The *MUC5B* Promoter Polymorphism Is Associated With Idiopathic Pulmonary Fibrosis in a Mexican Cohort but Is Rare Among Asian Ancestries

Anna L. Peljto, DrPH; Moises Selman, MD, FCCP; Dong Soon Kim, MD; Elissa Murphy, MS; Laura Tucker, BS; Annie Pardo, PhD; Jung Su Lee, MD; Wonjun Ji, MD; Marvin I. Schwarz, MD, FCCP; Ivana V. Yang, PhD; David A. Schwartz, MD; and Tasha E. Fingerlin, PhD

BACKGROUND: Polymorphisms in the *MUC5B* promoter, *TOLLIP*, and nine additional genetic loci have been associated with idiopathic pulmonary fibrosis (IPF) within non-Hispanic white populations. It is unknown whether these variants account for risk of IPF in other racial/ethnic populations. We conducted a candidate single nucleotide polymorphism (SNP) association study in cohorts of Mexican and Korean patients with IPF.

METHODS: We chose 12 SNPs from 11 loci that are associated with IPF among non-Hispanic whites and genotyped these SNPs in cohorts of Mexican (83 patients, 111 control subjects) and Korean (239 patients, 87 control subjects) people. Each SNP was tested for association with IPF, after adjusting for age and sex.

RESULTS: The *MUC5B* promoter SNP rs35705950 was associated with IPF in the Mexican ($OR = 7.36$, $P = .0001$), but not the Korean ($P = .99$) cohort. The SNP in *IVD* (chromosome15, rs2034650) was significantly associated with pulmonary fibrosis in both the Mexican ($OR = 0.40$, $P = .01$) and Korean ($OR = 0.13$, $P = .0008$) cohorts. In the Korean cohort, there were no other variants associated with disease. In the Mexican cohort, SNPs on chromosomes 3, 4, and 11 were also associated with disease.

CONCLUSIONS: The strongest identified genetic risk factor for IPF among the non-Hispanic white population, the *MUC5B* promoter polymorphism, is also a strong risk factor in a Mexican population, but is very rare in a Korean population. The majority of genetic variants that account for risk of IPF in groups other than non-Hispanic whites are unknown. Hispanic and Asian populations should be studied separately to identify genetic risk loci for IPF.

CHEST 2015; 147(2):460-464

Manuscript received April 8, 2014; revision accepted September 3, 2014; originally published Online First October 2, 2014.

ABBREVIATIONS: IPF = idiopathic pulmonary fibrosis; MAF = minor allele frequency; SNP = single nucleotide polymorphism

AFFILIATIONS: From the Department of Medicine (Drs Peljto, M. I. Schwarz, Yang, and D. A. Schwartz and Ms Murphy), Department of Immunology (Dr D. A. Schwartz), and Department of Epidemiology (Dr Fingerlin), University of Colorado Denver, Denver, CO; Instituto Nacional de Enfermedades Respiratorias (Dr Selman), Mexico City, Mexico; Asan Medical Center (Drs Kim, Lee, and Ji), University of Ulsan College of Medicine, Seoul, South Korea; Yale University (Ms Tucker), New Haven, CT; Universidad Nacional Autonoma de Mexico (Dr Pardo), Mexico City, Mexico; and Department of Medicine (Drs M. I. Schwarz and D. A. Schwartz), National Jewish Health, Denver, CO.

Drs D. A. Schwartz and Fingerlin contributed equally to the manuscript.

FUNDING/SUPPORT: This research was supported by the National Heart, Lung, and Blood Institute [R01-HL095393, R01-HL097163, P01-HL092870, and RC2-HL101715] and the Veterans Administration [1101BX001534].

CORRESPONDENCE TO: Anna L. Peljto, DrPH, University of Colorado, School of Medicine, 12631 E 17th Ave, MS B178, Aurora, CO 80045; e-mail: anna.peljto@ucdenver.edu

© 2015 AMERICAN COLLEGE OF CHEST PHYSICIANS. Reproduction of this article is prohibited without written permission from the American College of Chest Physicians. See online for more details.

DOI: 10.1378/chest.14-0867

In recent years, a number of studies have identified genetic risk factors for pulmonary fibrosis. In the largest, to our knowledge, genome-wide association study to date of fibrotic idiopathic interstitial pneumonias, we identified 10 genetic loci that were significantly associated with disease.¹ In another study, Noth et al² identified two additional single nucleotide polymorphisms (SNPs) that were associated with idiopathic pulmonary fibrosis (IPF). Among these recently reported associations, three were confirmations of previously identified risk loci for pulmonary fibrosis: *MUC5B* at 11p15, *TERT* at 5p15, and *TERC* at 3q26.³⁻⁵ Of note, a promoter polymorphism in *MUC5B* has been confirmed by multiple studies as a common risk factor with the largest genetic effect on development of pulmonary fibrosis, estimated to increase risk by six-fold for people who are heterozygous for the polymorphism and by 20-fold for those who are homozygous.^{1-3,6,7} The identification of these loci, and of the *MUC5B* polymorphism, in particular, have changed perspectives on the pathogenesis of pulmonary fibrosis, motivated additional research on the risk factors for pulmonary fibrosis, and generated new targets for pharmacologic therapies.^{3,8-10}

While research on the genetics of pulmonary fibrosis is making great progress, these studies share an important limitation: All were conducted in populations of non-Hispanic white subjects only. Therefore, it is unknown whether the variants identified in these

studies account for the risk of pulmonary fibrosis in other racial/ethnic populations. Distinct genetic risk factors would warrant separate genetic and epidemiologic studies for different populations, which may have implications for different treatment strategies. Conversely, shared genetic risk profiles could potentially allow the conclusions from previously reported studies to be applied to other racial/ethnic populations for the purposes of genetic counseling, prognosis, and patient care.

Understanding the contributions of genetic factors to risk of IPF in different populations is especially pertinent given advancements of our understanding of the role of the *MUC5B* polymorphism. Subsequent studies of the polymorphism have shown it also has prognostic value among patients with IPF⁹ and is associated with more mild interstitial lung abnormalities in the general population.¹⁰ These findings have important implications for potentially informing clinical trials and identifying IPF in the earlier stages of disease. Identification of shared and distinct genetic risk factors across racial and ethnic cohorts will determine whether these findings may be applied to other nonwhite populations. We conducted a candidate SNP association study in Mexican and Korean cohorts to explore the genetic risk variants for pulmonary fibrosis that may be similar to or distinct from those identified in non-Hispanic white populations.

Materials and Methods

Candidate SNPs

We chose 12 SNPs from 11 loci previously reported to be common in the non-Hispanic white population and associated with pulmonary fibrosis.¹⁻³ The *MUC5B* promoter polymorphism was originally reported to be associated with IPF by Seibold et al,³ nine loci outside of the *MUC5B* locus were reported by Fingerlin et al,¹ and two SNPs within *TOLLIP* were reported by Noth et al.² We chose not to genotype a third reported SNP within *TOLLIP*, since it was found to be in high linkage disequilibrium ($r^2 = 0.97$) with one of the other genotyped *TOLLIP* SNPs (rs111521887).²

Study Cohorts

We genotyped these 12 candidate SNPs in cohorts of Mexican (83 patients with IPF, 111 control subjects) and Korean (239 patients with IPF, 87 control subjects) people. Mexican subjects with IPF were all recruited from Instituto Nacional de Enfermedades Respiratorias in Mexico City, Mexico. Control subjects with at least 5 years of smoking history (former and current) and normal pulmonary function tests were

recruited from a smoking cessation program at the same institute, and nonsmoking control subjects were unrelated healthy volunteers. Korean subjects were recruited from the pulmonary department at Asan Medical Center in Seoul, South Korea. In both cohorts, IPF was diagnosed according to the American Thoracic Society/European Respiratory Society International Consensus Guidelines.¹¹ The study was approved by the Science and Bioethical Committee at the National Institute of Respiratory Diseases and the National Jewish Health Institutional Review Board (1441A). All subjects gave written informed consent.

Statistical Analysis

Logistic regression was used to test each SNP for association with IPF compared with control subjects of the same race/ethnicity. All SNPs were modeled additively. Each model was adjusted for sex and age at the time of diagnosis. We compared the minor allele frequencies (MAFs) and ORs for developing pulmonary fibrosis in each cohort to those previously reported for non-Hispanic white cohorts. In secondary analyses, tests of association were also adjusted for the effect of the *MUC5B* promoter polymorphism (rs35705950).

Results

The Mexican and Korean patients were similar in age, proportion of women, and diffusion capacity of carbon monoxide; however, Korean patients were more likely to have a history of smoking ($P = .004$) and better FVC at

the time of diagnosis ($P < .0001$) compared with the Mexican patients (Table 1).

The *MUC5B* promoter polymorphism was the most strongly associated variant in the Mexican cohort (rs35705950 in *MUC5B*; OR = 7.36, $P = .0001$), as has

TABLE 1 Characteristics of Study Patients With Idiopathic Pulmonary Fibrosis

Characteristic	Mexico (No. = 83)	Korea (No. = 239)
Female, No. (%)	24 (29)	60 (25)
Age at diagnosis, mean (SD), y	66.0 (7.7)	65.1 (7.7)
Smoking history (ever), No. (%)	41 (49)	161 (67)
FVC % predicted at diagnosis, mean (SD)	59.3 (15.2)	69.0 (19.9)
DLco % predicted at diagnosis, mean (SD)	51.0 (15.3)	55.5 (21.5)

DLco = diffusing capacity of the lung for carbon monoxide.

been reported for non-Hispanic white cohorts. Interestingly, all subjects who carried the *MUC5B* minor allele in the Korean cohort also had IPF; however, the allele was very rare (five of 650 alleles), and the association was not significant after adjusting for age and sex ($P = .99$).

In general, the identified genetic risk profile for pulmonary fibrosis in non-Hispanic whites was somewhat different from that in the Mexican cohort and almost completely distinct from that in the Korean cohort. The SNP in *IVD* (chromosome15, rs2034650) was significantly associated with pulmonary fibrosis in both the Mexican ($OR = 0.40$, $P = .01$) and Korean ($OR = 0.13$, $P = .0008$) cohorts (Table 2). In the Korean cohort, none

of the other variants was associated with disease. In the Mexican cohort, SNPs on chromosomes 3 (rs6793295 in *LRRC34*; $OR = 2.24$, $P = .01$) and 4 (rs2609255 in *FAM13A*; $OR = 2.09$, $P = .02$) were also associated with disease. Another SNP on chromosome 11 (rs111521887) in *TOLLIP* was also associated with disease in the Mexican cohort; however, this association was no longer significant after adjusting for the effect of the *MUC5B* SNP ($OR = 1.10$, $P = .88$) (Table 3). The SNP on chromosome 5 (rs2736100) in *TERT* appeared moderately significant in the Mexican cohort ($OR = 0.40$, $P = .03$) after adjustment for the *MUC5B* SNP. None of the conclusions for the other SNPs changed after adjustment for *MUC5B*. The ORs for all SNPs that were significantly associated with pulmonary fibrosis in the Mexican and Korean cohorts indicated that direction of the effect of the risk allele was the same as that identified in the non-Hispanic whites.

All the SNPs included in this study were common among the non-Hispanic whites ($MAF \geq 0.05$). While allele frequencies varied among the cases of the different populations, most SNPs were also common in the Mexican and Korean cohorts. The exceptions were the *MUC5B* SNP (rs35705950), which had an MAF of 1% in the Korean cohort, and the *TOLLIP* SNP (rs111521887), which was completely absent within the Korean cohort.

TABLE 2 Association of Candidate SNPs With Pulmonary Fibrosis in Mexican and Korean Cohorts, Adjusted for Age and Sex, Compared With Reported Associations in Non-Hispanic White Cohorts

Chr	SNP	Gene	Non-Hispanic White ^a				Mexican			Korean		
			Minor Allele ^b	AF Case	OR	P Value	AF Case	OR ^c	P Value	AF Case	OR	P Value
3	rs6793295	<i>LRRC34</i>	C	0.32	1.30	3.2×10^{-7}	0.74	2.24	.01	0.82	1.62	.36
4	rs2609255	<i>FAM13A</i>	G	0.26	1.29	5.3×10^{-6}	0.41	2.09	.02	0.60	2.10	.10
5	rs2736100	<i>TERT</i>	C	0.43	0.73	7.6×10^{-14}	0.22	0.50	.05	0.30	0.57	.17
6	rs2076295	<i>DSP</i>	G	0.54	1.43	1.1×10^{-16}	0.47	0.99	.98	0.64	1.44	.34
7	rs4727443	Intergenic	A	0.46	1.30	6.7×10^{-9}	0.51	0.88	.67	0.64	2.20	.07
10	rs11191865	<i>OBFC1</i>	G	0.45	0.80	2.8×10^{-7}	0.62	1.79	.07	0.71	1.32	.53
11	rs35705950	<i>MUC5B</i>	T	0.31	4.51	7.2×10^{-95}	0.25	7.36	1.1×10^{-4}	0.01	∞	.99
11	rs5743890	<i>TOLLIP</i>	G	0.11	0.68	2.5×10^{-3}	0.05	1.15	.83	0.06	0.18	.17
11	rs111521887	<i>TOLLIP</i>	G	0.29	1.54	5.0×10^{-7}	0.20	3.79	.01	0.00	...	...
13	rs1278769	<i>ATP11A</i>	A	0.20	0.79	9.1×10^{-7}	0.25	1.36	.41	0.33	0.48	.10
15	rs2034650	<i>IVD</i>	G	0.42	0.77	1.9×10^{-9}	0.31	0.40	.01	0.10	0.13	8.1×10^{-4}
19	rs1260495	<i>DPP9</i>	G	0.34	1.29	9.6×10^{-9}	0.28	0.85	.65	0.16	0.91	.87

∞ = infinity; AF Case = allele frequency in the case group; Chr = chromosome; SNP = single nucleotide polymorphism.

^aAssociations reported in Fingerlin et al,¹ except *TOLLIP* SNPs, reported in Noth et al.²

^bThe minor allele is defined as the minor allele in the combined case and control group of the non-Hispanic white cohort.

^cOR for the minor allele.

TABLE 3] Association of Candidate SNPs With Pulmonary Fibrosis in Mexican and Korean Cohorts, Adjusted for Age, Sex, and *MUC5B* SNP (rs35705950)

Chr	SNP	Mexican		Korean	
		OR	P Value	OR	P Value
3	rs6793295	2.22	.02	1.58	.39
4	rs2609255	2.79	7.3×10^{-3}	1.90	.16
5	rs2736100	0.40	.03	0.58	.18
6	rs2076295	0.76	.41	1.37	.41
7	rs4727443	1.00	.99	2.35	.06
10	rs11191865	1.92	.06	1.34	.50
11	rs5743890	1.48	.56	0.18	.18
11	rs111521887	1.10	.88	NA	NA
13	rs1278769	1.07	.88	0.44	.08
15	rs2034650	0.28	2.8×10^{-3}	0.14	1.4×10^{-3}
19	rs1260495	1.05	.89	0.91	.88

NA = not applicable. See Table 2 legend for expansion of other abbreviations.

Discussion

We found that the strongest genetic risk factor for pulmonary fibrosis, rs35705950 in the promoter of *MUC5B*, identified in non-Hispanic white populations, is also a risk factor in Hispanic populations, but is rare among Asian cases of IPF. This suggests that findings related to IPF and *MUC5B* or future treatments that target *MUC5B* may apply to Hispanic patients, in addition to non-Hispanic white patients. In the Korean cohort, the *MUC5B* promoter polymorphism was extremely rare among patients with IPF (MAF < 0.01) and completely absent among the control subjects. This observation is similar to what has been reported by the 1000 Genomes project for other Asian populations (MAF ≤ 0.01).¹² Since the strongest identified genetic risk factor for IPF is not common in Asian populations, there are certainly other important unidentified genetic and/or environmental factors influencing IPF risk in these populations. Whether expression of *MUC5B* is increased in IPF lungs among Asian patients over and above the effect of the promoter variant, as has been observed in non-Hispanic white patients, is an important question that will also inform the potential for *MUC5B*-targeted treatments for Asian patients with IPF.

Including the *MUC5B* variant, the Mexican cohort shared about one-third of the genetic risk factors that we investigated with non-Hispanic whites, while the Korean cohort only shared one risk variant with the non-Hispanic white and Mexican populations. These results are not surprising given that the population divergence due to genetic structure, as measured by the

fixation index, has been shown to be much greater between Asian and European populations (fixation index = 0.11) than between Mexican and European populations (fixation index = 0.03).¹³

The only association that persisted across all three cohorts was for the SNP on chromosome 15, which is in *IVD*. *IVD* is a mitochondrial matrix enzyme involved in leucine catabolism. Within the Mexican cohort, the SNP in *FAM13A* was associated with IPF. *FAM13A* is a signal transduction gene that has also been associated with COPD.¹⁴ The other associated SNP is within *LRCC34*, which is of unknown function, but just downstream of *TERC*. Variants in *TERC* have previously been reported to be associated with IPF.^{4,5}

It is likely that different genetic risk profiles could account for differences in risk between diverse racial/ethnic populations. A previous study has examined the prevalence of IPF in diverse populations, reporting that Hispanics have a higher prevalence of IPF than non-Hispanic whites, while African Americans have a lower prevalence.¹⁵ Interestingly, the *MUC5B* promoter polymorphism is observed less frequently in African and Asian populations than Hispanic and non-Hispanic white populations.¹² Additional studies of Hispanic cohorts will be needed to identify the genetic components that may account for the increased risk. Genetic differences may also explain variance in the severity or progression of disease across different populations. For example, the frequency of *MUC5B* promoter polymorphism differs dramatically across racial/ethnic populations and is also associated with improved survival.⁹ This is consistent

with reports that Hispanic patients with IPF are more likely to die of IPF than non-Hispanic whites.¹⁶ Another study also reported higher mortality rates among non-Hispanic black and Hispanic subjects with IPF who were listed on lung transplant waiting lists.¹⁶ Understanding whether genetic differences are driving these observed prevalence and disease-severity differences will have crucial implications for providing genetic counseling and identifying effective treatments for nonwhite individuals with IPF. This is especially important since IPF is a fatal disease with no therapies that improve symptoms or outcome.

The limitations of our study include samples sizes that have limited power to identify significant effects of the variants other than the *MUC5B* SNP, which have smaller effects on disease risk. Second, while the SNPs included in this study were associated with disease at a genome-wide significance level among non-Hispanic whites, none (with the exception of the *MUC5B* pro-

motor polymorphism) have been definitively identified as causal variants having functional effects. It is possible that different variants within some of the identified genes have significant effects in different racial/ethnic populations. Comprehensive analysis of variants within *MUC5B* is certainly warranted among Asian populations, since the rs35705950 SNP appears to be associated with IPF, but too rare to contribute significantly to the risk of disease in the population. Other variants in the promoter of *MUC5B* have been reported to be associated with panbronchiolitis in a Japanese population, providing support for this further investigation of *MUC5B* among Asian populations.¹⁷ Future studies will determine exactly which variants are driving the observed associations and their potential role in the mechanism of disease. These results suggest that these populations and other racial/ethnic populations will need to be studied individually to verify these findings and identify the full range of genetic components that account for IPF risk.

Acknowledgments

Author contributions: A. L. P. had full access to the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. A. L. P., M. I. S., I. V. Y., D. A. S., and T. E. F. contributed to the study design; M. S., D. S. K., A. P. J. S. L., and W. J. recruited patients and conducted clinical evaluations; E. M. and L. T. performed the laboratory work; A. L. P., L. T., and T. E. F. analyzed data; M. I. S., I. V. Y., D. A. S., and T. E. F. contributed to interpretation of results; A. L. P., M. S., M. I. S., I. V. Y., D. A. S., and T. E. F. wrote the manuscript; and A. L. P., M. S., D. S. K., E. M., L. T., A. P. J. S. L., W. J., M. I. S., I. V. Y., D. A. S., and T. E. F. contributed to the editing of the manuscript.

Financial/nonfinancial disclosures: The authors have reported to *CHEST* the following conflicts of interest: Dr Kim is a member of the steering committee and advisory board for Nintedanib (Boehringer Ingelheim GmbH). Dr Yang has been supported by one or more grants from the National Institutes of Health (NIH) and has one or more patents. Dr D. A. Schwartz has research grants currently funded by the NIH and Veterans Administration, holds numerous patents related to pulmonary fibrosis research, and has provided expert testimony on mesothelioma cases.

Role of sponsors: The sponsors had no role in the design of the study, the collection and analysis of the data, or the preparation of the manuscript.

References

1. Fingerlin TE, Murphy E, Zhang W, et al. Genome-wide association study identifies multiple susceptibility loci for pulmonary fibrosis. *Nat Genet.* 2013;45(6):613-620.
2. Noth I, Zhang Y, Ma SF, et al. Genetic variants associated with idiopathic pulmonary fibrosis susceptibility and mortality: a genome-wide association study. *Lancet Respir Med.* 2013;1(4):309-317.
3. Seibold MA, Wise AL, Speer MC, et al. A common *MUC5B* promoter polymorphism and pulmonary fibrosis. *N Engl J Med.* 2011;364(16):1503-1512.
4. Armanios MY, Chen JJ, Cogan JD, et al. Telomerase mutations in families with idiopathic pulmonary fibrosis. *N Engl J Med.* 2007;356(13):1317-1326.
5. Tsakiri KD, Cronkwhite JT, Kuan PJ, et al. Adult-onset pulmonary fibrosis caused by mutations in telomerase. *Proc Natl Acad Sci U S A.* 2007;104(18):7552-7557.
6. Zhang Y, Noth I, Garcia JG, Kaminski N. A variant in the promoter of *MUC5B* and idiopathic pulmonary fibrosis. *N Engl J Med.* 2011;364(16):1576-1577.
7. Stock CJ, Sato H, Fonseca C, et al. Mucin 5B promoter polymorphism is associated with idiopathic pulmonary fibrosis but not with development of lung fibrosis in systemic sclerosis or sarcoidosis. *Thorax.* 2013;68(5):436-441.
8. Boucher RC. Idiopathic pulmonary fibrosis—a sticky business. *N Engl J Med.* 2011;364(16):1560-1561.
9. Peljto AL, Zhang Y, Fingerlin TE, et al. Association between the *MUC5B* promoter polymorphism and survival in patients with idiopathic pulmonary fibrosis. *JAMA.* 2013;309(21):2232-2239.
10. Hunninghake GM, Hatabu H, Okajima Y, et al. *MUC5B* promoter polymorphism and interstitial lung abnormalities. *N Engl J Med.* 2013;368(23):2192-2200.
11. Raghu G, Collard HR, Egan JJ, et al; ATS/ERS/JRS/ALAT Committee on Idiopathic Pulmonary Fibrosis. 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(6):788-824.
12. Abecasis GR, Auton A, Brooks LD, et al; 1000 Genomes Project Consortium. An integrated map of genetic variation from 1,092 human genomes. *Nature.* 2012;491(7422):56-65.
13. Altshuler DM, Gibbs RA, Peltonen L, et al; International HapMap 3 Consortium. Integrating common and rare genetic variation in diverse human populations. *Nature.* 2010;467(7311):52-58.
14. Cho MH, Boutaoui N, Klanderman BJ, et al. Variants in *FAM13A* are associated with chronic obstructive pulmonary disease. *Nat Genet.* 2010;42(3):200-202.
15. Swigris JJ, Olson AL, Huie TJ, et al. Ethnic and racial differences in the presence of idiopathic pulmonary fibrosis at death. *Respir Med.* 2012;106(4):588-593.
16. Lederer DJ, Arcasoy SM, Barr RG, et al. Racial and ethnic disparities in idiopathic pulmonary fibrosis: A UNOS/OPTN database analysis. *Am J Transplant.* 2006;6(10):2436-2442.
17. Kamio K, Matsushita I, Hijikata M, et al. Promoter analysis and aberrant expression of the *MUC5B* gene in diffuse pan-bronchiolitis. *Am J Respir Crit Care Med.* 2005;171(9):949-957.