Pathogenesis of idiopathic pulmonary fibrosis: review of recent findings
Elisabetta Renzoni¹*, Veeraraghavan Srihari² and Piersante Sestini³

Abstract
Idiopathic pulmonary fibrosis (IPF) is likely to result from the interaction between environmental exposures, including cigarette smoke, and genetic predisposition. This review focuses on clues provided by recent genetic association studies and other selected data and hypotheses. In IPF, association with surfactant mutations has highlighted the importance of type II epithelial cells, while shortened telomeres in some patients suggest that accelerated aging may play a role in the pathogenesis of lung fibrosis, possibly by affecting the renewal/differentiation potential of epithelial cells. The finding that a common variant in mucin 5B predisposes individuals to both familial and sporadic IPF suggests a hitherto under-investigated role of bronchiolar cells and mucins. Although the pathogenetic link between mucins and lung fibrosis is not known, it is possible that MUC5B overexpression interferes with physiological mucosal host defense, with reduced clearance of micro-organisms or inorganic noxious agents, or induction of endoplasmic reticulum stress. Other components of innate and adaptive immunity are likely to be involved in IPF pathogenesis/progression. Finally, the importance of the clotting cascade in IPF pathogenesis has been confirmed by a recent epidemiological study, in which patients with IPF were almost five times more likely than general population controls to have at least one inherited or acquired clotting defect.

Introduction
Idiopathic pulmonary fibrosis (IPF) is a progressive fibrotic disease limited to the lungs, occurring in older individuals, more frequently men, and characterized by a dismal prognosis, with a median survival of 3 to 5 years since diagnosis, although a wide variability in disease course is increasingly recognized. According to the more commonly held and current pathogenetic model, IPF results from recurrent injury to epithelial cells caused by a variety of exposures. These include cigarette smoke, dusts, and other environmental agents, which, in a genetically predisposed individual, lead to activation of abnormal pathways, resulting in failed resolution of the wound-healing response. This article does not aim to provide an exhaustive review of pathogenetic pathways in IPF, but rather will focus on selected topics, in particular on contributions to the understanding of IPF pathogenesis provided by recent genetic association studies, and other selected data and hypotheses.

Genetic studies
Genetic predisposition to IPF is supported by familial clustering, the occurrence of lung fibrosis in genetic multi-system disorders, and differing susceptibilities in humans exposed to similar levels of fibrogenic agents. Genetic association studies are useful in identifying relevant molecules/cell types among the wide array of potential pathogenetic pathways and can highlight otherwise unexpected areas.
The role of type II cells

Mutations in the genes for surfactant protein C (SFTPC) and surfactant protein A2 (SFTPA2) have been described in association with familial lung fibrosis [1,2] and rarely with sporadic IPF [3,4]. Although they account for only approximately 1% of familial cases, the fact that both surfactant proteins are exclusively synthesized by type II alveolar epithelial cells points to type II cell dysfunction as a key element in IPF. The relevance of type II cells is also suggested by the prognostic value of serum surfactant proteins in sporadic IPF and other interstitial lung diseases (ILDs) [5]. In addition to synthesizing surfactant proteins, type II alveolar cells are responsible for the constant regeneration of the alveolar epithelium, as precursors of the terminally differentiated type I cells. Surfactant protein C and A mutations have been shown to cause incorrect protein folding/processing, thereby activating the cell’s endoplasmic reticulum (ER) stress response. This is a stereotypical protective mechanism in response to the accumulation of misfolded proteins, ultimately leading to activation of apoptosis pathways [6,7]. Increased staining for markers of ER stress response and apoptosis is also observed in sporadic IPF [8,9], highlighting a prominent role for activation of the ER stress response and alveolar cell apoptosis in pathogenesis.

Type II cells are also affected in the Hermansky-Pudlak syndrome (HPS), a genetic disease frequently associated with lung fibrosis. HPS is caused by mutations in genes encoding for proteins involved in lysosome-related intracellular trafficking, leading to giant lamellar body formation in type II alveolar cells [10]. Although alveolar epithelial ER stress has been described in HPS lung fibrosis [11], Young and colleagues report that impaired intracellular trafficking pathways do not directly result in ER stress, suggesting further studies are needed to evaluate mechanisms [12].

Telomere shortening

Pulmonary fibrosis occurs in a subset of patients with dyskeratosis congenita, a rare genetic disorder. Rapidly dividing tissues are affected, including epithelial cells and bone marrow, secondary to mutations in genes encoding for telomerase enzyme components. Telomerasases are responsible for the elongation of telomeres, repetitive DNA sequences capping chromosome ends. Telomeres shorten at each cell replication until a critical length is reached, triggering a DNA damage response, leading to cellular senescence or apoptosis. Loss of telomerase activity reduces the renewal potential of stem cell populations [13]. Rare mutations in the telomerase genes TERT and TERC have been identified in approximately 8% of familial cases and in 1% to 3% of sporadic cases [14,15]. Telomeres shorten with age, and telomere dysfunction may be involved in a number of age-related degenerative disorders [16]. Even in the absence of TERT/TERC mutations, a proportion of IPF patients have shorter telomeres compared with age-matched controls [17,18]. In aggregate, these findings suggest that IPF may be a disease of accelerated aging in the lung, at least in the subset characterized by shortened telomeres. However, shorter telomeres in circulating leukocytes compared with age-matched control populations are also observed in chronic obstructive pulmonary disease (COPD), in asthma, and in association with lower lung function in control individuals, suggesting that the phenomenon of premature aging is not specific to pulmonary fibrosis [19].

The mechanisms through which telomere/telomerase defects may lead to pulmonary fibrosis are not well known. As telomerase dysfunction is most likely to affect cells with high turnover, epithelial cells are obvious candidates. Indeed, IPF alveolar epithelium had shorter telomeres than normal control epithelium, independently of detectable telomerase mutations in one study [17], although further studies are needed to confirm this interesting observation. Telomerase activity may differ according to cell type as increased telomerase activity has been reported in IPF lung fibroblasts, which is in keeping with potential increased survival and resistance to apoptosis in these cells [20].

Bronchiolar cells

A prominent role for activated bronchiolar-type epithelial cells was first proposed by Chilosi and colleagues, in a study reporting activation of the Wnt/beta-catenin/matrylsin developmental pathway in hyperplastic bronchiolar lesions [21]. This was a feature seen in IPF but not in other ILD patterns, in contrast with nuclear beta-catenin expression observed in type II pneumocytes across ILDs. A particularly strong beta-catenin staining in areas of bronchiolization was confirmed by Konigshoff and colleagues [22].

In genetic studies, a role for bronchiolar cells has been suggested by the finding of a marked increase in a MUC5B gene common promoter polymorphism in familial interstitial pneumonia (FIP) and sporadic IPF compared with controls (minor allele present in 33.8% of FIP cases, 37.5% of sporadic IPF cases, and 9.1% of control subjects) [23,24]. This strong association suggests the pathogenetic significance of mucins/small airways in IPF. By contrast, no association was found between MUC5B and scleroderma-associated lung fibrosis or with fibrotic sarcoidosis, suggesting that this association may be specific to the idiopathic ILDs [25,26].

Although the MUC5B variants are associated with overexpression in lung tissue compared with non-carriers
in control subjects, this is not the case in IPF, where lung MUC5B overexpression is seen independently of genetic variants [23]. This observation suggests that, in the absence of (or in addition to) MUC5B genetic variation, other environmental or genetic factors (or both) linked to disease pathogenesis cause its upregulation. MUC5B appears to localize mainly to the distal airways and to the honeycomb cysts in what appear to be bronchiolar cells rather than type II alveolar cells [27]. Honeycombing lesions would therefore derive from the dilatation of the small airways, an observation that was originally made more than 50 years ago [28]. The mucin-producing cells in the honeycombed areas are characterized by a transcription factor and mucin profile suggestive of abnormal programming, with a phenotype similar to that of submucosal gland goblet cells, normally not found in the distal airspaces [29]. The mechanism by which MUC5B polymorphisms predispose individuals to pulmonary fibrosis is not known. One of the hypotheses is that MUC5B variants may lead to impaired clearance of inhaled micro-organisms and other particles and therefore an aberrant innate response, although a direct toxic effect on epithelial cells and/or ectopic expression with intracellular accumulation and resulting ER stress response are also possibilities [23,30].

Interestingly, while MUC5B overexpression is associated with IPF independently of the genetic polymorphisms, the MUC5B genetic variant actually appears to be associated with slower progression and better survival [26,31], suggesting that it marks a subset of the disease. There is a known variability in the clinical course of IPF, ranging from prolonged stability/ggradual worsening to rapid stepwise progression [32–34]. Disease subsets are also suggested by a recent whole-genome expression analysis in a large number of IPF biopsies [35]. Clustering of the genes differentially expressed compared with controls reveals a separation into two clear IPF subpopulations, based on a distinct molecular signature. One IPF group is characterized by overexpression (compared with the other) of a number of genes already known to be upregulated in IPF, including osteopontin, matrix metalloprotease 1 (MMP1), MMP7, and MUC5B. The same IPF group is also characterized by upregulation of a large number of transcripts associated with cilium genes. These include structural components such as axonemal dyneins involved in ciliary motility as well as transcription factors regulating cilium gene expression, including RFX3, FOXJ1, and RFX2. Patients with high cilium gene expression demonstrate more microscopic honeycombing but not more fibroblastic foci on lung biopsy [35]. These observations reinforce the role played by airway cells (at least in a subset) and for the first time provide molecular evidence that IPF may in fact consist of more than one disease, each with likely different pathogenesis, prognosis, and treatment responses. The authors do not mention whether these two subsets separate out according to the MUC5B polymorphisms or whether they are characterized by different longitudinal behavior, and no doubt further studies will shed light on this.

Two recent large genome-wide case-control association studies have confirmed the association with MUC5B [36,37], with the study by Fingerlin and colleagues [36] also identifying associations with common variants in the TERT and TERC genes. Both studies also report associations with several novel common gene variants. Among these, Fingerlin and colleagues report an association with genes involved in epithelial integrity and cell-cell adhesion, including desmoplakin (DSP) and DPP9, suggesting that defects in cell-cell adhesion or cellular cytoskeleton (or both) could predispose patients to injury in response to stimuli, including mechanical stretch. Both studies report an association with TOLLIP, a regulator of innate immune responses involved in modulating Toll-like receptor (TLR) signaling. The TOLLIP gene resides close to MUC5B, and variants in the two genes are in weak linkage disequilibrium. Once the MUC5B variants were included in the analysis, the association with the TOLLIP gene single-nucleotide polymorphisms (SNPs) only just reached statistical significance in the study by Noth and colleagues [37] (P = 0.05) and was no longer significant in the study by Fingerlin and colleagues, indicating the need for further studies. Interestingly, the minor allele of one of the TOLLIP SNPs (rs5743890) found to be protective toward lung fibrosis was associated with worse survival. This finding was maintained on adjusting for other gene variants, although MUC5B was not included in the multivariate analysis [37]. Again, whether the TOLLIP variant is a prognostic determinant, independently of MUC5B, remains to be determined.

**Gene expression and epigenetic regulation**

Transcriptional profiling studies of fibrotic lung tissues have identified key differences with control tissues and between different fibrotic ILD patterns [38–41]. Matriuin (MMP7), one of the downstream targets of Wnt/beta-catenin, is highly overexpressed in IPF biopsies [38] and has been investigated as a peripheral blood biomarker in ILD. In the largest prospective biomarker study performed so far in IPF, MMP7, together with intercellular adhesion molecule-1 (ICAM-1) and interleukin-8 (IL-8), was the strongest predictor of survival in combination with clinical characteristics [42]. Boon and colleagues reported on upregulation in pathways involved in cell growth, proliferation, and migration in lung tissue from rapidly progressive compared with relatively stable IPF [43], suggesting that there may be common pathways.
between lung fibrosis and cancer development [44]. Rapidly progressive fibrosis was also characterized by upregulation of PLUNC, a molecule expressed mainly by secretory bronchial columnar cells, again highlighting a potential role for airway/innate defense functions [43]. With regard to fibroblasts, these are the main effector cell types involved in the fibrotic process [45], regardless of the initiating triggers. Gene expression profiling studies of cultured lung fibroblasts show specific pathways of activation [46–48], which appear to be similar in IPF and in scleroderma-associated ILD [48].

Although the study of epigenetic regulation in IPF is relatively recent, analysis in lung tissue or in IPF fibroblasts (or both) of the global and individual gene expression regulation through methylation [49–51], acetylation [52,53], and regulatory microRNAs [54–61] is providing information on regulatory networks and should allow a better understanding of how particular genetic variants and environmental exposures interact to cause pulmonary fibrosis, ultimately leading to novel treatment strategies.

**Innate and adaptive immune responses**

The role played by conventional inflammation in IPF pathogenesis has been questioned, mainly in view of the lack of response to corticosteroids and immunosuppressants, which may indeed be harmful in patients with IPF [62]. However, components of the innate and adaptive immune response are thought to contribute to the pathogenesis of lung fibrosis.

The innate immune system plays a crucial role in initiating and terminating inflammatory responses to endogenous and exogenous noxious stimuli. The alveolar macrophage is a potential source of chemokines and growth factors that regulate the wound-healing response. Macrophages are characterized by at least two phenotypes: M1, induced by Th1 cytokines such as interferon, and expressing IL-12, tumor necrosis factor (TNF), and C-X-C motif chemokine (CXCL)10, and M2 (or alternatively activated M), which expresses CD163, mannose receptors, and secretes increased levels of IL-10, chemokine (C-C motif) ligand (CCL)18, and CCL22 [63]. Two carefully conducted large prospective studies have shown that bronchoalveolar lavage (BAL) and serum levels of CCL18, one of the markers of M2 macrophages, are significantly linked to extent of fibrosis and predict likelihood of progression, even after adjustment for disease severity, both in IPF and in systemic sclerosis-ILD [64,65].

Pattern recognition receptors, including TLRs, recognize pathogen and endogenous patterns, activate innate and adaptive immune responses, and participate in regulating wound healing [66]. Pathogen-associated molecular patterns (PAMPs) are suspected to play a role in the activation of fibroblasts through TLRs as well as nucleotide-binding oligomerization domain-like receptors (NLRs) and retinoic acid-inducible gene-1 (RIG-1) receptors [67]. Endogenous ligands can also be identified by TLRs, including components of the extracellular matrix such as hyaluronic acid [68], and fibrinogen degradation products [69]. Interestingly, the ligation of the latter to TLR4 has been shown to induce upregulation of mucin MUC5AC expression in both alveolar macrophages and epithelial cells [69].

TLR9 expression has been found to be increased in ILD lung tissue and in IPF fibroblasts compared with control subjects [70,71]. In particular, in both usual interstitial pneumonia (UIP) and non-specific interstitial pneumonia (NSIP), the increased expression is localized to areas of fibrosis in the epithelium and interstitium [70]. TLR9 activation by hypomethylated DNA significantly increased alpha smooth muscle actin [70] and promoted profibrotic cytokine and chemokine synthesis in IPF fibroblasts in culture [72]. TLR9 was particularly upregulated in the lung tissues of rapidly progressive IPF compared with more stable disease [73].

TLR3 has been shown to act as a sensor of tissue necrosis during acute inflammation. In addition, it recognizes virus-associated molecular patterns and activates pro-inflammatory cascades [74,75]. Recently, a TLR3 polymorphism (TLR3L412F) associated with defective function of the receptor was studied in patients with IPF. Fibroblasts homozygous for the variant allele had an impaired interferon-beta response, displaying enhanced proliferation after TLR3 stimulation, partially corrected after interferon-beta treatment. Furthermore, the variant was associated with disease progression or early mortality (or both) in two separate IPF cohorts, with an additive effect for each allele [76]. This study is another example of a genetic variant linked to outcome in IPF and highlights the possibility of differences in response to treatments in different subsets, as suggested by the editorial accompanying the article [77]. The activation of TLRs in epithelial and interstitial cells could be one of the links between exogenous/endogenous noxious signals and amplification of immune and fibroproliferative responses. Further studies are needed to assess their role in the pathogenesis of pulmonary fibrosis.

Adaptive immunity is also likely to be involved in IPF progression. We know that in hypersensitivity pneumonitis (HP) adaptive immunity is a crucial driver of
A proportion of patients with fibrotic HP have a UIP histology and can display an IPF-like disease course [78], although features such as airway centeredness of fibrosis and excess inflammation suggest the diagnosis of HP-associated UIP [79]. There may be at least a subset of patients with IPF in whom the fibrotic response is a response to an unknown antigen, with a role played by adaptive immunity. Indeed, Gilani and colleagues found that increased proportions of circulating CD28nullCD4+ T cells are associated with poor outcomes in patients with IPF [80]. CD28nullCD4+ cells are effector memory T cells, which develop as a result of repeated antigen-driven proliferations. These cells display increased production of cytokines and pro-inflammatory cytokines, in keeping with enhanced noxious potential. Interestingly, CD28null T-cell telomeres are shortened as a result of repetitive T-cell clonal proliferations [81], and T cells comprise 40% to 50% of peripheral blood monocytes. This could contribute to findings of telomere shortening in the peripheral blood leukocytes of IPF patients and other populations with chronic disease. The relevance of T-cell differentiation in IPF, again as a result of repetitive antigen stimulation, is also supported by a recent report by Herazo-Mayar and colleagues [82]. BAL and lung tissue-activated CD4+ T cells are observed in asymptomatic early familial ILD cases [83], and pro-inflammatory dendritic cells are found in advanced lung fibrosis [84]. In IPF lungs, CD8+ T lymphocytes correlate with degree of breathlessness and functional severity [85]. IPF lung-derived proteins induce proliferation of autologous CD4+ T cells from these patients [86]. Semaphorin 7a+ regulatory T cells are increased in the blood of patients with rapidly progressive IPF [87].

In addition to T cells, B cells may also be involved in IPF pathogenesis. Auto-antibodies against heat shock protein 70 (HSP 70) are associated with short-term deterioration in patients with IPF, independently of disease severity, suggesting a potential link to a pathogenetically significant process [88]. The finding of lymphoid follicles in IPF biopsies has been reported in a number of studies [89-91]. The presence of B-cell aggregates in injured tissues suggests a possible direct contribution of the significant pathogenetic potential of B cells to the fibrotic process [92]. Xue and colleagues observed an increase in circulating antigen-differentiated plasma cells in IPF patients and overexpression of a specific B-cell trophic factor (plasma B lymphocyte-stimulating factor), associated with poor short-term outcome [90]. A chemokine that mediates B-cell trafficking, CXCL13, was increased in the lung tissue and peripheral blood of IPF patients compared with healthy controls and COPD patients and was predictive of early mortality, further supporting a pathogenetic role for B cells [93].

It seems likely that both the innate and adaptive immune systems participate in lung fibrosis progression. However, the question of whether the hyperstimulation of the immune system described above is directly causative of lung fibrosis rather than a consequence of the loss of the normal lung architecture and therefore of natural defenses is more difficult to resolve. We should also consider the possibility of an interaction between the different pathogenetic mechanisms. For example, Agrawal and colleagues recently observed that an interaction between surfactant protein (SP)-A and TLR2 is involved in regulating secretion of pro-inflammatory mediators in macrophage lines both in vitro and in vivo [94].

**The clotting cascade**

The clotting cascade appears to be activated in pulmonary fibrosis, as suggested by in vitro and animal model studies [95]. Cleaved clotting factors have major pro-inflammatory and profibrogenic effects [96], and activated platelets/endothelial cells release fibrogenic mediators, including platelet-derived growth factor (PDGF) and transforming growth factor-beta (TGF3) [97]. Navaratnam and colleagues recently reported on a carefully conducted case-control study of 211 incident IPF cases to evaluate the incidence of clotting abnormalities identified through a pro-thrombotic panel of 11 clotting defects [98]. The authors found that patients with IPF have an almost fivefold greater chance of having at least one clotting abnormality, with incrementally increasing odds ratios for increasing numbers of clotting abnormalities. The highest odds ratios were seen for elevation of factor VIII levels (>165 IU/dL in 124 out of 194 patients with IPF versus 45 out of 211 controls). Furthermore, clotting abnormalities were associated with more severe disease at baseline and a worse survival. How this fits with the recent trial of warfarin treatment in IPF, which was stopped early because of an increased rate of pulmonary fibrosis deterioration in the warfarin-treated arm [99], will need to be further investigated, but it does support previous studies pointing to an important role played by the coagulation system in the lung fibrosis process.

**Physical factors**

Leslie recently proposed a unifying hypothesis for IPF as a disease of “recurrent stretch injury to the peripheral and basal lung occurring over many years in predisposed individuals” [100]. This theory attempts to reconcile a number of unexplained features, including the predominantly basal and subpleural distribution, the progression...
of the disease toward the apical regions, and the finding that fibroblastic foci are a continuum extending inward from the pleural surface (fibroblastic reticulum) [101]. IPF would start in the basal and peripheral regions of the lungs, where the alveoli are the smallest in the upright and supine positions and are therefore more likely to collapse in response to the tractional forces during respiration, as also suggested by Galvin and colleagues [102] and Dail [103]. Surfactant protein abnormalities could lead to increased surface tension and greater likelihood of alveolar collapse, while the reticular network of fibroblastic foci would occur along stress fractures/lines at the interface between the epithelium and the mesenchymal cells. A vicious circle of increasing alveolar surface tension causing further alveolar collapse and subsequent enlargement of the alveolar ducts leading to microscopic honeycombing would follow. Carloni and colleagues applied a mathematical model to verify whether the distribution of IPF lesions corresponds to the areas of greatest mechanical tension [104]. If these mechanical forces have a role in the origin or in the progression of the disease, they should also influence those cases with associated emphysema, which were initially described more than 20 years ago [105] and which only recently have attracted the attention of researchers [106]. In these cases, one would expect a lower traction in the basal zones because of reduced elastic recoil of the upper parts of the lung.

Conclusion
In conclusion, genetic association studies suggest a pivotal pathogenetic role played by abnormal regeneration/differentiation potential of respiratory epithelial cells in IPF (both alveolar and bronchiolar), which is likely to be determined by the interaction between environmental exposures such as cigarette smoke and gene variants, including surfactant protein and telomerase genes. The association with MUC5B suggests a role of airway epithelial cells and mucins. Further studies on how molecular phenotypes can be linked to genetic data and to clinical, radiological, and histological characteristics of the disease will be crucial to predict clinical course and drive novel treatments in IPF.

Abbreviations
BAL, bronchoalveolar lavage; COPD, chronic obstructive pulmonary disease; CCL, chemokine (C-C motif) ligand; CXCL, C-X-C motif chemokine; ER, endoplasmic reticulum; FIP, familial interstitial pneumonia; HP, hypersensitivity pneumonitis; HPS, Hermansky-Pudlak syndrome; IILD, interstitial lung disease; IPF, idiopathic pulmonary fibrosis; MMP, matrix metalloprotease; SNP, single-nucleotide polymorphism; TLR, Toll-like receptor; UIP, usual interstitial pneumonia.

Disclosures
The authors declare that they have no disclosures.

References


