

---

# Identifying the shared genes and their related microRNAs, metabolites, and pathways in ischemic stroke and epilepsy

---

Received: 8 August 2025

---

Accepted: 4 February 2026

---

Published online: 10 February 2026

Cite this article as: Chen Y., Man S., Li Q. *et al.* Identifying the shared genes and their related microRNAs, metabolites, and pathways in ischemic stroke and epilepsy. *Sci Rep* (2026). <https://doi.org/10.1038/s41598-026-39299-5>

**Yu Chen, Shuhong Man, Qinfeng Li, Yuelong Ji, Biwen Peng, Yansheng Ding & Jian Xu**

---

We are providing an unedited version of this manuscript to give early access to its findings. Before final publication, the manuscript will undergo further editing. Please note there may be errors present which affect the content, and all legal disclaimers apply.

If this paper is publishing under a Transparent Peer Review model then Peer Review reports will publish with the final article.

**Identifying the shared genes and their related microRNAs, metabolites, and pathways in ischemic stroke and epilepsy****Running title:** Common Mechanisms of IS and Epilepsy

Yu Chen<sup>1#</sup>, Shuhong Man<sup>2#</sup>, Qinfeng Li<sup>3</sup>, Yuelong Ji<sup>4</sup>, Biwen Peng<sup>5\*</sup>,  
Yansheng Ding<sup>6\*</sup>, Jian Xu<sup>7\*</sup>

<sup>1</sup>Department of Clinical Laboratory, Weifang Maternal and Child Health Hospital, Weifang, 261011, Shandong, China, mengcandy18@163.com

<sup>2</sup>Department of Obstetrics and Gynecology, Weifang People ' s Hospital, Weifang, 261000, Shandong, China, qifeng2003@126.com

<sup>3</sup>Department of Clinical Laboratory, Weifang Maternal and Child Health Hospital, Weifang, 261011, Shandong, China, Liqinfeng1974@163.com

<sup>4</sup>School of Public Health, Peking University, Beijing, 100191, China, yuelong.ji@bjmu.edu.cn

<sup>5</sup>Department of Physiology, School of Basic Medical Sciences, Wuhan University, Wuhan, 430071, China, pengbiwen@whu.edu.cn

<sup>6</sup>Department of Clinical Laboratory, Weifang Maternal and Child Health Hospital, Weifang, 261011, Shandong, China, 1548895227@qq.com

<sup>7</sup>Department of Clinical Laboratory, Weifang Maternal and Child Health Hospital, Wuhan University School of Basic Medical Sciences-Weifang Children's Neurological Diseases and Innovation Transformation Joint Research Center, Weifang, 261011, Shandong, China, ydukongjian@163.com

**#Co-first authors**

**\*Corresponding authors:** Jian Xu, Yansheng Ding, Biwen Peng

Email: ydukongjian@163.com, 1548895227@qq.com,  
pengbiwen@whu.edu.cn

ORCID: 0000-0002-3087-824X

Address: Department of Clinical Lab, Weifang Maternal and Child Health Hospital, No. 12007, Yingqian Street, High-tech Zone, Weifang, Shandong Province, 261011, China

**Highlights:**

Shared IS and epilepsy genes, including IL10RA, CD2, and C3AR1, were identified.

IL10RA, CD2, and C3AR1 had high diagnostic efficacy for IS and epilepsy.

hsa-let-7b-5p was predicted as an upstream miRNA of C3AR1.

C3AR1 may contribute to epilepsy via interaction with acetylcholine.

**Abstract**

**Background:** This study aimed to identify shared genes between ischemic stroke (IS) and epilepsy and explore underlying mechanisms.

**Methods:** Transcriptomic datasets from the GEO database were analyzed using differential expression and weighted gene co-expression network analysis (WGCNA). Hub-shared genes were

identified through protein-protein interaction networks, ROC analysis, and expression validation. Upstream miRNAs were predicted. Additionally, untargeted plasma metabolomics was performed on children with epilepsy and healthy controls, followed by differential metabolite analysis and metabolic pathway construction.

**Results:** WGCNA revealed 594 epilepsy-related and 2,623 IS-related DEGs, with 38 shared DEGs identified, including IL10RA, CD2, and C3AR1. These genes showed high diagnostic value, with their AUC value  $> 0.66$  in both training and validation datasets. Additionally, hsa-let-7b-5p was predicted to target C3AR1. Metabolomics identified 139 differential metabolites, and C3AR1 was implicated in synaptic vesicle cycle, taste transduction, and nicotine addiction pathways via acetylcholine.

**Conclusions:** The shared genes, especially C3AR1 may be a key regulator in the development IS and epilepsy, showing potential as a biomarker for both diseases. However, its diagnostic efficacy requires further clinical validation. Given the complexity of these diseases, future research may focus on identifying a panel of biomarkers rather than relying on a single gene.

**Keywords:** ischemic stroke; epilepsy; shared genes; microRNA; metabolites

## Declarations

***Ethics approval and consent to participate***

The study was approved by Weifang Maternal and Child Health Hospital, and the guardians of children with epilepsy who provided clinical data signed written informed consent. Experimental procedures involving humans were conducted under the World Medical Association Declaration of Helsinki (2000).

***Consent for publication***

Not applicable.

***Availability of data and materials***

Data is provided within the manuscript or supplementary information file. Further enquiries can be directed to the corresponding author.

***Competing interests***

The authors declare that they have no competing interests.

***Funding***

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by grants from the China Postdoctoral Science Foundation (Grant No. 2018M642618), the National Natural Science Foundation of China (Grant No. 81401230), the Natural Science Foundation of Shandong Province (Grant No. ZR2019BH056), and Shandong Provincial Medical and Health

Science and Technology Project[Grant No.202411001099].

### ***Authors' contributions***

Yu Chen and Shuhong Man contributed equally to the conceptualization and design of the study. Yu Chen performed the bioinformatics analysis and data interpretation. Shuhong Man was responsible for drafting the manuscript, supervising the data analysis, and conducting additional experiments and analyses as per the reviewer's comments. Qingfeng Li and Yuelong Ji assisted in the computational analysis and contributed to the discussion of results. Biwen Peng, Yansheng Ding, and Jian Xu provided critical revisions to the manuscript and helped in the integration of the results. Biwen Peng, Yansheng Ding, and Jian Xu are the corresponding authors, guiding the overall study design, manuscript preparation, and the revision process. All authors read and approved the final manuscript.

### ***Acknowledgments***

None.

## Introduction

Stroke is a major cause of disability worldwide and seriously affects living standards and quality of life. There are two types of stroke: ischemic stroke (IS) and hemorrhagic stroke. IS is the most common type of stroke and accounts for approximately 80% of all strokes <sup>1</sup>. Cardioembolic stroke is a subtype in which cerebral thromboembolic ischemia is caused by cardiogenic emboli, which accounts for 14%-30% of IS and is more likely to cause disability <sup>2</sup>. Stroke is recognized as one of the most common causes of epilepsy, with post-stroke epilepsy (PSE) accounting for 11% of all epilepsies in the adult population <sup>3,4</sup>. Epilepsy is a common neurological disorder characterized by repeated seizures. It affects approximately 65 million people worldwide and imposes a great burden on individuals and health systems <sup>5,6</sup>. To date, managing PSE does not completely match the treatment of stroke and epilepsy caused by other reasons, and neurologists lack guidance on managing PSE, such as when and how to treat patients. Therefore, investigating the pathogenesis and association between stroke and epilepsy is necessary.

The risk of seizures is high in many post-stroke patients, and the incidence of PSE has been increasing over the last decades <sup>7,8</sup>. Although its incidence after hemorrhagic stroke is higher than that after IS, the overall seizure burden in the IS group is greater because of the higher frequency of IS <sup>9</sup>. Common

pathophysiological mechanisms are assumed to underlie stroke and epilepsy <sup>10</sup>. Following a stroke, various inflammatory mediators, such as prostaglandins, cytokines, chemokines, damage-associated molecular patterns (DAMPs), and complements are released to repair brain damage. When sustained over time, neuroinflammation causes neuronal and astroglial dysfunctions, resulting in altered synaptic transmission, neuronal loss, hyperexcitability, and abnormal neurogenesis <sup>11</sup>. These mechanisms contribute to epilepsy development <sup>12-14</sup>. Moreover, some less common factors that may cause seizures, such as the central nervous system (CNS) infections, systemic metabolic abnormalities, such as hypocalcemia and hypo/hyperglycemia, and benzodiazepines and barbiturates, may also coexist in patients with strokes, suggesting a similar etiology of seizures and strokes <sup>15</sup>. Furthermore, early PSE is associated with high levels of the neural cell adhesion molecule and low levels of tumor necrosis factor-R1 <sup>16</sup>. A recent study has identified hub genes such as TPGS2, TMCC3, GADD45B, and KCNJ2 as key players in both IS and epilepsy, suggesting that these genes may play a key role in the co-pathogenesis of IS and epilepsy <sup>17</sup>. Although a strong association exists between stroke and epilepsy from the clinical and epidemiological perspectives, the shared genetic and metabolic mechanisms in IS and epilepsy are largely unknown.

In this study, we downloaded the IS- and epilepsy-related datasets from the Gene Expression Omnibus (GEO) database and

analyzed the shared genes of the two diseases using differential expression analysis and weighted gene co-expression network analysis (WGCNA). We identified hub-shared genes via protein-protein interaction (PPI) analysis, receiver-operating characteristic (ROC) curve analysis, and differential expression validation using validation datasets. Moreover, we predicted the miRNAs upstream of the hub-shared genes. Furthermore, we performed nontargeted metabolomics analysis of the plasma samples from children with epilepsy and healthy controls, analyzed key differential metabolites, and constructed a metabolic pathway network of the hub-shared genes. A flowchart of the study is shown in Supplementary Figure 1. Our study aimed to explore the shared genetic and metabolic mechanisms between IS and epilepsy, offering new insights for identifying potential targets that could be useful in treating PSE.

## Methods

### *Data acquisition and preprocessing*

The IS-related gene expression profiles GSE58294 (training dataset) and GSE16561 (validation dataset), as well as the epilepsy-related gene expression profiles GSE143272 (training dataset), GSE16969 (validation dataset) and GSE205661 (validation dataset), were downloaded from the NCBI GEO database <sup>18</sup>. The detailed information of these datasets is listed in Table 1.

The mRNA probe expression matrix of each dataset and the annotation file of the detection platform were downloaded. Probes

were individually converted into gene symbols. Probes that did not match the gene symbols were removed, and the average expression of multiple probes mapped to the same gene symbol was selected.

#### *Differential expression analysis*

Based on the training datasets (GSE143272 epilepsy dataset and GSE58294 IS dataset), the differentially expressed genes (DEGs) in the epilepsy and IS samples were compared with their respective control samples using the R limma package (version 3.48.3) <sup>19</sup>. The P value was adjusted by Benjamini-Hochberg method. The adjusted P (P.adj) value < 0.05 and |log fold change (FC)| > 0.263 (that was FC > 1.20 for upregulation and FC < 0.8 for downregulation)<sup>20</sup> were set as the cutoff values for DEG screening.

#### *Identifying disease-related genes using WGCNA*

The disease-related key modules and module genes were selected by conducting WGCNA on the DEGs identified from peripheral blood samples in the training datasets of two diseases to construct gene co-expression modules using the R WGCNA package (version 1.71) <sup>21</sup>. Briefly, an appropriate soft-thresholding power was selected to obtain a scale-free network. The Pearson correlation coefficient between the two diseases is determined through the use of formula 1:  $S_{ij} = \left| \frac{1 + \text{cor}(x_i, y_i)}{2} \right|$ , where  $i$  and  $j$  are the expression levels of the  $i$ -th gene and the  $j$ -th gene respectively.

Then, formula 2 and 3 are used to convert the gene expression similarity matrix into an adjacency matrix, with the network type

being signed, where  $\beta$  is the soft threshold, which is the square of the Pearson correlation coefficient  $\beta$  for each pair of genes. The adjacency matrix is converted into a topology matrix, and the topology overlap (Topological overlap measure, TOM) is used to describe the degree of association between genes:

$$a_{ij} = \left| \frac{1 + \text{cor}(x_i + y_i)}{2} \right|^\beta \quad \text{and} \quad = \frac{\sum_{u \neq ij} a_{iu} a_{uj} + a_{ij}}{\min(\sum_u a_{iu} + \sum_u a_{ju}) + 1 - a_{ij}}.$$

Gene co-expression modules were determined, and genes with similar expression patterns were clustered. Using formula 4 for clustering analysis, 1-TOM represents the degree of difference between genes  $i$  and  $j$ . Genes are hierarchically clustered, followed by module identification using dynamic cutting tree methods. The most representative gene in each module is called a feature vector gene, or ME for short. It represents the overall level of gene expression within the module and is the first principal component in each module. Here,  $i$  denotes the gene in module  $q$ , and  $l$  denotes the chip sample in module  $q$ :  $ME = \text{princomp}(x_{il}^{(q)})$ .

Module Identity (MM) is used to measure the association degree of a gene with a specific module. This is achieved by calculating the Pearson correlation coefficient between the expression profile of a gene in all samples and the eigenvector gene

$ME^q$  of module  $q$ . This process can be completed using formula 5:  $MM_i^q = \text{cor}(x_i, ME^q)$ . Here,  $x_i$  represents the expression profile of the  $i$ -th gene,  $ME^q$  is the eigenvector gene expression profile of module  $q$ , and  $MM_i^q$  is the Module Identity value of

gene  $i$  in module  $q$ . When  $MM_i^q$  equals 0, it indicates that the gene does not belong to module  $q$ ; if the value of  $MM_i^q$  approaches +1 or -1, it suggests that gene  $i$  has a stronger correlation with module  $q$ . A positive sign indicates a positive correlation between gene  $i$  and module  $q$ , while a negative sign indicates a negative correlation.

Finally, it is analyzed and visualized. The cluster tree was cut into branches to define the modules using a dynamic tree cut, and each module was assigned a different color for visualization. Each module contained at least 30 genes. The epilepsy- or IS-related key modules were identified using hierarchical clustering and correlation analysis of the module eigengene values and clinical traits. Disease-related DEGs were identified from key epilepsy- and IS-related modules.

#### *Analyzing the shared genes of epilepsy and IS*

The overlapping DEGs in the epilepsy- and IS-related key modules were considered as the shared genes of two diseases and were obtained using Venn diagram analysis using the R VennDetail package (version 1.2.0).

#### *Functional enrichment analysis for shared genes*

The shared genes of two diseases were imported for Gene Ontology (GO) functions and Kyoto Encyclopedia of Genes and Genomes (KEGG)<sup>22-24</sup> pathway enrichment analyses using the R clusterProfiler package (version 4.44)<sup>25</sup>. GO describes the overrepresented biological functions from three aspects: molecular

function (MF), cellular component (CC), and biological processes (BP).

#### *PPI network analysis and hub gene identification*

The PPI relationships of the shared genes of two diseases were analyzed using the STRING (version 11.0) database <sup>26</sup>, followed by the construction of a PPI network using Cytoscape (version 3.9.2) <sup>27</sup>. The interaction score was set to 0.15. The top five hub genes were analyzed using the cytoHubba plug-in <sup>28</sup>, and four topological analysis algorithms, including the maximal neighborhood component (MNC), maximal clique centrality (MCC), degree, and edge-percolated component (EPC). MCC calculated the importance of nodes by identifying the maximal cliques in the network. MNC evaluated node importance based on the strength of interactions with neighboring nodes. Degree assessed node importance by calculating its number of connections. EPC identified key edges and nodes in the network through edge percolation. These algorithms were widely utilized for selecting hub genes from PPI networks <sup>29,30</sup>. The intersecting genes of the top five genes from each topological analysis algorithm were considered candidate hub genes for the two diseases.

#### *Diagnostic efficacy evaluation of the hub-shared genes*

To evaluate the diagnostic efficacy of the candidate hub-shared gene, ROC curves were plotted using the R pROC package (version 1.18.0) <sup>31</sup> based on the expression data of the hub-shared genes in all the enrolled datasets. The higher the value of the area under the

curve (AUC), the stronger the diagnostic efficacy was.

#### *Validating the differential expression of the candidate hub-shared genes*

Using the validation datasets of IS (GSE16561) and epilepsy (GSE16969 and GSE205661), the differential expression of the candidate hub-shared genes in the two diseases was validated. Differential expression analysis of the candidate hub-shared genes was performed using Student's t-test. Genes with the same expression trends in the disease samples in the training and validation datasets were considered the hub-shared genes.

#### *PPI network of the hub-shared genes*

The possible mechanisms of the hub-shared genes were further studied by performing its PPI analysis with 20 interacting genes using the GeneMANIA database <sup>32</sup>. Thus, their co-localization, shared protein domains, co-expression, and pathways were explored.

#### *Prediction of the upstream miRNAs of the hub-shared genes*

The common miRNAs associated with the two diseases were identified using the Human microRNA Disease Database (HMDD) <sup>33</sup>. Moreover, miRNAs upstream of hub-shared genes were predicted using the miRWalk database, and an miRNA-gene network was established. Disease-related miRNAs of the hub-shared genes were identified using the intersection analysis of the disease-related miRNAs and upstream miRNAs of the hub-shared genes.

#### *Collection of clinical samples for nontargeted metabolomics*

*analysis*

Fourteen plasma samples from children with epilepsy and 20 plasma samples from healthy controls were collected from Weifang Maternal and Child Health Hospital. All patients with epilepsy met the diagnostic criteria for epilepsy. Exclusion criteria included: 1) other neurological diseases or genetic disorders; 2) somatic or psychiatric diseases that could potentially affect metabolism; 3) factors such as infections or immune deficiencies that significantly impact metabolism. All children in this study had not used any antiepileptic drugs, central nervous system medications, or immunosuppressants during the treatment process. The study was approved by Weifang Maternal and Child Health Hospital, and the guardians of children signed written informed consent.

*Nontargeted metabolomics analysis and differential metabolite identification*

Metabolomics analyses were conducted using ultra-high performance liquid chromatography (1290 Infinity LC, Agilent Technologies) coupled to a quadrupole time-of-flight (AB Sciex TripleTOF 6600) at Beijing Allwogene Technology Co., Ltd. (Beijing, China). In detail, the samples were added to cold methanol/acetonitrile/H<sub>2</sub>O (2:2:1) solution, shaken and extracted for 30 min. Then, the samples were centrifuged at 14000 g at 4°C for 20 min, and the supernatant were subjected to vacuum drying. The residue was dissolved in 100 µL acetonitrile solution and centrifuged at 14000 g at 4°C for 10 min, and the supernatant was

collected for liquid chromatography-mass spectrometry (LC-MS) analysis. The LC conditions were as follows: the Vanquish ultrahigh-performance liquid chromatograph (UHPLC) (Thermo) with Waters, ACQUITY UPLC BEH Amide 1.7  $\mu$ m, 2.1 mm  $\times$  100 mm column was used, with the column temperature of 25°C, flow rate of 0.3 mL/min, and sample size of 2  $\mu$ L. Mobile phase A consisted of H<sub>2</sub>O, 25 mM ammonium acetate, and 25 mM ammonia water. Mobile phase B was acetonitrile. The solvent gradient was set as follows: 98% B, 1.5 min; 98-2% B, 12 min; 2% B, 14 min; 2-98% B, 14.1 min; 98% B, 17 min. The samples were placed in a 4°C automatic injector during the entire analysis process. The quality control (QC) samples were inserted into the sample queue to monitor and evaluate the stability of the system and the reliability of the experimental data. Using the QExactive mass spectrometer, the primary and secondary spectra of samples were collected. This mass spectrometer provided electrospray ionization (ESI) positive and negative ion detection modes. ESI source conditions were as follows: the ion source gas 1: 60, ion source gas 2: 60, curtain gas (CUR): 30 psi, source temperature: 600°C; ion spray voltage floating (ISVF):  $\pm$  5500 V, primary mass-charge ratio detection range: 80-1200 Da, resolution: 60000, accumulation time: 100 ms, scanning range of secondary stage: 70-1200 Da, resolution: 30000, scan accumulation time: 50 ms, and dynamic exclusion time: 4 s. The raw data were subjected to peak alignment, retention time correction and peak area extraction using XCMS software.

Multivariate statistical analyses were performed. Subsequently, the differential metabolites were identified with the cutoff values of  $p < 0.05$  (Student's t-test),  $FC > 1.5$  or  $< 0.67$ <sup>34,35</sup>, and variable importance in the projection (the first principal component of orthogonal partial least squares discriminant analysis)  $> 1$ .

#### *Pathway enrichment analysis for the hub-shared genes and differential metabolites*

To clarify the biological significance of the hub-shared genes and differential metabolites, the pathways significantly enriched by the hub-shared genes and differential metabolites were analyzed using the IMPaLA<sup>36</sup> tool with the background of the KEGG database. The screening parameters were set as follows: pathway\_source, KEGG; num\_overlapping\_metabolites/genes  $> 0$ ; and  $p_{joint} < 0.05$ .

#### *Metabolic pathway enrichment analysis*

To further understand the overall affected pathways and corresponding differential metabolites, metabolic pathway enrichment analysis was performed to map a metabolic pathway network using FELLA, an R package to perform a network-based enrichment of a list of affected metabolites<sup>37</sup>. Enrichment was calculated based on the KEGG pathways using the diffusion method. According to the results of the pathway enrichment analysis, with the KEGG graph as a large background, a metabolite pathway background network was constructed, containing pathways, metabolites, reactions, enzymes, and modules. Using the diffusion

algorithm, heat was forced to flow from a given node (the metabolites in the significantly enriched pathway) and pass through the metabolite pathway background network, leading to the changes in the score of each node in the background network until the final rest. The temperatures (diffusion scores) were calculated as follows:

$$T = -K_I^{-1} \times G.$$

$G$  is the heat generation vector (1 is the differential metabolite and 0 otherwise);  $K_I$  is the conductance matrix, where  $K_I = L + B$ ,  $L$  is the unnormalized graph Laplacian,  $B$  is the diagonal adjacency matrix with  $B_{ii} = 1$  if node  $i$  is a pathway, and  $B_{ii} = 0$  otherwise. Matrix  $B$  ensures that the flow leaves the graph through the nodes of the pathways.

The p-value was obtained using a permutation test according to the diffusion scores. The network nodes were sorted according to the p-value, and nodes with p-value  $< 0.05$  were screened to construct the metabolic pathway network.

## Results

### *Identification of DEGs in epilepsy and IS*

We performed a differential expression analysis to identify DEGs in epilepsy and IS. A total of 549 upregulated and 739 downregulated DEGs between the epilepsy and control samples were identified based on the GSE143272 epilepsy dataset (Supplementary Table 1), and the volcano plot of these DEGs is shown in Figure 1A. In addition, 3593 upregulated and 1410

downregulated DEGs between the IS and control samples were screened based on the GSE58294 IS dataset (Supplementary Table 2). The volcano plot of these DEGs is shown in Figure 1B.

*Analysis of the co-expression module and identification of the disease-related genes*

To identify the disease-related genes, gene co-expression modules associated with epilepsy and IS were analyzed using the WGCNA of the DEGs. In this analysis, DEGs with similar expression patterns were assigned to the same module. Based on the GSE143272 epilepsy dataset, a power of 12 (scale-free  $R^2 = 0.85$ ) was chosen to ensure a scale-free network (Figure 2A), and three modules were obtained, with each color representing each module (Figure 2B). Subsequent correlation analysis revealed that two modules (MEblack and MEblue) were positively correlated with epilepsy ( $r > 0.8$ ;  $p < 0.001$ ) (Figure 2C) and were considered the epilepsy-related key modules containing 594 DEGs. Similarly, based on the GSE58294 IS dataset, a power of 10 (scale-free  $R^2 = 0.85$ ) was selected to ensure a scale-free network (Figure 2D), and seven modules were identified (Figure 2E). The MEblack module was positively correlated with IS ( $r = 0.88$ ,  $p = 2e-14$ ) (Figure 2F) and was considered an IS-related key module. A total of 2623 DEGs were included in this module.

*Analysis of the shared genes of epilepsy and IS*

Using Venn diagram analysis, 38 overlapping DEGs in the epilepsy- and IS-related key modules were obtained, which were

considered as the shared genes of two diseases.

#### *Functional enrichment analysis for the shared genes of epilepsy and IS*

Functional enrichment analyses were performed to elucidate the potential functions of the shared genes in the two diseases. The shared genes were remarkably enriched in 189 GO BP terms such as neutrophil activation, 20 GO CC terms such as ficolin-1-rich granule lumen, 17 GO MF terms such as hydrolase activity, and 9 KEGG pathways such as purine metabolism (Figure 3).

#### *PPI network analysis and hub gene identification*

Based on the information obtained from the STRING database, we constructed a PPI network including 26 shared genes of two diseases (Figure 4A). The top five genes were respectively identified based on four topological analysis algorithms, including MNC, MCC, degree, and EPC. Intersection analysis of the top five genes from each topological analysis algorithm identified three overlapping genes as candidate hub genes for the two diseases (Figure 4B), including interleukin 10 receptor subunit alpha (IL10RA), CD2 molecule (CD2), and complement C3a receptor 1 (C3AR1).

#### *Candidate hub-shared genes had high diagnostic efficacy for epilepsy and IS*

The diagnostic efficacies of three candidate hub-shared genes, including L10RA, CD2, and C3AR1 were evaluated based on the gene expression data from all the datasets. The results showed that

the AUC value of all these genes in training and validation datasets was  $> 0.66$  (Figure 4C). Therefore, all these genes had high diagnostic efficacies for epilepsy and IS.

*Validating the differential expression of the candidate hub-shared genes*

We first extracted the expression of candidate hub-shared genes from the epilepsy and IS samples based on the training datasets of epilepsy (GSE143272) and IS (GSE58294). Compared to the respective control samples, the expression of IL10RA, CD2, and C3AR1 was upregulated in epilepsy and IS samples (Figure 5A). Subsequently, we analyzed the differential expression of the candidate hub-shared genes based on the validation datasets of epilepsy (GSE16969 and GSE205661) and IS (GSE16561). The results showed that only C3AR1 expression was upregulated in the epilepsy and IS samples based on the validation datasets (Figure 5B), which was similar to the results obtained from the training datasets. Thus, C3AR1 was recognized as a hub-shared gene in the two diseases.

*Construction of the PPI network of C3AR1*

To explore the potential mechanisms of C3AR1 in the two diseases, a PPI network of C3AR1 was constructed using the GeneMANIA database (Figure 6A). In this network, 20 cooperators of C3AR1, including the complement C3 (C3), complement C4A (C4A), and complement C5a receptor 1 (C5AR1) were identified.

*Analysis of the upstream miRNAs of C3AR1*

Using the HMDD, we identified 32 miRNAs associated with epilepsy and 40 miRNAs associated with IS. Subsequently, 10 common miRNAs associated with two diseases were identified. Moreover, 125 upstream miRNAs of C3AR1 were predicted using the miRWalk database, and an upstream miRNA-C3AR1 network was constructed (Figure 6B). Notably, among the miRNAs upstream of C3AR1, only hsa-let-7b-5p was common in both diseases.

#### *Demographic data of clinical participants*

This study included 34 participants: 14 children with epilepsy (7 girls and 7 boys) and 20 healthy controls (13 girls and 7 boys). No significant age differences were observed between the epilepsy and healthy control groups (mean age:  $4.0 \pm 1.5$  years vs.  $3.4 \pm 0.8$  years,  $p > 0.05$ ), minimizing the impact of age as a confounding factor.

#### *Analysis of differential metabolites and pathway enrichment analysis*

Using nontargeted metabolomics analysis, 51 upregulated and 88 downregulated differential metabolites were identified between the plasma samples from children with epilepsy and healthy controls (Figure 7A). Pathway enrichment analysis was performed to clarify the biological significance of C3AR1 and its metabolites. A total of 22 significantly enriched pathways containing 20 metabolites were identified (Figure 7B).

#### *Metabolic pathway network analysis*

We identified 22 significantly enriched pathways from the KEGG database. A metabolite pathway background network was constructed including 11391 nodes, 33332 edges, and 17 metabolites. The diffusion scores of the nodes were calculated using a diffusion algorithm. Subsequently, a metabolic pathway network was constructed using nodes with  $p < 0.05$ . The metabolic pathway network included 81 nodes, 85 edges, and 12 metabolites (Figure 7C). Based on the diffusion algorithm, heat was believed to flow from the 12 differential metabolites and pass through the metabolite pathway background network. At rest, six pathways were identified: hsa04721, synaptic vesicle cycle; hsa04725, cholinergic synapses; hsa04742, taste transduction; hsa04913, ovarian steroidogenesis; hsa04925, aldosterone synthesis and secretion; and hsa05033, nicotine addiction. C3AR1 was involved in three pathways (hsa04721, hsa04742, and hsa05033) through interactions with the differential metabolite C01996 (acetylcholine). Acetylcholine was downregulated in the plasma samples from children with epilepsy.

## Discussion

IS has been identified as an important cause of epilepsy<sup>38</sup>. PSE is a common complication of IS that adversely affects the patients' prognosis. Currently, the risk of PSE cannot be predicted adequately using clinical and radiological parameters<sup>39</sup>. Therefore, elucidating the common mechanisms underlying IS and epilepsy is necessary to uncover the underlying mechanisms and potential

biomarkers of PSE. In this study, we investigated the shared genes of IS and epilepsy and explored their related mechanisms. The findings will improve our understanding of the common mechanisms underlying IS and epilepsy.

Using microarray technology, the expression data of thousands of genes for multiple diseases can be rapidly obtained, making it possible to elucidate the common pathogenesis of multiple diseases at the gene level <sup>40,41</sup>. WGCNA is a useful tool for discovering gene-gene and gene-disease relationships and has been widely applied to construct gene co-expression networks and identify key genes in disease-related networks <sup>42-44</sup>. In this study, we performed WGCNA and identified two epilepsy-related key modules containing 594 DEGs and an IS-related key module containing 2623 DEGs. PPI network analysis identified three candidate hub-shared genes from the 38 overlapping DEGs in the epilepsy- and IS-related key modules, including IL10RA, CD2, and C3AR1. These data suggest that these shared genes may be involved in the development of IS and epilepsy.

IL-10 is a multifunctional anti-inflammatory cytokine that plays a neuroprotective role in the brain <sup>45,46</sup>. It can inhibit IL-1 $\beta$  production and suppress inflammasome activation of microglia in epileptic seizures <sup>47</sup>. IL-1 $\beta$  can also predict seizure recurrence after the first epileptic seizure among IS patients and may serve as a promising prognostic biomarker for PSE <sup>48,49</sup>. Moreover, IL-10 is positively associated with the risk of stroke <sup>50,51</sup>. IL-10 activates

downstream signaling by binding to the IL-10 receptors (IL-10RA and IL10RB)<sup>52</sup>. Polymorphisms in IL10, IL10RA, and IL10RB genes are associated with IS in terms of hypertension<sup>53</sup>. CD2 belongs to the immunoglobulin superfamily and plays a key role in mediating T and natural killer (NK) cell activation<sup>54</sup>. Cluster of differentiation 8 (CD8+), cluster of differentiation 4 (CD4+), and NK T cells are recruited within 24 h of IS and play a crucial role in regulating the inflammatory response after injury<sup>55</sup>. T-cell numbers are associated with neuronal loss in medial temporal lobe epilepsy<sup>56</sup>. These findings suggest that CD2 contributes to epilepsy and IS by affecting T and NK cell activation. C3AR1 is a key regulator of neuronal tau pathogenesis and has been implicated in the immune network of the CNS<sup>57</sup>. Tau protein hyperphosphorylation has been observed in animal models of epilepsy<sup>58</sup> and in patients with epilepsy<sup>59</sup>. Tau protein also contributes to brain damage following stroke as seen in an animal model of stroke<sup>60</sup> and has been suggested as a potential therapeutic target for IS<sup>61</sup>. Moreover, Tau protein is reported to have a sensitivity of 100% and a specificity of 73% for predicting PSE<sup>39</sup>. These data suggest that C3AR1 contributes to epilepsy and IS by regulating the pathogenesis of the neuronal tau protein. In our study, ROC analysis revealed that the diagnostic efficacy of L10RA, CD2, and C3AR1 was high, with AUC values > 0.66 in the training and validation datasets. Therefore, these genes may serve as promising biomarkers for the diagnosis of IS and epilepsy. However, differential expression validation showed

that only the expression trend of C3AR1 was consistent in both the training and validation datasets. The differential expression and diagnostic value of L10RA, CD2, and C3AR1 need to be validated using additional datasets or clinical cohorts.

miRNAs are small noncoding RNA molecules of 22 nucleotides in length. Accumulating evidence has confirmed that miRNAs are key regulators of IS<sup>62</sup>, epilepsy<sup>63</sup>, and PSE<sup>64</sup>. In this study, we found that hsa-let-7b-5p is a common miRNA in epilepsy and IS and can target C3AR1. A previous study has shown that knockdown of let-7b-5p reverses the effect of repetitive transcranial magnetic stimulation on microglia phenotype, and play a key role in neurological recovery and preventing ischemic stroke<sup>65</sup>. Combined analyses of Fas and hsa-let-7b-5p expression have revealed them to be promising biomarkers for predicting poor neurological outcomes in patients with IS patients<sup>66</sup>. Although the role of hsa-let-7b-5p in epilepsy has not been reported, our results imply that hsa-let-7b-5p-C3AR1 axis may be a potential common regulatory mechanism mediating IS and epilepsy.

Several metabolites like lactate, glutamate, and citrate as well as metabolic pathways such as glycine, serine, and threonine metabolism were recently shown to be involved in epilepsy, providing a novel perspective for exploring promising biomarkers and therapeutic targets for epilepsy<sup>67</sup>. Thus, we conducted a nontargeted metabolomics analysis to identify differential metabolites in the plasma samples of children with epilepsy and

subsequently analyzed the key metabolites that may be regulated by C3AR1. We found that acetylcholine was downregulated in the plasma samples of children with epilepsy and that C3AR1 was involved in three pathways (hsa04721, synaptic vesicle cycle; hsa04742, taste transduction; and hsa05033, nicotine addiction) through interaction with acetylcholine. Acetylcholine is the main stimulant of the autonomic nervous system and plays a key role in signal transmission via the cholinergic and nicotinic receptors. Increasing evidence has highlighted the role of acetylcholine and cholinergic neurotransmission in the pathogenesis of epilepsy<sup>68-70</sup>. Moreover, the synaptic vesicle cycle controls neurotransmitter release and is implicated in epilepsy progression<sup>71</sup>. Nicotine addiction is also an important pathway that mediates the role of drugs in treating epilepsy<sup>72</sup>, suggesting a potential role of nicotine addiction in epilepsy. Overall, our data suggest that C3AR1 may contribute to PSE development by regulating acetylcholine and related pathways such as the synaptic vesicle cycle and nicotine addiction.

Several limitations should be considered when interpreting the findings of this study. Firstly, the hub-shared genes such as C3AR1 and its related miRNAs like hsa-let-7b-5p were obtained from bioinformatics analyses of publicly available datasets, and their expression was not validated in clinical samples. Validating their expression and investigating how C3AR1 interacts with other relevant molecules and its specific mechanisms in disease

pathology is crucial for understanding its potential as a diagnostic biomarker. Secondly, the functions of hub-shared genes, especially C3AR1 were not explored in depth. Further functional studies, such as gene knockout or overexpression experiments, are required to investigate the role of in epilepsy and ischemic stroke. Thirdly, fourteen plasma samples from children with epilepsy were used in this study. The sample size was small, which may reduce the statistical power. Lastly, this study only validated changes in metabolites in plasma samples from epilepsy patients, without assessing metabolite alterations in IS. Future studies should be conducted to investigate metabolic changes in IS patients as well to help clarify the similarities and differences between epilepsy and IS at the metabolic level, further supporting the potential of shared genes as biomarkers for both conditions. Furthermore, given the complexity of these diseases, future research may focus on identifying a panel of biomarkers rather than relying on a single gene.

## Conclusions

Our findings reveal that the key shared genes, especially C3AR1, may be implicated in the development of IS and epilepsy and could serve as potential diagnostic biomarkers for both diseases. Furthermore, C3AR1 may contribute to PSE development by interacting with hsa-let-7b-5p and acetylcholine. These findings may aid in the precise diagnosis and treatment of IS and epilepsy.

## List of abbreviations

IS: ischemic stroke; IL10RA: Interleukin 10 Receptor Subunit Alpha; CD2: cluster of differentiation 2; C3AR1: complement component 3a receptor 1; GEO: Gene Expression Omnibus; WGCNA: weighted gene co-expression network analysis; PPI: Protein-protein interaction; ROC: receiver-operating characteristic; DEG: differentially expressed genes; PSE: post-stroke epilepsy; DAMP: damage-associated molecular patterns; CNS: central nervous system; GO: Gene Ontology; KEGG: Kyoto Encyclopedia of Genes and Genomes; MF: molecular function; CC: cellular component; BP: biological processes; MNC: maximal neighborhood component; MCC: maximal clique centrality; EPC: edge-percolated component; AUC: Area under the curve; HMDD: Human microRNA Disease Database; NK: natural killer; CD8: cluster of differentiation 8; CD4: cluster of differentiation 4; CD2: cluster of differentiation 2

## References

- 1 Popiela, T. J. *et al.* The assessment of endovascular therapies in ischemic stroke: management, problems and future approaches. *Journal of Clinical Medicine* **11**, 1864 (2022).
- 2 Murtagh, B. & Smalling, R. W. Cardioembolic stroke. *Current atherosclerosis reports* **8**, 310-316 (2006).
- 3 Alet, M. *et al.* Predictive factors for the development of epilepsy after ischemic stroke. *J Stroke Cerebrovasc Dis* **31**, 4 (2022).
- 4 Hassani, M., Cooray, G., Sveinsson, O. & Cooray, C. Post-stroke epilepsy in an ischemic stroke cohort-Incidence and diagnosis. *Acta Neurol Scand* **141**, 141-147 (2020).
- 5 Kanner, A. M. & Bicchi, M. M. Antiseizure Medications for Adults With Epilepsy: A Review. *Jama* **327**, 1269-1281 (2022).
- 6 Chen, P., Chen, F. & Zhou, B. Understanding the Role of Glia-Neuron Communication in the Pathophysiology of Epilepsy: A Review. *Journal of Integrative Neuroscience* **21**, 102 (2022).
- 7 Hasan, T. F. *et al.* Diagnosis and Management of Acute Ischemic Stroke. *Mayo Clin Proc* **93**, 523-538 (2018).
- 8 Krueger, H. *et al.* Prevalence of individuals Experiencing the Effects of Stroke in Canada: Trends and Projections. *Stroke* **46**, 2226-2231 (2015).
- 9 Zhang, C. *et al.* Risk factors for post-stroke seizures: a systematic review and meta-analysis. *Epilepsy Res* **108**, 1806-1816 (2014).
- 10 Tröscher, A. R. *et al.* Inflammation mediated epileptogenesis as possible mechanism underlying ischemic post-stroke epilepsy. *Frontiers in aging neuroscience* **13**, 781174 (2021).
- 11 Zhao, H., Li, Y., Zhang, Y., He, W.-Y. & Jin, W.-N. Role of immune and inflammatory mechanisms in stroke: a review of current advances. *Neuroimmunomodulation* **29**, 255-268 (2022).
- 12 Henshall, D. C. & Engel, T. Contribution of apoptosis-associated signaling pathways to epileptogenesis: lessons from Bcl-2 family knockouts. *Front Cell Neurosci* **7** (2013).
- 13 Klein, P. *et al.* Commonalities in epileptogenic processes from different acute brain insults: Do they translate? *Epilepsia* **59**, 37-66 (2018).
- 14 Vezzani, A., Balosso, S. & Ravizza, T. Neuroinflammatory pathways as treatment targets and biomarkers in epilepsy. *Nat Rev Neurosci* **15**, 459-472 (2019).
- 15 Doria, J. W. & Forgacs, P. B. Incidence, Implications, and Management of Seizures Following Ischemic and Hemorrhagic Stroke. *Curr Neurol Neurosci Rep* **19**, 019-0957 (2019).
- 16 Abraira, L. *et al.* Correlation of blood biomarkers with early-onset seizures

after an acute stroke event. *Epilepsy Behav* **104**, 31 (2020).

17 Fu, L. *et al.* Negative regulation of angiogenesis and the MAPK pathway may be a shared biological pathway between IS and epilepsy. *PLoS One* **18**, e0286426, doi:10.1371/journal.pone.0286426 (2023).

18 Barrett, T. *et al.* NCBI GEO: archive for functional genomics data sets--update. *Nucleic Acids Res* **41**, 27 (2013).

19 Smyth, G. K. in *Bioinformatics and computational biology solutions using R and Bioconductor* 397-420 (Springer, NewYork, 2005).

20 Huang, C. *et al.* Combined transcriptomics and proteomics forecast analysis for potential biomarker in the acute phase of temporal lobe epilepsy. *Frontiers in neuroscience* **17**, 1145805, doi:10.3389/fnins.2023.1145805 (2023).

21 Langfelder, P. & Horvath, S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* **9**, 1471-2105 (2008).

22 Kanehisa, M. & Goto, S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* **28**, 27-30, doi:10.1093/nar/28.1.27 (2000).

23 Kanehisa, M. Toward understanding the origin and evolution of cellular organisms. *Protein Sci* **28**, 1947-1951, doi:10.1002/pro.3715 (2019).

24 Kanehisa, M., Furumichi, M., Sato, Y., Kawashima, M. & Ishiguro-Watanabe, M. KEGG for taxonomy-based analysis of pathways and genomes. *Nucleic Acids Res* **51**, D587-D592, doi:10.1093/nar/gkac963 (2023).

25 Yu, G., Wang, L. G., Han, Y. & He, Q. Y. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* **16**, 284-287, doi:10.1089/omi.2011.0118 (2012).

26 Szklarczyk, D. *et al.* The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. *Nucleic Acids Res* **45**, D362-D368, doi:10.1093/nar/gkw937 (2017).

27 Shannon, P. *et al.* Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* **13**, 2498-2504, doi:10.1101/gr.1239303 (2003).

28 Chin, C. H. *et al.* cytoHubba: identifying hub objects and sub-networks from complex interactome. *BMC Syst Biol* **4**, 1752-0509 (2014).

29 Yuan, Y., Zhu, H., Huang, S., Zhang, Y. & Shen, Y. Establishment of a diagnostic model based on immune-related genes in children with asthma. *Helion* **10**, e25735, doi:10.1016/j.heliyon.2024.e25735 (2024).

30 Lin, G. *et al.* Identification of key genes as potential diagnostic biomarkers in sepsis by bioinformatics analysis. *PeerJ* **12**, e17542, doi:10.7717/peerj.17542 (2024).

31 Robin, X. *et al.* pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics* **12**, 1471-2105 (2011).

32 Franz, M. *et al.* GeneMANIA update 2018. *Nucleic Acids Res* **46**, W60-W64 (2018).

33 Shen, Y. *et al.* MicroRNA-Disease Network Analysis Repurposes Methotrexate for the Treatment of Abdominal Aortic Aneurysm in Mice.

34 *Genomics Proteomics Bioinformatics* **24**, 002 (2022).

35 Bao, L. *et al.* Analysis of serum amino acids and tryptophan metabolites to predict hepatic encephalopathy in portal hypertension patients receiving a transjugular intrahepatic portal shunt (TIPS). *Frontiers in pharmacology* **16**, 1546665, doi:10.3389/fphar.2025.1546665 (2025).

36 Gao, X. *et al.* Plasma lipidomic fingerprinting enables high-accuracy biomarker discovery for Alzheimer's disease: a targeted LC-MRM/MS approach. *GeroScience*, doi:10.1007/s11357-025-01777-5 (2025).

37 Cavill, R. *et al.* Consensus-phenotype integration of transcriptomic and metabolomic data implies a role for metabolism in the chemosensitivity of tumour cells. *PLoS Comput Biol* **7**, 31 (2011).

38 Picart-Armada, S., Fernández-Albert, F., Vinaixa, M., Yanes, O. & Perera-Lluna, A. FELLA: an R package to enrich metabolomics data. *BMC bioinformatics* **19**, 1-9 (2018).

39 Pitkänen, A., Roivainen, R. & Lukasiuk, K. Development of epilepsy after ischaemic stroke. *The Lancet Neurology* **15**, 185-197 (2016).

40 Dev, P., Cyriac, M., Chakravarty, K. & Pathak, A. Blood and CSF biomarkers for post-stroke epilepsy: a systematic review. *Acta Epileptologica* **4**, 21 (2022).

41 Gong, F. C. *et al.* Identification of Potential Biomarkers and Immune Features of Sepsis Using Bioinformatics Analysis. *Mediators Inflamm* **9** (2020).

42 Zhou, J. *et al.* Weighted correlation network bioinformatics uncovers a key molecular biosignature driving the left-sided heart failure. *BMC Medical Genomics* **13**, 1-13 (2020).

43 Wang, M. *et al.* LncRNAs related key pathways and genes in ischemic stroke by weighted gene co-expression network analysis (WGCNA). *Genomics* **112**, 2302-2308 (2020).

44 Raghow, R. An 'omics' perspective on cardiomyopathies and heart failure. *Trends in molecular medicine* **22**, 813-827 (2016).

45 Daneshfrooz, N., Bagherzadeh Cham, M., Majidi, M. & Panahi, B. Identification of potentially functional modules and diagnostic genes related to amyotrophic lateral sclerosis based on the WGCNA and LASSO algorithms. *Scientific Reports* **12**, 20144 (2022).

46 Liesz, A. *et al.* Regulatory T cells are key cerebroprotective immunomodulators in acute experimental stroke. *Nature medicine* **15**, 192-199 (2009).

47 Ren, X. *et al.* Regulatory B cells limit CNS inflammation and neurologic deficits in murine experimental stroke. *Journal of Neuroscience* **31**, 8556-8563 (2011).

48 Sun, Y. *et al.* Interleukin-10 inhibits interleukin-1 $\beta$  production and inflammasome activation of microglia in epileptic seizures. *Journal of neuroinflammation* **16**, 1-13 (2019).

49 Zhang, Q. *et al.* Association between IL-1 $\beta$  and recurrence after the first

epileptic seizure in ischemic stroke patients. *Scientific reports* **10**, 13505 (2020).

49 Shen, L., Yang, J. & Tang, Y. Predictive values of the SeLECT score and IL-1 $\beta$  for post-stroke epilepsy. *Neuropsychiatric Disease and Treatment*, 2465-2472 (2021).

50 Li, S. *et al.* MicroRNA-4443 regulates monocyte activation by targeting tumor necrosis factor receptor associated factor 4 in stroke-induced immunosuppression. *European Journal of Neurology* **27**, 1625-1637 (2020).

51 Wang, H. *et al.*  $\beta$ -arrestin2 functions as a key regulator in the sympathetic-triggered immunodepression after stroke. *Journal of neuroinflammation* **15**, 1-11 (2018).

52 Donnelly, R. P., Sheikh, F., Kotenko, S. V. & Dickensheets, H. The expanded family of class II cytokines that share the IL-10 receptor-2 (IL-10R2) chain. *Journal of leukocyte biology* **76**, 314-321 (2004).

53 Park, H.-K., Kim, D. H., Yun, D. H. & Ban, J. Y. Association between IL10, IL10RA, and IL10RB SNPs and ischemic stroke with hypertension in Korean population. *Molecular biology reports* **40**, 1785-1790 (2013).

54 Crawford, K. *et al.* CD2 engagement induces dendritic cell activation: implications for immune surveillance and T-cell activation. *Blood* **102**, 1745-1752 (2003).

55 Selvaraj, U. M. & Stowe, A. M. Long-term T cell responses in the brain after an ischemic stroke. *Discovery medicine* **24**, 323 (2017).

56 Tröscher, A. R. *et al.* T cell numbers correlate with neuronal loss rather than with seizure activity in medial temporal lobe epilepsy. *Epilepsia* **62**, 1343-1353 (2021).

57 Litvinchuk, A. *et al.* Complement C3aR inactivation attenuates tau pathology and reverses an immune network deregulated in tauopathy models and Alzheimer's disease. *Neuron* **100**, 1337-1353. e1335 (2018).

58 Tian, F.-F. *et al.* Potential roles of Cdk5/p35 and tau protein in hippocampal mossy fiber sprouting in the PTZ kindling model. *Clinical laboratory* **56**, 127-136 (2010).

59 Tai, X. Y. *et al.* Hyperphosphorylated tau in patients with refractory epilepsy correlates with cognitive decline: a study of temporal lobe resections. *Brain* **139**, 2441-2455 (2016).

60 Bi, M. *et al.* Tau exacerbates excitotoxic brain damage in an animal model of stroke. *Nature communications* **8**, 473 (2017).

61 Zheng, G. Q., Wang, X. M., Wang, Y. & Wang, X. T. Tau as a potential novel therapeutic target in ischemic stroke. *Journal of cellular biochemistry* **109**, 26-29 (2010).

62 Vasudeva, K. & Munshi, A. miRNA dysregulation in ischaemic stroke: Focus on diagnosis, prognosis, therapeutic and protective biomarkers. *European Journal of Neuroscience* **52**, 3610-3627 (2020).

63 Ghafouri-Fard, S., Hussen, B. M., Abak, A., Taheri, M. & Jalili Khoshnoud, R. Aberrant expression of miRNAs in epilepsy. *Molecular biology reports* **49**,

5057-5074 (2022).

64 Zöllner, J. P. *et al.* Seizures and epilepsy in patients with ischaemic stroke. *Neurological Research and Practice* **3**, 1-17 (2021).

65 Hong, Y. *et al.* High-frequency repetitive transcranial magnetic stimulation (rTMS) protects against ischemic stroke by inhibiting M1 microglia polarization through let-7b-5p/HMGA2/NF- $\kappa$ B signaling pathway. *BMC neuroscience* **23**, 49 (2022).

66 Chi, N. F. *et al.* Hyperglycemia-related FAS gene and hsa-let-7b-5p as markers of poor outcomes for ischaemic stroke. *European Journal of Neurology* **27**, 1647-1655 (2020).

67 Lai, W., Du, D. & Chen, L. Metabolomics Provides Novel Insights into Epilepsy Diagnosis and Treatment: A Review. *Neurochemical Research* **47**, 844-859 (2022).

68 Ghasemi, M. & Hadipour-Niktarash, A. Pathologic role of neuronal nicotinic acetylcholine receptors in epileptic disorders: implication for pharmacological interventions. *Reviews in the Neurosciences* **26**, 199-223 (2015).

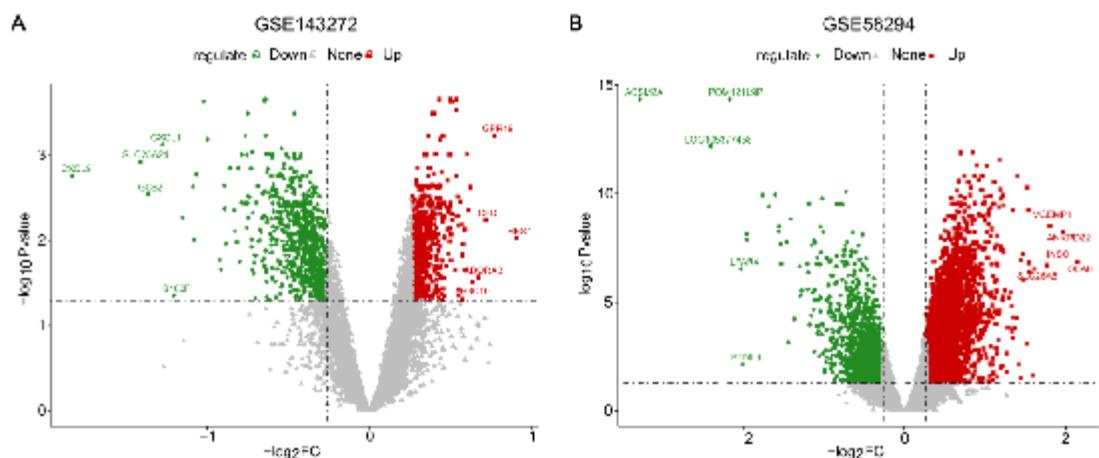
69 Biagioni, F. *et al.* Degeneration of cholinergic basal forebrain nuclei after focally evoked status epilepticus. *Neurobiology of disease* **121**, 76-94 (2019).

70 Meller, S., Brandt, C., Theilmann, W., Klein, J. & Löscher, W. Commonalities and differences in extracellular levels of hippocampal acetylcholine and amino acid neurotransmitters during status epilepticus and subsequent epileptogenesis in two rat models of temporal lobe epilepsy. *Brain research* **1712**, 109-123 (2019).

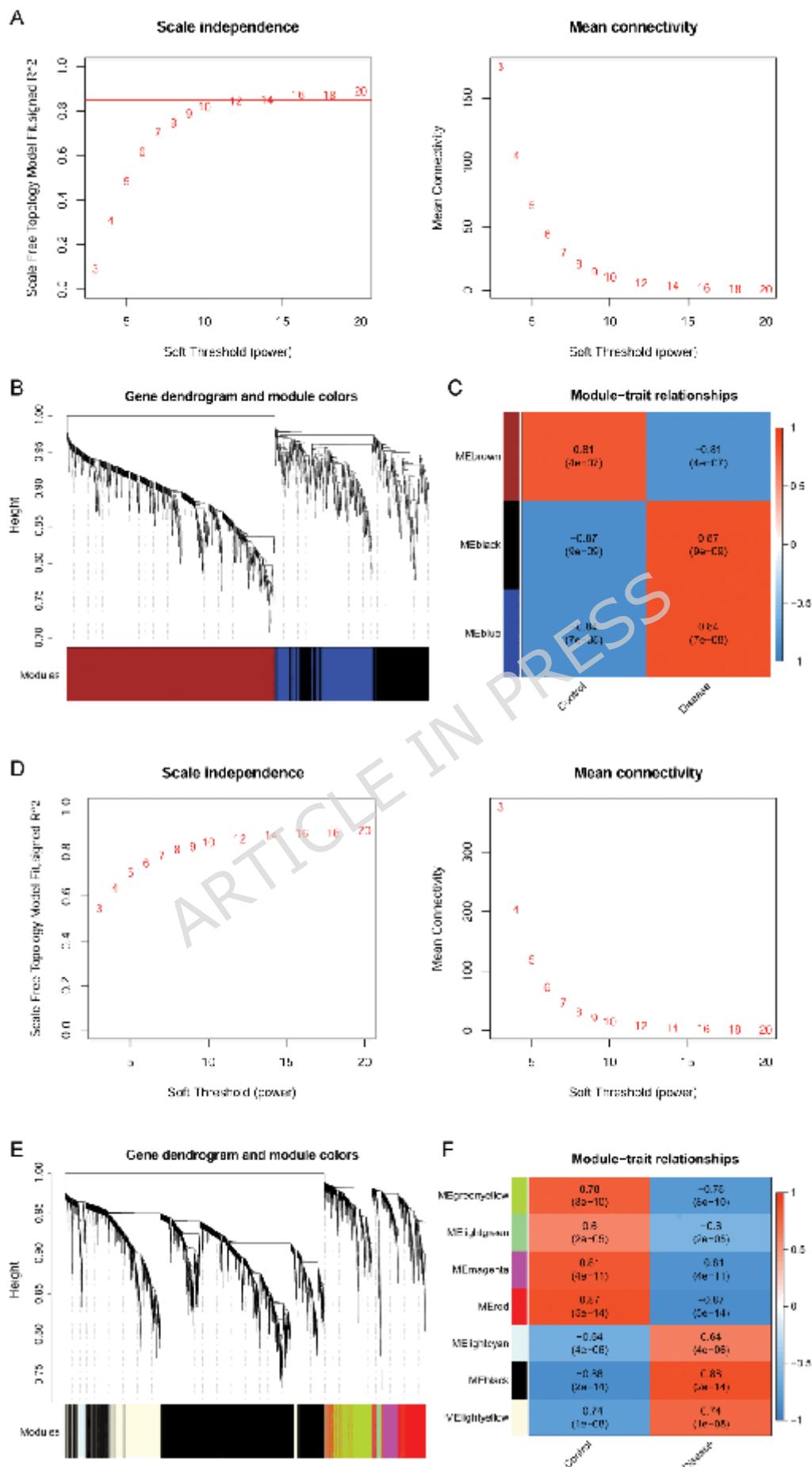
71 Xi, X.-J. *et al.* Dlg4 and Vamp2 are involved in comorbid epilepsy and attention-deficit hyperactivity disorder: A microarray data study. *Epilepsy & Behavior* **110**, 107192 (2020).

72 Huang, D. *et al.* Investigating the molecular mechanism of Compound Danshen Dropping Pills for the treatment of epilepsy by utilizing network pharmacology and molecular docking technology. *Annals of Translational Medicine* **10** (2022).

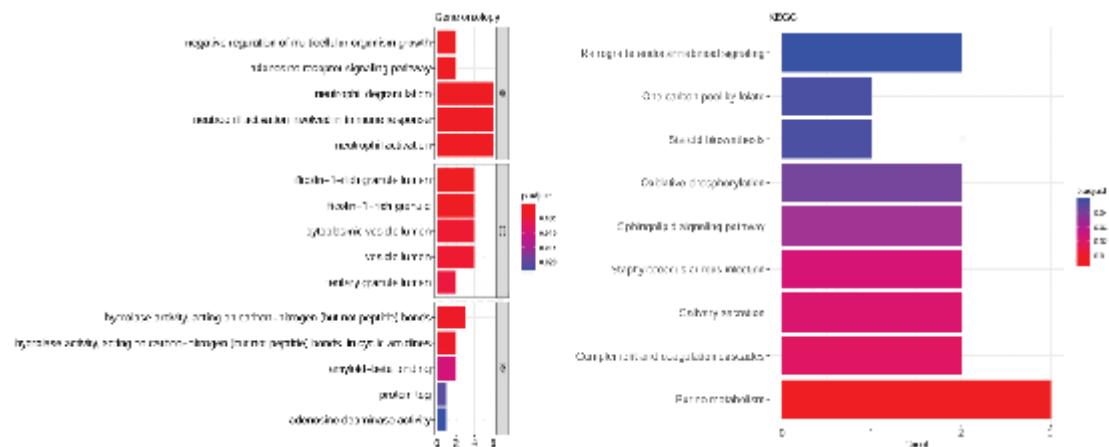
## Figure



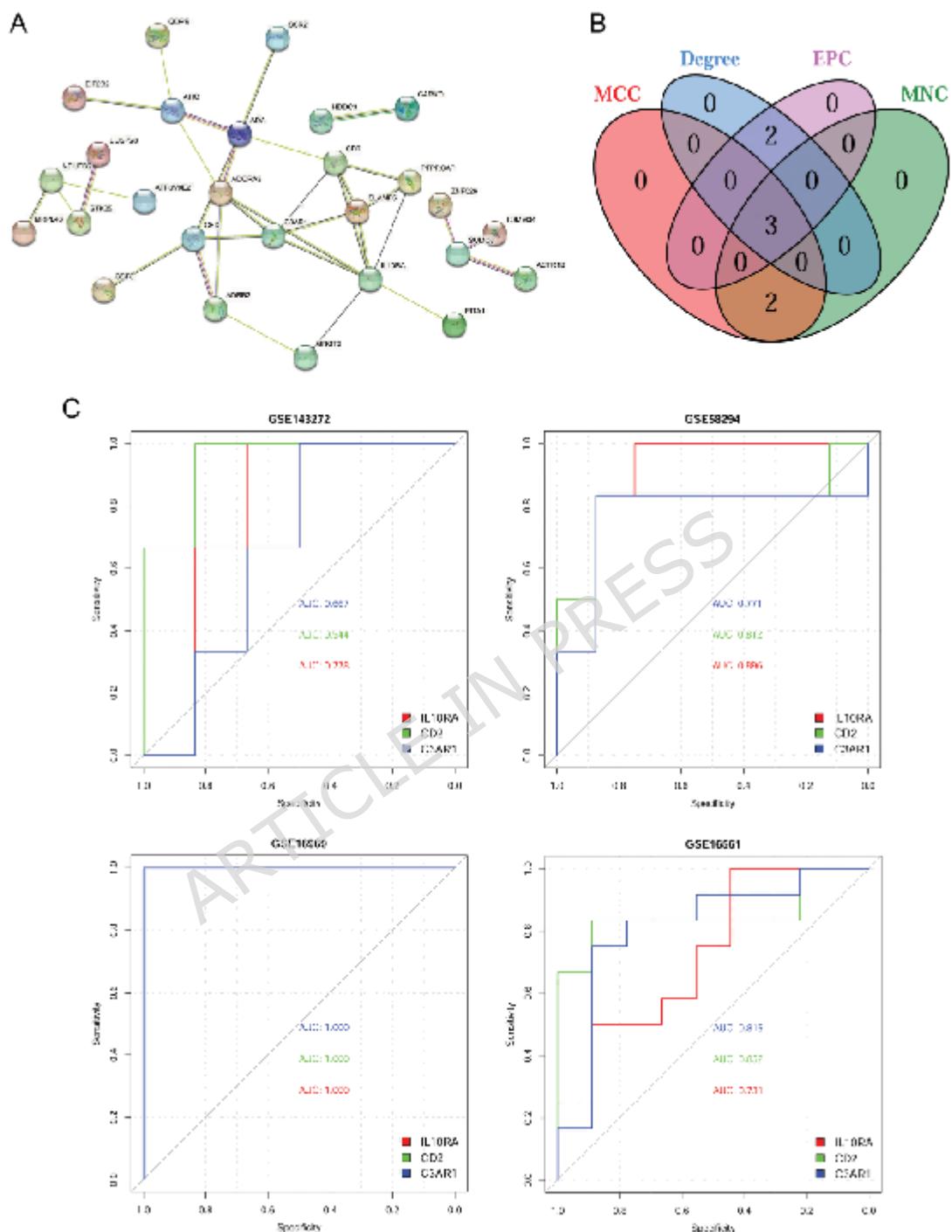
**Figure 1** Volcano plots of DEGs associated with epilepsy and IS. A: Volcano plot of DEGs between epilepsy and control samples based on the GSE143272 dataset. B: Volcano plot of DEGs between IS and control samples based on the GSE58294 dataset. The horizontal axis is  $-\log_{10}(p\text{-value})$ , and the vertical axis indicates  $\log_2(\text{FC})$ . Each node represents a gene. The red nodes indicate upregulated genes, the green nodes indicate downregulated genes and the gray nodes indicate genes without significant differences. IS: ischemic stroke; DEGs: differentially expressed genes; FC: fold change



**Figure 2** Analyzing epilepsy- and IS-related gene coexpression modules using WGCNA based on GSE143272 and GSE58294 datasets, respectively. A: Analysis of scale independence and mean connectivity to determine soft-thresholding power in WGCNA based on the GSE143272 dataset. B: Clustering dendrograms of co-expression genes in epilepsy. Genes were divided into various modules through hierarchical clustering, and different colors represent different modules. The gene that cannot be classified into any module by default is indicated in grey. C: The heat map of epilepsy-related module-trait relationships. D: Analysis of scale independence and mean connectivity based on the GSE58294 dataset. E: Clustering dendrograms of co-expression genes in IS. F: The heat map of IS-related module-trait relationships. IS: ischemic stroke; WGCNA: weighted gene co-expression network analysis.



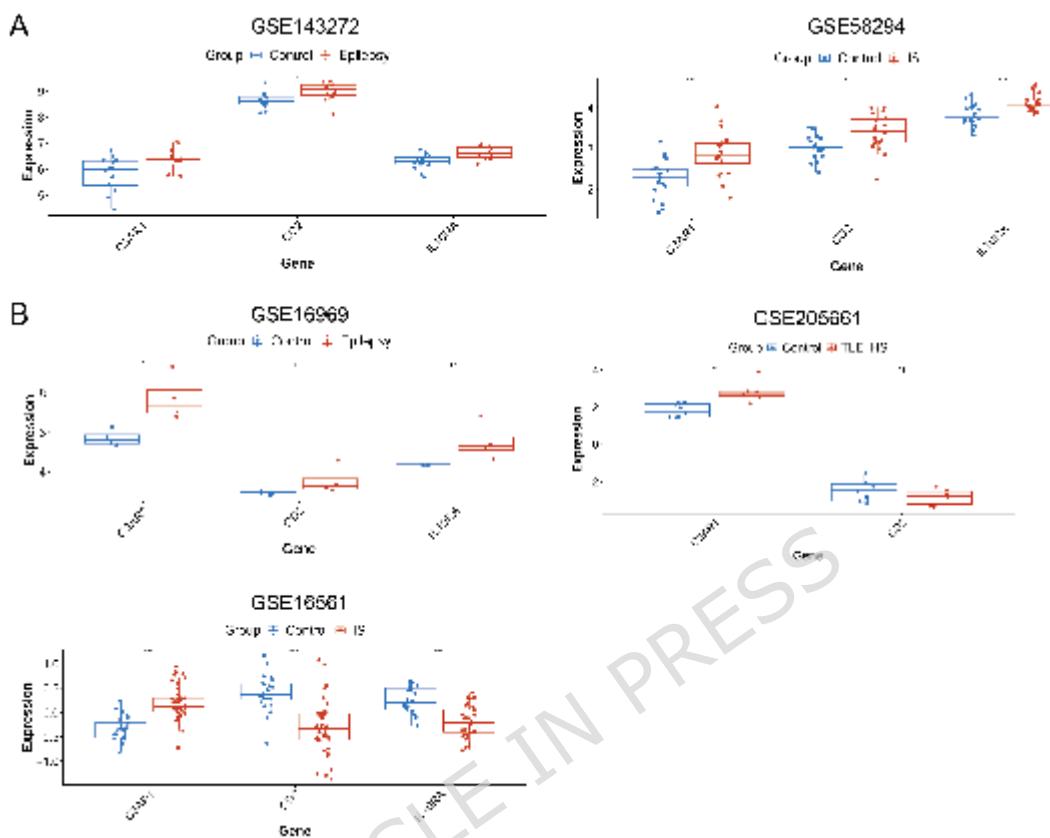
**Figure 3** GO and KEGG pathway enrichment analysis for shared genes of epilepsy and IS. IS: ischemic stroke; GO: Gene Ontology; KEGG: Kyoto Encyclopedia of Genes and Genomes.



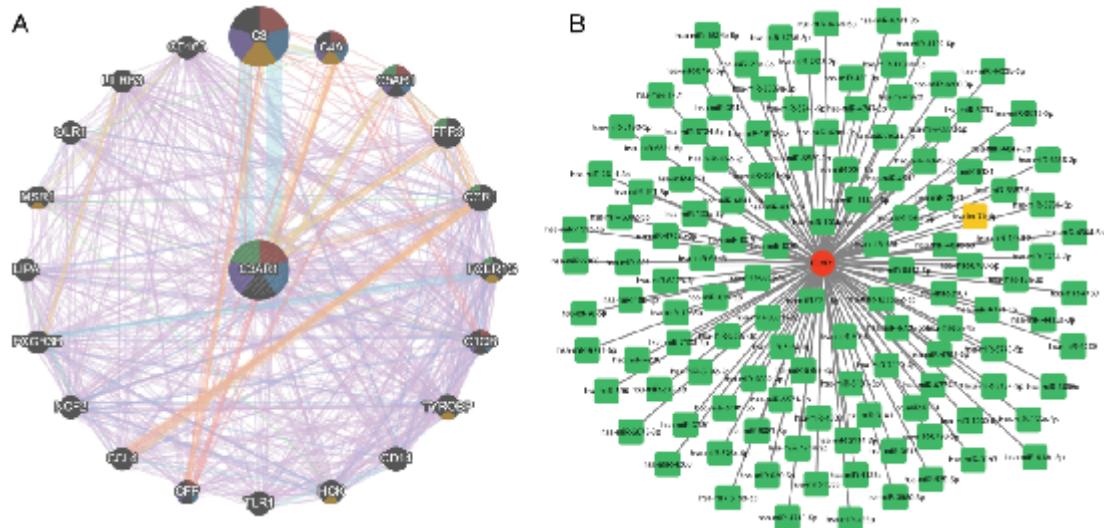
**Figure 4** PPI network analysis and evaluating diagnostic efficacy of candidate hub-shared genes for epilepsy and IS. A: PPI network constructed by the shared genes of two diseases. B: Venn diagram shows the number of overlapping top genes in the PPI network

based on four topological properties (MCC, MNC, degree, and EPC).

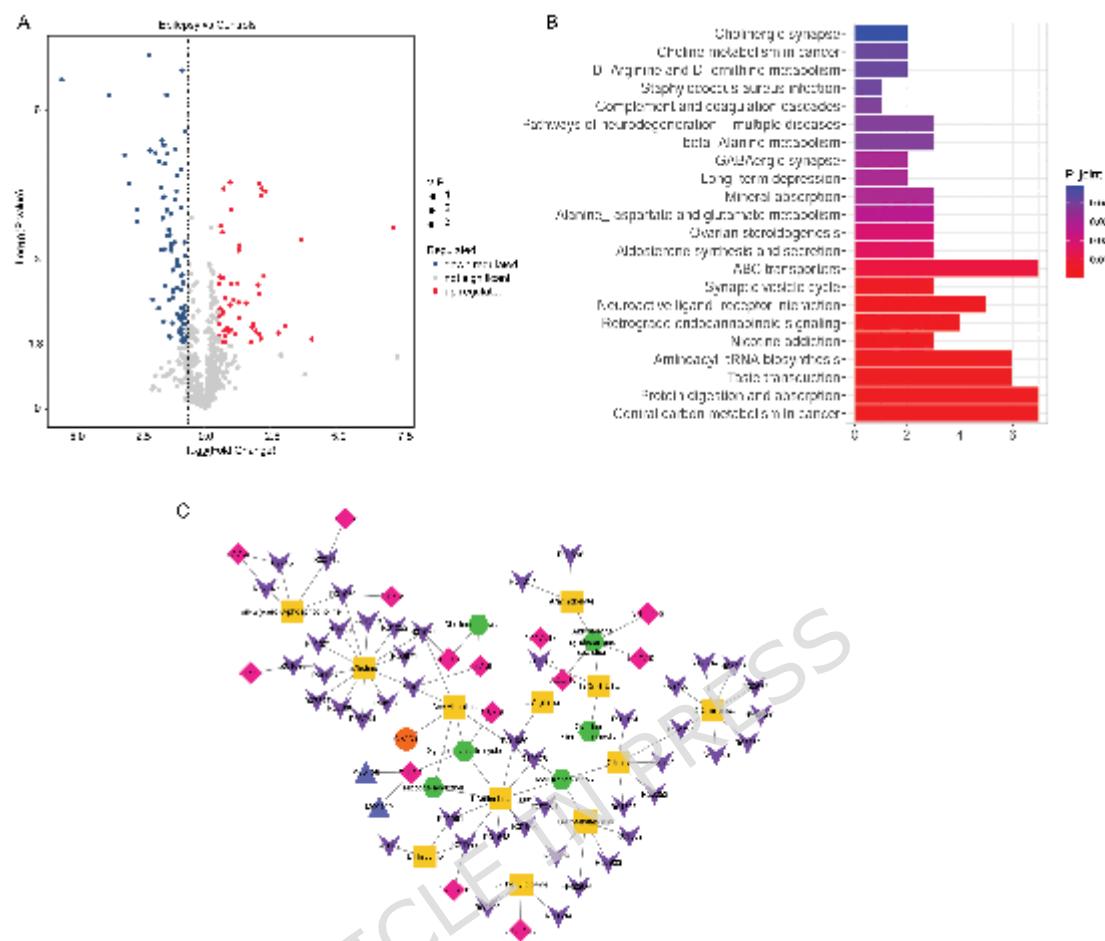
C: Evaluation of diagnostic efficacy of candidate hub-shared genes by ROC based on four datasets. IS: ischemic stroke; PPI: protein-protein interaction; MNC: maximal neighborhood component; MCC: maximal clique centrality; EPC: edge percolated component (EPC); ROC: receiver-operating characteristic; AUC: area under the curve.



**Figure 5** Validating the differential expression of the shared genes of epilepsy and IS. A: Analysis of the expression of hub-shared genes using training datasets (the GSE143272 epilepsy dataset and GSE58294 IS dataset). B: Analysis of the expression of hub-shared genes using validation datasets (the GSE16969 and GSE205661 epilepsy datasets and GSE16561 IS dataset). \*  $p < 0.05$ , \*\*  $p < 0.01$ , and \*\*\*  $p < 0.001$ . IS: ischemic stroke; TLE+HS: temporal lobe epilepsy with hippocampal sclerosis.



**Figure 6** The PPI network of C3AR1 and upstream miRNA-C3AR1 network. A: The PPI network of C3AR1 and its 20 cooperators was constructed using GeneMANIA. B: The upstream miRNA-C3AR1 network. In this network, the red circle node is the hub-shared gene, the green square node is the miRNA, and the yellow square node is the common miRNA of the two diseases.



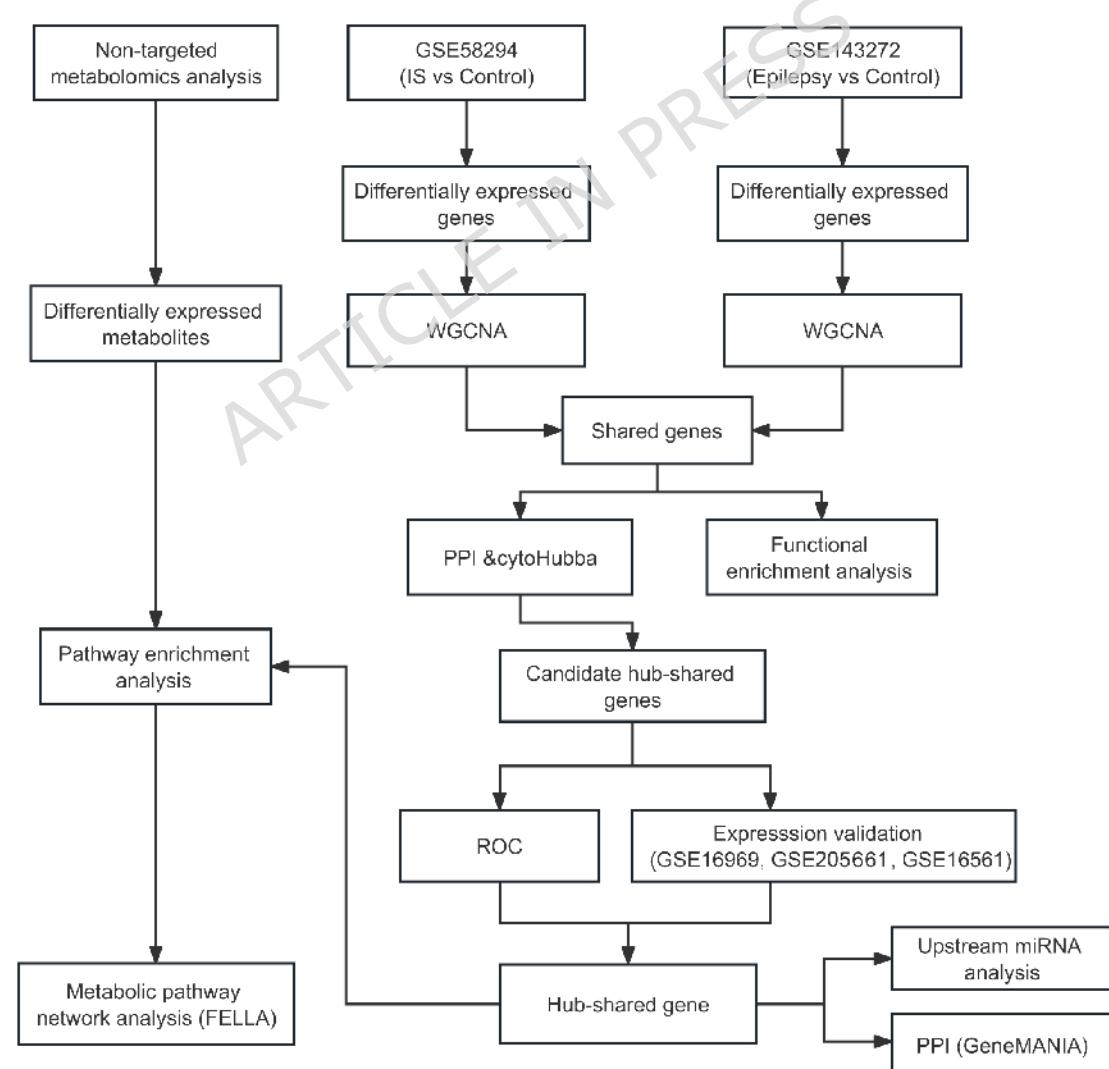
**Figure 7** Identifying the differential metabolites, pathway enrichment analysis, and metabolic pathway network analysis. A: Volcano plot of differential metabolites between the plasma samples from children with epilepsy and healthy controls. B: Pathway enrichment analysis for differential metabolites. C: Metabolic pathway network of C3AR1. In this network, the green hexagon node is the pathway, the blue triangle node is the module, the pink diamond node is the enzyme, the purple inverted triangle node is the reaction, the yellow square node is the differential metabolite, and the red circle node is the hub-shared gene.

Table

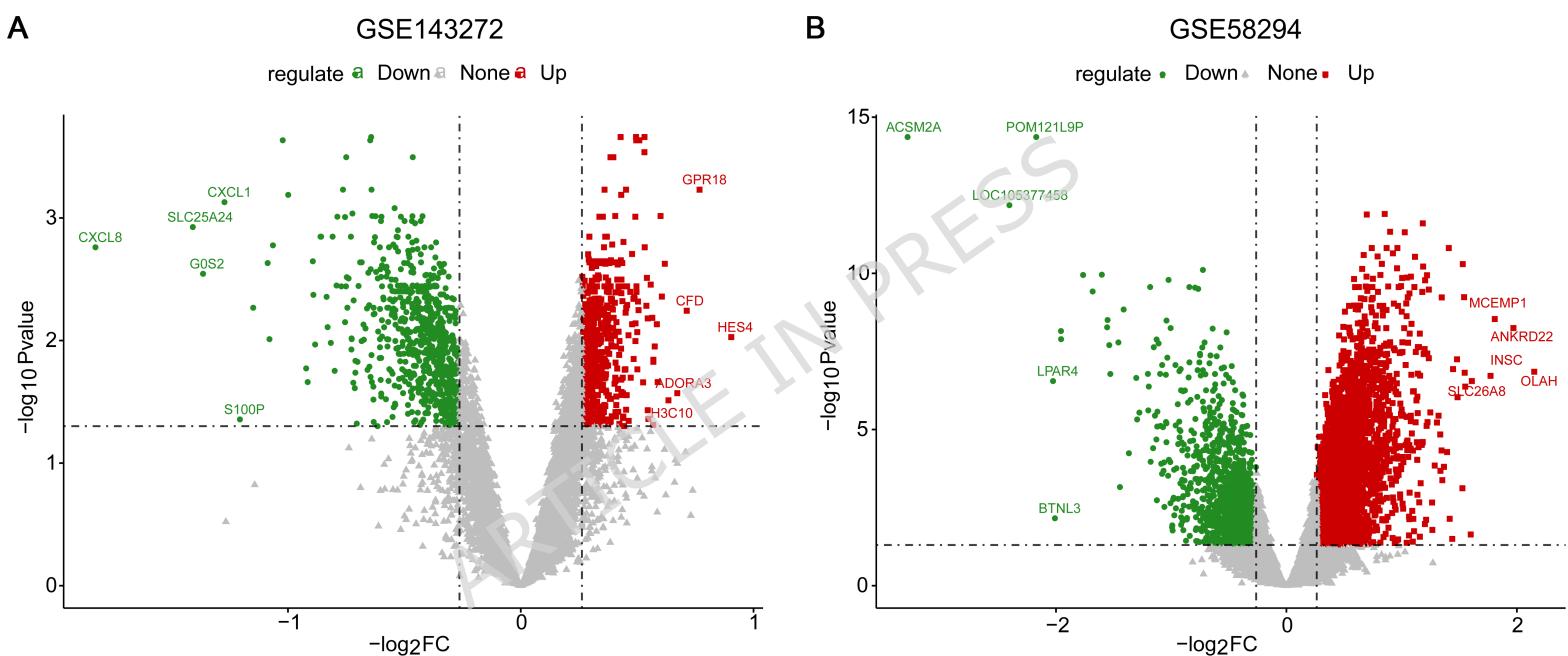
Table 1 The detail information about datasets

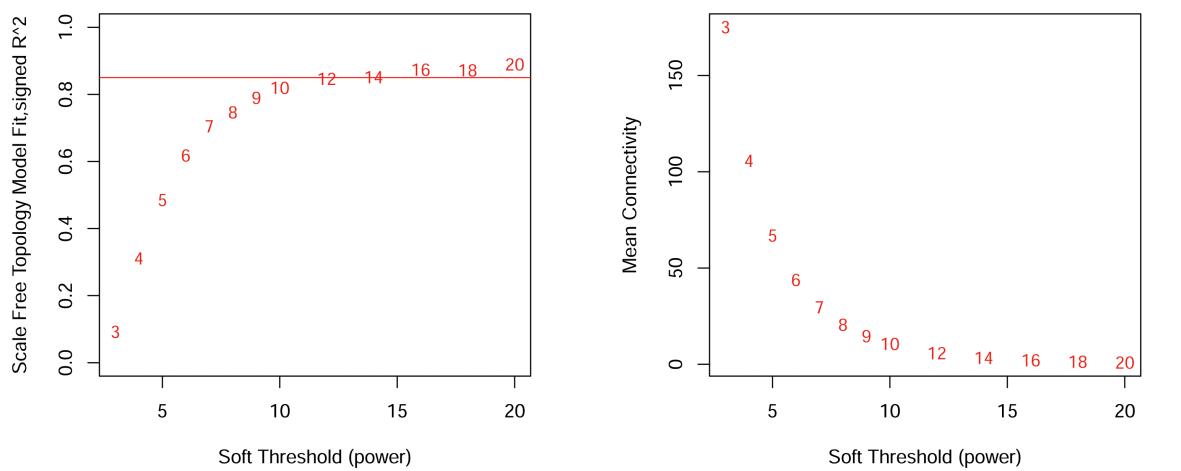
Disease	Datasets		Disease samples	Control samples	Platform
IS	GSE58294	Training	21 cardioembolic stroke (time after stroke: 3 h)	20 control	GPL570 [HG-U133_Plus_2] Affymetrix Human Genome U133 Plus 2.0 Array
	GSE16561	Validation	39 IS	24 control	GPL6883 Illumina HumanRef-8 v3.0 expression bead chip
Epilepsy	GSE143272	Training	12 epilepsy (4 idiopathic, symptomatic and cryptogenic epilepsy)	4 control	GPL570 [HG-U133_Plus_2] Affymetrix Human Genome U133 Plus 2.0 Array
	GSE16969	Validation	4 epilepsy	4 control	GPL570 [HG-U133_Plus_2] Affymetrix Human Genome U133 Plus 2.0 Array
	GSE205661	Validation	6 temporal lobe epilepsy with hippocampal sclerosis (TLE+HS)	9 control	GPL19072 Agilent-052909 CBC_lncRNAmRNA_V3 (Probe Name version)

Supplement Figure 1 Research flow chart

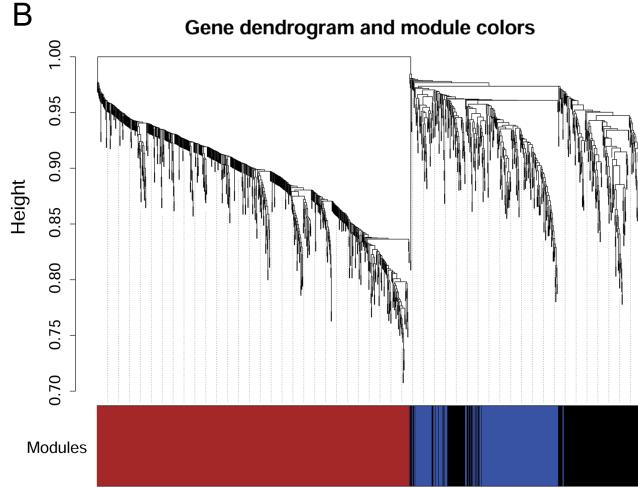


ARTICLE IN PRESS

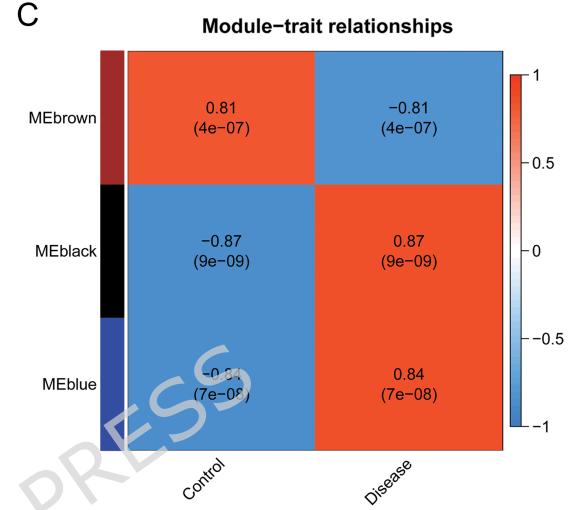




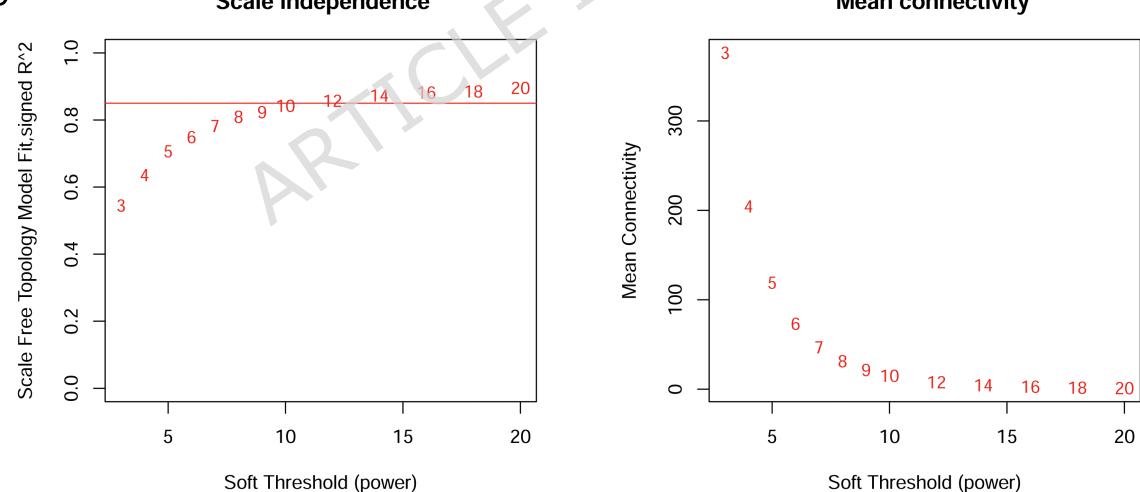
B



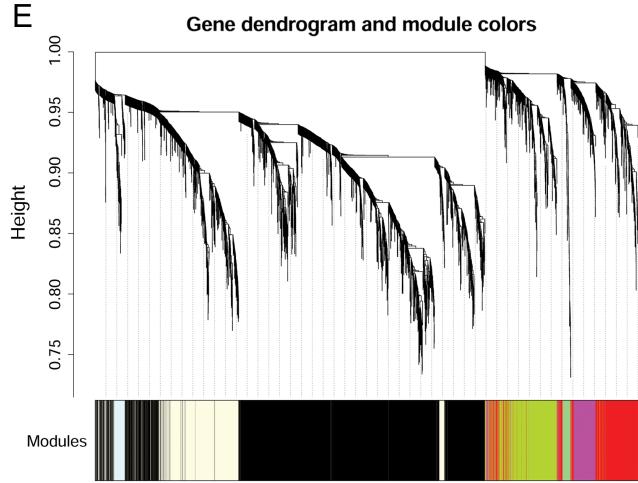
C



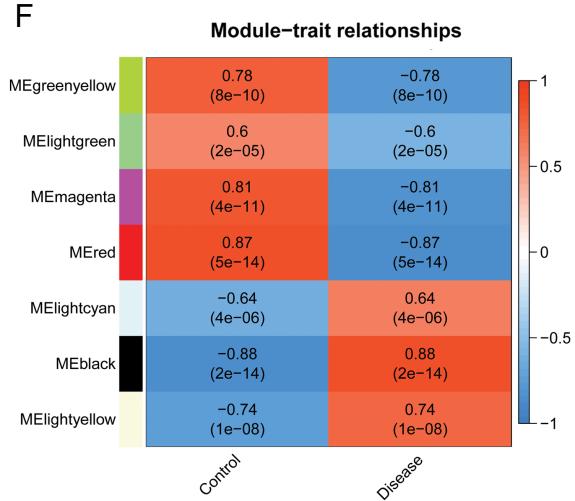
D

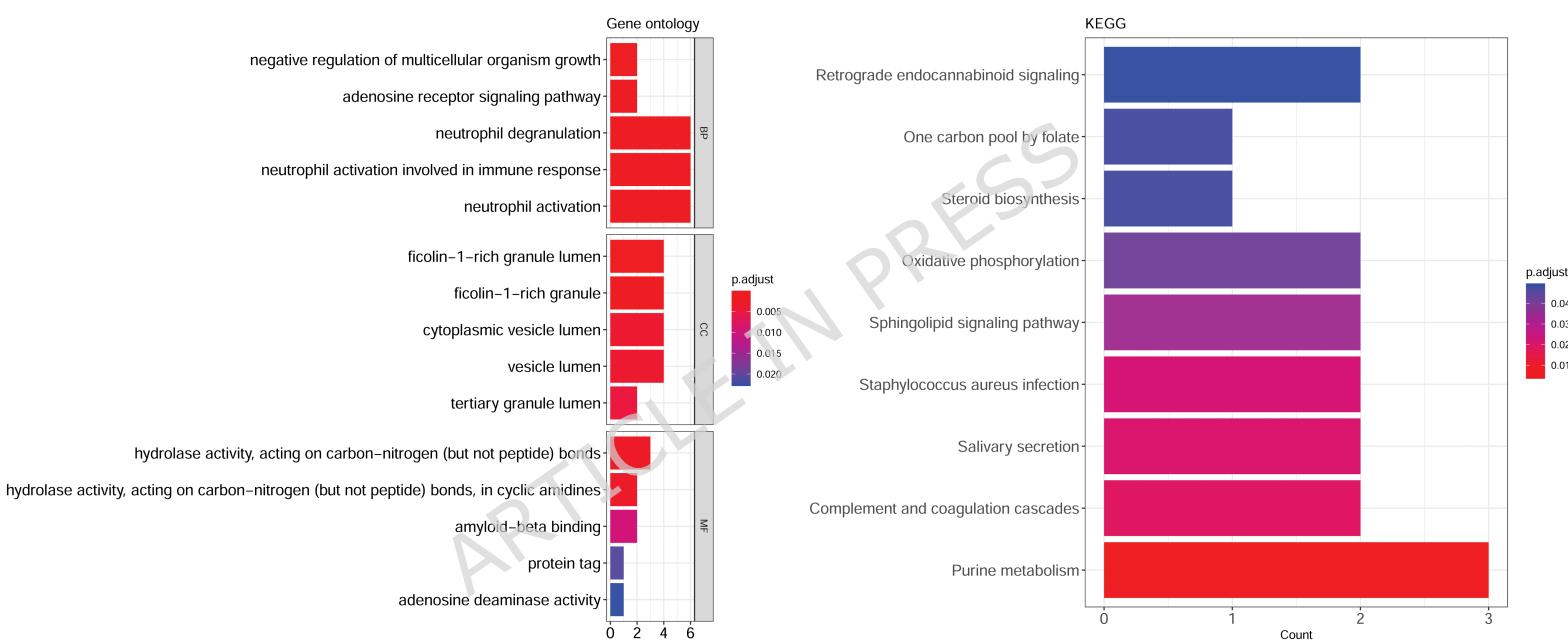


E

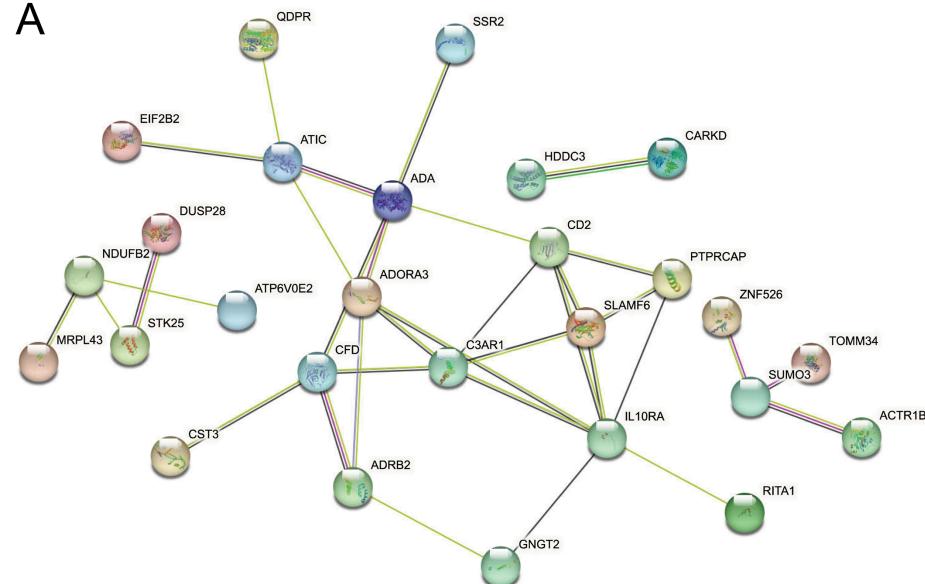


F

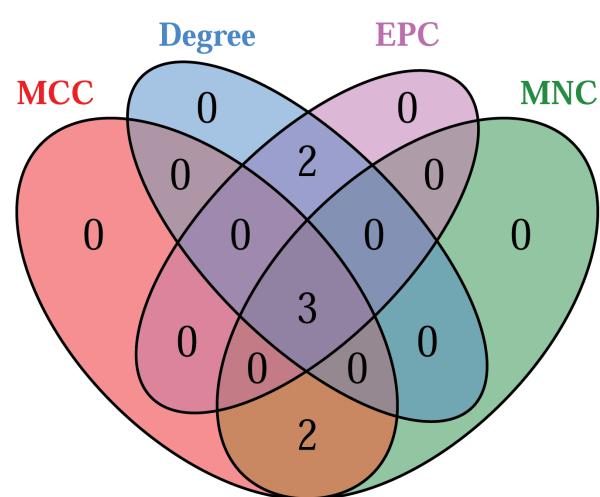




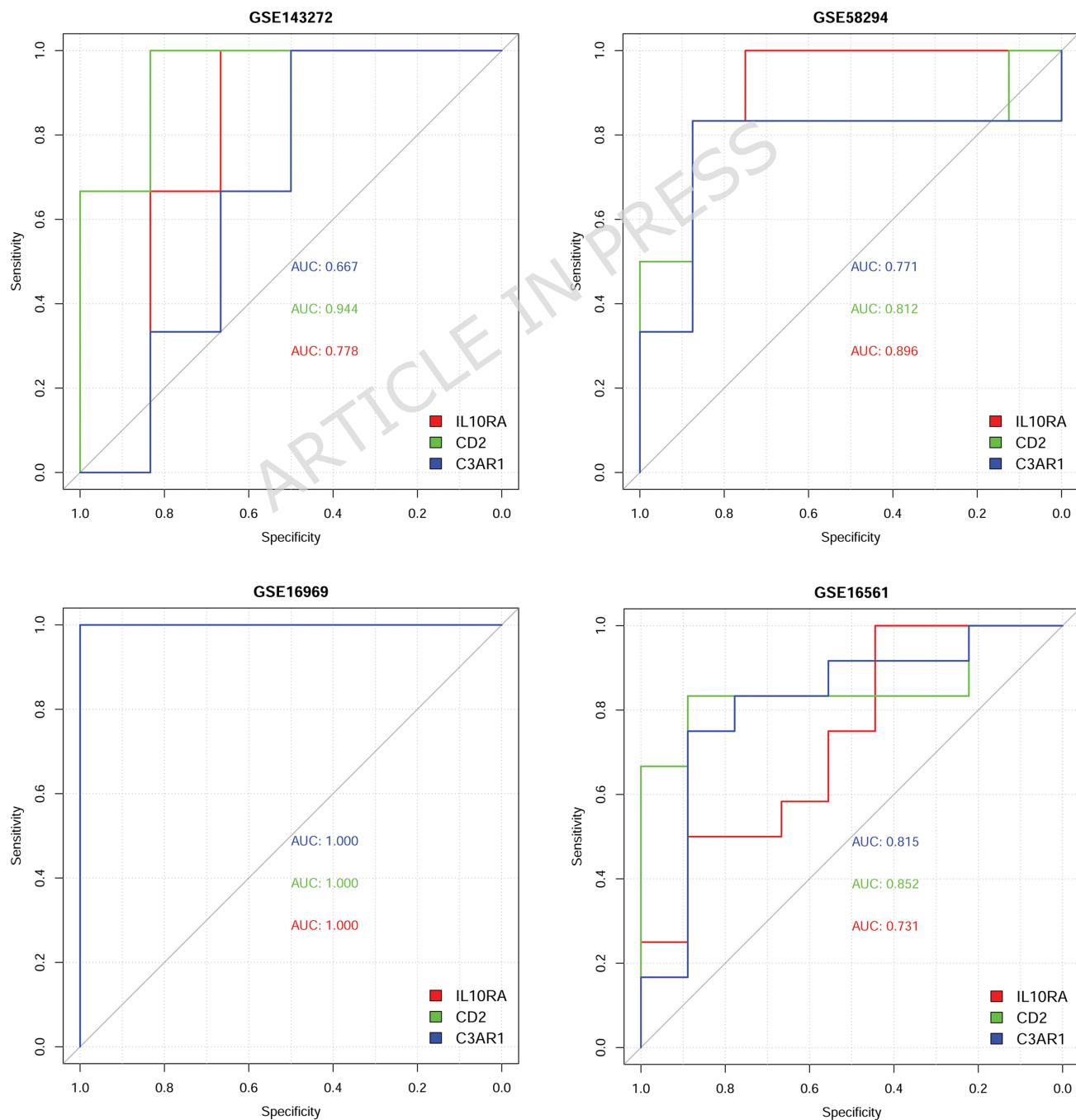
A

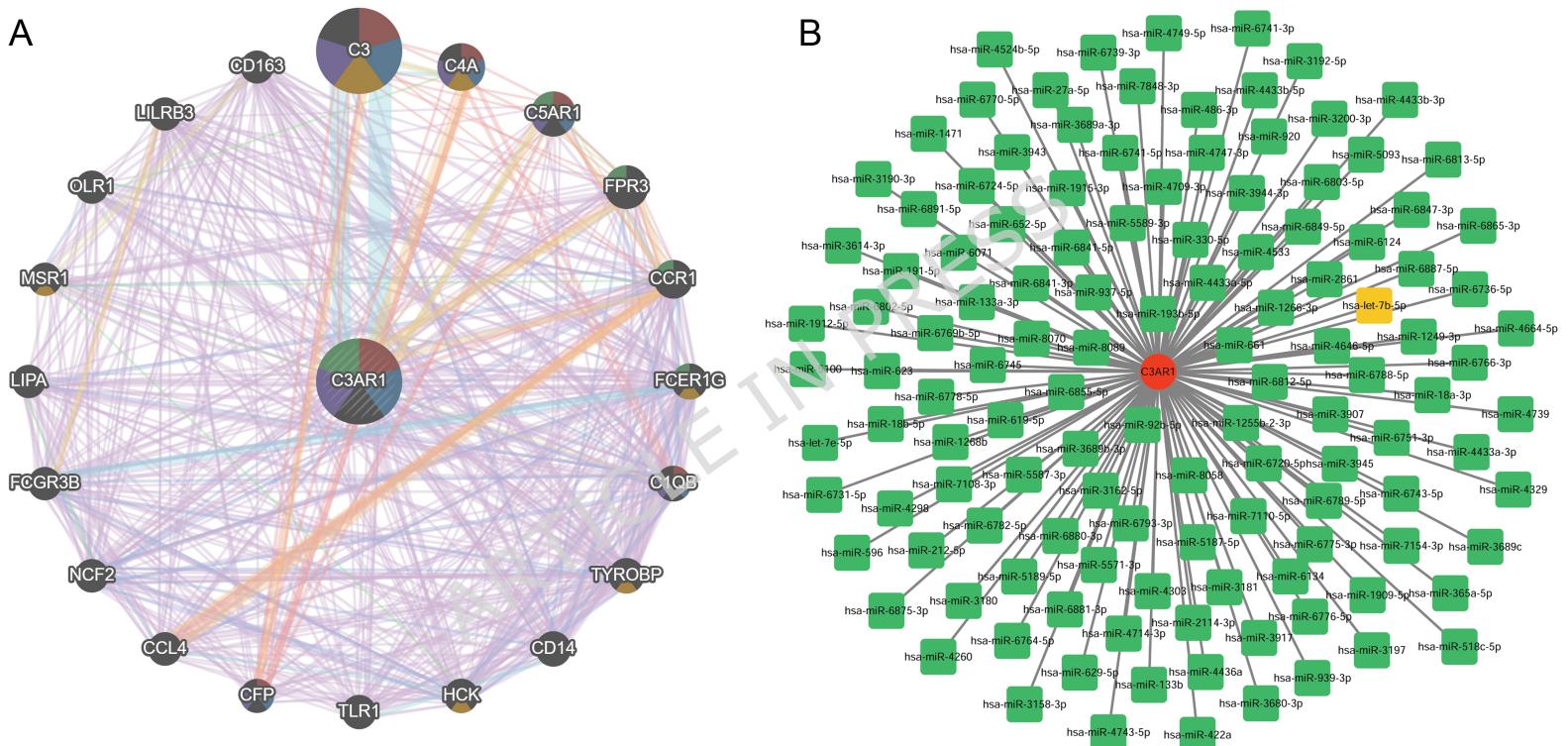


B

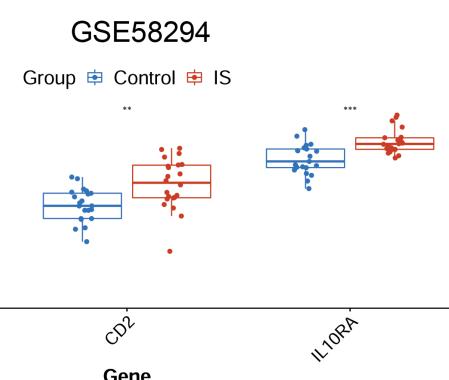
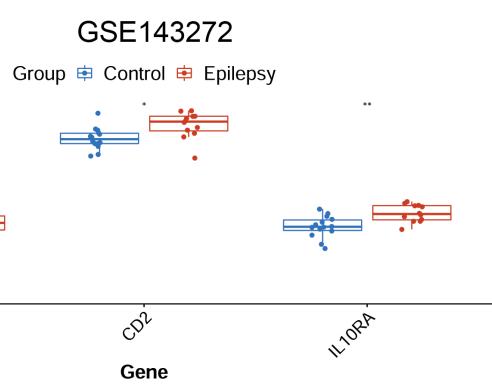


C

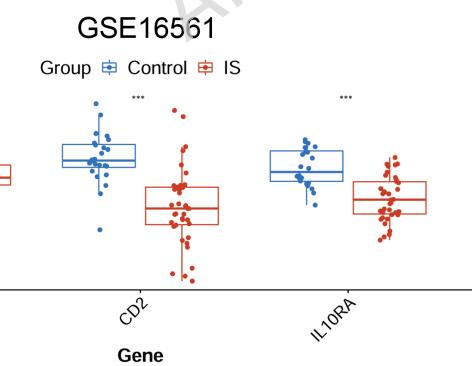
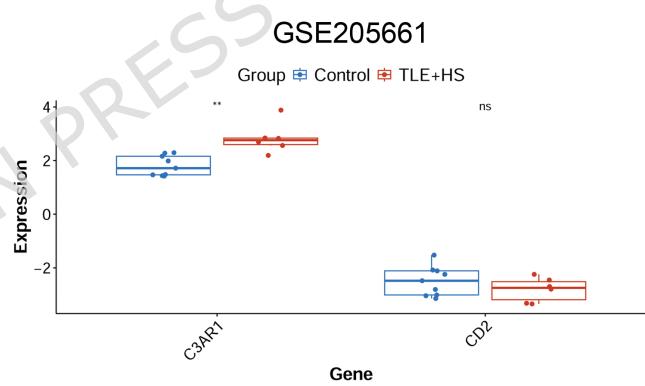
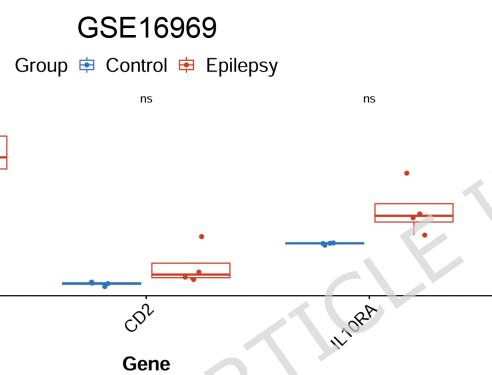




A



B



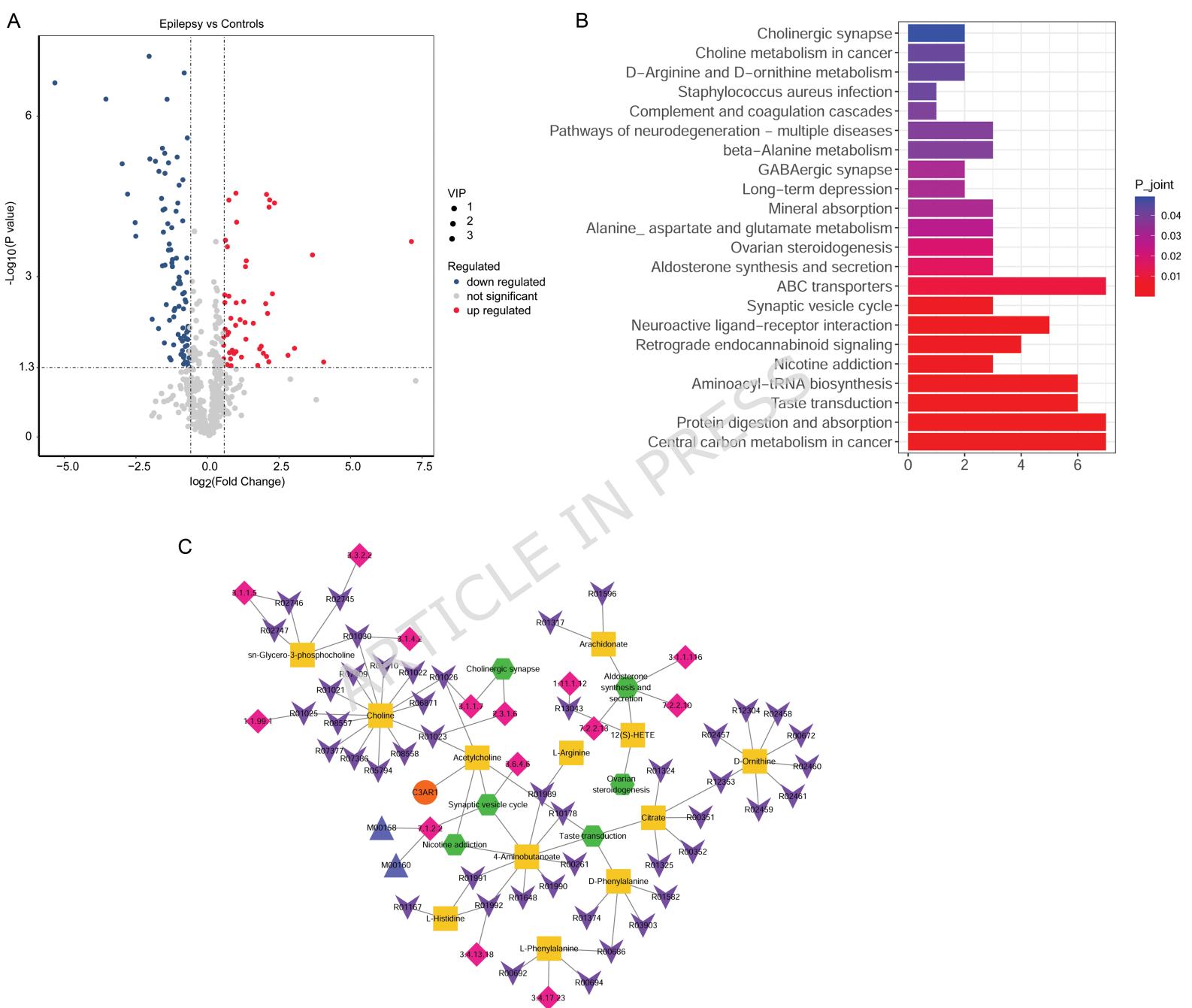


Table 1 The detail information about datasets

Disease	Datasets		Disease samples	Control samples	Platform
IS	GSE58294	Training	21 cardioembolic stroke (time after stroke: 3 h)	20 control	GPL570 [HG-U133_Plus_2] Affymetrix Human Genome U133 Plus 2.0 Array
	GSE16561	Validation	39 IS	24 control	GPL6883 Illumina HumanRef-8 v3.0 expression bead chip
Epilepsy	GSE143272	Training	12 epilepsy (4 idiopathic, 4 symptomatic and 4 cryptogenic epilepsy)	4 control	GPL570 [HG-U133_Plus_2] Affymetrix Human Genome U133 Plus 2.0 Array
	GSE16969	Validation	4 epilepsy	4 control	GPL570 [HG-U133_Plus_2] Affymetrix Human Genome U133 Plus 2.0 Array
	GSE205661	Validation	6 temporal lobe epilepsy with hippocampal sclerosis (TLE+HS)	9 control	GPL19072 Agilent-052909 CBC_lncRNAmRNA_V3 (Probe Name version)