


Identification of differentially expressed genes and signaling pathways in chronic obstructive pulmonary disease via bioinformatic analysis

Xinwei Huang^{1,2}, Yunwei Li^{2,3}, Xiaoran Guo², Zongxin Zhu², Xiangyang Kong² , Fubing Yu⁴ and Qiang Wang⁵

1 Faculty of Environmental Science and Engineering, Kunming University of Science and Technology, China

2 Medical School, Kunming University of Science and Technology, China

3 Department of Pharmacy, Kunming Children's Hospital, China

4 Department of Gastroenterology, Fourth Affiliated Hospital of Kunming Medical University, China

5 Physical Examination Center, Second People's Hospital of Yunnan Province, Kunming, China

Keywords

bioinformatic analysis; chronic obstructive pulmonary disease; differentially expressed gene; epidemiology; GEO data

Correspondence

Q. Wang, Physical Examination Center, Second People's Hospital of Yunnan Province, 176 Qingnian Road, Kunming, Yunnan 650500, China

Tel: +86 13668777585

E-mail: tjxwangqiang@163.com

F. Yu, Department of Gastroenterology, Fourth Affiliated Hospital of Kunming Medical University, Yunnan Province, 176 Qingnian Road, Kunming, Yunnan 650500, China

Tel: +86 15808875292

E-mail: yufb0402@163.com

X. Kong, Medical School, Kunming University of Science and Technology, 727 Jingming South Road, Kunming, Yunnan 650500, China

Tel: +86 18468257083

E-mail: 3517826707@qq.com

Xinwei Huang and Yunwei Li contributed equally to this work

Chronic obstructive pulmonary disease (COPD) is a multifactorial and heterogeneous disease that creates public health challenges worldwide. The underlying molecular mechanisms of COPD are not entirely clear. In this study, we aimed to identify the critical genes and potential molecular mechanisms of COPD by bioinformatic analysis. The gene expression profiles of lung tissues of COPD cases and healthy control subjects were obtained from the Gene Expression Omnibus. Differentially expressed genes were analyzed by integration with annotations from Gene Ontology and Kyoto Encyclopedia of Genes and Genomes, followed by construction of a protein-protein interaction network and weighted gene coexpression analysis. We identified 139 differentially expressed genes associated with the progression of COPD, among which 14 Hub genes were identified and found to be enriched in certain categories, including immune and inflammatory response, response to lipopolysaccharide and receptor for advanced glycation end products binding; in addition, these Hub genes are involved in multiple signaling pathways, particularly hematopoietic cell lineage and cytokine-cytokine receptor interaction. The 14 Hub genes were positively or negatively associated with COPD by *wGCNA* analysis. The genes *CX3CR1*, *PTGS2*, *FPR1*, *FPR2*, *S100A12*, *EGR1*, *CD163*, *S100A8* and *S100A9* were identified to mediate inflammation and injury of the lung, and play critical roles in the pathogenesis of COPD. These findings improve our understanding of the underlying molecular mechanisms of COPD.

Abbreviations

ARG1, arginase 1; BP, biological process; CC, cellular component; COPD, chronic obstructive pulmonary disease; CS, cigarette smoke; CX3CR1, C-X3-C motif chemokine receptor 1; DALY, disability-adjusted life year; DEG, differentially expressed gene; EGR1, early growth response 1; FC, fold change; FGG, fibrinogen gamma chain; FPR1, formyl peptide receptor 1; FPR2, formyl peptide receptor 2; GO, Gene Ontology; GS, gene significance; KEGG, Kyoto Encyclopedia of Genes and Genomes; MF, molecular function; MM, module membership; ORM1, orosomucoid 1; PPBP, proplatelet basic protein; PPI, protein-protein interaction; PTGS2, prostaglandin-endoperoxide synthase 2; S100A12, S100 calcium binding protein A12; S100A8, S100 calcium binding protein A8; S100A9, S100 calcium binding protein A9; VCAM1, vascular cell adhesion molecule 1; *wGCNA*, weighted gene coexpression network analysis; YLLs, years of life lost.

(Received 5 June 2019, revised 7 August 2019, accepted 13 August 2019)

doi:10.1002/2211-5463.12719

Chronic obstructive pulmonary disease (COPD), characterized by long-term poorly reversible airway limitation and persistent respiratory symptoms, is a common and preventable disease [1]. COPD is projected to become the third leading cause of all death by 2030 in the world [2]. Globally, COPD affected 299.4 million people in 2017, with a 71.2% increase in the prevalence rate compared with 2015, ranking it as the fifth leading cause of disability-adjusted life years (DALYs) and the seventh leading noncommunicable disease cause of years of life lost (YLLs) [3–6]. As shown in Fig. 1A, we observed a 12.3% increase in global all-age deaths caused by COPD from 2.85 million in 1990 to 3.20 million in 2017 [6–9] and a predicted increase of 60% by 2020 compared with 1990. Figure 1B indicated that the all-age standardized death rate of COPD in males, females, and both sexes separately decreased from 1990 to 2015 [6,8], which could be because of population growth and aging. Although the COPD death rate varies with different countries, more than 90% of COPD deaths occurred in low- and middle-income countries [10]. The global all-age YLLs with COPD showed a small increase of 7.5% and 3.6% for both sexes and males, respectively, as well as a 21% decrease for females from 1990 to 2015 (Fig. 1C) [6,8]. In addition, as shown in Fig. 1D [3–5,7,11–14], global all-age DALYs caused by COPD had a small increase of 4.2% during 1990–2015 and was projected to decline to 57.6 million by 2020. The age-standardized DALY rate caused by COPD in females was about twice as high as that of males, and that in low- and middle-income countries was 6.7 times higher than in some high-income countries [3]. We observed that the global all-age years lived with disability caused by COPD has grown 52.2% from 1990 to 2017. Taken together, COPD has presented a global public health challenge with high prevalence, mortality and disability rates, whereas the diagnosis of COPD is usually made based on spirometry values and clinical symptoms with a frequent underdiagnosis [15]. Thus, it is important to explore the underlying molecular mechanisms and identify more effective early diagnostic methods and reliable biomarkers for this disease.

As a large-scale and efficient technique for acquiring genomic data, microarray-based gene expression profiles have been widely used to seek new insights for

biomarkers in many human diseases [16]. Currently, many studies have been conducted on COPD gene expression profiles, and these studies have screened thousands of differentially expressed genes (DEGs) implicated in the development and progression of this disease [17,18]. However, the results for the identification of DEGs are discrepant among these studies due to sample heterogeneity and differences in technological detection platforms. In this study, we performed an integrated analysis on some of the gene expression profiling data based on lung tissues of COPD cases and control subjects using an unbiased approach aiming to identify the potential molecular mechanisms and biomarkers for COPD.

We selected two Gene Expression Omnibus microarray datasets on COPD (GSE27597 and GSE106986). DEGs were identified by R software (Auckland, New Zealand) and subsequently analyzed using bioinformatic methods including Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichments and construction of protein-protein interaction (PPI) and weighted gene coexpression analysis (WGCNA) networks. We screened the DEGs for potential association with the development and progression of COPD. Our work may further the understanding of the potential molecular mechanisms of COPD.

Materials and methods

Gene data

Two gene expression datasets, GSE27597 and GSE106986, were downloaded from the National Center for Biotechnology Information Gene Expression Omnibus database (<https://www.ncbi.nlm.nih.gov/geo/>). GSE27597 comprises the expression profile from 64 lung tissue samples from COPD cases and 16 samples from healthy donors [19]. GSE106986 includes molecular profiling of 19 lung tissue samples, containing 14 samples from COPD cases and 5 from smokers [20]. The experiments of GSE27597 and GSE106986 were conducted on the Affymetrix Human Exon 1.0 ST GeneChip (Affymetrix, Inc., Santa Clara, CA, USA; GPL5175 platform) and Agilent-026652 Whole Human Genome Microarray 4x44K v2 (Agilent Technologies, Inc., Palo Alto, CA, USA; GPL13497 platform), respectively. #AuthorQuery Rep

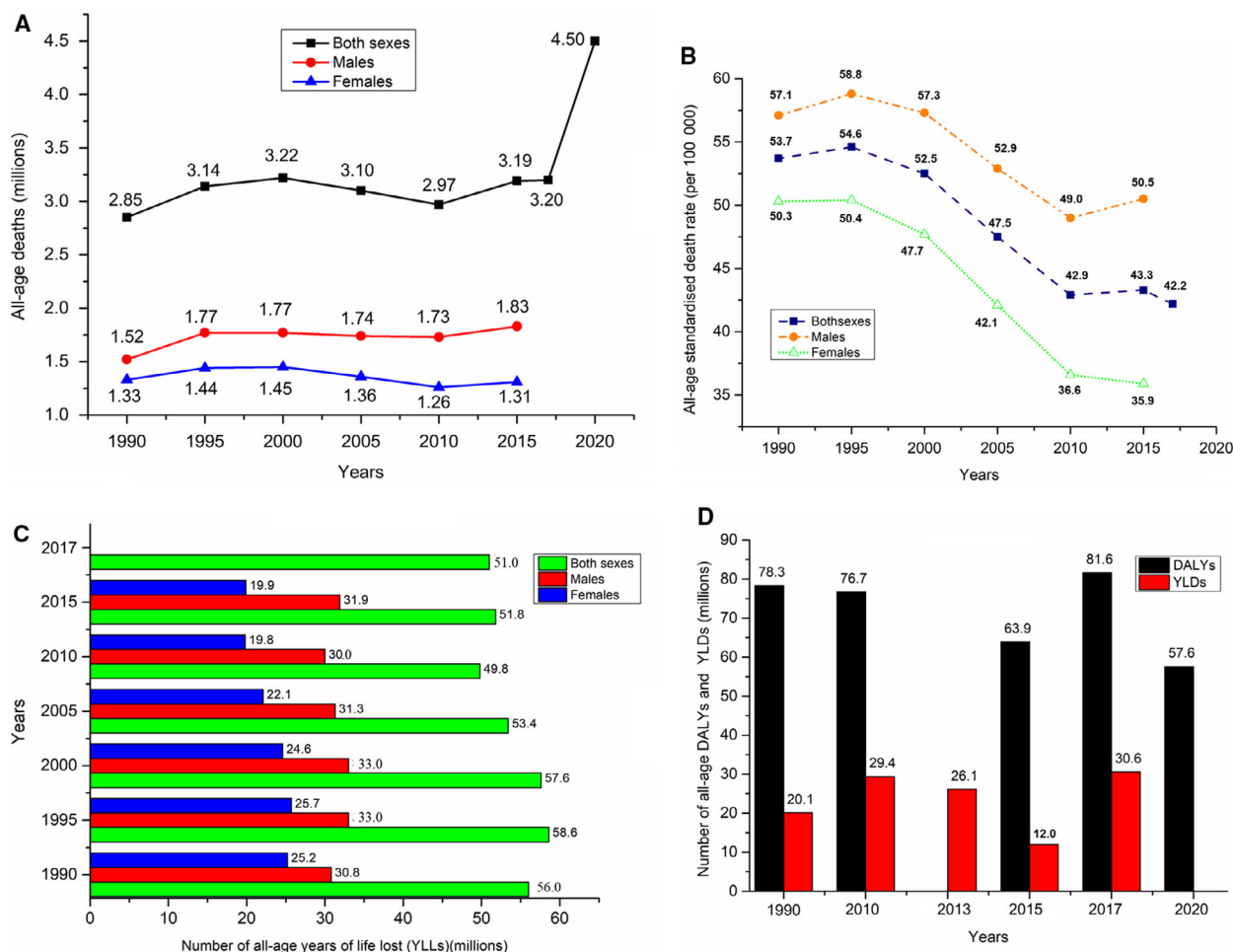


Fig. 1. The global death and burden caused by COPD. (A) Global age-related deaths (millions) caused by COPD in men and women, respectively, from 1990 to 2020 [6–9]. (B) Global age-related death rates (per 100 000) caused by COPD for both sexes, males and females, respectively, from 1990 to 2015 [6,8]. (C) Global age-related YLLs (millions) caused by COPD for both sexes, males and female, respectively, from 1990 to 2015 [6,8]. (D) Global age-related DALYs and years lived with disability (YLD) (millions) by COPD for both sexes from 1990 to 2020 [3–5,7,11–,14].

Table 1. Screening up-regulated and down-regulated DEGs in COPD by integrated microarray

DEGs	Gene symbol
Up-regulated (62 genes)	<i>PIEZO2, INHBA, VCAM1, RSPO2, CD1C, HSD17B6, AQP3, APLNR, GPD1, EGR2, KCNA, SLC18A2, FRZB, PTGS2, OLR1, EDNRA, SHISA2, ELF5, LUM, CYSLTR1, BMP5, HPGDS, MS4A2, DCC, FOSB, HHIP, MSMB, CD200, AMPD1, ICOS, RTKN2, CD83, FABP4, ISLR, GPR174, SLAMF7, WIF1, CTSW, CHIT1, CPA3, BMP1, EGR3, CX3CR1, EGR1, TREM2, ATF3, ICAM4, C8B, NR4A3, CCL8, HAS2, C4BPA, ITGB6, HBEGF, AGTR2, SERPIND1, CEACAM5, SFRP2, SELE, HLA-DRB1, HLA-DRB5, FGG</i>
Down-regulated (77 genes)	<i>IL1R2, CEACAM4, ST6GALNAC3, ITGA10, MERTK, FKBP5, CLEC4E, LAMB1, MGAM, RPS6KA2, FLT1, IRAK3, SULT1B1, AOX1, SH3PXD2B, RNASE2, IL18R1, CD163, IL1RL1, GRASP, MT1A, S100A12, MMP8, GLT1D1, TMTC1, S100A8, IL4R, IL18RAP, FLT3, ANPEP, MT1M, SIGLEC10, SPARCL1, SMAP2, TIMP4, ANGPTL1, HIF3A, PKHD1L1, LILRB3, FPR1, SLED1, LDLRAD3, FAM150B, FPR2, ZBTB16, GCA, ARG1, CXCR2P1, S100A9, TMEM204, TINAGL1, ABCC8, VCAN, APOLD1, DDIT4, ARRDC2, SERPINA3, PIK3R3, ADM, PNMT, BTNL9, CRISPLD2, SLCO2A1, VNN2, TUBB1, PSAT1, PPBP, DEFA4, AQP9, TTN, PDK4, MT1L, MT1X, ORM1, CHRM2, PTX3, EIF1AY</i>

Table 2. Screening 139 DEGs in COPD by integrated microarray. AveExpr, average expression; FDR, false discovery rate

Gene symbol	logFC	AveExpr	<i>t</i>	<i>P</i> -value	FDR
<i>IL1R2</i>	-3.724170635	4.910345092	-14.57070196	5.24E-25	8.38E-21
<i>CEACAM4</i>	-1.546190038	3.289938835	-14.11578934	3.65E-24	2.92E-20
<i>ST6GALNAC3</i>	-1.509348297	4.115922297	-13.72934365	1.93E-23	1.03E-19
<i>ITGA10</i>	-1.304452408	4.156071544	-13.3424854	1.05E-22	4.18E-19
<i>MERTK</i>	-1.371100619	5.743922084	-13.22669655	1.74E-22	4.99E-19
<i>FKBP5</i>	-2.553152066	5.964790869	-13.21010646	1.87E-22	4.99E-19
<i>CLEC4E</i>	-2.102592383	3.712865442	-12.2724444	1.22E-20	2.44E-17
<i>LAMB1</i>	-1.236900341	6.51036255	-12.11282186	2.51E-20	4.31E-17
<i>MGAM</i>	-1.918275604	3.743487329	-12.0971852	2.69E-20	4.31E-17
<i>RPS6KA2</i>	-1.051171425	6.235125074	-11.99287281	4.32E-20	6.28E-17
<i>FLT1</i>	-1.43735859	5.302047032	-11.6588263	1.98E-19	2.43E-16
<i>IRAK3</i>	-1.440362949	5.407265256	-11.53504535	3.48E-19	3.71E-16
<i>PIEZO2</i>	1.165120322	5.154801689	11.4846069	4.38E-19	4.38E-16
<i>SULT1B1</i>	-1.966128566	3.712292447	-11.44142282	5.34E-19	5.02E-16
<i>AOX1</i>	-1.498003567	4.380016656	-10.80239392	1.02E-17	8.14E-15
<i>SH3PXD2B</i>	-1.360072374	4.796109196	-10.45348254	5.15E-17	3.92E-14
<i>RNASE2</i>	-2.214352488	3.502846585	-10.41707921	6.11E-17	4.44E-14
<i>IL18R1</i>	-1.552709013	4.479557525	-10.21566115	1.56E-16	1.04E-13
<i>CD163</i>	-2.353331128	6.483508901	-10.18905044	1.77E-16	1.13E-13
<i>IL1RL1</i>	-2.008254697	5.747865421	-10.06701534	3.13E-16	1.88E-13
<i>INHBA</i>	1.437565757	4.698423572	10.06451042	3.17E-16	1.88E-13
<i>VCAM1</i>	1.58323448	4.54949487	9.975765691	4.80E-16	2.74E-13
<i>GRASP</i>	-1.019125221	5.351837423	-9.953904434	5.32E-16	2.93E-13
<i>RSPO2</i>	1.22508185	4.281279622	9.943524241	5.59E-16	2.98E-13
<i>CD1C</i>	1.320908918	4.048138575	9.847836259	8.75E-16	4.51E-13
<i>HSD17B6</i>	1.611303762	5.543589418	9.811225763	1.04E-15	5.03E-13
<i>MT1A</i>	-2.003809136	5.341700669	-9.802106469	1.08E-15	5.10E-13
<i>AQP3</i>	1.465444106	6.255672988	9.709215286	1.68E-15	7.66E-13
<i>S100A12</i>	-2.381911789	3.793860618	-9.658062064	2.13E-15	9.26E-13
<i>APLN</i>	1.462761779	4.060794485	9.514044214	4.19E-15	1.76E-12
<i>GPD1</i>	1.162408188	4.425391991	9.418745073	6.56E-15	2.65E-12
<i>EGR2</i>	1.7756242	5.312493191	9.415498145	6.66E-15	2.65E-12
<i>MMP8</i>	-1.963376342	1.850422771	-9.410996945	6.80E-15	2.65E-12
<i>KCNA3</i>	1.127720579	5.181587794	9.376872801	7.99E-15	2.98E-12
<i>SLC18A2</i>	1.155076678	4.861252305	9.375465524	8.04E-15	2.98E-12
<i>GLT1D1</i>	-1.354253187	4.134707596	-9.333414043	9.80E-15	3.33E-12
<i>TMTC1</i>	-1.221922174	4.907718006	-9.2268097	1.62E-14	5.28E-12
<i>S100A8</i>	-1.984436486	6.732435285	-9.226677546	1.62E-14	5.28E-12
<i>IL4R</i>	-1.016737627	6.13438495	-9.220088063	1.67E-14	5.34E-12
<i>IL18RAP</i>	-1.20706302	4.403623117	-9.181902429	2.00E-14	6.26E-12
<i>FRZB</i>	1.108947075	4.570621342	9.054280215	3.64E-14	1.08E-11
<i>PTGS2</i>	1.909599543	5.385732252	9.037623232	3.94E-14	1.14E-11
<i>FLT3</i>	-1.068728329	2.955444457	-9.011875495	4.44E-14	1.24E-11
<i>OLR1</i>	1.516024757	5.630156756	9.006280553	4.56E-14	1.24E-11
<i>ANPEP</i>	-1.451754311	4.687966342	-9.001861447	4.66E-14	1.24E-11
<i>MT1M</i>	-2.585708871	5.454959847	-8.990902454	4.90E-14	1.29E-11
<i>SIGLEC10</i>	-1.464287775	4.29164058	-8.960779525	5.65E-14	1.46E-11
<i>EDNRA</i>	1.016233852	5.691095587	8.939250683	6.25E-14	1.56E-11
<i>SHISA2</i>	1.00660372	5.288052754	8.886393339	8.02E-14	1.94E-11
<i>SPARCL1</i>	-1.096814178	7.567105051	-8.760826399	1.45E-13	3.30E-11
<i>SMAP2</i>	-1.030733342	6.515681918	-8.593715399	3.17E-13	6.73E-11
<i>TIMP4</i>	-1.656076369	4.229601784	-8.456527093	6.04E-13	1.19E-10
<i>ELF5</i>	1.521617319	4.673219557	8.448782695	6.26E-13	1.21E-10
<i>ANGPTL1</i>	-1.111036176	3.340161894	-8.413321275	7.39E-13	1.37E-10
<i>HIF3A</i>	-1.065614158	4.526630285	-8.386705336	8.38E-13	1.52E-10

Table 2. (Continued).

Gene symbol	logFC	AveExpr	<i>t</i>	<i>P</i> -value	FDR
<i>PKHD1L1</i>	-1.402091953	3.533210448	-8.281317436	1.37E-12	2.29E-10
<i>LUM</i>	1.008125744	6.388824425	8.261829056	1.50E-12	2.45E-10
<i>CYSLTR1</i>	1.017621995	4.334713339	8.231060286	1.74E-12	2.78E-10
<i>LILRB3</i>	-1.182035723	3.930324211	-8.201110624	2.00E-12	3.11E-10
<i>FPR1</i>	-1.497358339	5.123265733	-8.165184027	2.36E-12	3.63E-10
<i>BMP5</i>	1.082572	4.856866963	8.039056504	4.26E-12	6.14E-10
<i>HPGDS</i>	1.178976758	4.9661482	7.971057179	5.85E-12	8.07E-10
<i>SLED1</i>	-1.418643068	4.176703046	-7.93529438	6.92E-12	9.42E-10
<i>MS4A2</i>	1.127545704	4.501069641	7.934255426	6.95E-12	9.42E-10
<i>DCC</i>	1.107054924	4.066396917	7.874112823	9.20E-12	1.21E-9
<i>LDLRAD3</i>	-1.132183539	4.777010972	-7.827670594	1.14E-11	1.42E-9
<i>FAM150B</i>	-1.326763929	3.661484534	-7.801467295	1.29E-11	1.57E-9
<i>FOSB</i>	2.16754811	6.605909434	7.722677313	1.86E-11	2.20E-9
<i>FPR2</i>	-1.585779616	4.015151074	-7.702560188	2.04E-11	2.40E-9
<i>ZBTB16</i>	-1.704030574	5.808698383	-7.660799671	2.48E-11	2.87E-9
<i>GCA</i>	-1.036962474	4.831056628	-7.646120016	2.65E-11	3.03E-9
<i>HHIP</i>	1.361681455	6.338309941	7.624501498	2.93E-11	3.28E-9
<i>CD200</i>	1.045448111	4.105658515	7.579848564	3.61E-11	3.95E-9
<i>ARG1</i>	-1.363394352	2.024018729	-7.54254425	4.29E-11	4.60E-9
<i>AMPD1</i>	1.043721979	3.103940463	7.443735113	6.77E-11	6.85E-9
<i>ICOS</i>	1.222541378	4.277643563	7.41874473	7.59E-11	7.59E-9
<i>CXCR2P1</i>	-1.15616615	4.439580145	-7.399724607	8.29E-11	8.23E-9
<i>S100A9</i>	-1.545635339	5.747693368	-7.378920244	9.12E-11	9.01E-9
<i>RTKN2</i>	1.696929199	6.166705232	7.344341435	1.07E-10	1.03E-8
<i>TMEM204</i>	-1.048529786	6.26406167	-7.165974961	2.43E-10	2.11E-8
<i>CD83</i>	1.134658456	5.71435664	7.163495775	2.45E-10	2.12E-8
<i>FABP4</i>	1.58545563	4.778254935	7.12319994	2.95E-10	2.51E-8
<i>TINAGL1</i>	-1.026326535	5.532226715	-7.093847539	3.38E-10	2.84E-8
<i>ABCC8</i>	-1.278473261	3.740007718	-7.069781506	3.77E-10	3.09E-8
<i>VCAN</i>	-1.162194417	5.676393323	-7.050563618	4.11E-10	3.29E-8
<i>ISLR</i>	1.017944512	5.018753667	7.010420181	4.94E-10	3.80E-8
<i>APOLD1</i>	-1.385170588	5.098238511	-6.976021904	5.78E-10	4.35E-8
<i>DDIT4</i>	-1.347721118	5.257280573	-6.808572358	1.23E-9	8.33E-8
<i>ARRDC2</i>	-1.061269085	5.499220509	-6.762488227	1.52E-9	1.01E-7
<i>GPR174</i>	1.391314128	4.486280437	6.747714106	1.63E-9	1.07E-7
<i>SERPINA3</i>	-1.612709918	5.391772667	-6.745585747	1.64E-9	1.07E-7
<i>PIK3R3</i>	-1.190376757	5.660292101	-6.728449631	1.77E-9	1.14E-7
<i>SLAMF7</i>	1.083604552	4.392800911	6.727219831	1.78E-9	1.14E-7
<i>WIF1</i>	1.237156067	6.258872842	6.708747667	1.94E-9	1.24E-7
<i>ADM</i>	-1.075662424	5.196278444	-6.701236397	2.00E-9	1.28E-7
<i>PNMT</i>	-1.051631588	4.068574241	-6.553848929	3.88E-9	2.30E-7
<i>CTSW</i>	1.07957302	4.747480058	6.545368949	4.04E-9	2.38E-7
<i>BTNL9</i>	-1.132836349	5.578696855	-6.494296937	5.07E-9	2.82E-7
<i>CHIT1</i>	1.774209888	5.594573377	6.490067624	5.17E-9	2.87E-7
<i>CRISPLD2</i>	-1.078448778	5.7156417	-6.459148182	5.93E-9	3.26E-7
<i>CPA3</i>	1.002710147	5.278769406	6.398904433	7.75E-9	4.12E-7
<i>BMP1</i>	1.049814522	5.450758141	6.3426456	9.95E-9	5.12E-7
<i>SLCO2A1</i>	-1.051736858	6.743524273	-6.312993603	1.13E-8	5.69E-7
<i>VNN2</i>	-1.015536891	4.57616008	-6.274690529	1.34E-8	6.55E-7
<i>TUBB1</i>	-1.110536854	3.804924924	-6.264199999	1.41E-8	6.82E-7
<i>PSAT1</i>	-1.190066031	3.005323949	-6.258737184	1.44E-8	6.94E-7
<i>EGR3</i>	1.186935902	4.097335491	6.248573488	1.51E-8	7.22E-7
<i>CX3CR1</i>	1.284828947	5.212105592	6.144966898	2.38E-8	1.08E-6
<i>EGR1</i>	1.194905541	7.538328018	6.08564001	3.08E-8	1.35E-6
<i>TREM2</i>	1.128703086	3.987081843	6.03163178	3.90E-8	1.67E-6

Table 2. (Continued).

Gene symbol	logFC	AveExpr	<i>t</i>	<i>P</i> -value	FDR
<i>ATF3</i>	1.034162475	4.884273698	5.696410877	1.65E-7	5.92E-6
<i>PPBP</i>	-1.424569764	3.860588827	-5.646955582	2.04E-7	7.16E-6
<i>ICAM4</i>	1.390911425	5.019641081	5.625030952	2.24E-7	7.74E-6
<i>DEFA4</i>	-1.206319405	2.718323142	-5.614366702	2.34E-7	8.06E-6
<i>C8B</i>	1.199289417	3.724442441	5.595043977	2.54E-7	8.65E-6
<i>NR4A3</i>	1.153028606	4.646253713	5.56155394	2.93E-7	9.79E-6
<i>CCL8</i>	1.533954481	3.932933151	5.448623389	4.71E-7	1.48E-5
<i>AQP9</i>	-1.182290805	4.835881555	-5.430616006	5.08E-7	1.56E-5
<i>HAS2</i>	1.525864945	3.562127425	5.306311934	8.51E-7	2.44E-5
<i>TTN</i>	-1.043207507	3.817413024	-5.293238479	8.98E-7	2.56E-5
<i>C4BPA</i>	1.055176798	7.356691544	5.268476209	9.95E-7	2.79E-5
<i>ITGB6</i>	1.140444449	6.076736655	5.260091655	1.03E-6	2.86E-5
<i>PKD4</i>	-1.021408815	6.733080202	-5.134267178	1.72E-6	4.47E-5
<i>HBEGF</i>	1.032852155	5.680806271	5.061817586	2.31E-6	5.73E-5
<i>MT1L</i>	-1.182918274	2.526143932	-4.990697548	3.08E-6	7.38E-5
<i>MT1X</i>	-1.137570423	3.368378253	-4.974584006	3.29E-6	7.77E-5
<i>AGTR2</i>	1.364930071	3.524442993	4.842112874	5.57E-6	0.000119643
<i>ORM1</i>	-1.331793001	2.918107873	-4.800748967	6.56E-6	0.000138298
<i>CHRM2</i>	-1.011281909	2.379732249	-4.689445874	1.01E-5	0.000200001
<i>SERPIND1</i>	1.253485683	3.227485742	4.600275561	1.43E-5	0.000265554
<i>CEACAM5</i>	1.121861431	3.16229753	4.495719542	2.14E-5	0.000368719
<i>SFRP2</i>	1.129656517	4.14961175	4.486856217	2.21E-5	0.000377179
<i>SELE</i>	1.138034748	3.853029685	3.80214458	0.00026654	0.003035211
<i>PTX3</i>	-1.402248844	4.135304651	-3.574413207	0.000577009	0.005747801
<i>HLA-DRB1</i>	2.096264176	4.01015487	3.496867141	0.000745469	0.007027456
<i>HLA-DRB5</i>	2.134896841	3.007061623	3.473802838	0.000803932	0.007481531
<i>EIF1AY</i>	-1.506138016	3.020677733	-3.35747922	0.00117078	0.010029593
<i>FGG</i>	1.286056757	4.918716551	2.965432506	0.003905804	0.026040868
<i>MSMB</i>	1.936339164	3.146192242	2.931036227	0.00432076	0.028265267

Data preprocessing and screening for DEGs

The probe set IDs of two datasets were converted into gene symbols using the R software and annotation packages. The two datasets were merged into one array dataset and then batch normalized using R packages (*sva* and *limma* 3.40.6). The DEGs between COPD cases and control subjects were identified using the *limma* package in R 3.6.0. A *P*-value <0.05 after being adjusted by false discovery rate and $|\log_2FC| > 1$, where FC represents fold change, were applied together as the cutoff for DEGs screening.

GO and KEGG pathway enrichment analysis of DEGs

GO enrichment of the DEGs on the biological process (BP), molecular function (MF) and cellular component (CC) categories was performed using a DAVID online tool (<https://david.ncifcrf.gov/>) [21,22]. KEGG pathway enrichment analysis was performed by using the KOBAS 3.0 online analysis database (<http://kobas.cbi.pku.edu.cn/>) [23].

Construction of the PPI network

STRING database (<https://string-db.org/>) is frequently used for identifying the protein interactions [16,24].

STRING database contains huge amounts of experimental and text mining data [25]. CYTOSCAPE is an open source bioinformatic software platform used for integrating gene expression profiles and visualizing molecular interaction networks. CYTOSCAPE plugin cytoHubba provides multiple topological analysis methods on Hub genes, regulatory networks and drug targets for experimental biologists [26]. In this study, we used STRING database to identify the interactions between the identified DEGs. A confidence score >0.4 was used as the cutoff criterion. In addition, Hub DEGs were identified with CYTOSCAPE version 3.6.1 (1999 Free Software Foundation, Inc., Boston, MA, USA) with cytoHubba plugin according to the rank of connection degree (number) for each gene, which is represented by the different degrees of color (from red to yellow): the role of the gene is greater in the PPI network with the darker color of the gene [26].

WGCNA on COPD

WGCNA may be used for screening modules (clusters) of highly correlated genes and for calculating module membership (MM) measures, in which the module eigengene or an intramodular Hub gene is used to summarize such

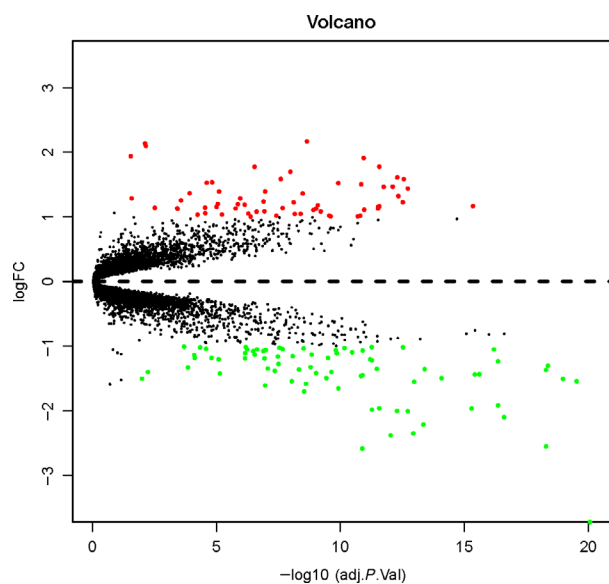


Fig. 3. Volcano plots of differential gene expression data between two sets of samples. Red represents the up-regulated DEGs, and green represents down-regulated DEGs. adj., *P*.Val, adjusted *P*-value.

Results

Identification of DEGs in COPD

After batch normalization on the integrated dataset from GSE27597 and GSE106986 by the *sva* and *limma* packages, 139 DEGs were identified using the *limma* package (corrected $P < 0.05$; $|\log_2FC| > 1$) (Tables 1 and 2). The cluster heatmap of the top 139 DEGs is shown in Fig. 2. Among them, 62 genes were up-regulated and 77 genes were down-regulated, which is shown in Fig. 3.

GO enrichment analysis of DEGs

GO analysis was done on the DEGs against BP, MF and CC terms. Biological annotation of the DEGs with COPD was identified using the DAVID online analysis tool. As shown in Fig. 4, GO functional enrichments of the DEGs with a P -value < 0.05 were obtained. Significant results of the GO enrichment analysis of the DEGs associated with COPD were shown in Table 3. In the BP category, the DEGs were mainly involved in inflammatory response, immune response and response to lipopolysaccharide. In the MF category, the DEGs were mainly enriched in receptor activity and receptor for advanced glycation end products (RAGE) receptor binding. In the CC category, the DEGs were mainly involved in the extracellular region and space, the integral component of the plasma membrane, the plasma membrane and the external side of the plasma membrane (Fig. 5).

KEGG pathway analysis of DEGs

We analyzed the DEGs using the KOBAS 3.0 online analysis database. As shown in Table 4, the DEGs were enriched in 48 pathways, especially hematopoietic cell lineage and cytokine-cytokine receptor interaction. The genes and pathway nodes are represented by CYTOSCAPE version 3.6.1 software (1999 Free Software Foundation, Inc., Boston, MA, USA) that was used to calculate the topological characteristics of the network and determine each node (Fig. 6).

PPI network analysis of DEGs

The 139 DEGs were applied for construction of PPI networks using STRING. After removing the discrete and partially connected nodes, the PPI network data of

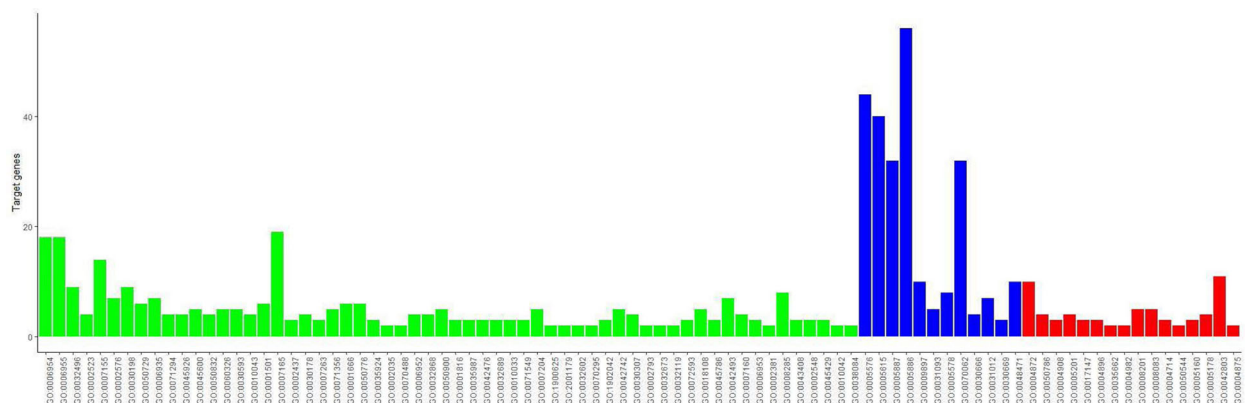


Fig. 4. GO enrichment analysis of DEGs in COPD. GO analysis divided DEGs into three functional groups: BPs, cell composition and MF. Green represents BP category, blue represents cell composition category and red represents MF category.

Table 3. GO analysis of DEGs associated with COPD

Term ID	Category	Description	Count	P-value	Bonferroni
GO:0006954	BP	Inflammatory response	18	9.11E-9	0.0000088
GO:0006955	BP	Immune response	18	4.26E-8	0.0000412
GO:0032496	BP	Response to lipopolysaccharide	9	0.0000513	0.04832
GO:0002523	BP	Leukocyte migration involved in inflammatory response	4	0.0000782	0.072795
GO:0007155	BP	Cell adhesion	14	0.0000787	0.073195
GO:0002576	BP	Platelet degranulation	7	0.000175	0.155654
GO:0030198	BP	Extracellular matrix organization	9	0.000178	0.157584
GO:0050729	BP	Positive regulation of inflammatory response	6	0.000292	0.245789
GO:0006935	BP	Chemotaxis	7	0.000438	0.344834
GO:0071294	BP	Cellular response to zinc ion	4	0.000438	0.345355
GO:0045926	BP	Negative regulation of growth	4	0.000438	0.345355
GO:0045600	BP	Positive regulation of fat cell differentiation	5	0.00053	0.400817
GO:0050832	BP	Defense response to fungus	4	0.001263	0.705093
GO:0060326	BP	Cell chemotaxis	5	0.001801	0.824787
GO:0030593	BP	Neutrophil chemotaxis	5	0.001906	0.841641
GO:0010043	BP	Response to zinc ion	4	0.002927	0.941058
GO:0001501	BP	Skeletal system development	6	0.004843	0.990808
GO:0007165	BP	Signal transduction	19	0.00503	0.992337
GO:0002437	BP	Inflammatory response to antigenic stimulus	3	0.0062	0.997541
GO:0030178	BP	Negative regulation of Wnt signaling pathway	4	0.008258	0.999668
GO:0007263	BP	Nitric oxide-mediated signal transduction	3	0.009889	0.999932
GO:0071356	BP	Cellular response to tumor necrosis factor	5	0.011676	0.999988
GO:0001666	BP	Response to hypoxia	6	0.012312	0.999994
GO:0050776	BP	Regulation of immune response	6	0.014105	0.999999
GO:0035924	BP	Cellular response to vascular endothelial growth factor stimulus	3	0.014331	0.999999
GO:0002035	BP	Brain renin-angiotensin system	2	0.015897	1
GO:0070488	BP	Neutrophil aggregation	2	0.015897	1
GO:0006952	BP	Defense response	4	0.016421	1
GO:0032868	BP	Response to insulin	4	0.016421	1
GO:0050900	BP	Leukocyte migration	5	0.016519	1
GO:0001816	BP	Cytokine production	3	0.016818	1
GO:0035987	BP	Endodermal cell differentiation	3	0.019474	1
GO:0042476	BP	Odontogenesis	3	0.019474	1
GO:0032689	BP	Negative regulation of interferon-gamma production	3	0.020864	1
GO:0010033	BP	Response to organic substance	3	0.020864	1
GO:0071549	BP	Cellular response to dexamethasone stimulus	3	0.022294	1
GO:0007204	BP	Positive regulation of cytosolic calcium ion concentration	5	0.022427	1
GO:1900625	BP	Positive regulation of monocyte aggregation	2	0.023751	1
GO:2001179	BP	Regulation of interleukin-10 secretion	2	0.023751	1
GO:0032602	BP	Chemokine production	2	0.023751	1
GO:0070295	BP	Renal water absorption	2	0.023751	1
GO:1902042	BP	Negative regulation of extrinsic apoptotic signaling pathway via death domain receptors	3	0.0284	1
GO:0042742	BP	Defense response to bacterium	5	0.028971	1
GO:0030307	BP	Positive regulation of cell growth	4	0.029638	1
GO:0002793	BP	Positive regulation of peptide secretion	2	0.031543	1
GO:0032673	BP	Regulation of interleukin-4 production	2	0.031543	1
GO:0032119	BP	Sequestering of zinc ion	2	0.031543	1
GO:0072593	BP	Reactive oxygen species metabolic process	3	0.031676	1
GO:0018108	BP	Peptidyl-tyrosine phosphorylation	5	0.034254	1
GO:0045786	BP	Negative regulation of cell cycle	3	0.035093	1
GO:0042493	BP	Response to drug	7	0.03518	1
GO:0007160	BP	Cell-matrix adhesion	4	0.035316	1
GO:0006953	BP	Acute-phase response	3	0.038646	1
GO:0002381	BP	Immunoglobulin production involved in immunoglobulin-mediated immune response	2	0.039273	1

Table 3. (Continued).

Term ID	Category	Description	Count	P-value	Bonferroni
GO:0008285	BP	Negative regulation of cell proliferation	8	0.039604	1
GO:0043408	BP	Regulation of mitogen-activated protein kinase cascade	3	0.042329	1
GO:0002548	BP	Monocyte chemotaxis	3	0.044219	1
GO:0045429	BP	Positive regulation of nitric oxide biosynthetic process	3	0.046139	1
GO:0010042	BP	Response to manganese ion	2	0.046942	1
GO:0038084	BP	Vascular endothelial growth factor signaling pathway	2	0.046942	1
GO:0005576	CC	Extracellular region	44	2.35E-14	3.22E-12
GO:0005615	CC	Extracellular space	40	4.29E-14	5.89E-12
GO:0005887	CC	Integral component of plasma membrane	32	2.71E-8	0.00000371
GO:0005886	CC	Plasma membrane	56	0.00000132	0.000181
GO:0009897	CC	External side of plasma membrane	10	0.0000284	0.003882
GO:0031093	CC	Platelet alpha-granule lumen	5	0.000713	0.093083
GO:0005578	CC	Proteinaceous extracellular matrix	8	0.003704	0.398509
GO:0070062	CC	Extracellular exosome	32	0.011945	0.807247
GO:0030666	CC	Endocytic vesicle membrane	4	0.012698	0.826358
GO:0031012	CC	Extracellular matrix	7	0.02228	0.954358
GO:0030669	CC	Clathrin-coated endocytic vesicle membrane	3	0.03651	0.993875
GO:0048471	CC	Perinuclear region of cytoplasm	10	0.039931	0.996238
GO:0004872	MF	Receptor activity	10	0.0000321	0.009231
GO:0050786	MF	RAGE receptor binding	4	0.0000611	0.017515
GO:0004908	MF	Interleukin-1 receptor activity	3	0.001097	0.271869
GO:0005201	MF	Extracellular matrix structural constituent	4	0.013171	0.97833
GO:0017147	MF	Wnt protein binding	3	0.02166	0.998215
GO:0004896	MF	Cytokine receptor activity	3	0.028658	0.999776
GO:0035662	MF	Toll-like receptor 4 binding	2	0.029063	0.999801
GO:0004982	MF	N-formyl peptide receptor activity	2	0.029063	0.999801
GO:0008201	MF	Heparin binding	5	0.030385	0.999866
GO:0008083	MF	Growth factor activity	5	0.031597	0.999907
GO:0004714	MF	Transmembrane receptor protein tyrosine kinase activity	3	0.031677	0.999909
GO:0050544	MF	Arachidonic acid binding	2	0.036196	0.999976
GO:0005160	MF	Transforming growth factor-beta receptor binding	3	0.03807	0.999987
GO:0005178	MF	Integrin binding	4	0.042233	0.999996
GO:0042803	MF	Protein homodimerization activity	11	0.042565	0.999997
GO:0004875	MF	Complement receptor activity	2	0.043278	0.999997

DEGs were imported into the cytoHubba of CYTOSCAPE version 3.6.1, and a complex network of the DEGs was constructed. As shown in Figs 7 and 8, 14 Hub DEGs were obtained, including *C-X3-C motif chemokine receptor 1 (CX3CR1)*, *proplatelet basic protein (PPBP)*, *prostaglandin-endoperoxide synthase 2 (PTGS2)*, *formyl peptide receptor 1 (FPR1)*, *formyl peptide receptor 2 (FPR2)*, *vascular cell adhesion molecule 1 (VCAM1)*, *S100 calcium binding protein A12 (S100A12)*, *arginase 1 (ARG1)*, *early growth response 1 (EGRI)*, *CD163*, *fibrinogen gamma chain (FGG)*, *orosomucoid 1 (ORM1)*, *S100 calcium binding protein A8 (S100A8)* and *S100 calcium binding protein A9 (S100A9)*.

WGCNA network construction in lung tissues

A WGCNA network was first constructed using lung tissue expression data from cohorts [GSE27597](#) and

[GSE106986](#), independent of COPD status and smoking status (ever/current smoking versus non-smoking). A total of 2942 DEGs were selected (a corrected $P < 0.05$) and subsequently used to identify modules of coexpressed genes using a hierarchical clustering procedure. The corresponding branches of the resulting dynamic clustering tree and module are shown as colored bands underneath the cluster tree. We then merged the highly similar dynamic clustering modules into the merged dynamic modules (cut height = 0.25) (Fig. 9). We identified nine modules ranging in size from 113 genes in the Purple module to 1081 in the Grey module. A module eigengene, a weighted average of the module gene expression profiles, was used to summarize the expression profiles of transcripts in a given module through their first principal component.

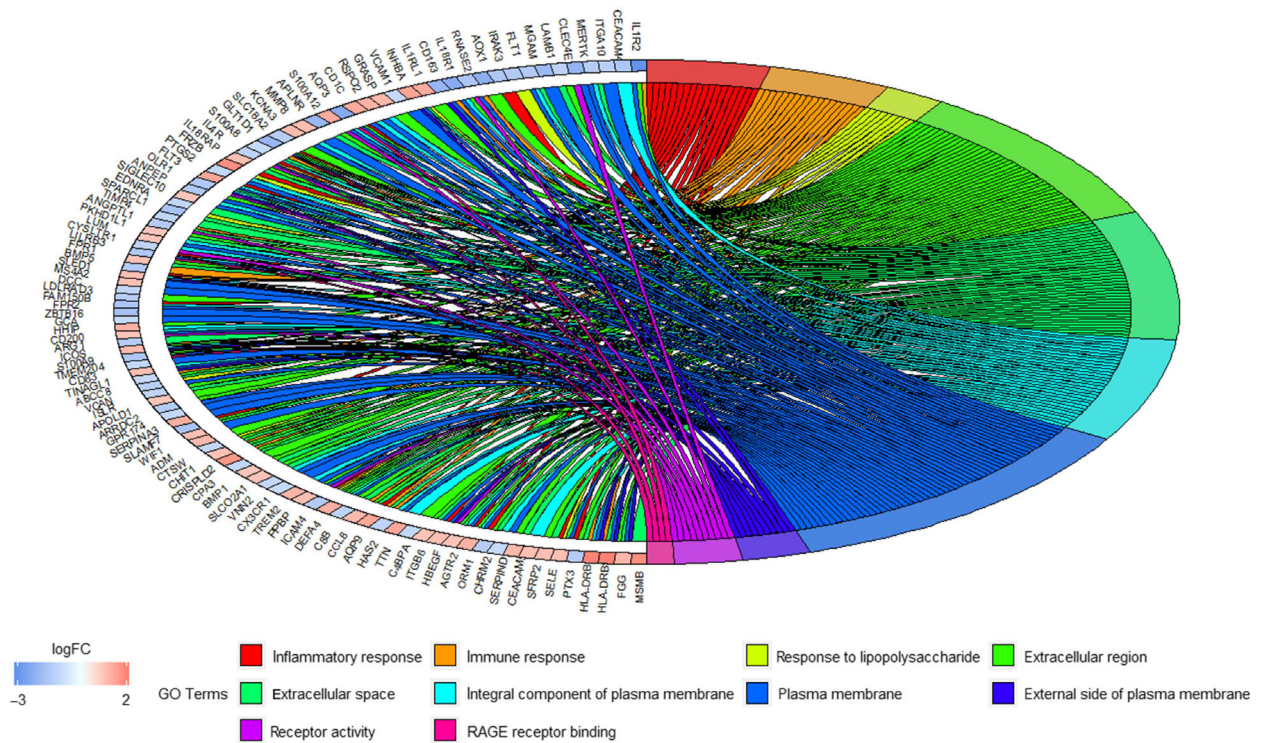


Fig. 5. Distribution of DEGs in COPD for the most significant GO-enriched functions.

Coexpression modules associated with COPD

To pinpoint modules associated with COPD and smoking status, we analyzed the association of each of the nine module eigengenes with the two traits. As shown in Fig. 10 and Table 5, all nine modules were significantly correlated with COPD and smoking status. Four modules were negatively associated with COPD and smoking status, marked Tan, Brown, Blue and Cyan, indicating that genes in these modules were predominantly down-regulated in COPD cases and those who had a history of smoking. However, five modules, in Green yellow, Purple, Black, Red and Grey, were positively associated with COPD cases and smoking status, showing that genes in these modules are predominantly up-regulated with the traits.

Four of these nine gene modules, in Cyan, Purple, Red and Grey, attracted our attention in that 14 Hub genes were identified as DEGs from the PPI analysis, including *CX3CR1*, *PPBP*, *PTGS2*, *FPR1*, *FPR2*, *S100A12*, *ARG1*, *EGR1*, *CD163*, *VCAM1*, *FGG*, *ORM1*, *S100A8* and *S100A9*. We calculated gene significance (GS) versus each MM. We found that the 14 Hub genes were also either positively or negatively associated with COPD (Table 6). *CX3CR1*, *PPBP*, *PTGS2*, *VCAM1*, *S100A12*, *ARG1*, *EGR1*, *CD163*, *S100A8* and *S100A9* were significantly associated with

each MM, whereas *FPR1*, *FPR2* and *ORM1* were correlated with each MM except Red MM, and *FGG* was correlated to each MM except the Purple MM. In addition, we found that the Purple (*CX3CR1*), Red (*EGR1*, *VCAM1* and *PTGS2*) and Grey (*ARG1*, *FGG*, and *PPBP*) genes most significantly correlated with GS for COPD were also the important MM elements (Fig. 11).

Discussion

In this study, we performed an integrated analysis on the gene expression profiles from lung tissues with or without COPD, aiming to identify the DEGs and related key signaling pathways for the disease. We identified 139 DEGs, including 62 up-regulated genes and 77 down-regulated genes. In addition, GO analysis showed that the 139 DEGs associated with COPD were involved in 60 BPs, 16 MFs and 12 CCs. Among these categories, the most important BPs included inflammatory response, immune response and response to lipopolysaccharide; the most important MFs included receptor activity and RAGE receptor binding; and the most important CCs included extracellular region and space, integral component of plasma membrane, plasma membrane and external side of plasma

Table 4. KEGG pathway analysis of DEGs associated with COPD

ID	Pathway	Gene count	Corrected <i>P</i> -value	DEGs
hsa04640	Hematopoietic cell lineage	7	4.07E−6	<i>FLT3, IL4R, HLA-DRB1, HLA-DRB5, IL1R2, CD1C, ANPEP</i>
hsa04060	Cytokine-cytokine receptor interaction	10	4.07E−6	<i>FLT3, FLT1, CCL8, IL4R, INHBA, CX3CR1, IL1R2, IL18RAP, PPBP, IL18R1</i>
hsa05150	<i>Staphylococcus aureus</i> infection	5	0.000147	<i>HLA-DRB1, FPR2, FGG, FPR1, HLA-DRB5</i>
hsa05166	Human T-lymphotropic virus type 1 infection	8	0.000174	<i>PIK3R3, HLA-DRB1, ATF3, HLA-DRB5, VCAM1, IL1R2, EGR2, EGR1</i>
hsa05321	Inflammatory bowel disease	5	0.000174	<i>HLA-DRB1, IL18RAP, IL4R, HLA-DRB5, IL18R1</i>
hsa04514	Cell adhesion molecules	6	0.000456	<i>SELE, HLA-DRB1, ICOS, HLA-DRB5, VCAM1, VCAN</i>
hsa04151	Phosphoinositide 3-kinase-Akt signaling pathway	8	0.000732	<i>CHRM2, PIK3R3, ITGA10, IL4R, LAMB1, DDIT4, ITGB6, FLT1</i>
hsa05146	Amoebiasis	5	0.000732	<i>C8B, PIK3R3, ARG1, LAMB1, IL1R2</i>
hsa04668	Tumor necrosis factor signaling pathway	5	0.001006	<i>VCAM1, PIK3R3, SELE, PTGS2, IL18R1</i>
hsa04080	Neuroactive ligand-receptor interaction	7	0.00112	<i>CHRM2, FPR2, AGTR2, APLNR, EDNRA, CYSLTR1, FPR1</i>
hsa05200	Pathways in cancer	8	0.001401	<i>DCC, FLT3, ZBTB16, LAMB1, PIK3R3, HHIP, PTGS2, EDNRA</i>
hsa04614	Renin-angiotensin system	3	0.00144	<i>CPA3, AGTR2, ANPEP</i>
hsa04610	Complement and coagulation cascades	4	0.002727	<i>FGG, C8B, C4BPA, SERPIND1</i>
hsa05310	Asthma	3	0.003037	<i>HLA-DRB1, HLA-DRB5, MS4A2</i>
hsa00750	Vitamin B ₆ metabolism	2	0.003805	<i>PSAT1, AOX1</i>
hsa05202	Transcriptional misregulation in cancer	5	0.005115	<i>FLT3, ZBTB16, FLT1, NR4A3, IL1R2</i>
hsa04933	AGE-RAGE signaling pathway in diabetic complications	4	0.005115	<i>VCAM1, PIK3R3, SELE, EGR1</i>
hsa04510	Focal adhesion	5	0.007653	<i>PIK3R3, ITGA10, ITGB6, LAMB1, FLT1</i>
hsa04672	Intestinal immune network for IgA production	3	0.007653	<i>HLA-DRB1, ICOS, HLA-DRB5</i>
hsa05145	Toxoplasmosis	4	0.007715	<i>HLA-DRB1, PIK3R3, LAMB1, HLA-DRB5</i>
hsa04978	Mineral absorption	3	0.007715	<i>MT1X, MT1A, MT1M</i>
hsa04923	Regulation of lipolysis in adipocytes	3	0.009035	<i>PIK3R3, PTGS2, FABP4</i>
hsa05221	Acute myeloid leukemia	3	0.009074	<i>FLT3, ZBTB16, PIK3R3</i>
hsa04380	Osteoclast differentiation	4	0.009515	<i>LILRB3, PIK3R3, FOSB, TREM2</i>
hsa01100	Metabolic pathways	12	0.00997	<i>PSAT1, HPGDS, ARG1, ST6GALNAC3, AOX1, PNMT, HSD17B6, AMPD1, PTGS2, CHIT1, MGAM, ANPEP</i>
hsa04145	Phagosome	4	0.015537	<i>HLA-DRB1, TUBB1, HLA-DRB5, OLR1</i>
hsa05140	Leishmaniasis	3	0.015831	<i>HLA-DRB1, PTGS2, HLA-DRB5</i>
hsa04512	Extracellular matrix-receptor interaction	3	0.02016	<i>ITGA10, ITGB6, LAMB1</i>
hsa05410	Hypertrophic cardiomyopathy	3	0.02016	<i>TTN, ITGA10, ITGB6</i>
hsa05222	Small-cell lung cancer	3	0.021463	<i>PIK3R3, PTGS2, LAMB1</i>
hsa05414	Dilated cardiomyopathy	3	0.023455	<i>TTN, ITGA10, ITGB6</i>
hsa05323	Rheumatoid arthritis	3	0.023455	<i>HLA-DRB1, FLT1, HLA-DRB5</i>
hsa04062	Chemokine signaling pathway	4	0.023572	<i>CX3CR1, PIK3R3, CCL8, PPBP</i>
hsa04915	Estrogen signaling pathway	3	0.027527	<i>FKBP5, PIK3R3, HBEGF</i>
hsa04024	cAMP signaling pathway	4	0.027527	<i>CHRM2, PIK3R3, HHIP, EDNRA</i>
hsa04810	Regulation of actin cytoskeleton	4	0.032602	<i>CHRM2, PIK3R3, ITGA10, ITGB6</i>
hsa00350	Tyrosine metabolism	2	0.032602	<i>AOX1, PNMT</i>
hsa05020	Prion diseases	2	0.032602	<i>C8B, EGR1</i>
hsa05143	African trypanosomiasis	2	0.032602	<i>VCAM1, SELE</i>
hsa05330	Allograft rejection	2	0.038438	<i>HLA-DRB1, HLA-DRB5</i>
hsa04722	Neurotrophin signaling pathway	3	0.038438	<i>IRAK3, PIK3R3, RPS6KA2</i>
hsa05332	Graft-versus-host disease	2	0.042328	<i>HLA-DRB1, HLA-DRB5</i>
hsa04940	Type I diabetes mellitus	2	0.04503	<i>HLA-DRB1, HLA-DRB5</i>
hsa04973	Carbohydrate digestion and absorption	2	0.047746	<i>MGAM, PIK3R3</i>

Table 4. (Continued).

ID	Pathway	Gene count	Corrected <i>P</i> -value	DEGs
hsa05322	Systemic lupus erythematosus	3	0.048708	<i>HLA-DRB1</i> , <i>C8B</i> , <i>HLA-DRB5</i>
hsa04930	Type II diabetes mellitus	2	0.049139	<i>ABCC8</i> , <i>PIK3R3</i>
hsa05144	Malaria	2	0.049139	<i>VCAM1</i> , <i>SELE</i>
hsa05030	Cocaine addiction	2	0.049139	<i>SLC18A2</i> , <i>FOSB</i>

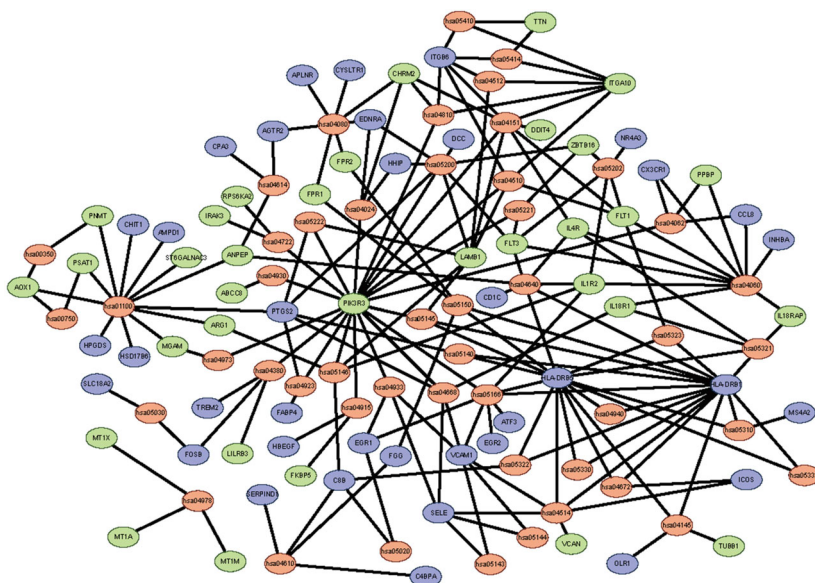


Fig. 6. The significant KEGG pathways enrichment of DEGs. Green represents down-regulated DEGs, blue represents up-regulated DEGs and red represents the signaling pathway.

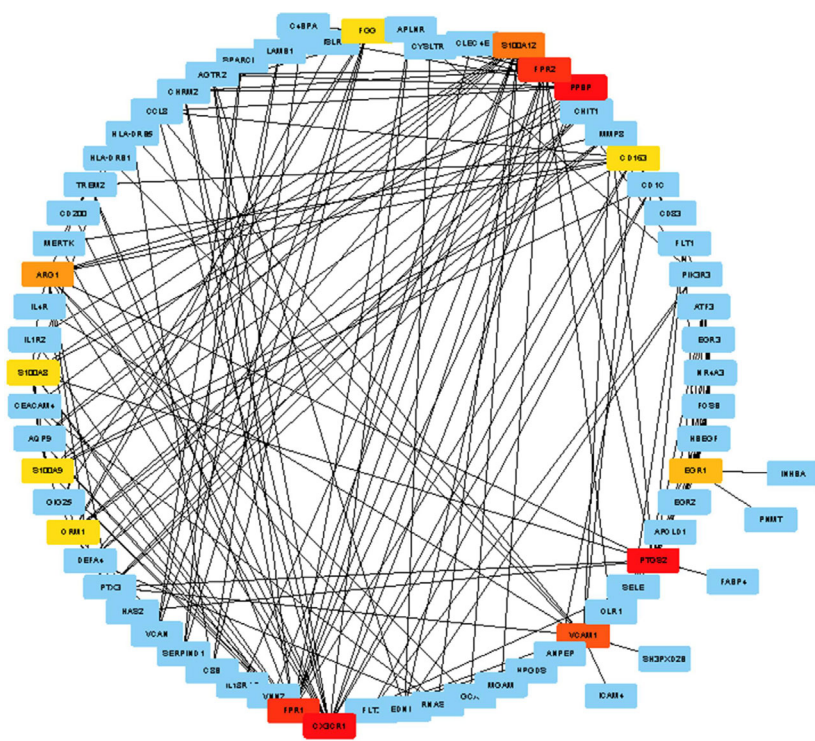


Fig. 7. PPI network and Hub DEGs. Hub DEGs were identified with CYTOSCAPE version 3.6.1 with cytoHubba plugin, according to the rank of connection degree (number) for each gene, which is represented by the different degrees of color (from red to yellow): the role of the gene is greater in the PPI network with the darker color of the gene. Red, saffron yellow and yellow represent Hub DEGs.

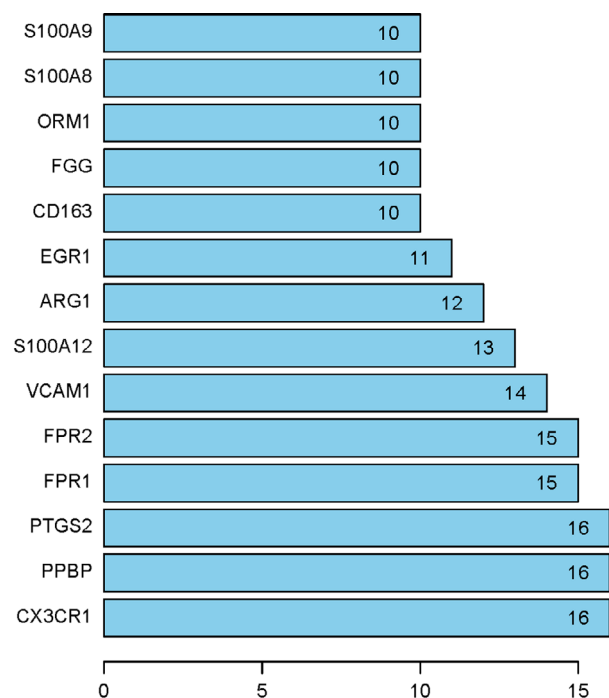


Fig. 8. PPI network identified Hub DEGs. Numbers represent connection points of the 14 Hub genes identified by the cytoHubba plugin.

membrane. This finding accords with the knowledge that COPD is characterized by chronic inflammation in the lung and airways [28,29]; immune response mediates the development of COPD caused by the harmful stimuli [30–32]; lipopolysaccharide may lead to increased airway and systemic inflammation, and contribute to the progressive deterioration of lung function [33,34]; and RAGE is a ‘driving force’ for

cigarette smoke (CS)-induced airway inflammation in COPD [35].

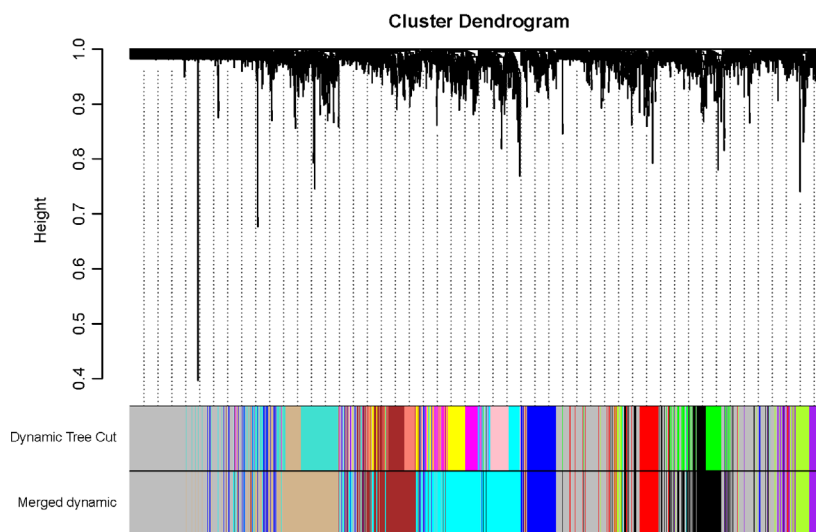
KEGG pathway analysis indicated that 48 pathways corresponded to these DEGs associated with COPD. Two pathways including hematopoietic cell lineage and cytokine-cytokine receptor interaction were most important. This finding is in line with the results from previous studies [18,32].

The PPI network of proteins encoded by DEGs identified 14 Hub DEGs associated with COPD, including *CX3CR1*, *PPBP*, *PTGS2*, *FPR1*, *FPR2*, *VCAM1*, *S100A12*, *ARG1*, *EGR1*, *CD163*, *FGG*, *ORM1*, *S100A8* and *S100A9*. All of these Hub genes were involved in the most important two BPs, two MFs or five CCs revealed by GO analysis, and were mainly implicated in multiple pathways identified by KEGG analysis in this study. Those results indicate that these Hub DEGs are involved in the development and progression of COPD by playing important biological roles in multiple signaling pathways.

Using WGCNA on the merged expression profile from two cohorts of lung tissues with COPD and healthy controls, we identified a set of gene signatures based on the 14 Hub genes. The increased expression of *CX3CR1*, *FGG*, *EGR1*, *VCAM1* and *PTGS2* is positively associated with COPD, and the underexpression of *PPBP*, *FPR1*, *FPR2*, *S100A12*, *ARG1*, *CD163*, *ORM1*, *S100A8* and *S100A9* is negatively associated with COPD.

CX3CR1 plays an important role in the development of chronic inflammatory lung diseases, such as COPD and emphysema, by contributing to structural destruction and remodeling. Chemoattractant inflammatory cells releasing *CX3CR1*, such as CD8⁺/CD4⁺, dendritic cells, $\gamma\delta$ T lymphocytes, natural killer cells and monocytes/macrophages, may lead to mononuclear cell

Fig. 9. Network construction identifies distinct modules of coexpressed genes. The network was constructed using the lung tissue expression dataset of GSE27597 and GSE106986. The cluster dendrogram was produced by average linkage hierarchical clustering of genes using 1 – topological overlap as dissimilarity measure. Modules (Dynamic Tree Cut) and similarly merged modules (Merged dynamic) of coexpressed genes were assigned colors corresponding to the branches indicated by the horizontal bar beneath the dendrogram (merged cut height = 0.25).



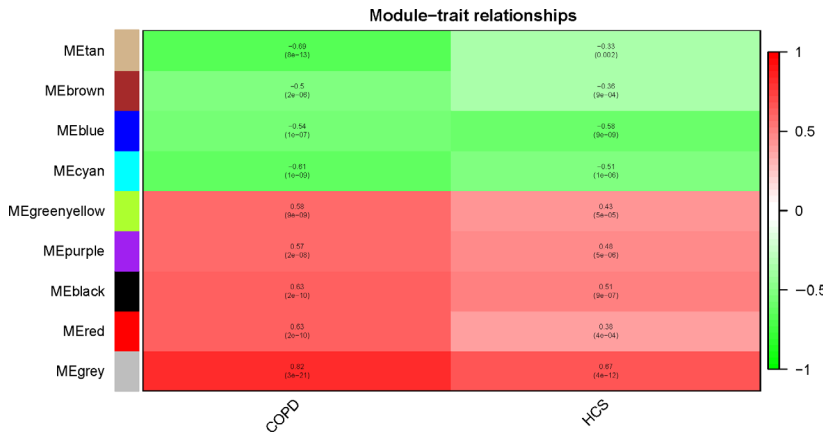


Fig. 10. WGCNA heatmap. Using the default parameter setting and all DEGs ($n = 2942$), we identified nine gene modules using WGCNA that were positively or negatively associated with COPD and smoking trait. Each row corresponds to a module eigengene and each column to a clinical trait (COPD and smoking status). Positive associations are red, and negative associations are green. HCS, history of smoking.

accumulation in the parenchyma and lung vessel walls, release mediators to induce injury, stimulate proliferation and chemoattractant inflammatory cells [36]. In addition, *CX3CR1*⁺ mononuclear phagocytes may induce an innate immune response to CS via producing interleukin-6 and tumor necrosis factor- α , and contribute to emphysema [37].

PTGS2 (COX-2), an important mediator of inflammation, was shown to be involved in inflammation response and associated with COPD pathogenesis [38–41]. The decreased activity of *PTGS2* may protect smokers against the development of COPD [40]. Furthermore, *FPR1* and *FPR2* were reported to be involved in recruitment and activation of inflammatory

Table 5. Correlation of module eigengene with COPD and smoking status traits

WGCNA modules	Gene number	Merged COPD dataset		Smoking status	
		Correlation	P-value	Correlation	P-value
Tan	353	-0.69	8E-13	-0.33	0.002
Brown	244	-0.5	2E-6	-0.36	9E-4
Blue	202	-0.54	1E-7	-0.58	9E-9
Cyan	467	-0.61	1E-9	-0.51	1E-6
Green yellow	105	0.58	9E-9	0.43	5E-5
Purple	113	0.57	2E-8	0.48	5E-6
Black	256	0.63	2E-10	0.51	9E-7
Red	121	0.63	2E-10	0.38	4E-4
Grey	1081	0.82	3E-21	0.67	4E-12

Table 6. Fourteen Hub genes positively or negatively associated with COPD and each MM. *P*, P-value for COPD or each MM

Gene	Located module	GS.COPD	P GS.COPD	P.MM Cyan	P.MM Grey	P.MM Purple	P.MM Red
<i>CD163</i>	Cyan	-0.739974023	1.33E-15	4.23E-21	8.88E-10	0.00249	2.57E-5
<i>FPR1</i>	Cyan	-0.662223538	9.25E-12	7.92E-23	5.84E-7	6.61E-5	0.07339
<i>FPR2</i>	Cyan	-0.63983063	7.43E-11	1.37E-21	4.51E-7	6.73E-5	0.07508
<i>ORM1</i>	Cyan	-0.459721444	1.23E-5	3.81E-20	0.00014	0.00196	0.70155
<i>S100A12</i>	Cyan	-0.721577872	1.41E-14	1.50E-14	2.82E-7	4.13E-5	0.00278
<i>S100A8</i>	Cyan	-0.70598989	9.04E-14	6.83E-18	4.01E-7	0.00026	0.01347
<i>S100A9</i>	Cyan	-0.623468037	3.07E-10	7.84E-20	4.68E-6	0.00049	0.04290
<i>ARG1</i>	Grey	-0.632480057	1.42E-10	0.02374	2.47E-10	0.00054	2.19E-9
<i>FGG</i>	Grey	0.304097261	0.00519	0.00646	0.005969	0.2425	3.39E-8
<i>PPBP</i>	Grey	-0.520314885	4.61E-7	5.44E-6	0.001842	0.0388	0.01072
<i>CX3CR1</i>	Purple	0.553212675	5.84E-8	0.00133	9.38E-10	1.81E-13	4.46E-5
<i>EGR1</i>	Red	0.549754835	7.33E-8	0.00127	2.09E-6	0.03903	1.80E-14
<i>PTGS2</i>	Red	0.698684808	2.07E-13	0.00341	2.42E-11	4.99E-5	1.10E-28
<i>VCAM1</i>	Red	0.7344482490	2.75E-15	0.00947	2.66E-12	5.70E-13	3.14E-16

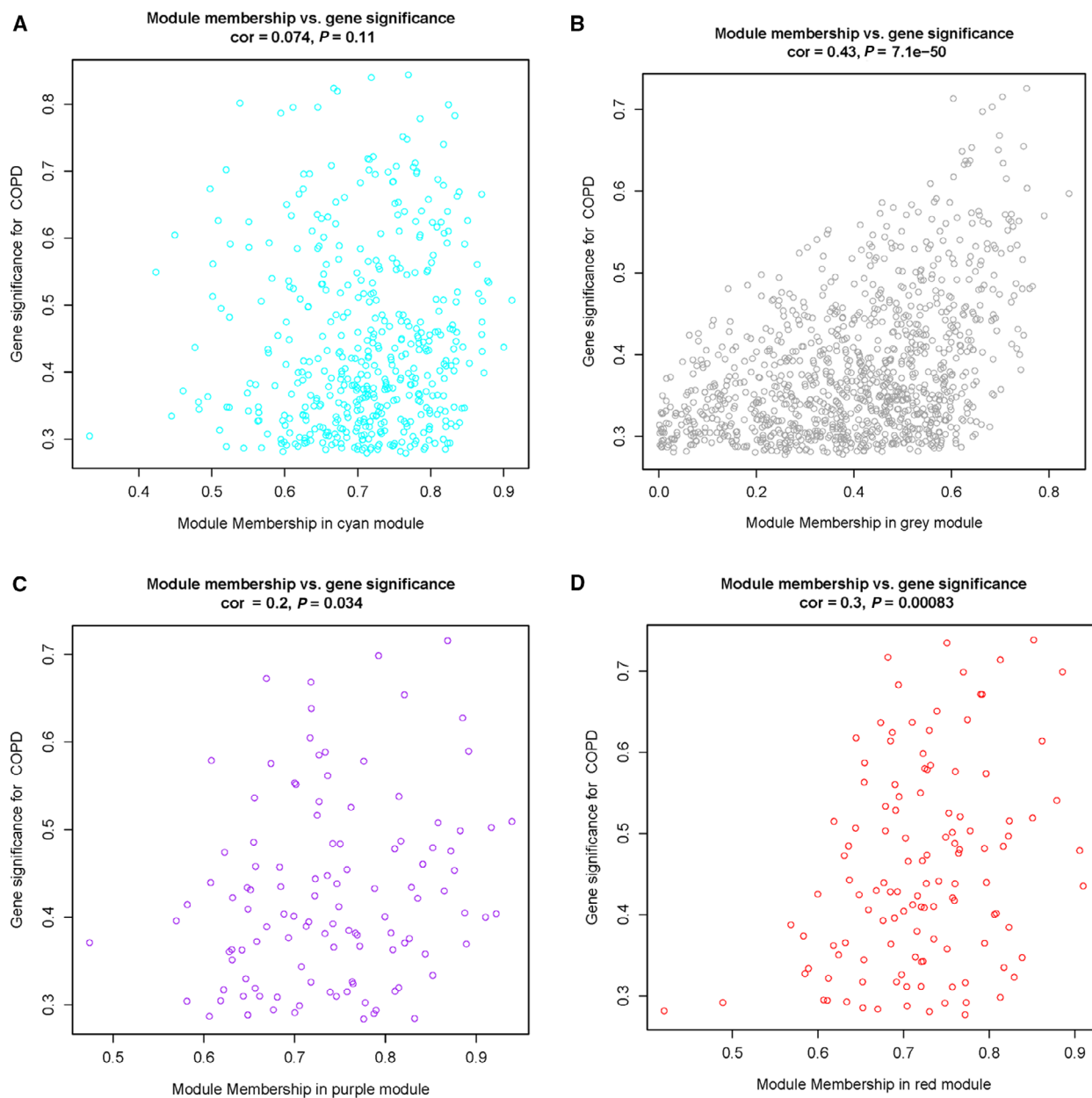


Fig. 11. COPD absolute GS versus MM. WGCNA calculation of GS to COPD versus MM. In oversimplified terms, MM is a measure of how 'tight' genes cluster within the module, or mathematically, how close gene expression is to the module eigenvalue. A gene with high MM and GS identifies Hub genes that are both key components to the underlying BP and highly associated with the trait of interest. The GS for COPD was plotted. (A) Cyan represents MM. (B) Grey represents MM. (C) Purple represents MM. (D) Red represents MM.

cells induced by CS, and play important roles in lung inflammation, injury and the pathogenesis of COPD [42–45].

S100A8, *S100A9*, and *S100A12* might induce neutrophil and monocyte chemotaxis, adhesion to fibrinogen and diapedesis, and neutrophil migration to inflammatory sites [46,47], and have been identified as key biomarkers in inflammatory diseases including

COPD and neutrophil-dominated infections [35,48,49]. The mRNA levels of *S100A8*, *S100A9* and *S100A12* may be regulated by RAGE, which was shown to contribute to CS-induced airway inflammation in COPD [35]. This is consistent with our result in this study that the RAGE pathway including *S100A8*, *S100A9* and *S100A12* is important in the development of COPD.

EGRI is an autophagy regulator gene that plays important roles in cellular homeostasis, airway remodeling and control of inflammatory immune response; it is also a significant risk factor for susceptibility to COPD [50–52]. *EGRI* may be induced by CS and involved in proinflammatory mechanisms that are likely associated with the development of COPD [51], whereas *Egr-1*^{-/-} mice were observed to resist CS-induced autophagy, apoptosis and emphysema [53]. These findings exhibit a critical role for *EGRI* in CS-induced inflammatory immune response and COPD, and effective inhibition of *EGRI* may attenuate airway remodeling and inflammation associated with the pathology of COPD.

CD163, a carefully regulated component of the innate immune response to infection and a macrophage receptor for bacteria, was shown to play important roles in functional pulmonary defense elements and the inflammatory immune response of the respiratory system [54,55]. Overexpression of *CD163* on lung alveolar macrophages may be implicated in the pathogenesis of COPD and poor lung function [56].

ARG1 was shown to contribute to asthma pathogenesis by inhibiting nitric oxide production, modulating fibrosis, regulating arginine metabolism and inhibiting T cell proliferation, and it involves the initiation of adaptive T helper 2 cell-mediated allergic lung inflammation by regulating group 2 innate lymphoid cells [57–60], whereas *ARG1* ablation in the lung may enhance peripheral lung function but have little effect on airway inflammation [61]. The role of *ARG1* in COPD needs to be studied in the future.

ORM1 appears to function in regulating the activity of the immune system during the acute-phase reaction and has been identified as an acute exacerbation of COPD-specific immunomodulatory mediator [62]. *PPBP* serves as a potent neutrophil chemoattractant and activator, and its elevated expression in the bronchial mucosa might be involved in the pathogenesis of COPD [63,64]. In addition, *VCAMI* was shown to express in endothelial cells of atopic asthma cases, but not COPD cases [65], and present an association with lung function [66]. *FGG* was found to be involved in blast lung injury resistance via promoting tissue-protective adenosine signaling [67]. In the lung tissues of COPD cases, its mRNA expression was reported to correlate with the burden of particulate matter in total lung and lung parenchyma [68].

In conclusion, we have identified 139 candidate DEGs associated with the progression of COPD. The results from bioinformatic analysis are in agreement with those from previous cell and animal models and human studies. Our results showed that nine Hub

genes, *CX3CR1*, *PTGS2*, *FPR1*, *FPR2*, *SI00A12*, *EGRI*, *CD163*, *SI00A8* and *SI00A9*, potentially mediated inflammation and injury of the lung, and play critical roles in the pathogenesis of COPD. The roles of five Hub genes, including *PPBP*, *ARG1*, *FGG*, *VCAMI* and *ORM1*, identified to be associated with COPD in this study need to be confirmed in the future. These findings could improve our understanding of the underlying molecular mechanisms of COPD and provide us with insights for drug target discovery for the disease.

Acknowledgements

This work was funded by the National Natural Science Foundation of China (81460007), Natural Science Foundation of Yunnan Province, China (2009CD130), and two funds of the Department of Education Scientific Research of Yunnan Province, China (09C0294 and 2015Z085).

Conflict of interest

The authors declare no conflict of interest.

Author contributions

XK, FY and QW conceived and designed the project. XH and YL acquired the data. XG and ZZ analyzed and interpreted the data. XH and YL wrote the paper.

References

- 1 Vogelmeier CF, Criner GJ, Martinez FJ, Anzueto A, Barnes PJ, Bourbeau J, Celli BR, Chen R, Decramer M, Fabbri LM *et al.* (2017) Global strategy for the diagnosis, management, and prevention of chronic obstructive lung disease 2017 report: GOLD executive summary. *Arch Bronconeumol* **53**, 128–149.
- 2 Song Q, Christiani DC, Wang X and Ren J (2014) The global contribution of outdoor air pollution to the incidence, prevalence, mortality and hospital admission for chronic obstructive pulmonary disease: a systematic review and meta-analysis. *Int J Environ Res Public Health* **11**, 11822–11832.
- 3 GBD 2015 Chronic Respiratory Disease Collaborators (2017) Global, regional, and national deaths, prevalence, disability-adjusted life years, and years lived with disability for chronic obstructive pulmonary disease and asthma, 1990–2015: a systematic analysis for the Global Burden of Disease Study 2015. *Lancet Respir Med* **5**, 691–706.
- 4 GBD 2017 Disease and Injury Incidence and Prevalence Collaborators (2018) Global, regional, and national

- incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet* **392**, 1789–1858.
- 5 GBD 2017 DALYs and HALE Collaborators (2018) Global, regional, and national disability-adjusted life-years (DALYs) for 359 diseases and injuries and healthy life expectancy (HALE) for 195 countries and territories, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet* **392**, 1859–1922.
 - 6 GBD 2017 Causes of Death Collaborators (2018) Global, regional, and national age-sex-specific mortality for 282 causes of death in 195 countries and territories, 1980–2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet* **392**, 1736–1788.
 - 7 Lopez AD and Murray CC (1998) The global burden of disease, 1990–2020. *Nat Med* **4**, 1241–1243.
 - 8 GBD 2015 Mortality and Causes of Death Collaborators (2016) Global, regional, and national life expectancy, all-cause mortality, and cause-specific mortality for 249 causes of death, 1980–2015: a systematic analysis for the Global Burden of Disease Study 2015. *Lancet* **388**, 1459–1544.
 - 9 GBD 2016 Causes of Death Collaborators (2017) Global, regional, and national age-sex specific mortality for 264 causes of death, 1980–2016: a systematic analysis for the Global Burden of Disease Study 2016. *Lancet* **390**, 1151–1210.
 - 10 Mathers CD and Loncar D (2006) Projections of global mortality and burden of disease from 2002 to 2030. *PLoS Med* **3**, e442.
 - 11 GBD 2015 DALYs and HALE Collaborators (2016) Global, regional, and national disability-adjusted life-years (DALYs) for 315 diseases and injuries and healthy life expectancy (HALE), 1990–2015: a systematic analysis for the Global Burden of Disease Study 2015. *Lancet* **388**, 1603–1658.
 - 12 Lopez AD, Mathers CD, Ezzati M, Jamison DT and Murray CJ (2006) Global and regional burden of disease and risk factors, 2001: systematic analysis of population health data. *Lancet* **367**, 1747–1757.
 - 13 Murray CJ, Vos T, Lozano R, Naghavi M, Flaxman AD, Michaud C, Ezzati M, Shibuya K, Salomon JA, Abdalla S *et al.* (2012) Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet* **380**, 2197–2223.
 - 14 Global Burden of Disease Study 2013 Collaborators (2015) Global, regional, and national incidence, prevalence, and years lived with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990–2013: a systematic analysis for the Global Burden of Disease Study 2013. *Lancet* **386**, 743–800.
 - 15 Cortopassi F, Gurung P and Pinto-Plata V (2017) Chronic obstructive pulmonary disease in elderly patients. *Clin Geriatr Med* **33**, 539–552.
 - 16 Yang X, Zhu S, Li L, Zhang L, Xian S, Wang Y and Cheng Y (2018) Identification of differentially expressed genes and signaling pathways in ovarian cancer by integrated bioinformatics analysis. *Onco Targets Ther* **11**, 1457–1474.
 - 17 Chen W, Hong YQ and Meng ZL (2014) Bioinformatics analysis of molecular mechanisms of chronic obstructive pulmonary disease. *Eur Rev Med Pharmacol Sci* **18**, 3557–3563.
 - 18 Zhao J, Cheng W, He X, Liu Y, Li J, Sun J, Li J, Wang F and Gao Y (2018) Chronic obstructive pulmonary disease molecular subtyping and pathway deviation-based candidate gene identification. *Cell J* **20**, 326–332.
 - 19 Campbell JD, McDonough JE, Zeskind JE, Hackett TL, Pechkovsky DV, Brandsma CA, Suzuki M, Gosselink JV, Liu G, Alekseyev YO *et al.* (2012) A gene expression signature of emphysema-related lung destruction and its reversal by the tripeptide GHK. *Genome Med* **4**, 67.
 - 20 Heinbockel L, Marwitz S, Schromm AB, Watz H, Kugler C, Ammerpohl O, Schnepf K, Rabe KF, Droemann D and Goldmann T (2018) Identification of novel target genes in human lung tissue involved in chronic obstructive pulmonary disease. *Int J Chron Obstruct Pulmon Dis* **13**, 2255–2259.
 - 21 da Huang W, Sherman BT and Lempicki RA (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* **4**, 44–57.
 - 22 Huang DW, Sherman BT and Lempicki RA (2009) Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res* **37**, 1–13.
 - 23 Xie C, Mao X, Huang J, Ding Y, Wu J, Dong S, Kong L, Gao G, Li CY and Wei L (2011) KOBAS 2.0: a web server for annotation and identification of enriched pathways and diseases. *Nucleic Acids Res* **39**, W316–W322.
 - 24 Szklarczyk D, Morris JH, Cook H, Kuhn M, Wyder S, Simonovic M, Santos A, Doncheva NT, Roth A, Bork P *et al.* (2017) The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. *Nucleic Acids Res* **45**, D362–D368.
 - 25 von Mering C, Huynen M, Jaeggi D, Schmidt S, Bork P and Snel B (2003) STRING: a database of predicted functional associations between proteins. *Nucleic Acids Res* **31**, 258–261.
 - 26 Chin CH, Chen SH, Wu HH, Ho CW, Ko MT and Lin CY (2014) cytoHubba: identifying hub objects and subnetworks from complex interactome. *BMC Syst Biol* **8** (Suppl 4), S11.

- 27 Langfelder P and Horvath S (2008) WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* **9**, 559.
- 28 Barnes PJ (2019) Inflammatory endotypes in COPD. *Allergy* **74**, 1249–1256.
- 29 Barnes PJ (2017) Cellular and molecular mechanisms of asthma and COPD. *Clin Sci (Lond)* **131**, 1541–1558.
- 30 Hopkins RJ and Young RP (2019) Mevalonate signaling, COPD and cancer: the statins and beyond. *J Investig Med* **67**, 711–714.
- 31 Bruno A, Cipollina C, Di Vincenzo S, Siena L, Dino P, Di Gaudio F, Gjomarkaj M and Pace E (2017) Ceftaroline modulates the innate immune and host defense responses of immunocompetent cells exposed to cigarette smoke. *Toxicol Lett* **279**, 9–15.
- 32 Bi H, Zhou J, Wu D, Gao W, Li L, Yu L, Liu F, Huang M, Adcock IM, Barnes PJ *et al.* (2015) Microarray analysis of long non-coding RNAs in COPD lung tissue. *Inflamm Res* **64**, 119–126.
- 33 Gupta V, Banyard A, Mullan A, Sriskantharajah S, Southworth T and Singh D (2015) Characterization of the inflammatory response to inhaled lipopolysaccharide in mild to moderate chronic obstructive pulmonary disease. *Br J Clin Pharmacol* **79**, 767–776.
- 34 D'Anna C, Cigna D, Di Sano C, Di Vincenzo S, Dino P, Ferraro M, Bini L, Bianchi L, Di Gaudio F, Gjomarkaj M *et al.* (2017) Exposure to cigarette smoke extract and lipopolysaccharide modifies cytoskeleton organization in bronchial epithelial cells. *Exp Lung Res* **43**, 347–358.
- 35 Li M, Guo L, Wang H, Wang T, Shen Y, Liao Z, Wen F and Chen L (2015) RAGE-ligands axis: a new 'driving force' for cigarette smoke-induced airway inflammation in COPD? *Respirology* **20**, 998–999.
- 36 Zhang J and Patel JM (2010) Role of the CX3CL1-CX3CR36 axis in chronic inflammatory lung diseases. *Int J Clin Exp Med* **3**, 233–244.
- 37 Xiong Z, Leme AS, Ray P, Shapiro SD and Lee JS (2011) CX3CR3⁺ lung mononuclear phagocytes spatially confined to the interstitium produce TNF- α and IL-6 and promote cigarette smoke-induced emphysema. *J Immunol* **186**, 3206–3214.
- 38 Wu L, Merrilees M, Young RP and Black PN (2011) The cyclooxygenase-2-765C promoter polymorphism protects against the development of chronic obstructive pulmonary disease. *Respir Med* **105**, 506–510.
- 39 Bonanno A, Albano GD, Siena L, Montalbano AM, Riccobono L, Anzalone G, Chiappara G, Gagliardo R, Profita M and Sala A (2016) Prostaglandin E(2) possesses different potencies in inducing vascular endothelial growth factor and interleukin-8 production in COPD human lung fibroblasts. *Prostaglandins Leukot Essent Fatty Acids* **106**, 11–18.
- 40 Pietras T, Szemraj J, Panek M, Witusik A, Banasiak M, Antczak A and Gorski P (2012) Functional polymorphism of cyclooxygenase-2 gene (G-765C) in chronic obstructive pulmonary disease patients. *Mol Biol Rep* **39**, 2163–2167.
- 41 Zago M, Sheridan JA, Traboulsi H, Hecht E, Zhang Y, Guerrina N, Matthews J, Nair P, Eidelman DH, Hamid Q *et al.* (2017) Low levels of the AhR in chronic obstructive pulmonary disease (COPD)-derived lung cells increases COX-2 protein by altering mRNA stability. *PLoS One* **12**, e0180881.
- 42 Chen YC, Lin MC, Lee CH, Liu SF, Wang CC, Fang WF, Chao TY, Wu CC, Wei YF, Chang HC *et al.* (2018) Defective formyl peptide receptor 2/3 and annexin A1 expressions associated with M2a polarization of blood immune cells in patients with chronic obstructive pulmonary disease. *J Transl Med* **16**, 69.
- 43 Cardini S, Dalli J, Fineschi S, Perretti M, Lungarella G and Lucattelli M (2012) Genetic ablation of the fpr1 gene confers protection from smoking-induced lung emphysema in mice. *Am J Respir Cell Mol Biol* **47**, 332–339.
- 44 Migeotte I, Communi D and Parmentier M (2006) Formyl peptide receptors: a promiscuous subfamily of G protein-coupled receptors controlling immune responses. *Cytokine Growth Factor Rev* **17**, 501–519.
- 45 Hartt JK, Barish G, Murphy PM and Gao JL (1999) N-formylpeptides induce two distinct concentration optima for mouse neutrophil chemotaxis by differential interaction with two N-formylpeptide receptor (FPR) subtypes. Molecular characterization of FPR2, a second mouse neutrophil FPR. *J Exp Med* **190**, 741–747.
- 46 Ryckman C, Vandal K, Rouleau P, Talbot M and Tessier PA (2003) Proinflammatory activities of S100: proteins S100A8, S100A9, and S100A8/A9 induce neutrophil chemotaxis and adhesion. *J Immunol* **170**, 3233–3242.
- 47 Rouleau P, Vandal K, Ryckman C, Poubelle PE, Boivin A, Talbot M and Tessier PA (2003) The calcium-binding protein S100A12 induces neutrophil adhesion, migration, and release from bone marrow in mouse at concentrations similar to those found in human inflammatory arthritis. *Clin Immunol* **107**, 46–54.
- 48 Kang JH, Hwang SM and Chung IY (2015) S100A8, S100A9 and S100A12 activate airway epithelial cells to produce MUC5AC via extracellular signal-regulated kinase and nuclear factor-kappaB pathways. *Immunology* **144**, 79–90.
- 49 Lorenz E, Muhlebach MS, Tessier PA, Alexis NE, Duncan Hite R, Seeds MC, Peden DB and Meredith W (2008) Different expression ratio of S100A8/A9 and S100A12 in acute and chronic lung diseases. *Respir Med* **102**, 567–573.
- 50 Kramer EL, Hardie WD, Mushaben EM, Acciani TH, Pastura PA, Korfhagen TR, Hershey GK, Whitsett JA

- and Le Cras TD (2011) Rapamycin decreases airway remodeling and hyperreactivity in a transgenic model of noninflammatory lung disease. *J Appl Physiol (1985)* **111**, 1760–1767.
- 51 Reynolds PR, Cosio MG and Hoidal JR (2006) Cigarette smoke-induced Egr-1 upregulates proinflammatory cytokines in pulmonary epithelial cells. *Am J Respir Cell Mol Biol* **35**, 314–319.
- 52 Chen CZ, Ou CY, Wang RH, Lee CH, Lin CC, Chang HY and Hsiue TR (2015) Association of Egr-1 and autophagy-related gene polymorphism in men with chronic obstructive pulmonary disease. *J Formos Med Assoc* **114**, 750–755.
- 53 Chen ZH, Kim HP, Sciruba FC, Lee SJ, Feghali-Bostwick C, Stolz DB, Dhir R, Landreneau RJ, Schuchert MJ, Yousem SA *et al.* (2008) Egr-1 regulates autophagy in cigarette smoke-induced chronic obstructive pulmonary disease. *PLoS One* **3**, e3316.
- 54 Weaver LK, Pioli PA, Wardwell K, Vogel SN and Guyre PM (2007) Up-regulation of human monocyte CD163 upon activation of cell-surface Toll-like receptors. *J Leukoc Biol* **81**, 663–671.
- 55 Fabrick BO, van Bruggen R, Deng DM, Ligtenberg AJ, Nazmi K, Schornagel K, Vloet RP, Dijkstra CD and van den Berg TK (2009) The macrophage scavenger receptor CD163 functions as an innate immune sensor for bacteria. *Blood* **113**, 887–892.
- 56 Kaku Y, Imaoka H, Morimatsu Y, Komohara Y, Ohnishi K, Oda H, Takenaka S, Matsuoka M, Kawayama T, Takeya M *et al.* (2014) Overexpression of CD163, CD204 and CD206 on alveolar macrophages in the lungs of patients with severe chronic obstructive pulmonary disease. *PLoS One* **9**, e87400.
- 57 Sy HY, Ko FW, Chu HY, Chan IH, Wong GW, Hui DS and Leung TF (2012) Asthma and bronchodilator responsiveness are associated with polymorphic markers of ARG1, CRHR2 and chromosome 17q21. *Pharmacogenet Genomics* **22**, 517–524.
- 58 Barron L, Smith AM, El Kasmī KC, Qualls JE, Huang X, Cheever A, Borthwick LA, Wilson MS, Murray PJ and Wynn TA (2013) Role of arginase 1 from myeloid cells in Th2-dominated lung inflammation. *PLoS One* **8**, e61961.
- 59 Halim TY, Steer CA, Matha L, Gold MJ, Martinez-Gonzalez I, McNagny KM, McKenzie AN and Takei F (2014) Group 2 innate lymphoid cells are critical for the initiation of adaptive T helper 2 cell-mediated allergic lung inflammation. *Immunity* **40**, 425–435.
- 60 Monticelli LA, Buck MD, Flamar AL, Saenz SA, Tait Wojno ED, Yudanin NA, Osborne LC, Hepworth MR, Tran SV, Rodewald HR *et al.* (2016) Arginase 1 is an innate lymphoid-cell-intrinsic metabolic checkpoint controlling type 2 inflammation. *Nat Immunol* **17**, 656–665.
- 61 Cloots RH, Sankaranarayanan S, de Theije CC, Poynter ME, Terwindt E, van Dijk P, Hakvoort TB, Lamers WH and Kohler SE (2013) Ablation of Arg1 in hematopoietic cells improves respiratory function of lung parenchyma, but not that of larger airways or inflammation in asthmatic mice. *Am J Physiol Lung Cell Mol Physiol* **305**, L364–L376.
- 62 Shi L, Zhu B, Xu M and Wang X (2018) Selection of AECOPD-specific immunomodulatory biomarkers by integrating genomics and proteomics with clinical informatics. *Cell Biol Toxicol* **34**, 109–123.
- 63 Smith NL, Bromley MJ, Denning DW, Simpson A and Bowyer P (2015) Elevated levels of the neutrophil chemoattractant pro-platelet basic protein in macrophages from individuals with chronic and allergic aspergillosis. *J Infect Dis* **211**, 651–660.
- 64 Di Stefano A, Caramori G, Gnemmi I, Contoli M, Bristot L, Capelli A, Ricciardolo FL, Magno F, D'Anna SE, Zanini A *et al.* (2009) Association of increased CCL5 and CXCL7 chemokine expression with neutrophil activation in severe stable COPD. *Thorax* **64**, 968–975.
- 65 Popper HH, Pailer S, Wurzing G, Feldner H, Hesse C and Eber E (2002) Expression of adhesion molecules in allergic lung diseases. *Virchows Arch* **440**, 172–180.
- 66 Bhatt SP, Nath HP, Kim YI, Ramachandran R, Watts JR, Terry NLJ, Sonavane S, Deshmane SP, Woodruff PG, Oelsner EC *et al.* (2018) Centrilobular emphysema and coronary artery calcification: mediation analysis in the SPIROMICS cohort. *Respir Res* **19**, 257.
- 67 Hagiwara K, Kinoshita M, Miyawaki H, Sato S, Miyazaki H, Takeoka S, Suzuki H, Iwaya K, Seki S, Shono S *et al.* (2016) Fibrinogen gamma-chain peptide-coated adenosine 5' diphosphate-encapsulated liposomes rescue mice from lethal blast lung injury via adenosine signaling. *Crit Care Med* **44**, e827–e837.
- 68 Ling SH, McDonough JE, Gosselink JV, Elliott WM, Hayashi S, Hogg JC and van Eeden SF (2011) Patterns of retention of particulate matter in lung tissues of patients with COPD: potential role in disease progression. *Chest* **140**, 1540–1549.