chemokines. The subjects consisted of 13 life-long nonsmokers, 13 smokers without COPD, and 14 patients with COPD, who required lung resection for small peripheral tumors. Only in terminal bronchiolar epithelium was catalase mRNA significantly decreased in patients with COPD compared with nonsmokers, although the difference did not reach statistical significance in smokers with and without COPD. In contrast, interleukin-8, macrophage inflammatory protein-1 α , and monocyte chemoattractant protein-1 mRNA levels were higher in smokers with COPD than those in nonsmokers or smokers without COPD. Furthermore, cDNA array revealed the overexpression of CC chemokine receptor 2 and the down-regulation of several heat shock protein (HSP) 70 family chaperones in terminal bronchiolar epithelium from patients with COPD. Considering that catalase is responsible for the majority of hydrogen peroxide consumption in the lung, a decrease in the expression of catalase along with HSP70 family chaperones specifically in terminal bronchiolar epithelium may lead to the oxidative stress and the up-regulation of inflammatory chemokines that occur in the same population of cells lining terminal bronchioles and may be relevant to the development of COPD. This study also demonstrates the use of LCM combined with RT-PCR and cDNA microarrays to investigate functional profiles of individual structural and inflammatory cells in human lungs.

Conflict of Interest Statement: None of the authors has a financial relationship with a commercial entity that has an interest in the subject of this manuscript.

(Received in original form March 15, 2006; accepted in final form March 21, 2006)

Correspondence and requests for reprints should be addressed to Tomoko Betsuyaku, First Department of Medicine, Hokkaido University School of Medicine, Kita-15, Nishi-7, Kita-ku, Sapporo 060-8638, Japan. E-mail: bytomoko@med.hokudai.ac.jp

Proc Am Thorac Soc Vol 3, 2006
DOI: 10.1513/pats.200603-032MS
Internet address: www.atsjournals.org

Microarray Data-Based Prioritization of Chronic Obstructive Pulmonary Disease Susceptibility Genes

Soumyaroop Bhattacharya, Sorachai Srisuma, Dawn L. DeMeo, John J. Reilly, Raphael Bueno, Edwin K. Silverman, and Thomas J. Mariani

Division of Pulmonary Medicine, Brigham and Women's Hospital; and The Lung Biology Center, Harvard Medical School, Boston, Massachusetts

One limitation of genome-wide linkage screens aiming to identify disease-susceptibility genes is the resolution at which they define individual candidates. Fine mapping of linked loci can often be laborious, protracted, and poorly directed. We reasoned that an integrative genomics approach could help to expedite candidate gene identification. We tested this hypothesis by prioritizing genes within three linkage regions (chromosomes 2q33-36, 8pter-22, and 12p13-12) previously identified in the Boston Early-Onset COPD cohort (1), using two independent microarray data sets. Genes within previously defined linkage regions were identified using Golden Path (University of California Santa Cruz, <http://genome.ucsc.edu>), and associated probes were assigned using NetAffx (Affymetrix, Santa Clara, CA). Probe sets were filtered by sequence verification using The Lung Transcriptome (<http://lungtranscriptome.bwh.harvard.edu>). The

first microarray data set studied consists of whole lung tissue samples derived from 20 patients who had undergone lung volume reduction surgery with severe emphysema and from 14 control subjects with mild to moderate obstruction (2). The second data set consists of whole lung tissue samples (uninvolved margin) derived from 31 patients undergoing surgery for solitary pulmonary nodules, with varying degrees of airflow obstruction (FEV₁% predicted: mean, 72; range, 10–133), and has not been previously described. Signal intensities were derived using both MAS5 and RMA algorithms. Unsupervised clustering with the nonparametric bootstrap was applied to check for undesirable and unanticipated structure or associations among the samples. Pearson and Spearman correlations were used to test for significant associations between gene expression and continuous phenotypic variables (e.g., TLC, FEV₁, FVC, FEV₁/FVC, DL_{CO}, FEF_{25–75}, smoking history). All analysis methods were exhaustively repeated for each gene/probe set and each data set. For each gene/probe set, analysis results were summarized where data implicated an association between gene expression and the disease variable. Finally, genes were rank-prioritized based upon their frequency of significant association. A number of genes within each locus were repeatedly associated with multiple phenotypic variables in each data set. As has been previously noted, there was limited consistency between data sets, which in this case may be due to distinctions in patient populations and/or technical limitations. However, this approach consistently identified a small number of genes, including the recently implicated candidate susceptibility gene SERPINE2, as targets for further study. We believe this provides a rational approach that is applicable to many diseases and model systems for which data may already exist. However, further investigation is necessary to assert causation between prioritized candidates and the phenotype being investigated.

Acknowledgment: The authors acknowledge Craig P. Hersh for providing additional linkage information.

Conflict of Interest Statement: S.B. does not have a financial relationship with a commercial entity that has an interest in the subject of this manuscript. S.S. does not have a financial relationship with a commercial entity that has an interest in the subject of this manuscript. D.L.D. does not have a financial relationship with a commercial entity that has an interest in the subject of this manuscript. J.J.R. does not have a financial relationship with a commercial entity that has an interest in the subject of this manuscript. R.B. does not have a financial relationship with a commercial entity that has an interest in the subject of this manuscript. E.K.S. received grant support, consulting fees, and honoraria from GlaxoSmithKline for studies of COPD genetics. He also received a speaker fee from Wyeth for a talk on COPD genetics. He also received honoraria from Bayer. T.J.M. does not have a financial relationship with a commercial entity that has an interest in the subject of this manuscript.

References

1. Silverman EK, Palmer LJ, Mosley JD, Barth M, Senter JM, Brown A, Drazen JM, Kwiatkowski DJ, Chapman HA, Campbell EJ, *et al*. Genomewide linkage analysis of quantitative spirometric phenotypes in severe early-onset chronic obstructive pulmonary disease. *Am J Hum Genet* 2002;70:1229–1239.
2. 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.

(Received in original form March 16, 2006; accepted in final form April 4, 2006)

These studies were supported by NIH GRANTS HL72303 (J.J.R.) and HL71885 (T.J.M.)

Correspondence and requests for reprints should be addressed to Thomas J. Mariani, Ph.D., Pulmonary Medicine, Thorn 8, Brigham and Women's Hospital, 75 Francis Street, Boston, MA 02115. E-mail: tmariani@partners.org

Proc Am Thorac Soc Vol 3, 2006
DOI: 10.1513/pats.200603-033MS
Internet address: www.atsjournals.org