Identification of immune-infiltrated hub genes as potential biomarkers of Moyamoya disease by bioinformatics analysis

Fa Jin

Zhujiang Hospital <https://orcid.org/0000-0002-2470-2942>

Chuanzhi Duan (✉ doctor_duanzj@163.com)

Zhujiang Hospital <https://orcid.org/0000-0002-2025-8637>

Research

Keywords: Moyamoya disease, Immune infiltration, Neutrophils, UNC13D, Bioinformatics analysis

Posted Date: March 29th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-324031/v1>

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

Version of Record: A version of this preprint was published at Orphanet Journal of Rare Diseases on February 23rd, 2022. See the published version at <https://doi.org/10.1186/s13023-022-02238-4>.

Abstract

Background Moyamoya disease (MMD) is a unique chronic progressive cerebrovascular disease. The molecular mechanism behind pathophysiology is still elusive. This study aims to determine the key genes and their roles in the immune infiltration of MMD.

Methods We download raw gene expression profiles (GSE157628, GSE141024) of cerebrovascular tissue from GEO database. Identify differentially expressed genes (DEGs) and perform functional enrichment analysis. The CIBERSORT deconvolution algorithm was used to analyze the proportion of immune cell infiltration between MMD and negative control group. We screened for neutrophil-associated DEGs, constructed a protein-protein interaction network (PPI) using STRING, and clarified hub genes using the Cytoscape plugin MCODE analysis. The receiver operating characteristic (ROC) curve is applied to test and filter the best gene signature.

Results A total of 570 DEGs were detected, including 212 downregulated and 358 up-regulated genes. Reactome and KEGG enrichment revealed that DEGs are involved in the cell cycle, molecular transport, and metabolic pathways. The immune infiltration profile demonstrates that MMD cerebrovascular tissues contained a higher proportion of neutrophils, monocytes, and NK cells than negative control group. PPI network and MCODE cluster identified 9 DEGs (*UNC13D*, *AZU1*, *PYCARD*, *ELANE*, *SDCBP*, *CCL11*, *CCL15*, *CCL20*, and *CXCL5*) associated with neutrophil infiltration. ROC results showed that *UNC13D* has good specificity and sensitivity (AUC = 0.7846).

Conclusions The characteristics of immune infiltration in the cerebrovascular tissues of MMD patients and abnormal expression of hub genes provide new insights for understanding MMD progression. *UNC13D* is promising to be one of the candidate molecules to determine neutrophil infiltration characteristics in MMD.

1 Background

Although moyamoya disease (MMD) is a rare disease, it is one of the leading causes of stroke in children and young adults[1]. Moyamoya disease's primary pathological condition is the chronic progressive stenosis of the terminal part of the internal carotid artery, which develops into an abnormal vascular network (moyamoya vessels) of collateral circulation at the base of the brain[2]. MMD has the highest incidence in Asian countries ($\leq 0.94/100000$), but the incidence in non-Asian countries increases, and there are ethnic differences[3, 4]. The onset age of the disease is bimodal, with the first peak in children under ten years old, mainly manifested as a focal cerebral lesion, seizures, and intellectual impairment; the second peak in adults aged 30 ~ 40 years old, mainly manifested as cerebral hemorrhage, hemiplegia, aphasia or other neurological damage; the elderly over 50 years of age often very serious, even life-threatening[5].

The pathogenesis of moyamoya disease remains unresolved, and current studies have shown that its caused by two major factors: genetic inheritance and acquired autoimmune changes[6]. Genetic factors,

a mutation in the *RNF213* (Ring finger 213) gene was identified as the most critical susceptibility gene for MMD using a genome-wide linkage and exome analysis[7–9]. In the Japanese population, 95% of familial MMD patients, 80% of sporadic MMD patients and 1.8% of the control population have found the R4810K (c.14576G > A) variant of the *RNF213* gene. However, the R4810K mutation of *RNF213* is not the MMD specific biomarker, and it is also found in patients with intracranial atherosclerosis[10]. Besides, several microRNAs (miRNAs) have been identified to participate in the pathophysiological development of MMD indirectly; for example, the single nucleotide polymorphism (SNP) of *miR-196a* is closely related to MMD, *miR-6722-3p*, and *miR-328-3p* is significantly down-regulated in MMD[11, 12]. On the other side, MMD has been linked with factors beyond genetic backgrounds, such as autoimmune inflammation, vessel wall damage, angiogenesis, and thrombogenic factors[13]. Some reports indicate that MMD is associated with autoimmune diseases, including autoimmune thyroid disease (Graves' disease), type 1 diabetes, and systemic lupus erythematosus (SLE)[14–16]. Suzuki summarized the literature and found that A large amount of IgG, IgM, and other immunoglobulins are deposited in the intimal thickening layer of MMD vessels[17]. Nevertheless, there are currently few studies on the molecular mechanisms in terms of immune infiltration in MMD.

Bioinformatics is an emerging cross-discipline combining molecular biology and information technology, and it plays an essential role in revealing the molecular mechanisms of disease[18]. On account of MMD is closely related to autoimmune diseases, and the study of its immune infiltration helps to understand the underlying pathophysiological mechanism. This paper used bioinformatics methods to analyze the MMD and control's cerebrovascular tissue microarray data. Aim to identify MMD immune infiltration characteristics and specific differential expressed genes, search for promising biomarkers or therapeutic targets, and provide new insights into the pathogenesis of MMD.

2 Results

Data preprocessing and DEG analysis

After the expression matrix normalized (Fig. 2a), we determined the grouping of sample data by principal component analysis (PCA), there was no overlap in data distribution between the MMD and EPI groups, and discrimination was evident; However, the IA group overlapped part of the data with both the EPI and MMD groups, and the IA and EPI groups were clearly divided into two groups, and thus were not suitable to be used together as negative controls (Fig. 2b). In order to find the differential genes more precisely, we used only EPI group as the negative control in the DEG analysis and excluded the IA group. 570 DEGs were screening out between case group and NC group, among which 358 genes were up-regulated and 212 genes were down-regulated. The screening criteria for DEGs were as follows: p -value < 0.01 and $|\log_2$ fold change| > 1. Based on the analysis of the samples' gene expression, the volcano plot was made, as shown in Fig. 2c. The top 200 genes with the most significant differences are shown in a heat map (Fig. 2d).

Functional and Pathway Enrichment

According to the Reactome gene sets enrichment analysis, there were ten down-regulated terms and 11 up-regulated terms, mainly involving the biological processes of small molecule metabolism and cell cycle. The results are shown in Fig. 3a, c. Interestingly, both Transport of small molecules and SLC-mediated transmembrane transport terms appeared in the up and down-regulated genes. The enrichment results of down-regulated DEGs in KEGG pathways were Pyrimidine metabolism, Purine metabolism, and Ferroptosis, et al. (Fig. 3b). The enriched KEGG pathways of up-regulated genes were Thyroid hormone synthesis, Neuroactive ligand-receptor interaction, Cytokine-cytokine receptor interaction, Alpha-linolenic acid metabolism et al. (Fig. 3d).

Immune infiltration analyses

Through the CIBERSORT algorithm, we obtained the differences in immune infiltration of 14 immune cell subgroups in MMD group and NC group. Compared with the NC group, the MMD group generally contained a higher proportion of Neutrophils, Monocytes, Dendritic cells, Gamma delta T cells, Follicular helper T cells, whereas the proportions of CD4 + T cells, CD8 + T cells, and eosinophils were relatively lower (Fig. 4a, b).

GO network analysis

By checking the database for potential sets of genes known to regulate neutrophils, we used the intersection to obtain DEGs associated with neutrophil infiltration (Fig. 5a). The expression characteristics of these Neutrophil-associated DEGs were visualized as a heat map. (Fig. 5b). Then, we further examined the biological processes that these neutrophil-associated DEGs are involved in regulating neutrophils in MMD and used the ClueGO plugin of Cytoscape software to analyze the network relationship between these biological processes (Fig. 5c). GO analysis showed that these neutrophils-associated DEGs mainly involved RNA polymerase II transcription preinitiation complex assembly, Regulation of neutrophil migration, Regulation of lymphocyte migration, and NK cell degranulation. The detailed results are shown in **Table 1**.

Identification of hub genes using PPI network and modular screening

In order to clarify which are hub genes, we use the STRING database to perform PPI network analysis on 30 neutrophils-associated genes, and the resulting network diagram which included 37 edges and 23 nodes (Fig. 5d). Moreover, this network was further assessed with the MCODE plugin of Cytoscape, Hub genes from the top 2 modules, including *UNC13D*, *AZU1*, *PYCARD*, *ELANE*, *SDCBP*, *CCL11*, *CCL15*, *CCL20*, and *CXCL5* (Fig. 5e). Subsequently, correlation analysis showed that *UNC13D*, *PYCARD*, *SDCBP* and *CCL20* were positively correlated with neutrophils ($r = 0.802$, $r = 0.690$, $r = 0.263$, $r = 0.107$); *AZU1*, *ELANE*, *CCL11*, and *CCL15* were negatively correlated with neutrophils ($r = -0.210$, $r = -0.385$, $r = -0.223$, $r = -0.146$) (Fig. 5f). The correlation results indicated that *UNC13D*, *PYCARD* have a strong positive correlation with neutrophil infiltration. Finally, the relative expressions were calculated to examine the differences between these nine Neutrophil-associated DEGs at the transcriptional level (Fig. 6a-i). The results showed that only *UNC13D*, *AZU1*, *PYCARD*, and *CCL15* were statistically significant in the EPI group versus the MMD group

($p < 0.05$). Interestingly, the IA group's expression levels were between the EPI and MMD groups, and the expression of all genes was not statistically significant compared to the MMD group ($p > 0.05$). In addition, *AZU1* and *PYCARD* were statistically significant both in the EPI group vs IA group and EPI vs MMD group ($p < 0.05$).

Diagnostic effectiveness evaluation

To assess the potential of four neutrophil-associated DEGs as biomarkers for MMD, we performed ROC analysis on these genes, and the four genes with results were used for visualization as ROC curves for comparison. ROC curves confirmed that they could distinguish MMD from non-MMD (all AUC > 0.5) and that the AUC was between 0.5846 and 0.7846 (Fig. 7a-b), while the gene with the highest predict power was *UNC13D* with an AUC of 0.7846. Detailed results data are shown in (Fig. 7c).

3 Discussion

Our study is the first to reveal the characteristics of immune infiltration in the cerebrovascular of MMD based on microarray data. Most studies used samples of body fluids (cerebrospinal fluid or peripheral blood) from MMD patients in recent relevant genetic studies. They identified some specific molecules such as autoimmune antibodies[19], Transfer RNA-derived fragments[20], Circular RNAs[21], Exosome-Derived miRNAs[22], etc. We compared the composition of immune infiltrating cells in the samples in order to narrow down and discover potential regulatory genes in MMD immune infiltration. Through further excavation, several hub genes involved in the regulation of neutrophil infiltration were identified. Thus, we conducted diagnostic ability tests on these newly discovered related genes to screen out the novel biomarker.

In the results of our enrichment analysis, Reactome gene sets analysis suggested that small-molecule transport and transmembrane transport were enriched in down-regulated differential genes, as well as cell cycle changes. Interestingly, among the KEGG enriched pathways, we found up-regulated activation of pathways associated with thyroid hormone synthesis, which in line with previous findings. Nakamura et al. emphasized that excess thyroid hormone is thought to alter cerebral hemodynamics, increase metabolism and oxygen consumption in the brain and be harmful to the arterial wall[23]. Moreover, thyrotoxicosis may lead to hyperhomocysteinemia or sympathetic nervous activity, which is associated with premature atherosclerosis and MMD[23, 24].

MMD is closely related to chronic inflammatory cell infiltration. Masuda et al. reported that the primary component of intimal hyperplasia in the main internal carotid artery vessels of MMD patients is smooth muscle cells (SMCs), and immune cells infiltration can promote the proliferation of SMCs in patients, leading to MMD vascular lesions[25], so immune-mediated pathological changes may be involved in the pathogenesis of MMD. However, due to limitations of technique or sample scarcity, the landscape of immune infiltration in MMD has not been revealed, particularly in subpopulations with a low abundance of cells. Using the CIBERSORT deconvolution algorithm, we first investigated the immune infiltration difference between MMD and non-MMD cerebrovascular tissues in 14 subpopulations of immune cells.

Among the several immune cell compositions up-regulated, the increased neutrophil proportion was most significant in the up-regulated immune cells, which is noteworthy because neutrophils play an important role both in promoting vaso-occlusion and inducing angiogenesis[26, 27]. Neutrophils are innate immune phagocytes with a central role in immune defence. Nonetheless, neutrophils also damage host tissues; their deployment is through tightly regulated strategies: phagocytosis, degranulation, and release of neutrophil extracellular traps (NETs)[28]. The stimulation of neutrophils forms NETs by several proteins; the serine protease neutrophil elastase (ELANE) is one of its activators[29]. NET release results in a cell death process called NETosis[30]. NET and participate in the pathological process of vascular occlusion, its mechanism of action is: neutrophils accumulate in the blood vessels in a P-selectin-dependent manner, followed by platelet recruitment; neutrophils promote platelet production of thromboxane A2, induce endothelial cells to express intercellular adhesion molecule 1 (ICAM1)[31], and strengthen the interaction between neutrophils and endothelium. The process triggers NETosis in vascular endothelial cells through a mechanism involving platelet-derived high mobility group box 1 (HMGB1), ROS, and integrin 1[32, 33]. On the other hand, owing to inflammatory signals initiate neutrophil recruitment, neutrophils begin to secrete growth factors (e.g., VEGF-A, prokineticin2), chemokines, and MMP-9 to recruit other leukocytes to the damaged areas[34]. Corey et al. study suggest that varying neutrophil phenotypes facilitate an inherent autoimmune state within MMD patients, and crucially promote angiogenesis[35]. Therefore, combined with microarray analysis results, we speculated that neutrophils might play a potential role in vascular occlusion and pathological changes of new capillaries in MMD.

The highlight of this study is to explore the differences in MMD gene expression and immune cell infiltration. We have identified nine hub neutrophils-associated genes in total, of which *UNC13D* is a gene with the highest diagnostic accuracy. *UNC13D* (Protein unc-13 homolog D), also known as munc13-4 gene, encoded and expressed Munc13-4 protein is one of the components of the neutrophil secretion mechanism. It regulates neutrophil exocytosis (degranulation) in a calcium-dependent and SNARE dependent manner. In general, the degranulation of neutrophils is tightly regulated because excessive particle content can cause tissue damage. In the present study, *UNC13D* was significantly up-regulated in MMD microarrays, implying that neutrophil-mediated immune damage occurs. We speculate it may be related to the increased degranulation of *UNC13D* regulated secretion, and the specific mechanism is necessary to be further experimentally verified. Additionally, the remaining eight genes are involved in regulating neutrophils, as shown in **Table 2**.

There are several limitations to our present study. The number of samples we obtained from GSE157628 and GSE141024 was small, generating some bias when analyzing the DEGs and calculating the AUC, and more data samples are needed for validation in further research. Although this study confirms that *UNC13D* has good diagnostic properties and can be used as a potential biomarker for MMD, genetic diagnosis by obtaining vascular tissue is not feasible. Cerebral angiography, CTA or MRA are more commonly used clinically for diagnosis. Due to the complex functions and molecular mechanisms of genes, these bioinformatics results need to be verified by experiments. The slow progression of MMD may have different gene expression at each time stage, as well as the genetic characteristics of children and adults may be distinct, which need to be further clarified.

4 Conclusions

In conclusion, we revealed the composition of immune cell infiltration in the cerebrovascular tissues of patients with MMD for the first time by bioinformatics analysis and screened nine hub genes related to regulating neutrophils, of which UNC13D could be used as a promising candidate biomarker for MMD, and this study provides new insights into the pathogenesis for MMD.

5 Methods

Data Collection and Preprocessing

The flowchart of this study is represented in Fig. 1. We searched for the keyword "Moyamoya disease" in the Gene Expression Omnibus (GEO) database to find and select datasets that met the requirements. Only data containing vascular tissue from MMD were collected, excluding small sample sizes, peripheral blood, and cerebrospinal fluid data. After removing duplicate subsets, the raw chip data of GSE157628 and GSE141024 were downloaded for our analysis, which were generated from the same microarray chip; the chip model is Agilent SurePrint G3 Human GE v2 8x60K. The GSE157628 contained middle cerebral artery vascular (MCA) wall tissue from 11 patients with MMD, six patients with internal carotid aneurysm (IA), and three patients with epilepsy (EPI); the GSE141024 contained superficial temporal artery (STA) vascular tissue from 4 patients with MMD and four patients with the internal carotid aneurysm. Detailed sample information is shown in **Supplementary Table 1**. All samples were background corrected and quantile normalized using the "limma" package[36] before comprehensive analysis. We used the "data table" package to clean and extract the data. For probes with duplicate gene symbols, we used the mean as its unique expression value.

Principal component analysis (PCA) and DEGs screening

In order to select appropriate subgroups for DEGs analysis, we first visualized the distribution of the samples using PCA to assess the overall data patterns. Due to MMD is a chronic vascular lesion with subtle variants in cellular composition from normal arteries, and vascular tissue from EPI patients was more representative of normal vascular tissues than IA patients. Therefore, we selected MCA samples from 11 MMD patients in the GSE157628 dataset as a case group and MCA samples from 3 epileptic patients as a negative control group. The "limma" package was performed to identify DEGs. Both log fold change (log FC) absolute value > 1 and P-value < 0.01 are regarded as threshold values.

Functional enrichment analyses of the DEGs

For functional enrichment analysis, Reactome and KEGG[37] pathway enrichment analyses of the DEGs were performed using Metascape[38] platform (<http://metascape.org/gp>). Upload the gene symbol lists of up-regulated and down-regulated differential genes to the server, respectively, and wait until the analysis is finished to download the returned results. All terms were selected with enrichment significance

evaluated at $p < 0.05$, count ≥ 5 , and enrichment factor > 1.5 based on their similarities automatically by the platform. The results were imported into R and plotted as dot plot.

CIBERSORT analysis

To quantify the proportion of immune cells in samples, we performed computational deconvolution using CIBERSORT[39] webpage tool. A normalized gene expression matrix was used as input uploaded to the CIBERSORT server(<https://cibersort.stanford.edu/>). Both absolute and relative modes were run, and quantile normalization was disabled. One thousand permutations were run for statistical testing. The result returned a percentage rate of immune cell for all samples, and the sum of each sample's immune cell ratio is 1.

Extraction of mRNA expression data involved in the regulation of neutrophils

In order to clarify which DEGs are involved in regulating neutrophils, we first obtained gene symbol lists associated with the regulation of neutrophils from the AmiGO 2 website (<http://amigo.geneontology.org/>) by searching the keyword "neutrophils". Remove duplicate gene symbols, and then these neutrophil-associated gene symbols were intersected with DEGs. Finally, we extracted an expression matrix based on intersection gene symbols from the normalized matrix.

GO enrichment analysis

We used Cytoscape software for network visualization (<http://cytoscape.org/>). Apply the ClueGO plugin to analyze the neutrophils-associated genes through the gene ontology biological process (BP) analysis, and establish a network relationship diagram.

Protein-protein interaction (PPI) network construction and hub gene identification

To gain insight into the genetic interactions related to the identified neutrophil infiltration, we used the STRING (<https://string-db.org/>) web-based tool[40] to analyze and construct a PPI network to reveal the molecular mechanism of neutrophil-associated genes involvement in MMD. The neutrophil-associated genes in the PPI network act as nodes and the line between two nodes represents the associated interaction. The more core genes, the more connections there are. We visualized the PPI network result using Cytoscape software and used the Molecular Complex Detection (MCODE)[41] plugin to screen highly interconnected gene clusters PPI network.

Correlation Analysis of Hub Genes and Infiltrating Immune Cells

After obtaining hub genes, we performed Spearman correlation analysis on hub genes and infiltrating immune cells by the "corrplot" package. So as to understand the correlation between the expression levels of these genes and the proportion of immune infiltrating cells.

Relative expression analysis of neutrophil-associated genes

To reduce the bias caused by normalization on the expression values of the data and the false positives of using P-value as DEGs cutoff. We extracted the expression values of these hub neutrophil-associated genes and the internal reference gene (GAPDH) directly from the raw data. By calculating the relative expressions to clarify whether they are statistically different. All MCA samples from GSE157628 were used, including 11 cases of MMD, 6 cases of IA, and 3 cases of EPI.

ROC analysis

To further test hub neutrophil-associated genes' ability to distinguish MMD, we divided the 28 samples from GSE157628 and GSE141024 into two groups (MMD, non-MMD) according to their diagnosis information. We applied the receiver operating characteristic (ROC) curve for testing and used the area under the curve (AUC) to evaluate their accuracy. Use the R package "pROC" [42] to analyze results and visualize the data. The perfect AUC value is 1. When AUC greater than 0.5 is considered to have predictive value. It is closer to 1, the better specificity and sensitivity.

Statistical analysis

All statistical analysis is performed by R 4.0.2. $p < 0.05$ was considered statistically significant. Box plots, volcano plot, principal component analysis plot, bubble plots, and the violin plot are drawn by the R package "ggplot2" [43]. Heat maps are generated using the R package "pheatmap" [44].

List Of Abbreviations

MMD: Moyamoya disease; GEO: Gene Expression Omnibus; DEGs: Differentially expressed genes; ROC: Receiver operating characteristic; AUC: Area under the ROC Curve; miRNAs: Micro RNAs; SNP: Single nucleotide polymorphism; SLE: Systemic lupus erythematosus; MCA: Middle cerebral artery vascular; STA: Superficial temporal artery; IA: Internal carotid aneurysm; EPI: Epilepsy; PCA: Principal component analysis; KEGG: Kyoto Encyclopedia of Genes and Genomes; GO enrichment: Gene Ontology Term Enrichment; PPI: Protein-protein interaction; NETs: Neutrophil extracellular traps; SMCs: Smooth muscle cells.

Declarations

Ethics approval and consent to participate

Not applicable. The GEO belongs to public databases. The patients involved in the database have obtained ethical approval. Users can download relevant data for free for research and publish relevant articles. Our study is based on open-source data, so there are no ethical issues and other conflicts of interest.

Consent for publication

All authors agreed on the manuscript.

Availability of data and materials

The datasets analyzed in this study were downloaded and accessed from the Gene Expression Omnibus (GEO) database: <https://www.ncbi.nlm.nih.gov/geo/>, with accession No: GSE157628, GSE141024.

Competing interests

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

Funding

This work was supported by a grant from the Key Project of Clinical Research of Southern Medical University (Grant number: LC2016ZD024).

Authors' contributions

Fa Jin conducted data analysis, data visualization, and writing manuscripts. Chuanzhi Duan conceived and revised the manuscript.

Acknowledgements

We thank GEO database for providing their platforms and contributors for uploading their meaningful datasets.

References

1. Acker G, Fekonja L, Vajkoczy P. Surgical Management of Moyamoya Disease. *Stroke*. 2018;49:476–82.
2. Suzuki J, Takaku A. Cerebrovascular “moyamoya” disease. Disease showing abnormal net-like vessels in base of brain. *Arch Neurol*. 1969;20:288–99.
3. Research Committee on the Pathology and Treatment of Spontaneous Occlusion of the Circle of Willis, Health Labour Sciences Research Grant for Research on Measures for Infractable Diseases. Guidelines for diagnosis and treatment of moyamoya disease (spontaneous occlusion of the circle of Willis). *Neurol Med Chir (Tokyo)*. 2012;52:245–66.
4. Acker G, Goerdes S, Schneider UC, Schmiedek P, Czabanka M, Vajkoczy P. Distinct clinical and radiographic characteristics of moyamoya disease amongst European Caucasians. *Eur J Neurol*. 2015;22:1012–7.
5. Baba T, Houkin K, Kuroda S. Novel epidemiological features of moyamoya disease. *J Neurol Neurosurg Psychiatry*. 2008;79:900–4.
6. Kronenburg A, Braun KPJ, van der Zwan A, Klijn CJM. Recent advances in moyamoya disease: pathophysiology and treatment. *Curr Neurol Neurosci Rep*. 2014;14:423.

7. Kamada F, Aoki Y, Narisawa A, Abe Y, Komatsuzaki S, Kikuchi A, et al. A genome-wide association study identifies RNF213 as the first Moyamoya disease gene. *J Hum Genet.* 2011;56:34–40.
8. Liu W, Morito D, Takashima S, Mineharu Y, Kobayashi H, Hitomi T, et al. Identification of RNF213 as a susceptibility gene for moyamoya disease and its possible role in vascular development. *PLoS One.* 2011;6:e22542.
9. Fujimura M, Sonobe S, Nishijima Y, Niizuma K, Sakata H, Kure S, et al. Genetics and Biomarkers of Moyamoya Disease: Significance of RNF213 as a Susceptibility Gene. *J Stroke.* 2014;16:65–72.
10. Kim HJ, Choi E-H, Chung J-W, Kim J-H, Kim YS, Seo W-K, et al. Role of the RNF213 Variant in Vascular Outcomes in Patients With Intracranial Atherosclerosis. *J Am Heart Assoc.* 2020;e017660.
11. Park YS, Jeon YJ, Lee BE, Kim TG, Choi J-U, Kim D-S, et al. Association of the miR-146aC > G, miR-196a2C > T, and miR-499A > G polymorphisms with moyamoya disease in the Korean population. *Neurosci Lett.* 2012;521:71–5.
12. Uchino H, Ito M, Kazumata K, Hama Y, Hamauchi S, Terasaka S, et al. Circulating miRNome profiling in Moyamoya disease-discordant monozygotic twins and endothelial microRNA expression analysis using iPS cell line. *BMC Med Genomics.* 2018;11:72.
13. Bang OY, Fujimura M, Kim S-K. The Pathophysiology of Moyamoya Disease: An Update. *J Stroke.* 2016;18:12–20.
14. Bower RS, Mallory GW, Nwojo M, Kudva YC, Flemming KD, Meyer FB. Moyamoya disease in a primarily white, midwestern US population: increased prevalence of autoimmune disease. *Stroke.* 2013;44:1997–9.
15. Kim SJ, Heo KG, Shin HY, Bang OY, Kim G-M, Chung C-S, et al. Association of thyroid autoantibodies with moyamoya-type cerebrovascular disease: a prospective study. *Stroke.* 2010;41:173–6.
16. Watanabe C, Oishi T, Yamamoto T, Sasaki K, Tosaka M, Sato T, et al. Chorea and Broca aphasia induced by diabetic ketoacidosis in a type 1 diabetic patient diagnosed as Moyamoya disease. *Diabetes Res Clin Pract.* 2005;67:180–5.
17. Suzuki J, Kodama N. Moyamoya disease—a review. *Stroke.* 1983;14:104–9.
18. Raut SA, Sathe SR, Raut A. Bioinformatics: Trends in gene expression analysis. 2010 International Conference on Bioinformatics and Biomedical Technology. 2010. p. 97–100.
19. Sigdel TK, Shoemaker LD, Chen R, Li L, Butte AJ, Sarwal MM, et al. Immune response profiling identifies autoantibodies specific to Moyamoya patients. *Orphanet J Rare Dis.* 2013;8:45.
20. Wang C, Zhao M, Wang J, Zhang D, Wang S, Zhao J. Expression analysis of transfer RNA–derived fragments in the blood of patients with moyamoya disease: A preliminary study. *Mol Med Rep.* 2019;19:3564–74.
21. Zhao M, Gao F, Zhang D, Wang S, Zhang Y, Wang R, et al. Altered expression of circular RNAs in Moyamoya disease. *J Neurol Sci.* 2017;381:25–31.
22. Wang G, Wen Y, Faleti OD, Zhao Q, Liu J, Zhang G, et al. A Panel of Exosome-Derived miRNAs of Cerebrospinal Fluid for the Diagnosis of Moyamoya Disease. *Front Neurosci.* 2020;14:548278.

23. Nakamura K, Yanaka K, Ihara S, Nose T. Multiple intracranial arterial stenoses around the circle of Willis in association with Graves' disease: report of two cases. *Neurosurgery*. 2003;53:1210–4. discussion 1214–1215.
24. Inaba M, Henmi Y, Kumeda Y, Ueda M, Nagata M, Emoto M, et al. Increased stiffness in common carotid artery in hyperthyroid Graves' disease patients. *Biomed Pharmacother*. 2002;56:241–6.
25. Masuda J, Ogata J, Yutani C. Smooth muscle cell proliferation and localization of macrophages and T cells in the occlusive intracranial major arteries in moyamoya disease. *Stroke*. 1993;24:1960–7.
26. Martinod K, Demers M, Fuchs TA, Wong SL, Brill A, Gallant M, et al. Neutrophil histone modification by peptidylarginine deiminase 4 is critical for deep vein thrombosis in mice. *Proc Natl Acad Sci U S A*. 2013;110:8674–9.
27. Seignez C, Phillipson M. The multitasking neutrophils and their involvement in angiogenesis. *Curr Opin Hematol*. 2017;24:3–8.
28. Papayannopoulos V. Neutrophil extracellular traps in immunity and disease. *Nat Rev Immunol*. 2018;18:134–47.
29. Urban CF, Ermert D, Schmid M, Abu-Abed U, Goosmann C, Nacken W, et al. Neutrophil extracellular traps contain calprotectin, a cytosolic protein complex involved in host defense against *Candida albicans*. *PLoS Pathog*. 2009;5:e1000639.
30. Fuchs TA, Abed U, Goosmann C, Hurwitz R, Schulze I, Wahn V, et al. Novel cell death program leads to neutrophil extracellular traps. *J Cell Biol*. 2007;176:231–41.
31. Rossaint J, Kühne K, Skupski J, Van Aken H, Looney MR, Hidalgo A, et al. Directed transport of neutrophil-derived extracellular vesicles enables platelet-mediated innate immune response. *Nat Commun*. 2016;7:13464.
32. Etulain J, Martinod K, Wong SL, Cifuni SM, Schattner M, Wagner DD. P-selectin promotes neutrophil extracellular trap formation in mice. *Blood*. 2015;126:242–6.
33. von Brühl M-L, Stark K, Steinhart A, Chandraratne S, Konrad I, Lorenz M, et al. Monocytes, neutrophils, and platelets cooperate to initiate and propagate venous thrombosis in mice in vivo. *J Exp Med*. 2012;209:819–35.
34. Tecchio C, Cassatella MA. Neutrophil-derived cytokines involved in physiological and pathological angiogenesis. *Chem Immunol Allergy*. 2014;99:123–37.
35. Corey S, Luo Y. Circular RNAs and neutrophils: Key factors in tackling asymptomatic moyamoya disease. *Brain Circ*. 2019;5:150–5.
36. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res*. 2015;43:e47.
37. Kanehisa M, Goto S. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res*. 2000;28:27–30.
38. Zhou Y, Zhou B, Pache L, Chang M, Khodabakhshi AH, Tanaseichuk O, et al. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat Commun*. 2019;10:1523.

39. Chen B, Khodadoust MS, Liu CL, Newman AM, Alizadeh AA. Profiling tumor infiltrating immune cells with CIBERSORT. *Methods Mol Biol.* 2018;1711:243–59.

40. von Mering C, Huynen M, Jaeggi D, Schmidt S, Bork P, Snel B. STRING: a database of predicted functional associations between proteins. *Nucleic Acids Res.* 2003;31:258–61.

41. Bader GD, Hogue CWV. An automated method for finding molecular complexes in large protein interaction networks. *BMC Bioinformatics.* 2003;4:2.

42. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez J-C, et al. pROC: an open-source package for R and S + to analyze and compare ROC curves. *BMC Bioinformatics.* 2011;12:77.

43. Wickham H. ggplot2. *WIREs Computational Statistics.* 2011;3:180–5.

44. Raivo Kolde. pheatmap: Pretty Heatmaps. R package [Internet]. 2015; Available from: <https://cran.r-project.org/web/packages/pheatmap/index.html>.

Tables

TABLE 1 | The enriched gene ontology terms for neutrophil-associated DEGs.

Term	Description	Count	Genes	P-value
GO:1990266	neutrophil migration	10	C3AR1, CCL11, CCL15, CCL20, CCL7, CXCL5, DAPK2, PDE4B, UMOD, XG	7.33E-15
GO:0030593	neutrophil chemotaxis	8	C3AR1, CCL11, CCL15, CCL20, CCL7, CXCL5, DAPK2, PDE4B	8.57E-12
GO:0072677	eosinophil migration	5	CCL11, CCL15, CCL7, DAPK2, EPX	1.05E-09
GO:0008009	chemokine activity	5	CCL11, CCL15, CCL20, CCL7, CXCL5	2.25E-08
GO:0048245	eosinophil chemotaxis	4	CCL11, CCL15, CCL7, DAPK2	7.79E-08
GO:2000403	positive regulation of lymphocyte migration	3	CCL20, CCL7, PYCARD	3.68E-05
GO:1902624	positive regulation of neutrophil migration	3	C3AR1, DAPK2, XG	2.89E-05
GO:1902622	regulation of neutrophil migration	3	C3AR1, DAPK2, XG	6.48E-05
GO:2000406	positive regulation of T cell migration	2	CCL20, PYCARD	0.001295116
GO:2000144	positive regulation of DNA-templated transcription, initiation	2	CAND1, PSMC3	0.001295116
GO:0090023	positive regulation of neutrophil chemotaxis	2	C3AR1, DAPK2	0.001138195
GO:0070944	neutrophil-mediated killing of bacterium	2	AZU1, ELANE	5.63E-05
GO:0060261	positive regulation of transcription initiation from RNA polymerase II promoter	2	CAND1, PSMC3	7.89E-04
GO:0060260	regulation of transcription initiation from RNA polymerase II promoter	2	CAND1, PSMC3	0.001377244
GO:0051123	RNA polymerase II preinitiation complex assembly	2	CAND1, PSMC3	0.001377244
GO:0045899	positive regulation of RNA polymerase II transcription preinitiation complex assembly	2	CAND1, PSMC3	1.20E-04
GO:0045898	regulation of RNA polymerase II transcription preinitiation complex assembly	2	CAND1, PSMC3	2.42E-04
GO:0045073	regulation of chemokine biosynthetic process	2	AZU1, ELANE	3.61E-04
GO:0043320	natural killer cell degranulation	2	UNC13D, VAMP7	7.50E-05
GO:0043304	regulation of mast cell degranulation	2	UNC13D, VAMP7	0.001138195
GO:0042104	positive regulation of activated T cell proliferation	2	EPX, PYCARD	7.89E-04
GO:0035584	calcium-mediated signaling using intracellular calcium source	2	AZU1, CCL20	9.21E-04
GO:0002693	positive regulation of cellular extravasation	2	ELANE, XG	7.27E-04

TABLE 2 | Characteristic of Hub genes involved in the regulation of neutrophils.

Gene symbol	Gene title	location	Function	P-value
UNC13D	unc-13 homolog D	Chromosome 17	Neutrophil degranulation	0.01615
AZU1	azurocidin 1	Chromosome 19	Neutrophil degranulation	0.04886
SDCBP	syndecan Binding Protein	Chromosome 8	Neutrophil degranulation	0.01608
ELANE	elastase, neutrophil expressed	Chromosome 19	A serine protease secreted by neutrophils	0.04862
PYCARD	PYD And CARD Domain Containing	Chromosome 16	Neutrophil activation and degranulation	0.01679
CCL11	C-C motif chemokine ligand 11	Chromosome 17	Neutrophil recruitment	0.04675
CCL15	C-C motif chemokine ligand 15	Chromosome 17	Neutrophil recruitment	0.01857
CCL20	C-C motif chemokine ligand 20	Chromosome 2	Neutrophil recruitment	0.01271
CXCL5	C-X-C Motif Chemokine Ligand 5	Chromosome 4	Neutrophil recruitment	0.02993

Figures

Figure 1

A workflow of bioinformatics analysis.

Figure 2

Normalization of all samples and DEG analysis. (a) Boxplots of sample expression, before normalization and after normalization. The x-axis is the sample information, and the y-axis is the gene expression value. All samples are middle cerebral artery vessel walls; blue (MMD) represents Moyamoya disease samples, green (IA) represents internal carotid aneurysm samples, and red (EPI) represents epilepsy samples. (a) PCA visualizes the grouping information of samples. (c) The volcano plot with all genes; each dot

represents a gene. Genes that meet the screening criteria are shown in blue (down-regulated) and red (up-regulated), and the remaining genes are designated in gray. (d) A heatmap of the top 200 DEGs includes 105 up-regulated genes and 95 down-regulated genes. The first column shows the grouping information. Each row shows one gene, one sample per column, including 11 MMD case samples and 3 EPI as negative control samples. Up-regulated genes are represented in bright, and down-regulated genes are represented in dark.

Figure 3

Functional enrichment analysis of DEGs. (a-b) Reactome pathway and KEGG pathway analysis of the down-regulated DEGs. (c-d) Pathway analysis of the down-regulated DEGs. The size of the dot represents the gene counts enriched in the term or signaling pathway. The color of each dot indicates the degree of significance.

Figure 4

The landscape of immune infiltration between MMD and NC. (a) Violin chart of the ratio of immune cells. Y-axis is the percentage of immune cells; the symbols at the top represent statistical significance between the two groups. ns: not significant; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$. (b) The bar graph on the left shows the percentage of immune cell examples (mean) in samples from the negative control and MMD groups. Different colors represent immune cell populations; the table on the right shows the specific percentage values. NC: negative control, MMD: Moyamoya disease, SD: standard deviation.

Figure 5

Identification of hub genes in neutrophil-associated DEGs. (a) Venn diagram represents the intersection genes taken between the differential gene set of MMD and neutrophil-associated gene set. A total of 30

intersection genes were identified. (b) Neutrophil-associated DEGs were extracted, and drawn as a heatmap. The heat map row names are the gene symbols intersected by the list of DEGs and neutrophil-associated genes. (c) Neutrophil-associated DEGs were subjected to GO functional enrichment analysis. Network diagram between enriched GO terms, each dot is a pathway of enriched biological processes, the color of the dot indicates the terms of the cluster, and the most significant terms in the cluster are identified in bold font with color. (d) PPI network relationship graph of neutrophil-associated DEGs predicted by STRING. Each dot represents the protein molecule ultimately transcribed by each DEG and is distinguished by a different color. Connecting lines between proteins indicate the presence of their interactions, and the thickness of lines indicates the strength of interactions. The bottom row of dots is protein molecules for which no network relationships are identified. (e) Identification of hub genes from the PPI network by MCODE. Cluster sub-networks were identified using Cytoscape's MCODE plugin, and the top two sub-networks are shown. (f) Heat map of the correlation between hub genes and immune cell ratios. The X-axis indicates the expression of hub genes; Y-axis shows the percentage of immune cell content, pie chart size represents the correlation coefficient, blue is positive correlation, red is negative correlation, color shade correlation coefficient is proportional to the absolute value size.

Figure 6

The relative expression of neutrophil-associated hub genes. (a-i) Expression values were extracted from the raw microarray data, and their ratio to the internal reference was calculated to compare the relative expression of hub genes in all samples of GSE157628. Three cases of EPI, six cases of IA, 11 cases of MMD, t-test, $p < 0.05$ was considered statistically significant.

Figure 7

Test the candidate hub genes by ROC. (a) ROCs for four candidate genes (UNC13D, AZU1, PYCARD, CCL15) are shown, respectively. Samples for ROC testing include all vascular tissue samples from the GSE157628 and GSE141024. The x-axis indicates specificity, the y-axis indicates sensitivity, error bars are CI, and AUC represents the area under the curve. (b) Combined ROCs for comparison. (c) Table of detailed ROC analysis data.

Supplementary Files

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

- [SupplementaryTable1.xlsx](#)