

Idiopathic pulmonary fibrosis: a paradigm of late-onset, single-gene human disease?

Stephen J Chapman,¹ Matt P Wise²

To cite: Chapman SJ, Wise MP. Idiopathic pulmonary fibrosis: a paradigm of late-onset, single-gene human disease?. *BMJ Open Res* 2014;1:e000070. doi:10.1136/bmjresp-2014-000070

Received 7 November 2014
Accepted 10 November 2014

Idiopathic pulmonary fibrosis (IPF) remains as one of the most devastating respiratory diseases. Incompletely understood and difficult to characterise clinically, there are only limited treatments of proven benefit available and the prognosis is typically poor. Early studies were hampered by a lack of clear classification and diagnostic precision among interstitial lung diseases, and until recently the pathogenesis of IPF has only been poorly understood. However, the last few years have seen major advances in our understanding of the molecular basis of this disease, and in particular a series of landmark studies have begun to uncover the genetic basis of IPF.

Familial forms of IPF are well recognised, although apparently account for only a minority (approximately 10%) of adult cases.¹ A number of genes have been implicated in familial IPF through linkage analysis, including genes involved in regulation of telomere length (eg, *TERT* and *TERC*) and those encoding surfactant proteins (eg, *SFTPA2* and *SFTPC*) and mucin 5B (*MUC5B*).^{1–3} Recently, sporadic (non-familial) IPF has been subject to extensive genome-wide analysis in an attempt to identify underlying genetic susceptibility variants. Genome-wide association studies (GWAS) comprise the genotyping of hundreds of thousands of common genetic variants (termed polymorphisms) spread at intervals across the genome, and the frequency of these variants are typically compared between disease cases and healthy controls. The attraction of such studies is that they are not based on existing assumptions about disease pathogenesis—which are usually incomplete and may be misleading—and hence have the potential to identify associations with previously unsuspected genes and pathways. The cost of this unbiased approach is that large sample sizes (typically in the thousands) with independent replication case–control collections are required in

order to power robust associations, given the very large number of statistical tests performed.

Such GWAS approaches have confirmed the role of genetic variation in *TERT* and *TERC* and *MUC5B* in susceptibility to sporadic as well as familial IPF.^{4–6} These studies have also identified novel associations, for example with polymorphism in the genes *TOLLIP* involved in innate immune signalling and *DSP* and *DPP9* involved in cell–cell adhesion.^{5 6} Furthermore they have extended the phenotypic spectrum by describing associations between *MUC5B* polymorphism and early-stage interstitial lung abnormalities in the general population as well as outcome from IPF.^{7 8} The role of *MUC5B* in particular in sporadic IPF is remarkable, as the functional *MUC5B* promoter polymorphism appears to exert a disproportionately large effect size (an approximately eightfold increase in risk) for such a common variant.⁹ Nevertheless, as the heritability of IPF is unknown, it is unclear how much of the genetic component of this disease is accounted for by these variants. For other complex human adult diseases, the common polymorphisms identified through GWAS together appear to account for only a small proportion of known heritability. A popular explanation for the remaining ‘missing heritability’ is multiple different rare mutations of individually large effect size, which sit below the frequency cut-off identifiable through GWAS: identification of such rarer variants requires gene resequencing approaches.¹⁰

Aaron Hamvas and colleagues now report a candidate gene resequencing study of adults with IPF, compared with a disease-control group of patients with chronic obstructive pulmonary disease (COPD) and population-based database controls.¹¹ Six genes were selected for resequencing based on their known involvement in cases of familial IPF and/or childhood interstitial lung



CrossMark

¹Oxford University Hospitals and Oxford Biomedical Research Centre, Oxford, UK
²University Hospital of Wales, Cardiff, UK

Correspondence to

Dr Stephen J Chapman;
schapman@well.ox.ac.uk

disease: genes encoding surfactant proteins (*SFTPA2*, *SFTPC*), ATP-binding cassette member A3 (*ABCA3*), telomerase (*TERT*), thyroid transcription factor (*NKX2-1*) and mucin 5B (*MUC5B*). Unlike association studies of common polymorphism, the analysis of single variants is probably not appropriate for rare variant analysis and lacks statistical power, and instead a collapsing analysis was performed, which aggregated rare variants to assess mutational load across genes. Remarkably, a significant excess of rare, computationally predicted deleterious mutations were identified in the genes *SFTPC* and *TERT* in adult patients with IPF when compared to patients with COPD and population-based controls.¹¹ Common polymorphisms in these genes were present at similar frequencies between the groups. Interestingly, the majority of patients with IPF with a rare, predicted deleterious mutation (11 of 14 cases) did not report a positive family history and had apparently sporadic, adult-onset IPF.¹¹ The association between IPF and the common *MUC5B* promoter polymorphism was again replicated; of note there was no demonstrable enrichment of rare coding mutations in *MUC5B* in IPF cases compared to controls, although exon 49 was not sequenced.

There did not appear to be any interaction between these rare variants and the common *MUC5B* promoter polymorphism, although it should be noted that the sample size is relatively small and the study lacks power to demonstrate such epistatic (gene–gene) interactions. The study has other limitations common to IPF research, in particular that the diagnosis of IPF can be difficult to make with absolute certainty and may be mistaken for other idiopathic interstitial pneumonias or other forms of interstitial lung disease, which may have led to misclassification of cases. The recruitment of candidates for lung transplantation into the current study has resulted in patients at the younger end of the disease spectrum (average age 54 years), who may arguably be enriched for host genetic susceptibility factors. Nevertheless, the findings raise the intriguing possibility that rare genetic mutations of individually large effect may account for a significant fraction of adults with apparently sporadic IPF. Taken together with the recently described major role of *MUC5B* polymorphism in IPF susceptibility, an emerging paradigm is that adult-onset IPF may be in the large part a genetically determined condition, with some of the longest latency periods observed in human single-gene disease. Why do such large-effect mutations not manifest as clear-cut familial disease? In some cases they may represent familial IPF but a clear family history of other affected members is not apparent; alternatively the size of effect may not be sufficient to result in classical Mendelian segregation, or in some cases they may even represent de novo mutations which have arisen in affected individuals.

As with all genetic association studies, the findings require confirmatory replication in an independent study group. In addition it would be worthwhile

extending the list of candidates to include other genes recently found to associate with IPF at genome-wide level, as in some cases an apparently disease-associated common polymorphism may simply be a marker which reflects a true association driven by causative rare variants on the same haplotype.⁵ Indeed, the study suggests that whole exome-based or genome-based sequencing approaches may be fruitful in this disease, and perhaps lead to the identification of novel mutations in previously unsuspected genes and pathways, some of which may be amenable to drug therapy. Such larger studies may also uncover the genetic architecture of IPF, in particular the relative contribution of common and rare variants, as well as possible epistatic interactions between variants of different frequency and potential gene–environment interactions, for example with cigarette smoking and ageing.

Even following the reclassification of interstitial lung disease, IPF is likely to be a rather crude umbrella term which contains a number of clinically and radiologically indistinguishable molecular subphenotypes within it. In addition to the use of genetic association studies, complementary approaches to subphenotype IPF include proteomics and microarray analysis of lung and peripheral blood cell gene expression and DNA methylation profile.^{12–15} Identification of these molecular subphenotypes will lend a new level of precision to clinical trials and outcome prediction, as well as valuable insights into the underlying pathophysiology and indeed aetiology—perhaps at last taking the ‘idiopathic’ out of pulmonary fibrosis.

Competing interests None.

Provenance and peer review Not commissioned; internally peer reviewed.

Open Access This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>

REFERENCES

1. Steele MP, Schwartz DA. Molecular mechanisms in progressive idiopathic pulmonary fibrosis. *Annu Rev Med* 2013;64:265–76.
2. Armanios MY, Chen JJ, Cogan JD, *et al.* Telomerase mutations in families with idiopathic pulmonary fibrosis. *N Engl J Med* 2007;356:1317–26.
3. Seibold MA, Wise AL, Speer MC, *et al.* A common *MUC5B* promoter polymorphism and pulmonary fibrosis. *N Engl J Med* 2011;364:1503–12.
4. Mushiroda T, Wattanapokayakit S, Takahashi A, *et al.* A genomewide association study identifies an association of a common variant in *TERT* with susceptibility to idiopathic pulmonary fibrosis. *J Med Genet* 2008;45:654–6.
5. Fingerlin TE, Murphy E, Zhang W, *et al.* Genome-wide association study identifies multiple susceptibility loci for pulmonary fibrosis. *Nat Genet* 2013;45:613–20.
6. 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;1:309–17.
7. Hunninghake GM, Hatabu H, Okajima Y, *et al.* *MUC5B* promoter polymorphism and interstitial lung abnormalities. *N Engl J Med* 2013;368:2192–200.

8. Peljto AL, Zhang Y, Fingerlin TE, *et al.* Association between the *MUC5B* promoter polymorphism and survival in patients with idiopathic pulmonary fibrosis. *JAMA* 2013;309:2232–9.
9. Putman RK, Rosas IO, Hunninghake GM. Genetics and early detection in idiopathic pulmonary fibrosis. *Am J Resp Crit Care Med* 2014;189:770–8.
10. Chapman SJ, Hill AVS. Human genetic susceptibility to infectious disease. *Nat Rev Genet* 2012;13:175–88.
11. Coghlan MA, Shifren A, Huang HJ, *et al.* Sequencing of idiopathic pulmonary fibrosis-related genes reveals independent single gene associations. *BMJ Open Respir Res* 2014;1:e000057.
12. Richards TJ, Kaminski N, Baribaud F, *et al.* Peripheral blood proteins predict mortality in idiopathic pulmonary fibrosis. *Am J Resp Crit Care Med* 2012;185:67–76.
13. Yang IV, Coldren CD, Leach SM, *et al.* Expression of cilium-associated genes defines novel molecular subtypes of idiopathic pulmonary fibrosis. *Thorax* 2013;68:1114–21.
14. Herazo-Maya JD, Noth I, Duncan SR, *et al.* Peripheral blood mononuclear cell gene expression profiles predict poor outcome in idiopathic pulmonary fibrosis. *Sci Transl Med* 2013;5:205ra136.
15. Sanders YY, Ambalavanan N, Halloran B, *et al.* Altered DNA methylation profile in idiopathic pulmonary fibrosis. *Am J Resp Crit Care Med* 2012;186:525–35.