R E V I E W

Candidate genes of idiopathic pulmonary fibrosis: current evidence and research

Wei Zhou[1,2]
Yaping Wang[1,2]

[1]Department of Medical Genetics, [2]Jiangsu Key Laboratory of Molecular Medicine, Nanjing University School of Medicine, Nanjing, People's Republic of China

Abstract: Idiopathic pulmonary fibrosis (IPF) is a group of common and lethal forms of idiopathic interstitial pulmonary disease. IPF is characterized by a progressive decline in lung function with a median survival of 2–3 years after diagnosis. Although the pathogenesis of the disease remains unknown, genetic predisposition could play a causal role in IPF. A set of genes have been identified as candidate genes of IPF in the past 20 years. However, the recent technological advances that allow for the analysis of millions of polymorphisms in different subjects have deepened the understanding of the genetic complexity of IPF susceptibility. Genome-wide association studies and whole-genome sequencing continue to reveal the genetic loci associated with IPF risk. In this review, we describe candidate genes on the basis of their functions and aim to gain a better understanding of the genetic basis of IPF. The discovered candidate genes may help to clarify pivotal aspects in the diagnosis, prognosis, and therapies of IPF.

Keywords: idiopathic pulmonary fibrosis, candidate genes, susceptibility

Introduction

Idiopathic pulmonary fibrosis (IPF) is a chronic lung disease characterized by progressively worsening dyspnea and pulmonary function. Although published data on the epidemiology of IPF in the world are limited, the majority of studies have shown an increase in its incidence and prevalence during recent decades. A systematic review assessed all the available population-based studies from 1968 to 2012 and estimated a conservative incidence range of three to nine cases per 100,000 per year in Europe and North America,[1] whereas the incidences were lower in East Asia[2–4] and South America.[1] Patients with IPF exhibited variable disease courses and worse prognosis and had a median survival of 2–3 years after the initial diagnosis.[5] In 2011, the American Thoracic Society/European Respiratory Society panel of experts issued a new classification scheme for IPF. It was defined as a specific form of chronic, progressive fibrosing interstitial pneumonia of unknown cause, occurring mainly in older adults and associated with the histopathological and/or radiological pattern of usual interstitial pneumonia.[6] During recent decades, although there have been significant advances in the understanding of the pathogenesis of IPF, the prognosis of the disease was barely improved.[7] Lung transplantation is still the only treatment option that successfully increases the survival.[8]

Since Haschek et al[9] proposed the "alveolar epithelial disrepair" hypothesis to challenge the traditional "inflammatory theory", the pathogenesis of lung fibrosis has remained controversial. The point of conflict was whether the alveolar impairment was caused by an inflammation reaction or by an abnormal repair. Although the exact

Correspondence: Yaping Wang
Department of Medical Genetics, Nanjing University School of Medicine, Nanjing 210093, People's Republic of China
Tel +86 25 8368 6495
Fax +86 25 8368 6451
Email wangyap@nju.edu.cn

etiology of pulmonary fibrosis is uncertain, the process of fibrosis is likely to involve an interaction between exogenous and endogenous.[10] Cigarette smoking is the most important environmental risk factor associated with IPF. In a multicenter case–control study, a higher proportion of IPF patients had a history of smoking compared with controls.[11] Exposure to several environment/occupational agents, particularly metal, wood dust and farming, or stone cutting/polishing, has also been observed to increase the risk of IPF.[12,13] It is also worth noting that the comorbid conditions, such as gastroesophageal reflux,[14] pulmonary hypertension,[15] emphysema,[16] viral or bacterial lung infections,[17] and radiotherapy, may be the risk factors for IPF.

Studies have also demonstrated the genetic basis for IPF. The most persuasive evidence is a familial clustering of the disease, familial pulmonary fibrosis (FPF), which accounts for <5% of all cases.[18] More recent data suggested that in up to 20% of IPF patients, the FPF could be underestimated.[19] FPF is clinically and histologically indistinguishable from sporadic IPF except that the age of onset tends to be earlier. Subsequently, several gene variants and dysfunctional proteins were detected to be associated with the susceptibility of FPF. Three alveolar stability-associated genes (surfactant protein C, *SFTPC*; surfactant protein A2, *SFTPA2*; and ATP-binding cassette member A3, *ABCA3*) were detected to concentrate in FPF. As the incidence and prevalence of IPF increased with age, the disease was considered as an age-related condition with the distinctive features of degenerative disorders, such as telomerase dysfunction and shortening.[20] Approximately 10% of FPF and 1%–3% of sporadic IPF patients carry variations in two major telomerase components: telomerase reverse transcriptase (*TERT*) and telomerase RNA component (*TERC*). Recent research has shown that dyskerin, dyskeratosis congenita 1 (DKC1), and regulator of telomere elongation helicase 1 (RTEL1) are also involved in the telomere shortening and may deteriorates the survival in IPF and supported the hypothesis that abnormal telomeres were involved in the disease pathogenesis and outcome.[21] In 2011, a common variant (rs35705950) in the promoter region of *MUC5B* was detected to be strongly associated with both FPF and sporadic IPF. Sporadic reports showed that variants in multiple genes related to immunity and inflammation (*TOLLIP, ELMOD2*, major histocompatibility complex [MHC], interleukins [ILs], tumor necrosis factor [TNF], and transforming growth factor β [TGF-β]) can be found in some cases of IPF patients.

In this review, we present the current evidence and research on these genes that may contribute to the disease in order to uncover the underlying pathophysiological changes in FPF and sporadic IPF.

The genes associated with alveolar stability
SFTPC

SFTPC has been widely reported to be associated with various interstitial lung diseases (ILDs), including IPF. The *SFTPC* gene is located on the short arm of chromosome 8 (gene ID: 6440) with six exons encoding a 197 amino acid precursor protein. In 2001, Nogee et al reported the first case who was an infant with a family history of ILD and carried a heterozygous substitution of A for G in the splice donor site of intron 4.[22] This nucleotide substitution resulted in the skipping of exon 4 and a 37 amino acid deletion in the C-terminal domain of the precursor protein. As a result, the mature surfactant protein C (SP-C) was absent in the lung tissue and bronchoalveolar lavage fluid. Later, two other heterozygous mutations, Leu188 to Gln (L188Q) and Ile73 to Thr (I73T), were identified in *SFTPC* gene.[23,24] These two mutations may result in abnormal intracellular trafficking of protein and an accumulation of aberrantly processed prosurfactant protein C (pro-SP-C) within alveoli.

Lawson et al showed ten single-nucleotide polymorphisms (SNPs) in the *SFTPC* sequence in 89 IPF patients but not in 104 controls.[25] These mutations were mainly detected in FPF. Markart et al reported two nonsynonymous variants in the *SFTPC* gene in a group of sporadic IPF cases, but their allele frequencies were similar to that of controls.[26] Subsequently, at least 40 pathogenic mutations in the human *SFTPC* gene have been discovered through genetic analyses on the DNA samples from infants, children, and adult patients with ILDs. Compared with FPF, the mutations of *SFTPC* gene are not frequently detected in sporadic cases of IPF. The association observed between the mutations and the ILD led several groups to investigate its molecular mechanism.

The pro-SP-C encoded by the *SFTPC* would undergo several rounds of proteolytic cleavage to form the mature 35-residue hydrophobic protein (SP-C) and then secrete to the alveolar space. The SP-C, together with surfactant phospholipids and other surfactant proteins, is essential for maintaining alveolar stability.[27] The abnormal SP-C proteins resulted from the *SFTPC* mutants can induce endoplasmic reticulum (ER) stress, inhibit the ubiquitin/proteasome system, and activate the apoptotic pathways.[28] The cultured type II alveolar epithelial cells (AECII) transfected with *SFTPC^{L188Q}* caused the development of abnormal lamellar bodies, ER stress, and unfold protein response.[29]

Some fatal mutations have been found to be located in the BRICHOS domain of pro-SP-C, a special domain detected in 12 protein families with a wide range of functions and disease associations.[30] For the pro-SP-C, the BRICHOS domain acts as a chaperone for SP-C during biosynthesis. It plays a pivotal role in the maturation of pro-SP-C, proper folding process and correct packaging, and eventually enters the lamellar bodies with other surfactant components for exocytosis. This process may explain why SP-C mutations can result in pulmonary fibrosis that occurs as a consequence of unfolded pro-SP-C but not due to deficiency, which is different from infant/adult respiratory distress. Epithelial to mesenchymal transition (EMT) has also been proven to occur in the process of collagen deposition in the alveolus wall and pulmonary fibrosis. Targeted expression of *SFTPC*[L188Q] in A549 cells resulted in decreased expression of E-cadherin and zonulaoccluden-I and increased expression of the mesenchymal marker α-smooth muscle actin, showing the relationship between the variants of *SFTPC* and EMT.[31]

SP-C, together with SP-A, SP-B, and SP-D, constitutes 10% of the protein components of pulmonary surfactant in humans, and the remaining 90% is lipids. The main function of the surfactant is to reduce the surface tension at the air–water interface and prevent alveolar collapse. Because the surfactant alterations and alveolar type II cell apoptosis had been implicated in the pathogenesis of IPF, the hereditary and acquired disorders associated with the surfactants attracted more attention in the etiology study of IPF.

SFTPA2

The identification of the *SFTPC* mutation led to a search for a susceptible gene in encoding the other surfactant proteins in IPF. Wang et al[32] used linkage analysis to map a susceptible locus in the 15.7 Mb region of chromosome 10q22 in a four-generation family with IPF and identified a transversion mutation (GGG to GTG) in codon 231 of one *SFTPA2* allele that encoded SP-A (SP-A2). This mutation is predicted to change a highly conserved glycine residue to a valine (G231V).[32] Subsequently, a heterozygous transversion mutation (TTC to TCC) in codon 198 that substituted a serine residue for a phenylalanine (F198S) in the same gene was found in a 45-year-old man with IPF and lung cancer in another family. Both of the mutations, G213V and F183S, occurred in the highly conserved carbohydrate-recognition domain of the protein and may disrupt cellular trafficking of the protein, probably by interfering with its proper folding.[32]

Unlike SP-C, SP-A is a large multimeric protein that belongs to a family of collectin proteins characterized by an NH2-terminal collagen-like region and a C-terminal lectin domain. Because it is homologous to collectin, SP-A has always been thought to play a role in the innate defense system and pulmonary immune response. It may interact with pathogen-associated molecular patterns anchored on membranes via the calcium-dependent carbohydrate-binding domains.[33] SP-A has been associated with SP-D in immune cells, in which it activates various cellular functions and acts as an opsonin in the modulation of macrophage phagocytosis and the production of reactive oxygen species.[34] Experimental data showed that the expression of mutant proteins increased the transcription of a BiP-reporter construct, expression of BiP protein, and production of an ER stress-induced XBP-1-spliced product.[35] The patients with a mutant protein of SP-A2 showed a higher level of TGF-β1 than that in the control subjects.[36] This finding suggested that the function lesion of SP-A2 may cause the disease through a change in the secretion of TGF-β, a cytokine known as a fibrogenic mediator in epithelial cells.

Although the pathogenic mechanism of the gene mutation associated with pulmonary fibrosis remains to be clarified, the linkage analysis and the candidate gene approaches indicated that the *SFTPA2* may be a susceptible gene in IPF.

ABCA3

ABCA3 has been identified as a neonatal respiratory distress syndrome or ILD-associated protein in childhood. It is a transporter protein predominantly expressed in the AECII and is involved in the transport of lipids across plasma membranes. The *ABCA3* gene encodes a 1,704 amino acid protein located on the limiting membrane of lamellar bodies, the secretory vesicle for surfactant, suggesting that it may play a role in surfactant metabolism and transport. There is a wide spectrum of variants in the *ABCA3* gene, and >150 distinct mutations have been found to cause genetic abnormalities in surfactant metabolism, which would lead to neonatal respiratory distress and pediatric ILD.[37–39] The lung disorder caused by *ABCA3* mutations occurs in an autosomal recessive pattern, and the compound heterozygote or homozygous mutations could be detected in both alleles of the patient. *ABCA3* may also act as a modifier of the phenotype associated with *SFTPC* mutation, as the heterozygous *ABCA3* mutation was found in severely affected infants with *SFTPC*[I73T], a commonly identified substitution of isoleucine by threonine in codon 73.[40] A sequence analysis in a full-term female baby who developed severe chronic lung disease revealed that she was a homozygous carrier of a mutation p.Trp308Arg (c.922T>C) at exon 9 of the *ABCA3* gene but without any

mutation in *SFTPC*.[41] A novel homozygous G>A transition at nucleotide 2891 in exon 21 of the *ABCA3* gene, resulting in G964D, was found in a large kindred of pulmonary fibrosis.[38] Kumar et al reported an IPF patient who carried a compound heterozygote for *ABCA3*, E292V, and G964D, further supporting the possibility that adult-onset fibrotic lung disease may be due to the mutations in *ABCA3*.[42] In our unreported study, a novel variant in exon 24 of *ABCA3* may be a risk factor for IPF in the Chinese population.

The genes associated with maintaining of telomere length
TERT and TERC

The *TERT* gene located at chromosome 5p15.33 (gene ID: 7015) encodes the catalytic component of telomerase, and the *TERC* at chromosome 3q26.2 (gene ID: 7012) is an RNA gene and acts as a component of telomerase to maintain telomere. In germ cells and other stem cells, the integrity of the telomere is maintained by telomerase complex, containing the following two key components: the TERT and the TERC. The telomerase complex uses the template encoded by *TERC* to add telomere repeats (TTAGGG) to the end of the chromosomes, which works as a "cap" to offset the telomere shortening with cell divisions. Evidence has indicated a relationship between chronological age and the incidence of IPF.[20] Similarly, telomere length could limit tissue renewal capacity, which was implicated in age-related disease. Indeed, IPF is one of the most prevalent clinical manifestations of human telomere syndromes, which have defined pathological ranges of telomere shortening.[43]

Recent research showed that short telomere length was detected in both the peripheral blood and the lung tissue cells of IPF patients.[44] Moreover, in bleomycin-induced lung fibrosis mice, telomerase mutation or telomere shortening caused emphysematous lesions.[45] Overall, the germ-line mutations in telomerase components, *TERT* and *TERC*, underlie 8%–15% of FPF and 1%–3% of sporadic pulmonary fibrosis patients.[46] In 2007, Armanios et al identified five heterozygous mutations in *TERT* and one mutation in *TERC* in six families, and none of these mutations were found in controls.[47] Later, a genome-wide association study (GWAS) with 159 IPF patients and 934 controls identified a susceptibility SNP (rs2736100) in intron 2 of the *TERT* gene in a Japanese population.[48] This variant (rs2736100) is located within a linkage disequilibrium block in *TERT*, suggesting that it could be associated with some other unidentified variations in IPF.

Polymorphisms in *TERT* and *TERC* might contribute to telomere attrition, which was reported as a hallmark of aging.[49] Increased mutations affecting telomere function have been reported in IPF patients, but the underlying mechanism for its pathogenesis still remains unclear. Alveolar stem cell senescence may be a driver of the pathogenesis. Evidence has shown that damaged telomere maintenance was related to epithelial cell senescence in lung injury response, which results in impaired lung epithelial homeostasis. Telomere dysfunction in AECs preferentially upregulated the DNA damage response, activated the p53 pathway, and induced cell senescence.[50] Another intriguing target of TERT regulation is the Wnt-β-catenin pathway. In the lung tissues of IPF patients, the Wnt-β-catenin pathway is activated, which correlated with increased expression of (TGF-β) to promote EMT.[51] Further investigations are needed to confirm the role of telomere maintenance and cellular senescence in lung fibrosis and elucidate the association between the aforementioned pathways in lung injury response and fibrotic remodeling.

DKC1 and RTEL1

Up to 25% of IPF patients and 40% of FPF patients exhibited the shortened telomeres, but less than half of these patients had the germline mutation in the *TERT* or *TERC* gene. The low frequency of these events suggested that there may be other genes responsible for the telomere length. The *DKC1* gene that encodes a component of the telomerase complex was reported to cause autosomal dominant dyskeratosis congenita (DC), a disorder in which 20% of affected individuals develop pulmonary fibrosis. A transversion mutation (c.145A>T) in *DKC1* exon 3, which predicted a missense substitution (p.Thr49Ser) near the nuclear localization signal of dyskerin, was identified in a IPF patient, but no features of DC developed in this patient.[52] Recently, Kropski et al screened a kindred of FPF with short telomeres and found a 69-year-old male patient carrying a novel variant, c.1213A>G, in the *DKC1* gene. No mutation was identified in those aforementioned FPF-associated genes (*TERT* and *TERC*) in this patient. These results suggested that the *DKC1* may be the third telomerase-related gene in the etiology of FPF.[53]

Using next-generation sequencing technologies, Stuart et al recently revealed two genes that were also associated with telomere length.[54] One was the *PARN* gene, which encoded an exoribonuclease. Five damaging mutations of PARN in heterozygous were identified in 78 European FPF patients but not in 2,816 controls (P=1.3×10^{-8}). The other was *RTEL1* gene, an established locus for DC, which harbored more novel damaging and missense variants at conserved residues in patients than in controls (P=1.6×10^{-6}). Four highly

detrimental mutations were identified in a group of French FPF patients using whole-exome sequencing analysis in this gene. Three of the four mutations were missense mutations (p.Thr49Met, p.Arg213Trp, and p.Phe964Leu). The fourth was one base pair insertion in exon 33 (c.3493dupC, pGln1165Profs*22), leading to a frameshift and a premature stop codon.[55] A function prediction showed that these variants may seriously modify the functional domains of the protein.

RTEL1 is a member of the subclass of iron–sulfur cluster-containing DNA helicases that participate in telomere maintenance, and shortened telomeres were detected in the cells from the patients with the *RTEL1* mutation. Other researchers found that telomere length in peripheral blood mononuclear cells from patients with *RTEL1* mutations was extremely short and similar to the peripheral blood mononuclear cell telomere length in other patients with known telomerase pathway mutations (*TERT*, *TETC*, and *DKC1*), suggesting that the rare loss-of-function variants in *RTEL1* represented a genetic risk of FPF.[56] These findings indicated that the *RTEL1*, as a novel candidate gene, was associated with IPF and confirmed the important role of telomerase dysfunction and shortened telomeres in the pathogenesis of pulmonary fibrosis.

Common variation and IPF risk
MUC5B

A genome-wide linkage screen in 82 multiplex families with familial interstitial pneumonia (FIP) demonstrated an association between idiopathic interstitial pneumonia and the 3.4 Mb region of chromosome 11p15 where the maximum multipoint LOD score was 3.3 (P=0.00004).[57] Consequently, fine mapping was performed by genotyping 306 tagging SNPs in the mucin gene cluster region on the p-terminus of chromosome 11 in the unrelated FIP subjects, IPF subjects, and healthy controls. Nineteen of them were detected to be associated with FIP and/or IPF.

Among these SNPs, the variant (rs35705950) at 3 kb upstream of the *MUC5B* transcription start site, in the putative promoter, had the most substantial effect on pulmonary fibrosis risk with a frequency of 38% among subjects with IPF, 34% among those with FIP, and 9% in controls. *MUC5B* (gene ID: 727897) is an evolutionarily conserved gene that encodes mucin glycoprotein, a principal macromolecule in airway mucus. Interestingly, the expression of the MUC5B protein in the lung tissues of IPF patients was ~14-fold higher than those who did not.[57] This variant of rs35705950 was associated with a 37.4-fold increase in MUC5B protein

expression in the lung tissues of unaffected subjects who were heterozygous/homozygous compared with the wild-type allele.[57] The minor allele of the promoter region exhibited a higher frequency in both familial and sporadic IPF patients. Therefore, this variant is currently the most common genetic predisposing humans to IPF. However, our previous study showed that the promoter variant rs35705950 was rare in the Chinese population, although an association was observed between this variant and a risk of ILD.[58] Peljto et al reported that the *MUC5B* promoter variant rs35705950 was associated with IPF in a Mexican population but not in Koreans.[59]

Recent evidence showed that patients with pulmonary fibrosis had positive MUC5B staining in the cells that were morphologically consistent with AECs, which normally do not produce mucins.[60] The mechanism by which the expression of MUC5B increases and it favors the development of IPF remains uncertain. Based on the relationship between this variant and excess production of MUC5B, the increased MUC5B expression may impair mucus clearance and alveolar repair and further interfere with host defense.[61] As a risk factor for this complex disease, this typically common variation may have potential predictive value for the development of this disease in a population. Although the previous data provide strong evidence for an association between the polymorphism of *MUC5B* gene and IPF, the molecular consequences of the variation and the contribution of the increased MUC5B protein expression contributes to the development of lung fibrosis require further explanation.

The genes related to immunity and inflammation
Toll-like receptor-associated genes

Despite the controversy on corticosteroids and immune suppressants, the innate and adaptive immune responses were thought to contribute to the pathogenesis of lung fibrosis. Toll-like receptors (TLRs), as a family of pattern recognition receptors, mainly express by immune-related cells. TLR3 specifically recognizes double-stranded RNA and induces multiple intracellular events responsible for innate antiviral immunity.[62] The *TLR3* polymorphism (L412F) that would defect the function of TLR3 was found in patients with IPF.[63] The further study showed that the lung fibroblasts from IPF patient with the variant L412F in TLR3 could reduce cytokine responses and deregulate fibroproliferation when compared with the wild genotype.[63]

TOLLIP is a ubiquitin-binding protein that interacts with several components of the TLR signaling cascade. It is an important regulator of innate immune responses mediated

by TLR and the TGF-β signaling pathway. In 2013, Noth et al[64] identified three SNPs, rs111521887, rs5743894, and rs5743890, in *TOLLIP* with different frequencies between IPF patients and controls through a three-stage GWAS on European Americans. The minor alleles of the three SNPs had effects in the same directions in all stages: two (rs111521887 and rs5743894) were associated with IPF risk, and one (rs5743890), a novel variant of *TOLLIP*, with protection from IPF of them. Notably, the minor allele of rs5743890 was also associated with mortality of IPF. And the reduced expression of TOLLIP in patient with IPF who carried *TOLLIP* SNPs emphasizes the importance of this gene in this disease.[64]

Hodgson et al[65] performed a genome-wide scan in six multiplex families with FPF and further hierarchical fine mapping with 24 families from southeastern Finland. They revealed a shared 110 kb to 13 Mb haplotype on chromosome 4q31, embodying two functionally uncharacterized genes, *ELMOD2* and *LOC152586*, of which only ELMOD2 is expressed in alveolar macrophages and AECII. The results suggested that *ELMOD2* could be included as a novel candidate gene for FPF. Pulkkinen et al used overexpression cell model to check genes networked with *ELMOD2* and found that TLR3 pathway is dependent on ELMOD2.[66] However, sequencing of *ELMOD2* gene failed to reveal any mutation in FPF patients. *ELMOD2*, as a candidate gene for IPF, remains further investigation.[65]

Major histocompatibility complex

The MHC is a set of genes that encode cell surface molecules that controls a major part of the immune system by determining histocompatibility. Several studies showed that the genetic variations within MHC loci contributed to the susceptibility to IPF.[67] Aquino-Galvez et al reported a cohort study with 80 sporadic IPF patients and 201 health controls and found a significant association between MHC class I chain-related gene A allele 1(MICA*001) and IPF.[68] The HLA class I and class II were further investigated by Xue et al, who found an increased prevalence of DRB1*1501 in IPF patients when compared to normal subjects.[69] Recently, three SNPs HSPA1B (rs1061581), HSPA1L (rs2227956), and HSPA1 (rs1043618) in HSP70 gene were indicated associated with a decreased risk of IPF. The MHC locus could represent a specific chromosomal area for genetic studies aimed at the identification of the various clinical IPF phenotypes.

Proinflammatory cytokines

The pathogenesis of IPF is complex and remains to be clarified. The proinflammatory cytokines may play an important role in the initiation and development of pulmonary fibrosis. Vasakova et al investigated the cytokines polymorphisms of the IL-1α, IL-1β, IL-1R, IL-1RA, IL-2, IL-4, IL-4RA, IL-6, IL-10, IL-12, TNF-α, IFN-γ, and TGF-β genes among 30 IPF patients and 103 healthy volunteers and found the CT genotypes at the promoter regions, the positions −590 and −33, of IL-4 gene was significantly associated with IPF susceptibility.[70] Ahn et al demonstrated that IL-8 rs4073 T allele is risk factor for IPF in the Korean population.[71] However, Alhamad et al[72] examined the allele distributions of polymorphisms in TNF-α, IFN-γ, IL-6, IL-10, and TGF-β1 and the serum levels of cytokines in 60 patients with IPF and 150 healthy controls. Their results indicated that the genotypes of the polymorphisms investigated in the cytokine genes did not predispose to the susceptibility of IPF but played an important role in disease severity.

The other candidate genes related to IPF *DSP and DPP9*

The GWAS by Fingerlin et al identified multiple susceptibility loci for pulmonary fibrosis, including two cell–cell adhesion molecules, *DSP* and *DPP9*, which suggested that defects in cell adhesion or the cellular cytoskeleton could predispose to injury in response to stimuli.[73] *DSP* encodes desmoplakin, a component of the desmosome that tightly links adjacent cells and forms a dynamic structure with other proteins that can bound the cytoskeleton to the cell membrane. The DPP9 is a member of the same protein family as fibroblast activation protein, which was shown to express in fibroblastic foci and in fibrotic interstitium but not in adjacent healthy lung of IPF patients. These results suggested that pulmonary fibrosis might be caused by defects in cell–cell adhesion or the cytoskeleton and *DSP* and *DPP9* could be considered as the candidate genes for IPF.

PINK1

Some age-related neurodegenerative disease has an etiology that is related to mitochondrial dysfunction. Bueno et al recently showed that AECII in the lungs of IPF patients exhibits marked accumulation of dysmorphic and dysfunctional mitochondria.[74] Their results also indicated that impaired mitochondria in IPF and aging lungs were associated with low expression of PTEN-induced putative kinase 1 (PINK1). Patel et al[75] found that TGF-β could induce lung epithelial cell mitochondrial ROS and depolarization and stabilizes PINK1 protein. PINK1 can ameliorate epithelial cell death and may be necessary to limit fibrogenesis. This evidence indicated that mitochondrial dysfunction may have a role in the pathogenesis of pulmonary fibrosis, and

mitochondrial-associated genes could be recognized as a novel genetic susceptibility loci.

Conclusion

In recent decades, tremendous progress has been made in understanding the genetic predisposition of IPF. Genetic association studies play a pivotal role in demonstrating the genetic cause of IPF and revealing the underlying mechanisms of pulmonary fibrosis. The gene variants mentioned in our review only identified in part of IPF patients, and the various pathological phenotypes presented in patients indicated a complex genetic basis of IPF. A increased list of genes involved in IPF pathogenesis would be obtained with high-throughput analysis and next-generation sequencing in the future. Moreover, more effort should be concentrated on investigating the roles of genetic variants and interaction between the genetic background and environmental factors. Elucidation of the genetic basis and pathogenesis of this disease may produce radical changes in the diagnosis and management of IPF.

Acknowledgment

Disclosure

The authors report no conflicts of interest in this work.

References

1. Hutchinson J, Fogarty A, Hubbard R, McKeever T. Global incidence and mortality of idiopathic pulmonary fibrosis: a systematic review. *Eur Respir J*. 2015;46(3):795–806.
2. Lai CC, Wang CY, Lu HM, et al. Idiopathic pulmonary fibrosis in Taiwan – a population-based study. *Respir Med*. 2012;106(11):1566–1574.
3. Gjonbrataj J, Choi WI, Bahn YE, Rho BH, Lee JJ, Lee CW. Incidence of idiopathic pulmonary fibrosis in Korea based on the 2011 ATS/ERS/JRS/ALAT statement. *Int J Tuberc Lung Dis*. 2015;19(6):742–746.
4. Ohno S, Nakaya T, Bando M, Sugiyama Y. Idiopathic pulmonary fibrosis – results from a Japanese nationwide epidemiological survey using individual clinical records. *Respirology*. 2008;13(6):926–928.
5. Ley B, Collard HR, King TE. Clinical course and prediction of survival in idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med*. 2011;183(4):431–440.
6. Raghu G, Collard HR, Egan JJ, et al; ATS/ERS/JRS/ALAT Committee on Idiopathic Pulmonary Fibrosis. An official ATS/ERS/JRS/ALAT statement: idiopathic pulmonary fibrosis: evidence-based guidelines for diagnosis and management. *Am J Respir Crit Care Med*. 2011;183(6):788–824.
7. Spagnolo P, Sverzellati N, Rossi G, et al. Idiopathic pulmonary fibrosis: an update. *Ann Med*. 2015;47(1):15–27.
8. Izumi S, Iikura M, Hirano S. Prednisone, azathioprine, and N-acetylcysteine for pulmonary fibrosis. *N Engl J Med*. 2012;367(9):870; author reply 870–871.
9. Haschek WM, Meyer KR, Ullrich RL, Witschi HP. Pulmonary fibrosis-a possible mechanism. *Toxicol Appl Pharmacol*. 1979;38(3):1155–1155.
10. Flaherty KR, Toews GB, Lynch JP 3rd, et al. Steroids in idiopathic pulmonary fibrosis: a prospective assessment of adverse reactions, response to therapy, and survival. *Am J Med*. 2001;110(4):278–282.
11. Baumgartner KB, Samet JM, Stidley CA, Colby TV, Waldron JA. Cigarette smoking: a risk factor for idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med*. 1997;155(1):242–248.
12. Miyake Y, Sasaki S, Yokoyama T, et al. Occupational and environmental factors and idiopathic pulmonary fibrosis in Japan. *Ann Occup Hyg*. 2005;49(3):259–265.
13. Baumgartner KB, Samet JM, Coultas DB, et al. Occupational and environmental risk factors for idiopathic pulmonary fibrosis: a multicenter case-control study. *Am J Epidemiol*. 2000;152(4):307–315.
14. Lee JS, Collard HR, Raghu G, et al. Does chronic microaspiration cause idiopathic pulmonary fibrosis? *Am J Med*. 2010;123(4):304–311.
15. Nadrous HF, Pellikka PA, Krowka MJ, et al. Pulmonary hypertension in patients with idiopathic pulmonary fibrosis. *Chest*. 2005;128(4):2393–2399.
16. Cottin V, Le Pavec J, Prevot G, et al; GERM"O"P. Pulmonary hypertension in patients with combined pulmonary fibrosis and emphysema syndrome. *Eur Respir J*. 2010;35(1):105–111.
17. Molyneaux PL, Maher TM. The role of infection in the pathogenesis of idiopathic pulmonary fibrosis. *Eur Respir Rev*. 2013;22(129):376–381.
18. Tsakiri KD, Cronkhite JT, Kuan PJ, et al. Adult-onset pulmonary fibrosis caused by mutations in telomerase. *Proc Natl Acad Sci U S A*. 2007;104(18):7552–7557.
19. Garce A. Political ideologies and party adaptation: the case of MLN-tupamaros (1985-2009). *Rev Cienc Polit Sant*. 2011;31(1):117–137.
20. Selman M, Pardo A. Revealing the pathogenic and aging-related mechanisms of the enigmatic idiopathic pulmonary fibrosis. an integral model. *Am J Respir Crit Care Med*. 2014;189(10):1161–1172.
21. Stuart BD, Lee JS, Kozlitina J, et al. Effect of telomere length on survival in patients with idiopathic pulmonary fibrosis: an observational cohort study with independent validation. *Lancet Respir Med*. 2014;2(7):557–565.
22. Nogee LM, Dunbar AE 3rd, 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(8):573–579.
23. Thomas AQ, Lane K, Phillips 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(9):1322–1328.
24. Brasch F, Griese M, Tredano M, et al. Interstitial lung disease in a baby with a de novo mutation in the SFTPC gene. *Eur Respir J*. 2004;24(1):30–39.
25. Lawson WE, Grant SW, Ambrosini V, et al. Genetic mutations in surfactant protein C are a rare cause of sporadic cases of IPF. *Thorax*. 2004;59(11):977–980.
26. Markart P, Ruppert C, Wygrecka M, et al. Surfactant protein C mutations in sporadic forms of idiopathic interstitial pneumonias. *Eur Respir J*. 2007;29(1):134–137.
27. Li J, Liepinsh E, Almlen A, et al. Structure and influence on stability and activity of the N-terminal propeptide part of lung surfactant protein C. *FEBS J*. 2006;273(5):926–935.
28. Maguire JA, Mulugeta S, Beers MF. Multiple ways to die: delineation of the unfolded protein response and apoptosis induced by Surfactant Protein C BRICHOS mutants. *Int J Biochem Cell Biol*. 2012;44(1):101–112.
29. Lawson WE, Cheng DS, Degryse AL, et al. Endoplasmic reticulum stress enhances fibrotic remodeling in the lungs. *Proc Natl Acad Sci U S A*. 2011;108(26):10562–10567.
30. Willander H, Hermansson E, Johansson J, Presto J. BRICHOS domain associated with lung fibrosis, dementia and cancer – a chaperone that prevents amyloid fibril formation? *FEBS J*. 2011;278(20):3893–3904.
31. Tanjore H, Cheng DS, Degryse AL, et al. Alveolar epithelial cells undergo epithelial-to-mesenchymal transition in response to endoplasmic reticulum stress. *J Biol Chem*. 2011;286(35):30972–30980.

32. Wang Y, Kuan PJ, Xing C, et al. Genetic defects in surfactant protein A2 are associated with pulmonary fibrosis and lung cancer. *Am J Hum Genet*. 2009;84(1):52–59.

33. El-Anwar MW, Hamed AA, Mohamed AE, Nofal AA, Mohamed MA, Abdel-Aziz HR. Surfactant protein a expression in chronic rhinosinusitis and atrophic rhinitis. *Int Arch Otorhinolaryngol*. 2015;19(2):130–134.

34. Sano H, Kuroki Y. The lung collectins, SP-A and SP-D, modulate pulmonary innate immunity. *Mol Immunol*. 2005;42(3):279–287.

35. Maitra M, Wang Y, Gerard RD, Mendelson CR, Garcia CK. Surfactant protein A2 mutations associated with pulmonary fibrosis lead to protein instability and endoplasmic reticulum stress. *J Biol Chem*. 2010;285(29):22103–22113.

36. Maitra M, Cano CA, Garcia CK. Mutant surfactant A2 proteins associated with familial pulmonary fibrosis and lung cancer induce TGF-beta1 secretion. *Proc Natl Acad Sci U S A*. 2012;109(51):21064–21069.

37. Bruder E, Hofmeister J, Aslanidis C, et al. Ultrastructural and molecular analysis in fatal neonatal interstitial pneumonia caused by a novel ABCA3 mutation. *Mod Pathol*. 2007;20(10):1009–1018.

38. Campo I, Zorzetto M, Mariani F, et al. A large kindred of pulmonary fibrosis associated with a novel ABCA3 gene variant. *Respir Res*. 2014;15:43.

39. Ciantelli M, Ghirri P, Presi S, et al. Fatal respiratory failure in a full-term newborn with two ABCA3 gene mutations: a case report. *J Perinatol*. 2011;31(1):70–72.

40. Bullard JE, Nogee LM. Heterozygosity for ABCA3 mutations modifies the severity of lung disease associated with a surfactant protein C gene (SFTPC) mutation. *Pediatr Res*. 2007;62(2):176–179.

41. Parappil H, Al Baridi A, ur Rahman S, et al. Respiratory distress syndrome due to a novel homozygous ABCA3 mutation in a term neonate. *BMJ Case Rep*. 2011;2011:bcr1020103427.

42. Kumar A, Dougherty M, Findlay GM, et al. Genome sequencing of idiopathic pulmonary fibrosis in conjunction with a medical school human anatomy course. *PLoS One*. 2014;9(9):e106744.

43. Armanios M. Telomeres and age-related disease: how telomere biology informs clinical paradigms. *J Clin Invest*. 2013;123(3):996–1002.

44. Alder JK, Chen JJ, Lancaster L, et al. Short telomeres are a risk factor for idiopathic pulmonary fibrosis. *Proc Natl Acad Sci U S A*. 2008;105(35):13051–13056.

45. Strong MA, Vidal-Cardenas SL, Karim B, Yu H, Guo N, Greider CW. Phenotypes in mTERT(+)/(-) and mTERT(-)/(-) mice are due to short telomeres, not telomere-independent functions of telomerase reverse transcriptase. *Mol Cell Biol*. 2011;31(12):2369–2379.

46. Armanios M. Telomerase and idiopathic pulmonary fibrosis. *Mutat Res*. 2012;730(1–2):52–58.

47. Armanios MY, Chen JJ, Cogan JD, et al. Telomerase mutations in families with idiopathic pulmonary fibrosis. *N Engl J Med*. 2007;356(13):1317–1326.

48. Mushiroda T, Wattanapokayakit S, Takahashi A, 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(10):654–656.

49. Lopez-Otin C, Blasco MA, Partridge L, Serrano M, Kroemer G. The hallmarks of aging. *Cell*. 2013;153(6):1194–1217.

50. Alder JK, Barkauskas CE, Limjunyawong N, et al. Telomere dysfunction causes alveolar stem cell failure. *Proc Natl Acad Sci U S A*. 2015;112(16):5099–5104.

51. Zhou BY, Liu YX, Kahn M, et al. Interactions between beta-catenin and transforming growth factor-beta signaling pathways mediate epithelial-mesenchymal transition and are dependent on the transcriptional co-activator cAMP-response element-binding protein (CREB)-binding protein (CBP). *J Biol Chem*. 2012;287(10):7026–7038.

52. Alder JK, Parry EM, Yegnasubramanian S, et al. Telomere phenotypes in females with heterozygous mutations in the dyskeratosis congenita 1 (DKC1) gene. *Hum Mutat*. 2013;34(11):1481–1485.

53. Kropski JA, Mitchell DB, Markin C, et al. A novel dyskerin (DKC1) mutation is associated with familial interstitial pneumonia. *Chest*. 2014;146(1):E1–E7.

54. Stuart BD, Choi JM, Zaidi S, et al. Exome sequencing links mutations in PARN and RTEL1 with familial pulmonary fibrosis and telomere shortening. *Nat Genet*. 2015;47(5):512–U114.

55. Kannengiesser C, Borie R, Ménard C, et al. Heterozygous RTEL1 mutations are associated with familial pulmonary fibrosis. *Eur Respir J*. 2015;46(2):474–485.

56. Cogan JD, Kropski JA, Zhao M, et al. Rare variants in RTEL1 are associated with familial interstitial pneumonia. *Am J Respir Crit Care Med*. 2015;191(6):646–655.

57. Seibold MA, Wise AL, Speer MC, et al. A common MUC5B promoter polymorphism and pulmonary fibrosis. *N Engl J Med*. 2011;364(16):1503–1512.

58. Wang CL, Zhuang Y, Guo WW, et al. Mucin 5B promoter polymorphism is associated with susceptibility to interstitial lung diseases in Chinese males. *PLoS One*. 2014;9(8):e104919.

59. Peljto AL, Selman M, Kim DS, 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(2):460–464.

60. Liptzin DR, Watson AM, Murphy E, et al. MUC5B expression and location in surfactant protein C mutations in children. *Pediatr Pulmonol*. Epub 2015 Apr 7.

61. Spagnolo P, Rossi G, Cavazza A. Pathogenesis of idiopathic pulmonary fibrosis and its clinical implications. *Expert Rev Clin Immunol*. 2014;10(8):1005–1017.

62. Zhou Y, Wang X, Liu M, et al. A critical function of toll-like receptor-3 in the induction of anti-human immunodeficiency virus activities in macrophages. *Immunology*. 2010;131(1):40–49.

63. O'Dwyer DN, Armstrong ME, Trujillo G, et al. The toll-like receptor 3 L412F polymorphism and disease progression in idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med*. 2013;188(12):1442–1450.

64. Noth I, Zhang Y, Ma SF, et al. Genetic variants associated with idiopathic pulmonary fibrosis susceptibility and mortality: a genome-wide association study. *Lancet Respir Med*. 2013;1(4):309–317.

65. Hodgson U, Pulkkinen V, Dixon M, et al. ELMOD2 is a candidate gene for familial idiopathic pulmonary fibrosis. *Am J Hum Genet*. 2006;79(1):149–154.

66. Pulkkinen V, Bruce S, Rintahaka J, et al. ELMOD2, a candidate gene for idiopathic pulmonary fibrosis, regulates antiviral responses. *FASEB J*. 2010;24(4):1167–1177.

67. Falfan-Valencia R, Camarena A, Juarez A, et al. Major histocompatibility complex and alveolar epithelial apoptosis in idiopathic pulmonary fibrosis. *Hum Genet*. 2005;118(2):235–244.

68. Aquino-Galvez A, Perez-Rodriguez M, Camarena A, et al. MICA polymorphisms and decreased expression of the MICA receptor NKG2D contribute to idiopathic pulmonary fibrosis susceptibility. *Hum Genet*. 2009;125(5–6):639–648.

69. Xue J, Gochuico BR, Alawad AS, et al. The HLA class II Allele DRB1*1501 is over-represented in patients with idiopathic pulmonary fibrosis. *PLoS One*. 2011;6(2):e14715.

70. Vasakova M, Striz I, Slavcev A, Jandova S, Kolesar L, Sulc J. Th1/Th2 cytokine gene polymorphisms in patients with idiopathic pulmonary fibrosis. *Tissue Antigens*. 2006;67(3):229–232.

71. Ahn MH, Park BL, Lee SH, et al. A promoter SNP rs4073T>A in the common allele of the interleukin 8 gene is associated with the development of idiopathic pulmonary fibrosis via the IL-8 protein enhancing mode. *Respir Res*. 2011;12:73.

72. Alhamad EH, Cal JG, Shakoor Z, Almogren A, AlBoukai AA. Cytokine gene polymorphisms and serum cytokine levels in patients with idiopathic pulmonary fibrosis. *BMC Med Genet*. 1;14:66.

73. Fingerlin TE, Murphy E, Zhang W, et al. Genome-wide association study identifies multiple susceptibility loci for pulmonary fibrosis. *Nat Genet*. 2013;45(6):613–620.

74. Bueno M, Lai YC, Romero Y, et al. PINK1 deficiency impairs mitochondrial homeostasis and promotes lung fibrosis. *J Clin Invest*. 2015;125(2):521–538.

75. Patel AS, Song JW, Chu SG, et al. Epithelial cell mitochondrial dysfunction and PINK1 are induced by transforming growth factor-beta1 in pulmonary fibrosis. *PLoS One*. 2015;10(3):e0121246.