A recurring theme in pulmonary fibrosis genetics

Paul J. Wolters

Affiliation: School of Medicine, University of California San Francisco, San Francisco, CA, USA.

Correspondence: Paul J. Wolters, University of California San Francisco, Box 0111, San Francisco, CA 94143-0111, USA. E-mail: paul.wolters@ucsf.edu

In 1980, the question “Is there a fibrotic gene?” was posed in an editorial [1] commenting on a report about pulmonary fibrosis in monozygotic twins. Over the next 37 years, the identification of kindreds of families with pulmonary fibrosis and advances in DNA sequencing provided answers to this question. An early example of a gene associated with pulmonary fibrosis involved a case of familial pulmonary fibrosis (FPF) that was linked to a heterozygous mutation in the surfactant protein C (SFTPC) gene, which is expressed predominantly by alveolar type-II (ATII) cells in the lung [2]. This was followed by reports that mutations in other ATII cell associated genes (ABCA3 and SFTPA) had been identified in patients with FPF. Similarly, candidate gene and whole-genome linkage scanning approaches identified mutations in TERT and TERC, two genes required for telomere maintenance, in kindreds of pulmonary fibrosis patients. These findings were followed by the association of other genes involved in telomere maintenance (DKC1, RTEL1, PARN, TIN2, and NAF1) to FPF. Overall, mutations in these ten ATII cell and telomere associated genes are predicted to account for up to 25% of genes associated with FPF [3, 4], which can present as different clinical syndromes including idiopathic pulmonary fibrosis (IPF), chronic hypersensitivity pneumonitis, pleuroparenchymal fibroelastosis, interstitial pneumonia with autoimmune features, or idiopathic nonspecific interstitial pneumonia [5, 6].

The genetic risk for developing sporadic IPF, an age-associated, progressive fibrotic lung disease characterised histopathologically by the usual interstitial pneumonia (UIP) pattern of lung fibrosis, has also been studied. A large genome-wide association study consisting of 1616 patients with idiopathic interstitial pneumonia, the majority of whom had IPF, linked common variants of several genes (including MUC5B, TERT, TERC, DSP, OBFC1 and DPP9) to the development of IPF [7]. A second genome-wide association study in IPF patients confirmed the association of MUC5B and identified variants to the TOLLIP gene. Similarly, mutations in the ABCA3, SFTPC and SFTPA genes were linked to cases of sporadic IPF [8]. Of these genes, MUC5B has the most robust linkage to IPF and has been confirmed in at least 10 independent studies [9]. Nevertheless, not all individuals with the common variants associated with IPF acquire the disease. This implies additional factors, such as environmental exposure, smoking [10] or aging [11], are also important for disease development.

Patients with rheumatoid arthritis (RA) are at increased risk for pulmonary fibrosis, with up to 40% developing the disease [12, 13]. The predominant radiographic and histopathologic pattern of fibrosis is the UIP pattern and, as with IPF patients, the survival rate of patients with RA-UIP is poor, averaging 3–5 years from diagnosis [14]. These similar histopathologic and epidemiologic features give rise to the
obvious question of whether the molecular mechanisms of pulmonary fibrosis in RA-UIP are similar, or even identical, to those occurring in patients with a UIP pattern of pulmonary fibrosis in other clinical contexts, such as IPF. Considering these relationships, Juge et al. [15] performed whole exome sequencing on 101 patients with RA-associated interstitial lung disease (RA-ILD), of whom 65% had a UIP pattern of disease, and their results are published in this issue of the European Respiratory Journal. Restricting their analysis to nine genes linked to FPF, they found mutations in the TERT, RTEL1, PARN or SFTP C coding regions in 11.9% of patients with RA-ILD. However, no mutations were found in the ABCA3, SFTPA2, TERC, DKC1 or TINF2 genes. Seven of the mutations were novel and, although their functional significance needs to be confirmed by biochemical assay, patients with mutations in the TERT, RTEL1 or PARN genes were also found to have short telomeres in their peripheral blood leukocytes, suggesting that the mutations are biologically relevant.

A pattern has emerged in the genes linked to cases of pulmonary fibrosis, whereby the linked genes are predominantly associated with lung epithelial cells and telomere maintenance. One limitation of the study by Juge et al. [15] is that analysis was restricted to candidates already associated with FPF, potentially missing other relevant genes. Therefore, whether additional variants in genes such as MUC5B, which lies outside the exome, can be linked to RA-ILD will require further analysis. Conversely, their findings are remarkable because they identified relevant mutations in nearly 12% of patients, a high percentage for such a limited analysis. Nonetheless, the highlighted genetic studies address the question posed 37 years ago of whether there is a fibrosis gene. Intriguingly, the answer appears not to be that “there is a fibrosis gene”, but rather that genetic abnormalities in epithelial and telomere-associated genes contribute to the development of fibrosis. These associations strongly suggest that although the clinical syndromes manifest as fibrosis, the diseases origins are epithelial cell and telomere dysfunction.

There are ongoing efforts to determine how the biological effects of rare mutations or common variants contribute to the development of pulmonary fibrosis. Mouse models have been used to test this. For example, modelling telomere dysfunction by isolated deletion of the shelterin proteins capping telomeres in mouse type-II alveolar epithelial cells (AECII) leads to AECII senescence, stem cell failure, lung remodelling and fibrosis, confirming that telomere dysfunction alone can cause lung fibrosis [16–18]. In contrast, modelling SFTP C gene mutations in mice [19] does not lead directly to lung fibrosis. Rather, it sensitises the animal to other profibrotic stimuli such as bleomycin. These findings suggest that some genetic mutations do not directly cause fibrosis, but rather increase an individual’s susceptibility to it. In all modelling of the genetic variants, one element that cannot be replicated in a similar fashion to that in humans is time. Patients with pulmonary fibrosis acquire their genetic risk at birth; however, lung fibrosis typically requires decades to develop. How the passage of time contributes to the development of fibrosis remains unknown. Possible explanations include a lifetime of exposure to environmental insults such as pollution, infections or tobacco smoke, which has been shown to increase the risk for development of FPF, IPF and RA-ILD [10, 20, 21]. In the context of autoimmune diseases such as RA, another possibility is that the pro-inflammatory background of these diseases synergize with the mutant proteins to cause lung fibrosis. It is also conceivable that molecular changes that occur during “normal” biological ageing, such as mitochondrial dysfunction, changes in DNA methylation or loss of proteostasis, synergize with the mutant proteins to cause lung fibrosis [11, 22]. These and other possibilities require further study, either in relevant patient populations or unique models of pulmonary fibrosis that reliably model gene-environment interactions.

An endotype is a disease subtype defined by a distinct functional or molecular mechanism. There has been recent interest in considering whether molecular endotyping can be used to classify cases of pulmonary fibrosis [23]. Although there is scant evidence supporting this approach, the article by Juge et al., reporting mutations in the genes of RA-ILD patients which have been identified in other forms of pulmonary fibrosis (including IPF, hypersensitivity pneumonitis and idiopathic nonspecific interstitial pneumonitis) [5, 6], provides additional evidence that it may be time to reconsider current classification schemes using molecular data such as genetic risk or telomere length. Asthma is an example where endotyping has been used successfully to dichotomise patients into the clinically meaningful TH2 high or TH2 low molecular subsets [24]. However, unlike asthma, which splits a uniform clinical phenotype into two molecular endotypes, the phenotype of pulmonary fibrosis has already been divided into “endotypes” using histopathologic and clinical criteria. Therefore, the molecular endotyping of pulmonary fibrosis patients would require the study of pulmonary fibrosis patients across consensus clinical diagnoses, seeking to identify whether they share common molecular origins. One collection of patients that could be studied as a group, regardless of clinical diagnosis, is those with a UIP pattern of lung fibrosis. Identifying molecular commonalities that extend beyond clinical classification schema, either in total or by histopathologic subsets, may lead to new insights into disease pathogenesis.
Genetic analysis of pulmonary fibrosis has expanded the understanding of risks for the development of these diseases and redirected attention to the importance of epithelial cell and telomere dysfunction as drivers of the disease process. Nevertheless, the mutations and variants identified to date do not fully explain the risk for developing pulmonary fibrosis. More work needs to be done to identify additional genes linked to FPF, IPF, hypersensitivity pneumonitis, or pulmonary fibrosis associated with RA or scleroderma. In the context of these variants, how environmental exposures or comorbidities such as RA contribute to the development of lung fibrosis needs to be investigated. Such studies will inform the debate on whether or not the theme of epithelial cell and telomere dysfunction persists as a central driver of lung fibrosis regardless of the clinical context, and whether or not there are distinct molecular subtypes that span the current classification of disease subsets.

References