bleomycin-induced pulmonary fibrosis. *J Clin Invest* 2011;121: 277–287.
14. Liang J, Zhang Y, Xie T, Liu N, Chen H, Geng Y, *et al*. Hyaluronan and TLR4 promote surfactant-protein-C-positive alveolar progenitor cell renewal and prevent severe pulmonary fibrosis in mice. *Nat Med* 2016;22:1285–1293.

15. Guttman M, Donaghey J, Carey BW, Garber M, Grenier JK, Munson G, *et al*. lincRNAs act in the circuitry controlling pluripotency and differentiation. *Nature* 2011;477:295–300.

Resequencing to Fine Map Known Idiopathic Pulmonary Fibrosis Risk Genes
Homing in on Causal Variants

Genetic determinants of familial interstitial pneumonia, the familial form of idiopathic pulmonary fibrosis (IPF), were initially recognized as far back as 2001 as rare variants in the genes encoding *SFTPC* (surfactant protein C) and *TERT* (telomerase reverse transcriptase) (1, 2). Subsequent targeted candidate gene and whole-exome sequencing studies found that rare variation in candidate genes for familial IPF in the telomere complex (*TERT*, *RTEL1* [regulator of telomere elongation helicase 1], and *PARN* [poly(A)-specific ribonuclease]) and surfactant protein pathways (*SFTPA2* [surfactant protein A2] and *SFTPC*) also determined risk for sporadic IPF in individuals without a family history of IPF (Figure 1) (3–6). A linkage study in 2011 was the first to identify a SNP on chromosome 11p15.5 within the promoter of the mucin 5 gene (rs35705950 in *MUC5B*), with large effects on both familial and sporadic IPF risk replicated in subsequent genome-wide association studies (GWASs) (7).

For nearly a decade, GWASs have discovered 17 common genetic variants (allele frequency >5%) scattered across the genome associated with risk for IPF (Figure 1) (8–12). Among the most successfully replicated of these IPF risk loci was rs35705950 on *MUC5B*, confirming previous linkage studies and the large effect of this locus on familial and sporadic IPF pathogenesis (7). This led to the hypothesis that IPF results from increased *MUC5B* expression, causing excess production of mucus in the airways and thus impairing lung defense (13). The risk allele frequency of rs35705950 mirrors the observed prevalence of IPF, with a risk allele frequency of 10% in European individuals, the group with the highest IPF prevalence, whereas the risk allele is nearly nonexistent in African descent populations, in which IPF is much less prevalent. Although the *MUC5B* promoter polymorphism explains approximately 30% of the observed genetic risk (14), IPF is determined by environmental interactions with variation in multiple genes related to different pathogenic pathways identified by GWAS discoveries, including genes related to host defense

(*TOLLIP* [Toll-interacting protein]), telomere maintenance (*TERT*, *TERC* [telomerase RNA component]), signaling (*AKAP13* [A-kinase anchor protein 13]), and cell–cell adhesion (*DSP* [desmoplakin]) (8–12).

GWASs are based on chip genotyping data from subsets of SNPs to tag whole genomes and, more recently, additional imputed genotypes. Hence, GWASs are not usually sufficient to identify the causal variant and have the potential to miss uncovered or previously unknown rare variants. GWASs are also underpowered to detect rare variant associations owing to low frequency and the large number of rare variants throughout the genome. Because of these limitations, it can be difficult to conclusively determine the total number of causal variants within a genomic region. For example, there are contrasting reports of additional association signals within the chromosome 11p15.5 region independent of the *MUC5B* promoter polymorphism (9, 10). Hence, deep resequencing of known candidate genes for familial and sporadic IPF is required to fully characterize the genetic architecture for risk loci and map independent causative common and rare variants, both known and novel.

In this issue of the *Journal* (pp. 199–208), Moore and colleagues (15) describe targeted DNA resequencing of 16 genomic regions surrounding loci previously associated with risk for familial or sporadic IPF. Common variants (allele frequency ≥3%) were investigated individually in 3,624 individuals with IPF and 4,442 control subjects, and rare variants were investigated using gene-level and region-based tests in a subgroup of 7,116 subjects with confirmed European ancestry. This is the largest resequencing study of known loci for this uncommon respiratory disease, which has been evaluated mostly in smaller IPF genetic studies. The size of this large IPF cohort in combination with the use of deep, targeted gene resequencing analyzed with conditional analyses and gene-level rare variant tests allowed for the most detailed estimates of the contribution of multiple common and rare variants to IPF risk currently possible.

This resequencing study demonstrated several important aspects of the genetic architecture of known IPF risk loci. First, this study confirmed reported associations for rare and common variations initially described for the genomic regions investigated. However, the top associations were not always at the initially reported sentinel variant (i.e., near *ZKSCAN1* [Zinc finger protein with KRAB and SCAN domains 1] and *IVD* [isovaleryl-coenzyme A dehydrogenase]), demonstrating the power of resequencing
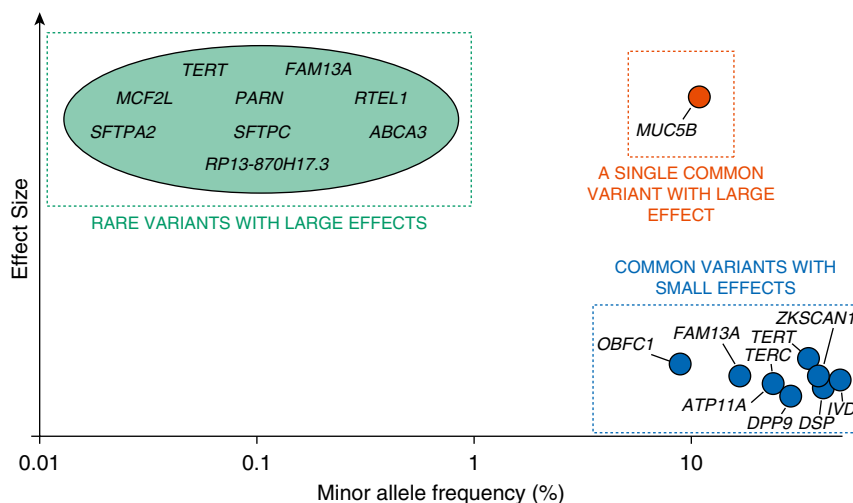
Figure 1. Allele frequencies and effect sizes of known risk loci for familial and sporadic idiopathic pulmonary fibrosis. Plot shows the distribution of allele frequencies for risk variants identified within the genes shown with estimated effect sizes based on prior studies.

to detect causal variation. Second, known common variant associations were determined by a single, independent signal, including 11p15.5, demonstrating that the multiple signals previously reported for *MUC5B* (including *TOLLIP*) were likely due to linkage disequilibrium with rs35705950 (9). Third, GWAS loci with common IPF risk variants have the potential to harbor rare variations independently associated with IPF. Specifically, rare variants in the 11p15.5 region downstream of RP13-870H17.3 (near *MUC2*), *TERT*, and *FAM13A* (family with sequence similarity 13 member A) were independently associated with IPF, of which rare variant associations in 11p15.5 and *FAM13A* were novel. Fourth, this study confirms that loci with rare variants associated with familial IPF potentially contain rare variants associated with sporadic IPF but do not harbor common genetic determinants. Finally, these analyses confidently confirmed prior causal variants, including the *MUC5B* promotor variant, which had an effect size similar to those in multiple independent studies (16).

Despite the novel and confirmatory aspects of this large resequencing study, there remain aspects of the genetics of IPF not fully addressed. First, the known loci evaluated in this study account for only a small proportion of the observed risk for sporadic IPF. Hence, additional IPF risk loci could be identified in larger cohorts with more detailed genotype data inclusive of whole-genome sequencing for future GWASs. Second, many genetic associations are regulatory loci in noncoding regions identified with mRNA expression data (expression quantitative trait loci [eQTL] analyses) from lung tissue (10, 12). The challenge of evaluating IPF risk loci outside of coding regions relates in part to the lack of available lung tissue from patients with IPF for eQTL and epigenetic studies to identify regulatory variation for a causal gene within an associated noncoding region. Third, most genetic studies of IPF have been performed in European white descent cohorts because of the high prevalence of IPF in white individuals and to minimize confounding by varying ancestral backgrounds. IPF does not appear exclusively in white individuals, and studies limited to a single race could miss IPF risk loci unique to different ancestries. Finally, this study compared variation between subjects with IPF and control subjects, which is not informative regarding effects on

severity, progression, or treatment response, which require genetic studies limited to IPF cases.

This resequencing study provides the most detailed characterization of the genetic architecture of known IPF risk loci, including the cumulative and independent genetic factors, both common and rare, constituting IPF risk in familial and sporadic cases with varying effect sizes (Figure 1). In this instance, deep sequencing of known loci rigorously confirmed causal variation from prior studies while identifying novel risk variation. As the cost of next-generation sequencing improves and the number of patients with IPF with genetic data and tissue samples for epigenetic and eQTL studies increases, future studies will improve understanding of this fatal disease without a known cure. ∎

Richard Allen, Ph.D.
Department of Health Sciences
University of Leicester
Leicester, United Kingdom

Victor E. Ortega, M.D., Ph.D.
Department of Internal Medicine
Center for Precision Medicine
Winston-Salem, North Carolina

ORCID ID: 0000-0002-8450-3056 (R.A.).

References

1. Nogee LM, Dunbar AE III, Wert SE, Askin F, Hamvas A, Whitsett JA. A mutation in the surfactant protein C gene associated with familial interstitial lung disease. *N Engl J Med* 2001;344: 573–579.
2. Armanios M, Chen JL, Chang YP, Brodsky RA, Hawkins A, Griffin CA, *et al*. Haploinsufficiency of telomerase reverse transcriptase leads to anticipation in autosomal dominant dyskeratosis congenita. *Proc Natl Acad Sci USA* 2005;102:15960–15964.

3. Coghlan MA, Shifren A, Huang HJ, Russell TD, Mitra RD, Zhang Q, *et al*. Sequencing of idiopathic pulmonary fibrosis-related genes reveals independent single gene associations. *BMJ Open Respir Res* 2014;1: e000057.

4. Stuart BD, Choi J, Zaidi S, Xing C, Holohan B, Chen R, *et al*. Exome sequencing links mutations in PARN and RTEL1 with familial pulmonary fibrosis and telomere shortening. *Nat Genet* 2015;47: 512–517.

5. van Moorsel CH, ten Klooster L, van Oosterhout MF, de Jong PA, Adams H, Wouter van Es H, *et al*. SFTPA2 mutations in familial and sporadic idiopathic interstitial pneumonia. *Am J Respir Crit Care Med* 2015; 192:1249–1252.

6. Petrovski S, Todd JL, Durheim MT, Wang Q, Chien JW, Kelly FL, *et al*. An exome sequencing study to assess the role of rare genetic variation in pulmonary fibrosis. *Am J Respir Crit Care Med* 2017;196: 82–93.

7. 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.

8. Mushiroda T, Wattanapokayakit S, Takahashi A, Nukiwa T, Kudoh S, Ogura T, *et al*.; Pirfenidone Clinical Study Group. A genome-wide association study identifies an association of a common variant in TERT with susceptibility to idiopathic pulmonary fibrosis. *J Med Genet* 2008;45:654–656.

9. 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.

10. 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.

11. Fingerlin TE, Zhang W, Yang IV, Ainsworth HC, Russell PH, Blumhagen RZ, *et al*. Genome-wide imputation study identifies novel HLA locus for pulmonary fibrosis and potential role for auto-immunity in fibrotic idiopathic interstitial pneumonia. *BMC Genet* 2016;17:74.

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 Resp Med* 2017;5:869–880.

13. 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.

14. Mathai SK, Newton CA, Schwartz DA, Garcia CK. Pulmonary fibrosis in the era of stratified medicine. *Thorax* 2016;71:1154–1160.

15. 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.

16. Zhu QQ, Zhang XL, Zhang SM, Tang SW, Min HY, Yi L, *et al*. Association between the MUC5B promoter polymorphism rs35705950 and idiopathic pulmonary fibrosis: a meta-analysis and trial sequential analysis in Caucasian and Asian populations. *Medicine (Baltimore)* 2015;94:e1901.

ᵔ Ironing Out the Roles of Macrophages in Idiopathic Pulmonary Fibrosis

Airway macrophages (AMs) act to remove inhaled particles, debris, allergens, and microbes, making them crucial to host defense and epithelial homeostasis (1). Cross-talk between AMs, dendritic cells, alveolar epithelial cells, and T cells regulates how the immune system responds to environmental lung stimuli. Thus, AMs are pivotal in the response to alveolar injury and the subsequent biological pathways regulating resolution or persistence of alveolar inflammation (2). Given this keystone role, AMs have been studied in both human disease and animal models of pulmonary fibrosis (3–5), where the prevailing notion is that AMs derived from circulating monocytes worsen disease (5, 6). However, the mechanism(s) responsible for the ability of AMs to promote pulmonary fibrogenesis are unclear.

In this issue of the *Journal* (pp. 209–219), Allden and colleagues advance knowledge about macrophages and idiopathic pulmonary fibrosis (IPF) (7). They used BAL acquired from two independent IPF patient cohorts from observational clinical studies and interrogated specific lung leukocyte phenotypes by complimentary techniques, including multiparametric flow

cytometry matched with immunohistochemistry. They show increased proportions of AMs lacking surface CD71 (the transferrin receptor) in human patients with IPF. The authors carefully interrogated the phenotype of these CD71⁻ AMs. Remarkably, CD71⁻ AMs demonstrated impaired function with defective phagocytosis, reduced markers of macrophage maturity, and profibrotic gene activation. In a further provocative turn, the authors demonstrated an association between reduced survival and higher proportions of CD71⁻ AMs in the BAL using a cohort of 97 patients with IPF. They conclude that CD71⁻ AM percentages may serve as a novel biomarker of IPF disease progression and that the CD71 pathway may represent a target for therapeutic manipulation.

CD71 is an integral membrane protein that binds diferric transferrin complexes to mediate uptake into the cell via receptor-mediated endocytosis. The majority of iron in circulation in the steady state is bound to transferrin. Transferrin–iron complexes bind CD71, which serves as a cellular receptor but also serves to limit the ability of iron to catalyze the formation of free radicals from reactive oxygen species, resulting in iron toxicity (8). Furthermore, control of free iron is an important host defense function because it limits availability of iron to potentially pathogenic bacteria *in vivo*.

Given these associations between macrophages, iron, bacteria, and fibrosis, it is interesting to speculate about how and why loss of CD71 on IPF macrophages may be important for disease pathogenesis. Previous studies have shown elevated levels of iron and altered iron metabolism in IPF lungs (9–11). Allden and