Molecular classification of idiopathic pulmonary fibrosis: 
Personalized medicine, genetics and biomarkers

NATHAN HAMBLY1,2,3 CHIKO SHIMBORI1,2,3 AND MARTIN KOLB1,2,3

1Firestone Institute for Respiratory Health, 2St. Joseph’s Healthcare and 3Department of Medicine, McMaster University Hamilton, Hamilton, ON, Canada

ABSTRACT
Idiopathic pulmonary fibrosis (IPF) is a chronic and progressive fibrotic lung disease associated with high morbidity and poor survival. Characterized by substantial disease heterogeneity, the diagnostic considerations, clinical course and treatment response in individual patients can be variable. In the past decade, with the advent of high-throughput proteomic and genomic technologies, our understanding of the pathogenesis of IPF has greatly improved and has led to the recognition of novel treatment targets and numerous putative biomarkers. Molecular biomarkers with mechanistic plausibility are highly desired in IPF, where they have the potential to accelerate drug development, facilitate early detection in susceptible individuals, improve prognostic accuracy and inform treatment recommendations. Although the search for candidate biomarkers remains in its infancy, attractive targets such as MUC5B and MPP7 have already been validated in large cohorts and have demonstrated their potential to improve clinical predictors beyond that of routine clinical practices. The discovery and implementation of future biomarkers will face many challenges, but with strong collaborative efforts among scientists, clinicians and the industry the ultimate goal of personalized medicine may be realized.

Key words: biomarker, genetics, idiopathic pulmonary fibrosis.

INTRODUCTION
Idiopathic pulmonary fibrosis (IPF) is a chronic and fatal fibrotic lung disease that typically affects adults over the age of 50. Characterized by the relentless progression of interstitial fibrosis and a progressive decline in gas exchange, the median survival from the time of diagnosis is approximately 3–5 years. The burden of disease is further compounded by the fact that IPF afflicts more patients than any other form of fibrotic parenchymal lung disease. Histologically, IPF is defined by the distinct pattern of usual interstitial pneumonia (UIP). The
pathological features of UIP include evidence of marked subpleural and paraseptal fibrosis, heterogeneous architectural distortion with or without honeycomb cysts, and in areas of active fibrosis clusters of fibroblasts and myofibroblasts, termed ‘fibroblastic foci’. The UIP pattern, however, is not unique to IPF and can be observed in numerous fibrotic lung diseases, including asbestosis, connective tissue disease-related interstitial lung disease (CTD-ILD), chronic hypersensitivity pneumonitis (HP) and drug reaction. Hence, as proposed by the updated 2011 American Thoracic Society/European Respiratory Society international consensus statement on IPF, the establishment of a diagnosis of IPF requires first the exclusion of known causes of ILD, and second compatible radiographical and pathological findings on high-resolution computed tomography and surgical lung biopsy, respectively. In the absence of a diagnostic gold standard, diagnostic accuracy is dependent on clinical, radiographical and histopathological correlation. This can be best accomplished with expert multidisciplinary discussion, particularly in instances of diagnostic uncertainty.

Once a consensus diagnosis of IPF is reached, a major challenge facing clinicians involves accurately predicting disease course. Significant heterogeneity exists among individual patients, as the clinical course can be variable and unpredictable. In some, the disease is characterized by a rapid deterioration with progressive functional disability, leading to death within months. Others experience a protracted course with little functional impairment. In some, periods of relative stability are punctuated by acute exacerbations that can precipitate dramatic respiratory failure. Traditional means of risk stratifying patients with fibrotic lung disease are dependent on clinical variables (history and physical examination, pulmonary function, exercise testing, radiographical findings, and histological features) that are poorly reflective of disease pathogenesis and provide insufficient power to accurately predict clinical outcome. Although multiple prognostic scoring systems that provide reasonable insight have been reported, they do not account for the distinct molecular mechanisms that drive the fibrotic cascade in individual patients. This issue represents a major roadblock to personalized patient care in IPF and is the primary focus of ongoing clinical research.

In this review, we examine candidate molecular biomarkers that have been investigated in IPF through the lens of the complex pathogenesis of pulmonary fibrosis (PF). We explore how these biomarkers play key roles in the progressive fibrotic cascade observed in IPF, and how their measurement can influence diagnostic, prognostic and treatment considerations.

PATHOGENESIS OF IDIOPATHIC PULMONARY FIBROSIS

Although the exact mechanisms responsible for the development of IPF remain largely unknown, significant strides have been made. Most notable is the pathogenic paradigm of IPF shifting from one of inflammation-driven fibrogenesis to one of aberrant wound healing following repetitive alveolar epithelial cell (AEC) injury. According to this concept, the alveolar epithelium is subject to clinically silent micro-injuries over a prolonged period. As injuries accumulate, a theoretical threshold is reached, and AECs are aberrantly activated. The dysfunctional epithelial cells then activate pro-fibrotic signalling pathways, involving growth factors and chemokines, such as TGF-β. Subsequent migration, proliferation and activation of mesenchymal cells leads to the formation of fibroblastic foci and exuberant extracellular matrix (ECM) deposition. The consequences of such signalling result in a ‘fibroproliferative disorder’, with fibroblastic foci being the primary sites of pro-fibrotic responses.

This paradigm shift is supported by the PANTHER-IPF clinical trial, a randomized, double-blind, placebo-controlled study that evaluated the safety and efficacy of a three-drug immunosuppressive regimen of oral prednisone, azathioprine and N-acetylcysteine in patients with IPF. In comparison to placebo, immunosuppressive therapy was associated with an increased risk of death and hospitalization. Conversely, therapeutic agents, such as pirfenidone and nintedanib, which target key pro-fibrotic signalling pathways, have recently been shown in large randomized trials to be safe and efficacious. Clearly, the importance of approaching IPF as a ‘fibroproliferative disorder’ cannot be understated.

Although inflammation is never a prominent histopathological finding in UIP, its role in IPF should not be fully abandoned, as it cannot be ignored that low-level inflammation and immune activation are relevant components of wound repair. A complex interplay between lymphocyte subsets exists, with regulatory T cells acting as pro-fibrotic elements. The role of CD4+ T cells is variable depending on the local microenvironment, while B-lymphocytes have been implicated in the production of auto-antibodies against AEC antigens and immune-complex mediated inflammatory reactions. Circulating B-lymphocyte stimulating factor (BlyS), a molecule necessary for B-cell development and survival, has been correlated with acute exacerbations and survival in IPF patients. Through pro-fibrotic cytokines and growth factors, macrophages regulate ECM composition and fibroblast function.

The accumulation of fibroblasts into fibroblastic foci, the differentiation and proliferation of myofibroblasts, and the subsequent synthesis of ECM are hallmarks of IPF. Myofibroblasts are not native to the lung and differentiate in response to tissue injury. They are mobile cells with contractile ability and are the primary producer of collagen and other ECM proteins. These cells can exhibit anchorage independent growth, analogous to that of malignancies. The role of the myofibroblast is further convoluted by the fact that several cell types have been proposed as their putative precursor: tissue-resident fibroblasts, epithelial cells (via epithelial to mesenchymal transition), pleural mesothelial cells, bone marrow-derived fibrocytes and pericytes.

© 2015 Asian Pacific Society of Respirology

Respirology (2015)
The lung matrix is a complex and dynamic network of collagens, elastin, proteoglycans and glycoproteins. Traditionally regarded as a mere structural support for the complex lung architecture, the ECM is now considered a key signalling entity in the development of progressive fibrosis. A stiff fibrotic matrix promotes fibroblast proliferation and reduces fibroblast apoptosis. Matrix stiffness can also transform fibroblasts from a quiescent state to a self-sustained activated state that drives progressive fibrosis.

The most important enzymes in ECM turnover are matrix metalloproteinases (MMP), which degrade multiple components of the ECM, activate and degrade biological mediators, and facilitate cell migration. Hence, the progressive fibrosis observed in IPF is dependent on a plethora of intricate signalling pathways activating numerous disparate effector cells (Fig. 1). This complexity complicates the development of reliable biomarkers and effective treatments (Fig. 1). We will use these core mechanistic pathways to exemplify the various candidate biomarkers that have been described to date (Table 1).

**MOLECULAR BIOMARKERS**

The term ‘molecular biomarker’ generally refers to any objectively quantifiable biological measurement (e.g. protein level in serum or bronchoalveolar lavage fluid (BALF), or specific genetic mutation or polymorphism) that may act as a surrogate marker for clinically meaningful variables. An ideal biomarker should be easily acquired through non-invasive means, have high validity and reliability, and be available for serial monitoring. Rather than represent epi-phenomenon, molecular biomarkers should reflect the pathobiological mechanisms driving progressive fibrosis. From a clinical perspective, a biomarker is particularly useful when the information it entails permits the provision of superior patient care beyond that of conventional practices.

In theory, molecular biomarkers have the potential for dramatic clinical impact in IPF, influencing multiple aspects of patient care. Validated biomarkers could be used as (i) predisposition biomarkers that identify patients at risk of developing IPF; (ii) screening biomarkers that identify individuals...
Table 1  Currently explored biomarkers for IPF

<table>
<thead>
<tr>
<th>Biomarker</th>
<th>Correlates with</th>
<th>Differentiates IPF from other ILD</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Predisposition</td>
<td>Diagnosis</td>
<td>Prognosis</td>
</tr>
<tr>
<td>Genomic markers</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SPA2</td>
<td>+/-</td>
<td>-</td>
<td>+/-</td>
</tr>
<tr>
<td>SPC</td>
<td>+/-</td>
<td>-</td>
<td>+/-</td>
</tr>
<tr>
<td>MUC5B</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>TERT/TERC</td>
<td>+/-</td>
<td>-</td>
<td>+/-</td>
</tr>
<tr>
<td>Telomere length</td>
<td>+/-</td>
<td>+/-</td>
<td>-</td>
</tr>
<tr>
<td>TLR3</td>
<td>+/-</td>
<td>-</td>
<td>+/-</td>
</tr>
<tr>
<td>VEGF</td>
<td>+/-</td>
<td>+/-</td>
<td>-</td>
</tr>
<tr>
<td>cCK18</td>
<td>+/-</td>
<td>+/-</td>
<td>-</td>
</tr>
<tr>
<td>Fibroproliferation and matrix deposition</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MMP7</td>
<td>+/-</td>
<td>+/-</td>
<td>+/-</td>
</tr>
<tr>
<td>MMP1</td>
<td>+/-</td>
<td>+/-</td>
<td>-</td>
</tr>
<tr>
<td>Periostin</td>
<td>+/-</td>
<td>+/-</td>
<td>-</td>
</tr>
<tr>
<td>Osteopontin</td>
<td>+/-</td>
<td>+/-</td>
<td>-</td>
</tr>
<tr>
<td>Circulating fibrocytes</td>
<td>+/-</td>
<td>+/-</td>
<td>-</td>
</tr>
<tr>
<td>Immune dysregulation</td>
<td>+/-</td>
<td>+/-</td>
<td>-</td>
</tr>
</tbody>
</table>

Key: The strength of evidence supporting each putative biomarker is denoted below. Boldfaced items have stronger evidence supporting their use.
+ , Consistent or strong evidence supporting candidate biomarker.
+/- , Conflicting or poor quality evidence to support candidate biomarker.
-, Consistent or strong evidence against a candidate biomarker.
Blank, lack of evidence.

BLyS, B-lymphocyte stimulating factor; cCK18, cleaved cytokeratin 18; CCL18, CC chemokine ligand 18; CXCL13, C-X-C motif chemokines 13; FVC, forced vital capacity; ILD, interstitial lung disease; IPF, idiopathic pulmonary fibrosis; KL-6, Krebs vol den Lungen 6; MMP, matrix metalloproteinase; MUC5B, mucin 5B promoter variant; NSIP, non-specific interstitial pneumonia; PCMI, personal clinical and molecular mortality prediction index; PH, pulmonary hypertension; SP, surfactant protein; SP-A, surfactant proteins A; SP-D, surfactant proteins D; TERC, telomerase RNA component; TERT, telomerase reverse transcriptase; TLR3, toll-like receptor 3; VEGF, vascular endothelial growth factor.
with subclinical disease; (iii) diagnostic biomarkers to assist in the diagnostic assessment and classification of patients with fibrotic lung disease; (iv) prognostic biomarkers that assist in appraising disease severity, predicting disease progression and determining overall prognosis; and (v) treatment efficacy biomarkers that have the ability to either isolate patients with the greatest likelihood of treatment response or substitute for clinically meaningful outcomes.64,65 This last point is critical as molecular biomarkers are at the forefront of the ‘personalized medicine’ movement, permitting the development of therapies that target the signalling pathways responsible for disease progression and the patients most likely to benefit.

Considering their wide range of potential clinical and research applications, the interest garnered by biomarkers is not surprising. ‘Hypothesis-driven’ research and high-throughput proteomic and genomic technologies have identified multiple candidate molecular biomarkers in patients with IPF. Unfortunately, despite these advances, no specific marker has been internationally accepted for widespread implementation. This disparity is related to the multiple challenges encountered with biomarker discovery, validation and clinical evaluation, particularly in a disease state as complex as IPF.

### Biomarkers associated with alveolar epithelial cell injury and dysfunction

**Krebs von den Lungen-6 (KL-6)**

Krebs von den Lungen-6 (KL-6) is a mucin-like glycoprotein expressed on the extracellular surface of regenerating alveolar type II epithelial cells (AEC-II) and bronchiolar epithelial cells.56,57 Upon epithelial damage, KL-6 leaks into the circulation where it can be measured. It acts as a chemotactic factor, promoting the migration, proliferation and survival of lung fibroblasts.58,59

Compared with healthy volunteers, serum KL-6 levels are significantly elevated in patients with IPF50,61. Unfortunately, the specificity of KL-6 in IPF is poor, with elevated levels observed in the setting of non-specific interstitial pneumonia (NSIP) and systemic sclerosis-related ILD.41 Moreover, elevated serum measurements are observed in lung cancer and tuberculosis.62,63 Hence, the value of KL-6 as a diagnostic biomarker is limited.

KL-6 has been studied as a prognostic marker in multiple forms of ILD. Satoh et al. prospectively evaluated serum KL-6 levels in 152 patients with idiopathic interstitial pneumonia and 67 patients with CTD-ILD.41 Patients with a KL-6 level >1000 U/mL experienced poorer survival compared with those with lower levels. In a smaller study of 14 patients, Yokoyama et al. found that serial increases in KL-6 concentration were associated with poor survival.64 Elevated levels have also been observed in patients experiencing an acute exacerbation of IPF. In the largest study involving IPF patients, baseline KL-6 levels did not improve the prediction of survival beyond traditional clinical predictors.65 Interestingly, the same authors report that the addition of KL-6 to a multivariate equation that included surfactant protein A (SP-A) and MMP7 significantly improved the prediction of prognosis. Unfortunately, when evaluated prospectively in randomized treatment trials, changes in KL-6 do not correlate with treatment response.62,66 These results emphasize that although increased KL-6 levels may indicate a worse prognosis in ILD patients, larger studies are needed before a definitive statement can be made on its value beyond that of routine clinical practice.

**Surfactant proteins**

Pulmonary surfactants are lipoprotein complexes synthesized by AEC-II and secreted into a liquid layer lining the epithelium, where they decrease the surface tension at the air–liquid interface. The reduction in surface tension allows the lung to inflate with lower transpulmonary pressures and prevents collapse with exhalation. In addition, surfactants play key roles in host defence against pathogens that reach the terminal airways.67 Surfactant proteins A (SP-A) and D (SP-D) have been studied extensively as diagnostic and prognostic biomarkers in patients with ILD. Abnormal surfactant proteins increase alveolar endoplasmic reticulum (ER) stress and trigger the unfolded protein response (UPR), both processes being involved in cell injury and dysfunction.68 Defects in the genes encoding SP-A1 and SP-A2 are associated with familial PF; suggesting that these loci may play a key role in IPF.69–77

Both serum SP-A and SP-D are significantly elevated in IPF patients compared with healthy controls.60,78 Elevated serum SP-A and SP-D levels are thought to reflect a combination of increased permeability of the epithelial barrier and increased secretion from AEC-II. This may also explain why SP-D levels are elevated in patients with acute exacerbations of IPF in comparison to stable disease.79 The utility of serum SP-A and/or SP-D levels in IPF versus other forms of ILD has been evaluated in numerous small trials.80 Although these studies suggest that elevated surfactant protein levels distinguish IPF from other ILD, there remains insufficient evidence to use these proteins as diagnostic markers.

An increased serum level of either surfactant protein at the time of diagnosis has been demonstrated to be independent predictor of death or requirement for lung transplantation.81–83 Interestingly, a prediction model recently published by Song et al. in 118 patients with IPF found that neither serum SP-A or SP-D provided greater prognostic discrimination in comparison to routine clinical assessment with regular pulmonary function testing.65 Similar to KL-6, prospective trials evaluating the efficacy of pirfenidone showed no difference in serum surfactant protein levels between treatment and placebo groups.62,66 Although both SP-A and SP-D have potential as prognostic biomarkers, further evidence is required before regular serum assessments are implemented into routine clinical practice.
Telomeres
The ends of each chromosome are capped by a region of non-coding repetitive nucleotide sequences called telomeres. Telomeres are disposable buffers at the ends of chromosomes that are truncated during cell division; their presence protects the downstream coding segments from being shortened. Telomerase, a ribonucleoprotein complex composed of a specialized reverse transcriptase enzyme (telomerase reverse transcriptase, TERT) and a RNA primer sequence (telomerase RNA component, TERC), acts to protect the terminal ends of chromosomes from degradation. When telomeres reach a critical length, apoptotic or cell senescence pathways are activated. Telomere attrition is a hallmark of cellular ageing and may contribute to impaired AEC integrity. Mutations in both TERT and TERC have been reported in 8–15% of familial PF cases, creating a pathogenic link between short telomeres and PF. Short telomeres are common in IPF, with approximately 25% of patients demonstrating blood leucocyte telomere lengths below the 10th percentile. The utility of telomere length as a diagnostic biomarker, however, is limited by the fact that short telomeres have also been associated with chronic obstructive pulmonary disease (COPD). A recent observational study found that blood leucocyte telomere length was predictive of transplant-free survival, independent of conventional markers. Larger prospective studies are required to further validate telomere length as a mechanistic biomarker of significance.

Mucin 5b
In 2011, Seibold et al. identified a single nucleotide polymorphism in the promoter region of the Mucin 5B (MUC5B) gene (rs35705950), on the short arm of chromosome 11, which was highly associated with both familial PF and sporadic IPF, with odds ratios of 6.3 and 8.3, respectively. What makes the association between the MUC5B promoter variant and PF unique is that this is a rare example of a common genetic variant with a very large genetic effect, as it has been reported in 8–15% of familial PF cases, creating a pathogenic link between short telomeres and PF. Short telomeres are common in IPF, with approximately 25% of patients demonstrating blood leucocyte telomere lengths below the 10th percentile. The utility of telomere length as a diagnostic biomarker, however, is limited by the fact that short telomeres have also been associated with chronic obstructive pulmonary disease (COPD). A recent observational study found that blood leucocyte telomere length was predictive of transplant-free survival, independent of conventional markers. Larger prospective studies are required to further validate telomere length as a mechanistic biomarker of significance.

Vascular endothelial growth factor
Vascular endothelial growth factor (VEGF) is a glycoprotein expressed in AEC that promotes vascular permeability and regulates angiogenesis. Angiogenesis may contribute to the development of IPF. Two separate studies found that BALF VEGF concentrations were significantly lower in patients with IPF in comparison to healthy individuals. Meyer et al. reported that BALF VEGF levels correlated with diffusing capacity in IPF and sarcoidosis. In a cohort of 41 IPF patients, Ando and colleagues found that an increase in serum VEGF levels was associated with poor gas exchange with a high alveolar–arterial oxygen difference (AaDO2). Although serum VEGF levels did not correlate with baseline pulmonary function, an inverse relationship was observed between baseline VEGF and change in vital capacity over time. In addition, IPF patients with serum VEGF levels greater than the median tended to have shorter survival (P = 0.075). Although these results are derived from a single-centre retrospective study, they suggest that serum VEGF may be an effective prognostic biomarker reflecting disease severity and outcome.

Cleaved cytokeratin 18
ER stress leading to activation of the UPR and subsequent AEC apoptosis may play a role in the
pathogenesis of IPF. During apoptosis, a cytoskeletal protein named cytokeratin 18 is cleaved by caspases yielding a fragment—cleaved cytokeratin 18 (cCK18). Using immunohistochemistry, Cha et al. confirmed the presence of cCK18 in AEC in IPF lung, but not in control. Furthermore, cCK18 and mediators of the UPR increased following thapsigargin-triggered ER stress, suggesting that cCK18 may be a suitable surrogate biomarker to assess UPR-induced AEC apoptosis. Although serum cCK18 levels have not been associated with disease severity or outcome, they are significantly elevated in the serum of IPF patients in comparison to normal controls, and patients with either HP or NSIP. This suggests that cCK18 may also be a valuable diagnostic biomarker. Further studies will be necessary to validate this promising molecule.

**Biomarkers associated with extracellular matrix remodelling and fibroproliferation**

**Matrix metalloproteinases**

MMP are a structurally and functionally related family of 23 known zinc-dependent proteases. They play an important role in the pathogenesis of fibrosis by ECM turnover regulation, chemokine metabolism, cell migration and mediator activation. Normally expressed at low levels in healthy tissue, MMP are highly expressed in IPF lungs. MMP1 is responsible for the degradation of fibrillar collagens and is increased in IPF lung tissue and BALF. Serum MMP1 levels are elevated in IPF compared with HP, sarcoidosis and COPD.

MMP7 is the smallest member of the MMP superfamily and is capable of degrading multiple components of the ECM (including signalling molecules and receptors), positioning itself as a key contributor to fibrosis. Using a MMP7 knockout model, Zuo et al. demonstrated that deficient mice were protected from bleomycin-induced PE. Elevated serum MMP7 levels are observed in IPF patients compared with patients with HP sarcoidosis and COPD. In comparison, MMP7 concentrations in BALF and lung tissue samples from IPF patients are similar to those observed in other forms of ILD.

By combining the serum measurements of both MMP1 and MMP7, Rosas et al. were able to distinguish IPF from HP with 96% sensitivity and 87% specificity. This study also demonstrated that serum MMP7 concentrations were elevated in patients with ILA and negatively correlated with forced vital capacity (FVC) and diffusing capacity. Thus, increased MMP7 levels may be relevant screening and prognostic biomarkers, indicative of asymptomatic disease and predictors of disease progression.

Richards and colleagues prospectively analysed the prognostic value of 95 potential biomarkers in a derivation cohort of 140 patients. A subsequent validation cohort of 101 patients demonstrated that high concentrations of MMP7, ICAM-1, IL-8, VCAM-1 and S100A12 were significantly associated with transplant-free survival. Interestingly, the longest transplant-free survival (4.3 years) was seen in patients with low MMP7. The group then developed a personal clinical and molecular mortality prediction index (PCMI) using gender, FVC, diffusing capacity and MMP7 concentration as variables. In the replication cohort, the PCMI was highly predictive of mortality, suggesting that the use of molecular markers may improve clinical predictions in IPF. Finally, in a clinical trial cohort of 438 patients, MMP7 was an independent predictor of survival in a model that included clinical parameters and MUC5B genotype, further solidifying its place as a reliable prognostic biomarker that may augment routine clinical care.

**Periostin**

Periostin is an ECM protein that promotes ECM deposition, mesenchymal cell proliferation and parenchymal fibrosis. Bronchial epithelial cells secrete periostin in response to interleukin-13 (IL-13). Immunohistochemical analyses suggest that expression of periostin is greater in IPF tissue when compared with controls and patients with other ILD. Serum periostin levels are elevated in IPF and correlate with physiological progression. Mouse models demonstrate that periostin is upregulated after bleomycin-induced lung injury, and that periostin-null mice are protected from fibrosis. This evidence has led to a randomized control trial evaluating the use of a monoclonal antibody directed against IL-13 in patients with IPF (NCT01872689). In this trial, the utility of periostin as a biomarker and therapeutic target will hopefully be clarified.

**Osteopontin**

Osteopontin is a phosphorylated glycoprotein that functions as an inflammatory cytokine involved in tissue repair. Mouse models of bleomycin-induced PF demonstrate that osteopontin promotes the migration, adhesion and proliferation of fibroblasts. Pardo et al. provided a mechanically plausible role for osteopontin in the development of fibrosis, with osteopontin playing a key role in fibroblast recruitment and ECM deposition. Osteopontin upregulates MMP7 expression and co-localizes with MMP7 in the AEC of IPF lungs. Osteopontin is elevated in the BALF and serum of IPF patients, but this does not differ from other ILD. In a small study cohort of 17 patients with ILD, osteopontin levels of 300–380 ng/mL distinguished between patients with ILD and controls with 100% specificity and sensitivity. Serum osteopontin concentrations were inversely correlated with PaO2, but not FVC or diffusing capacity. Further studies involving this intriguing molecule are warranted to clarify its utility in clinical practice.

**Circulating fibrocytes**

Fibrocytes are spindle-shaped circulating bone marrow-derived mesenchymal progenitor cells that produce ECM and have the ability to differentiate into fibroblasts and myofibroblasts during wound healing. A pathogenic role for fibrocytes in PF has been demonstrated in murine models where the blockage of fibrocyte recruitment protected against
fibrosis.119 Fibrocyte recruitment involves the CXCL12-CXCR4 and CCL2-CCR2 axes.120 In a cohort of 58 patients with IPF, an increased proportion of circulating fibrocytes was observed in patients with IPF in comparison to controls (2.72% vs 1% of peripheral blood leucocytes, respectively). Furthermore, circulating fibrocytes increase during acute exacerbations of IPF, and have been observed to return to baseline in instances of successful recovery.48 Moeller et al. went on to report that survival was poorer in patients with more than 5% circulating fibrocytes. The majority of those with elevated levels, however, had experienced an acute exacerbation, which in itself carries prognostic significance.

Recently, Trimble and colleagues evaluated whether the number and phenotype of circulating fibrocytes predicted outcome in patients with Hermansky–Pudlak syndrome, a genetic form of ILD that manifests in early adulthood.121 Blood fibrocyte counts were markedly elevated in a subset of subjects with Hermansky–Pudlak syndrome who had ILD but not subjects without lung disease. Elevations in fibrocyte counts were strongly associated with death from ILD. Further research is required to better understand the role of circulating fibrocytes in both the pathogenesis of IPF and as diagnostic and prognostic biomarkers.

**Biomarkers associated with immune dysfunction**

**YKL-40**

YKL-40 is a chitinase-like protein that regulates the proliferation and survival of many cell types. Increased levels are observed in many inflammatory disorders, including fibrotic liver disease, sarcoidosis, COPD and asthma.122–124 Although the exact role of YKL-40 in the pathogenesis of IPF is unclear, it appears to facilitate the release of fibrotic and inflammatory mediators from alveolar macrophages and has a mitogenic effect on lung fibroblasts.125 This signalling appears to be mediated through Th2-lymphocyte/IL-13 signalling pathways.126 In comparison to normal controls, YKL-40 is significantly elevated in the lung tissue, BALF and serum of patients with IPF.49,127 Although there is a poor correlation between serum and BALF YKL-40 levels, a small single-centre study demonstrated that both are associated with worse survival.127 Interestingly, in this same trial, patients with both low-serum and low-BALF YKL-40 levels demonstrated no IPF related mortality.25.51,52 Associated with death (hazard ratio 7.98, 95% CI: 2.49–25.51, P = 0.005). These data suggest that CCL18 may not only have potential as a prognostic marker, but that it could also possibly be used serially in the ongoing assessment of patients.

**Toll-like receptor 3**

Toll-like receptor 3 (TLR3) is a receptor that mediates the innate immune response to tissue injury, inflammation and infection. O’Dwyer et al. investigated the function of TLR3 in primary human lung fibroblasts from the lungs of patients with IPF who were either wild-type, heterozygous or homozygous for the TLR3 L412F polymorphism.40 Defective cytokine, type-1 interferon and fibroproliferative responses were observed in those carrying at least one copy of the L412F polymorphism. Furthermore, the study confirmed an association with early mortality and an accelerated decline in lung function in those carrying the mutation. Such information may be critical in identifying those patients with a rapidly progressive phenotype.

**Adaptive immunity and inflammation in idiopathic pulmonary fibrosis**

As previously described, inflammation is not a prominent histopathological finding in IPF. This being said, it cannot be ignored that inflammation and immune activation are found in IPF lungs, and that immune markers may provide useful information for patient stratification and innovative treatment strategies. For instance, IgG autoantibodies to heat shock protein 70 (HSP70) have been found in the serum of 25% of patients with IPF compared with 3% of controls.24 Animal models of bleomycin-induced PF propose that HSP70 plays a protective role and attenuates injury, inflammation and fibrosis. Preliminary evidence suggests that auto-antibodies targeted against HSP70 may play a role in IPF, and that positive anti-HSP70 serology is associated with worse lung function and survival.24 This is supported by evidence which suggests that circulating B cells in IPF subjects are more differentiated, with increased plasmablast proportions, in comparison to healthy controls.28 The degree of differentiation is inversely correlated to FVC. CD20+ B-cell aggregates, parenchymal and perivascular immune complexes, and complement deposition have also been shown to be increased in IPF patients. Plasma concentrations of circulating BLyS, a trophic factor necessary for B-cell survival,
maturation and antibody production, are significantly greater in IPF patients than in normal controls. BLyS concentrations were highest among those subjects with pulmonary hypertension and also predicted transplant-free survival at 1 year.25

Implicated in the pathogenesis of numerous immunological disorders, C-X-C chemokine 13 (CXCL13) is a critical agent for B-cell trafficking to lymphoid aggregates and inflammatory foci. Circulating CXCL13 levels have been demonstrated to be elevated in patients with IPF, particularly in the setting of pulmonary hypertension.31 Patients developing or experiencing an acute exacerbation also had significantly greater concentrations of CXCL13. Furthermore, both baseline CXCL13 levels and increases in CXCL13 over time were associated with reduced survival.31

T cells are the predominant mononuclear cell type isolated from IPF lungs, and have been associated with disease severity and survival.133 The level of expression of genes associated with the costimulatory signal during T-cell activation, including CD28, ICOS, LCK and ITK, was shown to predict prognosis in two IPF cohorts.22 Furthermore, increased CD8 + T-cell density in the pulmonary interstitium and the downregulation of serum CD28 CD4 + cells both correlate with progressive disease.52,132 These findings provide insight into the complex pathogenesis of IPF and the plethora of ‘inflammatory’ biomarkers that potentially have clinical relevance.

USING GENE EXPRESSION ANALYSIS TO ENHANCE PERSONALIZED MEDICINE

Gene expression profiling has been widely utilized to generate new hypotheses regarding molecular events that drive the natural history of IPF. Such analyses have the potential to identify novel and clinically relevant molecular biomarkers that can provide insight into the aberrant processes associated with disease pathology. For instance, using microarray analysis on autopsy-derived or explanted lung tissue, Konishi et al. found that IPF patients with stable disease or acute exacerbation exhibited overall similar gene expression signatures in comparison to control samples.133 However, on direct comparison of the IPF subgroups, more than 500 of the 30 000 on the array were differentially expressed. Interestingly, the gene expression profile in exacerbating patients did not exhibit an increase in inflammatory response, but rather a signal pointing to excessive AEC proliferation and apoptosis. Taken together, these results indicate the central role of the pulmonary epithelium in acute exacerbations. In another study, Boon et al. evaluated the differences between the molecular phenotypes of patients with stable versus progressive IPF.134 In patients with progressive disease, a total of 102 transcripts were noted to be upregulated, many of them also engaged in epithelial cell proliferation and activation.

Recently, DePianto and colleagues performed genome-wide transcriptomic analysis of lung biopsy tissue from 40 patients with IPF and 8 controls.135 Two clusters of co-regulated genes were found to be upregulated in IPF. The first cluster comprised genes related to the bronchiolar epithelium, and the second consisted of T- and B-cell markers, Fc receptor genes, and chemokines. These clusters of genes were termed the ‘bronchiolar signature’ and ‘lymphoid signature’, respectively. Within these signatures, the genes encoding MMP3 and CXCL13 were identified as being upregulated and found to be negatively correlated with survival over a 3-year follow-up period.

Many other research findings have been published in this area, but covering these would go beyond the scope of this review.

CONCLUSIONS

In the past two decades, numerous putative molecular biomarkers for IPF have been identified. These may provide insight into the activation and propagation of fibrosis and hold great promise in advancing the care of the patients who suffer from IPF. Considering the complexity of the pathogenesis of IPF, it is clear that a single biomarker assay is unlikely to have transformative effects on clinical practice. In contrast, implementation of multi-marker panels, assessing the many pathobiological processes involved in IPF, may provide clinicians with the information needed to improve patient care. Unfortunately, the majority of studies evaluating molecular biomarkers in IPF are underpowered, retrospective and lack necessary validation. These facts preclude the rapid incorporation of biomarkers into routine clinical practice. Which of these putative biomarkers will have the greatest clinical utility remains uncertain. To date, only MUC5B and MMP7 have been studied in large validation cohorts, and perhaps more importantly demonstrated clinical utility beyond that achieved with conventional clinical predictors.

Diagnostic and therapeutic biomarkers will continue to evolve as larger cohorts are studied in prospective interventional trials. Of great interest will be the isolation of specific ‘subphenotypes’ of IPF defined by a distinct molecular signature. Stratifying patients based on their unique biology may provide clinicians with the personalized information required to optimize diagnostic, prognostic and therapeutic assessments. Only then will the visionary goal of personalized medicine come to fruition and advance the care of all those who suffer from IPF.

REFERENCES


Molecular classification of IPF


Molecular classification of IPF


118 Bellini A, Mattoli S. The role of the fibrocyte, a bone marrow-derived mesenchymal progenitor, in reactive and reparative fibroses. Lab Invest 2007; 87: 858–70.


© 2015 Asian Pacific Society of Respirology

Respirology (2015)