Comprehensive molecular analyses of an autoimmune-related gene predictive model and immune infiltrations using machine learning methods in moyamoya disease

Shifu Li[1,2†], Ying Han[3,4†], Qian Zhang[1,2], Dong Tang[1,2], Jian Li[1,2,5] and Ling Weng[2,6,7,8,9]*

[1]Department of Neurosurgery, Xiangya Hospital, Central South University, Changsha, Hunan, China, [2]National Clinical Research Center for Geriatric Disorders, Central South University, Changsha, Hunan, China, [3]Department of Oral and Maxillofacial Surgery, Center of Stomatology, Xiangya Hospital of Central South University, Changsha, Hunan, China, [4]Center for Medical Genetics and Hunan Key Laboratory of Medical Genetics, School of Life Sciences, Central South University, Changsha, Hunan, China, [5]Hydrocephalus Center, Xiangya Hospital, Central South University, Changsha, Hunan, China, [6]Department of Neurology, Xiangya Hospital, Central South University, Changsha, Hunan, China, [7]Engineering Research Center of Hunan Province in Cognitive Impairment Disorders, Central South University, Changsha, China, [8]Hunan International Scientific and Technological Cooperation Base of Neurodegenerative and Neurogenetic Diseases, Changsha, China, [9]Key Laboratory of Hunan Province in Neurodegenerative Disorders, Central South University, Changsha, China

Background: Growing evidence suggests the links between moyamoya disease (MMD) and autoimmune diseases. However, the molecular mechanism from genetic perspective remains unclear. This study aims to clarify the potential roles of autoimmune-related genes (ARGs) in the pathogenesis of MMD.

Methods: Two transcription profiles (GSE157628 and GSE141025) of MMD were downloaded from GEO databases. ARGs were obtained from the Gene and Autoimmune Disease Association Database (GAAD) and DisGeNET databases. Differentially expressed ARGs (DEARGs) were identified using "limma" R packages. GO, KEGG, GSVA, and GSEA analyses were conducted to elucidate the underlying molecular function. There machine learning methods (LASSO logistic regression, random forest (RF), support vector machine-recursive feature elimination (SVM-RFE)) were used to screen out important genes. An artificial neural network was applied to construct an autoimmune-related signature predictive model of MMD. The immune characteristics, including immune cell infiltration, immune responses, and HLA gene expression in MMD, were explored using ssGSEA. The miRNA-gene regulatory network and the potential therapeutic drugs for hub genes were predicted.

Results: A total of 260 DEARGs were identified in GSE157628 dataset. These genes were involved in immune-related pathways, infectious diseases, and autoimmune diseases. We identified six diagnostic genes by overlapping the three machine learning algorithms: CD38, PTPN11, NOTCH1, TLR7, KAT2B, and

ISG15. A predictive neural network model was constructed based on the six genes and presented with great diagnostic ability with area under the curve (AUC) = 1 in the GSE157628 dataset and further validated by GSE141025 dataset. Immune infiltration analysis showed that the abundance of eosinophils, natural killer T (NKT) cells, Th2 cells were significant different between MMD and controls. The expression levels of HLA-A, HLA-B, HLA-C, HLA-DMA, HLA-DRB6, HLA-F, and HLA-G were significantly upregulated in MMD. Four miRNAs (mir-26a-5p, mir-1343-3p, mir-129-2-3p, and mir-124-3p) were identified because of their interaction at least with four hub DEARGs.

Conclusion: Machine learning was used to develop a reliable predictive model for the diagnosis of MMD based on ARGs. The uncovered immune infiltration and gene-miRNA and gene-drugs regulatory network may provide new insight into the pathogenesis and treatment of MMD.

Introduction

Moyamoya disease (MMD) is an uncommon, chronic cerebrovascular disorder characterized by progressive occlusion of the supraclinoid internal carotid artery (ICA) and its main branches within the circle of Willis. MMD, also known as an abnormal netlike vascular disease at the base of the brain, is a term coined by Suzuki and Takaku in 1969 to describe the classic angiographic appearance—a puff of cigarette smoke drifting in the air (Suzuki and Takaku, 1969). Important clinical features include ischemic stroke, often presented in childhood, and hemorrhagic stroke, generally observed in adults. The incidence rate is twice higher in females than males, and children around 5 years and adults in their mid-40s are particularly affected (Kuroda and Houkin, 2008). MMD is the most common pediatric cerebrovascular disease in Japan, affecting approximately three out of every 100,000 children (Scott and Smith, 2009). MMD conventionally refers to patients with above idiopathic pathology without a previously diagnosed condition. Distinct from the definitive MMD, Moyamoya (MM) syndrome (also named Quasi-Moyamoya disease, secondary Moyamoya disease, and akin-Moyamoya disease) is the occurrence of angiographic MM in association with acquired (i.e., autoimmune diseases) or inherited disorders [i.e., neurofibromatosis type 1, sickle cell anemia, Down syndrome (DS)] (Scott and Smith, 2009).

The exact etiology of MMD remains unknown; however, hereditary, immunogenic inflammatory, and hemodynamic factors are known to be responsible. The close relationship between patients with MM syndrome and autoimmune diseases, such as type 1 diabetes mellitus, thyroid disease, systemic lupus erythematosus (SLE), and DS has been reported (Huang et al., 2017). A study in a primarily white, midwestern United States population showed that the prevalence of autoimmune diseases was significantly higher in patients with MMD, particularly type 1 diabetes mellitus (8.5% *versus* 0.4% in

the general population), thyroid disease (17.0% *versus* 8.0% in the institutional general patient population), and hyperlipidemia (27.7% *versus* 16.3% in the general population). A meta-analysis conducted in 2014 revealed that elevated thyroid autoantibodies and elevated thyroid function are independently associated risk factors for MMD (Lei et al., 2014). Autoimmunity is the main link between SLE and MM syndrome since immune complexes lead to vasculitis and narrow or occluded vessels (El Ramahi and Al Rayes, 2000; Jeong et al., 2008). Chen et al. (2016) found that the overall prevalence of autoimmune diseases in patients with unilateral MMD was significantly higher than that in patients with bilateral MMD. Although the close relationships between autoimmune diseases and MDD diseases have been recognized, the underlying mechanisms remain to be clarified.

Advances in molecular biology and next-generation sequencing technologies have made it possible to study disease mechanisms at the genetic and mRNA levels. Gene expression profiling through methods such as microarray and RNA sequencing based on the Gene Expression Omnibus (GEO) database is widely used to explore differentially expressed genes (DEGs), analyze potential function pathways, and determine molecular mechanisms involved in various cerebrovascular diseases (Chen et al., 2022). A recent bioinformatics study identified the potential neutrophil-associated genes in MMD (Jin and Duan, 2022). However, the role of autoimmune-related genes (ARGs) in the pathophysiology of MMD is still unclear. In recent years, the development of machine learning algorithms has provided more choices for diagnostic models as precision medical predictive tools. Our study integrated least absolute shrinkage and selection operator (LASSO) logistic regression, random forest (RF), support vector machine-recursive feature elimination (SVM-RFE), and artificial neural network to screen and identify

diagnostic markers and construct an autoimmune-related signature predictive model of MMD. The immune characteristics, including immune cell infiltration, immune responses, and HLA gene expression in MMD, were explored. The miRNA-gene regulatory network and the potential therapeutic drugs for hub genes were predicted.

Materials and methods

Downloading and processing of data

Microarray data containing two transcription profiles (GSE157628 and GSE141025) were downloaded from the NCBI GEO database (https://www.ncbi.nlm.nih.gov/geo/). The dataset of GSE157628 was utilized as the exploratory dataset, and the GSE141025 profile acted as the validation dataset. The GSE157628 profile included micro-samples of the middle cerebral artery (MCA) collected from 11 patients with MMD and nine age- and gender-matched control samples (three from patients with epilepsy and six from patients with ICA aneurysms) at the platform of GPL16699. The expression profiles of MCA samples from four patients with MMD and four matched superficial temporal artery controls in GSE141025 were extracted for dataset validation. If multiple probes matched one gene, the probe with the maximal median expression values was annotated into the homologous gene symbol through the platform's annotation information.

Collection of autoimmune-related genes

ARGs were obtained from the Gene and Autoimmune Disease Association Database (GAAD) (Lu et al., 2018) and DisGeNET databases (Piñero et al., 2020) after deleting duplicate genes. GAAD contained 44,762 associations between 49 autoimmune-related diseases and 4,249 genes through text mining and manual curation. DisGeNET (v7.0), one of the largest publicly available collections of genes and variants associated with human diseases, contained 1,134,942 gene-disease associations (GDAs), between 21,671 genes and 30,170 diseases, disorders, traits, and clinical or abnormal human phenotypes.

Identification of differentially expressed autoimmune-related genes

The principal component analysis was conducted by using the factoextra R package. To identify the DEARGs, we performed differential expression analysis using the "limma" package in R software to detect DEGs between the MMD and control groups in the training dataset (GSE157628). The DEGs were screened with the criteria of $|log2FoldChange| > 1$ and $p < 0.05$. Volcano maps and clustering heatmaps were prepared to visualize the differences using the "ggplot2" and "ComplexHeatmap" packages in the R software. We intersected the DEGs with ARGs to identify DEARGs and visualized them with the "VennDiagram" package.

Functional enrichments between moyamoya disease tissues and controls

To uncover the biological function in MMD, gene set enrichment analysis (GSEA) was used to enrich the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways using the "clusterProfiler" package. Gene set variation analysis (GSVA) algorithm was used to calculate the Reactome ("c2.cp.reactome.v7.5.1.symbols" gene set from the Molecular Signatures Database) (Liberzon et al., 2015) processes score using the "GSVA" R package.

Authentication of the organ/tissue-specific expressed DEARGs

To understand the tissue/organ-specific expression of these DEARGs, we analyzed the gene distribution in tissues using the online tool BioGPS (http://biogps.org/) (Wu et al., 2009). The following criteria had to be met (Wang et al., 2020): 1) the expression level of transcripts mapped to a single organ system was >10 times the median, and 2) the second-highest level was not more than one-third of the highest expression level.

Protein-protein interaction network and functional annotation of DEARGs

The PPI network of DEARGs was prepared using the online tool STRING (https://string-db.org/) with a minimum required interaction score of 0.4. We downloaded the interaction information and visualized the PPI network using Cytoscape software (v3.8.2). ClueGO, a plugin app of Cytoscape for the function enrichment, was used to annotate the biological processes (BP) of Gene Ontology (GO) and KEGG pathways of this network and the genes participating in these terms. The R package "DOSE" was used to perform Disease Ontology (DO) analysis (Yu et al., 2015).

Screening for crucial DEARGs and candidate signatures

First, we applied five methods (Closeness, Degree, MCC(Maximal Clique Centrality), MNC (Maximum

neighborhood component), and Radiality) in cytoHubba to select the top 30 genes and intersected them through the Venn plot to find the common genes (Chin et al., 2014). Based on common genes, we constructed a co-expression network *via* GeneMANIA (http://www.genemania.org/) to identify internal associations (Warde-Farley et al., 2010).

Second, to further identify the crucial DEARGs and candidate signatures, three machine learning [LASSO logistic regression, RF, and support vector machine (SVM)] algorithms were adopted. The LASSO logistic regression model was used to select optimal variables using the penalty coefficient. RF is a machine learning algorithm with an ensemble of multiple decision trees that combines the knowledge generated by a collection of individual trees using randomness. The top 10 variables were selected as the most important features in the methods. SVM is a supervised machine learning technique widely utilized for classification and regression. To avoid over-fitting, the SVM-RFE requires training multiple classifiers on subsets of features of decreasing size to search for the best features. In the present study, we overlapped genes identified by the three methods and designated the communal gens as our candidate signatures for constructing the next diagnostic model.

Construction of the artificial neural network diagnostic model

We constructed a back propagation artificial neural network model using the "neuralnet" package. The expression profiles of the above screened signatures in the GSE157628 dataset were extracted and normalized. The min-max method was selected, and the data were mapped in the range of zero to one before training the neural network. The number of neurons was between the input and output layer sizes, usually two-thirds of the input size. A single hidden layer with four nodes was used. We calculated the classification score of the obtained disease neural network model as follows:

$$\text{Predicting score} = \sum gene\ expression * neural\ network\ weight$$

The diagnostic ability was evaluated through the receiver operating characteristic (ROC) curve and confusion matrix. Another external dataset, GSE141025, was used to validate this model.

Immune characteristics of the moyamoya disease microenvironment

xCell Aran et al. (2017), a novel gene signature-based method to identify 64 immune and stromal cell types, was used to score the abundance of immune cells in MMD. Single sample gene set enrichment analysis (ssGSEA) was used to analyze the immune cells and activities between MMD and controls. We also compared the expression levels of HLA molecules between the

two groups. The significantly different immune characteristics were depicted with boxplots and heatmaps.

The association of the crucial gene signatures with the scores of infiltrating immune cells, immune activities, and expression of HLA molecules was explored using Pearson's correlation analysis in R software. The resulting associations were visualized as a heatmap prepared with the "ggplot2" package.

Prediction of a miRNA-genes regulatory network and potential drugs

NetworkAnalyst is a user-friendly online tool to create PPI networks, cell-type or tissue-specific PPI networks, gene regulatory networks, gene co-expression networks, and networks for toxicogenomics and pharmacogenomics studies (Zhou et al., 2019). We used the NetworkAnalyst to predict the miRNA-genes regulatory network through the Tarbase database (Karagkouni et al., 2018).

The Drug-Gene Interaction Database (DGIdb) (http://www.dgidb.org/) is an online database to predict drug-gene interaction based on the data mined from DrugBank, PharmGKB, Chembl, Drug Target Commons, and TTD. The DGIdb was searched to make predictions on potential molecule-related drugs that interact with crucial DEARGs. Only the drugs with identified interaction types persisted, and the Drugs-Genes interactions were visualized through a Sankey diagram.

Sample collection and real-time quantitative polymerase chain reaction

We recruited ten patients who were diagnosed with MMD in Xiangya Hospital of Central South University for this research between June 2022 and October 2022. A total of 10 healthy controls with gender and age matched were also selected.

The detailed procedures of RT-qPCR were described in our previous studies (Li et al., 2022a; Li et al., 2022b). Birefly, peripheral blood monocytes (PBMCs) were isolated from the blood samples of patients and health persons. We extracted the total RNA from the PBMCs, performed reverse transcription reactions, and then amplified the cDNA. The results were analyzed using the $2^{-\Delta\Delta CT}$ method and expressed as ratio of the internal control, GAPDH. The primer sequences used for RT-qPCR are listed in Supplementary Table S1.

Results

Differential expression analysis

The study flowchart is depicted in Figure 1. Principal component analysis showed that the MMD tissues and
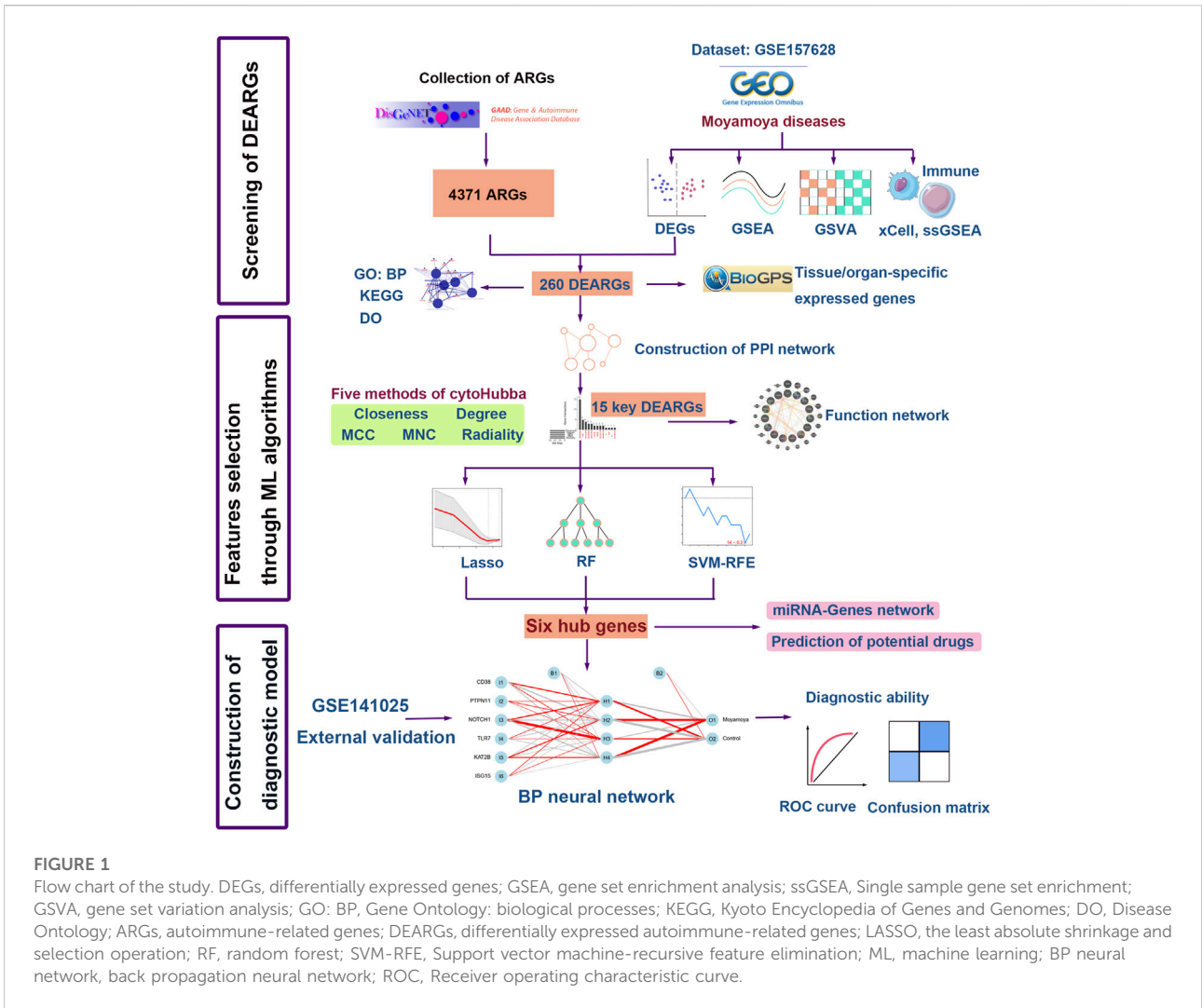
FIGURE 1
Flow chart of the study. DEGs, differentially expressed genes; GSEA, gene set enrichment analysis; ssGSEA, Single sample gene set enrichment; GSVA, gene set variation analysis; GO: BP, Gene Ontology: biological processes; KEGG, Kyoto Encyclopedia of Genes and Genomes; DO, Disease Ontology; ARGs, autoimmune-related genes; DEARGs, differentially expressed autoimmune-related genes; LASSO, the least absolute shrinkage and selection operation; RF, random forest; SVM-RFE, Support vector machine-recursive feature elimination; ML, machine learning; BP neural network, back propagation neural network; ROC, Receiver operating characteristic curve.
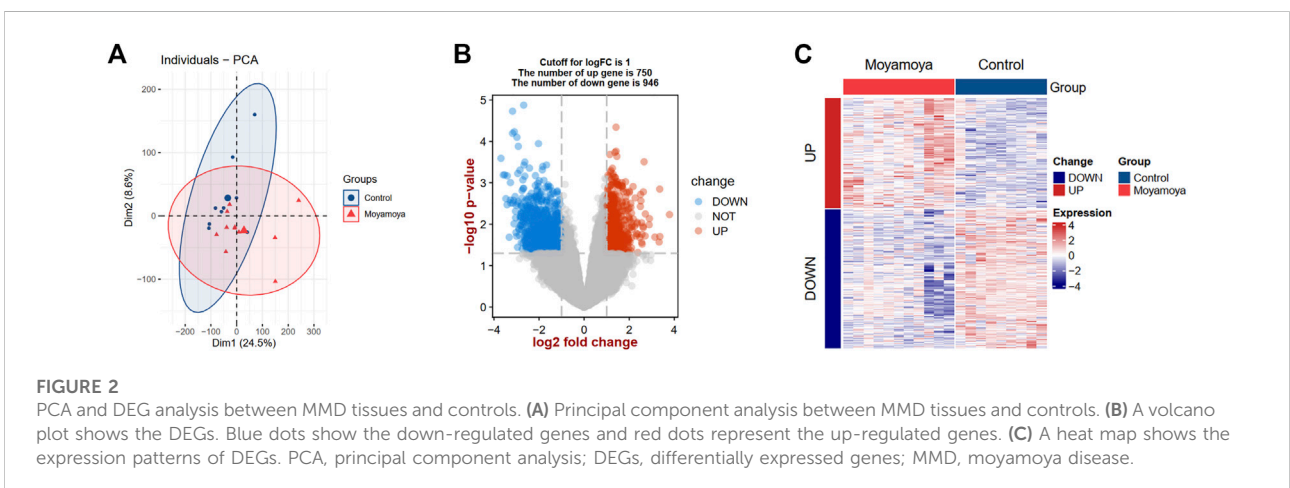
FIGURE 2
PCA and DEG analysis between MMD tissues and controls. **(A)** Principal component analysis between MMD tissues and controls. **(B)** A volcano plot shows the DEGs. Blue dots show the down-regulated genes and red dots represent the up-regulated genes. **(C)** A heat map shows the expression patterns of DEGs. PCA, principal component analysis; DEGs, differentially expressed genes; MMD, moyamoya disease.

FIGURE 3
Biological KEGG and Reactome pathways involved in MMD based on GSEA and GSVA. **(A)** and **(B)** Up-regulated and down-regulated KEGG pathways from GSEA results, respectively. **(C)** The top differentially regulated reactome pathways from GSVA results.

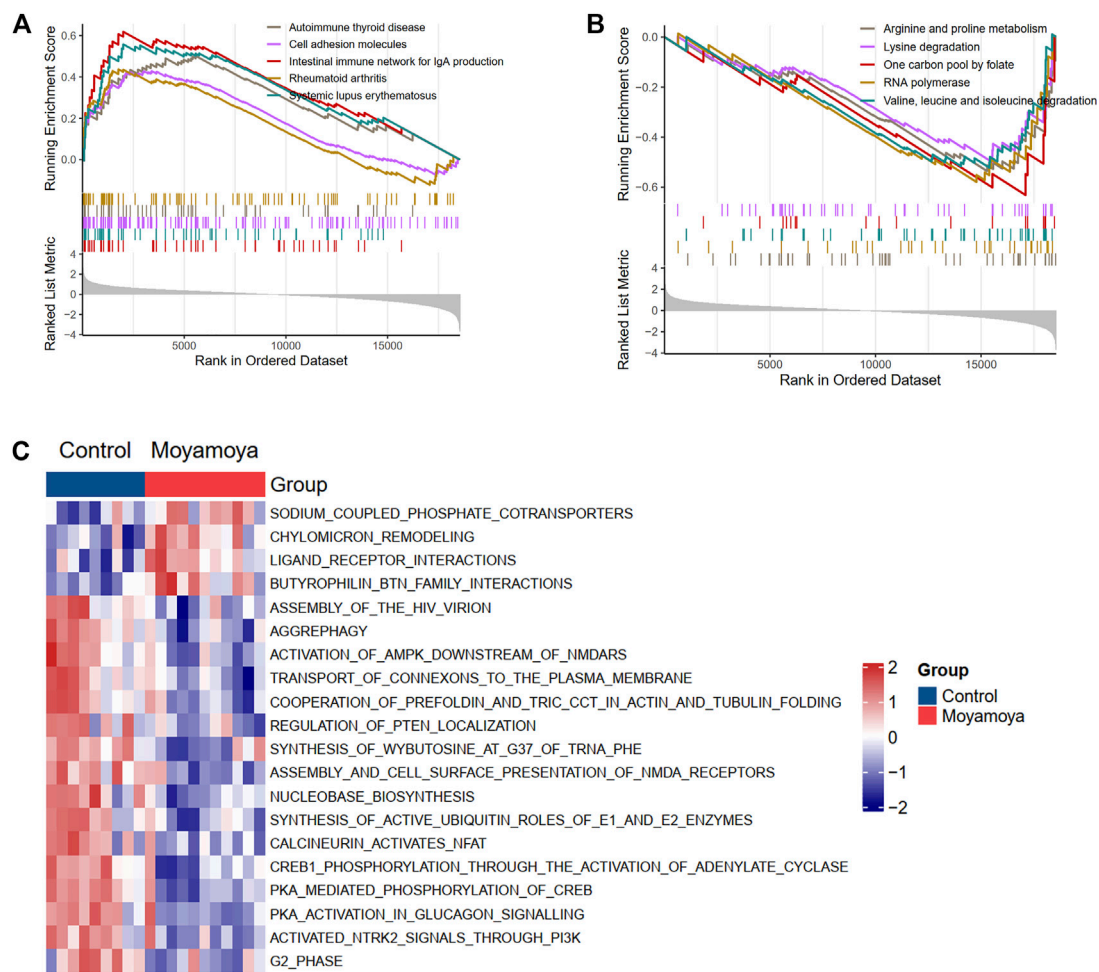controls could be clearly distinguished in the GSE157628 dataset (Figure 2A). Differential expression analysis was further performed to screen for DEGs. Based on the selection criteria, 1696 DEGs (750 upregulated and 946 downregulated) were identified (Figure 2B). The expression patterns of these DEGs were visualized through a hierarchical clustering heatmap (Figure 2C).

Functional enrichment analyses in moyamoya disease

GSEA and GSVA were performed to reveal the underlying biological pathways in MMD. The KEGG and Reactome analysis focus on the biological pathways. Compared with the comprehensive KEGG pathway, the Reactome pathway has more specific functions and focuses more on biochemical reactions. In our analysis, both methods were adopted. The KEGG analysis based on GSEA showed that the autoimmune thyroid disease, cell adhesion molecules, and rheumatoid arthritis pathways were upregulated (Figure 3A). In contrast, metabolism-related pathways (arginine and proline metabolism, lysine degradation, and one carbon pool by folate) were downregulated in MMD tissues (Figure 3B) compared with controls. The top 20 significantly differential Reactome pathways between MMD and controls are presented in Figure 3C, showing upregulated sodium-coupled phosphate cotransporters, chylomicron remodeling, and ligand-receptor interactions pathways. At the same time, aggrephagy and regulation of PTEN localization were downregulated in MMD tissues compared with controls (Figure 3C).

FIGURE 4
Identification and PPI network construction of DEARGs. **(A)** A venn plot show 260 DEARGs in MMD. **(B)** PPI network of DEARGs. The blue nodes represent the down-regulated genes and the red nodes represent the up-regulated genes. The dot size indicates the degree of the nodes. DEARGs, differentially expressed autoimmune-related genes.
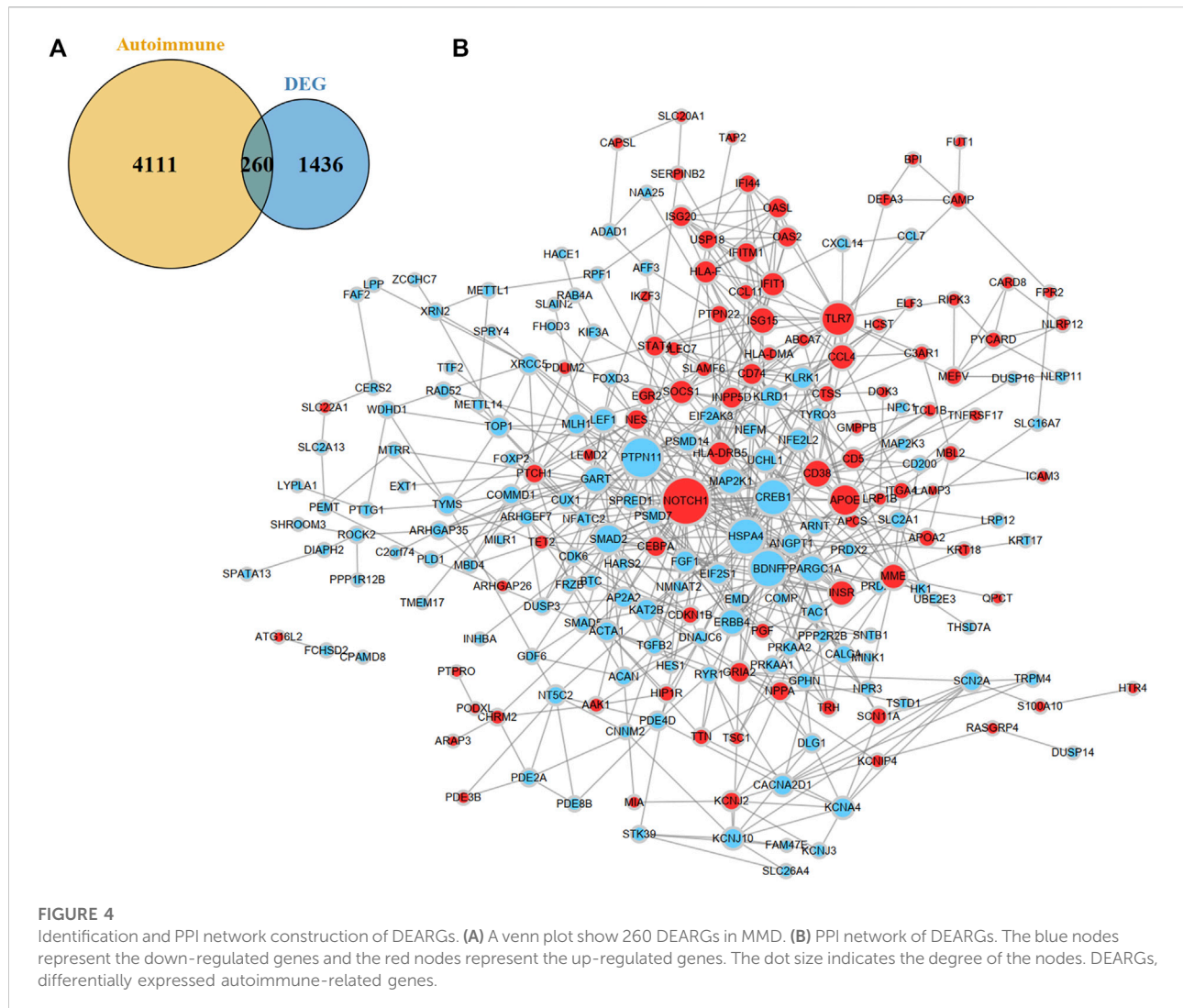
Construction of protein-protein interaction network of DEARGs

After combining the GAAD and DisGeNET databases, 4371 ARGs were obtained. We overlapped the DEGs and ARGs, resulting in 260 DEARGs in MMD (Figure 4A). To elucidate the molecules' functional associations, we imported these genes into the STRING database to construct a PPI network and further visualized it in Cytoscape. After removing nodes without interaction with other genes, a PPI network with 225 nodes and 602 edges was constructed (Figure 4B).

Functional enrichment analyses of DEARGs

For exploring the function and pathway of the DEARGs, we used ClueGO, a plugin of Cytoscape, and the "DOSE" package.

The biological processes-genes network showed that mononuclear cell migration and regulation of viral life cycle were most enriched for DEARGs (Figure 5A and Supplementary Figure S1A). The cAMP signaling pathway, PI3K-Akt signaling pathway, antigen processing and presentation, and microbial infection-related pathways were enriched in the KEGG analysis (Figure 5B and Supplementary Figure S1B). To uncover the diseases that these genes may be involved in, a DO analysis was conducted, which showed that these genes participated in the development of different cancers, infectious diseases, and autoimmune diseases.

Identification of tissue/organ-specific expressed genes

A total of 70 tissue/organ-specific expressed genes were identified for 260 DEARGs by BioGPS (Table 1). We observed

FIGURE 5
Biological annotations and DO analysis of DEARGs. **(A)** and **(B)** The enrichment network of biological processes and KEGG pathways with the participated genes in Cluego software. **(C)** The diseases of DEARGs involved through Disease Ontology analysis. DEARGs, differentially expressed autoimmune-related genes; DO, Disease Ontology.
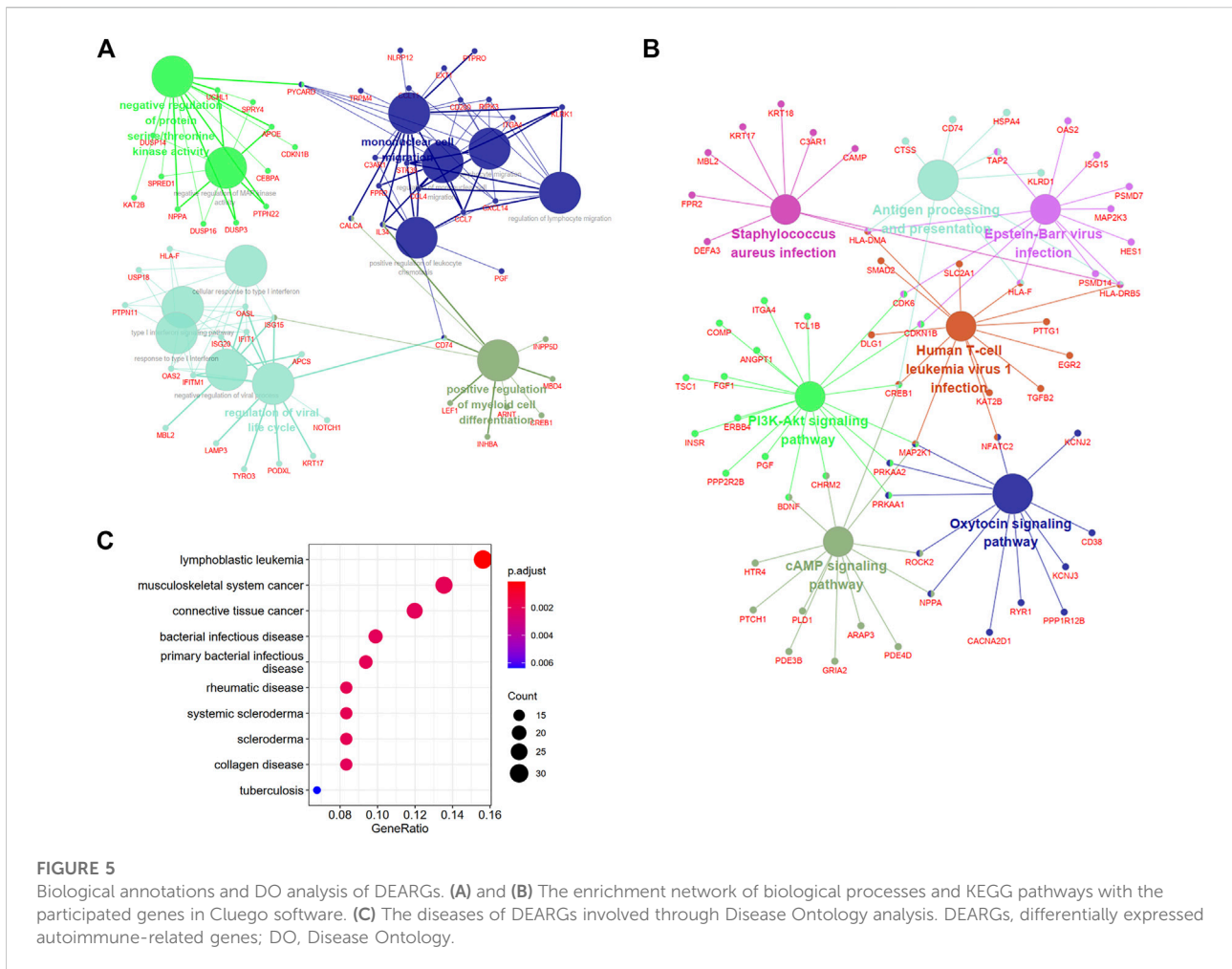
that the system with the greatest distribution of tissue-specific expressed genes was the hematologic/immune system (30/70, 42.90%). The second organ-specific expressed system was the nervous system, which included 11 genes (11/70, 15.70%), followed by the digestive system (7/70, 10.00%), endocrine system (5/70, 7.10%), and genital system (4/70, 5.70%).

Screening for crucial DEARGs and candidate signatures

First, we integrated five methods (Closeness, Degree, MCC, MNC, and Radiality) in cytoHubba and overlapped the top 30 genes from each method for robustness (Supplementary Figure S2 and Supplementary Table S2). A total of 15 shared genes were identified (Figure 6A). A co-expression network was constructed (Figure 6B), and 20 genes were identified that interacted with the 15 key

DEARGs. In the complex PPI network, the interaction of the physical interactions accounted for 38.45%, predicted for 35.98%, co-expression for 20.16%, and colocalization for 3.44%.

Second, we applied three machine learning algorithms (LASSO regression, RF, and SVM-RFE) to screen further the most important signatures. Seven genes were determined by LASSO regression with the optimal values (Figures 7A, B). We selected the top 10 genes ranked by the variable importance in RF (Figure 7C). The error fell to the lowest perigee, and the accuracy reached the peak when the number of features was set to 14 in the SVM-RFE method (Figures 7D, E). We identified six diagnostic genes by overlapping the three algorithms (Figure 7F): *CD38*, *PTPN11*, *NOTCH1*, *TLR7*, *KAT2B*, and *ISG15*. The detailed descriptions of the six diagnostic signatures are listed in Table 2. The correlation among these genes was studied by Pearson's method and presented with a heatmap (Supplementary Figure S3).

TABLE 1 70 Identified tissue/organ-specific expressed genes by BioGPS.

System/Organ	Counts	Frequency (%)	Genes
Haematologic/Immune cells	30	42.90	TNFRSF17, CTSS, FPR2, HLA-DMA, HLA-DRB5, ICAM3, ITGA4, KLRD1, STAT4, MBD4, CARD8, KLRK1, PTPN22, TLR7, CCL4, INPP5D, NLRP12, PYCARD, OAS2, ISG20, KCNJ2, FCHSD2, FUT1, TOP1, CBLN2, XRN2, C3AR1, ZCCHC7, SPATA13, MILR1
Nervous	11	15.70	FGF1, GRIA2, PPP2R2B, TAC1, UCHL1, PPP1R17, SERPINI1, DNAJC6, PDE2A, AAK1, NMNAT2
Digestive	7	10.00	MBL2, SLC22A1, PEMT, CEBPA, SPRY4, APCS, APOA2
Endocrine	5	7.10	CALCA, CA3, FAM47E, SLC26A4, SPINK1
Genital	4	5.70	PGF, PRKAA1, DIAPH2, ARHGAP35
Bonemarrow	3	4.30	BPI, CAMP, DEFA3
Smooth muscle	3	4.30	BDNF, CCL7, CCL11
Adipocyte	2	2.90	COMP, MME
Immune organs	2	2.90	LEF1, NLRP11
Heart	1	1.40	NPPA
Respiratory	1	1.40	LAMP3
Tongue/Skeletal Muscle	1	1.40	TTN

Construction and validation of the biological processes neural network diagnostic model
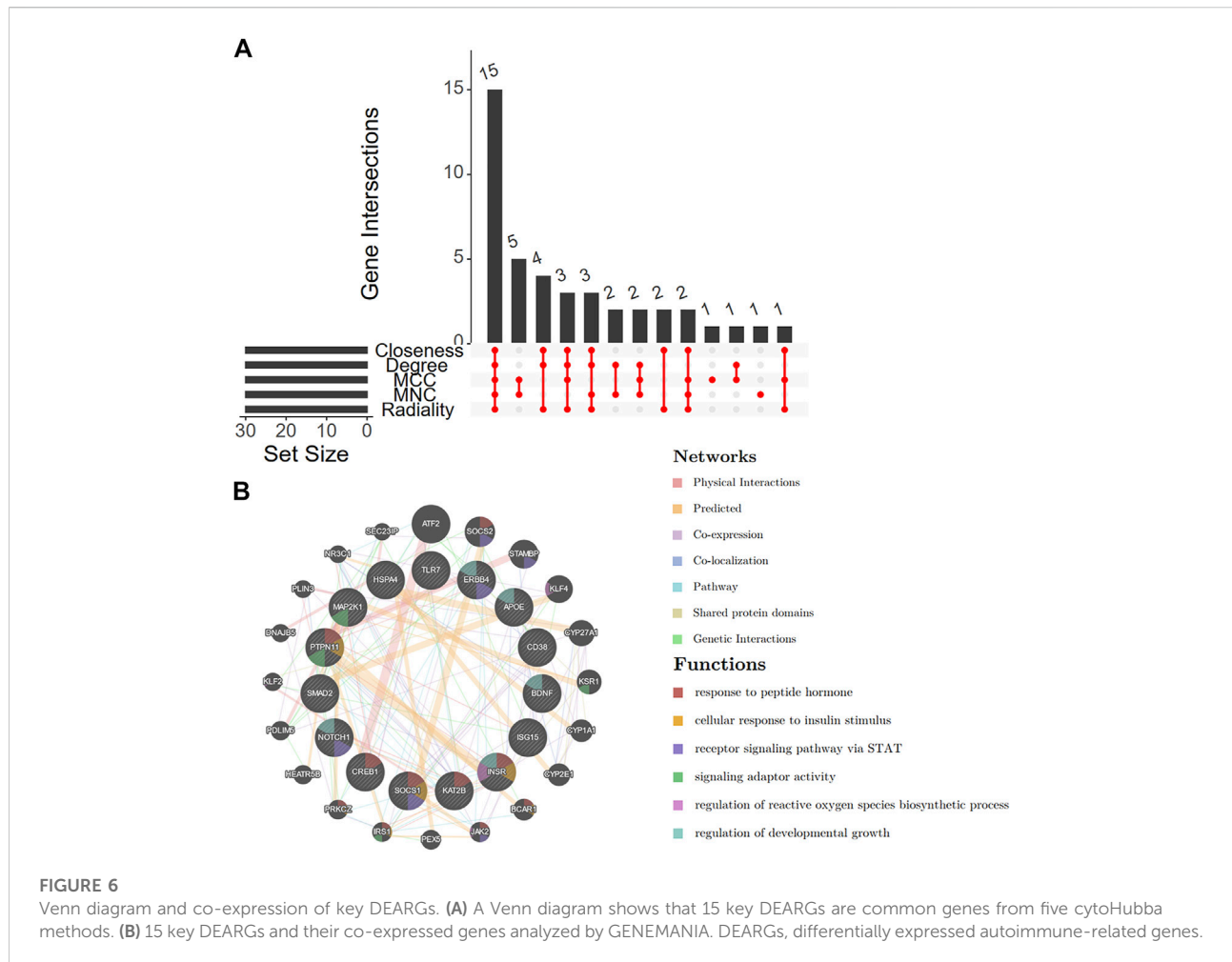
Based on the above screened six diagnostic signatures, we used the GSE157628 dataset as the training set to construct the back propagation artificial neural network model using the R package "neuralnet". After performing the preprocessing and scaling of this dataset, a neural network model with one hidden neuron layer was established. According to the output results of the neural network model (Figure 8A, Supplementary Table S3, and Supplementary Figure S4), the entire training was performed in 179 steps with an error of 0.021. Among the output results, the predicted weights of each hidden neuron layer were −14.758, −1.778, −5.728, and 13.268 for MMD. The confusion matrix and ROC curves show that the predicted scores (namely neuroMMD) present great diagnostic ability with AUC = 1 in the GSE157628 dataset (Figures 8B, C). The relative expression levels of these six diagnostic signatures are depicted through raincloud plots, indicating that CD38, NOTCH11, TLR7, and ISG15 were upregulated, and PTPN11 and KAT2B were downregulated (Figure 8D). We also recruited an external GSE141025 dataset that demonstrated the discriminatory performance of neuroMMD scores for MMD. The confusion matrix and ROC curves validated the great diagnostic ability of neuroMMD scores in MMD (Figures 8E, F). Although these genes from the verifying dataset showed the same expression trends as the training dataset, only CD38 and PTPN11 reached statistical significance (Figure 8G). Our RT-qPCR results also verified the same expression trends of

CD38 and PTPN11 as the GSE157628 and GSE141025 datasets, and we also found the significant up-regulation of NOTCH1 of MMD patients in our results (Supplementary Figure S5).

Identification of differential immune characteristics between moyamoya disease and controls

Based on xCell analysis results, we identified three immune cells (eosinophils, natural killer T cells, and Th2 cells) with significant infiltration differences between diseased tissues and controls. Among them, eosinophils and NKT were increased in diseased tissues, while Th2 cells were decreased. We also assessed the differences in immune activities and responses between MMD and controls using the ssGSEA algorithm (Figure 9A). Three immune activities (interleukins, interferon, and LCK molecules) showed significant differences between groups (Figure 9B). Seven differential expressed HLA molecules were also discovered (Figure 9C), and all were upregulated. The differential immune characteristics between MMD and controls are depicted through a heatmap (Figure 9D).

The correlation analysis between six hub DEARGs and differential immune characteristics showed that the upregulated genes were positively correlated with increasing immune characteristics and negatively with decreasing immune characteristics. In contrast, the downregulated genes showed inverse effects.

FIGURE 6
Venn diagram and co-expression of key DEARGs. **(A)** A Venn diagram shows that 15 key DEARGs are common genes from five cytoHubba methods. **(B)** 15 key DEARGs and their co-expressed genes analyzed by GENEMANIA. DEARGs, differentially expressed autoimmune-related genes.

Establishment of miRNA-gene and drug-gene regulatory networks

To identify the regulatory and therapeutic mechanisms for MMD, we predicted the miRNA and drugs targeting the six hub DEARGs through the Tarbase and DGIdb databases, respectively. A total of 131 miRNAs were found to potentially regulate the hub ARDEGs (Figure 10A). Genes with the most regulated miRNAs were identified as *KAT2B* (regulated by 56 miRNAs), followed by *PTPN11* (43 miRNAs), *NOTCH1* (35 miRNAs), *ISG15* (29 miRNAs), *CD38* (13 miRNAs), and *TLR7* (5 miRNAs) (Supplementary Table S4). We also discovered that four miRNAs (mir-26a-5p, mir-1343-3p, mir-129-2-3p, and mir-124-3p) interacted with four hub genes.

The DGIdb was applied to predict possible medicines or molecular compounds reacting with the hub ARDEGs. After excluding drugs/compounds for which the interaction types with genes were not clear, 13 potential target drugs/compounds remained for MMD treatment (Figure 10B and Supplementary Table S5). Of these, six drugs interacted with

TLR7; two targeted *CD38*, *KAT2B*, and *NOTCH1*; one drug targeted *ISG15*. *PTPN11* interacted with no potential drug.

Discussion

As the most common pediatric cerebrovascular disease in Japan, the incidence of frequently recurrent ischemic episodes (transient ischemic attacks or strokes) is 70%–80% in children with MMD (Currie et al., 2011). MMD cannot be effectively treated with pharmacological interventions alone; therefore, surgical procedures for revascularization (direct, indirect, and combined bypass) are required. There is much controversy surrounding the optimal treatment for this disorder (Jang et al., 2017; Deng et al., 2018). The incidence of perioperative complications such as stroke, hyperperfusion syndrome, and acute thrombogenesis is also concerning (Kim et al., 2016). It is of profound significance to understand the pathophysiologic processes of MMD and prevent the occurrence of this disorder. However, the molecular etiology of MMD remains unclear.
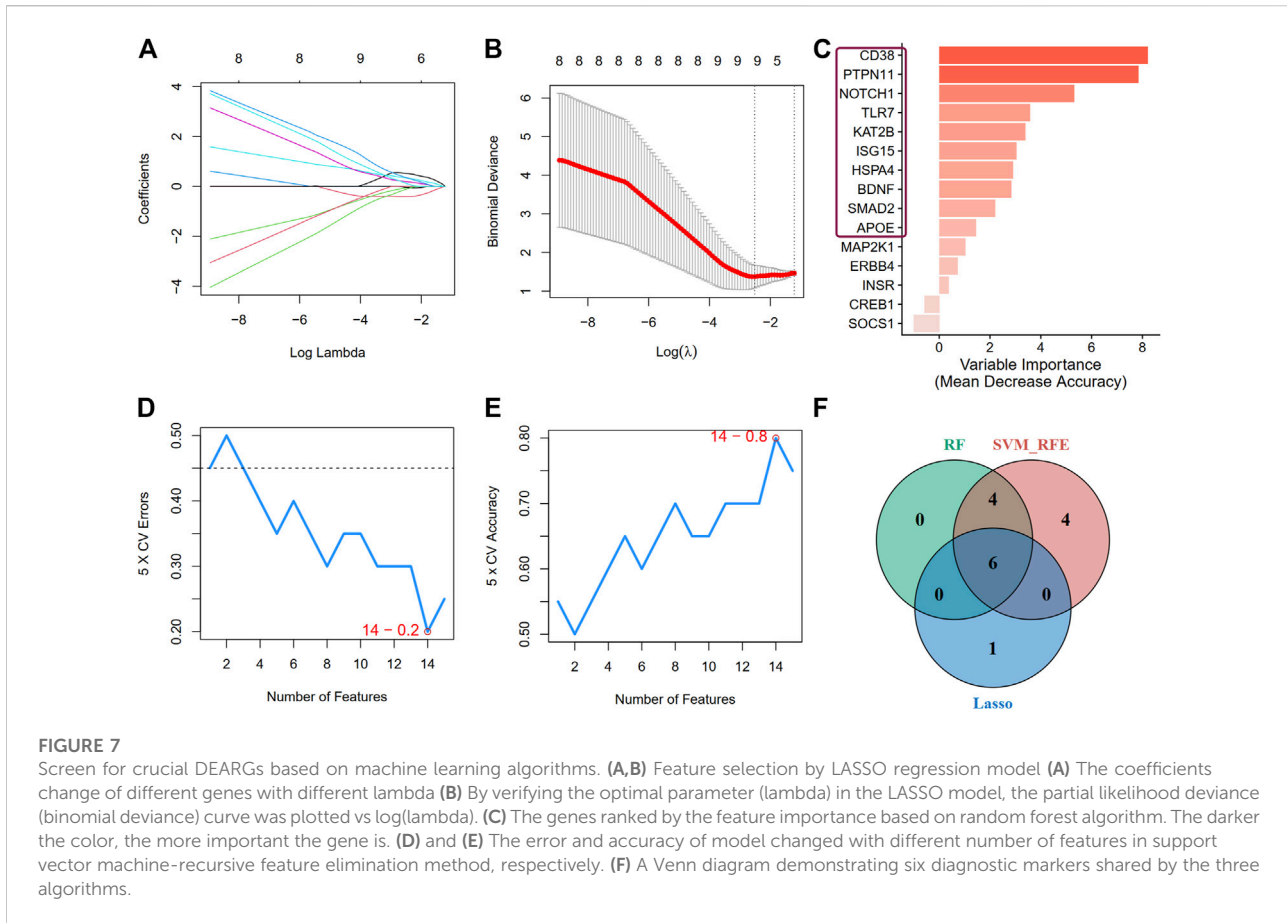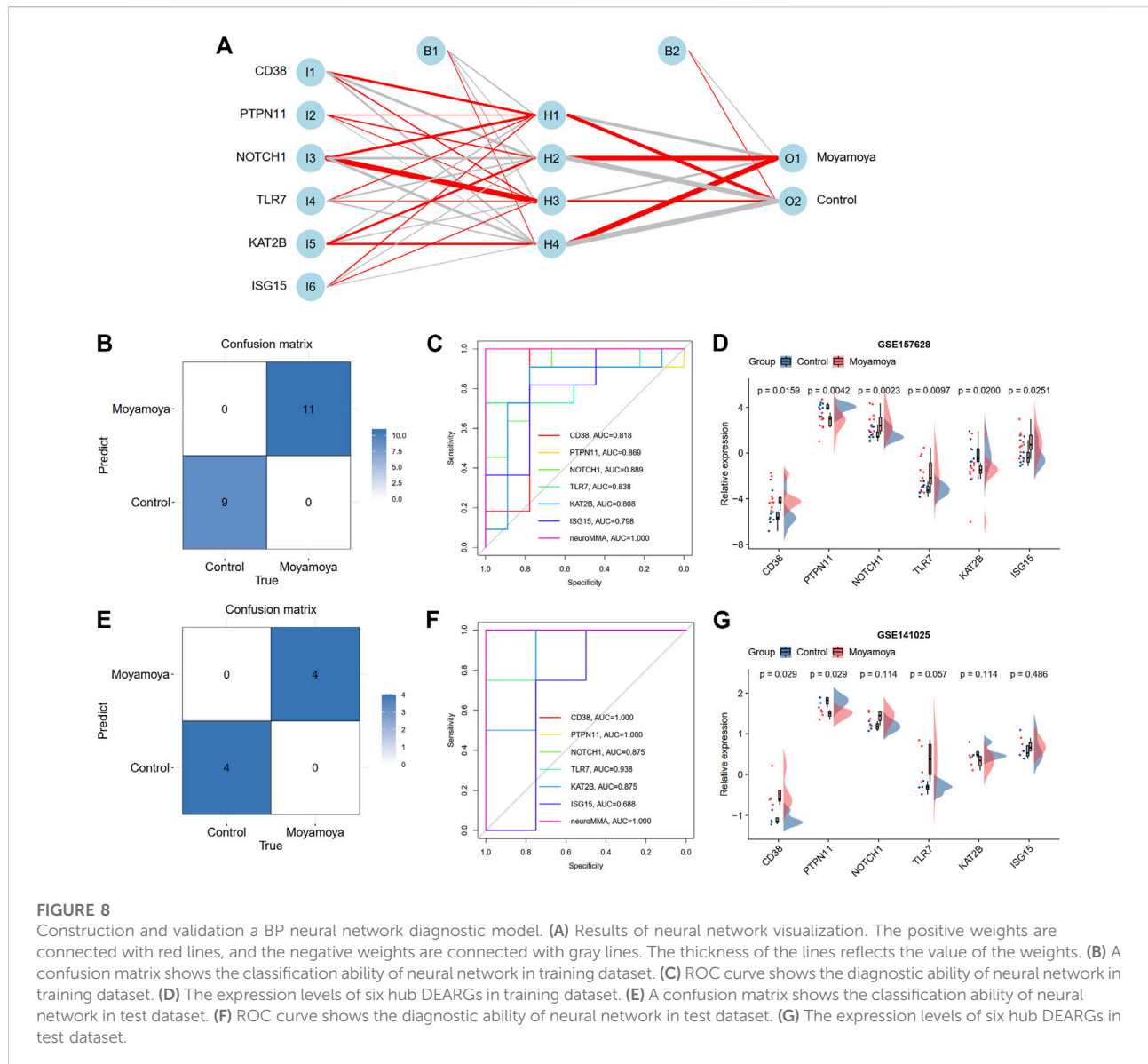
FIGURE 7
Screen for crucial DEARGs based on machine learning algorithms. **(A,B)** Feature selection by LASSO regression model **(A)** The coefficients change of different genes with different lambda **(B)** By verifying the optimal parameter (lambda) in the LASSO model, the partial likelihood deviance (binomial deviance) curve was plotted vs log(lambda). **(C)** The genes ranked by the feature importance based on random forest algorithm. The darker the color, the more important the gene is. **(D)** and **(E)** The error and accuracy of model changed with different number of features in support vector machine-recursive feature elimination method, respectively. **(F)** A Venn diagram demonstrating six diagnostic markers shared by the three algorithms.

TABLE 2 Detail information about the six hub genes identified by machine learning.

Gene	Description	Chromosome	logFC	P.Value	Change
CD38	CD38 molecule	4	1.257	0.029	UP
PTPN11	protein tyrosine phosphatase non-receptor type 11	12	−1.021	0.007	DOWN
NOTCH1	notch receptor 1	9	1.104	0.005	UP
TLR7	toll like receptor 7	X	1.417	0.006	UP
KAT2B	lysine acetyltransferase 2B	3	−1.435	0.046	DOWN
ISG15	ISG15 ubiquitin like modifier	1	1.072	0.023	UP

Previous studies have suggested the role of comorbidities as the link between autoimmune diseases and MMD (Kim et al., 2010; Chen et al., 2016). Therefore, analysis of ARGs may help determine the pathogenesis of MMD.

This study aimed to elucidate the critical processes and ARGs responsible for developing MMD by integrating bioinformatics and machine learning methods. Machine learning and artificial intelligence have become indispensable productivity tools for the 21st century for precision medicine. Machine learning and

artificial intelligence differ from traditional biomedical research because they use huge volumes of data to uncover natural laws, which are then applied to medical research. The field of bioinformatics involves the development of computational tools and approaches for acquiring, storing, visualizing, and interpreting medical or biological data. Combining machine learning and bioinformatics will facilitate the generation, analysis, maintenance, and interpretation of information derived from molecular genetics tests. Apart from

FIGURE 8
Construction and validation a BP neural network diagnostic model. **(A)** Results of neural network visualization. The positive weights are connected with red lines, and the negative weights are connected with gray lines. The thickness of the lines reflects the value of the weights. **(B)** A confusion matrix shows the classification ability of neural network in training dataset. **(C)** ROC curve shows the diagnostic ability of neural network in training dataset. **(D)** The expression levels of six hub DEARGs in training dataset. **(E)** A confusion matrix shows the classification ability of neural network in test dataset. **(F)** ROC curve shows the diagnostic ability of neural network in test dataset. **(G)** The expression levels of six hub DEARGs in test dataset.

the field of oncology research, the integrated approach is widely applied in cardiovascular diseases, such as myocardial infarction (Wu et al., 2022), heart failure (Tian et al., 2020), and aortic valve calcification (Xiong et al., 2022). Through machine learning and bioinformatics technology, our study revealed the differential expression of ARGs and their potential biological functions in MMD for the first time. A predictive model with hub ARGs was constructed by using an artificial neural network. Immune cell infiltration, immune activities, and HLA molecule expression were investigated in MMD. The elucidated correlations between ARGs and immune characteristics may help further explain the interplay of ARGs and the immune microenvironment in MMD.

In our study, a total of 1696 DEGs (750 upregulated genes and 946 downregulated genes) between MMD and controls were

screened. In biological function analysis, the pathways of autoimmune diseases (such as autoimmune thyroid disease, rheumatoid arthritis, the intestinal immune network for IgA production, and SLE), cell adhesion molecules (CAMs), and chylomicron remodeling were more enriched in MMD than controls. When overlapped with the ARGs, we further intersected 260 DEARGs involved in bacterial infectious disease, rheumatic disease, and collagen disease in DO analysis. Moyamoya vasculopathy in patients with the underlying causal condition is usually regarded as "MM syndrome". MMD concurrent with Graves' disease (GD) was first reported by Kushima et al. (1991). Over the past two decades, reports of these two concurrent diseases have increased (Tendler et al., 1997; Ni et al., 2014; Chen et al.,

FIGURE 9
The significantly differential immune characteristics between MMD and controls. (A) The significantly differential immune cells. (B) The significantly differential immune activities. (C) The significantly differential expressed HLA molecules. (D) A heat map shows the landscapes of immune characteristics between MMD and controls. (E) The correlation of immune characteristics and six hub DEARGs. Significance level was denoted by *p-value <0.05, **p-value <0.01, ***p-value <0.001.

FIGURE 10
The miRNA-Genes regulatory network (A) and Drug-Genes interactions (B).

2015). The associations between thyroid function and thyroid autoantibodies with MMD were also discovered (Kim et al., 2010). A case-control study found that compared with control subjects, the thyroid function and thyroid autoantibodies are elevated in pediatric patients with MMD (Li et al., 2011), further supported by another study without age stratification (Lei et al., 2014). Studies about SLE associated with MM syndrome have rarely been conducted and are published mostly in the form of case reports (El Ramahi and Al Rayes, 2000; Jeong et al., 2008). According to a recent review, MMD complicated with SLE mostly occurred in female patients [84.6% (11/13)], and most of these patients developed MMD by the age of 30 years (Tanaka et al., 2020). Among the 13 patients, 10 were from East Asian countries. Complications with rheumatoid arthritis and MMD are rare (Paciaroni et al., 2005). The common molecular characteristic of SLE and rheumatoid arthritis is the change in CAMs (da Rosa Franchi Santos et al., 2020), which was also identified by our enrichment analysis in MMD. CAMs, which are transmembrane proteins that facilitate cell-to-cell or cell-to-extracellular matrix binding, may be categorized into three different types named immunoglobulin supergene family members, selectins, and integrins. CAMs regulate the inflammatory response and endothelial function. Therefore, they may be targeted in cardiovascular disease (Kunutsor et al., 2017), such as atherosclerosis (Ling et al., 2012) and ischemic stroke (Yilmaz and Granger, 2008). DIAPH1 may be a novel MMD risk gene that impairs vascular cell actin remodeling that may cause neointimal expansion and progressive narrowing of the bilateral internal carotid arteries in MMD pathogenesis (Kundishora et al., 2021). Soriano et al. (2002). found significantly elevated levels of soluble CAMs in the cerebrospinal fluid of children with MM syndrome compared with the control group, suggesting the potential roles of CAMs in MMD. These results indicate the crucial association between MMD and autoimmune diseases. By examining the common pathogenesis of these disorders, we can clarify the etiology of MMD. CAMs may act as a bridge that triggers the common pathogenesis processes.

We also performed enrichment analysis of DEARGs in MMD, suggesting that infectious diseases (such as *Staphylococcus aureus* infection, Epstein-Barr virus infection, and *tuberculosis*), cAMP signaling pathway, and PI3K-Akt signaling pathway were involved. The infection hypothesis has occasionally been proposed as one of three mechanisms while investigating the pathogenesis of MMD, apart from autoimmune and HLA abnormality. Infections associated with MMD have been reported in many cases, including bacterial meningitis due to pneumococcus, tuberous infection, and viral infection by the varicella-zoster virus, Epstein-Barr virus, and *Leptospira* infection (Houkin et al., 2012). Czartoski proposed that the inflammation and subsequent post-infectious autoimmune response associated with meningitis can lead to a progressive vasculopathy, which may cause arterial occlusions in MM

syndrome after autopsy in a patient with pneumococcal meningitis (Czartoski et al., 2005). Despite suggesting a possible infectious cause in MMD, these results were only based on case studies, and no specific pathogen has been identified. A large-sample study is indispensable to finding the relationship between infections and MMD. The DEARGs we identified as associated with these diseases may provide a molecular-level explanation.

A PPI network was constructed based on the DEARGs to explore relationships among proteins. We found 15 potential genes overlapping the top 30 genes identified by six algorithms (Closeness, Degree, MCC, MNC, and Radiality) in cytoHubba. To further screen out the hub genes, three machine learning methods (LASSO regression, RF, and SVM-RFE) were applied, and six genes (*CD38*, *PTPN11*, *NOTCH1*, *TLR7*, *KAT2B*, and *ISG15*) were selected to construct a predictive model using BP artificial neural network in the GSE157628 dataset. The GSE141025 dataset also verified the predictive performance with great diagnostic ability, which proved the applicability of our model. Type I interferons (IFNs) induce the expression of over 500 genes, collectively referred to as IFN-stimulated genes (ISGs). ISG15 is a ubiquitin-like protein that can conjugate to substrate proteins (ISGylation) in response to microbial infection. This IFN-α/β-inducible ISG15 does not serve as a substrate for ISGylation-based antiviral immunity but for regulating IFN-α/β by USP18 and preventing IFN-α/β-dependent auto-inflammation (Zhang et al., 2015). The antiviral and antineoplastic roles of ISG15 have been extensively studied (Mustachio et al., 2018; Perng and Lenschow, 2018). RNF213 is an interferon-induced mega protein frequently mutated in MMD as a susceptibility gene (Liu et al., 2011). A recent study pointed out that RNF213, an ISG15 interactor, can act as a sensor for ISGylated proteins to counteract infection (Thery et al., 2021). In our immune infiltration analysis, we observed that the activity of IFNs and the expression of ISG15 genes were higher in MMD than in controls. Therefore, the overexpression of ISG15 induced by IFNs may be involved in the pathogenesis of MMD through its interaction with RNF213. This finding may provide a new direction for basic experiments in the future. *PTPN11*, the gene encoding the protein tyrosine phosphatase SHP2, is a ubiquitously expressed non-receptor tyrosine phosphatase that regulates cell survival, proliferation, differentiation, migration, and adhesion. Germline mutations in *PTPN11* cause Noonan syndrome, the clinically related LEOPARD syndrome (LS), and leukemogenesis (Tartaglia et al., 2006; Alfayez et al., 2021). Seki et al. found that the expression of SHP2 was markedly elevated in the thickened aortic intima in rats with balloon-induced injury in an atherosclerosis animal model (Seki et al., 2002). The inhibition of SHP2 can protect against atherosclerosis by inhibiting smooth muscle cell proliferation (Chen et al., 2018). The most prominent pathological change in MMD is the inner elastic lamellar's breakage and smooth muscle cells' destruction and

proliferation in the tunica media (Huang et al., 2017). There are some common pathogenesis links between MMD and atherosclerosis. MMD susceptibility variant RNF213 p. R4810K can increase the risk of recurrent cerebrovascular events, such as ischemic stroke caused by large-artery atherosclerosis (Okazaki et al., 2019; Kim et al., 2021). Considering no direct evidence connecting *CD38*, *PTPN11*, *NOTCH1*, *TLR7*, and *KAT2B* with the pathogenesis of MMD, their roles in atherosclerosis may also provide a new perspective and direction for future research on the molecular targeted therapy of MMD (Salagianni et al., 2012; Briot et al., 2015; Xu et al., 2016; Qi et al., 2021).

Considering the important roles of immune activities in MMD, we also studied the immune characteristics from the perspective of immune cell infiltration, activities of immune responses, and HLA molecule expression. Our results showed that the abundance of eosinophils and natural killer T (NKT) cells is significantly elevated while Th2 cells were decreased in MMD compared to controls. The expression levels of HLA-A, HLA-B, HLA-C, HLA-DMA, HLA-DRB6, HLA-F, and HLA-G were significantly upregulated in MMD. The abnormality of HLA is considered one of the molecular mechanisms leading to the occurrence of MMD. Hong et al. (2009) found that the phenotype frequencies of HLA-DRB1(*) 1302 and DQB1(*) 0609 were significantly increased in familial MMD compared to both controls and non-familial Korean patients with MMD. In a Japanese case-control study on MMD, the HLA-DRB1*04: 10 allele was found to be a predisposing genetic factor, and the frequency of autoimmune thyroid diseases was increased in HLA-DRB1*04:10-positive patients with MMD compared with that in HLA-DRB1*04:10-negative patients with MMD (Tashiro et al., 2019). Recent research in Chinese Han population indicated that the genetic polymorphism of HLA-DQA2 and HLA-B was identified as a risk factor for MMD (Wan et al., 2021).

Gene-miRNAs modify the appearance of proteins with the progression of diseases by targeting their main targets. In this study, we also constructed a gene-miRNA regulatory network, and four miRNAs (mir-26a-5p, mir-1343-3p, mir-129-2-3p, and mir-124-3p) were identified because of their interaction at least with four hub DEARGs. Mir-26a-5p can alleviate cardiac hypertrophy and dysfunction (Shi et al., 2021) and protect against myocardial ischemia/reperfusion injury (Xing et al., 2020). Mir-1343-3p plays a significant role in the development of human cancers such as lung cancer (Zhang et al., 2020), colorectal carcinoma (Bhat et al., 2022), and hepatocellular carcinoma (Mou and Sun, 2022). The involvement of mir-1343-3p in cardiovascular diseases is also reported (Sharma et al., 2020). The potential role of mir-129-2-3p in ischemic stroke was identified by suppressing *SYK* gene expression (Huang et al., 2019). Mir-124-3p contributes to the development of different cardiovascular diseases, such as atherosclerosis, myocardial infarction, and ischemic stroke (de

Ronde et al., 2017; Badacz et al., 2021). These four miRNAs may be used as interventional targets for examining the mechanisms of ARGs in MMD since they interact with at least four hub DEARGs. Moreover, a total of 13 potential drugs/compounds were predicted for MMD treatment by targeting the hub ARGs in our study.

Further studies are warranted to address some limitations of the present study. First, although the diagnostic model constructed by an artificial neural network performed well in the training and testing datasets, the sample size was very small, especially for the validation dataset, where only four samples were available in each group. Therefore, studies with larger sample sizes are essential. Second, in the GSE157628 dataset, six control samples were collected from the MCA of patients with ICA aneurysms. Considering the different hemodynamic and genetic effects, the normal artery of patients with aneurysms may differ from normal vessels at the transcriptional level. However, the collection of normal vessels from healthy control, in essence, is against medical ethics. Setting the normal artery from patients with aneurysms as the control group in studying MMD is acceptable (Kanamori et al., 2021). Third, the results were based on bioinformatics and conducted RT-qPCR, but *in vitro* and *in vivo* experiments should be conducted to verify the results.

Conclusion

In our analysis, a total of 260 DEARGs were identified in MMD, which were involved in autoimmune-related diseases and immune responses. Six ARGs (*CD38*, *PTPN11*, *NOTCH1*, *TLR7*, *KAT2B*, and *ISG15*) were selected by three machine learning methods (LASSO regression, RF, and SVM-RFE). They were finally used to construct a predictive model using BP artificial neural network that could be used to identify patients with MMD. Immune infiltration analysis showed that immune activities and HLA expression levels in MMD were enhanced. Finally, a gene-miRNA network was prepared, and pharmacological agents targeting hub genes were predicted as potentially effective in treating MMD.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author.

Author contributions

Conceptualization, SL and YH; methodology, SL; software, SL; validation, YH; formal analysis, SL; investigation, SL; writing—original draft preparation, SL and YH;

writing—review and editing, DT and QZ; visualization, YH and JL; supervision, LW.

Funding

Conflict of interest

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.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmolb. 2022.991425/full#supplementary-material

References

Alfayez, M., Issa, G. C., Patel, K. P., Wang, F., Wang, X., Short, N. J., et al. (2021). The Clinical impact of PTPN11 mutations in adults with acute myeloid leukemia. *Leukemia* 35, 691–700. doi:10.1038/s41375-020-0920-z

Aran, D., Hu, Z., and Butte, A. J. (2017). xCell: digitally portraying the tissue cellular heterogeneity landscape. *Genome Biol.* 18, 220. doi:10.1186/s13059-017-1349-1

Badacz, R., Kleczyński, P., Legutko, J., Żmudka, K., Gacoń, J., Przewłocki, T., et al. (2021). Expression of miR-1-3p, miR-16-5p and miR-122-5p as possible risk factors of secondary cardiovascular events. *Biomedicines* 9, 1055. doi:10.3390/biomedicines9081055

Bhat, I. P., Rather, T. B., Bhat, G. A., Maqbool, I., Akhtar, K., Rashid, G., et al. (2022). TEAD4 nuclear localization and regulation by miR-4269 and miR-1343-3p in colorectal carcinoma. *Pathol. Res. Pract.* 231, 153791. doi:10.1016/j.prp.2022.153791

Briot, A., Civelek, M., Seki, A., Hoi, K., Mack, J. J., Lee, S. D., et al. (2015). Endothelial NOTCH1 is suppressed by circulating lipids and antagonizes inflammation during atherosclerosis. *J. Exp. Med.* 212, 2147–2163. doi:10.1084/jem.20150603

Chen, J. B., Lei, D., He, M., Sun, H., Liu, Y., Zhang, H., et al. (2015). Clinical features and disease progression in moyamoya disease patients with Graves disease. *J. Neurosurg.* 123, 848–855. doi:10.3171/2014.10.Jns141140

Chen, J. B., Liu, Y., Zhou, L. X., Sun, H., He, M., and You, C. (2016). Increased prevalence of autoimmune disease in patients with unilateral compared with bilateral moyamoya disease. *J. Neurosurg.* 124, 1215–1220. doi:10.3171/2015.4.Jns142936

Chen, J., Cao, Z., and Guan, J. (2018). SHP2 inhibitor PHPS1 protects against atherosclerosis by inhibiting smooth muscle cell proliferation. *BMC Cardiovasc. Disord.* 18, 72. doi:10.1186/s12872-018-0816-2

Chen, B., Tao, W., Li, S., Zeng, M., Zhang, L., Huang, Z., et al. (2022). Medial gap: a structural factor at the arterial bifurcation aggravating hemodynamic insult. *J. Neuropathol. Exp. Neurol.* 81, 282–290. doi:10.1093/jnen/nlac017

Chin, C. H., Chen, S. H., Wu, H. H., Ho, C. W., Ko, M. T., and Lin, C. Y. (2014). cytoHubba: identifying hub objects and sub-networks from complex interactome. *BMC Syst. Biol.* 8, S11. doi:10.1186/1752-0509-8-s4-s11

Currie, S., Raghavan, A., Batty, R., Connolly, D. J., and Griffiths, P. D. (2011). Childhood moyamoya disease and moyamoya syndrome: a pictorial review. *Pediatr. Neurol.* 44, 401–413. doi:10.1016/j.pediatrneurol.2011.02.007

Czartoski, T., Hallam, D., Lacy, J. M., Chun, M. R., and Becker, K. (2005). Postinfectious vasculopathy with evolution to moyamoya syndrome. *J. Neurol. Neurosurg. Psychiatry* 76, 256–259. doi:10.1136/jnnp.2004.041046

da Rosa Franchi Santos, L. F., Costa, N. T., Maes, M., Simão, A. N. C., and Dichi, I. (2020). Influence of treatments on cell adhesion molecules in patients with systemic lupus erythematosus and rheumatoid arthritis: a review. *Inflammopharmacology* 28, 363–384. doi:10.1007/s10787-019-00674-6

de Ronde, M. W. J., Kok, M. G. M., Moerland, P. D., Van den Bossche, J., Neele, A. E., Halliani, A., et al. (2017). High miR-124-3p expression identifies smoking individuals susceptible to atherosclerosis. *Atherosclerosis* 263, 377–384. doi:10.1016/j.atherosclerosis.2017.03.045

Deng, X., Gao, F., Zhang, D., Zhang, Y., Wang, R., Wang, S., et al. (2018). Effects of different surgical modalities on the clinical outcome of patients with moyamoya disease: a prospective cohort study. *J. Neurosurg.* 128, 1327–1337. doi:10.3171/2016.12.Jns162626

El Ramahi, K. M., and Al Rayes, H. M. (2000). Systemic lupus erythematosus associated with moyamoya syndrome. *Lupus* 9, 632–636. doi:10.1191/096120300678828686

Hong, S. H., Wang, K. C., Kim, S. K., Cho, B. K., and Park, M. H. (2009). Association of HLA-DR and -DQ genes with familial moyamoya disease in Koreans. *J. Korean Neurosurg. Soc.* 46, 558–563. doi:10.3340/jkns.2009.46.6.558

Houkin, K., Ito, M., Sugiyama, T., Shichinohe, H., Nakayama, N., Kazumata, K., et al. (2012). Review of past research and current concepts on the etiology of moyamoya disease. *Neurol. Med. Chir.* 52, 267–277. doi:10.2176/nmc.52.267

Huang, S., Guo, Z. N., Shi, M., Yang, Y., and Rao, M. (2017). Etiology and pathogenesis of Moyamoya Disease: An update on disease prevalence. *Int. J. Stroke* 12, 246–253. doi:10.1177/1747493017694393

Huang, S., Lv, Z., Wen, Y., Wei, Y., Zhou, L., Ke, Y., et al. (2019). miR-129-2-3p directly targets SYK gene and associates with the risk of ischaemic stroke in a Chinese population. *J. Cell. Mol. Med.* 23, 167–176. doi:10.1111/jcmm.13901

Jang, D. K., Lee, K. S., Rha, H. K., Huh, P. W., Yang, J. H., Park, I. S., et al. (2017). Bypass surgery versus medical treatment for symptomatic moyamoya disease in adults. *J. Neurosurg.* 127, 492–502. doi:10.3171/2016.8.Jns152875

Jeong, H. C., Kim, Y. J., Yoon, W., Joo, S. P., Lee, S. S., and Park, Y. W. (2008). Moyamoya syndrome associated with systemic lupus erythematosus. *Lupus* 17, 679–682. doi:10.1177/0961203307087375

Jin, F., and Duan, C. (2022). Identification of immune-infiltrated hub genes as potential biomarkers of Moyamoya disease by bioinformatics analysis. *Orphanet J. Rare Dis.* 17, 80. doi:10.1186/s13023-022-02238-4

Kanamori, F., Yokoyama, K., Ota, A., Yoshikawa, K., Karnan, S., Maruwaka, M., et al. (2021). Transcriptome-wide analysis of intracranial artery in patients with moyamoya disease showing upregulation of immune response, and downregulation of oxidative phosphorylation and DNA repair. *Neurosurg. Focus* 51, E3. doi:10.3171/2021.6.Focus20870

Karagkouni, D., Paraskevopoulou, M. D., Chatzopoulos, S., Vlachos, I. S., Tastsoglou, S., Kanellos, I., et al. (2018). DIANA-TarBase v8: a decade-long collection of experimentally supported miRNA-gene interactions. *Nucleic Acids Res.* 46, D239–d45. doi:10.1093/nar/gkx1141

Kim, S. J., Heo, K. G., Shin, H. Y., Bang, O. Y., Kim, G. M., Chung, C. S., et al. (2010). Association of thyroid autoantibodies with moyamoya-type cerebrovascular disease: a prospective study. *Stroke* 41, 173–176. doi:10.1161/strokeaha.109.562264

Kim, T., Oh, C. W., Bang, J. S., Kim, J. E., and Cho, W. S. (2016). Moyamoya disease: Treatment and outcomes. *J. Stroke* 18, 21–30. doi:10.5853/jos.2015.01739

Kim, H. J., Choi, E. H., Chung, J. W., Kim, J. H., Kim, Y. S., Seo, W. K., et al. (2021). Role of the RNF213 variant in vascular outcomes in patients with intracranial atherosclerosis. *J. Am. Heart Assoc.* 10, e017660. doi:10.1161/jaha.120.017660

Kundishora, A. J., Peters, S. T., Pinard, A., Duran, D., Panchagnula, S., Barak, T., et al. (2021). DIAPH1 variants in non-east Asian patients with sporadic moyamoya disease. *JAMA Neurol.* 78, 993–1003. doi:10.1001/jamaneurol.2021.1681

Kunutsor, S. K., Bakker, S. J. L., and Dullaart, R. P. F. (2017). Soluble vascular cell adhesion molecules may be protective of future cardiovascular disease risk: Findings from the PREVEND prospective cohort study. *J. Atheroscler. Thromb.* 24, 804–818. doi:10.5551/jat.38836

Kuroda, S., and Houkin, K. (2008). Moyamoya disease: current concepts and future perspectives. *Lancet. Neurol.* 7, 1056–1066. doi:10.1016/s1474-4422(08)70240-0

Kushima, K., Satoh, Y., Ban, Y., TaniyaMaM.Ito, K., and Sugita, K. (1991). Graves' thyrotoxicosis and Moyamoya disease. *Can. J. Neurol. Sci.* 18, 140–142. doi:10.1017/s0317167100031607

Lei, C., Wu, B., Ma, Z., Zhang, S., and Liu, M. (2014). Association of moyamoya disease with thyroid autoantibodies and thyroid function: a case-control study and meta-analysis. *Eur. J. Neurol.* 21, 996–1001. doi:10.1111/ene.12425

Li, H., Zhang, Z. S., Dong, Z. N., Ma, M. J., Yang, W. Z., Han, C., et al. (2011). Increased thyroid function and elevated thyroid autoantibodies in pediatric patients with moyamoya disease: a case-control study. *Stroke* 42, 1138–1139. doi:10.1161/strokeaha.110.608471

Li, S., Zhang, Q., Chen, Z., Huang, Z., and Chen, F. (2022). Novel insight into ferroptosis-related genes, molecular subtypes, and immune characteristics in intracranial aneurysms. *Inflamm. Res.* 71, 1347–1364. doi:10.1007/s00011-022-01633-8

Li, S., Zhang, Q., Weng, L., Han, Y., and Li, J. (2022). Novel insight into m6A regulator-mediated methylation modification patterns and immune characteristics in intracranial aneurysm. *Front. Aging Neurosci.* 14, 973258. doi:10.3389/fnagi.2022.973258

Liberzon, A., Birger, C., Thorvaldsdóttir, H., Ghandi, M., Mesirov, J. P., and Tamayo, P. (2015). The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst.* 1, 417–425. doi:10.1016/j.cels.2015.12.004

Ling, S., Nheu, L., and Komesaroff, P. A. (2012). Cell adhesion molecules as pharmaceutical target in atherosclerosis. *Mini Rev. Med. Chem.* 12, 175–183. doi:10.2174/138955712798995057

Liu, W., Morito, D., Takashima, S., Mineharu, Y., Kobayashi, H., Hitomi, T., et al. (2011). Identification of RNF213 as a susceptibility gene for moyamoya disease and its possible role in vascular development. *PloS one* 6, e22542. doi:10.1371/journal.pone.0022542

Lu, G., Hao, X., Chen, W. H., and Mu, S. (2018). GAAD: A gene and autoimmiune disease association database. *Genomics Proteomics Bioinforma.* 16, 252–261. doi:10.1016/j.gpb.2018.05.001

Mou, Y., and Sun, Q. (2022). The long non-coding RNA ASMTL-AS1 promotes hepatocellular carcinoma progression by sponging miR-1343-3p that suppresses LAMC1 (laminin subunit gamma 1). *Bioengineered* 13, 746–758. doi:10.1080/21655979.2021.2012628

Mustachio, L. M., Lu, Y., Kawakami, M., Roszik, J., Freemantle, S. J., Liu, X., et al. (2018). Evidence for the ISG15-specific deubiquitinase USP18 as an antineoplastic target. *Cancer Res.* 78, 587–592. doi:10.1158/0008-5472.Can-17-1752

Ni, J., Zhou, L. X., Wei, Y. P., Li, M. L., Xu, W. H., Gao, S., et al. (2014). Moyamoya syndrome associated with Graves' disease: a case series study. *Ann. Transl. Med.* 2, 77. doi:10.3978/j.issn.2305-5839.2014.08.03

Okazaki, S., Morimoto, T., Kamatani, Y., Kamimura, T., Kobayashi, H., Harada, K., et al. (2019). Moyamoya disease susceptibility variant RNF213 p.R4810K increases the risk of ischemic stroke attributable to large-artery atherosclerosis. *Circulation* 139, 295–298. doi:10.1161/circulationaha.118.038439

Paciaroni, M., Micheli, S., Caso, V., Venti, M., Alberti, A., Milia, P., et al. (2005). Angiographic findings of moyamoya vessels in a patient with rheumatoid arthritis. *Cerebrovasc. Dis.* 20, 415–416. doi:10.1159/000088666

Perng, Y. C., and Lenschow, D. J. (2018). ISG15 in antiviral immunity and beyond. *Nat. Rev. Microbiol.* 16, 423–439. doi:10.1038/s41579-018-0020-5

Piñero, J., Ramírez-Anguita, J. M., Saüch-Pitarch, J., Ronzano, F., Centeno, E., Sanz, F., et al. (2020). The DisGeNET knowledge platform for disease genomics: 2019 update. *Nucleic Acids Res.* 48, D845–D855. doi:10.1093/nar/gkz1021

Qi, X., Wang, H., Xia, L., Lin, R., Li, T., Guan, C., et al. (2021). miR-30b-5p releases HMGB1 via UBE2D2/KAT2B/HMGB1 pathway to promote pro-inflammatory polarization and recruitment of macrophages. *Atherosclerosis* 324, 38–45. doi:10.1016/j.atherosclerosis.2021.02.016

Salagianni, M., Galani, I. E., Lundberg, A. M., Davos, C. H., Varela, A., Gavriil, A., et al. (2012). Toll-like receptor 7 protects from atherosclerosis by constraining "inflammatory" macrophage activation. *Circulation* 126, 952–962. doi:10.1161/circulationaha.111.067678

Scott, R. M., and Smith, E. R. (2009). Moyamoya disease and moyamoya syndrome. *N. Engl. J. Med.* 360, 1226–1237. doi:10.1056/NEJMra0804622

Seki, N., Hashimoto, N., Suzuki, Y., Mori, S., Amano, K., and Saito, Y. (2002). Role of SRC homology 2-containing tyrosine phosphatase 2 on proliferation of rat smooth muscle cells. *Arterioscler. Thromb. Vasc. Biol.* 22, 1081–1085. doi:10.1161/01.atv.0000022878.37277.ec

Sharma, A. R., Vohra, M., Shukla, V., Guddattu, V., Razak Uk, A., Shetty, R., et al. (2020). Coding SNPs in hsa-miR-1343-3p and hsa-miR-6783-3p target sites of CYP2C19 modulates clopidogrel response in individuals with cardiovascular diseases. *Life Sci.* 245, 117364. doi:10.1016/j.lfs.2020.117364

Shi, H., Li, H., Zhang, F., Xue, H., Zhang, Y., and Han, Q. (2021). MiR-26a-5p alleviates cardiac hypertrophy and dysfunction via targeting ADAM17. *Cell Biol. Int.* 45, 2357–2367. doi:10.1002/cbin.11685

Soriano, S. G., Cowan, D. B., Proctor, M. R., and Scott, R. M. (2002). Levels of soluble adhesion molecules are elevated in the cerebrospinal fluid of children with moyamoya syndrome. *Neurosurgery* 50, 544–549. doi:10.1097/00006123-200203000-00022

Suzuki, J., and Takaku, A. (1969). Cerebrovascular "moyamoya" disease. Disease showing abnormal net-like vessels in base of brain. *Arch. Neurol.* 20, 288–299. doi:10.1001/archneur.1969.00480090076012

Tanaka, R., Shimojima, Y., Ueno, K. I., Takasone, K., Ichikawa, T., Kishida, D., et al. (2020). Moyamoya syndrome related to systemic lupus erythematosus developing during pregnancy: a case-based review. *Clin. Rheumatol.* 39, 3861–3867. doi:10.1007/s10067-020-05246-6

Tartaglia, M., Martinelli, S., Stella, L., Bocchinfuso, G., Flex, E., Cordeddu, V., et al. (2006). Diversity and functional consequences of germline and somatic PTPN11 mutations in human disease. *Am. J. Hum. Genet.* 78, 279–290. doi:10.1086/499925

Tashiro, R., Niizuma, K., Khor, S. S., Tokunaga, K., Fujimura, M., Sakata, H., et al. (2019). Identification of HLA-DRB1*04:10 allele as risk allele for Japanese moyamoya disease and its association with autoimmune thyroid disease: A case-control study. *PloS one* 14, e0220858. doi:10.1371/journal.pone.0220858

Tendler, B. E., Shoukri, K., Malchoff, C., MacGillivray, D., DuckRow, R., Talmadge, T., et al. (1997). Concurrence of Graves' disease and dysplastic cerebral blood vessels of the moyamoya variety. *Thyroid* 7, 625–629. doi:10.1089/thy.1997.7.625

Thery, F., Martina, L., Asselman, C., Zhang, Y., Vessely, M., Repo, H., et al. (2021). Ring finger protein 213 assembles into a sensor for ISGylated proteins with antimicrobial activity. *Nat. Commun.* 12, 5772. doi:10.1038/s41467-021-26061-w

Tian, Y., Yang, J., Lan, M., and Zou, T. (2020). Construction and analysis of a joint diagnosis model of random forest and artificial neural network for heart failure. *Aging* 12, 26221–26235. doi:10.18632/aging.202405

Wan, J., Ling, W., Zhengshan, Z., Xianbo, Z., Lian, D., and Kai, W. (2021). Association of HLA-DQA2 and HLA-B with moyamoya disease in the Chinese han population. *Neurol. Genet.* 7, e592. doi:10.1212/nxg.0000000000000592

Wang, H., Zhu, H., Zhu, W., Xu, Y., Wang, N., Han, B., et al. (2020). Bioinformatic analysis identifies potential key genes in the pathogenesis of turner syndrome. *Front. Endocrinol.* 11, 104. doi:10.3389/fendo.2020.00104

Warde-Farley, D., Donaldson, S. L., Comes, O., Zuberi, K., Badrawi, R., Chao, P., et al. (2010). The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function. *Nucleic Acids Res.* 38, W214–W220. doi:10.1093/nar/gkq537

Wu, C., Orozco, C., Boyer, J., Leglise, M., Goodale, J., Batalov, S., et al. (2009). BioGPS: an extensible and customizable portal for querying and organizing gene annotation resources. *Genome Biol.* 10, R130. doi:10.1186/gb-2009-10-11-r130

Wu, Y., Chen, H., Li, L., Zhang, L., Dai, K., Wen, T., et al. (2022). Construction of novel gene signature-based predictive model for the diagnosis of acute myocardial infarction by combining random forest with artificial neural network. *Front. Cardiovasc. Med.* 9, 876543. doi:10.3389/fcvm. 2022.876543

Xing, X., Guo, S., Zhang, G., Liu, Y., Wang, X., Lu, Q., et al. (2020). miR-26a-5p protects against myocardial ischemia/reperfusion injury by regulating the PTEN/PI3K/AKT signaling pathway. *Braz. J. Med. Biol. Res.* 53, e9106. doi:10.1590/1414-431x20199106

Xiong, T., Han, S., Pu, L., Zhang, T. C., Zhan, X., Fu, T., et al. (2022). Bioinformatics and machine learning methods to identify FN1 as a novel biomarker of aortic valve calcification. *Front. Cardiovasc. Med.* 9, 832591. doi:10.3389/fcvm.2022.832591

Xu, X., Yuan, X., Li, N., Dewey, W. L., Li, P. L., and Zhang, F. (2016). Lysosomal cholesterol accumulation in macrophages leading to coronary

atherosclerosis in CD38(-/-) mice. *J. Cell. Mol. Med.* 20, 1001–1013. doi:10. 1111/jcmm.12788

Yilmaz, G., and Granger, D. N. (2008). Cell adhesion molecules and ischemic stroke. *Neurol. Res.* 30, 783–793. doi:10.1179/174313208x341085

Yu, G., Wang, L. G., Yan, G. R., and He, Q. Y. (2015). DOSE: an R/bioconductor package for disease ontology semantic and enrichment analysis. *Bioinforma. Oxf. Engl.* 31, 608–609. doi:10.1093/bioinformatics/btu684

Zhang, X., Bogunovic, D., Payelle-Brogard, B., Francois-Newton, V., Speer, S. D., Yuan, C., et al. (2015). Human intracellular ISG15 prevents interferon-α/β over-amplification and auto-inflammation. *Nature* 517, 89–93. doi:10.1038/nature13801

Zhang, X., Du, L., Han, J., Li, X., Wang, H., Zheng, G., et al. (2020). Novel long non-coding RNA LINC02323 promotes epithelial-mesenchymal transition and metastasis via sponging miR-1343-3p in lung adenocarcinoma. *Thorac. Cancer* 11, 2506–2516. doi:10.1111/1759-7714.13562

Zhou, G., Soufan, O., Ewald, J., Hancock, R. E. W., Basu, N., and Xia, J. (2019). NetworkAnalyst 3.0: a visual analytics platform for comprehensive gene expression profiling and meta-analysis. *Nucleic Acids Res.* 47, W234–W241. doi:10.1093/nar/gkz240