Identification of multiple sclerosis-related genes regulated by EBV-encoded microRNAs in B cells

Xinming Rang  
Second Affiliated Hospital of Harbin Medical University

Yuan Liu  
Second Affiliated Hospital of Harbin Medical University

Jingguo Wang  
Second Affiliated Hospital of Harbin Medical University

Yifei Wang  
Second Affiliated Hospital of Harbin Medical University

Chaohan Xu  
Harbin Medical College: Harbin Medical University

Jin Fu  
Second Affiliated Hospital of Harbin Medical University  
https://orcid.org/0000-0001-7532-784X

Research Article

Keywords: Multiple sclerosis, Genomics, Transcriptomics, MS-related genes, Epstein-Barr virus, microRNAs

DOI: https://doi.org/10.21203/rs.3.rs-338909/v1

License: ☝️  This work is licensed under a Creative Commons Attribution 4.0 International License.  
Read Full License
Abstract

Background: Multiple sclerosis (MS) is driven by the interaction between genetic susceptibility and environmental triggers, particularly to Epstein-Barr virus (EBV) infection. EBV-encoded microRNAs (miRNAs) are abundantly expressed in all stages of EBV infection and latency, which can target both viral and host cellular mRNAs, allowing EBV-infected B cells to evade the host immune response. However, it remains a big gap to understand the roles of EBV miRNAs and their target genes in MS pathogenesis.

Methods: We investigated the correlation between MS-related viruses infection and MS risk quantitatively by systematic analysis. All MS-related genes in B cells were obtained by integrating MS susceptibility genes and differentially expressed genes from B cells. In comparison with differentially expressed genes from B cells after EBV infection in vitro, we confirmed EBV-regulated, MS-related genes. Subsequently, we obtained target EBV miRNAs which can regulate these genes from several online databases. By constructing pathway-pathway, pathway-gene and protein-protein interaction networks, we further screened out MS-related genes and risk pathways regulated by EBV miRNAs. Finally, we identified target EBV miRNAs may directly regulate MS-related genes through bioinformatic prediction and experimental validation.

Results: EBV infection showed the strongest correlation with MS risk. A total of 873 MS-related genes and 52 risk pathways in B cells were obtained. We then identified 150 MS-related genes and 18 associated risk pathways that EBV was involved in. In addition, 42 human target genes regulated by 36 EBV miRNAs overlapped with EBV-regulated, MS-related genes. Finally, 15 target EBV miRNAs and their regulated, 6 MS-related genes (MALT1, BCL10, IFNGR2, STAT3, CDK6 and FOXP1) have been confirmed as crucial pathogenic molecules, which could promote the initiation and development of MS through NF-kappa B (MALT1 and BCL10) and PD-L1/PD-1 (IFNGR2 and STAT3) pathways. Surprisingly, ebv-miR-BHRF1-2-5p directly targeting MALT1 was confirmed by our experiments, and FOXP1 was identified as a target gene of ebv-miR-BART11.

Conclusions: This work identified the target EBV miRNAs and their regulated, MS-related genes and risk pathways, which may provide a novel insight into discovering diagnostic biomarkers and therapeutic targets for MS.

Background

Multiple sclerosis (MS) is an autoimmune disease characterized by demyelination of white matter in the central nervous system (CNS). It is a major cause of serious physical disability in young adults in developed countries[1]. MS involves the interaction between genetic susceptibility and environmental risk factors, and more than 200 gene loci related to MS have been identified through genome wide association studies (GWAS), which can account for 25% of the perceived heritability of MS[2–4]. Though a proportionately larger contribution was driven by non-genetic influences, such as Epstein-Barr virus (EBV) infection, vitamin D deficiency and smoking, less progress has been made in how environmental
triggers may function in MS pathogenesis[5]. Until now, the diagnosis of MS mainly relies on imaging features and clinical manifestations due to the limited understanding of MS pathogenesis, suggesting a tremendous need of specific diagnostic biomarkers.

Epidemiological researches reported that viruses infection are closely related to the pathogenesis and development of MS, including EBV, Cytomegalovirus (CMV), Human immunodeficiency virus (HIV), influenza virus and measles virus, but the correlation between these viruses and MS risk has not been determined quantitatively[6–10]. From a large longitudinal study, the seropositivity of anti-EBV nuclear antigen IgG and the history of infectious mononucleosis caused by EBV infection were the top two risk factors associated with MS[11]. Pathological studies indicated that EBV-infected B cells can migrate to the CNS of MS patients. At advanced stages of the disease (secondary progressive MS), it may be induced to organize themselves into B cell follicles, contributing to viral existence and reactivation[12]. Moreover, a significant additive interaction between EBV infection and major histocompatibility complex, class II, dr beta 1 (HLA-DRB1)*1501 has been observed in the risk of MS[13]. Collectively, EBV infection has been identified as a key regulator in the initiation and progress of MS, but its roles in the pathogenesis of MS have not been established.

MicroRNAs (miRNAs) are small non-coding RNAs that regulate gene expression at the posttranscriptional level, whose potential importance has been highlighted in the pathogenesis of MS. In contrast to human miRNAs, viral miRNAs have been confirmed as more active players in host-microbe interactions, but little is known about miRNAs encoded by EBV, the major environmental risk factor of MS. EBV encodes 44 mature miRNAs which are located in two main clusters, the BamHI fragment H rightward open reading frame 1 (BHRF1) and BamHI A rightward transcripts (BART) clusters[14, 15]. EBV miRNAs are abundantly expressed in all stages of EBV infection and latency (25% of miRNAs in Latency III), which can target both viral and host cellular mRNAs, allowing EBV-infected B cells to evade the host immune response[16, 17]. Albanese et al compared the primary B cells infected with EBV expressing or lacking viral miRNAs. In contrast to miRNAs-expressing EBV, primary B cells infected with miRNAs-deficient EBV were more susceptible to death when introducing EBV-specific CD8+ T cells[18]. Barth et al indicated that ebv-miR-BART2 inhibited transition from latent to lytic viral replication by downregulating the viral DNA polymerase BALF5, suggesting that EBV miRNAs had an important regulatory impact on host B cells[19]. Current researches of EBV miRNAs are mainly focusing on tumors, ebv-miR-BART8-3p has been confirmed as a potential biomarker for nasopharyngeal carcinoma, which can induce epithelial-mesenchymal transition and promote metastasis through directly targeting Ring Finger Protein 38 (RNF38) in nasopharyngeal carcinoma cells[20]. However, the molecular mechanism of EBV miRNAs may function in MS pathogenesis has not been systematically elucidated.

In the present study, we collected and investigated the currently available transcriptomics datasets of MS-related viruses and MS patients, the results of which showed the most significant correlation between EBV infection and MS risk by systematic analysis. Furthermore, we comprehensively screened out MS-related genes in B cells by integrating MS susceptibility genes and differentially expressed genes (DEGs) from B cells. In comparison with DEGs from B cells after EBV infection in vitro, we confirmed EBV-
regulated, MS-related genes. We then obtained target EBV miRNAs which can regulate these genes from several online databases. Finally, we identified target EBV miRNAs may directly regulate MS-related genes through bioinformatic prediction and experimental validation. These findings may provide a novel insight into discovering diagnostic biomarkers and therapeutic targets for MS.

Methods

MS-related viruses and MS patients datasets selection

A flowchart for this part was shown in Fig. 1a. All currently available microarray gene expression datasets of MS-related viruses and MS patients were downloaded from the Gene Expression Omnibus (GEO) database (http://www.ncbi.nlm.nih.gov/geo/). The selection criteria were as follows: (1) samples of MS-related viruses datasets were derived from viruses susceptible cells which were isolated from healthy individuals and then infected in vitro. (2) samples of MS patients datasets were derived from those viruses susceptible cells. (3) samples of MS patients datasets were immunosuppressive treatment naïve or had not received treatment for at least two months. (4) the sample size of both case and control groups should be no fewer than three. According to these criteria, seven viral gene expression profiling studies were included: two datasets for EBV (accession number: GSE45829 and GSE49628[21, 22]), two datasets for CMV (accession number: GSE17948 and GSE24238[23, 24]), one dataset for HIV (accession number: GSE33877[25]), one dataset for influenza virus (accession number: GSE32140[9]) and one dataset for measles virus (accession number: GSE980[10]). Meanwhile, a total of four MS patients datasets were taken from GSE117935, GSE66988, GSE41890 and GSE37750[26–29].

Data pre-processing and identification of the DEGs

Raw expression values were quantile normalized and log2 transformed. Probe sets mapping multiple genes were removed and the mean values were calculated as expression levels of those mapping on one gene. DEGs were screened out by using linear model in “limma” package in R[30]. The cut-off criteria for statistical significance were set as \( p \)-value less than 0.05.

Hypergeometric test

A hypergeometric test was used to assess the correlation between MS-related viruses infection and MS risk[31]. The \( p \)-value was derived from the hypergeometric function below:
Here, \( m \) represents the total number of human genome genes, \( j \) is the number of DEGs in each group of cells after MS-related viruses infection, \( n \) belongs to the number of DEGs between MS patients and healthy controls, and \( x \) denotes the number of genes belonged to both two groups of DEGs. The \( p \)-value less than 0.05 was considered statistically significant.

Acquisition of MS-related genes regulated by EBV

A flowchart for this part was shown in Fig. 1b. In order to comprehensively screen out MS-related genes in B cells, DEGs from B cells between MS patients and healthy individuals were integrated with MS susceptibility genes. Through comparing MS-related genes and DEGs between EBV-infected B cells and primary B cells, we identified EBV-regulated, MS-related genes.

Pathway enrichment analysis

Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis was performed for given genes through clusterProfiler package in R[32]. Gene symbols were transformed into EntrezID by means of org.Hs.eg.db package. These significant pathways were selected using Fisher’s exact test, with the threshold of 0.05 for \( p \)-value.

Pathway-pathway and pathway-gene networks construction

In order to further identify the involvement of EBV in MS pathogenesis, KEGG pathway enrichment analysis and hypergeometric test were used to construct pathway-pathway network, which was visualized by Cytoscape software[33]. In this hypergeometric function, \( m \) represents the total number of human genome genes, \( j \) and \( n \) represent the number of genes in the two MS risk pathways respectively, and \( x \) denotes the number of genes belonged to both two MS risk pathways. In addition, all MS risk pathways regulated by EBV and their mapped genes were considered and combined to construct a pathway-gene network. The hub genes were determined by calculating centrality degree.

Investigation of target EBV miRNAs

MiRTarbase (http://mirtarbase.mbc.nctu.edu.tw/) and TarBase (http://carolina.imis.athena-innovation.gr/diana_tools/web/index.php?r=tarbasev8%2Findex) were used to obtain human target
genes of 44 mature EBV miRNAs. All target genes should be experimentally validated, including western blot, luciferase reporter assay or high throughput sequencing[34, 35]. In comparison with EBV-regulated, MS-related genes, we identified MS-related genes regulated by EBV miRNAs. The pathway-pathway network was constructed according to KEGG pathway enrichment analysis and hypergeometric test. Subsequently, we obtained the hub genes in the MS risk pathway-gene network regulated by EBV miRNAs. In addition, a protein-protein interaction (PPI) network was constructed by STRING database to further investigate EBV miRNAs-regulated, MS-related genes[36]. Finally, we identified target EBV miRNAs may directly regulate MS-related genes through bioinformatic prediction and experimental validation.

**Results**

**Acquisition of MS-related viruses and MS patients DEGs**

Based on the threshold of 0.05 for \( p \)-value, we first screened out DEGs in viruses susceptible cells from healthy individuals after MS-related viruses infected in vitro. Subsequently, we identified DEGs in viruses susceptible cells between MS patients and healthy controls. The number of DEGs was shown in Table 1.
## Table 1

Detailed information of MS-related viruses and MS patients gene expression profiles.

<table>
<thead>
<tr>
<th>GEO ID</th>
<th>Features</th>
<th>Platforms</th>
<th>Cases/Controls</th>
<th>Organism</th>
<th>cut off criteria</th>
<th>DEGs</th>
</tr>
</thead>
<tbody>
<tr>
<td>GSE45829</td>
<td>EBV</td>
<td>Affymetrix Human Exon 1.0 ST Array [transcript (gene) version]</td>
<td>3/3</td>
<td>B cells</td>
<td>$p &lt; 0.05$</td>
<td>7218</td>
</tr>
<tr>
<td>GSE49628</td>
<td>EBV</td>
<td>Affymetrix Human Genome U133 Plus 2.0 Array</td>
<td>3/3</td>
<td>B cells</td>
<td>$p &lt; 0.05$</td>
<td>9044</td>
</tr>
<tr>
<td>GSE17948</td>
<td>CMV</td>
<td>Affymetrix Human Genome U95 Version 2 Array</td>
<td>4/4</td>
<td>Monocytes</td>
<td>$p &lt; 0.05$</td>
<td>931</td>
</tr>
<tr>
<td>GSE24238</td>
<td>CMV</td>
<td>Affymetrix Human Genome U95 Version 2 Array</td>
<td>3/3</td>
<td>Monocytes</td>
<td>$p &lt; 0.05$</td>
<td>2428</td>
</tr>
<tr>
<td>GSE33877</td>
<td>HIV</td>
<td>Illumina HumanHT-12 V3.0 expression beadchip</td>
<td>6/6</td>
<td>Peripheral blood leukocytes</td>
<td>$p &lt; 0.05$</td>
<td>2579</td>
</tr>
<tr>
<td>GSE32140</td>
<td>Influenza</td>
<td>Affymetrix Human Genome U133 Plus 2.0 Array</td>
<td>28/12</td>
<td>Peripheral blood leukocytes</td>
<td>$p &lt; 0.05$</td>
<td>2626</td>
</tr>
<tr>
<td>GSE980</td>
<td>Measles</td>
<td>Affymetrix Human Genome U95 Version 2 Array</td>
<td>3/4</td>
<td>Dendritic cells</td>
<td>$p &lt; 0.05$</td>
<td>3753</td>
</tr>
<tr>
<td>GSE117935</td>
<td>MS</td>
<td>Affymetrix Human Exon 1.0 ST Array [transcript (gene) version]</td>
<td>10/10</td>
<td>B cells</td>
<td>$p &lt; 0.05$</td>
<td>601</td>
</tr>
<tr>
<td>GSE66988</td>
<td>MS</td>
<td>Affymetrix Human Gene 1.0 ST Array [transcript (gene) version]</td>
<td>4/4</td>
<td>Monocytes</td>
<td>$p &lt; 0.05$</td>
<td>4227</td>
</tr>
<tr>
<td>GSE41890</td>
<td>MS</td>
<td>Affymetrix Human Gene 1.0 ST Array [transcript (gene) version]</td>
<td>14/24</td>
<td>Peripheral blood leukocytes</td>
<td>$p &lt; 0.05$</td>
<td>2117</td>
</tr>
<tr>
<td>GSE37750</td>
<td>MS</td>
<td>Affymetrix Human Genome U133 Plus 2.0 Array</td>
<td>9/8</td>
<td>Dendritic cells</td>
<td>$p &lt; 0.05$</td>
<td>1925</td>
</tr>
</tbody>
</table>

### Evaluation of correlation between MS-related viruses infection and MS risk

A hypergeometric test was conducted to explore the correlation between MS-related viruses infection and MS risk. We compared each group of MS-related viruses and MS patients DEGs derived from the same type of cells and found that four types of viruses were statistically significant with MS, apart from HIV. As anticipated, the correlation between EBV infection and MS risk was the most significant ($p < 1.0 \times 10^{-6}$), followed by measles virus, influenza virus and CMV (Fig. 2).
Acquisition of all MS-related genes and risk pathways in B cells

In recent years, a total of 284 MS susceptibility genes have been confirmed through GWAS\cite{2, 3}. By analyzing published transcriptomics data from B cells between relapsing-remitting MS (RRMS) patients and healthy controls\cite{26}, 601 DEGs were confirmed (322 upregulated-genes, 279 downregulated-genes), with only 12 overlapped genes between DEGs and MS susceptibility genes (Fig. 3a). This result suggested that hereditary susceptibility did not solely dominate the alteration of gene expression during MS development, environmental factors also played an important role, so that both susceptibility genes and DEGs should be regarded as MS-related genes. Therefore, we integrated 284 susceptibility genes and 601 DEGs from B cells, and a total of 873 MS-related genes were collected. By using KEGG pathway enrichment analysis, 52 MS risk pathways were obtained (Fisher's exact test, \( p < 0.05 \), Additional file 1: Table S1). As shown in Fig. 3b, MS-related genes in B cells were mainly mapped into immune signaling pathways, especially Epstein-Barr virus infection.

Identification of MS-related genes regulated by EBV

We further analyzed the impact of EBV infection in the pathogenesis and development of MS. A total of 3556 DEGs were confirmed in both GSE45829 and GSE49628, with 2726 being significantly upregulated and 830 being significantly downregulated in EBV-infected B cells compared with primary B cells. In comparison with 873 MS-related genes in B cells with the same trend, a significant overlap with 150 EBV-regulated, MS-related genes were screened out. By using KEGG pathway enrichment analysis, we obtained 23 significant MS risk pathways which were regulated by EBV (Fisher's exact test, \( p < 0.05 \), Additional file 1: Table S2). Furthermore, a total of 18 closely associated MS risk pathways were identified by hypergeometric test, and the pathway-pathway network was visualized by Cytoscape (Fig. 4a, b). In addition, we constructed a pathway-gene network among these 18 risk pathways and their enriched genes. In this network, 10 genes \( \text{HLA-DPB1, IFNGR2, STAT3, TNFRSF1A, CXCR4, IL10RA, LCK, MALT1, BCL10 and LTA} \) were determined as hub genes due to their high centrality degree (Fig. 4c, d).

Identification of the target EBV miRNAs

By searching miRTarbase and TarBase, 209 and 5962 EBV miRNAs-mRNAs pairs were downloaded, respectively. And a total of 3171 human target genes of 44 mature EBV miRNAs were obtained, out of which 42 target genes overlapped with 150 EBV-regulated, MS-related genes. According to the result, these 42 target genes were regulated by 36 target EBV miRNAs (Fig. 5a, Additional file 2: Figure S1). To further explore the impact of EBV miRNAs in the pathogenesis of MS, 19 closely associated MS risk pathways regulated by EBV miRNAs were screened out through KEGG pathway enrichment analysis and hypergeometric test (Additional file 1: Table S3). Subsequently, a pathway-gene network among these 19 MS risk pathways and their enriched genes were constructed, and a total of 5 hub genes \( \text{MALT1, BCL10, IFNGR2, STAT3 and CDK6} \) were obtained (Fig. 5b, Additional file 2: Figure S2). In addition, a PPI network of these 42 target proteins were established by STRING database, 10 hub genes of which were selected according to betweenness centrality (Fig. 5c). We found that forkhead box p1 \( \text{(FOXP1)} \) has a close
relationship with mucosa-associated lymphoid tissue lymphoma transport protein 1 (MALT1) and BCL10 immune signaling adaptor (BCL10) in the PPI network, which should also be a focus point in our study. Finally, we obtained 15 target EBV miRNAs and their regulated, 6 MS-related genes (Fig. 5d). Surprisingly, MALT1 has been confirmed as a target of ebv-miR-BHRF1-2-5p by luciferase reporter assay in our research[37]. Meanwhile, ebv-miR-BART11 has been shown to directly target FOXP1[38, 39].

Discussion

Over the past decades, epidemiological researches have shown that viruses infection are closely associated with the pathogenesis and development of MS, including EBV, CMV, HIV, measles virus and influenza virus[6–10]. In this study, we obtained the transcriptomics datasets of these viruses and MS patients from the GEO database, and the correlation between EBV infection and MS risk was the most significant according to systematic analysis. The result was in line with previous studies, Rosella et al found that three herpesviruses showed statistical significance through the analysis between MS susceptibility genes and 20 interactomes, with EBV showing higher levels of significance compared to Human Herpesvirus 8 and, more evidently, to CMV[40]. Therefore, we dived deep into the roles of EBV in MS pathogenesis.

MS is caused by the interaction between genetic susceptibility and environmental triggers. International Multiple Sclerosis Genetics Consortium (IMSGC) has identified 284 MS susceptibility genes through GWAS of more than 100,000 MS patients and healthy controls, which can account for 25% of the perceived heritability of MS[2–4]. Thus, hereditary susceptibility did not solely dominate the alteration of gene expression during MS development, environmental factors also played an important role, suggesting both susceptibility genes and DEGs should be regarded as MS-related genes. We then obtained a total of 873 MS-related genes by integrating 284 susceptibility genes and 601 DEGs from B cells. After KEGG pathway enrichment analysis, we found that MS-related genes in B cells were mainly mapped into immune signaling pathways, including T helper type 17 (Th17) cell differentiation, Epstein-Barr virus infection, janus kinase (JAK)-signal transducer and activator of transcription (STAT) signaling pathway and NF-kappa B (NF-κB) signaling pathway. Among them, Th17 cells have been confirmed as pathogenic cells in the pathogenesis of MS and experimental autoimmune encephalomyelitis (EAE), the main animal models of MS. The JAK/STAT and NF-κB signaling pathways played an important role in Th17 cells differentiation and the development of EAE[41, 42]. Inhibition of the JAK/STAT pathway can lead to amelioration of clinical symptoms in EAE models[42]. In addition, EBV infection is a necessary environmental factor in the pathogenesis of MS. EBV can establish latency in the host memory B cells after primary infection, and be induced to activation when the immune system is impaired. However, the roles of EBV infection in the pathogenesis of MS have remained unclear.

Subsequently, we characterized the crucial roles of EBV infection in the pathogenesis of MS. Smith et al performed transcriptome analyses on resting B cells and two types of B blasts at 7 days after EBV infection or (CD40 Ligand) CD40L/interleukin (IL)-4 stimulation. There were more dysregulated genes in B blasts with EBV infection when compared to the original resting B cells, reflecting EBV infection played a facilitative role in B cell activation[21]. Thus, we obtained a total of 3556 DEGs from primary B cells after
EBV infection in vitro. Through intersecting these 3556 DEGs and 873 MS-related genes in B cells by the same trend, we identified 150 EBV-regulated, MS-related genes. Subsequently, a total of 18 closely associated MS risk pathways regulated by EBV were obtained through KEGG pathway enrichment analysis and hypergeometric test, including Viral protein interaction with cytokine and cytokine receptor, NF-κB signaling pathway and PD-L1 expression and PD-1 checkpoint pathway in cancer. After pathway-gene network analysis, HLA-DRB1 was confirmed as a vital factor in the development of MS that EBV was involved in. Recent studies revealed that B cells could express MHC class II and serve as antigen-presenting cells for activation of CD4+ T cells, which was critically involved in MS pathogenesis[43]. Meanwhile, another two molecules in the MS risk pathway-gene network regulated by EBV, MALT1 and BCL10, were important parts of CARMA1-BCL10-MALT1 (CBM) complex, which was a central mediator of T cell receptor and B cell receptor-induced NF-κB activation, and in this way drove lymphocyte activation[44]. These results indicated that EBV-infected B cells might lead to the activation of T cells via NF-κB pathway, which may point to the crucial molecular processes in the pathogenesis of MS.

In order to explore the regulatory effects and associated signaling pathways of EBV miRNAs in MS pathogenesis, we identified 42 MS-related genes regulated by 36 target EBV miRNAs. By constructing pathway-pathway and pathway-gene networks, a total of 5 hub genes (MALT1, BCL10, IFNGR2, STAT3 and CDK6) were obtained. Subsequently, a PPI network of these 42 target proteins were established by STRING database, and we found that FOXP1 had a close relationship with MALT1 and BCL10, which should also be a focus point in our study. Finally, we identified 15 target EBV miRNAs and their regulated, 6 MS-related genes. It is worth noting that ebv-miR-BHRF1-2-5p and ebv-miR-BHRF1-3 belonged to the same class cluster and had a synergistic effect. In this study, we found that ebv-miR-BHRF1-2-5p can target both MALT1 and BCL10, the components of CBM complex. MALT1 was a key molecule that promoted activation of NF-κB pathway, regulated regulatory T cells (Tregs) function and maintained immune tolerance. MALT1 knockout mice showed absence of Treg, increased Th1 and Th2 cells, which consequently led to lymphocyte infiltration and multi-organ inflammation[45]. Surprisingly, we firstly detected ebv-miR-BHRF1-2-5p and ebv-miR-BHRF1-3 were significantly increased in the circulation of RRMS patients. By luciferase reporter assay, MALT1 has been confirmed as a target gene of ebv-miR-BHRF1-2-5p in our research[37]. The target gene of ebv-miR-BHRF1-3, phosphatase and tensin homolog (PTEN), has also been confirmed. PTEN was a key regulator of the phosphatidylinositol 3-kinase (PI3K)/AKT (protein kinase B) survival pathway[46]. Furthermore, FOXP1 was closely associated with MALT1 and BCL10 in the PPI network and identified as a target gene of ebv-miR-BART11[38, 39]. FOXP1 was a critical regulator in maintaining Treg homeostasis and suppressive function. Mice with FOXP1-deficient Treg developed spontaneous inflammatory disease with age, which can lead to more severe EAE[47]. In addition, we found another two EBV miRNAs-regulated, MS-related genes, interferon gamma receptor 2 (IFNGR2) and STAT3 were both the upstream genes of programmed cell death ligand 1 (PD-L1) and PD-1 in the PD-L1 expression and PD-1 checkpoint pathway in cancer. PD-L1 and PD-1 knockout mice showed more severe inflammatory in EAE models[48]. Thus, EBV miRNAs could directly target MS-related genes, which consequently altered the expression and function of B cells, changed the autoimmune
response of T cells and promoted the initiation and development of MS through NF-κB (MALT1 and BCL10) and PD-L1/PD-1 (IFNGR2 and STAT3) pathways (Fig. 6).

In this study, we have identified that 15 target EBV miRNAs and their regulated, 6 MS-related genes. Among them, Song et al reported that FOXP1 could be directly regulated by ebv-miR-BART11. Meanwhile, ebv-miR-BHRF1-2-5p directly targeting MALT1 was confirmed by our experiments. However, there is a paucity of studies about their roles in clinical application. In addition, future work is required to investigate other target EBV miRNAs and their regulated MS-related genes.

**Conclusions**

This work indicated EBV infection showed the strongest correlation with MS risk. A total of 873 MS-related genes in B cells were obtained, which were mainly mapped into immune signaling pathways, particularly to Epstein-Barr virus infection. We then identified 150 MS-related genes and 18 associated risk pathways that EBV was involved in. Finally, 15 target EBV miRNAs and their regulated, 6 MS-related genes (MALT1, BCL10, IFNGR2, STAT3, CDK6 and FOXP1) have been confirmed as crucial pathogenic molecules, which could promote the initiation and development of MS through NF-κB (MALT1 and BCL10) and PD-L1/PD-1 (IFNGR2 and STAT3) pathways. Surprisingly, ebv-miR-BHRF1-2-5p directly targeting MALT1 was confirmed by our experiments. These findings may provide a novel insight into discovering diagnostic biomarkers and therapeutic targets for MS.

**Abbreviations**

MS: Multiple sclerosis; CNS: central nervous system; GWAS: genome wide association studies; EBV: Epstein-Barr virus; CMV: Cytomegalovirus; HIV: Human immunodeficiency virus; HLA-DRB: major histocompatibility complex, class II, dr beta; miRNAs: microRNAs; BHRF1: BamHI fragment H rightward open reading frame 1; BART: BamHI A rightward transcripts; RNF38: Ring Finger Protein 38; DEGs: differentially expressed genes; GEO: Gene Expression Omnibus; KEGG: Kyoto Encyclopedia of Genes and Genomes; PPI: protein-protein interaction; RRMS: relapsing-remitting MS; FOXP1: forkhead box p1; MALT1: mucosa-associated lymphoid tissue lymphoma transport protein 1; BCL10: BCL10 immune signaling adaptor; IMSGC: International Multiple Sclerosis Genetics Consortium; Th: T helper; JAK: janus kinase; STAT: signal transducer activator of transcription; NF-κB: NF-kappa B; EAE: experimental autoimmune encephalomyelitis; CD40L: CD40 Ligand; IL: interleukin; CBM: CARMA1-BCL10-MALT1; Tregs: regulated regulatory T cells; PTEN: phosphatase and tensin homolog; PIK3: phosphatidylinositol 3-kinase; AKT: protein kinase B; IFNGR2: interferon gamma receptor 2; PD-L1: programmed cell death ligand 1

**Declarations**

**Ethics approval and consent to participate**

Not applicable.
Consent for publication

Not applicable.

Availability of data and materials

The datasets used during the current study are available as listed below:

- Processed gene expression datasets of MS-related virus and MS patients are available on the NCBI Gene Expression Omnibus (GEO; http://www.ncbi.nlm.nih.gov/geo) under accession number GSE45829, GSE49628, GSE17948, GSE24238, GSE33877, GSE32140, GSE980, GSE117935, GSE66988, GSE41890 and GSE37750[9, 10, 21-29].
- The human target genes of 44 mature EBV miRNAs are available on the MiRTarbase (http://mirtarbase.mbc.nctu.edu.tw/) and TarBase (http://carolina.imis.athena-innovation.gr/diana_tools/web/index.php?r=tarbasev8%2Findex)[34, 35].

Competing interests

The authors declare that they have no competing interests.

Funding

This work was supported by a grant from the Provincial Natural Science Foundation of Heilongjiang province (No. ZD2020H004). Funding bodies had no role in the design of the study, analysis, or interpretation of data or in writing the manuscript.

Authors' contributions

J. F, C. X, X. R, Y. L and J. W designed the project. X. R, Y. L, Y. L and Y. W collected and analyzed data. All authors contributed to the discussions. X. R, Y. L and J. W wrote the manuscript. Funding was supplied by grants to J. F and C. X. All authors read and approved the final manuscript.

Acknowledgements

Not applicable.

Authors' information

1 Department of Neurology, the Second Affiliated Hospital of Harbin Medical University, Harbin Medical University, Harbin, Heilongjiang Province, China. 2 College of Bioinformatics Science and Technology, Harbin Medical University, Harbin, Heilongjiang Province, China.

References


Figures
Figure 1

A study flowchart. a Evaluation of correlation between MS-related viruses infection and MS risk. b Identification of MS-related genes regulated by target EBV miRNAs.

Figure 2

The correlation between MS-related viruses infection and MS risk. This result was calculated by hypergeometric test, and the p-value was log processed.
Figure 3

MS-related genes and their risk pathways in B cells. a The integration of MS susceptibility genes and DEGs from B cells between MS patients and healthy controls. b The top 10 MS risk pathways based on KEGG pathway enrichment analysis, and the p-value was log processed.
Figure 4

The MS risk pathway-pathway and pathway-gene network which were regulated by EBV. a The pathway-pathway network between MS risk pathways regulated by EBV. The blue diamonds represent MS risk pathways, and the line between two diamonds represents a significant correlation between these two pathways. b The number of associated pathways of each MS risk pathway in the pathway-pathway network. c The pathway-gene network between MS risk pathways which were regulated by EBV and their enriched genes. The blue diamonds represent MS risk pathways, the yellow nodes represent enriched genes, and the red nodes represent hub genes. d The number of enriched pathways of each hub gene in the pathway-gene network.
Figure 5

Target EBV miRNAs and their regulated, MS-related genes. a A total of 36 target EBV miRNAs and the amount of MS-related genes regulated by each EBV miRNA. b The pathway-gene network between MS risk pathways regulated by EBV miRNAs and their enriched genes. The blue diamonds represent MS risk pathways, and the red nodes represent hub genes. c A PPI network of 42 EBV miRNAs-regulated, MS-related genes. The yellow nodes represent MS-related genes, and the red nodes represent hub genes. d 15 target EBV miRNAs and their regulated, 6 MS-related genes.
Figure 6

EBV miRNAs could directly target MS-related genes in NF-κB and PD-L1/PD-1 pathways. The red frames represent EBV miRNAs, the pink background color represent their regulated MS-related genes, and the white background color represent genes in NF-κB and PD-L1/PD-1 pathways.

Supplementary Files

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

- FigureS1.pdf
- FigureS2.pdf
- TableS1.xls
- TableS2.xls
• TableS3.xls