Title

Blood Gene Expression and Immune Cell Subtypes Associated with Chronic Obstructive Pulmonary Disease Exacerbations.

Journal

American Journal of Respiratory and Critical Care Medicine, 208(3)

Authors

Ryu, Min
Yun, Jeong
Morrow, Jarrett
et al.

Publication Date

2023-08-01

DOI

10.1164/rccm.202301-0085OC

Peer reviewed

ORIGINAL ARTICLE

Blood Gene Expression and Immune Cell Subtypes Associated with Chronic Obstructive Pulmonary Disease Exacerbations

Min Hyung Ryu[1,3], Jeong H. Yun[1,2,3], Jarrett D. Morrow[1,3], Aabida Saferali[1,3], Peter Castaldi[1,3], Robert Chase[1], Meryl Stav[1], Zhonghui Xu[1], Igor Barjaktarevic[4], MeiLan Han[5], Wassim Labaki[5], Yvonne J. Huang[5,6], Stephanie Christenson[7], Wanda O'Neal[8], Russell Bowler[9], Don D. Sin[10,11], Christine M. Freeman[5,12], Jeffrey L. Curtis[5,13], and Craig P. Hersh[1,2,3]

[1]Channing Division of Network Medicine and [2]Division of Pulmonary and Critical Care Medicine, Department of Medicine, Brigham and Women's Hospital, Boston, Massachusetts; [3]Department of Medicine, Harvard Medical School, Boston, Massachusetts; [4]Division of Pulmonary and Critical Care Medicine, Department of Medicine, David Geffen School of Medicine, University of California, Los Angeles, California; [5]Division of Pulmonary and Critical Care Medicine and [6]Department of Microbiology and Immunology, University of Michigan, Ann Arbor, Michigan; [7]Division of Pulmonary, Critical Care, Allergy, and Sleep Medicine, University of California, San Francisco, California; [8]Marsico Lung Institute, School of Medicine, University of North Carolina, Chapel Hill, North Carolina; [9]Division of Pulmonary and Critical Care Medicine, Department of Medicine, National Jewish Health, Denver, Colorado; [10]Centre for Heart and Lung Innovation, St. Paul's Hospital, Vancouver, British Columbia, Canada; [11]Division of Respiratory Medicine, Department of Medicine, University of British Columbia, Vancouver, British Columbia, Canada; and [12]Research Service and [13]Medical Service, VA Ann Arbor Healthcare System, Ann Arbor, Michigan

ORCID IDs: 0000-0003-3071-7363 (M.H.R.); 0000-0002-4361-8295 (J.H.Y.); 0000-0002-5670-5865 (J.D.M.); 0000-0002-7497-6597 (Y.J.H.); 0000-0003-4651-363X (R.B.); 0000-0002-1342-4334 (C.P.H.).

Abstract

Rationale: Acute exacerbations of chronic obstructive pulmonary disease (AE-COPDs) are associated with a significant disease burden. Blood immune phenotyping may improve our understanding of a COPD endotype at increased risk of exacerbations.

Objective: To determine the relationship between the transcriptome of circulating leukocytes and COPD exacerbations.

Methods: Blood RNA sequencing data ($n = 3{,}618$) from the COPDGene (Genetic Epidemiology of COPD) study were analyzed. Blood microarray data ($n = 646$) from the ECLIPSE (Evaluation of COPD Longitudinally to Identify Predictive Surrogate Endpoints) study were used for validation. We tested the association between blood gene expression and AE-COPDs. We imputed the abundance of leukocyte subtypes and tested their association with prospective AE-COPDs. Flow cytometry was performed on blood in SPIROMICS (Subpopulations and Intermediate Outcomes in COPD Study) ($n = 127$), and activation markers for T cells were tested for association with prospective AE-COPDs.

Measurements and Main Results: Exacerbations were reported 4,030 and 2,368 times during follow-up in COPDGene (5.3 ± 1.7 yr) and ECLIPSE (3 yr), respectively. We identified 890, 675, and 3,217 genes associated with a history of AE-COPDs, persistent exacerbations (at least one exacerbation per year), and prospective exacerbation rate, respectively. In COPDGene, the number of prospective exacerbations in patients with COPD (Global Initiative for Chronic Obstructive Lung Disease stage ≥ 2) was negatively associated with circulating $CD8^+$ T cells, $CD4^+$ T cells, and resting natural killer cells. The negative association with naive $CD4^+$ T cells was replicated in ECLIPSE. In the flow-cytometry study, an increase in CTLA4 on $CD4^+$ T cells was positively associated with AE-COPDs.

Conclusions: Individuals with COPD with lower circulating lymphocyte counts, particularly decreased $CD4^+$ T cells, are more susceptible to AE-COPDs, including persistent exacerbations.

Keywords: chronic obstructive pulmonary disease; RNA sequencing; immune phenotyping; circulating leukocytes; COPD exacerbation

(Received in original form January 14, 2023; accepted in final form June 6, 2023)

Supported by National Heart, Lung, and Blood Institute grants R01HL124233, R01HL147326, R01HL157879, P01HL114501, U01HL089856, U01HL089897, and K25 HL136846. The COPDGene study (NCT00608764) is also supported by the COPD Foundation through contributions made to an Industry Advisory Committee composed of AstraZeneca, Bayer Pharmaceuticals, Boehringer-Ingelheim, Genentech, GlaxoSmithKline, Novartis, Pfizer, and Sunovion.

Author Contributions: Concept and design: M.H.R., J.L.C, and C.P.H.; acquisition or processing of the data: J.H.Y., J.D.M., A.S., P.C., R.C., M.S., Z.X., I.B., M.H., W.L., Y.J.H., S.C., W.O'N., R.B., D.D.S., C.M.F., J.L.C., and C.P.H.; data analysis: M.H.R., J.H.Y., J.D.M., A.S., P.C., and C.P.H; statistical analysis and support: M.H.R., J.H.Y., J.D.M., A.S., P.C., and C.P.H.; manuscript writing, draft: M.H.R. and C.P.H. All authors reviewed, edited, and approved the final manuscript.

Am J Respir Crit Care Med Vol 208, Iss 3, pp 247–255, Aug 1, 2023

Originally Published in Press as DOI: 10.1164/rccm.202301-0085OC on June 7, 2023
Internet address: www.atsjournals.org

At a Glance Commentary

Current Scientific Knowledge on the Subject: Acute exacerbations of chronic obstructive pulmonary disease (COPD) are associated with a significant disease burden. Several blood-based predictors of future exacerbations have been identified, and changes in cellular composition in blood have been associated with COPD exacerbation in retrospective studies.

What This Study Adds to the Field: This omics-based investigation provides robust data unveiling the relationship between blood-based molecular and cellular phenotypes with COPD exacerbations, an important clinical endpoint. The work highlights that individuals with COPD who have lower circulating lymphocytes, particularly decreased $CD4^+$ T cells, are more susceptible to acute exacerbation of COPD, including persistent exacerbations.

Chronic obstructive pulmonary disease (COPD) is a highly prevalent, often progressive, and complex chronic disease of the lungs with immune-mediated aspects (1). Respiratory "flare-ups" accompanied by acute worsening symptoms requiring additional therapy, or acute exacerbations of COPD (AE-COPDs), are associated with increased morbidity, mortality, and medical costs (2–4). AE-COPDs are more frequent and more severe as the severity of COPD increases (5) and are associated with accelerated lung function loss in COPD, particularly in those with early/mild COPD (6). The frequency of AE-COPDs is variable from year to year (7), but an increased rate may represent an exacerbation-susceptible endotype. The best predictor of future exacerbations across all severities of COPD is a history of exacerbations (5, 8). Defining and understanding an exacerbation-susceptible endotype is an active area of research because it holds great potential to be translated into tools for improving clinical trials for new therapies and to be used in the long-term management of COPD.

Omics-based investigation in large longitudinal cohorts may provide important molecular insight into COPD endotypes and increase our understanding of susceptibility factors for important clinical endpoints such as AE-COPDs (9). Here, we report the results from blood transcriptome analysis in the COPDGene (Genetic Epidemiology of COPD) study to investigate the association between blood gene expression and AE-COPDs. Using whole-blood RNA sequencing (RNA-seq), we quantified mRNAs expressed in whole blood from former and current smokers with and without COPD from 21 centers across the United States. We provide further support to our RNA-seq findings with replication using the RNA microarray data from the ECLIPSE (Evaluation of COPD Longitudinally to Identify Predictive Surrogate Endpoints) study and the blood immune phenotyping study in SPIROMICS (Subpopulations and Intermediate Outcomes in COPD Study). Therefore, our study provides robust data to unveil the relationship of blood-based molecular and cellular phenotypes with COPD exacerbations. We hypothesized that blood gene expression and abundance of circulating leukocytes are associated with prospective AE-COPDs.

Some of the results of these studies have been previously reported in the form of an abstract (10).

Methods

Study Populations

Whole-blood transcriptome and prospective exacerbation data were collected in COPDGene (clinicaltrials.gov ID NCT00608764) and ECLIPSE (clinicaltrials.gov ID NCT00292552). Detailed descriptions of the study populations are provided in the online supplement.

COPDGene is a multicenter prospective study that enrolled 10,192 smokers with and without COPD in phase 1 (11). At enrollment, COPDGene participants were between 45 and 80 years old and had $\geq$10 pack-years of cigarette-smoking history. Here, blood RNA-seq data from 3,618 former and current smokers collected at phase 2 (5-yr follow-up visit) were analyzed.

ECLIPSE was a multicenter 3-year longitudinal prospective study that enrolled 2,747 participants (12). Whole-blood RNA microarray and 3-year prospective AE-COPD data were obtained from 646 study subjects in the ECLIPSE study (5, 13). ECLIPSE participants included in our analysis had an FEV_1/FVC ratio lower than 0.7 and $FEV_1 < 80\%$ predicted. All subjects in ECLIPSE had $\geq$10 pack-years of cigarette-smoking history.

Blood flow cytometry data from the SPIROMICS bronchoscopy substudy were analyzed. Details of this study were published previously (7, 14, 15). All subjects included in our SPIROMICS analysis had $\geq$20 pack-years of cigarette-smoking history.

All studies were approved by the institutional review boards at the participating centers, and all participants provided written informed consent.

Exacerbation Events

An acute exacerbation event was defined as any health care use (office visit, hospital admission, or emergency department visit) for a respiratory flare-up that involved treatment with antibiotics and/or systemic corticosteroids (6, 16). History of exacerbation (i.e., retrospective exacerbations) during the 1 year before the study visit was assessed by questionnaires.

Prospective exacerbations were documented during longitudinal follow-up for 5.3 ± 1.7 years in COPDGene, 3 years in ECLIPSE, and 1 year in SPIROMICS. Persistent exacerbations were noted in those with at least one prospective exacerbation per year (7). Data from each cohort were analyzed independently to account for differences in follow-up duration.

Gene-Expression Analysis

Detailed methods are provided in the online supplement. After quality control, RNA-seq data from 3,618 COPDGene subjects were available for analysis. ECLIPSE whole-blood

Correspondence and requests for reprints should be addressed to Craig P. Hersh, M.D., M.P.H., Channing Division of Network Medicine, Brigham and Women's Hospital, 181 Longwood Avenue, Boston, MA 02115. Email: craig.hersh@channing.harvard.edu.

This article has a related editorial.

This article has an online supplement, which is accessible from this issue's table of contents at www.atsjournals.org.

Table 1. Study Subjects in Gene Expression Analyses

Characteristic	COPDGene (*N* = 3,618)	ECLIPSE (*N* = 646)
Age, yr	64.90 [58.20–71.80]	64.00 [59.00–69.00]
Female sex	1,767 (48.8)	222 (34.4)
Race		
Non-Hispanic White	2,615 (72.3)	638 (98.8)
African American	1,003 (27.7)	8 (1.2)
BMI, kg/m^2, median [IQR]	28.08 [24.57–32.28]	26.68 [23.36–29.98]
Current smoker	1,360 (37.6)	160 (24.8)
Smoking history, pack-years	40.00 [27.50–54.80]	41.00 [28.00–56.00]
COPD status		
PRISm	473 (13.1)	0
Normal spirometry	1,571 (43.4)	0
GOLD stage		
1	363 (10.0)	0
2	703 (19.4)	270 (46.1)
3	360 (10.0)	243 (41.5)
4	148 (4.1)	73 (12.5)
FEV_1/FVC	0.72 [0.61–0.78]	0.46 [0.37–0.57]
FEV_1% predicted after bronchodilation	82.00 [63.30–95.77]	50.25 [37.88–65.65]
History of exacerbation	616 (17.0)	256 (39.6)
History of severe exacerbation	328 (9.1)	83 (12.8)
Persistent exacerbators	291 (8.9)*	278 (43.0)
Prospective exacerbation rate per year	0.00 [0.00–0.25]	0.67 [0.00–2.00]

Definition of abbreviations: COPD = chronic obstructive pulmonary disease; COPDGene = Genetic Epidemiology of COPD; ECLIPSE = Evaluation of COPD Longitudinally to Identify Predictive Surrogate Endpoints; PRISm = preserved ratio impaired spirometry (FEV_1/FVC $\geq$ 0.70, FEV_1% predicted $<$80%); GOLD = Global Initiative for Chronic Obstructive Lung Disease; NA = not available.

Values presented as number (percentage) or median [IQR]. History of exacerbation refers to self-reported exacerbation during 1 yr before study visit. Severe exacerbation was defined as any exacerbation requiring an emergency room visit or hospital admission. Persistent exacerbators were defined as those who have at least one exacerbation per year in the follow-up years.

*Excluding 356 subjects without longitudinal follow-up data.

samples were assayed using the Affymetrix Human Gene 1.1 ST array, and microarray processing details were reported previously (13). Microarray data from 646 participants were available. For differential gene expression analysis, we used limma and voom; this approach analyzes entire experiments as an integrated whole with linear models with covariate adjustment and is computationally efficient in epidemiologic-scale datasets (17). Adjustment for multiple testing controlled for a false discovery rate (FDR) $<$5%.

Cellular Deconvolution

Cellular deconvolution is a computational method that uses reference single-cell RNA-seq data or bulk RNA-seq data from purified cells to estimate the abundance of specific cell types in bulk RNA-seq data. Cellular deconvolution was performed using CIBERSORTx with the LM22 reference (18). For the imputation of cell fraction, we computed the abundance scores of 22 hematopoietic populations (cells listed in the online supplement). For the imputation of cell type-specific gene expression, we computed gene expression of 10 major leukocyte types (eosinophils, dendritic cells, mast cells, monocytes, neutrophils, plasma cells, natural killer [NK] cells, $CD4^+$ and $CD8^+$ T cells, and B cells).

SPIROMICS Flow-Cytometry Data

Blood samples in the SPIROMICS bronchoscopy substudy were analyzed by flow cytometry to determine percentages of $CD4^+$ and $CD8^+$ cells in $CD45^+$ cells and their activation markers (i.e., CCR5, CCR7, CTLA4, CXCR3, CD28, and ICOS) (15).

Statistical Analyses

Detailed methods for statistical analyses are included in the online supplement. To test the association between cell proportion scores and prospective exacerbations, zero-inflated Poisson regression analyses were performed using covariates listed in the table legends. *P* values were corrected for an FDR of 5%.

To test the association between cell type–specific gene expression and prospective exacerbation rate, we used linear regression models adjusted for sex, age, race, smoking status, smoking pack-years, FEV_1% predicted, exacerbation history, microarray batch (only in ECLIPSE), and white blood cell counts. *P* values were corrected for an FDR of 5%.

Results

Exacerbation Events

Table 1 summarizes participant characteristics and exacerbation events in COPDGene and ECLIPSE. In COPDGene, 4,030 prospective exacerbations were reported among 3,262 subjects with RNA-seq data (excluding those with missing data for phase 2 spirometry, complete blood count [CBC] with differentials, or longitudinal follow-up data). In ECLIPSE, 2,368 prospective exacerbations were reported in the 3 years of follow-up among 646 participants. Table E1 summarizes the SPIROMICS study subjects' characteristics and exacerbation variables.

Gene-Expression Associated with Exacerbation Events

Whole-blood RNA-seq data from 3,618 subjects in the COPDGene study were analyzed (Table 1 and Figure E1 in the online supplement). We identified differentially expressed genes (DEGs) associated with the history of exacerbation (at least one exacerbation in the 1 yr before the study visit), persistent exacerbator status (281 participants with at least one exacerbation per year during the longitudinal follow-up), and prospective exacerbation rate (defined as log [total number of exacerbations + 0.5] – log[years followed + 1]) (Figure 1 and Tables E2–E4). The prospective exacerbation rate was log-transformed to account for varying follow-up duration. Of the DEGs, 214 were associated with all three exacerbation phenotypes (Figure 2A and Table E5).

Enriched Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways were identified in Signature Overrepresentation Analysis (19). There were 20, 9, and 67 KEGG pathways overrepresented by the DEGs associated with a history of exacerbation, persistent exacerbations, and

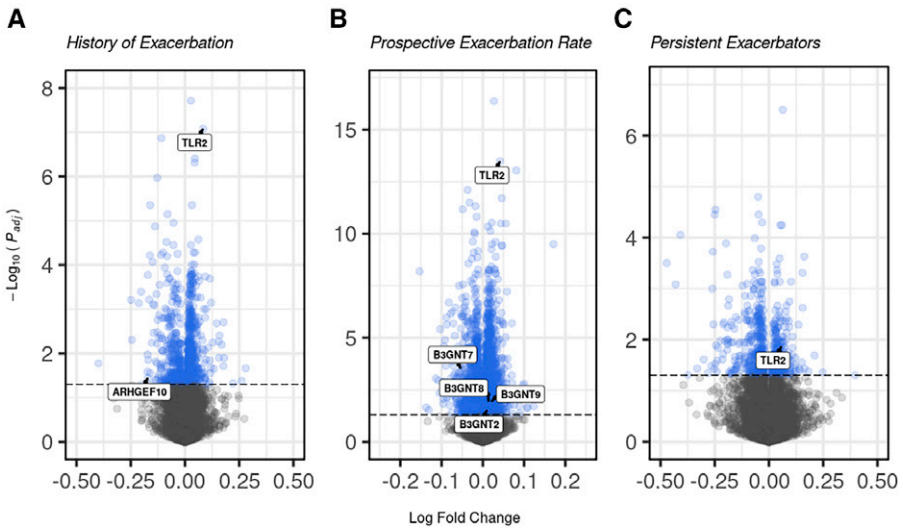

Figure 1. Volcano plots for differentially expressed genes in COPDGene (Genetic Epidemiology of COPD) whole-blood RNA-sequencing analysis. (*A*) History of exacerbation was defined as having one or more exacerbations in the year before the study visit. (*B*) Prospective exacerbation rate was defined as log(total number of exacerbation + 0.5) − log(years followed + 1). (*C*) Persistent exacerbators were participants with at least one prospective exacerbation per year in the follow-up years. For differential gene-expression analyses, linear models were adjusted for age, race, sex, current smoking status, smoking pack-years, library generation batch, and surrogate variables for technical variation.

prospective exacerbation rate, respectively (Figure 2B and Table E6).

The differential gene-expression analyses indicated that the cell composition of blood and COPD severity (as measured by FEV_1% predicted) accounted for a significant portion of the association between blood gene expression and the exacerbation phenotypes. This is because, when we additionally adjusted for FEV_1% predicted and blood leukocyte abundance and their relative percentages in our statistical models, a smaller number of DEGs remained significant (*see* online supplement).

In ECLIPSE, there were 36 DEGs associated with the prospective exacerbation rate. Of the 36 DEGs, 16 genes were also significantly associated with prospective exacerbation rate in COPDGene (*see* online supplement).

Persistent Exacerbators Were Characterized by an Increase in Plasma Cells and a Decrease in $CD4^+$ T-Cell Subtypes in the Blood

We optimized the cell-type deconvolution by benchmarking imputed cell fraction to CBC counts for neutrophils and monocytes (*see* online supplement). Of the 22 mature hematopoietic cell populations queried in the COPDGene and ECLIPSE datasets, 10 immune cell populations (neutrophils, monocytes, $CD8^+$ T cells, naive $CD4^+$ T cells, resting memory $CD4^+$ T cells, activated memory $CD4^+$ T cells, resting NK cells, naive B cells, memory B cells, and plasma cells) were successfully imputed in both COPDGene and ECLIPSE. Tables E7 and E8 summarize the mean imputed

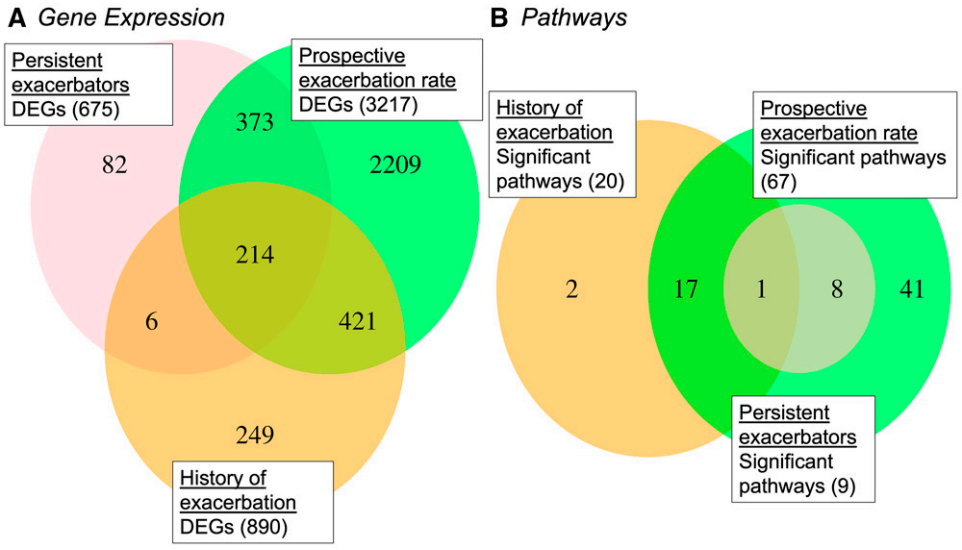

Figure 2. Numbers of genes and pathways significantly associated with exacerbation measures. (*A*) Venn diagram showing overlap between differentially expressed genes in the three separate analyses. Persistent exacerbators were participants with at least one prospective exacerbation per year in the follow-up years. History of exacerbation was defined as having one or more exacerbations in the year prior to the study visit. Prospective exacerbation rate was defined as log(total number of exacerbation + 0.5) − log(years followed + 1). For differential gene expression analyses, linear models were adjusted for age, race, sex, current smoking status, smoking pack-years, library generation batch, and surrogate variables for technical variation. (*B*) Venn diagram showing the number of overlaps between significantly enriched Kyoto Encyclopedia of Genes and Genomes pathways identified in Signature Overrepresentation Analysis. DEG = differentially expressed gene.

abundance scores stratified by having one or more prospective exacerbations. In COPDGene, persistent exacerbators were characterized by an increase in plasma cells ($P < 0.001$) and a decrease in naive $CD4^+$ T cells ($P = 0.006$) and resting memory $CD4^+$ T cells ($P < 0.001$) (*see* Table E7). In the ECLIPSE study, persistent exacerbators were characterized by increased plasma cells ($P = 0.04$) and decreased resting memory $CD4^+$ T cells ($P < 0.001$) (*see* Table E8).

The Number of Prospective Exacerbations Was Associated with the Abundance of Blood Lymphocyte Subtypes

Table 2 summarizes the association between the number of prospective exacerbations and blood lymphocyte subsets. In COPDGene, the number of prospective exacerbation events was negatively associated with $CD8^+$ T cells (incident risk ratio for each 1-SD increase [95% CI], 0.87 [0.83–0.91]; $P < 0.001$), activated memory $CD4^+$ T cells (0.82 [0.80–0.85]; $P < 0.001$), and resting NK cells (0.70 [0.68–0.73]; $P < 0.001$). The number of prospective exacerbation events was positively associated with naive B cells (1.10 [1.05–1.14]; $P < 0.001$). Including only those with COPD (Global Initiative for Chronic Obstructive Lung Disease stage 2–4), the number of prospective exacerbation events was additionally associated with decreases in naive $CD4^+$ T cells (0.91 [0.86–0.95]; $P < 0.001$) and memory B cells (0.83 [0.80–0.87]; $P < 0.001$). However, the association between the number of prospective exacerbations and naive B cells was no longer significant in the COPD-only analysis.

In ECLIPSE, the number of prospective exacerbations was positively associated with plasma cells (1.08 [1.01–1.19]; $P < 0.001$) and $CD8^+$ T cells (1.11 [1.04–1.18]; $P < 0.001$). The total number of exacerbations was negatively associated with naive $CD4^+$ T cells (0.91 [0.85–0.97]; $P < 0.001$) and resting memory $CD4^+$ T cells (0.89 [0.83–0.96]; $P < 0.001$).

In COPDGene, eosinophil and neutrophil counts from CBCs were associated with an increased risk of prospective exacerbations, whereas monocytes and lymphocytes were associated with a decreased risk (Table E9). These four cell types were all positively associated in ECLIPSE, but the effects were smaller.

Cell Type–Specific Gene Expression in Neutrophils and $CD8^+$ and $CD4^+$ T Cells Were Associated with Prospective Exacerbation Rate

Cell type–specific gene expression was imputed for 10 major blood cell types: neutrophils, eosinophils, $CD4^+$ and $CD8^+$ T cells, B cells, NK cells, monocytes, dendritic cells, plasma cells, and mast cells. Table 3 summarizes the variable genes in cell type–specific transcriptome and cell type–specific gene expression associated with prospective exacerbation rates in COPDGene and ECLIPSE. There were 32 neutrophil-specific, 41 $CD4^+$ T cell–specific, 10 $CD8^+$

Table 2. Association of Blood Leukocyte Subsets and Prospective Exacerbations

	IRR (95% CI)	*P*	Adjusted *P*
COPDGene ($n = 3,188$)			
$CD8^+$ T cells	0.87 (0.83–0.91)	<0.001	<0.001
Naive $CD4^+$ T cells	1.00 (0.97–1.04)	0.91	0.91
Resting memory $CD4^+$ T cells	0.96 (0.92–1.00)	0.05	0.08
Activated memory $CD4^+$ T cells	0.82 (0.79–0.85)	<0.001	<0.001
Resting NK cells	0.70 (0.68–0.73)	<0.001	<0.001
Naive B cells	1.10 (1.05–1.14)	<0.001	<0.001
Memory B cells	0.97 (0.93–1.00)	0.07	0.09
Plasma cells	0.98 (0.95–1.00)	0.10	0.12
COPDGene; GOLD 2–4 ($n = 1,066$)			
$CD8^+$ T cells	0.91 (0.86–0.96)	0.001	0.002
Naive $CD4^+$ T cells	0.91 (0.86–0.95)	<0.001	<0.001
Resting memory $CD4^+$ T cells	1.06 (1.00–1.12)	0.04	0.05
Activated memory $CD4^+$ T cells	0.89 (0.85–0.92)	<0.001	<0.001
Resting NK cells	0.70 (0.67–0.74)	<0.001	<0.001
Naive B cells	1.01 (0.96–1.07)	0.69	0.69
Memory B cells	0.83 (0.80–0.87)	<0.001	<0.001
Plasma cells	1.01 (0.98–1.04)	0.60	0.68
ECLIPSE ($N = 646$)			
$CD8^+$ T cells	1.11 (1.04–1.18)	<0.001	<0.001
Naive $CD4^+$ T cells	0.91 (0.85–0.97)	<0.001	<0.001
Resting memory $CD4^+$ T cells	0.89 (0.83–0.96)	<0.001	<0.001
Activated memory $CD4^+$ T cells	1.04 (0.96–1.11)	0.10	0.14
Resting NK cells	1.04 (0.98–1.14)	0.14	0.16
Naive B cells	0.96 (0.88–1.03)	0.11	0.14
Memory B cells	1.01 (0.95–1.06)	0.65	0.65
Plasma cells	1.08 (1.01–1.19)	<0.001	<0.001

Definition of abbreviations: CI = confidence interval; COPDGene = Genetic Epidemiology of COPD; ECLIPSE = Evaluation of COPD Longitudinally to Identify Predictive Surrogate Endpoints; IRR = incident risk ratio; NK = natural killer.
IRR is shown for each 1-SD increase of each predictor (abundance score from CIBERSORTx). Zero-inflated Poisson models were adjusted for sex, race, age, body mass index, smoking status, pack-years of smoking, FEV_1% predicted, exacerbation history in the 1 yr before the study visit, neutrophil count, and eosinophil count. For COPDGene, the number of surveys done in the longitudinal follow-up was used as an offset in the zero-inflated models. For the ECLIPSE dataset, 95% CI represents bootstrap bias-corrected CI calculated using 1,200 replicates.

Table 3. Cell Type–Specific Gene Expression Associated with Prospective Exacerbation Rates in COPDGene and ECLIPSE

	COPDGene		ECLIPSE		Replication Results	
	Variable Genes in Cell Type–Specific Transcriptome	Cell Type–Specific Genes Associated with Prospective Exacerbation Rates	Variable Genes in Cell Type–Specific Transcriptome	Cell Type–Specific Genes Associated with Prospective Exacerbation Rates	Replicated Cell Type–Specific Genes Associated with Prospective Exacerbation Rates	Replicated Cell Type–Specific Genes Also in the List of 214 Genes Associated with All 3 Exacerbation Phenotypes
Neutrophils	12,080	62 up, 763 down	10,804	437 up, 324 down	3 up, 29 down	↑TLR2
$CD4^+$ T cells	10,451	192 up, 2,191 down	5,944	115 up, 258 down	3 up, 38 down	↓*B3GALT2*, ↑*BTLA*, ↓CCR6, ↓*ZBTB25*
B cells	4,188	2 up, 2 down	5,729	25 up, 16 down	—	—
NK cells	3,451	113 up, 48 down	7,645	74 up, 31 down	3 up	—
$CD8^+$ T cells	2,881	28 up, 60 down	13,139	466 up, 1,438 down	10 down	—
Monocytes	3,130	24 up, 18 down	4,994	5 up	—	—
Dendritic cells	924	13 up, 1 down	508	17 up, 2 down	—	—
Plasma cells	1,590	14 up	1,342	—	—	—
Mast cells	1,793	—	1,046	—	—	—
Eosinophils	82	—	76	—	—	—

Definition of abbreviations: COPDGene = Genetic Epidemiology of COPD; ECLIPSE = Evaluation of COPD Longitudinally to Identify Predictive Surrogate Endpoints; NK = natural killer.

Cell-type–specific gene expression of 10 major leukocytes (eosinophils, dendritic cells, mast cells, monocytes, neutrophils, plasma cells, NK cells, $CD4^+$ and $CD8^+$ T cells, and B cells) were queried using CIBERSORTx. The COPDGene count matrix included 25,171 genes, and the ECLIPSE count matrix included 19,526 genes. In testing significant association with prospective exacerbation rate, linear models adjusted for sex, age, race, smoking status, smoking pack-years, FEV_1% predicted, exacerbation history, microarray batch ID (only in ECLIPSE), and white blood cell counts.

T cell–specific, and 3 NK cell–specific gene-expression levels associated with prospective exacerbation in COPDGene that were replicated in ECLIPSE (Table E10).

Upregulation of Cell-Surface Activation Markers in $CD4^+$ and $CD8^+$ T Cells Was Associated with Exacerbation Events

Blood-flow cytometry data from SPIROMICS were analyzed to test the association between T cells and prospective exacerbations (during the 1 yr after the sample collection). In this dataset, the percentages of $CD4^+$ and $CD8^+$ cells within the $CD45^+$ cell population were negatively associated with prospective exacerbation events, adjusting for sex, age, current smoking status, and FEV_1% predicted (Table 4). This association was still significant after additional adjustment for exacerbation history (Table E11). There was a trend toward significance for decreased percent $CD4^+$ cells in COPD compared with those without COPD ($P = 0.051$).

Expression of CTLA4 (cytotoxic T-lymphocyte antigen 4), as measured by mean fluorescent intensity values, on $CD4^+$ and $CD8^+$ cells was positively associated with the number of prospective exacerbations. Mean fluorescent intensity for CCR5 in $CD8^+$ T cells was positively associated with prospective exacerbations. These associations were still significant after additional adjustments for exacerbation history.

Discussion

We report in this multicenter, prospective observational study the blood gene expression and immune cell subtypes associated with COPD exacerbations. In COPDGene, we identified more than 3,500 unique genes in the blood transcriptome associated with at least one of three exacerbation measures: a history of exacerbation, persistent exacerbations, and prospective exacerbation rate. Among these genes, 214 were associated with all three exacerbation measures. Using signature overrepresentation analysis, we identified 69 enriched KEGG pathways, and the NOD-like receptor signaling pathway was associated with all three exacerbation phenotypes. Using cellular deconvolution of whole-blood bulk RNA-seq, we determined the abundance of eight immune cell subpopulations that were not captured by clinical CBC differentials. We

Table 4. Association between Prospective Exacerbation and T Cells and Their Activation Markers

Marker	IRR (95% CI)	Adjusted *P*
Percent $CD4^+$ cells in $CD45^+$ ($n = 125$)*	0.16 (0.06–0.46)	0.001
Percent $CD8^+$ cells in $CD45^+$ ($n = 126$)*	0.31 (0.13–0.70)	0.005
$CD4^+$ CCR5 MFI ($n = 127$)†	3.02 (0.85–10.7)	0.087
$CD4^+$ CTLA4 MFI ($n = 127$)†	3.45 (1.48–8.01)	0.004
$CD8^+$ CCR5 MFI ($n = 123$)†	4.23 (1.06–16.9)	0.041
$CD8^+$ CTLA4 MFI ($n = 124$)†	2.91 (1.04–8.11)	0.042

Definition of abbreviations: CI = confidence interval; IRR = incident risk ratio; MFI = mean fluorescence intensity.
IRRs are shown for each 1-SD increase of each predictor.
*Zero-inflated Poisson model adjusted for sex, age, current smoking status, and FEV_1% predicted.
†Zero-inflated Poisson model adjusted for sex, age, and control fluorescent bead batch.

found that the number of prospective exacerbations in subjects with COPD (Global Initiative for Chronic Obstructive Lung Disease stage ⩾2) was negatively associated with $CD8^+$ T cells, naive and activated memory $CD4^+$ T cells, and resting NK cells in COPDGene.

Our results extend the findings of previous works associating frequent exacerbations (at least two exacerbations per year) with gene expression in ECLIPSE ($n = 438$) (20) and in our previous analysis in the TESRA (Treatment of Emphysema with a γ-Selective Retinoid Agonist) study ($n = 248$) (21). Singh and colleagues found *B3GNT*, *LAF4*, and *ARHGEF10* to predict frequent exacerbators in the ECLIPSE study (20). In line with ECLIPSE, we found genes in the β 3-glycosyltransferase family (i.e., *B3GNT7*, *B3GNT9*, *B3GNT8*, and *B3GNT2*) associated with prospective exacerbation rates (*see* Table E4), and *ARHGEF10* was associated with a history of exacerbation (*see* Table E2) in COPDGene. However, *LAF4* was not found to be associated with any one of the three exacerbation measures in COPDGene. Similar to KEGG pathways associated with an exacerbation gene module in the TERSA study (21), we found KEGG pathways for B-cell receptor signaling and allograft rejection to be enriched by genes associated with AE-COPDs. Our findings suggest that altered B-cell function and autoimmunity may play a role in increasing the risk of future exacerbations.

We examined the association between the abundance of blood leukocytes and exacerbation risk prospectively. Consistent with previous studies (22, 23), blood neutrophil and eosinophil counts were positively associated with increased risk of AE-COPD in COPDGene and ECLIPSE. This finding was in contrast with previous SPIROMICS findings that stratification only by sputum eosinophil, and not by blood eosinophil count, was significantly associated with increased exacerbations (24). However, the present study focused on prospective exacerbation, rather than retrospective exacerbation, which was evaluated in SPIROMICS.

We found monocyte and lymphocyte counts to be negatively associated with AE-COPDs in COPDGene, but this association was opposite in direction in ECLIPSE. Monocyte and lymphocyte abundances are composites of several distinctive cell subtypes. Therefore, the overall direction of association with exacerbation may not be consistent when considering these cell types in broader categories. Nevertheless, a decrease in circulating naive $CD4^+$ T cells was associated with an increased risk of exacerbations in COPDGene and ECLIPSE.

Our $CD4^+$ finding is in agreement with a previous smaller study by Freeman and colleagues in which participants with COPD were immune-phenotyped in multiple stable study visits and during AE-COPDs; AE-COPDs were associated with decreased $CD4^+$ and $CD8^+$ T cells in the blood (25). Indeed, in our SPIROMICS flow-cytometry analysis, subjects with a lower percentage of $CD4^+$ T cells were associated with a higher 1-year prospective exacerbation risk. Lower $CD4^+$ T-cell numbers were accompanied by an increased activation marker, CTLA4, which is also a negative regulator of T-cell activation to balance the consequences of T-cell activation (26). This observation is in line with our finding that the majority of cell type–specific gene expression in $CD4^+$ T cells was negatively associated with prospective exacerbation rate. Our results show that individuals with lower circulating lymphocytes, particularly decreased $CD4^+$ T cells, are more susceptible to AE-COPDs, including persistent exacerbations. These results suggest that relative immunodeficiency may contribute to this important endotype. It would be informative to test if interventions boosting T cells would decrease the risk of AE-COPDs.

To date, several blood-based predictors of future exacerbations have been identified. In the ECLIPSE study, the exacerbation-susceptible phenotype was associated with increased white cell count (5). In addition to leukocyte count, blood levels of surfactant protein D, fetuin-A, adiponectin, CRP, and fibrinogen were predictive of future exacerbations (27–31). The increase in neutrophil-to-lymphocyte ratio was associated with AE-COPDs in numerous retrospective studies (32), and the neutrophil-to-lymphocyte ratio was a predictor of AE-COPDs (33).

SPIROMICS participants with consistent exacerbations (at least one exacerbation every year during the 3 yr of follow-up) had lower circulating levels of IL-15, which regulates the activation and proliferation of T cells (7). Increased IL-8, a cytokine for neutrophil recruitment and degranulation, was also associated with consistent exacerbators in SPIROMICS. In COPDGene and ECLIPSE, we found that neutrophils and $CD4^+$ T cells had the most cell type–specific genes associated with prospective exacerbation rates, suggesting that an exacerbation-susceptible endotype may be defined by alterations in neutrophils and $CD4^+$ T-cell abundance and functions.

We found that *TLR2* expression in neutrophils was positively associated with exacerbation rates in COPDGene and ECLIPSE, and *TLR2* expression level in the bulk RNA-seq was associated with all three exacerbation measures in COPDGene. Toll-like receptor 2, encoded by *TLR2*, is a pattern-recognition molecule involved in recognition of pathogen-associated molecular patterns and is an important regulator of neutrophil functions (34). The exacerbation-susceptible endotype might be characterized by the enhanced neutrophil response to pathogen-associated molecular patterns upon exposure to pathogens.

Strengths of our study include the large, multicenter sample and longitudinal study

design that captured exacerbation data over a 5-year period. Our study included large numbers of women and Black participants, which allows for adjustment for these important groups, which may influence the phenotypic expression of COPD (35). Our definition of AE-COPDs was rigorous in that we defined an exacerbation event as any healthcare utilization that involved additional treatment with antibiotics, systemic corticosteroids, or both. Moreover, we used the same definition for COPDGene, ECLIPSE, and SPIROMICS. Our cohorts included a large number of healthy controls, participants with mild COPD, and subjects with preserved-ratio impaired lung function (36).

Therefore, we captured exacerbation frequency in a wide range of spirometry values. However, using our definition, we might not have captured subjects with milder exacerbation events that were left untreated. Also, the severity of AE-COPDs was not recorded based on the extent of underlying physiological derangement, limiting our ability to investigate the relationship between blood gene expression and the severity of AE-COPDs instead of their frequency. Moreover, there is a possibility that factors such as COPD severity or socioeconomic status may have impacted the health-seeking decision that can impact the frequency of exacerbation captured (37). We also cannot rule out that some of our observations may have been the result of differences in treatment for COPD, such as the use of inhaled corticosteroids in those with frequent exacerbation.

With respect to deconvolution methodology, our method could not reliably impute cell type–specific gene expression of low-abundance cell types (e.g., eosinophils). Rare cell types may play important roles in the exacerbation-susceptible endotype, which will be missed by our deconvolution-based RNA-seq immune phenotyping. To overcome this technical issue, it is advisable to enrich these rare cell types in future studies using fluorescence-activated cell sorting (FACS) in combination with single-cell RNA-seq. Nevertheless, at the present time, immune phenotyping by computational deconvolution of whole-blood RNA-seq is advantageous over FACS or scRNA-seq because FACS or scRNA-seq analyses of large sample cohorts ($N > 1{,}000$) are not yet practical, and the impact of sample processing during FACS and scRNA-seq on the transcriptome is not fully understood.

One potential reason for poor replication of cell type–specific gene expression between COPDGene and ECLIPSE (Table 3) is the difference in gene-expression assays. COPDGene used RNA-seq, and ECLIPSE used microarrays. We suspect that ECLIPSE deconvolution overestimated $CD8^+$ T cells at the cost of underestimating $CD4^+$ cells. This may have been because $CD8^+$ T cells are transcriptionally similar to $CD4^+$ T cells (38), and deconvolution in microarray data may not be as robust as in RNA-seq data. This may explain the difference in the number of variable genes in the cell type–specific transcriptomes.

In conclusion, immune phenotyping using blood RNA-seq data found that individuals with lower circulating lymphocytes, particularly decreased $CD4^+$ T cells, are more susceptible to AE-COPDs, including persistent exacerbations. We validated the key associations between lymphocyte subtypes and exacerbation rate in ECLIPSE. In the SPIROMICS flow-cytometry study, alteration in the circulating $CD4^+$ T cell population was associated with an increased 1-year exacerbation risk. Neutrophils and $CD4^+$ T cells had the most cell type–specific genes associated with prospective exacerbation rates. Our findings support that circulating immune cells, specifically neutrophils and $CD4^+$ T cells, and their gene expression may be a potential biomarker for AE-COPDs, but further research and validation are needed. ■

Author disclosures are available with the text of this article at www.atsjournals.org.

Acknowledgment: The authors thank the COPDGene, ECLIPSE, and SPIROMICS participants and participating physicians, investigators, and staff for making this research possible. More information about the studies and how to access COPDGene and SPIROMICS data are available at http://www.copdgene.org/ and www.spiromics.org, respectively. Full acknowledgments for COPDGene and SPIROMICS are provided in the online supplement.

References

1. Bhat TA, Panzica L, Kalathil SG, Thanavala Y. Immune dysfunction in patients with chronic obstructive pulmonary disease. *Ann Am Thorac Soc* 2015;12:S169–S175.
2. Seemungal TAR, Donaldson GC, Paul EA, Bestall JC, Jeffries DJ, Wedzicha JA. Effect of exacerbation on quality of life in patients with chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 1998;157:1418–1422.
3. Singanayagam A, Schembri S, Chalmers JD. Predictors of mortality in hospitalized adults with acute exacerbation of chronic obstructive pulmonary disease. *Ann Am Thorac Soc* 2013;10:81–89.
4. Global Initiative for Chronic Obstructive Lung Disease. Global strategy for the diagnosis, management, and prevention of COPD: 2023 report. Fontana, WI: Global Initiative for Chronic Obstructive Lung Disease; 2022 [accessed 2022 Dec 8]. Available from: https://goldcopd.org/2023-gold-report-2/.
5. Hurst JR, Vestbo J, Anzueto A, Locantore N, Müllerova H, Tal-Singer R, *et al.*; Evaluation of COPD Longitudinally to Identify Predictive Surrogate Endpoints (ECLIPSE) Investigators. Susceptibility to exacerbation in chronic obstructive pulmonary disease. *N Engl J Med* 2010;363:1128–1138.
6. Dransfield MT, Kunisaki KM, Strand MJ, Anzueto A, Bhatt SP, Bowler RP, *et al.*; COPDGene Investigators. Acute exacerbations and lung function loss in smokers with and without chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2017;195:324–330.
7. Han MK, Quibrera PM, Carretta EE, Barr RG, Bleecker ER, Bowler RP, *et al.*; SPIROMICS Investigators. Frequency of exacerbations in COPD: an analysis of the SPIROMICS cohort. *Lancet Respir Med* 2017;5:619–626.
8. Hurst JR, Anzueto A, Vestbo J. Susceptibility to exacerbation in COPD. *Lancet Respir Med* 2017;5:e29.
9. Regan EA, Hersh CP, Castaldi PJ, DeMeo DL, Silverman EK, Crapo JD, *et al.* Omics and the search for blood biomarkers in chronic obstructive pulmonary disease. Insights from COPDGene. *Am J Respir Cell Mol Biol* 2019;61:143–149.
10. Ryu MH, Yun JH, Morrow J, Saferali A, Castaldi P, Chase R, *et al.* Blood gene expression and immune cell subtypes associated with COPD exacerbations. *Am J Respir Crit Care Med* 2023;207:A4809.
11. Regan EA, Hokanson JE, Murphy JR, Make B, Lynch DA, Beaty TH, *et al.* Genetic epidemiology of COPD (COPDGene) study design. *COPD* 2010;7:32–43.
12. Vestbo J, Anderson W, Coxson HO, Crim C, Dawber F, Edwards L, *et al.*; ECLIPSE Investigators. Evaluation of COPD Longitudinally to Identify Predictive Surrogate End-points (ECLIPSE). *Eur Respir J* 2008;31:869–873.
13. Obeidat M, Nie Y, Chen V, Shannon CP, Andiappan AK, Lee B, *et al.* Network-based analysis reveals novel gene signatures in peripheral

blood of patients with chronic obstructive pulmonary disease. *Respir Res* 2017;18:72.
14. Wells JM, Arenberg DA, Barjaktarevic I, Bhatt SP, Bowler RP, Christenson SA, *et al.* Safety and tolerability of comprehensive research bronchoscopy in chronic obstructive pulmonary disease. Results from the SPIROMICS bronchoscopy substudy. *Ann Am Thorac Soc* 2019;16:439–446.
15. Freeman CM, Crudgington S, Stolberg VR, Brown JP, Sonstein J, Alexis NE, *et al.* Design of a multi-center immunophenotyping analysis of peripheral blood, sputum and bronchoalveolar lavage fluid in the Subpopulations and Intermediate Outcome Measures in COPD Study (SPIROMICS). *J Transl Med* 2015;13:19.
16. Serban KA, Pratte KA, Bowler RP. Protein biomarkers for COPD outcomes. *Chest* 2021;159:2244–2253.
17. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, *et al.* limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* 2015;43:e47.
18. Newman AM, Steen CB, Liu CL, Gentles AJ, Chaudhuri AA, Scherer F, *et al.* Determining cell type abundance and expression from bulk tissues with digital cytometry. *Nat Biotechnol* 2019;37:773–782.
19. Foroushani ABK, Brinkman FSL, Lynn DJ. Pathway-GPS and SIGORA: identifying relevant pathways based on the over-representation of their gene-pair signatures. *PeerJ* 2013;1:e229.
20. Singh D, Fox SM, Tal-Singer R, Bates S, Riley JH, Celli B. Altered gene expression in blood and sputum in COPD frequent exacerbators in the ECLIPSE cohort. *PLoS One* 2014;9:e107381.
21. Morrow JD, Qiu W, Chhabra D, Rennard SI, Belloni P, Belousov A, *et al.* Identifying a gene expression signature of frequent COPD exacerbations in peripheral blood using network methods. *BMC Med Genomics* 2015;8:1.
22. Lonergan M, Dicker AJ, Crichton ML, Keir HR, Van Dyke MK, Mullerova H, *et al.* Blood neutrophil counts are associated with exacerbation frequency and mortality in COPD. *Respir Res* 2020;21:166.
23. Chan MC, Yeung YC, Yu ELM, Yu WC. Blood eosinophil and risk of exacerbation in chronic obstructive pulmonary disease patients: a retrospective cohort analysis. *Int J Chron Obstruct Pulmon Dis* 2020; 15:2869–2877.
24. Hastie AT, Martinez FJ, Curtis JL, Doerschuk CM, Hansel NN, Christenson S, *et al.*; SPIROMICS investigators. Association of sputum and blood eosinophil concentrations with clinical measures of COPD severity: an analysis of the SPIROMICS cohort. *Lancet Respir Med* 2017;5:956–967.
25. Freeman CM, Martinez CH, Todt JC, Martinez FJ, Han MK, Thompson DL, *et al.* Acute exacerbations of chronic obstructive pulmonary disease are associated with decreased CD4+ & CD8+ T cells and increased growth & differentiation factor-15 (GDF-15) in peripheral blood. *Respir Res* 2015;16:94.
26. Freimer JW, Shaked O, Naqvi S, Sinnott-Armstrong N, Kathiria A, Garrido CM, *et al.* Systematic discovery and perturbation of regulatory genes in human T cells reveals the architecture of immune networks. *Nat Genet* 2022;54:1133–1144.
27. Kirdar S, Serter M, Ceylan E, Şener AG, Kavak T, Karadağ F. Adiponectin as a biomarker of systemic inflammatory response in smoker patients with stable and exacerbation phases of chronic obstructive pulmonary disease. *Scand J Clin Lab Invest* 2009;69: 219–224.
28. Minas M, Mystridou P, Georgoulias P, Pournaras S, Kostikas K, Gourgoulianis KI. Fetuin-A is associated with disease severity and exacerbation frequency in patients with COPD. *COPD* 2013;10:28–34.
29. Ozyurek BA, Ulasli SS, Bozbas SS, Bayraktar N, Akcay S. Value of serum and induced sputum surfactant protein-D in chronic obstructive pulmonary disease. *Multidiscip Respir Med* 2013;8:36.
30. Miller BE, Tal-Singer R, Rennard SI, Furtwaengler A, Leidy N, Lowings M, *et al.*; Perspective of the Chronic Obstructive Pulmonary Disease Biomarker Qualification Consortium. Plasma fibrinogen qualification as a drug development tool in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2016;193:607–613.
31. Thomsen M, Ingebrigtsen TS, Marott JL, Dahl M, Lange P, Vestbo J, *et al.* Inflammatory biomarkers and exacerbations in chronic obstructive pulmonary disease. *JAMA* 2013;309:2353–2361.
32. Pascual-González Y, López-Sánchez M, Dorca J, Santos S. Defining the role of neutrophil-to-lymphocyte ratio in COPD: a systematic literature review. *Int J Chron Obstruct Pulmon Dis* 2018;13:3651–3662.
33. Lee H, Um SJ, Kim YS, Kim DK, Jang AS, Choi HS, *et al.* Association of the neutrophil-to-lymphocyte ratio with lung function and exacerbations in patients with chronic obstructive pulmonary disease. *PLoS One* 2016;11:e0156511.
34. Sabroe I, Dower SK, Whyte MKB. The role of Toll-like receptors in the regulation of neutrophil migration, activation, and apoptosis. *Clin Infect Dis* 2005;41:S421–S426.
35. Maselli DJ, Bhatt SP, Anzueto A, Bowler RP, DeMeo DL, Diaz AA, *et al.* Clinical epidemiology of COPD: insights from 10 years of the COPDGene Study. *Chest* 2019;156:228–238.
36. Wan ES, Castaldi PJ, Cho MH, Hokanson JE, Regan EA, Make BJ, *et al.*; COPDGene Investigators. Epidemiology, genetics, and subtyping of preserved ratio impaired spirometry (PRISm) in COPDGene. *Respir Res* 2014;15:89.
37. Celli BR, Fabbri LM, Aaron SD, Agusti A, Brook R, Criner GJ, *et al.* An updated definition and severity classification of chronic obstructive pulmonary disease exacerbations: the Rome Proposal. *Am J Respir Crit Care Med* 2021;204:1251–1258.
38. Szabo PA, Levitin HM, Miron M, Snyder ME, Senda T, Yuan J, *et al.* Single-cell transcriptomics of human T cells reveals tissue and activation signatures in health and disease. *Nat Commun* 2019;10:4706.