



Identification of immune-related diagnostic markers in primary Sjögren's syndrome based on bioinformatics analysis

Qingle Zeng, Jing Wen, Leting Zheng, Wen Zeng, Shuyuan Chen, Cheng Zhao

Department of Rheumatology and Immunology, the First Affiliated Hospital of Guangxi Medical University, Nanning, China

Contributions: (I) Conception and design: C Zhao; (II) Administrative support: J Wen; (III) Provision of study materials or patients: W Zeng; (IV) Collection and assembly of data: L Zheng, S Chen; (V) Data analysis and interpretation: Q Zeng; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Cheng Zhao. Department of Rheumatology and Immunology, the First Affiliated Hospital of Guangxi Medical University, 6 Shuangyong Road, Qingxiu District, Nanning 530021, China. Email: zhaochengx@163.com.

Background: Primary Sjögren's syndrome (pSS) is a relatively common diffuse connective tissue disease that often invades exocrine glands, such as the lacrimal and salivary glands, and manifests as dry eyes and dry mouth. At present, the molecular mechanism of pSS is not clear. This study was designed to explore the internal mechanism of pSS from the gene level and screen out the immune-related diagnostic markers of pSS.

Methods: The gene expression profiles GSE84844, GSE7451, and GSE40611 were downloaded from the Gene Expression Omnibus (GEO) database. The differentially expressed genes (DEGs) were identified with R software. Then, the DEGs were intersected with the immune genes obtained from the ImmPort database to acquire differentially expressed immune-related genes (DEIRGs), and functional enrichment analyses were performed. The DEIRGs were screened through the least absolute shrinkage and selection operator (LASSO) logistic regression algorithm to obtain the optimal immune-related genes (IRGs). Expression levels of the optimal IRGs were verified by quantitative reverse transcription-polymerase chain reaction (qRT-PCR) to obtain the key genes. Next, gene chips GSE7451 and GSE40611, from other tissues, were selected as the training sets to verify the sensitivity and specificity of the diagnosis of the key genes by receiver operating characteristic (ROC) analysis.

Results: A total of 54 DEIRGs were obtained. The functional enrichment analysis results showed that they play an important role in immune and inflammatory responses. Nine optimal IRGs were screened from the DEIRGs by the LASSO logistic regression algorithm. After qRT-PCR verification, eight out of nine optimal IRGs (*IL-18*, *JAK2*, *TBK1*, *EED*, *TNFSF10*, *TNFSF13B*, *CYSLTR1*, and *ICOS*) were significantly highly expressed in pSS patients and were defined as key genes. ROC analysis identified that *TNFSF13B* and *CYSLTR1* had high sensitivity and specificity. Finally, the lack of previous research on *EED* and *CYSLTR1* in pSS suggests that these IRGs may be regarded as new gateways to explore the diagnosis and pathogenesis of pSS.

Conclusions: The key DEIRGs play a decisive role during the occurrence and development of pSS.

Keywords: Primary Sjögren's syndrome (pSS); bioinformatics analysis; differentially expressed immune-related genes (DEIRGs); least absolute shrinkage and selection operator (LASSO); quantitative reverse transcription-polymerase chain reaction (qRT-PCR); receiver operating characteristic (ROC)

Submitted Jan 13, 2022. Accepted for publication Apr 20, 2022.

doi: 10.21037/atm-22-1494

View this article at: <https://dx.doi.org/10.21037/atm-22-1494>

Introduction

Primary Sjögren's syndrome (pSS) is a common chronic autoimmune rheumatic disease. It is characterized by abnormal proliferation of B lymphocytes and invasion of exocrine glands, such as the salivary and lacrimal glands, and can even lead to lymphoma. The increase of antinuclear antibodies, anti SSA, anti SSB antibodies as well as rheumatoid factors and hypergammaglobulinemia can often be detected in pSS patients by laboratory tests. PSS affects about 0.3–3% of the global population and is more common in women between 30 and 50 years of age, with a female to male ratio of approximately 9:1 (1,2). The main clinical manifestations of pSS are dry mouth and dry eyes. In addition to exocrine gland dysfunction, there may also be symptoms such as fatigue, joint pain, and rash. Approximately one in three patients may also have serious systemic damage, such as interstitial pneumonia, tubulointerstitial nephritis, neuropathy, or hemocytopenia, leading to a poor prognosis and a significant social burden (3,4).

The exact etiology and pathogenesis of pSS are still unclear. Many factors such as heredity, infection, and environment can induce an autoimmune response, which can lead to pSS. At present, the research on pSS lacks depth, especially regarding the molecular mechanism, and there is a paucity of sensitive and specific diagnostic indicators. Moreover, the initial symptoms of pSS are mild and often ignored, and there are no specific therapeutic drugs recommended at present. Due to the lack of targeted drugs, the treatment of patients is often delayed, and the curative effect is not good. Therefore, identifying disease-associated biomarkers is of utmost importance in diagnosing and treating pSS, especially in the early stage.

In recent years, much research has been undertaken to explore new diagnostic markers and therapeutic targets of pSS. Many studies have attempted to find the expression characteristics of a series of genes and identify key genes using various research methods (5,6). At present, it has been clearly found that epigenetic deregulation exists in pSS, especially DNA methylation, and is limited to specific cell subsets, such as lymphocytes and salivary gland epithelial cells. Notably, a comprehensive analysis of DNA methylation in circulating B cells and in small salivary glands biopsies from pSS patients, confirmed significant hypomethylation of IFN-regulatory genes in whole blood and in B cells (7,8).

A previous study showed that genes related to immunity and inflammation are upregulated, and genes related to salivary secretion and carbohydrate digestion and

absorption are downregulated in pSS patients, suggesting that pSS is closely associated with the immune response and abnormal salivary gland function (9). Some researchers have also found new biomarkers of pSS through proteomic research. Possible novel biomarkers were screened by liquid chromatography-mass spectrometry (LC-MS) of saliva and tears from pSS patients, solely and in combination with size-exclusion chromatography (10). Another study used laser microdissection to separate tissue samples enriched in subjects' inflammatory lesions for RNA-seq analysis and found that the expression of two chemokines, *CCR7* and *CCL21*, was significantly increased (11). Recent research data has confirmed that Th17 subsets of CD4(+) CXCR5(+) T cells may participate in the antibody-related immune reaction in the blood, and CD4(+) CXCR5(+) CCR6(+) Tfh cells have a high-frequency expression in blood, suggesting that they may be appropriate biomarkers for pSS (12). By comparing the differential expression of long non-coding RNAs (LncRNAs) and mRNAs from peripheral blood mononuclear cells (PBMCs) in pSS patients and healthy controls, Peng *et al.* validated that *LINC00426*, *TPTEP1-202*, *CYTOR*, *NRIR*, and *BISPR* were aberrantly expressed and were closely related to the disease activity of pSS. The functional enrichment analysis showed that they were mainly concentrated in the immune response and immune system processes (13).

With the development of molecular biology, many studies on pSS have been conducted, and many biomarkers beneficial to diagnosis and treatment have been found (14,15). However, there is still significant variability among the research findings. Current research suggests that the onset of pSS is possibly a result of a complex interplay between an activated immune system and epithelial cells (targets and players of the autoimmune response), indicating that the immune mechanism is closely related to the pathogenesis of pSS. T lymphocytes are the critical factors in the pathogenesis of pSS immunity, and the activation of B cells accelerates the progress of the disease. Although the innate immune system plays an important role in the early stage of pSS, the exact role of regulatory cells and their related genes in pSS remains unclear (14-16).

Our study is the first to use bioinformatics to investigate immune-related diagnostic markers of pSS and explore the differential expression of immune-related genes between pSS patients and healthy people. In order to avoid the false-positive problem caused by a single study, this study analyzed the data of one gene chip to obtain the key differential immune-related genes, and then carried

out experimental verification, used the other two gene chips from different tissue sources for receiver operating characteristic (ROC) analysis, and finally obtained the key genes most closely related to the onset of pSS. We consider that providing a bioinformatics method to study the possible pathogenesis of pSS plays an essential role in improving pSS diagnosis and treatment (17,18). We present the following article in accordance with the STREGA reporting checklist (available at <https://atm.amegroups.com/article/view/10.21037/atm-22-1494/rc>).

Methods

Microarray data

Gene expression profiles of GSE84844, GSE7451, and GSE40611 were downloaded from the Gene Expression Omnibus (GEO) database (<https://www.ncbi.nlm.nih.gov/geo/>) (GPL570 platform, Affymetrix Human Genome U133 Plus 2.0 Array). GSE84844 contained 60 samples, including 30 whole blood samples from healthy controls and 30 from pSS patients. GSE7451 contained 20 samples, including 10 saliva samples from healthy controls and 10 from pSS patients. GSE40611 contained 49 samples, but 14 samples from non-pSS sicca patients were excluded, leaving a total of 17 gland tissues samples from pSS patients and 18 from healthy controls (17).

Identification of DEIRGs

The gene expression matrices were obtained by R software (version 3.6.3; <https://www.r-project.org/>), and the “Bioconductor” R package (<http://www.bioconductor.org/>) was used for the data analyses. The linear models for the microarray data (LIMMA; <https://www.bioconductor.org/packages/release/bioc/html/limma.html>) package in “Bioconductor” was applied to screen the differentially expressed genes (DEGs). Only genes that met the cut-off value of $P < 0.05$ and fold change (FC) $= 1.5$ ($|\log_2 \text{FC}| \geq 0.5$) were chosen as DEGs. Differentially expressed immune-related genes (DEIRGs) were acquired by matching 1,793 immune-related genes (IRGs) from the ImmPort database (<https://www.immport.org/shared/genelists>) IRG dataset to DEGs using Venn online software (<http://bioinformatics.psb.ugent.be/webtools/Venn/>) (19). The heatmap for the DEIRGs was constructed with the “heatmap” package in R software (<https://bioconductor.org/packages/release/bioc/html/heatmaps.html>). The volcano map of DEIRGs

was created to show the differential expression by using the “ggplot2” package (20,21).

Functional and pathway enrichment analysis of DEIRGs

Using Gene Ontology (GO) enrichment analysis, we aimed to determine the cellular components (CCs), biological processes (BPs), and molecular functions (MFs) of the DEIRGs. The gene routes and related functions of the DEIRGs were determined with the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis. The Database for Annotation, Visualization, and Integrated Discovery (DAVID; <https://david.ncifcrf.gov>) was used to perform the KEGG pathway and GO enrichment analyses. $P < 0.05$ was considered to be statistically significant (22,23).

Screening of the optimal IRGs

We used the least absolute shrinkage and selection operator (LASSO) with the “glmnet” package in R to screen DEIRGs, to find the optimal diagnostic marker IRGs of pSS. $P < 0.05$ was considered to be statistically significant (24,25).

Quantitative reverse transcription-polymerase chain reaction (qRT-PCR)

Nine optimal IRGs were selected as candidate whole blood biomarkers of pSS for further qRT-PCR verification. For this, a total of 22 participants, including 13 pSS patients and 9 healthy volunteers, were recruited from inpatient and outpatient departments of the First Affiliated Hospital of Guangxi Medical University. There was no significant difference in age and gender between the disease and control groups. The pSS patients fulfilled the 2016 American College of Rheumatology (ACR)/European League Against Rheumatism (EULAR) classification criteria (26) or the 2012 ACR criteria (27). Patients who had hereditary diseases, cancer, or other autoimmune diseases were excluded. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). All participants signed written informed consent before starting the study. This research was approved by the First Affiliated Hospital of Guangxi Medical University Ethics Committee [approval No. 2022-KY-E-(095)].

Whole blood samples were collected from each participant before initial therapy. Subsequently, red blood cell lysate (Solarbio) was added, lysed on ice for 10 min, and the precipitate was collected after centrifugation. Total RNA

Table 1 Oligonucleotides used in this study

Gene_ID	Gene name	Sequences (5'-3')	Product length (bp)
3606	<i>IL-18</i>	Forward: CTGACTGTAGAGATAATGCACCCC Reverse: TGATGTTATCAGGAGGATTCATTC	171
3717	<i>JAK2</i>	Forward: TGCTACAGTGCTGGTCGGC Reverse: CCATACCCTTGCCATATCTGAGATG	145
29110	<i>TBK1</i>	Forward: ATGTGGTGGGTGGAATGAATC Reverse: ACATATCAGGGTGCAAATATTCTTC	201
8726	<i>EED</i>	Forward: TGTAGGAAGCAACAGAGTTACCTTG Reverse: CCAGCAGAGGATGGCTCGTAT	146
8743	<i>TNFSF10</i>	Forward: GAGCTGAAGCAGATGCAGGAC Reverse: GCCACTTGACTTGCCAGCA	127
10673	<i>TNFS13B</i>	Forward: CCTCACGGTGGTGTCTTTCTAC Reverse: GGTTCAAAGATTTTCAGTCCCG	171
10800	<i>CYSLTR1</i>	Forward: TCGAATTTACTGAAGACTTGGAGC Reverse: TGAATACACTTGATTGCGGAA	189
29851	<i>ICOS</i>	Forward: GGCTTACAAAAAGAAGTATTCATCC Reverse: GGGTGCCAGAGTCCATATTATA	131
9034	<i>CCRL2</i>	Forward: CATAGAAGGTGAACTGGAGAGCG Reverse: AGGATAAGCACAACCAGGAGATT	144
2597	<i>GAPDH</i>	Forward: AATCAAGTGGGGCGATGCTG Reverse: GCAAATGAGCCCCAGCCTTC	86

was extracted from each sample using TRIzol (Invitrogen), and then the reverse transcription was conducted using the MonScript™ 1st Strand cDNA Synthesis Kit (Monad) at 55 °C for 15 minutes and 85 °C for 5 minutes. Next, qPCR was performed with a QuantiNova SYBR Green PCR Kit (QIAGEN) at a temperature of 95 °C for 2 minutes, followed by 40 cycles with a temperature of 95 °C for 5 seconds, 60 °C for 40 cycles for 30 seconds. Finally, the $2^{-\Delta\Delta C_t}$ method was used to determine the relative expression of each selected mRNA between cases and controls and "ggplot2" package was used to draw violin diagram to visualize the results. The primer sequences used in the study are shown in *Table 1* (28).

Verification of optimal IRGs by ROC analysis in training sets from other organizations

The key genes were verified by qRT-PCR. To evaluate the

role of key genes in the other tissues of pSS, we validated the key genes above by performing ROC analysis with the training dataset GSE7451 and GSE40611 using the "pROC" R package. The key genes with an area under the curve (AUC) >0.7 and a P value <0.05 were deemed to be meaningful (29).

Statistical Analysis

SPSS 17.0 statistical software was used to process the obtained data, non-parametric test was used for the obtained non normal distribution data, and Mann-Whitney U test was used to analyze the significance of the difference between the two groups. Through these processes, the expression difference of the optimal IRGs in pSS peripheral blood was obtained. The difference was statistically significant (P<0.05).

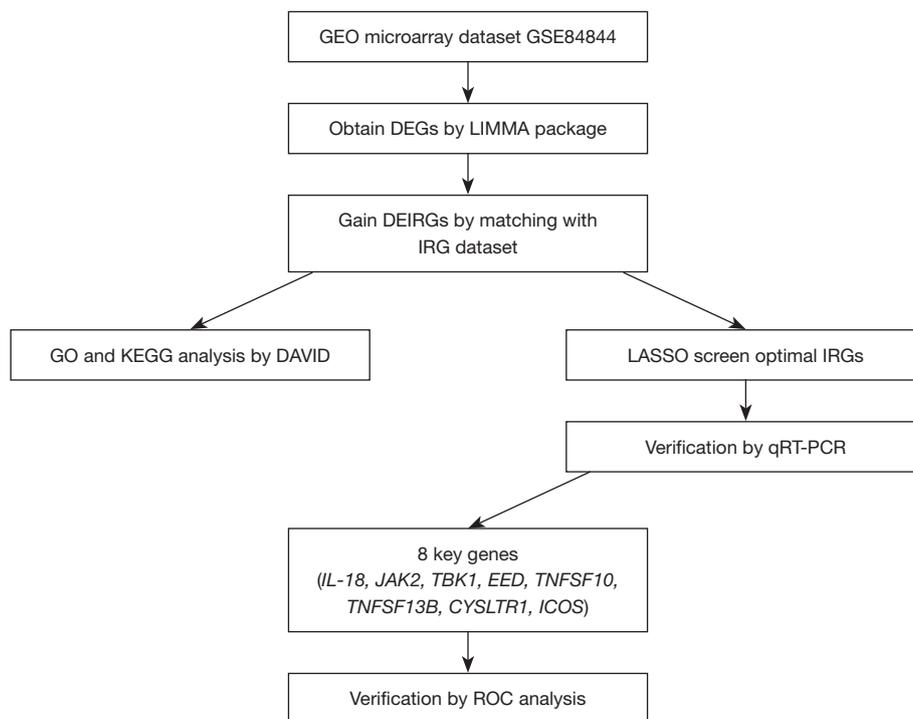


Figure 1 Summary of research workflow. LASSO, Least absolute shrinkage and selection operator; GEO, Gene Expression Omnibus; DEGs, differentially expressed genes; DEIRGs, differentially expressed immune-related genes; IRG, immune-related gene; GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; qRT-PCR, quantitative reverse transcription-polymerase chain reaction; ROC, receiver operating characteristic.

Results

Identification of DEIRGs

The workflow of this study is shown in *Figure 1*. A total of 552 DEGs, including 29 downregulated and 523 upregulated genes, were detected. By matching 1,793 IRGs from the ImmPort database, 54 DEIRGs were identified, including 3 downregulated genes and 51 significantly upregulated genes (*Figure 2A*). *Figure 2B* shows the DEIRGs expression heatmap. The results were proved by a volcano plot of all downregulated genes and upregulated genes (*Figure 2C*).

Enrichment analysis of the DEIRGs

All DEIRGs were uploaded to the online software DAVID to identify the GO categories. The different GO classification and expression changes of the DEIRGs impacted the results of the GO enrichment analysis terms (*Figure 3*). “Immune

response”, “inflammatory response”, “defense response to virus”, “innate immune response”, and “negative regulation of viral genome replication” exhibited highly significant enrichment within the GO BP category. For the CC category, DEIRGs were significantly enriched mostly in “cytoplasm”, “extracellular space”, “integral component of plasma membrane”, “external side of plasma membrane”, and “extracellular region”. In the MF category, DEIRGs were significantly enriched in “ATP binding”, “double-stranded RNA binding”, “chemokine activity”, “single-stranded RNA binding”, and “hydrolase activity, acting on acid anhydrides”. KEGG pathway enrichment analysis was used to clarify the signaling pathway enrichment of DEIRGs (*Figure 4*). With $P < 0.05$, the top enriched biological pathways of pSS included “Influenza A”, “Herpes simplex infection”, “Measles”, “Cytokine-cytokine receptor interaction”, “Hepatitis B”, “Hepatitis C”, “TNF signaling pathway”, “RIG-I-like receptor signaling pathway”, “Toll-like receptor signaling pathway”, and “Cytosolic DNA-sensing pathway” (14).

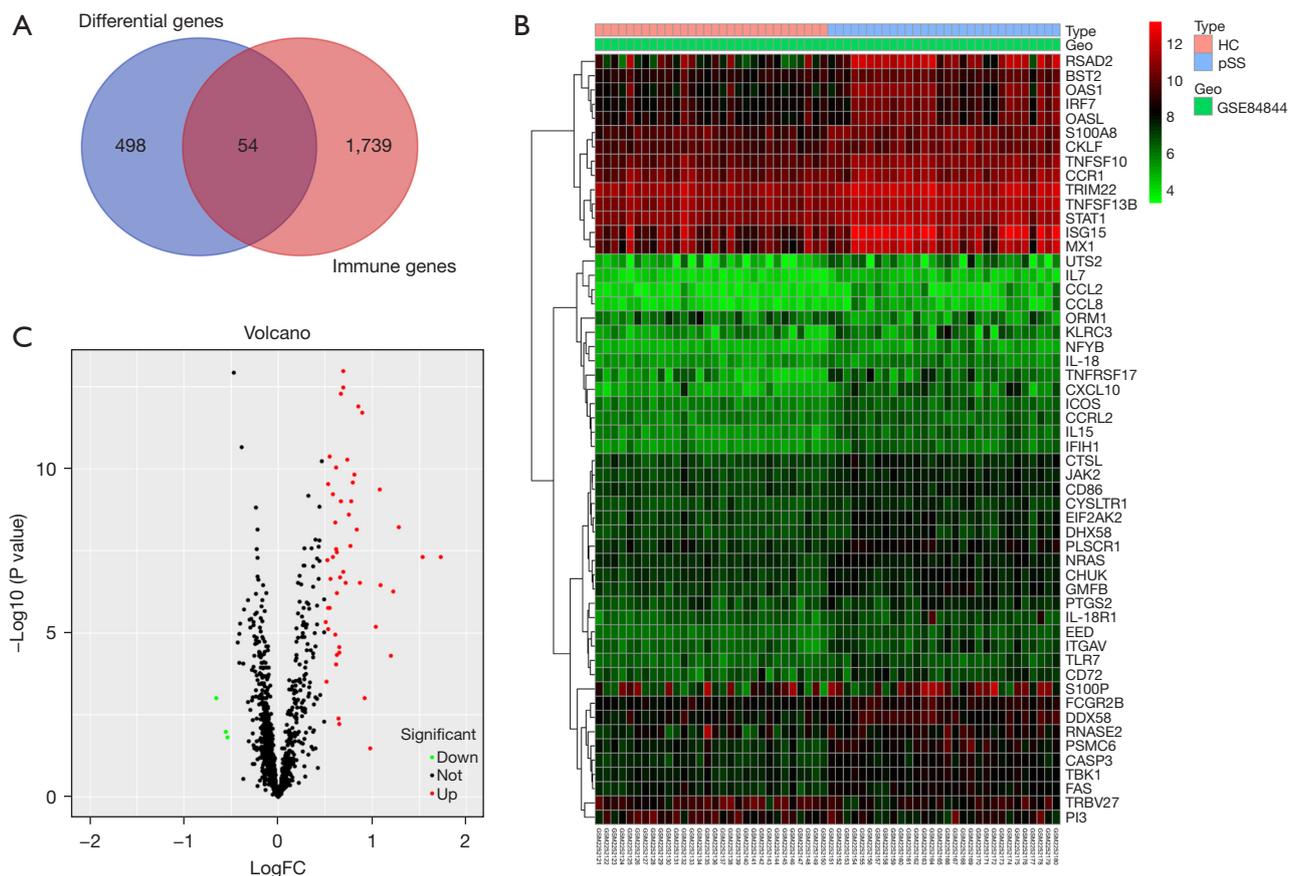


Figure 2 Venn diagram (A) shows the intersection of DEGs and IRGs. Heatmap (B) and volcano plot (C) showing the DEIRGs between pSS patients and normal controls. Red dots represent upregulated, green dots represent downregulated, and black dots represent no difference in DEIRGs (the differences are set as $P < 0.05$ and $|\log_2FC| \geq 0.5$). DEGs, differentially expressed genes; IRGs, immune-related genes; DEIRGs, differentially expressed immune-related genes; pSS, primary Sjögren's syndrome; FC, fold change.

Screening of potential diagnostic markers

A LASSO logistic regression algorithm was applied to identify nine optimal IRGs from all DEIRGs as potential diagnostic markers for pSS (Figure 5): *IL-18*, *JAK2*, *TBK1*, *EED*, *TNFSF10*, *TNFSF13B*, *CYSLTR1*, *ICOS*, and *CCRL2*.

Verification of potential biomarker expression by qRT-PCR

To further investigate the actual expression level of nine optimal IRGs (*IL-18*, *JAK2*, *TBK1*, *EED*, *TNFSF10*, *TNFSF13B*, *CYSLTR1*, *ICOS*, and *CCRL2*) in whole blood cells between pSS patients and healthy volunteers, we determined the gene expression in whole blood cells of pSS patients and healthy volunteers by quantitative real-time PCR (qRT-PCR). The results of the violin plot showed

that the expression of eight of the nine genes was in accord with the predicted results, and all of them were upregulated (Figure 6). These eight optimal IRGs were defined as key genes (*IL-18*, *JAK2*, *TBK1*, *EED*, *TNFSF10*, *TNFSF13B*, *CYSLTR1*, and *ICOS*).

Further verification of optimal IRGs by ROC analysis

To further test the diagnostic efficacy of the eight key genes, we performed ROC analysis in R Studio to validate it with the GSE7451 and GSE40611 datasets as the training sets (Figure 7). These two training datasets, GSE7451 and GSE40611, were from the saliva and parotid glands, respectively. Finally, the results showed that two of the key genes, *TNFSF13B* and *CYSLTR1*, had an AUC > 0.70 and could diagnose pSS patients with excellent specificity and

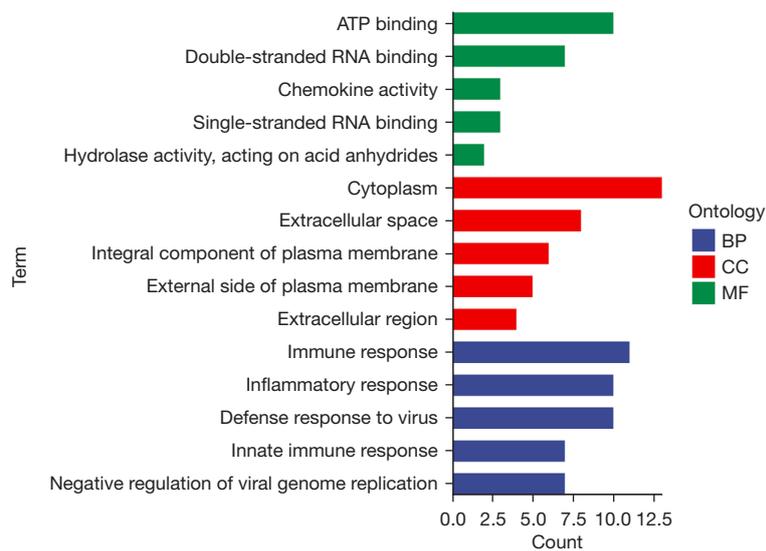


Figure 3 GO analysis results of DEIRGs between pSS patients and normal controls (top 5 according to P values in BP, CC, and MF, respectively). GO, Gene Ontology; BP, biological processes; CC, cellular components; MF, molecular functions; DEIRGs, differentially expressed immune-related genes; pSS, primary Sjögren's syndrome.

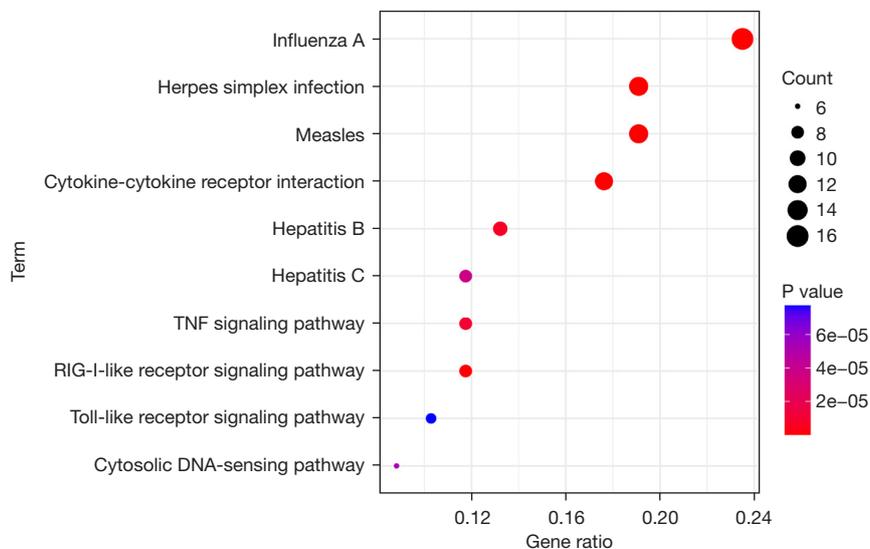


Figure 4 KEGG analysis results of DEIRGs between pSS patients and normal controls (top 10 according to P values). KEGG, Kyoto Encyclopedia of Genes and Genomes; DEIRGs, differentially expressed immune-related genes; pSS, primary Sjögren's syndrome.

sensitivity in whole blood, saliva, or the parotid gland.

Discussion

PSS is one of the most prevalent autoimmune connective tissue diseases in middle-aged and older women. Although

many scholars have found new biomarkers and therapeutic targets of pSS from molecular biology and genetics, the possible etiology and potential pathogenesis of pSS are still unclear, and there are many differences of opinion regarding the explicit molecular mechanism and biological diagnostic criteria for pSS. Therefore, a molecular biology

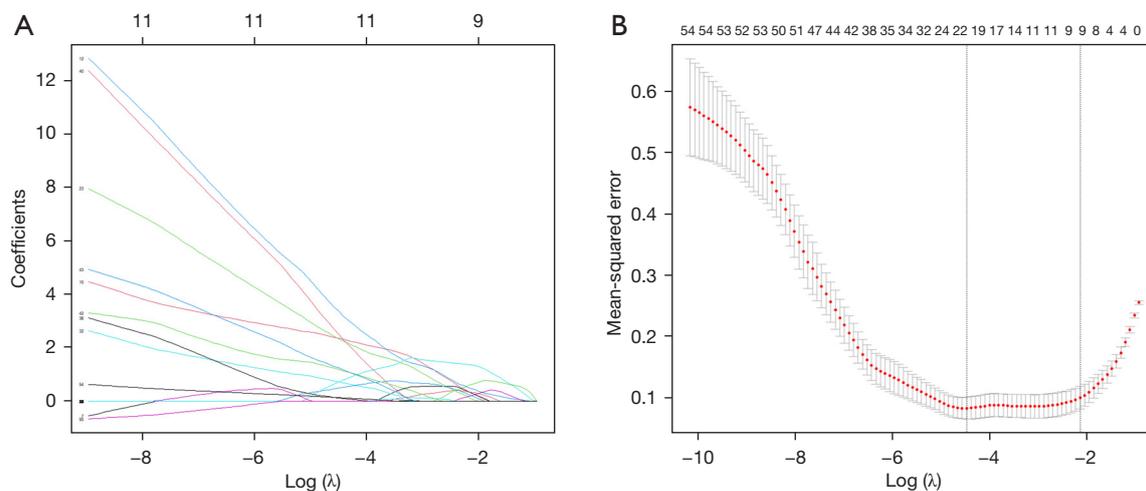


Figure 5 LASSO regression algorithm to screen the optimal IRGs in DEIRGs. LASSO, Least absolute shrinkage and selection operator; IRGs, immune-related genes; DEIRGs, differentially expressed immune-related genes.

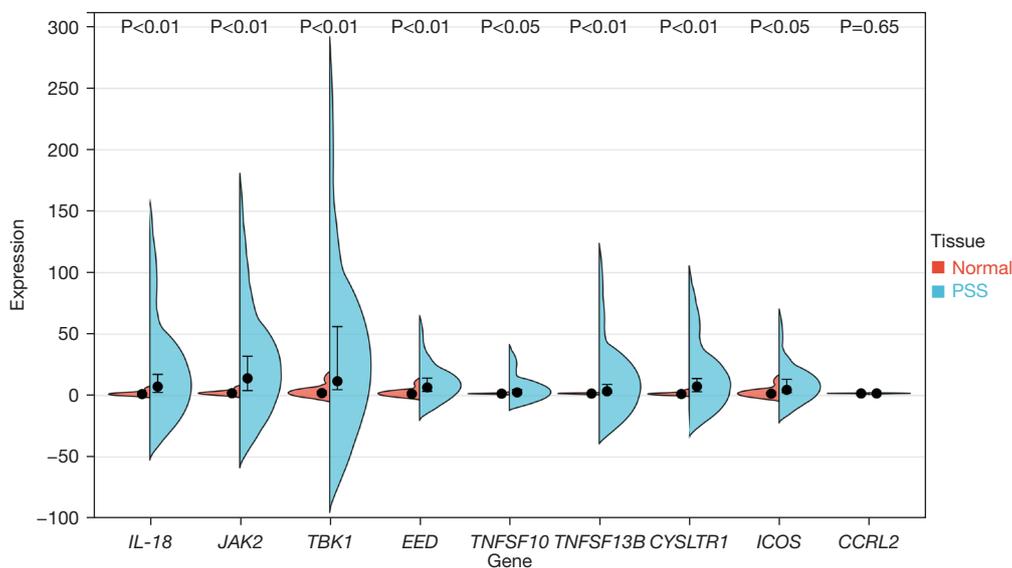


Figure 6 Violin diagram of the gene expression proportion of nine optimal IRGs between pSS patients and normal controls. IRGs, immune-related genes; pSS, primary Sjögren's syndrome.

approach is necessary to accurately identify biomarkers related to disease pathogenesis to study the occurrence and development of diseases, allowing the pathogenic mechanism to be explored from the root. Anti SSA and anti SSB antibodies are the most important biomarkers found so far and have been included in the classification criteria of pSS. Antinuclear antibody and rheumatoid factor are also commonly used parameters in pSS as clinical and diagnostic tools, but they only occupy part of the classification

criteria. It is worth noting that after so many years, no new biomarkers have been established (30,31).

The pathogenesis of autoimmune diseases is very closely related to immune-related mechanisms. For this reason, we should start with the immune mechanism as the basis for further research on pSS pathogenesis. Thus, it can be seen that mining of IRGs is crucial to more accurately explore pathogenesis, find effective treatment strategies, and provide early diagnosis and accurate treatment of pSS.

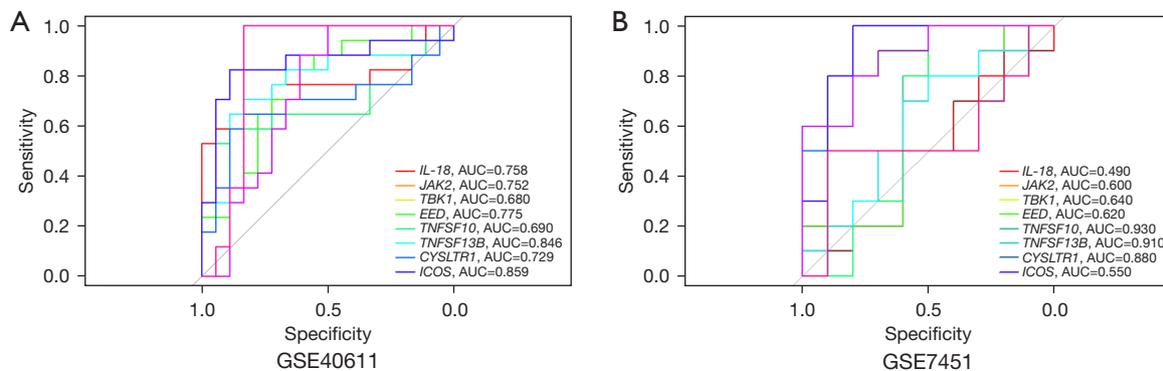


Figure 7 The ROC curves of the eight key genes in the GSE40611 and GSE7451 training sets. ROC, receiver operating characteristic.

As far as we know, this study is the first to explore pSS IRGs in combination with clinical verification to identify more convincing biomarkers (32,33). In this study, a total of 54 DEIRGs were identified. The impact of DEIRGs is often to initiate immune-related fundamental molecular changes.

To further clarify the underlying mechanism of pSS, functional enrichment analysis of the DEIRGs was performed. The GO terms implicated in pSS pathogenesis included immune response, inflammatory response, defense response to virus, and innate immune response. Furthermore, in the KEGG enrichment analysis, the DEIRGs were closely associated with immune-related pathways and infection-related pathways, including the virus infection-related signal pathways, such as Influenza A, Cytokine-cytokine receptor interaction, TNF signaling pathway, RIG-I-like receptor signaling pathway, and Toll-like receptor (TLR) signaling pathway. Among them, one of the important triggers of pSS onset appears to be viral infection. Some studies have indicated that the viral-induced autoimmune process is triggered by various mechanisms (34,35). It has been reported that higher antibody responses in untreated patients with pSS following non-adjuvant viral antigen exposure and viral infections play an important part in environmental risk factors for autoimmune disease (36). A related study has showed that patients with pSS respond to immunization with increased antibody responses compared with healthy people (37). The cytokine-cytokine receptor interaction pathway is an immune-related signal pathway, which plays an important role in inflammatory host defense, cell growth, and differentiation (38). TLRs, especially TLRs 7–9, are important innate immune receptors. Among them, TLR-7 signaling for type I IFNs is significantly expressed in ducts and mononuclear cells from pSS patients (39).

Currently, one study shows that TLR7 pathway regulation might be a potential new treatment option for patients with pSS-associated thrombocytopenia (40). Based on the results of our GO and KEGG analyses, it can be explained that these DEIRGs are closely related to immune-related pathways and can serve as potential biomarkers for pSS.

We screened nine optimal IRGs from DEIRGs using a LASSO logistic regression algorithm. We identified eight key genes by qRT-PCR, including *IL-18*, *JAK2*, *TBK1*, *EED*, *TNFSF10*, *TNFSF13B*, *CYSLTR1*, and *ICOS*, which were all upregulated in the previous prediction. The expression of these key genes is consistent with our prediction. Their expression was significantly higher than that of the healthy control group in our study. Some of them have been reported as potential biomarkers and are possibly involved in the pathogenesis of pSS.

IL-18, also known as *IGIF*, *IL-1g*, and *IL1F4*, acts as a proinflammatory cytokine and plays a role in producing other inflammatory cytokines, such as IFN- γ , TNF- α , and IL-1, and leads to the aggregation of B lymphocytes. Researchers have found elevated *IL-18* serum levels in patients with pSS, which may be related to the occurrence of the disease (41–43). *JAK2* is also known as *JAK10*. Current research shows that *JAKs* participate in different inflammatory, autoimmune diseases, and malignancies by activating the JAK/STAT signaling pathway. The JAK/STAT pathway is an intracellular signal transduction pathway that can be triggered by numerous cytokines and leads to the additional production of pro-inflammatory, anti-inflammatory cytokines, and locally destructive enzymes. Current clinical studies show that JAK inhibitors have been administered to many patients with autoimmune diseases, including pSS, with satisfactory curative effects

(44-46). Relevant research also indicates that increased gene expression of *TBK1* in whole blood was observed in pSS patients compared to healthy controls. *TBK1* is also called *NAK*, *T2K*, *IIAE8*, and *FTDALS4*. IFN-I has shown to be upregulated in pSS patients, and *TBK1* can induce transcription and eventually lead to the production of IFN-I, which suggests that *TBK1* as a signaling hub might be a latent treatment target of pSS (47). *TNFSF10* is also known as *TRAIL*. A study of pSS minor salivary glands indicated that *TRAIL* protein expression on infiltrating mononuclear cells and *TRAIL* receptors on ductal epithelial cells confirmed that the *TRAIL* system's role in regulating autoimmunity in pSS salivary glands is non-negligible. *TRAIL* has a regulatory role in the autoimmune process not only in the salivary glands but also in pSS B lymphocytes (48). *ICOS* is also known as *AILIM*, *CD278*, and *CVID1*. According to current research findings, *ICOS* is involved in the occurrence of pSS by promoting immune infiltration and is considered one of the promising biomarkers for detecting pSS (49,50).

The high expression of these above-mentioned genes in pSS is consistent with the experimental results of our study. Among the eight key genes, there is currently little research on the involvement of *EED* and *CYSLTR1* in the pathogenesis of pSS. Therefore, it is reasonable to envisage that these two genes could become novel gateways for our future research on pSS.

In addition, we tested the diagnostic effects of the eight key genes by using two data sets from saliva and parotid tissue as validation sets. The results showed that *TNFSF13B* and *CYSLTR1* had high sensitivity and specificity in whole blood, saliva, and parotid tissue of pSS patients. Many studies have shown that *TNFSF13B* participates in the pathogenesis of pSS through B lymphocyte immunity (51,52). *TNFSF13B* has another common name, *BAFF*. The protein encoded by this gene is a cytokine, which is expressed in B cell lineage cells and acts as a potent B cell activator. Current research shows that it plays an essential role in the proliferation and differentiation of B cells. Elevated *BAFF* levels are often present in patients with pSS, supporting autoimmunity and potentially blocking therapeutic elimination of pathogenic B cell clones, thus playing an important role in the immune-related pathogenesis of pSS (53,54). Present research shows that *BAFF* levels in circulating and salivary gland tissue are elevated in pSS patients and closely related to increased disease activity. Therefore, it can be inferred that *BAFF* could be a potential target for diagnosis and therapy in pSS since

it is one of the links between innate and adaptive immune responses (55-57).

CYSLTR1 is also known as *CYSLT1*, *CYSLTR*, *CYSLT1R*, and *HMTMF81*. Recent studies indicate that this gene is involved in the pathogenesis of inflammatory and immune diseases. It is reported that *CYSLTR1* plays a role in the pathogenic process of experimental autoimmune encephalomyelitis. Clinical studies have also shown that multiple sclerosis patients have a higher expression of *CYSLTR1* in the circulatory and central nervous system. But to date, there is no relevant research exploring the relationship between *CYSLTR1* and pSS (58,59). In our research, both *TNFSF13B* and *CYSLTR1* were not only confirmed by PCR but also had high diagnostic ability in whole blood, saliva, and the parotid gland, justifying that these two genes are both potential diagnostic markers for pSS. *CYSLTR1* may become a new gateway to explore the pathogenic mechanism of pSS in future research.

Conclusions

In summary, our study used transcriptome sequencing and bioinformatics analysis combined with clinical data to identify that eight key genes (*IL-18*, *JAK2*, *TBK1*, *EED*, *TNFSF10*, *TNFSF13B*, *CYSLTR1*, and *ICOS*) are highly expressed in pSS patients. These eight genes may be involved in the pathogenesis of pSS and have potential value for diagnosing pSS. Several studies have confirmed that *BAFF* is a good diagnostic marker of pSS (53-57), and some studies have also shown that *IL-18*, *JAK2*, *TBK1*, *TNFSF10*, and *ICOS* are related to the occurrence and development of pSS (41-50). But there is currently a lack of research exploring the role of *EED* and *CYSLTR1* in the pathogenesis of pSS. Our study shows that *EED* and *CYSLTR1* can be regarded as new avenues to explore the pathogenesis of pSS.

Acknowledgments

We acknowledge the GEO database for providing their platform and thank the contributors for uploading their meaningful datasets. And we thank all participants involved in this study.

Funding: This study was supported by the Guangxi Medical and Health Appropriate Technology Research and Development Project (No. S201304-03) and the Guangxi Medical and Health Appropriate Technology Development, Promotion, and Application Project (No. S2020039).

Footnote

Reporting Checklist: The authors have completed the STREGA reporting checklist. Available at <https://atm.amegroups.com/article/view/10.21037/atm-22-1494/rc>

Data Sharing Statement: Available at <https://atm.amegroups.com/article/view/10.21037/atm-22-1494/dss>

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at <https://atm.amegroups.com/article/view/10.21037/atm-22-1494/coif>). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). All participants signed written informed consent before starting the study. This research was approved by the First Affiliated Hospital of Guangxi Medical University Ethics Committee [Approval No. 2022-KY-E-(095)].

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: <https://creativecommons.org/licenses/by-nc-nd/4.0/>.

References

1. Parisi D, Chivasso C, Perret J, et al. Current State of Knowledge on Primary Sjögren's Syndrome, an Autoimmune Exocrinopathy. *J Clin Med* 2020;9:2299.
2. Ramos-Casals M, Brito-Zerón P, Kostov B, et al. Google-driven search for big data in autoimmune geoepidemiology: analysis of 394,827 patients with systemic autoimmune diseases. *Autoimmun Rev* 2015;14:670-9.
3. Luo J, Liao X, Zhang L, et al. Transcriptome Sequencing Reveals Potential Roles of ICOS in Primary Sjögren's Syndrome. *Front Cell Dev Biol* 2020;8:592490.
4. Fox RI. Sjögren's syndrome. *Lancet* 2005;366:321-31.
5. Fox RI, Fox CM, Gottenberg JE, et al. Treatment of Sjögren's syndrome: current therapy and future directions. *Rheumatology (Oxford)* 2021;60:2066-74.
6. Stefanski AL, Tomiak C, Pleyer U, et al. The Diagnosis and Treatment of Sjögren's Syndrome. *Dtsch Arztebl Int* 2017;114:354-61.
7. Soret P, Le Dantec C, Desvaux E, et al. A new molecular classification to drive precision treatment strategies in primary Sjögren's syndrome. *Nat Commun* 2021;12:3523.
8. Pontarini E, Lucchesi D, Bombardieri M. Current views on the pathogenesis of Sjögren's syndrome. *Curr Opin Rheumatol* 2018;30:215-21.
9. Zhang L, Xu P, Wang X, et al. Identification of differentially expressed genes in primary Sjögren's syndrome. *J Cell Biochem* 2019;120:17368-77.
10. Aqrabi LA, Galtung HK, Vestad B, et al. Identification of potential saliva and tear biomarkers in primary Sjögren's syndrome, utilising the extraction of extracellular vesicles and proteomics analysis. *Arthritis Res Ther* 2017;19:14.
11. Tandon M, Perez P, Burbelo PD, et al. Laser microdissection coupled with RNA-seq reveal cell-type and disease-specific markers in the salivary gland of Sjögren's syndrome patients. *Clin Exp Rheumatol* 2017;35:777-85.
12. Li XY, Wu ZB, Ding J, et al. Role of the frequency of blood CD4(+) CXCR5(+) CCR6(+) T cells in autoimmunity in patients with Sjögren's syndrome. *Biochem Biophys Res Commun* 2012;422:238-44.
13. Peng Y, Luo X, Chen Y, et al. LncRNA and mRNA expression profile of peripheral blood mononuclear cells in primary Sjögren's syndrome patients. *Sci Rep* 2020;10:19629.
14. Zhu N, Hou J, Wu Y, et al. Identification of key genes in rheumatoid arthritis and osteoarthritis based on bioinformatics analysis. *Medicine (Baltimore)* 2018;97:e10997.
15. García-Carrasco M, Pizcueta P, Cervera R, et al. Circulating concentrations of soluble L-selectin (CD62L) in patients with primary Sjögren's syndrome. *Ann Rheum Dis* 2000;59:297-9.
16. Mariette X, Criswell LA. Primary Sjögren's Syndrome. *N Engl J Med* 2018;378:931-9.
17. Cai W, Li H, Zhang Y, et al. Identification of key biomarkers and immune infiltration in the synovial tissue of osteoarthritis by bioinformatics analysis. *PeerJ* 2020;8:e8390.
18. Li R, Liu X, Zhou XJ, et al. Identification of a Prognostic Model Based on Immune-Related Genes of Lung Squamous Cell Carcinoma. *Front Oncol* 2020;10:1588.

19. Liu J, Sun GL, Pan SL, et al. Identification of hub genes in colon cancer via bioinformatics analysis. *J Int Med Res* 2020;48:300060520953234.
20. Xiong Y, Mi BB, Liu MF, et al. Bioinformatics Analysis and Identification of Genes and Molecular Pathways Involved in Synovial Inflammation in Rheumatoid Arthritis. *Med Sci Monit* 2019;25:2246-56.
21. Li S, Jiang Z, Chao X, et al. Identification of key immune-related genes and immune infiltration in atrial fibrillation with valvular heart disease based on bioinformatics analysis. *J Thorac Dis* 2021;13:1785-98.
22. Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* 2009;4:44-57.
23. Huang da W, Sherman BT, Lempicki RA. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res* 2009;37:1-13.
24. Deng YJ, Ren EH, Yuan WH, et al. GRB10 and E2F3 as Diagnostic Markers of Osteoarthritis and Their Correlation with Immune Infiltration. *Diagnostics (Basel)* 2020;10:171.
25. Chang Z, Huang R, Fu W, et al. The Construction and Analysis of ceRNA Network and Patterns of Immune Infiltration in Colon Adenocarcinoma Metastasis. *Front Cell Dev Biol* 2020;8:688.
26. Shiboski CH, Shiboski SC, Seror R, et al. 2016 American College of Rheumatology/European League Against Rheumatism Classification Criteria for Primary Sjögren's Syndrome: A Consensus and Data-Driven Methodology Involving Three International Patient Cohorts. *Arthritis Rheumatol* 2017;69:35-45.
27. Shiboski SC, Shiboski CH, Criswell L, et al. American College of Rheumatology classification criteria for Sjögren's syndrome: a data-driven, expert consensus approach in the Sjögren's International Collaborative Clinical Alliance cohort. *Arthritis Care Res (Hoboken)* 2012;64:475-87.
28. Yi XH, Zhang B, Fu YR, et al. STAT1 and its related molecules as potential biomarkers in Mycobacterium tuberculosis infection. *J Cell Mol Med* 2020;24:2866-78.
29. Ren C, Li M, Du W, et al. Comprehensive Bioinformatics Analysis Reveals Hub Genes and Inflammation State of Rheumatoid Arthritis. *Biomed Res Int* 2020;2020:6943103.
30. Conti P, Stellin L, Caraffa A, et al. Advances in Mast Cell Activation by IL-1 and IL-33 in Sjögren's Syndrome: Promising Inhibitory Effect of IL-37. *Int J Mol Sci* 2020;21:4297.
31. Zhao X, Gan Y, Jin Y, et al. Interleukin 17E associates with haematologic involvement and autoantibody production in primary Sjögren's syndrome. *Clin Exp Rheumatol* 2021;39:378-84.
32. Inamo J, Suzuki K, Takeshita M, et al. Identification of novel genes associated with dysregulation of B cells in patients with primary Sjögren's syndrome. *Arthritis Res Ther* 2020;22:153.
33. Li N, Li L, Wu M, et al. Integrated Bioinformatics and Validation Reveal Potential Biomarkers Associated With Progression of Primary Sjögren's Syndrome. *Front Immunol* 2021;12:697157.
34. Smatti MK, Cyprian FS, Nasrallah GK, et al. Viruses and Autoimmunity: A Review on the Potential Interaction and Molecular Mechanisms. *Viruses* 2019;11:762.
35. Triantafyllopoulou A, Moutsopoulos H. Persistent viral infection in primary Sjögren's syndrome: review and perspectives. *Clin Rev Allergy Immunol* 2007;32:210-4.
36. Bjork A, Thorlacius GE, Mofors J, et al. Viral antigens elicit augmented immune responses in primary Sjögren's syndrome. *Rheumatology (Oxford)* 2020;59:1651-61.
37. Brauner S, Folkersen L, Kvarnstrom M, et al. H1N1 vaccination in Sjögren's syndrome triggers polyclonal B cell activation and promotes autoantibody production. *Ann Rheum Dis* 2017;76:1755-63.
38. Proell M, Riedl SJ, Fritz JH, et al. The Nod-like receptor (NLR) family: a tale of similarities and differences. *PLoS One* 2008;3:e2119.
39. Shimizu T, Nakamura H, Takatani A, et al. Activation of Toll-like receptor 7 signaling in labial salivary glands of primary Sjögren's syndrome patients. *Clin Exp Immunol* 2019;196:39-51.
40. Zhang S, Qu J, Wang L, et al. Activation of Toll-Like Receptor 7 Signaling Pathway in Primary Sjögren's Syndrome-Associated Thrombocytopenia. *Front Immunol* 2021;12:637659.
41. Liuqing W, Liping X, Hui S, et al. Elevated IL-37, IL-18 and IL-18BP serum concentrations in patients with primary Sjögren's syndrome. *J Investig Med* 2017;65:717-21.
42. Blokland SLM, Flessa CM, van Roon JAG, et al. Emerging roles for chemokines and cytokines as orchestrators of immunopathology in Sjögren's syndrome. *Rheumatology (Oxford)* 2019.
43. Olsson P, Skogstrand K, Nilsson A, et al. Smoking, disease characteristics and serum cytokine levels in patients with primary Sjögren's syndrome. *Rheumatol Int* 2018;38:1503-10.

44. Musumeci F, Greco C, Giacchello I, et al. An Update on JAK Inhibitors. *Curr Med Chem* 2019;26:1806-32.
 45. Simon LS, Taylor PC, Choy EH, et al. The Jak/STAT pathway: A focus on pain in rheumatoid arthritis. *Semin Arthritis Rheum* 2021;51:278-84.
 46. Aota K, Yamanoi T, Kani K, et al. Inhibition of JAK-STAT Signaling by Baricitinib Reduces Interferon-gamma-Induced CXCL10 Production in Human Salivary Gland Ductal Cells. *Inflammation* 2021;44:206-16.
 47. Bodewes ILA, Huijser E, van Helden-Meeuwse CG, et al. TBK1: A key regulator and potential treatment target for interferon positive Sjogren's syndrome, systemic lupus erythematosus and systemic sclerosis. *J Autoimmun* 2018;91:97-102.
 48. Imgenberg-Kreuz J, Sandling JK, Bjork A, et al. Transcription profiling of peripheral B cells in antibody-positive primary Sjogren's syndrome reveals upregulated expression of CX3CR1 and a type I and type II interferon signature. *Scand J Immunol* 2018;87:e12662.
 49. Brokstad KA, Fredriksen M, Zhou F, et al. T follicular-like helper cells in the peripheral blood of patients with primary Sjogren's syndrome. *Scand J Immunol* 2018:e12679.
 50. Verstappen GM, Meiners PM, Corneth OBJ, et al. Attenuation of Follicular Helper T Cell-Dependent B Cell Hyperactivity by Abatacept Treatment in Primary Sjogren's Syndrome. *Arthritis Rheumatol* 2017;69:1850-61.
 51. Nocturne G, Mariette X. B cells in the pathogenesis of primary Sjogren syndrome. *Nat Rev Rheumatol* 2018;14:133-45.
 52. Gandolfo S, De Vita S. Double anti-B cell and anti-BAFF targeting for the treatment of primary Sjogren's syndrome. *Clin Exp Rheumatol* 2019;37 Suppl 118:199-208.
 53. Dorner T, Posch MG, Li Y, et al. Treatment of primary Sjogren's syndrome with ianalumab (VAY736) targeting B cells by BAFF receptor blockade coupled with enhanced, antibody-dependent cellular cytotoxicity. *Ann Rheum Dis* 2019;78:641-7.
 54. Carrillo-Ballesteros FJ, Palafox-Sanchez CA, Franco-Topete RA, et al. Expression of BAFF and BAFF receptors in primary Sjogren's syndrome patients with ectopic germinal center-like structures. *Clin Exp Med* 2020;20:615-26.
 55. Jiang W, Zhang L, Zhao Y, et al. The efficacy and mechanism for action of igitumab in primary Sjogren's syndrome patients. *Int Ophthalmol* 2020;40:3059-65.
 56. Bodewes ILA, Björk A, Versnel MA, et al. Innate immunity and interferons in the pathogenesis of Sjogren's syndrome. *Rheumatology (Oxford)* 2019. [Epub ahead of print].
 57. Mielle J, Tison A, Cornec D, et al. B cells in Sjogren's syndrome: from pathophysiology to therapeutic target. *Rheumatology (Oxford)* 2019. [Epub ahead of print].
 58. Dong H, Liu F, Ma F, et al. Montelukast inhibits inflammatory response in rheumatoid arthritis fibroblast-like synoviocytes. *Int Immunopharmacol* 2018;61:215-21.
 59. Han B, Zhang YY, Ye ZQ, et al. Montelukast alleviates inflammation in experimental autoimmune encephalomyelitis by altering Th17 differentiation in a mouse model. *Immunology* 2021;163:185-200.
- (English Language Editor: D. Fitzgerald)

Cite this article as: Zeng Q, Wen J, Zheng L, Zeng W, Chen S, Zhao C. Identification of immune-related diagnostic markers in primary Sjogren's syndrome based on bioinformatics analysis. *Ann Transl Med* 2022;10(8):487. doi: 10.21037/atm-22-1494