Genome-Wide Association Studies in Idiopathic Pulmonary Fibrosis: Bridging the Gap between Sequence and Consequence

Genome-wide association studies (GWASs) have been developed with the aim of improving our understanding of disease biology and discovering novel therapeutic targets through the identification of sequence variants associated with the disease or trait of interest across the genome of affected individuals (1). Since 2005, GWASs have identified several thousands of loci associated with hundreds of complex diseases (i.e., those determined by variations across multiple genes, often interacting with environmental factors, and each contributing an effect of varying yet relatively modest magnitude). However, contrary to early expectations that GWASs would identify functional (i.e., protein-disrupting) variants, ~90% of GWAS loci linked to disease risk lie in noncoding regions of the genome. The functions of many such disease-associated risk variants have remained elusive, although it is generally presumed that these variants play *cis*- or *trans*-regulatory roles. Rapid, high-throughput approaches for defining the function of such regulatory variants are clearly needed to accelerate the translation of these genetic discoveries to disease-relevant biological understanding.

Idiopathic pulmonary fibrosis (IPF) is a chronic, progressive, and almost invariably fatal interstitial lung disease (ILD) of unknown origin that occurs primarily in older adults (2). Although the mechanisms of fibrosis in IPF remain incompletely understood, the disease is believed to result from aberrant repair of the alveolar epithelium after repetitive microinjuries, with smoking, viral infection, environmental pollutants, and chronic microaspiration of gastric content representing plausible putative triggers of the fibrotic response (3). Excessive unopposed extracellular matrix synthesis by myofibroblasts—cells that express features of both fibroblasts and smooth muscle cells—leads to progressive scarring of the lung, parenchymal distortion, and irreversible loss of function (4). The role of genetic factors in both familial and sporadic cases of IPF is increasingly appreciated (5). Specifically, familial studies have identified associations with genes related to telomere biology and surfactant production, whereas GWASs of sporadic cases, including high-resolution resequencing of implicated loci (6), have reported associations with loci containing genes related to lung defense, telomere maintenance, and cell–cell adhesion (7). However, these genetic abnormalities have been estimated to account for only ~30% of the genetic risk of IPF, and the molecular mechanisms through which they promote disease development are largely unknown (8).

In this issue of the *Journal*, Allen and colleagues (pp. 564–574) report findings from a collaborative effort to perform the largest GWAS of IPF susceptibility to date (9). The study population included all patients and control subjects of European ancestry who had been recruited to any previously reported IPF GWAS (10–12) (i.e., 2,668 patients with IPF and 8,591 control subjects). The authors conducted a meta-analysis of the results of these GWASs (the discovery cohort). Two independent case/control datasets were included as a replication cohort (1,467 patients with IPF and 11,874 control subjects). The study confirmed associations at 11 previously reported loci, and conditional analyses confirmed that risk at the 11p15 locus is driven by the *MUC5B* promoter variant. Three novel association signals near *KIF15* (in 3p21.31), *MAD1L1* (in 7p22.3), and *DEPTOR* (DEP domain-containing mTOR-interacting protein) (in 8q24.12) were also identified and remained significant after adjustment for multiplicity, and were replicated in all of the discovery and replication datasets. Using three expression quantitative trait loci datasets, the authors determined that the intronic variant near *KIF15* is associated with reduced gene expression in the brain, the intronic *MAD1L1* variant is associated with reduced gene expression in the heart, and the intronic variant near *DEPTOR* is associated with reduced gene expression in the lung, colon, and skin. *DEPTOR* encodes a protein that interacts with mTOR and inhibits its kinase activity, and the authors suggest that this variant could contribute to IPF pathogenesis by reducing the inhibition of mTOR signaling, thus favoring TGF-β_1–induced collagen synthesis and fibrogenesis (13). Finally, to estimate the cumulative contributions of the identified DNA sequence variants to IPF susceptibility, Allen and colleagues generated polygenic risk scores, which are individual-level metrics of genetic risk derived from individual signals across the genome combined with mathematical modeling. After exclusion of the 14 IPF risk variants (as well as variants within 1 Mb), which explain a relatively large proportion of disease risk, the polygenic risk scores were significantly associated with increased susceptibility to IPF but explained only ~2% of the remaining risk. Although this indicates that additional common genetic variants contribute to IPF risk, this analysis also suggests that the cumulative impact of these variants is modest. Together, these data shed new light on the genetic architecture of IPF risk. However, much more work integrating genetic epidemiological and bioinformatic methods, as well as *in vitro* and *in vivo* experiments, will be required to unravel the biological significance of the identified associations in disease pathogenesis.

This study represents a considerable accomplishment for the field, and also highlights areas where further study is required. Genotyping for the discovery cohort was based on array platforms and imputation, which can reliably measure most low-frequency (1–5% minor allele frequency [MAF]) and common (>5% MAF) variants, but cannot address the role of low-frequency single-nucleotide variants (i.e., <1% MAF), copy number variations, or other structural variants. Notably, this study was conducted in individuals of European ancestry, so the generalizability of these findings to diverse populations is unclear. For instance, 60–70% of non-Hispanic white individuals carry at least one copy of *MUC5B* rs35705950 T (14), the strongest and most well-replicated single IPF risk variant, but this variant is very rare among Koreans with IPF (15). There is a clear need to develop a more integrated understanding of genetic risk for IPF across groups of diverse ancestry. Further study is also required to better elucidate how genetic factors contribute to disease risk across the spectrum of ILD phenotypes. GWASs to date have focused

Supported by grants from the Department of Cardiac, Thoracic, Vascular Sciences and Public Health, University of Padova, Padova, Italy (BIRD163522 to P.S.), the NIH (R01HL145372 and K08130595 to J.A.K.), the Department of Defense (W81XWH1910415 to J.A.K.), and the Doris Duke Charitable Foundation (J.A.K.).

Originally Published in Press as DOI: 10.1164/rccm.201911-2286ED on December 20, 2019

heavily on IPF; however, targeted studies analyzing rheumatoid arthritis–associated ILD and chronic hypersensitivity pneumonitis suggest that the prevalence of both *MUC5B* promoter polymorphism (16, 17) and rare variants in telomere-related genes is similar to that observed in IPF (18, 19). The most powerful insights that could come from future genetic studies in ILD would be links between genetic risk factors and specific ILD features or phenotypes. Such studies will require large and diverse cohorts, high-resolution phenotyping, and high-quality longitudinal data. This study by Allen and colleagues gives us reason to be optimistic that the large-scale collaborative efforts required for such investigations may be possible in the near future. ■

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

Paolo Spagnolo, M.D., Ph.D.
Department of Cardiac, Thoracic, Vascular Sciences and Public Health
University of Padova
Padova, Italy

Jonathan A. Kropski, M.D.
Department of Medicine
and
Department of Cell and Developmental Biology
Vanderbilt University Medical Center
Nashville, Tennessee
and
Department of Veterans Affairs Medical Center
Nashville, Tennessee

ORCID ID: 0000-0002-1096-0596 (P.S.).

References

1. Manolio TA. Bringing genome-wide association findings into clinical use. *Nat Rev Genet* 2013;14:549–558.
2. Lederer DJ, Martinez FJ. Idiopathic pulmonary fibrosis. *N Engl J Med* 2018;378:1811–1823.
3. Kropski JA, Blackwell TS. Progress in understanding and treating idiopathic pulmonary fibrosis. *Annu Rev Med* 2019;70:211–224.
4. Selman M, Pardo A. Revealing the pathogenic and aging-related mechanisms of the enigmatic idiopathic pulmonary fibrosis: an integral model. *Am J Respir Crit Care Med* 2014;189:1161–1172.
5. Kropski JA, Blackwell TS, Loyd JE. The genetic basis of idiopathic pulmonary fibrosis. *Eur Respir J* 2015;45:1717–1727.
6. Moore C, Blumhagen RZ, Yang IV, Walts A, Powers J, Walker T, *et al*. Resequencing study confirms that host defense and cell senescence gene variants contribute to the risk of idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med* 2019;200:199–208.
7. Evans CM, Fingerlin TE, Schwarz MI, Lynch D, Kurche J, Warg L, *et al*. Idiopathic pulmonary fibrosis: a genetic disease that involves mucociliary dysfunction of the peripheral airways. *Physiol Rev* 2016;96:1567–1591.
8. Mathai SK, Newton CA, Schwartz DA, Garcia CK. Pulmonary fibrosis in the era of stratified medicine. *Thorax* 2016;71:1154–1160.
9. Allen RJ, Guillen-Guio B, Oldham JM, Ma S-F, Dressen A, Paynton ML, *et al*. Genome-wide association study of susceptibility to idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med* 2020;201:564–574.
10. Noth I, Zhang Y, Ma SF, Flores C, Barber M, Huang Y, *et al*. Genetic variants associated with idiopathic pulmonary fibrosis susceptibility and mortality: a genome-wide association study. *Lancet Respir Med* 2013;1:309–317.
11. Fingerlin TE, Murphy E, Zhang W, Peljto AL, Brown KK, Steele MP, *et al*. Genome-wide association study identifies multiple susceptibility loci for pulmonary fibrosis. *Nat Genet* 2013;45:613–620.
12. Allen RJ, Porte J, Braybrooke R, Flores C, Fingerlin TE, Oldham JM, *et al*. Genetic variants associated with susceptibility to idiopathic pulmonary fibrosis in people of European ancestry: a genome-wide association study. *Lancet Respir Med* 2017;5:869–880.
13. Woodcock HV, Eley JD, Guillotin D, Platé M, Nanthakumar CB, Martufi M, *et al*. The mTORC1/4E-BP1 axis represents a critical signaling node during fibrogenesis. *Nat Commun* 2019;10:6.
14. Seibold MA, Wise AL, Speer MC, Steele MP, Brown KK, Loyd JE, *et al*. A common MUC5B promoter polymorphism and pulmonary fibrosis. *N Engl J Med* 2011;364:1503–1512.
15. Peljto AL, Selman M, Kim DS, Murphy E, Tucker L, Pardo A, *et al*. The MUC5B promoter polymorphism is associated with idiopathic pulmonary fibrosis in a Mexican cohort but is rare among Asian ancestries. *Chest* 2015;147:460–464.
16. Ley B, Newton CA, Arnould I, Elicker BM, Henry TS, Vittinghoff E, *et al*. The MUC5B promoter polymorphism and telomere length in patients with chronic hypersensitivity pneumonitis: an observational cohort-control study. *Lancet Respir Med* 2017;5:639–647.
17. Juge PA, Lee JS, Ebstein E, Furukawa H, Dobrinskikh E, Gazal S, *et al*. MUC5B promoter variant and rheumatoid arthritis with interstitial lung disease. *N Engl J Med* 2018;379:2209–2219.
18. Ley B, Torgerson DG, Oldham JM, Adegunsoye A, Liu S, Li J, *et al*. Rare protein-altering telomere-related gene variants in patients with chronic hypersensitivity pneumonitis. *Am J Respir Crit Care Med* 2019;200:1154–1163.
19. Juge PA, Borie R, Kannengiesser C, Gazal S, Revy P, Wemeau-Stervinou L, *et al*.; FREX consortium. Shared genetic predisposition in rheumatoid arthritis-interstitial lung disease and familial pulmonary fibrosis. *Eur Respir J* 2017;49:1602314.

Positive Airway Pressure in Obesity Hypoventilation: Getting to the Heart of the Matter

The past decade has seen an increasing interest in the early identification and treatment of obesity hypoventilation syndrome (OHS). This has largely arisen from a greater recognition that untreated OHS is associated with significantly higher levels of morbidity and mortality compared with obstructive sleep apnea (OSA) alone (1). Individuals with OHS experience poorer quality of life, increased health resource use, and greater adverse socioeconomic impacts than equally obese eucapnic individuals (2, 3). In addition, more than 50% of those with OHS have echocardiographic evidence of pulmonary hypertension and left ventricular hypertrophy (4, 5). Positive airway pressure (PAP) remains the mainstay of therapy for OHS, aiming to correct sleep disordered breathing, reverse awake

Originally Published in Press as DOI: 10.1164/rccm.201911-2162ED on November 20, 2019