What's in a genotype? Unpicking genetic links in complex disease

S63

GENOME-WIDE SEX-BY-SNP INTERACTION ANALYSIS OF SUSCEPTIBILITY TO IDIOPATHIC PULMONARY FIBROSIS

¹OC Leavy, ¹RJ Allen, ²JM Oldham, ¹B Guillen-Guio, ³A Stockwell, ⁴R Braybrooke, ⁴RB Hubbard, ⁵S Ma, ⁶TE Fingerlin, ⁷N Kaminski, ⁸Y Zhang, ⁶DA Schwartz, ³B Yaspán, ⁹TM Maher, ¹⁰PL Molyneaux, ¹¹C Flores, ¹I Noth, ¹⁰RG Jenkins, ¹LV Wain. ¹University of Leicester, Leicester, UK; ²University of California at Davis, Sacramento, USA; ³Genentech Inc, San Francisco, USA; ⁴University of Nottingham, Nottingham, UK; ⁵University of Virginia, Charlottesville, USA; ⁶National Jewish Health, Denver, USA; ⁷Yale School of Medicine, New Haven, USA; ⁸The University of Pittsburgh, Pittsburgh, USA; ⁹University of Southern California, Los Angeles, USA; ¹⁰Royal Brompton Hospital, London, UK; ¹¹Hospital Universitario Nuestra Señora de Candelaria, Santa Cruz de Tenerife, Spain

10.1136/thorax-2021-BTSabstracts.69

Idiopathic pulmonary fibrosis (IPF) is a complex and heterogeneous fibrotic lung disease with median survival of 3 years. The disease has a genetic component, with genome-wide association studies (GWAS) identifying 15 genetic loci that influence IPF risk. The rs35705950 SNP (single nucleotide polymorphism), in the promoter region of the *MUC5B* gene, is the largest common genetic risk factor of IPF (OR=4.84 [95%CI: 4.37, 5.36], minor allele (T) frequency in cases=29.1%). IPF is more prevalent in males than females but the reasons behind this are not understood. We hypothesised that the genetic determinants of IPF might help to identify biological mechanisms that drive differences in IPF susceptibility and presentation between males and females. To address this, we performed a genome-wide sex-by-SNP interaction analysis of IPF susceptibility.

A genome-wide sex-by-SNP interaction meta-analysis was performed that included four independent IPF case-control data sets comprising a discovery of 3,432 cases (2,421 males & 1,011 females) and 18,559 controls (12,096 males & 6,463 females) of European descent. We sought replication in an independent dataset comprising 664 cases (488 males &

176 females) and 1,874 controls (507 males & 1,367 females).

We defined suggestive statistical significance as independent variants with $P < 5 \times 10^{-5}$, and consistent direction of effect with $P < 0.05$ in at least 3 of the 4 studies. Eight independent variants meeting these criteria were identified in the meta-analysis. Five variants had $P < 0.05$ for IPF risk in both males and females when analysed separately; all five were observed to have an opposite direction of effect on males compared to females. None of the eight significant variants were associated ($P < 0.05$) with IPF susceptibility when males and females were combined, and none reached our replication significance threshold in the independent dataset.

Our findings suggest that there may be differences in genetic determinants of IPF risk between males and females, but larger studies are needed to confirm the reported variants. Understanding the mechanisms behind these could provide further insight into different biological pathways that underlie pulmonary fibrosis.

S64

APPLICATION OF OPEN VIRTUAL PHEWAS TOOLS FOR EXPLORATORY ANALYSES OF RISK ALLELES IN IDIOPATHIC PULMONARY FIBROSIS

S Moss, I Stewart, RG Jenkins. Margaret Turner Warwick Centre for Fibrosing Lung Disease, National Lung and Heart Institute, Imperial College London, London, UK

10.1136/thorax-2021-BTSabstracts.70

Introduction and Objectives Publicly available databases are becoming increasingly valuable resources for exploratory analyses of multi-omics data. This virtual PheWAS method demonstrates the application of open databases of GWAS summary data in identifying genetic disease-associated phenotypes, which can support further investigation into disease pathways and novel treatment targets.

We aimed to assess the application of open virtual PheWAS tools to identify significant phenotypic associations with established genomic risk loci in idiopathic pulmonary fibrosis (IPF).

Abstract S64 Table 1 Grouped phenotypic associations with IPF risk loci

Gene/locus - SNP	GWAS p-value	Phenotypic associations ($p < 5 \times 10^{-8}$)				
		Blood cell traits	Lung function	Cancer	IPF	General health and physical measurements
<i>KIF15</i> - rs78238620	4.05x10 ⁻¹⁴	No	No	No	No	No
<i>MAD1L1</i> - rs12699415	9.38x10 ⁻²⁰	No	No	No	No	Yes**
<i>DEPTOR</i> - rs28513081	1.93x10 ⁻¹¹	No	No	No	No	Yes**
<i>RTEL1</i> - rs41308092	2.24x10 ⁻¹⁰	No	No	No	No	No
<i>HECTD2</i> - rs537322302	3.43x10 ⁻⁸	No	No	No	No	No
<i>TERC</i> - rs12696304	7.09x10 ⁻¹³	Yes**	No	No	Yes	Yes*
<i>FAM13A</i> - rs2013701	3.30x10 ⁻¹³	Yes*	Yes**	No	No	No
<i>TERT</i> - rs7725218	1.54x10 ⁻²⁰	Yes**	No	Yes**	No	Yes*
<i>DSP</i> - rs2076295	2.79x10 ⁻³⁰	No	Yes*	No	Yes	No
<i>MUC5B</i> - rs35705950	1.18x10 ⁻²⁰³	No	No	No	Yes	No
<i>ATP11A</i> - rs9577395	1.34x10 ⁻¹⁰	Yes**	No	No	No	No
<i>IVD</i> - rs59424629	7.30x10 ⁻¹⁶	Yes*	Yes*	No	No	Yes*
<i>AKAP13</i> - rs62023891	1.27x10 ⁻¹⁰	Yes*	No	No	No	Yes*
<i>MAPT</i> - rs2077551	2.83x10 ⁻¹⁶	Yes**	Yes**	Yes	No	Yes**
<i>DPP9</i> - rs12610495	2.92x10 ⁻¹²	No	No	No	Yes	No
7q22.1 - rs2897075	3.10x10 ⁻¹⁴	Yes*	Yes*	No	No	Yes*

Key	
Phenoscaner
IEU OpenGWAS
Replicated in both databases

Table 1: Grouped phenotypic associations with IPF risk loci.

Methods A total of 16 GWAS-established single nucleotide polymorphisms (SNPs) associated with IPF that meet genome-wide significance ($p < 5 \times 10^{-8}$) in published studies were identified for evaluation. The Phenoscanner and IEU Open GWAS databases were queried via the *phenoscanner* and *ieugwasr* packages, respectively, in R (R Version 4.0.5). To maximise data retrieval, no p-value threshold filter was applied in initial SNP-trait association data collection. Subsequent statistical analysis and data plotting was performed in R to characterise significant phenotypic associations ($p < 5 \times 10^{-8}$).

Results Statistically significant ($p < 5 \times 10^{-8}$) phenotype associations were identified for 13/16 IPF risk loci (table 1). Associated phenotypes included lung function measures (FEV1/FVC, FVC), overall health, blood cell traits (eosinophil percentage of granulocytes, red blood cell (erythrocyte) count, neutrophil count) and multiple types of cancer (ovarian, prostate, breast, and lung). SNP-trait associations with blood cell traits and cancers observed for three SNPs (*TERT*, *TERC* and *MAPT*) were replicated in both databases. Some well-established genomic risk loci in IPF did not replicate phenotypic associations in both databases (rs35705950, *MUC5B*; rs12610495, *DPP9*), highlighting the advantage of complementary dataset analyses to maximise identification of pleiotropic associations.

Conclusions Virtual PheWAS identified phenotypes associated with IPF risk loci that include blood cell traits and multiple cancers. Combining PheWAS results across multiple datasets is necessary to detect established phenotypes, demonstrating the importance of complementary sources in detecting novel pleiotropy. Open virtual analysis tools offer an efficient exploratory approach which can identify SNP-phenotype associations that warrant focused investigation through large scale PheWAS with disease-agnostic data.

REFERENCE

- Allen RJ, et al. Genome-wide association study of susceptibility to idiopathic pulmonary fibrosis. *Am. J. Respir. Crit. Care Med.* 2020;201:564–574.

S65

GENOME-WIDE ASSOCIATION STUDY OF SURVIVAL TIMES AFTER DIAGNOSIS OF IDIOPATHIC PULMONARY FIBROSIS

¹RJ Allen, ²JM Oldham, ³JM Lorenzo-Salazar, ⁴PL Molyneaux, ⁵SF Ma, ⁶A Stockwell, ⁷C Joseph, ⁸JS Kim, ⁹B Guillen-Guio, ¹⁰T Hernandez-Beefink, ¹¹J Kropski, ¹²Y Huang, ¹³CT Lee, ¹⁴A Adegunsoye, ¹⁵JV Pugashetti, ¹⁶A Linderholm, ¹⁷V Vo, ¹⁸M Strek, ¹⁹R Hubbard, ²⁰N Hirani, ²¹MKB Whyte, ²²S Hart, ²³A Nicholson, ²⁴H Parfrey, ²⁵D Rassl, ²⁶W Wallace, ²⁷WA Fahy, ²⁸E Valenzi, ²⁹Y Zhang, ³⁰N Kaminski, ³¹P Wolters, ³²M Molina-Molina, ³³FJ Martinez, ³⁴I Hall, ³⁵MD Tobin, ³⁶TM Maher, ³⁷T Blackwell, ³⁸B Yaspan, ³⁹RG Jenkins, ⁴⁰C Flores, ⁴¹LV Wain, ⁴²I Noth. ¹University of Leicester, Leicester, UK; ²University of California, Davis, Sacramento, USA; ³ITER, Santa Cruz de Tenerife, Spain; ⁴Imperial College London, London, UK; ⁵University of Virginia, Charlottesville, USA; ⁶Genentech, San Francisco, USA; ⁷University of Nottingham, Nottingham, UK; ⁸Hospital Universitario Nuestra Señora de Candelaria, Santa Cruz de Tenerife, Spain; ⁹Vanderbilt University, Nashville, USA; ¹⁰University of Chicago, Chicago, USA; ¹¹University of Edinburgh, Edinburgh, UK; ¹²Hull York Medical School, Hull, UK; ¹³Royal Brompton Hospital, London, UK; ¹⁴Royal Papworth Hospital, Cambridge, UK; ¹⁵GlaxoSmithKline, Stevenage, UK; ¹⁶University of Pittsburgh, Pittsburgh, USA; ¹⁷Yale School of Medicine, New Haven, USA; ¹⁸University of California, San Francisco, San Francisco, USA; ¹⁹Universitat de Barcelona, Barcelona, Spain; ²⁰Weill Cornell Medical Centre, New York, USA; ²¹University of Southern California, Los Angeles, USA

10.1136/thorax-2021-BTSAbstracts.71

Introduction Idiopathic pulmonary fibrosis (IPF) is a devastating lung disease where the lungs become progressively scarred. The median survival time after diagnosis is three years, albeit

disease progression varies greatly between individuals. Recently, large genome-wide association studies (GWAS) have shown there are a number of DNA regions associated with disease risk, however these variants are generally not associated with disease progression.

Aim To identify genetic variants associated with disease progression.

Methods We performed a two-stage GWAS of transplant-free survival times after diagnosis of IPF. In stage 1, genome-wide analyses were performed in three separate studies using a Cox proportional hazards models adjusting for age, sex, study centre and genetic principal components and the results were meta-analysed across the studies. In stage 2, variants with $p < 5 \times 10^{-5}$ in the stage 1 meta-analysis and $p < 0.05$ in each separate study with consistent direction of effects were tested for their association with IPF survival in independent samples. Variants were deemed associated with IPF survival if they were genome-wide significant in a meta-analysis of stages 1 and 2 ($p < 5 \times 10^{-8}$). Immunohistochemistry and transcriptomics were performed to follow-up genes implicated by the genome-wide analysis.

Results A total of 1,481 IPF cases and 9 million genetic variants were included in the stage 1 analysis, and 397 individuals in stage 2. One variant, rs35647788 in an intron of *PCSK6*, was genome-wide significant in the stage 1 and 2 meta-analysis, and showed nominal significance ($p < 0.05$) in each study with consistent direction of effects across all studies. Gene-prioritisation analyses identified *PCSK6*, of which rs35647788 lies in an intron of, as the most likely gene of interest. Immunohistochemistry showed that *PCSK6* was highly expressed in lung tissue samples from IPF cases compared to controls, specifically in the lung epithelium. In IPF cases, higher levels of circulating *PCSK6* were associated with poorer survival times and increased *PCSK6* gene expression was associated with a greater annual decline in lung capacity. There was little overlap between the genetic determinants of disease risk and survival times.

Conclusions This is the first GWAS of IPF survival which has identified novel important biological processes involved in the progression of IPF.

S66

IRON DEFICIENCY: COMPLICATIONS, COMPENSATIONS AND TREATMENTS BY HEREDITARY HAEMORRHAGIC TELANGIECTASIA MOLECULAR GENOTYPE

L Sharma, A Rizvi, N Badiani, A Soni, CL Shovlin. Imperial College London, London, UK

10.1136/thorax-2021-BTSAbstracts.72

Introduction and objectives Patients with pulmonary arteriovenous malformations (PAVMs) require normal iron levels in order to compensate appropriately for hypoxaemia. Low serum iron has also been shown to be associated with ischaemic stroke and venous thromboembolism risk in the population. Those with hereditary haemorrhagic telangiectasia (HHT) are at risk of developing iron deficiency if the iron lost during nosebleeds and/or gastrointestinal haemorrhage is not adequately replaced. The aims of this study were to examine associations between the HHT molecular genotype and iron deficiency indices.

Methods A database containing repeated measurements from 426 genotyped HHT and PAVM patients was retrospectively analysed to compare iron deficiency rates, complications,