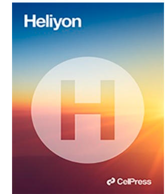
Heliyon 10 (2024) e29651

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon



Research article

Comprehensive analysis of lncRNA-miRNA-mRNA ceRNA network in ischemic stroke

Lin Cheng [a,1], Yun Zhao [a,1], Hong Ke [b,*]

[a] Department of Emergency, Shandong Provincial Third Hospital, Jinan, Shandong, 250031, China
[b] Department of Neurology, The Fourth People's Hospital of Jinan, Jinan, Shandong, 250031, China

ARTICLE INFO

Keywords:
Ischemic stroke
Differentially expressed RNAs
ceRNA network

ABSTRACT

Objective: Competitive endogenous RNA (ceRNA) networks have uncovered a novel mode of RNA interaction, and are implicated in various biological processes and the pathogenesis of IS. This study aimed to explore the potential mechanisms underlying the ceRNA network in IS.
Methods: Four public datasets containing lncRNA and mRNA (GSE22255 and GSE16561) and miRNA (GSE55937 and GSE43618) expression profiles from the GEO database were systematically analyzed to explore the role of RNAs in ischemic stroke (IS). Differentially expressed mRNAs (DEmRNAs), lncRNAs (DElncRNAs), and miRNAs (DEmiRNAs) between IS and normal control samples were identified. LncRNA-miRNA and miRNA-mRNA interactions were predicted, and the competing endogenous RNA (ceRNA) regulatory network was constructed using the Cytoscape software. The correlation between the RNAs in the ceRNA network and the clinical features of the samples was evaluated. Finally, principal component analysis was performed on the RNAs that constitute the ceRNA regulatory network, and their differential expression and principal component relationships among different types of samples were observed.
Results: A total of 224 DEmRNAs, 7 DEmiRNAs, and four DElncRNAs related to IS in four datasets were identified. Then, through target gene prediction, a lncRNA–miRNA–mRNA ceRNA network that contained 3 DElncRNAs, 2 DEmiRNAs, and 24 DEmRNAs was constructed. Correlations of the clinical characteristics showed that PART1 and SERPINH1 were related to clinical diseases, WNK1 was related to lifestyle, and seven RNAs were related to age. PCA results indicate that three principal components of PC1, PC2, and PC3 can clearly distinguish between control and IS samples.
Conclusion: Overall, we constructed a ceRNA network in IS, which could offer insights into the molecular mechanism and potential prognostic biomarkers for further research.

1. Introduction

Stroke is a destructive cerebrovascular event caused by the interruption of cerebral blood flow due to a blocked or burst blood vessel, which leads to physical disability and multiple functional damage [1]. Ischemic stroke (IS) accounts for 60–70 % of the total number of strokes and has the characteristics of high morbidity, high disability, high mortality, and a high recurrence rate [2]. IS occurs due to the interruption of the blood supply to the brain, which is usually due to an embolism, arterial thrombosis or changes in

* Corresponding author.
E-mail address: kehong862@163.com (H. Ke).
[1] Lin Cheng and Yun zhao contribute equally to the article.

https://doi.org/10.1016/j.heliyon.2024.e29651
Received 23 November 2023; Received in revised form 9 April 2024; Accepted 11 April 2024
Available online 20 April 2024
2405-8440/© 2024 Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

arteriosclerosis [3]. IS triggers complex processes, including neuroinflammation, oxidative stress, excitotoxicity, and apoptosis, leading to neuronal death [4]. IS can have multiple risk factors, including high blood pressure, high cholesterol, diabetes, smoking, obesity, physical inactivity, and a family history of stroke [5]. Other factors such as age, gender, race, and genetics can also contribute to the risk of IS [6]. The etiology and pathogenesis of IS are very complex, and the occurrence and development of IS cannot be fully explained. Therefore, it is urgent to find a more effective method of preventing IS.

Non-coding RNA (ncRNA) is a type of RNA that has no protein-coding function [7]. MicroRNAs (miRNAs) are a type of short (approximately 22 nt) ncRNA that can promote the degradation of targeted mRNAs and suppress translation, as well as negatively regulating gene expression at the post-transcriptional level [8,9]. Long non-coding RNAs (lncRNAs) are more than 200 nt in length and participate in a variety of physiological and pathological processes [10,11]. It has been reported that lncRNA is essential in brain development and may be related to brain damage [12]. The endogenous RNA (ceRNA) network combines the functions of protein-encoding mRNAs with the functions of ncRNAs [13]. When mRNA competes to bind miRNA, its stability is reduced, the translation process is blocked, gene expression is affected, and a variety of ncRNAs participate in the regulation of mRNA coding functions [14]. The ceRNA network plays a role in a variety of physiological and pathological processes.

In this study, the GEO database was used to identify DEmRNA, DElncRNA, and DEmiRNA between IS and control tissues, and a ceRNA network related to IS was constructed. This study aims to provide new insights on the role of IS-related RNAs and related ceRNA regulation mechanisms for the diagnosis and treatment of IS.

2. Materials and methods

2.1. Microarray datasets

The keywords "Ischemic Stroke, *Homo sapiens*" were used to search the NCBI GEO [15] (https://www.ncbi.nlm.nih.gov/) database for datasets that meet certain conditions. Datasets were required to meet the following criteria: (1) tissue samples are of the same type (the same entity/blood/other tissue samples). (2) samples have detected at least one lncRNA, mRNA or miRNA. (3) samples include IS disease and normal control samples, and each type of sample has at least two duplicate samples. According to these standards, four IS datasets (GSE22255, GSE16561, GSE55937 and GSE43618) were downloaded from the GEO database. The details of each microarray study are provided in Table 1.

2.2. Functional Re-annotation of lncRNAs and mRNAs

To obtain lncRNA and mRNA expression profiles, we re-annotated the obtained microarray data. Firstly, detailed annotation information of the Affymetrix human genome U133 Plus 2.0 array and the Illumina HumanRef-8 v3.0 expression bead chip platform were downloaded from Ensembl Genome Browser 96 (http://asia.ensembl.org/index.html). The annotation information included probe set ID, Ensemble ID, probe sequence, and other data. Probe sets were then assigned to the Ensemble gene IDs and GENCODE annotation was obtained. Human coding and non-coding gene annotation files from the Ensemble database were used to extract the matching information of the gene ID and gene symbol. When multiple probes are mapped to the same gene, the expression value of the mRNA or lncRNA was calculated by determining the average expression value of all corresponding probes. Finally, 1708 lncRNAs and 17326 mRNAs were obtained after re-annotation.

Differentially expressed mRNAs (DEmRNAs), miRNAs (DEmiRNAs) and lncRNAs (DElncRNAs) and functional enrichment analysis.

The Limma Version 3.34.0 [16] (https://bioconductor.org/packages/release/bioc/html/limma.html) in R3.6.1 language was used to screen differentially expressed RNAs (DERNAs) (i.e. DElncRNAs, DEmRNAs, and DEmiRNAs) for each dataset. The FDR value < 0.05 and $|\log2FC| > 0.5$ were used as the threshold for screening DElncRNAs, DEmRNAs, and DEmiRNAs. A Venn diagram was used to identify the intersection of DERNAs, based on DAVID version 6.8 [17,18] (https://david.ncifcrf.gov/) and the GO (biology process) and KEGG pathway enrichment analysis annotations for genes in the intersection DERNAs. *P*-value$<$0.05 was selected as the threshold for enrichment significance.

2.3. Construction of the ceRNA network

The lncRNA-miRNA-mRNA ceRNA network is based on the hypothesis that lncRNAs can regulate mRNA activity by directly interacting with miRNA sponges [19]. The construction of the ceRNA network involved two steps: (1) Relationships between the target DElncRNA and DEmiRNAs were searched from the DIANA-LncBasev2 database [20] (http://carolina.imis.athena-innovation.gr/diana_tools/web/index.php?r=lncbasev2%2Findex-experimental). (2) Then, the starBase Version 2.0 database [18] (http://

Table 1
Associated microarray datasets from the gene expression omnibus (GEO) database.

PMID	Record	Tissue	Platform	Normal	IS
22453632	GSE22255	IS	Affymetrix Human Genome U133 Plus 2.0 Array	20	20
20837969/28446746/29263821	GSE16561	IS	Illumina HumanRef-8 v3.0 expression beadchip	24	39
24911610	GSE55937	IS	Affymetrix Multispecies miRNA-3 Array	24	24
23559260	GSE43618	IS	Agilent-027233 *Homo sapiens* miRNA array	2	2

starbase.sysu.edu.cn/), which integrates the target gene prediction information from targetScan, picTar, RNA22, PITA, and miRanda databases, was used to predict the putative mRNA targets of DEmiRNAs that have a relationship with DElncRNA in step (1). The regulatory relationship contained in at least one of the databases was selected as the DEmiRNA-target gene pair. The mRNAs obtained in step (2) were mapped to the target genes regulated by DEmiRNAs to construct a miRNA-mRNA regulatory relationship. Finally, the DElncRNA-DEmiRNA and DEmiRNA-DEmRNA connection network was established, and the lncRNA-miRNA-mRNA ceRNA network was built and visualized using Cytoscape Version 3.6.1 (https://cytoscape.org/). Based on DAVID 6.8 online software, biological process and KEGG pathway enrichment annotation were performed for genes in the ceRNA network. *P*-value<0.05 was selected as the significance threshold.

2.4. Functional enrichment analysis

GO analysis and KEGG pathway enrichment of DEmRNAs were performed for annotation, visualization, and integrated discovery to screen the potential biological processes and signaling pathways in which the mRNAs are involved. A *P*-value <0.05 was considered to be significant.

2.5. Clinical feature analysis of the ceRNAs network

Clinical information related to the samples was extracted from the GSE22255 dataset, which extracts relatively more clinical data than the other datasets. Hmisc Version 4.4-1 in R3.6.1 language was used to calculate the correlation between the expression levels of intersecting DElncRNAs, DEmiRNAs, and DEmRNAs in the ceRNA network and various clinical features. *P*-value<0.05 was considered statistically significant.

2.6. Principal component analysis for the ceRNAs network

Principal component analysis (PCA) is a statistical process that uses orthogonal transformation to transform a set of possibly related variable observations into a set of linearly independent variable values called principal components [21]. PCA is widely used for data processing and dimensionality reduction [22]. The contribution rate of 80 % or more of the main component is considered significant. The R3.6.1 language psychology software package version 1.7.8 was used to perform PCA on the intersecting DElncRNA, DEmiRNA, and DEmRNA in the ceRNA network. In addition, sample identification based on the expression levels of RNAs constituting the ceRNA regulatory network was utilized using the receiver operating characteristic (ROC) curve method in the R language pROC. ROC curves are typically used to determine the ability of a specific factor to identify a sample, and are generally demarcated by 0.7, with an the area under receiver operating characteristic curve (AUC) higher than 0.7 indicating a better ability to identify a sample.

2.7. Sample collection

This study was approved by the ethical committee of The Fourth People's Hospital of Jinan. Prior to the study, written informed consent was obtained from all patients. The peripheral blood samples were obtained from six patients with ischemic stroke. Six healthy individuals were used as control samples. Clinical characteristics of the two groups of samples were comparable.

2.8. Key RNA expression validation

To further validate the expression levels of the ceRNAs with large differential expression in the lncRNA-miRNA-mRNA network, we

Table 2
Primers for quantitative real-time PCR (qRT-PCR).

Gene names	Primer sequence (5′–3′)
has-miR-24-RT	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGACCTGTTC
has-miR-24-F	GCGTGGCTCAGTTCAGCAG
has-miR-19a-RT	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGACTCAGTT
has-miR-19a-F	GCCGTGTGCAAATCTATGCAA
U6–F	CTCGCTTCGGCAGCACA
U6-R	AACGCTTCACGAATTTGCGT
PART1-F	TCACCCTACCACCACCCA
PART1-R	GAGGCTCTTTCCACTATCTTCA
WNK1–F	GCCGTCAGATCCTTAAAGGTC
WNK1-R	CCAGTAGGGCCGGTGATAA
PINK1-AS-F	AGATGGAGGTGTTGGGTTAT
PINK1-AS-R	AGAGGTGAGTCCCGAGTG
SERPINH1-F	GCCGTCAGATCCTTAAAGGTC
SERPINH1-R	CCAGTAGGGCCGGTGATAA
GAPDH-F	TGACAACTTTGGTATCGTGGAAGG
GAPDH-R	AGGCAGGGATGATGTTCTGGAGAG

used two datasets from the GEO database: GSE180470 and GSE199942. These two datasets both performed RNA sequencing on IS and CTRL group samples using the Illumina HiSeq platform (*Homo sapiens*). GSE180470 included 3 IS group and 3 CTRL group samples, using the Illumina HiSeq 4000 platform. GSE199942 included 5 IS group and 5 CTRL group samples, using the Illumina HiSeq 2500 platform.

Then, the quantitative real-time PCR (qRT-PCR) was also performed to further verify the expression of the ceRNAs with large differential expression in the lncRNA-miRNA-mRNA network. Briefly, RNA was extracted from blood samples of IS disease and normal controls using a TRIzol reagent (Thermo Fisher, Waltham, MA, USA). Total RNA was reversed to cDNA using a PrimeScriptTM RT Reagent kit (TaKaRa) following the manufacturer's protocol. Quantitative PCR analysis was conducted utilizing an ABI Prism 7300HT Real-Time PCR amplification system. Subsequent data analysis was performed using the $2^{-\Delta\Delta Ct}$ method to determine relative gene expression levels. The sequences are shown in Table 2.

3. Results

3.1. Identification of DERNAs

In the GSE22255, GSE16561, GSE55937, and GSE43618 data sets, 1708 lncRNAs and 17326 mRNAs were annotated. Then, the R3.6.1 limma software package was used to screen DERNAs. In the GSE22255 and GSE16561 datasets, 1489 DERNAs (including 114 DElncRNAs and 1375 DEmRNAs) and 997 DERNAs (including nine DElncRNAs and 988 DEmRNAs) were identified, respectively (Fig. 1A, Tables S1 and S2). Meanwhile, 63 and 48 DEmiRNAs were identified in the GSE55937 and GSE43618 data sets, respectively (Fig. 1C). Then, according to the intersection of DEmRNA and DElncRNA in the GSE22255 and GSE16561 data sets, 373 overlapping DERNAs were determined, of which 228 overlapping DERNAs (four DElncRNAs and 224 DEmRNAs) were significantly different in the same direction (Fig. 1B). We also took the intersection of the DEmiRNAs of the two data sets and identified 11 overlapping DEmiRNAs, of which seven DEmiRNAs were significantly different in the same direction (Fig. 1D).

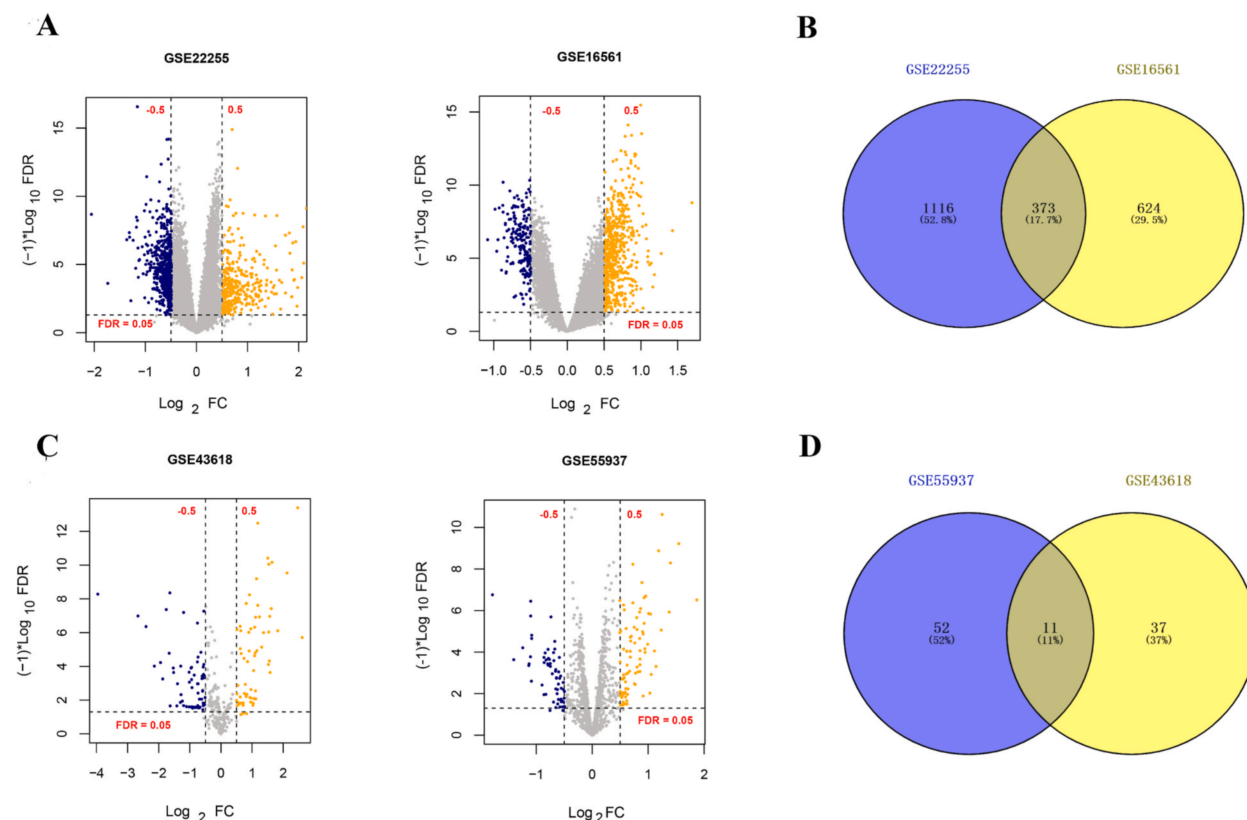


Fig. 1. Volcano plots of the distributions of DERNAs in different datasets. A: Volcano plots of the distributions of DElncRNAs and DEmRNAs in the GSE22255 and GSE16561 datasets; orange plot represented upregulation, and blue plot represented downregulation. B: Venn diagram of the DERNAs among the lncRNA and mRNA expression datasets GSE22255 and GSE16561. C: Volcano plots of the distributions of DEmiRNAs in the GSE43618 and GSE55937 datasets; orange plot represented upregulation, and blue plot represented downregulation. D: Venn diagram of the DEmiRNAs among the miRNA expression datasets GSE43618 and GSE55937.

3.2. Functional analysis of DEmRNAs

The biological process and enrichment annotation analysis of the KEGG signal pathway were performed on the DEmRNAs with the same significant difference direction (Table 3). The results indicated that these DEmRNAs are mainly related to biological pathways such as "mitochondrial translational elongation", "mitochondrial translational termination" and "negative regulation of cell proliferation". KEGG pathway analysis showed that the most important pathways were "Spliceosome", "TNF signaling pathway", "Alzheimer's disease", and "Ribosome".

3.3. Construction of the ceRNA network

First of all, the DIANA-LncBasev2 database was searched for lncRNA-miRNA interactions. A total of three DElncRNA-DEmiRNA interaction pairs were determined. Then, the starBase Version 2.0 database was used to predict the target mRNAs of miRNAs in DElncRNA-DEmiRNA interaction pairs. The mRNAs predicted by DEmiRNAs in the starBase database were mapped to the target genes regulated by DEmiRNAs to construct a miRNA-mRNA regulatory relationship. Finally, 24 miRNA-mRNA interaction pairs were obtained. According to the interaction between lncRNA-miRNA and miRNA-mRNA, a ceRNA network composed of three DElncRNAs, two DEmiRNAs, and 24 DEmRNAs was constructed (Fig. 2A and B).

3.4. Functional analysis of the ceRNA network-associated DEmRNAs

Next, the potential biological processes and KEGG pathways of 26 DEmRNAs in the ceRNA network were analyzed (Fig. 3). The GO function enrichment analysis was carried out using DAVID database and nine critical GO biological processes were identified. The results showed that DEmRNAs were significantly enriched in the biology process of "GO:0070125: mitochondrial translational elongation", "GO:0070126: mitochondrial translational termination" and "GO:0006415: translational termination". The results indicate that the most significant KEGG pathway for DEmRNAs is "hsa03040: Spliceosome".

3.5. Validation of the ceRNA network RNAs

Two lncRNAs (PART1 and PINK1-AS), two miRNAs (has-miR-24 and has-miR-19a) and two mRNAs (WNK1 and SERPINH1) were selected for validation by datasets (Fig. 4A) and qPCR (Fig. 4B). The expression of the six selected RNAs was consistent with the results in the ceRNA network (Fig. 4). has-miR-24, PART1 and *WNK1* showed up-regulated expression in the ceRNA network, while has-miR-19a, PINK1-AS and *SERPINH1* showed down-regulated expression in the ceRNA network, which was confirmed by qRT-PCR, GSE180470 and GSE199942 datasets.

3.6. Clinical feature analysis of ceRNAs

The correlation between DElncRNAs, DEmiRNAs, and DEmRNAs involved in the ceRNA network and clinical features, including clinical diseases (high blood sugar, hypertension, and high blood cholesterol), lifestyle (smoking and drinking), and age, in the GSE22255 dataset were analyzed (Table S1). The results showed that *PART1* and *SERPINH1* were significantly related to clinical diseases (high blood sugar, hypertension and high blood cholesterol) (*PART1*: *P*-value = 0.01378, Cor. = 0.3864, *SERPINH1*: *P*-value = 0.03782, Cor. = −0.3296); *WNK1* was significantly related to lifestyle (smoking and drinking) (*P*-value = 0.04661, Cor. = −0.2935); *M6PR*, *IL1RAP*, *C12orf43*, *ITM2C*, *WDR26*, *ANKRD46,* and *GOSR2* were significantly related to age (*P*-value <0.05). *M6PR*, *C12orf43*, *ITM2C,* and *ANKRD46* are negatively correlated with age, indicating that the risk of IS decreases with age. *IL1RAP*, *WDR26,* and *GOSR2* are positively correlated with age, indicating that the risk of IS increases with age (Table S3).

Table 3
GO enrichment analysis and KEGG pathway analysis of DEmRNAs with the same significant difference direction in IS.

Category	Term	Count	PValue
Biology Process	GO:0070125~mitochondrial translational elongation	8	7.57E-05
	GO:0070126~mitochondrial translational termination	8	8.15E-05
	GO:0006415~translational termination	3	2.11E-03
	GO:0008285~negative regulation of cell proliferation	12	8.95E-03
	GO:0032259~methylation	5	1.18E-02
	GO:0000398~mRNA splicing, via spliceosome	8	1.85E-02
	GO:2000349~negative regulation of CD40 signaling pathway	2	2.40E-02
	GO:0030488~tRNA methylation	3	3.91E-02
	GO:0006369~termination of RNA polymerase II transcription	4	4.24E-02
	GO:0006366~transcription from RNA polymerase II promoter	12	4.74E-02
	GO:0048662~negative regulation of smooth muscle cell proliferation	3	4.77E-02
KEGG Pathway	hsa03040:Spliceosome	7	5.21E-03
	hsa04668:TNF signaling pathway	5	3.99E-02
	hsa05010:Alzheimers disease	6	4.52E-02
	hsa03010:Ribosome	5	4.81E-02

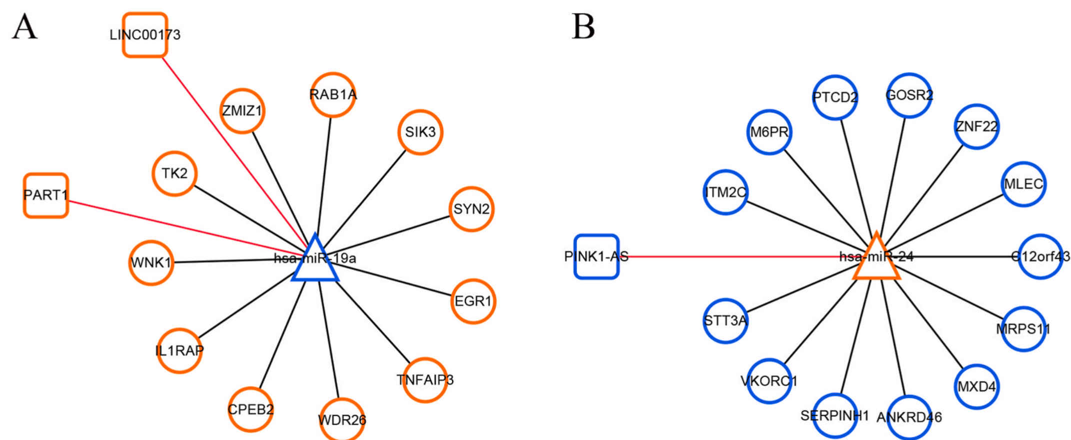


Fig. 2. Construction of the ceRNA regulatory network. A. LncRNA-has-miR-19a-mRNA ceRNA network. B. LncRNA-has-miR-24 -mRNA ceRNA network. Squares indicate lncRNAs, triangles indicate miRNAs, and circles indicate mRNA. Orange and blue represent RNAs that are consistently up-regulated and down-regulated in the data set, respectively.

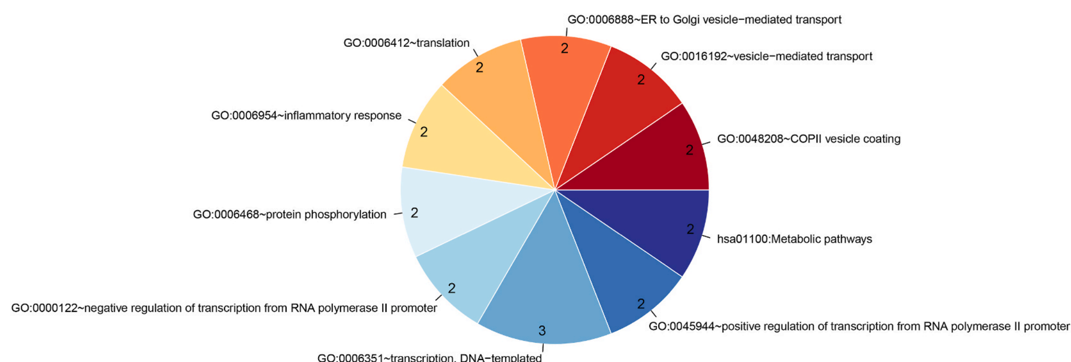


Fig. 3. Functional analysis of DEmRNAs in the ceRNA network. The graph displays only significantly enriched biological process terms and KEGG terms ($P < 0.05$). The number indicates the number of genes involved, and the color indicates significance. The closer the color is to red, the greater the significance.

4. Results of PCA for ceRNAs

To further explore the characteristics of the expression level of RNAs in the ceRNA network, PCA was performed on the DElncRNA and DEmRNA in the ceRNA network. Through the calculation and analysis of the expression values of DERNAs that constitute the ceRNA network in the GSE22255 and GSE16561 data sets, a total of ten principal components (PC) were obtained through fitting (Fig. 5A and B). Among them, the cumulative contribution rate of PC1, PC2, and PC3 was greater than 80 %, which indicates that PC1, PC2, and PC3 contain a lot of vital information about the original variables (RNAs expression value).

To validate the ability of the first three principal components (PC1, PC2, and PC3) to reflect the characteristics of each dataset sample, we plotted the three-dimensional sample graphs and the sample classification ROC curves based on these three principal components. As shown in Fig. 5C and D, the AUC values of the ROC curves for the GSE22255 and GSE16561 datasets were 0.920 and 0.905, respectively, both higher than 0.7, indicating that using only these three principal components could effectively distinguish the samples of the control group and the IS disease group, suggesting that these three principal components contained most of the information of the original variable RNAs expression.

However, this study still has some limitations. First, the potential lncRNA–miRNA–mRNA interactions were identified by bioinformatics analysis. The target relationships and functions of the lncRNA–miRNA–mRNA interactions in IS development lack experimental validation at the molecular and cellular levels. Secondly, given the intricate and multifaceted nature of IS, it is prudent to acknowledge that outcomes derived from an exclusive focus on blood tissue may not comprehensively encapsulate the entirety of the disease pathophysiology.

5. Discussion

In this study, four microarray data sets downloaded from the GEO database were fully analyzed. Firstly, DEmRNAs, DEmiRNAs, and

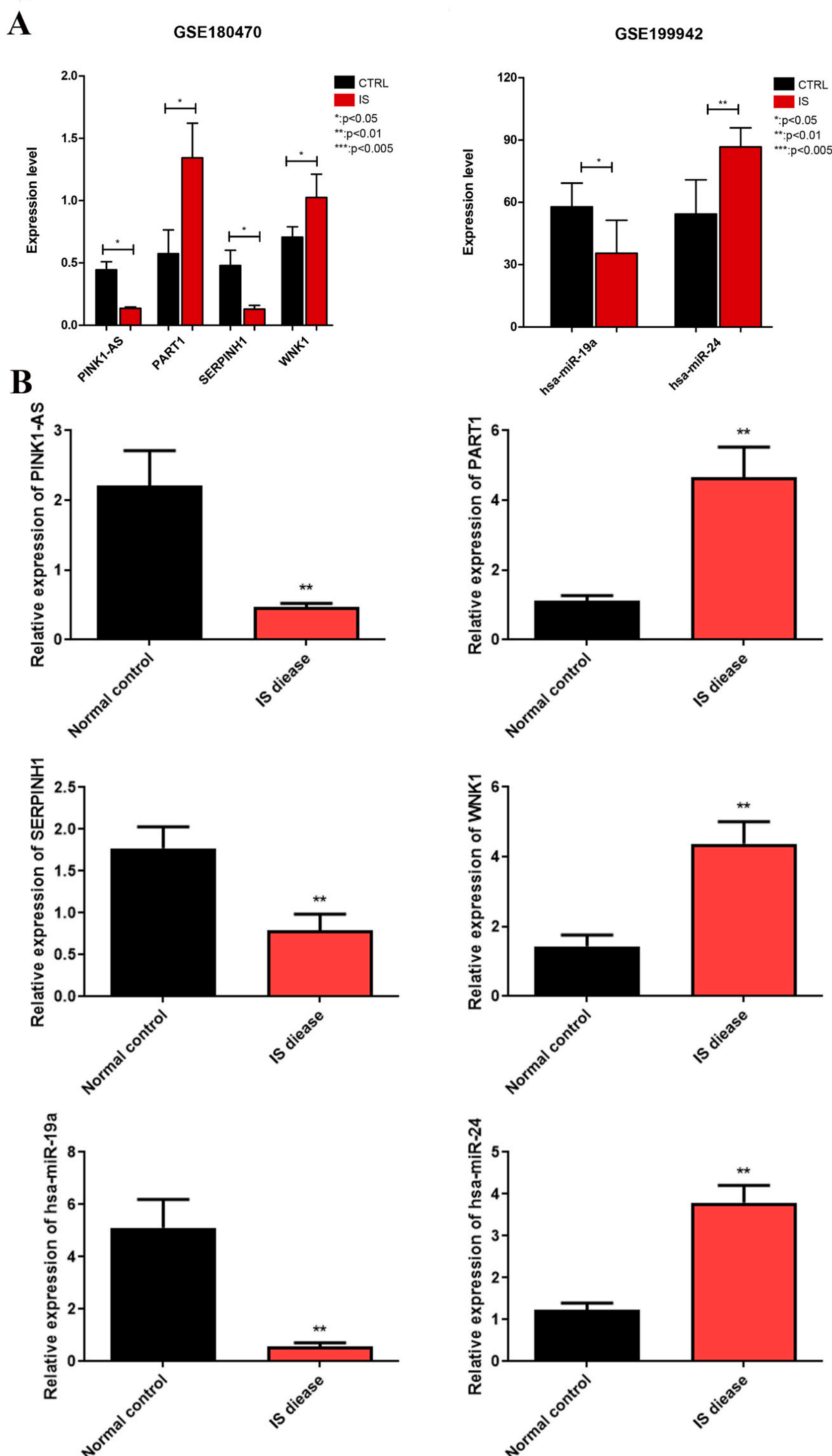


Fig. 4. RNA expression in the IS group compared with the normal control samples. A. Validated by GSE180470 and GSE199942 datasets. B. Validated by RT-PCR. $^{**}P < 0.01$ versus normal control group.

DElncRNAs were identified from four data sets, and the IS-related ceRNA network was constructed. Then, the mRNA in the network was enriched by GO and KEGG pathways. Through the analysis of the correlation between DERNAs in the ceRNA network and clinical features, it was found that *PART1* and *SERPINH1* were related to clinical diseases (high blood sugar, hypertension and high blood cholesterol), *WNK1* was related to lifestyle, and seven RNAs, including *M6PR*, *IL1RAP*, and *C12orf43*, were significantly related to age.

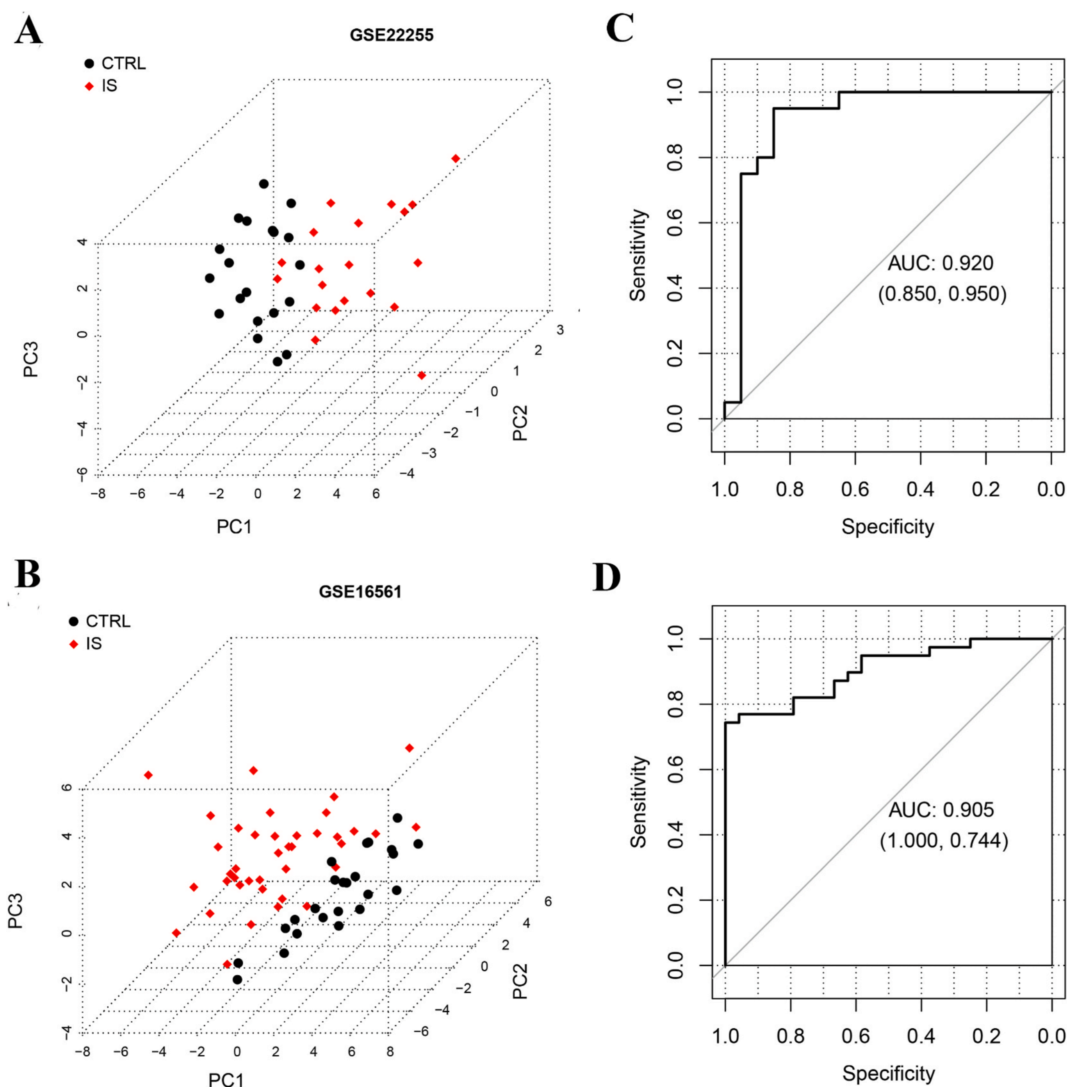


Fig. 5. The 3D result of the principal component analysis for DERNAs in GSE22255 (A) and GSE16561 (B) datasets. The area under the curve (AUC) of the ROC curve in the GSE22255 (C) and GSE16561 (D) datasets. Black dots represent normal control samples, and red dots represent IS samples. The X, Y, and Z axes represent PC1, PC2, and PC3, respectively.

In addition, PCA was performed on DERNAs in the ceRNA network to find that three principal components of PC1, PC2, and PC3 can clearly distinguish between control and IS samples.

The DEmRNAs related to the ceRNA network were analyzed using GO and KEGG. Through GO annotation, it is found that DEmRNAs may be involved in several biological processes, including "GO:0070125: mitochondrial translational elongation", "GO:0070126: mitochondrial translational termination", and "GO:0006415: translational termination". DEmRNAs in the ceRNA network were significantly enriched in the "hsa03040: Spliceosome". In fact, more and more evidence shows that these enrichment pathways are related to IS. Mitochondrial dysfunction is one of the hallmarks of IS, contributing to the pathophysiology of ischemia and reperfusion [23]. Andrabi et al. found that pramipexole prevents ischemic cell death via mitochondrial pathways in ischemic stroke [24]. In addition, Agarwal et al. showed that spliceosome and synaptic transmission are involved in the pathology of stroke [25]. Therefore, enrichment findings could imply that the lncRNA-miRNA-mRNA ceRNA network plays a crucial role in the development of IS via these pathways.

In this study, we found that *PART1* and *SERPINH1* were related to clinical diseases (high blood sugar, hypertension and high blood cholesterol), *WNK1* was related to lifestyle, and seven RNAs including *M6PR*, *IL1RAP*, and *C12orf43* were significantly related to age. *SERPINH1* is a collagen-specific chaperone, which is capable of co-translational folding and assembly of procollagens in the endoplasmic reticulum [26]. It has been reported that *SERPINH1* is closely related with cerebrovascular disease [27,28]. Lee et al. found that *SERPINH1* is expressed in cerebrovascular smooth muscle cells and is related to atherosclerosis [29]. The expression of *SERPINH1* has been found to change in the acute phase of stroke, indicating its targeting effect in the combined biomarkers covering environment

and genetic factors [30]. Previous studies have shown that SERPINH1 is down-regulated in IS. LncRNA PART1 (ENST00000152931) is a poorly characterized long non-coding RNA located on chromosome 5: 59783540-59,843,484. Lin et al. first identified it as a regulator of the androgen receptor-regulated gene network in the human prostate, which may be involved in the pathogenesis of prostate cancer [31]. Recent studies have shown that lncRNA PART1 protects against myocardial ischemia/reperfusion injury by inhibiting miR-503-5p/BIRC5 mediated mitochondrial apoptosis [32]. The results of this study are consistent with previous studies, indicating that *SERPINH1* may be a potential therapeutic target for IS. In addition, Yagil et al. found that *SERPINH1* is associated with hypertension [33]. *SERPINH1* is a member of the heat shock protein family, which has been implicated in hypertension [34]. Hypertension is the most common modifiable risk factor for stroke in the stroke population [33]. In this study we found that SERPINH1 was related to clinical diseases, high blood sugar, hypertension and high blood cholesterol, indicating that SERPINH1 and PART1 may play an important role in IS by regulating these three systems.

WNK1 is one of the four serine-threonine protein kinase family members [35]. *WNK1* is found in dorsal root ganglion cell bodies, the developing nervous system, and brain vasculature [36]. Studies have found that high expression of *WNK1* has been linked to variations in blood pressure and increased vulnerability to hypertension, a risk factor for stroke [37,38]. Consistent with results from previous studies, in this study we also found that *WNK1* is up-regulated in IS. We also found that *WNK1* is related to lifestyle (such as smoking and drinking). Drinking is one of the most important factors affecting the incidence of stroke. A previous study found that alcohol consumption was found to be negatively related to non-fatal CHD risk but positively related to the risk of various stroke subtypes [39]. Accordingly, the 2017 update of The Canadian Stroke Best Practice Recommendations for the Secondary Prevention of Stroke provides guidance for the prevention of IS recurrence through the identification and management of modifiable vascular risk factors (such as lifestyle, smoking, and hypertension) [40]. These results indicate that *WNK1* may play an important role in the prevention of IS.

M6PR (Mannose-6-Phosphate Receptor) is a protein-coding gene. It has been reported that *M6PR* plays a role in protein endocytosis and intracellular transport [41]. *IL1RAP* is an important part of the receptor complex in mediating the immune response to IL-1 family cytokines, and organizes neuronal synapse formation in the brain [42]. Studies have found that IL1RAP regulates neuronal gene expression and IL-1 induces neuroinflammatory responses [43]. Inflammation plays an important role in the pathogenesis of IS and other forms of ischemic brain injury [44]. Systemic inflammation could influence the susceptibility of patients to stroke and their subsequent prognosis [45]. *ITM2C* is a protein-coding gene, which is mainly related to cerebral amyloid angiopathy and Itm2b-related cerebral amyloid angiopathy. The *ITM2C* genotype was a significant and independent determinant of subarachnoid hemorrhage, indicating that *ITM2C* is closely related to stroke [46]. *M6PR* and *ITM2C* are down-regulated in IS, whilst *IL1RAP* is up-regulated in IS. Furthermore, our study also found that *M6PR*, *ITM2C,* and *IL1RAP* are related to age. Aging is a major risk factor for stroke, stroke-related mortality, and post-stroke complications [47]. It is reported that the incidence of IS increases with age [48]. These results indicate that *M6PR*, *ITM2C,* and *IL1RAP* may play an important role in IS by regulating risk factors (age) of IS.

6. Conclusion

In short, this study established a ceRNA regulatory network related to IS; three DElncRNAs, two DEmiRNAs, and 24 DEmRNAs were determined. Moreover, a clinical information correlation analysis and PCA were performed to further determine the main regulatory factors in the network. These may help to explore the biological mechanism of IS and identify potential therapeutic targets for IS. Nevertheless, molecular experiments are required to further verify the specific pathogenesis and molecular targets of IS.

Funding

None.

Data accessibility

The data that support the findings of this study are publicly available in the GEO database (https://www.ncbi.nlm.nih.gov/geo/) under the accession number GSE22255, GSE16561, GSE55937 and GSE43618.

CRediT authorship contribution statement

Lin Cheng: Writing – original draft, Data curation, Conceptualization. **Yun Zhao:** Writing – original draft, Resources, Formal analysis, Conceptualization. **Hong Ke:** Writing – review & editing, Project administration, Methodology, Formal analysis.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2024.e29651.

References

[1] D. Lloyd-Jones, R. Adams, M. Carnethon, G. De Simone, T.B. Ferguson, K. Flegal, E. Ford, K. Furie, A. Go, K. Greenlund, N. Haase, S. Hailpern, M. Ho, V. Howard, B. Kissela, S. Kittner, D. Lackland, L. Lisabeth, A. Marelli, M. McDermott, J. Meigs, D. Mozaffarian, G. Nichol, C. O'Donnell, V. Roger, W. Rosamond, R. Sacco, P. Sorlie, R. Stafford, J. Steinberger, T. Thom, S. Wasserthiel-Smoller, N. Wong, J. Wylie-Rosett, Y. Hong, Heart disease and stroke statistics–2009 update: a report from the American heart association statistics committee and stroke statistics subcommittee, Circulation 119 (3) (2009) 480–486.

[2] R. Rodrigo, R. Fernández-Gajardo, R. Gutiérrez, J.M. Matamala, R. Carrasco, A. Miranda-Merchak, W. Feuerhake, Oxidative stress and pathophysiology of ischemic stroke: novel therapeutic opportunities, CNS Neurol. Disord.: Drug Targets 12 (5) (2013) 698–714.

[3] K.V. Bulygin, N.M. Beeraka, A.R. Saitgareeva, V.N. Nikolenko, I. Gareev, O. Beylerli, L.R. Akhmadeeva, L.M. Mikhaleva, L.F. Torres Solis, A. Solís Herrera, M. F. Avila-Rodriguez, S.G. Somasundaram, C.E. Kirkland, G. Aliev, Can miRNAs Be considered as diagnostic and therapeutic molecules in ischemic stroke pathogenesis?-current status, Int. J. Mol. Sci. 21 (18) (2020).

[4] S.E. Khoshnam, W. Winlow, M. Farzaneh, Y. Farbood, H.F. Moghaddam, Pathogenic mechanisms following ischemic stroke, Neurol. Sci. 38 (7) (2017) 1167–1186.

[5] A.S. Alharbi, M.S. Alhayan, S.K. Alnami, R.S. Traad, M.A. Aldawsari, S.A. Alharbi, A.O.A. Sharif, S.T. Alboqami, S.A. Alshammari, M.K.S. Alzeer, Epidemiology and risk factors of stroke, Arch. Pharm. Pract. 10 (4) (2019).

[6] A.K. Boehme, C. Esenwa, M.S. Elkind, Stroke risk factors, genetics, and prevention, Circ. Res. 120 (3) (2017) 472–495.

[7] M. Esteller, Non-coding RNAs in human disease, Nat. Rev. Genet. 12 (12) (2011) 861–874.

[8] V.N. Kim, J.W. Nam, Genomics of microRNA, Trends Genet. 22 (3) (2006) 165–173.

[9] J. Li, S. Tan, R. Kooger, C. Zhang, Y. Zhang, MicroRNAs as novel biological targets for detection and regulation, Chem. Soc. Rev. 43 (2) (2014) 506–517.

[10] A. Bhan, S.S. Mandal, Long noncoding RNAs: emerging stars in gene regulation, epigenetics and human disease, ChemMedChem 9 (9) (2014) 1932–1956.

[11] X. Li, Z. Wu, X. Fu, W. Han, Long noncoding RNAs: insights from biological features and functions to diseases, Med. Res. Rev. 33 (3) (2013) 517–553.

[12] G. Barry, Integrating the roles of long and small non-coding RNA in brain function and disease, Mol Psychiatry 19 (4) (2014) 410–416.

[13] M.S. Ebert, J.R. Neilson, P.A. Sharp, MicroRNA sponges: competitive inhibitors of small RNAs in mammalian cells, Nat. Methods 4 (9) (2007) 721–726.

[14] Y. Cheng, Y. Su, S. Wang, Y. Liu, L. Jin, Q. Wan, Y. Liu, C. Li, X. Sang, L. Yang, C. Liu, Z. Wang, Identification of circRNA-lncRNA-miRNA-mRNA competitive endogenous RNA network as novel prognostic markers for acute myeloid leukemia, Genes 11 (8) (2020).

[15] T. Barrett, D.B. Troup, S.E. Wilhite, P. Ledoux, D. Rudnev, C. Evangelista, I.F. Kim, A. Soboleva, M. Tomashevsky, R. Edgar, NCBI GEO: mining tens of millions of expression profiles–database and tools update, Nucleic Acids Res. 35 (Database issue) (2007) D760–D765.

[16] M.E. Ritchie, B. Phipson, D. Wu, Y. Hu, C.W. Law, W. Shi, G.K. Smyth, Limma powers differential expression analyses for RNA-sequencing and microarray studies, Nucleic Acids Res. 43 (7) (2015) e47.

[17] W. Huang da, B.T. Sherman, R.A. Lempicki, Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists, Nucleic Acids Res. 37 (1) (2009) 1–13.

[18] W. Huang da, B.T. Sherman, R.A. Lempicki, Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources, Nat. Protoc. 4 (1) (2009) 44–57.

[19] N. Landeros, P.M. Santoro, G. Carrasco-Avino, A.H. Corvalan, Competing endogenous RNA networks in the epithelial to mesenchymal transition in diffuse-type of gastric cancer, Cancers 12 (10) (2020).

[20] M.D. Paraskevopoulou, I.S. Vlachos, D. Karagkouni, G. Georgakilas, I. Kanellos, T. Vergoulis, K. Zagganas, P. Tsanakas, E. Floros, T. Dalamagas, A. G. Hatzigeorgiou, DIANA-LncBase v2: indexing microRNA targets on non-coding transcripts, Nucleic Acids Res. 44 (D1) (2016). D231-238.

[21] C.C. David, D.J. Jacobs, Principal component analysis: a method for determining the essential dynamics of proteins, Methods Mol. Biol. 1084 (2014) 193–226.

[22] E. Alizadeh, S.M. Lyons, J.M. Castle, A. Prasad, Measuring systematic changes in invasive cancer cell shape using Zernike moments, Integr. Biol. 8 (11) (2016) 1183–1193.

[23] Z. He, N. Ning, Q. Zhou, S.E. Khoshnam, M. Farzaneh, Mitochondria as a therapeutic target for ischemic stroke, Free Radic. Biol. Med. 146 (2020) 45–58.

[24] S.S. Andrabi, M. Ali, H. Tabassum, S. Parveen, S. Parvez, Pramipexole prevents ischemic cell death via mitochondrial pathways in ischemic stroke, Dis Model Mech 12 (8) (2019).

[25] A. Agarwal, S. Park, S. Ha, J.S. Kwon, M.R. Khan, B.G. Kang, T.M. Dawson, V.L. Dawson, S.A. Andrabi, S.U. Kang, Quantitative mass spectrometric analysis of the mouse cerebral cortex after ischemic stroke, PLoS One 15 (4) (2020) e0231978.

[26] Y. Kishimoto, M. Yamashita, A. Wei, Y. Toya, S. Ye, C. Kendziorski, N.V. Welham, Reversal of vocal fold mucosal fibrosis using siRNA against the collagen-specific chaperone Serpinh1, Mol. Ther. Nucleic Acids 16 (2019) 616–625.

[27] Y. Li, G. Deng, Y. Qi, H. Zhang, L. Gao, H. Jiang, Z. Ye, B. Liu, Q. Chen, Bioinformatic profiling of prognosis-related genes in malignant glioma microenvironment, Med Sci Monit 26 (2020) e924054.

[28] S.P. Pantazatos, Y.Y. Huang, G.B. Rosoklija, A.J. Dwork, V. Arango, J.J. Mann, Whole-transcriptome brain expression and exon-usage profiling in major depression and suicide: evidence for altered glial, endothelial and ATPase activity, Mol Psychiatry 22 (5) (2017) 760–773.

[29] S.J. Lee, S. Kwon, J.R. Gatti, E. Korcari, T.E. Gresser, P.C. Felix, S.G. Keep, K.C. Pasquale, T. Bai, S.A. Blanchett-Anderson, N.W. Wu, C. Obeng-Nyarko, K. M. Senagbe, K.C. Young, S. Maripudi, B.C. Yalavarthi, D. Korcari, A.Y. Liu, B.C. Schaffler, R.F. Keep, M.M. Wang, Large-scale identification of human cerebrovascular proteins: inter-tissue and intracerebral vascular protein diversity, PLoS One 12 (11) (2017) e0188540.

[30] Y. Wang, J. Liu, H. Wu, Y. Cai, Combined biomarkers composed of environment and genetic factors in stroke, Biosci Trends 12 (4) (2018) 360–368.

[31] B. Lin, J.T. White, C. Ferguson, R. Bumgarner, C. Friedman, B. Trask, W. Ellis, P. Lange, L. Hood, P.S. Nelson, PART-1: a novel human prostate-specific, androgen-regulated gene that maps to chromosome 5q12, Cancer Res. 60 (4) (2000) 858–863.

[32] Z. Guo, M. Zhao, G. Jia, R. Ma, M. Li, LncRNA PART1 alleviated myocardial ischemia/reperfusion injury via suppressing miR-503-5p/BIRC5 mediated mitochondrial apoptosis, Int. J. Cardiol. 338 (2021) 176–184.

[33] C. Yagil, N. Hubner, J. Monti, H. Schulz, M. Sapojnikov, F.C. Luft, D. Ganten, Y. Yagil, Identification of hypertension-related genes through an integrated genomic-transcriptomic approach, Circ. Res. 96 (6) (2005) 617–625.

[34] M. Gupta, C. Vavasis, W.H. Frishman, Heat shock proteins in cardiovascular disease a new therapeutic target, Cardiol. Rev. 12 (1) (2004) 26–30.

[35] B. Xu, J.M. English, J.L. Wilsbacher, S. Stippec, E.J. Goldsmith, M.H. Cobb, WNK1, a novel mammalian serine/threonine protein kinase lacking the catalytic lysine in subdomain II, J. Biol. Chem. 275 (22) (2000) 16795–16801.

[36] C. Delaloy, J. Hadchouel, M. Imbert-Teboul, M. Clemessy, A.M. Houot, X. Jeunemaitre, Cardiovascular expression of the mouse WNK1 gene during development and adulthood revealed by a BAC reporter assay, Am. J. Pathol. 169 (1) (2006) 105–118.

[37] S.J. Newhouse, C. Wallace, R. Dobson, C. Mein, J. Pembroke, M. Farrall, D. Clayton, M. Brown, N. Samani, A. Dominiczak, J.M. Connell, J. Webster, G. M. Lathrop, M. Caulfield, P.B. Munroe, Haplotypes of the WNK1 gene associate with blood pressure variation in a severely hypertensive population from the British Genetics of Hypertension study, Hum. Mol. Genet. 14 (13) (2005) 1805–1814.

[38] F.Z. Marques, C.R. Mackay, D.M. Kaye, Beyond gut feelings: how the gut microbiota regulates blood pressure, Nat. Rev. Cardiol. 15 (1) (2018) 20–32.

[39] C. Ricci, A. Wood, D. Muller, M.J. Gunter, A. Agudo, H. Boeing, Y.T. van der Schouw, S. Warnakula, C. Saieva, A. Spijkerman, I. Sluijs, A. Tjonneland, C. Kyro, E. Weiderpass, T. Kuhn, R. Kaaks, M.J. Sanchez, S. Panico, C. Agnoli, D. Palli, R. Tumino, G. Engstrom, O. Melander, F. Bonnet, J.M.A. Boer, T.J. Key, R. C. Travis, K. Overvad, W.M.M. Verschuren, J.R. Quiros, A. Trichopoulou, E.M. Papatesta, E. Peppa, C.M. Iribas, D. Gavrila, A.S. Forslund, J.H. Jansson, G. Matullo, L. Arriola, H. Freisling, C. Lassale, I. Tzoulaki, S.J. Sharp, N.G. Forouhi, C. Langenberg, R. Saracci, M. Sweeting, P. Brennan, A.S. Butterworth, E. Riboli, N.J. Wareham, J. Danesh, P. Ferrari, Alcohol intake in relation to non-fatal and fatal coronary heart disease and stroke: EPIC-CVD case-cohort study, BMJ 361 (2018) k934.

[40] T. Wein, M.P. Lindsay, R. Côté, N. Foley, J. Berlingieri, S. Bhogal, A. Bourgoin, B.H. Buck, J. Cox, D. Davidson, D. Dowlatshahi, J. Douketis, J. Falconer, T. Field, L. Gioia, G. Gubitz, J. Habert, S. Jaspers, C. Lum, D. McNamara Morse, P. Pageau, M. Rafay, A. Rodgerson, B. Semchuk, M. Sharma, A. Shoamanesh, A. Tamayo,

E. Smitko, D.J. Gladstone, Canadian stroke best practice recommendations: Secondary prevention of stroke, sixth edition practice guidelines, update 2017, Int. J. Stroke 13 (4) (2018) 420–443.

[41] S. Kornfeld, Structure and function of the mannose 6-phosphate/insulinlike growth factor II receptors, Annu. Rev. Biochem. 61 (1992) 307–330.

[42] T. Yoshida, T. Shiroshima, S.J. Lee, M. Yasumura, T. Uemura, X. Chen, Y. Iwakura, M. Mishina, Interleukin-1 receptor accessory protein organizes neuronal synaptogenesis as a cell adhesion molecule, J. Neurosci. 32 (8) (2012) 2588–2600.

[43] D.E. Smith, B.P. Lipsky, C. Russell, R.R. Ketchem, J. Kirchner, K. Hensley, Y. Huang, W.J. Friedman, V. Boissonneault, M.M. Plante, S. Rivest, J.E. Sims, A central nervous system-restricted isoform of the interleukin-1 receptor accessory protein modulates neuronal responses to interleukin-1, Immunity 30 (6) (2009) 817–831.

[44] Y. Samson, B. Lapergue, H. Hosseini, [Inflammation and ischaemic stroke: current status and future perspectives], Rev. Neurol. (Paris) 161 (12 Pt 1) (2005) 1177–1182.

[45] R. Jin, L. Liu, S. Zhang, A. Nanda, G. Li, Role of inflammation and its mediators in acute ischemic stroke, J Cardiovasc Transl Res 6 (5) (2013) 834–851.

[46] T. Yoshida, K. Kato, K. Yokoi, M. Oguri, S. Watanabe, N. Metoki, H. Yoshida, K. Satoh, Y. Aoyagi, Y. Nozawa, Y. Yamada, Association of genetic variants with hemorrhagic stroke in Japanese individuals, Int. J. Mol. Med. 25 (4) (2010) 649–656.

[47] L. Bentsen, L. Christensen, A. Christensen, H. Christensen, Outcome and risk factors presented in old patients above 80 years of age versus younger patients after ischemic stroke, J. Stroke Cerebrovasc. Dis. 23 (7) (2014) 1944–1948.

[48] A. Navis, R. Garcia-Santibanez, M. Skliut, Epidemiology and outcomes of ischemic stroke and transient ischemic attack in the adult and geriatric population, J. Stroke Cerebrovasc. Dis. 28 (1) (2019) 84–89.