Huang *et al.*
European Journal of Medical Research (2025) 30:195
https://doi.org/10.1186/s40001-025-02454-1

European Journal of Medical Research

RESEARCH

Open Access

Identification of novel inflammatory response-related biomarkers in patients with ischemic stroke based on WGCNA and machine learning

Chenyi Huang[1], Dengxuan Wu[1], Guifen Yang[1], Chuchu Huang[1] and Li Li[1*]

Abstract

Background Ischemic stroke (IS) is one of the most common causes of disability in adults worldwide. This study aimed to identify key genes related to the inflammatory response to provide insights into the mechanisms and management of IS.

Methods Transcriptomic data for IS were downloaded from the Gene Expression Omnibus (GEO) database. Weighted gene co-expression network analysis (WGCNA) and differential expression analysis were used to identify inflammation-related genes (IRGs) associated with IS. Hub IRGs were screened using Lasso, SVM-RFE, and random forest algorithms, and a nomogram diagnostic model was constructed. The diagnostic performance of the model was assessed using receiver operating characteristic (ROC) curves and calibration plots. Additionally, immune cell infiltration and potential small molecule drugs targeting IRGs were analyzed. The expression of IRG was verified by qRT-PCR in healthy controls and IS patients.

Results Nine differentially expressed IRGs were identified in IS, including NMUR1, AHR, CD68, OSM, CDKN1A, RGS1, BTG2, ATP2C1, and TLR3. Machine learning algorithms selected three hub IRGs (AHR, OSM, and NMUR1). A diagnostic model based on these three genes showed excellent diagnostic performance for IS, with an area under the curve (AUC) greater than 0.9 in both the training and validation sets. Immune infiltration analysis revealed higher levels of neutrophils and activated CD4 + T cells, and lower levels of CD8 + T cells, activated NK cells, and naive B cells in IS patients. The hub IRGs exhibited significant correlations with immune cell infiltration. Furthermore, small molecule drugs targeting hub IRGs were identified, including chrysin, piperine, genistein, and resveratrol, which have potential therapeutic effects for IS. qRT-PCR evaluation demonstrated that the levels of blood biomarkers (AHR, OSM, and NMUR1) in IS patients could serve as distinguishing indicators between IS patients and healthy controls ($P < 0.05$).

Conclusion This study confirmed the significant impact of IRGs on the progression of IS and provided new diagnostic and therapeutic targets for personalized treatment of IS.

Keywords Ischemic stroke, Inflammatory response, Nomogram, Immune infiltration, Personalized treatment

Introduction

Stroke stands as a significant global health challenge, being the second leading cause of death and the primary contributor to disability worldwide [1]. Every year, approximately 15 million individuals are affected, with 5

*Correspondence:
Li Li
lily07172024@163.com
[1] Rehabilitation Department, Tongde Hospital of Zhejiang Province, Hangzhou 310012, Zhejiang, China

million fatalities and an equal number left with permanent physical impairments [2]. Recent data indicate a rising trend in stroke incidence, with one in four individuals expected to experience a stroke during their lifetime [3]. This increase is partly attributed to an aging population and the accumulation of risk factors, as well as the growing prevalence of these risk factors among younger populations in developing countries [4]. Ischemic stroke (IS), caused by a blood vessel obstruction reducing cerebral blood flow, accounts for over 80% of all stroke cases [5]. Despite extensive research on the mechanisms of neuronal injury and the development of targeted therapies, the clinical efficacy of such treatments remains suboptimal [6]. Early diagnosis of IS is crucial, but it is often delayed due to the reliance on multiple time-consuming procedures, including CT scans, MRI, and angiography, which can hinder timely intervention [7]. Additionally, current treatments, including drug and interventional thrombolysis, face limitations due to a narrow therapeutic window and associated risks [8]. Given these challenges, there is an urgent need to identify novel biomarkers with high sensitivity and specificity for early diagnosis. Such biomarkers could significantly reduce the time to diagnosis and improve patient outcomes, addressing the ongoing public health challenge posed by stroke [9, 10].

The intricate relationship between stroke and inflammation has emerged as a critical area of investigation. Post-stroke inflammation plays a pivotal role in both the acute phase and long-term recovery, influencing the extent of tissue damage and repair processes [11]. Inflammation begins almost immediately after a stroke event, as ischemia triggers a cascade of immune responses that involve both local glial cells and infiltrating leukocytes [12]. The activation of microglia and astrocytes, along with the recruitment of peripheral immune cells, leads to the release of various pro-inflammatory cytokines and chemokines, which can exacerbate tissue injury [13]. However, the inflammatory response also facilitates tissue repair and regeneration through the release of growth factors and other neuroprotective agents [14]. Despite this dual-edged sword, the precise molecular mechanisms underlying the inflammatory response in stroke remain incompletely understood. Identifying specific genes involved in this process could provide valuable insights into the pathophysiology of stroke and potentially lead to new therapeutic targets [15]. Moreover, the discovery of reliable biomarkers related to the inflammatory response could enable more accurate prognosis and personalized treatment strategies.

Weighted Gene Coexpression Network Analysis (WGCNA) is a powerful systems biology approach that constructs coexpression networks to identify gene modules and their relationships with biological pathways. This method has been widely utilized to predict alterations in disease-associated signaling pathways and to explore biomarkers for IS and other diseases [16–18]. In this study, we used WGCNA and differential expression analysis to identify inflammation response-related genes (IRGs) differentially expressed in IS. To select hub IRGs, we applied Lasso, support vector machine analysis with recursive feature elimination (SVM-RFE), and random forest algorithms. Using these hub IRGs, we constructed a nomogram for diagnosing IS. This approach not only holds promise for improving stroke diagnostic accuracy but also for identifying potential therapeutic targets that could mitigate the inflammatory response and enhance patient outcomes. Our findings may contribute to a more comprehensive understanding of the molecular mechanisms underlying stroke and pave the way for more effective interventions. The study's workflow diagram is shown in Fig. 1.

Materials and methods

Data collection and processing

Gene expression profiles were downloaded from the public database Gene Expression Omnibus (GEO; https://www.ncbi.nlm.nih.gov/geo/, accessed date: July 1, 2024) for the datasets GSE16561 (Last update date: Oct 04, 2019), GSE58294 (Last update date: Mar 25, 2019) and GSE22255 (Last update date: Mar 25, 2019). The GSE58294 dataset comprises 69 IS samples and 23 healthy controls, while the GSE22255 dataset includes 20 IS samples and 20 healthy controls. Both datasets were generated using the GPL570 platform (Affymetrix Human Genome U133 Plus 2.0 Array) and consist of peripheral blood samples. The GSE16561 dataset was generated using the GPL6883 Illumina HumanRef-8 v3.0 expression beadchip and consists of whole blood obtained from 39 IS patients and 24 healthy control subjects. Gene expression data were normalized using the limma package (version 3.60.4) [19], and batch effects were removed using the sva package (version 3.52.0) [20]. Additionally, 200 IRGs (Supplementary materials: Table S1) were collected from the Molecular Signatures Database (MSigDB; https://www.gsea-msigdb.org/gsea/msigdb, version 7.4, accessed date: July 1, 2024).

Weighted gene co-expression network analysis

WGCNA analysis was performed on the GSE22255 dataset using the WGCNA R package (version 1.73) [21] to identify candidate biomarkers or therapeutic targets related to IS. First, a sample clustering dendrogram was constructed based on cutHeight to remove outliers. A soft-thresholding power β was defined to ensure scale-free topology of the network. The

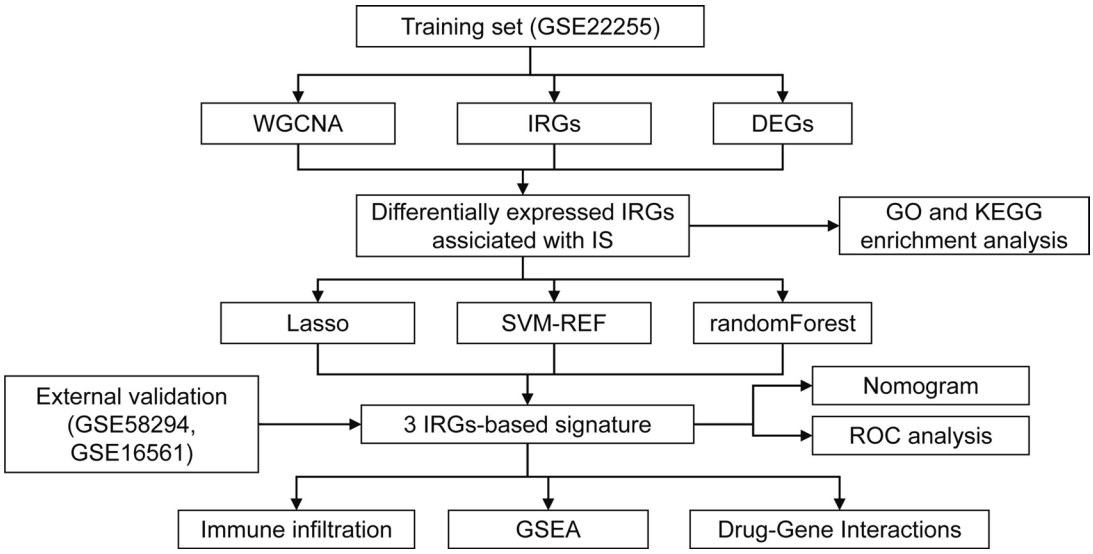

Fig. 1 The workflow of the study design

adjacency matrix was transformed into a topological overlap matrix (TOM), and dynamic tree cutting was applied to cluster genes into distinct modules. Modules significantly associated with IS phenotype were selected based on their correlation.

Differential expression analysis

Differential expression analysis between IS patients and healthy controls was conducted using the limma package in R. Genes with a p-value < 0.05 and a log2 fold change (logFC) of 0 were considered differentially expressed. Volcano plots were generated using the ggplot2 package (version 3.5.1) in R.

Enrichment analysis

The intersection of genes from the WGCNA-selected modules and differentially expressed genes (DEGs) underwent enrichment analysis for Gene Ontology (GO, https://geneontology.org/, accessed date: July 1, 2024) and Kyoto Encyclopedia of Genes and Genomes (KEGG, https://www.kegg.jp/, accessed date: July 1, 2024) pathways using the clusterProfiler package (version 4.12.6) [22]. Bar plots and dot plots were created to visualize the results. The GSE22255 cohort was divided into high and low expression groups based on the median expression of marker genes, and differential expression analysis was performed using limma. Gene set enrichment analysis (GSEA) for KEGG pathways was also conducted using clusterProfiler.

Feature selection

Three machine learning algorithms were employed for feature selection: (1) Lasso: Lasso analysis was performed using the glmnet package (version 4.1–8) [23] with tenfold cross-validation, and features with nonzero coefficients were selected. (2) SVM-RFE: Support vector machine analysis with SVM-RFE was implemented using the mRFE package (version 0.0.0.9000) [24] with tenfold cross-validation, and features with the lowest error were selected. (3) Random Forest: The randomForest package (version 4.7–1.2) [25] was used for random forest analysis, and the top five features were selected based on importance scores. Venn diagrams were used to identify overlapping features, which were then validated for differential expression in the GSE58294, GSE16561, and GSE22255 cohorts.

Nomogram construction and evaluation

A nomogram was constructed using the selected features in the GSE22255 cohort with the rms package (version 6.8–2). Calibration curves and receiver operating characteristic (ROC) curves were plotted to evaluate the performance of the IRGs and the nomogram, and the area under the curve (AUC) was calculated. Validation of the IRGs and the nomogram was performed in the GSE58294 cohort.

Immune infiltration assessment

Immune cell infiltration was assessed using the IOBR package (version 0.99.8) [26] and the CIBERSORT

algorithm to calculate the infiltration proportions of 22 immune cell types. Differences in immune cell infiltration between IS and control groups were compared using the Wilcoxon test. Correlations between marker IRGs and immune cell infiltration were calculated, and a heatmap was generated using the corrplot package (version 0.95).

Small-molecule drug prediction

Potential drugs targeting hub genes were identified using the Drug Gene Interaction Database (DGIdb, http://www.dgidb.org, accessed date: July 4, 2024) [27]. The determined target network was visualized using Cytoscape software (version 3.10.2, usage date: July 4, 2024).

Collection of whole blood samples

Participants were recruited from Zhejiang Provincial Tongde Hospital. All IS patients underwent comprehensive neurological examinations. IS diagnosis was established according to the International Classification of Diseases, Ninth Revision (ICD-9) criteria. Exclusion criteria included histories of hematological disorders, type 1 diabetes mellitus, autoimmune diseases, thyroid dysfunction, malignancies, or hepatic/renal pathologies. Whole blood samples were collected from 15 healthy controls and 15 IS patients using residual clinical specimens scheduled for routine diagnostic disposal. All samples were aliquoted and stored at − 80 °C until subsequent analysis. This study protocol adhered to the ethical principles outlined in the Declaration of Helsinki (2008 revision) and received approval from the Institutional Review Board of Zhejiang Provincial Tongde Hospital (Approval No. 2023-106 K). Written informed consent was waived by the ethics committee as only anonymized residual specimens were utilized.

RT-qPCR

RNAprep Pure High Efficiency Total RNA Extraction Kit (Cat no. DP443, TianGen, Beijing, China) was used to extract total RNA in blood serum. Briefly, 800 ng total RNA of each sample was used to perform reverse transcription to synthesis the first chain cDNA using TAKARA PrimeScript RT reagent Kit (Cat no. RR037, TAKARA, Japanese). Then, SYBR green reagent (TAKARA, Japanese) was used to determine the expression of target genes during process of amplification. GAPDH was used as reference to determine loading controls, while 2-detadeta T formula was used to calculate the relative expression of target genes. Primers used for the present study was shown as following: 5′-CGT CCCTGCATCCCACTACTT-3′ (AHR forward primer), 5′ - GGACATGGCCCCAGCATAG-3' (AHR reverse primer), 5'-GCACGGGCCAGAGTACCAGGAC-3' (OSM forward primer), 5'-CTGGTGTTGTAGTGGACC GTGAG-3' (OSM reverse primer), 5'-GGCTCCAGC AGCACGATC-3' (NMUR1 forward primer), 5'-GCA GATGCCAAACACCACG-3' (NMUR1 reverse primer), 5'-CGGAGTCAACGGATTTGGTCGTAT-3' (GAPDH forward primer), and 5'- AGCCTTCTCCATGGTGGT GAAGAC-3' (GAPDH reverse primer).

Statistical analysis

Statistical analysis was performed using SPSS (V. 27.0; ICM Corp., Armonk, NY, USA.) and R software (V. 3.6.2). Comparisons between groups were analyzed using unpaired t-tests, Wilcoxon test or one-way ANOVA, as appropriate. For all analyses, a two-sided $P < 0.05$ was considered to indicate statistical significance.

Results

Identification of significant module genes in IS via *WGCNA*

An unweighted scale-free co-expression network was established to identify key modules relevant to IS. Initially, we clustered samples from the GSE22255 dataset based on Euclidean distances of gene expression values to detect outliers; Fig. 2A shows that three outlier samples were identified. After removing these outliers, we re-clustered the remaining samples, where white represented control samples and red represented IS samples (Fig. 2B). Subsequently, when the soft threshold (power) was set to 6, the R^2 value reached 0.85, indicating that connectivity tended towards zero (Fig. 2C and D). Following this, the dynamic tree cut algorithm identified eight modules within the co-expression network (Fig. 2E). Based on the module-trait relationships depicted in Fig. 2F, we selected three modules (black, red, and turquoise) with correlations greater than 0.2 for further analysis. In total, 4,467 IS-related genes were identified across these three modules for subsequent analyses.

Identification of differentially expressed IRGs in IS

To explore the extent of gene expression differences between IS and normal conditions, we identified DEGs. We screened 1,049 genes between the IS and control groups, including 564 upregulated and 485 downregulated genes (Fig. 3A). Intersecting these DEGs with the 4,467 genes identified through WGCNA resulted in 383 IS-related DEGs. Enrichment analysis revealed that these genes were involved in processes such as regulation of translation, nuclear speckles, protein kinase activity, and the NF-kappa B signaling pathway (Fig. 3B and C). Ultimately, Venn diagram analysis indicated that nine of these IS-related DEGs were IRGs, which were further utilized for feature selection (Fig. 3D).

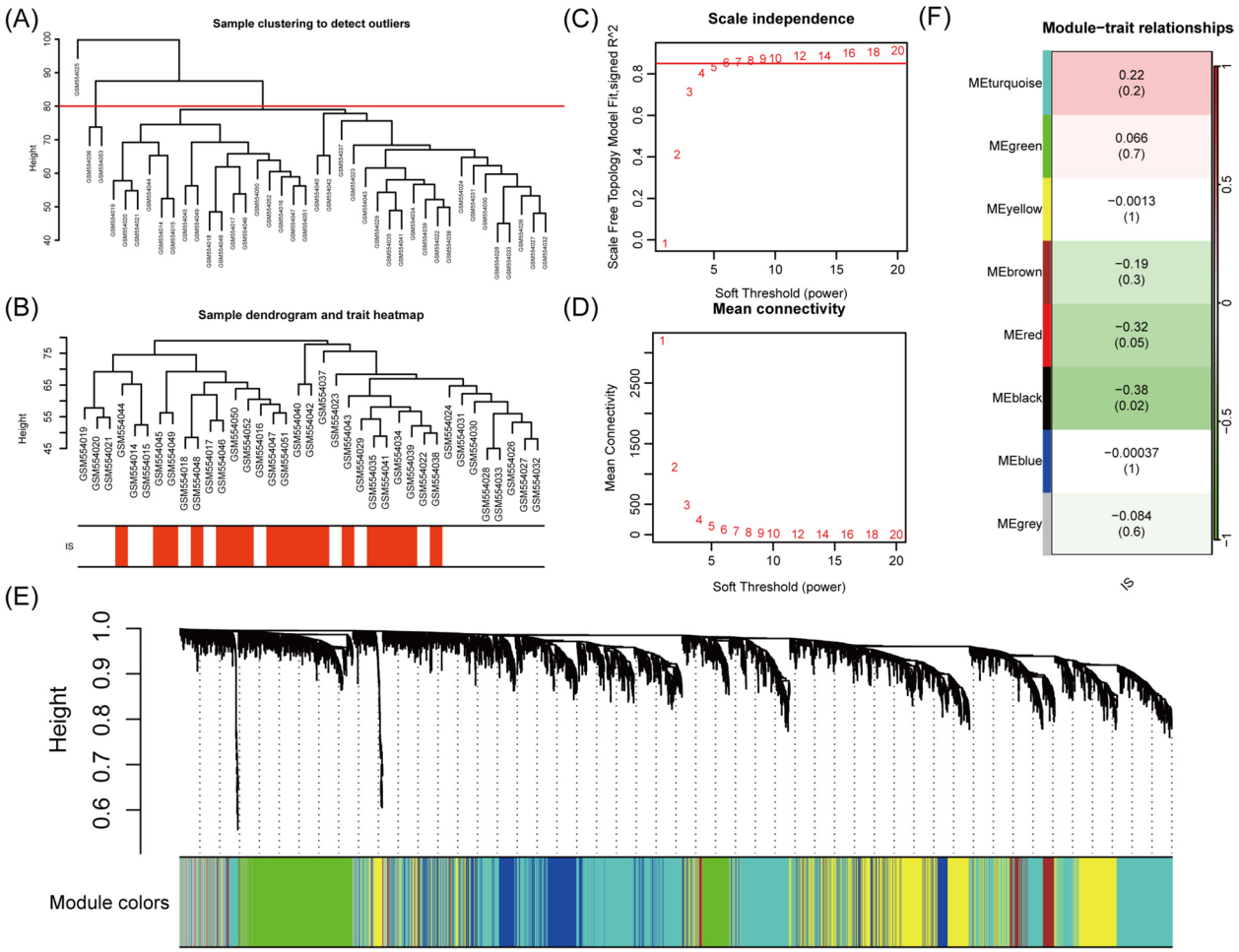

Fig. 2 Identification of Gene Modules Related to Ischemic Stroke via WGCNA. **A** Clustering of samples to detect outliers. **B** Sample dendrogram and trait heatmap. **C**, **D** Determination of the optimal soft-thresholding power. **E** Identification of modules based on the co-expression network. **F** Relevance of modules to ischemic stroke

Identification of hub IRGS using machine learning algorithms

Three different machine learning algorithms were employed to screen for reliable candidate hub genes in IS. LASSO regression identified five genes—BTG anti-proliferation factor 2 (BTG2), aryl hydrocarbon receptor (AHR), oncostatin M (OSM), ATPase secretory pathway Ca2+transporting 1 (ATP2C1), and neuromedin U receptor 1 (NMUR1)—as diagnostic markers for IS (Fig. 4A and B). Additionally, SVM-RFE selected the top eight feature genes with the minimum error, which included OSM, AHR, NMUR1, cyclin-dependent kinase inhibitor 1A (CDKN1A), BTG2, CD69, ATP2C1, and regulator of G protein signaling 1 (RGS1) (Fig. 4C). Random Forest analysis was employed to evaluate the importance of the differentially expressed IRGs in IS (Fig. 4D and E), and the top five genes identified were NMUR1, AHR, CD69, OSM, and CDKN1A. Finally, after overlaying key genes using a Venn diagram, the AHR, OSM, and NMUR1 were chosen as common potential hub genes in IS (Fig. 4F).

Expression characteristics of Hub IRGs

We further investigated the expression levels of hub IRGs in IS patients. In the GSE22255 cohort, AHR was significantly upregulated, NMUR1 was significantly downregulated, and the expression difference of OSM was not significant in IS patients compared to normal controls (Fig. 5A). In the GSE58294 and GSE16561 cohorts, OSM was significantly upregulated, NMUR1 was significantly downregulated, and the expression difference of AHR was not significant (Fig. 5B and C). Correlation analysis showed that OSM had a negative correlation with NMUR1, while the correlation of AHR with other genes was inconsistent across all cohorts. Furthermore, validation analyses in clinical specimens demonstrated

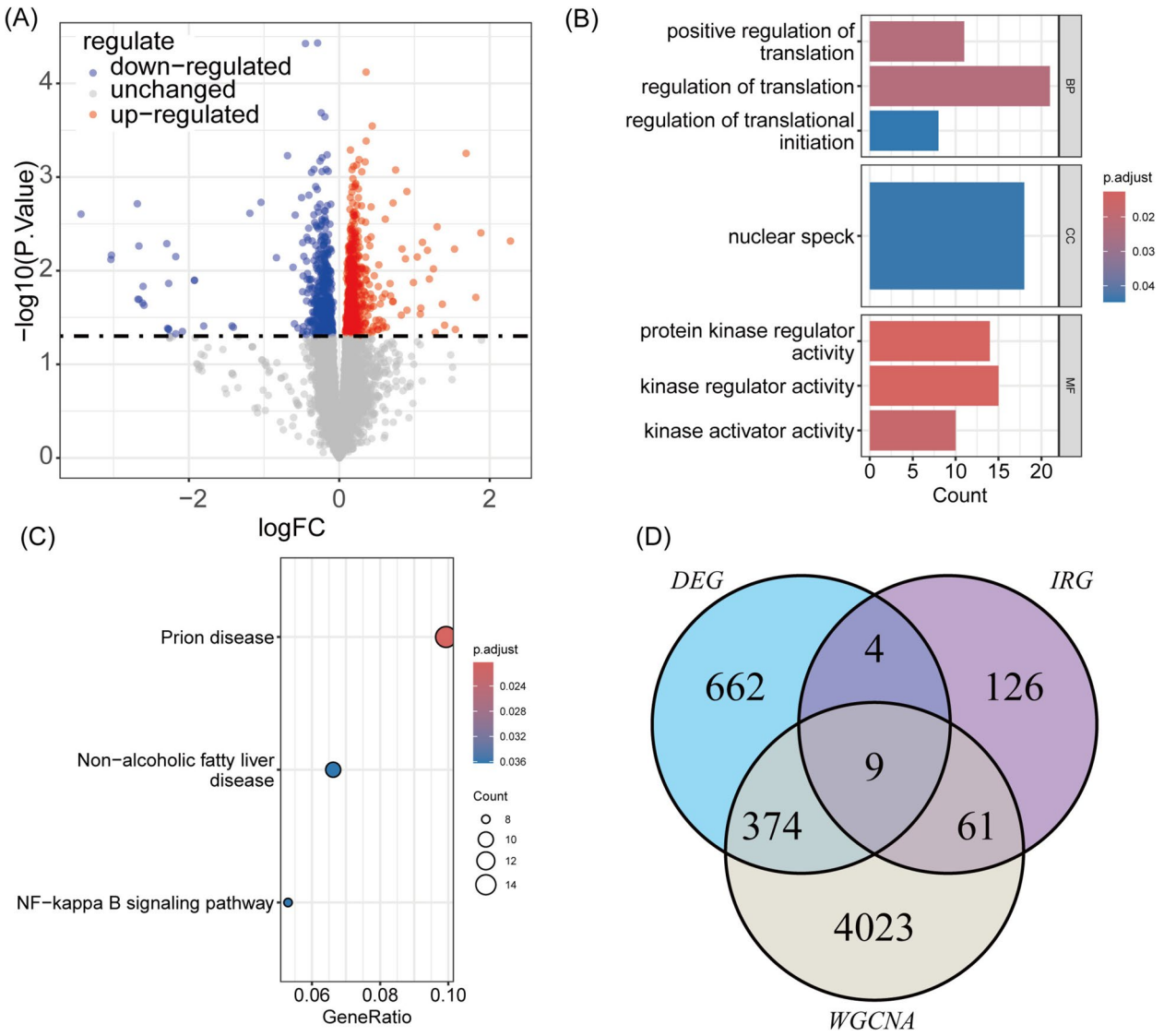

Fig. 3 Selection of Differentially Expressed IRGs Related to Ischemic Stroke. **A** Volcano plot illustrating the distribution of differentially expressed genes (DEGs) between the ischemic stroke and control groups. **B** Gene Ontology (GO) annotation of DEGs related to ischemic stroke. **C** Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis of DEGs related to ischemic stroke. **D** Venn diagram identifying overlapping genes among DEGs, IRGs, and WGCNA module genes

statistically significant elevations in AHR and OSM levels, coupled with a marked reduction in NMUR1 expression within peripheral blood samples of IS patients compared to healthy controls (Fig. 5D–F). These findings substantiate the clinical relevance of these hub genes as potential diagnostic and prognostic biomarkers for IS pathogenesis.

Development and evaluation of nomogram

Using the identified feature genes (AHR, OSM, and NMUR1; Fig. 6A), we developed a diagnostic nomogram for ischemic stroke and assessed its predictive ability using calibration curves. The calibration curves demonstrated minimal differences between actual and predicted risks of ischemic stroke, indicating excellent accuracy of the diagnostic nomogram (Fig. 6B). In the GSE22255 cohort, we performed ROC analyses to evaluate the predictive performance of the nomogram and the gene expressions of AHR, OSM, and NMUR1 for IS risk. The results showed that the AUCs were 0.906 for the nomogram (Fig. 6C), 0.705 for AHR (Fig. 6D), 0.398 for OSM (Fig. 6E), and 0.734 for NMUR1 (Fig. 6F), respectively. In the GSE58294 cohort, the corresponding AUC values for predicting IS risk were 0.944 for the nomogram (Fig. 6G),

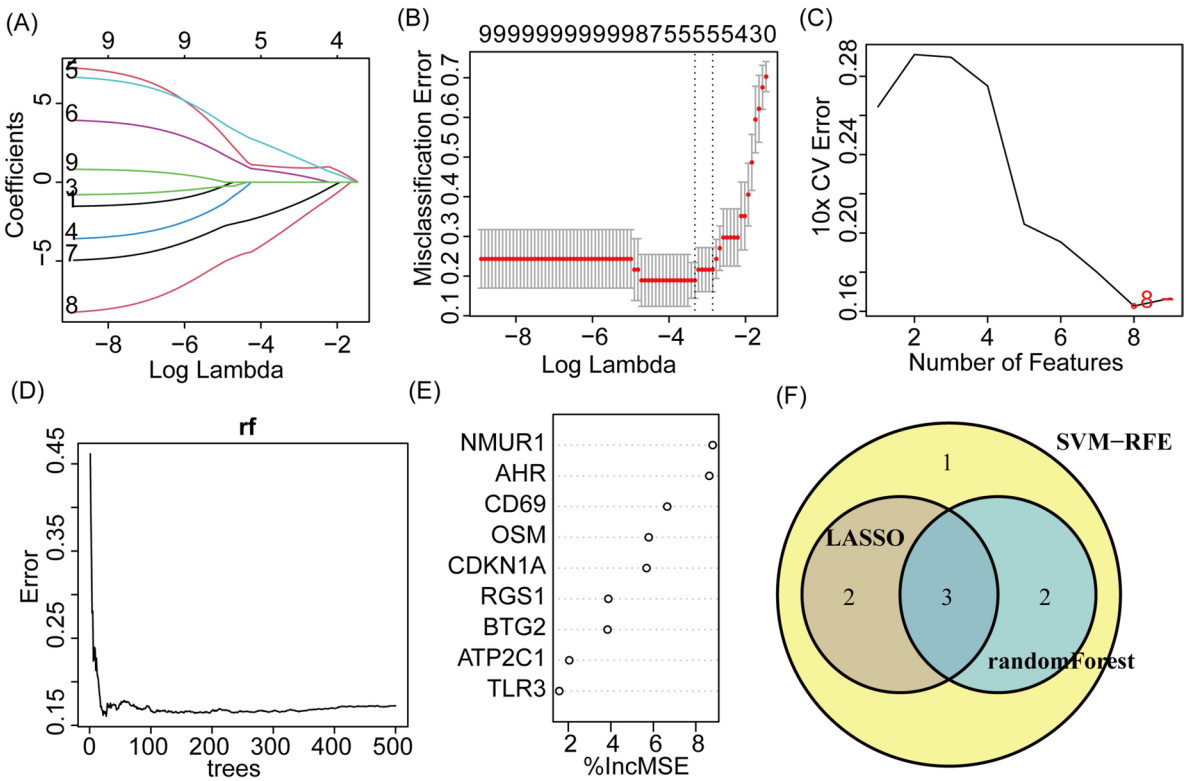

Fig. 4 Identification of Hub IRGs Using Machine Learning Algorithms. **A** Trajectory of independent variables in LASSO regression. **B** Confidence intervals at different lambda values in LASSO regression. **C** SVM-RFE algorithm for feature gene selection. **D** Error rate in the ischemic stroke and control groups using the random forest algorithm. **E** Genes ranked based on importance scores from the random forest algorithm. **F** Venn diagram showing the feature genes shared by LASSO, SVM-RFE, and random forest algorithms

(See figure on next page.)

Fig. 5 Validation of Expression Differences for Hub IRGs. **A** Boxplot and correlation analysis of hub IRGs in the GSE22255 cohort. **B** Boxplot and correlation analysis of hub IRGs in the GSE58294 cohort. **C** Boxplot and correlation analysis of hub IRGs in the GSE16561 cohort. **D** Comparison of AHR mRNA levels in blood samples from IS patients and healthy controls, as determined by qRT-PCR. **E** Comparison of OSM mRNA levels in blood samples from IS patients and healthy controls, as determined by qRT-PCR. Comparison of NMUR1 mRNA levels in blood samples from IS patients and healthy controls, as determined by qRT-PCR. *ns, not significant, *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, ****$p < 0.0001$

0.563 for AHR (Fig. 6H), 0.903 for OSM (Fig. 6I), and 0.834 for NMUR1 (Fig. 6J). These findings suggested that the nomogram demonstrated high precision in predicting IS risk, whereas the gene expressions of AHR, OSM, and NMUR1 exhibit inconsistent accuracy in this context.

Association of IRGs with IS immune infiltration

We further evaluated immune cell infiltration in IS patients and found that, compared to controls, IS patients had lower naive B cell, CD8 T cell, resting CD4 memory T cell, and activated NK cell infiltration, but higher plasma cell, activated CD4 memory T cell, resting NK cell, and neutrophil infiltration (Fig. 7A). Further analysis revealed that CD8 T cells had a significant negative correlation with AHR, OSM, and NMUR1. Additionally, the IRGs exhibited complex and diverse significant correlations with naive B cells, CD4 T cells, NK cells, monocytes, macrophages, dendritic cells, and neutrophils (Fig. 7B). These findings suggest that IRGs play a role in the formation and development of the immune microenvironment in IS patients.

Screening of small molecule drugs

For the treatment of IS patients, we used DGIdb to identify 76 potential drugs (Fig. 8), including 16 approved and 60 not approved (purple edges). Approved drugs included chrysin, methylcellulose, piperine, genistein, resveratrol, clioquinol, carbaryl, levothyroxine, niclosamide, phenazopyridine hydrochloride, tapinarof, nitazoxanide, thiabendazole, romiplostim, olanzapine, and

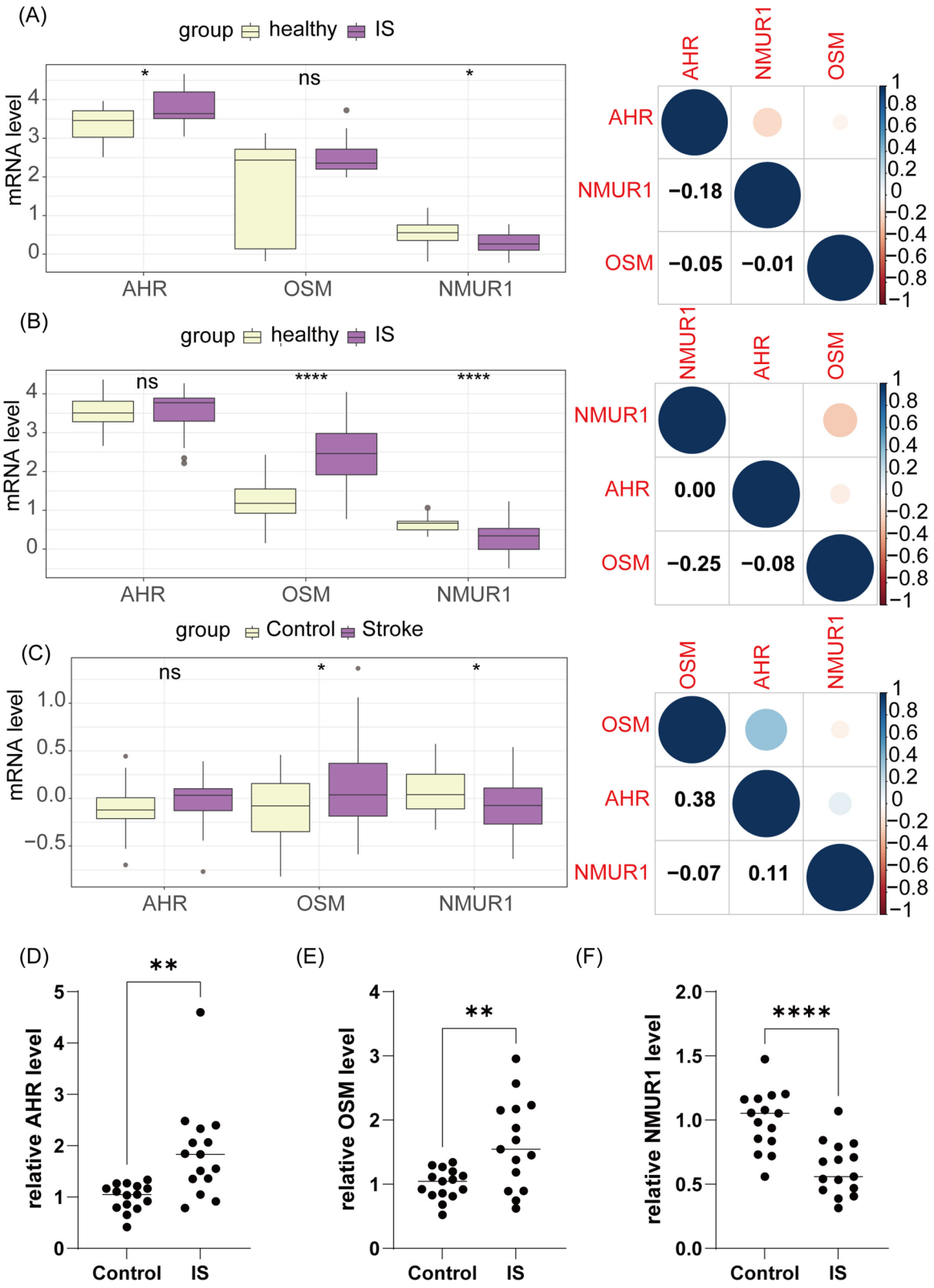

Fig. 5 (See legend on previous page.)

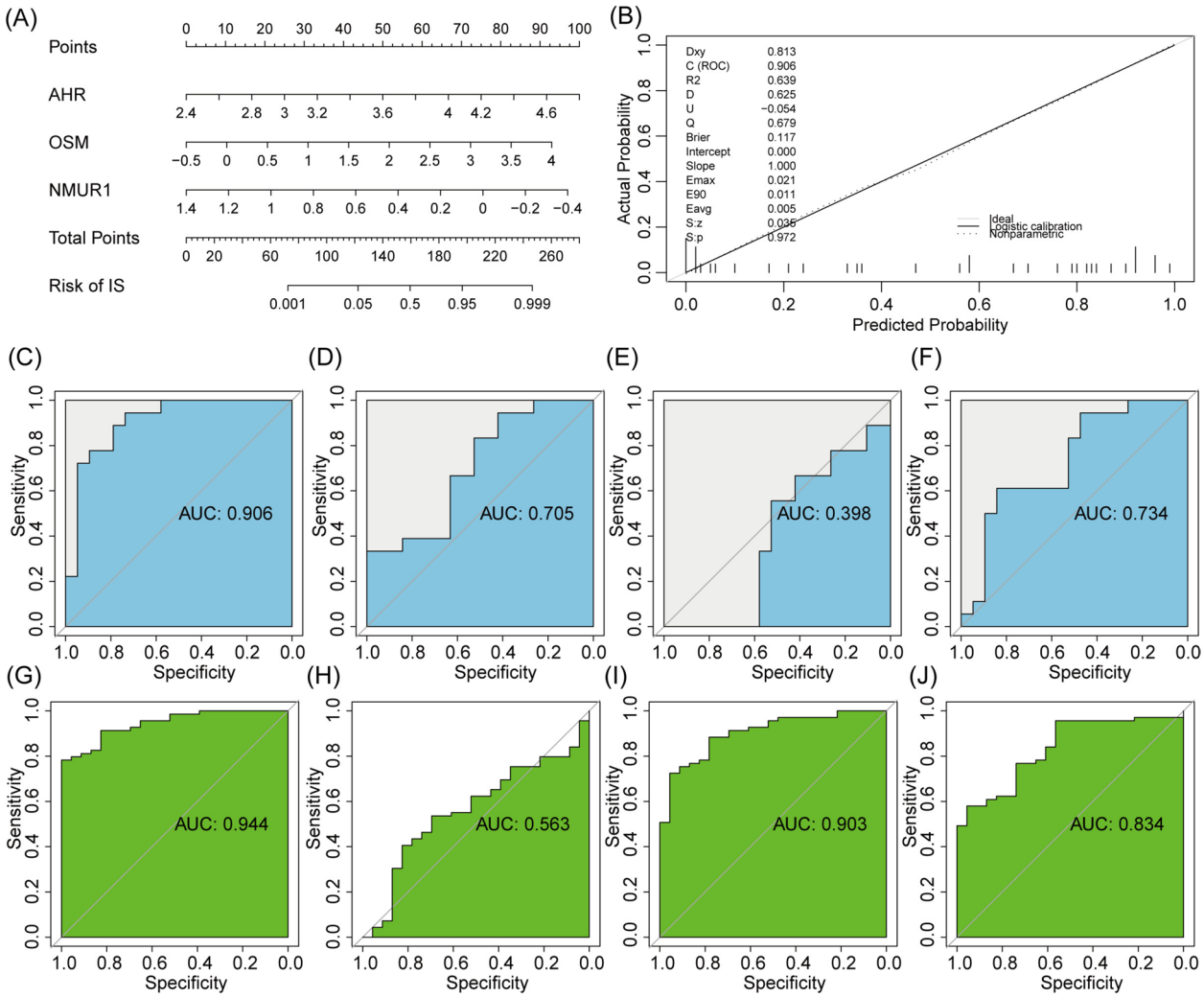

Fig. 6 Development and Evaluation of the Nomogram. **A** Nomogram for the diagnostic model of ischemic stroke. **B** Calibration curves to evaluate the predictive accuracy of the nomogram. **C** Receiver operating characteristic (ROC) curves for the nomogram in the GSE22255 cohort. **D** ROC curves for the AHR in the GSE58294 cohort. **E** ROC curves for the OSM in the GSE58294 cohort. ROC curves for the NMUR1 in the GSE58294 cohort. **G** ROC curves for the nomogram in the GSE58294 cohort. **H** ROC curves for the AHR in the GSE58294 cohort. **I** ROC curves for the OSM in the GSE58294 cohort. ROC curves for the NMUR1 in the GSE58294 cohort

omeprazole. A drug-gene network was constructed using Cytoscape, where all approved drugs targeted AHR, but there were no drugs targeting NMUR1 and OSM.

Pathways associated with hub IRGs

We analyzed the potential mechanisms of action of hub IRGs in IS using GSEA. It was found that genes in the high-expression cohorts of AHR, OSM, and NMUR1 were highly enriched in the ribosome pathway. In the high-expression cohort of AHR, genes were enriched in the C-type lectin receptor signaling pathway and the Fanconi anemia pathway, whereas in the low-expression cohort, genes were enriched in the TNF signaling pathway and thyroid hormone synthesis, secretion, and action (Fig. 9A). Genes in the high-expression cohort of OSM were enriched in the AGE-RAGE signaling pathway, aminoacyl-tRNA biosynthesis, and the C-type lectin receptor signaling pathway, while genes in the low-expression cohort were mainly enriched in the IL-17 and PPAR signaling pathways (Fig. 9B). Genes in the high-expression cohort of NMUR1 were primarily enriched in antigen processing and presentation, glycosaminoglycan biosynthesis, and natural killer-mediated cytotoxicity pathways, whereas genes in the low-expression cohort were primarily enriched in T cell receptor and VEGF signaling pathways (Fig. 9C).

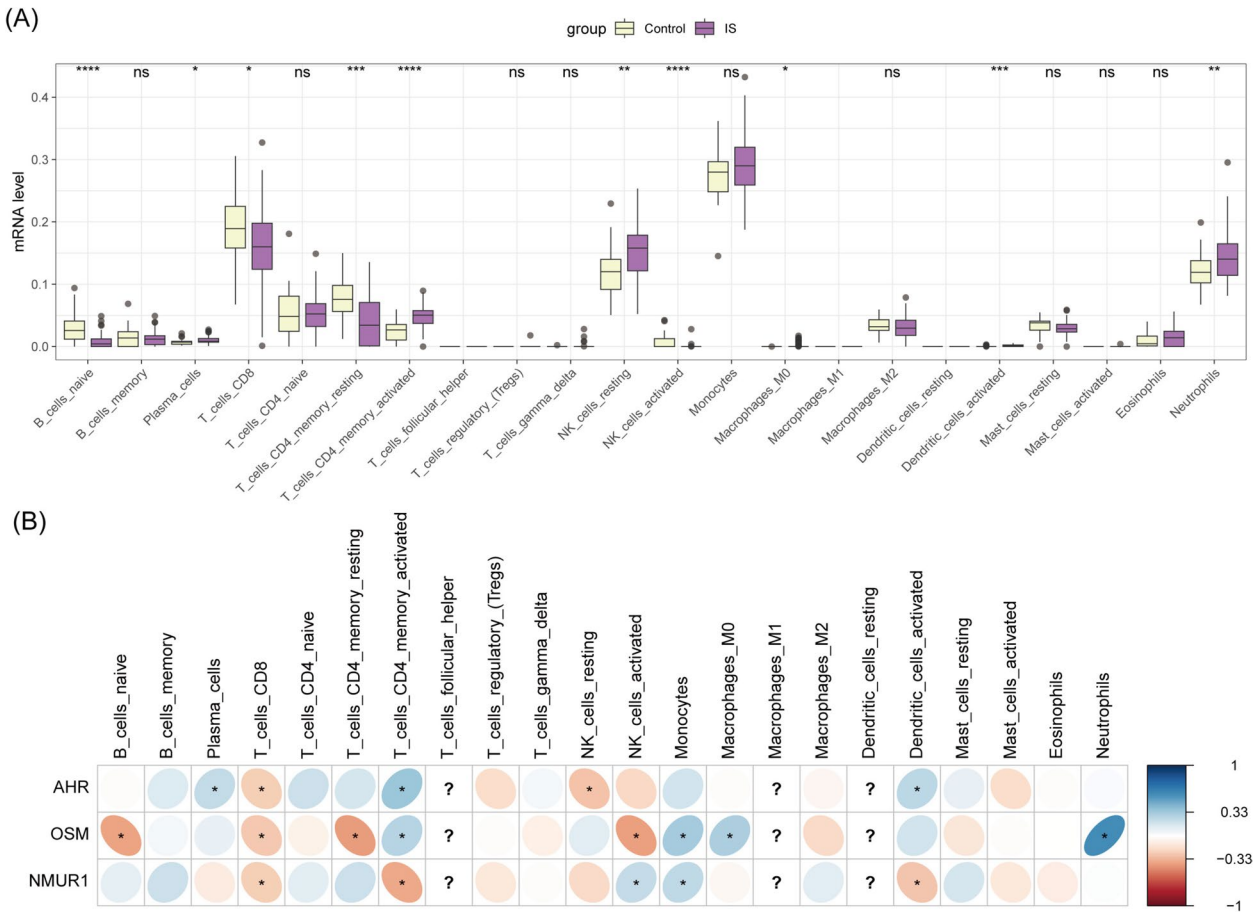

Fig. 7 Analysis of Immune Cell Infiltration. **A** Comparison of immune cell infiltration differences between ischemic stroke and control groups. **B** Heatmap showing the correlations between hub IRGs and immune cell infiltration. *ns, not significant, *p < 0.05, **p < 0.01, ***p < 0.001, ****p < 0.0001

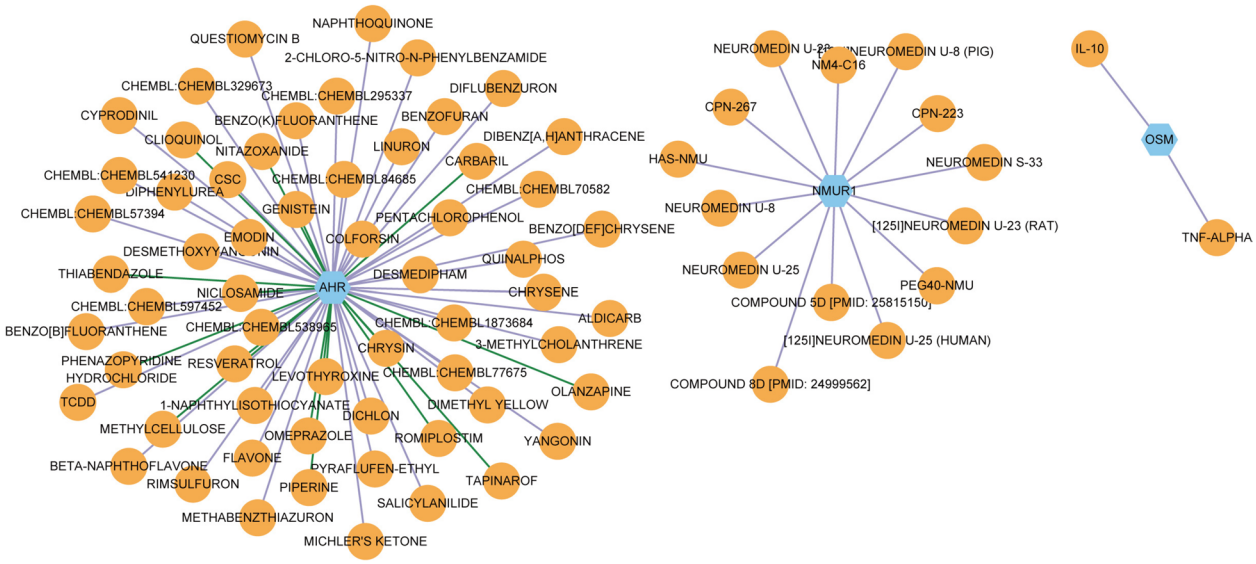

Fig. 8 Drug-Gene Networks Constructed by Cytoscape. Blue nodes represent hub genes, and green nodes represent targeted drugs. Green edges indicate approved drugs, and purple edges indicate drugs that are not approved

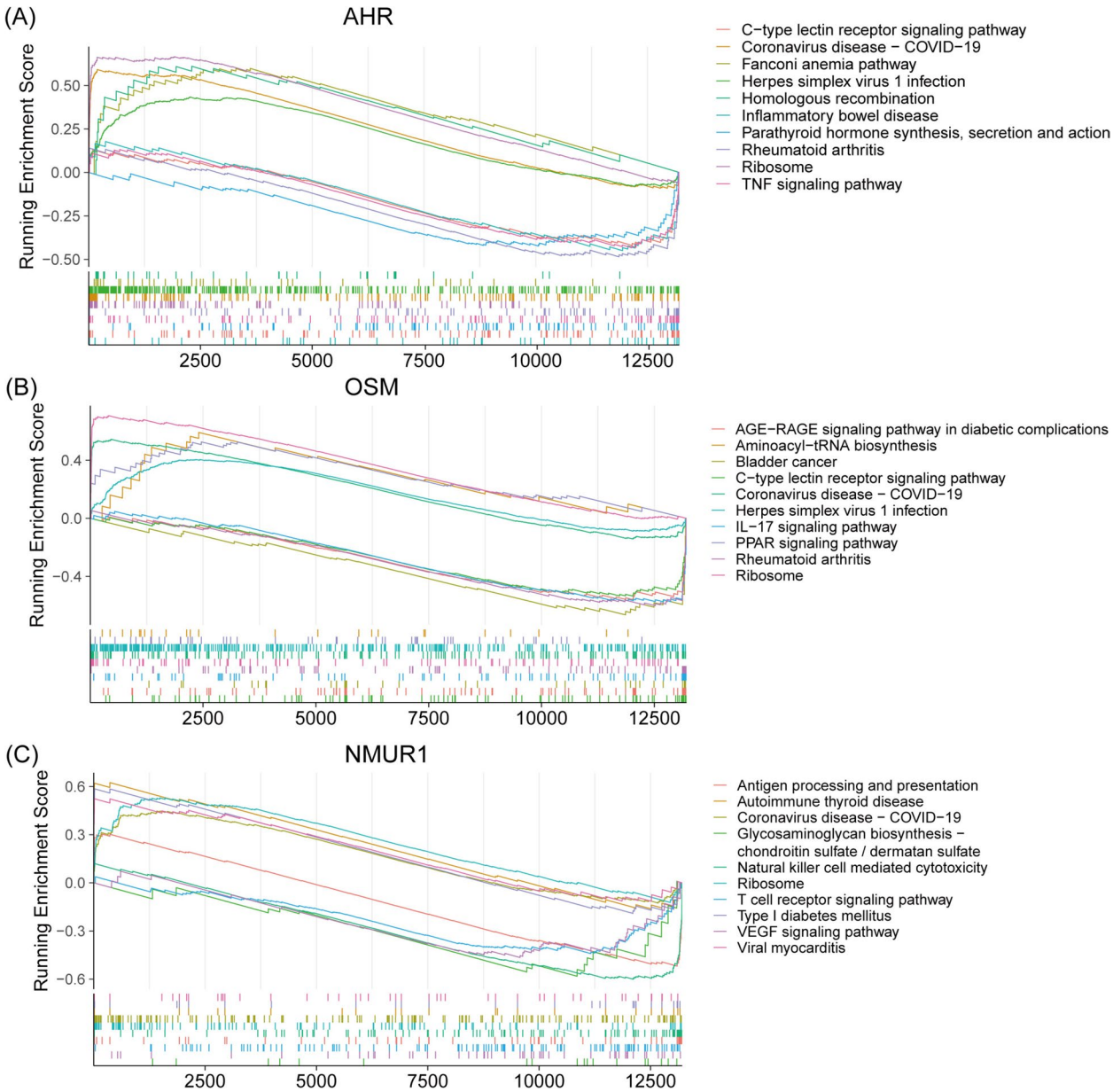

Fig. 9 GSEA Results Related to the Function of Hub Genes. **A** GSEA results for AHR. **B** GSEA results for OSM. **C** GSEA results for NMUR1

Discussion

Inflammation plays a crucial role in the pathogenesis and progression of IS [28]. Ischemia-induced cell death releases cellular contents that can activate the immune system, triggering the production of local inflammatory mediators such as cytokines, chemokines, and free radicals, which exacerbate brain tissue damage, including disruption of the blood–brain barrier and thrombosis [29, 30]. WGCNA identified a series of modules and genes related to IS, including 70 IRGs, which constitute one-third of all IRGs, further confirming the critical role of inflammation-related genes in the formation and progression of IS. In this study, we leveraged the complementary strengths of three machine learning algorithms to screen IRG features for constructing a nomogram diagnostic model for IS and identified three hub IRGs (AHR, OSM, NMUR1). The model demonstrated excellent performance in both the training and validation sets, suggesting potential clinical utility.

The AHR gene encodes the aryl hydrocarbon receptor, a ligand-activated transcription factor initially recognized for its binding capacity to environmental pollutants such

as polychlorinated biphenyls (PCBs) and dioxins. Beyond its role in environmental toxicology, increasing evidence indicates that AhR is also involved in immune system regulation and inflammatory processes. Activation of AhR regulates immune cell functions and influences inflammation-related signaling pathways, playing a key role in maintaining immunological homeostasis and controlling inflammatory responses [31]. Rzemieniec et al. found in a rat model of perinatal asphyxia that the AHR signaling pathway mediates the neuroprotective effects of 3,3'-diindolylmethane, suggesting that compounds capable of inhibiting AHR signaling are promising therapeutic tools for stroke prevention [32]. Additionally, AHR has been shown to regulate nuclear factor kappa B (NF-κB) signaling activation through tumor necrosis factor receptor-associated factor-6 (TRAF6), thereby modulating inflammatory responses [33]. Furthermore, it can protect neurons from ischemia–reperfusion injury by regulating toll-like receptor 4 (TLR4)-mediated inflammation [34].

OSM is a pleiotropic cytokine involved in various inflammatory responses, such as wound healing, liver regeneration, and bone remodeling. As a member of the IL-6 cytokine family, OSM binds to the shared receptor gp130, recruits OSMRβ or LIFRβ, and activates multiple signaling pathways, including JAK/STAT, MAPK, JNK, and PI3K/AKT [35]. Recent studies have found elevated serum OSM expression in IS patients, which correlates with stroke severity and serves as an independent predictor of poor prognosis in ischemic stroke patients. Moreover, OSM is significantly associated with several risk factors for acute IS, including age, low-density lipoprotein, non-high-density lipoprotein, prothrombin time, and systolic blood pressure [36–40]. Bone marrow mesenchymal stem cells (BMSCs) have regenerative potential in brain injury, and OSM is highly expressed in the brains of middle cerebral artery occlusion (MCAO) stroke rats, upregulating SDF-1 and promoting BMSC migration, suggesting that OSM combined with BMSC therapy can improve BMSC transplantation efficiency and neurological recovery [41]. Additionally, reduced neuronal expression of OSMRβ leads to worse stroke outcomes, while overexpression of OSMRβ in neurons has neuroprotective effects. OSM exerts neuroprotective effects by recruiting OSMRβ and activating the JAK2/STAT3 pro-survival signaling pathway. These data support the notion that human OSM may represent a promising candidate for stroke therapy [42].

NMUR1 is the receptor for the neuropeptide NMU, and NMU-NMUR1 signaling regulates inflammatory responses [43]. In the brain tissue of ischemic stroke patients and focal ischemic mice, innate lymphoid cells type 2 (ILC2s) accumulate in the peri-infarct region. Adoptive transfer and expansion of ILC2s reduce infarct size. Importantly, infiltrating ILC2s produce IL-4, reducing the severity of stroke damage [44]. Studies have shown that ILC2s specifically express NMUR1 in inflammatory environments, mediating the activation of ILC2s by neuromedin U (NMU). Loss of NMU-NMUR1 signaling reduces ILC2 frequency and effector function [45]. Therefore, further investigation of NMUR1-mediated ILC2 activation in IS is warranted.

The complex relationship between IRG gene expression and immune cell infiltration may be a critical mechanism underlying their involvement in IS. Blood-derived neutrophils and neutrophil extracellular traps (NETs) are observed in the brains of ischemic stroke patients and corresponding animal models [46, 47], with NETs impairing vascular remodeling [48]. Post-ischemia, neutrophils rapidly migrate to the damaged brain area, releasing reactive oxygen species and proteases that can disrupt the vascular endothelium and surrounding tissues. Correlation analysis reveals a significant positive association between OSM expression and neutrophil infiltration, suggesting its potential role in influencing IS outcomes via neutrophils. Neutrophils are a major source of OSM, and most OSM + neutrophils express arginase 1, indicating an N2 phenotype that plays a crucial role in repair processes and inflammation resolution [49–51]. However, these findings contradict the clinical significance of OSM in IS. Additionally, CD8 +, CD4 +, and NK cells have been implicated in IS [52–54], and given the correlations between IRGs and these cells, they may modulate the immune response in IS.

Through database searches, we identified a series of drugs targeting hub IRGs that may have potential therapeutic efficacy for IS. Chrysin has been shown to reduce brain edema after ischemic stroke [55]. It can also reduce the expression of pro-inflammatory cytokines (TNF-α and IL-10), decrease pro-apoptotic (Bax) and increase anti-apoptotic (Bcl2) proteins, further alleviating post-ischemic injury, exerting neuroprotective effects [55, 56]. Piperine, a primary active component isolated from Piper nigrum L., improves brain injury in ischemic stroke rats by modulating the PI3K/AKT/mTOR pathway [57]. Clinical studies show that supplementation with curcumin-piperine improves carotid intima-media thickness, serum hs-CRP, total cholesterol, triglycerides, total antioxidant capacity, as well as systolic and diastolic blood pressure in patients during the rehabilitation phase of IS [58]. Additionally, genistein [59] and resveratrol [60] have been shown to benefit brain damage induced by ischemic stroke.

Future clinical translation will require the establishment of a multi-omics biomarker scoring system, such as integrating AHR activity, OSM serum concentrations, and NMUR1 expression profiles for the construction of

prognostic models. At the therapeutic level, cross-pathway synergistic interventions could be explored, such as using AHR inhibitors to control early inflammatory storms, OSM-enhanced BMSC transplantation to promote tissue repair, combined with NMUR1 agonists to modulate the immune microenvironment, thereby forming a temporally precise combination therapy regimen. Furthermore, the development of portable detection devices for rapid quantification of these biomarkers would facilitate optimized treatment decision-making in emergency settings.

While this study provides valuable insights, it is important to acknowledge several limitations. Firstly, the retrospective nature of the data derived from publicly available databases may limit the representativeness and comprehensiveness of the findings, necessitating further large-scale prospective studies to validate their reliability and applicability. Secondly, while the biomarkers identified in this study were validated in multiple retrospective cohorts, clinical samples and animal models are required to confirm their clinical relevance.

Conclusion

In summary, this study systematically analyzed IRGs associated with IS and identified hub IRGs to construct a nomogram for precise diagnosis of IS. By evaluating the relationship between IRGs and immune cell infiltration in IS, we provide a novel perspective that may aid in the identification of potential therapeutic targets. This work contributes to advancing treatment strategies for cerebrovascular diseases by offering new insights into the molecular mechanisms underlying IS and suggesting potential drug candidates for targeted therapies.

Abbreviations

IS	Ischemic stroke
WGCNA	Weighted gene co-expression network analysis
IRG	Inflammation response-related genes
SVM-RFE	Support vector machine analysis with recursive feature elimination
GEO	Gene expression omnibus
TOM	Topological overlap matrix
DEGs	Differentially expressed genes
GO	Gene ontology
KEGG	Kyoto Encyclopedia of Genes and Genomes
GSEA	Gene set enrichment analysis
ROC	Receiver operating characteristic
AUC	Area under the curve
DGIdb	Drug Gene Interaction Database
BTG2	BTG anti-proliferation factor 2
AHR	Aryl hydrocarbon receptor
OSM	Oncostatin M
NMUR1	Neuromedin U receptor 1
ATP2C1	ATPase secretory pathway Ca2 + transporting 1
RGS1	Regulator of G protein signaling 1
PCB	Polychlorinated biphenyls
TRAF6	Tumor necrosis factor receptor-associated factor-6
TLR4	Toll-like receptor 4
BMSC	Bone marrow mesenchymal stem cells
MCAO	Middle cerebral artery occlusion
ILC2	Innate lymphoid cells type 2
NMU	Neuromedin U
NET	Neutrophil extracellular traps

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s40001-025-02454-1.

Additional file1

Acknowledgements

Not applicable.

Author contributions

Chenyi Huang, Dengxuan Wu, Guifen Yang contribute to the Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Resources, Software, Validation, Visualization and Writing-original draft. Chuchu Huang, Li Li contribute to Funding acquisition, Project administration, Supervision and Writing-review editing.

Funding

This research was supported by the Project of Administration of Traditional Chinese Medicine of Zhejiang Province of China under Grant No.2025ZR089.

Availability of data and materials

All data in this study are included in this article and its supplementary information files.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Received: 2 August 2024 Accepted: 10 March 2025
Published online: 22 March 2025

References

1. Feigin VL, Brainin M, Norrving B, et al. World stroke organization (wso): global stroke fact sheet 2022. Int J Stroke. 2022;17:18–29. https://doi.org/10.1177/17474930211065917.
2. Maida CD, Norrito RL, Daidone M, et al. Neuroinflammatory mechanisms in ischemic stroke: focus on cardioembolic stroke, background, and therapeutic approaches. Int J Mol Sci. 2020. https://doi.org/10.3390/ijms21186454.
3. Feigin VL, Nguyen G, Cercy K, et al. Global, regional, and country-specific lifetime risks of stroke, 1990 and 2016. N Engl J Med. 2018;379:2429–37. https://doi.org/10.1056/NEJMoa1804492.
4. Katan M, Luft A. Global burden of stroke. Semin Neurol. 2018;38:208–11. https://doi.org/10.1055/s-0038-1649503.
5. Bamford J, Sandercock P, Dennis M, et al. Classification and natural history of clinically identifiable subtypes of cerebral infarction. Lancet. 1991;337:1521–6. https://doi.org/10.1016/0140-6736(91)93206-o.
6. Zhao Y, Zhang X, Chen X, et al. Neuronal injuries in cerebral infarction and ischemic stroke: from mechanisms to treatment (review). Int J Mol Med. 2022. https://doi.org/10.3892/ijmm.2021.5070.
7. Hu S, Cai J, Chen S, et al. Identification of novel biomarkers and immune infiltration characteristics of ischemic stroke based on comprehensive bioinformatic analysis and machine learning. Biochem Biophys Rep. 2024;37: 101595. https://doi.org/10.1016/j.bbrep.2023.101595.

8. Bindal P, Kumar V, Kapil L, et al. Therapeutic management of ischemic stroke. Naunyn Schmiedebergs Arch Pharmacol. 2024;397:2651–79. https://doi.org/10.1007/s00210-023-02804-y.
9. Zhang H, Sun J, Zou P, et al. Identification of hypoxia- and immune-related biomarkers in patients with ischemic stroke. Heliyon. 2024;10: e25866. https://doi.org/10.1016/j.heliyon.2024.e25866.
10. Chinnappan J, Ramu A, V VR, et al. Integrative bioinformatics approaches to therapeutic gene target selection in various cancers for nitroglycerin. Sci Rep. 2021;11:22036. https://doi.org/10.1038/s41598-021-01508-8.
11. Simats A, Liesz A. Systemic inflammation after stroke: implications for post-stroke comorbidities. Embo Mol Med. 2022;14:e16269. https://doi.org/10.15252/emmm.202216269.
12. Kim JY, Park J, Chang JY, et al. Inflammation after ischemic stroke: the role of leukocytes and glial cells. Exp Neurobiol. 2016;25:241–51. https://doi.org/10.5607/en.2016.25.5.241.
13. Kaur D, Sharma V, Deshmukh R. Activation of microglia and astrocytes: a roadway to neuroinflammation and alzheimer's disease. Inflammopharmacology. 2019;27:663–77. https://doi.org/10.1007/s10787-019-00580-x.
14. Choi B, Lee C, Yu JW. Distinctive role of inflammation in tissue repair and regeneration. Arch Pharm Res. 2023;46:78–89. https://doi.org/10.1007/s12272-023-01428-3.
15. Lin W, Zhao XY, Cheng JW, et al. Signaling pathways in brain ischemia: mechanisms and therapeutic implications. Pharmacol Ther. 2023;251: 108541. https://doi.org/10.1016/j.pharmthera.2023.108541.
16. Lin W, Wang Y, Chen Y, et al. 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;2021:8060477. https://doi.org/10.1155/2021/8060477.
17. Zhang J, Zhang B, Li T, et al. Exploring the shared biomarkers between cardioembolic stroke and atrial fibrillation by wgcna and machine learning. Front Cardiovasc Med. 2024;11:1375768. https://doi.org/10.3389/fcvm.2024.1375768.
18. Zhang X, Wang X, Wang S, et al. Machine learning algorithms assisted identification of post-stroke depression associated biological features. Front Neurosci. 2023;17:1146620. https://doi.org/10.3389/fnins.2023.1146620.
19. Ritchie ME, Phipson B, Wu D, et al. Limma powers differential expression analyses for rna-sequencing and microarray studies. Nucleic Acids Res. 2015;43: e47. https://doi.org/10.1093/nar/gkv007.
20. Leek JT, Johnson WE, Parker HS, et al. The sva package for removing batch effects and other unwanted variation in high-throughput experiments. Bioinformatics. 2012;28:882–3. https://doi.org/10.1093/bioinformatics/bts034.
21. Langfelder P, Horvath S. Wgcna: an r package for weighted correlation network analysis. BMC Bioinform. 2008;9:559. https://doi.org/10.1186/1471-2105-9-559.
22. Yu G, Wang LG, Han Y, et al. Clusterprofiler: an r package for comparing biological themes among gene clusters. OMICS. 2012;16:284–7. https://doi.org/10.1089/omi.2011.0118.
23. Engebretsen S, Bohlin J. Statistical predictions with glmnet. Clin Epigenetics. 2019;11:123. https://doi.org/10.1186/s13148-019-0730-1.
24. Sanz H, Valim C, Vegas E, et al. Svm-rfe: selection and visualization of the most relevant features through non-linear kernels. BMC Bioinform. 2018;19:432. https://doi.org/10.1186/s12859-018-2451-4.
25. Ahn S, Lee SE, Kim MH. Random-forest model for drug-target interaction prediction via kullbeck-leibler divergence. J Cheminform. 2022;14:67. https://doi.org/10.1186/s13321-022-00644-1.
26. Zeng D, Ye Z, Shen R, et al. Iobr: multi-omics immuno-oncology biological research to decode tumor microenvironment and signatures. Front Immunol. 2021;12: 687975. https://doi.org/10.3389/fimmu.2021.687975.
27. Cotto KC, Wagner AH, Feng YY, et al. Dgidb 3.0: a redesign and expansion of the drug-gene interaction database. Nucleic Acids Res. 2018;46:D1068–73. https://doi.org/10.1093/nar/gkx1143.
28. DeLong JH, Ohashi SN, O'Connor KC, et al. Inflammatory responses after ischemic stroke. Semin Immunopathol. 2022;44:625–48. https://doi.org/10.1007/s00281-022-00943-7.
29. Candelario-Jalil E, Dijkhuizen RM, Magnus T. Neuroinflammation, stroke, blood-brain barrier dysfunction, and imaging modalities. Stroke. 2022;53:1473–86. https://doi.org/10.1161/STROKEAHA.122.036946.
30. Maïer B, Tsai AS, Einhaus JF, et al. Neuroimaging is the new "spatial omic": multi-omic approaches to neuro-inflammation and immuno-thrombosis in acute ischemic stroke. Semin Immunopathol. 2023;45:125–43. https://doi.org/10.1007/s00281-023-00984-6.
31. Neavin DR, Liu D, Ray B, et al. The role of the aryl hydrocarbon receptor (ahr) in immune and inflammatory diseases. Int J Mol Sci. 2018. https://doi.org/10.3390/ijms19123851.
32. Rzemieniec J, Bratek E, Wnuk A, et al. Neuroprotective effect of 3,3'-diindolylmethane against perinatal asphyxia involves inhibition of the ahr and nmda signaling and hypermethylation of specific genes. Apoptosis. 2020;25:747–62. https://doi.org/10.1007/s10495-020-01631-3.
33. Li L, He G, Shi M, et al. Edaravone dexborneol ameliorates cognitive impairment by regulating the nf-κb pathway through ahr and promoting microglial polarization towards the m2 phenotype in mice with bilateral carotid artery stenosis (bcas). Eur J Pharmacol. 2023;957: 176036. https://doi.org/10.1016/j.ejphar.2023.176036.
34. Liu J, Mu D, Xu J, et al. Inhibition of tlr4 signaling by isorhapontigenin targeting of the ahr alleviates cerebral ischemia/reperfusion injury. J Agric Food Chem. 2023;71:13270–83. https://doi.org/10.1021/acs.jafc.3c00152.
35. Wolf CL, Pruett C, Lighter D, et al. The clinical relevance of osm in inflammatory diseases: a comprehensive review. Front Immunol. 2023;14:1239732. https://doi.org/10.3389/fimmu.2023.1239732.
36. Angerfors A, Brännmark C, Lagging C, et al. Proteomic profiling identifies novel inflammation-related plasma proteins associated with ischemic stroke outcome. J Neuroinflammation. 2023;20:224. https://doi.org/10.1186/s12974-023-02912-9.
37. Christian M, Long B, Tian Z, et al. Correlation between oncostatin m and acute ischemic stroke: a case-control study. Cureus. 2023;15: e50297. https://doi.org/10.7759/cureus.50297.
38. Hazelwood HS, Frank JA, Maglinger B, et al. Plasma protein alterations during human large vessel stroke: a controlled comparison study. Neurochem Int. 2022;160: 105421. https://doi.org/10.1016/j.neuint.2022.105421.
39. Stanne TM, Angerfors A, Andersson B, et al. Longitudinal study reveals long-term proinflammatory proteomic signature after ischemic stroke across subtypes. Stroke. 2022;53:2847–58. https://doi.org/10.1161/STROKEAHA.121.038349.
40. Abraira L, López-Maza S, Quintana M, et al. Exploratory study of blood biomarkers in patients with post-stroke epilepsy. Eur Stroke J. 2024. https://doi.org/10.1177/23969873241244584.
41. Han J, Feng Z, Xie Y, et al. Oncostatin m-induced upregulation of sdf-1 improves bone marrow stromal cell migration in a rat middle cerebral artery occlusion stroke model. Exp Neurol. 2019;313:49–59. https://doi.org/10.1016/j.expneurol.2018.09.005.
42. Guo S, Li ZZ, Gong J, et al. Oncostatin m confers neuroprotection against ischemic stroke. J Neurosci. 2015;35:12047–62. https://doi.org/10.1523/JNEUROSCI.1800-15.2015.
43. Moriyama M, Fukuyama S, Inoue H, et al. The neuropeptide neuromedin u activates eosinophils and is involved in allergen-induced eosinophilia. Am J Physiol Lung Cell Mol Physiol. 2006;290:L971–7. https://doi.org/10.1152/ajplung.00345.2005.
44. Zheng P, Xiu Y, Chen Z, et al. Group 2 innate lymphoid cells resolve neuroinflammation following cerebral ischaemia. Stroke Vasc Neurol. 2023;8:424–34. https://doi.org/10.1136/svn-2022-001919.
45. Wallrapp A, Riesenfeld SJ, Burkett PR, et al. The neuropeptide nmu amplifies ilc2-driven allergic lung inflammation. Nature. 2017;549:351–6. https://doi.org/10.1038/nature24029.
46. Laridan E, Denorme F, Desender L, et al. Neutrophil extracellular traps in ischemic stroke thrombi. Ann Neurol. 2017;82:223–32. https://doi.org/10.1002/ana.24993.
47. Perez-de-Puig I, Miró-Mur F, Ferrer-Ferrer M, et al. Neutrophil recruitment to the brain in mouse and human ischemic stroke. Acta Neuropathol. 2015;129:239–57. https://doi.org/10.1007/s00401-014-1381-0.
48. Kang L, Yu H, Yang X, et al. Neutrophil extracellular traps released by neutrophils impair revascularization and vascular remodeling after stroke. Nat Commun. 2020;11:2488. https://doi.org/10.1038/s41467-020-16191-y.
49. Cuartero MI, Ballesteros I, Moraga A, et al. N2 neutrophils, novel players in brain inflammation after stroke: modulation by the ppary agonist rosiglitazone. Stroke. 2013;44:3498–508. https://doi.org/10.1161/STROKEAHA.113.002470.
50. Ma Y, Yabluchanskiy A, Iyer RP, et al. Temporal neutrophil polarization following myocardial infarction. Cardiovasc Res. 2016;110:51–61. https://doi.org/10.1093/cvr/cvw024.

51. Mishalian I, Bayuh R, Levy L, et al. Tumor-associated neutrophils (tan) develop pro-tumorigenic properties during tumor progression. Cancer Immunol Immunother. 2013;62:1745–56. https://doi.org/10.1007/s00262-013-1476-9.
52. Rolfes L, Ruck T, David C, et al. Natural killer cells are present in rag1(-/-) mice and promote tissue damage during the acute phase of ischemic stroke. Transl Stroke Res. 2022;13:197–211. https://doi.org/10.1007/s12975-021-00923-3.
53. Wang YR, Cui WQ, Wu HY, et al. The role of t cells in acute ischemic stroke. Brain Res Bull. 2023;196:20–33. https://doi.org/10.1016/j.brainresbull.2023.03.005.
54. Zhang Z, Duan Z, Cui Y. Cd8(+) t cells in brain injury and neurodegeneration. Front Cell Neurosci. 2023;17:1281763. https://doi.org/10.3389/fncel.2023.1281763.
55. Rashno M, Ghaderi S, Nesari A, et al. Chrysin attenuates traumatic brain injury-induced recognition memory decline, and anxiety/depression-like behaviors in rats: insights into underlying mechanisms. Psychopharmacology. 2020;237:1607–19. https://doi.org/10.1007/s00213-020-05482-3.
56. Li TF, Ma J, Han XW, et al. Chrysin ameliorates cerebral ischemia/reperfusion (i/r) injury in rats by regulating the pi3k/akt/mtor pathway. Neurochem Int. 2019;129: 104496. https://doi.org/10.1016/j.neuint.2019.104496.
57. Zhang Y, Yang M, Yuan Q, et al. Piperine ameliorates ischemic stroke-induced brain injury in rats by regulating the pi3k/akt/mtor pathway. J Ethnopharmacol. 2022;295: 115309. https://doi.org/10.1016/j.jep.2022.115309.
58. Boshagh K, Khorvash F, Sahebkar A, et al. The effects of curcumin-piperine supplementation on inflammatory, oxidative stress and metabolic indices in patients with ischemic stroke in the rehabilitation phase: a randomized controlled trial. Nutr J. 2023;22:69. https://doi.org/10.1186/s12937-023-00905-1.
59. Nabavi SF, Daglia M, Tundis R, et al. Genistein: a boon for mitigating ischemic stroke. Curr Top Med Chem. 2015;15:1714–21. https://doi.org/10.2174/1568026615666150427122709.
60. Hou Y, Wang K, Wan W, et al. Resveratrol provides neuroprotection by regulating the jak2/stat3/pi3k/akt/mtor pathway after stroke in rats. Genes Dis. 2018;5:245–55. https://doi.org/10.1016/j.gendis.2018.06.001.

Publisher's Note