Original research article

Dynamic analysis of gene signatures in the progression of chronic obstructive pulmonary disease

Junchao Jiang, Shengsong Chen, Tao Yu, Chenli Chang, Jixiang Liu, Xiaoxia Ren, Hongtao Niu, Ke Huang, Baicun Li, Chen Wang, Ting Yang

Dynamic analysis of gene signatures in the progression of chronic obstructive pulmonary disease

Junchao Jiang[1,2,3,4], Shengsong Chen[1,2,3,4], Tao Yu[1,2,3,4], Chenli Chang[1,2,3,4], Jixiang Liu[1,2,3,4], Xiaoxia Ren[2,3,4], Hongtao Niu[2,3,4], Ke Huang[2,3,4], Baicun Li[2,3,4], Chen Wang[1,2,3,4,5] and Ting Yang[1,2,3,4.5].

[1] China-Japan Friendship Hospital (Institute of Clinical Medical Sciences), Chinese Academy of Medical Sciences & Peking Union Medical College.

[2] National Center for Respiratory Medicine, Beijing, China.

[3] Institute of Respiratory Medicine, Chinese Academy of Medical Sciences, Beijing, China.

[4] National Clinical Research Center for Respiratory Diseases, Beijing, China.

[5] These authors contributed equally.

*Correspondence author: Chen Wang. No 2, East Yinghua Road, Chaoyang District, Beijing 100029, China.

Tel [+86-01-8420 6276]

Email: cyh-birm@263.net.

Abstract

Purpose: Oxidative stress is an important amplifying mechanism in chronic obstructive pulmonary disease (COPD), however, it is unclear how oxidative stress changes and its exact amplification mechanism in pathological process. We aimed to dynamically analyze the progression of COPD and further elucidate the characteristics of each developmental stage and unveil the underlying mechanisms.

Methods: We performed a holistic analysis by integrating GEO microarray datasets related to smoking, emphysema, and GOLD classification based on the concept of gene, environment and time (GET). Gene Ontology (GO), protein-protein interaction (PPI) networks and gene set enrichment analysis (GSEA) were used to explore the changing characteristics and potential mechanisms. Using lentivirus to promote HIF3A overexpression.

Results: In smokers vs nonsmokers, GO term mainly enriched in "negative regulation of apoptotic process". In later transitions between stages, the main enriched terms were continuous progression of "oxidation-reduction process", and "cellular response to hydrogen peroxide". Logistic regression showed that these core DEGs had diagnostic accuracy in test (AUC=0.823) and validation (AUC=0.750) sets. GSEA and PPI network showed one of the core DEGs, HIF3A, strongly interacted with the ubiquitin-mediated proteolysis pathway. Overexpression of HIF3A restored SOD levels and alleviated the ROS accumulation caused by CSE treatment.

Conclusion: Oxidative stress was continuously intensified from mild emphysema to GOLD-4, thus, special attention should be paid to the identification of emphysema. Furthermore, the downregulated HIF3A may play an important role in the intensified oxidative

stress in COPD.

Key words: COPD, oxidative stress, emphysema, hypoxia, HIF3A

Introduction

Chronic obstructive pulmonary disease (COPD) is a common chronic disease characterized by persistent respiratory symptoms and airflow limitation, which remains a significant global health problem due to its high prevalence and mortality[1]. A Global Burden of Disease (GBD) study reported that the disease resulted in the death of an estimated 3.2 million people worldwide in 2017[2]. In China, the prevalence of COPD among adults aged 20 years and older in China was 8.6%, and the prevalence among adults aged 40 years and older was 13.7%. This indicates that nearly 100 million people have COPD in China[3], thus, improving the prevention and treatment of this disease is critical.

In most patients with COPD, the disease develops over decades and involves etiology, pathology, and the worsening of clinical symptoms [4-6]. Several studies have shown the late diagnosis of COPD is the most important factor underlying poor prognosis among a significant proportion of patients[7-9]. Therefore, it is not sufficient to confine the study of COPD to the stage of clinical symptoms; the focus should be shifted forward to the period of development of pathological lesions, such as emphysema (the typical pathogenic development phase of COPD), and the period involving the risk factor of smoking. A holistic and dynamic analysis of COPD, taking into account of smoking, emphysema and Global Initiative for Chronic Obstructive Lung Disease (GOLD) classification, which reflects the clinical symptoms is needed.

Recently, a new concept has emerged, which integrates gene (G), environment (E) and clinical phenotype information from basic transcriptomics over time (T). GET was a novel approach, by integrating information from basic and clinical omics with exposures over the

lifespan, attempt to uncover novel opportunities for prevention and early treatment of COPD, leading to improved understanding of underlying pathogenic mechanisms and identification of novel targets（10）. There have been many transcriptomics studies comparing COPD patients and healthy individuals(11-13). However, thus far, no study has combined data related to etiology, pathology, and GOLD classification to conduct a holistic and dynamic analysis of COPD. In this study we conducted a GET analysis toward COPD by integrate the transcriptomics microarray data including smoking (the etiology and the environment pollution of COPD), 0emphysema (the clinical phenotype and progression stage of COPD), and the GOLD classification (indicate the worsening of clinical symptoms of COPD), in order to elucidate the characteristics in each of the progression stage and unveiling the mechanisms underlying.

Materials and Methods

Dataset Preparation

Microarray datasets were screened from Gene Expression Omnibus (GEO, http://www.ncbi.nlm.nih.gov/geo). The selection criteria were as follows: 1.Lung tissue samples from nonsmokers, smokers, emphysema and COPD patients were included; 2. Former smokers had quit smoking when participating the program were included; 3. Diffusion Capacity for Carbon Monoxide of the Lung (DLCO) < 80% Ref and/or emphysema shown by CT were included; 4. The classification of COPD should according to the GOLD guidelines; 5. Patients with asthma who had a persistent airflow obstruction were excluded: Based on these criteria(1, 14-16), the GSE37768 (smokers vs nonsmokers), GSE119040 (emphysema vs

non-emphysema patients), GSE1650 (severe vs mild emphysema patients), and GSE69818

(GOLD-4 vs GOLD-1 COPD patients) were obtained. The sample types and platform

information of each dataset are described in Table1.

DEG identification and enrichment analysis

To identify the DEGs, GEO2R (http://www.ncbi.nlm.nih.gov/geo/geo2r/) was used with

each dataset. GEO2R is an interactive online tool that compares two groups of samples

collected under the same experimental conditions[17]. Adjusted P<0.05 and |log2fold

change| ≥0.5 were used as cut-off values. GO analysis was used to investigate the biological

functions of the DEGs, which involved determining the enriched biological process (BP),

cellular component (CC), and molecular function (MF) terms using the functional annotation

tool in the DAVID website[18, 19]. The top 10 terms were visualized using ggplot2 v3.3.2.

Logistic regression prediction model based on the core DEGs

We used GSE37768 data as the training set and GSE8581 data as the test set which

including 18. lung tissues from control subjects (FEV$_1$/FVC＞0.7, FEV$_1$%pred＞80%) and 15

lung tissues from COPD subjects (FEV$_1$/FVC＜0.7, FEV$_1$%pred＜70%). Receiver operating

characteristic (ROC) curves were used to visualize the model, area under the curve (AUC)

was used to evaluate its diagnostic ability, and the Youden index was used to evaluate its

sensitivity and specificity.

Construction of the protein-protein interaction (PPI) networks

A PPI network of each of the four sets of DEGs was constructed using the STRING

website v11.0 (https://www.string-db.org/) and exported to Cytoscape software v3.7.1 for

visualization[20]. Molecular Complex Detection (MCODE) plugin in Cytoscape was used to

identifies clusters that are highly interconnected in a network (21), the top three scored sub-networks were then subjected to GO analysis.

Gene set enrichment analysis (GSEA) of DEGs

GSEA was used to ascertain the enriched Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways for the DEGs identified in each of the four comparisons. The normalized enrichment score (NES) was used to quantify the enrichment magnitude(22, 23).

Participants and ethical approval

Participants, comprising COPD patients and healthy participants (controls) were recruited from the China-Japan Friendship Hospital. COPD patients were diagnosed based on the Global Initiative for Chronic Obstructive Lung Disease (GOLD) clinical criteria. Lung tissues were collected from the COPD patients and controls scheduled for pulmonary lobectomy. The lung function and basic information of the participants are shown in (Table 2).

All experimental work with humans was approved by the Ethics Committee of China-Japan Friendship Hospital. All participants provided written informed consent.

Cell culture and treatment with cigarette smoke extract (CSE)

The murine lung epithelial (MLE) cell line MLE-12 was maintained in DMEM (Gibco, USA) containing 10% fetal bovine serum, 100 U/mL penicillin, and 100 μg/mL streptomycin. CSE was prepared by a modification of a previously published method(24). Briefly, one 3R4F reference cigarette containing 9.4 mg tar and 0.73 mg nicotine (University of Kentucky, USA) was bubbled into 10 mL of high-glucose Dulbecco's Modified Eagle Medium (DMEM; Gibco, USA), and it was subsequently diluted to 5% in use. The MLE-12 cells were then treated with 5% CSE for 24 hours.

Immunohistochemical (IHC) analysis

Lung tissues were fixed in 4% paraformaldehyde solution (Beyotime, China) for 24 h and then dehydrated, embedded in paraffin, and sectioned following routine methods. The sections were incubated with anti-HIF-3α antibody (1:100 dilution, Proteintech, China) overnight at 4°C, followed by incubating with anti-HRP antibody (1:200 dilution, SeraCare, China) at room temperature for 30 min. 3,3'-diaminobenzidine (DAB) was used for color development. The results were scored as the integrated optical density (IOD)/area as detected by Image-Pro Plus.

Protein lysis and western blotting

Equal amounts of protein (20 μg) were separated by 10% sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and transferred onto polyvinylidene fluoride (PVDF) membranes (Immobilon-P membrane, Millipore). After 1 h of blocking with 10% milk, the membranes were incubated with anti-HIF-3α (ABclonal, China) or anti-GAPDH (ABclonal, China) antibody at 4℃ overnight. After incubation with horseradish peroxidase (HRP)-conjugated secondary antibody (ABclonal, China) for 1 h, protein bands were developed in an automatic exposure machine (ProteinSimple, USA).

The reactive oxygen species (ROS) Measurement

MLE-12 cells were incubated with 10 μM DCFH-DA (Solarbio, China) at 37 °C for 30 min. The supernatant was removed, and the cells were washed with DMEM. Images viewed under a fluorescence microscope were obtained.

Superoxide Dismutase Assay

The relative superoxide dismutase (SOD) concentration in cell lysates was assessed

using a superoxide dismutase (SOD) assay kit (Beyotime, China) according to the manufacturer's instructions.

CCK-8 Assay

A CCK-8 assay was performed to detect cell viability according to the manufacturer's protocol (Beyotime, China). Approximately 5×10^3 cells were seeded in 96-well plates. The cells were treated with 10 μL of CCK-8 solution for 24 h and incubated in the dark for another 2 h. The absorbance was measured at a wavelength of 450 nm.

Statistical analysis

All experiments were performed independently at least three times. The data are presented as mean ± standard error of the mean (SEM). Comparisons between the two groups were performed using two-tailed Student's t-tests. $P<0.05$ was considered statistically significant.

Results

Identification of DEGs during the progression of COPD

We identified DEGs (adjusted $P<0.05$ and |log2fold change| ≥0.5) in each of the four datasets. There were 577 upregulated genes and 482 downregulated genes in the GSE37768 dataset (smokers vs nonsmokers), 1832 upregulated genes and 1798 downregulated genes in the GSE119040 dataset (emphysema vs non-emphysema patients), 1757 upregulated genes and 2851 downregulated genes in the GSE1650 dataset (severe vs mild emphysema patients), and 654 upregulated genes and 998 downregulated genes in the GSE69818 dataset (GOLD-4 vs GOLD-1 COPD patients).

GO analyses of DEGs

Regarding the GO analysis of the GSE37768 dataset (smokers vs nonsmokers), the DEGs were mainly enriched in "negative regulation of apoptotic process", "immune response", and "inflammatory response", functioned in "nuclear" and "perinuclear of cytoplasm" through "protein binding" and "receptor activity" (Figure 1a) (Figure S1a). Regarding the GSE119040 dataset (emphysema vs non-emphysema patients), the DEGs were mainly enriched in "negative regulation of the apoptotic process", "innate immune response", and "extracellular matrix organization", functioned in "cytoplasm" through "protein binding" (Figure 1b) (Figure S1b). Regarding the GSE1650 dataset (severe vs mild emphysema patients), the DEGs were mainly enriched in "apoptotic process", "oxidation-reduction process", "protein phosphorylation", and "inflammatory response", functioned in "cytoplasm" through "protein binding" (Figure 1c) (Figure S1c). Lastly, regarding the GSE69818 dataset (GOLD-4 vs GOLD-1 patients), the DEGs were mainly enriched in "immune response", "inflammatory response", and "oxidation-reduction process", functioned in "extracellular exosome" through "protein binding" (Figure 1d) (Figure S1d). The cellular component results show, in the progression of COPD, the location where DEGs functions shift from the nuclear and perinuclear at the beginning, then to the cytoplasm and finally to the extracellular exosome.

GO analyses of PPI networks

To understand the relationships among the four sets of DEGs at the protein level, i.e. the relationships among the differentially expressed proteins (DEPs), and to analyze their functions in detail, we constructed four PPI networks. MCODE was applied for further analysis and the top three densely connected sub-networks in each PPI networks were identified, and

then GO analyses were conducted for each of them (21).

Regarding smoking vs non-smoking, the functions of the three sub-networks were mainly enriched in "positive regulation of endothelial cell proliferation", and "cellular response to zinc, cadmium, metal ions and erythropoietin" (Figure 2a). In the progression to emphysema, the functions of the three sub-networks were mainly enriched in "cytoplasmic translation", "negative regulation of apoptotic process", and "positive regulation of NF-kappaB signaling" (Figure 2b). To further explore the functions of the sub-networks, we analyzed the sub-networks in parts. The first sub-network consisted of parts a and b (Figure 2b.1). The enriched functions of part a included "mRNA splicing, via spliceosome". The enriched functions of part b included "nuclear-transcribed mRNA catabolic process", and "nonsense-mediated decay". Parts a and b interacted through NHP2L1. The second sub-network consisted of parts a, b, and c (Figure 2b.2). The enriched functions of part a included "ephrin receptor signaling pathway". The enriched functions of part b included "protein ubiquitination". The enriched functions of part c included "retrograde vesicle-mediated transport, Golgi to ER". Parts a and c interacted through UBE2N. The third sub-network consisted of parts a and b (Figure 2b.3). The enriched functions of part a included "mitochondrial electron transport", and "ubiquinol to cytochrome c". The enriched functions of part b included "interferon-γ-mediated signaling pathway".

Regarding the progression from mild to severe emphysema, the functions of the three sub-networks were mainly enriched in "extracellular matrix organization", "extracellular matrix disassembly", "inflammatory response", and "negative regulation of endodermal cell differentiation" (Figure 2c). Regarding the progression from GOLD-1 to GOLD-4, the functions

of the three sub-networks were mainly enriched in "inflammatory and immune response", "positive regulation of leukocyte chemotaxis", "fat cell differentiation", and "blood coagulation, fibrin clot formation" (Figure 2d). The second sub-network consisted of parts a and b (Figure 2d.2). The enriched functions of part a included "innate immune response" and "cellular response to hydrogen peroxide". The enriched functions of part b included "adaptive immune response". The third sub-network consisted of parts a and b (Figure 2d.3). The enriched functions of part a included "positive regulation of RNA polymerase II promoter transcription" and "fat cell differentiation". The enriched functions of part b included "extracellular matrix organization" and "positive regulation of ERK1 and ERK2 cascade".

Along with the GO analysis above, the results demonstrated a persistently enhanced oxidation-reduction process, immune and inflammatory response in the development of COPD. Besides, in smoker vs nonsmoker and emphysema vs non-emphysema patients, it showed a negative regulation of apoptotic process. Interestingly, in severe vs mild emphysema patients and GOLD-4 vs GOLD-1 patients, it showed a positive regulation of apoptotic process and intensified immune and inflammatory responses, suggesting that there may be a compensation threshold in the development of COPD, when the damage reach the threshold, the organism exhibits an injury response, which may be related to the lung inflammation persists even after smoking cessation in patients with COPD(1).

Identification of core DEGs

Taken together, the GO and PPI network analyses indicated that continuously immune and inflammatory responses and oxidative stress are exhibited in the development of COPD. We further identified and annotated fourteen DEGs that were consistently either upregulated

or downregulated in the four comparisons (Table S1-4), 6 of which were upregulated (Figure 3a) (C10orf10, MSMB, IGLV1-44, CYP1B1, ZNF385D, and MUC4) and 8 of which were downregulated (Figure 3b) (PPFIBP1, DDX17, FOSB, SVEP1, RABGAP1L, HIF3A, PTCH1, and PTPRD). By annotating these 14 DEGs, we found that they were associated with the comorbidities of COPD, such as anxiety and depression, atherosclerosis, and NSCLC, which would provide evidence for future in-depth studies on the comorbidities of COPD (Table S3, S4).

Logistic regression prediction model based on the core DEGs

A logistic regression prediction model based on the core DEGs was constructed to distinguish between COPD patients and non-COPD patients, using GSE37768 data as the training set and GSE8581 data as the test set. The AUC for the training set was 0.823 (Figure 3c), while the sensitivity and specificity were 71.7% and 80%, respectively. The AUC for the test set was 0.750 (Figure 3d), while the sensitivity and specificity were 74.1% and 60.2% respectively. The results above demonstrated the core DEGs were not only meaningful in biological analysis, but also have efficacy in distinguishing COPD patients from non-COPD patients in clinical practice.

GSEA of DEGs

We performed a GSEA of the four sets of DEGs in order to identify core pathways that (1) were consistently upregulated or downregulated in all four comparisons and (2) had a NES $\geqslant$ 1 in at least three of the comparisons. We identified 5 core pathways; the following 2 were continuously upregulated: "alanine aspartate and glutamate metabolism", and "pathogenic escherichia coli infection", while the following 3 were continuously downregulated:

"phosphatidylinositol signaling system", "oocyte meiosis", and "ubiquitin mediated proteolysis" (Figure 4b) (Figure S2).

PPI network of core DEPs and core pathways

To investigate the interactions between the 5 core pathways and 14 core DEPs, we constructed a PPI network (the flow of the construction was shown as Figure S3). The results indicated that 42 proteins out of a total of 394 proteins in the 5 pathways interacted with 11 of the core DEPs (Figure 4a). Among the 42 proteins, there were 19 proteins belonged to the ubiquitin mediated proteolysis pathway, followed by 11, 6, 5, and 1 belonging to the oocyte meiosis, pathogenic escherichia coli infection, phosphatidylinositol signaling system, and alanine aspartate and glutamate metabolism pathways, respectively (Table S5). Among the 14 core DEGs, HIF3A exhibited specialized interactions with 6 proteins that all belonged to the ubiquitin mediated proteolysis pathway, which further suggesting the interaction between HIF3A and ubiquitin mediated proteolysis pathway may play a key role in the progression of COPD.

Validation of HIF-3α in vitro and in vivo.

To validate our findings, we evaluated HIF-3α expression by IHC analysis and western blotting of lung tissue from COPD patients and healthy participants. IHC showed that HIF-3α protein expression (concentrated in alveolar epithelial cells) was decreased in COPD (Figures 5a and 5b). Western blotting confirmed the HIF-3α protein downregulation in COPD (Figures 5c and 5d). We also established a COPD cell model, and western blotting confirmed that HIF-3α protein expression was downregulated in CSE-exposed MLE-12 cells (Figures 5e and 5f). The results above indicate the down-regulation of HIF-3α expression in COPD tissues in vivo

and CSE-exposed MLE-12 cells in vitro.

The HIF3A overexpression mitigates CSE induced ROS accumulation.

Oxidative stress arises as a result of endogenous antioxidant defenses being genetically impaired and/or overwhelmed by the presence of ROS(25). SOD plays an important role in antioxidant activity and prevent ROS-initiated reactions(26). To further explore the interactions between HIF3A and oxidative stress, a lentivirus vector was used to promote HIF3A overexpression in CSE treated MLE-12 cells. Overexpression of HIF3A (Figure 6 a-b) decreased ROS production (Figure 6 e) and restored SOD level (Figure 6 c). Besides, the low cell viability (Figure 6 d) caused by CSE were alleviated after the HIF3A overexpression. The results above suggest that the HIF3A overexpression mitigates oxidative stress and further enhanced cell viability.

Discussion

As late diagnosis of COPD is a major factor in the poor prognosis of many patients(7-9), therefore, research on COPD should shifted to the period of pathological changes rather than the period of clinical symptoms. In this study, we conducted a GET analysis and divided COPD development into four developmental stages based on etiology, pathology, and the worsening of clinical symptoms (exposure to smoking, from non-emphysema to emphysema, from mild to severe emphysema, and from GOLD-1 to GOLD-4) and analyzed GEO microarray datasets related to these stages, conducted a holistic analysis by integrating these four datasets to uncover the gene function characteristics in each of the progression stage

and unveiling the mechanisms underlying.

GET analysis is a novel approach, and it may lead to improved understanding of underlying pathogenic mechanisms and identification of novel targets(10). By conducting a GET analysis in this study, we revealed that, during the progression of COPD, immune and inflammatory response was continuously enriched, accompanied by intensified oxidation-reduction process, positive regulation of apoptotic and leukocyte chemotaxis, and cellular response to hydrogen peroxide. These functions were especially intensified from mild to severe emphysema and from GOLD-1 to GOLD-4. This suggests that patients exhibit persistently increased inflammation and oxidative stress after the emphysema stage develops, which further highlights, in clinical practice, special attention should be paid to identifying emphysema and to monitoring and controlling oxidative stress, which provides more precise insights into the GOLD Report.

Given the sustained progression of oxidative stress and sustained immune and inflammatory responses, we identified 14 core DEGs that were continuously upregulated or downregulated in all four comparisons. Of these, 6 were continuously upregulated (C10orf10, MSMB, IGLV1-44, CYP1B1, ZNF385D, and MUC4) and 8 were continuously downregulated (PPPFIBP1, DDX17, FOSB, SVEP1, RABGAP1L, HIF3A, PTCH1, and PTPRD). These core DEGs were associated with the comorbidities of COPD such as anxiety and depression, atherosclerosis, and NSCLC (Table S3, S4). The results above further support the idea that COPD is a systemic syndrome rather than a uniform disease entity(27, 28).

In addition, to identify the potential pathways involved in the development of COPD, we performed GSEA to each of the four dataset, and identified 5 core pathways. Two of them

were continuously upregulated ("alanine aspartate and glutamate metabolism", "pathogenic escherichia coli infection"), and three of them were continuously downregulated ("oocyte meiosis", "phosphatidylinositol signaling system", and "ubiquitin mediated proteolysis").

To elucidate the potential mechanisms underlying the development of COPD, the proteins encoded by the core DEGs and the enriched core pathways were used to construct a PPI network. Among the core DEPs, HIF-3α attracted our attention, as it exhibited specialized interactions with 6 proteins that all belonged to the ubiquitin mediated proteolysis pathway. HIF3A is a member of the hypoxia-inducible transcription factor (HIF) family, which are master regulators of the adaptive cell response to decreased oxygen levels, controlling the expression of many genes in an oxygen-dependent manner, the product of these genes were involved in hematopoiesis, angiogenesis, iron transport, oxidation stress, and extracellular matrix synthesis(29). There are three members in this family; HIF1A, HIF2A (EPAS1) and HIF3A. Among them, HIF3A is the most recently identified member. Compared with HIF-1α and HIF-2α, HIF-3α has dual functions: Inhibition of the activities of HIF-1α and HIF-2α, and regulation of its own target genes in a response to hypoxia(30, 31). It can upregulate genes involved in glucose and amino acid metabolism, apoptosis, proteolysis, p53 signaling, PPAR signaling, Jak-STAT signaling and NOD-like receptor signaling(31). In addition, it can inhibit the production of ROS and reduce oxidative stress levels(31-34). Therefore, decreased HIF3A may be a major factor that causes and aggravates the oxidative stress in COPD.

In the PPI analysis of core DEPs and core pathways, the interaction between HIF3A and von Hippel-Lindau (VHL) attracted our attention (Figure 5). Under normal oxygen partial pressure, HIF-3α can be degraded by the pVHL ubiquitin-proteasome due to the shared

common oxygen-dependent degradation (ODD) domain in HIF-3α. The ability of VHL to degrade HIF-3α is dependent on the proline 490 residue of HIF-3α and this is increased in the presence of prolyl hydroxylase (PHD), a cellular sensor for low oxygen(35). PHD catalyzes the hydroxylation of key amino acid residues in the HIF-α ODD domain(36, 37). This is followed by VHL binding to HIF-α and inducing degradation via the ubiquitin-proteasome pathway(38, 39). A previous study showed that HIF-3α expression was elevated during acute hypoxia and decreased slowly after 14 days of hypoxia, accompanied by sustained HIF-1α and HIF-2α upregulation(40, 41). Combined with these results above, the transient elevated HIF-3α may be due to the inhibited PHD and VHL under hypoxia, and in chronic hypoxia (which is often occurred in COPD), HIF-3α downregulation may be due to the competition with HIF-1α and HIF-2α in binding to the HIF-1β subunits（42）. In vitro and in vivo experiments showed that, under hypoxia, HIF-1α and HIF-2α upregulation can lead to ROS accumulation（32）, HIF1A and HIF2A knockdown reduces the oxidative stress(33, 34). According to the GOLD-2022 report(1), "oxidative stress may be an important amplifying mechanism in COPD". Oxidative stress is an imbalance between oxidants and antioxidants in favor of the oxidants, leading to a disruption of redox signaling and control and/or molecular damage(43), and as HIF3A has transcriptional activation ability under hypoxia and is a negative regulator of HIF1A and HIF2A(31, 42), the downregulated HIF3A in our study may be one of the reasons for the persistent elevated oxidative stress level and the amplifying mechanism in COPD.

　　Another of the core DEPs, the PTCH1 (patched 1) protein, also attracted our attention. According to the results of a GWAS on COPD, PTCH1 showed the strongest positive

associations with FEV_1/FVC(44). PTCH1 encodes a member of the patched family of proteins and a component of the hedgehog signaling pathway. Similarly, the 2022 GOLD report also indicated that HHIP, which is a part of the same hedgehog signaling pathway as PTCH1, is highly correlated with COPD phenotype, but the exact pathogenic mechanism remains unclear, suggesting that our analytical approach can indeed identify key factors in the pathogenesis of COPD.

In this study, we applied GETomics approach to conduct a holistic analysis toward COPD, identified oxidative stress was especially intensified from mild to severe emphysema and from GOLD-1 to GOLD-4, which provides more precise insights into the GOLD Report and clinical practice. According to the continuously intensified oxidative stress, immune and inflammatory response, we identified and annotated 14 core DEGs that were consistently either upregulated or downregulated in the four comparisons, and found these core DEGs were related to the comorbidities of COPD. By constructing and validating the prediction model, we discovered these 14 core DEGs were not only meaningful in biological analysis, but also have efficacy in differentiating COPD patients from healthy individuals. Through further GSEA and PPI analysis, we found the highly interacted HIF3A and ubiquitin mediated proteolysis pathway may contribute to the oxidative stress in COPD pathogenesis. Finally, a lentivirus was used to promote the HIF3A overexpression, the results suggest overexpression of HIF3A mitigates oxidative stress and further enhanced cell viability in CSE treated MLE-12 cells. There were two earlier studies have reported the HIF3A in COPD, one of the study demonstrated the HIF3A mRNA was downregulated in the tibialis anterior muscle of severe COPD patients,

which was consistence with our study(45). However, another study reported HIF-3α was upregulated in the lung of mice for 15 weeks smoking(46). As the HIF3A is the most recently identified member of the HIF family, the specific regulatory mechanisms of HIF3A still need further research. This current study still has some limitations. We only assessed HIF3A expression in lung and MLE-12 cells, so its expression in serum and sputum remains unknown. Serum and sputum are more easily assessed and therefore more widely used in clinical practice to assess biomarkers, which we intend to carry out in the future.

Conclusion

In conclusion, the results of the GET analysis strongly indicate that special attention should be paid by clinicians to identifying of emphysema and monitoring and controlling COPD patients' oxidative stress level. Besides, the 14 core DEGs have certain efficacy in differentiating COPD patients form healthy individuals, and may related to the comorbidities. Furthermore, the overexpression of HIF3A can alleviate the oxidative stress in COPD.

Declarations

Data availability statement

Publicly available datasets were analyzed in this study. These data can be found here: GEO database, accession number: GSE37768, GSE119040, GSE1650, GSE69818, GSE8581.

Ethics approval and consent to participate

This study was approved by the Ethics Committee of China-Japan Friendship Hospital.

Author contributions

JJ, SC, TY participated and conceived the study design. JJ, TY and CC collected the data. JJ

performed the experiments, analyzed the data, and wrote the manuscript. JJ, JL, XR, CW, and YT interpreted and discussed the data. CW, TY, KH, HN, BL refined the final draft and revised the manuscript. All the authors reviewed the final version of the manuscript.

Competing interests

Conflict of interest: Junchao Jiang has nothing to disclose.

Conflict of interest: Shengsong Chen has nothing to disclose.

Conflict of interest: Tao Yu has nothing to disclose.

Conflict of interest: Chenli Chang has nothing to disclose.

Conflict of interest: Jixiang Liu has nothing to disclose.

Conflict of interest: Xiaoxia Ren has nothing to disclose.

Conflict of interest: Hongtao Niu has nothing to disclose.

Conflict of interest: Ke Huang has nothing to disclose.

Conflict of interest: Baicun Li has nothing to disclose.

Conflict of interest: Ting Yang has nothing to disclose.

Conflict of interest: Chen Wang has nothing to disclose.

References:

1. 2022 GOLD Reports - Global Initiative for Chronic Obstructive Lung Disease - GOLD.

2. Global, regional, and national age-sex-specific mortality for 282 causes of death in 195 countries and territories, 1980-2017: a systematic analysis for the Global Burden of Disease Study 2017. LANCET. [Journal Article; Research Support, N.I.H., Extramural; Research Support, Non-U.S. Gov't; Research Support, U.S. Gov't, Non-P.H.S.]. 2018 2018-11-10;392(10159):1736-88.

3. Wang C, Xu J, Yang L, Xu Y, Zhang X, Bai C, Kang J, Ran P, Shen H, Wen F, Huang K, Yao W, Sun T, Shan G, Yang T, Lin Y, Wu S, Zhu J, Wang R, Shi Z, Zhao J, Ye X, Song Y, Wang Q, Zhou Y, Ding L, Yang T, Chen Y, Guo Y, Xiao F, Lu Y, Peng X, Zhang B, Xiao D, Chen CS, Wang Z, Zhang H, Bu X, Zhang X, An L, Zhang S, Cao Z, Zhan Q, Yang Y, Cao B, Dai H, Liang L, He J. Prevalence and risk factors of chronic obstructive pulmonary disease in China (the China Pulmonary Health [CPH] study): a national cross-sectional study. LANCET. [Journal Article; Research Support, Non-U.S. Gov't]. 2018 2018-04-28;391(10131):1706-17.

4. Hogg JC, Timens W. The pathology of chronic obstructive pulmonary disease. Annu Rev Pathol. [Journal Article; Research Support, N.I.H., Extramural; Research Support, Non-U.S. Gov't; Review]. 2009 2009-01-20;4:435-59.

5. Rennard SI, Drummond MB. Early chronic obstructive pulmonary disease: definition, assessment, and prevention. LANCET. [Journal Article; Review]. 2015 2015-05-02;385(9979):1778-88.

6. Hu WP, Zeng YY, Zuo YH, Zhang J. Identification of novel candidate genes involved in the progression of emphysema by bioinformatic methods. Int J Chron Obstruct Pulmon Dis. [Comparative Study; Journal Article]. 2018 2018-01-20;13:3733-47.

7. Lange P, Ahmed E, Lahmar ZM, Martinez FJ, Bourdin A. Natural history and mechanisms of COPD. RESPIROLOGY. [Journal Article; Review]. 2021 2021-04-01;26(4):298-321.

8. Martinez FJ, Han MK, Allinson JP, Barr RG, Boucher RC, Calverley P, Celli BR, Christenson SA, Crystal RG, Fageras M, Freeman CM, Groenke L, Hoffman EA, Kesimer M, Kostikas K, Paine RR, Rafii S, Rennard SI, Segal LN, Shaykhiev R, Stevenson C, Tal-Singer R, Vestbo J, Woodruff PG, Curtis JL, Wedzicha JA. At the Root: Defining and Halting Progression of Early Chronic Obstructive Pulmonary Disease. Am J Respir Crit Care Med. [Comparative Study; Journal Article]. 2018 2018-06-15;197(12):1540-51.

9. Jones RC, Price D, Ryan D, Sims EJ, von Ziegenweidt J, Mascarenhas L, Burden A, Halpin DM, Winter R, Hill S, Kearney M, Holton K, Moger A, Freeman D, Chisholm A, Bateman ED. Opportunities to diagnose chronic obstructive pulmonary disease in routine care in the UK: a retrospective study of a clinical cohort. Lancet Respir Med. [Journal Article; Research Support, Non-U.S. Gov't]. 2014 2014-04-01;2(4):267-76.

10. Agusti A, Melen E, DeMeo DL, Breyer-Kohansal R, Faner R. Pathogenesis of chronic obstructive pulmonary disease: understanding the contributions of gene-environment interactions across the lifespan. Lancet Respir Med. [Journal Article; Review; Research Support, N.I.H., Extramural; Research Support, Non-U.S. Gov't]. 2022 2022-05-01;10(5):512-24.

11. Yang D, Yan Y, Hu F, Wang T. CYP1B1, VEGFA, BCL2, and CDKN1A Affect the Development of Chronic Obstructive Pulmonary Disease. Int J Chron Obstruct Pulmon Dis. [Journal Article]. 2020 2020-01-20;15:167-75.

12. Duan R, Niu H, Yu T, Cui H, Yang T, Hao K, Wang C. Identification and Bioinformatic Analysis

of Circular RNA Expression in Peripheral Blood Mononuclear Cells from Patients with Chronic Obstructive Pulmonary Disease. Int J Chron Obstruct Pulmon Dis. [Journal Article]. 2020 2020-01-20;15:1391-401.

13.   Sun X, Shang J, Wu A, Xia J, Xu F. Identification of dynamic signatures associated with smoking-related squamous cell lung cancer and chronic obstructive pulmonary disease. J CELL MOL MED. [Journal Article; Research Support, Non-U.S. Gov't]. 2020 2020-01-01;24(2):1614-25.

14.   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. [Journal Article; Research Support, Non-U.S. Gov't]. 2019 2019-10-01;23(10):6989-99.

15.   Peinado VI, Gomez FP, Barbera JA, Roman A, Angels MM, Ramirez J, Roca J, Rodriguez-Roisin R. Pulmonary vascular abnormalities in chronic obstructive pulmonary disease undergoing lung transplant. J Heart Lung Transplant. [Journal Article; Research Support, Non-U.S. Gov't]. 2013 2013-12-01;32(12):1262-9.

16.   Faner R, Cruz T, Casserras T, Lopez-Giraldo A, Noell G, Coca I, Tal-Singer R, Miller B, Rodriguez-Roisin R, Spira A, Kalko SG, Agusti A. Network Analysis of Lung Transcriptomics Reveals a Distinct B-Cell Signature in Emphysema. Am J Respir Crit Care Med. [Journal Article; Research Support, Non-U.S. Gov't]. 2016 2016-06-01;193(11):1242-53.

17.   Barrett T, Wilhite SE, Ledoux P, Evangelista C, Kim IF, Tomashevsky M, Marshall KA, Phillippy KH, Sherman PM, Holko M, Yefanov A, Lee H, Zhang N, Robertson CL, Serova N, Davis S, Soboleva A. NCBI GEO: archive for functional genomics data sets--update. NUCLEIC ACIDS RES. [Journal Article; Research Support, N.I.H., Intramural]. 2013 2013-01-01;41(Database issue):D991-5.

18.   Sherman BT, Hao M, Qiu J, Jiao X, Baseler MW, Lane HC, Imamichi T, Chang W. DAVID: a web server for functional enrichment analysis and functional annotation of gene lists (2021 update). NUCLEIC ACIDS RES. [Journal Article]. 2022 2022-03-23.

19.   Huang DW, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. NAT PROTOC. [Journal Article; Research Support, N.I.H., Extramural]. 2009 2009-01-20;4(1):44-57.

20.   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. [Journal Article; Research Support, Non-U.S. Gov't; Research Support, U.S. Gov't, Non-P.H.S.; Research Support, U.S. Gov't, P.H.S.]. 2003 2003-11-01;13(11):2498-504.

21.   Bader GD, Hogue CW. An automated method for finding molecular complexes in large protein interaction   networks. BMC BIOINFORMATICS. [Evaluation Study; Journal Article; Research Support, Non-U.S. Gov't]. 2003 2003-01-13;4:2.

22.   Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, Mesirov JP. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A. [Journal Article]. 2005 2005-10-25;102(43):15545-50.

23.   Mootha VK, Lindgren CM, Eriksson KF, Subramanian A, Sihag S, Lehar J, Puigserver P, Carlsson E, Ridderstrale M, Laurila E, Houstis N, Daly MJ, Patterson N, Mesirov JP, Golub TR, Tamayo P, Spiegelman B, Lander ES, Hirschhorn JN, Altshuler D, Groop LC. PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. NAT GENET. [Journal Article; Research Support, Non-U.S. Gov't]. 2003 2003-07-01;34(3):267-73.

24. Carp H, Janoff A. Possible mechanisms of emphysema in smokers. In vitro suppression of serum elastase-inhibitory capacity by fresh cigarette smoke and its prevention by antioxidants. Am Rev Respir Dis. [Journal Article; Research Support, U.S. Gov't, P.H.S.]. 1978 1978-09-01;118(3):617-21.

25. Kirkham PA, Barnes PJ. Oxidative stress in COPD. CHEST. [Journal Article; Research Support, Non-U.S. Gov't; Review]. 2013 2013-07-01;144(1):266-73.

26. Barrera G. Oxidative stress and lipid peroxidation products in cancer progression and therapy. ISRN Oncol. [Journal Article]. 2012 2012-01-20;2012:137289.

27. Koo HK, Vasilescu DM, Booth S, Hsieh A, Katsamenis OL, Fishbane N, Elliott WM, Kirby M, Lackie P, Sinclair I, Warner JA, Cooper JD, Coxson HO, Pare PD, Hogg JC, Hackett TL. Small airways disease in mild and moderate chronic obstructive pulmonary disease: a cross-sectional study. Lancet Respir Med. [Journal Article; Research Support, Non-U.S. Gov't]. 2018 2018-08-01;6(8):591-602.

28. Janssen R, Piscaer I, Franssen F, Wouters E. Emphysema: looking beyond alpha-1 antitrypsin deficiency. Expert Rev Respir Med. [Journal Article; Review]. 2019 2019-04-01;13(4):381-97.

29. Heikkila M, Pasanen A, Kivirikko KI, Myllyharju J. Roles of the human hypoxia-inducible factor (HIF)-3alpha variants in the hypoxia response. CELL MOL LIFE SCI. [Journal Article; Research Support, Non-U.S. Gov't]. 2011 2011-12-01;68(23):3885-901.

30. Tolonen JP, Heikkila M, Malinen M, Lee HM, Palvimo JJ, Wei GH, Myllyharju J. A long hypoxia-inducible factor 3 isoform 2 is a transcription activator that regulates erythropoietin. CELL MOL LIFE SCI. [Journal Article]. 2020 2020-09-01;77(18):3627-42.

31. Zhang P, Yao Q, Lu L, Li Y, Chen PJ, Duan C. Hypoxia-inducible factor 3 is an oxygen-dependent transcription activator and regulates a distinct transcriptional response to hypoxia. CELL REP. [Journal Article; Research Support, Non-U.S. Gov't; Research Support, U.S. Gov't, Non-P.H.S.]. 2014 2014-03-27;6(6):1110-21.

32. Zou Y, Palte MJ, Deik AA, Li H, Eaton JK, Wang W, Tseng YY, Deasy R, Kost-Alimova M, Dancik V, Leshchiner ES, Viswanathan VS, Signoretti S, Choueiri TK, Boehm JS, Wagner BK, Doench JG, Clish CB, Clemons PA, Schreiber SL. A GPX4-dependent cancer cell state underlies the clear-cell morphology and confers sensitivity to ferroptosis. NAT COMMUN. [Journal Article; Research Support, N.I.H., Extramural; Research Support, Non-U.S. Gov't]. 2019 2019-04-08;10(1):1617.

33. Jiang Y, Mao C, Yang R, Yan B, Shi Y, Liu X, Lai W, Liu Y, Wang X, Xiao D, Zhou H, Cheng Y, Yu F, Cao Y, Liu S, Yan Q, Tao Y. EGLN1/c-Myc Induced Lymphoid-Specific Helicase Inhibits Ferroptosis through Lipid Metabolic Gene Expression Changes. THERANOSTICS. [Journal Article]. 2017 2017-01-20;7(13):3293-305.

34. Singhal R, Mitta SR, Das NK, Kerk SA, Sajjakulnukit P, Solanki S, Andren A, Kumar R, Olive KP, Banerjee R, Lyssiotis CA, Shah YM. HIF-2alpha activation potentiates oxidative cell death in colorectal cancers by increasing cellular iron. J CLIN INVEST. [Journal Article; Research Support, N.I.H., Extramural; Research Support, Non-U.S. Gov't; Research Support, U.S. Gov't, Non-P.H.S.]. 2021 2021-06-15;131(12):e143691.

35. Huang LE, Arany Z, Livingston DM, Bunn HF. Activation of hypoxia-inducible transcription factor depends primarily upon redox-sensitive stabilization of its alpha subunit. J BIOL CHEM. [Journal Article; Research Support, Non-U.S. Gov't; Research Support, U.S. Gov't, P.H.S.]. 1996 1996-12-13;271(50):32253-9.

36. Pientka FK, Hu J, Schindler SG, Brix B, Thiel A, Johren O, Fandrey J, Berchner-Pfannschmidt U,

Depping R. Oxygen sensing by the prolyl-4-hydroxylase PHD2 within the nuclear compartment and the influence of compartmentalisation on HIF-1 signalling. J CELL SCI. [Journal Article; Research Support, Non-U.S. Gov't]. 2012 2012-11-01;125(Pt 21):5168-76.

37. Niecknig H, Tug S, Reyes BD, Kirsch M, Fandrey J, Berchner-Pfannschmidt U. Role of reactive oxygen species in the regulation of HIF-1 by prolyl hydroxylase 2 under mild hypoxia. Free Radic Res. [Journal Article; Research Support, Non-U.S. Gov't]. 2012 2012-06-01;46(6):705-17.

38. Groulx I, Lee S. Oxygen-dependent ubiquitination and degradation of hypoxia-inducible factor requires nuclear-cytoplasmic trafficking of the von Hippel-Lindau tumor suppressor protein. MOL CELL BIOL. [Journal Article; Research Support, Non-U.S. Gov't]. 2002 2002-08-01;22(15):5319-36.

39. Ivan M, Kondo K, Yang H, Kim W, Valiando J, Ohh M, Salic A, Asara JM, Lane WS, Kaelin WJ. HIFalpha targeted for VHL-mediated destruction by proline hydroxylation: implications for O2 sensing. SCIENCE. [Journal Article; Research Support, Non-U.S. Gov't; Research Support, U.S. Gov't, P.H.S.]. 2001 2001-04-20;292(5516):464-8.

40. Chen YR, Dai AG, Hu RC, Jiang YL. Differential and reciprocal regulation between hypoxia-inducible factor-alpha subunits and their prolyl hydroxylases in pulmonary arteries of rat with hypoxia-induced hypertension. Acta Biochim Biophys Sin (Shanghai). [Journal Article; Research Support, Non-U.S. Gov't]. 2006 2006-06-01;38(6):423-34.

41. Li QF, Wang XR, Yang YW, Lin H. Hypoxia upregulates hypoxia inducible factor (HIF)-3alpha expression in lung epithelial cells: characterization and comparison with HIF-1alpha. CELL RES. [Comparative Study; Journal Article]. 2006 2006-06-01;16(6):548-58.

42. Yang SL, Wu C, Xiong ZF, Fang X. Progress on hypoxia-inducible factor-3: Its structure, gene regulation and biological function (Review). MOL MED REP. [Journal Article; Review]. 2015 2015-08-01;12(2):2411-6.

43. Sies H, Berndt C, Jones DP. Oxidative Stress. ANNU REV BIOCHEM. [Journal Article; Review; Research Support, N.I.H., Extramural; Research Support, Non-U.S. Gov't]. 2017 2017-06-20;86:715-48.

44. Obeidat M, Hao K, Bosse Y, Nickle DC, Nie Y, Postma DS, Laviolette M, Sandford AJ, Daley DD, Hogg JC, Elliott WM, Fishbane N, Timens W, Hysi PG, Kaprio J, Wilson JF, Hui J, Rawal R, Schulz H, Stubbe B, Hayward C, Polasek O, Jarvelin MR, Zhao JH, Jarvis D, Kahonen M, Franceschini N, North KE, Loth DW, Brusselle GG, Smith AV, Gudnason V, Bartz TM, Wilk JB, O'Connor GT, Cassano PA, Tang W, Wain LV, Soler AM, Gharib SA, Strachan DP, Sin DD, Tobin MD, London SJ, Hall IP, Pare PD. Molecular mechanisms underlying variations in lung function: a systems genetics analysis. Lancet Respir Med. [Journal Article; Research Support, N.I.H., Extramural; Research Support, N.I.H., Intramural; Research Support, Non-U.S. Gov't]. 2015 2015-10-01;3(10):782-95.

45. Jatta K, Eliason G, Portela-Gomes GM, Grimelius L, Caro O, Nilholm L, Sirjso A, Piehl-Aulin K, Abdel-Halim SM. Overexpression of von Hippel-Lindau protein in skeletal muscles of patients with chronic obstructive pulmonary disease. J CLIN PATHOL. [Journal Article; Research Support, Non-U.S. Gov't]. 2009 2009-01-01;62(1):70-6.

46. Gu W, Wang L, Deng G, Gu X, Tang Z, Li S, Jin W, Yang J, Guo X, Li Q. Knockdown of long noncoding RNA MIAT attenuates cigarette smoke-induced airway remodeling by downregulating miR-29c-3p-HIF3A axis. TOXICOL LETT. [Journal Article]. 2022 2022-03-01;357:11-9.
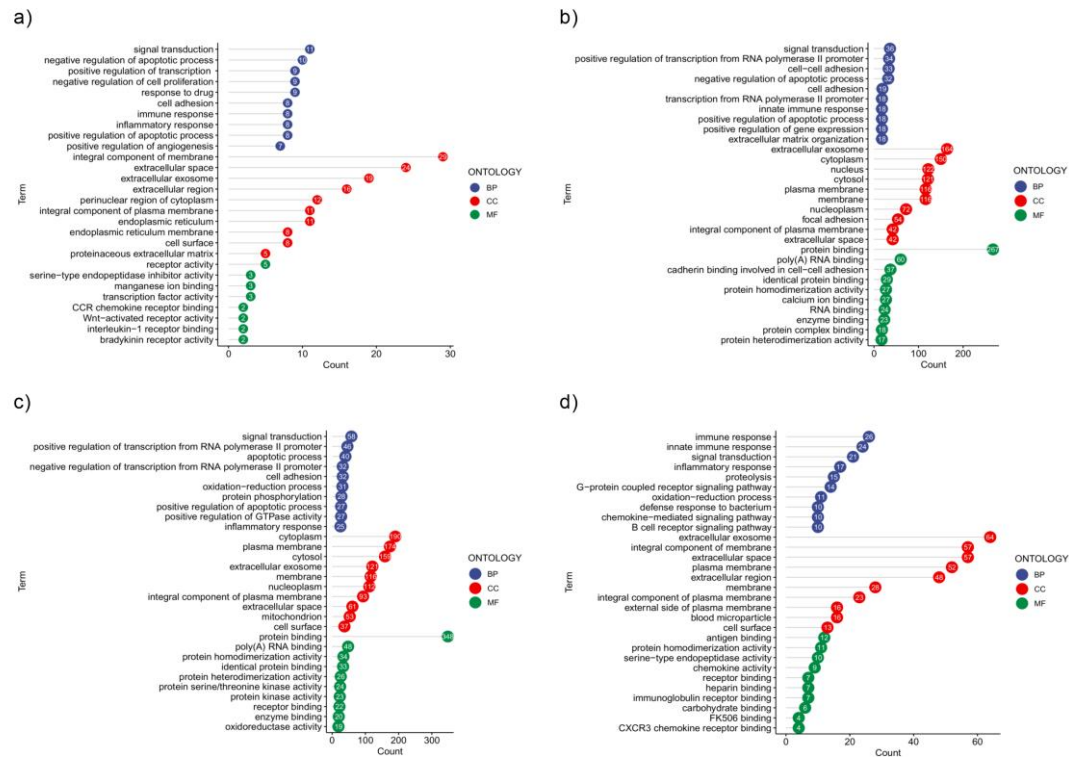
Figure 1. Gene Ontology (GO) enrichment analyses for upregulated genes. GO enrichment analyses for a) smokers vs nonsmokers ; b) emphysema vs non emphysema patients; c) severe vs mild emphysema patients; and d) GOLD-4 vs GOLD-1 patients.
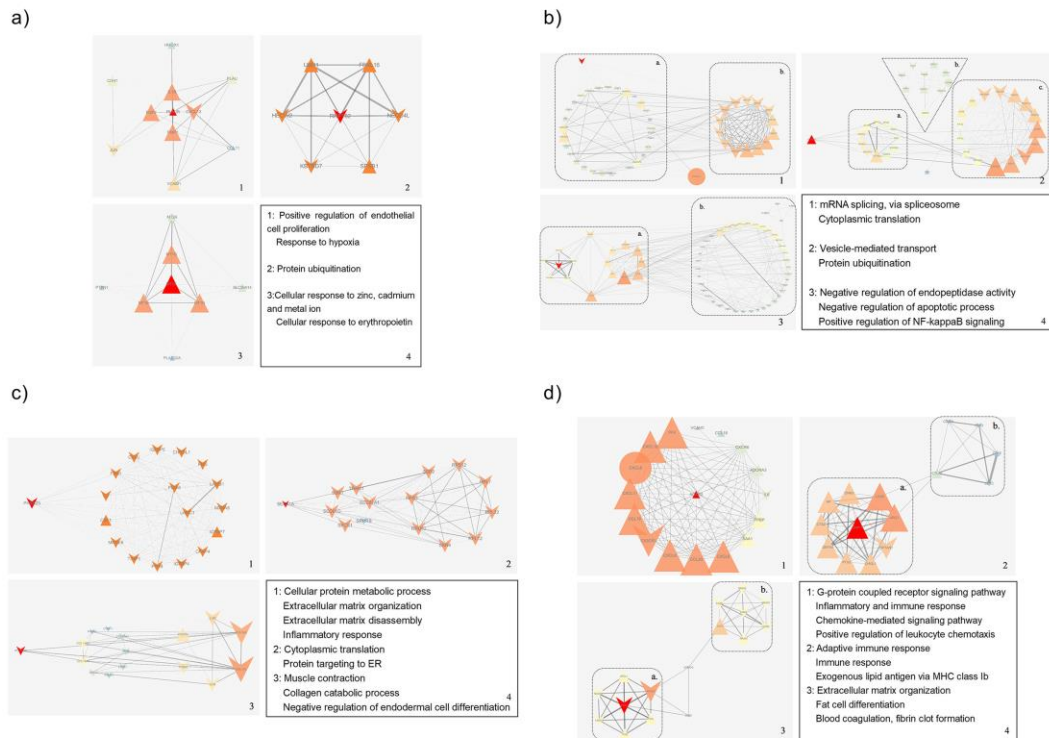
Figure 2. Protein–protein interaction (PPI) network of differentially expressed proteins (DEPs). Graphs 1, 2, and 3 shows the top three sub-networks of DEGs and graph 4 shows the enriched Gene Ontology (GO) biological process (BP) terms for a) smoking vs non-smoking, b) emphysema vs non-emphysema patients, c) severe vs mild emphysema patients, and d) GOLD-4 vs GOLD-1 COPD patients. Triangles represent upregulated DEPs, "V" shapes represent downregulated DEPs, and circles represent proteins produced by the string database, have no state in our data. Larger node size and a more orange color indicate higher degree value. Thicker edges represent stronger interactions (higher combined_score of DEPs).
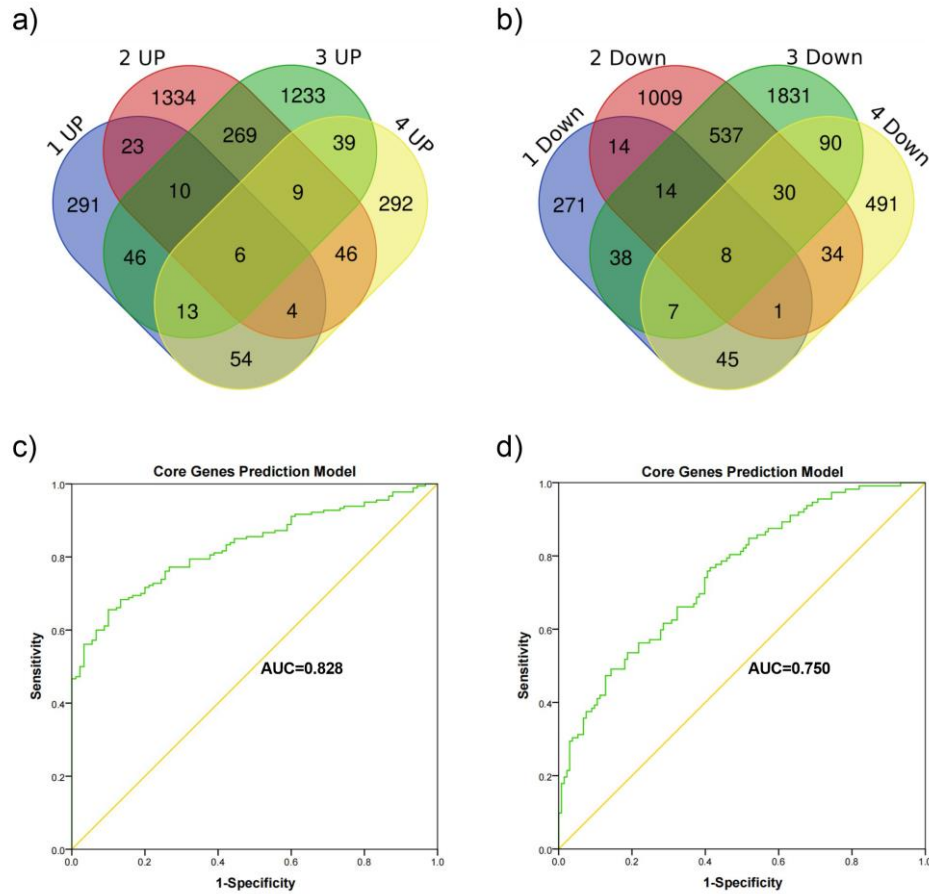
Figure 3. Venn and Logistic regression analysis. Venn diagrams showing a) 6 consistently upregulated DEGs and b) 8 consistently downregulated DEGs. ROC curve analysis of the Core Genes prediction model. c) Construction of the logistic prediction model. d) Validation of the logistic prediction model.
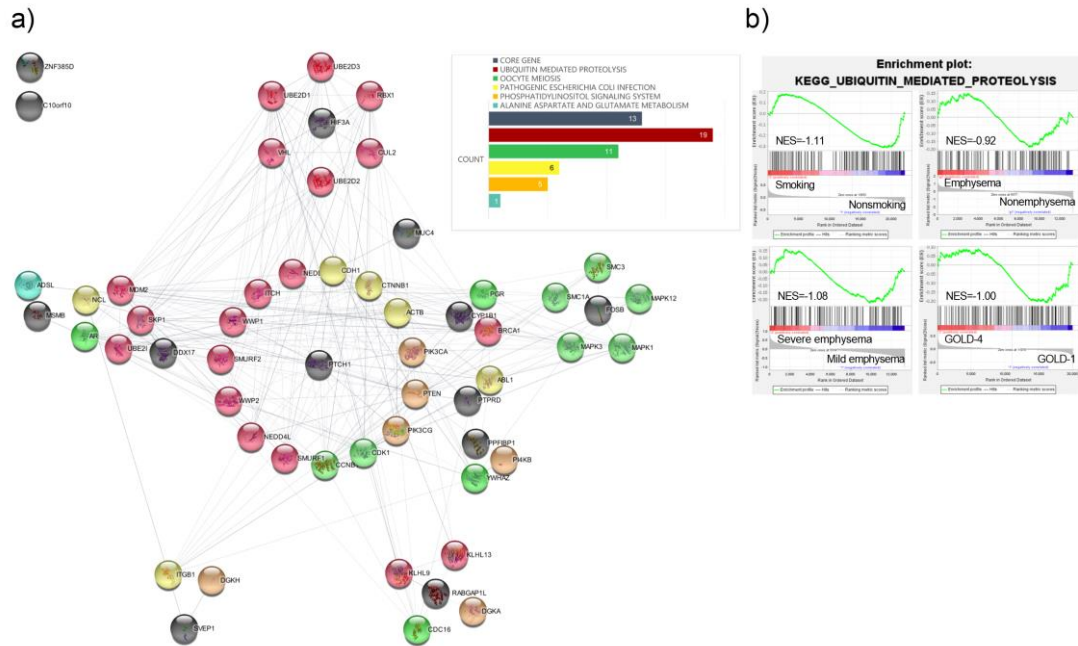
Figure 4. PPI network of consistently upregulated or downregulated DEPs and consistently enriched pathways. a) Black solid circles indicate that the protein is encoded by a core DEG. Red, green, yellow, orange, and blue solid circles indicate that the protein belongs to the UBIQUITIN MEDIATED PROTEOLYSIS pathway (continuously downregulated), OOCYTE MEIOSIS pathway (continuously downregulated), PATHOGENIC ESCHERICHIA COLI INFECTION pathway (continuously upregulated), PHOSPHATIDYLINOSITOL SIGNALING SYSTEM pathway (continuously downregulated), and ALANINE ASPARTATE AND GLUTAMATE METABOLISM pathway (continuously upregulated), respectively. b) The codownregulation pathway in strongest interaction with core DEPs.
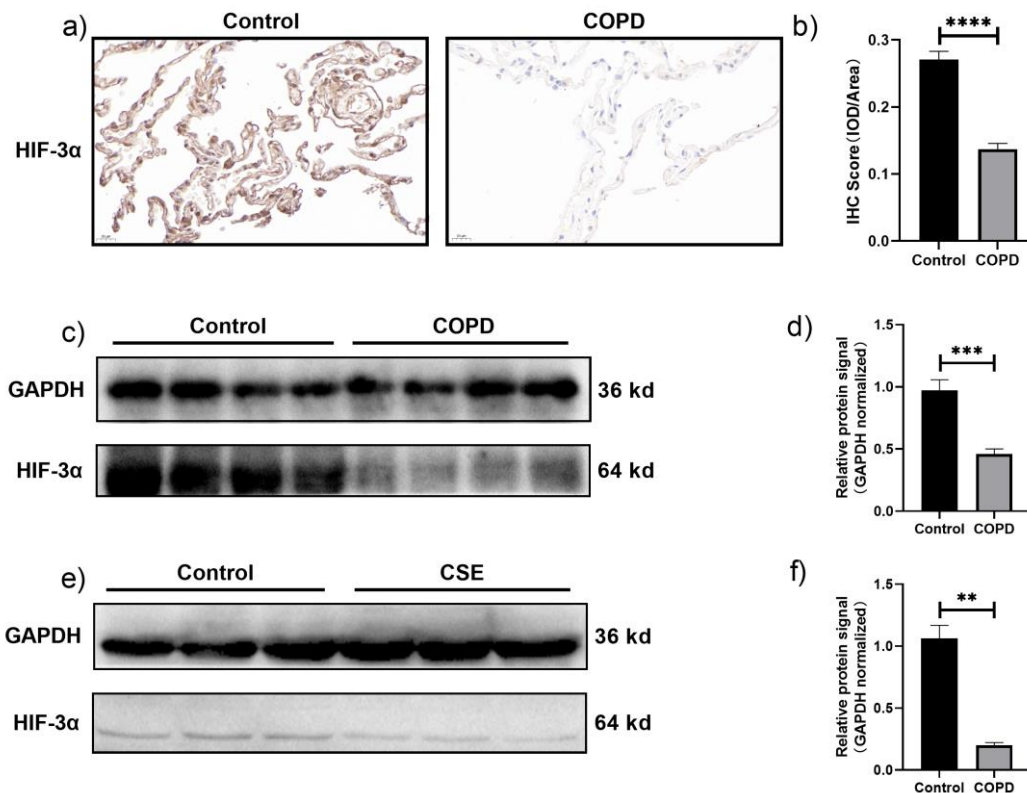
Figure 5. HIF-3a is reduced in the lungs of COPD patients and cigarette smoke extract (CSE)-treated murine lung epithelial (MLE)-12 cells. a) IHC showed that HIF-3a was markedly lower in COPD patients than controls. Original magnification ×40. Scale bar: 20 μm. b) IHC score (mean ± SEM) in airway epithelial cells based on 10 image fields (40X) per sample is shown in the right panel. c, d) Western blots of human lung homogenates from COPD and control patients that were probed with an anti-HIF-3a antibody (normalized to GAPDH). e, f) Western blots of CSE-exposed MLE-12 cells (exposed for 24 h) that were probed with an anti- HIF-3a antibody (normalized to GAPDH) **P < 0.01; ***P < 0.005; ****P < 0.001.
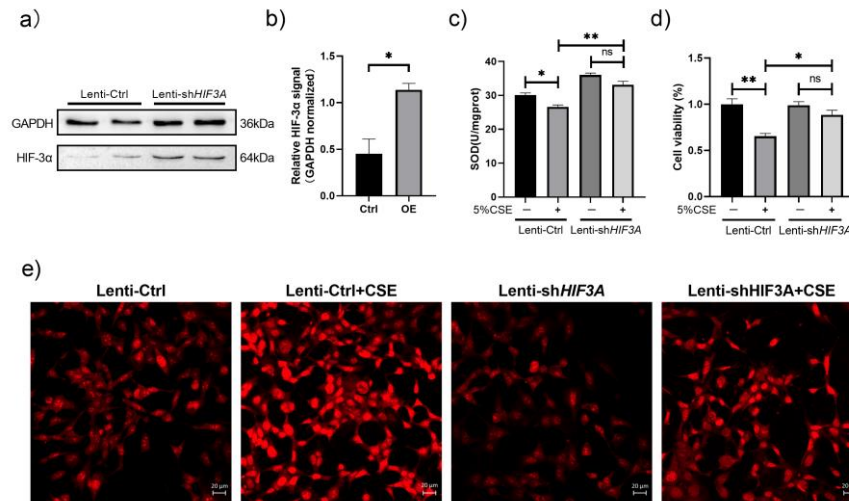
Figure 6. HIF3A overexpression mitigates CSE-induced ROS accumulation and the decreased cell viability in MLE-12. a–b) MLE-12 cells were transfected with HIF3A overexpression vector (HIF3A OE) or control vector. HIF-3a protein levels were detected by Western blot analysis and normalized to GAPDH levels. c) SOD levels in each group. d) Cell viability of each group was analyzed by the CCK-8 assay, and all values were normalized to those of the control group. e) Images of intracellular ROS stained by DCFH-DA were captured with a fluorescence microscope. Bar: 20µm. The results were presented as the means ± S.E.M. for three independent experiments. *P< 0.05; **P < 0.01; ns: P>0.05.

Table S1. List of consistently upregulated genes in the four comparisons.

Names	total	elements
	6	C10orf10
1 UP GSE37768		MSMB
2 UP GSE119040		IGLV1-44
3 UP GSE1650		CYP1B1
4 UP GSE69818		ZNF385D
		MUC4

Table S2. List of consistently downregulated genes in the four comparisons.

Names	total	elements
	8	PPFIBP1
		DDX17
1 DOWN GSE37768		FOSB
2 DOWN		SVEP1
GSE119040		RABGAP1L
3 DOWN GSE1650		HIF3A
4 DOWN GSE69818		PTCH1
		PTPRD

Table S3. Annotation of co-differentially upregulated genes in the four comparisons.

Name	Full name	Annotation
C10orf10	DEPP1 autophagy regulator	A transcriptional target of FOXO3, impairs cellular reactive oxygen species (ROS) detoxification. C10orf10 overexpression elevated cellular ROS levels and sensitized cells to H2O2- and etoposide-induced neuronal cell death[2].
MSMB	microseminoprotein beta	MSMB is synthesized by the epithelial cells of the prostate gland and secreted into the seminal plasma. A previous study identified the increased expression of MSMB was accompanied by the characteristic of small airway epithelial secretory cell changes in COPD[34].
IGLV1-44	immunoglobulin lambda variable 1-44	IGLV1-44 was significantly elevated in POEMS (polyneuropathy, organomegaly, endocrinopathy, monoclonal gammopathy, and skin changes) syndrome compared to the IGLV1-40 (the only two germlines which produce IGLV)[5]
CYP1B1	cytochrome P450 family 1 subfamily B member 1	CYP1B1 protein is a member of the cytochrome P450 superfamily of enzymes. Cytochrome P450 proteins are monooxygenases that catalyze many reactions involved in drug metabolism and the synthesis of cholesterol, steroids, and other lipids. Lung macrophages stimulated by fine particulate matter (PM2.5) can upregulate CYP1B1[6]
ZNF385D	zinc finger protein 385D	ZNF385D may influence several of the negative symptoms of schizophrenia (according to a recent meta-analysis of two genome-wide association studies [GWAS][7]). Therefore, elevated ZNF385D may be associated with the anxiety and depressive symptoms that often occur in patients with COPD[8].
MUC4	mucin 4, cell surface associated	MUC4 encodes a cell-surface glycoprotein that is a transmembrane protein with a large extracellular polypeptide core, and it is upregulated after e-cigarette exposure[9].

Table S4. Annotation of co-differentially downregulated genes in the four comparisons.

PPFIBP1	PPFIA binding protein 1	PPFIBP1 protein is a member of the LAR protein-tyrosine phosphatase-interacting protein (liprin) family, and is an oncogene, based on in vitro focus formation assays and in vivo tumorigenicity assays[10].
DDX17	DEAD-box helicase 17	DDX17 protein was involved in embryogenesis, spermatogenesis, cellular growth and division. Gefitinib-resistant cells have higher DDX17 expression than gefitinib-sensitive cells in non-small-cell lung cancer (NSCLC) patients, DDX17 upregulation enhances gefitinib resistance, while DDX17-silenced cells exhibit partially restored gefitinib sensitivity[11]
FOSB	FosB proto-oncogene	FOSB protein has been implicated as a regulators of cell proliferation, differentiation, and transformation. FOSB was downregulated in NSCLC patients and was negatively correlated with pathological grade[12].
SVEP1	sushi, von Willebrand factor type A, EGF and pentraxin domain containing 1	SVEP1 protein increases leukocyte recruitment to atherosclerotic plaques and promotes atherosclerotic plaque formation[13].
RABGAP1L	RAB GTPase activating protein 1 like	RABGAP1L deletion at the 1q25.1 locus increased the risk of systemic lupus erythematosus in Korean women[14].
HIF3A	hypoxia-inducible factor 3 subunit alpha	HIF3A is a member of the hypoxia-inducible transcription factor (HIF) family that regulate many adaptive responses to hypoxia. HIF3A was downregulated in muscle in patients with COPD GOLD-4 compared to GOLD-1[15].
PTCH1	patched 1	PTCH1 protein is a member of the patched family and is a component of the hedgehog signaling pathway. According to the results of a GWAS toward COPD, PTCH1 showed the strongest positive associations with FEV_1/FVC[16].
PTPRD	protein tyrosine phosphatase receptor type D	PTPRD protein is a member of the protein tyrosine phosphatase (PTP) family. PTPs are signaling molecules that regulate cell growth, differentiation and oncogenic transformation. PTPRD is downregulated in lung adenocarcinoma tissue and is significantly correlated with lung adenocarcinoma staging[17].

Table S5. List of proteins (belonging to key pathways) that have direct interactions with core differentially expressed proteins (DEPs).

Pathway / Core Genes	UBIQUITIN MEDIATED PROTEOLYSIS	OOCYTE MEIOSIS	PATHOGENIC ESCHERICHIA COLI INFECTION	PHOSPHATID YLINOSITOL SIGNALING SYSTEM	ALANINE ASPARTATE AND GLUTAMATE METABOLISM	CORE GENE
PTCH1	NEDD4 ITCH WWP1 SMURF2 WWP2 NEDD4L SMURF1	CCNB1 CDK1	CDH1 ACTB CTNNB1	PIK3CA PIK3CG PTEN		
DDX17	SKP1 MDM2 UBE2I WWP2	AR	NCL CTNNB1			
HIF3A	RBX1 VHL UBE2D3 UBE2D1 CUL2					

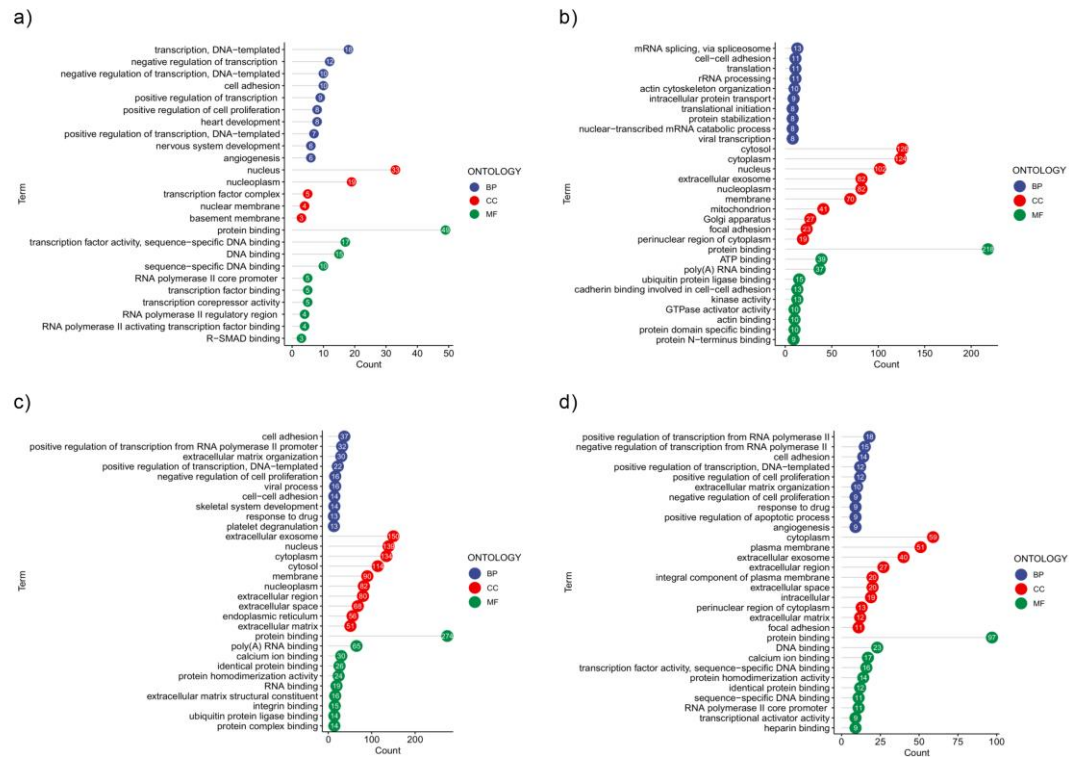|						
	UBE2D2					
RABGAP1L	KLHL9 KLHL13	CDC16		DGKA		
CYP1B1	BRCA1	PGR	ACTB			
FOSB		MAPK12 MAPK1 MAPK3 SMC3 SMC1A				
PTPRD			ABL1	PIK3CA PTEN		PPFIBP1
PPFIBP1		YWHAZ		PI4KB		PTPRD
SVEP1			ITGB1	DGKH		
MSMB		AR			ADSL	
MUC4			CDH1			
ZNF385D						
C10orf10						

Figure S1. Gene Ontology (GO) enrichment analyses for downregulated genes. GO enrichment analyses for a) smokers vs nonsmokers ; b) emphysema vs non emphysema patients; c) severe vs mild emphysema patients; and d) GOLD-4 vs GOLD-1 patients.
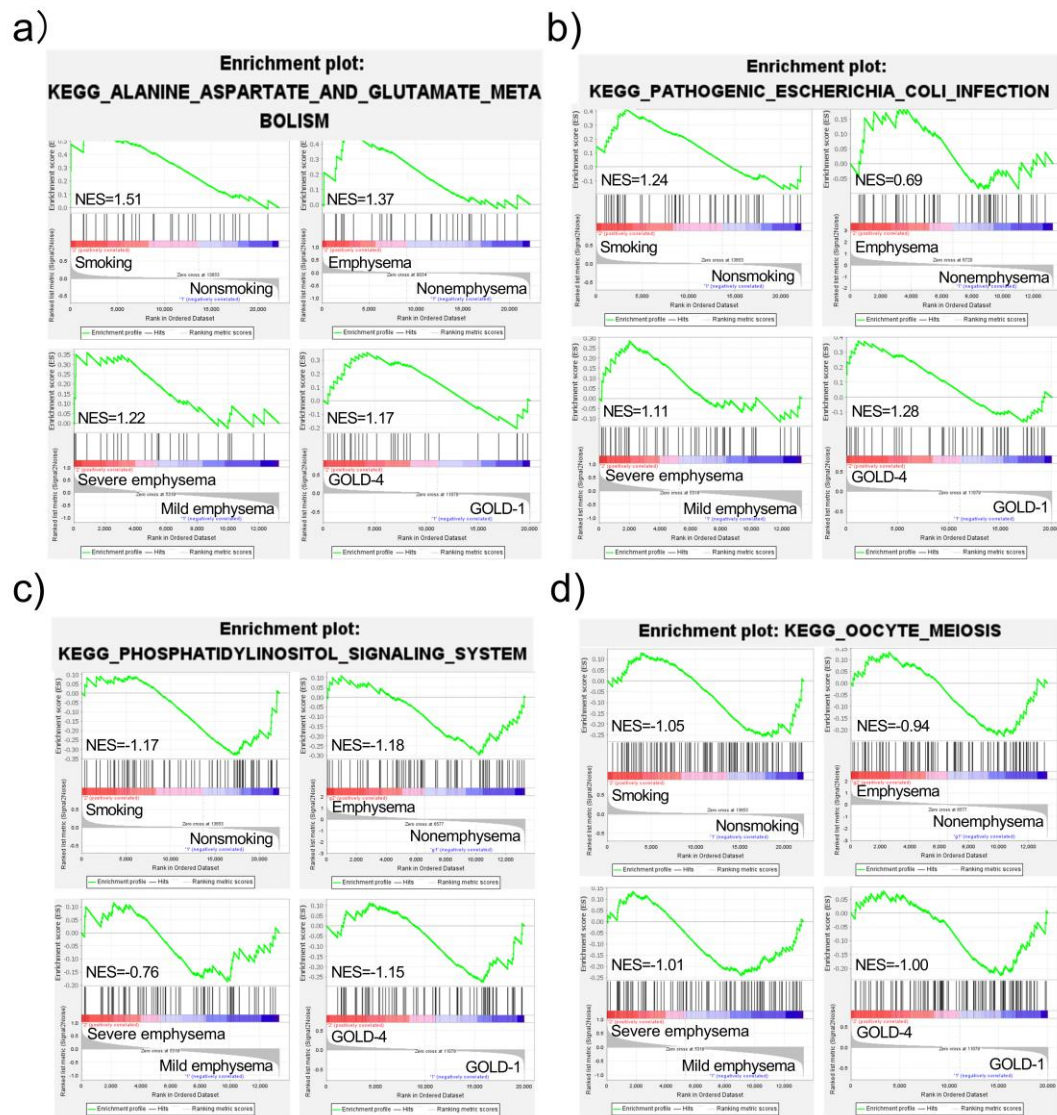
Figure S2. Gene Set Enrichment Analysis (GSEA) regarding the four stages in COPD. Identified 2 co-upregulated pathways (a, b), and 3 co-downregulated pathways (c, d) (FIgure 5b).
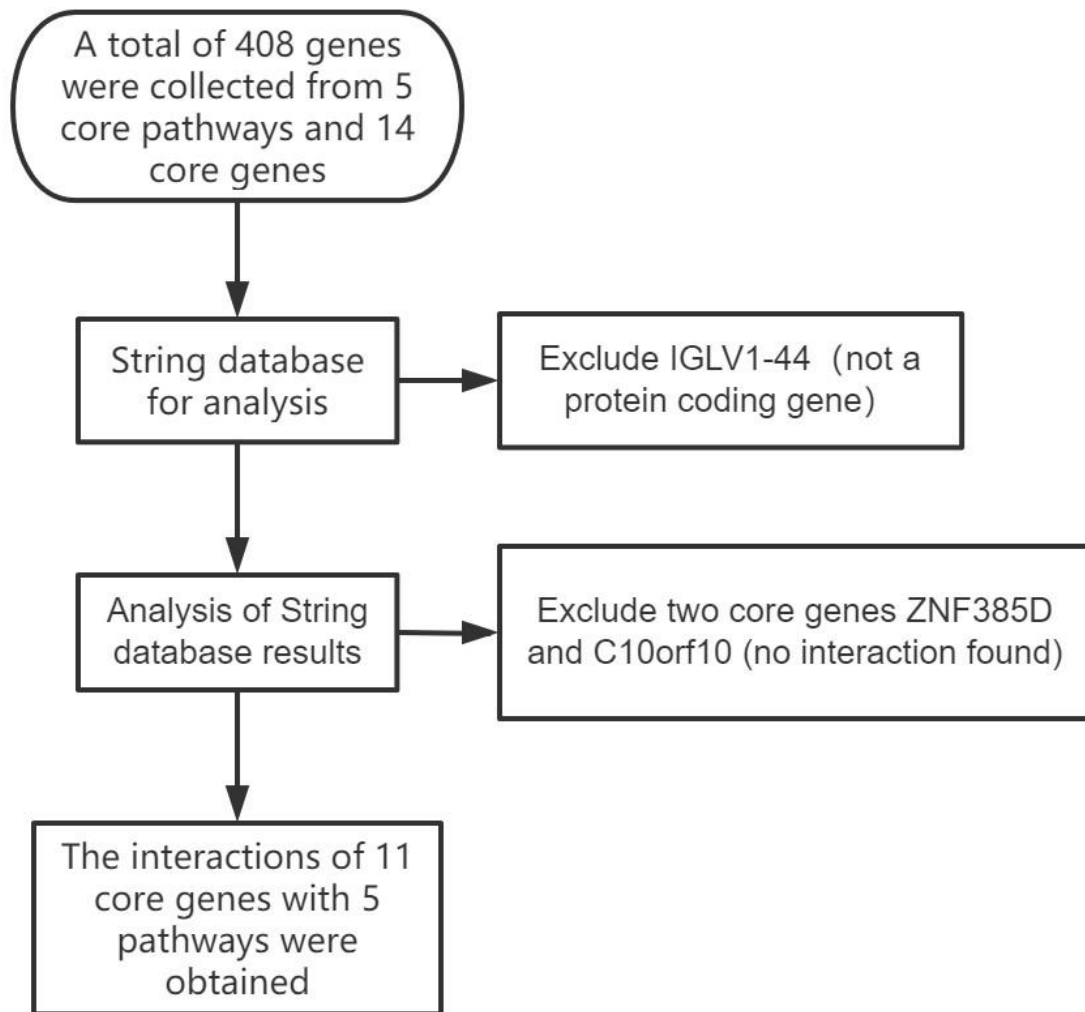
Figure S3. The flow of the core DEGs and core pathways PPI analysis procedure.

References:

1. Zou Y, Palte MJ, Deik AA et al. A GPX4-dependent cancer cell state underlies the clear-cell morphology and confers sensitivity to ferroptosis. *Nat Commun*. 2019;10(1):1617

2. Salcher S, Hagenbuchner J, Geiger K et al. C10ORF10/DEPP, a transcriptional target of FOXO3, regulates ROS-sensitivity in human neuroblastoma. *Mol Cancer*. 2014;13:224

3. Wang G, Lou HH, Salit J et al. Characterization of an immortalized human small airway basal stem/progenitor cell line with airway region-specific differentiation capacity. *Respir Res*. 2019;20(1):196

4. Chen G, Korfhagen TR, Xu Y et al. SPDEF is required for mouse pulmonary goblet cell differentiation and regulates a network of genes associated with mucus production. *J Clin Invest*. 2009;119(10):2914-24

5. Kawajiri-Manako C, Mimura N, Fukuyo M et al. Clonal immunoglobulin lambda light-chain gene rearrangements detected by next generation sequencing in POEMS syndrome. *Am J Hematol*. 2018;93(9):1161-1168

6. Martin PJ, Heliot A, Tremolet G et al. Cellular response and extracellular vesicles characterization of human macrophages exposed to fine atmospheric particulate matter. *Environ Pollut*. 2019;254(Pt A):112933

7. Xu C, Aragam N, Li X et al. BCL9 and C9orf5 are associated with negative symptoms in schizophrenia: meta-analysis of two genome-wide association studies. *Plos One*. 2013;8(1):e51674

8. Yohannes AM, Alexopoulos GS. Depression and anxiety in patients with COPD. *Eur Respir Rev*. 2014;23(133):345-9

9. Ghosh A, Coakley RC, Mascenik T et al. Chronic E-Cigarette Exposure Alters the Human Bronchial Epithelial Proteome. *Am J Respir Crit Care Med*. 2018;198(1):67-76

10. Takeuchi K, Soda M, Togashi Y et al. Pulmonary inflammatory myofibroblastic tumor expressing a novel fusion, PPFIBP1-ALK: reappraisal of anti-ALK immunohistochemistry as a tool for novel ALK fusion identification. *Clin Cancer Res*. 2011;17(10):3341-8

11. Li K, Mo C, Gong D et al. DDX17 nucleocytoplasmic shuttling promotes acquired gefitinib resistance in non-small cell lung cancer cells via activation of beta-catenin. *Cancer Lett*. 2017;400:194-202

12. Ting CH, Lee KY, Wu SM et al. FOSB(-)PCDHB13 Axis Disrupts the Microtubule Network in Non-Small Cell Lung Cancer. *Cancers (Basel)*. 2019;11(1)

13. Winkler MJ, Muller P, Sharifi AM et al. Functional investigation of the coronary artery disease gene SVEP1. *Basic Res Cardiol*. 2020;115(6):67

14. Kim JH, Jung SH, Bae JS et al. Deletion variants of RABGAP1L, 10q21.3, and C4 are associated with the risk of systemic lupus erythematosus in Korean women. *Arthritis Rheum*. 2013;65(4):1055-63

15. Jatta K, Eliason G, Portela-Gomes GM et al. Overexpression of von Hippel-Lindau protein in skeletal muscles of patients with chronic obstructive pulmonary disease. *J Clin Pathol*. 2009;62(1):70-6

16. Obeidat M, Hao K, Bosse Y et al. Molecular mechanisms underlying variations in lung function: a systems genetics analysis. *Lancet Respir Med*. 2015;3(10):782-95

17. Wu A, Yang X, Zhang B, Wang S, Li G. miR-516a-3p promotes proliferation, migration, and invasion and inhibits apoptosis in lung adenocarcinoma by targeting PTPRD. *Int J Clin Exp Pathol*.