Heliyon 10 (2024) e28085

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

Identification of Y–linked biomarkers and exploration of immune infiltration of normal-appearing gray matter in multiple sclerosis by bioinformatic analysis

Shaoru Zhang [1], Mengjie Zhang [1], Lei Zhang, Zijie Wang, Shi Tang, Xiaolin Yang, Zhizhong Li, Jinzhou Feng [**], Xinyue Qin [*]

Department of Neurology, The First Affiliated Hospital of Chongqing Medical University, 1st Youyi Road, Yuzhong District, Chongqing, 400016, China

ARTICLE INFO

Keywords:
Multiple sclerosis
Normal-appearing cortical gray matter
Y–linked gene
Immune infiltration
Bioinformatic analysis

ABSTRACT

Background: The knowledge of normal–appearing cortical gray matter (NAGM) in multiple sclerosis (MS) remains unclear. In this study, we aimed to identify diagnostic biomarkers and explore the immune infiltration characteristics of NAGM in MS through bioinformatic analysis and validation *in vivo*.

Methods: Differentially expressed genes (DEGs) were analyzed. Subsequently, the functional pathways of the DEGs were determined. After screening the overlapping DEGs of MS with two machine learning methods, the biomarkers' efficacy and the expression levels of overlapping DEGs were calculated. Quantitative reverse transcription polymerase chain reaction (qRT–PCR) identified the robust diagnostic biomarkers. Additionally, infiltrating immune cell populations were estimated and correlated with the biomarkers. Finally, the characteristics of immune infiltration of NAGM from MS were evaluated.

Results: A total of 98 DEGs were identified. They participated in sensory transduction of the olfactory system, synaptic signaling, and immune responses. Nine overlapping genes were screened by machine learning methods. After verified by ROC curve, four genes, namely *HLA–DRB1*, *RPS4Y1*, *EIF1AY* and *USP9Y*, were screened as candidate biomarkers. The mRNA expression of *RPS4Y1* and *USP9Y* was significantly lower in MS patients than that in the controls. They were selected as the robust diagnostic biomarkers for male MS patients. *RPS4Y1* and *USP9Y* were both positively correlated with memory B cells. Moreover, naive $CD4^+$ T cells and monocytes were increased in the NAGM of MS patients compared with those in controls.

List of abbreviations: AUC, area under the ROC curve; BP, biological process; CC, cellular component; CNS, central nervous system; CSF, cerebrospinal fluid; DEGs, differentially expressed genes; EAE, experimental allergic encephalomyelitis; FA, fractional anisotropy; GEO, Gene Expression Omnibus; GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; LASSO, least absolute shrinkage and selection operator; MF, molecular function; MS, multiple sclerosis; NAGM, normal–appearing cortical gray matter; NAWM, normal–appearing white matter; qRT–PCR, quantitative reverse transcription polymerase chain reaction; ROC, receiver operator characteristic; SVM–RFE, support vector machine–recursive feature elimination.

* Corresponding author.
** Corresponding author.
E-mail addresses: fengjinzhou@hotmail.com (J. Feng), qinxinyuecqmu@163.com (X. Qin).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.heliyon.2024.e28085
Received 15 May 2023; Received in revised form 3 March 2024; Accepted 12 March 2024
Available online 13 March 2024

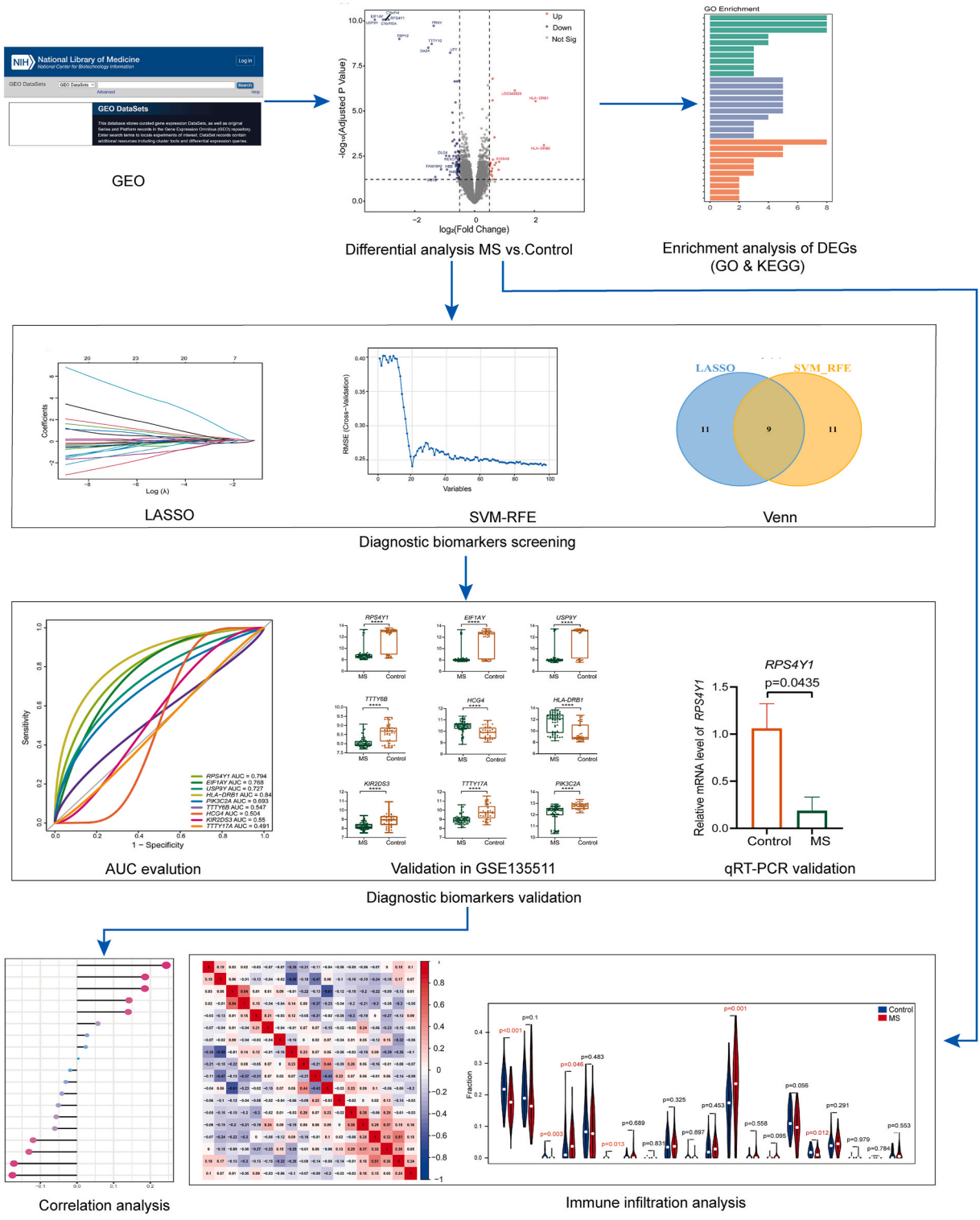

Fig. 1. The flowchart of bioinformatics analysis and experimental validation. Abbreviations: DEGs, differentially expressed genes; MS, multiple sclerosis; GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; LASSO, least absolute shrinkage and selection operator; SVM−RFE, support vector machine-recursive feature elimination; AUC, area under the receiver-operator characteristic curve; qRT−PCR, quantitative reverse transcription polymerase chain reaction.

Conclusions: Low expressed Y−linked genes, *RPS4Y1* and *USP9Y*, were identified as diagnostic biomarkers for MS in male patients. The inhomogeneity of immune cells in NAGM might exacerbate intricate interplay between the CNS and the immune system in the MS.

1. Introduction

More prevalent in females, multiple sclerosis (MS) is a neurological inflammatory and degenerative disease with an increasing prevalence worldwide, with approximately three million people estimated to be affected [1,2]. Timely diagnosis of MS enables effective therapy and minimizes unpredictable dysfunction [3]. Efforts have been made to improve diagnostic criteria for MS recently [4]. The latest standard for MS diagnosis comprises clinical symptoms, magnetic resonance imaging lesions, cerebrospinal fluid (CSF) specific oligoclonal bands and exclusion of other diagnostic possibilities [5]. While the criteria are often misunderstood and inappropriately applied [6]. Therefore, recent researches have been attempted to explore novel diagnostic biomarkers to improve diagnosis, monitoring and treatment for MS [3]. Serum glial fibrillary acid protein readily distinguishes MS from healthy controls [7]. Kappa free light chains is demonstrated as a sensitive CSF marker of MS and already in commercial use [8]. Autophagy markers show promise to predict MS [9]. Besides, many other markers are impacted by treatment effects or lack specificity. Although developing reliable biomarkers remains challenging, continuing studies are making effort to address the unmet need for an accessible and accurate diagnostic approach for MS [3,4]. Notwithstanding specific triggers are still unexplicit, epidemiologic studies have considered that genetically predisposed individuals may develop MS for some certain environmental factors [10]. Diagnostic biomarkers at the genetic level might drive further investigation of the diagnosis of MS.

There is burgeoning evidence that primitive demyelination and astrocytic scarring in the central nervous system (CNS) are the most distinctive pathological hallmarks of MS [11]. Moreover, pathology of normal−appearing white matter (NAWM) and normal−appearing gray matter (NAGM) has recently been recognized as a pathogenisis in MS [12]. On MRI scans, fractional anisotropy (FA) is found to be increased in the NAGM of MS, NAGM−FA has a high correlation with the progression of physical disability in MS patients [13]. Furthermore, the NAGM exhibits microstructural damage, the severity of which is related to cognitive impairment and neurodegeneration in MS [12,14]. NAMG has gained attention in the field of MS research. However, the current understanding of NAGM relies on the differences in imaging and is still short of generalizing the unique relationship between NAGM and genomics.

Another limitation is that a considerable number of studies focus on the etiological effects of single or several immune cells in MS [10,15]. Despite their explanatory power, the overall comprehensive immune cell infiltration in MS deserves additional attention and appraisal. Resident brain cells, particularly immune cells, are altered to different degrees in MS. The regulation of astrocyte−neuronal lactate shuttling and glutamatergic-glutamine cycle genes of NAGM in MS is changed by immune signaling, which may underlie the cerebral cortical dysfunction encountered in MS patients [16]. Due to the potential influence of NAGM and immune responses in MS, gaining a better understanding of the complex immune cellular interactions in NAGM is important for MS progression.

In light of the above two limitations, our study draws on bioinformatics analysis and extends it to the context of NAGM in MS. In recent studies on transcriptomics, microarray data allow simultaneous analysis of multiple genes against the background of a certain disease or state [17]. Studies based on bioinformatics analysis have elucidated how the transcriptome varies among distinct phenotypes [18]. Integrative analysis of bioinformatics methods and genomics studies, including transcriptomics, proteomics, high-throughput sequencing, and single−cell sequencing has identified a group of diagnostic genes, which provide insight into the molecular mechanism of diseases progression [19,20]. Following their lead, we comprehensively utilized microarray data, bioinformatics analysis, machine learning methods and experimental validation to identify biomarkers and explore the immune features of NAGM in MS.

This article was organized into five sections. We began with a differential analysis of the microarray dataset to analyze the differentially expressed genes (DEGs) of NAGM in MS patients. We then analyzed the enrichment pathways of the DEGs. The third and fourth sections identified the diagnostic biomarkers and verified their expression in MS patients, respectively. We subsequently determined the correlations of the infiltrating immune cells and biomarkers and explored the immune infiltration of NAGM in MS (Fig. 1).

2. Materials and methods

2.1. Data processing

We searched in Gene Expression Omnibus (GEO) database to download microarray datasets [21]. The GSE131281 dataset comprises 64 postmortem brain tissue samples from 25 MS patients as well as 42 tissue samples from 14 controls. The GSE135511 dataset includes gene expression profile analysis of 40 postmortem motor cortices from 20 MS patients and 10 samples from 10 non-neurological controls. GSE135511 was used as the validation set.

Raw idat files were subjected to outlier exclusion, background correction and normalization. The "beadarraynormalizeIllumina (method = quantile)" followed by $\log_2$ transformation for further analyses [22] was used. The platform offered the annotation files which guided converting probes into gene symbols.

Table 1
Primer sequences used for qRT−PCR.

Genes	Forward primer sequence	Reverse primer sequence
HLA−DRB1	CGGGGTTGGTGAGAGCTTC	AACCACCTGACTTCAATGCTG
RPS4Y1	GGTTCGAGTGGATGTCACATAC	CGGAAATGTTCACCTGTCTTCTC
EIF1AY	CGGAAATGTTCACCTGTCTTCTC	TAGTCCCGTAGACCAACCAAT
USP9Y	ACTCATGGCTCTCCAGTAGGA	GGAAATGCAGGCTCTTCATCT
GAPDH	GGAGCGAGATCCCTCCAAAAT	GGCTGTTGTCATACTTCTCATGG

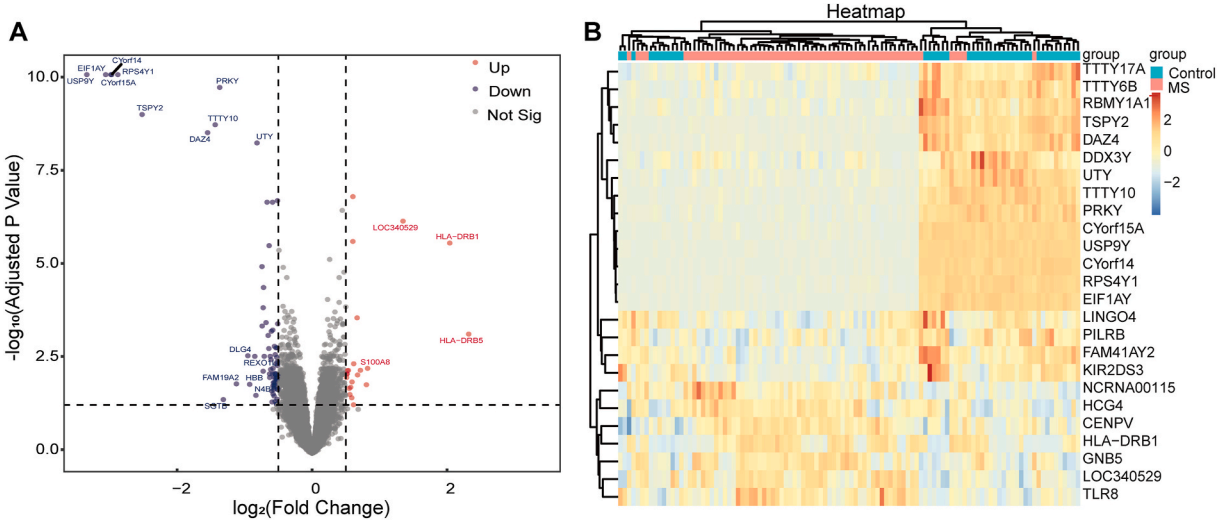

Fig. 2. Identification of differentially expressed genes (DEGs) between MS and control samples. (A)Volcano plot of DEGs between MS patients and controls in the GSE131281 dataset. The dotted lines represented $|\log_2$ fold change$| = 0.5$. The dots showed the number of upregulated (red) and downregulated (blue) genes. (B) Heatmap of top 25 DEGs in the GSE131281 dataset. Orange represented high expression, and blue represented low expression. Abbreviation: Not Sig, not significant. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

2.2. DEGs identification

DEGs between MS patients and controls analyzed by "limma" package [23]. The adjusted P value less than 0.05 along with the $|\log_2$ fold change$|$ more than 0.5 were set as threshold. The differential expression of DEGs was visualized by the volcano plot.

2.3. Enrichment analysis

Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) were performed by "clusterProfiler" [24] and "org.Hs.eg.db" packages. Subsequently, "ggplot2" package was used to visualize the targets.

2.4. Overlapping DEGs screening

Least absolute shrinkage and selection operator (LASSO) is a shrinking methodology minimizing regression coefficients which reduces the potential of overfitting and selects pertinent predictor covariates for the data [25]. Support vector machines (SVM) have exhibited high superiority in classification but have failed to assess the importance of predictors. SVM−recursive feature elimination (SVM−RFE) is a more elegant feature-filtering method [26]. Thus, we respectively performed LASSO and SVM−RFE to select featured DEGs for MS. The crossover featured DEGs selected by the two algorithms were used for subsequent validations. The "glmnet" package performed LASSO [27] and the generalization error was generated by 100−fold cross validation. SVM−RFE was performed with "mlbench" and "caret" packages and validated by 10-fold cross-validations.

2.5. Verification the effectiveness of the overlapping DEGs

The GSE135511 dataset was applied as the external validation dataset. The "pROC" package was applied to evaluate the prediction potency of overlapping DEGs in the GSE131281 and GSE135511 datasets and the receiver-operator characteristic (ROC) curve of each overlapping DEGs were estimated. The area under the ROC curve (AUC) was used for determining DEGs' efficiency. The expression

Fig. 3. The Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis of the DEGs. (A) The top 10 enrichment items for 3 categories in the GO enrichment analysis. Green bars represented biological process (BP) terms, blue bars represented cell component (CC) terms, and red bars represented molecular function (MF) terms. (B) The top 20 enrichment items in the KEGG pathway analysis of DEGs. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

levels of the overlapping DEGs were calculated. The DEGs with fine AUC (AUC>0.7) and differentially expressed in GSE135511 dataset were chose as candidate genes for *in vivo* validation.

2.6. Patients recruitment and samples collection

The quantitative reverse transcription polymerase chain reaction (qRT−PCR) experiments was applied to validate the biomarkers in peripheral blood of MS patients. The study was approved by the Ethics Committee of the First Affiliated Hospital of Chongqing Medical University (Ethics approval number: 2023−180; Ethics approval date April 26th, 2023). Criteria for enrollment of MS patients: (1) fulfilled the 2017 McDonald criteria; (2) at least 4 weeks without steroid treatment before enrollment; (3) signed written consent.

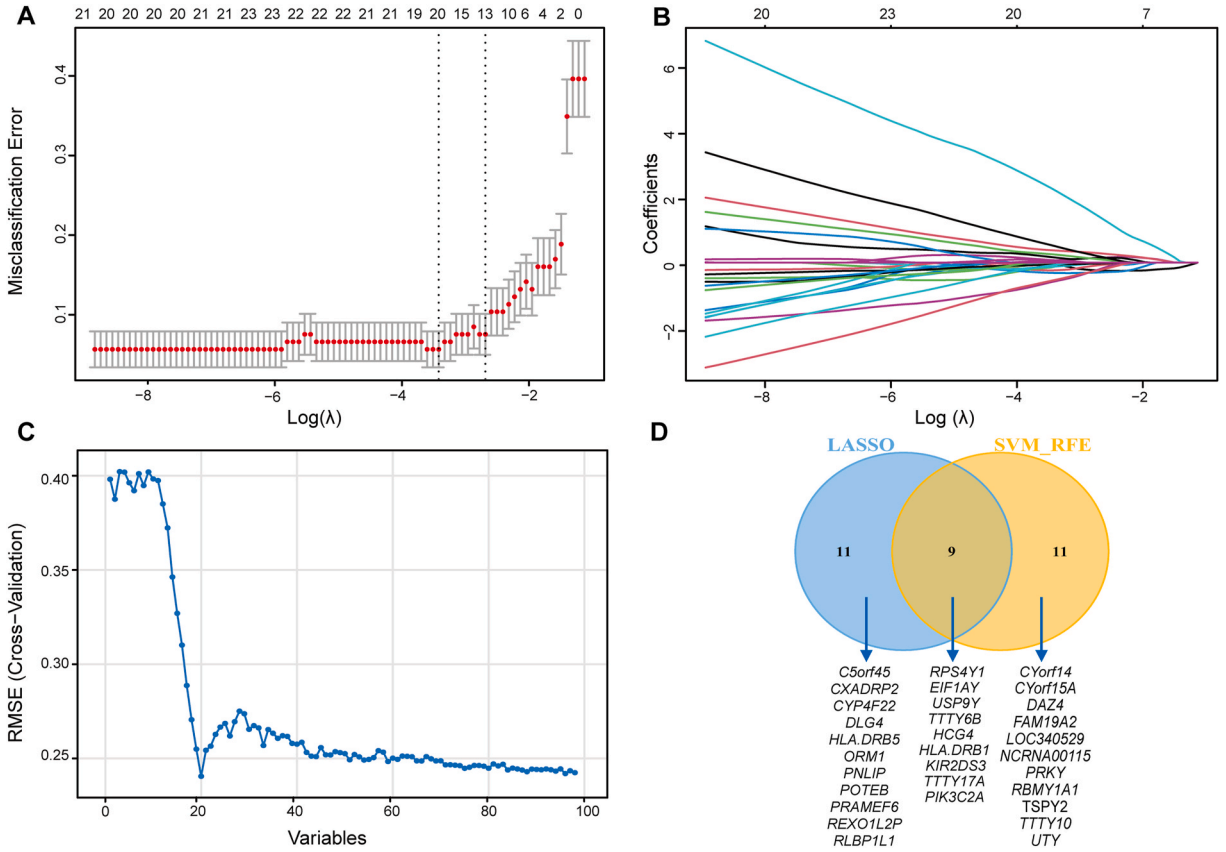

Fig. 4. Overlapping DEGs screened via the least absolute shrinkage and selection operator (LASSO) logistic regression algorithm and support vector machine-recursive feature elimination (SVM-RFE) algorithm. (A) The misclassification error curve plotted versus log(λ), where λ was the tuning parameter. 100−fold cross−validation was used to tunning parameter selection in the LASSO model. (B) LASSO coefficient profiles of DEGs. (C) The SVM−RFE algorithm used RMSE (root mean square error) to filter the DEGs and to select the optimal feature genes for MS. (D) Venn diagram depicted nine overlapping genes as defined by the LASSO and SVM−RFE algorithms.

All participants enrolled after obtaining written consent.

Whole blood samples (approximately 2 mL) were collected into an EDTA−treated tube and then divided into 1.5 mL tubes (n = 20, 100 μL/tube) with anticoagulant. Each 100 μL blood sample was mixed with 1 mL of Trizol and stored at −80 °C.

2.7. qRT−PCR experiment to verify the expression of diagnostic biomarkers

Total RNA was extracted from blood sample based on standard protocols of RNA Extraction Reagent (9109, Takara, Japan). The kit (R333, Vazyme, China) was applied for cDNA synthesis. The Mix (Q711, Vazyme, China) was provided for quantifying the gene expression. The $2^{-\triangle\triangle Ct}$ method was used to calculated the expression levels. Gene expression with P value less than 0.05 was considered as the biomarker of NAGM for MS. The primers we used are listed in Table 1.

2.8. Correlation analysis

The Spearman correlation analysis and CIBERSORT were performed to analyze the correlation between the infiltrating immune cells and biomarkers in NAGM from MS patients. CIBERSORT is a gene expression-based arithmetic deconvoluting expression profiles along with estimating immunocyte subtypes abundance of samples [28]. The immunocytes in the GSE131281 dataset were determined by CIBERSORT with 1000 permutations. As CIBERSORT P value < 0.05, the inferred portion of the subtypes was considered reliable. The correlations were visualized by using "ggplot2" and "psych".

2.9. Immunocyte subtypes evaluation

CIBERSORT was further implemented to explored the immune characteristics of NAGM from MS patients. The difference analysis of immunocyte proportions of NAGM between the MS and control samples in the GSE131281 dataset were visualized utilizing "vioplot" package. The correlation degree of each subtype was visualized by the heatmap plot via the "corrplot" package.

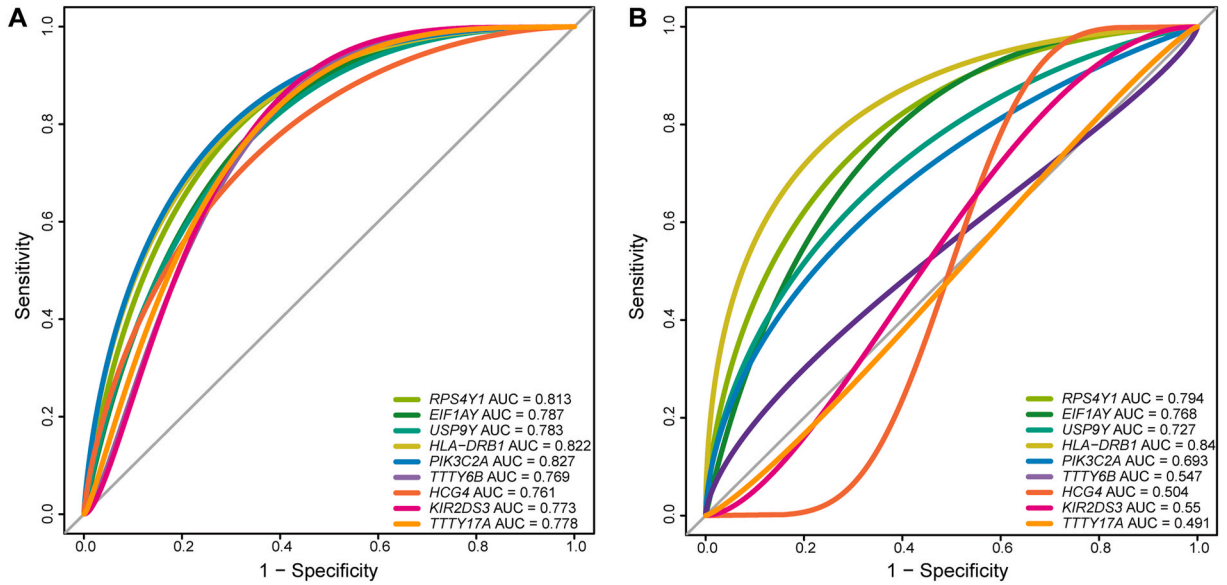

Fig. 5. Validation of overlapping genes in the training dataset and the validation dataset. (A) ROC curve evaluating the diagnostic effectiveness of the nine overlapping genes in the GSE131281 dataset. (B) ROC curve for diagnostic efficacy verification of the nine overlapping genes in the GSE135511 dataset.

2.10. Statistical analysis

We applied R (version 4.1.2) to analyze the datasets. Wilcoxon rank−sum test and t−test were utilized for the difference analysis. The correlation analysis was evaluated by the Spearman correlation analysis. The qRT−PCR data were analyzed by t−test using Prism 8.0 (GraphPad, USA). Significant differences were considered to be $P < 0.05$ or adjusted $P < 0.05$, which defined as the raw P values adjusted for multiple testing using Bonferroni correction in this study.

3. Results

3.1. The DEGs in NAGM of MS

The DEGs of 64 MS tissue samples and 42 control cortex tissues from the GSE131281 dataset were analyzed. Ultimately, 22 upregulated genes and 76 downregulated genes were identified (Fig. 2A). Table S1 listed the DEGs. The overall expression patterns of the top 25 DEGs were visualized in a clustering heatmap. Most genes in MS samples were expressed at lower levels (Fig. 2B).

3.2. Functional correlation analysis

According to the GO analysis, 33 GO−biological process (BP) terms, 18 GO−cellular component (CC) terms and 13 GO−molecular function (MF) terms were statistically significantly enriched in DEGs. The top 10 enrichment terms in each category were displayed in Fig. 3A. The top 3 enriched GO−BP terms included detection of chemical stimuli involved in sensory perception, sensory perception of smell and detection of chemical stimuli involved in sensory perception of smell. The five most enriched GO−CC terms showed that the DEGs were significantly associated with endocytic vesicles, vesicle lumens, cytoplasmic vesicle lumens, secretory granule lumens, and coated vesicles. The three most enriched terms in the GO-MF were olfactory receptor activity, antigen binding and immune receptor activity.

KEGG pathway analysis also revealed that DEGs associated strongly with olfactory transduction, graft-versus-host disease and antigen processing and presentation (Fig. 3B).

3.3. Screening, validation and effectiveness of feature DEGs in NAGM of MS

LASSO regression identified 20 genes (Fig. 4A and B). A subset of 20 items were finally selected by the SVM−RFE algorithm (Fig. 4C). Eventually, nine overlapping genes were defined by the two machine learning algorithms: *RPS4Y1*, *EIF1AY*, *USP9Y*, *TTTY6B*, *HCG4*, *HLA−DRB1*, *KIR2DS3*, *TTTY17A* and *PIK3C2A* (Fig. 4D).

The nine overlapping DEGs for MS were further verified in the GSE135511 dataset to obtain more precise and credible biomarkers. The AUC of each overlapping gene in the GSE131281 dataset was 0.813 for *RPS4Y1*, 0.787 for *EIF1AY*, 0.783 for *USP9Y*, 0.769 for *TTTY6B*, 0.761 for *HCG4*, 0.822 for *HLA-DRB1*, 0.773 for *KIR2DS3*, 0.778 for *TTTY17A* and 0.827 for *PIK3C2A* (Fig. 5A). In the

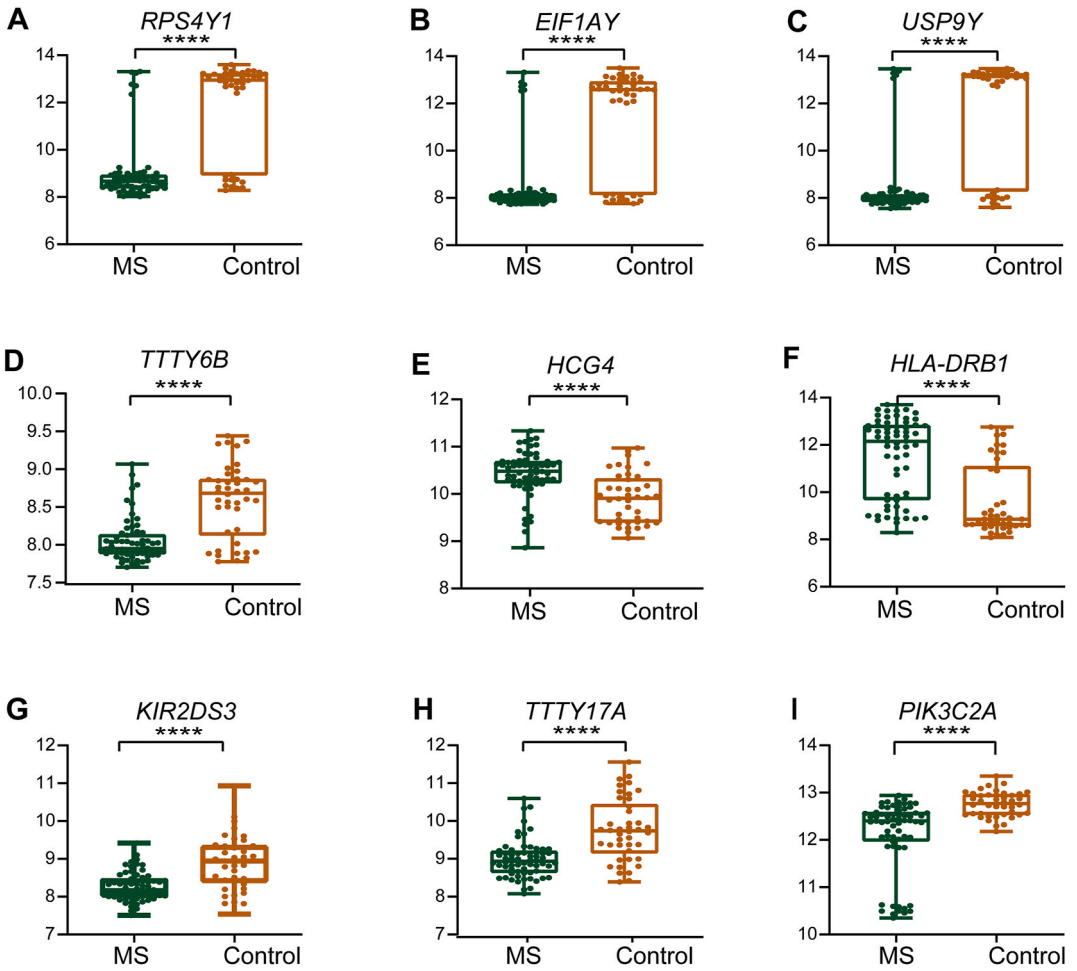

Fig. 6. The expression of nine overlapping genes in the MS and control samples from the GSE131281 dataset (****P < 0.001).

GSE135511 dataset, *RPS4Y1* (AUC = 0.794), *EIF1AY* (AUC = 0.768), *USP9Y* (AUC = 0.727) and *HLA−DRB1* (AUC = 0.84) also exhibited good prediction efficiency (Fig. 5B).

Furthermore, the expression of each gene of the MS were significantly different from those of the controls in the GSE131281 dataset (Fig. 6A–I). We also examined the expression of these biomarkers in the GSE135511 validation dataset. The expression of *RPS4Y1*, *EIF1AY*, *USP9Y* and *HLA−DRB1* of MS patients were significantly different from that of the controls in the GSE135511 dataset (Fig. 7A, B, 7C, 7F). While the expression levels of *TTTY6B* (Fig. 7D), *HCG4* (Fig. 7E), *KIR2DS3* (Fig. 7G), *TTTY17A* (Fig. 7H) and *PIK3C2A* (Fig. 7I) were not statistically different between MS patients and controls. Therefore, *RPS4Y1*, *EIF1AY*, *USP9Y* and *HLA−DRB1* were considered as the candidate biomarkers for further validation in peripheral blood of MS patients.

3.4. Expression level of RPS4Y1, EIF1AY, USP9Y and HLA−DRB1 in peripheral blood samples

Three MS patients and three paired age- and gender-matched controls were enrolled. Details of demographic characteristics were presented Table 2. The mRNA expression of *RPS4Y1* and *USP9Y* were downregulated in the MS compared to that of the control group (Fig. 8A and B). The expression of *EIF1AY and HLA−DRB1* exhibited no differences between MS patients and controls (Fig. 8C and D).

3.5. Correlation analysis of biomarkers between immunocytes of NAGM

Both *RPS4Y1* (r = 0.24, P = 0.011) and *USP9Y* (r = 0.31, P = 0.001) had positive correlation with memory B cells (Fig. 9A and B). *USP9Y* also exhibited a positive correlation with resting dendritic cells (r = 0.22, P = 0.026) and a negative correlation with M0 macrophages (r = −0.22, P = 0.021) and activated mast cells (r = −0.25, P = 0.009) (Fig. 9B).

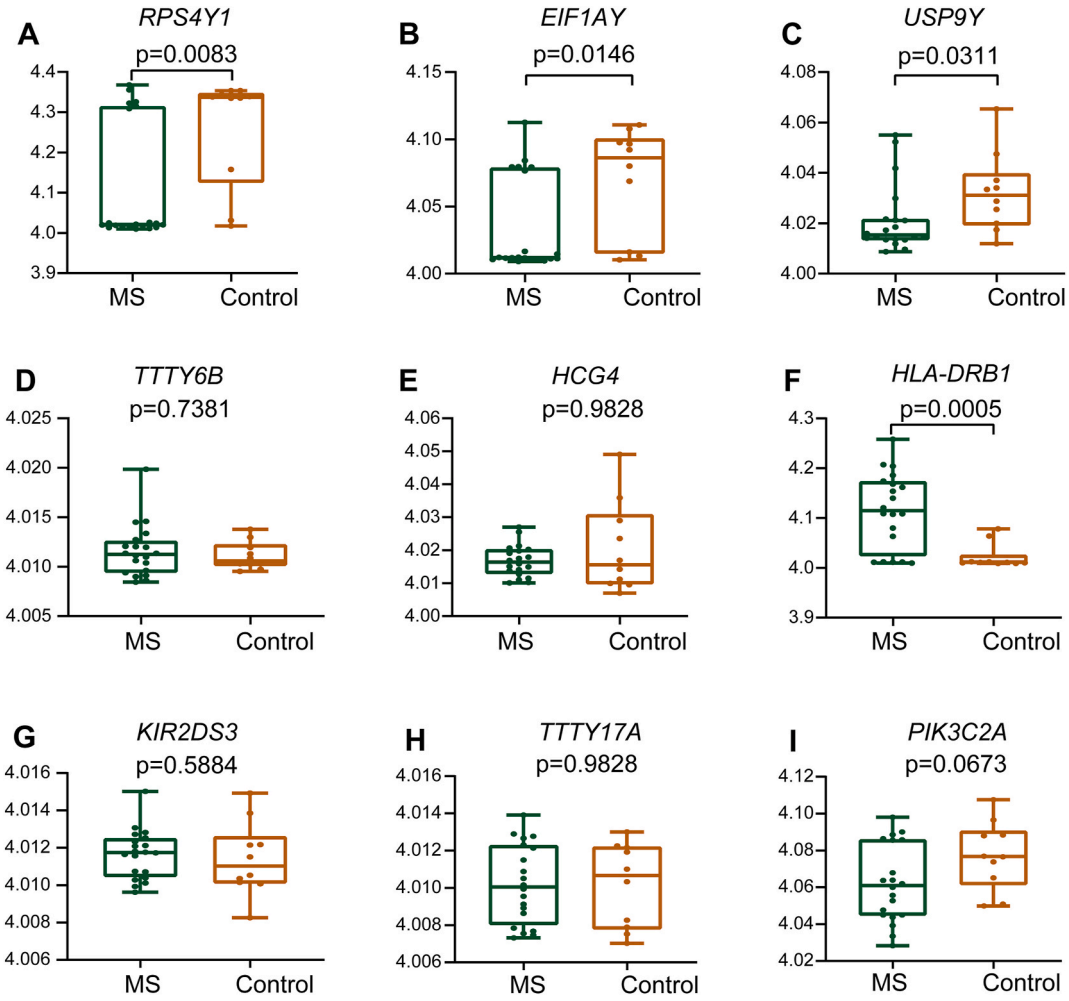

Fig. 7. Gene expression of nine overlapping genes in MS and control samples from the GSE135511 dataset.

Table 2
Demographic characteristics of MS patients and controls.

Variables	MS	Control
Cases	3	3
Age(years)	36.0 ± 16.7	35.7 ± 15.6
Gender	Male	Male
Smoke status	Never	Never

3.6. Immune infiltration analysis

The immunocyte proportions in NAGM of MS contained higher proportions of monocytes and naive $CD4^+$ T cells, but smaller proportions of naive B cells, $CD8^+$ T cells, activated memory $CD4^+$ T cells and activated dendritic cells versus controls (Fig. 10A).

The connections among subtypes of immunocytes in NAGM from MS patients were investigated via a correlation matrix (Fig. 10B). $CD8^+$ T cells were negatively associated with resting dendritic cells ($r = -0.58$, $P < 0.001$). Monocytes exhibited positive correlation with M2 macrophages ($r = 0.51$, $P < 0.001$). Activated memory $CD4^+$ T cells negatively correlated with resting memory $CD4^+$ T cells ($r = -0.61$, $P < 0.001$).

4. Discussion

MS is a demyelinating inflammatory disorder that affects not only white matter lesions but normal appearing matter lesions [29]. The complex etiopathogenesis of MS means that there is still room for improvement of its diagnosis. This study was focus on identifying

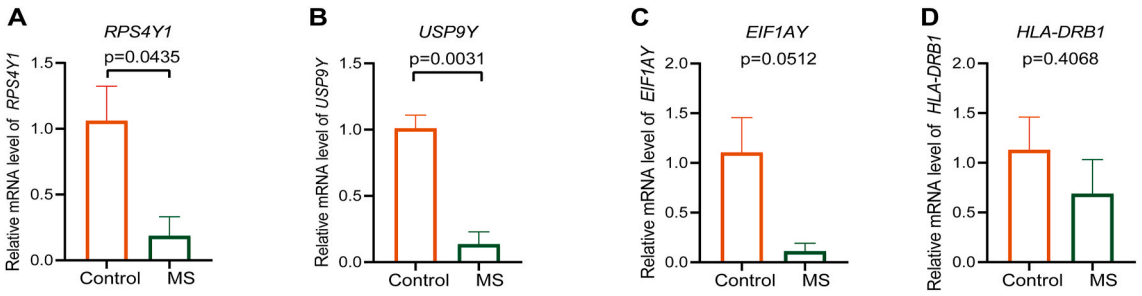

Fig. 8. Quantitative reverse transcription polymerase chain reaction (qRT–PCR) analyses of the expression levels of (A) *RPS4Y1*, (B) *EIF1AY*, (C) *USP9Y*, and (D) *HLA–DRB1* in peripheral blood samples isolated from controls and MS patients (n = 3).

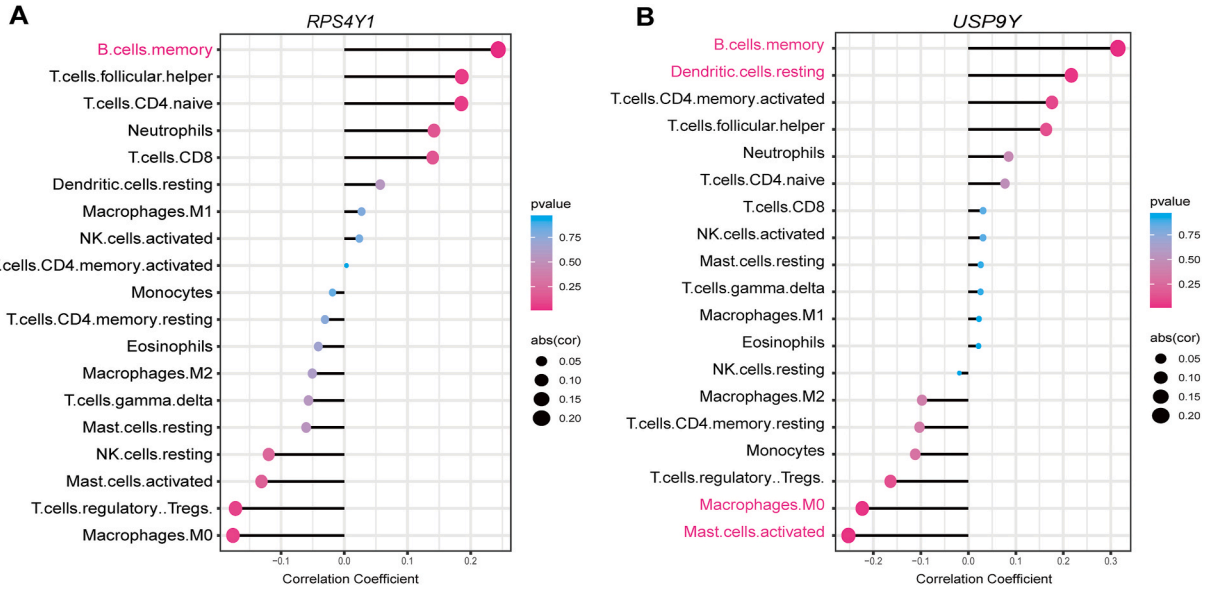

Fig. 9. Correlation between diagnostic biomarkers and infiltrating immunocyte abundance in MS patients. (A) Correlation between *RPS4Y1* and infiltrating immunocytes. (B) Correlation between *USP9Y* and infiltrating immunocytes.

diagnostic biomarkers for MS from NAGM and improving better understanding of the immune infiltration of NAGM in MS.

First, we identified DEGs of NAGM between controls and MS patients. Then, we analyzed the DEGs' enrichment pathways. The DEGs enriched in olfactory transduction, synaptic signaling mainly in vesicle communication, and immune responses. The latter included antigen processing and presentation and graft-versus-host disease. Previous findings have indicated that the olfactory threshold reflects the relapse of MS [30]. Synaptic dysfunction can cause primary neuronal death in MS [31]. Meanwhile, the impairment of vesicle communication might affect the axonal integrity and synaptic signaling pathways, which may lead to the onset of MS [32]. As an acquired autoimmune condition, MS is involved in diverse immune responses and inflammatory infiltration [33]. Our current enrichment analysis implied that DEGs of MS contributed to MS pathogenesis.

Next, we identified the diagnostic biomarkers of MS and determined their association with immune cells. After selecting overlapping candidate DEGs, we validated the candidate genes by ROC curve, external validation dataset and *in vivo* experiment to identify robust diagnostic biomarkers of MS. An unexpected observation in this study was that the diagnostic biomarkers were linked to the Y chromosome. *RPS4Y*1 encodes the 40S ribosomal protein S4, which may function in ribosomal assembly and aging–related processes [34,35]. *USP9Y* encodes a protein that resembles ubiquitin–specific protease. Ubiquitin and ubiquitin-like proteins regulate key features to buffer stimulus imbalances caused by immune attack [36,37]. The immune cell correlation analysis suggested that low levels of *RPS4Y1* and *USP9Y* were involved in the immune dysfunction of MS to some extent.

Extremely downregulated expression of *RPS4Y1* and *USP9Y* has been observed in cancers [38]. The downregulated expression of *RPS4Y1* and *USP9Y* was first reported in MS in this study. The Y chromosome is believed to contain polymorphic genes that may reduce susceptibility to the experimental allergic encephalomyelitis (EAE) model of MS [39]. MS occurs primarily in females; however, males may experience more progressive MS [40]. Our results provided evidence that *RPS4Y1* and *USP9Y* may have implications for the pathogenesis of male MS.

Finally, we explored the immune infiltration of NAGM in MS. There was an increase abundance of naive $CD4^+$ T cells and

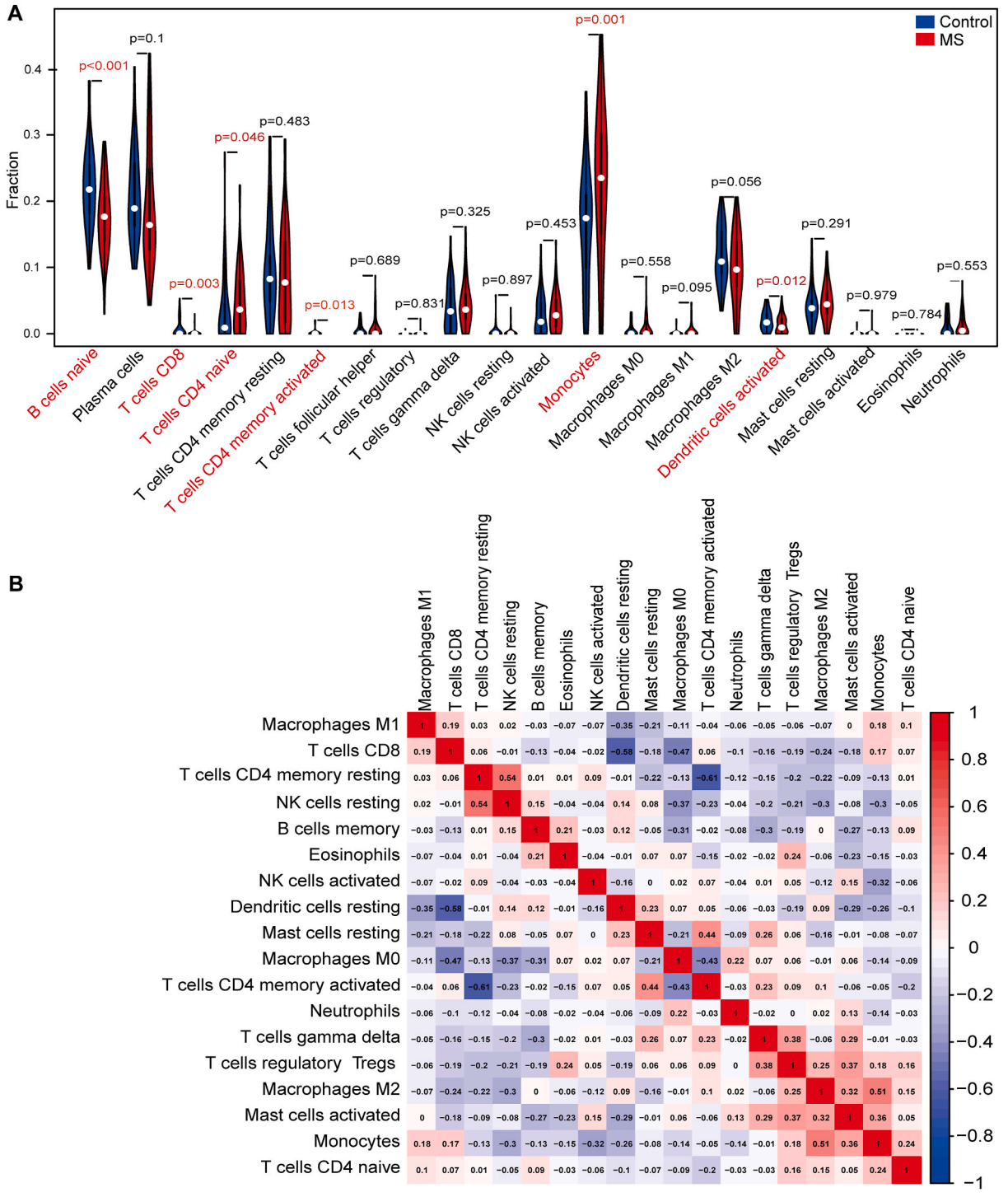

Fig. 10. Infiltration of immune cells in the GSE131281 dataset. (A) Violin plot of the differences in the abundance of infiltrating immune cell subtypes between the two groups. Blue represented control samples, and red represented MS samples. Horizontal and vertical axes represented subsets of immune cells and relative percentages, respectively. (B) Heatmap of correlations of the abundance of immune cells between MS patients and controls. The horizontal and vertical axes represented subsets of immune cells. Subtypes of immune cells with higher, lower, and equivalent correlation levels were presented in red, blue, and white, respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

monocytes, with a concomitant reduction of $CD8^+$ T cells, activated dendritic cells, activated memory $CD4^+$ T cells and naive B cells of NAGM in MS relative to the controls. Naive $CD4^+$ T cells are the important responders in MS when antigen stimulates [33]. Monocytes implicate in the pathogenesis of MS by damaging myelin [41]. During the acute and relapsing phases of MS, B cells, $CD4^+$ T cells and $CD8^+$ T cells attack myelin, causing resultant demyelination [33,42,43]. Causing axonal loss by releasing cytolytic granules, the amount of $CD8^+$ T cells are far higher than that of other immune cells in MS lesions [43]. The potential role of dendritic cells in MS and EAE pathogenesis is heterogeneous because they can be with pro-inflammatory or tolerogenic phenotype under certain conditions [44]. However, naive B cells have been shown to downregulate proinflammatory immune responses [45]. Our findings indicated that the altered abundance of immune cell subtypes in NAGM from MS patients might possess both protective and deleterious effects eliciting complicated processes.

We found that the infiltrating immune cell subpopulations also exhibited complex correlations with each other. Monocytes infiltrate into the CNS and participate in demyelination and tissue destruction by secreting proinflammatory cytokines, disrupting the blood–brain barrier, and allowing the entry of activated T cells into the CNS [46]. M2 macrophages are recruited to the site of inflammation and promote repair and remyelination in MS [47]. MS involves in a complex interplay among distinct immune cell populations, which either ameliorate or exacerbate the disease progress [48,49]. The immune infiltration results confirmed the enrichment analysis and were in accordance with the earlier investigations.

Shortcomings of this study could not be neglected. Firstly, the small blood sample size of our participants may be the reason for the failure validation of *HLA–DRB1* and *EIF1AY*. Secondly, most of the donor samples of the dataset were from progressive forms of MS, which limits the impact of the findings, as the majority of MS patients manifest relapsing–remitting disease. Therefore, the interpretations of this study should be performed cautiously.

5. Conclusions

DEGs of NAGM in MS were enriched in olfactory transition, synaptic signaling, and the immune response. Low expression of the Y–linked genes, *RPS4Y1* and *USP9Y*, may be important in the pathogenesis of male MS patients. Immune cells not only infiltrated lesions but NAGM. The inhomogeneity of immune cells in NAGM exerted either ameliorating or exacerbating roles which exacerbated intricate interactions in MS. We hope that this study may further the knowledge of NAGM in MS. Further studies are needed to reveal effective diagnostic biomarkers, explore novel therapeutic targets, and serve as prognostic tools.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Ethics declarations

This study was reviewed and approved by the Ethics Committee of the First Affiliated Hospital of Chongqing Medical University, with the ethics approval number of 2023–180 and ethics approval date in April 26th, 2023. All participants/patients provided informed consent to participate in the study.

Consent for publication

All authors agree to publish the manuscript.

Data availability statement

The datasets supporting the conclusions of this article are available in the Gene Expression Omnibus (GEO) repository. The source datasets are available at https://www.ncbi.nlm.nih.gov/gds/?term=GSE131281 and https://www.ncbi.nlm.nih.gov/gds/?term=GSE135511, respectively. Further inquiries can be directed to the corresponding author.

CRediT authorship contribution statement

Shaoru Zhang: Data curation, Formal analysis, Writing – original draft, Software. **Mengjie Zhang:** Data curation, Methodology, Formal analysis. **Lei Zhang:** Writing – original draft. **Zijie Wang:** Validation. **Shi Tang:** Validation. **Xiaolin Yang:** Methodology. **Zhizhong Li:** Methodology. **Jinzhou Feng:** Conceptualization. **Xinyue Qin:** Conceptualization, Supervision, Writing – review & editing.

Declaration of competing interest

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

Acknowledgments

We acknowledge Lukas Simon Enz, Roberta Magliozzi and their colleagues for aggregating the gene expression datasets of the MS and the control groups. We would also like to acknowledge the contributors of whom created databases and developed algorithms.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2024.e28085.

References

[1] C. Walton, R. King, L. Rechtman, W. Kaye, E. Leray, R.A. Marrie, N. Robertson, N. La Rocca, B. Uitdehaag, I. van der Mei, et al., Rising prevalence of multiple sclerosis worldwide: insights from the Atlas of MS, third edition. Multiple sclerosis (Houndmills, Basingstoke, England) 26 (14) (2020) 1816–1821, https://doi.org/10.1177/1352458520970841.
[2] M. Charabati, M.A. Wheeler, H.L. Weiner, F.J. Quintana, Multiple sclerosis: neuroimmune crosstalk and therapeutic targeting, Cell 186 (7) (2023) 1309–1327, https://doi.org/10.1016/j.cell.2023.03.008.
[3] M. Kaisey, G. Lashgari, J. Fert-Bober, D. Ontaneda, A.J. Solomon, N.L. Sicotte, An update on diagnostic laboratory biomarkers for multiple sclerosis, Curr. Neurol. Neurosci. Rep. 22 (10) (2022) 675–688, https://doi.org/10.1007/s11910-022-01227-1.
[4] Z. Khan, G.D. Gupta, S. Mehan, Cellular and molecular evidence of multiple sclerosis diagnosis and treatment challenges, J. Clin. Med. 12 (13) (2023), https://doi.org/10.3390/jcm12134274.
[5] W.M. Carroll, McDonald MS diagnostic criteria: evidence-based revisions, Mult. Scler. 24 (2) (2017) 92–95, https://doi.org/10.1177/1352458517751861, 2018.
[6] A.J. Solomon, R. Pettigrew, R.T. Naismith, S. Chahin, S. Krieger, B. Weinshenker, Challenges in multiple sclerosis diagnosis: misunderstanding and misapplication of the McDonald criteria, Mult. Scler. 27 (2) (2021) 250–258, https://doi.org/10.1177/1352458520910496.
[7] A. Abdelhak, A. Huss, J. Kassubek, H. Tumani, M. Otto, Serum GFAP as a biomarker for disease severity in multiple sclerosis, Sci. Rep. 8 (1) (2018) 14798, https://doi.org/10.1038/s41598-018-33158-8.
[8] R.S. Saadeh, S.C. Bryant, A. McKeon, B. Weinshenker, D.L. Murray, S.J. Pittock, M.A.V. Willrich, CSF kappa free light chains: cutoff validation for diagnosing multiple sclerosis, Mayo Clin. Proc. 97 (4) (2022) 738–751, https://doi.org/10.1016/j.mayocp.2021.09.014.
[9] O. Joodi Khanghah, A. Nourazarian, F. Khaki-Khatibi, M. Nikanfar, D. Laghousi, A.M. Vatankhah, S. Moharami, Evaluation of the diagnostic and predictive value of serum levels of ANT1, ATG5, and parkin in multiple sclerosis, Clin. Neurol. Neurosurg. 197 (2020) 106197, https://doi.org/10.1016/j.clineuro.2020.106197.
[10] S. Rodríguez Murúa, M.F. Farez, F.J. Quintana, The immune response in multiple sclerosis, Annual review of pathology 17 (2022) 121–139, https://doi.org/10.1146/annurev-pathol-052920-040318.
[11] H. Lassmann, Pathogenic mechanisms associated with different clinical courses of multiple sclerosis, Front. Immunol. 9 (2018) 3116, https://doi.org/10.3389/fimmu.2018.03116.
[12] C.A. Roman, P.A. Arnett, Structural brain indices and executive functioning in multiple sclerosis: a review, J. Clin. Exp. Neuropsychol. 38 (3) (2016) 261–274, https://doi.org/10.1080/13803395.2015.1105199.
[13] M. Calabrese, F. Rinaldi, D. Seppi, A. Favaretto, L. Squarcina, I. Mattisi, P. Perini, A. Bertoldo, P. Gallo, Cortical diffusion-tensor imaging abnormalities in multiple sclerosis: a 3-year longitudinal study, Radiology 261 (3) (2011) 891–898, https://doi.org/10.1148/radiol.11110195.
[14] X. Han, X. Wang, L. Wang, Z. Zheng, J. Gu, D. Tang, L. Liu, S. Liu, Investigation of grey matter abnormalities in multiple sclerosis patients by combined use of double inversion recovery sequences and diffusion tensor MRI at 3.0 Tesla, Clin. Radiol. 73 (9) (2018) 834.e817–834.e823, https://doi.org/10.1016/j.crad.2018.04.016.
[15] S. Dybowski, S. Torke, M.S. Weber, Targeting B cells and microglia in multiple sclerosis with bruton tyrosine kinase inhibitors: a review, JAMA Neurol. 80 (4) (2023) 404–414, https://doi.org/10.1001/jamaneurol.2022.5332.
[16] T. Zeis, I. Allaman, M. Gentner, K. Schroder, J. Tschopp, P.J. Magistretti, N. Schaeren-Wiemers, Metabolic gene expression changes in astrocytes in Multiple Sclerosis cerebral cortex are indicative of immune-mediated signaling, Brain Behav. Immun. 48 (2015) 313–325, https://doi.org/10.1016/j.bbi.2015.04.013.
[17] M.L. Elkjaer, R. Röttger, J. Baumbach, Z. Illes, A systematic review of tissue and single cell transcriptome/proteome studies of the brain in multiple sclerosis, Front. Immunol. 13 (2022) 761225, https://doi.org/10.3389/fimmu.2022.761225.
[18] J.H. Malone, B. Oliver, Microarrays, deep sequencing and the true measure of the transcriptome, BMC Biol. 9 (2011) 34, https://doi.org/10.1186/1741-7007-9-34.
[19] M. Wang, W.M. Song, C. Ming, Q. Wang, X. Zhou, P. Xu, A. Krek, Y. Yoon, L. Ho, M.E. Orr, et al., Guidelines for bioinformatics of single-cell sequencing data analysis in Alzheimer's disease: review, recommendation, implementation and application, Mol. Neurodegener. 17 (1) (2022) 17, https://doi.org/10.1186/s13024-022-00517-z.
[20] K. Song, C. Liu, J. Zhang, Y. Yao, H. Xiao, R. Yuan, K. Li, J. Yang, W. Zhao, Y. Zhang, Integrated multi-omics analysis reveals miR-20a as a regulator for metabolic colorectal cancer, Heliyon 8 (3) (2022) e09068, https://doi.org/10.1016/j.heliyon.2022.e09068.
[21] T. Barrett, S.E. Wilhite, P. Ledoux, C. Evangelista, I.F. Kim, M. Tomashevsky, K.A. Marshall, K.H. Phillippy, P.M. Sherman, M. Holko, et al., NCBI GEO: archive for functional genomics data sets–update, Nucleic Acids Res. 41 (Database issue) (2013) D991–D995, https://doi.org/10.1093/nar/gks1193.
[22] M.J. Dunning, M.L. Smith, M.E. Ritchie, S. Tavaré, beadarray: R classes and methods for Illumina bead-based data, Bioinformatics 23 (16) (2007) 2183–2184, https://doi.org/10.1093/bioinformatics/btm311.
[23] M.E. Ritchie, B. Phipson, D. Wu, Y. Hu, C.W. Law, W. Shi, G.K. Smyth, Limma powers differential expression analyses for RNA-sequencing and microarray studies, Nucleic Acids Res. 43 (7) (2015) e47, https://doi.org/10.1093/nar/gkv007.
[24] G. Yu, L.G. Wang, Y. Han, Q.Y. He, clusterProfiler: an R package for comparing biological themes among gene clusters, OMICS A J. Integr. Biol. 16 (5) (2012) 284–287, https://doi.org/10.1089/omi.2011.0118.
[25] A.J. McEligot, V. Poynor, R. Sharma, A. Panangadan, Logistic LASSO regression for dietary intakes and breast cancer, Nutrients 12 (9) (2020), https://doi.org/10.3390/nu12092652.
[26] H. Sanz, C. Valim, E. Vegas, J.M. Oller, F. Reverter, SVM-RFE: selection and visualization of the most relevant features through non-linear kernels, BMC Bioinf. 19 (1) (2018) 432, https://doi.org/10.1186/s12859-018-2451-4.
[27] N. Simon, J. Friedman, T. Hastie, R. Tibshirani, Regularization paths for cox's proportional hazards model via coordinate descent, J. Stat. Software 39 (5) (2011) 1–13, https://doi.org/10.18637/jss.v039.i05.
[28] A.M. Newman, C.L. Liu, M.R. Green, A.J. Gentles, W. Feng, Y. Xu, C.D. Hoang, M. Diehn, A.A. Alizadeh, Robust enumeration of cell subsets from tissue expression profiles, Nat. Methods 12 (5) (2015) 453–457, https://doi.org/10.1038/nmeth.3337.

[29] R. Rahmanzadeh, M. Weigel, P.J. Lu, L. Melie-Garcia, T.D. Nguyen, A. Cagol, F. La Rosa, M. Barakovic, A. Lutti, Y. Wang, et al., A comparative assessment of myelin-sensitive measures in multiple sclerosis patients and healthy subjects, NeuroImage Clinical 36 (2022) 103177, https://doi.org/10.1016/j.nicl.2022.103177.
[30] H. Wood, Neurophysiological and olfactory biomarkers for multiple sclerosis, Nat. Rev. Neurol. 15 (1) (2019) 1, https://doi.org/10.1038/s41582-018-0119-3.
[31] M.C. Geloso, N. D'Ambrosi, Microglial pruning: relevance for synaptic dysfunction in multiple sclerosis and related experimental models, Cells 10 (3) (2021), https://doi.org/10.3390/cells10030686.
[32] T. Bergaglio, A. Luchicchi, G.J. Schenk, Engine failure in axo-myelinic signaling: a potential key player in the pathogenesis of multiple sclerosis, Front. Cell. Neurosci. 15 (2021) 610295, https://doi.org/10.3389/fncel.2021.610295.
[33] M.J. Mansilla, S. Presas-Rodríguez, A. Teniente-Serra, I. González-Larreategui, B. Quirant-Sánchez, F. Fondelli, N. Djedovic, D. Iwaszkiewicz-Grześ, K. Chwojnicki, Miljković Đ, et al., Paving the way towards an effective treatment for multiple sclerosis: advances in cell therapy, Cell. Mol. Immunol. 18 (6) (2021) 1353–1374, https://doi.org/10.1038/s41423-020-00618-z.
[34] A.A. Johnson, M.N. Shokhirev, Pan-tissue aging clock genes that have intimate connections with the immune system and age-related disease, Rejuvenation Res. 24 (5) (2021) 377–389, https://doi.org/10.1089/rej.2021.0012.
[35] R. Chang, L. Chen, G. Su, L. Du, Y. Qin, J. Xu, H. Tan, C. Zhou, Q. Cao, G. Yuan, et al., Identification of Ribosomal Protein S4, Y-Linked 1 as a cyclosporin A plus corticosteroid resistance gene, J. Autoimmun. 112 (2020) 102465, https://doi.org/10.1016/j.jaut.2020.102465.
[36] M. Trujillo, Ubiquitin signalling: controlling the message of surface immune receptors, New Phytol. 231 (1) (2021) 47–53, https://doi.org/10.1111/nph.17360.
[37] G. Çetin, S. Klafack, M. Studencka-Turski, E. Krüger, F. Ebstein, The ubiquitin-proteasome system in immune cells, Biomolecules 11 (1) (2021), https://doi.org/10.3390/biom11010060.
[38] F. Pecori Giraldi, M.F. Cassarino, A. Sesta, M. Terreni, G. Lasio, M. Losa, Sexual dimorphism in cellular and molecular features in human ACTH-secreting pituitary adenomas, Cancers 12 (3) (2020), https://doi.org/10.3390/cancers12030669.
[39] B. Angeloni, R. Bigi, G. Bellucci, R. Mechelli, C. Ballerini, C. Romano, E. Morena, G. Pellicciari, R. Reniè, V. Rinaldi, et al., A case of double standard: sex differences in multiple sclerosis risk factors, Int. J. Mol. Sci. 22 (7) (2021), https://doi.org/10.3390/ijms22073696.
[40] C. Ramien, A. Taenzer, A. Lupu, N. Heckmann, J.B. Engler, K. Patas, M.A. Friese, S.M. Gold, Sex effects on inflammatory and neurodegenerative processes in multiple sclerosis, Neurosci. Biobehav. Rev. 67 (2016) 137–146, https://doi.org/10.1016/j.neubiorev.2015.12.015.
[41] J.W. Prineas, J.D.E. Parratt, Multiple sclerosis: microglia, monocytes, and macrophage-mediated demyelination, J. Neuropathol. Exp. Neurol. 80 (10) (2021) 975–996, https://doi.org/10.1093/jnen/nlab083.
[42] J.A. Lopez, M. Denkova, S. Ramanathan, R.C. Dale, F. Brilot, Pathogenesis of autoimmune demyelination: from multiple sclerosis to neuromyelitis optica spectrum disorders and myelin oligodendrocyte glycoprotein antibody-associated disease, Clinical & translational immunology 10 (7) (2021) e1316, https://doi.org/10.1002/cti2.1316.
[43] C. Baecher-Allan, B.J. Kaskow, H.L. Weiner, Multiple sclerosis: mechanisms and immunotherapy, Neuron 97 (4) (2018) 742–768, https://doi.org/10.1016/j.neuron.2018.01.021.
[44] F. Piacente, M. Bottero, A. Benzi, T. Vigo, A. Uccelli, S. Bruzzone, G. Ferrara, Neuroprotective potential of dendritic cells and sirtuins in multiple sclerosis, Int. J. Mol. Sci. 23 (8) (2022), https://doi.org/10.3390/ijms23084352.
[45] A. Bar-Or, R. Li, Cellular immunology of relapsing multiple sclerosis: interactions, checks, and balances, Lancet Neurol. 20 (6) (2021) 470–483, https://doi.org/10.1016/s1474-4422(21)00063-6.
[46] M. Carstensen, T. Christensen, M. Stilund, H.J. Møller, E.L. Petersen, T. Petersen, Activated monocytes and markers of inflammation in newly diagnosed multiple sclerosis, Immunol. Cell Biol. 98 (7) (2020) 549–562, https://doi.org/10.1111/imcb.12337.
[47] H.J. Fischer, T.L.K. Finck, H.L. Pellkofer, H.M. Reichardt, F. Lühder, Glucocorticoid therapy of multiple sclerosis patients induces anti-inflammatory polarization and increased chemotaxis of monocytes, Front. Immunol. 10 (2019) 1200, https://doi.org/10.3389/fimmu.2019.01200.
[48] J.E. Lengfeld, S.E. Lutz, J.R. Smith, C. Diaconu, C. Scott, S.B. Kofman, C. Choi, C.M. Walsh, C.S. Raine, I. Agalliu, et al., Endothelial Wnt/β-catenin signaling reduces immune cell infiltration in multiple sclerosis, Proc. Natl. Acad. Sci. U.S.A. 114 (7) (2017) E1168–e1177, https://doi.org/10.1073/pnas.1609905114.
[49] C. Veroni, B. Serafini, B. Rosicarelli, C. Fagnani, F. Aloisi, C. Agresti, Connecting immune cell infiltration to the multitasking microglia response and TNF receptor 2 induction in the multiple sclerosis brain, Front. Cell. Neurosci. 14 (2020) 190, https://doi.org/10.3389/fncel.2020.00190.