Original Article

Integrated analysis reveals lung fibrinogen gamma chain as a biomarker for chronic obstructive pulmonary disease

Hai Zhang[1,2#], Chenfei Li[2#], Xiaomin Song[2#], Lei Cheng[2], Qi Liu[2], Na Zhang[3], Liangyu Wei[2], Kianfan Chung[4], Ian M. Adcock[4], Chunhua Ling[1], Feng Li[2]

[1]Department of Pulmonary and Critical Care Medicine, The First Affiliated Hospital of Soochow University, Suzhou, China; [2]Department of Pulmonary and Critical Care Medicine, Shanghai Chest Hospital, Shanghai Jiao Tong University, Shanghai, China; [3]Central Laboratory, Shanghai Chest Hospital, Shanghai Jiao Tong University, Shanghai, China; [4]Section of Airways Disease, National Heart and Lung Institute, Imperial College London, London, UK

Contributions: (I) Conception and design: C Ling, F Li; (II) Administrative support: C Ling, F Li; (III) Provision of study materials or patients: H Zhang, C Li, X Song; (IV) Collection and assembly of data: H Zhang, C Li, X Song; (V) Data analysis and interpretation: All authors; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

[#]These authors contributed equally to this work.

Correspondence to: Chunhua Ling. Department of Pulmonary and Critical Care Medicine, The First Affiliated Hospital of Soochow University, Suzhou 215006, China. Email: linchunhua88@hotmail.com; Feng Li. Department of Pulmonary and Critical Care Medicine, Shanghai Chest Hospital, Shanghai Jiao Tong University, Shanghai 200030, China. Email: lifeng741@aliyun.com.

Background: Chronic obstructive pulmonary disease (COPD) is a common, preventable, and treatable airway disease. This study aimed to identify key genes related to COPD pathogenesis through an integrated transcriptomic and proteomic analysis of lung tissue from COPD subjects undergoing lung resection for malignancy.

Methods: We performed RNA sequencing, gene expression analysis, and gene set enrichment analysis (GSEA) on lung tissue in 13 non-smokers, 16 smokers, and 16 COPD patients. Key genes were verified by RT-qPCR, immunohistochemistry and Western blot in human lung tissues, cigarette smoke extract (CSE)-exposed human bronchial epithelial cell line (BEAS-2B cells), and a cigarette smoke (CS)-induced mouse model.

Results: There were 521 differentially expressed genes between non-smokers and smokers, 57 genes between smokers and COPD patients, and 860 genes between non-smokers and COPD patients. Fibrinogen gamma chain (FGG) was highly expressed in COPD patients versus smokers and in COPD patients versus healthy controls. GSEA of the COPD patients with the highest FGG expression were enriched in the B cell receptor signaling pathway, pantothenate and CoA biosynthesis, Fc epsilon RI signaling pathway, and the Toll-like receptor (TLR) signaling pathway. RT-PCR analysis confirmed enhanced FGG mRNA levels in the lungs of both smokers and COPD patients compared to non-smokers and in CSE-exposed cells compared to control cells. FGG protein levels were elevated in the lungs of COPD patients and smokers compared to non-smokers and in the lungs of CS-exposed mice compared to control mice.

Conclusions: FGG may serve as a biomarker for COPD and may play an important role in its pathogenesis.

Keywords: Chronic obstructive pulmonary disease (COPD); fibrinogen gamma chain (FGG); RNA-sequence; smoke

Submitted Oct 19, 2021. Accepted for publication Dec 14, 2021.
doi: 10.21037/atm-21-5974
View this article at: https://dx.doi.org/10.21037/atm-21-5974

Introduction

Chronic obstructive pulmonary disease (COPD) is a common chronic disease involving airflow obstruction which can develop into cor pulmonale and respiratory failure (1). Five million people die each year from COPD, making it the third leading cause of death worldwide in 2020 (2). Treatments that effectively halt the disease progression, reduce the frequency and/or severity of exacerbations, or reverse the disease course are urgently required. COPD results from the complex interaction of many immune and structural cells exposed to toxic gases and particles such as cigarette smoke (CS) and particulate matter (PM) combined with a variety of host factors including genetics, airway hyper-responsiveness, and poor lung growth during childhood (3,4). The pathophysiological processes of COPD include effects on innate and adaptive Th1-type immunity in the airways and/or alveoli leading to diversity in clinical presentations, response to treatment, and progression patterns (5). Although the mechanistic understanding of COPD pathophysiology has improved, the identification of pathways that could be used for its diagnosis, treatment, are lacking.

COPD is an inflammatory disease and is associated with an increased risk of lung cancer. Lipopolysaccharide (LPS)-mediated chronic inflammation creates a favorable immunosuppressive microenvironment for tumor progression (6). In COPD patients, Th1 cell populations are expanded in both the lung and tumor microenvironment, and non-small cell lung cancer (NSCLC) patients with COPD have a longer progression-free survival after treatment with immune checkpoint inhibitors (7).

A recent study identified 599 different genes in lung tissue between non-smokers and current smokers. Most genes returned to near normal levels with age, while some remained upregulated after quitting smoking (8). Many genomic studies have been conducted in peripheral blood, lung, and airway tissues of COPD patients. Some studies choose to download genetic data for analysis because it is difficult to obtain lung tissue. The Gene Expression Omnibus (GEO) database contains data on COPD patients from other countries, but we chose to collect lung tissues from Chinese patients for gene expression profile analysis. We hypothesized that expression profiling of lung tissue from non-smokers, smokers, and COPD patients could identify distinct sets of smoking- and COPD-related genes and pathways, and using statistical and gene set enrichment methods, we aimed to identify gene-sets and signal pathways that may play a key role in COPD pathogenesis. We used lung tissue samples and two different COPD models to test our hypothesis.

We present the following article in accordance with the ARRIVE reporting checklist (available at https://dx.doi.org/10.21037/atm-21-5974).

Methods

Study subjects and specimens

We obtained lung tissues from two groups of subjects. The first group, consisting of 13 non-smokers, 16 smokers, and 16 COPD patients aged between 53 and 75 admitted to Shanghai Chest Hospital for lung cancer surgery between July and August, 2019 were enrolled for transcriptomics study. Lung tissues from the second group of 18 subjects consisting of six non-smokers, six smokers, and six COPD patients were enrolled for PCR study, immunohistochemistry, and Western blot study, respectively, for validation purposes. The diagnosis of COPD was based on the criteria of the Global Initiative for Chronic Obstructive Lung Disease (GOLD) and forced expiratory volume in 1 s (FEV_1) and forced vital capacity (FVC) were detected by post-bronchodilator spirometry (FEV_1/FVC <70%) using a body plethysmograph (MasterScreen Body/Diff, Jaeger, Hoechberg, Germany). Immediately after the lung resection, lung tissues regarded as macroscopically normal were dissected far away from the cancerous tissues then stored in liquid nitrogen tanks. Complete clinical and follow-up data was obtained on all subjects. All procedures performed in this study involving human participants were in accordance with the Declaration of Helsinki (as revised in 2013). This study was approved by the institutional review board of Shanghai Chest Hospital (No. KS1969), and written informed consent was obtained from all subjects.

RNA sequencing and data analysis

Total RNA of each lung tissue was extracted using a RNAmini kit (Qiagen, Germany). Enrichment of mRNA, fragmentation, reverse transcription, library construction, Illumina Novaseq 6000, and data analysis were performed by Genergy Biotechnology, Co., Ltd. (Shanghai, China). Raw data was managed by Skewer and data quality was checked by FastQC v0.11.2 (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/).

Differentially expressed genes (DEGs) analysis

The differential analysis was performed by limma packages in R software (version 3.6.1, R foundation, Austria). The DEGs were identified using a selecting criteria P value <0.05 between non-smokers and smokers, smokers and COPD patients, and between non-smokers and COPD patients. DEGs were visualized by heatmap according to $\log_2$(fold-change).

Gene set enrichment analysis (GSEA)

To further investigate the biological pathways involved in COPD pathogenesis, GSEA was performed using GSEA version 4.1.0 from the Broad Institute at MIT. GSEA identified differentially activated pathways between COPD patients in the top and bottom quartiles for specific gene expression were labelled as high and low expressors respectively. KEGG gene sets in the biological process database from the Molecular Signatures Database-Msig DB (http://www.broad.mit.edu/gsea/msigdb/index.jsp) were used for enrichment analysis with default settings and 1,000 gene set permutations. A P value <0.05 was used to screen the enriched pathways in each group and for the analysis results, the gene set under the pathway of |NES| >1, NOM P value <0.05 was regarded as meaningful.

Human airway epithelial cell model

Bronchial epithelial (BESA-2B) cells (Shanghai Institutes for Biological Sciences, China Academy of Science, Shanghai, China) were cultured in high glucose Dulbecco's Modified Eagle's Medium (Hyclone, Logan, UT, USA) with 10% fetal bovine serum (Gibco, MA, USA), 100 U/mL penicillin, and 100 μg/mL streptomycin (Thermo Fisher Scientific, Waltham, MA, USA) in a cell incubator (5% CO_2, 37 ℃). CS extract (CSE) was sucked from two burning Marlboro cigarettes (Marlboro Red Label, Longyan Tobacco Industrial, Co., Ltd., Fujian, China, licensed by Philip Morris Switzerland) at a constant flow rate (8 mL/s) by a syringe into 10 mL serum-free medium. The medium mixed with CSE sterilized using a 0.22-μm filter (Millex-GS, Millipore; MA, USA) was considered 100% CSE solution. After adjusting the pH to 7.4, 100% CSE solution was further diluted with serum-free medium to 5% CSE and used to treat cells for 24 h to induce the cell injury model.

Animal model

Specific pathogen free male C57/BL6 mice aged 8–10 weeks were purchased from the Shanghai SLAC Laboratory Animal, Co., Ltd. (Shanghai, China) and randomly divided into a control group and smoke-exposed group with five mice in each group. The CS-exposed mice were placed in an exposure chamber (90 cm × 55 cm × 40 cm) 4 times a day for 1 h each time with a 1-h smoke-free interval, for 10 weeks. During the exposure, an iron frame holding five lit Marlboro cigarettes (10 mg of tar and 0.8 mg of nicotine) was placed beside the cage. On days 1 and 14, the model mice received an intranasal instillation of 10 μg of LPS dissolved in saline (0.1 mg/mL, Beyotime Biotechnology, Haimen, Jiangsu, China), while the control mice were exposed to filtered air and intranasally instilled with saline only. If the mouse exhibits significant weight loss, shortness of breath, and slow movement during the tobacco exposure process (long-term), the mouse is considered to be pain, suffering, or distress. In this case, the mouse should be euthanized and no longer accepted exposed. All experimental studies involving animals were approved by the Shanghai Science and Technology Committee (STCSM) (No. SYXK 2018-0016), and animals were handled in accordance with the hospital laboratory management requirements for the care and use of animals. A protocol was prepared before the study without registration.

Quantitative real-time PCR

Total RNA from lung tissues and BEAS-2B cells was extracted from cells using TRIzol reagent (TaKaRa, Dalian, Liaoning, China). Each cDNA was then amplified and used for reverse transcription following the manufacturer's instructions. The reaction mixture contained 1 μL cDNA, 2× ChamQ Universal SYBR qPCR Master Mix (Vazyme Biotech, Co., Ltd., Nanjing, China), 0.4 μL of each primer, and 9.2 μL ddH_2O, in a total of 20 μL. The expression of RNAs were quantified using an ABI ViiATM 7 System at 50 ℃ for 2 min, 95 ℃ for 10 min, followed by 50 cycles at 95 ℃ for 15 s and 60 ℃ for 1 min. Threshold cycle (Ct) values were obtained from SDS, Version 2.3 software (Applied Biosystems) to calculate $2^{-\Delta\Delta Ct}$ for analysis, and β-actin served as the internal control of fibrinogen gamma chain (FGG). All samples were tested in triplicate. Primers of FGG provided by BioTNT (Shanghai, China) were as follows:

- Human FGG: forward primer (5'-GCATTGCAGATTTCCTGTCTAC-3'); reverse primer (5'-TCAAAGTAGCAGCGTCTATCAT-3');
- Mouse FGG: forward primer

Table 1 Demographic characteristics of study patients

Characteristics	Non-smokers (N=13)	Smokers (N=16)	COPD (N=16)	F or χ^2	P value
Men (%)	46.15	100.0	100.0	30.259[a]	<0.001
Age	61.54±5.47	64.19±5.71	63.94±5.14	1.002	0.376
BMI (kg/m^2)	23.16±2.19	23.83±4.05	22.78±2.95	0.433	0.651
Smoking (pack-years)	0	43.52±23.23	42.50±20.58	0.229[b]	0.635[b]
FEV_1/pred (%)	105.40±9.64	94.79±12.48	76.24±13.87	47.221	<0.001
FEV_1/FVC ratio (%)	83.87±5.62	79.80±5.56	66.54±4.12	21.229	<0.001
CAT score	0.62±0.77	5.81±4.31	10.06±6.07	16.033	<0.001
Pathological type	AC [11], BL [2]	AC [8], SC [4], SCLC [2], BL [2]	AC [11], SC [2], SCLC [2], NSCLC [1]	–	–

[a], Fisher's exact test; [b], COPD *vs.* smokers. COPD, chronic obstructive pulmonary disease; BMI, body mass index; FEV_1/pred (%), forced expiratory volume in 1s percentage predicted; FVC, forced vital capacity; CAT score, COPD assessment test score; AC, adenocarcinoma; SC, squamous cell carcinoma; SCLC, small cell lung cancer; NSCLC, non-small cell lung cancer; BL, benign lesion.

(5'-GGACAACGACAACGATAAGTTT-3'); reverse primer (5'-CATAGAATACCAGCGGCTT TTC-3');

- β-actin: forward primer (5'-GGCCAACCGCGAGAAGATGAC-3'); reverse primer (5'-GGATAGCACAGCCTGGATAGCA AC-3').

Western blotting

Lung tissues and BEAS-2B cells were homogenized and lysed using RIPA buffer (Beyotime Biotechnology, Haimen, China) to extract total proteins. Thirty micrograms of protein quantified by BCA reagent (Beyotime) was loaded onto SDS-PAGE and transferred to 0.45 μm PVDF membrane. After being blocked for 2 h at 25 ℃ in 5% fat-free milk, the membrane was incubated with the primary FGG antibody (Sigma-Aldrich; St. Louis, MO, USA) (1:500) overnight at 4 ℃ and anti-rabbit IgG/HRP secondary antibody (Cell Signaling Technology, Danvers, MA, USA) (1:10,000) for 2 h at room temperature. Immunoblots were developed using a chemiluminescent agent (ECL, Millipore; MA, USA).

Immunohistochemistry

The localization and expression of FGG in lung tissues were examined by immunohistochemical staining. Human lung tissue samples were fixed in formalin after the section and embedded in paraffin, before 4-μm-thick sections were cut on slides for staining. Antigen retrieval was performed in pH 6.0 citrate-phosphate buffer at 100 ℃ for 40 min. The sections were incubated with anti-FGG antibody (Proteintech, Wuhan, China) overnight at 4 ℃ then incubated with biotinylated secondary antibodies for 1 h followed by visualization with diaminobenzidine (DAB) solution and counter-staining with hematoxylin. The average optical density of FGG staining was measured by Image J (https://imagej.nih.gov/ij/).

Statistical analysis

Clinical statistical analyses were performed using SPSS v25.0 (SPSS Inc., Chicago, IL, USA) and data were presented as mean value ± standard deviation (mean ± SD). FGG mRNA expression levels were first evaluated by the one-sample Kolmogorov-Smirnov goodness-of-fit test to determine whether they followed a normal distribution pattern, and depending on the results, Pearson's or the non-parametric Spearman's rank correlation were used to examine their association with continuous variables (pulmonary function). GraphPad Prism 8.0 (La Jolla, CA, USA) was used for statistical analysis and multiple comparisons were conducted by the one-way ANOVA test. According to the Dunnett T3 adjustment, statistical significance for multiple comparison testing was set at P<0.017, and a P value less than 0.05 was considered statistically significant.

Results

Study subjects

There were 13 non-smokers, 16 smokers, and 16 COPD patients in this study (*Table 1*), and their mean post-

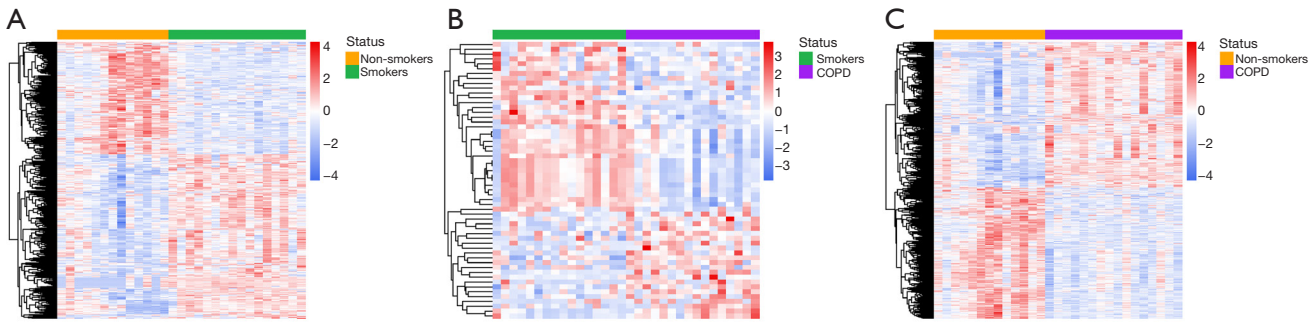

Figure 1 Heat maps of gene expression in the lung tissues. (A) A heat map of smokers *vs*. non-smokers. (B) A heat map of COPD patients *vs*. smokers. (C) A heat map of COPD patients *vs*. non-smokers. COPD, chronic obstructive pulmonary disease.

bronchodilator FEV_1 values were 105.4%, 94.79%, and 76.24%, respectively. There were no significant differences in age or BMI among the three groups of subjects, nor in in the smoking index between the smoker group and COPD group (P=0.991). The FEV_1/FVC ratio of the three groups was 83.87%, 79.80%, and 66.54%, respectively, and the CAT scores were 0.62, 5.81, and 10.06, with the COPD group having the highest score. Of the 45 cases of surgical pathology, 30 cases were adenocarcinoma (66.67%) and six cases were squamous carcinoma (13.33%). The pathological types of lung surgery in patients with COPD included adenocarcinoma (11, 68.75%), squamous cell carcinoma (2, 12.5%), small cell lung cancer (2, 12.5%), and NSCLC (1, 6.25%).

DEGs

There were 521 DEGs between non-smokers and smokers, 57 genes between smokers and COPD patients, and 860 genes between non-smokers and COPD patients (table available at https://cdn.amegroups.cn/static/public/10.21037atm-21-5974-1.pdf). Heat maps indicating the genes that were differentially expressed among the three groups are shown in *Figure 1* (P<0.05).

DEGs in lung tissue samples between smokers and non-smokers

A total of 521 DEGs were identified between smokers and non-smokers (P<0.05; *Figure 1A*) with three hundred and nine genes upregulated and 212 genes downregulated in smokers compared with non-smokers. The top 10 up- and down-regulated genes according to fold change and their known biological function are shown in Table S1.

DEGs in lung tissue samples between COPD patients and smokers

Fifty-seven DEGs were identified between COPD patients and smokers (P<0.05; *Figure 1B*) with 22 genes upregulated and 35 genes downregulated in COPD patients compared with smokers. Table S2 shows the top 10 up- and down-regulated genes respectively according to fold change together with their known biological function.

DEGs in lung tissue between COPD patients and non-smokers

Eight hundred and sixty DEGs were identified between COPD patients and non-smokers (P<0.05; *Figure 1C*) with 465 genes upregulated and 407 genes downregulated in COPD patients compared with non-smokers. Table S3 shows the top 10 up- and down-regulated genes according to fold change and their known biological function.

FGG was the most upregulated coding gene differentiating COPD from smokers and the fourth most DEG between COPD and healthy subjects. A schematic analysis of the DEGs across the three comparisons also identified a reduced number of COPD, smoking, and healthy non-smoker DEGs (*Figure 2*). This schematic analysis also identified FGG as the only common gene among the three gene sets (*Figure 2*).

FGG expression was significantly elevated in COPD patients versus non-smokers (*Figure 3*), and a significant spread of FGG expression across COPD patients was seen. Among the cases, the clusters analogous to non-smokers, smokers, and COPD patients had median (interquartile range) FGG expression scores of 0.87 (0.21–1.32), 1.95 (1.05–2.52), and 3.23 (1.87–4.51), respectively (F=11.353,

P<0.001).

FGG level is associated with the severity of pulmonary function

Scatter plot analysis of FGG levels [log_2fragments per kilobase of transcript per million fragments mapped (FPKM)] in relation to pulmonary function indices showed that non-smokers, smokers, and COPD patients form distinct clusters, resulting in a negative association of FGG levels with FEV_1/FVC (r=–0.495, P=0.001, *Figure 4A*). In contrast, no statistically significant association was observed between FGG levels with FEV_1 (r=–0.291, P=0.052, *Figure 4B*) and FVC (r =–0.201, P=0.187, *Figure 4C*).

Identification of FGG gene associated biological pathways

To determine whether this spread reflected different molecular phenotypes underlying separate COPD subtypes, we examined the pathways that distinguished FGG^{high} versus FGG^{low} COPD patients using the upper and lower quartiles of expression. GSEA was performed using data from

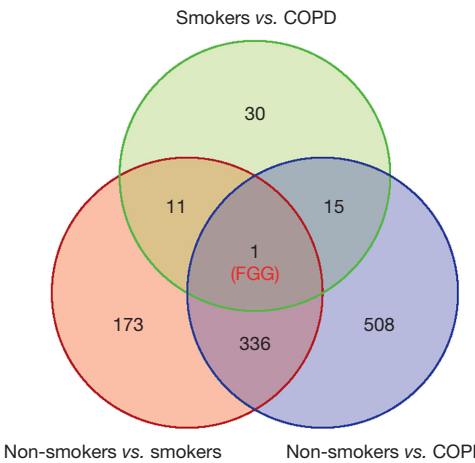

Figure 2 Schematic Venn diagram of the number of DEGs in the three subject groups. DEGs, differentially expressed genes; COPD, chronic obstructive pulmonary disease; FGG, fibrinogen gamma chain.

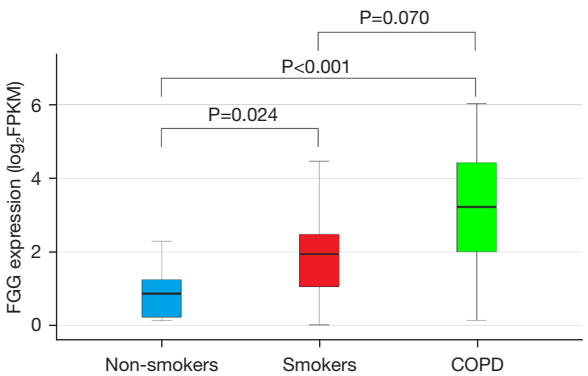

Figure 3 Boxplot showing the FGG expression for the three groups. Non-smokers are found to be at the lowest FGG expression, while COPD patients are at the highest FGG expression (P<0.001). FGG, fibrinogen gamma chain; COPD, chronic obstructive pulmonary disease; FPKM, fragments per kilobase of transcript per million fragments mapped.

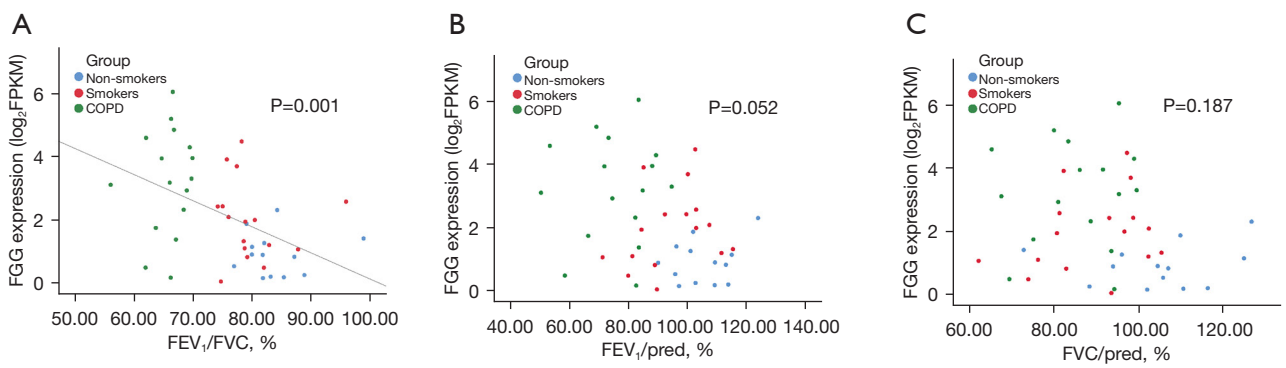

Figure 4 FGG expression level (log_2FPKM) in relation to pulmonary function in non-smokers (blue), smokers (red), and COPD patients (green). There is a significant correlation between FGG level and FEV_1/FVC (P=0.001). FGG, fibrinogen gamma chain; COPD, chronic obstructive pulmonary disease; FEV_1, forced expiratory volume in 1 s; FVC, forced vital capacity; FPKM, fragments per kilobase of transcript per million fragments mapped.

Table 2 Gene sets enriched in FGG^{high} compared to FGG^{low} COPD patients

GeneSet name	NES	Nominal P value	FDR
B cell receptor signaling pathway	1.62	<0.001	0.384
Pantothenate and CoA biosynthesis	1.55	0.037	0.604
Fc epsilon RI signaling pathway	1.47	0.020	0.993
Toll-like receptor signaling pathway	1.46	0.040	0.806

FGG, fibrinogen gamma chain; COPD, chronic obstructive pulmonary disease; NES, normalized enrichment score; FDR, false discovery rate; CoA, coenzyme A.

the COPD FGG^{high} and FGG^{low} subjects and a pathway was classified as enriched by setting the P value to <0.05. FGG^{high} subjects were enriched in B cell receptor signaling, pantothenate and coenzyme A (CoA) biosynthesis, Fc epsilon RI signaling pathway, and Toll-like receptor (TLR) signaling pathways (*Table 2*, Figure S1), indicating the potential role of FGG in the development of COPD.

mRNA expression of FGG in lung tissue and cells

To validate the RNA-sequencing data, we used RT-qPCR and showed a significant increase in FGG mRNA expression in the lungs of both smokers and COPD patients compared to non-smokers (P<0.001, *Figure 5A*). FGG mRNA levels were higher in COPD patients than in smokers (P<0.001, *Figure 5A*).

CS exposure induced a significantly higher expression of FGG mRNA in mice (P<0.001, *Figure 5B*), and CSE-exposed BEAS-2B cells showed a similar pattern (P<0.001, *Figure 5C*). Taken together, we reason that CS could be the cause of the increased expression of FGG in smokers and COPD patients. By monitoring the body weight and observing the behavior of the mice, no adverse reactions were found in any mice in this experiment.

Relative protein expression of FGG in lung tissue

FGG protein expression was detected in human and mouse lung tissues by Western blot. The results showed the protein levels of FGG in COPD patients were significantly elevated compared to non-smokers (P<0.01, *Figure 5D*), and there was an increasing trend between smokers and non-smokers. In addition, the protein levels of FGG were higher in CS-exposed mice compared to control mice (P<0.01, *Figure 5E*).

Localization and expression of FGG in lung tissue

The localization of FGG was mainly on airway epithelium and alveolar epithelium (*Figure 5F-5I*), while the protein expression of FGG indicated as average of optical density (AOD) was increased in COPD patients compared to non-smokers (P<0.01, *Figure 5I*). Similar to Western blot results, there was no difference in FGG protein between smokers and non-smokers and between COPD patients and smokers.

Discussion

There have been many studies on COPD biomarkers in the past. The sampling and processing of sputum and saliva have the advantages of practicality, non-invasiveness and low cost. Tao Dong *et al.* reviewed the clinical value of interleukin (IL)-6 and IL-8, matrix metalloproteinase (MMP)-8 and MMP-9, C-reactive protein (CRP), tumor necrosis factor-alpha (TNF-α), and neutrophil elastase (NE) in saliva and sputum (9). Blood-based biomarkers need to prove that they are derived from the lung or are related to the COPD disease mechanism. Using proteomics methods, it was found that SERPINA3 in plasma was significantly up-regulated in AECOPD compared with the recovery period (10). Serum resistin may be used as a biomarker of inflammation, risk assessment, severity and prognosis of COPD disease progression (11). A study of serological assessment suggested that type VI collagen turnover and elastin degradation by NE were associated with COPD-related inflammation and emphysema (12). Extracellular matrix (ECM) remodeling of the lung tissue releases protein fragments into the blood, Serum biomarkers of ECM turnover were significantly associated with disease severity and clinically relevant outcomes in patients with COPD. Serum levels of C3M, C6M, Pro-C3, Pro-C6, and EL-NE were associated with lung function (13). In the COPDGene study, protein biomarkers related to the severity and progress of COPD were identified, such as soluble receptor for advanced glycation endproducts (sRAGE) and IL-6, IL-8, ceramide and sphingomyelin (14). Spira *et al.* used high-density microarrays to measure gene expression to determine 102 genes in the lung tissues of 34 patients, which can accurately distinguish severe emphysema and non/mild emphysema lung tissues, and then use whole-genome linkage analysis to analyze the

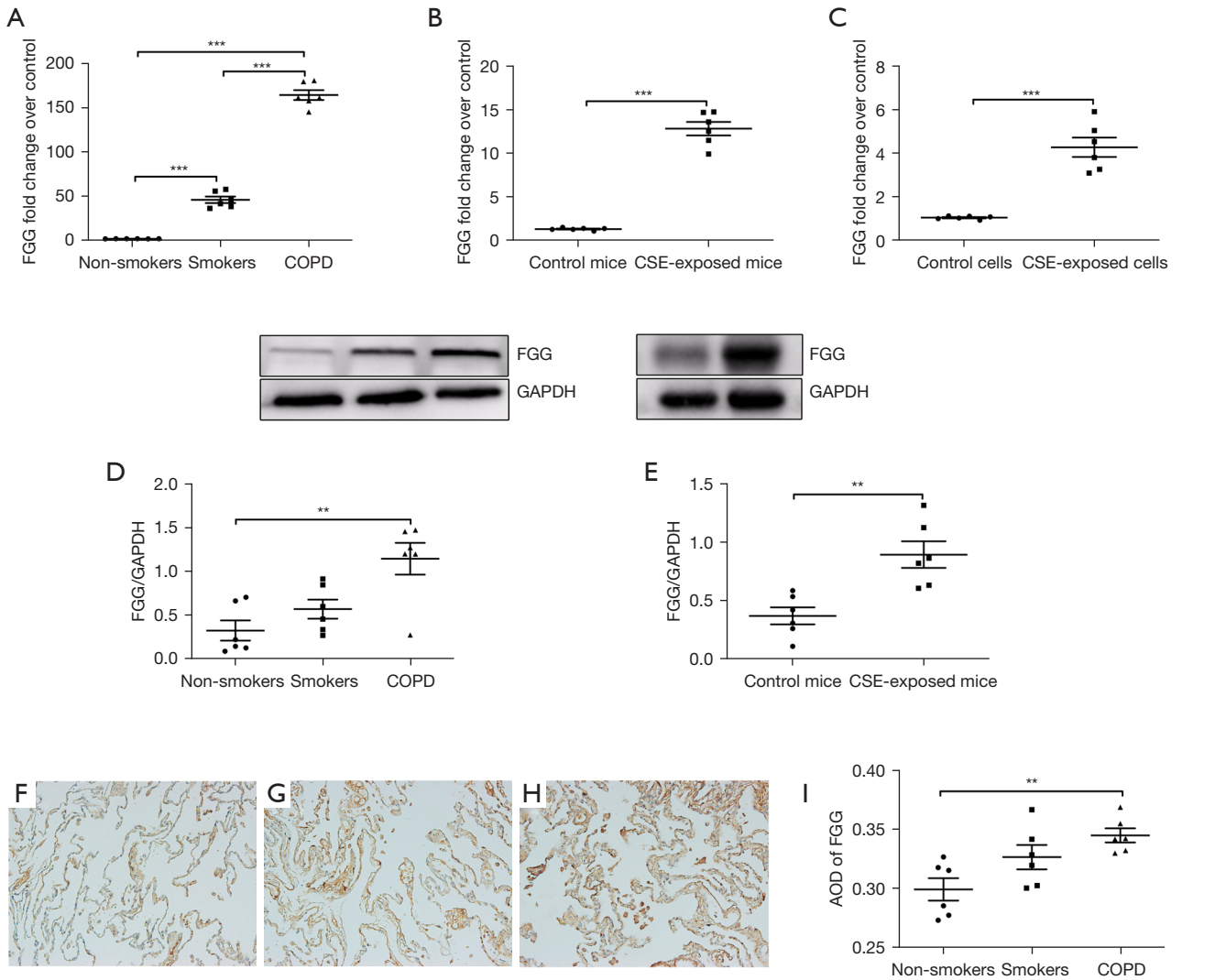

Figure 5 The results of FGG expression in lung tissues and cells. The mRNA expression of FGG in lung tissues from non-smokers, smokers, and COPD patients (A), from control mice and CS-exposed mice (B), and from control BEAS-2B cells and CSE-exposed BEAS-2B cells (C). The protein expression of FGG in lung tissue of human (D) and mice (E). Western blot analysis of the relative protein expression of FGG to GAPDH in (D,E). Representative immunohistochemical (IHC) staining for FGG in lung tissue of non-smoker (F), smoker (G), and COPD patient (H) (magnification, ×200). Individual and mean AOD of FGG measured in lung immunohistochemical sections. (I) Average optical density of FGG in the lung tissue. The FGG expression levels in COPD patients were significantly increased compared with non-smokers. **, P<0.01; ***, P<0.001, compared with non-smokers, control mice, and control cells. FGG, fibrinogen gamma chain; COPD, chronic obstructive pulmonary disease; GAPDH, glyceraldehyde-3-phosphate dehydrogenase; CS, cigarette smoke; CSE, cigarette smoke extract; AOD, average of optical density.

correlation between these genes and COPD (15). Faner *et al.* performed transcriptomics (Affymetrix array) analysis on the lung tissue of 70 former smokers with COPD to compare the differential gene expression of bronchiolitis and emphysema (16). It was found that B cell recruitment and immunoglobulin transcription genes (*CXCL13*, *CCL19* and *POU2AF1*) were associated with the severity of emphysema (16).

Transcriptomic analysis in lung tissue identified FGG as a novel COPD-associated gene which was confirmed using RT-qPCR and Western blotting. FGG was elevated to a lesser extent in healthy smokers and CS exposure enhanced

FGG mRNA and protein expression in a mouse model of COPD and FGG expression in human airway epithelial cells. GSEA of FGG high versus low expressing COPD patients demonstrated a potential role for B cells (B cell receptor signaling pathway), CoA synthesis (pantothenate and CoA biosynthesis), and mast cells (MC) (Fc epsilon RI signaling pathway) in driving COPD in these patients. A clinicopathological analysis of liver cancer showed that the up-regulation of FGG expression in cells was significantly associated with increased vascular infiltration, increased satellite nodules, and increased TNM staging (17). Our results show that FGG is related to immune cells. Yu *et al.* also found that FGG is highly expressed in COPD and is significantly related to activated dendritic cells, which may affect COPD's immune regulation mechanism and COPD through dendritic cells (18).

Smoking is considered the leading cause of lung dysfunction, and the lung function decline of smokers is much greater than expected with age (19). Epidemiological data demonstrate that long-term smokers are at increased risk of developing COPD and lung cancer compared to non-smokers. Genetic changes caused by smoking can alter susceptibility to respiratory diseases and alter gene expression profiles which may accelerate the progress of respiratory disorders (20). The gaseous and particulate phases of cigarettes contain more than 4,500 components such as harmful particles, reactive chemicals, and free radicals, which can adjust the function of airway structural and immune cells (21). CS results in the activation of structural cells, the accumulation of neutrophils, monocytes, lymphocytes and dendritic cells, and the release of proinflammatory mediators, reactive oxygen species (ROS), and proteolytic enzymes considered as crucial players in the development of COPD (22,23).

FGG is downregulated during epithelial-mesenchymal transition and encodes the γ chain of fibrinogen (24). A recent study showed that FGG expression was higher in former smokers than in non-smokers in people undergoing lung adenocarcinoma surgery. In addition, FGG was upregulated among former smokers compared with those who never smoked without cancer (25). This upregulation of FGG by smoking in lung cancer patients is consistent with our data. To reduce the heterogeneity, Ham *et al.* deconvoluted COPD samples and found that acute-phase proteins such as FGG and FGA were always upregulated in COPD (26). Kim *et al.* analyzed 189 samples in the GEO database and in finding 98 COPD samples including 36 cases of adenocarcinoma and 62 cases of squamous cell carcinoma, confirmed several new COPD-related genes including FGG (27). In a bioinformatics analysis, Huang *et al.* obtained 14 hub DEGs and found five genes including FGG were upregulated in COPD (28). The mechanism for this induction is unclear and how FGG may regulate COPD is still unknown. GSEA revealed that FGG overexpression was positively correlated with the B cell receptor signaling pathway, pantothenate and CoA biosynthesis, Fc epsilon RI signaling pathway, and TLR signaling pathway. This is consistent with the changes in lung tissue protection mechanisms caused by smoking toxicity (20). Mouse models may help our understanding of these processes. The adaptive immune response includes B cell-mediated humoral immune response, and co-expression analysis of lung tissues of patients with COPD revealed several B-cell related genes which were positively correlated with emphysema (29). DEGs analysis determined that several B-cell-related genes were upregulated in COPD patients with emphysema (16). Some elements of the B-cell receptor signaling pathway were also upregulated in severe emphysema regions (30). Igα (CD79A) and Igβ (CD79B) are the main signal transduction molecules and transduce the signal generated by antigen binding to B cell receptor, and there was a significant correlation between the volume fraction of airway wall and the volume fraction of CD79A positive alveolar tissue and the increase of linear intercept (Lm) (30).

CoA biosynthesis in humans requires cysteine, pantothenic acid, and adenosine triphosphate (ATP) and is mainly involved in the metabolism of fatty acids and pyruvate (31). Mitochondria, especially fatty acid disorders, have attracted more and more attention in COPD (32-34) and several mitochondrial-directed drugs are being developed to treat the disease (35). This pathway may also be actively involved in inducing the metabolic changes seen in COPD, which is currently an ongoing hot research area (36).

The FcεRI-mediated signal transduction pathway in MC is initiated by the interaction of antigen and IgE with the extracellular domain of the FcεRI α chain, and there are many similarities in the signal transduction pathways in MC, T cells, and B cells. The immunoreceptor tyrosine-based activation motif (ITAM) exists in the β and γ subunits of FcεRI and in the Igα and Igβ of the B cell receptor. ITAM is essential for cell activation (37), and epithelial integrity and thickness of the laminin layer was related to MC density (38). In addition, compared with the healthy control group, the density of tryptase-positive MC (MC-T) in the subepithelial area of the central airways of COPD patients was lower (39). The density of MC-T was also positively correlated with

FEV_1/FVC (39). It has been reported that MCs may participate in the functional effect of enhanced IL-17A expression in COPD. IL-17A induces the release of the proangiogenic factor fibroblast growth factor (FGF)-2 and vascular endothelial growth factor (VEGF) through IL-17 receptor on MCs, thereby mediating vascular remodeling (40).

Several TLRs are implicated in COPD pathogenesis. TLRs are the main components of the innate immune system, and their main function is to recognize the innate molecular structure of pathogens. TLR3, TLR7, and TLR9 are involved in the inflammatory response of COPD virus infection (41-46), whereas TLR2 and TLR4 are the most important receptors in response to bacterial infection and the pathogenesis of COPD (47-50). Evidence suggests that TLR2 and TLR4 have a complex inter-relationship in COPD, with TLR4 being detrimental and TLR2 activation being of some benefit in disease. Future research may determine the benefit of deactivating TLR2 and/or inhibit TLR4 to treat COPD (51).

We suspected that FGG may play its biological function in COPD by regulating B cell receptor signaling, mitochondrial and free fatty metabolism, MC, and TLR signaling. One study reported that FGG mRNA expression correlated with the total lung burden of PM and its burden in the lung parenchyma (52). We did not analyze PM in our study, but this may help link the exposure to environmental pollution and the upregulation of FGG in smokers and in COPD patients, and stimulation of key immune and metabolic pathways that possibly reflects enhanced immune surveillance/response to noxious stimuli. Animal models of COPD will help resolve this issue and enable proof of concept studies using FGG antagonists.

Although our study has many strengths including the use of human lung samples and two different models of COPD to demonstrate enhanced FGG expression in response to smoking with a higher expression in COPD patients, there are nevertheless some limitations. Determination of genetic characteristics, immune characteristics and transcriptional regulation mechanisms of COPD is complex and challenging. Technical methods such as exome sequencing, whole genome sequencing, and RNA sequencing have played an important role in gene identification and gene pathways. The potential for FGG to be truly causal of disease is uncertain due to the upregulation with CS exposure alone. FGG upregulation may alter the immune system to respond differentially to subsequent challenges driving the onset of disease. Further experiments in animal models may help resolve this. In addition, all subjects in our study had lung cancer and FGG as the intersect gene may reflect an impact of cancer since it has been implicated in many cancers (17,24,53). We did not determine the cellular site of FGG upregulation. The Human Protein Atlas does not show any information on FGG expression in the lung whilst the COPD Cell Atlas (https://p2med.shinyapps.io/copd-cell-atlas/) showed enhanced expression in mesothelial and alveolar epithelial cells. Another possible limitation of our study is that the sample size was small, and we are continuing to collect COPD lung tissue samples for further study. We also need to verify the specific mechanism of FGG influencing COPD through their signaling pathways. In the follow-up research, we will study the regulatory network of the FGG gene, identify its upstream and downstream genes, and explore its mechanism of action in COPD.

In summary, we found that FGG was highly expressed in the lung tissue of COPD, and the expression level of FGG mRNA was related to the pulmonary function of patients. Accordingly, inhibition of FGG in lung tissue may be an effective therapeutic strategy. Our study provides for the first time that FGG might regulate these biological progresses through the B cell receptor signaling pathway, pantothenate and CoA biosynthesis, Fc epsilon RI signaling pathway, and TLR signaling pathway, and may represent a useful target for therapy.

Acknowledgments

Funding: This work was supported by Project of National Nature Science Foundation of China (No. 82070041) and Shanghai Health Commission (No. 2020YJZX0115).

Footnote

Reporting Checklist: The authors have completed the ARRIVE reporting checklist. Available at https://dx.doi.org/10.21037/atm-21-5974

Data Sharing Statement: Available at https://dx.doi.org/10.21037/atm-21-5974

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at https://dx.doi.org/10.21037/atm-21-5974). The authors have no conflicts

of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. All experimental studies involving animals were approved by the STCSM (No. SYXK 2018-0016), and animals were handled in accordance with the hospital laboratory management requirements for the care and use of animals. All procedures performed in this study involving human participants were in accordance with the Declaration of Helsinki (as revised in 2013). This study was approved by the institutional review board of Shanghai Chest Hospital (No. KS1969), and written informed consent was obtained from all subjects.

References

1. Global Initiative for Chronic Obstructive Lung Disease Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease Global Initiative for Chronic Obstructive Lung Disease. 2020. Available online: http://www.goldcopd.org/. Accessed Oct 25 2020.
2. López-Campos JL, Tan W, Soriano JB. Global burden of COPD. Respirology 2016;21:14-23.
3. Keely S, Talley NJ, Hansbro PM. Pulmonary-intestinal cross-talk in mucosal inflammatory disease. Mucosal Immunol 2012;5:7-18.
4. Lange P, Celli B, Agustí A, et al. Lung-Function Trajectories Leading to Chronic Obstructive Pulmonary Disease. N Engl J Med 2015;373:111-22.
5. Riley CM, Sciurba FC. Diagnosis and Outpatient Management of Chronic Obstructive Pulmonary Disease: A Review. JAMA 2019;321:786-97.
6. Liu CH, Chen Z, Chen K, et al. Lipopolysaccharide-Mediated Chronic Inflammation Promotes Tobacco Carcinogen-Induced Lung Cancer and Determines the Efficacy of Immunotherapy. Cancer Res 2021;81:144-57.
7. Mark NM, Kargl J, Busch SE, et al. Chronic Obstructive Pulmonary Disease Alters Immune Cell Composition and Immune Checkpoint Inhibitor Efficacy in Non-Small Cell Lung Cancer. Am J Respir Crit Care Med 2018;197:325-36.
8. Bossé Y, Postma DS, Sin DD, et al. Molecular signature of smoking in human lung tissues. Cancer Res 2012;72:3753-63.
9. Dong T, Santos S, Yang Z, et al. Sputum and salivary protein biomarkers and point-of-care biosensors for the management of COPD. Analyst 2020;145:1583-604.
10. Kim SH, Ahn HS, Park JS, et al. A Proteomics-Based Analysis of Blood Biomarkers for the Diagnosis of COPD Acute Exacerbation. Int J Chron Obstruct Pulmon Dis 2021;16:1497-508.
11. Kumor-Kisielewska A, Kierszniewska-Stępień D, Pietras T, et al. Assessment of leptin and resistin levels in patients with chronic obstructive pulmonary disease. Pol Arch Med Wewn 2013;123:215-20.
12. Bihlet AR, Karsdal MA, Sand JM, et al. Biomarkers of extracellular matrix turnover are associated with emphysema and eosinophilic-bronchitis in COPD. Respir Res 2017;18:22.
13. Stolz D, Leeming DJ, Kristensen JHE, et al. Systemic Biomarkers of Collagen and Elastin Turnover Are Associated With Clinically Relevant Outcomes in COPD. Chest 2017;151:47-59.
14. Zemans RL, Jacobson S, Keene J, et al. Multiple biomarkers predict disease severity, progression and mortality in COPD. Respir Res 2017;18:117.
15. Spira A, Beane J, Pinto-Plata V, et al. Gene expression profiling of human lung tissue from smokers with severe emphysema. Am J Respir Cell Mol Biol 2004;31:601-10.
16. Faner R, Cruz T, Casserras T, et al. Network Analysis of Lung Transcriptomics Reveals a Distinct B-Cell Signature in Emphysema. Am J Respir Crit Care Med 2016;193:1242-53.
17. Zhang X, Wang F, Huang Y, et al. FGG promotes migration and invasion in hepatocellular carcinoma cells through activating epithelial to mesenchymal transition. Cancer Manag Res 2019;11:1653-65.
18. Yu H, Guo W, Liu Y, et al. Immune Characteristics Analysis and Transcriptional Regulation Prediction Based on Gene Signatures of Chronic Obstructive Pulmonary Disease. Int J Chron Obstruct Pulmon Dis 2021;16:3027-39.
19. Corbin RP, Loveland M, Martin RR, et al. A four-year

follow-up study of lung mechanics in smokers. Am Rev Respir Dis 1979;120:293-304.
20. Zhou Z, Chen P, Peng H. Are healthy smokers really healthy? Tob Induc Dis 2016;14:35.
21. Sarir H, Mortaz E, Janse WT, et al. IL-8 production by macrophages is synergistically enhanced when cigarette smoke is combined with TNF-alpha. Biochem Pharmacol 2010;79:698-705.
22. Reynolds PR, Cosio MG, Hoidal JR. Cigarette smoke-induced Egr-1 upregulates proinflammatory cytokines in pulmonary epithelial cells. Am J Respir Cell Mol Biol 2006;35:314-9.
23. Xu X, Wang H, Wang Z, et al. Plasminogen activator inhibitor-1 promotes inflammatory process induced by cigarette smoke extraction or lipopolysaccharides in alveolar epithelial cells. Exp Lung Res 2009;35:795-805.
24. Wang H, Meyer CA, Fei T, et al. A systematic approach identifies FOXA1 as a key factor in the loss of epithelial traits during the epithelial-to-mesenchymal transition in lung cancer. BMC Genomics 2013;14:680.
25. Pintarelli G, Noci S, Maspero D, et al. Cigarette smoke alters the transcriptome of non-involved lung tissue in lung adenocarcinoma patients. Sci Rep 2019;9:13039.
26. Ham S, Oh YM, Roh TY. Evaluation and Interpretation of Transcriptome Data Underlying Heterogeneous Chronic Obstructive Pulmonary Disease. Genomics Inform 2019;17:e2.
27. Kim DY, Kim WJ, Kim JH, et al. Identification of Putative Regulatory Alterations Leading to Changes in Gene Expression in Chronic Obstructive Pulmonary Disease. Mol Cells 2019;42:333-44.
28. Huang X, Li Y, Guo X, et al. Identification of differentially expressed genes and signaling pathways in chronic obstructive pulmonary disease via bioinformatic analysis. FEBS Open Bio 2019;9:1880-99.
29. Qin J, Yang T, Zeng N, et al. Differential coexpression networks in bronchiolitis and emphysema phenotypes reveal heterogeneous mechanisms of chronic obstructive pulmonary disease. J Cell Mol Med 2019;23:6989-99.
30. Campbell JD, McDonough JE, Zeskind JE, et al. A gene expression signature of emphysema-related lung destruction and its reversal by the tripeptide GHK. Genome Med 2012;4:67.
31. Naquet P, Kerr EW, Vickers SD, et al. Regulation of coenzyme A levels by degradation: the 'Ins and Outs'. Prog Lipid Res 2020;78:101028.
32. Wiegman CH, Michaeloudes C, Haji G, et al. Oxidative stress-induced mitochondrial dysfunction drives inflammation and airway smooth muscle remodeling in patients with chronic obstructive pulmonary disease. J Allergy Clin Immunol 2015;136:769-80.
33. Prihandoko R, Kaur D, Wiegman CH, et al. Pathophysiological regulation of lung function by the free fatty acid receptor FFA4. Sci Transl Med 2020;12:eaaw9009.
34. Haji G, Wiegman CH, Michaeloudes C, et al. Mitochondrial dysfunction in airways and quadriceps muscle of patients with chronic obstructive pulmonary disease. Respir Res 2020;21:262.
35. Manevski M, Muthumalage T, Devadoss D, et al. Cellular stress responses and dysfunctional Mitochondrial-cellular senescence, and therapeutics in chronic respiratory diseases. Redox Biol 2020;33:101443.
36. Michaeloudes C, Bhavsar PK, Mumby S, et al. Role of Metabolic Reprogramming in Pulmonary Innate Immunity and Its Impact on Lung Diseases. J Innate Immun 2020;12:31-46.
37. Zhang J, Berenstein EH, Evans RL, et al. Transfection of Syk protein tyrosine kinase reconstitutes high affinity IgE receptor-mediated degranulation in a Syk-negative variant of rat basophilic leukemia RBL-2H3 cells. J Exp Med 1996;184:71-9.
38. Ekberg-Jansson A, Amin K, Bake B, et al. Bronchial mucosal mast cells in asymptomatic smokers relation to structure, lung function and emphysema. Respir Med 2005;99:75-83.
39. Gosman MM, Postma DS, Vonk JM, et al. Association of mast cells with lung function in chronic obstructive pulmonary disease. Respir Res 2008;9:64.
40. Roos AB, Mori M, Gura HK, et al. Increased IL-17RA and IL-17RC in End-Stage COPD and the Contribution to Mast Cell Secretion of FGF-2 and VEGF. Respir Res 2017;18:48.
41. Mortaz E, Adcock IM, Ito K, et al. Cigarette smoke induces CXCL8 production by human neutrophils via activation of TLR9 receptor. Eur Respir J 2010;36:1143-54.
42. Koarai A, Yanagisawa S, Sugiura H, et al. Cigarette smoke augments the expression and responses of toll-like receptor 3 in human macrophages. Respirology 2012;17:1018-25.
43. Nadigel J, Préfontaine D, Baglole CJ, et al. Cigarette smoke increases TLR4 and TLR9 expression and induces cytokine production from CD8(+) T cells in chronic obstructive pulmonary disease. Respir Res 2011;12:149.
44. Freeman CM, Martinez FJ, Han MK, et al. Lung CD8+ T cells in COPD have increased expression of bacterial TLRs. Respir Res 2013;14:13.

45. Pomerenke A, Lea SR, Herrick S, et al. Characterization of TLR-induced inflammatory responses in COPD and control lung tissue explants. Int J Chron Obstruct Pulmon Dis 2016;11:2409-17.
46. Gimenes-Junior J, Owuar N, Vari HR, et al. FOXO3a regulates rhinovirus-induced innate immune responses in airway epithelial cells. Sci Rep 2019;9:18180.
47. Baines KJ, Simpson JL, Gibson PG. Innate immune responses are increased in chronic obstructive pulmonary disease. PLoS One 2011;6:e18426.
48. Zuo L, Lucas K, Fortuna CA, et al. Molecular Regulation of Toll-like Receptors in Asthma and COPD. Front Physiol 2015;6:312.
49. Hansbro PM, Haw TJ, Starkey MR, et al. Toll-like receptors in COPD. Eur Respir J 2017;49:1700739.
50. McGrath JJC, Stampfli MR. The immune system as a victim and aggressor in chronic obstructive pulmonary disease. J Thorac Dis 2018;10:S2011-7.
51. Haw TJ, Starkey MR, Pavlidis S, et al. Toll-like receptor 2 and 4 have opposing roles in the pathogenesis of cigarette smoke-induced chronic obstructive pulmonary disease. Am J Physiol Lung Cell Mol Physiol 2018;314:L298-317.
52. Ling SH, McDonough JE, Gosselink JV, et al. Patterns of retention of particulate matter in lung tissues of patients with COPD: potential role in disease progression. Chest 2011;140:1540-9.
53. Liu YL, Yan ZX, Xia Y, et al. Ligustrazine reverts anthracycline chemotherapy resistance of human breast cancer by inhibiting JAK2/STAT3 signaling and decreasing fibrinogen gamma chain (FGG) expression. Am J Cancer Res 2020;10:939-52.

(English Language Editor: B. Draper)

Cite this article as: Zhang H, Li C, Song X, Cheng L, Liu Q, Zhang N, Wei L, Chung K, Adcock IM, Ling C, Li F. Integrated analysis reveals lung fibrinogen gamma chain as a biomarker for chronic obstructive pulmonary disease. Ann Transl Med 2021;9(24):1765. doi: 10.21037/atm-21-5974

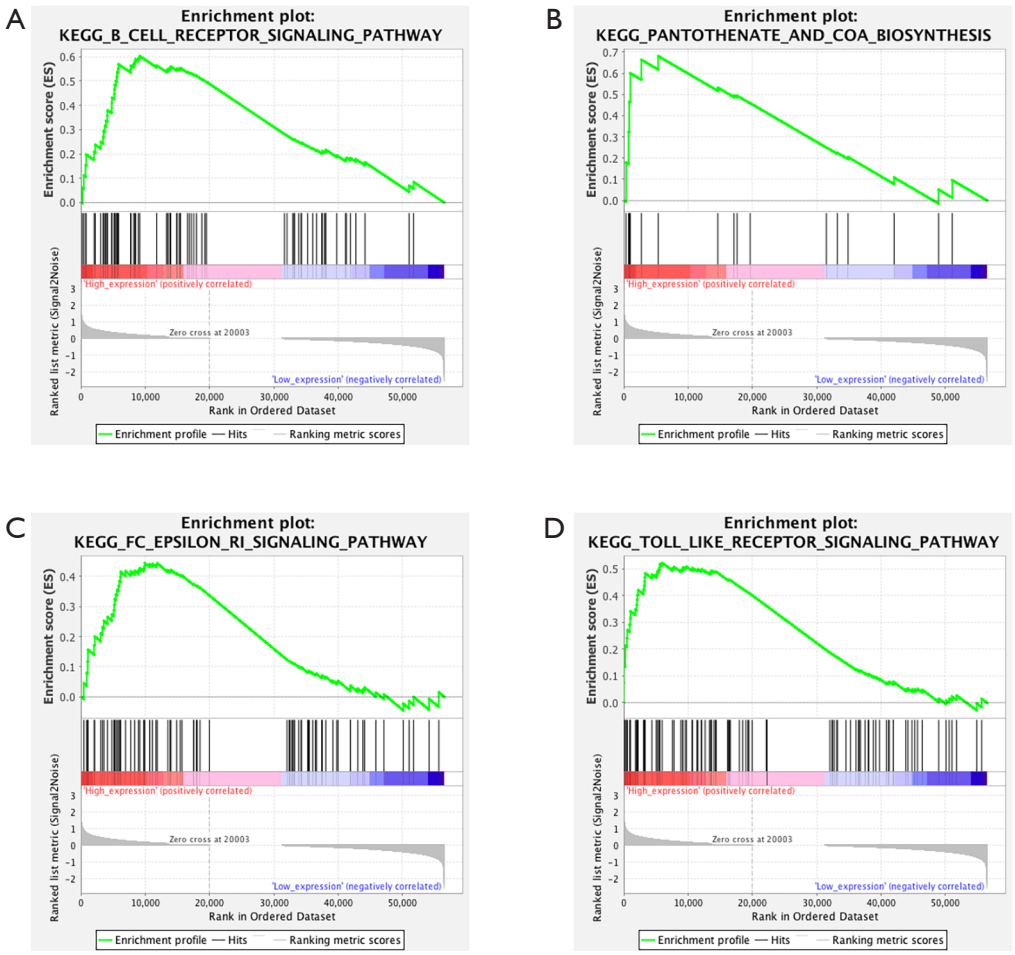

Figure S1 Gene set enrichment analysis (GSEA) associated with FGG expression. The gene sets of "B cell receptor signaling pathway" (A), "pantothenate and CoA biosynthesis" (B), "Fc epsilon RI signaling pathway" (C) and "toll-like receptor signaling pathway" (D) were enriched in COPD samples with FGG highly expressed. FGG, fibrinogen gamma chain; CoA, coenzyme A; COPD, chronic obstructive pulmonary disease.

https://dx.doi.org/10.21037/atm-21-5974

Table S1 Top ten up- and down-regulated genes in smokers compared with non-smokers

No.	Gene symbol	Gene function	Log_2FC	P value	FDR
Up					
1	*RPS4Y1*	Ribosomal protein S4, Y-linked 1	4.082734	2.93×10^{-6}	0.007796
2	*DDX3Y*	DEAD-box helicase 3, Y-linked	3.110421	3.97×10^{-7}	0.003271
3	*EIF1AY*	Eukaryotic translation initiation factor 1A, Y-linked	2.297136	2.52×10^{-7}	0.003271
4	*RNU2-2P*	RNA, U2 small nuclear 2, pseudogene	1.62925	0.007246	0.109298
5	*KDM5D*	Lysine demethylase 5D	1.548177	1.18×10^{-5}	0.013452
6	*SNORA73A*	Small nucleolar RNA, H/ACA box 73A	1.520264	0.044346	0.297499
7	*TXLNGY*	Taxilin gamma pseudogene, Y-linked	1.494942	7.38×10^{-7}	0.003785
8	*RN7SKP203*	RNA, 7SK small nuclear pseudogene 203	1.480157	0.012755	0.149845
9	*PSMA6P1*	Proteasome subunit alpha 6 pseudogene 1	1.425072	1.63×10^{-6}	0.006117
10	*PRKY*	Protein kinase, Y-linked, pseudogene	1.40373	1.39×10^{-5}	0.013677
Down					
1	*XIST*	X inactive specific transcript (non-protein coding)	–2.47777	3.51×10^{-6}	0.007927
2	*MIR3687-2*	microRNA 3687-2	–2.33861	0.000831	0.042056
3	*MIR3687-1*	microRNA 3687-1	–2.1256	0.001734	0.055803
4	*MIR3648-1*	microRNA 3648-1	–1.93688	0.007883	0.114649
5	*MIR3648-2*	microRNA 3648-2	–1.78911	0.017309	0.177343
6	*AC018638.1*	Unknown	–1.39712	0.001046	0.046365
7	*PRKCD*	Protein kinase C delta	–1.22466	0.002458	0.064787
8	*ADAMTS7P3*	ADAMTS7 pseudogene 3	–1.13843	0.000429	0.03273
9	*EGFL7*	EGF like domain multiple 7	–1.09507	0.000117	0.021783
10	*TSPO*	Translocator protein 2	–1.03523	0.006953	0.10673

FC, fold change; FDR, false discovery rate.

https://dx.doi.org/10.21037/atm-21-5974

Table S2 Top ten up- and down-regulated genes in COPD patients compared with smokers

No.	Gene symbol	Gene function	Log_2FC	P value	FDR
Up					
1	*MIR3648-2*	microRNA 3648-2	1.327612	8.59×10^{-5}	0.173248
2	*FGG*	Fibrinogen gamma chain	1.155104	0.028642	0.689009
3	*IGHV3-23*	Immunoglobulin heavy variable 3-23	0.935122	0.045358	0.746359
4	*MIR3648-1*	microRNA 3648-1	0.870674	0.000674	0.383676
5	*AC026369.3*	Unknown	0.781978	0.002713	0.560834
6	*FGA*	Fibrinogen alpha chain	0.744763	0.024844	0.681361
7	*ZNF672*	Zinc finger protein 672	0.725588	0.020114	0.665167
8	*AC008738.7*	Unknown	0.711877	0.001453	0.484761
9	*IGKV1D-13*	Immunoglobulin kappa variable 1D-13	0.69582	0.022259	0.671114
10	*PTX3*	Pentraxin 3	0.688336	0.040471	0.733427
Down					
1	*RNU2-2P*	RNA, U2 small nuclear 2, pseudogene	–1.73202	3.44×10^{-5}	0.121012
2	*SCARNA5*	Small Cajal body-specific RNA 5	–1.27381	9.44×10^{-5}	0.177608
3	*AC247036.6*	Immunoglobulin heavy variable 5-10-1	–1.19032	0.038312	0.730368
4	*RNU4-2*	RNA, U4 small nuclear 2	–1.17173	3.58×10^{-5}	0.121012
5	*SLC39A1*	Solute carrier family 39 member 1	–1.1014	0.012167	0.637104
6	*SNORD17*	Small nucleolar RNA, C/D box 17	–1.08461	4.01×10^{-5}	0.121012
7	*RNVU1-7*	RNA, variant U1 small nuclear 7	–1.02956	8.05×10^{-6}	0.121012
8	*SNORD97*	Small nucleolar RNA, C/D box 97	–0.94319	0.010285	0.621664
9	*RNU4-1*	RNA, U4 small nuclear 1	–0.94101	4.72×10^{-5}	0.126794
10	*RNVU1-18*	RNA, variant U1 small nuclear 18	–0.92276	1.90×10^{-5}	0.121012

COPD, chronic obstructive pulmonary disease; FC, fold change; FDR, false discovery rate.

https://dx.doi.org/10.21037/atm-21-5974

Table S3 Top ten up- and down-regulated genes in COPD patients compared with non-smokers

No.	Gene symbol	Gene function	Log_2FC	P value	FDR
Up					
1	*RPS4Y1*	Ribosomal protein S4, Y-linked 1	4.194222	1.78×10^{-6}	0.002284
2	*DDX3Y*	DEAD-box helicase 3, Y-linked	3.143001	4.28×10^{-7}	0.002284
3	*EIF1AY*	Eukaryotic translation initiation factor 1A, Y-linked	2.247875	3.16×10^{-7}	0.002284
4	*FGG*	Fibrinogen gamma chain	2.015582	0.000285	0.012016
5	*CXCL8*	C-X-C motif chemokine ligand 8	1.911682	0.008309	0.080009
6	*SNORA73A*	Small nucleolar RNA, H/ACA box 73A	1.784756	0.008717	0.082324
7	*CXCL1*	C-X-C motif chemokine ligand 1	1.689804	0.012565	0.102645
8	*CYP1B1*	Cytochrome P450 family 1 subfamily B member 1	1.561504	2.92×10^{-7}	0.002284
9	*IGHG4*	Immunoglobulin heavy constant gamma 4 (G4m marker)	1.541758	0.017562	0.126753
10	*CCL2*	C-C motif chemokine ligand 2	1.540555	0.04216	0.217972
Down					
1	*XIST*	X inactive specific transcript (non-protein coding)	–2.48934	3.26×10^{-6}	0.002284
2	*MIR3687-2*	microRNA 3687-2	–2.18868	0.001302	0.026981
3	*MIR3687-1*	microRNA 3687-1	–1.94183	0.003179	0.044999
4	*AC018638.1*	Unknown	–1.57748	0.000181	0.009589
5	*AC073850.1*	Unknown	–1.38813	1.96×10^{-5}	0.003997
6	*SCARNA5*	Small Cajal body-specific RNA 5	–1.3045	0.000138	0.008421
7	*SOSTDC1*	Sclerostin domain containing 1	–1.17014	0.013221	0.105953
8	*ADAMTS7P3*	ADAMTS7 pseudogene 3	–1.16421	2.49×10^{-5}	0.004371
9	*SNX18P12*	Sorting nexin 18 pseudogene 12	–1.14717	0.00166	0.031134
10	*RPL23AP32*	Ribosomal protein L23a pseudogene 32	–1.11901	5.82×10^{-5}	0.005802

COPD, chronic obstructive pulmonary disease; FC, fold change; FDR, false discovery rate.

https://dx.doi.org/10.21037/atm-21-5974