



# International Journal of Chronic Obstructive Pulmonary Disease

ISSN: 1176-9106 (Print) 1178-2005 (Online) Journal homepage: [www.tandfonline.com/journals/dcop20](http://www.tandfonline.com/journals/dcop20)

## Gene expression profile of human lung in a relatively early stage of COPD with emphysema

Ina Jeong, Jae-Hyun Lim, Dong Kyu Oh, Woo Jin Kim & Yeon-Mok Oh

To cite this article: Ina Jeong, Jae-Hyun Lim, Dong Kyu Oh, Woo Jin Kim & Yeon-Mok Oh (2018) Gene expression profile of human lung in a relatively early stage of COPD with emphysema, International Journal of Chronic Obstructive Pulmonary Disease, , 2643-2655, DOI: [10.2147/COPD.S166812](https://doi.org/10.2147/COPD.S166812)

To link to this article: <https://doi.org/10.2147/COPD.S166812>



© 2018 Jeong et al. This work is published and licensed by Dove Medical Press Limited



Published online: 28 Aug 2018.



Submit your article to this journal



Article views: 375



View related articles



View Crossmark data



Citing articles: 8 View citing articles

 Open Access Full Text Article

## ORIGINAL RESEARCH

# Gene expression profile of human lung in a relatively early stage of COPD with emphysema

Ina Jeong<sup>1,\*</sup>  
Jae-Hyun Lim<sup>2,\*</sup>  
Dong Kyu Oh<sup>3</sup>  
Woo Jin Kim<sup>4</sup>  
Yeon-Mok Oh<sup>3</sup>

<sup>1</sup>Department of Internal Medicine, National Medical Center, Seoul, Republic of Korea; <sup>2</sup>Biomedical Informatics Research Center, Division of Biomedical Informatics, Seoul National University College of Medicine, Seoul, Republic of Korea; <sup>3</sup>Department of Pulmonary and Critical Care Medicine, Asan Medical Center, University of Ulsan College of Medicine, Seoul, Republic of Korea; <sup>4</sup>Department of Internal Medicine and Environmental Health Center, Kangwon National University Hospital, School of Medicine, Kangwon National University, Chuncheon, Republic of Korea

\*These authors contributed equally to this work

**Purpose:** As only some smokers develop COPD with emphysema, we explored the molecular pathogenesis of early-stage COPD with emphysema using gene expression profiling of human lung tissues.

**Patients and methods:** First, 110 subjects who had smoked more than ten pack-years were classified into three groups: COPD with emphysema, COPD without emphysema, and healthy smokers. COPD and emphysema were confirmed by post-bronchodilator forced expiratory volume in 1 second/forced vital capacity <0.7 and by chest computed tomography. Lung tissues obtained surgically from the 110 subjects were processed and used for RNA-Seq analysis.

**Results:** Among the 110 subjects, 29 had COPD with emphysema, 21 had COPD without emphysema, and 60 were healthy smokers; their mean post-bronchodilator forced expiratory volume in 1 second values were 78%, 80%, and 94%, respectively. Using RNA-Seq, we evaluated 16,676 genes expressed in lung tissues. Among them, 1,226 genes in the COPD with emphysema group and 434 genes in the COPD without emphysema group were differentially expressed genes compared to the expression in healthy smokers. In the COPD with emphysema group, ACER2 and LMAN2L were markedly increased and decreased, respectively. In the COPD without emphysema group, the CHRM3 gene, previously reported to be associated with COPD, and HDAC10 were markedly increased and decreased, respectively.

**Conclusion:** Our study identified differences in gene expression in subjects with COPD according to emphysema status using RNA-Seq transcriptome analysis. These findings may have mechanistic implications in COPD.

**Keywords:** COPD, emphysema, RNA-Seq, transcriptome, gene

## Introduction

COPD is characterized by persistent airflow limitation and is currently the fourth leading cause of death worldwide. COPD is an inflammatory lung disorder with complex pathogenesis and pathological features. Although the mechanisms of COPD are not completely understood, emphysema and small airway disease are the two main characteristics of COPD.<sup>1</sup> While emphysema can exist in individuals who do not have airflow obstruction, it is more common among patients who have moderate or severe airflow obstruction and is associated with increased mortality.<sup>1–3</sup> The mechanisms involved in emphysema development include protease and anti-protease imbalance, persistent airway inflammation, increased oxidative stress, apoptosis of alveolar cells, and ineffective repair.<sup>1,4</sup> However, current therapeutic approaches to COPD mostly focus on targeting chronic bronchiolitis, which appears to be ineffective for emphysema. While cigarette smoking is one of the most well-studied risk factors for COPD, only about 25% of long-term cigarette smokers develop COPD, while the remainder do not.<sup>5</sup> The development of COPD is strongly influenced by genetic factors, and genetic variation

Correspondence: Yeon-Mok Oh  
Department of Pulmonary and Critical Care Medicine, Asan Medical Center, University of Ulsan College of Medicine, 88, Olympic-ro 43-gil, Songpa-gu, Seoul 05535, Republic of Korea  
Tel +82 23 010 3136  
Fax +82 23 010 6968  
Email yeonmok.oh@gmail.com

is also an important determinant of emphysema and airway disease,<sup>6</sup> as well as environmental factors.<sup>7</sup>

Gene expression studies of diseased lungs can provide high-throughput results to better understand the molecular mechanism of COPD. Recent gene expression studies in the lung tissue of emphysema or COPD patients demonstrated and suggested not only several top differentially expressed genes (DEGs) such as known COPD genome-wide association study genes or B-cell-related genes, but also information on pathways related to tissue remodeling and wound repair due to chronic inflammation of COPD.<sup>8–10</sup> Next-generation sequencing technology was applied to transcriptomics recently. RNA-Seq technology provides read counts of RNA fragments in each gene.<sup>11</sup> Background and cross-hybridization are not issues in RNA-Seq, and the technology can quantify both lowly and highly abundant transcripts.<sup>12</sup> The RNA-Seq method has been found to have a high correlation with microarray analyses in gene expression profiles. It was superior at detecting low-abundance transcripts, differentiating biologically critical isoforms, and identifying genetic variants. RNA-Seq was also demonstrated to have a broader dynamic range than microarray, allowing for the detection of more DEGs with higher fold changes.<sup>13</sup> In recent studies on lung diseases using RNA-Seq data, the number of COPD subjects was relatively small. In a previous transcriptome study, we reported a method for identifying DEGs between COPD patients and individuals without the disease, suggesting that oxidative phosphorylation, protein degradation, and chromatin modification were dysregulated pathways in the lung tissues of COPD patients.<sup>14,15</sup> However, COPD patients diagnosed using decreased lung function parameters can have various phenotypes and are a heterogeneous population. Therefore, in this study, we identified subjects with parenchymal emphysema using computed tomography (CT) imaging and performed gene expression profiling using RNA-Seq of lung tissues. The objective of the current study was to investigate the genetic differences between subjects with emphysema COPD and with non-emphysema COPD.

## Patients and methods

### Study subjects and specimens

Subjects had been registered at an in-house tissue storage system, the Asan Biobank, between January 2008 and November 2011. The lungs from the subjects were resected due to malignancy. Immediately after the resection, lung tissues were obtained as far away as possible from the cancer tissues and stored in liquid nitrogen tanks in the Asan Biobank.

We selected only subjects who had smoked more than ten pack-years. We defined past smokers as subjects who quit smoking at least 1 year before lung resection surgery. All current smokers stopped smoking at least 1 month before lung resection surgery. On the contrary, we excluded the following subjects: nonsmokers or smokers with less than ten pack-years of smoking, no CT images, or no post-bronchodilator spirometry data. We also excluded subjects in cases where emphysema was uncertain as evaluated by CT (eg, borderline emphysema with a modified Goddard score <1). In addition, we excluded subjects who had been using inhaled or systemic corticosteroids because corticosteroids may affect the expression of genes related to inflammatory and immune reactions.<sup>16</sup> The flowchart of the study subjects selection is shown in Figure 1.

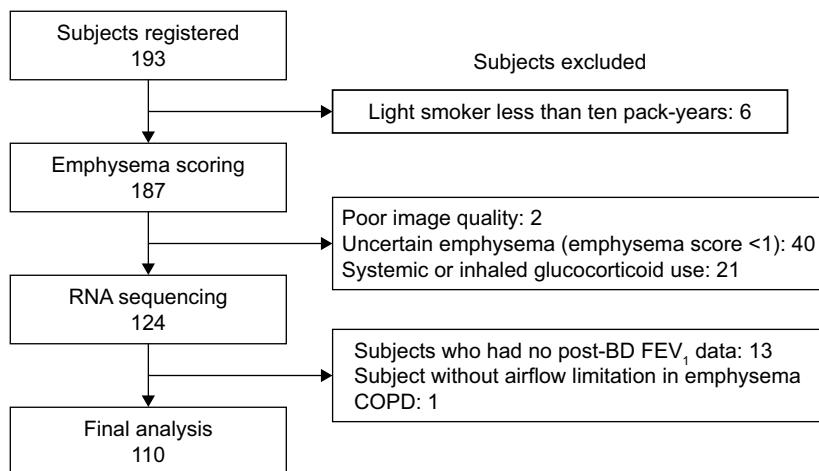
This retrospective study was approved by the institutional review board of Asan Medical Center (protocol no: 2014-0369), and written informed consent was obtained from all subjects.

### Classification of groups

We classified the subjects into three groups: COPD with emphysema, COPD without emphysema, and healthy smokers without emphysema or airflow limitation, because these three groups can have DEGs and different involved pathways. We defined the “COPD with emphysema group” as subjects with emphysema and airflow limitation and the “COPD without emphysema group” as subjects without emphysema but with airflow limitation. Airflow limitation was defined as a post-bronchodilator forced expiratory volume in 1 second ( $FEV_1$ )/forced vital capacity ratio of <0.7 in accordance with the American Thoracic Society/European Respiratory Society criteria.<sup>17</sup>

### Emphysema evaluation

The severity of emphysema was semi-quantitatively assessed by two independent pulmonologists according to the modified Goddard scoring system, a visual scale, in which areas of vascular disruption and low attenuation value were scored for each lung field.<sup>18</sup> Six images were evaluated in three slices of the apex, base, and middle of the lungs. Each image was classified as normal (score 0), ≤5% affected (score 0.5), ≤25% affected (score 1), ≤50% affected (score 2), ≤75% affected (score 3), or >75% affected (score 4), giving a minimum score of 0 and maximum of 4 (Figure 2). Moreover, the sum of the six image scores was converted to an average score that was considered the representative value of the severity of emphysema in each person. All images were blindly evaluated



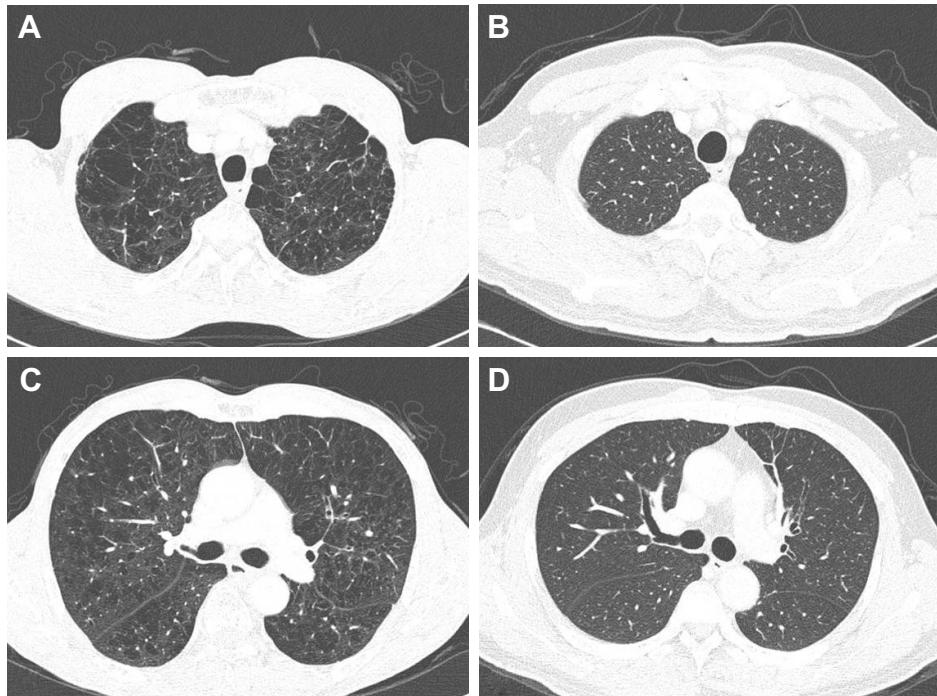
**Figure 1** Selection of study subjects.

**Notes:** Among the 193 subjects who had all been smokers and registered in the Asan Biobank, we excluded 6 subjects for being light smokers, 2 subjects for poor-quality CT images, 40 subjects for uncertainty of emphysema, 21 subjects for current use of systemic or inhaled glucocorticoids, 13 subjects for the absence of post-bronchodilator spirometry data, and 1 subject for no airflow limitation. A total of 110 subjects were included in the final analysis.

**Abbreviations:** BD, bronchodilator; CT, computed tomography; FEV<sub>1</sub>, forced expiratory volume in 1 second.

twice by two educated readers who are pulmonologists. Because an interclass correlation analysis showed that the emphysema score had good agreement between the two physicians (Cronbach's alpha value was 0.967), emphysema scoring by one principal physician was finally applied.

We categorized the subjects into three groups according to emphysema score. Subjects with a final emphysema score of  $>1$  were defined as having “emphysema”; subjects with an emphysema score  $\leq 0.5$  were defined as having “no emphysema”; and the remaining subjects were defined as



**Figure 2** Comparison of CT images of two subjects with high emphysema score and low emphysema score.

**Notes:** (A) Upper chest images of the emphysema subject, showing 75%–100% of LAA in the right upper lobe, modified Goddard classification 4, and showing 50%–75% of LAA in the left upper lobe, modified Goddard classification 3. (B) Upper chest images of the “no emphysema” subject, showing 0%–5% of LAA in the right upper lobe, modified Goddard classification 0.5. (C) Middle chest images of the emphysema subject, showing 25%–50% of LAA in the right middle lobe, modified Goddard classification 2. (D) Middle chest images of the “no emphysema” subject, showing 0% of LAA in the right middle lobe, modified Goddard classification 0.

**Abbreviations:** CT, computed tomography; LAA, low attenuation area.

“uncertainty of emphysema”. We decided to exclude the 40 “uncertainty of emphysema” subjects from the genetic analysis to avoid confusion, particularly with subjects with borderline emphysema.

## RNA preparation and sequencing

Total RNA was isolated from apparently normal fresh frozen lung tissue that was remote from the lung cancer. RNA integrity was assessed using an Agilent Bioanalyzer, and RNA purity was assessed using a NanoDrop spectrophotometer. One microgram of total RNA was used to generate cDNA libraries using the TruSeq RNA library kit. The protocol consisted of poly A-selected RNA extraction, RNA fragmentation, reverse transcription using random hexamer primers, and 100 bp paired-end sequencing using the Illumina HiSeq 2000 system. All data have been deposited in the National Center for Biotechnology Information (NCBI) Gene Expression Omnibus public repository and can be accessed through the accession number GSE57148.

## Quality control and data management

For quality control, read quality was verified using FastQC and read alignment was verified using Picard. DEG analysis was performed using TopHat and Cufflinks software.<sup>19</sup> To estimate the expression levels, the RNA-seq reads were mapped to the human genome using TopHat (version 1.4.1)<sup>20</sup> and quantified using Cufflinks software 2.0.0.<sup>21</sup> Cufflinks software was run with the UCSC hg19 human genome and transcriptome references. The numbers of isoform and gene transcripts were calculated, and the relative abundance of transcripts was measured in fragments per kilobase of exon per million fragments mapped (FPKM). Expression levels were extracted as an FPKM value for each gene of each sample using Cufflinks software. Genes with FPKM values of 0 across all samples were excluded. Filtered data were subjected to upper quantile normalization. Statistical significance was determined using Student’s *t*-test. The false discovery rate was controlled by adjusting values using the Benjamini–Hochberg algorithm. The analysis steps used are shown in Figure S1.

## Pathway analysis

Functional enrichment analysis was performed using gene set enrichment analysis (GSEA; version 2.0.8), which combines information from previously defined gene sets obtained from the Molecular Signature Database (version 6.1). Biological gene functional annotation analysis was performed using DAVID (version 6.7) with a list of DEGs.

## Statistical analysis

Clinical statistical analyses were performed using SPSS v18.0 (SPSS Inc., Chicago, IL, USA). The results are expressed as the mean  $\pm$  SD. One-way analysis of variance (ANOVA) was performed using R software version 3.3.3 to identify DEGs between patients with COPD with emphysema, with COPD without emphysema, and healthy smokers. For genes with adjusted *P*-values  $<0.01$  in ANOVA, Tukey’s honestly significant difference post hoc test was used for two-group comparisons.

## Results

### Study subjects

Among the 110 subjects, 29 had COPD with emphysema, 21 had COPD without emphysema, and 60 were healthy smokers. Their mean post-bronchodilator FEV<sub>1</sub> values were 78%, 80%, and 94%, respectively (Table 1). Among the three groups of subjects, the group of healthy smokers was significantly younger (67.0 vs 65.7 vs 60.0 years; *P*<0.001) and smoked less (43.6 vs 49.1 vs 37.0 pack-years; *P*=0.039). On the contrary, the pre-bronchodilator and post-bronchodilator FEV<sub>1</sub> values and the ratio of FEV<sub>1</sub> to forced vital capacity were significantly higher in healthy smokers. There were no significant differences in comorbidities or inhaled bronchodilator use among the three groups of subjects (Table 1).

### Differentially expressed genes

Data about quality control of transcript analysis of RNA-Seq were collected (see Supplementary material). After filtering for genes with zero counts in the whole samples, noncoding genes, and low-variance genes, a total of 16,676 genes were analyzed. Out of these genes, 4,087 genes were differentially expressed among the three groups (Benjamini–Hochberg adjusted). In the COPD with emphysema group, 1,588 genes were upregulated and 1,519 genes were downregulated, compared with those in healthy smokers. In the COPD without emphysema group, 1,257 genes were upregulated and 1,058 genes were downregulated, compared with those in healthy smokers. In both COPD groups, 939 genes were upregulated and 942 genes were downregulated compared with those in healthy smokers, and no genes were reciprocally regulated. Thus, 649 genes were upregulated in the COPD with emphysema group, but not in the COPD without emphysema group, compared with those in healthy smokers, and 577 genes were downregulated in the COPD with emphysema group, but not in the COPD without emphysema group, compared with those in healthy smokers. Moreover, 318 genes were upregulated in the COPD without emphysema group, but not in the COPD with emphysema group, compared with those in healthy smokers, and 116 genes were downregulated in the

**Table 1** Characteristics of the subjects in the three groups

	COPD with emphysema (n=29)	COPD without emphysema (n=21)	Healthy smokers (n=60)	P-value
Age, years	67.0±6.2	65.7±6.3	60.0±9.3	<0.001
Pack-years	43.6±17.1	49.1±23.5	37.0±18.8	0.039
FEV <sub>1</sub> , L (pre-bronchodilator)	2.26±0.56	2.39±0.47	3.00±0.57	<0.001
FEV <sub>1</sub> , % (pre-bronchodilator)	73.4±14.9	75.7±10.8	90.9±12.1	<0.001
FEV <sub>1</sub> , L (post-bronchodilator)	2.39±0.57	2.51±0.40	3.10±0.59	<0.001
FEV <sub>1</sub> , % (post-bronchodilator)	77.5±14.9	79.7±8.8	94.0±13.0	<0.001
FEV <sub>1</sub> /FVC (post-bronchodilator)	59.6±8.6	64.1±3.8	76.9±4.2	<0.001
Current smokers, n (%)	13 (44.8)	2 (9.5)	19 (31.7)	0.037
COPD	29	21	0	
GOLD1, n (%)	13 (44.8)	13 (61.9)	N/A	0.233
GOLD2, n (%)	15 (51.7)	8 (38.1)	N/A	0.340
GOLD3, n (%)	1 (3.4)	0	N/A	1
DM medication, n (%)	6 (20.7)	7 (33.3)	8 (13.3)	0.129
HT medication, n (%)	12 (41.4)	13 (61.9)	20 (33.3)	0.072
Inhaled long-acting bronchodilator use, n (%)	4 (13.8)	1 (4.8)	2 (3.3)	0.157
Inhaled short-acting bronchodilator use, n (%)	8 (27.6)	8 (38.1)	18 (30)	0.711

**Abbreviations:** DM, diabetes mellitus; FEV<sub>1</sub>, forced expiratory volume in 1 second; FVC, forced vital capacity; GOLD, Global initiative for chronic Obstructive Lung Disease; HT, hypertension; N/A, not applicable.

COPD without emphysema group, but not in the COPD with emphysema group, compared with those in healthy smokers (Table 2; Figure S2). Therefore, a total of 1226 genes in the COPD with emphysema group and 434 genes in the COPD without emphysema group were differentially expressed genes respectively compared to expression in healthy smokers. Among the top DEGs in the COPD with emphysema group, ACER2 and LMAN2L were markedly increased and decreased, respectively. In addition, the CHRM3 gene and HDAC10 were markedly increased and decreased, respectively, in the COPD without emphysema group (Table 3).

We performed EdgeR analysis to confirm the results, and the results were consistent with 95% or more of the *t*-test. We also performed two-way ANOVA using the presence and absence of emphysema and current smoking status as independent variables. Moreover, we performed three-way ANOVA

including emphysema presence and current smoking interaction term because current smoking status has an important effect on emphysema development. However, for the DEG data, the two-way ANOVA results showed about 90% agreement with the one-way ANOVA results, and the three-way ANOVA results were similar to the two-way ANOVA results.

## Pathway analysis

In the pathway analysis using DAVID, mitochondrion genes, response to redox state, and cell redox homeostasis-related pathways decreased in expression in the COPD with emphysema group, compared to those in healthy smokers. Mitochondrion genes, response to oxidative stress, positive regulation of DNA repair, and mismatch repair-related pathways showed decreased expression in the COPD without emphysema group, compared to those in healthy smokers. Lamellipodium, cell–cell adherens junction, cell adhesion, cadherin binding involved in cell–cell adhesion, and focal adhesion-related pathways increased in expression in the COPD without emphysema group, compared to those in healthy smokers (Table 4). Heat maps indicating the genes that were differentially expressed between each group of COPD subjects and healthy smokers are shown in Figures 3 and 4. Also, the additional analysis on the GSEA as well as the pathway analysis showed no significant relation with cancer gene ontology term.

## Discussion

In this study, we identified new genes related to emphysema by using RNA-Seq to analyze the gene expression profiles

**Table 2** Brief summary of DEGs results

COPD subgroup	Number of upregulated genes	Number of downregulated genes
COPD with emphysema <sup>a</sup>	1,588	1,519
COPD without emphysema <sup>a</sup>	1,257	1,058
Both types of COPD <sup>a</sup>	939	942
COPD with emphysema group only <sup>a,b</sup>	649	577
COPD without emphysema group only <sup>a,c</sup>	318	116

**Notes:** <sup>a</sup>Compared with those in healthy smokers. <sup>b</sup>Not in COPD without emphysema group. <sup>c</sup>Not in COPD with emphysema group.

**Abbreviation:** DEG, differentially expressed gene.

**Table 3** Top ten genes with increased and decreased expression in COPD with emphysema subjects and in COPD without emphysema subjects compared with healthy smokers

Gene	Gene function	Fold change $\log_2$ (emphysema COPD/control)	P-value	Expression levels ( $\log_2$ [FPKM])
<b>Increased expression in COPD with emphysema subjects compared with healthy smokers</b>				
CCDC134	Coiled-coil domain containing 134	0.61	$1.16 \times 10^{-7}$	1.91
FAT1	FAT atypical cadherin 1	0.69	$8.97 \times 10^{-7}$	3.40
LONRF2	LON peptidase N-terminal domain and ring finger 2	0.62	$9.19 \times 10^{-7}$	1.21
XYLB	Xylulokinase	0.59	$9.77 \times 10^{-7}$	0.50
HDLBP	High-density lipoprotein binding protein	0.32	$9.79 \times 10^{-7}$	48.08
DUSP28	Dual specificity phosphatase 28	0.59	$1.30 \times 10^{-6}$	1.62
ACER2	Alkaline ceramidase 2	0.71	$1.63 \times 10^{-6}$	1.36
MTFR1	Mitochondrial fission regulator 1	0.32	$1.76 \times 10^{-6}$	5.73
B3GALNT2	Beta-1,3-N-acetylgalactosaminyltransferase 2	0.26	$2.67 \times 10^{-6}$	5.32
CNNM2	Cyclin and CBS domain divalent metal cation transport mediator 2	0.47	$4.83 \times 10^{-6}$	2.34
<b>Decreased expression in COPD with emphysema subjects compared with healthy smokers</b>				
TSEN15	TRNA splicing endonuclease subunit 15	-0.36	$6.33 \times 10^{-8}$	18.24
C7orf36	Yael domain containing 1	-0.44	$6.9 \times 10^{-7}$	12.17
N6AMT2	N-6 adenine-specific DNA methyltransferase 2	-0.35	$1.45 \times 10^{-6}$	10.72
LMAN2L	Lectin, mannose binding 2 like	-0.32	$2.90 \times 10^{-6}$	8.76
IFI27	Interferon alpha inducible protein 27	-0.72	$3.17 \times 10^{-6}$	316.14
NIINJ2	Ninjurin 2	-0.60	$3.92 \times 10^{-6}$	14.89
PGAP2	Post-GPI attachment to proteins 2	-0.47	$8.05 \times 10^{-6}$	10.16
RAB9A	RAB9A, member RAS oncogene family	-0.26	$8.34 \times 10^{-6}$	19.93
CNPPD1	Cyclin Pas1/PHO80 domain containing 1	-0.27	$8.39 \times 10^{-6}$	29.09
LAMTOR3	Late endosomal/lysosomal adaptor, MAPK and MTOR activator 3	-0.26	$1.15 \times 10^{-5}$	10.13
<b>Increased expression in COPD without emphysema subjects compared with healthy smokers</b>				
MYO10	Myosin X	0.56	$8.61 \times 10^{-6}$	7.85
ZNF432	Zinc finger protein 432	0.42	$9.89 \times 10^{-6}$	3.75
ZNF532	Zinc finger protein 532	0.28	$1.29 \times 10^{-5}$	9.21
CHRM3	Cholinergic receptor muscarinic 3	0.75	$2.03 \times 10^{-5}$	0.73
PPFIBP1	PPFIA binding protein 1	0.52	$2.38 \times 10^{-5}$	17.08
SLIT2	Slit guidance ligand 2	0.84	$4.84 \times 10^{-5}$	14.47
PAXIPI	PAX interacting protein 1	0.23	$5.16 \times 10^{-5}$	3.15
PTPRG	Protein tyrosine phosphatase, receptor type G	0.74	$5.36 \times 10^{-5}$	4.03
C7orf49	Chromosome 7 open reading frame 49	0.26	$6.14 \times 10^{-5}$	10.72
DOCK9	Dedicator of cytokinesis 9	0.43	$6.17 \times 10^{-5}$	15.11
<b>Decreased expression in COPD without emphysema subjects compared with healthy smokers</b>				
MRPL45	Mitochondrial ribosomal protein L45	-0.18	$2.47 \times 10^{-4}$	13.87
DHRS4	Dehydrogenase/reductase 4	-0.40	$2.60 \times 10^{-4}$	16.70
SWI5	Homologous recombination repair protein	-0.43	$3.81 \times 10^{-4}$	9.26
AHCY	Adenosylhomocysteinase	-0.28	$3.97 \times 10^{-4}$	21.32
NUDT2	Nudix hydrolase 2	-0.44	$4.97 \times 10^{-4}$	13.69
HDAC10	Histone deacetylase 10	-0.71	$5.91 \times 10^{-4}$	6.74
C4orf52	Small integral membrane protein 20	-0.31	$6.82 \times 10^{-4}$	22.33
STOML2	Stomatin like 2	-0.24	$7.23 \times 10^{-4}$	27.10
BAX	BCL2 associated X, apoptosis regulator	-0.45	$7.40 \times 10^{-4}$	45.84
ECH1	Enoyl-CoA hydratase 1	-0.30	$7.57 \times 10^{-4}$	76.13

**Abbreviation:** FPKM, fragments per kilobase of exon per million fragments mapped.

of human lung tissues from COPD subjects with or without emphysema and comparing them with those from healthy smokers. For proper analysis, we classified the COPD subjects according to emphysema phenotype into two groups,

one with emphysema and the other without emphysema, using visual emphysema scoring on CT scans. We found different gene sets in the two groups of COPD with emphysema and without emphysema.

**Table 4** Representative DAVID results for pathways with decreased expression in COPD with emphysema subjects and with increased and decreased expression in COPD without emphysema subjects compared with healthy smokers

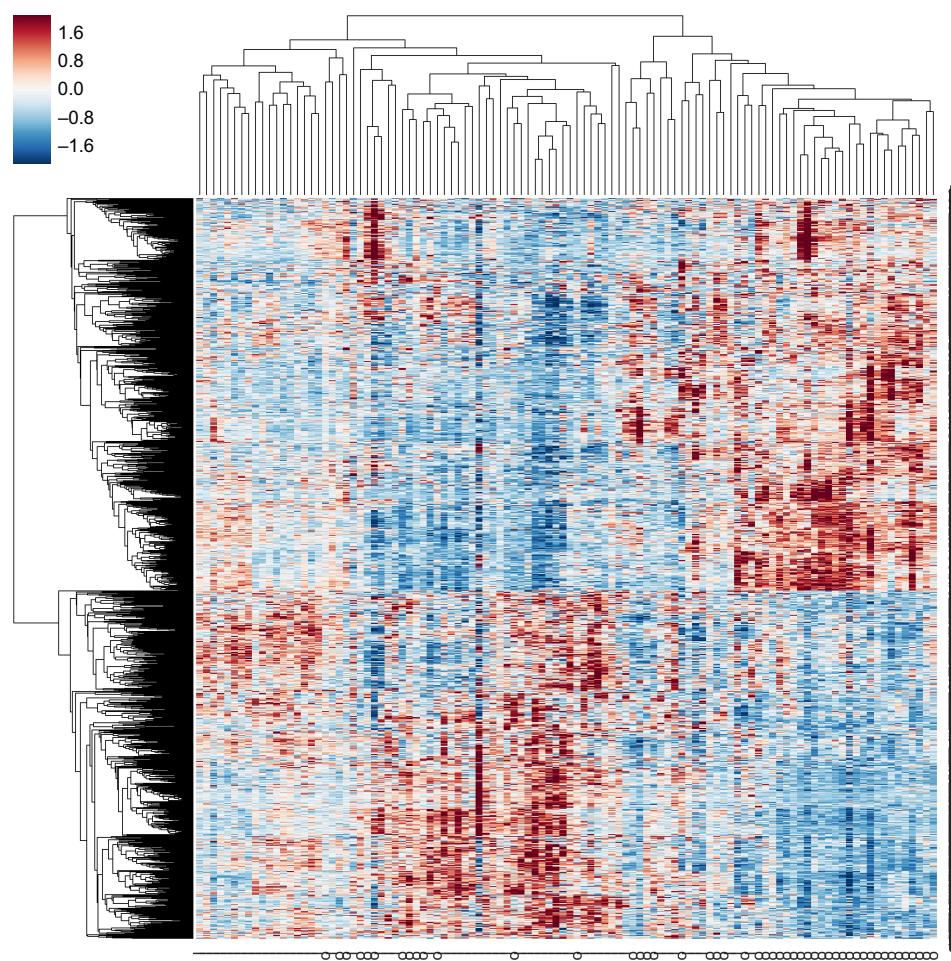
Terms	Count of genes involved	Fold enrichment	FDR
<b>Representative DAVID results for pathways with decreased expression in COPD with emphysema subjects compared with healthy smokers</b>			
GO:0005739-mitochondrion	71	1.96	1.03×10 <sup>-4</sup>
GO:0070062-extracellular exosome	120	1.57	4.14×10 <sup>-4</sup>
GO:0005515-protein binding	289	1.18	0.03
GO:0005829-cytosol	124	1.37	0.17
GO:0051775-response to redox state	4	14.54	3.54
GO:0006695-cholesterol biosynthetic process	6	5.74	5.95
GO:0036109-alpha-linolenic acid metabolic process	4	11.18	7.78
GO:0045454-cell redox homeostasis	8	3.78	8.37
GO:0000502-proteasome complex	7	4.28	7.70
GO:0031902-late endosome membrane	9	3.27	8.50
<b>Representative DAVID results for pathways with increased expression in COPD without emphysema subjects compared with healthy smokers</b>			
GO:0030027-lamellipodium	13	5.36	7.81×10 <sup>-3</sup>
GO:0005913-cell–cell adherens junction	17	3.48	0.05
GO:0007156-homophilic cell adhesion via plasma membrane adhesion molecules	12	4.74	0.07
GO:0098641-cadherin binding involved in cell–cell adhesion	16	3.40	0.11
GO:0005925-focal adhesion	18	3.04	0.13
GO:0007010-cytoskeleton organization	11	4.27	0.44
GO:0005911-cell–cell junction	11	4.22	0.39
GO:0005737-cytoplasm	106	1.34	0.50
GO:0005856-cytoskeleton	16	2.85	0.73
GO:0005509-calcium ion binding	25	2.15	0.89
<b>Representative DAVID results for pathways with decreased expression in COPD without emphysema subjects compared with healthy smokers</b>			
GO:0005739-mitochondrion	19	2.55	0.43
GO:0005829-cytosol	33	1.78	0.90
GO:0034644-cellular response to UV	4	16.07	2.81
GO:0005515-protein binding	61	1.30	2.96
GO:0005743-mitochondrial inner membrane	9	3.65	3.80
GO:0006979-response to oxidative stress	5	8.03	4.94
GO:0003723-RNA binding	9	3.09	10.18
GO:0070062-extracellular exosome	26	1.65	11.28
GO:0045739-positive regulation of DNA repair	3	17.68	16.76
GO:0006298-mismatch repair	3	15.15	21.88

Note: FDR=false positive/total positive.

Abbreviations: FDR, false discovery rate; GO, gene ontology; UV, ultraviolet.

In total, 4,087 genes were identified with a Benjamini–Hochberg adjusted *P*-value <0.01. We previously validated a subset of RNA-Seq data with quantitative real-time polymerase chain reaction and the results were in good agreement.<sup>15</sup> We attempted to identify notable genes related to emphysema or COPD among the 649 upregulated and 577 downregulated genes that were differentially expressed in the COPD with emphysema group. However, the COPD with emphysema group did not exhibit any previously known remarkable DEG. Although emphysema is defined as enlargement of the respiratory acinus, the pattern of enlargement may not be uniform, and different genetic or epigenetic etiologies of emphysema

tend to produce different patterns of acinar enlargement. Moreover, some subjects with COPD with emphysema may also overlap with COPD subjects without emphysema, and finally, it is difficult to differentiate exactly between the two groups of COPD, as genes associated with the COPD without emphysema group might also be associated and, therefore, expressed in the COPD with emphysema group. Nonetheless, there are some genes associated with emphysema among the top-rated genes showing increased or decreased expression in the COPD with emphysema group. All new genes related to emphysema have relatively low expression levels, and it could be possible that those of DEGs in the present study have



**Figure 3** Heat map of gene expression in the lung tissues on comparing COPD with emphysema vs healthy smokers.

**Notes:** Hierarchical clustering of COPD with emphysema and healthy smoker is shown. The small signs at the bottom indicate the status of the subjects, where “o (case)” indicates subjects with COPD with emphysema and “— (control)” indicates healthy smokers.

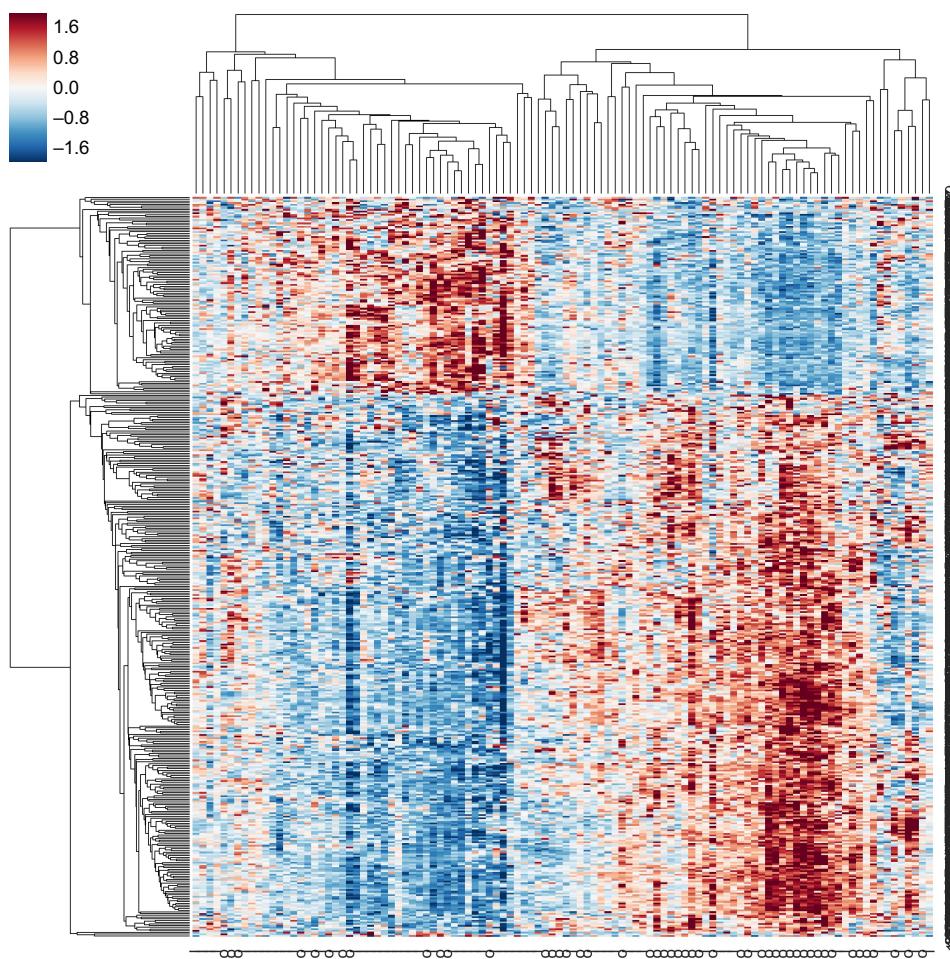
not been found in previous other next-generation sequencing technology experiments other than RNA-seq since the expression level was relatively low.

ACER2, which was one of the genes with a high expression in the COPD with emphysema group, is a critical sphingolipid metabolizing enzyme and has been shown to be markedly upregulated and involved in apoptosis in response to various stimuli such as DNA damage.<sup>22,23</sup> Ceramide upregulation was suggested to be a possible crucial mediator of alveolar destruction in an animal emphysema model experiment.<sup>23</sup> It is not known whether ACER2 with increased expression was associated with apoptosis of the emphysema in the present study. FAM13A, known as a COPD-related gene, was significantly upregulated in COPD with emphysema group ( $P$ -value 0.00096), and this is consistent with many previous genetic studies.<sup>9,10</sup>

Mannose-binding lectin protein is a known key component of innate immunity that promotes macrophage phagocytosis

and clears apoptotic cells. It was shown to decrease in the airway following oxidative stress and be associated with clinical outcomes in COPD in recent studies.<sup>24,25</sup> Because LMAN2L also showed decreased expression in the COPD with emphysema group in the present study, it is a possible emphysema-related downregulated gene.

Meanwhile, among the 318 upregulated and 116 downregulated genes that were differentially expressed in the COPD without emphysema group, several genes were remarkable and previously reported to be associated with respiratory disease. The CHRM3 gene showed increased expression in this group of COPD without emphysema subjects compared to that in the healthy smoker group. M3-cholinoreceptors are expressed on macrophages, one of the key cells involved in inflammation, and contribute to their chemotaxis. Although little is known about the role of CHRM3 gene in the development of airway obstruction or COPD, it was recently suggested to be a druggable target candidate gene.<sup>26</sup> According to



**Figure 4** Heat map of gene expression in the lung tissues on comparing COPD without emphysema vs healthy smokers.

**Notes:** Hierarchical clustering of COPD without emphysema and healthy smoker is shown. The small signs at the bottom indicate the status of the subjects, where "o (case)" indicates subjects with COPD without emphysema and "— (control)" indicates healthy smokers.

a recent study, the differential expression of the CHRM3 gene might be associated with the clinical phenotypes of COPD with bronchial hyperresponsiveness.<sup>27</sup> Significantly lower levels of CHRM3 mRNA were observed in patients with COPD compared with those in asthma patients, and CHRM3 gene expression was significantly elevated in COPD patients with bronchial hyperresponsiveness compared with that in patients without bronchial hyperresponsiveness.<sup>27</sup> The SLIT2 gene also showed increased expression in COPD without emphysema subjects compared with that in healthy smokers. SLIT2 is present in murine fetal lungs and adult rat lungs and is expressed by club cells in the adult mouse. However, little is known about the expression of SLIT2 in human lung tissue during inflammation and fibrosis. In another study, SLIT2, which inhibits fibrocyte differentiation, was shown to be produced by fibroblasts and epithelial cells in the lung tissue from patients with relatively normal lung function, whereas

less SLIT2 was found in the fibrotic lesions of patients with advanced disease. Whether the upregulated SLIT2 level is augmented by fibroblasts, which modulate the immune system, in a relatively early stage of the disease (COPD without emphysema rather than COPD with emphysema) or there are factors present that upregulate SLIT2 production is unclear. SLIT2 levels seem to be important in the regulation of lung inflammation and fibrosis, and therapeutic use of SLIT2 may be beneficial for regulating fibrosis.<sup>28,29</sup>

Among the top ten genes with decreased DEGs in the COPD without emphysema group, the expression levels of HDAC10 and STOML2 were noticeable. HDAC2 is known to play a critical role in suppressing inflammatory gene expression in the airways, lung parenchyma, and alveolar macrophages, and HDAC activity was found to be decreased in patients with COPD.<sup>29</sup> Histone deacetylation modulates chromatin structure and plays an important role

in transcriptional regulation, cell cycle progression, and developmental events. HDAC10 is a novel class IIB histone deacetylase found in both the nucleus and the cytoplasm, and it was reported to function as a transcriptional repressor.<sup>30</sup> It was reported that the knockdown of HDAC10 resulted in the accumulation of reactive oxygen species (ROS) and cell death, indicating its importance in the response to oxidative stress in gastric cancer cell lines.<sup>31</sup> STOML2 is a gene that plays a key role in mitochondrial function and T-cell activation. Recently, the expression of STOML2 was negatively correlated with that of the IL6R, which has been implicated in asthma pathophysiology, and the increased expression of STOML2 was associated with an increased asthma risk by RNA-seq analysis.<sup>32</sup> The exact roles and mechanisms of HDAC10, STOML2, and other genes in COPD, especially in early stages of the disease, need to be clarified in future studies. BAX, an apoptosis regulator, is a marker of upregulated apoptosis in COPD, and it was found to downregulated in the COPD without emphysema group compared with that in healthy smokers in the present study, contrary to a previous report that patients with emphysema showed a significant increase in the Bax index.<sup>33</sup> Alterations in the expression of apoptosis-related proteins might be important in determining progression to apoptosis in emphysema pathogenesis.

The major analytic strategy for interpreting hundreds of statistically significant differences in gene expression involves pathway analysis. Previous studies on the COPD transcriptome reported a high degree of overlap in the biological processes affected. In the current study, the most altered pathway in the COPD without emphysema group was lamellipodium organization and cell–cell adherens junctions, possibly providing a clue for the pathogenesis of emphysema. A study found that the genes involved in focal adhesion and lamellipodium organization showed decreased expression with disease severity in emphysema patients.<sup>34</sup> Lamellipodium is involved in cell migration, which suggests cytoskeleton reorganization related to epithelial wound repair.<sup>35</sup> Cell–cell adherens junctions in airway epithelium can function as a barrier in the airway tract. Impairment of the barrier function may increase susceptibility to infection and promote exaggerated and prolonged innate immune responses to pathogens, resulting in chronic inflammation.<sup>35</sup> Pathways related to focal adhesion, cytoskeleton or organization, and cell–cell junctions also showed increased expression in subjects with COPD without emphysema. The repair process of airway epithelium after injury in the early stage of COPD might be augmented as a defense mechanism. There was a study about genes decreasing in expression such as transforming growth factor beta pathway, actin organization,

and integrin signaling were enriched in tissue repair processes in the lung tissue of emphysematous destruction.<sup>9</sup> Also, another recent report about the differential expression of genes associated with tissue repair between small airway obstructive and emphysematous destructive lesions suggested a mechanism for the development of peripheral lesions of COPD.<sup>36</sup> The expression of mitochondrial genes was previously shown to be reduced in the lung tissues of COPD subjects using a serial analysis of gene expression,<sup>37</sup> which is consistent with the results of the current study for both COPD groups. Pathways related to cytosol, cellular responses to ultraviolet, responses to oxidative stress, positive regulation of DNA repair, and mismatch repair showed decreased expression in the COPD without emphysema group. Mitochondria are multifunctional cellular organelles and play an important role in not only numerous aspects of cell morphology and physiology, but also the innate immune system.<sup>38</sup> Oxidant–antioxidant imbalance is recognized as one of the significant factors in COPD pathogenesis. One endogenous source of ROS is mitochondria. Although the leakage of electrons from the electron transport chain and the formation of ROS result from the physiological functioning of mitochondria, there are various intra- and extracellular factors that may increase these amounts and significantly contribute to oxidative–antioxidative imbalance.<sup>39</sup> Proteasome complex-related genes were also decreased in expression in the COPD with emphysema group in the present study. The proteasome is not only a multicatalytic enzyme complex that is responsible for the maintenance of cellular proliferation, cell cycling, gene transcription, apoptosis, and antioxidant responses, but also serves as a protein quality control system. Impaired proteasome function can result in the accumulation of altered proteins, resulting in cellular dysfunction and cell death. Recently, decreased proteasomal function was shown to accelerate cigarette smoke-induced pulmonary emphysema in a mouse model.<sup>40</sup> Impaired proteasomal expression might be important in the pathogenesis of emphysema.

Pathways with increased expression in the COPD with emphysema group included genes related to adenosine triphosphate binding, helicase activity, nucleoplasm, adenosine triphosphate-dependent RNA helicase activity, membranes, and the endoplasmic reticulum, among others, which could be nonspecifically upregulated in both COPD groups. In addition, we did not find any B-cell-related genes whose expression increased with emphysematous destruction in recently reported studies,<sup>8,9</sup> probably because the subjects were in relatively early stages of the disease. The genes with increased expression in the COPD without emphysema group might be associated with, or included among, the genes that increased

in the COPD with emphysema group, since small airway obstruction occurred first and emphysema followed. We performed two-way ANOVA and three-way ANOVA to assess the effects of current smoking on gene expression, because it was previously reported that gene expression is associated with smoking in the oxidative stress response, antiapoptosis, and cell death signaling and can be altered by smoking.<sup>41</sup> It is possible that genes with increased expression related to current smoking status are more likely to be genes related to emphysema or COPD. However, the overall results after correction for the smoking effect were similar to those from the one-way ANOVA. The total number of patients included in the study was small and the difference by the presence of COPD is evident, so smoking effects may not be reflected. Ultimately, our findings will require replication, ideally in additional large cohorts that include a range of COPD severities.

## Limitations

There are several limitations of this study. First, all subjects were men who currently smoked or were former smokers. Healthy smokers also included current or former smokers, and although they did not have emphysema or airflow limitation, it is possible that they had aberrant genetic modifications and expression. According to previous reports regarding sex differences in emphysema and airway disease, male smokers had more emphysema than female smokers, but female smokers did not show increased airway wall thickness compared with men.<sup>42,43</sup> However, to date, whether sex-specific genetic risk factors might underlie the differences in COPD susceptibility and severity is largely unknown. According to a recent genetic association study by genome-wide association study, a higher expression of CELSR1, a fetal lung development gene, was observed in female compared with male samples of humans in fetal lung tissue dataset.<sup>44</sup> Further studies are needed to explore sex-specific association in the development of emphysema or COPD. Additional gene expression studies that include never smokers and female subjects would improve the RNA-seq analysis of COPD. Second, we used normal-looking tissue adjacent to the lung cancer tissue for RNA-seq analysis. Lung tissue consists of many cell types, including macrophages, epithelial cells, and endothelial cells. Microdissection of lung tissues or single-cell sequencing would be required to determine whether the differential expression is present in all lung cells or only in a specific subset of cells. There might be possible gene expression patterns related to lung cancer. Among DEGs with significant *P*-values, there are several genes known to be associated with cancer, including lung cancer. PTPRG, which is thought to be a candidate tumor suppressor gene,<sup>45</sup> was one of the top ten

increased DEG in COPD without emphysema group. IFI27 gene, which is known to be associated with cancer cell apoptosis and lung carcinoma,<sup>46</sup> showed decreased expression in COPD with emphysema group (fifth in *P*-value). However, not only the number of known cancer-related genes is small, but also it is difficult to distinguish whether these genes are associated with emphysema or only cancer. Also, the additional results in the GSEA as well as the pathway analysis showed no relation with cancer gene ontology term.

Finally, it is difficult to determine whether the dysregulated pathways identified in this study are a cause or a consequence of the pathogenesis of emphysema or COPD. Experiments in which the increase/decrease in DEGs is reversed and is shown to slow disease progression are needed to confirm that these pathways are causally involved in the pathogenesis of COPD.

## Conclusion

Our study identified differences in gene expression in subjects with COPD according to emphysema status using RNA-Seq transcriptome analysis. These findings may have mechanistic implications in COPD.

## Acknowledgments

This work was supported by the Korean Health Technology R&D Project, Ministry of Health and Welfare, Republic of Korea (grant number HC16C2187) and the Global Research Development Center Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science and ICT (MSIT; NRF-2015K1A4A3046807). The biospecimens and data used in this study were provided by Asan Bio-Resource Center, Korea Biobank Network (2012-9[52]). The authors express their sincere gratitude to Dr Sang-min Lee (Department of Radiology, Asan Medical Center) for his assistance with the radiological analysis.

## Author contributions

Study design: YMO and JHK. Clinical data collection and primary analysis: IJ and DKO. Genetic data quality control and analysis: JHL. All authors made substantial contributions to interpretation of data; and took part in drafting the article or revising it critically for important intellectual content.

## Disclosure

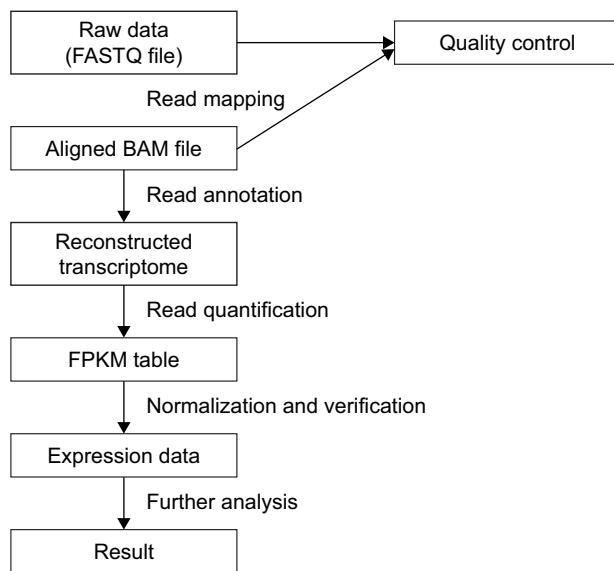
The authors report no conflicts of interest in this work.

## References

1. Global Initiative for Chronic Obstructive Lung Disease (GOLD). Global strategy for the diagnosis, management and prevention of chronic obstructive pulmonary disease; 2017. Available from: <http://goldcopd.org/>. Accessed February 10, 2017.

2. McDonough JE, Yuan R, Suzuki M, et al. Small-airway obstruction and emphysema in chronic obstructive pulmonary disease. *N Engl J Med.* 2011;365(17):1567–1575.
3. Haruna A, Muro S, Nakano Y, et al. CT scan findings of emphysema predict mortality in COPD. *Chest.* 2010;138(3):635–640.
4. Macnee W, Tudor RM. New paradigms in the pathogenesis of chronic obstructive pulmonary disease I. *Proc Am Thorac Soc.* 2009;6(6):527–531.
5. Løkke A, Lange P, Scharling H, Fabricius P, Vestbo J. Developing COPD: a 25 year follow up study of the general population. *Thorax.* 2006;61(11):935–939.
6. Cho MH, Castaldi PJ, Hersh CP, et al. A genome-wide association study of emphysema and airway quantitative imaging phenotypes. *Am J Respir Crit Care Med.* 2015;192(5):559–569.
7. Kim WJ, Lee CY. Environmental exposures and chronic obstructive pulmonary disease. *Mol Cell Toxicol.* 2017;13(3):251–255.
8. 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(11):1242–1253.
9. 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(8):67.
10. Morrow JD, Zhou X, Lao T, et al. Functional interactors of three genome-wide association study genes are differentially expressed in severe chronic obstructive pulmonary disease lung tissue. *Sci Rep.* 2017;7:44232.
11. Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat Methods.* 2008;5(7):621–628.
12. Ozsolak F, Milos PM. RNA sequencing: advances, challenges and opportunities. *Nat Rev Genet.* 2011;12(2):87–98.
13. Zhao S, Fung-Leung WP, Bittner A, Ngo K, Liu X. Comparison of RNA-Seq and microarray in transcriptome profiling of activated T cells. *PLoS One.* 2014;9(1):e78644.
14. Kim WJ, Lim JH, Hong Y, et al. Altered miRNA expression in lung tissues of patients with chronic obstructive pulmonary disease. *Mol Cell Toxicol.* 2017;13(2):207–212.
15. Kim WJ, Lim JH, Lee JS, Lee SD, Kim JH, Oh YM. Comprehensive analysis of transcriptome sequencing data in the lung tissues of COPD subjects. *Int J Genomics.* 2015;2015:206937.
16. Adcock IM, Ito K, Barnes PJ. Glucocorticoids: effects on gene transcription. *Proc Am Thorac Soc.* 2004;1(3):247–254.
17. Miller MR, Hankinson J, Brusasco V, et al. Standardisation of spirometry. *Eur Respir J.* 2005;26(2):319–338.
18. Makita H, Nasuhara Y, Nagai K, et al. Characterisation of phenotypes based on severity of emphysema in chronic obstructive pulmonary disease. *Thorax.* 2007;62(11):932–937.
19. Trapnell C, Roberts A, Goff L, et al. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat Protoc.* 2012;7(3):562–578.
20. Trapnell C, Pachter L, Salzberg SL. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics.* 2009;25(9):1105–1111.
21. Trapnell C, Williams BA, Pertea G, et al. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat Biotechnol.* 2010;28(5):511–515.
22. Coant N, Sakamoto W, Mao C, Hannun YA. Ceramidases, roles in sphingolipid metabolism and in health and disease. *Adv Biol Regul.* 2017;63:122–131.
23. Petracche I, Natarajan V, Zhen L, et al. Ceramide upregulation causes pulmonary cell apoptosis and emphysema-like disease in mice. *Nat Med.* 2005;11(5):491–498.
24. Mandal J, Malla B, Steffensen R, et al. Mannose-binding lectin protein and its association to clinical outcomes in COPD: a longitudinal study. *Respir Res.* 2015;16:150.
25. Tran HB, Ahern J, Hodge G, et al. Oxidative stress decreases functional airway mannose binding lectin in COPD. *PLoS One.* 2014;9(6):e98571.
26. Wain LV, Shrine N, Artigas MS, et al. Genome-wide association analyses for lung function and chronic obstructive pulmonary disease identify new loci and potential druggable targets. *Nat Genet.* 2017;49(3):416–425.
27. Selivanova PA, Kulikov ES, Kozina OV, et al. Differential expression of the β2-adrenoreceptor and M3-cholinoreceptor genes in bronchial mucosa of patients with asthma and chronic obstructive pulmonary disease. *Ann Allergy Asthma Immunol.* 2012;108(1):39–43.
28. Pilling D, Zheng Z, Vakil V, Gomer RH. Fibroblasts secrete Slit2 to inhibit fibrocyte differentiation and fibrosis. *Proc Natl Acad Sci U S A.* 2014;111(51):18291–18296.
29. Ito K, Ito M, Elliott WM, et al. Decreased histone deacetylase activity in chronic obstructive pulmonary disease. *N Engl J Med.* 2005;352(19):1967–1976.
30. Koeneke E, Witt O, Oehme I. HDAC family members intertwined in the regulation of autophagy: a druggable vulnerability in aggressive tumor entities. *Cells.* 2015;4(2):135–168.
31. Lee JH, Jeong EG, Choi MC, et al. Inhibition of histone deacetylase 10 induces thioredoxin-interacting protein and causes accumulation of reactive oxygen species in SNU-620 human gastric cancer cells. *Mol Cells.* 2010;30(2):107–112.
32. Revez JA, Matheson MC, Hui J, et al. Identification of STOML2 as a putative novel asthma risk gene associated with IL6R. *Allergy.* 2016;71(7):1020–1030.
33. Yokohori N, Aoshiba K, Nagai A; Respiratory Failure Research Group in Japan. Increased levels of cell death and proliferation in alveolar wall cells in patients with pulmonary emphysema. *Chest.* 2004;125(2):626–632.
34. Kantrowitz JJ, Kinose D, Vasilescu DM, et al. Transcriptomic alterations associated with emphysematous lung destruction in centrilobular and panlobular emphysema. *Am J Respir Crit Care Med.* 2016;193:A4075.
35. Puchelle E, Zahm JM, Tournier JM, Coraux C. Airway epithelial repair, regeneration, and remodeling after injury in chronic obstructive pulmonary disease. *Proc Am Thorac Soc.* 2006;3(8):726–733.
36. Gosselink JV, Hayashi S, Elliott WM, et al. Differential expression of tissue repair genes in the pathogenesis of chronic obstructive pulmonary disease. *Am J Respir Crit Care Med.* 2010;181(12):1329–1335.
37. Ning W, Li CJ, Kaminski N, et al. Comprehensive gene expression profiles reveal pathways related to the pathogenesis of chronic obstructive pulmonary disease. *Proc Natl Acad Sci U S A.* 2004;101(41):14895–14900.
38. Aravamudan B, Thompson MA, Pabelick CM, Prakash YS. Mitochondria in lung diseases. *Expert Rev Respir Med.* 2013;7(6):631–646.
39. Bialas AJ, Sitarek P, Milkowska-Dymanowska J, Piotrowski WJ, Górska P. The role of mitochondria and oxidative/antioxidative imbalance in pathobiology of chronic obstructive pulmonary disease. *Oxid Med Cell Longev.* 2016;2016:7808576.
40. Yamada Y, Tomaru U, Ishizu A, et al. Decreased proteasomal function accelerates cigarette smoke-induced pulmonary emphysema in mice. *Lab Invest.* 2015;95(6):625–634.
41. Charlesworth JC, Curran JE, Johnson MP, et al. Transcriptomic epidemiology of smoking: the effect of smoking on gene expression in lymphocytes. *BMC Med Genomics.* 2010;3:29.
42. Camp PG, Coxson HO, Levy RD, et al. Sex differences in emphysema and airway disease in smokers. *Chest.* 2009;136(6):1480–1488.
43. Grydeland TB, Dirksen A, Coxson HO, et al. Quantitative computed tomography: emphysema and airway wall thickness by sex, age and smoking. *Eur Respir J.* 2009;34(4):858–865.
44. Hardin M, Cho MH, Sharma S, et al. Sex-based genetic association study identifies CELSR1 as a possible chronic obstructive pulmonary disease risk locus among women. *Am J Respir Cell Mol Biol.* 2017;56(3):332–341.
45. Cheung AK, Ip JC, Chu AC, et al. PTPRG suppresses tumor growth and invasion via inhibition of Akt signaling in nasopharyngeal carcinoma. *Oncotarget.* 2015;6(15):13434–13447.
46. Cheon H, Borden EC, Stark GR. Interferons and their stimulated genes in the tumor microenvironment. *Semin Oncol.* 2014;41(2):156–173.

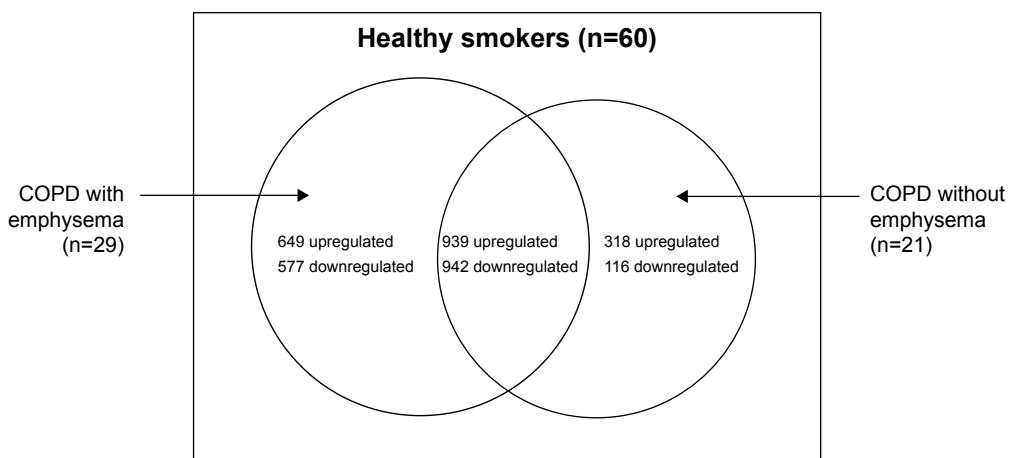
## Supplementary materials



**Figure S1** Schematic overview of the transcript analysis of the RNA-seq experiment.

**Notes:** Briefly, we used TopHat to align raw FASTQ files and used Cufflinks read annotation and quantification. FastQC was used to check read quality.

**Abbreviation:** FPKM, fragments per kilobase of exon per million fragments mapped.



**Figure S2** Schematic Venn diagram of the number of DEGs in the three subjects groups.

**Notes:** In the COPD with emphysema group, 1,588 genes were upregulated and 1,519 genes were downregulated, compared with those in healthy smokers. In the COPD without emphysema group, 1,257 genes were upregulated and 1,058 genes were downregulated, compared with those in healthy smokers. In both COPD groups, 939 genes were upregulated and 942 genes were downregulated, compared with those in healthy smokers.

**Abbreviation:** DEGs, differentially expressed genes.