Int J Clin Exp Pathol 2026;19(1):54-68
www.ijcep.com /ISSN:1936-2625/IJCEP0169465

Original Article

A seven-gene panel with diagnostic and prognostic value for idiopathic pulmonary fibrosis and candidate herbal drugs identified based on these genes

Yu Song, Xianglin Meng, Jiaqi Tian, Yunhe Hao, Kai Zeng, Likun Zhang, Jiannan Zhang, Kaili Zhang, Yu Xin, Changsong Wang, Kaijiang Yu

Key Laboratory of Critical Care Medicine of Heilongjiang Province, Department of Critical Care Medicine, The First Affiliated Hospital of Harbin Medical University, Harbin 150001, Heilongjiang, China

Received October 22, 2025; Accepted December 26, 2025; Epub January 15, 2026; Published January 30, 2026

Abstract: Background: The objective of this study was to identify key idiopathic pulmonary fibrosis (IPF) related genes, thereby establishing a novel IPF diagnostic/warning panel and proposing drugs against IPF based on the strategy of targeting key genes. Methods: The GEO datasets GSE245965, GSE279637, and GSE235435 were used to select IPF-related genes, as well as the IPF associated genes from the GeneCards and DisGeNET databases. The DEGs were used for enrichment analysis, PPI network construction, and targeted therapeutic value analysis. Results: An intersection analysis yielded 60 commonly up-regulated genes and 16 commonly down-regulated genes. GO/KEGG/Reactome/Immunologic Signature terms that were novel and interesting were found to be enriched. In the interaction network, WDR90 and ANKRD1 were identified as hub genes. Among the 60 common up-regulated genes, seven (namely SERPINB3, TUBB3, SERPINB4, CHTF18, BAX, WDR90 and ITGAX) were shared by the disease sets. In the Symmap database, we found some herbs with the most targets, such as Lygodii Spora, Smilacis Glabrae Rhizoma, and Aloe. Conclusions: A panel comprising seven key IPF genes was identified, which may have diagnostic and prognostic value for IPF. A comprehensive analysis of the Dgidb database revealed potential drugs that may be antitumor agents against IPF, such as Lygodii Spora, Smilacis Glabrae Rhizoma, and Aloe.

Keywords: Idiopathic pulmonary fibrosis, pneumonia, IPF, bioinformatics, enrichment analysis

Introduction

Idiopathic pulmonary fibrosis (IPF) is a progressive disease. It is the most common and fatal type of idiopathic interstitial pneumonia [1]. IPF is characterized by cicatricial fibrosis. The median survival after diagnosis is 2-3 years, and its progression is highly unpredictable [2-4]. At present, the etiology and molecular mechanisms of IPF are still not fully understood.

The diagnosis of IPF requires the integration of clinical, radiologic, and, in selected cases, pathologic findings, following the exclusion of other known causes of interstitial lung disease (ILD). The cornerstone of diagnosis is high-resolution computed tomography (HRCT). According to contemporary international guidelines, the diagnosis is based on identifying a usual interstitial pneumonia (UIP) pattern. The diagnostic process involves a multidisciplinary discussion (MDD) among pulmonologists, radiologists, and pathologists. Blood tests are primarily used to rule out connective tissue diseases or other secondary causes of ILD and are not definitive for IPF diagnosis [5-7]. The core pathologic feature of IPF is the UIP pattern, which is characterized by temporal heterogeneity, fibroblastic foci, and honeycomb lung changes. The core pathophysiologic process is repetitive alveolar epithelial injury, which leads to abnormal epithelial cell activation and release of pro-fibrotic factors (especially TGF-β), which drives fibroblast/myofibroblast activation and proliferation, resulting in excessive extracellular matrix deposition and aberrant remodeling, and the development of progressive, irreversible pulmonary fibrosis and structural destruction. These lead to clinical manifestations such as pro-

https://doi.org/10.62347/STNF1901

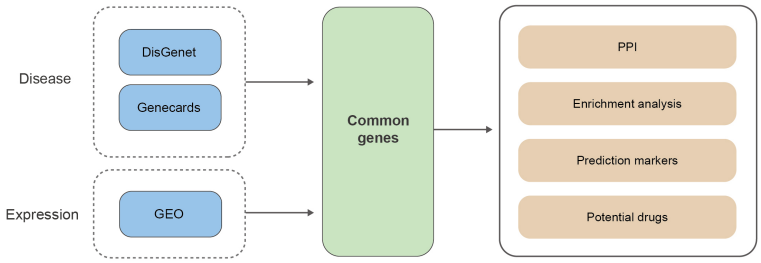

Figure 1. Overall flowchart of this study.

gressive dyspnea, intractable hypoxemia, and respiratory failure [8-14]. Currently, there have been several studies focusing on genes associated with IPF survival [15-20]. Treatment strategies for IPF encompass pharmacologic therapy to slow disease progression and comprehensive supportive care. For disease-modifying pharmacotherapy, Pirfenidone and Nintedanib are the two foundational anti-fibrotic drugs. Supportive and non-pharmacologic managements include pulmonary rehabilitation, oxygen therapy, and lung transplantation [21-23]. However, there is still a lack of effective therapeutic strategies due to the high difficulty in treating IPF. Given that IPF is a progressive, irreversible, and rapidly advancing disease, and once diagnosed, the survival period is short, early diagnosis of IPF has tremendous clinical significance. In this regard, biomarker panels are still in short supply. Previously, in order to explore effective markers, some studies have explored key genes for IPF in terms of genes related to a certain mechanism (e.g., telomerase- or autophagy-associated genes) [15, 24]. However, this strategy tends to have a high subjective bias, and unbiased alveolar lavage fluid/blood markers are still limited. In addition, there are few clearly effective drugs for IPF. In addition, there are nearly no clearly effective drugs for IPF. Chinese herbal medicines, due to their safety and rich variety, can often have unexpected effects on intractable progressive diseases. There are very few studies that have explored the adjunctive therapeutic effects of herbal medicines in IPF, for example angelicae sinensis radix [25], oxytropis falcata bunge [26], and Astragalus radix [27]. Nevertheless, these studies involved very few herbal medicines and did not meet clinical needs. In this study, we obtained a novel 7-gene IPF diagnostic/warning panel based on bioinformatic methods and proposed several candidate herbal medicines against IPF.

Materials and methods

IPF-related datasets and genes

The overall flowchart of this study was shown in **Figure 1**. First, we searched for IPF-associated datasets in the GEO datasets. The criteria were: (1) comparison of two groups: IPF vs. Control (the IPF group should include lung tissue samples from the confirmed diagnosed patients; and the control samples are from healthy normal or non-IPF patients); (2) the transcriptome sequencing data, focusing on the mRNA expression; (3) at least 3 samples in each group; and (4) focusing on different types of cells and looking for common changes in different cells in lung tissues, which can help predict biomarkers of IPF. Together, three types of samples were found: alveolar epithelial type II cells, exosomes derived from IPF lung fibroblasts (SK-MES-1 cells incubated with exosomes), and lung resident mesenchymal stem cells (as an auxiliary reference). The details of each set were as follows.

GSE245965: Alveolar epithelial type II cells purified from primary human normal and IPF samples that underwent bulk RNAseq and bulk ATACseq profiling.

GSE279637: A study about exosomes derived from IPF lung fibroblasts (DHLF-exosomes), RNA-seq analysis was performed on SK-MES-1 cells incubated with/without DHLF-exosomes.

GSE235435: A transcriptomic analysis to characterize lung resident mesenchymal stem cells from IPF patients.

Finally, the Disgenet (the term: GDA_CURATED_C1800706) and GeneCards (only coding genes and RNA genes) were used to screen for IPF-associated genes.

Differentially expressed genes

All the datasets were analyzed with the online GEO2R tool to acquire differentially expressed genes (DEGs). The fold change (FC) and *P*-value were obtained for each transcript. In this study, the up-regulated DEGs were those with an adjusted $P < 0.05$ and $Log_2 > 1$; and the down-regulated DEGs were those with an adjusted

$P < 0.05$ and $Log_2 > -1$. For visualization of DEGs, the wellplot was drawn by the Yangbo studio online visualization tool (http://yangbostudio.cn). Wellplot is an improved map that combines the characteristics of volcano plot and heat map, which is shaped like a well. The left part of the well bottom is a blue/green colored section with the number of the down-regulated genes, the middle part of the well bottom is a grey colored section with the number of unchanged genes, and the right part of the well bottom is a red/orange colored section with the number of the up-regulated genes. The cut-off criteria for DEGs are labeled at the boundaries of the three sections. The walls of the well are made of two layers of bricks showing the top 20 down-regulated and 20 up-regulated genes (by the fold-change or *p* values), respectively. Wellplot embodies both statistics of the DEGs as the volcano plot and the presentation of the top genes of greatest interest as the heatmap. In addition, Venn plots were used to analyze the co-upregulated genes, co-downregulated genes, and the intersection of DEGs and IPF-related genes in disease databases (Disgenet and Genecards).

Protein-protein interaction (PPI) analysis

The protein-protein interaction of common up-regulated or down-regulated genes was analyzed using the STRING database (https://string-db.org/). All interactions with confidence score > 0.4 were used to establish the PPI network.

Enrichment analysis

The Metascape tool was used to explore the enriched Gene Ontology (GO) terms, Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways, Reactome sets, and Immunologic Signatures. Accumulative hypergeometric *p*-values and enrichment factors were calculated and used for filtering. Terms with a *p*-value < 0.05, a minimum count of 2, and an enrichment factor > 1.5 were considered enriched. The top terms of each enrichment were shown as bar plots. The bar graphs were drawn by the Yangbo studio online visualization tool (http://yangbostudio.cn). If there were no more than 20 enriched terms, all terms were presented; otherwise, the top 20 ones were shown.

The key IPF marker panel

There were 60 up-regulated genes shared by different GEO datasets. Following this, the intersection of this 60-gene set, and the disease databases (Disgenet and Genecards) were analyzed. The common ones were considered as key IPF markers, useful to propose an IPF warning panel.

Candidate drugs

Based on the seven key up-regulated markers, we explored therapeutic drugs that simultaneously target most of these upregulated genes. We used the Drug-Gene Interaction database (DGIdb) database to mine therapeutic chemicals and the Symmap database for herbs targeting the key markers. A drug (chemical or herb) was selected if it had as many targets (among the seven key markers) as possible. The drugs with the most targets, as well as the number of drugs for each target, were shown by a horizontal bar graph drawn by the Yangbo studio online visualization tool (http://yangbostudio.cn).

Results

IPF-related DEGs

In the GSE245965 dataset, there were 1559 up-regulated and 1232 down-regulated genes (**Figure 2A**); and in the GSE279637 dataset, there were 363 up-regulated and 148 down-regulated genes (**Figure 2B**). As an auxiliary reference, there were 8 up-regulated and 32 down-regulated genes in the GSE240470 dataset (**Figure 2C**). Among these up-regulated genes, there were 58 common genes shared by GSE245965-up and GSE279637-up, one by GSE245965-up and GSE240470-up, and one by GSE279637-up and GSE240470-up (**Figure 2D**). Among the down-regulated genes, 15 ones were shared by GSE245965-down and GSE279637-down, and one was shared by GSE245965-down and GSE240470-down. Collectively, there were 60 common up-regulated genes and 16 common down-regulated genes.

PPI and enrichment analysis

First, based on the 60 up-regulated proteins (coding genes), a PPI network was generated

Idiopathic pulmonary fibrosis panel and candidate drugs

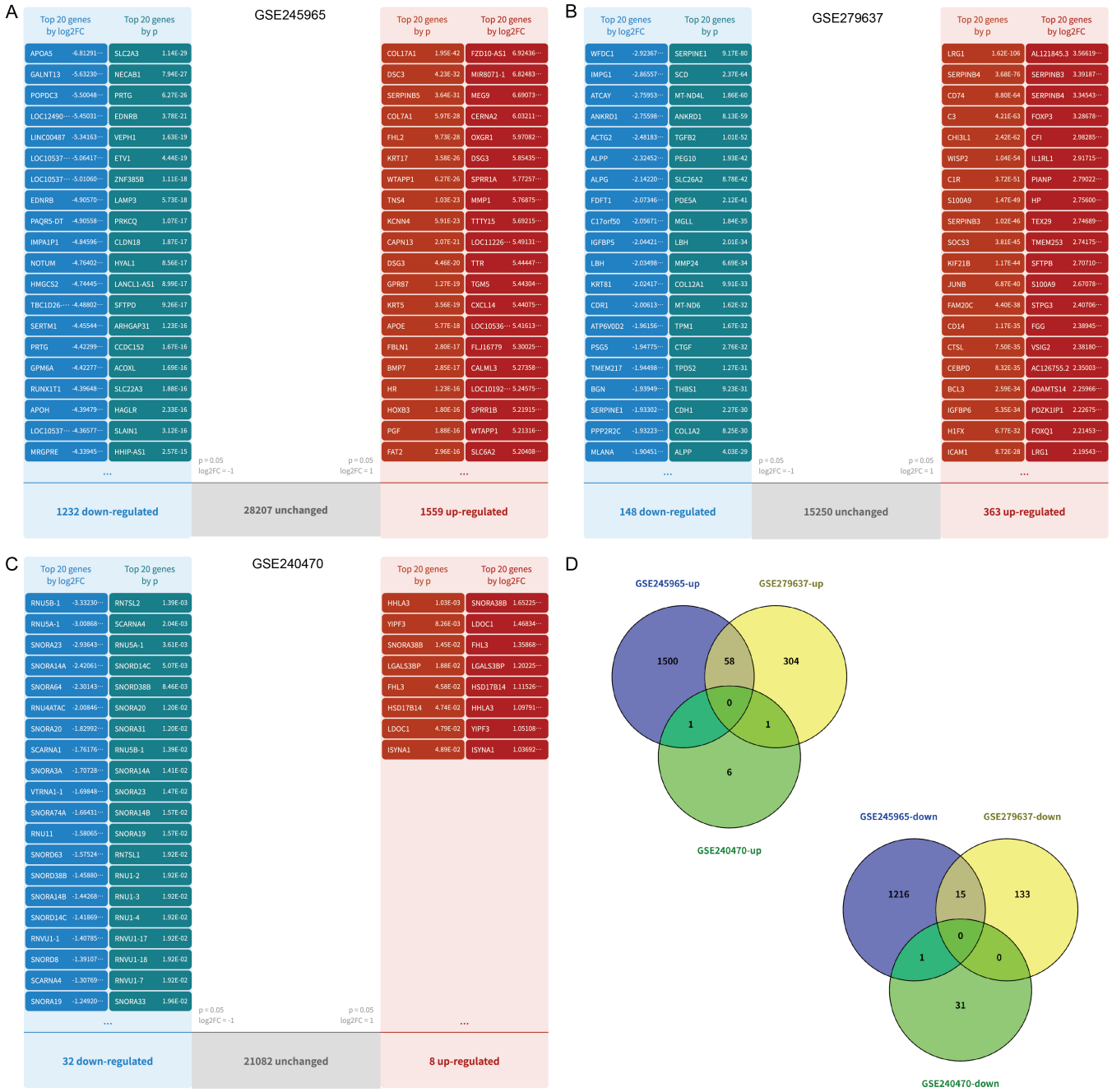

Figure 2. Differentially expressed genes (DEGs) in GEO datasets. A. The wellplot of diagram of DEGs in the GSE245965 dataset. B. The wellplot of diagram of DEGs in the GSE279637 dataset. C. The wellplot of diagram of DEGs in the GSE240470 dataset. D. Among up-regulated genes, there were 58 common genes shared by GSE245965-up and GSE279637-up, one by GSE245965-up and GSE240470-up, and one by GSE279637-up and GSE240470-up. Among down-regulated genes, 15 were shared by GSE245965-down and GSE279637-down, and one was shared by GSE245965-down and GSE240470-down.

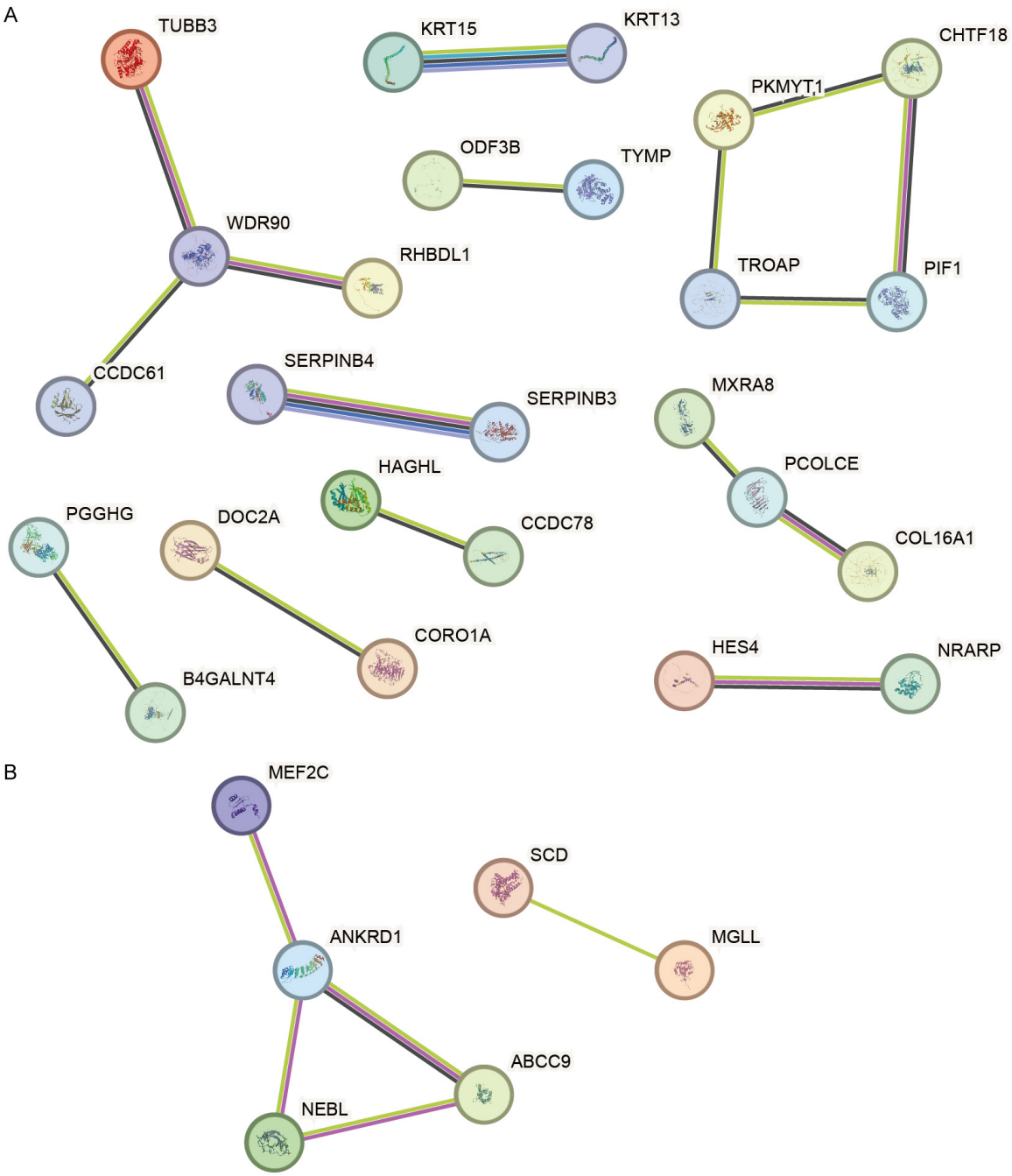

Figure 3. The protein-protein interaction network. A. Protein-protein interaction (PPI) network of the STRING database based on the 60 common up-regulated proteins. B. PPI network based on the 16 down-regulated proteins.

(**Figure 3A**), in which WDR90 was the most significant hub gene. Next, within the PPI network based on the 16 down-regulated proteins, ANKRD1 was the hub gene (**Figure 3B**).

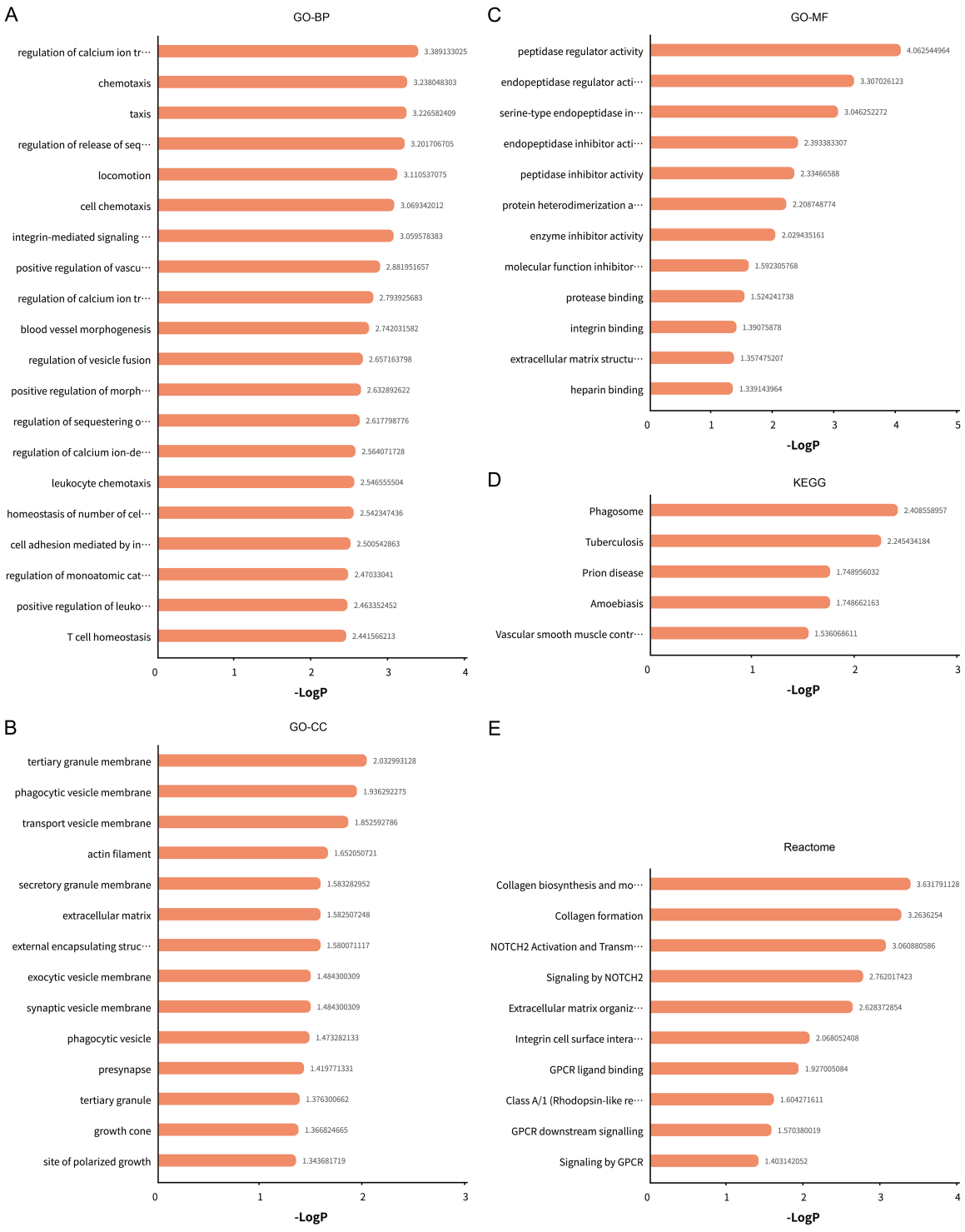

Figure 4. GO, KEGG, and reactome enrichment analysis based on the 60 common up-regulated genes. A. The enriched GO-BP pathways. B. The enriched GO-BP terms. C. The enriched GO-MF terms. D. The enriched KEGG pathways. E. The enriched reactome terms.

Subsequently, enrichment analysis was performed based on the 60 up-regulated genes. The enriched GO Biological Processes (BP) terms were presented in **Figure 4A**, and the top enriched terms include regulation of calcium ion transmembrane transport, chemotaxis,

taxis, regulation of release of sequestered calcium ion into cytosol, locomotion, cell chemotaxis, integrin-mediated signaling pathway, positive regulation of vascular endothelial cell proliferation, regulation of calcium ion transport, blood vessel morphogenesis, regulation of vesicle fusion, positive regulation of morphogenesis of an epithelium, regulation of sequestering of calcium ion, regulation of calcium ion-dependent exocytosis, leukocyte chemotaxis, homeostasis of number of cells within a tissue, cell adhesion mediated by integrin, regulation of monoatomic cation transmembrane transport, positive regulation of leukocyte migration, and T cell homeostasis. The enriched GO Cellular Components (CC) terms are actin filament, extracellular matrix, external encapsulating structure, phagocytic vesicle membrane, phagocytic vesicle, tertiary granule membrane, secretory granule membrane, tertiary granule, transport vesicle membrane, synaptic vesicle membrane, exocytic vesicle membrane, presynapse, growth cone, and site of polarized growth (**Figure 4B**). The enriched GO Molecular Functions (MF) terms are peptidase regulator activity, endopeptidase regulator activity, serine-type endopeptidase inhibitor activity, endopeptidase inhibitor activity, peptidase inhibitor activity, enzyme inhibitor activity, molecular function inhibitor activity, protease binding, extracellular matrix structural constituent, protein heterodimerization activity, integrin binding, and heparin binding (**Figure 4C**). There were five enriched KEGG pathways: Amoebiasis, Phagosome, Tuberculosis, Prion disease, and Vascular smooth muscle contraction (**Figure 4D**). The enriched reactome terms include Collagen biosynthesis and modifying enzymes, Collagen formation, NOTCH2 Activation and Transmission of Signal to the Nucleus, Signaling by NOTCH2, Extracellular matrix organization, Integrin cell surface interactions, GPCR ligand binding, Class A/1 (Rhodopsin-like receptors), GPCR downstream signaling, and Signaling by GPCR (**Figure 4E**).

Based on the down-regulated genes, the enriched terms were as follows. The top 20 enriched GO BP terms were cellular response to transforming growth factor beta stimulus, response to transforming growth factor beta, striated muscle tissue development, muscle tissue development, skeletal muscle tissue development, skeletal muscle organ development, response to muscle stretch, regulation of smooth muscle cell proliferation, muscle structure development, cardiac muscle tissue development, cellular response to xenobiotic stimulus, cellular response to growth factor stimulus, response to growth factor, heart development, striated muscle cell differentiation, heart morphogenesis, regulation of vascular associated smooth muscle cell migration, carboxylic acid biosynthetic process, organic acid biosynthetic process, and muscle cell differentiation (**Figure 5A**). The enriched GO CC terms were sarcomere, myofibril, contractile muscle fiber, I band, transcription regulator complex, actin cytoskeleton, focal adhesion, cell-substrate junction, and receptor complex (**Figure 5B**). The enriched GO CC terms were SMAD binding, histone deacetylase binding, RNA polymerase II-specific DNA-binding transcription factor binding, DNA-binding transcription factor binding, transcription factor binding, protein kinase activity, phosphotransferase activity, and alcohol group as acceptor (**Figure 5C**). The following KEGG pathways were enriched: Regulation of lipolysis in adipocytes, Biosynthesis of cofactors, Oxytocin signaling pathway, cGMP-PKG signaling pathway, Cytoskeleton in muscle cells, and Thermogenesis (**Figure 5D**). The enriched reactome terms were Metabolism of lipids, Cardiac conduction, Epigenetic regulation of gene expression by MLL3 and MLL4 complexes, MLL4 and MLL3 complexes regulate expression of PPARG target genes in adipogenesis and hepatic steatosis, Epigenetic regulation of adipogenesis genes by MLL3 and MLL4 complexes, Metabolism of steroids, Epigenetic regulation by WDR5-containing histone modifying complexes, Muscle contraction, and Epigenetic regulation of gene expression (**Figure 5E**).

A key IPF marker panel

A total of 1464 and 30 genes associated with IPF were identified in the GeneCards and DisGeNET databases, respectively. Of the 60 common up-regulated genes, seven (SERPINB3, TUBB3, SERPINB4, CHTF18, BAX, WDR90, and ITGAX) were shared by the GeneCards set (**Figure 6A**). Furthermore, an analysis of the 16 down-regulated genes revealed that one of these genes (DDR2) was common to the

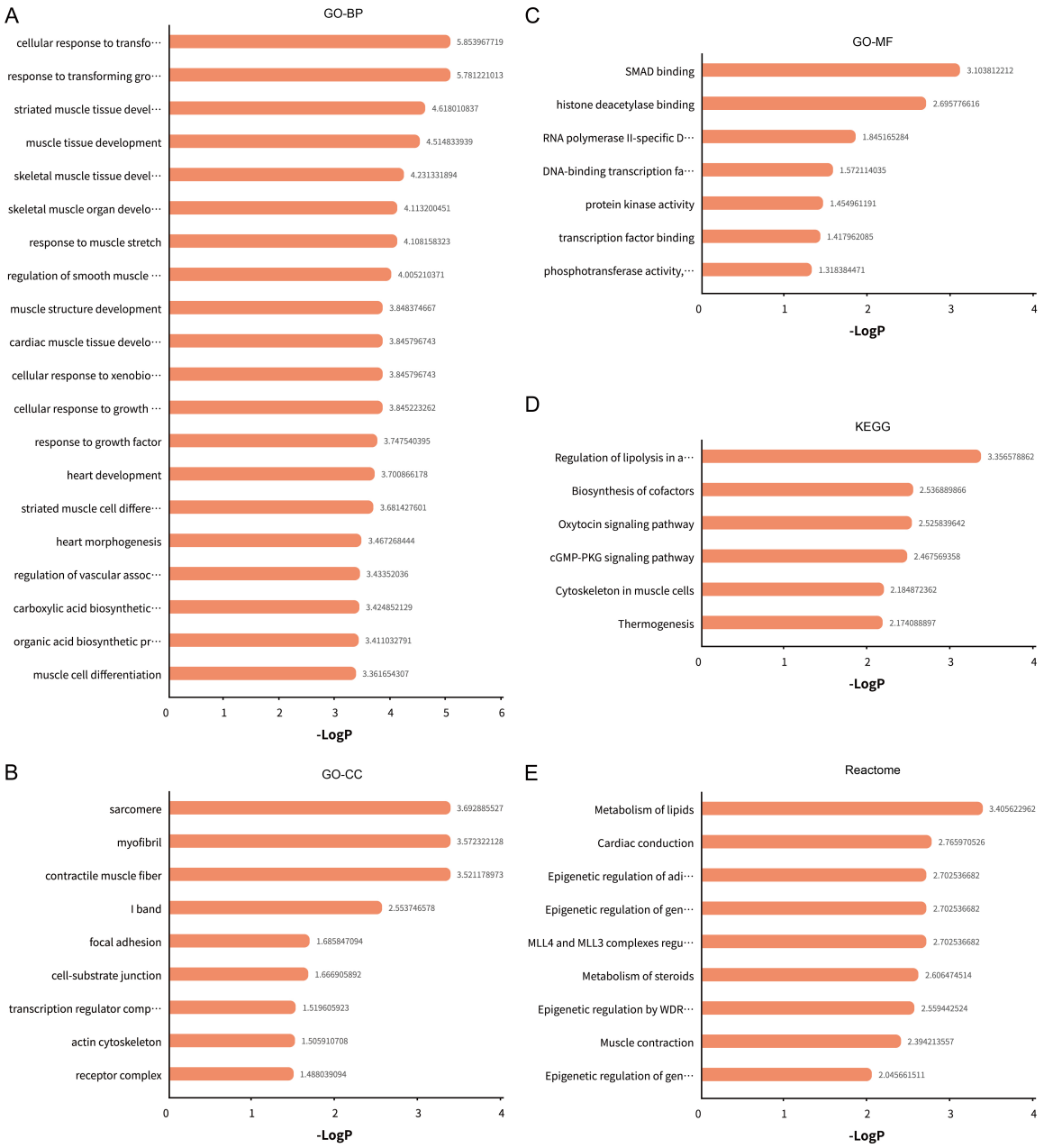

Figure 5. GO, KEGG, and reactome enrichment analysis based on the 16 common down-regulated genes. A. The enriched GO-BP pathways. B. The enriched GO-BP terms. C. The enriched GO-MF terms. D. The enriched KEGG pathways. E. The enriched reactome terms.

GeneCards set (**Figure 6B**). It was observed that no genes were shared by the DisGeNET set. It is important to note that the activation of the DDR2 signal was negatively correlated with IPF (and may even be positively correlated instead). For the time being, this was not used as a negative indicator of IPF. In this study, we proposed seven commonly upregulated genes as key markers of IPF. The use of this panel in the future may facilitate the identification of IPF development through the examination of alveolar lavage fluid or peripheral blood.

An enrichment analysis was performed based on these markers (**Figure 7**). Top enriched GO terms (**Figure 7A** and **7B**) included negative regulation of endopeptidase activity, negative regulation of molecular function, negative regu-

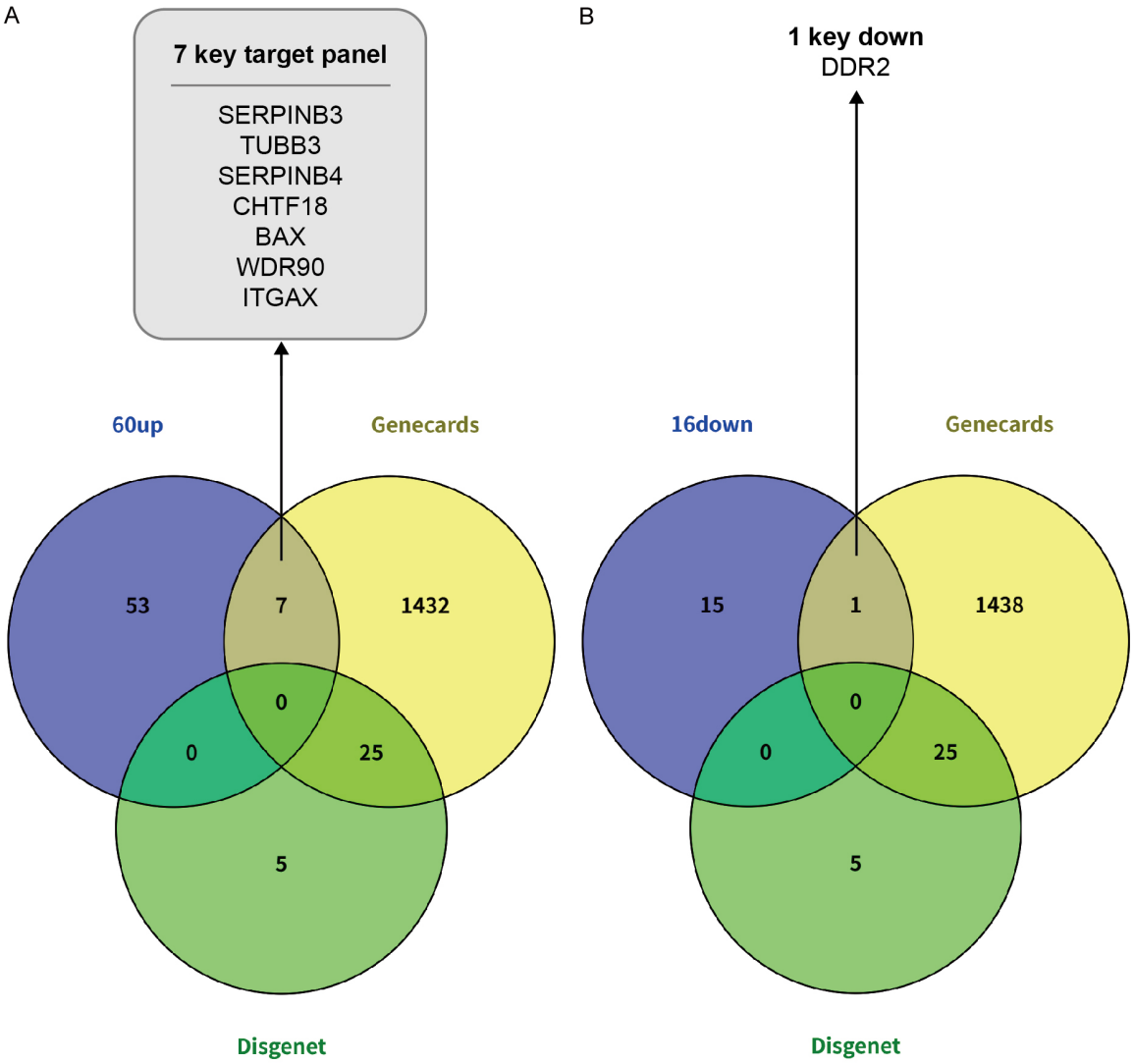

Figure 6. Key IPF marker panel. A. Among the 60 common up-regulated genes, 7 (SERPINB3, TUBB3, SERPINB4, CHTF18, BAX, WDR90, and ITGAX) were shared by the GeneCards set. B. Among the 16 down-regulated genes, one gene (DDR2) was shared by the GeneCards set.

lation of peptidase activity, serine-type endopeptidase inhibitor activity, protease binding, negative regulation of hydrolase activity, endopeptidase inhibitor activity, peptidase inhibitor activity, and endopeptidase regulator activity. The enriched KEGG pathways included Amoebiasis, Pathogenic Escherichia coli infection, Salmonella infection, Tuberculosis, Parkinson disease, Prion disease, Huntington disease, Amyotrophic lateral sclerosis, Pathways of neurodegeneration (**Figure 7C**). The enriched reactome included Hemostasis, Cytokine Signaling in Immune system, Neutrophil degranulation, and Cell Cycle (**Figure 7D**). The most related immune cells were shown to be CD4, CD8, TH17, and Treg cells (**Figure 7E**). Subsequently, we used the seven targets of this panel to propose drugs against IPF.

Candidate drugs for IPF

Inhibitory chemicals and herbs towards the seven targets presented in the aforementioned panel were screened. In the Dgidb database, four targets (of the seven genes) had corresponding chemicals: The drug numbers for TUBB3, BAX, ITGAX and SERPINB3 are shown in **Figure 8A**. It was evident that four chemicals have two targets: docetaxel anhydrous, fosbretabulin disodium, recombinant interleu-

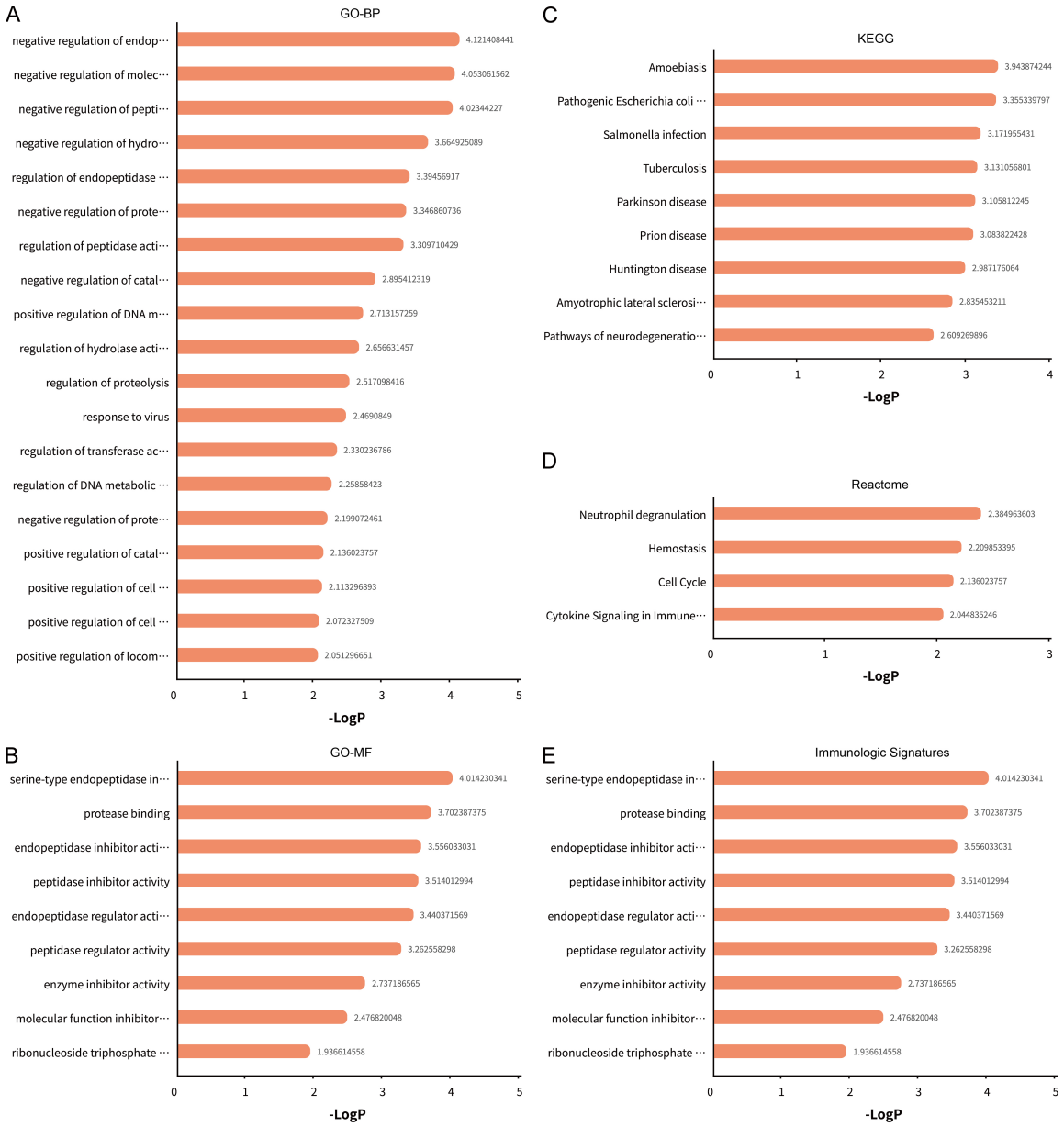

Figure 7. Enrichment analysis based on the 7 key targets in the panel. A. The enriched GO-BP pathways. B. The enriched GO-MF terms. C. The enriched KEGG pathways. D. The enriched reactome terms. E. The enriched immunologic signatures.

kin-1, and vinorelbine (**Figure 8B**). These pharmaceutical agents are all anti-tumor and possess a high level of toxicity. Furthermore, an investigation was conducted into the potential safety of various herbs by means of the Symmap database. BAX had the most corresponding herbs (420), followed by TUBB3 (12), SERPINB4 (11), ITGAX (9), SERPINB3 (8), and CHTF18 (1) (**Figure 8C**). The following herbs were identified as the most effective in terms of their targeting capabilities, andthe following ingredients were used in the formula: Lygodii Spora (spore of Japanese climbing fern), Smilacis Glabrae Rhizoma (glabrous greenbrier rhizome), Aloe (aloe), Tamaricis Cacumen (Chinese tamarisk twig), Polygoni Cuspidati Rhizoma Et Radix (rhizome of giant knotweed), Erodii Herba Geranii Herba (all grass of common heron's bill), Ilicis Cornutae Folium (Folium Ilicis Cornutae) and Vespae Nidus (**Figure 8D**).

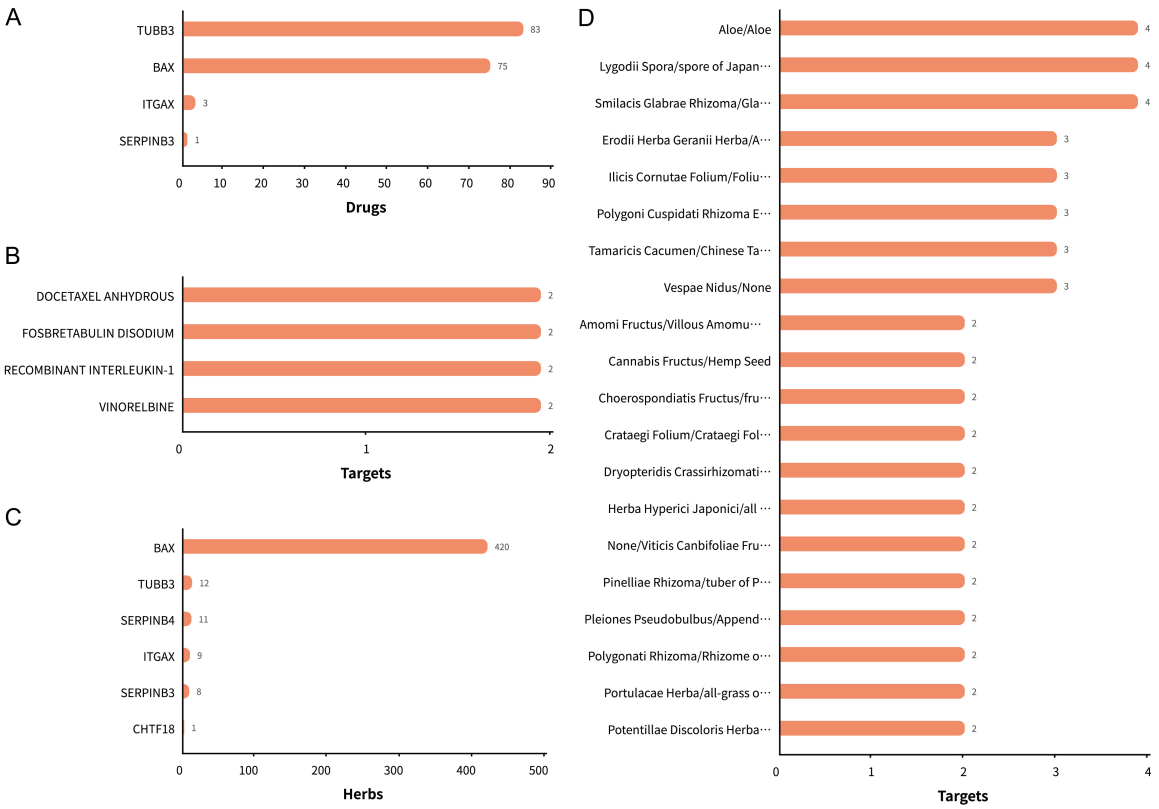

Figure 8. Candidate new drugs. A. In the Dgidb database, four targets (of the seven genes) have corresponding chemicals: TUBB3, BAX, ITGAX and SERPINB3. B. In the Dgidb database, four chemicals have two targets: docetaxel anhydrous, fosbretabulin disodium, recombinant interleukin-1, and vinorelbine. C. In the Symmap database, number of herbs of all targets. D. In the Symmap database, herbs with the most targets.

Discussion

This study was one of a few to explore key markers and potential herbal medicines for Idiopathic pulmonary fibrosis (IPF) through a bioinformatic approach. The main findings were as follows. We discovered some common DEGs in different cells in the lung microenvironment. A panel comprising seven key markers (namely SERPINB3, TUBB3, SERPINB4, CHTF18, BAX, WDR90 and ITGAX) was proposed for the identification of IPF warning signs. Several herbs may be beneficial for IPF treatment by targeting the key markers (e.g., Lygodii Spora, Smilacis Glabrae Rhizoma, and Aloe).

One of the innovations of this study was that we proposed a panel of 7 key markers. This IPF warning panel may aid in the early clinical detection of IPF, particularly for patients exhibiting suspected IPF symptoms. In addition, the panel facilitates regular disease monitoring for existing patients. However, at present, we have only proposed this as a novel panel concept and have not yet conducted clinical studies to validate its specific diagnostic and prognostic value. Although all of these key genes were derived from GeneCards and GEO differential comparisons, there are still very few studies that propose they have a direct association with IPF. Tubulin Beta 3 (TUBB3) is a class III member of the beta-tubulin protein family. This protein is primarily expressed in neurons and may be involved in neurogenesis and axon guidance and maintenance. Mutations in this gene are the cause of congenital fibrosis of the extraocular muscles type 3 [28-30]. TUBB3 is involved early in the fibrotic process [30]. In 2023, Chinese scholars developed a three-gene random forest model for diagnosing idiopathic pulmonary fibrosis based on circadian rhythm-related genes in lung tissue, and the TUBB3 expression was one of the important features [31]. SERPINB4 is a member of the serpin family of serine protease inhibitors, which is highly expressed in tumor cells and can

inactivate granzyme M. SERPINB4, along with serpin B3, can be processed into smaller fragments that aggregate to form an autoantigen in psoriasis, probably by causing chronic inflammation [32-36]. Similarly, a study in 2023 constructed an extracellular matrix-related risk model by machine learning for IPF, and SERPINB4 was one of the seven genes in the model [19]. BAX is a widely recognized pro-apoptotic gene belonging to the BCL2 protein family. A link between BAX and IPF has been repeatedly reported. In 2014, it was observed that the BAX inhibitor-1-associated V-ATPase glycosylation can enhance collagen degradation in IPF [37]. In animal models, the m6A methyltransferase ZC3H13 can improve IPF through regulating Bax expression [38]. In 2024, a novel senolytic drug (BTSA1) for IPF was proposed, that targets apoptosis of senescent myofibroblasts by activating BAX [39]. However, the following markers remain under-investigated in IPF: SERPINB3, CHTF18, WDR90, and ITGAX; and we believe that these genes could be preferred in subsequent studies. In addition, we found a core down-regulated gene DDR2 (discoidin domain receptor tyrosine kinase 2). Tyrosine kinases are involved in the regulation of tissue remodeling; and DDR2 functions as a cell surface receptor for fibrous collagen and regulates cell differentiation, extracellular matrix remodeling, migration, and proliferation [40-43]. Therefore, the decrease in DDR2 may be a homeostatic protective effect of the lung tissue, and we did not use this gene as an important marker for IPF.

Based on the key markers, we proposed some candidate chemicals/herbs for IPF. However, compounds with multiple targets, all of which are anticancer agents, have high cytotoxicity and are not clinically available. Among the therapeutic herbs, Lygodii Spora, Smilacis Glabrae Rhizoma, and Aloe target four markers. Lygodii Spora (spore of Japanese climbing Fern) drug has antibacterial and diuretic effects; it can be used for upper respiratory tract infections, mumps, and urinary tract infections [44-46]. Our study was the first to suggest that Lygodii Spora may contribute to the control of IPF. Smilacis Glabrae Rhizoma (Glabrous Greenbrier Rhizome) has clear anti-inflammatory and immunomodulatory effects. It can alleviate the oxidative stress caused by hyperuricemia by upregulating catalase expression [47]. The total glucosides of Rhizoma Smilacis Glabrae have a therapeutic role in psoriasis by regulating Th17/Treg balance [48]. Moreover, Smilacis Glabrae Rhizoma inhibits pathogen-induced upper genital tract inflammation through suppression of the NF-κB pathway [49]. Additionally, it has a protective role for kidney injury against oxidative stress-induced apoptosis by inhibiting caspase-3 activation [50]. Therefore, theoretically, this is consistent with its known bioactive function against IPF, and this medicinal value deserves subsequent validation. Aloe is protective against inflammatory injuries [51], and this effect has been proven in intestinal barrier damage [52], mastitis [53], cecal ligation and puncture-induced sepsis [54]. In addition, it has an inhibitory effect on a wide range of pathogenic microorganisms [55]. Therefore, Aloe may also have a role in the prevention and treatment of IPF. However, at present, there is not any direct evidence, subsequent preclinical and clinical studies are needed to validate this role.

Certain limitations should be noted in this study. First, our proposed warning panel (containing 7 key targets) remains at the theoretical level and has not yet been validated in clinical cases. The diagnostic and prognostic value of these targets will be clarified subsequently. Also, we have tentatively regarded the up-regulated genes as pathogenic targets, and this is indirectly supported by other previous studies. However, upregulation of expression does not necessarily imply that a gene (or protein) is a pathogenic factor promoting pulmonary fibrosis, but it could also be a protective strategy for homeostasis. Therefore, subsequent causality studies based on the regulation of gene expression are still necessary.

Conclusion

We uncovered a panel of 7 key IPF genes that may have a diagnostic and prognostic value for IPF. Some new herbal drugs may have therapeutic value against IPF, namely Lygodii Spora, Smilacis Glabrae Rhizoma, and Aloe.

Disclosure of conflict of interest

None.

Address correspondence to: Changsong Wang and Kaijiang Yu, Key Laboratory of Critical Care

Medicine of Heilongjiang Province, Department of Critical Care Medicine, The First Affiliated Hospital of Harbin Medical University, Harbin 150001, Heilongjiang, China. E-mail: changsongwang@hrbmu.edu.cn (CSW); drkaijiang@163.com (KJY)

References

[1] Tian Y, Li H, Gao Y, Liu C, Qiu T, Wu H, Cao M, Zhang Y, Ding H, Chen J and Cai H. Quantitative proteomic characterization of lung tissue in idiopathic pulmonary fibrosis. Clin Proteomics 2019; 16: 6.

[2] Benegas Urteaga M, Ramírez Ruz J and Sánchez González M. Idiopathic pulmonary fibrosis. Radiologia (Engl Ed) 2022; 64 Suppl 3: 227-239.

[3] Glass DS, Grossfeld D, Renna HA, Agarwala P, Spiegler P, DeLeon J and Reiss AB. Idiopathic pulmonary fibrosis: current and future treatment. Clin Respir J 2022; 16: 84-96.

[4] Koudstaal T and Wijsenbeek MS. Idiopathic pulmonary fibrosis. Presse Med 2023; 52: 104166.

[5] Hou Z, Dong B, Yao Q, Chen H, Shao Q, Li M, Wang J, Chen K, Zhu Z, Peng F, Wei S, Hu X, Li J, Liu M, Xu B, Zheng S, Bi N, Zheng S, Xu Q, Chen B, Wu C, Li R, Chen W, Liu X, Tian Y, Li X, Guo S, Zhao L, Zhu Y, Cai L, Li Q, Li L, Zhang H, Hu C, Wang L, Li Q, Chen B and Chen M. Pirfenidone for grade 2 and grade 3 radiation-induced lung injury: a multicentre, open-label, randomised, phase 2 trial. Lancet Oncol 2025; 26: 1552-1562.

[6] Renner A, Vertanen E, Sutinen E, Ainola M, Myllärniemi M and Hollmén M. Characterisation of patients with antifibrotic-treated non-idiopathic pulmonary fibrosis progressive pulmonary fibrosis: a retrospective real-life study. BMJ Open 2025; 15: e097246.

[7] Rich RL, Peterson A and Brown J. Diagnosis and management of acute exacerbation of interstitial lung disease in the intensive care unit. AACN Adv Crit Care 2025; 36: 374-388.

[8] Amaral AF, Colares PFB and Kairalla RA. Idiopathic pulmonary fibrosis: current diagnosis and treatment. J Bras Pneumol 2023; 49: e20230085.

[9] Bartold K, Iskierko Z, Sharma PS, Lin HY and Kutner W. Idiopathic pulmonary fibrosis (IPF): diagnostic routes using novel biomarkers. Biomed J 2024; 47: 100729.

[10] Bonella F, Spagnolo P and Ryerson C. Current and future treatment landscape for idiopathic pulmonary fibrosis. Drugs 2023; 83: 1581-1593.

[11] Mei Q, Liu Z, Zuo H, Yang Z and Qu J. Idiopathic pulmonary fibrosis: an update on pathogenesis. Front Pharmacol 2021; 12: 797292.

[12] Muri J, Durcová B, Slivka R, Vrbenská A, Makovická M, Makovický P, Škarda J, Delongová P, Kamarád V and Vecanová J. Idiopathic pulmonary fibrosis: review of current knowledge. Physiol Res 2024; 73: 487-497.

[13] Podolanczuk AJ, Thomson CC, Remy-Jardin M, Richeldi L, Martinez FJ, Kolb M and Raghu G. Idiopathic pulmonary fibrosis: state of the art for 2023. Eur Respir J 2023; 61: 2200957.

[14] Wang J, Hu K, Cai X, Yang B, He Q, Wang J and Weng Q. Targeting PI3K/AKT signaling for treatment of idiopathic pulmonary fibrosis. Acta Pharm Sin B 2022; 12: 18-32.

[15] Fan G, Liu J, Wu Z, Li C and Zhang Y. Development and validation of the prognostic model based on autophagy-associated genes in idiopathic pulmonary fibrosis. Front Immunol 2022; 13: 1049361.

[16] Fu C, Tian X, Wu S, Chu X, Cheng Y, Wu X and Yang W. Role of telomere dysfunction and immune infiltration in idiopathic pulmonary fibrosis: new insights from bioinformatics analysis. Front Genet 2024; 15: 1447296.

[17] Li Y, He Y, Chen S, Wang Q, Yang Y, Shen D, Ma J, Wen Z, Ning S and Chen H. S100A12 as biomarker of disease severity and prognosis in patients with idiopathic pulmonary fibrosis. Front Immunol 2022; 13: 810338.

[18] Lin K, Wang T, Tang Q, Chen T, Lin M, Jin J, Cao J, Zhang S, Xing Y, Qiao L and Liang Y. IL18R1-related molecules as biomarkers for asthma severity and prognostic markers for idiopathic pulmonary fibrosis. J Proteome Res 2023; 22: 3320-3331.

[19] Luo H, Yan J and Zhou X. Constructing an extracellular matrix-related prognostic model for idiopathic pulmonary fibrosis based on machine learning. BMC Pulm Med 2023; 23: 397.

[20] Xu F, Tong Y, Yang W, Cai Y, Yu M, Liu L and Meng Q. Identifying a survival-associated cell type based on multi-level transcriptome analysis in idiopathic pulmonary fibrosis. Respir Res 2024; 25: 126.

[21] Dowman LM and Holland AE. Pulmonary rehabilitation in idiopathic pulmonary fibrosis. Curr Opin Pulm Med 2024; 30: 516-522.

[22] Gao F, Pan L, Liu W, Chen J, Wang Y, Li Y, Liu Y, Hua Y, Li R, Zhang T, Zhu T, Jin F and Gao Y. Idiopathic pulmonary fibrosis microenvironment: novel mechanisms and research directions. Int Immunopharmacol 2025; 155: 114653.

[23] Noble PW, Albera C, Bradford WZ, Costabel U, Glassberg MK, Kardatzke D, King TE Jr, Lancaster L, Sahn SA, Szwarcberg J, Valeyre D and du Bois RM; CAPACITY Study Group. Pirfenidone in patients with idiopathic pulmonary fibrosis (CAPACITY): two randomised trials. Lancet 2011; 377: 1760-1769.

[24] Armanios M. Telomerase and idiopathic pulmonary fibrosis. Mutat Res 2012; 730: 52-58.
[25] Wu X, Xiao X, Fang H, He C, Wang H, Wang M, Lan P, Wang F, Du Q and Yang H. Elucidating shared biomarkers in gastroesophageal reflux disease and idiopathic pulmonary fibrosis: insights into novel therapeutic targets and the role of angelicae sinensis radix. Front Pharmacol 2024; 15: 1348708.
[26] Li XZ, Wang XL, Wang YJ, Liang QK, Li Y, Chen YW and Ming HX. Total flavonoids of Oxytropis falcata Bunge have a positive effect on idiopathic pulmonary fibrosis by inhibiting the TGF-β1/Smad signaling pathway. J Ethnopharmacol 2022; 285: 114858.
[27] Sun C, Liu H, Chi B, Han J, Koga Y, Afshar K and Liu X. Improvement of idiopathic pulmonary fibrosis through a combination of Astragalus radix and Angelica sinensis radix via mammalian target of rapamycin signaling pathway-induced autophagy in rat. J Thorac Dis 2024; 16: 1397-1411.
[28] Katarzyna L, Kyriakos O, Linda V, Ingrid S, Petra W and Karin Ö. Evaluation of tubulin β-3 and 5 hydroxy-methyl cytosine as diagnostic and prognostic markers in malignant melanoma. Ann Diagn Pathol 2024; 72: 152332.
[29] Orfanidis K, Wäster P, Lundmark K, Rosdahl I and Öllinger K. Evaluation of tubulin β-3 as a novel senescence-associated gene in melanocytic malignant transformation. Pigment Cell Melanoma Res 2017; 30: 243-254.
[30] Wawro ME, Sobierajska K, Ciszewski WM, Wagner W, Frontczak M, Wieczorek K and Niewiarowska J. Tubulin beta 3 and 4 are involved in the generation of early fibrotic stages. Cell Signal 2017; 38: 26-38.
[31] He J, Hu J and Liu H. A three-gene random forest model for diagnosing idiopathic pulmonary fibrosis based on circadian rhythm-related genes in lung tissue. Expert Rev Respir Med 2023; 17: 1307-1320.
[32] de Koning PJ, Kummer JA, de Poot SA, Quadir R, Broekhuizen R, McGettrick AF, Higgins WJ, Devreese B, Worrall DM and Bovenschen N. Intracellular serine protease inhibitor SERPINB4 inhibits granzyme M-induced cell death. PLoS One 2011; 6: e22645.
[33] Izuhara K, Yamaguchi Y, Ohta S, Nunomura S, Nanri Y, Azuma Y, Nomura N, Noguchi Y and Aihara M. Squamous cell carcinoma antigen 2 (SCCA2, SERPINB4): an emerging biomarker for skin inflammatory diseases. Int J Mol Sci 2018; 19: 1102.
[34] Ouyang XM, Lin JH, Lin Y, Zhao XL, Huo YN, Liang LY, Huang YD, Xie GJ, Mi P, Ye ZY and Guleng B. The SERPINB4 gene mutation identified in twin patients with Crohn's disease impaires the intestinal epithelial cell functions. Sci Rep 2025; 15: 2638.
[35] Riaz N, Havel JJ, Kendall SM, Makarov V, Walsh LA, Desrichard A, Weinhold N and Chan TA. Recurrent SERPINB3 and SERPINB4 mutations in patients who respond to anti-CTLA4 immunotherapy. Nat Genet 2016; 48: 1327-1329.
[36] Wang X, Wang J and Tian L. Icariin ameliorates TNF-α/IFN-γ-induced oxidative stress, inflammatory response and apoptosis of human immortalized epidermal cells through the WTAP/SERPINB4 axis. Arch Dermatol Res 2024; 316: 557.
[37] Lee MR, Lee GH, Lee HY, Kim DS, Chung MJ, Lee YC, Kim HR and Chae HJ. BAX inhibitor-1-associated V-ATPase glycosylation enhances collagen degradation in pulmonary fibrosis. Cell Death Dis 2014; 5: e1113.
[38] Guan J, Yin L, Huang Q, Chen J, Liu H and Li J. m(6)A methyltransferase ZC3H13 improves pulmonary fibrosis in mice through regulating Bax expression. Exp Cell Res 2024; 442: 114255.
[39] Shen M, Fu J, Zhang Y, Chang Y, Li X, Cheng H, Qiu Y, Shao M, Han Y, Zhou Y and Luo Z. A novel senolytic drug for pulmonary fibrosis: BTSA1 targets apoptosis of senescent myofibroblasts by activating BAX. Aging Cell 2024; 23: e14229.
[40] Chen Y, Yu Y, Fan Y, Lu W, Wei Y, Wu J, Ruan B, Wan Z, Zhao Y, Xie K, Jie W and Zheng S. Biomimetic nanoplatforms for combined DDR2 inhibition and photothermal therapy in dense breast cancer treatment. Biomaterials 2025; 324: 123497.
[41] Huang G, Cong Z, Zhao Y, Zhu T, Yuan R, Li Z, Wang X and Qi J. DDR2-mediated autophagy inhibition contributes to angiotensin II-induced adventitial remodeling. Clin Transl Med 2025; 15: e70361.
[42] Vessella T, Rozen EJ, Shohet J, Wen Q and Zhou HS. Investigation of cell mechanics and migration on DDR2-expressing neuroblastoma cell line. Life (Basel) 2024; 14: 1260.
[43] Vessella T, Xiang S, Xiao C, Stilwell M, Fok J, Shohet J, Rozen E, Zhou HS and Wen Q. DDR2 signaling and mechanosensing orchestrate neuroblastoma cell fate through different transcriptome mechanisms. FEBS Open Bio 2024; 14: 867-882.
[44] Elliott ML and Rayamajhi MB. First report of bipolaris sacchari causing leaf spot on lygodium japonicum and L. microphyllum in Florida. Plant Dis 2005; 89: 1244.
[45] Ireland KB, Noor NAHM, Aitken EAB, Schmidt S and Volin JC. First report of glomerella cingulata (colletotrichum gloeosporioides) causing anthracnose and tip dieback of lygodium mi-

crophyllum and L. japonicum in Australia. Plant Dis 2008; 92: 1369.
[46] Rayamajhi MB, Pemberton RW, Van TK and Pratt PD. First report of infection of lygodium microphyllum by Puccinia lygodii, a potential biocontrol agent of an invasive fern in Florida. Plant Dis 2005; 89: 110.
[47] Hong Q, Yu S, Mei Y, Lv Y, Chen D, Wang Y, Geng W, Wu D, Cai G and Chen X. Smilacis Glabrae Rhizoma reduces oxidative stress caused by hyperuricemia via upregulation of catalase. Cell Physiol Biochem 2014; 34: 1675-1685.
[48] Tang Y, Yu J, Zhao W, Liu J, Peng H, Zhang H, Jiang Z, Yu Q and Zhang L. Total glucosides of Rhizoma Smilacis Glabrae: a therapeutic approach for psoriasis by regulating Th17/Treg balance. Chin J Nat Med 2023; 21: 589-598.
[49] Zou W, Zhou H, Hu J, Zhang L, Tang Q, Wen X, Xiao Z and Wang W. Rhizoma Smilacis Glabrae inhibits pathogen-induced upper genital tract inflammation in rats through suppression of NF-κB pathway. J Ethnopharmacol 2017; 202: 103-113.
[50] Liu C, Kang Y, Zhou X, Yang Z, Gu J and Han C. Rhizoma smilacis glabrae protects rats with gentamicin-induced kidney injury from oxidative stress-induced apoptosis by inhibiting caspase-3 activation. J Ethnopharmacol 2017; 198: 122-130.
[51] Khan A, Ullah S, Ramzan F, Rehman AU, Rehman S, Kholik K, Sukri A, Munawaroh M and Sucipto TH. Combined effects of Aloe vera leaf extract and vitamin E on wound healing in Oryctolagus cuniculus (Rabbits). Open Vet J 2025; 15: 700-708.
[52] Zhang D, Chen K, Yu Y, Feng R, Cui SW, Zhou X and Nie S. Polysaccharide from Aloe vera gel improves intestinal stem cells dysfunction to alleviate intestinal barrier damage via 5-HT. Food Res Int 2025; 214: 116675.
[53] de Aguiar SC, Cottica SM, Dos Santos ST, da Fonseca JM, da Silva Leite L and da Silva ML. Antioxidant activity, phenolic acid, and flavonoid composition of an antiseptic ointment based on aloe and green propolis and its potential for preventing mastitis in dairy cows. Vet Sci 2025; 12: 248.
[54] Su J, Deng X, Hu S, Lin X, Xie L, Ye H, Lin C, Zhou F, Wu S and Zheng L. Aloe-emodin plus TIENAM ameliorate cecal ligation and puncture-induced sepsis in mice by attenuating inflammation and modulating microbiota. Front Microbiol 2024; 15: 1491169.
[55] Farajnabi R, Hojjati Bonab Z and Shagayegh N. The in Vitro antimicrobial effect of ethanolic extracts of aloe vera on the H-pylori and inhibitory effect of secondary metabolites of aloe vera by molecular docking. Adv Biomed Res 2025; 14: 10.