

Identification and Experimental Validation of Biomarkers Related to MiR-125a-5p in Chronic Obstructive Pulmonary Disease

Xia Jing, Yueqin Li

Department of General Medical, Shanxi Bethune Hospital, Shanxi Academy of Medical Sciences, Tongji Shanxi Hospital, Third Hospital of Shanxi Medical University, Taiyuan, Shanxi, People's Republic of China

Correspondence: Yueqin Li, Department of General Medical, Shanxi Bethune Hospital, Shanxi Academy of Medical Sciences, Tongji Shanxi Hospital, Third Hospital of Shanxi Medical University, Taiyuan, Shanxi, 030000, People's Republic of China, Email 13934571745@163.com

Purpose: The miR-125a-5p has been reported influence the development of lung cancer, however, the link between it and chronic obstructive pulmonary disease (COPD) is still not well understood. Hence, this study was designed to investigate the molecular pathway by which miR-125a-5p related biomarkers were involved in COPD.

Patients and Methods: The differentially expressed genes (DEGs) and module genes related to COPD in GSE100153 were screened out by differential analysis and weighted gene co-expression network analysis, respectively. Then, the target genes of miR-125a-5p obtained from miRWalk database were intersected with DEGs and module genes, followed by identification of biomarkers through SVM-RFE algorithms. Moreover, the gene set enrichment analysis, immune infiltration analysis, construction of regulatory network, single-cell analysis and Mendelian randomization (MR) analysis were performed. At last, the expression levels of the biomarkers were further validated in GSE100153 and GSE146560 as well as in qRT-PCR.

Results: A total of 10 genes were acquired by intersecting the 126 DEGs, the 3989 module genes, and 2329 target genes, of which PITHD1, CNTNAP2 and GUCD1 were identified as biomarkers. Enrichment analysis showed their roles in various cellular functions. In addition, significant associations were identified between 9 distinct cells and biomarkers. Subsequently, 5 TFs and 63 therapeutic agents were predicted as biomarkers. Moreover, GUCD1 and PITHD1 were significantly different between case and control in T cells and Alveolar cells. In COPD, GUCD1 and PITHD1 were significantly down-regulated in GSE100153 and GSE146560 datasets and confirmed by qRT-PCR.

Conclusion: In our study, PITHD1, CNTNAP2, and GUCD1 were recognized as biomarkers related to miR-125a-5p-related genes in COPD, providing new references for treatment of COPD.

Keywords: drug prediction, immune infiltration analysis, MiR-125a-5p, single-cell analysis

Introduction

Chronic obstructive pulmonary disease (COPD) is a heterogeneous lung disorder marked by persistent respiratory symptoms such as dyspnea, cough, and sputum production. It results from airway (bronchitis, bronchiolitis) and alveolar (emphysema) abnormalities, leading to ongoing and often worsening airflow blockage.¹ COPD ranks among the top three global mortality causes, with 90% of these fatalities occurring in low- and middle-income countries.^{2,3} COPD may result from gene-environment interactions that impair lung function or disrupt their typical development and aging processes.⁴ Primary environmental risk factors include tobacco smoke and exposure to harmful particles and gases from both household and outdoor air pollution. Additionally, factors like atypical lung development and rapid aging of the lungs may contribute.⁴⁻⁶ The genetic risk factors identified to date that are most associated with COPD (although epidemiologically rare) mutations was SERPINA1 gene, which cause α 1-antitrypsin deficiency, but other genetic variants with smaller individual effects, such as genomic regions near FAM13A, RIN3, CYP2A6, and DSP, are also linked with reduced lung function and COPD risk.⁷ Currently, the diagnosis of COPD mainly relies on spirometry to identify airflow

obstructions that are non-fully reversible. However, this method is full of limitations, such as it is difficult to diagnose early detection of COPD.⁸ Therefore, the identification of novel biomarkers and exploration of their potential molecular mechanisms will help to parse the pathogenesis of COPD, provide a basis for early clinical diagnosis and a new target for treatment.

MiRNAs (miRNAs) are tiny non-coding RNA strands, about 22 nucleotides long, that control gene expression post-transcriptionally either by degrading or by attaching to target mRNAs.⁹ The miR-125a is localized to chromosome 19q13,¹⁰ and miR-125a-5p is its mature body. A number of studies have shown that miR-125a has an inhibitory effect in a variety of tumors.¹¹ Recent studies have also highlighted the involvement of miR-125a-5p in autoimmune diseases, suggesting its role in modulating immune responses.¹² For instance, our previous research demonstrated that miR-125a is highly expressed in human bronchial epithelial cells exposed to cigarette smoke extract. Inhibition of this miRNA was found to reduce both apoptosis and inflammation in these cells, underscoring its importance in inflammatory responses.¹³ Furthermore, Wang et al collected 141 local patients with COPD, isolated peripheral leukocytes, and analyzed them by qRT-PCR, showing that the level of miR-125-5p increased in all COPD groups, and gradually decreased in groups A, B, and C, but increased in group D compared with group C.¹⁴ This suggests that miR-125a-5p may play a dynamic role throughout the progression of COPD. Additionally, another study indicated that in COPD patients and mice exposed to cigarette smoke who were infected with influenza A viruses, the upregulation of miR-125a resulted in decreased expression of A20 and mitochondrial antiviral signaling pathways. This led to exaggerated inflammation and impaired antiviral responses, further supporting the hypothesis that miR-125a plays a crucial role in regulating inflammatory processes.¹⁵ Given these findings, exploring the mechanisms by which miR-125a-5p-related genes influence COPD pathogenesis could provide valuable insights for clinical diagnosis and treatment strategies for this disease.

Mendelian randomisation (MR) functions similarly to a natural experiment, using genetic variation as an instrumental variable to determine the causal link between exposure and outcome. This method is effective in reducing the effects of possible confounders and reverse causality. MR analyses using pooled data can identify potential correlations.¹⁶ In this study, biomarkers linked with miR-125a-5p were identified using COPD transcriptome data from the Gene Expression Omnibus (GEO) public database. Single-cell analysis was conducted to investigate the potential mechanisms of these biomarkers at the single-cell level. Subsequently, MR analysis was employed to explore the causal relationship between the identified biomarkers and COPD, which is crucial for further unraveling the pathogenesis of the disease.

Materials and Methods

Data Source and Study Design

The GSE100153 as a training set, containing 19 COPD and 24 control samples,¹⁷ and the GSE146560 as a validation set, including 8 COPD and 8 control samples,¹⁸ the scRNA-seq dataset, GSE227691, included samples from 8 COPD patients and 4 normal lung tissues, were mined from the GEO database (<https://www.ncbi.nlm.nih.gov/geo/>). Besides, the related genes targeted miR-125a-5p were predicted via the miRWalk database (<http://mirwalk.umm.uni-heidelberg.de/>),¹⁹ and a sum of 2329 target genes of miR-125a-5p were obtained with “Score>0.95, Position=3UTR” as the filtering condition. Additionally, genome-wide association study (GWAS) data for COPD and expression quantitative trait loci (eQTL) data for exposure factors, used in Mendelian Randomization (MR) analyses, were acquired from the Integrated Epidemiology Unit’s (IEU) Open GWAS database (<https://gwas.mrcieu.ac.uk/>). The dataset finn-b-COPD_OPPORTUNIST_INFECTIONS comprised 186,957 samples (ncase:ncontrol=234:186,723) and 16,380,363 single-nucleotide polymorphisms (SNPs). All samples are of European descent. [Figure 1](#) shows the flow chart of this study.

Weighted Gene Co-Expression Network Analysis (WGCNA)

In the training set, the module genes associated with COPD were screened by WGCNA.²⁰ First, the samples were clustered and outliers removed. To best match a scale-free topology, we established an optimal soft threshold. Furthermore, we defined the minimum number of genes required for each module at 150, following the hybrid dynamic tree cutting method. Finally, Pearson analysis was conducted to find the gene modules strongly associated with COPD, in which the genes were as module genes.

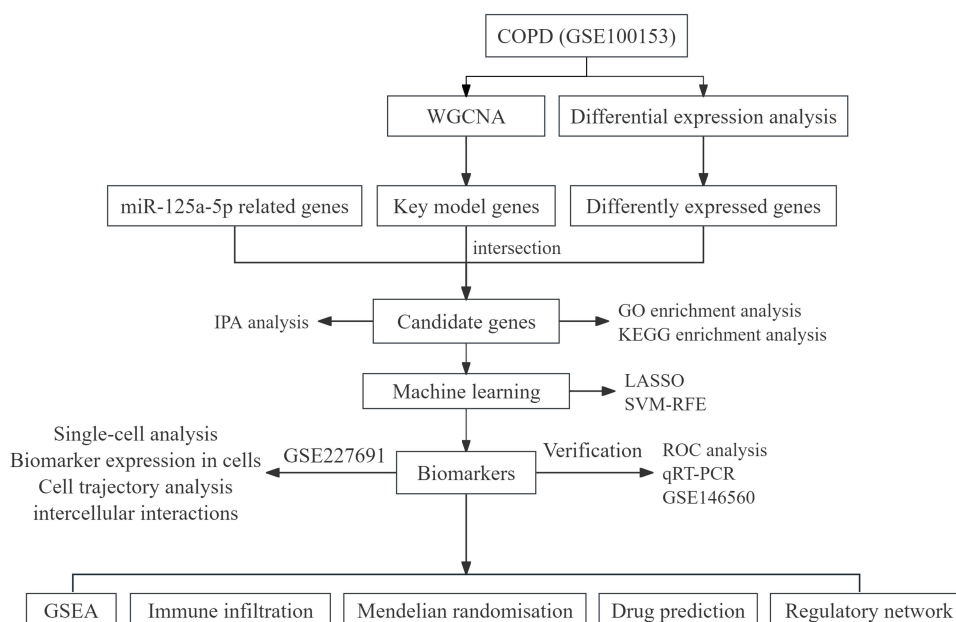


Figure 1 Flowchart of this study.

Identification and Functional Enrichment Analysis of Candidate Genes

The limma package was performed to compare the differences in gene expression levels between COPD and control samples in GSE100153, thereby screening for differentially expressed genes (DEGs) with the screening criteria of $|\log_2FC| > 0.5$ and $P < 0.05$.²¹ Volcano plot was created using ggplot2, while heat map was generated with pheatmap.²² Subsequently, candidate genes were obtained by overlapping DEGs, module genes, and target genes using the eulerr package. The GO and KEGG enrichment analyses were executed via the clusterProfiler package,²³ so as to look for common functions and related pathways between the candidate genes ($P < 0.05$).

Screening Biomarkers by Machine Learning Algorithms

Candidate genes were applied for Lasso regression analysis and SVM-RFE analysis using the glmnet package and e1071 package, respectively,²⁴ where the feature genes were overlapped to identify key feature genes. ROC curves were plotted for key feature genes in the GSE100153 and GSE146560 datasets using the pROC package, and genes with AUC values greater than 0.7 in both datasets were defined as biomarkers.²⁵ By the way, the expression of biomarkers in the control and COPD groups in the training and validation sets was analysed by Wilcoxon test.²⁶

Functional Enrichment of Biomarkers

To further understand the classical pathway, and function analyses of DEGs were first performed using QIAGEN (www.qiagen.com/ingenuity). Then, the upstream regulators, downstream target genes and possible pathways of action of the biomarkers were analysed using ingenuity pathway analysis (IPA).²⁷ At last, Spearman correlation analyses between each biomarker and the all genes were performed separately using the psych in R package,²⁸ and gene set enrichment analysis (GSEA) was conducted via clusterProfiler package based on the KEGG background gene set in MSigDB database.²⁹

Immune Infiltration Analysis

Based on the immune-related genes obtained as background gene sets from previous studies,³⁰ the cluster analysis of infiltrating immune cells in control and COPD samples was conducted via Pheatmap package in R.³¹ Then, the Wilcoxon test was conducted to monitor differences in enrichment scores for each immune cell.²⁶ Where after, based on the set of genes for 13 immune-related functions obtained from previous studies,³² the ssGSEA algorithm was employed to

measure immune-related function scores for the COPD and control samples in the training set and to analyse the differences in each immune-related function using the Wilcoxon test.³³ At last, correlation analyses between differential immune cells, between biomarkers and differential immune cells, and between biomarkers and differential immune-related functions were performed by Spearman.²⁸

Construction of Gene-Gene Interaction Network (GGI) Network and Competing Endogenous RNA (CeRNA) Network

To search for key genes associated with biomarkers, the GGI network of biomarkers was constructed based on GeneMANIA (<https://genemania.org/>), in which the top 20 genes and the top 7 pathways were visualized.³⁴ Next, in order to further understand the gene expression regulation pattern, the corresponding lncRNAs of miR-125a-5p were predicted by the Starbase database with the filtering criterion of “clipExpNum>10”,³⁵ and the Cytoscape software was used for visualisation of the ceRNA network of lncRNA-miRNA-mRNA.³⁶

Analysis of Biomarkers for Transcription Factors (TFs), Drugs and Related Diseases

TFs targeting biomarkers was screened out with “library=ENCODE, FET p-value<0.05” as the screening condition based on the ChEA3 database (<https://maayanlab.cloud/chea3/>).³⁷ Then, DNA binding sites were predicted for the top three TFs based on the JASPAR database (<https://jaspar.genereg.net/>).³⁸ In addition, biomarker target therapeutic drugs were predicted via the CTD database (<http://ctdbase.org/>)³⁹ and the results were visualised by Cytoscape software. Finally, diseases associated with biomarkers were analyzed via the disgenet2r package based on the DisGeNET database, and the results were similarly visualized using Cytoscape.^{36,40}

Single-Cell Sequencing Analysis

The R package Seurat (v 4.1.0)⁴¹ systematically processed and analyzed single-cell transcriptome sequencing data. First, we constructed Seurat objects and filtered out data that appeared in fewer than 3 cells or detected fewer than 200 genes per cell. Subsequently, quality control was performed to filter out cells with nFeature_RNA less than 200 or more than 4000, nCount_RNA less than 200 or more than 20,000, and percent.mt higher than 20%. Then, the FindVariableFeatures function was used to select the top 2000 highly variable genes. The JackStrawPlot function was used to identify the principal components for analysis. After completing the principal components analysis (PCA) dimensionality reduction, we used the Seurat (v 4.1.0) package to cluster the cells using the uniform manifold approximation and projection (UMAP) method. Then different cell types were annotated utilizing CellMarker database. Finally, the key cells were obtained by Wilcoxon test comparing the differences in biomarker expression in annotated cells.

Analysis of Cell Trajectory Analysis and Intercellular Interactions

To explore the developmental trajectories of key cells, simulation analysis was executed using the R package Monocle2 (v 2.18.0).⁴² Finally, cell communication between the key cells and other annotated cell types was analyzed using the R package CellChat (v 1.5.0).⁴³

MR Analysis

To further investigate the causal relationship between biomarkers and COPD, we performed MR analysis. Instrumental variables (IVs) used for MR analysis were required to fulfill the following three assumptions. 1) IVs must be strongly associated with exposure, 2) IVs must be independent of confounders, 3) associations between IVs and outcomes are mediated only by exposure factors. Exposure factors were read through extract_instruments in the R-package TwoSampleMR (v 0.5.7).⁴⁴ In order to screen IVs significantly related to exposure factors, the SNPs with linkage disequilibrium (LD) were finally removed under the conditions of $p < 5 \times 10^{-5}$, clump = TRUE, r2 = 0.01 and kb = 10. The F-statistic for all IVs was then calculated to exclude weak IVs using $F = \frac{N-k-1}{k} \times \frac{R^2}{1-R^2}$ ($F > 10$).

Next, the R package two-sample MR function `harmonise_data` was used to standardize effect alleles and effect sizes, aligning them with the exposure factor-IV-outcome relationship. And MR analysis was performed by `mr` function combined with 5 algorithms (MR Egger, Weighted median, Inverse variance weighted (IVW), Simple mode, Weighted mode). In addition, scatter plot, forest plot and funnel plot were drawn to display MR results directly.

Sensitivity Analysis

To assess the robustness of the MR analysis results, we performed a sensitivity analysis that included a heterogeneity test, a horizontal pleiotropy test, and a leave-one-out analysis. The heterogeneity was evaluated using Cochran's Q test to determine if there was significant variation among the samples ($P > 0.05$). We also used the MR-Egger intercept to check for the presence of horizontal pleiotropy ($P > 0.05$). Finally, we conducted a leave-one-out analysis, sequentially omitting each SNP to reassess the effects of the remaining SNPs on the outcome variable.

Quantitative Real-Time Polymerase Chain Reaction (qRT-PCR)

In this study, a total of 8 pairs blood samples from Shanxi Bethune Hospital were utilized qRT-PCR, including 8 COPD samples and 8 control samples. Basic patient characteristics are shown in [Supplementary Table 1](#). The study was approved by the ethics committee of Shanxi Bethune Hospital (YXLL-2023-148). All patients had signed an informed consent form and the study complied with the Declaration of Helsinki.

First, the total RNA of the frozen COPD and control samples was extracted a by TRIzol reagent, after which RNA detection was performed based on the NanoPhotometer N50. Then, equal volumes of mRNA were reverse transcribed using the SureScript First-Strand cDNA Synthesis Kit, and the resulting cDNA was diluted between 5 to 20 times with RNase/DNase-free ddH₂O. Next, the qPCR reaction was performed with 2 x Universal Blue SYBR Green qPCR Master Mix kit according to the following reaction system: predenaturation at 95°C for 1 min; Denatured 20s at 95°C, annealed 20s at 55°C, extended 30s at 72°C; 40 cycles. Finally, gene expression level was quantified using the $2^{-\Delta\Delta CT}$ method, with primer sequences detailed in [Supplementary Table 2](#).

Statistical Analysis

Bioinformatics analyses were performed utilizing the R programming language (v 4.2.2). Wilcoxon rank sum test were used to compare the differences between two groups. Fisher's exact test was used to calculate the significance p-value. $P < 0.05$ was considered statistically significant.

Results

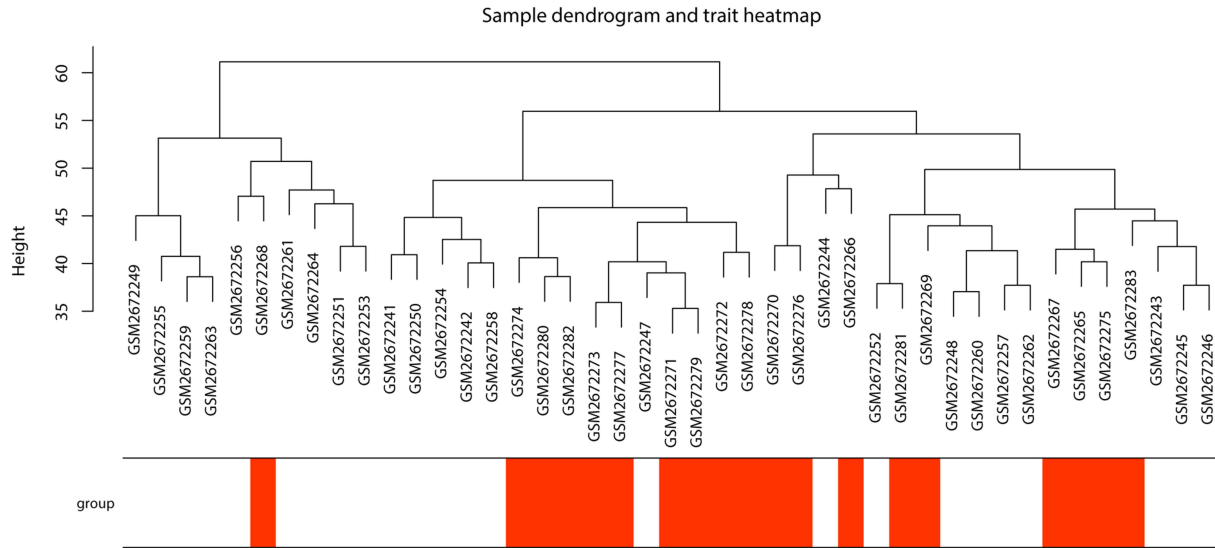
A Total of 3989 Modular Genes were Identified by WGCNA

This study was designed to identify module genes using the WGCNA. Firstly, cluster analysis of the samples revealed no outliers in this dataset ([Figure 2A](#)). When the soft threshold was 6, to ensure that the plot was as close as possible to a scale-free distribution, the R² approached the 0.9 threshold (red line) ([Figure 2B](#)). Then, a total of 12 gene modules were obtained ([Figure 2C](#)), of which the brown module ($\text{cor} = -0.51$, $p = 5 \times 10^{-4}$) and black were selected as key modules ($\text{cor} = 0.54$, $p = 2 \times 10^{-4}$) ([Figure 2D](#)). The brown module contained 2668 genes and the black module contained 1321 genes, which were identified as module genes.

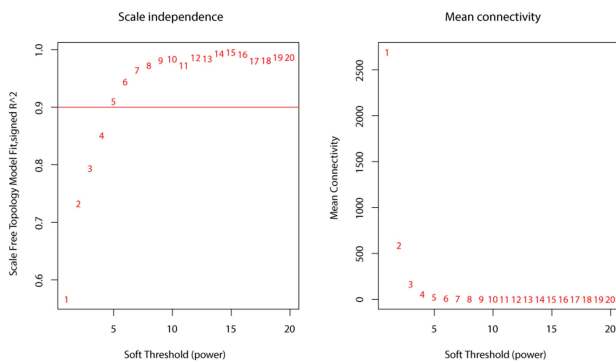
Identification and Functional Enrichment Analysis of Candidate Genes

First, 126 DEGs were selected in GSE100153, of which 17 were up-regulated and 109 down-regulated ([Figure 3A and B](#)). Then, 10 candidate genes were generated from intersections between DEGs, target genes and module genes ([Figure 3C](#)). GO enrichment analysis revealed enrichment of candidate genes mainly related to positive regulation of myeloid cell differentiation, platelet α -granule membrane, etc ([Figure 3D](#)). In regard to KEGG enrichment analysis, these genes played important roles in bladder cancer, melanoma, etc ([Figure 3E](#)).

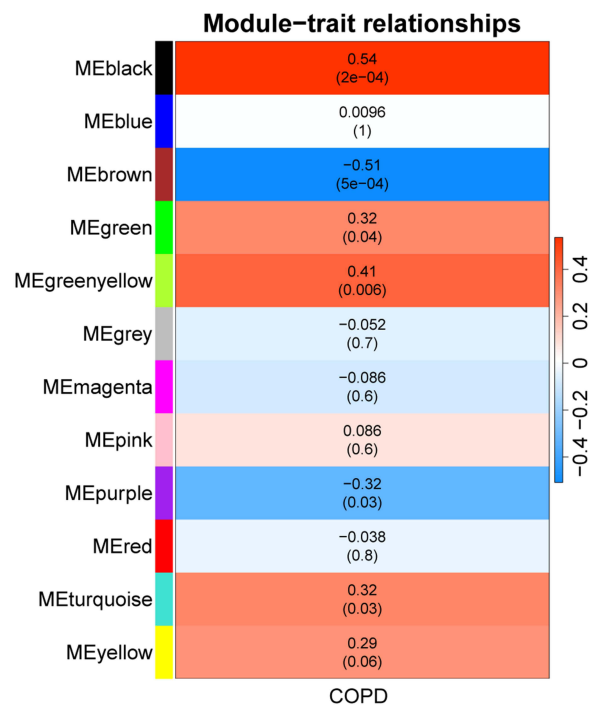
A



B



D



C

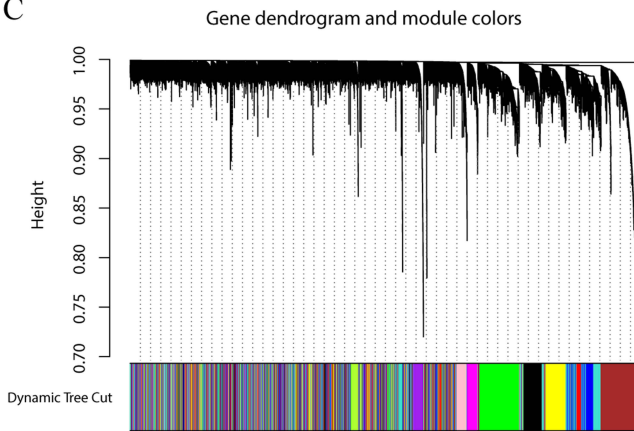


Figure 2 WGCNA identification analysis. (A) Sample clustering tree. (B) Soft threshold filtering. (C) Hierarchical clustering tree. (D) Module-trait correlation heat map.

PITHDI, CNTNAP2 and GUCD1 were Identified as Biomarkers Associated with COPD

First of all, LASSO regression analysis yielded 8 feature genes, including NFIX, GUCD1, FAM210B, CLEC12A, PITHDI, RNF10, SPARC, and CNTNAP2 (Figure 4A and B). Next, 3 feature genes (PITHDI, CNTNAP2, and GUCD1) were also

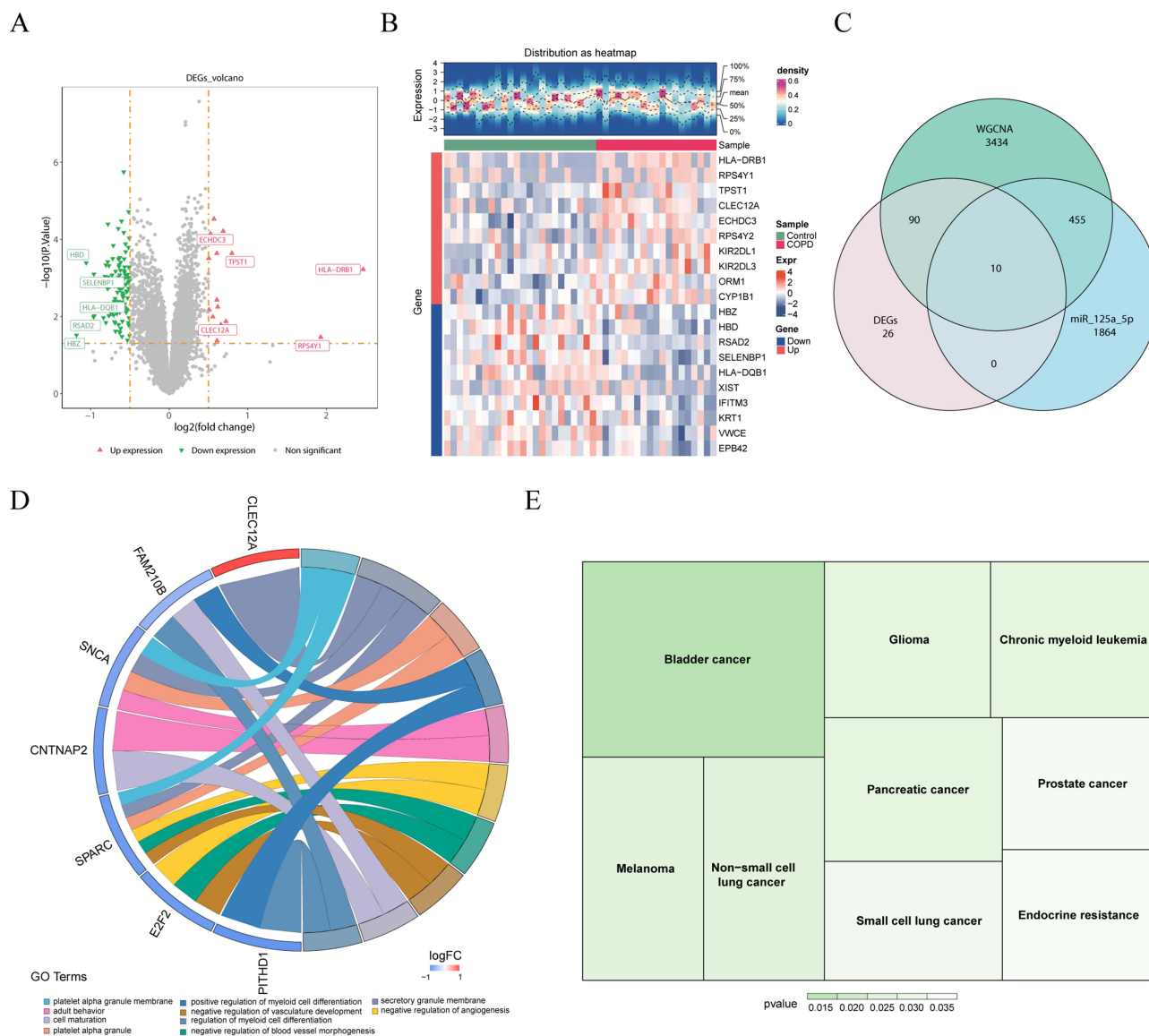


Figure 3 Identification and functional enrichment analysis of candidate genes. (A) Volcano map of gene differential expression between COPD and control group. (B) Heatmap of gene differential expression between COPD and control group. (C) Genes with the same trend of difference. (D) GO annotation results. (E) KEGG enrichment results.

obtained by SVM-RFE (Figure 4C). Three key feature genes, PITHD1, CNTNAP2 and GUCD1, were identified at the intersection of feature genes derived from two different machine learning models (Figure 4D), and their AUC values in both the training and testing sets were greater than 0.7. Therefore, they were defined as biomarkers (Figure 4E and F). Finally, in GSE100153 and the validation set GSE146560, the Wilcoxon test showed that the expression trend of the three biomarkers was consistent (Figure 4G and H).

Biomarkers were Mainly Concentrated in Ribosome, Proteasome, Endocytosis

A total of 256 pathways and 46 major disease modules associated with the DEGs were identified by IPA analysis. The top 10 pathways were illustrated in Figure 5A, eg interferon signaling, antigen presentation pathway and so on, and the top 12 disease modules could be seen in Figure 5B, including immunological disease, inflammatory disease, etc. In addition, the activation and inhibition of 46 disease modules can be seen in Figure 5C. Analysis of

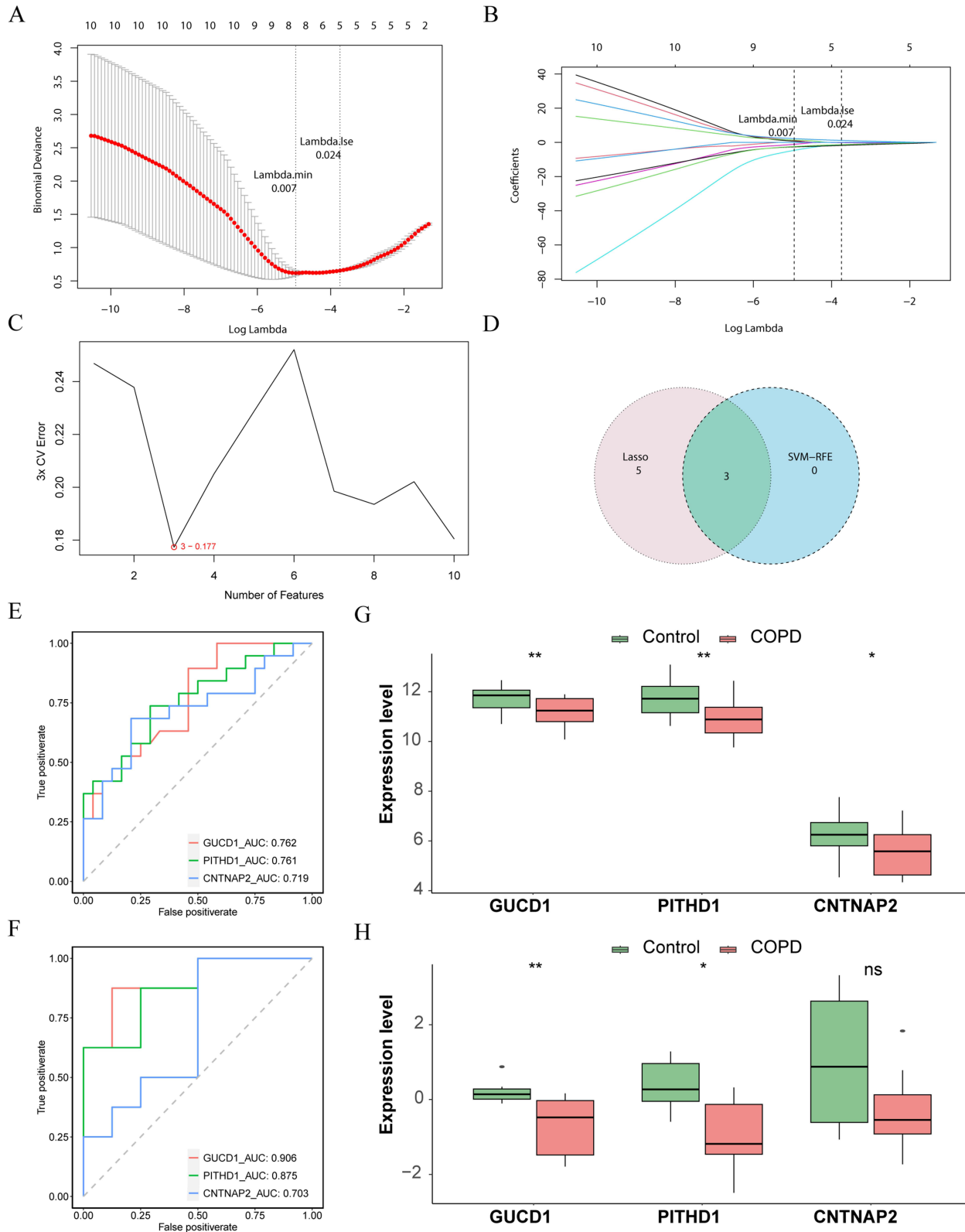


Figure 4 Screening of biomarkers. (A) The variation characteristics of the coefficient of variables. (B) The selection process of the optimum value of the parameter λ in the Lasso regression model by cross-validation method. (C) Screening characteristic genes by SVM-RFE algorithm. (D) Genes with the same trend of difference. (E and F) Training and testing sets analysis. (G and H) Differential expression between COPD and control group. *indicates $p < 0.05$; **indicates $p < 0.01$.

the upstream regulators, downstream target genes and the possible pathways among three biomarkers indicated that PITHD1 might indirectly inhibit PSMC2, and CNTNAP2 might be indirectly inhibited by SAFB and SAFB2 (Figure 5D). In GSEA pathway enrichment analysis, the main signaling pathways of PITHD1 included ribosome, spliceosome, etc (Figure 5E). GUCD1 was involved in biological processes such as ribosome, proteasome, etc (Figure 5F). CNTNAP2 played an important role in endocytosis, proteasome and so on (Figure 5G).

The Biomarkers were Significantly Correlated with CD56dim Natural Killer Cell, MDSC, Monocyte and Other Immune Cells

Firstly, heat maps elucidated abundance of immune cell infiltration in training set (Figure 6A). There were immune cells, including activated B cells and CD56dim natural killer cells, etc., had significantly different enrichment scores between the COPD and control groups (Figure 6B). Correlation analysis revealed that immature B cells had a strong positive correlation with activated B cells ($r = 0.720$, $p < 0.0001$), and a strong negative correlation with monocytes in differential immune cells ($r = -0.410$, $p < 0.001$) (Figure 6C). Besides, PITHD1 was significantly associated with three types of immune cells: CD56dim natural killer cell, Myeloid-derived suppressor cell (MDSC) and Monocyte (Figures 6D and S1). There were significant correlations between GUCD1 and three immune cells: MDSC, CD56dim natural killer cell, and Monocyte (Figures 6E and S2), while CNTNAP2 was significantly correlated with Activated B cell, Immature B cell, and Type 17 T helper cell (Figures 6F and S3).

Moreover, Supplementary Figure S4A illustrated marked differences in the enrichment scores for two types of immune functions between the COPD and control groups, specifically in Type-I and Type-II interferon responses. Spearman analysis indicated there was a positive correlation relationship between these two immune-related functions ($r = 0.44$, $p = 0.0036$) (Supplementary Figure S4B). Finally, two biomarkers were significantly positively correlated with two differential immune-related functions, among which GUCD1 was significantly correlated with Type-II-IFN-Response, and PITHD1 was significantly correlated with Type-I-IFN-Response and Type-II-IFN-Response (Supplementary Figure S4C and D).

Regulatory Networks of Biomarkers were Constructed

By constructing GGI networks, key genes and associated pathways interacting with biomarkers were predicted. The key genes included CNTN2, CPE, etc., and the related pathways mainly included main axon, cell-cell junction assembly, etc (Figure 7A). Next, 12 lncRNAs targeting miR-125a-5p were predicted based on the Starbase database, which including AC007780.1, AL162258.1, AC018628.1, etc., (Figure 7B). Besides, a total of 5 TFs were identified, namely TAL1, GABPA, FOSL1, ETS1 and NFE2. Among them, 2 binding sites were predicted for both TAL1 and GABPA, and 7 binding sites were predicted for FOSL1 (Table 1 and Figure 7C, D). In order to explore the diseases associated with the biomarkers, 30 diseases associated with CNTNAP2, which contained Epilepsy, Language Delay, Bipolar Disorder and so on, and 2 diseases (Malignant neoplasm of Prostate and Prostatic Neoplasms) associated with GUCD1 were predicted in this study (Figure 7E). Finally, a total of 63 therapeutic drugs were predicted for 3 biomarkers based on the CTD database, such as Valproic Acid, Bisphenol A, Acetaminophen and so on. This result was visualized via Cytoscape and the results showed 74 interacting relationship pairs, including PITHD1-Copper Sulfate, GUCD1-Acrolein, CNTNAP2-Atrazine, etc (Figure 7F).

Totally Six Cell Clusters were Annotated

Through quality control, a total of 20,738 genes and 37,190 cells were acquired for subsequent analysis (Supplementary Figure S5A). After standard data processing, we selected 2000 highly variable genes. The top 10 genes exhibiting the most pronounced intercellular expression changes were SLPI, CCL18, TPSD2, FABP4, SFTPA1, SFTPA2, SCGB3A1, SCGB3A2, SFTPC, and SCGB1A1 (Supplementary Figure S5B). And the linearly optimal dimensionality value for cell clustering was 30 (Supplementary Figure S6A). Then, a sum of 11 different cell clusters were identified (Supplementary Figure S6B), and only 6 cell clusters were annotated, namely Epithelial cells, B cells, Macrophage, Alveolar cells, Neutrophils, and T cells (Supplementary Figure S6C). The bubble plot demonstrated that the marker

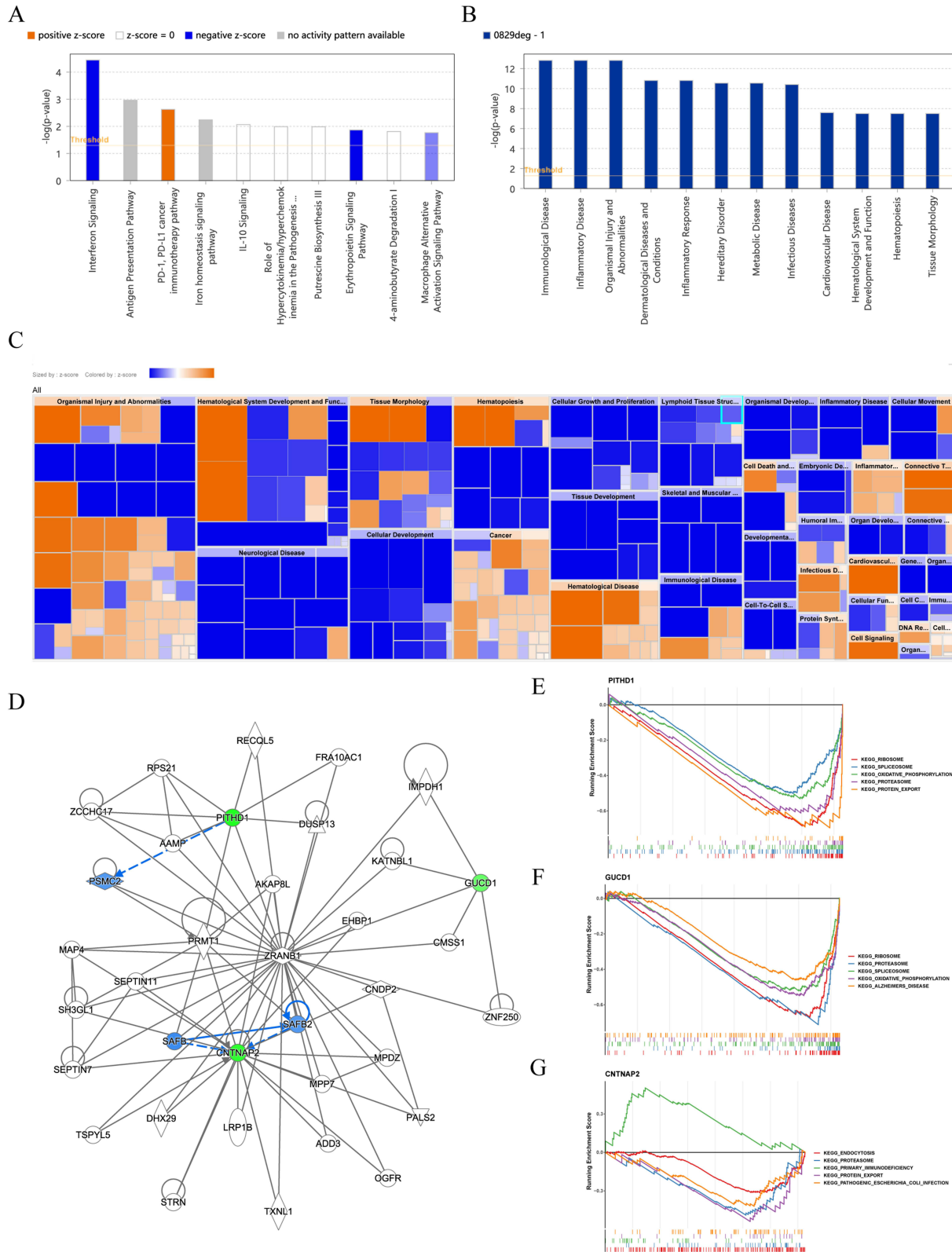


Figure 5 IPA analysis. **(A)** Identification of the top 10 pathways associated with DEGs. **(B)** Identification of top 12 disease modules associated with DEGs. **(C)** Heatmap of activation inhibition of disease modules. **(D)** IPA network analysis. **(E–G)** KEGG running Enrichment Score.

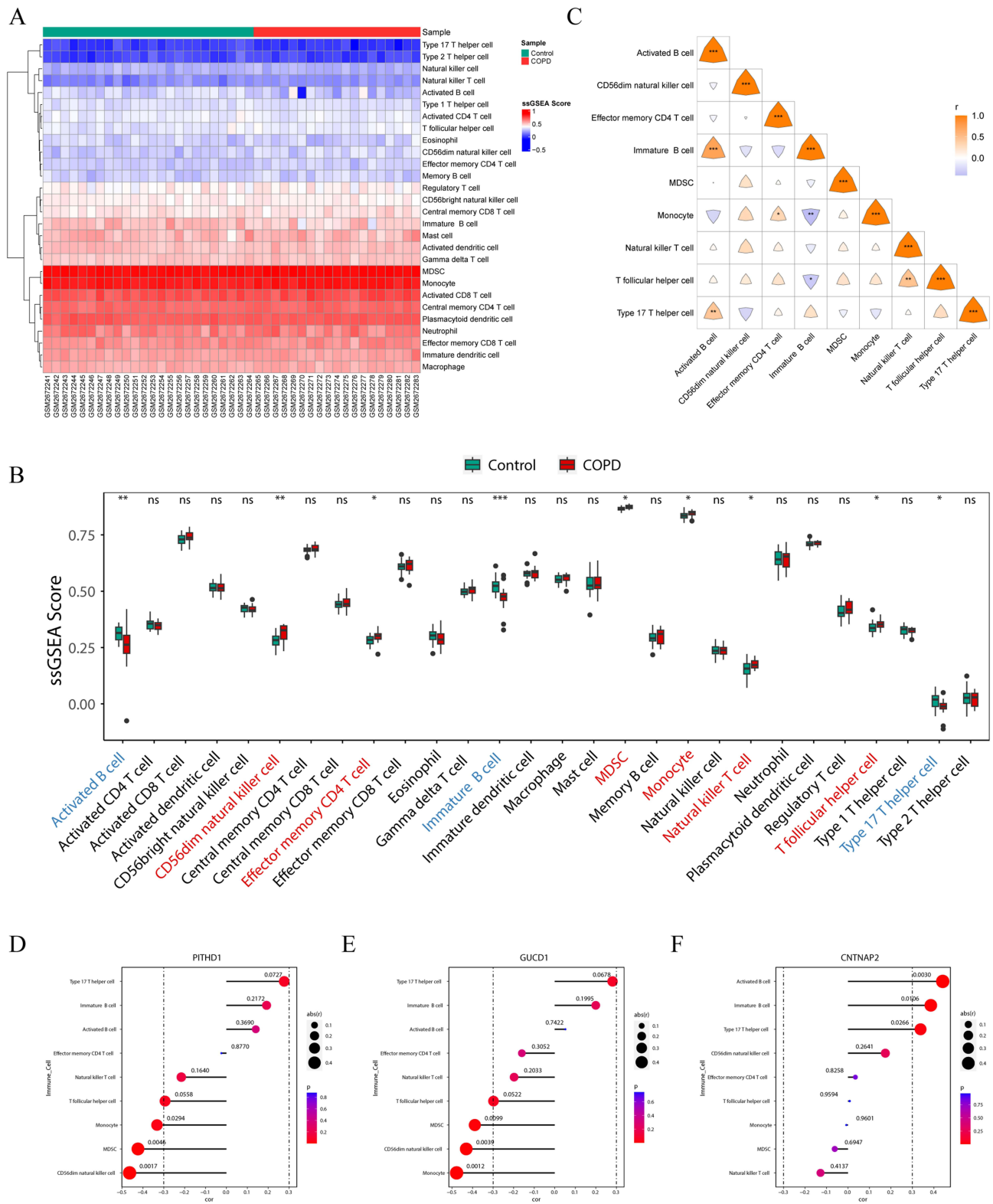


Figure 6 Immune infiltration analysis. **(A)** ssGSEA scores of different immune cells in control and COPD groups. **(B)** Differential analysis of ssGSEA scores in control and COPD groups. **(C)** Correlation analysis of different immune cells. **(D)** Correlation analysis of PITHD1 with different immune cells. **(E)** Correlation analysis of GUCD1 with different immune cells. **(F)** Correlation analysis of CNTNAP2 with different immune cells. *indicates $p < 0.05$; **indicates $p < 0.01$; ***indicates $p < 0.001$.

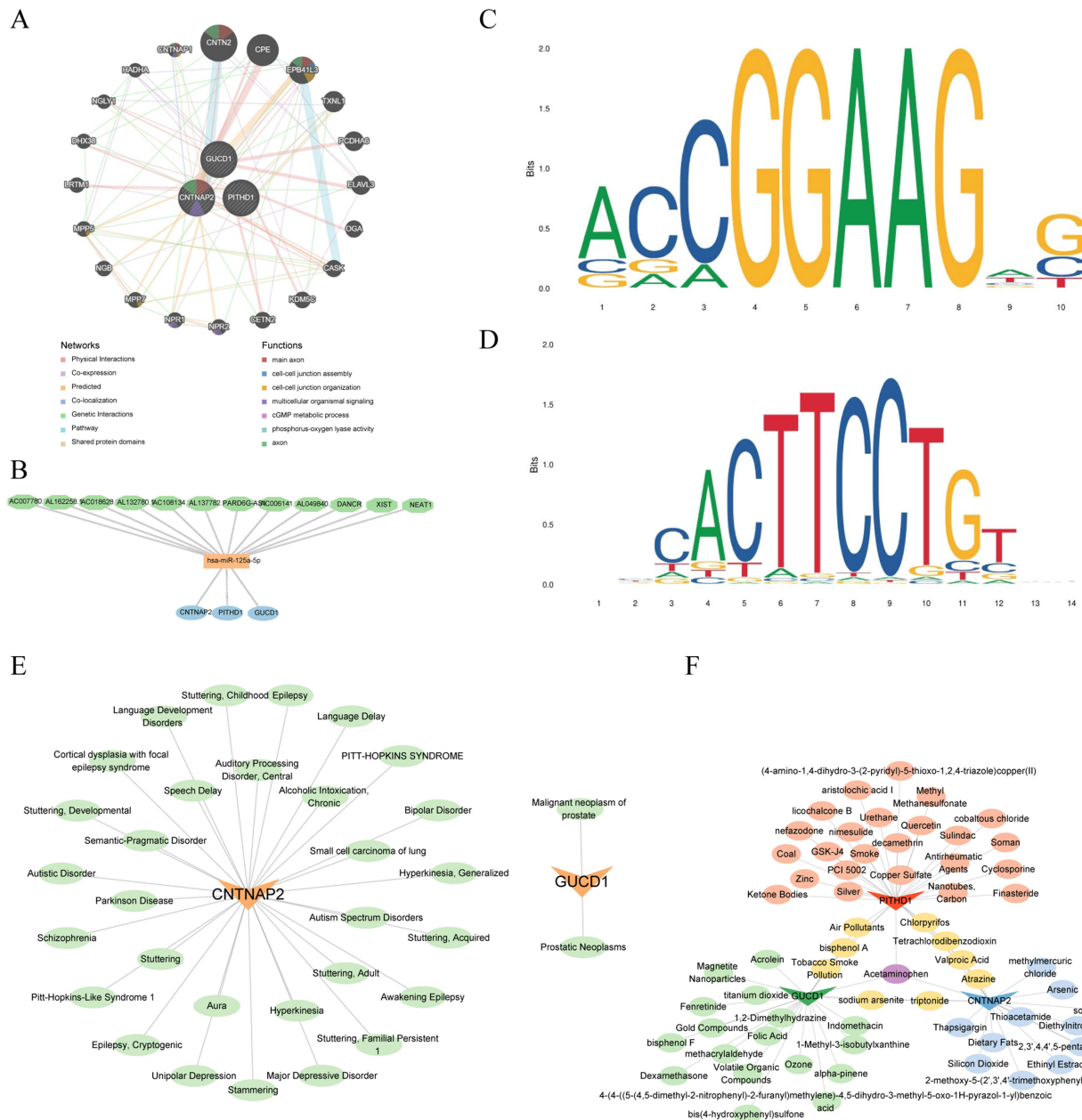


Figure 7 GGI networks construction. **(A)** Biomarkers and their predictive gene interaction networks. **(B)** LncRNR targeting network map. **(C and D)** Identification of TFs. **(E)** Predicted 30 diseases associated with CNTNAP2. **(F)** Predicting therapeutic agents associated with 3 biomarkers.

genes exhibited high specificity, thus the cells were named based on the marker genes ([Supplementary Figure S6D](#)). [Supplementary Figure S6E](#) showed the expression of biomarkers in different cells, where GUCD1 and PITHD1 were significantly different between case and control in T cells and Alveolar cells. However, CNTNAP2 was not expressed in the annotated cells.

Relationship Between T Cells and Other Annotated Cells

We further analyzed the differentiation trajectories of T cells and found that in the disease group, T cells were distributed across both the early and late stages of differentiation, whereas in the control group, they were predominantly found in

Table 1 The TFs of Biomarkers

Rank	TF	FET p-value	Overlapping_Genes
1	TALI	0.01039	GUCD1, PITHD1
2	GABPA	0.02014	PITHD1
3	FOSL1	0.03674	CNTNAP2
4	ETSI	0.04787	GUCD1, PITHD1
5	NFE2	0.04998	PITHD1

the late stages ([Supplementary Figure S7A–C](#)). Additionally, both GUCD1 and PITHD1 were expressed at higher levels during the early and late stages of differentiation ([Supplementary Figure S7D](#)).

To uncover the interactions between annotated cell types, we conducted a cellular communication analysis. This analysis revealed that alveolar cells in the control group exhibited strong interactions with other cell types, whereas T cells and neutrophils showed no interaction with macrophages ([Supplementary Figure S7E](#) and [F](#)). Interestingly, in the case group, macrophages did not interact with epithelial cells or T cells ([Supplementary Figure S7G](#) and [H](#)), suggesting a potential disruption or alteration in cellular communication pathways associated with the disease state. In addition, we found more receptor ligands in the Alveolar cells → B cells and B cells → Alveolar cells interactions in the disease group ([Supplementary Figure S7I](#) and [J](#)).

Causal Relationship Between Biomarkers and COPD

To further explore the causal relationship between biomarkers and COPD, we performed MR analysis. A total of 54 (GUCD1) and 40 (PITHD1) SNPs were used as IVs by screening ([Supplementary Table 3](#)). The MR analysis revealed no causal relationship between the biomarkers examined and COPD ([Table 2](#) and [Supplementary Figure S8A–C](#)), suggesting that other biological or environmental mechanisms might be influencing the development of this disease. Moreover, the robustness of these findings was confirmed through sensitivity analysis ([Supplementary Tables 4, 5](#) and [Supplementary Figure S8D](#)), which accounted for potential confounders and biases, ensuring the reliability of the results. This emphasizes the intricate nature of COPD pathogenesis and highlights the necessity for continued research to uncover additional factors and pathways contributing to the disease.

The Expression of Biomarkers was Validated

At last, in GSE100153 and GSE146560, the expression levels of GUCD1 and PITHD1 were significantly down-regulated in COPD group ([Figure 8A](#) and [B](#)). In addition, the expression of biomarkers was further verified by qRT-PCR, and the solubilisation curves of biomarkers and internal parameters were shown in [Supplementary Figure 9A–D](#). The results displayed that the expression of GUCD1 and PITHD1 were also significantly decreased in COPD group ([Figure 8C](#)).

Table 2 The MR Results of Biomarkers and COPD

Outcome	Exposure	Method	nsnp	b	se	pval	or	or_ici95	or_uci95
COPD	id: eqtl-a-ENSG00000138867	MR Egger	54	-0.154	0.354	0.665	0.857	0.428	1.717
COPD	id: eqtl-a-ENSG00000138867	Weighted median	54	0.093	0.216	0.667	1.098	0.719	1.676
COPD	id: eqtl-a-ENSG00000138867	Inverse variance weighted (fixed effects)	54	0.057	0.154	0.710	1.059	0.783	1.432
COPD	id: eqtl-a-ENSG00000138867	Simple mode	54	-0.553	0.388	0.160	0.575	0.269	1.231
COPD	id: eqtl-a-ENSG00000138867	Weighted mode	54	0.169	0.242	0.487	1.185	0.737	1.904
COPD	id: eqtl-a-ENSG00000057757	MR Egger	40	0.107	0.654	0.871	1.113	0.309	4.011
COPD	id: eqtl-a-ENSG00000057757	Weighted median	40	0.106	0.385	0.783	1.112	0.522	2.367
COPD	id: eqtl-a-ENSG00000057757	Inverse variance weighted (fixed effects)	40	0.351	0.235	0.135	1.420	0.897	2.248
COPD	id: eqtl-a-ENSG00000057757	Simple mode	40	-0.113	0.832	0.893	0.893	0.175	4.561
COPD	id: eqtl-a-ENSG00000057757	Weighted mode	40	-0.044	0.937	0.963	0.957	0.153	6.003

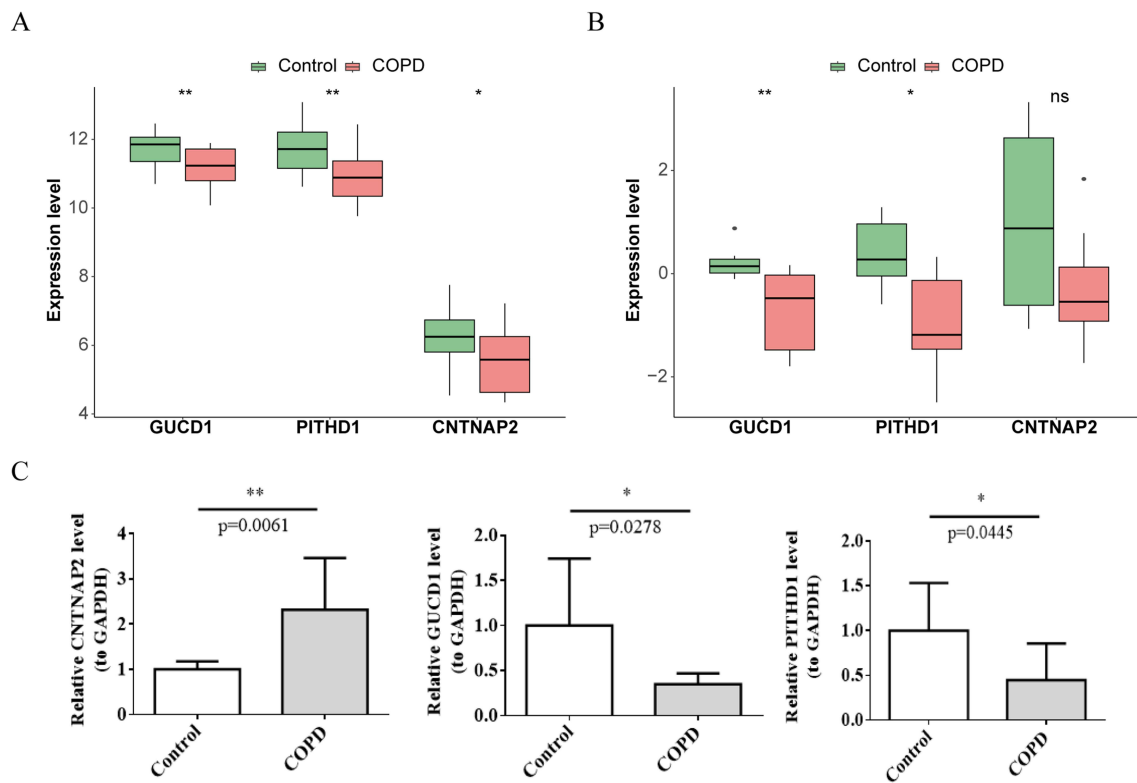


Figure 8 Differential expression analysis of different biomarkers. **(A and B)** Differential expression analysis of different biomarkers in control and COPD groups. **(C)** qRT-PCR analysis of different biomarkers in the control and COPD groups. *indicates $p < 0.05$; **indicates $p < 0.01$.

Discussion

COPD is defined by consistent airflow limitation and is an irreversible, yet preventable, condition.⁴⁵ Although the death rate of COPD declined from 2005 to 2017,⁴⁶ COPD remains a major health problem worldwide.⁸ However, the diagnosis of COPD predominantly depends on spirometry tests and clinical manifestations, and there is persistent underdiagnosis.⁴⁷ Recent findings suggest miR125a's involvement in inflammatory responses across various pathological conditions.^{48–50} Consequently, it is crucial to deepen our understanding of its molecular processes and identify new, trustworthy biomarkers associated with miR-125a-5p in COPD. In this study, a total of 10 candidate genes were obtained by overlapping the DEGs, the module genes, and target genes of miR-125a-5p, of which PITHD1, CNTNAP2 and GUCD1 were identified as biomarkers. The expression of biomarkers was further verified by qRT-PCR, and the expression of PITHD1 and GUCD1 was significantly decreased in COPD group.

Very little is known about PITH domain containing 1 (PITHD1). Recent research indicates that PITHD is down-regulated in leukemia, potentially regulating RUNX1 expression, promoting megakaryocyte differentiation, and activating internal ribosomal entry site.⁵¹ Additionally, while PITHD1 may act as an oncogene in epithelial ovarian carcinoma, its protein expression correlates with a favorable prognosis.⁵² PITHD1 is associated with immunoproteasome in the testis, while it shows no association with thymoproteasomes in the thymus. Mice lacking PITHD1 showed severe male infertility, accompanied by abnormal morphology and impaired spermatozoa motility.⁵³ Expression of PITHD in the olfactory bulb escalates with age in wild-type (WT) mice but diminishes in Tg2576 Alzheimer's disease mice during advanced stages. Stimulating olfactory neuroepithelial (ON) cells with PITHD1 modifies their phosphoproteome, affects cell proliferation rates, and triggers a pro-inflammatory phenotype.⁵⁴ In this study, we have demonstrated for the first time that PITHD1 can serve as a potential biomarker for COPD, with its expression significantly down-regulated in patients with the disease. This finding is particularly noteworthy as it highlights the possibility of utilizing PITHD1 levels as an indicator for COPD diagnosis or progression. The down-regulation of PITHD1 may reflect underlying pathological changes associated with COPD, such as chronic inflammation and tissue remodeling. Our analysis revealed a strong

correlation between PITHD1 expression and several immune cell types, including CD56dim natural killer cells and MDSCs. The significant association of PITHD1 with these immune cells suggests that it may play a role in modulating immune responses in COPD, potentially influencing disease severity through its interaction with inflammatory pathways. Furthermore, our GSEA pathway enrichment analysis revealed that PITHD1 is involved in the ribosomal signaling pathway, indicating that miR-125a-5p may promote inflammatory responses by inhibiting this pathway. Given that ribosomal signaling is crucial for protein synthesis and cellular stress responses, the inhibition of PITHD1 could lead to dysregulation of these processes, potentially exacerbating inflammatory pathways commonly observed in COPD.

A study on hepatocellular carcinoma (HCC) showed that inhibition of guanylyl cyclase domain containing 1 (GUCD1) expression in Hep3B cells was followed by up-regulation of E-cadherin expression and down-regulation of N-cadherin and vimentin expression, suggesting that down-regulation of GUCD1 can inhibit invasion and migration of HCC cells.⁵⁵ Another study showed that GUCD1 was highly expressed in regenerating liver after partial hepatectomy. NEDD4-1 controlled the degradation of GUCD1 through the ubiquitin-proteasome system.⁵⁶ GUCD1 was found to be under-expressed in COPD patients and showed significant correlations with immune functions related to type I and type II interferon responses. These findings suggest that GUCD1 may also be implicated in the pathogenesis of COPD through its involvement in immune regulation. Together, our results establish PITHD1 and GUCD1 as novel biomarkers for COPD and open avenues for further research into their functional roles and therapeutic potential within this complex disease landscape. However, the expression of CNTNAP2 exhibited inconsistencies when comparing our qPCR with those from the validation group. This discrepancy may be attributed to several factors, including differences in sample characteristics, variations in experimental conditions, or potential biological variability among subjects. Further investigation is warranted to understand the underlying mechanisms driving this inconsistency, as CNTNAP2 has been implicated in neurodevelopmental processes and could play a role in respiratory function.

GSEA enrichment analysis displayed that the down-regulated genes were linked to the ribosome pathway. The primary function of ribosomes is to synthesize proteins by translating mRNAs in the nucleolus,⁵⁷ and to participate in the process of cell proliferation, growth, and survival.⁵⁸ Abnormal overexpression of the ribosome pathway in polycystic ovary syndrome can lead to the accumulation of misfolded proteins, cause ovarian endoplasmic reticulum stress, inhibit oocyte maturation and impair embryonic development.⁵⁹ Another finding suggested that upregulation of the ribosomal pathway played a critical role for histone deacetylase 4, which can improve chondrocyte survival rate and biofunction.⁶⁰ Recent studies have shown that blocking ribosome biogenesis can lead to the stabilization and activation of p53, primarily via the ribosomal proteins (RPs)-MDM2-p53 pathway.⁶¹ Polymorphisms in the MDM2 and p53 are linked to apoptotic signaling and the emphysematous changes seen in smokers' lungs. Additionally, cigarette smoke triggers p53-mediated apoptosis in various lung cell types such as pulmonary endothelial cells, alveolar epithelial cells, and terminal bronchiolar cells.⁶² Combining the above information, we conclude that miR-125a-5p may cause COPD susceptibility via the ribosome-p53 pathway.

The proteasome system is considered to be an important regulatory mechanism in cell cycle and growth.⁶³ Prior research indicates that elevated proteasome activity contributes to the breakdown of contractile proteins, resulting in weakened diaphragm function in animal models of COPD.⁶⁴ In individuals with mild to moderate COPD, increased ubiquitin-proteasome pathway activity in the diaphragm leads to damage to diaphragm fibers.⁶⁵ In normal lung fibroblasts, cigarette smoke degrades Akt through the proteasome system, inducing cytotoxicity.⁶³ Acute exposure to cigarette smoke directly stimulates proteasome activity in both mouse lungs and human epithelial cells.⁶⁶ Degradation of the proteasome by methyltransferase proteolysis significantly mitigated cigarette smoke-mediated lung epithelial inflammation and cell apoptosis.⁶⁷ Sestrin 2 interferes with lung injury repair by positively regulating the proteasome and inhibiting platelet-derived growth factor receptor β .⁶⁸ In conclusion, the proteasome pathway is crucial for mediating inflammation, apoptosis and muscle injury in COPD. The miR-125a-5p may aggravate COPD through the proteasome pathway.

Immune infiltration analysis showed that the two down-regulated genes were significantly linked with three types of immune cells: CD56dim natural killer (NK) cell, myeloid-derived suppressor cell (MDSC) and monocyte. NK cells have direct cytotoxicity and influence immunoregulatory secretion of other immune cells,⁶⁹ contributing to their involvement in inflammatory conditions like COPD.⁷⁰ The highly differentiated CD56dimCD16+ NK cells are the dominant

population of NK cells in the human lung.⁶⁹ NK cell levels in the lungs of both current and former smokers were notably lower than those in non-smokers.⁶⁹ MDSCs, composed of a varied group of both mature and immature granulocytic and monocytic cells from the bone marrow, suppress T cell responses and regulate immune responses in inflammatory (non-malignant) conditions.⁷¹ Current smoking has upregulation and activation of circulating MDSCs in patients with COPD, accompanied by significant downregulation of the T-cell receptor (TCR) ζ chain, and this effect persists after smoking cessation.⁷² A notable link was found between the presence of MDSCs in bronchoalveolar lavage fluid (BALF) and airflow limitation severity in COPD patients, independent of their smoking status. This indicates that airway MDSCs also play a key role in local inflammation in COPD.⁷³ In COPD, the function of monocyte-derived macrophages is altered, leading to increased release of pro-inflammatory mediators.⁷⁴ During acute COPD exacerbations, systemic monocyte priming leads to the recruitment and differentiation of these cells into hyper-activated macrophages. These macrophages exhibit increased inflammatory activity compared to their stable state⁷⁴ and are involved in host defense, airway remodeling, and parenchymal destruction.⁷⁵ CD56dimCD16+ NK, MDSC and monocyte all modulate the inflammatory response, suggesting that miR-125a-5p plays an important role in the inflammatory mechanism of COPD.

Single-cell analysis identified six cell clusters, among which PITHD1 and GUCD1 exhibited significant differences between disease and control groups in T cells and alveolar cells. T cells and alveolar cells are key components of the pulmonary immune response, and their functional dysregulation is closely related to the inflammatory responses in COPD. Studies have shown that a reduction in regulatory T cells leads to a rapid decline in lung function, which in turn triggers the onset of COPD.⁷⁶ Moreover, in patients with COPD, an increase in T lymphocytes in the lung parenchyma, airways, and vasculature, along with a weakened memory CD4+ T cell response, has been reported. This immune imbalance may exacerbate the invasiveness of SARS-CoV-2 infection, leading to a higher prevalence of the disease.⁷⁷ Given this, we hypothesize that PITHD1 and GUCD1 may influence the occurrence and progression of COPD by regulating the quantity and function of T cells. Further research into the specific roles of these biomarkers in T cells and alveolar cells could help elucidate their pathogenic mechanisms in COPD and may offer important insights for early intervention and personalized treatment.

The type I IFN response is crucial for the innate immune system.⁷⁸ This response primarily involves the IFN-induced pathway and the JAK/STAT pathway, with the latter facilitating the expression of IFN-stimulated genes.⁷⁸ Type I IFN is a family of cytokines that includes at least 13 IFN- α isotypes and one IFN- β isotype⁷⁹ and is involved in autoinflammatory diseases.⁸⁰ Type I IFN is derived primarily from alveolar macrophages and promotes natural killer cell functions while inhibiting pro-inflammatory pathways and cytokine production.⁸¹ IFN- β expression was reduced by 40–65% in lung cells from patients with stable COPD.⁸¹ The type II IFN response serves as a defensive mechanism, primarily driven by the molecule IFN- γ .⁸² IFN- γ is produced by Th1 lymphocytes.⁸³ There is evidence of elevated IFN- γ levels in the airways and lungs of patients with COPD.^{84–86} IFN- γ leads to phosphorylation of the inflammatory promoter region STAT1⁸⁷ through phosphorylation of JAK2, activating the JAK/STAT signal pathway in macrophages.⁸³ COPD patients have an increased number of alveolar macrophages that secrete chemokines, cytokines and proteases.⁸³ Our study indicated that miR-125a-5p played a central role in disease pathophysiology by regulating the monocyte-macrophage system through type I IFN response and type II IFN response.

In COPD, airway compartment neutrophils release a variety of proteases, including prolyl endopeptidase (PE), serine proteinase, and matrix metalloproteinase (MMPs).⁸⁸ In COPD patients, MMP-8, MMP-9, and PE can be detected in both serum and sputum,^{89,90} and their initial cleavage of collagen produces PGP.⁹¹ PGP acts as a neutrophil chemoattractant and induces features of COPD such as neutrophil inflammation, alveolar enlargement, and right ventricular hypertrophy when introduced into the airways of mice.⁹² One study showed that valproic acid (VPA), a PE inhibitor, modifies the secondary structure of PE in a way that makes substrate binding at the catalytic side of PE impossible, thereby reducing the severity of chronic inflammation. VPA notably suppresses PGP production in vitro and effectively reduces neutrophil influx induced by cigarette smoke in vivo.⁹¹ Our study also predicts VPA as a potential therapeutic agent for COPD based on the CTD database, suggesting that VPA has great potential in the treatment of chronic inflammation in COPD. However, there are still many limitations in this study. For example, the clinical application of biomarkers needs data support from more samples, and the application of targeted drugs needs further clinical trials. Therefore, in order to

further verify the functions and mechanisms of miR-125-5p and its target genes PITHD1, CNTNAP2 and GUCD1 in COPD, further experimental studies are needed.

Conclusion

In summary, our study is the first to analyze the role of miR-125a-5p in COPD. In this study, we found that PITHD1, CNTNAP2 and GUCD1 may be novel biomarker related to miR-125a-5p for COPD. Our research enhances knowledge about the molecular mechanisms of COPD, potentially aiding in its clinical diagnosis and treatment.

Data Sharing Statement

The datasets analyzed in this study, GSE100153 and GSE146560, are available in the GEO database at (<https://www.ncbi.nlm.nih.gov/geo/>).

Ethics Approval and Informed Consent

The study was approved by the ethics committee of Shanxi Bethune Hospital (YXLL-2023-148). All patients had signed an informed consent form.

Consent for Publication

All authors have read and approved the final manuscript and consent to its publication.

Acknowledgment

We thank the entire team for their support and assistance with this study.

Author Contributions

All authors made a significant contribution to the work reported, whether that is in the conception, study design, execution, acquisition of data, analysis and interpretation, or in all these areas; took part in drafting, revising or critically reviewing the article; gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agree to be accountable for all aspects of the work.

Funding

Scientific Research Project of the Shanxi Provincial Health Commission (201601010), the Doctoral Foundation of Shanxi Bethune Hospital (2019YJ01), and the Fundamental Research Program of Shanxi Province (20210302123492).

Disclosure

The authors report no conflicts of interest in this work.

References

1. Celli B, Fabbri L, Criner G, et al. Definition and nomenclature of chronic obstructive pulmonary disease: time for its revision. *Am J Respir Crit Care Med.* 2022;206:1317–1325. doi:10.1164/rccm.202204-0671PP
2. Halpin DMG, Celli BR, Criner GJ, et al. The GOLD Summit on chronic obstructive pulmonary disease in low- and middle-income countries. *Int J Tuberc Lung Dis.* 2019;23:1131–1141. doi:10.5588/ijtld.19.0397
3. Meghji J, Mortimer K, Agusti A, et al. Improving lung health in low-income and middle-income countries: from challenges to solutions. *Lancet.* 2021;397:928–940. doi:10.1016/S0140-6736(21)00458-X
4. Agusti A, Melen E, DeMeo DL, et al. Pathogenesis of chronic obstructive pulmonary disease: understanding the contributions of gene-environment interactions across the lifespan. *Lancet Respir Med.* 2022;10:512–524. doi:10.1016/S2213-2600(21)00555-5
5. Sin DD, Doiron D, Agusti A, et al. Air pollution and COPD: GOLD 2023 committee report. *Eur Respir J.* 2023;61:2202469. doi:10.1183/13993003.02469-2022
6. Yang IA, Jenkins CR, Salvi SS. Chronic obstructive pulmonary disease in never-smokers: risk factors, pathogenesis, and implications for prevention and treatment. *Lancet Respir Med.* 2022;10:497–511. doi:10.1016/S2213-2600(21)00506-3
7. Cho MH, Hobbs BD, Silverman EK. Genetics of chronic obstructive pulmonary disease: understanding the pathobiology and heterogeneity of a complex disorder. *Lancet Respir Med.* 2022;10:485–496. doi:10.1016/S2213-2600(21)00510-5
8. Zhao Y, Li M, Yang Y, et al. Identification of macrophage polarization-related genes as biomarkers of chronic obstructive pulmonary disease based on bioinformatics analyses. *Biomed Res Int.* 2021;2021:9921012. doi:10.1155/2021/9921012

9. Hirai K, Shirai T, Shimoshikiryo T, et al. Circulating microRNA-15b-5p as a biomarker for asthma-COPD overlap. *Allergy*. 2021;76:766–774. doi:10.1111/all.14520
10. Guo S, Lu J, Schlanger R, et al. MicroRNA miR-125a controls hematopoietic stem cell number. *Proc Natl Acad Sci U S A*. 2010;107:14229–14234. doi:10.1073/pnas.0913574107
11. Zhang C, Zhao Y, Wang Q, et al. Overexpression of miR-125a-5p inhibits hepatocyte proliferation through the STAT3 regulation in vivo and in vitro. *Int J Mol Sci*. 2022;23:8661.
12. Zhan JL, Huang YL, Liang QW, et al. Anti-inflammatory effect of miR-125a-5p on experimental optic neuritis by promoting the differentiation of Treg cells. *Neural Regen Res*. 2023;18:451–455. doi:10.4103/1673-5374.346462
13. Jing X, Huo J, Li L, et al. Baicalin relieves airway inflammation in COPD by inhibiting miR-125a. *Appl Biochem Biotechnol*. 2023;196:3374–3386. doi:10.1007/s12010-023-04671-y
14. Wang R, Xu J, Liu H, et al. Peripheral leukocyte microRNAs as novel biomarkers for COPD. *Int J Chron Obstruct Pulmon Dis*. 2017;12:1101–1112. doi:10.2147/COPD.S130416
15. Hsu AC, Dua K, Starkey MR, et al. MicroRNA-125a and -b inhibit A20 and MAVS to promote inflammation and impair antiviral response in COPD. *JCI Insight*. 2017;2:e90443. doi:10.1172/jci.insight.90443
16. Li N, Yi H, Sun W, et al. Revealing genes associated with cervical cancer in distinct immune cells: a comprehensive Mendelian randomization analysis. *Int J Cancer*. 2024;155:149–158. doi:10.1002/ijc.34911
17. Altman MC, Rinchai D, Baldwin N, et al. Development of a fixed module repertoire for the analysis and interpretation of blood transcriptome data. *Nat Commun*. 2021;12:4385. doi:10.1038/s41467-021-24584-w
18. Trivedi A, Bade G, Madan K, et al. Effect of smoking and its cessation on the transcript profile of peripheral monocytes in COPD patients. *Int J Chron Obstruct Pulmon Dis*. 2022;17:65–77. doi:10.2147/COPD.S337635
19. Zhu H, Tan J, Wang Z, et al. Bioinformatics analysis constructs potential ferroptosis-related ceRNA network involved in the formation of intracranial aneurysm. *Front Cell Neurosci*. 2022;16:1016682. doi:10.3389/fncel.2022.1016682
20. Lin W, Wang Y, Chen Y, et al. Role of calcium signaling pathway-related gene regulatory networks in ischemic stroke based on multiple WGCNA and single-cell analysis. *Oxid Med Cell Longev*. 2021;2021:8060477. doi:10.1155/2021/8060477
21. Zhang Q, Li J, Weng L. Identification and validation of aging-related genes in Alzheimer's disease. *Front Neurosci*. 2022;16:905722. doi:10.3389/fnins.2022.905722
22. Zhang MY, Huo C, Liu JY, et al. Identification of a five autophagy subtype-related gene expression pattern for improving the prognosis of lung adenocarcinoma. *Front Cell Dev Biol*. 2021;9:756911. doi:10.3389/fcell.2021.756911
23. Chen L, Zhang YH, Lu G, et al. Analysis of cancer-related lncRNAs using gene ontology and KEGG pathways. *Artif Intell Med*. 2017;76:27–36. doi:10.1016/j.artmed.2017.02.001
24. Liu Z, Mi M, Li X, et al. A lncRNA prognostic signature associated with immune infiltration and tumour mutation burden in breast cancer. *J Cell Mol Med*. 2020;24:12444–12456. doi:10.1111/jcmm.15762
25. Hillyar CR, Kanabar SS, Pufal KR, et al. A systematic review and meta-analysis of the diagnostic effectiveness of human papillomavirus methylation biomarkers for detection of cervical cancer. *Epigenomics*. 2022;14:1055–1072. doi:10.2217/epi-2022-0160
26. Hallstrom AP. A modified Wilcoxon test for non-negative distributions with a clump of zeros. *Stat Med*. 2010;29:391–400. doi:10.1002/sim.3785
27. Liu X, Lin L, Lu L, et al. Comparative transcriptional analysis identified characteristic genes and patterns in HIV-infected immunological non-responders. *Front Immunol*. 2022;13:807890. doi:10.3389/fimmu.2022.807890
28. Liu X, Zhang W, Wang H, et al. Decreased expression of ACADSB predicts poor prognosis in clear cell renal cell carcinoma. *Front Oncol*. 2021;11:762629. doi:10.3389/fonc.2021.762629
29. Liu Z, Li H, Pan S. Discovery and validation of key biomarkers based on immune infiltrates in Alzheimer's disease. *Front Genet*. 2021;12:658323. doi:10.3389/fgene.2021.658323
30. Charoentong P, Finotello F, Angelova M, et al. Pan-cancer Immunogenomic analyses reveal genotype-immunophenotype relationships and predictors of response to checkpoint blockade. *Cell Rep*. 2017;18:248–262. doi:10.1016/j.celrep.2016.12.019
31. Zheng X, Ma Y, Bai Y, et al. Identification and validation of immunotherapy for four novel clusters of colorectal cancer based on the tumor microenvironment. *Front Immunol*. 2022;13:984480. doi:10.3389/fimmu.2022.984480
32. Liang JY, Wang DS, Lin HC, et al. A novel ferroptosis-related gene signature for overall survival prediction in patients with hepatocellular carcinoma. *Int J Biol Sci*. 2020;16:2430–2441. doi:10.7150/ijbs.45050
33. Jin Y, Wang Z, He D, et al. Identification of novel subtypes based on ssGSEA in immune-related prognostic signature for tongue squamous cell carcinoma. *Cancer Med*. 2021;10:8693–8707. doi:10.1002/cam4.4341
34. Yue T, Chen S, Zhu J, et al. The aging-related risk signature in colorectal cancer. *Aging*. 2021;13:7330–7349. doi:10.18632/aging.202589
35. Yang JH, Li JH, Shao P, et al. starBase: a database for exploring microRNA-mRNA interaction maps from Argonaute CLIP-Seq and Degradome-Seq data. *Nucleic Acids Res*. 2011;39:D202–9. doi:10.1093/nar/gkq1056
36. Shannon P, Markiel A, Ozier O, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res*. 2003;13:2498–2504. doi:10.1101/gr.1239303
37. Fu W, Kadeer G, He Y, et al. The regulatory network of potential transcription factors and MiRNAs of mitochondria-related genes for sarcopenia. *Front Genet*. 2022;13:975886. doi:10.3389/fgene.2022.975886
38. Castro-Mondragon JA, Riudavets-Puig R, Rauluseviciute I, et al. JASPAR 2022: the 9th release of the open-access database of transcription factor binding profiles. *Nucleic Acids Res*. 2022;50:D165–D173. doi:10.1093/nar/gkab1113
39. Davis AP, Wiegiers TC, Johnson RJ, et al. Comparative Toxicogenomics Database (CTD): update 2023. *Nucleic Acids Res*. 2023;51:D1257–D1262. doi:10.1093/nar/gkac833
40. Zhang L, Han L, Wang X, et al. Exploring the mechanisms underlying the therapeutic effect of *Salvia miltiorrhiza* in diabetic nephropathy using network pharmacology and molecular docking. *Biosci Rep*. 2021;41. doi:10.1042/BSR20203520
41. Satija R, Farrell JA, Gennert D, et al. Spatial reconstruction of single-cell gene expression data. *Nat Biotechnol*. 2015;33:495–502. doi:10.1038/nbt.3192
42. Du J, Yuan X, Deng H, et al. Single-cell and spatial heterogeneity landscapes of mature epicardial cells. *J Pharm Anal*. 2023;13:894–907. doi:10.1016/j.jpha.2023.07.011

43. Jin S, Guerrero-Juarez CF, Zhang L, et al. Inference and analysis of cell-cell communication using CellChat. *Nat Commun.* 2021;12:1088. doi:10.1038/s41467-021-21246-9
44. Hemani G, Zheng J, Elsworth B, et al. The MR-Base platform supports systematic causal inference across the human phenome. *Elife.* 2018;7. doi:10.7554/eLife.34408
45. Vogelmeier CF, Criner GJ, Martinez FJ, et al. Global strategy for the diagnosis, management, and prevention of chronic obstructive lung disease 2017 report. GOLD executive summary. *Am J Respir Crit Care Med.* 2017;195:557–582. doi:10.1164/rccm.201701-0218PP
46. Calverley PMA, Walker PP. Contemporary concise review 2022: chronic obstructive pulmonary disease. *Respirology.* 2023;28:428–436. doi:10.1111/resp.14489
47. Morrow JD, Chase RP, Parker MM, et al. RNA-sequencing across three matched tissues reveals shared and tissue-specific gene expression and pathway signatures of COPD. *Respir Res.* 2019;20:65. doi:10.1186/s12931-019-1032-z
48. Curtale G, Renzi TA, Mirolo M, et al. Multi-Step regulation of the TLR4 pathway by the miR-125a~99b~let-7e Cluster. *Front Immunol.* 2018;9:2037. doi:10.3389/fimmu.2018.02037
49. Hwang SJ, Ahn BJ, Shin MW, et al. miR-125a-5p attenuates macrophage-mediated vascular dysfunction by targeting Ninjurin1. *Cell Death Differ.* 2022;29:1199–1210. doi:10.1038/s41418-021-00911-y
50. Zhaolin Z, Jiaojiao C, Peng W, et al. OxLDL induces vascular endothelial cell pyroptosis through miR-125a-5p/TET2 pathway. *J Cell Physiol.* 2019;234:7475–7491. doi:10.1002/jcp.27509
51. Lu B, Sun X, Chen Y, et al. Novel function of PITH domain-containing 1 as an activator of internal ribosomal entry site to enhance RUNX1 expression and promote megakaryocyte differentiation. *Cell Mol Life Sci.* 2015;72:821–832. doi:10.1007/s00018-014-1704-2
52. Engqvist H, Parris TZ, Kovacs A, et al. Immunohistochemical validation of COL3A1, GPR158 and PITHD1 as prognostic biomarkers in early-stage ovarian carcinomas. *BMC Cancer.* 2019;19:928. doi:10.1186/s12885-019-6084-4
53. Kondo H, Matsumura T, Kaneko M, et al. PITHD1 is a proteasome-interacting protein essential for male fertilization. *J Biol Chem.* 2020;295:1658–1672. doi:10.1074/jbc.RA119.011144
54. Lachen-Montes M, Mendizuri N, Ausin K, et al. Smelling the dark proteome: functional characterization of PITH domain-containing protein 1 (Clorf128) in olfactory metabolism. *J Proteome Res.* 2020;19:4826–4843. doi:10.1021/acs.jproteome.0c00452
55. He Y, He X. MicroRNA-370 regulates cellearthelial-mesenchymal transition, migration, invasion, and prognosis of hepatocellular carcinoma by targeting GUCD1. *Yonsei Med J.* 2019;60:267–276. doi:10.3349/yjm.2019.60.3.267
56. Bellet MM, Piobbico D, Bartoli D, et al. NEDD4 controls the expression of GUCD1, a protein upregulated in proliferating liver cells. *Cell Cycle.* 2014;13:1902–1911. doi:10.4161/cc.28760
57. Brogna S, Sato TA, Rosbash M. Ribosome components are associated with sites of transcription. *Mol Cell.* 2002;10:93–104. doi:10.1016/S1097-2765(02)00565-8
58. Turi Z, Lacey M, Mistrik M, et al. Impaired ribosome biogenesis: mechanisms and relevance to cancer and aging. *Aging.* 2019;11:2512–2540. doi:10.18632/aging.101922
59. Nakanishi N, Osuka S, Kono T, et al. Upregulated ribosomal pathway impairs follicle development in a polycystic ovary syndrome mouse model: differential gene expression analysis of oocytes. *Reprod Sci.* 2023;30:1306–1315. doi:10.1007/s43032-022-01095-7
60. Guo L, Guo H, Zhang Y, et al. Upregulated ribosome pathway plays a key role in HDAC4, improving the survival rate and biofunction of chondrocytes. *Bone Joint Res.* 2023;12:433–446. doi:10.1302/2046-3758.127.BJR-2022-0279.R2
61. Zhang Y, Zhang J, Chen X, et al. Polymeric immunoglobulin receptor (PIGR) exerts oncogenic functions via activating ribosome pathway in hepatocellular carcinoma. *Int J Med Sci.* 2021;18:364–371. doi:10.7150/ijms.49790
62. Mizuno S, Ishizaki T, Kadowaki M, et al. p53 signaling pathway polymorphisms associated with emphysematous changes in patients with COPD. *Chest.* 2017;152:58–69. doi:10.1016/j.chest.2017.03.012
63. Kim SY, Lee JH, Huh JW, et al. Cigarette smoke induces Akt protein degradation by the ubiquitin-proteasome system. *J Biol Chem.* 2011;286:31932–31943. doi:10.1074/jbc.M111.267633
64. van Hees H, Ottenheijm C, Ennen L, et al. Proteasome inhibition improves diaphragm function in an animal model for COPD. *Am J Physiol Lung Cell Mol Physiol.* 2011;301:L110–6. doi:10.1152/ajplung.00396.2010
65. Ottenheijm CA, Heunks LM, Li YP, et al. Activation of the ubiquitin-proteasome pathway in the diaphragm in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med.* 2006;174:997–1002. doi:10.1164/rccm.200605-721OC
66. van Rijt SH, Keller IE, John G, et al. Acute cigarette smoke exposure impairs proteasome function in the lung. *Am J Physiol Lung Cell Mol Physiol.* 2012;303:L814–23. doi:10.1152/ajplung.00128.2012
67. Li T, He X, Luo L, et al. F-box protein FBXW17-mediated proteasomal degradation of protein methyltransferase PRMT6 exaggerates CSE-induced lung epithelial inflammation and apoptosis. *Front Cell Dev Biol.* 2021;9:599020. doi:10.3389/fcell.2021.599020
68. Tomasovic A, Kurrle N, Surun D, et al. Sestrin 2 protein regulates platelet-derived growth factor receptor beta (Pdgfrbeta) expression by modulating proteasomal and Nrf2 transcription factor functions. *J Biol Chem.* 2015;290:9738–9752. doi:10.1074/jbc.M114.632133
69. Marquardt N, Kekalainen E, Chen P, et al. Human lung natural killer cells are predominantly comprised of highly differentiated hypofunctional CD69(-)CD56(dim) cells. *J Allergy Clin Immunol.* 2017;139:1321–1330e4. doi:10.1016/j.jaci.2016.07.043
70. Freeman CM, Stolberg VR, Crudgington S, et al. Human CD56+ cytotoxic lung lymphocytes kill autologous lung cells in chronic obstructive pulmonary disease. *PLoS One.* 2014;9:e103840. doi:10.1371/journal.pone.0103840
71. Scrimini S, Pons J, Agusti A, et al. Expansion of myeloid-derived suppressor cells in chronic obstructive pulmonary disease and lung cancer: potential link between inflammation and cancer. *Cancer Immunol Immunother.* 2015;64:1261–1270. doi:10.1007/s00262-015-1737-x
72. Scrimini S, Pons J, Agusti A, et al. Differential effects of smoking and COPD upon circulating myeloid derived suppressor cells. *Respir Med.* 2013;107:1895–1903. doi:10.1016/j.rmed.2013.08.002
73. Brajer-Luftmann B, Nowicka A, Kaczmarek M, et al. Myeloid-derived suppressor cells in bronchoalveolar lavage fluid in patients with chronic obstructive pulmonary disease. *Pol Arch Med Wewn.* 2016;126:980–988. doi:10.20452/pamw.3718
74. Singh R, Belchamber KBR, Fenwick PS, et al. Defective monocyte-derived macrophage phagocytosis is associated with exacerbation frequency in COPD. *Respir Res.* 2021;22:113. doi:10.1186/s12931-021-01718-8
75. Ravi AK, Plumb J, Gaskell R, et al. COPD monocytes demonstrate impaired migratory ability. *Respir Res.* 2017;18:90. doi:10.1186/s12931-017-0569-y

76. Eriksson Ström J, Pourazar J, Linder R, et al. Airway regulatory T cells are decreased in COPD with a rapid decline in lung function. *Respir Res.* 2020;21:330. doi:10.1186/s12931-020-01593-9
77. Yang CX, Tomchaney M, Landecho MF, et al. Lung spatial profiling reveals a T cell signature in COPD patients with fatal SARS-CoV-2 infection. *Cells.* 2022;11:1864.
78. Xue W, Ding C, Qian K, et al. The interplay between coronavirus and type I IFN response. *Front Microbiol.* 2021;12:805472. doi:10.3389/fmicb.2021.805472
79. Kornstadt L, Pierre S, Weigert A, et al. Bacterial and fungal toll-like receptor activation elicits type I IFN responses in mast cells. *Front Immunol.* 2020;11:607048. doi:10.3389/fimmu.2020.607048
80. Morell M, Varela N, Castillejo-Lopez C, et al. SIDT1 plays a key role in type I IFN responses to nucleic acids in plasmacytoid dendritic cells and mediates the pathogenesis of an imiquimod-induced psoriasis model. *EBioMedicine.* 2022;76:103808. doi:10.1016/j.ebiom.2021.103808
81. Garcia-Valero J, Olloquequi J, Montes JF, et al. Deficient pulmonary IFN-beta expression in COPD patients. *PLoS One.* 2019;14:e0217803. doi:10.1371/journal.pone.0217803
82. Huang R, Li Z, Zhu X, et al. Collagen type III alpha 1 chain regulated by GATA-binding protein 6 affects type II IFN response and propanoate metabolism in the recurrence of lower grade glioma. *J Cell Mol Med.* 2020;24:10803–10815. doi:10.1111/jcmm.15705
83. Southworth T, Metryka A, Lea S, et al. IFN-gamma synergistically enhances LPS signalling in alveolar macrophages from COPD patients and controls by corticosteroid-resistant STAT1 activation. *Br J Pharmacol.* 2012;166:2070–2083. doi:10.1111/j.1476-5381.2012.01907.x
84. Hens G, Vanaudenaerde BM, Bullens DM, et al. Sinonasal pathology in nonallergic asthma and COPD: ‘united airway disease’ beyond the scope of allergy. *Allergy.* 2008;63:261–267. doi:10.1111/j.1398-9995.2007.01545.x
85. Panzner P, Lafitte JJ, Tscipoulos A, et al. Marked up-regulation of T lymphocytes and expression of interleukin-9 in bronchial biopsies from patients with chronic bronchitis with obstruction. *Chest.* 2003;124:1909–1915. doi:10.1378/chest.124.5.1909
86. Reeves EP, Williamson M, Byrne B, et al. IL-8 dictates glycosaminoglycan binding and stability of IL-18 in cystic fibrosis. *J Immunol.* 2010;184:1642–1652. doi:10.4049/jimmunol.0902605
87. Gough DJ, Levy DE, Johnstone RW, et al. IFNgamma signaling-does it mean JAK-STAT? *Cytokine Growth Factor Rev.* 2008;19:383–394. doi:10.1016/j.cytogfr.2008.08.004
88. Gaggar A, Jackson PL, Noerager BD, et al. A novel proteolytic cascade generates an extracellular matrix-derived chemoattractant in chronic neutrophilic inflammation. *J Immunol.* 2008;180:5662–5669. doi:10.4049/jimmunol.180.8.5662
89. O’Reilly P, Jackson PL, Noerager B, et al. N-alpha-PGP and PGP, potential biomarkers and therapeutic targets for COPD. *Respir Res.* 2009;10:38. doi:10.1186/1465-9921-10-38
90. O’Reilly PJ, Hardison MT, Jackson PL, et al. Neutrophils contain prolyl endopeptidase and generate the chemotactic peptide, PGP, from collagen. *J Neuroimmunol.* 2009;217:51–54. doi:10.1016/j.jneuroim.2009.09.020
91. Abdul Roda M, Sadik M, Gaggar A, et al. Targeting prolyl endopeptidase with valproic acid as a potential modulator of neutrophilic inflammation. *PLoS One.* 2014;9:e97594. doi:10.1371/journal.pone.0097594
92. Weathington NM, van Houwelingen AH, Noerager BD, et al. A novel peptide CXCR ligand derived from extracellular matrix degradation during airway inflammation. *Nat Med.* 2006;12:317–323. doi:10.1038/nm1361

International Journal of Chronic Obstructive Pulmonary Disease

Publish your work in this journal

The International Journal of COPD is an international, peer-reviewed journal of therapeutics and pharmacology focusing on concise rapid reporting of clinical studies and reviews in COPD. Special focus is given to the pathophysiological processes underlying the disease, intervention programs, patient focused education, and self management protocols. This journal is indexed on PubMed Central, MedLine and CAS. The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.

Submit your manuscript here: <https://www.dovepress.com/international-journal-of-chronic-obstructive-pulmonary-disease-journal>

Dovepress
Taylor & Francis Group