NEWS AND COMMENTARY

Cystic Fibrosis

• •

Using genetic association to identify modifiers of disease variability in cystic fibrosis

Garry R Cutting

. . . .

One of the greatest challenges in modern medicine is to understand why patients with the same disease can have different outcomes. Identifying the causes of disease variation can be of considerable prognostic value. Furthermore, factors that modify a disease may be more amenable to therapeutic manipulation than the underlying cause of the disease. Thus, the recent report that severity of life limiting pulmonary disease in cystic fibrosis (CF) patients is associated with variation in the transforming growth factor $\beta 1$ (TGF$\beta 1$) gene has generated considerable attention.[1] This interest is warranted on several grounds. First, genetic variation in TGF$\beta 1$ was estimated to account for 15–20% of the variability in lung disease (as measured in their study) suggesting that it could be an important prognostic factor. Second, although the CF transmembrane conductance regulator (CFTR) was identified as the molecular defect in CF 16 years ago, the disease has proven to be challenging to treat at the molecular level. Finally, the TGFβ signaling pathway is a critical determinant of lung homeostasis, and therapies are already available to manipulate this pathway.

The study design used by Drumm *et al* is similar to case–control association studies that test whether selected genetic variants alter susceptibility to complex traits. Variants in 10 candidate modifier genes previously shown to be associated with CF disease outcome were tested for association with patients grouped according to 'mild' or 'severe' lung disease. Only variants -509C/T and codon 10 T/C in the TGF$\beta 1$ gene showed significant differences in genotype distribution among the two groups of patients. Three features distinguish this study from preceding work; the criteria used to define disease severity, the number of subjects studied, and replication of association.

Classifying patients according to severity of a chronic disease manifestation can be challenging, especially if progression of the illness is nonlinear, as is the case for CF lung disease. To address this issue, Drumm *et al* used an objective measure of pulmonary mechanics, the volume of air in the first second of a forced exhalation (forced exhalation volume in one second or FEV$_1$). This cross-sectional measure is predictive of CF survival and is used for monitoring progress and response to treatment of CF patients.[2,3] Patients were classified as 'mild' or 'severe' if their age-adjusted FEV$_1$ at the time of enrollment fell in the upper 25% or the lower 25%, respectively, of FEV$_1$ values of patients in the US CF Registry in 1999. To control for allelic effects at the CFTR locus, only FEV$_1$ values from patients homozygous for the common CFTR mutation ΔF508 in the CF Registry were utilized thereby matching the CFTR genotype of the enrolled patients. As cross-sectional measures of an on-going process may lead to misclassification, the authors also utilized FEV$_1$ measures over the preceding five years and a regression model[4] to predict pulmonary function at 20 years of age for each subject. The aforementioned measure was found to be consistent with the initial classification for 97.3% of the enrolled patients. The use of cross-sectional and longitudinal measures to define disease severity is a major strength of the paper and sets a worthy precedent for genetic association studies of chronic disorders.

The other distinguishing features are not unique to this study but are highly desirable in association studies. Adequate power to detect associations between alleles of moderate effect and phenotype require substantial numbers of well-characterized patients. Multicenter collaboration enabled enrollment of over 800 patients for the Drumm study. While the authors did not provide an analysis of the power of their study to detect association of single or multiple modifiers, the number of individuals enrolled far exceeds most candidate gene modifier studies of single gene disorders in humans. Second, validation of a genetic association via replication is an important step to excluding false positive associations.[5] Constructive replication studies performed by a different set of investigators, using a separate set of patients with different genotyping methods and a different statistical model provide the most objective form of replication.[6] The replication study by Drumm *et al* had a number of these attributes. An independent set of patients were recruited with a broad range of CFTR genotypes and without bias regarding disease severity classification. The authors used both cross-sectional and longitudinal measures of FEV$_1$ but divided patients into 'severe' and 'mild' groups using a different method than the initial study.

What are the next steps in evaluating genetic modifiers of CF lung disease and in particular, the role of genetic variation in TGF $\beta 1$? Although the study by Drumm *et al* suggests that TGF$\beta 1$ alleles appear to account for a reasonable fraction of variance in CF lung function, the relative contribution of genetic and non-genetic factors to this trait is unknown. Hence, it is not possible to determine whether TGF$\beta 1$ variation accounts for all or part of the genetic contribution to FEV$_1$ variation. To address this issue, association studies searching for modifiers of CF lung disease should assess the contribution of novel candidates to trait variance relative to TGF$\beta 1$ alleles. Second, as noted in the

editorial accompanying the Drumm paper,[7] patients recruited according to lung disease severity also had a significant difference in body mass index (BMI). Thus, TGFβ1 alleles could be associated with FEV$_1$, BMI or a combination of the two. Intriguingly, analysis of CF twins and CF siblings revealed that genetic factors contributed to a composite measure of FEV$_1$ and BMI, but not to FEV$_1$ alone.[8] Finally, stratification of patient groups still lingers as a possible explanation for association observed by Drumm et al.[9] Stratification has been observed in Caucasian Americans that can be traced to different European ancestry.[10] In the case of CF, the frequency of the CF alleles differs considerably depending upon the geographic and ethnic origin of European patients. The frequency of TGFβ1 alleles might also be stratified in European Americans. The replication study by Drumm et al demonstrated association between TGFβ1 genotypes and age-adjusted FEV$_1$ above or below 68% predicted that was independent of CFTR genotype and independent of substantial variation in environment (patients were enrolled at several centers in North America). These findings suggest that adequately powered studies of independent collections of CF patients should detect association between TGFβ1 alleles and CF lung disease severity as defined in the replication study of Drumm et al. Replication of the findings of Drumm et al in other CF patient populations or in family based transmission disequilibrium studies could reduce or eliminate the concern of false positive association due to population stratification. Such studies could also be structured to determine if TGFβ1 variants affect pulmonary status, body mass index or a combination of both. Although the path from association to causation may be tortuous, affirmation of the findings of Drumm, Knowles and the CF Gene Modifier Group would provide a new tool in the fight against CF ■

Dr GR Cutting is at Johns Hopkins Medical Institutions, BRB 559, 733 N. Broadway, Baltimore, MD 21287, USA. E-mail: gcutting@jhmi.edu

References

1 Drumm ML, Konstan MW, Schluchter MD et al: Genetic modifiers of lung disease in cystic fibrosis. N Engl J Med 2005; **353**: 1443–1453.
2 Kerem E, Reisman J, Corey M, Canny GJ, Levison H: Prediction of mortality in patients with cystic fibrosis. N Engl J Med 1992; **326**: 1187–1191.
3 Ramsey BW, Boat TF, Accurso FJ et al: Outcome measures for clinical trials in cystic fibrosis. Summary of Cystic Fibrosis Foundation Consensus Conference. J Pediatr 1994; **124**: 177–192.
4 Schluchter MD, Konstan MW, Davis PB: Jointly modelling the relationship between survival and pulmonary function in cystic fibrosis patients. Stat Med 2002; **21**: 1271–1287.
5 Ioannidis JP, Ntzani EE, Trikalinos TA, Contopoulos-Ioannidis DG: Replication validity of genetic association studies. Nat Genet 2001; **29**: 306–309.
6 Page GP, George V, Go RC, Page PZ, Allison DB: Are we there yet? Deciding when one has demonstrated specific genetic causation in complex diseases and quantitative traits. Am J Hum Genet 2003; **73**: 711–719.
7 Haston CK, Hudson TJ: Finding genetic modifiers of cystic fibrosis. N Engl J Med 2005; **353**: 1509–1511.
8 Mekus F, Ballmann M, Bronsveld I, Bijman J, Veeze H, Tummler B: Categories of deltaF508 homozygous cystic fibrosis twin and sibling pairs with distinct phenotypic characteristics. Twin Res 2000; **3**: 277–293.
9 Freedman ML, Reich D, Penney KL et al: Assessing the impact of population stratification on genetic association studies. Nat Genet 2004; **36**: 388–393.
10 Campbell CD, Ogburn EL, Lunetta KL et al: Demonstrating stratification in a European American population. Nat Genet 2005; **37**: 868–872.

Association Studies

• •

A genome-wide association approach to mapping the genetic determinants of the transcriptome in human populations

Eric E Schadt

Human genetic approaches to elucidate common human diseases include segregation studies to assess heritability of disease, family-based linkage studies to identify disease susceptibility loci, and genetic association studies to identify genes for disease. Common human diseases typically involve a number of genetic and environmental factors, and human genetic approaches in this context have not met with huge success. However, with the completion of the sequencing of the human and other genomes, the completion of the first phase of the HapMap Project,[1] and with the development of a number of high-throughput functional genomic technologies able to provide unprecedented looks at molecular processes underlying disease, human genetics is poised to have an impact on common diseases like never before. One exciting variation to the classic human genetics approaches that has recently emerged seeks to integrate molecular profiling data (eg, gene expression) with genotypic and clinical trait data to elucidate the network of molecular interactions that underlie complex traits.[2–4] Identifying variations in DNA that lead to variations in tran-