Published in final edited form as:
Lancet Respir Med. 2013 June ; 1(4): 309–317. doi:10.1016/S2213-2600(13)70045-6.

Genetic variants associated with idiopathic pulmonary fibrosis susceptibility and mortality: a genome-wide association study

Imre Noth[#], **Yingze Zhang**[#], **Shwu-Fan Ma**[#], **Carlos Flores**, **Mathew Barber**, **Yong Huang**, **Steven M Broderick**, **Michael S Wade**, **Pirro Hysi**, **Joseph Scuirba**, **Thomas J Richards**, **Brenda M Juan-Guardela**, **Rekha Vij**, **MeiLan K Han**, **Fernando J Martinez**, **Karl Kossen**, **Scott D Seiwert**, **Jason D Christie**, **Dan Nicolae**[#], **Naftali Kaminski**[#], and **Joe G N Garcia**[#]

Section of Pulmonary and Critical Care Medicine (Prof I Noth MD, S-F Ma PhD, M Barber PhD, Y Huang MS, S M Broderick BS, R Vij MD) and Section of Genetic Medicine (Prof D Nicolae PhD), University of Chicago, Chicago, IL, USA; Dorothy P and Richard P Simmons Center for Interstitial Lung Disease, Division of Pulmonary, Allergy and Critical Care Medicine, University of Pittsburgh, Pittsburgh, PA, USA (Y Zhang PhD, J Scuirba BS, T J Richards PhD, B M Juan-Guardela MD, Prof N Kaminski MD); CIBER de Enfermedades Respiratorias, Instituto de Salud Carlos III, Madrid, Spain (C Flores PhD); Research Unit, Hospital Universitario Nuestra Señora de Candelaria, Tenerife, Spain (C Flores); Institute for Personalized Respiratory Medicine, University of Illinois at Chicago, Chicago, IL, USA (M S Wade MS, Prof J G N Garcia MD); Department of Twin Research and Genetic Epidemiology, King's College London, London, UK (P Hysi PhD); Division of Pulmonary and Critical Care Medicine, University of Michigan, Ann Arbor, MI, USA (M K Han MD, Prof F J Martinez MD); InterMune, Brisbane, CA, USA (K Kossen PhD, S D Seiwert PhD); and Pulmonary, Allergy and Critical Care Division, University of Pennsylvania Medical Center, Pennsylvania, PA, USA (J D Christie MD)

[#] These authors contributed equally to this work.

Correspondence to: Prof Imre Noth, Section of Pulmonary and Critical Care Medicine, University of Chicago, Chicago, IL 60637, USA inoth@medicine.bsd.uchicago.edu.

Contributors IN, YZ, and S-FM collected data and wrote each draft of the report. CF, MB, YH, PH, TJR, and DN analysed and interpreted data. SMB, MSW, JS, and BMJ-G made technical contributions. IN, RV, MKH, FJM, KK, SDS, JDC, and NK recruited patients and obtained consent, and maintained phenotype databases at each medical centre. IN, NK, and JGNG were the principal investigators and conceived the project. All authors read and commented on the report.

Conflicts of interest IN is a paid consultant for ImmuneWorks, Stromedix, and Sunovion; has received speaker's honoraria from GlaxoSmithKline; is an inventor on patent applications for the use of peripheral blood proteins in prediction of IPF outcomes and for the use of an mRNA signature in the prediction of IPF outcomes; and has contracts for clinical studies of IPF with Stromedix, Sanofi, and Boehringer Ingelheim. FJM has received payment for participation in advisory boards for chronic obstructive pulmonary disease or IPF topics for Able Associates, Actelion, Almirall, Bayer, GlaxoSmithKline, Ikaria, Janssen, MedImmune, Merck, Pearl, Pfizer, and Vertex; has been a paid consultant for the American Institute for Research, AstraZeneca, Bayer, Carden Jennings, CardioMEMS, Grey Healthcare, HealthCare Research and Consulting, Janssen, Merion, Nycomed/Takeda, and Sudler and Hennessey; been a member of steering committees for studies sponsored by Actelion, Centocor, Forest, GlaxoSmithKline, Gilead, Mpex, Nycomed/Takeda; has participated in Food and Drug Administration mock panels for Boehringer Ingelheim, Forest, and GlaxoSmithKline; served on speaker's bureaus or in continuing medical education activities sponsored by the American College of Chest Physicians, American Lung Association, AstraZeneca, Bayer, William Beaumont Hospital, Boehringer Ingelheim, Benter for Health Care Education, CME Incite, Forest, France Foundation, GlaxoSmithKline, Lovelace, MedEd, MedScape/WebMD, National Association for Continuing Education, Network for Continuing Medical Education, Nycomed/Takeda, Projects in Knowledge, St Luke's Hospital, University of Illinois at Chicago, University of Texas Southwestern, University of Virginia, and UpToDate; and has served on data and safety monitoring boards for Biogen and Novartis; and received royalties from Castle Connolly and Informa. FJM's institution has received funds from the National Institutes of Health for studies of chronic obstructive pulmonary disease and IPF. KK and SDS are employees of InterMune and own stock in the company. NK has been a paid consultant for Sanofi, Stromedix, Vertex, and InterMune; is a recipient of investigator-initiated grants from Celgene and Gilead; and is an inventor on a patent application on the use of peripheral blood proteins in prediction of IPF outcomes. The other authors declare that they have no conflicts of interest.

For more on the **database of genotypes and phenotypes** see http://www.ncbi.nlm.nih.gov/gap

For more on the **Lung Genomics Research Consortium** see http://www.lung-genomics.org

Summary

Background—Idiopathic pulmonary fibrosis (IPF) is a devastating disease that probably involves several genetic loci. Several rare genetic variants and one common single nucleotide polymorphism (SNP) of *MUC5B* have been associated with the disease. Our aim was to identify additional common variants associated with susceptibility and ultimately mortality in IPF.

Methods—First, we did a three-stage genome-wide association study (GWAS): stage one was a discovery GWAS; and stages two and three were independent case-control studies. DNA samples from European-American patients with IPF meeting standard criteria were obtained from several US centres for each stage. Data for European-American control individuals for stage one were gathered from the database of genotypes and phenotypes; additional control individuals were recruited at the University of Pittsburgh to increase the number. For controls in stages two and three, we gathered data for additional sex-matched European-American control individuals who had been recruited in another study. DNA samples from patients and from control individuals were genotyped to identify SNPs associated with IPF. SNPs identified in stage one were carried forward to stage two, and those that achieved genome-wide significance ($p<5 \times 10^{-8}$) in a meta-analysis were carried forward to stage three. Three case series with follow-up data were selected from stages one and two of the GWAS using samples with follow-up data. Mortality analyses were done in these case series to assess the SNPs associated with IPF that had achieved genome-wide significance in the meta-analysis of stages one and two. Finally, we obtained gene-expression profiling data for lungs of patients with IPF from the Lung Genomics Research Consortium and analysed correlation with SNP genotypes.

Findings—In stage one of the GWAS (542 patients with IPF, 542 control individuals matched one-by-one to cases by genetic ancestry estimates), we identified 20 loci. Six SNPs reached genome-wide significance in stage two (544 patients, 687 control individuals): three *TOLLIP* SNPs (rs111521887, rs5743894, rs5743890) and one *MUC5B* SNP (rs35705950) at 11p15.5; one *MDGA2* SNP (rs7144383) at 14q21.3; and one *SPPL2C* SNP (rs17690703) at 17q21.31. Stage three (324 patients, 702 control individuals) confirmed the associations for all these SNPs, except for rs7144383. Linkage disequilibrium between the *MUC5B* SNP (rs35705950) and *TOLLIP* SNPs (rs111521887 [r^2=0.07], rs5743894 [r^2=0.16], and rs5743890 [r^2=0.01]) was low. 683 patients from the GWAS were included in the mortality analysis. Individuals who developed IPF despite having the protective *TOLLIP* minor allele of rs5743890 carried an increased mortality risk (meta-analysis with fixed-effect model: hazard ratio 1.72 [95% CI 1.24–2.38]; p=0.0012). *TOLLIP* expression was decreased by 20% in individuals carrying the minor allele of rs5743890 (p=0.097), 40% in those with the minor allele of rs111521887 ($p=3.0 \times 10^{-4}$), and 50% in those with the minor allele of rs5743894 ($p=2.93 \times 10^{-5}$) compared with homozygous carriers of common alleles for these SNPs.

Interpretation—Novel variants in *TOLLIP* and *SPPL2C* are associated with IPF susceptibility. One novel variant of *TOLLIP*, rs5743890, is also associated with mortality. These associations and the reduced expression of *TOLLIP* in patients with IPF who carry *TOLLIP* SNPs emphasise the importance of this gene in the disease.

Funding—National Institutes of Health; National Heart, Lung, and Blood Institute; Pulmonary Fibrosis Foundation; Coalition for Pulmonary Fibrosis; and Instituto de Salud Carlos III.

Introduction

Idiopathic pulmonary fibrosis (IPF) has a low incidence (4.6–16.3 cases per 100 000 person-years),[1,2] but is a devastating disease of unknown aetiology that is characterised by an interstitial fibrotic process and high mortality.[3] Lung transplantation is the only treatment option that successfully improves survival; immunosuppression regimens have been shown to be harmful.[4] Several rare genetic variants mostly associated with familial IPF and a

common single nucleotide polymorphism (SNP) of *MUC5B* in sporadic IPF have been identified, suggesting that genetic factors contribute to disease.[5,6] As yet, no genetic variant has been associated with IPF outcomes. Therefore, identification of genetic variants associated with susceptibility to IPF and alleles involved in the heterogeneity of disease course and mortality is an important research focus.

Mutations in *TERT* or *TERC* result in telomere shortening and are associated with both familial and non-familial IPF.[7] A previous genome-wide association study (GWAS) examining about 250 000 SNPs in only 159 IPF cases[8] supported this association. Rare heterozygous variants in *SFTPA2* and *SFTPC* have also been implicated in familial IPF.[9] These findings suggest that the aetiology of IPF might involve several genetic loci.

GWAS continues to be the method of choice for identification of common genetic variants associated with complex diseases.[10] In view of the power of GWAS to detect expected effects for common variants with realistic sample sizes,[8] we postulated that an independent GWAS of IPF would identify novel polymorphisms associated with disease susceptibility and allow assessment of whether these loci were also associated with mortality. Although new knowledge of IPF susceptibility could provide targets for novel treatment strategies, identification of genes associated with mortality in patients with IPF might draw attention to potential therapeutic targets and enable prognostication.

Methods

Study design and samples

First, we did a three-stage GWAS (figure 1). Stage one was a discovery GWAS in European Americans for susceptibility to IPF. Identified susceptibility loci were followed up for replication in two independent case-control association studies (stages two and three). Second, we assessed the association between susceptibility loci and mortality in three case series with follow-up data selected from stages one and two (figure 1). Finally, we assessed the correlation between SNP genotypes and gene expression in lung tissue with paired data (figure 1). Complete methods are provided in the appendix.

Patients with IPF in stage one of the discovery GWAS were clinically characterised at the University of Chicago and University of Pittsburgh, via the Lung Tissue Research Consortium (LTRC), and from the Correlating Outcomes with biomedical Markers to Estimate Time-progression in IPF (COMET) study. DNA samples were obtained from each individual. Independent patients with IPF in stage two were obtained from the University of Chicago, InterMune,[11] Lung Transplant Outcomes Group (LTOG) cohort,[12] and LTRC. Independent patients for stage three were from LTOG and the AntiCoagulant Effectiveness in Idiopathic Pulmonary Fibrosis (ACE-IPF) study.[13] DNA samples were obtained from each individual. The timeframe of sample collection varied by cohort. All DNA samples of individuals with IPF used in each stage for association studies were of European-American descent.

InterMune, ACE-IPF, COMET, and LTRC had guidelines for diagnosis of IPF, all of which were adapted from 2000 guidelines from the American Thoracic Society and European Respiratory Society.[14] All patients from the University of Chicago and University of Pittsburgh underwent similar diagnostic review in accordance with the 2000 and 2011 guidelines,[3,14] with each institution engaging in the recommended multidisciplinary (radiology, pathology, and clinical) approach to exclude an alternative diagnosis, as recommended by the 2011 guidelines.[3] Site investigators individually diagnosed patients with IPF in LTOG.

All eligible patients were at least 35-years-old and reported having symptoms of idiopathic interstitial pneumonia for at least 3 months. A high-resolution CT scan that showed definite or probable usual interstitial pneumonitis was necessary for inclusion.[3] A surgical lung biopsy sample to confirm usual interstitial pneumonitis could be obtained if the diagnosis was in doubt. Patients with clinically significant exposure to known fibrogenic agents and those with other known causes of interstitial lung disease were excluded before study entry.

For control individuals in stage one of the GWAS, we gathered data from European Americans from the database of genotypes and phenotypes; additional healthy individuals were recruited from the University of Pittsburgh to increase the available sample size. We selected a subset of control individuals that matched individual IPF cases by means of genome-wide genetic ancestry estimates for downstream case-control analysis.

For control individuals of the replication studies in stages two and three, we gathered data for additional sex-matched European-American individuals of any age who were recruited from 2005 to 2012, as part of the Translational Research in the Department of Medicine study at the University of Chicago. Although the source was the same in all stages, all control individuals were of European-American descent and were independent of each other.

For the mortality analysis, we included all patients with IPF in stages one and two who were genotyped successfully and for whom follow-up data were available. Finally, we obtained gene-expression profiling data for lungs of patients with IPF from the Lung Genomics Research Consortium (LGRC; appendix p 5). We used samples that had paired data for gene expression and genotypes of SNPs that we had identified in the GWAS.

The study was approved by institutional review boards and ethics committees at participating centres. All participants provided written informed consent.

Genotyping and statistical analysis

Genotyping in stage one of the GWAS was done with the Genome-Wide Human SNP 6·0 array (Affymetrix, Santa Clara, CA, USA). Genotyping in stages two and three was done with the iPLEX Gold Platform (Sequenom, San Diego, CA, USA). IMPUTE software (version 2) was used for haplo type-based inference (ie, imputation) of missing genotypes in regions for which genotype information was not covered by the SNP array, with data from the 1000 Genomes Project as a reference. Association testing was done with SNPTEST software (version 2.3; appendix p 3). To assess population stratification—ie, whether one ance stral group was over-represented—we under took inflation of test statistics (SNPTEST; appendix p 3).

SNPs compiled from the loci identified in stage one were carried forward to stage two (appendix p 4). Because these selected SNPs were all based on imputed data, results were validated by regenotyping stage one samples with the iPLEX Gold Platform. SNPs that achieved genome-wide significance[15] (overall $p<5 \times 10^{-8}$) were carried forward to stage three (appendix p 4).

Linkage disequilibrium between SNPs in identified regions was assessed with pairwise r^2 measures.[16] The mode of inheritance for the SNPs (ie, dominant or recessive) was established by comparing the odds ratios of the heterozygous and at-risk homozygous genotypes. A conditional regression-based analysis of the effects of the SNPs on IPF susceptibility was implemented in R (version 2.15.1) to provide evidence for the independence of association signals.[17]

We did Cox regression analyses using the Survival (version 2.36–10) and cmprsk (version 2.2–4) packages for R (version 2.15.0) on the IPF susceptibility loci that were significant in the meta-analysis of stages one and two of the GWAS. Time at risk was defined as the interval between the date of enrolment in a specific study and date of last follow-up, lung transplantation, or death. Patients who underwent lung transplantation were censored at time of transplant, because the procedure could have been a confounder of survival.

We did univariate and multivariable analyses, including relevant demographic and clinical variables. We selected the best-fitting model on the basis of Akaike information criteria. One aggregate result for each SNP was obtained with METASOFT (version 2.0.1), by means of a meta-analysis applying both fixed-effects and random-effects models[18] to account for the different available follow-up data in the case series studied. We did additional sensitivity analyses, including patients who had undergone trans plantation as non-survivors (transplant-free survival), or restricting the analyses to non-transplanted cases only or treating transplantation as a competing risk for mortality (appendix pp 4–5). Additionally, we did an assessment of drug–gene interactions (appendix pp 7–8).

In the analysis of data from LTRC, we stratified gene expression into two groups according to presence or absence of the minor allele for SNPs identified by the GWAS. We then compared the two groups with unequal variance (Welch's) *t* test with the stat package in R.

Role of funding source

The sponsors of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report. The corresponding author had full access to all the data in the study and had final responsibility for the decision to submit for publication.

Results

Table 1 shows the sample sizes for each stage of the GWAS. As expected, similar to a large single-centre experience for first presentation of IPF,[19] the cases included in stage one had a wide range of disease severity and age (table 2). Patients in stage two had similar overall severity to those in stage one. Patients in stage three had more severe disease than did those in stages one and two, as assessed by forced vital capacity (FVC) and diffusing capacity for carbon monoxide (D_LCO). However all IPF cases met diagnostic criteria,[3] and patients in each stage were of similar ages and similar proportions were men.

After completion of sample quality control and genotype filtering (appendix p 3), we retained 542 European Americans with IPF for stage one (table 2) and 542 control individuals who matched individual patients with IPF by means of genome-wide genetic ancestry estimates. We used 555 432 high-quality genotyped variants for imputation, resulting in 10 601 812 accurately imputed variants with minor allele frequencies of greater than 5% for downstream association studies. Inflation was modest (λ=1.06; appendix p 18), indicating that the results were not significantly confounded by population stratification. We identified 20 genomic loci in stage one, from which we compiled 52 SNPs, consisting of genotyped SNPs, imputed SNPs, and tagging SNPs (appendix pp 3–5, 9–13).

Testing of SNPs in stage two nominally replicated many of the associations with IPF susceptibility detected in stage one. Five imputed SNPs and the previously identified *MUC5B* susceptibility SNP[5,6] reached genome-wide significance in a joint analysis of the two stages: three *TOLLIP* SNPs (rs111521887, rs5743894, rs5743890) and one *MUC5B* SNP (rs35705950) at 11p15.5; one *MDGA2* SNP(rs7144383) at 14q21.3; and one *SPPL2C* SNP (rs17690703) at 17q21.31 (table 3). We positively validated these results by regenotyping available samples in stage one (appendix p 6).

Stage three replicated the associations identified in stage two, except that with rs7144383 in *MDGA2* (table 3). Patients who carried minor alleles of the SNPs had effects in the same directions (ie, increased risk of IPF or protection against the disorder) in all stages: minor alleles rs35705950_T, rs111521887_G, and rs5743984_G constituted a risk for IPF; and rs5743890_G and rs17690703_T protected against the disorder (table 3). In the 11p15.5 locus, linkage disequilibrium between the *MUC5B* promoter SNP (rs35705950) and *TOLLIP* SNPs (rs111521887 [r^2=0.07], rs5743894 [r^2=0.16], and rs5743890 [r^2=0.01]) was low, indicating that the signals of association for *TOLLIP* SNPs are independent from that of *MUC5B* (appendix p 4). When we assessed linkage disequilibrium between *TOLLIP* SNPs, we established that at least two are independent (r^2=0.97 between rs111521887 and rs5743894; r^2=0.04 between rs111521887 or rs5743894 and rs5743890). Moreover, the mode of effect for the *MUC5B* SNP (dominant) was different from that for the *TOLLIP* SNPs (additive or recessive), providing additional evidence that these are independent signals. Lastly, in a conditional regression-based analysis, we combined genotypes according to the mode of inheritance and, although the *MUC5B* SNP had the strongest signal ($p=2 \times 10^{-16}$), the *TOLLIP* SNPs were also significantly associated with IPF (p=0.05 for all three).

683 patients from the GWAS were subsequently included in the mortality analysis (table 4). Enrolment criteria in the InterMune study meant that patients from this study included in our analysis had better pulmonary function (as assessed by FVC) and less heterogeneity of disease severity (as assessed by SD for lung function) than did those from the University of Chicago or the University of Pittsburgh (table 4). Additionally, InterMune had a shorter follow-up (table 4). Because follow-up varied widely, we decided to assess the effect of the newly identified susceptibility alleles in association with mortality both separately and jointly through a meta-analysis.

Univariate meta-analysis of the three case series showed that the carriers of the minor allele rs5743890_G that decreased susceptibility to IPF had increased mortality (table 5). Conversely, carriers of the major allele rs5743890_A that increased susceptibility to IPF had decreased mortality (figure 2). Various sensitivity analysis models sustained the findings for rs5743890, but had mixed results for rs17690703 (appendix pp 7, 14–15).

The best fitting multivariate Cox regression model with adjustment for covariates, such as demographic variables (age and sex), health status (smoking habits), and lung function (FVC and D_LCO), individually for data from InterMune, University of Chicago, and University of Pittsburgh, and collectively in meta-analysis also showed that the minor allele rs5743890_G was associated with increased mortality (table 5). Cochran's *Q* test for heterogeneity was done to ensure consistency of data. Although we detected only one instance of heterogeneity across all the comparisons (rs17690703; *Q*=8.9, p=0.011; table 5), we did meta-analyses with both fixed-effects and random-effects models to increase confidence in the findings. The association between rs5743890_G and mortality was sustained in both models (table 5). Three additional multivariate sensitivity analyses maintained these findings (appendix pp 7, 15–16).

In an assessment of treatment interactions, only InterMune had a treatment group (interferon gamma) with sufficient numbers and study duration for investigation of any level of drug–gene interaction. We did not record evidence for an interaction (appendix pp 7–8). Although we did not have a sufficient number for formal analysis, all cases in the ACE-IPF study with genotyping data and a mortality event carried the *MUC5B* polymorphism, which suggests an interaction (six given warfarin; one given placebo).

We next correlated the genotypes of the six SNPs that were associated with IPF susceptibility with gene-expression profiling data for IPF lung tissues from 67 individuals from the LGRC. *TOLLIP* expression was decreased by 20% in individuals carrying the minor allele rs5743890_G (p=0.097), 40% in those carrying the minor allele rs111521887_G ($p=3.0\times10^{-4}$), and 50% in those carrying the minor allele rs5743894_G ($p=2.93\times10^{-5}$) compared with homozygous carriers of common alleles for these SNPs. The other three associated SNPs identified in the GWAS located in *MUC5B*, *SPPL2C*, and *MDGA2* did not correlate with the expression of their respective genes or with *TOLLIP* (data not shown).

Discussion

We have identified novel genetic variants residing in two genetic loci—in *TOLLIP* at 11p15.5 and in *SPPL2C* at 17q21.31—that are associated with IPF susceptibility in European Americans, and have replicated the previously identified association with an SNP in *MUC5B*. As far as we are aware, we have shown for the first time that one of the novel genetic variants for susceptibility—rs5743890 in *TOLLIP*—is also associated with IPF mortality (panel).

MUC5B and *TOLLIP* reside at the same genetic locus. The difference in inheritability (dominant *vs* additive) and the linkage disequilibrium structure of the region suggest that the association of *TOLLIP* genetic variants with IPF susceptibility is independent from that of the previously reported *MUC5B* promoter SNP. Notably, the minor allele rs5743890_G in *TOLLIP* was a protective allele, because it was associated with reduced susceptibility to IPF. However, our mortality analysis showed that individuals who developed IPF despite having the protective rs5743890_G allele had increased mortality. *TOLLIP* SNPs rs111521887 and rs5743894 were in linkage disequilibrium with each other, but not with rs5743890. By contrast, individuals with the susceptibility *MUC5B* rs35705950_T allele who develop IPF have decreased mortality (unpublished). The independence of the association signals at 11p15.5 suggest that more than one variant at this locus might be modifying disease susceptibility and course. Additional functional studies are needed to address the causality of variants, and to fully establish either the individual contributions or the inter-relation between these two genes.

The region that contains *MUC5B* and *TOLLIP* exemplifies the complexity of expected association patterns in the genetic architecture of complex diseases. Patients with IPF carrying the minor alleles of associated *TOLLIP* SNPs had decreased *TOLLIP* expression in lung tissues, suggesting that this gene might modify disease course. As previously described, although rs35705950 genotypes have been associated with differential gene expression of *MUC5B* in healthy lung tissues, this correlation has been absent in patients with IPF.[5] Similarly, in our study, stratification by rs35705950 genotype in *MUC5B* did not correlate with expression of *MUC5B* or *TOLLIP* in patients with IPF. Although this evidence is compelling, it does not allow individual contributions or the inter-relation of these two genes in IPF susceptibility and mortality to be assessed.

TOLLIP is an important regulator of innate immune responses mediated by Toll-like receptor and the transforming growth factor β (TGF-β) signalling pathway. TOLLIP activates MYD88-dependent NF-κB to modulate Toll-like receptor signalling and membrane trafficking;[20] interacts with SMAD7 to modulate intracellular trafficking; negatively regulates the TGF-β signalling pathway by degrading ubiquitinated TGF-β type 1 receptor;[21] and interacts with caveolin-1 interacting protein in monocytes, regulating signalling in antigen-presenting cells to induce antigen-specific proliferation of T cells or B cells, or both.[22]*TOLLIP* polymorphisms are involved in regulation of TLR2 and TLR4 and

are associated with susceptibility to tuberculosis,[23] atopic dermatitis,[24] and sepsis.[25] Additionally, *TOLLIP* is differentially hypomethylated in lungs affected by IPF.[26] Toll-like receptor signalling affects TLR9, which can be used to identify rapidly progressive cases of IPF.[27] Lastly, failure to upregulate *TOLLIP* expression in inflammatory bowel disease can lead to chronic inflammation.[28] These studies, in context of the debate about the role of the immune system in IPF and its regulation of TGF-β signalling, make *TOLLIP* biologically plausible as both a gene of susceptibility and modifier of disease course.

Microsatellite instability and loss of heterozygosity in the 17q21 region has been reported in patients with IPF and sarcoidosis in previous cytological sputum studies,[29,30] and has been associated with the response to inhaled corticosteroids in asthma and chronic obstructive pulmonary disease.[31] We recorded that the minor allele rs17690703_T in the 17q21.31 region was associated with decreased susceptibility for IPF. This region—spanning 440 kb—partly or entirely involves five genes (*CRHR1*, *SPPL2C*, *MAPT*, *STH*, and *KANSL1*),[32] includes a known inversion in a large region of conserved linkage disequilibrium[33] that is positively selected in Europeans,[34] harbours many copy number variants, and has been associated with a microdeletion syndrome. The complexity of copy number variants and structural variants can affect genotyping; however, the location of rs17690703 in the 17q21.31 region is not near the commonly ascribed region known for copy number variants, which mostly includes *KANSL1*.[35] Although we detected many variants with nominal significance in the region in stage one of our GWAS, we focused on only the top SNPs. Additional sequencing and functional studies will be needed to establish causal variants, presence of insertions or deletions, effects of copy number variants in this region, and the effect of the inversion (if any).

We recorded a novel association between an SNP in *MDGA2*, which resides at 14q21.3, and IPF susceptibility, but did not replicate the finding in stage three of the GWAS, which could be because of our fairly small sample size. *MDGA2* is a paralogue for *ICAM*, which has been shown to be a potential biomarker of IPF disease activity.[36] This gene is intriguing and potentially important in IPF.

Our study has several limitations. First, IPF is a heterogeneous disease and by definition is a diagnosis of exclusion. As such, misdiagnoses are possible, which might lead to a reduction in power. However, all participants met accepted criteria for diagnosis[3] and many were checked with core pathology and radiology as in InterMune, ACE-IPF, and other studies. Analysis of the differences in minor allele frequencies of the susceptibility SNPs across the IPF case series supported the homogeneity of these cohorts, confirming consistency of the diagnostic criteria applied by each acquisition study. Second, additional SNPs could remain undiscovered secondary to poor information available from the 1000 Genomes Project, as the source of the enhanced number of SNPs by imputation. Third, as for most complex traits,[37] the effects reported for the common variants involved in IPF susceptibility are modest. Similarly, the hazard ratio for rs5743890 and IPF mortality was modest, and further studies are necessary to establish its interactive effect. Lastly, clinical characteristics are always limited by uncertainty about the date of onset of the disease, specifically relating to severity of illness with time. However, adjustment for disease severity by pulmonary function did not appreciably change the association between the *TOLLIP* SNP and mortality.

Although the genetic mechanisms underlying IPF are still not completely understood, our GWAS identified novel genetic variants associated with IPF susceptibility. Furthermore, one of the novel susceptibility alleles was associated with mortality and others were significantly associated with gene expression in lung tissue. This combination of biology and outcome emphasises the importance of *TOLLIP* in IPF; the gene warrants further investigation as a possible site of intervention. Overall, identification of common genetic variants associated

with IPF might provide insight into this complex disease and will hopefully lead to earlier detection, more predictable prognosis, and personalised therapeutic strategies.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This study was funded by the National Institutes of Health (P50 HL0894932, R01 HL095397, RC1 HL09961, and RC2 HL096845 to NK; RO1 HLl081619, RO1 HL087115, and R01 HL096845 to JDC; RC1 HL09961, HL101740, and HL080513 to IN; P01 HL98050, HL101740 and U01 HL105371 to JGNG); National Heart, Lung, and Blood Institute; Pulmonary Fibrosis Foundation; Coalition for Pulmonary Fibrosis; and Instituto de Salud Carlos III. S-FM received Core Subsidy Mini Awards of the Insitute of Translational Medicine and Clinical and Translational Science Award (UL1 RR024999). CF received a Spanish National Health Institute grant cofinanced by the European Regional Development Fund (PI11/00623). Gene-expression data were extracted from the LGRC. Spanish National Health Institute PI11/00623 grant, co-financed by the European Regional Development Funds.

References

1. Gribbin J, Hubbard RB, Le Jeune I, Smith CJ, West J, Tata LJ. Incidence and mortality of idiopathic pulmonary fibrosis and sarcoidosis in the UK. Thorax. 2006; 61:980–85. [PubMed: 16844727]
2. Raghu G, Weycker D, Edelsberg J, Bradford WZ, Oster G. Incidence and prevalence of idiopathic pulmonary fibrosis. Am J Respir Crit Care Med. 2006; 174:810–16. [PubMed: 16809633]
3. Raghu G, Collard HR, Egan JJ, et al. An official ATS/ERS/JRS/ALAT statement: idiopathic pulmonary fibrosis: evidence-based guidelines for diagnosis and management. Am J Respir Crit Care Med. 2011; 183:788–824. [PubMed: 21471066]
4. Raghu G, Anstrom KJ, King TE Jr, Lasky JA, Martinez FJ. Prednisone, azathioprine, and N-acetylcysteine for pulmonary fibrosis. N Engl J Med. 2012; 366:1968–77. [PubMed: 22607134]
5. Seibold MA, Wise AL, Speer MC, et al. A common MUC5B promoter polymorphism and pulmonary fibrosis. N Engl J Med. 2011; 364:1503–12. [PubMed: 21506741]
6. Zhang Y, Noth I, Garcia JG, Kaminski N. A variant in the promoter of MUC5B and idiopathic pulmonary fibrosis. N Engl J Med. 2011; 364:1576–77. [PubMed: 21506748]
7. Alder JK, Chen JJ, Lancaster L, et al. Short telomeres are a risk factor for idiopathic pulmonary fibrosis. Proc Natl Acad Sci USA. 2008; 105:13051–56. [PubMed: 18753630]
8. 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; 45:654–56. [PubMed: 18835860]
9. 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; 285:22103–13. [PubMed: 20466729]
10. Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature. 2007; 447:661–78. [PubMed: 17554300]
11. Raghu G, Brown KK, Bradford WZ, et al. A placebo-controlled trial of interferon gamma-1b in patients with idiopathic pulmonary fibrosis. N Engl J Med. 2004; 350:125–33. [PubMed: 14711911]
12. Lederer DJ, Kawut SM, Wickersham N, et al. Obesity and primary graft dysfunction after lung transplantation: the Lung Transplant Outcomes Group Obesity Study. Am J Respir Crit Care Med. 2011; 184:1055–61. [PubMed: 21799077]
13. Noth I, Anstrom KJ, Calvert SB, et al. A placebo-controlled randomized trial of warfarin in idiopathic pulmonary fibrosis. Am J Respir Crit Care Med. 2012; 186:88–95. [PubMed: 22561965]

14. American Thoracic Society, European Respiratory Society. Idiopathic pulmonary fibrosis: diagnosis and treatment—international consensus statement. Am J Respir Crit Care Med. 2000; 161:646–64. [PubMed: 10673212]
15. Pe'er I, Yelensky R, Altshuler D, Daly MJ. Estimation of the multiple testing burden for genomewide association studies of nearly all common variants. Genet Epidemiol. 2008; 32:381–85. [PubMed: 18348202]
16. Gabriel SB, Schaffner SF, Nguyen H, et al. The structure of haplotype blocks in the human genome. Science. 2002; 296:2225–29. [PubMed: 12029063]
17. R Development Core Team. R: a language and environment for statistical computing. R Foundation for Statistical Computing; Vienna: 2012.
18. Han B, Eskin E. Random-effects model aimed at discovering associations in meta-analysis of genome-wide association studies. Am J Hum Genet. 2011; 88:586–98. [PubMed: 21565292]
19. Nathan SD, Shlobin OA, Weir N, et al. Long-term course and prognosis of idiopathic pulmonary fibrosis in the new millennium. Chest. 2011; 140:221–29. [PubMed: 21729893]
20. Bulut Y, Faure E, Thomas L, Equils O, Arditi M. Cooperation of Toll-like receptor 2 and 6 for cellular activation by soluble tuberculosis factor and *Borrelia burgdorferi* outer surface protein A lipoprotein: role of Toll-interacting protein and IL-1 receptor signaling molecules in Toll-like receptor 2 signaling. J Immunol. 2001; 167:987–94. [PubMed: 11441107]
21. Zhu L, Wang L, Luo X, et al. Tollip, an intracellular trafficking protein, is a novel modulator of the transforming growth factor-beta signaling pathway. J Biol Chem. 2012; 287:39653–63. [PubMed: 23027871]
22. Kimman TG, Banus S, Reijmerink N, et al. Association of interacting genes in the toll-like receptor signaling pathway and the antibody response to pertussis vaccination. PLoS One. 2008; 3:e3665. [PubMed: 18987746]
23. Shah JA, Vary JC, Chau TT, et al. Human TOLLIP regulates TLR2 and TLR4 signaling and its polymorphisms are associated with susceptibility to tuberculosis. J Immunol. 2012; 189:1737–46. [PubMed: 22778396]
24. Schimming TT, Parwez Q, Petrasch-Parwez E, Nothnagel M, Epplen JT, Hoffjan S. Association of toll-interacting protein gene polymorphisms with atopic dermatitis. BMC Dermatol. 2007; 7:3. [PubMed: 17362526]
25. Song Z, Yin J, Yao C, et al. Variants in the Toll-interacting protein gene are associated with susceptibility to sepsis in the Chinese Han population. Crit Care. 2011; 15:R12. [PubMed: 21219635]
26. Rabinovich EI, Kapetanaki MG, Steinfeld I, et al. Global methylation patterns in idiopathic pulmonary fibrosis. PLoS One. 2012; 7:e33770. [PubMed: 22506007]
27. Trujillo G, Meneghin A, Flaherty KR, et al. TLR9 differentiates rapidly from slowly progressing forms of idiopathic pulmonary fibrosis. Sci Transl Med. 2010; 10:57ra82.
28. Steenholdt C, Andresen L, Pedersen G, Hansen A, Brynskov J. Expression and function of toll-like receptor 8 and Tollip in colonic epithelial cells from patients with inflammatory bowel disease. Scand J Gastroenterol. 2009; 44:195–204. [PubMed: 18985539]
29. Vassilakis DA, Sourvinos G, Markatos M, et al. Microsatellite DNA instability and loss of heterozygosity in pulmonary sarcoidosis. Am J Respir Crit Care Med. 1999; 160:1729–33. [PubMed: 10556148]
30. Vassilakis DA, Sourvinos G, Spandidos DA, Siafakas NM, Bouros D. Frequent genetic alterations at the microsatellite level in cytologic sputum samples of patients with idiopathic pulmonary fibrosis. Am J Respir Crit Care Med. 2000; 162:1115–19. [PubMed: 10988139]
31. Kim WJ, Sheen SS, Kim TH, et al. Association between CRHR1 polymorphism and improved lung function in response to inhaled corticosteroid in patients with COPD. Respirology. 2009; 14:260–63. [PubMed: 19210659]
32. Tan TY, Aftimos S, Worgan L, et al. Phenotypic expansion and further characterisation of the 17q21.31 microdeletion syndrome. J Med Genet. 2009; 46:480–89. [PubMed: 19447831]
33. Zody MC, Garber M, Adams DJ, et al. DNA sequence of human chromosome 17 and analysis of rearrangement in the human lineage. Nature. 2006; 440:1045–49. [PubMed: 16625196]

34. Boettger LM, Handsaker RE, Zody MC, McCarroll SA. Structural haplotypes and recent evolution of the human 17q21.31 region. Nat Genet. 2012; 44:881–85. [PubMed: 22751096]

35. [(accessed Feb 1, 2013)] Database of Genomic Variants. About the project. http://projects.tcag.ca/variation/project.html

36. Richards TJ, Kaminski N, Baribaud F, et al. Peripheral blood proteins predict mortality in idiopathic pulmonary fibrosis. Am J Respir Crit Care Med. 2012; 185:67–76. [PubMed: 22016448]

37. Hindorff LA, Sethupathy P, Junkins HA, et al. Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. Proc Natl Acad Sci USA. 2009; 106:9362–67. [PubMed: 19474294]

Panel: Research in context

Systematic review

We searched PubMed for reports of genome-wide association studies (GWAS) of idiopathic pulmonary fibrosis (IPF) with the search terms “IPF and GWAS”, “IPF and genetic variants”, “IPF and genetic association”, and “IPF and genetic association and mortality”. We used no date or language restrictions. We did not identify any GWAS of a similar scale to ours. Genetic variants of *MUC5B* had been identified in other cohorts,[5,6] but none had been associated with IPF mortality.

Interpretation

Our three-stage GWAS provides solid evidence that genetic variations underpin IPF. The association between a susceptibility allele in a novel genetic variant—*TOLLIP*—and mortality could provide insight into the complex disease process and will hopefully lead to earlier detection, more predictable prognosis, and personalised therapeutic strategies.

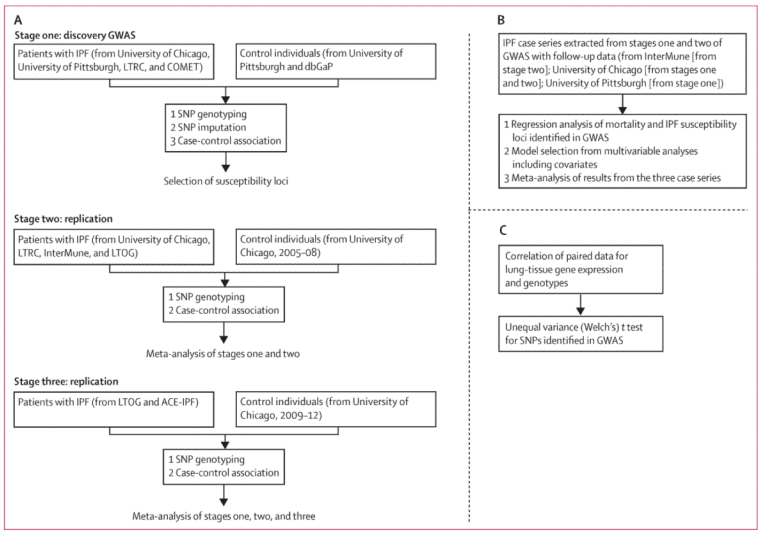

Figure 1. Study design

(A) Three-stage GWAS. Samples at each stage are independent. (B) Mortality analyses of data from three case series. (C) Assessment in correlation between genotypes and gene expression of lung tissue from LTRC. GWAS=genome-wide association study. IPF=idiopathic pulmonary fibrosis. LTRC=Lung Tissue Research Consortium. COMET=Correlating Outcomes with Biomedical Markers to Estimate Time-progression in IPF study. dbGAP=database of genotypes and phenotypes. SNP=single nucleotide polymorphism. LTOG=Lung Transplant Outcomes Group. ACE-IPF=AntiCoagulant Effectiveness in Idiopathic Pulmonary Fibrosis study.

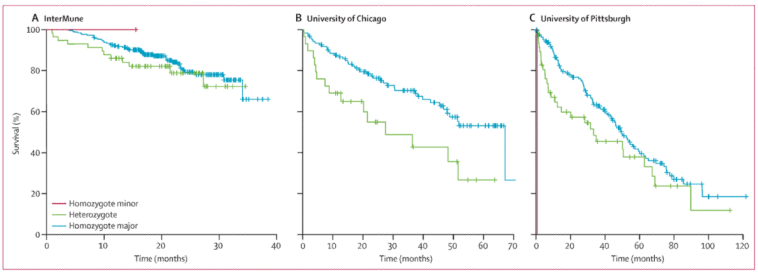

Figure 2. Survival analysis by rs5743890 genotype

Vertical lines indicate censoring. No homozygote minor patients from University of Chicago. Homozygote minor=carries two minor alleles (rs5743890_G). Heterozygote=carriers one minor allele (r5743890_G). Homozygous major=carries two major alleles (r5743890_A) and no minor alleles.

Table 1

Sample size in each stage of the genome-wide association study

	Sample size
Stage one: cases (n=633) *	
University of Chicago	187
University of Pittsburgh	254
COMET	89
LTRC	103
Stage one: control individuals (n=1545) †	
University of Pittsburgh	103
dbGAP	1442
Stage two: cases (n=544) †	
University of Chicago	14
LTOG	174
LTRC	42
InterMune	314
Stage two: control individuals (n=687) †	
University of Chicago	687
Stage three: cases (n=324) †	
LTOG	234
ACE-IPF	90
Stage three: control individuals (n=702) †	
University of Chicago	702

COMET=Correlating Outcomes of Biomedical Markers to Estimate Time-progression in IPF. LTRC=Lung Tissue Research Consortium. dbGAP=database of genotypes and phenotypes. LTOG=Lung Transplant Outcomes Group. ACE-IPF=AntiCoagulant Effectiveness in Idiopathic Pulmonary Fibrosis.

*617 European Americans.

†All European American.

Table 2

Characteristics of patients with idiopathic pulmonary fibrosis and control individuals in all stages of the genome-wide association study

	Patients				Control individuals*		
	Stage one (n=633)	Passed quality control in stage one (n=542)	Stage two (n=544)	Stage three (n=324)	Stage one (n=103)†	Stage two (n=687)†	Stage three (n=702)†
Age at diagnosis (years)	67 (61–73)	68 (61–73)	65 (59–70)	68 (63–72)	63 (40–70)	56 (46–65)	62 (55–68)
Sex							
Unknown	16	0	174	234	0	0	0
Known	617	542	370	90	103	687	702
Men	440 (71%)	385 (71%)	262 (71%)	68 (76%)	42 (41%)	393 (57%)	497 (71%)
Women	177 (29%)	157 (29%)	108 (29%)	22 (24%)	61 (59%)	294 (43%)	205 (29%)
Smoking status							
Unknown	146	132	174	234	0	..	..
Known	487	410	370	90	103	..	..
Never smoker	149 (31%)	116 (28%)	119 (32%)	20 (22%)	60 (58%)	..	..
Ever smoker	338 (69%)	294 (72%)	251 (68%)	70 (78%)	43 (42%)	..	..
Lung biopsies‡	227 (36%)	202 (37%)	96 (18%)	NA	..	..	..
Percentage of predicted FVC	64.9% (18.1)	64.9% (17.9)	70.3% (13.6)	58.4% (16.1)	..	..	..
Percentage of predicted D_LCO	46.5% (17.8)	46.5% (17.3)	47.1% (11.0)	35.0% (12.7)	..	..	..

Data are median (IQR), n (%), or mean (SD), unless otherwise stated. NA=not available. FVC=forced vital capacity. D_LCO=diffusion capacity of lung for carbon monoxide.

*Individuals with genome-wide data acquired from the database of genotypes and phenotypes were not included.

†Data for not obtained for all variables.

‡Lung Transplant Outcomes Group cohort in stages two and three, and cohort from AntiCoagulant Effectiveness in Idiopathic Pulmonary Fibrosis study in stage three were excluded because data were not available.

Table 3

Signals associated with IPF susceptibility in each stage of the genome-wide association study

	Minor allele	Major allele	Gene (genes to left and right)	Location	Stage one: discovery				Stage two: replication				Combined p value	Stage three: replication				Overall OR (95% CI)*	Overall p value
					MAF in patients with IPF (n=542)	MAF in control individuals (n=542)	OR (95% CI)*	P value	MAF in patients with IPF (n=544)	MAF in control individuals (n=687)	OR (95% CI)*	P value		MAF in patients with IPF (n=324)	MAF in control individuals (n=702)	OR (95% CI)*	P value		
rs35705950	T	G	*MUC5B* (*MUC5AC, TOLLIP*)	11p15.5	14%†	9%	1.57 (1.20–2.05)	6.31×10^{-6}	33%	12%	3.72 (3.02–4.58)	2.50×10^{-40}	2.38×10^{-38}	31%	14%	3.08 (2.37–3.99)	5.14×10^{-16}	2.43 (2.13–2.77)	2.40×10^{-50}
rs111521887	G	A	*TOLLIP* (*MUC5B, BRSK2*)	11p15.5	29%	21%	1.54 (1.26–1.87)	5.01×10^{-7}	25%	18%	1.58 (1.29–1.94)	7.90×10^{-6}	7.10×10^{-12}	23%	19%	1.24 (0.98–1.56)	8.03×10^{-2}	1.48 (1.32–1.66)	2.20×10^{-12}
rs5743894	G	A	*TOLLIP* (*MUC5B, BRSK2*)	11p15.5	29%	21%	1.58 (1.30–1.93)	1.26×10^{-7}	24%	17%	1.52 (1.25–1.86)	1.60×10^{-5}	2.40×10^{-12}	20%	18%	1.19 (0.94–1.50)	1.57×10^{-1}	1.49 (1.33–1.68)	1.35×10^{-12}
rs5743890	G	A	*TOLLIP* (*MUC5B, BRSK2*)	11p15.5	11%	15%	0.68 (0.53–0.88)	2.51×10^{-3}	9%	15%	0.57 (0.44–0.73)	4.00×10^{-6}	7.68×10^{-8}	10%	17%	0.60 (0.45–0.80)	3.49×10^{-4}	0.61 (0.52–0.71)	3.43×10^{-11}
rs7144383	G	A	*MDGA2* (NA)	14q21.3	13%	8%	1.82(1.37–2.41)	2.00×10^{-5}	11%	7%	1.57 (1.18–2.08)	1.00×10^{-3}	3.50×10^{-8}	13%	11%	1.16 (0.88–1.54)	2.99×10^{-1}	1.44 (1.23–1.69)	3.71×10^{-6}
rs17690703	T	C	*SPPL2C* (*CRHR1, MAPI*)	17q21.31	18%	26%	0.62 (0.51–0.76)	5.01×10^{-6}	20%	26%	0.74 (0.61–0.89)	1.60×10^{-3}	4.20×10^{-8}	20%	24%	0.78 (0.62–0.98)	3.07×10^{-2}	0.70 (0.62–0.79)	5.70×10^{-9}

IPF=idiopathic pulmonary fibrosis. OR=odds ratio. MAF=minor allele frequency. NA=not applicable (no genes present within a predefined distance in either direction.

*Calculated on basis of MAF.

†MAF from imputation; physical genotyped MAF=0.34. The MAF discrepancy could be due to limited available information in the 1000 Genomes Project reference.

Table 4

Characteristics of patients in mortality analysis

	InterMune (n=314)	University of Chicago (n=157)	University of Pittsburgh (n=212)	Overall (n=683)
Follow-up (months)	22 (7)	30 (20)	36 (29)	28 (21)
Participants who were alive at last known follow-up	257 (82%)	94 (60%)	91 (43%)	444 (65%)
Participants who had died at last known follow-up	57 (18%)	63 (40%)	121 (57%)	239 (35%)
Age at diagnosis (years)	67 (61–72)	69 (63–75)	68 (62–74)	68 (62–73)
Men	228 (73%)	118 (75%)	143 (67%)	489 (72%)
Women	86 (27%)	39 (25%)	69 (33%)	194 (28%)
Never smoker*	97 (31%)	41 (29%)	23 (24%)	161 (29%)
Ever smoker*	217 (69%)	102 (71%)	74 (76%)	393 (71%)
Had undergone transplantation before last known follow-up	6 (2%)	45 (29%)	10 (5%)	61 (9%)
Percentage of predicted FVC	71.56% (12.68)	65.17% (18.29)†	65.27% (19.72)‡	68.32% (16.39)
Percentage of predicted D_LCO	47.24% (8.86)	48.17% (17.92)	47.55% (18.77)	47.61% (14.33)

Data are mean (SD), n (%), or median (IQR). Patients who had undergone lung transplantation were censored. FVC=forced vital capacity. DLCO=diffusion capacity of lung for carbon monoxide.

*Percentages calculated with number of known phenotypes (314 InterMune; 143 University of Chicago; 97 University of Pittsburgh; 554 overall.

†Significantly different from that in InterMune ($p=2.06 \times 10^{-5}$).

‡Significantly different from that in InterMune ($p=4.48 \times 10^{-5}$).

Table 5

Cox regression models for mortality

	InterMune (n=314)		University of Chicago (n=157)		University of Pittsburgh (n=212)		Meta-analysis with random-effects model (n=683)		Meta-analysis with fixed-effects model (n=683)	
	HR (95% CI)	p value	HR (95% CI)	p value	HR (95% CI)	p value	HR (95% CI)	p value	HR (95% CI)	p value
Univariate Cox regression model										
rs111521887_G	0.94 (0.61–1.46)	0.800	0.90 (0.59–1.39)	0.650	0.88 (0.63–1.23)	0.460	0.90(0.72–1.13)	0.433	0.90(0.72–1.13)	0.382
rs5743894_G	0.96 (0.60–1.51)	0.840	0.95 (0.60–1.51)	0.810	0–85 (0.61–1.18)	0.320	0.90 (0.71–1.13)	0.407	0.90 (0.71–1.13)	0.358
rs5743890_G	1.28 (0.68–2.39)	0.440	2.18 (1.22–3.89)	0.0083	1.60 (1.06–2.41)	0.025	1.65 (1.23–2.21)	0.0012	1.65 (1.23–2.21)	0.0009
rs7144383_G	0.78 (0.38–1.60)	0.500	1.06 (0.59–1.89)	0.840	1.10 (0.75–1.63)	0.630	1.03 (0.77–1.38)	0.891	1.03 (0.77–1.38)	0.847
rs17690703_T	1.51 (1.01–2.24)	0.044	1.57 (1.05–2.34)	0.030	0.76 (0.54–1.07)	0.120	1.20 (0.75–1.94)*	0.034	1.15 (0.93–1.43)	0.198
Multivariable Cox regression model										
rs111521887_G	0.92 (0.59–1.43)	0.720	0.96 (0.59–1.56)	0.860	1.06 (0.76–1.49)	0.730	1.00 (0.79–1.26)	0.985	1.00 (0.79–1.26)	0.972
rs5743894_G	0.98 (0.61–1.56)	0.930	1.00 (0.14–7.09)	1.000	1.03 (0.74–1.45)	0.840	1.01 (0.77–1.33)	0.948	1.01 (0.77–1.33)	0.916
rs5743890_G	1.78 (0.93–3.38)	0.080	2.09 (1.05–4.18)	0.036	1.55 (0.99–2.45)	0.057	1.72 (1.24–2.38)	0.0015	1.72 (1.24–2.38)	0.0012
rs7144383_G	0.76 (0.37–1.56)	0.460	1.87 (0.99–3.53)	0.054	1.00 (0.65–1.53)	0.990	1.12 (0.71–1.79)	0.584	1.11 (0.81–1.52)	0.531
rs17690703_T	1.45 (0.97–2.16)	0.069	1.42 (0.81–2.49)	0.220	0.83 (0.57–1.21)	0.340	1.17 (0.80–1.72)	0.250	1.14 (0.89–1.45)	0.307

Lung transplant individuals were censored. Time at risk was defined as the interval between data of enrolment in each cohort and date of the last follow-up, lung transplant, or death. HR=hazard ratio.

*Effects associated with significant heterogeneity across studies (Q=8.9; p=0.011).