which can activate the non-canonical transcription factor nuclear factor-kappaB (NF- κ B) and then transcribe NF- κ B-dependent target genes, secreting inflammatory factors such as tumor necrosis factor- α (TNF- α) and IL-1B. Although many *NOD2* studies have focused on inflammatory bowel disease, it has been shown that *NOD2* is involved in the inflammatory response after cerebral ischemia,

triggering an excessive inflammatory response and exacerbating brain injury (Kuban et al., 2017). This study is the first to suggest that *NOD2* may be a key gene in the development of ICH. Our GSEA results suggest a high correlation of *NOD2* with ANTIGEN PROCESSING AND PRESENTATION.

5 Conclusion

In this study, the existing ICH patient data in the GEO database were analyzed by combining autophagy-related genes in the GENE database, and 11 potential pathogenic genes were finally obtained. Finally, with diagnostic and predictive value, *IL-1B*, *STAT3*, *NLRP3* and *NOD2* were obtained through PPI analysis and ROC curve analysis. Then, based on the database and R package, we found that these 4 key genes cause immune cell infiltration into ICH lesions. GSEA revealed the specific signaling pathways involved in key genes, and we explored the possibility that these pathways might influence the development of ICH. The demonstration of TFs and ceRNA networks affecting key genes provides a theoretical basis for TFs and ncRNA in the regulation of the expression of these key genes. The identification of four key genes contributes to the understanding of the mechanism of ICH and provides potential targets and directions for the clinical treatment of ICH.

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

YX: Conceptualization, Methodology, Investigation, Software, Formal analysis. CW: Methodology, Investigation, Software, Formal analysis. YZ: Methodology, Writing—Original Draft. YG: Validation, Formal analysis. JG: Writing—Review & Editing, Resources. TH: Visualization, Project administration, Funding acquisition.

Funding

The present study was supported by the National Natural Science Foundation of China (grant nos. 82001170 and 82172190).

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.2023.1032639/full#supplementary-material>

SUPPLEMENTARY FIGURE S1

The ROC curves of key genes suggest that the genes have good predictive performance for ICH. The first row is the GSE149317 dataset, and the second row is the GSE24265 dataset.

SUPPLEMENTARY FIGURE S2

(A–D) Pearson correlation between key genes and immune cells. IL-1B, STAT3, NLRP3 and NOD2 are listed in alphabetical order. (E) Pearson

correlation analysis of 4 key genes and 20 kinds of immune cells; purple indicates a negative correlation, and red indicates a positive correlation.

SUPPLEMENTARY FIGURE S3

(A–D) Pearson correlations of key genes and various immune factors; purple indicates a negative correlation, and red indicates a positive correlation. Chemokine-related genes, receptor-related genes, immunoinhibitor-related genes, MHC-related genes and immunostimulator-related genes are listed in alphabetical order. The p value of the level of correlation is indicated by an asterisk: * for $p < 0.05$, ** for $p < 0.01$, and *** for $p < 0.001$.

SUPPLEMENTARY FIGURE S4

(A) Differences in the expression of ICH disease-regulating genes; blue indicates healthy controls, and yellow indicates ICH patients. (B) The middle panel shows the Pearson correlation analysis of ICH disease-regulating genes and key genes. Blue indicates a negative correlation, and red indicates a positive correlation.

SUPPLEMENTARY FILE S1

All the enriched motifs and corresponding transcription factors of core genes are displayed in the document.

References

- Aran, D., Hu, Z., and Butte, A. J. (2017). xCell: digitally portraying the tissue cellular heterogeneity landscape. *Genome Biol.* 18, 220. doi:10.1186/s13059-017-1349-1
- Biffi, A., Bailey, D., Anderson, C. D., Ayres, A. M., Gurol, E. M., Greenberg, S. M., et al. (2016). Risk factors associated with early vs delayed dementia after intracerebral hemorrhage. *JAMA Neurol.* 73, 969–976. doi:10.1001/jamaneurol.2016.0955
- Bobinger, T., Burkhardt, P., B Huttner, H., and Manaenko, A. (2018). Programmed cell death after intracerebral hemorrhage. *Curr. Neuropharmacol.* 16, 1267–1281. doi:10.2174/1570159X15666170602112851
- Cheng, X., Ander, B. P., Jickling, G. C., Zhan, X., Hull, H., Sharp, F. R., et al. (2020). MicroRNA and their target mRNAs change expression in whole blood of patients after intracerebral hemorrhage. *J. Cereb. Blood flow metab. Off. J. Int. Soc. Cereb. Blood Flow Metab.* 40, 775–786. doi:10.1177/0271678X19839501
- Duan, X., Wen, Z., Shen, H., Shen, M., and Chen, G. (2016). Intracerebral hemorrhage, oxidative stress, and antioxidant therapy. *Oxid. Med. Cell. Longev.* 2016, 1203285. doi:10.1155/2016/1203285
- Edgar, R., Domrachev, M., and Lash, A. E. (2002). Gene expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic acids Res.* 30, 207–210. doi:10.1093/nar/30.1.207
- Feigin, V. L., Lawes, C. M., Bennett, D. A., Barker-Collo, S. L., and Parag, V. (2009). Worldwide stroke incidence and early case fatality reported in 56 population-based studies: A systematic review. *Lancet Neurol.* 8, 355–369. doi:10.1016/S1474-4422(09)70025-0
- Fu, K., Xu, W., Lenahan, C., Mo, Y., Wen, J., Deng, T., et al. (2022). Autophagy regulates inflammation in intracerebral hemorrhage: Enemy or friend? *Front. Cell Neurosci.* 16, 1036313. doi:10.3389/fncel.2022.1036313
- Hua, Y., Wu, J., Keep, R. F., Nakamura, T., Hoff, J. T., and Xi, G. (2006). Tumor necrosis factor- α increases in the brain after intracerebral hemorrhage and thrombin stimulation. *Neurosurgery* 58, 542–550; discussion 542–550. doi:10.1227/01.NEU.0000197333.55473.AD
- Huang, J., Liu, B., Yang, C., Chen, H., Eunice, D., and Yuan, Z. (2013). Acute hyperglycemia worsens ischemic stroke-induced brain damage via high mobility group box-1 in rats. *Brain Res.* 1535, 148–155. doi:10.1016/j.brainres.2013.08.057
- Huysmans, M., Buono, R. A., Skorzinski, N., Radio, M. C., De Winter, F., Parizot, B., et al. (2018). NAC transcription factors ANAC087 and ANAC046 control distinct aspects of programmed cell death in the arabidopsis columella and lateral root cap. *Plant Cell* 30, 2197–2213. doi:10.1105/tpc.18.00293
- Illanes, S., Liesz, A., Sun, L., Dalpke, A., Zorn, M., and Veltkamp, R. (2011). Hematoma size as major modulator of the cellular immune system after experimental intracerebral hemorrhage. *Neurosci. Lett.* 490, 170–174. doi:10.1016/j.neulet.2010.11.065
- Ji, X. C., Shi, Y. J., Zhang, Y., Chang, M. Z., and Zhao, G. (2020). Reducing suppressors of cytokine signaling-3 (SOCS3) expression promotes M2 macrophage polarization and functional recovery after intracerebral hemorrhage. *Front. Neurol.* 11, 586905. doi:10.3389/fneur.2020.586905
- Ji, N., Wu, L., Shi, H., Li, Q., Yu, A., and Yang, Z. (2022). VSIG4 attenuates NLRP3 and ameliorates neuroinflammation via JAK2-STAT3-A20 pathway after intracerebral hemorrhage in mice. *Neurotox. Res.* 40, 78–88. doi:10.1007/s12640-021-00456-5
- Jia, J., Wang, Z., Zhang, M., Huang, C., Song, Y., Xu, F., et al. (2020). SQR mediates therapeutic effects of H(2)S by targeting mitochondrial electron transport to induce mitochondrial uncoupling. *Sci. Adv.* 6, eaaz5752. doi:10.1126/sciadv.aaz5752
- Kim-Han, J. S., Kopp, S. J., Dugan, L. L., and Diringer, M. N. (2006). Perihematomal mitochondrial dysfunction after intracerebral hemorrhage. *Stroke* 37, 2457–2462. doi:10.1161/01.STR.0000240674.99945.4e
- Kuban, K. C., Joseph, R. M., O'Shea, T. M., Heeren, T., Fichorova, R. N., Douglass, L., et al. (2017). Circulating inflammatory-associated proteins in the first month of life and cognitive impairment at age 10 Years in children born extremely preterm. *J. Pediatr.* 180, 116–123.e1. doi:10.1016/j.jpeds.2016.09.054
- Lee, S. T., Chu, K., Sinn, D. I., Jung, K. H., Kim, E. H., Kim, S. J., et al. (2006). Erythropoietin reduces perihematomal inflammation and cell death with eNOS and STAT3 activations in experimental intracerebral hemorrhage. *J. Neurochem.* 96, 1728–1739. doi:10.1111/j.1471-4159.2006.03697.x
- Li, Z., Li, M., Shi, S. X., Yao, N., Cheng, X., Guo, A., et al. (2020). Brain transforms natural killer cells that exacerbate brain edema after intracerebral hemorrhage. *J. Exp. Med.* 217, e20200213. doi:10.1084/jem.20200213
- Li, J., Wu, X., He, Y., Wu, S., Guo, E., Feng, Y., et al. (2021). PINK1 antagonizes intracerebral hemorrhage by promoting mitochondrial autophagy. *Ann. Clin. Transl. neurology* 8, 1951–1960. doi:10.1002/acn3.51425
- Liu, Z., Wu, X., Yu, Z., and Tang, X. (2021). Reconstruction of circRNA-miRNA-mRNA associated ceRNA networks reveal functional circRNAs in intracerebral hemorrhage. *Sci. Rep.* 11, 11584. doi:10.1038/s41598-021-91059-9
- Ma, N., Tie, C., Yu, B., Zhang, W., and Wan, J. (2020). Identifying lncRNA-miRNA-mRNA networks to investigate Alzheimer's disease pathogenesis and therapy strategy. *Aging* 12, 2897–2920. doi:10.18632/aging.102785
- Mangan, M. S. J., Olhava, E. J., Roush, W. R., Seidel, H. M., Glick, G. D., and Latz, E. (2018). Targeting the NLRP3 inflammasome in inflammatory diseases. *Nat. Rev. Drug Discov.* 17, 588–606. doi:10.1038/nrd.2018.97
- Martini, S. R., Flaherty, M. L., Brown, W. M., Haverbusch, M., Comeau, M. E., Sauerbeck, L. R., et al. (2012). Risk factors for intracerebral hemorrhage differ according to hemorrhage location. *Neurology* 79, 2275–2282. doi:10.1212/WNL.0b013e318276896f
- Mayer, S. A., and Rincon, F. (2005). Treatment of intracerebral haemorrhage. *Lancet Neurol.* 4, 662–672. doi:10.1016/S1474-4422(05)70195-2
- Moujalled, D., Strasser, A., and Liddell, J. R. (2021). Molecular mechanisms of cell death in neurological diseases. *Cell Death Differ.* 28, 2029–2044. doi:10.1038/s41418-021-00814-y
- Ohsumi, Y. (2014). Historical landmarks of autophagy research. *Cell Res.* 24, 9–23. doi:10.1038/cr.2013.169
- Qi, X., Zhang, D. H., Wu, N., Xiao, J. H., Wang, X., and Ma, W. (2015). ceRNA in cancer: possible functions and clinical implications. *J. Med. Genet.* 52, 710–718. doi:10.1136/jmedgenet-2015-103334
- 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
- Rosell, A., Vilalta, A., García-Berrocó, T., Fernández-Cadenas, I., Domingues-Montanari, S., Cuadrado, E., et al. (2011). Brain perihematoma genomic profile following spontaneous human intracerebral hemorrhage. *PloS one* 6, e16750. doi:10.1371/journal.pone.0016750
- Schett, G., Dayer, J. M., and Manger, B. (2016). Interleukin-1 function and role in rheumatic disease. *Nat. Rev. Rheumatol.* 12, 14–24. doi:10.1038/nrrheum.2016.166

- Sheth, K. N., and Rosand, J. (2014). Targeting the immune system in intracerebral hemorrhage. *JAMA neurol.* 71, 1083–1084. doi:10.1001/jamaneurol.2014.1653
- Shi, H., Wang, J., Wang, J., Huang, Z., and Yang, Z. (2018). IL-17A induces autophagy and promotes microglial neuroinflammation through ATG5 and ATG7 in intracerebral hemorrhage. *J. Neuroimmunol.* 323, 143–151. doi:10.1016/j.jneuroim.2017.07.015
- Stelzer, G., Rosen, N., Plaschkes, I., Zimmerman, S., Twik, M., Fishilevich, S., et al. (2016). The GeneCards suite: From gene data mining to disease genome sequence analyses. *Curr. Protoc. Bioinforma.* 54, 1. doi:10.1002/cpbi.5
- Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., et al. (2005). Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U. S. A.* 102, 15545–15550. doi:10.1073/pnas.0506580102
- 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
- Wang, L., Tian, M., and Hao, Y. (2020). Role of p75 neurotrophin receptor in neuronal autophagy in intracerebral hemorrhage in rats through the mTOR signaling pathway. *Cell cycle (Georget. Tex)* 19, 376–389. doi:10.1080/15384101.2019.1711318
- Wang, H., Cao, X., Wen, X., Li, D., Ouyang, Y., Bao, B., et al. (2021). Transforming growth factor- β 1 functions as a competitive endogenous RNA that ameliorates intracranial hemorrhage injury by sponging microRNA-93-5p. *Mol. Med. Rep.* 24, 499. doi:10.3892/mmr.2021.12138
- Wang, J. (2010). Preclinical and clinical research on inflammation after intracerebral hemorrhage. *Prog. Neurobiol.* 92, 463–477. doi:10.1016/j.pneurobio.2010.08.001
- Xue, M., and Del Bigio, M. R. (2000). Intracerebral injection of autologous whole blood in rats: Time course of inflammation and cell death. *Neurosci. Lett.* 283, 230–232. doi:10.1016/s0304-3940(00)00971-x
- Yang, C., Wu, J., Lu, X., Xiong, S., and Xu, X. (2022). Identification of novel biomarkers for intracerebral hemorrhage via long noncoding RNA-associated competing endogenous RNA network. *Mol. omics* 18, 71–82. doi:10.1039/d1mo00298h
- You, Y., Borgmann, K., Edara, V. V., Stacy, S., Ghorpade, A., and Ikezu, T. (2020). Activated human astrocyte-derived extracellular vesicles modulate neuronal uptake, differentiation and firing. *J. Extracell. Vesicles* 9, 1706801. doi:10.1080/20013078.2019.1706801
- Yu, G., Wang, L. G., Han, Y., and He, Q. Y. (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. *Omics a J. Integr. Biol.* 16, 284–287. doi:10.1089/omi.2011.0118
- Yuan, J. J., Chen, Q., Xiong, X. Y., Zhang, Q., Xie, Q., Huang, J. C., et al. (2020). Quantitative profiling of oxylipins in acute experimental intracerebral hemorrhage. *Front. Neurosci.* 14, 777. doi:10.3389/fnins.2020.00777
- Zhang, C., Qian, C., Yang, G., Bao, Y. X., and Qian, Z. M. (2021). Hepcidin inhibits autophagy in intracerebral hemorrhage models *in vitro* and *in vivo*. *Mol. Cell. Neurosci.* 111, 103589. doi:10.1016/j.mcn.2021.103589
- Zhou, Y., Zhou, B., Pache, L., Chang, M., Khodabakhshi, A. H., Tanaseichuk, O., et al. (2019). Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat. Commun.* 10, 1523. doi:10.1038/s41467-019-09234-6
- Zhu, H., Jian, Z., Zhong, Y., Ye, Y., Zhang, Y., Hu, X., et al. (2021). Janus kinase inhibition ameliorates ischemic stroke injury and neuroinflammation through reducing NLRP3 inflammasome activation via JAK2/STAT3 pathway inhibition. *Front. Immunol.* 12, 714943. doi:10.3389/fimmu.2021.714943

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

