ORIGINAL ARTICLE

Rare and Common Variants in *KIF15* Contribute to Genetic Risk of Idiopathic Pulmonary Fibrosis

David Zhang[1], Gundula Povysil[2], Philippe H. Kobeissy[1], Qi Li[1], Binhan Wang[1], Mason Amelotte[1], Hager Jaouadi[1], Chad A. Newton[3], Toby M. Maher[4,5], Philip L. Molyneaux[5], Imre Noth[6], Fernando J. Martinez[7], Ganesh Raghu[8], Jamie L. Todd[9,10], Scott M. Palmer[9,10], Carolina Haefliger[11], Adam Platt[12], Slavé Petrovski[11,13], Joseph A. Garcia[1], David B. Goldstein[1,2], and Christine Kim Garcia[1,2]

[1]Department of Medicine and [2]Institute for Genomic Medicine, Irving Medical Center, Columbia University, New York, New York; [3]Department of Medicine, University of Texas Southwestern Medical Center, Dallas, Texas; [4]Keck School of Medicine, University of Southern California, Los Angeles, California; [5]National Heart and Lung Institute, Imperial College London, London, United Kingdom; [6]Department of Medicine, School of Medicine, University of Virginia, Charlottesville, Virginia; [7]Department of Medicine, Weill-Cornell Medical Center, New York, New York; [8]Department of Medicine, University of Washington Medical Center, Seattle, Washington; [9]Department of Medicine, Duke University Medical Center, Durham, North Carolina; [10]Duke Clinical Research Institute, Durham, North Carolina; [11]Centre for Genomics Research, Discovery Sciences, and [12]Translational Science and Experimental Medicine, Research and Early Development, Respiratory and Immunology, BioPharmaceuticals R&D, AstraZeneca, Cambridge, United Kingdom; and [13]Department of Medicine, University of Melbourne, Melbourne, Victoria, Australia

ORCID IDs: 0000-0002-2983-9796 (D.Z.); 0000-0003-4625-5909 (G.P.); 0000-0002-6806-0072 (H.J.); 0000-0001-5256-9029 (C.A.N.); 0000-0001-7192-9149 (T.M.M.); 0000-0002-2412-3182 (F.J.M.); 0000-0001-7506-6643 (G.R.); 0000-0003-4247-3693 (J.L.T.); 0000-0002-5095-5716 (C.H.); 0000-0002-3455-1789 (A.P.); 0000-0002-0771-1249 (C.K.G.).

Abstract

Rationale: Genetic studies of idiopathic pulmonary fibrosis (IPF) have improved our understanding of this disease, but not all causal loci have been identified.

Objectives: To identify genes enriched with rare deleterious variants in IPF and familial pulmonary fibrosis.

Methods: We performed gene burden analysis of whole-exome data, tested single variants for disease association, conducted KIF15 (kinesin family member 15) functional studies, and examined human lung single-cell RNA sequencing data.

Measurements and Main Results: Gene burden analysis of 1,725 cases and 23,509 control subjects identified heterozygous rare deleterious variants in *KIF15*, a kinesin involved in spindle separation during mitosis, and three telomere-related genes (*TERT* [telomerase reverse transcriptase], *RTEL1* [regulator of telomere elongation helicase 1], and *PARN* [poly(A)-specific ribonuclease]). KIF15 was implicated in autosomal-dominant models of rare deleterious variants (odds ratio [OR], 4.9; 95% confidence interval [CI], 2.7–8.8; $P = 2.55 \times 10^{-7}$) and rare protein-truncating variants (OR, 7.6; 95% CI, 3.3–17.1; $P = 8.12 \times 10^{-7}$). Meta-analyses of the discovery and replication cohorts, including 2,966 cases and 29,817 control subjects, confirm the involvement of *KIF15* plus the three telomere-related genes. A common variant within a *KIF15* intron (rs74341405; OR, 1.6; 95% CI, 1.4–1.9; $P = 5.63 \times 10^{-10}$) is associated with IPF risk, confirming a prior report. Lymphoblastoid cells from individuals heterozygous for the common variant have decreased KIF15 and reduced rates of cell growth. Cell proliferation is dependent on KIF15 in the presence of an inhibitor of Eg5/KIF11, which has partially redundant function. *KIF15* is expressed specifically in replicating human lung cells and shows diminished expression in replicating epithelial cells of patients with IPF.

Conclusions: Both rare deleterious variants and common variants in *KIF15* link a nontelomerase pathway of cell proliferation with IPF susceptibility.

Keywords: IPF; KIF15; spindle; cell proliferation; genetics

(Received in original form October 28, 2021; accepted in final form April 12, 2022)

Supported by NIH grant R01HL093096 (C.K.G.), NHLBI grant K23HL148498, and National Center for Advancing Translational Sciences grant UL1TR001105 (C.A.N.); the Stony Wold-Herbert Fund (D.Z.); and AstraZeneca (C.H., A.P., and S.P.). This paper was prepared using research materials obtained from the NHLBI Biologic Specimen and Data Repository Information Coordinating Center and does not necessarily reflect the opinions of PANTHER-IPF, ACE, or the NHLBI.

Author Contributions: Study design: C.K.G., D.B.G., C.H., S.P., and J.A.G.; patient recruitment or data sharing: C.K.G., C.A.N., T.M.M., P.L.M., I.N., F.J.M., G.R., J.L.T., S.M.P., A.P., and S.P.; genetic analysis: D.Z., G.P., Q.L., and B.W.; investigation of gene function: J.A.G., P.H.K., D.Z., M.A., and H.J.; interpretation of results: D.Z., G.P., P.H.K., J.A.G., D.B.G., and C.K.G.; and manuscript preparation: D.Z. and C.K.G. All authors contributed to the manuscript review and approved the submitted draft.

Am J Respir Crit Care Med Vol 206, Iss 1, pp 56–69, Jul 1, 2022

Originally Published in Press as DOI: 10.1164/rccm.202110-2439OC on April 13, 2022
Internet address: www.atsjournals.org

At a Glance Commentary

Scientific Knowledge on the Subject: Identification of rare pathologic variants provides personalized genetic information for patients and their family members. Studies to find these rare variants are hampered by their low frequencies and the need for whole-exome sequencing data and specialized analytic tools.

What this Study Adds to the field: In this study, collapsing analysis of patients with idiopathic pulmonary fibrosis (IPF) and familial pulmonary fibrosis from large, international, multisite cohorts provides evidence of rare variants in *KIF15* (kinesin family member 15) exceeding genome- and study-wide significance. A nearby common variant was previously implicated in a genome-wide association study. *In vitro* studies demonstrate that the common and rare variants lead to decreased KIF15 protein expression and reduced rates of cell proliferation. KIF15 is a kinesin that, together with KIF11, is involved with spindle separation during mitosis. Single-cell RNA sequencing of human lung demonstrates that *KIF15* is specifically expressed in replicating cells, with decreased expression in proliferating epithelial cells from patients with IPF. These findings suggest that decreased replicative reserves, especially in epithelial cells, underlies genetic susceptibility to IPF. Evidence of joint effects of rare and common variants in the same gene, although rare in medicine, provides a strong framework of integrated effects from a telomerase-independent pathway of cell proliferation to IPF susceptibility.

Idiopathic pulmonary fibrosis (IPF) is the prototypic fibrotic interstitial lung disease that affects older adults (1). It is characterized by unrelenting progression and reduced life expectancy (2). Antifibrotic medications decrease the rate of disease progression but do not stop or reverse the fibrotic process. Better understanding of the genetic basis of IPF provides insights into disease pathogenesis that may lead to more effective treatments targeting its underlying etiology.

Genome-wide association studies (GWASs) have reported associations between a number of different loci and IPF (3, 4). Of these, the common variant with the largest effect is the SNP residing in the *MUC5B* (mucin 5B, oligomeric mucus/gel-forming) promoter (rs35705950). Another SNP located between *KIF15* (kinesin family member 15) and an adjacent gene, *TMEM42* (transmembrane protein 42) (rs78238620), was recently identified (4). Like other GWAS hits in noncoding regions, the function of this disease-associated risk allele is unclear.

Genetic predisposition to pulmonary fibrosis is strongly supported by families with multiple individuals affected with various forms of pulmonary fibrosis, typically demonstrating an autosomal-dominant pattern of inheritance and incomplete penetrance (5). Early studies identified mutations in the surfactant genes (*SFTPC* [surfactant protein C], *SFTPA1/2* [surfactant protein A1/A2]), but these explain a very small percentage of familial pulmonary fibrosis (FPF) or IPF cases. Rare coding, heterozygous, deleterious variants in telomere-related genes (*TERT* [telomerase reverse transcriptase], *TERC* [telomerase RNA component], *RTEL1* [regulator of telomere elongation helicase 1], and *PARN* [poly(A)-specific ribonuclease], among others) are found in ~25% of FPF probands and ~10% of sporadic cases (6, 7). Both rare and common variants in telomerase (*TERT* or *TERC*) result in telomere shortening and have been linked to IPF susceptibility in linkage, candidate gene, genome-wide association, Mendelian randomization, and next-generation sequencing studies (3, 6–10).

By focusing on protein-coding variants, a more direct link can be made between variants and their functional consequences. The contribution of rare variants to human disease has been revolutionized by next-generation sequencing and refined methods for analyzing these data (11, 12). In this study, which represents the largest IPF whole-exome sequencing (WES) study to date, we identify *KIF15* as a novel IPF gene. This finding is robust, as it is seen in two different genetic models, exceeds genome- and study-wide significance, and is confirmed by the analysis of both discovery and replication cohorts. A common variant within a *KIF15* intron (rs74341405) is associated with IPF risk, confirming a prior study. We evaluated the functional consequences of both common and rare *KIF15* risk alleles and found that both result in decreased cellular proliferation. *KIF15* is specifically expressed in replicating cells in the human lung; thus, KIF15 links a nontelomerase pathway of cell proliferation with IPF susceptibility.

Methods

More detailed methods are provided in the online supplement.

Patient Cohorts

The institutional review board at Columbia University Medical Center (CUMC; #AAAS0753 and #AAAAS7495) approved

Data sharing: The position of the DNA and protein variants are described using KIF15 (kinesin family member 15) NM_020242.3 and NP_064627.1. Accession numbers for whole-genome sequencing data from the Database of Genotypes and Phenotypes are as follows: phs002692 (portion of discovery cohort), phs001607.v2 (pulmonary fibrosis), phs0001416.v2 (MESA), and phs000974.v4 (Framingham Heart Study). Accession numbers for the human lung single-cell data are GSE132771 and GSE135893. Data from third parties are available under the terms of a data-use agreement compliant with ethical and legal requirements.

Correspondence and requests for reprints should be addressed to Christine Kim Garcia, M.D., Ph.D., Division of Pulmonary, Allergy, and Critical Care, Department of Medicine, Columbia University Medical Center, 630 West 168th Street, New York, NY 10032. E-mail: ckg2116@cumc.columbia.edu.

This article has a related editorial.

This article has an online supplement, which is accessible from this issue's table of contents at www.atsjournals.org.

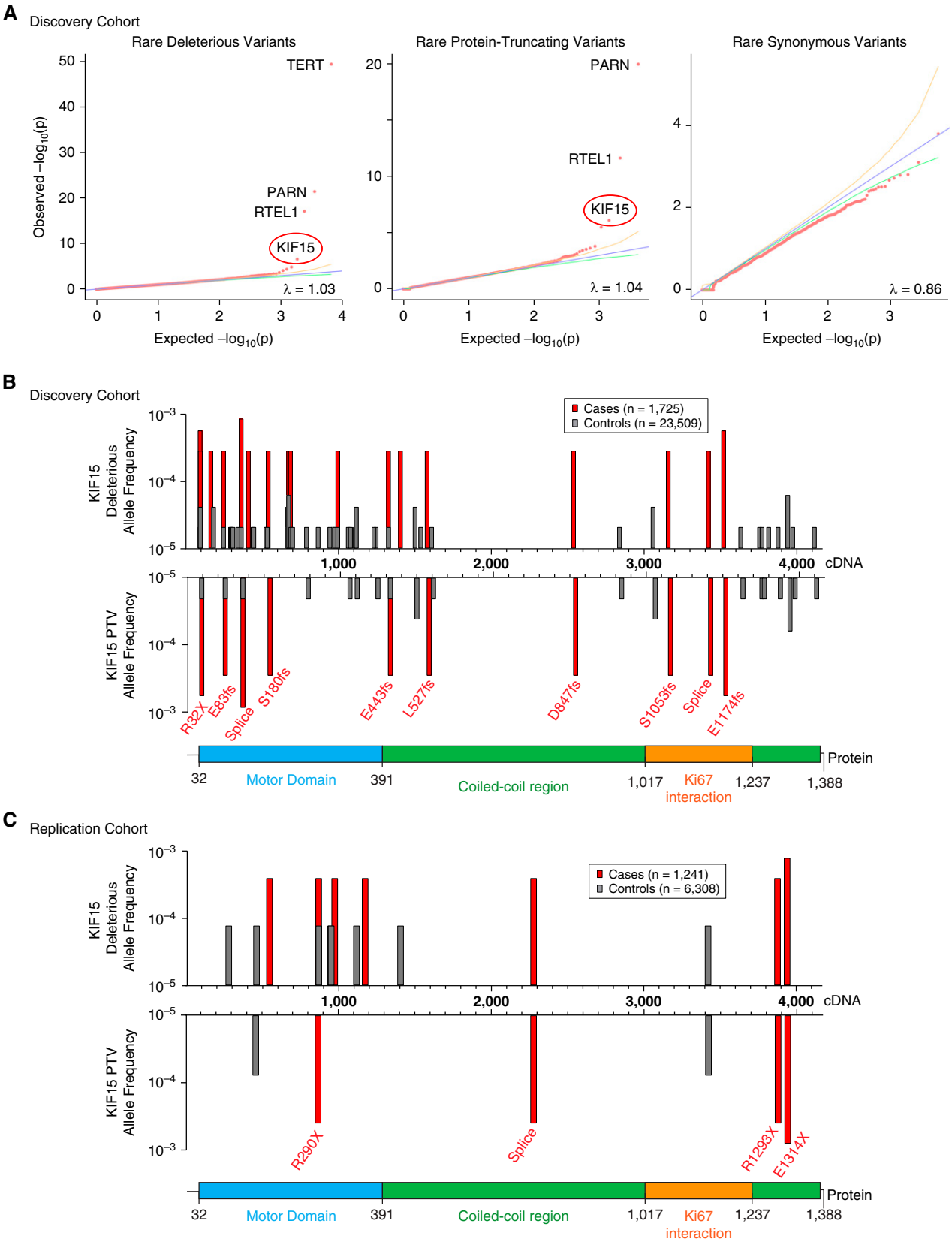

Figure 1. Deleterious *KIF15* (kinesin family member 15) variants found in patients with familial pulmonary fibrosis and those with idiopathic pulmonary fibrosis (IPF). (*A*) Quantile–quantile plot of observed versus expected *P* values comparing the burden of rare deleterious variants in protein-coding genes in familial pulmonary fibrosis and IPF cases and control subjects from the discovery cohort. Collapsing analysis of 1,725

this study. Patients collected by the University of Texas Southwestern (2003–2020), CUMC (2019–2020), Imperial College London (2009–2019), Duke University Medical Center (2000–2015), and the IPF Clinical Research Network PANTHER-IPF (Evaluating the Effectiveness of Prednisone, Azathioprine, and N-acetylcysteine in Patients With IPF; NCT 00650091) and ACE (Anticoagulant Effectiveness in Idiopathic Pulmonary Fibrosis; NCT 00957242) clinical trials (2009–2011; including only those individuals who consented to participate in both the parent study and the optional genetic substudy) constituted the discovery cohort (*see* Figure E1 in the online supplement). Written informed consent was obtained from all participants. At the time of recruitment, all subjects carried a diagnosis of IPF or FPF. Individuals were followed longitudinally at two of the sites; some diagnoses changed over time because of the subjects' clinical manifestations or because of revised guidelines (*see* Table E3). Genomic DNA was isolated from blood leukocytes. Cases ($n = 989$) underwent whole-genome sequencing (WGS) by the Institute for Genomic Medicine (IGM) at CUMC according to standard protocols using Illumina's NovaSeq 6000 platform with 150-bp paired-end reads. WES of patients with IPF from Duke (*see* Table E1) was previously performed (6), and the data were reanalyzed for this study. IGM discovery cohort control subjects underwent WGS or WES by the IGM as control subjects, healthy family members, or individuals participating in nonpulmonary studies (*see* Table E2) using Illumina's HiSeq 2000, HiSeq 2500, and NovaSeq 6000.

The replication cohorts were obtained as WGS data with permission from the Database of Genotypes and Phenotypes. The replication cases include an IPF case cohort (phs001607); the replication control subjects include participants of MESA (Multi-Ethnic Study of Atherosclerosis) (phs001416) and the Framingham Heart Study (phs000974).

Sequence Data Quality Control, Variant Calling, Ancestral Clustering, and Collapsing Analysis

Sequence data for cases and control subjects were processed using the same bioinformatics pipeline for quality control and variant calling (*see* the online supplement for details specific to the discovery and replication cohorts). We performed variant annotation by including population-based allele frequencies, variant pathogenicity predictors, and clinical annotations.

To account for population substructure, we performed clustering using principal-component analysis and community detection methods to determine ancestral clusters of cases and control subjects. Cluster-specific site-based pruning was performed for coverage harmonization between cases and control subjects to reduce the influence of coverage differences on the analysis (12). Extensive sample pruning was performed (described in the online supplement) to avoid inclusion of the same individual or related individuals more than once in any of the cohorts included in this study.

Gene-based rare-variant collapsing was performed to find genes with differences in the proportion of cases versus control subjects in the discovery cohort that carry at least one qualifying variant (QV) in the gene. We used three different autosomal-dominant models to evaluate gene burden analysis of QVs: *1*) a rare deleterious (ensemble) model, 2) a rare protein-truncating variant (PTV) model, and *3*) a rare synonymous model as a negative control. Each gene was tested for an association between QV status and the pulmonary fibrosis phenotype using the exact two-sided Cochran-Mantel-Haenszel test. Genomic inflation was assessed by means of a quantile–quantile plot for each genetic model. Meta-analysis of the collapsing analyses from the discovery and replication cohorts was performed using the Cochran-Mantel-Haenszel test across all ancestry-matched clusters in both cohorts (*see* the online supplement for details).

Single-Variant Association Study

To confirm the contributions of common variants to risk of IPF, we performed a targeted single-variant association study in the genetic region surrounding *KIF15*. We included analyses of WGS data from discovery and replication cohorts to identify significant single variants, both rare and common, associated with IPF. We focused on the 1.1 MB region of chromosome 3 surrounding *KIF15* spanning 3:44,303,435–45,394,208.

Variant selected for inclusion met the following criteria: *1*) single nonmultiallelic variants, *2*) variants called by both discovery and replication data sets, *3*) variants that passed NHLBI Trans-Omics for Precision Medicine (TOPMed) Freeze 8 hard filters (support vector machine quality filter, excess heterozygosity filter if Hardy-Weinberg disequilibrium $P < 1 \times 10^{-6}$, and Mendelian discordance filter), and *4*) $P > 1 \times 10^{-15}$ on Hardy-Weinberg equilibrium score test after correcting for population structure. A total of 7,205 single variants met inclusion criteria and were analyzed.

We used principal-component analysis to adjust for population substructure and performed an association study of IPF susceptibility for discovery and replication cohorts separately. Meta-analysis was performed by combining discovery and replication analyses using a fixed-effect, inverse variance–weighted analysis (*see* the online supplement for details).

Cell Lines

Lymphocytes were isolated from blood drawn in acid citrate dextrose tubes using lymphocyte separation medium (MP Biomedicals). Epstein-Barr virus–transformed lymphoblastoid cell lines (LCLs) were established by incubating $\sim 1 \times 10^6$ lymphocytes, either freshly isolated or previously stored in liquid nitrogen, with Epstein-Barr virus

Figure 1. (*Continued*). cases and 23,509 matched control subjects identifies study-wide significant enrichment of rare deleterious variants in *TERT*, *PARN*, *RTEL1*, and *KIF15* using a model of autosomal-dominant inheritance (left). The distribution of observed *P* values for each gene was compared with the distribution of expected *P* values (*see* Methods for more details). Analysis of models analyzing rare PTVs (center) and rare synonymous variants (right) is also shown. (*B* and *C*) Distribution and frequency of rare deleterious variants found in cases (red) and control subjects (gray) from the discovery (*B*) and replication (*C*) cohorts. The variants distribute across the full length of the cDNA and predict alterations that affect the motor domain, the coiled-coil region, and the Ki67 interaction domain of the protein. The allele frequencies are displayed on a logarithmic scale. Details regarding individual deleterious variants are shown in Table E5. cDNA = complementary DNA; *PARN* = poly(A)-specific ribonuclease; PTV = protein-truncating variant; *RTEL1* = regulator of telomere elongation helicase 1; *TERT* = telomerase reverse transcriptase.

Table 1. Increased Burden of Rare Deleterious Variants in Familial Pulmonary Fibrosis and Sporadic Idiopathic Pulmonary Fibrosis

		Discovery Cohort						Replication Cohort						Meta-Analysis	
		Cases		Control Subjects				Cases		Control Subjects					
Model	Gene	QV Carriers	No QV	QV Carriers	No QV	OR	*P* Value	QV Carriers	No QV	QV Carriers	No QV	OR	*P* Value	OR	*P* Value
Deleterious	*TERT*	64	1,661	21	23,488	48.8 (26.7–92.5)	$\mathbf{4.53 \times 10^{-50}}$	63	1,178	7	6,301	38.1 (15.8–111.9)	$\mathbf{2.73 \times 10^{-31}}$	45.4 (27.6–77.4)	$\mathbf{6.37 \times 10^{-80}}$
Deleterious	*PARN*	32	1,693	24	23,485	23.1 (11.8–46.1)	$\mathbf{4.09 \times 10^{-22}}$	18	1,223	1	6,307	77.2 (9.8–3822.4)	$\mathbf{2.18 \times 10^{-10}}$	26.4 (14.3–50.4)	$\mathbf{3.90 \times 10^{-31}}$
Deleterious	*RTEL1*	40	1,685	63	23,446	8.2 (5.1–13.1)	$\mathbf{7.89 \times 10^{-18}}$	36	1,205	17	6,291	8.6 (4.3–18.1)	$\mathbf{4.54 \times 10^{-12}}$	8.4 (5.7–12.3)	$\mathbf{1.98 \times 10^{-28}}$
Deleterious	*KIF15*	21	1,704	57	23,452	4.9 (2.7–8.8)	$\mathbf{2.55 \times 10^{-7}}$	8	1,233	8	6,300	6.8 (1.8–28.0)	1.87×10^{-3}	5.2 (3.0–8.8)	$\mathbf{1.73 \times 10^{-9}}$
PTV	*TERT*	5	1,720	4	23,505	21.0 (3.5–144.6)	2.39×10^{-4}	13	1,228	0	6,308	—	$\mathbf{9.79 \times 10^{-11}}$	62.6 (15.0–329.9)	$\mathbf{4.35 \times 10^{-13}}$
PTV	*PARN*	25	1,700	8	23,501	52.5 (19.5–157.1)	$\mathbf{1.06 \times 10^{-20}}$	13	1,228	0	6,308	—	$\mathbf{5.40 \times 10^{-8}}$	59.9 (23.4–172.8)	$\mathbf{1.02 \times 10^{-27}}$
PTV	*RTEL1*	22	1,703	24	23,485	12.0 (5.9–24.7)	$\mathbf{2.33 \times 10^{-12}}$	22	1,219	6	6,302	12.5 (4.5–42.9)	$\mathbf{6.38 \times 10^{-9}}$	12.2 (6.8–22.1)	$\mathbf{5.76 \times 10^{-20}}$
PTV	*KIF15*	14	1,711	22	23,487	7.6 (3.3–17.1)	$\mathbf{8.12 \times 10^{-7}}$	5	1,236	3	6,305	9.4 (1.4–83.3)	8.31×10^{-3}	7.9 (3.8–16.5)	$\mathbf{1.73 \times 10^{-8}}$

Definition of abbreviations: *KIF15* = kinesin family member 15; LOFTEE = loss-of-function transcript effect estimator; OR = odds ratio; *PARN* = poly(A)-specific ribonuclease; PTV = protein-truncating variant; QV = qualifying variant; REVEL = rate exome variant ensemble learner; *RTEL1* = regulator of telomere elongation helicase 1; *TERT* = telomerase reverse transcriptase.

The deleterious model includes heterozygous variants with an Exome Aggregation Consortium/Genome Aggregation Database allele frequency of less than 0.0005 predicted to be protein-truncating (stop, frameshift, splice site), insertion/deletion, or deleterious missense variants (PolyPhen damaging, REVEL score ≥ 0.5, and PrimateAI score ≥ 0.8) (*see* Methods for more details). The PTV model includes heterozygous variants with an allele frequency of less than 0.001 predicted to encode a stop, frameshift, or splice site mutation (LOFTEE high confidence). *P* values in boldface type are less than a study-wide Bonferroni multiplicity-adjusted significance value of $P < 8.94 \times 10^{-7}$, accounting for the inclusion of three different models (deleterious, PTV, and synonymous) and the examination of ~18,650 genes.

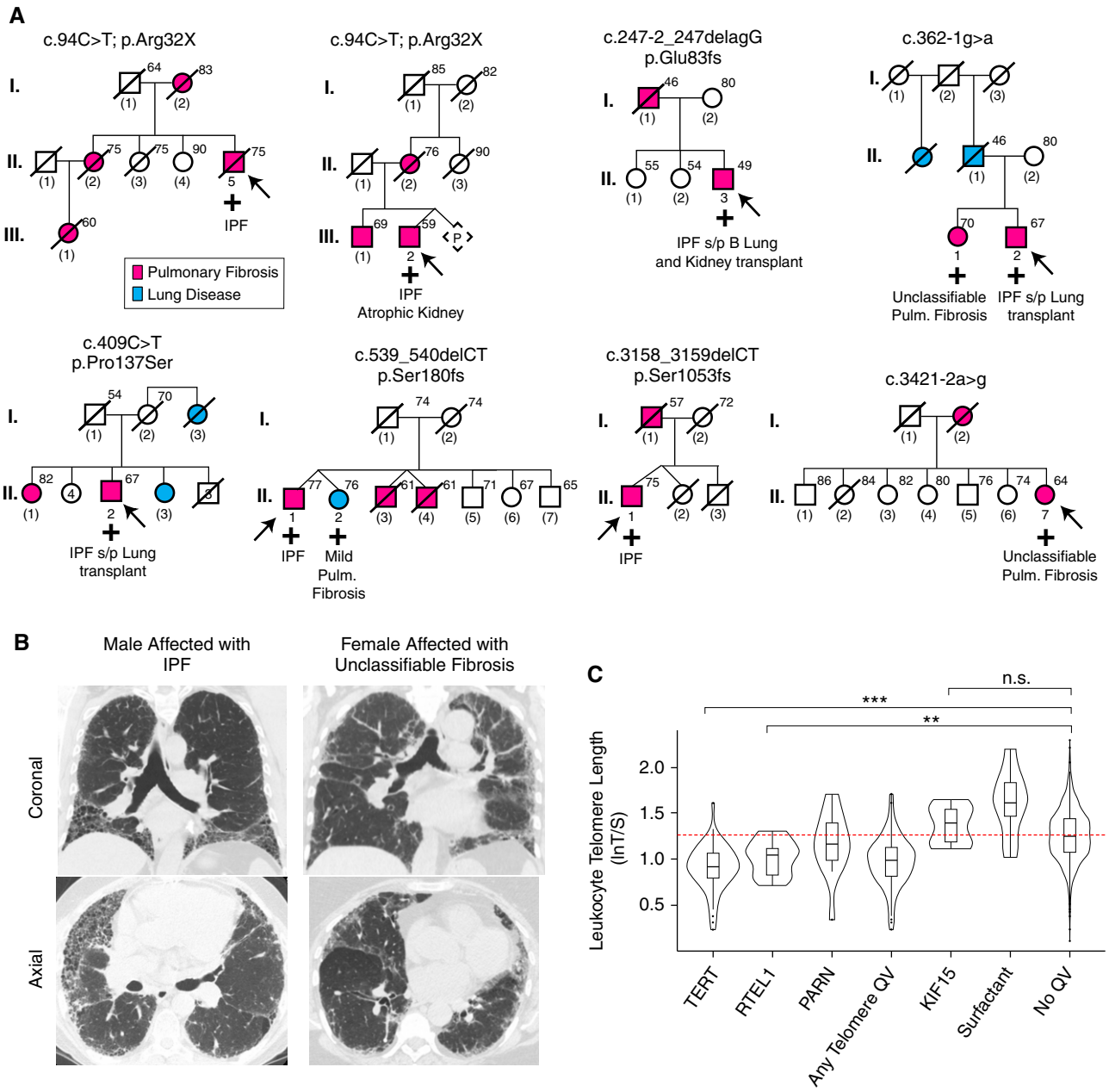

Figure 2. Clinical characterization of individuals from the discovery cohort with deleterious *KIF15* (kinesin family member 15) variants. (*A*) Abridged pedigrees of eight kindreds with familial pulmonary fibrosis and deleterious rare variants in *KIF15*. Each complementary DNA variant and predicted amino acid change is listed above the family. Individuals with pulmonary fibrosis or lung disease not otherwise specified are indicated by red and blue symbols, respectively. Open symbols represent individuals with no self-reported lung disease. Arrows indicate probands. Numbers in parentheses indicate individuals for whom no DNA samples were available. The presence of the rare variant is indicated by the plus sign. Age (in years) at the time of blood draw or death is indicated to the upper right of each symbol. Two kindreds were found to have the same loss-of-function variant (c.94C>T) predicting a substitution of arginine at amino acid 32 with a stop codon. There is no known or cryptic relatedness between these two probands. (*B*) Coronal and axial computed tomography imaging of the chest from a patient with IPF (left, proband with p.Ser180fs variant) and unclassifiable pulmonary fibrosis (right, sister of proband with c.362-1g>a variant). (*C*) Telomere length of individuals as measured by quantitative PCR assay. Individuals without a rare deleterious variant ($n = 805$) are compared with those with rare deleterious variants in *TERT* ($n = 60$), *RTEL1* ($n = 18$), *PARN* ($n = 14$), any one of six telomere-related genes (*TERT*, *RTEL1*, *PARN*, and *TINF2* [$n = 1$]; *DKC1* [$n = 2$]; and *NAF1* [$n = 4$]), *KIF15* ($n = 10$), and the surfactant genes (*SFTPC* [$n = 7$] and *SFTPA2* [$n = 1$]). Analysis was performed using the Kruskal-Wallis rank sum test; $^{**}P < 0.005$ and $^{***}P < 0.0001$. B = bilateral; *DKC1* = dyskerin pseudouridine synthase 1; IPF = idiopathic pulmonary fibrosis; *NAF1* = nuclear assembly factor 1 ribonucleoprotein; n.s. = not significant; *PARN* = poly(A)-specific ribonuclease; Pulm. = pulmonary; *RTEL1* = regulator of telomere elongation helicase 1; QV = qualifying variant; *SFTPA2* = surfactant protein A2; s/p = status post; *TERT* = telomerase reverse transcriptase; *TINF2* = TERF1 interacting nuclear factor 2.

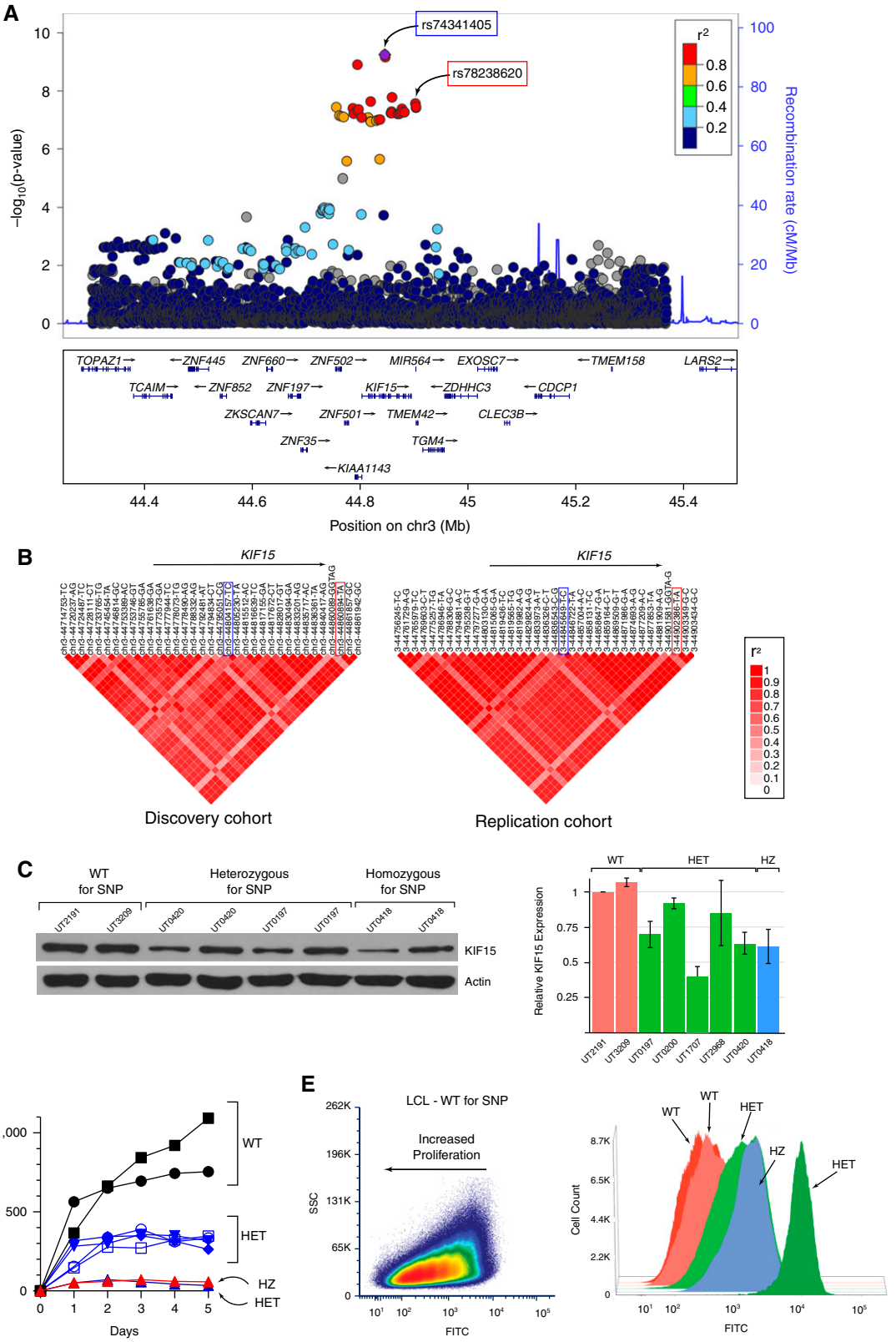

Figure 3. *KIF15* (kinesin family member 15) common variant association with IPF. (*A*) Region plot of an IPF association signal near *KIF15* from the meta-analysis of the discovery and replication cohorts. The *x*-axis shows chromosomal position and the *y*-axis the $-\log_{10}$(*P* value). (*B*) Haplotype block of the top 35 SNPs with $P_{association} < 1 \times 10^{-5}$ in the region of *KIF15* from the meta-analysis of the discovery and replication cohorts.

(American Type Culture Collection) and the mitogen phytohemagglutinin (Sigma-Aldrich). The cell lines were cultured in RPMI 1640 supplemented with 15% fetal bovine serum (FBS), 25 mM *N*-2-hydroxyethylpiperazine-*N*′-ethane sulfonic acid (HEPES), 2 mM L-glutamine, 1 mM pyruvate, penicillin (100 U/ml), and streptomycin (100 μg/ml) and expanded to a suspension culture volume of at least 50 ml before storage and use.

Hap1 cells deficient in KIF15 (HZGHC002439c010) and wild-type control cells (C631) were obtained from Horizon Discovery and cultured by serial passage according to the manufacturer's protocol in Iscove's modified Dulbecco's medium (IMDM) (Thermo Fisher Scientific) with 10% FBS, penicillin (100 U/ml), and streptomycin (100 μg/ml). Flow cytometry using propidium iodine–stained cells was used to isolate diploid cells, which were used for all subsequent experiments.

Crystal Violet Assay

Hap1 cells (3×10^5) were plated in 2 ml IMDM plus 10% FBS per well of a six-well plate on Day 0. On Day 1, cells were transfected with either mock plasmid or a plasmid expressing KIF15 (3 μg) using TurboFectin 8 (OriGene) (6 μl) in 250 μl Opti-MEM (Thermo Fisher Scientific) media. After 7 hours, either vehicle (DMSO) or 2 nM Arry-520 trifluoroacetate (catalog #4676; TOCIRS Bioscience) was added to each well, and cells were stained with crystal violet 48 hours later.

Proliferation Assays

Aliquots of LCL cells (2×10^5) in 2 ml media were plated per well of a 12-well culture plate on Day 0. At each time point, triplicate aliquots were removed, and the number of viable cells was determined using trypan blue exclusion. To determine cell proliferation by carboxyfluorescein succinimidyl ester (CFSE), 5×10^5 cells were cultured in 5 ml RPMI medium overnight. Thymidine was added to a final concentration of 2 mM, and cells were incubated for 18 hours at 37°C, then washed with Dulbecco's phosphate-buffered saline and released in fresh medium for 9 hours before incubating the cells with 2 nM thymidine for 18 hours. Cells were stained with 5 μM CFSE (CellTrace CFSE Cell Proliferation Kit, catalog #C34554; Invitrogen) in warm Dulbecco's phosphate-buffered saline for 7 minutes at 37°C, washed twice with warm medium, incubated for 1 and 72 hours in complete medium, and sorted by flow cytometry (LSRFortessa Flow Cytometer; BD Biosciences).

Aliquots of Hap1 cells (4×10^3) were plated per well of a 96-well microtiter plate in 200 μl IMDM plus 10% FBS. A transfection mix including either a mock plasmid or a plasmid expressing KIF15 (0.3 μg) and TurboFectin 8 (0.6 μl) in 9 μl Opti-MEM media was added to each well. After 7 hours, either vehicle (DMSO) or the Eg5 inhibitor Arry-520 trifluoroacetate was added to each well. Cell proliferation was measured by adding 20 μl 3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl)-2H-tetrazolium (MTS) reagent (Abcam) to each well at the indicated time point, including blank wells containing media alone. After incubation at 37°C for 2 hours, optical absorbance was determined using a plate reader (BioTek) at 490 nm. Data were plotted either as fold change in MTS activity or percentage MTS activity.

Cell Cycle Assay

Hap1 cells (3×10^5) were plated in 2 ml IMDM plus 10% FBS in six-well plates on Day 0. On Day 1, cells were incubated in medium containing vehicle (DMSO) or Arry-520 trifluoroacetate for 2 hours, followed by a 2-hour 5-ethynyl-2′-deoxyuridine (EdU) pulse labeling (Click-iT EdU flow cytometry assay kit; Thermo Fisher Scientific). Cells were labeled with DAPI, and the proliferating cells that had incorporated EdU were visualized per the manufacturer's protocol and flow sorted using a LSRFortessa Flow Cytometer. Data were analyzed using FCS 7 Express software (De Novo Software). When cells were transfected with either mock plasmid or a plasmid expressing KIF15 on Day 1, the treatment with vehicle or Arry-520 trifluoroacetate and the EdU/DAPI labeling was performed on Day 2.

Human Lung Single-Cell RNA Sequencing Analysis

Publicly available single-cell RNA sequencing data sets in Gene Expression Omnibus of human lung cells derived from fibrotic and control lung tissue were used for analysis (GSE132771 and GSE135893). For the data set derived from GSE135893, analysis was performed using cell-type designations assumed from the primary report. Unsupervised clustering and analysis of independent data from GSE132771 was performed to identify a distinct lung cell cluster with coexpression of *KIF15* with *KIF11*, *KIF14*, and cell-division genes. Specific *KIF15* expression from proliferating epithelial cells, macrophages, and lymphocytes was compared among control subjects and those with IPF, chronic hypersensitivity pneumonitis, and nonspecific interstitial pneumonitis. Cell-type proportion of proliferating cell types was compared between diseased and control subjects (*see* the online supplement).

Figure 3. (*Continued*). The sentinel SNP identified by this study (rs74341405) is indicated by the blue box; the sentinel SNP previously linked to IPF (rs78238620) is indicated by the red box. The boundary of *KIF15* is indicated by the arrow (hg38 coordinates 3:44,803,209–44,894,753 on the basis of the canonical *KIF15* transcript, ENST00000326047). (*C*) Representative immunoblot (left) and analysis (right) of KIF15 expression in lymphoblastoid cell lines (LCLs) derived from control subjects who do not have the risk *KIF15* common variant (rs78238620) (wild type [WT]), as well as those heterozygous (HET) or homozygous (HZ) for the *KIF15* common variant. (*D*) Cell proliferation of LCLs derived from control subjects who are WT, HET, and HZ for the *KIF15* rs7823860 as determined by counting cells over 5 days. (*E*) Proliferation of LCLs derived from control subjects who are WT, HET, and HZ for the *KIF15* rs7823860 as measured by *in vitro* labeling of cells using CFSE dye dilution and flow cytometry. CFSE = carboxyfluorescein succinimidyl ester; *CDCP1* = CUB domain containing protein 1; chr3 = chromosome 3; *CLEC3B* = C-type lectin domain family 3 member B; *EXOSC7* = exosome component 7; FITC = fluorescein isothiocyanate; IPF = idiopathic pulmonary fibrosis; *KIAA1143* = uncharacterized protein KIAA1143; *LARS2* = leucyl-TRNA synthetase 2, mitochondrial; *MIR564* = microRNA 564; SSC = side scatter; *TCAIM* = T cell activation inhibitor, mitochondrial; *TGM4* = transglutaminase 4; *TMEM* = transmembrane protein; *TOPAZ1* = testis and ovary specific TOPAZ 1; *ZDHHC3* = zinc finger DHHC-type palmitoyltransferase 3; *ZKSCAN7* = zinc finger with KRAB and SCAN domains 7; *ZNF* = zinc finger protein.

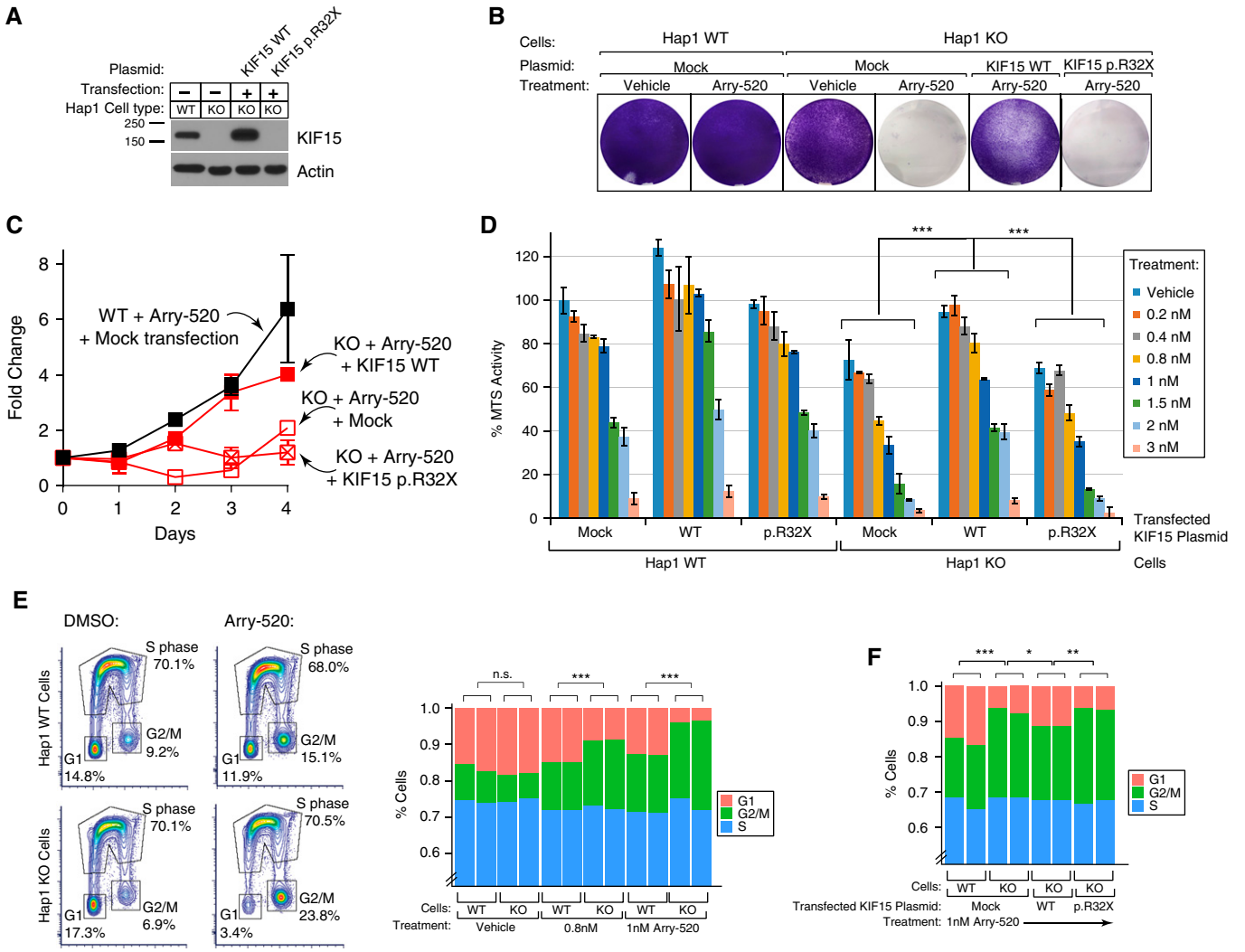

Figure 4. Cellular proliferation impaired in Hap1 cells lacking KIF15 (kinesin family member 15). (*A*) Immunoblot analysis of Hap1 wild-type (WT) and KIF15 knockout (KO) cell lines. The Hap1 KO cell line was transiently transfected with a plasmid expressing either the *KIF15* WT cDNA or the *KIF15* cDNA harboring a rare stop variant (c.94C>T; p.Arg32*) and analyzed 48 hours after transfection. Actin is shown as a loading control. (*B*) Crystal violet staining of Hap1 WT or KO cells transiently transfected with the indicated plasmids (mock, *KIF15* WT, or *KIF15* p.Arg32*) and grown in the presence of vehicle or 2 nM Arry-520. (*C*) Growth of Hap1 cells transiently transfected with the indicated plasmids and grown in the presence of 1 nM Arry-520 as determined by the 3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl)-2H-tetrazolium (MTS) cell proliferation assay. (*D*) Growth of Hap1 cells transiently transfected with the indicated plasmids (mock, *KIF15* WT, or *KIF15* p.Arg32*) in the presence of vehicle (DMSO) or Arry-520 at the indicated concentrations, as measured on Day 3 by the MTS assay. Data of triplicates are shown as mean ± SD. Comparison of all data except those at 3 nM Arry-520 was made using ordinary one-way ANOVA with multiple-comparison testing. (*E*) Cell cycle analysis of cells treated with vehicle (DMSO) or Arry-520 at the indicated concentration. Quantification of 5-ethynyl-2′-deoxyuridine (Alexa Fluor allophycocyanin fluorescence) on the *y*-axis and DNA content (DAPI) on *x*-axis, which allows resolution of G1, S, and G2/M cell cycle phases. Comparison of the percentage of G1 and G2/M cells was made by two-way ANOVA with multiple-comparison testing. (*F*) Cell cycle analysis of cells transiently transfected with the indicated plasmids (mock, *KIF15* WT, or *KIF15* p.Arg32*), treated with 1 nM Arry-520, and analyzed for cell cycle phases as described in (*E*). Comparison of the percentage of G1 and G2/M cells was made by two-way ANOVA with multiple comparison testing. $^{*}P < 0.05$, $^{**}P < 0.005$, and $^{***}P < 0.0001$. cDNA = complementary DNA; n.s. = not significant.

Results

Collapsing Analysis

WGS was performed on genomic DNA samples from patients with FPF or sporadic IPF ($n = 1{,}518$) and analyzed with WES data ($n = 251$) (*see* Table E1 and Figure E1). Given the population-based frequency of these diseases (13, 14), we expected that dominant alleles of large effect would be rare, with a population-based allele frequency of less than 0.001. To enrich for variants likely to be deleterious, we identified those predicted to be protein truncating (premature termination, splice site, and frameshift) or to cause a deleterious missense change as assessed using *in silico* tools. After sample-level pruning, ancestry

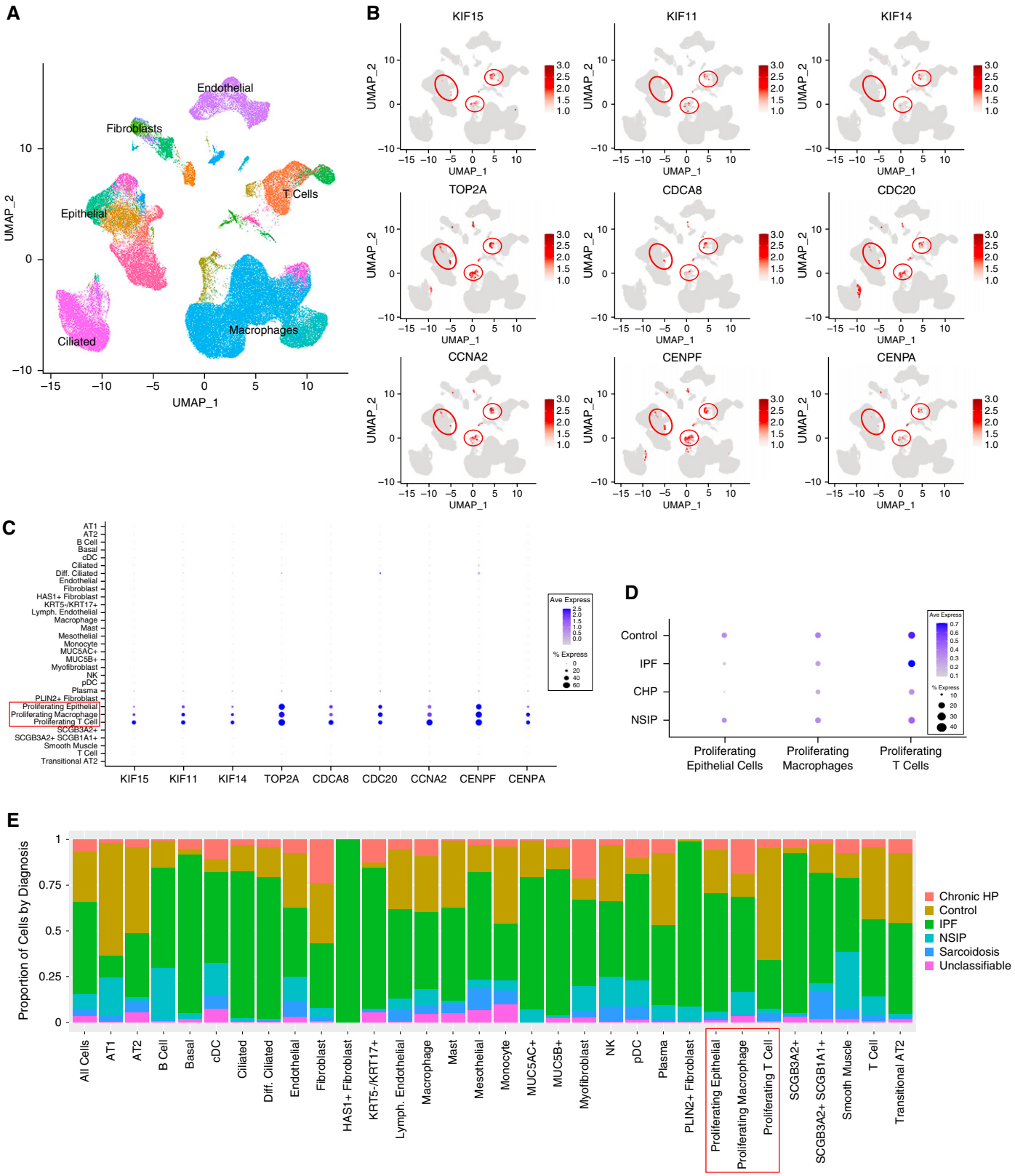

Figure 5. *KIF15* (kinesin family member 15) expression in human lung. (*A*) Human lung single-cell RNA sequencing data UMAP plot with annotated cell types (40). (*B* and *C*) *KIF15* is specifically expressed in replicating epithelial cells, T cells, and macrophages (*B*) and is coexpressed with several genes, including *KIF11* (encoding Eg5), and other cell cycle genes (*C*). (*D*) Relative expression of *KIF15* in

matching (*see* Figure E2 and Table E3), and sequence coverage harmonization, we performed collapsing analysis of cases ($n = 1{,}769$; $n = 1{,}725$ after pruning) and control subjects ($n = 26{,}714$; $n = 23{,}509$ after pruning) (*see* Table E2). As expected from prior studies (6, 7, 15, 16), we find significant enrichment of rare deleterious variants in three telomere-related genes, *TERT*, *PARN*, and *RTEL1* (Figure 1A), using an autosomal-dominant model. In addition, we make the novel observation of an excess number of rare deleterious *KIF15* variants (odds ratio [OR], 4.9; 95% confidence interval [CI], 2.7–8.8; $P = 2.55 \times 10^{-7}$) (Table 1). A separate model of rare PTVs (representing a subset of deleterious variants) demonstrates the novel observation of study-wide enrichment of *KIF15* alleles (OR, 7.6; 95% CI, 3.3–17.1; $P = 8.12 \times 10^{-7}$), as well as *PARN* and *RTEL1* alleles. There was no enrichment of genes using a rare synonymous model (Figure 1A).

After sample-level pruning and ancestry matching (*see* Figure E3), we performed collapsing analysis of WGS data from independent IPF ($n = 1{,}241$ after pruning) and control ($n = 6{,}308$ after pruning) cohorts obtained from the Database of Genotypes and Phenotypes (*see* Table E1). Rare deleterious and PTVs in *KIF15* were enriched in cases, to a lesser degree than *TERT*, *RTEL1*, and *PARN* variants (Table 1 and Figure E4). We observe a trend in enrichment of rare *KIF15* variants in the replication cohort for both the deleterious and PTV models (ORs, 6.8 [95% CI, 1.8–28.0], and 9.4 [95% CI, 1.4–83.3], respectively) in the same direction as the discovery cohort. Meta-analysis across both discovery and replication cohorts, including a total of 2,966 cases and 29,817 control subjects, implicates four genes: *KIF15* plus *TERT*, *PARN*, and *RTEL1* (Table 1). No other genes exceeded study-wide significance in the meta-analysis (*see* Table E4), although two genes, *SFTPC* (surfactant protein C) and *NAF1* (nuclear assembly factor 1 ribonucleoprotein), have been linked previously to FPF (17, 18).

In the autosomal-dominant model, qualifying rare *KIF15* PTVs and deleterious missense alleles span the length of the protein, disrupting the N-terminal motor domain and the C-terminal coiled-coil region (Figures 1B and 1C and Table E5). We find that some unrelated affected cases share the same rare loss-of-function variant (p.Arg32*, c.362-1g>a splice acceptor, p.Glu1174fs, and p.Glu1314*), even though there is no known or detectable interrelatedness between them. In the discovery and replication cohorts, there are 8.3- and 12.7-fold, respectively, more PTVs in cases than control subjects (Figure 1B and Table E7) as well as 3.1- and 3.0-fold, respectively, more deleterious missense alleles in cases than control subjects. Familial cases have a higher prevalence of rare deleterious variants (1.9%) and PTVs (1.6%) in *KIF15* than those with sporadic disease (0.90% and 0.30%, respectively) (*see* Table E8).

Most individuals with rare deleterious *KIF15* variants from the discovery cohort are male, are former smokers, and carry a diagnosis of IPF (*see* Table E9). FPF pedigrees are notable for a high frequency of twinning, with a twin birth rate of 10.9% (7 twin births per 64 births) (Figure 2A). Within these kindreds, female subjects are more likely to carry a diagnosis of unclassifiable pulmonary fibrosis and to have less severe disease (Figure 2B). Patients carrying rare deleterious alleles do not have evidence of telomere shortening (Figure 2C). One of 29 cases heterozygous for a rare deleterious *KIF15* variant was also found to carry a rare variant in *TERT* (*see* Table E5), suggesting possible oligogenic contributions.

Common Variant Association

An IPF GWAS previously reported the association of a SNP (rs78238620) residing between *KIF15* and *TMEM42* (transmembrane protein 42) (4). We verify that this SNP and 34 linked SNPs ($R^2 > 0.5$) are overrepresented in IPF cases from the discovery and replication cohorts with an OR of more than 1.4 (Figure 3A and Table E10). The SNP with the most significant association, rs74341405 (OR, 1.6; 95% CI, 1.4–1.9; $P = 5.63 \times 10^{-10}$), is located within a *KIF15* intron, has an allele frequency of 6.2–6.9% in cases and 4.0–4.1% in control subjects, and is in linkage disequilibrium with the previously reported sentinel SNP (Figures 3B and E6). The effects of the rare and common variants appear to be independent of each other (*see* Figures E7 and E8) in the discovery and replication cohorts.

KIF15 Expression and Cellular Replication

Analysis of Genotype-Tissue Expression data did not indicate a high likelihood that the SNPs linked to IPF risk affect expression of *KIF15* or the adjacent genes in the lung, although analysis of brain showed colocalization (*see* Figures E9 and E10) (4). To test for an association of the SNPs with KIF15 protein expression, we generated LCLs from control subjects who have no lung disease but who carry the *KIF15* SNP rs78238620. These LCLs show reduced amounts of KIF15 protein (Figure 3C) and reduced rates of cell proliferation (Figure 3D) compared with control subjects. Of nine attempts to derive LCLs from individuals with *KIF15* rare PTVs, none was successful.

We obtained Hap1 cells deficient in KIF15 and confirmed absence of the protein by immunoblot analysis (Figure 4A). Consistent with the partially redundant function of KIF15 and Eg5 (kinesin-5, KIF11), cell proliferation of the knockout Hap1 cells was sensitive to an Eg5 inhibitor, Arry-520, and was restored after transfection with the wild-type KIF15 complementary DNA (cDNA) but not the KIF15 cDNA encoding a premature stop codon p.Arg32* (Figures 4B–4D). Eg5 inhibitor–treated cells deficient in KIF15 exhibited a block in

Figure 5. (*Continued*). proliferating epithelial cells, proliferating macrophages, and proliferating T cells from control samples or samples from patients with IPF, CHP, or NSIP. (*E*) Relative proportion of each cell type by patient diagnosis. AT1 = type I alveolar cells; AT2 = type II alveolar cells; *CCNA2* = cyclin A2; cDC = classical dendritic cell; *CDC20* = cell division cycle 20; *CDCA8* = cell division cycle associated 8; *CENPA* = centromere protein A; *CENPF* = centromere protein F; CHP = chronic hypersensitivity pneumonitis; Diff. = differentiated; *HAS1* = hyaluronan synthase 1; HP = hypersensitivity pneumonitis; IPF = idiopathic pulmonary fibrosis; KRT = keratin; Lymph. = lymphatic; Mast = Mast cells; *MUC5AC* = mucin 5AC, oligomeric mucus/gel-forming; *MUC5B* = mucin 5B, oligomeric mucus/gel-forming; NK = natural killer; NSIP = nonspecific interstitial pneumonitis; pDC = plasmacytoid dendritic cell; *PLIN2* = perilipin 2; *SCGB* = secretoglobin family; *TOP2A* = DNA topoisomerase II α; UMAP = uniform manifold approximation and projection.

mitosis with an accumulation of cells in G2/M and a reduction of cells in G1 (Figures 4E and 4F). This mitotic block was alleviated after transfection of wild-type KIF15 cDNA but not the mutant KIF15 with the premature stop.

Analysis of human lung single-cell RNA sequencing data revealed that *KIF15* is expressed specifically in replicating epithelial cells, macrophages, and T cells (Figures 5A–5C and E11) and is coexpressed with other kinesins (*KIF11* and *KIF14*), DNA topoisomerase (*TOP2A*), cyclin A (*CCNA2*), and proteins required for mitotic spindle stability (*CDCA8* [cell division cycle associated 8]) or chromosomal segregation (*CDC20*, *CENPF* [centromere protein F] and *CENPA* [centromere protein A]). Patients with IPF and chronic hypersensitivity pneumonitis (another subtype of pulmonary fibrosis) express less *KIF15* in proliferating epithelial cells, despite abundant representation of this cell type in diseased lung tissue (Figures 5D and 5E).

Discussion

KIF15 (Hklp2, kinesin-12) is a microtubule-associated kinesin first identified by its interaction with the cell proliferation marker Ki-67 in a yeast two-hybrid screen (19). Both KIF15 and Eg5/KIF11 are capable of sliding microtubules apart, promoting spindle elongation, and ensuring segregation of chromatids during mitosis (20, 21). However, each of the chromosomal kinesins has nonredundant functions, which include distinct spatial and temporal interactions with the mitotic spindle apparatus (22) and different mechanical force–generating behaviors (23). Of the known chromokinesins, KIF15 has the longest coiled-coil domain that extends into a C-terminal leucine zipper, which is essential for dimerization (24).

Both qualifying rare and common genetic variants in *KIF15* lead to decreased protein expression and reduced cell proliferation. Our analyses indicate that *KIF15* is expressed specifically by replicating cells in the lung that coexpress other cell cycle genes. Proliferating epithelial cells isolated from patients with two subtypes of fibrotic interstitial lung disease appear to have reduced *KIF15* expression. Data from this study and others demonstrate the direct relationship between KIF15 protein expression and cell proliferation. Increased KIF15 leads to increased proliferation and, in cancer, is a poor prognostic marker (25, 26), whereas reduction of KIF15 activity leads to reduced proliferation (27–29).

Male individuals heterozygous for rare deleterious *KIF15* alleles develop IPF, whereas female individuals appear to have less severe disease and atypical radiographic features, at least in familial kindreds. The pedigrees are notable for the relatively high prevalence of fraternal twinning, which has not been observed in other FPF kindreds. Observations regarding twinning and sex-specific differences in severity are based on observations of a few kindreds and may not be generalizable to the entire cohort. Future studies will be needed to define the sex- and age-adjusted penetrance of disease, although the frequency of rare qualified variants in control populations suggests that there will be incomplete penetrance, similar to surfactant- and telomere-related FPF (15, 30).

The collapsing analysis approach used in this study has a number of limitations, including the need for large sample sizes. We note that the discovery cohort, but not the replication cohort, surpassed study-wide significance, even though the meta-analysis of both cohorts demonstrated increased ORs and more significant enrichment of KIF15 variants than the discovery cohort alone. Differences in cohort size and composition, especially regarding family history, may have led to reduced power of the replication cohort to detect the rare variant associations. There was not sufficient power to address whether *KIF15* QVs are enriched in sex- or ancestry-defined subgroups of each cohort. Another limitation includes the potential overlap of cases included in this study and one that previously identified common variants near *KIF15* (4), leading to redundant findings.

We cannot rule out the possibility that the common variants are affecting expression of another gene, although the most parsimonious explanation suggests that they affect *KIF15* expression, given the position of the top SNPs overlapping the KIF15 gene boundary, their linkage disequilibrium, the Genotype-Tissue Expression colocalization studies, the lymphoblastoid immunoblot data, and the discovery of loss-of-function rare variants that directly affect KIF15 protein concentrations. In addition, we acknowledge that despite finding *KIF15* expression restriction to replicating cells in two independent single-cell data sets, accurate quantification of its expression may be limited by cell sparsity, cell cycle phase, and heterogeneous effects from different cell types. Analyses of KIF15 function in replicating lung cells will be needed, as the human lymphoblastoid and Hap1 cells were used to interrogate specific phenotypes relevant to cell proliferation and mitosis.

Postnatal type II alveolar cells have the ability to self-renew and differentiate into epithelial type 1 cells (31), thus restoring the expansive alveolar surface area needed to support gas exchange after injury (32). Type II alveolar cells from patients with IPF demonstrate reduced capacity to proliferate (33), impaired differentiation into type I alveolar cells (34), and persistence of an intermediate transitional state expressing DNA damage response genes and markers of cellular senescence (35). A relative decrease of *KIF15* expression in proliferating lung epithelial cells in IPF and chronic hypersensitivity pneumonitis suggests a biologic mechanism for this decreased proliferative capacity. KIF15 deficiency, in the setting of impaired Eg5 function, leads to a block in mitosis through its maintenance of the bipolar spindle assembly (21, 36). The recent association between single variants in two spindle assembly genes, *MAD1L1* (mitotic arrest deficient 1 like 1) and *SPDL1* (spindle apparatus coiled-coil protein 1) (4, 37), and IPF provides collateral evidence of impaired mitosis underpinning IPF susceptibility.

Conclusions

There are a number of parallels between KIF15- and telomerase-related pulmonary fibrosis. Rare and common variants in telomerase (*TERT* and *TERC*) result in telomere shortening and have been linked to IPF susceptibility. Both rare and common variants in *KIF15* increase IPF susceptibility. Rare variants are exceedingly rare and include PTVs or missense variants predicted to impair protein function. Common variants lead to decreased gene expression. Patients

with heterozygous rare variants present after the fifth decade of life, whereas homozygous rare variants lead to pediatric presentations (38, 39). All genes function to maintain chromosomal integrity during mitosis and are necessary for robust cell division. Thus, the convergence of rare and common *KIF15* variants underscores the relevance of this nontelomere pathway of impaired cell proliferation in IPF. ■

Author disclosures are available with the text of this article at www.atsjournals.org.

Acknowledgment: The authors thank the probands and their families for their participation, Lesley Vickers for technical excellence, Deanna Rivas for help with blood sample collections, Nitin Bhardwaj and King-Tung Chan for assistance with data transfer and storage, Ming Nguyen for cloning expertise, and the staff of the IGM for generating the sequencing data. The authors thank the investigators and institutions who supported the NHLBI cohorts (IPF, MESA, and the Framingham Heart Study) and those who deposited human lung single-cell data into the Gene Expression Omnibus.

References

1. Lederer DJ, Martinez FJ. Idiopathic pulmonary fibrosis. *N Engl J Med* 2018;378:1811–1823.
2. Raghu G, Chen SY, Yeh WS, Maroni B, Li Q, Lee YC, *et al.* 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. Fingerlin TE, Murphy E, Zhang W, Peljto AL, Brown KK, Steele MP, *et al.* Genome-wide association study identifies multiple susceptibility loci for pulmonary fibrosis. *Nat Genet* 2013;45:613–620.
4. Allen RJ, Guillen-Guio B, Oldham JM, Ma SF, Dressen A, Paynton ML, *et al.* Genome-wide association study of susceptibility to idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med* 2020;201: 564–574.
5. Garcia CK, Talbert JL. Pulmonary fibrosis predisposition overview. In: Adam MP, Ardinger HH, Pagon RA, Wallace SE, editors. GeneReviews. Seattle: University of Washington; 1993–2021 [updated 2022 May 12; accessed 2022 May 12]. Available from: https://www.ncbi.nlm.nih.gov/books/NBK1230/.
6. Petrovski S, Todd JL, Durheim MT, Wang Q, Chien JW, Kelly FL, *et al.* An exome sequencing study to assess the role of rare genetic variation in pulmonary fibrosis. *Am J Respir Crit Care Med* 2017;196:82–93.
7. Dressen A, Abbas AR, Cabanski C, Reeder J, Ramalingam TR, Neighbors M, *et al.* Analysis of protein-altering variants in telomerase genes and their association with MUC5B common variant status in patients with idiopathic pulmonary fibrosis: a candidate gene sequencing study. *Lancet Respir Med* 2018;6:603–614.
8. Armanios MY, Chen JJ, Cogan JD, Alder JK, Ingersoll RG, Markin C, *et al.* Telomerase mutations in families with idiopathic pulmonary fibrosis. *N Engl J Med* 2007;356:1317–1326.
9. Tsakiri KD, Cronkhite JT, Kuan PJ, Xing C, Raghu G, Weissler JC, *et al.* Adult-onset pulmonary fibrosis caused by mutations in telomerase. *Proc Natl Acad Sci USA* 2007;104:7552–7557.
10. Duckworth A, Gibbons MA, Allen RJ, Almond H, Beaumont RN, Wood AR, *et al.* Telomere length and risk of idiopathic pulmonary fibrosis and chronic obstructive pulmonary disease: a mendelian randomisation study. *Lancet Respir Med* 2021;9:285–294.
11. Wang Q, Dhindsa RS, Carss K, Harper AR, Nag A, Tachmazidou I, *et al.*; AstraZeneca Genomics Initiative. Rare variant contribution to human disease in 281,104 UK Biobank exomes. *Nature* 2021;597:527–532.
12. Povysil G, Petrovski S, Hostyk J, Aggarwal V, Allen AS, Goldstein DB. Rare-variant collapsing analyses for complex traits: guidelines and applications. *Nat Rev Genet* 2019;20:747–759.
13. Marshall RP, Puddicombe A, Cookson WO, Laurent GJ. Adult familial cryptogenic fibrosing alveolitis in the United Kingdom. *Thorax* 2000;55: 143–146.
14. Ley B, Collard HR. Epidemiology of idiopathic pulmonary fibrosis. *Clin Epidemiol* 2013;5:483–492.
15. Stuart BD, Choi J, Zaidi S, Xing C, Holohan B, Chen R, *et al.* Exome sequencing links mutations in PARN and RTEL1 with familial pulmonary fibrosis and telomere shortening. *Nat Genet* 2015;47: 512–517.
16. Moore C, Blumhagen RZ, Yang IV, Walts A, Powers J, Walker T, *et al.* Resequencing study confirms that host defense and cell senescence gene variants contribute to the risk of idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med* 2019;200:199–208.
17. Nogee LM, Dunbar AE III, Wert SE, Askin F, Hamvas A, Whitsett JA. A mutation in the surfactant protein C gene associated with familial interstitial lung disease. *N Engl J Med* 2001;344:573–579.
18. Stanley SE, Gable DL, Wagner CL, Carlile TM, Hanumanthu VS, Podlevsky JD, *et al.* Loss-of-function mutations in the RNA biogenesis factor NAF1 predispose to pulmonary fibrosis-emphysema. *Sci Transl Med* 2016;8:351ra107.
19. Sueishi M, Takagi M, Yoneda Y. The forkhead-associated domain of Ki-67 antigen interacts with the novel kinesin-like protein Hklp2. *J Biol Chem* 2000;275:28888–28892.
20. Rath O, Kozielski F. Kinesins and cancer. *Nat Rev Cancer* 2012;12: 527–539.
21. Tanenbaum ME, Macůrek L, Janssen A, Geers EF, Alvarez-Fernández M, Medema RH. Kif15 cooperates with eg5 to promote bipolar spindle assembly. *Curr Biol* 2009;19:1703–1711.
22. Vanneste D, Takagi M, Imamoto N, Vernos I. The role of Hklp2 in the stabilization and maintenance of spindle bipolarity. *Curr Biol* 2009;19: 1712–1717.
23. Drechsler H, McAinsh AD. Kinesin-12 motors cooperate to suppress microtubule catastrophes and drive the formation of parallel microtubule bundles. *Proc Natl Acad Sci USA* 2016;113:E1635–E1644.
24. Wittmann T, Boleti H, Antony C, Karsenti E, Vernos I. Localization of the kinesin-like protein Xklp2 to spindle poles requires a leucine zipper, a microtubule-associated protein, and dynein. *J Cell Biol* 1998;143: 673–685.
25. Ding L, Li B, Yu X, Li Z, Li X, Dang S, *et al.* KIF15 facilitates gastric cancer via enhancing proliferation, inhibiting apoptosis, and predict poor prognosis. *Cancer Cell Int* 2020;20:125.
26. Wang J, Guo X, Xie C, Jiang J. KIF15 promotes pancreatic cancer proliferation via the MEK-ERK signalling pathway. *Br J Cancer* 2017; 117:245–255.
27. Qiao Y, Chen J, Ma C, Liu Y, Li P, Wang Y, *et al.* Increased KIF15 expression predicts a poor prognosis in patients with lung adenocarcinoma. *Cell Physiol Biochem* 2018;51:1–10.
28. Wang Q, Han B, Huang W, Qi C, Liu F. Identification of KIF15 as a potential therapeutic target and prognostic factor for glioma. *Oncol Rep* 2020;43:1035–1044.
29. Gao X, Zhu L, Lu X, Wang Y, Li R, Jiang G. KIF15 contributes to cell proliferation and migration in breast cancer. *Hum Cell* 2020;33: 1218–1228.
30. Diaz de Leon A, Cronkhite JT, Katzenstein AL, Godwin JD, Raghu G, Glazer CS, *et al.* Telomere lengths, pulmonary fibrosis and telomerase (TERT) mutations. *PLoS ONE* 2010;5:e10680.
31. Desai TJ, Brownfield DG, Krasnow MA. Alveolar progenitor and stem cells in lung development, renewal and cancer. *Nature* 2014; 507:190–194.
32. Kotton DN, Morrisey EE. Lung regeneration: mechanisms, applications and emerging stem cell populations. *Nat Med* 2014;20: 822–832.
33. Liang J, Zhang Y, Xie T, Liu N, Chen H, Geng Y, *et al.* Hyaluronan and TLR4 promote surfactant-protein-C-positive alveolar progenitor cell renewal and prevent severe pulmonary fibrosis in mice. *Nat Med* 2016; 22:1285–1293.
34. Jiang P, Gil de Rubio R, Hrycaj SM, Gurczynski SJ, Riemondy KA, Moore BB, *et al.* Ineffectual type 2-to-type 1 alveolar epithelial cell differentiation in idiopathic pulmonary fibrosis: persistence of the KRT8^{hi} transitional state. *Am J Respir Crit Care Med* 2020;201:1443–1447.

35. Kobayashi Y, Tata A, Konkimalla A, Katsura H, Lee RF, Ou J, *et al.* Persistence of a regeneration-associated, transitional alveolar epithelial cell state in pulmonary fibrosis. *Nat Cell Biol* 2020;22:934–946.
36. Milic B, Chakraborty A, Han K, Bassik MC, Block SM. KIF15 nanomechanics and kinesin inhibitors, with implications for cancer chemotherapeutics. *Proc Natl Acad Sci USA* 2018;115:E4613–E4622.
37. Dhindsa RS, Mattsson J, Nag A, Wang Q, Wain LV, Allen R, *et al.*; FinnGen Consortium. Identification of a missense variant in SPDL1 associated with idiopathic pulmonary fibrosis. *Commun Biol* 2021; 4:392.
38. Sleiman PMA, March M, Nguyen K, Tian L, Pellegrino R, Hou C, *et al.* Loss-of-function mutations in KIF15 underlying a Braddock-Carey genocopy. *Hum Mutat* 2017;38:507–510.
39. Savage SA, Dokal I, Armanios M, Aubert G, Cowen EW, Domingo DL, *et al.* Dyskeratosis congenita: the first NIH clinical research workshop. *Pediatr Blood Cancer* 2009;53:520–523.
40. Habermann AC, Gutierrez AJ, Bui LT, Yahn SL, Winters NI, Calvi CL, *et al.*, Single-cell RNA sequencing reveals profibrotic roles of distinct epithelial and mesenchymal lineages in pulmonary fibrosis. *Sci Adv*, 2020;6:eaba1972.