

## Original articles

## Potential function exploration of lncRNAs in idiopathic pulmonary fibrosis: insights from whole transcriptome sequencing data analysis

Xinran Ma<sup>a</sup>, Dong Chen<sup>b</sup>, Chenghai Li<sup>c</sup>, Wenjuan Wu<sup>a,d,\*</sup><sup>a</sup> People's Hospital of Zhengzhou University, Zhengzhou, China<sup>b</sup> Wuhan Ruixing Biotechnology Co. Ltd., Wuhan, China<sup>c</sup> Stem Cell Program at Henan Provincial People's Hospital and People's Hospital of Zhengzhou University, China<sup>d</sup> Department of Geriatric Medicine, Henan Provincial People's Hospital and Provincial Hospital Clinical Medical College of Zhengzhou University, Zhengzhou, China

## ARTICLE INFO

## Keywords:

Idiopathic pulmonary fibrosis

RNA-seq

LncRNAs

Co-expression

Inflammatory response

## ABSTRACT

**Background:** Research has shown that long noncoding RNAs (lncRNAs) play a role in Idiopathic Pulmonary Fibrosis (IPF), but their specific functions and patterns of expression are still unclear.**Method:** A diagnostic study was conducted by utilizing analysis techniques. RNA sequencing (RNA-seq) data from 12 IPF patients and 5 controls was used to study lncRNA functions in IPF. The authors identified Differentially Expressed lncRNAs (DELncRNAs) and explored co-expression networks in a transient manner, as well as using Weighted Gene Co-expression Network Analysis (WGCNA) to identify modules associated with IPF.**Results:** The study found 541 differentially expressed lncRNAs in IPF, with 201 up-regulated and 340 down-regulated. DELncRNAs, especially the up-regulated, were significantly correlated with DEMRNAs at their expression levels. DEMRNAs showed extracellular matrix-related biological functions in addition to increased lncRNAs. WGCNA results demonstrated that Module Eigengene green (MEgreen) and MEred modules indicated the highest negative correlation significance with IPF phenotype, and eigengene patterns of both modules were downregulated in IPF samples. The authors identified six significant lncRNAs in these two modules, including FAM13A-AS1, RP11-180C16.1, MYO16-AS1, AC007278.2, BACH1-IT2, and RP11-153M7.5, and their co-expressed DE mRNAs were enriched in inflammatory response pathways. The authors used single-cell RNA sequencing (scRNA-seq) data to investigate dysregulated lncRNAs and their co-expressed mRNAs. The authors found that five DEMRNAs that were co-expressed with DELncRNAs exhibited dysregulated expression patterns in multiple cell types of the IPF samples.**Conclusion:** LncRNAs are functionally active and potentially involved in the inflammatory response in pathological processes of IPF. It is also important to consider some specific lncRNAs as potential diagnostic biomarkers or therapeutic targets for preclinical and clinical studies with IPF in the future.

## Introduction

Idiopathic Pulmonary Fibrosis (IPF) is a chronic lung disease characterized by progressive scarring and causes epithelial cells to die, fibroblasts to proliferate, excess Extracellular Matrix (ECM) to accumulate, the destruction of pulmonary vessels, and loss of alveolar gas exchange units.<sup>1,2</sup> IPF is the most common disease in elderly individuals, with a prevalence of 494 cases per 100,000 in adults over 65 in 2011, which was twice as much as that in the prevalence recorded 10-years earlier.<sup>3</sup> Patients with IPF are affected both physically and emotionally, and patient numbers are rising in recent years.<sup>4,5</sup> Although many pathobiological concepts are emerging, including senescence,

oxidative stress and cellular plasticity, the exact ways in which IPF works are not completely known.<sup>6</sup>

Long Noncoding RNA (lncRNA) is a diverse type of RNA over 200 nucleotides in length and does not code for proteins. LncRNAs do not code for proteins but can interact with different cellular molecules, including mRNAs, to control cell fate.<sup>7,8</sup> LncRNAs play a role in the development of different diseases, including tumors, infection, and other extensive diseases, through regulating gene expression at different stages, including transcription and post-transcription.<sup>7</sup> Interestingly, lncRNAs are supposed to be critical regulators and prognostic markers in IPF. For example, Fibrosis is induced by lnc949 mediated by FKBP5.<sup>9</sup> Targeting lncRNA H19 can enhance IPF treatment by acting as a sponge

\* Corresponding author.

E-mail address: [18503862016@163.com](mailto:18503862016@163.com) (W. Wu).<https://doi.org/10.1016/j.clinsp.2025.100732>

Received 11 April 2024; Received in revised form 23 June 2025; Accepted 13 July 2025

1807-5932/© 2025 HCFMUSP. Published by Elsevier España, S.L.U. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

for miR-196a.<sup>10</sup> lncRNA ZEB1-AS1 promotes pulmonary fibrosis by inducing epithelial-mesenchymal transition through ZEB1.<sup>11</sup> Age-related IPF is predicted by upregulation of lncRNA AP003419.16.<sup>12</sup> By targeting miR-138, PFAR promotes lung fibrogenesis and activation of the YAP1-Twist axis.<sup>13</sup> IPF progression is prevented by lncRNA CTD-2528L19.6 by alleviating fibroblast activation.<sup>14</sup> lncRNA Hoxaas3 promotes lung fibrosis via miR-450b-5p-mediated regulation of Runx1.<sup>15</sup> Alveolar epithelial cells are induced to undergo EMT by ChRF through inhibiting miR-146a and upregulating L1CAM.<sup>16</sup> IPF is suppressed when microRNA-369-3p/TRIM2 is knocked down by DLEU2 knockdown.<sup>17</sup> CTGF promotes lung fibrosis by competing with lncRNA PFAL for miR-18a.<sup>18</sup> Inhibiting lung fibroblast proliferation through reduced levels of  $\beta$ -Catenin is achieved by lncRNA FENDRR.<sup>19</sup> As mentioned above, the main process has been observed to understand the basic mechanisms of IPF and find targets for treatment. Therefore, lncRNAs have drawn attention for their roles in IPF and the authors believe that there are still a large number of unidentified lncRNAs to be implicated in the progression of IPF.

In this study, the authors suggest that certain unidentified lncRNAs may play a role in the development of IPF. The authors analyzed high-throughput RNA-seq and single-cell RNA-seq (scRNA-seq) datasets from the GEO database, including GSE138283<sup>20</sup> and GSE122960,<sup>21</sup> respectively, to compare lung tissue samples from IPF patients and healthy controls. Using the ribosome elimination method to extract RNAs, the authors constructed a transcriptome library and identified lncRNAs that may be involved in inflammation by interacting with differentially expressed mRNAs. This study is one of the first studies linking several novel lncRNAs to IPF inflammatory pathways using both bulk and single-cell transcriptomics and may provide new insights into potential therapeutic targets for DE lncRNAs in IPF in future studies.

## Materials and methods

### Obtaining and processing of the public transcriptome sequencing data

By using the NCBI SRA Tool fastq-dump and FASTX-Toolkit, Run files of the public dataset GSE138283 from the Sequence Read Archive (SRA) were downloaded and converted to fastq format after pruning low-quality bases (v.0.0.13; [http://hannonlab.cshl.edu/fastx\\_toolkit/](http://hannonlab.cshl.edu/fastx_toolkit/)). FastQC was performed for a clean read evaluation (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc>). Due to the absence of metadata, clinical variations were not controlled in this study. This study adheres to the relevant guidelines for reporting research. The authors followed the STROBE Statement to ensure the rigor and transparency of the present study design, analysis, and reporting.

Meanwhile, a single cell RNA-seq (scRNA-seq) dataset PRJNA507000<sup>21</sup> of lung transplantation samples from three elderly controls and three young controls, as well as four IPF patients, was determined by Unique Molecular Identifier (UMI) (<https://www.ncbi.nlm.nih.gov/bioproject/PRJNA507000>). The UMI count matrix was transformed into a Seurat matrix with the Seurat R package (version 4.0.4).<sup>22</sup> Low-quality cells were excluded from the study if their UMI count exceeded 1000, or their detectable gene count exceeded 500, or the mitochondrial-derived UMI count exceeded 15 %.

### Data pre-processing and quality control for the scRNA-seq analysis

To generate possible anchors, the top 2000 variable genes were utilized in Seurat's Find Integration Anchors feature, followed by data integration using the Integrate Data function. The authors used Seurat's Elbow plot function to reduce the dimensionality of scRNA-Seq datasets by performing PCA on them. In Seurat's Find Clusters function, the default resolution was set in order to identify the main clusters of cells (res = 0.4). Using tSNE or UMAP plots, the authors visualized 25 major cell types. In the Seurat package, the "Find Markers" function was used to identify genes for each cluster of cells (v4.0.4), and the ScType tool<sup>23</sup>

was used to annotate previously published pulmonary fibrosis markers.<sup>24</sup>

### Analyzing differentially expressed genes (DEGs) based on read alignment

The alignment of clean data to GRCh38 was permitted with four mismatches.<sup>25</sup> Gene read numbers and FPKM were calculated based on uniquely mapped reads. Gene expression levels were assessed by using FPKM. In order to identify DEGs, the authors detected RNA-seq data using DESeq2,<sup>26</sup> a tool specifically developed for analyzing gene expression levels.

The authors assessed gene expression differences by examining fold change ( $FC \geq 2$  or  $\leq 0.5$ ) and false discovery rate ( $FDR \leq 0.05$ ) to identify genes that were differentially expressed.

### Predictions and identifications of lncRNA directions

The present analysis of lncRNA expression patterns was driven by Cufflinks software, which utilizes a pipeline for identifying lncRNAs similar to previous studies.

### Analysis of co-expression networks and weighted gene co-expression networks (WGCNA)

Gene expression patterns were recognized by the WGCNA package (version 1.73)[27] by clustering genes with mimicking expression patterns with default parameters. Module preservation statistics and sensitivity analyses were run. Clustered modules are represented by eigengenes, which are genes whose expression patterns are similar. Automatic network construction with an optimal soft thresholding power of  $\beta = 8$  and a minimum of 30 genes in one module was used. According to Pearson's Correlation Coefficients (PCCs), the lncRNAs and their host mRNAs were classified into three classes: correlations, both positively and negatively, as well as non-correlations.

### Analyzing functional enrichment

By using GO terms and KEGG pathways from the KOBAS 2.0 server, the authors performed a systematic functional classification of DEGs.<sup>28</sup> Using Hypergeometric Testing and Benjamini-Hochberg FDR controls, each term was enriched. Functional enrichment was analyzed using Reactome's (<http://reactome.org>) pathway profiling.

### Other statistical analysis

The authors used factoextra in R to visualize sample clustering with PCA's first two components (<https://cloud.r-project.org/package=factoextra>). The R script heatmap visualized normalized NGS data and genomic annotations(<https://cran.r-project.org/web/packages/pheatmap/index.html>). By using the Student's *t*-test, two groups were compared.

## Results

### Genome-wide profiling of the lncRNA expression in IPF

To study lncRNA expression patterns and functions in IPF, a systematic analysis of RNA-seq dataset (GSE138283) between the IPF and the controls<sup>20</sup> was performed in this study. The lung biopsy samples were obtained from 12 IPF patients and 5 controls. By depleting ribosomal RNAs from biopsy samples, the RNA-seq dataset was prepared to analyze lncRNA expression and functions. After acquiring high-quality data and aligning them to the human genome, the lncRNA prediction procedure and following differential expression and co-expression analysis between mRNAs and lncRNAs were analyzed according to the methods previously published.<sup>29</sup> The overall design of this manuscript

was presented in Fig. 1A. The Pearson correlation coefficients among different samples were calculated and presented in the diagonal heatmap (Supplemental Fig. 1A). It was shown that two control samples were distributed among IPF samples and the other three control samples were clustered together, indicating the individual variation of tissue samples. PCA was used to analyze gene expression levels in both the IPF and control groups to study sample clustering. The confidence ellipses of the two groups showed no overlap (Fig. 1B). With the criteria of  $|\log_2 \text{fold change}| > 1$  and False Discovery Rate (FDR)  $< 0.05$ , 1743 genes were differentially expressed in the IPF group, including 1126 up-regulated and 617 down-regulated genes (Fig. 1C). Among 1743 DEGs, 1109 were protein-coding genes; 337 were lncRNAs and 297 were other types (Fig. 1D). The protein-coding genes and lncRNAs account for 83.0 % of the total DEGs, indicating their important roles in IPF. The functions of the DE mRNAs were analyzed using GO enrichment analysis. The up-regulated mRNAs were mainly enriched in extracellular matrix organization and tumor necrosis factor-mediated signaling pathways. The decreased mRNAs were primarily involved in immune response to fungus, inflammation regulation, and cell signaling pathways (Supplemental Fig. 1B–C). These studies showed that extracellular matrix and inflammation-related signal pathways play important roles in the pathogenesis of IPF. In the following part of this research, the authors will focus on the expression patterns and functions of lncRNAs.

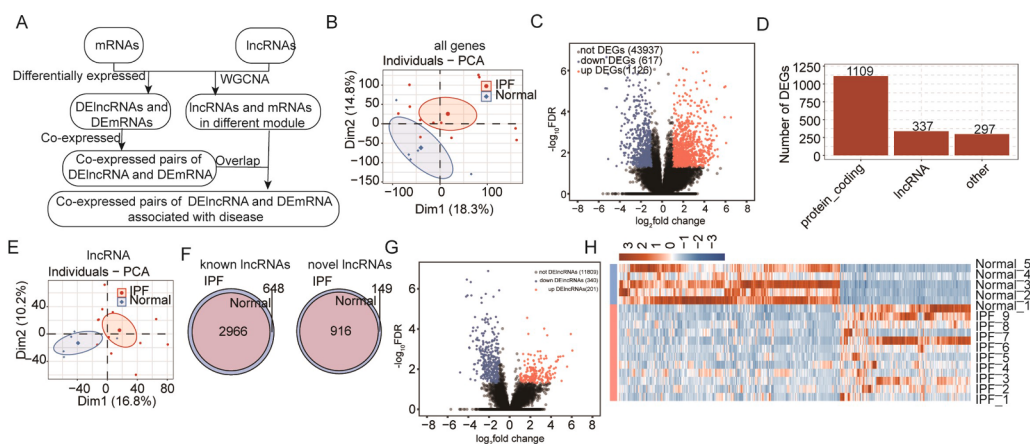
Further exploration of lncRNAs in IPF revealed separation of detected lncRNAs in both IPF and control groups based on PCA results (Fig. 1E). As known, 4679 lncRNAs were detected in total, including 3614 known and 1065 novel lncRNAs. All known lncRNAs in the IPF group were included in the control group, and 648 novel lncRNAs were specific to the control group. A similar result was observed for novel lncRNAs, with 916 common lncRNAs in both groups, and 149 specific lncRNAs in the control group (Fig. 1F). In IPF, 541 common lncRNAs showed differential expression, with 201 up-regulated and 340 down-regulated (Fig. 1G). The detailed number of the DE known and novel lncRNAs can be acquired in Supplemental Figure 1D. The DE lncRNAs heatmap displayed strong consistency across the samples (Fig. 1H). The results indicate that lncRNAs are actively expressed and differ between IPF and control samples, suggesting their potential importance in IPF.

### DE lncRNAs correlated with DE mRNAs in a trans manner

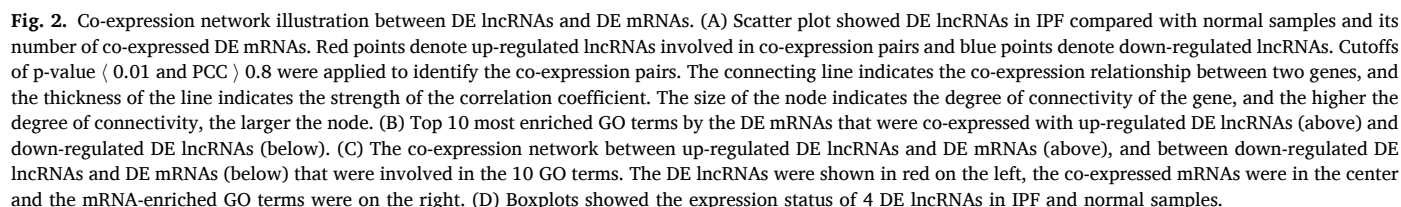
A co-expression network analysis was used to study the involvement of lncRNAs in IPF by analyzing the relationship between differentially expressed lncRNAs and mRNAs. Genes with strong correlations may have similar functions and regulate each other in a trans manner. The criteria for the co-expression relationship were  $|\text{Pearson's correlation coefficients}| > 0.8$  and  $p\text{-value} < 0.01$ . Under this criteria, 527 DE lncRNAs were correlated with 1043 DE mRNAs, with 41,858 DE lncRNA-mRNA pairs, suggesting the close relationship between the two RNA types and the close relationship between DE lncRNAs and mRNAs. Compared to the down-regulated lncRNAs, the up-regulated DE lncRNAs were more highly correlated with DE mRNAs, with a correlating number of more than 120 in some cases (Fig. 2A). Some DE lncRNAs, including LINC01207, RP11-28A20.1 and LINC00939, were significantly up-regulated ( $-\log_{10}(p\text{-value}) > 2.5$ ) and correlated with abundant DE mRNAs ( $> 150$  DE mRNAs) (Fig. 2A). GO analysis showed that DE mRNAs co-expressed with up-regulated lncRNAs were enriched in extracellular matrix organization, tumor necrosis factor signaling, and extracellular matrix disassembly (Fig. 2B), and cytokine-cytokine receptor interaction by KEGG enrichment analysis (Supplemental Fig. 2A). DE mRNAs co-expressed with down-regulated lncRNAs were enriched in chronic inflammation response, inflammatory response regulation, TNF-mediated signaling, wound response, extracellular matrix organization, and NF-kappa B activity regulation (Fig. 2B). They were also enriched in cytokine-cytokine receptor and ECM-receptor interactions (Supplemental Fig. 2B). These signal pathways which are related to extracellular matrix and inflammation have been reported to be important pathogenesis in IPF. The up- and down-regulated lncRNAs-DE mRNAs regulatory networks were illustrated in Fig. 2C. Several DE lncRNAs were significantly changed in IPF and co-expressed with abundant DE mRNAs, including RP11-275114.4, CTB-51J22.1, RP11-412H8.2 and XLOC.302,021 (Fig. 2D). Taken together, the result showed that the up- and down-regulated lncRNAs might function in the extracellular matrix and inflammation response-related pathways in IPF.

### lncRNA-mRNA co-expression network analysis

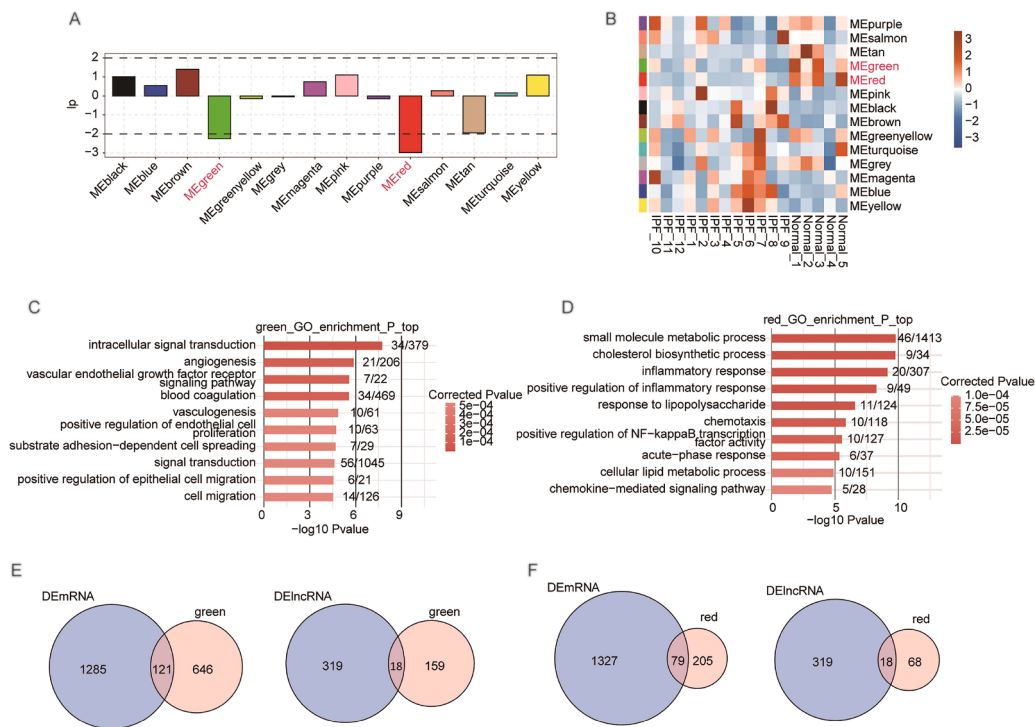
WGCNA was utilized to analyze the co-expression of lncRNAs and



**Fig. 1.** Abundant lncRNAs detected in Idiopathic Pulmonary Fibrosis (IPF) samples and the control samples. (A) The flow chart showed the construction process of the mRNA and lncRNA network. (B) Principal Component Analysis (PCA) of 12 IPF and 5 normal samples based on normalized expression levels of all known genes. (C) Volcano plot presented Differentially Expressed Genes (DEGs) in IPF samples compared with normal samples in each group as experimental replicates. Red indicates up-regulated genes ( $FC \geq 2$  and  $FDR \leq 0.05$ ) and blue indicates down-regulated genes ( $FC \leq 0.5$  and  $FDR \leq 0.05$ ). (D) The bar chart showed the number of different types of genes (including protein-coding genes, annotated lncRNAs, and others). (E) Principal Component Analysis (PCA) of 12 IPF and 5 normal samples based on normalized lncRNAs (including predicted novel lncRNAs) expression levels. (F) Venn diagram of detected annotated lncRNAs (left) and predicted novel lncRNAs (right) in IPF and normal samples. At least two samples with  $FPKM \geq 0.2$  were considered to be detected in the group. (G) Volcano plot presented differentially expressed lncRNAs in IPF compared with normal samples in each group as experimental replicates. Red indicates up-regulated genes ( $FC \geq 2$  and  $FDR \leq 0.05$ ) and blue indicates down-regulated genes ( $FC \leq 0.5$  and  $FDR \leq 0.05$ ). H. Expression heatmap of all significant DE lncRNAs in IPF and normal samples.



Eigengene patterns of MEgreen and MERed were also consistently downregulated in IPF samples (Fig. 3B). The potential functions of genes from the two modules were further explored. The MEgreen module genes showed enrichment in the VEGF receptor signaling pathway according to GO analysis (Fig. 3C). The MERed module genes were enriched in inflammation-related pathways, such as positive regulation of NF- $\kappa$ B transcription factor activity and chemokine-mediated



**Fig. 3.** The co-expression pattern of differentially expressed genes by WGCNA. (A) Bar plot presented the significance of the correlation in the IPF samples among each pattern. Red indicates the two modules with the highest negative correlation significance. (B) The eigengene pattern of each gene co-expression module was presented by a heatmap. (C) Top 10 most enriched GO BP terms by mRNA genes from green module. (D) Top 10 most enriched GO BP terms by mRNA genes from red module. (E) Venn diagram showed the overlapped genes between DE mRNAs and mRNAs from green module (left), and DE lncRNAs and lncRNAs from green module (right). (F) Venn diagram showed the overlapped genes between DE mRNAs and mRNAs from red module (left), and DE lncRNAs and lncRNAs from red module (right).

signaling pathway (Fig. 3D). The pathways in the MERed module were related to inflammation, an important mechanism in IPF.

There were 121 DE mRNAs and 18 DE lncRNAs in the MEgreen module (Fig. 3E), 79 DE mRNAs and 18 DE lncRNAs in the MERed module (Fig. 3F). Of the 177 lncRNAs in the MEgreen module, 66 were annotated (known) and 111 were newly predicted (novel) (Fig. 4A). In the MERed module, 40 were known and 46 were novel (Fig. 4A). For DE lncRNAs, 18 were known and 69 were novel in the MEgreen module, and 18 were known and 36 were novel in the MERed module (Fig. 4B). Supplemental Figure 4 shows expression levels of known DE lncRNAs and mRNAs in two modules. GO terms were used to analyze potential functions of DE lncRNAs, revealing enrichment in inflammatory response and NF-kappa B transcription factor activity in the MEgreen module (Fig. 4C). Then, the hub known as co-expressed lncRNAs and mRNAs was extracted, and the regulatory network was constructed. The lncRNAs and mRNAs in the network included FAM13A-AS1, RP11-180C16.1, MYO16-AS1, and IL1RL1, S100A12, S100A8 (Fig. 4D). The co-expressed DE mRNAs in the MERed module were enriched in pathways related to inflammation, NF-kB activity, and immune response (Fig. 4E). The hub known co-expressed lncRNAs and mRNAs, including AC007278.2, BACH1-IT2, RP11-153M7.5, and CTH, TLR2, CCL26 (Fig. 4F), were also extracted. Their regulatory network was also constructed. Using WGCNA, IPF-related modules and the specific lncRNAs and mRNAs were detected, indicating that DE lncRNAs might function in inflammation-related signal pathways in IPF.

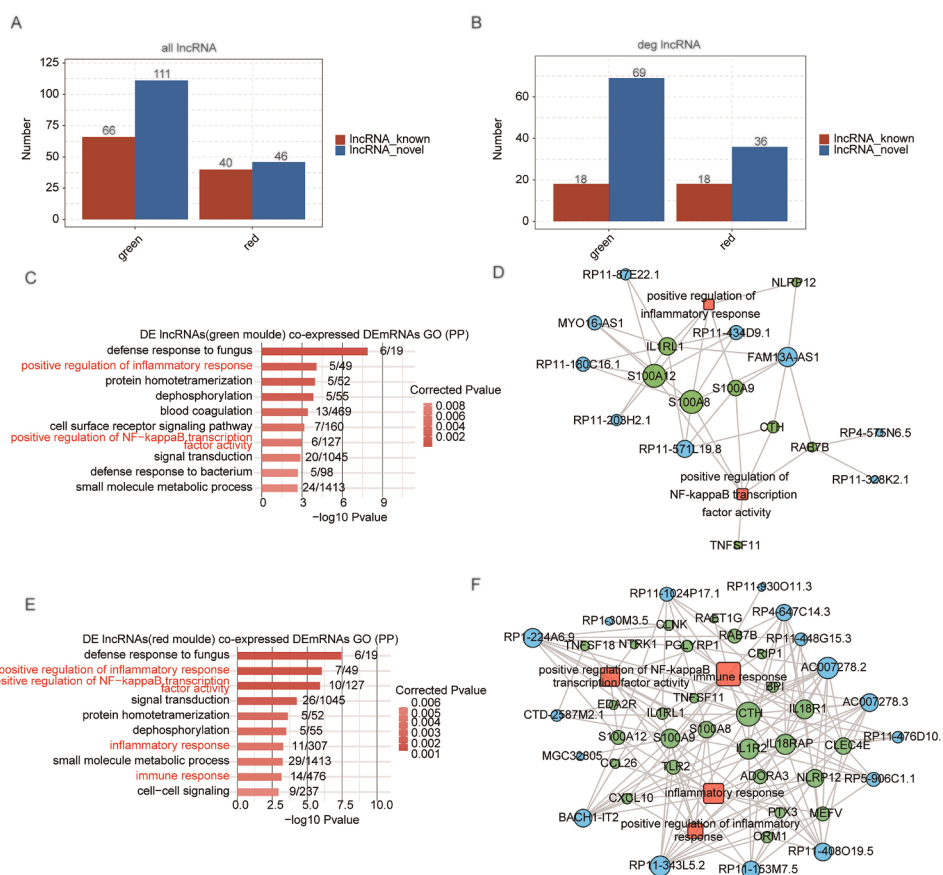
To assess the level of co-expression between different lung histiocytes of DEmRNA and DELncRNA, the authors conducted an analysis of the single-cell sequencing database. No co-expressed DELncRNAs were detected. Interestingly, the present findings indicate that the expression of DEmRNA S100A12 and S100A9 in the green module related to interstitial pulmonary fibrosis was abnormal, S100A9 expression was reduced in T/NKT cells and B-cells (Fig. 5A), while S100A12 expression

was increased in monocytes (Fig. 5B). Special focus on enhancing the positive regulation of inflammation and NF-kappa B pathway activity. Furthermore, the authors found unusual levels of certain genes in the red module associated with idiopathic pulmonary fibrosis, specifically S100A12, S100A9, CLEC4E, CRIP1, and IL1R2. These genes are linked to inflammatory and immune responses, CLEC4E expression was reduced in macrophages (Fig. 5C), IL1R2 expression was reduced in dendritic cells and monocytes (Fig. 5D), and CRIP1 expression was elevated in rod-shaped cells and fibroblasts (Fig. 5E).

**Discussion**

As the authors know, many factors, including injury or pathogen infection, increased resistance of myofibroblasts to apoptosis, and extracellular matrix deposition, contribute to the pathogenesis and development of IPF, which is enhanced by aging and genetic alterations.<sup>30</sup> An in-depth examination of the expression patterns and potential regulatory roles of dysregulated lncRNAs in IPF was conducted in this study, suggesting that lncRNAs may be involved in ECM and immune/inflammatory pathways by co-expressing with related genes and, consequently, the dysregulation of lncRNAs can be closely related to IPF development. Importantly, the results of PCA showed that DE lncRNAs can be well separated in the IPF and control specimens, thus suggesting that lncRNAs could also be potentially diagnostic biomarkers in IPF.

The present hypothesis that lncRNAs contribute to the pathology in IPF is well supported by the previous multiple studies. There is increasing evidence of lncRNAs' role in pulmonary fibrosis, such as lnc949,<sup>31</sup> ZEB1-AS1,<sup>11</sup> and AP003419.16.<sup>12</sup> Interestingly, one previous study reported that 272 lncRNAs, including the 42 newly identified lncRNAs, were differentially expressed in bleomycin-induced pulmonary fibrosis and 59 lncRNAs and 131 mRNAs were co-expressed in the



**Fig. 4.** Differentially expressed lncRNAs in IPF-related modules and co-expressed DE mRNAs are enriched in immune and inflammation-related pathways. (A) Bar chart showed the number of all annotated (known) and newly predicted (novel) lncRNAs in IPF-related modules. (B) The bar chart showed the number of DE annotated (known) and newly predicted (novel) lncRNAs in IPF-related modules. (C) Top 10 enriched GO BP terms of mRNAs that were co-expressed with DE lncRNAs in the IPF-related green module. (D) The co-expression network between known DE lncRNAs (blue circle) in red modules and co-expressed DE mRNAs (green circle) in immune and inflammation-related pathways (red square). (E) Top 10 enriched GO BP terms of mRNAs co-expressed with DE lncRNAs in the IPF-related red module. (F) The co-expression network between known DE lncRNAs (blue circle) in green modules and co-expressed DE mRNAs (green circle) in immune and inflammation-related pathways (red square).

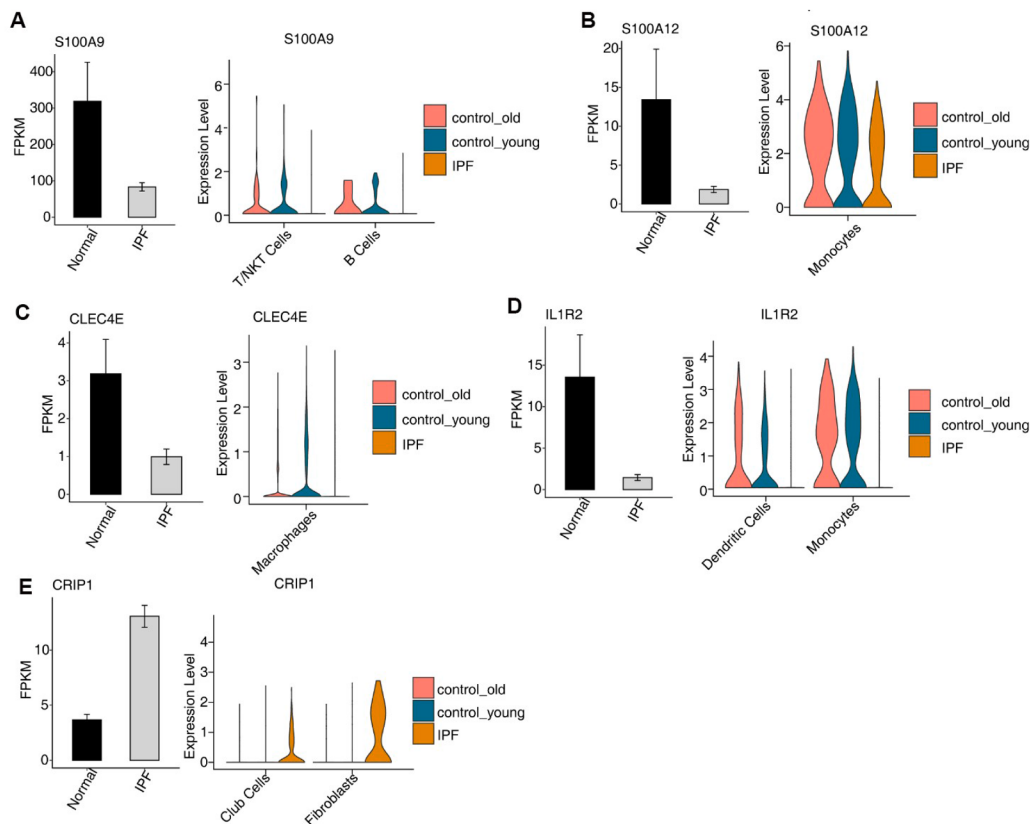
lung fibrosis.<sup>9</sup> In the present study, the authors identified 541 differentially expressed lncRNAs in IPF, including 201 up-regulated and 340 down-regulated lncRNAs. These results may provide new understanding of how lncRNAs contribute to pulmonary fibrosis development and treatment.

Recent findings indicate that lncRNAs can influence gene expression through either nearby (cis-acting) or distant (trans-acting) mechanisms.<sup>32-34</sup> IPF-associated mRNAs and lncRNAs were predicted by analyzing co-expression networks. Finally, 527 DE lncRNAs were identified and co-expressed with 1043 DE mRNAs using their expression values from three groups. Co-expressed DE mRNAs of up-regulated DE lncRNAs indicated that they were enriched in pathways associated with extracellular matrix remodeling, which is related to the pathogenetic involvement in pulmonary fibrosis.<sup>35</sup> DE mRNAs that co-expressed with down-regulated lncRNAs can primarily initiate multiple inflammatory responses in this study, including chronic inflammation response, positive regulation of chronic inflammation response, and NF- $\kappa$ B transcription factor activity, which are among the factors affected by chronic inflammation.

Additionally, four potential lncRNAs were newly identified with significant expression changes in IPF, including RP11-275114.4, CTB-51J22.1, RP11-412H8.2 and XLOC\_302,021. Noticeably, these lncRNAs were co-expressed with abundant DE mRNAs. Their expression levels and identified functions are associated with a range of diseases (e. g., cancer), albeit the DE lncRNAs involved in pulmonary fibrosis have not previously been reported. It is known that RP11-275,114.4 and CTB-

51J22.1 are highly expressed in endometrial carcinoma, and both lncRNAs may play a role in cancer progression.<sup>36</sup> However, no previous studies mentioned the important functions of RP11-412H8.2 and XLOC\_302,021, which were identified in the present study, indicating their potential functions in IPF. The aforementioned lncRNAs have such regulatory roles in cancers and other diseases, as well as two lncRNAs with unknown function but significantly dysregulated in IPF, implying they may have similar roles in pulmonary fibrosis development and progression.

DE lncRNAs are linked to DE mRNAs through co-expression analysis, aiding in the prediction of lncRNA functions.<sup>37</sup> In this study, WGCNA was used to study the relationship between lncRNA and mRNA in IPF.<sup>27</sup> IPF-related modules were identified, with DE mRNAs in the MEgreen module showing enrichment in inflammatory response and NF- $\kappa$ B pathways. The pathways in the MERed module were related to the inflammation response, an important mechanism involved in IPF. IPF-related modules and specific lncRNAs and mRNAs were detected by WGCNA, indicating that DE lncRNAs might function in inflammation-related signal pathways in IPF. Therefore, these findings suggest that certain lncRNAs in the co-expression network play a role in pulmonary fibrosis by influencing extracellular matrix remodeling and inflammation. The hub of co-expressed lncRNAs and mRNAs was identified, and a regulatory network was created. The lncRNAs and mRNAs in the network included FAM13A-AS1, MYO16-AS1, IL1RL1, S100A12, S100A8, AC007278.2 and RP11-153M7.5, respectively. FAM13A-AS1 is a novel biomarker for the prognosis of thyroid cancer,<sup>38</sup> and



**Fig. 5.** Differentially expressed DEmRNAs undergo differential changes in the specific cellular groups by scRNA-seq analysis. (A) Bar chart showed the expression of S100A9 mRNA in normal and IPF groups (left) with the decrease of S100A9 expression in T/NKT cells and B cells (right). (B) Bar chart showed the expression of S100A12 mRNA in normal and IPF groups (left) with the decrease of S100A12 expression in macrophages (right). (C) Bar chart showed the expression of CLEC4E mRNA in normal and IPF groups (left) with CLEC4E expression decreases in dendritic cells and monocytes (right). (D) Bar chart showing the expression of IL-1R2 mRNA in normal and IPF groups (left) with the decrease of IL-1R2 expression occurs in Dendritic Cells and monocytes (right). (E) Bar chart showed the expression of CRIP1 mRNA in normal and IPF groups (left) with the increase of CRIP1 expression in club cells and fibroblasts (right).

MYO16-AS1 acts as an oncogenic lncRNA in bladder cancer,<sup>39</sup> AC007278.2 and RP11-153M7.5 are associated with SLE.<sup>40,41</sup> Through network analysis, the hub lncRNAs may regulate the expression of co-expressed mRNAs through multiple mechanisms, such as acting as a sponge of miRNAs, modulating mRNA by transcriptional regulation, or chromatin modification.<sup>32</sup>

At the conclusion of this study, the authors observed aberrant expression of DEmRNAs in the red and green modules associated with IPF through the analysis of the scRNA-seq dataset. Notably, genes such as S100A12, S100A9, CLEC4E, CRIP1, and IL1R2 exhibited abnormal expression patterns. GO analysis revealed enrichment of genes in inflammatory and immune response processes, as well as positive regulation of inflammatory factor activity and NF- $\kappa$ B pathways. Furthermore, the present findings demonstrated that CLEC4E expression was reduced in macrophages, IL1R2 expression was reduced in dendritic cells and monocytes, and CRIP1 expression was increased in club cells and fibroblasts. S100A9 expression was decreased in T/NKT cells and B-cells, whereas the expression of S100A12 was observed to be increased in monocytes. These observed changes in signaling pathways align with the physiological functions associated with immune cells expressing these mRNA molecules. Previous research has demonstrated the significant involvement of S100A8/A9 in initiating inflammation and facilitating the development of a self-perpetuating inflammatory cycle that contributes to fibrosis, primarily through the activation of neutrophils and macrophages.<sup>42</sup> Additionally, S100A12 exhibited predominant and elevated expression levels in monocytes. There have been limited prior investigations conducted on the other three genes. Notably, the co-expression of DElncRNAs with these five DEmRNAs (e.g.,

RP11-434D9.1, RP11-203H2.1, MYO16-AS1, RP11-180C16.1, RP11-87E22.1, RP11-571L19.8, FAM13A-AS1, etc.) was not observed, indicating that these DEmRNAs may be regulated by other factors or indirectly by lncRNAs. Thus, identifying DElncRNAs that regulate mRNA targets could provide new research opportunities for diagnosing and treating IPF.

Existing studies have reported divergent roles of certain lncRNAs in IPF compared to these findings. For instance, some research suggests that specific lncRNAs primarily influence IPF through the modulation of cell proliferation-related pathways. LncRNA NONMMUT060091 (PFI) was found to be significantly downregulated in mouse models of lung fibrosis, while it was upregulated in human lung fibroblasts. This lncRNA plays a critical role in the pathogenesis of pulmonary fibrosis by interacting with and modulating the expression of Serine/arginine-Rich Splicing Factor-1 (SRSF1), a protein primarily involved in RNA splicing processes.<sup>43</sup> LncRNA DACH1 is significantly down-regulated in the lungs of patients with IPF as well as in experimental mouse models of pulmonary fibrosis. Overexpression of lncRNA DACH1 has been shown to attenuate the aberrant activation, collagen deposition, and differentiation of mouse lung fibroblasts induced by TGF $\beta$ 1.<sup>44</sup> This discrepancy may be attributed to several factors, including variations and bias in sample sources. The samples utilized in different studies may differ in terms of disease stages and patient-specific characteristics, such as age, comorbidities, and treatment history. More samples should be considered in future studies.

The present study identified dysregulated lncRNAs and their potential targets in pulmonary fibrosis, expanding the knowledge of their functional properties. However, the limited sample size and

undetermined clinical characteristics of patients such as age, smoking status, and treatment history, of the RNA-seq dataset were shortcomings in this study to identify the highly confident DE lncRNAs in IPF, and more samples are needed for validation in the future. Another limitation of the study is the necessity of further experiments, such as expression validation or knockdown cellular or animal models, to validate the functions of differentially co-expressed lncRNA-mRNA pairs. These pairs have validated functions in other diseases, indicating their potential relevance in IPF progression. Further research is required to validate the interaction between these DE lncRNAs and mRNAs at various levels using larger and ethnically diverse cohorts. Meanwhile, predictive modeling or biomarker panels should be constructed on these DE lncRNAs to validate their potential in IPF diagnosis. The study results offer insights for studying the molecular mechanisms of pulmonary fibrosis and suggest new directions for clinical research on potential therapeutic targets.

## Conclusion

This study analyzed the genome-wide profiling of lncRNAs in IPF, detecting abundant lncRNAs and building co-expression networks with DE mRNAs. The study found that DE lncRNAs and their co-expressed DE mRNAs were enriched in pathways related to inflammation, providing insight into the pathogenetic mechanisms of pulmonary fibrosis and, importantly, this study offers new avenues for investigating the development of IPF with lncRNAs in future preclinical and clinical studies.

## Declarations

Ethics approval and consent to participate: The public data sets used in this study have passed the original research ethics review and all data have been anonymized in compliance with the Helsinki Declaration.

## Informed consent and patient details

Not applicable.

## Consent for publication

Not applicable.

**Availability of data and materials:** The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

## Funding

Henan Province Medical Science and Technology Key Project (Jointly Constructed by the Province and the Ministry) (SBGJ202402013).

## Conflicts of interest

The authors declare no conflicts of interest.

## Acknowledgements

No.

## Supplementary materials

Supplementary material associated with this article can be found, in the online version, at [doi:10.1016/j.clinsp.2025.100732](https://doi.org/10.1016/j.clinsp.2025.100732).

## References

- Lederer DJ, Martinez FJ. Idiopathic pulmonary Fibrosis. *N Engl J Med*. 2018;378(19):1811–1823.
- Burgy O, Bellay P-S, Beltramo G, Goirand F, Bonniaud P. Pathogenesis of fibrosis in interstitial lung disease. *Curr Opin Pulm Med*. 2020;26(5):429–435.
- Kondoh Y, Suda T, Hongo Y, Yoshida M, Hiroi S, Iwasaki K. Prevalence of idiopathic pulmonary fibrosis in Japan based on a claims database analysis. *Respir Res*. 2022;23(1):24.
- Harari S, Davi M, Biffi A, Caminati A, Ghirardini A, Lovato V, et al. Epidemiology of idiopathic pulmonary fibrosis: a population-based study in primary care. *Intern Emerg Med*. 2020;15(3):437–445.
- Mortimer KM, Bartels DB, Hartmann N, Capapey J, Yang J, Gately R, et al. Characterizing health outcomes in idiopathic pulmonary fibrosis using US health claims data. *Respiration*. 2020;99(2):108–118.
- Ding Q, Luckhardt T, Hecker L, Zhou Y, Liu G, Antony VB, et al. New insights into the pathogenesis and treatment of idiopathic pulmonary fibrosis. *Drugs*. 2011;71(8):981–1001.
- Smith CM, Catchpole D, Hutvagner G. Non-coding RNAs in pediatric solid tumors. In: *Front Genet*. 10. 2019:798.
- Niederer RO, Hass EP, Zappulla DC. Long noncoding RNAs in the Yeast *S. cerevisiae*. *Adv Exp Med Biol*. 2017;1008:119–132.
- Li C, Wang Z, Zhang J, Zhao X, Xu P, Liu X, et al. Crosstalk of mRNA, miRNA, lncRNA, and circRNA and their regulatory pattern in pulmonary fibrosis. *Mol Ther Nucleic Acids*. 2019;18:204–218.
- Lu Q, Guo Z, Xie W, Jin W, Zhu D, Chen S, et al. The lncRNA H19 mediates pulmonary fibrosis by regulating the miR-196a/COL1A1 axis. *Inflammation*. 2018;41(3):896–903.
- Qian W, Cai X, Qian Q, Peng W, Yu J, Zhang X, et al. lncRNA ZEB1-AS1 promotes pulmonary fibrosis through ZEB1-mediated epithelial-mesenchymal transition by competitively binding miR-141-3p. *Cell Death Dis*. 2019;10(2):129.
- Hao X, Du Y, Qian L, Li D, Liu X. Upregulation of long noncoding RNA AP003419.16 predicts high risk of aging-associated idiopathic pulmonary fibrosis. *Mol Med Rep*. 2017;16(6):8085–8091.
- Zhao X, Sun J, Chen Y, Su W, Shan H, Li Y, et al. lncRNA PFAR promotes lung fibroblast activation and fibrosis by targeting miR-138 to regulate the YAP1-twist axis. *Mol Ther*. 2018;26(9):2206–2217.
- Chen T, Guo Y, Wang J, Ai L, Ma L, He W, et al. lncRNA CTD-2528L19.6 prevents the progression of IPF by alleviating fibroblast activation. *Cell Death Dis*. 2021;12(6):600.
- Lin S, Zhang R, Xu L, Ma R, Xu L, Zhu L, et al. lncRNA Hoxaas3 promotes lung fibroblast activation and fibrosis by targeting miR-450b-5p to regulate Runx1. *Cell Death Dis*. 2020;11(8):706.
- Li J, Jiang Z-Z, Li Y-Y, Tang W-T, Yin J, Long X-P. lncRNA CHRF promotes TGF-β1 induced EMT in alveolar epithelial cells by inhibiting miR-146a up-regulating L1CAM expression. *Exp Lung Res*. 2021;47(4):198–209.
- Yi H, Luo D, Xiao Y, Jiang D. Knockdown of long non-coding RNA DLEU2 suppresses idiopathic pulmonary fibrosis by regulating the microRNA-369-3p/TRIM2 axis. *Int J Mol Med*. 2021;47(5):80.
- Li X, Yu T, Shan H, Jiang H, Sun J, Zhao X, et al. lncRNA PFAL promotes lung fibrosis through CTGF by competitively binding miR-18a. *FASEB J*. 2018;32(10):5285–5297.
- Senavirathna LK, Liang Y, Huang C, Yang X, Bamunuarachchi G, Xu D, et al. Long noncoding RNA FENDRR inhibits lung fibroblast proliferation via a reduction of beta-catenin. *Int J Mol Sci*. 2021;22(16):8536.
- Yin Q, Strong MJ, Zhuang Y, Flemington EK, Kaminski N, Andrade JA, et al. Assessment of viral RNA in idiopathic pulmonary fibrosis using RNA-seq. *BMC Pulm Med*. 2020;20(1):81.
- Reyffman PA, Walter JM, Joshi N, Anekalla KR, McQuattie-Pimentel AC, Chiu S, et al. Single-cell transcriptomic analysis of Human lung provides insights into the pathobiology of pulmonary fibrosis. *Am J Respir Crit Care Med*. 2019;199(12):1517–1536.
- Hao Y, Hao S, Andersen-Nissen E, Mauck 3rd WM, Zheng S, Butler A, et al. Integrated analysis of multimodal single-cell data. *Cell*. 2021;184(13):3573–3587.e29.
- Ianevski A, Giri AK, Aittokallio T. Fully-automated and ultra-fast cell-type identification using specific marker combinations from single-cell transcriptomic data. *Nat Commun*. 2022;13(1):1246.
- Reyffman PA, Walter JM, Joshi N, Anekalla KR, McQuattie-Pimentel AC, Chiu S, et al. Single-cell transcriptomic analysis of Human lung reveals complex multicellular changes during pulmonary fibrosis. *Am J Respir Crit Care Med*. 2019;199(12):1517–1536.
- Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL, et al. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol*. 2013;14(4):R36.
- Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014;15(12):550.
- Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinform*. 2008;9:559.
- Xie C, Mao X, Huang J, Ding Y, Wu J, Dong S, et al. KOBAS 2.0: a web server for annotation and identification of enriched pathways and diseases. *Nucleic Acids Res*. 2011;39(Web Server issue):W316–W322.
- Liu S, Wang Z, Chen D, Zhang B, Tian R-R, Wu J, et al. Annotation and cluster analysis of spatiotemporal- and sex-related lncRNA expression in Rhesus macaque brain. *Genome Res*. 2017;27(9):1608–1620.

30. Moss BJ, Ryter SW, Rosas IO. Pathogenic mechanisms underlying idiopathic pulmonary fibrosis. *Annu Rev Pathol.* 2022;17:515–546.
31. Zhang J, Chen X, Chen H, Li R, Xu P, Lv C, et al. Engeletin ameliorates pulmonary fibrosis through endoplasmic reticulum stress depending on lnc949-mediated TGF- $\beta$ 1-Smad2/3 and JNK signalling pathways. *Pharm Biol.* 2020;58(1):1105–1114.
32. Mattick JS, Amaral PP, Carninci P, Carpenter S, Chang HY, Chen L-L, et al. Long non-coding RNAs: definitions, functions, challenges and recommendations. *Nat Rev Mol Cell Biol.* 2023;24(6):430–447.
33. Herman AB, Tsitsipatis D, Gorospe M. Integrated lncRNA function upon genomic and epigenomic regulation. *Mol Cell.* 2022;82(12):2252–2266.
34. Yao RW, Wang Y, Chen LL. Cellular functions of long noncoding RNAs. *Nat Cell Biol.* 2019;21(5):542–551.
35. Junior C, Ulldemolins A, Narciso M, Almendros I, Farré R, Navajas D, et al. Multi-step extracellular matrix remodelling and stiffening in the development of idiopathic pulmonary fibrosis. *Int J Mol Sci.* 2023;24(2):1708.
36. Sun Y, Zou X, He J, Mao Y. Identification of long non-coding RNAs biomarkers associated with progression of endometrial carcinoma and patient outcomes. *Oncotarget.* 2017;8(32):52604–52613.
37. Statello L, Guo C-J, Chen L-L, Huarte M. Gene regulation by long non-coding RNAs and its biological functions. *Nat Rev Mol Cell Biol.* 2021;22(2):96–118.
38. Rao Y, Liu H, Yan X, Wang J. In silico analysis identifies differently expressed lncRNAs as novel biomarkers for the prognosis of thyroid cancer. *Comput Math Methods Med.* 2020;2020, 3651051.
39. Shen D, Zhang Y, Zheng Q, Yu S, Xia L, Cheng S, et al. A competing endogenous RNA network and an 8-lncRNA prognostic signature identify MYO16-AS1 as an oncogenic lncRNA in bladder cancer. *DNA Cell Biol.* 2021;40(1):26–35.
40. You Y, Zhao X, Wu Y, Mao J, Ge L, Guo J, et al. Integrated transcriptome profiling revealed that elevated long non-coding RNA-AC007278.2 expression repressed CCR7 transcription in systemic Lupus Erythematosus. *Front Immunol.* 2021;12, 615859.
41. Cao H, Li D, Lu H, Sun J, Li H. Uncovering potential lncRNAs and nearby mRNAs in systemic lupus erythematosus from the Gene Expression Omnibus dataset. *Epigenomics.* 2019;11(16):1795–1809.
42. Araki K, Kinoshita R, Tomonobu N, Gohara Y, Tomida S, Takahashi Y, et al. The heterodimer S100A8/A9 is a potent therapeutic target for idiopathic pulmonary fibrosis. *J Mol Med (Berl).* 2021;99(1):131–145.
43. Sun J, Jin T, Su W, Guo Y, Niu Z, Guo J, et al. The long non-coding RNA PFI protects against pulmonary fibrosis by interacting with splicing regulator SRSF1. *Cell Death Differ.* 2021;28(10):2916–2930.
44. Sun J, Jin T, Niu Z, Guo J, Guo Y, Yang R, et al. lncRNA DACH1 protects against pulmonary fibrosis by binding to SRSF1 to suppress CTNNB1 accumulation. *Acta Pharm Sin B.* 2022;12(9):3602–3617.