Journal of Inflammation Research

Dovepress
Taylor & Francis Group

Open Access Full Text Article

ORIGINAL RESEARCH

Identification of Glycosylation-Related Biomarkers in COPD and IPF Through Integrated Machine Learning and WGCNA Analysis

Xiao Ling Yin[1], Ying Zhai[2], Lei Wang[3]

[1]Department of Respiratory, Zibo Hospital of Integrated Traditional Chinese and Western Medicine, Zibo City, Shandong Province, 255022, People's Republic of China; [2]Department of Neurology, Zibo Hospital of Integrated Traditional Chinese and Western Medicine, Zibo City, Shandong Province, 255022, People's Republic of China; [3]Department of Oncology, Zibo Hospital of Integrated Traditional Chinese and Western Medicine, Zibo City, Shandong Province, 255022, People's Republic of China

Correspondence: Lei Wang, Zibo Hospital of Integrated Traditional Chinese and Western Medicine, No. 8, Jinjing Avenue, Zhangdian District, Zibo City, Shandong Province, 255022, People's Republic of China, Email WangleiZXYJH@163.com

Objective: This research aimed to explore key glycosylation-related genes (signature genes) and associated molecular mechanism on chronic obstructive pulmonary disease (COPD) and idiopathic pulmonary fibrosis (IPF), which further providing new perspectives for disease prognosis and diagnose.

Patients and Methods: The gene expression profiles were obtained from the public GEO database. The glycosylation-related genes were identified based co-DEGs from COPD vs normal samples and IPF vs normal samples, module genes by weighted gene co-expression network analysis (WGCNA), as well as glycosylation genes from database. Signature genes were screened using machine learning methods, followed by immune infiltration, function analysis, drug-gene and transcriptional regulatory network analysis. Finally, validation analysis based on tissue samples from COPD/IPF patients were performed to test the expression of signature genes.

Results: A total of 35 differentially expressed glycosylation-related genes for both COPD and IPF were explored. By three kinds of machine learning analyses, totally three signature genes including SULF1, ST8SIA1 and FCN3 were explored. In COPD, the AUC values for FCN3, ST8SIA1, and SULF1 were 0.643, 0.722, and 0.719, respectively; while in IPF, they were 0.955, 0.792, and 0.943, respectively. Immune infiltration and GSEA analysis showed that signature genes were dramatically correlated with activated B cell and extracellular matrix (ECM)-associated functions ($P < 0.05$). Bisphenol A and Valproic acid were common drugs for both three signature genes. Validation analysis found that the ST8SIA1 expression was related to the disease state ($P < 0.001$). Functional experiments revealed that it affected cell behaviors, including proliferation, apoptosis, migration, and invasion (all, $P < 0.01$), and drug treatments could regulate its expression and glycosylation levels (all, $P < 0.05$), providing crucial evidence for biomarker research and pathogenesis exploration of the two diseases.

Conclusion: We identified SULF1, ST8SIA1 and FCN3 as shared glycosylation-related biomarkers in COPD and IPF. These genes bridge fibrosis, inflammation, and immune dysregulation, offering potential diagnostic and therapeutic targets.

Keywords: chronic obstructive pulmonary disease, glycosylation-related gene, idiopathic pulmonary fibrosis, immune cell infiltration, machine learning, validation analysis

Introduction

Chronic obstructive pulmonary disease (COPD) and idiopathic pulmonary fibrosis (IPF) are both progressive, debilitating respiratory disorders that significantly affect the quality of life of patients.[1,2] COPD is primarily characterized by chronic inflammation and irreversible airflow limitation, resulting from long-term exposure to harmful particles such as tobacco smoke. On the other hand, IPF is a form of interstitial lung disease marked by the accumulation of scar tissue (fibrosis) within the lungs, leading to respiratory failure. Although distinct in their underlying pathophysiology, these two diseases share common features, particularly in their chronic inflammatory components and fibrotic remodeling of lung tissue. Interestingly, some patients exhibit overlapping characteristics of both conditions,[3] known as the COPD-IPF overlap

https://doi.org/10.2147/JIR.S537967
Received: 1 May 2025
Accepted: 26 November 2025
Published: 22 January 2026

syndrome. These patients tend to have a more complex clinical course and a worse prognosis due to the combined impact of inflammation and fibrosis on lung function. Understanding the molecular mechanisms driving both COPD and IPF, particularly those that overlap, is crucial for improving diagnostic and therapeutic strategies.

Glycosylation is a post-translational modification process, an enzymatic process that links a sugar group to a protein or lipid through the catalysis of glycosyltransferases.[4] Glycosylation can be broadly divided into two types: O-linked glycosylation (for serine or threonine amino acids) and N-linked glycosylation (for asparagine residues).[5,6] Glycosylation plays a multifaceted role in the pathogenesis of lung diseases, including cell signal transduction, cell adhesion and immune regulation, promoting epithelial-mesenchymal transition, and reducing drug sensitivity.[6,7] In both COPD and IPF, glycosylation is intricately linked to disease progression, influencing inflammation, tissue remodeling, and fibrosis.[8,9] In IPF, alterations in the glycosylation of extracellular matrix (ECM) components such as collagen and fibronectin have been shown to enhance their deposition and crosslinking, which contributes to the thickening of lung tissue and the progression of fibrosis.[10] Moreover, abnormal glycosylation of cell surface glycoproteins and glycolipids in alveolar epithelial cells impairs their regenerative capacity, exacerbating lung damage and fibrosis. In COPD, the role of glycosylation is similarly significant, particularly in the context of mucus hypersecretion. COPD patients often have elevated levels of mucins, glycoproteins that play a crucial role in airway defense.[11] The degree and type of glycosylation on mucins can influence their viscosity and clearance from the airways. Altered glycosylation of these mucins may increase mucus thickness, impairing mucociliary clearance and contributing to airway obstruction. Interleukin-8 (IL-8), a pro-inflammatory cytokine renowned for its neutrophil chemotactic activity, has been found to be elevated in patients with allergic rhinitis and chronic rhinosinusitis with nasal polyps.[12] IL-8 recruits inflammatory neutrophils to infiltrate the lungs of COPD patients and correlates with disease severity in COPD[13]. In IPF, IL-8 simultaneously promotes senescence of mesenchymal progenitor cells and upregulates PD-L1 expression, thereby evading immune surveillance.[14] Additionally, the glycosylation of inflammatory mediators such as IL-8 can impact their biological activity and stability, thereby sustaining chronic inflammation in the airways.[15] As with IPF, sustained inflammation in COPD leads to structural lung damage and remodeling, driving the disease's progression. Glycosylation modification sustains chronic airway inflammation in COPD while potentially promoting the formation and maintenance of a fibrotic microenvironment in IPF.[10,16,17] Although the specific mechanistic pathways involved in glycosylation overlap between these two diseases, its effects are more focused on mucus secretion and neutrophilic inflammation in COPD, whereas in IPF they predominantly involve ECM deposition and epithelial-mesenchymal interactions. Notably, in COPD patients, increased inflammation has been observed, characterized by reduced galactosylation and sialylation. Plasma IgG glycosylation may serve as a potential biomarker for early diagnosis of COPD.[16] The reduction of ST6GAL1 and α2-6 sialylation enhances IL-6 secretion in human bronchial epithelial cells (HBECs) and is associated with adverse clinical outcomes in COPD.[18] Additionally, a study identified serum levels of AGE, AOPP, and MMP7 as potential diagnostic markers for IPF.[9] While numerous studies have highlighted the importance of glycosylation in the pathophysiology of diseases, there remains a need for comprehensive investigations to identify key glycosylation-related biomarkers that could serve as diagnostic or prognostic tools.[19,20] Furthermore, the complex interactions between glycosylation and disease mechanisms remain poorly understood, particularly when considering overlapping features in COPD and IPF. We hypothesize that in COPD-IPF overlap syndrome, glycosylation patterns may exhibit unique characteristics: on one hand, glycosylation alterations associated with mucus hypersecretion and neutrophil inflammation (such as mucin glycoforms) may resemble those observed in COPD; on the other hand, aberrant glycosylation of ECM components and profibrotic glycan signaling may share similarities with IPF. The convergence and divergence of glycosylation features in this overlap syndrome suggest their potential as molecular markers for distinguishing disease phenotypes, while providing insights into the interplay between these pathological mechanisms.

This study innovatively integrates machine learning with weighted gene co-expression network analysis (WGCNA) to systematically identify hub genes related to glycosylation that are shared between COPD and IPF. This integrated strategy not only leverages WGCNA to extract co-expression modules highly correlated with phenotypes from high-throughput data but also employs machine learning algorithms to further identify biomarkers with high discriminatory power, thereby enabling more precise molecular feature mining compared to previous single-method approaches. This study explored common glycosylation biomarkers shared between COPD and IPF through bioinformatics analysis, followed by functional validation

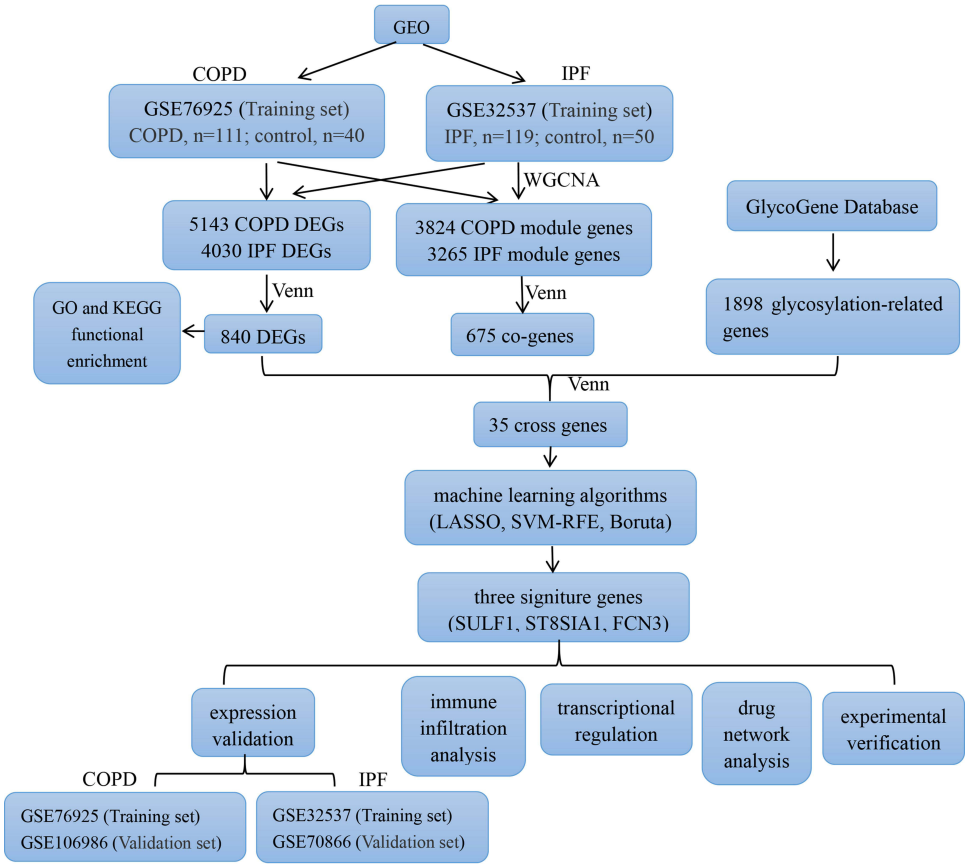

Figure 1 Workflow diagram of this research.

of key characteristic genes. These findings will enhance our understanding of shared mechanisms underlying these two diseases and lay the foundation for developing novel diagnostic and therapeutic strategies. The study workflow is illustrated in Figure 1.

Materials and Methods

Microarray Data and Data Preprocessing

Three datasets including GSE76925 (111 COPD samples and 40 control samples) and GSE106986 (14 COPD samples and 5 control samples) were download from Gene Expression Omnibus (GEO) database as the training data. In addition, GSE32537 (169 IPF samples and 50 control samples) and GSE70866 (112 IPF samples and 20 control samples) was enrolled as validation dataset in current study. For the two datasets (GSE76925 and GSE32537) used in the analysis, background correction, quantile normalization, and summarization were performed using the RMA (Robust Multi-array Average) method to ensure comparability across samples. Quality control was conducted as follows: boxplots were used to examine the expression value distribution of samples in GSE76925 and GSE32537, with no outlier samples identified. Principal component analysis (PCA) and hierarchical clustering were performed to visualize sample grouping and detect outliers. In preliminary analysis, no significant outlier samples were observed in either dataset (Supplementary Figure 1A–J). Using the probe expression matrix and annotation file, probes that did not correspond to a gene symbol were excluded. For genes with multiple corresponding probes, the average expression value of these probes was calculated and used as the expression value for that gene.

Investigation of Common DEGs for COPD and IPF

By using limma package in R (version: 3.58.1),[21] the DEGs for COPD vs control and IPF vs control was revealed based on training dataset GSE76925 and GSE32537, respectively. Briefly, the significance analysis for the expression of all

https://doi.org/10.2147/JIR.S537967

genes was performed based on log2 fold change (FC) and P value. The selection threshold for DEGs in COPD vs control were $P < 0.05$ and $|logFC| > 0.263$. Meanwhile, the selection threshold for DEGs in IPF vs control were $P < 0.05$ and $|logFC| > 0.263$.[22] Then, the common DEGs (co-DEGs) were explored by using VENN plot analysis. Furthermore, GO function analyses was performed based on the co-DEGs using clusterProfiler package of R. The GO function including biological process (BP), cellular component (CC) and molecular function (MF). The adj $P < 0.05$ was considered as the thresholds for current enrichment analysis.

WGCNA

The analysis started by performing a variance analysis on the expression matrix of HTS samples to identify the top 5000 genes with the highest variability across the samples. Next, WGCNA (version 1.72–5) was applied to detect gene modules with significant co-expression patterns. Initially, the soft threshold was used to convert the adjacency matrix into a continuous scale between 1 and 30, ensuring the resulting network adhered to a power-law distribution, which more accurately reflects the properties of biological networks. In this study, a soft threshold of 0.9 was chosen for network construction, as it was the first instance to meet the required criteria (minModuleSize = 50). Following this, the scale-free network was generated using the blockwiseModules function, and module partition analysis was conducted. For each module, the module membership and gene significance were calculated. Finally, to explore common module genes, the VENN plot analysis was performed on all module genes in both COPD and IPF by using clusterProfiler package of R.

Enrichment Analysis and Signature Gene Investigation

The GO enrichment analysis was performed on common module genes using clusterProfiler package of R. Moreover, glycosylation-related genes were obtained from common genes in GlycoGene Database (GGDB, https: //acgg.asia/ggdb2/) and HGNC (https://www.genenames.org/data/genegroup/#!/group/424), followed by removal of duplicates. The cross genes among DEGs, modules genes in WGCNA and glycosylation-related genes were revealed by using VENN plot analysis. Furthermore, LASSO Cox regression analysis was performed with 10-fold cross-validation using the glmnet package (version 4.1.6) in R. SVM analysis was conducted using the recursive feature elimination (RFE) algorithm in R to analyze co-genes. The Boruta algorithm (version: 8.0.0), a supervised classification feature selection method in R, was utilized to accurately pinpoint all relevant features in current study. The common genes identified by all three algorithms were considered as the signature genes for COPD and IPF, respectively. Finally, the VENN plot analysis was used to reveal key genes based on signature genes for both COPD and IPF.

Validation and Immune Infiltration Analysis

The Wilcoxon signed rank test was performed to assess the differentially expression for key signature genes based on all training datasets and validation dataset. The Receiver Operating Characteristic (ROC) curves and Area Under Curve (AUC) values were calculated based on the expression of SULF1, ST8SIA1, and FCN3 in the GSE76925 (COPD) and GSE32537 (IPF) datasets using the pROC package (version 1.18.5) in R, to evaluate the ability of these genes to distinguish disease from control groups. Then, the infiltration of 28 immune cells was assessed based on gene expression data from the training set, utilizing the ssGSEA algorithm[23] in R. In addition, the correlation between signature genes and 28 immune cells was analyzed by Pearson correlation coefficient using ggplot2 (version: 3.5.0) package of R. The results were visualized using the pheatmap.

Transcriptional Regulation and Drug-Gene Network Analysis

Transcription factors (TFs) regulating diagnostic markers were investigated using the online database hTFtarget. MiRNA-mRNA interactions were scanned using miRWalk 3.0 software,[24] which served as the default database. Additionally, the TransmiR v2.0 database[25] was used to predict the TFs associated with miRNAs in mRNA-miRNA interactions. Then, miRNA-mRNA-TF interactions regulated by the same miRNA were identified and visualized using Cytoscape software. Furthermore, the drug directly related to disease were explored based on Comparative Toxicogenomics Database (CTD) database.[26] Then, the drug-gene interaction network was constructed. The TOP 5 drugs according to the interaction count were visualized by using Cytoscape software.

Patients and Sample Collection

The study cohort included clinically diagnosed patients with Chronic Obstructive Pulmonary Disease (COPD; n = 15) and Idiopathic Pulmonary Fibrosis (IPF; n = 10), along with age- and sex-matched healthy controls (COPD controls: n = 15; IPF controls: n = 10). Peripheral blood samples were collected via venipuncture in EDTA tubes, and lung tissue biopsies were obtained during clinically indicated procedures. All participants provided written informed consent, and this study was approved by the Medical Ethics Committee of Zibo TCM-Integrated Hospital (2023-KY-002). The research was conducted in accordance with the Declaration of Helsinki. Samples were immediately processed for RNA and protein extraction or stored at −80 °C until analysis.

Quantitative Real-Time PCR (qRT-PCR) Analysis

Total RNA was isolated from lung tissues using TRIzol reagent (Invitrogen), followed by quality assessment via Nanodrop spectrophotometry (A260/A280 > 1.8). The cDNA was synthesized from 1 μg RNA using the PrimeScript™ RT Reagent Kit (Takara), and qRT-PCR was performed with SYBR® Premix Ex Taq™ II (Takara) on a LightCycler® 480 system (Roche). The primer sequences for target genes (SULF1, ST8SIA1, FCN3) and the internal control (β-ACTIN) are listed in Supplementary Table 1. Thermal cycling conditions included an initial denaturation at 95°C for 30 sec, followed by 40 cycles of 95°C for 5 sec and 60°C for 30 sec. Relative gene expression was calculated using the 2-ΔΔCT method,[27] with normalization to β-ACTIN.

Our bioinformatics analysis revealed three signature genes including sulfatase 1 (SULF1), ST8 alpha-N-acetyl-neuraminide alpha-2,8-sialyltransferase 1 (ST8SIA1) and ficolin 3 (FCN3), among which ST8SIA1 was prioritized for functional validation due to its divergent expression patterns in COPD (down-regulated) and IPF (up-regulated) ($P < 0.05$). This opposing trend suggests ST8SIA1 may served as a pivotal node in disease-specific pathogenesis. Moreover, as a key regulator of glycosylation, a process critically implicated in lung disease progression, ST8SIA1's functional characterization could elucidate mechanistic links between glycosylation and COPD/IPF phenotypes. ST8SIA1 was prioritized due to its opposing expression in COPD and IPF, implicating disease-specific roles in glycosylation-mediated pathogenesis.[28] Thus, we systematically investigated ST8SIA1 via knockdown/overexpression and pharmacological interventions to uncover its therapeutic potential.

ELISA Detection of Glycosylation Markers

Advanced glycation end products (AGEs) in lung tissue homogenates were quantified using a commercial ELISA kit (YFXEH00915, Nanjing Yifeixue Biotech) according to the manufacturer's protocol. Briefly, samples and standards were incubated in pre-coated 96-well plates for 2 h at 37°C, followed by detection with horseradish peroxidase (HRP)-conjugated antibodies and tetramethylbenzidine (TMB) substrate. Absorbance was measured at 450 nm using a BioTek microplate reader, and AGEs concentrations were interpolated from standard curves. Data were normalized to total protein content determined by BCA assay (Pierce).

Functional Validation in Disease-Specific Cell Models

To investigate the functional role of ST8SIA1, we performed gene overexpression and knockdown experiments in cell models relevant to COPD and IPF, respectively. Based on previous study,[29] A549 alveolar epithelial cells (ATCC, Manassas, VA, USA) were seeded in culture wells and incubated at 37°C for 24 hours to allow cell adherence. Subsequently, 10% cigarette smoke extract (CSE) was added to the wells at a density of 10^6 cells/mL per well, and the cells were further incubated at 37°C for 12 hours to induce cellular injury. This 12 - hour exposure of adherent A549 cells to 10% CSE was designed to mimic the pulmonary cell damage observed in chronic obstructive pulmonary disease (COPD). Following successful establishment of the COPD cell model, ST8SIA1 overexpression plasmid (GenScript, Piscataway, NJ, USA) was transfected using Lipofectamine™ 3000 reagent (Thermo Fisher Scientific, Waltham, MA, USA). For IPF, primary human lung fibroblasts (Lonza, Basel, Switzerland) were transfected with small interfering RNA (siRNA) targeting ST8SIA1 (Invitrogen, Carlsbad, CA, USA). Transfection was carried out according to the manufacturer's instructions.[29]

Cellular Functional Assays

The functional impact of ST8SIA1 modulation was assessed through proliferation, migration, invasion, and apoptosis assays. For proliferation analysis, transfected cells were seeded in 96-well plates, and viability was measured at 24 and 72 h using the CCK-8 assay (Dojindo). Cell migration and invasion were evaluated using Transwell chambers (Corning), with or without Matrigel coating, respectively. After incubation, migrated/invaded cells were fixed with 4% paraformaldehyde, stained with 0.1% crystal violet, and quantified by counting five random fields per well under a microscope. Apoptosis was analyzed using Annexin V-FITC/PI staining (BD Biosciences) followed by flow cytometry (BD FACSCanto™). Briefly, cells were harvested, washed with PBS, and resuspended in binding buffer containing Annexin V-FITC and PI. After incubation for 15 min in the dark, apoptotic cells were quantified using FlowJo software (version 10.8.1).

Statistical Analysis

The statistical analysis of bioinformatics were performed using R software. All statistical analyses were two-sided and $P < 0.05$ indicated statistically significant. The analytical and statistical software for the experiments was ImageJ and Graphpad prism.

Results

DEGs Investigation and Enrichment Analysis

The differentially expression analysis revealed totally 2561 up-regulated genes and 2582 down-regulated genes between COPD samples and control samples (Figure 2A). In addition, there were a total of 2252 up-regulated genes and 1778 down-regulated genes between IPF samples and control samples (Figure 2B). The VENN plot analysis on DEGs for both COPD and IPF revealed totally 840 co-DEGs (Figure 2C). Furthermore, the GO enrichment analysis on co-DEGs showed that these genes were mainly assembled in functions like chemotaxis (BP, GO:0006935), collagen-containing extracellular matrix (CC, GO:0062023) and glycosaminoglycan binding (MF, GO:0005539). The TOP 10 enrichment results for GO-BP, GO-CC and GO-MF was showed in Figure 2D.

WGCNA Investigation

The hierarchical clustering results based on the TOP5000 genes demonstrated good data quality and reliability, providing a solid foundation for subsequent analysis (Figure 3A). Then, WGCNA analysis was conducted using a soft-threshold of 4 and a fitting degree of 0.9 (Figure 3B). Thus, 13 modules were obtained based on the combination results of dynamic tree cutting (Figure 3C). The correlation between module and different groups was visualized with a heatmap (Figure 3D). The analysis revealed that the turquoise module (r =- 0.33, $P < 0.001$), blue module (r = −0.34, $P < 0.001$) and pink module (r = 0.3, $P < 0.001$) exhibited the strongest associations with COPD (totally 3824 module genes). The WCGNA analysis was also performed on genes in IPF datasets (Figure 3E–G). The analysis revealed that turquoise module (r = −0.66, $P < 0.001$), blue module (r = −0.74, $P < 0.001$) and magenta module (r = −0.56, $P < 0.001$) had the strongest positive correlations with IPF (totally 3265 module genes) (Figure 3H). Finally, the VENN plot analysis revealed totally 675 co-genes between 3824 COPD module genes and 3265 IPF module genes (Supplementary Figure 2), which used for following analysis.

Enrichment Analysis and Signature Gene Investigation

GO function analysis performed co-genes showed that these genes were mainly assembled in functions like cell junction assembly (BP, GO:0034329), collagen-containing extracellular matrix (CC, GO:0062023) and extracellular matrix structural constituent (MF, GO:0005201) (Supplementary Figure 3). VENN plot analysis that performed on 675 module genes, 840 co-DEGs and 1898 glycosylation-related genes obtained from GGDB totally revealed 35 cross genes (Figure 4A). Then, three advanced machine learning algorithms were employed to analyze the 35 key genes. The result showed that there were totally nine key genes in COPD (Figure 4B–H) and 10 key genes in IPF (Figure 4I–O) revealed by three machine learning algorithms. In addition, the VENN plot analysis revealed totally three genes

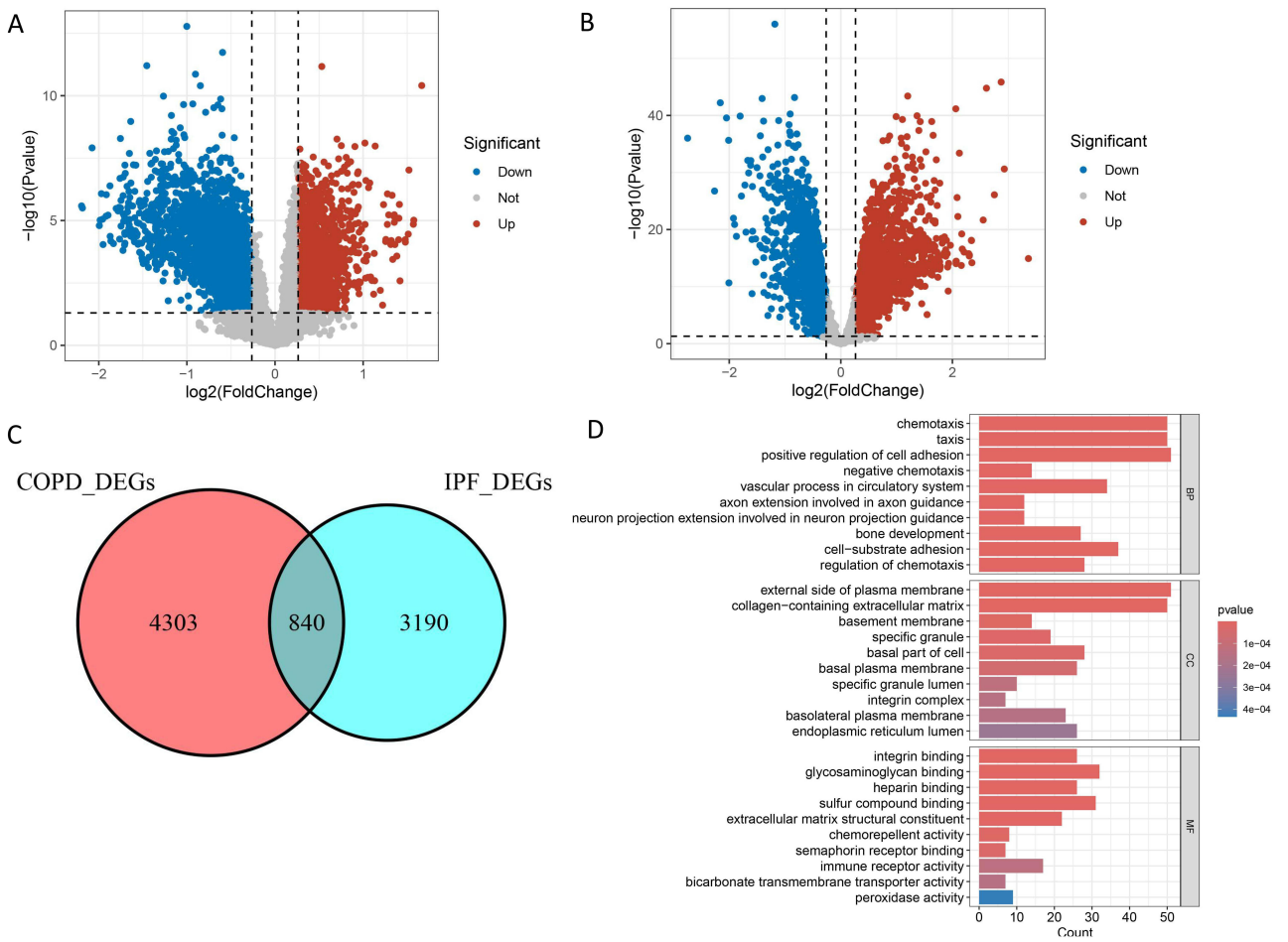

Figure 2 The differentially expressed genes (DEGs) investigation and enrichment analysis. (**A**) the volcano plot showed the DEGs between chronic obstructive pulmonary disease (COPD) samples and control samples. (**B**) the volcano plot showed the DEGs between idiopathic pulmonary fibrosis (IPF) samples and control samples. (**C**) the VENN plot analysis showed the common-DEGs (co-DEGs) between COPD and IPF. (**D**) the GO enrichment analysis showed the TOP 10 BP, CC and MF function significantly assembled by co-DEGs.

including SULF1, ST8SIA1 and FCN3, which were deemed signature genes for investigation into COPD and IPF (Figure 4P).

Validation Analysis for Signature Genes

The evaluation analysis was performed on three signature genes. The results showed that SULF1 and FCN3 were significantly up-regulated, but ST8SIA1 was dramatically down-regulated in COPD group when compared to control group in training dataset (Figure 5A). However, the difference for both three signature genes between two groups in validation dataset was not significant (Figure 5B). Moreover, SULF1 and ST8SIA1 were both significantly up-regulated in IPF group when compared to control group either in training dataset or validation dataset (Figure 5C and D). Compared with the control group, FCN3 was significantly down-regulated in the idiopathic pulmonary fibrosis (IPF) group of the training set, but it showed a significant up-regulation in the IPF group of the validation set (Figure 5C and D). Based on the expression levels of SULF1, ST8SIA1, and FCN3 in the GSE76925 (COPD) and GSE32537 (IPF) datasets, ROC curves were plotted and AUC values were calculated. The results revealed that in COPD, the AUC values for ST8SIA1 and SULF1 were greater than 0.7 (Supplementary Figure 4A). In IPF, the AUC values for FCN3 and SULF1 both exceeded 0.9, while the AUC for ST8SIA1 was 0.792 (Supplementary Figure 4B).

https://doi.org/10.2147/JIR.S537967

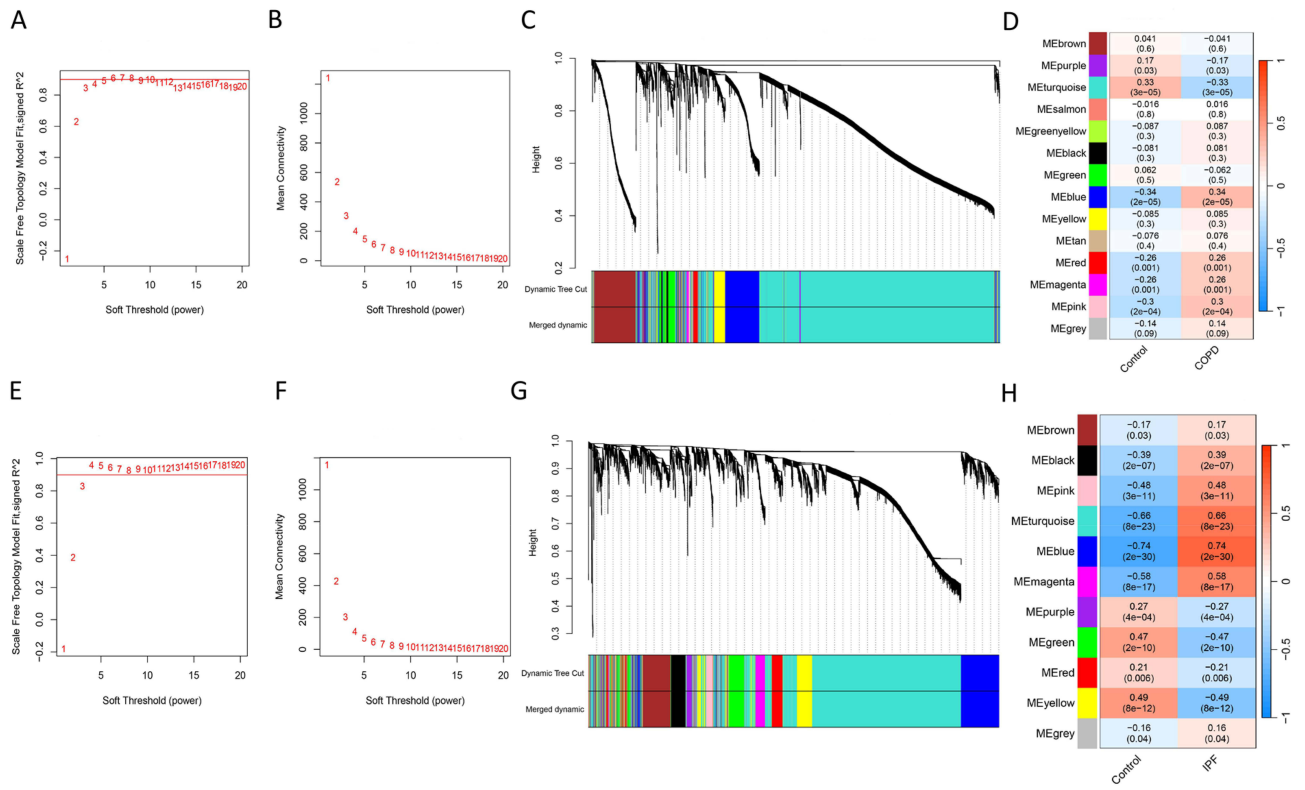

Figure 3 The result of WCGNA analysis. (**A**) the hierarchical clustering analysis on COPD dataset based on the TOP5000 genes. (**B**) the scale free soft-threshold distribution on COPD dataset. (**C**) clustering analysis for models on COPD dataset: dynamic tree cut and merged dynamic represented the module before and after the merge module; the different colors in the figure represented different modules. (**D**) the heatmap for correlation between modules and COPD traits. (**E**) the hierarchical clustering analysis on IPF dataset based on the TOP5000 genes. (**F**) the scale free soft-threshold distribution on IPF dataset. (**G**) clustering analysis for models on IPF dataset: dynamic tree cut and merged dynamic represented the module before and after the merge module; the different colors in the figure represented different modules. (**H**) the heatmap for correlation between modules and IPF traits.

Immune Infiltration Analysis

The ssGSEA analysis was performed on 28 immune cells to reveal the difference of expression between disease group and normal group. The result showed that when compared to control group, activated B cell, was significantly up-regulated in both COPD group (Figure 6A) and IPF group (Figure 6B) (all $P < 0.05$). Moreover, the immune correlation analysis was performed based on all signature genes and 28 immune cells. The result showed that immune cell activated B cell was significantly positively correlated with SULF1, but significantly negatively correlated with FCN3 in COPD (all $P < 0.05$) (Figure 6C). In addition, three signature genes were all dramatically correlated with immune cells such as activated B cell and neutrophil cell in IPF (all $P < 0.05$), which collectively contribute to the immune microenvironment (Figure 6D).

Transcriptional Regulation and Drug Network Analysis

A total of 107 TF and 16 miRNAs were revealed in current study. Then, combined with three signature genes, a transcriptional regulation network was constructed. The result showed that there were 162 interactions and 126 nodes in current network (Figure 7A). Furthermore, based on CTD database, a drug-gene interaction network was established with 3 signature genes and 10 drugs (Figure 7B). The result showed that Bisphenol A and Valproic acid were common drugs for both three signature genes, indicated that these two drugs may potentially play crucial roles in the biological processes related to these signature genes.

Dysregulation of Glycosylation-Related Genes in COPD and IPF

The qRT-PCR analysis revealed that ST8SIA1 expression was significantly downregulated in COPD patients compared to healthy controls ($P < 0.001$), while FCN3 expression was upregulated ($P < 0.001$) (Figure 8A). In contrast, IPF

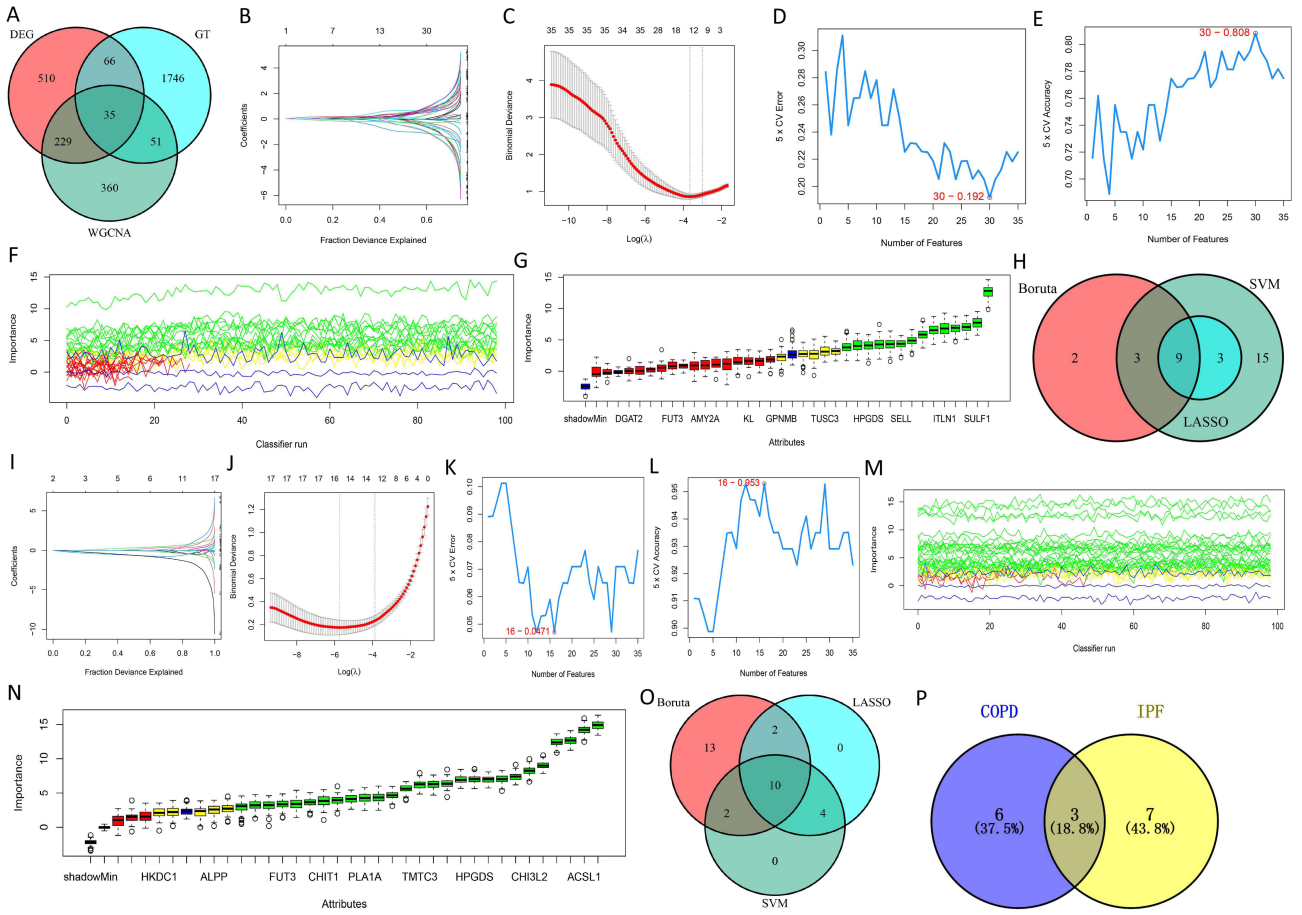

Figure 4 Enrichment analysis and signature genes investigation. (**A**) VENN plot revealed the cross genes among co-DEGs, module genes and glycosylation-related genes. (**B** and **C**) LASSO Cox analysis revealed optimal genes for COPD: the Y-axis in the (**B**) represented the coefficient of the variable, while the X-axis represented the value of log (lambda); the two dotted lines in (**C**) represented two special lambda values: lambda.min on the left and lambda.1se on the right; the lambda between these two values was considered appropriate. (**D**) the accuracy for SVM-RFE analysis for COPD. (**E**) the error rate for SVM-RFE analysis for COPD. (**F** and **G**) Boruta algorithm revealed the optimal genes for COPD. (**H**) the VENN plot analysis explored the common genes for COPD based on three algorithms. (**I** and **J**) LASSO Cox analysis revealed optimal genes for IPF. (**K** and **L**) SVM-RFE analysis revealed optimal genes for IPF. (**M** and **N**) Boruta algorithm revealed the optimal genes for IPF. (**O**) the VENN plot analysis explored the common genes for IPF based on three algorithms. (**P**) the VENN plot analysis revealed signature genes for both COPD and IPF.

patients exhibited significant upregulation of both ST8SIA1 and SULF1 ($P < 0.001$) (Figure 8A). ELISA further demonstrated an inverse correlation between ST8SIA1 expression and AGEs levels in COPD ($r = -0.75$, $P < 0.05$), whereas a positive correlation was observed in IPF ($r = 0.82$, $P < 0.05$) (Figure 8B). These results indicate that ST8SIA1 exhibited disease-specific expression patterns and was closely associated with glycosylation alterations in both COPD and IPF.

ST8SIA1 Influenced Cell Proliferation, Apoptosis, and Metastatic Potential in COPD and IPF

Functional characterization revealed that ST8SIA1 overexpression in COPD significantly enhanced cellular proliferation (CCK-8 assay, 24 h and 72 h, $P < 0.01$) (Figure 8C), while markedly increasing both migration (Transwell, $P < 0.01$) and invasion (Transwell, $P < 0.01$) capacities (Figure 8D). Conversely, ST8SIA1 knockdown in IPF substantially attenuated cell proliferation ($P < 0.01$) (Figure 8E), migratory ($P < 0.01$), and invasive ($P < 0.01$) capabilities (Figure 8F). Additionally, overexpression of ST8SIA1 reduced the apoptosis rate of COPD cells ($P < 0.01$), and knockout of ST8SIA1 increased the apoptosis rate of IPF cells ($P < 0.01$) (Figure 8G). These findings demonstrated that ST8SIA1 differentially regulated proliferative, apoptotic, and metastatic phenotypes in COPD and IPF pathogenesis, suggesting its disease-specific functional roles in disease progression.

https://doi.org/10.2147/JIR.S537967

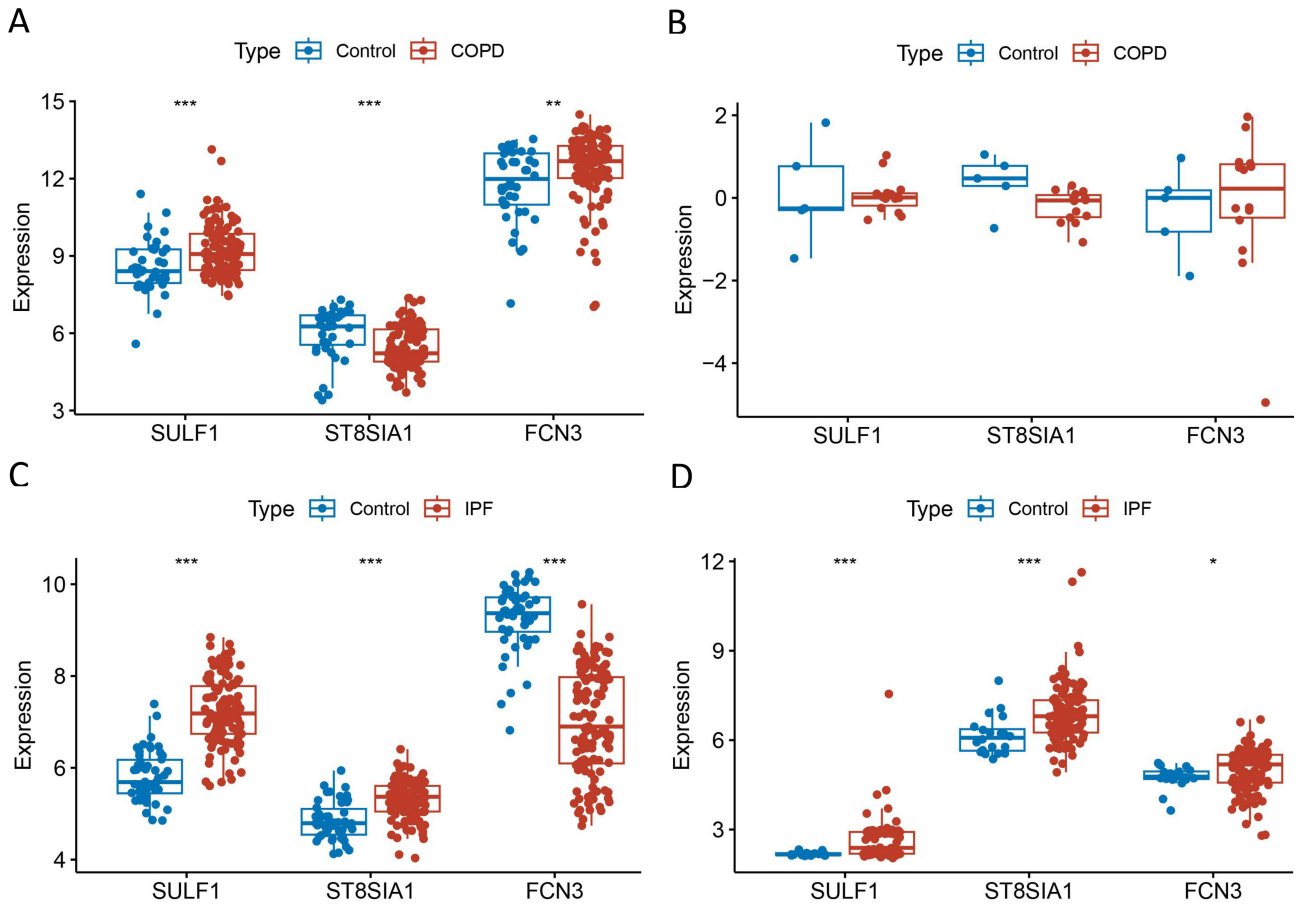

Figure 5 The evaluation analysis for signature genes in both COPD and IPF. (**A** and **B**) the validation analysis for signature genes in COPD based on training dataset and validation dataset: box plot showed the expression of signature genes between COPD group and control group. (**C** and **D**) the validation analysis for signature genes in IPF based on training dataset and validation dataset: box plot showed the expression of signature genes between IPF group and control group. *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$.

Discussion

COPD and IPF are two common and severe respiratory diseases, which share similarities in immune responses and fibrosis progression.[30] Glycosylation, as an important post-translational modification, has been shown to play a key role in the onset and progression of various fibrotic diseases. However, the specific role of glycosylation and its underlying molecular mechanisms in COPD and IPF remain insufficiently explored. This study represented the a integrative exploration of glycosylation-related biomarkers shared by COPD and IPF, combining multi-omics bioinformatics analyses with experimental validation. We identified SULF1, ST8SIA1, and FCN3 as signature genes dysregulated in both diseases, implicating aberrant glycosylation in their pathogenesis.

GO enrichment analysis of the DEGs shared between COPD and IPF revealed that 840 common genes were significantly enriched in several key biological functions, including chemotaxis, collagen-containing extracellular matrix, and glycosaminoglycan binding. This provides important clues for understanding the shared pathological mechanisms between COPD and IPF. Specifically, in COPD and IPF, the enrichment of chemotaxis-related genes may reflect the recruitment of immune cells, driving chronic inflammation and fibrosis. For instance, the upregulation of chemokines such as CCL18 and CCL17 is associated with disease progression.[31,32] The enrichment of ECM-related genes suggests an imbalance in collagen deposition. In COPD, ECM dysregulation can lead to lung tissue remodeling, affecting all lung compartments (ie, airway wall fibrosis and emphysema).[33] In IPF, a persistent myofibroblastic phenotype drives excessive ECM deposition and aberrant lung repair, resulting in tissue scarring, distorted alveolar architecture, and irreversible loss of lung function.[34] Glycosaminoglycans are essential components of the ECM. Among them,

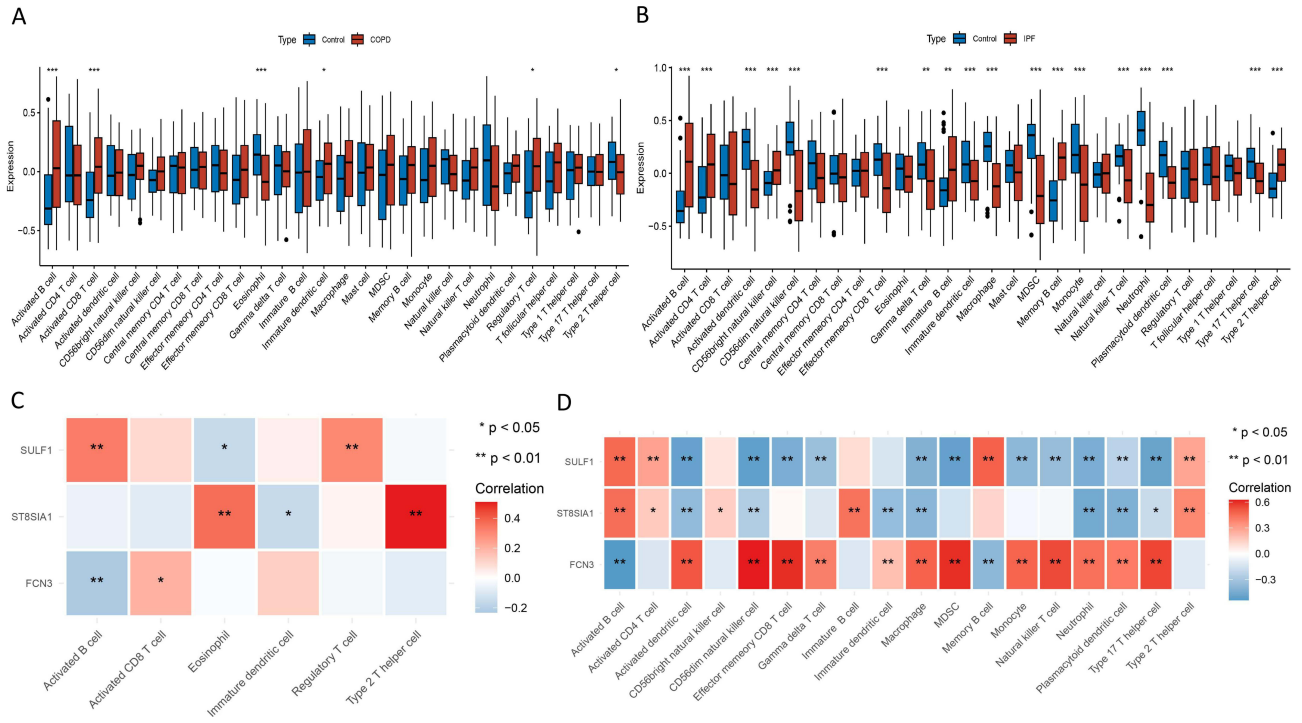

Figure 6 The immune correlation analysis for signature genes. (**A**) the heatmap analysis for the expression of 28 immune cells between COPD group and control group. (**B**) the heatmap analysis for the expression of 28 immune cells between IPF group and control group. (**C**) the correlation between signature genes and immune cells based on COPD dataset. (**D**) the correlation between signature genes and immune cells based on IPF dataset. *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$.

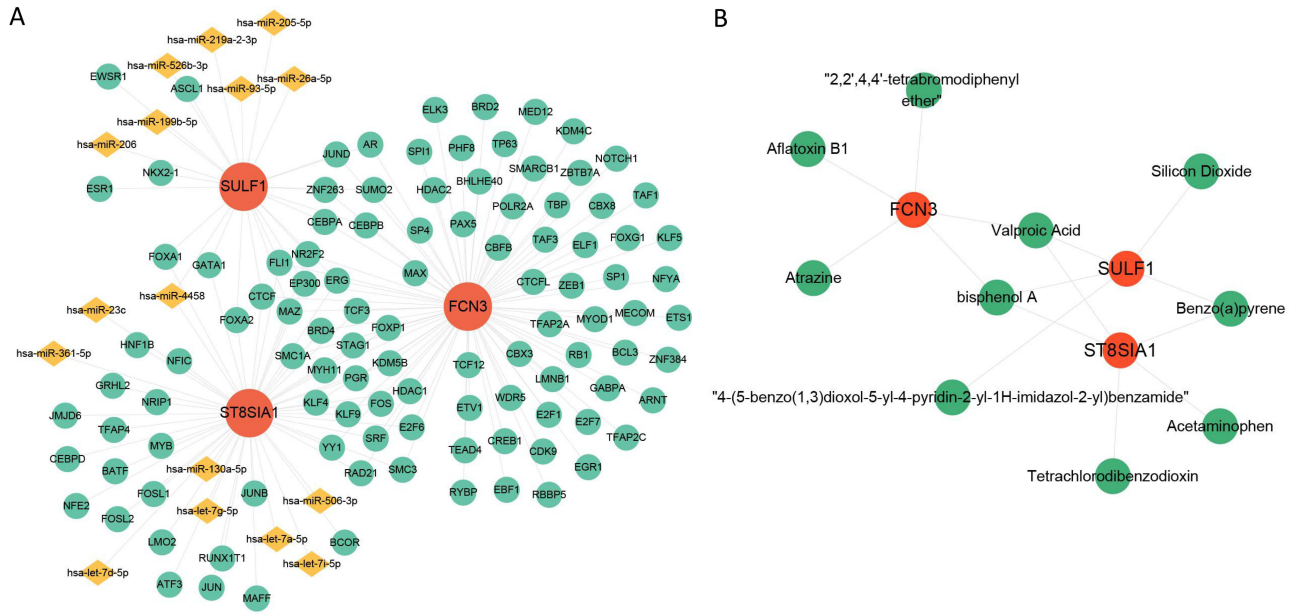

Figure 7 The transcription regulatory network and drug-gene interaction network analysis. (**A**) the transcription regulatory network constructed by 3 signature genes, 107 TF, 16 miRNAs and 162 interactions. (**B**) a drug-gene interaction network established with 3 signature genes and 10 drugs.

chondroitin sulfate is a major Glycosaminoglycans that binds to multiple factors within the ECM and participates in tissue remodeling across various diseases.[35] In summary, the enrichment in key biological functions such as chemotaxis, collagen matrix binding, and glycosaminoglycan binding reveals shared pathological mechanisms including immune cell

https://doi.org/10.2147/JIR.S537967

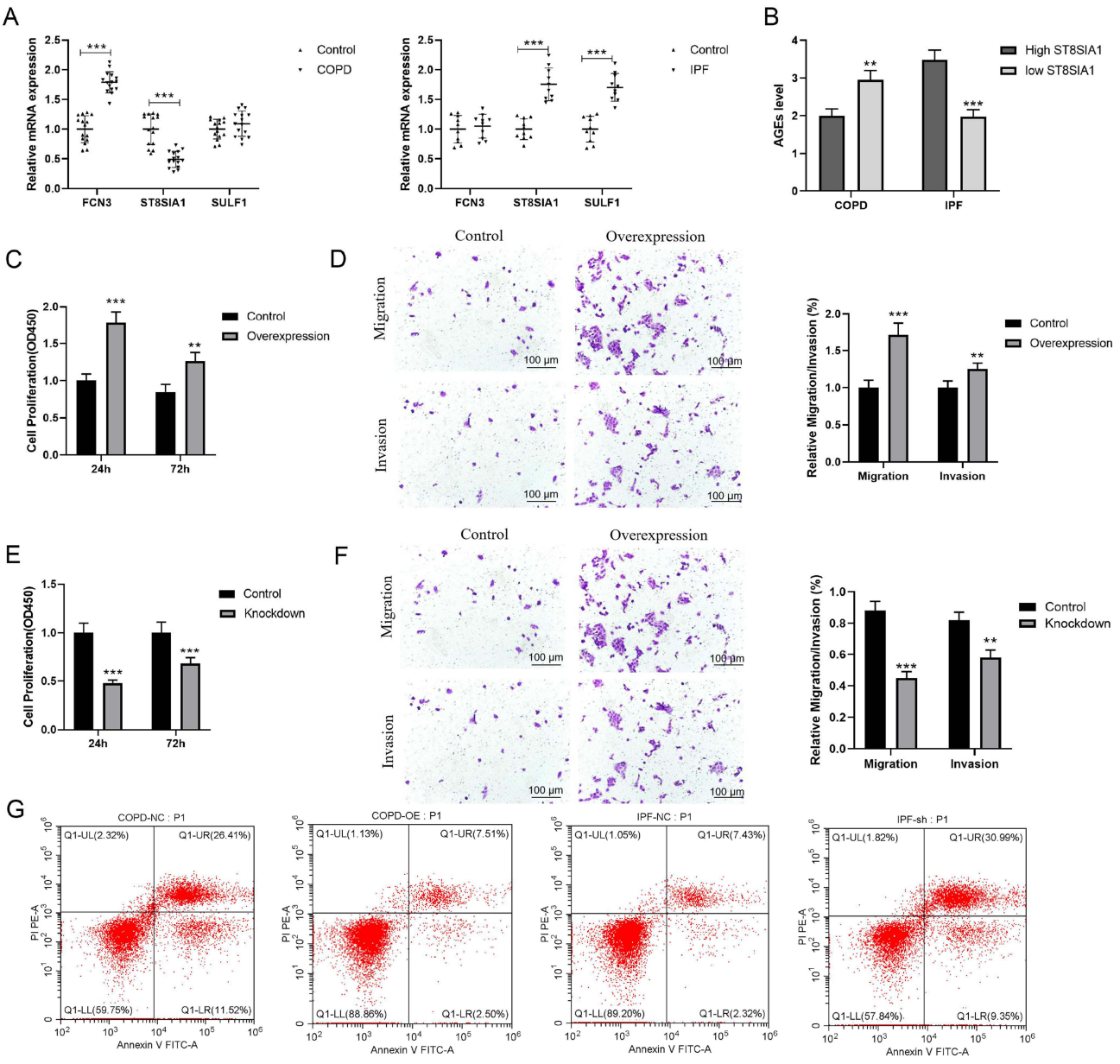

Figure 8 Dysregulation of glycosylation-related genes and functional roles of ST8SIA1 in COPD and IPF. (**A**) qRT-PCR analysis showed the expression levels of *ST8SIA1*, *FCN3*, and *SULF1* in COPD patients, IPF patients, and healthy controls. *ST8SIA1* expression was significantly downregulated in COPD patients, while *FCN3* was upregulated compared to healthy controls. In contrast, IPF patients exhibited significant upregulation of both *ST8SIA1* and *SULF1*. (**B**) ELISA results demonstrated the correlation between *ST8SIA1* expression and AGEs levels in COPD and IPF. An inverse correlation was observed in COPD, whereas a positive correlation was found in IPF. (**C**–**G**) functional characterization of *ST8SIA1* in COPD and IPF. Overexpression of *ST8SIA1* in COPD significantly enhanced cellular proliferation, migration, and invasion capabilities, while reducing apoptosis rates. Conversely, *ST8SIA1* knockdown in IPF substantially attenuated cell proliferation, migration, and invasion, and increased apoptosis rates. All data are presented as mean ± SD. **, $P < 0.01$ vs respective control; ***, $P < 0.001$ vs respective control. Scale bars: 100 μm.

recruitment-driven chronic inflammation, dysregulated collagen deposition, and ECM remodeling. These findings provide important insights into understanding the common molecular pathways shared by these two diseases.

SULF1 is a key enzyme involved in modulating the activity of heparan sulfate, which in turn regulates the function of growth factors and cytokines critical for tissue remodeling and inflammation.[36] The up-regulation of SULF1 with prior studies that suggested its role in promoting fibroblast proliferation and extracellular matrix (ECM) deposition, which are hallmarks of diseases.[37] A previous study found that SULF1 may serve as a potential biomarker for identifying smokers at risk of COPD, and its role may be related to airway wall thickening and fibrosis, thereby contributing to the characteristic progressive airflow limitation in COPD.[38] SULF1 is upregulated in IPF and primarily increased in

fibroblasts. SULF1 promotes pulmonary fibrosis through the TGF-β1/SMAD pathway, suggesting that targeting SULF1 may offer a promising therapeutic strategy for IPF.[39] Interestingly, in our study, the positive correlation between SULF1 and activated B cells in both diseases could indicate a role in modulating immune responses. B cells are integral to the pathogenesis of autoimmune and fibrotic diseases, suggesting that SULF1 may not only influence tissue remodeling but also affect immune cell recruitment and activation. Targeting SULF1 could thus have dual benefits: inhibiting fibrosis while modulating aberrant immune responses. ST8SIA1 exerts antitumor effects in bladder cancer by inhibiting the JAK/STAT signaling pathway, making it a novel diagnostic and therapeutic target for this disease[40]. Notably, ST8SIA1 is a sialyltransferase enzyme, plays a pivotal role in the biosynthesis of sialic acid,[41] with our functional assays revealing its disease-specific impacts. In COPD, ST8SIA1 overexpression promoted epithelial proliferation and migration, reduced apoptosis, and decreased AGEs, likely through sialylation of mucins and growth factor receptors (eg, EGFR). In IPF, its knockdown suppressed fibroblast activation, increased apoptosis, and elevated AGEs, suggesting ST8SIA1 drives fibrosis via glycosylation-dependent ECM crosslinking. The observed anti-apoptotic effects of ST8SIA1 in COPD and pro-apoptotic effects in IPF highlight its dual role in regulating cell survival, which may contribute to disease-specific pathological remodeling. FCN3 is a pattern recognition receptor that participates in the innate immune response, particularly in the activation of the lectin-complement pathway.[42] FCN3 may serve as a potential biomarker for prognosis and treatment of various diseases, including lung squamous cell carcinoma,[43] heart failure,[44] and ischemic cardiomyopathy.[45] For example, high expression of FCN3 in lung squamous cell carcinoma is associated with poor prognosis and correlates with immune cell infiltration, immune-related pathways, and immune-related molecules.[43] In this study, FCN3 was up-regulated in COPD training set but down-regulated in IPF training set. This dichotomy suggested that FCN3 may exert opposing effects in these diseases by modulating immune responses. In COPD, the up-regulation of FCN3 could be a compensatory response to chronic inflammation, potentially dampening excessive immune activation and limiting tissue damage. Conversely, the down-regulation of FCN3 in IPF may impair the innate immune response, thereby promoting persistent inflammation and fibrosis. The observation that FCN3 is negatively correlated with activated B cells in COPD provides additional evidence that FCN3 may play a role in modulating immune cell activation and immune suppression in chronic inflammatory diseases. This finding aligns with previous research showing that immune dysregulation is a key feature of fibrosis-related immune landscape.[46] The opposing regulation of FCN3 in these diseases highlights its potential as a biomarker for distinguishing between different stages or phenotypes of fibrotic lung disease. A contradiction that we cannot overlook is the inconsistent expression trend of FCN3 between the IPF training set (GSE32537) and the validation set (GSE70866). The reasons for this discrepancy are as follows. First, GSE32537 and GSE70866 originate from different research platforms and populations. GSE32537 includes 119 IPF cases and 50 controls, whereas GSE70866 comprises 112 IPF cases and 20 controls. Differences in sample size and control proportion may have influenced the results. Second, IPF is a heterogeneous disease, and different subtypes or disease stages may exhibit distinct gene expression patterns. FCN3 might be upregulated under certain conditions (eg, during the inflammatory phase) but downregulated under others. In the future, we will utilize larger sample datasets to validate the expression of FCN3. Thus, the three glycosylation-related genes identified in this study offered promising insights into the shared pathophysiology of COPD and IPF. Their differential expression patterns not only enhanced our understanding of the molecular underpinnings of these diseases, but also suggested novel biomarkers for both diagnosis and prognosis.

Glycosylation is a post-translational modification that critically influences various biological processes, including cell signaling, immune modulation, and ECM remodeling.[47] Dysregulated glycosylation appears to play a significant role in the pathogenesis of COPD and IPF.[9,48] Our findings suggested that altered glycosylation patterns, driven by genes such as SULF1, ST8SIA1, and FCN3, may contribute to key disease features such as airway obstruction in COPD and fibrosis in IPF. In COPD, changes in mucin glycosylation could promote mucus hypersecretion and impaired mucociliary clearance, while in IPF, altered glycosylation of ECM proteins may drive excessive matrix deposition and fibrosis.[11,49] The immune infiltration analysis revealed that activated B cells were significantly upregulated in both diseases, further implicating glycosylation-related genes in modulating the immune microenvironment. The correlation of SULF1 with activated B cells suggested that it may contribute to the inflammatory cascade, whereas the negative correlation of FCN3 could point to an immunosuppressive role. Our validation experimental data confirmed that ST8SIA1-mediated

sialylation directly modulates AGEs levels, with opposing effects in COPD (reduced AGEs) and IPF (elevated AGEs). This suggests that targeting glycosylation enzymes like ST8SIA1 could normalize aberrant glycoprotein profiles-reducing mucus viscosity in COPD or ECM stiffness in IPF.

This study is the first to integrate multi-omics bioinformatics with functional validation of glycosylation-related genes in both COPD and IPF, revealing ST8SIA1 as a dual-role regulator through disease-specific assays. However, several limitations must be acknowledged. The relatively small sample size of the clinical validation experiments may limit the generalizability of the findings. Additionally, the lack of in vivo validation experiments and the use of primary human bronchial epithelial cells (HBECs) are notable limitations. Future studies should employ primary human bronchial epithelial cells (HBECs) differentiated at the air-liquid interface (ALI) (eg, Lonza) to better reflect in vivo airway physiology. These biomarkers should also be further validated in larger cohorts to assess their clinical applicability. Moreover, while our study focused on glycosylation-related genes, other post-translational modifications, such as phosphorylation and acetylation, may also play critical roles in the pathogenesis of COPD and IPF. Integrating proteomic and metabolomic data would offer a more comprehensive view of the molecular landscape of these diseases.

Conclusion

Our study suggests that ST8SIA1 may serve as potential glycosylation-related biomarkers shared between COPD and IPF. This gene appear to connect fibrosis, inflammation, and immune dysregulation pathways, warranting further investigation for their diagnostic and therapeutic implications. Moreover, the findings of this study unveil pivotal directions for future investigation. Regarding diagnostics, the expression signatures of SULF1, ST8SIA1, and FCN3 hold promise for developing more sensitive and specific non-invasive diagnostic tools. Therapeutically, they provide crucial clues for designing novel interventions targeting glycosylation pathways. Furthermore, unraveling how these signature genes precisely regulate ECM remodeling, immune cell activation, and cellular behaviors (proliferation, apoptosis, migration) through glycosylation modifications will substantially deepen our understanding of the shared pathogenesis of COPD and IPF. This mechanistic insight will also establish a conceptual framework for developing precision therapies based on glycosylation-based interventions.

Abbreviations

COPD, chronic obstructive pulmonary disease; IPF, idiopathic pulmonary fibrosis; WGCNA, weighted gene co-expression network analysis; ECM, extracellular matrix; IL-8, interleukin-8; GEO, Gene Expression Omnibus; FC, fold change; BP, biological process; CC, cellular component; MF, molecular function; CTD, Comparative Toxicogenomics Database; qRT-PCR Quantitative real-time PCR; PBMCs, peripheral blood mononuclear cells; AGEs, Advanced glycation end products; HRP, horseradish peroxidase; TMB, tetramethylbenzidine.

Funding

This research was funded by the Zibo Medical and Health Research Project (20231502096).

Disclosure

The authors report no conflicts of interest in this work.

References

1. Christenson SA, Smith BM, Bafadhel M, Putcha N. Chronic obstructive pulmonary disease. *Lancet*. 2022;399(10342):2227–2242. doi:10.1016/s0140-6736(22)00470-6
2. Sankari A, Chapman K, Ullah S. Idiopathic pulmonary fibrosis. In: *StatPearls*. StatPearls Publishing StatPearls Publishing LLC; 2025.
3. van Moorsel CHM. Trade-offs in aging lung diseases: a review on shared but opposite genetic risk variants in idiopathic pulmonary fibrosis, lung cancer and chronic obstructive pulmonary disease. *Curr Opin Pulm Med*. 2018;24(3):309–317. doi:10.1097/mcp.0000000000000476
4. De Vellis C, Pietrobono S, Stecca B. The role of glycosylation in melanoma progression. *Cells*. 2021;10(8):2136. doi:10.3390/cells10082136
5. Anghelescu GDC, Mernea M, Mihăilescu DF. Mapping O- and N-glycosylation in transmembrane and interface regions of proteins: insights from a database search study. *Int J Mol Sci*. 2025;26(1):327. doi:10.3390/ijms26010327
6. Drzewicka K, Zasłona Z. Metabolism-driven glycosylation represents therapeutic opportunities in interstitial lung diseases. *Front Immunol*. 2024;15:1328781. doi:10.3389/fimmu.2024.1328781

7. Huang Y, Song J, Chen J, Li F, Zhang Y, Yang JH. The significance of protein N-glycosylation in the pathogenesis of lung cancer and its clinical implications. *Cancer Lett*. 2025;631:217849. doi:10.1016/j.canlet.2025.217849
8. Venkatakrishnan V, Thomsson KA, Padra M, et al. Protein N-glycosylation in the bronchoalveolar space differs between never-smokers and long-term smokers with and without COPD. *Glycobiology*. 2023;33(12):1128–1138. doi:10.1093/glycob/cwad071
9. Cabrera Cesar E, Lopez-Lopez L, Lara E, et al. Serum biomarkers in differential diagnosis of idiopathic pulmonary fibrosis and connective tissue disease-associated interstitial lung disease. *J Clin Med*. 2021;10(14):3167. doi:10.3390/jcm10143167
10. Hewitt RJ, Puttur F, Gaboriau DCA, et al. Lung extracellular matrix modulates KRT5(+) basal cell activity in pulmonary fibrosis. *Nat Commun*. 2023;14(1):6039. doi:10.1038/s41467-023-41621-y
11. Bechtella L, Chunsheng J, Fentker K, et al. Ion mobility-tandem mass spectrometry of mucin-type O-glycans. *Nat Commun*. 2024;15(1):2611. doi:10.1038/s41467-024-46825-4
12. Cergan R, Berghi O, Dumitru M, et al. Interleukin 8 molecular interplay in allergic rhinitis and chronic rhinosinusitis with nasal polyps: a scoping review. *Life*. 2025;15(3). doi:10.3390/life15030469
13. Norheim KL, Ben Ezra M, Heckenbach I, et al. Effect of nicotinamide riboside on airway inflammation in COPD: a randomized, placebo-controlled trial. *Nat Aging*. 2024;4(12):1772–1781. doi:10.1038/s43587-024-00758-1
14. Yang L, Xia H, Gilbertsen A, et al. IL-8 concurrently promotes idiopathic pulmonary fibrosis mesenchymal progenitor cell senescence and PD-L1 expression enabling escape from immune cell surveillance. *Am J Physiol Lung Cell Mol Physiol*. 2023;324(6):L849–l862. doi:10.1152/ajplung.00028.2023
15. Huang S, Thomsson KA, Jin C, et al. Truncated lubricin glycans in osteoarthritis stimulate the synoviocyte secretion of VEGFA, IL-8, and MIP-1α: interplay between O-linked glycosylation and inflammatory cytokines. *Front Mol Biosci*. 2022;9:942406. doi:10.3389/fmolb.2022.942406
16. Zhou J, Tan Y, Wu W, et al. Plasma IgG glycosylation profiling reveals the biological features of early chronic obstructive pulmonary disease. *J Proteome Res*. 2025;24(4):1804–1816. doi:10.1021/acs.jproteome.4c00819
17. Xie X, Kong S, Cao W. Targeting protein glycosylation to regulate inflammation in the respiratory tract: novel diagnostic and therapeutic candidates for chronic respiratory diseases. *Front Immunol*. 2023;14:1168023. doi:10.3389/fimmu.2023.1168023
18. Krick S, Helton ES, Easter M, et al. ST6GAL1 and α2-6 sialylation regulates IL-6 expression and secretion in chronic obstructive pulmonary disease. *Front Immunol*. 2021;12:693149. doi:10.3389/fimmu.2021.693149
19. Yang S, Xia J, Yang Z, Xu M, Li S. Lung cancer molecular mutations and abnormal glycosylation as biomarkers for early diagnosis. *Cancer Treatment Res Commun*. 2021;27:100311. doi:10.1016/j.ctarc.2021.100311
20. Balbisi M, Sugár S, Turiák L. Protein glycosylation in lung cancer from a mass spectrometry perspective. *Mass Spectrom Rev*. 2024. doi:10.1002/mas.21882
21. Liu S, Wang Z, Zhu R, Wang F, Cheng Y, Liu Y. Three differential expression analysis methods for RNA sequencing: limma, EdgeR, DESeq2. *J Visualiz Exper*. 2021;(175). doi:10.3791/62528
22. Zang JCS, Hohoff C, Van Assche E, et al. Immune gene co-expression signatures implicated in occurence and persistence of cognitive dysfunction in depression. *Prog Neuropsychopharmacol Biol Psychiatry*. 2023;127:110826. doi:10.1016/j.pnpbp.2023.110826
23. Xiao B, Liu L, Li A, et al. Identification and verification of immune-related gene prognostic signature based on ssGSEA for osteosarcoma. *Front Oncol*. 2020;10:607622. doi:10.3389/fonc.2020.607622
24. Dweep H, Gretz N. miRWalk2.0: a comprehensive atlas of microRNA-target interactions. *Nat Methods*. 2015;12(8):697. doi:10.1038/nmeth.3485
25. Tong Z, Cui Q, Wang J, Zhou Y. TransmiR v2.0: an updated transcription factor-microRNA regulation database. *Nucleic Acids Res*. 2019;47(D1):D253–d258. doi:10.1093/nar/gky1023
26. Mattingly CJ, Colby GT, Forrest JN, Boyer JL. The comparative toxicogenomics database (CTD). *Environ Health Perspect*. 2003;111(6):793–795. doi:10.1289/ehp.6028
27. ST LKJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-delta delta C(T)) method. *Methods*. 2001;25(4):402–408. doi:10.1006/meth.2001.1262
28. Chopra A, Mueller R, Weiner J, et al. BACH1 binding links the genetic risk for severe periodontitis with ST8SIA1. *J Dental Res*. 2022;101(1):93–101. doi:10.1177/00220345211017510
29. Zhou L, Jian T, Wan Y, et al. Luteolin alleviates oxidative stress in chronic obstructive pulmonary disease induced by cigarette smoke via modulation of the TRPV1 and CYP2A13/NRF2 signaling pathways. *Int J Mol Sci*. 2023;25(1):369. doi:10.3390/ijms25010369
30. Mahler DA, Petrone RA, Krocker DB, Cerasoli F. A perspective on web-based information for patients with chronic lung disease. *Ann Am Thoracic Soc*. 2015;12(7):961–965. doi:10.1513/AnnalsATS.201502-104PS
31. Dilektasli AG, Demirdogen Cetinoglu E, Uzaslan E, et al. Serum CCL-18 level is a risk factor for COPD exacerbations requiring hospitalization. *Int J Chron Obstruct Pulmon Dis*. 2017;12:199–208. doi:10.2147/copd.S118424
32. Enomoto T, Takeda Y, Shirai Y, et al. Serum C-C motif chemokine ligand 17 as a predictive biomarker for the progression of non-idiopathic pulmonary fibrosis interstitial lung disease. *Respir Res*. 2025;26(1):157. doi:10.1186/s12931-025-03237-2
33. Brandsma CA, Van den Berge M, Hackett TL, Brusselle G, Timens W. Recent advances in chronic obstructive pulmonary disease pathogenesis: from disease mechanisms to precision medicine. *J Pathol*. 2020;250(5):624–635. doi:10.1002/path.5364
34. Mei Q, Liu Z, Zuo H, Yang Z, Qu J. Idiopathic pulmonary fibrosis: an update on pathogenesis. *Front Pharmacol*. 2021;12:797292. doi:10.3389/fphar.2021.797292
35. Kai Y, Yoneyama H, Yoshikawa M, Kimura H, Muro S. Chondroitin sulfate in tissue remodeling: therapeutic implications for pulmonary fibrosis. *Respir Investig*. 2021;59(5):576–588. doi:10.1016/j.resinv.2021.05.012
36. da Costa FH B, Lewis MS, Truong A, Carson DD, Farach-Carson MC. SULF1 suppresses Wnt3A-driven growth of bone metastatic prostate cancer in perlecan-modified 3D cancer-stroma-macrophage triculture models. *PLoS One*. 2020;15(5):e0230354. doi:10.1371/journal.pone.0230354
37. Wade A, Engler JR, Tran VM, Phillips JJ. Measuring sulfatase expression and invasion in glioblastoma. *Methods Mole Biolog*. 2022;2303:415–425. doi:10.1007/978-1-0716-1398-6_33
38. Chen L, Zhu D, Huang J, Zhang H, Zhou G, Zhong X. Identification of hub genes associated with COPD through integrated bioinformatics analysis. *Int J Chronic Obstr*. 2022;17:439–456. doi:10.2147/copd.S353765
39. Tu M, Lu C, Jia H, et al. SULF1 expression is increased and promotes fibrosis through the TGF-β1/SMAD pathway in idiopathic pulmonary fibrosis. *J Transl Med*. 2024;22(1):885. doi:10.1186/s12967-024-05698-3

40. Yu S, Wang S, Sun X, et al. ST8SIA1 inhibits the proliferation, migration and invasion of bladder cancer cells by blocking the JAK/STAT signaling pathway. *Oncol Lett*. 2021;22(4):736. doi:10.3892/ol.2021.12997
41. Ramos RI, Bustos MA, Wu J, et al. Upregulation of cell surface GD3 ganglioside phenotype is associated with human melanoma brain metastasis. *Mol oncol*. 2020;14(8):1760–1778. doi:10.1002/1878-0261.12702
42. Babaha F, Abolhassani H, Hamidi Esfahani Z, Yazdani R, Aghamohammadi A. A new case of congenital ficolin-3 deficiency with primary immunodeficiency. *Expert Rev Clin Immunol*. 2020;16(7):733–738. doi:10.1080/1744666x.2020.1792779
43. Li W, Zu L, Xu S. FCN3 can serve as a potential biomarker for prognosis and immunotherapy of lung squamous cell carcinoma. *Zhongguo Fei Ai Za Zhi*. 2025;28(2):114–130. doi:10.3779/j.issn.1009-3419.2025.105.01
44. Jiang Y, Zhang Y, Zhao C. Integrated gene expression profiling analysis reveals SERPINA3, FCN3, FREM1, MNS1 as candidate biomarkers in heart failure and their correlation with immune infitration. *J Thorac Dis*. 2022;14(4):1106–1119. doi:10.21037/jtd-22-22
45. Wang J, Xie S, Cheng Y, Li X, Chen J, Zhu M. Identification of potential biomarkers of inflammation-related genes for ischemic cardiomyopathy. *Front Cardiovasc Med*. 2022;9:972274. doi:10.3389/fcvm.2022.972274
46. Ye C, Zhu J, Wang J, et al. Single-cell and spatial transcriptomics reveal the fibrosis-related immune landscape of biliary atresia. *Clin Translational Med*. 2022;12(11):e1070. doi:10.1002/ctm2.1070
47. Nah EH, Cho S, Kim S, Kim HS, Cho HI. Diagnostic performance of Mac-2 binding protein glycosylation isomer (M2BPGi) in screening liver fibrosis in health checkups. *J Clin Lab Analysis*. 2020;34(8):e23316. doi:10.1002/jcla.23316
48. Ohkawa Y, Harada Y, Taniguchi N. Keratan sulfate-based glycomimetics using Langerin as a target for COPD: lessons from studies on Fut8 and core fucose. *Biochem Soc TranS*. 2021;49(1):441–453. doi:10.1042/bst20200780
49. Merl-Pham J, Basak T, Knüppel L, et al. Quantitative proteomic profiling of extracellular matrix and site-specific collagen post-translational modifications in an in vitro model of lung fibrosis. *Matrix Biology Plus Feb*. 2019;1:100005. doi:10.1016/j.mbplus.2019.04.002