

Exploring biomarkers and therapeutic targets for ischemic stroke through integrated microarray data analysis

Miao Lv

Guangxi Medical University

Wanting He

Guangxi Medical University

Tian Liang

Guangxi Medical University

Jialei Yang

Guangxi Medical University

Xiaolan Huang

Guangxi Medical University

Shengying Liu

Guangxi Medical University

Xueying Liang

Guangxi Medical University

Jianxiong Long

Guangxi Medical University

Li Su (✉ suli2018@hotmail.com)

Guangxi Medical University <https://orcid.org/0000-0002-1435-9669>

Research Article

Keywords: Ischemic Stroke, LASSO, WGCNA, Hub genes, Diagnostic biomarkers, Targeted drugs

Posted Date: December 1st, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-1105546/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Abstract

Stroke was the third leading cause of global disability-adjusted life years and its treatment drugs were limited. Researchers still need to find reliable diagnostic biomarkers and therapeutic targets. This study aimed to explore hub genes associated with ischemic stroke (IS) through bioinformatics analysis and identify its potential diagnostic biomarkers and therapeutic targets. The microarray data sets of GSE58294 and GSE16561 related to IS were downloaded from the Gene Expression Omnibus (GEO) database. Weighted co-expression network analysis (WGCNA) and differential expression analysis were integrated to find overlapping genes, and then hub genes were screened by Least absolute shrinkage and selection operator (LASSO) regression. NetworkAnalyst database was used to construct the TF-gene network and miRNA-TF regulatory network of the hub genes. DGIdb and CMap databases were used to predict targeted drugs and small molecule compounds. Through the Comprehensive analysis of GSE58294 and GSE16561, 10 hub genes were screened by LASSO regression, namely ARG1, LY96, ABCA1, SLC22A4, CD163, TPM2, SLC25A42, ID3, FAM102A and CD79B. NetworkAnalyst database analysis showed that ABCA1 was the hub gene most regulated by miRNA and TF. TRRUST and NetworkAnalyst databases showed that SP1 (P=0.0246) was a key transcription factor for hub genes. Finally, DGIdb database analysis identified 7 therapeutic drugs targeted with ABCA1 and SLC22A4, and 4 small molecule compounds targeted with ID3 and SLC22A4 were obtained from CMap database. We identified IS-related diagnostic biomarkers and therapeutic drugs, which can provide new insights for the diagnosis and treatment of IS.

Introduction

Stroke was the third leading cause of disability-adjusted life years globally, and the second leading cause of disability-adjusted life years for the age groups over 50 years old (Collaborators. 2020). Stroke was the leading cause of China's 2017 life lost years and disability-adjusted life years (Zhou et al. 2019). Among them, 85% of strokes were caused by cerebral ischemia, which were caused by obstruction of blood circulation in the brain. There are an estimated 11 million stroke cases and 1.1 million stroke-related deaths in China each year. The report of China's National Stroke Screening Survey conducted from 2013 to 2014 showed that compared with the values reported in the past ten years, the prevalence and morbidity of stroke have increased significantly (Wu et al. 2019). The treatment strategies for IS are limited. The only approved drug therapy to restore blood flow is intravenous injection of tissue-type plasminogen (tPA) (Group. 1995), However, it must be treated within 4.5 hours of the onset of symptoms. Therefore, the treatment time window is quite narrow, and many patients cannot receive treatment in time (Hacke et al. 2008; Henderson et al. 2018). Therefore, we still need to find reliable diagnostic biomarkers and therapeutic targets for IS.

Microarray technology makes genome-wide research possible, can perform genome functional annotations, and helps to explore the potential molecular mechanisms of many diseases (Ayyildiz and Piazza 2019). weighted correlation network analysis (WGCNA), a systematic biological method, is increasingly being used to analyze gene expression data. WGCNA analysis can identify the interaction between different expression modules, and find highly related gene modules, which promotes gene screening based on the construction of the network and can be used to identify candidate biomarkers or therapeutic targets (Langfelder et al. 2008). Recent studies have shown that hub genes related to the occurrence and development of diseases, candidate biomarkers and therapeutic drugs have been identified through bioinformatics technology. Rui et al used WGCNA and LASSO analysis to identify hub genes that regulate the prognosis of prostate cancer patients through immune infiltration (Rui et al. 2019). Qu et al also screened acute myeloid leukemia's hub gene by WGCNA and LASSO regression analysis (Qu et al. 2020). Hub genes and potential therapeutic drugs of substance use disorders patients in suicidal and non-suicidal contexts were identified by Cabrera-Mendoza et al through WGCNA analysis and DGIdb database (Cabrera-Mendoza et al. 2021).

In this study, we identified genes associated with IS through the integration of WGCNA and gene differential expression analysis. LASSO regression screened hub genes, and then the TF-gene network and miRNA-TF regulatory network were constructed through the NetworkAnalyst database. Finally, we predicted IS-targeted therapeutic drugs and small molecule compounds through DGIdb and CMap databases.

Material And Methods

Data collection and Preprocessing

Microarray datasets GSE58294 and GSE16561 were downloaded from the (GEO) database (<https://www.ncbi.nlm.nih.gov/geo/>). The onset of 69 ischemic stroke samples includes 3 time points (3h, 5h, and 24h). In this study, ischemic stroke samples in the time point of 24h were selected for analysis. The data set GSE16561 includes 24 control samples and 39 ischemic stroke samples. The details of two microarray datasets were listed in Table 1. Normalization (limma package), log₂ transformation and gene name conversion were performed for each data set using Rstudio4.0.5 software.

Construction of weighted correlation network analysis (WGCNA)

The "sva" package was used to remove the batch effect on the combined data sets of GSE16561 and GSE58294. After detecting and removing the missing values of the gene expression matrix, a weighted co-expression network was constructed through the WGCNA package. WGCNA's pick Soft Threshold function was used to calculate soft-threshold power. The correlation of module genes between the case group and control group was limited to be greater than 0.5 to screen and obtain module characteristic genes related to ischemic stroke. In addition, we used the "clusterProfiler" and "enrichplot" packages to perform functional enrichment analysis (Gene Ontology, GO and Kyoto Encyclopedia of Genes and Genomes, KEGG).

Identification of Differentially expressed genes (DEGs)

The standard of differential expression analysis was $|\log FC| > 1$, and the P value < 0.05 . Genes with $\log FC > 1$ and the p value < 0.05 were up-regulated, while $\log FC < -1$ and the p value < 0.05 were down-regulated.

Identification of overlapping genes

Overlapping genes were selected between modular genes by the weighted coexpression network and DEGs of GSE16561 and GSE58294. An online site (<http://www.ehbio.com/test/venn/#/>) was used to plot Venn diagrams.

Screening of hub genes

Overlap genes were further screened by LASSO regression, which was performed through the "glmnet" package. Then, the "ggstatsplot" package and the "ggpubr" package were used to make a violin plot. Finally, the ROC curve of the hub gene was made by the "pROC" package.

TF-gene interaction network and TF-miRNA regulatory network

Building a network of TF target genes and genetic data from ENCODE ChIP-seq (<https://www.encodeproject.org/>) database. A comprehensive analysis of TF-gene interaction was conducted to evaluate the effects of TF on the expression and functional pathways of hub genes (Ye et al. 2019). The NetworkAnalyst (<https://www.networkAnalyst.ca/>) was used to visualize the interaction network between TF genes and hub genes. Relationships in TF-miRNA co-regulatory networks were collected from the RegNetwork library, which contained transcriptional and post-transcriptional regulatory relationships, and interactions between TF/ miRNAs and their targets could be obtained from this database (Liu et al. 2015). The NetworkAnalyst website was also used to visualize the TF-miRNA coordinated regulation network.

TRRUST database

The TRRUST database (<https://www.grnpedia.org/trrust/>) is currently the most comprehensive human-mouse TF target interaction database based on manual management. Researchers can search for key TF genes by submitting target genes (Han et al. 2018).

HPA database

To determine the expression and localization of hub genes in normal brain and blood, we downloaded immunohistochemical slices of hub genes in the brain and RNA expression data of the hub genes in the brain and blood from the HPA database (<https://www.proteinatlas.org/>).

Identification of candidate drugs and small molecule compounds

In this study, the Drug-Gene Interaction Database (DGIdb, <http://www.dgldb.org>) was used to explore effective targeting drugs for hub genes, which parameter select approval filter conditions, and the default database. The Touchtone function of Connection map (CMap, <https://clue.io/>) database was used to search for small molecule compounds related to the input hub genes, and small molecule compounds with a gene selection score of less than -90 play a reversing role in gene upregulation. The gene selection score of small molecule compounds is greater than 90, promoting the protective effect of genes (Musa et al. 2018).

Results

WGCNA gene modules results

After detecting the missing values of 12694 gene expression matrices, 9520 gene expression matrices were obtained, which were used to construct a weighted co-expression network. When 0.85 was used as the correlation coefficient threshold, we chose a soft threshold of 9 (Fig.1a). A total of 17 co-expression modules (Except for the gray module) were identified by WGCNA, the size of which ranged from 91 to 791 genes. Genes with different characteristics were divided into different modules to obtain gene cluster trees and different module colors (Fig. 1b). The heat map showed the relationship of module's characteristic genes between the case group and the control group with the correlation coefficients and significance P -values of each module (Fig. 1c). A summary of the significance of ischemic stroke-related genes in each module was shown in Fig. 1d.

Identification of DEGs

Genes with $\log FC > 1$ and the p value < 0.05 were up-regulated, while $\log FC < -1$ and the p value < 0.05 were down-regulated. A total of 11 differential expression genes (DEGs) were identified in the data set GSE16561, including 6 up-regulated genes and 5 down-regulated genes. At the same time, a total of 304 DEGs were identified in the data set GSE58294, including 157 up-regulated genes and 147 down-regulated genes. The DEGs of the two data sets were merged to obtain 313 IS-related DEGs.

Functional enrichment of module genes associated with IS

The modules with high correlation with ischemic stroke were screened from 18 modules (except gray). The correlation between the screening module and ischemic stroke was greater than 0.5, three modules with high correlation with ischemic stroke were obtained, namely green (355), brown (449) and red (306). The three modules were merged, and 1110 module genes were obtained. Meanwhile, 1100 modular genes were included for GO and KEGG function enrichment analysis (Fig. 2a, b). GO enrichment analysis showed that biologic processes (BP) were mainly enriched in the T cell activation, neutrophil

degranulation and neutrophil activation involved in immune response. For cellular component (CC), modular genes were mainly enriched in the organellar ribosome, mitochondrial ribosome and mitochondrial protein complex. For the molecular function (MF), modular genes were mostly enriched in immune receptor activity, ribonucleoprotein complex binding and protein N-terminus binding. KEGG pathway analyses were mainly enriched in the NF- κ B signaling pathway, apoptosis, lipid and atherosclerosis pathway.

Integrated analysis of WGCNA and DEGs

Integration analysis on the selected three modular genes and IS-related DEGs were performed. A total of 52 genes were obtained by taking the intersections of 313 IS-related module genes and 1100 module genes (Fig. 3a).

Screening of hub genes

52 overlapping genes were further screened in the merged data (control=47, ischemic stroke=62) by LASSO regression analysis, and 10 hub genes were obtained (Fig. 4a). The violin map showed that ARG1, LY96, ABCA1, SLC22A4, CD163 were up-regulated, while TPM2, SLC25A42, ID3, FAM102A and CD79B were down-regulated, and P values were all less than 0.05 (Fig. 4b). Fig. 4c showed the ROC curve of 10 hub genes, Values of sensitivity, specificity and AUC (area under the curve) were displayed in Table S1. Among them, FAM102A had the highest diagnostic value and the value of its AUC, sensitivity and specificity were 0.973576, 0.919 and 0.936, respectively. LY96 had the lowest diagnostic value, with an AUC of 0.841, a sensitivity of 0.726, and a specificity of 0.894.

Construction of TF-gene interaction network and TF-miRNA regulatory network

Ten hub genes were submitted to the NetworkAnalyst database for analysis, and network for TF-gene interaction (Fig. 5a) and TF-miRNA regulatory network (Fig. 5b) were obtained. In the TF-gene interaction network, ID3 had the highest interaction rate with other TFs, up to 75 degrees. Among the regulatory factors, TFDP1 had significant interaction, up to 4 degrees. In the TF-miRNA regulatory network, ABCA1 had a high interaction rate with miRNAs and TFs with a degree of 92. Among the regulatory factors of miRNA and TF genes, transcription factors SP1, hsa-miR-340, hsa-miR-27a, hsa-miR-27b and hsa-miR-96 had significant interactions with hub genes. Meanwhile, SP1 (P=0.0246) was the key transcription factor for the six hub genes ABCA1, ARG1, CD163, ID3, LY96, and SLC22A4 in the TRRUST database.

HPA database analysis

The ten hub genes were submitted to the HPA database, and the immunohistochemical slices of the eight genes in the brain (Fig. 5c) and the gene expression levels in the brain and blood were obtained (Fig. 5d). Among them, the expression levels of ABCA1, SLC25A42, SLC22A4, ID3 and FAM102A in the brain were higher than those in the blood.

Potential drug analysis of hub genes

The 5 hub genes including ABCA1, SLC25A42, SLC22A4, ID3 and FAM102A were uploaded to the DGIdb Database, and seven targeted drugs with hub genes were found. As shown in Table 2, there were six therapeutic drugs targeted with ABCA1, namely simvastatin, probucol, fenofibrate, cyclosporine, pravastatin and atorvastatin. Imatinib was a therapeutic drug targeted with SLC22A4. We also submitted these 5 hub genes to the CMap database for analysis. Four small molecule compounds targeted with ID3 and SLC22A4 were listed in Table 3. There are three small molecule compounds targeted with ID3, namely 2', 5'-dideoxyadenosine, DR-2313 (Fig. S1) and otenzepad. Dubinidine was a small molecule compound targeted with SLC22A4.

Discussion

In this study, 10 hub genes related to the occurrence of IS were identified through WGCNA and LASSO regression analysis based on the merged datasets of GSE16561 and GSE58294. Values of the AUC for the 10 hub genes ranged from 0.841 to 0.974. ID3 had the highest interaction rate with other TF-genes in the TF-gene interaction network. ABCA1, transcription factors SP1 and hsa-miR-340 had a high interaction rate in the miRNA-TF regulatory network. HPA database analysis showed that ABCA1, SLC25A42, SLC22A4, ID3 and FAM102A had higher gene expression levels in the cerebral cortex than those in blood. We also predicted seven therapeutic drugs targeted with ABCA1 and SLC22A4, and four small molecule compounds targeted with ID3 and SLC22A4.

KEGG analysis showed that modular genes were mainly enriched in 13 pathways. Among them, the NF- κ B signaling pathway, apoptosis, lipid and atherosclerosis pathways may be related to the pathogenesis and development of IS. As for the NF- κ B signaling pathway, evidence showed that ischemic preconditioning could activate the NF- κ B signaling pathway, and enhance the expression of NF- κ B to protein in neurons after lethal ischemia-reperfusion (Lee et al. 2016). Study had shown that the NF- κ B gene was the downstream target of the Notch signaling pathway (Schwarzer et al. 2012). The Notch signal partly participated in the activation and regulation of microglia after hypoxia through the interaction of the TLR4/MyD88/TRAF6/NF- κ B signaling pathway (Yao et al. 2013). Fann et al reported that the NF- κ B signaling pathway played a key role in the expression and activation of NLRP1 and NLRP3 inflammasomes in primary cortical neurons and brain tissue under ischemic conditions (Fann et al. 2018). For apoptosis pathway, current study manifested that the internal and external pathways for cerebral ischemia to trigger apoptosis were derived from the release of cytochrome C from mitochondria, the activation of caspase-3 and cell surface death receptors, which lead to the stimulation of caspase-8 (Broughton et al. 2009). Overexpression of the apoptotic protein bcl-2 can reduce cerebral infarction tissue in the middle cerebral artery occlusion model (Martinou et al. 1994). By preventing injury-induced down-regulation of the apoptotic protein bcl-2 expression, estradiol had a profound protective effect on cerebral ischemia

caused by permanent middle cerebral artery occlusion (Dubal et al. 1999). The overexpression of a variety of pro-apoptotic and anti-apoptotic proteins in the penumbra after ischemic stroke may be used as potential therapeutic targets for stroke (Uzdensky 2019). For the lipid and atherosclerosis pathway, lipids accumulate on the arterial wall in the form of fatty deposits and cholesterol plaques. It was the main component of early atherosclerotic lesions (Kauw et al. 2018). Atherosclerosis was a chronic inflammatory disease characterized by the formation of lipid plaques on the blood vessel wall, which can increase the risk and recurrence of stroke (Chow et al. 2020).

Ten hub genes were identified in this study, which were ARG1, TPM2, LY96, ABCA1, SLC25A42, SLC22A4, ID3, FAM102A, CD79B and CD163. Previous experimental studies have revealed that 6 hub genes were related to the occurrence and development of stroke. Study have shown that STAT6/ARG1 regulated the efferocytosis of the microglia/macrophage phenotype, accelerated the resolution of inflammation, and improved the prognosis of stroke (Cai et al. 2019). Barr et al reported that LY96 and ARG1 had been identified as stroke expression profile genes, which may be biomarkers of stroke (Barr et al. 2010). As a major reverse cholesterol transporter, ABCA1 plays a key role in the formation of high-density lipoprotein (HDL) cholesterol in the brain. ABCA1/ApoE/HDL pathway was at least partially involved in the process of nerve recovery after stroke induced by gw3965 (Cui et al. 2017). In addition, ABCA1 rs2230806 GG was significantly associated with an increased risk of IS (Au et al. 2017). SLC22A4's rs273909 was significantly associated with IS (P=0.0123) (Yamase et al. 2015), which increased the susceptibility of IS. ID3, an Inhibitor Of DNA Binding 3, was involved in stress response, neural plasticity, neural circuits, and other processes (Avecilla et al. 2017). CD163 is a scavenger receptor expressed in the innate immune cell population. CD163 served as a potential biomarker for monocyte activation in patients with ischemic stroke (Greco et al. 2021).

Adamski et al showed that the transcriptome AUC containing 7 genes PLBD1, PYGL, BST1, DUSP1, FOS, VCAN and FCGR1A was 0.854 for identification of stroke by the high throughput next-generation qPCR verification (Adamski et al. 2014). In our research, we finally got 10 hub genes, 9 genes had AUC values higher than 0.854, the AUC of 7 genes had higher than 0.9. FAM102A had the highest diagnostic value with the AUC of 0.974, sensitivity of 0.919 and specificity of 0.936, respectively.

NetworkAnalyst database analysis showed that transcription factor SP1 encodes a zinc finger transcription factor, which was involved in many cellular processes such as cell differentiation, apoptosis, and immunity. It has been reported that Curcumin up-regulated Peroxiredoxin 6 through SP1 to induce ischemic oxidative damage after stroke in rats (Jia et al. 2017). H₂O can be used as a signal molecule by neurons to quickly induce Sp1 translation through an ires-dependent translation pathway. Through feedforward transcription activation, with the late rise of Sp1, neurons can be protected from ischemic damage (Yeh et al. 2011). Yoo et al indicated that miR-340-5p regulated immune response by targeting ARG1 protein after acute ischemic stroke (Yoo et al. 2019).

HPA database analysis showed that the 10 hub genes in this study were expressed in the brain and blood. Studies have found a correlation between brain gene expression profiles and blood gene expression profiles, and blood gene expression profiles had the potential to evaluate various disease states in terms of diagnosis, mechanism and treatment (Tang et al. 2001, 2002). A brain tissue gene expression prediction model based on whole blood gene expression data proved that only whole blood expression profile data could accurately predict the gene expression of unsampled brain tissue (Wenjian and Wei 2019). Iturria-medina Y et al showed that blood gene expression strongly reflected the severity of neuropathology, clinical deterioration and disease progression in vivo (Iturria-Medina et al. 2020). Therefore, we hypothesize that gene expression in the blood may reflect brain conditions in some extent. Among them, ABCA1, SLC25A42, SLC22A4, ID3 and FAM102A had higher gene expression levels in the cerebral cortex than in blood. Compared with the other five genes, we speculated that these five genes with high expression in brain tissue had more important value in the pathological mechanism of IS occurrence and development, and may be used as a target for drug therapy.

Next, we used DGIdb and cMap databases to predict 5 gene-targeted therapeutic drugs. DGIdb database analysis showed that there were 6 therapeutic drugs targeted with ABCA1, namely simvastatin, probucol, fenofibrate, cyclosporine, pravastatin and atorvastatin. Imatinib was a therapeutic drug targeted with SLC22A4. CMap database analysis showed that there were 3 small molecule compounds targeted with ID3, namely 2', 5'-dideoxyadenosine, DR-2313 and otenzepad. And dubinidine was a small molecule compound targeted with SLC22A4. Among the drugs targeted ABCA1, simvastatin, pravastatin and atorvastatin are all statins. According to previous studies, the use of statins for the treatment of stroke has long been widely accepted and used by related researchers, which has been promoted in the continuously updated guidelines. According to the 2013 AHA/ASA guidelines, statins were recommended to be applied in patients presumed to be atherosclerotic ischemic stroke or TIA to reduce the risk of stroke and cardiovascular events (Kernan et al. 2014). Statins can significantly reduce the infarct volume and neurological deficits in tMCAO model mice. At the same time, compared with atorvastatin, simvastatin and pravastatin were more significant in reducing infarct volume (Christophe et al. 2020). Statins seemed to reduce the risk of recurrence of ischemic stroke and other cardiovascular events in patients (Tramacere et al. 2019).

Probucol inhibited LPS-induced microglia activation and improved cerebral ischemic injury in mice through anti-neuro-inflammatory effects, and it may be a potential therapeutic drug for the treatment of ischemic patients (Jung et al. 2016). Fenofibrate could significantly reduce the volume of cerebral infarction, the activation of microglia and the infiltration of neutrophils in the ischemic area to play a protective effect on the brain, which may be related to the acute effect and preconditioning mechanism (Ouk et al. 2009). Fenofibrate played a preventive and protective effect on cerebral ischemia by reducing the infarct volume, and the related mechanism may be related to the genome regulation in the apoptosis pathway (Altintas et al. 2017). Cyclosporine A had potential neuroprotective properties in ischemic stroke, which interfered with the typical NF- κ B signaling pathway, and the Akt pathway may also participate in the role of Cyclosporine A (Deng et al. 2020; Nighoghossian et al. 2016). Deng et al showed that DR2313 exerted neuroprotective effects through its powerful PARP inhibitory effect, and could easily penetrate the blood-brain barrier after peripheral to the medicine. It may be more useful than free radical scavengers in the treatment of acute stroke, which may be a promising drug for stroke treatment (Nakajima et al. 2005).

Conclusion

In summary, we finally screened out 10 hub genes that may play a key role in the occurrence and development of IS through comprehensive bioinformatics analysis. Among them, SLC25A42 and FAM102A may be new candidate biomarkers or therapeutic targets. In addition, we also predicted gene-targeted therapeutic drugs, and many previous studies have supported the value of these drugs in the treatment of IS.

Declarations

Funding

This work was supported by National Natural Science Foundation of China (NO. 81874395), Natural Science Foundation of Guangxi Zhuang Autonomous Region (NO. 2017GXNSFDA198013 and NO. 2018GXNSFAA281224).

Conflicts of interest

All authors declare that they have no conflicts of interest.

Ethics approval

Not applicable.

Consent to participate (include appropriate statements)

Not applicable.

Consent for publication

Not applicable.

Availability of data and material

Not applicable.

Data availability statements

Data sharing not applicable to this article as no datasets were generated or analysed during the current study.

Code availability

Not applicable.

Authors' contributions

All authors contributed to the conception and design of the research. Xueying Liang and Shengying Liu were responsible for the collection of materials. Miao Lv and Wanting He analyzed and interpreted the data. Jialei Yang and Miao Lv wrote the manuscript. Tian Liang and Xiaolan Huang conceived and designed the study. Jianxiong Long and Li Su supervised the study and provided financial support. All authors approved the final manuscript and agreed to publish the Journal of Molecular Neuroscience.

References

1. Adamski MG, Li Y, Wagner E et al (2014) Expression profile based gene clusters for ischemic stroke detection. *Genomics* 104:163–169
2. Altintas O, Altintas MO, Aydin MS et al (2017) Neuroprotective Effects of Chronic Fenofibrate Treatment via Modulating the Immunoreactivity of Cleaved Caspase-3 in Stroke Induced by Transient Middle Cerebral Artery Occlusion Rat Model. *Turk Neurosurg* 27:522–529
3. Au A, Griffiths LR, Irene L, Kooi CW, Wei LK (2017) The impact of APOA5, APOB, APOC3 and ABCA1 gene polymorphisms on ischemic stroke: Evidence from a meta-analysis. *Atherosclerosis* 265:60–70
4. Avecilla V, Doke M, Felty Q (2017) Contribution of Inhibitor of DNA Binding/Differentiation-3 and Endocrine Disrupting Chemicals to Pathophysiological Aspects of Chronic Disease. *Biomed Res Int* 2017: 6307109
5. Ayyildiz D, Piazza S (2019) Introduction to Bioinformatics. *Methods Mol Biol* 1986: 1-15
6. Broughton BR, Reutens DC, Sobey CG (2009) Apoptotic mechanisms after cerebral ischemia. *Stroke* 40:e331–e339
7. Barr TL, Conley Y, Ding J, Dillman A, Warach S, Singleton A, Matarin M (2010) Genomic biomarkers and cellular pathways of ischemic stroke by RNA gene expression profiling
8. Cui X, Chopp M, Zhang Z et al (2017) ABCA1/ApoE/HDL Pathway Mediates GW3965-Induced Neurorestoration After Stroke. *Stroke* 48:459–467
9. Cai W, Dai X, Chen J et al (2019) STAT6/Arg1 promotes microglia/macrophage efferocytosis and inflammation resolution in stroke mice. *JCI Insight* 4

10. Chow YL, Teh LK, Chyi LH, Lim LF, Yee CC, Wei LK (2020) Lipid Metabolism Genes in Stroke Pathogenesis The Atherosclerosis
11. Christophe B, Karatela M, Sanchez J, Pucci J, Connolly ES (2020) Statin Therapy in Ischemic Stroke Models: A Meta-Analysis. *Transl Stroke Res* 11:590–600
12. Cabrera-Mendoza B, Martínez-Magaña JJ, Monroy-Jaramillo N et al (2021) Candidate pharmacological treatments for substance use disorder and suicide identified by gene co-expression network-based drug repositioning. *Am J Med Genet B Neuropsychiatr Genet* 186:193–206
13. Dubal DB, Shughrue PJ, Wilson ME, Merchenthaler I, Wise PM (1999) Estradiol Modulates bcl-2 in Cerebral Ischemia. A Potential Role for
14. Deng H, Zhang S, Ge H et al (2020) The effect of cyclosporin a on ischemia-reperfusion damage in a mouse model of ischemic stroke. *Neurol Res* 42:721–729
15. Fann DY, Lim YA, Cheng YL et al (2018) Evidence that NF-kappaB and MAPK Signaling Promotes NLRP Inflammasome Activation in Neurons Following Ischemic Stroke. *Mol Neurobiol* 55:1082–1096
16. GBD 2019 Diseases and Injuries Collaborators (2020) Global burden of 369 diseases and injuries in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. *Lancet* 396:1204–1222
17. Greco R, Demartini C, Zanaboni AM et al (2021) CD163 as a Potential Biomarker of Monocyte Activation in Ischemic Stroke Patients. *Int J Mol Sci* 22
18. Hacke W, Kaste M, Bluhmki E et al (2008) Thrombolysis with alteplase 3 to 4.5 hours after acute ischemic stroke
19. Han H, Cho JW, Lee S et al (2018) TRRUST v2: an expanded reference database of human and mouse transcriptional regulatory interactions. *Nucleic Acids Res* 46:D380–DD86
20. Henderson SJ, Weitz JI, Kim PY (2018) Fibrinolysis: strategies to enhance the treatment of acute ischemic stroke. *J Thromb Haemost* 16:1932–1940
21. Iturria-Medina Y, Khan AF, Adewale Q, Shirazi AH, Alzheimer's Disease Neuroimaging Initiative (2020) Blood and brain gene expression trajectories mirror neuropathology and clinical deterioration in neurodegeneration. *Brain* 143:661–673
22. Jung YS, Park JH, Kim H et al (2016) Probucol inhibits LPS-induced microglia activation and ameliorates brain ischemic injury in normal and hyperlipidemic mice. *Acta Pharmacol Sin* 37:1031–1044
23. Jia G, Tan B, Ma J, Zhang L, Jin X, Li C (2017) Prdx6 Upregulation by Curcumin Attenuates Ischemic Oxidative Damage via SP1 in Rats after Stroke. *Biomed Res Int* 2017: 6597401
24. Kernan WN, Ovbiagele B, Black HR et al (2014) Guidelines for the prevention of stroke in patients with stroke and transient ischemic attack: a guideline for healthcare professionals from the American Heart Association/American Stroke Association. *Stroke* 45:2160–2236
25. Kauw F, Takx RAP, de Jong HWAM, Velthuis BK, Kappelle LJ, Dankbaar JW (2018) Clinical and Imaging Predictors of Recurrent Ischemic Stroke: A Systematic Review and Meta-Analysis. *Cerebrovasc Dis* 45:279–287
26. Langfelder P, Horvath S (2008) WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* 9:559
27. Liu ZP, Wu C, Miao H, Wu H (2015) RegNetwork: an integrated database of transcriptional and post-transcriptional regulatory networks in human and mouse. *Database* 2015
28. Lee JC, Tae HJ, Kim IH et al (2016) Roles of HIF-1 α , VEGF, and NF- κ B in Ischemic Preconditioning-Mediated Neuroprotection of Hippocampal CA1 Pyramidal Neurons Against a Subsequent Transient Cerebral Ischemia. *Mol Neurobiol* 54:6984–6998
29. Martinou JC, Dubois-Dauphin M, Staple JK et al (1994) Overexpression of BCL-2 in Transgenic Mice Protects Neurons from Naturally Occurring Cell Death and Experimental Ischemia
30. Musa A, Ghoraie LS, Zhang SD et al (2018) A review of connectivity map and computational approaches in pharmacogenomics. *Brief Bioinform* 19:506–523
31. National Institute of Neurological Disorders and Stroke rt-PA Stroke Study Group (1995) Tissue plasminogen activator for acute ischemic stroke
32. Nakajima H, Kakui N, Ohkuma K, Ishikawa M, Hasegawa T (2005) A newly synthesized poly(ADP-ribose) polymerase inhibitor, DR2313 [2-methyl-3,5,7,8-tetrahydrothiopyrano[4,3-d]pyrimidine-4-one]: pharmacological profiles, neuroprotective effects, and therapeutic time window in cerebral ischemia in rats. *J Pharmacol Exp Ther* 312:472–481
33. Nighoghossian N, Ovize M, Mewton N, Ong E, Cho TH (2016) Cyclosporine A, a Potential Therapy of Ischemic Reperfusion Injury. A Common History for Heart and Brain. *Cerebrovasc Dis* 42:309–318
34. Ouk T, Laprais M, Bastide M, Mostafa K, Gautier S, Bordet R (2009) Withdrawal of fenofibrate treatment partially abrogates preventive neuroprotection in stroke via loss of vascular protection. *Vascul Pharmacol* 51:323–330
35. Qu Y, Zhang S, Qu Y et al (2020) Novel Gene Signature Reveals Prognostic Model in Acute Myeloid Leukemia. *Front Genet* 11:566024
36. Rui X, Shao S, Wang L, Leng J (2019) Identification of recurrence marker associated with immune infiltration in prostate cancer with radical resection and build prognostic nomogram. *BMC Cancer* 19:1179
37. Schwarzer R, Dörken B, Jundt F (2012) Notch is an essential upstream regulator of NF-kappaB and is relevant for survival of Hodgkin and Reed-Sternberg cells. *Leukemia* 26:806–813
38. Tang Y, Lu A, Aronow BJ, Sharp FR (2001) Blood genomic responses differ after stroke, seizures, hypoglycemia, and hypoxia: blood genomic fingerprints of disease. *Ann Neurol* 50:699–707
39. Tang Y, Lu A, Aronow BJ, Wagner KR, Sharp FR (2002) Genomic responses of the brain to ischemic stroke, intracerebral haemorrhage, kainate seizures, hypoglycemia, and hypoxia. *Eur J Neurosci* 15:1937–1952

40. Tramacere I, Boncoraglio GB, Banzi R et al (2019) Comparison of statins for secondary prevention in patients with ischemic stroke or transient ischemic attack: a systematic review and network meta-analysis. *BMC Med* 17:67
41. Uzdensky AB (2019) Apoptosis regulation in the penumbra after ischemic stroke: expression of pro- and antiapoptotic
42. Wu S, Wu B, Liu M et al (2019) Stroke in China: advances and challenges in epidemiology, prevention, and management. *Lancet Neurol* 18:394–405
43. Xu Wenjian, Li W (2019) Building a prediction model of brain tissues gene expression based on whole blood gene expression profiles
44. Yeh SH, Yang WB, Gean PW et al (2011) Translational and transcriptional control of Sp1 against ischaemia through a hydrogen peroxide-activated internal ribosomal entry site pathway. *Nucleic Acids Res* 39:5412–5423
45. Yao L, Kan EM, Kaur C et al (2013) Notch-1 signaling regulates microglia activation via NF-kappaB pathway after hypoxic exposure in vivo and in vitro. *PLoS ONE* 8:e78439
46. Yamase Y, Horibe H, Ueyama C et al (2015) Association of TOMM40 and SLC22A4 polymorphisms with ischemic stroke. *Biomed Rep* 3:491–498
47. Ye Z, Wang F, Yan F et al (2019) Bioinformatic identification of candidate biomarkers and related transcription factors in nasopharyngeal carcinoma. *World J Surg Oncol* 17:60
48. Yoo H, Kim J, Lee AR et al (2019) Alteration of microRNA 340-5p and Arginase-1 Expression in Peripheral Blood Cells during Acute Ischemic Stroke. *Mol Neurobiol* 56:3211–3221
49. Zhou M, Wang H, Zeng X et al (2019) Mortality, morbidity, and risk factors in China and its provinces, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet* 394:1145–1158

Tables

Table 1 Basic information of microarray Data

GEO Accession ID	Type	Platform	Tissue	Total of sample	Sample of ischemic stroke samples	Sample of controls
GSE16561	mRNA	GPL6883	Whole blood	63	39	24
GSE58294	mRNA	GPL570	Whole blood	46	23(24h)	23

Table 2 Targeted drugs predicted by the Hub gene in the DGldb database

No.	Gene	Medicine	Sources	PMIDs
1	ABCA1	simvastatin	PharmGKB	20031551
2	ABCA1	probucol	ChemblInteractions GuideToPharmacology TTD	19075491 19458386 15514211 1868500 20075420 19427182
3	ABCA1	fenofibrate	PharmGKB	-
4	ABCA1	cyclosporine	NCI	15084924
5	ABCA1	pravastatin	PharmGKB	19673941 20031551
6	ABCA1	atorvastatin	PharmGKB	20031551
7	SLC22A4	Imatinib	PharmGKB	23127916 22875622

Table 3 Small molecule compounds predicted by Hub gene in CMap database

No.	Gene	Score	Name	Description
1	ID3	99.21	2',5'-dideoxyadenosine	Adenylyl cyclase inhibitor
2	ID3	98.71	DR-2313	PARP inhibitor
3	ID3	98.33	otenzepad	Acetylcholine receptor antagonist
4	SLC22A4	-95.16	dubinitidine	Anti-epileptic

Figures

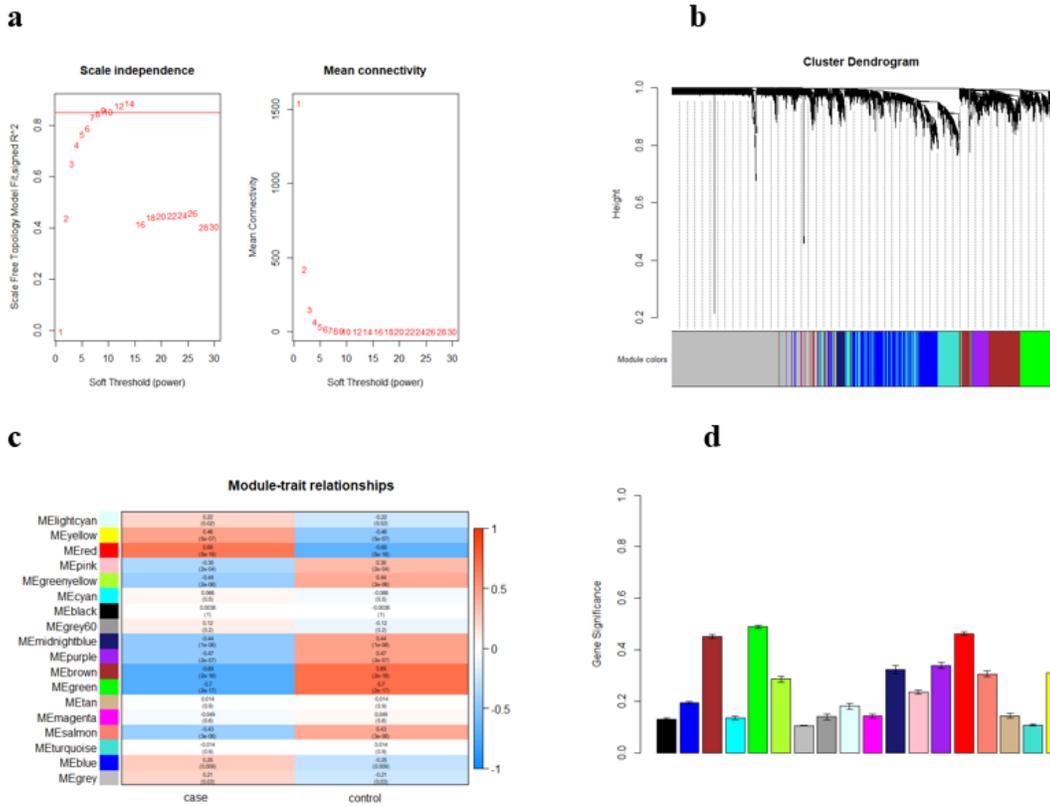
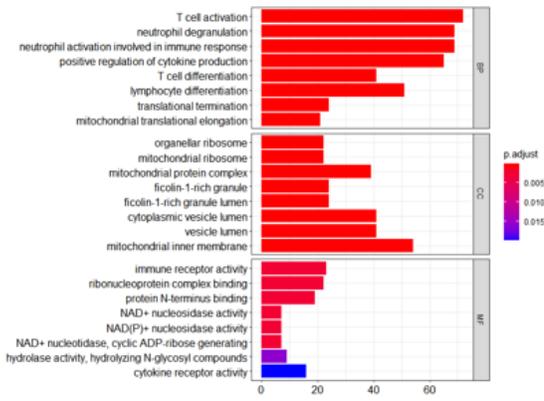


Figure 1

WGCNA analysis of genes after merging two data sets. (a) Power value screening of ischemic stroke co-expression module genes. (b) Gene Cluster Tree Based on different level clustering of adjacencies (Cluster Tree). (c) Heat maps of the case and control groups concerning each module. (d) Modular significance values of ischemic stroke-related co-expression modules.

a



b

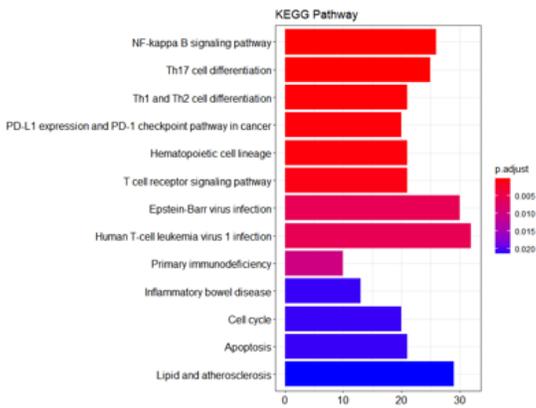


Figure 2

GO and KEGG analysis of modular genes associated with IS. (a) GO functional analysis of modular genes associated with IS. (b) KEGG pathway enrichment analysis of modular genes associated with IS.

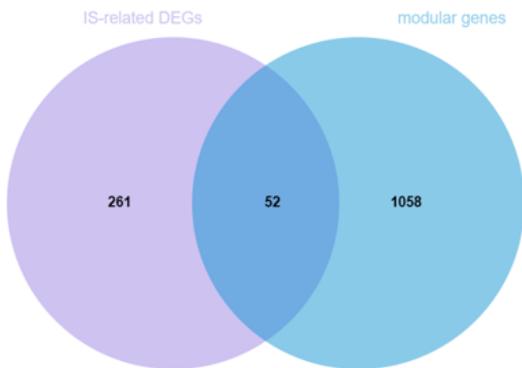


Figure 3

Integrated analysis of WGCNA and DEGs. Venn diagram of module genes and IS-related genes.

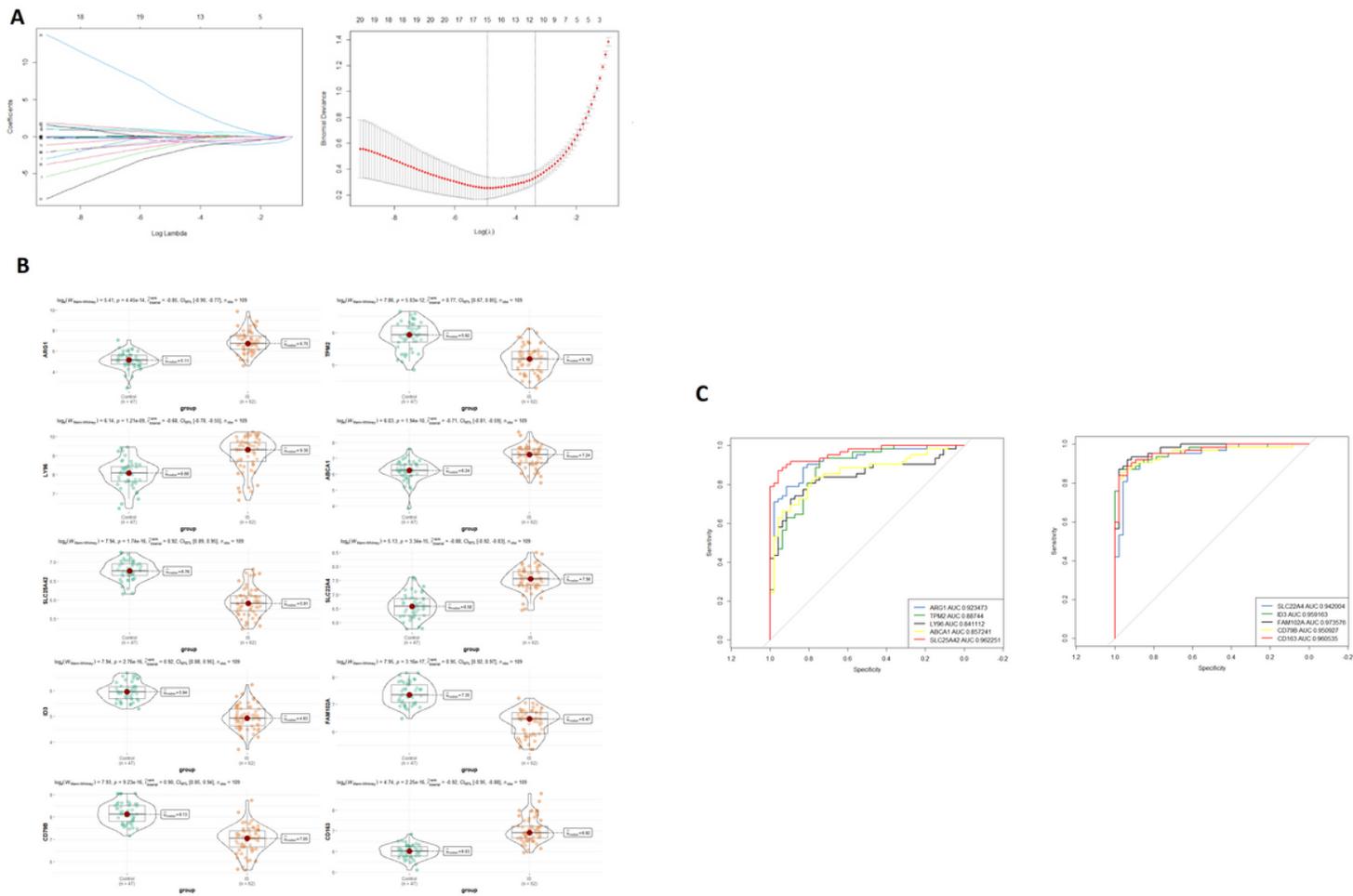


Figure 4

Screening of hub genes. (a) LASSO regression was performed to further screen hub genes. (b) Violin plots were performed for 10 hub genes. (c) ROC curves of 10 hub genes.

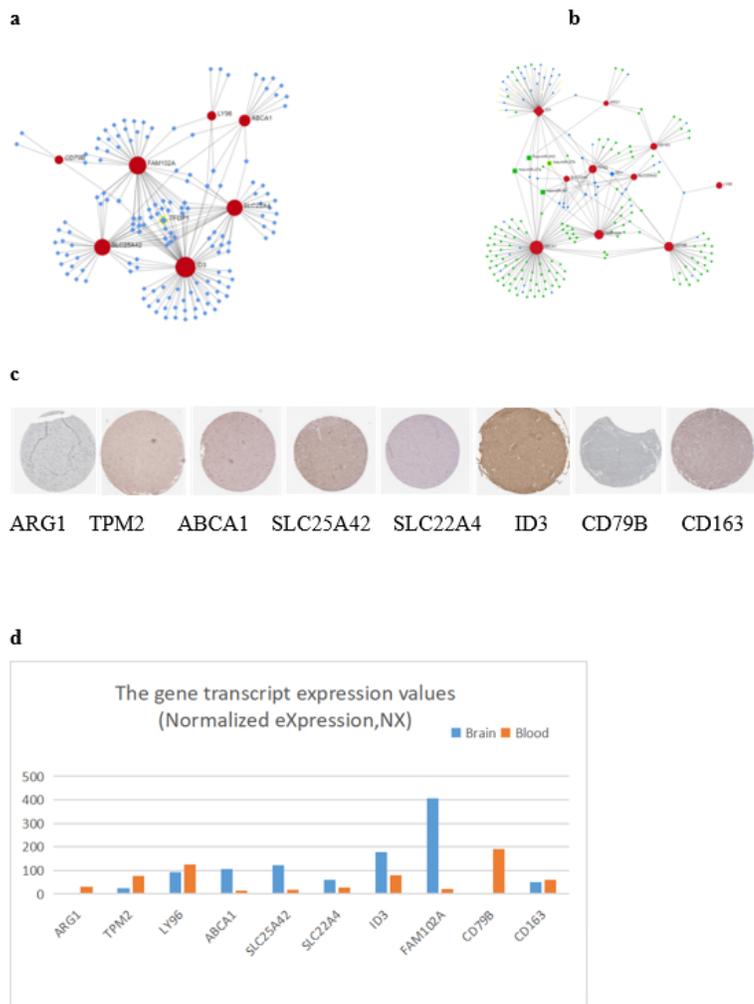


Figure 5
 the TF-gene interaction network, TF-miRNA regulatory network and the expression of hub genes. (a) Network for TF-gene interaction with hub genes. Red nodes represent hub genes and blue nodes represent TF genes. The network consists of 153 nodes and 221 edges. (b) Network for TF-miRNA coregulatory network with hub genes. Red nodes represent hub genes, green nodes represent miRNAs, blue nodes represent TF genes, and yellow nodes represent proteins. The network consists of 216 nodes and 272 edges. (c) immunohistochemical slices of 8 hub genes were found in the normal brain. (d) Expression levels of hub genes in the cerebral cortex and blood.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryInformation1.docx](#)
- [SupplementaryInformation2.docx](#)