Identification of Potential ceRNA Network and Patterns of Immune Cell Infiltration in Systemic Sclerosis-Associated Interstitial Lung Disease

Qiuhong Wu[1], Yang Liu[2], Yan Xie[1], Shixiong Wei[1] and Yi Liu[1,3,4*]

[1] Department of Rheumatology and Immunology, West China Hospital, Sichuan University, Chengdu, China, [2] Lung Cancer Center, West China Hospital, Sichuan University, Chengdu, China, [3] Rare Diseases Center, West China Hospital, Sichuan University, Chengdu, China, [4] Institute of Immunology and Inflammation, Frontiers Science Center for Disease-Related Molecular Network, West China Hospital, Sichuan University, Chengdu, China

Purpose: Systemic sclerosis-associated interstitial lung disease (SSc-ILD) is one of the most severe complications of systemic sclerosis (SSc) and is the leading cause of SSc-related deaths. However, the precise pathogenesis of pulmonary fibrosis in SSc-ILD remains unknown. This study aimed to evaluate the competing endogenous RNA (ceRNA) regulatory network and immune cell infiltration patterns in SSc-ILD.

Methods: One microRNA (miRNA) and three messenger RNA (mRNA) microarray datasets were obtained from the Gene Expression Omnibus (GEO) database. Then, the differentially expressed miRNAs (DEmiRs) and mRNAs (DEMs) between SSc-ILD patients and normal controls were identified, respectively, followed by the prediction of the target genes and target lncRNAs of DEmiRs. The overlapping genes between DEmiRs target genes and DEMs were identified as core mRNAs to construct the ceRNA network. In addition, the "Cell Type Identification by Estimating Relative Subsets of Known RNA Transcripts (CIBERSORT)" algorithm was used to analyze the composition of infiltrating immune cells in lung tissues of SSc-ILD patients and controls, and differentially expressed immune cells were recognized. The correlation between immune cells and core mRNAs was evaluated by Pearson correlation analysis.

Results: Totally, 42 SSc-ILD lung tissues and 18 normal lung tissues were included in this study. We identified 35 DEmiRs and 142 DEMs and predicted 1,265 target genes of DEmiRs. Then, 9 core mRNAs related to SSc-ILD were recognized, which were the overlapping genes between DEmiRs target genes and DEMs. Meanwhile, 9 DEmiRs related to core mRNAs were identified reversely, and their target lncRNAs were predicted. In total, 9 DEmiRs, 9 core mRNAs, and 51 predicted lncRNAs were integrated to construct the ceRNA regulatory network of SSc-ILD. In addition, 9 types of immune cells were differentially expressed in lung tissues between SSc-ILD patients and controls. Some core mRNAs, such as *COL1A1*, *FOS*, and *EDN1*, were positively or negatively correlated with the number of infiltrating immune cells.

Conclusion: This is the first comprehensive study to construct the potential ceRNA regulatory network and analyze the composition of infiltrating immune cells in lung tissues of SSc-ILD patients, which improves our understanding of the pathogenesis of SSc-ILD.

Keywords: systemic sclerosis, interstitial lung disease, ceRNA, immune cell infiltration, GEO

INTRODUCTION

Systemic sclerosis (SSc) is a systemic autoimmune disease characterized by immune dysregulation, vasculopathy, and localized or diffused skin fibrosis. As a rare disease, the incidence of SSc is very low (10–50 per million people per year), but the mortality rate of SSc is relatively higher than other rheumatic diseases (Allanore et al., 2015; Denton and Khanna, 2017). Systemic sclerosis-associated interstitial lung disease (SSc-ILD) is a common SSc complication, which is the main cause of SSc-related deaths (Perelas et al., 2020). However, the treatment options of SSc-ILD are limited, partly because of the fact that the precise mechanisms of the development of pulmonary fibrosis in SSc-ILD have not been clearly elucidated. Thus, exploring potential pathogenesis in SSc-ILD and finding effective biomarkers for the diagnosis and treatment of SSc-ILD are quite necessary.

Non-coding RNAs (ncRNAs) account for nearly 98% of the human genome. Although ncRNAs do not code for proteins, they play significant roles in maintaining normal development and physiological functions by regulating gene expression at both the transcriptional and post-transcriptional levels (Esteller, 2011; Cech and Steitz, 2014). MicroRNAs (miRNAs), as a subgroup of short ncRNAs, are evolutionarily conserved single-stranded RNA molecules, which approximately contain 19–25 nucleotides. They can negatively regulate post-transcriptional gene expression by pairing with complementary sequences in the $3'$-untranslated region ($3'$-UTR) of the target messenger RNAs (mRNAs), resulting in target gene degradation or translation inhibition. miRNAs are involved in a variety of cellular biological processes, such as cell proliferation, apoptosis, and differentiation (Bartel, 2004, 2009; Valencia-Sanchez et al., 2006). And accumulating evidences have demonstrated that the dysregulation of miRNAs was associated with the development of many human diseases, such as cancers, neurological disorders, cardiovascular disorders, as well as rheumatic diseases (Nakasa et al., 2008; Esteller, 2011; Christmann et al., 2016; Young et al., 2017). For instance, Nakasa et al. (2008) demonstrated that miR-146a/b was highly expressed in synovial tissues of rheumatoid arthritis (RA). Christmann et al. (2016) found that the expression of miR-155 was related to progressive SSc-ILD. Therefore, miRNAs can be identified as potential biomarkers for the diagnosis, treatment and prognosis of diseases.

Long non-coding RNAs (lncRNAs) are defined as a type of ncRNAs that are longer than 200 nucleotides in length, and they perform multilevel regulatory functions in gene expression, such as transcription, translation and epigenetics. In addition, lncRNAs play pivotal roles in maintaining cellular homeostasis, and differential expression patterns of lncRNAs have been found in many human diseases (Mercer et al., 2009; Li et al., 2013; Shi et al., 2013). Salmena et al. (2011) put forward a competing endogenous RNA (ceRNA) hypothesis, in which lncRNAs, mRNAs, and other RNAs were able to compete with each other as ceRNAs to bind to miRNAs through sharing one or more miRNA response elements (MREs). lncRNAs can serve as endogenous molecular sponges for miRNAs to indirectly regulate the expression of mRNAs, and the ceRNA network links the function of ncRNAs to the function of protein-coding mRNAs (Salmena et al., 2011; Tay et al., 2014). In recent years, many studies have indicated that the dysregulation of lncRNA-miRNA-mRNA networks was closely related to the pathogenesis of many human diseases, such as cancers (Wang et al., 2016; Zhang et al., 2016). Nevertheless, the potential ceRNA regulatory network in SSc-ILD has not been explored.

To be mentioned, the occurrence and progression of SSc-ILD is a complicated process involving multiple biological mechanisms, such as autoimmunity, fibrosis, inflammation, and vascular injury (Ciechomska et al., 2015; Furue et al., 2017; Brown and O'Reilly, 2019; Perelas et al., 2020). Immune cell infiltration is closely associated with the fibrosis of SSc patients. Many immune cells, such as T cells, B cells, monocytes, dendritic cells (DCs), macrophages, and mast cells, have been indicated in the pathogenesis of fibrosis in SSc (Brown and O'Reilly, 2019). For instance, Kalogerou et al. (2005) found that the expression levels of CD69, the early T cell activation marker, were increased in skin biopsy samples of SSc patients, implying that T cells might be included in the process of skin fibrosis. Francois et al. (2013) indicated that B cells played important roles in the production of collagens by dermal fibroblasts and could be the inducers of fibrosis in SSc. However, in SSc-ILD, the landscape of immune cell infiltration has not been fully elucidated.

In this study, we download one miRNA microarray dataset (GSE81293) and three mRNA microarray datasets (GSE81292, GSE48149, and GSE76808) from the Gene Expression Omnibus (GEO) database and identified differentially expressed miRNAs (DEmiRs) and differentially expressed mRNAs (DEMs) in lung tissues between SSc-ILD patients and normal controls. Then, the target genes and target lncRNAs of DEmiRs were predicted. The overlapping genes between DEmiRs target genes and DEMs were identified as core mRNAs to construct the potential ceRNA regulatory network of SSc-ILD. Meanwhile, the algorithm "Cell Type Identification by Estimating Relative Subsets of Known RNA Transcripts (CIBERSORT)" was used to analyze the different patterns of immune cell infiltration between SSc-ILD lung tissues and normal lung tissues. Moreover, the co-expression patterns of immune cells and core mRNAs were also evaluated. This study aimed to explore the potential ceRNA-related

pathogenesis of SSc-ILD and identify the underlying immune gene signature of SSc-ILD.

MATERIALS AND METHODS

Microarray Data Download and Processing

A total of one miRNA microarray dataset GSE81293 (Christmann et al., 2016) and three mRNA microarray datasets GSE81292 (Christmann et al., 2016), GSE48149 (Hsu et al., 2011; Renaud et al., 2020), and GSE76808 (Christmann et al., 2014) from lung biopsy tissues of SSc-ILD patients and controls were downloaded from the National Center for Biotechnology Information (NCBI) GEO database[1]. The miRNA microarray dataset GSE81293 was based on the GPL16384 platform ([miRNA-3] Affymetrix Multispecies miRNA-3 Array) and included 15 SSc-ILD samples and 5 normal control samples. And a total of 42 SSc-ILD lung specimens and 18 normal lung tissues were enrolled from three mRNA microarray datasets GSE81292, GSE48149, and GSE76808, whose detection platforms were GPL18991 ([HG-U133A_2] Affymetrix Human Genome U133A 2.0 Array), GPL16221 (Illumina HumanRef-8 v3.0 expression beadchip), and GPL571 ([HG-U133A_2] Affymetrix Human Genome U133A 2.0 Array), respectively (**Table 1**).

Each probe expression matrix was matched with its platform annotation, and the probe IDs were transformed into homologous gene symbols. All the expression data were normalized and log2 transformed. If one gene had more than one probe expression data, the average value was taken for further analysis. Then, the expression matrixes of three mRNA microarray datasets GSE81292, GSE48149, and GSE76808 were merged into one, and the heterogeneity caused by different experimental platforms and batches was eliminated, which was conducted using "sva" package of R software (v.4.0.5) (Leek et al., 2020).

Differential Expression Analysis

The miRNAs and mRNAs that were differentially expressed between SSc-ILD lung tissues and normal lung tissues were identified by "limma" package of R software (Ritchie et al., 2015). *P*-values were adjusted by the Benjamini-Hochberg (BH) false discovery rate (FDR) method (Glickman et al., 2014). RNAs that meet the screening criteria [| log2 fold change (FC)| > 1, and FDR adjusted *P*-value (*q*-value) < 0.05] were considered as DEmiRs or DEMs. Then, heat maps and volcano plots were drawn using "pheatmap" and "ggplot2" packages of R software to depict these differentially expressed RNAs.

Gene Ontology (GO) and Pathway Enrichment Analysis

In order to explore the potential biological mechanisms of DEMs in SSc-ILD, GO enrichment analysis [including biological process (BP), cellular component (CC), and molecular

function (MF)] and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis were conducted using R "ClusterProfiler" and "org.Hs.eg.db" packages (Yu et al., 2012). As for DEmiRs, the GO and biological pathway enrichment analysis were carried out using the FunRich software (v.3.1.3) (Pathan et al., 2015). The *q*-value < 0.05 was considered statistically significant.

Construction of the Protein-Protein Interaction (PPI) Network and Hub Gene Analysis

The Search Tool for the Retrieval of Interacting Genes (STRING) online database[2] was used to analyze the PPI network for the DEMs with interaction score > 0.7 (Szklarczyk et al., 2019). Then, the PPI network was constructed and visualized using Cytoscape software (v.3.8.1), and the top 10 hub genes were identified by the Cytoscape's cytoHubba plugin according to the number of connections (Chin et al., 2014).

Prediction of DEmiRs Target Genes and Identification of Core mRNAs

The DEmiRs target genes were predicted using three databases including TargetScan (Agarwal et al., 2015), miRTarBase (Huang et al., 2020), and miRDB (Chen and Wang, 2020). Only those genes that concurrently existed in all three databases were considered as the candidate targets of DEmiRs. Subsequently, the intersection of DEmiRs target genes and DEMs was analyzed using the Venny 2.1.0 online database[3], and the overlapping genes were identified as core mRNAs related to SSc-ILD.

Prediction of DEmiRs Target lncRNAs and Construction of the ceRNA Network

The DIANA-LncBase v.2 prediction module[4] was used to predict the target lncRNAs of DEmiRs related to core mRNAs (Paraskevopoulou et al., 2016). The lncRNAs that meet the filter criteria (prediction interaction score > 0.99 and lncRNA tissue type: lung) were selected as potential DEmiRs target lncRNAs. Then, DEmiRs regulated for both lncRNAs and core mRNAs were selected for the construction of ceRNA regulatory network, which was visualized by Cytoscape software (v.3.8.1).

Analysis of Immune Cell Infiltration

To analyze the proportion of 22 infiltrating immune cells in lung tissues of SSc-ILD patients and normal controls, the mRNA expression data were uploaded to the online platform of CIBERSORT[5], with 1,000 permutations applied to the algorithm (Newman et al., 2015). The CIBERSORT results with output *P*-values < 0.05 were considered eligible and accurate for further analysis. Subsequently, the Wilcoxon rank-sum test was performed to assess the differential composition of infiltrating

immune cells between SSc-ILD patients and controls, which was visualized by "pheatmap" and "vioplot" packages in R software. Two-sided $P < 0.05$ was considered statistically significant. Furthermore, the correlation among 22 immune cell subtypes and the relationship between differentially expressed immune cells and core mRNAs were evaluated using Pearson correlation analysis.

RESULTS

Identification of DEmiRs and DEMs in SSc-ILD

In total, 42 SSc-ILD lung samples and 18 normal lung tissues from GEO database were included in this study, and the details of these datasets (one miRNA microarray dataset and three mRNA microarray datasets) were shown in **Table 1**. In the miRNA microarray dataset GSE81293, 35 DEmiRs were identified in lung tissues of SSc-ILD patients compared with matched normal lung tissues, of which 11 DEmiRs were upregulated and 24 DEmiRs were downregulated in SSc-ILD. The volcano plot and heat map of these DEmiRs were shown in **Figures 1A,B**. As for the merged mRNA microarray dataset, a total of 142 DEMs were obtained between SSc-ILD patients and normal controls, including 53 upregulated DEMs and 89 downregulated DEMs in SSc-ILD, which were shown by volcano plot and heat map in **Figures 1C,D**.

GO and Pathway Enrichment Analysis of DEmiRs and DEMs in SSc-ILD

In order to explore the potential mechanisms of DEmiRs and DEMs in SSc-ILD, GO and pathway enrichment analysis were performed. **Figure 2** showed the top 10 enriched GO terms including BP, CC, and MF and the top 10 enriched biological pathways of DEmiRs. The BP terms of these DEmiRs were significantly enriched in cell communication (q-value < 0.01) and signal transduction (q-value < 0.01). Nucleus (q-value < 0.05) and cytoplasm (q-value < 0.05) were two significantly enriched CC terms. And in the MF group, DEmiRs were primarily enriched in transcription factor activity (q-value < 0.01), cell adhesion molecule activity (q-value < 0.05), and receptor signaling protein serine/threonine kinase activity (q-value < 0.05) (**Figures 2A–C**). As for the biological pathway enrichment analysis of DEmiRs, they were significantly enriched in some fibrosis-related signaling pathways, such as integrin family cell surface interactions, beta1 integrin cell surface

interactions, TRAIL signaling pathway, and VEGF and VEGFR signaling network (all q-values < 0.001) (**Figure 2D**).

The results of GO and KEGG enrichment analysis of DEMs were shown in **Figure 3**. In the BP category, DEMs were mainly enriched in leukocyte migration, extracellular structure organization, extracellular matrix organization, response to lipopolysaccharide, and response to molecule of bacterial origin. Collagen-containing extracellular matrix, endoplasmic reticulum lumen, and collagen trimer were the top three enriched terms of CC. Regarding MF, DEMs were significantly enriched in receptor ligand activity, cytokine activity, G protein-coupled receptor binding, cytokine receptor binding, and cytokine binding. In addition, the KEGG analysis indicated that DEMs were significantly enriched in 14 pathways, such as cytokine-cytokine receptor interaction, IL-17 signaling pathway, viral protein interaction with cytokine and cytokine receptor, AGE-RAGE signaling pathway in diabetic complications, and TNF signaling pathway.

PPI Network Construction and Hub Genes Identification

The PPI network containing 80 nodes and 164 edges was constructed based on the STRING online database and visualized by Cytoscape software. Subsequently, the top 10 hub genes, including *IL6* (interleukin 6), *JUN* (Jun proto-oncogene), *CCL2* (C-C motif chemokine ligand 2), *FOS* (Fos proto-oncogene), *IGF1* (insulin like growth factor 1), *CXCL2* (C-X-C motif chemokine ligand 2), *ATF3* (activating transcription factor 3), *EDN1* (endothelin 1), *COL1A1* (collagen type I alpha 1 chain), and *SPP1* (secreted phosphoprotein 1), were identified using the Cytoscape's cytoHubba plugin according to the number of connections (**Figure 4**).

Identification of Core mRNAs and Construction of the ceRNA Network in SSc-ILD

We used three databases, TargetScan, miRTarBase, and miRDB, to predict the target genes of DEmiRs. There were a total of 1,265 target genes that coexisted in all the three databases (**Supplementary Table 1**). Then, the Venn method was used to analyze the intersection between DEMs and these DEmiRs target genes. As a result, 9 overlapping core mRNAs associated with SSc-ILD were identified (**Figure 5A**). According to these 9 core mRNAs, 9 DEmiRs were recognized reversely. Subsequently, we

TABLE 1 | Details of miRNA and mRNA datasets of patients with SSc-ILD.

Type	GEO accession	Platform	Sample organism	Experiment type	Samples (lung tissues), n	
					SSc-ILD patients	Controls
miRNA	GSE81293	GPL16384	Homo sapiens	Non-coding RNA profiling by array	15	5
mRNA	GSE81292	GPL18991	Homo sapiens	Expression profiling by array	15	5
	GSE48149	GPL16221	Homo sapiens	Expression profiling by array	13	9
	GSE76808	GPL571	Homo sapiens	Expression profiling by array	14	4

miRNA, microRNA; mRNA, messenger RNA; SSc-ILD, systemic sclerosis-associated interstitial lung disease.

FIGURE 1 | DEmiRs and DEMs in lung tissues between SSc-ILD patients and normal controls. **(A)** The volcano plot of DEmiRs. **(B)** The heat map of DEmiRs. **(C)** The volcano plot of DEMs. **(D)** The heat map of DEMs. DEmiRs, differentially expressed microRNAs; DEMs, differentially expressed messenger RNAs; SSc-ILD, systemic sclerosis-associated interstitial lung disease.
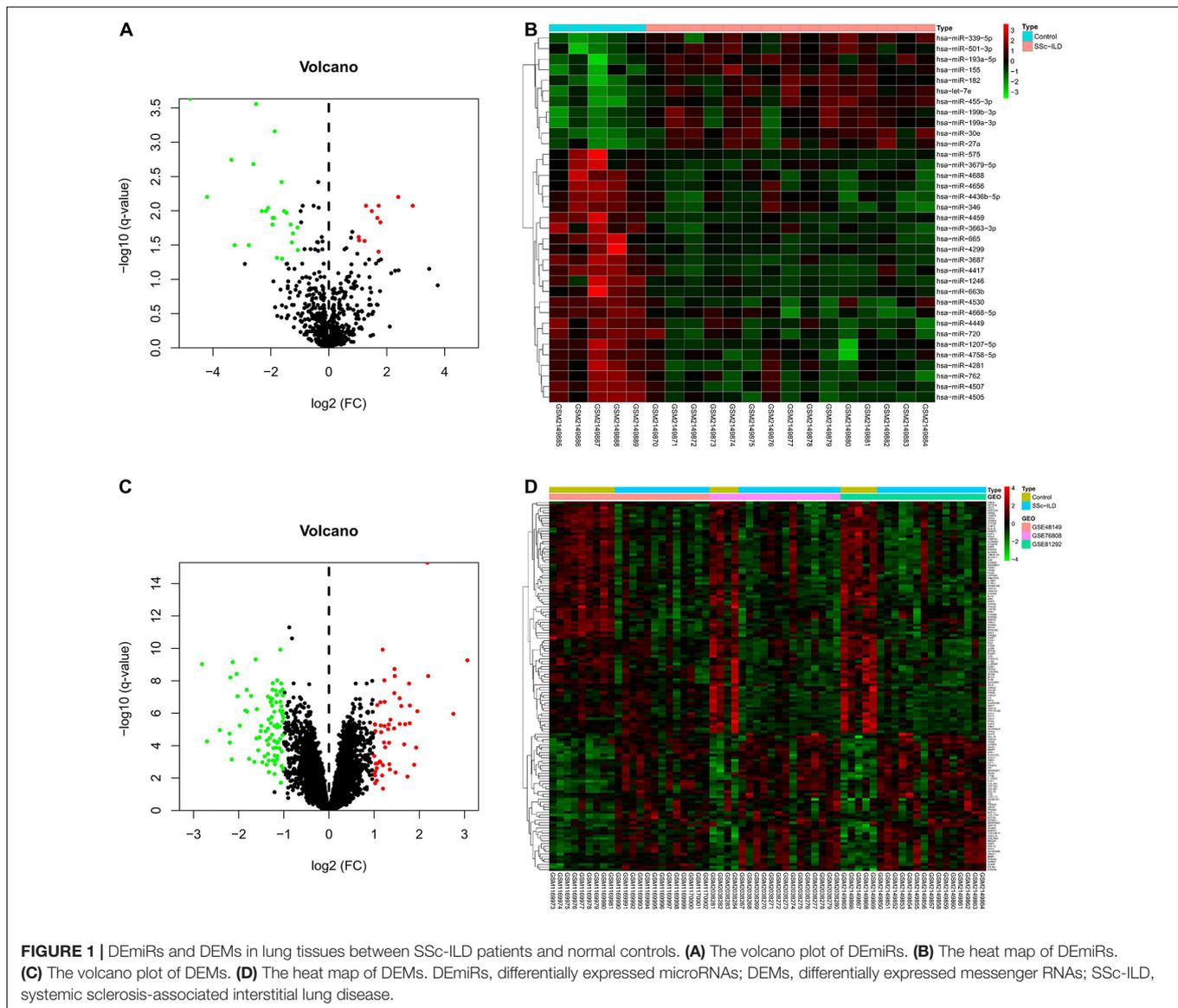
used the DIANA-LncBase v.2 prediction module to explore the target lncRNAs that could potentially bind to these 9 DEmiRs related to core mRNAs. Thes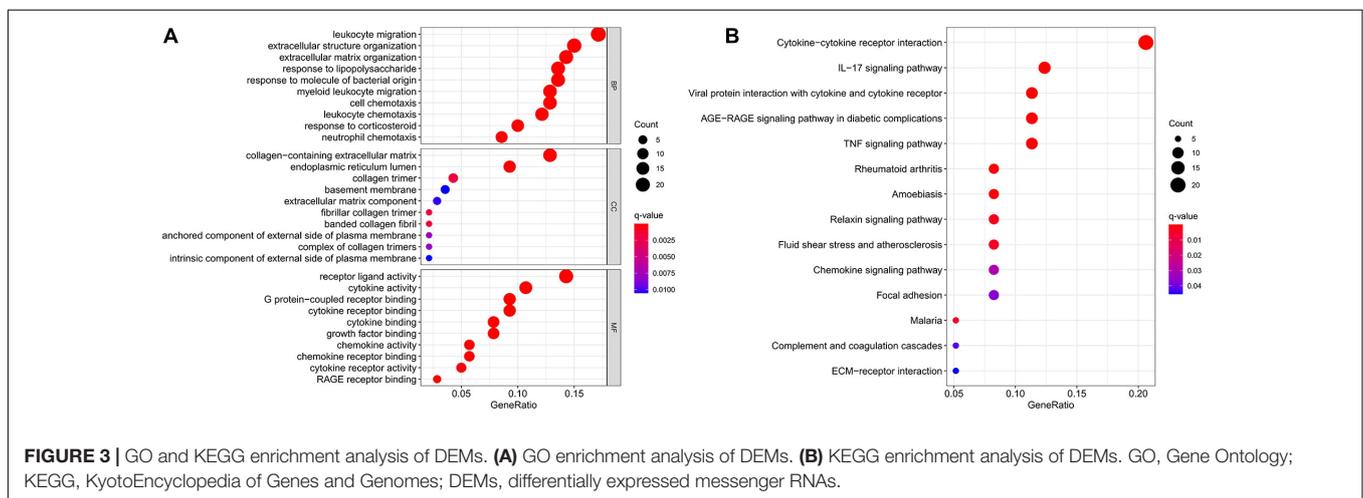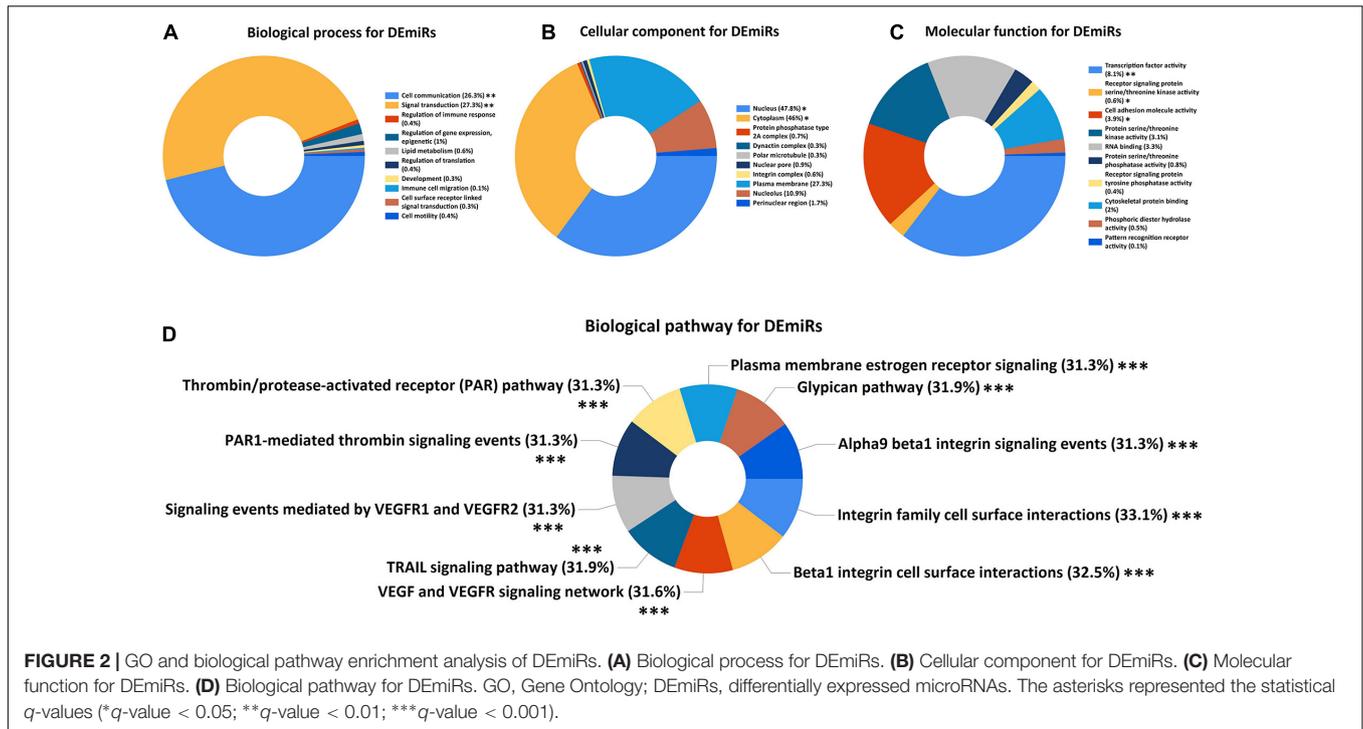e 9 DEmiRs could interact not only with lncRNAs but also with core mRNAs, which were selected for the construction of ceRNA regulatory network. In total, 9 DEmiRs, 9 core mRNAs, and 51 predicted lncRNAs were integrated to construct the potential ceRNA network of SSc-ILD, which included 69 nodes and 73 edges (**Table 2** and **Figure 5B**).

Composition of Infiltrating Immune Cells in SSc-ILD

The composition of 22 infiltrating immune cells in lung tissues of SSc-ILD patients and normal controls were estimated by the CIBERSORT algorithm (**Figure 6A**). Since the output P-value of GSM2038281 was greater than 0.05 ($P = 0.077$), it was excluded for further analysis. The distribution of 22 immune

cell types in each sample varied significantly, of which M2 macrophages accounted for the most enriched proportion. The relationships among 22 immune cells were indicated in **Figure 6B**. Activated mast cells were negatively correlated with resting mast cells ($R = -0.55$), whereas positively correlated with activated DCs ($R = 0.54$). Naive CD4$^+$ T cells had a negative correlation with resting memory CD4$^+$ T cells ($R = -0.51$), and monocytes were negatively correlated with resting mast cells ($R = -0.57$). Other immune cell subpopulations were weakly to moderately correlated.

The differential proportion of infiltrating immune cells between SSc-ILD patients and controls was analyzed by Wilcoxon rank-sum test, and the results showed that 9 types of immune cells, plasma cells, resting NK cells, activated NK cells, monocytes, M2 macrophages, resting DCs, resting mast cells, activated mast cells, and neutrophils, were differentially expressed. Specifically, plasma cells, activated NK cells, M2 macrophages, resting DCs,

FIGURE 2 | GO and biological pathway enrichment analysis of DEmiRs. **(A)** Biological process for DEmiRs. **(B)** Cellular component for DEmiRs. **(C)** Molecular function for DEmiRs. **(D)** Biological pathway for DEmiRs. GO, Gene Ontology; DEmiRs, differentially expressed microRNAs. The asterisks represented the statistical q-values (*q-value < 0.05; **q-value < 0.01; ***q-value < 0.001).

FIGURE 3 | GO and KEGG enrichment analysis of DEMs. **(A)** GO enrichment analysis of DEMs. **(B)** KEGG enrichment analysis of DEMs. GO, Gene Ontology; KEGG, KyotoEncyclopedia of Genes and Genomes; DEMs, differentially expressed messenger RNAs.

and resting mast cells were upregulated in lung tissues of SSc-ILD patients (**Figures 6C,D**).

Co-expression Patterns of Immune Cells and Core mRNAs

The correlation between core mRNAs and differentially expressed immune cells was estimated via Pearson correlation analysis, and the significantly correlated pairs with correlation coefficient > 0.5 and $P < 0.001$ were shown in **Figure 7**. The *ACVRL1* (activin A receptor like type 1) gene was positively correlated with neutrophils and resting NK cells ($R = 0.53$ and $R = 0.56$, respectively), but negatively correlated with plasma cells ($R = -0.53$). The *DUSP1* (dual specificity phosphatase 1) gene

and *FOS* gene were both positively correlated with activated mast cells ($R = 0.6$ and $R = 0.51$, respectively). And neutrophils could be negatively regulated by the *COL1A1* gene ($R = -0.51$) but be positively regulated by the *FSTL3* (follistatin like 3) gene ($R = 0.54$). In addition, both the *EDN1* gene and the *STC2* (stanniocalcin 2) gene had a positive correlation with activated mast cells ($R = 0.6$ and $R = 0.64$, respectively) and resting NK cells ($R = 0.51$ and $R = 0.54$, respectively).

DISCUSSION

SSc-ILD is one of the most severe complications of SSc and is the leading cause of SSc-related deaths. Although significant

FIGURE 4 | The PPI network of DEMs in SSc-ILD. The red dots represent the upregulated genes, and the green dots represent the downregulated genes. PPI, protein-protein interaction; DEMs, differentially expressed messenger RNAs; SSc-ILD, systemic sclerosis-associated interstitial lung disease.

progress has been achieved in the study of the pathogenesis of SSc-ILD, the detailed molecular mechanisms of the lung fibrosis in SSc-ILD remain largely unknown. Therefore, in order to improve the diagnosis and treatment of SSc-ILD, it is of vital significance to explore the potential molecular mechanisms and identify effective biomarkers with high specificity and sensitivity. In recent years, the hypothesis of ceRNA network has greatly aroused the interest of researchers, and it makes the link between protein-coding mRNAs and non-coding RNAs. Based on the ceRNA hypothesis, lncRNAs can act as miRNA sponges to influence the expression of mRNAs (Salmena et al., 2011; Tay et al., 2014). In this study, in order to explore the potential pathogenesis of SSc-ILD, we constructed the ceRNA regulatory network of SSc-ILD based on 4 microarray datasets.

miRNAs, as post-transcription regulators, play important roles in affecting the expression of downstream target genes, which in turns to participate in a variety of physiological and pathological processes (Valencia-Sanchez et al., 2006; Bartel, 2009). In recent years, the dysregulation of miRNA expression has been found to be related to the occurrence and development of many diseases including SSc-ILD. The expression of miRNAs can be used as an efficient biomarker for the diagnosis and prognosis of diseases. In fact, several miRNAs have been found to be associated with the pathogenesis of SSc fibrosis (Henry et al., 2019). Christmann et al. demonstrated that the expression levels of miR-155 in lung tissues were associated with progressive SSc-ILD. In the bleomycin-induced pulmonary fibrosis model, miR-155 knockout mice developed milder lung fibrosis when compared with wild-type mice (Christmann et al., 2016). Makino et al. indicated that let-7a expression levels were significantly decreased in both the skin and the serum of SSc patients, and the overexpression or inhibition of the expression levels of let-7a in human or mouse skin fibroblasts could regulate the type I collagen expression. Moreover, the overexpression of let-7a in the

FIGURE 5 | The core mRNAs identification and ceRNA regulatory network in SSc-ILD. **(A)** The intersection between DEMs and the target genes of DEmiRs. **(B)** The ceRNA regulatory network. The red dots represent DEmiRs, the green dots represent core mRNAs, and the blue dots represent predicted lncRNAs. mRNAs, messenger RNAs; ceRNA, competing endogenous RNA; SSc-ILD, systemic sclerosis-associated interstitial lung disease; DEMs, differentially expressed mRNAs; DEmiRs, differentially expressed microRNAs; lncRNAs, long non-coding RNAs.

skin could attenuate the skin fibrosis induced by bleomycin in mice (Makino et al., 2013). Maurer et al. (2010) found that miR-29a had a strong relationship with SSc fibrosis, in which miR-29a was significantly downregulated in SSc skins and fibroblasts, and it could also regulate the expression levels of type I and type III collagens. In the current study, we identified some DEmiRs related to SSc-ILD. Additionally, these DEmiRs were significantly enriched in some well-known signaling pathways that have been proved to be associated with fibrosis, such as integrin family cell surface interactions, beta1 integrin cell surface interactions, TRAIL signaling pathway, and VEGF and VEGFR signaling network (all q-values < 0.001), indicating our results were credible.

Furthermore, we integrated three mRNA microarray data to identify the DEMs and uncovered the underlying PPI network and hub genes in SSc-ILD. Meanwhile, the DEmiRs target genes were predicted using three distinct prediction databases. Then, 9 overlapping core mRNAs associated with SSc-ILD were identified, which were not only the target genes of DEmiRs, but also the DEMs between SSc-ILD and controls. Among them, four core mRNAs, *COL1A1*, *IGF1*, *EDN1*, and *FOS*, were also the hub genes existed in the PPI network.

COL1A1 gene is a fibrosis-related gene that encodes the alpha1 chains of type I collagen, which is the most common extracellular matrix (ECM) protein in tissues. Although the precise mechanisms of fibrosis have not been elucidated, type

TABLE 2 | The DEmiRs, core mRNAs, and predicted lncRNAs in the ceRNA network.

miRNA	Core mRNA	lncRNA
hsa-let-7e	DUSP1, EDN1, IGF1	XLOC_012138, TPTEP1
hsa-miR-155	FOS, GPM6B	LOC100190986, LINC00506, MIR155HG, XLOC_012997, LINC01915, CASC2
hsa-miR-199a-3p	STC2	XLOC_009222, TUG1, XLOC_010279
hsa-miR-199b-3p	STC2	XLOC_009222, TUG1, XLOC_010279
hsa-miR-4505	ACVRL1	CYP4F26P, UCA1, RP11-34P13.13, XLOC_009222, RP11-206L10.3, XLOC_010279, XLOC_007253, XLOC_000782, LINC00665, IGFL2-AS1, RP11-1348G14.4
hsa-miR-4530	FOS	XLOC_000667, LNCPRESS1, LINC00662, LINC00641
hsa-miR-4656	COL1A1	UCA1, RP11-81H3.2, DGCR5, XLOC_009222, RP11-391M1.4, LINC01659, AC068039.4, XLOC_002214, LOC284581, RP11-1275H24.3, XLOC_002216, XLOC_010972, LINC00475, RP11-680F8.1, XLOC_012444, XLOC_012351, XLOC_011448
hsa-miR-4688	FSTL3	UCA1, XLOC_012083, LINC00894, ARHGEF26-AS1, FAM239A, FAM239B, XLOC_013420, LINC00662
hsa-miR-762	COL1A1	UCA1, AC093627.9, XLOC_002214, ZNF571-AS1, AC009404.2, LINC00707, AC017002.4

DEmiRs, differentially expressed microRNAs; mRNAs, messenger RNAs; ceRNA, competing endogenous RNA.

I collagens play significant roles in the development of tissue fibrosis in SSc (Hitraya and Jimenez, 1996; Jimenez and Saitta, 1999). In fact, a lot of studies have found that the expression levels of collagen corresponding genes including *COL1A1* were increased in SSc fibroblasts, and inhibiting collagen genes expression or collagen proteins production could be an effective treatment for the tissue fibrosis of SSc (Hitraya and Jimenez, 1996; Jimenez and Saitta, 1999; Sandorfi et al., 2005; Piera-Velazquez et al., 2020; Shi et al., 2020).

IGF1 gene encodes insulin-like growth factor 1 that is similar to insulin in function and structure. IGF1 participates in the regulation of a variety of physiological processes, such as the regulation of cell cycle, mitosis, apoptosis, inflammation, and immunity (Guevara-Aguirre, 1996; Stewart and Rotwein, 1996; Heemskerk et al., 1999). Many studies have indicated that IGF1 protein was involved in the development of SSc (Hamaguchi et al., 2008; Winsz-Szczotka et al., 2016; Tabata et al., 2021). For instance, Hamaguchi et al. reported that patients with SSc had significantly increased levels of serum IGF1 compared with patients with systemic lupus erythematosus or healthy controls. In addition, the expression levels of serum IGF1 were even higher in patients with high skin thickness scores or with severe pulmonary fibrosis (Hamaguchi et al., 2008).

The EDN1, encoded by *EDN1* gene, is a secreted peptide that belongs to the endothelin/sarafotoxin family, and numbers of studies have illustrated the active role of EDN1 in fibrogenesis. EDN1 can lead to the conversion of fibroblasts to myofibroblasts and increase the deposition of excessive collagens, which are two key processes of tissue fibrosis (Shi-Wen et al., 2004, 2007; Soldano et al., 2009). Moreover, EDN1 can induce the apoptosis resistance in lung fibroblasts and can mediate the transforming growth factor-β (TGF-β)-induced fibrosis (Kulasekaran et al., 2009). In patients with SSc-ILD, the expression of EDN1 is increased in lung tissues, and the blockade of EDN1 could be an effective method to limit the progression of lung fibrosis (Abraham et al., 1997; Swigris and Brown, 2010).

FOS gene encodes c-FOS protein, which can dimerize with JUN family proteins to form transcription factor complex activator protein-1 (AP-1). The c-FOS proteins are included in many physiological processes, such as proliferation, differentiation, and transformation, and can act as regulators in tissue fibrosis (Angel and Karin, 1991; Sullivan et al., 2009). Importantly, in SSc, Avouac et al. found the upregulation of c-FOS in SSc animal models and the skin and dermal fibroblasts of SSc patients. Selective inhibition of AP-1 could decrease the production of collagens in SSc fibroblasts and prevent the development of experimental dermal fibrosis (Avouac et al., 2012).

To be mentioned, the lung fibrosis in SSc-ILD is caused by the interactions of multiple factors, such as immune cell infiltration, autoimmunity, inflammation, and vascular injury. Both the innate and adaptive immunity are involved in the pathogenesis of SSc-ILD, and immune cell infiltration is a hallmark of SSc-ILD (Furue et al., 2017; Brown and O'Reilly, 2019). In the current study, we evaluated the composition of infiltrating immune cells in lung tissues of SSc-ILD patients and identified differentially expressed immune cells between SSc-ILD patients and normal controls, such as M2 macrophages, DCs, and mast cells.

The polarization of M2 macrophages is closely related to the pathogenesis of tissue fibrosis in SSc (Chia and Lu, 2015; Manetti, 2015; Stifano and Christmann, 2016). M2 macrophages can produce profibrotic cytokine TGF-β, which plays a pivotal role in tissue fibrosis (Hu et al., 2018; Shapouri-Moghaddam et al., 2018). Christmann et al. (2014) compared the gene expression levels of lung tissues between SSc-ILD patients and controls, and the results revealed that M2 macrophage markers, such as *CD163* and *CCL18*, were upregulated in SSc patients, which were related to the progression of pulmonary fibrosis. Pechkovsky et al. (2010) illustrated the importance of M2 macrophage phenotype in the pathogenesis of lung fibrosis, and IL-4 and IL-10 could induce the shift of alveolar macrophages to M2. Our results in this study also found that the number of M2 macrophages was upregulated in lung tissues of SSc-ILD patients, which was consistent with those previous studies.

DCs are potent antigen-presenting cells (APCs) and play important roles in both innate and adaptive immune system (Audiger et al., 2017). Recently, the relationship between the pathogenesis of SSc and the expression of DCs, especially plasmacytoid DCs (pDCs), has been widely discussed (Lu, 2012; Affandi et al., 2018). The pDCs, as a special subset of DCs, could mediate the development of SSc by producing type I interferon (IFN-I) and other inflammatory mediators, such as
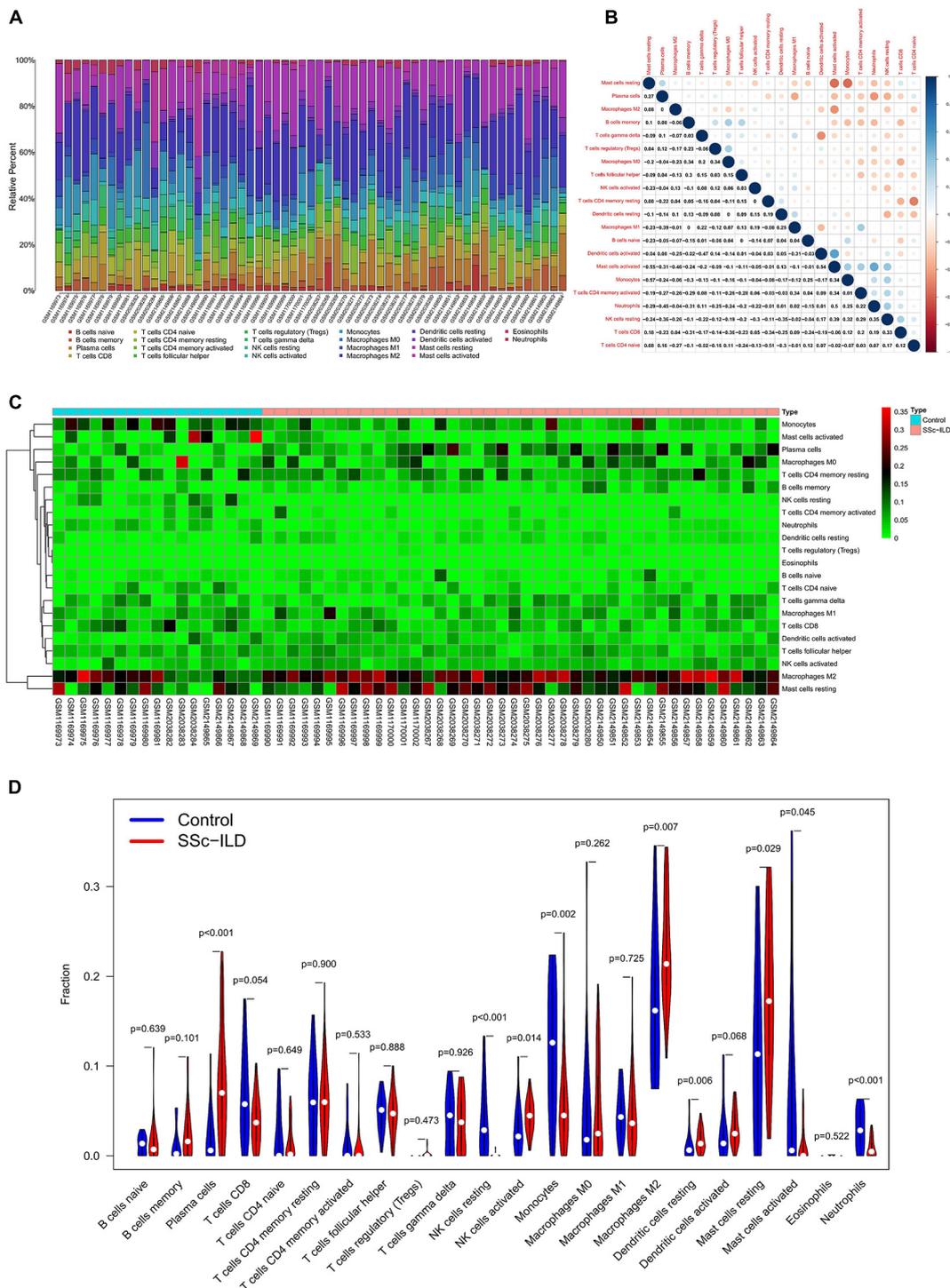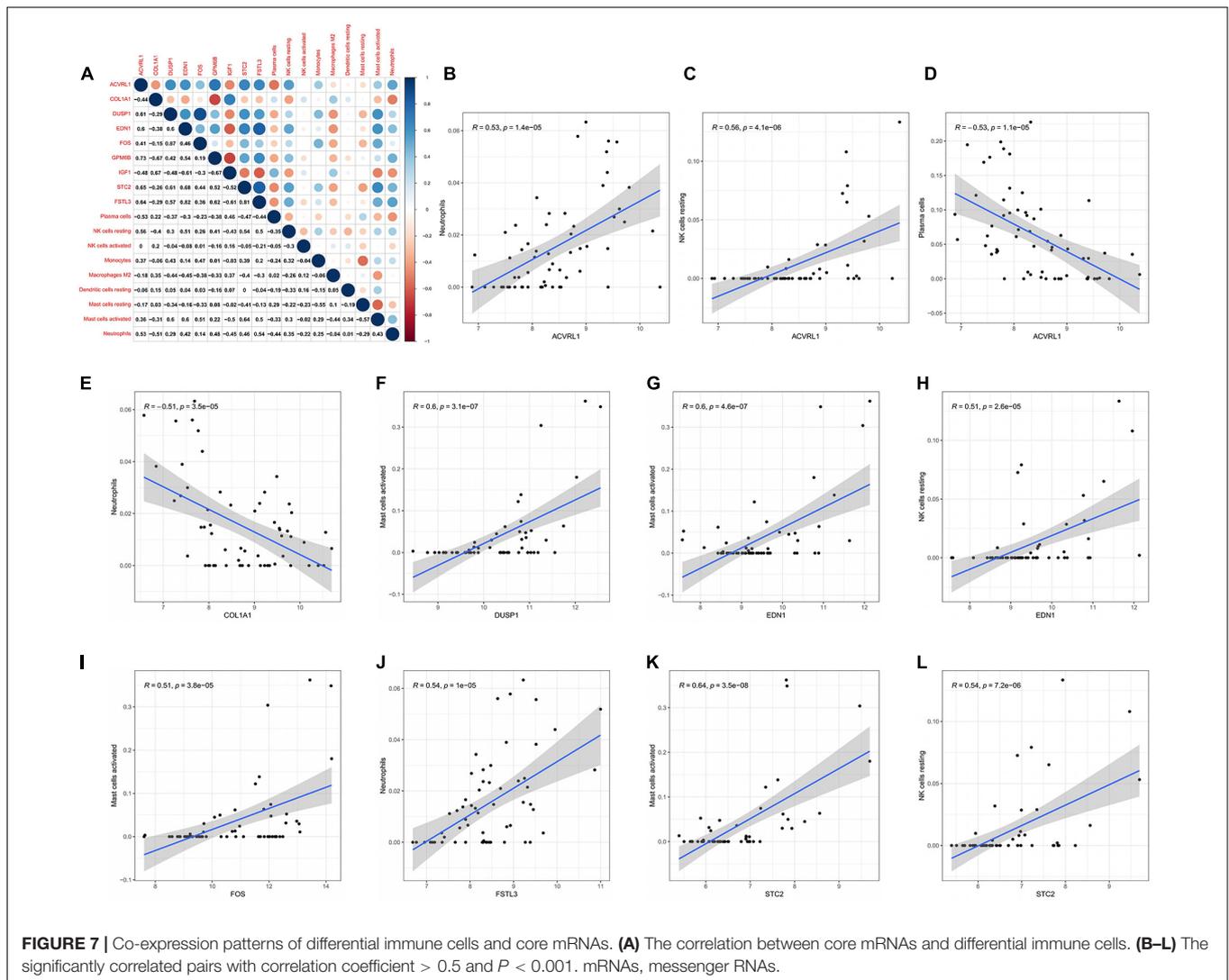
FIGURE 6 | Composition of infiltrating immune cells in lung tissues. **(A)** The distribution of immune cell types in each sample. **(B)** The correlation among immune cell types. **(C)** The heat map of infiltrating immune cells. **(D)** The violin plot of infiltrating immune cells.

chemokine C-X-C motif ligand 4 (CXCL4). To be mentioned, pDCs isolated from SSc patients perform the increased secretion of CXCL4, and the expression levels of CXCL4 in the plasma of SSc patients are correlated with the severity of fibrosis

(van Bon et al., 2014; Ah Kioon et al., 2018). Moreover, Kafaja et al. (2018) found that the pDC levels were increased in the lung tissues but reduced in the peripheral blood of SSc patients compared with normal controls. In the bleomycin-induced

FIGURE 7 | Co-expression patterns of differential immune cells and core mRNAs. **(A)** The correlation between core mRNAs and differential immune cells. **(B–L)** The significantly correlated pairs with correlation coefficient > 0.5 and $P < 0.001$. mRNAs, messenger RNAs.

fibrosis models, the depletion of pDCs could reduce the collagen deposition in lung and skin tissues and downregulate the expression of profibrotic genes, such as *TGFβ1* (Kafaja et al., 2018). Overall, pDCs are involved in the pathogenesis of tissue fibrosis in SSc and may act as a potential therapeutic target for SSc patients.

Mast cells have also been reported in the pathogenesis of tissue fibrosis, and the skin of SSc patients has a relatively increased number of mast cells (Seibold et al., 1990). In addition, the mast cell numbers are also elevated in lung tissues of patients with fibrotic lung diseases, such as idiopathic pulmonary fibrosis (IPF) (Pesci et al., 1993; Veerappan et al., 2013). Mast cells can not only produce but also transfer the TGF-β, thus participating in the process of fibrosis (Hugle et al., 2011). Besides, other immune cells have also been identified to be involved in the complex immunopathogenesis of SSc, such as NK cells and neutrophils (Barnes et al., 2012; Almeida et al., 2015; Benyamine et al., 2018). However, in this study, the analysis of immune cell infiltration in lung tissues of SSc-ILD patients

was based on the CIBERSORT algorithm, which only included 22 immune cell types. Considering the heterogeneity and complexity of the immune microenvironment, further studies are required to investigate the complete landscape of infiltrating immune cells in lung tissues of SSc-ILD. Nowadays, the rapid development of single-cell RNA-sequencing (scRNA-seq) technology provides us a novel method to explore the complex immune microenvironment. For instance, by conducting scRNA-seq analysis, Valenzi et al. (2019) identified multiple immune cell populations from 5 healthy control and 8 SSc-ILD lung tissue samples and demonstrated the fibroblast heterogeneity in SSc-ILD. Uncovering the heterogeneity of immune cell subpopulations in SSc-ILD and healthy lungs may provide new insights into the pathological mechanisms driving fibrosis in SSc-ILD.

In order to explore the potential regulatory mechanisms of genes on infiltration immune cells, we performed correlation analysis between core mRNAs and differential immune cells in lung tissues of SSc-ILD patients. The results showed that some

fibrosis-related genes were significantly related to infiltrating immune cells ($P < 0.001$ and correlation coefficient > 0.5). For instance, the expression of *ACVRL1* gene was positively correlated with neutrophils and resting NK cells but negatively correlated with plasma cells. Interestingly, the correlation analysis presented in the current study also showed that the expression of *ACVRL1* gene was negatively correlated with the expression of *COL1A1* gene ($R = -0.44$) and *IGF* gene ($R = -0.48$), although these three genes were all fibrosis-related genes. *ACVRL1* gene encodes activin receptor-like kinase 1 (ALK1), which is one of the type I cell surface receptors for TGF-β (Munoz-Felix et al., 2013). Previous studies have proved that TGF-β is a profibrotic cytokine involved in numerous fibrotic disorders including SSc. In most cell types, TGF-β plays its profibrotic roles via TGF-β/ALK5/Smad2/Smad3 signaling pathway, which is the canonical TGF-β signaling pathway (Verrecchia and Mauviel, 2007; Lafyatis, 2014). Specifically, the fibrotic reaction is characterized by the increased synthesis of ECM proteins including type I collagen, the activation of fibroblasts, and the epithelial-mesenchymal transition (EMT). Smad 3 is indispensable for these processes, and *Smad3*-knockout mice are resistant to TGF-β- and bleomycin-mediated lung fibrosis (Willis and Borok, 2007; Santibanez et al., 2011). Besides, recent studies have also demonstrated that TGF-β could participate in the process of fibrosis via another ALK1/Smad1/Smad5 signaling pathway. However, the role of this pathway in fibrosis remains controversial (Munoz-Felix et al., 2013). On the one hand, ALK1/Smad1/Smad5 signaling pathway can regulate the profibrotic phenotype in SSc fibroblasts, leading to the increased production of ECM components, such as type I collagen and connective tissue growth factor (CTGF) (Pannu et al., 2007, 2008). On the other hand, ALK1 signaling can oppose the ALK5 pathway and inhibit the synthesis of ECM proteins in endothelial cells (Goumans et al., 2003; Finnson et al., 2008; Vorstenbosch et al., 2017). Therefore, the regulatory effect of ALK1/Smad1/Smad5 signaling on fibrosis might depend on the types of organs and cells, as well as other circumstances. This may partly explain the negative correlation between *ACVRL1* gene and some other fibrosis-related genes, but the detailed mechanisms need to be further studied. In addition, other core mRNAs, such as *COL1A1*, *FOS*, and *EDN1*, were positively or negatively correlated with the number of infiltrating immune cells including neutrophils, NK cells and mast cells. This correlation analysis gives us a novel insight into the gene regulatory mechanisms on immune system, and further studies are needed to investigate them in depth.

However, there are several limitations in this study. First, this study was conducted only based on the GEO database. Although we integrated four different microarray datasets of SSc-ILD, the sample size was still limited, partly because of the low morbidity of SSc-ILD. And not all the patients with SSc-ILD experienced lung biopsy, which could be another reason for the small sample size. Second, no lncRNA microarray dataset related to SSc-ILD was existed in the GEO database, so the potential target lncRNAs of DEmiRs were predicted using the online tool, which needs to be experimentally validated. Third, the analysis of immune cell infiltration was based on the CIBERSORT algorithm, which

only included 22 immune cell types. Since the heterogeneity and complexity of the immune microenvironment were not taken into account, further studies are needed to investigate the complete landscape of infiltrating immune cells in lung tissues of SSc-ILD. Finally, considering that this is a pure bioinformatics analysis, the underlying regulatory mechanisms of the ceRNA network and gene-immune cells were not clearly elucidated, and further functional biological experiments with larger sample sizes are required.

To our best knowledge, this is the first comprehensive study to construct the potential ceRNA regulatory network and analyze the composition of infiltrating immune cells in lung tissues of SSc-ILD patients. Based on one miRNA and three mRNA microarray datasets, we identified DEmiRs and DEMs between SSc-ILD lung tissues and normal lung tissues, which were used for the construction of potential ceRNA regulatory network. Moreover, we also evaluated the differentially expressed infiltrating immune cells in lung tissues of SSc-ILD compared with normal controls and their correlation with core mRNAs, which provide new insights into gene-immune cells regulatory mechanisms in SSc-ILD. Our findings improve our understanding of the pathogenesis of SSc-ILD and may provide potential therapeutic targets for patients with SSc-ILD, which need to be further studied.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/**Supplementary Material**.

AUTHOR CONTRIBUTIONS

FUNDING

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcell.2021.622021/full#supplementary-material

REFERENCES

Abraham, D. J., Vancheeswaran, R., Dashwood, M. R., Rajkumar, V. S., Pantelides, P., Xu, S. W., et al. (1997). Increased levels of endothelin-1 and differential endothelin type A and B receptor expression in scleroderma-associated fibrotic lung disease. *Am. J. Pathol.* 151, 831–841.

Affandi, A. J., Carvalheiro, T., Radstake, T., and Marut, W. (2018). Dendritic cells in systemic sclerosis: Advances from human and mice studies. *Immunol. Lett.* 195, 18–29. doi: 10.1016/j.imlet.2017.11.003

Agarwal, V., Bell, G. W., Nam, J. W., and Bartel, D. P. (2015). Predicting effective microRNA target sites in mammalian mRNAs. *Elife* 4:e05005. doi: 10.7554/eLife.05005

Ah Kioon, M. D., Tripodo, C., Fernandez, D., Kirou, K. A., Spiera, R. F., Crow, M. K., et al. (2018). Plasmacytoid dendritic cells promote systemic sclerosis with a key role for TLR8. *Sci. Transl. Med.* 10:eaam8458. doi: 10.1126/scitranslmed.aam8458

Allanore, Y., Simms, R., Distler, O., Trojanowska, M., Pope, J., Denton, C. P., et al. (2015). Systemic sclerosis. *Nat. Rev. Dis. Primers* 1:15002. doi: 10.1038/nrdp.2015.2

Almeida, I., Silva, S. V., Fonseca, A. R., Silva, I., Vasconcelos, C., and Lima, M. (2015). T and NK Cell Phenotypic Abnormalities in Systemic Sclerosis: a Cohort Study and a Comprehensive Literature Review. *Clin. Rev. Allergy Immunol.* 49, 347–369. doi: 10.1007/s12016-015-8505-8

Angel, P., and Karin, M. (1991). The role of Jun, Fos and the AP-1 complex in cell-proliferation and transformation. *Biochim. Biophys. Acta* 1072, 129–157. doi: 10.1016/0304-419x(91)90011-9

Audiger, C., Rahman, M. J., Yun, T. J., Tarbell, K. V., and Lesage, S. (2017). The Importance of Dendritic Cells in Maintaining Immune Tolerance. *J. Immunol.* 198, 2223–2231. doi: 10.4049/jimmunol.1601629

Avouac, J., Palumbo, K., Tomcik, M., Zerr, P., Dees, C., Horn, A., et al. (2012). Inhibition of activator protein 1 signaling abrogates transforming growth factor beta-mediated activation of fibroblasts and prevents experimental fibrosis. *Arthritis Rheum.* 64, 1642–1652. doi: 10.1002/art.33501

Barnes, T. C., Anderson, M. E., Edwards, S. W., and Moots, R. J. (2012). Neutrophil-derived reactive oxygen species in SSc. *Rheumatology* 51, 1166–1169. doi: 10.1093/rheumatology/ker520

Bartel, D. P. (2004). MicroRNAs. *Cell* 116, 281–297. doi: 10.1016/s0092-8674(04)00045-5

Bartel, D. P. (2009). MicroRNAs: target recognition and regulatory functions. *Cell* 136, 215–233. doi: 10.1016/j.cell.2009.01.002

Benyamine, A., Magalon, J., Sabatier, F., Lyonnet, L., Robert, S., Dumoulin, C., et al. (2018). Natural Killer Cells Exhibit a Peculiar Phenotypic Profile in Systemic Sclerosis and Are Potent Inducers of Endothelial Microparticles Release. *Front. Immunol.* 9:1665. doi: 10.3389/fimmu.2018.01665

Brown, M., and O'Reilly, S. (2019). The immunopathogenesis of fibrosis in systemic sclerosis. *Clin. Exp. Immunol.* 195, 310–321. doi: 10.1111/cei.13238

Cech, T. R., and Steitz, J. A. (2014). The noncoding RNA revolution-trashing old rules to forge new ones. *Cell* 157, 77–94. doi: 10.1016/j.cell.2014.03.008

Chen, Y., and Wang, X. (2020). miRDB: an online database for prediction of functional microRNA targets. *Nucleic Acids Res.* 48, D127–D131. doi: 10.1093/nar/gkz757

Chia, J. J., and Lu, T. T. (2015). Update on macrophages and innate immunity in scleroderma. *Curr. Opin. Rheumatol.* 27, 530–536. doi: 10.1097/BOR.0000000000000218

Chin, C. H., Chen, S. H., Wu, H. H., Ho, C. W., Ko, M. T., and Lin, C. Y. (2014). cytoHubba: identifying hub objects and sub-networks from complex interactome. *BMC Syst. Biol.* 8(Suppl. 4):S11. doi: 10.1186/1752-0509-8-S4-S11

Christmann, R. B., Sampaio-Barros, P., Stifano, G., Borges, C. L., de Carvalho, C. R., Kairalla, R., et al. (2014). Association of Interferon- and transforming growth factor beta-regulated genes and macrophage activation with systemic sclerosis-related progressive lung fibrosis. *Arthritis Rheumatol.* 66, 714–725. doi: 10.1002/art.38288

Christmann, R. B., Wooten, A., Sampaio-Barros, P., Borges, C. L., Carvalho, C. R., Kairalla, R. A., et al. (2016). miR-155 in the progression of lung fibrosis in systemic sclerosis. *Arthritis Res. Ther.* 18:155. doi: 10.1186/s13075-016-1054-6

Ciechomska, M., van Laar, J., and O'Reilly, S. (2015). Current frontiers in systemic sclerosis pathogenesis. *Exp. Dermatol.* 24, 401–406. doi: 10.1111/exd.12673

Denton, C. P., and Khanna, D. (2017). Systemic sclerosis. *Lancet* 390, 1685–1699. doi: 10.1016/s0140-6736(17)30933-9

Esteller, M. (2011). Non-coding RNAs in human disease. *Nat. Rev. Genet.* 12, 861–874. doi: 10.1038/nrg3074

Finnson, K. W., Parker, W. L., ten Dijke, P., Thorikay, M., and Philip, A. (2008). ALK1 opposes ALK5/Smad3 signaling and expression of extracellular matrix components in human chondrocytes. *J. Bone Miner Res.* 23, 896–906. doi: 10.1359/jbmr.080209

Francois, A., Chatelus, E., Wachsmann, D., Sibilia, J., Bahram, S., Alsaleh, G., et al. (2013). B lymphocytes and B-cell activating factor promote collagen and profibrotic markers expression by dermal fibroblasts in systemic sclerosis. *Arthritis Res. Ther.* 15:R168. doi: 10.1186/ar4352

Furue, M., Mitoma, C., Mitoma, H., Tsuji, G., Chiba, T., Nakahara, T., et al. (2017). Pathogenesis of systemic sclerosis-current concept and emerging treatments. *Immunol. Res.* 65, 790–797. doi: 10.1007/s12026-017-8926-y

Glickman, M. E., Rao, S. R., and Schultz, M. R. (2014). False discovery rate control is a recommended alternative to Bonferroni-type adjustments in health studies. *J. Clin. Epidemiol.* 67, 850–857. doi: 10.1016/j.jclinepi.2014.03.012

Goumans, M.-J., Valdimarsdottir, G., Itoh, S., Lebrin, F., Larsson, J., Mummery, C., et al. (2003). Activin Receptor-like Kinase (ALK)1 Is an Antagonistic Mediator of Lateral TGFβ/ALK5 Signaling. *Mol. Cell* 12, 817–828. doi: 10.1016/s1097-2765(03)00386-1

Guevara-Aguirre, J. (1996). Insulin-like growth factor I–an important intrauterine growth factor. *N. Engl. J. Med.* 335, 1389–1391. doi: 10.1056/NEJM199610313351810

Hamaguchi, Y., Fujimoto, M., Matsushita, T., Hasegawa, M., Takehara, K., and Sato, S. (2008). Elevated serum insulin-like growth factor (IGF-1) and IGF binding protein-3 levels in patients with systemic sclerosis: possible role in development of fibrosis. *J. Rheumatol.* 35, 2363–2371. doi: 10.3899/jrheum.080340

Heemskerk, V. H., Daemen, M. A., and Buurman, W. A. (1999). Insulin-like growth factor-1 (IGF-1) and growth hormone (GH) in immunity and inflammation. *Cytokine Growth Factor Rev.* 10, 5–14. doi: 10.1016/s1359-6101(98)00022-7

Henry, T. W., Mendoza, F. A., and Jimenez, S. A. (2019). Role of microRNA in the pathogenesis of systemic sclerosis tissue fibrosis and vasculopathy. *Autoimmun. Rev.* 18:102396. doi: 10.1016/j.autrev.2019.102396

Hitraya, E. G., and Jimenez, S. A. (1996). Transcriptional activation of the alpha 1(I) procollagen gene in systemic sclerosis dermal fibroblasts. Role of intronic sequences. *Arthritis Rheum.* 39, 1347–1354. doi: 10.1002/art.1780390812

Hsu, E., Shi, H., Jordan, R. M., Lyons-Weiler, J., Pilewski, J. M., and Feghali-Bostwick, C. A. (2011). Lung tissues in patients with systemic sclerosis have gene expression patterns unique to pulmonary fibrosis and pulmonary hypertension. *Arthritis Rheum.* 63, 783–794. doi: 10.1002/art.30159

Hu, H. H., Chen, D. Q., Wang, Y. N., Feng, Y. L., Cao, G., Vaziri, N. D., et al. (2018). New insights into TGF-beta/Smad signaling in tissue fibrosis. *Chem. Biol. Interact.* 292, 76–83. doi: 10.1016/j.cbi.2018.07.008

Huang, H. Y., Lin, Y. C., Li, J., Huang, K. Y., Shrestha, S., Hong, H. C., et al. (2020). miRTarBase 2020: updates to the experimentally validated microRNA-target interaction database. *Nucleic Acids Res.* 48, D148–D154. doi: 10.1093/nar/gkz896

Hugle, T., Hogan, V., White, K. E., and van Laar, J. M. (2011). Mast cells are a source of transforming growth factor beta in systemic sclerosis. *Arthritis Rheum.* 63, 795–799. doi: 10.1002/art.30190

Jimenez, S. A., and Saitta, B. (1999). Alterations in the regulation of expression of the alpha 1(I) collagen gene (COL1A1) in systemic sclerosis (scleroderma). *Springer Semin Immunopathol.* 21, 397–414. doi: 10.1007/BF00870302

Kafaja, S., Valera, I., Divekar, A. A., Saggar, R., Abtin, F., Furst, D. E., et al. (2018). pDCs in lung and skin fibrosis in a bleomycin-induced model and patients with systemic sclerosis. *JCI Insight* 3:98380. doi: 10.1172/jci.insight.98380

Kalogerou, A., Gelou, E., Mountantonakis, S., Settas, L., Zafiriou, E., and Sakkas, L. (2005). Early T cell activation in the skin from patients with systemic sclerosis. *Ann. Rheum. Dis.* 64, 1233–1235. doi: 10.1136/ard.2004.027094

Kulasekaran, P., Scavone, C. A., Rogers, D. S., Arenberg, D. A., Thannickal, V. J., and Horowitz, J. C. (2009). Endothelin-1 and transforming growth factor-beta1 independently induce fibroblast resistance to apoptosis via AKT activation. *Am. J. Respir. Cell Mol. Biol.* 41, 484–493. doi: 10.1165/rcmb.2008-0447OC

Lafyatis, R. (2014). Transforming growth factor beta–at the centre of systemic sclerosis. *Nat. Rev. Rheumatol.* 10, 706–719. doi: 10.1038/nrrheum.2014.137

Leek, J. T., Johnson, W. E., Parker, H. S., Fertig, E. J., Jaffe, A. E., Zhang, Y., et al. (2020). *sva: Surrogate Variable Analysis. R package version 3.38.0.* Vienna: R Core Team.

Li, X., Wu, Z., Fu, X., and Han, W. (2013). Long Noncoding RNAs: Insights from Biological Features and Functions to Diseases. *Med. Res. Rev.* 33, 517–553. doi: 10.1002/med.21254

Lu, T. T. (2012). Dendritic cells: novel players in fibrosis and scleroderma. *Curr. Rheumatol. Rep.* 14, 30–38. doi: 10.1007/s11926-011-0215-5

Makino, K., Jinnin, M., Hirano, A., Yamane, K., Eto, M., Kusano, T., et al. (2013). The downregulation of microRNA let-7a contributes to the excessive expression of type I collagen in systemic and localized scleroderma. *J. Immunol.* 190, 3905–3915. doi: 10.4049/jimmunol.1200822

Manetti, M. (2015). Deciphering the alternatively activated (M2) phenotype of macrophages in scleroderma. *Exp. Dermatol.* 24, 576–578. doi: 10.1111/exd.12727

Maurer, B., Stanczyk, J., Jungel, A., Akhmetshina, A., Trenkmann, M., Brock, M., et al. (2010). MicroRNA-29, a key regulator of collagen expression in systemic sclerosis. *Arthritis Rheum.* 62, 1733–1743. doi: 10.1002/art.27443

Mercer, T. R., Dinger, M. E., and Mattick, J. S. (2009). Long non-coding RNAs: insights into functions. *Nat. Rev. Genet.* 10, 155–159. doi: 10.1038/nrg2521

Munoz-Felix, J. M., Gonzalez-Nunez, M., and Lopez-Novoa, J. M. (2013). ALK1-Smad1/5 signaling pathway in fibrosis development: friend or foe? *Cytokine Growth Fact. Rev.* 24, 523–537. doi: 10.1016/j.cytogfr.2013.08.002

Nakasa, T., Miyaki, S., Okubo, A., Hashimoto, M., Nishida, K., Ochi, M., et al. (2008). Expression of microRNA-146 in rheumatoid arthritis synovial tissue. *Arthritis Rheum.* 58, 1284–1292. doi: 10.1002/art.23429

Newman, A. M., Liu, C. L., Green, M. R., Gentles, A. J., Feng, W., Xu, Y., et al. (2015). Robust enumeration of cell subsets from tissue expression profiles. *Nat. Methods* 12, 453–457. doi: 10.1038/nmeth.3337

Pannu, J., Asano, Y., Nakerakanti, S., Smith, E., Jablonska, S., Blaszczyk, M., et al. (2008). Smad1 pathway is activated in systemic sclerosis fibroblasts and is targeted by imatinib mesylate. *Arthritis Rheum.* 58, 2528–2537. doi: 10.1002/art.23698

Pannu, J., Nakerakanti, S., Smith, E., ten Dijke, P., and Trojanowska, M. (2007). Transforming growth factor-beta receptor type I-dependent fibrogenic gene program is mediated via activation of Smad1 and ERK1/2 pathways. *J. Biol. Chem.* 282, 10405–10413. doi: 10.1074/jbc.M611742200

Paraskevopoulou, M. D., Vlachos, I. S., Karagkouni, D., Georgakilas, G., Kanellos, I., Vergoulis, T., et al. (2016). DIANA-LncBase v2: indexing microRNA targets on non-coding transcripts. *Nucleic Acids Res.* 44, D231–D238. doi: 10.1093/nar/gkv1270

Pathan, M., Keerthikumar, S., Ang, C. S., Gangoda, L., Quek, C. Y., Williamson, N. A., et al. (2015). FunRich: An open access standalone functional enrichment and interaction network analysis tool. *Proteomics* 15, 2597–2601. doi: 10.1002/pmic.201400515

Pechkovsky, D. V., Prasse, A., Kollert, F., Engel, K. M., Dentler, J., Luttmann, W., et al. (2010). Alternatively activated alveolar macrophages in pulmonary fibrosis-mediator production and intracellular signal transduction. *Clin. Immunol.* 137, 89–101. doi: 10.1016/j.clim.2010.06.017

Perelas, A., Silver, R. M., Arrossi, A. V., and Highland, K. B. (2020). Systemic sclerosis-associated interstitial lung disease. *Lancet Respir. Med.* 8, 304–320. doi: 10.1016/s2213-2600(19)30480-1

Pesci, A., Bertorelli, G., Gabrielli, M., and Olivieri, D. (1993). Mast cells in fibrotic lung disorders. *Chest* 103, 989–996. doi: 10.1378/chest.103.4.989

Piera-Velazquez, S., Fertala, J., Huaman-Vargas, G., Louneva, N., and Jimenez, S. A. (2020). Increased expression of the transforming growth factor beta-inducible gene HIC-5 in systemic sclerosis skin and fibroblasts: a novel antifibrotic therapeutic target. *Rheumatology* 59, 3092–3098. doi: 10.1093/rheumatology/keaa200

Renaud, L., da Silveira, W. A., Takamura, N., Hardiman, G., and Feghali-Bostwick, C. (2020). Prominence of IL6, IGF, TLR, and Bioenergetics Pathway Perturbation in Lung Tissues of Scleroderma Patients With Pulmonary Fibrosis. *Front. Immunol.* 11:383. doi: 10.3389/fimmu.2020.00383

Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., et al. (2015). limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* 43:gkv007. doi: 10.1093/nar/gkv007

Salmena, L., Poliseno, L., Tay, Y., Kats, L., and Pandolfi, P. P. (2011). A ceRNA hypothesis: the Rosetta Stone of a hidden RNA language? *Cell* 146, 353–358. doi: 10.1016/j.cell.2011.07.014

Sandorfi, N., Louneva, N., Hitraya, E., Hajnoczky, G., Saitta, B., and Jimenez, S. A. (2005). Inhibition of collagen gene expression in systemic sclerosis dermal fibroblasts by mithramycin. *Ann. Rheum. Dis.* 64, 1685–1691. doi: 10.1136/ard.2005.037515

Santibanez, J. F., Quintanilla, M., and Bernabeu, C. (2011). TGF-beta/TGF-beta receptor system and its role in physiological and pathological conditions. *Clin. Sci.* 121, 233–251. doi: 10.1042/CS20110086

Seibold, J. R., Giorno, R. C., and Claman, H. N. (1990). Dermal mast cell degranulation in systemic sclerosis. *Arthritis Rheum.* 33, 1702–1709. doi: 10.1002/art.1780331114

Shapouri-Moghaddam, A., Mohammadian, S., Vazini, H., Taghadosi, M., Esmaeili, S. A., Mardani, F., et al. (2018). Macrophage plasticity, polarization, and function in health and disease. *J. Cell Physiol.* 233, 6425–6440. doi: 10.1002/jcp.26429

Shi, X., Sun, M., Liu, H., Yao, Y., and Song, Y. (2013). Long non-coding RNAs: a new frontier in the study of human diseases. *Cancer Lett.* 339, 159–166. doi: 10.1016/j.canlet.2013.06.013

Shi, Y., Tang, B., Yu, J., Luo, Y., Xiao, Y., Pi, Z., et al. (2020). Aryl hydrocarbon receptor signaling activation in systemic sclerosis attenuates collagen production and is a potential antifibrotic target. *Int. Immunopharmacol.* 88:106886. doi: 10.1016/j.intimp.2020.106886

Shi-Wen, X., Chen, Y., Denton, C. P., Eastwood, M., Renzoni, E. A., Bou-Gharios, G., et al. (2004). Endothelin-1 promotes myofibroblast induction through the ETA receptor via a rac/phosphoinositide 3-kinase/Akt-dependent pathway and is essential for the enhanced contractile phenotype of fibrotic fibroblasts. *Mol. Biol. Cell* 15, 2707–2719. doi: 10.1091/mbc.e03-12-0902

Shi-Wen, X., Renzoni, E. A., Kennedy, L., Howat, S., Chen, Y., Pearson, J. D., et al. (2007). Endogenous endothelin-1 signaling contributes to type I collagen and CCN2 overexpression in fibrotic fibroblasts. *Matrix. Biol.* 26, 625–632. doi: 10.1016/j.matbio.2007.06.003

Soldano, S., Montagna, P., Villaggio, B., Parodi, A., Gianotti, G., Sulli, A., et al. (2009). Endothelin and sex hormones modulate the fibronectin synthesis by cultured human skin scleroderma fibroblasts. *Ann. Rheum. Dis.* 68, 599–602. doi: 10.1136/ard.2008.097378

Stewart, C. E., and Rotwein, P. (1996). Growth, differentiation, and survival: multiple physiological functions for insulin-like growth factors. *Physiol. Rev.* 76, 1005–1026. doi: 10.1152/physrev.1996.76.4.1005

Stifano, G., and Christmann, R. B. (2016). Macrophage Involvement in Systemic Sclerosis: Do We Need More Evidence? *Curr. Rheumatol. Rep.* 18:2. doi: 10.1007/s11926-015-0554-8

Sullivan, D. E., Ferris, M., Nguyen, H., Abboud, E., and Brody, A. R. (2009). TNF-alpha induces TGF-beta1 expression in lung fibroblasts at the transcriptional level via AP-1 activation. *J. Cell Mol. Med.* 13, 1866–1876. doi: 10.1111/j.1582-4934.2009.00647.x

Swigris, J. J., and Brown, K. K. (2010). The role of endothelin-1 in the pathogenesis of idiopathic pulmonary fibrosis. *BioDrugs* 24, 49–54. doi: 10.2165/11319550-000000000-00000

Szklarczyk, D., Gable, A. L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., et al. (2019). STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.* 47, D607–D613. doi: 10.1093/nar/gky1131

Tabata, K., Mikita, N., Yasutake, M., Matsumiya, R., Tanaka, K., Tani, S., et al. (2021). Up-regulation of IGF-1, RANTES and VEGF in patients with anti-centromere antibody-positive early/mild systemic sclerosis. *Mod. Rheumatol.* 31, 171–176. doi: 10.1080/14397595.2020.1726599

Tay, Y., Rinn, J., and Pandolfi, P. P. (2014). The multilayered complexity of ceRNA crosstalk and competition. *Nature* 505, 344–352. doi: 10.1038/nature12986

Valencia-Sanchez, M. A., Liu, J., Hannon, G. J., and Parker, R. (2006). Control of translation and mRNA degradation by miRNAs and siRNAs. *Genes Dev.* 20, 515–524. doi: 10.1101/gad.1399806

Valenzi, E., Bulik, M., Tabib, T., Morse, C., Sembrat, J., Trejo Bittar, H., et al. (2019). Single-cell analysis reveals fibroblast heterogeneity and myofibroblasts in systemic sclerosis-associated interstitial lung disease.

Ann. Rheum. Dis. 78, 1379–1387. doi: 10.1136/annrheumdis-2018-214865

van Bon, L., Affandi, A. J., Broen, J., Christmann, R. B., Marijnissen, R. J., Stawski, L., et al. (2014). Proteome-wide analysis and CXCL4 as a biomarker in systemic sclerosis. *N. Engl. J. Med.* 370, 433–443. doi: 10.1056/NEJMoa1114576

Veerappan, A., O'Connor, N. J., Brazin, J., Reid, A. C., Jung, A., McGee, D., et al. (2013). Mast cells: a pivotal role in pulmonary fibrosis. *DNA Cell Biol.* 32, 206–218. doi: 10.1089/dna.2013.2005

Verrecchia, F., and Mauviel, A. (2007). Transforming growth factor-beta and fibrosis. *World J. Gastroenterol.* 13, 3056–3062. doi: 10.3748/wjg.v13.i22.3056

Vorstenbosch, J., Nguyen, C. M., Zhou, S., Seo, Y. J., Siblini, A., Finnson, K. W., et al. (2017). Overexpression of CD109 in the Epidermis Differentially Regulates ALK1 Versus ALK5 Signaling and Modulates Extracellular Matrix Synthesis in the Skin. *J. Invest. Dermatol.* 137, 641–649. doi: 10.1016/j.jid.2016.09.039

Wang, S. H., Ma, F., Tang, Z. H., Wu, X. C., Cai, Q., Zhang, M. D., et al. (2016). Long non-coding RNA H19 regulates FOXM1 expression by competitively binding endogenous miR-342-3p in gallbladder cancer. *J. Exp. Clin. Cancer Res.* 35:160. doi: 10.1186/s13046-016-0436-6

Willis, B. C., and Borok, Z. (2007). TGF-beta-induced EMT: mechanisms and implications for fibrotic lung disease. *Am. J. Physiol. Lung Cell Mol. Physiol.* 293, L525–L534. doi: 10.1152/ajplung.00163.2007

Winsz-Szczotka, K., Kuznik-Trocha, K., Komosinska-Vassev, K., Kucharz, E., Kotulska, A., and Olczyk, K. (2016). Relationship between adiponectin, leptin, IGF-1 and total lipid peroxides plasma concentrations in patients with systemic sclerosis: possible role in disease development. *Int. J. Rheum. Dis.* 19, 706–714. doi: 10.1111/1756-185X.12332

Young, N. A., Valiente, G. R., Hampton, J. M., Wu, L. C., Burd, C. J., Willis, W. L., et al. (2017). Estrogen-regulated STAT1 activation promotes TLR8 expression to facilitate signaling via microRNA-21 in systemic lupus erythematosus. *Clin. Immunol.* 176, 12–22. doi: 10.1016/j.clim.2016.12.005

Yu, G., Wang, L. G., Han, Y., and He, Q. Y. (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* 16, 284–287. doi: 10.1089/omi.2011.0118

Zhang, K., Li, Q., Kang, X., Wang, Y., and Wang, S. (2016). Identification and functional characterization of lncRNAs acting as ceRNA involved in the malignant progression of glioblastoma multiforme. *Oncol. Rep.* 36, 2911–2925. doi: 10.3892/or.2016.5070