Journal of Advanced Research 24 (2020) 13-27



Contents lists available at ScienceDirect

Journal of Advanced Research

journal homepage: www.elsevier.com/locate/jare

Integrated analysis of the proteome and transcriptome in a MCAO mouse model revealed the molecular landscape during stroke progression



Litao Li^{a,1}, Lipeng Dong^{a,1}, Zhen Xiao^{b,1}, Weiliang He^a, Jingru Zhao^a, Henan Pan^{a,c}, Bao Chu^a, Jinming Cheng^a, Hebo Wang^{a,*}

^a Department of Neurology, Hebei General Hospital, Shijiazhuang 050051, Hebei, China ^b College of Life Sciences, Shanghai Normal University, Shanghai 200234, China ^c North China University of Science and Technology, Tangshan 063210, Hebei, China

HIGHLIGHTS

G R A P H I C A L A B S T R A C T

- DIA proteomics was applied to MCAO mice detection for the first time.
- Proteomics and bioinformatics revealed relationship between stroke process and immunity, especially inflammation.
- C3, Apoa4 and S100a9 were highlighted as a marker or drug targets for stroke.



ARTICLE INFO

Article history: Received 8 October 2019 Revised 10 January 2020 Accepted 10 January 2020 Available online 12 January 2020

Keywords: Stroke MCAO Proteome Immunity

ABSTRACT

Strokes usually results in long-term disability and death, and they occur worldwide. Recently, increased research on both on the physiopathological mechanisms and the transcriptome during stroke progression, have highlighted the relationship between stroke progression and immunity, with a special focus on inflammation. Here, we applied proteome analysis to a middle carotid artery occlusion (MCAO) mouse model at 0 h, 6 h, 12 h and 24 h, in which proteome profiling was performed with 23 samples, and 41 differentially expressed proteins (DEPs) were identified. Bioinformatics studies on our data revealed the importance of the immune response and particularly identified the inflammatory response, cytokine- cytokine receptor interactions, the innate immune response and reactive oxygen species (ROS) during stroke progression. In addition, we compared our data with multiple gene expression omnibus (GEO) datasets with and without a time series, in which similar pathways were identified, and three proteins, C3, Apoa4 and S100a9, were highlighted as markers or drug targets for stroke; these three

Peer review under responsibility of Cairo University.

https://doi.org/10.1016/j.jare.2020.01.005

2090-1232/© 2020 THE AUTHORS. Published by Elsevier BV on behalf of Cairo University.

This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

^{*} Corresponding author at: Department of Neurology, Hebei General Hospital, 348 Heping Xi Road, Shijiazhuang 050051, China. *E-mail address:* wanghbhope@hebmu.edu.cn (H. Wang).

¹ These authors contributed equally to the work.

proteins were significantly upregulated in the MCAO model, both in our proteomic data and in the GEO database.

© 2020 THE AUTHORS. Published by Elsevier BV on behalf of Cairo University. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Introduction

Ischemic stroke is a cardiovascular disease, and it represents a leading cause of morbidity and mortality worldwide, leading to a major economic burden [1]. Prevention and therapy need to be improved to tackle this public health problem and to improve ischemic stroke outcome. Until now, there have been few therapeutic options used in clinical treatment for stroke patients. It is worth continuing to research novel therapeutic interventions for the prevention and recovery of stroke.

Recently, increasing evidence has shown that inflammation and the immune response after stroke are closely associated with stroke severity and outcome [2,3], which has been confirmed in animal studies [4]. More concretely, the suppression of innate and adaptive immune system pathways caused after a stroke increases the risk of poststroke infection, contributing to increased mortality and poor outcome [5]. Microarray studies demonstrated that after an ischemic stroke, global mRNA and non-coding RNA expression profiles are altered extensively in the blood or brain [6–8], providing important new insight into this disease at the transcriptional and translational levels. However, to deepen stroke understanding and uncover novel mechanisms for treatment, scientific investigation of the identified gene products at the protein level is eagerly needed [9]. Proteomics, the analysis of protein expression in biosamples, can improve our understanding of pathophysiological mechanisms and aid in diagnostic tool development [10]. As the earliest proteomic study on stroke, Sironi's group conducted an analysis of serum and urine samples from spontaneously hypertensive stroke-prone rats using 2D gel electrophoresis [11]. It was found that thiostatin, a marker of the inflammatory response, could be detected in the serum at least 4 weeks before the occurrence of a stroke. Wen et al. used a proteomics method to study the rat cerebral cortex in the subacute to long-term phases of cerebral ischemia-reperfusion injury [12]. They found that with the elongation of reperfusion time to the long-term phase, a tendency of recovery was detected in the cytoskeleton, while inflammation pathways that were different than those in the subacute phase were activated [12]. In addition to the study of animal models, many clinical proteomic analyses have been carried out. A proteomic substrate profiling study indicated that there was an alteration in matrix metalloproteinase protein (MMP) profiles in stroke patients who received rtPA treatment when compared to the profiles the control group [13]. More recently, proteomic analysis showed that inflammatory proteins and a marker (thrombospondin-1) for barrier dysfunction were increased when cerebral endothelial cells responded to oxidative stress in vitro in humans. This result was confirmed with plasma from stroke patients [14]. Similar results were also reported for the experimental ischemic response in both mouse and rat brain endothelial cells [15,16].

Although an increasing number of both clinical and preclinical studies support that the inflammatory immune response is important after stroke, no proteomics study has systemically revealed the mechanism of inflammation and the immune response after stroke.

Here, we established an MCAO mouse model and performed proteome profiling on cerebral cortex samples at 0 h, 6 h, 12 h and 24 h after operation. Several samples were obtained at 0 h, and 6 samples were obtained at 6 h, 12 h, and 24 h. As a result, a landscape of proteome profiling with 23 samples was obtained. The different time points were compared with 0 h and revealed that 41 genes were significantly different. Gene enrichment analysis of the 41 genes highlighted the importance of the immune response with an emphasis on the following: inflammatory response, cytokine-cytokine receptor interaction, and innate immune response. Furthermore, through various bioinformatics methods, including Short Time-series Expression Miner (STEM) analysis, weighted gene co-expression network analysis (WGCNA) and gene set enrichment analysis (GSEA), co-expressing gene networks were established. Genes related to immunity were generally upregulated, while genes related to neurodevelopment were significantly downregulated throughout stroke progression. In addition, we compared our data with multiple GEO datasets with or without time series, in which similar pathways were identified. Three proteins, C3, Apoa4 and S100a9, were highlighted as markers or drug targets for stroke; they were significantly upregulated in the MCAO model proteomic data and GEO database.

Materials and methods

Animals

Eight- to ten-week-old male C57BL/6 mice were purchased from Charles River Laboratories (Beijing, China) and used for construction of a middle carotid artery occlusion model. All animal experiments were approved by the Ethics Committee of Hebei General Hospital (approval no. 201921). Animals were grouped and housed at a controlled temperature ($20 \pm 2 \, ^{\circ}$ C) with a 12 h light-dark cycle, and they had free access to food and water.

Transient middle cerebral artery occlusion

A MCAO mouse model was constructed in our study, as described in a previous study [17]. Transient focal cerebral ischemia was induced by right middle cerebral artery occlusion (MCAO). Mice were anesthetized with 10% hydral (350 mg/kg). Body temperature was monitored and maintained at 36.5–37.5. Briefly, after making an incision in the midline skin, the right common carotid artery, external carotid artery (ECA), and internal carotid artery (ICA) were exposed. Then, a 6–0 nylon monofilament with a rounded tip was inserted into the right ICA through the broken end of the ECA to block the origin of the MCA. Cerebral ischemia through the intraluminal suture was maintained for 60 min, and it was followed by removal of the monofilament and reperfusion. At 6, 12, 24 h hours after MCAO, neurological deficit scores were assessed, and then the brains were obtained by decapitation.

Analysis of neurological deficit

All behavioral tests in mice were conducted in a quiet, dimly lit room by an experimenter who was blinded with respect to the experimental groups; analysis was performed at 6, 12, and 24 h after MCAO. Neurologic deficits were scored as follows: 0, no deficits; 1, difficulty in fully extending the contralateral forelimb; 2, unable to extend the contralateral forelimb; 3, mild circling to the contralateral side; 4, severe circling; and 5, falling to the contralateral side. A higher neurological deficit score indicated a more severe impairment of motor motion injury.

Determination of infarct volume

At 0, 6, 12, and 24 h after MCAO, mice were anesthetized with chloral hydrate. Brains were dissected and cut into 4 coronal slices; the slices were incubated in 2% 2, 3,5-triphenyltetrazolium chloride (TTC) for 15 min at 37.0 °C, which was followed by immersion-fixation in 4% paraformaldehyde.

Sample preparation

Twenty-three MCAO mice were randomly divided into four groups (5, 6, 6, and 6/group), and the ischemic cerebral cortex was removed at 6 h, 12 h or 24 h after operation, and cerebral cortex of 0 h was the control. The samples were ground in liquid nitrogen, and then lysis buffer (1% SDS, 7 M urea, 1x Protease Inhibitor Cocktail (Roche Ltd. Basel, Switzerland)) was added to the samples, which was followed by vortexing and crushing three times for 400 s. The samples were then lysed on ice for 30 min and centrifuged at 12000 rpm for 20 min at 4. The supernatant was collected and transferred to a new Eppendorf tube. The protein concentration of the supernatant was determined using the bicinchoninic acid (BCA) protein assay, and 20 μ g of protein per condition was used for SDS-PAGE.

Protein digestion

One hundred micrograms of protein per condition was transferred into a tube, and the final volume was adjusted to 100 μ L with 8 M urea. Then, 2 μ L of 0.5 M TCEP was added, and the sample was incubated at 37 for 1 h. Next, 4 μ L of 1 M iodoacetamide was added to the sample, and it was incubated for 40 min in the dark at room temperature. Next, five volumes of prechilled (-20) acetone were added to precipitate the proteins overnight at -20. The precipitates were washed twice with 1 mL of prechilled 90% acetone aqueous solution and then were redissolved in 100 μ L of 100 mM TEAB. Sequence grade modified trypsin (Promega, Madison, WI) was added at a ratio of 1:50 (enzyme: protein, weight: weight) to digest the proteins at 37 overnight. The peptide mixture was desalted with a C18 ZipTip, quantified by a Pierce^{IM} Quantitative Colorimetric Peptide Assay (#23275) and then lyophilized with a SpeedVac.

Data-dependent acquisition (DDA) mass spectrometry

The peptide mixture was redissolved in buffer A (buffer A: 20 mM ammonium formate in water, pH 10.0, adjusted with ammonium hydroxide) and then fractionated by an Ultimate 3000 system (Thermo Fisher Scientific, MA, USA) connected to a reversed-phase column (XBridge C18 column, 4.6 mm \times 250 mm, 5 µm, Waters Corporation, MA, USA). High pH separation was performed using a linear gradient, ranging from 5% B to 45% B, over 40 min (B: 20 mM ammonium formate in 80% ACN, pH 10.0, adjusted with ammonium hydroxide). The column was reequilibrated at the initial condition for 15 min. The column flow rate was maintained at 1 mL/min, and the column temperature was maintained at 30 . Ten fractions were collected, and each fraction was dried in a vacuum concentrator. The fractions were redissolved in solvent A (A: 0.1% formic acid in water) and were analyzed by on-line nanospray LC-MS/MS on an Orbitrap Fusion Lumos Tribrid (Thermo Fisher Scientific, MA, USA) coupled to a Waters nano ACQUITY UPLC system (Waters, MA, USA). A 2 µL peptide sample was loaded onto an analytical column (Acclaim Pep-Map C18, 75 μ m \times 25 cm) and separated over 120 min with a gradient ranging from 3% to 30% B (B: 0.1% formic acid in ACN). The electrospray voltage of 2.1 kV versus the inlet of the mass spectrometer was used. The mass spectrometer was run in a data-dependent acquisition mode, and it automatically switched between MS and MS/MS mode. The parameters used were as follows: (1) MS: scan range (m/z) = 350-1550; resolution = 60,000; AGC target = 4e5; maximum injection time = 50 ms; dynamic exclusion = 30 s; (2) HCD-MS/MS: resolution = 15,000; AGC target = 5e4; maximum injection time = 35 ms; collision energy = 30. DDA Raw data were processed and analyzed bv Spectronaut X (Biognosys AG, Switzerland) with default settings to generate an initial target list. Spectronaut was used to search the UniProt-Proteome Mus musculus 201,711 database, assuming trypsin as the digestion enzyme. Carbamidomethyl (C) was specified as the fixed modification. Oxidation (M) was specified as the variable modifications. A Q value (FDR) cutoff on the precursor and protein level was applied at 1%.

Quantification of proteins using DIA mass spectrometry

The mass spectrometer was run with a data-independent acquisition mode and automatically switched between MS and MS/MS mode. The parameters used were as follows: (1) MS: scan range (m/z) = 350-1500; resolution = 60,000; AGC target = 4e5; maximum injection time = 50 ms; (2) HCD-MS/MS: resolution = 30,00 0; AGC target = 1e6; collision energy = 32; stepped CE = 5%; (3) DIA was performed with variable isolation window, each window overlapped by 1 m/z, and the window number was 45; total cycle time was 3.98 s. Raw DIA data were processed and analyzed by Spectronaut X (Biognosys AG, Switzerland) using the default settings. The retention time prediction type was set to dynamic iRT. Data extraction was determined by Spectronaut X based on extensive mass calibration. Spectronaut Pulsar X was used to dynamically identify the ideal extraction window depending on iRT calibration and gradient stability. A Q value (FDR) cutoff on the precursor and protein level was applied at 1%. Decoy generation was set to mutate, which applied a random number of AA position swamps (min = 2, max = length/2). All selected fragment ions passing the filters were used for quantification. MS2 interference removed all interfering fragment ions except for the 3 with the least interference. The average top 3 filtered peptides that passed the 1% Q value cutoff were used to calculate major group quantities.

RNA extraction and qRT-PCR

Total RNA was extracted from cerebral cortex by using TRIzol reagent (Invitrogen, Thermo Fisher Scientific, Inc., Waltham, MA, USA). RNA was converted into cDNA using the Reverse Transcription Kit (Takara, Dalian, China), after which cRNA was synthesized from total RNA using SYBR Premix Dimmer Eraser kit (Takara) according to the manufacturer's instructions. Primer sequences are summarized in Table 1. The relative mRNA levels of C3, Apoa4 and S100a9 were calculated by the $2^{-\Delta\Delta Ct}$ method and normalized to the internal control GAPDH.

Weighted gene coexpression network analysis

WGCNA uses the topological overlapping measurement to identify corresponding expression modules. These expression modules were analyzed based on their eigengene correlation with each diffusion parameter. WGCNA is a robust tool for integrative network analysis, and it has been used in several recent studies [18–20]. In addition, a permutation-based preranked GSEA was applied to each expression module to verify its pathways [21]. The predefined gene sets from the Molecular Signature Database v5.1 were used. Networks were exported to Cytoscape 2.0 for further visualization. The WGCNA integrated function (exportNetworkToCytoscape) was used to calculate a weighted network.

Table 1Primers used for quantitative RT-qPCR analysis.

Gene name	Sequence
GAPDH	F: GGAGCGAGATCCCTCCAAAAT R: GGCTGTTGTCATACTTCTCATGG
C3	F: TCCAACAAGAACACCCTCA
APOA4	F: CACCTGAAGCCCTATGCC
S100A9	R: CTCCTTGATCGTGGTCTGC F: ATGGTGGAAGCACAGTTG R: TGGTTTGTGTCCAGGTCC

Gene set enrichment analysis

Permutation-based gene set enrichment analysis was performed for each expression module to find specifically enriched biological functions and related pathways [21]. Preranked GSEA was performed with 1000 permutations. P-values were calculated by the familywise error rate (FWER), which is a robust method for testing multiple samples [22]. The Molecular Signatures Database version 5.0 was used, and it included pathway gene sets (C2) (http://www.broadinstitute.org/gsea) as input databases for analysis. GSEA plots were visualized using the limma R-package (barcodeplot function).

Temporal gene expression profiles

We used the Short Time-series Expression Miner program [23] to analyze differentially expressed genes and identify temporal expression profiles. Genes whose expression levels met the 1.5-fold change criterion at any time point were used, and STEM profiles were clustered with all parameters set to the default value. Temporal expression profiles that showed statistically significant variation from the time series were corrected using a false discovery rate (FDR) calculation that was performed on 1000 randomly selected permutations.

Bioinformatic analysis

To identify differentially expressed proteins between different MCAO stages, differential expression analyses were performed using cuffdiff. P-values < 0.05 and FC > 1.5 were selected as the criteria for identifying significantly differentially expressed proteins. To obtain an overview of the characteristics of DEPs, the R package was used to generate a Venn diagram and heatmaps, and hierarchical clustering analysis was performed based on the normalized values of all proteins. Gene enrichment analysis was performed based on KEGG and Gene ontology, and a p value of < 0.05 was set as the cutoff for significantly enriched functional GO term or KEGG pathways. Protein-protein interaction was assessed by String database and visualized by Cytoscape.

GEO dataset download and raw data preprocessing

The raw data of mRNA expression dataset GSE61616, GSE78731, GSE23160, GSE58294 and GSE119121 (Table 2) were screened from the National Center of Biotechnology Information

Table	22	
0.00		

GEO datasets used in this study.

(NCBI) Gene Expression Omnibus (GEO, http://www.ncbi.nlm.nih.gov/gds/). The GSE61616 and GSE78731 datasets are each composed of 10 samples, including 5 MCAO and 5 control rat brains. For dataset GSE23160, there are 32 mRNA expression data from mouse brain samples. For dataset GSE119121, there are 47 mRNA expression data from rat blood samples at 6 different time point. For dataset GSE58294, there are 92 mRNA expression data from human blood samples at 4 different time point. The R software package was used to process the downloaded files and to convert and reject the unqualified data. The data were calibrated, standardized, and log2 transformed. The differently expressed mRNA were screened using Limma package, with the criterion of |log 2(fold change [FC]| > 2 and *P*-value < 0.05.

Results

The neurological deficit and brain infarct volume in middle cerebral artery occlusion

To examine the neurological deficit effect of cerebral ischemia, neurologic deficits were examined and scored at 0, 6, 12, and 24 h after MCAO. Compared with mice in the 0 h group, 6–24 h mice showed significant neurological deficits after MCAO (Fig. 1A, *p < 0.05), suggesting nerve injury in acute stroke. The cerebral infarction was detected by triphenyltetrazolium chloride (TTC) staining and is displayed in Fig. 1B. No infarction was observed in the 0 h group animals, while extensive lesions developed in the animals in the 6–24 h groups (Fig. 1B, *p < 0.05).

Proteome analysis of 4 time points following MCAO

To describe the development features of stroke, we established a mouse MCAO model according to a previous protocol [17]. The cerebral cortexes at 0 h, 6 h, 12 h and 24 h after operation were obtained and subjected to LC-MS/MS analysis through a Data-Independent Acquisition (DIA) strategy to obtain their proteome profiles. As a result, 4995 proteins were successfully identified.

Then, we analyzed the proteome features at 4 time points following stroke development in the MCAO model. In total, 109 proteins, 87 proteins and 92 proteins were identified as differentially expressed proteins (DEPs, p < 0.05, FC > 1.5) at 6 h, 12 h and 24 h compared with 0 h. In addition, 41 proteins were consistently highly expressed during stroke development in the MCAO model (Fig. 2A). The unsupervised hierarchical clustering of the 41 DEPs among 23 experiments at 4 time points identified the features that were most different from those of the 0 h group (Fig. 2B). To describe the overall expression trend of 41 DEPs during stoke development, we mapped the protein expression intensity to the 4 time points (Fig. 2B), and the result suggested that the expression of most genes was upregulated after MCAO stimulation (Fig. 2B).

To identify the biological function and network of the 41 DEPs, we performed enrichment analysis with ClueGo plug-in of Cytoscape software. As a result, 4 terms were discovered: acute inflammatory response, complement and coagulation cascades, cellular oxidant detoxification and protein activation cascade (Fig. 2C). Furthermore, we conducted enrichment analysis of the DEPs at 6 h,

Dataset	Platform	Species	Samples	Tissue type	Targets	Time series
GSE61616	GPL1355	Rat	5 MCAO + 5 Sham	Brain	RNA	No
GSE78731	GPL15084	Rat	5 MCAO + 5 saline	Brain	RNA	No
GSE23160	GPL6885	Mouse	(8)*(0 h, 2 h, 8 h, 24 h)	Brain	RNA	Yes
GSE119121	GPL6247	Rat	(8)*(0 h, 1 h, 2 h, 3 h, 24 h)+(7)*(6 h)	Blood	RNA	Yes
GSE58294	GPL570	Human	(23)*(0 h, 3 h ,5h, 24 h)	Blood	RNA	Yes



Fig. 1. The neurologic deficit scores and TTC staining. (A) The neurologic deficit scores were performed 0, 6, 12, and 24 h after MCAO; (B) Brain slices stained with TTC 24 h after MCAO. The white section of the slices represents the ischemic area. The red section of the slices represents normal tissue.



Fig. 2. Proteomics investigation of gene expression at 0 h, 6 h, 12 h and 24 h following the introduction of a middle cerebral artery occlusion. (A) A total of 109, 87 and 92 proteins were identified as differentially expressed genes (DEGs) at 6 h, 12 h and 24 h, respectively, and 41 proteins were identified in all three DEG groups; (B) Unsupervised hierarchical clustering of the 41 DEPs among 23 experiments at 4 time points, 0 h group was used as a control; (C) The biological function of 41 DEGs was annotated, and a gene-biological process network was constructed. (D) A total of 41 genes were subjected to enrichment analysis through GO-BP and KEGG databases, and the significance of genes at 6 h, 12 h, and 24 h are shown.

12 h and 24 h compared with those at 0 h using the GO-BP and KEGG databases and showed significantly enriched terms at the 3 time points (Fig. 2D). Interestingly, immunity-related functions, including innate immune response, response to cytokines, acute phase response and complement and coagulation cascades, were significantly enriched, suggesting a fundamental role for the immune response in stroke development after MCAO stimulation.

The expression patterns of 41 DEPs in the MCAO model were consistent with the data from GEO datasets

To verify the expression level of the 41 DEPs identified above, 2 GEO datasets, GSE78731 and GSE61616, were used. The datasets were all collected on samples of MCAO rat models, and 5 MCAO samples were compared to 5 sham samples (Table 2). Then, we compared the expression of 41 DEGs in the MCAO and sham samples in the 2 datasets. The expression trends of most genes in the 2 datasets were consistent with our proteome profiling (Fig. 3A, B). For example, expression of Atp1a4 was significantly downregulated in MCAO samples in both GEO datasets. In addition, expression of Apoa4, Fga, Hpx, Plg, Pzp, Serpina3 and serpinc1 in GSE78731 and the expression trends of most genes in the expression of the expression of Apoa4, Fga, Hpx, Plg, Pzp, Serpina3 and serpinc1 in GSE78731 and the expression of Apoa4.

sion of Ahsg, Alb, Apoa1, Gc and Orm1 in GSE61616 were upregulated in MCAO samples, and our proteome profiling showed the same results (Fig. 2D). Furthermore, C3 and S100a9, which were related to immune response and complement and coagulation cascades, were upregulated in MCAO samples in both datasets and proteomes. QRT-PCR was used to validate the expression in the cerebral cortex of 4 Sham, 4 MCAO samples, and the results were consistent with our proteomics and GEO dataset results (Fig. 3C).

The results verified and strengthened the credibility of the 41 DEPs identified in normal and stroke samples (Fig. 2B, C). Furthermore, the integrated analysis also highlighted the function of the immune response during stroke progression.

STEM analysis of proteome profiling

To explore the gene regulation network during stroke development, we next conducted STEM analysis for proteome profiling of the 4 time points of the MCAO model. As a result, two clusters were identified significantly, including cluster 13 with 99 proteins (with p value = 2e-3) and cluster 0 with 42 proteins (with p value = 7e-11) (Fig. 4A). We portrayed their expression trends for the 2 clusters



Fig. 3. Validation of the expression of 41 DEGs. GSE78731 (A) and GSE61616 (B) were applied to verify the expression of 41 DEGs. Then we used qRT-PCR to validate our results(C). *, p < 0.05, ** p < 0.01, and *** p < 0.001.



Fig. 4. Short Time-series Expression Miner (STEM) analysis identified two expression modules. Short Time-series Expression Miner analysis was employed to explore gene regulation networks during stroke progression. Two significant clusters were identified, namely, cluster 13 and cluster 0 (A-B). (C) Gene expression in cluster 13 and cluster 0 at 4 time points following MCAO. (D) GO-BP and KEGG pathway enrichment of genes in cluster 13. (E) The biological function of genes in cluster 13 was determined, and a gene-biological process network was constructed. (F) Gene expression patterns of cluster 13 in 23 experiments are shown.

during stoke development in the MCAO model among the 23 experiments (Fig. 4B and C). To research the main biological function of genes in the 2 clusters, we performed enrichment analysis using GO or KEGG databases. As expected, immune-related functions, such as complement and coagulation cascades, negative regulation of peptidase activity and innate immune response, were significantly enriched for genes in cluster 13, which were upregulated during stroke development. In addition, mainly related functions of the brain, such as neuron projection and central nervous system neuron development, were significantly enriched for genes in cluster 0, which were downregulated during stroke development. The results suggested that many genes participating in immune regulation were activated, while the genes related to brain development were downregulated during stroke progression.

To verify this conclusion, the genes in cluster 13 were subjected to Cluego analysis, and 5 critical biological processes were significantly identified: negative regulation of hydrolase activity, negative regulation of lipase activity, acute inflammatory response, complement activation and neuron projection regeneration (Fig. 4D). The gene expression patterns of genes in Cluster 13 for the 23 experiments are shown in Fig. 4F.

The STEM protein profiling signature was confirmed by GEO datasets

To test the reliability of the results in STEM analysis, 3 GEO datasets that used different time series and species to study stroke development after MCAO stimulation (GSE23160, GSE58294 and GSE119121) were applied and subjected to STEM analysis. GSE23160 covered 32 experiments, including 8 experiments of each 0 h, 2 h, 8 h and 24 h of MCAO models of mice. The STEM analysis of GSE23160 identified 2 significant clusters: cluster 12 (p value = 4e-75) with 131 proteins and cluster 13 (p value = 1e-4) with 16 proteins. The expression of genes in the two clusters was upregulated during stroke development (Fig. 5A, B). To test the main function of genes in two clusters, we performed enrichment analysis on the two clusters. As a result, critical terms related to immunity, such as inflammatory response, cytokine-cytokine receptor interaction, innate immune response and neutrophil chemotaxis, were significantly enriched for genes in cluster 12. Similarly, immunity-related functions, such as response to stress and negative regulation of the inflammatory response, were enriched for genes in cluster 13. The results support the important role of the immune response in the stroke process. In addition, negative regulation of protein kinase activity and regulation of transcription from the RNA polymerase II promoter were enriched for genes in cluster 13, indicating possible roles for kinases and transcription regulation during stroke development (Fig. 5C).

Similarly, we performed STEM analysis on the GSE58294 and GSE119121 datasets. GSE58294 contained 23 S patient samples collected at 0 h, 3 h, 5 h and 24 h, which formed a large dataset that possessed 96 blood samples in total (Table 2). STEM analysis for GSE58294 identified 8 significant clusters: cluster 48 (p value = 3e-247) with 495 proteins, cluster 12 (p value = 4e-148) with 485 proteins, cluster 49 (p value = 1e-130) with 366 proteins, cluster 0 (p value = 2e-89) with 161 proteins, cluster 42 (p value = 6e-40) with 55 proteins, cluster 1 (p value = 1e-6) with 43 proteins, cluster 2 (*p* value = 1e-6) with ** proteins and cluster 23 (p value = 6e-4) with ** proteins (Fig. 6A). The expression patterns of genes in the first 6 clusters at different time points of stroke development are illustrated in Fig. 6B. In summary, genes were upregulated in cluster 48, cluster 49 and cluster 42, while they were downregulated in cluster 12, cluster 0 and cluster 1 during stroke progression. Enrichment analysis for the genes in the first 4 clusters indicated the same results as what was found in previous analysis, in which immune-related functions were significantly enriched for upregulated clusters. For example, genes in cluster 48 were significantly enriched in the innate immune response and in the interferon-gamma-mediated signaling pathway, genes in cluster 49 were significantly enriched in T cell proliferation and the inflammatory response, and genes in cluster 42 were significantly enriched in positive regulation of T cell proliferation. Genes in cluster 12, which were downregulated during stroke development, were mostly enriched in transcription regulation (Fig. 6C).

For dataset GSE119121, which contained 8 samples collected at 0 h, 1 h, 2 h, 3 h, and 24 h and 7 samples collected at 6 h from rat MCAO models, STEM analysis identified 4 significant clusters: cluster 24 (*p* value = 2e-209), cluster 4 (*p* value = 8e-118), cluster 8 (*p* value = 2e-34) and cluster 3 (*p* value = 5e-10) (Fig. 7A). A total of 342 and 119 genes in clusters 24 and 8, respectively, were upregulated during stroke development, while 485 and 95 genes in clusters 4 and 3 were downregulated during stoke development (Fig. 7B). The genes in cluster 24 and cluster 8 were significantly enriched in the innate immune response, the apoptotic process and cytokine-cytokine receptor interaction (Fig. 7C), indicating a similar result from STEM analysis of proteome data that was found in the other two GEO datasets.

A gene co-expression network generated from proteome profiling of stroke progression

To further identify the co-expression patterns of genes during stroke development, we performed weighted gene co-repression analysis (WGCNA) on our proteome profiles through the R package WGCNA algorithm. First, we performed sample clustering analysis to detect variations of 23 samples and outliers. As a result, no outlier was identified, and all of the samples were used in next step analysis (Fig. 8A). Next, we set the power as 7 and the threshold as 9, which made the network similar to a scale free network (Fig. 8B). Then, Person correlation coefficients were calculated for pairwise genes to yield a similarity matrix, which was transformed into an adjacency matrix using the threshold and power values listed above. Average linkage hierarchical clustering was then performed to identify modules with densely interconnected genes. and genes that were not assigned to specific modules were colored gray. As a result, we found one module with significantly coexpressed genes (turquoise, p value = 0.008) (Fig. 8C) that were tightly related to stroke development (Fig. 8D), indicating an important gene set for further study of stroke.

We next established the PPI network for the 68 genes in the turquoise module, which identified 458 PPIs (Fig. 8E). Then, we matched the nodes in the network with genes in cluster 13 in the STEM analysis of proteome profiling. As shown in Fig. 4E, the genes, such as Apla1, Kng1, Pzp, Plg and Olf4613, which were included both in cluster 13 and the turquoise module, possessed a tight connection with other genes and acted as the central hub in the network. To detect the fundamental function of the gene in the turquoise module, we performed enrichment analysis using GO and KEGG datasets. As expected, immune-related functions, such as complement and coagulation cascades, acute phase response, and hemostasis, were significantly enriched, further indicating the importance of the immune response during stroke development (Table 3).

GSEA of hallmark genes in the proteome profiling

To search for the basic function of the overall genes in our proteome profiling, gene set enrichment analysis was performed, and the results are summarized in Fig. 9A. The genes in the first five terms with high enrichment scores (ES), including coagulation, complement, oxygen species pathway, interferon gamma response and apoptosis, were upregulated during the stroke process



Profile	Ierm	Count	Gene Ratio	Enrichment Score
	Inflammatory Response	21	16.76%	14.26
	Cytokine-Cytokine Receptor Interaction	13	10.37%	7.22
	Innate Immune Response	13	10.37%	5.54
	Neutrophil Chemotaxis	12	9.58%	12.78
	Immune Response	12	9.58%	6.37
	Immune System Process	12	9.58%	4.93
	Positive Regulation Of Angiogenesis	11	8.78%	8.76
	Positive Regulation Of Gene Expression	11	8.78%	4.03
	Chemotaxis	10	7.98%	7.63
12	Angiogenesis	10	7.98%	5.04
	Positive Regulation Of Inflammatory Response	9	7.18%	8.60
	Positive Regulation Of Erk1 And Erk2 Cascade	9	7.18%	4.92
	Response To Lipopolysaccharide	9	7.18%	4.77
	Chemokine Signaling Pathway	9	7.18%	4.29
	Chemokine-Mediated Signaling Pathway	8	6.38%	7.59
	Response To Virus	8	6.38%	6.30
	Cellular Response To Tumor Necrosis Factor	8	6.38%	5.51
	Response To Drug	8	6.38%	2.49
	Cell Adhesion	8	6.38%	1.68
	Positive Regulation Of Cell Proliferation	8	6.38%	1.45
	Response To Stress	3	14./1%	3.49
	Negative Regulation Of Inflammatory Response	3	14.71%	2.70
	Negative Regulation Of Protein Kinase Activity	3	14.71%	2.62
	Cellular Oxidant Detoxification	2	9.81%	2.41
	Activation Of Mapkkk Activity	2	9.81%	2.21
13	Positive Regulation Of P38Mapk Cascade	2	9.81%	1.88
	Erythrocyte Development	2	9.81%	1.68
	Activation Of Mapkk Activity	2	9.81%	1.62
	Positive Regulation Of Apoptotic Process	3	14./1%	1.57
	Positive Regulation Of Cell Differentiation	2	9.81%	1.46
	Regulation Of Transcription From Rna Polymerase li Promoter	3	14./1%	1.43
	African Trypanosomiasis	2	9.81%	1.40

Fig. 5. GSE23160 datasets were used to validate the STEM analysis results. (A) Two significant clusters, cluster 12 with 131 proteins and cluster 13 with 16 proteins, were identified. (B) Gene expression patterns of cluster 12 and cluster 13 from the 32 samples are illustrated. (C) Functional enrichment results of genes in cluster 12 and cluster 13 are shown.

(Fig. 9B), indicating a fundamental function of the immune response during stroke development. Finally, we used the GO term (Fig. 9C), miRNA (Fig. 9D) and transcription factors (Fig. 9E) of GSEA to compare 6 h vs 0 h, 12 h vs 6 h and 24 h vs 12 h, and the enrichment scores are shown by a heatmap. Furthermore, detailed information of GO terms, miRNAs and transcription factors for the 3 comparisons are shown in a Venn diagram (Figs. S1, S2, S3).

Discussion

Strokes occur worldwide, and they usually result in long-term disability and death. Brain inflammation contributes to secondary injuries following ischemia and stroke [18–20], involving several

factors triggering the inflammatory response process. In the case of the ischemic brain, chemokines were upregulated to stimulate inflammatory cell entry into the brain, especially around the penumbra, or the infarct's border. Adhesion molecules on activated endothelia in turn lead to the adhesion of circulating leukocytes, causing microvascular occlusions and an infiltration of immune cells into the brain parenchyma [21,22]. Then, numerous inflammation-associated substances are secreted from activated immune cells, including ROS, NO, cytokines, chemokines, prostaglandins, leukotrienes, and platelet-activating factor (PAF) [23– 26]. They directly or indirectly participate in the inflammatory response, resulting in worse outcomes.

When the brain suffers acute insults, including stroke, cytokine production increases in the brain. Although it is clear that some of the cytokines are produced by microglia and infiltrating leuko-



C							
0	Profile	Term	Count	Gene Ratio	Enrichme		
	. rome		oount	(100%)	nt Score		
		Innate Immune Response	23	4.80%	5.08		
		Proteolysis	17	3.54%	1.79		
		Response To Lipopolysaccharide	9	1.88%	2.06		
		Blood Coagulation	9	1.88%	1.78		
		Interferon-Gamma-Mediated Signaling Pathway	8	1.67%	3.60		
		T Cell Receptor Signaling Pathway	8	1.67%	1.80		
		Extracellular Matrix Disassembly	7	1.46%	2.65		
		Leukocyte Migration	7	1.46%	1.68		
	48	Viral Entry Into Host Cell	6	1.25%	1.87		
		Positive Regulation of Catalytic Activity	6	1.25%	1.85		
		Positive Regulation of Peptidyl-Tyrosine Phosphorylation	6	1.25%	1.83		
		Response To Virus	6	1.25%	1 34		
		Protein Secretion	5	1.04%	2.64		
		Myd88-Dependent Toll-Like Receptor Signaling Pathway	5	1.04%	2.59		
		Triglyceride Metabolic Process	5	1.04%	2.49		
		Copii Vesicle Coating	5	1.04%	1.65		
		Cell Morphogenesis	5	1.04%	1.62		
		Regulation of Inflammatory Response	28	1.04%	2.07		
		Protein Phosphorylation	17	4.32%	3.05		
		Apoptotic Process	15	3.81%	1.47		
		Innate Immune Response	13	3.31%	1.67		
		Intracellular Signal Transduction	12	3.05%	1.51		
		Inflammatory Response	11	2.80%	1.34		
		Viral Process	10	2.54%	1.55		
		Response To Drug	10	2.54%	1.51		
		Mapk Cascada	9	2.29%	3.30		
	49	Protein Autophosphorylation	8	2.29%	1.46		
		Movement of Cell Or Subcellular Component	7	1.78%	2.84		
		Positive Regulation of Peptidyl-Serine Phosphorylation	6	1.53%	2.49		
		Activation of Mapk Activity	6	1.53%	1.73		
		Regulation of Gtpase Activity	5	1.27%	1.86		
		Cellular Response To Mechanical Stimulus	5	1.27%	1.73		
		Ephrin Receptor Signaling Pathway	5	1.27%	1.46		
		I Cell Proliferation	4	1.02%	2.03		
		Thymus Development	4	102%	1.63		
		Ubiguitin Mediated Proteolysis	4	7.78%	3.08		
		Positive Regulation of T Cell Proliferation	3	5.84%	2.35		
		Spermatid Development	3	5.84%	2.14		
		Regulation of Cell Communication By Electrical Coupling	2	3.89%	2.00		
		Regulation of Cardiac Muscle Contraction By Regulation of The Release of Sequestered	2	3.89%	1.51		
	42	Response To Atp	2	3.89%	1.46		
		Calcium Ion Homeostasis	2	3.89%	1.44		
		Negative Regulation of Apoptotic Process	4	7.78%	1.41		
		Glycosaminoglycan Catabolic Process	2	3.80%	1.40		
		Negative Regulation of Insulin Receptor Signaling Pathway	2	3.89%	1.33		
		Transcription, DNA-Templated	27	11.78%	1.86		
		Regulation of Transcription, DNA-Templated	21	9.17%	1.52		
		Positive Regulation of Transcription From RNA Polymerase II Promoter	16	6.98%	1.70		
		Transcription From Rna Polymerase II Promoter	14	6.11%	3.35		
		Cell Adhesion	10	4.36%	1.77		
		Positive Regulation of Transcription, DNA-Templated	10	4.36%	1.49		
		Somatic Stem Cell Population Maintenance	6	3.00%	4.73		
	12	Negative Regulation of Neuron Apoptotic Process	3	1 31%	2.20		
		Odontogenesis	3	1.31%	1.66		
		Sarcomere Organization	3	1.31%	1.60		
		Negative Regulation of Notch Signaling Pathway	3	1.31%	1.60		
		Negative Regulation of Osteoblast Differentiation	3	1.31%	1.36		
		Chondrocyte Differentiation Involved In Endochondral Bone Morphogenesis	2	0.87%	1.60		
		Sarcomerogenesis	2	0.87%	1.60		
		Neuron Recognition	2	0.07%	1.47		
		i touron i toooginaon	-	- OF 70			

Fig. 6. The **CSE58294** dataset was used to validate the STEM analysis results. (A) Eight significant clusters were identified. (B) Gene expression patterns for 495, 485, 366, 161, 55, and 43 proteins were identified for the first 6 clusters, and the expression patterns at different steps are illustrated. (C) Functional enrichment results for genes in clusters 48, 49, 42 and 12 are shown.



Profile	Term	Count	Gene Ratio	Enrichment Score
	Inflammatory Response	25	7.44%	11.92
	Response To Lipopolysaccharide	24	7.14%	11.58
	Innate Immune Response	20	5.95%	8.76
	Aging	16	4 76%	4 73
	Response To Drug	16	4 76%	2 37
	Cellular Response To Lipopolycaccharide	15	4.46%	7.36
	Negative Regulation of Anontotic Process	15	4.46%	207
	Response To Organic Cuello Compound	14	4 170/	4 19
	Response To Organic Cyclic Compound	19	4.17.70	4.10
	Associatio Dessess	13	3.87%	0.73
24	Apoptotic Process	13	3.87%	2.461
	Positive Regulation of Apoptotic Process	12	3.5/%	2.29
	Intracellular Signal Transduction	12	3.5/%	1.92
	Proteolysis	12	3.57%	1.32
	Positive Regulation of Gene Expression		3.21%	1.60
	Cellular Response To Tumor Necrosis Factor	10	2.98%	4.18
	Response To Estradio	10	2.98%	2.86
	Immune Response	10	2.98%	2.05
	Response To Hypoxia	10	2.98%	2.02
	Mapk Signaling Pathway	10	2.98%	1.75
	Positive Regulation of Angiogenesis	9	2.68%	3.81
	Positive Regulation of Transcription From RNA Polymerase II Promoter	21	7.41%	1.13
	Regulation of Transcription, DNA-Templated	17	6.00%	1.22
	Positive Regulation of Transcription, DNA-Templated	15	5.29%	1.28
	Transcription, DNA-Templated	15	5.29%	1.28
	Negative Regulation of Transcription From RNA Polymerase II Promoter	14	4.94%	1.31
	Transcription From RNA Polymerase II Promoter	11	3.88%	1.41
	Hematopoietic Cell Lineage	9	3.18%	1.50
	Cell Differentiation	9	3.18%	1.50
	B Cell Receptor Signaling Pathway	8	2.82%	1.55
	Innate Immune Response	8	2.82%	1.55
-	Cell Migration	7	2.47%	1.61
	Tuberculosis	7	2.47%	1.61
	Immune Response	7	2.47%	1.61
	Response To Hypoxia	7	2.47%	1.61
	Primary Immunodeficiency	6	2.12%	1.67
	Negative Regulation of Cell Migration	6	2.12%	1.67
	Cellular Response To Lipopolysaccharide	6	2 12%	1.67
	Protein Autophosphorylation	6	2 12%	1.67
	T Cell Costimulation	5	1 76%	1.75
	Humoral Immune Response	5	1 76%	1.75
	Metabolic Pathways	18	13.06%	2 21
	Cytokine-Cytokine Receptor Interaction	8	5.80%	3.33
	Cellular Response To Lipopolysaccharide	7	5.08%	3.88
	Axon Guidance	7	5.08%	3.57
	Response To Drug	7	5.08%	1.43
	Cytokine-Mediated Signaling Pathway	6	4.35%	3.33
	Sphingolipid Signaling Pathway	6	4 35%	2.75
	Response To Lipopolysaccharide	6	4.35%	2.00
	Positive Regulation of Gene Expression	6	4 35%	1.52
	Chagas Disease (American Trypanosomiasis)	5	3.63%	2 17
8	Amoebiasis	5	3.63%	2 10
	Toxoplasmosis	5	3.63%	2.06
	Negative Regulation of Gana Expression	5	3.63%	1.76
	Champking Signaling Pathway	5	3.63%	1.44
	Semaphorin Playin Signaling Pathway		2 90%	2.21
	Glutathione Metabolic Process		2 90%	2.73
	Bespense To Bestide Hormone		2.00%	1.65
	Corb Signaling Bathway		2.00%	1.65
	Network Killer Cell Mediated Ordetevicity		2.90%	1.00
	Inflammater: Mediater Regulation of Tro Channels	-	2.00%	1.40
	Initianimatory Mediator Regulation of Trp Channels	9	2.30%	1.32
	Hematenelatis Cell Lineage	7	10 70%	6.24
	Hematopoletic Cell Lineage	-	12.7270	6.34
	measles	7	12.12%	5.07
	Immune Response	/	12.72%	4.13
	Delense Response To Virus	0	10.90%	4.04
	Cellular Response To Cytokine Stimulus	5	9.08%	5.25
	T Only Dependency Olegalian Dethumu		9.06%	4.17
	T Cell Receptor Signaling Pathway	0	9.08%	3.23
	Cell Adhesion Molecules (Cams)	5	9.08%	2.44
3	Cytokine-Cytokine Receptor Interaction	5	9.08%	2.22
-	Inflammatory Response	5	9.08%	2.09
	Positive Regulation of Gene Expression	5	9.08%	1.76
	Negative Regulation of Apoptotic Process	5	9.08%	1.31
	Cellular Response To Interferon-Beta	4	7.27%	3.95
	Primary Immunodeficiency	4	7.27%	3.47
	Positive Regulation of T Cell Proliferation	4	7.27%	3.13
	Allograft Rejection	4	7.27%	2.66
	Autoimmune Thyroid Disease	4	7.27%	2.54
	Systemic Lunus Forthematosus	4	7.27%	1.89
	Cysternic Edgas Erythematosas			

Fig. 7. GSE119121 dataset was used to validate the STEM analysis results. (A) Four significant clusters, namely, cluster 24, cluster 4, cluster 8 and cluster 3, were identified. (B-C) Gene expression patterns and functional enrichment results of genes in 4 clusters are shown.

cytes, some studies report that cytokines can also be produced by resident brain cells when ischemia occurs [27–29]. As the prominent factors, IL-1, TNF- α , IL-6, IL-10 and TGF- β have been the most

С

studied in brain ischemia [30]. Among these star factors, IL-1 is a key mediator in ischemic, excitotoxic and traumatic brain injury; TNF- α might have dual roles in both neuronal injury and protec-



Fig. 8. Weighted Gene Coexpression analysis (WGCNA) identified a MCAO progression-related module. (A) Outlier detection was performed for the 23 samples; (B) A soft threshold was calculated, and an optimal threshold was identified; (C) gene expression modules were constructed; (D) relationships between the gene expression modules and the MCAO time series were determined; (E) A protein–protein interaction network was constructed for the turquoise module (**p* < 0.05).

Table 3

Functional enrichment of pathways and biological processes.

Term	Count	Gene ratio	Enrichment score
Complement and coagulation cascades	14	0.194932	0.710117
Acute-phase response	10	0.139237	0.856245
Negative regulation of peptidase activity	13	0.181008	0.742302
Hemostasis	7	0.097466	1.011147
Response to cytokine	8	0.11139	0.953155
Blood coagulation	8	0.11139	0.953155
Fibrinolysis	5	0.069618	1.157275
Response to peptide hormone	7	0.097466	1.011147
Erythrocyte development	5	0.069618	1.157275
Phospholipid efflux	4	0.055695	1.254185
Complement activation, classical pathway	5	0.069618	1.157275
Innate immune response	9	0.125313	0.902003
Regulation of intestinal cholesterol absorption	3	0.041771	1.379124
Blood coagulation, fibrin clot formation	3	0.041771	1.379124
Cholesterol efflux	4	0.055695	1.254185
Lipoprotein metabolic process	4	0.055695	1.254185
Positive regulation of peptide hormone secretion	3	0.041771	1.379124
High-density lipoprotein particle assembly	3	0.041771	1.379124
Immune system process	8	0.11139	0.953155
Positive regulation of cholesterol esterification	3	0.041771	1.379124
Positive regulation of ERK1 and ERK2 cascade	6	0.083542	1.078094
Positive regulation of vasoconstriction	4	0.055695	1.254185
Plasminogen activation	3	0.041771	1.379124

tion; IL-10 and TGF- β may play a role in neuroprotection [31]. However, the role of IL-6 in stroke [32–35] is conflicting. In our study, after constructing the MCAO mouse model of ischemic stroke, a whole proteome wide profiling of the ischemic cerebral cortex area at 0, 6, 12 and 24 h was conducted. After analyzing the DEPs in 4 groups, a total of 41 genes were identified, implicat-



Fig. 9. Gene set enrichment analysis (GSEA) of hallmark genes in the protein profiling. (A) GSEA was performed to explore the basic function of the overall genes in our proteomic datasets, and (B) the first five items with higher enrichment scores (ES), coagulation, complement, oxygen species pathway, interferon gamma response and apoptosis, are illustrated. (C) Go terms, (D) miRNAs and (E) transcription factors of GSEA when comparing 6 h vs 0 h, 12 h vs 6 h and 24 h vs 12 h, and the enrichment scores are shown by heatmap.

ing their potential role in stroke pathogenesis. Though proteins were usually translated by mRNAs, where the significant differential expression exists in mRNA level was a puzzle. To verify our finding, two independent gene expression datasets were applied. As shown in Fig. 3, ATP1A4, C3 and S100A9 also showed significant differential expression in both datasets, in mRNA level. With different analysis method, C3 was also enriched in the hub network of stroke (Fig. 8). Complement C3 has been reported as stroke biomarkers in several studies [36-38]. Additionally, C3 was involved in cerebral ischemia/reperfusion Injury [39], neuroinflammation and brain damage [40,41] and microglia reaction [42] in stroke. Similarly, studies pointed that S100A9 was involved in excess inflammation after ischemic stroke [43], and vaccine against S100A9 inhibits thrombosis without increasing the risk of bleeding [44], supporting its potential value of stroke target. As a Na/K-ATPase, ATP1A4 was well-documented in sperm function [45,46]. while there role in stroke was firstly identified in this study, and its role in stroke will be studied in the future. In sum, we identified three genes with both differential expression change in mRNA and protein level, and the first identified ATP1A4 may indicate certain function or biomarker candidates' value in stroke.

To identify genes clusters with certain expression pattern in stoke progression, a Short Time-series Expression Miner (STEM) was applied to identify genes in different time points of MCAO. The STEM analysis was conducted first in our dataset, and then verified in another three GEO datasets. Though the number of identified clusters was different, the function of cluster showed certain similarity. In our dataset, cluster 13 was the most significantly enriched item, with the 5 core function of hydrolase activity, lipase activity, acute inflammatory response, complement activation and neuron projection regeneration. The functions can be generally categorized into 3 classes of fibrolysis (hydrolase activity, lipase activity), inflammation and immune response, and neurocognition. Fibrolysis has been well known for its role in stroke progression [47], and tissue plasminogen activator (tPA) has been applied for stroke treatment [48,49]. With stroke induced cognitive impairment becoming a tricky clinical problem, more attention are been paid on neurointervention [50]. Studies pointed that neuroinflammation play a pivotal role in stroke, and microglia was reported in this process [51,52]. In the validation dataset, inflammation and immune response was the most observed functional item, while fibrolysis (proteolysis) and neurocognition (axon guidance, neuron recognition) observed in GSE58294 (Fig. 6) and GSE119121 (Fig. 7). It's a pity that we didn't observe a significant dominating or subordinate role of these three function items at different time points. The possible reason, we think, was the complex interaction of genes and cell types.

At last, we conducted GSEA analysis to enrich gene sets, instead of individual genes to look into the functional items and hub regulators in stroke. In our proteomic study, genes related to ROS, which were upregulated in the MCAO module, were also identified in GSEA with high scores. These genes are worthy of further research. ROS generation is an intrinsic event that occurs during the onset of the inflammatory cascade via several enzyme systems [53,54]. Until now, ROS have been identified as an important underlying factor involved in delayed neuronal death. During reperfusion, robust oxidants are produced and directly participate in the damage to cellular macromolecules, such as lipids, proteins, and nucleic acids, and eventually lead to cell death [55]. In addition, we identified a set of microRNAs and transcription factors in different time points. Transcription factors and microRNAs have been widely reported in stroke, in transcriptional [56,57] and posttranscriptional level [58,59], respectively. Furthermore, miRNAs and transcription factor were reported to regulate each other in ischemic stroke [60]. The expression pattern change indicates that, with the progression of stroke, targets of microRNAs and transcription factors undergo certain regulation to exert their functions. For example, at 6 h, targets of transcription factors STATs (Signal Transducers and Activators of Transcription) increased, and at 12 h, decreased dramatically. The role of STATs have been reported in neuronal injury and protection [61,62], and neuroinflammation [63,64] in stroke. Interestingly, targets of miR-296 was elevated at 6 h, downregulated at 12 h, and at last, increased at 24 h. MiR-296 has been reported to regulated growth factor receptor overexpression in angiogenic endothelial cells [65], but its role in stroke remains obscure.

Taken together, we performed an integrated analysis of the proteome and transcriptome, revealing the landscape during stroke progression in an MCAO mouse model, suggesting potential genes and hub regulators during stroke progression.

Conclusions

In conclusion, the results of the present study reveal that the inflammatory response, cytokine-cytokine receptor interaction, innate immune response and ROS are very important pathways that are active following a stroke. C3, Apoa4 and S100a9 may be biomarkers or drug targets for stroke.

Compliance with ethics requirments

All the animal experiments were approved by the Ethics Committee of Hebei General Hospital (approval no. 201921).

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 material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jare.2020.01.005.

References

- Volny O et al. MicroRNAs in cerebrovascular disease. Adv Exp Med Biol 2015;888:155–95.
- [2] Chamorro A et al. The immunology of acute stroke. Nat Rev Neurol 2012;8 (7):401-10.
- [3] Iadecola C, Anrather J. The immunology of stroke: from mechanisms to translation. Nat Med 2011;17(7):796–808.

- [4] Muir KW et al. Inflammation and ischaemic stroke. Curr Opin Neurol 2007;20 (3):334–42.
- [5] Westendorp WF et al. Post-stroke infection: a systematic review and metaanalysis. BMC Neurol 2011;11:110.
- [6] Vemuganti R. All's well that transcribes well: non-coding RNAs and poststroke brain damage. Neurochem Int 2013;63(5):438–49.
- [7] Zhang J et al. Altered long non-coding RNA transcriptomic profiles in brain microvascular endothelium after cerebral ischemia. Exp Neurol 2016;277:162–70.
- [8] He W et al. Altered long non-coding RNA transcriptomic profiles in ischemic stroke. Hum Gene Ther 2018;29(6):719–32.
- [9] Asano S, Chantler PD, Barr TL. Gene expression profiling in stroke: relevance of blood-brain interaction. Curr Opin Pharmacol 2016;26:80–6.
- [10] Huang JT et al. Disease biomarkers in cerebrospinal fluid of patients with firstonset psychosis. PLoS Med 2006;3(11):e428.
- [11] Sironi L et al. Acute-phase proteins before cerebral ischemia in stroke-prone rats: identification by proteomics. Stroke 2001;32(3):753–60.
- [12] Wen M et al. Proteomic analysis of rat cerebral cortex in the subacute to longterm phases of focal cerebral ischemia-reperfusion injury. J Proteome Res 2019;18(8):3099–118.
- [13] Ning M et al. Proteomic protease substrate profiling of tPA treatment in acute ischemic stroke patients: a step toward individualizing thrombolytic therapy at the bedside. Transl Stroke Res 2010;1(4):268–75.
- [14] Ning M et al. Proteomic temporal profile of human brain endothelium after oxidative stress. Stroke 2011;42(1):37–43.
- [15] Murugesan N et al. Analysis of mouse brain microvascular endothelium using laser capture microdissection coupled with proteomics. Methods Mol Biol 2011;686:297–311.
- [16] Lu Q et al. Analysis of mouse brain microvascular endothelium using immunolaser capture microdissection coupled to a hybrid linear ion trap with Fourier transform-mass spectrometry proteomics platform. Electrophoresis 2008;29 (12):2689–95.
- [17] Longa EZ et al. Reversible middle cerebral artery occlusion without craniectomy in rats. Stroke 1989;20(1):84–91.
- [18] Iancu OD et al. Cosplicing network analysis of mammalian brain RNA-Seq data utilizing WGCNA and Mantel correlations. Front Genet 2015;6.
- [19] Holtman IR et al. Induction of a common microglia gene expression signature by aging and neurodegenerative conditions: a co-expression meta-analysis. Acta Neuropathologica Commun. 2015:3.
- [20] Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. BMC Bioinf 2008:9.
- [21] Subramanian A et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. PNAS 2005;102 (43):15545–50.
- [22] Benjamini Y, Hochberg Y. Controlling the false discovery rate a practical and powerful approach to multiple testing. J. Royal Statist. Soc. Ser. B-Statist. Methodol. 1995;57(1):289–300.
- [23] Ernst J, Bar-Joseph Z. STEM: a tool for the analysis of short time series gene expression data. BMC Bioinf 2006:7.
- [24] Stanimirovic D et al. The induction of ICAM-1 in human cerebromicrovascular endothelial cells (HCEC) by ischemia-like conditions promotes enhanced neutrophil/HCEC adhesion. J Neuroimmunol 1997;76(1–2):193–205.
- [25] Stanimirovic D, Satoh K. Inflammatory mediators of cerebral endothelium: a role in ischemic brain inflammation. Brain Pathol 2000;10(1):113–26.
- [26] Wang X et al. Upregulation of intercellular adhesion molecule 1 (ICAM-1) on brain microvascular endothelial cells in rat ischemic cortex. Brain Res Mol Brain Res 1994;26(1–2):61–8.
- [27] Hurn PD et al. T- and B-cell-deficient mice with experimental stroke have reduced lesion size and inflammation. J Cereb Blood Flow Metab 2007;27 (11):1798–805.
- [28] Liu T et al. Tumor necrosis factor-alpha expression in ischemic neurons. Stroke 1994;25(7):1481–8.
- [29] Sairanen T et al. Evolution of cerebral tumor necrosis factor-alpha production during human ischemic stroke. Stroke 2001;32(8):1750–8.
- [30] Han HS, Yenari MA. Cellular targets of brain inflammation in stroke. Curr Opin Invest Drugs 2003;4(5):522–9.
- [31] Allan SM, Rothwell NJ. Cytokines and acute neurodegeneration. Nat Rev Neurosci 2001;2(10):734–44.
- [32] Smith, C.W., Leukocyte-endothelial cell interactions. Semin Hematol, 1993. 30 (4 Suppl 4): p. 45-53; discussion 54-5.
- [33] Yamashita T et al. Blockade of interleukin-6 signaling aggravates ischemic cerebral damage in mice: possible involvement of Stat3 activation in the protection of neurons. J Neurochem 2005;94(2):459–68.
- [34] Herrmann O et al. Regulation of body temperature and neuroprotection by endogenous interleukin-6 in cerebral ischemia. J Cereb Blood Flow Metab 2003;23(4):406–15.
- [35] Clark WM et al. Lack of interleukin-6 expression is not protective against focal central nervous system ischemia. Stroke 2000;31(7):1715–20.
- [36] Palikhe A et al. Serum complement C3/C4 ratio, a novel marker for recurrent cardiovascular events. Am J Cardiol 2007;99(7):890–5.
- [37] Engstrom G et al. Complement C3 and C4 in plasma and incidence of myocardial infarction and stroke: a population-based cohort study. Eur J Cardiovasc Prev Rehabil 2007;14(3):392–7.
- [38] Stokowska A et al. Plasma C3 and C3a levels in cryptogenic and large-vessel disease stroke: associations with outcome. Cerebrovasc Dis 2011;32 (2):114–22.

- [39] Lin Z et al. Complement component C3 Promotes cerebral ischemia/ reperfusion injury mediated by TLR2/NFkappaB activation in diabetic mice. Neurochem Res 2018;43(8):1599–607.
- [40] Zhang X et al. Salidroside reduces inflammation and brain injury after permanent middle cerebral artery occlusion in rats by regulating PI3K/PKB/ Nrf2/NFkappaB signaling rather than complement C3 activity. Inflammation 2019;42(5):1830–42.
- [41] Mocco J et al. Complement component C3 mediates inflammatory injury following focal cerebral ischemia. Circ Res 2006;99(2):209–17.
- [42] Surugiu R et al. Intracortical administration of the complement C3 receptor antagonist trifluoroacetate modulates microglia reaction after brain injury. Neural Plast 2019;2019:1071036.
- [43] Shichita T et al. MAFB prevents excess inflammation after ischemic stroke by accelerating clearance of damage signals through MSR1. Nat Med 2017;23 (6):723–32.
- [44] Kawano T et al. Therapeutic vaccine against \$100A9 (\$100 calcium-binding protein A9) inhibits thrombosis without increasing the risk of bleeding in ischemic stroke in mice. Hypertension 2018;72(6):1355–64.
- [45] Rajamanickam GD, Kastelic JP, Thundathil JC. Content of testis-specific isoform of Na/K-ATPase (ATP1A4) is increased during bovine sperm capacitation through translation in mitochondrial ribosomes. Cell Tissue Res 2017;368 (1):187–200.
- [46] Rajamanickam GD et al. Testis-specific isoform of Na/K-ATPase (ATP1A4) regulates sperm function and fertility in dairy bulls through potential mechanisms involving reactive oxygen species, calcium and actin polymerization. Andrology 2017;5(4):814–23.
- [47] Thiebaut AM et al. The role of plasminogen activators in stroke treatment: fibrinolysis and beyond. Lancet Neurol 2018;17(12):1121–32.
- [48] Lekoubou A, Awoumou JJ, Kengne AP. Incidence of seizure in stroke patients treated with recombinant tissue plasminogen activator: a systematic review and meta-analysis. Int J Stroke 2017;12(9):923–31.
- [49] Kim JT et al. Treatment with tissue plasminogen activator in the golden hour and the shape of the 4.5-hour time-benefit curve in the national united states get with the guidelines-stroke population. Circulation 2017;135(2):128–39.
- [50] Jacquin GJ, van Adel BA. Treatment of acute ischemic stroke: from fibrinolysis to neurointervention. J Thromb Haemost 2015;13(Suppl 1):S290–6.

- [51] Wang J et al. Treatment targets for M2 microglia polarization in ischemic stroke. Biomed Pharmacother 2018;105:518–25.
- [52] Kanazawa M et al. Microglia and monocytes/macrophages polarization reveal novel therapeutic mechanism against stroke. Int J Mol Sci 2017;18(10).
- [53] Nizyaeva N et al. Expression analysis of tollip, the inhibitor of inflammatory signaling cascade, in placenta samples in cases of early onset preeclampsia. Virchows Arch 2017;471:S96–7.
- [54] Chan PH. Reactive oxygen radicals in signaling and damage in the ischemic brain. J Cereb Blood Flow Metab 2001;21(1):2–14.
- [55] Chan PH. Mitochondria and neuronal death/survival signaling pathways in cerebral ischemia. Neurochem Res 2004;29(11):1943–9.
- [56] Jin X et al. Association of sterol regulatory element-binding transcription factor gene polymorphisms with ischaemic stroke. J Int Med Res 2012;40 (1):157–66.
- [57] Tang X et al. Genetic deletion of kruppel-like factor 11 aggravates ischemic brain injury. Mol Neurobiol 2018;55(4):2911–21.
- [58] Yuan Y et al. MiRNA-27b regulates angiogenesis by targeting AMPK in mouse ischemic stroke model. Neuroscience 2019;398:12–22.
- [59] Wang Z et al. Inhibition of miRNA-27b enhances neurogenesis via AMPK activation in a mouse ischemic stroke model. FEBS Open Bio 2019;9 (5):859–69.
- [60] Nampoothiri SS, Fayaz SM, Rajanikant GK. A novel five-node feed-forward loop unravels miRNA-gene-tf regulatory relationships in ischemic stroke. Mol Neurobiol 2018;55(11):8251–62.
- [61] Jung JE et al. STAT-dependent upregulation of 12/15-lipoxygenase contributes to neuronal injury after stroke. J Cereb Blood Flow Metab 2015;35(12):2043–51.
- [62] Zhou H et al. Activation of STAT3 is involved in neuroprotection by electroacupuncture pretreatment via cannabinoid CB1 receptors in rats. Brain Res 2013;1529:154–64.
- [63] Dinapoli VA et al. Age exaggerates proinflammatory cytokine signaling and truncates signal transducers and activators of transcription 3 signaling following ischemic stroke in the rat. Neuroscience 2010;170(2):633–44.
- [64] Chen S et al. Homocysteine exaggerates microglia activation and neuroinflammation through microglia localized STAT3 overactivation following ischemic stroke. J Neuroinflammation 2017;14(1):187.
- [65] Wurdinger T et al. miR-296 regulates growth factor receptor overexpression in angiogenic endothelial cells. Cancer Cell 2008;14(5):382–93.