The Advent of High-Throughput Sequencing Studies of Chronic Obstructive Pulmonary Disease

A genetic basis for chronic obstructive pulmonary disease (COPD) has long been suspected on the basis of marked interindividual variability in lung function among cigarette smokers (1). Severe alpha-1 antitrypsin deficiency is a known, strong genetic risk factor for COPD (2, 3), but it explains only a small proportion of cases. Findings from studies of siblings or relatives of subjects with severe COPD unrelated to severe alpha-1 antitrypsin deficiency suggest that most such cases are heritable and likely caused by multiple genes interacting with the detrimental effects of cigarette smoking (4, 5).

Several approaches have been used to identify susceptibility genes for COPD unrelated to severe alpha-1 antitrypsin deficiency, most recently including candidate-gene and genome-wide association studies (GWAS) focusing on common (minor allele frequency $\geqslant 0.05$) single-nucleotide polymorphisms. Candidate-gene studies identified potential susceptibility loci for COPD or COPD-related phenotypes in inflammatory, oxidant–antioxidant, and protease–antiprotease gene pathways, a few of which have been subsequently replicated using genome-wide approaches (e.g., matrix metalloproteinase 12, *MMP12)* (6–8). Over the last 6 years, GWAS have identified multiple susceptibility genes for COPD or its intermediate phenotypes (e.g., lung function), confirming the multifactorial etiology of most cases (8–10).

In spite of advances in gene discovery, the effect sizes of common single-nucleotide polymorphisms identified by GWAS are small and account for only a modest proportion of COPD heritability. One potential explanation for the "missing heritability" of COPD is rare variants with strong effects, which could be theoretically identified using novel sequencing approaches.

In a report published in this issue of the *Journal* (pp. 1353–1363), Qiao and colleagues attempted to discover rare susceptibility variants for an extreme COPD phenotype (severe disease of early onset), which may be influenced by rare coding variants with large effects. The investigators first conducted whole-exome sequencing to identify rare variants in 347 members of 49 extended pedigrees in the Boston Early-Onset COPD (EOCOPD) Study (11). They then conducted a segregation analysis, which identified 69 genes with rare deleterious variants that segregated in at least two pedigrees in the EOCOPD Study. Rare functional variants in these 69 genes were next tested for association with severe COPD in 399 participants (204 cases and 195 control subjects ["resistant smokers"]) in the COPDGene Study, in which four genes (*DNAH8*, *ALCAM*, *RARS*, and *GBF1*) were shown to have a non–statistically significant trend for a higher frequency of rare deleterious variants in cases than in control subjects. Variants for one of these four genes (activated leukocyte cell adhesion molecule, *ALCAM*) were only found in three cases in the COPDGene Study, and a second gene (*RARS* [arginyl-tRNA synthetase, cytoplasmic]) was shown to be differentially expressed in lung tissue from 111 severe COPD cases and 40 control subjects from an unrelated study cohort.

In addition to attempting to identify rare susceptibility variants in novel genes, Qiao and colleagues tried to discover rare variants in known candidate genes for COPD or lung function (11). Supportive but nondefinitive evidence for such rare variants was shown for some candidate genes, including *FAM13A* and *CCDC38*.

The study by Qiao and colleagues is the first report of a whole-exome sequencing analysis of COPD, and thus a novel contribution to the literature (11). As acknowledged by the authors, their results are inconclusive due to lack of statistical significance or functional validation and must thus be cautiously interpreted pending further work. Moreover, the study findings cannot be generalized to members of other racial or ethnic groups such as African Americans, in whom rare variants are more frequently detected than in subjects of European descent (12). Nonetheless, the report suggests that rare causal variants for severe COPD are likely to be found in multiple genes and further illustrates the genetic heterogeneity of COPD, even within an extreme phenotype such as severe disease of early onset.

The study by Qiao and colleagues provides important lessons for future next-generation sequencing studies of COPD (11). First, this report and recent sequencing studies of complex cardiovascular diseases (13, 14) suggests that large sample size (e.g., thousands of subjects) is likely needed to detect rare susceptibility variants for COPD because of genetic heterogeneity as well as potential gene-by-gene or gene-by-smoking interactions. Second, replication of positive findings in the discovery cohort(s) in adequately powered replication studies, followed by functional validation of replicated results, are critical to exclude false positives due to multiple testing. Third, whole-genome sequencing would be preferred to whole-exome sequencing, which does not fully capture variation (within the 5′ promoter, 3′ untranslated region, and intronic regions of genes) that may affect the development or severity of COPD through regulation of transcription and expression. Fourth, new methods are needed to analyze potential rare susceptibility variants for COPD and other complex lung diseases. Current analytical methods have been developed for gene-level burden testing, which account for varying genetic effects of different rare variants and appropriate allele frequency thresholds for variants from different genes (15). Despite the power of these analytical approaches, the selection of arbitrary allele frequencies can exclude variants with deleterious effects.

The common goal of human genetics and precision medicine research is to identify and validate genetic variants predictive of disease risk or treatment response. As we continue to identify genetic variation with next-generation sequencing, variants relevant to human disease should be incorporated into a larger set of molecular and genetic biomarker panels, with the ultimate goal of stratifying individuals for personalized interventions for the prevention and early detection of disease and development of targeted therapeutics. To develop genetic

Am J Respir Crit Care Med Vol 193, Iss 12, pp 1323–1336, Jun 15, 2016
Internet address: www.atsjournals.org

profiles for precision medicine in COPD, future studies must be able to analyze the expanding volume of data from next-generation whole-genome DNA and RNA sequencing, along with other "omics" data, to comprehensively characterize ethnically diverse cohorts with an adequately broad range of age, sex, cigarette smoke exposure, and disease severity. ■

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

Victor E. Ortega, M.D.
Center for Genomics and Personalized Medicine
Wake Forest School of Medicine
Winston-Salem, North Carolina

Juan C. Celedón, M.D., Dr.P.H.
Children's Hospital of Pittsburgh of UPMC
University of Pittsburgh
Pittsburgh, Pennsylvania

References

1. Burrows B, Knudson RJ, Cline MG, Lebowitz MD. Quantitative relationships between cigarette smoking and ventilatory function. *Am Rev Respir Dis* 1977;115:195–205.
2. Eriksson S. Pulmonary emphysema and alpha1-antitrypsin deficiency. *Acta Med Scand* 1964;175:197–205.
3. Yoshida A, Lieberman J, Gaidulis L, Ewing C. Molecular abnormality of human alpha1-antitrypsin variant (Pi-ZZ) associated with plasma activity deficiency. *Proc Natl Acad Sci USA* 1976;73:1324–1328.
4. McCloskey SC, Patel BD, Hinchliffe SJ, Reid ED, Wareham NJ, Lomas DA. Siblings of patients with severe chronic obstructive pulmonary disease have a significant risk of airflow obstruction. *Am J Respir Crit Care Med* 2001;164:1419–1424.
5. 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.
6. Celedón JC, Lange C, Raby BA, Litonjua AA, Palmer LJ, DeMeo DL, Reilly JJ, Kwiatkowski DJ, Chapman HA, Laird N, *et al*. The transforming growth factor-beta1 (TGFB1) gene is associated with chronic obstructive pulmonary disease (COPD). *Hum Mol Genet* 2004; 13:1649–1656.
7. Hunninghake GM, Cho MH, Tesfaigzi Y, Soto-Quiros ME, Avila L, Lasky-Su J, Stidley C, Melén E, Söderhäll C, Hallberg J, *et al*. MMP12, lung function, and COPD in high-risk populations. *N Engl J Med* 2009; 361:2599–2608.
8. Cho MH, McDonald ML, Zhou X, Mattheisen M, Castaldi PJ, Hersh CP, Demeo DL, Sylvia JS, Ziniti J, Laird NM, *et al*.; NETT Genetics, ICGN, ECLIPSE and COPDGene Investigators. Risk loci for chronic obstructive pulmonary disease: a genome-wide association study and meta-analysis. *Lancet Respir Med* 2014;2:214–225.
9. Hancock DB, Eijgelsheim M, Wilk JB, Gharib SA, Loehr LR, Marciante KD, Franceschini N, van Durme YM, Chen TH, Barr RG, *et al*. Meta-analyses of genome-wide association studies identify multiple loci associated with pulmonary function. *Nat Genet* 2010;42:45–52.
10. Cho MH, Boutaoui N, Klanderman BJ, Sylvia JS, Ziniti JP, Hersh CP, DeMeo DL, Hunninghake GM, Litonjua AA, Sparrow D, *et al*. Variants in FAM13A are associated with chronic obstructive pulmonary disease. *Nat Genet* 2010;42:200–202.
11. Qiao D, Lange C, Beaty TH, Crapo JD, Barnes KC, Bamshad M, Hersh CP, Morrow J, Pinto-Plata VM, Marchetti N, *et al*.; Lung GO; NHLBI Exome Sequencing Project; COPDGene Investigators. Exome sequencing analysis in severe, early-onset chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2016;193: 1353–1363.
12. Abecasis GR, Auton A, Brooks LD, DePristo MA, Durbin RM, Handsaker RE, Kang HM, Marth GT, McVean GA; 1000 Genomes Project Consortium. An integrated map of genetic variation from 1,092 human genomes. *Nature* 2012;491:56–65.
13. Crosby J, Peloso GM, Auer PL, Crosslin DR, Stitziel NO, Lange LA, Lu Y, Tang ZZ, Zhang H, Hindy G, *et al*.; TG and HDL Working Group of the Exome Sequencing Project, National Heart, Lung, and Blood Institute. Loss-of-function mutations in APOC3, triglycerides, and coronary disease. *N Engl J Med* 2014;371:22–31.
14. Jørgensen AB, Frikke-Schmidt R, Nordestgaard BG, Tybjærg-Hansen A. Loss-of-function mutations in APOC3 and risk of ischemic vascular disease. *N Engl J Med* 2014;371:32–41.
15. Lin DY, Tang ZZ. A general framework for detecting disease associations with rare variants in sequencing studies. *Am J Hum Genet* 2011;89:354–367.

Copyright © 2016 by the American Thoracic Society

A Cough Is a Cough, Is It Not?
Neurophenotypes Define Patients with Chronic Cough

From a physiological perspective, coughing represents an important pulmonary defense mechanism, inasmuch as it serves to clear the airways of inhaled or aspirated substances and excessive secretions, which can compromise airway patency and/or promote pulmonary damage and infection (1, 2). Acute (or transient) coughing, typically associated with an upper respiratory tract infection, is defined as cough lasting less than 3 weeks, and although it can be troubling to the patient at the time, it generally resolves with minimal therapeutic intervention (3). Chronic cough, in contrast, is defined as cough lasting greater than 8 weeks, and in a significant number of patients, the duration of coughing can be many months, and even years (3, 4). Perhaps surprising to many is that chronic cough is very prevalent in the community, affecting 9–33% of the population, including children (5). Indeed, patients presenting with persistent cough represent up to 38% of a chest specialist outpatient practice, making cough the most common complaint for which patients seek medical attention (5).

Chronic cough is common in patients with both pulmonary and nonpulmonary conditions, including asthma and cough-variant asthma, eosinophilic bronchitis, gastroesophageal reflux disease, rhinosinusitis, chronic obstructive pulmonary disease, and bronchiectasis (3, 5). However, up to 40% of coughers attending specialized cough clinics are diagnosed with chronic "idiopathic" (unexplained or refractory) cough, as an underlying cause for their cough cannot be identified (5, 6). Indeed, in a recent survey of people with chronic cough, 47% of respondents indicated the cause of their cough was undiagnosed, and only 7% reported that the