

MUC5B and Idiopathic Pulmonary Fibrosis

Ivana V. Yang^{1,2,3}, Tasha E. Fingerlin^{3,4}, Christopher M. Evans^{1,5}, Marvin I. Schwarz¹, and David A. Schwartz^{1,3,5}

¹Department of Medicine, University of Colorado Denver, Aurora, Colorado; ²Department of Epidemiology, Colorado School of Public Health, Aurora, Colorado; ³Center for Genes, Environment, and Health, National Jewish Health, Denver, Colorado; ⁴Department of Biostatistics and Informatics, Colorado School of Public Health, Aurora, Colorado; and ⁵Department of Immunology, University of Colorado Denver, Aurora, Colorado

Abstract

Idiopathic pulmonary fibrosis (IPF), a fatal disease that is a result of complex interactions between genetics and the environment, has limited treatment options. We have identified the *MUC5B* promoter polymorphism and other common genetic variants that in aggregate explain roughly one-third of disease risk. The *MUC5B* promoter polymorphism is the strongest and the most replicated genetic risk factor for IPF, appears to be protective and predictive in this disease, and is likely involved in disease pathogenesis through an increase in *MUC5B* expression in

terminal bronchi and honeycombed cysts. Expression of *MUC5B* is also highly correlated with expression of cilium genes in IPF lung. Our work suggests that mucociliary dysfunction in the distal airway may play a role in the development of progressive fibroproliferative lung disease. In addition, our work has important implications for secondary prevention, early detection, and future early and personalized treatment based on genetic profiles.

Keywords: *MUC5B*; mucociliary dysfunction; pulmonary fibrosis; genetic variant; gene expression profile

(Received in original form March 1, 2015; accepted in final form April 13, 2015)

Correspondence and requests for reprints should be addressed to Ivana V. Yang, Ph.D., University of Colorado Denver, 12700 East 19th Avenue, 8611, Aurora, CO 80045. E-mail: ivana.yang@ucdenver.edu

Ann Am Thorac Soc Vol 12, Supplement 2, pp S193–S199, Nov 2015

Copyright © 2015 by the American Thoracic Society

DOI: 10.1513/AnnalsATS.201503-110AW

Internet address: www.atsjournals.org

Idiopathic pulmonary fibrosis (IPF) is characterized by progressive scarring of the pulmonary parenchyma that leads to progressive loss of lung function with dyspnea and hypoxemia, and ultimately respiratory failure and death. The prevalence of IPF is currently estimated at 63 individuals in 100,000 in the United States (1), the median survival is 3 years, and only limited treatment options exist (2, 3). IPF increases in prevalence with age (1) and as our population ages, IPF will affect a larger proportion of individuals. Disease prevalence is highly dependent on the racial/ethnic background of the individual; whites appear to be at higher risk of developing IPF than Hispanics and Asians, and this disease is rare in populations of African descent (1).

It is currently believed that IPF results from aberrant activation of injured alveolar epithelial cells, which produce mediators

that lead to proliferation of resident fibroblasts, fibrocyte recruitment, and epithelial–mesenchymal transition (EMT) resulting in the formation of myofibroblastic foci, followed by accumulation of extracellular matrix (ECM), dysregulated wound repair, and lung remodeling (4). Fibroblasts from these sources are thought to be the key cell type in the disease pathophysiology (4); however, the extensive work that has been performed on fibroblast biology has not resulted in many promising therapeutic strategies so far. It is becoming increasingly clear that the disease process underlying the IPF phenotype is heterogeneous and many different molecular processes may be involved. In addition to EMT (5–7), other potential processes include growth factor regulation (8), apoptosis (9, 10), oxidative stress (11), endoplasmic reticulum (ER) stress (12,

13), cellular senescence associated with aging/telomere shortening (14–16), epithelial stem cell exhaustion (17), intraalveolar coagulation (18), and aberrant recapitulation of developmental pathways (19).

Genetic Basis of Pulmonary Fibrosis

IPF is likely the result of complex interactions between genetic and environmental factors, such as cigarette smoke (20, 21). Evidence of the genetic basis for pulmonary fibrosis is substantial; familial aggregation has been confirmed through a variety of studies in twins, siblings raised apart, and multigenerational families (22, 23). Pulmonary fibrosis is associated with mutations in surfactant protein C (24–26), surfactant protein A2

(27), and in genes that maintain telomere length (14, 28–31).

Beginning with a genome-wide linkage screen in 82 families with IPF, our group identified a 3.4-Mb region of chromosome 11p15 that was associated with the presence of the disease. To narrow the genomic interval to the gene(s) that harbor genetic variants associated with disease, we performed fine mapping using association analysis in 492 subjects with IPF and 322 control subjects. This led to the discovery of a common variant in the promoter region of *MUC5B* (allelic frequency of 33.8% of familial idiopathic interstitial pneumonia [FIP] case subjects, 37.5% of IPF case subjects, and 9.1% of control subjects) that has a profound effect on the risk of developing either familial or sporadic forms of idiopathic interstitial pneumonia (IIP) (32). Odds ratios for disease in subjects heterozygous (GT) and homozygous (TT) for the minor allele of this *MUC5B* polymorphism (rs35705950) were 6.8 (95% confidence interval [CI], 3.9–12.0) and 20.8 (95% CI, 3.8–113.7) for FIP ($P = 1.2 \times 10^{-15}$), and 9.0 (95% CI, 6.2–13.1) and 21.8 (95% CI, 5.1–93.5) for IPF ($P = 2.5 \times 10^{-37}$) in the original study. This finding has been validated in seven independent non-Hispanic white (NHW) cohorts (33–39), and this promoter polymorphism remains the strongest risk factor (genetic and otherwise) for pulmonary fibrosis.

The association of the *MUC5B* promoter polymorphism appears to be specific to pulmonary fibrosis; we have not detected significant associations with other lung diseases, specifically systemic sclerosis interstitial lung disease (37, 40), asbestosis, sarcoidosis, acute lung injury/acute respiratory distress syndrome, chronic obstructive pulmonary disease, and asthma (our unpublished data). The relationship of the *MUC5B* promoter single-nucleotide polymorphism (SNP) and cough and mucus production has been investigated, and patients with IPF with the rs35705950 allele are more likely to have cough-related symptomatology (41).

Analogous to disease prevalence, the association of the *MUC5B* promoter polymorphism depends greatly on racial/ethnic background. In addition to being the strongest genetic risk factor for pulmonary fibrosis among NHW patients, this SNP is also a strong genetic risk factor for Mexican patients with IPF (odds ratio, 7.36; $P = 0.0001$) (42). Similar to the findings from

the 1000 Genomes Project (43), we found that the *MUC5B* promoter SNP is rare among Korean case subjects with IPF (42) and was absent in nondiseased Korean control subjects. Other studies showed a slightly higher prevalence of the SNP among Japanese case subjects with IPF (3.4%) compared with healthy Japanese subjects (0.8%) (39) as well as Chinese case subjects with IPF (3.3%) compared with control subjects (0.7%) (44), extending the risk of the *MUC5B* promoter allele to Asian populations. However, the prevalence of rs35705950 is clearly higher in both the NHW control and case populations, suggesting a selective advantage of the allele. Interestingly, the *MUC5B* polymorphism is not present in African populations, also mirroring the fact that populations of African descent are rarely diagnosed with IPF (1).

***MUC5B* Promoter Polymorphism and Clinical Outcomes of Pulmonary Fibrosis**

The *MUC5B* promoter polymorphism appears to be predictive (45) and prognostic (46) in IPF. Using the Framingham Heart Study population, we have shown that rs35705950 is associated with interstitial lung abnormalities on chest computed tomography (CT) scan; for carriers of the T (minor) allele, we observed a 2.8-fold increase in the odds of having interstitial lung abnormalities (95% CI, 2.0–3.9; $P < 0.001$) and a 6.3-fold increase in the odds of having definite CT evidence of pulmonary fibrosis (95% CI, 3.1–12.7; $P < 0.001$) (45), suggesting that the *MUC5B* promoter SNP may prove useful to identify individuals with preclinical forms of IPF. Moreover, in two independent cohorts of patients with IPF, we have demonstrated that the *MUC5B* promoter polymorphism is associated with twofold improved survival (46), suggesting that the *MUC5B* genotype (and possibly other genotypes) may be helpful in identifying unique clinical phenotypes of this heterogeneous disease. The survival differences may be explained by enhanced host defense, as suggested by our collaborative work in mice (47) and humans (48) that suggests that enhanced mucin production is associated with improved microbial host defense. Although these findings may help explain the survival advantage associated with the *MUC5B*

promoter variant, these findings also suggest that enhanced host defense may positively select for individuals with this allele and explain the high prevalence of this allele in NHW populations. Moreover, the *MUC5B* promoter variant appears to be identifying a subtype of IPF that can be diagnosed earlier and has a more benign course, further affirming the genotype–phenotype relationship.

***MUC5B* Promoter Polymorphism and Expression of *MUC5B* in Pulmonary Fibrosis**

rs35705950 is located approximately 3 kb upstream of the *MUC5B* transcription start site in a highly conserved sequence region that is predicted to be regulatory (49). The G-to-T transversion was predicted in our original publication (32) to disrupt an E2F binding site and create two new binding sites (HOX9 and PAX2). This area of the promoter was also demonstrated to bind NF- κ B in airway epithelial cells (50) and has been more recently shown to bind the glucocorticoid receptor in dexamethasone-stimulated A549 cells in ENCODE (Encyclopedia of DNA Elements) data (51). Moreover, ENCODE analysis of 125 cell lines shows the SNP to be in an area of open chromatin as evidenced by DNase I hypersensitivity analysis as well as synthesized results of FAIRE (Formaldehyde-Assisted Isolation of Regulatory Elements)-seq and ChIP (Chromatin Immunoprecipitation)-seq analyses of chromatin marks. Finally, ChromHMM analysis (52) of ENCODE data predicted the SNP to be in the area of the promoter that is poised for transcription on stimulation. Collectively, these observations suggest that the *MUC5B* promoter SNP rs35705950 is highly likely to be regulatory in IPF lung. Consistent with these observations, we found that the *MUC5B* promoter variant is associated with a 34.1-fold increase in *MUC5B* expression in lung tissue among unaffected subjects (32, 53) and with a 5.3-fold increase among patients with IPF (53), with patients with IPF expressing 14.1-fold more *MUC5B* than unaffected control subjects (32). *MUC5B* immunohistochemical staining showed dense accumulation of *MUC5B* in terminal bronchioles as well as the pseudostratified bronchial epithelium and lumen of honeycomb cysts (Figure 1)

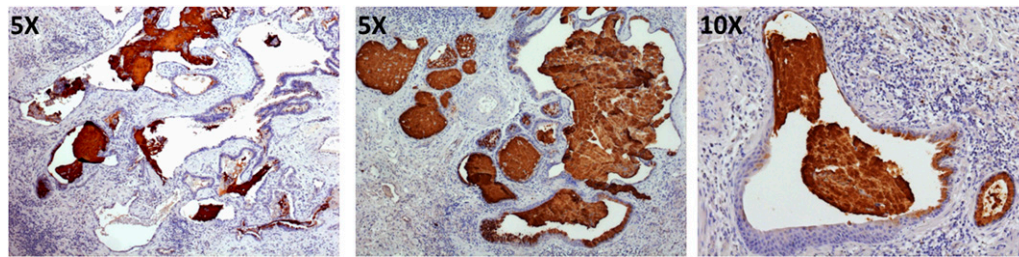


Figure 1. Immunohistochemical staining for MUC5B protein (*brown*) in cystic structures in idiopathic pulmonary fibrosis (IPF) lung. *Left:* An area of IPF lung tissue containing histologically normal airway and honeycomb cysts. *Middle and right:* Honeycomb cysts, exclusively. In all three panels, honeycomb cysts are filled with mucus. Tissue was counterstained with hematoxylin.

(32, 54), a prominent pathologic lesion in IPF lung.

We also identified two molecular subtypes of IPF based on a strong gene expression signature of cilium-associated genes (Gene Ontology [GO] category 0005929; Benjamini-corrected $P = 3.7 \times 10^{-11}$) and their structural components (axoneme, $P = 3.9 \times 10^{-11}$; dynein, $P = 9.4 \times 10^{-7}$) (55) in IPF lung tissue (Figure 2). The most significant network of genes contains cilium genes as well as transcription factors that regulate expression of cilium genes (*FOXJ1*, *RFX2*, and *RFX3*). The cilium gene signature was associated with the presence of microscopic honeycombing ($P < 0.0001$), the pathologic lesion in the IPF lung in which MUC5B is overproduced (32, 54), but not with fibroblastic foci, another characteristic histological feature of IPF. Importantly, the signature was validated in multiple lung specimens from the same subjects and in an independent cohort of subjects with IPF. Moreover, the signature is unique to IPF (not present in chronic obstructive pulmonary disease) and it contains both motile and nonmotile primary cilium genes. MUC5B expression is highly correlated to expression of cilium genes ($r = 0.76-0.90$; $P = 6.3 \times 10^{-33}-8.4 \times 10^{-45}$). This is also the case for *MMP7* ($r = 0.71-0.83$; $P = 1.1 \times 10^{-25}-8.4 \times 10^{-45}$), an extracellular matrix gene that has emerged as the main expression biomarker for IPF (56–58) and was shown to play a role in attenuating ciliated cell differentiation during wound repair (59). A follow-up study to our publication demonstrated that BPIFB1/LPLUNC1, the most differentially expressed gene in two molecular subtypes of IPF, colocalizes with MUC5B to the bronchiolized epithelium in

the honeycomb cysts in IPF lung (60). Fifteen cilium genes from the GO category 0005929 also have a small but significant ($P < 0.05$) change in expression in the lung tissue of individuals with IPF who are polymorphic for the MUC5B promoter SNP (GT and TT) compared with wild-type (GG) individuals (our unpublished data).

MUC5B Promoter Polymorphism and Pathogenesis of Pulmonary Fibrosis

In the absence of direct evidence from cell or animal models, we can only speculate on the role that MUC5B plays in the pathogenesis of IPF. We believe that at least two distinct mechanisms fit the current evidence, that they are not mutually exclusive, and that they could both contribute to the pathogenesis of IPF. One possible mechanism is that MUC5B overexpression results in chronic mucus hypersecretion and accumulation in the peripheral airspace, which in turn impairs mucociliary transport, results in mucus adhesion in the bronchoalveolar region, and consequently induces and potentiates chronic inflammation and injury (61). Supporting this, MUC5B accumulates in terminal bronchioles and areas of microscopic honeycombing in IPF lung (Figure 1) (32, 54). Given the critical role of the *Muc5b* gene in acute microbial host defense in mice (47), it is feasible that excess mucus adversely alters host defense response to chronic, persistent exposure to microorganisms. These could include gammaherpesviral infections that have been associated with IPF (62) or other persistent infections that have not yet been associated with IPF (63).

The other distinct possibility is that MUC5B overexpression results in an aberrant repair process after injury to the bronchoalveolar regions of the lung. In this case, honeycomb cysts in IPF lung would represent a failed regenerative process after injury. A body of evidence demonstrates that various populations of cells in the distal lung may contribute to fibroproliferation after injury (64). Using lineage tracing experiments in mice, Rock and colleagues showed that airway epithelial cells give rise to alveolar type I and alveolar type II epithelial cells in bleomycin-induced pulmonary fibrosis (65). Another study demonstrated that distal airway stem cells proliferate into interbronchial regions of alveolar ablation after H1N1 infection and assemble into alveolus-like structures (66). Finally, it was shown that distinct stem/progenitor cell pools repopulate injured tissue depending on the extent of the injury (67). This study identified a rare, undifferentiated lineage-negative epithelial stem/progenitor (LNEP) cell population that is the major responder in distal lung after severe damage and whose activation and differentiation state is modulated by Notch signaling. Moreover, Notch signaling promoted the presence of abnormal parenchymal structures that derive from LNEPs in mice after H1N1 infection that resembled honeycomb cysts in human IPF lung.

The cilium gene expression signature we identified contains both motile and primary cilium genes (55). It is feasible that aberrant expression of cilium genes in the multiciliated cells in distal airways and terminal bronchioles and/or primary cilium genes in any of the cells in the distal lung may play a role in the pathogenesis of IPF. A role for cilium gene expression and

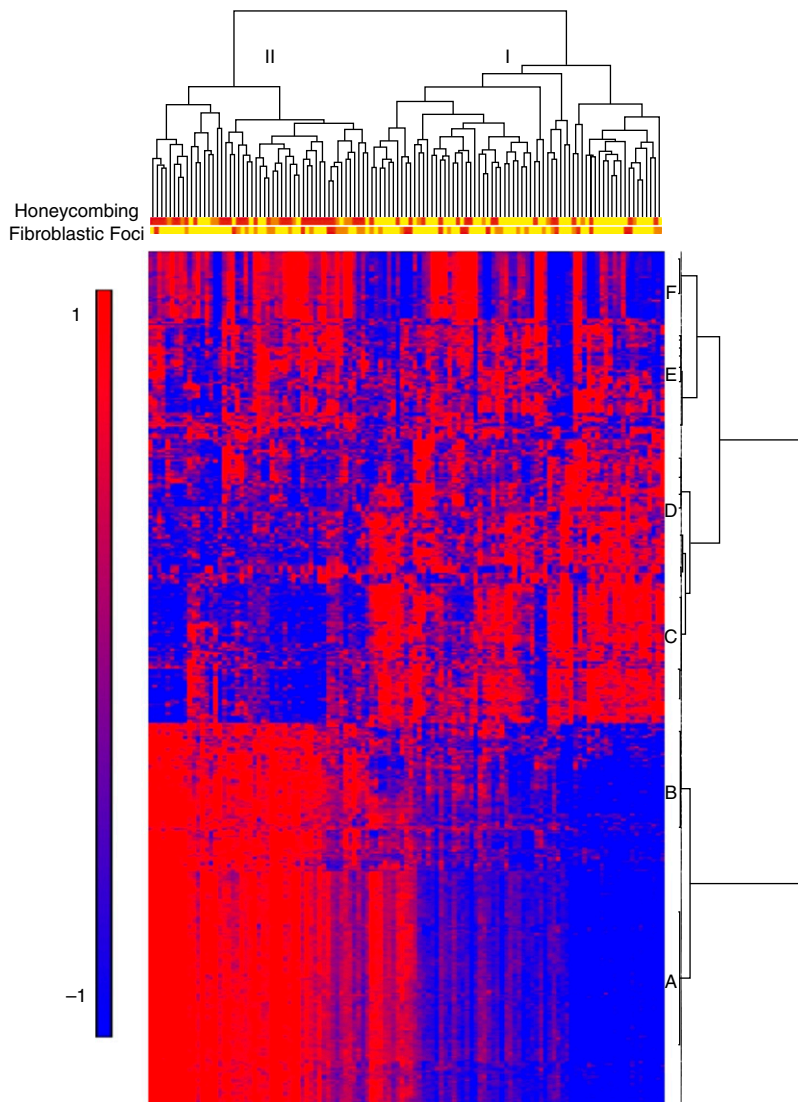


Figure 2. Gene expression profiling identifies two subtypes of idiopathic pulmonary fibrosis/usual interstitial pneumonia (IPF/UIP). mRNA profiles from 119 IPF/UIP lungs were subject to hierarchical clustering based on the expression of 472 transcripts that are differentially expressed at a 5% false discovery rate and with a greater than 2-fold change in IPF/UIP compared with control lung. The distance metric is Euclidean, with complete linkage across samples and Ward's linkage across genes. Extent of honeycombing and fibroblastic foci in each sample as assessed by pathology is depicted by the color: yellow (unscored/not present), orange (rare), red (present). Reprinted by permission from Reference 55.

function in multiciliated airway epithelial cells in the distal airways and bronchioles would be consistent with the idea that accumulation of mucus in the bronchoalveolar region induces and potentiates chronic inflammation and injury. On the other hand, developmental pathways have been shown to signal in the primary cilium (68), and their aberrant recapitulation is prominent in IPF lung (19). It has been shown that primary ciliogenesis precedes motile cilia during

development and that primary cilia are transiently present in adult mouse airway during repair after respiratory virus injury (69). A role for primary cilia in the pathogenesis of IPF would be consistent with the aberrant repair process after injury to the bronchoalveolar regions of the lung.

Regardless of the mechanism(s) by which *MUC5B* overexpression influences the development of IPF, our findings suggest that the pseudostratified airway epithelium in the distal lung may play a role in this

disease. It is therefore important to start examining the role of injury and repair in bronchoalveolar regions of the lung in the pathogenesis of pulmonary fibrosis.

Additional Genetic Risk Loci for Pulmonary Fibrosis

Although the *MUC5B* promoter variant is a strong risk factor for IPF, it is likely that other genetic variants, alone or in combination with each other (gene \times gene) or environmental exposures (gene \times environment), contribute to the development of this disease. This is supported by the fact that the minor allele of *MUC5B* is present in the heterozygote or homozygote state in 19% of the NHW population and that IPF occurs in far less than 1% of the population (70, 71). With the goal of identifying additional genetic risk factors for pulmonary fibrosis, we conducted a case-control genome-wide association study (GWAS) in 1,616 NHW case subjects of fibrotic idiopathic interstitial pneumonias (fIIP) (IPF, 1,251 [77%]) and 4,683 control subjects followed by a replication study of 876 NHW case subjects (IPF, 774 [88%]) and 1,890 control subjects (35). We performed a meta-analysis of the discovery and replication phases to identify the SNPs with genome-wide evidence for association with fIIP ($P < 5 \times 10^{-8}$). In addition to confirming the association with *TERT* at 5p15, *MUC5B* at 11p15, and the 3q26 region near *TERC*, we also identified seven new loci, including *FAM13A* (4q22), *DSP* (6p24), *OBFC1* (10q24), *ATP11A* (13q34), *DPP9* (19p13), and chromosomal regions 7q22 and 15q14–15. In aggregate, the common risk variants associated with IIP in our GWAS point to the role of host defense (*MUC5B*, *ATP11A*), cell-cell adhesion (*DSP* and *DPP9*), and DNA repair (*TERT*, *TERC*, and *OBFC1*) in the initiation of this disease. We also identified an association at 17q21, but the association at that locus is confounded by the H2 1-Mb inversion polymorphism haplotype (72, 73) that is present at high frequencies among Europeans. Among those that did not carry the H2 haplotype, we did not observe enough variation in SNPs in this region variation to allow robust tests of association, and it was therefore not possible to determine whether it is truly a risk factor for pulmonary fibrosis (35).

Importantly, the 10 genetic risk loci we identified (excluding rs35705950) in aggregate account for approximately one-third of disease risk, demonstrating the importance of common genetic variation in the development of IIPs. There were no substantial differences in odds ratios for familial and sporadic forms of the disease or for IPF and non-IPF IIPs, suggesting that the genetic etiology of various forms of IIP is similar.

Implications for Personalized Medicine

Our work on understanding the genetic basis of pulmonary fibrosis has resulted in several important observations. First, pulmonary fibrosis is a complex genetic disease with common variants identified so far accounting for approximately one-third of the disease

risk. The odds ratios for many of the variants we identified and especially the *MUC5B* promoter SNP are substantially higher than those observed in other complex diseases. Second, familial and sporadic forms of pulmonary fibrosis have similar genetic risks, suggesting that siblings of individuals with either familial or sporadic IPF could be detected at an earlier stage. Moreover, the *MUC5B* promoter variant appears to be predictive of IPF in the general population, opening up an additional potential avenue for early disease detection. Third, although mechanistic studies on the role of the genetic variants in disease pathogenesis are so far limited, they are beginning to point to the role for *MUC5B* overexpression, driven by the promoter polymorphism, in the development of pulmonary fibrosis. Finally, localization of *MUC5B* expression in IPF lung and the cilium gene expression fingerprint suggest a prominent role of injury

and disrupted repair to the distal airway epithelium in IPF, a condition that has traditionally been viewed as a disease of the alveolar epithelium. These findings suggest that mucociliary dysfunction in the distal airway may play a role in the development of progressive fibroproliferative lung disease.

Collectively, our findings have important implications for future diagnosis and treatment in a fatal disease with limited treatment options. We envision that genetic variants, alone or in combination with gene expression fingerprints and/or peripheral blood biomarkers, will allow us to treat patients on the basis of the genetic subtype of IPF they have, and moreover to initiate treatment in the preclinical stages of their disease rather than at a late stage when prognosis is poor. ■

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

References

- Ley B, Collard HR. Epidemiology of idiopathic pulmonary fibrosis. *Clin Epidemiol* 2013;5:483–492.
- King TE Jr, Bradford WZ, Castro-Bernardini S, Fagan EA, Glaspole I, Glassberg MK, Gorina E, Hopkins PM, Kardatzke D, Lancaster L, et al.; ASCEND Study Group. A phase 3 trial of pirfenidone in patients with idiopathic pulmonary fibrosis. *N Engl J Med* 2014;370:2083–2092.
- Richeldi L, du Bois RM, Raghu G, Azuma A, Brown KK, Costabel U, Cottin V, Flaherty KR, Hansell DM, Inoue Y, et al.; INPULSIS Trial Investigators. Efficacy and safety of nintedanib in idiopathic pulmonary fibrosis. *N Engl J Med* 2014;370:2071–2082.
- King TE Jr, Pardo A, Selman M. Idiopathic pulmonary fibrosis. *Lancet* 2011;378:1949–1961.
- Willis BC, Liebler JM, Luby-Phelps K, Nicholson AG, Crandall ED, du Bois RM, Borok Z. Induction of epithelial–mesenchymal transition in alveolar epithelial cells by transforming growth factor- β : potential role in idiopathic pulmonary fibrosis. *Am J Pathol* 2005;166:1321–1332.
- Kim KK, Kugler MC, Wolters PJ, Robillard L, Galvez MG, Brumwell AN, Sheppard D, Chapman HA. Alveolar epithelial cell mesenchymal transition develops *in vivo* during pulmonary fibrosis and is regulated by the extracellular matrix. *Proc Natl Acad Sci USA* 2006;103:13180–13185.
- Tanjore H, Xu XC, Polosukhin VV, Degryse AL, Li B, Han W, Sherrill TP, Plieth D, Neilson EG, Blackwell TS, et al. Contribution of epithelial-derived fibroblasts to bleomycin-induced lung fibrosis. *Am J Respir Crit Care Med* 2009;180:657–665.
- Farkas L, Gauldie J, Voelkel NF, Kolb M. Pulmonary hypertension and idiopathic pulmonary fibrosis: a tale of angiogenesis, apoptosis, and growth factors. *Am J Respir Cell Mol Biol* 2011;45:1–15.
- Fattman CL. Apoptosis in pulmonary fibrosis: too much or not enough? *Antioxid Redox Signal* 2008;10:379–385.
- Thannickal VJ, Horowitz JC. Evolving concepts of apoptosis in idiopathic pulmonary fibrosis. *Proc Am Thorac Soc* 2006;3:350–356.
- Hecker L, Vittal R, Jones T, Jagirdar R, Luckhardt TR, Horowitz JC, Pennathur S, Martinez FJ, Thannickal VJ. NADPH oxidase-4 mediates myofibroblast activation and fibrogenic responses to lung injury. *Nat Med* 2009;15:1077–1081.
- Tanjore H, Cheng DS, Degryse AL, Zoz DF, Abdolrasulnia R, Lawson WE, Blackwell TS. Alveolar epithelial cells undergo epithelial-to-mesenchymal transition in response to endoplasmic reticulum stress. *J Biol Chem* 2011;286:30972–30980.
- Lawson WE, Cheng DS, Degryse AL, Tanjore H, Polosukhin VV, Xu XC, Newcomb DC, Jones BR, Roldan J, Lane KB, et al. Endoplasmic reticulum stress enhances fibrotic remodeling in the lungs. *Proc Natl Acad Sci USA* 2011;108:10562–10567.
- Tsakiri KD, Cronkhite JT, Kuan PJ, Xing C, Raghu G, Weissler JC, Rosenblatt RL, Shay JW, Garcia CK. Adult-onset pulmonary fibrosis caused by mutations in telomerase. *Proc Natl Acad Sci USA* 2007;104:7552–7557.
- Alder JK, Chen JJ, Lancaster L, Danoff S, Su SC, Cogan JD, Vulto I, Xie M, Qi X, Tudor RM, et al. Short telomeres are a risk factor for idiopathic pulmonary fibrosis. *Proc Natl Acad Sci USA* 2008;105:13051–13056.
- Cronkhite JT, Xing C, Raghu G, Chin KM, Torres F, Rosenblatt RL, Garcia CK. Telomere shortening in familial and sporadic pulmonary fibrosis. *Am J Respir Crit Care Med* 2008;178:729–737.
- Chilosi M, Doglioni C, Murer B, Poletti V. Epithelial stem cell exhaustion in the pathogenesis of idiopathic pulmonary fibrosis. *Sarcoidosis Vasc Diffuse Lung Dis* 2010;27:7–18.
- Scotton CJ, Krupiczkoj MA, Königshoff M, Mercer PF, Lee YC, Kaminski N, Morser J, Post JM, Maher TM, Nicholson AG, et al. Increased local expression of coagulation factor X contributes to the fibrotic response in human and murine lung injury. *J Clin Invest* 2009;119:2550–2563.
- Selman M, Pardo A, Kaminski N. Idiopathic pulmonary fibrosis: aberrant recapitulation of developmental programs? *PLoS Med* 2008;5:e62.
- Taskar VS, Coultas DB. Is idiopathic pulmonary fibrosis an environmental disease? *Proc Am Thorac Soc* 2006;3:293–298.
- Ding Q, Luckhardt T, Hecker L, Zhou Y, Liu G, Antony VB, deAndrade J, Thannickal VJ. New insights into the pathogenesis and treatment of idiopathic pulmonary fibrosis. *Drugs* 2011;71:981–1001.
- Steele MP, Speer MC, Loyd JE, Brown KK, Herron A, Slifer SH, Burch LH, Wahidi MM, Phillips JA III, Sporn TA, et al. Clinical and pathologic features of familial interstitial pneumonia. *Am J Respir Crit Care Med* 2005;172:1146–1152.
- Mathai SK, Schwartz DA, Warg LA. Genetic susceptibility and pulmonary fibrosis. *Curr Opin Pulm Med* 2014;20:429–435.
- 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.

- 25 Thomas AQ, Lane K, Phillips J III, Prince M, Markin C, Speer M, Schwartz DA, Gaddipati R, Marney A, Johnson J, *et al.* Heterozygosity for a surfactant protein C gene mutation associated with usual interstitial pneumonitis and cellular nonspecific interstitial pneumonitis in one kindred. *Am J Respir Crit Care Med* 2002;165:1322–1328.
- 26 Lawson WE, Grant SW, Ambrosini V, Womble KE, Dawson EP, Lane KB, Markin C, Renzoni E, Lympany P, Thomas AQ, *et al.* Genetic mutations in surfactant protein C are a rare cause of sporadic cases of IPF. *Thorax* 2004;59:977–980.
- 27 Wang Y, Kuan PJ, Xing C, Cronkhite JT, Torres F, Rosenblatt RL, DiMaio JM, Kinch LN, Grishin NV, Garcia CK. Genetic defects in surfactant protein A2 are associated with pulmonary fibrosis and lung cancer. *Am J Hum Genet* 2009;84:52–59.
- 28 Armanios MY, Chen JJ, Cogan JD, Alder JK, Ingersoll RG, Markin C, Lawson WE, Xie M, Vulto I, Phillips JA III, *et al.* Telomerase mutations in families with idiopathic pulmonary fibrosis. *N Engl J Med* 2007;356:1317–1326.
- 29 van Moersel CH, van Oosterhout MF, Barlo NP, de Jong PA, van der Vis JJ, Ruven HJ, van Es HW, van den Bosch JM, Grutters JC. Surfactant protein C mutations are the basis of a significant portion of adult familial pulmonary fibrosis in a dutch cohort. *Am J Respir Crit Care Med* 2010;182:1419–1425.
- 30 Cogan JD, Kropski JA, Zhao M, Mitchell DB, Rives L, Markin C, Garnett ET, Montgomery KH, Mason WR, McKean DF, *et al.*; University of Washington Center for Mendelian Genomics. Rare variants in *RTEL1* are associated with familial interstitial pneumonia. *Am J Respir Crit Care Med* 2015;191:646–655.
- 31 Alder JK, Stanley SE, Wagner CL, Hamilton M, Hanumanthu VS, Armanios M. Exome sequencing identifies mutant *TINF2* in a family with pulmonary fibrosis. *Chest* 2015;147:1361–1368.
- 32 Seibold MA, Wise AL, Speer MC, Steele MP, Brown KK, Loyd JE, Fingerlin TE, Zhang W, Gudmundsson G, Groshong SD, *et al.* A common *MUC5B* promoter polymorphism and pulmonary fibrosis. *N Engl J Med* 2011;364:1503–1512.
- 33 Zhang Y, Noth I, Garcia JG, Kaminski N. A variant in the promoter of *MUC5B* and idiopathic pulmonary fibrosis. *N Engl J Med* 2011;364:1576–1577.
- 34 Stock CJ, Sato H, Fonseca C, Banya WA, Molyneaux PL, Adamali H, Russell AM, Denton CP, Abraham DJ, Hansell DM, *et al.* Mucin 5B promoter polymorphism is associated with idiopathic pulmonary fibrosis but not with development of lung fibrosis in systemic sclerosis or sarcoidosis. *Thorax* 2013;68:436–441.
- 35 Fingerlin TE, Murphy E, Zhang W, Peljto AL, Brown KK, Steele MP, Loyd JE, Cosgrove GP, Lynch D, Groshong S, *et al.* Genome-wide association study identifies multiple susceptibility loci for pulmonary fibrosis. *Nat Genet* 2013;45:613–620.
- 36 Noth I, Zhang Y, Ma S-F, Flores C, Barber M, Huang Y, Broderick SM, Wade MS, Hysi P, Scuirba J, *et al.* Genetic variants associated with idiopathic pulmonary fibrosis susceptibility and mortality: a genome-wide association study. *Lancet Respir Med* 2013;1:309–317.
- 37 Borie R, Crestani B, Dieude P, Nunes H, Allanore Y, Kannengiesser C, Airo P, Matucci-Cerinic M, Wallaert B, Israel-Biet D, *et al.* The *MUC5B* variant is associated with idiopathic pulmonary fibrosis but not with systemic sclerosis interstitial lung disease in the European Caucasian population. *PLoS One* 2013;8:e70621.
- 38 Wei R, Li C, Zhang M, Jones-Hall YL, Myers JL, Noth I, Liu W. Association between *MUC5B* and *TERT* polymorphisms and different interstitial lung disease phenotypes. *Transl Res* 2014;163:494–502.
- 39 Horimasu Y, Ohshimo S, Bonella F, Tanaka S, Ishikawa N, Hattori N, Kohno N, Guzman J, Costabel U. *MUC5B* promoter polymorphism in Japanese patients with idiopathic pulmonary fibrosis. *Respirology* 2015;20:439–444.
- 40 Peljto AL, Steele MP, Fingerlin TE, Hinchcliff ME, Murphy E, Podlasky S, Carns M, Schwarz M, Varga J, Schwartz DA. The pulmonary fibrosis-associated *MUC5B* promoter polymorphism does not influence the development of interstitial pneumonia in systemic sclerosis. *Chest* 2012;142:1584–1588.
- 41 Scholand MB, Wolff R, Crossno PF, Sundar K, Winegar M, Whipple S, Carey P, Sunchild N, Coon H. Severity of cough in idiopathic pulmonary fibrosis is associated with *MUC5B* genotype. *Cough* 2014;10:3.
- 42 Peljto AL, Selman M, Kim DS, Murphy E, Tucker L, Pardo A, Lee JS, Ji W, Schwarz MI, Yang IV, *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.
- 43 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.
- 44 Wang C, Zhuang Y, Guo W, Cao L, Zhang H, Xu L, Fan Y, Zhang D, Wang Y. Mucin 5B promoter polymorphism is associated with susceptibility to interstitial lung diseases in Chinese males. *PLoS One* 2014;9:e104919.
- 45 Hunninghake GM, Hatabu H, Okajima Y, Gao W, Dupuis J, Latourelle JC, Nishino M, Araki T, Zazueta OE, Kurugol S, *et al.* *MUC5B* promoter polymorphism and interstitial lung abnormalities. *N Engl J Med* 2013;368:2192–2200.
- 46 Peljto AL, Zhang Y, Fingerlin TE, Ma SF, Garcia JG, Richards TJ, Silveira LJ, Lindell KO, Steele MP, Loyd JE, *et al.* Association between the *MUC5B* promoter polymorphism and survival in patients with idiopathic pulmonary fibrosis. *JAMA* 2013;309:2232–2239.
- 47 Roy MG, Livraghi-Butrico A, Fletcher AA, McElwee MM, Evans SE, Boerner RM, Alexander SN, Bellinghausen LK, Song AS, Petrova Y, *et al.* *Muc5b* is required for airway defence. *Nature* 2014;505:412–416.
- 48 Molyneaux PL, Cox MJ, Willis-Owen SA, Mallia P, Russell KE, Russell AM, Murphy E, Johnston SL, Schwartz DA, Wells AU, *et al.* The role of bacteria in the pathogenesis and progression of idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med* 2014;190:906–913.
- 49 Siepel A, Bejerano G, Pedersen JS, Hinrichs AS, Hou M, Rosenbloom K, Clawson H, Spieth J, Hillier LW, Richards S, *et al.* Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. *Genome Res* 2005;15:1034–1050.
- 50 Fujisawa T, Chang MM, Velichko S, Thai P, Hung LY, Huang F, Phuong N, Chen Y, Wu R. NF- κ B mediates IL-1 β - and IL-17A-induced *MUC5B* expression in airway epithelial cells. *Am J Respir Cell Mol Biol* 2011;45:246–252.
- 51 Dunham I, Kundaje A, Aldred SF, Collins PJ, Davis CA, Doyle F, Epstein CB, Frieze S, Harrow J, Kaul R, *et al.*; ENCODE Project Consortium. An integrated encyclopedia of DNA elements in the human genome. *Nature* 2012;489:57–74.
- 52 Ernst J, Kellis M. ChromHMM: automating chromatin-state discovery and characterization. *Nat Methods* 2012;9:215–216.
- 53 Helling BA, Steele MP, Brown KK, Loyd JE, Cosgrove GP, Fingerlin TE, Yang IV, Groshong SD, Markin C, Talbert JL, *et al.* A common *MUC5B* promoter polymorphism (rs35705950) and risk of interstitial lung disease [abstract]. *Am J Respir Crit Care Med* 2013;187:A3815.
- 54 Seibold MA, Smith RW, Urbanek C, Groshong SD, Cosgrove GP, Brown KK, Schwarz MI, Schwartz DA, Reynolds SD. The idiopathic pulmonary fibrosis honeycomb cyst contains a mucociliary pseudostratified epithelium. *PLoS One* 2013;8:e58658.
- 55 Yang IV, Coldren CD, Leach SM, Seibold MA, Murphy E, Lin J, Rosen R, Neidermyer AJ, McKean DF, Groshong SD, *et al.* Expression of cilium-associated genes defines novel molecular subtypes of idiopathic pulmonary fibrosis. *Thorax* 2013;68:1114–1121.
- 56 Konishi K, Gibson KF, Lindell KO, Richards TJ, Zhang Y, Dhir R, Biscaglia M, Gilbert S, Yousem SA, Song JW, *et al.* Gene expression profiles of acute exacerbations of idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med* 2009;180:167–175.
- 57 Richards TJ, Kaminski N, Baribaud F, Flavin S, Brodmerkel C, Horowitz D, Li K, Choi J, Vuga LJ, Lindell KO, *et al.* Peripheral blood proteins predict mortality in idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med* 2012;185:67–76.
- 58 Zuo F, Kaminski N, Eugui E, Allard J, Yakhini Z, Ben-Dor A, Lollini L, Morris D, Kim Y, DeLustro B, *et al.* Gene expression analysis reveals matrilysin as a key regulator of pulmonary fibrosis in mice and humans. *Proc Natl Acad Sci USA* 2002;99:6292–6297.
- 59 Gharib SA, Altermeier WA, Van Winkle LS, Plopper CG, Schlesinger SY, Buell CA, Brauer R, Lee V, Parks WC, Chen P. Matrix

- metalloproteinase-7 coordinates airway epithelial injury response and differentiation of ciliated cells. *Am J Respir Cell Mol Biol* 2013; 48:390–396.
- 60 Bingle CD, Araujo B, Wallace WA, Hirani N, Bingle L. What is top of the charts? BPIFB1/LPLUNC1 localises to the bronchiolised epithelium in the honeycomb cysts in UIP. *Thorax* 2013;68:1167–1168.
- 61 Boucher RC. Idiopathic pulmonary fibrosis—a sticky business. *N Engl J Med* 2011;364:1560–1561.
- 62 Williams KJ. Gammaherpesviruses and pulmonary fibrosis: evidence from humans, horses, and rodents. *Vet Pathol* 2014;51:372–384.
- 63 Molyneaux PL, Maher TM. The role of infection in the pathogenesis of idiopathic pulmonary fibrosis. *Eur Respir Rev* 2013;22:376–381.
- 64 Kotton DN, Morrisey EE. Lung regeneration: mechanisms, applications and emerging stem cell populations. *Nat Med* 2014;20:822–832.
- 65 Rock JR, Barkauskas CE, Cronic MJ, Xue Y, Harris JR, Liang J, Noble PW, Hogan BL. Multiple stromal populations contribute to pulmonary fibrosis without evidence for epithelial to mesenchymal transition. *Proc Natl Acad Sci USA* 2011;108:E1475–E1483.
- 66 Kumar PA, Hu Y, Yamamoto Y, Hoe NB, Wei TS, Mu D, Sun Y, Joo LS, Dagher R, Zielonka EM, *et al.* Distal airway stem cells yield alveoli *in vitro* and during lung regeneration following H1N1 influenza infection. *Cell* 2011;147:525–538.
- 67 Vaughan AE, Brumwell AN, Xi Y, Gotts JE, Brownfield DG, Treutlein B, Tan K, Tan V, Liu FC, Looney MR, *et al.* Lineage-negative progenitors mobilize to regenerate lung epithelium after major injury. *Nature* 2015;517:621–625.
- 68 Goetz SC, Anderson KV. The primary cilium: a signalling centre during vertebrate development. *Nat Rev Genet* 2010;11:331–344.
- 69 Jain R, Pan J, Driscoll JA, Wisner JW, Huang T, Gunsten SP, You Y, Brody SL. Temporal relationship between primary and motile ciliogenesis in airway epithelial cells. *Am J Respir Cell Mol Biol* 2010;43:731–739.
- 70 Raghu G, Weycker D, Edelsberg J, Bradford WZ, Oster G. Incidence and prevalence of idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med* 2006;174:810–816.
- 71 Olson AL, Swigris JJ, Lezotte DC, Norris JM, Wilson CG, Brown KK. Mortality from pulmonary fibrosis increased in the United States from 1992 to 2003. *Am J Respir Crit Care Med* 2007;176:277–284.
- 72 Boettger LM, Handsaker RE, Zody MC, McCarroll SA. Structural haplotypes and recent evolution of the human 17q21.31 region. *Nat Genet* 2012;44:881–885.
- 73 Steinberg KM, Antonacci F, Sudmant PH, Kidd JM, Campbell CD, Vives L, Malig M, Scheinfeldt L, Beggs W, Ibrahim M, *et al.* Structural diversity and African origin of the 17q21.31 inversion polymorphism. *Nat Genet* 2012;44:872–880.