

## Research Article

# IRF9 and XAF1 as Diagnostic Markers of Primary Sjogren Syndrome

Lu Xiao , Zhou Yang, and Shudian Lin 

Department of Rheumatology, Hainan General Hospital (Hainan Affiliated Hospital of Hainan Medical University), Hainan 570311, China

Correspondence should be addressed to Lu Xiao; xiaolu\_230@163.com

Received 27 March 2022; Revised 11 August 2022; Accepted 18 August 2022; Published 12 September 2022

Academic Editor: Huiling Chen

Copyright © 2022 Lu Xiao et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

**Objective.** Primary Sjogren syndrome (pSS) is characterized by lymphocytic infiltration of the salivary and lacrimal glands. It is a chronic systemic autoimmune disease. Genetic contributions and disturbed biological systems are the two major causes of pSS, but its etiology is unclear. This study is aimed at identifying potential pSS diagnostic markers and mechanisms at the transcriptome level. **Methods.** Whole blood datasets of patients with pSS were downloaded from the Gene Expression Omnibus (GEO) database. Differentially expressed genes (DEGs) were identified using the online tool, GEO2R. R software was used to perform enrichment analyses to understand the functions and enriched pathways of the DEGs. A protein-protein interaction network was constructed to identify hub genes and significant gene clusters. The least absolute shrinkage and selection operator logistic regression was used to screen pSS diagnostic markers. The expression level and diagnostic performance of the identified genes were tested using another GEO dataset. **Results.** A total of 221 DEGs were screened from the whole blood samples of 161 patients with pSS and 59 healthy controls. Functional enrichment analysis demonstrated that DEGs were mostly enriched in defense response to virus, response to virus, and type I interferon signaling pathway. Cytoscape identified 10 hub genes and two gene clusters. IRF9 (AUC = 0.799) and XAF1 (AUC = 0.792) were identified as pSS diagnostic markers. The expression levels of the two identified genes were validated by GSE51092. **Conclusion.** IRF9 and XAF1 were identified as diagnostic markers. The potential underlying molecular mechanism of pSS was explored.

## 1. Introduction

Primary Sjogren syndrome (pSS) is characterized by lymphocytic infiltration of the salivary and lacrimal glands. It is a chronic systemic autoimmune disease that is overwhelmingly dominated by dry mouth (xerostomia), dry eyes (keratoconjunctivitis sicca), and autoantibody production. It may be accompanied by multiorgan systemic manifestations, such as pulmonary fibrosis and B cell lymphoma [1]. The prevalence of pSS in the general population is 0.33%–0.77%. Its occurrence rate is 10 times higher in females than in males [2]. The mortality rate is high among patients with cryoglobulinemic vasculitis, parotid enlargement, and lymphoma [3]. Genetic contributions and disturbed biological systems are the two major causes of pSS, but the etiology of pSS remains unclear.

Differentially expressed genes (DEGs) among different groups of people can be discovered with the rapid progress of

microarray techniques. In the past few years, several bioinformatic studies have revealed potential biomarkers in patients with pSS, which helped better elucidate the biological mechanisms of the disease. Chen et al. demonstrated that the expression levels of STAT1, MNDA, IL10RA, and CCR1 in the minor labial gland biopsy of pSS can serve as potential biological indicators for the disease [4]. Another study reported that PTPRC, CD86, and LCP2 are hub genes in the gland tissues of patients with pSS [5]. Li et al. identified seven key genes, namely, MS4A1, CD19, TCL1A, CCL19, CXCL9, CD3G, and CD3D, in gland tissues that have potential values for evaluating pSS severity [6]. Most past studies focused on gland tissues, but few studies focused on whole blood. In 2019, Yao et al. applied weighted gene coexpression network analysis to identify the potential pathways and hub genes that may be involved in pSS development [7]. Moreover, gene coexpression modules and hub genes in peripheral blood and parotid tissue associated

with B cell levels in pSS were identified by Lei and Zhang [8]. The substantial pathogenesis-related processes of pSS have been studied, but researchers still need to elucidate the exact pathogenesis and the key pathogenic factor of the disease. Useful algorithms and machine learning methods help with the easy discovery of diagnostic markers to better understand disease pathogenesis [9]. Wang et al. established the value of information feedback models and proved that hybrid adaptive differential evolution algorithm is a good state-of-the-art algorithm [10]. Therefore, the potential key genes and pathway networks related to pSS development can be explored through the combination of microarray and bioinformatic algorithm. Whole blood sample is easier to obtain than gland tissues; hence, studying the hub genes that may have diagnostic and treatment value from whole blood is of great importance for better patient outcomes.

In this study, pSS-related sequencing data from a public database were downloaded to discover the potential diagnostic markers and biological functions of the disease. The microarray datasets for pSS and HC, namely, GSE84844 (GPL570) and GSE66795 (GPL10558), respectively, were downloaded from the Gene Expression Omnibus (GEO) database. Then, the data were integrated and reanalyzed. A total of 221 common DEGs were identified between pSS and healthy control (HC). DEGs were clustered with functional enrichment analysis and gene set enrichment analysis (GSEA). Furthermore, a protein-protein interaction (PPI) network was constructed using the online tool, Search Tool for the Retrieval of Interacting Genes/Proteins (STRING). Genes identified by the least absolute shrinkage and selection operator (LASSO) logistic regression and Cytohubba, a Cytoscape plugin to detect hub genes, were combined for further analysis. Subsequently, the two identified hub genes were validated using GEO dataset GSE51092 (GPL6884). The diagnostic accuracy of the identified hub genes for pSS was evaluated with the area under the receiver operating characteristic curve (AUC).

## 2. Materials and Methods

**2.1. Data Download and Preprocessing.** “Primary Sjogren syndrome” was used as the search keyword for the expression profiles of pSS in the GEO database [11]. Datasets containing the sequencing information of whole blood from patients with pSS and HCs were included. Finally, two datasets, namely, GSE84844 (GPL570) and GSE66795 (GPL10558), were selected as the test sets [12, 13]. Dataset GSE51092 (GPL6884), which includes samples from 190 patients with pSS and 32 HCs, was selected as the validation set [14]. Basic information, including platform, samples, source tissue, and attribution, is listed in Table 1. The overall flowchart is shown in Supplementary Figure 1.

**2.2. DEG Screening.** Online web-based tool GEO2R was applied to discover the DEGs between pSS and HC. Adjusted  $P < 0.01$  was regarded as statistically significant. The overlapping DEGs from the two datasets were detected by the online tool Draw Venn Diagram (<http://bioinformatics.psych.ugent.be/webtools/Venn/>).

**2.3. Functional Enrichment Analysis.** Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analyses of the identified DEGs were performed by R packages (clusterProfiler, ggplot2, and GOplot) [15]. clusterProfiler was applied to analyze the DEGs. ggplot2 and GOplot were used to visualize the results.

**2.4. GSEA.** GSEA was conducted for all genes from the two gene datasets (GSE84844 and GSE66795) using R studio. ggplot2 was used to visualize the results. The gene set arrangement was performed 1000 times per analysis. Gene sets were considered significantly enriched when false discovery rate (FDR)  $< 0.25$ , adjusted  $P < 0.05$ , and  $|\text{normalized enrichment score (NES)}| > 1$ .

**2.5. PPI Network Construction.** A PPI network was constructed by the online tool STRING [16]. The cutoff standard was set as a combined score of  $> 0.4$  [17]. Then, the results were visualized by Cytoscape software. Molecular Complex Detection (MCODE, V1.5.1), a Cytoscape plugin, was used to identify significant modules (MCODE score  $\geq 4$ ) [18]. Moreover, the hub genes were chosen by Cytohubba based on a high number of associations with other genes in the PPI network [19].

**2.6. Screening and Verification of Diagnostic Markers.** LASSO logistic regression was used to perform feature selection to identify diagnostic markers for pSS. The expression matrix of the GSE66795 dataset was subjected to LASSO logistic regression. The LASSO regression analyses were conducted using the glmnet R package [20]. The genes from LASSO regression analysis and the hub genes identified by Cytohubba were combined for further analysis.

**2.7. Statistical Analysis.** RStudio was used to perform the statistical analysis. GSE51092 was used to validate the expression levels of the identified diagnostic markers. Wilcoxon rank sum test was used as the statistical method when the data were not normally distributed. Independent sample  $t$ -test was used when the data were normally distributed. Moreover, AUC, which was used to represent the diagnostic performance of the identified genes, was determined by RStudio. GLM function was used to build the logistic regression model. PROC package was used to analyze the receiver operating characteristic (ROC) curves. ggplot2 package was used to visualize the results.  $P < 0.05$  was considered significant.

## 3. Results

**3.1. Data Preprocessing and DEG Screening.** The overall workflow of the study is shown in Supplementary Figure 1. Two public pSS datasets, namely, GSE66795 and GSE84844, were downloaded to investigate the common DEGs involved in pSS. The principal component analysis (PCA) clusters for GSE66795 and GSE84844 are presented in Figures 1(a) and 1(b), respectively. PCA demonstrated that almost all variations were represented by HC and pSS in GSE66795 (16.1% and 10.1%, respectively) and GSE84844 (21.9% and 8.1%, respectively). In addition, Figures 1(c) and 1(d) show the heat maps and Figures 1(e) and 1(f) present the uniform manifold approximation and projection (UMAP) of the two

TABLE 1: Information for selected microarray datasets.

GEO accession	Platform	Samples		Source tissue	Attribution
		pSS	HC		
GSE84844	GPL570	30	30	Whole blood	Test set
GSE66795	GPL10558	131	29	Whole blood	Test set
GSE51092	GPL6884	190	32	Whole blood	Validation set

pSS: primary Sjogren syndrome; HC: healthy control.

gene datasets. Two clusters could be identified by UMAP analysis in GSE84844. As shown in Figure 1(g), 486 and 5466 DEGs were found in GSE66795 and GSE84844, respectively. After these, DEGs were integrated, and 221 common DEGs were discovered (Figure 1(g)).

**3.2. Biological Functions of Common DEGs.** GO, KEGG, and GSEA were used to analyze the 221 common DEGs (Figure 2, Tables 2 and 3). Based on GO enrichment, the biological process acted on defense response to virus, response to virus, and type I interferon signaling pathway. Regarding molecular functions, these proteins played roles in double-stranded RNA binding, NAD<sup>+</sup> ADP-ribosyltransferase activity, and single-stranded RNA binding. According to the KEGG pathway analysis, these proteins were involved in influenza A and NOD-like receptor signaling pathway (Figures 2(a)–2(d)). GSEA is designed to analyze the association between gene sets and biological signals in one dataset. It was conducted to identify the possible pathological processes of pSS. The top three significantly enriched gene sets were cytokine signaling in immune system, interferon signaling, and interferon alpha beta signaling in GSE66795 (Figure 2(e) and Table 3). Considering that the adjusted *P* value was greater than 0.05 in GSE84844, the enriched gene sets were not significant in this dataset (Table 3, adjusted *P* = 0.055).

**3.3. PPI Network Analysis, MCODE Cluster Modules, and Hub Gene Identification.** The PPI network for the 221 DEGs was constructed after the common DEGs were imported into STRING (Figure 3(a)). The top 10 hub genes were identified among the DEGs based on the information in the STRING database and Cytohubba. The top 10 hub genes were detected, including E3 ISG15-protein ligase HERC5, interferon alpha-inducible protein 6, interferon-induced transmembrane protein 1, probable E3 ubiquitin-protein ligase HERC6, interferon-induced protein with tetratricopeptide repeats 5, interferon-induced transmembrane protein 3, interferon alpha-inducible protein 27, interferon regulatory factor 9 (IRF9), interferon-stimulated gene 20 kDa protein, and XIAP-associated factor 1 (XAF1, Figure 3(b)). All the 10 hub genes were upregulated. The significant modules were identified by MCODE with an MCODE score of  $\geq 4$  as the threshold. Two modules with MCODE scores  $\geq 4$  are displayed in Figures 3(c) and 3(d). Cluster 1 (MCODE score = 44.96) had 51 nodes and 1124 edges (Figure 3(c)). Cluster 2 (MCODE score = 4) had 4 nodes and 6 edges (Figure 3(d)).

**3.4. Screening and Verification of Diagnostic Markers.** The LASSO logistic regression algorithm identified 19 genes from the DEGs as diagnostic markers for pSS (Figures 4(a)

and 4(b)). The gene markers obtained by LASSO algorithms and the hub genes overlapped. Finally, two diagnosis-related genes were obtained (Figure 4(c)). GSE51092 was used to validate the diagnostic efficacy of IRF9 and XAF1. The AUCs of IRF9 and XAF1 were 0.799 and 0.792, respectively, which indicated that IRF9 and XAF1 had certain diagnostic values (Figure 4(d)). When IRF9 and XAF1 were fitted into one variable, the diagnostic efficiency reached a higher level in the validation set (AUC = 0.822, Figure 4(d)). The identified logistic regression model was  $-27.6726 + (2.0467 \times \text{IRF9}) + (0.6087 \times \text{XAF1})$ . In addition, the GSE51092 dataset was used to verify the expression of the two identified genes. The mRNA expression levels of IRF9 and XAF1 were significantly increased in the pSS samples compared with the HCs ( $P < 0.01$ , Figures 4(e) and 4(f)).

## 4. Discussion

Although pSS is among the intractable autoimmune diseases, the number of studies focusing on the identification of important genes and pathways associated with pSS is far less than that on systemic lupus erythematosus and rheumatoid arthritis. The molecular and cellular events that occur during the pathogenesis of pSS need to be characterized. As the treatment of patients with pSS is still a clinical challenge, identifying the susceptibility genes of pSS is essential in studying the cause of pSS and find potential treatments.

In the present study, datasets containing the sequencing information of whole blood from pSS and HC were downloaded from the GEO database, and 221 DEGs were identified. The biological functions of these common DEGs were investigated. Defense response to virus, response to virus, and type I interferon signaling pathway were considerably enriched in pSS samples. In addition, the pathways enriched by GSEA involve cytokine signaling in immune system, interferon signaling, and interferon alpha beta signaling. The interferon pathway is activated in the pathogenesis of pSS as shown in previous research [21]. The first indication of interferon activation in pSS dated back to the late 1970s. In 2019, one study reported that salivary gland epithelial cells show upregulated interferon signaling pathway according to RNA-sequencing analysis [21]. Beyond salivary gland tissues, the presence of interferon signature was also evaluated at systemic level. The increased expression of interferon signaling pathway-associated genes or proteins has been detected in whole blood [22, 23], which is consistent with our findings. Interferons are mediators of innate immune defense mechanisms. The activation of innate immune pathways is a central pathogenetic contributor to pSS. Therefore, interferon blockade seems to be an attractive therapeutic target for this disease.

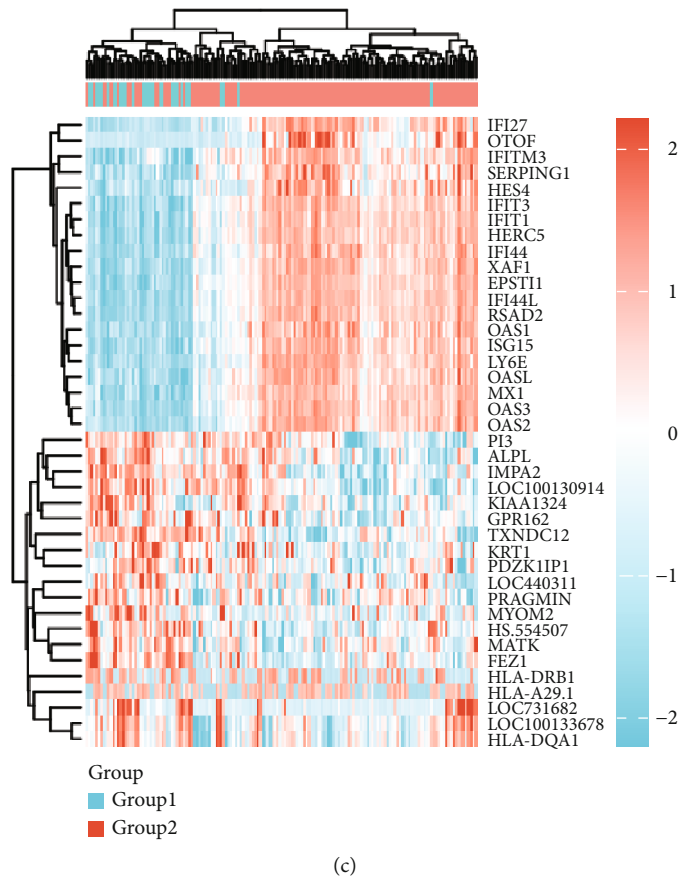
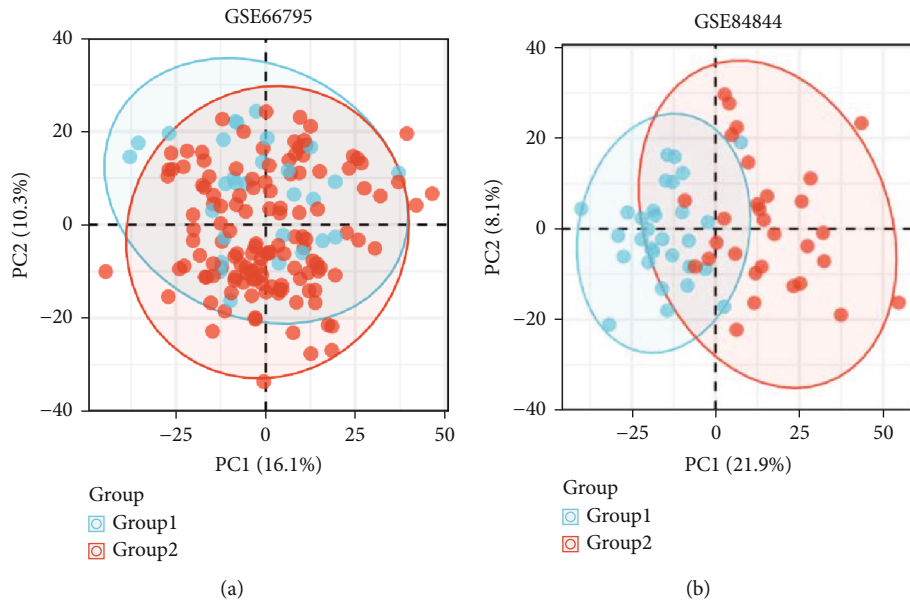


FIGURE 1: Continued.

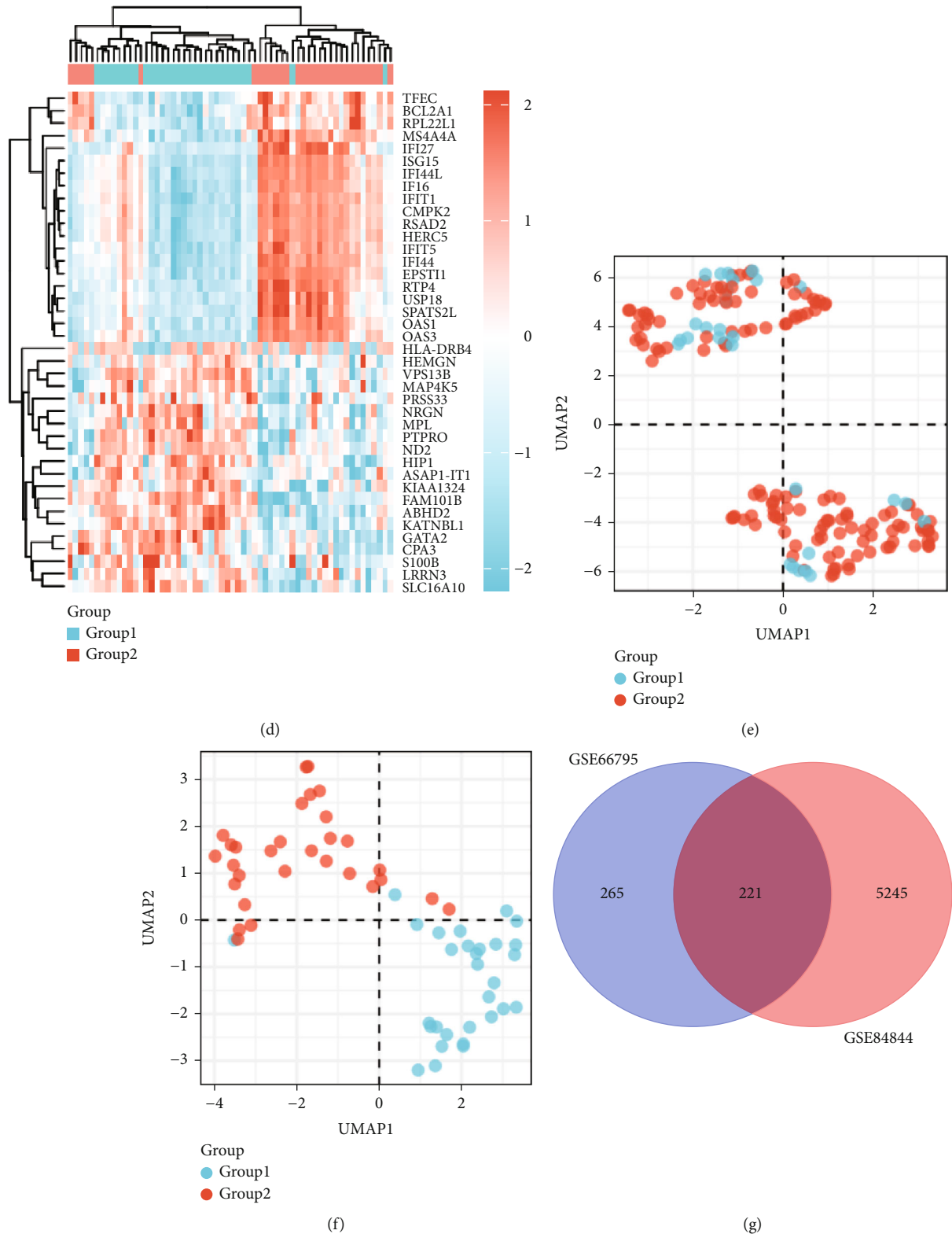


FIGURE 1: Identification of DEGs. (a) Principal component analysis (PCA) plot generated from DEGs in GSE66795. (b) PCA plot generated from DEGs in GSE84844. (c) Heat map of GSE66795. (d) Heat map of GSE84844. (e) Uniform manifold approximation and projection (UMAP) showing distinct clusters of DEGs in GSE66795. (f) UMAP showing distinct clusters of DEGs in GSE84844. (g) Venn diagram of common DEGs from the two datasets. Data points in red represent upregulated genes, and those in blue represent downregulated genes. Group 1 indicates HC and group 2 indicates pSS.



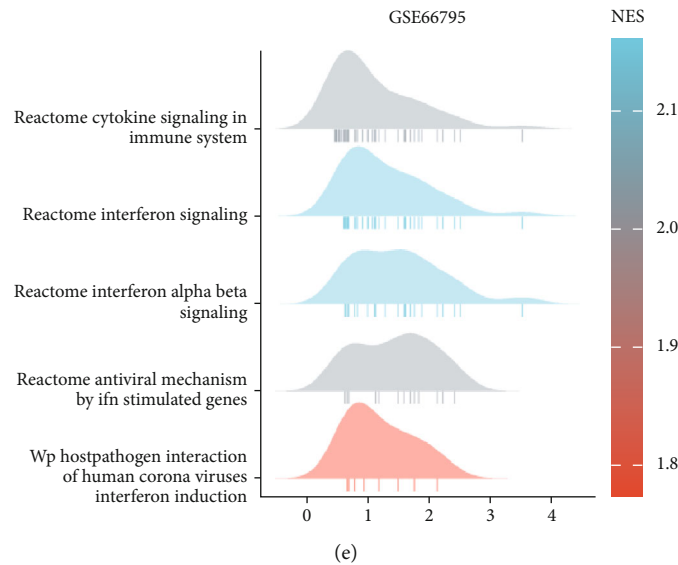


FIGURE 2: Functional enrichment of DEGs. Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analyses of DEGs. (a) Histogram, where the horizontal axis represents the number of DEGs under the GO or KEGG term. (b, c) Chordal graph, enrichment analysis combined logFC. On the basis of enrichment analysis, logFC of the molecule provided is used to calculate the corresponding Z score of each item and preliminarily determine whether the corresponding item is positively regulated (Z score is positive) or negatively regulated (Z score is negative). (d) Visual network of GO enrichment and KEGG pathway analysis of the overlapping DEGs. (e) Gene set enrichment analysis of GSE66795.

TABLE 2: GO and KEGG analyses of DEGs.

Ontology	ID	Description	Gene ratio	Adjusted $P$	$q$ value
BP	GO:0051607	Defense response to virus	46/199	$1.40e-41$	$1.29e-41$
BP	GO:0009615	Response to virus	49/199	$2.82e-39$	$2.59e-39$
BP	GO:0060337	Type I interferon signaling pathway	28/199	$2.78e-30$	$2.55e-30$
MF	GO:0003725	Double-stranded RNA binding	10/201	$5.37e-06$	$4.91e-06$
MF	GO:0003950	NAD <sup>+</sup> ADP-ribosyltransferase activity	4/201	0.015	0.014
MF	GO:0003727	Single-stranded RNA binding	7/201	0.015	0.014
KEGG	hsa05164	Influenza A	21/117	$5.07e-12$	$4.82e-12$
KEGG	hsa04621	NOD-like receptor signaling pathway	16/117	$5.37e-07$	$5.10e-07$

CC: cellular component group; MF: molecular function group; BP: biological process group; KEGG: Kyoto Encyclopedia of Genes and Genomes.

TABLE 3: Gene set enrichment analysis.

ID	NES	Adjusted $P$	FDR
GSE84844			
Reactome cytokine signaling in immune system	1.614	0.055	0.051
Reactome translation	1.74	0.055	0.051
Reactome interferon signaling	2.468	0.055	0.051
Reactome regulation of expression of SLITs and ROBOs	1.883	0.055	0.051
Reactome influenza infection	2.016	0.055	0.051
GSE66795			
Reactome cytokine signaling in immune system	2.085	0.01***	0.007
Reactome interferon signaling	2.185	0.01***	0.007
Reactome interferon alpha beta signaling	2.18	0.01***	0.007
Reactome antiviral mechanism by IFN-stimulated genes	2.077	0.01***	0.007
WP host pathogen interaction of human coronaviruses' interferon induction	1.777	0.025*	0.018

\*\*\* Adjusted  $P \leq 0.01$ ; \* adjusted  $P \leq 0.05$ .





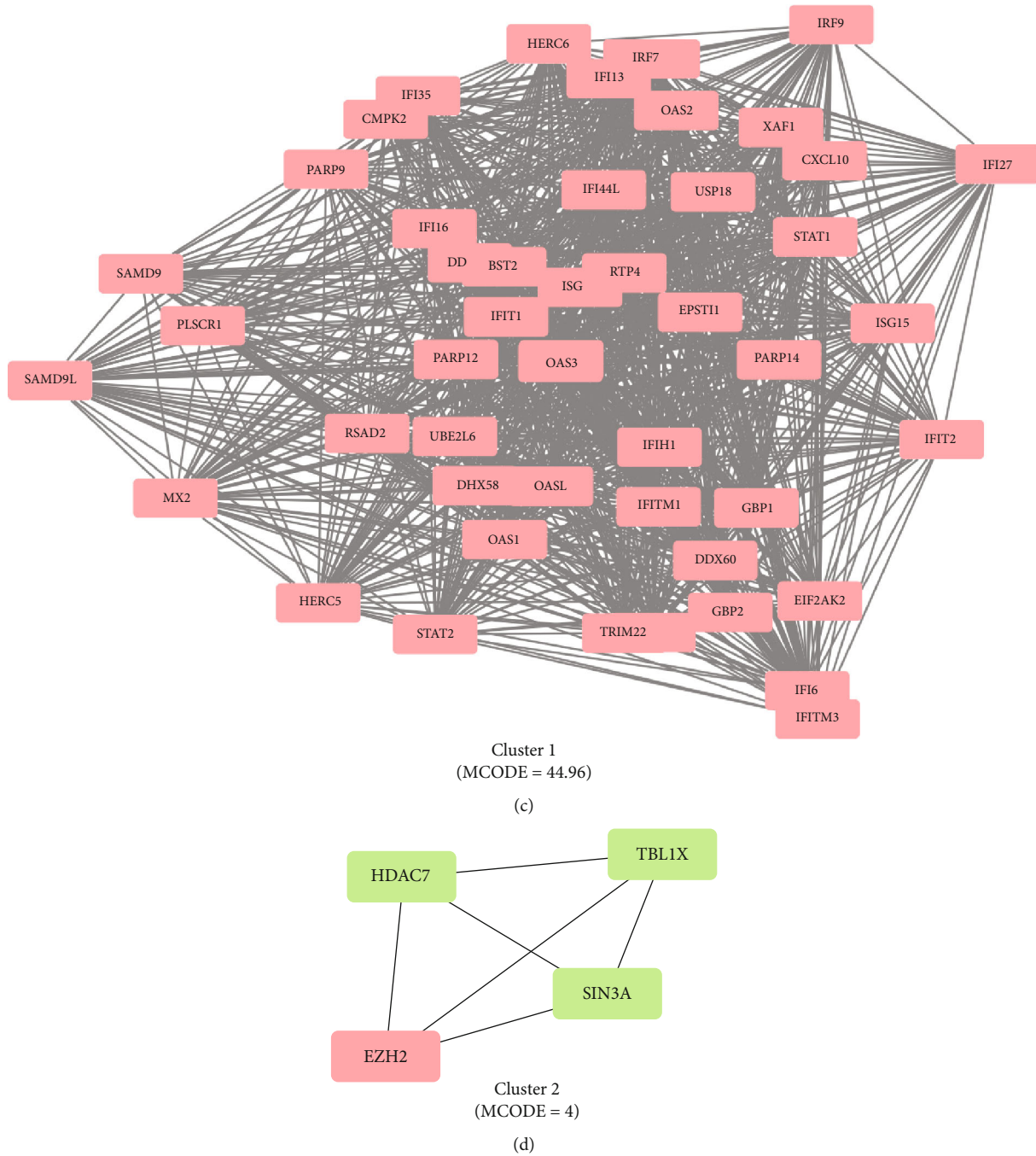


FIGURE 3: PPI network of DEGs and three cluster modules extracted by MCODE. (a) The interaction network between proteins coded by DEGs. (b) The interaction between the ten identified hub genes. Each node represents a protein, whereas each edge represents one protein-protein association. The purple rectangles represent the upregulated gene, and the yellow ones represent the downregulated gene. Two cluster modules extracted by MCODE. (c) Cluster 1 had higher cluster score (MCODE score = 44.96), followed by (d) cluster 2 (MCODE score = 4).

LASSO logistic regression was first proposed by Robert Tibshirani in 1996 [24]. It is a compressed estimation that determines the variable by finding  $\lambda$  when the classification error is the smallest. This algorithm is used to screen feature variables and build the best classification model. In this study, two diagnostic markers, namely, IRF9 and XAF1, were obtained by combining the genes identified by the LASSO

logistic regression and the hub genes identified by Cytohubba. A hub gene plays a critical role in biological processes and is often influenced by the regulation of other genes in related pathways. Therefore, these two identified genes may play important roles in the pathogenesis of pSS and have good diagnostic performance. The expression level and diagnostic performance of IRF9 and XAF1 were tested by another gene

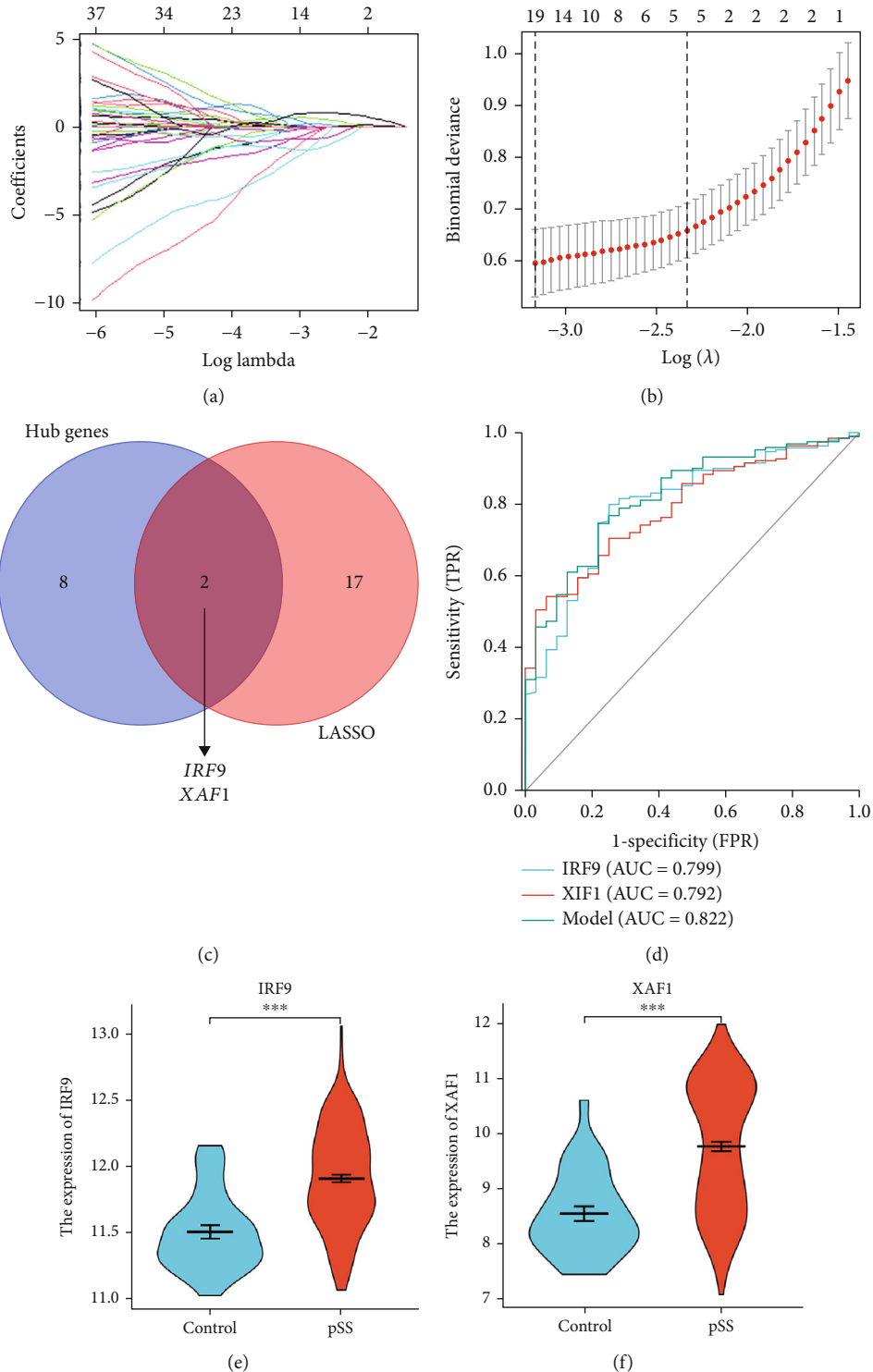


FIGURE 4: Screening and verification of diagnostic markers. (a) Least absolute shrinkage and selection operator (LASSO) logistic regression algorithm to screen diagnostic markers. Different colors represent different genes. Each curve corresponds to an independent variable in the full model prior to optimization. Curves indicate the path of each variable coefficient as  $\lambda$  varies. Lambda.min corresponds to  $\lambda$  which minimizes mean squared error in the model and was used for the selection of the seven predictor variables. (b) LASSO coefficient profiles of the 19 candidates in GSE66795. Plot of nonzero variable fit after cross-validation. Representation of the 10-fold cross-validation performed in LASSO that chooses the optimal  $\lambda$ . Lambda.min corresponds to  $\lambda$  which minimizes mean squared error and was used for variable selection. Lambda.1se corresponds to  $\lambda$  that is one standard error from Lambda.min. (c) Venn diagram shows the intersection of diagnostic markers obtained by LASSO and hub genes. (d) The ROC curve of the diagnostic efficacy tested by GSE51092. (e) The expression level of IRF9 between patients with pSS and HCs in GSE51092. (f) The expression level of XAF1 between patients with pSS and HCs in GSE51092. \*\*\*  $P < 0.001$ .

dataset, GSE51092. IRF9 and XAF1 showed certain diagnostic accuracy with AUCs of 0.799 and 0.792, respectively. Moreover, the AUC of the combination of the two genes reached 0.822, indicating a good diagnostic performance. The identified logistic regression model was  $-27.6726 + (2.0467 \times \text{IRF9}) + (0.6087 \times \text{XAF1})$ . This model would be of great value in future clinical practice to diagnose pSS. IRF9 is an important transcription factor involved in type I interferon production [25]. The elevated level of IRF9 in patients with pSS has been reported in the past [26]. Its binding with the signal transducer and activator of transcription 1 and 2 heterodimer leads to the formation of a complex, namely, interferon-stimulated gene factor (ISG) 3 [27]. The encoded ISG proteins have several immunomodulatory functions, including the induction of B cell activating factor, immunoglobulin switching, and increased antigen presentation, which all have vital roles in the pathogenesis of pSS [28, 29]. XAF1, an interferon type I inducible gene, encodes a zinc-finger proapoptotic protein. The role of XAF1 is discussed in different kinds of cancers. XAF1 works as a modifier of p53 function and cancer susceptibility [30]. Limited studies focused on the role of XAF1 in pSS. Two studies reported its elevation in patients with pSS [7, 31]. Our result showed that IRF9 and XAF1 had certain diagnostic abilities. Therefore, future precisely designed studies are necessary to verify these potential genes.

In the interpretation of our results, the following limitations require careful discussion. On the one hand, the clinical samples included in this study were from different datasets. The clinical activity, severity, or gender may be different among groups. Therefore, heterogeneity and confounding factors may have distorted the analysis. On the other hand, our study is only a bioinformatic study, and the identified hub genes were not confirmed through in vitro assays or in vivo models. Hence, further precisely designed studies are necessary to verify the two identified genes.

In conclusion, the biological analyses provided an overview of the differential gene expression between pSS and HC, which determined 221 DEGs. The functional analysis of these DEGs indicated that defense response to virus, response to virus, and type I interferon signaling pathway were considerably enriched in the whole blood of patients with pSS. In addition, two diagnostic markers, IRF9 and XAF1, were obtained and proven to have certain diagnostic accuracy. Our analysis revealed previously unknown transcriptional changes in pSS and demonstrated the role of microarray-based expression profiling in characterizing biomarkers in diseases. However, the exact diagnostic values of the identified models still need to be tested in future large-scale investigations. The results of our study may provide new treatment targets for pSS. Hence, this analysis may guide future experimental research and clinical transformation.

## Data Availability

The datasets generated and/or analyzed during the current study are available in the GEO repository. GEO is a free, publicly available repository database that stores a large number of gene functions and expression. The working links are as follows: GSE84844 (<https://www.ncbi.nlm.nih>

[.gov/geo/query/acc.cgi?acc=GSE84844](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE84844)), GSE66795 (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE66795>), and GSE51092 (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE51092>).

## Conflicts of Interest

The authors declare no competing interests.

## Authors' Contributions

Lu Xiao designed the study. Lu Xiao and Zhou Yang wrote the manuscript. Shudian Lin revised the manuscript. All authors read and approved the final manuscript.

## Acknowledgments

The authors thank all participants for their cooperation in this study and the staff for their dedication. This research was supported by the Natural Science Foundation of Hainan Province of China (820QN386) and the Hainan Province Clinical Medical Center.

## Supplementary Materials

Supplementary Figure 1: a flowchart of the study. (*Supplementary Materials*)

## References

- [1] I. Thorne and N. Sutcliffe, "Sjögren's syndrome," *British Journal of Hospital Medicine (London, England)*, vol. 78, no. 8, pp. 438–442, 2017.
- [2] T. Witte, "Sjögren's syndrome," *Zeitschrift für Rheumatologie*, vol. 78, no. 6, pp. 511–517, 2019.
- [3] M. Ramos-Casals, P. Brito-Zerón, A. Sisó-Almirall, and X. Bosch, "Primary Sjogren syndrome," *BMJ*, vol. 344, no. -jun14 1, article e3821, 2012.
- [4] L. Chen, D. Lu, K. Yu et al., "Bioinformatics analysis for identification of key genes in salivary gland and the potential of a combination of biomarkers for the diagnosis of Ss," *Journal of Inflammation Research*, vol. 14, pp. 4143–4153, 2021.
- [5] Y. Lin, X. Yao, M. Yan et al., "Integrated analysis of transcriptomics to identify hub genes in primary Sjögren's syndrome," *Oral Diseases*, 2021.
- [6] N. Li, L. Li, M. Wu et al., "Integrated bioinformatics and validation reveal potential biomarkers associated with progression of primary Sjögren's syndrome," *Frontiers in Immunology*, vol. 12, article 697157, 2021.
- [7] Q. Yao, Z. Song, B. Wang, Q. Qin, and J. A. Zhang, "Identifying key genes and functionally enriched pathways in Sjögren's syndrome by weighted gene co-expression network analysis," *Frontiers in Genetics*, vol. 10, p. 1142, 2019.
- [8] S. Lei and Y. Zhang, "Identification of the key genes and pathways involved in B cells in primary Sjögren's syndrome," *Bioengineered*, vol. 12, no. 1, pp. 2055–2073, 2021.
- [9] J. Zhang, Y. Jianhong, F. Suna, and X. Tian, "Adoption value of deep learning and serological indicators in the screening of atrophic gastritis based on artificial intelligence," *The Journal of Supercomputing*, vol. 77, no. 8, pp. 8674–8693, 2021.

- [10] G. G. Wang, D. Gao, and W. Pedrycz, "Solving multi-objective fuzzy job-shop scheduling problem by a hybrid adaptive differential evolution algorithm," *IEEE Transactions on Industrial Informatics*, pp. 1–1, 2022.
- [11] E. Clough and T. Barrett, "The gene expression omnibus database," *Methods in Molecular Biology*, vol. 1418, pp. 93–110, 2016.
- [12] S. Tasaki, K. Suzuki, A. Nishikawa et al., "Multiomic disease signatures converge to cytotoxic Cd8 T cells in primary Sjögren's syndrome," *Annals of the Rheumatic Diseases*, vol. 76, no. 8, pp. 1458–1466, 2017.
- [13] K. James, S. Al-Ali, J. Tarn et al., "A transcriptional signature of fatigue derived from patients with primary Sjögren's syndrome," *PLoS One*, vol. 10, no. 12, article e0143970, 2015.
- [14] C. J. Lessard, H. Li, I. Adrianto et al., "Variants at multiple loci implicated in both innate and adaptive immune responses are associated with Sjogren's syndrome," *Nature Genetics*, vol. 45, no. 11, pp. 1284–1292, 2013.
- [15] G. Yu, L. G. Wang, Y. Han, and Q. Y. He, "Clusterprofiler: an R package for comparing biological themes among gene clusters," *OmicS: a journal of integrative biology*, vol. 16, no. 5, pp. 284–287, 2012.
- [16] D. Szklarczyk, A. L. Gable, K. C. Nastou et al., "The String database in 2021: customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets," *Nucleic Acids Research*, vol. 49, no. D1, pp. D605–D612, 2021.
- [17] D. Szklarczyk, A. L. Gable, D. Lyon et al., "String V11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets," *Nucleic Acids Research*, vol. 47, no. D1, pp. D607–D613, 2019.
- [18] W. P. Bandettini, P. Kellman, C. Mancini et al., "Multicontrast delayed enhancement (Mcode) improves detection of subendocardial myocardial infarction by late gadolinium enhancement cardiovascular magnetic resonance: a clinical validation study," *Journal of Cardiovascular Magnetic Resonance*, vol. 14, no. 1, p. 83, 2012.
- [19] C. H. Chin, S. H. Chen, H. H. Wu, C. W. Ho, M. T. Ko, and C. Y. Lin, "Cytohubba: identifying hub objects and sub-networks from complex interactome," *BMC Systems Biology*, vol. 8, Supplement 4, p. S11, 2014.
- [20] J. Friedman, T. Hastie, and R. Tibshirani, "Regularization paths for generalized linear models via coordinate descent," *Journal of Statistical Software*, vol. 33, no. 1, pp. 1–22, 2010.
- [21] E. Rivière, J. Pascaud, N. Tchitchek et al., "Salivary gland epithelial cells from patients with Sjögren's syndrome induce B-lymphocyte survival and activation," *Annals of the Rheumatic Diseases*, vol. 79, no. 11, pp. 1468–1477, 2020.
- [22] R. Davies, I. Sarkar, D. Hammenfors et al., "Single cell based phosphorylation profiling identifies alterations in toll-like receptor 7 and 9 signaling in patients with primary Sjögren's syndrome," *Frontiers in Immunology*, vol. 10, p. 281, 2019.
- [23] A. Nezos, F. Gravani, A. Tassidou et al., "Type I and II interferon signatures in Sjogren's syndrome pathogenesis: contributions in distinct clinical phenotypes and Sjogren's related lymphomagenesis," *Journal of Autoimmunity*, vol. 63, pp. 47–58, 2015.
- [24] Y. P. Zhao, J. J. Wang, X. Y. Li, G. J. Peng, and Z. Yang, "Extended least squares support vector machine with applications to fault diagnosis of aircraft engine," *ISA Transactions*, vol. 97, pp. 189–201, 2020.
- [25] M. Sjöstrand, A. Ambrosi, S. Brauner et al., "Expression of the immune regulator tripartite-motif 21 is controlled by Ifn regulatory factors," *Journal of Immunology*, vol. 191, no. 7, pp. 3753–3763, 2013.
- [26] L. A. Aqrabi, J. L. Jensen, G. Øijordsbakken et al., "Signalling pathways identified in salivary glands from primary Sjögren's syndrome patients reveal enhanced adipose tissue development," *Autoimmunity*, vol. 51, no. 3, pp. 135–146, 2018.
- [27] L. B. Ivashkiv, "Ifn $\gamma$ : signalling, epigenetics and roles in immunity, metabolism, disease and cancer immunotherapy," *Nature Reviews. Immunology*, vol. 18, no. 9, pp. 545–558, 2018.
- [28] P. Z. Brohawn, K. Streicher, B. W. Higgs et al., "Type I interferon gene signature test-low and -high patients with systemic lupus erythematosus have distinct gene expression signatures," *Lupus*, vol. 28, no. 13, pp. 1524–1533, 2019.
- [29] M. Gursel and I. Gursel, "Development of CpG Odn based vaccine adjuvant formulations," *Methods in Molecular Biology*, vol. 1404, pp. 289–298, 2016.
- [30] E. M. Pinto, B. C. Figueiredo, W. Chen et al., "Xaf1 as a modifier of P53 function and cancer susceptibility," *Science Advances*, vol. 6, no. 26, p. eaba3231, 2020.
- [31] Z. Brkic, N. I. Maria, C. G. van Helden-Meeuwse et al., "Prevalence of interferon type I signature in Cd14 monocytes of patients with Sjogren's syndrome and association with disease activity and Baff gene expression," *Annals of the Rheumatic Diseases*, vol. 72, no. 5, pp. 728–735, 2013.