Transcriptomic Studies of the Airway Field of Injury Associated with Smoking-Related Lung Disease

Adam C. Gower¹, Katrina Steiling^{2,3}, John F. Brothers II¹, Marc E. Lenburg^{1,2}, and Avrum Spira^{1,2}

¹Bioinformatics Program, Boston University, ²Section of Computational Biomedicine, Department of Medicine, and ³Pulmonary Center, Boston University School of Medicine, Boston, Massachusetts

The “field of injury” hypothesis proposes that exposure to an inhaled insult such as cigarette smoke elicits a common molecular response throughout the respiratory tract. This response can therefore be quantified in any airway tissue, including readily accessible epithelial cells in the bronchus, nose, and mouth. High-throughput technologies, such as whole-genome gene expression microarrays, can be employed to catalog the physiological consequences of such exposures in the airway epithelium. Pulmonary diseases such as chronic obstructive pulmonary disease, lung cancer, and asthma are also thought to be associated with a field of injury, and in patients with these diseases, airway epithelial cells can be a useful surrogate for diseased tissue that is often difficult to obtain. Global measurement of mRNA and microRNA expression in these cells can provide useful information about the molecular pathogenesis of such diseases and may be useful for diagnosis and for predicting prognosis and response to therapy. In this review, our aim is to summarize the history and state of the art of such “transcriptomic” studies in the human airway epithelium, especially in smoking and smoking-related lung diseases, and to highlight future directions for this field.

Keywords: epithelium; lung neoplasms; chronic obstructive pulmonary disease; asthma; tobacco

Lung diseases associated with cigarette smoking, such as lung cancer and chronic obstructive pulmonary disease (COPD), are leading causes of morbidity and mortality in the United States and throughout the world. Much remains to be learned about the molecular pathogenesis of these diseases, partly because of the relative difficulty of obtaining diseased tissue from distal sites within the lung and the risks associated with invasive procedures in individuals with impaired lung function. If disease processes that occur in lung tissue are reflected at the molecular level in the airway epithelium, it may be possible to use this more readily accessible population of cells as a surrogate to study these diseases.

The “field of injury” hypothesis proposes that the entire respiratory system of an individual reacts to environmental insults such as tobacco smoke in a similar manner and that common aspects of this response can be measured in airway epithelium. Genome-wide gene expression profiling can be used to obtain a detailed molecular portrait of this physiological response. Similarly, pulmonary diseases may also create a field of injury throughout the respiratory system, and disease-specific changes in the expression of mRNAs and microRNAs in the airway epithelium may therefore provide insight into the molecular mechanisms of these diseases and yield biomarkers for diagnosis, prognosis, and response to therapy. In this review,

our aim is to describe the state of the art of studies (summarized in Table 1 and Figure 1) that use genome-wide technologies to measure gene expression changes in the human airway epithelium that are associated with smoking and smoking-related lung diseases.

TRANSCRIPTOMIC STUDIES OF SMOKING-RELATED GENE EXPRESSION CHANGES IN THE AIRWAY

The first transcriptomic study that used microarrays to measure changes in airway gene expression (1) involved bronchial brushings obtained from the third-generation bronchi of volunteer smokers and nonsmokers. Although bronchial epithelium was profiled with an early “whole genome” array platform, the authors focused their analysis on a curated set of 44 antioxidant-related genes. Several classes of these genes, especially those involved in glutathione metabolism and redox balance, were both up-regulated and highly variable within smokers. Our group also profiled bronchial brushings of the large airway, expanding on the list of genes with smoking-associated differential expression and confirming that there is significant variability within the up-regulation of antioxidant-related genes in smokers (2). This variability may offer a biological explanation as to why only a subset of current and former smokers appears to be susceptible to smoking-induced lung diseases. These reports were followed by a similar study (3) that found that many of the same genes were also differentially expressed in the small airway epithelium (brushings of 10–12th-generation bronchi) of smokers, suggesting that the response to cigarette smoke is similar throughout the bronchial epithelium.

Smoking-induced lung diseases affect similar percentages of both current and former smokers, suggesting that genomic changes that occur in the airway of smokers persist after smoking cessation and may be associated with the development of these diseases. The question of which changes in airway gene expression persist after smoking cessation was first addressed in a 2004 study (2) and more fully addressed in a subsequent study (4). In both studies, our group profiled gene expression in brushings of the right mainstem bronchus of healthy current, former, and never-smokers. These studies found that the expression levels of most genes that are differentially expressed in active smokers rapidly revert to the levels seen in nonsmokers within a few months of smoking cessation. However, there is a subset of genes whose expression levels remained irreversibly altered for more than 10 years after smoking cessation. This set of genes includes a cluster of metallothioneins located at 16q13 that is irreversibly repressed in smokers, suggesting that smoking may have long-term effects on the antioxidant response of the airway epithelium. These genes were also found to be coordinately differentially expressed in the bronchial epithelium in an independent set of current and former smokers, suggesting that irreversible smoking-associated changes in gene expression could be used as a biomarker of prior exposure to tobacco smoke (4). The finding of irreversibly altered changes in bronchial airway gene expression among former smokers has been confirmed by two other groups using

(Received in original form November 1, 2010; accepted in final form December 30, 2010)

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

Proc Am Thorac Soc Vol 8, pp 173–179, 2011

DOI: 10.1513/pats.201011-066MS

Internet address: www.atsjournals.org

TABLE 1. TRANSCRIPTOMIC STUDIES OF AIRWAY EPITHELIUM

GEO ID(s)	Type(s)	Platform(s)	Phenotypes	Reference(s)
GSE5056	Bronchial brushings, large airway	HuGeneFL	Nonsmokers and smokers	1, 43
GSE994	Bronchial brushings, large airway	HG-U133A	Never-, former, and current smokers	2, 6, 44, 45
—	Nasal brushings (inferior turbinate)	HG-U133 Plus 2.0	Children without asthma; children with stable or acute asthma	32
—	Oral mucosa	Custom spotted cDNA array	Nonsmokers and smokers	46
—	Microdissected, matched bronchial epithelium, alveolar tissue, and NSCLC tumor cells	HG-U133A	Nonsmokers with or without cancer; smokers with or without cancer	15
GSE3320	Bronchial brushings, small airway	HG-U133A	Nonsmokers and smokers	3, 43
GSE5057	Bronchial brushings, large airway	HG-U133A	Nonsmokers and smokers	43
GSE5058	Bronchial brushings, small airway	HG-U133 Plus 2.0	Nonsmokers and smokers; early COPD, COPD	29, 43
GSE5059	Bronchial brushings, large airway	HG-U133 Plus 2.0	Nonsmokers and smokers	43
GSE4498	Bronchial brushings, small airway	HG-U133 Plus 2.0	Nonsmokers and smokers	3, 29
GSE4115, GSE12815	Bronchial brushings, histologically normal large airway epithelium	HG-U133A	Current and former smokers, with and without lung cancer	16, 17, 19
—	Bronchial brushings	HG-U133A, HG-U133B	Healthy smokers and nonsmokers; smokers with COPD	28
GSE5473	Bronchial brushings, peripheral airway	SAGE	Never-, former, and current smokers	5
GSE7895, GSE12815	Bronchial brushings, large airway	HG-U133A	Never-, former, and current smokers	4, 19
GSE4302	Bronchial brushings	HG-U133 Plus 2.0	Nonsmokers with asthma (untreated, placebo, or fluticasone treated); healthy nonsmokers; smokers with mild to moderate COPD	30, 31
—	Bronchial brushings, large airway	HG-U133 Plus 2.0	Former and current smokers	6
GSE8987	Buccal epithelium	HG-U133A	Healthy never- and current smokers	8
—	Nasal epithelium	HG-U133A 2.0		
GSE8545	Bronchial brushings, small airway	HG-U133 Plus 2.0	Normal smokers and nonsmokers; smokers with COPD	44
GSE10006	Bronchial brushings, large and small airways	HG-U133 Plus 2.0	Nonsmokers and smokers	47
GSE14633	Bronchial brushings, large airway	Human Exon 1.0ST	Never- and current smokers	7
GSE14634		NCode miRBase V2		
GSE7832	Bronchial brushings, small airway	HG-U133 Plus 2.0	Normal smokers and nonsmokers; smokers with COPD	29
GSE4635	Bronchial brushings, large airway	HG-U133A	Never- and current smokers	48
GSE10135	Bronchial brushings, large airway	HuGeneFL, HG-U133A, HG-U133 Plus 2.0	Nonsmokers and smokers	45
GSE12428, GSE12472	Microdissected sections: histologically normal bronchial epithelium and SCC tumor cells	Whole Human Genome G4112A	Normal bronchial epithelium and SCC from current and former smokers	14, 49
GSE11952	Bronchial brushings, small airway	HG-U133 Plus 2.0	Nonsmokers and smokers	11
GSE13933	Tracheal brushings	HG-U133 Plus 2.0	Nonsmokers and smokers	12
—	Bronchial brushings, small airway			
GSE13931	Matched pairs: small airway bronchial brushings and bronchoalveolar macrophages	HG-U133 Plus 2.0	Nonsmokers and smokers	33
GSE17913	Oral mucosa	HG-U133 Plus 2.0	Never- and current smokers	10
GSE16149	Buccal epithelium	HG-U133 Plus 2.0	Nonsmokers and smokers	50
GSE16008	Bronchial (large airway) and nasal brushings	Human Exon 1.0 ST	Never- and current smokers	9
GSE12815	Bronchial brushings, large airway	HG-U133A	COPD, no COPD	19
GSE12815	Bronchial brushings, large airway	Human Exon 1.0 ST	Individuals with dysplasia, untreated or before/after <i>myo</i> -inositol; healthy smokers without dysplasia	19

Definition of abbreviations: COPD = chronic obstructive pulmonary disease; GEO ID = Gene Expression Omnibus identification number; NSCLC = non-small cell lung cancer; SAGE = Serial Analysis of Gene Expression; SCC = squamous cell carcinoma.

Data sets are listed in chronological order according to their initial date of publication. Gene Expression Omnibus (GEO) series IDs are listed where available.

both microarray and SAGE (Serial Analysis of Gene Expression) platforms (5, 6).

The discovery of the role of microRNA in regulating key biological processes and the availability of whole-genome microRNA arrays has prompted investigation into the potential role of microRNAs in regulating the transcriptional response to cigarette smoke. We profiled mRNA and microRNA in parallel from bronchial brushings of the large airway epithelium of current and never-smokers (7). The majority of differentially expressed microRNAs are down-regulated in smokers, whereas the majority of mRNAs that are differen-

tially expressed in smokers are up-regulated, which is intriguing given the role of microRNAs largely as negative regulators of mRNA levels. The expression of one microRNA in particular, miR-218, was inversely correlated with the expression of several of its computationally predicted mRNA targets, including MAFG (v-maf musculoaponeurotic fibrosarcoma oncogene homolog G, avian), a transcription factor whose binding site is overrepresented among genes that are differentially expressed in response to smoking. We further found that, *in vitro*, overexpression of miR-218 in cultured bronchial epithelial cells attenuated the up-regulation of MAFG in

Figure 1. Timeline of selected studies involving transcriptomic profiling of airway epithelial cells. Studies are shaded according to the main experimental condition(s) of interest. COPD = chronic obstructive pulmonary disease; PI3K = phosphoinositide-3-kinase; Th2 = helper T-cell type 2.

response to treatment with cigarette smoke condensate. These results suggest that miR-218 may be an important regulator of the transcriptional response of the airway epithelium to tobacco smoke.

Although the aforementioned studies have shown that changes in the expression of mRNAs and microRNAs occur throughout the bronchial epithelium of smokers, sampling of the bronchial airway via bronchoscopy is an invasive procedure that is not a feasible method for screening large populations at risk for smoking-related lung disease. Several groups have performed studies to determine whether the field of injury extends to upper, extrathoracic airway epithelia that could be collected less invasively than bronchial epithelium. Our group identified concordant smoking-associated changes in gene expression in brushings of the large airway and nasal epithelium in both cross-sectional (8) and paired (9) studies. Another study (10) demonstrated that a number of genes previously reported to be differentially expressed in the bronchus of smokers, especially xenobiotic detoxification genes, are also differentially expressed in the buccal mucosa of smokers obtained via biopsy, which is consistent with our findings of a modest smoking-related buccal gene-expression signature (8). These results suggest that there is a common response to tobacco smoke throughout the airway, and that changes in gene expression at any of these sites may be representative of changes in the bronchial epithelium in response to cigarette smoke. This point is illustrated in a side-by-side comparison of the differential expression of smoking-associated genes in epithelial cells of the small and large airway, trachea, and nose (Figure 2). Genes that were identified as differentially expressed in the bronchial airway of smokers in one data set (4) behave similarly in the

small airway, trachea, and nose of smokers profiled in other studies (9, 11, 12).

DOES THE COMMON AIRWAY FIELD OF INJURY EXTEND TO TOBACCO-RELATED DISEASE?

The consistency of the physiological response to tobacco smoke exposure across airway epithelia, as measured by similar patterns of tobacco smoke exposure-associated changes in gene expression, may reflect the direct response of cells with a common physiology to compounds in tobacco smoke that are present throughout the respiratory tree of smokers. An important question, therefore, is whether signatures of processes that contribute to the molecular pathogenesis of pulmonary disease in general, and tobacco-related pulmonary diseases in particular, can also be detected globally in tissues throughout the respiratory system. The ability to detect the pathogenesis of disease in readily accessible tissue that is removed from the site of disease would not only facilitate studies focused on better understanding the mechanisms of disease formation, but might also be clinically useful for disease detection and monitoring. For pulmonary diseases in general, the potential to detect disease-specific changes in apparently nondiseased tissue might reflect aspects of disease that affect the entire respiratory system. For smoking-related pulmonary diseases in particular, it is additionally possible that in those smokers who develop a specific disease, cells throughout the respiratory system respond to tobacco smoke exposure in a manner that is consistent with disease formation. This scenario would suggest that not only might this prodisease response be detected throughout the respiratory system of smokers with clinical

Figure 2. Transcriptomic profiling reveals a common transcriptional response to cigarette smoke in epithelia throughout the airway. Four sets of gene expression profiles of epithelial cells from the small (10th–12th generation) or large (third generation) bronchus, trachea, or nose of healthy smokers were retrieved from the Gene Expression Omnibus (GEO). The Robust Multichip Average (RMA) (50) and an Entrez Gene-specific probeset mapping (51) were used to estimate the levels of 74 genes whose expression has previously been reported to be differentially expressed in response to cigarette smoke (4). Genes that are up- or down-regulated in smokers (red and blue sidebars, respectively) were clustered within separate groups using hierarchical clustering. In the same manner, samples from never-smokers (yellow) and current smokers (blue) were clustered separately. Samples from former smokers (green) were ordered (and shaded) according to time since smoking cessation. Gene (row) order was established using the small airway set (GSE11952) and maintained in the remaining heatmaps.

disease manifestation, but also that it might be possible to detect this prodisease response to tobacco exposure before disease manifestation.

TRANSCRIPTOMIC STUDIES OF DIFFERENTIAL GENE EXPRESSION IN THE AIRWAY OF SMOKERS WITH LUNG CANCER

The significant heterogeneity in smoking-associated differential gene expression observed in the normal airway of smokers has led several authors to postulate that certain smokers may have a qualitatively different response to tobacco smoke that predisposes them to the development of lung diseases such as lung cancer (2, 13) or COPD (1). Further evidence of a relationship between the response to smoking and lung tumorigenesis has

been provided by two genome-wide gene expression profiling studies of lung tumors and normal bronchial epithelium (14, 15). These studies identified changes in gene expression associated with current smoking status in the normal bronchial epithelium that were also present in lung tumors versus normal bronchial epithelium (irrespective of smoking status). These changes included the up-regulation of genes involved in the oxidant and xenobiotic response and the down-regulation of genes encoding tumor suppressors.

These observations suggest that there is a connection between airway gene expression and lung cancer and that gene expression in the normal airway epithelium might serve as a diagnostic or screening tool for this disease. Our group has investigated this possibility and has developed a gene expression profile that can serve as an early diagnostic biomarker for

lung cancer, using bronchial brushings of cytologically normal large airway epithelium from individuals undergoing bronchoscopy for suspect lung cancer (16). This biomarker improves diagnostic accuracy when combined with traditional diagnostic bronchoscopy (16), and is independent of clinical factors as a diagnostic risk factor (17). The concept that gene expression in normal airway epithelium can be used as a biomarker of lung cancer was also used by Blomquist and colleagues to identify 14 genes (including antioxidant and DNA repair genes) whose expression was altered in brushings of normal bronchial epithelium of patients with lung cancer versus those without lung cancer (18).

The finding that changes in gene expression in the normal bronchial epithelium are associated with the presence of lung cancer raises the possibility that preneoplastic disease may be associated with similar changes. To test this hypothesis, our group used whole-genome gene expression microarrays to profile cytologically normal bronchial epithelium in smokers with and without dysplastic airway lesions (19). A gene expression signature of phosphoinositide-3-kinase signaling pathway activation was coordinately differentially expressed in smokers with these dysplastic lesions, as well as in the cytologically normal large airway epithelium of smokers with lung cancer from our previous study (16). Furthermore, pharmacological inhibition of this pathway led to the regression of dysplastic lesions and a concomitant reversal of the perturbation of the phosphoinositide-3-kinase gene expression signature in the subset of smokers with regression. This suggests that smoking-induced changes in airway gene expression may reflect early, potentially reversible deregulation of specific molecular pathways related to lung carcinogenesis, and that monitoring of these changes might assist in the selection of specific chemopreventive agents for smokers at highest risk for lung cancer.

TRANSCRIPTOMIC STUDIES OF THE AIRWAY FIELD OF INJURY IN COPD

COPD is clinically defined as obstruction to airflow that is not completely reversible with bronchodilator treatment. However, COPD is a heterogeneous disease with a variable natural history, and individuals with COPD exhibit varying degrees of small airway inflammation, bronchitis, and emphysema. Early transcriptomic studies of gene expression in COPD relied on lung tissue specimens from individuals undergoing surgery (20–25). These studies were generally hampered by small sample sizes, limited phenotyping of COPD, and confounding variables such as smoking status and concomitant lung cancer, and the overlap between sets of COPD-related differentially expressed genes reported by each study was small. However, meta-analysis of these studies by Gene Set Enrichment Analysis (GSEA) (26) has shown that these studies demonstrate reproducible COPD-related gene expression changes in COPD-affected lung tissue (27).

If a COPD-associated field of injury extends to the bronchial airway, it may be possible to measure COPD pathogenesis in airway epithelium. As current studies of COPD have been limited by the availability of surgical samples, this approach would facilitate expanded studies of the mechanisms of COPD pathogenesis in primary human samples. Two studies have demonstrated the feasibility of this approach. By profiling bronchial epithelial brushings, Pierrou and colleagues found that genes that were differentially expressed in the airway epithelium of smokers with and without COPD were enriched in genes from curated oxidative stress response pathways (28). Similarly, Tilley and colleagues identified that a subset of genes from the Notch signaling pathway were down-regulated in the

small airway epithelium of smokers with COPD compared with smokers without COPD (29). As we described in an abstract presented at the 2010 Aspen Lung Conference, genome-wide gene expression profiling of bronchial brushings from large numbers of matched subjects with and without COPD has identified concordant differential expression of genes whose expression is known to be perturbed in lung tissue of smokers with COPD. This suggests that airway gene expression in COPD reflects molecular processes occurring in more distal diseased lung tissue, and that the bronchial airway epithelium may therefore serve as a surrogate that can be sampled less invasively to study COPD.

TRANSCRIPTOMIC STUDIES OF THE AIRWAY FIELD OF INJURY IN ASTHMA

Asthma is defined clinically as reversible obstruction of the airway. It is caused by exposure to inhaled allergens, which leads to helper T-cell type 2 (Th2)–mediated inflammation of the airways and results in inflammatory cell infiltration, airway hyperresponsiveness, and tissue remodeling. Because asthma is associated with an airway-wide inhaled exposure, it might also induce an airway-wide field of injury reflecting the presence and severity of this disease. If so, an airway epithelium transcriptomic signature of asthma might be a useful tool for developing personalized therapies, discovering new drugs, and creating prognostic biomarkers.

The first genome-wide study of airway epithelial gene expression in asthma used microarrays to profile bronchial brushings from nonsmokers with and without asthma (30). This study identified changes in gene expression associated with the presence of asthma, as well as those associated with treatment with the inhaled corticosteroid fluticasone. A follow-up study identified a set of genes induced by IL-13 and IL-4, whose expression was increased in a subset of individuals with asthma (31). This “Th2-high” pattern of gene expression was reversed after 4 to 8 weeks of treatment with fluticasone, accompanied by improvements in FEV₁. These observations suggest that asthma is associated with a field of injury in the airway, and that airway gene expression profiling of individuals with asthma has the potential to identify individuals with distinct subtypes of asthma that may preferentially respond to particular therapies.

As described previously regarding smoking and smoking-associated diseases, the field of injury related to asthma may also extend to the epithelia of the upper airway. Asthma-related biomarkers measured in these epithelia would allow for less invasive specimen collection, which is especially important for studies involving children who are affected by this disease. Guajardo and colleagues established a nasal asthma-associated field of injury by profiling nasal mucosal brushings from children with stable or acute asthma as well as from healthy control subjects (32). Several functional categories of genes were differentially expressed in both asthmatic groups, including genes related to immune function and signal transduction. Distinct gene expression profiles were also associated with either stable or acute asthma, suggesting that the nasal field of injury reflects disease heterogeneity and may be useful in identifying subgroups of individuals with asthma who are refractory to particular therapies or who might be at higher risk for more frequent exacerbations of this disease.

FUTURE DIRECTIONS

As described in this review, significant progress has been made in establishing that an airway-wide field of injury occurs in response to tobacco smoke and in establishing that a similar

field of injury exists in response to smoking-related diseases such as lung cancer and COPD, or inhaled exposure-related diseases such as asthma. Through the use of genome-wide gene expression profiling technologies, it has been possible to create detailed molecular portraits of these fields of injury that are useful not only for understanding mechanisms of disease pathogenesis but also for developing biomarkers capable of monitoring disease formation and progression in readily accessible airway tissues. The rapid emergence of next-generation sequencing of RNA promises to provide us with an even more comprehensive view of the transcriptional changes associated with this airway field of injury.

The airway epithelium is the first barrier to environmental exposures in the lung, and as such, patterns of gene expression in these cells are likely to be the most informative about the response of the lung to inhaled insults. However, other cell types have potential advantages as sites for studies of airway disease. Several studies (e.g., References 33 and 34) have measured genome-wide gene expression in alveolar macrophages collected by bronchoalveolar lavage, samples enriched in inflammatory cells that might participate directly in some disease processes within the lung. Sputum is also a potential biosample for gene expression studies, although this is technically challenging because of mRNA degradation. Peripheral blood is a routinely collected clinical specimen that is also a source of inflammatory cells that are exposed to inhaled insults because of the high degree of lung vascularization. Studies have measured the effect of chronic smoking on gene expression in peripheral blood mononuclear cells (35, 36) and the use of peripheral blood mononuclear cell gene expression as a biomarker of lung cancer has also been reported (37). Ongoing work in these and other cell types should be especially helpful for understanding the systemic effects of tobacco smoke exposure and tobacco-related disease and for creating additional biomarkers of pulmonary disease that may complement those measured in airway epithelium.

Although tobacco smoke is the most well-characterized chronic insult to the airway epithelium, other environmental exposures contribute to disease, including ozone, diesel exhaust, unvented coal smoke, and asbestos. Although these agents have been studied using cell culture and animal models, there are limited transcriptomic studies that have measured the effect of these exposures on gene expression in human airway epithelium. Understanding the similarities and differences between the fields of injury induced by these different exposures may help to better define the repertoire of physiological responses to inhaled exposures and allow the identification of aspects of these responses that contribute to disease formation.

Several other pulmonary diseases, including interstitial lung disease, acute respiratory distress syndrome, and sarcoidosis, may also be associated with a field of injury in the lung and respiratory tract. Genome-wide gene expression has been measured in these diseases in lung tissue or in surrogates such as blood and bronchoalveolar lavage fluid (38–42). Airway-based transcriptomic studies of these diseases could complement these studies and be useful for elucidating molecular pathologies, identifying clinically relevant subclasses of disease, and creating diagnostic and prognostic biomarkers. They may also serve to more generally define how disease processes that may be less directly connected with inhaled exposures affect the physiology of the entire respiratory system and thereby give rise to an airway field of injury.

Author Disclosure: A.C.G. received grant support from the NIH. K.S. does not have a financial relationship with a commercial entity that has an interest in the subject of this manuscript. J.F.B. received grant support from the NIH. M.E.L. was a consultant for and owns stock in Allegro Diagnostics Inc., and he received grant

support from the NIH. A.S. was a consultant for and owns stock in Allegro Diagnostics Inc., and he received grant support from the NIH.

References

- Hackett NR, Heguy A, Harvey BG, O'Connor TP, Luettich K, Flieder DB, Kaplan R, Crystal RG. Variability of antioxidant-related gene expression in the airway epithelium of cigarette smokers. *Am J Respir Cell Mol Biol* 2003;29:331–343.
- 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.
- Harvey BG, Heguy A, Leopold PL, Carolan BJ, Ferris B, Crystal RG. Modification of gene expression of the small airway epithelium in response to cigarette smoking. *J Mol Med* 2007;85:39–53.
- 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.
- Chari R, Lonergan KM, Ng RT, MacAulay C, Lam WL, Lam S. Effect of active smoking on the human bronchial epithelium transcriptome. *BMC Genomics* 2007;8:297.
- Zhang L, Lee JJ, Tang H, Fan YH, Xiao L, Ren H, Kurie J, Morice RC, Hong WK, Mao L. Impact of smoking cessation on global gene expression in the bronchial epithelium of chronic smokers. *Cancer Prev Res* 2008;1:112–118.
- Schembri F, Sridhar S, Perdomo C, Gustafson AM, Zhang X, Ergun A, Lu J, Liu G, Zhang X, Bowers J, et al. MicroRNAs as modulators of smoking-induced gene expression changes in human airway epithelium. *Proc Natl Acad Sci USA* 2009;106:2319–2324.
- Sridhar S, Schembri F, Zeskind J, Shah V, Gustafson AM, Steiling K, Liu G, Dumas YM, Zhang X, Brody JS, et al. Smoking-induced gene expression changes in the bronchial airway are reflected in nasal and buccal epithelium. *BMC Genomics* 2008;9:259.
- 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 nasal and bronchial epithelium. *Physiol Genomics* 2010;41:1–8.
- Boyle JO, Gumus ZH, Kacker A, Choksi VL, Bocker JM, Zhou XK, Yantiss RK, Hughes DB, Du B, Judson BL, et al. Effects of cigarette smoke on the human oral mucosal transcriptome. *Cancer Prev Res* 2010;3:266–278.
- Hubner RH, Schwartz JD, De BP, Ferris B, Omberg L, Mezey JG, Hackett NR, Crystal RG. Coordinate control of expression of Nrf2-modulated genes in the human small airway epithelium is highly responsive to cigarette smoking. *Mol Med* 2009;15:203–219.
- Turetz ML, O'Connor TP, Tilley AE, Strulovici-Barel Y, Salit J, Dang D, Teater M, Mezey J, Clark AG, Crystal RG. Trachea epithelium as a “canary” for cigarette smoking-induced biologic phenotype of the small airway epithelium. *Clin Transl Sci* 2009;2:260–272.
- Willey JC, Coy EL, Frampton MW, Torres A, Apostolakis MJ, Hoehn G, Schuermann WH, Thilly WG, Olson DE, Hammersley JR, et al. Quantitative RT-PCR measurement of cytochromes P450 1A1, 1B1, and 2B7, microsomal epoxide hydrolase, and NADPH oxidoreductase expression in lung cells of smokers and nonsmokers. *Am J Respir Cell Mol Biol* 1997;17:114–124.
- Boelens MC, van den Berg A, Fehrmann RS, Geerlings M, de Jong WK, te Meerman GJ, Sietsma H, Timens W, Postma DS, Groen HJ. Current smoking-specific gene expression signature in normal bronchial epithelium is enhanced in squamous cell lung cancer. *J Pathol* 2009;218:182–191.
- Woenckhaus M, Klein-Hitpass L, Grepmeier U, Merk J, Pfeifer M, Wild P, Bettstetter M, Wuensch P, Blaszyk H, Hartmann A, et al. Smoking and cancer-related gene expression in bronchial epithelium and non-small-cell lung cancers. *J Pathol* 2006;210:192–204.
- 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.
- Beane J, Sebastiani P, Whitfield TH, Steiling K, Dumas YM, Lenburg ME, Spira A. A prediction model for lung cancer diagnosis that integrates genomic and clinical features. *Cancer Prev Res* 2008;1:56–64.
- Blomquist T, Crawford EL, Mullins D, Yoon Y, Hernandez DA, Khuder S, Ruppel PL, Peters E, Oldfield DJ, Austermiller B, et al. Pattern of antioxidant and DNA repair gene expression in normal airway epithelium associated with lung cancer diagnosis. *Cancer Res* 2009;69:8629–8635.

19. 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;2:26ra25.
20. Bhattacharya S, Srisuma S, Demeo DL, Shapiro SD, Bueno R, Silverman EK, Reilly JJ, Mariani TJ. Molecular biomarkers for quantitative and discrete COPD phenotypes. *Am J Respir Cell Mol Biol* 2009;40:359–367.
21. Francis SM, Larsen JE, Pavey SJ, Bowman RV, Hayward NK, Fong KM, Yang IA. Expression profiling identifies genes involved in emphysema severity. *Respir Res* 2009;10:81.
22. Golpon HA, Coldren CD, Zamora MR, Cosgrove GP, Moore MD, Tudor RM, Geraci MW, Voelkel NF. Emphysema lung tissue gene expression profiling. *Am J Respir Cell Mol Biol* 2004;31:595–600.
23. 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.
24. 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.
25. Wang TM, Stephanians 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;177:402–411.
26. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, Mesirov JP. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci USA* 2005;102:15545–15550.
27. Zeskind JE, Lenburg ME, Spira A. Translating the COPD transcriptome: insights into pathogenesis and tools for clinical management. *Proc Am Thorac Soc* 2008;5:834–841.
28. 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.
29. 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.
30. Woodruff PG, Boushey HA, Dolganov GM, Barker CS, Yang YH, Donnelly S, Ellwanger A, Sidhu 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.
31. Woodruff PG, Modrek B, Choy DF, Jia G, Abbas AR, Ellwanger A, Koth LL, Arron JR, Fahy JV. T-helper type 2-driven inflammation defines major subphenotypes of asthma. *Am J Respir Crit Care Med* 2009;180:388–395.
32. Guajardo JR, Schleifer KW, Daines MO, Ruddy RM, Aronow BJ, Wills-Karp M, Hershey GK. Altered gene expression profiles in nasal respiratory epithelium reflect stable versus acute childhood asthma. *J allergy clin immunol* 2005;115:243–251.
33. Carolan BJ, Harvey BG, Hackett NR, O'Connor TP, Cassano PA, Crystal RG. Disparate oxidant gene expression of airway epithelium compared to alveolar macrophages in smokers. *Respir Res* 2009;10:111.
34. Shaykhiev R, Krause A, Salit J, Strulovici-Barel Y, Harvey BG, O'Connor TP, Crystal RG. Smoking-dependent reprogramming of alveolar macrophage polarization: implication for pathogenesis of chronic obstructive pulmonary disease. *J Immunol* 2009;183:2867–2883.
35. Lampe JW, Stephanians SB, Mao M, Radich JP, Dai H, Linsley PS, Friend SH, Potter JD. Signatures of environmental exposures using peripheral leukocyte gene expression: tobacco smoke. *Cancer Epidemiol Biomarkers Prev* 2004;13:445–453.
36. van Leeuwen DM, van Agen E, Gottschalk RW, Vlietinck R, Gielen M, van Herwijnen MH, Maas LM, Kleinjans JC, van Delft JH. Cigarette smoke-induced differential gene expression in blood cells from monozygotic twin pairs. *Carcinogenesis* 2007;28:691–697.
37. Showe MK, Vachani A, Kossenkoff AV, Yousef M, Nichols C, Nikonova EV, Chang C, Kucharczuk J, Tran B, Wakeam E, *et al.* Gene expression profiles in peripheral blood mononuclear cells can distinguish patients with non-small cell lung cancer from patients with nonmalignant lung disease. *Cancer Res* 2009;69:9202–9210.
38. Konishi K, Gibson KF, Lindell KO, Richards TJ, Zhang Y, Dhir R, Bisceglia M, Gilbert S, Yousem SA, Song JW, *et al.* Gene expression profiles of acute exacerbations of idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med* 2009;180:167–175.
39. Schischmanoff PO, Naccache JM, Carrere A, Richardson S, Kambouchner M, Raphael M, Valeyre D, Fagard R. Progressive pulmonary sarcoidosis is associated with over-expression of TYK2 and p21Waf1/Cip1. *Sarcoidosis Vasc Diffuse Lung Dis* 2006;23:101–107.
40. Selman M, Pardo A, Barrera L, Estrada A, Watson SR, Wilson K, Aziz N, Kaminski N, Zlotnik A. Gene expression profiles distinguish idiopathic pulmonary fibrosis from hypersensitivity pneumonitis. *Am J Respir Crit Care Med* 2006;173:188–198.
41. Ukena SN, Koenecke C, Geffers R, Fuehner T, Welte T, Ganser A, Buer J, Franzke A. T helper type 2 differentiation is associated with induction of antibacterial defense mechanisms in blood lymphocytes of patients with sarcoidosis. *Immunol Invest* 2009;38:49–66.
42. Wang Z, Beach D, Su L, Zhai R, Christiani DCA. Genome-wide expression analysis in blood identifies pre-elafin as a biomarker in ARDS. *Am J Respir Cell Mol Biol* 2008;38:724–732.
43. Carolan BJ, Heguy A, Harvey BG, Leopold PL, Ferris B, Crystal RG. Up-regulation of expression of the ubiquitin carboxyl-terminal hydrolase L1 gene in human airway epithelium of cigarette smokers. *Cancer Res* 2006;66:10729–10740.
44. Ammous Z, Hackett NR, Butler MW, Raman T, Dolgalev I, O'Connor TP, Harvey BG, Crystal RG. Variability in small airway epithelial gene expression among normal smokers. *Chest* 2008;133:1344–1353.
45. Vanni H, Kazeros A, Wang R, Harvey BG, Ferris B, De BP, Carolan BJ, Hubner RH, O'Connor TP, Crystal RG. Cigarette smoking induces overexpression of a fat-depleting gene AZGP1 in the human. *Chest* 2009;135:1197–1208.
46. Smith RV, Schlecht NF, Childs G, Prystowsky MB, Belbin TJ. Pilot study of mucosal genetic differences in early smokers and non-smokers. *Laryngoscope* 2006;116:1375–1379.
47. Carolan BJ, Harvey BG, De BP, Vanni H, Crystal RG. Decreased expression of intelectin 1 in the human airway epithelium of smokers compared to nonsmokers. *J Immunol* 2008;181:5760–5767.
48. Steiling K, Kadar AY, Bergerat A, Flanigan J, Sridhar S, Shah V, Ahmad QR, Brody JS, Lenburg ME, Steffen M, *et al.* Comparison of proteomic and transcriptomic profiles in the bronchial airway epithelium of current and never smokers. *PLoS One* 2009;4:e5043.
49. Boelens MC, Gustafson AM, Postma DS, Kok K, van der Vries V, van der Vlies V, Spira A, Lenburg ME, Geerlings M, Sietsma H, *et al.* A chronic obstructive pulmonary disease related signature in squamous cell lung cancer. *Lung Cancer*. Epub 2010 December 9, 2010.
50. Kupfer DM, White VL, Jenkins MC, Burian D. Examining smoking-induced differential gene expression changes in buccal mucosa. *BMC Med Genomics* 2010;3:24.
51. Irizarry RA, Hobbs B, Collin F, Beazer-Barclay YD, Antonellis KJ, Scherf U, Speed TP. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* 2003;4:249–264.
52. Dai M, Wang P, Boyd AD, Kostov G, Athey B, Jones EG, Bunney WE, Myers RM, Speed TP, Akil H, *et al.* Evolving gene/transcript definitions significantly alter the interpretation of GeneChip data. *Nucleic Acids Res* 2005;33:e175.