ORIGINAL RESEARCH

Transcriptomics Analysis Identifies the Presence of Upregulated Ribosomal Housekeeping Genes in the Alveolar Macrophages of Patients with Smoking-Induced Chronic Obstructive Pulmonary Disease

Li Han [1]
Jing Wang [1]
Xiao-Bin Ji [1]
Zai-Yan Wang [1]
Yi Wang [2]
Li-Yue Zhang [2]
Hong-Peng Li [2]
Ze-Ming Zhang [1]
Qing-Yun Li [2]

[1]Department of Respiratory Medicine, Shanghai University of Medicine & Health Science Afiliated Zhoupu Hospital, Shanghai, People's Republic of China; [2]Department of Respiratory and Critical Care Medicine, Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, People's Republic of China

Background and Aims: Alveolar macrophages (AM) play a crucial role in the development of chronic obstructive pulmonary disease (COPD). The role that AM plays in the molecular pathways and clinical phenotypes associated with tobacco-related emphysema remain poorly understood. Thus, we investigated the transcriptomic profile of AM in COPD patients with a history of smoking and explored the molecular mechanisms associated with enriched pathways and hub genes.

Methods: Four data sets (GSE2125, GSE8823, GSE13896 and GSE130928) were retrieved from the GEO Database. A total of 203 GEO samples (GSM) were collated for this study. About 125 of these cases were classified as smokers (91 as healthy non-COPD smokers and 34 as COPD smokers). Based on the bioinformatics obtained using the R3.6.1 program, the data were successively adopted for differential genetic expression analysis, enrichment analysis (EA), and then protein–protein interaction analysis (PPI) in a STRING database. Finally, Cytoscape 3.8 software was used to screen the hub genes. A further data analysis was performed using a set of 154 cases, classified as 64 healthy non-smokers and 91 as healthy smokers. The same procedures were used as for the COPD dataset.

Results: When comparing the data pertaining to COPD-smokers and non-COPD smokers, the top ten genes with the greatest transcriptional differences were found to be NADK, DRAP1, DEDD, NONO, KLHL12, PRKAR1A, ITGAL, GLE1, SLC8A1, SVIL. A GSEA (Gene Set Enrichment Analysis) revealed that these genes manifested an up-regulated ribosomal pathway in contrast with other genes that exhibited an extensive down-regulated pathway. The hub genes were mainly genes encoding ribosomal subunits through PPI. Furthermore, it was found that there is a narrow transcriptional difference between healthy non-smokers and non-COPD smokers and the hub genes identified here are mainly members of the chemokines, including CCL5, CCR5, CXCL9 and CXCL11.

Conclusion: An elevated activity of the ribosome pathway in addition to the increased expression of ribosomal housekeeping genes (also known as hub genes) were identified with COPD-smokers, and these have the potential to cause a wide range of downstream pathogenetic effects. As for the preclinical phase, non-COPD smokers were found to be characterized by enriched pathways of several chemokines in AM.

Keywords: cigarette smoke, alveolar macrophages, transcriptomics, bioinformatics, chronic obstructive pulmonary disease

Correspondence: Ze-Ming Zhang
Department of Respiratory Medicine, Shanghai University of Medicine & Health Science affiliated Zhoupu Hospital, Shanghai, 201318, People's Republic of China
Email zhangzemingwang@sohu.com

Qing-Yun Li
Department of Respiratory and Critical Care Medicine, Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, 200025, People's Republic of China
Email liqingyun68@hotmail.com

Introduction

The prevalence of chronic obstructive pulmonary disease (COPD) is increasing worldwide, resulting in a higher mortality in patients with this disease. The global prevalence of COPD has been found to be 10.1%. In 2017 alone, this represented 3.2 million deaths.[1] It has become the third largest cause of death, and the fifth largest global medical burden.[1–3] Long-term smoking, from the consumption of tobacco, and biomass smoke, is the most common and major cause of COPD. As a heterogeneous and complex airway disease causing chronic inflammation, the pathogenesis of COPD is characterized by the irreversible limitation of airflow, accompanied with hypersecretion, remodeling and an impaired immune response of the mucous membrane. Proteolytic enzymes and reactive oxygen species contribute to tissue damage, and results in other debilitating conditions such as emphysema. The available data shows that only 15–20% of smokers develop severe COPD, but the mechanisms behind the development of COPD are yet to be elucidated.[3–5]

Alveolar macrophages (AM) participate in the phagocytosis and removal of foreign particles or pathogens and plays a vital role in maintaining alveolar structure and function, as well as the non-specific immunity in lung tissue. The impact of smoking on AM has been reported in several studies.[6–9] The transcriptomic alterations of AM were, however, only investigated in a few small sample studies.[5,10,11] Although COPD has been well known as chronic inflammation, the exact pathways or cytokines implied in the process of COPD are still not well understood. Thus, using bioinformatics we retrieved and sorted the datasets of similar studies available in the GEO database in the last 15 years, explored the transcriptomic influence of cigarette smoking exposure on COPD patients and healthy people, seeking the hub genes. Providing the greatest correlation in the relevant gene module or interaction network, hub genes are often key players in a number of clinical traits. The results will provide a greater insight and increased understanding of the pathogenesis and biological features associated with smoking-induced COPD at the transcription level.

Materials and Methods

Keywords such as "COPD", "smoking", "cigarette", "tobacco", "macrophage" were used as terms for searching the GEO database (https://www.ncbi.nlm.nih.gov/geo/), and the species was limited to "homo sapiens", preferring to select the same popular GPL platform, in accordance with case-control trials, until August 2020. A total of 4 datasets: GSE2125 (n=45), GSE8823 (n=24), GSE13896 (n=88), GSE130928 (n=70), representing a total of 203 cases of GSM, and all on an Affymetrix HG-U133 Plus 2.0 (GPL570 microarray), were obtained (Figure 1).

The total of 188 cases obtained (for smoking capacity calculation) were comprised of 63 non-smokers (22 female, 41 male) with an average age of 40.35±7.95 years, 91 non-COPD smokers (23 female, 68 male) with an average age of 43.68±8.56 years, and 34 COPD smokers (6 female, 28 male) with an average age of 54.25 ±8.35 years. A difference in smoking history existed within the COPD and non-COPD smokers, where non-COPD smokers consumed 26.91±17.20 pack-years (n=76), which was less than that of the COPD smokers, who consumed 45.40±27.87 pack-years (n=34, P<0.01), which is determined using the t-test, with the exception of 15 smokers in GSE2125 that had no information concerning their smoking history. AM were collected by bronchoalveolar lavage (BAL), as previously described.[12]

One hundred and twenty-five cases of GSM were collated as the smoking population, and the healthy smokers (non-COPD, n=91) were compared with the COPD smokers (n=34). After data cleaning using R 3.6.1 software, irrelevant or duplicate data were excluded (some data within the GSE13896 set came from GSE8823, and GSE2125 included some asthmatic data) and 54,675 probes were obtained in total. The combined expression matrix, grouping information and GPL annotation information were also extracted. Following the process of the batch effect (Figure 2), analysis and visualization of the differentially expressed genes (DEGs), including a PCA (principal component analysis) chart, heat-map and volcano plot were performed using "means" of the corresponding R package "limma"; The above data then underwent enrichment analysis by Gene Ontology (GO) and Gene Set Enrichment Analysis (GSEA). The interacting genes were retrieved by STRING platform (www.string-db.org/) based on DEGs and the protein–protein interaction (PPI), which was constructed to screen for hub genes using Cytoscape 3.8.1 software. A gene possessing high degrees in a PPI network would be defined as a hub gene in the module. The hub genes of interest were also identified and analyzed in comparison using the t-test. In addition, 155 cases of GSM were collated as the non-COPD population, the healthy control (non-smokers,
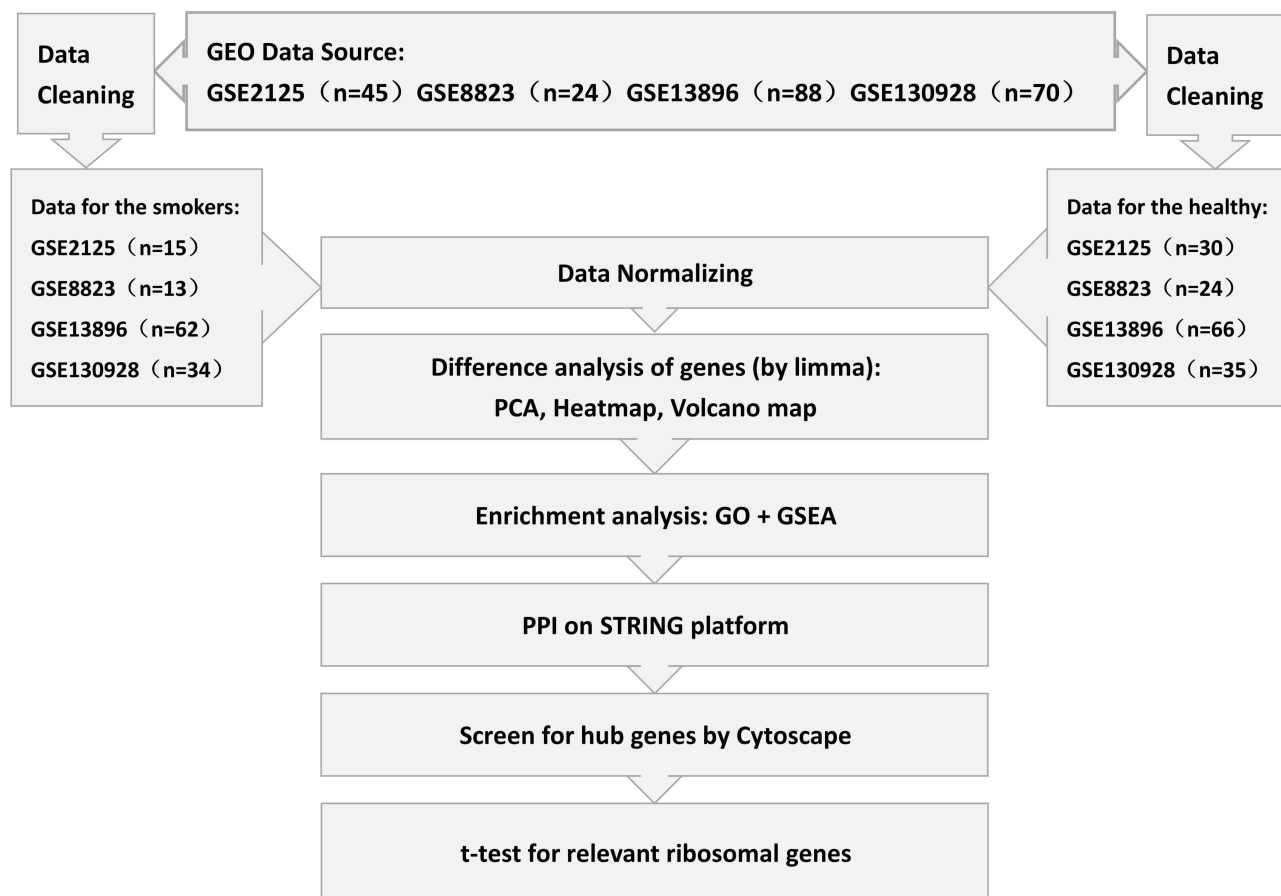
Figure 1 Flow Chart of Data Processing for 4 GSEs.

n=64) versus the healthy smokers (non-COPD, n=91). The dataset was processed using the same analysis process described above (Figure 1).

R package "limma", which is usually used for the analysis of a gene expression microarray or sequencing data using a generalized linear model and empirical Bayes method, can fit a linear model to the expression data for each gene and "retrain" the false positives, making the overall analysis stable.[13] GO enrichment analysis is designed to calculate the P value of each pathway (pre-defined according to GO) with a differential gene expression, which is part of an over-representation analysis based on hypergeometric distribution.[14] As a computational method, GESA is used to determine whether a priori defined set of genes (obtained from KEGG) shows statistically significant, concordant differences between two biological states. GSEA first ranks all the genes according to their various expression by means of functional class scoring (based on a Kolmogorov–Smirnov test). The ranked list is then used to assess how the genes within

each gene set are distributed across the entire ranked list.[15] Adjusted P<0.05 and |log FC| (fold change) > 1.5 were considered to be statistically significant.

Results

Principal Component Analysis (PCA)

A PCA highlighted that there was a significant difference between the genes expressed by the COPD smokers and non-COPD smokers (Figure 2B), whereas there was only a limited difference between the genes expressed by the non-smokers and non-COPD smokers (Figure 2D).

Differential Gene Expression

When comparing the COPD smokers and non-COPD smokers, the top ten significant gene differences (sorted by adj. P<0.05) were: NADK, DRAP1, DEDD, NONO, KLHL12, PRKAR1A, ITGAL, GLE1, SLC8A1 and SVIL (Table 1). Most of these are down-regulated genes, with KPNA2 being the only up-regulated genes (in the top 20, not shown). When comparing the non-COPD smokers and
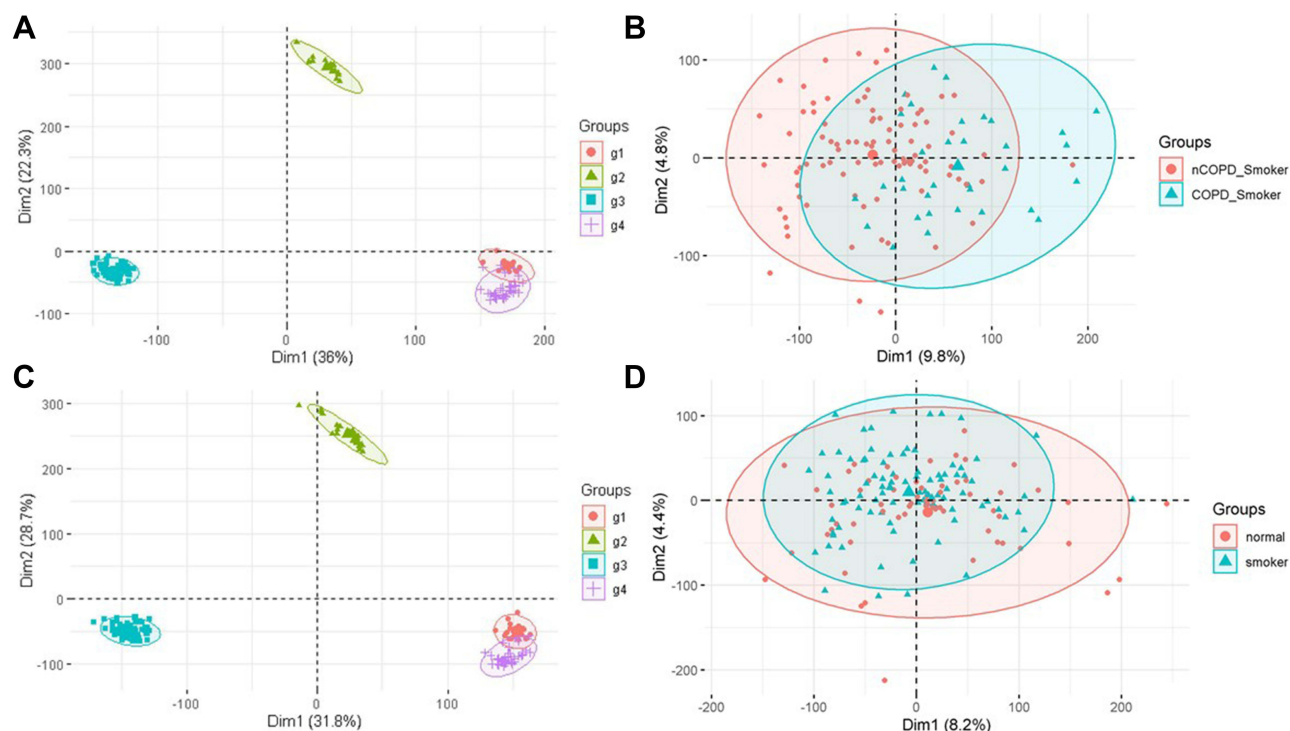
Figure 2 PCA of 4 GSEs before and after Batch Effect Adjustment. The 4 datasets unprocessed (against Batch Effect) are shown divergently by means of PCA (**A**); After the adjustment allowing for the Batch Effect, the datasets are integrated by the COPD smokers and the non-COPD smokers and the inherent difference can be sharpened (**B**). The colored dots represent the probes corresponding to relevant genes symbolically and quite a number of genes are differently distributed between the groups (**B**). The identical process of PCA was applied to the non-COPD groups with the unprocessed data shown (**C**). As the normalized data is displayed, there is only a limited but definite difference of gene expression between the non-smokers and the non-COPD smokers (**D**).

non-smokers, the top ten genes with significant differences (sorted by adj.P<0.05) were: VGLL3, SLC26A11, NOL4L, ATP6V0D2, ITGAE, CADPS2, LPAR3, PDE4B, IFI27 and SEMA4D (Table 1). The five most distinct down-regulated genes were VGLL3, NOL4L, CADPS2, LPAR3 and PDE4B, while the three most significant up-regulated genes were SLC26A11, ATP6V0D2 and ITGAE. With the default logFC (0.585) and adj.P (<0.05), the differently expressed genes have been visualized in the heatmap and volcano plot presented in Figure 3.

The Enrichment Analysis

For the COPD and non-COPD smokers, a GO enrichment analysis shows that there is a remarkable enrichment of cell components, biological processes and molecular functions. Cellular components include intercellular and cytoplasmic adhesion components, lysosomal membranes and other structures; Biological processes include neutrophil activation, migration, degranulation and related immune processes. Molecular functions consist mainly of binding

entities such as cell adhesion molecules (CAM) including cadherin and others (Table 2, Figure 4A–C).

It was also found that more than 120 enriched pathways were acquired by GSEA (P<0.05, minGeneset Size: 20, maxGeneset size: 200). Ribosome and RNA transport are the only two up-regulated pathways (red-colored in Figure 5A) in contrast with the majority of down-regulated pathways. The latter includes processes such as the inflammation signaling pathways of various viral, bacterial and even protozoal infections, pinocytosis processes, cytoskeleton and connection structures, cellular adhesion molecules (CAMs) and chemokine signaling pathways, pattern recognition receptor (PRR) signaling including NLR and TLR, NF-κB pathways; PI3K-Akt pathways, mTOR signaling; phospholipase D signaling, TNF signaling; lysosome and phagosome, cell senescence and necroptosis, osteoclast differentiation; a variety of carcinomas and related processes including relevant misregulating transcription and microRNAs, viral carcinogenesis, PD-L1 and PD-1 pathways, HER2 signaling, tumor proteoglycans; EGFR and VEGF pathways; insulin-related signaling including relaxin and FoxO signaling as well as

Table I The Top 10 Genes with Most Significant Difference of Transcription Between Relevant Groups

Group	Entrez ID	Symbol	Description	Up/Down	logFC	Adj.P
The COPD vs non-COPD smokers	65220	NADK	NAD kinase	Down	−1.016090856	1.18E-09
	10589	DRAP1	DR1 Associated Protein 1	Down	−0.70553346	2.40E-09
	9191	DEDD	Death Effector Domain Containing	Down	−0.834479281	3.74E-09
	4841	NONO	Non-POU Domain Containing Octamer Binding	Down	−0.751792587	6.67E-09
	59349	KLHL12	Kelch Like Family Member 12	Down	−0.611660528	6.67E-09
	5573	PRKAR1A	Protein Kinase CAMP-Dependent Type 1 Regulatory Subunit Alpha	Down	−0.881787568	6.67E-09
	3683	ITGAL	Integrin Subunit Alpha L	Down	−0.939391902	6.67E-09
	2733	GLE1	GLE1 RNA Export Mediator	Down	−0.762754132	7.34E-09
	6546	SLC8A1	Solute Carrier Family 8 Member A1	Down	−0.860251027	1.54E-08
	6840	SVIL	Supervillin	Down	−0.839944317	1.87E-08
The non-COPD smokers vs non-smokers	227399	VGLL3	Vestigial Like Family Member 3	Down	−2.194799911	1.49E-22
	226679	SLC26A11	Solute Carrier Family 26 Member 11	Up	0.980624248	2.78E-19
	225224	NOL4L	Nucleolar Protein 4 Like	Down	−0.840477441	2.36E-18
	1553151	ATP6V0D2	ATPase H+ Transporting V0 Submit D2	Up	1.759028091	1.37E-16
	205055	ITGAE	Integrin Subunit Alpha E	Up	0.864863449	1.63E-16
	219572	CADPS2	Calcium Dependent Secretion Activator 2	Down	−1.635076486	2.05E-16
	231192	LPAR3	Lysophosphatidic Acid Receptor 3	Down	−1.818363959	2.05E-16
	203708	PDE4B	Phosphodiesterase 4B	Down	−1.918284538	2.24E-16
	202411	IFI27	Interferon Alpha Inducible Protein 27	Down	−2.107723118	2.90E-16
	203528	SEMA4D	Semaphorin 4D	Down	−0.954989951	4.82E-16

RAGE-AGE signaling, Type II diabetes processes and carbohydrate metabolism; adrenergic signaling, hypertrophic and dilated cardiomyopathy; JAK-STAT signaling and hematopoietic lineage; T and NK lymphocyte activation and differentiation, BCR and FcεRI signaling; leukocyte migration, platelet activation; atherosclerosis process; adipocyte metabolism; neurotrophin signaling and cholinergic synapses; prolactin, growth hormone, parathyroid and thyroxine processes; Hippo signaling, sphingolipid signaling, HIF-1 signaling, Notch signaling, Hedgehog signaling, etc. As in Figure 5A, only pathways with P<0.0005 are displayed for the figure size.

For the non-COPD smokers and non-smokers (ie, the healthy control), the GO enrichment analysis highlighted evident enrichment on molecular function (MF) and biological processes (BP). The cytokine activity and its receptor binding pathways were found to be the most significant processes for molecular functions. As for biological processes, leukocyte migration, inflammatory cells releasing chemokines, and non-specific immune responses to pathogen components, such as LPS are the most prevalent (Table 2, Figure 4D–F). Likewise, using GSEA, 50 enrichment pathways were acquired (P<0.05, minGeneset Size: 20, maxGeneset size: 200) with quite a proportion of them being analogous to the above, include the relevant

pathways or mechanisms of viral or bacterial or protozoal infections, T lymphocyte differentiation and activation pathways, and other various steps associated with inflammatory processes (Figure 5B).

Protein–Protein Interaction (PPI) Analysis

The PPI analysis of genetic differential expression between the COPD and non-COPD smokers highlighted that ribosomal protein-coding genes, such as RPS11, RPL14, RPL27A, RPS27L, RPL37A and RPL38, are the hub genes, which are also related to PPP2R1A, EIF4B, TUBB, SMURF2, ACTB, and are closely connected to each in a network (Figure 6A). All of the ribosomal genes (above) in COPD smokers have been remarkably lifted, as shown in Figure 6C, with the most significant being RPL38. The result is in accordance with the up-regulated tendency in GSEA.

As for the non-COPD smokers and the non-smokers, the PPI analysis highlighted that the hub genes present were dominantly CCL5, CCR5, CXCL9 and CXCL11 and other chemokines (or receptors) of the network, which are related to ITGAE, LPAR3, EMR1, GBP2/5 and other G protein coupling or related components, as well as CD69/80 and other T lymphocyte activation receptors and TNF superfamily members. Out of these, the most
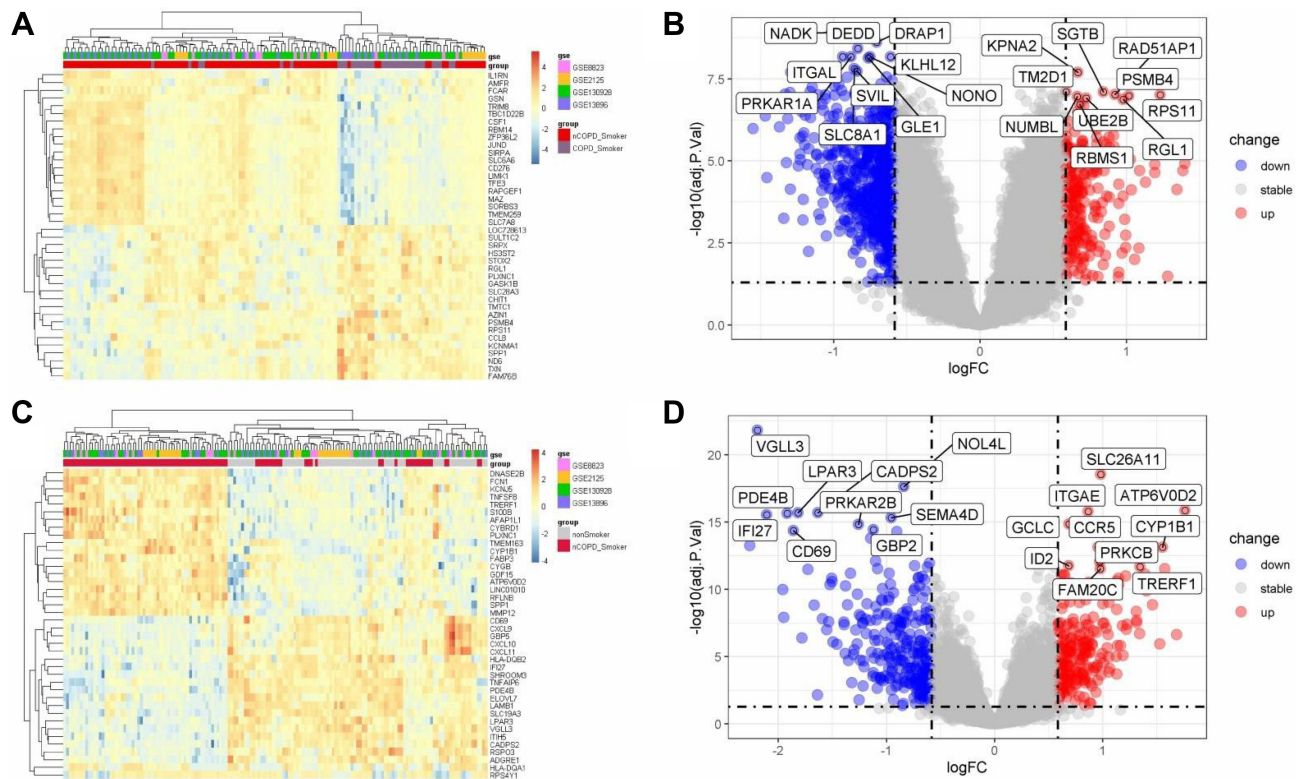
Figure 3 The Visualization of Differently Expressed Genes Between Relevant Groups. According to logFC, the Heatmap (**A**) displays the top 40 of differential genes expressed (both upregulated and downregulated) between the COPD smokers and the non-COPD Smokers. With the default logFC (0.585) and P value (0.05), the volcano plot (**B**) shows the top 20 differential expression of relevant genes. Likewise, the Heatmap (**C**) displays the top 40 differential genes upregulated or downregulated between the non-COPD smokers and the non-Smokers. With the same default of logFC and P value, the volcano plot (**D**) shows the top 20 differential expression of relevant genes.

prominent were the aforementioned chemokines ([Figure 6B](#)).

Discussion

Although it has been well-recognized that cigarette smoke is the most common cause of COPD, the specific mechanisms responsible have not yet been elucidated.[16,17] This study analyzed the last 15 years of datasets (until August 2020) available from the GEO database, and found that there was an extensive and significant degree of gene differential expression of AM between the COPD smoking and non-COPD smoking group. It was found that people with COPD have a number of significantly differently expressed genes related to energy metabolism or oxidative stress than those without COPD. For example, NADK catalyzes the phosphorylation of nicotinamide adenine dinucleotide (NAD) into NADP. The latter is widely considered as an electron donor in the process of biosynthesis, especially the most basic oxidative phosphorylation of cells and oxidative stress. SLC8A1 encoding protein NCX can affect the process of oxidative phosphorylation by regulating the Ca^{2+} concentration in mitochondria.[18]

GSEA manifests a definite trend of up-regulation of the ribosomal pathway as well as a substantial down-regulated trend of most other pathways, which is coincident with the transcriptional down-regulation of massive genes. A PPI analysis indicated that quite a few hub genes are encoding proteins of ribosomal subunits, also known as housekeeping genes. It was also found that the expression of ribosomal genes in COPD smokers were all significantly elevated ([Figure 6C](#), sorted by adj.P<0.05), which suggests that the process of protein synthesis has been affected. Nevertheless, the results obtained using the GO enrichment analysis were found to be dissimilar to the results obtained using GSEA. GO enrichment analysis is based on differential gene expression, whereas GSEA is designed to be not limited in the DEGs and could reveal whether the accumulation of subtle changes in a priori defined set of genes could be statistically significant and reflect the difference between the two biological states or phenotypes.

Studies have found that the AMs in smokers lose a certain degree of their pro-inflammatory properties, such as phagocytosis, their release of inflammatory mediators, and reduction of NO, which highlights the

Table 2 The Top 10 of Most Significant Pathways Between Relevant Groups

Group	GO: ID	Description	Gene Ratio	Pathway	q value	Adj. P
The COPD vs non-COPD smokers	0005924	Cell-substrate adherens junction	64/913	CC	2.37E-15	2.98E-15
	0005925	Focal adhesion	64/913	CC	2.37E-15	2.98E-15
	0030055	Cell-substrate junction	64/913	CC	2.61E-15	3.28E-15
	0050839	Cell adhesion molecule binding	60/885	MF	3.28E-07	3.55E-07
	0002446	Neutrophil mediated immunity	57/885	BP	1.04E-06	1.26E-06
	0043312	Neutrophil degranulation	56/885	BP	1.04E-06	1.26E-06
	0042119	Neutrophil activation	57/885	BP	1.04E-06	1.26E-06
	0002283	Neutrophil activation involved in immune response	57/885	BP	1.04E-06	1.26E-06
	0045296	Cadherin binding	44/885	MF	2.02E-06	2.19E-06
	0098852	Lytic vacuole membrane	41/913	CC	8.27E-06	1.04E-05
The non-COPD smokers vs non-smokers	0097529	Myeloid leukocyte migration	29/447	BP	1.03E-10	1.38E-10
	0005125	Cytokine activity	29/446	MF	2.37E-10	2.91E-10
	0030595	Leukocyte chemotaxis	29/447	BP	2.80E-10	3.76E-10
	0050900	Leukocyte migration	42/447	BP	2.12E-09	2.84E-09
	0060326	Cell chemotaxis	32/447	BP	2.38E-09	3.19E-09
	0071621	Granulocyte chemotaxis	20/447	BP	1.09E-08	1.46E-08
	0097530	Granulocyte migration	21/447	BP	1.35E-08	1.81E-08
	0032496	Response to lipopolysaccharide	32/447	BP	1.35E-08	1.81E-08
	0002237	Response to molecule of bacterial origin	32/447	BP	2.93E-08	3.93E-08
	0005126	Cytokine receptor binding	29/446	MF	8.20E-08	1.01E-07

Abbreviations: CC, Cellular Components; MF, Molecular Function; BP, Biological Process.

imbalance between the so-called M1 pro-inflammatory phenotype and the variant M2 anti-inflammatory phenotype.[7,8] This imbalance is in accordance with the results of this study, which highlights the presence of vast numbers of down-regulated pathways. This imbalance not only raises the instances of mortality through certain pathogenic infections, but also enhances the occurrence and development of tumors.[4] Several transcriptomics studies of AM have focused on COPD.[5,11,19] An earlier study found that the expression of osteopontin (encoded by the gene SPP1) was up-regulated (4 times, P=0.006) in the AMs of healthy smokers, with significant differences in CCR5, CD69 and ITIH5. This also suggest that AM is activated by exposure to smoking, however the sample size was limited.[10] Another study found that MERTK, the receptor responsible for recognizing apoptotic bodies on AM surfaces of healthy asymptomatic smokers, was significantly augmented, even more in patients with mild or moderate COPD.[11] Using an RNA sequencing method, Morrow et al showed the enrichment of amyloid and telomere-related pathways across the three tissues (AM, peripheral blood and airway epithelium) in a study comparing 11 COPD smokers and 6 non-COPD smokers. In this study, the SCGB1A1 (secretoglobin family 1A

member 1), which encodes for CC16 (Club Cell Secretory Protein), a blood biomarker of COPD, was differently expressed,[20] which as well implied somewhat systemic effect of COPD.

Shaykhiev et al obtained a different result, whereby M1 inflammation-related chemokines, such as CXCL9, CXCL10, CXCL11 and CCL5 (obtained by analysis of the AMs of healthy asymptomatic smokers), were relatively down-regulated, while M2-related species such as MMP2, MMP9 and ADORA3 were up-regulated.[5] A later study, which adopted a multi-factor mixed method to stimulate AM (multiple cytokines and different combinations), not only confirmed the existence of the M1 or M2 polarized inflammation phenotypes, but illustrated there were even up to 9 different expression patterns of the inflammatory response existing, highlighting the complexity of chronic inflammation.[21] Not only an increased number of AM of patients with COPD, but the plasticity of AM to change phenotypes, have been already discovered. Different phenotypes and the distribution of AM largely depend on their microenvironment, including various inflammatory mediators or cytokines and even the loci of small airways.[22] In another research, it was found that AMs in patients exposed to smoking
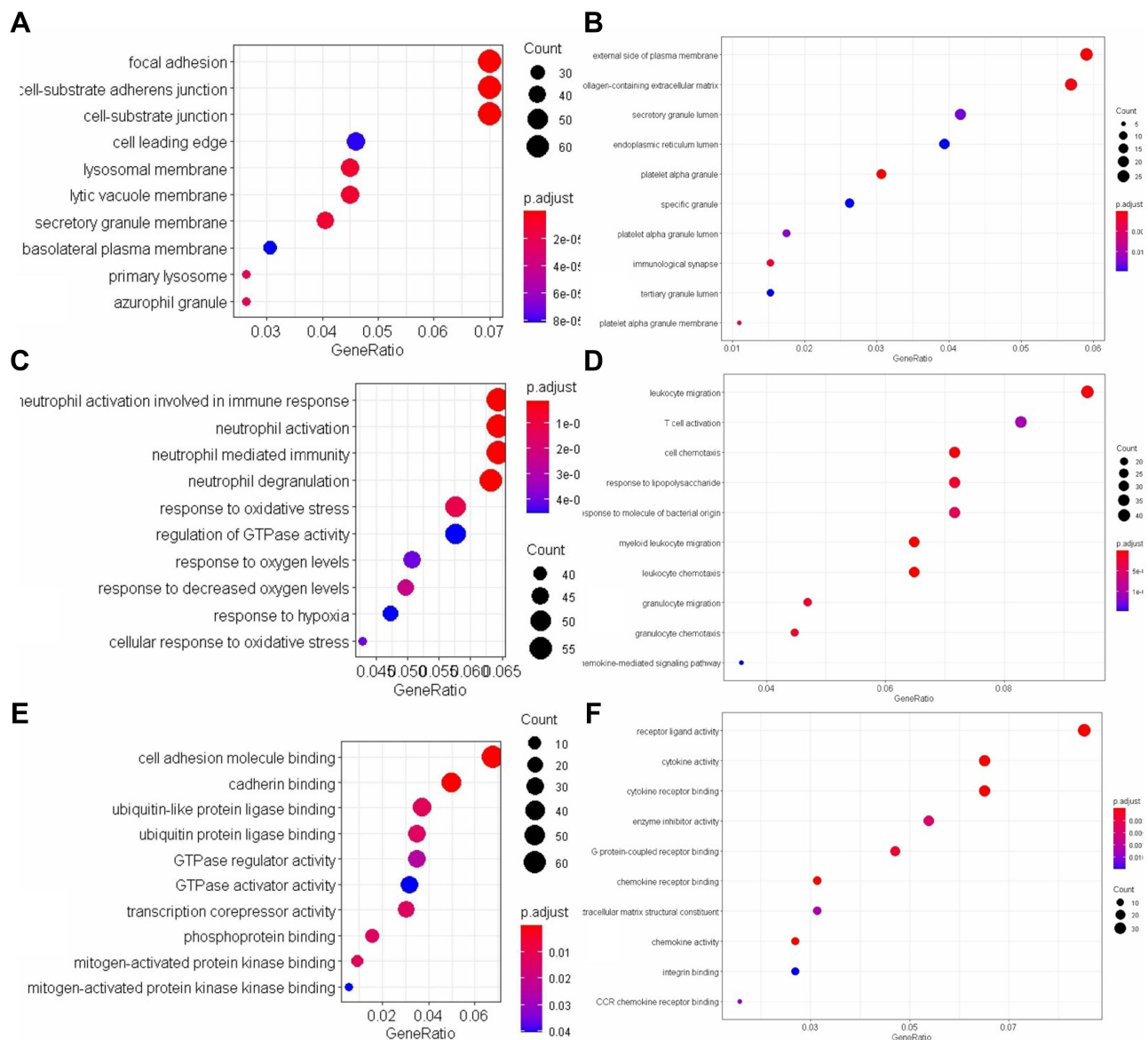
Figure 4 GO Enrichment Analysis of AM Between Relevant Groups. (**A–C**) GO comparison of the COPD Smokers and the non-COPD Smokers; (**A**) for Cell Components (Top 10), (**B**) for Biological Process (Top 10), (**C**) for Molecular Function (Top 10). (**D–F**) GO comparison of the non-COPD Smokers and the non-Smokers; (**D**) for Cell Components (Top 10), (**E**) for Biological Process (Top 10), (**F**) for Molecular Function (Top 10).

could simultaneously express both markers of M1 (iNOS) and M2 (arginase), while dual positive (both M1 and M2 marker-positive) macrophages were more frequently distributed in lung tissues of subjects with severe COPD. These results allow the postulation that AM populations with M1 and M2 signatures do not exclude each other, but can often coexist.[23]

Our study further compared the transcriptomic profile of AM between the non-smokers (the healthy control) and the non-COPD smokers (the healthy smokers). A PPI analysis indicated that CCL5, CCR5, CXCL9 and CXCL11 were the hub genes. For COPD patients with a smoking history, the

levels of chemokines, such as IL-6, CCL26 of blood and IL8, CXCL1, CCL5 of AM had increased significantly, which could also be correlated with the degree of airflow limitation.[24–26] The results obtained through the PPI and enrichment analysis performed in this study suggests that related chemokines could play a crucial role, which is consistent with previous studies and can therefore be regarded as a verification of the bioinformatic analysis contained within this paper.

The main challenge with this work is that the relevant differential genes or hub genes determined as being significant in this study have not yet been verified in the
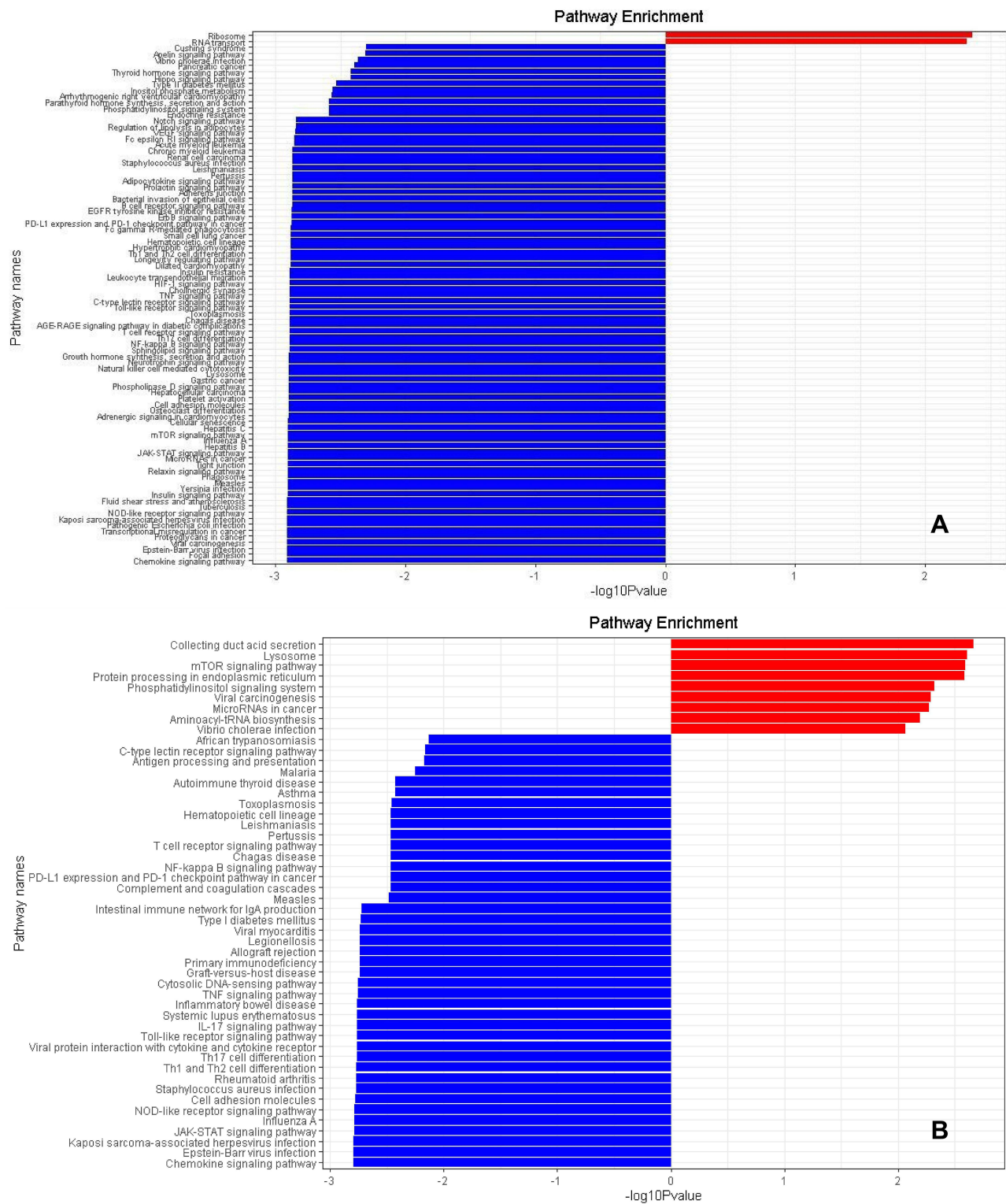
Figure 5 GSEA of AM Between Relevant Groups. (**A**) GESA for the non-COPD Smokers vs COPD Smokers. The blue-colored part on the left represents the relatively down-regulated pathways according to the enrichment analysis by GSEA, while a few of the red-colored parts represent the relatively up-regulated pathways. The default cutoff of P value is 0.05, with a total of more than 120 pathways being achieved, thus only pathways with P value <0.005 are shown here. (**B**) GSEA for the non-Smokers vs non-COPD Smokers. The blue-colored part on the left represents the relatively down-regulated pathways according to the enrichment analysis by GSEA, while a few of the red-colored part represent the relatively up-regulated pathways. The default cutoff of P value is 0.05, with a total of 50 pathways being achieved. The trend of up- or down-regulated is based on the algorithm of functional class scoring (FCS).
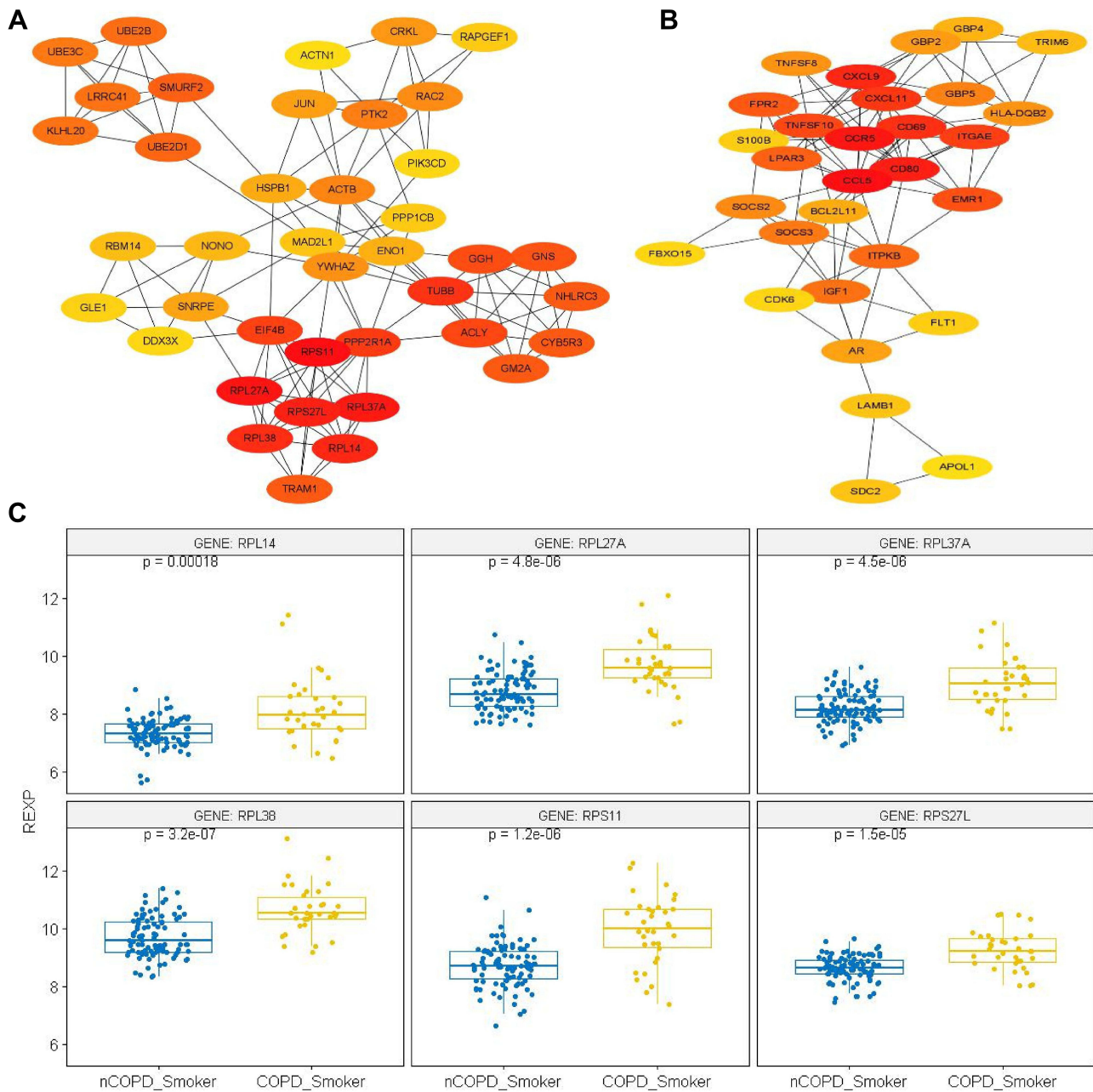
Figure 6 Protein–Protein Interaction Network and Hub gene Expression. All of the nodes could be regarded as hub-genes of the relevant network in (**A** and **B**). (**A**) PPI comparison of the Cytoscape between the COPD and non-COPD Smokers. The Hub genes have been displayed as the color nodes. The most deep-colored are RPL27A, RPL14, RPL37A, RPL38, RPS11 and RPS27L, signifying the most concentrated genes of the network. (**B**) PPI comparison of the non-COPD smokers and non-smokers. The most deep-colored are CCL5, CCR5, CXCL9 and CXCL11, signifying the most concentrated genes of the network. (**C**) The expression of ribosome-encoding genes such as RPS11, RPL14, RPL27A, RPS27L, RPL37A and RPL38 between the COPD and non-COPD Smokers. All of these genes in COPD Smokers have been remarkably up-regulated, with RPL38 being most significant, with a P value of 3.2×10^{-7}.

laboratory. The sample size of the smoking COPD patients was relatively small compared to that of the non-smoking COPD patients (34:91) and this may have lead to a statistical bias, however the results obtained are consistent with a number of other similar studies. Though ICS (inhaled corticosteroid) usage has an impact to alter AM

phenotype and function,[27] there is no relevant information provided within the data set.

In conclusion, it is clear that continuous exposure to cigarette smoking could trigger the onset of COPD. This study has revealed the comprehensive trend of repressed pathways and elevated ribosome proteins, each associated

with the development of the disease. A large number of hub genes are up-regulated ribosomal genes, suggesting that exposure to cigarettes could adversely affect the synthesis of relevant proteins, and hence increase the subsequent pathogenic risks.

Acknowledgments

Disclosure

The authors declare that they have no conflicts of interest in this work.

References

1. Celli BR, Wedzicha JA. Update on clinical aspects of chronic obstructive pulmonary disease. *N Engl J Med*. 2019;381(13):1257–1266. PMID: 31553837. doi:10.1056/NEJMra1900500
2. Sullivan J, Pravosud V, Mannino DM, et al. National and state estimates of COPD morbidity and mortality - United States, 2014–2015. *Chronic Obstr Pulm Dis*. 2018;5(4):324–333. PMID: 30723788. doi:10.15326/jcopdf.5.4.2018.0157
3. Barmes PJ, Burney PG, Silverman EK, et al. Chronic obstructive pulmonary disease. *Nat Rev Dis Primers*. 2015;1:15076. PMID: 27189863. doi:10.1038/nrdp.2015.76
4. Cosio MG, Saetta M. Evasion of COPD in smokers: at what price? *Eur Respir J*. 2012;39(6):1298–1303. doi:10.1183/09031936.00135711
5. Shaykhiev R, Krause A, Salit J, et al. Smoking-dependent reprogramming of alveolar macrophage polarization: implication for pathogenesis of chronic obstructive pulmonary disease. *J Immunol*. 2009;183(4):2867–2883. PMID: 19635926. doi:10.4049/jimmunol.0900473
6. Barnawi J, Tran H, Jersmann H, et al. Potential Link between the Sphingosine-1-phosphate (S1P) system and defective alveolar macrophage phagocytic function in chronic obstructive pulmonary disease (COPD). *PLoS One*. 2015;10(10):e0122771. PMID: 26485657. doi:10.1371/journal.pone.0122771
7. Bewley MA, Preston JA, Mohasin M, et al. Impaired mitochondrial microbicidal responses in chronic obstructive pulmonary disease macrophages. *Am J Respir Crit Care Med*. 2017;196(7):845–855. PMID: 28557543. doi:10.1164/rccm.201608-1714OC
8. Dewhurst JA, Lea S, Hardaker E, et al. Characterisation of lung macrophage subpopulations in COPD patients and controls. *Sci Rep*. 2017;7(1):7143. PMID: 28769058. doi:10.1038/s41598-017-07101-2
9. Gleeson LE, O'Leary SM, Ryan D, et al. Cigarette smoking impairs the bioenergetic immune response to mycobacterium tuberculosis infection. *Am J Respir Cell Mol Biol*. 2018;59(5):572–579. PMID: 29944387. doi:10.1165/rcmb.2018-0162OC
10. Woodruff PG, Koth LL, Yang YH, et al. A distinctive alveolar macrophage activation state induced by cigarette smoking. *Am J Respir Crit Care Med*. 2005;172(11):1383–1392. PMID: 16166618. doi:10.1164/rccm.200505-686OC
11. Kazeros A, Harvey BG, Carolan BJ, et al. Overexpression of apoptotic cell removal receptor MERTK in alveolar macrophages of cigarette smokers. *Am J Respir Cell Mol Biol*. 2008;39(6):747–757. PMID: 18587056. doi:10.1165/rcmb.2007-0306OC
12. Russi TJ, Crystal RG. Use of bronchoalveolar lavage and airway brushing to investigate the human lung. In: Crystal RG, West JB, Weibel ER, Barnes PJ, editors. *The Lung: Scientific Foundations*. 2nd ed. Philadelphia: Lippincott-Raven Publishers; 1997:371–382.
13. Phipson B, Lee S, Majewski IJ, et al. Robust hyperparameter estimation protects against hypervariable genes and improves power to detect differential expression. *Ann Appl Stat*. 2016;10(2):946–963. PMID: 28367255. doi:10.1214/16-AOAS920
14. Boyle EI, Weng S, Gollub J, et al. GO::TermFinder–open source software for accessing Gene Ontology information and finding significantly enriched Gene Ontology terms associated with a list of genes. *Bioinformatics*. 2004;20(18):3710–3715. PMID: 15297299. doi:10.1093/bioinformatics/bth456
15. Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *PNAS*. 2005;102(43):15545–15550. PMID: 16199517. doi:10.1073/pnas.0506580102
16. Xu Y, Hu B, Alnajm SS, et al. SEGEL: a web server for visualization of smoking effects on human lung gene expression. *PLoS One*. 2015;10(5):e0128326. PMID: 26010234. doi:10.1371/journal.pone.0128326
17. Park EJ, Park YJ, Lee SJ, et al. Cigarette smoke extract may induce lysosomal storage disease-like adverse health effects. *J Appl Toxicol*. 2019;39(3):510–524. PMID: 30485468. doi:10.1002/jat.3744
18. Khananshvil D. The SLC8 gene family of sodium–calcium exchangers (NCX) – structure, function, and regulation in health and disease. *Mol Aspects Med*. 2013;24(2–3):220–235. PMID: 23506867. doi:10.1016/j.mam.2012.07.003
19. O'Beirne SL, Kikkers SA, Oromendia C, et al. Alveolar macrophage immunometabolism and lung function impairment in smoking and chronic obstructive pulmonary disease. *Am J Respir Crit Care Med*. 2020;201(6):735–739. PMID: 31751151. doi:10.1164/rccm.201908-1683LE
20. Morrow JD, Chase RP, Parker MM, et al. RNA-sequencing across three matched tissues reveals shared and tissue-specific gene expression and pathway signatures of COPD. *Respir Res*. 2019;20(1):65. PMID: 30940135. doi:10.1186/s12931-019-1032-z
21. Xue J, Schmidt SV, Sander J, et al. Transcriptome-based network analysis reveals a spectrum model of human macrophage activation. *Immunity*. 2014;40(2):274–288. PMID: 24530056. doi:10.1016/j.immuni.2014.01.006
22. Yamasaki K, Eeden SFV. Lung macrophage phenotypes and functional responses: role in the pathogenesis of COPD. *Int J Mol Sci*. 2018;19(2):582. PMID: 29462886. doi:10.3390/ijms19020582
23. Bazzan E, Turato G, Tinè M, et al. Dual polarization of human alveolar macrophages progressively increases with smoking and COPD severity. *Respir Res*. 2017;18(1):40. PMID: 28231829. doi:10.1186/s12931-017-0522-0
24. Bradford E, Jacobson S, Varasteh J, et al. The value of blood cytokines and chemokines in assessing COPD. *Respir Res*. 2017;18(1):180. PMID: 29065892. doi:10.1186/s12931-017-0662-2
25. Kaur M, Singh D. Neutrophil chemotaxis caused by chronic obstructive pulmonary disease alveolar macrophages: the role of CXCL8 and the receptors CXCR1/CXCR2. *J Pharmacol Exp Ther*. 2013;347(1):173–180. PMID: 23912333. doi:10.1124/jpet.112.201855

26. Lea S, Plumb J, Metcalfe H, et al. The effect of peroxisome prolif-erator-activated receptor-γ ligands on in vitro and in vivo models of COPD. *Eur Respir J.* 2014;43(2):409–420. PMID: 23794466. doi:10.1183/09031936.00187812

27. Higham A, Scott T, Li J, et al. Effects of corticosteroids on COPD lung macrophage phenotype and function. *Clin Sci (Lond).* 2020;134 (7):751–763. PMID: 32227160. doi:10.1042/CS20191202