Identification of Potential Biomarkers in PBMC of Systemic Lupus Erythematosus: Results from Bioinformatic Analysis

Yan Sun
  Huazhong University of Science and Technology Tongji Medical College
Chen-chen Wang
  Huazhong University of Science and Technology Tongji Medical College
Fu-quan Wang
  Huazhong University of Science and Technology Tongji Medical College
Rui Chen
  Huazhong University of Science and Technology Tongji Medical College
Chun-lin Yao
  Huazhong University of Science and Technology Tongji Medical College
Yan Wu
  Wuhan Union Hospital
Yun Lin (✉ franklinyun@hust.edu.cn)
  Wuhan Union Hospital   https://orcid.org/0000-0001-9705-6972

Abstract

Background

The discovery of biomarkers has become an attractive field in studying autoimmune diseases. For example, in the study of systemic lupus erythematosus (SLE), various biomarkers such as genes and miRNAs have been identified for the diagnosis of SLE and its organ involvement.

Results

The expression data of gene microarray GSE50772 was downloaded from the GEO, and 257 differentially expressed genes (DEGs) were obtained by using limma plug-in for R software. The tissue-specific gene expression analyses were performed in BioGPS database. Then, a protein-protein interaction (PPI) network was constructed with STRING and visualized in Cytoscape. Whereafter, top twenty hub genes derived from the PPI network, could basically differentiate the SLE samples from the non-SLE samples, were ascertained through CytoHubba. What is noticeable is that the five novel hub genes (ORM1, SLPI, OLFM4, TCN1 and CRISP3) and a related miRNA (hsa-let-7e-5p) may be considered as candidate biomarkers of SLE.

Conclusions

Five genes (ORM1, SLPI, OLFM4, TCN1 and CRISP3) and a miRNA(hsa-let-7e-5p) in this discovery-driven study may become potential biomarkers for diagnosing SLE and assessing its organ damage, and they also will provide valuable information on the pathogenesis of SLE.

Background

SLE is one of the most prevalent autoimmune diseases and often causes tremendous sufferings to patients. The risk of morbidity and mortality of SLE patients is still significantly high[1, 2]. Thus, making a confirmed diagnosis early of SLE has always been essential for initiating the appropriate therapy, but there is no single clinical symptoms or lab abnormality for diagnosing lupus definitely[3], even if plenty of biomarkers for diagnostic use have been found, such as antinuclear antibodies (ANAs), in particular anti-dsDNA antibodies, anti-Sm antibodies, as well as SLE-associated loci and genes, miRNAs, and other molecules[4–6]. As we all know, SLE is characterized by the protean clinical course and a broad spectrum of organ or system manifestations[7, 8], which has an important impact on prognosis of patients[9]. But the manifestations are usually non-specific at onset, making it easy to confuse lupus with a variety of other diseases [10]. Therefore, discovering more biomarkers with relatively high specificity and sensitivity is one of the most crucial and urgent problems for auxiliary diagnosis of SLE.

Over the years, with the wide application of DNA microarray technology and bioinformatics analysis, genetics and epigenetics have attracted extensive attention in SLE researches, especially the expression levels of genes and miRNAs acting as biomarkers [11, 12]. Previous studies have declared that some genes are susceptible to SLE, such as IRF7, STAT4, BLK, etc [13, 14], and many genes even have been regarded as good biomarkers for diagnosing SLE, like OASL, ISG15 MX1, etc [15]. Most expression products of SLE-associated gene participate in immune response [16], and many genes are also related with damaged target organs. In addition, as critical regulators in regulating post-transcriptional target gene expression, miRNAs can interrupt intercellular signal pathways, perturb immune homeostasis and produce autoantibodies, and eventually trigger the occurrence of autoimmune responses [17–20]. Strong evidences for the correlation between dysregulated miRNAs and the pathogenesis and adverse complications of SLE have been provided by published literatures [21–25].

Bioinformatics faces huge-volume heterogeneous biological data [26], including fundamental biology and the biology that underlies disease [27]. During identifying biomarkers for SLE diagnosis, large datasets, through bioinformatic analysis, can be obtained to screen out virtual genetic or epigenetic alternations. In this paper, data quality analysis was performed on the GSE50772, which was downloaded from the GEO public database. In order to identify DEGs, gene expression data of SLE patients and normal controls were extracted by the limma package of R software. Furthermore, DEGs were analyzed for tissue-specific gene expression and the identification of hub genes by respectively using BioGPS and CytoHubba. Subsequently, the functional enrichment of DEGs was analyzed by DAVID, the PPI network of DEGs was constructed through STRING. Moreover, by means of starBase v2.0, the genes, which were involved in the top 10 biological processes with statistical significance, were selected to make miRNAs prediction and gene-miRNA interaction network analysis. Our results will provide new biological information to improve the understanding of the pathogenesis of SLE, and novel biomarkers may be helpful for diagnosiof the disease in early time.

Materials And Methods

Microarray data

The Gene Expression Omnibus Database (GEO, https://www.ncbi.nlm.nih.gov/geo/) is an international public repository for the collection and distribution of high-throughput microarrays and next-generation sequenced functional genomic data sets [28]. The microarray expression dataset GSE50772, uploaded by Kennedy and Maciuca et al., was retrieved and downloaded from the GEO. The selected species was Homo sapiens, the type of data was microarray expression profiles, and the dataset was based on the GPL570 [HG-U133_Plus_2] Affymetrix Human Genome U133 Plus 2.0 Array platform. This research contains 81 samples consisting of 61 subjects with SLE and 20 healthy controls. In addition, the annotation file for GPL570 was also obtained from the GEO.

Estimation of the RNA quality of samples and preprocessing

RNA degradation, proceeding from the 5' end to the 3' end, plays an crucial role in modulating gene expression and correcting systematic biases [29, 30]. RNA degradation measurement presents the best correlation of the RNA integrity number (RIN) and an independent RNA integrity measurement, and therefore is

able to be a valuable tool for quality control [31]. We used Affy package and affyPLM package of R software (R Foundation for Statistical Computing, Vienna, Austria) to estimate the quality of GSE50772 dataset, and the RNA degradation plot was to display the results of the analysis. RMA and KNN methods were used to preprocess the data of each sample in the dataset.

Differential expression analysis

R software was used to normalize and process the original expression matrix, and DEGs were screened via the limma package. The P- values were calculated by adopting the T- test methods, and the adjusted P- values were computed by applying the Benjamini and Hochberg's method. The DEGs were screened out by the following selection criteria: 1) | log2 (fold-change) | >1, and 2) the adjusted P-values < 0.05. The heatmap and volcano map for the DEGs were created by SangerBox software (http://sangerbox.com/).

Tissue-specific gene expression analysis

We analyzed the tissue specific expression of the DEGs by the online resource BioGPS (http://biogps.org). If two following criteria were satisfied, transcripts mapped to the single tissue would be identified as highly tissue specific: 1) The tissue-specific expression of the transcripts was 10 times higher than its median level, 2) The second highest expression level was lower than 1/3 of the highest expression level [32].

Functional enrichment analysis of DEGs

We used Database for annotation, visualization and integrated discovery (DAVID) v6.8 (https://david.ncifcrf.gov/tools.jsp) to conduct the functional enrichment analyses of DEGs, including Gene Ontology (GO) terms and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways. The GO, which are used to predict protein functions, includes cell composition (CC), molecular function (MF) and biological process (BP) [33]. The KEGG pathway analysis, which is used to allot a series of DEGs on specific pathways, constructs the molecular reaction, interaction and relationship [34]. Thus, we conducted pathway analysis to identify which key pathways might be associated with DEGs. In addition, P- values < 0.05 and enriched gene count > 5 were chosen as the criteria for significance.

Protein-protein interaction (PPI) network analysis and hub genes identification.

The DEGs were uploaded to STRING(https://string-db.org/) to produce the PPI network diagram. These protein-protein interactions involve both functional and physical connections, with data derived primarily from high-throughput experiments, computational predictions, co-expression networks and automated text mining. In addition, the PPI network constructed from STRING analysis was imported into Cytoscape v.3.8.0 software, thereby to make a visual design of PPI network. And CytoHubba was used to process the network data to identify the top 20 hub genes. Subsequently, immune-related hub genes were identified by intersection between top 20 hub genes and immune genes, which were downloaded from the Immport immune database (https://www.immport.org/). The GO analyses for these hub genes were performed on Metascape (https://metascape.org/).

Prediction of pivotal miRNAs and identification of hub genes

Based on the results of functional enrichment analysis of DEGs, genes enriched in top 10 statistically significant biological processes were selected and then performed with miRWalk 2.0 software (http://mirwalk.umm.uni-heidelberg.de/), then their targeted miRNAs were further predicted [35]. Parallelly, in order to verify the accuracy of the results, we used miRWalk, miRDB and TargetScan for the intersection. Therefore, miRNAs targeting at more than two genes were screened out [36]. Furthermore, the GeneCards database (https://www.genecards.org/) was used to identify hub genes which were found to serve as candidate biomarkers in our study.

Results

Study resign

The workflow chart of our study design is shown in Fig. 1. Our original goal is to identify more useful biomarkers involved in SLE pathogenesis. Above all, RNA quality analysis was performed on the GSE50772 dataset, which was downloaded from the GEO public database. Extracted from the limma package of R software, gene expression datas of SLE patients and normal controls were used to filter DEGs. Then, the selected DEGs were analyzed for tissue-specific gene expression and the functional enrichment. The PPI network of DEGs was constructed followed, and through which, the top 20 hub genes were identified. In addition, miRNAs prediction and gene-miRNA interaction network analysis were performed on the genes which involved in the top 10 biological processes with statistical significance. Finally, the GeneCards database was used to verify SLE - related hub genes.

Assessment of the RNA quality of samples

To assess the quality of the available data of GSE50772 in SLE, We applied Affy package and affyPLM package of R software to draw the RNA degradation plot for all sample arrays. In Fig. 2, the analysis result demonstrates that 81 nearly parallel curves representing 81 samples have appropriate slopes, which doesn't display the bias of probe-positional intensity associated with RNA integrity, indicating that the RNA quality of samples is relatively ideal.

Differentially expressed genes

$|Log_2FC|$ greater than 1 and adjusted P-values less than 0.05 were considered as criteria to screen the DEGs out. A total of 257 DEGs are obtained, among which 227 are down-regulated and 30 up-regulated. As shown in Table 1, the RPS4Y1, EIF1AY and KDM5D are the most up-regulated. Similarly, the three most down-regulated genes are CXCL8, ANXA3 and IFI27, whose logᵢFC values are − 3.675, -3.186 and − 3.173 respectively (Additional file 1). The volcano plot and heatmap of the DEGs are seen in Fig. 3.

Table 1
Top 10 up-regulated and down-regulated genes.

Gene	logFC †	Adjusted P-value
Up-regulated		
RPS4Y1	2.539	2.09E-07
EIF1AY	2.179	5.96E-07
KDM5D	1.800	3.07E-09
USP9Y	1.711	6.81E-08
KLRC4	1.543	6.22E-14
GCSAML	1.542	5.68E-16
ZNF850	1.476	7.35E-31
PDCD4	1.365	3.27E-19
ZNF566	1.325	7.51E-17
FAM169A	1.315	1.31E-19
Down-regulated		
CXCL8	-3.675	3.99E-13
ANXA3	-3.186	1.85E-12
IFI27	-3.173	5.71E-08
OLFM4	-2.922	8.10E-06
IL1R2	-2.886	6.19E-09
CXCL1	-2.841	1.32E-19
MMP9	-2.789	2.70E-12
EGR1	-2.689	1.30E-13
G0S2	-2.635	3.06E-10
S100P	-2.628	1.62E-12

† logFC: log-transformed fold change of gene expression.

Tissue-specific expression of genes

We used the BioGPS database, an online tool, to identify 57 genes expressed in specific tissues or organ systems. As shown in Table 2, the system with the most highly tissue-specific expression is the hematologic/immune system (59.6%, 34/57), followed by the digestive system (15.8%, 9/57). The respiratory and skin/skeletal muscle systems have similar levels of enrichment (about 8.8%, 5/57), while the urinary and reproductive systems have the lowest enrichment levels (about 3.5%, 2/57).

Table 2
Tissue-specific expressed genes identified by BioGPS.

System	Genes
Hematologic/immune	CAMP,ELANE,MMP8,DEFA4,CEACAM8,RSAD2,IFIT1, FFAR2,CXCR2, RETN,SLC25A37,CXCR1, SELENBP1, MZB1,FPR2,SUCNR1,SLC22A4, CMPK2,IFIT2,FOSB, HBM,AQP9,FPR1,PPP1R15A,MXD1, CCL4,NFE2, AZU1, KCNJ2,MGAM,RNASE2,RNASE3,SH2D1B,EIF1AY
Skin/skeletal muscle	CXCL1,IL1B,CCL2,TNNT1,CXCL3
Respiratory	TCN1,ANXA3,C1QC,AREG,COL17A1
Digestive	OLFM4,HP,ORM1,LRG1,ARG1,CYP4F3,ANG,C15orf48, KRT23
Urinary	TCN2,ALOX15B
Reproductive	ADM,S100P

Functional and pathway enrichment of DEGs

To investigate the biological function of the DEGs, we performed functional enrichment analyses of 257 DEGs, GO annotation and KEGG pathway analyses were conducted by using the DAVID online tool. The top 10 GO items, including BPs, CCs and MFs, and the top 10 KEGG pathways, are listed in Fig. 4A-4D. Moreover, the top 10 biological processes of DEGs and their related details are shown in Table 3, among which biological pathways with P-value < 0.05 are statistically significant. The results suggest that the biological pathways with DEGs significantly enriching are immune system−related pathways. Furthermore, the extracellular exosome, cytosol and extracellular space, and integral component of plasma membrane account for the majority of CC. And the most abundant MFs are protein binding. KEGG pathway analysis reveals that the DEGs mainly enrich in Cytokine-cytokine receptor interaction, TNF signaling pathway and Chemokine signaling pathway (Fig. 4D).

Table 3
The top10 biological process (BP) of DEGs.

Term	Count	P-Value	Genes
inflammatory response	32	1.38E-16	OLR1,C3AR1,PROK2,CXCL8,TNFAIP6,CXCL1,CXCL3,PTGS2,HCK,IL1B,CHI3L1,PTX3,SIGLEC1,TLR2,TNF,CXCL
immune response	25	9.87E-10	IFITM3,CXCL8,AQP9,CXCL1,CXCL3,CXCL2,FCGR3B,CXCL10,NFIL3,CCL4,CCL2,PGLYRP1,FCGR1B,CCR1,CCL2
innate immune response	22	1.64E-07	C1QB,DEFA4,CRISP3,MX2,MX1,SH2D1B,BMX,HERC5,TLR2,IGLL5,CLEC4D,SLPI,GPER1,LCN2,S100A12,PTX3,
G-protein coupled receptor signaling pathway	18	0.070008171	CXCL8,CCL20,FPR1,FPR2,CXCL3,CXCL2,AREG,CCL4,CCL2,CXCL10,HCAR3,CXCR1,GPER1,C3AR1,PROK2,FFA SUCNR1
Type I interferon signaling pathway	15	5.67E-14	IFITM3,EGR1,RSAD2,MX2,MX1,IFI6,ISG15,IFI35,IFIT1,IFIT3,IFIT2,OASL,IFI27,OAS1, OAS3
chemotaxis	15	5.27E-10	CCR1,CXCL8,CCL20,FPR1,CXCL1,FPR2,RNASE2,CXCL2,CXCL10,CXCR1,CXCR2,C3AR1,PROK2,CCL2,CMTM2
defense response to virus	15	2.75E-08	IFITM3,RSAD2,MX2,MX1,ISG15,AZU1,IFIT1,IFIT3,OASL,IFIT2,HERC5,CXCL10,OAS1,OAS3,IFI44L
positive regulation of cell proliferation	15	0.002819349	IRS2,ADM,AREG,CDC20,CXCL10,HCK,HLX,CEACAM6,NAMPT,CXCR2,MZB1,PROK2,PRTN3,CAMP,GPER1
response to virus	14	1.54E-09	IFITM3,RSAD2,MX2,MX1,IFI44,IFIT1,TNF,IFIT3,IFIT2,OASL,OAS1,OAS3,CCL4,LCN2
response to lipopolysaccharide	13	1.36E-06	JUN,CXCL1,FOS,CXCL3,PTGS2,SOD2,MPO,CXCL2,CXCL10,SLPI,ELANE,TLR2,ADM

PPI network construction, hub genes selection and analysis

The 257 DEGs were inputted to the STRING tool for further analysis, and a PPI network with 224 nodes and 1485 edges were visualized with Cytoscape. The interaction score of PPI network is greater than 0.4. The nodes correspond to genes, and the edges represent the links between genes. Green nodes represent down-regulated genes, red nodes represent up-regulated genes. The local clustering coefficient is 0.532 and PPI enrichment P-value is less than 1.0e-16. Then, the data file was processed with Cytoscape (Fig. 5A). CytoHubba was used to process the network data, and then to identify hub genes, the top 20 hub genes (CXCL1, CAMP, HP, PTX3, ARG1, ELANE, LCN2, RETN, MMP8, SLPI, PGLYRP1, LTF, OLFM4, ORM1, TCN1, LRG1, CRISP3, CHI3L1, MMP9 and DEFA4) were identified (Fig. 5B). The color of a node in the network reflects the rank of hub genes. Clustering shows that the hub genes could basically differentiate the SLE samples from the non-SLE samples. Most hub genes are highly expressed in SLE samples, while relatively low in non-SLE samples (Fig. 5C). In addition, functional enrichment analysis indicates that these hub genes mainly enrich in immune system process as shown in Fig. 5D. In order to further explore and confirm the nature of hub genes, as shown in the supplementary materials for this study, 2482 immune genes downloaded from the Immport immune database were used to intersect with the top 20 hub genes. As expected, we found that 13 hub genes, including CXCL1, CAMP, PTX3, ARG1, ELANE, LCN2, RETN, SLPI, PGLYRP1, LTF, ORM1, MMP9 and DEFA4, were immune-related genes (Additional file 2).

Further miRNA mining and identification of key genes

Among the top 10 biological processes, eighty-six genes, associated with statistically significant biological processes, were selected, and the gene-miRNA analysis was conducted with miRWalk 2.0 software. The intersection of miRNA results predicted by miRWalk, TargetScan and miRDB databases was considered as the result. The selection condition was set as P-value < 0.05, the target gene binding region was 3'UTR. Therefore, hsa-let-7e-5p with high number of gene cross-links (≥ 2) is identified, it targets at OLR1 and IRS2 as shown in Fig. 6. The score of it is 1, which means that it has high reliability.

In addition, using the GeneCards database, SLE - related genes were manually identified (Additional file 3), and three of our novel hub genes (ORM1, SLPI and TCN1) were verified to be potentially involved in the pathogenesis of SLE (Fig. 7).

Table 4
Fifteen hub genes that have been reported in previous SLE studies.

Gene symbol	Full name	Role in SLE	References
HP	haptoglobin	display immunosuppressive abilities to consist in the host defence responses to inflammation and infection	[73–75]
RETN, MMP8	Resistin, Matrix metalloprotein-ase-8	associate with the presence of coronary artery calcium and increase vulnerability to atherosclerotic plaque	[76, 77]
ARG1, CXCL1	arginase 1, chemokine with C-X-C motif ligand 1	manipulate type 17 T helper cells (Th17) pathway	[78, 79]
CAMP, LTF	Cathelicidin antimicrobial peptide, Lactotransferr-in	estrogen exerts powerful effects on the immune response by affecting the expression of CAMP and LTF in B cells	[80]
DEFA4	Defensin alpha 4	strongly link to immune function and numerous autoimmune diseases	[80]
PTX3	Pentraxin 3	a biomarker or therapeutic target of SLE	[81]
ELANE	Elastase	participate into end-organ damage	[82]
LCN2	Lipocalin 2	a nephritis-associated inflammatory mediator	[83]
LRG1	Leucine-rich alpha-2-glycoprotein 1	reflect specific pathologic lesions in kidney and activity of lupus nephritis	[84]
PGLYRP1	Peptidoglycan recognition protein 1	might perturb the cytotoxic effect of autoantibodies, reduce tissue injury by competitively binding with autoantibodies	[85]
CHI3L1	Chitinase 3 like 1	evaluate the activity of lupus	[86]
MMP9	Matrix metallopeptid-ase 9	participate in pathways and immune system responses associated with SLE	[87, 88]

Discussion

SLE is one of the most common systemic autoimmune diseases that seriously endangers human health. There are many factors causing the pathogenesis of SLE, among which the abnormal expression of important genes may contribute to lupus pathogenesis by participating in critical pathways, including immune complex processing, type I interferon producing, toll-like receptor signaling, and so on [37]. However, the accurate mechanism of SLE caused these microenvironmental factors has not been completely elucidated, so more attention should be paid to the detection and evaluation of the expression level of lupus - related genes in lupus researches. In present study, we are committed to discover possible SLE - causing molecules, and 257 DEGs (30 up-regulated genes and 227 down-regulated genes) are screened out from GSE50772 expression dataset. They certainly has laid a foundation for our subsequent analyses, and that may be able to illuminate the initiation and progression of SLE.

Considering that multiple organs or systems involvement caused by autoantibodies is the feature of SLE, we performed tissue - specific expression analysis on DEGs. As revealed by the result, the most highly enriched system is the hematologic/immune system, which is in line with the pathogenesis and clinical manifestations of SLE, and this seems to explain the underlying molecular mechanisms of a self-aimed immune response in SLE patients. Besides, skin/skeletal muscle system, respiratory system, digestive system, urinary and reproductive systems are also enriched by DEGs. Some studies have also suggested that ANXA3 [38, 39], TCN1 [40], C1QC [41], AREG [42] and COL17A1 [43] are associated with respiratory injury. ALOX15B [44] may be related to kidney diseases as reported in the literature. ADM [45] and S100P [46] possibly play important roles in lesions of the reproductive system. And changes in expression of OLFM4 [47], CYP4F3 [48], HP [49], ORM1 [50], ANG [51], LRG1 [52], C15orf48 [53], KRT23 [54] and ARG1 [55] are involved with digestive system diseases. Even some of them have been thought to be classic markers of a particular tissue injury. While whether those tissue - specific DEGs above-mentioned are essential for the development of complications of SLE remains inconclusive, and we postulate that the abnormal expression of those genes probably can indicate organ involvement in SLE patients. However, in our results, there are other relatively common organs and tissues that are not significantly enriched by DEGs, such as the central nervous system, cardiovascular and circulatory systems. Limited gene expression microarray data with insufficient samples may be a by-no-means negligible cause.

In order to understand disease machinery more deeply and to visualize the overview of the functional connections between all DEGs, we constructed PPI networks, the vital tools for analysis by identifying subnetworks or modules that display specific topology and/or functional characteristics [56]. Afterwards, on the basis of DEGs' PPI networks, the top 20 hub genes (CXCL1, CAMP, HP, PTX3, ARG1, ELANE, LCN2, RETN, MMP8, SLPI, PGLYRP1, LTF, OLFM4, ORM1, TCN1, LRG1, CRISP3, CHI3L1, MMP9 and DEFA4) were selected. According to cluster analysis results on them, it is obvious that most hub genes are up-expressed in SLE patients, while relatively low-expressed in normal subjects, which highlights the importance and representativeness of these hub genes in SLE disease. And further functional enrichment analysis on them manifests that immune system processes are dominant. Furthermore, 13 hub genes verified by Immport database are thought to be immune-related genes as expected, namely CXCL1, CAMP, PTX3, ARG1, ELANE, LCN2, RETN, SLPI, PGLYRP1, LTF, ORM1, MMP9 and DEFA4. The result exactly supports the idea that these 20 hub genes probably play essential roles in immune-related pathways which can trigger autoimmune dysfunction in patients, and resulting in the pathogenesis and development of SLE.

In prior studies on 20 hub genes mentioned above, the aberrant expression levels of 15 genes have been investigated that they may have various and crucial influences for different processes of SLE development. Given the roles of five novel genes in SLE as shown in Table 4, several novel genes, including ORM1, SLPI, OLFM4, TCN1 and CRISP3 may also have diagnostic value in the condition. ORM1 (orosomucoid 1) is an acute phase plasma protein known to activate NFκB, p38 and JNK pathways in macrophages, and it has been reported in rheumatoid arthritis (RA) [57], sarcoidosis and other immune diseases [58]. In experimental autoimmune encephalomyelitis, SLPI (secretory leukocyte peptidase inhibitor) exerted potent pro-inflammatory actions by regulating T cell activity, a process that might benefit the patient [59]. OLFM4(olfactomedin 4) could mediate the autoimmune inflammatory responses of generalized pustular psoriasis, a severe inflammatory skin disease [60]. Low expression of TCNI (transcobalamin I) involved in innate immunity might be partly responsible for the pathogenesis of IgG4-related disease, due to impairments in the innate immune system [61]. Since CRISP3 (cysteine-rich secretory protein 3) was detected to be significantly elevated in RA, the researchers hypothesized that it was implicated in the development of RA [62]. In addition, using the GeneCards database, three novel hub genes (ORM1, SLPI, and TCN1) were confirmed to be potentially involved in the pathogenesis of SLE. Almost all of these genes, either high or low expression, are associated in the development of immune diseases. Consequently, chances are that the five hub genes play pivotal roles in the molecular mechanism of SLE pathogenesis, and we reasonably confer that these novel hub genes may be used as biomarkers to help improve the diagnostic rate of SLE and to provide valuable information for the evaluation of organ or system involvement.It is well known that miRNAs can interfere with the transcription and regulate gene expression [63]. Altered miRNA expression has been regarded as another important factor to the pathogenesis of immune-related diseases, such as SLE. And because of the nature of stability of miRNA, measuring effective miRNA levels may be conducive to disease detectionin [64]. In immune cells, aberrant miRNAs can disturb immune homeostasis, produce massive autoantibodies and induce autoimmunity [65]. Following GO terms, we performed miRNA mining and interaction network analysis. MiRNA hsa-let-7e-5p targeting at OLR1(oxidized low- density lipoprotein receptor 1) and IRS2 (insulin receptor substrate 2) was identified. The miRNA let-7e is a member of the let-7 family, and it plays a key role in inhibiting or promoting inflammatory response by regulating cytokine expression in various inflammatory and autoimmune diseases [66]. In an animal experiment, down-regulating the expression of hsa-let-7e-5p and other two miRNAs, 17β-estradiol could amplify the activation of IFN-α signaling in B cells to contribute to the sex bias in SLE [67−69]. Moreover, as target genes of has-let-7e-5p in this study, OLR1 and IRS2 are down-regulated in SLE patients compared with non-SLE subjects. When it comes to biological processes, OLR1 is associated with inflammatory response, while ISR2 is related to positive regulation of cell proliferation. Regarding molecular function, OLR1 and IRS2 both participate in exerting protein binding. Recent studies show that OLR1 is an inflammation-induced receptor. Together with a host of other reactions, an increase in OLR1 can trigger the formation of neutrophil extracellular traps, which can promote systemic inflammation, vascular damage and lung injury. Elevated expression level of OLR1 has been recognized as a possible indicator of high risk of SLE - related cardiovascular disease [70, 71]. Targeted by MiR-203a, IRS2 regulates the proliferation and apoptosis of pancreatic β cell [72], which implies the expression of IRS2 is closely related to type 1 diabetes mellitus (T1DM), an autoimmune disease. These results rend us to speculate that has-let-7e-5p may be a potential molecule to induce and deteriorate the SLE even though there have been rare relevant studies are published on this subject. Thus, we propose that hsa-let-7e-5p probably acts as another novel latent biomarker of SLE, and we hope it could provide new insights into molecular mechanism underlying the development and progression SLE.

Additionally, since the GSE50772 is a public dataset, patient consent or ethics committee approval is not required, but the information on individuals' age, gender and health status, as well as medication use, is absent, which appears to be an underlying limitation.

Conclusions

In conclusion, some DEGs specifically expressed in a tissue or system might be a signal to estimate organ involvement in SLE, and novel candidate biomarkers including hub genes (ORM1, SLPI, OLFM4, TCN1 and CRISP3) and has-let-7e-5p were identified to assist in diagnosing SLE through comprehensive bioinformatic analyses. Our point will provide new and meaningful reference for later SLE studies. Since the current finding is limited by the lack of experimental validation in vivo and in vitro, it is necessary to conduct multiple in-depth studies to detect and verify those potential biomarkers.

Declarations

Availability of data and materials

The datasets used and analyzed during the current study are available from the corresponding author on reasonable request.

Authors' contributions

YS, CW and FW conceived the study and participated in the study design, performance, coordination and manuscript writing. YS, CW, FW, RC, and CY carried out the analysis. YL and YW revised the manuscript. All authors reviewed and approved the final manuscript.

Ethics approval and consent to participate

Not applicable.

Consent for publication

All the authors have consented for the publication.

Contributor Information

Yan Sun, Email: m201975873@hust.edu.cn.

Chen-chen Wang, Email: chenchen763795289@163.com.

Fu-quan Wang, Email: 544335721@qq.com.

Rui Chen, Email: m202075984@hust.edu.cn.

Chun-lin Yao, Email: m159274454661@163.com.

Yan Wu, Phone: +8602785351606, Email: wuyan5151@hust.edu.cn

Yun Lin, Phone: +8602785351606, Email: franklinyun@hust.edu.cn.

References

1. Frieri M. Mechanisms of disease for the clinician: systemic lupus erythematosus. Ann Allergy Asthma Immunol. 2013;110:228–32.

2. Basta F, Fasola F, Triantafyllias K, et al. Systemic Lupus Erythematosus (SLE) Therapy: The Old and the New. Rheumatol Ther. 2020;7:433–46.

3. Justiz Vaillant AA, Goyal A, Bansal P, et al. Systemic Lupus Erythematosus. StatPearls. Treasure Island (FL)2020.

4. Schiffer L, Worthmann K, Haller H, et al. CXCL13 as a new biomarker of systemic lupus erythematosus and lupus nephritis - from bench to bedside? Clin Exp Immunol. 2015;179:85–9.

5. Liu T, Son M, Diamond B. HMGB1 in Systemic Lupus Erythematosus. Front Immunol. 2020;11:1057.

6. Kim BS, Jung JY, Jeon JY, et al. Circulating hsa-miR-30e-5p, hsa-miR-92a-3p, and hsa-miR-223-3p may be novel biomarkers in systemic lupus erythematosus. HLA. 2016;88:187–93.

7. Kuhn A, Bonsmann G, Anders HJ, et al. The Diagnosis and Treatment of Systemic Lupus Erythematosus. Dtsch Arztebl Int. 2015;112:423–32.

8. Sebastiani GD, Prevete I, Iuliano A, et al. The Importance of an Early Diagnosis in Systemic Lupus Erythematosus. Isr Med Assoc J. 2016;18:212–5.

9. Wang H, Ren YL, Chang J, et al. A Systematic Review and Meta-analysis of Prevalence of Biopsy-Proven Lupus Nephritis. Arch Rheumatol. 2018;33:17–25.

10. Maddison PJ. Is it SLE? Best Pract Res Clin Rheumatol. 2002;16:167–80.

11. International Consortium for Systemic Lupus. Erythematosus G, Harley JB, Alarcon-Riquelme ME, et al. Genome-wide association scan in women with systemic lupus erythematosus identifies susceptibility variants in ITGAM, PXK, KIAA1542 and other loci. Nat Genet 2008;40:204–10.

12. Omidi F, Hosseini SA, Ahmadi A, et al. Discovering the signature of a lupus-related microRNA profile in the Gene Expression Omnibus repository. Lupus. 2020;29:1321–35.

13. Fu Q, Zhao J, Qian X, et al. Association of a functional IRF7 variant with systemic lupus erythematosus. Arthritis Rheum. 2011;63:749–54.

14. Siena S, Villa S, Bonadonna G, et al. Specific ex-vivo depletion of human bone marrow T lymphocytes by an anti-pan-T cell (CD5) ricin A-chain immunotoxin. Transplantation. 1987;43:421–6.

15. Feng X, Huang J, Liu Y, et al. Identification of interferon-inducible genes as diagnostic biomarker for systemic lupus erythematosus. Clin Rheumatol. 2015;34:71–9.

16. Deng Y, Tsao BP. Genetic susceptibility to systemic lupus erythematosus in the genomic era. Nat Rev Rheumatol. 2010;6:683–92.

17. Calixto SM, Mohan C. Lupus genes at the interface of tolerance and autoimmunity. Expert Rev Clin Immunol. 2007;3:603–11.

18. Ou ZL, Luo Z, Wei W, et al. Hypoxia-induced shedding of MICA and HIF1A-mediated immune escape of pancreatic cancer cells from NK cells: role of circ_0000977/miR-153 axis. RNA Biol. 2019;16:1592–603.

19. Ando Y, Yang GX, Kenny TP, et al. Overexpression of microRNA-21 is associated with elevated pro-inflammatory cytokines in dominant-negative TGF-beta receptor type II mouse. J Autoimmun. 2013;41:111–9.

20. Simpson LJ, Ansel KM. MicroRNA regulation of lymphocyte tolerance and autoimmunity. J Clin Invest. 2015;125:2242–9.

21. Saito Y, Saito H, Liang G, et al. Epigenetic alterations and microRNA misexpression in cancer and autoimmune diseases: a critical review. Clin Rev Allergy Immunol. 2014;47:128–35.

22. Hedrich CM, Tsokos GC. Epigenetic mechanisms in systemic lupus erythematosus and other autoimmune diseases. Trends Mol Med. 2011;17:714–24.

23. Zhang H, Huang X, Ye L, et al. B Cell-Related Circulating MicroRNAs With the Potential Value of Biomarkers in the Differential Diagnosis, and Distinguishment Between the Disease Activity and Lupus Nephritis for Systemic Lupus Erythematosus. Front Immunol. 2018;9:1473.

24. Carlsen AL, Schetter AJ, Nielsen CT, et al. Circulating microRNA expression profiles associated with systemic lupus erythematosus. Arthritis Rheum. 2013;65:1324–34.

25. Smith S, Wu PW, Seo JJ, et al. IL-16/miR-125a axis controls neutrophil recruitment in pristane-induced lung inflammation. JCI Insight 2018;3.

26. Chen M, Hofestadt R, Taubert J. Integrative Bioinformatics: History and Future. J Integr Bioinform 2019;16.

27. Rothberg J, Merriman B, Higgs G. Bioinformatics. Introduction Yale J Biol Med. 2012;85:305–8.

28. Barrett T, Wilhite SE, Ledoux P, et al. NCBI GEO: archive for functional genomics data sets–update. Nucleic Acids Res. 2013;41:D991-5.

29. Viljoen KS, Blackburn JM. Quality assessment and data handling methods for Affymetrix Gene 1.0 ST arrays with variable RNA integrity. BMC Genom. 2013;14:14.

30. Tisen X, Xuegui L, Dejie J, et al. [Mechanism of 5'-to-3' degradation of eukaryotic and prokaryotic mRNA]. Yi Chuan. 2015;37:250–8.

31. Fasold M, Binder H. Estimating RNA-quality using GeneChip microarrays. BMC Genom. 2012;13:186.

32. Massingham LJ, Johnson KL, Scholl TM, et al. Amniotic fluid RNA gene expression profiling provides insights into the phenotype of Turner syndrome. Hum Genet. 2014;133:1075–82.

33. Ashburner M, Ball CA, Blake JA, et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet. 2000;25:25–9.

34. Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res. 2000;28:27–30.

35. Dweep H, Sticht C, Pandey P, et al. miRWalk–database: prediction of possible miRNA binding sites by "walking" the genes of three genomes. J Biomed Inform. 2011;44:839–47.

36. Yi XH, Zhang B, Fu YR, et al. STAT1 and its related molecules as potential biomarkers in Mycobacterium tuberculosis infection. J Cell Mol Med. 2020;24:2866–78.

37. Cui Y, Sheng Y, Zhang X. Genetic susceptibility to SLE: recent progress from GWAS. J Autoimmun. 2013;41:25–33.

38. Maghsoudloo M, Azimzadeh Jamalkandi S, Najafi A, et al. Identification of biomarkers in common chronic lung diseases by co-expression networks and drug-target interactions analysis. Mol Med. 2020;26:9.

39. Wu XM, Ji KQ, Wang HY, et al. MicroRNA-339-3p alleviates inflammation and edema and suppresses pulmonary microvascular endothelial cell apoptosis in mice with severe acute pancreatitis-associated acute lung injury by regulating Anxa3 via the Akt/mTOR signaling pathway. J Cell Biochem. 2018;119:6704–14.

40. Li X, Xie Y, Zhang D, et al. Identification of Significant Genes in HIV/TB via Bioinformatics Analysis. Ann Clin Lab Sci. 2020;50:600–10.

41. Shi L, Zhu B, Xu M, et al. Selection of AECOPD-specific immunomodulatory biomarkers by integrating genomics and proteomics with clinical informatics. Cell Biol Toxicol. 2018;34:109–23.

42. Fukumoto J, Harada C, Kawaguchi T, et al. Amphiregulin attenuates bleomycin-induced pneumopathy in mice. Am J Physiol Lung Cell Mol Physiol. 2010;298:L131-8.

43. Jeong I, Lim JH, Park JS, et al. Aging-related changes in the gene expression profile of human lungs. Aging. 2020;12:21391–403.

44. Kugita M, Nishii K, Morita M, et al. Global gene expression profiling in early-stage polycystic kidney disease in the Han:SPRD Cy rat identifies a role for RXR signaling. Am J Physiol Renal Physiol. 2011;300:F177-88.

45. Li L, Tang F, O WS. Coexpression of adrenomedullin and its receptor component proteins in the reproductive system of the rat during gestation. Reprod Biol Endocrinol. 2010;8:130.

46. Zhang SY, Lin XN, Song T, et al. [Gene expression profiles of peri-implantation endometrium in natural and superovulation cycles]. Zhonghua Yi Xue Za Zhi. 2008;88:2343–6.

47. Levinsky NC, Mallela J, Opoka AM, et al. The olfactomedin-4 positive neutrophil has a role in murine intestinal ischemia/reperfusion injury. FASEB J. 2019;33:13660–8.

48. Sen A, Stark H. Role of cytochrome P450 polymorphisms and functions in development of ulcerative colitis. World J Gastroenterol. 2019;25:2846–62.

49. Lisi S, Gamucci O, Vottari T, et al. Obesity-associated hepatosteatosis and impairment of glucose homeostasis are attenuated by haptoglobin deficiency. Diabetes. 2011;60:2496–505.

50. Zhu HZ, Zhou WJ, Wan YF, et al. Downregulation of orosomucoid 2 acts as a prognostic factor associated with cancer-promoting pathways in liver cancer. World J Gastroenterol. 2020;26:804–17.

51. Liu WX, Gu SZ, Zhang S, et al. Angiopoietin and vascular endothelial growth factor expression in colorectal disease models. World J Gastroenterol. 2015;21:2645–50.

52. Zhang J, Zhu L, Fang J, et al. LRG1 modulates epithelial-mesenchymal transition and angiogenesis in colorectal cancer via HIF-1alpha activation. J Exp Clin Cancer Res. 2016;35:29.

53. Zhang GL, Pan LL, Huang T, et al. The transcriptome difference between colorectal tumor and normal tissues revealed by single-cell sequencing. J Cancer. 2019;10:5883–90.

54. Zhang N, Zhang R, Zou K, et al. Keratin 23 promotes telomerase reverse transcriptase expression and human colorectal cancer growth. Cell Death Dis. 2017;8:e2961.

55. Coburn LA, Horst SN, Allaman MM, et al. L-Arginine Availability and Metabolism Is Altered in Ulcerative Colitis. Inflamm Bowel Dis. 2016;22:1847–58.

56. Vella D, Marini S, Vitali F, et al. MTGO: PPI Network Analysis Via Topological and Functional Module Identification. Sci Rep. 2018;8:5499.

57. Higuchi H, Kamimura D, Jiang JJ, et al. Orosomucoid 1 is involved in the development of chronic allograft rejection after kidney transplantation. Int Immunol. 2020;32:335–46.

58. Fan C, Nylander PO, Stendahl U, et al. Synergistic interaction between ORM1 and C3 types in disease associations. Exp Clin Immunogenet. 1995;12:92–5.

59. Muller AM, Jun E, Conlon H, et al. Inhibition of SLPI ameliorates disease activity in experimental autoimmune encephalomyelitis. BMC Neurosci. 2012;13:30.

60. Shao S, Fang H, Zhang J, et al. Neutrophil exosomes enhance the skin autoinflammation in generalized pustular psoriasis via activating keratinocytes. FASEB J. 2019;33:6813–28.

61. Nakamura T, Satoh-Nakamura T, Nakajima A, et al. Impaired expression of innate immunity-related genes in IgG4-related disease: A possible mechanism in the pathogenesis of IgG4-RD. Mod Rheumatol. 2020;30:551–7.

62. Chapman EA, Lyon M, Simpson D, et al. Caught in a Trap? Proteomic Analysis of Neutrophil Extracellular Traps in Rheumatoid Arthritis and Systemic Lupus Erythematosus. Front Immunol. 2019;10:423.

63. Filipowicz W, Bhattacharyya SN, Sonenberg N. Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight? Nat Rev Genet. 2008;9:102–14.

64. Cortez MA, Bueso-Ramos C, Ferdin J, et al. MicroRNAs in body fluids–the mix of hormones and biomarkers. Nat Rev Clin Oncol. 2011;8:467–77.

65. Liu A, La Cava A. Epigenetic dysregulation in systemic lupus erythematosus. Autoimmunity. 2014;47:215–9.

66. Gui L, Zhang Q, Cai Y, et al. Effects of let-7e on LPS-Stimulated THP-1 Cells Assessed by iTRAQ Proteomic Analysis. Proteomics Clin Appl. 2018;12:e1700012.

67. Dong G, Fan H, Yang Y, et al. 17beta-Estradiol enhances the activation of IFN-alpha signaling in B cells by down-regulating the expression of let-7e-5p, miR-98-5p and miR-145a-5p that target IKKepsilon. Biochim Biophys Acta. 2015;1852:1585–98.

68. Korganow AS, Knapp AM, Nehme-Schuster H, et al. Peripheral B cell abnormalities in patients with systemic lupus erythematosus in quiescent phase: decreased memory B cells and membrane CD19 expression. J Autoimmun. 2010;34:426–34.

69. Fortuna G, Brennan MT. Systemic lupus erythematosus: epidemiology, pathophysiology, manifestations, and management. Dent Clin North Am. 2013;57:631–55.

70. Sagar D, Gaddipati R, Ongstad EL, et al. LOX-1: A potential driver of cardiovascular risk in SLE patients. PLoS One. 2020;15:e0229184.

71. Ishikawa M, Ito H, Furu M, et al. Plasma sLOX-1 is a potent biomarker of clinical remission and disease activity in patients with seropositive RA. Mod Rheumatol. 2016;26:696–701.

72. Duan X, Zhao L, Jin W, et al. MicroRNA-203a regulates pancreatic beta cell proliferation and apoptosis by targeting IRS2. Mol Biol Rep. 2020;47:7557–66.

73. Lange V. [Haptoglobin polymorphism–not only a genetic marker]. Anthropol Anz. 1992;50:281–302.

74. Dobryszycka W. Biological functions of haptoglobin–new pieces to an old puzzle. Eur J Clin Chem Clin Biochem. 1997;35:647–54.

75. Marquez L, Shen C, Cleynen I, et al. Effects of haptoglobin polymorphisms and deficiency on susceptibility to inflammatory bowel disease and on severity of murine colitis. Gut. 2012;61:528–34.

76. Chung CP, Solus JF, Oeser A, et al. Genetic variation and coronary atherosclerosis in patients with systemic lupus erythematosus. Lupus. 2014;23:876–80.

77. Nguyen HT, Le VL, Nguyen TMH, et al. Temperature dependence of the dielectric function and critical points of alpha-SnS from 27 to 350 K. Sci Rep. 2020;10:18396.

78. Pan HF, Leng RX, Feng CC, et al. Expression profiles of Th17 pathway related genes in human systemic lupus erythematosus. Mol Biol Rep. 2013;40:391–9.

79. Amorim WW, Passos LC, Oliveira MG. Potentially inappropriate medications in older adults: a commentary on the study by Roux et al. Fam Pract; 2020.

80. Fan H, Dong G, Zhao G, et al. Gender differences of B cell signature in healthy subjects underlie disparities in incidence and course of SLE related to estrogen. J Immunol Res. 2014;2014:814598.

81. Wu Q, Guan SY, Dan YL, et al. Circulating pentraxin-3 levels in patients with systemic lupus erythematosus: a meta-analysis. Biomark Med. 2019;13:1417–27.

82. Maleknia S, Salehi Z, Rezaei Tabar V, et al. An integrative Bayesian network approach to highlight key drivers in systemic lupus erythematosus. Arthritis Res Ther. 2020;22:156.

83. Chalmers SA, Glynn E, Garcia SJ, et al. BTK inhibition ameliorates kidney disease in spontaneous lupus nephritis. Clin Immunol. 2018;197:205–18.

84. Yang Y, Luo R, Cheng Y, et al. Leucine-rich alpha2-glycoprotein-1 upregulation in plasma and kidney of patients with lupus nephritis. BMC Nephrol. 2020;21:122.

85. Sharapova TN, Romanova EA, Soshnikova NV, et al. Autoantibodies from SLE patients induce programmed cell death in murine fibroblast cells through interaction with TNFR1 receptor. Sci Rep. 2020;10:11144.

86. Yan C, Yu L, Zhang XL, et al. Cytokine Profiling in Chinese SLE Patients: Correlations with Renal Dysfunction. J Immunol Res. 2020;2020:8146502.

87. Gorji AE, Roudbari Z, Alizadeh A, et al. Investigation of systemic lupus erythematosus (SLE) with integrating transcriptomics and genome wide association information. Gene. 2019;706:181–7.

88. Leal T, Carvalho S, Costa JM. A granular cell tumor: an unusual colon polyp. Rev Esp Enferm Dig. 2019;111:329.

Figures
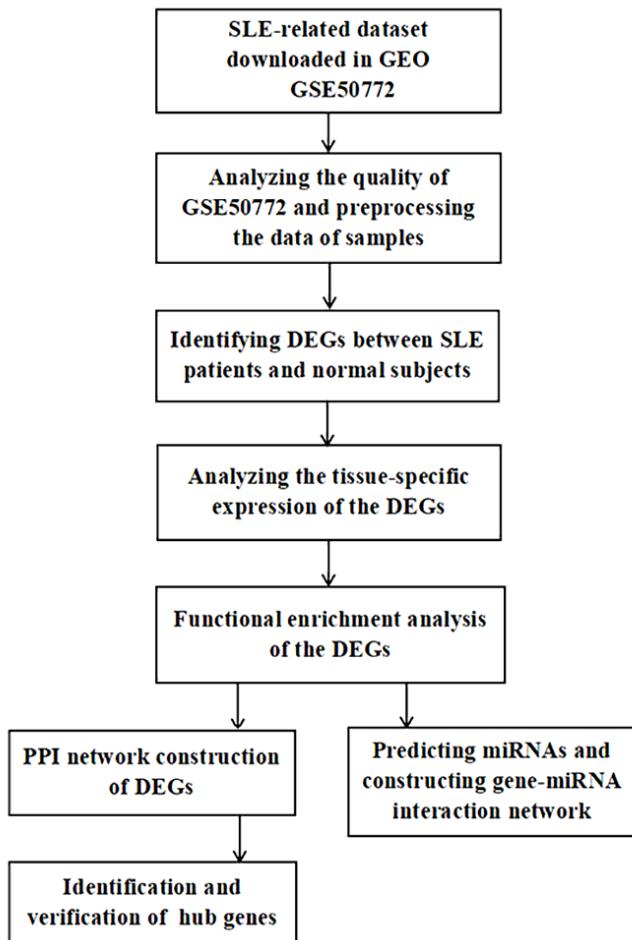
Figure 1

Flow chart of data preparation, processing, analysis, and validation. The gene expression profiles of GSE50772 were downloaded from the GEO database. Tissue-specific expression of genes, the functional enrichment and PPI networks were used to investigate potential biomarkers associated with the pathogenesis and clinical manifestation of SLE. In addition, key miRNA and hub genes were were further identified, and some SLE-related hub genes were validated based on data from the GeneCards database.
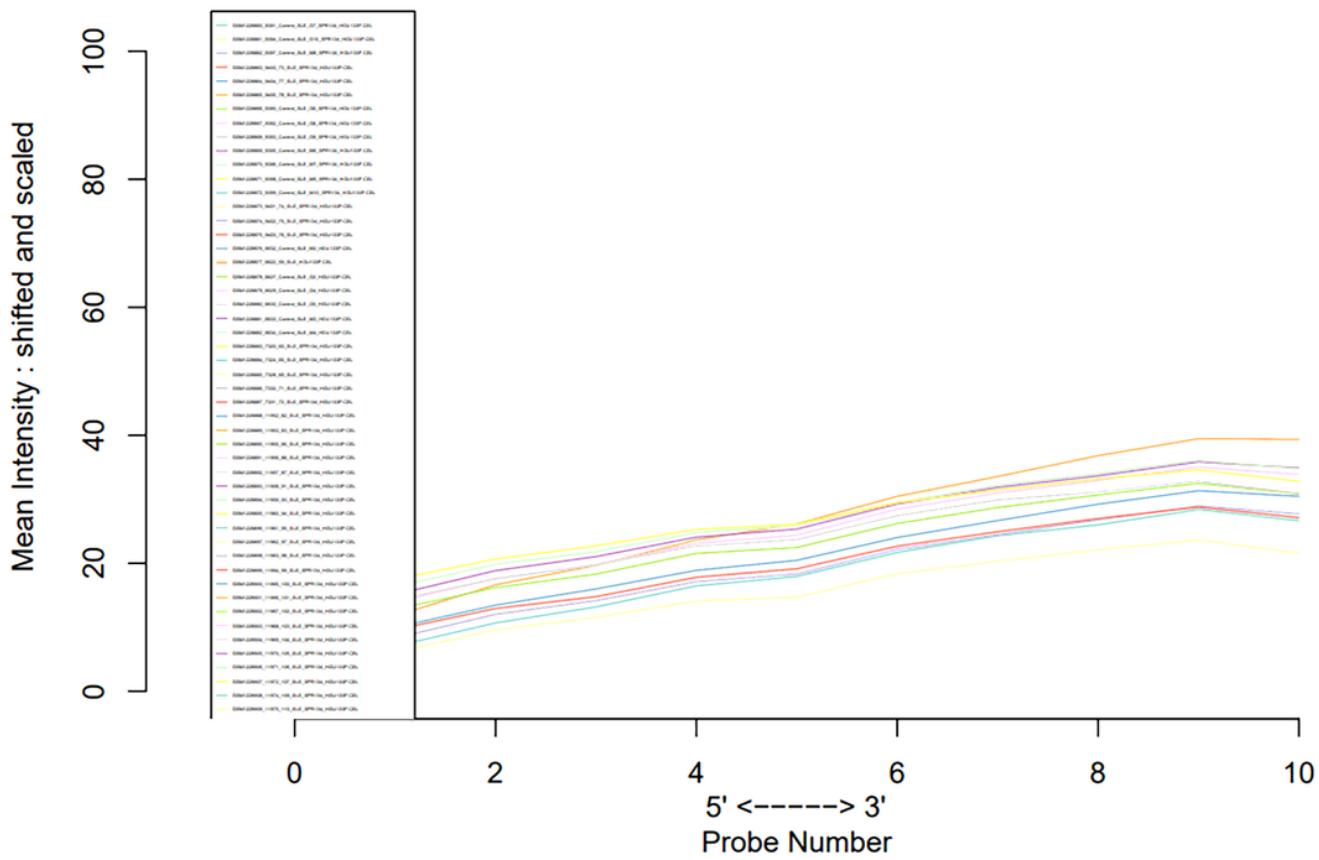
Figure 2

the RNA degradation plot.
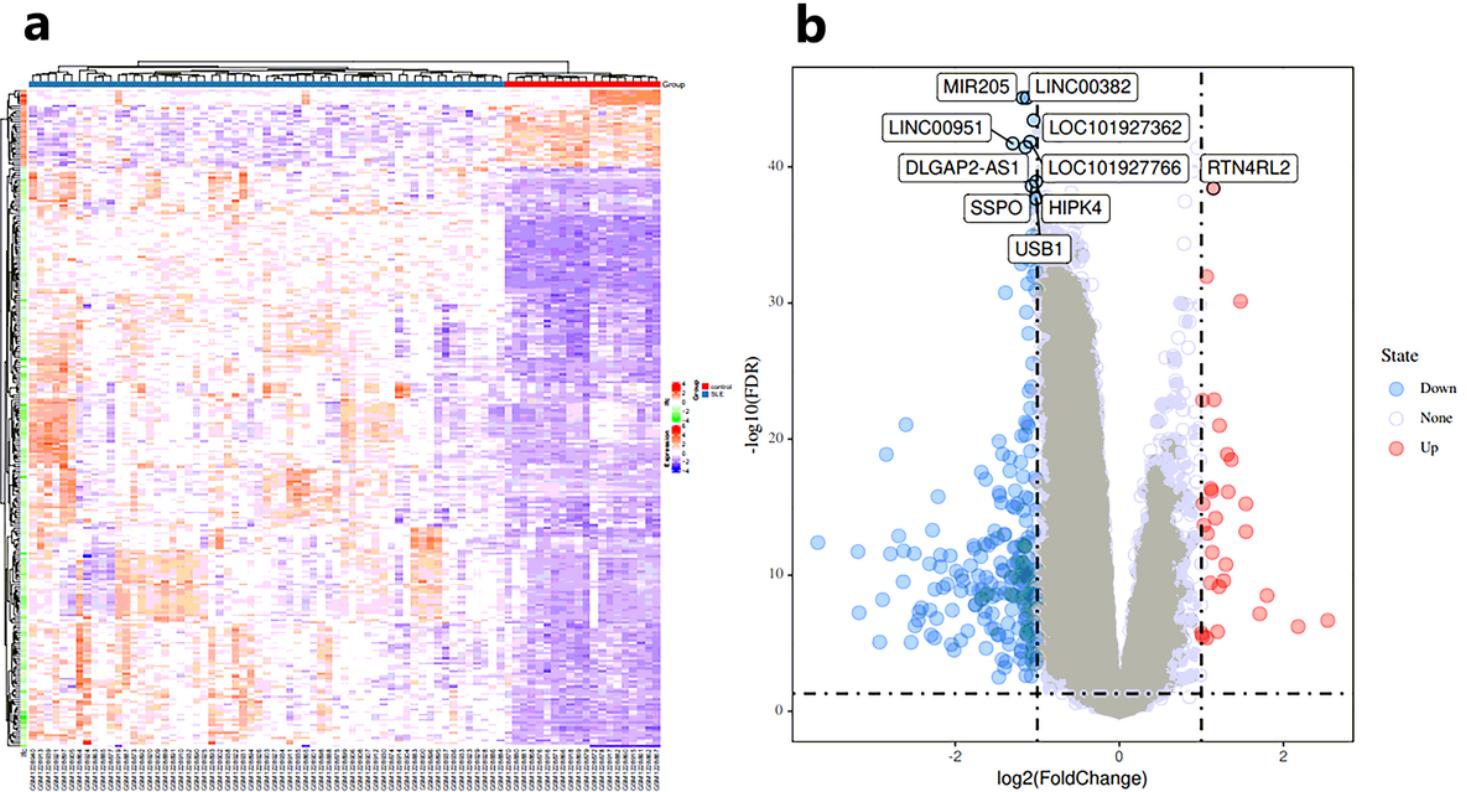
Figure 3

DEGs in 61 SLE patients and 20 normal individuals. (A) The heatmap: the potential DEGs between SLE samples and normal samples in GSE50772. (B) The volcano plot: blue dots represent significantly down-regulated genes, and red dots represent significantly up-regulated genes.
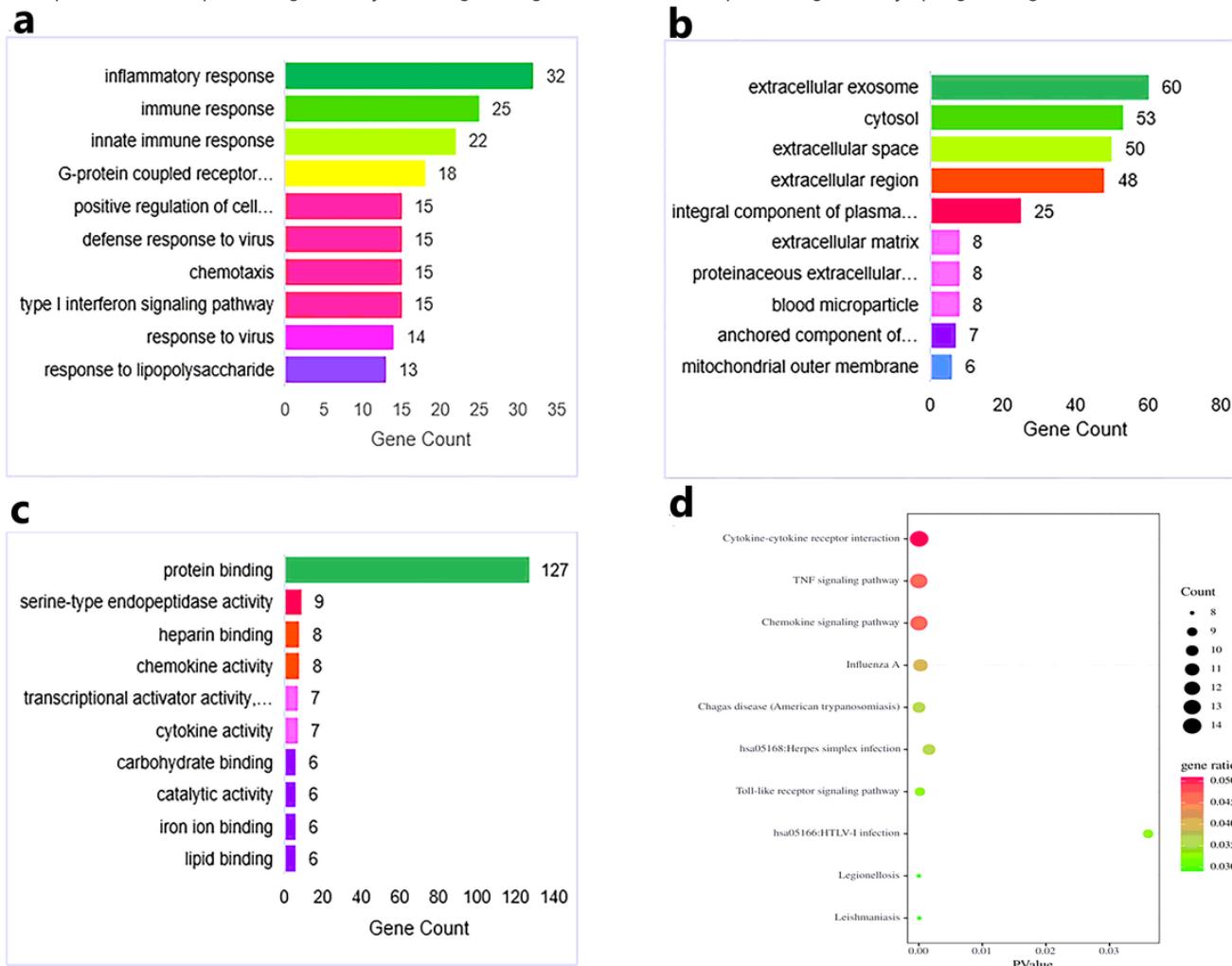

Figure 4

A bar chart of top 10 GO and KEGG terms of DEGs based on the count of genes. (A)BP: biological process; (B) CC: cellular component; (C) MF: molecular function; (D) KEGG: kyoto encyclopedia of genes and genomes.
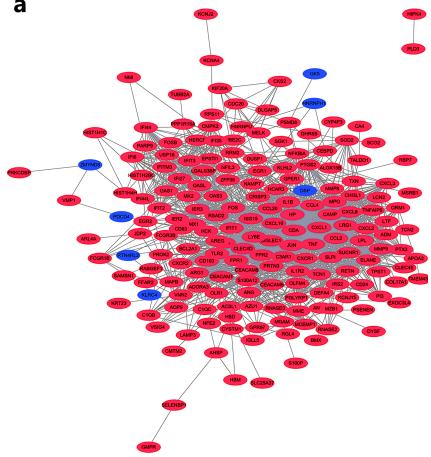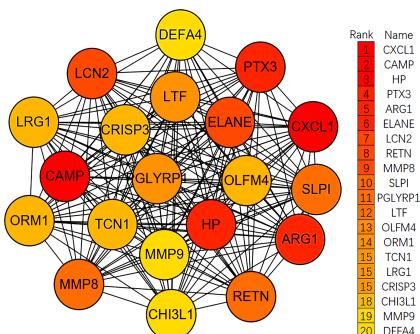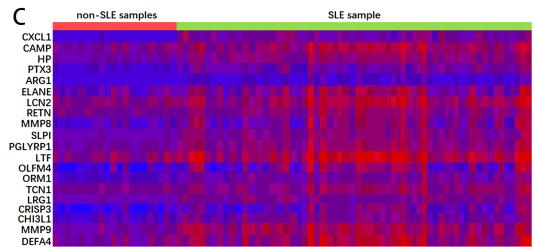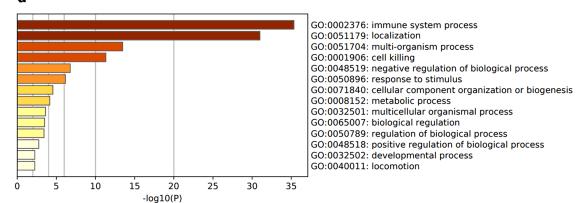
Figure 5

The PPI networks and the most significant module of DEGs. (A) The PPI network of DEGs was constructed using Cytoscape. The red dot represents up-regulated gene and the green dot represents down-regulated gene. (B) Top 20 hub genes in network ranked by MCC method, and the color of a node in the network reflects the rank of hub genes. (C) Hierarchical clustering of hub genes was constructed using Excel. The samples under the pink bar are non-SLE samples and the samples under the glue bar are SLE samples. High expression of genes is marked in red; lower expression of genes is marked in blue. (D)The top-level Gene Ontology biological processes of hub genes were performed using Metascape.
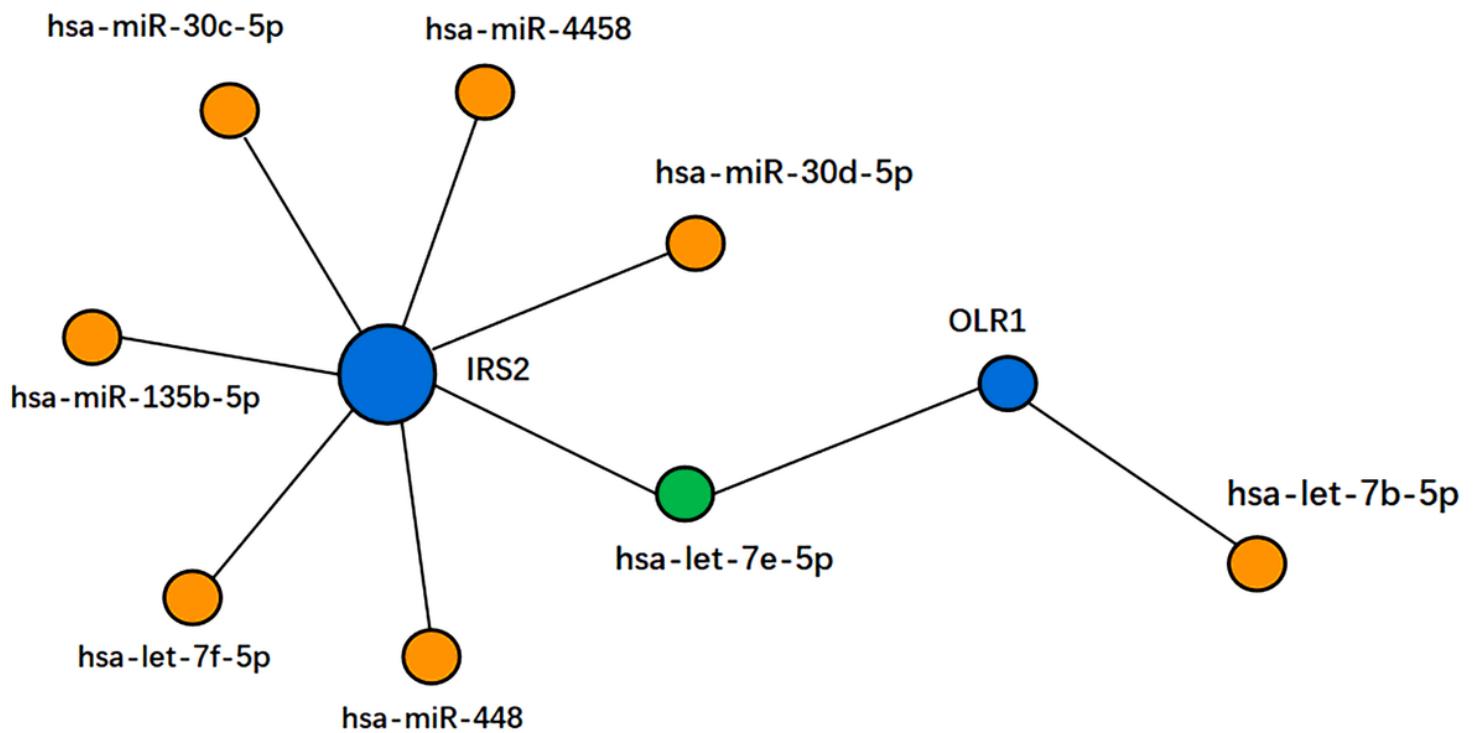
Figure 6

Interaction network between genes involved in top 10 biological processes and its targeted miRNAs. Genes are coloured in blue, and node size is adjusted according to number of targeted miRNAs; miRNAs are coloured in red; miRNAs targeting more than two genes simultaneously are coloured in green.
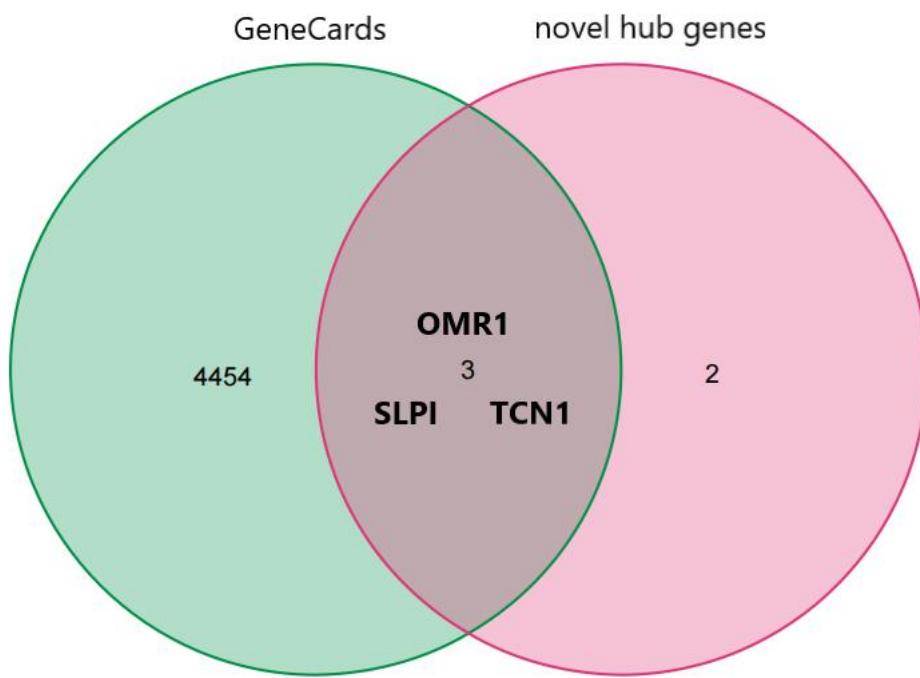

Figure 7

Venn diagram of key genes between five novel hub genes in our study and SLE-related genes in GeneCards. ORM1,SPLI and TCN1 are identified.

Supplementary Files

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

- Supplementarymaterial.docx