Check for updates

*CORRESPONDENCE
Ze-Hua He
✉ hezehua2000@126.com
Qing-Feng Chen
✉ qingfeng@gxu.edu.cn

Identification of ribosome biogenesis genes and subgroups in ischaemic stroke

Xi Wang[1], Xiao-Yu Zhang[2], Nan-Qing Liao[1], Ze-Hua He[3]*
and Qing-Feng Chen[4]*

[1]School of Medicine, Guangxi University, Nanning, China, [2]The College of Life Sciences, Northwest
University, Xian, China, [3]Department of General Surgery, Guangxi Hospital Division of The First
Affiliated Hospital, Sun Yat-sen University, Nanning, China, [4]School of Computer, Electronics and
Information, Guangxi University, Nanning, China

Background: Ischaemic stroke is a leading cause of death and severe disability worldwide. Given the importance of protein synthesis in the inflammatory response and neuronal repair and regeneration after stroke, and that proteins are acquired by ribosomal translation of mRNA, it has been theorised that ribosome biogenesis may have an impact on promoting and facilitating recovery after stroke. However, the relationship between stroke and ribosome biogenesis has not been investigated.

Methods: In the present study, a ribosome biogenesis gene signature (RSG) was developed using Cox and least absolute shrinkage and selection operator (LASSO) analysis. We classified ischaemic stroke patients into high-risk and low-risk groups using the obtained relevant genes, and further elucidated the immune infiltration of the disease using ssGSEA, which clarified the close relationship between ischaemic stroke and immune subgroups. The concentration of related proteins in the serum of stroke patients was determined by ELISA, and the patients were divided into groups to evaluate the effect of the ribosome biogenesis gene on patients. Through bioinformatics analysis, we identified potential IS-RSGs and explored future therapeutic targets, thereby facilitating the development of more effective therapeutic strategies and novel drugs against potential therapeutic targets in ischaemic stroke.

Results: We obtained a set of 12 ribosome biogenesis-related genes (EXOSC5, MRPS11, MRPS7, RNASEL, RPF1, RPS28, C1QBP, GAR1, GRWD1, PELP1, UTP, ERI3), which play a key role in assessing the prognostic risk of ischaemic stroke. Importantly, risk grouping using ribosome biogenesis-related genes was also closely associated with important signaling pathways in stroke. ELISA detected the expression of C1QBP, RPS28 and RNASEL proteins in stroke patients, and the proportion of neutrophils was significantly increased in the high-risk group.

Abbreviations: GEO, Gene Expression Omnibus; RSG, Ribosome Biogenesis Gene; LASSO, least absolute shrinkage and selection operator; GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; DEGs, differentially expressed genes; ELISA, enzyme-linked immunosorbent assay.

Conclusions: The present study demonstrates the involvement of ribosomal biogenesis genes in the pathogenesis of ischaemic stroke, providing novel insights into the underlying pathogenic mechanisms and potential therapeutic strategies for ischaemic stroke.

Introduction

Stroke is a group of diseases with symptoms of ischemic and hemorrhagic damage to the brain as the main clinical manifestation, including ischemic stroke and hemorrhagic stroke (1). Among them, ischemic stroke (IS) accounts for about 80% of all stroke cases, occurring in middle-aged and elderly people, and is characterized by high mortality and disability rates. Its treatment and management have been viewed as a serious medical and public health challenge (2).

Ischemic stroke is primarily caused by a sudden disruption of blood flow due to thrombosis or embolism. The IS pathological process involves a complex temporal and spatial cascade of responses and is the result of multiple cellular pathways. Ischemia caused by stroke restricts blood flow to specific regions of the brain, inducing a series of pathological reactions that culminate in the infiltration of immune cells (3, 4). Furthermore, the immune microenvironment and inflammatory response play a crucial role in the development of IS and are closely linked to the severity and prognosis of the condition (5). Thus, the discovery of new therapeutic targets for IS can be facilitated by gaining a deeper understanding of stroke pathogenesis and immune microenvironment changes, which may offer fresh prospects and avenues for exploration. In general, the standard pathological change after the onset of IS is the disruption of the blood-brain barrier, leading to the gathering of various infiltrating immune cells, such as T cells, B cells, neutrophils, dendritic cells, and macrophages, in the edema area (6). These immune cells play a dual role (7). Also, immune cells release a significant number of inflammatory cytokines, which can promote secondary neuroinflammation. Many immune cells remove necrotic cell debris, reduce the inflammatory response and play a protective role in repairing the blood-brain barrier and promoting angiogenesis (8). Immunomodulation after stroke has therefore become an important area of research and therapy, aiming to reduce damage and promote recovery by controlling the intensity and nature of the immune response.

Ribosome biogenesis, also known as RiboSis, is a highly intricate process responsible for producing the ribosomes necessary for protein synthesis during various cellular processes such as proliferation, differentiation, apoptosis, development, and transformation (9–12). In stroke, especially in ischaemic stroke, changes in RiboSis may have an important impact on cell fate and disease progression. Further, after cerebral ischaemia, cells undergo metabolic reprogramming to adapt to changes in energy demand (13). This metabolic adjustment affects multiple steps in ribosome biogenesis. Ribosomes are a common component of autophagy. After a stroke, cells may break down existing ribosomes and initiate autophagy mechanisms to recycle damaged cellular components and gain energy (14, 15). Ribosomes are not only the site of conventional protein synthesis, but are also involved in the synthesis of certain neuroprotective proteins such as growth factors and antioxidant proteins. For example, certain neuroprotective strategies may enhance neuronal resistance to ischaemic injury or accelerate its recovery by promoting the expression of specific ribosomal proteins or modulating certain steps in ribosome biogenesis. Further, as mentioned earlier, the inflammatory response to stroke involves the synthesis of a large number of immune-related proteins, and their production is directly dependent on ribosome function. Precision medicine has become the main development trend in the clinical treatment of many chronic diseases, especially the accurate classification and hierarchical management of patients. Given the significance of Ribosis in the cellular response to stroke, to construct a classification model of Ribosis in the cellular response to stroke, and to study its regulatory mechanism may find new therapeutic methods. Previous studies have reported that ribosomal biogenesis plays a central role in cancer (16, 17). Ribosome biogenetic risk scores have been reported in cancer (18). To date, few classification models constructed with ribosome biogenesis-related genes have been applied to stroke research.

As a member of the predictive modelling and data analysis toolbox:lasso regression, this technique is appreciated for its accuracy and versatility in various fields, and is also widely used in medicine (19). As part of linear regression, Lasso regression addresses the infamous "curse of dimensionality", where the number of predictors exceeds the observed value (20, 21). It contains penalties for large coefficients to simplify the model and keep it comprehensibility (21). It has been used to construct classification models in diseases such as cervical cancer and breast cancer (22–24). Our study can be briefly described by Figure 1, in which we constructed a high-risk group model with genes related to ribosomal biogenesis and assessed the immune infiltration of IS patients by the CIBERSORT method. We also analysed the characteristics of observed immune cell subtype specificity by

FIGURE 1
Research flowchart.
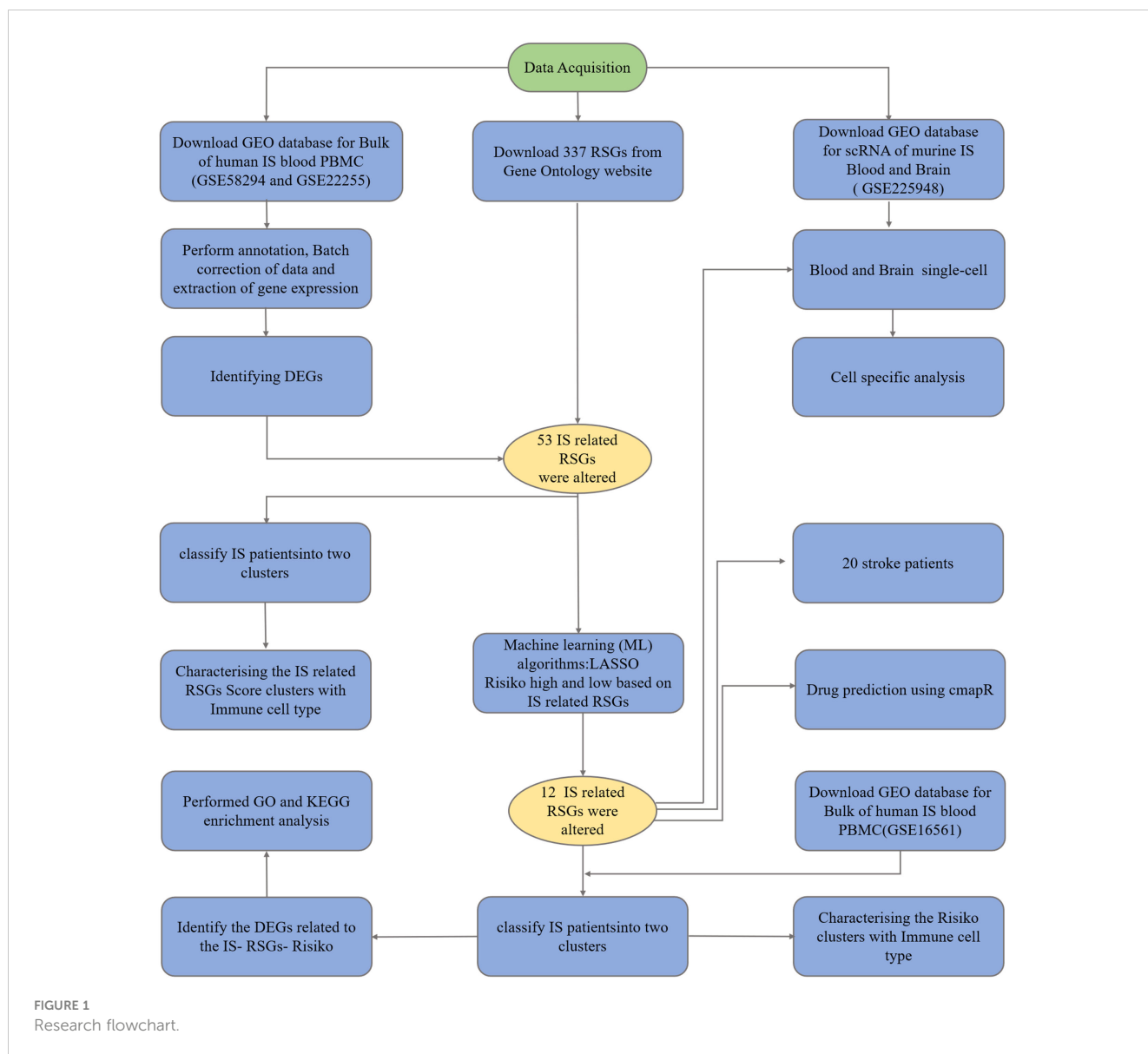
scRNA sequencing. In conclusion, we used ribosome biogenesis-related genes to classify and observe immune cell infiltration in stroke patients. The proposed strategy may provide further ideas for future clinical studies on prevention and treatment of high and low risk groups. It is expected to provide a unique perspective for exploring new ways of stroke prevention or treatment.

Materials and methods

Datasets and quality control

The RNA-seq gene expression files was obtained from Gene Expression Omnibus (GEO) database, which can be accessed at (GEO, https://www.ncbi.nlm.nih.gov/geo/) GSE58294 (25) (containing 20 Control and 69 Stroke samples), GSE22255 (26) (containing 20 Control and 20 Stroke samples) GSE16561 (27) (containing 24 Control and 39 Stroke samples).

All statistical analyses and graphical representations in this study were conducted using the R 4.2.2 software. During the data preprocessing stage, probes without corresponding genes were removed, and the average was taken when multiple probes corresponded to one gene. Standardization of the data was performed, followed by merging the two series. The limma package in R was utilized to standardize the data and eliminate the batch effect. It is important to note that this research adhered to the data access rules and release principles of the database in question.

Ribosome biogenetic genes related gene collection

We obtained the seven ribosome biological genetic related pathways from the GO database (https://www.geneontology.org/) (RIBOSOME BIOGENESIS, RRNA TRANSCRIPTION, obsolete cleavage involved in rRNA processing, RRNA MODIFICATION,

RIBOSOME ASSEMBLY,RIBOSOMAL SMALL SUBUNIT EXPORT FROM NUCLEUS, RIBOSOMAL LARGE SUBUNIT EXPORT FROM NUCLEUS). 337 ribosome biogenetic genes were eventually sorted out after removing duplicate genes.

Identification of IS-RSGs in stroke

The DESeq2 package was utilized to identify differentially expressed genes (DEGs) between control samples and stroke samples based on the following criteria: log2(Fold-Change) > 0.232 and adjusted p-value < 0.001. Heatmaps and volcano plots of the differential analysis results were generated using the pheatmap and ggplot2 packages, respectively. The analysis involved the examination of the intersection of differentially expressed genes and RSGs. Subsequently, ssGSEA scores were conducted using GSVA packaging, and patients were classified into high and low score groups based on the median score.

Estimation of immune cells

The CIBERSORT algorithm (Stanford University's <https://cibersortx.stanford.edu/>) was employed to determine the proportion of 22 immune cell types in GC samples, using the relative mode and 1000 permutations. Specifically, we utilized CIBERSORT R script v1.04 for the immune infiltration analysis of the samples. The comparison between the composition of immune cells in the high-score and low-score ischemic stroke groups was made. Additionally, the comparison between the immune cells from ischemic stroke patients in the high-risk and low-risk groups was also carried out. The box diagram shows the results obtained through the ggplot2 package in R.

Gene ontology and Kyoto encyclopedia of genes and genomes analysis

Functional enrichment analysis, including biological processes (BP), cellular compositions (CC), and molecular functions (MF), was conducted for IS-RSGs characteristics using the clusterProfiler package in R. We utilized this package to examine the functional enrichment of the clusterProfiler in KEGG pathways.

Machine learning to identify RSG biomarkers

Looking at the intersection of differential genes and RSGs. The best prognostic model was obtained by utilizing the lasso Cox technique in the glmnet package with a ten-fold cross-validation set. The prognostic risk model was determined by multiplying all risk factors and factor-related coefficients. The patient cohort was then divided into high-risk and low-risk groups based on the optimal cut-off value calculated by the maxstat package. coef is in Supplementary Table 1.

$$RiskScore = \sum_{i=1}^{N}[G_i * coef_i]$$

G = Gene expression, N = 11

Single-cell data download and analysis

GSE225948 and conduct the analysis using the Seurat (28). Principal component analysis (PCA) with 10 principal components (PCs) selected and visualized through Uniform Manifold Approximation and Projection (UMAP). The expression of IS-RSGs gene in different clusters was analysed by cluster dimensionality reduction, and different subsets of cells were labelled by their unique marker genes.

RNASEL,RPS28 and C1QBP1 protein expression by ELISA

100 μL of serum was added into the micropores pre-coated with the trapped antibodies (RNASEL,RPS28 and C1QBP1), covered with the sealing plate membrane and incubated for 1 hour at 37°C. The serum was then discarded, biotinylated antibody was added to 100 μL per well, and the sealing plate was covered at 37°C for 1 hour. After the liquid is discarded, add 300 μL 1X washing solution to each well and let it stand for 1 minute. Then shake off the washing solution and pat dry on absorbent paper. Repeat this for 3 times. Add 100 μL enzyme conjugate working liquid, cover the sealing plate and incubate for 30 minutes at 37°C After discarding the liquid, the plate was washed for 5 times, and 90 μL substrate was added to the plate for 15 minutes at 37°C away from light. After taking out the specimen plate, 50 μL was added to each well, and the OD value was measured at 450nm wavelength immediately on the Thermo Fischer Multiscan Go Reader. The concentrations of RNASEL,RPS28 and C1QBP1 in the sample were determined by the standard curve of recombinant protein production.

Statistical analysis and visualization of results: R packages and significance assessment

This study carried out statistical analyses and generated visual plots using R (version 4.2.2), with the help of the ggplot2 package for visualization purposes. (version 3.3.6). Additionally, Seurat analysis was utilized with the Seurat package (version 5.0.1), while UMAP (Uniform Manifold Approximation and Projection) analysis utilized the umap package (version 0.2.7.0). PCA (Principal Component Analysis) analysis was conducted using the R package stats (version 3.6.0). The Wilcoxon rank sum test was used to evaluate discrepancies between groups, with a statistical significance level set at p < 0.05. Patients' baseline data were analysed using Fisher's exact test. The p-values associated with this test were *p < 0.05, **p < 0.01, ***p < 0.001, and ****p < 0.0001.
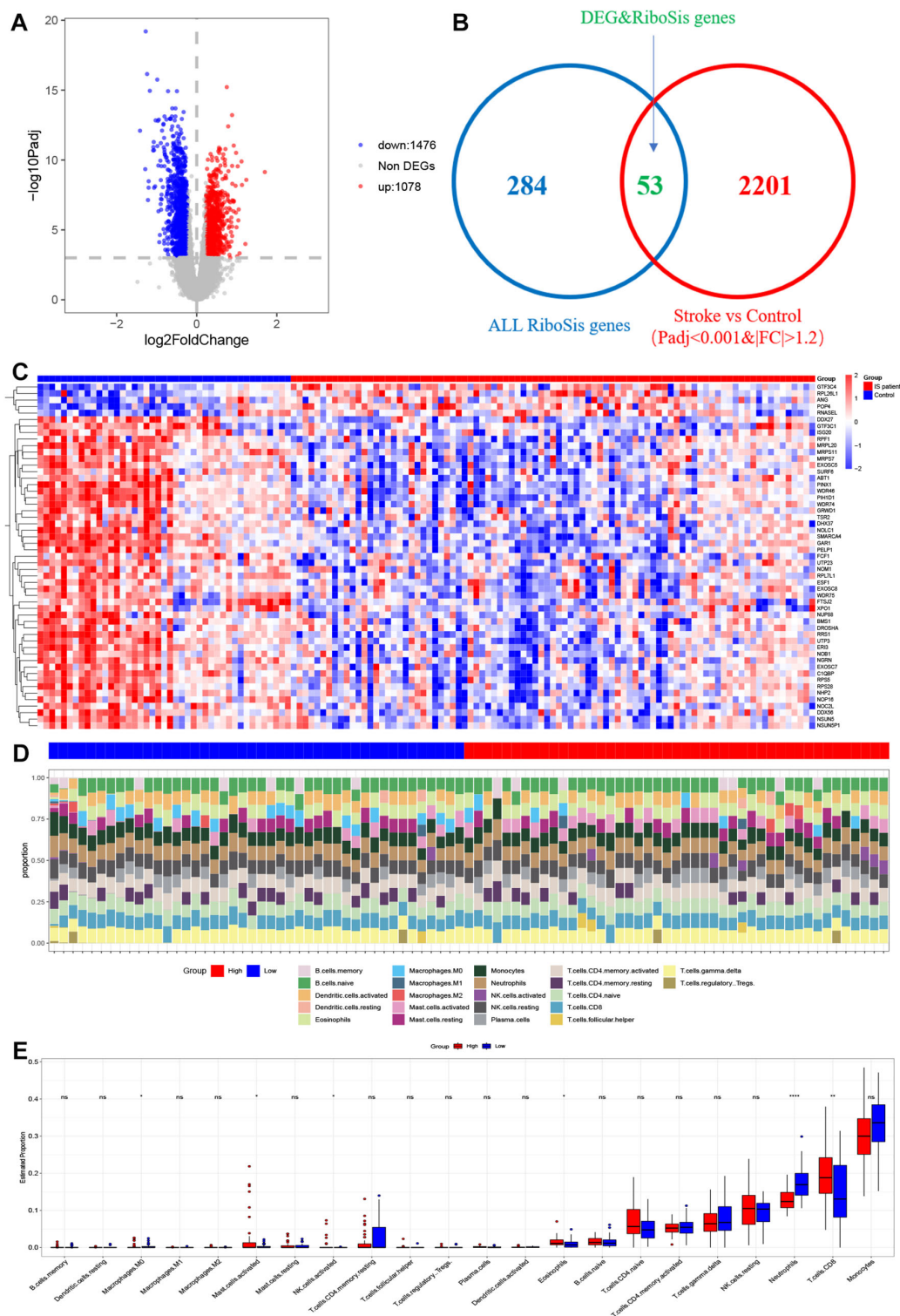
FIGURE 2

Visualization of deg and DEGS & RSGS in data sets and immune cell infiltration in IS high-score and IS low-score groups **(A)** Volcano plot of DEGs in GSE58294 and GSE16561. **(B)** Venn diagram of common genes between DEGs and ribosome biogenesis genes. **(C)** Heatmap for DEGs&RSGs in Stroke dataset. Red represents high gene expression and blue represents low expression. **(D)** The relative content of 22 kinds of immune cells high score stroke patient group and low score stroke patient group were showed in the histograms. **(E)** The BarPlot illustrated the difference in immune cell infiltration between high score stroke patient group and low score stroke patient group. ns(no significance) p≥0.05; * p< 0.05; ** p< 0.01; **** p< 0.0001.

Results

Biogenetic identification of differences in ribosome expression between Ischemic stroke and normal subjects and assessment of subgroup of immune cell infiltration

Ribosome biogenesis is the process by which ribosomes are produced and plays an important role in cell proliferation, differentiation, apoptosis, development and transformation.

Therefore, it is urgent to know which ribosome biogenesis genes are differentially expressed between Ischemic stroke and normal samples. We obtained 89 Ischemic stroke samples and 40 normal controls after normalizing the data derived from 2 published Ischemic strokes. We performed differential gene counting on the expression profiles and obtained 2,554 differential genes, including 1,078 up-regulated genes and 1,476 down-regulated genes. DEGs results are presented in the form of volcano plots (Figure 2A). By intersecting 2554 DEGs with 337 ribosomal biogenesis genes obtained in GO, 53 key genes were finally identified (Figure 2B). We found in the heat map (Figure 2C) that most of these genes were highly expressed in normal tissues. We utilized the
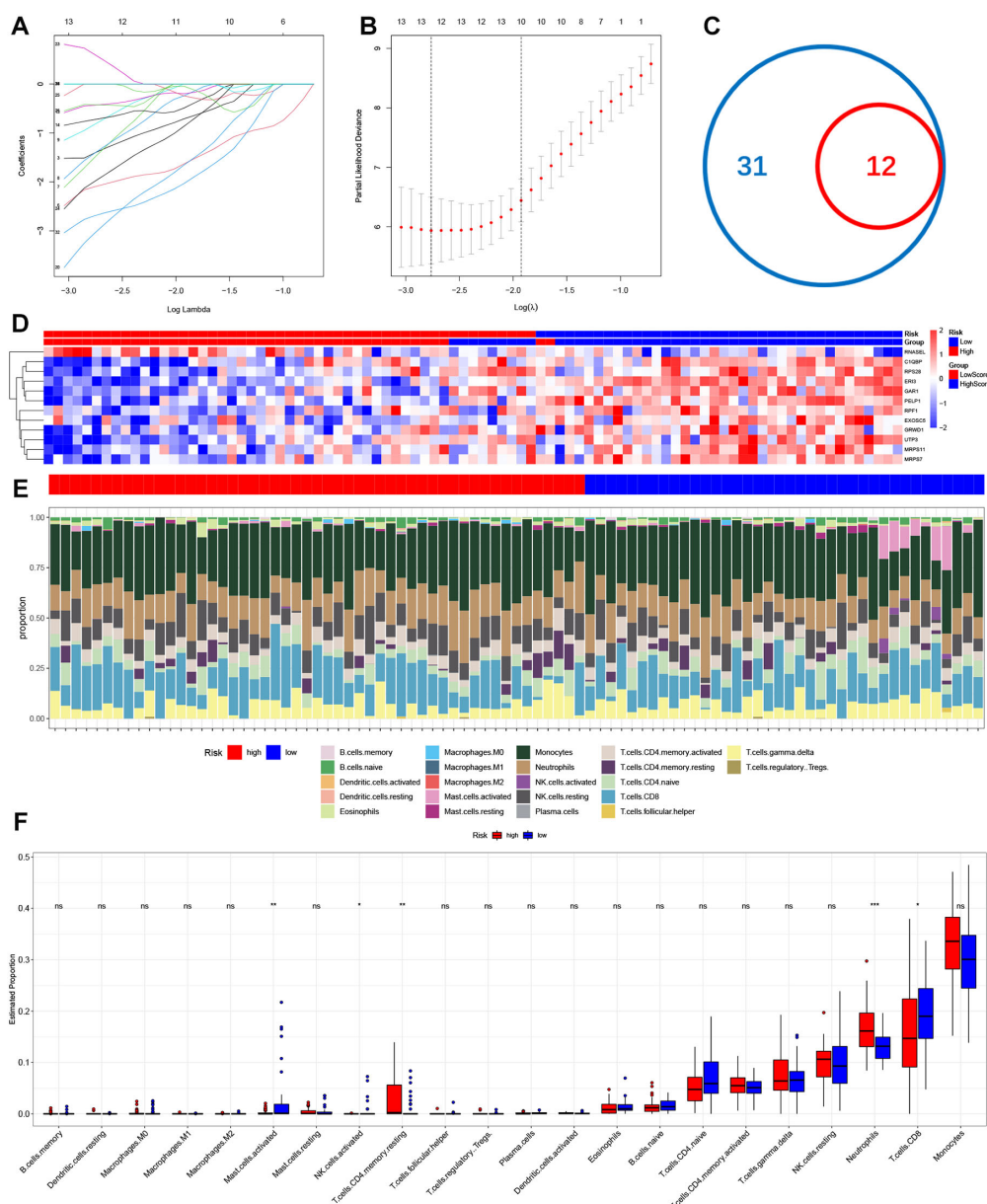
FIGURE 3

Construction of ribosome biogenic risk signal and infiltration of immune cells in high-risk groups and low-risk groups. **(A, B)** The process of Lasso regression analysis. **(C)** Venn diagram of common genes between DEGs&RSGs and hub genes. **(D)** Heatmap for hub genes in Stroke dataset. **(E)** The relative content of 22 kinds of immune cells high risk stroke patient group and low risk stroke patient group were showed in the histograms. **(F)** The BarPlot illustrated the difference in immune cell infiltration between high risk stroke patient group and low risk stroke patient group. ns(no significance) p≥0.05; * p< 0.05; ** p< 0.01; *** p< 0.001.

ssgsea algorithm for immune infiltration analysis to explore differences in immune system between low- and high-score subgroups, revealing disparities in the proportions of 22 infiltrating immune cell types. (Figure 2D). The differences in immune cell content between Ischemic stroke and healthy individuals were further explored (Figure 2E). The proportion of neutrophils was found to be significantly higher in the low group than in the high group, whereas the proportion of Eosinophils, CD8$^+$T cells, NK cells activated, Mast cells activated and Macrophages M0 were higher in the high-score group.

Identification of RSG signalling constructs and immune cell infiltration in high- and low-risk populations by machine learning

To replace our scoring model with fewer genes, we used LASSO Cox regression to build a ribosome biogenesis genes-related feature (RSG) named ribosome biogenesis genes score (Figures 3A, B). We ultimately selected 12 genes to calculate the score for each individual (Figure 3D) and observed the expression of these 12 genes in the
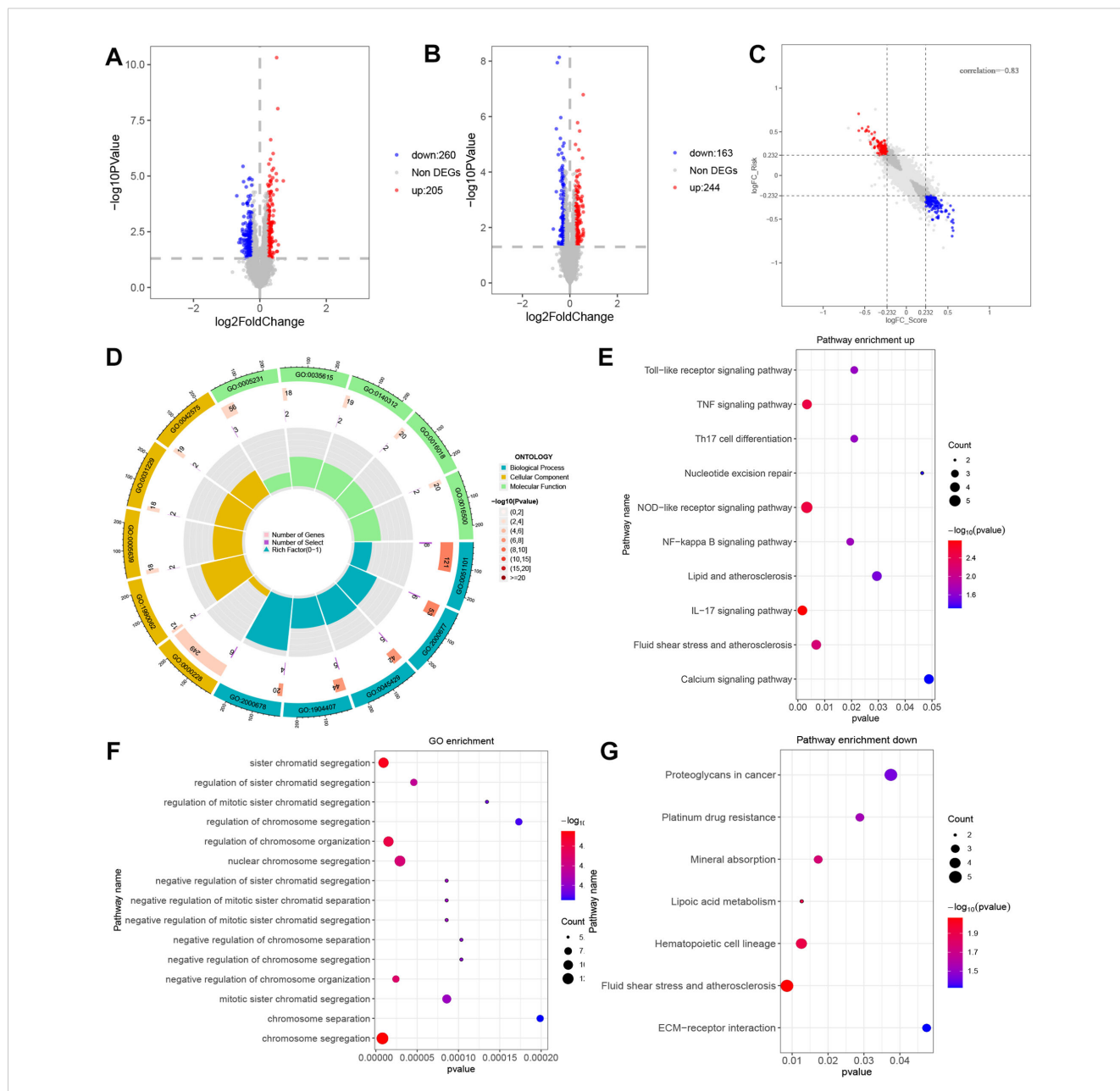
FIGURE 4
Differential expression analysis and functional enrichment analysis between RiskTypes. **(A)** Volcano plot of differentially expressed genes between high- and low-risk groups; the red dots represent upregulated genes in the high-risk group. **(B)** Volcano plot of differentially expressed genes between high- and low-score groups; the red dots represent upregulated genes in the high-risk group. **(C)** Scatter plot of 9-quadrant associate analyses of DEGs from log2 FC Scoretypes and RiskTypes. **(D)** GO functional enrichment analysis of up DEGs. **(E)** KEGG pathway analysis of up DEGs. **(F)** GO functional enrichment analysis of down DEGs. **(G)** KEGG pathway analysis of down DEGs.

Ischemic stroke samples (Figure 3D). Subsequently, we calculated IS-RSGs scores for each individual, allowing us to categorize all individuals into high or low risk categories based on median IS-RSGs scores. Ssgsea algorithm is applied to examine immune infiltration, and our findings revealed that the percentage of 22 distinct immune cell types varied between high-risk and low-risk groups (Figure 3E). The individuals in the high-risk group exhibited a significantly higher ratio of neutrophils to T-cell CD4 memory cells compared to those in the low-risk group. Conversely, the low-risk group showed a higher proportion of activated CD8+T cells, NK cells, and Mast cells (Figure 3F). At the same time, we downloaded GSE16561 for grouping and divided them into two sub-groups using our scoring method. We also used the ssgsea algorithm to analyse the proportion difference of 22 infiltrating immune cell types (Supplementary Figure 1A). To further investigate the difference of immune cell content between ischemic stroke and healthy individuals (Supplementary Figure 1B). It was also found that the proportion of neutrophils in the low-risk group was significantly higher than that in the high-risk group.

Differential expression analysis and functional enrichment analysis between risk types

To further understand the differences between these two subgroups, our analysis involved conducting a differential expression study, which resulted in the identification of 465 DEGs, comprising 260 up-regulated and 205 down-regulated genes. The distribution of these genes is depicted in the volcano plot (Figure 4A). Compared with the differential genes obtained from the previous high and low DEG groupings (Figure 4B) had a significant correlation coefficient of -0.83 (Figure 4C). The next step was to analyse the DEGs for their GO functional annotation and KEGG pathway enrichment. The GO analysis (BP) revealed gene enrichment in regulation of DNA binding, regulation of transcription regulatory region DNA binding, positive regulation of nitric oxide biosynthetic process, positive regulation of nitric oxide metabolic process, negative regulation of transcription regulatory region DNA binding (CC) shows gene clustering in nuclear chromosome,RPAP3/R2TP/prefoldin-like complex, integral component of nuclear inner membrane, intrinsic component of nuclear inner membrane, DNA polymerase complex (MF) shows excitatory extracellular ligand-gated ion channel activity, clathrin adaptor activity, cargo adaptor activity, cyclosporin A binding protein-hormone receptor activity (Figure 4D). KEGG enrichment analysis showed that these genes were found to be predominantly enriched in genes such as IL-17 signalling pathway, NOD-like receptor signalling pathway, TNF signalling pathway, Fluid shear stress and atherosclerosis, NF-kappa B signalling pathway, Toll-like receptor signalling pathway, Th17 cell differentiation, Lipid and atherosclerosis, Nucleotide excision repair, Calcium signalling pathway (Figure 4E). GO analysis of down-regulated genes revealed that the functions of chromosome segregation, sister chromatid segregation, regulation of chromosome organization, negative regulation of chromosome organization, negative regulation of chromosome organization, and nuclear chromosome segregation were annotated. KEGG

enrichment analysis showed that these down-regulated genes were mainly enriched for Fluid shear stress and atherosclerosis, Hematopoietic cell lineage, Lipoic acid metabolism, Mineral absorption, Platinum drug resistance, Proteoglycans in cancer and ECM-receptor interaction (Figure 4G).

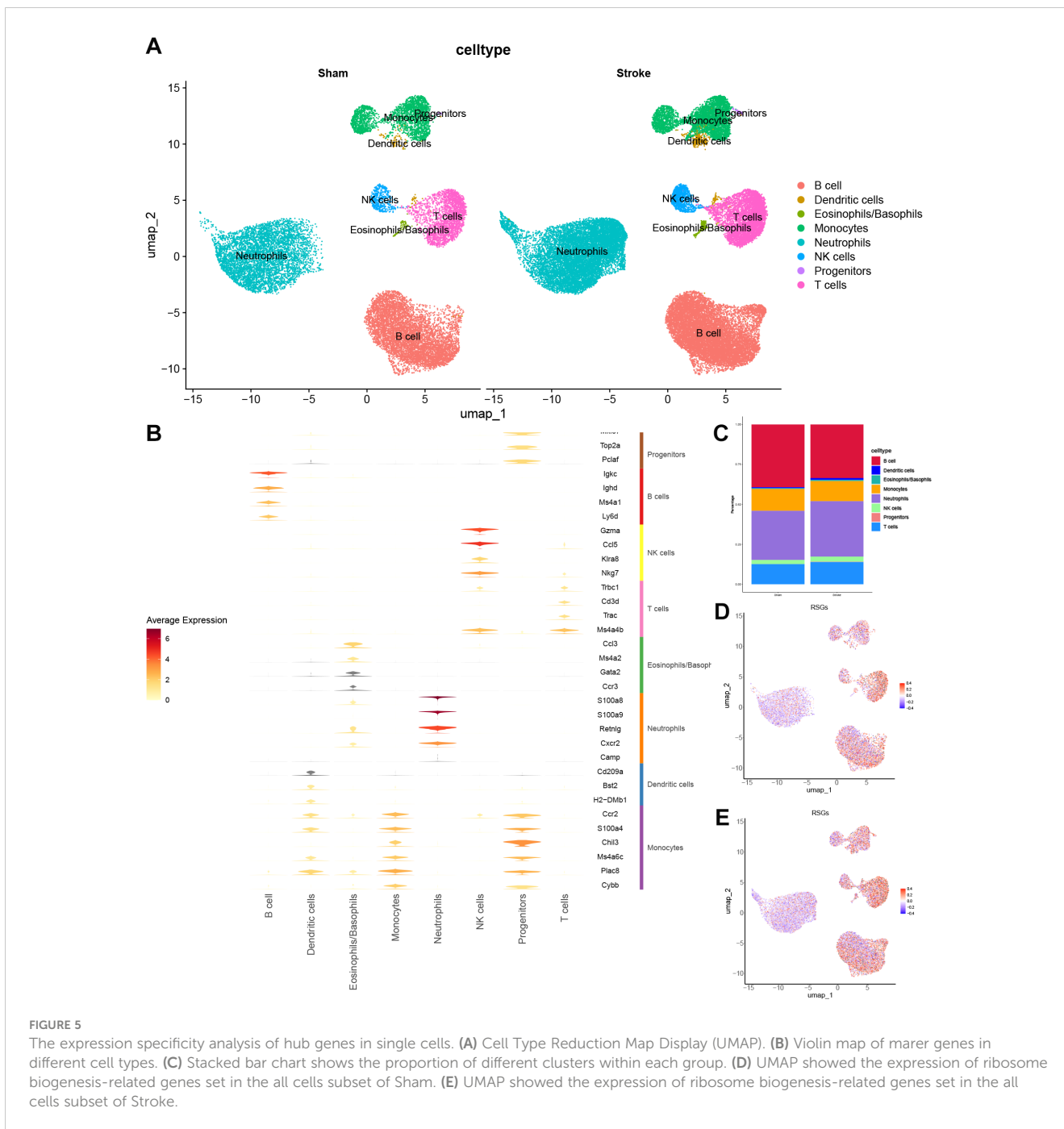## The expression specificity analysis of hub genes in single cells

The analysis of scRNA-seq data included 4 Ischemic stroke samples and 7 sham surgery samples. Following quality filtering, a total of 51,509 cells were screened. Subsequently, the top 10 principal components were selected through principal component analysis. The neighbourhoods between cells were then calculated at a resolution of 0.7, followed by downscaling and projection using UMAP. UMAP downscaling showed that these cells were categorized into 17 clusters (Supplementary Figure 2A) These 17 cell clusters expressed markers for 8 known cell types as shown (Supplementary Figure 2B). These 17 cell clusters were further divided into 8 known cell lineages (Supplementary Figure 2C), Distribution and cell proportion of 8 known cell types in sham operation and Ischemic stroke (Figure 5A). Marker expression of these 8 cell types (Figure 5B) and their proportion in Sham and Ischemic stroke (Figure 5C). We observed the IS-RSGs gene set score in all cells, Neutrophils in both the sham and ischemic stroke groups had lower scores (Figures 5D, E).

Expression specificity analysis of hub genes in neutrophil cells

In earlier times, we discovered disparities in neutrophils among the high-risk group. We subsequently isolated a total of 17,181 cells from the neutrophils subgroup and selected the top ten principal components following principal component analysis. The neighbourhood relationships between cells were calculated at a resolution of 0.3, followed by downscaling and projection using UMAP. The UMAP downscaling revealed that these cells were grouped into 6 clusters (Figure 6A). In Figures 6B, C, we found an increase in neutrophils in cluster 0, cluster 1, cluster 5 in the ischemic stroke group compared to the sham operation group. The IS-RSGs score of Cluster1 in neutrophils was lower than that of cluster0 and Cluster5 (Figure 6D, Supplementary Figure 2D).

Expression specificity analysis of hub genes in CD4+T cells and CD8+T cells

In the previous analysis, we found that there were also significant differences in several T cell ratios (Figure 3F), here we analysed the 6928 cells of the T cell subgroup of Figure 5A by principal component analysis and selected the top 10 principal components. Neighbourhood relationships between cells were then calculated at a resolution of 0.3, followed by downscaling and projection using UMAP. UMAP downscaling showed that these cells were classified

FIGURE 5

The expression specificity analysis of hub genes in single cells. **(A)** Cell Type Reduction Map Display (UMAP). **(B)** Violin map of marer genes in different cell types. **(C)** Stacked bar chart shows the proportion of different clusters within each group. **(D)** UMAP showed the expression of ribosome biogenesis-related genes set in the all cells subset of Sham. **(E)** UMAP showed the expression of ribosome biogenesis-related genes set in the all cells subset of Stroke.

into 9 clusters (Supplementary Figure 3A). Combined with the expression of Cd4, Cd8a, Ccr7, S100a4 (Supplementary Figure 3B), we divided the T cell subgroups into two main subgroups CD4$^+$T (Cluster0, Cluster2, Cluster6, Cluster7, Cluster8), CD8$^+$T (Cluster1, Cluster3, Cluster4, Cluster5). Finally, 3735 CD4$^+$ T cells were obtained, and after principal component analysis, the top 10 principal components were selected. Then the neighbour relationship between the cells was calculated at a resolution of 0.4, and then UMAP was used for downscaling and projection. UMAP downscaling showed that these cells were divided into 8 clusters (Figure 7A). Also, we observed whether these cells were from sham or from Ischemic stroke (Figure 7B). We previously found that the

proportion of T cells CD4 memory resting cells increased, and we combined with the Cd44 fractionation (Figure 7C) to find that Cluster1 was T cells CD4 memory resting, and we found that subgroup 5 of T cells CD4 memory resting was significantly higher in the Ischemic stroke than in the sham-operated group, based on the distribution of cell proportions (Figure 7E). T cells CD4 memory resting in cluster 5 was significantly higher in the Ischemic stroke than in the sham-operated group (Supplementary Figures 3C, D). Similarly, we observed the IS-RSGs gene set score in CD4$^+$T cells, and the expression of subgroup 5 was lower in T cells CD4 memory resting (Figure 7D). CD8$^+$T consisted of 3193 cells. After principal component analysis, the top 10 principal components were selected.
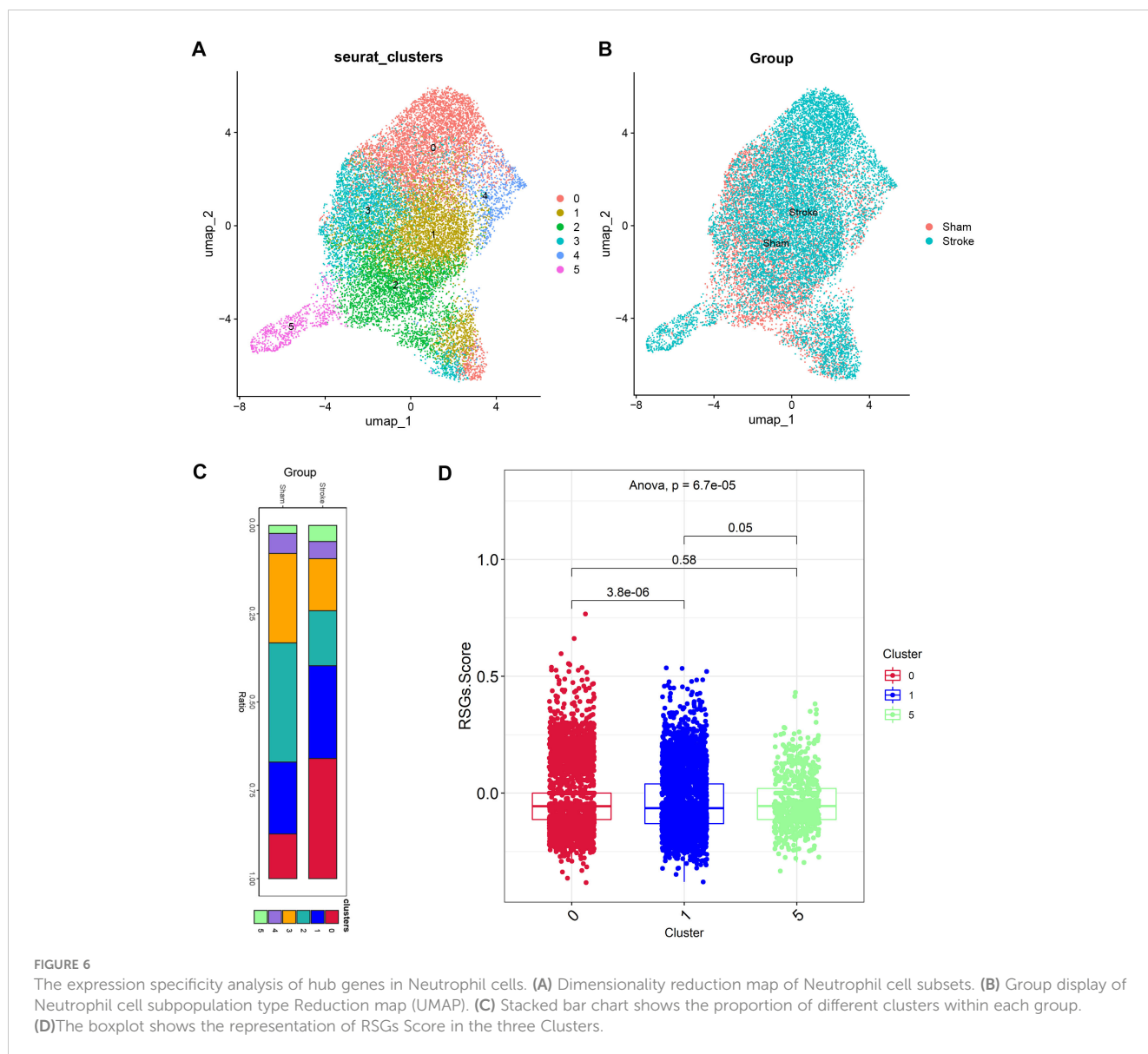
FIGURE 6

The expression specificity analysis of hub genes in Neutrophil cells. **(A)** Dimensionality reduction map of Neutrophil cell subsets. **(B)** Group display of Neutrophil cell subpopulation type Reduction map (UMAP). **(C)** Stacked bar chart shows the proportion of different clusters within each group. **(D)**The boxplot shows the representation of RSGs Score in the three Clusters.
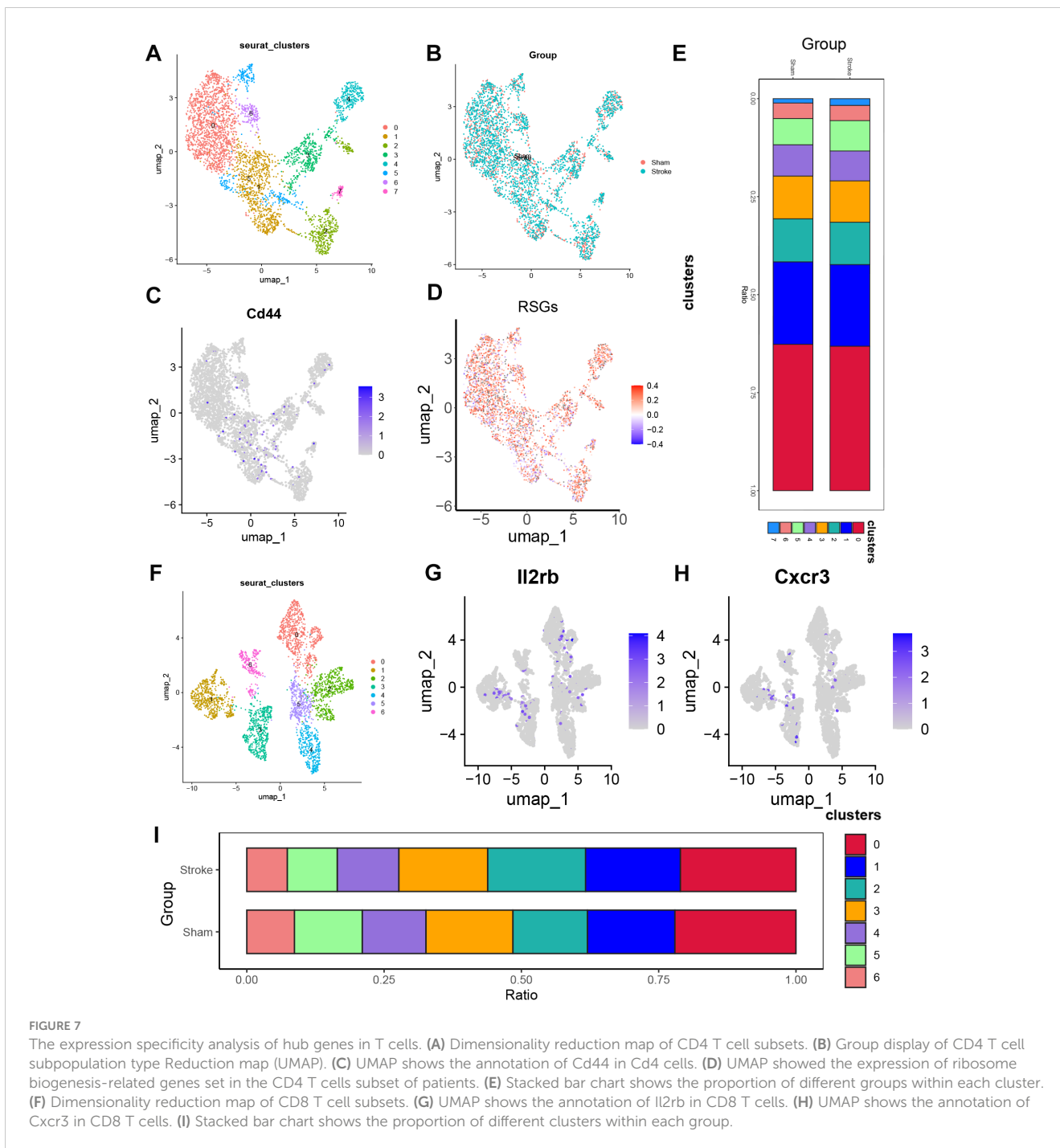
Neighbourhood relationships between cells were then calculated at a resolution of 0.3, followed by downscaling and projection using UMAP.UMAP downscaling showed that these cells were classified into 7 clusters. (Figure 7F). Among them subclusters 1, 2, 3 were increased in Ischemic stroke compared to the sham group (Figure 7I), combined with Figures 7G, H the three subclusters increased subclusters 1 and 3 were probably CD8(TRLs) cells, where the IS-RSGs gene set score in the subclusters was not seen to be significantly reduced (Supplementary Figure 3D).

Expression specificity analysis of hub genes in microglia cells

Microglia showed different functional states at different stages after stroke. We obtained 50,773 cells from single cells of mouse brain tissue after filtration, and then selected the first 5 principal components through principal component analysis, calculated the

neighbourhood relationship between cells with a resolution of 0.8, and then used UMAP for downscaling and projection. UMAP reduction showed that these cells were divided into 20 clusters (Supplementary Figure 4A), and we classified the subpopulations by markers specific to each subpopulation, and eventually we divided these 20 subpopulations into 9 cell types (Supplementary Figures 4B, C). We further selected 22,241 microglia, selected the first 5 principal components by principal component analysis, calculated the neighbourhood relationships between the cells at a resolution of 0.5, and then reduced and projected them using UMAP. UMAP showed that the cells were divided into 9 clusters. (Figure 8A) We found an increase in microglia in clusters 3, 4, 6, and 7 in the ischemic stroke group compared to the sham surgery group (Figures 8B, C). The IS-RSGs scores of Cluster3 and Cluster7 in microglia were significantly lower than those of Cluster4 and Cluster6 (Figure 8D). We also found that the expressions of P2ry12, Siglech and Tmem119 were lower in Cluster3 and Cluster7 than in other clusters (Figures 8E–G).

FIGURE 7

The expression specificity analysis of hub genes in T cells. **(A)** Dimensionality reduction map of CD4 T cell subsets. **(B)** Group display of CD4 T cell subpopulation type Reduction (UMAP). **(C)** UMAP shows the annotation of Cd44 in Cd4 cells. **(D)** UMAP showed the expression of ribosome biogenesis-related genes set in the CD4 T cells subset of patients. **(E)** Stacked bar chart shows the proportion of different groups within each cluster. **(F)** Dimensionality reduction map of CD8 T cell subsets. **(G)** UMAP shows the annotation of Il2rb in CD8 T cells. **(H)** UMAP shows the annotation of Cxcr3 in CD8 T cells. **(I)** Stacked bar chart shows the proportion of different clusters within each group.

Expression of serum RSG protein in patients with ischemic stroke and presentation of patient baseline data

We collected serum from 20 stroke patients, and serum ELISA results showed RNASEL, RPS28andC1QBP1.Patients were divided into high and low risk groups according to our scoring method. We analysed the gender, age, hypertension, hyperglycemia, neutrophil ratio, lymphocyte ratio, smoking history and drinking history of these 20 patients. The results showed that neutrophil ratio and

smoking history were different between the high and low risk groups (Figure 9).

Screening of relevant small molecule drugs

We predicted the differentially expressed genes between RSGs genes combined with RiskTypes via the CMAp (https://clue.io/query) website. After analysing and screening for 2131 relevant small molecule drugs (Figure 10A). We presented the small
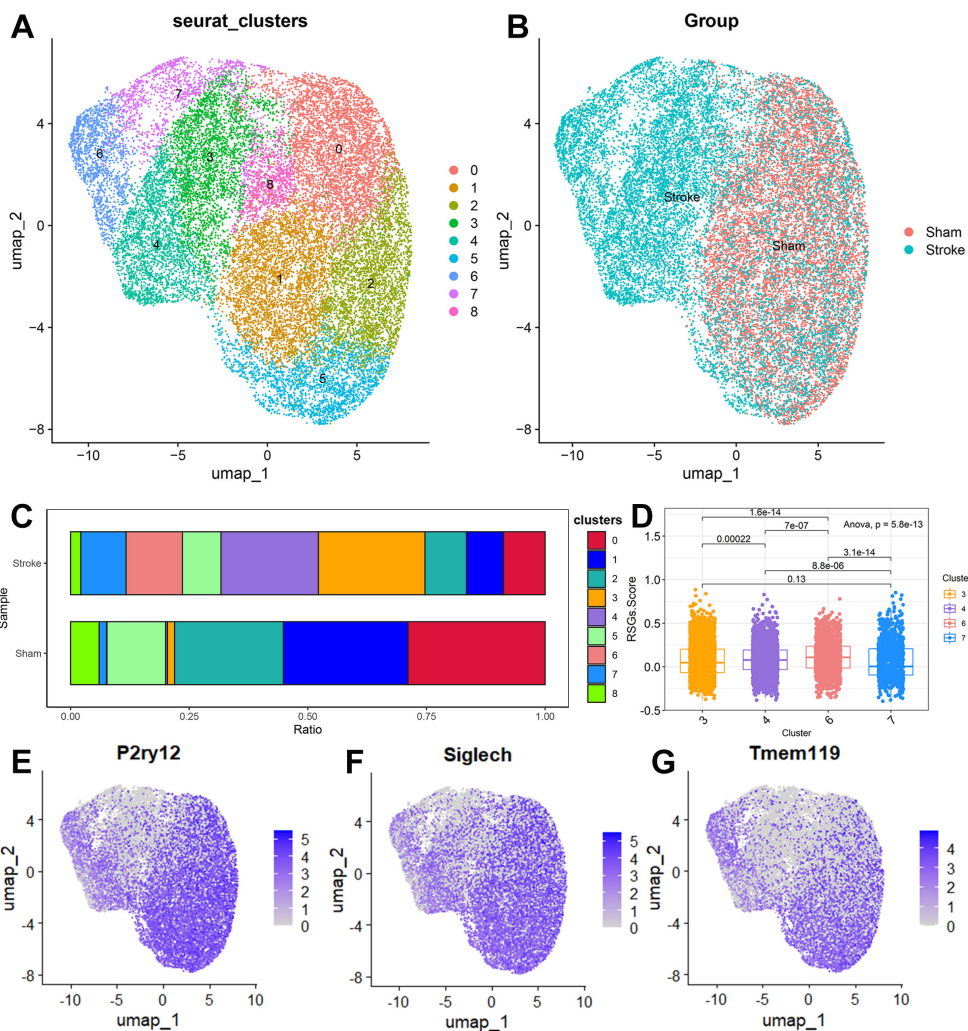
FIGURE 8

The expression specificity analysis of hub genes in Microglia cells. **(A)** Dimensionality reduction map of Microglia cell subsets. **(B)** Group display of Microglia cell subpopulation type Reduction map (UMAP). **(C)** Stacked bar chart shows the proportion of different clusters within each group. **(D)** The boxplot shows the representation of RSGs Score in the four Clusters. **(E)** UMAP shows the annotation of P2ry12 in Microglia cells. **(F)** UMAP shows the annotation of Siglech in Microglia cells. **(G)** UMAP shows the annotation of Tmem119 in Microglia cells.
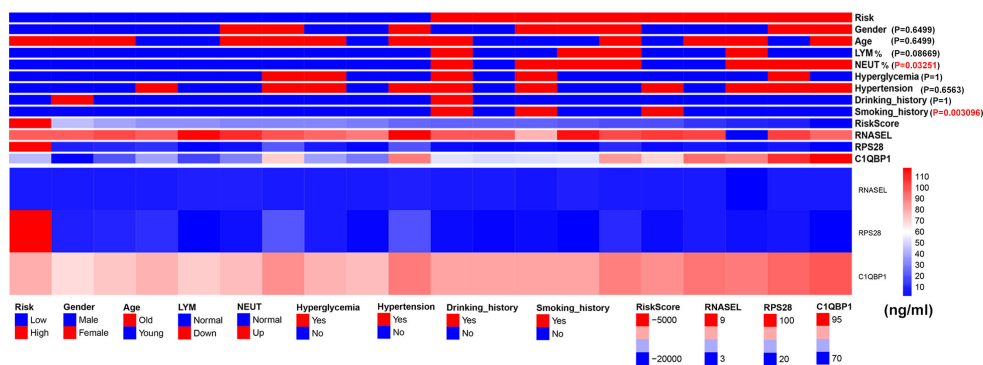

FIGURE 9

Expression of serum RSG protein in patients with ischemic stroke and presentation of patient baseline data.
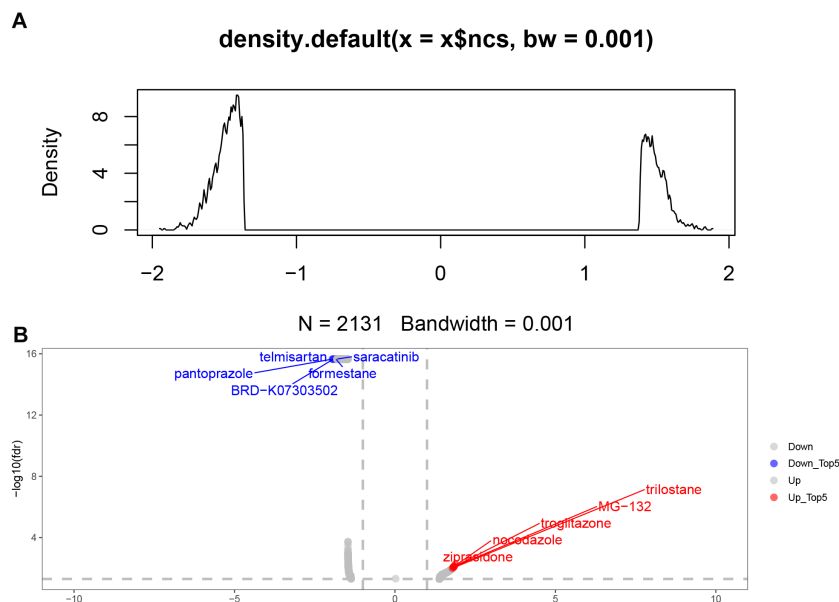
FIGURE 10
Screening of relevant small molecule drugs. **(A)** 2131 related small molecule drugs were screened. **(B)** The top and bottom 5 small molecule drugs were displayed.

molecule drugs with the top 5 inhibition rankings and the top 5 promotion rankings (Figure 10B).

Discussion

Stroke is the second leading cause of death and a major cause of disability worldwide (29). Rehabilitation for stroke can minimize the impact of disability on normal life, and early diagnosis and effective treatment of ischemic stroke are key to improving clinical outcomes (30, 31). Therefore, our study established a link between ribosome biogenesis genes and the pathogenesis of ischemic stroke. We identified potential IS-RSGs and explored future therapeutic targets via data analysis, thereby facilitating the development of more effective therapeutic strategies and novel drugs against potential therapeutic targets in ischemic stroke.

In our study, we finalized the set of 12 ribosome biogenesis-related genes (EXOSC5, MRPS11, MRPS7, RNASEL, RPF1, RPS28, C1QBP, GAR1, GRWD1, PELP1, UTP, ERI3). These sets of ribosome biogenesis-associated genes play a key role in stroke pathogenesis. These sets of ribosome biogenesis-associated genes play a key role in stroke pathogenesis. Among them, Exosc5 is a component of the RNA exosome complex and is involved in many cellular processes related to RNA processing and degradation. Importantly, mutations in the EXOSC5 gene are associated with an increased risk of sudden cardiac death (32). In a previous study, EXOSC5 was suggested to play an important role in stroke (33). MRPS11 and MRPS7 are members of the mitochondrial ribosomal proteins (MRPs) family, which is essential for the structural and functional integrity of the mitochondrial ribosomal complex (34). It has been shown that MRPS11 shows significant downregulation in the peripheral blood of ischemic stroke patients (35). It has been

reported that RNASEL plays an important role in the development of stroke (36, 37), and RNASE also has the functions of regulating the cell cycle (38) and apoptosis (39). Our analysis of single-cell data found that neutrophils increased in the ischemic stroke group. Neutrophils are the first cells to enter the brain after stroke, and they aggravate brain damage through a variety of mechanisms (40). Neutrophil extracellular traps (NETs) that can be released by neutrophils have been reported to be associated with poorer brain injury and stroke outcomes by impairs revascularization and vascular remodelling (41–43). This is consistent with the increase in Cluster0, Cluster1, and Cluster5 in the neutrophil subsets we found in the ischemic stroke group. The RSGs Score of Cluater1 is significantly lower than that of Cluater0 and Cluater5. This subgroup may be specific to the high-risk group we have identified. Microglia show different functional states at different stages after stroke. In our study, two groups of microglia were found. They also showed low expression of homeostasis related genes P2ry12 (44), Siglech (45) and Tmem119 (46) with low ribosome-related gene set scores. Microglia P2ry12 form special somatic junctions where ATP is released from the neuron's cell body, allowing continuous contact and monitoring of the neuron. After ischemia, the connectome area of microglia covering neurons increases, and the degree of neuronal damage can be reduced by regulating the concentration of calcium ions in neurons and enhancing the activity of mitochondria (47). Decreased P2ry12 expression may affect this function. This is consistent with the hypothesis that we found that the high-risk group had a poor prognosis. Studies have reported that By predicting small molecule drugs through CMAp, we found rilostane, MG-132, troglitazone, nocodazole and ziprasidone, the top 5 up drugs. It has been reported that Nocodazole binds to beta-tubulin and disrupts microtubule assembly/disassembly dynamics, thereby preventing

mitosis and inducing tumour cell apoptosis (48). MG-132 effectively blocks the proteolytic activity of 26S proteasome complex. MG-132 is a peptide aldehyde and an autophagy activator. MG-132 can also induce apoptosis (49, 50). These small molecule drugs are expected to improve outcomes in high-risk patients.

In summary, by using mass transcription and single-cell transcription techniques, we revealed the association between Ribosome biogenesis genes and infiltrating immune cells, and selected the features based on 12 Ribosome biogenesis genes as the best machine learning model for constructing a high and low risk group score for grouping patients with moderate scores. The expression of immune cells was different between the high-risk group and the low-risk group. At the single-cell level, we found that the proportion of neutrophils increased, and the score of IS-RSGs related gene set significantly decreased in neutrophils, while the neutrophils subsets of the high-risk group cluster0 and cluster5 were significantly increased in the analysis of neutrophils separately. In microglia, we found that the scores of IS-RSGs related gene sets in cluster3 and cluster7 were significantly reduced. At the stroke patient level, we found that patients in the high-risk group had a higher proportion of neutrophils and more patients with a history of smoking. Our findings reveal the role of Ribosome biogenesis genes in the progression of ischemic stroke and provide new insights into the underlying pathogenic processes and therapeutic strategies for ischemic stroke.

Limitations and outlook

The limitation of this study was that it was not possible to collect brain tissue of ischemic stroke patients for single-cell sequencing, and only relevant data of mouse brain tissue could be obtained through mouse models for verification. The mechanisms of how the 13 genes we identified function require further research, and we need to collect more samples to explore the differences between high and low risk patients. This study is based on the characteristics of 12 ribosome biogenetic genes as the optimal machine learning model for constructing high - and low-risk group scores. The high-risk group is expected to improve patient outcomes by increasing ribosome-related gene activity.

Data availability statement

The RNA-seq gene expression files was obtained from Gene Expression Omnibus (GEO) database, which can be accessed at (GEO, https://www.ncbi.nlm.nih.gov/geo/) GSE58294 (25) (containing 20 Control and 69 Stroke samples), GSE22255 (26) (containing 20 Control and 20 Stroke samples) GSE16561 (27) (containing 24 Control and 39 Stroke samples).

Ethics statement

The studies involving human participants were reviewed and approved by the Guangxi Hospital Division Of The First Affiliated Hospital, Sun Yat-Sen University(KY-LW-2024(008)). The patients/participants provided their written informed consent to participate in this study. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study. The manuscript presents research on animals that do not require ethical approval for their study.

Author contributions

XW: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Resources, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. X-YZ: Conceptualization, Data curation, Formal analysis, Funding acquisition, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. N-QL: Project administration, Writing – original draft, Data curation, Formal analysis, Investigation, Supervision. Q-FC: Funding acquisition, Resources, Writing – original draft, Writing – review & editing, Conceptualization, Project administration. Z-HH: Formal analysis, Funding acquisition, Investigation, Resources, Software, Supervision, Writing – original draft, Writing – review & editing.

Funding

Conflict of interest

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

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fimmu.2024.1449158/full#supplementary-material

References

1. Campbell BCV, Khatri P. Stroke. *Lancet (London England)*. (2020) 396:129–42. doi: 10.1016/S0140-6736(20)31119-X

2. Pu L, Wang L, Zhang R, Zhao T, Jiang Y, Han L. Projected global trends in ischemic stroke incidence, deaths and disability-adjusted life years from 2020 to 2030. *Stroke*. (2023) 54(5):1330–9. doi: 10.1161/STROKEAHA.122.040073

3. Deb P, Sharma S, Hassan KM. Pathophysiologic mechanisms of acute ischemic stroke: An overview with emphasis on therapeutic significance beyond thrombolysis. *Pathophysiology*. (2010) 17(3):197–218. doi: 10.1016/j.pathophys.2009.12.001

4. Davidson SM, Adameová A, Barile L, Cabrera-Fuentes HA, Lazou A, Pagliaro P, et al. Mitochondrial and mitochondrial-independent pathways of myocardial cell death during ischaemia and reperfusion injury. *J Cell Mol Med*. (2020) 24(7):3795–806. doi: 10.1111/jcmm.15127

5. Zera KA, Buckwalter MS. The local and peripheral immune responses to stroke: Implications for therapeutic development. *Neurotherapeutics*. (2020) 17(2):414–35. doi: 10.1007/s13311-020-00844-3

6. Hu S, Cai J, Chen S, Wang Y, Ren L. Identification of novel biomarkers and immune infiltration characteristics of ischemic stroke based on comprehensive bioinformatic analysis and machine learning. *Biochem biophysics Rep*. (2024) 37:101595. doi: 10.1016/j.bbrep.2023.101595

7. Amantea D, Micieli G, Tassorelli C, Cuartero MI, Ballesteros I, Certo M, et al. Rational modulation of the innate immune system for neuroprotection in ischemic stroke. *Front Neurosci*. (2015) 9:147. doi: 10.3389/fnins.2015.00147

8. Yang C, Hawkins KE, Doré S, Candelario-Jalil E. Neuroinflammatory mechanisms of blood-brain barrier damage in ischemic stroke. *Am J Physiol Cell Physiol*. (2019) 316(2):C135–c53. doi: 10.1152/ajpcell.00136.2018

9. Khatter H, Myasnikov AG, Natchiar SK, et al. Structure of the human 80S ribosome. *Nature*. (2015) 520:640–5. doi: 10.1038/nature14427

10. Darosa PA, Penchev I, Gumbin SC. UFM1 E3 ligase promotes recycling of 60S ribosomal subunits from the ER. *Nature*. (2024) 627(8003):445–52. doi: 10.1038/s41586-024-07073-0

11. Shen PS, Park J, Qin Y, Li X, Parsawar K, Larson MH, et al. Protein synthesis. Rqc2p and 60S ribosomal subunits mediate mRNA-independent elongation of nascent chains. *Sci (New York NY)*. (2015) 347:75–8. doi: 10.1126/science.1259724

12. Jiao L, Liu Y, Yu XY, Pan X, Zhang Y, Tu J, et al. Ribosome biogenesis in disease: new players and therapeutic targets. *Signal Transduct Target Ther*. (2023) 8(1):15. doi: 10.1038/s41392-022-01285-4

13. Li L, Cheng S-Q, Sun Y-Q, Yu JB, Huang XX, Dong YF, et al. Resolvin D1 reprograms energy metabolism to promote microglia to phagocytize neutrophils after ischemic stroke. *Cell Rep*. (2023) 42:112617. doi: 10.1016/j.celrep.2023.112617

14. Yang Z, Lin P, Chen B, Zhang X, Xiao W, Wu S, et al. Autophagy alleviates hypoxia-induced blood-brain barrier injury via regulation of CLDN5 (claudin 5). *Autophagy*. (2021) 17:3048–67. doi: 10.1080/15548627.2020.1851897

15. Wang M, Liang X, Cheng M, Yang L, Liu H, Wang X, et al. Homocysteine enhances neural stem cell autophagy in *in vivo* and in *in vitro* model of ischemic stroke. *Cell Death Dis*. (2019) 10:561. doi: 10.1038/s41419-019-1798-4

16. Derenzini M, Montanaro L, Trer D. Ribosome biogenesis and cancer. *Acta histochimica*. (2017) 119:190–7. doi: 10.1016/j.acthis.2017.01.009

17. Elhamamsy AR, Metge BJ. Ribosome biogenesis: A central player in cancer metastasis and therapeutic resistance. *Cancer Res*. (2022) 82(13):2344–53. doi: 10.1158/0008-5472.CAN-21-4087

18. Zang Y, Ran X, Yuan J, Wu H, Wang Y, Li H, et al. Genomic hallmarks and therapeutic targets of ribosome biogenesis in cancer. *Briefings Bioinf*. (2024) 25(2):bbae023. doi: 10.1093/bib/bbae023

19. Kourou K, Exarchos TP, Exarchos KP, Karamouzis MV, Fotiadis DI. Machine learning applications in cancer prognosis and prediction. *Comput Struct Biotechnol J*. (2015) 13:8–17. doi: 10.1016/j.csbj.2014.11.005

20. Tibshirani R. Regression shrinkage and selection via the lasso. *J R Stat Society: Ser B (Methodological)*. (2018) 58:267–88. doi: 10.1111/j.2517-6161.1996.tb02080.x

21. James GWD, Hastie T, Tibshiranir. *An Introduction to Statistical Learning*. New York: Springer (2013).

22. Li S, Han F, Qi N, Wen L, Li J, Feng C, et al. Determination of a six-gene prognostic model for cervical cancer based on WGCNA combined with LASSO and Cox-PH analysis. *World J Surg Oncol*. (2021) 19(1):277. doi: 10.1186/s12957-021-02384-2

23. Maurya NS, Kushwaha S, Chawade A, Mani A. Transcriptome profiling by combined machine learning and statistical R analysis identifies TMEM236 as a potential novel diagnostic biomarker for colorectal cancer. *Sci Rep*. (2021) 11:14304. doi: 10.1038/s41598-021-92692-0

24. Pei S, Zhang P, Yang L, Kang Y, Chen H, Zhao S, et al. Exploring the role of sphingolipid-related genes in clinical outcomes of breast cancer. *Front Immunol*. (2023) 14:1116839. doi: 10.3389/fimmu.2023.1116839

25. Stamova B, Jickling GC, Ander BP, Zhan X, Liu D, Turner R, et al. Gene expression in peripheral immune cells following cardioembolic stroke is sexually dimorphic. *PLoS One*. (2014) 9:e102550. doi: 10.1371/journal.pone.0102550

26. Krug T, Gabriel JP, Taipa R, Fonseca BV, Domingues-Montanari S, Fernandez-Cadenas I, et al. TTC7B emerges as a novel risk factor for ischemic stroke through the convergence of several genome-wide approaches [J]. *J Cereb Blood Flow metabolism: Off J Int Soc Cereb Blood Flow Metab*. (2012) 32:1061–72. doi: 10.1038/jcbfm.2012.24

27. O'connell GC, Treadway MB, Petrone AB, Tennant CS, Lucke-Wold N, Chantler PD, et al. Peripheral blood AKAP7 expression as an early marker for lymphocyte-mediated post-stroke blood brain barrier disruption. *Sci Rep*. (2017) 7:1172. doi: 10.1038/s41598-017-01178-5

28. Garcia-Bonilla L, Shahanoor Z, Sciortino R, Nazarzoda O, Racchumi G, Iadecola C, et al. Analysis of brain and blood single-cell transcriptomics in acute and subacute phases after experimental stroke. *Nat Immunol*. (2024) 25:357–70. doi: 10.1038/s41590-023-01711-x

29. Global, regional, and national burden of neurological disorders during 1990-2015: a systematic analysis for the Global Burden of Disease Study 2015. *Lancet Neurol*. (2017) 16:877–97. doi: 10.1016/S1474-4422(17)30299-5

30. Nair R, Wagner AN, Buck BH. Advances in the management of acute ischemic stroke. *Curr Opin Neurol*. (2023) 36:147–54. doi: 10.1097/WCO.0000000000001136

31. Ja DEG, Visser-Meily JM, Schepers VP, Baars A, Kappelle LJ, Passier PE, et al. Comparison between EQ-5D-5L and PROMIS-10 to evaluate health-related quality of life 3 months after stroke: a cross-sectional multicenter study. *Eur J Phys Rehabil Med*. (2021) 57:337–46. doi: 10.23736/S1973-9087.21.06335-8

32. Calame DG, Herman I, Fatih JM, Du H, Akay G, Jhangiani SN, et al. Risk of sudden cardiac death in EXOSC5-related disease. *Am J Med Genet Part A*. (2021) 185:2532–40. doi: 10.1002/ajmg.a.62352

33. Yu H, Ji X, Ouyang Y. Unfolded protein response pathways in stroke patients: a comprehensive landscape assessed through machine learning algorithms and experimental verification. *J Trans Med*. (2023) 21:759. doi: 10.1186/s12967-023-04567-9

34. Cheong A, Lingutla R, Mager J. Expression analysis of mammalian mitochondrial ribosomal protein genes. *Gene Expression patterns: GEP*. (2020) 38:119147. doi: 10.1016/j.gep.2020.119147

35. Tao J, Xie X, Luo M, Sun Q. Identification of key biomarkers in ischemic stroke: single-cell sequencing and weighted co-expression network analysis. *Aging*. (2023) 15:6346–60. doi: 10.18632/aging.v15i13

36. Walberer M, Tschernatsch M, Fischer S, Ritschel N, Volk K, Friedrich C, et al. RNase therapy assessed by magnetic resonance imaging reduces cerebral edema and infarction size in acute stroke. *Curr neurovascular Res*. (2009) 6:12–9. doi: 10.2174/156720209787466037

37. Xu Y, Hu Y, Xu S, Liu F, Gao Y. Exosomal microRNAs as potential biomarkers and therapeutic agents for acute ischemic stroke: new expectations. *Front Neurol*. (2021) 12:747380. doi: 10.3389/fneur.2021.747380

38. Al-Ahmadi W, Al-Haj L, Al-Mohanna FA, Silverman RH, Khabar KS. RNase L downmodulation of the RNA-binding protein, HuR, and cellular growth. *Oncogene*. (2009) 28:1782–91. doi: 10.1038/onc.2009.16

39. Brennan-Laun SE, Li XL, Ezelle HJ, Venkataraman T, Blackshear PJ, Wilson GM, et al. RNase L attenuates mitogen-stimulated gene expression via transcriptional and post-transcriptional mechanisms to limit the proliferative response. *J Biol Chem*. (2014) 289:33629–43. doi: 10.1074/jbc.M114.589556

40. Zhao Z, Pan Z, Zhang S, Ma G, Zhang W, Song J, et al. Neutrophil extracellular traps: A novel target for the treatment of stroke. *Pharmacol Ther*. (2023) 241:108328. doi: 10.1016/j.pharmthera.2022.108328

41. Mu Q, Yao K, Syeda MZ, Wan J, Cheng Q, You Z, et al. Neutrophil targeting platform reduces neutrophil extracellular traps for improved traumatic brain injury and stroke theranostics. *Advanced Sci (Weinheim Baden-Wurttemberg Germany)*. (2024) 11:e2308719. doi: 10.1002/advs.202308719

42. Kang L, Yu H. Neutrophil extracellular traps released by neutrophils impair revascularization and vascular remodeling after stroke. *Nat Commun*. (2020) 11:2488. doi: 10.1038/s41467-020-16191-y

43. Wang R, Zhu Y, Liu Z, Chang L, Bai X, Kang L, et al. Neutrophil extracellular traps promote tPA-induced brain hemorrhage via cGAS in mice with stroke. *Blood*. (2021) 138:91–103. doi: 10.1182/blood.2020008913

44. Buttgereit A, Lelios I, Yu X, Vrohlings M, Krakoski NR, Gautier EL, et al. Sall1 is a transcriptional regulator defining microglia identity and function. *Nat Immunol*. (2016) 17:1397–406. doi: 10.1038/ni.3585

45. Kumari A, Ayala-Ramirez R, Zenteno JC, Huffman K, Sasik R, Ayyagari R, et al. Single cell RNA sequencing confirms retinal microglia activation associated with early onset retinal degeneration. *Sci Rep*. (2022) 12:15273. doi: 10.1038/s41598-022-19351-w

46. Bennett ML, Bennett FC, Liddelow SA, Ajami B, Zamanian JL, Fernhoff NB, et al. New tools for studying microglia in the mouse and human CNS. *Proc Natl Acad Sci*. (2016) 113:E1738–E46. doi: 10.1073/pnas.1525528113

47. Fan Y, Xie L, Chung CY. Signaling pathways controlling microglia chemotaxis. *Molecules Cells*. (2017) 40:163–8. doi: 10.14348/molcells.2017.0011

48. Marquis C, Fonseca CL, Queen KA, Wood L, Vandal SE, Malaby HLH, et al. Chromosomally unstable tumor cells specifically require KIF18A for proliferation. *Nat Commun*. (2021) 12:1213. doi: 10.1038/s41467-021-21447-2

49. Harhouri K, Navarro C, Depetris D, Mattei MG, Nissan X, Cau P, et al. MG132-induced progerin clearance is mediated by autophagy activation and splicing regulation. *EMBO Mol Med*. (2017) 9:1294–313. doi: 10.15252/emmm. 201607315

50. Matsumoto Y, Miyamoto Y, Cabral H, Matsumoto Y, Nagasaka K, Nakagawa S, et al. Enhanced efficacy against cervical carcinomas through polymeric micelles physically incorporating the proteasome inhibitor MG132. *Cancer Sci*. (2016) 107:773–81. doi: 10.1111/cas.12926