Machine learning algorithms assisted identification of post-stroke depression associated biological features

Xintong Zhang[1†], Xiangyu Wang[2†], Shuwei Wang[3†], Yingjie Zhang[1†], Zeyu Wang[4†], Qingyan Yang[1], Song Wang[5], Risheng Cao[6]*, Binbin Yu[1]*, Yu Zheng[1]* and Yini Dang[7]*

[1]Department of Rehabilitation Medicine, The First Affiliated Hospital of Nanjing Medical University, Nanjing, Jiangsu, China, [2]Department of Rehabilitation Medicine, The Affiliated Lianyungang Oriental Hospital of Kangda College of Nanjing Medical University, Lianyungang, Jiangsu, China, [3]Department of Critical Care Medicine, Taizhou Hospital of Zhejiang Province Affiliated to Wenzhou Medical University, Taizhou, Zhejiang, China, [4]Department of Rehabilitation Medicine, Shanghai Ruijin Rehabilitation Hospital, Shanghai, China, [5]Department of Neurological Rehabilitation, Wuxi Yihe Rehabilitation Hospital, Wuxi, Jiangsu, China, [6]Department of Science and Technology, The First Affiliated Hospital of Nanjing Medical University, Nanjing, Jiangsu, China, [7]Department of Gastroenterology, The First Affiliated Hospital of Nanjing Medical University, Nanjing, Jiangsu, China

Objectives: Post-stroke depression (PSD) is a common and serious psychiatric complication which hinders functional recovery and social participation of stroke patients. Stroke is characterized by dynamic changes in metabolism and hemodynamics, however, there is still a lack of metabolism-associated effective and reliable diagnostic markers and therapeutic targets for PSD. Our study was dedicated to the discovery of metabolism related diagnostic and therapeutic biomarkers for PSD.

Methods: Expression profiles of GSE140275, GSE122709, and GSE180470 were obtained from GEO database. Differentially expressed genes (DEGs) were detected in GSE140275 and GSE122709. Functional enrichment analysis was performed for DEGs in GSE140275. Weighted gene co-expression network analysis (WGCNA) was constructed in GSE122709 to identify key module genes. Moreover, correlation analysis was performed to obtain metabolism related genes. Interaction analysis of key module genes, metabolism related genes, and DEGs in GSE122709 was performed to obtain candidate hub genes. Two machine learning algorithms, least absolute shrinkage and selection operator (LASSO) and random forest, were used to identify signature genes. Expression of signature genes was validated in GSE140275, GSE122709, and GSE180470. Gene set enrichment analysis (GSEA) was applied on signature genes. Based on signature genes, a nomogram model was constructed in our PSD cohort (27 PSD patients vs. 54 controls). ROC curves were performed for the estimation of its diagnostic value. Finally, correlation analysis between expression of signature genes and several clinical traits was performed.

Results: Functional enrichment analysis indicated that DEGs in GSE140275 enriched in metabolism pathway. A total of 8,188 metabolism associated genes were identified by correlation analysis. WGCNA analysis was constructed to obtain 3,471 key module genes. A total of 557 candidate hub genes were identified by interaction analysis. Furthermore, two signature genes (SDHD and FERMT3)

were selected using LASSO and random forest analysis. GSEA analysis found that two signature genes had major roles in depression. Subsequently, PSD cohort was collected for constructing a PSD diagnosis. Nomogram model showed good reliability and validity. AUC values of receiver operating characteristic (ROC) curve of SDHD and FERMT3 were 0.896 and 0.964. ROC curves showed that two signature genes played a significant role in diagnosis of PSD. Correlation analysis found that SDHD ($r = 0.653$, $P < 0.001$) and FERM3 ($r = 0.728$, $P < 0.001$) were positively related to the Hamilton Depression Rating Scale 17-item (HAMD) score.

Conclusion: A total of 557 metabolism associated candidate hub genes were obtained by interaction with DEGs in GSE122709, key modules genes, and metabolism related genes. Based on machine learning algorithms, two signature genes (SDHD and FERMT3) were identified, they were proved to be valuable therapeutic and diagnostic biomarkers for PSD. Early diagnosis and prevention of PSD were made possible by our findings.

Introduction

Stroke remains the second leading cause of death and may lead to long-term disability in adults (GBD 2019 Stroke Collaborators, 2021; Sun et al., 2021). After the acute stage, most of stroke patients suffer from physical and mental disabilities of varying degrees, including hemiplegia, reduced energy, and disturbed sleep (Zhang et al., 2013; Dong et al., 2021). Previous studies have shown that about 30–40% of stroke patients develop post-stroke depression (PSD) which is a mood disorder characterized by depression and anhedonia, and is associated with decreased rehabilitation motivation, reduced quality of life, poor functional outcome, as well as increased cost of treatment and burden of family caregiver (Li et al., 2020). One meta-analysis concluded that a hazard ratio for post-stroke depression and all-cause mortality was 1.59 (Cai et al., 2019). However, PSD is often concealed due to unrecognized depressive symptoms and their decreased willingness of treatment attendance (Klinedinst et al., 2012). Diagnosis of PSD is currently based on clinical symptoms, and there is no reliable objective parameter. Therefore, it is necessary to explore the new diagnostic and therapeutic biomarkers for PSD in subacute period of stroke.

There is accumulating evidence that PSD and metabolism are intimately related. Compared with non-PSD, stroke patients with PSD have higher glutamate levels in the frontal lobe (Wang et al., 2012). Previous studies found that a high level of homocysteine has been identified as the qualifiable risk factor for ischemic stroke, and elevated serum level of homocysteine is also significantly associated with depression (Li et al., 2017; Zhou et al., 2018). Jiang et al. (2021) demonstrated that gut microbiome may participate in the development of PSD, the discriminating fecal metabolites were mainly involved in lipid metabolism, amino acid metabolism, carbohydrate metabolism and nucleotide metabolism. These results indicated that metabolism plays an important role in the pathological process of PSD.

Recently with the assistant of advanced sequencing technologies and machine learning algorithms, intelligent hub gene and signaling pathway detection becomes realistic. Several studies based on weighted gene co-expression network analysis (WGCNA) have reported changes in relevant key pathways and differential expression of key related genes in post-stroke patients (Li et al., 2020; Wang et al., 2020; Lin et al., 2021). Furthermore, Liu et al. (2022) used WGCNA combined with the random forest model and the least absolute shrinkage and selection operator (LASSO) analysis to identify 10 key genes in patients with Alzheimer's Disease. However, these techniques have not been widely applied in the investigation of metabolism biomarkers of PSD.

Upon the above concerns, this study employed multiple bioinformatic approaches to find possible biomarkers. Firstly, three gene expression profiles of stroke were obtained from GEO database. Differentially expressed genes (DEGs) were detected. WGCNA was constructed to identify disease related module genes. Then, correlation analysis was performed to obtain metabolism related genes. Interaction analysis was performed to obtain candidate hub genes. Subsequently, signature genes were identified by LASSO and random forest analysis. Gene set enrichment analysis was applied on signature genes. Finally, a diagnosis model was built in PSD cohort. In general, the findings of this research may assist in the diagnosis and treatment of PSD as well as increase our understanding of etiology of PSD.

Materials and methods

Data sources and processing

Three datasets (GSE140275, GSE122709, GSE180470) were downloaded from Gene Expression Omnibus (GEO).[1] The GSE140275 dataset contained six patients, including three healthy controls (HC) and three stroke patients. The GSE122709 dataset

---

1 http://www.ncbi.nlm.nih.gov/geo/

FIGURE 1

Flow chart. GEO, Gene Expression Omnibus; WGCNA, weighted gene co-expression network analysis; PSD, post-stroke depression; ROC, receiver operating characteristic; GO, gene ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; GSEA, gene set enrichment analysis.
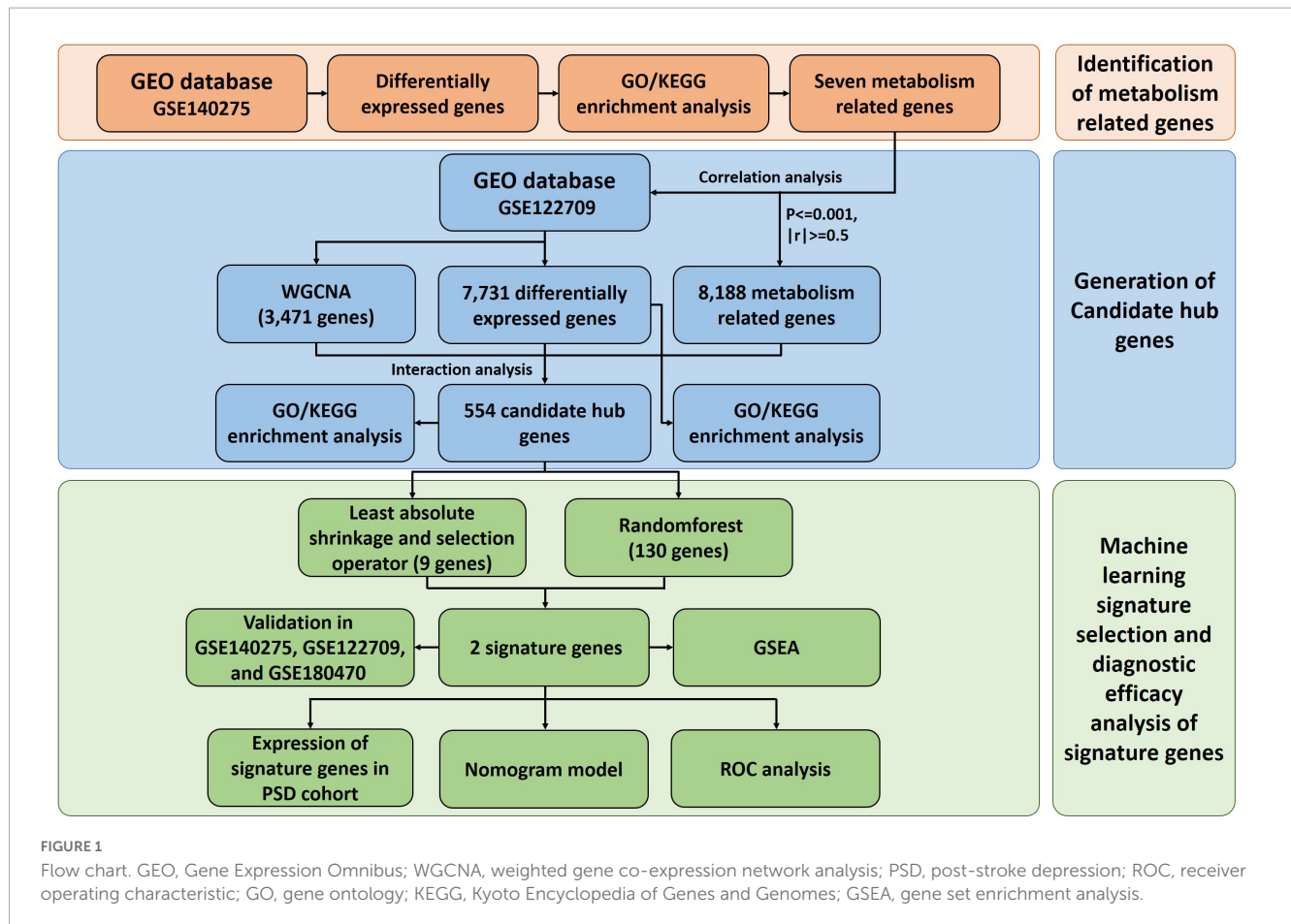
included five HC and ten stroke patients. GSE180470 dataset included three HC and three stroke patients. Expressions of three datasets were all derived from human blood tissue. "Limma" and "edgeR" package in R software was used to investigate differentially expressed genes (DEGs) (Robinson et al., 2010; Ritchie et al., 2015), which was specified as "P-value < 0.05 and $\log_2$ (fold change) > 1 or $\log_2$ (fold change) < -1." For visualization, the volcano plots were generated to show DEGs, while the top 25 upregulated and the top 25 downregulated DEGs were displayed by heatmaps.

Functional enrichment analysis

Functional enrichment analysis was conducted to evaluate major biological attributes of DEGs, specifically including Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis using "ClusterProfiler" package in R software. Threshold was set at P-value < 0.05. GO categories comprised biological processes (BP), molecular functions (MF), and cellular components (CC) (Zhu et al., 2022).

Weighted gene co-expression network analysis (WGCNA)

Based on scale-free topology criterion, co-expression network in GSE122709 dataset was constructed using "WGCNA" package

in R software to identify co-expression gene modules (Langfelder and Horvath, 2008). Briefly, genes with read counts less than 10 and "NA" were filtered out, top 5,000 variant genes were selected. Pearson's correlations between each gene pair were calculated to build an adjacency matrix. Afterward, a "soft" threshold power (β) was estimated according to the criteria of scale-free topology to construct a biologically important scale-free network. Dynamic Tree Cut algorithm was then used to identify gene modules (Lin et al., 2021). Module membership (MM) and gene significance (GS) were estimated to connect modules with clinical characteristics. Hub gene modules were designated as those with the highest Pearson module membership correlation and P-value < 0.05 (Liu et al., 2021).

Screening for candidate hub genes

Based on R software, "WGCNA" package was used for correlation analysis for genes in GSE122709 and seven genes associated with metabolism with the following parameters: | R| > 0.5, $P < 0.001$. Then, we identified candidate hub genes by the intersection of DEGs, key module genes and metabolism related genes. Finally, results were visualized by Venn diagram *via* online tool Venny 2.1.0² (He et al., 2021).

---

2   https://bioinfogp.cnb.csic.es/tools/venny/index.html

Identification for signature genes in patients with stroke

We screened candidate hub genes by the intersection of DEGs, key module genes and metabolism related genes. Subsequently, two machine learning algorithms, least absolute shrinkage and selection operator (LASSO) and random forest, were used to identify hub gene. LASSO, a dimension reduction approach, shows superiority in evaluating high-dimensional data in comparison to regression analysis (Kang et al., 2021). The "glmnet" package was used to implement LASSO analysis with a turning/penalty parameter utilizing a 10-fold cross-validation. Furthermore, the "random forest" package was used for performing the random forest analysis which determined the optimal number of variables by computing average error rate of candidate hub genes (Mantero and Ishwaran, 2021). A random forest tree model was built and the importance scores of each candidate hub genes were identified. Genes with importance value >0.25 were determined. The intersection genes of LASSO and random forest analysis were used to pick signature genes of patients with stroke.

Establishment of nomogram

The "rms" package was applied for incorporating signature genes to establish a nomogram. The "score" is the score of the relevant item below, and the "total score" is the sum of all the elements above. Calibration curves were used for assessing the predictive power of the model. Clinical usefulness of nomogram was evaluated by decision curve analysis, which determines clinical practicability of nomogram by quantifying the net benefits under different threshold probabilities in the validation set. Furthermore, we performed clinical impact curves to evaluate clinical utility of the model (Xu et al., 2021).

Curve analysis of receiver operating characteristics (ROC)

The "pROC" package was applied to create Receiver Operating Characteristic (ROC) curves to determine the area under the curve (AUC) for screening signature genes and evaluating their diagnostic value (Robin et al., 2011). AUCs of 0.5–0.7 were considered with low diagnostic accuracy, 0.7–0.9 were considered with moderate accuracy, and >0.9 indicates high accuracy.

Gene set enrichment analysis (GSEA)

To functionally investigate the biological significance of signature genes, GSEA (version 4.1.0) was performed in different subgroups. KEGG gene sets were chosen as the gene set database (Subramanian et al., 2005). Normalized enrichment score (NES) and false discovery rate (FDR) were used to determine if differences were statistically significant and cut-off values were FDR < 0.25, $P < 0.05$, and | NES| > 1.

PSD validation cohort

This was a cohort study enrolled at the First Affiliated Hospital of Nanjing Medical University from September 2020 to April 2022. It was approved by the Committee of Institutional Ethics (Institutional Review Board, 2018-SR-339) and all participants provided written informed consent prior to participation. Patients eligible for inclusion in the study were: (1) aged older than 18 years; (2) diagnosed with ischemic stroke on brain MRI; (3) with stable vital signs (Luft et al., 2004; Upreti et al., 2019). Patients were excluded if (1) presence of severe cognitive impairment; (2) participated in other clinical trials within 6 months (Shi et al., 2021; Yu et al., 2022).

All participants underwent an initial clinical assessment, including the collection of clinical and demographic information. Depression symptoms in post-stroke patients were evaluated by the Hamilton Depression Rating Scale 17-item (HAMD) at 1 month after stroke by a trained neurologist (Lin et al., 2020; Qiao et al., 2020). A score of 0–7 was considered normal, while a HAMD score ≥8 is indicative of depression. Stroke severity was measured using the National Institute of Health Stroke Scale (NIHSS) (He et al., 2020). Modified Rankin Scale (mRS) was used to estimate the functional disability (Liu et al., 2018). Independence and level of activities of daily life (ADL) were evaluated with the Barthel index (Kamal et al., 2020). For research purposes, a blood sample (10 ml) was taken from each subject for further ELISA assessment when they completed the HAMD assessment.

ELISA analysis

Concentration of signature genes in serum of stroke patients were measured using ELISA kit (antibodies-online, Philadelphia, PA, USA). Briefly, 100 µL standard or sample were added to each well and incubated for 90 min at 37°C. After washing two times, 100 µL Biotin-labeled antibody working solution was added and incubated for 60 min for 37°C. After plates were washed three times. A total of 100 µL SABC Working Solution was added and incubated for 30 min at 37°C. Subsequently, 90 µL TMB Substrate Solution was added and incubated 20 min at 37°C. After the incubation, 50 µL stop solution was added into each well to stop the reaction. Finally, Absorbance value at 450 nm was read immediately and calculation (Kaida et al., 2013; Zhou et al., 2015).

Statistical analysis

All statistical analyses in our study were implemented using R software (version 4.1.2). The difference between the two groups was analyzed by Student's *t*-test. The correlation between genes in GSE122709 and metabolism related genes was determined using Pearson's correlation test. All statistical *P*-values were two-sided, and statistical significance was considered with *P*-value < 0.05.

Results

Detailed procedure of our study is shown in **Figure 1**.

FIGURE 2

DEGs screening and functional enrichment analysis. **(A)** Volcano plot of differentially expressed genes in GSE140275. **(B)** Heatmap of differentially expressed genes in GSE140275. **(C)** KEGG pathway analyses of upregulated mRNAs in GSE140275. **(D)** GO functional analyses of upregulated mRNAs in GSE140275. **(E)** Volcano plot of differentially expressed genes in GSE122709. **(F)** Heatmap of differentially expressed genes in GSE122709. **(G)** KEGG pathway analyses of downregulated mRNAs in GSE140275. **(H)** GO functional analyses of downregulated mRNAs in GSE140275. **(I)** KEGG pathway analyses of mRNAs in GSE122709. **(J)** GO functional analyses of mRNAs in GSE122709. GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; DEGs, differentially expressed genes.

Identification of DEGs between HC and stroke patients

To identify potential DEGs, expression profiles of GSE140275 and GSE122709 in GEO database were performed using "Limma" package with $P < 0.05$ and $|logFC| > 1$ as threshold. A total of 1,724 DEGs were screened in GSE140275 including 861 upregulated genes and 863 downregulated genes (Supplementary Table 1). A total of 7,731 DEGs were obtained, of which 3,516 genes presented upregulation and 4,215 genes presented downregulation in GSE122709 (Supplementary Table 2). The volcano plots were demonstrated in Figures 2A, E. The heatmap showed the top 25 upregulated and top 25 downregulated DEGs between healthy control and stroke patients (Figures 2B, F).
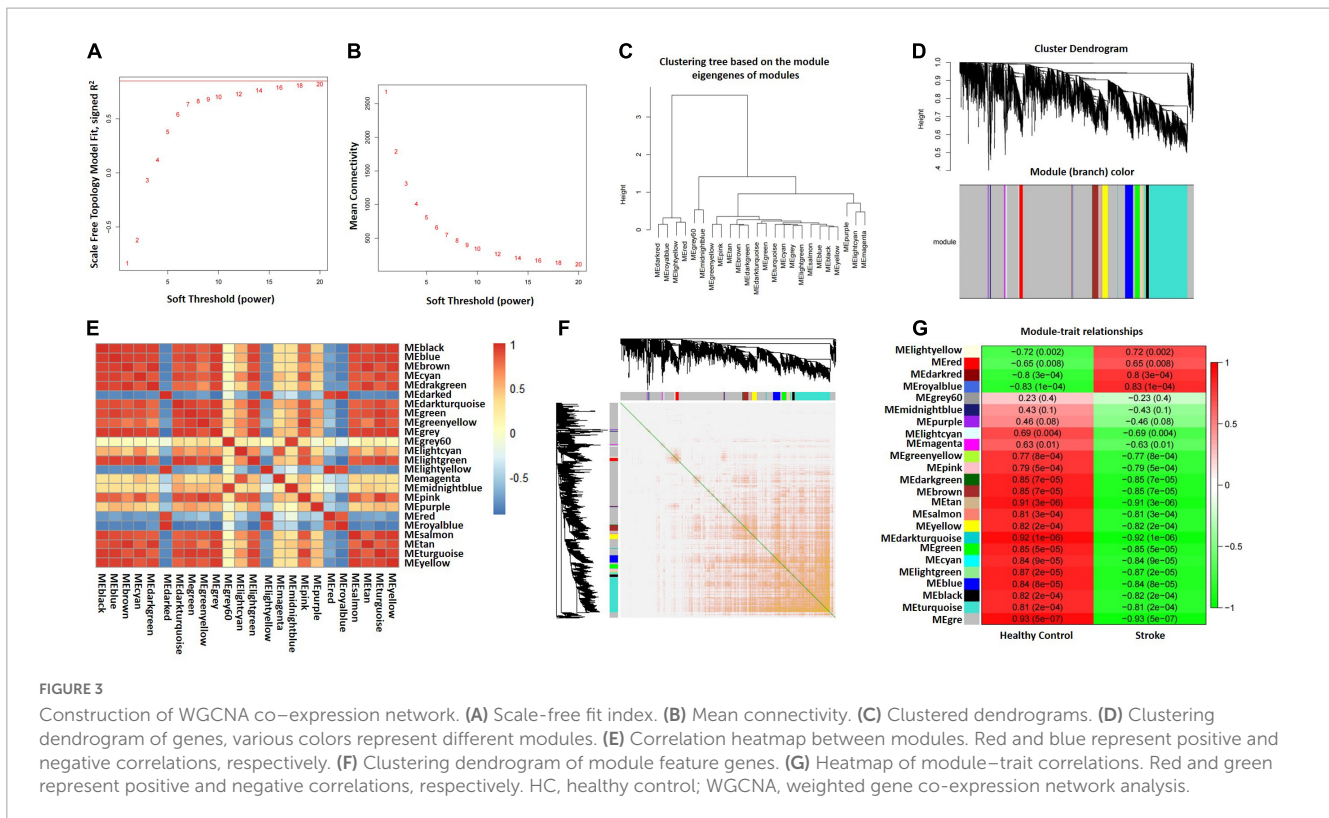
Functional enrichment analysis of DEGs in GSE140275

Functional enrichment analysis was carried out to investigate the biological functions of DEGs in GSE140275. Among upregulated DEGs, KEGG enrichment analysis demonstrated that "autophagy," "porphyrin metabolism," and "glycine, serine and threonine metabolism" were highly enriched (Figure 2C); GO analysis showed that multiple metabolic pathways were also significantly enriched in biological processes, such as

"monoacylglycerol metabolic process," "acylglycerol metabolic process," and "glycerolipid metabolic process" (Figure 2D). The results of KEGG showed downregulated DEGs were especially enriched in "ribosome," "protein export," and "T cell receptor signaling pathway" (Figure 2G). Additional GO analysis suggested downregulated DEGs were significantly enriched in "structural constituent of ribosome" in MFs, "ribosome" in CCs, and "regulation of leukocyte mediated immunity" in BPs (Figure 2H). Similarly, KEGG pathways analysis of GSE122709 showed that "porphyrin metabolism", and "glycine, serine and threonine metabolism" were significantly enriched (Figures 2I, J), indicating that metabolism played an important role in stroke.

Construction of the weighted gene co-expression network

The GSE122709 dataset (five HC and 10 stroke patients) was obtained for WGCNA analysis to identify modules of highly correlated genes. A scale-free co-expression network was constructed with the soft threshold to 20 and the mean connectivity is relatively favorable (Figures 3A, B). We selected 0.25 as clustering height limit to merge the strongly associated modules (Figure 3C). Subsequently, 24 signature modules were identified and labeled with different colors (Figure 3D). The correlation between modules was computed, and the results were showed

FIGURE 3
Construction of WGCNA co−expression network. **(A)** Scale-free fit index. **(B)** Mean connectivity. **(C)** Clustered dendrograms. **(D)** Clustering dendrogram of genes, various colors represent different modules. **(E)** Correlation heatmap between modules. Red and blue represent positive and negative correlations, respectively. **(F)** Clustering dendrogram of module feature genes. **(G)** Heatmap of module−trait correlations. Red and green represent positive and negative correlations, respectively. HC, healthy control; WGCNA, weighted gene co-expression network analysis.

in Figure 3E. In addition, transcription correlation analysis was performed and demonstrated that there was no substantive connection between modules (Figure 3F). Finally, we calculated the correlation between each module and clinical features. Results indicated that the MEroyalblue module was negatively correlated with HC ($r = -0.83$, $P = 1e-04$) and positively correlated with stroke ($r = 0.83$, $P = 1e-04$), while the Megrey module was negatively correlated with stroke ($r = -0.93$, $P = 5e-07$) and positively correlated with healthy control ($r = 0.93$, $P = 5e-07$) (Figure 3G and Supplementary Table 3). Therefore, Meroyalblue and Megrey modules were identified as clinically meaningful modules.

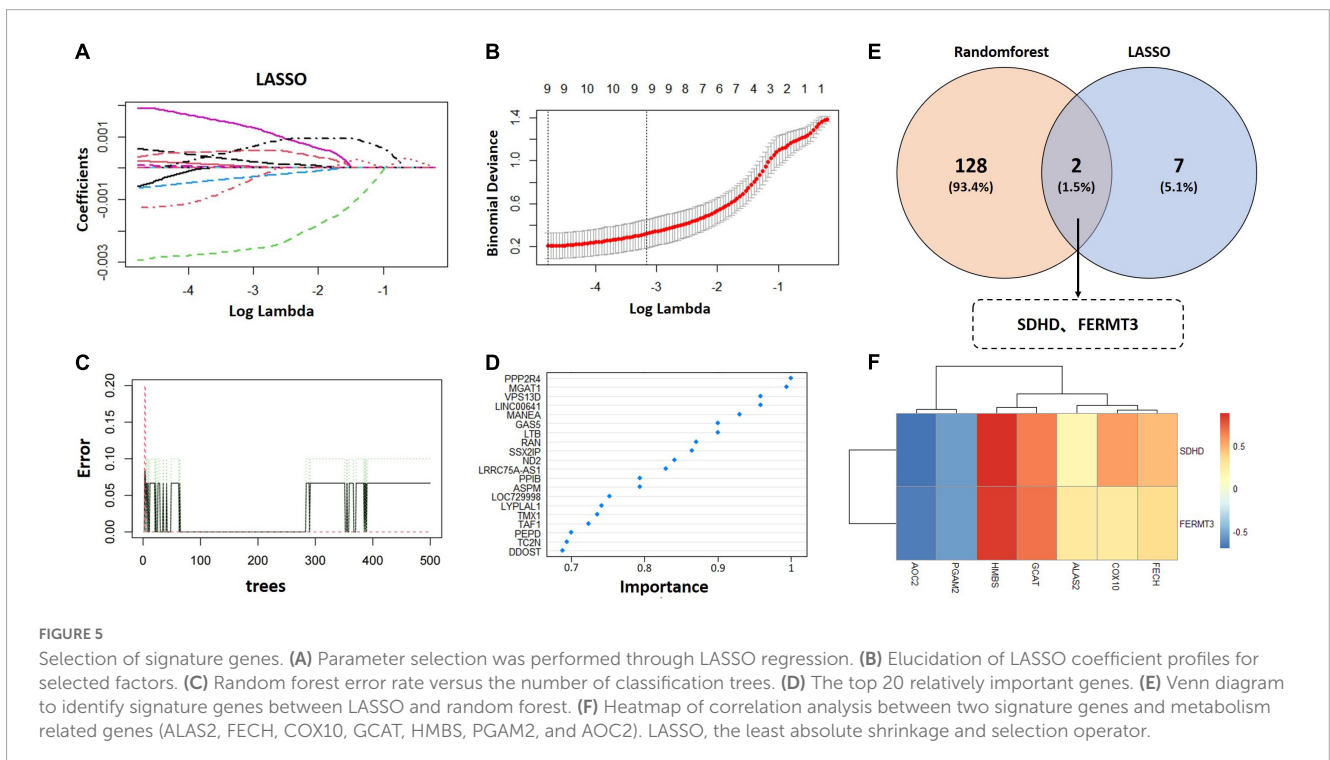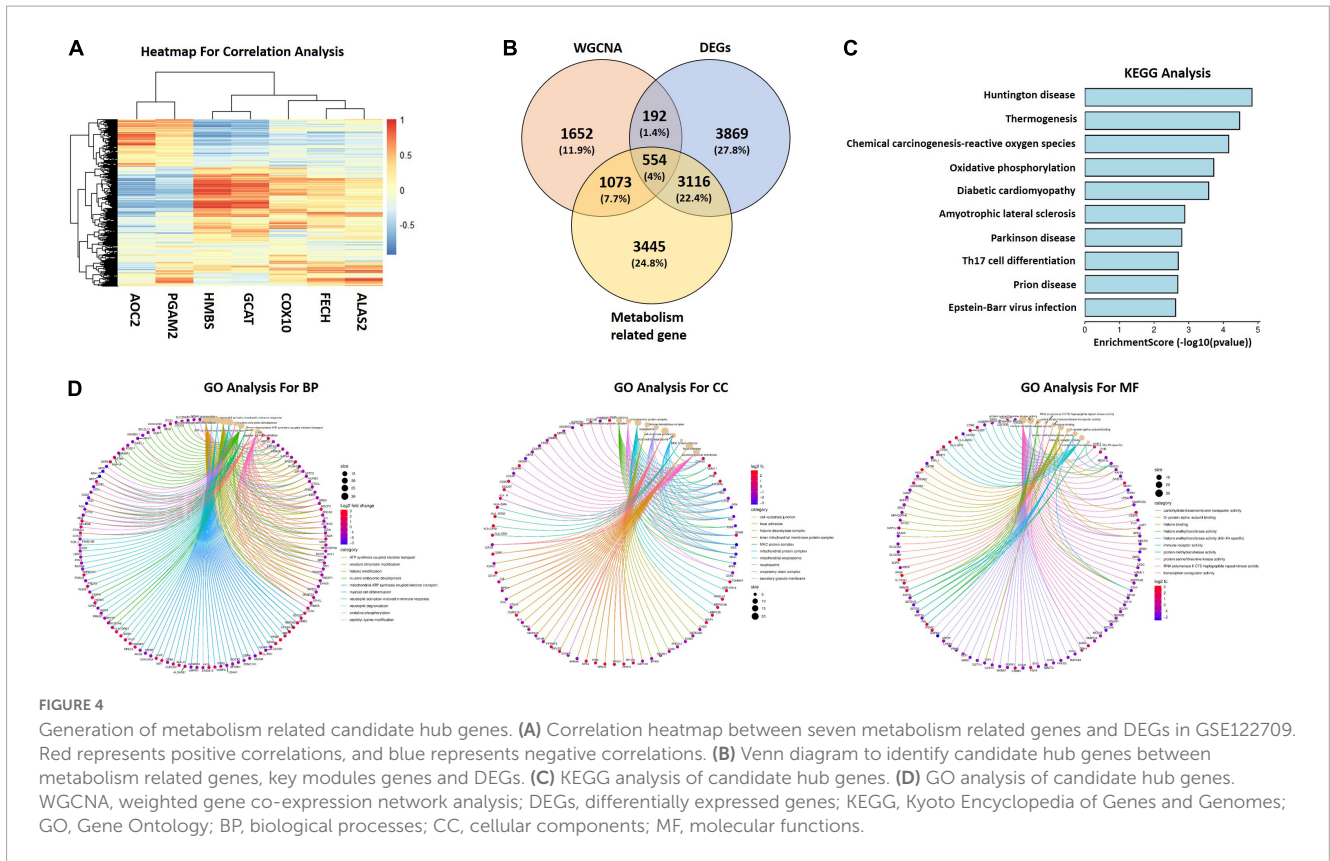## Identification of metabolism related candidate hub genes

Based on KEGG pathway analysis in GSE140275, we extracted porphyrin metabolism and glycine, serine and threonine metabolism related genes (ALAS2, FECH, COX10, GCAT, HMBS, PGAM2, and AOC2). Correlation analysis between seven genes and all genes in GSE122709 dataset was conducted. A total of 8,188 metabolism related genes were identified ($|r| \geq 0.5$, $P \leq 0.001$). The heatmap of correlation analysis were shown in Figure 4A. Subsequently, we interacted DEGs in GSE122709, genes in Meroyalblue and Megrey modules, and metabolism related genes, 554 common genes were obtained as metabolism related candidate hub genes (Figure 4B). Functional enrichment analysis revealed that metabolism related candidate hub genes were enriched in "oxidative phosphorylation," "ATP synthesis coupled electron transport," "cell-substrate junction," and "carbohydrate transmembrane transporter activity" (Figures 4C, D).

Selection of signature genes *via* machine learning algorithms

Least absolute shrinkage and selection operator and random forest algorithms were applied to identify signature genes from 554 metabolism related candidate hub genes. For LASSO analysis, nine signature genes were selected from statistically significant univariate variables (Figures 5A, B and Supplementary Table 4). For random forest analysis, we set importance value to 0.25 as the threshold and 130 signature genes were determined (Figures 5C, D and Supplementary Table 5). The interaction analysis of LASSO and random forest indicated that two signature genes were finally screened out, including succinate dehydrogenase complex subunit D (SDHD) and fermitin family member 3 (FERMT3) (Figure 5E). Finally, correlation analysis of two signature genes and metabolism related genes (ALAS2, FECH, COX10, GCAT, HMBS, PGAM2, and AOC2) demonstrated that SDHD and FERMT3 were significantly correlated with metabolism (Figure 5F).

Validation of signature genes

We further investigated the role of SDHD and FERMT3. The expression of SDHD and FERMT3 was verified in GSE140275 and GSE122709. The results showed that SDHD was substantially upregulated in the stroke group, while the same trend was seen in expression of FERMT3 (Figures 6A, B). To further confirm the reliability of our results, validation dataset (GSE180470) was used to validate expression of SDHD and FERMT3. SDHD and FERMT3 were highly expressed in the stroke group (Figure 6C), suggesting that these genes may play a significant role in stroke.

FIGURE 4
Generation of metabolism related candidate hub genes. **(A)** Correlation heatmap between seven metabolism related genes and DEGs in GSE122709. Red represents positive correlations, and blue represents negative correlations. **(B)** Venn diagram to identify candidate hub genes between metabolism related genes, key modules genes and DEGs. **(C)** KEGG analysis of candidate hub genes. **(D)** GO analysis of candidate hub genes. WGCNA, weighted gene co-expression network analysis; DEGs, differentially expressed genes; KEGG, Kyoto Encyclopedia of Genes and Genomes; GO, Gene Ontology; BP, biological processes; CC, cellular components; MF, molecular functions.

FIGURE 5
Selection of signature genes. **(A)** Parameter selection was performed through LASSO regression. **(B)** Elucidation of LASSO coefficient profiles for selected factors. **(C)** Random forest error rate versus the number of classification trees. **(D)** The top 20 relatively important genes. **(E)** Venn diagram to identify signature genes between LASSO and random forest. **(F)** Heatmap of correlation analysis between two signature genes and metabolism related genes (ALAS2, FECH, COX10, GCAT, HMBS, PGAM2, and AOC2). LASSO, the least absolute shrinkage and selection operator.

GSEA analysis of signature genes

Gene set enrichment analysis was performed for evaluating signaling pathways involved in the signature genes. The results showed that SDHD was significantly correlated with "emotional lability," "depression," and "abnormal fear anxiety related behavior" (Figure 6D). Meanwhile, "depression," "emotional blunting," and "abnormal fear anxiety related behavior" were detected for FERMT3 (Figure 6E). The results indicated that SDHD and FERMT3
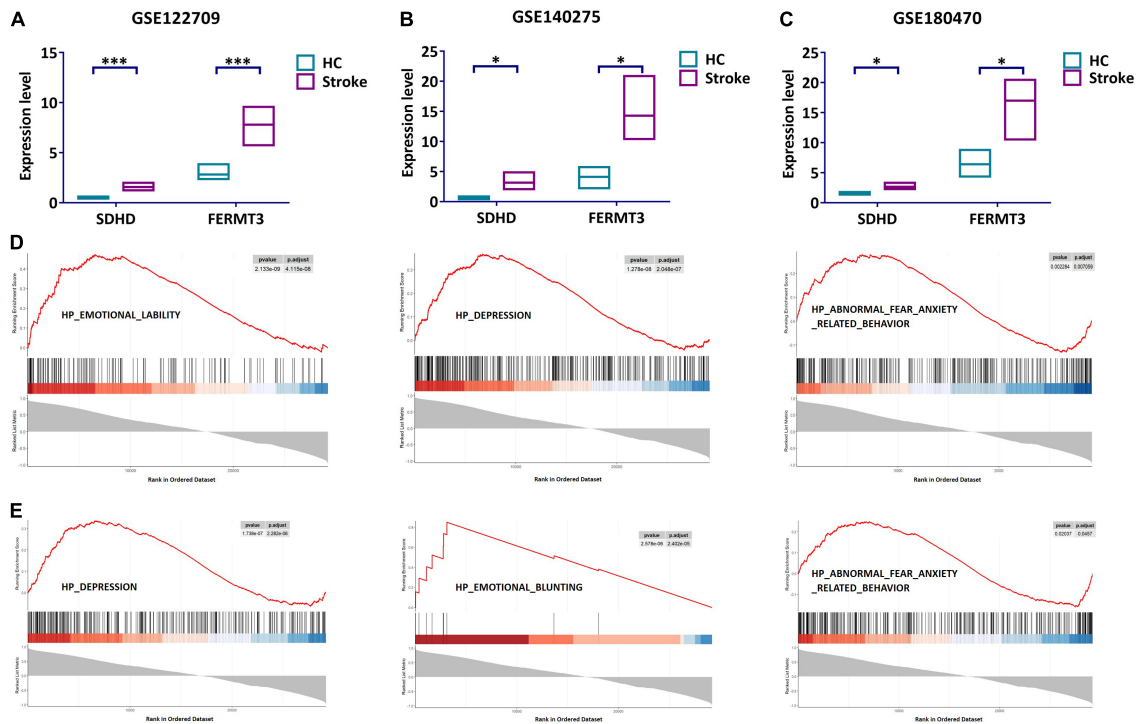
FIGURE 6

Validation and GSEA analysis of signature genes. **(A)** Expression level of SDHD and FERMT3 in GSE122709. **(B)** Expression level of SDHD and FERMT3 in GSE140275. **(C)** Expression level of SDHD and FERMT3 in GSE180470. **(D)** GSEA analysis of SDHD. **(E)** GSEA analysis of FERMT3. *$P < 0.05$; ***$P < 0.001$. HC, healthy control; GSEA, gene set enrichment analysis.

played a key role for diagnosis of psychosocial state in stroke patients.

Diagnostic efficacy of signature genes in PSD patients

Based on GSEA analysis of two signature genes (SDHD and FERMT3), we found that they have a significant correlation with depression. Therefore, we collected 81 stroke patients who were assigned into the PSD group (mean HAMD score = 14.74) and non-PSD group (mean HAMD score = 3.41). There was no difference in baseline clinical features between groups (**Supplementary Table 6**). Meanwhile, expression of serum SDHD and FERMT3 in all patients were detected by ELISA kit. SDHD and FERMT3 presented higher expression in the PSD group than the non-PSD group (**Figures 7A, B**), indicating their potential roles in diagnosis of depression in stroke patients. To predict diagnostic performance of signature genes in stroke patients with depression, the nomogram model for the signature genes (SDHD and FERMT3) was built using "rms" package (**Figure 7C**). Calibration curves demonstrated that the difference between the real and predicted depression risks was very minimal, indicating the nomogram model enabled an accurate estimation (**Figure 7D**). In addition, decision curves analysis demonstrated that the nomogram provided a greater clinical benefit (**Figure 7E**). The ROC curve also showed that the model was able to help clinicians accurately diagnose depression of stroke patients (**Figure 7F**). Additionally, correlation analysis between two signature genes and several clinical traits (HAMD,

NIHSS, BI, and mRS) indicated that SDHD ($r = 0.653$, $P < 0.001$) and FERM3 ($r = 0.728$, $P < 0.001$) were positively related HAMD, while SDHD also displayed a negative association with Barthel index ($r = -0.224$, $P = 0.044$) (**Figures 7G, H**).

Discussion

In this study, we included three datasets (GSE140275, GSE122709, GSE180470) with 27 patients for data analysis. We first screened 1,724 DEGs in GSE140275 including 861 upregulated genes and 863 downregulated genes. Subsequent KEGG enrichment analysis showed "porphyrin metabolism" and "glycine, serine and threonine metabolism" were highly enriched. Recent research reveals that stroke causes systemic complications, including hyperlipemia, high blood viscosity, dysfunctional gut microbiota, and a leaky gut (Yamashiro et al., 2017; Chen et al., 2019a). Chen et al. (2019b) demonstrated that stroke would cause gut microbiota dysbiosis, translocation of gut microbiota, and disruption to the gut barrier. And supplementation of short chain fatty acids (SCFAs), especially butyric acid, could remodel the gut microbiota and treat stroke (Chen et al., 2019a). Moreover, with the development of biology, metabolomics was applied to explore biomarkers and mechanisms of stroke by identifying metabolic alterations. Several studies reported the increase in ketone bodies levels in rats with stroke compared with sham group (Chen et al., 2019c; Wang et al., 2019). Fu et al. (2019) reported a decrease in β-hydroxybutyric acid level in serum but an increase in brain tissue in stroke rats, providing more energy for brain. These

FIGURE 7
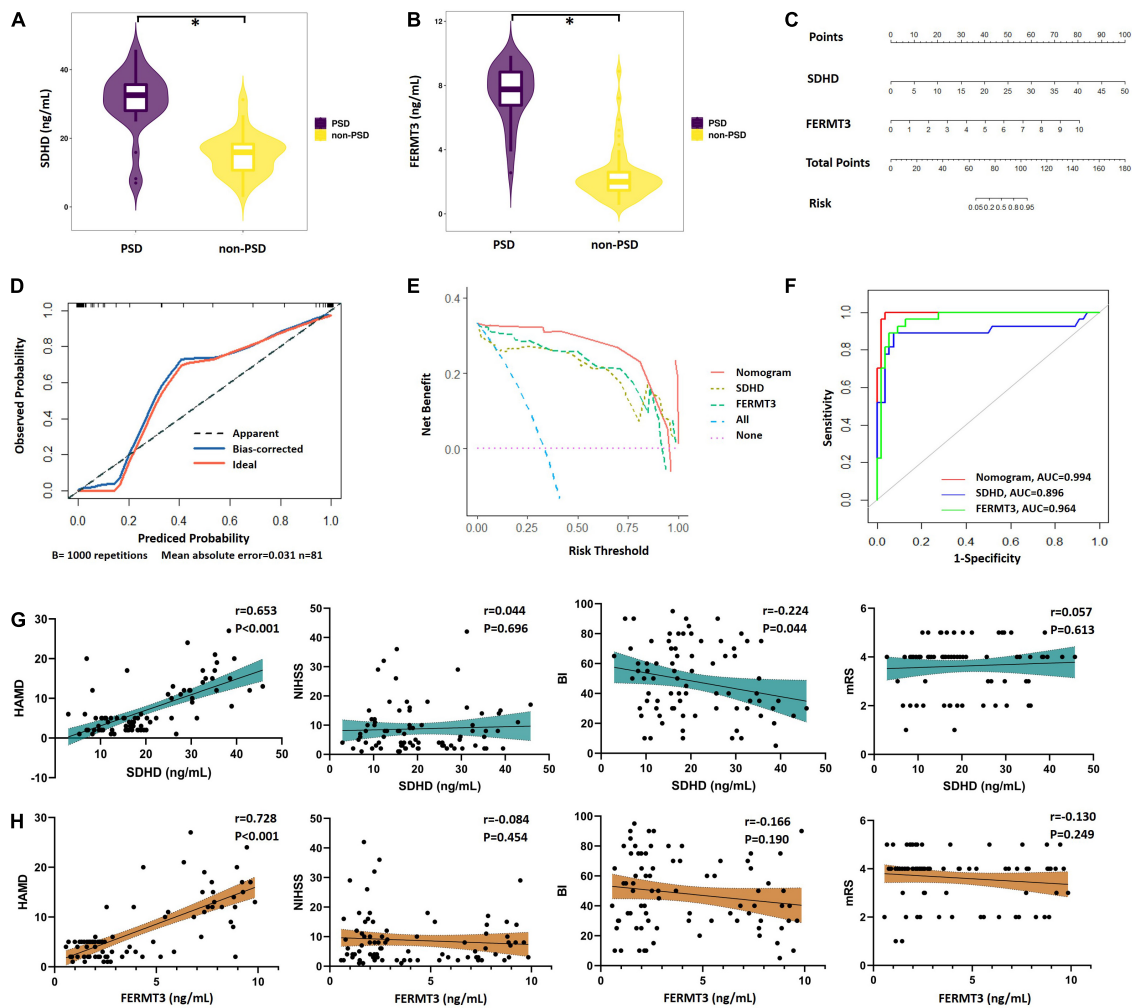
Performance of signature genes in PSD patients. **(A)** Expression level of SDHD in PSD and non-PSD groups. **(B)** Expression level of FERMT3 in PSD and non-PSD groups. **(C)** Nomograms for the prediction of the PSD risk. **(D)** Scatter diagram of calibration plot for internal verification of the nomogram model. **(F)** DCA curves of the nomogram model. **(E)** ROC curves of the nomogram model. **(G)** Correlation between expression of SDHD and four clinical traits (HAMD, NIHSS, BI, and mRS). **(H)** Correlation between expression of FERMT3 and four clinical traits (HAMD, NIHSS, BI, and mRS). PSD, post-stroke depression; DCA, decision curve analysis; ROC, receiver operating characteristic; AUC, area under the curve; HAMD, Hamilton Depression Rating Scale 17-item; NIHSS, National Institute of Health Stroke Scale; BI, Barthel index; mRS, Modified Rankin Scale. *$P < 0.05$.

studies suggest that metabolism features strongly correlate with prevention, diagnosis and treatment of stroke. Based on the role of metabolism in stroke, we extracted seven genes related to the "porphyrin metabolism" and "glycine, serine and threonine metabolism" pathways, including ALAS2, FECH, COX10, GCAT, HMBS, PGAM2, and AOC2. We then performed correlation analysis between these genes enriched in these two pathways in GSE140275 and all genes in GSE122709 to identify metabolism related genes. A total of 8,188 metabolism related genes were identified. Nevertheless, with the help of advanced bioinformatic approaches genetic information could be further derived.

Weighted gene co-expression network analysis (WGCNA) is a frequently applied method to identify co-expression pattern at whole transcriptome level. Wang et al. (2019) performed WGCNA analysis to investigate co-expression modules related with osteosarcoma and found genes in brown module might be related with carcinogenesis of osteosarcoma. In addition, there were several studies screened key module genes related to stroke by WGCNA

analysis (Fan et al., 2022; Zheng et al., 2022). However, metabolism related pathways and key genes in stroke are seldomly identified. Therefore, we performed WGCNA analysis of GSE122709 to identify 24 gene modules. No significant correlation between dividing modules was found. Module-traits relationship analysis indicated that Meroyalblue and Megrey modules were significantly associated with stroke disease. After this step, we interacted DEGs in GSE122709, genes in Meroyalblue and MEgrey modules, and metabolism related gene and showed 554 metabolism related candidate hub genes. Nonetheless, a single WGCNA analysis had significant limitations and inaccuracies (Tzimas et al., 2019). Currently, studies applied WGCNA were normally combined with multiple machine learning algorithms to identify biomarkers for disease prognosis and diagnosis. Zhao et al. (2022) identified four core genes (BTN3A2, CYFIP2, ST8SIA1, and TYMS) as biomarker for diagnosis of rheumatoid arthritis *via* comprehensive analysis of WGCNA, LASSO, random forest, and support vector machine analysis. By WGCNA, LASSO, and random forest algorithms,

Fan et al. (2022) obtained five signature genes (UPP1, S100A9, KIF1B, S100A12, SLC26A8) and emerged remarkable diagnostic performance in pediatric septic shock. In the current study, LASSO regression analysis and random forest algorithms found two signature genes, then three validation datasets, including GSE140275, GSE122709, and GSE180470, confirmed that SDHD and FERMT3 were highly expressed in the stroke group.

SDHD, one subunit of succinate dehydrogenase (SDH), dual roles in respiration by transferring electrons from succinate to ubiquinone in the mitochondrial electron transport chain (ETC) and catalyzing oxidation of succinate to fumarate in the mitochondrial Krebs cycle (Cecchini, 2003; Sun et al., 2005). Researchers reported mutations of SDHD in patients with hereditary pheochromocytomas and hereditary paragangliomas (Baysal et al., 2000). *In vitro* experiment performed by Bandara et al. (2021) demonstrated that mutation of SDHD *via* CRISPR/Cas9 approach could suppress glycolysis and overall ATP synthesis in HEK293. Overexpression of SDHD could significantly suppressed cell proliferation *in vitro* and tumor growth of HCC cells *in vivo* (Yuan et al., 2022). FERMT3 is a member of the kindlin family of binding proteins containing the FERM domain (Rognoni et al., 2016). FERMT3 mediates integrin activation and integrin-ligand binding. Therefore, FERMT3 is closely related to various biological activities, including cell adhesion, spreading, cell survival, proliferation and differentiation (Rognoni et al., 2016). Mutations of FERMT3 gene could cause leukocyte adhesion deficiency type III (LAD III) (Kuijpers et al., 2009). Liu et al. (2021) performed RNA sequencing in patients with triple-negative breast cancer and identified FERMT3 as protective gene in compound kushen injection treatment. Nonetheless, correlations of FERMT3 and SDHD with stroke have not been previously reported.

Post-stroke depression (PSD), the most common psychiatric problem after stroke, is an independent risk factor of stroke mortality (D'Anci et al., 2019). PSD is closely associated with worse outcomes of physical and cognitive recovery, functioning, and health related quality of life (Villa et al., 2018). It is worth noting that PSD might halt or impede rehabilitation treatments. However, the complex pathophysiology of PSD is still only partly known till now. The current evidence indicates genetic factors as major aetiopathological predictors for PSD. Yang et al. (2010) reported that IL-18 level in serum on day 7 after admission might predict the risk of PSD. Plasma levels of glutamate and glutamate oxaloacetate transaminase at admission were also reported to be closely related PSD within 3 months (Cheng et al., 2014). To further probe the role of hub genes in stroke, we performed a GSEA analysis of signature genes. The results demonstrated that SDHD and FERMT3 were significantly enriched in depression. Then we validated our findings in stroke patients with and without depression. We found increase expression levels of SDHD and FERMT3 in stroke patients with depression, compatible with our previous research inferences. In addition, based on the two signature genes (SDHD and FERMT3) that we identified, we successfully established a PSD diagnosis for evaluating diagnosis value of SDHD and FERMT3 in our PSD cohort. Nomogram model showed great predictive ability and clinical usefulness. Meanwhile, AUC values of SDHD and FERMT3 were 0.896 and 0.964. Our results suggested that SDHD and FERMT3 might play essential roles in diagnosis of PSD. Finally, we performed correlation analysis of two signature genes

and several clinical traits. We found that the SDHD and FERM3 were positively correlated with depression, which suggested that SDHD and FERMT3 had certain therapeutic predictive value in PSD. Moreover, SDHD was also found a negative correlation with activities of daily living in this study. Considering the feature of this parameters, it suggested that these two signature genes may also serve as biomarkers to monitor the mental functional prognosis in patients with PSD (van Hulsteijn et al., 2013).

The present study also has certain shortcomings. Firstly, we collected data from public databases with small samples. There could have been a selection bias. Large datasets of stroke patients are limited, so we tried to minimize the bias of our results by validating signature genes across multiple datasets. Secondly, the metabolism related-pathways and -hub genes in stroke lack literature support and required further confirmation. Thirdly, although two metabolism related signature genes were identified as potential predictors for PSD, larger patient cohorts should be examined in the future to validate the correlation between two signature genes (SDHD and FERMT3) with PSD. Then further *in vivo* or *in vitro* studies should be carried out to validate diagnostic value and potential therapeutic value.

Conclusion

In conclusion, we identified two signature genes (SDHD and FERMT3) in peripheral blood of stroke patients by machine learning. SDHD and FERMT3 were found to be significantly associated with depression, and were identified as diagnostic and therapeutic signatures by our stroke cohorts with and without PSD, which could be a valuable reference for future clinical practice.

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 authors.

Ethics statement

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

Author contributions

YD, YZhe, BY, and RC designed the current study. XZ, XW, and ShuweiW collected the clinical information. YZha and SongW completed data downloading and processing. XZ, ZW, and YZha performed bioinformatics analysis. QY performed ELISA testing. YZhe, XZ, and YZha drafted the manuscript. YD supervised and modified the drafting process. All authors contributed to the article and approved the submitted version.

Funding

This study was funded by the National Natural Science Foundation of China (82000529 and 82272595), the Natural Science Foundation of Jiangsu Province (SBK2020042595), and the National Key Research & Development Program of the Ministry of Science and Technology of the People's Republic of China (2018YFC2002300 and 2018YFC2002301).

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.2023.1146620/full#supplementary-material

SUPPLEMENTARY TABLE 1
Differentially expressed genes in GSE140275.

SUPPLEMENTARY TABLE 2
Differentially expressed genes in GSE122709.

SUPPLEMENTARY TABLE 3
Correlation analysis results of modules and clinical features.

SUPPLEMENTARY TABLE 4
Signature genes selected by least absolute shrinkage and selection operator (LASSO) analysis.

SUPPLEMENTARY TABLE 5
Signature genes determined by random forest analysis.

References

Bandara, A., Drake, J., and Brown, D. (2021). Complex II subunit SDHD is critical for cell growth and metabolism, which can be partially restored with a synthetic ubiquinone analog. *BMC Mol. Cell Biol.* 22:35. doi: 10.1186/s12860-021-00370-w

Baysal, B., Ferrell, R., Willett-Brozick, J., Lawrence, E., Myssiorek, D., Bosch, A., et al. (2000). Mutations in SDHD, a mitochondrial complex II gene, in hereditary paraganglioma. *Science* 287, 848–851.

Cai, W., Mueller, C., Li, Y., Shen, W., and Stewart, R. (2019). Post stroke depression and risk of stroke recurrence and mortality: A systematic review and meta-analysis. *Ageing Res. Rev.* 50, 102–109.

Cecchini, G. (2003). Function and structure of complex II of the respiratory chain. *Annu. Rev. Biochem.* 72, 77–109.

Chen, R., Wu, P., Cai, Z., Fang, Y., Zhou, H., Lasanajak, Y., et al. (2019a). Puerariae lobatae radix with Chuanxiong rhizoma for treatment of cerebral ischemic stroke by remodeling gut microbiota to regulate the brain-gut barriers. *J. Nutr. Biochem.* 65, 101–114. doi: 10.1016/j.jnutbio.2018.12.004

Chen, R., Xu, Y., Wu, P., Zhou, H., Lasanajak, Y., Fang, Y., et al. (2019b). Transplantation of fecal microbiota rich in short chain fatty acids and butyric acid treat cerebral ischemic stroke by regulating gut microbiota. *Pharmacol. Res.* 148:104403. doi: 10.1016/j.phrs.2019.104403

Chen, S., Sun, M., Zhao, X., Yang, Z., Liu, W., Cao, J., et al. (2019c). Neuroprotection of hydroxysafflor yellow A in experimental cerebral ischemia/reperfusion injury via metabolic inhibition of phenylalanine and mitochondrial biogenesis. *Mol. Med. Rep.* 19, 3009–3020. doi: 10.3892/mmr.2019.9959

Cheng, S., Zhao, Y., Li, J., Chen, X., Wang, R., and Zeng, J. (2014). Plasma levels of glutamate during stroke is associated with development of post-stroke depression. *Psychoneuroendocrinology* 47, 126–135. doi: 10.1016/j.psyneuen.2014.05.006

D'Anci, K., Uhl, S., Oristaglio, J., Sullivan, N., and Tsou, A. (2019). Treatments for post-stroke motor deficits and mood disorders: A systematic review for the 2019 U.S. department of Veterans affairs and U.S. department of defense guidelines for stroke rehabilitation. *Ann. Intern. Med.* 171, 906–915. doi: 10.7326/M19-2414

Dong, L., Brown, D., Chervin, R., Case, E., Morgenstern, L., and Lisabeth, L. (2021). Pre-stroke sleep duration and post-stroke depression. *Sleep Med.* 77, 325–329. doi: 10.1016/j.sleep.2020.04.025

Fan, J., Shi, S., Qiu, Y., Liu, M., and Shu, Q. (2022). Analysis of signature genes and association with immune cells infiltration in pediatric septic shock. *Front. Immunol.* 13:1056750. doi: 10.3389/fimmu.2022.1056750

Fan, X., Chen, H., Xu, C., Wang, Y., Yin, P., Li, M., et al. (2022). S1PR3, as a core protein related to ischemic stroke, is involved in the regulation of blood-brain barrier damage. *Front. Pharmacol.* 13:834948. doi: 10.3389/fphar.2022.834948

Fu, X., Wang, J., Liao, S., Lv, Y., Xu, D., Yang, M., et al. (2019). (1)H NMR-based metabolomics reveals refined-Huang-Lian-Jie-Du-decoction (BBG) as a potential ischemic stroke treatment drug with efficacy and a favorable therapeutic window. *Front. Pharmacol.* 10:337. doi: 10.3389/fphar.2019.00337

GBD 2019 Stroke Collaborators (2021). Global, regional, and national burden of stroke and its risk factors, 1990-2019: A systematic analysis for the Global Burden of Disease Study 2019. *Lancet Neurol.* 20, 795–820.

He, D., Li, Q., Du, G., Meng, G., Sun, J., and Chen, S. (2021). An integration of network pharmacology and experimental verification to investigate the mechanism of guizhi to treat nephrotic syndrome. *Front. Pharmacol.* 12:755421. doi: 10.3389/fphar.2021.755421

He, L., Wang, J., Zhang, L., Wang, F., Dong, W., and Yang, H. (2020). Admission heart rate variability is associated with poststroke depression in patients with acute mild-moderate ischemic stroke. *Front. Psychiatry* 11:696. doi: 10.3389/fpsyt.2020.00696

Jiang, W., Gong, L., Liu, F., Ren, Y., and Mu, J. (2021). Alteration of gut microbiome and correlated lipid metabolism in post-stroke depression. *Front. Cell. Infect. Microbiol.* 11:663967. doi: 10.3389/fcimb.2021.663967

Kaida, Y., Fukami, K., Matsui, T., Higashimoto, Y., Nishino, Y., Obara, N., et al. (2013). DNA aptamer raised against AGEs blocks the progression of experimental diabetic nephropathy. *Diabetes* 62, 3241–3250.

Kamal, A., Khoja, A., Usmani, B., Magsi, S., Malani, A., Peera, Z., et al. (2020). Effect of 5-minute movies shown via a mobile phone app on risk factors and mortality after stroke in a low- to middle-income country: Randomized controlled trial for the stroke caregiver dyad education intervention (Movies4Stroke). *JMIR mHealth uHealth* 8:e12113. doi: 10.2196/12113

Kang, J., Choi, Y., Kim, I., Lee, H., Kim, H., Baik, S., et al. (2021). LASSO-based machine learning algorithm for prediction of lymph node metastasis in T1 colorectal cancer. *Cancer Res. Treat.* 53, 773–783. doi: 10.4143/crt.2020.974

Klinedinst, N., Dunbar, S., and Clark, P. (2012). Stroke survivor and informal caregiver perceptions of poststroke depressive symptoms. *J. Neurosci. Nurs.* 44, 72–81.

Kuijpers, T., van de Vijver, E., Weterman, M., de Boer, M., Tool, A., van den Berg, T., et al. (2009). LAD-1/variant syndrome is caused by mutations in FERMT3. *Blood* 113, 4740–4746. doi: 10.1182/blood-2008-10-182154

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

Li, G., Miao, J., Sun, W., Song, X., Lan, Y., Zhao, X., et al. (2020). Lower serum uric acid is associated with post-stroke depression at discharge. *Front. Psychiatry* 11:52. doi: 10.3389/fpsyt.2020.00052

Li, Y., Cao, L., Liu, L., and Qi, Q. (2017). Serum levels of homocysteine at admission are associated with post-stroke depression in acute ischemic stroke. *Neurol. Sci.* 38, 811–817. doi: 10.1007/s10072-017-2848-2

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.* 18:296.

Lin, S., Luan, X., He, W., Ruan, Y., Yuan, C., Fan, A., et al. (2020). Post-stroke depression and estimated glomerular filtration rate: A prospective stroke cohort. *Neuropsychiatr. Dis. Treat.* 16, 201–208. doi: 10.2147/NDT.S225905

Lin, W., Wang, Y., Chen, Y., Wang, Q., Gu, Z., and Zhu, Y. (2021). Role of calcium signaling pathway-related gene regulatory networks in ischemic stroke based on multiple WGCNA and single-cell analysis. *Oxid. Med. Cell Longev.* 2021:8060477. doi: 10.1155/2021/8060477

Liu, C., Zhang, X., Chai, H., Xu, S., Liu, Q., Luo, Y., et al. (2022). Identification of immune cells and key genes associated with Alzheimer's disease. *Int. J. Med. Sci.* 19, 112–125.

Liu, K., Chen, S., and Lu, R. (2021). Identification of important genes related to ferroptosis and hypoxia in acute myocardial infarction based on WGCNA. *Bioengineered* 12, 7950–7963. doi: 10.1080/21655979.2021.1984004

Liu, M., Li, G., Tang, J., Liao, Y., Li, L., Zheng, Y., et al. (2018). The influence of sex in stroke thrombolysis: A systematic review and meta-analysis. *J. Clin. Neurol.* 14, 141–152.

Liu, X., Wu, Y., Zhang, Y., Bu, D., Wu, C., Lu, S., et al. (2021). High throughput transcriptome data analysis and computational verification reveal immunotherapy biomarkers of compound Kushen injection for treating triple-negative breast cancer. *Front. Oncol.* 11:747300. doi: 10.3389/fonc.2021.747300

Luft, A., McCombe-Waller, S., Whitall, J., Forrester, L., Macko, R., Sorkin, J., et al. (2004). Repetitive bilateral arm training and motor cortex activation in chronic stroke: A randomized controlled trial. *JAMA* 292, 1853–1861.

Mantero, A., and Ishwaran, H. (2021). Unsupervised random forests. *Stat. Anal. Data Min.* 14, 144–167.

Qiao, J., Sui, R., Zhang, L., and Wang, J. (2020). Construction of a risk model associated with prognosis of post-stroke depression based on magnetic resonance spectroscopy. *Neuropsychiatr. Dis. Treat.* 16, 1171–1180. doi: 10.2147/NDT.S245129

Ritchie, M., Phipson, B., Wu, D., Hu, Y., Law, C., 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

Robin, X., Turck, N., Hainard, A., Tiberti, N., Lisacek, F., Sanchez, J., et al. (2011). pROC: An open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics* 12:77. doi: 10.1186/1471-2105-12-77

Robinson, M., McCarthy, D., and Smyth, G. (2010). edgeR: A bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26, 139–140.

Rognoni, E., Ruppert, R., and Fässler, R. (2016). The kindlin family: Functions, signaling properties and implications for human disease. *J. Cell Sci.* 129, 17–27. doi: 10.1242/jcs.161190

Shi, Y., Guo, L., Chen, Y., Xie, Q., Yan, Z., Liu, Y., et al. (2021). Risk factors for ischemic stroke: Differences between cerebral small vessel and large artery atherosclerosis aetiologies. *Folia Neuropathol.* 59, 378–385.

Subramanian, A., Tamayo, P., Mootha, V., Mukherjee, S., Ebert, B., Gillette, M., 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.

Sun, F., Huo, X., Zhai, Y., Wang, A., Xu, J., Su, D., et al. (2005). Crystal structure of mitochondrial respiratory membrane protein complex II. *Cell* 121, 1043–1057.

Sun, P., Ma, F., Xu, Y., Zhou, C., Stetler, R., and Yin, K. (2021). Genetic deletion of endothelial microRNA-15a/16-1 promotes cerebral angiogenesis and neurological recovery in ischemic stroke through Src signaling pathway. *J. Cereb. Blood Flow Metab.* 41, 2725–2742. doi: 10.1177/0271678X211010351

Tzimas, C., Rau, C., Buergisser, P., Jean-Louis, G. Jr., Lee, K., Chukwuneke, J., et al. (2019). WIPI1 is a conserved mediator of right ventricular failure. *JCI Insight* 5:e122929. doi: 10.1172/jci.insight.122929

Upreti, H., Kasmani, J., Dane, K., Braunstein, E., Streiff, M., Shanbhag, S., et al. (2019). Reduced ADAMTS13 activity during TTP remission is associated with stroke in TTP survivors. *Blood* 134, 1037–1045. doi: 10.1182/blood.2019001056

van Hulsteijn, L., Louisse, A., Havekes, B., Kaptein, A., Jansen, J., Hes, F., et al. (2013). Quality of life is decreased in patients with paragangliomas. *Eur. J. Endocrinol.* 168, 689–697.

Villa, R., Ferrari, F., and Moretti, A. (2018). Post-stroke depression: Mechanisms and pharmacological treatment. *Pharmacol. Ther.* 184, 131–144.

Wang, D., Wang, Q., Chen, R., Yang, S., Li, Z., and Feng, Y. (2019). Exploring the effects of Gastrodia elata Blume on the treatment of cerebral ischemia-reperfusion injury using UPLC-Q/TOF-MS-based plasma metabolomics. *Food Funct.* 10, 7204–7215. doi: 10.1039/c9fo01729a

Wang, J., Wang, Y., Zhong, Y., Li, X., Du, S., Xie, P., et al. (2019). Identification of co-expression modules and pathways correlated with osteosarcoma and its metastasis. *World J. Surg. Oncol.* 17:46.

Wang, M., Wang, L., Pu, L., Li, K., Feng, T., Zheng, P., et al. (2020). LncRNAs related key pathways and genes in ischemic stroke by weighted gene co-expression network analysis (WGCNA). *Genomics* 112, 2302–2308. doi: 10.1016/j.ygeno.2020.01.001

Wang, X., Li, Y., Li, M., Lu, J., Zhao, J., Sun, X., et al. (2012). Glutamate level detection by magnetic resonance spectroscopy in patients with post-stroke depression. *Eur. Arch. Psychiatry Clin. Neurosci.* 262, 33–38. doi: 10.1007/s00406-011-0209-3

Xu, D., Wang, Y., Liu, X., Zhou, K., Wu, J., Chen, J., et al. (2021). Development and clinical validation of a novel 9-gene prognostic model based on multi-omics in pancreatic adenocarcinoma. *Pharmacol. Res.* 164:105370. doi: 10.1016/j.phrs.2020.105370

Yamashiro, K., Tanaka, R., Urabe, T., Ueno, Y., Yamashiro, Y., Nomoto, K., et al. (2017). Gut dysbiosis is associated with metabolism and systemic inflammation in patients with ischemic stroke. *PLoS One* 12:e0171521. doi: 10.1371/journal.pone.0171521

Yang, L., Zhang, Z., Sun, D., Xu, Z., Zhang, X., and Li, L. (2010). The serum interleukin-18 is a potential marker for development of post-stroke depression. *Neurol. Res.* 32, 340–346. doi: 10.1179/016164110X12656393665080

Yu, B., Zhang, X., Cheng, Y., Liu, L., YanJiang, Wang, J., et al. (2022). The effects of the biceps brachii and brachioradialis on elbow flexor muscle strength and spasticity in stroke patients. *Neural Plast.* 2022:1295908. doi: 10.1155/2022/1295908

Yuan, T., Zhou, T., Qian, M., Du, J., Liu, Y., Wang, J., et al. (2022). SDHA/B reduction promotes hepatocellular carcinoma by facilitating the deNEDDylation of cullin1 and stabilizing YAP/TAZ. *Hepatology* doi: 10.1002/hep.32621

Zhang, Y., Zhang, P., Shen, X., Tian, S., Wu, Y., Zhu, Y., et al. (2013). Early exercise protects the blood-brain barrier from ischemic brain injury via the regulation of MMP-9 and occludin in rats. *Int. J. Mol. Sci.* 14, 11096–11112. doi: 10.3390/ijms140611096

Zhao, Z., He, S., Yu, X., Lai, X., Tang, S., Mariya, M., et al. (2022). Analysis and experimental validation of rheumatoid arthritis innate immunity gene CYFIP2 and pan-cancer. *Front. Immunol.* 13:954848. doi: 10.3389/fimmu.2022.954848

Zheng, P., Chen, L., Liu, P., Pan, H., Fan, W., and Liu, Z. (2022). Identification of immune-related key genes in the peripheral blood of ischaemic stroke patients using a weighted gene coexpression network analysis and machine learning. *J. Transl. Med.* 20:361. doi: 10.1186/s12967-022-03562-w

Zhou, Y., Wu, W., Lindholt, J., Sukhova, G., Libby, P., Yu, X., et al. (2015). Regulatory T cells in human and angiotensin II-induced mouse abdominal aortic aneurysms. *Cardiovasc. Res.* 107, 98–107.

Zhou, Z., Liang, Y., Qu, H., Zhao, M., Guo, F., Zhao, C., et al. (2018). Plasma homocysteine concentrations and risk of intracerebral hemorrhage: A systematic review and meta-analysis. *Sci. Rep.* 8:2568. doi: 10.1038/s41598-018-21019-3

Zhu, M., Zeng, Q., Fan, T., Lei, Y., Wang, F., Zheng, S., et al. (2022). Clinical significance and immunometabolism landscapes of a novel recurrence-associated lipid metabolism signature in early-stage lung adenocarcinoma: A comprehensive analysis. *Front. Immunol.* 13:783495. doi: 10.3389/fimmu.2022.783495