Identification of hub genes and key pathways in the emphysema phenotype of COPD

Qiunan Zuo[1,2,*], Youyu Wang[3,*], Deqing Yang[1,*], Shujin Guo[2], Xiaohui Li[2], Jiajia Dong[1], Chun Wan[1], Yongchun Shen[1], Fuqiang Wen[1]

[1]Department of Respiratory and Critical Care Medicine, West China Hospital of Sichuan University and Division of Pulmonary Diseases, State Key Laboratory of Biotherapy of China, Chengdu 610041, China
[2]Respiratory Ward, Department of Geriatrics, Sichuan Provincial People's Hospital, University of Electronic Science and Technology of China, Chengdu 611731, China
[3]Department of Thoracic Surgery, Sichuan Provincial People's Hospital, University of Electronic Science and Technology of China, Chengdu 611731, China
*Co-first authors

Correspondence to: Yongchun Shen, Fuqiang Wen; email: shenyongchun@wchscu.cn, wenfuqiang@scu.edu.cn
Keywords: chronic obstructive pulmonary disease, emphysema, weighted gene co-expression network analysis
Received: September 21, 2020   Accepted: December 10, 2020   Published: February 1, 2021

ABSTRACT

Chronic obstructive pulmonary disease (COPD) is a heterogeneous condition associated with high morbidity and mortality. This study aimed to use weighted gene co-expression network analysis (WGCNA) to explore the molecular pathogenesis of the emphysema phenotype of COPD. After obtaining lung mRNA expression profiles from ten patients with the emphysema phenotype of COPD and eight controls, emphysema-associated gene modules were identified with WGCNA. Among 13 distinct modules, the green-yellow and brown modules showed the strongest correlations with emphysema severity and lung function and were thus selected as hub modules. On gene ontology analysis, these two modules were mainly enriched in immune response, B cell receptor (BCR) signaling pathway, extracellular matrix (ECM) organization, and collagen fibril organization. Pathway analysis primarily showed enrichment in BCR signaling pathways, ECM receptor interaction, and NF-κB and TGF-β signaling pathways for the two hub modules. Several genes, including FCRLA, MS4A1, CD19, FKBP10, C1S and HTRA1, among others, were identified as hub genes. Our results shed light on the potential genetic mechanisms underlying the pathogenesis of the emphysema phenotype of COPD. However, further research will be needed to confirm the involvement of the identified genes and to determine their therapeutic relevance.

INTRODUCTION

Chronic obstructive pulmonary disease (COPD) is a chronic airway inflammatory disorder characterized by high morbidity and mortality, which represents a severe public health problem worldwide [1]. COPD is a heterogeneous disease and is thus increasingly categorized into different phenotypes based on clinical features, frequency of exacerbations, treatment responses, and clinical prognosis. Chronic bronchiolitis and emphysema represent the most notable phenotypes of COPD [2]. Patients with the chronic bronchiolitis phenotype present chronic airway inflammation, airway goblet cell hyperplasia, and airway mucus hyper-secretion. By contrast, emphysema, another common phenotype of COPD, is characterized by elastolytic destruction of the alveolar wall and no evidence of obvious fibrosis, has drawn in recent years increased attention in both clinical and basic research [3–5].

Patients with phenotypes of emphysema show also some distinct clinical features and treatment responses. The former includes worse pulmonary function, lower body mass index (BMI), and more severe of dyspnea than bronchiolitis patients [6]. A study analyzing 342 hospitalized patients with COPD for the first time of acute exacerbation showed that those with emphysema presented a more severe condition and a worse clinical situation than bronchiolitis patients [7]. As a result of these different mechanisms and clinical characteristics, the pharmacotherapy for each COPD phenotype is different. COPD patients with bronchiolitis phenotype are generally administered inhaled corticosteroids (ICS), while results from a prospective study revealed that COPD patients with emphysema phenotype show less improvement in lung function and dyspnea on ICS and long-acting beta-agonist combination intervention [8]. The pathogenesis of the emphysema phenotype of COPD is highly complex, and significant gaps remain about the molecular mechanisms responsible for the different sub-phenotypes. There is thus a clear need to unveil the specific disease mechanisms in order to advance new treatments for these patients.

Transcriptomic studies based on pathological specimens may help find novel insights into the molecular mechanisms of human disease and identify new treatment targets to develop patient-centered treatments. Previous work has profiled gene expression in patients with COPD, but simple analyses of differential expressed genes are not able to obtain full heterogeneity of pathophysiology for COPD, leaving the molecular mechanisms underlying emphysema unclear [9]. Weighted gene co-expression network analysis (WGCNA) is used to identify potential correlations among genes across sequencing or microarray specimens, allowing detection of clusters (modules) of highly correlated genes [9–11]. WGCNA defines gene clusters and reveal relationships among modules and between modules and external clinical traits by using an intramodular hub gene or the module eigengene [9–11]. In this way, WGCNA may be useful to understand molecular mechanisms of disease and to identify reliable biomarkers for more effective diagnosis, prognosis, and treatment. Using new RNA-Seq data, the current study aimed to address the paucity of WGCNA studies related to the characterization of key genes and signaling pathways involved in the pathogenesis of the emphysema phenotype of COPD.

RESULTS

Demographic and clinical characteristics of study patients

The current study enrolled 18 patients, with 15 males and 3 females. The mean ages of all patients were 60.6

years. The severity of emphysema was assessed using the Goddard scoring method, based on the HRCT features of each patient (Supplementary File 1). The demographic and clinical data of patients and controls was summarized in Table 1.

Coexpression networks and modules construction

The top 50% most varying genes (n = 9,550) were used for WGCNA network construction. Thirteen types of clinical data, including disease status (COPD vs. control), sex, age, history of smoking, white blood cell count, neutrophil count, smoking (packages per years), FEV_1/FVC, FEV_1/Pre, maximum mid-expiratory flow (MMEF), CAT score, and Goddard score were evaluated. Genes were divided into modules through hierarchical average linkage clustering if they exhibited similar patterns of expression (Figure 1). To assure relatively balanced mean connectivity and scale independence, the network topology was constructed with various soft threshold powers in order and a soft threshold power of 9 was finally chosen to analyze network topology. Using a dynamic tree-cutting algorithm and 0.25 as the merging threshold function, 13 modules were identified in the dataset (Figure 2). A topological overlap matrix (TOM) plot of the gene network was constructed (Figure 3).

Relationships between clinical traits and modules

Next, the correlations between the 13 module eigengenes (MEs) and traits of interest were evaluated (Figure 4). The brown module was negatively associated with FEV_1/FVC (r=-0.74, p=5e-04), FEV_1/Pre (r=-0.52, p=0.03), positively associated with Goddard score (r=0.63, p=0.005), CAT score (r=0.57, p=0.01), and COPD status (r=0.46, p=0.05). The green-yellow module was negatively associated with FEV_1/FVC (r=-0.70, p=0.001), FEV_1/Pre (r=-0.59, p=0.01), positively associated with Goddard score (r=0.79, p=9e-05), CAT score (r=0.70, p=0.001), and COPD status (r=0.56, p=0.02). We focused on the relationship between emphysema severity (Goddard score), severity of lung function impairment, CAT score traits, and MEs, which led us to select the brown and green-yellow modules for next analysis.

Functional annotation of genes in highly correlated modules

There were 47 genes in green-yellow module and 1,402 genes in brown module, respectively. The main biological processes of genes in green-yellow module contained were immune response, adaptive immune response, B cell receptor (BCR) signaling pathway, and primary immunodeficiency; Kyoto Encyclopedia of

Table 1. Demographic and clinical information of the study subjects.

Characteristic	COPD (n=10)	Non-COPD (n=8)	P
Age (years)	62.4±5.06	58.3±7.9	0.41
Sex (F/M)	1/9	2/6	0.55
Smoking history (packages per year)	41.75±38.33	14.4±12.08	0.19
White blood cell count (/10^9)	6.55±1.95	4.95±0.88	0.041
Neutrophil count (/10^9)	4.34±1.91	2.91±0.96	0.076
CAT score	16.7±1.25	9.87±9.35	0.20
FEV1/FVC (%)	52.86±21.28	80.7±3.80	<0.001
FEV1/Pre (%)	57.04±25.34	107.7±17.07	0.002
MMEF (L)	0.70±0.43	2.98±1.10	<0.001
Goddard score	11.1±5.92	0	<0.001

CAT, Chronic obstructive pulmonary disease assessment test; FEV1, Forced expiratory volume in the first second; FVC, Forced vital capacity; MMEF, Maximum mid-expiratory flow. Values are n or mean ± SD, unless otherwise noted.

Genes and Genomes (KEGG) analysis found that the main enriched signaling pathways were BCR signaling pathway, cytokine-cytokine receptor interaction, and NF-κB activity. The genes in the brown module were mainly enriched in the biological processes of extracellular matrix (ECM) organization, ECM disassembly, collagen fibril organization, signaling pathways of the protein digestion and absorption, complement and coagulation cascades, focal adhesion, and TGF-β signaling pathways. The Gene Ontology
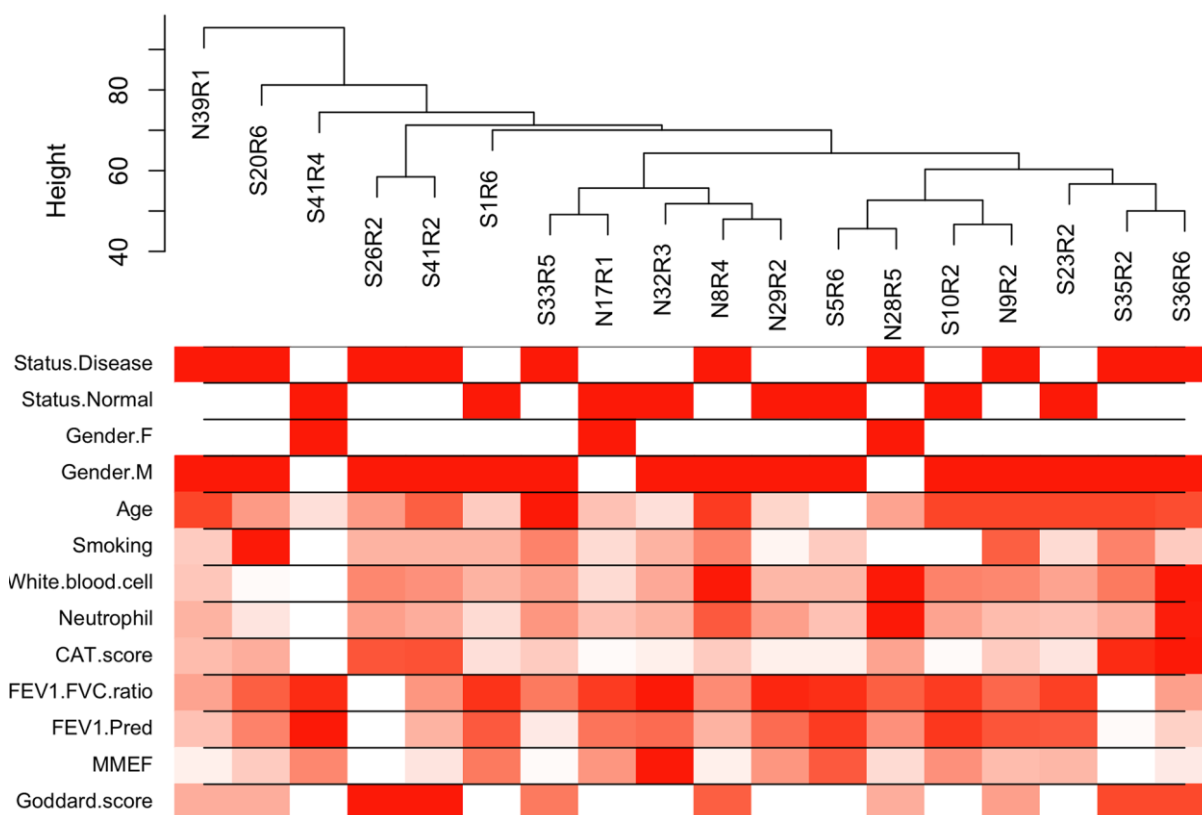

Figure 1. Cluster tree of clinical samples. The leaves of the tree correspond to the samples. Color bands represent the numeric values of the physiological traits.

(GO) terms for biological processes, cellular component, molecular function, as well as the KEGG signaling pathways enriched in genes in the two modules are shown in Figures 5, 6.

Identification and visualization of intramodular hub genes

The Cytoscape was used to visualize the top 40 gene connections and to find the hub genes in each module. The top ten hub genes in the green-yellow module were FCRLA, MS4A1, CD19, FCRL5, CD79A, CD79B, PNOC, BLK, POU2AF1, and VPREB3 (Figure 7A). Genes of FKBP10, C1S, HTRA1, ADAMTS2, EMILIN1, LOXL1, COL6A1, PCOLCE, COL5A1, and COL1A2 were the top ten hub genes in the brown module (Figure 7B). In turn, differential expression

analysis indicated that the expression of CD79A, VPREB3, LOXL1, BLK, FKBP10, FCRL5, COL1A2, HTRA1, and CD19 was significantly increased in patients with COPD, relative to controls (all P<0.05, Supplementary Figure 1).

DISCUSSION

In this study, we used WGCNA to analyze the lung transcriptome of ten patients with emphysema phenotype of COPD and eight controls without COPD. We identified a series of significant pathways, including BCR signaling pathway, ECM-related pathways, NF-κB and TGF-β signaling pathways, and several hub genes (i.e. CD19, BLK, MS4A1, POU2AF1, and COL6A1) that may be involved in the emphysema phenotype of COPD.
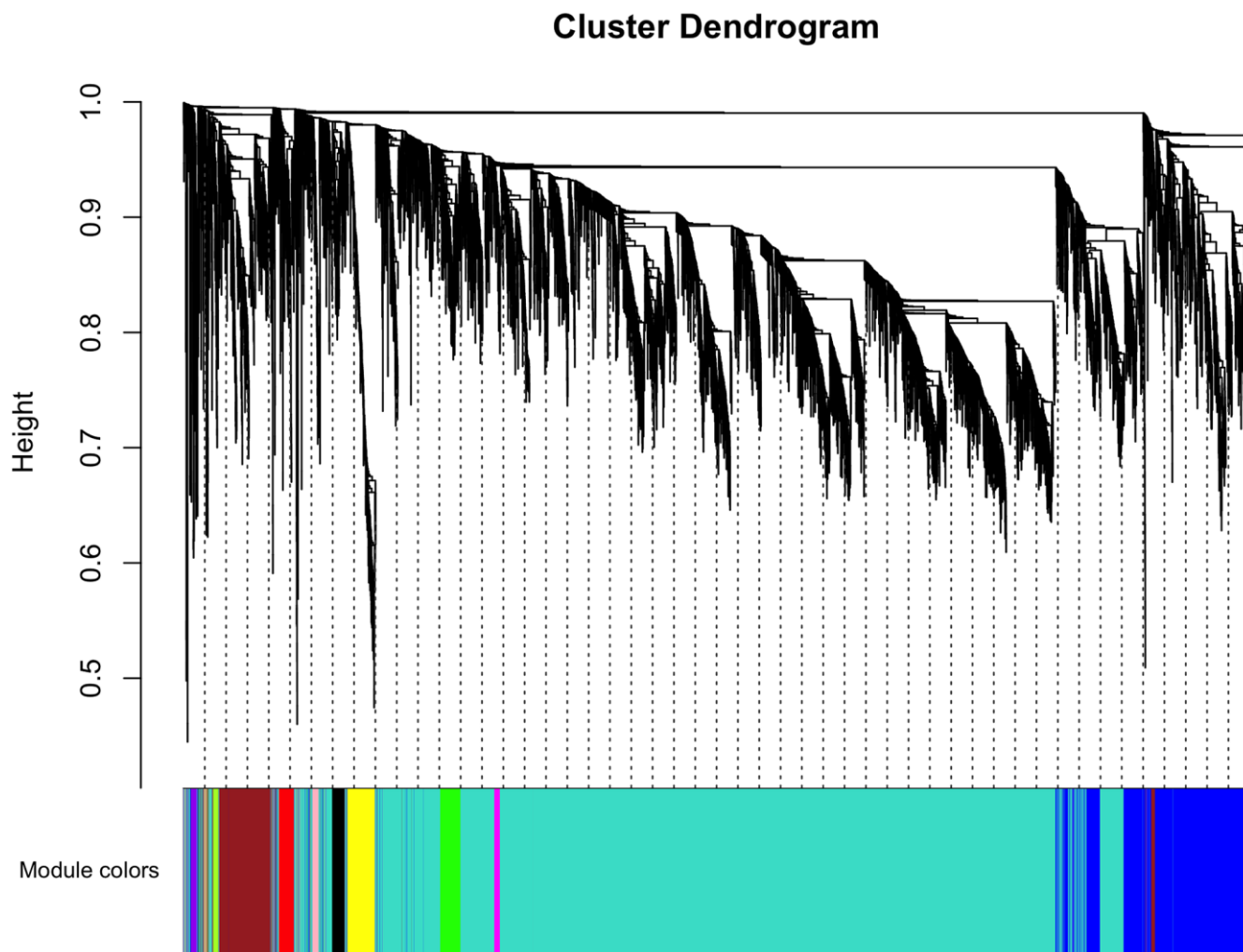
Cluster Dendrogram

Figure 2. Cluster dendrogram of gene coexpression and functional modules. Heatmap plot of the topological overlap matrix (TOM) supplemented by hierarchical clustering dendrograms and module colors. A total of 13 distinct co-expression modules were identified containing tan to turquoise genes. Another 30 uncorrelated genes were assigned to a grey module that was not included in subsequent analyses.

In our previous WGCNA study using GSE69818 dataset, IFT88, Bik, MMP10, and CCDC103 were identified as hub genes and were consistently expressed in patients with emphysema, the emphysema related genes were mainly enriched for apoptotic mitochondrial changes, proteolysis, and functions of cilium assembly and movement [12]. However, there was insufficient patient data to comprehensively assess correlations
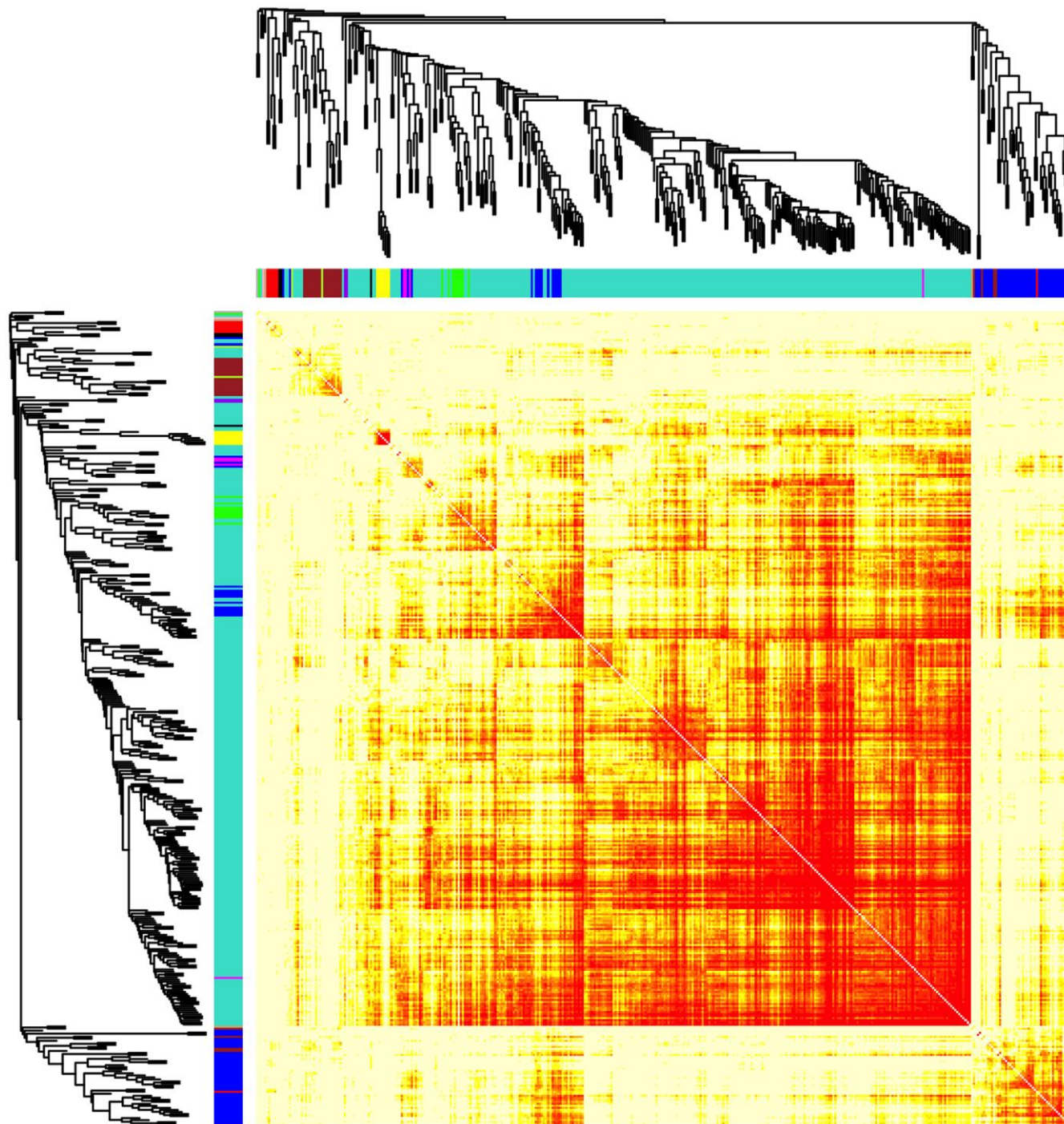

Figure 3. Heatmap of the gene network. TOM plot for all the genes in the analysis. Light colors represent low overlap and progressively darker (red) colors represent higher overlap. Darker blocks along the diagonal correspond to modules. The gene dendrogram and module assignment are also shown along the left and top sides.

between WGCNA gene modules and clinical characteristics in that study [12]. By contrast, in the present study we were able to correlate lung transcriptome data, obtained with the latest sequencing platform, with emphysema severity (Goddard score), severity of lung function (FEV1/FVC, FEV1/Pre), CAT score, and other clinical traits. This comprehensive analysis allowed us to select the most correlated gene modules to obtain new insights on the molecular mechanisms potentially affecting the emphysema phenotype of COPD. Attesting to the validity of our research, the current findings overlap partially with those of Faner et al. [13], who also identified MS4A1, POU2AF1, FCRL5, COL1A2 as potential hub genes of emphysema and indicated also the potential involvement of BCR signaling pathway and ECM organization in its development.

Mounting evidence revealed a prominent role of BCR signaling in the pathogenesis of emphysema and COPD.

In blood and lung tissue of patients with COPD, increased B cell products (autoantibodies) and B cell numbers were observed, and the B cell-rich lymphoid follicles has been demonstrated in COPD patients with more severe stages [14–16]. In a previous study, a prominent B cell molecular signature occurs preferentially in emphysema; upregulated genes are enriched in ontologies related to B-cell homing and activation, while the immune coexpression network shows a central core of B cell-related genes, which is absent in bronchiolitis cases [13]. The current study found several B cell–related genes (BLK, CD19, MS4A1, CD79B, CD79A, POU2AF1) that were enriched in the green-yellow module. The BCR signaling pathway was also enriched on KEGG analysis, suggesting that BCR is an attractive research target for the emphysema phenotype of COPD.

CD19 is a characteristic marker of B cells. Cigarette smoking, one most important risk factor for COPD, is
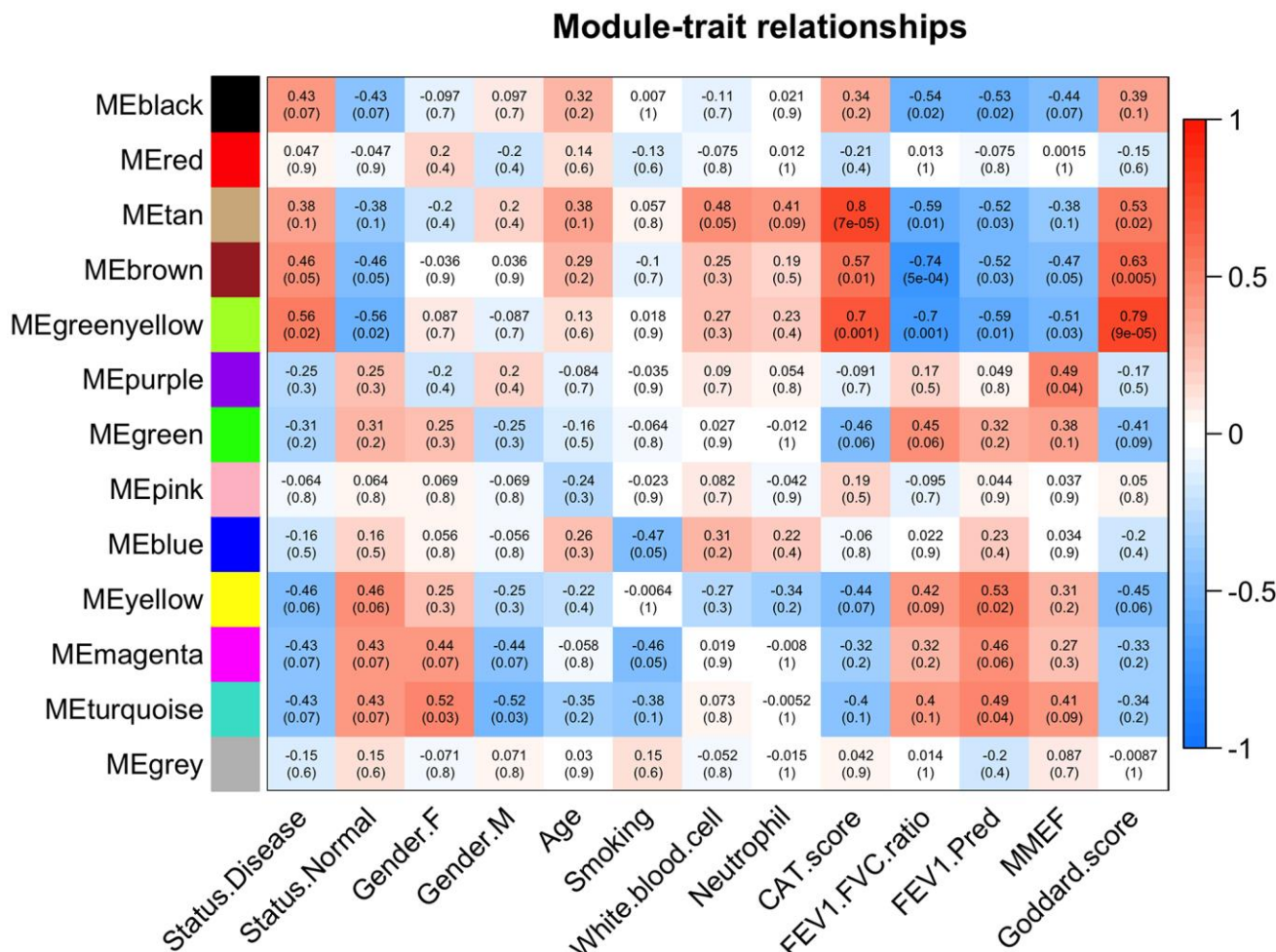

Figure 4. Module-trait associations. Each row corresponds to a module eigengene and each column to a trait. Each cell contains the corresponding correlation and p values. The table is color-coded based on correlation values.
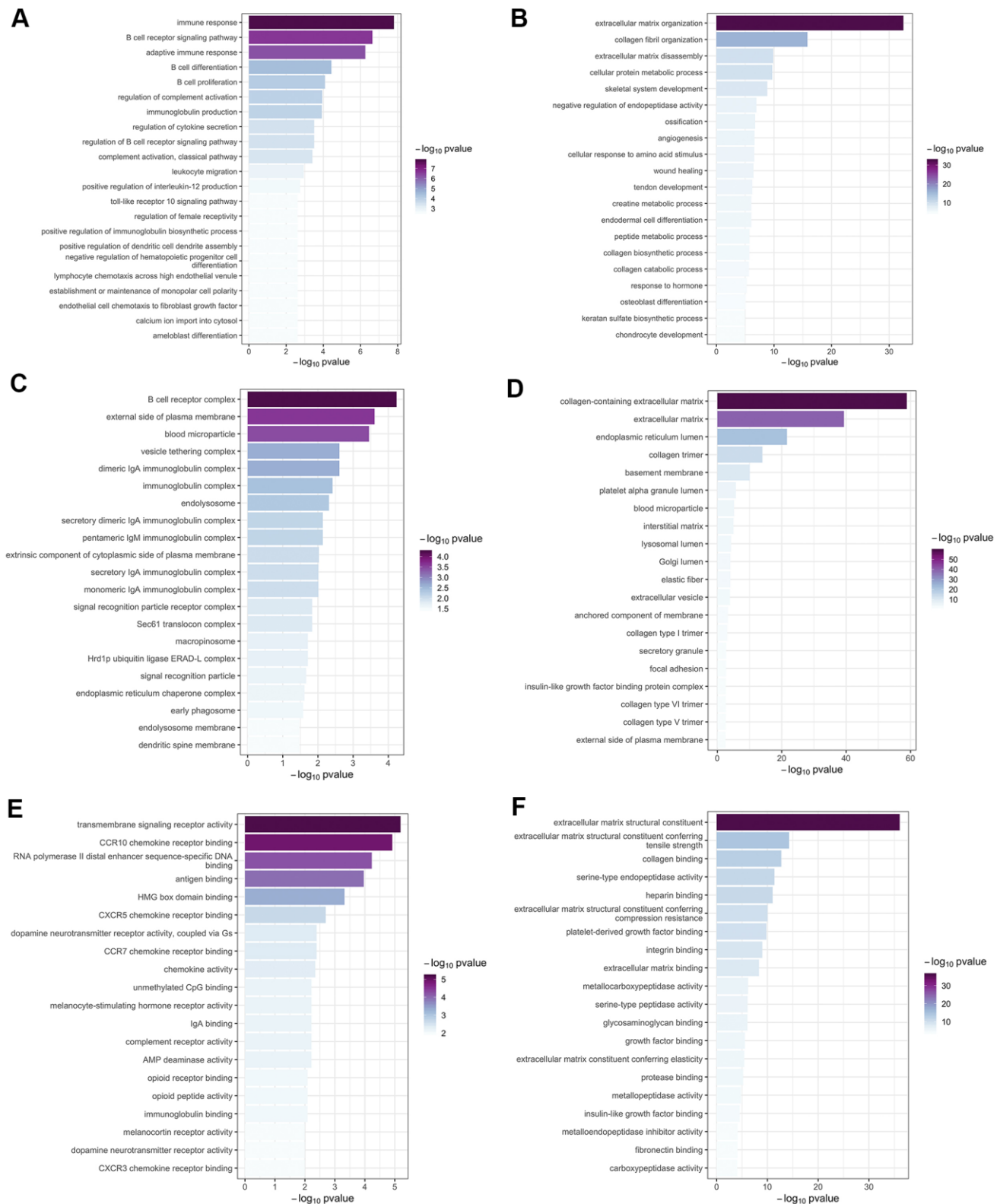
Figure 5. Gene ontology enrichment analysis. The top 20 significant (p < 0.05) GO terms were screened in the green-yellow and brown modules. Column color was used to map the p value of specific functional items: darker colors indicate smaller p values (greater significance) for the corresponding enrichment. Results are shown for (**A**) biological process in the yellow-green module; (**B**) biological process in the brown module; (**C**) cellular component in the yellow-green module; (**D**) cellular component in the brown module; (**E**) molecular function in the yellow-green module; and (**F**) molecular function in the brown module.

associated with percentage increases in CD19+ B cells, as well as activated B cells, among peripheral blood mononuclear cells [17]. MS4A1, also named as CD20, is an activated-glycosylated phosphoprotein that interacts physically with BCR and other membrane proteins and regulates the development and differentiation of B cells into plasma cells [18]. LncRNA COPDA1 can increase the expression of MS4A1, then to increase store-operated calcium entry in human bronchial smooth muscle cells, promoting their proliferation and airway remodeling [19].

POU2AF1 expression correlates with emphysema severity, suggesting a potential contributing role [16]. Zhou et al reported that enhancing expression of POU2AF1 alleviated the suppression of host defense genes by smoking in human airway epithelial cells [20]. However, the mechanism of POU2AF1 in emphysema needs further research. Although functional studies are needed to validate the present findings, our study provides further evidence that B cell-associated genes and pathways are involved in the emphysema phenotype of COPD.
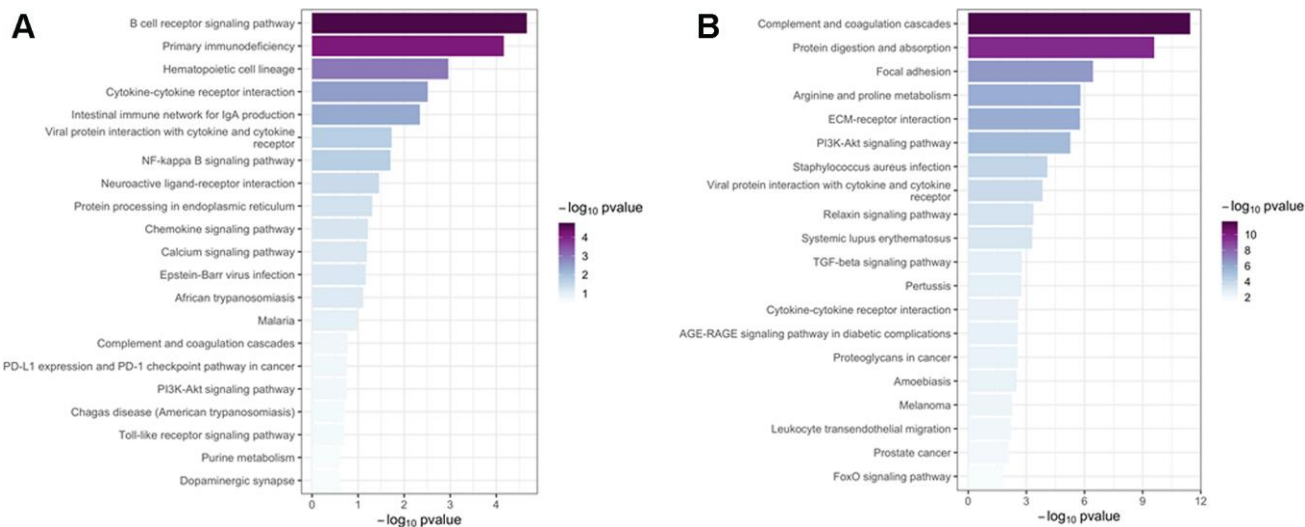

Figure 6. Kyoto Encyclopedia of Genes and Genomes pathway analysis. The top 20 significant (p < 0.05) KEGG signaling pathways were screened in the green-yellow (**A**) and brown (**B**) modules. Column color was used to map the p value of specific functional items: darker columns indicate smaller p values (greater significance) for the corresponding enrichment.
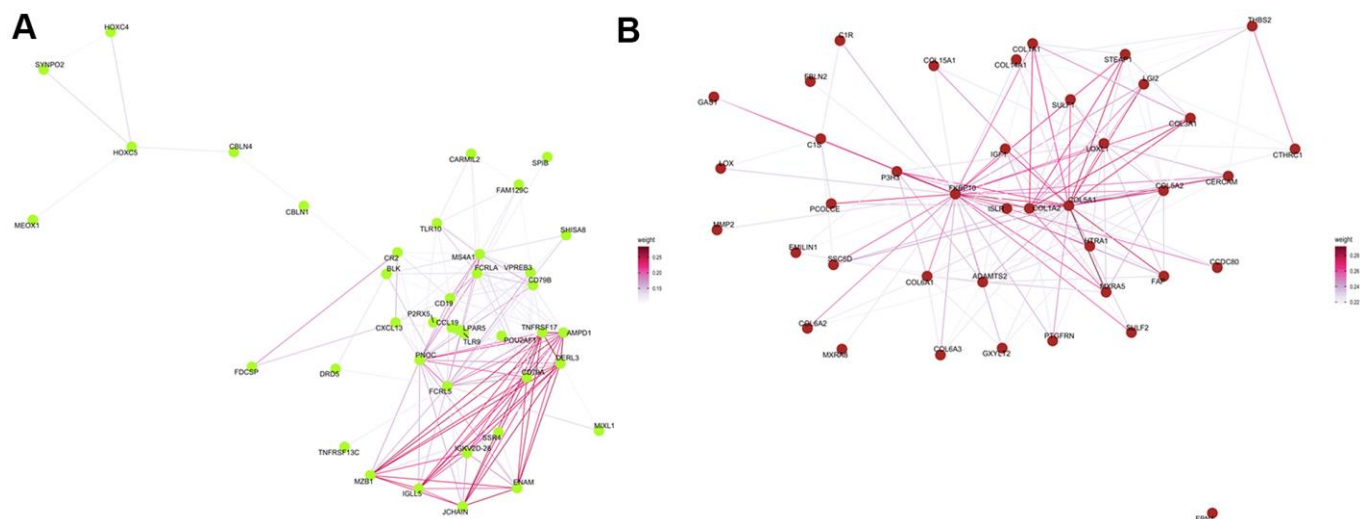

Figure 7. Protein–protein interaction network analysis. Visualization of the network connections among the most highly connected genes within the green-yellow (**A**) and brown (**B**) modules. Edge weights represent similarity between nodes.

Disturbance of the ECM can lead to lung tissue remodeling (including airway wall fibrosis and emphysema) that affects all lung compartments in COPD. Accordingly, several studies have shown that ECM composition is altered in COPD and emphysema patients [21]. For instance, the disruption and remodeling of collagen and elastic fibers surrounding the alveolar airspaces represents a ubiquitous pathological finding in emphysematous human lung samples [22]. Matrix metalloproteinases (MMPs), including MMP 1, 2, 7, 9, and 12, may contribute to the process of emphysematous lung tissue destruction in COPD, which shifts the balance of ECM synthesis and degradation towards degradation [23]. This pheno-menon is consistent with the enrichment observed in current study for genes of the brown module in the ECM GO term. MMP-2 and MMP-9 have already been implicated in the pathogenesis of emphysema [24, 25], so they may be potential therapeutic targets in the emphysema phenotype of COPD.

A previous genome-wide association study identified 234 genes that may be associated with lung function in COPD, and several ECM-related processes, including ECM organization, extracellular structure organization, and elastic fiber formation were enriched from these genes [26]. In our study, COL6A1 and EMILIN1, two genes involved in ECM organization, were also enriched in emphysema, and therefore their function in emphysema and COPD merits further research. Increased deposition of collagen in the small airway walls and changes structures denoted mainly by disorganized collagen fibrils are also critical ECM changes in the development of COPD [27]. Changes in collagen and elastic fibers with loss of alveolar attachments result in loose airway walls that deficit alveolar support, which contribute to increased collapsibility of airway, airflow limitation, and alveolar enlargement. Our WGCNA results identified the hub genes in the brown module (i.e. ADAMTS2 and COL5A1) that were enriched in collagen fibril organization. These results support that altered ECM homeostasis is associated with the pathogenesis of COPD.

NF-κB is another potential therapeutic target for the emphysema phenotype of COPD. NF-κB promotes the production of inflammatory mediators, including tumor necrosis factor-α, interleukin-6, and interleukin-1β, growing evidence suggests that NF-κB plays a crucial role in airway inflammation and remodeling processes in COPD [28]. Overexpressed protein arginine methyltransferase-6 may suppress the ability of cigarette smoke extract to activate NF-κB and induce pro-inflammatory gene expression in a murine emphysema model [29]. NF-κB is also involved in the

protective process by which erythromycin attenuates MMP/anti-MMP imbalance in cigarette smoke-induced emphysema [30].

We detected enrichment of TGF-β signaling pathway genes in the brown module. TGF-β is necessary for lung organogenesis and homeostasis, and its dysregulation impacts many respiratory diseases, including COPD [31]. TGF-β1 overexpression has been observed in small airway epithelial cells among smokers and COPD patients, suggesting that activated TGF-β signaling is involved in the pathogenesis of COPD [32]. TGF-β overexpression is associated with early events during the emphysematous process [33], and TGF-β dysregulation due to acetylated Smad7 was shown to contribute to emphysema [34]. Future researches are required to determine the mechanism of TGF-β in the emphysema phenotype of COPD.

Despite impressive progress in COPD treatments, no 'one-size-fits-all' pharmacological treatment has been discovered. Instead, COPD treatment is relied on clinical symptoms and severity of disease, and focuses on each patient's individual pathology [35]. The discovery of key pathways involved in COPD, such as B cell-related inflammation, disturbed ECM changes and collagen fibers, updates our perspective on the pathological mechanisms and identifies potential therapeutic targets for COPD with the emphysema phenotype. Still, the lack of functional characterization of some COPD hub genes, like VPREB3, LOXL1, BLK, FKBP10, FCRL5, and HTRA1, suggest that extensive research needs to be undertaken to arrive at more precise and effective therapies.

The findings of the present study should be explained carefully due to potential limitations. First of all, the sample was relatively small and data on some clinical traits like history of exacerbation were not presented, so we were unable to carry out a more comprehensive WGCNA. Second, the Goddard score is a relatively crude measure of emphysema severity, which may bias our results to a certain extent. Third, since lung tissue specimens are hard to obtain, we included lung samples from only 18 patients. Such limited number of specimens and the heterogeneity among them may have a significant effect on the final results. Additionally, we included three cases of patients with comorbidities of lung cancer; since our gene data showed enrichment in the BCR, ECM, NF-κB, and TNF signaling pathways, which are all related to inflammation and tumor, we can't exclude the possibility of false positive results due to tumor progression [36–38]. Finally, our study did not include mechanistic studies to verify and deepen the insights obtained from our *in silico* analysis. Future work involving a larger patient sample and strict

inclusion criteria should build on our results to better understand and treat the emphysema phenotype of COPD.

In summary, we used WGCNA based on lung transcriptome data to build a coexpression network associated with the emphysema phenotype of COPD. We identified two modules and a series of hub genes in association with the emphysema phenotype, which supply new insights into the underlying molecular mechanisms. The findings of this study should be confirmed *in vitro* and in clinical studies.

MATERIALS AND METHODS

Lung tissue collection

Lung tissue specimens were prospectively collected from eighteen patients who required thoracic surgery due to a solitary pulmonary nodule (n=12) or lung volume reduction (n=6). The cases were defined by both high-resolution computed tomography (HRCT) identified emphysema and spirometer-detected COPD with the diagnostic criteria set by Global Initiative for Chronic Obstructive Lung Disease [1]. The study included six cases of lung volume reduction, three cases of pulmonary nodule due to lung cancer, and one case of inflammatory pseudotumor. The controls were patients with benign inflammatory nodules without COPD and signs of emphysema on HRCT. Normal lung tissue under microscopy examination was taken at least 5 cm from the tumor if the patient underwent resection surgery for lung tumor or nodule. The collected lung tissue was immediately stored at -80° C. All samples were histologically analyzed using standard hematoxylin and eosin staining [39, 40].

All participants provided written informed consent for their anonymized clinical data to be collected and published for scientific research purposes. This research was approved by the Ethics Committee of Sichuan Provincial People's Hospital (Number: 2020-94).

Clinical data collection and emphysema evaluation

Demographic and clinical data including age, sex, history of smoking, white blood cell count, neutrophil count, COPD Assessment Test (CAT) score, and spirometry examination were collected for all subjects on admission. Emphysema was detected by HRCT (Siemens, Germany), which is part of routine clinical management before surgery at our hospital. Technical parameters of HRCT were the following: 1-mm collimation, 1–2 mm slice thickness, 120–140 kV, 75–350 mA, and 0.75–1 s scanning time.

The Goddard score was used to analyze the severity of emphysema based on lung HRCT [41]. The interpretations of chest image were carried out at window levels of 2600-2900 Hounsfield units (HU) and at 800–1500 HU window widths, which are the best conditions for detecting emphysema. Six images in three levels of both lung slices (the aortic arch, carina, and 1-2 cm above the highest hemidiaphragm) were analyzed, and the images were scored as normal (0), 5% affected (0.5), 25% affected (1), 50% affected (2), 75% affected (3), or >75% affected (4). The total score ranged from 0 to 24, and the scores were determined independently by two pulmono-logists.

RNA sequencing

Total RNA in lung tissues was extracted using TRIzolH reagent (Invitrogen, Burlington, ON, Canada), following the manufacturer's protocols. A Qubit Fluorometer (Thermo Fisher Scientific, Waltham, MA, USA) was chosen to assess the yield and purity of the RNA. The integrity of RNA was examined in a 1% agarose gel with RNA 6000 Nano Assay Kit and an Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA). RNA from each specimens (5 μg) was sent to isolate poly(A) mRNA, then a non-directional Illumina RNA-Seq library was created with an mRNA-Seq Sample Prep Kit (Illumina, San Diego, CA, USA). A Bioanalyzer Chip DNA 1000 series II (Agilent Technologies) was used for Library quality control and quantification. Every RNA-Seq library had an insert size of 2 base pairs (bp), and 150-bp sequences were created using an Illumina HiSeq 4000 system.

WGCNA

First, a weighted correlation network was generated by constructing a matrix of pairwise correlations between all pairs of genes chosen according to variance. Comparability was assessed by correlation analysis of connectivity and average gene expression. Connectivity was assessed with the function of soft connectivity in the software R with WGCNA package [42], which was also used to construct the WGCNA in this study. Cluster analysis using the Flash Cluster package in R was carried out. A module was defined as group of genes that show similar patterns of connection strengths with all the other genes in the network, and these groups of gens often have similar functions. Modules were determined using the cut-tree Hybrid function in the software R with Dynamic Tree Cut package.

To identify MEs, the module eigengenes function in the WGCNA package was chosen for calculation. The

signed module membership was used to evaluate the correlations between individual genes and MEs. The correlation was used to assess whether a given gene belonged to a specific module. Correlations between MEs and interested clinical traits were calculated to find gene modules whose expression patterns were associated with specific clinical traits. Gene significance (GS) for individual gene was set as the correlation between its expression level and the presence of a specific clinical trait, using the plot module significance in the WGCNA package. The clinical traits evaluated in our study included: disease status (COPD vs. control), age, sex, history of smoking, white blood cell count, neutrophil count, CAT score, and Goddard score. Functional annotation of all genes in each module was performed with the Cluster Profiler based on GO and KEGG databases. The top 20 most enriched GO and KEGG function terms (ordered by p value) were visualized graphically. Intramodular hub genes were identified as genes with the maximum module membership values in protein-protein interaction (PPI) network analysis. The top 40 gene connections based on topological overlap was visualized using Cytoscape software [43]. The expression levels of top 10 intramodular hub genes were compared between COPD and control using Wilcoxon Rank Sum Test, and a P value < 0.05 was set as significant.

Statistical analysis

Clinical data of COPD patients and controls are presented as means±standard deviation. Comparisons between COPD and control groups were carried out using nonparametric statistics and chi-squared test. Differences associated with $p<0.05$ were considered significant.

AUTHOR CONTRIBUTIONS

Qiunan Zuo: Conceptualization, research, data curation and analysis, and writing (original draft). Youyu Wang and Deqing Yang: Conceptualization, research, data curation and analysis, and writing (review and editing). Shujing Guo and Xiaohui Li: Methodology, research, and data curation. Jiajia Dong and Chun Wan: Methodology, research, data curation, and funding acquisition. Yongchun Shen and Fuqiang Wen: Conceptualization, project administration and supervision, funding acquisition, and writing (review and editing).

CONFLICTS OF INTEREST

The authors declare that they have no conflicts of interest.

REFERENCES

1. Singh D, Agusti A, Anzueto A, Barnes PJ, Bourbeau J, Celli BR, Criner GJ, Frith P, Halpin DM, Han M, López Varela MV, Martinez F, Montes de Oca M, et al. Global strategy for the diagnosis, management, and prevention of chronic obstructive lung disease: the GOLD science committee report 2019. Eur Respir J. 2019; 53:1900164.
https://doi.org/10.1183/13993003.00164-2019
PMID:30846476

2. Siafakas N, Corlateanu A, Fouka E. Phenotyping before starting treatment in COPD? COPD. 2017; 14:367–74.
https://doi.org/10.1080/15412555.2017.1303041
PMID:28388265

3. Lopez-Campos JL, Centanni S. Current approaches for phenotyping as a target for precision medicine in COPD management. COPD. 2018; 15:108–17.
https://doi.org/10.1080/15412555.2018.1443064
PMID:29558165

4. 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; 250:624–35.
https://doi.org/10.1002/path.5364 PMID:31691283

5. Han MK, Tayob N, Murray S, Woodruff PG, Curtis JL, Kim V, Criner G, Galban CJ, Ross BD, Hoffman EA, Lynch DA, Kazerooni E, Martinez FJ, and COPDGene and SPIROMICS Investigators. Association between emphysema and chronic obstructive pulmonary disease outcomes in the COPDGene and SPIROMICS cohorts: a post hoc analysis of two clinical trials. Am J Respir Crit Care Med. 2018; 198:265–67.
https://doi.org/10.1164/rccm.201801-0051LE
PMID:29485901

6. Izquierdo-Alonso JL, Rodriguez-Gonzálezmoro JM, de Lucas-Ramos P, Unzueta I, Ribera X, Antón E, Martín A. Prevalence and characteristics of three clinical phenotypes of chronic obstructive pulmonary disease (COPD). Respir Med. 2013; 107:724–31.
https://doi.org/10.1016/j.rmed.2013.01.001
PMID:23419828

7. Garcia-Aymerich J, Gómez FP, Benet M, Farrero E, Basagaña X, Gayete À, Paré C, Freixa X, Ferrer J, Ferrer A, Roca J, Gáldiz JB, Sauleda J, et al, and PAC-COPD Study Group. Identification and prospective validation of clinically relevant chronic obstructive pulmonary disease (COPD) subtypes. Thorax. 2011; 66:430–37. https://doi.org/10.1136/thx.2010.154484 PMID:21177668

8. Lee JH, Lee YK, Kim EK, Kim TH, Huh JW, Kim WJ, Lee JH, Lee SM, Lee S, Lim SY, Shin TR, Yoon HI, Sheen SS, et al. Responses to inhaled long-acting beta-agonist and corticosteroid according to COPD subtype. Respir Med. 2010; 104:542–49. https://doi.org/10.1016/j.rmed.2009.10.024 PMID:19926461

9. Liu W, Li L, Ye H, Tu W. [Weighted gene co-expression network analysis in biomedicine research]. Sheng Wu Gong Cheng Xue Bao. 2017; 33:1791–801. https://doi.org/10.13345/j.cjb.170006 PMID:29202516

10. Xue Z, Huang K, Cai C, Cai L, Jiang CY, Feng Y, Liu Z, Zeng Q, Cheng L, Sun YE, Liu JY, Horvath S, Fan G. Genetic programs in human and mouse early embryos revealed by single-cell RNA sequencing. Nature. 2013; 500:593–97. https://doi.org/10.1038/nature12364 PMID:23892778

11. Duan H, Ge W, Zhang A, Xi Y, Chen Z, Luo D, Cheng Y, Fan KS, Horvath S, Sofroniew MV, Cheng L, Yang Z, Sun YE, Li X. Transcriptome analyses reveal molecular mechanisms underlying functional recovery after spinal cord injury. Proc Natl Acad Sci U S A. 2015; 112:13360–65. https://doi.org/10.1073/pnas.1510176112 PMID:26460053

12. Qin J, Yang T, Zeng N, Wan C, Gao L, Li X, Chen L, Shen Y, Wen F. Differential coexpression networks in bronchiolitis and emphysema phenotypes reveal heterogeneous mechanisms of chronic obstructive pulmonary disease. J Cell Mol Med. 2019; 23:6989–99. https://doi.org/10.1111/jcmm.14585 PMID:31419013

13. Faner R, Cruz T, Casserras T, López-Giraldo A, Noell G, Coca I, Tal-Singer R, Miller B, Rodriguez-Roisin R, Spira A, Kalko SG, Agustí A. Network analysis of lung transcriptomics reveals a distinct B-cell signature in emphysema. Am J Respir Crit Care Med. 2016; 193:1242–53. https://doi.org/10.1164/rccm.201507-1311OC PMID:26735770

14. Gosman MM, Willemse BW, Jansen DF, Lapperre TS, van Schadewijk A, Hiemstra PS, Postma DS, Timens W, Kerstjens HA, and Groningen and Leiden Universities Corticosteroids in Obstructive Lung Disease Study Group. Increased number of B-cells in bronchial biopsies in COPD. Eur Respir J. 2006; 27:60–64.

https://doi.org/10.1183/09031936.06.00007005 PMID:16387936

15. Núñez B, Sauleda J, Antó JM, Julià MR, Orozco M, Monsó E, Noguera A, Gómez FP, Garcia-Aymerich J, Agustí A, and PAC-COPD Investigators. Anti-tissue antibodies are related to lung function in chronic obstructive pulmonary disease. Am J Respir Crit Care Med. 2011; 183:1025–31. https://doi.org/10.1164/rccm.201001-0029OC PMID:21097696

16. Lee SH, Goswami S, Grudo A, Song LZ, Bandi V, Goodnight-White S, Green L, Hacken-Bitar J, Huh J, Bakaeen F, Coxson HO, Cogswell S, Storness-Bliss C, et al. Antielastin autoimmunity in tobacco smoking-induced emphysema. Nat Med. 2007; 13:567–69. https://doi.org/10.1038/nm1583 PMID:17450149

17. Burchiel SW, Lauer FT, Factor-Litvak P, Liu X, Santella RM, Islam T, Eunus M, Alam N, Islam T, Rahman M, Ahmed A, Ahsan H, Graziano J, Parvez F. An increase in circulating B cells and B cell activation markers in peripheral blood is associated with cigarette smoking in a male cohort in Bangladesh. Toxicol Appl Pharmacol. 2019; 384:114783. https://doi.org/10.1016/j.taap.2019.114783 PMID:31669812

18. Eon Kuek L, Leffler M, Mackay GA, Hulett MD. The MS4A family: counting past 1, 2 and 3. Immunol Cell Biol. 2016; 94:11–23. https://doi.org/10.1038/icb.2015.48 PMID:25835430

19. Zheng M, Hong W, Gao M, Yi E, Zhang J, Hao B, Liang C, Li X, Li C, Ye X, Liao B, He F, Zhou Y, et al. Long noncoding RNA COPDA1 promotes airway smooth muscle cell proliferation in chronic obstructive pulmonary disease. Am J Respir Cell Mol Biol. 2019; 61:584–96. https://doi.org/10.1165/rcmb.2018-0269OC PMID:31050548

20. Zhou H, Brekman A, Zuo WL, Ou X, Shaykhiev R, Agosto-Perez FJ, Wang R, Walters MS, Salit J, Strulovici-Barel Y, Staudt MR, Kaner RJ, Mezey JG, et al. POU2AF1 functions in the human airway epithelium to regulate expression of host defense genes. J Immunol. 2016; 196:3159–67. https://doi.org/10.4049/jimmunol.1502400 PMID:26927796

21. Bihlet AR, Karsdal MA, Sand JM, Leeming DJ, Roberts M, White W, Bowler R. Biomarkers of extracellular matrix turnover are associated with emphysema and eosinophilic-bronchitis in COPD. Respir Res. 2017; 18:22. https://doi.org/10.1186/s12931-017-0509-x PMID:28103932

22. Abraham T, Hogg J. Extracellular matrix remodeling of lung alveolar walls in three dimensional space identified using second harmonic generation and multiphoton excitation fluorescence. J Struct Biol. 2010; 171:189–96.
https://doi.org/10.1016/j.jsb.2010.04.006
PMID:20412859

23. Abboud RT, Vimalanathan S. Pathogenesis of COPD. Part I. The role of protease-antiprotease imbalance in emphysema. Int J Tuberc Lung Dis. 2008; 12:361–67.
PMID:18371259

24. Dharwal V, Sandhir R, Naura AS. PARP-1 inhibition provides protection against elastase-induced emphysema by mitigating the expression of matrix metalloproteinases. Mol Cell Biochem. 2019; 457:41–49.
https://doi.org/10.1007/s11010-019-03510-1
PMID:30993494

25. Ren S, Guo LL, Yang J, Liu DS, Wang T, Chen L, Chen YJ, Xu D, Feng YL, Wen FQ. Doxycycline attenuates acrolein-induced mucin production, in part by inhibiting MMP-9. Eur J Pharmacol. 2011; 650:418–23.
https://doi.org/10.1016/j.ejphar.2010.10.034
PMID:21036164

26. Wain LV, Shrine N, Artigas MS, Erzurumluoglu AM, Noyvert B, Bossini-Castillo L, Obeidat M, Henry AP, Portelli MA, Hall RJ, Billington CK, Rimington TL, Fenech AG, et al, and Understanding Society Scientific Group, and Geisinger-Regeneron DiscovEHR Collaboration. Genome-wide association analyses for lung function and chronic obstructive pulmonary disease identify new loci and potential druggable targets. Nat Genet. 2017; 49:416–25.
https://doi.org/10.1038/ng.3787
PMID:28166213

27. Tjin G, Xu P, Kable SH, Kable EP, Burgess JK. Quantification of collagen I in airway tissues using second harmonic generation. J Biomed Opt. 2014; 19:36005.
https://doi.org/10.1117/1.JBO.19.3.036005
PMID:24604535

28. Edwards MR, Bartlett NW, Clarke D, Birrell M, Belvisi M, Johnston SL. Targeting the NF-kappaB pathway in asthma and chronic obstructive pulmonary disease. Pharmacol Ther. 2009; 121:1–13.
https://doi.org/10.1016/j.pharmthera.2008.09.003
PMID:18950657

29. He X, Li T, Luo L, Zeng H, Chen Y, Cai S. PRMT6 mediates inflammation via activation of the NF-κB/p65 pathway on a cigarette smoke extract-induced murine emphysema model. Tob Induc Dis. 2020; 18:8.
https://doi.org/10.18332/tid/116413
PMID:32047419

30. Zhou X, Gu D, Hou G. Erythromycin attenuates metalloprotease/anti-metalloprotease imbalance in cigarette smoke-induced emphysema in rats via the mitogen-activated protein kinase/nuclear factor-κB activation pathway. Mol Med Rep. 2017; 15:2983–90.
https://doi.org/10.3892/mmr.2017.6416
PMID:28358431

31. Aschner Y, Downey GP. Transforming growth factor-β: master regulator of the respiratory system in health and disease. Am J Respir Cell Mol Biol. 2016; 54:647–55.
https://doi.org/10.1165/rcmb.2015-0391TR
PMID:26796672

32. Takizawa H, Tanaka M, Takami K, Ohtoshi T, Ito K, Satoh M, Okada Y, Yamasawa F, Nakahara K, Umeda A. Increased expression of transforming growth factor-beta1 in small airway epithelium from Tobacco smokers and patients with chronic obstructive pulmonary disease (COPD). Am J Respir Crit Care Med. 2001; 163:1476–83.
https://doi.org/10.1164/ajrccm.163.6.9908135
PMID:11371421

33. Koenders MM, Wismans RG, Starcher B, Hamel BC, Dekhuijzen RP, van Kuppevelt TH. Fibrillin-1 staining anomalies are associated with increased staining for TGF-beta and elastic fibre degradation; new clues to the pathogenesis of emphysema. J Pathol. 2009; 218:446–57.
https://doi.org/10.1002/path.2548 PMID:19373854

34. Su BH, Tseng YL, Shieh GS, Chen YC, Wu P, Shiau AL, Wu CL. Over-expression of prothymosin-α antagonizes TGFβ signalling to promote the development of emphysema. J Pathol. 2016; 238:412–22.
https://doi.org/10.1002/path.4664 PMID:26496995

35. Wouters EF, Wouters BB, Augustin IM, Franssen FM. Personalized medicine and chronic obstructive pulmonary disease. Curr Opin Pulm Med. 2017; 23:241–46.
https://doi.org/10.1097/MCP.0000000000000377
PMID:28257315

36. Sarvaria A, Madrigal JA, Saudemont A. B cell regulation in cancer and anti-tumor immunity. Cell Mol Immunol. 2017; 14:662–74.
https://doi.org/10.1038/cmi.2017.35 PMID:28626234

37. Qu X, Tang Y, Hua S. Immunological approaches towards cancer and inflammation: a cross talk. Front Immunol. 2018; 9:563.
https://doi.org/10.3389/fimmu.2018.00563
PMID:29662489

38. Pickup MW, Mouw JK, Weaver VM. The extracellular matrix modulates the hallmarks of cancer. EMBO Rep. 2014; 15:1243–53.

https://doi.org/10.15252/embr.201439246
PMID:25381661

39. Liao Z, Dong J, Hu X, Wang T, Wan C, Li X, Li L, Guo L, Xu D, Wen F. Enhanced expression of human β-defensin 2 in peripheral lungs of patients with chronic obstructive pulmonary disease. Peptides. 2012; 38:350–56.
https://doi.org/10.1016/j.peptides.2012.09.013
PMID:23000304

40. Ladjemi MZ, Martin C, Lecocq M, Detry B, Nana FA, Moulin C, Weynand B, Fregimilicka C, Bouzin C, Thurion P, Carlier F, Serré J, Gayan-Ramirez G, et al. Increased IgA expression in lung lymphoid follicles in severe chronic obstructive pulmonary disease. Am J Respir Crit Care Med. 2019; 199:592–602.
https://doi.org/10.1164/rccm.201802-0352OC
PMID:30339768

41. Goddard PR, Nicholson EM, Laszlo G, Watt I. Computed tomography in pulmonary emphysema. Clin Radiol. 1982; 33:379–87.
https://doi.org/10.1016/s0009-9260(82)80301-2
PMID:7083738

42. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics. 2008; 9:559.
https://doi.org/10.1186/1471-2105-9-559
PMID:19114008

43. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. 2003; 13:2498–504.
https://doi.org/10.1101/gr.1239303 PMID:14597658

SUPPLEMENTARY MATERIALS

Supplementary Figure
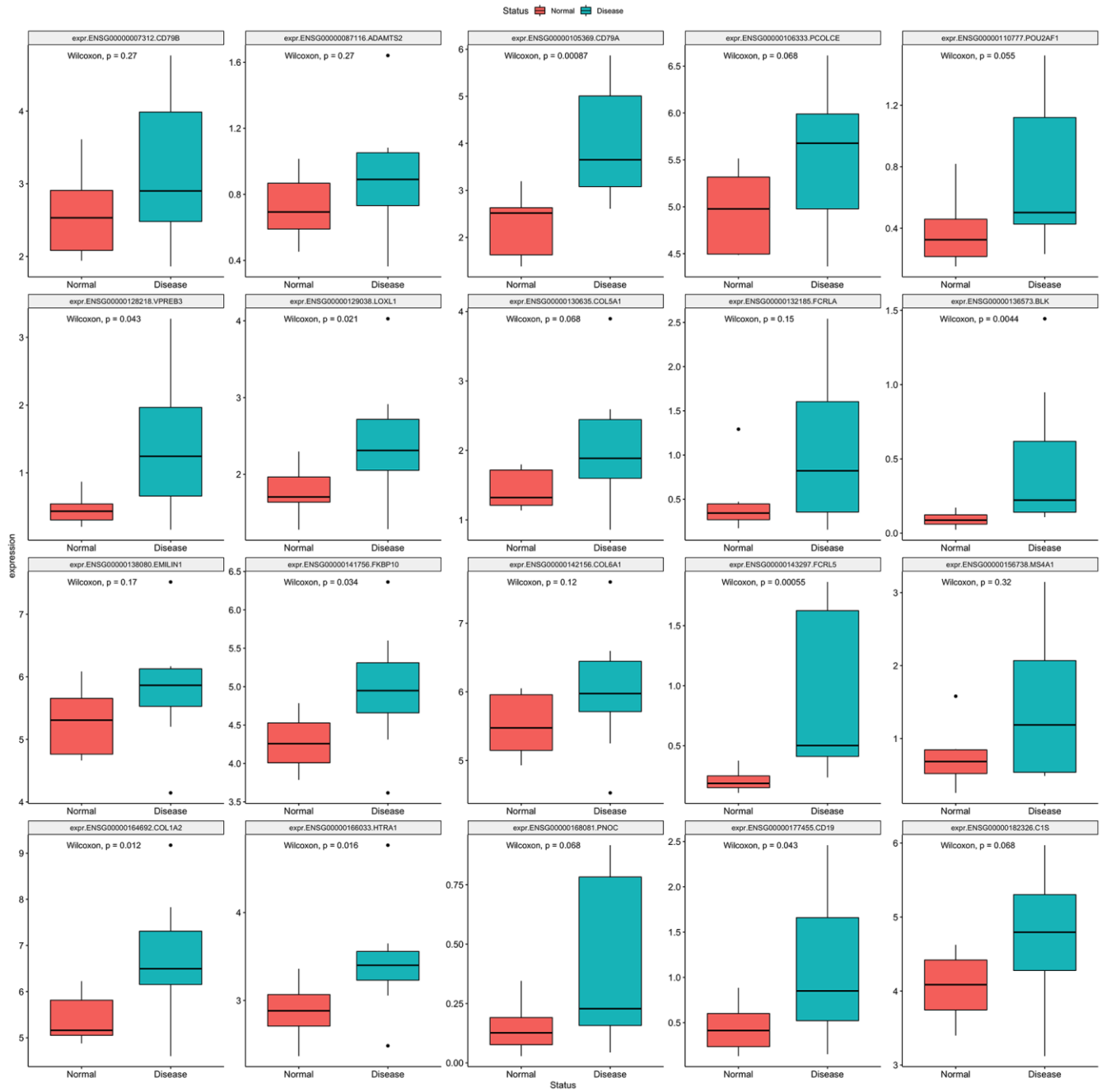
Supplementary Figure 1. Expression of top 10 hub genes in the yellow-green and brown modules.

Supplementary File

Please browse Full Text version to see the data of Supplementary File 1.

Supplementary File 1. CT figure of included patients.