Genes & Genomics (2022) 44:97–108
https://doi.org/10.1007/s13258-021-01173-1

Online ISSN 2092-9293
Print ISSN 1976-9571

RESEARCH ARTICLE

Dysregulated lncRNA and mRNA may promote the progression of ischemic stroke via immune and inflammatory pathways: results from RNA sequencing and bioinformatics analysis

Yingshuang Wang[1] · Feifei Feng[2] · Pingping Zheng[3] · Lijuan Wang[4] · Yanjun Wang[5] · Yaogai Lv[1] · Li Shen[1] · Kexin Li[1] · Tianyu Feng[1] · Yang Chen[1] · Zhigang Liu[6] · Yan Yao[1]

Received: 17 May 2021 / Accepted: 27 September 2021 / Published online: 26 October 2021

Abstract

Background Long non-coding RNAs (lncRNAs) are widely involved in gene transcription regulation and which act as epigenetic modifiers in many diseases.

Objective To determine whether lncRNAs are involved in ischemic stroke (IS), we analyzed the expression profile of lncRNAs and mRNAs in IS.

Methods RNA sequencing was performed on the blood of three pairs of IS patients and healthy controls. Differential expression analysis was used to identify differentially expressed lncRNAs (DElncRNAs) and mRNAs (DEmRNAs). Based on the co-expression relationships between lncRNA and mRNA, a series of bioinformatics analysis including GO and KEGG enrichment analysis and PPI analysis, were conducted to predict the function of lncRNA.

Results RNA sequencing produced a total of 5 DElncRNAs and 144 DEmRNAs. Influenza A pathway and Herpes simplex infection pathway were the most significant pathways. EP300 and NFKB1 were the most important target proteins, and Human leucocyte antigen (HLA) family were the key genes in IS.

Conclusions Analysis of this study revealed that dysregulated lncRNAs in IS may lead to IS by affecting the immune and inflammation system.

Keywords Ischemic stroke · Long non-coding RNA · RNA-sequencing · Bioinformatics · Expression profiles

Introduction

Ischemic stroke (IS) is one of the cardiovascular diseases and the leading cause of death and disability worldwide, especially in developing countries (Strong et al. 2007). Cardiovascular and cerebrovascular diseases, diabetes, dyslipidemia, obesity, smoking, drinking, and anticoagulants may be the main factors affecting the occurrence of IS (Lavados et al. 2005; O'Donnell et al. 2010). Factors may lead to IS through vascular and neuronal damage, dysfunction of molecular signaling pathways, inflammatory cytokine damage, and oxidative stress damage (Chaitanya et al. 2013; Jin et al. 2013; Li et al. 2018). Focusing on the mechanisms

Yingshuang Wang and Feifei Feng contributed equally to this article.

✉ Zhigang Liu
liuzhigang@jlu.edu.cn

✉ Yan Yao
yaoyan@jlu.edu.cn

[1] Key Laboratory of Organ Regeneration and Transplantation of Ministry of Education, Department of Epidemiology and Biostatistics, School of Public Health, Jilin University, Changchun 130021, China

[2] Jilin Province of Jilin Combine Traditional Chinese and Western Medicine Hospital, Jilin 132000, China

[3] Futian District Center for Disease Prevention and Control, Shenzhen 518040, China

[4] Department of Neurology, The Neuroscience Center, The First Hospital of Jilin University, Jilin University, Changchun 130021, China

[5] Nursing Department, The Second Hospital of Jilin University, Changchun 130021, Jilin, China

[6] Department of Pain Management, The Second Hospital of Jilin University, No.218, Ziqiang Street, Nanguan District, Changchun 130041, Jilin, China

of IS injury, repair and inflammation, and its possible molecular mechanisms may provide a more theoretical basis for IS screening, prevention and treatment.

Biomarkers associated with stroke are mainly a variety of proteins associated with pathological processes such as inflammation-related damage, neuronal apoptosis, and vascular endothelial dysfunction (Serena et al. 2005; An et al. 2013). However, due to the complex protein composition, there are various post-translational modifications, hydrolysis and denaturation, which makes it difficult to select the appropriate method for accurate detection (Calligaris et al. 2011; Lam et al. 2016). The physical and chemical properties of traditional biomarkers are unstable and poorly conserved, and their content is related to the regulation of gene expression. Therefore, rational application of gene expression variation can be used for diagnostic prediction in the early stages of the disease (Ebert et al. 2006; Sepramaniam et al. 2014).

lncRNA is a group of non-protein coding RNA molecules, and their transcripts are more than 200 nucleotides in length (Yang et al. 2014). LncRNA regulates DNA methylation, histone modification or chromosome remodeling on multiple layers through diverse molecular regulation mechanisms, and participates in many important biological regulation processes such as genomic imprinting, transcriptional interference and nuclear transport (Wilusz et al. 2009; Chen and Carmichael 2010; Marchese et al. 2016, 2017). Current lncRNA research covers many fields, such as tumor, blood system diseases, cardiovascular and cerebrovascular diseases, especially in ischemic and hypoxic diseases (Harries 2012; Li and Chen 2013; Sánchez and Huarte 2013). LncRNA UCA1 can be used as a potential diagnostic marker and therapeutic target for acute myocardial infarction (Yan et al. 2016), lncRNA ANRIL is associated with the sensitivity of atherosclerotic disease and can be used as a marker for the diagnosis (Holdt and Teupser 2018). Changes in the expression of certain lncRNA after ischemia-reperfusion injury may be biomarkers of ischemia-reperfusion injury during liver surgery or transplantation (Chen et al. 2013). In addition, lncRNA is also considered effective in the treatment of ischemic diseases by promoting stem cell differentiation and preventing erythrocyte apoptosis (Yang and Lu 2016; Zhang et al. 2016).

Although the regulatory mechanism of lncRNA and its relationship with some ischemic and hypoxic diseases have been preliminarily understood, the regulatory network of gene expression mediated by lncRNA and its molecular mechanism in IS remain to be further explored. To determine the functional significance of lncRNAs in the pathophysiological regulation of IS, RNA sequencing technology, which is superior to microarrays, was used to analyze lncRNA and mRNA expression profiles in this study.

Materials and methods

Participants and study design

Three patients first diagnosed IS and three healthy controls, Han Chinese males, 40–60 years old (Supplementary Table 1), were recruited from the hospital of Jilin University from July to December 2017. IS patients were diagnosed for the first time based on the "*Chinese guidelines for diagnosis and treatment of acute ischemic stroke 2014*", and all of them had no history of using antiplatelet or antidiabetic agents. All objectives with history of diabetes mellitus, atrial fibrillation, myocardial infarction, tumor, acute infectious disease, immunity disease, blood disease, renal or liver failure, and hemorrhagic stroke or recurrent stroke were excluded. Written informed consent was obtained. The study was approved by the Ethics Committee of School of Public Health, Jilin University. The design flow chart of this study was shown in Fig. 1.

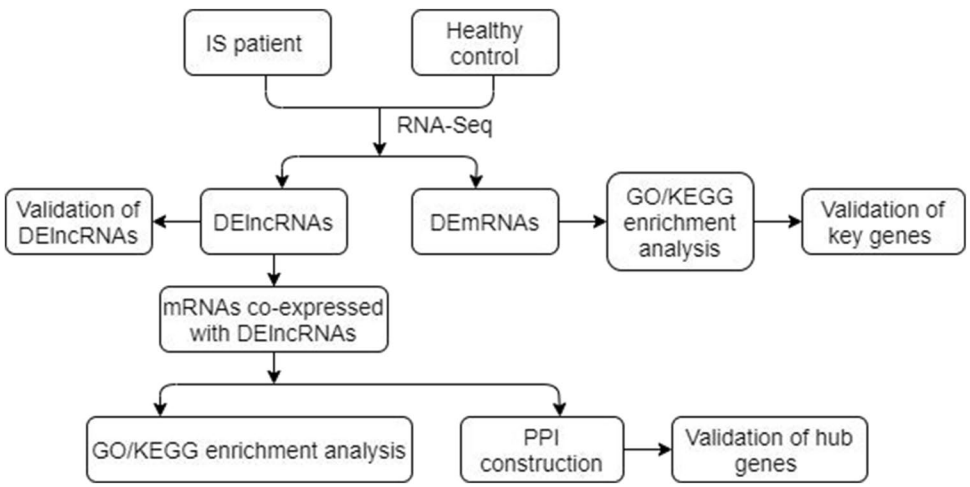

Fig. 1 Flow diagram of the study design

Sample preparation and RNA sequencing

Peripheral blood samples were collected in the next morning after the participants had fasted for ten hours or overnight. Total RNA was isolated and purified using standard TRIzol Reagent (Invitrogen, Carlsbad, CA) according to the manual. RNA concentration and purity, RNA integrity were then assessed using the Nanodrop2000 spectrophotometer (Thermo Fisher Scientific, Wilmington, DE) and an Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA). The cDNA library was constructed using the Truseq RNA sample Prep Kit (Illumina, Inc., San Diego, CA). RNA sequencing ($2^{*}150$ bp) was conducted on the Illumina HiSeq 4000 sequencing system (Illumina, Inc.).

Quality control and quantification

Quality-control filtering was performed by removing the adapter sequences and low-quality bases, and high-quality clean reads were generated. The clean reads were mapped to the reference using HISAT2 v2.0.4 (Pertea et al. 2016) and the reference-based assembly of transcripts was performed using Stringtie v1.3.1 (Pertea et al. 2015) (reference genome sequence and annotation files were downloaded from GENCODE (GRCh38, https://www.gencodegenes.org/)). The lncRNAs were screened out according to the number of exons, length, known annotation and coding potential of transcripts. Then transcripts including mRNAs, lncRNAs and transcripts of unknown coding potential (TUCPs) were quantified by StringTie-eB.

Differential expression analysis

To identify differentially expressed lncRNAs and mRNAs between IS patients and controls, differential expression analyses were performed by "ballgown" package of R software (Pertea et al. 2016). From the perspective of statistical significance, different types of transcripts (lncRNA, TUCPs and mRNA) were analyzed as a whole, so that the results had no preference for molecular types. Transcripts with a P-value < 0.05 and $|\log2\text{foldchange}| > 1$ were assigned as differentially expressed.

GO and KEGG enrichment analysis

Gene Ontology (GO) enrichment analysis of differentially expressed genes or lncRNAs target genes were implemented by the "GOseq" R package, in which gene length bias was corrected (Young et al. 2010). Corrected *P*-value < 0.05 were considered to be significantly enriched. Kyoto Encyclopedia of Genes and Genomes (KEGG) is a database resource for understanding high-level functions and utilities of the biological system (Kanehisa et al. 2008). We used KOBAS software to test the statistical enrichment of differential expression genes or lncRNA target genes in KEGG pathways (Mao et al. 2005). *P*-value was used to determine whether the pathways were significant.

PPI analysis

Protein-Protein Interaction (PPI) analysis of differentially expressed genes was based on the online database resource which known as STRING (https://string-db.org/cgi/input.pl) and predicted interactions between proteins. We used mRNAs co-expressed with DElncRNAs to construct the PPI network and visualized it in Cytoscape 3.6.1 (Shannon et al. 2003).

Network construction of GO/KEGG pathways of DEmRNAs

GO/KEGG Enrichment Pathways Network was constructed using all differentially expressed mRNAs with a plug-in called "ClueGO" in Cytoscape 3.6.1 (Bindea et al. 2009). ClueGO integrates the terms of Gene Ontology (GO) and the KEGG/BioCarta pathways, it could achieve comprehensive visualization by creating a GO/KEGG pathways network.

Cross validation

Gene expression profiling of GSE22255 dataset was performed in peripheral blood mononuclear cells of 20 IS patients and 20 sex- and age-matched controls using GPL570 [HG-U133_Plus_2] Affymetrix Human Genome U133 Plus 2.0 Array. In order to be consistent with the design of this study, we only downloaded the expression matrix of male cases and male healthy controls (10 vs. 10), and used R software "limma" package (Ritchie et al. 2015) for standardization and differential expression analysis, and obtained the log2(foldchange) value and P value to validate significance of RNAs that we have found.

Results

Quality control and mapping

After filtering out the reads with adapter, the reads with N base more than 10 % and the low quality reads from raw reads, 97,077,054, 86,224,578 and 104,855,436 clean reads were produced in IS patients, and 97,740,922, 97,583,408 and 100,893,034 clean reads were produced in the control group (Supplementary Table 2). The percentage of total mapped reads was 94.96%, 95.37% and 96.44% in IS patients group, and 95.34%, 97.23% and 96.32% in control group, respectively. The total mapped rate was more than

90%, indicating that the selection of reference genome was appropriate and there was no pollution in the experiment (Supplementary Table 3). A total of 86,983 transcripts were detected, including 80,863 mRNAs, 2469 novel lncRNAs, 2369 annotated lncRNAs and 1282 TUCPs. A total of 24,566 genes were detected, including 19,627 mRNAs, 2137 novel lncRNAs, 1854 annotated lncRNAs and 948 TUCPs.

Differentially expressed lncRNAs and mRNAs in IS

A total of five lncRNAs were differentially expressed in IS, including 4 up-regulated lncRNAs and 1 down-regulated lncRNA (Table 1). There were 144 differentially expressed mRNAs, among which 70 were up-regulated and 74 were down-regulated, each top 10 were listed in Table 2, respectively.

LncRNA/mRNA co-expression analysis and GO/KEGG annotation to predict the probable functions

LncRNA can regulate the expression of target coding genes to achieve function (Bazin et al. 2017). We analyzed the function of mRNA co-expressed with lncRNA to clarify the biological role of DElncRNA. The lncRNA-mRNA co-expression network (Supplementary Fig. 1) was established to show co-expression relationships of them, the network comprised 946 nodes and 1009 connections. The most crucial sub-network was constructed by the transcripts with

Table 1 Differentially expressed lncRNAs in ischemic stroke

Transcript ID	Gene symbol	Status	Regulation	Average FPKM		log2 (foldchange)	P-value	Number of co-expression mRNAs
				IS group	Control group			
LNC_000015	–	Novel	Up	176.714	19.325	3.193	0.0223	118
ENST00000602573.1	SNHG8	Annotated	Up	2.808	0.938	1.582	0.01583	56
ENST00000381051.6	MIRLET7BHG	Annotated	Up	3.013	1.150	1.390	0.00674	97
LNC_001727	–	Novel	Up	11.639	5.802	1.004	0.00158	95
ENST00000574212.1	AF001548.5	Annotated	Down	1.707	4.790	− 1.488	0.00649	643

Table 2 Top 10 up- and downregulated mRNAs in ischemic stroke

Transcript ID	Gene name	Regulation	Average FPKM		log2 (foldchange)	P-value
			IS group	Control group		
ENST00000468133	MAPK14	Up	2.255	0	Inf	5.78E−05
ENST00000254066	RARA	Up	2.644	0	Inf	0.00181
ENST00000443308	SHISA5	Up	3.858	0	Inf	0.00754
ENST00000475880	WDR45	Up	9.328	0	Inf	0.01142
ENST00000367026	TRAF3IP3	Up	2.743	0.001	12.329	0.00429
ENST00000616058	SUCO	Up	2.894	0.008	8.484	0.02382
ENST00000620123	TAP2	Up	3.560	0.010	8.454	0.03042
ENST00000596676	SMCO4	Up	2.029	0.079	4.689	0.00839
ENST00000337907	RERE	Up	2.720	0.110	4.626	0.017
ENST00000381733	ASAH1	Up	4.050	0.210	4.270	0.00272
ENST00000434651	HLA-DQB1	Down	0	12.974	− Inf	0.00101
ENST00000400908	RERE	Down	0	4.746	− Inf	0.00089
ENST00000557140	NFKBIA	Down	0	4.301	− Inf	0.00355
ENST00000343139	HLA-DQA1	Down	0	4.082	− Inf	0.01339
ENST00000394690	STX5	Down	0	4.060	− Inf	0.04904
ENST00000351660	DHPS	Down	0	2.554	− Inf	0.0465
ENST00000396009	NFATC2	Down	0	2.319	− Inf	0.002
ENST00000531372	IGLL5	Down	0	2.123	− Inf	0.00743
ENST00000507711	FRYL	Down	0	2.113	− Inf	0.00036
ENST00000542643	RAP1GAP	Down	0	1.810	− Inf	0.03751

Inf maximum, represents the highest degree of up regulation, − *Inf* minimum, represents the highest degree of down regulation

a high degree, which would be the core regulatory modules of the entire co-expression network. Thus, lncRNA AF001548.5, which has the most co-expressed mRNAs was the most important lncRNA in the network. GO enrichment analysis showed the top 20 terms of three types, biological process, cellular component, and molecular function (Supplementary Fig. 2 and Supplementary Table 4). KEGG enrichment analysis showed that the DElncRNAs play roles mainly in Influenza A pathway and Herpes simplex infection pathway (Fig. 2 and Supplementary Table 5).

PPI analysis of mRNAs co-expressed with DElncRNAs

The PPI network (Fig. 3) contains 417 nodes and 810 edges, protein interactions were evaluated with the highest confidence (0.9). Degree was used to assess the importance of genes, and the genes with top 30° were listed, EP300 and NFKB1 were regarded as the most important top two target genes that regulated protein expression (Fig. 4).

Network construction of GO/KEGG pathways of DEmRNAs

We utilized all DEmRNAs for the GO and KEGG analysis and constructed a GO/KEGG pathway network to illustrate the critical mRNAs in the process of IS. The network built 13 groups composed of GO terms and KEGG pathways, and these functional groups with target genes of mRNAs were shown in Table 3. HLA-DQB1, HLA-DQA1 and HLA-DRB5 were the key genes which had overlapped in at least four pathway groups, and all of them were belonged to the HLA family (Fig. 5). In addition, these three genes were mainly involved in the inflammatory bowel disease (IBD) pathway and asthma pathway, both in Th17 cell differentiation group.

Cross validation

Five valuable coding genes and three DElncRNAs in this study were verified. Because the early microarray sequencing was not enough to detect all genes in this study, we can only verify the expression of seven IS related genes. EP300, NFKB1, SNHG8 and MIRLET7BHG met the differential expression criteria of $P < 0.1$ and HLA-DQB1, HLA-DQA1, HLA-DRB5 met the criteria |log2foldchange| > 1 (Table 4).

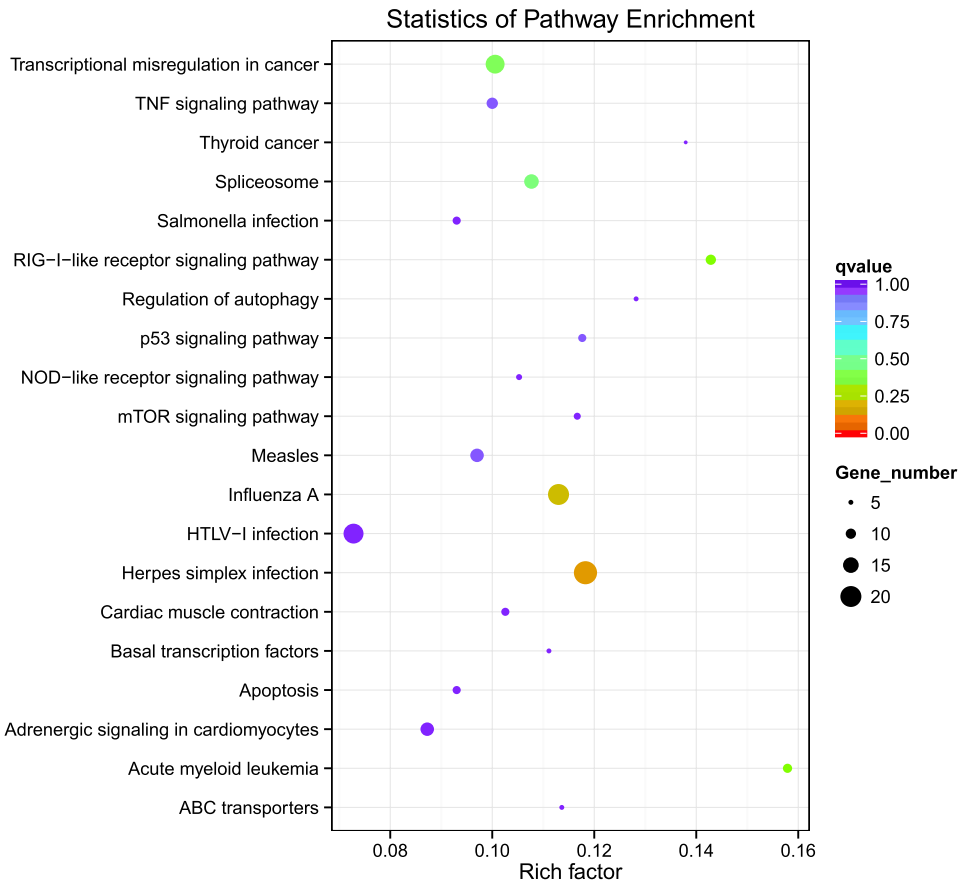

Fig. 2 Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analyses of mRNAs co-expressed with DElncRNAs in ischemic stroke. The size of the circle represents the number of genes enriched, and the color represents the significance

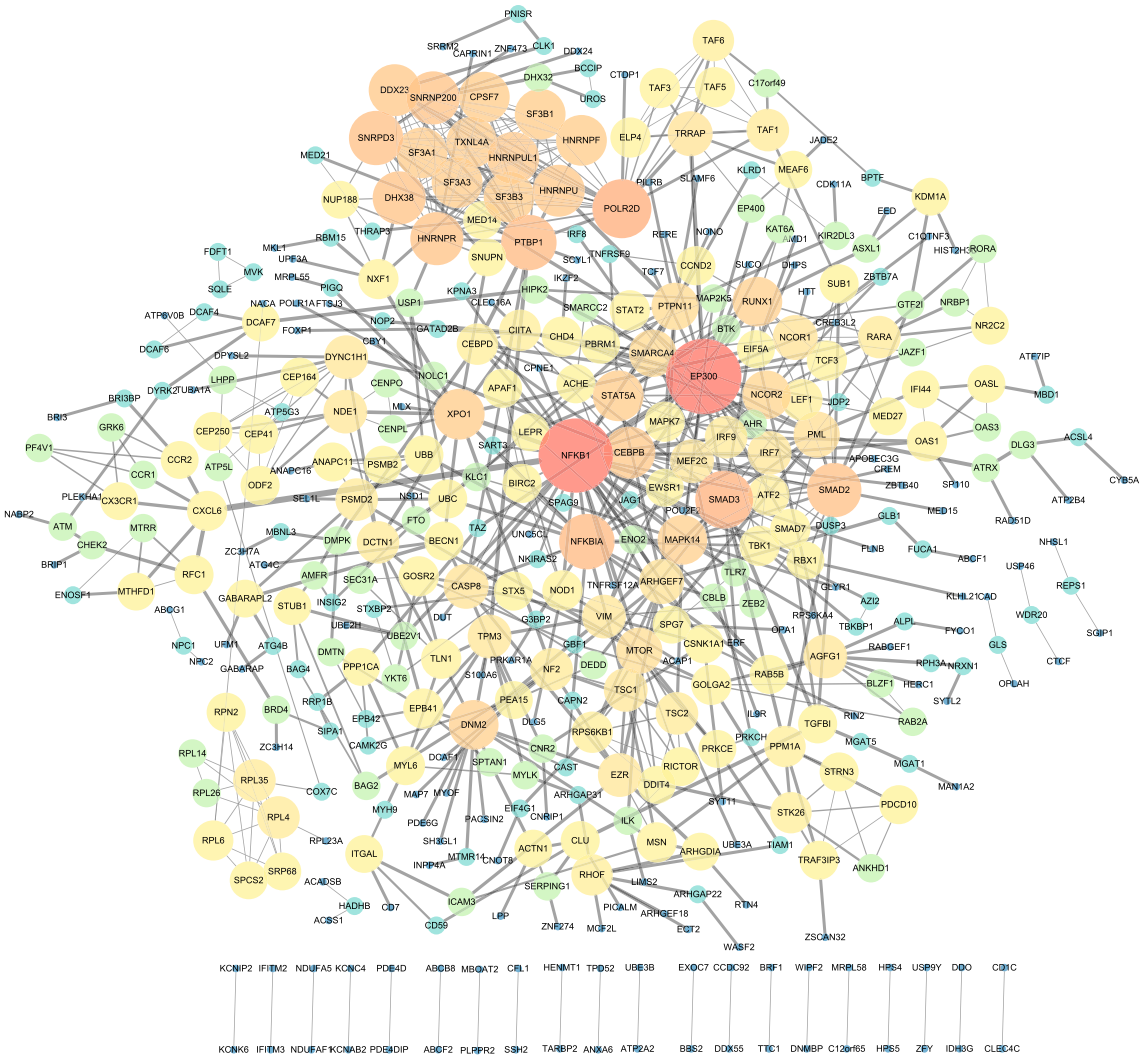

Fig. 3 Protein–protein interaction networks constructed by Search Tool for the Retrieval of Interacting Genes/Proteins (STRING) based on the mRNAs co-expressed with DElncRNAs. The larger the circle and the redder the color, representing the greater degree of the gene, and the more genes that are connected to it. Conversely, the smaller the circle and the bluer the color, representing the smaller degree of the gene, and the fewer number of genes connected to it

Discussion

In this study, 5 DElncRNAs were identified that differentially expressed in IS, among which LNC_000015, LNC_001727 were novel lncRNAs, SNHG8, MIRLET7BHG, AF001548.5 are annotated lncRNA in the database. After a series of bioinformatics analysis, we mainly found that dysregulated lncRNAs and mRNAs in IS vs. controls may lead to IS by affecting the immune system of the body.

There are few published RNA-Seq studies of IS patients, and some of these studies have not strictly controlled variables, so a large number of differentially expressed lncRNAs have been screened. Some of the DE-genes may not be related to IS, but related to other IS parallel diseases due to selection bias. In our study, only five differentially expressed lncRNAs were screened by setting strict inclusion and exclusion criteria of patients and control group. Among the five differentially expressed lncRNAs, the lncRNA SNHG8 has been widely reported. Up regulation of lncRNA SNHG8 is a risk factor for many diseases. It plays an important regulatory role in the occurrence and development of acute myocardial infarction (Zhuo et al. 2019), a variety of cancers (such as liver cancer, pancreatic cancer, nasopharyngeal carcinoma, endometrial cancer, esophageal squamous cell

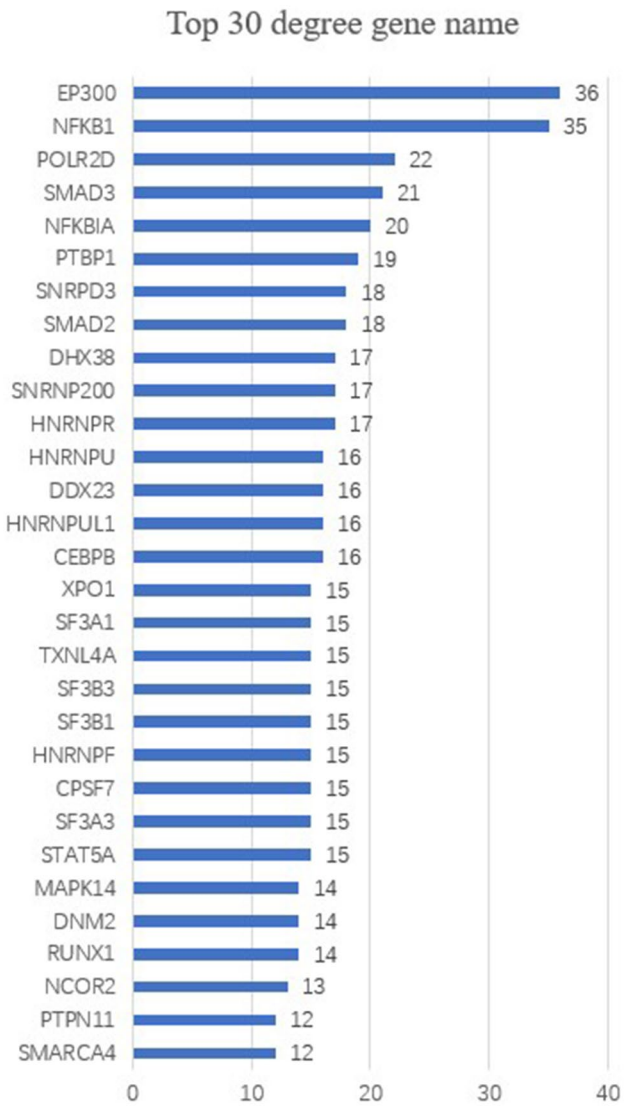

Fig. 4 Top 30 degree genes involved in the PPI network

carcinoma, ovarian cancer, gastric cancer, breast cancer, cervical cancer, non-small cell lung cancer, prostate cancer, etc.) (Yuan et al. 2021) and ischemic diseases (Liu et al. 2019; Tian et al. 2020). Liu et al. (2019) verified that Snhg8/miR-384/Hoxa13/FAM3A axis regulating neuronal apoptosis in ischemic mice model, and then Tian et al. (2020) proved that LncRNA Snhg8 attenuates microglial inflammation response and blood-brain barrier damage in IS through regulating miR-425-5p mediated SIRT1/NF-κB signaling. Similarly, we found the hub genes EP300 and NFKB1 in the constructed PPI network of DElncRNAs co-expression mRNA. The conclusion of Tian et al. (2020) is consistent with our study. Although the other four DElncRNAs have not been reported to be associated with IS. They can be used as potential biomarkers and therapeutic targets of IS, which providing direction for future research.

Stroke is still a severe challenge for either developed countries or developing countries, poses a substantial socio-economic burden (Addo et al. 2012). It's urgent to find its pathogenesis and control it effectively. KEGG analysis found that Influenza A and Herpes simplex infection pathway were the most significant two pathways in this study, both related to immune and inflammation. Inflammation is considered involved in all forms of brain damage, and immune mechanisms play an important role in the risk and progression of stroke and in cerebral ischemia (Smith et al. 2013; Fu et al. 2015). Recent data from clinical and experimental research clearly show that systemic inflammatory diseases such as atherosclerosis, obesity, diabetes or infection are associated with dysregulated immune responses, which can profoundly contribute to cerebrovascular inflammation and injury in the central nervous system (Ling et al. 2015). Some findings have shown that several immune-mediated diseases (IMDs) are linked to cerebrovascular diseases, and many hospitalizations of IMDs have been proved to be related to the increased risk of ischemic or hemorrhagic stroke (Zöller et al. 2012; Cho et al. 2014).

Evidence suggests that acute bacterial and viral infections are prime factors for an increased risk of IS (Urbanek et al. 2010), and the mortality from vascular disease and hospitalization for stroke increased during and after the influenza pandemic. Influenza vaccination can reduce hospitalization and mortality in the elderly and prevent incapacity in working-aged adults (Madjid et al. 2009). Herpes Simplex Virus (HSV) Type 2, a class of influenza virus, was regarded as a cause of IS, and researchers found that untreated HSV-2 meningitis could lead to vascular inflammation and IS ultimately (Snider et al. 2014; Zis et al. 2016). The pathophysiological mechanism of stroke caused by varicella zoster virus (VZV) infection is believed to be similar to the pathophysiological mechanism of IS caused by HSV central nervous system infection, the prevention of VZV infection is considered as a treatable factor for transient IS (Nagel and Gilden 2014).

In this study, Ep300 and NFKB1 were the hub target genes in PPI network, which regulate protein expression in IS. EP300 was known as Histone acetyltransferase p300 or p300 which is an enzyme encoded by EP300 gene (Eckner et al. 1994). It regulates transcription of genes via chromatin remodeling, which has been found to play an essential role in the biological function of regulatory T cells and is expected to be used in cancer immunotherapy in the future (Liu et al. 2013; Ghosh et al. 2016). Unfortunately, no studies have

Table 3 ClueGO results: functional groups with genes

Function	Groups	Group genes
Apoptosis	Group01	ACTG1,AKT1,BIRC2,CAPN2,GZMB,NFKBIA,SPTAN1,TUBA1A
Influenza A	Group09	ACTG1,AKT1,ARHGDIA,CAMK2G,CCND2,**HLA-DQA1**, **HLA-DQB1,HLA-DRB5**,IL2RG,MAPK14,MX1,NFKBIA,OAS1, OSBPL8,STAT5A
Pertussis	Group02	C4B,CFL1,MAPK14,SERPING1
S100 protein binding	Group00	AHNAK,EZR,S100A6
Th17 cell differentiation	Group13	ACTG1,AKT1,ARHGDIA,BIRC2,C4B,CAMK2G,CCND2,CD2, CD37,GLG1,GZMB,HIST1H2BD,HIST1H2BE,**HLA-DQA1**, **HLA-DQB1,HLA-DRB5**,IL2RG,MAPK14,MX1,NFATC2, NFKBIA,OAS1,OASL,RARA,STAT5A,STX5,TAP2,TBKBP1, TMBIM6,VCAN
Toxoplasmosis	Group10	AKT1,ARHGDIA,BIRC2,CAMK2G,**HLA-DQA1,HLA-DQB1, HLA-DRB5**,MAPK14,NFATC2, NFKBIA,OSBPL8,TBKBP1
Immunological synapse	Group07	ARHGDIA,CD37,ELF1,EZR,GZMB,PTPRJ
Negative regulation of protein kinase B signaling	Group04	AKT1,CD2,PTPRJ
Positive regulation of myoblast fusion	Group12	CAPN2,MAPK14,NFATC2,RIPOR2
Protein deacylation	Group08	CHD4,HDAC6,JDP2,NIPBL,SIRT5,TAP2
Regulation of GTP binding	Group05	C9orf72,CSNK1A1,RAP1GAP,RIPOR2
Regulation of nuclease activity	Group11	AKT1,CAMK2G,**HLA-DQA1,HLA-DQB1,HLA-DRB5**,IFI6, IFITM2,MX1,OAS1,OASL,TMBIM6
Tertiary granule lumen	Group06	ARMC8,ASAH1,SPTAN1
Transferase activity, transferring pentosyl groups	Group03	PARP8,SIRT5,TYMP

Bold genes were the key genes which had overlapped in at least four pathway groups, and all of them were belonged to the HLA family

found the mechanism of action of EP300 related to IS up to now. However, given the important role of EP300 in PPI network, we believe that EP300 should be paid more attention in future studies when studying the molecular mechanism of IS. As for NFKB1, which is a factor that inhibits inflammation, aging and cancer, is thought to be associated with cerebral ischemia-reperfusion injury, and a study of Korean adults found that genetic polymorphisms of NFKB1 are associated with stroke susceptibility (Cartwright et al. 2016; Kim et al. 2018; Zhu et al. 2018). NFKB1 has been confirmed as a potential biomarker for the diagnosis and treatment of IS in the study of Liang (2015), indicating that PPI network in this study has certain accuracy to predict the hub genes.

From the network of GO/KEGG pathways, we found that HLA-DQB1, HLA-DQA1A and HLA-DRB5 were the key genes which had overlapped in at least four pathway groups, in addition, both of them were belonged to the human leucocyte antigen (HLA) family. A previous study on IS patients in China has shown that the HLA-DRB1*04, HLA-DRB1*03, and the HLA-DRB1*12 alleles have protective effects on stroke (Liu et al. 2011). Similarly, a study on South Indian patients presented the association of human leucocyte antigen HLA-DRB1/DQB1 alleles and haplotypes with IS (Murali et al. 2016). Interestingly, we found that the HLA alleles (HLA-DQB1, HLA-DQA1, HLA-DRB5) at the same time in the channel were connected to the inflammatory bowel disease (IBD) pathway. It is worthy to note that IBD was considered to increase the risk of IS in a retrospective cohort study of a Taiwanese population, because of the systemic inflammatory burden that IBD led to may be a key determinant of atherosclerotic thrombosis (Huang et al. 2014). This finding was supported by previous studies that have shown a disease severity-dependent could increase the risk of developing stroke and myocardial infarction among patients with other chronic inflammatory diseases such as rheumatoid arthritis and psoriasis (Solomon et al. 2010; Ahlehoff et al. 2012).

As this study is a preliminary exploratory study, further experiments are expected to be conducted to validate the sequencing results, and we included a total of 6 subjects (3 patients vs. 3 controls), the sample size of RNA sequencing

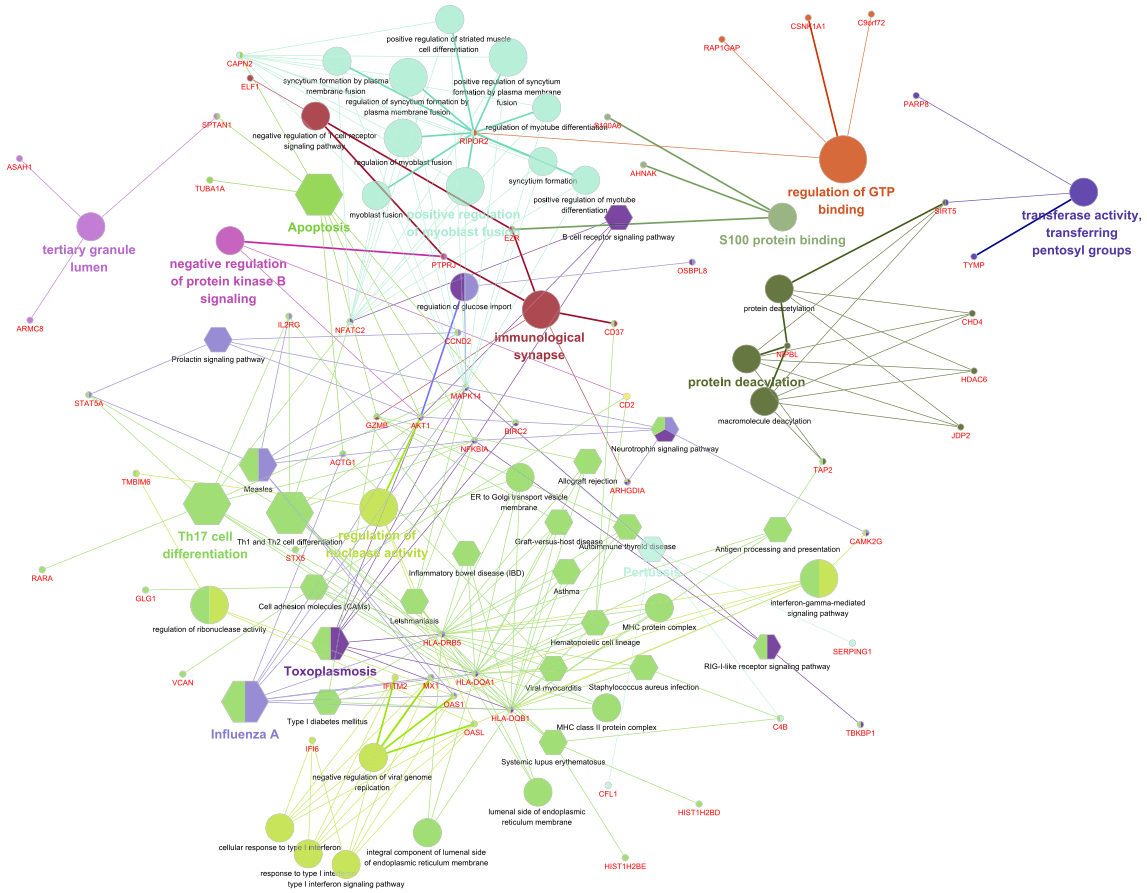

Fig. 5 GO/KEGG pathways network constructed by ClueGO app in Cytoscape. Hexagons represent significantly enriched KEGG pathways; big ellipses indicate significantly enriched GO pathways; small ellipses indicate mRNAs that act as the link hinge between the pathways. Hexagons and big ellipses of the same color represent a functional group, while a hexagon or a big ellipse contains different colors, meaning that it belongs to different functional groups at the same time. When a small ellipse consists of different colors, it is included in a different path (colour figure online)

Table 4 Expression of valuable genes validated in the GSE22255

Gene name	Type	GSE22255		
		t	P	Log2FC
EP300	Important target genes of DElncRNAs	− 2.1664	0.0417	− 0.30725
NFKB1	Important target genes of DElncRNAs	− 2.93576	0.00779	− 0.35679
HLA-DQB1	Key genes	− 1.69671	0.104	− 1.08079
HLA-DQA1	Key genes	− 1.58918	0.127	− 1.63289
HLA-DRB5	Key genes	− 1.69671	0.104	− 1.08079
SNHG8	DElncRNAs	− 1.72965	0.0981	− 0.34551
MIRLET7BHG	DElncRNAs	1.915236	0.0689	0.185295
AF001548.5	DElncRNAs	–	–	–

may be small, so expanding of the sample size is expected to verify the results of this study.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s13258-021-01173-1.

Acknowledgements We would like to express our gratitude to the Novogene Co., LTD for providing the high-throughput sequencing technology helps.

Author contributions Conceptualization: YY, ZL; methodology: Yingshuang W, PZ; formal analysis and investigation: TF, KL, YC; validation: YL, LS; writing—original draft preparation: Yingshuang W, PZ; writing—review and editing: FF; funding acquisition: YY; resources: LW, Yanjun W; supervision: ZL.

Funding This work was supported by the Bethune program of Jilin University (No. 470110000715). The funders had no role in the design of the study or in the collection, analysis, and interpretation of data or in writing the manuscript.

Data availability All raw and processed sequencing data generated in this study have been submitted to the NCBI Gene Expression Omnibus (GEO; https://www.ncbi.nlm.nih.gov/geo/) under accession number GSE180470.

Declarations

Conflict of interest Author Yingshuang Wang, author Feifei Feng, author Pingping Zheng, author Lijuan Wang, author Yanjun Wang, author Yaogai Lv, author Li Shen, author Kexin Li, author Tianyu Feng, author Yang Chen, author Zhigang Liu and author Yan Yao declare that they have no conflict of interest.

Ethics approval This study was approved by the Ethics Committee of Jilin University, School of Public Health.

Consent to participate Written informed consent was obtained from all subjects and/or their guarantors.

References

Addo J, Ayerbe L, Mohan KM, Crichton S, Sheldenkar A, Chen R, Wolfe CD, McKevitt C (2012) Socioeconomic status and stroke: an updated review. Stroke 43:1186–1191. https://doi.org/10.1161/STROKEAHA.111.639732

Ahlehoff O, Gislason GH, Jørgensen CH, Lindhardsen J, Charlot M, Olesen JB, Abildstrøm SZ, Skov L, Torp-Pedersen C, Hansen PR (2012) Psoriasis and risk of atrial fibrillation and ischaemic stroke: a Danish Nationwide Cohort Study. Eur Heart J 33:2054–2064. https://doi.org/10.1093/eurheartj/ehr285

An SA, Kim J, Kim OJ, Kim JK, Kim NK, Song J, Oh SH (2013) Limited clinical value of multiple blood markers in the diagnosis of ischemic stroke. Clin Biochem 46:710–715. https://doi.org/10.1016/j.clinbiochem.2013.02.005

Bazin J, Baerenfaller K, Gosai SJ, Gregory BD, Crespi M, Bailey-Serres J (2017) Global analysis of ribosome-associated noncoding RNAs unveils new modes of translational regulation. Proc Natl Acad Sci USA 114:E10018–E10027. https://doi.org/10.1073/pnas.1708433114

Bindea G, Mlecnik B, Hackl H, Charoentong P, Tosolini M, Kirilovsky A, Fridman WH, Pagès F, Trajanoski Z, Galon J (2009) ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. Bioinformatics 25:1091–1093. doi:https://doi.org/10.1093/bioinformatics/btp101

Calligaris D, Villard C, Lafitte D (2011) Advances in top-down proteomics for disease biomarker discovery. J Proteom 74:920–934. https://doi.org/10.1016/j.jprot.2011.03.030

Cartwright T, Perkins ND, L WC (2016) NFKB1: a suppressor of inflammation, ageing and cancer. Febs J 283:1812–1822. doi:https://doi.org/10.1111/febs.13627

Chaitanya GV, Cromer WE, Parker CP, Couraud PO, Romero IA, Weksler B, Mathis JM, Minagar A, Alexander JS (2013) A recombinant inhibitory isoform of vascular endothelial growth factor164/165 aggravates ischemic brain damage in a mouse model of focal cerebral ischemia. Am J Pathol 183:1010–1024. doi:https://doi.org/10.1016/j.ajpath.2013.06.009

Chen LL, Carmichael GG (2010) Decoding the function of nuclear long non-coding RNAs. Curr Opin Cell Biol 22:357–364. doi:https://doi.org/10.1016/j.ceb.2010.03.003

Chen Z, Jia S, Li D, Cai J, Tu J, Geng B, Guan Y, Cui Q, Yang J (2013) Silencing of long noncoding RNA AK139328 attenuates ischemia/reperfusion injury in mouse livers. Plos One 8:e80817. doi:https://doi.org/10.1371/journal.pone.0080817

Cho HJ, Kim SS, Sung SM, Jung DS (2014) Impact of thyroid autoantibodies on functional outcome in patients with acute ischemic stroke. J Stroke Cerebrovasc Dis 23:1915–1920. doi:https://doi.org/10.1016/j.jstrokecerebrovasdis.2014.02.008

Ebert MP, Korc M, Malfertheiner P, Röcken C (2006) Advances, challenges, and limitations in serum-proteome-based cancer diagnosis. J Proteome Res 5:19–25. doi:https://doi.org/10.1021/pr050271e

Eckner R, Ewen ME, Newsome D, Gerdes M, DeCaprio JA, Lawrence JB, Livingston DM (1994) Molecular cloning and functional analysis of the adenovirus E1A-associated 300-kD protein (p300) reveals a protein with properties of a transcriptional adaptor. Genes Dev 8:869–884. doi:https://doi.org/10.1101/gad.8.8.869

Fu Y, Liu Q, Anrather J, Shi FD (2015) Immune interventions in stroke. Nat Rev Neurol 11:524–535. doi:https://doi.org/10.1038/nrneurol.2015.144

Ghosh S, Taylor A, Chin M, Huang HR, Conery AR, Mertz JA, Salmeron A, Dakle PJ, Mele D, Cote A, Jayaram H, Setser JW, Poy F, Hatzivassiliou G, DeAlmeida-Nagata D, Sandy P, Hatton C, Romero FA, Chiang E, Reimer T, Crawford T, Pardo E, Watson VG, Tsui V, Cochran AG, Zawadzke L, Harmange JC, Audia JE, Bryant BM, Cummings RT, Magnuson SR, Grogan JL, Bellon SF, Albrecht BK, Sims RR, Lora JM (2016) Regulatory T Cell Modulation by CBP/EP300 Bromodomain Inhibition. J Biol Chem 291:13014–13027. doi:https://doi.org/10.1074/jbc.M115.708560

Harries LW (2012) Long non-coding RNAs and human disease. Biochem Soc Trans 40:902–906. doi:https://doi.org/10.1042/BST20120020

Holdt LM, Teupser D (2018) Long noncoding RNA ANRIL: Lncing genetic variation at the chromosome 9p21 locus to molecular mechanisms of atherosclerosis. Front Cardiovasc Med 5:145. https://doi.org/10.3389/fcvm.2018.00145

Huang WS, Tseng CH, Chen PC, Tsai CH, Lin CL, Sung FC, Kao CH (2014) Inflammatory bowel diseases increase future ischemic stroke risk: a Taiwanese population-based retrospective cohort study. Eur J Intern Med 25:561–565. https://doi.org/10.1016/j.ejim.2014.05.009

Jin R, Liu L, Zhang S, Nanda A, Li G (2013) Role of inflammation and its mediators in acute ischemic stroke. J Cardiovasc Transl Res 6:834–851. doi:https://doi.org/10.1007/s12265-013-9508-6

Kanehisa M, Araki M, Goto S, Hattori M, Hirakawa M, Itoh M, Katayama T, Kawashima S, Okuda S, Tokimatsu T, Yamanishi Y (2008) KEGG for linking genomes to life and the environment. Nucleic Acids Res 36:D480–D484. doi:https://doi.org/10.1093/nar/gkm882

Kim SK, Jang HM, Kim DY (2018) The promoter polymorphism of NFKB1 gene contributes to susceptibility of ischemic stroke in Korean population. J Exerc Rehabil 14:1096–1100. doi:https://doi.org/10.12965/jer.1836592.296

Lam MP, Ping P, Murphy E (2016) Proteomics research in cardiovascular medicine and biomarker discovery. J Am Coll Cardiol 68:2819–2830. https://doi.org/10.1016/j.jacc.2016.10.031

Lavados PM, Sacks C, Prina L, Escobar A, Tossi C, Araya F, Feuerhake W, Galvez M, Salinas R, Alvarez G (2005) Incidence, 30-day case-fatality rate, and prognosis of stroke in Iquique, Chile: a 2-year community-based prospective study (PISCIS project). Lancet 365:2206–2215. doi:https://doi.org/10.1016/S0140-6736(05)66779-7

Li CH, Chen Y (2013) Targeting long non-coding RNAs in cancers: progress and prospects. Int J Biochem Cell Biol 45:1895–1910. doi:https://doi.org/10.1016/j.biocel.2013.05.030

Li P, Stetler RA, Leak RK, Shi Y, Li Y, Yu W, Bennett M, Chen J (2018) Oxidative stress and DNA damage after cerebral ischemia: potential therapeutic targets to repair the genome

and improve stroke recovery. Neuropharmacology 134:208–217. https://doi.org/10.1016/j.neuropharm.2017.11.011
Liang K, Zhu L, Tan J, Shi W, He Q, Yu B (2015) Identification of autophagy signaling network that contributes to stroke in the ischemic rodent brain via gene expression. Neurosci Bull 31:480–490. doi:https://doi.org/10.1007/s12264-015-1547-3
Ling MR, Chapple IL, Matthews JB (2015) Peripheral blood neutrophil cytokine hyper-reactivity in chronic periodontitis. Innate Immun 21:714–725. doi:https://doi.org/10.1177/1753425915589387
Liu SY, Chang Y, Zhang SY, Wu W, Sun BJ, Han XM (2011) Association of atherosclerotic cerebral infarction and human leukocyte antigen-DRB in a North Chinese Han population. Neurol India 59:521–526. doi:https://doi.org/10.4103/0028-3886.84330
Liu Y, Wang L, Predina J, Han R, Beier UH, Wang LC, Kapoor V, Bhatti TR, Akimova T, Singhal S, Brindle PK, Cole PA, Albelda SM, Hancock WW (2013) Inhibition of p300 impairs Foxp3+ T regulatory cell function and promotes antitumor immunity. Nat Med 19:1173–1177. https://doi.org/10.1038/nm.3286
Liu J, An P, Xue Y, Che D, Liu X, Zheng J, Liu Y, Yang C, Li Z, Yu B (2019) Mechanism of Snhg8/miR-384/Hoxa13/FAM3A axis regulating neuronal apoptosis in ischemic mice model. Cell Death Dis 10:441. doi:https://doi.org/10.1038/s41419-019-1631-0
Madjid M, Curkendall S, Blumentals WA (2009) The influence of oseltamivir treatment on the risk of stroke after influenza infection. Cardiology 113:98–107. doi:https://doi.org/10.1159/000172796
Mao X, Cai T, Olyarchuk JG, Wei L (2005) Automated genome annotation and pathway identification using the KEGG Orthology (KO) as a controlled vocabulary. Bioinformatics 21:3787–3793. doi:https://doi.org/10.1093/bioinformatics/bti430
Marchese FP, Grossi E, Marín-Béjar O, Bharti SK, Raimondi I, González J, Martínez-Herrera DJ, Athie A, Amadoz A, Brosh RJ, Huarte M (2016) A long noncoding RNA regulates sister chromatid cohesion. Mol Cell 63:397–407. https://doi.org/10.1016/j.molcel.2016.06.031
Marchese FP, Raimondi I, Huarte M (2017) The multidimensional mechanisms of long noncoding RNA function. Genome Biol 18:206. doi:https://doi.org/10.1186/s13059-017-1348-2
Murali V, Rathika C, Ramgopal S, Padma MR, Arun KM, Neethi AV, Jeyaram IK, Balakrishnan K (2016) Susceptible and protective associations of HLA DRB1*/DQB1* alleles and haplotypes with ischaemic stroke. Int J Immunogenet 43:159–165. doi:https://doi.org/10.1111/iji.12266
Nagel MA, Gilden D (2014) Update on varicella zoster virus vasculopathy. Curr Infect Dis Rep 16:407. doi:https://doi.org/10.1007/s11908-014-0407-z
O'Donnell MJ, Xavier D, Liu L, Zhang H, Chin SL, Rao-Melacini P, Rangarajan S, Islam S, Pais P, McQueen MJ, Mondo C, Damasceno A, Lopez-Jaramillo P, Hankey GJ, Dans AL, Yusoff K, Truelsen T, Diener HC, Sacco RL, Ryglewicz D, Czlonkowska A, Weimar C, Wang X, Yusuf S (2010) Risk factors for ischaemic and intracerebral haemorrhagic stroke in 22 countries (the INTERSTROKE study): a case-control study. Lancet 376:112–123. doi:https://doi.org/10.1016/S0140-6736(10)60834-3
Pertea M, Pertea GM, Antonescu CM, Chang TC, Mendell JT, Salzberg SL (2015) StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. Nat Biotechnol 33:290–295. doi:https://doi.org/10.1038/nbt.3122
Pertea M, Kim D, Pertea GM, Leek JT, Salzberg SL (2016) Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. Nat Protoc 11:1650–1667. doi:https://doi.org/10.1038/nprot.2016.095
Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK (2015) limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res 43:e47. doi:https://doi.org/10.1093/nar/gkv007
Sánchez Y, Huarte M (2013) Long non-coding RNAs: challenges for diagnosis and therapies. Nucleic Acid Ther 23:15–20. doi:https://doi.org/10.1089/nat.2012.0414
Sepramaniam S, Tan JR, Tan KS, DeSilva DA, Tavintharan S, Woon FP, Wang CW, Yong FL, Karolina DS, Kaur P, Liu FJ, Lim KY, Armugam A, Jeyaseelan K (2014) Circulating microRNAs as biomarkers of acute stroke. Int J Mol Sci 15:1418–1432. doi:https://doi.org/10.3390/ijms15011418
Serena J, Blanco M, Castellanos M, Silva Y, Vivancos J, Moro MA, Leira R, Lizasoain I, Castillo J, Dávalos A (2005) The prediction of malignant cerebral infarction by molecular brain barrier disruption markers. Stroke 36:1921–1926. doi:https://doi.org/10.1161/01.STR.0000177870.14967.94
Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 13:2498–2504. doi:https://doi.org/10.1101/gr.1239303
Smith CJ, Lawrence CB, Rodriguez-Grande B, Kovacs KJ, Pradillo JM, Denes A (2013) The immune system in stroke: clinical challenges and their translation to experimental research. J Neuroimmune Pharmacol 8:867–887. doi:https://doi.org/10.1007/s11481-013-9469-1
Snider SB, Jacobs CS, Scripko PS, Klein JP, Lyons JL (2014) Hemorrhagic and ischemic stroke secondary to herpes simplex virus type 2 meningitis and vasculopathy. J Neurovirol 20:419–422. doi:https://doi.org/10.1007/s13365-014-0253-7
Solomon DH, Kremer J, Curtis JR, Hochberg MC, Reed G, Tsao P, Farkouh ME, Setoguchi S, Greenberg JD (2010) Explaining the cardiovascular risk associated with rheumatoid arthritis: traditional risk factors versus markers of rheumatoid arthritis severity. Ann Rheum Dis 69:1920–1925. doi:https://doi.org/10.1136/ard.2009.122226
Strong K, Mathers C, Bonita R (2007) Preventing stroke: saving lives around the world. Lancet Neurol 6:182–187. doi:https://doi.org/10.1016/S1474-4422(07)70031-5
Tian X, Liu Y, Wang Z, Wu S (2020) lncRNA SNHG8 promotes aggressive behaviors of nasopharyngeal carcinoma via regulating miR-656-3p/SATB1 axis. Biomed Pharmacother 131:110564. doi:https://doi.org/10.1016/j.biopha.2020.110564
Urbanek C, Palm F, Grau AJ (2010) Influenza and stroke risk: a key target not to be missed? Infect Disord Drug Targets 10:122–131. doi:https://doi.org/10.2174/187152610790963474
Wilusz JE, Sunwoo H, Spector DL (2009) Long noncoding RNAs: functional surprises from the RNA world. Genes Dev 23:1494–1504. doi:https://doi.org/10.1101/gad.1800909
Yan Y, Zhang B, Liu N, Qi C, Xiao Y, Tian X, Li T, Liu B (2016) Circulating long noncoding RNA UCA1 as a novel biomarker of acute myocardial infarction. Biomed Res Int 2016:8079372. https://doi.org/10.1155/2016/8079372
Yang L, Lu ZN (2016) Long non-coding RNA HOTAIR promotes ischemic infarct induced by hypoxia through up-regulating the expression of NOX2. Biochem Biophys Res Commun 479:186–191. doi:https://doi.org/10.1016/j.bbrc.2016.09.023
Yang L, Froberg JE, Lee JT (2014) Long noncoding RNAs: fresh perspectives into the RNA world. Trends Biochem Sci 39:35–43. doi:https://doi.org/10.1016/j.tibs.2013.10.002
Young MD, Wakefield MJ, Smyth GK, Oshlack A (2010) Gene ontology analysis for RNA-seq: accounting for selection bias. Genome Biol 11:R14. doi:https://doi.org/10.1186/gb-2010-11-2-r14
Yuan X, Yan Y, Xue M (2021) Small nucleolar RNA host gene 8: A rising star in the targets for cancer therapy. Biomed Pharmacother 139:111622. doi:https://doi.org/10.1016/j.biopha.2021.111622

Zhang J, Yuan L, Zhang X, Hamblin MH, Zhu T, Meng F, Li Y, Chen YE, Yin KJ (2016) Altered long non-coding RNA transcriptomic profiles in brain microvascular endothelium after cerebral ischemia. Exp Neurol 277:162–170. doi:https://doi.org/10.1016/j.expneurol.2015.12.014

Zhu JR, Lu HD, Guo C, Fang WR, Zhao HD, Zhou JS, Wang F, Zhao YL, Li YM, Zhang YD, Yang CQ, Sun JG (2018) Berberine attenuates ischemia-reperfusion injury through inhibiting HMGB1 release and NF-κB nuclear translocation. Acta Pharmacol Sin 39:1706–1715. doi:https://doi.org/10.1038/s41401-018-0160-1

Zhuo LA, Wen YT, Wang Y, Liang ZF, Wu G, Nong MD, Miao L (2019) LncRNA SNHG8 is identified as a key regulator of acute myocardial infarction by RNA-seq analysis. Lipids Health Dis 18:201. doi:https://doi.org/10.1186/s12944-019-1142-0

Zis P, Stritsou P, Angelidakis P, Tavernarakis A (2016) Herpes simplex virus type 2 encephalitis as a cause of ischemic stroke: case report and systematic review of the literature. J Stroke Cerebrovasc Dis 25:335–339. https://doi.org/10.1016/j.jstrokecerebrovasdis.2015.10.002

Zöller B, Li X, Sundquist J, Sundquist K (2012) Risk of subsequent ischemic and hemorrhagic stroke in patients hospitalized for immune-mediated diseases: a nationwide follow-up study from Sweden. Bmc Neurol 12:41. doi:https://doi.org/10.1186/1471-2377-12-41