The genetic basis of idiopathic pulmonary fibrosis

Jonathan A. Kropski1, Timothy S. Blackwell1,2,3,4 and James E. Loyd1

Affiliations:
1Division of Allergy, Pulmonary and Critical Care Medicine, Dept of Medicine, Vanderbilt University School of Medicine, Nashville, TN, USA.
2Dept of Cell and Developmental Biology, Vanderbilt University School of Medicine, Nashville, TN, USA.
3Dept of Cancer Biology, Vanderbilt University School of Medicine, Nashville, TN, USA.
4Department of Veterans Affairs Medical Center, Nashville, TN, USA.

Correspondence:
Jonathan A. Kropski MD, Division of Allergy, Pulmonary and Critical Care Medicine, Dept of Medicine, Vanderbilt University School of Medicine, 1611 21st Avenue South, T-1218 Medical Center North, Nashville, TN 37232, USA.
Email: jon.kropski@vanderbilt.edu

ABSTRACT Throughout the past decade, there have been substantial advances in understanding the pathogenesis of idiopathic pulmonary fibrosis (IPF). Recently, several large genome-wide association and linkage studies have identified common genetic variants in more than a dozen loci that appear to contribute to IPF risk. In addition, family-based studies have led to the identification of rare genetic variants in genes related to surfactant function and telomere biology, and mechanistic studies suggest pathophysiological derangements associated with these rare genetic variants are also found in sporadic cases of IPF. Current evidence suggests that rather than existing as distinct syndromes, sporadic and familial cases of IPF (familial interstitial pneumonia) probably reflect a continuum of genetic risk. Rapidly evolving bioinformatic and molecular biology techniques, combined with next-generation sequencing technologies, hold great promise for developing a comprehensive, integrated approach to defining the fundamental molecular mechanisms that underlie IPF pathogenesis.
Novel ideas

Through the past decade, rapid advances in genetic and genomic technologies have begun to reshape our understanding of the "idiopathic" interstitial pneumonias. Genome-wide association studies have identified more than a dozen common genetic variants associated with idiopathic pulmonary fibrosis (IPF) risk, and may be linked to altered disease progression and survival. Rare genetic variants in eight genes have been implicated in familial interstitial pneumonia, the familial form of IPF, which broadly fall into two categories: genes related to surfactant protein processing and trafficking, and those linked to telomere biology. In addition to genetic links, unique disease phenotypes based on transcriptomic changes have been identified. As we go forward, we anticipate that advances in these genetic and genomic technologies will result in a re-organisation of the way we define and classify interstitial lung disease based on molecular characterisation. As we evolve from a system of diagnosis based on histopathology to one based on a specific genetic/genomic signature reflecting the fundamental biology of the disease, there will be unique opportunities to develop and test therapies in specific patient populations based on the molecular profiles. Coupled with advances in detection of early disease, the coming decade offers an unprecedented opportunity to dramatically change the lives of patients with IPF.

Introduction

Idiopathic pulmonary fibrosis (IPF), the most common of the idiopathic interstitial pneumonias (IIPs), is characterised by clinical symptoms of cough and dyspnoea, restrictive pulmonary function tests with impaired gas exchange, and progressive lung scarring [1]. Recently, two modestly effective drugs for treating IPF have been identified [2, 3]; however, the prognosis of IPF remains grave, emphasising a need for a more complete understanding of the mechanisms of disease pathogenesis. Available data indicate that both genetic and environmental factors contribute to the risk of IPF and other IIPs [1, 4, 5]. The first insights into IIP genetics came from studies of families with heritable cases of IIP, a syndrome termed familial interstitial pneumonia (FIP). As early as the 1950s, it was recognised that, on occasion, IIP cases clustered in families [6, 7], suggesting a genetic basis to at least a subset of disease. By the 1990s, it was reported that FIP represented a rare subset of IIP, comprising 3–5% of cases [8]. More recently, estimates from several independent groups have suggested that as many as 20% of IIP cases are familial [9–12]. Studies in families have uncovered rare genetic variants in eight genes that are linked to FIP, including three surfactant-related proteins (surfactant protein C, SFTPC [13–16]; surfactant protein A2, SFTPA2 [17]; and ATP-binding cassette member A3, ABCA3 [18, 19]) as well as five genes linked to telomere function (telomerase reverse transcriptase, TERT [20, 21]; human telomerase RNA component, hTR [20, 21]; dyskerin, DKC1 [22–24]; telomere interacting factor 2, TINF2 [25–27]; and regulator of telomere elongation helicase, RTEL1 [28]). Rare genetic variants in FIP-associated genes can be found in some cases of sporadic IPF [12], and investigations into the mechanisms through which these mutated genes contribute to disease have uncovered common underlying pathobiological changes that probably contribute to progressive fibrotic remodelling in FIP and sporadic IPF [4, 5].

Common genetic variants in IPF

Several studies since the early 2000s have investigated the role of functional polymorphisms in a variety of genes in relationship to IPF risk (table 1). Variants in several genes related to inflammation and immune response, including transforming growth factor beta-1 (TGFB1) [44, 45], interleukin-1 receptor alpha (IL1RN) [29–31], interleukin 8 (IL8) [33], toll-like receptor 3 (TLR3) [34], HLA DRB1*1501 [37], as well as cell cycle progression related genes CDKN1A and TP53 [36], have been nominally associated with IPF risk or progression. However, results from these small studies have not yet been validated in independent cohorts. More recently, several large genome-wide linkage and association studies have been completed and have identified numerous additional loci that appear to confer risk for IPF.

MUC5B

In 2011, a genome-wide linkage study identified a locus on chromosome 11 that was significantly associated with IPF risk [38]. Resequencing of this region subsequently identified a common single nucleotide polymorphism (SNP) (rs35705950) in the promoter of the gene encoding mucin 5B (Muc5B) that was associated with a six- to eight-fold increased risk for IPF. The association of this MUC5B promoter polymorphism and IPF has since been confirmed in several independent cohorts, predominantly in Caucasians [32, 39–42, 46]. Interestingly, it appears the MUC5B SNP has a similar frequency in FIP and sporadic IPF cases [38]. This association, however, may be specific to IIP among interstitial lung diseases (ILDs) since reports indicate rs35705950 does not confer increased risk of scleroderma-related ILD or sarcoidosis [39, 41, 47]. This association of rs35705950 with IPF was confirmed in a cohort of Mexican patients [48]; however, rs35705950 was found to be rare in a Korean
cohort of IPF patients. Similarly, in a Chinese population, rs35705950 was rare in IPF patients but different MUC5B polymorphisms were associated with disease [49]. While the rs35705950 MUC5B SNP was associated with increased MUC5B mRNA expression in lungs of control subjects, MUC5B expression was uniformly increased in lungs of IPF patients compared to controls, regardless of whether the MUC5B SNP was present [38]. Consistent with this observation, increased numbers of MUC5B expressing cells have been detected in the distal airways of IPF patients [50]. MUC5B rs35705950 has also been reported as a risk factor for asymptomatic interstitial lung abnormalities detected on computed tomography (CT) scan among subjects over the age of 50 years in the Framingham cohort [51]. Surprisingly, although minor allele carriers of rs35705950 have increased risk of developing disease, IPF patients who carry the minor (risk) allele appear to have improved survival compared to noncarriers [43]. Previous animal studies have suggested that MUC5B regulates airway host defence [52], but the mechanisms by which MUC5B influences fibrotic remodelling are uncertain at present.

**Insights from genome-wide association studies**

A major advance of the past several years has been the development of large, robust datasets with sufficient statistical power for genome-wide association studies. Two large independent genome-wide association studies of IPF patients have now been conducted and identified numerous genetic loci that confer IPF risk. The first, published in 2013 [32], evaluated 1616 IIP cases (the vast majority of which were IPF) and 4683 control subjects with replication in an additional 876 cases and 1890 controls. In addition to confirming the previously reported association with MUC5B, nine additional loci were significantly associated with IIP, predominantly IPF (table 1), including SNPs near TERT and hTR. 10 SNPs on chromosome 11p15 nominally met genome-wide significance, but after controlling for MUC5B rs35705950 these loci no longer met genome-wide significance, suggesting that weak linkage disequilibrium with the MUC5B promoter polymorphism was largely responsible for this association.

Results of a second genome-wide association study again implicated a locus on chromosome 11p15 as significantly associated with IPF [40], but did not replicate other risk loci identified by FINGERLIN et al. [32]. This genome-wide association study performed a three-stage analysis, including a discovery and two replication cohorts, comprising in total 1410 IPF cases and 2934 control individuals. Five loci achieved

<table>
<thead>
<tr>
<th>Locus</th>
<th>Gene</th>
<th>SNP</th>
<th>IPF risk</th>
<th>IPF survival</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>2q14</td>
<td>IL1RN</td>
<td>rs408392</td>
<td>Yes</td>
<td></td>
<td>[29–31]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>rs419598</td>
<td>Yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>rs2637988</td>
<td>Yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3q26</td>
<td>hTR</td>
<td>rs6793295</td>
<td>Yes</td>
<td></td>
<td>[32]</td>
</tr>
<tr>
<td>4q13</td>
<td>IL8</td>
<td>rs4073</td>
<td>Yes</td>
<td></td>
<td>[33]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>rs227307</td>
<td>Yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4q22</td>
<td>FAM13A</td>
<td>rs2609255</td>
<td>Yes</td>
<td></td>
<td>[32]</td>
</tr>
<tr>
<td>4q35</td>
<td>TLR3</td>
<td>rs3775291</td>
<td></td>
<td>Harmful</td>
<td>[34]</td>
</tr>
<tr>
<td>5p15</td>
<td>TERT</td>
<td>rs2736100</td>
<td>Yes</td>
<td></td>
<td>[32, 36]</td>
</tr>
<tr>
<td>6p21</td>
<td>CDKKN1A</td>
<td>rs2395655</td>
<td>Yes</td>
<td>Harmful</td>
<td>[36]</td>
</tr>
<tr>
<td>6p21</td>
<td>HLA-DRB1</td>
<td></td>
<td>Yes</td>
<td></td>
<td>[37]</td>
</tr>
<tr>
<td>6q24</td>
<td>DSP</td>
<td>rs2076295</td>
<td>Yes</td>
<td></td>
<td>[32]</td>
</tr>
<tr>
<td>7q22</td>
<td>Intergenic</td>
<td>rs47274443</td>
<td>Yes</td>
<td></td>
<td>[32]</td>
</tr>
<tr>
<td>10q24</td>
<td>OBFC1</td>
<td>rs1119185</td>
<td>Yes</td>
<td></td>
<td>[32]</td>
</tr>
<tr>
<td>11p15</td>
<td>MUC5B</td>
<td>rs35705950</td>
<td>Yes</td>
<td>Protective</td>
<td>[32, 38–43]</td>
</tr>
<tr>
<td></td>
<td>MUC2</td>
<td>rs37934686</td>
<td>Yes</td>
<td></td>
<td>[32]</td>
</tr>
<tr>
<td></td>
<td>TOLLIP</td>
<td>rs11521887</td>
<td>Yes</td>
<td></td>
<td>[40]</td>
</tr>
<tr>
<td></td>
<td>TOLLIP</td>
<td>rs5743994</td>
<td>Yes</td>
<td></td>
<td>[40]</td>
</tr>
<tr>
<td></td>
<td>TOLLIP</td>
<td>rs2743890</td>
<td>Yes</td>
<td>Protective</td>
<td>[40]</td>
</tr>
<tr>
<td>13q34</td>
<td>ATP11A</td>
<td>rs1278769</td>
<td>Yes</td>
<td></td>
<td>[32]</td>
</tr>
<tr>
<td>14q21</td>
<td>MDG42</td>
<td>rs7144383</td>
<td>Yes</td>
<td></td>
<td>[40]</td>
</tr>
<tr>
<td>15q14–15</td>
<td>Intergenic</td>
<td>rs2034650</td>
<td>Yes</td>
<td></td>
<td>[32]</td>
</tr>
<tr>
<td>17q13</td>
<td>TPS3</td>
<td>rs12951053</td>
<td>No</td>
<td>Harmful</td>
<td>[36]</td>
</tr>
<tr>
<td></td>
<td>TPS3</td>
<td>rs12602273</td>
<td>No</td>
<td>Harmful</td>
<td>[36]</td>
</tr>
<tr>
<td>17q21</td>
<td>MAPT</td>
<td>rs1981997</td>
<td>Yes</td>
<td></td>
<td>[32]</td>
</tr>
<tr>
<td>17q21</td>
<td>SPP1LC2</td>
<td>rs17690703</td>
<td>Yes</td>
<td></td>
<td>[40]</td>
</tr>
<tr>
<td>19q13</td>
<td>DPP9</td>
<td>rs12610495</td>
<td>Yes</td>
<td></td>
<td>[32]</td>
</tr>
<tr>
<td>19q13</td>
<td>T0F81</td>
<td>rs1800470</td>
<td>No</td>
<td>Harmful</td>
<td>[44, 45]</td>
</tr>
</tbody>
</table>

DOI: 10.1183/09031936.00163814
genome-wide significance, including four SNPs on chromosome 11p15 and one on 17q21. Among the 11p15 SNPs were MUC5B rs35705950 and three SNPs within the Toll-interacting protein (TOLLIP) locus. Linkage disequilibrium was reported to be low with rs35705950, suggesting TOLLIP may represent an independent risk locus. Similar to MUC5B rs35705950, IPF cases with the TOLLIP risk allele (the major allele) had decreased mortality compared to minor allele carriers.

Deciphering the biological effects of common genetic variants identified by genome-wide association studies has proven challenging so far. It is possible that the relevant biological effect of most individual SNPs is subtle or manifests only in the context of unique additional genetic or environmental factors to confer disease risk. Despite challenges, future studies are needed to clarify the biological role of disease-associated common genetic variants.

**Rare genetic variants in FIP and IPF**

FIP and sporadic IPF share many clinical and histopathological features [53], which has led to the hypothesis that similar mechanisms underlie the pathogenesis of sporadic and familial disease. Additionally, Scholander et al. [54] employed an extensive genealogical database and found unexpected relatedness among patients who died of what was believed to be sporadic IPF, further supporting the idea that the genetic landscapes of sporadic IPF and FIP overlap considerably. Together, studies to date suggest that sporadic and familial disease reflect a spectrum of genetic risk for pulmonary fibrosis (fig. 1). In this model, genetic risk factors of small and large effects interact with rare and common environmental stimuli to produce the phenotype of pulmonary fibrosis. Most FIP kindreds appear to have an autosomal dominant inheritance pattern with incomplete penetrance, suggesting an important role for genetic rare variants of large effect. In contrast, sporadic IPF may occur more often in the setting of de novo or low penetrance rare variants, or a combination of more common, less severe genetic risk alleles. As described below, in some cases genetic rare variants of large effect may be found in genes that lie within loci containing common variants associated with IPF risk. Focusing on familial disease offers the ability to use Mendelian approaches to identify disease-associated rare variants. This represents a promising approach to enhance mechanistic understanding of the impact of genetic risk factors on development of FIP and potentially sporadic IPF.

**Telomerase and short telomeres**

Pulmonary fibrosis occurs in approximately 20% of patients with dyskeratosis congenita [55], a rare inherited genetic disorder characterised by leukoplakia, bone marrow failure and dystrophic nails that typically affects young males. Rare variants in genes related to telomere biology have been implicated in dyskeratosis congenita [56]. In 2007, using candidate gene approaches, two groups identified heterozygous loss-of-function rare variants in telomere-related genes in 7–15% of FIP families who did not have a history of dyskeratosis congenita [20, 21]. These variants in TERT and hTR lead to short telomeres in peripheral blood and in the lung [20, 21, 57, 58]. To date, TERT rare variants are the most commonly identified mutations linked to FIP; however, TERT rare variants are rarely identified in sporadic cases of IPF [57]. In addition to variants in TERT and hTR, two recent reports identified FIP patients with rare variants in the gene encoding for dyskerin (DKC1) [22, 23], another component of the telomerase complex. Several reports have also identified pulmonary fibrosis in families with dyskeratosis congenita associated with rare variants in TINF2 [26, 27]. In one of these families, there was evidence of somatic or acquired mosaicism for a deletion which abolished expression of the missense variant [25]. This interesting observation suggests acquired genetic variation may represent one mechanism regulating the clinical spectrum of disease linked to telomere pathway rare variants.

Using whole-exome sequencing in a cohort of >180 FIP kindreds, our group has identified heterozygous loss-of-function rare variants in another telomere-related gene, regulator of telomere elongation helicase (RTEL1) in nine families with IPF [28]. Similar to TERT, hTR and DKC1 rare variants, these RTEL1 rare variants are associated with short telomeres in peripheral blood. In addition to a role in telomere...
maintenance, RTEL1 appears to play a more general role in genome stability, DNA repair and replication [59–61], suggesting it may confer disease risk through additional mechanisms. As TERT deficiency has also been associated with abnormal DNA repair [62], it is possible that this mechanism, rather than having a direct effect on telomere length, could be an important in mediating disease risk associated with telomerase pathway rare variants. Cumulatively, rare variants in these five telomere-related genes are reported in approximately 15–20% of FIP families.

Notably, the short telomere phenotype in peripheral blood mononuclear cells (PBMCs) is not limited to patients with loss of function rare variants in telomerase complex genes. Approximately one-third of sporadic IPF and FIP patients have short telomeres (<10th percentile for age) in PBMCs [57, 58]. In addition, it appears that the majority of IPF patients have short telomeres in alveolar epithelial cells [23, 57], suggesting that additional factors (besides genetic risk) contribute to telomere shortening in lungs of patients with IPF and FIP. Interestingly, asymptomatic first-degree relatives of FIP patients have decreased alveolar epithelial cell telomere length compared to controls, and alveolar epithelial cell telomere length is significantly associated with the presence of interstitial changes on high-resolution chest CT [63]. In addition, it appears that PBMC telomere length within families with TERT rare variants can be inherited, at least in part independently of a known rare variant, producing a unique scenario of inherited genetic risk without the risk allele [21, 57, 64]. PBMC telomere length appears to be predictive of survival among patients with IPF, wherein IPF patients with short telomeres have reduced survival compared to those with “normal” length PBMC telomeres [65]. In addition to rare genetic variants in telomere related genes, common genetic variants in loci near TERT, hTR and telomere gene OBFC1 have been linked to sporadic IPF by a genome-wide association study [32] and may be an important factor in determining telomere length. Environmental factors, including cigarette smoke exposure, may also play a role in telomere shortening in FIP and sporadic IPF [66].

Genetic and clinical evidence provide a compelling association between lung fibrosis and telomere biology. Although the mechanisms through which telomerase pathway rare variants lead to lung fibrosis are uncertain, it has been suggested that these loss of function variants disrupt lung epithelial repair mechanisms [67]. Murine models of telomerase dysfunction have been developed but present a number of challenges that limit their utility for mechanistic studies [4]. In spite of these limitations, several studies have reported attempts to model telomerase deficiency in the lung. Tert null mice have decreased numbers of alveolar epithelial cells and modest architectural changes in the lung [68]. These mice have increased susceptibility to cigarette smoke induced-emphysema [69], but do not develop lung fibrosis. Studies using pro-fibrotic stimuli such as bleomycin to investigate fibrotic susceptibility in Tert and Terc null mice have yielded conflicting results [70, 71]. Together, it appears that recapitulating the biology of telomere dysfunction in humans using mouse models is problematic. Therefore, new approaches are needed in this area.

**Surfactant protein-related genes**

Nogee et al. [13] first described a heterozygous mutation in the gene encoding surfactant protein C (SFTPC) in a young woman and her child with IIP in 2001. Soon after, we identified the first association between SFTPC and FIP in a large family with 11 affected individuals [14]. Subsequently, other groups have reported heterozygous rare variants in SFTPC in 1–2% of FIP [9, 12, 14–16, 72]. Although one group reported SFTPC rare variants in 25% of FIP kindreds [15], this high frequency was probably related to founder effects. The mechanisms through which SFTPC rare variants contribute to disease pathogenesis have recently been reviewed elsewhere [4, 73]. In brief, it appears that C-terminal BRICHOS domain mutants result in defects in folding of the propeptide within the endoplasmic reticulum, leading to endoplasmic reticulum stress and activation of the unfolded protein response [74–78]. Linker domain mutants (such as the I73T rare variant) appear to alter trafficking of the pro-peptide [79] and lead to dysregulated proteostasis [80]. Animal modelling suggests induction of endoplasmic reticulum stress in alveolar epithelial cells is not sufficient to induce spontaneous fibrosis but results in an exaggerated fibrotic response following low-dose bleomycin challenge [81]. Induction of endoplasmic reticulum stress in alveolar epithelial cells increases susceptibility to apoptotic stimuli [81, 82], increases expression of mesenchymal markers, and enhances production of profibrotic mediators [83, 84]. In addition, rare variants in another surfactant protein (SFTPA2) [17] has been linked to FIP. SFTPA2 rare variants also result in endoplasmic reticulum stress and may increase latent TGFβ activation [85, 86].

While the frequency of surfactant protein rare variants in sporadic IPF appears to be low [87, 88], several groups have reported that endoplasmic reticulum stress and unfolded protein response activation is a common feature of FIP and sporadic IPF [75, 89], suggesting that environmental factors (such as herpesviruses [75, 90] and tobacco smoke [91–93]) may contribute to this phenotype. Promisingly, it has recently been shown that pharmacological chaperones might improve processing of mutant surfactant
proteins in alveolar epithelial cell lines [94]. These exciting developments raise the possibility of targeted therapies for at least a subset of patients with pulmonary fibrosis.

In addition, rare variants in another gene involved in surfactant processing, ATP-binding cassette-type 3 (ABCA3) (previously linked to paediatric interstitial lung disease) have been reported in several FIP families [18, 19, 95], as well as in sporadic cases of IPF [12]. In one consanguineous family [18], homozygous rare variants in ABCA3 were identified. A heterozygous rare variant in ABCA3 was also reported in a patient with “combined pulmonary fibrosis and emphysema” [19]. In another family carrying the I73T SFTPc rare variant, a second heterozygous rare variant in ABCA3 modified disease penetrance [95]. The exact mechanisms by which ABCA3 variants confer FIP risk are unclear at present, but presumably relate to epithelial cell dysfunction.

ELMOD2

A report using linkage in a cohort of Finnish FIP families suggested ELMOD2 as a candidate FIP gene [96] and in vitro studies suggest ELMOD2 may play a role in antiviral responses [97]. Whether ELMOD2 plays an important role in lung fibrosis has not yet been elucidated.

Missing heritability and future genetic discovery

Cumulatively, available literature suggests that rare variants in FIP genes SFTPC, SFTPA2, ABCA3, TERT, hTR, DKC1, TINF2 and RTEL1 comprise 15–20% of FIP cases (table 2). However, it is possible that selection of families for these candidate-based genetic studies, as well as founder effects in certain populations, may have overestimated the frequency of some variants among all FIP families. While common genetic variants also confer FIP risk and may explain as much as 30% of FIP risk [32], there remains substantial “missing heritability”. In an effort to identify novel FIP genes, recently, our group has performed extensive whole-exome sequencing of subjects from FIP kindreds. While this approach has implicated RTEL1 as an FIP gene, ongoing analysis suggests that rare variants in a single gene (or small group of related genes) do not account for disease in a majority of families [28]. Compared to other genetic lung diseases, such as cystic fibrosis and familial pulmonary arterial hypertension, it appears the genetic basis of FIP is substantially more heterogenous. Given the lack of a dominant gene in FIP, the principal challenge is one of power, both within families and across subjects. Linkage or co-segregation based approaches rely upon the power of large, multigenerational pedigrees with numbers of affected individuals, a well-established inheritance model, and ability to clearly ascertain affection status. Considering that each individual typically harbours 200 rare genetic variants in their exomes [98] and most FIP kindreds are small, dozens of rare variants will be shared among affected individuals, making it difficult to identify the culprit rare variant by this approach. In light of what appears to be substantial allelic heterogeneity, similar to what has been observed for other familial disorders including hypercholesterolaemia [99], neurodegenerative diseases [100] and cardiomyopathies [101], functional testing and validation in cell and/or animal models will be critical to determine the pathogenicity of specific variants and genes tentatively linked to disease.

These challenges suggest that novel and creative approaches will be necessary to further genetic discovery in FIP and IPF. As elaborated further below, we anticipate that evolving bioinformatic approaches to variant prioritisation [102, 103], coupled with network- and pathway-based analysis [104] hold promise for identifying, validating and integrating disease-associated variants. In addition to rare coding variants, future studies investigating intronic variants or variants in more distant cis- and trans- regulatory regions should further inform understanding of the spectrum of genetic risk for FIP.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Reported % of FIP</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>TERT</td>
<td>8–15%</td>
<td>[20, 21]</td>
</tr>
<tr>
<td>RTEL1</td>
<td>5%</td>
<td>[28]</td>
</tr>
<tr>
<td>hTR</td>
<td>&lt;1%</td>
<td>[20, 21]</td>
</tr>
<tr>
<td>DKC1</td>
<td>&lt;1%</td>
<td>[23, 24]</td>
</tr>
<tr>
<td>TINF2</td>
<td>&lt;1%</td>
<td>[25–27]</td>
</tr>
<tr>
<td>SFTPc</td>
<td>2–25%</td>
<td>[13–16, 87, 88]</td>
</tr>
<tr>
<td>SFTPA2</td>
<td>&lt;1%</td>
<td>[17]</td>
</tr>
<tr>
<td>ABCA3</td>
<td>&lt;1%</td>
<td>[18, 19]</td>
</tr>
<tr>
<td>Unknown</td>
<td>75–85%</td>
<td></td>
</tr>
</tbody>
</table>

TABLE 2 Rare genetic variants linked to familial interstitial pneumonia (FIP)
Implications for other IIPs
The evolving understanding of the genetic landscape of FIP and sporadic IPF leads to the question of whether the same genes and genetic variants confer risk for other ILDs, including other forms of IIP. Genetic predisposition for IIPs other than IPF has been poorly characterised; however, several clues suggest there are both conserved and distinct genetic risk factors. Within FIP pedigrees, it has been well recognised that individuals may harbour the same disease-associated rare variant yet present with different histopathologies [13, 53]. This finding indicates that differential environmental exposures overlaid on a common set of genetic risk factors may play a role in determining IIP phenotype. No study to date has been adequately powered to assess whether common genetic variants linked to IPF also confer risk to other IIPs. Future studies are required to address this important issue.

Implications for clinical management
Impact on clinical trial design and treatment
Although no clinical trials to date have stratified patients based on genetic risk, this strategy could prove useful in light of the now recognised associations of the 11p15 risk variants (MUC5B and TOLLIP) with reduced disease progression and improved survival. As additional genes and risk alleles are identified, more complex stratification schemes may be developed to enhance study design. With the recent publication of randomised controlled trials of two pharmacological agents, pirfenidone [2] and nintedanib [3], that reduce lung function decline among IPF patients, a reasonable question for future study is whether genetic factors influence response to treatment with one or both of these medications.

Whether genetic factors influence outcomes after lung transplant remains an underexplored question. One small series described successful lung transplant in eight patients with TERT rare variants, but a high rate of haematological and renal toxicities were noted [105]. The role of other rare or common genetic variants on outcome after lung transplant is not known.

Genetic testing
As our understanding of the influence of genetic factors on risk of IPF and its natural history, a potential role for clinical genetic testing emerges. At present, we suggest that evidence is not sufficient to recommend routine genetic testing for rare or common genetic variants for patients with sporadic IPF. In families with FIP, clinical testing for variants in SFTPC, SFTPA2, and telomerase-related genes including TERT, hTR, DKCI and RTEL1 is available from commercial and academic sources. Our practice is to offer genetic counselling and consideration of genetic testing to patients with FIP and a family history suggestive of a telomerase dysfunction syndrome (including diagnoses of aplastic anaemia, cryptogenic cirrhosis or premature graying). Decisions to undergo genetic testing are complex and highly individual, and no studies have evaluated the impact of genetic testing on patients/families in the context of FIP. Extrapolating from our experience and literature from other disorders [106–108], continued close follow-up of patients who undergo genetic testing appears essential regardless of whether they are found to carry a disease-associated rare variant. In light of the incomplete penetrance and variable expressivity of FIP-associated rare variants, we recommend patients and their families confer with genetic counsellors before consideration of genetic tests that are currently available or may become available in the future. Over time, we anticipate there may be a role for broader screening of common and rare genetic variants associated with IPF.

Overcoming challenges and future directions
To date, investigations of the genetic basis of FIP and sporadic IPF have provided crucial insights into underlying mechanisms of progressive pulmonary fibrosis. We suggest that there is probably not a single “road to IPF” but rather there are “multiple paths” that converge to a common phenotype. The development of exciting next-generation sequencing capabilities, along with continuously evolving bioinformatic approaches to analysis of large datasets and rapidly improving molecular biological techniques, offer unique possibilities to identify additional novel genes and pathways that contribute to the pathogenesis of progressive pulmonary fibrosis (fig. 2). Whole-genome sequencing is rapidly becoming feasible, and although it presents new and greater bioinformatic challenges, the possibility of analysing noncoding variants, as well as interactions among variants, holds promise in identifying yet unexplained heritability of FIP and sporadic IPF. In addition, the rapidly evolving field of stem cell biology offers the possibility of studying the effects of genetic variants in primary human cell types of interest using inducible pluripotent stem cell differentiation strategies [109, 110]. The development of improved gene editing technologies, including the CRISPR-Cas9-based system, should facilitate enhanced ability to characterise genetic variants with functional testing in vitro and in vivo [111, 112].

We anticipate that by understanding the biological mechanisms through which individual genetic variants contribute to disease pathogenesis, key pathways will be identified that will clarify the crucial molecular mediators of IPF pathogenesis. We anticipate that a role for molecular genetics in the classification of IIPs...
Genetic determinants of pulmonary fibrosis: evolving concepts.


DOI: 10.1183/09031936.00163814


DOI: 10.1183/09031936.00163814