ORIGINAL RESEARCH

Chronic Obstructive Pulmonary Disease Is Associated with Epigenome-Wide Differential Methylation in BAL Lung Cells

Jonas Eriksson Ström[1]*, Simon Kebede Merid[2]*, Jamshid Pourazar[1], Anders Blomberg[1], Anne Lindberg[1], Mikael V. Ringh[3,4], Michael Hagemann-Jensen[5], Tomas J. Ekström[4,6], Annelie F. Behndig[1‡], and Erik Melén[2,7‡]

[1]Department of Public Health and Clinical Medicine, Section of Medicine, Umeå University, Umeå, Sweden; [2]Department of Clinical Sciences and Education, [3]Department of Clinical Neuroscience, [4]Center for Molecular Medicine, [5]Department of Cell and Molecular Biology, and [6]Department of Molecular Medicine and Surgery, Karolinska Institute, Stockholm, Sweden; and [7]Sachs Children's Hospital, Stockholm, Sweden

ORCID IDs: 0000-0002-3434-988X (J.E.S.); 0000-0002-3292-7471 (A.L.).

Abstract

DNA methylation patterns in chronic pulmonary obstructive disease (COPD) might offer new insights into disease pathogenesis. To assess methylation profiles in the main COPD target organ, we performed an epigenome-wide association study on BAL cells. Bronchoscopies were performed in 18 subjects with COPD and 15 control subjects (ex- and current smokers). DNA methylation was measured using the Illumina MethylationEPIC BeadChip Kit, covering more than 850,000 CpGs. Differentially methylated positions (DMPs) were examined for *1*) enrichment in pathways and functional gene relationships using the Kyoto Encyclopedia of Genes and Genomes and Gene Ontology, *2*) accelerated aging using Horvath's epigenetic clock, *3*) correlation with gene expression, and *4*) colocalization with genetic variation. We found 1,155 Bonferroni-significant ($P < 6.74 \times 10^{-8}$) DMPs associated with COPD, many with large effect sizes. Functional analysis identified biologically plausible pathways and gene relationships, including enrichment for transcription factor activity. Strong correlation was found between DNA methylation and chronological age but not between COPD and accelerated aging. For 79 unique DMPs, DNA methylation correlated significantly with gene expression in BAL cells. Thirty-nine percent of DMPs were colocalized with COPD-associated SNPs. To the best of our knowledge, this is the first epigenome-wide association study of COPD on BAL cells, and our analyses revealed many differential methylation sites. Integration with mRNA data showed a strong functional readout for relevant genes, identifying sites where DNA methylation might directly affect expression. Almost half of DMPs were colocated with SNPs identified in previous genome-wide association studies of COPD, suggesting joint genetic and epigenetic pathways related to disease.

Keywords: chronic obstructive pulmonary disease; DNA methylation; epigenetics; BAL cells; gene expression

Clinical Relevance

This study integrates novel findings of widespread differential methylation in BAL cells from subjects with chronic obstructive pulmonary disease (COPD) compared with control subjects with data from previous transcriptomic and genome-wide association studies. Taken together, results suggest that at some of the COPD-associated sites, DNA methylation might serve as a mediator between genetics and disease and at others be acquired alterations with functional relevance. As BAL collects cells from the parts of the lung most affected by the disease, this study might contribute to insights into COPD pathogenesis.

Chronic obstructive pulmonary disease (COPD) is a heterogeneous chronic inflammatory airway disease, characterized clinically by airflow limitation, chronic respiratory symptoms, and many systemic comorbidities. It has long been known that the key environmental risk factor for COPD is tobacco smoking, with smoke from biomass fuels and other pollutants as other major contributors. More recently, it has become clear that a failure to reach the predicted degree of peak lung function in early adulthood might also lead to chronic airflow limitation, even in the absence of environmental exposure (1, 2). According to

(Received in original form September 13, 2021; accepted in final form March 10, 2022)

*These authors contributed equally to this work and share first authorship.

‡These authors contributed equally to this work and share last authorship.

Am J Respir Cell Mol Biol Vol 66, Iss 6, pp 638–647, June 2022

Originally Published in Press as DOI: 10.1165/rcmb.2021-0403OC on March 14, 2022
Internet address: www.atsjournals.org

the World Health Organization, COPD is now the third leading cause of death in the world (3).

Why some, but not all, long-time smokers develop COPD is a major yet largely unresolved question in COPD research. Large-scale genome-wide association studies (GWASs) have revealed several genetic variants related to COPD (4), and according to twin studies, ~40–60% of COPD susceptibility is explained by genetic factors (5, 6). What additional role epigenetics in general, and DNA methylation in particular, might play in COPD pathogenesis has not been thoroughly investigated. Previous epigenome-wide association studies (EWAS) are scarce, especially those based on cells from the main target organ of the disease. In a systematic review from 2017, Machin and colleagues found no consistent differences in DNA methylation in peripheral blood (7), whereas the few existing EWASs of lung tissue and bronchial brushings have more consistently reported aberrant methylation in COPD (8–10). Different methylation patterns between target organ and surrogate tissue would be in line with what has previously been shown in transcriptomic studies of COPD (11).

The distal airways represent the main site of airflow limitation in COPD and are the parts of the lung most affected by the disease (12). To investigate DNA methylation patterns in cells from distal airways, we performed an EWAS on BAL cells from subjects with COPD and control subjects. We subsequently investigated potential functional effects of differential methylation through integration with GWAS and gene expression data and also tested the hypothesis of accelerated aging in COPD, using Horvath's epigenetic clock (Figure 1).

Methods

Study Subjects

Twenty-two subjects with COPD and 15 ever-smokers with normal lung function participated in this cross-sectional study. All subjects were recruited from the longitudinal OLIN (Obstructive Lung Disease in Northern Sweden) COPD study using predetermined criteria (13); details are provided in the data supplement.

Bronchoscopy

All subjects underwent bronchoscopy with BAL, as described in the supplement. BAL could not be performed in three subjects with COPD, because of problems tolerating bronchoscopy. In one subject with COPD, BAL recovery was too low to perform analysis.

Statistical Analyses

We performed EWAS analysis to estimate association between DNA methylation β values and COPD using robust linear regression (rlm in the MASS R package [14]). The analyses were adjusted for sex, age (years), inhaled corticosteroid use, pack-years, smoking intensity, and measured cell type count (macrophages, lymphocytes, neutrophils, and eosinophils). Covariate selection was aided by a principal-component regression analysis (*see* Figure E1 in the data supplement). We applied Bonferroni correction ($P < 6.74 \times 10^{-8}$) to account for multiple testing in EWAS results. Differentially methylated regions (DMRs) were identified using DMRcate (15). Input parameters for the DMR calling algorithms are provided in the data supplement.

Functional Gene Relationship and Pathway Enrichment Analysis

Functional enrichment analysis in gene ontology and pathway enrichment in Kyoto Encyclopedia of Genes and Genomes (KEGG) was performed using the ConsensusPathDB (16, 17) overrepresentation analysis tool (http://consensuspathdb.org). *P* values for enrichment were adjusted for multiple testing using the false discovery rate method.

Epigenetic Age

DNA methylation age was calculated using the new DNA Methylation Age Calculator by Horvath (18). Details on preprocessing and calculator options can be found in the data supplement.

Integration with GWAS

COPD-associated differentially methylated positions (DMPs) were compared with the 82 SNPs from a recent large-scale GWAS meta-analysis (4). We investigated if DMPs in the present study were located within a window 500 kb upstream and 500 kb downstream of these SNPs. Moreover, we checked whether SNPs and DMPs were located in the same gene.

DNA Methylation in Relation to Gene Expression

Correlations between DNA methylation and amounts of gene expression were tested using a previously published paired DNA methylation and gene expression BAL cell data set (19). We tested transcript concentrations of genes within a 500-kb region of the 1,155 DMPs (250 kb upstream and 250 kb downstream). The mRNA gene expression (RNA sequencing [RNA-seq]) and methylation (Illumina Infinium MethylationEPIC BeadChip Kit) were measured in BAL samples from 7 smokers and 12 nonsmokers (19). First, we created residuals for mRNA expression and residuals for DNA methylation and used Pearson correlation to evaluate correlations between expression residuals and DNA methylation

Supported by Hjärt-Lungfonden, Vetenskapsrådet, Region Västerbotten, Region Stockholm (ALF Project), the Visare Norr Fund/Northern County Councils Regional Federation, Umeå Universitet, King Gustaf V's and Queen Victoria's Freemason Foundation, the Arnerska Research Foundation, and Kempestiftelserna.

Author Contributions: J.E.S., S.K.M., A.F.B., and E.M. developed the research idea. S.K.M. conducted the main bioinformatic analyses. J.E.S. drafted the manuscript. A.L. was responsible for subject recruitment. A.B. and A.F.B. were responsible for bronchoscopy procedures and collection of samples. J.P. extracted DNA and conducted flow cytometry analysis. M.V.R. and M.H.-J. performed additional DNA methylation and RNA sequencing analyses. T.J.E. provided intellectual input for the interpretation of data. All authors critically reviewed and approved the manuscript.

Correspondence and requests for reprints should be addressed to Jonas Eriksson Ström, M.Sc., Department of Public Health and Clinical Medicine, Umeå University, 901 87 Umeå, Sweden.. E-mail: jonas.eriksson.strom@umu.se.

This article has a related editorial.

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

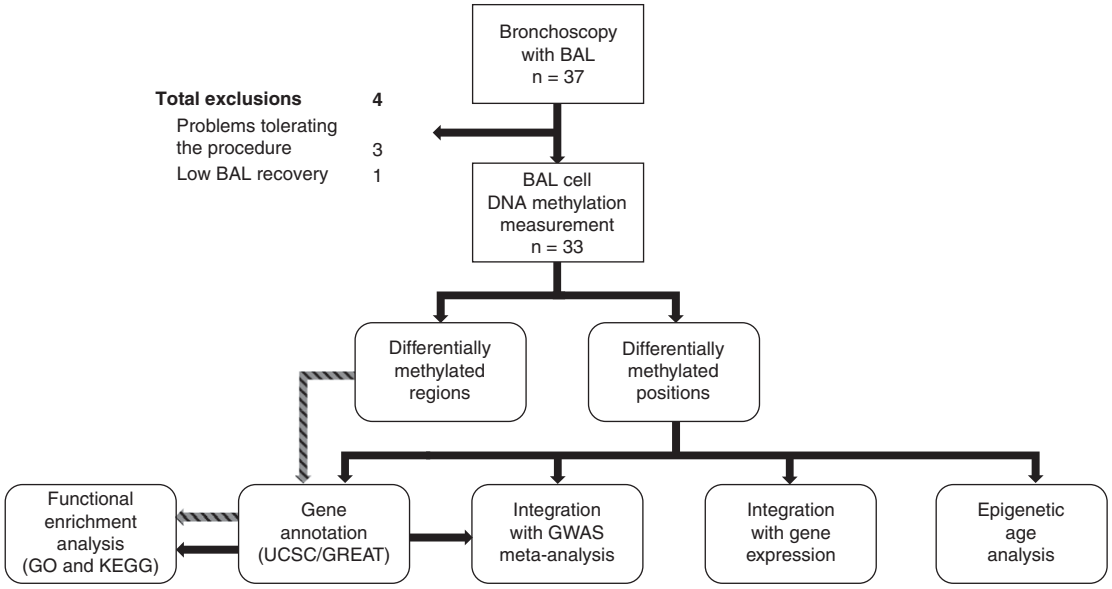

Figure 1. Study overview. GO = Gene ontology; GREAT = Genomic Regions Enrichment of Annotations Tool; GWAS = genome-wide association study; KEGG = Kyoto Encyclopedia of Genes and Genomes; UCSC = The University of California, Santa Cruz, Genome Browser Database.

residuals. These residual models were adjusted for covariates (age, sex, and smoking status). We corrected these analyses for multiple testing using Bonferroni correction. In addition, we explored the Immunological Genome Project (ImmGen) RNA-seq data set to evaluate if differentially methylated genes were likely to originate from macrophages or other cell types (20).

Results

DNA methylation data were available for 18 subjects with COPD and 15 control subjects with normal lung function (Table 1). There were no significant differences between cases and control subjects by age or body mass index. As expected, inhaled corticosteroids were used only by subjects with COPD. Pack-years were higher in the COPD group, and consistent with the results of previous studies (21), BAL recovery was lower. Both groups consisted of current and ex-smokers. Among current smokers, smoking intensity was higher in subjects with COPD. Differential cell counts revealed no significant differences between groups, and as expected, the predominant cell type in BAL was macrophages (Table 1).

Differential Methylation: Positions and Regions

DNA methylation data were analyzed for DMPs and DMRs.

We found 1,155 Bonferroni-significant DMPs associated with COPD, spanning all chromosomes (chrs) (*see* Figure 2 and Figure E3). DMPs were enriched for localization to CpG islands (23% of COPD-associated CpGs, 19% of all CpGs on the array; $P_{enrichment} = 0.0004$) and relatively depleted in shores (15% of COPD-associated CpGs, 18% of all CpGs; $P_{enrichment} = 0.01$; *see* Figure E3). Fifty-five percent of DMPs (638 of 1,155) displayed higher mean methylation in subjects with COPD relative to control subjects.

The DMP with the lowest *P* value was annotated to *CLTA* (clathrin light chain A) on chr 9 ($P = 6.00 \times 10^{-39}$ for cg00202130; Table 2) and had a 1.1% higher mean methylation degree in subjects with COPD relative to control subjects. The largest effect size was observed for cg14640588, annotated to *ZDHHC14* (zinc finger DHHC-type palmitoyltransferase 14) on chr 6, with 28.9% higher mean methylation in COPD (Table 3). Among DMPs with lower mean methylation in COPD, cg15969227 (nearest gene *ZNF322* [zinc finger protein 322] on chr 6) had the largest effect size, with a mean methylation difference of 27.6% (Table 3). *NUB1* (negative regulator of ubiquitin like proteins 1) on chr 7 was the gene with the largest proportion of significant DMPs annotated to it (3 differentially methylated CpGs/75 total CpGs = 4.0%), all with higher mean methylation in COPD.

COPD-associated DMPs were annotated to 1,089 genes. Using the ImmGen RNA-seq data set (20), we identified a total of 5,328 genes expressed exclusively by nonmacrophage cell types (i.e., genes expressed by innate lymphocytes, αβ and γδ T cells, B cells, stem cells, dendritic cells, granulocytes, mast cells, basophils, eosinophils, and stromal cells but not by macrophages; *see* the data supplement for details). The overlap between these two gene sets was 8.4%.

Using DMRcate, we identified 7,097 DMRs. The most significant DMR (false discovery rate–adjusted $P = 2.58 \times 10^{-43}$) was a region on chr 15 spanning five CpG sites and overlapping the promotor of *ZNF609* (zinc finger protein 609) (*see* Table E3).

Functional Gene Relationship and Pathway Enrichment Analysis

Enrichment analyses was conducted separately for DMPs (annotated to 1,089 genes) and DMRs (annotated to 5,491 genes); results below are those overlapping in DMP and DMR analyses (for complete results, *see* Tables E4–E7). Gene ontology analysis showed significant enrichment for 502 terms, including transcription factor activity (molecular functions), nervous system and anatomical structure development, and regulation of metabolic processes (biological processes). KEGG pathway analyses revealed seven significant

Table 1. Study Subjects

	COPD (*n* = 18)	Smokers with Normal LF (*n* = 15)	*P* Value
Female:male	4:14	8:7	NS
Age, y	63 (62–72)	65 (63–71.5)	NS
BMI, kg/m^2	26.2 (23.7–29.0)	26.0 (24.2–27.8)	NS
Current smokers:ex-smokers	10:8	3:12	NS
Smoking intensity (among current smokers), cigarettes/d	14 (10–20)	8 (7–10)	0.049
Pack-years	36 (30–40)	15 (13.5–21.5)	<0.0001
FEV_1% predicted	68.3 (44.1–72.3)	110.2 (94.4–117.9)	<0.0001
FEV_1/VC	0.53 (0.43–0.62)	0.74 (0.70–0.77)	<0.0001
Use of inhaled corticosteroids, yes:no	6:12	0:15	0.02
BAL recovery, %	43.6 (32.2–55.6)	61.1 (56.1–66.7)	0.001
BAL macrophages, %	89.5 (86.4–92.8)	86.4 (81.3–92.6)	NS
BAL lymphocytes, %	8.3 (5.2–13.2)	11.8 (6.1–18.2)	NS
BAL neutrophils, %	1.0 (0.4–1.8)	0.6 (0.3–1.2)	NS
BAL eosinophils, %	0.3 (0.0–0.6)	0.2 (0.0–0.3)	NS
BAL mast cells, %	0.2 (0.0–0.5)	0.0 (0.0–0.1)	NS

Definition of abbreviations: BMI = body mass index; COPD = chronic obstructive pulmonary disease; FEV_1 = forced expiratory volume in 1 second; LF = lung function; NS= not significant; VC = vital capacity (defined as the highest value of forced and slow vital capacity). Values are given as median and interquartile range unless indicated differently. Statistical comparisons: for ratios, the Fisher exact test was used, and for other comparisons, the Mann-Whitney *U* test was used; a *P* value of less than 0.05 was considered to indicate significance. Pack-years were calculated as (number of cigarettes smoked per day/20) × number of years smoked.

terms, many related to various cancers but also to cAMP and Rap1 signaling pathways as well as different types of cell junctions.

Epigenetic Aging

Using the DNA methylation age calculator of Horvath (18), we calculated the subjects' DNA methylation age. As expected, there was a strong correlation between DNA methylation and chronological age ($r = 0.71$; *see* Figure E4), but no association was found between COPD status and the difference in DNA methylation age and chronological age (*see* the data supplement), nor were COPD-associated DMPs enriched in the CpGs included in Horvath's epigenetic clock (data not shown).

Correlation of DNA Methylation and Gene Expression

Using a data set with paired BAL cell DNA methylation and gene expression data from Ringh and colleagues (19) (*see* also Table E10), we assessed the correlation between degrees of gene expression and CpG methylation. Of the 1,155 COPD-associated DMPs in the present study, 1,065 were available for analysis in Ringh and colleagues' data set (within a ±250-kb window of a transcript). Of these, 79 unique DMPs (101 total CpG–transcript associations) correlated significantly with gene expression (Bonferroni $P < 0.05$; *see* Table E9). Fifty-four percent of all correlations were negative, with the lowest *P* value as well as the largest effect estimate ($P = 4.81 \times 10^{-5}$; $\rho = -0.80$) reported for cg18196647 and *CPD* (carboxypeptidase D) (Figure 3). Among positive correlations, the one between cg13267718 and *FLI1* (friend leukemia integration 1) transcription factor had both the lowest *P* value and the largest effect estimate ($P = 1.24 \times 10^{-5}$; $\rho = 0.83$; Figure 3).

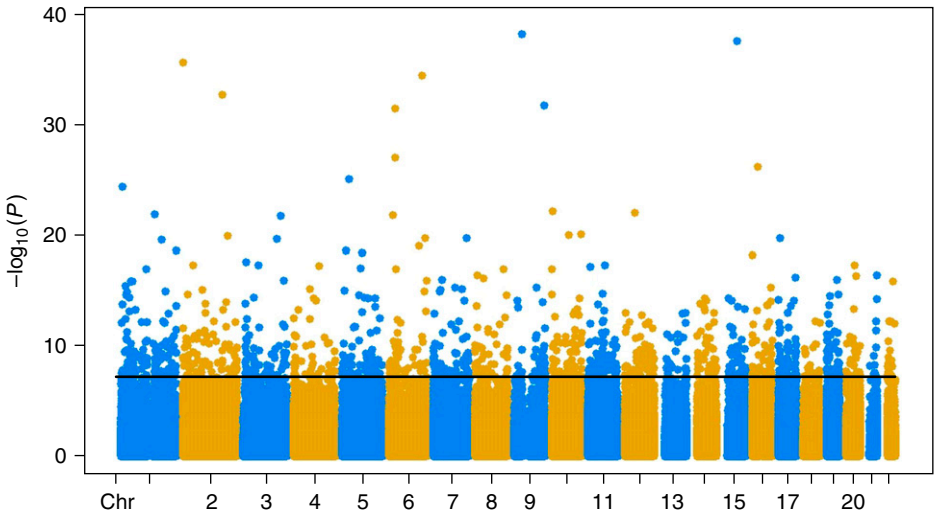

Figure 2. Manhattan plot indicating the associations between COPD and DNA methylation. Black line represents the Bonferroni-corrected threshold ($P < 0.05$/[number of CpGs analyzed] ≈ 6.74×10^{-8}). Chr = chromosome; COPD = chronic obstructive pulmonary disease.

Table 2. The Top 10 Bonferroni-Significant Differentially Methylated Positions Associated with Chronic Obstructive Pulmonary Disease, Ranked by *P* Value

CpG ID	Chr	Genomic Coordinate	Annotated to Gene	Coefficient	*P* Value
cg00202130	9	36190707	*CLTA1**	−0.0113	6.00×10^{-39}
cg27079832	15	64790751	*ZNF609**	0.0207	2.72×10^{-38}
cg07088935	2	1426845	*TPO**	0.1035	2.25×10^{-36}
cg06958223	6	150499199	*PPP1R14C**	0.0392	3.39×10^{-35}
cg16928994	2	178977002	*RBM45**	−0.0150	1.93×10^{-33}
cg11619181	9	136500618	*DBH**	−0.0416	1.88×10^{-32}
cg05738196	6	26577821	*ABT1*†	0.0306	3.47×10^{-32}
cg15969227	6	26745279	*ZNF322*†	0.2759	9.23×10^{-28}
cg14948709	16	25248396	*ZKSCAN2**	0.0468	6.77×10^{-27}
cg15803210	5	31757766	*PDZD2*†	−0.0445	8.31×10^{-26}

Definition of abbreviations: *ABT1* = activator of basal transcription 1; Chr = chromosome; *CLTA1* = clathrin light chain A sequence 1; *DBH* = dopamine β-hydroxylase; ID = identifier; *PDZD2* = PDZ domain containing 2; *PPP1R14C* = protein phosphatase 1 regulatory inhibitor subunit 14C; *RBM34* = RNA binding motif protein 45; *TPO* = thyroid peroxidase; *ZKSCAN2* = zinc finger with KRAB and SCAN domains 2; *ZNF322* = zinc finger protein 322; *ZNF609* = zinc finger protein 609.
Adjusted for sex, age, cell type, inhaled corticosteroid use, pack-years, and smoking intensity.
*Illumina annotation based on The University of California, Santa Cruz (UCSC) database.
†Nearest gene as reported by Genomic Regions Enrichment of Annotations Tool (GREAT) version 4.0.4.

Integration with GWAS

To investigate colocalization of genetic and epigenetic variation, DMPs in the present study were mapped to the 82 SNPs associated with COPD in a recent GWAS meta-analysis of 35,735 cases and 222,076 control subjects (4). We found that 447 of the 1,155 DMPs were within a ±500-kb window of one or more COPD-associated SNPs (*see* Table). Of these, 116 DMPs were within 100 kb of one or more SNPs. Ten colocalized SNPs and DMPs were annotated to the same gene, and of these, *ADGRG6*/*GPR126* (adhesion G protein–coupled receptor G6) was the only gene with more than one DMP annotated to it (Table 4). Conversely, 80 of 82 COPD-associated SNPs had one or more DMPs located within a ±500-kb window. Sixty-four SNPs were within 100 kb of one or more DMPs.

Discussion

Differential Methylation

To the best of our knowledge, this is the first EWAS of COPD on BAL cells. Measurement of more than 850,000 CpGs revealed genome-wide differential methylation in COPD. Effect sizes were large for many DMPs, comparable to those previously reported for tobacco smoking (22). As for effect direction, roughly as many DMPs had higher as lower mean methylation in subjects with COPD relative to control subjects. This is in contrast to smoking, where a large majority of DMPs are hypomethylated (19). One advantage of using BAL fluid is that it collects lung immune cells from the most distal airways, which also are the parts most affected in COPD (12). Our results may therefore provide new insights about COPD pathogenesis, although it should be noted that distal lung epithelial cells, which are also

Table 3. The Top 10 Bonferroni-Significant Differentially Methylated Positions Associated with Chronic Obstructive Pulmonary Disease, Ranked by the Magnitude of Positive and Negative Effects (5 Differentially Methylated Positions Each)

CpG ID	Chr	Genomic Coordinate	Annotated to Gene	Coefficient	*P* Value
cg15969227	6	26745279	*ZNF322*†	0.2759	9.23×10^{-28}
cg07464716	5	157079520	*SOX30**	0.1973	4.18×10^{-8}
cg02757970	2	25427350	*POMC*†	0.1659	7.02×10^{-10}
cg02567750	1	247569605	*NLRP3*†	0.1654	3.67×10^{-12}
cg21834645	6	10409467	*TFAP2A**	0.1625	2.03×10^{-9}
cg14640588	6	158028202	*ZDHHC14**	−0.2891	1.91×10^{-8}
cg08486986	16	1512134	*CLCN7**	−0.2090	6.16×10^{-8}
cg02023276	7	151049248	*NUB1**	−0.1973	1.43×10^{-11}
cg20095851	12	6486701	*SCNN1A**	−0.1743	1.01×10^{-13}
cg21627955	13	46952839	*KIAA0226L**	−0.1621	1.02×10^{-8}

Definition of abbreviations: *CLCN7* = chloride voltage-gated channel 7; *KIAA0226L* = Rubicon like autophagy enhancer; *NLRP3* = NLR family pyrin domain containing 3; *NUB1* = negative regulator of ubiquitin like proteins 1; *POMC* = proopiomelanocortin; *SCNN1A* = sodium channel epithelial 1 subunit α; *SOX30* = SRY-box transcription factor 30; *TFAP2A* = transcription factor AP-2 α; *ZDHHC14* = zinc finger DHHC-type palmitoyltransferase 14.
Adjusted for sex, age, cell type, inhaled corticosteroid use, pack-years, and smoking intensity.
*Illumina annotation based on the UCSC database.
†Nearest gene as reported by GREAT version 4.0.4.

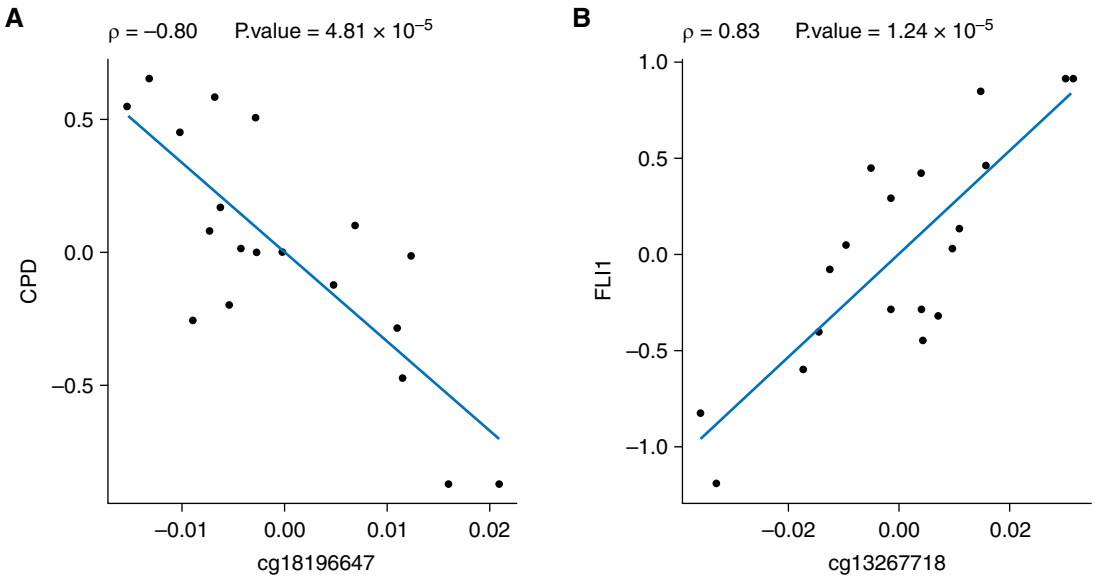

Figure 3. Correlations between methylation and predicted gene expression levels for selected pairs. (*A*) CPD = carboxypeptidase D. (*B*) FLI1 = friend leukemia integration 1. We created residuals for mRNA expression and DNA methylation and then used Pearson correlation to evaluate correlations between expression residuals and methylation residuals. These residual models were adjusted for age, sex, and smoking status.

believed to play an important role, are not collected using BAL.

Among genes annotated to the 10 DMPs with the largest effect sizes, four have previously been associated with COPD and/or disease severity: *NLRP3* (NLR family pyrin domain containing 3), *SOX30* (SRY-box transcription factor 30), *POMC* (proopiomelanocortin), and *ZNF322*. Of these, *NLRP3* (23) and *SOX30* (4) have also been put forward as candidate drug targets. *NLRP3* encodes for a cytoplasmic pattern recognition receptor that, as part of the NLRP3 inflammasome, is a key regulator of the inflammatory response. In mouse models, *NLRP3* has been shown to be essential for the development of COPD (23), and there is increasing evidence of an important role for the NLRP3 inflammasome in COPD pathogenesis also in humans (24, 25). *SOX30* is a transcription factor previously associated with progression of emphysema in patients with COPD (26). It was recently also identified as a potential causal/effector gene in an integrated analysis of COPD GWAS, gene expression, and gene regulation data (4). In the present study, the DMP annotated to *SOX30* (cg07464716) correlates significantly with the expression of U7 snRNA-associated Sm-like protein LSm10, a protein previously found to be involved in chronic mucus hypersecretion in COPD (27).

The predominant BAL fluid cell type in our data set was macrophages (≈90%), and it has been shown previously that alveolar macrophage function can be modified epigenetically in COPD (28). To investigate if differential methylation in the present study was related to macrophages, we explored the ImmGen platform and available cell-specific expression profiles. First, we compiled a list

Table 4. Differentially Methylated Positions and SNPs Annotated to the Same Gene

CpG ID	Genomic Coordinate	rs ID	Genomic Coordinate	Chr	Annotated to Gene	Base Pair Difference
cg15497249	239485905	rs11579382	239737706	1	*CHRM3*†	251801
cg07386720	11676054	rs2442776	11599127	3	*VGLL4**	−76927
cg13029151	29685706	rs13073544	29430921	3	*RBMS3**	−254785
cg15257376	141121677	rs7650602	141428572	3	*ZBTB38**	306895
cg24836671	148029901	rs10037493	148475407	5	*HTR4**	445506
cg14755119	19753466	rs9350191	19842430	6	*ID4*†	88964
cg18128350	142706817	rs9399401	142347764	6	*ADGRG6/GPR126**	−359053
cg18131185	142623588	rs9399401	142347764	6	*ADGRG6/GPR126**	−275824
cg04461802	142623433	rs9399401	142347764	6	*ADGRG6/GPR126**	−275669
cg06280150	18328109	rs9617650	18006117	22	*MICAL3**	−321992

Definition of abbreviations: *ADGRG6/GPR126* = adhesion G protein–coupled receptor G6; *CHRM3* = cholinergic receptor muscarinic 3; *HTR4* = 5-hydroxytryptamine receptor 4; *ID4* = inhibitor of DNA binding 4, HLH protein; *MICAL3* = microtubule associated monooxygenase, calponin and LIM domain containing 3; *RBMS3* = RNA binding motif single stranded interacting protein 3; rs = reference SNP; *VGLL4* = vestigial like family member 4; *ZBTB38* = zinc finger and BTB domain containing 38.
*Illumina annotation based on the UCSC database.
†Nearest gene as reported by GREAT version 4.0.4.

of genes expressed exclusively by cell types other than macrophages (i.e., genes expressed by innate lymphocytes, αβ and γδ T cells, B cells, stem cells, dendritic cells, granulocytes, mast cells, basophils, eosinophils, and stromal cells, but not by macrophages); we then compared this list with our list of differentially methylated COPD genes. The overlap between the nonmacrophage list from ImmGen and our COPD gene list was only 8.4%, indicating that the vast majority of identified changes likely originate from macrophages (and/or unmeasured cell types that share expression profile with macrophages).

COPD, Tobacco Smoking, or Both?

Tobacco smoking has been shown to have a clear and widespread effect on the methylome, not only in peripheral blood but also in BAL cells (19, 29). As the subjects with COPD in the present study are ex- or current smokers, one challenge was to separate the signal of smoking from that of COPD. We approached this in two ways: *1*) by comparing smokers with COPD to smokers without COPD (the major characteristic separating groups was thus only disease status, not smoking *and* disease status as in studies in which never-smokers are used as control subjects); and *2*) through statistical adjustments for effects of smoking, both cumulative (pack-years) and acute (smoking intensity; i.e., number of cigarettes per day at the time of bronchoscopy). In a recent systematic review, smoking-related changes in DNA methylation were deemed most documented for four genes: *AHRR* (aryl-hydrocarbon receptor repressor), *F2RL3* (F2R like thrombin), *DAPK* (death associated protein kinase 1), and *p16*/*CDKN2A* (cyclin dependent kinase inhibitor 2A) (30). Notably, none of these was among the 1,089 genes annotated to DMPs in the present study, possibly indicating a COPD-specific signal.

Besides smoking, we also adjusted for use of inhaled corticosteroids. Epigenetic mechanisms, including DNA methylation, have long been believed to play a central role in mediating the antiinflammatory effect of glucocorticoids (31). In COPD, systemic corticosteroid use has been associated with differential methylation (32), and recently van Nijnatten and colleagues presented data indicating that this might be the case also for inhaled corticosteroids (33).

Enrichment for Biological Processes and Pathways

KEGG and gene ontology enrichment analysis identified several biologically plausible pathways and functional gene relationships. In the latter category, we found significant enrichment for known transcription factor activity, replicating the finding of one previous EWAS of lung tissue from patients with COPD (8). Of significant pathways, many were related to malignancies, which is not unexpected, as COPD is a known risk factor, independently of smoking, for several types of cancer (34). Interestingly, the two significant pathways not evidently cancer related, cAMP and Rap1 signaling, are both targeted by drugs commonly used in COPD. β-Agonists increase cAMP concentrations through activation of adenylate cyclase, which initiates a signal transduction pathway important for bronchodilation (35) and might also affect airway epithelial barrier function through the EPAC (exchange protein directly activated by cAMP)/Rap1 pathway (36). Anticholinergics enable cAMP-mediated relaxation of smooth muscle cells by inhibiting M_2 muscarinic receptors (37). Corticosteroids (38) and roflumilast (39) both stimulate cAMP production, and the latter has also been suggested to exert its effect via actions on neutrophil Rap1 (39).

Epigenetic Clock

Many features of normal aging are prominent in COPD lungs (e.g., increased cellular senescence, oxidative stress, stem cell exhaustion [40]), and it has been proposed that COPD is a condition of accelerated aging (41). DNA methylation age, calculated using Horvath's epigenetic clock, has been validated to correlate with chronological age, not only in peripheral blood but also in BAL cells from healthy subjects (42). As epigenetic alteration is one of the hallmarks of aging (43), we hypothesized that subjects with COPD would have an increased difference between DNA methylation age and chronological age compared with control subjects. However, no significant difference was found, suggesting that if epigenetic alteration is indeed involved in accelerated aging of BAL cells, it is likely through mechanisms other than DNA methylation. Histone deacetylase enzyme activity, for example, regulates histone acetylation and has previously been associated with disease severity in COPD. Interestingly, one of the transcription factors identified in the ontology analysis in the current study, sirtuin 1, is a histone deacetylase associated not only with COPD but also with aging (44).

Integration with Gene Expression

To identify potential functional effects of differential methylation, we investigated correlation with gene expression in a separate BAL cell data set. We found many DMPs for which methylation was strongly correlated with gene expression, indicating that these CpGs may have a direct functional effect on disease manifestation.

Among the 10 genes with the highest correlations between DMP methylation and gene expression, 7 have previously been associated with COPD and/or lung function (*FLI1*, *C11orf45* [chromosome 11 open reading frame 45], *ASF1A* [anti-silencing function 1A histone chaperone], *GOSR1* [Golgi SNAP receptor complex member 1], *NSRP1* [nuclear speckle splicing regulatory protein 1], *SHH* [sonic hedgehog signaling molecule], and *PCDHAC1* [protocadherin α subfamily C, 1]) and 4 with lung cancer (*FLI1*, *ASF1A*, *CPD*, and *SHH*), whereas *PCDHB3* (protocadherin β 3) and *PGAM2* (phosphoglycerate mutase 2) have not previously been linked to lung disease or lung function.

The most significant positive correlation was found for cg13267718 and *FLI1*. This protooncogene encodes for a transcription factor believed to play a role in both small-cell lung cancer (45) and non–small-cell lung cancer (46), both comorbidities of COPD. *FLI1* has also been associated with lung function (47) and, specifically in COPD, has been found to be upregulated in activated neutrophils (48). Dysregulated neutrophilic inflammation has long been considered a prominent feature of COPD, contributing to tissue destruction and the emblematic progressive decline in lung function (49).

Integration with GWAS

One major question in COPD research is why some, but not all, long-time smokers develop COPD. Genetic determinants have been proposed to partially explain this variation in disease susceptibility, and in a large-scale GWAS meta-analysis, Sakornsakolpat and colleagues recently identified 82 loci associated with the disease (4). In the present study, 38.7% of DMPs were colocalized with one or more of these COPD-associated SNPs, indicating that these

might be sites where *cis*-acting DNA motifs and/or DNA-binding factors shape DNA methylation and, conversely, that the remaining 61.3% of DMPs are more likely influenced by environmental determinants. Correlation with gene expression was found among DMPs in both these groups, suggesting that in COPD, DNA methylation might at some sites serve as a mediator between genetics and disease and at others be acquired alterations with functional relevance.

Among genes annotated to sites of colocalized epigenetic and genetic variation, *ADGRG6*/*GPR126* stood out as the only one with more than one DMP annotated to it. This gene was differentially methylated in COPD also in one previous EWAS on epithelial cells from bronchial brushings (10). In GWASs and gene expression studies, *ADGRG6*/*GPR126* has been associated not only with COPD (50) but also independently with D_{LCO}/V_A (51). Technically, D_{LCO}/V_A is an index of how efficiently carbon monoxide is transferred over the alveolar–capillary membrane. In practice, D_{LCO}/V_A is often low in subjects with emphysema, and it is used together with other measurements to distinguish among different pulmonary pathologies (52). *ADGRG6*/*GPR126* codes for a G protein–coupled receptor that modulates the VEGF (vascular endothelial growth factor) cascade in hypoxia-induced angiogenesis (53), which in turn might affect D_{LCO}/V_A through decreased perfusion. In summary, findings from multiple tissues and several omics layers indicate a possible role for *ADGRG6*/*GPR126* in COPD pathogenesis that might be mediated through epigenetic mechanisms.

Strengths and Limitations

One strength of the present study is its well-characterized study population. All subjects were recruited from the longitudinal and population-based OLIN COPD study, which provided granular and reliable data on smoking habits, use of inhaled corticosteroids, and other possible confounders (13). In the comprehensive statistical analysis, these data were then used as covariates to sift out a signal as COPD specific as possible. Another strength is that we investigated cells from the lungs, not from peripheral blood. Even though COPD is a systemic disease, when trying to uncover its underlying epigenetic mechanisms, investigating the main target organ is likely to produce more relevant data than using surrogate tissue. In addition, to account for potential cell type effects, we adjusted our models using measured cell count.

One limitation of the present study is that it included relatively few participants. Although this is true for most studies dependent on bronchoscopies, a larger study population would have been desirable. Also, although we did detect widespread differential methylation associated with COPD, previous studies of lung diseases other than COPD indicate that a larger sample size might be needed to identify a stable and replicable signature (54). As this is the first published EWAS on COPD BAL cells, *1*) one major limitation was that replication in an independent COPD BAL data set was not possible; and *2*) correlation between BAL cell DNA methylation and gene expression had to be assessed in a non-COPD data set, and results are thus dependent on the assumption that methylation–expression patterns are fairly stable among individuals and groups. Although replication in peripheral blood data sets would have been possible, transcriptomic studies of COPD have shown that gene expression patterns differ between target organ and surrogate tissue such as peripheral blood (11). And because DNA methylation is one of the main mechanisms regulating tissue-specific gene expression (55), it seems reasonable to assume that DNA methylation patterns would also differ among compartments. Thus, results of a replication analysis in a peripheral blood data set, whether positive or negative, would be hard to interpret with any certainty. Finally, a cross-sectional study such as the present one cannot determine the causal direction of an association. In other words, it cannot conclude whether the identified epigenetic variation contributed to COPD susceptibility, or if it was a result of the disease, postdisease processes, or even disease-associated drug interventions.

Conclusions

We found epigenome-wide differential methylation in COPD lung cells from BAL. For many DMPs, effect sizes were large, and some also displayed a strong correlation with gene expression. Almost half of DMPs were located in close proximity to sites where genetic variation (SNPs) is known to be associated with COPD, sites where DNA sequence and/or DNA-binding factors might dictate methylation status. Conversely, remaining DMPs are more likely influenced by external factors such as tobacco smoke and air pollutants. Our interpretation is that genetic and acquired epigenetic factors both might affect COPD susceptibility. As this is the first ever EWAS of COPD BAL cells, the results need to be validated in future studies. ■

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

Acknowledgment: The authors thank Eva Rönnmark, Viktor Strandkvist, Helena Backman, Annika Johansson, Frida Holmström, Dirk Albrecht, the OLIN studies, and the Division of Respiratory Medicine and Allergy, Department of Medicine, Sunderby Central Hospital of Norrbotten, Luleå, Sweden, for their contributions to the project. Methylation profiling was performed by the SNP&SEQ Technology Platform in Uppsala (www.genotyping.se). The facility is part of the National Genomics Infrastructure supported by the Swedish Research Council for Infrastructures and Science for Life Laboratory, Sweden.

References

1. Agustí A, Hogg JC. Update on the pathogenesis of chronic obstructive pulmonary disease. *N Engl J Med* 2019;381:1248–1256.
2. Melén E, Guerra S, Hallberg J, Jarvis D, Stanojevic S. Linking COPD epidemiology with pediatric asthma care: implications for the patient and the physician. *Pediatr Allergy Immunol* 2019;30:589–597.
3. World Health Organization. Global health estimates 2019: disease burden by cause, age, sex, by country and by region, 2000–2019. Geneva, Switzerland: World Health Organization; 2019 [accessed 2021 Jun 3]. Available from: https://www.who.int/data/gho/data/themes/mortality-and-global-health-estimates.
4. Sakornsakolpat P, Prokopenko D, Lamontagne M, Reeve NF, Guyatt AL, Jackson VE, *et al.*; SpiroMeta Consortium; International COPD Genetics Consortium. Genetic landscape of chronic obstructive pulmonary disease identifies heterogeneous cell-type and phenotype associations. *Nat Genet* 2019;51:494–505.
5. Ingebrigtsen T, Thomsen SF, Vestbo J, van der Sluis S, Kyvik KO, Silverman EK, *et al.* Genetic influences on chronic obstructive pulmonary disease—a twin study. *Respir Med* 2010;104:1890–1895.

6. Meteran H, Backer V, Kyvik KO, Skytthe A, Thomsen SF. Comorbidity between chronic obstructive pulmonary disease and type 2 diabetes: a nation-wide cohort twin study. *Respir Med* 2015;109:1026–1030.
7. Machin M, Amaral AFS, Wielscher M, Rezwan FI, Imboden M, Jarvelin M-R, *et al.*; ALEC study. Systematic review of lung function and COPD with peripheral blood DNA methylation in population based studies. *BMC Pulm Med* 2017;17:54.
8. Morrow JD, Cho MH, Hersh CP, Pinto-Plata V, Celli B, Marchetti N, *et al.* DNA methylation profiling in human lung tissue identifies genes associated with COPD. *Epigenetics* 2016;11:730–739.
9. Morrow JD, Glass K, Cho MH, Hersh CP, Pinto-Plata V, Celli B, *et al.* Human lung DNA methylation quantitative trait loci colocalize with chronic obstructive pulmonary disease genome-wide association loci. *Am J Respir Crit Care Med* 2018;197:1275–1284.
10. Vucic EA, Chari R, Thu KL, Wilson IM, Cotton AM, Kennett JY, *et al.* DNA methylation is globally disrupted and associated with expression changes in chronic obstructive pulmonary disease small airways. *Am J Respir Cell Mol Biol* 2014;50:912–922.
11. Faner R, Morrow JD, Casas-Recasens S, Cloonan SM, Noell G, López-Giraldo A, *et al.* Do sputum or circulating blood samples reflect the pulmonary transcriptomic differences of COPD patients? A multi-tissue transcriptomic network META-analysis. *Respir Res* 2019;20:5.
12. Burgel P-R, Bourdin A, Chanez P, Chabot F, Chaouat A, Chinet T, *et al.* Update on the roles of distal airways in COPD. *Eur Respir Rev* 2011; 20:7–22.
13. Lindberg A, Linder R, Backman H, Eriksson Ström J, Frølich A, Nilsson U, *et al.* From COPD epidemiology to studies of pathophysiological disease mechanisms: challenges with regard to study design and recruitment process: respiratory and cardiovascular effects in COPD (KOLIN). *Eur Clin Respir J* 2017;4:1415095.
14. Venables WN, Ripley BD. *Modern applied statistics with S.* New York: Springer Science & Business Media; 2013.
15. Peters TJ, Buckley MJ, Statham AL, Pidsley R, Samaras K, V Lord R, *et al.* De novo identification of differentially methylated regions in the human genome. *Epigenetics Chromatin* 2015;8:6.
16. Kamburov A, Wierling C, Lehrach H, Herwig R. ConsensusPathDB—a database for integrating human functional interaction networks. *Nucleic Acids Res* 2009;37:D623–D628.
17. Kamburov A, Pentchev K, Galicka H, Wierling C, Lehrach H, Herwig R. ConsensusPathDB: toward a more complete picture of cell biology. *Nucleic Acids Res* 2011;39:D712–D717.
18. Horvath S. DNA methylation age of human tissues and cell types. *Genome Biol* 2013;14:R115.
19. Ringh MV, Hagemann-Jensen M, Needhamsen M, Kular L, Breeze CE, Sjöholm LK, *et al.* Tobacco smoking induces changes in true DNA methylation, hydroxymethylation and gene expression in bronchoalveolar lavage cells. *EBioMedicine* 2019;46:290–304.
20. Heng TSP, Painter MW; Immunological Genome Project Consortium. The Immunological Genome Project: networks of gene expression in immune cells. *Nat Immunol* 2008;9:1091–1094.
21. Röpcke S, Holz O, Lauer G, Müller M, Rittinghausen S, Ernst P, *et al.* Repeatability of and relationship between potential COPD biomarkers in bronchoalveolar lavage, bronchial biopsies, serum, and induced sputum. *PLoS One* 2012;7:e46207.
22. Joehanes R, Just AC, Marioni RE, Pilling LC, Reynolds LM, Mandaviya PR, *et al.* Epigenetic signatures of cigarette smoking. *Circ Cardiovasc Genet* 2016;9:436–447.
23. Yang W, Ni H, Wang H, Gu H. NLRP3 inflammasome is essential for the development of chronic obstructive pulmonary disease. *Int J Clin Exp Pathol* 2015;8:13209–13216.
24. Rovina N, Koutsoukou A, Koulouris NG. Inflammation and immune response in COPD: where do we stand? *Mediators Inflamm* 2013; 2013:413735.
25. Birrell MA, Eltom S. The role of the NLRP3 inflammasome in the pathogenesis of airway disease. *Pharmacol Ther* 2011;130: 364–370.
26. Hu W-P, Zeng Y-Y, Zuo Y-H, Zhang J. Identification of novel candidate genes involved in the progression of emphysema by bioinformatic methods. *Int J Chron Obstruct Pulmon Dis* 2018;13:3733–3747.
27. Tasena H, Faiz A, Timens W, Noordhoek J, Hylkema MN, Gosens R, *et al.* microRNA-mRNA regulatory networks underlying chronic mucus hypersecretion in COPD. *Eur Respir J* 2018;52:1701556.
28. Barnawi J, Jersmann H, Haberberger R, Hodge S, Meech R. Reduced DNA methylation of sphingosine-1 phosphate receptor 5 in alveolar macrophages in COPD: a potential link to failed efferocytosis. *Respirology* 2017;22:315–321.
29. Monick MM, Beach SRH, Plume J, Sears R, Gerrard M, Brody GH, *et al.* Coordinated changes in AHRR methylation in lymphoblasts and pulmonary macrophages from smokers. *Am J Med Genet B Neuropsychiatr Genet* 2012;159B:141–151.
30. Kaur G, Begum R, Thota S, Batra S. A systematic review of smoking-related epigenetic alterations. *Arch Toxicol* 2019;93:2715–2740.
31. Kagoshima M, Wilcke T, Ito K, Tsaprouni L, Barnes PJ, Punchard N, *et al.* Glucocorticoid-mediated transrepression is regulated by histone acetylation and DNA methylation. *Eur J Pharmacol* 2001;429:327–334.
32. Wan ES, Qiu W, Baccarelli A, Carey VJ, Bacherman H, Rennard SI, *et al.* Systemic steroid exposure is associated with differential methylation in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2012;186:1248–1255.
33. van Nijnatten J, Brandsma C-A, Hiemstra P, Timens W, van den Berge M, Faiz A. Changes in DNA methylation after corticosteroids treatment in COPD patients [abstract]. *Eur Respir J* 2020;56:A4783.
34. Ahn SV, Lee E, Park B, Jung JH, Park JE, Sheen SS, *et al.* Cancer development in patients with COPD: a retrospective analysis of the National Health Insurance Service–National Sample Cohort in Korea. *BMC Pulm Med* 2020;20:170.
35. Proskocil BJ, Fryer AD. Beta2-agonist and anticholinergic drugs in the treatment of lung disease. *Proc Am Thorac Soc* 2005;2:305–310, discussion 311–312.
36. Oldenburger A, Maarsingh H, Schmidt M. Multiple facets of cAMP signalling and physiological impact: cAMP compartmentalization in the lung. *Pharmaceuticals (Basel)* 2012;5:1291–1331.
37. Belmonte KE. Cholinergic pathways in the lungs and anticholinergic therapy for chronic obstructive pulmonary disease. *Proc Am Thorac Soc* 2005;2:297–304, discussion 311–312.
38. Nuñez FJ, Johnstone TB, Corpuz ML, Kazarian AG, Mohajer NN, Tliba O, *et al.* Glucocorticoids rapidly activate cAMP production via $G_{\alpha s}$ to initiate non-genomic signaling that contributes to one-third of their canonical genomic effects. *FASEB J* 2020;34:2882–2895.
39. Dunne AE, Kawamatawong T, Fenwick PS, Davies CM, Tullett H, Barnes PJ, *et al.* Direct inhibitory effect of the PDE4 inhibitor roflumilast on neutrophil migration in chronic obstructive pulmonary disease. *Am J Respir Cell Mol Biol* 2019;60:445–453.
40. Mercado N, Ito K, Barnes PJ. Accelerated ageing of the lung in COPD: new concepts. *Thorax* 2015;70:482–489.
41. MacNee W. Is chronic obstructive pulmonary disease an accelerated aging disease? *Ann Am Thorac Soc* 2016;13:S429–S437.
42. McEwen LM, Jones MJ, Lin DTS, Edgar RD, Husquin LT, MacIsaac JL, *et al.* Systematic evaluation of DNA methylation age estimation with common preprocessing methods and the Infinium MethylationEPIC BeadChip array. *Clin Epigenetics* 2018;10:123.
43. López-Otín C, Blasco MA, Partridge L, Serrano M, Kroemer G. The hallmarks of aging. *Cell* 2013;153:1194–1217.
44. Brandsma C-A, de Vries M, Costa R, Woldhuis RR, Königshoff M, Timens W. Lung ageing and COPD: is there a role for ageing in abnormal tissue repair? *Eur Respir Rev* 2017;26:170073.
45. Li L, Li W, Chen N, Zhao H, Xu G, Zhao Y, *et al. FLI1* exonic circular RNAs as a novel oncogenic driver to promote tumor metastasis in small cell lung cancer. *Clin Cancer Res* 2019;25:1302–1317.
46. Lin S-F, Wu C-C, Chai C-Y. Increased FLI-1 expression is associated with poor prognosis in non-small cell lung cancers. *Appl Immunohistochem Mol Morphol* 2016;24:556–561.
47. Kichaev G, Bhatia G, Loh P-R, Gazal S, Burch K, Freund MK, *et al.* Leveraging polygenic functional enrichment to improve GWAS power. *Am J Hum Genet* 2019;104:65–75.
48. Oudijk E-JD, Nijhuis EHJ, Zwank MD, van de Graaf EA, Mager HJ, Coffer PJ, *et al.* Systemic inflammation in COPD visualised by gene profiling in peripheral blood neutrophils. *Thorax* 2005;60:538–544.
49. Hoenderdos K, Condliffe A. The neutrophil in chronic obstructive pulmonary disease. *Am J Respir Cell Mol Biol* 2013;48:531–539.
50. Hobbs BD, de Jong K, Lamontagne M, Bossé Y, Shrine N, Artigas MS, *et al.*; COPDGene Investigators; ECLIPSE Investigators; LifeLines Investigators; SPIROMICS Research Group; International COPD Genetics Network Investigators; UK BiLEVE Investigators; International

COPD Genetics Consortium. Genetic loci associated with chronic obstructive pulmonary disease overlap with loci for lung function and pulmonary fibrosis. *Nat Genet* 2017;49:426–432.
51. Hobbs BD, Cho MH. Dissecting respiratory disease heterogeneity through the genetics of diffusing capacity. *Eur Respir J* 2018;52: 1801468.
52. Hughes JMB, Pride NB. Examination of the carbon monoxide diffusing capacity (D_{LCO}) in relation to its KCO and V_A components. *Am J Respir Crit Care Med* 2012;186:132–139.
53. Cui H, Wang Y, Huang H, Yu W, Bai M, Zhang L, *et al.* GPR126 protein regulates developmental and pathological angiogenesis through modulation of VEGFR2 receptor signaling. *J Biol Chem* 2014;289:34871–34885.
54. Konigsberg IR, Maier LA, Yang IV. Epigenetics and sarcoidosis. *Eur Respir Rev* 2021;30:210076.
55. Anastasiadi D, Esteve-Codina A, Piferrer F. Consistent inverse correlation between DNA methylation of the first intron and gene expression across tissues and species. *Epigenetics Chromatin* 2018;11:37.