

RESEARCH

Open Access



Integrated bioinformatics analysis for the identification of idiopathic pulmonary fibrosis–related genes and potential therapeutic drugs

Zhenzhen Zhang^{1,2}, Qingzhou Guan^{1,2*}, Yange Tian^{1,2}, Xuejie Shao^{1,2}, Peng Zhao^{1,2}, Lidong Huang^{1,2} and Jiansheng Li^{2,3}

Abstract

Objective The pathogenesis of idiopathic pulmonary fibrosis (IPF) remains unclear. We sought to identify IPF-related genes that may participate in the pathogenesis and predict potential targeted traditional Chinese medicines (TCMs).

Methods Using IPF gene-expression data, Wilcoxon rank-sum tests were performed to identify differentially expressed genes (DEGs). Protein–protein interaction (PPI) networks, hub genes, and competitive endogenous RNA (ceRNA) networks were constructed or identified by Cytoscape. Quantitative polymerase chain reaction (qPCR) experiments in TGF- β 1-induced human fetal lung (HFL) fibroblast cells and a pulmonary fibrosis mouse model verified gene reliability. The SymMap database predicted potential TCMs targeting IPF. The reliability of TCMs was verified in TGF- β 1-induced MRC-5 cells.

Materials Multiple gene-expression profile data of normal lung and IPF tissues were downloaded from the Gene Expression Omnibus database. HFL fibroblast cells and MRC-5 cells were purchased from Wuhan Procell Life Science and Technology Co., Ltd. (Wuhan, China). C57BL/12 mice were purchased from Beijing Vital River Laboratory Animal Technology Co., Ltd. (Beijing, China).

Results In datasets GSE134692 and GSE15197, DEGs were identified using Wilcoxon rank-sum tests (both $p < 0.05$). Among them, 1885 DEGs were commonly identified, and 87% (1640 genes) had identical dysregulation directions (binomial test, $p < 1.00E-16$). A PPI network with 1623 nodes and 8159 edges was constructed, and 18 hub genes were identified using the Analyze Network plugin in Cytoscape. Of 18 genes, *CAV1*, *PECAM1*, *BMP4*, *VEGFA*, *FYN*, *SPP1*, and *COL1A1* were further validated in the GeneCards database and independent dataset GSE24206. ceRNA networks of *VEGFA*, *SPP1*, and *COL1A1* were constructed. The genes were verified by qPCR in samples of TGF- β 1-induced HFL fibroblast cells and pulmonary fibrosis mice. Finally, Sea Buckthorn and Gnaphalium Affine were predicted as potential TCMs for IPF. The TCMs were verified by qPCR in TGF- β 1-induced MRC-5 cells.

Conclusion This analysis strategy may be useful for elucidating novel mechanisms underlying IPF at the transcriptome level. The identified hub genes may play key roles in IPF pathogenesis and therapy.

*Correspondence:

Qingzhou Guan
a5410980@163.com

Full list of author information is available at the end of the article



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Keywords Idiopathic pulmonary fibrosis, Gene expression, Differentially expressed genes, ceRNA network, Traditional Chinese medicine

Introduction

Idiopathic pulmonary fibrosis (IPF), an unexplained progressive and fatal interstitial lung disease, is the most common form of idiopathic interstitial pneumonia and predominantly affects the health of men and smokers [1]. In recent years, evidence has suggested that the morbidity and mortality rates of IPF patients are increasing [2]. IPF primarily occurs in middle-aged or elderly people with a median age of approximately 65 years at the time of diagnosis, and the median survival time is 3–5 years following diagnosis [3, 4]. In addition, people with IPF usually experience increasingly debilitating symptoms, including cough, dyspnea, fatigue, and weight loss. The physical, psychological, and socio-economic burdens of IPF are substantial [5, 6], yet treatment options for patients with IPF remain limited. Lung transplantation is a fundamental treatment that clearly improves survival in carefully selected patients, but the restricted supply of donor organs and the limitations of chronic allograft rejection mean that only a few patients can receive this intervention [7]. Two drugs, pirfenidone and nintedanib, are also commonly used for IPF and can slow down the progression of pulmonary fibrosis and reduce mortality. However, their effects are not ideal, and they trigger complications like nausea, diarrhea, dyspepsia, and rash [8, 9].

The etiology of IPF remains unclear. Numerous pathophysiological factors have been implied in the genesis and progression of the disease, but most are not yet fully elucidated [10]. The underlying pathomechanisms of IPF, with its complex immunological and inflammatory processes and external impacts, have been the focus of recent research. Smoking, pneumotoxic medications, and the inhalation of dusts are well-known risk factors [11]. Moreover, genetic factors play a vital role in the pathogenesis of IPF [12]. For example, in patients with IPF, *GDF15* messenger RNA (mRNA) expression in lung tissue is significantly increased and correlates with pulmonary function [13]. Son et al. [14] suggested that *ROR2* expression correlates with the development and severity of IPF. *LAMA1* mRNA expression was significantly increased in lung tissue obtained from IPF patients and was identified to be a genetic modifier of transforming growth factor TGF- β 1 effector responses that significantly affect the development of pulmonary fibrosis [15]. Therefore, it is of great significance to explore the pathogenesis of IPF at the gene level.

In recent years, high-throughput genetic testing technology has become increasingly advanced, and many transcriptomics features have become more common in disease research [16–18]. In addition, competitive endogenous RNA (ceRNA) regulatory networks have been found to be involved in the transcription and regulation of various disease-related genes in a variety of studies [19, 20]. By combining high-throughput microarray/RNA-sequencing data and bioinformatics analysis algorithms, pathogenic genes can be identified, providing some guidance for pathogenesis research and the clinical treatment of these diseases.

In this study, we analyzed the gene-expression profile data of IPF from the Gene Expression Omnibus (GEO) database and identified differentially expressed genes (DEGs). A protein–protein interaction (PPI) network was constructed based on the STRING database (<https://www.string-db.org/>), and the Analyze Network plugin for Cytoscape (Institute for Systems Biology, Seattle, WA, USA) was applied to verify hub genes. The above hub genes were then optimized using the GeneCards database. Finally, an independent GEO dataset was employed as a validation group to further validate the hub genes. Furthermore, target microRNAs (miRNAs) of identified hub genes were predicted, a co-expressed network was constructed by Cytoscape, and ceRNA networks were constructed based on the prediction results of long non-coding RNAs (lncRNAs). Traditional Chinese medicine (TCM) plays a significant part in treating IPF, boasting strengths like hypotoxicity and multi-level and multi-target actions [9]. The potential therapeutic effect of TCM for IPF was predicted by SymMap database. Our findings provide evidence for the mechanisms of disease development in IPF at the transcriptome level and explore potential hub genes for the treatment of IPF.

Materials and methods

Data acquisition and identification of DEGs

First, the keywords “idiopathic pulmonary fibrosis”, “tissue”, “gene expression”, “biopsy or surgical resection” and “homo sapiens” were searched in the GEO database (<https://www.ncbi.nlm.nih.gov/geo/>) [21], and the samples from lung transplant patient were excluded in this study. Then, a total of 81 IPF and 62 lung normal samples from 6 datasets were found. Among the above datasets, GSE134692 (including 46 IPF and 26 normal samples) and GSE15197 (including 8 IPF and 13 normal samples) has the largest sample size, as shown in Table 1. Thus,

Table 1 Normal lung and IPF tissue data used in this study

GEO Acc	Platform	Normal	IPF
GSE15197	Agilent GPL6480	13	8
GSE134692 ^a	Illumina HiSeq 2000 GPL16791	26	46

^a Denotes the RNA sequencing data

these two datasets were downloaded on January 4, 2021 for the following high-throughput analysis. Processed data detected by microarray (Agilent Technologies, Santa Clara, CA, USA) and an RNA-sequencing platform (Illumina, San Diego, CA, USA) were directly downloaded. The probe identifier was mapped to the Entrez Gene identifier with the corresponding platform file for the array-based data, while, for the sequencing-based data, gene symbols were mapped to the Entrez Gene identifier with the Database for Annotation, Visualization and Integrated Discovery (DAVID) database (accessed on 28 January 2021). Then, for the data from public databases, the Wilcoxon rank-sum test was used to identify DEGs ($p < 0.05$).

PPI network construction and hub gene identification

A PPI network was constructed based on the genes of interest using information from the STRING database (<https://www.string-db.org/>) (accessed on 30 August 2021) with a combined score of > 0.4 points. Then, the interaction information was downloaded and Cytoscape (version 3.8.2) was used to visualize the PPI network. Degree is an important analytical parameter in network analysis, which represents the number of genes that directly interact with the node gene in the network. The greater the degree value of a node gene was, the more genes that could directly interact with the node gene, and the degree value is also commonly used to identify hub genes [22–25].

Based on information from the DAVID database (<https://david.ncifcrf.gov/>) (accessed on 8 October 2021), Gene Ontology (GO) [26, 27] and Kyoto Encyclopedia of Genes and Genomes (KEGG) [28–30] pathway enrichment analyses were used to explore the function of hub genes, and the threshold was set as a false-discovery rate (FDR) of < 0.05 . GeneCards is a comprehensive database of human genes, providing concise genomic, proteomic, transcriptional, genetic, and functional information on all known and predicted human genes. Information mainly includes relationships that point to disease, gene expression and gene function, etc. [31]. Then, the hub genes were further optimized and validated in the GeneCards database (<https://www.genecards.org/>) (accessed on 18 October 2021) and an independent GEO dataset (GSE24206).

Prediction of target miRNAs and the construction of ceRNA networks

Five online miRNA databases (namely miRDIP, the Encyclopedia of RNA Interactomes [ENCORI], TargetScan, DIANA-micro T, and miRWalk) (accessed on 19 October 2021) were used to predict the target miRNAs of genes of interest with the default parameters [20, 32]. In this study, if a gene (such as gene A) could be targeted by a certain miRNA in at least four of the five databases, then it was defined as the target miRNA of gene A. Based on the miRNA predict result of the interested genes by the five miRNA databases, the miRNA–mRNA pairs were obtained. Then, those gene pairs were imported in Cytoscape software, and the visualization results of miRNA-mRNA co-expressed network was generated.

ENCORI (<https://starbase.sysu.edu.cn/index.php>) (accessed on 26 October 2021) was used to predict lncRNAs that interacted with the miRNAs of interest using the following screening criteria: mammalian, human h19 genome, strict stringency (≥ 5) of CLIP-Data, and the presence or absence of degradome data [33]. The ceRNA networks of mRNAs, miRNAs, and lncRNAs were constructed using Cytoscape.

Herbal medicines

Sea Buckthorn was purchased from Tongrentang (Beijing, China) (No.221109004) and Gnaphalium Affine was purchased from KANGMEI PHARMACEUTICAL CO., LTD. (Guangdong, China) (No.3014258). The proportion of Chinese medicine and water is 100 g: 1200 ml water boil to 100 ml for decocting medicine, and then use EYELA freeze-drying machine FDU-2110 (TOKYO RIKAKIKAI CO.,LTD., SER.NO. 12211071) to freeze dry into powder, and finally mix into 25 mg/ml concentration.

Cell culture

Human fetal lung fibroblasts cells were purchased from Wuhan Procell Life Science and Technology Co., Ltd. (Wuhan, China). HFL fibroblast cells were cultured in Ham's F-12 K medium containing 10% fetal bovine serum and 1% penicillin–streptomycin solution at 37 °C and 5% CO₂. Then, the cultured cells were divided into a control group and a model group. A pulmonary fibrosis cell model was established through HFL fibroblast cells induced by TGF- β 1 (5 ng/mL) to activate fibroblasts, while the control group was not treated. The changes in HFL cells induced by TGF- β 1 are similar to the pathological changes in human pulmonary fibrotic disease, suggesting that a TGF- β 1-induced fibroblast activation model has been successfully established, which has also been used as a model for interstitial

pneumonia in other studies [34, 35]. After 6 h of induction, two groups of cells were harvested for quantitative polymerase chain reaction (qPCR) experiments.

MRC-5 cells were purchased from Wuhan Procell Life Science and Technology Co., Ltd. (Wuhan, China) (No. CL-0161). The cells were cultured in medium containing procell MEM + 5% double antibody + 10% lonsera in a 5% CO₂ incubator at 37 °C. MRC-5 cells were activated by TGF-β1 (5 ng/mL) to construct the pulmonary fibrosis cell model [36] and divided into control group, model group, Sea Buckthorn group, and Gnaphalium Affine group. The treatment groups were given Sea Buckthorn and Gnaphalium Affine (25 mg/mL), respectively. After 3~4 h of drug action, TGF-β1 was added to induce for 6 h except the control group. The cells were collected by QIAzol, and the mRNA expressions of *VEGFA*, *SPP1* and *COL1A1* were detected by qPCR experiments.

Animal experiment

Six C57BL/12 mice (6–8 weeks, 18–22 g) were purchased from Beijing Vital River Laboratory Animal Technology Co., Ltd. (Beijing, China). This experiment was approved by the Ethics Committee of Laboratory Animal Welfare of the Henan University of Chinese Medicine (DWLL202203014). The mice were randomly divided into a model group and control group; then, the model group mice were anesthetized with 1% pentobarbital sodium (50 mg/kg), and pulmonary fibrosis mice were established by injection with bleomycin (5 mg/kg) through an endotracheal drip [35]. The mice in the control group were dripped with the same volume of sterile saline. On day 28, the mice in each group were killed, and their lung tissues (50 mg) were harvested for qPCR experimentation.

Quantitative real-time polymerase chain reaction

Total RNA was extracted using QIAzol (Qiagen, Hilden, Germany) according to the manufacturer's instructions. Reverse transcription was performed using HiScript® II Q RT SuperMix (Vazyme, Nanjing, China). The complementary DNA after reverse transcription were amplified using a real-time fluorescent qPCR instrument. *GAPDH* and *MUBB* were used as housekeeping genes in HFL fibroblast cells and mice lung tissue for qPCR experiments, respectively (the primers are shown in Tables 2 and 3). The $2^{-\Delta\Delta CT}$ method was used to assess the gene-expression level. Finally, statistical analysis was performed by one-way analysis of variance testing in SPSS version 22.0 (IBM Corporation, Armonk, NY, USA).

Table 2 Primers for human genes of interest

Gene Symbol		Primer sequence (5'–3')
<i>VEGFA</i>	F	GAGGGCAGAATCATCACGAAG
	R	TGTGCTGTAGGAAGCTCATCTCTC
<i>COL1A1</i>	F	GAGGGCCAAGACGAAGACATC
	R	CAGATCACGTCATCGCACAAAC
<i>SPP1</i>	F	GGAGTTGAATGGTGCATACAAGG
	R	CCACGGCTGTCCCAATCAG
<i>GAPDH</i>	F	GGAGCGAGATCCCTCCAAAAT
	R	GGCTGTTGCATACTTCTCATGG

Table 3 Primers for mouse genes of interest

Gene Symbol		Primer sequence (5'–3')
<i>Vegfa</i>	F	GATCTGCTCCCTCCCTCTACA
	R	TTGACCCCTTCCCTTTCCTCG
<i>Spp1</i>	F	AGCAAGAAACTCTTCCAAGCAA
	R	GTGAGATTGCTCAGATTCATCCG
<i>Mubb</i>	F	TGGCTATTAATTATTCGGTCTGCA
	R	GCAAGTGGCTAGAGTGCAGAGTAA

Prediction of potential TCMs based on the SymMap database

SymMap (<http://www.symmap.org/>) (accessed on 29 December 2021) integrates TCM with modern medicine through both internal molecular mechanisms and external symptom mapping, thus providing a massive amount of information on herbs/ingredients, targets, and the clinical symptoms and diseases, they are being used to treat for drug-screening efforts. In order to obtain the more reliable TCMs prediction results, we choose a stricter statistical threshold (FDR < 0.01) [37]. Moreover, when all the genes of interest could be targeted by a single herb, this herb was defined as a target medicine of IPF.

Results

Identification of DEGs

For the 46 IPF and 26 normal tissue samples from GSE134692, 7550 DEGs were identified using Wilcoxon rank-sum tests ($p < 0.05$). Similarly, for the eight IPF and 13 normal tissue samples from GSE15197, 3557 DEGs were identified. Among those DEGs, 1885 DEGs were commonly identified, and 87% (1640 genes, including 940 up-regulated and 700 down-regulated genes) had identical dysregulation directions, which could not occur by chance (binomial test, $p < 1.00E-16$).

PPI network construction and hub gene identification

Based on the 1640 DEGs, a PPI network composed of 1623 nodes and 8159 edges was constructed (interaction score > 0.4 points). Then, the network was visualized by Cytoscape, and the Analyze Network plugin in Cytoscape was used to identify hub genes. In this study, most genes have higher degree values, as shown in Table 4. Thus, a relatively strict degree threshold (≥ 100) was set to identify hub genes. Finally, 18 genes (including *IGF1*, *SPP1*, *BMP4*, *CDH2*, *PECAM1*, *ITGA2B*, *CDH5*, *KDR*, *SOX2*, *TJP1*, *FYN*, *BDNF*, *EGF*, *VEGFA*, *DLG4*, *KRAS*, *CAV1*, and *COL1A1*) were identified, as shown both in Table 4 and in Table S1 in the [Supplementary Material](#).

GO term and KEGG pathway enrichment analyses

GO and KEGG analyses of the 18 hub genes were performed using the DAVID database. The GO functional

analysis was divided into the following three parts: biological processes (BPs), molecular functions (MFs), and cell components (CCs). The results were considered statistically significant if $FDR < 0.05$. The top five most significant GO terms and KEGG pathways are shown in Fig. 1 and Fig. 2. And the complete list of enriched GO terms and KEGG pathways were shown in supplementary Table S2 and Table S3.

The GO analysis results revealed that some of these pathways were reported to be related with IPF. For example, in BPs, positive regulation of cell migration, positive regulation of MAPK cascade and cell–cell adhesion are related to the pathogenesis of IPF. Pulmonary fibrosis is typically characterized by activated fibroblast proliferation and migration, and its abnormality is a major concern for treating pulmonary fibrosis [38]. Studies have shown that TGF- β 1 plays a central role in the occurrence and development of IPF, and it may promote the fibrotic process of IPF through MAPK signaling pathway [39]. A genome-wide association study of susceptibility to IPF showed that cell–cell adhesion was associated with increased sensitivity to IPF [40]. IPF is characterized primarily by excessive deposition of extracellular matrix (ECM) proteins by activated lung fibroblasts and myofibroblasts, resulting in reduced gas exchange and impaired pulmonary function [41], therefore extracellular space, extracellular region in CCs and extracellular matrix binding in MFs are related to the pathogenesis of IPF. Some enriched KEGG pathways, such as PI3K-Akt signaling pathway and Ras signaling pathway were reported to be related with IPF. Studies showed that the PI3K/Akt/mTOR axis are implicated in fibrosis, the pan-PI3K/mTOR inhibition are currently under clinical evaluation for IPF [42]. Studies have shown that RAB6, a member of the RAS family, and its knockout could inhibit pulmonary fibrosis, oxidative stress, and AEC2 cell death in PM2.5-damaged mice [43].

Table 4 The degree value of gene in the network analysis

Gene Symbol	Degree
<i>VEGFA</i>	250
<i>EGF</i>	214
<i>KRAS</i>	198
<i>PECAM1</i>	154
<i>KDR</i>	152
<i>CAV1</i>	146
<i>FYN</i>	142
<i>BDNF</i>	142
<i>DLG4</i>	140
<i>COL1A1</i>	136
<i>IGF1</i>	128
<i>CDH2</i>	124
<i>SOX2</i>	122
<i>BMP4</i>	116
<i>CDH5</i>	116
<i>ITGA2B</i>	104
<i>TJP1</i>	102
<i>SPP1</i>	102
<i>BMP2</i>	96
<i>COL1A2</i>	96
<i>RUNX2</i>	94
<i>WNT3A</i>	94
<i>CRK</i>	94
<i>CSF2</i>	90
<i>COL3A1</i>	90
<i>NFKB1A</i>	88
<i>CAT</i>	88
<i>PRKCA</i>	88
<i>ARRB1</i>	86
<i>HSPG2</i>	86

This table only lists the degree of the top 30 genes due to textual limitation

Optimization and verification of the hub genes

By searching the 18 candidate hub genes in GeneCards database, we found that *CAV1* and *PECAM1* were specifically expressed in lung tissues, *BMP4* and *VEGFA* were related to angiogenesis, *FYN* and *SPP1* were related to immunity or inflammation, and *COL1A1* was closely associated with pulmonary fibrosis. Thus, *CAV1*, *PECAM1*, *BMP4*, *VEGFA*, *FYN*, *SPP1*, and *COL1A1* genes were selected for further investigation.

To further test the reliability of the above seven genes, they were validated in an independent GEO dataset (GSE24206) in which eight IPF and six normal tissue samples were included. The expression levels of *BMP4*, *COL1A1*, and *SPP1* were increased and those of *PECAM1*, *FYN*, *VEGFA*, and *CAV1* were decreased in

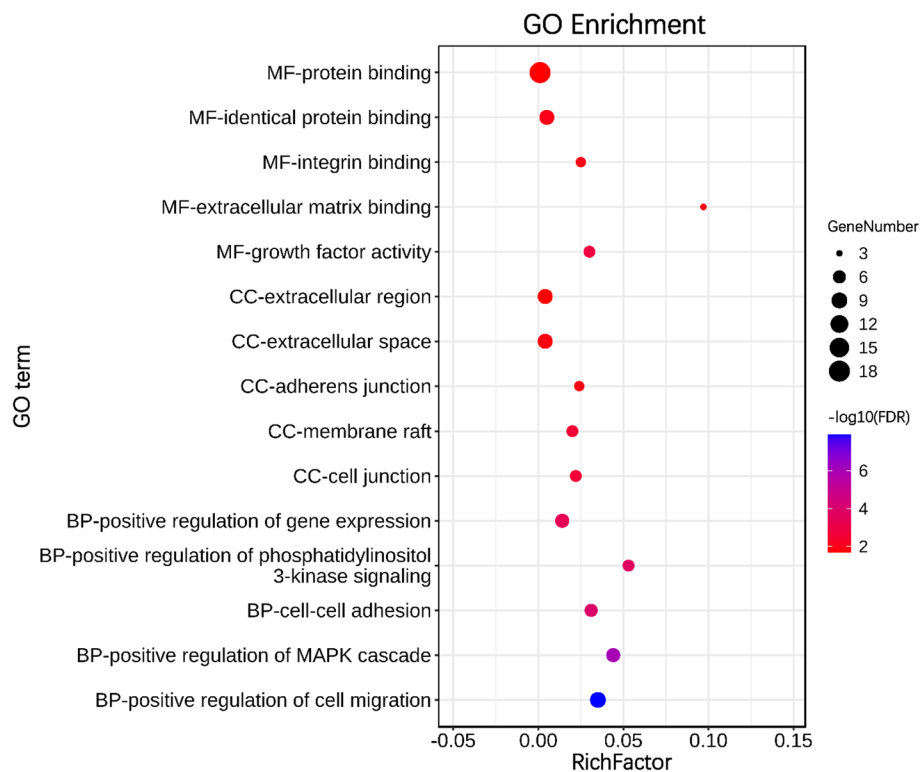


Fig. 1 GO enrichment analyses of 18 hub genes. The top 5 terms in each GO category (MF, CC, and BP). All significant GO terms were performed using the DAVID online tool with the cutoff criterion of FDR < 0.05. The color of each bubble represents the FDR for that term, with blue indicating greater significance. The rich factor refers to the proportion of enriched genes for each term

IPF tissues compared to normal tissues in the GEO dataset GSE24206, and all of these genes exhibited dysregulation directions identical to those in the training dataset. These results proved that the aforementioned seven genes were reliable and deserving of further study.

Target miRNA prediction and ceRNA network construction

Five online miRNA databases were used to predict the target miRNAs of the seven genes of interest. For each gene, the target miRNA of the five miRNA databases was obtained, as shown in Tables S4–S10 in the [Supplementary Material](#). When a particular miRNA could target a gene of interest in at least four databases, the miRNA was defined as the target of this gene. Finally, 39 target miRNAs of three genes (*VEGFA*, *COL1A1*, and *SPP1*) and 42 miRNA–mRNA pairs were obtained. An interaction network of mRNAs and miRNAs, including 42 nodes and 42 edges, was constructed by Cytoscape, as shown in Fig. 3.

It is well known that miRNAs can bind mRNAs and induce gene silencing, further reducing the expression level of a gene, whereas lncRNAs can combine miRNA response elements and increase the expression level of the corresponding gene. This interaction between RNAs is known as a ceRNA network [19]. We also used the

ENCORI database to predict the lncRNAs that interact with the selected miRNAs, and the results of ceRNA network are shown in Fig. 4.

COL1A1, *SPP1*, and *VEGFA* were verified in qPCR

As shown in Fig. 5, for TGF- β 1-induced HFL fibroblast cells, compared to the control group, the expression levels of *COL1A1* and *SPP1* in the model group were significantly up-regulated ($p < 0.01$ and $p < 0.05$, respectively), displaying an identical trend to the results of our data analysis. Compared to the control group, the expression level of *VEGFA* in the model group was also up-regulated ($p < 0.05$); however, the direction of this gene was opposite that found during data analysis in this study, which might be ascribed to the dual role of *VEGFA* in IPF [44].

For pulmonary fibrosis mice, compared to the control group, the expression of *Spp1* in the model group was significantly increased ($p < 0.05$), consistent with the results of data analysis and HFL fibroblast cell qPCR experimentation. Compared to the control group, the expression of *Vegfa* in the model group was also significantly increased ($p < 0.01$), which was also consistent with the results of HFL fibroblast cell qPCR experimentation. And the results are shown in Fig. 6.

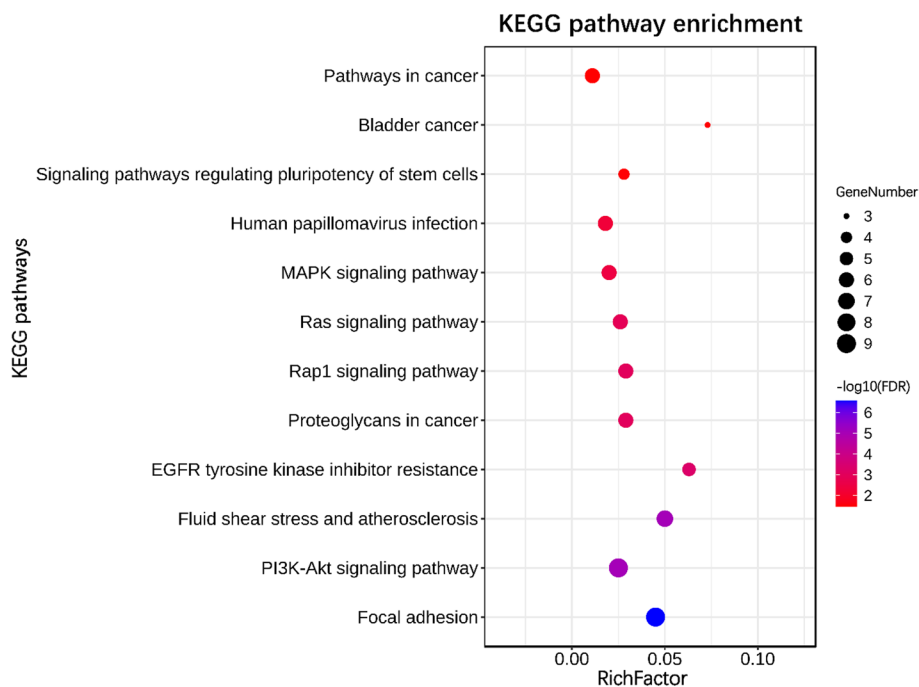


Fig. 2 KEGG enrichment analyses of 18 hub genes. All significant KEGG pathways were performed using the DAVID online tool with the cutoff criterion of $\text{FDR} < 0.05$. The color of each bubble represents the FDR for that pathway, with blue indicating greater significance. The rich factor refers to the proportion of enriched genes for each pathway

Prediction of potential therapeutic TCM results

TCMs have a unique advantage in the present situation of treating IPF. Potential therapeutic TCMs that target hub genes were predicted using the SymMap database [9]. Three key genes were mapped to screen potential TCMs for IPF under the condition of $\text{FDR} < 0.01$, and some TCMs were identified, as shown in Tables S11–S13 in the [Supplementary Material](#). In this study, if all three genes could be targeted by a single herb, then this herb was defined as a target medicine of IPF. Then, Sea Buckthorn and Gnaphalium Affine were predicted to be potential therapeutic TCMs for IPF, and the intervention mechanism of these herbs on the hub genes were verified in TGF- β 1-induced MRC-5 cells. The qPCR result showed that the expression level of *VEGFA* and *SPP1* was significantly up-regulated in the model group ($p < 0.001$ and $p < 0.05$, respectively), which was identical with the result of TGF- β 1-induced HFL cell and animal experiment. After the intervention of Sea Buckthorn, the direction of *VEGFA* and *SPP1* were significantly reversed ($p < 0.05$). Similarly, compared with the model group, after the intervention of Gnaphalium Affine, the

direction of *VEGFA* was significantly reduced ($p < 0.001$), while the direction of *SPP1* had a decline trend with weak statistical significance ($p = 0.073$), as shown in Fig. 7. The above results indicated that Sea Buckthorn and Gnaphalium Affine have a certain curative effect in the treatment of IPF.

Discussion

Based on multiple gene-expression data from human lung tissues, we successfully identified DEGs by comparing IPF tissues with normal lung tissues. Then, following PPI network and Cytoscape software analyses, 18 hub genes were identified. Among those genes, we found that *CAV1* and *PECAM1* were specifically expressed in lung tissues, *BMP4* and *VEGFA* were related to angiogenesis, *FYN* and *SPP1* were related to immunity or inflammation, and *COL1A1* was closely associated with pulmonary fibrosis according to the GeneCards database. All of these seven genes were well-verified in the independent GEO dataset GSE24206 (Wilcoxon rank-sum test, $p < 0.05$).

miRNAs are single-stranded, endogenous and non-coding RNAs that function as negative regulators of gene

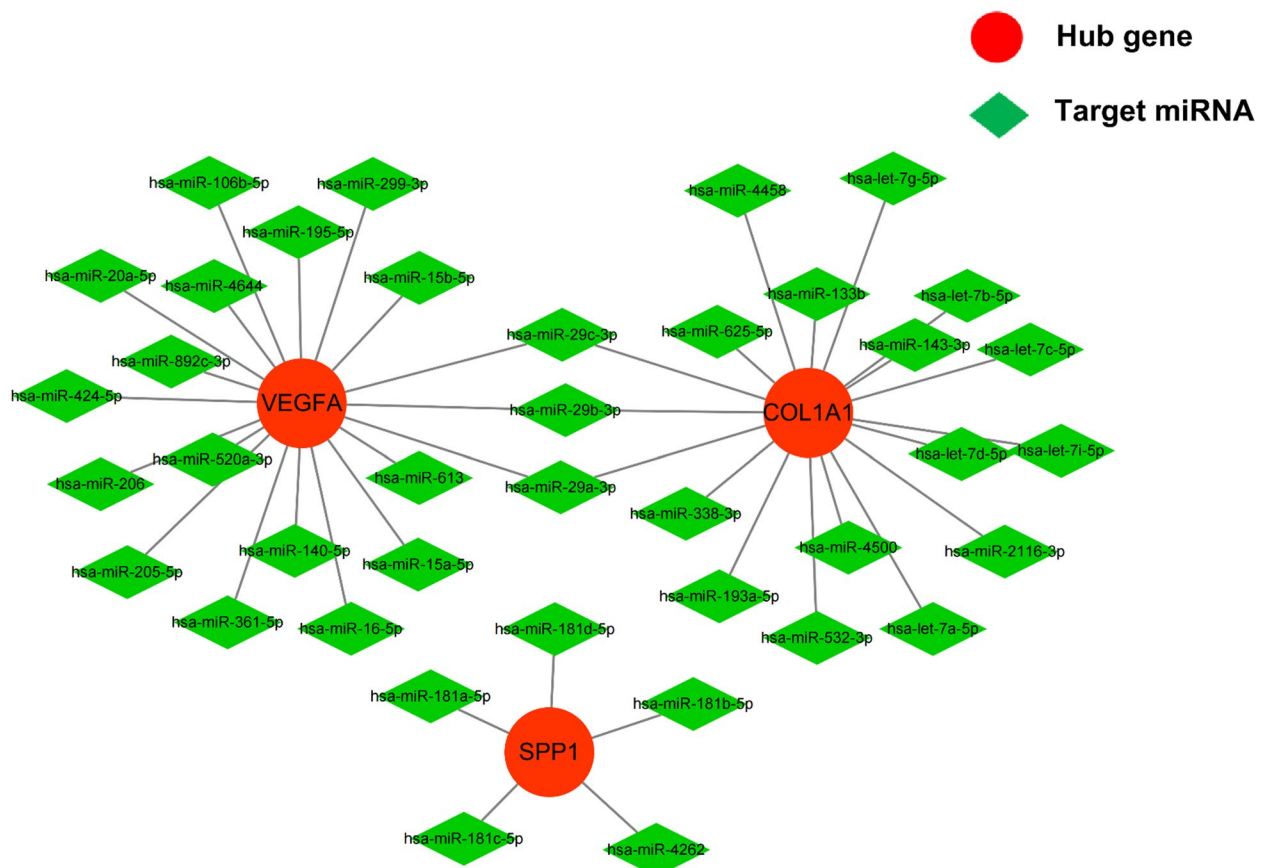


Fig. 3 The co-expressed network of mRNAs and target miRNAs. The miRNA–mRNA co-expressed network was constructed by Cytoscape, and included 42 nodes, 42 edges. Red circles represent the hub genes, while green diamonds represent miRNAs

expression at the post-transcriptional level, which could fully or partially bind with complementary sequences in the 3'UTR of the target mRNA, thereby degrading the mRNA or inhibiting its translation [45]. Previous study showed that there was about 10% of miRNAs differ between IPF and control lungs for miRNA microarrays data, miRNAs and the identified potential target genes may contribute to understanding the complex transcriptional programs of pulmonary fibrosis [46]. The network of miRNAs–mRNAs is involved in many important pathways, and there is an urgent need to elucidate the exact mechanisms, which may provide a new approach to identify gene functions and mechanisms involved in IPF pathogenesis [47]. As an endogenous competitive RNA, lncRNA plays an important role in lung disease, tumor, etc. Studies have found that lncRNA could regulate the binding of miRNA and its target mRNA, and is an emerging key regulator of different cellular processes [48, 49]. Therefore, the ceRNA network of lncRNA–miRNA–mRNA was constructed to explore the pathogenesis of IPF. Among the aforementioned seven genes, the ceRNA networks of *SPP1*, *COL1A1*, and *VEGFA* were

constructed, and the three genes were identified as IPF hub genes.

The identified hub genes were verified by qPCR in samples of TGF- β 1-induced HFL fibroblast cells and pulmonary fibrosis mice. Finally, Sea Buckthorn and Gnaphalium Affine were predicted to be potential therapeutic TCMs that target IPF hub genes by the SymMap database, and the intervention mechanism of these herbs were verified by qPCR in TGF- β 1-induced MRC-5 cells. We found that these herbs could reverse the expression level of hub gene *VEGFA* and *SPP1* in TGF- β 1-induced MRC-5 cells. These results showed that the hub genes and predicted herbs have higher reliability and might play vital role in IPF. Further investigation of these genes and TCMs might provide ideas for the research of IPF mechanism and the intervention of TCMs in IPF patient treatment, and perhaps also serve for clinical scenario in the future.

Studies have shown that *SPP1*, *COL1A1*, and *VEGFA* are closely related to IPF. *SPP1* is a multifunctional protein that plays a vital part in cell–ECM interactions and the pathogenesis of IPF; its silencing is very

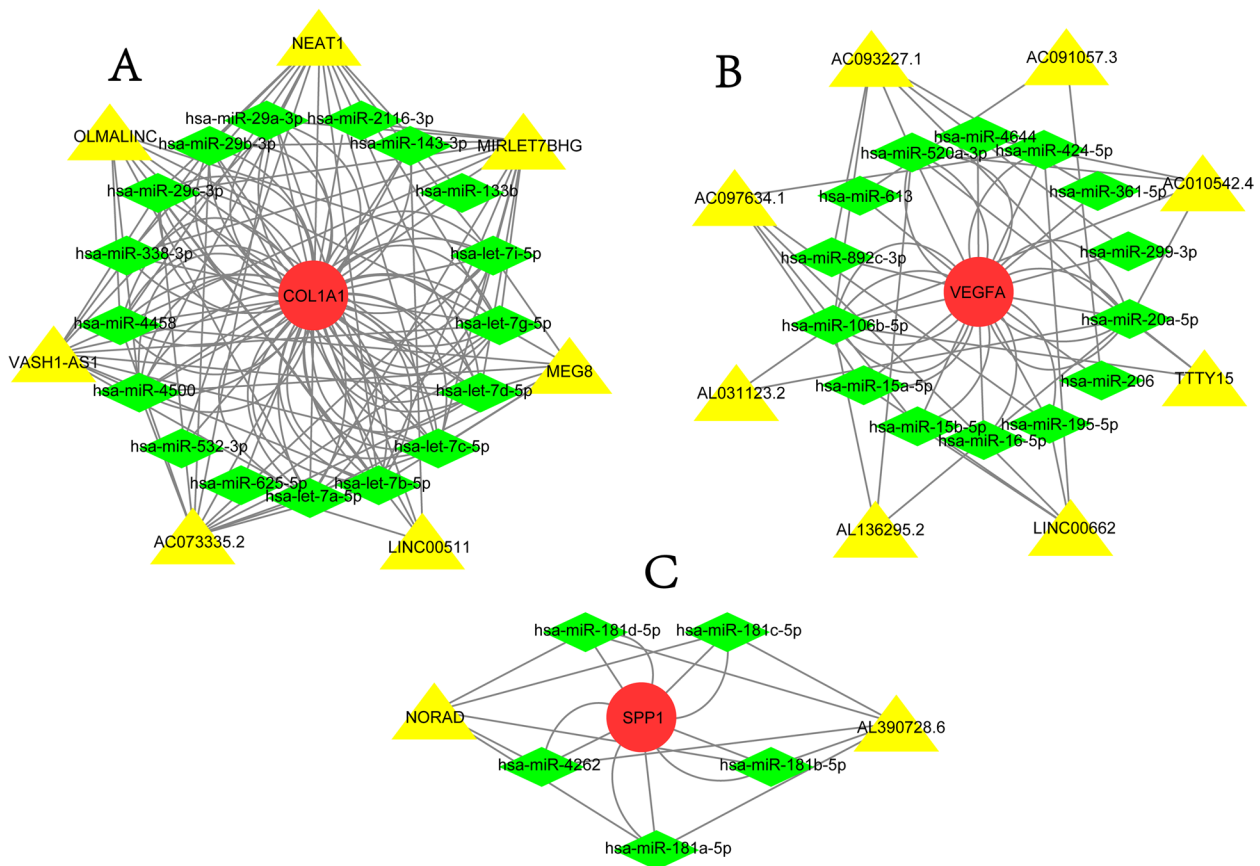


Fig. 4 The ceRNA networks of (A) *COL1A1* (25 nodes and 142 edges), (B) *VEGFA* (23 nodes and 60 edges), and (C) *SPP1* (8 nodes and 20 edges). Red circles represent the hub genes, green diamonds represent miRNAs, and yellow triangles represent lncRNAs

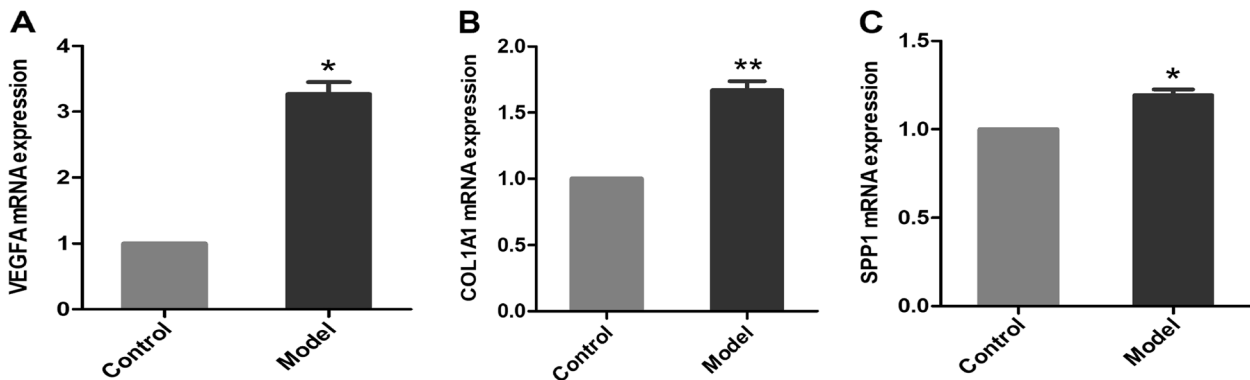


Fig. 5 The mRNA expression levels of (A) *VEGFA*, (B) *COL1A1*, and (C) *SPP1* in the HFL fibroblast cells model group and control group ($n=3$ for each group). The mean \pm standard deviation format is used for the expression of values. Compared to the control group, * denotes $p < 0.05$ and ** denotes $p < 0.01$

useful in the treatment of IPF [50]. Moreover, *SPP1* has been shown to have both pro-inflammatory and pro-fibrotic properties, and the *SPP1* gene-expression level was found to increase in mouse models of pulmonary fibrosis and patients with IPF [51]. Lung fibrosis

is characterized by the deposition of ECM proteins, including *COL1A1*, combined with collagen I as the most abundant ECM component in both normal and fibrotic conditions. Collagen I has a triple helix structure that arises from two α -1 and one α -2 chains, which

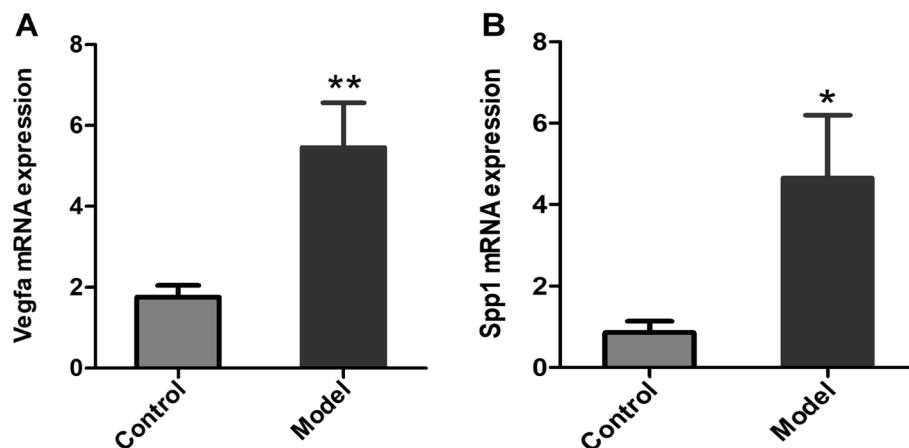


Fig. 6 The mRNA expression levels of (A) *Vegfa* and (B) *Spp1* in the mouse model group and control group ($n=3$ for each group). Compared to the control group, * denotes $p < 0.05$ and ** denotes $p < 0.01$

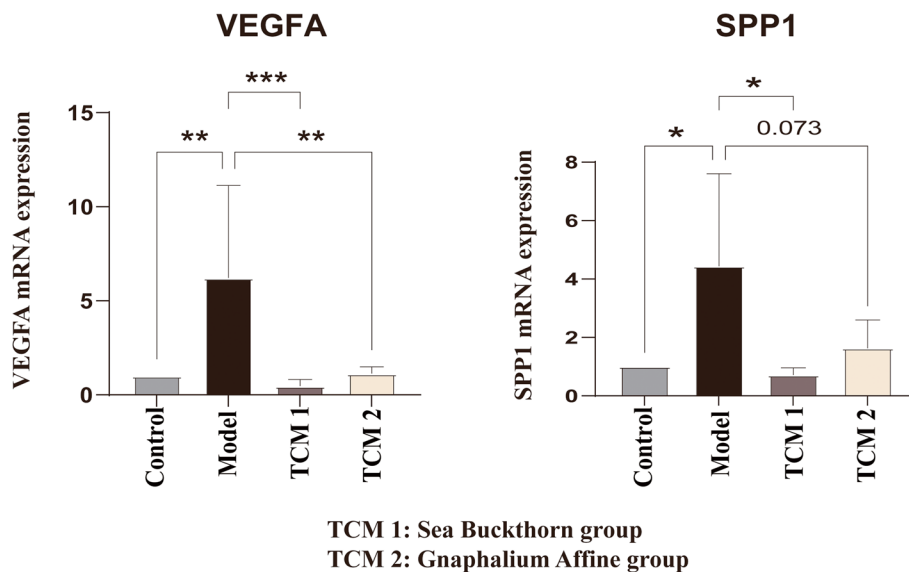


Fig. 7 The mRNA expression levels of *VEGFA*, *SPP1* in the MRC-5 cells control group, model group, TCM1 group (Sea Buckthorn group) and TCM2 group (Gnaphalium Affine group) ($n=3$ for each group). The mean \pm standard deviation format is used for the expression of values. * denotes $p < 0.05$, ** denotes $p < 0.001$, *** denotes $p < 0.0001$

are the products of the *COL1A1* and *COL1A2* genes, respectively; down-regulation of *COL1A1* alleviates the accumulation of ECM and then inhibits pulmonary fibrosis [52–54]. *VEGFA* is a growth factor that is active in angiogenesis and endothelial cell growth [55, 56]. Dysregulation of *VEGFA* bioavailability is associated with the progression of lung injury/fibrosis [57]. Several clinical studies have documented a reparative effect of *VEGFA* in the lungs, and it has also been shown that *VEGFA* has a protective effect against the development of excessive pulmonary fibrosis caused by lung injury [44, 58]. Thus, *SPP1*, *COL1A1*, and *VEGFA*

play crucial roles in the occurrence and development of IPF and deserve further study.

Compared to control samples, we found that the expression levels of *COL1A1*, *SPP1*, and *VEGFA* were significantly up-regulated in TGF- β 1-induced HFL fibroblast cells. Similarly, the expression levels of *Vegfa* and *Spp1* were significantly increased in the pulmonary fibrosis mouse model. The results of *SPP1* were identical in our data analysis, cell experiments, and mouse experiments. In the cell and mouse experiments, consistent results for *VEGFA* were obtained, and the dysregulated direction of this gene was

significantly opposite that in the data analysis, which might due to the fact that *VEGFA* has a dual role in the occurrence and development of IPF [44]. *COL1A1* showed the same results in our data analysis and cell experiments, and there was no statistically significant finding in the mouse experiments.

Studies have demonstrated that Seabuckthorn Wuwei Pulvis (SWP) could improve lung inflammation and chronic bronchitis in mice, and it could also alleviate chronic obstructive pulmonary disease in rats, improve lung function and inhibit inflammatory response through intestinal microbiota short-chain fatty acid axis [59]. A clinical study suggested that Sea Buckthorn could reduce the levels of collagen types III and IV in patients with liver cirrhosis and is an effective drug for the prevention and treatment of liver fibrosis [60]. Other studies have shown that it may suppress the synthesis of collagen and other ECM components [61, 62]. Gnaphalium Affine is an important TCM, and it was used to treat asthma, hyperuricemia, rheumatic arthritis, antitussive, expectorant and cardiovascular in folk medicine due its anti-inflammatory and anti-oxidant activity [63, 64], which might also work for IPF. Some species of Gnaphalium Affine are used as folk medicines in Mexico to treat various respiratory diseases, such as grippe, fever, asthma, cough, cold, bronchitis, expectoration, and bronchial affections [65]. *GnafC* is derived from Gnaphalium Affine. When adding ascorbate to the culture, *gnafC* enhanced the expression of the *ALPase* and *MMP13* genes from the early stage of differentiation, leading to maturation of the collagenous ECM [66]. IPF is a chronic, progressive, and ultimately fatal interstitial lung disease characterized by intensive ECM accumulation; therefore, Sea Buckthorn and Gnaphalium Affine have a certain therapeutic effect in IPF from the perspective of ECM management [67].

Our study has certain limitations. Relatively lax thresholds ($p < 0.05$) were used in DEG analysis of IPF data from a public database because DEGs are not recognized in GSE15197 when strict thresholds ($FDR < 0.05$) are used. However, in the DEGs of the two datasets (GSE15197 and GSE134692) with $p < 0.05$, it was found that 1885 genes were common, and 87% of these genes (1640) had the same dysregulation direction, which would not happen by chance (binomial test, $p < 1.00E-16$). The results indicate that the DEGs identified at the 0.05 threshold were reliable and the cutoff used was reasonable. Currently, there are few IPF samples of high-throughput data available in public databases. In future research, with greater accumulation of IPF data from public databases, an improved algorithm or the analysis of more IPF data should be considered.

Conclusions

We identified three genes of interest (*SPP1*, *COL1A1*, and *VEGFA*) by bioinformatics analysis that were verified by qPCR in samples of TGF- β 1-induced HFL fibroblast cells and pulmonary fibrosis mice. Based on these genes, Sea Buckthorn and Gnaphalium Affine were predicted as potential therapeutic TCMs for IPF, qPCR results showed that these herbs could reverse the expression level of hub gene *VEGFA* and *SPP1* in TGF- β 1-induced MRC-5 cells. Further experiments are needed to verify the key genes and related Chinese medicines in the future.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12890-023-02678-z>.

Additional file 1: Table S1. The analyze network results of 1640 DEGs. **Table S2.** GO terms of the 18 hub genes. **Table S3.** KEGG pathways of the 18 hub genes. **Table S4.** Target microRNAs of *SPP1* based on five online miRNA databases. **Table S5.** Target microRNAs of *VEGFA* based on five online miRNA databases. **Table S6.** Target microRNAs of *COL1A1* based on five online miRNA databases. **Table S7.** Target microRNAs of *CAV1* based on five online miRNA databases. **Table S8.** Target microRNAs of *PECAM1* based on five online miRNA databases. **Table S9.** Target microRNAs of *BMP4* based on five online miRNA databases. **Table S10.** Target microRNAs of *FYN* based on five online miRNA databases. **Table S11.** Traditional Chinese medicine prediction results of *COL1A1*. **Table S12.** Traditional Chinese medicine prediction results of *VEGFA*. **Table S13.** Traditional Chinese medicine prediction results of *SPP1*.

Acknowledgements

Not applicable.

Authors' contributions

ZZZ and QZG conceived the study, analyzed the data, made figures, performed the statistical analysis, and drafted the manuscript. YGT, XJS, PZ and LDH searched the data and participated in the statistical analysis. JSL participated in discussing and revising the manuscript. QZG and JSL conceived of the study, participated in its design and coordination, helped to draft the manuscript, and supervised the work. All authors read and approved the final manuscript.

Funding

This study was funded by the China Postdoctoral Science Foundation (grant no. 2020M682314) and the Foundation of the Department of Science and Technology of Henan Province (grant no. 222102310535).

Availability of data and materials

The datasets presented in this study can be found in online repositories (<https://www.ncbi.nlm.nih.gov/>). The names of the repository/repositories and accession number(s) can be found in the article.

Declarations

Ethics approval and consent to participate

GEO belongs to public databases. The patients involved in the database have obtained ethical approve. Users can download data for free for research and publish relevant articles. This animal experiment was approved by the Ethics Committee of Laboratory Animal Welfare of the Henan University of Chinese Medicine (DWLL202203014).

We confirmed that all methods were carried out in accordance with relevant guidelines and regulations. We confirmed that all methods are reported

in accordance with ARRIVE guidelines (<https://arriveguidelines.org>) for the reporting of animal experiments.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

¹Academy of Chinese Medical Sciences, Henan University of Chinese Medicine, Zhengzhou 450046, China. ²Henan Key Laboratory of Chinese Medicine for Respiratory Disease, Collaborative Innovation Center for Chinese Medicine and Respiratory Diseases Co-Constructed By Henan Province and Education Ministry of P.R. China, Henan University of Chinese Medicine, Zhengzhou 450046, China. ³Department of Respiratory Diseases, The First Affiliated Hospital of Henan University of Chinese Medicine, Zhengzhou 450000, China.

Received: 19 April 2023 Accepted: 26 September 2023

Published online: 04 October 2023

References

- Hewlett JC, Kropski JA, Blackwell TS. Idiopathic pulmonary fibrosis: Epithelial-mesenchymal interactions and emerging therapeutic targets. *Matrix Biol.* 2018;71–72:112–27.
- Durheim MT, Judy J, Bender S, Baumer D, Lucas J, Robinson SB, Mohamedaly O, Shah BR, Leonard T, Conoscenti CS, et al. In-Hospital Mortality in Patients with Idiopathic Pulmonary Fibrosis: A US Cohort Study. *Lung.* 2019;197(6):699–707.
- Raghu G, Weycker D, Edelsberg J, Bradford WZ, Oster G. Incidence and prevalence of idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med.* 2006;174(7):810–6.
- Fernández Pérez ER, Daniels CE, Schroeder DR, St Sauver J, Hartman TE, Bartholmai BJ, Yi ES, Ryu JH. Incidence, prevalence, and clinical course of idiopathic pulmonary fibrosis: a population-based study. *Chest.* 2010;137(1):129–37.
- Nakamura Y, Suda T. Idiopathic Pulmonary Fibrosis: Diagnosis and Clinical Manifestations. *Clin Med Insights Circ Respir Pulm Med.* 2015;9(Suppl 1):163–71.
- Diamantopoulos A, Wright E, Vlahopoulou K, Cornic L, Schoof N, Maher TM. The Burden of Illness of Idiopathic Pulmonary Fibrosis: A Comprehensive Evidence Review. *Pharmacoeconomics.* 2018;36(7):779–807.
- Kumar A, Kapnadak SG, Girgis RE, Raghu G. Lung transplantation in idiopathic pulmonary fibrosis. *Expert Rev Respir Med.* 2018;12(5):375–85.
- Cerri S, Monari M, Guerrieri A, Donatelli P, Bassi I, Garuti M, Luppi F, Betti S, Bandelli G, Carpano M, et al. Real-life comparison of pirfenidone and nintedanib in patients with idiopathic pulmonary fibrosis: A 24-month assessment. *Respir Med.* 2019;159:105803.
- Zhang Y, Lu P, Qin H, Zhang Y, Sun X, Song X, Liu J, Peng H, Liu Y, Nwafor EO, et al. Traditional Chinese medicine combined with pulmonary drug delivery system and idiopathic pulmonary fibrosis: Rationale and therapeutic potential. *Biomed Pharmacother.* 2021;133:111072.
- Phan THG, Paliogiannis P, Nasrallah GK, Giordo R, Eid AH, Fois AG, Zinellu A, Mangoni AA, Pintus G. Emerging cellular and molecular determinants of idiopathic pulmonary fibrosis. *Cell Mol Life Sci.* 2021;78(5):2031–57.
- Schäfer SC, Funke-Chambour M, Berezowska S. Idiopathic pulmonary fibrosis-epidemiology, causes, and clinical course. *Pathologie.* 2020;41(1):46–51.
- Mei Q, Liu Z, Zuo H, Yang Z, Qu J. Idiopathic Pulmonary Fibrosis: An Update on Pathogenesis. *Front Pharmacol.* 2021;12:797292.
- Zhang Y, Jiang M, Nourai M, Roth MG, Tabib T, Winters S, Chen X, Sembrat J, Chu Y, Cardenas N, et al. GDF15 is an epithelial-derived biomarker of idiopathic pulmonary fibrosis. *Am J Physiol Lung Cell Mol Physiol.* 2019;317(4):L510–21.
- Son J-H, Lee J-U, Chin S, Go E-S, Park J-S, Shin H-K, Chang HS, Park J-S, Park C-S. Upregulation of receptor tyrosine kinase-like orphan receptor 2 in idiopathic pulmonary fibrosis. *Korean J Intern Med.* 2021;36(4):914–23.
- Lee CM, Cho SJ, Cho WK, Park JW, Lee JH, Choi AM, Rosas IO, Zheng M, Peltz G, Lee CG, et al. Laminin $\alpha 1$ is a genetic modifier of TGF- $\beta 1$ -stimulated pulmonary fibrosis. *JCI Insight.* 2018;3(18):e99574.
- Yang Z, Zhuan B, Yan Y, Jiang S, Wang T. Identification of gene markers in the development of smoking-induced lung cancer. *Gene.* 2016;576(1 Pt 3):451–7.
- Panebianco F, Mazzanti C, Tomei S, Aretini P, Franceschi S, Lessi F, Di Coscio G, Bevilacqua G, Marchetti I. The combination of four molecular markers improves thyroid cancer cytologic diagnosis and patient management. *BMC Cancer.* 2015;15:918.
- Wu Y, Deng J, Lai S, You Y, Wu J. A risk score model with five long non-coding RNAs for predicting prognosis in gastric cancer: an integrated analysis combining TCGA and GEO datasets. *PeerJ.* 2021;9:e10556.
- Salmena L, Poliseno L, Tay Y, Kats L, Pandolfi PP. A ceRNA hypothesis: the Rosetta Stone of a hidden RNA language? *Cell.* 2011;146(3):353–8.
- Guan Q, Tian Y, Zhang Z, Zhang L, Zhao P, Li J. Identification of Potential Key Genes in the Pathogenesis of Chronic Obstructive Pulmonary Disease Through Bioinformatics Analysis. *Front Genet.* 2021;12:754569.
- Edgar R, Domrachev M, Lash AE. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res.* 2002;30(1):207–10.
- Jeong H, Mason SP, Barabási AL, Oltvai ZN. Lethality and centrality in protein networks. *Nature.* 2001;411(6833):41–2.
- Chin CH, Chen SH, Wu HH, Ho CW, Ko MT, Lin CY. cytoHubba: identifying hub objects and sub-networks from complex interactome. *BMC Syst Biol.* 2014;8 Suppl 4(Suppl 4):S11.
- Matenga JA, Russell J, Paul B, Thomas JE. Suicide due to caustic soda—a case report with some unusual complications. *Cent Afr J Med.* 1987;33(3):77–9.
- Chin CS, Samanta MP. Global snapshot of a protein interaction network—a percolation based approach. *Bioinformatics.* 2003;19(18):2413–9.
- Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet.* 2000;25(1):25–9.
- Aleksander SA, Balhoff J, Carbon S, Cherry JM, Drabkin HJ, Ebert D, Feuerman M, Gaudet P, Harris NL, Hill DP, et al. The Gene Ontology knowledgebase in 2023. *Genetics.* 2023;224(1):iyad031.
- Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* 2000;28(1):27–30.
- Kanehisa M. Toward understanding the origin and evolution of cellular organisms. *Protein Sci.* 2019;28(11):1947–51.
- Kanehisa M, Furumichi M, Sato Y, Kawashima M, Ishiguro-Watanabe M. KEGG for taxonomy-based analysis of pathways and genomes. *Nucleic Acids Res.* 2023;51(D1):D587–92.
- Fishilevich S, Zimmerman S, Kohn A, Iny Stein T, Olender T, Kolker E, Safran M, Lancet D. Genic insights from integrated human proteomics in GeneCards. *Database (Oxford).* 2016;2016:baw030.
- Shi ZE, Zhang MY, Liu JY, Zhang WD, Hu DM, Wang QX, Ji XL, Jiang YY, Qu YQ. Autophagy Induced by BCL2-Related ceRNA Network Participates in the Occurrence of COPD. *Int J Chron Obstruct Pulmon Dis.* 2022;17:791–808.
- Li JH, Liu S, Zhou H, Qu LH, Yang JH. starBase v2.0: decoding miRNA-ceRNA, miRNA-ncRNA and protein-RNA interaction networks from large-scale CLIP-Seq data. *Nucleic Acids Res.* 2014;42(Database issue):D92–7.
- Li J, Li K, Tian Y, Zhao P, Liu X, Li M, Bai Y. Effective-compounds of Jinshui Huanxian formula ameliorates fibroblast activation in pulmonary fibrosis by inhibiting the activation of mTOR signaling. *Phytomedicine.* 2023;109:154604.
- Shao D, Liu X, Wu J, Zhang A, Bai Y, Zhao P, Li J. Identification of the active compounds and functional mechanisms of Jinshui Huanxian formula in pulmonary fibrosis by integrating serum pharmacochimistry with network pharmacology. *Phytomedicine.* 2022;102:154177.
- Bai Y, Li J, Zhao P, Li Y, Li M, Feng S, Qin Y, Tian Y, Zhou T. A Chinese Herbal Formula Ameliorates Pulmonary Fibrosis by Inhibiting Oxidative Stress via Upregulating Nrf2. *Front Pharmacol.* 2018;9:628.
- Wu Y, Zhang F, Yang K, Fang S, Bu D, Li H, Sun L, Hu H, Gao K, Wang W, et al. SymMap: an integrative database of traditional Chinese medicine enhanced by symptom mapping. *Nucleic Acids Res.* 2019;47(D1):D1110–7.

38. Chen H, Wang J, Li R, Lv C, Xu P, Wang Y, Song X, Zhang J. Astaxanthin attenuates pulmonary fibrosis through IncITPF and mitochondria-mediated signal pathways. *J Cell Mol Med*. 2020;24(17):10245–50.
39. Ye Z, Hu Y. TGF- β 1: Gentlemanly orchestrator in idiopathic pulmonary fibrosis (Review). *Int J Mol Med*. 2021;48(1):132.
40. Allen RJ, Guillen-Guio B, Oldham JM, Ma SF, Dressen A, Paynton ML, Kraven LM, Obeidat Me, Li X, Ng M et al. Genome-Wide Association Study of Susceptibility to Idiopathic Pulmonary Fibrosis. *Am J Respir Crit Care Med* 2020;201(5):564–574.
41. Chanda D, Otoupalova E, Smith SR, Volckaert T, De Langhe SP, Thannickal VJ. Developmental pathways in the pathogenesis of lung fibrosis. *Mol Aspects Med*. 2019;65:56–69.
42. Woodcock HV, Eley JD, Guillotin D, Platé M, Nanthakumar CB, Martufi M, Peace S, Joberty G, Poecckel D, Good RB, et al. The mTORC1/4E-BP1 axis represents a critical signaling node during fibrogenesis. *Nat Commun*. 2019;10(1):6.
43. Wang J, Hu K, Cai X, Yang B, He Q, Wang J, Weng Q. Targeting PI3K/AKT signaling for treatment of idiopathic pulmonary fibrosis. *Acta Pharmaceutica Sinica B*. 2022;12(1):18–32.
44. Barratt SL, Blythe T, Jarrett C, Ourradi K, Shelley-Fraser G, Day MJ, Qiu Y, Harper S, Maher TM, Oltean S, et al. Differential Expression of VEGF-A(xxx) Isoforms Is Critical for Development of Pulmonary Fibrosis. *Am J Respir Crit Care Med*. 2017;196(4):479–93.
45. Sun XF, Li YP, Pan B, Wang YF, Li J, Shen W. Molecular regulation of miR-378 on the development of mouse follicle and the maturation of oocyte in vivo. *Cell Cycle*. 2018;17(18):2230–42.
46. Li S, Geng J, Xu X, Huang X, Leng D, Jiang D, Liang J, Wang C, Jiang D, Dai H. miR-130b-3p Modulates Epithelial-Mesenchymal Crosstalk in Lung Fibrosis by Targeting IGF-1. *PLoS ONE*. 2016;11(3):e0150418.
47. Wang L, Huang W, Zhang L, Chen Q, Zhao H. Molecular pathogenesis involved in human idiopathic pulmonary fibrosis based on an integrated microRNA-mRNA interaction network. *Mol Med Rep*. 2018;18(5):4365–73.
48. Wang Z, Qu S, Zhu J, Chen F, Ma L. Comprehensive analysis of lncRNA-associated competing endogenous RNA network and immune infiltration in idiopathic pulmonary fibrosis. *J Thorac Dis*. 2020;12(5):1856–65.
49. Sun J, Su W, Zhao X, Shan T, Jin T, Guo Y, Li C, Li R, Zhou Y, Shan H, et al. lncRNA PFAR contributes to fibrogenesis in lung fibroblasts through competitively binding to miR-15a. *Biosci Rep*. 2019;39(7):BSR20190280.
50. Hatipoglu OF, Uctepe E, Opoku G, Wake H, Ikemura K, Ohtsuki T, Inagaki J, Gunduz M, Gunduz E, Watanabe S, et al. Osteopontin silencing attenuates bleomycin-induced murine pulmonary fibrosis by regulating epithelial-mesenchymal transition. *Biomed Pharmacother*. 2021;139:11633.
51. Pardo A, Gibson K, Cisneros J, Richards TJ, Yang Y, Becerril C, Yousem S, Herrera I, Ruiz V, Selman M, et al. Up-regulation and profibrotic role of osteopontin in human idiopathic pulmonary fibrosis. *PLoS Med*. 2005;2(9):e251.
52. Dong J, Ma Q. Myofibroblasts and lung fibrosis induced by carbon nanotube exposure. *Part Fibre Toxicol*. 2016;13(1):60.
53. Burgess JK, Mauad T, Tjin G, Karlsson JC, Westergren-Thorsson G. The extracellular matrix - the under-recognized element in lung disease? *J Pathol*. 2016;240(4):397–409.
54. Giménez A, Duch P, Puig M, Gabasa M, Xaubet A, Alcaraz J. Dysregulated Collagen Homeostasis by Matrix Stiffening and TGF- β 1 in Fibroblasts from Idiopathic Pulmonary Fibrosis Patients: Role of FAK/Akt. *Int J Mol Sci*. 2017;18(11):2431.
55. Pedram A, Razandil M, Hu RM, Levin ER. Vasoactive peptides modulate vascular endothelial cell growth factor production and endothelial cell proliferation and invasion. *J Biol Chem*. 1997;272(27):17097–103.
56. Evans IM, Britton G, Zachary IC. Vascular endothelial growth factor induces heat shock protein (HSP) 27 serine 82 phosphorylation and endothelial tubulogenesis via protein kinase D and independent of p38 kinase. *Cell Signal*. 2008;20(7):1375–84.
57. Barratt SL, Blythe T, Ourradi K, Jarrett C, Welsh GI, Bates DO, Millar AB. Effects of hypoxia and hyperoxia on the differential expression of VEGF-A isoforms and receptors in Idiopathic Pulmonary Fibrosis (IPF). *Respir Res*. 2018;19(1):9.
58. Stockmann C, Kerdiles Y, Nomaksteinsky M, Weidemann A, Takeda N, Doedens A, Torres-Collado AX, Iruela-Arispe L, Nizet V, Johnson RS. Loss of myeloid cell-derived vascular endothelial growth factor accelerates fibrosis. *Proc Natl Acad Sci USA*. 2010;107(9):4329–34.
59. Wang J, Ren C, Jin L, Batu W. Seabuckthorn Wuwei Pulvis attenuates chronic obstructive pulmonary disease in rat through gut microbiota-short chain fatty acids axis. *J Ethnopharmacol*. 2023;314:116591.
60. Gao ZL, Gu XH, Cheng FT, Jiang FH. Effect of sea buckthorn on liver fibrosis: a clinical study. *World J Gastroenterol*. 2003;9(7):1615–7.
61. Liu F, Wang T, Li X, Jia J, Lin Q, Xue Y. Involvement of NF- κ B in the reversal of CYP3A down-regulation induced by sea buckthorn in BCG-induced rats. *PLoS ONE*. 2020;15(9):e0238810.
62. Liu X, Xing Y, Yuen M, Yuen T, Yuen H, Peng Q. Anti-Aging Effect and Mechanism of Proanthocyanidins Extracted from Sea buckthorn on Hydrogen Peroxide-Induced Aging Human Skin Fibroblasts. *Antioxidants (Basel, Switzerland)*. 2022;11(10):1900.
63. Huang D, Chen Y, Chen W, Liu Y, Yao F, Xue D, Sun L. Anti-inflammatory effects of the extract of *Gnaphalium affine* D. Don in vivo and in vitro. *J Ethnopharmacol*. 2015;176:356–64.
64. Meng XM, He CX, Chen X, Yang XS, Liu C. The extract of *Gnaphalium affine* D. Don protects against H₂O₂-induced apoptosis by targeting PI3K/AKT/GSK-3 β signaling pathway in cardiomyocytes. *J Ethnopharmacol*. 2021;268:113579.
65. Zheng X, Wang W, Piao H, Xu W, Shi H, Zhao C. The genus *Gnaphalium* L. (Compositae): phytochemical and pharmacological characteristics. *Molecules (Basel, Switzerland)*. 2013;18(7):8298–318.
66. Aoshima Y, Hasegawa Y, Hasegawa S, Nagasaka A, Kimura T, Hashimoto S, Torii Y, Tsukagoshi N. Isolation of GnafC, a polysaccharide constituent of *Gnaphalium affine*, and synergistic effects of GnafC and ascorbate on the phenotypic expression of osteoblastic MC3T3-E1 cells. *Biosci Biotechnol Biochem*. 2003;67(10):2068–74.
67. Heukels P, Moor CC, von der Thüsen JH, Wijsenbeek MS, Kool M. Inflammation and immunity in IPF pathogenesis and treatment. *Respir Med*. 2019;147:79–91.

Publisher's Note

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

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

