Am J Transl Res 2022;14(2):772-787
www.ajtr.org /ISSN:1943-8141/AJTR0139607

Original Article

Integrated analysis of differentially expressed genes and a ceRNA network to identify hub lncRNAs and potential drugs for multiple sclerosis

Tianfeng Wang*, Si Xu*, Li Liu*, Shuang Li, Huixue Zhang, Xiaoyu Lu, Xiaotong Kong, Danyang Li, Jianjian Wang, Lihua Wang

*Department of Neurology, The Second Affiliated Hospital of Harbin Medical University, Harbin 150081, Heilongjiang, China. *Co-first authors.*

Received October 6, 2021; Accepted January 16, 2022; Epub February 15, 2022; Published February 28, 2022

Abstract: Objective: Multiple sclerosis (MS) is an autoimmune neuroinflammatory disease of the nervous system. However, the precise molecular mechanisms underlying MS have yet to be fully elucidated. In this study, our aim was to provide novel insight into the pathogenesis of MS and provide a resource for identifying new biomarkers and therapeutics for MS. Methods: In this study, we analyzed the gene expression profiles (GSE21942) and miRNA expression profiles (GSE61741) of MS patient samples that were downloaded from the GEO database and identified differentially expressed mRNAs and miRNAs (DEmRNAs, DEmiRNAs). Next, we constructed a protein-protein interaction (PPI) network and a MS-specific ceRNA network (MCEN) by integrating expression profiles, interaction pairs of mRNA-miRNAs and lncRNA-miRNAs. Then, according to the modular structure of the PPI network, we identified hub DEmRNAs and generated a ceRNA subnetwork so that we could analyze the key lncRNAs that were associated with MS. Results: We first identified 4 modules by constructing a PPI network using DEmRNAs. Functional enrichment analysis showed these modules were enriched in immune-related pathways. Then, we constructed the MCEN and the hub gene-associated ceRNA subnetwork using a comprehensive computational approach. We identified three key lncRNAs (LINC00649, TP73-AS1 and MALAT1) and further identified key lncRNA-mediated ceRNAs within the subnetwork. Finally, by analyzing LINC00649-miR-1275-CD20, we identified 6 drugs that may represent novel drugs for MS. Conclusion: Collectively, our results provide novel insight for the discovery of biomarkers and therapeutics for MS and provide a suitable foundation from which to design future investigations of the pathogenic mechanisms associated with MS.

Keywords: Multiple sclerosis, ceRNA, lncRNA, ceRNA network, biomarker

Introduction

Multiple sclerosis (MS) is a neuroinflammatory autoimmune disease that is characterized by white matter demyelination of the central nervous system (CNS), predominantly driven by myelin-specific immune T cells [1]. The main clinical features of this disease include the distribution of multiple lesions within the white matter and characteristic episodes of relapse and remission throughout the course of disease [2]. The mean global prevalence of MS is 33 per 100,000 individuals, with a higher prevalence in North America and Europe (140 and 108 per 100,000 individuals) than in Asian and sub-Saharan African countries (2.2 and 2.1 per 100,000 individuals, respectively) [3]. Statistics show that the global incidence of MS is increasing on an annual basis [4].

The underlying cause of MS has yet to be elucidated, although it is generally accepted that an interplay of both genetic and environmental factors may affect an individual's disease predisposition for MS. Within the genetically susceptible population, the occurrence of disease is primarily determined by major histocompatibility complex, while modifiable environmental factors, for example, smoking, Epstein-Barr virus infection, increased body mass index (BMI) dur-

ing adolescence, and low level of vitamin D, may affect whether an individual will develop MS [5]. The first genetic risk factor was discovered decades ago at the human leukocyte antigen (HLA) locus which is known to encode molecules that are involved in vital immune functions [6]. Notably, it has been suggested that certain environmental factors, such as smoking, can amplify pathogenic gene expression patterns in the CNS or immune cells *via* epigenetic mechanisms, thus synergizing with risk loci for MS [7].

Over recent years, a number of studies have supported the important role of non-coding RNAs, such as miRNAs and lncRNAs, in terms of the differentiation, dysfunction, and disproportionality of immune cells, as well as in their autoimmune and inflammatory response [8, 9], thus suggesting that these molecules may play key roles in MS. Moreover, lncRNAs are known to function as miRNA 'sponges' to compete with mRNAs and thereby regulate their activity [10]. A growing body of evidence now supports the fact that lncRNAs function as ceRNAs to regulate the expression levels of mRNAs, thereby participating in the progression of numerous diseases [11]. For example, lncRNA Gm15575 and PVT1 were shown to be aberrantly expressed in MS patients; furthermore, lncRNA Gm15575 and PVT1 have been shown to affect the functionality of Th17 in MS *via* ceRNA patterns [12, 13]. This indicates that lncRNA-ceRNA networks exert influence on the immune response in MS. A subsequent study also supported this notion by demonstrating that TUG1/miR-9-5p/NFKB1 (p50) ceRNET helped to regulate the immune response in MS [14]. The lncRNA taurine upregulation gene 1 (TUG1) was also shown to be up-regulated in serum samples taken from patients with MS [9]. Another study showed that NFKB1 was regulated by miR-9-5p by a sponging mechanism [14]. Studies have also shown that the down-regulation of TUG1 reduced the level of pro-inflammatory cytokines *in vivo* and increased the level of the anti-inflammatory cytokine IL-10 in EAE mice. Disturbances in the regulatory mechanisms associated with ceRNAs are known to cause immunological disorders and other diseases. It is evident that ceRNA networks are multifactorial and may represent a useful resource for identifying new biomarkers and therapeutic agents for different diseases.

In this study, we analyzed gene expression profiles and miRNA expression profiles acquired from two GEO datasets featuring patients with MS. This allowed us to determine differentially expressed mRNAs and miRNAs (DEmRNAs and DEmiRNAs, respectively). Next, we used bioinformatics to analyze gene functionality, pathway enrichment, PPI network, and network modular structure, to identify hub genes related to the immunological and inflammatory status of patients with MS. Next, we used a multistep computational approach that was based on mRNA-miRNA and miRNA-lncRNA interactions and the 'ceRNA hypothesis' to construct an MS-specific lncRNA-mediated ceRNA network (MCEN). Next, according to the modularization and topological properties of the PPI network, we constructed a ceRNA subnetwork. Then, we analyzed hub lncRNAs-mediated ceRNAs and identified potential therapeutic drugs that could be used to treat MS. Finally, we dissected the action of these hub lncRNA-mediated ceRNAs in the pathogenesis of MS (refer to the flowchart shown in **Figure 1**). Our aim was to provide new insight into the pathogenesis of MS and provide a resource for identifying new biomarkers and therapeutics for MS.

Materials and methods

Data acquisition

Gene expression (GSE21942) and miRNA expression (GSE61741) datasets were downloaded from the Gene Expression Omnibus (GEO, https://www.ncbi.nlm.nih.gov/geo/) database. GSE21942 featured data acquired from the peripheral blood mononuclear cells (PBMCs) of 12 MS patients and 15 controls (platform: GPL570 Affymetrix Human Genome U133 Plus 2.0 Array) while the GSE61741 dataset featured data acquired from the PBMCs of 23 MS patients and 94 controls (platform: GPL9040 *Homo sapiens* miRBase 13.0). Then, we identified mRNA-miRNA interaction pairs using the miRwalk database (version 2.0; zmf.umm. uni-heidelberg.de/apps/zmf/mirwalk2) [15]; this software predicts miRNAs that are likely to interact with differentially expressed genes. mRNA-miRNA interactions were then filtered by DEmiRNAs for MS. Finally, lncRNA-miRNA interactions were extracted from starBase [16], DIANA-LncBase [17], and LncACTdb [18]; these databases all feature miRNA-lncRNA

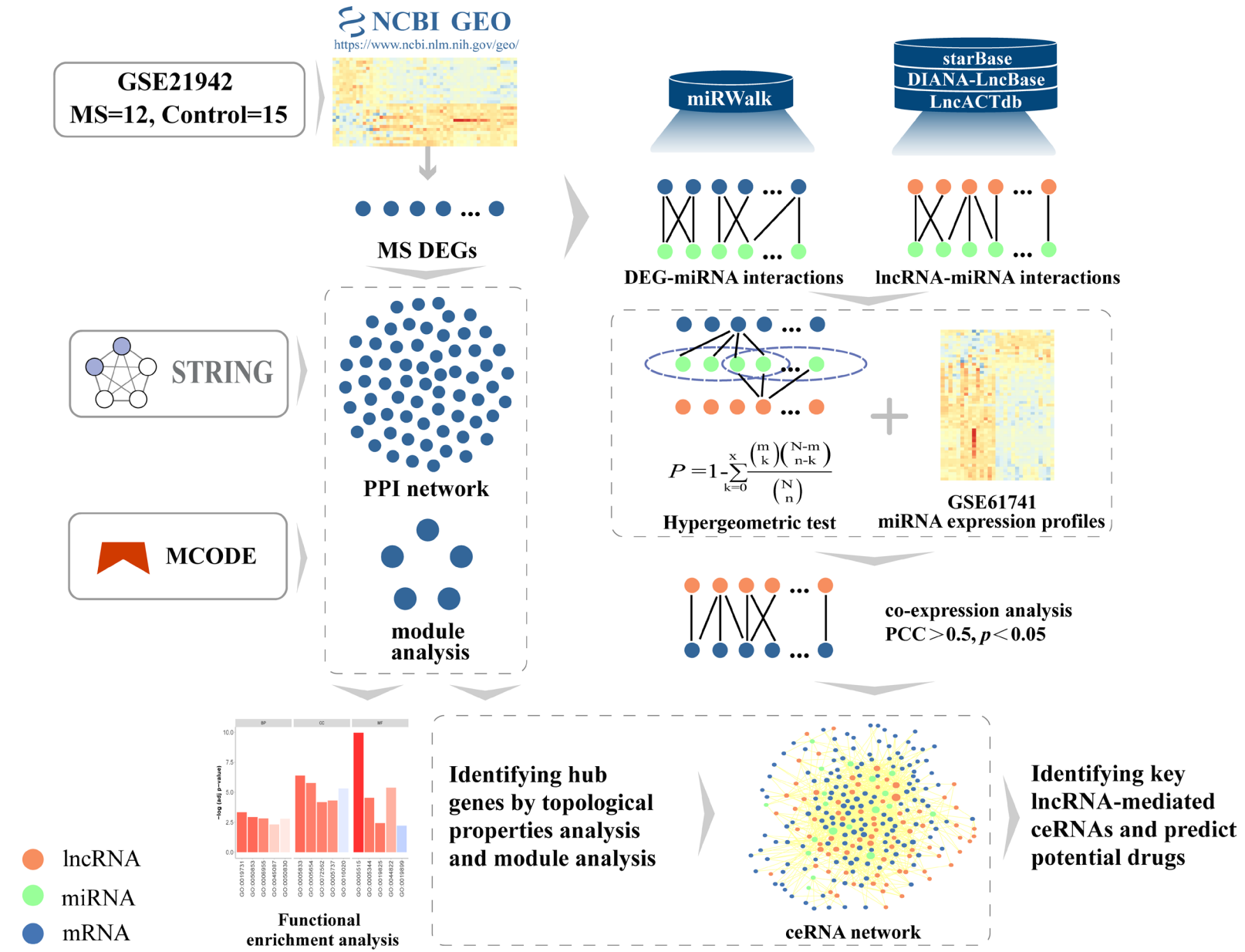

Figure 1. The workflow of this study showing steps containing the construction of PPI and ceRNA networks and identification of hub lncRNAs and potential drugs.

interactions that have been validated experimentally.

Data processing and the analysis of differential expression

GEO2R (http://www.ncbi.nlm.nih.gov/geo/geo2r/) online software was used to analyze the raw submitter-supplied data from microarrays and subsequently identify DEmRNAs and DEmiRNAs in MS. GEO2R is an interactive web tool that allows users to compare different groups of samples across experimental environments in a GEO series to identify genes that are differentially expressed. We used *P*-value <0.05 and |lgFC| >1.2 as the cut-off criteria to identify DEmRNAs and DEmiRNAs.

Functional enrichment analyses

The DAVID Database (https://david.ncifcrf.gov/) was used to perform functional and pathway enrichment analysis. The DAVID resource provides systematic and integrated functional annotation tools for researchers to determine the biological significance of different genes [19]. We used DAVID to analyze the DEmRNAs we identified in MS patients and performed specific analyses for gene ontology (GO), including biological process (BP), cellular component (CC), and molecular function (MF) [20]. We also performed Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis [21]. *P*-value <0.05 was considered to be statistically significant.

Protein-protein interaction (PPI) network construction and module analysis

Functional PPI analysis is essential if we are to interpret the molecular mechanisms of key cellular activities. In this study, we used the Search Tool for the Retrieval of Interacting Genes (STRING; http://string.embl.de/) database [22] and Cytoscape v3.8.1 (http://www.cytoscape.org/) software to construct a PPI network featuring the DEmRNAs we identified in patients with MS. An interaction score of 0.4 was regarded as the cut-off criterion and a PPI was generated. Cytoscape software was used to analyze a range of topological features for the nodes in the PPI network; this allowed us to identify hub genes, including degree and betweenness centrality. Then, we used the Molecular Complex Detection (MCODE; version 1.4.2; http://apps.cytoscape.org/apps/mcode) tool in Cytoscape software to select highly interconnected modules from the PPI network using specific selection criteria (MCODE degree cutoff =2; node score cutoff =0.2; k-core =2; max. depth =100).

Hypergeometric test

Next, we identified competing mRNA-lncRNA interaction pairs that shared the same miRNA. To do this, we used hypergeometric tests to identify competing pairs based on the common miRNAs of any pair of mRNAs and lncRNAs [23]; *P* values were computed using the formula given in Equation (1).

$$P = 1 - \sum_{k=0}^{x} \frac{\binom{m}{k}\binom{N-m}{n-k}}{\binom{N}{n}} \quad (1)$$

For each interaction pair, *N* denotes the total number of miRNAs in the interaction data, n and m denote the number of miRNAs that were associated with one mRNA and one lncRNA, and x represents the number of miRNAs shared with the mRNA and lncRNA. mRNA-lncRNA competitive interaction pairs with *P*-value <0.05 were considered to represent potential ceRNA pairs.

Co-expression correlation analysis for ceRNA interactions

Next, to identify lncRNA-mRNA interaction pairs, we applied co-expression correlation analysis to the lncRNA-mRNAs by using Pearson correlation coefficients (PCC) and by considering the expression of potential ceRNA interactions. PCC were calculated using the formula given in Equation (2).

$$\rho_{X,Y} = \frac{cov(X,Y)}{\sigma_X \sigma_Y} \quad (2)$$

In Equation (2), cov *(X, Y)* referred to the covariance of variables *X* and *Y*. σ_X and σ_Y represented standard deviations for *X* and *Y*. As a result, the final co-expressed lncRNA-mRNA pairs that met the hypergeometric test threshold (*P*-value <0.05) and crossed the PCC threshold (PCC >0.5 and *P*-value <0.05) were considered to be statistically significant pairs.

Construction of a ceRNA network

Next, we constructed a lncRNA-mediated ceRNA network based on the ceRNA interactions that were identified by the two-step method described above and identified 773 lncRNA-mediated ceRNA pairs. These were subsequently used to establish a MS-specific ceRNA network (MCEN) that was visualized Cytoscape software. Using the Cytoscape Network Analyzer tool, we analyzed a range of topological features with regards to the nodes of the MCEN, including degree and betweenness distribution analysis.

Identifying ceRNA subnetworks and potential drugs in MS

The hub immune-related DEmRNAs in MS patients that were identified by PPI network and module analysis were subsequently overlapped within an overall ceRNA network. The ceRNA subnetwork that included hub immune-related DEmRNAs was then systematically extracted to define a potential MS immune-relevant ceRNA subnetwork. Then, drugs and target genes were downloaded from DrugBank [24] (https://www.drugbank.ca/).

Results

Identification of DEmRNAs and DEmiRNAs in MS

Each set of array data was analyzed separately by the GEO2R system to identify DEmRNAs or DEmiRNAs (Supplementary Tables 1, 2). We successfully identified 290 DEmRNAs from the GSE21942 dataset; of these, 190 DEmRNAs were up-regulated and 100 DEmRNAs were down-regulated. We also identified 119 DEmiRNAs from the GSE61741 dataset; of these, 57 DEmiRNAs were up-regulated and 62 DEmiRNAs were down-regulated. DEmRNAs and DEmiRNAs were identified by applying specific criteria: a |lgFC| >1.2 and P<0.05. A heatmap was then generated to depict the expression tendencies of these DEmRNAs between MS and healthy controls (**Figure 2A**). In addition, volcano plots were generated to demonstrate the distribution of DEmRNAs and DEmiRNAs (**Figure 2B**).

Functional enrichment analysis of DEmRNAs

To identify the potential biological functions of the DEmRNAs, we performed GO functional enrichment analysis for BP, CC, and MF. We also used the DAVID database to perform KEGG analysis on the DEmRNAs [25]. The top five significant terms for GO functional enrichment (BP, CC and MF) and KEGG pathways (P<0.05) that might play pivotal roles in the immunological mechanisms associated with MS were shown in (**Figure 3**; Supplementary Table 3). GO analysis showed that the DEmRNAs were mainly enriched in biological processes associated with antibacterial humoral responses, the B cell receptor signaling pathway, and immune responses. KEGG analysis showed that the DEmRNAs were mainly enriched in viral myocarditis, the B cell receptor signaling pathway and influenza A. We demonstrated that the majority of GO terms were involved in immune response and many pathways, including the B cell receptor signaling pathway and influenza A, have been reported to be relevant with MS.

Generation of a PPI network, module analysis and hub gene selection

The STRING database was used to identify potential interactions of the DEmRNAs. A total of 680 PPI interactions and 217 DEmRNAs were used to construct a PPI network (**Figure 4A**). Twenty genes were identified as hub genes based on two vital topological features: degree and betweenness centrality (BC) (**Table 1**). Next, we used the MCODE plug-in to analyze the network and identified the top four modules (**Figure 4B-E**). The majority of the hub genes (CD19, B2M, CD79A, HNRNPH1, ICAM1, HBB, CD20, DDX3X, ALAS2, BLK, UBE2M and DDX17) were enriched in these four significant modules. Next, functional analysis of the genes associated with key modules revealed that they were primarily associated in the activation of neutrophils, the immune response, B cell activation, B cell proliferation, and the B cell receptor signaling pathway (**Figure 5**; Supplementary Tables 4, 5); it was evident that all of these functions are associated with immunological status. Collectively, these findings suggest that these hub genes may play a role in the immunological and inflammatory pathogenesis of MS.

Construction of a MS-related lncRNA-associated ceRNA network and topological analysis

It is well known that lncRNAs can act as 'sponges' for miRNAs to regulate mRNAs in various diseases. To identify the regulatory roles of

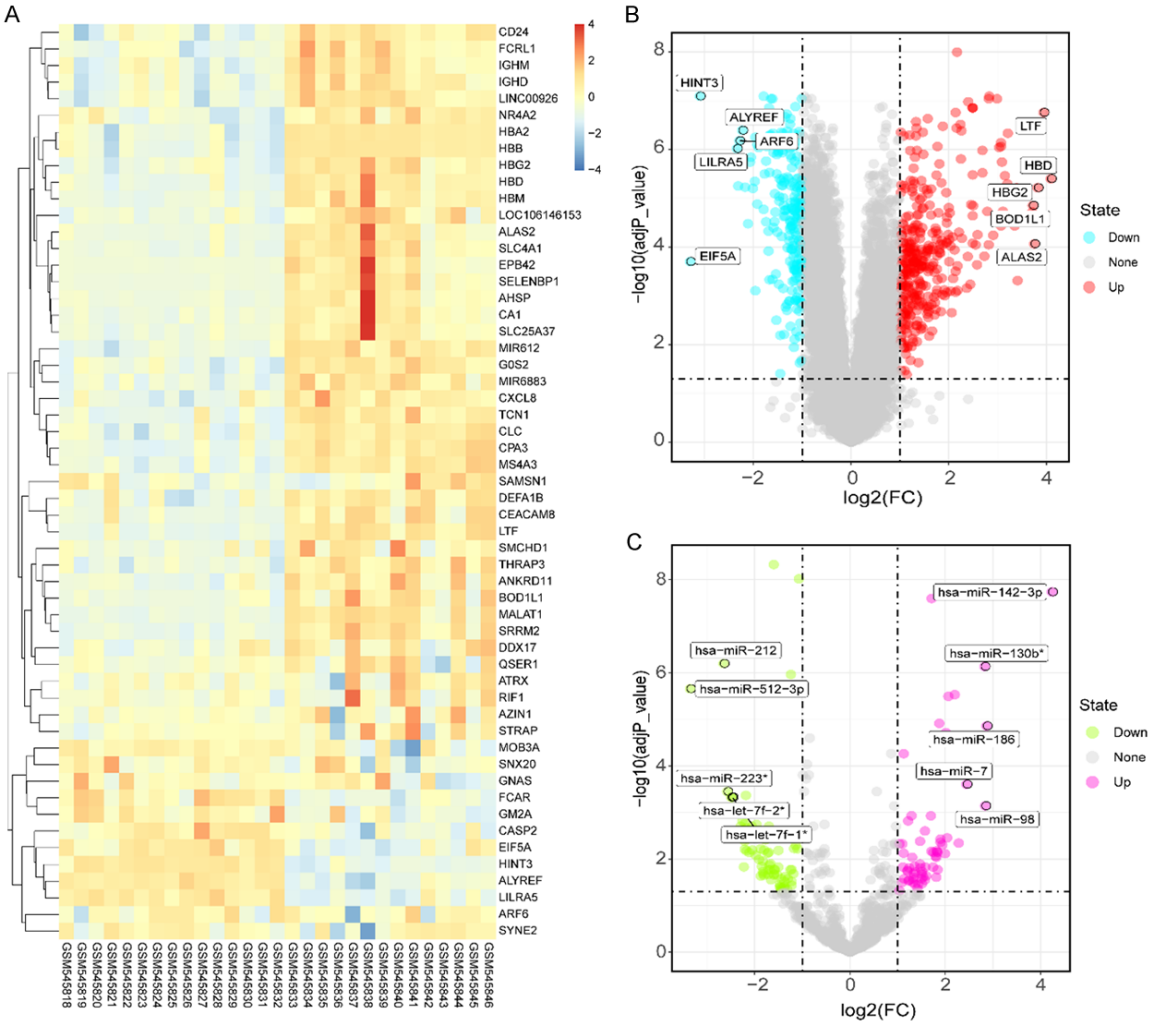

Figure 2. Analysis of DEmRNAs and DEmiRNAs. (A) Hierarchical heatmap for the top 55 differentially expressed mRNAs in multiple sclerosis. Volcano plots of mRNAs (B), miRNAs (C) with |lgFC| >2 (adjust *P* value <0.05).

lncRNAs in MS, we assembled 773 lncRNA-mediated ceRNA interactions to construct a MCEN that consisted of 250 nodes and 1158 edges (including 78 lncRNAs, 153 mRNAs, 19 miRNAs, 700 lncRNA-mRNA interactions, 99 lncRNA-miRNA interactions, and 359 mRNA-miRNA interactions) (**Figure 6A**; Supplementary Table 6). Subsequently, we analyzed the topological properties of the MCEN. We found that the degree distribution of the nodes in the MCEN closely followed a power law distribution that was defined by $f(x) = 130.55x^{-1.288}$ (R^2=0.8759), thus suggesting that the MCEN was a scale-free network (**Figure 6B**). We also calculated the betweenness of nodes in the MCEN (**Figure 6C**) and found that the higher the betweenness of a node, the more significant the node was in terms of maintaining tight connectivity within the network.

Furthermore, comparative analysis revealed significant differences in terms of degree distribution when compared between the mRNAs, miRNAs and lncRNAs (P<0.05, Kruskal-Wallis test) (**Figure 6D**). Collectively, these findings revealed that lncRNAs and miRNAs both exhibited significant high degrees, thereby indicating that they play vital roles in the MCEN.

Identification and analysis of a hub lncRNA-mediated ceRNA subnetwork in MS

Next, we used hub immune-related mRNAs from our PPI and modules to reconstruct a ceRNA subnetwork that contained 23 lncRNAs,

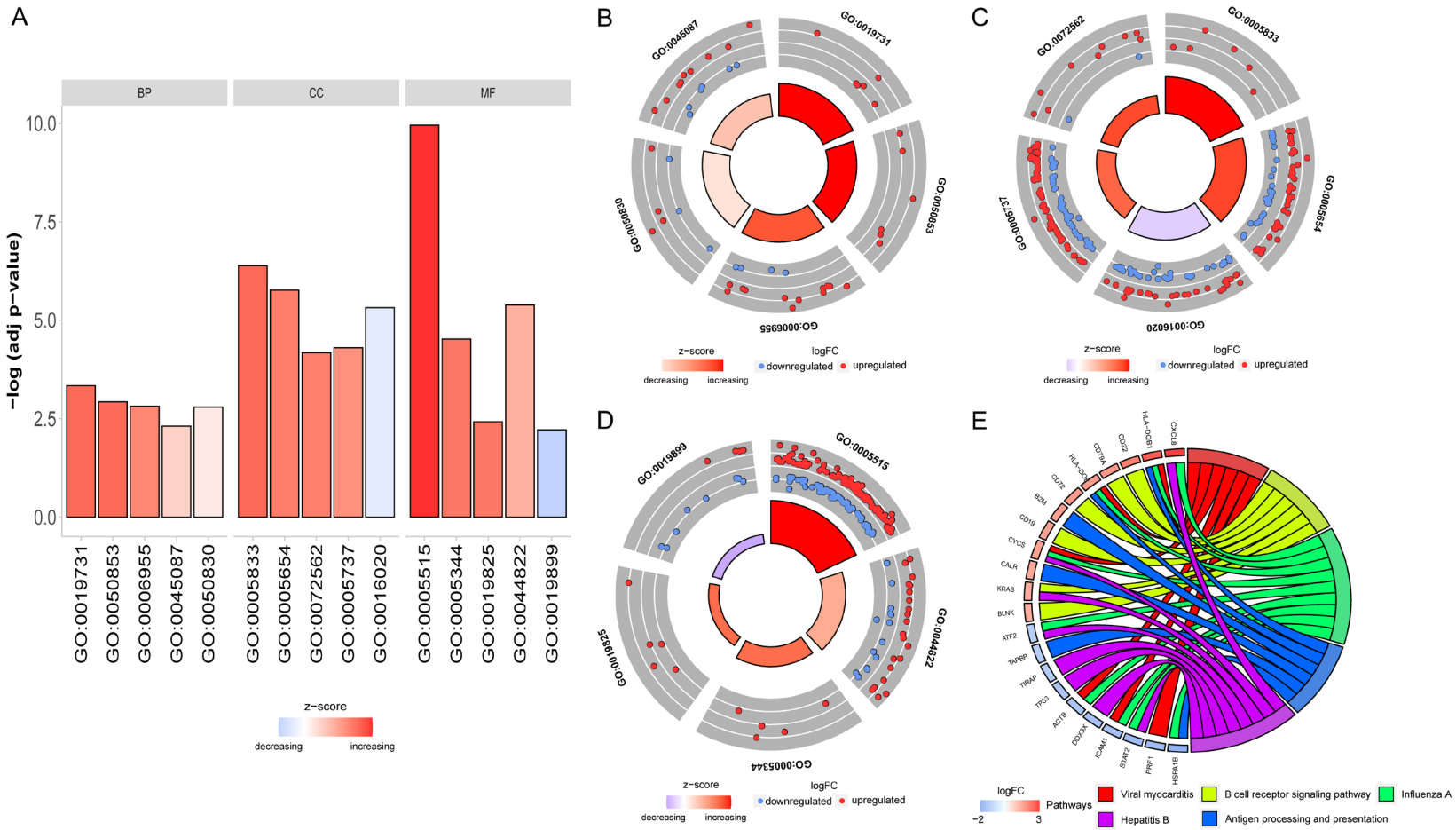

Figure 3. GO functional and KEGG pathway analysis of DEmRNAs. (A) Bar graph of BP, CC, MF and KEGG payhway. Distribution of DEmRNAs in each GO term of BP (B), CC (C) and MF (D). Red denotes upregulated expression genes; Blue denotes downregulated expression genes. (E) KEGG pathway analysis of DEmRNAs. Different colors represent different pathways.

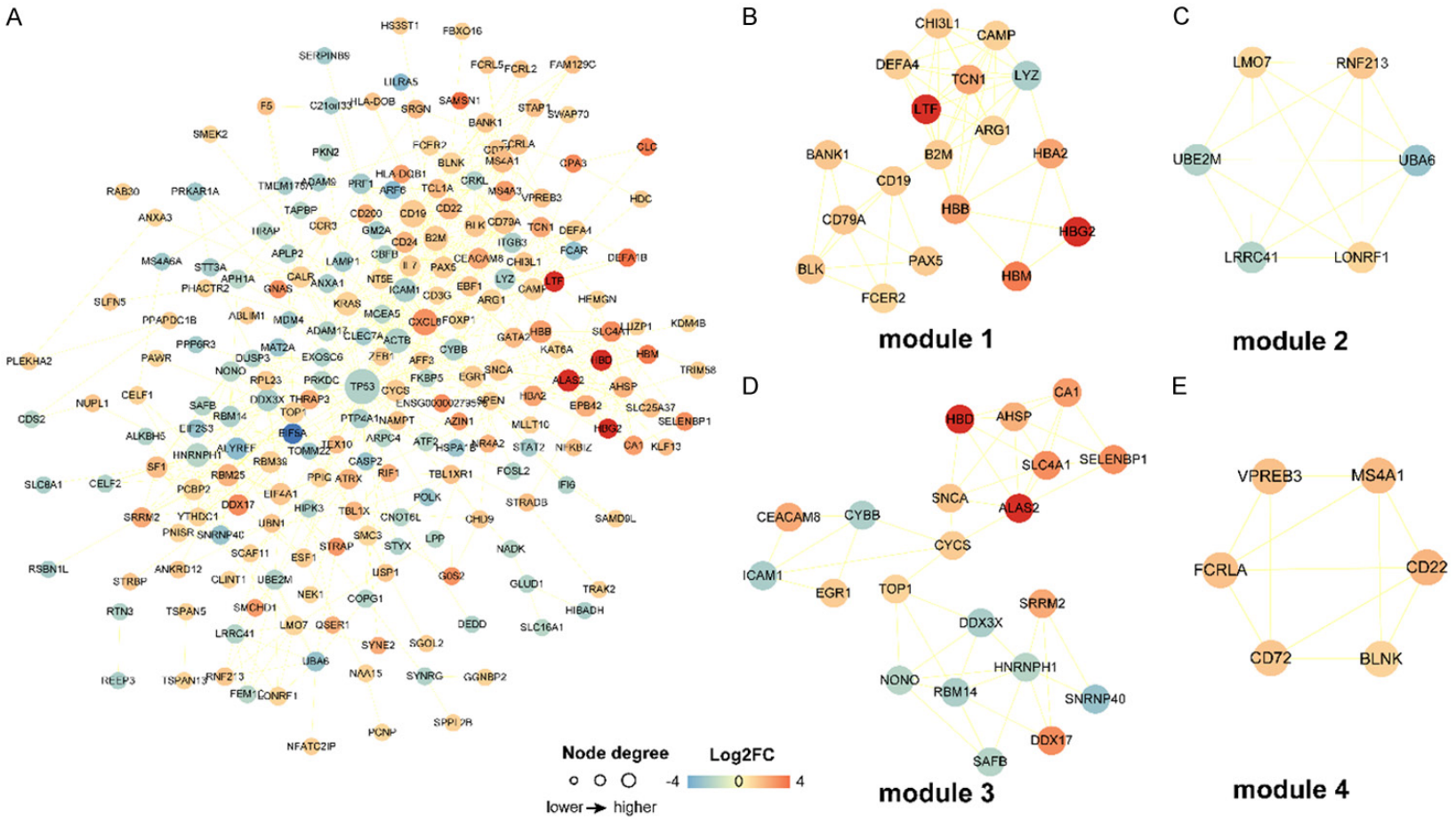

Figure 4. PPI network and modules of DEmRNAs in multiple sclerosis. (A) The nodes denote mRNAs in PPI network (confidence score >0.4). The bigger and darker the node is, the higher degree the node has. The expression level of mRNAs is shown using different colors; red circle represents upregulated mRNAs and blue circle represents downregulated mRNAs. The top four modules, module1 (B), module 2 (C), module 3 (D), module 4 (E) from PPI network are modularized by MCODE.

Table 1. The top 20 hub genes identified in the protein-protein interaction network from differentially expressed genes in MS

Gene	Degree	Gene	Betweenness Centrality
TP53	51	TP53	0.448331
CD19	28	B2M	0.083457
CXCL8	26	CXCL8	0.071215
B2M	22	ACTB	0.068447
ACTB	22	CD19	0.063355
CD79A	19	HBB	0.054128
HNRNPH1	19	KRAS	0.053122
ICAM1	19	HNRNPH1	0.050197
KRAS	19	CALR	0.050077
HBB	16	NAMPT	0.048897
MS4A1	15	DDX3X	0.044766
RBM25	15	UBE2M	0.044695
DDX3X	15	SMC3	0.042802
ALAS2	15	CD79A	0.038105
BLK	15	DDX17	0.038031

seven miRNAs and four mRNAs (**Figure 7A**). The 23 lncRNAs involved in this subnetwork were identified as key lncRNAs. For further analysis, all lncRNA nodes in the ceRNA subnetwork were ranked in descending order in terms of their degree and BC, respectively. The higher degree and BC of a node in the subnetwork, the more likely the node was related to MS. Next, we identified the top three lncRNAs that exhibited the highest degree and the highest BC; these were determined to be key lncRNA ceRNAs: LINC00649, TP73-AS1, MALAT1 (**Table 2**). We found that these three hub lncRNAs were related to 90% of the DEmRNAs and 60% of the DEmiRNAs within the subnetwork. These findings suggested that these hub lncRNA ceRNAs are important regulators that can influence ceRNA mechanisms during the immunological and inflammatory pathogenesis of MS.

Previous studies reported that these three lncRNAs play crucial roles in a variety of different diseases by acting on ceRNA, including cancer and autoimmune diseases. For instance, Guo et al. showed that LINC00649 acts as a ceRNA for miR-424-5p during the progression of acute myeloid leukemia [26]. In another study, MALAT1 was shown to act as a ceRNA by sponging the miRNA miR-338-3p, indirectly inducing MSL2 expression in MG [27]. A more recent study showed that TP73-AS1 could inhibit the growth of colorectal cancer cells by functioning as a ceRNA to regulate the expression levels of PTEN [28]. Collectively, these studies clearly demonstrated that the hub lncRNAs identified in our present study are clearly important regulators in a range of diseases.

The prediction of potential drugs for the treatment of MS

Emerging evidence has demonstrated that ncRNAs could be used as drug targets for the treatment of various diseases [29]. Next, we performed drug analysis using our hub lncRNA-mediated ceRNA subnetwork and the three key lncRNA ceRNAs identified previously. To do this, we analyzed the co-expression of the three lncRNAs (LINC00649, DPP9-AS1, TP73-AS1) with mRNAs from the subnetwork were to identify potential drugs in the DrugBank database. We discarded illegal drugs that are not available on the market, had not been approved by the FDA, need to be mixed with other products or drugs, lack APRD, or are vaginally administered. Following the application of these criteria, we screened 7 potential drugs. Finally, we identified these 7 drugs that involved in LINC00649 mediated ceRNAs (**Figure 7B**). Of these 7 drugs, rituximab, ibritumomab tiuxetan, ofatumumab, obinutuzumab, nofetumomab and ocrelizumab, were shown to be targeted by LINC00649-miR-1275-CD20 regulator pairs. A growing body of evidence now supports the fact that that rituximab can be used as a treatment option for MS [30] and that ofatumumab, as an anti-CD20 monoclonal antibody, could selectively deplete B cells [31] as a novel therapy for the treatment of MS treatment. Collectively, our analysis identified a number of novel therapeutic targets and drugs for the treatment of MS.

Discussion

MS is a complex neuroinflammatory autoimmune disease of the CNS. However, the pathogenic mechanisms underlying this condition have yet to be fully elucidated [32]. A recent study reported that interleukin (IL)-9 and transforming growth factor (TGF)-β were both expressed by immune cells and play key roles in regulating the pathogenesis of MS [33]. In

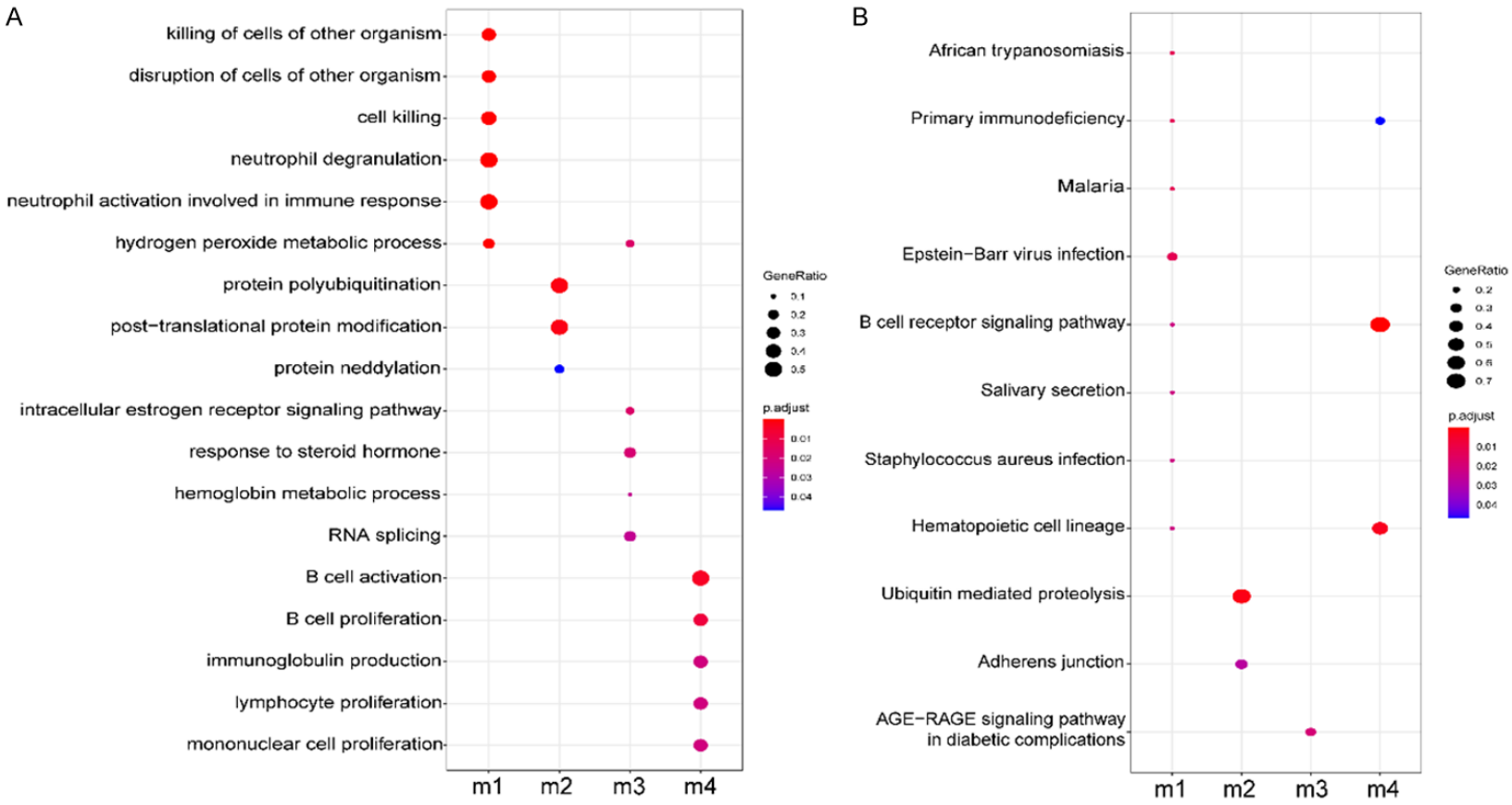

Figure 5. Functional enrichment analysis of top 4 modules. A. GO biological process analysis of modules. The bigger circles suggest that a more proportion of MS DEmRNAs in a module among GO function genes. B. KEGG pathway analysis of modules. The bigger circles suggest that a more proportion of MS DEmRNAs in a module among KEGG pathway genes.

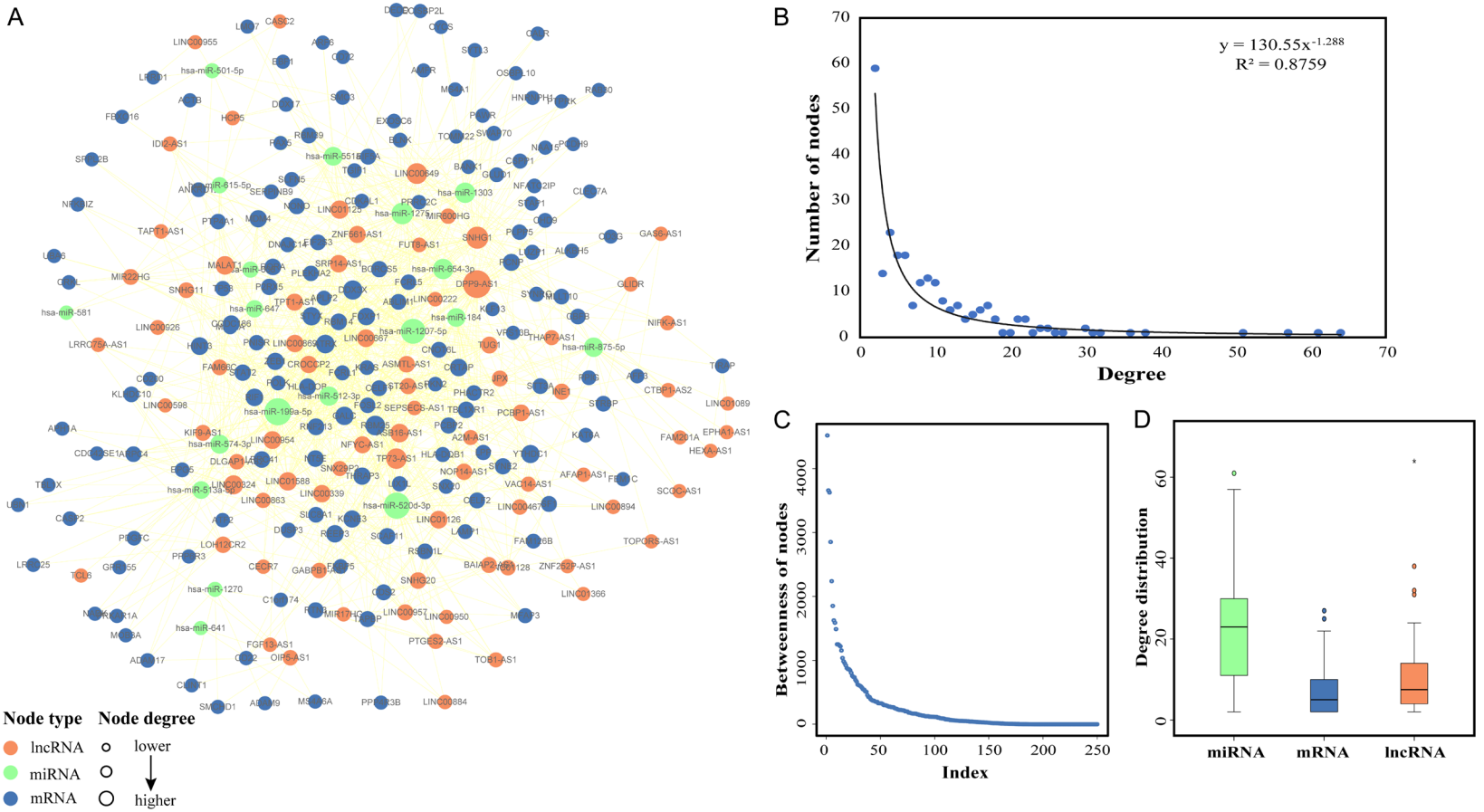

Figure 6. Construction of MCEN by DEmRNAs and analysis of topological properties. A. The MS specific ceRNA network (MCEN). Bule nodes represent mRNAs, green nodes represent miRNAs, orange nodes represent lncRNAs and lines between them represent their interactions, the bigger the node is, the higher degree the node has. B. The nodes degree distribution of the MCEN. C. The nodes betweenness distribution of the MCEN. D. The degree distribution of miRNAs, mRNAs and lncRNAs in MCEN.

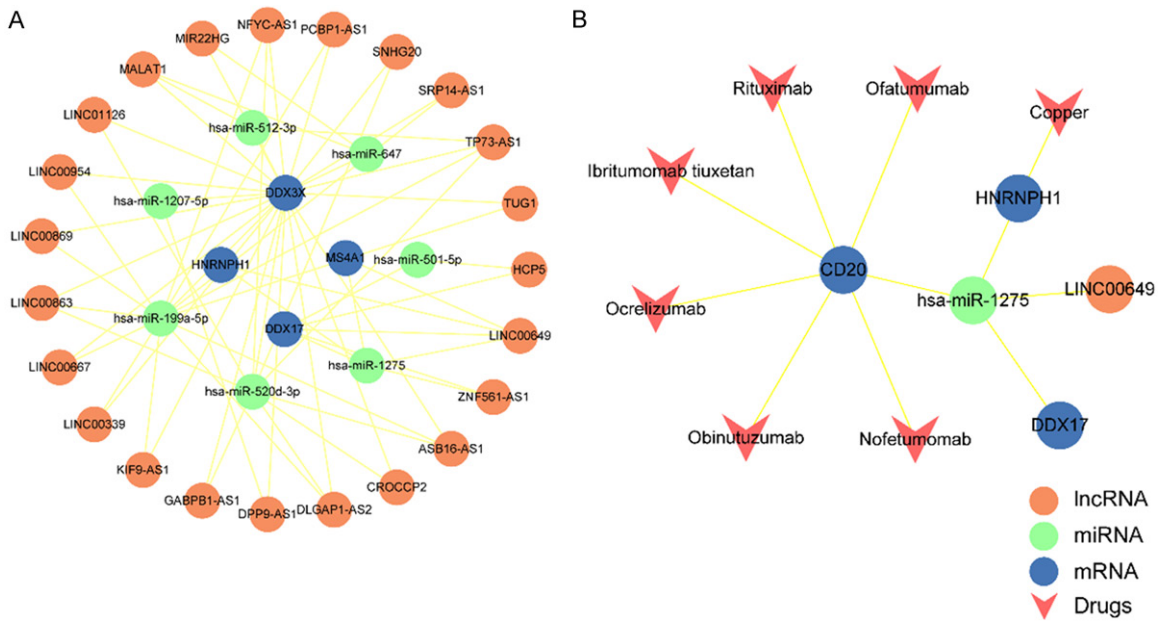

Figure 7. Construction of ceRNA subnetwork and associated drugs. A. The ceRNA subnetwork based on modularization of PPI network. B. Screened drugs targeted LINC00649-mediated ceRNAs. Blue circle represents mRNA, green circle represents miRNA, orange circle represents lncRNA and pink "V" represents drug.

Table 2. The top 3 lncRNAs ceRNAs ranked by degree and BC

LncRNA	Degree	LncRNA	Betweenness Centrality
LINC00649	4	LINC00649	0.1667
TP73-AS1	4	TP73-AS1	0.0039
MALAT1	3	MALAT1	0.0017

addition, an extensive body of evidence now supports the fact that lncRNA-mediated ceRNA regulation represents a new post-transcriptional layer of gene regulation that plays important roles in the immune system and can exert influence on the pathological processes of numerous autoimmune diseases [34, 35]. However, we have very limited understanding of the molecular mechanisms underlying the effects of lncRNA ceRNAs on MS. Elucidating the precise mechanisms responsible for the occurrence and development of MS will be highly beneficial to the diagnosis and treatment of MS patients as well as predicting their prognosis.

In this study, we integrated high-throughput expression profiles from MS patients and used the STRING tool to establish a PPI network to screen modules and extract hub DEmRNAs that might be associated with immune response and inflammatory status. Next, we constructed a MCEN that allowed us to investigate the regulatory role of lncRNAs and ceRNAs in the development of MS, as determined by miRNA-target interactions from multiple databases. According to the 'ceRNA hypothesis', the expression of mRNAs could be regulated by lncRNAs, thus indicating that the biological function of lncRNAs may be similar to their co-expressed mRNAs. Then, we applied PPI network and module analysis to identify hub DEmRNAs and construct a subnetwork from an entire ceRNA network. Based on further analysis of the subnetwork, we identified three hub lncRNAs ceRNAs (LINC00649, TP73-AS1, MALAT1) that might act as key lncRNAs in the immunological pathogenesis of MS. Moreover, recent studies have also demonstrated that MALAT1 can regulate both the expression of splicing factors and MS-related alternative splicing events, thus suggesting that MALAT1 participates in the pathogenesis of MS and could be applied as a new therapeutic agent for patients with MS [36, 37].

GO and KEGG functional analysis further showed that the significant modules were closely related to B cells, immune response, and the proliferation of lymphocytes; these processes play a key role in the pathology of MS. Moreover, B cells are becoming increasingly investigated as mediators of inflammation in MS [38]. Modularization can dissect complex

network into specific modules for further research; the genes within these modules may play vital functional roles. In this study, we found that lncRNAs were associated with the hub genes in modules *via* the ceRNA network and may play key regulatory roles in terms of biological function. Therefore, we could select the lncRNAs involved in the ceRNA subnetwork as important lncRNAs for further analysis. By analyzing the subnetwork, we identified three hub lncRNAs-mediated ceRNA interaction pairs featuring four hub genes (CD20, HNRNPH1, DDX17 and DDX3X). Several previous studies have shown that these hub genes, particularly HNRNPH1 and CD20, participate in the immune pathogenesis of MS [39, 40]. In addition, an anti-CD20 monoclonal antibody has been used as a novel therapeutic for the treatment of MS as it can selectively deplete B cells [41]. Identification of the functional roles of these genes and co-expressed lncRNAs may provide us with new insights into the pathogenesis of MS.

Next, we dissected the topological characteristics of the network and the interrelationships of the regulatory ceRNAs in MS, particularly with regards to the potential ceRNA mechanism for LINC00649. We identified that the dysregulation of LINC00649-mediated ceRNA interactions, containing three mRNAs and one miRNA, might participate in the immunological pathogenesis of MS either individually or synergistically. It is possible that the up-regulated LINC00649 was targeted by a greater number of miRNAs; therefore, there were fewer free miRNAs available to combine with mRNAs, thus leading to the up-regulation of mRNAs. The mRNAs that were affected by LINC00649 might be associated with an increased risk of MS. Therefore, the LINC00649-miR-1275-CD20 ceRNA network is highly likely to play specific roles in MS. To further highlight the importance of this ceRNA mechanism, we provided an example of how the inhibitory effect of miRNAs on lncRNAs and mRNAs expression (**Figure 8**) might be involved in the pathogenesis of MS by acting on mRNAs and other key pathways. As shown, the up-regulation of lncRNA would lead to a greater extent of miRNA binding, thus resulting in the down-regulation of miRNA; this would reduce the free miRNA repressed expression of mRNA, thus enhancing the expression of mRNA. For instance, as a gene target, CD20 received regulation through this ceRNA, which can be enriched in different pathways, such as B cell receptor signaling pathway and Epstein-Barr virus infection. This indicates an ongoing relationship between the mechanisms responsible for cell activation and proliferation that influence the immune response in the autoimmune conditions of MS. Furthermore, we used the ceRNA subnetwork to screen potential drug-targeting genes within the network. We identified six drugs that might be used to treat MS: rituximab, ibritumomab tiuxetan, ofatumumab, obinutuzumab, nofetumomab and ocrelizumab. It is well known that rituximab is an anti-CD20 monoclonal antibody that is used to treat many autoimmune diseases, including MS [30]. Moreover, based on the fact that the LINC00649-miR-1275-CD20 RNAs are known to be related to the immunological pathogenesis of MS, this LINC00649-mediated ceRNA pair could represent a novel and highly promising diagnostic biomarker and therapeutic target for MS.

Previous research has shown that LINC00649 is dysregulated in many diseases. For example, one study showed that LINC00649 was down-regulated in patients with acute myeloid leukemia (AML) and might represent an unfavorable prognostic biomarker for patients with AML [42]. In another study, Chen et al. found that LINC00649 was a target for miR-15a-5p and increased the expression of HMGA1, thus leading to the progression of malignancy in patients with bladder cancer [43]. Wang et al. suggested that LINC00649 could promote the progression of gastric cancer by binding to the miR-16-5p/YAP1/Hippo signaling pathway [44]. Collectively, these previous studies, and the data generated in the present study, demonstrated that the dysregulation of LINC00649-mediated ceRNA might play a crucial to role in the pathogenesis of MS.

In conclusion, in this study, we systematically identified several key immune-related DE-mRNAs in MS, integrated various RNAs to explore lncRNA/ceRNA regulatory mechanisms, and identified six potential drugs, thus providing new insight into the pathogenesis and treatment of MS. Our study focused on ceRNAs as a regulator because there is a dynamic balance of biological molecules within the human body that regulate physiological functions in a cooperative manner. Therefore, we focused on ceRNAs instead of single bio-

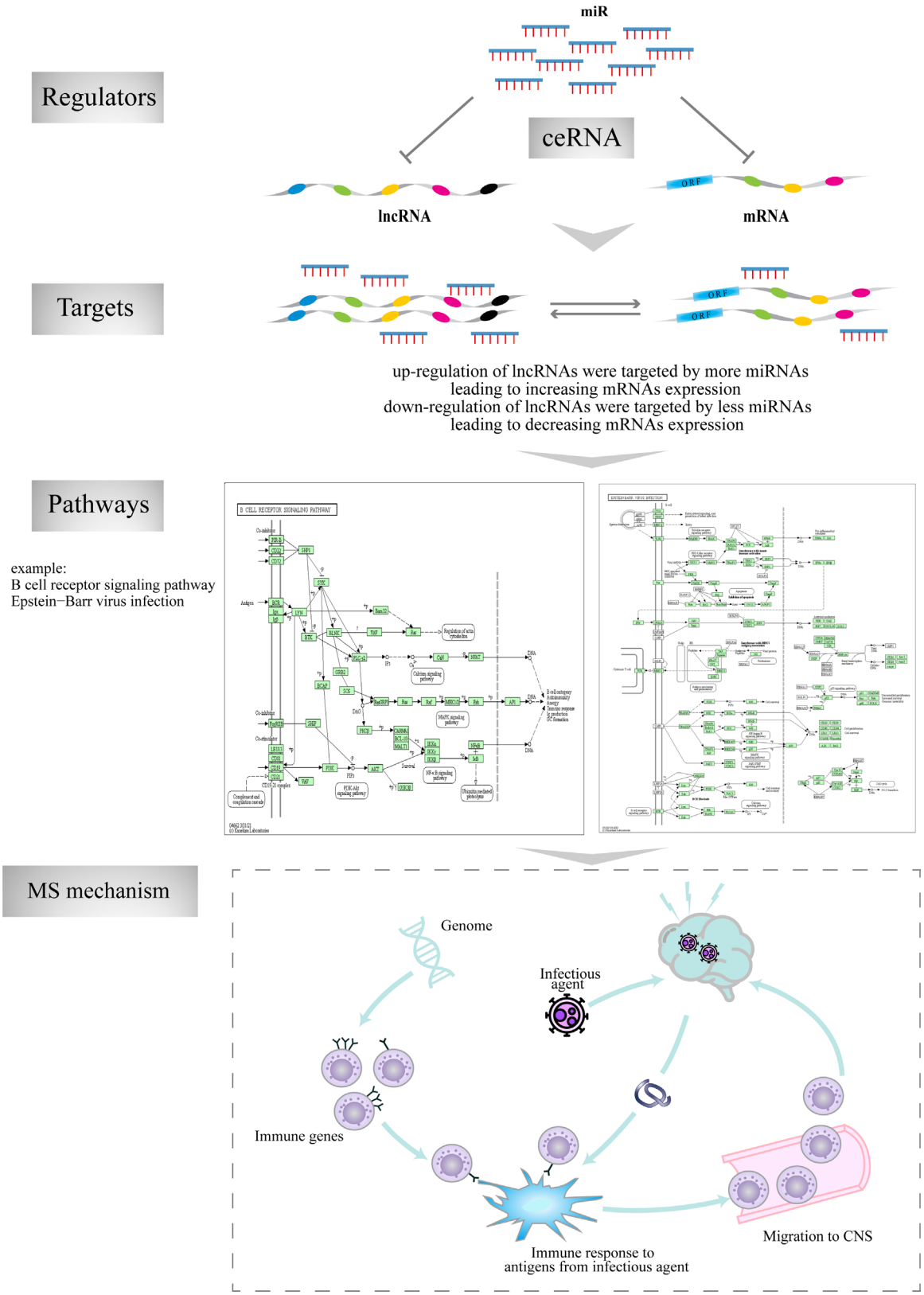

Figure 8. A model illustrating the ceRNA regulatory process. LncRNAs, miRNAs and genes cooperatively mediate pathways dysregulation that maybe influence the immune response and activation of immune cells in MS.

molecules to provide important evidence for lncRNAs as gene regulators in the pathogenesis of MS. New drugs screened through network will give us a novel insight in MS treatment. Thus, illustrating the ceRNAs mechanism will make a big difference in research for MS.

Acknowledgements

This study was supported by the National Natural Science Foundation of China (NSFC grant numbers 82071407, 81701190 and 81820108014).

Disclosure of conflict of interest

None.

Address correspondence to: Lihua Wang and Jianjian Wang, Department of Neurology, The Second Affiliated Hospital of Harbin Medical University, 246 Xuefu Road, Harbin 150081, Heilongjiang, China. E-mail: wanglh211@163.com (LHW); wangjian_427@163.com (JJW)

References

[1] Dendrou CA, Fugger L and Friese MA. Immunopathology of multiple sclerosis. Nat Rev Immunol 2015; 15: 545-558.

[2] Compston A and Coles A. Multiple sclerosis. Lancet 2008; 372: 1502-1517.

[3] Belbasis L, Bellou V, Evangelou E, Ioannidis JP and Tzoulaki I. Environmental risk factors and multiple sclerosis: an umbrella review of systematic reviews and meta-analyses. Lancet Neurol 2015; 14: 263-273.

[4] Koch-Henriksen N and Sørensen PS. The changing demographic pattern of multiple sclerosis epidemiology. Lancet Neurol 2010; 9: 520-532.

[5] Amato MP, Derfuss T, Hemmer B, Liblau R, Montalban X, Soelberg Sorensen P and Miller DH; 2016 ECTRIMS Focused Workshop Group. Environmental modifiable risk factors for multiple sclerosis: report from the 2016 ECTRIMS focused workshop. Mult Scler 2018; 24: 590-603.

[6] Jersild C, Svejgaard A and Fog T. HL-A antigens and multiple sclerosis. Lancet 1972; 1: 1240-1241.

[7] Olsson T, Barcellos LF and Alfredsson L. Interactions between genetic, lifestyle and environmental risk factors for multiple sclerosis. Nat Rev Neurol 2017; 13: 25-36.

[8] Sigdel KR, Cheng A, Wang Y, Duan L and Zhang Y. The emerging functions of long noncoding rna in immune cells: autoimmune diseases. J Immunol Res 2015; 2015: 848790.

[9] Santoro M, Nociti V, Lucchini M, De Fino C, Losavio FA and Mirabella M. Expression profile of long non-coding RNAs in serum of patients with multiple sclerosis. J Mol Neurosci 2016; 59: 18-23.

[10] Salmena L, Poliseno L, Tay Y, Kats L and Pandolfi PP. A ceRNA hypothesis: the Rosetta Stone of a hidden RNA language? Cell 2011; 146: 353-358.

[11] Zhou M, Wang X, Shi H, Cheng L, Wang Z, Zhao H, Yang L and Sun J. Characterization of long non-coding RNA-associated ceRNA network to reveal potential prognostic lncRNA biomarkers in human ovarian cancer. Oncotarget 2016; 7: 12598-12611.

[12] Bian Z, Lei W, Li Q, Xue W, Gao Y, Zeng Y, Wang Y, Tang L, Tang T, Chen C, Gao X and Guo W. Gm15575 functions as a ceRNA to up-regulate CCL7 expression through sponging miR-686 in Th17 cells. Mol Immunol 2020; 125: 32-42.

[13] Wu L, Xia J, Li D, Kang Y, Fang W and Huang P. Mechanisms of M2 macrophage-derived exosomal long non-coding RNA PVT1 in regulating Th17 cell response in experimental autoimmune encephalomyelitisa. Front Immunol 2020; 11: 1934.

[14] Yue P, Jing L, Zhao X, Zhu H and Teng J. Down-regulation of taurine-up-regulated gene 1 attenuates inflammation by sponging miR-9-5p via targeting NF-kappaB1/p50 in multiple sclerosis. Life Sci 2019; 233: 116731.

[15] Dweep H and Gretz N. miRWalk2.0: a comprehensive atlas of microRNA-target interactions. Nat Methods 2015; 12: 697.

[16] Li JH, Liu S, Zhou H, Qu LH and Yang JH. starBase v2.0: decoding miRNA-ceRNA, miRNA-ncRNA and protein-RNA interaction networks from large-scale CLIP-Seq data. Nucleic Acids Res 2014; 42: D92-97.

[17] Paraskevopoulou MD, Georgakilas G, Kostoulas N, Reczko M, Maragkakis M, Dalamagas TM and Hatzigeorgiou AG. DIANA-LncBase: experimentally verified and computationally predicted microRNA targets on long non-coding RNAs. Nucleic Acids Res 2013; 41: D239-245.

[18] Wang P, Li X, Gao Y, Guo Q, Wang Y, Fang Y, Ma X, Zhi H, Zhou D, Shen W, Liu W, Wang L, Zhang Y, Ning S and Li X. LncACTdb 2.0: an updated database of experimentally supported ceRNA interactions curated from low- and high-throughput experiments. Nucleic Acids Res 2019; 47: D121-D127.

[19] Dennis G Jr, Sherman BT, Hosack DA, Yang J, Gao W, Lane HC and Lempicki RA. DAVID: database for annotation, visualization, and integrated discovery. Genome Biol 2003; 4: P3.

[20] Gene Ontology Consortium. The Gene Ontology (GO) project in 2006. Nucleic Acids Res 2006; 34: D322-326.

[21] Kanehisa M and Goto S. KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res 2000; 28: 27-30.

[22] Szklarczyk D, Franceschini A, Wyder S, Forslund K, Heller D, Huerta-Cepas J, Simonovic M, Roth A, Santos A, Tsafou KP, Kuhn M, Bork P, Jensen LJ and von Mering C. STRING v10: protein-protein interaction networks, integrated over the tree of life. Nucleic Acids Res 2015; 43: D447-452.

[23] Zhang G, Sun H, Zhang Y, Zhao H, Fan W, Li J, Lv Y, Song Q, Li J, Zhang M and Shi H. Characterization of dysregulated lncRNA-mRNA network based on ceRNA hypothesis to reveal the occurrence and recurrence of myocardial infarction. Cell Death Discov 2018; 4: 35.

[24] Wishart DS, Feunang YD, Guo AC, Lo EJ, Marcu A, Grant JR, Sajed T, Johnson D, Li C, Sayeeda Z, Assempour N, Iynkkaran I, Liu Y, Maciejewski A, Gale N, Wilson A, Chin L, Cummings R, Le D, Pon A, Knox C and Wilson M. DrugBank 5.0: a major update to the DrugBank database for 2018. Nucleic Acids Res 2018; 46: D1074-D1082.

[25] Huang da W, Sherman BT and Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc 2009; 4: 44-57.

[26] Guo C, Ju QQ, Zhang CX, Gong M, Li ZL and Gao YY. Overexpression of HOXA10 is associated with unfavorable prognosis of acute myeloid leukemia. BMC Cancer 2020; 20: 586.

[27] Kong X, Wang J, Cao Y, Zhang H, Lu X, Wang Y, Bo C, Wang T, Li S, Tian K, Liu Z and Wang L. The long noncoding RNA MALAT-1 functions as a competing endogenous RNA to regulate MSL2 expression by sponging miR-338-3p in myasthenia gravis. J Cell Biochem 2019; 120: 5542-5550.

[28] Jia Z, Peng J, Yang Z, Chen J, Liu L, Luo D and He P. Long non-coding RNA TP73AS1 promotes colorectal cancer proliferation by acting as a ceRNA for miR103 to regulate PTEN expression. Gene 2019; 685: 222-229.

[29] Matsui M and Corey DR. Non-coding RNAs as drug targets. Nat Rev Drug Discov 2017; 16: 167-179.

[30] Ineichen BV, Moridi T, Granberg T and Piehl F. Rituximab treatment for multiple sclerosis. Mult Scler 2020; 26: 137-152.

[31] Hauser SL, Bar-Or A, Cohen JA, Comi G, Correale J, Coyle PK, Cross AH, de Seze J, Leppert D, Montalban X, Selmaj K, Wiendl H, Kerloeguen C, Willi R, Li B, Kakarieka A, Tomic D, Goodyear A, Pingili R, Haring DA, Ramanathan K, Merschhemke M, Kappos L and Asclepios I; ASCLEPIOS I and ASCLEPIOS II Trial Groups. Ofatumumab versus teriflunomide in multiple sclerosis. N Engl J Med 2020; 383: 546-557.

[32] Rogers KA and MacDonald M. Therapeutic yoga: symptom management for multiple sclerosis. J Altern Complement Med 2015; 21: 655-659.

[33] Donninelli G, Saraf-Sinik I, Mazziotti V, Capone A, Grasso MG, Battistini L, Reynolds R, Magliozzi R and Volpe E. Interleukin-9 regulates macrophage activation in the progressive multiple sclerosis brain. J Neuroinflammation 2020; 17: 149.

[34] Das S, Ghosal S, Sen R and Chakrabarti J. lnCeDB: database of human long noncoding RNA acting as competing endogenous RNA. PLoS One 2014; 9: e98965.

[35] Chen YG, Satpathy AT and Chang HY. Gene regulation in the immune system by long noncoding RNAs. Nat Immunol 2017; 18: 962-972.

[36] Shaker OG, Mahmoud RH, Abdelaleem OO, Ibrahem EG, Mohamed AA, Zaki OM, Abdelghaffar NK, Ahmed TI, Hemeda NF, Ahmed NA and Mansour DF. LncRNAs, MALAT1 and lncDC as potential biomarkers for multiple sclerosis diagnosis. Biosci Rep 2019; 39: BSR20181335.

[37] Dinescu S, Ignat S, Lazar AD, Constantin C, Neagu M and Costache M. Epitranscriptomic signatures in lncRNAs and their possible roles in cancer. Genes (Basel) 2019; 10: 52.

[38] Torke S and Weber MS. Inhibition of Bruton's tyrosine kinase as a novel therapeutic approach in multiple sclerosis. Expert Opin Investig Drugs 2020; 29: 1143-1150.

[39] von Essen MR, Ammitzboll C, Hansen RH, Petersen ERS, McWilliam O, Marquart HV, Damm P and Sellebjerg F. Proinflammatory CD20+ T cells in the pathogenesis of multiple sclerosis. Brain 2019; 142: 120-132.

[40] Cardamone G, Paraboschi EM, Solda G, Cantoni C, Supino D, Piccio L, Duga S and Asselta R. Not only cancer: the long non-coding RNA MALAT1 affects the repertoire of alternatively spliced transcripts and circular RNAs in multiple sclerosis. Hum Mol Genet 2019; 28: 1414-1428.

[41] Hauser SL and Cree BAC. Treatment of multiple sclerosis: a review. Am J Med 2020; 133: 1380-1390, e2.

[42] Guo C, Gao YY, Ju QQ, Zhang CX, Gong M and Li ZL. LINC00649 underexpression is an adverse prognostic marker in acute myeloid leukemia. BMC Cancer 2020; 20: 841.

[43] Chen X and Chen S. LINC00649 promotes bladder cancer malignant progression by regulating the miR15a5p/HMGA1 axis. Oncol Rep 2021; 45: 8.

[44] Wang H, Di X, Bi Y, Sun S and Wang T. Long non-coding RNA LINC00649 regulates YES-associated protein 1 (YAP1)/Hippo pathway to accelerate gastric cancer (GC) progression via sequestering miR-16-5p. Bioengineered 2021; 12: 1791-1802.

Supplementary Table 2. List of differentially expressed miRNAs

miRNA	logFC
hsa-miR-142-3p	4.2357085
hsa-miR-186	2.8655732
hsa-miR-98	2.83486
hsa-miR-130b	2.8221545
hsa-miR-7	2.4432054
hsa-miR-629	2.2825346
hsa-let-7c	2.1986221
hsa-let-7b	2.0665432
hsa-miR-619	2.0400192
hsa-let-7i	2.0122884
hsa-miR-140-5p	1.9960535
hsa-miR-34a	1.9393444
hsa-miR-892a	1.9002699
hsa-miR-615-5p	1.8999755
hsa-miR-145	1.8747742
hsa-miR-942	1.8390816
hsa-miR-148a	1.832381
hsa-miR-126	1.8322806
hsa-miR-1303	1.820051
hsa-miR-501-5p	1.814394
hsa-miR-608	1.8105159
hsa-miR-513a-3p	1.7980756
hsa-let-7e	1.7511411
hsa-miR-758	1.7443687
hsa-miR-223	1.7135023
hsa-miR-939	1.7101038
hsa-let-7g	1.6996091
hsa-miR-135a	1.5998915
hsa-miR-625	1.5803306
hsa-miR-1275	1.5745042
hsa-miR-885-5p	1.5727329
hsa-miR-658	1.5363959
hsa-miR-516b	1.5155584
hsa-miR-133a	1.5154452
hsa-miR-1207-5p	1.5098694
hsa-miR-624	1.5051318
hsa-miR-744	1.5032509
hsa-miR-451	1.4904226
hsa-miR-643	1.4800798
hsa-miR-22	1.4712502
hsa-miR-30d	1.4712471
hsa-miR-765	1.4316639
hsa-miR-1252	1.4127972
hsa-miR-365	1.3680369
hsa-miR-1911	1.3673465
hsa-miR-29b	1.35631
hsa-miR-200a	1.3394042
hsa-miR-572	1.3369357
hsa-miR-130a	1.3309718
hsa-miR-943	1.3189531
hsa-miR-199a-5p	1.3058443
hsa-miR-30e	1.2926271
hsa-miR-1287	1.2442879
hsa-miR-10b	1.2429673
hsa-miR-633	1.2279873
hsa-miR-1908	1.2208933
hsa-miR-659	1.2076403
hsa-miR-551b	-1.2224948

hsa-miR-371-5p	-1.2258024
hsa-miR-138	-1.2268572
hsa-miR-551a	-1.2494338
hsa-miR-647	-1.3045325
hsa-miR-211	-1.3164134
hsa-miR-302f	-1.3610912
hsa-miR-95	-1.3696805
hsa-miR-513a-5p	-1.3711852
hsa-miR-1910	-1.3731038
hsa-miR-654-3p	-1.3948593
hsa-miR-1270	-1.438118
hsa-miR-1256	-1.4412986
hsa-miR-548m	-1.4456289
hsa-miR-924	-1.4534229
hsa-miR-181d	-1.4737541
hsa-miR-641	-1.4997038
hsa-miR-1827	-1.5029866
hsa-miR-548l	-1.5206597
hsa-miR-10a	-1.5798142
hsa-miR-513b	-1.588105
hsa-miR-196a	-1.5955017
hsa-miR-19b	-1.5961783
hsa-miR-574-3p	-1.5991436
hsa-miR-30c	-1.6551129
hsa-miR-369-3p	-1.6664258
hsa-miR-513c	-1.6797069
hsa-miR-585	-1.6874115
hsa-miR-767-3p	-1.698606
hsa-miR-202	-1.7138339
hsa-miR-488	-1.7141297
hsa-miR-105	-1.726449
hsa-miR-99a	-1.7748024
hsa-miR-1249	-1.8006722
hsa-miR-644	-1.8014929
hsa-miR-379	-1.8200307
hsa-miR-1204	-1.8259435
hsa-miR-556-3p	-1.8261462
hsa-miR-609	-1.8485651
hsa-miR-382	-1.8701316
hsa-miR-218	-1.8763228
hsa-miR-299-3p	-1.9087321
hsa-miR-373	-1.9617412
hsa-miR-190b	-1.9789102
hsa-miR-548h	-2.0524013
hsa-miR-620	-2.0609552
hsa-miR-581	-2.114451
hsa-miR-610	-2.1596633
hsa-miR-875-5p	-2.170514
hsa-miR-651	-2.1708898
hsa-miR-125b	-2.1779881
hsa-miR-586	-2.1830611
hsa-miR-184	-2.2018869
hsa-miR-1306	-2.205356
hsa-miR-520d-3p	-2.2263496
hsa-miR-593	-2.2555876
hsa-miR-455-5p	-2.2946039
hsa-let-7a	-2.4571552
hsa-miR-1914	-2.4584416
hsa-let-7f	-2.4922172
hsa-miR-212	-2.654097
hsa-miR-512-3p	-3.3559523

Supplementary Table 3. Functional enriched analysis of the DEmRNAs

Category	ID	Term	Genes	*p*-value
BP	GO:0019731	antibacterial humoral response	IGHM, DEFA4, DEFA1B, B2M, CAMP, LTF	0.000458014
BP	GO:0050853	B cell receptor signaling pathway	BLK, IGHM, CD79A, CD19, IGHV3-23, IGHD	0.001183328
BP	GO:0006955	immune response	CXCL8, IGHV3-23, IFI6, SERPINB9, FCAR, TAPBP, VPREB3, ZEB1, IL7, IGHD, CEACAM8, DEFA1B, HLA-DOB, B2M, CD22, HLA-DQB1	0.001529446
BP	GO:0050830	defense response to Gram-positive bacterium	ADAM17, DEFA4, DEFA1B, LYZ, B2M, TIRAP, CAMP	0.001614498
BP	GO:0045087	innate immune response	BLK, IGHM, ANXA1, DDX3X, DEFA4, IGHV3-23, CYBB, TIRAP, LILRA5, CLEC7A, PCBP2, IGHD, DEFA1B, B2M, CAMP	0.004876241
CC	GO:0005833	hemoglobin complex	HBM, HBG2, HBB, HBA2, AHSP, HBD	4.15122E-07
CC	GO:0005654	nucleoplasm	ATF2, RBM25, RIF1, CELF1, CHD9, SPPL2B, KIAA1033, RORA, SMC3, AFF3, CLINT1, SYNE2, ACTB, NAMPT, PPP6R3, POLK, TGIF1, DDX17, SRRM2, PNISR, RBM14, ANXA1, DUSP3, SCAF11, TEX10, PPP4R3B, SAFB, SETSIP, TIRAP, FOXP1, SGO2, ZEB1, TBL1XR1, THRAP3, HNRNPH1, KAT6A, PPIG, TOP1, TP53, SF1, FKBP5, HEMGN, PRKDC, ANKRD11, BOD1L1, ANKRD12, GATA2, EXOSC6, CLEC7A, UBN1, PCBP2, USP1, FEM1C, TBL1X, SPEN, EGR1, RBM39, KDM4B, ALYREF, NONO, STAT2, DNAJC14, FOSL2, NR4A2, MLLT10, SNRNP40, ESF1, PKN2, MDM4, HSPA1B	1.72215E-06
CC	GO:0016020	membrane	CELF1, SPPL2B, PRF1, CD3G, PTPRK, CLINT1, SLC8A1, ACTB, TOMM22, ICAM1, LAMP1, CNTRL, CASP2, B2M, DDX17, EIF5A, ANXA3, APLP2, KRT1, SERPINB9, HBA2, CDKAL1, LRRC41, F5, TAPBP, ADAM17, PRKAR1A, HNRNPH1, HLA-DQB1, FKBP5, ARF6, KCNE3, CBFB, LUZP1, PRKDC, SELENBP1, APH1A, NT5E, RNF213, PCBP2, SNCA, MGEA5, SLC16A1, PLEKHA2, AMFR, ALYREF, NONO, PRRC2C, DNAJC14, TSPAN13, REEP3, GNAS, STT3A, KRAS, NAA15, CALR, CD24, CDS2	4.83176E-06
CC	GO:0005737	cytoplasm	ATF2, RIF1, SMC3, ACTB, LAMP1, CSPP1, NAMPT, NFATC2IP, NEK1, B2M, PNISR, SWAP70, LMO7, PPP4R3B, PLPP5, RAB30, ADAM17, PRKAR1A, PPIG, TP53, CDC42SE1, GGNBP2, KCNE3, UBA6, ANKRD11, ANKRD12, LPP, CD79A, RNF213, CLEC7A, PDGFC, PCBP2, STAP1, SCN3A, KLHDC10, EGR1, SYNRG, STYX, AZIN1, NR4A2, MLLT10, GLUD1, EIF2S3, GNAS, CALR, UBE2M, RBM25, DDX3X, CHD9, CELF1, CELF2, CD-CA7L, AFF3, SYNE2, OSBPL10, FCRLA, AKAP12, CA1, HINT3, CNTRL, PPP6R3, BLNK, CASP2, EPG5, EIF5A, RBM14, ANXA1, DUSP3, CNOT6L, ANXA3, ARG1, STRBP, TEX10, STRADB, CEP295, LIX1L, PAWR, SERPINB9, SETSIP, TIRAP, LRRC41, PTP4A1, ZEB1, HNRNPH1, CHI3L1, STRAP, LTF, ARPC4-TTLL3, TRAK2, FGD4, NT5E, ABLIM1, FEM1C, DEDD, SNCA, KDM4B, PLEKHA2, STAT2, HIPK3, P2RX5, SNRNP40, KRAS, PKN2, NAA15, HSPA1B	5.00075E-05
CC	GO:0072562	blood microparticle	IGHM, HBG2, KRT1, IGHV3-23, IGHD, HBB, HBA2, SLC4A1, HBD, HSPA1B, ACTB	6.68121E-05
MF	GO:0005515	protein binding	ATF2, SPPL2B, PRF1, SLC4A1, RORA, SMC3, ACTB, SLC8A1, ICAM1, CRKL, TOMM22, TCL1A, LAMP1, NAMPT, NEK1, B2M, BORCS5, DDX17, SCAF11, SWAP70, APLP2, KRT1, EBF1, CYBB, CDKAL1, FOXP1, SGO2, ADAM17, TBL1XR1, PRKAR1A, KAT6A, ADAM9, PPIG, TP53, SF1, FKBP5, BLK, RTN3, KCNE3, CBFB, EPB42, YTHDC1, UBA6, PRKDC, GATA2, LPP, APH1A, CD79A, CLEC7A, CD19, PDGFC, PCBP2, STAP1, LONRF1, NADK, KLHDC10, EGR1, SYNRG, ALYREF, NONO, STYX, ARPC4, AZIN1, FOSL2, NR4A2, MLLT10, GLUD1, EIF2S3, IL7, STT3A, GNAS, CYCS, MDM4, CALR, CD24, CD22, CD200, UBE2M, IGHM, RBM25, ALAS2, DDX3X, MS4A3, CXCL8, CELF1, ITGB3, CDCA7L, HBB, C1ORF174, PTPRK, HBD, CLINT1, SYNE2, AKAP12, CA1, CNTRL, NFKBIZ, PPP6R3, BLNK, PHACTR2, CASP2, CCR3, TGIF1, SRGN, EIF5A, FCRL2, RBM14, ANXA1, CNOT6L, KLF13, SECISBP2L, STRBP, TEX10, STRADB, ATRX, PAWR, SERPINB9, HBA2, PAX5, SAFB, FAM126B, TIRAP, F5, TAPBP, FCER2, ZEB1, HNRNPH1, THRAP3, STRAP, CEACAM8, TOP1, MS4A1, NUP58, ARF6, LTF, HEMGN, RUSC1-AS1, SDC-CAG3, HDC, IFI6, TRAK2, PCNP, SELENBP1, ABLIM1, MAT2A, UBN1, USP1, FEM1C, AHSP, TBL1X, DEDD, SNCA, LRRC25, SPEN, RBM39, CD72, PLEKHA2, AMFR, STAT2, G0S2, SNX20, SNRNP40, KRAS, PKN2, NAA15, COPG1, CLC, HSPA1B	1.10951E-10

MF	GO:0044822	poly(A) RNA binding	RBM25, DDX3X, PRKDC, CELF1, YTHDC1, CELF2, SAMSN1, CRKL, EXOSC6, ALKBH5, PCBP2, EIF5A, DDX17, SRRM2, SPEN, PNISR, RBM39, RBM14, SCAF11, SECISBP2L, ALYREF, NONO, STRBP, PRRC2C, SAFB, SNRNP40, THRAP3, HNRNPH1, STRAP, ESF1, PKN2, PPIG, NAA15, CALR, TOP1, SF1	4.11234E-06
MF	GO:0005344	oxygen transporter activity	HBM, HBG2, HBB, HBA2, HBD	3.01877E-05
MF	GO:0019825	oxygen binding	HBM, HBG2, HBB, HBA2, HBD	0.003784713
MF	GO:0019899	enzyme binding	LAMP1, PRKDC, ITGB3, PCBP2, TSPAN5, PAWR, CASP2, MDM4, ARPC4, TP53, HSPA1B, TRAK2	0.006107352
KEGG	hsa05416	Viral myocarditis	PRF1, CYCS, HLA-DOB, ACTB, ICAM1, HLA-DQB1	0.001825969
KEGG	hsa04662	B cell receptor signaling pathway	CD79A, CD72, CD19, BLNK, KRAS, CD22	0.004230988
KEGG	hsa05164	Influenza A	ATF2, CXCL8, STAT2, CYCS, HLA-DOB, HSPA1B, ACTB, ICAM1, HLA-DQB1	0.005537628
KEGG	hsa04612	Antigen processing and presentation	CALR, B2M, HLA-DOB, HSPA1B, TAPBP, HLA-DQB1	0.006385287
KEGG	hsa05161	Hepatitis B	ATF2, DDX3X, CXCL8, STAT2, CYCS, KRAS, TP53, TIRAP	0.007212606

Supplementary Table 4. GO biological process terms for top 4 modules

Mode	GO biological process	Description	Genes	adjust *P*-value
mode1	GO:0031640	killing of cells of other organism	FCER2/LTF/LYZ/CAMP/DEFA4/ARG1	5.85E-09
mode1	GO:0044364	disruption of cells of other organism	FCER2/LTF/LYZ/CAMP/DEFA4/ARG1	5.85E-09
mode1	GO:0001906	cell killing	FCER2/B2M/LTF/LYZ/CAMP/DEFA4/ARG1	1.85E-08
mode1	GO:0043312	neutrophil degranulation	HBB/B2M/LTF/CHI3L1/TCN1/LYZ/CAMP/DEFA4/ARG1	2.57E-08
mode1	GO:0002283	neutrophil activation involved in immune response	HBB/B2M/LTF/CHI3L1/TCN1/LYZ/CAMP/DEFA4/ARG1	2.57E-08
mode1	GO:0042119	neutrophil activation	HBB/B2M/LTF/CHI3L1/TCN1/LYZ/CAMP/DEFA4/ARG1	2.57E-08
mode1	GO:0002446	neutrophil mediated immunity	HBB/B2M/LTF/CHI3L1/TCN1/LYZ/CAMP/DEFA4/ARG1	2.57E-08
mode1	GO:0015671	oxygen transport	HBB/HBA2/HBM/HBG2	6.92E-08
mode1	GO:0015669	gas transport	HBB/HBA2/HBM/HBG2	1.74E-07
mode1	GO:0035821	modification of morphology or physiology of other organism	FCER2/LTF/LYZ/CAMP/DEFA4/ARG1	4.39E-07
mode1	GO:0006959	humoral immune response	FCER2/CD19/LTF/LYZ/CAMP/PAX5/DEFA4	1.20E-06
mode1	GO:0042744	hydrogen peroxide catabolic process	HBB/HBA2/HBM/HBG2	1.20E-06
mode1	GO:0015893	drug transport	HBB/HBA2/TCN1/HBM/HBG2/ARG1	2.01E-06
mode1	GO:0017001	antibiotic catabolic process	HBB/HBA2/HBM/HBG2	1.04E-05
mode1	GO:0042743	hydrogen peroxide metabolic process	HBB/HBA2/HBM/HBG2	1.04E-05
mode1	GO:0051187	cofactor catabolic process	HBB/HBA2/HBM/HBG2	2.00E-05
mode1	GO:0050829	defense response to Gram-negative bacterium	LTF/LYZ/CAMP/DEFA4	5.08E-05
mode1	GO:0002227	innate immune response in mucosa	LTF/CAMP/DEFA4	5.59E-05
mode1	GO:0097237	cellular response to toxic substance	HBB/HBA2/HBM/HBG2/ARG1	7.86E-05
mode1	GO:0098869	cellular oxidant detoxification	HBB/HBA2/HBM/HBG2	7.86E-05
mode1	GO:1990748	cellular detoxification	HBB/HBA2/HBM/HBG2	8.74E-05
mode1	GO:0019730	antimicrobial humoral response	LTF/LYZ/CAMP/DEFA4	0.000133
mode1	GO:0002385	mucosal immune response	LTF/CAMP/DEFA4	0.000138
mode1	GO:0050853	B cell receptor signaling pathway	CD79A/CD19/BLK/PAX5	0.000138
mode1	GO:0098754	detoxification	HBB/HBA2/HBM/HBG2	0.000138
mode1	GO:0002251	organ or tissue specific immune response	LTF/CAMP/DEFA4	0.000147
mode1	GO:0050832	defense response to fungus	LTF/DEFA4/ARG1	0.000163
mode1	GO:0051291	protein heterooligomerization	HBB/HBA2/HBM/HBG2	0.000163
mode1	GO:0042737	drug catabolic process	HBB/HBA2/HBM/HBG2	0.000199
mode1	GO:0019731	antibacterial humoral response	LTF/CAMP/DEFA4	0.000257
mode1	GO:0016999	antibiotic metabolic process	HBB/HBA2/HBM/HBG2	0.000257
mode1	GO:0009620	response to fungus	LTF/DEFA4/ARG1	0.000318
mode1	GO:0031343	positive regulation of cell killing	FCER2/B2M/ARG1	0.000612

mode1	GO:0051673	membrane disruption in other organism	LTF/DEFA4	0.000766
mode1	GO:0061844	antimicrobial humoral immune response mediated by antimicrobial peptide	LTF/CAMP/DEFA4	0.000789
mode1	GO:0001895	retina homeostasis	B2M/LTF/LYZ	0.000833
mode1	GO:0051851	modification by host of symbiont morphology or physiology	LTF/CAMP/ARG1	0.000915
mode1	GO:0051712	positive regulation of killing of cells of other organism	FCER2/ARG1	0.000979
mode1	GO:0051873	killing by host of symbiont cells	CAMP/ARG1	0.000979
mode1	GO:0051702	interaction with symbiont	LTF/CAMP/ARG1	0.000986
mode1	GO:0031341	regulation of cell killing	FCER2/B2M/ARG1	0.001222
mode1	GO:0051709	regulation of killing of cells of other organism	FCER2/ARG1	0.001222
mode1	GO:0051852	disruption by host of symbiont cells	CAMP/ARG1	0.001222
mode1	GO:0050830	defense response to Gram-positive bacterium	LYZ/CAMP/DEFA4	0.001423
mode1	GO:0050851	antigen receptor-mediated signaling pathway	CD79A/CD19/BLK/PAX5	0.001423
mode1	GO:0051883	killing of cells in other organism involved in symbiotic interaction	CAMP/ARG1	0.001505
mode1	GO:0072593	reactive oxygen species metabolic process	HBB/HBA2/HBM/HBG2	0.001742
mode1	GO:0051818	disruption of cells of other organism involved in symbiotic interaction	CAMP/ARG1	0.001837
mode1	GO:0051817	modification of morphology or physiology of other organism involved in symbiotic interaction	LTF/CAMP/ARG1	0.002191
mode1	GO:0042742	defense response to bacterium	LTF/LYZ/CAMP/DEFA4	0.002855
mode1	GO:0050855	regulation of B cell receptor signaling pathway	BLK/PAX5	0.003108
mode1	GO:0002705	positive regulation of leukocyte mediated immunity	FCER2/B2M/ARG1	0.003278
mode1	GO:0002237	response to molecule of bacterial origin	B2M/LTF/DEFA4/ARG1	0.003321
mode1	GO:0002724	regulation of T cell cytokine production	B2M/ARG1	0.003726
mode1	GO:0042542	response to hydrogen peroxide	HBB/HBA2/ARG1	0.003847
mode1	GO:0002822	regulation of adaptive immune response based on somatic recombination of immune receptors built from immunoglobulin superfamily domains	FCER2/B2M/ARG1	0.00402
mode1	GO:0002706	regulation of lymphocyte mediated immunity	FCER2/B2M/ARG1	0.004112
mode1	GO:0043900	regulation of multi-organism process	FCER2/LTF/CAMP/ARG1	0.004374
mode1	GO:0002819	regulation of adaptive immune response	FCER2/B2M/ARG1	0.004914
mode1	GO:0002369	T cell cytokine production	B2M/ARG1	0.006292
mode1	GO:0002429	immune response-activating cell surface receptor signaling pathway	CD79A/CD19/BLK/PAX5	0.006292
mode1	GO:0002697	regulation of immune effector process	FCER2/CD19/B2M/ARG1	0.007715
mode1	GO:0002768	immune response-regulating cell surface receptor signaling pathway	CD79A/CD19/BLK/PAX5	0.007988
mode1	GO:0015701	bicarbonate transport	HBB/HBA2	0.008015
mode1	GO:0002703	regulation of leukocyte mediated immunity	FCER2/B2M/ARG1	0.008508
mode1	GO:0071222	cellular response to lipopolysaccharide	LTF/DEFA4/ARG1	0.008755
mode1	GO:0043903	regulation of symbiosis, encompassing mutualism through parasitism	LTF/CAMP/ARG1	0.009136
mode1	GO:0071219	cellular response to molecule of bacterial origin	LTF/DEFA4/ARG1	0.009391
mode1	GO:0002699	positive regulation of immune effector process	FCER2/B2M/ARG1	0.00965
mode1	GO:0001894	tissue homeostasis	B2M/LTF/LYZ	0.011173
mode1	GO:0000302	response to reactive oxygen species	HBB/HBA2/ARG1	0.011452
mode1	GO:0001912	positive regulation of leukocyte mediated cytotoxicity	B2M/ARG1	0.011635
mode1	GO:0051896	regulation of protein kinase B signaling	CD19/CHI3L1/BANK1	0.011708
mode1	GO:0071216	cellular response to biotic stimulus	LTF/DEFA4/ARG1	0.011708
mode1	GO:0050854	regulation of antigen receptor-mediated signaling pathway	BLK/PAX5	0.011991
mode1	GO:0046686	response to cadmium ion	B2M/ARG1	0.01443
mode1	GO:0043491	protein kinase B signaling	CD19/CHI3L1/BANK1	0.014971
mode1	GO:0001818	negative regulation of cytokine production	LTF/BANK1/ARG1	0.01522
mode1	GO:0002709	regulation of T cell mediated immunity	B2M/ARG1	0.01522
mode1	GO:0001910	regulation of leukocyte mediated cytotoxicity	B2M/ARG1	0.019855
mode1	GO:0002718	regulation of cytokine production involved in immune response	B2M/ARG1	0.020121
mode1	GO:0055072	iron ion homeostasis	B2M/LTF	0.023585
mode1	GO:0060333	interferon-gamma-mediated signaling pathway	B2M/ARG1	0.026678
mode1	GO:0046677	response to antibiotic	HBB/HBA2/ARG1	0.027351
mode1	GO:0032496	response to lipopolysaccharide	LTF/DEFA4/ARG1	0.027351
mode1	GO:0002367	cytokine production involved in immune response	B2M/ARG1	0.028622

mode1	GO:0002824	positive regulation of adaptive immune response based on somatic recombination of immune receptors built from immunoglobulin superfamily domains	FCER2/B2M	0.031253
mode1	GO:0002449	lymphocyte mediated immunity	FCER2/B2M/ARG1	0.031253
mode1	GO:0006898	receptor-mediated endocytosis	HBB/B2M/HBA2	0.031407
mode1	GO:0002456	T cell mediated immunity	B2M/ARG1	0.031425
mode1	GO:0002460	adaptive immune response based on somatic recombination of immune receptors built from immunoglobulin superfamily domains	FCER2/B2M/ARG1	0.031838
mode1	GO:0001909	leukocyte mediated cytotoxicity	B2M/ARG1	0.031838
mode1	GO:0002708	positive regulation of lymphocyte mediated immunity	FCER2/B2M	0.031838
mode1	GO:0002821	positive regulation of adaptive immune response	FCER2/B2M	0.031838
mode1	GO:0000041	transition metal ion transport	B2M/TCN1	0.03449
mode1	GO:0002920	regulation of humoral immune response	FCER2/CD19	0.044423
mode1	GO:0002831	regulation of response to biotic stimulus	LTF/ARG1	0.046705
mode1	GO:0002700	regulation of production of molecular mediator of immune response	B2M/ARG1	0.048312
mode1	GO:0055076	transition metal ion homeostasis	B2M/LTF	0.049451
mode1	GO:0060249	anatomical structure homeostasis	B2M/LTF/LYZ	0.049451
mode1	GO:0051250	negative regulation of lymphocyte activation	BANK1/ARG1	0.049636
mode2	GO:0000209	protein polyubiquitination	RNF213/LMO7/LONRF1	0.001492
mode2	GO:0043687	post-translational protein modification	LRRC41/UBE2M/LMO7	0.002048
mode2	GO:0045116	protein neddylation	UBE2M	0.046956
mode3	GO:0042743	hydrogen peroxide metabolic process	SNCA/HBD/CYBB	0.016945
mode3	GO:0030520	intracellular estrogen receptor signaling pathway	RBM14/DDX17/SAFB	0.016945
mode3	GO:0048545	response to steroid hormone	RBM14/CYBB/DDX17/SAFB/3ARG1	0.018909
mode3	GO:0020027	hemoglobin metabolic process	ALAS2/AHSP	0.025107
mode3	GO:0008380	RNA splicing	HNRNPH1/DDX17/NONO/SNRNP40/SRRM2	0.025472
mode3	GO:0071383	cellular response to steroid hormone stimulus	RBM14/DDX17/SAFB/3ARG1	0.028159
mode3	GO:0072593	reactive oxygen species metabolic process	SNCA/HBD/CYBB/3ARG1	0.028904
mode3	GO:0034599	cellular response to oxidative stress	SNCA/CYBB/4841/CYCS	0.028904
mode3	GO:1903409	reactive oxygen species biosynthetic process	SNCA/CYBB/3ARG1	0.028904
mode3	GO:0043280	positive regulation of cysteine-type endopeptidase activity involved in apoptotic process	SNCA/DDX3X/CYCS	0.031828
mode3	GO:0001666	response to hypoxia	ALAS2/EGR1/CYBB/3ARG1	0.031828
mode3	GO:0036293	response to decreased oxygen levels	ALAS2/EGR1/CYBB/3ARG1	0.031828
mode3	GO:0030518	intracellular steroid hormone receptor signaling pathway	RBM14/DDX17/SAFB	0.031828
mode3	GO:2001056	positive regulation of cysteine-type endopeptidase activity	SNCA/DDX3X/CYCS	0.031828
mode3	GO:0070482	response to oxygen levels	ALAS2/EGR1/CYBB/3ARG1	0.031828
mode3	GO:0042744	hydrogen peroxide catabolic process	SNCA/HBD	0.031828
mode3	GO:0016999	antibiotic metabolic process	SNCA/HBD/CYBB	0.031828
mode3	GO:0070555	response to interleukin-1	SNCA/EGR1/3ARG1	0.031828
mode3	GO:0000377	RNA splicing, via transesterification reactions with bulged adenosine as nucleophile	HNRNPH1/DDX17/SNRNP40/SRRM2	0.031828
mode3	GO:0000398	mRNA splicing, via spliceosome	HNRNPH1/DDX17/SNRNP40/SRRM2	0.031828
mode3	GO:0071456	cellular response to hypoxia	EGR1/CYBB/3ARG1	0.031828
mode3	GO:0000375	RNA splicing, via transesterification reactions	HNRNPH1/DDX17/SNRNP40/SRRM2	0.031828
mode3	GO:0036294	cellular response to decreased oxygen levels	EGR1/CYBB/3ARG1	0.034505
mode3	GO:0010950	positive regulation of endopeptidase activity	SNCA/DDX3X/CYCS	0.034505
mode3	GO:0015701	bicarbonate transport	SLC4A1/759	0.036528
mode3	GO:0046688	response to copper ion	SNCA/3ARG1	0.036528
mode3	GO:0022900	electron transport chain	SNCA/CYBB/CYCS	0.036528
mode3	GO:0007157	heterophilic cell-cell adhesion via plasma membrane cell adhesion molecules	3ARG1/CEACAM8	0.036528
mode3	GO:0071453	cellular response to oxygen levels	EGR1/CYBB/3ARG1	0.036755
mode3	GO:0010952	positive regulation of peptidase activity	SNCA/DDX3X/CYCS	0.036755
mode3	GO:0043401	steroid hormone mediated signaling pathway	RBM14/DDX17/SAFB	0.036755
mode3	GO:0006979	response to oxidative stress	SNCA/CYBB/NONO/CYCS	0.039168

mode3	GO:0043281	regulation of cysteine-type endopeptidase activity involved in apoptotic process	SNCA/DDX3X/CYCS	0.043551
mode3	GO:0032722	positive regulation of chemokine production	EGR1/DDX3X	0.043551
mode3	GO:0017001	antibiotic catabolic process	SNCA/HBD	0.047378
mode3	GO:0042108	positive regulation of cytokine biosynthetic process	EGR1/CYBB	0.047378
mode3	GO:2001234	negative regulation of apoptotic signaling pathway	3ARG1/NONO/DDX3X	0.047378
mode4	GO:0042113	B cell activation	BLNK/MS4A1/CD22	0.004031
mode4	GO:0042100	B cell proliferation	MS4A1/CD22	0.008512
mode4	GO:0002377	immunoglobulin production	VPREB3/CD22	0.021385
mode4	GO:0046651	lymphocyte proliferation	MS4A1/CD22	0.021385
mode4	GO:0032943	mononuclear cell proliferation	MS4A1/CD22	0.021385
mode4	GO:0002440	production of molecular mediator of immune response	VPREB3/CD22	0.021385
mode4	GO:0070661	leukocyte proliferation	MS4A1/CD22	0.021385
mode4	GO:0002638	negative regulation of immunoglobulin production	CD22	0.021385
mode4	GO:0006959	humoral immune response	BLNK/MS4A1	0.02702
mode4	GO:0051023	regulation of immunoglobulin secretion	CD22	0.027968
mode4	GO:0048305	immunoglobulin secretion	CD22	0.029651
mode4	GO:0050855	regulation of B cell receptor signaling pathway	CD22	0.029852
mode4	GO:0050858	negative regulation of antigen receptor-mediated signaling pathway	CD22	0.029852
mode4	GO:0050849	negative regulation of calcium-mediated signaling	CD22	0.035143
mode4	GO:0002701	negative regulation of production of molecular mediator of immune response	CD22	0.035143

Supplementary Table 5. KEGG enriched pathways for top 4 modules

Mode	Pathway ID	Description	Genes	adjust *P*-value
mode1	hsa05143	African trypanosomiasis	HBB/HBA2	0.011927
mode1	hsa05340	Primary immunodeficiency	CD79A/CD19	0.011927
mode1	hsa05144	Malaria	HBB/HBA2	0.011927
mode1	hsa05169	Epstein-Barr virus infection	FCER2/CD19/B2M	0.011927
mode1	hsa04662	B cell receptor signaling pathway	CD79A/CD19	0.020777
mode1	hsa04970	Salivary secretion	LYZ/CAMP	0.020777
mode1	hsa05150	Staphylococcus aureus infection	CAMP/DEFA4	0.020777
mode1	hsa04640	Hematopoietic cell lineage	FCER2/CD19	0.020777
mode2	hsa04120	Ubiquitin mediated proteolysis	UBA6/UBE2M	0.001756
mode2	hsa04520	Adherens junction	LMO7	0.026044
mode3	hsa04933	AGE-RAGE signaling pathway in diabetic complications	EGR1/CYBB/ICAM1	0.018172
mode4	hsa04662	B cell receptor signaling pathway	CD72/BLNK/CD22	2.77E-05
mode4	hsa04640	Hematopoietic cell lineage	MS4A1/CD22	0.00305
mode4	hsa05340	Primary immunodeficiency	BLNK	0.043444