Matrix metalloproteinase 7 is a candidate biomarker in systemic sclerosis-associated interstitial lung disease

Xu Z[1], Chen W[2], Chen C[1]

ACTA REUMATOL PORT. 2020;45:191-200

ABSTRACT

Background: Pulmonary complications, including pulmonary fibrosis, are the leading causes of death in systemic sclerosis (SSc). However, the aetiology and pathophysiologic mechanisms of the disease have not been comprehensively investigated, and drugs for treating systemic sclerosis-associated interstitial lung disease (SSc-ILD) are limited. The objective of this study was to identify key novel genes and pathways linked to SSc-ILD and decipher the molecular mechanisms involved in the disease.

Methods: We compared three microarray datasets in the GEO database including 42 SSc-ILD samples and 18 normal samples to obtain differentially expressed genes (DEGs). Gene Ontology (GO) analysis and the Kyoto Encyclopaedia of Genes and Genomes (KEGG) pathway analysis were performed, and a protein-protein interaction network was constructed. After validation, gene set enrichment analysis (GSEA) was applied to obtain further insights into the function of the selected hub genes.

Results: A total of 25 DEGs were filtered. The GO analysis revealed genes that were mainly enriched in immune response, chemokine activity, and extracellular regions. KEGG pathway analysis of the DEGs revealed that SSc-ILD was associated with the tumour necrosis factor (TNF) signalling pathway and cytokine-cytokine receptor interaction. Matrix metallopeptidase 7 (MMP7) expression was consistently increased in all the three datasets, and results of the GSEA indicated that MMP7 might play a role in the regulation of the G-protein coupled amine receptor activity.

Conclusions: In summary, the novel DEGs, especially MMP7 and the SSc-ILD pathway genes identified in this study might provide further insights into potential molecular mechanism of the disease.

Keywords: Interstitial lung disease; Systemic sclerosis; Biomarker.

1. Department of Pulmonary and Critical Care Medicinet, The First Affiliated Hospital of Wenzhou Medical University;
2. Endocrinology department, The First Affiliated Hospital of Wenzhou Medical University.

INTRODUCTION

Systemic sclerosis (SSc) is a rare chronic progressive autoimmune disease of the connective tissue, characterized by immune dysfunction, vascular disease, cellular inflammation, and skin fibrosis[1]. SSc could affect multiple internal organs leading to disorders, such as interstitial lung disease (ILD)[1, 2], but in SSc, the lung involvement might occur without skin involvement[3]. Likewise, ILD is a group of diseases associated with connective tissue disease, which is regarded as a potential inflammation caused by the thickening of the interstitial tissue around the alveolar wall[4]. According to a review by Denton and Khanna (2017), four-fifths of SSc patients have pulmonary fibrosis or interstitial lung disease, but only 25-30% develop progressive ILD[5], often accompanied by a histopathological pattern of non-specific interstitial pneumonia (NSIP)[6]. The results of the EULAR Scleroderma Trials and Research (EUSTAR) cohort showed that lung diseases were the main cause of death in SSc patients, with pulmonary fibrosis, accounting for more than one-third of mortality[7]. In a Multinational Systemic Sclerosis Inception Cohort study, including patients from Australia, Canada, and Spain, ILD accounted for 20.7% of SSc-related deaths[8]. In addition to increasing the utilization of medical resources and related direct costs, systemic sclerosis-associated interstitial lung disease (SSc-ILD) also placed a heavy burden on the medical system with patients having poor health-related quality of life (HRQoL) in Australia[9]. The course of SSc-ILD from localized pulmonary infiltration to severe disease might usually be non-pro-

gressive, with the event occurring primarily in the first years of the disease onset, which might eventually progress to respiratory failure or death[10].

Due to the widespread use of angiotensin-converting enzyme (ACE) inhibitors, the mortality of scleroderma nephropathy and its sequelae might be markedly reduced[11], but effective SSc-ILD treatment is still being explored. Although many SSc-ILD treatments have been evaluated, the current management is largely limited to immune regulation, and non-selective immunosuppressants are still the most commonly used[10]. However, non-selective immunosuppressants, particularly cyclophosphamide, might have been linked to acute toxic accidents such as hematopoietic suppression[12], and could increase the risk of diseases like other rheumatism, cancer, and gonadal failure with cumulative doses[13]. Although mycophenolate mofetil could be identified as a safe drug, unfortunately it has a long-term effect, with unaffordable management costs for many patients in limited resourced economies [14]. Likewise, biologicals such as belimumab, tocilizumab, and abatacept appeared promising as they exhibited some skin benefit[15], while Nintedanib showed a positive effect in ILD[16].

High-throughput platform-based microarrays have become a reliable and efficient tool for exploring vital genetic or epigenetic changes in disease and identifying promising biomarkers for disease diagnosis and prognosis. Efficient microarray and bioinformatics analysis could help unravel the molecular mechanisms of pathogenesis and disease progression. This is a recommended pathway to explore genetic variation and identify potential diagnostic biomarkers. Therefore, we aimed to identify differentially expressed genes (DEGs) and pathways, as well as decipher the potential molecular mechanisms involved in SSc-ILD.

MATERIALS AND METHODS

GENE EXPRESSION DATA AND DATA PREPROCESSING

The transcription profiles for the lung tissue of SSc-ILD and normal samples dataset were retrieved from the NCBI GEO database (http://www.NCBI.nlm.nih.gov/GEO/) using the keywords “systemic sclerosis interstitial lung disease”, “systemic sclerosis”, and “interstitial lung disease”. “Homo sapiens”, “lung tissue” and “Expression profiling by array” were included in the next round of screening. The dataset included GSE48149, GSE76808, and GSE81292. The microarray data of GSE48149 was based on GPL16221 platform [Illumina HumanRef-8 v3.0 expression bead chip (Search Key version)], GSE81292 was based on GPL18991 platform [HG-U133A_2; Affymetrix Human Genome U133A 2.0 Array; HGU133A2_Hs_ENTREZG_16.0.0], and GSE76808 microarray data was based on GPL571 platform [HG-U133A_2; Affymetrix Human Genome U133A 2.0 Array]. The extracted datasets were from the total lung RNA and were for the RNA gene expression profiling. The samples included 42 SSc-ILD lung tissues and 18 normal controls (all from cancer-free patients).

The R software (version 3.4.2; https://www.R-project.org/) was used for data mining and statistical analysis. If multiple probes corresponded to the same gene symbol, the “aggregation” package (https://CRAN.R-project.org/package=aggregation) in R was used as the expression value of that specific gene to calculate the average. If there was no expression value for the probe, it was supplemented with the nearest neighbour mean (KNN) of the impute package[17] in R. Finally, the “Limma” package[18] was used for background correction and normalization.

IDENTIFICATION OF DEGS

The DEGs for the SSc-ILD samples compared with the normal controls were screened using the t-test method in the R “Limma” package. The $|\log2\ FC| > 1$ and the adjusted P-values <0.05 were selected as the cut-off criteria. Also, we compared 10 samples with SSc without ILD in the GSE48149 with the normal controls.

DEGS ENRICHMENT ANALYSIS

The Database for Annotation, Visualization, and Integrated Discovery (DAVID) 6.8 (https://david.ncifcrf.gov/)[19] was used for Gene Ontology (GO) analysis and the Kyoto Encyclopaedia of Genes and Genomes (KEGG) pathway analysis. The GO analysis included molecular function (MF), biological processes (BP), and cellular component (CC). The KOBAS 3.0[20] (http://KOBAS.cbi.pku.edu.cn) was used to evaluate the KEGG pathway analysis of the DEGs.

PROTEIN-PROTEIN INTERACTION (PPI) NETWORK ANALYSIS

The PPI network analysis is a crucial tool for understanding biological responses in health and disease. DEGs were imported into the STRING (version: 10.5) online database (http://string-db.org), a biological

database and network resource of protein-protein interaction, to analyse the functional interactions between proteins (parameters were the default in the STRING database). This was combined with the GeneMANIA online platform (http://GeneMANIA.org), and the file was subsequently imported into the Cytoscape software (version 3.7.1; http://www.Cytoscape.org/). The hub gene was identified through the CytoHubba plug with a maximum degree.

HUB GENES VALIDATION

The hub genes of interest were further validated in the GSE48149, GSE81292, and GSE76808. Additionally, the curve was plotted with the "ggstatsplot" package (https://CRAN.R-project.org/package=ggstatsplot) to evaluate the capability of the selected genes to distinguish the SSc-ILD patients and controls.

GENE SET ENRICHMENT ANALYSIS

To investigate the potential functions of the selected hub genes in SSc-ILD, the gene set enrichment analysis (GSEA) was performed. The GSEA is a free chip data analysis tool, which is based on the existing genesets. The median expression level of the hub gene was divided into two groups in the three datasets. The c5.all.v7.0.symbols.gmt was selected as the reference gene setting in the Molecular Signatures Database (MSigDB), and P adjusted value <0.05, while the FDR < 0.25 was selected as the cut-off standard.

RESULTS

IDENTIFICATION OF DEGS

In the SSc-ILD compared with normal samples, the DEGs were identified in the three transcript profiles obtained from the GEO database, including GSE48149, GSE81292 and GSE76808. There were 25 abnormally expressed DEGs, including "TNFAIP3", "CDH3", "LIF", "CX3CL1", "APOLD1", "IL13RA2", "MMP19", "CCL2", "SLCO4A1", "IL1RL1", "CHI3L2", "CA4", "COL15A1", "MEOX1", "PTGS2", "DUSP5", "COL3A1", "HYAL1", "RGS1", "MMP7", "IGFBP2", "AGER", "EDNRB", "CXCL14", "VIPR1" (Figure 1A). Likewise, 9 DEGs ("TNFAIP3", "IL13RA2", "DUSP5", "IL1RL1", "APOLD1", "RGS1", "IGFBP2", "COL15A1", "LIF") which were also abnormally expressed were found in the SSc without ILD samples (Figure 1B). The 25 DEGs were used for subsequent analysis because the SSc without ILD samples came from a single dataset. The calculation criteria were P <0.05 and the absolute value of log2FC> 1.

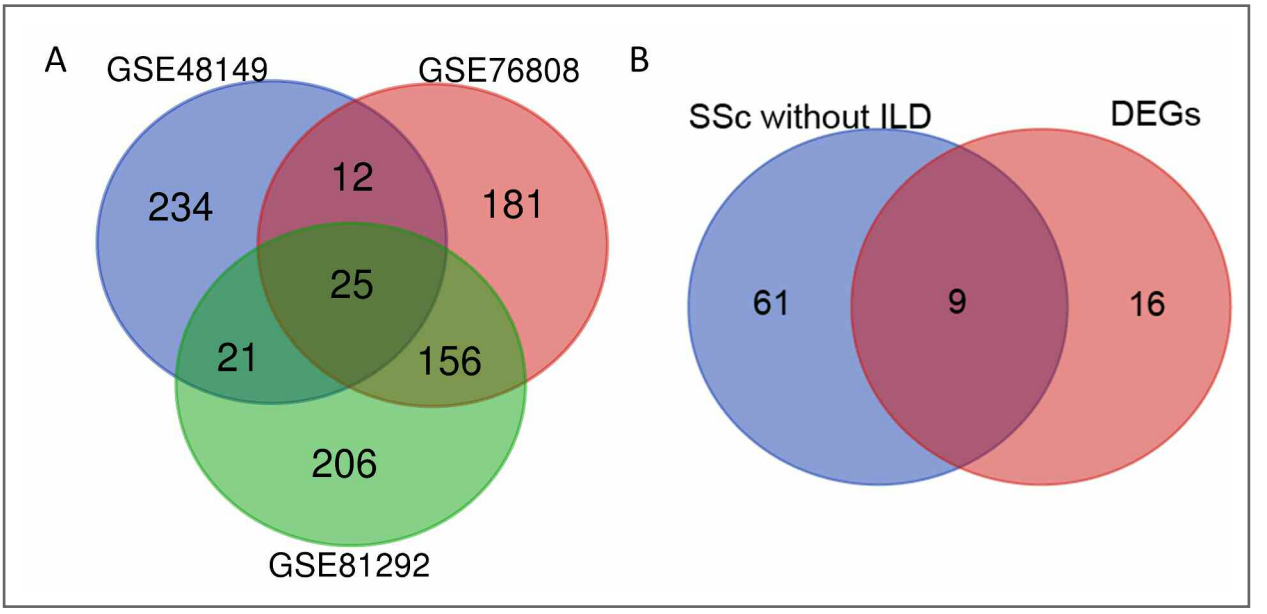

FIGURE 1. (A) The DEGs between the SSc-ILD and normal samples using Bioconductor package VennDiagram. (B) The genes between the SSc without ILD and the DEGs (between SSc-ILD and normal samples). The DEGs were identified with the classical t-test. Statistically significant DEGs were defined at P < 0.05 and absolute log2 fold change > 1 as the cut-off criterion. GSE,gene expression series.

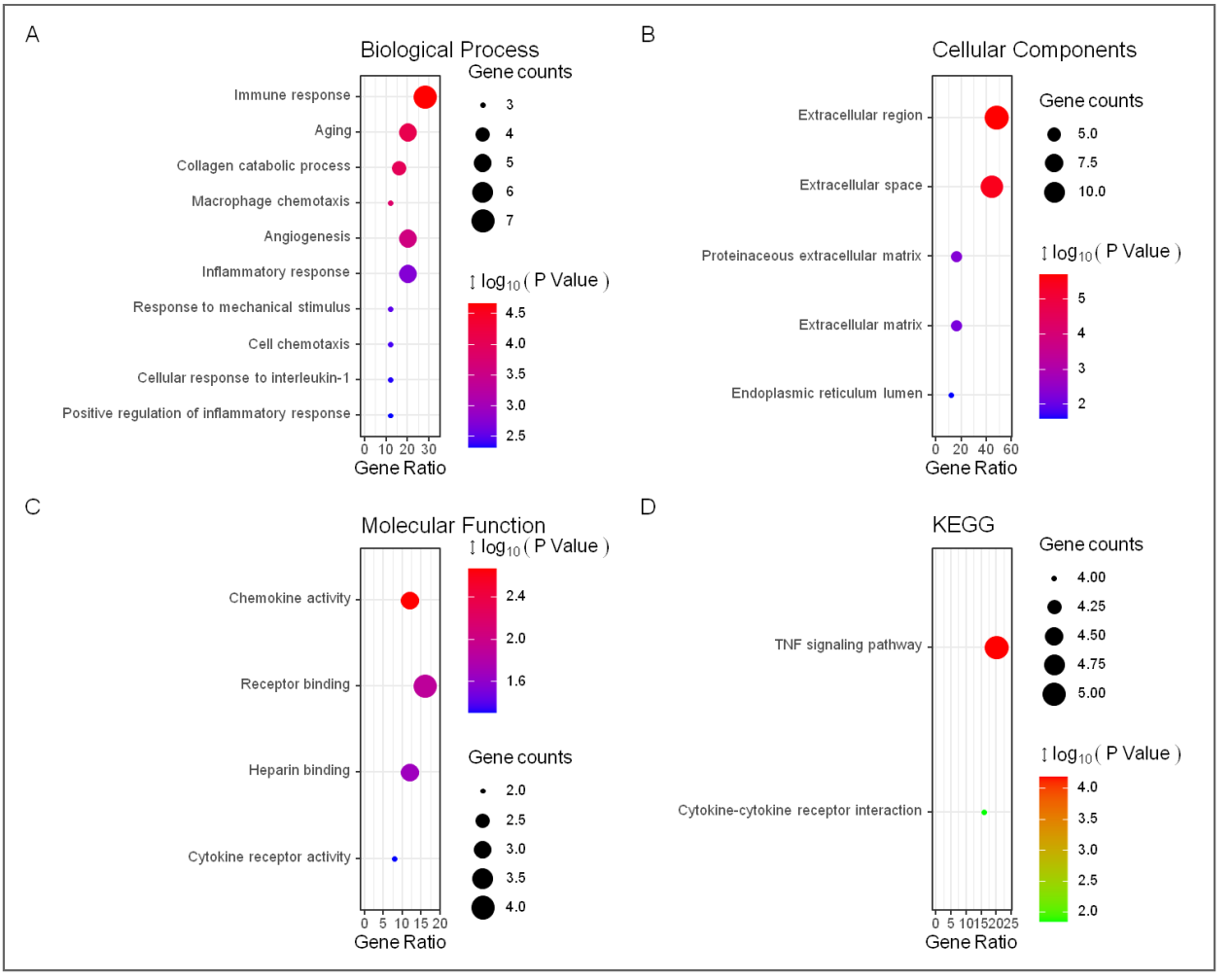

FIGURE 2. The GO analyses of the common DEGs. (A) The GO analyses according to the biological process. (B) The GO analyses according to cellular component. (C) The GO analyses according to molecular function. (D) The Kyoto Encyclopaedia of Genes and Genomes analyses of the common DEGs to identify the canonical pathways.TNF, tumour necrosis factor.

KEGG AND GO ENRICHMENT ANALYSES OF DEGS

The enrichment analyses of 25 DEGs were carried out to unravel the biological classification of DEGs using DAVID (Figure 2 and Table I). The results of the GO analysis showed that variations in the BP of common DEGs were significantly enriched in immune response, aging, collagen catabolic process, macrophage chemotaxis, angiogenesis, and inflammatory response (Figure 2A). For the CC, the DEGs were observably enriched in the extracellular region, extracellular space, and proteinaceous extracellular matrix (Figure 2B). Changes in the MF were mainly enriched in chemokine activity, receptor binding, and heparin-binding (Figure 2C). The analysis of the KEGG pathway revealed that tumour necrosis factor (TNF) signalling pathway and cytokine-cytokine receptor interaction were the top canonical pathways associated with the DEGs (Figure 2D). We imported the results into R software to draw the graphics, and $P < 0.05$ was considered as the threshold value of significance.

PPI NETWORK ANALYSIS

The online database STRING, GeneMANIA, and Cytoscape software were used to analyse the DEGs. The 25 DEGs showed a complex DEGs-PPI network with a co-expression of 86.48%, a shared protein domain of 3.19%, a co-localization rate of 8.14%, and genetic interactions of 1.19% (Figure 3A). We identified the first six genes with the highest degree of interaction, including C-C motif chemokine ligand 2 (CCL2),

TABLE I. THE BIOLOGICAL PROCESS IN THE ENRICHED ANALYSIS OF THE DIFFERENTIALLY EXPRESSED GENES BETWEEN THE SSC-ILD AND NORMAL SAMPLES (TOP 20 ACCORDING TO P-VALUE)

Term	Description	Count	P Value
GO:0006955	immune response	7	2.20E-05
GO:0007568	aging	5	8.19E-05
GO:0030574	collagen catabolic process	4	1.01E-04
GO:0048246	macrophage chemotaxis	3	1.51E-04
GO:0001525	angiogenesis	5	2.61E-04
GO:0006954	inflammatory response	5	0.001898
GO:0009612	response to mechanical stimulus	3	0.003187
GO:0060326	cell chemotaxis	3	0.003854
GO:0071347	cellular response to interleukin-1	3	0.004581
GO:0050729	positive regulation of inflammatory response	3	0.004836
GO:0007166	cell surface receptor signalling pathway	4	0.006749
GO:0071356	cellular response to tumour necrosis factor	3	0.010683
GO:0071222	cellular response to lipopolysaccharide	3	0.011247
GO:0010226	response to lithium ion	2	0.012793
GO:0006935	chemotaxis	3	0.013017
GO:0019221	cytokine-mediated signalling pathway	3	0.0149
GO:0002523	leukocyte migration involved in inflammatory response	2	0.015615
GO:0071318	cellular response to ATP	2	0.018428
GO:0007165	signal transduction	6	0.022087
GO:0001666	response to hypoxia	3	0.024837

prostaglandin-endoperoxide synthase 2 (PTGS2), matrix metallopeptidase 7 (MMP7), TNF alpha-induced protein 3 (TNFAIP3), leukaemia inhibitory factor (LIF), and C-X3-C motif chemokine ligand 1 (CX3CL1) (Figure 3B).

VALIDATION OF THE HUB GENES

The expression levels of the six hub genes including "CCL2", "PTGS2", "MMP7", "TNFAIP3", "LIF", and "CX3CL1" in the GSE48149, GSE76808, and GSE81292 datasets were evaluated. The P-value < 0.05 was taken as the cut off standard, only three genes were selected which were the CX3CL1, MMP7, and TNFAIP3. The expression of MMP7 in all three datasets increased significantly, whereas the CX3Cl1 and TNFAIP3 expression decreased in the GSE76808 and GSE81292 datasets but increased in the GSE48149 dataset. We observed that MMP7 was the most significant (Figure 4).

HUB GENE GSEA ANALYSIS

We executed a GSEA analysis of the MMP7 hub gene to further elucidate the possible mechanism of the SSc-ILD-related genes. The samples were divided into high expression and low expression groups according to the median value of gene expression. The GSEA analysis showed that low expression of MMP7 was enriched in the G-protein-coupled amine receptor activity (GSE48149) (Figure 4E), and the high expression was enriched in the establishment of protein localization to the endoplasmic reticulum (GSE76808) (Figure 4D).

DISCUSSION

At present, the etiopathogenesis of the SSc-ILD is not entirely clear; nevertheless, the principal pathogenesis of SSc-ILD was thought to sustained damage to lung cells[21]. The sustained damage induced fibrotic stimulation, leading to fibroblast activation and myofibroblast transformation[21]. Lung injury might result in the release of fibrotic mediators, such as transforming growth factor β1 (TGF-β1) and connective tissue growth factor (CTGF)[21]. Furthermore, TGF-β promoted the increase of the myofibroblasts in the damaged tissues and the transformation of pericytes[22]. Thus, un-

TABLE II. THE KYOTO ENCYCLOPAEDIA OF GENES AND GENOMES ANALYSES OF THE DIFFERENTIALLY EXPRESSED GENES ACCORDING TO KOBAS 3.0 ANALYSIS (CORRECTED P-VALUE < 0.95 WAS CONSIDERED AS THRESHOLD VALUES OF SIGNIFICANT DIFFERENCE)

ID	Term	Number	Corrected P-Value
hsa04668	TNF signalling pathway	112	1.20E-06
hsa04060	Cytokine-cytokine receptor interaction	294	1.20E-06
hsa04657	IL-17 signalling pathway	93	0.001286
hsa04061	Viral protein interaction with cytokine and cytokine receptor	100	0.001286
hsa04933	AGE-RAGE signalling pathway in diabetic complications	100	0.001286
hsa04062	Chemokine signalling pathway	190	0.005848
hsa05163	Human cytomegalovirus infection	225	0.007972
hsa04974	Protein digestion and absorption	90	0.02281
hsa04064	NF-kappa B signalling pathway	100	0.023513
hsa04926	Relaxin signalling pathway	130	0.033497
hsa04630	Jak-STAT signalling pathway	162	0.036993
hsa04621	NOD-like receptor signalling pathway	178	0.036993
hsa00910	Nitrogen metabolism	17	0.047452

TNF, tumour necrosis factor. IL, Interleukin. AGE-RAGE, Advanced glycated end-products- receptor. NF, nuclear factor. Jak-STAT, Janus kinase-signal transducers and activators of transcription. NOD, Nucleotide-binding and oligomerization domain.

derstanding the pathology and molecular mechanism of the SSc-ILD could facilitate and enhance clinical diagnosis and treatment. However, analysis of the data as a single microarray dataset produced higher false-positive rates, and one-sided results were observed. In this study, we performed a comprehensive analysis of three RNA microarray datasets (GSE48149, GSE81292, and GSE76808) from which 25 DEGs were identified. Then the online website was used further to cluster the DEGs through function and pathway enrichment analyses. Furthermore, the DEGs-PPI network was constructed through the STRING database and Cytoscape, which revealed the molecular mechanism of the SSc-ILD.

Therefore, the DEGs reflected the key biomarkers related to the pathogenesis and progression of SSc-ILD, some of which had been explored and used for therapeutic purposes. The CX3CL1 was associated with ILD progression, and its concentration in the lung tissue and serum of patients with SSc increase[23]. The decreased expression of serum chemokine (C-X-C motif) ligand 14 (CXCL14) helped to maintain abnormal immune function and neovascularization disorders, both of which might be fundamental to the development of SSc. Furthermore, using intravenous cyclophosphamide pulse therapy, the serum CXCL14 levels in SSc-ILD patients who successfully treated significantly increased compared with the baseline[24]. It was necessary to identify more DEGs and explore whether possible targeted genes affected the occurrence and development of SSc-ILD. Through molecular experiments, it was possible to identify biomarkers in SSc-ILD.

In the DEGs enrichment analysis, biological processes included immune response, aging, collagen catabolic process, macrophage chemotaxis, angiogenesis, and inflammatory response. As in previously published studies, SSc-ILD was shown to result from an interaction among fibrotic, self-immunity, inflammation, and vascular injury[25]. The cellular components included the extracellular region, extracellular space, and the proteinaceous extracellular matrix. The onset of SSc-ILD damaged the alveolar epithelium or vascular system, or both, followed by abnormal activation of the immune system[25, 26]. This process promoted the recruitment and activation of fibroblasts, the excessive production of the extracellular matrix, and the normal lung structure eventually replaced by scars[25, 26]. The increased release of the fibrotic cytokines, growth factors, peptides, and bioactive proteins, resulted in stronger signalling[26].

The SSc-ILD signalling pathways were different from

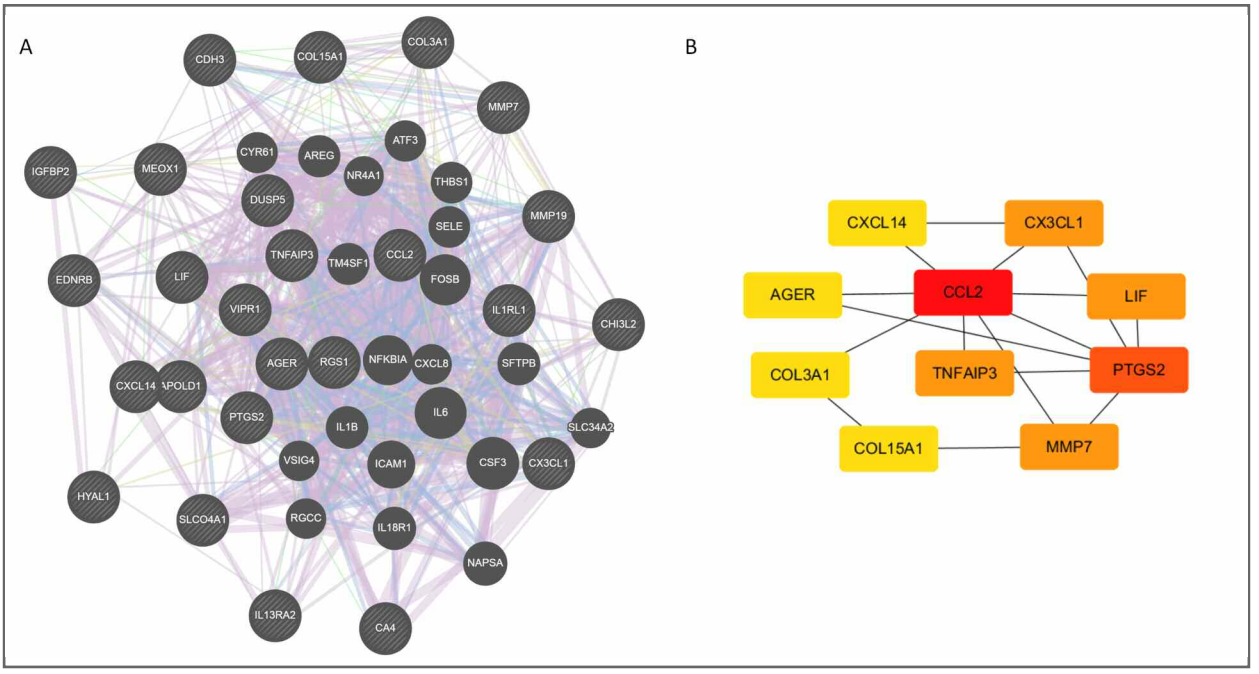

FIGURE 3. Based on the STRING database, GeneMANIA, and Cytoscape software, the protein-protein interaction networks of the differentially expressed genes were constructed along with the modular analyses. (A) The result of the protein-protein interaction networks according to GeneMANIA. (B) Different degree scores have different colours displayed on the graph. As the colour turned red, the score become higher.

the signal transduction pathways of primary tumours, such as the TNF signalling pathway, and cytokine-cytokine receptor interaction. Scleroderma was an inflammatory disease, and TNF played a critical role in the pathogenesis of inflammatory diseases. Previous studies showed that high TNF-alpha was the best predictor of progressive disease through univariate analysis[27]. Cytokine (Interleukin-17) was necessary for the pathogenesis of lung impairment in patients with systemic sclerosis[28]. Although some pathways were identified, a series of molecular experiments might help to provide more detailed and robust evidence for the likely phenotypes and pathway regulation of these predicted SSC-ILD genes.

Besides, the six hub genes of the SSc-ILD scored highest in the protein network. Previous research had pointed out that when the experimental group included ILD patients, some genes including CCL2 and male factors, could predict PAH with very high accuracy[29]. Alveolar lavage fluid from SSc-ILD patients showed higher CCL2 (30). In our study, the logFC of MMP7 was greater than 2 in the three datasets. Elevated serum MMP7 levels in patients with SSc might correlate with pulmonary involvement[31], and serum MMP7 was significantly higher in both the IPF and SSc-ILD patients[32]. The plasma MMP7 levels were negatively correlated with forced vital capacity (FVC) and carbon monoxide (DLCO) decline[33], together with FVC and DLCO (PCMI) contribution to mortality prediction[34]. Zuo *et al.* (2002), confirmed that MMP7 protein expression in epithelial cells increased in the bleomycin-induced fibrosis model, and also that MMP7-/- knockout could prevent bleomycin-induced fibrosis[35]. Unfortunately, most previous studies focused on IPF, not SSc-ILD. The co-expression and interaction between the hub genes are a new and exciting border area, requiring further research.

Studies of genetically targeted mice of pulmonary fibrosis revealed that most MMPs promoted the development of pulmonary fibrosis and found multiple mechanisms. The MMP7 could promote the epithelial-to-mesenchymal transition, increase the activity of fibrotic mediators, or reduce the levels of anti-fibrotic mediators in the lung tissue[36]. Besides, elevated MMP7 might indicate asymptomatic ILD and reflect disease progression in patients with idiopathic pulmonary fibrosis[33]. The MMP7 found on the surface of the lung epithelial cells was one of several MMPs that activated

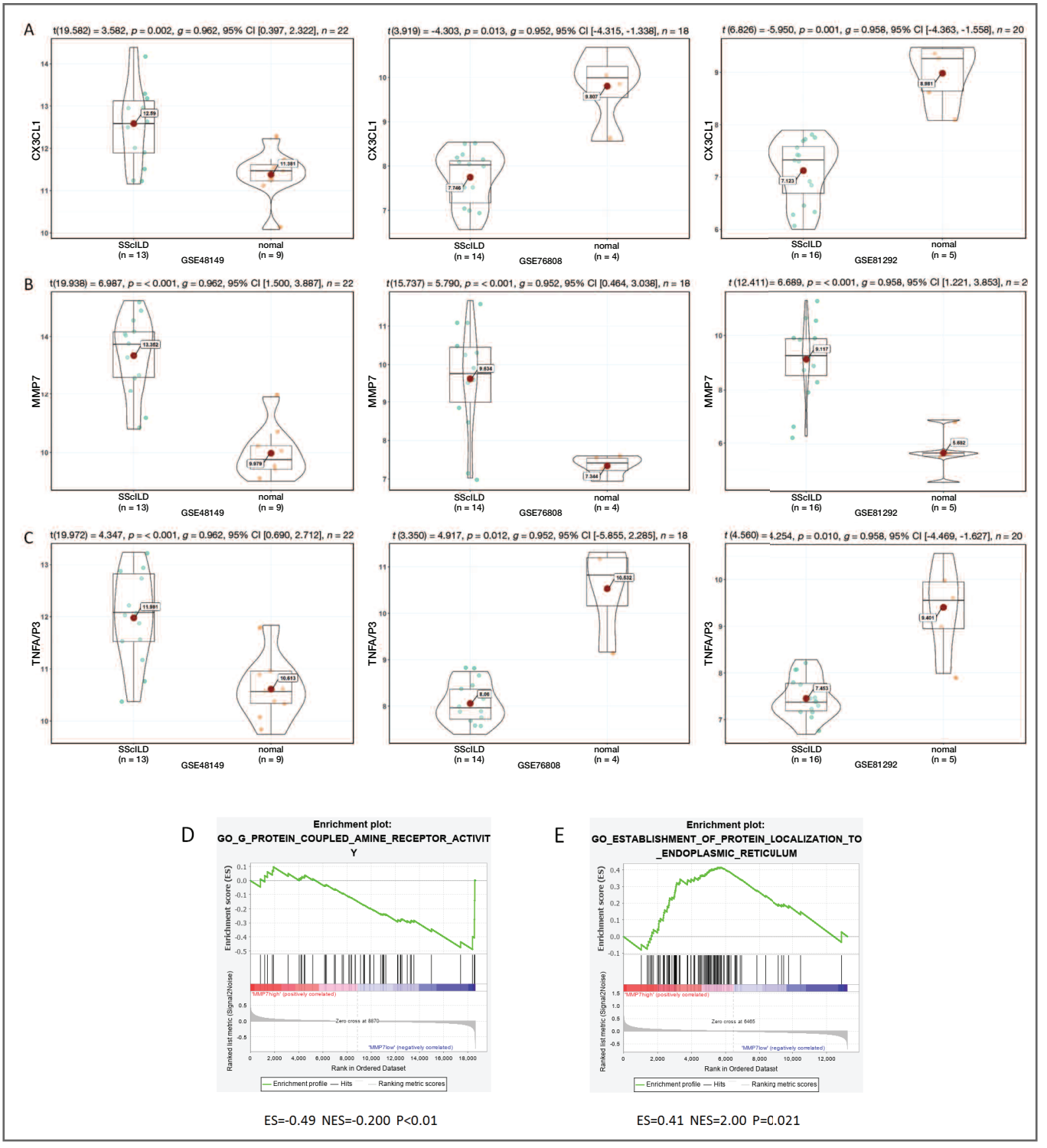

FIGURE 4. Validation of the hub genes, including MMP7(A), CX3CL1(B), and TNFAIP3(C) in the datasets (GSE48149, GSE76808, and GSE81292). (D) The GSEA analysis of high expression of MMP7 in GSE76808. (E) The GSEA analysis of low expression of MMP7 in GSE48149. Profile of the Running ES Score & Positions of GeneSet Members on the Rank Ordered List in (D) and (E). ES: enrichment score; NES: normalized enrichment score.

TGF-β[37,38], and the TGF-β signalling pathway was a key mediator of fibroblast activation[39]. Nintedanib had great potential in the treatment of progressive fibrotic interstitial lung disease. It did not only increase the expression and activity of MMPs in TGF-β1 activated pericytes but also inhibited TGF-β-induced fibroblast to myofibroblast transformation[16], although the specific relationship between MMP7 and Nintedanib still needs further research. In the colonic mucosa, the MMP7 controlled the transepithelial influx of neutrophils by generating a chemokine gradient, which could cause tissue damage if in excess[40]. However, there are few studies on the related mechanisms of SSc-ILD. Since MMP7 consistently increased in all three datasets, we chose MMP7 for the GSEA analysis. It was found that the low expression of MMP7 was rich in the G-protein coupled amine receptor activity, and the high expression was enriched in the establishment of protein localization in the endoplasmic reticulum, which might pave the way for subsequent research. The G-protein-coupled receptor kinase-2 could regulate the MMP7 levels and play a key role in TNF -induced wound epithelial cell healing in the colon epithelial cells[41]. Many classical heterotrimeric G protein-coupled receptors (GPCRs) have been shown to activate the classical Wnt pathway and ultimately stabilize β-catenin[42], while MMP7 might regulate β-catenin localization and signalling activation in injured lung epithelium[43]. Some agonists activate of GPCR activated MMPs and then activate the EGFR, thereby promoting vasoconstriction and growth[44].

Our approach improved our understanding of the potential targets of SSc-ILD treatment. However, our study had some limitations. Firstly, to fully unfold the molecular mechanism of SSc-ILD occurrence and development, microarray samples need to be extracted from patients with varying degrees of ILD; therefore, more samples are needed. Secondly, many SSc-ILD-related biomarkers remain uncharacterized, and further bioinformatics analysis and experimental confirmation are needed to clarify the biological functions of these predicted genes in SSc-ILD.

In summary, our current study provides a reliable comprehensive analysis by combining three gene expression datasets to study the differential expression of RNAs related to SSc-ILD progression. Twenty-five DEGs and hub genes were identified, and GO and KEGG analyses were performed. MMP7 was crucial in systemic sclerosis-associated interstitial lung disease and might play a role by regulating the G-protein coupled amine receptor activity. These results might help develop new strategies for the diagnosis and treatment of SSc-ILD.

CORRESPONDENCE TO
Chengshui Chen
Department of Pulmonary and Critical Care Medicine
The First Affiliated Hospital of Wenzhou Medical University
Ouhai District, Wenzhou, China.
E-mail: wzchencs@163.com

REFERENCES

1. Cottin V, Brown KK. Interstitial lung disease associated with systemic sclerosis (SSc-ILD). Respir Res. 2019;20(1):13.
2. Caron M, Hoa S, Hudson M, Schwartzman K, R S. Pulmonary function tests as outcomes for systemic sclerosis interstitial lung disease. Eur Respir Rev. 2018;27(148):170102.
3. Toya SP, Tzelepis GE. The many faces of scleroderma sine scleroderma: a literature review focusing on cardiopulmonary complications. Rheumatol Int. 2009;29(8):861-868.
4. Khanna D, Tashkin DP, Denton CP, Renzoni EA, Desai SR, Varga J. Aetiology, Risk Factors, and Biomarkers in Systemic Sclerosis with Interstitial Lung Disease. Am J Respir Crit Care Med. 2020;201(6):650-660.
5. Denton CP, Khanna D. Systemic sclerosis. Lancet. 2017;390(10103):1685-1699.
6. Bouros D, Wells AU, Nicholson AG, Colby TV, Polychronopoulos V, Pantelidis P, et al. Histopathologic Subsets of Fibrosing Alveolitis in Patients with Systemic Sclerosis and Their Relationship to Outcome. Am J Re Crit Care Med. 2002;165 (12):1581-1586.
7. Tyndall AJ, Bannert B, Vonk M, Airo P, Cozzi F, Carreira PE, et al. Causes and risk factors for death in systemic sclerosis: a study from the EULAR Scleroderma Trials and Research (EUSTAR) database. Ann Rheum Dis. 2010;69(10):1809-1815.
8. Hao Y, Hudson M, Baron M, Carreira P, Stevens W, Rabusa C, et al. Early Mortality in a Multinational Systemic Sclerosis Inception Cohort. Arthritis Rheumatol. 2017;69(5):1067-1077.
9. Morrisroe K, Stevens W, Sahhar J, Ngian GS, Ferdowsi N, Hansen D, et al. The clinical and economic burden of systemic sclerosis related interstitial lung disease. Rheumatology. 2019:kez532.
10. Giacomelli R, Liakouli V, Berardicurti O, Ruscitti P, Di Benedetto P, Carubbi F, et al. Interstitial lung disease in systemic sclerosis: current and future treatment. Rheumatol Int. 2017;37(6): 853-863.
11. Sottile PD, Iturbe D, Katsumoto TR, Connolly MK, Collard HR, Leard LA, et al. Outcomes in Systemic Sclerosis–Related Lung Disease After Lung Transplantation. Transplantation. 2013;95(7):975-980.
12. Furst DE, Tseng CH, Clements PJ, Strange C, Tashkin DP, Roth MD, et al. Adverse events during the Scleroderma Lung Study. Am J Med. 2011;124(5):459-467.
13. Martinez F, McCune W. Cyclophosphamide for Scleroderma Lung Disease. N Engl J Med. 2006;354(25):2707–2709.
14. Yang L, Wang Q, Hou Y, Zhao J, Li M, Xu D, et al. The Chinese herb Tripterygium wilfordii Hook F for the treatment of systemic sclerosis-associated interstitial lung disease: data from a Chinese EUSTAR Center. Clin rheumatol. 2019;39(3):813-821.
15. Katsiari CG, Simopoulou T, Alexiou I, Sakkas LI. Immunotherapy of systemic sclerosis. Hum Vacc & immunother. 2018;14(11):2559-2567.

16. Wollin L, Distler JHW, Redente EF, Riches DWH, Stowasser S, Schlenker-Herceg R, et al. Potential of nintedanib in treatment of progressive fibrosing interstitial lung diseases. Eur Respir J. 2019;54(3):1900161.
17. Troyanskaya O, Cantor M, Sherlock G, Brown P, Hastie T, Tibshirani R, et al. Missing value estimation methods for DNA microarrays. Bioinformatics. Bioinformatics. 2001;17(6):520-525.
18. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 2015;43(7):e47.
19. Huang da W, Sherman BT, RA. L. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res. 2009;37(1):1–13.
20. Xie C, Mao X, Huang J, Ding Y, Wu J, Dong S, et al. KOBAS 2.0: a web server for annotation and identification of enriched pathways and diseases. Nucleic Acids Res. 2011;39(Web Server issue):W316-W22.
21. Tochimoto A, Kawaguchi Y, Yamanaka H. Genetic Susceptibility to Interstitial Lung Disease Associated with Systemic Sclerosis. Clin Med Insights Circ, Respir Pulm Med. 2016;9(Suppl 1):135-140.
22. Cutolo M, Soldano S, Smith V. Pathophysiology of systemic sclerosis: current understanding and new insights. Expert Rev Clin Immunol. 2019;15(7):753-764.
23. Hoffmann-Vold AM, Weigt SS, Palchevskiy V, Volkmann E, Saggar R, Li N, et al. Augmented concentrations of CX3CL1 are associated with interstitial lung disease in systemic sclerosis. PloS One. 2018;13(11):e0206545.
24. Fukui Y, Miyagawa T, Hirabayashi M, Yamashita T, Saigusa R, Miura S, et al. Possible association of decreased serum CXCL14 levels with digital ulcers in patients with systemic sclerosis. J Dermatol. 2019;46(7):584-589.
25. Perelas A, Silver RM, Arrossi AV, Highland KB. Systemic sclerosis-associated interstitial lung disease. Lancet Respir Med. 2020;8(3):304-320.
26. Akter T, Silver R, Bogatkevich G. Recent Advances in Understanding the Pathogenesis of Scleroderma-Interstitial Lung Disease. Curr Rheumatol Rep. 2014;16(4):411.
27. Ates O, Musellim B, Ongen G, Topal-Sarikaya A. Analysis of TNF polymorphisms in Turkish systemic sclerosis patients with interstitial lung involvement. Biochem Genet. 2008;46(11-12):696-701.
28. Olewicz-Gawlik A, Danczak-Pazdrowska A, Kuznar-Kaminska B, Gornowicz-Porowska J, Katulska K, Trzybulska D, et al. Interleukin-17 and interleukin-23: importance in the pathogenesis of lung impairment in patients with systemic sclerosis. Int J Rheum Dis. 2014;17(6):664-670.
29. Moll M, Christmann RB, Zhang Y, Whitfield ML, Wang YM, Rice L, et al. Patients with systemic sclerosis-associated pulmonary arterial hypertension express a genomic signature distinct from patients with interstitial lung disease. J Scleroderma Relat Disord. 2018;3(3):242-248.
30. Schmidt K, Martinez-Gamboa L, Meier S, Witt C, Meisel C, Hanitsch LG, et al. Bronchoalveoloar lavage fluid cytokines and chemokines as markers and predictors for the outcome of interstitial lung disease in systemic sclerosis patients. Arthritis Res Ther. 2009;11(4):R111.
31. Moinzadeh P, Krieg T, Hellmich M, Brinckmann J, Neumann E, Müller-Ladner U, et al. Elevated MMP-7 levels in patients with systemic sclerosis: correlation with pulmonary involvement. Exp Dermatol. 2011;20(9):770-773.
32. Kennedy B, Branagan P, Moloney F, Haroon M, O'Connell OJ, O'Connor TM, et al. Biomarkers to identify ILD and predict lung function decline in scleroderma lung disease or idiopathic pulmonary fibrosis. Sarcoidosis Vasc Diffuse Lung Dis. 2015;32(3):228-236.
33. Rosas I, Richards T, Konishi K, Zhang Y, Gibson K, Lokshin A, et al. MMP1 and MMP7 as potential peripheral blood biomarkers in idiopathic pulmonary fibrosis. PLoS Med. 2008;5(4):e93.
34. Richards TJ, Kaminski N, Baribaud F, Flavin S, Brodmerkel C, Horowitz D, et al. Peripheral blood proteins predict mortality in idiopathic pulmonary fibrosis. Am J Respir Crit Care Med. 2012;185(1):67-76.
35. Zuo F, Kaminski N, Eugui E, Allard J, Yakhini Z, Ben-Dor A, et al. Gene expression analysis reveals matrilysin as a key regulator of pulmonary fibrosis in mice and humans. Proc Nat Acad Sci U S A. 2002;99(9):6292-6297.
36. Craig VJ, Zhang L, Hagood JS, CA O. Matrix metalloproteinases as therapeutic targets for idiopathic pulmonary fibrosis. Am J Respir Cell Mol Biol. 2015;53(5):585-600.
37. Chen P, Abacherli LE, Nadler ST, Wang Y, Li Q, Parks WC. MMP7 shedding of syndecan-1 facilitates re-epithelialization by affecting alpha(2)beta(1) integrin activation. PloS One. 2009;4(8):e6565.
38. Crosby LM, Waters CM. Epithelial repair mechanisms in the lung. Am J Physiol Lung Cell Mol Physiol. 2010;298(6):L715-L731.
39. Akhmetshina A, Palumbo K, Dees C, Bergmann C, Venalis P, Zerr P, et al. Activation of canonical Wnt signalling is required for TGF- -mediated fibrosis. Nat Commun. 2012;3:735.
40. Swee M, Wilson CL, Wang Y, McGuire JK, Parks WC. Matrix metalloproteinase-7 (matrilysin) controls neutrophil egress by generating chemokine gradients. J Leukoc Biol. 2008;83(6):1404-1412.
41. Steury MD, Lucas PC, McCabe LR, Parameswaran N. G-protein-coupled receptor kinase-2 is a critical regulator of TNF signaling in colon epithelial cells. Biocheml J. 2017;474(14):2301-2313.
42. Shevtsov SP, Haq S, Force T. Activation of beta-catenin signaling pathways by classical G-protein-coupled receptors: mechanisms and consequences in cycling and non-cycling cells. Cell Cycle. 2006;5(20):2295-2300.
43. Rims CR, McGuire JK. Matrilysin (MMP-7) catalytic activity regulates -catenin localization and signaling activation in lung epithelial cells. Exp Lung Res. 2014;40(3):126-136.
44. Hao L, Du M, Lopez-Campistrous A, Fernandez-Patron C. Agonist-induced activation of matrix metalloproteinase-7 promotes vasoconstriction through the epidermal growth factor-receptor pathway. Circ Res. 2004;94(1):68-76.