scientific reports

OPEN

LncRNA and mRNA expression profile of peripheral blood mononuclear cells in primary Sjögren's syndrome patients

Yu Peng[1,2,4], Xuan Luo[1,2,4], Yingying Chen[1,2], Linyi Peng[1,2], Chuiwen Deng[1,2], Yunyun Fei[1,2,3✉], Wen Zhang[1,2,3✉] & Yan Zhao[1,2,3✉]

The aim of this study was to elucidate the expression profile and the potential role of long non-coding RNA (LncRNA) in the peripheral blood mononuclear cells of primary Sjögren's syndrome (pSS) patients. RNA-seq technology was used to detect the differentially expressed LncRNAs and mRNAs between five age-and sex-matched paired pSS patients and healthy control PBMCs. The selected LncRNAs were detected in the validation study by RT-qPCR in 16 paired pSS patients and healthy controls. The GO, KEGG, co-localization, and co-expression analysis were performed to enrich the potential gene functions and pathways. In this study, 44 out of 1772 LncRNAs and 1034 out of 15,424 mRNAs were expressed differentially in the PBMCs of pSS patients. LINC00426, TPTEP1-202, CYTOR, NRIR, and BISPR were validated as aberrantly expressed, and these LncRNAs strongly correlated with disease activity of pSS. GO and KEGG pathway analysis revealed the significant enrichment of biological processes, cellular components, and molecular function of the up and down-regulated mRNAs, which were mainly concentrated in the immune response and immune system processes. Co-localization and co-expression analysis also revealed that differentially expressed LncRNAs in the PBMCs of pSS were strongly correlated to the mRNA functioning associated with immune response and cell metastasis. Numerous LncRNAs and mRNAs were found differentially expressed in the PBMCs of pSS patients, especially NRIR and BISPR; they interacted with the co-localized and co-expressed mRNAs, which might participate in the pathogenesis of pSS through the NF-κB, JAK-STAT, and other signaling pathways that regulate cell metastasis.

Primary Sjögren's syndrome (pSS) is a complex and heterogeneous systemic autoimmune disease with unknown etiology; and is characterized by the dysfunction of exocrine glands (mainly the salivary and lacrimal glands) and the infiltration of lymphocytes into the affected tissues[1,2]. Patients with pSS are commonly manifested by dry mouth and dry eyes, and other symptoms of organs involvement including lung, kidney and others[3,4]. pSS is a multifactorial disease resulting from genetic, environmental, and inflammatory factors; and aberrantly regulated innate and adaptive immunity pathways that were thought to play crucial roles in the disease[5].

However, the inflammatory factors need precise regulation at the transcriptional and post-transcriptional levels. Recently, some studies have revealed that long non-coding RNA (LncRNA), which is transcribed by RNA polymerase II, consists of more than 200 nucleotides that cannot translate into proteins that play important roles in regulating inflammatory pathways[6,7]. LncRNA was widely expressed in organisms and participated in many important biological processes, including genomic imprinting, chromatin modification, intracellular signaling and so on[8]. Recent studies have revealed that LncRNAs are involved in the regulation of innate and adaptive immunity. LncRNA-COX2[9], Lethe[10], and linc1992[11] can interact with RNA binding proteins (RBPs) and promote the release of corresponding cytokines, including IL-6, TNF-α, and IL-1β, which play a significant role in the regulation of innate immune responses. The expression of LncRNAs was highly specific at different

[1]Department of Rheumatology, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing, China. [2]Key Laboratory of Ministry of Health, National Clinical Research Center for Dermatologic and Immunologic Diseases (NCRC-DID), Beijing, China. [3]Department of Rheumatology, Clinical Immunology Center, Peking Union Medical College Hospital, Beijing, China. [4]These authors contributed equally: Yu Peng and Xuan Luo. ✉email: feiyunyun@pumch.cn; zhangwen91@sina.com; zhaoyan_pumch2002@aliyun.com

	pSS patients (n = 21)	Healthy controls (n = 21)
Demographics		
Age at onset	47.86 ± 11.55	45.86 ± 8.55
Sex (female)	21	21
Fever (n%)	2 (9.52%)	0
Joint pain (n%)	8 (38.10%)	0
ESSDAI (mean, S.D)	4.19 ± 2.54	N.D
Laboratory examinations		
Anti-SSA antibody (+/−)	18/3	Negative
Anti-SSB antibody (+/−)	8/13	Negative
Anti-Ro-52 antibody (+/−)	17/4	Negative
IgG (g/L, mean ± S.D)	21.59 ± 7.05	10.04 ± 2.72
IgA (g/L, mean ± S.D)	3.02 ± 1.14	2.06 ± 0.69
IgM (g/L, mean ± S.D)	1.62 ± 0.69	0.99 ± 0.37
C3 (g/L, mean ± S.D)	0.98 ± 0.23	1.04 ± 0.16
C4 (g/L, mean ± S.D)	0.19 ± 0.19	0.19 ± 0.07
ESR (mm/h, mean ± S.D)	28.05 ± 1.57	5.93 ± 3.13
RF (IU/mL, mean ± S.D)	126.51 ± 129.68	5.13 ± 4.06
CRP (mg/L, mean ± S.D)	2.35 ± 3.38	N.D

Table 1. Demographic and clinical features of pSS patients and healthy controls in this study. N.D. represented for "Not detect".

stages of CD4$^+$T cell differentiation, most of which could bind the transcription factors related to T cell functions (T-bet, GATA-3, and STAT4). LincR-Ccr2-5'AS[12] and Lnc-MAF-4[13] could regulate the specific Th2 differentiation. LncRNAs were recently reported to have important regulatory roles in autoimmune diseases (AID). NETA1 (activated MAPK signaling pathway through Toll like receptor (TLR) 4) was found to be elevated in systemic lupus erythematosus (SLE) patients[14]. Some inflammatory LncRNAs involved in the pathogenesis of autoimmune diseases were also found elevated in RA[15], MS[16], and other diverse autoimmune diseases. LncRNAs has significant role in the pathogenesis of autoimmune diseases. Some previous literatures have reported that LncRNAs had important roles in the pathogenesis (Immune regulation and cell metabolism) of labial gland of pSS patients[17,18]. While the different expression and function have not been found in pSS patients' peripheral blood mononuclear cells (PBMCs) yet. Thus, in the present study, we used the RNA-seq technology to analyze the expression profile of LncRNAs and mRNAs in pSS patients' PBMCs to investigate and further improve the potential roles of LncRNAs in the pathogenesis of pSS.

Methods

Patients and healthy controls.

Twenty-one pSS patients and 21 age-sex matched healthy controls were enrolled in this study. All the pSS patients were recruited from the Peking Union Medical College Hospital and met the American-European consensus criteria for diagnosis[19] and have not received any medication treatment. This study was approved by the Ethics Committee of Peking Union Medical College Hospital (Approval number JS-2038). This study was performed in accordance with the Declaration of Helsinki, all methods were performed in accordance with the relevant guidelines and regulations. Patients with other autoimmune diseases, active/severe infection or malignant diseases were excluded from the study. All the pSS patients and matched healthy controls in this study have signed the informed consent. The RNA-seq analysis was carried out on 5 pSS patient samples and 5 matched healthy controls while the following validation study consisted of 16 pSS patients and 16 matched healthy controls. After the patient enrolling stage, we enrolled 21 female pSS patients who met our inclusion criteria. Patient's disease activity was determined by EULAR Sjögren's syndrome disease activity index (ESSDAI)[20], and serological examinations, including Immunoglobulin G (IgG), immunoglobulin A (IgA), immunoglobulin M (IgM), anti-SSA and SSB antibody, erythrocyte sedimentation rate (ESR), rheumatoid factor (RF), and so on, were performed for the study subjects. The detailed demographic and clinical features of patients and healthy controls are listed in Table 1.

Sample processing.

Peripheral blood (10 mL) was obtained from each subject and all the samples were collected in the ethylene diamine tetraacetic acid (EDTA) tubes. Peripheral blood mononuclear cells (PBMCs) were isolated using the Ficoll density gradient centrifugation and were counted using Cellometer Auto T4 (Nexcelom Bioscience, USA). Total RNA was extracted using TRIzol reagent (Invitrogen, Carlsbad, CA, USA) and the concentration of total RNA was measured by NanoDrop2000c spectrophotometer (NanoDrop Technologies, Wilmington, DE, USA). Then, the total RNA was reverse transcribed into complementary DNA (cDNA) using PrimeScript RT Master Mix (TaKaRa, Dalian, China) and was stored at − 80 °C for the following study.

Library preparation, LncRNA sequencing. A total amount of 20 ng RNA per sample was used as input material for the RNA sample preparations. Firstly, ribosomal RNA was removed by Epicentre Ribo-zero rRNA Removal Kit (Epicentre, USA), and rRNA free residues was cleaned up by ethanol precipitation. Subsequently, sequencing libraries were generated using the rRNA-depleted RNA by NEBNext Directional RNA Library Prep Kit for Illumina (NEB, USA) following manufacturer's recommendations. Then, we used the Qubit RNA Assay Kit and Qubit 2.0 Flurometer (Life Technologies, CA, USA) to quantify RNAs. RNA integrity was assessed using the RNA Mano 6000 Assay Kit of the Bioanalyzer 2100 system (Agilent Technologies, CA, USA).

Clustering and sequencing. The clustering of the index-coded samples was performed on a cBot Cluster Generation System using TruSeq PE Cluster Kit v3-cBot-HS (Illumina, USA) according to the manufacturer's instructions. After cluster generation, the libraries were sequenced on an Illumina Hiseq 2500 platform and 125 bp paired-end reads were generated.

Quality control. Raw data (raw reads) of fastq format were firstly processed through in-house perl scripts. In this step, clean data (clean reads) were obtained by removing reads containing adapter, reads on containing poly-N and low-quality reads from raw data. At the same time, Q20, Q30 and GC content of the clean data were calculated. And all the down stream analyses were based on the high-quality clean data.

Mapping to reference genome. Reference genome and gene model annotation files were downloaded from genome website directly. Index of reference genome was built using HISAT2 v2.0.4 and paired-end clean reads were aligned to the reference genome using HISAT2 v2.0.4.

Transcriptome assembly. The mapped reads of each sample were assembled by StringTie v1.3.3 in a reference-based approach. String tie used a novel network flow algorithm as well as an optional de novo assembly step to assemble and quantitate full-length transcripts representing multiple splice variants for each gene locus.

Differentially expressed LncRNAs. Differentially expressed LncRNA were analyzed by DESeq package based on the negative binomial distribution test. A threshold FDR < 0.05 and |log$_2$Fold change| ≥ 1.

Real-time quantitative polymerase chain reaction (RT-qPCR). The stored cDNAs were processed by TB Green Premix Taq II kit (TaKaRa, Dalian, China) and 7900 HT Fast Real-Time PCR System (ABI, Foster City, CA, USA). SYBR Green Buffer, other reagents in this kit, and the cDNAs were in a total volume of 10 uL and the RT-qPCR was performed in 384-well microplates. The conditions we used to select LncRNAs were: (a) The fold change of the LncRNAs must be greater than 1.5 fold compared with healthy controls and also P-value must be less than 0.05; (b) Genes with LncRNA-mRNA repeated seruences and LncRNAs which could not find information in databases were not selected in the RT-qPCR validation study. And based on the results of the RNA-seq, we selected 11 differentially expressed LncRNAs meeting the above conditions for validation study. The selected LncRNAs' primer sequences were listed in supplementary Table 1. The expression of each LncRNA was represented as fold changes using the $^{\Delta\Delta}$Ct method to obtain RT-qPCR results.

GO and KEGG enrichment analysis. Gene ontology (GO) enrichment analysis of differentially expressed genes or LncRNA target genes were implemented by the GOseq R package, with the gene length bias being corrected. GO terms with corrected P-value less than 0.05 were considered significantly enriched by differentially expressed genes. KEGG was a database resource for understanding high-level functions and utilities of the biological system, and KOBAS software was used to test the statistical enrichment of differential expression genes or LncRNA target genes in KEGG pathways[21,22].

Co-expression and co-location analysis. Cis role is LncRNA acting on neighboring target genes. We searched coding genes 10k/100k upstream and downstream of LncRNA, found the adjacent mRNAs and then analyzed their function next. Trans role is LncRNA to identify each other by the expression level. We calculated the expressed correlation between LncRNAs with R function to find out the co-expressed mRNAs and then predicted the potential functions of the LncRNAs in pSS.

Statistical analysis. The results of the RT-qRCR were verified through t-test if satisfied normality distribution or through Nonparametric test if did not satisfy normality distribution. And the correlation between clinical features and validated LncRNAs were tested through Spearman's correlation tests. All the data were analyzed using R Statistical Package (version 3.6.3; https://www.r-project.org). A P-value (two-tailed) < 0.05 was thought to be statistically significant.

Results

Differentially expressed LncRNAs and mRNAs in PBMCS of pSS patients.
Five pSS patients and 5 sex-age matched healthy controls participated in the RNA sequencing study. The volcano plot (Fig. 1A,B) showed the aberrantly expressed LncRNAs and mRNAs in pSS patients as compared to the healthy controls. Among all the 1772 tested LncRNAs, 44 LncRNAs were expressed aberrantly as found in this study, including 26 up-regulated and 18 down-regulated, and the top 10 up-regulated and down-regulated LncRNAs are listed in Table 2. Out of 15,424 tested mRNAs in this study, a total of 1034 mRNAs were found to express aberrantly, including 562 up-regulated and 472 down-regulated. The distribution of these differentially expressed LncRNAs
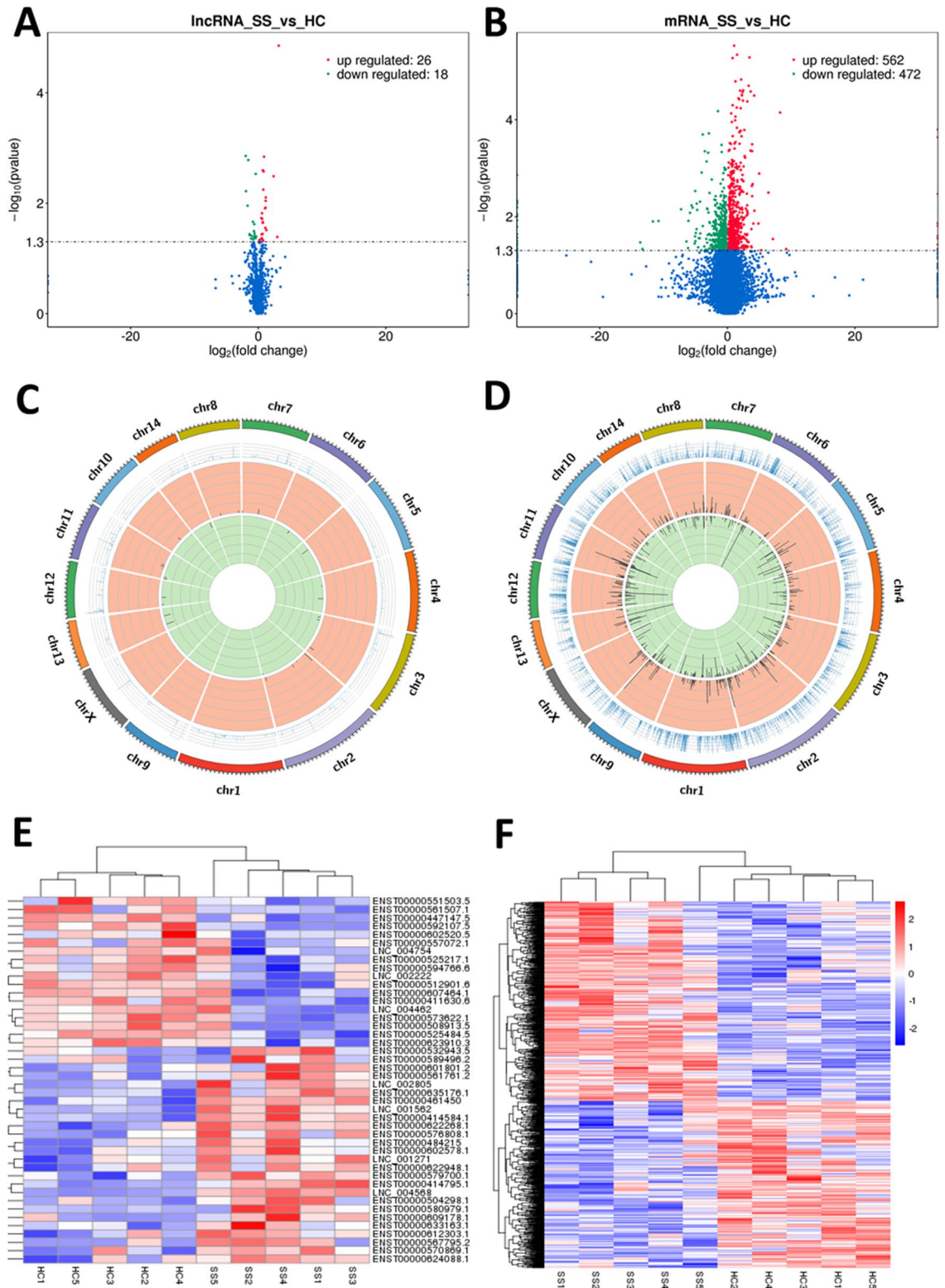
Figure 1. The LncRNA and mRNA expression profile of PBMCs in pSS patients. (**A,B**) The volcano plots for the differentially expressed LncRNAs and mRNAs between pSS patients and matched healthy controls, red stands for up-regulated and green stands for down-regulated. (**C,D**) The distribution of differentially expressed LncRNAs and mRNAs on chromosomes. (**E,F**) The heat maps for the hierarchical clustering of differentially expressed LncRNAs and mRNAs, red stands for up-regulated and blue stands for down-regulated. (R Statistical Package (version 3.6.3; https://www.r-project.org) was used to analyze the data of this figure and produce the figure).

Up-regulated			Down-regulated		
Gene ID	Fold change	P-value	Gene ID	Fold change	P-value
XLOC_167004	9.25	< 0.001***	ENSG00000238121.5	−3.92	0.001**
ENSG00000260114.2	7.89	0.041*	ENSG00000215068.7	−3.78	0.006**
ENSG00000225964.5	5.28	0.003**	XLOC_078623	−3.18	0.011*
ENSG00002726666.1	2.41	0.031*	ENSG00000100181.21	−2.95	0.002**
ENSG00000265479.5	2.26	0.008**	ENSG00000246223.8	−2.58	0.036*
XLOC_102167	2.24	0.009**	ENSG00000269893.6	−2.51	0.038*
ENSG00000282851.1	2.20	0.012*	ENSG00000258511.1	−2.17	0.044*
ENSG00000243398	2.16	0.027*	ENSG00000257261.5	−2.07	0.039*
ENSG00000262979.1	2.08	0.012*	ENSG00000255026.1	−1.78	0.022*
ENSG00000277999.1	1.88	0.001**	ENSG00000203875.10	−1.74	0.039*

Table 2. The Top 10 up-regulated and down-regulated LncRNAs of PBMCs in pSS patients that are significantly and differentially expressed. Gene ID were the gene in which the transcript resides. *P-value < 0.05, **P-value < 0.01, ***P-value < 0.001.

and mRNAs was shown in Fig. 1C,D. We did not observe any special enrichment of the differentially expressed LncRNAs or mRNAs on the specific chromosomes. The hierarchical clustering analysis (Fig. 1E,F) showed distinct expression signatures for both LncRNAs and mRNAs. All the RNA-seq data were stored in the GEO database (Accession number GSE145065).

The results of RT-qPCR of 11 selected LncRNAs and the correlation between differentially expressed LncRNAs and clinical features. Sixteen pSS patients and 16 matched healthy controls were involved at this stage. Eleven significantly differentially expressed LncRNAs were chosen for validation (the selection condition and primer sequences were listed above). Among the 11 selected LncRNAs, we conformed 5 LncRNAs differentially expressed through the RT-qPCR (Fig. 2), including BISPR (Fig. 2A, P < 0.001), CYTOR (Fig. 2B, P = 0.026), LINC00426 (Fig. 2D, P = 0.008), NRIR (Fig. 2G, P < 0.001), and TPTEP1-202 (Fig. 2K, P = 0.033). It was noted that NRIR and BISPR were significantly up-regulation in pSS patients, which were almost absent to lowly expressed in healthy controls but highly expressed in pSS patients. Then we analyzed the correlation between the differentially expressed LncRNAs and pSS patients' clinical features. The significant correlation between LncRNAs and patient's clinical features are shown in Fig. 3. BISPR and CYTOR were positively correlated with ESSDAI (r = 0.6734, P = 0.0042 with BISPR and r = 0.6724, P = 0.0043 with CYTOR) and serum IgG (r = 0.5106, P = 0.0433 with BISPR and r = 0.6352, P = 0.0082 with CYTOR). LINC00426 was negatively correlated with C-reactive protein (CRP) (r = −0.5427, P = 0.0366). NRIR was positively correlated with ESSDAI (r = 0.7704, P = 0.0005) and C4 (r = 0.6826, P = 0.0071). TPTEP1-202 was negatively correlated with ESSDAI (r = −0.6129, P = 0.0116) and IgG (r = −0.5646, P = 0.0227), while the correlation between NRIR and C4, LINC00426 and CRP seemed to be driven by some particular high outliers (Fig. 3E,F), thus LncRNA-NRIR and LINC00426 only had the trend of positive correlation with C4 and CRP, these correlation were still needed to be further tested in a lager sample size study.

Co-localization and co-expression between differentially expressed LncRNAs and mRNAs. Tables 3 and 4 show the results of co-localization and co-expression analysis. LncRNA might regulate the adjacent protein-coding genes. Among the 22 co-localized mRNAs, HMGB1, RSAD2, CMPK2, MVB12A, and BST2 were reported to be strongly related to the development and differentiation of immune cells and regulation of the release of the inflammatory cytokines. PGLS, MRPL34, SLC27A1, and ABHD8 were reported to be significantly related to the mitochondrial function, glycolysis, and fatty acid metabolism. Other co-localized mRNAs were related to the degradation and modification of DNAs and RNAs. Among the 21 co-expressed mRNAs, IFI44, RAPGEF5, HDHD2, CRIP2, MIB2, and NF1 were strongly related to the immune-related signaling pathways and the release of the inflammatory cytokines. LOXL2, CALCOCO1, CRIP2, and HSD17B11 were associated with glycolysis, fatty acid metabolism, and cell metastasis. Other co-localized mRNAs were related to DNA repair, the regulation of tumor and cell morphogenesis. The gene network of the co-localized and co-expressed mRNAs are shown in Fig. 4. The differentially expressed LncRNAs might have been involved in the pathogenesis of pSS through interaction with co-expressed and co-localized mRNAs.

The results of GO and KEGG analysis for the differentially expressed mRNAs. GO and KEGG pathway analysis was mainly used to discover the potential pathogenic mechanism in pSS patients. Figure 5 shows the results of the Go and KEGG analysis. We found that the top significantly enriched biological process (BP) for the down-regulated mRNAs were cellular macromolecule metabolic process, cellular macromolecule biosynthetic process, intracellular transport, cytoplasmic transport, and translation while the top significantly enriched BP for the up-regulated mRNAs were immune system process, response to stress, immune response, defense response, and immune effector process. The most significantly enriched cellular component (CC) for the down-regulated mRNAs were intracellular parts, intracellular organelle, membrane-bounded organelle, and intercellular organelle. In case of the up-regulated mRNAs, it was cell part, cell, intracellular part, intercellular,
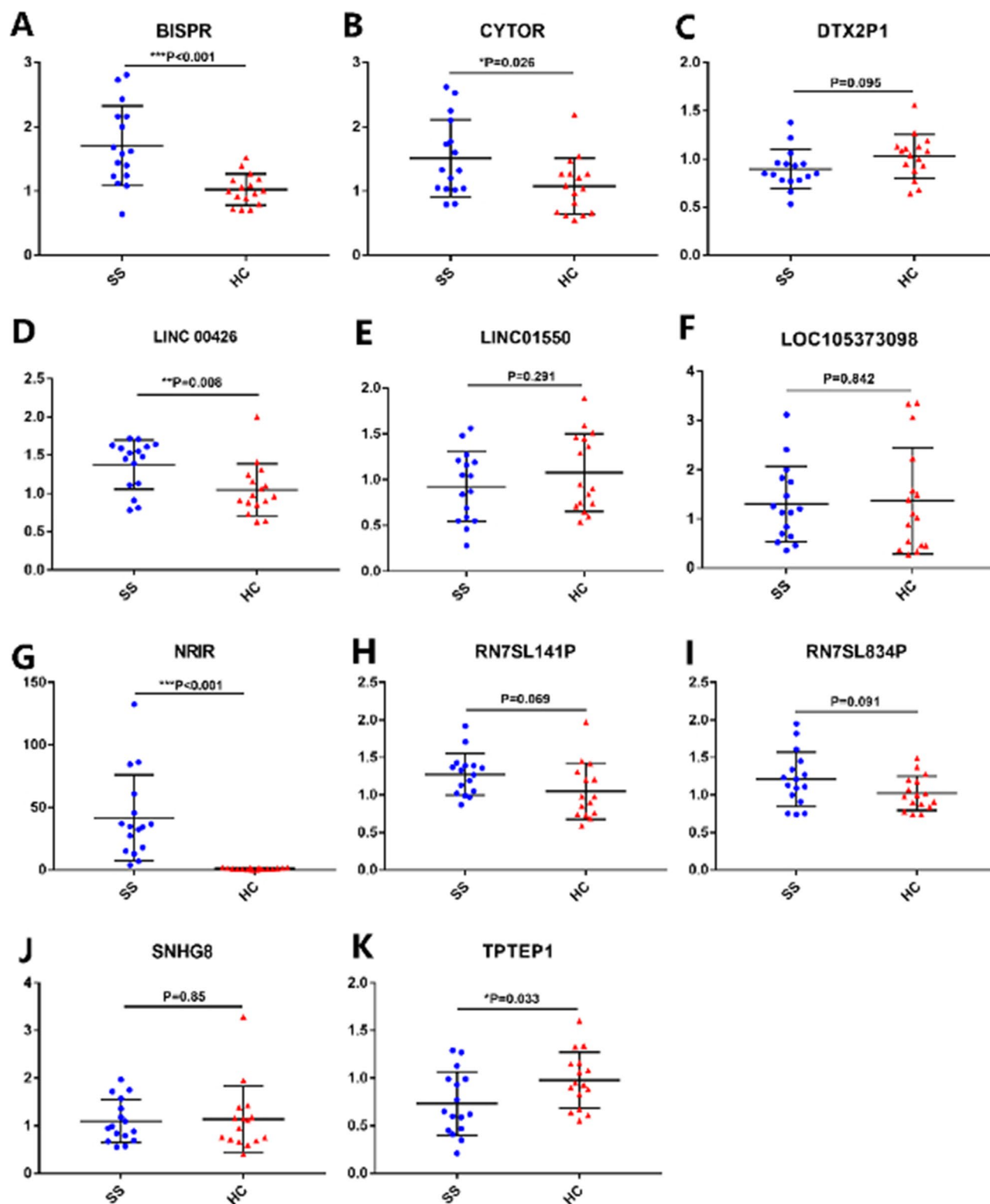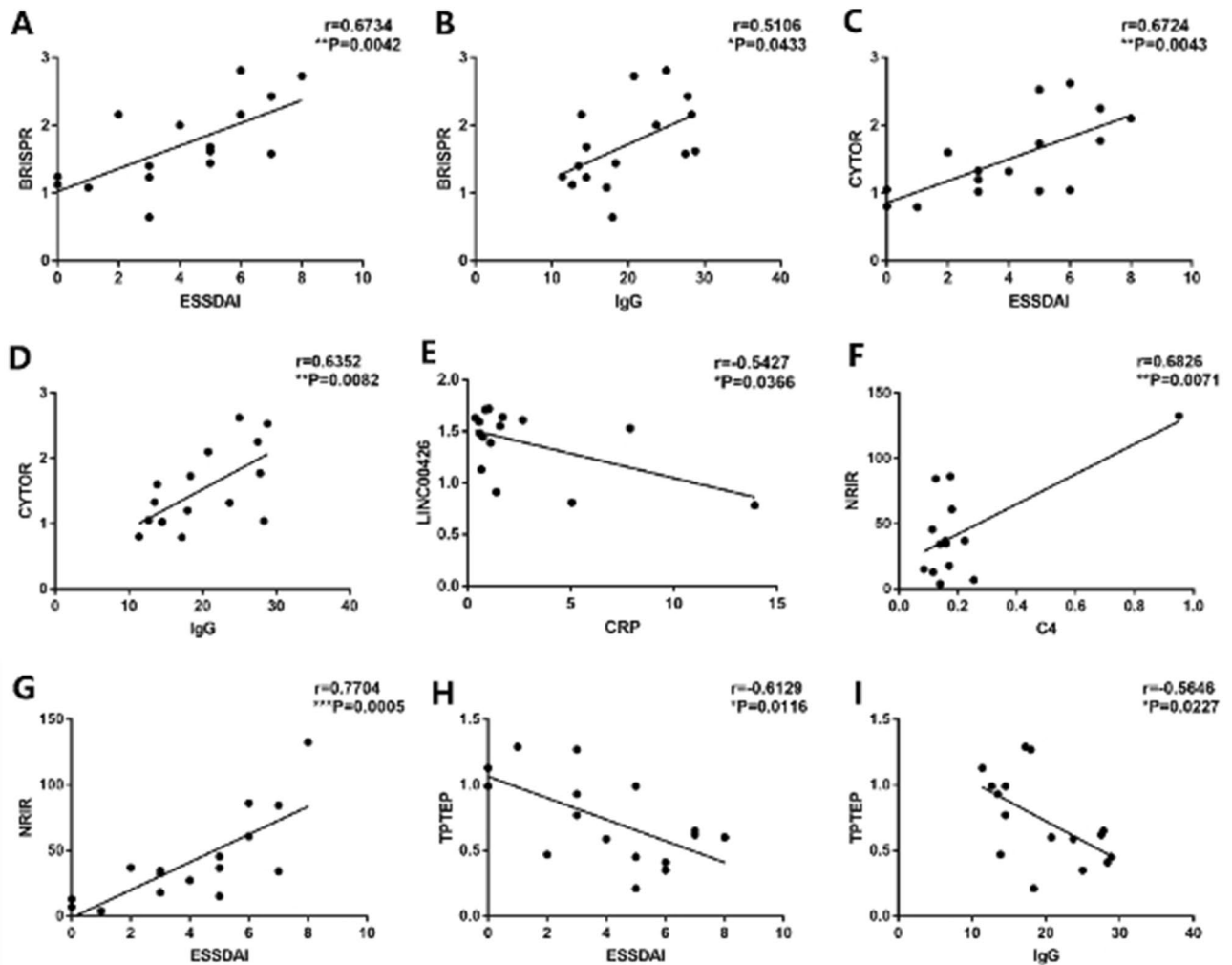
Figure 2. (**A–K**) The results of the validation study for the selected LncRNAs. BISPR (**A**), CYTOR-235 (**B**), DTX2P1-UPK3BP1-PMS2P11 (**C**), LINC00426 (**D**), LINC01550 (**E**), LOC105373098 (**F**), NRIR (**G**), RN7SL141P (**H**), RN7SL834P (**I**), SNHG8 (**J**), TPTEP1-202 (**K**) were selected for the validation study. Only BISPR (**A**), CYTOR (**B**), LINC00426 (**D**), NRIR (**G**), TPTEP1-202 (**K**) were significantly and differentially expressed.

Figure 3. The correlations between the differentially expressed LncRNAs and patients' clinical features. (**A,B**) The significant correlation between BISPR and ESSDAI, IgG. (**C,D**) The significant correlation between CYTOR and ESSDAI and IgG. (**E**) The significant correlation between LINC00426 and CRP. (**F,G**) The significant correlation between NRIR and ESSDAI, C4. (**H,I**) The significant correlation between TPTEP1-202 and ESSDAI, IgG.

and organelle. The significantly enriched molecular function (MF) for down-regulated mRNAs were molecular function, binding, protein binding, organic cyclic compound binding and heterocyclic compound binding. The same MFs were enriched for the up-regulated mRNAs. The results of KEGG pathway analysis for the differentially expressed mRNAs indicated that the down-regulated mRNAs in PBMCs of pSS patients were significantly enriched in ribosome, acute myeloid leukemia, legionellosis, chronic myeloid leukemia, and colorectal cancer; whereas the up-regulated mRNAs were significantly enriched in primary immunodeficiency, antigen processing and presentation, herpes simplex infection, RIG-I-like receptor signaling pathway, and influenza A, which were all closely related to the immunity.

Discussion

With the development of the sequencing technology, LncRNAs have increasingly being found to be involved in various biological processes. Through its interaction with DNA, RNA, and proteins, LncRNAs could influence the activity of immune cells at transcriptional or post-transcriptional levels[23]. Nowadays, an increasing number of aberrantly expressed LncRNAs and their functions are being reported in the autoimmune diseases and infectious diseases[24,25]. Although the potential role of LncRNAs in pSS has been speculated previously, the detailed role of LncRNAs in the pathogenesis of pSS remains unknown. In this study, we used the RNA-seq technology to identify the differentially expressed LncRNAs and mRNAs in PBMCs of pSS patients, and to find out the potential function of these genes.

In this study, we enrolled 21 paired pSS patients and healthy controls. Through sequencing analysis we found that 44 out of 1772 LncRNAs and 1034 out of 15,424 mRNAs were differentially expressed in the PBMCs of pSS patients.

LncRNA gene name	P-value	Fold change	Regulation	Co-located mRNA name	Location
LINC00426	0.001	0.26	Down	KATNAL1	Upstream
				HMGB1	Downstream
				UBE2L5	Upstream
TPTEP1-202	0.002	0.34	Down	CCT8L2	Upstream
NRIR	0.003	5.28	Up	RSAD2	Upstream
				CMPK2	Downstream
				RNF144A	Upstream
BISPR	0.012	2.2	Up	PGLS	Upstream
				GTPBP3	Downstream
				MRPL34	Downstream
				ACO10319.2	Downstream
				PLVAP	Upstream
				CCDC194	Upstream
				AN08	Upstream
				SLC27A1	Upstream
				MVB12A	Upstream
				ABHD8	Upstream
				NXNL1	Downstream
				DDA1	Downstream
				BST2	Upstream
				TMEM221	Downstream

Table 3. The results of co-localization analysis between differentially expressed LncRNAs and mRNAs.

LncRNA gene name	P-value	Fold change	Regulation	Correlated mRNA name	r	P-value
LINC00426	0.001	0.26	Down	RAD52	0.953	0.00002
				TPRG1	0.955	0.00002
				LOXL2	0.957	0.00001
				CALCOCO1	0.957	0.00001
				LRWD1	0.951	0.00002
TPTEP1-202	0.002	0.34	Down	CRIP2	0.961	0.00001
				ZNF286B	0.953	0.00002
CYTOR	0.002	1.64	Up	FRMD3	0.967	0.000005
				MCM8	0.954	0.00002
				HSD17B11	0.951	0.00002
				HDHD2	0.952	0.00002
				TEX46	0.950	0.00003
				GOLGA80	0.952	0.00002
				ARMC9	0.951	0.00002
NRIR	0.003	5.28	Up	RAPGEF5	0.953	0.00002
				IFI44	0.959	0.00001
				MIB2	0.961	0.00001
				FAM241B	0.963	0.000008
				NF1	0.961	0.000009
BISPR	0.012	2.2	Up	ATXN7	0.951	0.00002
				FAM149B1	0.954	0.00001

Table 4. The detailed results of the co-expression analysis between the differentially expressed LncRNAs and mRNAs.

The up-regulated mRNAs were significantly enriched in the immune system process, response to stress, and immune response, indicating that a variety of immune-related genes were aberrantly up-regulated and participated in the pathogenesis of pSS. And the KEGG pathway analysis revealed that up-regulated mRNAs were enriched in primary immunodeficiency, antigen processing and presentation, herpes simplex infection, RIG-I-like receptor signaling pathway, and influenza A. Similar to other autoimmune diseases, pSS patients also had a
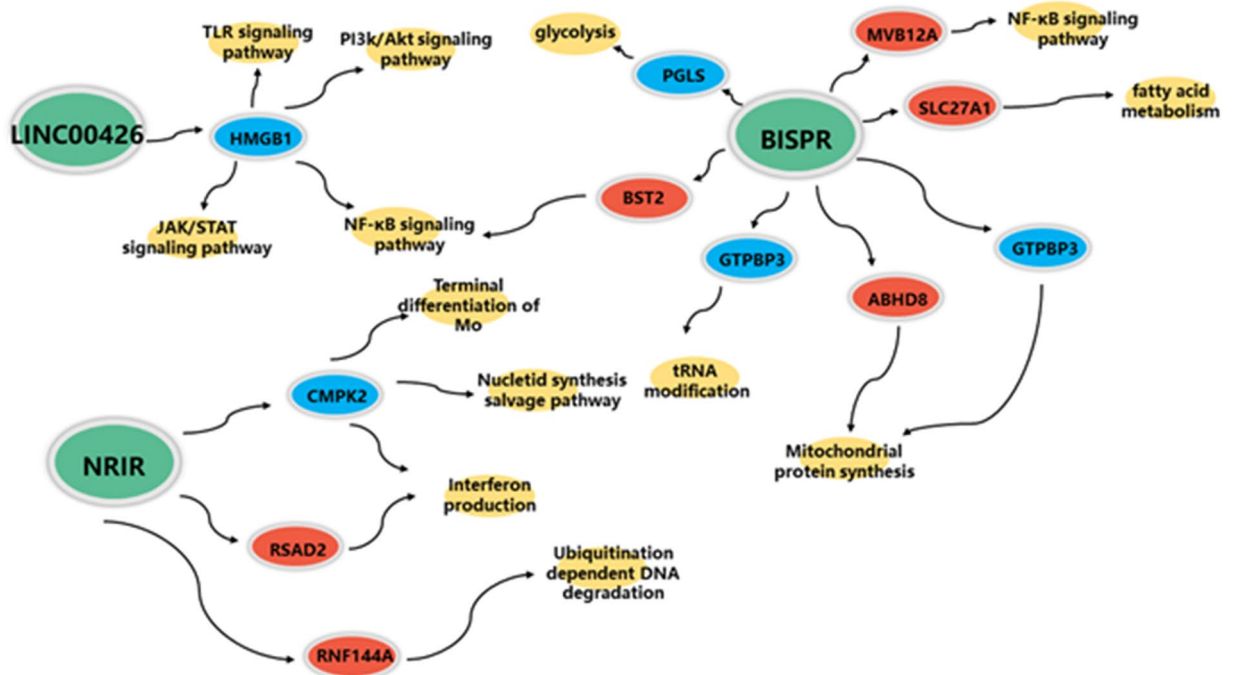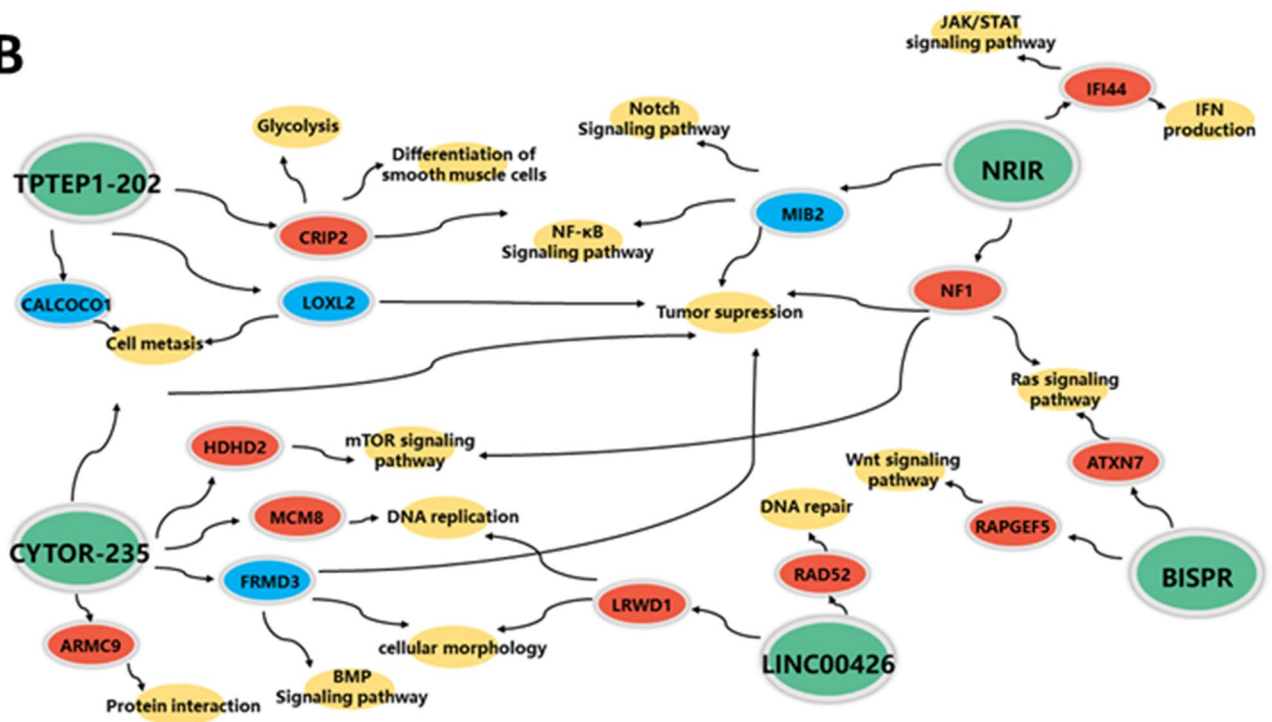
Figure 4. The gene pathway analysis of the co-expression and co-localization of mRNAs in this study. (**A**) The gene pathway network of the co-location mRNAs, green ellipses represented for LncRNA, red ellipses represent for the upstream mRNA, blue ellipses represent for the downstream mRNA and the yellow ellipses represent for the gene functions or signaling pathways. (**B**) The gene pathway network of the co-expression mRNAs, green ellipses represent for LncRNA, red/blue ellipses represent for the co-expression mRNA (up-regulated/down-regulated); yellow ellipses represent the gene functions or signaling pathways.

variety of autoantigens (mainly SSA and SSB antigen), which could be presented to the T cells or be recognized by B cells to trigger the release of autoantibodies to promote disease progression[26,27], and this attached importance
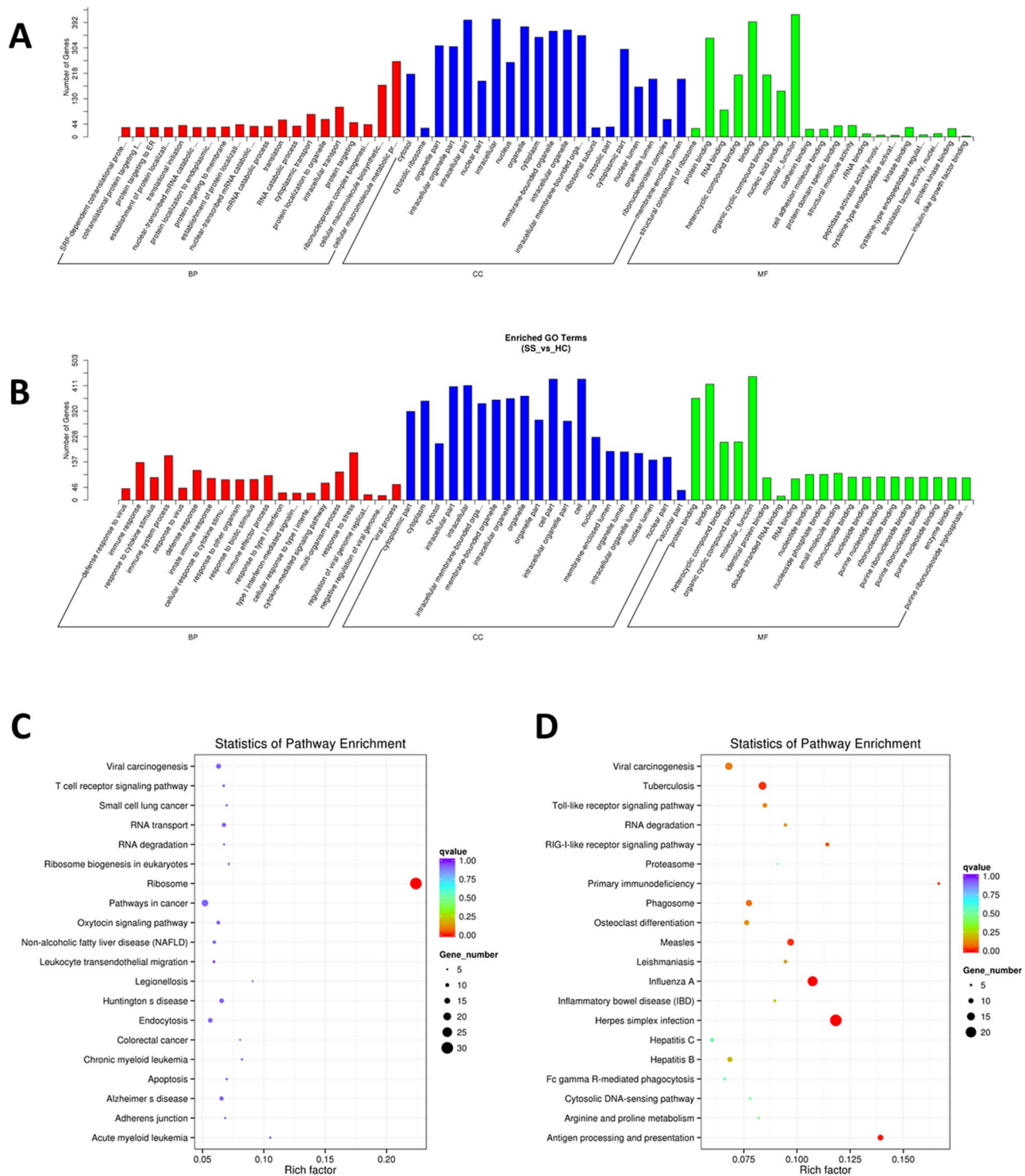
Figure 5. The results of GO and KEGG analysis of the differentially expressed mRNAs in the study. (**A**) The GO results for the down-regulated mRNAs. (**B**) The GO results for the up-regulated mRNAs. (**C**) The KEGG analysis results for the down-regulated mRNAs[21,22]. (**D**) The KEGG analysis results for the up-regulated mRNAs[21,22].

to the enrichment function of antigen presentation, which indicated that LncRNAs were also involved in the production of pSS autoantibodies. Previous study of the LncRNA in the monor salivary glands of pSS patients also had the similar results[28], and some immune-related signaling pathways like NF-κB signaling pathway, TNF signaling pathway, and natural killer cell mediated cytotoxicity were significantly enriched in the minor salivary gland of pSS patients. This indicated that LncRNAs participated in the immune disorder of pSS. Meanwhile, the analysis of the down-regulated mRNAs also enriched in immune function and metabolic function of PBMCs,

indicating that LncRNAs were related to the regulation of the immunity and metastasis of PBMCs; thus, contributing to the pathogenesis of pSS.

Eleven selected LncRNAs were involved in the RT-qPCR validation stage. Among them, LINC00426, TPTEP1-202, CYTOR, NRIR, and BISPR were conformed to be aberrantly expressed. These differentially expressed LncRNAs were strongly related to the ESSDAI and serum IgG, CRP, and C4 (The correlation of CRP and C4 seemed to be affected by the particularly high outlier samples and were still needed the larger sample size study), which were closely correlated with disease activity of pSS. Thus, these differentially expressed LncRNAs were thought to affect pSS disease activity, which would be further verified in the subsequent experimental study. It was noted that the NRIR and BISPR were significantly up-regulated in the pSS patients as compared to the healthy controls. LncRNA-NRIR was located closely to the protein-coding interferon stimulated genes (ISGs), and CMPK2 was strongly correlated to the JAK-STAT signaling pathway, which acts as a negative regulator of IFN response[29]. It was reported that the knockdown of the LncRNA resulted in an increase in the Type I IFN-stimulated transcription. While BISPR was a BST2 (IFN-α stimulation) promotor-sharing LncRNA[30] that was dependent on the JAK-STAT signaling pathway, overexpression of BISPR could up-regulate BST2 and induce the release of IFN-α. The activation of Type I IFN pathway has been found to play a vital role in the pathogenesis of pSS patients before[31,32]. Various factors like viral infection or immunity disorders can promote the release of type I IFN in pSS patients and activate the immune cells to release the inflammatory cytokines, antigen presentation, and production of autoantibodies, which was consistent with our LncRNA sequencing results. The differentially expressed LncRNA-NRIR and BISPR also attached importance to the Type I IFN signaling pathway in the pathogenesis of pSS. Wang et al. in their study found that LncRNA-TMEVPG1, which could regulate the release of IFN-γ, elevated in the pSS patients' CD4+ T cells[33] and were related to patients Th1 cells, ESR, serum IgG and anti-SSA antibody, this results just similar to the result of our study that the LncRNAs, especially the IFN related LncRNAs, were significantly differently expressed in the pSS patients PBMCs. While Some novel LncRNAs were also found differentially expressed in the sequencing study, like XLOC_167004, XLOC_102167, and XLOC_078623, their fold changes showed that they were highly expressed and their potential function could be verified in the future experiments.

LncRNAs could regulate the function of adjacent or correlated mRNAs. Therefore, in our study, we used the co-expression and co-localization analysis between the conformed differentially expressed LncRNAs and mRNAs to elaborate the potential mechanism of the LncRNAs in PBMCs of pSS patients. The results revealed that many immune-related genes were involved in PBMCs of pSS. HMGB1, which is located downstream of LINC00426, is a DNA-binding nuclear protein that has been implicated in some inflammatory disorders[34,35]. It could activate the innate immune cells (macrophages/monocytes) through interaction with TLR-2/4 and induce the release of cytokines like IL-8 through the NF-κB or MAPK signaling pathways. RASD2 (located upstream of NRIR) and IFI44 (correlated with NRIR) were reported to be the IFN inducers dependent on the JAK-STAT signaling pathway[36], which can activate the innate immune cells and promote the release of IFN. CMPK2 (located downstream of NRIR) was thought to be related with the differentiation of monocytes and the release of I-IFN through the JAK-STAT signaling pathway[37]. BST2 (located upstream of BISPR) was reported to promote the innate immune release of inflammatory cytokines through the NF-κB signaling pathways[38]. NF1 (correlated with NRIR) could promote the development and differentiation of the immune cells through mTOR signaling pathway[39]. These immune cells, inflammatory cytokines, and signaling pathways were reported to be important in the pathogenesis of pSS. LOXL2 (correlated with LINC00426), CALCOCO1 (correlated with LINC00426), CRIP2 (correlated with TPTEP1-202), HSD17B11 (correlated with CYTOR), and SLC27A1 (upstream of BISPR) could regulate the metastasis of the PBMCs to regulate the corresponding cell function and the release of cytokines[40–44], which could provide new insights into the pathogenesis of pSS. LncRNA-PVT1 participated in the regulation of cell metabolism through reprogramming glycolysis in CD4+ T cells to regulate immunity have been reported recently[45], and similar to our results, LncRNAs regulating immunity through cell metabolism provided new direction for the pathogenesis of pSS. The results of the co-localization and co-expression analysis indicated that LncRNAs which could regulate the immune response of PBMCs were involved in the pSS disease progression. However, the mechanism of LncRNA still needs further studies, and animal models should also be used to further confirm the function of LncRNAs.

References

1. Bowman, S. J. Primary Sjogren's syndrome. *Lupus* **27**, 32–35. https://doi.org/10.1177/0961203318801673 (2018).
2. Fox, R. I. Sjögren's syndrome. *Lancet* **366**(9482), 321–331. https://doi.org/10.1016/S0140-6736(05)66990-5 (2005).
3. Brito-Zerón, P. *et al.* Influence of geolocation and ethnicity on the phenotypic expression of primary Sjögren's syndrome at diagnosis in 8310 patients: A cross-sectional study from the Big Data Sjögren Project Consortium. *Ann. Rheum. Dis.* **76**, 1042–1050. https://doi.org/10.1136/annrheumdis-2016-209952 (2017).
4. Kau, C. K. *et al.* Primary Sjögren's syndrome complicated with cryoglobulinemic glomerulonephritis, myocarditis, and multi-organ involvement. *J. Formosan Med. Assoc.* **103**, 707–710 (2004).
5. Ferro, F., Marcucci, E., Orlandi, M., Baldini, C. & Bartoloni-Bocci, E. One year in review 2017: Primary Sjögren's syndrome. *Clin. Exp. Rheumatol.* **35**, 179–191 (2017).
6. Fitzgerald, K. A. & Caffrey, D. R. Long noncoding RNAs in innate and adaptive immunity. *Curr. Opin. Immunol.* **26**, 140–146. https://doi.org/10.1016/j.coi.2013.12.001 (2014).
7. Bhat, S. A. *et al.* Long non-coding RNAs: Mechanism of action and functional utility. *Non-Coding RNA Res.* **1**, 43–50. https://doi.org/10.1016/j.ncrna.2016.11.002 (2016).

8. Engreitz, J. M., Ollikainen, N. & Guttman, M. Long non-coding RNAs: Spatial amplifiers that control nuclear structure and gene expression. *Nat. Rev. Mol. Cell Biol.* **17**, 756–770. https://doi.org/10.1038/nrm.2016.126 (2016).
9. Carpenter, S. *et al.* A long noncoding RNA mediates both activation and repression of immune response genes. *Science (New York, N.Y.)* **341**, 789–792. https://doi.org/10.1126/science.1240925 (2013).
10. Rapicavoli, N. A. *et al.* A mammalian pseudogene lncRNA at the interface of inflammation and anti-inflammatory therapeutics. *eLife* **2**, e00762. https://doi.org/10.7554/eLife.00762 (2013).
11. Li, Z. *et al.* The long noncoding RNA THRIL regulates TNFalpha expression through its interaction with hnRNPL. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 1002–1007. https://doi.org/10.1073/pnas.1313768111 (2014).
12. Hu, G. *et al.* Expression and regulation of intergenic long noncoding RNAs during T cell development and differentiation. *Nat. Immunol.* **14**, 1190–1198. https://doi.org/10.1038/ni.2712 (2013).
13. Ranzani, V. *et al.* The long intergenic noncoding RNA landscape of human lymphocytes highlights the regulation of T cell differentiation by linc-MAF-4. *Nat. Immunol.* **16**, 318–325. https://doi.org/10.1038/ni.3093 (2015).
14. Zhang, F. *et al.* Identification of the long noncoding RNA NEAT1 as a novel inflammatory regulator acting through MAPK pathway in human lupus. *J. Autoimmunol.* **75**, 96–104. https://doi.org/10.1016/j.jaut.2016.07.012 (2016).
15. Plenge, R. M. *et al.* TRAF1-C5 as a risk locus for rheumatoid arthritis—A genomewide study. *New Engl. J. Med.* **357**, 1199–1209. https://doi.org/10.1056/NEJMoa073491 (2007).
16. Sayad, A., Taheri, M., Arsang-Jang, S., Glassy, M. C. & Ghafouri-Fard, S. Hepatocellular carcinoma up-regulated long non-coding RNA: A putative marker in multiple sclerosis. *Metab. Brain Dis.* **34**, 1201–1205 (2019).
17. Fu, J. *et al.* LncRNA PVT1 links Myc to glycolytic metabolism upon CD4(+) T cell activation and Sjögren's syndrome-like autoimmune response. *J. Autoimmunol.* https://doi.org/10.1016/j.jaut.2019.102358 (2019).
18. Shi, H. *et al.* Long non-coding RNA expression profile in minor salivary gland of primary Sjögren's syndrome. *Arthritis Res. Ther.* **18**, 109–109. https://doi.org/10.1186/s13075-016-1005-2 (2016).
19. Vitali, C. *et al.* Classification criteria for Sjogren's syndrome: A revised version of the European criteria proposed by the American-European Consensus Group. *Ann. Rheum. Dis.* **61**, 554–558. https://doi.org/10.1136/ard.61.6.554 (2002).
20. Seror, R. *et al.* EULAR Sjogren's syndrome disease activity index (ESSDAI): A user guide. *RMD Open* **1**, e000022. https://doi.org/10.1136/rmdopen-2014-000022 (2015).
21. Kanehisa, M. & Goto, S. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* **28**, 27–30 (2000).
22. Kanehisa, M., Sato, Y., Furumichi, M., Morishima, K. & Tanabe, M. New approach for understanding genome variations in KEGG. *Nucleic Acids Res.* **47**, D590–D595 (2019).
23. Quinn, J. J. & Chang, H. Y. Unique features of long non-coding RNA biogenesis and function. *Nat. Rev. Genet.* **17**, 47–62. https://doi.org/10.1038/nrg.2015.10 (2016).
24. Winterling, C. *et al.* Evidence for a crucial role of a host non-coding RNA in influenza A virus replication. *RNA Biol.* **11**, 66–75. https://doi.org/10.4161/rna.27504 (2014).
25. Ouyang, J. *et al.* NRAV, a long noncoding RNA, modulates antiviral responses through suppression of interferon-stimulated gene transcription. *Cell Host Microbe* **16**, 616–626. https://doi.org/10.1016/j.chom.2014.10.001 (2014).
26. Sandhya, P., Joshi, K. & Scaria, V. Long noncoding RNAs could be potential key players in the pathophysiology of Sjogren's syndrome. *Int. J. Rheum. Dis.* **18**, 898–905. https://doi.org/10.1111/1756-185x.12752 (2015).
27. Tzioufas, A. G., Hantoumi, I., Polihronis, M., Xanthou, G. & Moutsopoulos, H. M. Autoantibodies to La/SSB in patients with primary Sjogren's syndrome (pSS) are associated with upregulation of La/SSB mRNA in minor salivary gland biopsies (MSGs). *J. Autoimmunol.* **13**, 429–434. https://doi.org/10.1006/jaut.1999.0333 (1999).
28. Shi, H. *et al.* Long non-coding RNA expression profile in minor salivary gland of primary Sjogren's syndrome. *Arthritis Res. Therapy* **18**, 109. https://doi.org/10.1186/s13075-016-1005-2 (2016).
29. Kambara, H. *et al.* Negative regulation of the interferon response by an interferon-induced long non-coding RNA. *Nucleic Acids Res.* **42**, 10668–10680. https://doi.org/10.1093/nar/gku713 (2014).
30. Kambara, H. *et al.* Regulation of interferon-stimulated gene BST2 by a lncRNA transcribed from a shared bidirectional promoter. *Front. Immunol.* **5**, 676. https://doi.org/10.3389/fimmu.2014.00676 (2014).
31. Vakaloglou, K. M. & Mavragani, C. P. Activation of the type I interferon pathway in primary Sjogren's syndrome: An update. *Curr. Opin. Rheumatol.* **23**, 459–464. https://doi.org/10.1097/BOR.0b013e328349fd30 (2011).
32. Nordmark, G., Alm, G. V. & Ronnblom, L. Mechanisms of disease: Primary Sjogren's syndrome and the type I interferon system. *Nat. Clin. Pract. Rheumatol.* **2**, 262–269. https://doi.org/10.1038/ncprheum0173 (2006).
33. Wang, J. *et al.* Upregulation of long noncoding RNA TMEVPG1 enhances T helper type 1 cell response in patients with Sjögren syndrome. *Immunol. Res.* **64**, 489–496. https://doi.org/10.1007/s12026-015-8715-4 (2016).
34. Yu, M. *et al.* HMGB1 signals through toll-like receptor (TLR) 4 and TLR2. *Shock (Augusta)* **26**, 174–179. https://doi.org/10.1097/01.shk.0000225404.51320.82 (2006).
35. Sims, G. P., Rowe, D. C., Rietdijk, S. T., Herbst, R. & Coyle, A. J. HMGB1 and RAGE in inflammation and cancer. *Annu. Rev. Immunol.* **28**, 367–388. https://doi.org/10.1146/annurev.immunol.021908.132603 (2010).
36. Hallen, L. C. *et al.* Antiproliferative activity of the human IFN-alpha-inducible protein IFI44. *J. Interferon Cytokine Res.* **27**, 675–680. https://doi.org/10.1089/jir.2007.0021 (2007).
37. Valadkhan, S. & Gunawardane, L. S. lncRNA-mediated regulation of the interferon response. *Virus Res.* **212**, 127–136. https://doi.org/10.1016/j.virusres.2015.09.023 (2016).
38. Tokarev, A. *et al.* Stimulation of NF-kappaB activity by the HIV restriction factor BST2. *J. Virol.* **87**, 2046–2057. https://doi.org/10.1128/jvi.02272-12 (2013).
39. Johannessen, C. M. *et al.* The NF1 tumor suppressor critically regulates TSC2 and mTOR. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 8573–8578. https://doi.org/10.1073/pnas.0503224102 (2005).
40. Barker, H. E. *et al.* LOXL2-mediated matrix remodeling in metastasis and mammary gland involution. *Cancer Res.* **71**, 1561–1572. https://doi.org/10.1158/0008-5472.can-10-2868 (2011).
41. Yu, G. *et al.* Long noncoding RNA LINC00052 inhibits colorectal cancer metastasis by sponging microRNA-574-5p to modulate CALCOCO1 expression. *J. Cell. Biochem.* **120**, 17258–17272. https://doi.org/10.1002/jcb.28988 (2019).
42. Zhou, L. *et al.* HOXA9 inhibits HIF-1alpha-mediated glycolysis through interacting with CRIP2 to repress cutaneous squamous cell carcinoma development. *Nat. Commun.* **9**, 1480. https://doi.org/10.1038/s41467-018-03914-5 (2018).
43. Brereton, P. *et al.* Pan1b (17betaHSD11)-enzymatic activity and distribution in the lung. *Mol. Cell. Endocrinol.* **171**, 111–117. https://doi.org/10.1016/s0303-7207(00)00417-2 (2001).
44. Ordovás, L., Roy, R., Zaragoza, P. & Rodellar, C. Structural and functional characterization of the bovine solute carrier family 27 member 1 (SLC27A1) gene. *Cytogenet. Genome Res.* **115**, 115–122. https://doi.org/10.1159/000095230 (2006).
45. Fu, J. *et al.* LncRNA PVT1 links Myc to glycolytic metabolism upon CD4(+) T cell activation and Sjögren's syndrome-like autoimmune response. *J. Autoimmunol.* **107**, 102358. https://doi.org/10.1016/j.jaut.2019.102358 (2020).

Acknowledgements

Author contributions

Y.P. and X.L. designed the research, analysed data and wrote the manuscript. Y.C. helped in inputting patient data, L.P., Ch.D., W.Zh., Y.F. and Y.Zh. collected data, analysed data and revised the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41598-020-76701-2.

Correspondence and requests for materials should be addressed to Y.F., W.Z. or Y.Z.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.