Overlap of genetic risk between interstitial lung abnormalities and idiopathic pulmonary fibrosis

*Brian D. Hobbs[1,2], *Rachel K. Putman[2], Tetsuro Araki[3,4], Mizuki Nishino[3,4], Gunnar Gudmundsson[5], Vilmundur Gudnason[6,7], Gudny Eiriksdottir[7], Nuno Rodrigues Zilhao Nogueira[7], Josée Dupuis[8,9], Hanfei Xu[8], George T. O'Connor[9,10], Ani Manichaikul[11,12] , Jennifer Nguyen[11], Anna J. Podolanczuk[13], Purnema Madahar[13], Jerome I. Rotter[14,15], David J. Lederer[13,16], R. Graham Barr[13,16], Stephen S. Rich[11,12], Elizabeth J. Ampleford[17], Victor E. Ortega[17], Stephen P. Peters[17], Wanda K. O'Neal[18], John D. Newell Jr[19,20], Eugene R. Bleecker[21], Deborah A. Meyers[21], Richard J. Allen[22], Justin M. Oldham[23], Shwu-Fan Ma[24], Imre Noth[24], R. Gisli Jenkins[25], Toby M. Maher[26,27], Richard B. Hubbard[25,28], Louise V. Wain[29], Tasha E. Fingerlin[30,31], David A. Schwartz[31,32,33] , George R. Washko[2,4], Ivan O. Rosas[2], Edwin K. Silverman[1,2], Hiroto Hatabu[3,4], *Michael H. Cho[1,2], *Gary M. Hunninghake[2,3], COPDGene Investigators, ECLIPSE Investigators, SPIROMICS Research Group, and UK ILD Consortium

[1]Channing Division of Network Medicine, Brigham and Women's Hospital, Boston, MA, USA; [2]Division of Pulmonary and Critical Care Medicine, Brigham and Women's Hospital, Boston, MA, USA; [3]Department of Radiology, Brigham and Women's Hospital, Boston, MA, USA; [4]Center for Pulmonary Functional Imaging, Brigham and Women's Hospital, Boston, MA, USA; [5]Dept of Respiratory Medicine, Landspital University Hospital, University of Iceland, Reykjavik, Iceland; [6]Faculty of Medicine, University of Iceland, Reykjavik, Iceland; [7]Icelandic Heart Association, Kopavogur, Iceland; [8]Department of Biostatistics, Boston University School of Public Health, Boston, MA, USA; [9]National Heart, Lung, and Blood Institute's Framingham Heart Study, Framingham, MA, USA; [10]Pulmonary Center, Department of Medicine, Boston University, Boston, MA, USA; [11]Center for Public Health Genomics, University of Virginia, Charlottesville, VA, USA; [12]Department of Public Health Sciences, University of Virginia, Charlottesville, VA, USA; [13]Department of Medicine, College of Physicians and Surgeons, Columbia University, New York, NY, USA; [14]Institute for Translational Genomics and Population Sciences and Department of Pediatrics, Los Angeles Biomedical Research Institute at Harbor-UCLA Medical Center, Torrance, CA, USA; [15]Division of Genomic Outcomes, Departments of Pediatrics and Medicine, Harbor-UCLA Medical Center, Torrance, CA, USA; [16]Department of Epidemiology, Mailman School of Public Health, Columbia University, New York, NY, USA; [17] Department of Internal Medicine, Center for Precision Medicine, Wake Forest School of Medicine, Winston-Salem, NC, USA; [18]Marsico Lung Institute, Department of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA; [19]Division of Cardiovascular and Pulmonary Imaging, Department of Radiology, University of Iowa Carver College of Medicine, Iowa City, IA, USA; [20]Department of Radiology, University of Washington, Seattle, WA, USA; [21]Division of Genetics, Genomics and Precision Medicine, Department of Medicine, University of Arizona, Tucson, AZ, USA; [22]Department of Health Sciences, University of Leicester, Leicester, UK; [23]Department of Internal Medicine, University of California Davis, Davis, CA, USA; [24]Divsion of Pulmonary and Critical Care Medicine, University of Virginia, Charlottesville, VA, USA; [25]NIHR Biomedical Research Centre, Respiratory Research Unit, School of Medicine, University of Nottingham, Nottingham, UK;[26]NIHR Respiratory Biomedical Research Unit, Royal Brompton Hospital, London, UK; [27]Fibrosis Research Group, Inflammation, Repair and Development Section, National Heart and Lung Institute, Imperial College, London, UK; [28]Division of Epidemiology and Public Health, University of Nottingham, Nottingham, UK; [29]National Institute for Health Research, Leicester Respiratory Biomedical Research Centre, Glenfield Hospital, Leicester, UK; [30]Center for Genes, Environment and Health, National Jewish Health, Denver, CO, USA; [31]Department of Biostatistics and Informatics, University of Colorado Denver, Aurora, CO, USA; [32]Department of Medicine, School of Medicine, University of Colorado Denver, Aurora, CO,

USA; [33]Department of Immunology, School of Medicine, University of Colorado Denver, Aurora, CO, USA

* These authors contributed equally to this work

Support: Dr. Hobbs is supported by NIH grant K08 HL136928. Dr. Putman is supported by NIH grant K08 HL140087. Dr. Nishino is supported by NIH grant R01 CA203636. Dr. Gudmundsson is supported by Oddur Olafsson Fund, project grant 141513-051 from the Icelandic Research Fund and Landspitali Scientific Fund A-2015-030, A-2016-023, A-2017-030, A-2018-022 and A-2018-025. Dr. Gudnason is supported by NIA grant: 27120120022C and project grant 141513-051 from the Icelandic Research Fund. Dr. O'Connor is supported by NIH grant OT2 OD026553. Dr. Manichaikul is supported by NIH grant R01 HL131565. Dr. Podolanczuk is supported by NIH grant K23 HL140199. Dr. Rotter is supported by NIH grants R01 HL142302, R01 EY009052, R01EY023704, and P30 DK063491. Dr. Lederer is supported by NIH grants K24 HL131937, R01 HL103676 and R01 HL137234. Dr. Barr is supported by NIH grants R01 HL077612, R01 HL093081, R01 HL121270, and R01 HL142028. Dr. Rich is supported by NIH grants U01 HL120393, DP3 DK111906, and P01 HL136275. Dr. O'Neal is supported by NIH grants R01 HL117843 and U24 HL141762. Dr. Ortega is supported by NIH grants K08 HL118128 and R01 HL142992. Dr. Bleecker is supported by NIH grants UG1 HL1390534 and U01 HL109164. Dr. Meyers is supported by NIH grants R01 NR013700 and U01 HL109164. Dr. Allen is supported by an Action Pulmonary Fibrosis Mike Bray Fellowship. Dr. Oldham is supported by NIH grant K23 HL138190. Dr. Noth is supported by NIH grant R01 HL130796. Dr. Jenkins is supported by MRC Grant G0901226. Dr. Maher is supported by an NIHR Clinician Scientist Fellowship (NIHR reference CS-2013-13-017). Dr. Wain holds a GlaxoSmithKline/British Lung Foundation Chair in Respiratory Research. The research was partially supported by the NIHR Leicester Biomedical Research Centre; the views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health. Dr. Fingerlin is supported by NIH grants R01 HL114587, R01 HL097163, and P01 HL132821. Dr. Schwartz is supported by NIH grants P01 HL092870, R01 HL097163, and R33 HL120770, R33 CA182360, and UH3 HL123442. Washko is supported by NIH grants R01 HL116473 and R01 HL122464. Dr. Rosas is supported by NIH grants U01 HL133232 and R01 HL130974. Dr. Silverman is supported by NIH grants U01 HL089856, R01 HL113264, R01 HL137927, R01 HL133135, and P01 HL114501. Dr. Cho is supported by NIH grants R01 HL135142, R01 HL113264, and R01 HL137927. Dr. Hunninghake is supported by NIH grants R01 HL111024, R01 HL130974, R01 135142, and project grant 141513-051 from the Icelandic Research Fund. The Framingham Heart Study is supported by NIH contracts N01-HC-25195 and HHSN268201500001I. COPDGene is supported by NIH Grant Numbers U01 HL089897 and U01 HL089856. The COPDGene project is also supported by the COPD Foundation through contributions made to an Industry Advisory Board comprised of AstraZeneca, Boehringer Ingelheim, GlaxoSmithKline, Novartis, Pfizer, Siemens and Sunovion. The Age, Gene/Environment Susceptibility-Reykjavik Study was supported by NIA grant: 27120120022C, NIH contracts N01-AG-1-2100 and HHSN27120120022C, the NIA Intramural Research Program, Hjartavernd (the Icelandic Heart Association), and the Althingi (the Icelandic Parliament). MESA and the MESA SHARe project are conducted and supported by the National Heart, Lung, and Blood Institute (NHLBI) in collaboration with MESA investigators. Support for MESA is provided by contracts HHSN268201500003I, N01-HC-95159, N01-HC-95160, N01-HC-95161, N01-HC-95162, N01-HC-95163, N01-HC-95164, N01-HC-95165, N01-HC-95166, N01-HC-95167, N01-HC-95168, N01-HC-95169, UL1-TR-000040, UL1-TR-001079, and UL1-TR-001420, and by the National Center for Advancing Translational Sciences, CTSI grant ULTR001881, and the National Institute of Diabetes and Digestive and Kidney Disease

Diabetes Research Center (DRC) grant DK063491 to the Southern California Diabetes Endocrinology Research Center. SPIROMICS (Subpopulations and Intermediate Outcome Measures in COPD Study) is funded by contracts from the NHLBI (HHSN268200900013C, HSN268200900014C, HHSN268200900015C, HHSN268200900016C, HHSN268200900017C, HHSN268200900018C, HHSN268200900019C, and HHSN268200900020C) and a grant from the NIH/NHLBI (U01 HL137880), and supplemented by contributions made through the Foundation for the NIH and the COPD Foundation from AstraZeneca/MedImmune; Bayer; Bellerophon Therapeutics; Boehringer-Ingelheim Pharmaceuticals Inc.; Chiesi Farmaceutici; Forest Research Institute Inc.; GlaxoSmithKline; Grifols Therapeutics Inc.; Ikaria Inc.; Novartis Pharmaceuticals Corporation; Nycomed GmbH; ProterixBio; Regeneron Pharmaceuticals Inc.; Sanofi; Sunovion; Takeda Pharmaceutical Company; and Theravance Biopharma and Mylan.

The authors thank the SPIROMICS participants and participating physicians, investigators and staff for making this research possible. More information about the study and how to access SPIROMICS data is at www.spiromics.org. We would like to acknowledge the following current and former investigators of the SPIROMICS sites and reading centers: Neil E Alexis, MD; Wayne H Anderson, PhD; Mehrdad Arjomandi, MD; Igor Barjaktarevic, MD, PhD; R Graham Barr, MD, DrPH; Lori A Bateman, MSc; Surya P Bhatt, MD; Richard C Boucher, MD; Russell P Bowler, MD, PhD; Stephanie A Christenson, MD; Alejandro P Comellas, MD; Christopher B Cooper, MD, PhD; David J Couper, PhD; Gerard J Criner, MD; Ronald G Crystal, MD; Jeffrey Curtis, MD; Claire M Doerschuk, MD; Mark T Dransfield, MD; Brad Drummond, MD; Christine M Freeman, PhD; Craig Galban, PhD; MeiLan K Han, MD, MS; Nadia N Hansel, MD, MPH; Annette T Hastie, PhD; Eric A Hoffman, PhD; Yvonne Huang, MD; Robert J Kaner, MD; Richard E Kanner, MD; Eric C Kleerup, MD; Jerry A Krishnan, MD, PhD; Lisa M LaVange, PhD; Stephen C Lazarus, MD; Fernando J Martinez, MD, MS; Wendy C Moore, MD; John D Newell Jr, MD; Robert Paine, III, MD; Laura Paulin, MD, MHS; Cheryl Pirozzi, MD; Nirupama Putcha, MD, MHS; Elizabeth C Oelsner, MD, MPH; Sanjeev Raman, MBBS, MD; Stephen I. Rennard, MD; Donald P Tashkin, MD;; J Michael Wells, MD; Robert A Wise, MD; and Prescott G Woodruff, MD, MPH. The project officers from the Lung Division of the National Heart, Lung, and Blood Institute were Lisa Postow, PhD, and Lisa Viviano, BSN.

This paper is subject to the NIH public access policy:

http://www.nih.gov/about/publicaccess/Finalpublicaccessimplementation031505.htm.

Corresponding Author:
Gary M. Hunninghake, M.D., M.P.H.
Pulmonary and Critical Care Division
Department of Medicine
Brigham and Women's Hospital
75 Francis Street
Boston, MA 02115
Email: ghunninghake@bwh.harvard.edu
Phone: 617-525-9687

Author Contributions:

Study Design: V. Gudnason, G.T. O'Connor, D.J. Lederer, R.G. Barr, E.R. Bleecker, D.A. Meyers, I. Noth, R.G. Jenkins, L.V. Wain, D.A. Schwartz, E.K. Silverman, M.H. Cho, G.M. Hunninghake

Acquisition, analysis, or interpretation of the data: B.D. Hobbs, R.K. Putman, T Araki, M. Nishino, G. Gudmundsson, V. Gudnason, G Eiriksdottir, N.R.Z. Nogueira, J. Dupuis, H. Xu, G.T. O'Connor, A. Manichaikul, J. Nguyen, A.J. Podolanczuk, J.I. Rotter, D.J. Lederer, R.G. Barr, S.S. Rich, E.J. Ampleford, V.E. Ortega, S.P. Peters, W.K. O'Neal, J.D. Newell Jr., E.R. Bleecker, D.A. Meyers, R.J. Allen, J.M. Oldham, S-F. Ma, I. Noth, R.G. Jenkins, T.M. Maher, R.B. Hubbard, L.V. Wain, T.E. Fingerlin, D.A. Schwartz, G.R. Washko, I.O. Rosas, E.K. Silverman, H. Hatabu, M.H. Cho, G.M. Hunninghake

Critical revision of the manuscript for important intellectual content: B.D. Hobbs, R.K. Putman, T Araki, M. Nishino, G. Gudmundsson, V. Gudnason, G Eiriksdottir, N.R.Z. Nogueira, J. Dupuis, H. Xu, G.T. O'Connor, A. Manichaikul, J. Nguyen, A.J. Podolanczuk, J.I. Rotter, D.J. Lederer, R.G. Barr, S.S. Rich, E.J. Ampleford, V.E. Ortega, S.P. Peters, W.K. O'Neal, J.D. Newell Jr., E.R. Bleecker, D.A. Meyers, R.J. Allen, J.M. Oldham, S-F. Ma, I. Noth, R.G. Jenkins, T.M. Maher, R.B. Hubbard, L.V. Wain, T.E. Fingerlin, D.A. Schwartz, G.R. Washko, I.O. Rosas, E.K. Silverman, H. Hatabu, M.H. Cho, G.M. Hunninghake

Statistical analysis: B.D. Hobbs, R.K. Putman, N.R.Z. Nogueira, G. Eirksdottir, V. Gudnason, J. Dupuis, H. Xu, A. Manichaikul, J. Nguyen, E.J. Ampleford, V.E. Ortega, R.J. Allen, I. Noth, R.G. Jenkins, L.V. Wain, T.E. Fingerlin, D.A. Schwartz, M.H. Cho, G.M. Hunninghake

Obtained funding: V. Gudnason, G.T. O'Connor, D.J. Lederer, R.G. Barr, E.R. Bleecker, D.A. Meyers, I. Noth, R.G. Jenkins, L.V. Wain, D.A. Schwartz, E.K. Silverman, M.H. Cho, G.M. Hunninghake

9.23 Interstitial Lung Disease

Key words: genetics, genome-wide association study, interstitial lung abnormalities, idiopathic pulmonary fibrosis, single-nucleotide polymorphism

Running title:
Word count:
Abstract: 232
Manuscript: 3921

This article has an online data supplement, which is accessible from this issue's table of content

online at www.atsjournals.org.

At a Glance Commentary:

Scientific Knowledge on the Subject: Individuals with interstitial lung abnormalities (ILA) share a clinical syndrome similar to that observed in idiopathic pulmonary fibrosis (IPF) including physiologic decrements, radiologic progression, accelerated lung function decline, and an increased risk of death. ILA are associated with the most common and highest genetic risk locus in IPF, the *MUC5B* promoter polymorphism rs35705950. However, the extent to which there is additional overlap in the genetic risk of ILA and IPF is not known.

What This Study Adds to the Field: In a genome-wide association study of ILA, we confirmed findings at the *MUC5B* locus, and identified 3 novel loci for ILA and subpleural-predominant ILA. These novel loci were not associated with IPF. Additionally, of 12 distinct prior IPF GWAS loci, we identified 11 directionally consistent associations with ILA, of which 7 were at least nominally significant and 5 (near *DPP9*, *DSP*, *FAM13A*, *IVD,* and *MUC5B*) were significantly associated after adjustment for multiple testing.

ABSTRACT

Rationale

Interstitial lung abnormalities (ILA) are associated with the highest genetic risk locus for IPF; however, the extent to which there is additional overlap with IPF, or unique associations among those with ILA is not known.

Objectives

To perform a genome-wide association study (GWAS) of ILA.

Methods: ILA and the subpleural-predominant subtype were assessed on chest computed tomography (CT) scans in the AGES, COPDGene, Framingham Heart, ECLIPSE, MESA, and SPIROMICS studies. We performed a GWAS of ILA in each cohort and combined the results using a meta-analysis. We assessed for overlapping associations in independent GWASs of IPF.

Measurements and Main Results

Genome-wide genotyping data were available in 1,699 ILA cases and 10,274 controls. The *MUC5B* promoter variant rs35705950 was significantly associated with both ILA ($p=2.6x10^{-27}$) and subpleural ILA ($p=1.6x10^{-29}$). We discovered novel genome-wide associations near *IPO11* (rs6886640, $p=3.8x10^{-8}$) and *FCF1P3* (rs73199442, $p=4.8x10^{-8}$) with ILA, and *HTRE1* (rs7744971, $p=4.2x10^{-8}$) with subpleural-predominant ILA. These novel associations were not associated with IPF. Of 12 previously reported IPF GWAS loci, 5 (*DPP9*, *DSP*, *FAM13A*, *IVD,* and *MUC5B*) were significantly associated ($p<0.05/12$) with ILA.

Conclusions

In a GWAS of ILA in six studies, we confirmed the association with a *MUC5B* promoter variant and found strong evidence for an effect of previously described IPF loci; however, novel ILA associations were not associated with IPF. These findings highlight common and suggest distinct genetically-driven biologic pathways between ILA and IPF.

Introduction

Idiopathic pulmonary fibrosis (IPF), the most common and severe form of interstitial lung disease (ILD)(1), is a disorder of lung scarring that is reported to affect 1 in 200 adults >age 65(2) and results in a high rate of mortality(3, 4). A strong genetic basis for pulmonary fibrosis (PF) has been demonstrated in studies of familial aggregation(5), as well as in genome-wide linkage and association studies that have provided replicable evidence for associations between common genetic variants and IPF(6-11). Most consistently, IPF has been associated with increased copies of a common variant (rs35705950) in the promoter of the *mucin 5B* (*MUC5B*) gene(6, 8-11), a finding that may explain up to 30% of the risk of the disease(6).

Due to the severity of physiologic decrements, and the high rate of mortality at the time of diagnosis, recent efforts have sought to identify IPF, and other forms of pulmonary fibrosis, in their earliest stages(12-14). These efforts include chest computed tomography (CT) image characterization of undiagnosed research participants that have classified imaging features not only suggestive of IPF specifically(15-17), but also of the broader set of ILDs (termed interstitial lung abnormalities [ILA])(13). Evidence supporting a correlation between some undiagnosed research participants with ILA and patients with IPF include a shared association with increased copies of the *MUC5B* promoter variant(12, 18). However, the extent to which undiagnosed research participants with ILA and patients with IPF share common, and unique, genetic etiologies remains unclear.

We hypothesized that comparisons between research participants with and without ILA would identify findings of genetic association shared with those identified in patients with IPF and, based on the diversity of ILA phenotypes(18) and ILDs in general(19), unique associations. To test this hypothesis, common genome-wide single nucleotide variants were genotyped, with

additional genotypes imputed from reference panels, and then tested for association with visually assessed ILA, and the subpleural predominant ILA subtype, in populations of research participants from six unique cohorts. Based on the results, further comparisons were performed to examine the overlap of top genetic associations in research participants with ILA (and subpleural predominant ILA) with genetic associations previously reported in patients with IPF(8). Some of the results of this study have been previously reported in the form of an abstract.

Methods

Study population

Protocols for participant enrollment and phenotyping in the Framingham Heart Study (FHS), the Age Gene/Environment Susceptibility (AGES)-Reykjavik, the Genetic Epidemiology of COPD Study (COPDGene), the Evaluation of COPD Longitudinally to Identify Predictive Surrogate End-points (ECLIPSE), the Multi-Ethnic Study of Atherosclerosis (MESA)-Lung Study, and the SubPopulations and InteRmediate Outcome Measures In COPD Study (SPIROMICS) studies have been described previously.(12-14, 20-22) Each cohort obtained approval from appropriate ethical/regulatory bodies and informed consent was obtained for all individuals. More detailed cohort information, including cohort-specific methods, can be found in the **online supplement**.

ILA Characterization

ILA characterization in all cohorts is the result of visual assessments of chest CT scans. In FHS, AGES, COPDGene, and ECLIPSE, CT scans were evaluated for ILA using a sequential reading method by up to three readers (radiologists and pulmonologists), who were blind to all participant specific information, as previously described(23). ILA in these cohorts were defined as nondependent changes affecting greater than 5% of any lung zone. These abnormalities include ground glass, reticular abnormalities, diffuse centrilobular nodularity, multiple

nonemphysematous cysts, traction bronchiectasis, or honeycombing. Chest CT scans with either focal or unilateral ground glass attenuation, focal or unilateral reticulation, or patchy ground glass abnormalities were indeterminate for ILA.(12, 13) In MESA, ILA were assessed by a radiologist using the above criteria, as previously described(24). In SPIROMICS, ILA were classified as present or absent using the following criteria: bilateral, non-dependent, peripheral (but not necessarily subpleural) ground glass and/or reticular opacities and/or honeycombing needed to be present (see the **online supplement** for further details).

Based on prior data on the genetics of ILA(18), and to provide consistency of subtyping across cohorts, an additional ILA subset was created that excluded ILA limited to those with centrilobular nodules alone(13) and includes all those with ILA with predominantly subpleural imaging findings (Subpleural Predominant).

Genotyping and imputation

Details of genotyping in each cohort can be found in the **online supplement**. The genotypes of individuals from European-ancestry studies were imputed via the Michigan Imputation Server using minimac3 with the Haplotype Reference Consortium (HRC v1.1) reference panel(25). The COPDGene and MESA individuals of African-American ancestry were imputed using the 1000 Genomes Project (Phase 3, version 5)(26). The *MUC5B* promoter polymorphism (rs35705950) was poorly imputed (imputation $R^2 < 0.5$) in the AGES study and we instead used direct genotyping data, which was available in 3,209 individuals.

Genome-wide association study and meta-analysis

Given the case-control imbalance, Firth bias-corrected logistic regression(27, 28) was performed in each ancestry-subset of each study, adjusting for age, sex, pack-years of smoking, and ancestry-based principal components, as appropriate for each study. In the Framingham

Heart Study, to allow application of the Firth bias-corrected logistic regression, a subset of unrelated participants was selected for analysis, preferentially choosing ILA cases. Summary statistics from individual studies, including chromosome and position (hg19), effect allele and other allele oriented to the + strand, effect allele frequency, and imputation quality were uploaded to a secure site at the Brigham and Women's Hospital / Channing Division of Network Medicine.

The summary statistics from each study were assessed using EasyQC(29) version 10.1. Quality control assessments included allele frequency comparisons to either a Haplotype Reference Consortium or 1000 Genomes reference panel, standard error versus sample size checking, and quantile-quantile plot visualization. Variants with an imputation quality metric of < 0.5, a minor allele count (MAC) of < 10 (using the effective sample size or the number of cases adjusted for imputation quality, where appropriate), were set to missing. Variant names were all normalized to hg19 chromosome and position. Only the highest frequency alternate allele was retained for multi-allelic variants.

Following summary statistic quality control, we performed an inverse-variance weighted fixed effects meta-analysis in METAL (version 2011-03-25) (30, 31) for both the ILA and the subpleural predominant ILA analyses. In a set of secondary analyses, we performed a meta-analysis restricted to ILA and subpleural ILA results from European ancestry subpopulations and a smoking stratified (ever compared to never smokers) meta-analysis of our genome-wide significant variants. Only variants present in at least half of the cohort subpopulations in each meta-analysis were further evaluated. Genome-wide significance for all associations was considered to be $p < 5x10^{-8}$. To identify distinct results at each locus in European ancestry subpopulations, we used GCTA-COJO(32, 33) on all results with $P < 5x10^{-6}$ using the default

distance of 10Mb. COPDGene non-Hispanic whites (the largest representative population) were the reference population for the GCTA-COJO analysis.

Overlap of ILA genetic loci with idiopathic pulmonary fibrosis, high attenuation areas, smoking behaviors, and connective tissue disease

We evaluated the overlap of top ILA-associated genetic variants with idiopathic pulmonary fibrosis (IPF) in two ways: 1) lookup of IPF GWAS loci from the EBI-NHGRI GWAS Catalog (34) (downloaded 06/04/2018) in our ILA GWAS results; 2) lookup of our top ILA-associated variants ($p < 5e-7$ with either ILA or subpleural ILA in European [EUR] ancestry subpopulation) with IPF in a recent EUR ancestry IPF GWAS and meta-analysis (see **online supplement**)(11). In the lookup of prior IPF GWAS variants in our results, we restricted the lookup to 12 distinct IPF GWAS loci (reported results are for the variant demonstrating the greatest statistical significance at each locus). Significance for association of IPF GWAS variants with ILA or subpleural predominant ILA was set to $p < 0.05/12$. Additional analyses using logistic regression conditioning on the *MUC5B* promoter polymorphism (rs35705950) were done to assess for independence of the multiple SNPs previously identified at the 11p15 locus. To evaluate the overlap between ILA and high attenuation areas (HAA), which have been associated with early or subclinical ILD and future ILA(14, 24), we did a lookup of the previously reported genome wide significant variants associated with HAA(35). To evaluate the overlap between ILA and smoking behaviors we performed a look up of the previously reported genome wide significant variants associated with smoking behaviors(36). To assess a potential overlap of the genetic susceptibility to ILA (and subpleural predominant ILA) with connective tissue disease associated ILD (CTD-ILD) and sarcoidosis, we searched the NHGRI-EBI GWAS Catalog (34) for genome-wide significant SNPs in European-ancestry association studies related to rheumatoid arthritis, sarcoidosis, systemic lupus erythematosus, inflammatory myopathies, and systemic sclerosis. We assessed the p value for association of the genome-wide significant rheumatologic disease

SNPs with ILA and subpleural predominant ILA. We used Bonferroni p values as a significance threshold to correct for multiple testing (number of genome-wide significant SNPs) within each unique trait.

Expression quantitative trait lookups of top GWAS variants

We assessed whether our significant ILA genetic risk variants were expression quantitative trait loci (eQTLs) in the lung and blood using multiple available datasets including GTEx lung and blood eQTLs(37, 38), Westra et al. blood eQTLs(39), Hao et al. lung eQTLs(40), Jansen et al. NESDA/NTR conditional blood eQTLs(41), and the eQTL Consortium blood eQTL meta-analysis data(42). Only cis-eQTLs were assessed; significant associations were determined using the adjusted p-values reported in each available eQTL dataset.

Results

We performed a genome-wide association study (GWAS) and meta-analysis of 1699 ILA cases and 10274 controls in six cohorts, where individuals in each study were stratified into subpopulations according to European, African, and Hispanic ancestry. A secondary GWAS and meta-analysis was performed using the subset of 1287 subpleural predominant ILA cases (**Figure 1**). Baseline characteristics of each cohort, and subpopulation, stratified by ILA status are included in **Table 1** (and in those with ILA limited to the subpleural predominant subtype in, **Table E1**). Similar to prior studies, participants with ILA tended to be older(15) and generally had greater exposure to tobacco smoke than those without ILA.

Genome-wide Association

We identified three genome-wide significant variants associated with ILA including one at 11p15, at the known *MUC5B* promoter polymorphism, rs35705950, (odds ratio [OR] 1.97, 95%

confidence interval [CI] 1.74-2.22, p = $2.6x10^{-27}$) as well as two novel loci including rs6886640 at 5q12, near *IPO11* (OR 1.28, 95% CI 1.18-1.41, p=$3.8x10^{-8}$), and rs73199442 at 3q13, near the lncRNA *FCF1P3* (OR 1.68, 95% CI 1.39-2.02, p=$4.8x10^{-8}$) (**Table 2, Figure 2, supplemental figures E1 and E2**). In the subpleural predominant ILA analysis, in addition to the association with the *MUC5B* variant rs35705950 (OR 2.22, 95% CI 1.93-2.55, p = $1.6x10^{-29}$), we identified a novel genetic association at the 6q15 locus with rs7744971, near *HTR1E* (OR 1.32, 95% CI 95% CI 1.19-1.45, p=$4.2x10^{-8}$) (**Table 2** and **Figure 2**). The ILA risk variant at 3q13.1 (rs73199442) was missing in the African and Hispanic ancestry subpopulations (due to low minor allele frequency), though showed a consistent direction of effect in all European (EUR) ancestry subpopulations (see forest plots in **Figure 2**). Similar results were noted in meta-analyses of ILA and subpleural predominant ILA limited to individuals of European ancestry; however, the 5q12 locus was not significantly associated with ILA and the 6q15 locus was genome-wide significant in association with both ILA and subpleural-predominant ILA (**Supplemental Tables E2** and **E3)**. For each variant demonstrating genome-wide significance we tested for genotype-by-smoking (ever smokers compared to never smokers) interactions. There was no evidence of a significant interaction between smoking status and either of the four genome-wide significant variants (**Supplemental Table E4** and **supplemental figure E3**). To assess whether these novel ILA risk loci overlapped with IPF, we attempted to replicate our genome-wide significant associations with ILA and subpleural ILA associations in an European ancestry GWAS and meta-analysis of 2,668 IPF cases and 8,591 controls(11)(see the **online supplement**). Aside from the known overlap at rs35705950 (IPF p value = $1.2x10^{-203}$), none of our top ILA loci were significantly associated with IPF (**Table 2**).

Assessment of Replication for Prior IPF, HAA, Smoking Behaviors, and Connective Tissue Disease Genetic Loci

We examined the overlap of ILA and subpleural predominant ILA genetic associations with 12 previously reported, distinct idiopathic pulmonary fibrosis (IPF) GWAS loci from the NHGRI-EBI GWAS Catalog. There was a substantial enrichment of the 12 IPF GWAS loci in our ILA association results, where 5 SNPs near *DPP9*, *DSP*, *FAM13A*, *IVD,* and *MUC5B* were significantly associated ($p < 4.2x10^{-3}$) with ILA and an additional 2 SNPs at *MAPT* and *LRRC34*, were nominally significant ($p < 0.05$, but did not meet the threshold for significance after adjustment for multiple testing) in association with ILA (**Table 3** and **supplemental table E5**). All but one of the 12 IPF GWAS SNPs had a consistent direction of risk effect in IPF and ILA. Despite the smaller sample size in the subpleural predominant ILA association analysis, IPF genetic risk loci were generally more strongly associated (larger odds ratios and smaller p values) with risk to subpleural ILA than with ILA.

We assessed the top 21 loci reported in a GWAS of several HAA phenotypes (35) in our ILA and subpleural GWAS results and found no overlap of genetic loci between HAA and ILA; however, the direction of effect between risk to HAA and risk to ILA and subpleural ILA was generally consistent (**Supplemental Table E6**).

To evaluate the overlap in the genetic susceptibility to ILA and smoking behaviors we performed a look up of our 4 genome-wide significant ILA variants in a recent GWAS of four smoking behaviors: smoking initiation, age of smoking initiation, smoking cessation, and cigarettes per day(36) The p value was > 0.05 for association of the 4 top ILA SNPs with any smoking behavior (**Supplemental Table E7**). We also assessed the genome-wide significant loci reported in the smoking GWAS (36); after correction for multiple testing, there was no significant overlap of smoking-behavior SNPs with ILA or subpleural predominant ILA (**Supplemental Table E8**).

In a search for CTD-ILD and sarcoidosis associated SNPs in the NHGRI-EBI GWAS Catalog, we found 357 SNPs associated with 17 traits in reported in 39 publications. No SNPs were associated with ILA. Only one connective tissue disease SNP (rs13389408, intronic to *STAT4* on chromosome 2) met the threshold for Bonferroni significance in association with subpleural ILA (OR 1.3, 95% CI 1.1-1.5, $p=9.7x10^{-4}$). The SNP rs13389408 was discovered in a meta-analysis of "systemic seropositive rheumatic diseases" (including systemic sclerosis, systemic lupus erythematosus, and idiopathic inflammatory myopathies)(43).

Logistic Regression Conditioning on *MUC5B* at 11p15

Our conditional analysis did not identify any conditionally distinct signals at 3q13, 5q12, or 6p15. However, a previous GWAS of idiopathic pulmonary fibrosis (IPF) reported variant associations in *TOLLIP* (rs5743894, rs5743890, rs111521887) that – despite proximity to *MUC5B* in the 11p15 region – were reported independent of the rs35705950 association with IPF due to minimal linkage disequilibrium ($R^2 < 0.2$)(9). In the COPDGene study NHW and AA participants, we performed a meta-analysis of the association of previously reported *TOLLIP* SNPs with ILA and subpleural ILA and conditioned each *TOLLIP* SNP association on the *MUC5B* rs35705950 genotype. When each *TOLLIP* SNP association was adjusted for the rs35705950 genotype, the *TOLLIP* SNPs' effect sizes and strengths of association were diminished (**Supplemental Table E9**). These data suggest the *TOLLIP* SNP associations with ILA and subpleural predominant ILA in COPDGene are not independent of *MUC5B* rs35705950.

eQTL Assessments for Identified Loci

We evaluated if our four genome-wide significant ILA and subpleural ILA-associated variants have been reported as lung or blood expression quantitative trait loci (eQTLs). The ILA risk variants at 5q12 (rs6886640) and 3q13 (rs73199442) were not reported as lung or blood eQTLs

in any of the examined data. The *MUC5B* promoter polymorphism (rs35705950) T allele was associated with increased expression of *MUC5B* in the lung (q value = $3.99x10^{-9}$) in the GTEx lung eQTL data, but not in the Hao et al. lung eQTL data. Further, rs35705950 T allele was associated with decreased expression of *CD151* in blood in the eQTL Consortium cis-eQTL data(42). The subpleural ILA 6q15 variant rs7744971 risk allele (G) was significantly associated with decreased blood expression of *AKIRIN2*, *SLC35A1*, *C6orf164*, and *RP1-102H19.6* and increased blood expression of *ZNF292* in the eQTL Consortium cis-eQTL data (42). In the NESDA NTR Conditional eQTL Catalog (41) the rs7744971 G allele was also associated with decreased *AKIRIN2* expression as well as with decreased *C6orf162* expression in blood.

Discussion

Our study, which presents the first GWAS of visually assessed ILA, includes 1699 ILA cases and 10274 controls and demonstrates several notable findings. First, we provide the most comprehensive data to date demonstrating the links between the genetic association findings of patients with IPF and those of research participants with ILA. For example, these findings demonstrate at least nominal, and directionally consistent, evidence for association between ILA and most of the common genetic variants previously demonstrated to be associated with IPF.(7-11) In addition, our results demonstrate genome-wide significant evidence for association with 2 new genetic risk loci for ILA overall (3q13 and 5q12) and one loci for subpleural predominant ILA specifically (6q15). These new loci do not demonstrate evidence for association with IPF; while these could represent false positive associations, this data may also be consistent with some cases of ILA representing early stages of diverse forms of ILD, associated with genetic risk factors distinct from IPF.

Multiple lines of evidence now demonstrate shared genetic risk between some undiagnosed research participants with ILA and patients with clinically diagnosed IPF. Comparable to all prior genome-wide association studies of IPF that included the gain-of-function *MUC5B* promoter variant(8-11), our study demonstrates that the *MUC5B* promoter variant rs35705950 is the most significant finding of association with ILA. In addition, of 12 loci demonstrating prior genome-wide evidence for association with IPF(7-11) in at least one study, we present evidence for directionally consistent associations in 11 of these 12 loci, of which 5 were significant after adjustment for multiple testing ($p < 0.05/12$). More specifically, this study provides support for the fact that common genetic variants in 7 genomic regions are associated with early and/or mild stages of pulmonary fibrosis in addition to their known association with more advanced stages of disease. Data in support of the latter statement includes the fact that most of these genetic association findings were stronger when the ILA phenotype was limited to those with subpleural reticular involvement (an imaging phenotype we have previously demonstrated to be associated with subpleural fibrosis)(16).

Our study adds to the growing body evidence that demonstrates that increasing copies of the gain-of-function minor allele of the *MUC5B* promoter variant (rs35705950) increases the risk of IPF,(6, 8-11) ILA(12, 18), and other forms of pulmonary fibrosis,(44) perhaps as the result of an increased expression *MUC5B* in the distal airspaces(6, 45-47). Additionally, our conditional logistic regression analyses are consistent with prior analyses(8) which demonstrate that there do not appear to be additional distinct 11p15 genetic variants associated with either ILA, or IPF, after accounting for the effects of the *MUC5B* promoter genotype (rs35705950). Recent studies in mice demonstrate that *Muc5b* overexpression leads to impaired mucociliary clearance and the persistence of pulmonary fibrosis in response to bleomycin challenge which could be mitigated by targeted mucolytic agents.(47) Future studies will be needed to determine if

intervention in those with early stages of *MUC5B* associated interstitial changes(12, 18) could help to prevent progression to more advanced forms of pulmonary fibrosis.

While our study demonstrates that genetic risk factors between patients with IPF and some research participants with ILA are shared, our study identifies genetic loci of some research participants with ILA that are distinct. These results may be consistent with the fact that ILA identify diverse imaging features associated with ILDs in general,(23) and it is possible that some of the interstitial imaging findings of these unique diseases are contributed to by genetic associations that reflect more diverse pathobiologic and clinical processes than are found in patients with IPF.(48) For example, given the fact that many of these cohorts include large populations of smokers, some of these findings could be consistent with a genetic risk to develop smoking-related interstitial fibrosis.(49) Although our results do not demonstrate statistically significant evidence for a genotype-by-smoking interaction, the power for these analyses was limited, and should be evaluated in future studies. In addition, there is some evidence to suggest that genetic risk to develop ILD in some connective tissue diseases may be distinct from those processes leading to IPF.(50-52) Until our findings can be tested in sufficiently large cohorts of these other important populations of ILD, they should be viewed as preliminary.

The association between the minor allele (G) of the SNP rs7744971 on chromosome 6q14 and subpleural ILA deserves some mention as the G allele of this SNP has previously been demonstrated to be an expression quantitative trait locus (eQTL) that is expected to result in a decreased expression of the gene *AKIRIN2*.(41) Basic research has implicated the importance of Akirin-2, known to be expressed in the lung(53), as a critical factor in the innate immune system that helps to regulate inflammatory gene transcription,(54) and B-cell activation.(55)

While Akirin-2 is required for embryonic development,(56) selective knockdown of *akirin2* in myeloid cells in mice results in impaired inflammatory cytokine production in macrophages in response to Toll-like receptor stimulation.(56) Some of the effect of Akirin-2 is felt to be mediated through its importance as a bridge between some transcription factors (e.g. NF-kB(56) and Twist(57) that have also been implicated in pulmonary inflammatory(58) and fibrosis(59) development) and gene transcription. Future work to confirm the role of this variant on expression of *AKIRIN2*, and potentially testing the role of Akirin-2 deficiency in some forms of ILD mediated by pulmonary inflammation may be warranted.

In addition to the lack of statistically significant replication for the novel loci presented in this manuscript, our findings have some additional limitations. First, even though our study includes nearly all the available cohort where ILA characterization and genetic testing have been performed, larger sample sizes may still be required to more adequately characterize the genetic risk of ILA. Second, we cannot exclude the possibility that the generalizability of our findings may be limited by the omission of some participants from these analyses. Third, while we were able to demonstrate genetic associations with ILA and subpleural predominant ILA, sample size may have limited our ability to detect associations with additional specific radiologic features and patterns. In addition, we cannot rule out the possibility that inter-cohort differences in the methods of ILA characterization, or subclassification, could have introduced phenotypic heterogeneity, thus influencing our power to detect genetic associations. Efforts to develop standards of ILA characterization across different research populations could help to minimize this concern. Finally, although our findings demonstrate that the genetic risk to ILA in research participants overlaps with the genetic risk to IPF, it is important to note that we do not know the extent to which genetic risk of IPF is shared with the risk to other forms of pulmonary fibrosis (e.g. ILD in the setting of rheumatoid arthritis),(44) or could result in a more mild or less progressive form of pulmonary fibrosis.

In conclusion, our study demonstrates that while the *MUC5B* promoter variant is the dominant variant that is common between ILA and IPF, there are nominally significant associations between ILA and the majority of genetic loci previously associated with IPF. In addition, our findings demonstrate evidence for novel genetic loci associated with ILA, but not with IPF. These findings provide further evidence that the *DPP9, DSP, FAM13A, IVD,* and *MUC5B* loci may be important in the risk of both early and later stages of pulmonary fibrosis. Our findings also suggest that ILA characterization may also help to identify the genetic risk of developing imaging abnormalities, that may represent the early stage of other diverse forms ILD.

References

1. American Thoracic Society/European Respiratory Society International Multidisciplinary Consensus Classification of the Idiopathic Interstitial Pneumonias. This joint statement of the American Thoracic Society (ATS), and the European Respiratory Society (ERS) was adopted by the ATS board of directors, June 2001 and by the ERS Executive Committee, June 2001. *Am J Respir Crit Care Med* 2002; 165: 277-304.
2. Raghu G, Chen SY, Yeh WS, Maroni B, Li Q, Lee YC, Collard HR. Idiopathic pulmonary fibrosis in US Medicare beneficiaries aged 65 years and older: incidence, prevalence, and survival, 2001-11. *Lancet Respir Med* 2014; 2: 566-572.
3. Raghu G, Collard HR, Egan JJ, Martinez FJ, Behr J, Brown KK, Colby TV, Cordier JF, Flaherty KR, Lasky JA, Lynch DA, Ryu JH, Swigris JJ, Wells AU, Ancochea J, Bouros D, Carvalho C, Costabel U, Ebina M, Hansell DM, Johkoh T, Kim DS, King TE, Jr., Kondoh Y, Myers J, Muller NL, Nicholson AG, Richeldi L, Selman M, Dudden RF, Griss BS, Protzko SL, Schunemann HJ, Fibrosis AEJACoIP. 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.
4. Raghu G, Rochwerg B, Zhang Y, Garcia CA, Azuma A, Behr J, Brozek JL, Collard HR, Cunningham W, Homma S, Johkoh T, Martinez FJ, Myers J, Protzko SL, Richeldi L, Rind D, Selman M, Theodore A, Wells AU, Hoogsteden H, Schunemann HJ, American Thoracic S, European Respiratory s, Japanese Respiratory S, Latin American Thoracic A. An Official ATS/ERS/JRS/ALAT Clinical Practice Guideline: Treatment of Idiopathic Pulmonary Fibrosis. An Update of the 2011 Clinical Practice Guideline. *Am J Respir Crit Care Med* 2015; 192: e3-19.
5. Steele MP, Speer MC, Loyd JE, Brown KK, Herron A, Slifer SH, Burch LH, Wahidi MM, Phillips JA, 3rd, Sporn TA, McAdams HP, Schwarz MI, Schwartz DA. Clinical and pathologic features of familial interstitial pneumonia. *Am J Respir Crit Care Med* 2005; 172: 1146-1152.
6. Seibold MA, Wise AL, Speer MC, Steele MP, Brown KK, Loyd JE, Fingerlin TE, Zhang W, Gudmundsson G, Groshong SD, Evans CM, Garantziotis S, Adler KB, Dickey BF, du Bois RM, Yang IV, Herron A, Kervitsky D, Talbert JL, Markin C, Park J, Crews AL, Slifer SH, Auerbach S, Roy MG, Lin J, Hennessy CE, Schwarz MI, Schwartz DA. A common MUC5B promoter polymorphism and pulmonary fibrosis. *The New England journal of medicine* 2011; 364: 1503-1512.
7. Mushiroda T, Wattanapokayakit S, Takahashi A, Nukiwa T, Kudoh S, Ogura T, Taniguchi H, Kubo M, Kamatani N, Nakamura Y, Pirfenidone Clinical Study G. A genome-wide association study identifies an association of a common variant in TERT with susceptibility to idiopathic pulmonary fibrosis. *J Med Genet* 2008; 45: 654-656.
8. Fingerlin TE, Murphy E, Zhang W, Peljto AL, Brown KK, Steele MP, Loyd JE, Cosgrove GP, Lynch D, Groshong S, Collard HR, Wolters PJ, Bradford WZ, Kossen K, Seiwert SD, du Bois RM, Garcia CK, Devine MS, Gudmundsson G, Isaksson HJ, Kaminski N, Zhang Y, Gibson KF, Lancaster LH, Cogan JD, Mason WR, Maher TM, Molyneaux PL, Wells AU, Moffatt MF, Selman M, Pardo A, Kim DS, Crapo JD, Make BJ, Regan EA, Walek DS, Daniel JJ, Kamatani Y, Zelenika D, Smith K, McKean D, Pedersen BS, Talbert J, Kidd RN, Markin CR,

Beckman KB, Lathrop M, Schwarz MI, Schwartz DA. Genome-wide association study identifies multiple susceptibility loci for pulmonary fibrosis. *Nat Genet* 2013; 45: 613-620.

9. Noth I, Zhang Y, Ma SF, Flores C, Barber M, Huang Y, Broderick SM, Wade MS, Hysi P, Scuirba J, Richards TJ, Juan-Guardela BM, Vij R, Han MK, Martinez FJ, Kossen K, Seiwert SD, Christie JD, Nicolae D, Kaminski N, Garcia JGN. Genetic variants associated with idiopathic pulmonary fibrosis susceptibility and mortality: a genome-wide association study. *Lancet Respir Med* 2013; 1: 309-317.
10. Fingerlin TE, Zhang W, Yang IV, Ainsworth HC, Russell PH, Blumhagen RZ, Schwarz MI, Brown KK, Steele MP, Loyd JE, Cosgrove GP, Lynch DA, Groshong S, Collard HR, Wolters PJ, Bradford WZ, Kossen K, Seiwert SD, du Bois RM, Garcia CK, Devine MS, Gudmundsson G, Isaksson HJ, Kaminski N, Zhang Y, Gibson KF, Lancaster LH, Maher TM, Molyneaux PL, Wells AU, Moffatt MF, Selman M, Pardo A, Kim DS, Crapo JD, Make BJ, Regan EA, Walek DS, Daniel JJ, Kamatani Y, Zelenika D, Murphy E, Smith K, McKean D, Pedersen BS, Talbert J, Powers J, Markin CR, Beckman KB, Lathrop M, Freed B, Langefeld CD, Schwartz DA. Genome-wide imputation study identifies novel HLA locus for pulmonary fibrosis and potential role for auto-immunity in fibrotic idiopathic interstitial pneumonia. *BMC genetics* 2016; 17: 74.
11. Allen RJ, Porte J, Braybrooke R, Flores C, Fingerlin TE, Oldham JM, Guillen-Guio B, Ma SF, Okamoto T, John AE, Obeidat M, Yang IV, Henry A, Hubbard RB, Navaratnam V, Saini G, Thompson N, Booth HL, Hart SP, Hill MR, Hirani N, Maher TM, McAnulty RJ, Millar AB, Molyneaux PL, Parfrey H, Rassl DM, Whyte MKB, Fahy WA, Marshall RP, Oballa E, Bosse Y, Nickle DC, Sin DD, Timens W, Shrine N, Sayers I, Hall IP, Noth I, Schwartz DA, Tobin MD, Wain LV, Jenkins RG. Genetic variants associated with susceptibility to idiopathic pulmonary fibrosis in people of European ancestry: a genome-wide association study. *Lancet Respir Med* 2017; 5: 869-880.
12. Hunninghake GM, Hatabu H, Okajima Y, Gao W, Dupuis J, Latourelle JC, Nishino M, Araki T, Zazueta OE, Kurugol S, Ross JC, San Jose Estepar R, Murphy E, Steele MP, Loyd JE, Schwarz MI, Fingerlin TE, Rosas IO, Washko GR, O'Connor GT, Schwartz DA. MUC5B promoter polymorphism and interstitial lung abnormalities. *The New England journal of medicine* 2013; 368: 2192-2200.
13. Washko GR, Hunninghake GM, Fernandez IE, Nishino M, Okajima Y, Yamashiro T, Ross JC, Estepar RS, Lynch DA, Brehm JM, Andriole KP, Diaz AA, Khorasani R, D'Aco K, Sciurba FC, Silverman EK, Hatabu H, Rosas IO, Investigators CO. Lung volumes and emphysema in smokers with interstitial lung abnormalities. *The New England journal of medicine* 2011; 364: 897-906.
14. Lederer DJ, Enright PL, Kawut SM, Hoffman EA, Hunninghake G, van Beek EJ, Austin JH, Jiang R, Lovasi GS, Barr RG. Cigarette smoking is associated with subclinical parenchymal lung disease: the Multi-Ethnic Study of Atherosclerosis (MESA)-lung study. *Am J Respir Crit Care Med* 2009; 180: 407-414.
15. Putman RK, Hatabu H, Araki T, Gudmundsson G, Gao W, Nishino M, Okajima Y, Dupuis J, Latourelle JC, Cho MH, El-Chemaly S, Coxson HO, Celli BR, Fernandez IE, Zazueta OE, Ross JC, Harmouche R, Estepar RS, Diaz AA, Sigurdsson S, Gudmundsson EF, Eiriksdottir G, Aspelund T, Budoff MJ, Kinney GL, Hokanson JE, Williams MC, Murchison JT, MacNee

W, Hoffmann U, O'Donnell CJ, Launer LJ, Harrris TB, Gudnason V, Silverman EK, O'Connor GT, Washko GR, Rosas IO, Hunninghake GM, Evaluation of CLtIPSEI, Investigators CO. Association Between Interstitial Lung Abnormalities and All-Cause Mortality. *JAMA* 2016; 315: 672-681.
16. Miller ER, Putman, R.K., Vivero, M., Hung, Y., Araki, T., Nishino, M., Hatabu, H., Sholl, L., Hunninghake, G.M. . Interstitial Lung Abnormalities and Histopathologic Correlates in Patients Undergoing Lung Nodule Resection [abstract]. *Am J Respir Crit Care Med 2017; 195: A1120* 2017.
17. Podolanczuk AJ, Oelsner EC, Barr RG, Bernstein EJ, Hoffman EA, Easthausen IJ, Hinckley Stukovsky K, RoyChoudhury A, Michos ED, Raghu G, Kawut SM, Lederer DJ. High Attenuation Areas on Chest CT and Clinical Respiratory Outcomes in Community-dwelling Adults. *Am J Respir Crit Care Med* 2017.
18. Putman RK, Gudmundsson G, Araki T, Nishino M, Sigurdsson S, Gudmundsson EF, Eiriksdottir G, Aspelund T, Ross JC, San Jose Estepar R, Miller ER, Yamada Y, Yanagawa M, Tomiyama N, Launer LJ, Harris TB, El-Chemaly S, Raby BA, Cho MH, Rosas IO, Washko GR, Schwartz DA, Silverman EK, Gudnason V, Hatabu H, Hunninghake GM. The MUC5B promoter polymorphism is associated with specific interstitial lung abnormality subtypes. *The European respiratory journal* 2017; 50.
19. American Thoracic S, European Respiratory S. American Thoracic Society/European Respiratory Society International Multidisciplinary Consensus Classification of the Idiopathic Interstitial Pneumonias. This joint statement of the American Thoracic Society (ATS), and the European Respiratory Society (ERS) was adopted by the ATS board of directors, June 2001 and by the ERS Executive Committee, June 2001. *Am J Respir Crit Care Med* 2002; 165: 277-304.
20. Hurst JR, Vestbo J, Anzueto A, Locantore N, Mullerova H, Tal-Singer R, Miller B, Lomas DA, Agusti A, Macnee W, Calverley P, Rennard S, Wouters EF, Wedzicha JA, Evaluation of CLtIPSEI. Susceptibility to exacerbation in chronic obstructive pulmonary disease. *The New England journal of medicine* 2010; 363: 1128-1138.
21. Harris TB, Launer LJ, Eiriksdottir G, Kjartansson O, Jonsson PV, Sigurdsson G, Thorgeirsson G, Aspelund T, Garcia ME, Cotch MF, Hoffman HJ, Gudnason V. Age, Gene/Environment Susceptibility-Reykjavik Study: multidisciplinary applied phenomics. *Am J Epidemiol* 2007; 165: 1076-1087.
22. Couper D, LaVange LM, Han M, Barr RG, Bleecker E, Hoffman EA, Kanner R, Kleerup E, Martinez FJ, Woodruff PG, Rennard S, Group SR. Design of the Subpopulations and Intermediate Outcomes in COPD Study (SPIROMICS). *Thorax* 2014; 69: 491-494.
23. Washko GR, Lynch DA, Matsuoka S, Ross JC, Umeoka S, Diaz A, Sciurba FC, Hunninghake GM, San Jose Estepar R, Silverman EK, Rosas IO, Hatabu H. Identification of early interstitial lung disease in smokers from the COPDGene Study. *Academic radiology* 2010; 17: 48-53.
24. Podolanczuk AJ, Oelsner EC, Barr RG, Hoffman EA, Armstrong HF, Austin JH, Basner RC, Bartels MN, Christie JD, Enright PL, Gochuico BR, Hinckley Stukovsky K, Kaufman JD, Hrudaya Nath P, Newell JD, Jr., Palmer SM, Rabinowitz D, Raghu G, Sell JL, Sieren J, Sonavane SK, Tracy RP, Watts JR, Williams K, Kawut SM, Lederer DJ. High attenuation

areas on chest computed tomography in community-dwelling adults: the MESA study. *The European respiratory journal* 2016; 48: 1442-1452.
25. the Haplotype Reference C. A reference panel of 64,976 haplotypes for genotype imputation. *Nature genetics* 2016; 48: 1279.
26. Genomes Project C, Abecasis GR, Auton A, Brooks LD, DePristo MA, Durbin RM, Handsaker RE, Kang HM, Marth GT, McVean GA. An integrated map of genetic variation from 1,092 human genomes. *Nature* 2012; 491: 56-65.
27. Firth D. Bias reduction of maximum likelihood estimates. *Biometrika* 1993; 80: 27-38.
28. Ma C, Blackwell T, Boehnke M, Scott LJ, Go TDi. Recommended joint and meta-analysis strategies for case-control association testing of single low-count variants. *Genetic epidemiology* 2013; 37: 539-550.
29. Winkler TW, Day FR, Croteau-Chonka DC, Wood AR, Locke AE, Magi R, Ferreira T, Fall T, Graff M, Justice AE, Luan J, Gustafsson S, Randall JC, Vedantam S, Workalemahu T, Kilpelainen TO, Scherag A, Esko T, Kutalik Z, Heid IM, Loos RJ, Genetic Investigation of Anthropometric Traits C. Quality control and conduct of genome-wide association meta-analyses. *Nature protocols* 2014; 9: 1192-1212.
30. Abecasis G, Li Y, Willer C. METAL MetaAnalysis Helper. Version release 2011-03-25. URL: http://genome.sph.umich.edu/wiki/METAL_Program. 2011.
31. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* 2010; 26: 2190-2191.
32. Yang J, Lee SH, Goddard ME, Visscher PM. GCTA: a tool for genome-wide complex trait analysis. *Am J Hum Genet* 2011; 88: 76-82.
33. Yang J, Ferreira T, Morris AP, Medland SE, Genetic Investigation of ATC, Replication DIG, Meta-analysis C, Madden PA, Heath AC, Martin NG, Montgomery GW, Weedon MN, Loos RJ, Frayling TM, McCarthy MI, Hirschhorn JN, Goddard ME, Visscher PM. Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. *Nature genetics* 2012; 44: 369-375, S361-363.
34. Welter D, MacArthur J, Morales J, Burdett T, Hall P, Junkins H, Klemm A, Flicek P, Manolio T, Hindorff L, Parkinson H. The NHGRI GWAS Catalog, a curated resource of SNP-trait associations. *Nucleic Acids Res* 2014; 42: D1001-1006.
35. Manichaikul A, Wang XQ, Sun L, Dupuis J, Borczuk AC, Nguyen JN, Raghu G, Hoffman EA, Onengut-Gumuscu S, Farber EA, Kaufman JD, Rabinowitz D, Stukovsky KDH, Kawut SM, Hunninghake GM, Washko GR, O'Connor GT, Rich SS, Barr RG, Lederer DJ. Genome-wide association study of subclinical interstitial lung disease in MESA. *Respir Res* 2017; 18: 97.
36. Liu M, Jiang Y, Wedow R, Li Y, Brazel DM, Chen F, Datta G, Davila-Velderrain J, McGuire D, Tian C, Zhan X, andMe Research T, Psychiatry HA-I, Choquet H, Docherty AR, Faul JD, Foerster JR, Fritsche LG, Gabrielsen ME, Gordon SD, Haessler J, Hottenga JJ, Huang H, Jang SK, Jansen PR, Ling Y, Magi R, Matoba N, McMahon G, Mulas A, Orru V, Palviainen T, Pandit A, Reginsson GW, Skogholt AH, Smith JA, Taylor AE, Turman C, Willemsen G, Young H, Young KA, Zajac GJM, Zhao W, Zhou W, Bjornsdottir G, Boardman JD, Boehnke M, Boomsma DI, Chen C, Cucca F, Davies GE, Eaton CB, Ehringer MA, Esko T, Fiorillo E, Gillespie NA, Gudbjartsson DF, Haller T, Harris KM, Heath AC, Hewitt JK, Hickie IB, Hokanson JE, Hopfer CJ, Hunter DJ, Iacono WG, Johnson EO, Kamatani Y, Kardia SLR,

Keller MC, Kellis M, Kooperberg C, Kraft P, Krauter KS, Laakso M, Lind PA, Loukola A, Lutz SM, Madden PAF, Martin NG, McGue M, McQueen MB, Medland SE, Metspalu A, Mohlke KL, Nielsen JB, Okada Y, Peters U, Polderman TJC, Posthuma D, Reiner AP, Rice JP, Rimm E, Rose RJ, Runarsdottir V, Stallings MC, Stancakova A, Stefansson H, Thai KK, Tindle HA, Tyrfingsson T, Wall TL, Weir DR, Weisner C, Whitfield JB, Winsvold BS, Yin J, Zuccolo L, Bierut LJ, Hveem K, Lee JJ, Munafo MR, Saccone NL, Willer CJ, Cornelis MC, David SP, Hinds DA, Jorgenson E, Kaprio J, Stitzel JA, Stefansson K, Thorgeirsson TE, Abecasis G, Liu DJ, Vrieze S. Association studies of up to 1.2 million individuals yield new insights into the genetic etiology of tobacco and alcohol use. *Nat Genet* 2019; 51: 237-244.

37. Consortium GT. The Genotype-Tissue Expression (GTEx) project. *Nature genetics* 2013; 45: 580-585.
38. Consortium GT. Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. *Science* 2015; 348: 648-660.
39. Westra HJ, Peters MJ, Esko T, Yaghootkar H, Schurmann C, Kettunen J, Christiansen MW, Fairfax BP, Schramm K, Powell JE, Zhernakova A, Zhernakova DV, Veldink JH, Van den Berg LH, Karjalainen J, Withoff S, Uitterlinden AG, Hofman A, Rivadeneira F, t Hoen PA, Reinmaa E, Fischer K, Nelis M, Milani L, Melzer D, Ferrucci L, Singleton AB, Hernandez DG, Nalls MA, Homuth G, Nauck M, Radke D, Volker U, Perola M, Salomaa V, Brody J, Suchy-Dicey A, Gharib SA, Enquobahrie DA, Lumley T, Montgomery GW, Makino S, Prokisch H, Herder C, Roden M, Grallert H, Meitinger T, Strauch K, Li Y, Jansen RC, Visscher PM, Knight JC, Psaty BM, Ripatti S, Teumer A, Frayling TM, Metspalu A, van Meurs JB, Franke L. Systematic identification of trans eQTLs as putative drivers of known disease associations. *Nature genetics* 2013; 45: 1238-1243.
40. Hao K, Bosse Y, Nickle DC, Pare PD, Postma DS, Laviolette M, Sandford A, Hackett TL, Daley D, Hogg JC, Elliott WM, Couture C, Lamontagne M, Brandsma CA, van den Berge M, Koppelman G, Reicin AS, Nicholson DW, Malkov V, Derry JM, Suver C, Tsou JA, Kulkarni A, Zhang C, Vessey R, Opiteck GJ, Curtis SP, Timens W, Sin DD. Lung eQTLs to help reveal the molecular underpinnings of asthma. *PLoS genetics* 2012; 8: e1003029.
41. Jansen R, Hottenga JJ, Nivard MG, Abdellaoui A, Laport B, de Geus EJ, Wright FA, Penninx B, Boomsma DI. Conditional eQTL analysis reveals allelic heterogeneity of gene expression. *Human molecular genetics* 2017; 26: 1444-1451.
42. Võsa U, Claringbould A, Westra H-J, Bonder MJ, Deelen P, Zeng B, Kirsten H, Saha A, Kreuzhuber R, Kasela S, Pervjakova N, Alvaes I, Fave M-J, Agbessi M, Christiansen M, Jansen R, Seppälä I, Tong L, Teumer A, Schramm K, Hemani G, Verlouw J, Yaghootkar H, Sönmez R, Andrew AA, Kukushkina V, Kalnapenkis A, Rüeger S, Porcu E, Kronberg-Guzman J, Kettunen J, Powell J, Lee B, Zhang F, Arindrarto W, Beutner F, , Brugge H, , Dmitrieva J, Elansary M, Fairfax BP, Georges M, Heijmans BT, Kähönen M, Kim Y, Knight JC, Kovacs P, Krohn K, Li S, Loeffler M, Marigorta UM, Mei H, Momozawa Y, Müller-Nurasyid M, Nauck M, Nivard M, Penninx B, Pritchard J, Raitakari O, Rotzschke O, Slagboom EP, Stehouwer CDA, Stumvoll M, Sullivan P, 't Hoen PAC, Thiery J, Tönjes A, van Dongen J, van Iterson M, Veldink J, Völker U, Wijmenga C, Swertz M, Andiappan A, Montgomery GW, Ripatti S, Perola M, Kutalik Z, Dermitzakis E, Bergmann S, Frayling T, van Meurs J, Prokisch H, Ahsan H, Pierce B, Lehtimäki T, Boomsma D, Psaty BM, Gharib

SA, Awadalla P, Milani L, Ouwehand WH, Downes K, Stegle O, Battle A, Yang J, Visscher PM, Scholz M, Gibson G, Esko T, Franke L. Unraveling the polygenic architecture of complex traits using blood eQTL meta-analysis. *bioRxiv* 2018.

43. Acosta-Herrera M, Kerick M, Gonzalez-Serna D, Myositis Genetics C, Scleroderma Genetics C, Wijmenga C, Franke A, Gregersen PK, Padyukov L, Worthington J, Vyse TJ, Alarcon-Riquelme ME, Mayes MD, Martin J. Genome-wide meta-analysis reveals shared new loci in systemic seropositive rheumatic diseases. *Ann Rheum Dis* 2019; 78: 311-319.
44. Juge PA, Lee JS, Ebstein E, Furukawa H, Dobrinskikh E, Gazal S, Kannengiesser C, Ottaviani S, Oka S, Tohma S, Tsuchiya N, Rojas-Serrano J, Gonzalez-Perez MI, Mejia M, Buendia-Roldan I, Falfan-Valencia R, Ambrocio-Ortiz E, Manali E, Papiris SA, Karageorgas T, Boumpas D, Antoniou K, van Moorsel CHM, van der Vis J, de Man YA, Grutters JC, Wang Y, Borie R, Wemeau-Stervinou L, Wallaert B, Flipo RM, Nunes H, Valeyre D, Saidenberg-Kermanac'h N, Boissier MC, Marchand-Adam S, Frazier A, Richette P, Allanore Y, Sibilia J, Dromer C, Richez C, Schaeverbeke T, Liote H, Thabut G, Nathan N, Amselem S, Soubrier M, Cottin V, Clement A, Deane K, Walts AD, Fingerlin T, Fischer A, Ryu JH, Matteson EL, Niewold TB, Assayag D, Gross A, Wolters P, Schwarz MI, Holers M, Solomon J, Doyle T, Rosas IO, Blauwendraat C, Nalls MA, Debray MP, Boileau C, Crestani B, Schwartz DA, Dieude P. MUC5B Promoter Variant and Rheumatoid Arthritis with Interstitial Lung Disease. *N Engl J Med* 2018.
45. Kropski JA, Pritchett JM, Zoz DF, Crossno PF, Markin C, Garnett ET, Degryse AL, Mitchell DB, Polosukhin VV, Rickman OB, Choi L, Cheng DS, McConaha ME, Jones BR, Gleaves LA, McMahon FB, Worrell JA, Solus JF, Ware LB, Lee JW, Massion PP, Zaynagetdinov R, White ES, Kurtis JD, Johnson JE, Groshong SD, Lancaster LH, Young LR, Steele MP, Phillips Iii JA, Cogan JD, Loyd JE, Lawson WE, Blackwell TS. Extensive Phenotyping of Individuals At-risk for Familial Interstitial Pneumonia Reveals Clues to the Pathogenesis of Interstitial Lung Disease. *Am J Respir Crit Care Med* 2014.
46. Nakano Y, Yang IV, Walts AD, Watson AM, Helling BA, Fletcher AA, Lara AR, Schwarz MI, Evans CM, Schwartz DA. MUC5B Promoter Variant rs35705950 Affects MUC5B Expression in the Distal Airways in Idiopathic Pulmonary Fibrosis. *Am J Respir Crit Care Med* 2016; 193: 464-466.
47. Hancock LA, Hennessy CE, Solomon GM, Dobrinskikh E, Estrella A, Hara N, Hill DB, Kissner WJ, Markovetz MR, Grove Villalon DE, Voss ME, Tearney GJ, Carroll KS, Shi Y, Schwarz MI, Thelin WR, Rowe SM, Yang IV, Evans CM, Schwartz DA. Muc5b overexpression causes mucociliary dysfunction and enhances lung fibrosis in mice. *Nat Commun* 2018; 9: 5363.
48. Travis WD, Costabel U, Hansell DM, King TE, Jr., Lynch DA, Nicholson AG, Ryerson CJ, Ryu JH, Selman M, Wells AU, Behr J, Bouros D, Brown KK, Colby TV, Collard HR, Cordeiro CR, Cottin V, Crestani B, Drent M, Dudden RF, Egan J, Flaherty K, Hogaboam C, Inoue Y, Johkoh T, Kim DS, Kitaichi M, Loyd J, Martinez FJ, Myers J, Protzko S, Raghu G, Richeldi L, Sverzellati N, Swigris J, Valeyre D, Pneumonias AECoII. An official American Thoracic Society/European Respiratory Society statement: Update of the international multidisciplinary classification of the idiopathic interstitial pneumonias. *Am J Respir Crit Care Med* 2013; 188: 733-748.

49. Katzenstein AL. Smoking-related interstitial fibrosis (SRIF): pathologic findings and distinction from other chronic fibrosing lung diseases. *Journal of clinical pathology* 2013; 66: 882-887.
50. Johnson C RP, Lloyd T, McGready J, Horton M, Hall J, Mammen A, Danoff S. MUC5B Promoter Variant Is Not Associated With Myositis-Related Interstitial Lung Disease. Chest. Atlanta: Chest Journal 2012.
51. Peljto AL, Steele MP, Fingerlin TE, Hinchcliff ME, Murphy E, Podlusky S, Carns M, Schwarz M, Varga J, Schwartz DA. The pulmonary fibrosis-associated MUC5B promoter polymorphism does not influence the development of interstitial pneumonia in systemic sclerosis. *Chest* 2012; 142: 1584-1588.
52. Borie R, Crestani B, Dieude P, Nunes H, Allanore Y, Kannengiesser C, Airo P, Matucci-Cerinic M, Wallaert B, Israel-Biet D, Cadranel J, Cottin V, Gazal S, Peljto AL, Varga J, Schwartz DA, Valeyre D, Grandchamp B. The MUC5B variant is associated with idiopathic pulmonary fibrosis but not with systemic sclerosis interstitial lung disease in the European Caucasian population. *PLoS One* 2013; 8: e70621.
53. Bosch PJ, Fuller LC, Weiner JA. An essential role for the nuclear protein Akirin2 in mouse limb interdigital tissue regression. *Sci Rep* 2018; 8: 12240.
54. Tartey S, Takeuchi O. Chromatin Remodeling and Transcriptional Control in Innate Immunity: Emergence of Akirin2 as a Novel Player. *Biomolecules* 2015; 5: 1618-1633.
55. Tartey S, Matsushita K, Imamura T, Wakabayashi A, Ori D, Mino T, Takeuchi O. Essential Function for the Nuclear Protein Akirin2 in B Cell Activation and Humoral Immune Responses. *J Immunol* 2015; 195: 519-527.
56. Goto A, Matsushita K, Gesellchen V, El Chamy L, Kuttenkeuler D, Takeuchi O, Hoffmann JA, Akira S, Boutros M, Reichhart JM. Akirins are highly conserved nuclear proteins required for NF-kappaB-dependent gene expression in drosophila and mice. *Nat Immunol* 2008; 9: 97-104.
57. Nowak SJ, Aihara H, Gonzalez K, Nibu Y, Baylies MK. Akirin links twist-regulated transcription with the Brahma chromatin remodeling complex during embryogenesis. *PLoS Genet* 2012; 8: e1002547.
58. Christman JW, Sadikot RT, Blackwell TS. The role of nuclear factor-kappa B in pulmonary diseases. *Chest* 2000; 117: 1482-1487.
59. Tan J, Tedrow JR, Nouraie M, Dutta JA, Miller DT, Li X, Yu S, Chu Y, Juan-Guardela B, Kaminski N, Ramani K, Biswas PS, Zhang Y, Kass DJ. Loss of Twist1 in the Mesenchymal Compartment Promotes Increased Fibrosis in Experimental Lung Injury by Enhanced Expression of CXCL12. *J Immunol* 2017; 198: 2269-2285.

Table 1. Baseline characteristics of participants stratified by interstitial lung abnormality (ILA) status in each cohort.

	AGES-Reykjavik		COPDGene Non-Hispanic Whites		COPDGene African-Americans		ECLIPSE		Framingham Heart Study		MESA Non-Hispanic Whites		MESA Hispanics		MESA African-Americans		SPIROMICS	
	No ILA (1785, 86%)	ILA (279, 14%)	No ILA (3771, 88%)	ILA (497, 12%)	No ILA (1717, 89%)	ILA (214, 11%)	No ILA (494, 77%)	ILA (151, 23%)	No ILA (530, 79%)	ILA (138, 21%)	No ILA (675, 85%)	ILA (116, 15%)	No ILA (385, 88%)	ILA (54, 12%)	No ILA (467, 88%)	ILA 66, 12%)	No ILA (450, 71%)	ILA (184, 29%)
Age – yrs, mean (SD*)	76 (5)	77 (5)	61 (9)	64 (9)	54 (7)	55 (8)	62 (7)	64 (8)	58 (11)	71 (11)	69 (9)	75 (10)	67 (9)	74 (9)	69 (9)	74 (8)	65 (8)	68 (8)
Sex – no. female (%)	1078 (60)	125 (45)	1814 (48)	226 (45)	701 (41)	116 (54)	166 (34)	40 (26)	278 (53)	67 (49)	344 (51)	60 (52)	209 (54)	29 (54)	247 (53)	41 (62)	206 (46)	83 (45)
Body Mass Index, mean (SD)	27 (4)	27 (5)	29 (6)	29 (6)	29 (7)	29 (7)	27 (6)	26 (5)	28 (5)	28 (5)	28 (5)	28 (6)	30 (6)	30 (6)	30 (6)	30 (5)	27 (5)	29 (5)
Pack-years Smoking, median (IQ†)	2 (0, 26)	21 (0, 51)	40 (29, 56)	45 (34, 63)	34 (22, 46)	35 (24, 47)	45 (33, 62)	43 (29, 61)	0 (0, 12)	8 (0, 23)	14 (3, 34)	15 (2, 33)	3 (0, 11)	13 (2, 28)	11 (3, 24)	18 (6, 33)	41 (30, 60)	50 (37, 65)
Smoking Status – no. (%) Current Former Never	211 (12) 756 (42) 818 (46)	46 (16) 162 (58) 71 (26)	1426 (38) 2345 (62) --	263 (53) 234 (47) --	1377 (80) 340 (20) --	178 (83) 36 (17) --	189 (38) 305 (62) --	69 (46) 82 (54) --	33 (6) 238 (45) 259 (49)	11 (8) 74 (54) 53 (38)	43 (6) 322 (48) 310 (46)	8 (7) 68 (59) 40 (35)	20 (5) 158 (41) 207 (54)	6 (11) 25 (46) 23 (43)	50 (11) 206 (44) 211 (45)	8 (12) 38 (58) 20 (30)	126 (28) 294 (65) 30 (7)	59 (32) 119 (65) 6 (3)
History of COPD‡ – no. (%)	--	--	1527 (40)	171 (34)	380 (22)	57 (27)	494 (100)	151 (100)	46 (9)	18 (13)	129 (22)	21 (21)	62 (17)	11 (23)	98 (24)	14 (25)	290 (64)	116 (63)

*SD is standard deviation
†IQ is interquartile interval
‡COPD is chronic obstructive pulmonary disease and defined as FEV1/FVC ratio < 70 on spirometry
Missing data: MESA Non-Hispanic Whites COPD status: No ILA – 76, ILA – 17; MESA African-Americans COPD Status: No ILA – 58, ILA – 9; Framingham Heart Study Body Mass Index – 1; AGES-Reykjavik Body Mass Index – 1.

Table 2. Genome wide significant variants associated with interstitial lung abnormalities (ILA) and subpleural predominant ILA, and replication in an idiopathic pulmonary fibrosis (IPF) cohort.

Chromosome/ Location	Position	rsID	Risk Allele	Risk Allele Frequency	Nearest Gene	ILA* vs No ILA		Subpleural ILA vs No ILA		Replication in IPF† Cohort	
						Odds Ratio‡ (95% CI§)	P-Value	Odds Ratio (95% CI)	P-Value	Odds Ratio (95% CI)	P-Value
3q13	106571023	rs73199442	T	0.06	*FCF1P3*	1.68 (1.39, 2.02)	5×10^{-8}	1.61 (1.31, 1.99)	7×10^{-6}	0.98 (0.85, 1.12)	0.73
5q12	62172476	rs6886640	G	0.62	*IPO11*	1.28 (1.18, 1.41)	4×10^{-8}	1.27 (1.14, 1.40)	8×10^{-6}	1.06 (0.99, 1.14)	0.11
6q15	87737841	rs7744971	G	0.28	*HTR1E*	1.26 (1.16, 1.37)	1×10^{-7}	1.32 (1.19, 1.45)	4×10^{-8}	1.01 (0.94, 1.09)	0.75
11p15	1241221	rs35705950	T	0.11	*MUC5B*	1.97 (1.74, 2.22)	3×10^{-27}	2.22 (1.93, 2.55)	2×10^{-29}	4.84 (4.37, 5.36)	1×10^{-203}

* ILA is interstitial lung abnormalities
†IPF is idiopathic pulmonary fibrosis
‡Odds Ratios are per copy of the risk allele
§CI is confidence interval

Table 3. Association of 12 prior idiopathic pulmonary fibrosis (IPF) genome-wide association loci with ILA and subpleural ILA*

Chromosome/ Location	rsID	IPF Risk Allele	Nearest Gene	Studies	IPF Odds Ratio (95% CI†)	ILA vs No ILA		Subpleural ILA vs No ILA		Direction of Effect Consistent with Prior Reports
						Odds Ratio (95% CI)	P-Value	Odds Ratio (95% CI)	P-Value	
4q22	rs2609255	G	*FAM13A*	Fingerlin, NG‡, 2013	1.29 (1.18, 1.42)	1.18 (1.07, 1.29)	$5x10^{-4}$	1.22 (1.09, 1.35)	$3x10^{-4}$	Yes
6p24	rs2076295	G	*DSP*	Fingerlin, NG, 2013 Allen, LRMǁ, 2017	1.44 (1.35, 1.54)	1.14 (1.05, 1.2)	0.001	1.18 (1.08, 1.29)	$3x10^{-4}$	Yes
11p15	rs35705950	T	*MUC5B*	Fingerlin, NG, 2013	2.43 (2.13, 2.77)	1.97 (1.74, 2.22)	$3x10^{-27}$	2.22 (1.93, 2.55)	$2x10^{-29}$	Yes
15q15	rs2034650	A	*IVD*	Fingerlin, NG, 2013	1.30 (1.19, 1.41)	1.08 (0.99, 1.17)	0.07	1.15 (1.05, 1.26)	0.003	Yes
19p13	rs12610495	G	*DPP9*	Fingerlin, NG, 2013	1.29 (1.18, 1.41)	1.14 (1.03, 1.26)	0.01	1.23 (1.10, 1.37)	$2x10^{-4}$	Yes
3q26	rs6793295	C	*LRRC34*	Fingerlin, NG, 2013	1.30 (1.19 1.42)	1.06 (0.97, 1.15)	0.20	1.12 (1.01, 1.24)	0.03	Yes
17q21	rs1981997	G	*MAPT*	Fingerlin, NG, 2013	1.41 (1.28, 1.56)	1.16 (1.03, 1.30)	0.01	1.19 (1.05, 1.36)	0.009	Yes
5p15	rs2736100	A	*TERT*	Fingerlin, NG, 2013 Mushiroda, JMG**, 2008	2.11 (1.61, 2.78)	1.03 (0.95, 1.12)	0.44	1.06 (0.96, 1.16)	0.23	Yes
10q24	rs11191865	A	*OBFC1*	Fingerlin, NG, 2013	1.25 (1.15, 1.35)	1.03 (0.95, 1.12)	0.46	1.03 (0.94, 1.13)	0.56	Yes

13q34	rs1278769	G	*ATP11A*	Fingerlin, NG, 2013	1.27 (1.14, 1.39)	1.04 (0.95, 1.15)	0.37	1.04 (0.94, 1.16)	0.45	Yes
15q25	rs62025270	A	*AKAP13*	Allen, LRM, 2017	1.27 (1.18, 1.37)	1.09 (0.99, 1.20)	0.08	1.07 (0.96, 1.20)	0.23	Yes
7q22	rs4727443	C	*LOC100128334/ LOC105375423*	Fingerlin, NG, 2013	1.30 (1.19, 1.41)	0.95 (0.87, 1.03)	0.19	0.93 (0.84, 1.02)	0.12	No

*IPF is idiopathic pulmonary fibrosis, ILA is interstitial lung abnormalities. For each loci only the single nucleotide polymorphism is reported.

† CI is confidence interval

‡ Fingerlin TE, Murphy E, Zhang W, et.al. Genome-wide association study identifies multiple susceptibility loci for pulmonary fibrosis. *Nat Genet* 2013; 45: 613-620.

‖Allen RJ, Porte J, Braybrooke R, et.al. Genetic variants associated with susceptibility to idiopathic pulmonary fibrosis in people of European ancestry: a genome-wide association study. *Lancet Respir Med* 2017; 5: 869-88.

**Mushiroda T, Wattanapokayakit S, Takahashi A, et.al. A genome-wide association study identifies an association of a common variant in TERT with susceptibility to idiopathic pulmonary fibrosis. *J Med Genet* 2008; 45: 654-656.

Figure 1. Flowchart depicting the participants included and excluded from the genome wide association analysis by cohort and interstitial lung abnormality (ILA) status.

Figure 2. Locus zoom and forest plots for the genome wide significant loci associated with interstitial lung abnormalities (ILA) and subpleural predominant ILA. Panel A is the comparison of ILA to those without ILA, A1 is the locus zoom plot demonstrating the genome wide significant association at rs35705950 (nearest gene *MUC5B*), A2 is the forest plot demonstrating the results in each individual cohort and the overall meta-analysis, with x-axis on the log odds scale. Panel B is the comparison of ILA to those without ILA, B1 is the locus zoom plot demonstrating the genome wide significant association at rs6886640 (nearest gene *IPO11*), B2 is the forest plot demonstrating the results in each individual cohort and the overall meta-analysis, with x-axis on the log odds scale. Panel C is the comparison of ILA to those without ILA, C1 is the locus zoom plot demonstrating the genome wide significant association at rs73199442 (nearest gene *FCF1P3*), C2 is the forest plot demonstrating the results in each individual cohort and the overall meta-analysis, with x-axis on the log odds scale. Panel D is the comparison of subpleural predominant ILA to those without ILA, D1 is the locus zoom plot demonstrating the genome wide significant association at rs7744971 (nearest gene *HTR1E*), D2 is the forest plot demonstrating the results in each individual cohort and the overall meta-analysis, with x-axis on the log odds scale.