



OPEN

Identification of hypoxia-related genes and exploration of their relationship with immune cells in ischemic stroke

Kai Yang¹, Zhaoqi Zhang², Xiaoju Liu², Tong Wang², Zhicheng Jia², Xin Li^{3,5}✉ & Wei Liu^{2,4,5}✉

Ischemic stroke (IS) is a major threat to human health, and it is the second leading cause of long-term disability and death in the world. Impaired cerebral perfusion leads to acute hypoxia and glucose deficiency, which in turn induces a stroke cascade response that ultimately leads to cell death. Screening and identifying hypoxia-related genes (HRGs) and therapeutic targets is important for neuroprotection before and during brain recanalization to protect against injury and extend the time window to further improve functional outcomes before pharmacological and mechanical thrombolysis. First, we downloaded the GSE16561 and GSE58294 datasets from the NCBI GEO database. Bioinformatics analysis of the GSE16561 dataset using the limma package identified differentially expressed genes (DEGs) in ischemic stroke using adj. p. values < 0.05 and a fold change of 0.5 as thresholds. The Molecular Signature database and Genecards database were pooled to obtain hypoxia-related genes. 19 HRGs associated with ischemic stroke were obtained after taking the intersection. LASSO regression and multivariate logistic regression were applied to identify critical biomarkers with independent diagnostic values. ROC curves were constructed to validate their diagnostic efficacy. We used CIBERSORT to analyze the differences in the immune microenvironment between IS patients and controls. Finally, we investigated the correlation between HRGs and infiltrating immune cells to understand molecular immune mechanisms better. Our study analyzed the role of HRGs in ischemic stroke. Nineteen hypoxia-related genes were obtained. Enrichment analysis showed that 19 HRGs were involved in response to hypoxia, HIF-1 signaling pathway, autophagy, autophagy of mitochondrion, and AMPK signaling pathway. Because of the good diagnostic properties of SLC2A3, we further investigated the function of SLC2A3 and found that it is closely related to immunity. We have also explored the relevance of other critical genes to immune cells. Our findings suggest that hypoxia-related genes play a crucial role in the diversity and complexity of the IS immune microenvironment. Exploring the association between hypoxia-related critical genes and immune cells provides innovative insights into the therapeutic targets for ischemic stroke.

Abbreviations

IS	Ischemic stroke
HRG	Hypoxia-related gene
NK	Natural killer
BBB	Blood–brain barrier
GEO	Gene Expression Omnibus
MSigDB	Molecular signature database
DEGs	Differentially expressed genes
LASSO	Least absolute shrinkage and selection operator
ROC	Receiver operating characteristic

¹Acupuncture and Moxibustion and Massage College, Shandong University of Traditional Chinese Medicine, Jinan, China. ²Shandong University of Traditional Chinese Medicine, Jinan, China. ³Department of Neurology, The Affiliated Hospital of Shandong University of Traditional Chinese Medicine, Jinan, China. ⁴Department of Cerebral Disease, The Second Affiliated Hospital of Shandong University of Traditional Chinese Medicine, Jinan, China. ⁵These authors contributed equally: Xin Li and Wei Liu. ✉email: lixin639396@163.com; weiliu800109@163.com

BP	Biological process
CC	Cellular component
MF	Molecular function
KEGG	Kyoto Encyclopedia of Genes and Genomes
ROS	Reactive oxygen species
D.C.s	Dendritic cells
MACO	Middle cerebral artery occlusion
Tregs	Regulatory T cells
TIA	Transient ischemic attack

Stroke is a neurological condition that can be fatal and primarily affects older people. Stroke is the second leading cause of disability and death globally, with low and middle-income countries bearing the most tremendous burden of disease¹. Stroke can be broadly classified into hemorrhagic stroke and ischemic stroke, and the most common type of stroke in clinical practice is ischemic stroke. Arterial occlusion-related ischemic stroke is the leading cause of most strokes, accounting for 87% of stroke cases and nearly half of all deaths². According to the latest statistics, the global ischemic stroke population grew from 2.04 million to 3.29 million between 1990 and 2019 and is predicted to reach 4.9 million by 2030³.

When the arterial blockage occurs, the insufficient supply of oxygen and glucose can produce a series of harmful events⁴, including inadequate ATP supply, lactic acidosis, extracellular excitotoxicity, mitochondrial degradation, neuroinflammation, disruption of the blood–brain barrier (BBB), irreversible neuronal cell death. Impaired cerebral perfusion leads to acute hypoxia and glucose deprivation, which directly results in reduced adenosine triphosphate (ATP) synthesis, leading to lactic acidosis and disturbed cellular homeostasis resulting in cellular damage⁵. In addition, ATP deficiency leads to the failure of ATP-dependent ion transport pumps⁶, which induces extracellular excitability and cytotoxic edema⁶. The activation of associated proteases and lipases leads to free radical release and mitochondrial degradation. Hypoxia/reperfusion can also cause neuronal damage, while damaged neurons can activate resident microglia to move to the infarct core and penumbra⁵, with different subtypes of microglia exerting protective and damaging effects⁷. In addition to this, hypoxia enhances the expression of transcription factors, which in turn promote the synthesis of inflammatory proteins, amplify neuroinflammation, and promote infiltration of peripheral immune cells⁵. Invading neutrophils secrete matrix metalloproteinases, leading to disruption of tight junctions and disruption of the BBB, which in turn exacerbates brain edema, inducing cell death and neuronal loss^{5,8}.

Studies have shown that mortality and recurrence rates of ischemic stroke increase with increasing duration of hypoxia^{9,10}. Early embolization and pharmacological thrombolysis shorten reperfusion time to improve hypoxia, possibly reducing complications, disability, and mortality¹¹. Intravenous systemic thrombolysis with fibrinolytic drugs, mainly alteplase (rt-PA), is the standard of care in acute ischemic stroke. Rt-PA converts fibrinogen into the protein hydrolase fibrinolytic enzyme that cleaves fibrinogen and fibrin. Fibrinolysis leads to the breakdown of blood clots, allowing blood flow to return to the brain¹². However, the numerous contraindications and narrow time window for intravenous thrombolysis prevent many patients from receiving thrombolytic therapy^{13,14}.

Hypoxia plays a vital role in ischemic stroke, which runs through the occurrence and development of ischemic stroke. Based on the fact that a large proportion of the population cannot receive thrombolytic therapy in the prescribed time window. Finding targets related to hypoxia is essential to protect brain cells from ischemia and reperfusion injury and expand the treatment time window.

This study first identified differentially expressed hypoxia-related genes (HRGs) between IS and control samples from metadata cohorts and hypoxia-related datasets. It is helpful to clarify the signal pathway mechanism related to hypoxia and find effective targets for treating IS. Secondly, after integrated analysis among least absolute shrinkage and selection operator (LASSO) regression and multivariate logistic regression, we included critical genes in the receiver operating characteristic (ROC) curve and compared their diagnostic power for IS. Importantly, we explored the relationship between hypoxia-related genes and immune cells, providing a new perspective for developing immunomodulatory therapeutic options for IS.

Methods

Dataset preparation and data processing. Transcriptome profiles of IS patients containing the messenger RNA (mRNA) expression profiles and clinical information were obtained from Gene Expression Omnibus (GEO) database (<https://www.ncbi.nlm.nih.gov/geo>) on 11 August 2022. We included whole or peripheral blood gene expression profiles from patients with IS or control samples. This study used R to download ischemic stroke-related mRNA expression profile data and clinical data from GEO database.

The GSE16561¹⁵ dataset (GPL6883, Illumina HumanRef-8 v3.0 expression bead chip, array, Homo sapiens) contained 39 control subjects and 24 patients with ischemic stroke, serving as the training set. The GSE58294¹⁶ dataset (GPL570, Affymetrix Human Genome U133 Plus 2.0 Array, Homo sapiens) contains a total of 23 control samples and 69 ischemic stroke samples blood samples from less than 3 h, 5 h, and 24 h after the onset of ischemic stroke, serving as a validation cohort. More details about these two datasets are shown in Table 1.

In addition, 259 hypoxia-related genes were obtained according to the Molecular Signature database¹⁷ and Genecards (<https://www.genecards.org/>). There were 200 hypoxia-related genes in the Molecular Signature database. In the Genecards database, we searched for hypoxia-related genes using "hypoxia" as the search term and obtained 68 genes with an association score > 7 as the cut-off value. After removing overlapping genes, 259 HRGs were obtained (Supplementary Table 1).

Dataset	Platform	Control	IS	Author	Country	Submission	Sample	Application
GSE16561	GPL6883	24	39	Barr TL	USA	2010	Peripheral whole blood	Identification
GSE58294	GPL570	23	69	Stamova B	USA	2014	Peripheral whole blood	Validation

Table 1. Detailed information of the studied gene expression profiles.

Differential expression analysis. Differentially expressed genes (DEGs) between ischemic stroke and control samples in the GSE16561 were identified by the package of limma in R language (adj. p. values < 0.05 and $|\log_2FC| \geq 0.5$)^{18,19}. DEGs overlapped with hypoxia-related genes were regarded as differentially expressed hypoxia-related genes (HRGs). Meanwhile, the volcano map and cluster heatmap were represented by the package "pheatmap (v1.0.12)" and "ggplot2(v3.3.6)" in R to visualize the differences.

Functional enrichment analysis. Biological functions were analyzed using the clusterProfiler package (v4.4.4), which includes GO and KEGG. The above HRGs were comprehensively investigated using the MF, CC, BP, and KEGG, which refer to molecular function, biological process, cellular component, and the Kyoto Encyclopedia of Genes and Genomics pathway (v97.0)^{20,21}. We used the adjust the error discovery rate FDR to represent significant results. $P < 0.05$ as the screening criterion.

Identification of diagnostic markers and construction of the diagnosis model. Minimum absolute shrinkage selection operator (LASSO) logistic regression was used to reduce the number of genes in the model and solve the multicollinearity problem in regression analysis. The LASSO algorithm was applied using the "glmnet" package²² to determine feature selection and screening of IS diagnostic markers. We then used multivariate logistic regression and inverse methods to identify independent diagnostic biomarkers.

Diagnostic model internal and external validation. The "pROC" (v1.18.0) package²³ was used to access the diagnostic performance of selected HRG biomarkers in the training dataset (GSE16561). We then validated the diagnostic performance of crucial biomarkers that differentiate IS patients from controls in an external validation cohort (GSE58294). To assess the reliability and generalizability of the genetic characteristics, we selected the GSE58294 dataset containing samples from different time points within 24 h after the onset of ischemic stroke as the validation set. We evaluated the diagnostic model using the same parameters.

Immune infiltration analyses. At the same time, to compare the relative percentages of immune cells in each sample, CIBERSORT was used based on the mRNA expression²⁴. Herein, we performed "e1071 (v1.7-11)", "parallel" (v4.2.0), "preprocess Core (v1.48.0)" and "CIBERSORT" packages in R to evaluate. The bar graph provides an overview of each individual's relative percentage of the 22 immune cell subsets. The associations of all cell subsets in the form of correlated heat maps were shown by the corrplot (v0.90) package in R. And the differences in infiltration between IS and control samples were reflected in the form of violin maps using the ggplot2(v3.3.0) package. The critical value was $P < 0.05$.

Correlation analysis between key HRGs and immune cells. Spearman correlation analysis was then performed for key diagnostic biomarkers and immune cell subsets in IS to estimate their relationship using the "ggstatsplot(v0.9.3)" package in R. Meanwhile, the results of the correlational analysis was visualized by the "tidyverse (v1.3.2)", "ggsci(v2.9)" and "ggplot2 (v3.3.6)" packages in R.

Statistical analyses. Statistical analysis and drawing work were performed using R software (version 4.2.0). The ROC analysis was then visualized using the R's "pROC" package (v1.18.0). Continuous variables were represented by mean \pm SD, Student's t-test represented normal distribution variables, abnormal distribution variables were represented by Mann–Whitney U test. The cut-off thresholds of adj.p.value < 0.05 and $|\log_2FC| \geq 0.5$ were established as the cut-off thresholds in the differential expression analysis. Of all the studies, we considered $P < 0.05$ as a meaningful difference.

Result

Identification of 19 differentially expressed HRGs. The flow chart in Fig. 1 shows an overview of this study. The data for analysis was obtained from the online GEO database. The data of 39 whole IS blood samples and 24 control whole IS blood samples in the GSE16561 dataset were obtained, including gene expression profiles and clinical information, according to the following standards identified 555 DEGs: $|\log_2FC| \geq 0.5$ or higher, adj. p. value < 0.05 . Four hundred twenty-seven up-regulated genes (76.9%) and 128 down-regulated genes (23.1%) among DEGs were illustrated in the volcano plot (Fig. 2A,B). By intersecting 259 HRGs, Venn diagrams revealed 19 differentially expressed HRGs (HIF1A, HIGD2A, CREB1, FOS, ADM, BNIP3L, DUSP1, GAA, GLRX, HK2, IRS2, MYH9, NDST1, NFIL3, RRAGD, SAP30, SIAH2, SLC2A3, and WSB1) (Fig. 2C). Figure 2D showed the expression of these 19 genes in GSE16561. It can be seen from the data in the picture that, except for HIGD2A, other genes were both up-regulated in the ischemic stroke sample. Additional details about them are provided in Table 2.

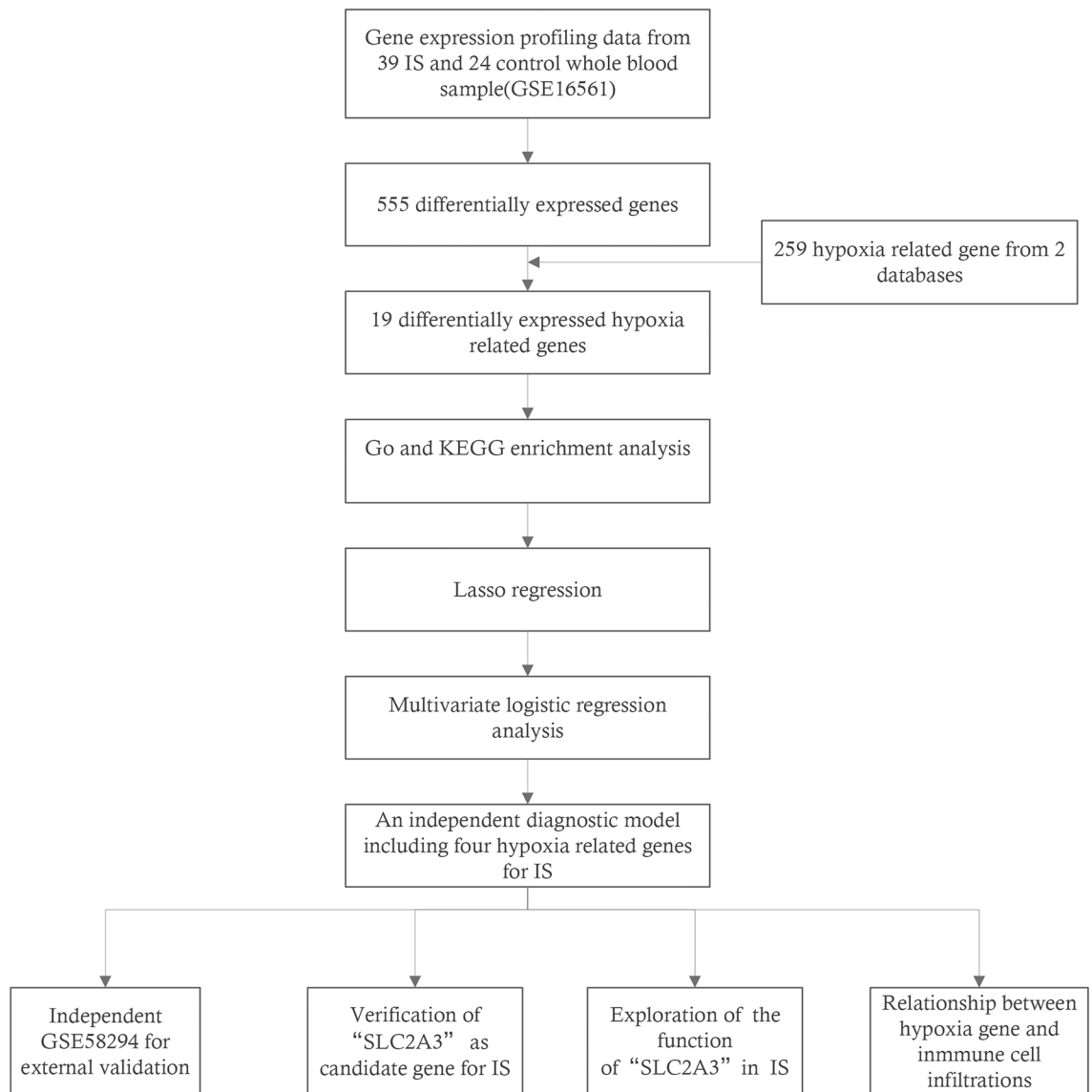


Figure 1. The flowchart of the overall study procedures.

Functional enrichment analysis of 19 HRGs. KEGG enrichment revealed the 19 HRGs predominantly participated in autophagy, HIF-1 signaling pathway, AMPK signaling pathway, TNF signaling pathway, growth hormone synthesis, secretion and action, Th17 cell differentiation, TNF signaling pathway, Relaxin signaling pathway (Fig. 3A,B). BP showed that the selected genes were principally enriched in the regulation of autophagy of mitochondrion, response to hypoxia, response to decreased oxygen levels, response to oxygen levels, and positive regulation of myeloid cell differentiation (Fig. 3C). In CC, HRGs were mainly associated with the RNA polymerase II transcription factor complex, nuclear transcription factor complex, transcription factor complex, and tertiary granule membrane (Fig. 3D). MF showed that 19 HRGs were mainly related to transcription corepressor activity, monosaccharide binding, glucose binding, monosaccharide binding, and RNA polymerase II-specific DNA-binding transcription factor binding (Fig. 3E). The other detailed information is listed in Supplemental Table 2.

Construction of the diagnostic risk model. From the graph in Fig. 4A, we can see the correlation matrix analysis of 19 genes, and partly hypoxia genes had a tough correlation with each other. In order to deal with the multicollinearity problem in regression analysis, LASSO analysis was applied to further zoom out the scope of HRGs as possible diagnostic biomarkers for IS. 19 genes were included in the LASSO model. Primarily, 9 genes (CREB1, FOS, GAA, GLRX, HIF1A, HIGD2A, MYH9, SAP30, SLC2A3) were identified based on λ_{1se} (Fig. 4B,C). Then, multivariate logistic regression analysis was used to determine the independent candidate diagnostic biomarkers. Four genes (SLC2A3, MYH9, CREB, and SAP30) were obtained (Fig. 4D). Besides, the four biomarkers demonstrated advantageous diagnostic ability and good diagnostic value in discriminating IS

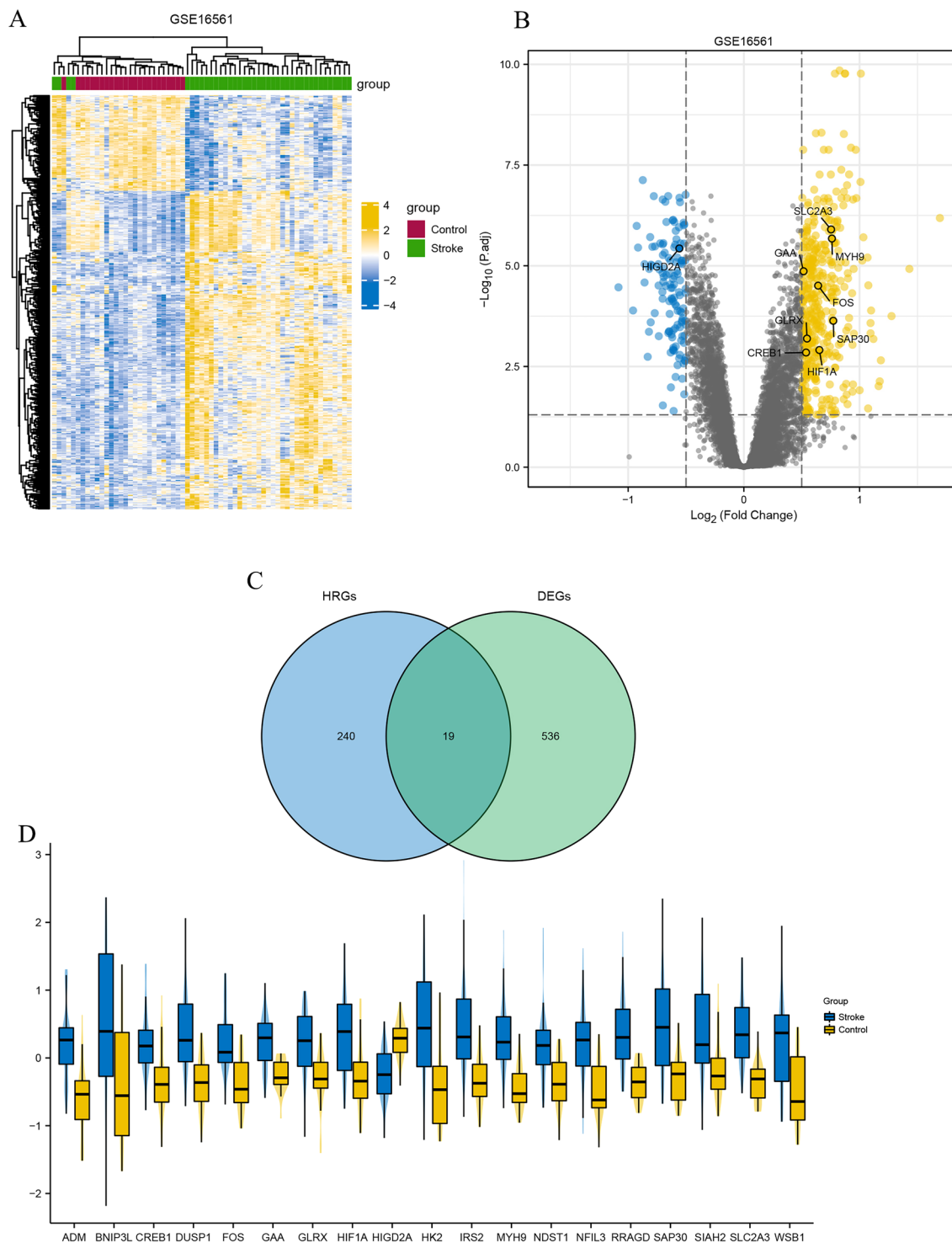


Figure 2. Identification of DEGs and hypoxia-related DEGs. **(A)** The heatmap of DEGs in GSE16561. **(B)** Volcano map of GSE16561. **(C)** Venn diagrams show the overlap between the significant differentially expressed genes and hypoxia-related genes. **(D)** Boxplots of the expression levels of 19 differentially expressed HRGs in IS and healthy controls. The blue box plots above the corresponding gene name represent the expression in IS, whereas the yellow box plots represent the expression in healthy controls.

from controls. As shown in Fig. 5, the value of SLC2A3, MYH9, CREB1, and SAP30 was 0.918 (95% CI 0.850, 0.986), 0.904 (95% CI 0.830, 0.978), 0.779 (95% CI 0.650, 0.908), 0.826 (95% CI 0.726, 0.926), respectively.

Gene	Full name	logFC	adj.P value
HIF1A	Hypoxia inducible factor 1 subunit alpha	0.652461	0.0012319
HIGD2A	HIG1 hypoxia inducible domain family member 2A	- 0.55777	3.708E-06
CREB1	CAMP responsive element binding protein 1	0.537547	0.0014228
FOS	Fos proto-oncogene, AP-1 transcription factor subunit	0.641672	3.112E-05
ADM	Adrenomedullin	0.837845	8.063E-06
BNIP3L	BCL2 interacting protein 3 like	0.943066	0.0088373
DUSP1	Dual specificity phosphatase 1	0.760383	1.776E-05
GAA	Alpha glucosidase	0.51596	1.367E-05
GLRX	Glutaredoxin	0.545609	0.0006402
HK2	Hexokinase 2	0.869945	0.0006451
IRS2	Insulin receptor substrate 2	0.804123	4.475E-05
MYH9	Myosin heavy chain 9	0.760886	2.106E-06
NDST1	N-deacetylase and N-sulfotransferase 1	0.608013	0.0002353
NFIL3	Nuclear Factor, Interleukin 3 Regulated	0.770197	8.058E-06
RRAGD	Ras related GTP binding D	0.733595	5.264E-06
SAP30	Sin3A associated protein 30	0.772688	0.0002318
SIAH2	Siah E3 ubiquitin protein ligase 2	0.552496	0.0135577
SLC2A3	Solute carrier family 2 member 3	0.753791	1.253E-06
WSB1	WD repeat and SOCS box containing 1	0.72695	0.0005869

Table 2. Information on the 19 differentially expressed hypoxia-related genes.

External validation of diagnostic model performance. To assess the clinical benefit of the four genes, we used ROC curves to demonstrate their value in discriminating IS diagnoses. And the GSE58294 dataset was used to validate the diagnostic accuracy of the four HRGs biomarkers. As can be seen from the Fig. 5, SLC2A3, MYH9, CREB1 and SAP30 were also indicated admirable diagnostic accuracy with the value were 0.921(95% CI=0.857, 0.984),0.743(95% CI=0.635, 0.851),0.866(95% CI=0.789, 0.943) and 0.959(95% CI=0.917, 1.000), respectively (Fig. 5).

SLC2A3 may represent a new candidate gene in IS. As Fig. 5 shows, SLC2A3 had the best performance among the four hypoxia genes in distinguishing IS patients from control samples. In order to further detect the character of SLC2A3 in IS, the GSE58294 dataset was analyzed. The SLC2A3 levels in IS patients were expressively higher than those in control samples (Fig. 6A). The results obtained from the preliminary analysis stand up for that SLC2A3 may represent a remarkable candidate gene in IS "Guilt-by-association" method was applied to explore the biological function of SLC2A3 in IS²⁵. In order to identify the related genes of SLC2A3, Spearman correlation analysis was performed between SLC2A3 and 555 differentially expressed genes in the patient sample. There were 296 differentially expressed genes that were remarkably associated with SLC2A3 ($p < 0.05$). In addition, Metascape was applied to explore the functional enrichment analysis. From the chart, it can be seen that these genes were remarkably enriched in Neutrophil degranulation and immune response-regulating signaling pathway, which were closely related to IS²⁶ (Fig. 6B,C).

Immune cell infiltration results. The proportions of NK cells, neutrophils, activated dendritic cells, eosinophils, and memory B cells in IS patients were significantly higher than those in control samples. However, the proportion of Activated CD8 T cells and Activated B cells in IS was substantially lower than that in control tissues (Fig. 7A). Correlation heatmap between immune cell subpopulations in ischemic stroke disclosed that Plasma cells were negatively correlated with B cells naive. Neutrophils represented an apparent relationship with T cells CD8, B cells naive, and T cells CD4 memory activated were negatively correlated with T cells CD8. NK cells resting were positively correlated with T cells CD4 memory activated and T cells CD8 (Fig. 7B).

Relationship between critical HRGs and immune cell. SLC2A3 had a positive correlation with Neutrophils ($r=0.635$, $P < 0.001$) and Macrophages M0 ($r=0.395$, $P=0.012$) (Fig. 8A), and MYH9 was not only positively correlated with Neutrophils ($r=0.35$, $P=0.026$) but also with NK cells resting ($r=0.585$, $P < 0.001$) (Fig. 8B). Besides, CREB1 had a positive correlation with Dendritic cells activated ($r=0.449$, $P=0.004$) (Fig. 8C). And the Gamma delta T cells was associated with SAP30 ($r=0.350$, $P=0.029$) (Fig. 8D). These results suggested that the condition of the brain microenvironment in ischemic stroke could be partially reflected by SLC2A3/CREB1/MYH9/SAP30.

Discussion

Up to now, several reports have shown that HRGs could serve as diagnostic or prognostic markers in many types of tumors, such as osteosarcoma²⁷, bladder cancer²⁸, and triple-negative breast cancer²⁹. However, little research has reported whether hypoxia genes or proteins could be applied as biomarkers in ischemic stroke for

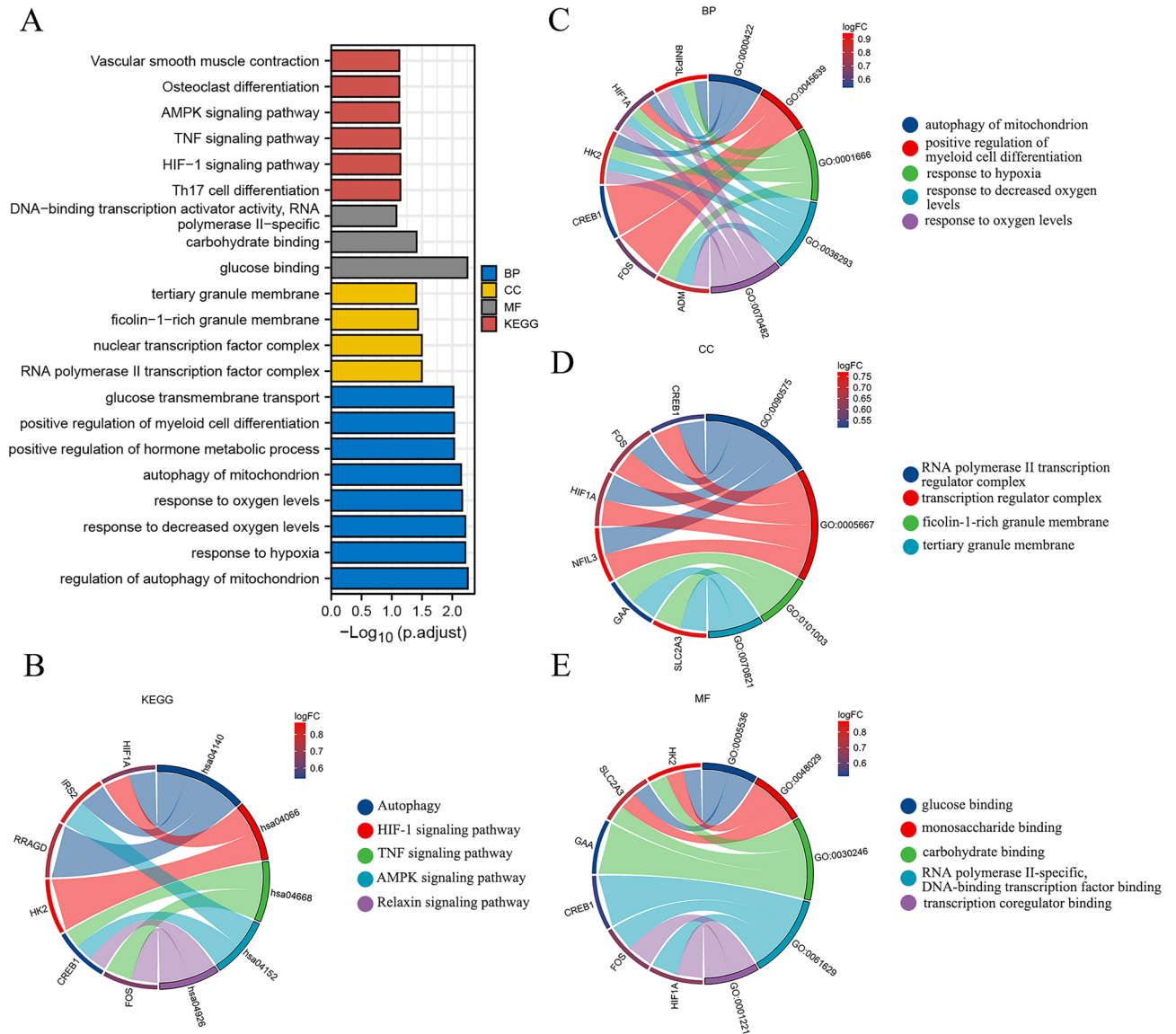


Figure 3. Functional enrichment for HRGs. (A,B) GO and KEGG analysis of 19 HRGs. (B–E) The enriched items in GO and KEGG analysis. "BP" stands for "biological process", "CC" stands for "cellular component", "MF" stands for "molecular function," and "KEGG" stands for "Kyoto Encyclopedia of Genes and Genomes".

early prevention. HRGs play a vital role in the overall development of ischemic stroke. In this study, we have investigated the role of HRGs in the diagnosis of IS. Four HRGs (SLC2A3, CREB1, MYH9, and SAP30) have been identified as crucial biomarkers for IS, and they have established the diagnostic model. External validation analysis also demonstrated that the diagnostic model has an excellent capacity to differentiate IS patients from controls.

We first explored hypoxia-related DEGs in ischemic stroke versus normal controls using GSE16561. As a result, a total of 19 HRGs were identified. In addition, GO and KEGG analysis revealed that these 19 hypoxia-related DEGs were mainly enriched in response to hypoxia, HIF-1 signaling pathway, autophagy, autophagy of mitochondrion, and AMPK signaling pathway. Hypoxia induces peripheral neutrophils to move toward the brain and release proteases, leading to impaired blood–brain barrier integrity and exacerbating oxidative stress³⁰. Pericyte HIF-1 can also cause blood–brain barrier disruption while worsening the poor prognosis of stroke³¹. The exact role of autophagy in the pathogenesis of ischemic stroke is controversial. The appropriate autophagy can exert neurological and excessive autophagy leads to neuronal death³². Some studies have confirmed that hypoxia and glucose deficiency will activate the AMPK pathway to trigger autophagy. The application of autophagy inhibitors can reduce the area of cerebral ischemia³³. Interestingly, metformin, as the first-line treatment for diabetes, can play a protective role in ischemic stroke by activating autophagy³⁴. Mitochondrial autophagy is a selective autophagy that prevents the production of excess reactive oxygen species (ROS) and subsequent cell death after the onset of ischemic stroke through mitochondrial autophagy³⁵.

Immunity plays an essential role in the progression of ischemic stroke, and there are growing studies of neuroprotective effects through neuroinflammatory modulation. The inflammatory mechanisms can exacerbate

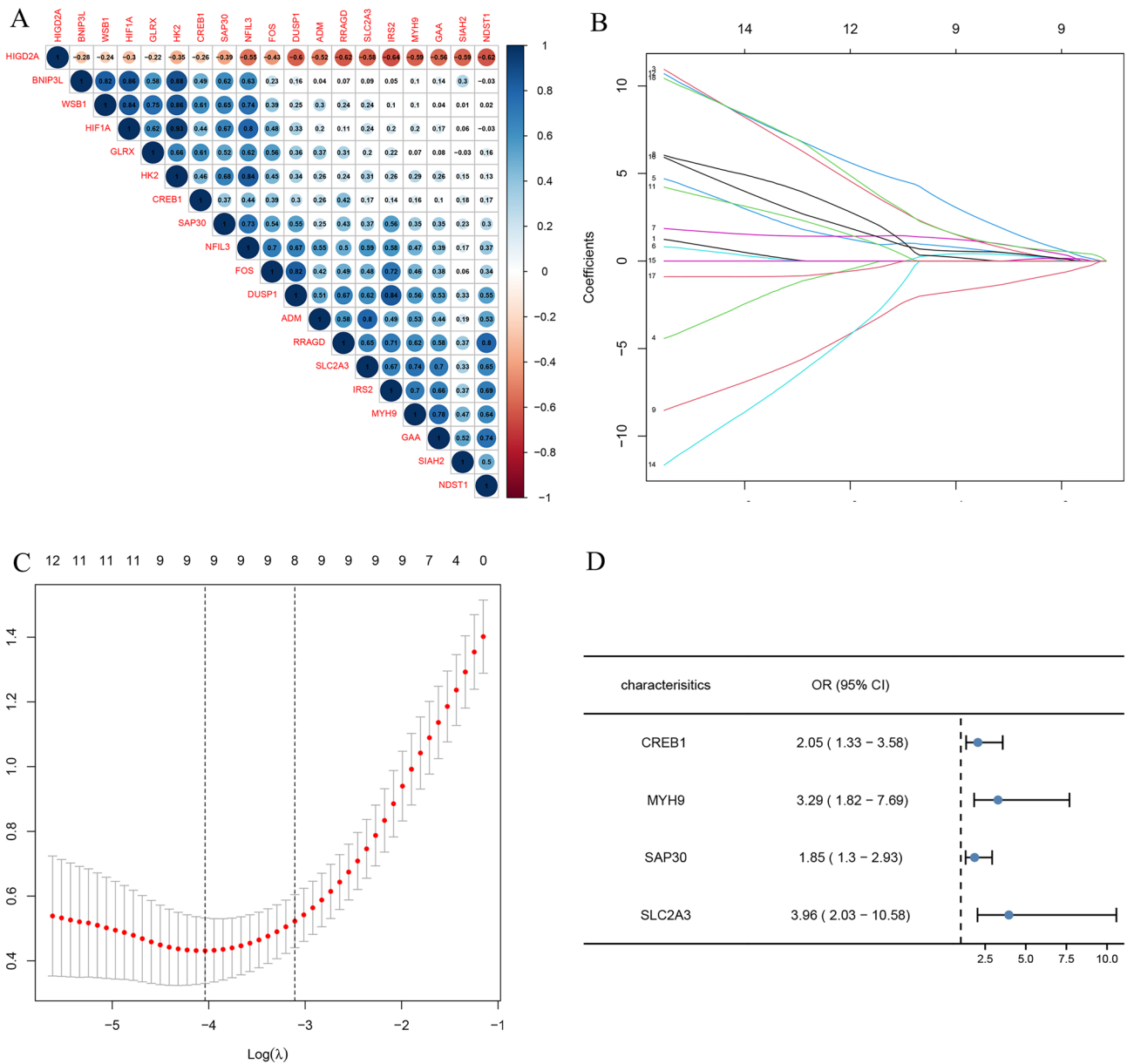


Figure 4. Screening of the optimal HRGs used for the final construction of the diagnostic model. **(A)** Correlation heat map of 19 differentially expressed hypoxia-related genes. The size of the colored squares represents the strength of the correlation; blue represents a positive correlation, and red represents a negative correlation. The darker the color is, the stronger correlation is. **(B,C)** Screening of the optimal parameter (using lambda.1se as the best lambda) at which the vertical lines were drawn. LASSO coefficient profiles of the 19 differentially expressed hypoxia-related genes. **(D)** Multivariate logistic regression determined independent candidate diagnostic biomarkers.

ischemic tissue damage and worsen clinical outcomes related to excessive immune response³⁶. Our results found that the proportions of NK cells, neutrophils, activated dendritic cells, eosinophils, and memory B cells were expressively higher than the control samples. Different immune cells play different roles in ischemic stroke. NK cells can pass the blood–brain barrier and play a dual role in the regulation of the nervous system and immune system.

On the one hand, NK cells exert toxic effects on injured neuronal cells, mediating inflammatory responses involved in cerebral ischemic injury and affecting stroke progression³⁷. On the other hand, the injured brain suppresses NK cells-mediated peripheral immune defense to avoid further damage³⁶, which will increase susceptibility to infection and cause increased mortality³⁸. Peripheral blood neutrophils can be used as an early indicator of stroke outcome. N2 neutrophils also promote macrophage phagocytosis, which is less damaging to ischemic neurons and is expected to be a new treatment³⁹. Dendritic cells (DCs) act cooperatively with other immune cells to influence brain injury in ischemic stroke. Animal experiments confirmed that type 2 DCs rapidly infiltrate the brain after MACO and promote neutrophil infiltration by inducing IL-17 production by $\gamma\delta$ T

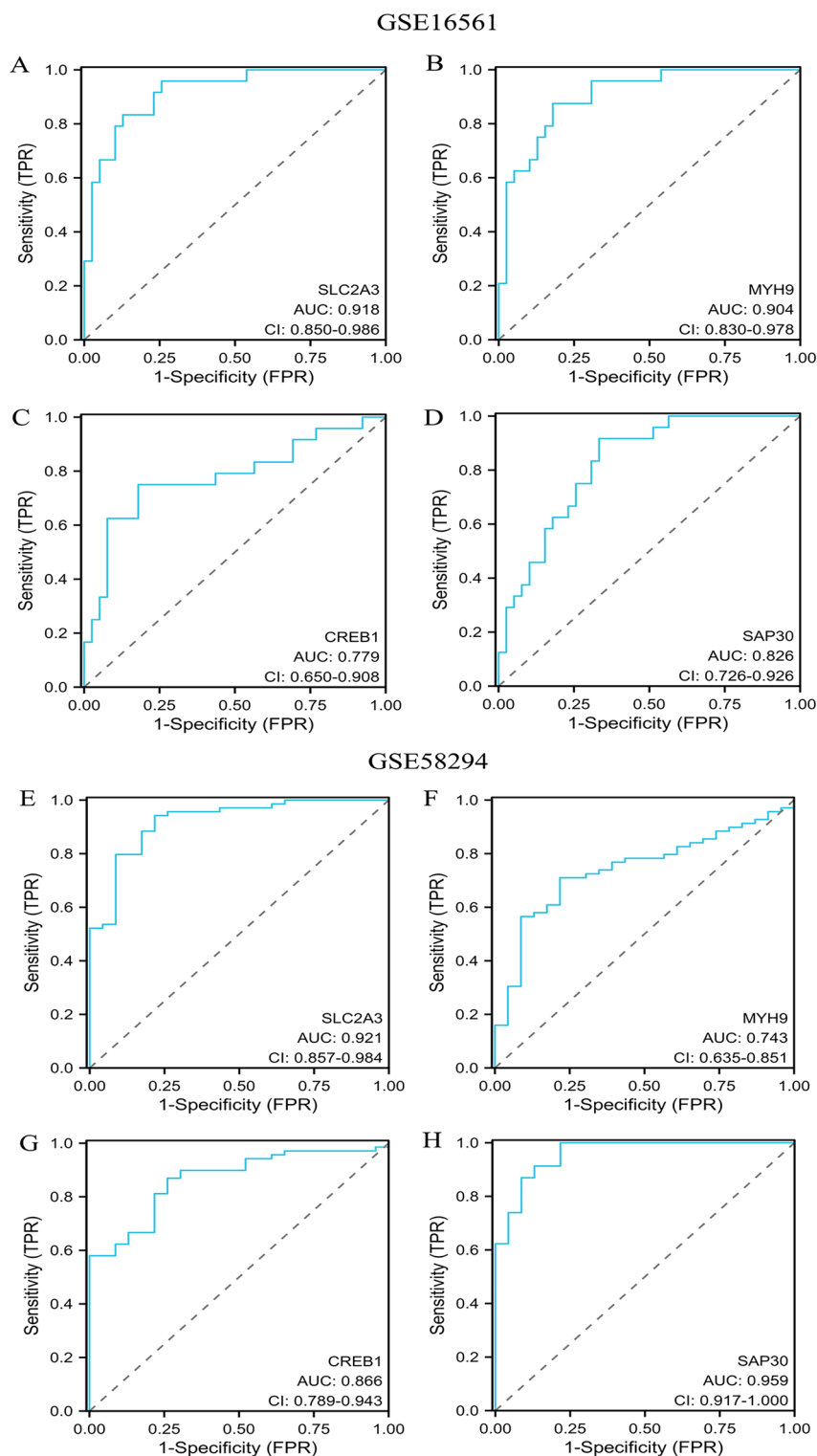


Figure 5. ROC analysis revealed the diagnostic value of hypoxia-related genes in ischemic stroke. (A–D) The AUC of hypoxia-related biomarkers in the diagnosis of ischemic stroke in ROC analysis in the GSE16561 dataset. (E–H) The AUC of hypoxia-related biomarkers in the diagnosis of ischemic stroke in ROC analysis in the GSE58294 dataset.

cells⁴⁰. In addition, DCs initiate Regulatory T cells (Tregs) that suppress immune responses and protect against brain injury caused by MACO⁴¹.

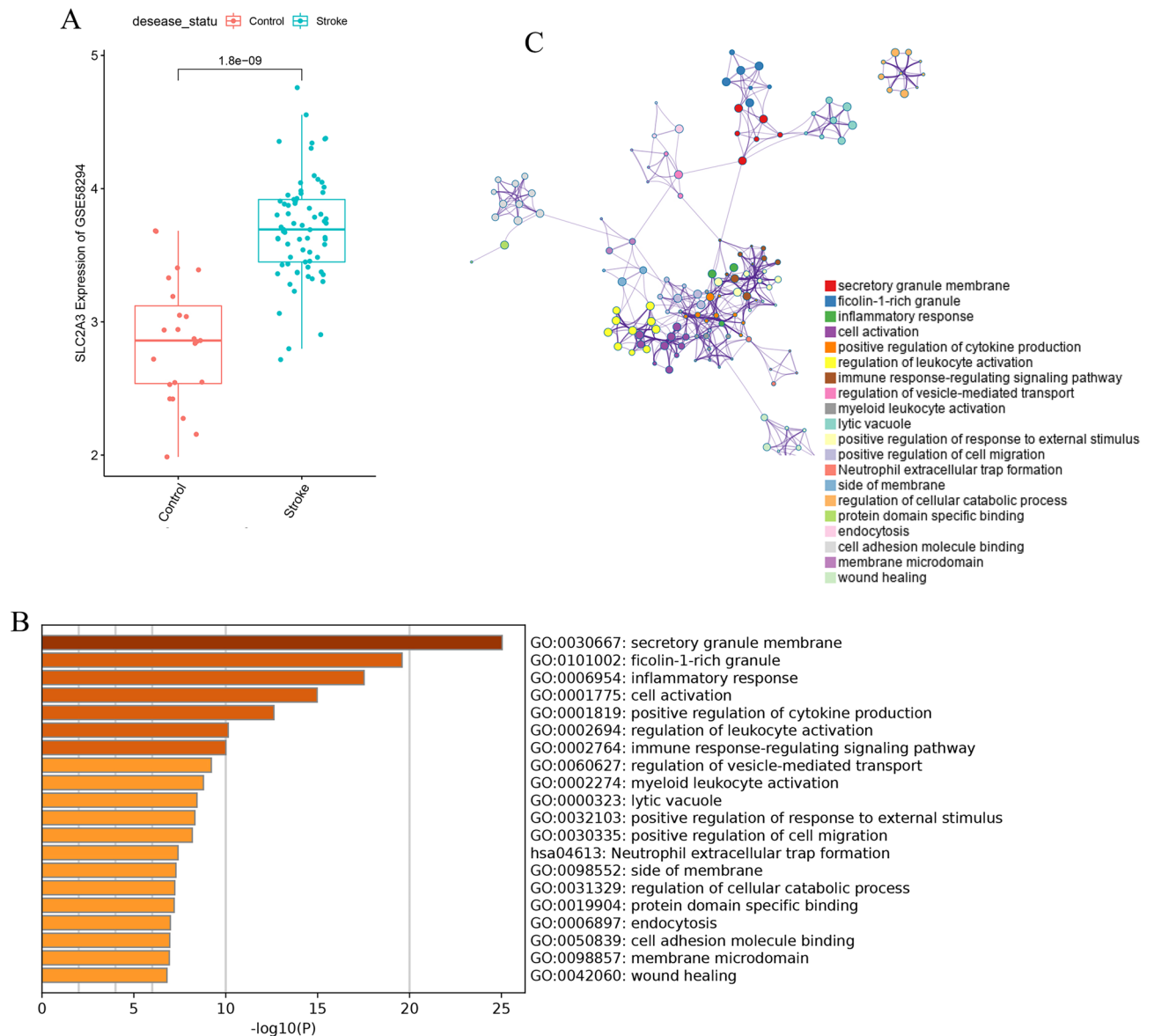


Figure 6. SLC2A3 may represent a new candidate gene in IS. **(A)** SLC2A3 mRNA levels are up-regulated in Ischemic stroke. **(B,C)** Enrichment analyses based on GO and KEGG pathways to predict the potential function of SLC2A3 using Metascape.

We applied a machine learning approach for screening essential genes for ischemic stroke using its better predictive performance, lower error rate, and higher reliability⁴². In our study, we screened four signature genes, namely SLC2A3, CREB1, MYH9, and SAP30, by LASSO algorithm and multivariate logistic regression, and these four signature genes have good diagnostic value. These four signature genes had good diagnostic value in the training dataset (GSE16561). Similarly, their diagnostic value results in the validation dataset (GSE58294) were satisfactory. The early prediction and diagnosis of ischemic stroke through diagnostic models are of great importance, especially for patients with a high risk of ischemic strokes, such as TIA⁴³, hypertension⁴⁴, and atherosclerosis⁴⁵. Early screening by molecular markers of hypoxia in peripheral blood allows effective stratification of high- and low-risk populations for management, thereby reducing the risk of stroke and improving poor prognosis.

CAMP Responsive Element Binding Protein 1 (CREB1), encoding a transcription factor that belongs to the leucine zipper family of DNA binding proteins. Prior studies have noted that pharmacological intervention can protect ischemic brain tissue by affecting CREB. The application of SMND-309 exerted significant efficacy in protecting neuronal cells from death by triggering the PI3K/Akt/CREB signaling pathway and inducing a considerable increase in BDNF⁴⁶. Improvement of neuroplasticity by interfering with CREB expression has also been reported, resulting in improved disease prognosis^{47,48}. CREB1 has multiple roles in neuroimmune. Its phosphorylation inhibits NF- κ B activation and suppresses pro-inflammatory responses⁴⁹. At the same time, CREB1 is involved in macrophage generation and differentiation and plays an essential role in myelin regeneration^{50,51}.

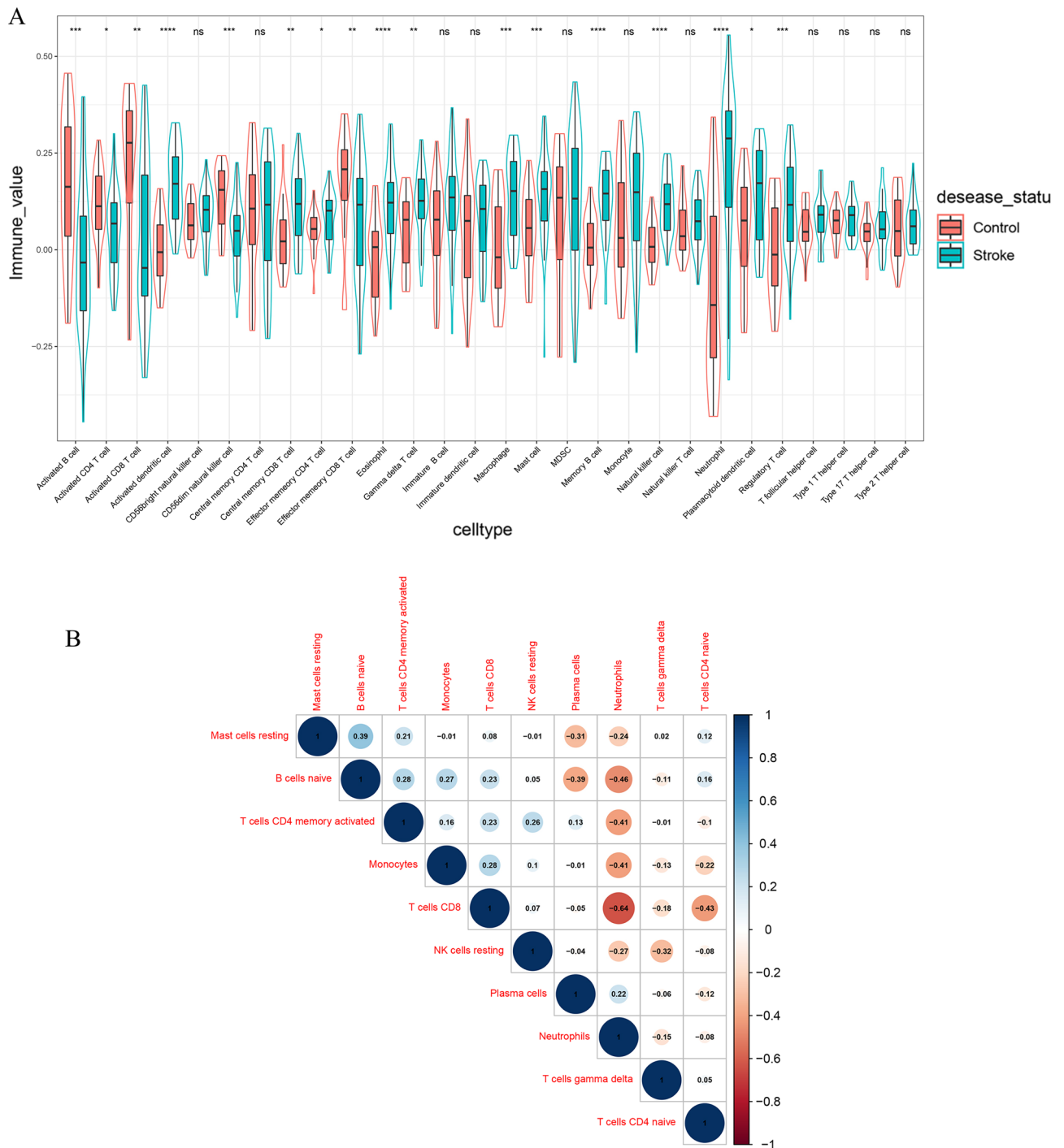


Figure 7. Correlation of infiltrating immune cells. **(A)** The violin plot exhibits the differences in CIBERSOFT immune cell fractions between IS and healthy controls. **(B)** Correlation heat map of immune cells. The size of the colored squares represents the strength of the correlation; blue represents a positive correlation, and red represents a negative correlation. The darker the color is, the stronger correlation is.

Myosin heavy chain 9 (MYH9) is one of the members that make up the cytoskeleton⁵². It is involved in blood–brain barrier dysfunction in ischemic stroke⁵³ and interacts with actin to mediate oxidative stress-induced neuronal apoptosis⁵⁴. Studies have demonstrated that inhibition or knockdown of MYH9 can slow down the blood–brain barrier damage and exert neuroprotective effects^{55,56}. In recent years, attention has also been paid to the vital role of MYH9 in the immune response⁵⁷, which has a positive correlation with NK cells in our study. It promotes cytotoxicity by binding to NK cell lysis granules⁵⁸. In innate immunity, MYH9 has an essential and fundamental role in neutrophil transport, which is closely associated with ischemic stroke⁵⁹. It provides a new biomarker and therapeutic target for hypoxia-induced blood–brain barrier damage.

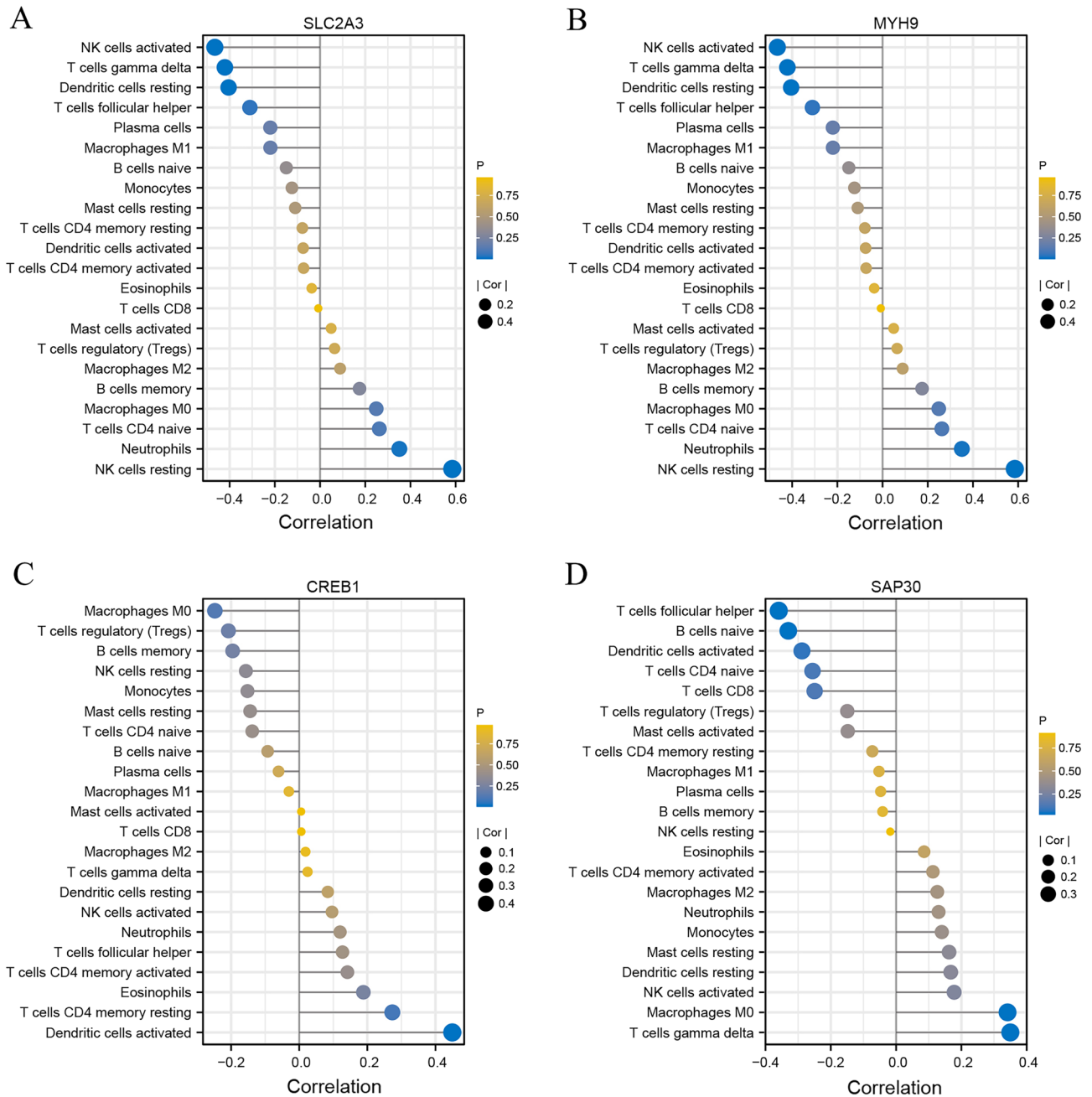


Figure 8. Correlation between CREB1, MYH9, SAP30, and SLC2A3 and immune cells. (A) Correlation between SLC2A3 and infiltrating immune cells. (B) Correlation between MYH9 and infiltrating immune cells. (C) Correlation between CREB1 and infiltrating immune cells. (D) Correlation between SLC2A3 and infiltrating immune cells. The size of the dots represents the strength of the correlation between genes and immune cells, the larger dots represent the stronger the correlation, and the smaller dots represent the weaker the correlation.

Sin3A-related protein 30 (SAP30), a gene encoding a protein of the histone deacetylase complex, contains SIN3, SAP18, HDAC1, HDAC2, RbAp46, RbAp48, and other polypeptides. There are few reports on SAP30 in ischemic stroke, while HDAC1 has been reported to play a vital role in neurological diseases. Improving cerebral ischemic neuroinflammation by modulating SAP30 levels and inducing M1 microglia polarization⁶⁰ provides a new therapeutic idea for neuroprotective effects.

SLC2A3 (Solute Carrier Family 2 Member), a member of the Solute Carrier Family, plays a crucial part in the central nervous system. SLC2A3 can be involved in glucose transport across the plasma membrane and blood–brain barrier transport⁶¹, which is essential for hypoxia and glucose deprivation after the onset of ischemic stroke. The adaptive upregulation of SLC2A3 may slow neuronal apoptosis by increasing glucose transport⁶², thus improving prognosis. In addition, SLC2A3 can affect neuroinflammation after ischemic stroke by influencing

the conversion of CD4T cells to Treg cells^{63,64}. Based on the adaptive upregulation of SLC2A3 under hypoxia and glucose-deficient conditions, it may be applied to the early diagnosis and target intervention of ischemic stroke.

Since SLC2A3 has the best diagnostic effect, therefore, we further explored the biological role of SLC2A3 in ischemic stroke. Spearman correlation analysis was performed between SLC2A3 and 555 differentially expressed genes in patient samples. And a total of 296 differentially expressed genes were found, and they were significantly associated with SLC2A3 ($p < 0.05$). Furthermore, we applied the Metascape online tools to conduct the analyses of complete functional and pathway enrichment. These genes associated with SLC2A3 were significantly associated with immune response and immune cell activation and interaction (Fig. 6B). Neuronal damage mitigation by modulation of SLC2A3 may be an effective therapeutic strategy that would benefit individuals who suffer from or are at high risk for ischemic stroke.

It should not be overlooked that this study has several limitations. First, our current study was conducted based on a public dataset with profiles from blood samples rather than brain tissue, which is heterogeneous from the actual immune infiltration in ischemic stroke. Second, the number of published data is limited, and the clinicopathological parameters are not comprehensive, which may lead to potential errors or bias. Third, it is that the training and validation datasets are from different platforms. The effect of differences between platforms on the results cannot be excluded. Fourth, microarray data are not clinically practical, so research and detection of critical markers for immune cells will be an essential part of future work. The last drawback is that the causal relationship between hypoxia-related biomarkers and immune cells and stroke can only be predicted by theoretical analysis rather than prospective studies. Therefore, we will continue to monitor the progress of stroke studies, and further investigation is needed to confirm whether our novel biomarkers and immune cells are potential prognostic predictors or therapeutic targets for stroke.

In conclusion, our studies indicate that the selected four genes may serve as new key biomarkers to assist the diagnosis of ischemic stroke patients from the people at the early stage and serve as new potential therapeutic targets. Meanwhile, it is apparent from the results that immune cells have strong connections with SLC2A3, CREB1, MYH9, and SAP30, which may serve as the new therapeutic target. More biological and clinical experiments are needed to confirm further the function of the selected genes and immune cells.

Data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Received: 12 January 2023; Accepted: 27 June 2023

Published online: 29 June 2023

References

- Saini, V., Guada, L. & Yavagal, D. R. Global epidemiology of stroke and access to acute ischemic stroke interventions. *Neurology* **97**, S6–S16 (2021).
- GBD 2016 Neurology Collaborators. Global, regional, and national burden of neurological disorders, 1990–2016: A systematic analysis for the Global Burden of Disease Study 2016. *Lancet Neurol.* **18**, 459–480 (2019).
- Fan, J. *et al.* Global burden, risk factors analysis, and prediction study of ischemic stroke, 1990–2030. *Neurology* <https://doi.org/10.1212/WNL.000000000000207387> (2023).
- Zhang, G.-L., Zhu, Z.-H. & Wang, Y.-Z. Neural stem cell transplantation therapy for brain ischemic stroke: Review and perspectives. *World J. Stem Cells* **11**, 817–830 (2019).
- Kuriakose, D. & Xiao, Z. Pathophysiology and treatment of stroke: Present status and future perspectives. *Int. J. Mol. Sci.* **21**, 7609 (2020).
- Dirnagl, U., Iadecola, C. & Moskowitz, M. A. Pathobiology of ischaemic stroke: An integrated view. *Trends Neurosci.* **22**, 391–397 (1999).
- Qin, C. *et al.* Dual functions of microglia in ischemic stroke. *Neurosci. Bull.* **35**, 921–933 (2019).
- Woodruff, T. M. *et al.* Pathophysiology, treatment, and animal and cellular models of human ischemic stroke. *Mol. Neurodegener.* **6**, 11 (2011).
- Jahan, R. *et al.* Association between time to treatment with endovascular reperfusion therapy and outcomes in patients with acute ischemic stroke treated in clinical practice. *JAMA* **322**, 252–263 (2019).
- Yafasova, A. *et al.* Time to thrombolysis and long-term outcomes in patients with acute ischemic stroke: A nationwide study. *Stroke* **52**, 1724–1732 (2021).
- Saver, J. L. *et al.* Time to treatment with endovascular thrombectomy and outcomes from ischemic stroke: A meta-analysis. *JAMA* **316**, 1279–1288 (2016).
- Haupt, M., Gerner, S. T., Bähr, M. & Doeppner, T. R. Neuroprotective strategies for ischemic stroke-future perspectives. *Int. J. Mol. Sci.* **24**, 4334 (2023).
- Berge, E. *et al.* European Stroke Organisation (ESO) guidelines on intravenous thrombolysis for acute ischaemic stroke. *Eur. Stroke J.* **6**, 1–LXII (2021).
- Ma, H., Jiang, Z., Xu, J., Liu, J. & Guo, Z.-N. Targeted nano-delivery strategies for facilitating thrombolysis treatment in ischemic stroke. *Drug Deliv.* **28**, 357–371 (2021).
- Barr, T. L. *et al.* Genomic biomarkers and cellular pathways of ischemic stroke by RNA gene expression profiling. *Neurology* **75**, 1009–1014 (2010).
- Stamova, B. *et al.* Gene expression in peripheral immune cells following cardioembolic stroke is sexually dimorphic. *PLoS ONE* **9**, e102550 (2014).
- Liberzon, A. *et al.* Molecular signatures database (MSigDB) 3.0. *Bioinformatics* **27**, 1739–1740 (2011).
- Zhu, W., Nan, Y., Wang, S. & Liu, W. Bioinformatics analysis of gene expression profiles of sex differences in ischemic stroke. *Biomed. Res. Int.* **2019**, 1–7 (2019).
- Liang, W., Sun, F., Zhao, Y., Shan, L. & Lou, H. Identification of susceptibility modules and genes for cardiovascular disease in diabetic patients using WGCNA analysis. *J. Diabetes Res.* **2020**, 4178639 (2020).
- Kanehisa, M., Furumichi, M., Sato, Y., Ishiguro-Watanabe, M. & Tanabe, M. KEGG: Integrating viruses and cellular organisms. *Nucleic Acids Res.* **49**, D545–D551 (2021).
- Kanehisa, M. & Goto, S. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* **28**, 27–30 (2000).

22. Engebretsen, S. & Bohlin, J. Statistical predictions with glmnet. *Clin. Epigenet.* **11**, 123 (2019).
23. Robin, X. *et al.* pROC: An open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinform.* **12**, 77 (2011).
24. Newman, A. M. *et al.* Robust enumeration of cell subsets from tissue expression profiles. *Nat. Methods* **12**, 453–457 (2015).
25. Tian, W. *et al.* Combining guilt-by-association and guilt-by-profiling to predict *Saccharomyces cerevisiae* gene function. *Genome Biol.* **9**(Suppl 1), S7 (2008).
26. Zhou, Y. *et al.* Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat. Commun.* **10**, 1523 (2019).
27. Liu, D. *et al.* Five hypoxia and immunity related genes as potential biomarkers for the prognosis of osteosarcoma. *Sci. Rep.* **12**, 1617 (2022).
28. Sun, X. *et al.* Identification and validation of a hypoxia-related prognostic and immune microenvironment signature in bladder cancer. *Cancer Cell Int.* **21**, 251 (2021).
29. Yang, X. *et al.* A combined hypoxia and immune gene signature for predicting survival and risk stratification in triple-negative breast cancer. *Aging* **13**, 19486–19509 (2021).
30. Justicia, C. *et al.* Neutrophil infiltration increases matrix metalloproteinase-9 in the ischemic brain after occlusion/reperfusion of the middle cerebral artery in rats. *J. Cereb. Blood Flow Metab.* **23**, 1430–1440 (2003).
31. Tsao, C.-C. *et al.* Pericyte hypoxia-inducible factor-1 (HIF-1) drives blood-brain barrier disruption and impacts acute ischemic stroke outcome. *Angiogenesis* **24**, 823–842 (2021).
32. Shi, Q., Cheng, Q. & Chen, C. The role of autophagy in the pathogenesis of ischemic stroke. *Curr. Neuropharmacol.* **19**, 629–640 (2021).
33. Zhou, X.-Y. *et al.* Inhibition of autophagy blocks cathepsins-tBid-mitochondrial apoptotic signaling pathway via stabilization of lysosomal membrane in ischemic astrocytes. *Cell Death Dis.* **8**, e2618 (2017).
34. Jiang, T. *et al.* Acute metformin preconditioning confers neuroprotection against focal cerebral ischaemia by pre-activation of AMPK-dependent autophagy. *Br. J. Pharmacol.* **171**, 3146–3157 (2014).
35. Shen, L. *et al.* Mitophagy in cerebral ischemia and ischemia/reperfusion injury. *Front. Aging Neurosci.* **13**, 687246 (2021).
36. Iadecola, C. & Anrather, J. The immunology of stroke: From mechanisms to translation. *Nat. Med.* **17**, 796–808 (2011).
37. Chen, C., Ai, Q.-D., Chu, S.-F., Zhang, Z. & Chen, N.-H. NK cells in cerebral ischemia. *Biomed. Pharmacother.* **109**, 547–554 (2019).
38. Dirnagl, U. *et al.* Stroke-induced immunodepression: Experimental evidence and clinical relevance. *Stroke* **38**, 770–773 (2007).
39. Cai, W. *et al.* Functional dynamics of neutrophils after ischemic stroke. *Transl. Stroke Res.* **11**, 108–121 (2020).
40. Gelderblom, M. *et al.* IL-23 (Interleukin-23)-producing conventional dendritic cells control the detrimental IL-17 (Interleukin-17) response in stroke. *Stroke* **49**, 155–164 (2018).
41. Chen, C., Chencheng, Z., Cuiying, L. & Xiaokun, G. Plasmacytoid dendritic cells protect against middle cerebral artery occlusion induced brain injury by priming regulatory T cells. *Front. Cell Neurosci.* **14**, 8 (2020).
42. Wang, J. *et al.* Identification of immune cell infiltration and diagnostic biomarkers in unstable atherosclerotic plaques by integrated bioinformatics analysis and machine learning. *Front. Immunol.* **13**, 956078 (2022).
43. Mendelson, S. J. & Prabhakaran, S. Diagnosis and management of transient ischemic attack and acute ischemic stroke: A review. *JAMA* **325**, 1088–1098 (2021).
44. Ferrari, F. & Villa, R. F. Brain bioenergetics in chronic hypertension: Risk factor for acute ischemic stroke. *Biochem. Pharmacol.* **205**, 115260 (2022).
45. Lin, S., Li, X., Zhang, J. & Zhang, Y. Omentin-1: Protective impact on ischemic stroke via ameliorating atherosclerosis. *Clin. Chim. Acta* **517**, 31–40 (2021).
46. Wang, Y. *et al.* SMND-309 promotes neuron survival through the activation of the PI3K/Akt/CREB-signalling pathway. *Pharm. Biol.* **54**, 1982–1990 (2016).
47. Mohamed, S. K., Ahmed, A. A. E., El Morsy, E. M. & Nofal, S. The protective effect of zeranol in cerebral ischemia reperfusion via p-CREB overexpression. *Life Sci.* **217**, 212–221 (2019).
48. Liu, Y. *et al.* Agonistic analog of growth hormone-releasing hormone promotes neurofunctional recovery and neural regeneration in ischemic stroke. *Proc. Natl. Acad. Sci. USA* **118**, e2109600118 (2021).
49. Wen, A. Y., Sakamoto, K. M. & Miller, L. S. The role of the transcription factor CREB in immune function. *J. Immunol.* **185**, 6413–6419 (2010).
50. Chen, Y. *et al.* Lnc-M2 controls M2 macrophage differentiation via the PKA/CREB pathway. *Mol. Immunol.* **124**, 142–152 (2020).
51. Mishra, M. K. *et al.* Harnessing the benefits of neuroinflammation: Generation of macrophages/microglia with prominent remyelinating properties. *J. Neurosci.* **41**, 3366–3385 (2021).
52. Titus, M. A. Growing, splitting and stacking myosin II filaments. *Nat. Cell Biol.* **19**, 77–79 (2017).
53. Lv, Y. & Fu, L. The potential mechanism for Hydroxysafflor yellow A attenuating blood-brain barrier dysfunction via tight junction signaling pathways excavated by an integrated serial affinity chromatography and shotgun proteomics analysis approach. *Neurochem. Int.* **112**, 38–48 (2018).
54. Wang, Y. *et al.* Myosin IIA-related actomyosin contractility mediates oxidative stress-induced neuronal apoptosis. *Front. Mol. Neurosci.* **10**, 75 (2017).
55. Wang, G.-Y. *et al.* NMMHC IIA inhibition ameliorates cerebral ischemic/reperfusion-induced neuronal apoptosis through caspase-3/ROCK1/MLC pathway. *DDDT* **14**, 13–25 (2020).
56. Gong, S. *et al.* Endothelial conditional knockdown of NMMHC IIA (nonmuscle myosin heavy chain IIA) attenuates blood-brain barrier damage during ischemia-reperfusion injury. *Stroke* **52**, 1053–1064 (2021).
57. Fernandez-Prado, R., Carriazo-Julio, S. M., Torra, R., Ortiz, A. & Perez-Gomez, M. V. MYH9-related disease: It does exist, may be more frequent than you think and requires specific therapy. *Clin. Kidney J.* **12**, 488–493 (2019).
58. Sanborn, K. B. *et al.* Phosphorylation of the myosin IIA tailpiece regulates single myosin IIA molecule association with lytic granules to promote NK-cell cytotoxicity. *Blood* **118**, 5862–5871 (2011).
59. Zehrer, A. *et al.* A fundamental role of Myh9 for neutrophil migration in innate immunity. *J. Immunol.* **201**, 1748–1764 (2018).
60. Wang, J. *et al.* Long noncoding RNA H19 promotes neuroinflammation in ischemic stroke by driving histone deacetylase 1-dependent M1 microglial polarization. *Stroke* **48**, 2211–2221 (2017).
61. Peng, W. *et al.* Glucose transporter 3 in neuronal glucose metabolism: Health and diseases. *Metabolism* **123**, 154869 (2021).
62. Herbet, M. *et al.* Altered expression of genes involved in brain energy metabolism as adaptive responses in rats exposed to chronic variable stress; changes in cortical level of glucogenic and neuroactive amino acids. *Mol. Med. Rep.* <https://doi.org/10.3892/mmr.2019.9865> (2019).
63. Zhang, Y. *et al.* RAGE-mediated T cell metabolic reprogramming shapes T cell inflammatory response after stroke. *J. Cereb. Blood Flow Metab.* **42**, 952–965 (2022).
64. Macintyre, A. N. *et al.* The glucose transporter Glut1 is selectively essential for CD4 T cell activation and effector function. *Cell Metab.* **20**, 61–72 (2014).

Acknowledgements

The research was supported and approved by the Taishan Scholars Project (NO. tsqn201909186).

Author contributions

K.Y. performed data analysis; Z.Z. and X.L. conducted data collection. K.Y. and Z.Z. analyzed the results. T.W. and Z.J. conducted the statistical analysis. W.L. and X.L. proposed the research idea of this paper. All authors discussed the findings and reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-023-37753-2>.

Correspondence and requests for materials should be addressed to X.L. or W.L.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023