



OPEN ACCESS

EDITED BY

Tingxiang Yan,
Mayo Clinic Florida, United States

REVIEWED BY

Weilin Pu,
Fudan University, China
Manuel De Jesús Castillejos López,
National Institute of Respiratory
Diseases-Mexico (INER), Mexico

*CORRESPONDENCE

Heng Yang,
✉ johnnelyang@hotmail.com
Wen Ouyang,
✉ ouyangwen133@vip.sina.com

SPECIALTY SECTION

This article was submitted to RNA,
a section of the journal
Frontiers in Genetics

RECEIVED 30 September 2022

ACCEPTED 20 February 2023

PUBLISHED 27 February 2023

CITATION

Zeng Y, Huang J, Guo R, Cao S, Yang H
and Ouyang W (2023), Identification and
validation of metabolism-related hub
genes in idiopathic pulmonary fibrosis.
Front. Genet. 14:1058582.
doi: 10.3389/fgene.2023.1058582

COPYRIGHT

© 2023 Zeng, Huang, Guo, Cao, Yang and
Ouyang. This is an open-access article
distributed under the terms of the
[Creative Commons Attribution License
\(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or
reproduction in other forums is
permitted, provided the original author(s)
and the copyright owner(s) are credited
and that the original publication in this
journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

Identification and validation of metabolism-related hub genes in idiopathic pulmonary fibrosis

Youjie Zeng¹, Jun Huang², Ren Guo², Si Cao¹, Heng Yang^{3*} and Wen Ouyang^{1*}

¹Department of Anesthesiology, Third Xiangya Hospital, Central South University, Changsha, China,

²Department of Pharmacy, Third Xiangya Hospital, Central South University, Changsha, China,

³Department of Neurology, Third Xiangya Hospital, Central South University, Changsha, China

Background: Idiopathic pulmonary fibrosis (IPF) is a fatal and irreversible interstitial lung disease. The specific mechanisms involved in the pathogenesis of IPF are not fully understood, while metabolic dysregulation has recently been demonstrated to contribute to IPF. This study aims to identify key metabolism-related genes involved in the progression of IPF, providing new insights into the pathogenesis of IPF.

Methods: We downloaded four datasets (GSE32537, GSE110147, GSE150910, and GSE92592) from the Gene Expression Omnibus (GEO) database and identified differentially expressed metabolism-related genes (DEMRGs) in lung tissues of IPF by comprehensive analysis. Then, we performed GO, KEGG, and Reactome enrichment analyses of the DEMRGs. Subsequently, key DEMRGs were identified by machine-learning algorithms. Next, miRNAs regulating these key DEMRGs were predicted by integrating the GSE32538 (IPF miRNA dataset) and the miRWalk database. The Cytoscape software was used to visualize miRNA-mRNA regulatory networks. In addition, the relative levels of immune cells were assessed by the CIBERSORT algorithm, and the correlation of key DEMRGs with immune cells was calculated. Finally, the mRNA expression of the key DEMRGs was validated in two external independent datasets and an *in vivo* experiment.

Results: A total of 101 DEMRGs (51 upregulated and 50 downregulated) were identified. Six key DEMRGs (ENPP3, ENTPD1, GPX3, PDE7B, PNMT, and POLR3H) were further identified using two machine-learning algorithms (LASSO and SVM-RFE). In the lung tissue of IPF patients, the expression levels of ENPP3, ENTPD1, and PDE7B were upregulated, and the expression levels of GPX3, PNMT, and POLR3H were downregulated. In addition, the miRNA-mRNA regulatory network of key DEMRGs was constructed. Then, the expression levels of key DEMRGs were validated in two independent external datasets (GSE53845 and GSE213001). Finally, we verified the key DEMRGs in the lung tissue of bleomycin-induced pulmonary fibrosis mice by qRT-PCR.

Conclusion: Our study identified key metabolism-related genes that are differentially expressed in the lung tissue of IPF patients. Our study emphasizes the critical role of metabolic dysregulation in IPF, offers potential therapeutic targets, and provides new insights for future studies.

KEYWORDS

IPF, metabolism, metabolic, bioinformatics, hub genes, gene expression omnibus, differentially expressed genes, biomarker

1 Introduction

Idiopathic pulmonary fibrosis (IPF) is a progressive, life-threatening, chronic interstitial lung disease of unknown etiology (Noble et al., 2012). It is characterized by progressive scarring of the lung parenchyma, accompanied by a continuous deterioration of respiratory symptoms and a decline in lung function, ultimately leading to death (Raghu et al., 2018). Approximately two to 3 years is the median survival time for patients with IPF after diagnosis (Ley et al., 2011). There is a higher prevalence of IPF in the elderly, and the mean age of patients with IPF is around 65–70 years (Maher et al., 2021). The FDA currently approves two antifibrotic drugs (nintedanib and pirfenidone) for IPF, which only slow, not stop, fibrosis progression (Saito et al., 2019). IPF is currently curable only through lung transplantation (Shenderov et al., 2021). Despite identifying several candidate biomarkers for IPF, none of these markers have yet been translated into clinical practice (Ley et al., 2014). Thus, there is an urgent need to explore the pathophysiological mechanisms of IPF further and develop new targeted therapeutic strategies.

An increasing number of studies have recently demonstrated the role of metabolic dysregulation in IPF. For instance, Kang et al. reported altered glycolysis and glutamine metabolism in human lungs with severe IPF (Kang et al., 2016). Furthermore, proteomics studies revealed dysregulated levels of transcription factors NF- κ B, PPAR γ , and c-myc in bronchoalveolar lavage fluid (BALF) from IPF patients compared to healthy controls (Landi et al., 2014). Interestingly, these transcription factors have been reported to participate in numerous metabolic dysregulation mechanisms (Kauppinen et al., 2013; Jiang et al., 2017; Botta et al., 2018). In addition, lung fibroblasts and alveolar epithelial cells have been observed to display profibrotic phenotypes due to dysregulated lipid metabolism (Mamazhakypov et al., 2019). A recent review summarized the proteins dysregulated in IPF involving the renin-angiotensin-aldosterone system, hypoxia, oxidative stress, iron metabolism, dysregulated lipid metabolism, and mitochondrial alterations, highlighting the potential impact of metabolic dysregulation in IPF (Bargagli et al., 2020). Conclusively, there is an inescapable relationship between metabolic dysregulation and IPF, and the search for novel metabolism-related markers can help further understand the metabolism-related pathological molecular mechanisms of IPF. Rectifying these metabolic alterations is emerging as a promising new strategy for antifibrotic therapy.

Our study first analyzed GSE32537, GSE110147, GSE150910, and GSE92592 from the Gene Expression Omnibus (GEO) database and identified differentially expressed metabolism-related genes (DEMREGs) in the lung tissue of IPF patients. Subsequently, we conducted a functional enrichment analysis of DEMREGs. Then, we used two machine-learning methods, least absolute shrinkage and selection operator (LASSO) regression and support vector machine recursive feature elimination (SVM-RFE), to identify six IPF signature genes as key DEMREGs: ENPP3, ENTPD1, PDE7B, GPX3, PNMT, and POLR3H. The expression of ENPP3, ENTPD1, and PDE7B was significantly upregulated in IPF patients' lung tissue, and the expression of GPX3, PNMT, and POLR3H was significantly downregulated. Afterward, we predicted miRNAs regulating key DEMREGs using the miRWalk database, combining it with the GSE32538 dataset (miRNA

microarray expression profiles of IPF) to construct a miRNA-mRNA regulatory network. Next, the relative levels of immune cells were assessed by the CIBERSORT algorithm, and the correlation of key DEMREGs with immune cells was calculated. Finally, we validated the expression patterns of six key DEMREGs by analyzing the external independent dataset GSE53845 and performing qRT-PCR.

2 Materials and methods

2.1 Study design

Figure 1 shows the overall flow chart of this study. First, we performed differential expression analysis on four GEO gene expression profile datasets to identify common differentially expressed metabolism-related genes (DEMREGs) in IPF lung tissues. Subsequently, we performed a functional enrichment analysis for these common DEMREGs. Then, we identified key DEMREGs using two machine-learning algorithms. Finally, we constructed potential miRNA-mRNA regulatory networks for key DEMREGs, calculated the correlation of key DEMREGs with immune cell levels, and validated the expression of key DEMREGs in external GEO datasets and animal models.

2.2 Gene expression profile data

Our study obtained publicly available datasets from the Gene Expression Omnibus (GEO; <https://www.ncbi.nlm.nih.gov/geo/>) database (Edgar et al., 2002). The GEO database is a public database that stores extensive publicly available high-throughput gene expression and other functional genomics datasets (Clough and Barrett, 2016). All datasets downloaded in this study stored gene expression at the mRNA level (array or high-throughput sequencing), and samples in the dataset were obtained from the lung tissue of IPF patients and healthy control individuals. First, we performed a comprehensive bioinformatics analysis of GSE32537, GSE110147, GSE150910, and GSE92592 to identify key DEMREGs. Then, using GSE32538, we constructed a miRNA-mRNA regulatory network of key DEMREGs. Finally, we validated the key DEMREGs in GSE53845 and GSE213001. Table 1 provides details of all the GEO datasets used in our study.

2.3 Screening of differentially expressed metabolism-related genes (DEMREGs)

The metabolism-related genes (MRGs) were obtained from the Molecular Signatures database (MSigDB, <https://www.gsea-msigdb.org/gsea/msigdb>) (Liberzon et al., 2011). Specifically, we first downloaded the KEGG gene set (c2_cp.kegg.v7.5.1.symbols.gmt) from the MSigDB, then searched for the keyword "metabolism" to obtain metabolism-related terms, and finally we integrated the genes within these selected metabolism-related terms, which were defined as MRGs to be used for subsequent analysis in our study. All metabolism-related terms and the genes within each term are shown in Supplementary Table S1.

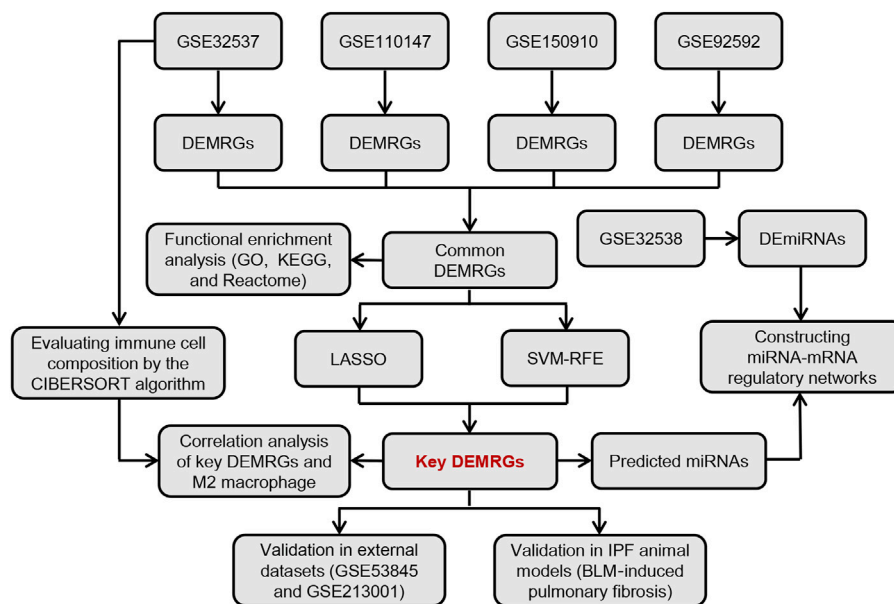


FIGURE 1
The overall flow chart of this study.

TABLE 1 Details of all the GEO datasets used in this study.

Accession number	Platform	Samples	Experiment type
GSE32537	GPL6244	119 IPF lung tissues vs 50 healthy lung tissues	Array
GSE110147	GPL6244	22 IPF lung tissues vs 11 healthy lung tissues	Array
GSE150910	GPL24676	103 IPF lung tissues vs 103 healthy lung tissues	High throughput sequencing
GSE92592	GPL11154	20 IPF lung tissues vs 19 healthy lung tissues	High throughput sequencing
GSE32538	GPL8786	106 IPF lung tissues vs 50 healthy lung tissues	Array (miRNA)
GSE53845	GPL6480	40 IPF lung tissues vs 8 healthy lung tissues	Array
GSE213001	GPL21290	62 IPF lung tissues vs 41 healthy lung tissues	High throughput sequencing

GSE32537 and GSE110147 were based on the GPL6244 platform of Affymetrix Human Gene 1.0 ST Array. We used the RMA algorithm *via* the “oligo” R package for background correction and normalization of the raw data in the two datasets. Subsequently, differentially expressed genes were identified using the “limma” R package. GSE150910 and GSE92592 were RNA-seq datasets that were generated using the Illumina platform. We first downloaded their raw gene count matrix files. Then, we performed differential expression analysis on the gene expression matrix normalized by the *vst* function of the “Deseq2” R package. An adjusted *p*-value <0.05 was set as the threshold for identifying differentially expressed genes.

After acquiring the DEMRGs from each of the four GEO datasets, we used the Venn diagram to search for the common upregulated DEMRGs and the common downregulated DEMRGs. The “ggvenn” R package was applied to plot the Venn diagrams of common DEMRGs of the four datasets.

2.4 Functional enrichment analysis of DEMRGs

We performed an enrichment analysis of the common DEMRGs using the Database for Annotation, Visualization and Integrated Discovery (DAVID database, <https://david.ncifcrf.gov/>) (Huang *et al.*, 2009; Sherman *et al.*, 2022). We performed three categories of enrichment analysis: Gene ontology (GO) enrichment analysis, Kyoto Encyclopedia of Genes and Genome (KEGG) pathway enrichment analysis, and Reactome pathway enrichment analysis. In addition, the GO enrichment analysis includes three sections: biological process (BP), cellular component (CC), and molecular function (MF). We downloaded the enrichment analysis results and defined the false discovery rate (FDR) < 0.05 as the significant enrichment threshold. In addition, we selected the top 10 most significantly enriched terms in each category and imported these results into the SangerBox platform to generate dot plots for visualization (Shen *et al.*, 2022).

2.5 Screening of IPF key DEMRGs

To identify the most critical DEMRGs, we used two machine-learning algorithms: least absolute shrinkage and selection operator (LASSO) regression and support vector machine recursive feature elimination (SVM-RFE). The LASSO algorithm is a regression analysis method that minimizes regression coefficients through successive shrinkage operations to reduce the possibility of overfitting, thereby reducing redundancy and eliminating uncorrelated genes from these analyses (Friedman et al., 2010). The SVM-RFE algorithm is a method for feature selection based on SVM that defines the minimum classification error and avoids overfitting and thus is frequently used to select the optimal genes (Duan et al., 2005). The LASSO and SVM-RFE algorithms were implemented respectively by the “glmnet” package and the “e1071” package in R software. By using the two machine-learning algorithms, two sets of DEMRGs can be obtained, and the overlapping genes of these two sets of DEMRGs will be identified as the IPF key DEMRGs.

2.6 Construction of miRNA-mRNA regulatory networks for key DEMRGs

We intend to investigate further the miRNAs that regulate these key DEMRGs, so we first performed a differential expression analysis of GSE32538 (IPF miRNA expression profile microarray) to obtain differentially expressed miRNAs (DEmiRNAs). The significance threshold was set at an adjusted *p*-value <0.05. Since the IDs of the miRNAs in this dataset were derived from an older version of miRBase, we updated the miRNA IDs using the miEAA 2.0 database (Kern et al., 2020). Subsequently, we predicted miRNAs that interacted with key DEMRGs using the miRWalk database (<http://mirwalk.umm.uni-heidelberg.de/>) (Sticht et al., 2018). If a DEmiRNA was present in the predicted miRNAs from miRWalk, it would be included in the final miRNA-mRNA regulatory network. Therefore, the upregulated DEmiRNAs were then intersected with the predicted miRNAs that interact with downregulated key DEMRGs, while the downregulated DEmiRNAs were intersected with the predicted miRNAs that interact with upregulated key DEMRGs. Finally, we visualized the miRNA-mRNA regulatory network in the Cytoscape software (v 3.9.1) (Shannon et al., 2003).

2.7 Immune infiltration analysis

We assessed the relative content of immune cells of each sample in the GSE32537 dataset using the CIBERSORT algorithm in R software (Newman et al., 2015). The CIBERSORT algorithm calculates the relative expression of 22 immune cells based on the “LM22” matrix downloaded from the CIBERSORT portal (<http://cibersort.stanford.edu/>). First, we evaluated the relative expression of immune cells in all samples and plotted a histogram of immune cell content for each sample. Subsequently, we compared the content of each immune cell between IPF patients and healthy controls and plotted a boxplot for visualization. The Shapiro-Wilk test was

TABLE 2 qRT-PCR primer sequences.

Gene	Primer sequence (5' → 3')
ENPP3	F: CAGCAACGGTGAAAGCAAAT
	R: CTGATGTAGTCCCTGTGGTAAAG
PDE7B	F: ACTCTGTGTGTACACCTCTTC
	R: GGTGTGACCGTGGTAATCT
ENTPD1	F: AACTGTCCACCGAACTGATAC
	R: CCGATTGTTCCGTTTCCATTC
PNMT	F: GGGACGGTCTCTCATTGATATT
	R: CTGACGGTGACTIONCCAAGAA
POLR3H	F: CCAGGGCCTCTTTCATGTT
	R: CTGCTCTGCCACCACTATT
GPX3	F: CCTTTTAAGCAGTATGCAGGCA
	R: CAAGCCAAATGGCCCAAGTT
GAPDH	F: GAGCATCTCCCTCACAATTC
	R: GGGTGCAGCGAACTTTAT

performed to examine the normality of data, and the t-test or Mann-Whitney Wilcoxon test was used to conduct comparisons between groups based on the results of normality test (Supplementary Table S2). Finally, we calculated the correlation between 6 key DEMRGs and M2 macrophage content in 119 IPF patients. All results were visualized using the “ggplot2” R package.

2.8 Validation of key DEMRGs in independent external datasets

To improve the confidence of the results, we validated the expression of key DEMRGs in two independent external datasets (GSE53845 and GSE213001). We compared the mRNA expression levels of the key DEMRGs between IPF patients and control groups. We performed the Shapiro-Wilk test to check the normality of the data before making comparisons between groups. Based on the normality results (Supplementary Table S3, S4), we used the t-test or the Mann-Whitney Wilcoxon test to compare differences between groups. A *p*-value of <0.05 was considered statistically significant. Gene expression comparisons between groups were analyzed and visualized using the “ggplot2” package in R software (Wickham, 2016).

2.9 Construction of IPF animal models

The animal study was approved by the Laboratory Animal Welfare Ethics Committee of Central South University. Mice of the C57BL/6 strain (Adult male, 20 ± 2 g) were purchased from Hunan SJA Laboratory Animal Co., Ltd. (Hunan, China). Mice were housed in pathogen-free conditions with a 12 h dark/light cycle and were given access to food and water without restriction.

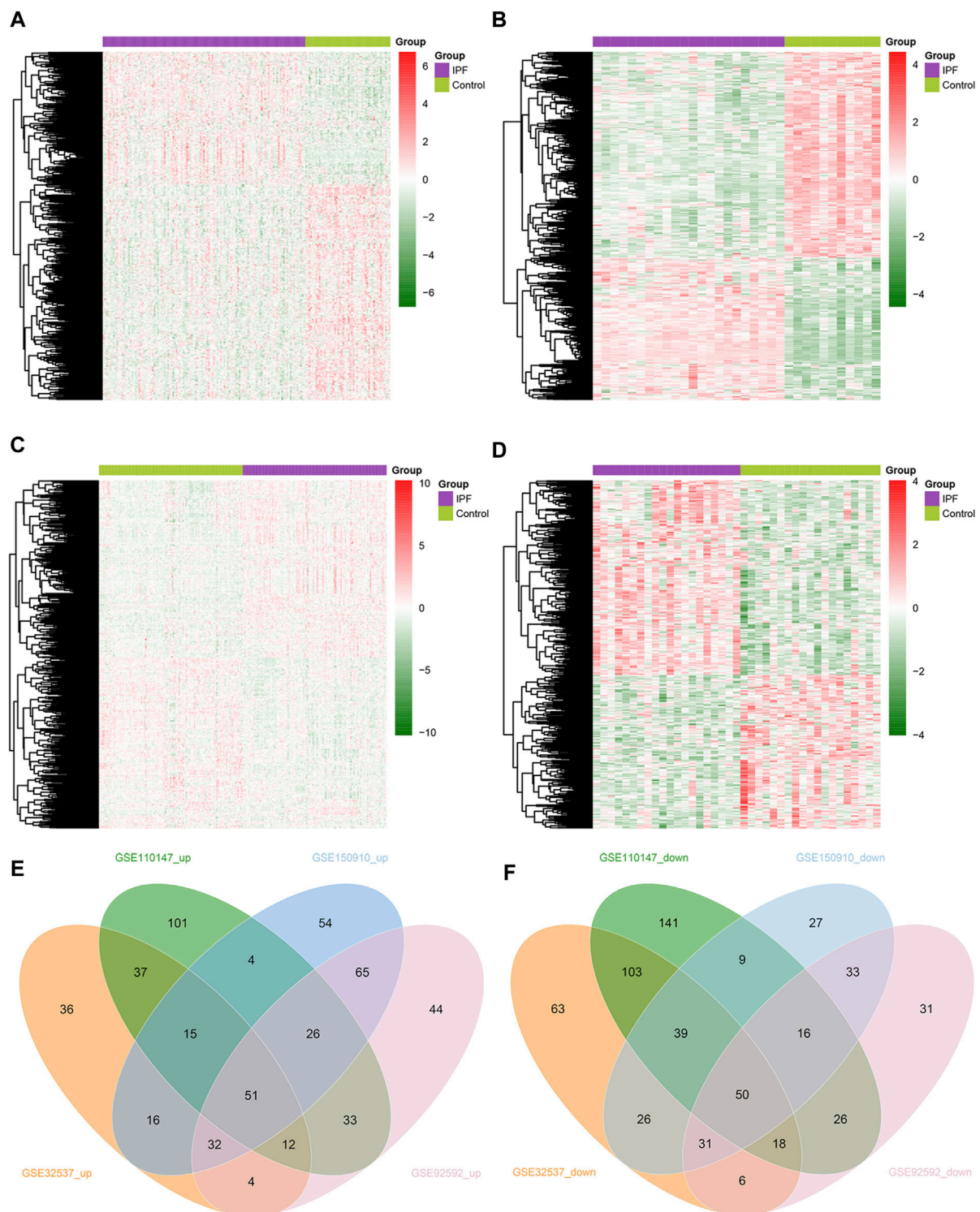
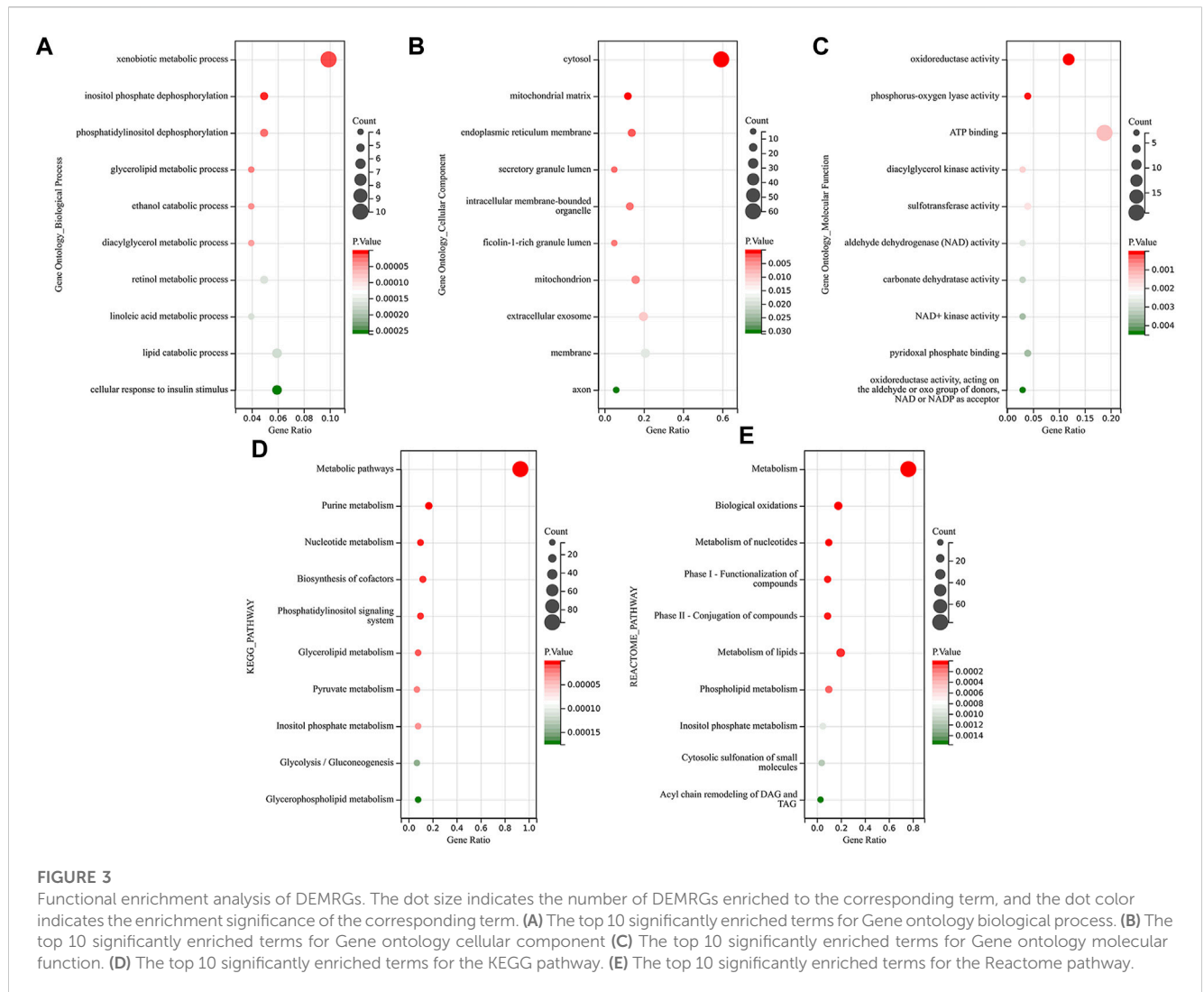


FIGURE 2

Identification of DEMRGs in IPF. **(A)** Heatmap of DEMRGs in GSE32537 (203 upregulated and 336 downregulated DEMRGs). **(B)** Heatmap of DEMRGs in GSE110147 (279 upregulated and 402 downregulated DEMRGs). **(C)** Heatmap of DEMRGs in GSE150910 (263 upregulated and 231 downregulated DEMRGs). **(D)** Heatmap of DEMRGs in GSE92592 (267 upregulated and 211 downregulated DEMRGs). **(E)** The Venn diagram identified fifty-one commonly upregulated DEMRGs. **(F)** The Venn diagram identified fifty commonly downregulated DEMRGs.

Single tracheal instillation of bleomycin (BLM) was applied to construct the pulmonary fibrosis model (Moeller et al., 2008). Mice were randomly divided into two groups: 1) Sham group (n = 6);

intra-tracheal instillation of 50 μ L saline alone; 2) BLM group (n = 6): intra-tracheal instillation of 50 μ L saline containing BLM (5 mg/kg). Before surgery, mice were anesthetized by



intraperitoneal injection of 1% sodium pentobarbital (50 mg/kg). All mice were euthanized 2 weeks after surgery, and their lung tissue was harvested.

2.10 Validation of key DEMRGs by qRT-PCR

Total RNA was extracted from lung tissue using TRIzol reagent (Invitrogen, Carlsbad, CA, United States), and qRT-PCR was performed using the ABI ViiA 7 real-time PCR system. GAPDH mRNA was used as an internal control for the key DEMRGs, and the relative fold differences were calculated using the $2^{-\Delta\Delta CT}$ method. Triplicates of all experiments were performed. Table 2 presents the qRT-PCR primer sequences utilized in our study.

Relative expression levels of the key DEMRGs were plotted in a barplot using the GraphPad Prism 8 software. Based on the normality results calculated by Shapiro-Wilk (Supplementary Table S5), the differences between groups were calculated using the t-test or Mann-Whitney Wilcoxon test, and p -values <0.05 were considered statistically significant.

3 Results

3.1 Identification of differentially expressed metabolism-related genes (DEMRGs)

We obtained 949 unique MRGs through MSigDB. Subsequently, we performed differential expression analysis on lung tissue samples from IPF patients and healthy control individuals from four GEO datasets (GSE32537, GSE110147, GSE150910, and GSE92592), and thus obtained the differentially expressed MRGs (DEMRGs) between IPF patients and healthy controls in each dataset. As a result, in GSE32537, GSE110147, GSE150910, and GSE92592, we detected 203, 279, 263, and 267 upregulated DEMRGs, respectively. In addition, we identified 336, 402, 231, and 211 DEMRGs that were downregulated in GSE32537, GSE110147, GSE150910, and GSE92592. The heat map shows the distribution of DEMRGs in the four datasets (Figures 2A–D). The red part of the heat map indicates the upregulated DEMRGs in IPF lung tissues, while the green part indicates the downregulated DEMRGs in IPF lung tissues. The Venn diagram shows that there were 51 common upregulated DEMRGs and 50 common downregulated DEMRGs in the four

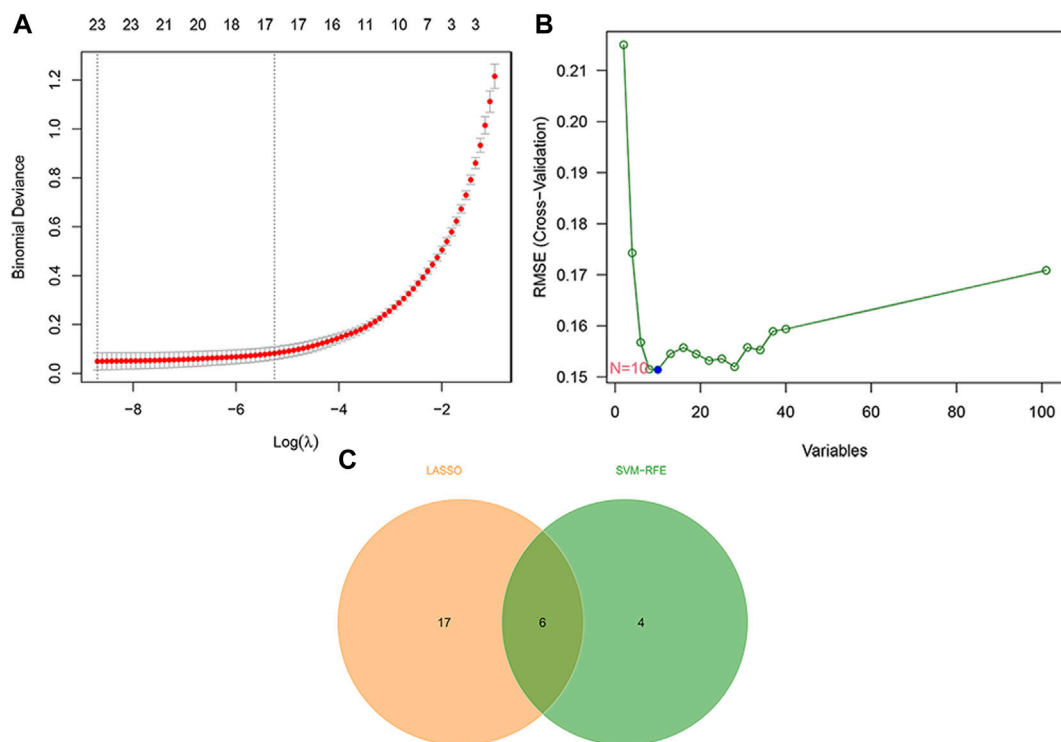


FIGURE 4

Identification of IPF key DEMRGs by using two machine-learning algorithms. (A) Twenty-three gene signatures were extracted via LASSO regression. (B) Ten gene signatures were extracted via SVM-RFE. (C) The Venn diagram identified six overlapping DEMRGs shared by LASSO and SVM-RFE. Therefore, the six overlapping DEMRGs were identified as key DEMRGs.

datasets (Figures 2E,F). These 101 common DEMRGs were used for subsequent analysis.

3.2 Gene ontology, KEGG pathway, and reactome pathway enrichment analysis

We performed a functional enrichment analysis of these 101 common DEMRGs through the DAVID database. Figure 3 shows the top 10 significantly enriched GO, KEGG, and Reactome pathway terms. The dot size indicates the number of DEMRGs enriched to the corresponding term, and the dot color indicates the enrichment significance of the corresponding term. In the BP category of the GO enrichment analysis, DEMRGs were mainly enriched in items such as “xenobiotic metabolic process”, “inositol phosphate dephosphorylation”, and “phosphatidylinositol dephosphorylation” (Figure 3A). In the CC category of the GO enrichment analysis, these genes were mainly enriched in items such as “cytosol”, “mitochondrial matrix”, and “endoplasmic reticulum membrane” (Figure 3B). In the MF category of the GO enrichment analysis, these genes were mainly enriched in items such as “oxidoreductase activity”, “phosphorus-oxygen lyase activity”, and “ATP binding” (Figure 3C). KEGG analysis showed that DEMRGs were likely related to “metabolic pathways”, “purine metabolism”, and “nucleotide metabolism” (Figure 3D). Reactome analysis indicated that DEMRGs were significantly enriched in “metabolism”, “biological oxidations”, and “metabolism of nucleotides” (Figure 3E).

3.3 Identification of IPF key DEMRGs

In order to identify key DEMRGs, the LASSO regression analysis was used to screen the gene signatures for the 101 common DEMRGs (Figure 4A), yielding 23 gene signatures. Furthermore, ten gene signatures were obtained using the SVM-RFE for the 101 common DEMRGs (Figure 4B). Finally, the Venn diagram showed that there were six overlapping DEMRGs (ENPP3, ENTPD1, GPX3, PDE7B, PNMT, and POLR3H) among the 23 genes identified by LASSO and the ten genes identified by SVM-RFE, and thus these six overlapping DEMRGs were defined as key DEMRGs (Figure 4C).

3.4 Identification of miRNA-mRNA regulatory networks of key DEMRGs

After updating the miRNA IDs by the miEAA 2.0 database, we identified 59 upregulated miRNAs and 103 downregulated miRNAs in IPF from the GSE32538 dataset (Figure 5A). Using the miRWalk database, we identified 1,295 miRNAs predicted to interact with the upregulated key DEMRGs (ENPP3, ENTPD1, and PDE7B), and they had 42 overlapping miRNAs with the 103 downregulated DEMiRNAs (Figure 5B). In addition, 1,103 miRNAs predicted to interact with the downregulated key DEMRGs (GPX3, PNMT, and POLR3H) were identified through the miRWalk database, and they had 18 overlapping

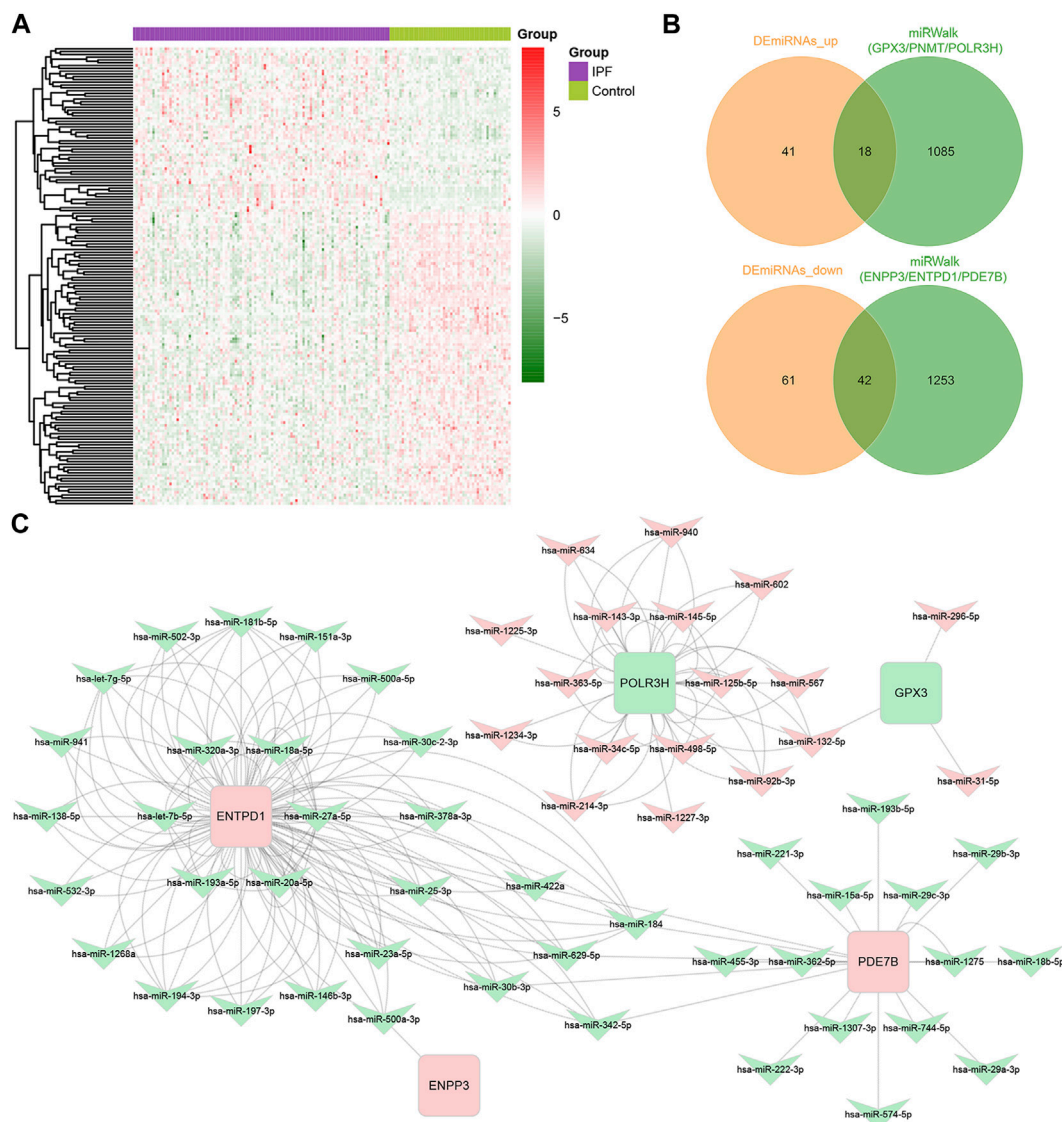


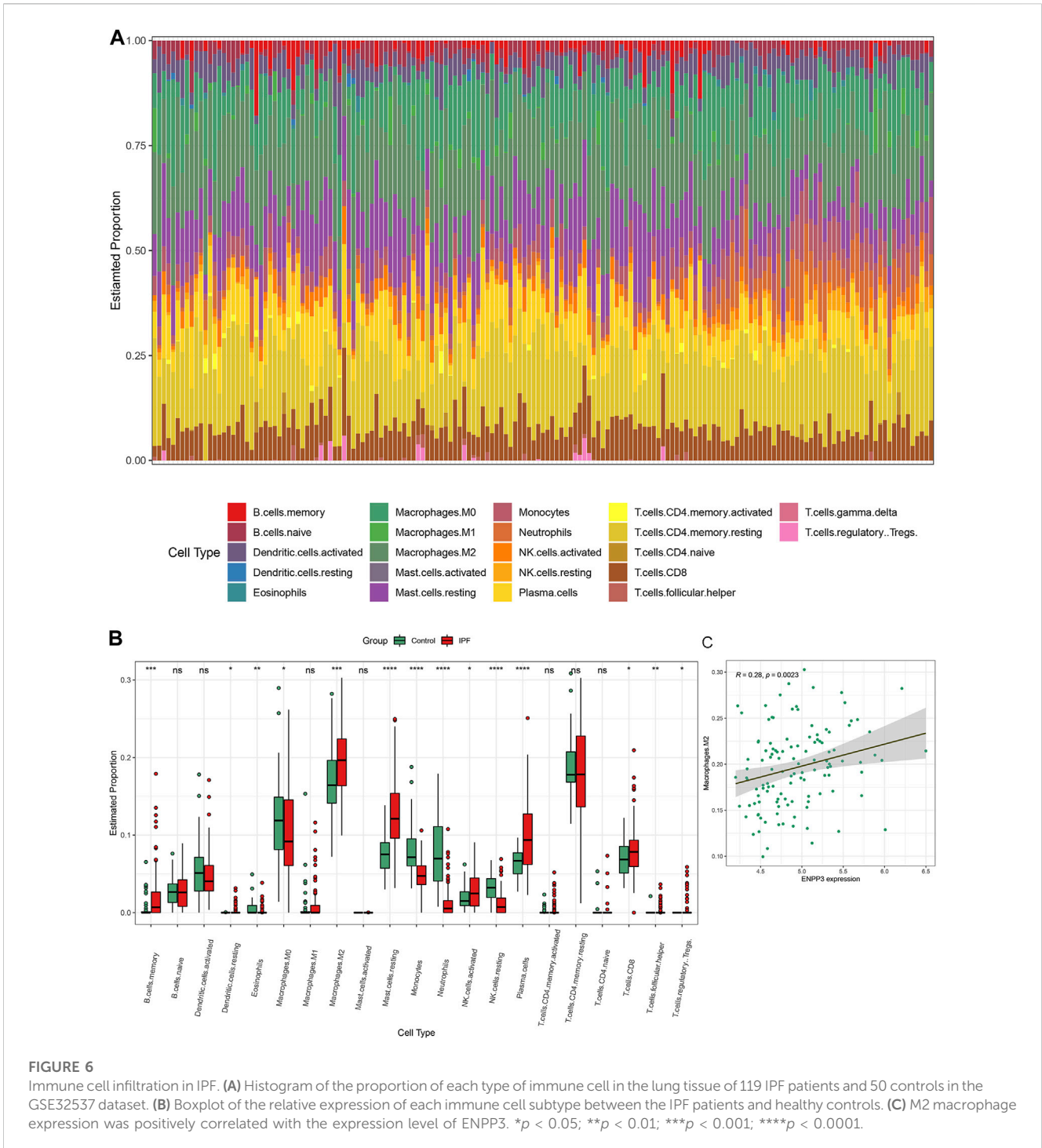
FIGURE 5 Identification of miRNA-mRNA regulatory networks of Key DEMRGs. **(A)** Heatmap of DE miRNAs in GSE32538 (59 upregulated and 103 downregulated DE miRNAs). **(B)** Venn diagram showing the intersecting miRNAs between DE miRNAs and the predicted miRNAs. **(C)** The metabolism-related miRNA-mRNA regulatory network contained 60 DE miRNAs and 5 DEMRGs. Red nodes represent upregulated key DEMRGs or DE miRNAs in IPF lung tissue, and green nodes represent downregulated key DEMRGs or DE miRNAs in IPF lung tissue.

miRNAs with the 59 upregulated DE miRNAs (Figure 5B). Except for PNMT, all other key DEMRGs have interactions with one or more overlapping miRNAs. Ultimately, a miRNA-mRNA regulatory network of 60 DE miRNAs and 5 DEMRGs was constructed by Cytoscape software, with red representing upregulation and green representing downregulation (Figure 5C).

3.5 Immune infiltration features of IPF

Figure 6A presents the distribution of immune cells in the lung tissue of 119 IPF patients and 50 healthy controls in the

GSE32537 dataset. The relative levels of many immune cells differed significantly between IPF and controls (Figure 6B). We focused on M2 macrophages because they contribute to the fibrotic phenotype exacerbation (Wynn and Vannella, 2016). A significant increase in M2 macrophages was found in the lung tissue of patients with IPF (Figure 6B). In addition, we calculated the correlation between the expression levels of six key DEMRGs and the expression levels of M2 macrophages. To minimize the false positive rate, correlation analysis was conducted on only 119 IPF patients. Figure 6C indicates that M2 macrophage expression was positively correlated with the expression level of ENPP3 ($R = 0.28, p = 0.0023$). Therefore, ENPP3 might be potentially associated with increased levels of M2 macrophages in the IPF process.



3.6 Key DEMRGs exhibited the same expression pattern in the external datasets

We compared the expression levels of six key DEMRGs in IPF patients and controls in two independent external datasets (GSE53845 and GSE213001). According to the results, ENPP3, ENTPD1, and PDE7B were significantly upregulated in the lung tissues of IPF patients (Figure 7), while GPX3, PNMT, and POLR3H were significantly downregulated (Figure 7). These results were in accordance with those in the previous four datasets.

3.7 Validation of the key DEMRGs by qRT-PCR

According to the results of qRT-PCR, the expression levels of ENPP3, PDE7B, and ENTPD1 were elevated, while the expression levels of PNMT, GPX3, and POLR3H were decreased in the lung tissues of bleomycin-induced pulmonary fibrosis mice compared with the sham group (Figure 8). The results of qRT-PCR remained consistent with the bioinformatics analysis; therefore, these key DEMRGs may play an essential role in the progression of IPF.

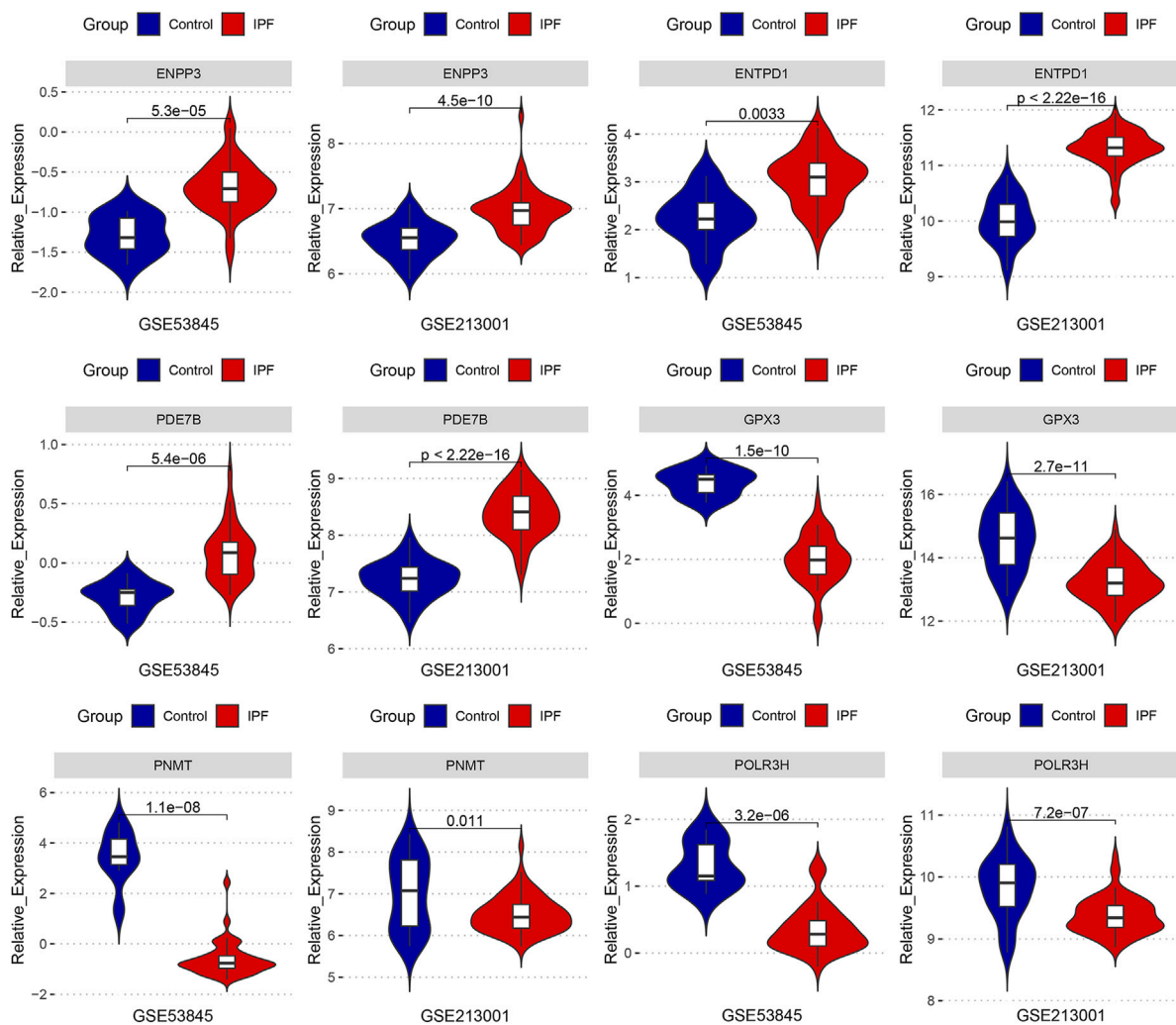


FIGURE 7

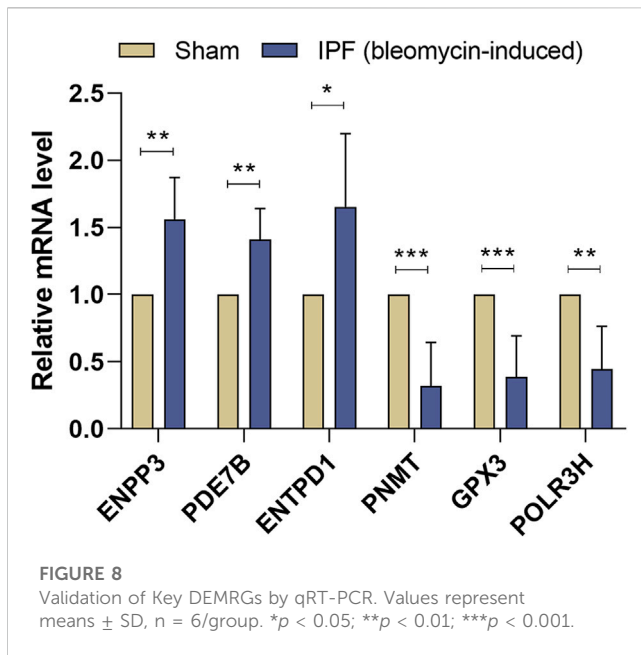
The expression levels of six key DEMRGs were validated in two independent external datasets (GSE53845 and GSE213001): ENPP3, ENTPD1, and PDE7B were significantly upregulated in IPF lung tissue ($p < 0.05$), while GPX3, PNMT, and POLR3H were significantly downregulated in IPF lung tissue ($p < 0.05$).

4 Discussion

Our study aims to identify key metabolism-related genes of IPF. First, we performed the differential analysis of the four GEO public datasets (GSE32537, GSE110147, GSE150910, and GSE92592) and integrated metabolism-related genes from the MSigDB dataset, resulting in 51 DEMRGs that were commonly upregulated and 50 DEMRGs that were commonly downregulated in the four IPF datasets. Subsequently, we performed functional enrichment analysis on these 101 DEMRGs, and the results indicated that these genes were involved in various metabolism-related terms. Then, two machine-learning algorithms were utilized to screen the key DEMRGs, resulting in six genes (ENPP3, ENTPD1, PDE7B, GPX3, PNMT, and POLR3H) as key DEMRGs. We further combined the miRNA expression profile dataset of IPF and the predicting results of the miRWalk database to construct the miRNA-mRNA network regulating the key DEMRGs. Next, we performed an immune infiltration analysis and identified an elevated M2 macrophage level

in IPF patients, which reflects the enhanced M2 polarization-mediated fibrosis phenotype. In addition, the mRNA expression of the key DEMRGs was validated in two external independent datasets (GSE53845 and GSE213001). Finally, the gene expression pattern was validated by qRT-PCR, demonstrating that the key DEMRGs might have potentially significant roles in IPF.

The immune cell infiltration results showed increased levels of M2 macrophages in the lung tissue of IPF patients. As the most abundant immune cells in the lung (approximately 70%), macrophages play a critical role in pulmonary fibrosis-related airway remodeling (Cai et al., 2014). Activated macrophages are usually divided into two categories, M1 macrophages (pro-inflammatory) and M2 macrophages (anti-inflammatory/pro-fibrotic) (Vasse et al., 2021). The ENPP3 and ENTPD1 encoded products can hydrolyze ATP. Thus the elevated levels of ENPP3 and ENTPD1 observed in our study lead to a decreased ATP level. Extracellular ATP increases the global inflammation level (Cauwels et al., 2014). Besides, we identified that the M2 macrophage expression was positively correlated with the



expression level of ENPP3. Taken together, we have reason to believe that ENPP3 and ENTPD1 may play a role in macrophages. The enhanced macrophage M2 polarization might be partly through the upregulation of ENPP3 and ENTPD1, leading to a decrease in ATP levels, which produces an anti-inflammatory and pro-fibrotic phenotype and ultimately exacerbates IPF. However, the specific mechanism needs to be validated in further studies.

GPX3 encodes glutathione peroxidase 3, which is expressed mainly in the lung and kidney (Lubos et al., 2011). Recent studies have shown a strong link between reactive oxygen species and fibrosis (Richter and Kietzmann, 2016). NADPH oxidase 4-derived ROS has been reported to regulate TGF- β 1-induced myofibroblast differentiation, extracellular matrix production, and contractility. A recent study uncovered a therapeutic effect of ROS-responsive liposomes in IPF, further suggesting the significance of anti-oxidative stress in IPF treatment (Liu et al., 2022). Our study shows that GPX3 expression levels are decreased in IPF lung tissue, which leads to increased levels of oxidative stress and thus exacerbates the fibrotic phenotype. Therefore, GPX3 is expected to be a potential novel target for the anti-oxidative stress treatment of IPF.

PDE7B encodes a phosphodiesterase that hydrolyzes cAMP and downregulates its signaling effects (Sasaki et al., 2000). In addition, the products of PNMT increase adrenaline production, and activation of adrenoceptors increases cAMP synthesis (Torphy, 1994; Martin et al., 2001). The decrease in cAMP results in a reduction of PKA activity and an increase in PFK activity, leading to increased F2,6BP levels. In response to the rise in F2,6BP levels, gluconeogenesis is suppressed, and glycolysis is stimulated (Pernicova and Korbonits, 2014). The increasing cellular cAMP level inhibits pulmonary fibroblast proliferation and collagen synthesis (Liu et al., 2004). In addition, glycolysis is increased early and sustainably during myofibroblast differentiation (Xie et al., 2015). The glucose transporter protein 1-dependent glycolytic phenotype was significantly increased in the lungs of aged mice, which was essential for pulmonary fibrosis (Cho et al., 2017). Actually, β -adrenergic agonists/cAMP play a key role in IPF, and β -adrenergic receptor agonists/cAMP have been shown to have

beneficial effects on alveolar injury, including protection from epithelial and endothelial cell damage, restoration of alveolar fluid clearance, and reduction of fibrotic remodeling (Sriram et al., 2021). Overall, the upregulation of PDE7B and downregulation of PNMT in the lung tissues of IPF patients identified in our study might conjointly result in decreased β 2-AR agonist/cAMP levels, decreased PKA activity, and enhanced glycolysis, which induced excessive collagen production and fibrosis formation.

The advantage of this study is that we have identified key metabolism-related genes that are commonly differentially expressed in IPF lung tissue using multiple bioinformatics approaches and validation in animal models. These genes may be a potential focus for future research on IPF metabolic disorders. However, several shortcomings of our study need to be acknowledged. First, the general profile of the IPF population cohort and the healthy control population cohort in the original dataset was not identical; for example, the mean age of the case group in the original study of GSE32537 was 62.6 years, whereas the mean age of the control group was 47.5 years. Therefore, it is unclear whether these differential gene expressions could be influenced by age. Nevertheless, our findings were obtained based on the analysis and validation of multiple datasets, thus minimizing the effect of potential confounding factors. The second limitation of this study is that although the identified key DEMRGs are commonly differentially expressed in IPF lung tissues, the specific degree of their impact on IPF needs to be clarified. Therefore, it will be important to interpret the findings with caution until they are validated by functional experimental research, despite the fact that they were based on reliable bioinformatics data.

5 Conclusion

Overall, through a comprehensive analysis of public datasets and experimental validation, we identified key metabolism-related genes that are differentially expressed in the lung tissue of IPF patients. Our study emphasizes the critical role of metabolic dysregulation in IPF, offers potential therapeutic targets, and provides new insights for future studies.

Data availability statement

Publicly available datasets were analyzed in this study. This data can be found here: Gene Expression Omnibus (GEO) database (<https://www.ncbi.nlm.nih.gov/geo/>; accession number: GSE32537, GSE110147, GSE150910, GSE92592, GSE32538, GSE53845, GSE213001).

Ethics statement

The animal study was reviewed and approved by the Laboratory Animal Welfare Ethics Committee of Central South University.

Author contributions

YZ designed the study, analyzed the data, and wrote the manuscript. JH and RG performed validation in animal models.

SC assisted in analyzing the data and revising the manuscript. HY and WO critically read and edited the manuscript.

Funding

This research was funded by the National Natural Science Foundation of China (81501128).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

- Bargagli, E., Refini, R. M., D'Alessandro, M., Bergantini, L., Cameli, P., Vantaggiato, L., et al. (2020). Metabolic dysregulation in idiopathic pulmonary fibrosis. *Int. J. Mol. Sci.* 21, 5663. doi:10.3390/ijms21165663
- Botta, M., Audano, M., Sahebkar, A., Sirtori, C. R., Mitro, N., and Ruscica, M. (2018). PPAR agonists and metabolic syndrome: An established role? *Int. J. Mol. Sci.* 19, 1197. doi:10.3390/ijms19041197
- Cai, Y., Sugimoto, C., Arainga, M., Alvarez, X., Didier, E. S., and Kuroda, M. J. (2014). *In vivo* characterization of alveolar and interstitial lung macrophages in rhesus macaques: Implications for understanding lung disease in humans. *J. Immunol.* 192, 2821. doi:10.4049/jimmunol.1302269
- Cauwels, A., Rogge, E., Vandendriessche, B., Shiva, S., and Brouckaert, P. (2014). Extracellular ATP drives systemic inflammation, tissue damage and mortality. *Cell Death Dis.* 5, e1102. doi:10.1038/cddis.2014.70
- Cho, S. J., Moon, J. S., Lee, C. M., Choi, A. M., and Stout-Delgado, H. W. (2017). Glucose transporter 1-dependent glycolysis is increased during aging-related lung fibrosis, and phloretin inhibits lung fibrosis. *Am. J. Respir. Cell Mol. Biol.* 56, 521. doi:10.1165/rcmb.2016-0225OC
- Clough, E., and Barrett, T. (2016). The gene expression omnibus database. *Methods Mol. Biol.*, 1418, 93. doi:10.1007/978-1-4939-3578-9_5
- Duan, K. B., Rajapakse, J. C., Wang, H., and Azuaje, F. (2005). Multiple SVM-RFE for gene selection in cancer classification with expression data. *IEEE Trans. Nanobioscience* 4, 228–234. doi:10.1109/tnb.2005.853657
- Edgar, R., Domrachev, M., and Lash, A. E. (2002). Gene expression omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res.* 30, 207–210. doi:10.1093/nar/30.1.207
- Friedman, J., Hastie, T., and Tibshirani, R. (2010). Regularization paths for generalized linear models via coordinate descent. *J. Stat. Softw.*, 33, 1. doi:10.18637/jss.v033.i01
- Huang Da, W., Sherman, B. T., and Lempicki, R. A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* 4, 44–57. doi:10.1038/nprot.2008.211
- Jiang, Y., Mao, C., Yang, R., Yan, B., Shi, Y., Liu, X., et al. (2017). EGLN1/c-Myc induced lymphoid-specific helicase inhibits ferroptosis through lipid metabolic gene expression changes. *Theranostics* 7, 3293–3305. doi:10.7150/thno.19988
- Kang, Y. P., Lee, S. B., Lee, J. M., Kim, H. M., Hong, J. Y., Lee, W. J., et al. (2016). Metabolic profiling regarding pathogenesis of idiopathic pulmonary fibrosis. *J. Proteome Res.* 15, 1717–1724. doi:10.1021/acs.jproteome.6b00156
- Kauppinen, A., Suuronen, T., Ojala, J., Kaarniranta, K., and Salminen, A. (2013). Antagonistic crosstalk between NF- κ B and SIRT1 in the regulation of inflammation and metabolic disorders. *Cell Signal* 25, 1939–1948. doi:10.1016/j.cellsig.2013.06.007
- Kern, F., Fehlmann, T., Solomon, J., Schwed, L., Grammes, N., Backes, C., et al. (2020). miEAA 2.0: integrating multi-species microRNA enrichment analysis and workflow management systems. *Nucleic Acids Res.* 48, W521–W528. doi:10.1093/nar/gkaa309
- Landi, C., Bargagli, E., Carleo, A., Bianchi, L., Gagliardi, A., Prasse, A., et al. (2014). A system biology study of BALF from patients affected by idiopathic pulmonary fibrosis (IPF) and healthy controls. *Proteomics Clin. Appl.* 8, 932–950. doi:10.1002/prca.201400001
- Ley, B., Brown, K. K., and Collard, H. R. (2014). Molecular biomarkers in idiopathic pulmonary fibrosis. *Am. J. Physiol. Lung Cell Mol. Physiol.* 307, L681–L691. doi:10.1152/ajplung.00014.2014
- Ley, B., Collard, H. R., and King, T. E., JR. (2011). Clinical course and prediction of survival in idiopathic pulmonary fibrosis. *Am. J. Respir. Crit. Care Med.* 183, 431–440. doi:10.1164/rccm.201006-0894CI
- Liberzon, A., Subramanian, A., Pinchback, R., Thorvaldsdottir, H., Tamayo, P., and Mesirov, J. P. (2011). Molecular signatures database (MSigDB) 3.0. *Bioinformatics* 27, 1739–1740. doi:10.1093/bioinformatics/btr260
- Liu, J., Wu, Z., Liu, Y., Zhan, Z., Yang, L., Wang, C., et al. (2022). ROS-responsive liposomes as an inhaled drug delivery nanopatform for idiopathic pulmonary fibrosis treatment via Nrf2 signaling. *J. Nanobiotechnology* 20, 213. doi:10.1186/s12951-022-01435-4
- Liu, X., Ostrom, R. S., and Insel, P. A. (2004). cAMP-elevating agents and adenylyl cyclase overexpression promote an antifibrotic phenotype in pulmonary fibroblasts. *Am. J. Physiol. Cell Physiol.* 286, C1089–C1099. doi:10.1152/ajpcell.00461.2003
- Lubos, E., Loscalzo, J., and Handy, D. E. (2011). Glutathione peroxidase-1 in health and disease: From molecular mechanisms to therapeutic opportunities. *Antioxid. Redox Signal* 15, 1957–1997. doi:10.1089/ars.2010.3586
- Maher, T. M., Bendstrup, E., Dron, L., Langley, J., Smith, G., Khalid, J. M., et al. (2021). Global incidence and prevalence of idiopathic pulmonary fibrosis. *Respir. Res.*, 22, 197. doi:10.1186/s12931-021-01791-z
- Mamazhakypov, A., Schermuly, R. T., Schaefer, L., and Wygrecka, M. (2019). Lipids - two sides of the same coin in lung fibrosis. *Cell Signal* 60, 65–80. doi:10.1016/j.cellsig.2019.04.007
- Martin, J. L., Begun, J., Mcleish, M. J., Caine, J. M., and Grunewald, G. L. (2001). Getting the adrenaline going: Crystal structure of the adrenaline-synthesizing enzyme PNMT. *Structure* 9, 977–985. doi:10.1016/s0969-2126(01)00662-1
- Moeller, A., Ask, K., Warburton, D., Gaudie, J., and Kolb, M. (2008). The bleomycin animal model: A useful tool to investigate treatment options for idiopathic pulmonary fibrosis? *Int. J. Biochem. Cell Biol.*, 40, 362. doi:10.1016/j.biocel.2007.08.011
- Newman, A. M., Liu, C. L., Green, M. R., Gentles, A. J., Feng, W., Xu, Y., et al. (2015). Robust enumeration of cell subsets from tissue expression profiles. *Nat. Methods* 12, 453–457. doi:10.1038/nmeth.3337
- Noble, P. W., Barkauskas, C. E., and Jiang, D. (2012). Pulmonary fibrosis: Patterns and perpetrators. *J. Clin. Invest.* 122, 2756–2762. doi:10.1172/JCI60323
- Pernicova, I., and Korbonits, M. (2014). Metformin--mode of action and clinical implications for diabetes and cancer. *Nat. Rev. Endocrinol.* 10, 143–156. doi:10.1038/nrendo.2013.256
- Raghu, G., Remy-Jardin, M., Myers, J. L., Richeldi, L., Ryerson, C. J., Lederer, D. J., et al. (2018). Diagnosis of idiopathic pulmonary fibrosis. An official ATS/ERS/RSR/ALAT clinical practice guideline. *Am. J. Respir. Crit. Care Med.*, 198, e44. doi:10.1164/rccm.201807-1255ST
- Richter, K., and Kietzmann, T. (2016). Reactive oxygen species and fibrosis: Further evidence of a significant liaison. *Cell Tissue Res.* 365, 591–605. doi:10.1007/s00441-016-2445-3
- Saito, S., Alkhatib, A., Kolls, J. K., Kondoh, Y., and Lasky, J. A. (2019). Pharmacotherapy and adjunctive treatment for idiopathic pulmonary fibrosis (IPF). *J. Thorac. Dis.* 11, S1740–S1754. doi:10.21037/jtd.2019.04.62
- Sasaki, T., Kotera, J., Yuasa, K., and Omori, K. (2000). Identification of human PDE7B, a cAMP-specific phosphodiesterase. *Biochem. Biophys. Res. Commun.* 271, 575–583. doi:10.1006/bbrc.2000.2661

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2023.1058582/full#supplementary-material>

- Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., et al. (2003). Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Res.* 13, 2498–2504. doi:10.1101/gr.1239303
- Shen, W., Song, Z., Zhong, X., Huang, M., Shen, D., Gao, P., et al. (2022). Sangerbox: A comprehensive, interaction-friendly clinical bioinformatics analysis platform. *iMeta*, 1, e36, doi:10.1002/imt2.36
- Shenderov, K., Collins, S. L., Powell, J. D., and Horton, M. R. (2021). Immune dysregulation as a driver of idiopathic pulmonary fibrosis. *J. Clin. Invest.* 131, e143226. doi:10.1172/JCI143226
- Sherman, B. T., Hao, M., Qiu, J., Jiao, X., Baseler, M. W., Lane, H. C., et al. (2022). David: A web server for functional enrichment analysis and functional annotation of gene lists (2021 update). *Nucleic Acids Res.* 50, W216–W221. doi:10.1093/nar/gkac194
- Sriram, K., Insel, M. B., and Insel, P. A. (2021). Inhaled beta2 adrenergic agonists and other cAMP-elevating agents: Therapeutics for alveolar injury and acute respiratory disease syndrome? *Pharmacol. Rev.* 73, 488–526. doi:10.1124/pharmrev.121.000356
- Sticht, C., De La Torre, C., Parveen, A., and Gretz, N. (2018). miRWalk: An online resource for prediction of microRNA binding sites. *PLoS One* 13, e0206239. doi:10.1371/journal.pone.0206239
- Torphy, T. J. (1994). Beta-adrenoceptors, cAMP and airway smooth muscle relaxation: Challenges to the dogma. *Trends Pharmacol. Sci.* 15, 370–374. doi:10.1016/0165-6147(94)90157-0
- Vasse, G. F., Nizamoglu, M., Heijink, I. H., Schleputz, M., Van Rijn, P., Thomas, M. J., et al. (2021). Macrophage-stroma interactions in fibrosis: Biochemical, biophysical, and cellular perspectives. *J. Pathol.* 254, 344–357. doi:10.1002/path.5632
- Wickham, H. (2016). “Data analysis,” in *ggplot2: Elegant graphics for data analysis*. Editor H. WICKHAM (Cham: Springer International Publishing).
- Wynn, T. A., and Vannella, K. M. (2016). Macrophages in tissue repair, regeneration, and fibrosis. *Immunity* 44, 450–462. doi:10.1016/j.immuni.2016.02.015
- Xie, N., Tan, Z., Banerjee, S., Cui, H., Ge, J., Liu, R. M., et al. (2015). Glycolytic reprogramming in myofibroblast differentiation and lung fibrosis. *Am. J. Respir. Crit. Care Med.*, 192, 1462, doi:10.1164/rccm.201504-0780OC