CORRESPONDENCE

Genome-wide Enrichment of *TERT* Rare Variants in Idiopathic Pulmonary Fibrosis Patients of Latino Ancestry

To the Editor:

Idiopathic pulmonary fibrosis (IPF) is a rare lung disease characterized by progressive fibrosis and irreversible decline in lung function. Genome-wide rare variant studies have either included only patients of European ancestry (1, 2) or have been disproportionally enriched with patients of European ancestry so that rare variant genetics of patients of non-European ancestry with IPF have been understudied. In this report, we evaluate the enrichment of rare genetic variants in patients of non-European ancestry with IPF.

Patients
This study was approved by the Institutional Review Board at Columbia University, and written consent was obtained from all participants. Detailed information regarding patient and control cohorts, including enrollment and sequencing sites, was previously published (3). Briefly, all case subjects carried a diagnosis of IPF or familial pulmonary fibrosis. The full cohort included multiethnic, unrelated cases (*n* = 2,966) and control subjects (*n* = 29,817) (3).

Methods
Sequence data for cases and control subjects from either Columbia or TOPMed (Trans-Omics for Precision Medicine) were processed using the program-specific bioinformatic pipelines for variant calling and variant annotation (3). Qualifying variants were defined as rare (mean allele frequency < 0.0005 in global gnomADv2 Genomes and all ExAC or gnomADv2 exome ancestry-specific subgroups) and predicted to be damaging (stop gained, start lost, frameshift, splice acceptor/donor site, or missense with consensus *in silico* predicted damaging effect by PolyPhen, REVEL [Rare Exome Variant Ensemble Learner], and PrimateAI). We used unsupervised Louvain clustering using principal components of ancestry, obtained using a set of predefined variants (4), to group cases and control subjects into 11 ethnically similar clusters for the Columbia cohorts and 7 ethnically similar clusters for the TOPMed cohorts, as previously described. Self-identified ancestry was available for the Columbia cohorts, against which predicted ancestry of unsupervised clusters was compared. *Peddy* (5) was used to infer ancestry of the TOPMed cohorts with a probability cutoff of 75% used for ancestry prediction. Individuals not meeting this cutoff for any ancestry were categorized as "Other Admixed". Individuals classified as "Latino" in the Columbia cohort were grouped with those classified as "American (AMR)" in the TOPMed cohort. Enrichment of rare

variants in each gene was tested for an association with IPF through an exact two-sided Cochran-Mantel-Haenszel test after grouping of the case and control clusters by ancestry. Genomic inflation was assessed for the non-European group using a permutation approach for estimating expected *P* values, as previously described (3).

Results
A total of 241 unrelated patients, which represents 8% of the total, were grouped into non-European cases. Slightly less than half (*n* = 120) of the non-European cases represent individuals of Latino ancestry; the remainder represents those of African, South Asian, East Asian, and Other Admixed ancestry. The remainder of the cases were grouped into European ancestry clusters (*n* = 2,725, 92% of the total cohort).

Gene burden analysis of deleterious rare (protein-truncating and missense) variants for 241 cases and 12,509 non-European control subjects showed no evidence of genomic inflation (λ = 0.996) (Figure 1A). We found an excess of *TERT* rare damaging variants that exceeded genome-wide significance (odds ratio [OR], 67.1; 95% confidence interval [CI], 23.1–195.0; $P = 9.4 \times 10^{-14}$). Analysis by ancestry demonstrated an excess of rare, damaging *TERT* variants in the Latino subgroup (OR, 80.9; 95% CI, 17.3–383.8; $P = 2.6 \times 10^{-8}$) (Figures 1B and 2). African (OR, 28.7; 95% CI, 0.6–300.6; $P = 4.3 \times 10^{-2}$) and South Asian (OR, 32.4; 95% CI, 4.5–362.0; $P = 1.7 \times 10^{-4}$) ancestry groups showed a trend toward enrichment that did not exceed study-wide significance ($P < 2.63 \times 10^{-6}$ accounting for ~19,000 genes). There were too few rare variant carriers of East Asian or Other Admixed ancestry to calculate meaningful ORs. All qualifying *TERT* rare variants discovered in non-European cases were submitted to ClinVar (SCV002520369).

Although the non-European group did not show enrichment of *PARN*, *RTEL1*, and *KIF15* rare deleterious variants, these groups all showed a trend in the same direction as the European ancestry group. No other genes exceeded study-wide significance. For *TERT* and *KIF15*, the inclusion of patients of non-European ancestry with IPF led to higher ORs and increased evidence in favor of rare deleterious variant contributions, compared with the analysis of the European ancestry group alone.

Discussion
Despite over 75% of the world's population being of non-European ancestry (6), relatively few genetic studies have focused on these populations. The lack of diversity in genetic studies is problematic because genetic variant effect sizes and polygenic risk prediction scores cannot be readily extrapolated to populations of different ancestries (7). This knowledge gap exacerbates health inequities in the modern era of personalized medical care.

We report the first genome-wide assessment of rare, deleterious genetic variants in patients of non-European ancestry with IPF. Enrichment of *TERT* rare deleterious variants was found to exceed genome-wide significance for patients of non-European ancestry, specifically, the Latino subgroup. We find that for at least two genes (*TERT* and *KIF15*), the inclusion of patients of non-European ancestry led to a greater OR in the meta-analysis. Like prior studies (8, 9), this approach demonstrated the increased power of multiethnic studies over single-ethnicity studies for improving genetic discoveries.

Here, we adjusted for population substructure in cases and ancestry-matched control subjects and used an exact test to assess the contributions of ultrarare variants in cases and control subjects.
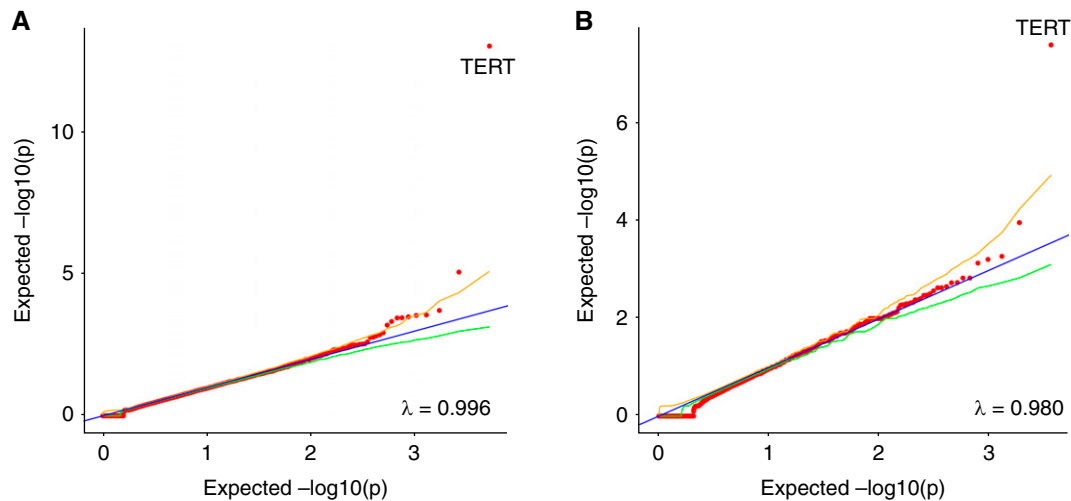
Figure 1. Analysis of non-European idiopathic pulmonary fibrosis and familial pulmonary fibrosis cases and control subjects. Quantile–quantile plot of observed versus expected *P* values comparing the burden of rare deleterious variants in protein-coding genes in cases and control subjects. (*A*) Collapsing analysis of 241 cases and 12,509 ancestry-matched non-European control subjects identify significant enrichment of rare deleterious variants in TERT using a model of autosomal dominant inheritance. (*B*) Similarly, collapsing analysis of 120 Latino cases and 4,959 ancestry-matched control subjects showed significant enrichment of *TERT* rare variants. The genetic model includes heterozygous variants (with an ExAC/gnomAD allele frequency < 0.0005 predicted to be a protein-truncating (stop, frameshift, splice site), insertion/deletion, or deleterious missense variant (PolyPhen damaging, REVEL ⩾ 0.5, and PrimateAI ⩾ 0.8).

Although rare and ultrarare variants are less influenced by population structure than common variants, decreased representation of non-European subjects in reference cohorts may lead to inflated rates of rare qualifying variants. We found no evidence of genomic inflation using our methodology. We acknowledge the uncertainty in comparing the frequency of qualifying variants for subjects of different non-European ancestries, given the smaller sample sizes and wide but overlapping CIs. However, this study justifies this approach in identifying rare variant genetic contributions for individuals of all ancestries.

To our knowledge, this is the first study that confirms the involvement of rare deleterious *TERT* variants surpassing genome-wide significance for patients of Latino and non-European ancestry with IPF. Indeed, short telomere lengths have been described for a subset of patients from Spain and Mexico with IPF, and short telomeres are associated with earlier onset and more rapidly progressive disease (10).
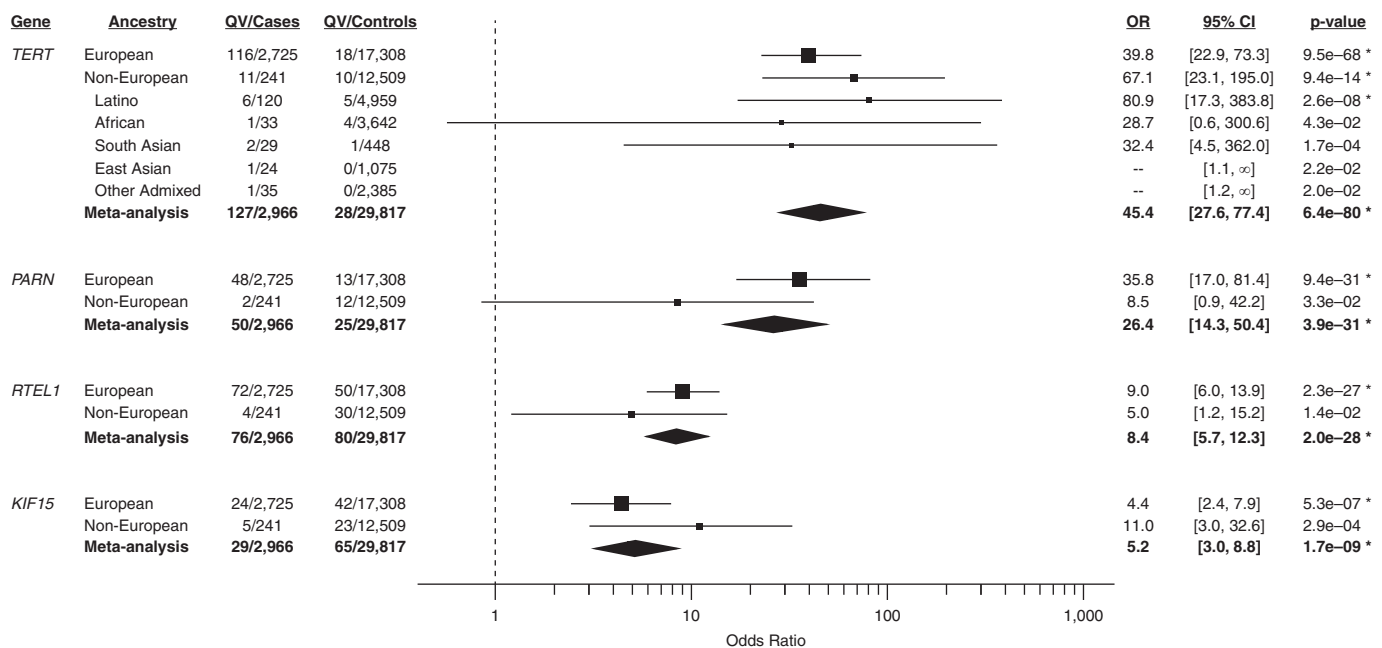

Figure 2. Forest plot of top significant genes from ancestry-stratified gene burden analysis. Odds ratio derived from Cochran-Mantel-Haenszel test of cases and control subjects in European and non-European clusters. Subsets of non-European cases and control subjects are indicated only for *TERT*. *P* values marked by the asterisk are less than a study-wide Bonferroni multiplicity-adjusted significance of $P < 2.63 \times 10^{-6}$.

To better understand the genetic underpinnings of patients of all ancestries with IPF, additional work will be needed to broaden patient recruitment to normalize imbalances. ■

David Zhang, M.D.
Gundula Povysil, M.D., Ph.D.
Columbia University Irving Medical Center
New York, New York

Chad A. Newton, M.D.
University of Texas Southwestern Medical Center
Dallas, Texas

Toby M. Maher, M.D.
University of Southern California
Los Angeles, California

and

Imperial College London
London, United Kingdom

Philip L. Molyneaux, M.D.
Imperial College London
London, United Kingdom

Imre Noth, M.D.
University of Virginia School of Medicine
Charlottesville,Virginia

Fernando J. Martinez, M.D.
Weill-Cornell Medical Center
New York, New York

Ganesh Raghu, M.D.
University of Washington Medical Center
Seattle, Washington

Jamie L. Todd, M.D.
Scott M. Palmer, M.D.
Duke University Medical Center and Duke Clinical Research Institute
Durham, North Carolina

Adam Platt, Ph.D.
AstraZeneca
Cambridge, United Kingdom

Slavé Petrovski, Ph.D.
AstraZeneca
Cambridge, United Kingdom

and

University of Melbourne
Melbourne, Victoria, Australia

David B. Goldstein, Ph.D.
Christine Kim Garcia, M.D., Ph.D.*
Columbia University Irving Medical Center
New York, New York

ORCID IDs: 0000-0002-2983-9796 (D.Z.); 0000-0003-4625-5909 (G.P.); 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.); 0000-0002-3455-1789 (A.P.); 0000-0002-0771-1249 (C.K.G.).

*Corresponding author (ckg2116@cumc.columbia.edu).

References

1. 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.
2. 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.
3. Zhang D, Povysil G, Kobeissy PH, Li Q, Wang B, Amelotte M, et al. Rare and common variants in *KIF15* contribute to genetic risk of idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med* 2022;206: 903–905.
4. Cameron-Christie S, Wolock CJ, Groopman E, Petrovski S, Kamalakaran S, Povysil G, et al. Exome-based rare-variant analyses in CKD. *J Am Soc Nephrol* 2019;30:1109–1122.
5. Pedersen BS, Quinlan AR. Who's who? Detecting and resolving sample anomalies in human DNA sequencing studies with peddy. *Am J Hum Genet* 2017;100:406–413.
6. 7 continents. World population share by continent [accessed 2022 Mar.]. Available from: https://www.worldometers.info/geography/7-continents/.
7. Martin AR, Gignoux CR, Walters RK, Wojcik GL, Neale BM, Gravel S, et al. Human demographic history impacts genetic risk prediction across diverse populations. *Am J Hum Genet* 2017;100:635–649.
8. Mensah-Ablorh A, Lindstrom S, Haiman CA, Henderson BE, Marchand LL, Lee S, et al. Meta-analysis of rare variant association tests in multiethnic populations. *Genet Epidemiol* 2016;40:57–65.
9. Wojcik GL, Graff M, Nishimura KK, Tao R, Haessler J, Gignoux CR, et al. Genetic analyses of diverse populations improves discovery for complex traits. *Nature* 2019;570:514–518.
10. Planas-Cerezales L, Arias-Salgado EG, Buendia-Roldán I, Montes-Worboys A, López CE, Vicens-Zygmunt V, et al. Predictive factors and prognostic effect of telomere shortening in pulmonary fibrosis. *Respirology* 2019;24:146–153.

Check for updates

Polarization-Sensitive Endobronchial Optical Coherence Tomography for Microscopic Imaging of Fibrosis in Interstitial Lung Disease

To the Editor:

Fibrotic interstitial lung diseases (fILDs), including idiopathic pulmonary fibrosis (IPF), are characterized by excessive collagen deposition and fibrotic remodeling. Early, precise diagnosis and monitoring of disease progression are essential to strategize