Uncovering potential new biomarkers and immune infiltration characteristics in primary Sjogren's syndrome by integrated bioinformatics analysis

Naidan Zhang

Peoples Hospital of Deyang City

Chaixia Ji

Peoples Hospital of Deyang City

Xiao Bao

Peoples Hospital of Deyang City

Xinyin Peng

Chengdu University of Chinese Medicine

Maoju Tang

North Sichuan Medical College

Chengliang Yuan (✉ 283072302@qq.com)

Peoples Hospital of Deyang City

Research Article

Keywords: primary Sjogren's syndrome, differentially expressed genes, immune infiltration, receiver operating characteristic curve, disease ontology annotation

Posted Date: March 2nd, 2022

DOI: <https://doi.org/10.21203/rs.3.rs-1372447/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background: Primary Sjogren's syndrome (pSS) is known as autoimmune disease of endocrine system. This study aimed to identify potential biomarkers for pSS using integrated bioinformatics analysis and explore the relationship between differentially expressed genes (DEGs) and immune infiltration.

Methods: Three pSS datasets (GSE7451, GSE23117 and GSE40611) from the Gene Expression Omnibus (GEO) database were integrated. All the datasets were processed by affy in R (version 4.0.3).

Results: A total of 16 immune cells and 13 immune functions infiltration scores were obtained. The top immune cell and immune function were "activated" dendritic cells (aDCs) and major histocompatibility complex (MHC) class I. Correlation analysis showed the top correlation among 16 immune cells were B cells and tumor infiltrating lymphocytes (TIL), check-point and T cell co-stimulation, respectively. In comparisons of immune score, aDCs (0.657 vs 0.594, $P < 0.001$), B cells (0.492 vs 0.434, $P = 0.004$), macrophages (0.631 vs 0.601, $P = 0.010$), inflammation-promoting (0.545 vs 0.478, $P < 0.001$), Type I interferon (IFN) Reponse (0.728 vs 0.625, $P < 0.001$) and so on were higher in pSS than control group. In correlation analysis, the up-regulation of IFIT1 gene was strongly correlated with Type I IFN Reponse with a correlation coefficient of 0.87. The ROC curve of 5 genes showed that the area under curve (AUC) was 0.891. In the verification model, the AUC was 0.881. In addition, DO analysis supported the association between DEGs and pSS.

Conclusions: In summary, pSS has a variety of DEGs in immune infiltration, which is worthy of the attention from clinicians.

Background

Sjogren's syndrome (SS) is also known as autoimmune exocrine disease. Xerostomia and xerophthalmia are the main clinical symptoms of SS. The disease is divided into primary and secondary. Primary Sjogren's syndrome (pSS) is a global disease [1]. Due to its insidious onset, it is difficult to diagnose in the early stage. As there is no radical cure for pSS, the clinical treatment is mainly to replace and relieve the symptoms and delay the progression [2]. At present, the pathogenesis of pSS is not clear. Genetic background and immune factors are considered to be the main basis of the occurrence and development of pSS [3, 4].

Differentially expressed genes (DEGs) play important roles in pathogenesis of pSS. Studies have shown that there are correlations between central genes in DEGs and immune cells [5]. DEGs are involved in immune and inflammatory responses in pSS, including interferon-mediated signal transduction pathway, antigen processing and presentation and carbohydrate metabolism pathway [6]. In addition, total exome sequencing (WES) of pSS families and sporadic patients indicated that there were common DEGs, and DEGs were associated with abnormalities in immune-related signaling pathways [7]. However, the study of salivary glands (SGs) of biopsy-positive and biopsy-negative pSS suggested that although DEGs could effectively separate biopsy-positive pSS from non-SS patients, it could not effectively distinguish biopsy-

negative pSS from non-SS patients [8]. The above studies suggested that personalized treatment methods should be selected for pSS according to the expression of SGs immune pathways.

In recent years, with the study of the pathogenesis driving immunopathology, the study of immune infiltration and heterogeneity of pSS is attracting more and more scholars' attention [9]. Studies that combine DEGs with immune infiltration will be more conducive to identifying potential biomarkers and their roles in pSS, so as to further explore potential therapeutic approaches [10].

As far as we know, there are few studies on the DEGs with immune cell and immune function infiltration score in SGs damage of pSS, especially using DEGs to build models and predict the occurrence of pSS. Therefore, we try to study from the perspective of DEGs and immune infiltration score as well as pSS prediction models. This is a new understanding of DEGs and immune infiltration in pSS, and may also help identify potential biomarkers for prediction and diagnosis of pSS.

Material And Methods

Microarray data

Study flowchart is shown in Fig. 1. Three pSS datasets (GSE7451, GSE23117, GSE40611) were downloaded from the NCBI Gene Expression Omnibus (GEO; <https://www.ncbi.nlm.nih.gov/geo/>). All the datasets were gene expression array and based on GPL570 (Affymetrix Human Genome U133 Plus 2.0 Array) platform. The dataset of GSE7451 included 10 saliva samples of pSS and 10 controls; GSE23117 contained 11 salivary gland samples of pSS and 4 controls; GSE40611 consisted of 49 parotid gland tissue samples including 17 pSS, 14 non-pSS sicca patients and 18 controls. All the datasets were pre-processed by affy in R (version 4.0.3), including data merge, normalization, and log2 transformation. When multiple probes corresponded to the same gene, the average value was used as the expression value. The surrogate variable analysis (SVA) package was used to eliminate batch effect, and LIMMA package was used to screen for differentially expressed genes (DEGs). A adjust P -value < 0.05 and $|\log_2$ Fold change FC| > 0.5 were set as cut-off points for DEGs screening.

Infiltration analysis of immune cell and immune function

In order to evaluate the changes of immune microenvironment in pSS, we predicted the composition and function of immune cells in the salivary and parotid tissue through gene expression matrix. The ssGSEA method was used to analyze the combined expression datasets and calculate immune cell infiltration and functional changes. The GSVA package was used to calculate the enrichment of 16 common immune cells and 13 immune functions in gene concentrations in each sample. The pheatmap package was used to draw the heatmap. The correlation of 16 infiltrating immune cells and 13 immune functions were analysed by using the corplot package, respectively. Correlation coefficient was used to evaluate the correlation between immune cells and immune functions, it was divided into extremely strong correlation, strong correlation, moderate correlation, weak correlation and very weak correlation. A plus sign indicated a positive correlation and a minus sign indicated a negative correlation.

Difference in immune score between controls and pSS

Statistical Package for the Social Sciences software (version 22.0) was used to analyze the difference of immune cell and immune function scores between pSS and healthy controls. Normally distributed data were presented as mean $\pm$ standard deviation (SD), non-normal variables were expressed as median (interquartile range, IQR). Comparison between groups was evaluated with independent-samples t test with normal distribution or wilcoxon rank sum test with non-normal distribution. The ggpubr package was used to plot boxplots of differences in immune cells and immune functions expression between pSS patients and controls. A P -value < 0.05 was considered statistically significant.

Correlation analysis of DEGs with immune cells and immune functions

According to the $|\log_2 \text{FC}|$, 5 up-regulated and 5 down-regulated genes were screened from DEGs. The psych package was used for data dimensionality reduction. The ggcorrplot package was used to plot the correlation heatmap between DEGs and immune cells/functions. Correlation coefficient was used to evaluate the correlation between immune cells and immune functions.

Differential genes construct ROC curves

In order to evaluate the correlation between DEGs with immune infiltration, we screened 5 highest associations with immune cells and immune functions. Logistic regression was used to calculate the correlation between the intensity of gene expression and risk of pSS. The receiver operating characteristic (ROC) curve was constructed to predict the risk of pSS. Area under curve (AUC) was obtained by ROC curve. The caret package was used for data segmentation, and 62 samples were selected as verification sets. AUC was used to evaluate the predictive efficiency of the model. A P -value < 0.05 was considered statistically significant.

Disease ontology (DO) and gene ontology (GO) annotation of DEGs

The DOSE package and the enrichplot package were used for DO and GO annotation of DEGs, respectively. The threshold for DO analyses was a P -value < 0.05 and count ≥ 5 . The barplot and dotplot packages were used to draw bar and bubble charts. The enriched GO groups were divided into biological process (BP), cellular component (CC), and molecular function (MF) ontologies. The threshold for GO analyses was a P -value < 0.05 and count ≥ 5 .

Results

Identification of DEGs

A total of 2337 genes were differentially expressed, of which 1221 genes were up-regulated and 1116 genes were down-regulated. The top 5 up-regulated genes were interferon induced protein with tetratricopeptide repeats 1 (*IFIT1*), interferon induced protein 44 like (*IFI44L*), cytidine/uridine

monophosphate kinase 2 (*CMPK2*), C-X-C motif chemokine ligand 9 (*CXCL9*), and Hieracium procerum (*SELL*) while the top 5 down-regulated genes were uncharacterized LOC100506446 (*LOC100506446*), SOX21 antisense divergent transcript 1 (*SOX21-AS1*), solute carrier family 27 member 4 (*SLC27A4*), ubiquitin specific peptidase 9 Y-linked (*USP9Y*), long intergenic non-protein coding RNA 1448 (*LINC01448*) (Table S1). The volcano plot of DEGs was shown in Fig. 2A.

Immune infiltration score and correlation analysis

A total of 16 immune cells and 13 immune functions infiltration scores of pSS were obtained by using the ssGSEA method. The top 5 immune cells were "activated" dendritic cells (aDCs), macrophages, regulatory T cells (Treg), CD8⁺ T cells and neutrophils while the top 5 immune functions were major histocompatibility complex (MHC) class I, human leukocyte antigen (HLA), type I interferon (IFN) reponse, parainflammation and cytolytic activity. The heatmap of immune cells and immune functions were shown in Fig. 2B.

Correlation analysis showed the top 3 correlation among 16 immune cells were B cells and tumor infiltrating lymphocytes (TIL), follicular helper T (Tfh) and TIL, Tfh and B cells. The top 3 correlation among 13 immune functions were check-point and T cell co-stimulation, check-point and CC chemokine receptors (CCR), check-point and T cell co-inhibition. The correlation plot of immune cells and immune functions were shown in Fig. 3A and Fig. 3B, respectively.

Comparison between immune score in pSS and controls

Comparisons of immune score between different groups were illustrated in Fig. 4. In immune cells, aDCs (0.657 vs 0.594, $P < 0.001$), B cells (0.492 vs 0.434, $P = 0.004$), macrophages (0.631 vs 0.601, $P = 0.010$), neutrophils (0.550 vs 0.489, $P = 0.001$), pDCs (0.478 vs 0.445, $P = 0.001$), T helper cells (0.399 vs 0.327, $P = 0.006$), Th2 cells (0.500 vs 0.470, $P = 0.004$), TIL (0.549 vs 0.492, $P < 0.001$) and Treg (0.621 vs 0.598, $P < 0.001$) were higher in pSS group than control group. In immune functions, antigen presenting cell (APC) co inhibition (0.657 vs 0.466, $P = 0.009$), cytolytic activity (0.614 vs 0.559, $P = 0.014$), HLA (0.793 vs 0.743, $P < 0.001$), inflammation-promoting (0.545 vs 0.478, $P < 0.001$), MHC class I (0.916 vs 0.873, $P < 0.001$), parainflammation (0.645 vs 0.601, $P < 0.001$), T cell co-inhibition (0.466 vs 0.439, $P = 0.032$) and type I IFN reponse (0.728 vs 0.625, $P < 0.001$) were higher in pSS group than control group. Other data showed no difference between the different groups ($P > 0.05$). The boxplots of differences between immune score were shown in Fig. 4A and Fig. 4B, respectively.

Correlation between DEGs and immune cell/function

The top 5 up-regulated genes we selected for correlation analysis were *IFIT1*, *IFI44L*, *CMPK2*, *CXCL9* and *SELL*. And the top 5 down-regulated genes were *LOC100506446*, *SOX21-AS1*, *SLC27A4*, *USP9Y* and *LINC01448*. The results showed that the DEGs were mainly related to the up-regulation of various immune cells and immune functions, and the up-regulation of *SELL* gene was related to the abnormal expression of many immune cells and immune functions. In correlation analysis, the up-regulation of

IFIT1 gene was strongly correlated with type I IFN reponse with a correlation coefficient of 0.87, followed by *CMPK2* and type I IFN reponse with a correlation coefficient of 0.86, as shown in Fig. 5A.

Differential genes construct ROC curves

According to the DEGs, the top 5 genes were selected to construct ROC curve, and the risk of pSS between DEGs were predicted. According to the median expression value of each gene, the genes were divided into low expression and high expression, and the column map was drawn. The results suggested that when the total score was greater than 160, the probability of pSS was close to 80%, and the high expressions of *CMPK2* and *IFI44L* genes were strongly correlated with the risk of pSS, as shown in Fig. 5B.

ROC curve indicated that the area under curve (AUC) of the combined 5 genes was the largest with the sensitivity was 80.0%, the specificity was 88.2% and AUC was 0.891, as shown in Fig. 5C. The next was *CMPK2* gene. When the score of *CMPK2* was 10.49, the sensitivity was 84.0%, the specificity was 82.4% and AUC was 0.875. The next was *IFI44L* gene. When the score of *IFI44L* was 9.06, the sensitivity was 68.0%, the specificity was 91.1% and AUC was 0.816, suggesting that the multi-gene combination ROC curve had the highest diagnostic efficiency for pSS occurrence, followed by *CMPK2* and *IFI44L*, as shown in Table 1 and Fig. 5C.

Table 1
ROC curve of DEGs.

Predictor variable	AUC	Sensitivity	Specificity	95%CI	Standard Error	Youden Index
CMPK2	0.875	0.840	0.824	0.795–0.956	0.181	0.664
IFI44L	0.816	0.680	0.912	0.727–0.905	0.270	0.592
CXCL9	0.756	0.740	0.706	0.650–0.862	0.301	0.446
IFIT1	0.811	0.740	0.794	0.714–0.908	0.189	0.534
SELL	0.780	0.640	0.824	0.679–0.881	0.304	0.464
Combined 5 genes	0.891	0.800	0.882	0.822–0.959	0.176	0.682
ROC, receiver operating characteristic; DEGs, differentially expressed genes; AUC, area under curve.						

In the verification cohort, the AUC was 0.881, indicating that the model had a good predictive ability for salivary gland injury in pSS patients, as shown in Table 2.

Table 2
ROC curve in verification sets.

Predictor variable	AUC	Sensitivity	Specificity	95%CI	Standard Error	Youden Index
CMPK2	0.868	0.872	0.826	0.765–0.972	0.208	0.698
IFI44L	0.821	0.718	0.913	0.717–0.924	0.295	0.631
CXCL9	0.814	0.744	0.826	0.705–0.923	0.366	0.570
IFIT1	0.821	0.795	0.870	0.705–0.936	0.214	0.664
SELL	0.798	0.641	0.870	0.684–0.912	0.355	0.511
Combined 5 genes	0.881	0.792	0.875	0.802–0.939	0.179	0.689
ROC, receiver operating characteristic; AUC, area under curve.						

DO and GO analyses of the DEGs

Through DO analysis of 2337 DEGs screened in this study, the top 10 diseases with the highest frequency of differential genes were explored, including “hepatitis”, “Human immunodeficiency virus infectious disease”, “Graves' disease”, “hyperthyroidism” and “autoimmune disease of endocrine system”, as shown in Fig. 6. The top 5 functional enrichment analysis in BP group were “T cell activation”, “response to virus”, “leukocyte cell-cell adhesion”, “regulation of T cell activation” and “leukocyte proliferation”. The top 5 functional enrichment analysis in CC group were “external side of plasma membrane”, “endocytic vesicle”, “cytoplasmic vesicle lumen”, “endocytic vesicle membrane” and “phagocytic vesicle”. MF analysis showed that the top 5 functions were “immune receptor activity”, “cytokine receptor activity”, “peptide antigen binding”, “MHC class II protein complex binding” and “T cell receptor binding”, as shown in Fig. 7.

Discussion

pSS is a chronic inflammatory autoimmune disease characterized by specific pathological lesions. Previous studies on DEGs mainly focused on the screening, the correlation analysis between DEGs and signaling pathways, and protein interaction networks [11–13]. However, few studies focused on DEGs and immune infiltration score, as well as the construction of prediction models. In this study, genes such as *IFIT1*, *IFI44L*, *CMPK2*, *CXCL9* and *SELL* were significantly upregulated whereas *LOC100506446*, *SOX21-AS1*, *SLC27A4*, *USP9Y* and *LINC01448* were significantly downregulated.

Studies showed that *IFIT1* and *IFI44L* were associated with the overactivation of the type I IFN in pSS, which was of great value in guiding treatment decisions, monitoring and patient stratification [14, 15]. As the ligand of chemokine receptor CXCR3, CXCL9 was induced by IFN γ and mainly acted on monocytes/macrophages and lymphocytes, mediating inflammatory responses [16, 17]. Studies suggested that IFN γ (MIG)/CXCL9 antagonist could improve SS salivary adenitis, suggesting that CXCL9 was a monitoring target for SS targeted therapy [18]. Another study showed that CXCL9 was not elevated in inflammatory modules in all pSS patients, suggesting that molecular phenotypes of SS should be characterized and clustered to solve the heterogeneity problem [19].

SELL encoded L-selectin, a cell surface adhesion molecule, and played an important role in migration of leukocytes to lymphoid organs. The *SELL* gene was associated with abnormal activation of a variety of immune cells, including activated CD4 memory T cells, plasma cells, γ - δ T cells, M2 macrophages and NK cells, as well as with the severity of inflammation [20, 21]. In this study, DEGs analysis found that *CXCL9* and *SELL* were significantly overexpressed in pSS, and a variety of immune cells were associated with immune function disorders. However, in the follow-up study of prediction models, it was found that *CXCL9* and *SELL* had less predictive value for the risk of pSS than *CMPK2* and *IFI44L*. In addition, a significantly down-regulated gene was found in pSS. In the correlation analysis between DEGs and immune infiltration, there was little correlation between down-regulated DEGs and immune infiltration. This study suggested that DEGs, especially up-regulated DEGs, were associated with multiple immune cell infiltrates, and monitoring the abnormal expression of DEGs was helpful to improve the level of immunoregulatory therapy in pSS.

In recent years, studies on immune infiltration and the pathogenesis of pSS attracted more and more attention. The analysis of infiltrating immune cells in pSS showed that not only DEGs but also interaction among various immune cell subsets played important roles in the occurrences and developments of pSS. These specific immune infiltration patterns of pSS may provide important clues for immunotherapy. On the basis of summarizing and analyzing the results of previous studies, this study adopted ssGSEA method to expand the analysis of immune cell infiltration to 16 common immune cells, and also added the analysis of 13 common immune functions.

We also explore the correlation between these significantly different immune cells and immune functions, which was an innovative attempt to analyses immune infiltration in pSS. In this study, we found that aDCs, macrophages, Treg, CD8 + T cells and neutrophils were closely correlated with pSS, while B cells, Tfh and TIL cells were highly correlated. The infiltration of macrophages and DC were increased in pSS. The increased IL-18 secretion by macrophages stimulated the activation of primary B cells, which played an important role in the expansion of invasive injury [22]. Tumour-associated macrophages-associated macrophages (TAMs) could control tumor growth and had anti-tumour immunity [23, 24]. However, studies on the function of macrophages associated with autoimmune diseases are not clear at present [25, 26]. This study suggested that pSS had abnormal activation of macrophages. It was still worthy of further investigation that how this activation participate in the development of pSS and expresses more inflammatory cytokines.

The research on CD8 + T cells suggested that the infiltrating CD8 + T cells were significantly more than CD4 + T cells in the murine SS model [27]. Reducing the infiltration of CD8 + T cells could effectively reduce the destruction of SGs [28]. CD4 + T cells were mainly composed of T helper cells (Th) and Treg. Tfh led to the formation of germinal center-like structures, which induced the activation of B cells in pSS [29]. The imbalance of Tfh/Treg in the blood and SGs of pSS suggested that immunotherapy to induce Treg while suppressing T-helper (Th) cells was one of the treatments for pSS [30, 31].

Immune functions in pSS showed abnormalities in MHC class I, HLA, type I IFN response, parainflammation and cytolytic activity. This observation has also been reported in the previous literature. Exocrine epithelial cells from pSS patients high levels of HLA-DR, inducing activation of infiltrating CD4 + T cells [32]. Recent studies using microarrays to analyze gene expression in small SGs also found that gene expression patterns in pSS involve multiple chronic inflammatory pathways, including chemokines, cytokines, MHC and IFN [33–35]. Unlike previous studies, this study also found that parainflammation was significantly up-regulated in the immune functions. It was suggested that the infiltration characteristics of parainflammation were similar to type I IFN response. Yet, parainflammation lacked all the characteristics of canonical inflammation, and remained autonomous to the primary disease [36, 37]. Furthermore, in tumors cell lines, parainflammation had a capacity to fulfill certain normal macrophage functions. It could mimic different immune subsets depending on the needs of the tumor [38]. In addition, the study on immune infiltrating cells also found that macrophages were significantly up-regulated, which was consistent with the study on immune functions. It was suggested that the study on parainflammation provided important clues for the treatment of pSS.

In order to further understand the association between DEGs and immune infiltration, we selected 10 genes that were significantly up-regulated and down-regulated, and analyzed the association between them and immune infiltration. The results showed that the DEGs were mainly related to the up-regulation immune cells and immune functions. The up-regulation of *SELL* gene was related to the infiltration of various immune cells and immune functions. However, the prediction value of *SELL* to pSS was lower than that of *CMPK2* and *IFI44L* in the analysis of lrogram and ROC curves. *CMPK2* and *IFI44L* combined diagnosis had the highest probability of pSS. *CMPK2* encoded the mitochondrial protein cytidyl/uridine monophosphate kinase 2, which was present in mitochondria and cytoplasm [39]. IFN- α induced *CMPK2* expression, which further regulated production of mtROS and activation of inflammasome and played an important role in proatherogenic events [39, 40]. Recent study showed that *CMPK2* function was closely related to NLRP3 inflammasome activation, and levels of IL-18, IL-1 β , and cleaved-caspase-1 are reduced after inhibition of *CMPK2* [40]. However, there were still few studies on *CMPK2* gene and protein expression in the pathogenesis of pSS. There were many researches on *IFI44L* in pSS. *IFI44L* encoded IFN-induced protein 44-like [41]. A number of studies showed that *IFI44L* was abnormally expressed in blood and SGs of pSS, which was related to the activation of IFN-I pathway [42]. Studies on prediction models of pSS suggested that *IFI44L* had important clinical value as an independent or combined prediction model, indicating that *IFI44L* was an important regulatory gene of IFN-I pathway [43, 44]. Type I IFNs decrease miRNA with anti-inflammatory roles in pSS, leading to upregulation of mucin 1 (MUC1)

and Toll-like receptor 4 (TLR4). It suggested that type I IFNs contributes to the inflammation in SGs from pSS.

Previous GO analysis of pSS showed that DEGs were mainly enriched in antigen presentation/processing, protein transport, immune/defense response and protein localization [45, 46]. In this study, we expanded the sample size for GO analysis based on previous research [47]. The results indicated that in addition to enrichment in the above pathways, DEGs were further significantly enriched in T cell activation, immune cell proliferation and cytoplasm endocytic vesicle, especially in T cell activation and cytoplasmic vesicle transport. It was suggested that neonates after exposure to anti-RO/La autoantibodies demonstrate innate immune activation, which was mainly related to IFN- γ produced by NK cells, CD4 + T cells and CD8 + T cells [48]. Additional study indicated that anti-centromere antibodies (ACAs) in the cytoplasm were detected in pSS and were associated with the accumulation of SGs centromere antibody secreting cells (ASCs) [49]. However, one study proposed that GO analysis was significantly different in salivary and lacrimal glands in time series [45]. It was suggested that we should try to select samples from the same tissue or organ for GO analysis, so as to avoid differences in results caused by different samples. We also used DO analysis to conduct enrichment analysis on DEGs involved in different diseases. The results indicated that DEGs were mainly concentrated in hepatitis, collagen disease and rheumatic disease. As one of the common and high incidence diseases for rheumatic disease, this study suggested that DEGs was associated with pSS, which was consistent with the results of DO analysis.

Our study had certain limitations. We focused on several important enrichment results and genes related to immune infiltration, and may ignore the potential genes related to pSS. In addition, further studies on immune infiltration need to be conducted to expand the sample size and in vitro and in vivo studies.

Conclusion

We conjectured that DEGs might be up-regulated by affecting immune cells and immune functions and further regulate SGs injury in pSS. *SELL* gene was significantly associated with abnormal expression of immune infiltration, and may serve as a important gene for inflammatory regulation of pSS. *CMPK2* and *IFI44L* genes have high predictive value for the risk of pSS. In addition, aDCs, macrophages, Treg, CD8⁺ T cells, MHC class I, HLA, type I IFN reponse and parainflammation may be involved in the pathogenesis of pSS.

Declarations

Ethical Approval and Consent to participate

Not applicable.

Supplementary Information

Additional Supporting Information may be found online in the supporting information Table S1 for this article. **Additional file 1: Table S1.** The top 5 up-regulated and down-regulated genes.

Acknowledgments

We wanted to express our gratitude to the providers of the datasets GSE7451, GSE23117 and GSE40611 for sharing the datasets.

Authors' contributions

Conceptualization: CLY. Formal analysis: NDZ, CXJ. Funds: NDZ, CLY. Investigation: NDZ, XYP, MJT. Methodology: CXJ, CLY. Project administration: NDZ, CLY. Resources: XB. Software: NDZhang. Supervision: XB, XYP. Writing - original draft: NDZ. Writing - review & editing: CLY, NDZ.

Funding

This work was supported by the Sichuan Medical Association (grant number: 2019HR35) and Chengdu University of Traditional Chinese Medicine Research Foundation (grant number: YYZX2019026).

Availability of data and materials

The datasets used and/or analyzed during the current study were available from the NCBI Gene Expression Omnibus (GEO; <https://www.ncbi.nlm.nih.gov/geo/>).

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Orcid

Naidan Zhang orcid: 0000-0002-3522-4544.

Chaixia Ji orcid: 0000-0002-3730-7618.

Xiao Bao orcid: 0000-0001-7844-0661.

Chengliang Yuan orcid: 0000-0002-2187-3498.

References

1. Brito-Zeron P, Retamozo S, Gheitasi H, Ramos-Casals M. Treating the Underlying Pathophysiology of Primary Sjogren Syndrome: Recent Advances and Future Prospects. *Drugs*. 2016;76(17):1601–1623.

<https://doi.org/10.1007/s40265-016-0659-z>.

2. Brown S, Navarro Coy N, Pitzalis C, et al. The TRACTISS protocol: a randomised double blind placebo controlled clinical trial of anti-B-cell therapy in patients with primary Sjogren's Syndrome. *BMC Musculoskelet Disord*. 2014;15:21. <https://doi.org/10.1186/1471-2474-15-21>.
3. Huang AF, Su LC, Jia H, Liu Y, Xu WD. No association of single nucleotide polymorphisms within H19 and HOX transcript antisense RNA (HOTAIR) with genetic susceptibility to systemic lupus erythematosus, rheumatoid arthritis, and primary Sjogren's syndrome in a Chinese Han population. *Clin Rheumatol*. 2017;36(11):2447–2453. <https://doi.org/10.1007/s10067-017-3833-0>.
4. Sogkas G, Atschekzei F, Adriawan IR, Dubrowskaja N, Witte T, Schmidt RE. Cellular and molecular mechanisms breaking immune tolerance in inborn errors of immunity. *Cell Mol Immunol*. 2021;18(5):1122–1140. <https://doi.org/10.1038/s41423-020-00626-z>.
5. Li N, Li L, Wu M, et al. Integrated Bioinformatics and Validation Reveal Potential Biomarkers Associated With Progression of Primary Sjogren's Syndrome. *Front Immunol*. 2021;12:697157. <https://doi.org/10.3389/fimmu.2021.697157>.
6. Zhang L, Xu P, Wang X, et al. Identification of differentially expressed genes in primary Sjogren's syndrome. *J Cell Biochem*. 2019;120(10):17368–17377. <https://doi.org/10.1080/21655979.2021.2000245>.
7. Qi X, Wang XQ, Jin L, Gao LX, Guo HF. Uncovering potential single nucleotide polymorphisms, copy number variations and related signaling pathways in primary Sjogren's syndrome. *Bioengineered*. 2021;12(2):9313–9331. <https://doi.org/10.1080/21655979.2021.2000245>.
8. Verstappen GM, Gao L, Pringle S, et al. The Transcriptome of Paired Major and Minor Salivary Gland Tissue in Patients With Primary Sjogren's Syndrome. *Front Immunol*. 2021;12:681941. <https://doi.org/10.3389/fimmu.2021.681941>.
9. Verstappen GM, Pringle S, Bootsma H, Kroese FGM. Epithelial-immune cell interplay in primary Sjogren syndrome salivary gland pathogenesis. *Nat Rev Rheumatol*. 2021;17(6):333–348. <https://doi.org/10.1038/s41584-021-00605-2>.
10. Xu H, Chen J, Wang Y, Wu Y, Liang Y. SELL and IFI44 as potential biomarkers of Sjogren's syndrome and their correlation with immune cell infiltration. *Genes Genet Syst*. 2021;96(2):71–80. <https://doi.org/10.1266/ggs.20-00053>.
11. Chen L, Lu D, Yu K, 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. *J Inflamm Res*. 2021;14:4143–4153. <https://doi.org/10.2147/JIR.S322318>.
12. Jang SI, Tandon M, Teos L, Zheng C, Warner BM, Alevizos I. Dual function of miR-1248 links interferon induction and calcium signaling defects in Sjogren's syndrome. *EBioMedicine*. 2019;48:526–538. <https://doi.org/10.1016/j.ebiom.2019.09.010>.
13. Johar AS, Mastronardi C, Rojas-Villarraga A, et al. Novel and rare functional genomic variants in multiple autoimmune syndrome and Sjogren's syndrome. *J Transl Med*. 2015;13:173. <https://doi.org/10.1186/s12967-015-0525-x>.

14. Imgenberg-Kreuz J, Sandling JK, Norheim KB, et al. DNA Methylation-Based Interferon Scores Associate With Sub-Phenotypes in Primary Sjogren's Syndrome. *Front Immunol.* 2021;12:702037. <https://doi.org/10.3389/fimmu.2021.702037>.
15. Bjork A, Richardsdotter Andersson E, Imgenberg-Kreuz J, et al. Protein and DNA methylation-based scores as surrogate markers for interferon system activation in patients with primary Sjogren's syndrome. *RMD Open.* 2020;6(1). <https://doi.org/10.1136/rmdopen-2019-000995>.
16. Ruffilli I. Sjogren's syndrome and chemokines. *Clin Ter.* 2014;165(6):e464-469. <https://doi.org/10.7417/CT.2014.1793>.
17. Ogawa N, Ping L, Zhenjun L, Takada Y, Sugai S. Involvement of the interferon-gamma-induced T cell-attracting chemokines, interferon-gamma-inducible 10-kd protein (CXCL10) and monokine induced by interferon-gamma (CXCL9), in the salivary gland lesions of patients with Sjogren's syndrome. *Arthritis Rheum.* 2002;46(10):2730–2741. <https://doi.org/10.1002/art.10577>.
18. Ruffilli I. Sjogren syndrome and MIG. *Clin Ter.* 2019;170(6):e478-e482. <https://doi.org/10.7417/CT.2019.2179>.
19. James JA, Guthridge JM, Chen H, et al. Unique Sjogren's syndrome patient subsets defined by molecular features. *Rheumatology (Oxford).* 2020;59(4):860–868. <https://doi.org/10.1093/rheumatology/kez335>.
20. Anvari S, Watkin L, Rajapakshe K, et al. Memory and naive gamma delta regulatory T-cell gene expression in the first 24-weeks of peanut oral immunotherapy. *Clin Immunol.* 2021;230:108820. <https://doi.org/10.1016/j.clim.2021.108820>.
21. Dashti N, Mahmoudi M, Gharibdoost F, et al. Evaluation of ITGB2 (CD18) and SELL (CD62L) genes expression and methylation of ITGB2 promoter region in patients with systemic sclerosis. *Rheumatol Int.* 2018;38(3):489–498. <https://doi.org/10.1007/s00296-017-3915-y>.
22. Manoussakis MN, Boiu S, Korkolopoulou P, et al. Rates of infiltration by macrophages and dendritic cells and expression of interleukin-18 and interleukin-12 in the chronic inflammatory lesions of Sjogren's syndrome: correlation with certain features of immune hyperactivity and factors associated with high risk of lymphoma development. *Arthritis Rheum.* 2007;56(12):3977–3988. <https://doi.org/10.1002/art.23073>.
23. Zhou Z, Yang Z, Wang J, Wang Q, Zhang H, Ju X. Research progress on tumour associated macrophages in gastric cancer (Review). *Oncol Rep.* 2021;45(4). <https://doi.org/10.3892/or.2021.7986>.
24. Quaranta V, Schmid MC. Macrophage-Mediated Subversion of Anti-Tumour Immunity. *Cells.* 2019;8(7). <https://doi.org/10.3390/cells8070747>.
25. Di Benedetto P, Ruscitti P, Vadasz Z, Toubi E, Giacomelli R. Macrophages with regulatory functions, a possible new therapeutic perspective in autoimmune diseases. *Autoimmun Rev.* 2019;18(10):102369. <https://doi.org/10.1016/j.autrev.2019.102369>.
26. Ma WT, Gao F, Gu K, Chen DK. The Role of Monocytes and Macrophages in Autoimmune Diseases: A Comprehensive Review. *Front Immunol.* 2019;10:1140. <https://doi.org/10.3389/fimmu.2019.01140>.

27. Ishimaru N, Saegusa K, Yanagi K, Haneji N, Saito I, Hayashi Y. Estrogen deficiency accelerates autoimmune exocrinopathy in murine Sjogren's syndrome through fas-mediated apoptosis. *Am J Pathol.* 1999;155(1):173–181. [https://doi.org/10.1016/S0002-9440\(10\)65111-5](https://doi.org/10.1016/S0002-9440(10)65111-5).
28. Gao CY, Yao Y, Li L, et al. Tissue-Resident Memory CD8 + T Cells Acting as Mediators of Salivary Gland Damage in a Murine Model of Sjogren's Syndrome. *Arthritis Rheumatol.* 2019;71(1):121–132. <https://doi.org/10.1002/art.40676>.
29. Pontarini E, Murray-Brown WJ, Croia C, et al. Unique expansion of IL-21 + Tfh and Tph cells under control of ICOS identifies Sjogren's syndrome with ectopic germinal centres and MALT lymphoma. *Ann Rheum Dis.* 2020;79(12):1588–1599. <https://doi.org/10.1136/annrheumdis-2020-217646>.
30. 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(9):1850–1861. <https://doi.org/10.1002/art.40165>.
31. Rios-Rios WJ, Sosa-Luis SA, Torres-Aguilar H. T Cells Subsets in the Immunopathology and Treatment of Sjogren's Syndrome. *Biomolecules.* 2020;10(11). <https://doi.org/10.3390/biom10111539>.
32. Hu S, Zhou M, Jiang J, et al. Systems biology analysis of Sjogren's syndrome and mucosa-associated lymphoid tissue lymphoma in parotid glands. *Arthritis Rheum.* 2009;60(1):81–92. <https://doi.org/10.1002/art.24150>.
33. Shah NR, Noll BD, Stevens CB, Brennan MT, Mougeot FB, Mougeot JC. Biosemantics guided gene expression profiling of Sjogren's syndrome: a comparative analysis with systemic lupus erythematosus and rheumatoid arthritis. *Arthritis Res Ther.* 2017;19(1):192. <https://doi.org/10.1186/s13075-017-1400-3>.
34. Yao Q, Song Z, Wang B, Qin Q, Zhang JA. Identifying Key Genes and Functionally Enriched Pathways in Sjogren's Syndrome by Weighted Gene Co-Expression Network Analysis. *Front Genet.* 2019;10:1142. <https://doi.org/10.3389/fgene.2019.01142>.
35. Fang TJ, Li RN, Lin YZ, et al. Association of F11R polymorphisms and gene expression with primary Sjogren's syndrome patients. *Int J Rheum Dis.* 2021;24(5):681–686. <https://doi.org/10.1111/1756-185X.14095>.
36. Aran D, Lasry A, Zinger A, et al. Widespread parainflammation in human cancer. *Genome Biol.* 2016;17(1):145. <https://doi.org/10.1186/s13059-016-0995-z>.
37. Lasry A, Aran D, Butte AJ, Ben-Neriah Y. Cancer Cell-Autonomous Parainflammation Mimics Immune Cell Infiltration. *Cancer Res.* 2017;77(14):3740–3744. <https://doi.org/10.1158/0008-5472.CAN-16-3383>.
38. Wang X, Undi RB, Ali N, Huycke MM. It takes a village: microbiota, parainflammation, paligenosis and bystander effects in colorectal cancer initiation. *Dis Model Mech.* 2021;14(5). <https://doi.org/10.1242/dmm.048793>.
39. Lai JH, Hung LF, Huang CY, Wu DW, Wu CH, Ho LJ. Mitochondrial protein CMPK2 regulates IFN alpha-enhanced foam cell formation, potentially contributing to premature atherosclerosis in SLE. *Arthritis*

- Res Ther. 2021;23(1):120. <https://doi.org/10.1186/s13075-021-02470-6>.
40. Luo Y, Zheng D, Mou T, et al. CMPK2 accelerates liver ischemia/reperfusion injury via the NLRP3 signaling pathway. *Exp Ther Med*. 2021;22(6):1358. <https://doi.org/10.3892/etm.2021.10793>.
41. Altorok N, Coit P, Hughes T, et al. Genome-wide DNA methylation patterns in naive CD4 + T cells from patients with primary Sjogren's syndrome. *Arthritis Rheumatol*. 2014;66(3):731–739. <https://doi.org/10.1002/art.38264>.
42. Huijser E, Gopfert J, Brkic Z, et al. Serum IFNalpha2 measured by single-molecule array associates with systemic disease manifestations in Sjogren's syndrome. *Rheumatology (Oxford)*. 2021. <https://doi.org/10.1093/rheumatology/keab688>.
43. Jara D, Carvajal P, Castro I, et al. Type I Interferon Dependent hsa-miR-145-5p Downregulation Modulates MUC1 and TLR4 Overexpression in Salivary Glands From Sjogren's Syndrome Patients. *Front Immunol*. 2021;12:685837. <https://doi.org/10.3389/fimmu.2021.685837>.
44. Khuder SA, Al-Hashimi I, Mutgi AB, Altorok N. Identification of potential genomic biomarkers for Sjogren's syndrome using data pooling of gene expression microarrays. *Rheumatol Int*. 2015;35(5):829–836. <https://doi.org/10.1007/s00296-014-3152-6>.
45. Wang D, Xue L, Yang Y, Hu J, Li G, Piao X. Temporal gene expression analysis of Sjogren's syndrome in C57BL/6.NOD-Aec1Aec2 mice based on microarray time-series data using an improved empirical Bayes approach. *Mol Biol Rep*. 2014;41(9):5953–5960. <https://doi.org/10.1007/s11033-014-3471-4>.
46. Yin J, Zheng J, Deng F, et al. Gene Expression Profiling of Lacrimal Glands Identifies the Ectopic Expression of MHC II on Glandular Cells as a Presymptomatic Feature in a Mouse Model of Primary Sjogren's Syndrome. *Front Immunol*. 2018;9:2362. <https://doi.org/10.3389/fimmu.2018.02362>.
47. Lin Y, Yao X, Yan M, et al. Integrated analysis of transcriptomics to identify hub genes in primary Sjogren's syndrome. *Oral Dis*. 2021. <https://doi.org/10.1111/odi.13943>.
48. Ivanchenko M, Thorlacius GE, Hedlund M, et al. Natural killer cells and type II interferon in Ro/SSA and La/SSB autoantibody-exposed newborns at risk of congenital heart block. *Ann Rheum Dis*. 2021;80(2):194–202. <https://doi.org/10.1136/annrheumdis-2019-216786>.
49. Kajio N, Takeshita M, Suzuki K, et al. Anti-centromere antibodies target centromere-kinetochore macrocomplex: a comprehensive autoantigen profiling. *Ann Rheum Dis*. 2021;80(5):651–659. <https://doi.org/10.1136/annrheumdis-2020-218881>.

Figures

Figure 1

Study flowchart. Abbreviations: pSS, primary Sjogren's syndrome; DEGs, differentially expressed genes.

Figure 2

Volcano plot and Heat map for DEGs and immune infiltration. (A) Red plot points represent up-regulated DEGs, blue plot points show down-regulated DEGs. (B) Light blue represents 16 types of immune cell, orange represents 13 types of immune function, each row represents a DEG, each column represents a sample, either control or pSS. Red represents DEGs are significantly related to immune infiltration, green represents DEGs are not related to immune infiltration. Abbreviations: DEGs, differentially expressed genes; pSS, primary Sjögren's syndrome; imc, immune cell; imf, immune function.

Figure 3

Correlation analysis of immune cells and immune functions. (A) Correlation analysis of 16 types of immune cells, horizontal and vertical axes demonstrate immune cell subtypes. (B) Correlation analysis of 13 types of immune functions, horizontal and vertical axes demonstrate immune function subtypes (red and blue represent strong positive and negative correlation respectively, while white represents no correlation). Abbreviations: aDCs, "activated" dendritic cells; Tfh, follicular helper T; Th, helper T cell; TIL, tumor infiltrating lymphocytes; Treg, regulatory T cells; APC, antigen presenting cell; CCR, CC chemokine receptors; HLA, human leukocyte antigen; MHC, major histocompatibility complex; IFN, interferon.

Figure 4

Comparison between immune score in pSS and controls. (A) Comparison of 16 types of immune cells, horizontal axes demonstrates immune cell subtypes, vertical axes demonstrates immune score (green represent control and red represent pSS). (B) Comparison of 13 types of immune functions, horizontal axes demonstrates immune functions, vertical axes demonstrates immune score (blue represent control and red represent pSS). Abbreviations: aDCs, "activated" dendritic cells; Tfh, follicular helper T; Th, helper T cell; TIL, tumor infiltrating lymphocytes; Treg, regulatory T cells; APC, antigen presenting cell; CCR, CC chemokine receptors; HLA, human leukocyte antigen; MHC, major histocompatibility complex; IFN, interferon. * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$.

Figure 5

Analysis of DEGs and immune infiltration, diagnostic efficacy. (A) Correlation of DEGs and immune infiltration, horizontal axes demonstrates immune infiltration, vertical axes demonstrates DEGs. (B) Column map of top 5 genes, horizontal axes demonstrates risk of pSS, vertical axes demonstrates DEGs. (C) ROC curve of DEGs. Abbreviations: pSS, primary Sjogren's syndrome; AUC, area under curve.

Figure 6

DO analyses of the DEGs. Horizontal axes demonstrates generatio (count), vertical axes demonstrates the top 10 diseases with the highest frequency of DEGs. Abbreviations: DO, disease ontology; DEGs, differentially expressed genes.

Figure 7

GO analyses of the DEGs. Horizontal axes demonstrates generatio, vertical axes demonstrates functional enrichment analysis in BP, CC and MF, respectively. Abbreviations: pSS, primary Sjogren's syndrome; AUC, area under curve. Abbreviations: GO, gene ontology; DEGs, differentially expressed genes; BP, biological process; CC, cellular component; MF, molecular function.

Supplementary Files

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

- [TableS1.docx](#)