A Dynamic Bronchial Airway Gene Expression Signature of Chronic Obstructive Pulmonary Disease and Lung Function Impairment

Katrina Steiling[1,2], Maarten van den Berge[3], Kahkeshan Hijazi[2], Roberta Florido[4], Joshua Campbell[2], Gang Liu[1], Ji Xiao[1], Xiaohui Zhang[1], Grant Duclos[1], Eduard Drizik[1], Huiqing Si[1], Catalina Perdomo[1], Charles Dumont[5], Harvey O. Coxson[6], Yuriy O. Alekseyev[7], Don Sin[8], Peter Pare[8], James C. Hogg[8], Annette McWilliams[9], Pieter S. Hiemstra[10], Peter J. Sterk[11], Wim Timens[12], Jeffrey T. Chang[13], Paola Sebastiani[14], George T. O'Connor[5], Andrea H. Bild[15], Dirkje S. Postma[3], Stephen Lam[9], Avrum Spira[1,2,7]*, and Marc E. Lenburg[1,2,7]*

[1]Division of Computational Biomedicine, [4]Department of Medicine, [5]Division of Pulmonary, Allergy, and Critical Care Medicine, and [7]Department of Pathology and Laboratory Medicine, Boston University School of Medicine, Boston, Massachusetts; [2]Bioinformatics Program, Boston University, Boston, Massachusetts; [3]Department of Pulmonary Diseases and [12]Department of Pathology, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands; [6]Department of Radiology, Vancouver General Hospital, University of British Columbia, Vancouver, British Columbia, Canada; [8]The James Hogg Research Centre, St. Paul's Hospital, Vancouver, British Columbia, Canada; [9]British Columbia Cancer Agency, Vancouver, British Columbia, Canada; [10]Department of Pulmonology, University Medical Center, Leiden, The Netherlands; [11]Department of Respiratory Medicine, University of Amsterdam, Amsterdam, The Netherlands; [13]Department of Integrative Biology and Pharmacology, The University of Texas Health Science Center in Houston, Houston, Texas; [14]Department of Biostatistics, Boston University School of Public Health, Boston, Massachusetts; and [15]Department of Pharmacology and Toxicology, University of Utah, Salt Lake City, Utah

Rationale: **Molecular phenotyping of chronic obstructive pulmonary disease (COPD) has been impeded in part by the difficulty in obtaining lung tissue samples from individuals with impaired lung function.**
Objectives: **We sought to determine whether COPD-associated processes are reflected in gene expression profiles of bronchial airway epithelial cells obtained by bronchoscopy.**
Methods: **Gene expression profiling of bronchial brushings obtained from 238 current and former smokers with and without COPD was performed using Affymetrix Human Gene 1.0 ST Arrays.**
Measurements and Main Results: **We identified 98 genes whose expression levels were associated with COPD status, FEV_1% predicted, and FEV_1/FVC. *In silico* analysis identified activating transcription factor 4 (ATF4) as a potential transcriptional regulator of genes with COPD-associated airway expression, and ATF4 overexpression in airway epithelial cells *in vitro* recapitulates COPD-associated gene expression changes. Genes with COPD-associated expression in the bronchial airway epithelium had similarly altered expression profiles in prior studies performed on small-airway epithelium and lung parenchyma, suggesting that transcriptomic alterations in the bronchial airway epithelium reflect molecular events found at more distal sites of disease activity. Many of the airway COPD-associated gene expression changes revert toward baseline after therapy with the inhaled corticosteroid fluticasone in independent cohorts.**
Conclusions: **Our findings demonstrate a molecular field of injury throughout the bronchial airway of active and former smokers with COPD that may be driven in part by ATF4 and is modifiable with therapy. Bronchial airway epithelium may ultimately serve as a relatively accessible tissue in which to measure biomarkers of disease activity for guiding clinical management of COPD.**

AT A GLANCE COMMENTARY

Scientific Knowledge on the Subject

Molecular phenotyping of chronic obstructive pulmonary disease (COPD) has been hindered by the limited availability and the heterogeneity of lung tissue specimens from individuals with impaired lung function.

What This Study Adds to the Field

This study demonstrates a COPD-associated "field of injury" measurable by gene expression profiling of the bronchial airway epithelium. The bronchial airway signature of COPD is similarly altered in lung tissue and reverts toward baseline after inhaled fluticasone. These findings suggest that the bronchial airway may ultimately serve as a relatively accessible tissue in which to measure biomarkers to guide clinical management of COPD.

(*Received in original form August 13, 2012; accepted in final form January 22, 2013*)

* These authors contributed equally as co–senior authors.

Supported by NIH/NHLBI 1R01 HL095388 (A.S. and M.E.L.); KL2RR025770 (K.S.); and U01CA96109, PO1 CA096964-01A1, and N01-CN 35000 (S.L.).

Author Contributions: Study design and experimental planning, A.S., M.E.L., D.S., G.T.O'C., J.C.H., K.S., P.P., S.L., and A.H.B. Execution of experiments and data collection, A.M., D.S., D.S.P., G.L., J.X., X.Z., M.v.d.B., P.J.S., P.S.H., S.L., W.T., Y.O.A., G.D., C.P., E.D., H.S., and H.O.C. Data analysis and interpretation, A.S., M.E.L., G.T.O'C., J.C., K.H., K.S., M.v.d.B., P.S., R.F., G.D., C.P., E.D., A.H.B., J.T.C., and C.D. Preparation of the manuscript, A.S., K.S., and M.E.L.

Correspondence and requests for reprints should be addressed to Avrum Spira, M.D., M.Sc., Boston University School of Medicine, Division of Computational Biomedicine, Department of Medicine, 72 East Concord Street, Evans 631, Boston, MA 02118. E-mail: aspira@bu.edu; or Marc E. Lenburg, Ph.D., Boston University School of Medicine, Division of Computational Biomedicine, Department of Medicine, 72 East Concord Street, Evans 631, Boston, MA 02118. E-mail: mlenburg@bu.edu

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

Am J Respir Crit Care Med Vol 187, Iss. 9, pp 933–942, May 1, 2013

Originally Published in Press as DOI: 10.1164/rccm.201208-1449OC on March 7, 2013
Internet address: www.atsjournals.org

Keywords: chronic obstructive pulmonary disease; gene expression profiling; biologic markers

Chronic obstructive pulmonary disease (COPD) affects 14.8 million individuals in the United States alone (1) and is the third leading cause of death (2). Although biologic processes, such as proteinase-antiproteinase imbalance, chronic inflammation, apoptosis, and oxidative stress, have been proposed to play a role in COPD pathogenesis, knowledge remains limited about how these molecular processes impact the clinical presentation and progression of COPD.

Genome-wide gene expression profiling provides a powerful way to survey COPD-associated molecular alterations, but this

approach has been hindered by the limited availability of lung tissue samples from individuals with impaired lung function. As a result, studies of whole-genome gene expression profiling of lung tissue in COPD (3–7) have been limited by small sample sizes and confounding variables, such as the presence of adjacent lung cancer. The development of a less invasive method for measuring COPD-associated cellular and molecular processes would allow for the study of large cohorts and the potential for identifying molecular subtypes of COPD and clinically useful predictors of prognosis and response to therapy.

Our group has shown that there are alterations in airway epithelial gene expression among current and former smokers that can serve as a tool for the early detection of lung cancer. Specifically, we have found that the expression levels of genes in cytologically normal large airway epithelial cells can serve as a sensitive and specific diagnostic biomarker for lung cancer (8). Airway gene expression also reflects PI3K pathway activation in smokers with airway epithelial cell dysplasia that is reversible with the candidate lung cancer chemoprevention agent *myo*-inositol (9). Importantly, PI3K is also activated in tumors, suggesting that the airway can potentially serve as a surrogate for assessing some disease-associated processes. The impact of lung cancer on airway gene expression suggests that the airway epithelium might also be impacted by other smoking-related diseases, such as COPD. Two small studies have demonstrated COPD-associated expression differences in airway epithelium, but focused on the expression of a limited number of genes hypothesized to be involved in the pathogenesis of COPD (10, 11). Moreover, the relationship of these airway gene expression changes to those that occur with COPD in lung tissue remains unstudied (10, 11), and it is unclear if the bronchial airway can be used as a more readily available biospecimen for identifying and measuring the activity of distal COPD-associated processes to guide clinical decisions in COPD management.

In the present study, whole-genome gene expression profiling was performed on bronchial brushings in a cohort of 238 current and former smokers with and without COPD. We investigated (*1*) the association of airway epithelial gene expression with COPD and continuous measures of lung function, (*2*) the relationship of COPD-associated gene expression changes in airway epithelium with those that occur in lung tissue, and (*3*) the reversibility of COPD-associated alterations in airway gene expression by inhaled corticosteroids. We have, for the first time, identified a COPD-associated field of injury that can be measured by gene expression profiling of the airway epithelium and that is modifiable with therapy.

Some of the results reported here have been previously published in abstract form (12).

METHODS

See the online supplement for a more detailed description of the methods.

Patient Population and Sample Processing

Bronchial airway brushings were obtained during bronchoscopy from subjects who were being followed longitudinally for the development of lung cancer at the British Columbia Cancer Research Agency between June 2000 and May 2009 as part of the British Columbia Lung Health Study (13) and the Pan-Canadian Lung Health Study. A total of 267 bronchial brushing samples were selected to ensure matching for covariates (Table 1). All subjects provided written informed consent. Institutional review board approval was obtained from participating institutions. High-molecular-weight RNA isolated from the bronchial brushings using the miRNeasy mini kit (Qiagen, Valencia, CA) was processed and hybridized to Affymetrix Human Gene 1.0 ST Arrays (Affymetrix, Santa Clara, CA).

TABLE 1. CLINICAL CHARACTERISTICS OF THE STUDY POPULATION

	COPD ($n = 87$)*	No COPD ($n = 151$)*	P Value†
Age, yr	65 (6)	64 (6)	0.25
Sex			0.5
Male	52	83	
Female	35	68	
Smoking status			0.1
Current	30	69	
Former	57	82	
Pack-years	51 (25)‡	47 (19)‡	0.11
FEV_1% predicted	60 (14)	93 (13)	$<10^{-4}$
FEV_1/FVC	0.56 (0.09)	0.75 (0.06)	$<10^{-4}$
Years since smoking cessation	11.84 (9.86)	11.11 (6.73)	0.52
Inhaled corticosteroid use	18 (21%)	7 (5%)	$<10^{-3}$
Inhaled bronchodilator use	21 (24%)	11 (7%)	$<10^{-3}$
Statin use	23 (26%)	23 (15%)	0.041
Nonsteroidal antiinflammatory drug use	21 (24%)	46 (30%)	0.37

Definition of abbreviation: COPD = chronic obstructive pulmonary disease. The mean and standard deviation are shown for continuous variables.
* A total of 97% of the subjects were white.
† P values calculated using a Student's t test or Fisher exact test.
‡ Missing pack-years for 5 subjects with COPD and 11 subjects without COPD.

Data Acquisition, Probe Set Summarization and Normalization, and Data Preprocessing

A total of 269 arrays from 267 samples including two samples run in duplicate were used for the generation of gene expression levels. The array data for two subjects were excluded because of sample annotation concerns, leaving a total of 265 samples. To minimize the potential confounding effect of lung cancer, data from 19 subjects with a diagnosis of lung cancer as of January 2010 were excluded as were data from eight subjects who lacked lung function testing within 1 year of their study bronchoscopy, leaving a total of 238 samples. All of the 238 remaining samples were of adequate quality for subsequent analysis (*see* Figure E1 in the online supplement).

Determination of Gene Expression Associated with COPD and Continuous Measures of Lung Function

Using spirometry measurements obtained within 1 year of bronchoscopy, COPD was defined as the presence of both an FEV_1/FVC less than or equal to 0.7 and FEV_1% predicted less than 80, based on standard reference equations (13, 14). Genes whose expression levels were associated with COPD and/or continuous measures of lung function were identified by an analysis of variance false discovery rate (FDR) of less than 0.05 and a linear fold change (FC) of greater than 1.25 between COPD and no COPD after controlling for major demographic variables and risk factors for COPD.

Enrichment Analysis

Functional enrichment analysis for genes associated with COPD was performed using DAVID 6.7b (15). Transcription factor binding site enrichment analysis was performed using GATHER (16). Additional predicted targets of selected transcription factors were identified with patser using Transfac version 12.1 (17, 18). Gene set enrichment analysis (GSEA) was used to determine the relationship between our results and previously published studies as detailed in the online supplement METHODS section and Table E1. A false-discovery rate threshold of FDR less than 0.05 was used to determine significant enrichment by GSEA.

Real-Time Polymerase Chain Reaction Validation of Gene Expression Associated with COPD

Quantitative real-time polymerase chain reaction was performed on nine genes to confirm COPD-associated expression changes.

Overexpression of Activating Transcription Factor 4 in Cultured Airway Epithelial Cells

To determine the relationship between activating transcription factor 4 (ATF4) expression and the airway COPD signature, we overexpressed ATF4 in immortalized human bronchial epithelial cells (BEAS2B). Total RNA was isolated from cells transfected with ATF4 (n = 3) or empty vector (n = 3), processed, labeled, and hybridized to Affymetrix Human Gene 1.0 ST Arrays. Gene expression levels associated with ATF4 overexpression were ranked according to the *t* statistic, and enrichment of the airway COPD signature was determined using GSEA.

RESULTS

Characteristics of the Study Population

There were no significant differences in age, cumulative smoking exposure, or smoking status between the 87 subjects with COPD and the 151 subjects without COPD (Table 1). Subjects with COPD had lower FEV_1% predicted and FEV_1/FVC than the control group. The FEV_1 across subjects with COPD ranged from 15–79% of the reference value, with most subjects with COPD having moderate disease (Global Initiative for Chronic Obstructive Lung Disease [GOLD] Grade 2) (19) as expected from a bronchoscopy-based cohort. A minority of the study population used inhaled corticosteroids or inhaled bronchodilators, with a statistically significant association with COPD status. Of the 14 subjects without COPD taking an inhaled medication, 3 had a history of asthma. A total of 17 subjects (eight with COPD and nine without COPD) reported a history of asthma. Groups also differed with respect to statin use. There was no significant difference in the use of nonsteroidal antiinflammatory drugs.

Bronchial Epithelial Gene Expression Associated with COPD and Continuous Measures of Lung Function

The expression levels of 107 genes were associated with COPD (FDR < 0.05; FC > 1.25) after adjusting for major demographic variables and risk factors for COPD including age, sex, smoking status, and cumulative smoke exposure. The expression levels of 110 genes were associated with FEV_1% predicted, and 102 with FEV_1/FVC as continuous measures. The expression profiles of 98 genes were associated with all three measures; 54 of these genes were increased, and 44 were decreased in COPD (Figure 1). This bronchial airway signature of COPD includes dihydropyrimidinase-like 3, CEACAM5, Sushi-repeat containing protein X-linked, and enoyl CoA delta isomerase 2, four genes described in prior studies as irreversibly altered by cigarette smoke even decades after smoking cessation (20, 21). Among individuals with COPD, cluster membership in Figure 1 was significantly associated with FEV1% predicted but no other clinical covariates or RNA quality (*see* Table E2). Further analysis of potential sources of the gene expression variability within classes is presented next.

To determine whether asthma, inhaled medications, statin use, or the method of COPD classification affected this analysis, we repeated the analysis excluding individuals with a self-reported history of asthma (n = 17), individuals using an inhaled corticosteroid or bronchodilator (n = 37), individuals using a statin medication (n = 46), or individuals with mild decreases in FEV_1% predicted (range, 70–80%; n = 49). We identified a consistent relationship between COPD-associated changes in airway gene expression in each of these analyses, with 80–99% of the 98 COPD-associated genes also showing an FDR less than 0.05 and FC greater than 1.25 in these analyses (*see* Table E3). We did not detect significant correlation of a metagene summarizing the COPD airway gene expression signature with years since quitting smoking among former smokers ($P > 0.05$). We also failed to detect significant association between COPD status and a metagene representing inflammatory cell-specific gene expression (21). When the inflammatory cell metagene was included as a covariate in the linear model, all 98 COPD-associated genes remained significant at FDR less than 0.05 and FC greater than 1.25.

To computationally validate the association of these genes with COPD, we performed GSEA using a publicly available whole-genome gene expression dataset of small airway epithelium (10th–12th generation bronchi) that included 12 healthy smokers and 4 smokers with COPD in GOLD Grade 1–2 severity (GSE5058) (11). We identified a concordant relationship between the 98 genes whose expression patterns were associated with COPD in the present study and COPD-associated gene expression differences observed in this dataset ($FDR_{GSEA} < 0.05$) (*see* Figure E2). We also experimentally validated the COPD-associated expression pattern of nine genes by quantitative real-time polymerase chain reaction (*see* Figure E3). Together, these data suggest a COPD-associated bronchial airway field of injury that reflects the presence and severity of COPD and that is consistent with COPD-associated gene expression changes in small airway epithelium.

To explore the biologic function of the 98 genes whose expression levels were associated with COPD, FEV_1% predicted, and FEV_1/FVC, genes were subdivided into two groups: higher expression in COPD and lower expression in COPD (Figure 1). Both lists were significantly enriched for genes belonging to a variety of functional categories (*see* Table E4) including glycoproteins (up-regulated), proteins involved in the acute inflammatory response (up-regulated), and epidermal growth factor (EGF)-like domains (down-regulated). These findings suggest that these gene expression changes reflect COPD-associated alterations in processes related to the inflammatory response and regulation of cell growth in bronchial airway epithelium.

ATF4 as a Mediator of Airway Gene Expression Alterations Associated with COPD

To explore potential regulators of COPD-associated changes in gene expression, we used GATHER to identify transcription factor binding sites enriched in the regulatory regions of differentially expressed genes. We identified enrichment of binding sites for ATF4 and CREB1 among the 98 genes with COPD-associated expression differences ($P < 0.001$) (*see* Table E5). To explore a potential mechanistic role for ATF4 in regulating COPD-associated gene expression differences, we examined the effects of overexpressing ATF4 in the BEAS2B bronchial epithelium cell line and found that this resulted in an increase in many of the same genes that are expressed at higher levels in the bronchial airway of individuals with COPD (Figure 2A; *see* Figure E4). Furthermore, all 13 of the core enrichment genes increased by both ATF4 overexpression and in the airway COPD signature are predicted targets of ATF4 (Figures 2B and 2C) (17, 18). These findings suggest that overexpression of ATF4 is sufficient to recapitulate a component of the airway gene expression differences associated with the presence of COPD *in vivo*, and that ATF4 might be a mediator of these changes.

The Relationship between COPD-associated Gene Expression in the Bronchial Airway Epithelium and in Lung Parenchyma

We next examined whether COPD-associated gene expression changes in the bronchial airway reflect disease-associated processes in lung parenchyma. By GSEA, we found concordant enrichment of gene expression changes in bronchial airway and lung tissue in three previously published COPD datasets ($FDR_{GSEA} < 0.05$) (Figure 3). Genes whose expression levels

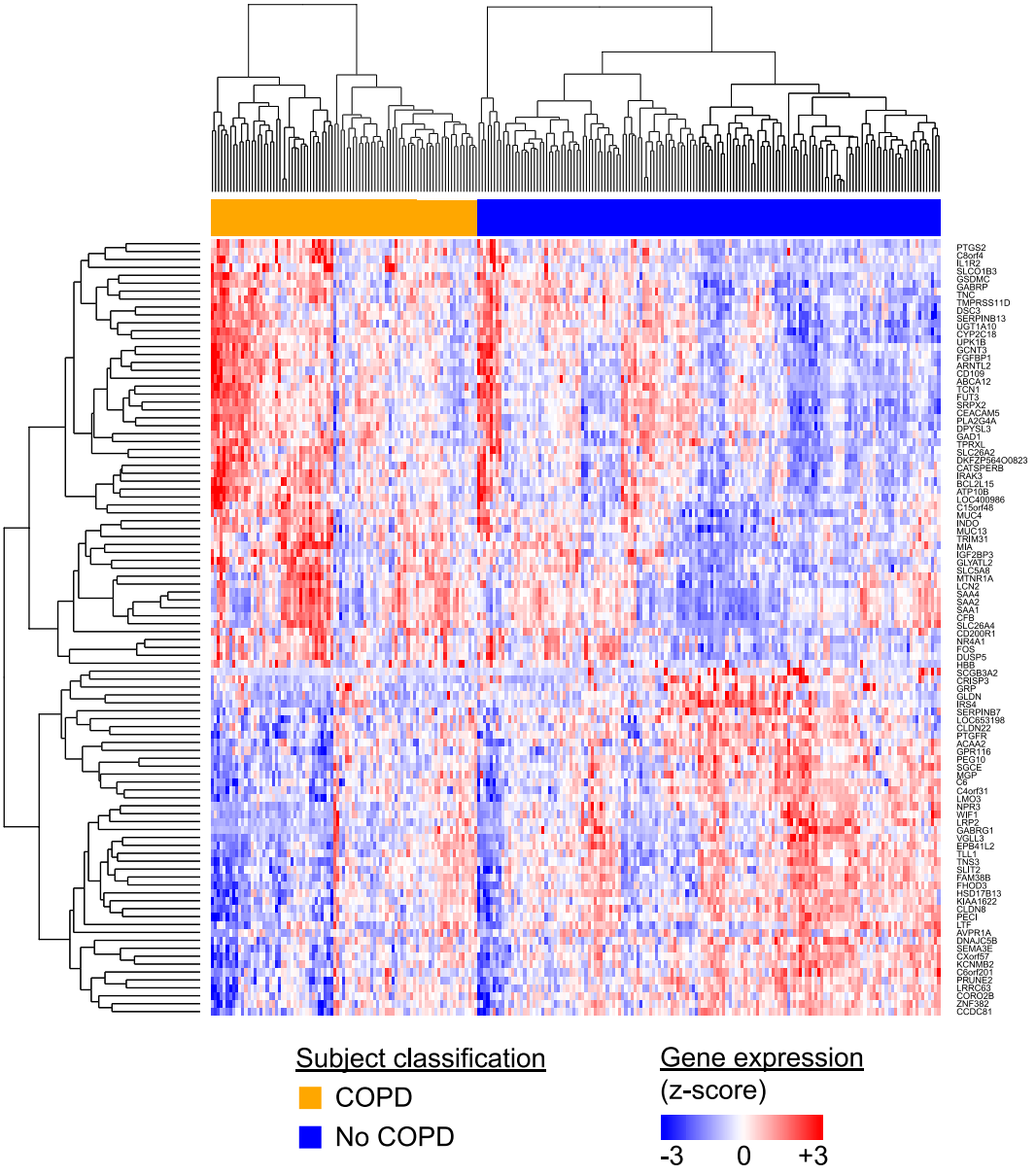

Figure 1. Semisupervised heatmap of the 98 genes associated with chronic obstructive pulmonary disease (COPD) and continuous measures of lung function. A total of 107 genes were associated with COPD, 110 genes with FEV_1% predicted, and 101 genes with FEV_1/FVC (false discovery rate < 0.05; fold change > 1.25). Ninety-eight genes were in common to all of these measures. These results demonstrate that airway epithelial gene expression reflects the presence of COPD and the severity of lung function impairment.

were increased in the lung tissue of GOLD Grade 2 subjects compared with GOLD Grade 0 subjects (5) or negatively correlated with $FEF_{25–75\%}$ (7) were enriched among genes whose expression was increased in the bronchial airway with COPD. Similarly, genes down-regulated in the lung tissue of cases of COPD compared with control subjects (6) or in lung tissue from subjects with worse lung function (6) were enriched among genes whose expression was decreased in the bronchial airway epithelium in COPD. There were also similarities between COPD-associated airway gene expression and lung parenchymal gene expression when gene expression profiles from these previously published datasets were ranked according to the strength of association with COPD or COPD-related traits ($FDR_{GSEA} < 0.05$) (*see* Figure E2) and interrogated with the disease-associated genes we identified in the bronchial airway. These findings demonstrate that similar changes in gene expression occur in the airway epithelium and lung tissue, suggesting that the COPD-associated airway gene expression differences mirror aspects of disease processes occurring in lung tissue.

To further explore the relationship between bronchial epithelial and lung tissue gene expression related to COPD, we used GSEA to examine the distribution of the 98 genes whose expression levels were associated with COPD in a ranking of all genes according to their expression change in lung parenchyma as a function of mean linear intercept, a morphologic measure of emphysema (GSE27597) (22). Lung parenchymal genes whose expression levels increased with regional emphysema severity were enriched for bronchial epithelial genes whose expression was increased in COPD ($FDR_{GSEA} < 0.05$) (Figure 4). The genes contributing most strongly to this enrichment included SERPINB13, a serine peptidase inhibitor, and TMPRSS11D, a trypsin-like protease. These findings support the biologic relevance of the bronchial epithelial gene expression signature of COPD by linking it to clinical and pathologic measures of disease severity.

Reversibility of COPD-associated Changes in Airway Epithelial Gene Expression with Treatment

Because inhaled corticosteroids are commonly used to treat COPD, we next sought to determine whether COPD-associated changes in airway gene expression were modifiable by fluticasone therapy in

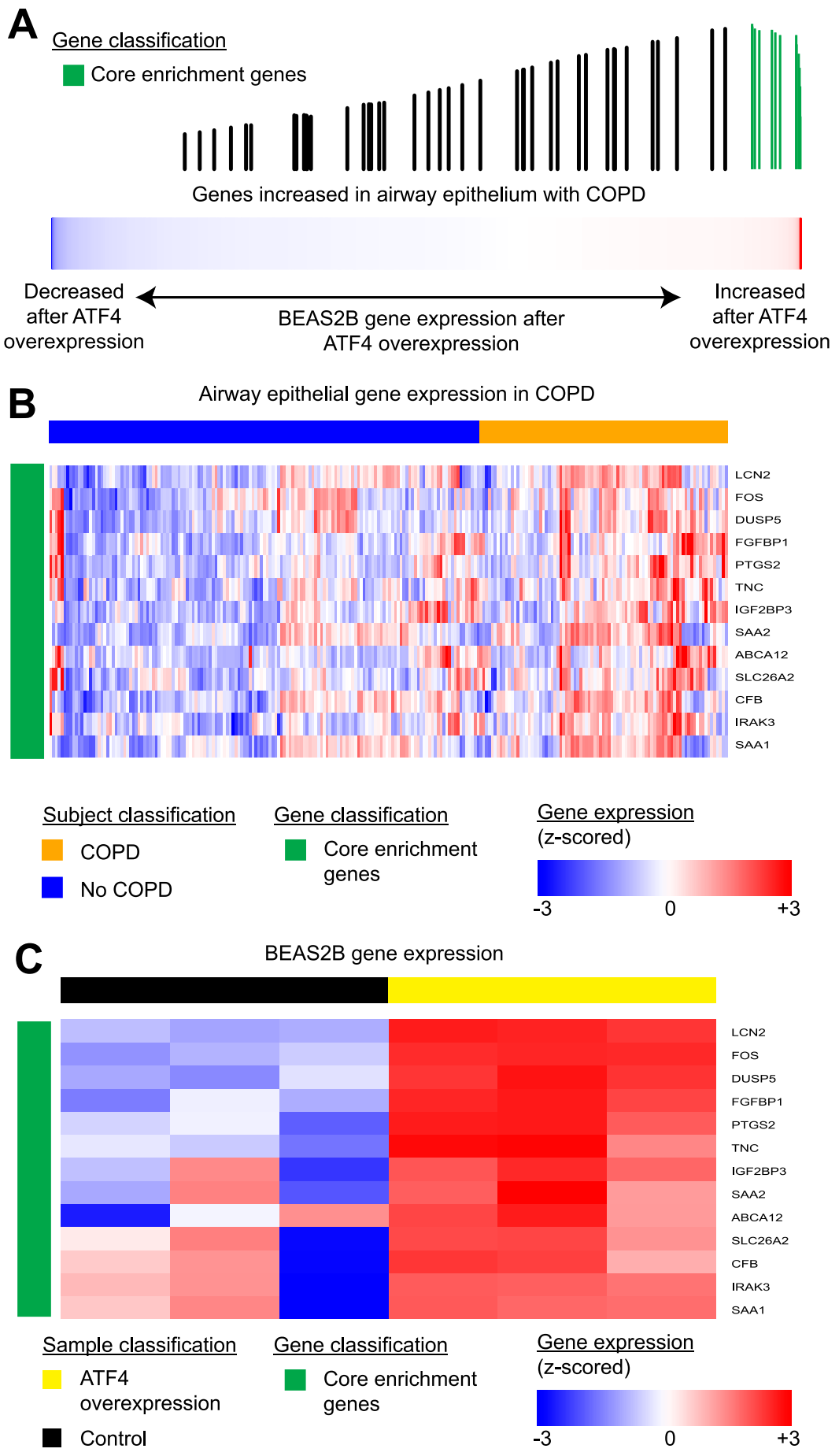

Figure 2. ATF4 overexpression in BEAS2B cells *in vitro* recapitulates the *in vivo* airway gene expression signature of chronic obstructive pulmonary disease (COPD). (*A*) Gene set enrichment analysis demonstrates enrichment of genes with increased expression in airway epithelium from individuals with COPD among genes whose expression is increased with ATF4 overexpression in BEAS2B cells (false discovery rate $<$ 0.05). Genes are ranked from left to right based on their ATF4-associated expression pattern *in vitro*. The position of each *vertical bar* indicates the position of a gene with COPD-associated gene expression in airway epithelium within this ranked list. The height of this bar represents the running gene set enrichment analysis enrichment score. Core enrichment genes are highlighted in *green*. (*B*) Expression levels of the core enrichment genes (*green*) in the bronchial brushing samples, all of which are predicted targets of ATF4 ($P < 0.001$) (17, 18), are shown in this heatmap supervised by COPD status (*orange*, COPD; *blue*, normal). (*C*) Expression levels of the core enrichment genes (*green*) with ATF4 overexpression in airway epithelium *in vitro* (*black*, negative control; *yellow*, ATF4 overexpression).

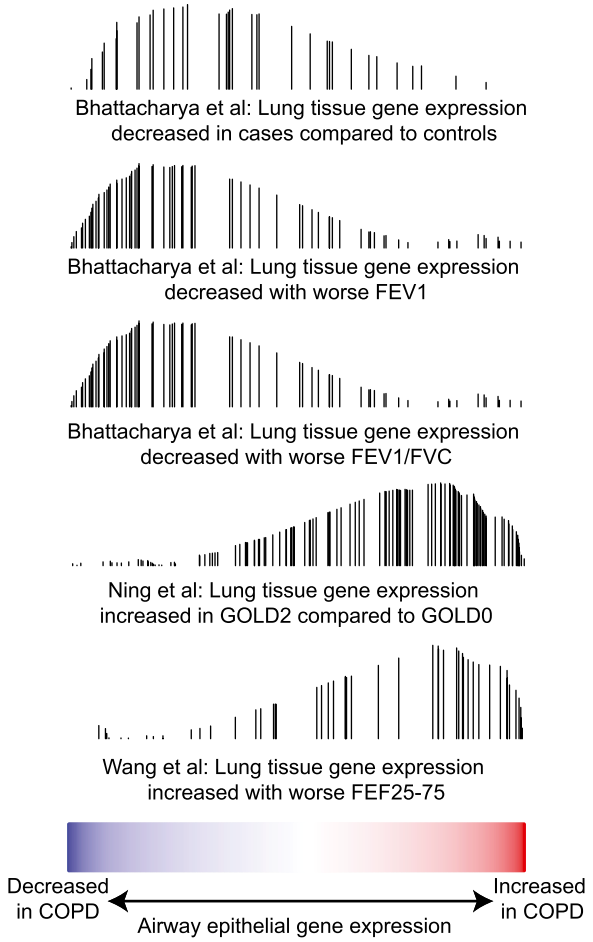

Figure 3. Airway epithelial gene expression associated with chronic obstructive pulmonary disease (COPD) is concordant with previously published microarray datasets of COPD lung tissue. Airway gene expression associated with COPD was compared with gene lists identified in previous studies of lung tissue gene expression in COPD using gene set enrichment analysis. The *color bar* indicates the strength of association of airway epithelial gene expression with COPD as measured by the *t* statistic for the COPD term after adjusting for covariates. The position of each *vertical bar* from left to right indicates the position of a gene from one of the previously published lung parenchyma gene sets (genes whose expression was previously identified to be associated with a COPD-related trait) within the ranked airway gene list. The height of this bar represents the running gene set enrichment analysis enrichment score. This analysis identified concordant enrichment of previously reported COPD-associated gene expression changes in lung tissue and COPD-associated changes in gene expression in the bronchial airway (false discovery rate < 0.05), and suggests that there is a common COPD effect in both tissues.

patients with COPD. We used GSEA to examine the expression of the bronchial epithelial COPD signature in a ranking of gene expression profiles derived from bronchial biopsies obtained from a subset of subjects from the GLUCOLD trial (ClinicalTrials.gov registration number NCT00158847) (GSE36221), an independent longitudinal study of subjects with COPD randomized to fluticasone with or without salmeterol, or placebo (23). Expression levels of the 54 genes increased in the bronchial epithelial COPD signature were enriched among genes whose expression decreased after treatment containing fluticasone ($FDR_{GSEA} < 0.05$) (Figure 5A). Similarly, expression levels of the 44 genes decreased with COPD in the bronchial airway signature were enriched among genes whose expression levels increased after treatment with fluticasone in the GLUCOLD cohort ($FDR_{GSEA} < 0.05$) (Figure 5A). The genes contributing most strongly to the this enrichment included DUSP5, a key regulator of cell proliferation and differentiation; TMPRSS11D, which serves a key role in host defense in the airway; and claudin 8 (CLDN8), which functions in tight junctions between epithelial cells (Figure 5B). These results suggest that a subset of airway gene expression changes associated with COPD can be reversed by inhaled corticosteroids.

To validate our findings in the GLUCOLD cohort, we examined the relationship between the airway gene expression signature of COPD and fluticasone-related gene expression differences from an independent dataset in which gene expression in bronchial epithelium samples from before and after fluticasone treatment was profiled using microarrays (24). Using a linear mixed-effects model, genes were ranked according to their change with fluticasone over time. Using GSEA, we found that the 54 genes up-regulated in the airway COPD signature were enriched among the genes decreased by fluticasone treatment and that the 44 genes down-regulated in the airway COPD signature were enriched among the genes increased by fluticasone treatment ($FDR_{GSEA} < 0.05$) (*see* Figure E5). This finding suggests that fluticasone reverts genes that are altered in the airways of patients with COPD. Taken together with our observations in the GLUCOLD cohort, these data suggest that COPD-associated gene expression patterns are potentially dynamic with therapy.

DISCUSSION

By performing whole-genome gene expression profiling of bronchial brushings in a study of individuals with and without COPD, we have identified a COPD-related bronchial airway field of injury that is defined by gene expression alterations and has several important characteristics. First, the gene expression alterations in this field of injury are associated with COPD and continuous COPD-related measures of lung function. Second, the COPD-associated gene expression field of injury measured in the bronchial airway epithelium is similar to COPD-associated gene expression differences occurring in lung parenchyma. Third, the COPD-associated gene expression field of injury is modifiable with treatment.

We have validated the COPD-associated airway-epithelium gene expression differences we identified in several previously published studies including one study of small-airway gene expression (11), and six studies of lung parenchyma (3–7, 22). These observations suggest a reliable COPD-associated pattern of gene expression in the bronchial airway that is similar to distal COPD-associated gene expression differences. Although the COPD-associated gene expression similarities between the bronchial airway and whole-lung tissue could be caused by similarities between the bronchial airway and either the lung parenchyma or the terminal bronchioles, our data suggest that the accessible bronchial airways reflect disease-associated processes occurring deep in the lung. Importantly, many of the previous studies of COPD-associated gene expression have involved lung tissue that is adjacent to lung cancer. In this study, by leveraging bronchoscopy samples from a lung cancer screening cohort where the prevalence of cancer is low, we were able to profile samples exclusively from a large number of patients without lung cancer. Taken together, these findings suggest that the bronchial airway might serve as a readily accessible biospecimen to measure COPD-related processes in research and clinical settings.

The specific genes within the COPD airway epithelial gene expression signature support the biologic plausibility of this signature. For example, TMPRSS11D, also called human airway trypsin-like protease, localizes to ciliated bronchial epithelial cells and was first isolated from the sputum of patients with

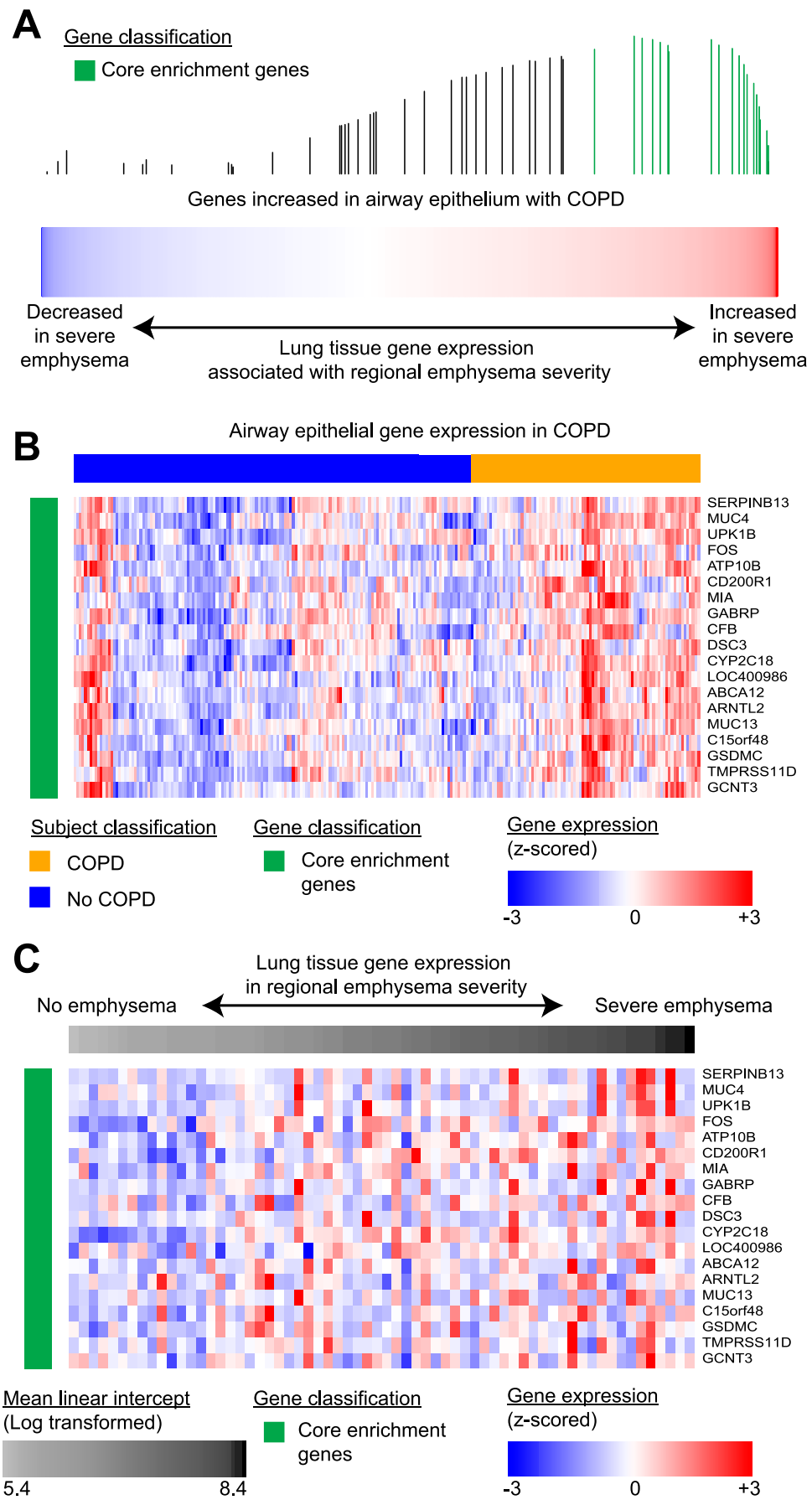

Figure 4. The airway transcriptomic alterations in chronic obstructive pulmonary disease (COPD) reflect gene expression changes associated with emphysema severity in lung tissue. (*A*) Gene set enrichment analysis demonstrates enrichment of genes whose expression levels in the airway epithelium significantly increased in COPD among genes whose expression is increased with worsening emphysema severity in lung tissue (false discovery rate < 0.05). Genes are ranked from left to right based on their emphysema-associated expression pattern in lung tissue. The position of each *vertical bar* indicates the position of a gene whose expression in airway epithelium is associated with COPD within this ranked list. The height of this bar represents the running gene set enrichment analysis enrichment score. The core enrichment genes are highlighted in *green*. (*B*) Expression of the core enrichment genes (*green*) in the bronchial brushing samples is shown in this heatmap supervised by COPD status (*orange*, COPD; *blue*, normal). (*C*) Expression of the core enrichment genes (*green*) in lung tissue samples is shown in this heatmap supervised by emphysema severity (*light gray*, no emphysema; *black*, severe emphysema).

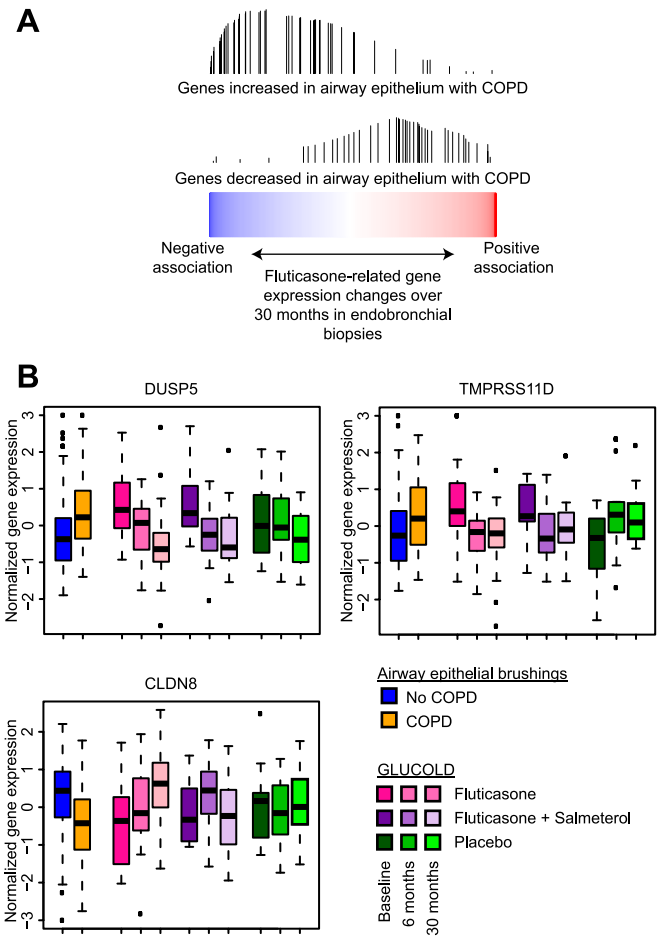

Figure 5. Gene expression changes in the airway of subjects with chronic obstructive pulmonary disease (COPD) are modulated by inhaled corticosteroids. (*A*) Using gene set enrichment analysis, we identified enrichment of airway gene expression associated with COPD in an independent gene expression dataset of endobronchial biopsies obtained at 0, 6, and 30 months from individuals with COPD randomized to receive fluticasone (n = 25), salmeterol and fluticasone (n = 20), or placebo (n = 23). Many genes increased in COPD decreased with fluticasone, and genes decreased in COPD increased with fluticasone. Genes are ranked from left to right based on their association with the time by treatment interaction effect. The position of each *vertical bar* indicates the position of a gene whose expression in airway epithelium is associated with COPD within this ranked list (the *upper plot* includes genes increased in COPD; the *lower plot* includes genes decreased in COPD). The height of this bar represents the running gene set enrichment analysis score. (*B*) *Boxplots* illustrate the expression levels of three core enrichment genes in the bronchial airway epithelium of subjects with COPD (n = 87) compared with subjects without COPD (n = 151) and in an independent cohort of subjects randomized to receive fluticasone-containing therapies or placebo (n = 55 subjects with ⩾1 time point). The *y* axis represents the *z* score normalized residual matrix after adjusting for RNA integrity number, treatment, time, and patient effect.

chronic airway diseases (25). The increased levels of TMPRSS11D gene expression in the airway epithelium of individuals with COPD are consistent with the hypothesis that this protein plays a key role in the biologic defense against inhaled substances (25). SERPINB13 is a serine peptidase inhibitor increased in airway and lung parenchyma in association with COPD. Our finding that TMPRSS11D and SERPINB13 are increased in the airway of patients with COPD and our finding that these genes are decreased with fluticasone suggests the protease-antiprotease imbalance that is thought to play a key role in COPD pathogenesis is also reflected in airway epithelial cells, and that restoration of this balance could be useful for monitoring response to COPD therapies, such as inhaled corticosteroids. Prostaglandin-endoperoxidase synthase 2 is a proinflammatory mediator increased in the bronchial airway of individuals with COPD and is a potential target for novel anti-inflammatory therapies. CLDN8 is a member of the claudin family, which plays a key role in tight junctions and paracellular permeability (26). Our finding that CLDN8 is decreased in the airway epithelium of subjects with COPD and increased after treatment with fluticasone suggests a potentially reversible impairment in the airway epithelium's critical barrier function (27), and this finding is consistent with the previously observed down-regulation of claudins and other tight junction genes in bronchial epithelial cells from smokers with COPD (28, 29).

Our observations about the potential role of ATF4 in mediating COPD-associated gene expression differences in bronchial epithelium is intriguing given the role of ATF4 in mediating the unfolded protein response (30, 31). Endoplasmic reticulum (ER) stress from acute cigarette smoke exposure leads to an unfolded protein response, which is proposed to play a role in the development of COPD (32, 33). An increase in ER stress markers has been described in the lungs of patients with COPD (34), and administration of acrolein, an aldehyde in cigarette smoke, leads to an increase in ER stress markers and airspace enlargement in mice, suggesting that ER stress and the unfolded protein response play key roles in the development of emphysema (35). This is the first study to our knowledge to identify ATF4-driven gene expression differences in individuals with COPD. We have validated predicted targets of ATF4 in the airway COPD signature, and have further demonstrated significant enrichment of genes increased in the airway COPD signature among genes increased by ATF4. Although we identified this potential regulatory relationship in airway epithelium, further studies are necessary to examine the extent of this response in lung tissue and its importance for disease development.

The potential clinical relevance of the COPD-associated field of injury is supported by its reversal with inhaled corticosteroids in the GLUCOLD cohort (23). This aspect of the airway signature of COPD indicates that the constituent gene expression differences reflect more than differences caused by demographic or smoking-related factors, but rather an aspect of the disease process that is modifiable with therapy. Moreover, further studies should be conducted to determine whether heterogeneity in the extent to which the airway signature of COPD is reversed by therapy is associated with differences in the clinical benefit obtained by patients. Similarly, it is important to determine whether gene expression heterogeneity among patients with COPD reflects underlying biologic differences that can be used to develop markers that predict aspects of the clinical heterogeneity of COPD, such as therapeutic response or rate of lung function decline.

As with other distal lung diseases, there are several potential mechanisms that might account for the similarity between lung tissue and bronchial airway gene expression (36). The COPD-associated transcriptomic alterations may reflect specific physiologic responses to the toxins in cigarette smoke that in turn contribute to COPD pathogenesis. The relationship between the airway signature of COPD and gene expression differences associated with regional emphysema severity within an individual, and the reversal of the signature after therapy, suggest that the etiology of the COPD-associated gene expression differences is not solely caused by an individual's physiologic response to tobacco smoke.

Other potential mechanisms for the airway field of injury are related to cell-cell communication. For example, inflammatory cells recruited into the airway and lungs of smokers with COPD and the cytokines they produce may induce gene expression alterations throughout the airway epithelium. This hypothesis is consistent with our finding of specific inflammatory-related pathways enriched among the genes in our signature (Figure 1; *see* Table E3). However, *in silico* analysis of white blood cell–specific gene expression in these samples did not reveal significant proportions of inflammatory cells or differences in the proportion of inflammatory cells in smokers with and without COPD, and thus we do not believe that our signature directly reflects changing numbers of inflammatory cells within our airway brushings in individuals with COPD. Nonetheless, infiltration of the airway wall with inflammatory cells in smokers with COPD (23) could produce changes in the adjacent epithelial layer lining that airway.

Through analysis of the largest cohort of bronchial airway gene expression in COPD, we have identified a COPD-associated airway field of injury despite several potentially important limitations to our study design. Because of the nature of this lung cancer screening cohort, characterization of COPD-related phenotypes was limited, and we defined COPD as airflow obstruction on prebronchodilator spirometry. However, the similarity with previously published lung tissue gene expression datasets suggests that these COPD-associated changes in bronchial airway gene expression are reproducible and reflective of disease activity. Although spirometry remains the standard for diagnosing COPD (19), the association of airway gene expression with subphenotypes of COPD was not evaluated including quantitative imaging of airway remodeling and emphysema, gas transfer capacity, chronic bronchitis, previous respiratory illness, frequency of exacerbations, and/or quality of life metrics. Given the clinical heterogeneity among smokers with COPD, it is possible that different clinical subphenotypes of disease impact airway gene expression differently and might contribute to the heterogeneity seen in the gene expression signature. Furthermore, given that we were leveraging a bronchoscopy-based cohort for this study in which most subjects with COPD had mild to moderate disease, it is unclear if our findings generalize to smokers with later stage disease or if there are alterations specific to more severe disease. However, the enrichment of our airway gene expression signature among genes that change with regional emphysema severity in the lungs of smokers with severe COPD suggests that our gene expression signature is also relevant in more severe disease. Finally, although fluticasone-containing therapy has not been consistently linked with a clinical benefit, the decrease in the COPD airway gene expression signature after fluticasone therapy in two independent cohorts suggests that the COPD-associated airway field of injury is not a static consequence of disease but rather is dynamic.

In summary, we have shown that COPD induces a field of injury that extends from the lung parenchyma into the bronchial airway, and that some of the COPD-associated alterations in airway gene expression may be mediated by ATF4. We have also shown that a subset of these COPD-associated airway gene expression changes is reversed by fluticasone in a COPD cohort where that treatment resulted in improvement in lung function. These data suggest that gene expression profiling of the airway epithelium, which can be sampled by bronchoscopy, may serve as a surrogate biomarker of disease activity. Further studies are needed to evaluate whether this field of injury in COPD extends to epithelial cells that can be more readily sampled from the nose (37). However, our findings of an airway-wide field of injury in the bronchial airway of smokers with COPD and the identification of a reversible component of this COPD-specific gene expression signature will promote the study and development of clinically useful markers of disease activity, molecular subtypes, prognosis, and response to therapy.

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

Acknowledgment: The authors thank Adam Gower for a customized R package for generating heatmaps.

References

1. National Heart Lung and Blood Institute. NHLBI Fiscal Year 2010 Fact Book [2010; accessed 2013 Apr 10]. Available from: http://www.nhlbi.nih.gov/about/factbook-10/FactBook_2010.pdf
2. Murphy SL, Xu J, Kochanek KD. Deaths: preliminary data for 2012. National Vital Statistics Reports. Vol 60, No 4. Hyattsville, MD: National Center for Health Statistics; 2012.
3. Spira A, Beane J, Pinto-Plata V, Kadar A, Liu G, Shah V, Celli B, Brody JS. Gene expression profiling of human lung tissue from smokers with severe emphysema. *Am J Respir Cell Mol Biol* 2004;31:601–610.
4. Golpon HA, Coldren CD, Zamora MR, Cosgrove GP, Moore MD, Tuder RM, Geraci MW, Voelkel NF. Emphysema lung tissue gene expression profiling. *Am J Respir Cell Mol Biol* 2004;31:595–600.
5. Ning W, Li CJ, Kaminski N, Feghali-Bostwick CA, Alber SM, Di YP, Otterbein SL, Song R, Hayashi S, Zhou Z, *et al*. Comprehensive gene expression profiles reveal pathways related to the pathogenesis of chronic obstructive pulmonary disease. *Proc Natl Acad Sci USA* 2004;101:14895–14900.
6. Bhattacharya S, Srisuma S, DeMeo DL, Shapiro SD, Bueno R, Silverman EK, Reilly JJ, Mariani TJ. Molecular biomarkers for quantitative and discrete COPD phenotyes. *Am J Respir Cell Mol Biol* 2009;40:359–367.
7. Wang IM, Stepaniants S, Boie Y, Mortimer JR, Kennedy B, Elliott M, Hayashi S, Loy L, Coulter S, Cervino S, *et al*. Gene expression profiling in patients with chronic obstructive pulmonary disease and lung cancer. *Am J Respir Crit Care Med* 2008;177:411.
8. Spira A, Beane JE, Shah V, Steiling K, Liu G, Schembri F, Gilman S, Dumas YM, Calner P, Sebastiani P, *et al*. Airway epithelial gene expression in the diagnostic evaluation of smokers with suspect lung cancer. *Nat Med* 2007;13:361–366.
9. Gustafson AM, Soldi R, Anderlind C, Scholand MB, Qian J, Zhang X, Cooper K, Walker D, McWilliams A, Liu G, *et al*. Airway PI3K pathway activation is an early and reversible event in lung cancer development. *Sci Transl Med* 2010;2:26a25.
10. Pierrou S, Broberg P, O'Donnell RA, Pawlowski K, Virtala R, Lindqvist E, Richter A, Wilson SJ, Angco G, Moller S, *et al*. Expression of genes involved in oxidative stress responses in airway epithelial cells of smokers with chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2007;175:577–586.
11. Tilley AE, Harvey BG, Heguy A, Hackett NR, Wang R, O'Connor TP, Crystal RG. Down-regulation of the Notch pathway in human airway epithelium in association with smoking and chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2009;179:457–466.
12. Steiling K, Lenburg ME, Spira A. Airway gene expression in chronic obstructive pulmonary disease. *Proc Am Thorac Soc* 2009;6:697–700.
13. Tammemagi MC, Lam SC, McWilliams AM, Sin DD. Incremental value of pulmonary function and sputum DNA image cytometry in lung cancer risk prediction. *Cancer Prev Res (Phila)* 2011;4:552–561.
14. Crapo RO, Morris AH, Gardner RM. Reference spirometric values using techniques and equipment that meet ATS recommendations. *Am Rev Respir Dis* 1981;123:659–664.
15. Dennis G, Sherman BT, Hosack DA, Yang J, Gao W, Lane HC, Lempicki RA. DAVID: database for annotation, visualization, and integrated discovery. *Genome Biol* 2003;4:3.
16. Chang JT, Nevins JR. GATHER: a systems approach to interpreting genomic signatures. *Bioinformatics* 2006;22:2926–2933.
17. Matys V, Kel-Margoulis OV, Fricke E, Liebich I, Land S, Barre-Dirrie A, Reuter I, Chekmenev D, Krull M, Hornischer K, *et al*. TRANSFAC and its module TRANSCompel: transcriptional gene regulation in eukaryotes. *Nucleic Acids Res* 2006;34:D108–D110.

18. Hertz GZ, Stormo GD. Identifying DNA and protein patterns with statistically significant alignments of multiple sequences. *Bioinformatics* 1999;15:563–577.
19. Global Initiative for Chronic Obstructive Lung Disease. Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease [2011; accessed 2013 Apr 10]. Available from: http://www.goldcopd.org/uploads/users/files/GOLDReport_April112011.pdf
20. Beane J, Sebastiani P, Liu G, Brody JS, Lenburg ME, Spira A. Reversible and permanent effects of tobacco smoke exposure on airway epithelial gene expression. *Genome Biol* 2007;8:R201.
21. Spira A, Beane J, Shah V, Liu G, Schembri F, Yang X, Palma J, Brody JS. Effects of cigarette smoke on the human airway epithelial cell transcriptome. *Proc Natl Acad Sci USA* 2004;101:10143–10148.
22. Campbell JD, McDonough JE, Zeskind JE, Hackett TL, Pechkovsky DV, Bradsma CA, Suzuki M, Gosselink JV, Liu G, Alekseyev YO, *et al.* A gene expression signature of emphysema-related lung destruction and its reversal by the tripeptide GHK. *Genome Med* 2012;4:67.
23. Lapperre TS, Snoeck-Stroband JB, Gosman MM, Jansen DF, van Schadewijk A, Thiadens HA, Vonk JM, Boezen HM, Ten Hacken NH, Sont JK, *et al.*; Groningen Leiden Universities Corticosteroids in Obstructive Lung Disease Study Group. Effect of fluticasone with and without salmeterol on pulmonary outcomes in chronic obstructive pulmonary disease: a randomized trial. *Ann Intern Med* 2009;151:517–527.
24. Woodruff PG, Boushey HA, Dolganov GM, Barker CS, Yang YH, Donnelly S, Ellwanger A, Sidu SS, Dao-Pick TP, Pantoja C, *et al.* Genome-wide profiling identifies epithelial cell genes associated with asthma and with treatment response to corticosteroids. *Proc Natl Acad Sci USA* 2007;104:15858–15863.
25. Takahashi M, Sano T, Yamaoka K, Kamimura T, Umemoto N, Nishitani H, Yasuoka S. Localization of human airway trypsin-like protease in the airway: an immunohistochemical study. *Histochem Cell Biol* 2001;115:181–187.
26. Lal-Nag M, Morin PJ. The claudins. *Genome Biol* 2009;10:235.
27. Heijink IH, Brandenburg SM, Noordhoek JA, Postma DS, Siebos DJ, van Oosterhout AJM. Characterization of cell adhesion in airway epithelial cell types using electric cell-substrate impedance sensing. *Eur Respir J* 2010;35:894–903.
28. Shaykhiev R, Otaki F, Bonsu P, Dang DT, Teater M, Strulovici-Barel Y, Salit J, Harvey BG, Crystal RG. Cigarette smoking reprograms apical junctional complex molecular architecture in the human airway epithelium in vivo. *Cell Mol Life Sci* 2011;68:877–892.
29. Soini Y. Claudins in lung diseases. *Respir Res* 2011;12:70.
30. Wek RC, Cavener DR. Translational control and the unfolded protein response. *Antioxid Redox Signal* 2007;9:2357–2371.
31. Rzymski T, Milani M, Singleton DC, Harris AL. Role of ATF4 in regulation of autophagy and resistance to drugs and hypoxia. *Cell Cycle* 2009;8:3838–3847.
32. Kelsen SG, Duan X, Ji R, Perez O, Liu C, Merali S. Cigarette smoke induces an unfolded protein response in the human lung: a proteomic approach. *Am J Respir Cell Mol Biol* 2008;38:541–550.
33. Geraghty P, Wallace A, D'Armiento JM. Induction of the unfolded protein response by cigarette smoke is primarily an activating transcription factor 4-C/EBP homologous protein mediated process. *Int J Chron Obstruct Pulmon Dis* 2011;6:309–319.
34. Malhotra D, Thimmulappa R, Vij N, Navas-Acien A, Sussan T, Merali S, Zhang L, Kelsen SG, Myers A, Wise R, *et al.* Heightened endoplasmic reticulum stress in the lungs of patients with chronic obstructive pulmonary disease: the role of Nrf2-regulated protesasomal activity. *Am J Respir Crit Care Med* 2009;180:1196–1207.
35. Kitaguchi Y, Taraseviciene-Stewart L, Hanaoka M, Natarajan R, Kraskauskas D, Voelkel NF. Acrolein induces endoplasmic reticulum stress and causes airspace enlargement. *PLoS ONE* 2012;7:e38038.
36. Steiling K, Ryan J, Brody JS, Spira A. The field of tissue injury in the lung and airway. *Cancer Prev Res (Phila)* 2008;1:396–403.
37. Zhang X, Sebastiani P, Liu G, Schembri F, Zhang X, Dumas YM, Langer EM, Alekseyev Y, O'Connor GT, Brooks DR, *et al.* Similarities and differences between smoking-related gene expression in the nasal and bronchial epithelium. *Physiol Genomics* 2010;41:1–8.