Native American Ancestry, Lung Function, and COPD in Costa Ricans

Wei Chen, PhD; John M. Brehm, MD; Nadia Boutaoui, PhD; Manuel Soto-Quiros, MD; Lydiana Avila, MD; Bartolome R. Celli, MD, FCCP; Shannon Bruse, PhD; Yohannes Tesfaigzi, PhD; and Juan C. Celedón, MD, DrPH, FCCP

Background: Whether Native American ancestry (NAA) is associated with COPD or lung function in a racially admixed Hispanic population is unknown.

Methods: We recruited 578 Costa Ricans with and without COPD into a hybrid case-control/family-based cohort, including 316 members of families of index case subjects. All participants completed questionnaires and spirometry and gave a blood sample for DNA extraction. Genome-wide genotyping was conducted with the Illumina Human610-Quad and HumanOmniExpress BeadChip kits (Illumina Inc), and individual ancestral proportions were estimated from these genotypic data and reference panels. For unrelated individuals, linear or logistic regression was used for the analysis of NAA and COPD (GOLD [Global Initiative for Chronic Obstructive Lung Disease] stage II or greater) or lung function. For extended families, linear mixed models and generalized estimating equations were used for the analysis. All models were adjusted for age, sex, educational level, and smoking behavior; models for FEV₁ were also adjusted for height.

Results: The average proportion of European, Native American, and African ancestry among participants was 62%, 35%, and 3%, respectively. After adjustment for current smoking and other covariates, NAA was inversely associated with COPD (OR per 10% increment, 0.55; 95% CI, 0.41-0.75) but positively associated with FEV₁, FVC, and FEV₁/FVC. After additional adjustment for pack-years of smoking, the association between NAA and COPD or lung function measures was slightly attenuated. We found that about 31% of the estimated effect of NAA on COPD is mediated by pack-years of smoking.

Conclusions: NAA is inversely associated with COPD but positively associated with FEV₁ or FVC in Costa Ricans. Ancestral effects on smoking behavior partly explain the findings for COPD but not for FEV₁ or FVC.

CHEST 2014; 145(4):704-710

Abbreviations: BD = bronchodilator; HGDP = Human Genome Diversity Project; LAMP = Local Ancestry in Admixed Populations; SNP = single-nucleotide polymorphism

COPD is a common respiratory disease and a leading cause of morbidity and mortality worldwide.¹ Cigarette smoking is a major risk factor for COPD, but other lifestyle (eg, exposure to wood smoke) or genetic factors influence disease pathogenesis or severity.²

COPD is an important public health problem in Latin America. Among adults aged ≥ 40 years living in five large Latin American cities (Mexico City, Mexico; Caracas, Venezuela; São Paulo, Brazil; Santiago, Chile; and Montevideo, Uruguay), the prevalence of COPD (defined using GOLD [Global Initiative for Chronic Obstructive Lung Disease] criteria following spirometry) was highest in Montevideo and lowest in Mexico City.^{3,4} Given that the average Native American ancestry

of Mexicans living in Mexico City (about 56%)⁴ is markedly higher than that of Uruguayans living in Montevideo (about 1%),⁵ these findings could be partly explained by protective effects of Native American ancestry against nicotine addiction (eg, reduced intensity of smoking) and the detrimental effects of cigarette smoking on lung function or COPD. Because no Costa Rican cities were included in the study, estimates of a relatively low prevalence of COPD in Costa Rica (by self-report) may be a result of underdiagnosis.⁶

A study of adult smokers living in New Mexico found that Hispanic ethnicity or Native American ancestry (estimated with genetic markers) was positively associated with lung function measures (FEV₁ and FEV₁/FVC) but inversely associated with COPD.⁷

Most of the participants in the study (as well as their parents and grandparents) were born in the United States; thus, a “healthy migrant effect” is an unlikely explanation for the findings.⁸ However, a key limitation of the study was an inability to adequately examine the relation between Native American ancestry and COPD or lung function only in racially admixed (Hispanic) subjects because of insufficient statistical power due to small sample size.⁷ Thus, Native American ancestry was largely a surrogate marker of Hispanic ethnicity.

To our knowledge, there has been no study to date of Native American ancestry and lung function or COPD in an exclusively Hispanic subgroup. In addition, there has been no published study of Native American ancestry and lung function or COPD in adults living in a Latin American country. We hypothesized that Native American ancestry would be inversely associated with COPD but positively associated with lung function in Costa Ricans, who are known to have predominantly European (about 60%-65%) and Native American (about 30%-35%) racial ancestry. We examined this hypothesis in our ongoing study of COPD in 578 Costa Ricans.

MATERIALS AND METHODS

Study Population

We recruited 578 subjects between April 2003 and November 2010, including 316 members of 13 families of probands (index case subjects [e-Table 1]) with COPD, 68 case subjects with COPD, 159 control subjects, and 35 individuals who could not be classified as either case or control subjects after review of their spirometric measures (see next). Probands and case subjects were recruited from four major public hospitals in San José, Costa Rica, and through newspaper ads. Control subjects were recruited from a smoking cessation clinic in San José and through newspaper ads. All study participants had to have at least six great-grandparents

born in the Central Valley of Costa Rica (to ensure the subject's descent from the founder population comprising primarily Europeans and Native Americans). All probands and case and control subjects also had to be aged ≥ 21 years and have a history of at least 10 pack-years smoking. Members of families of probands had to be aged ≥ 12 years. Other inclusion criteria for case subjects or probands were physician-diagnosed COPD and a post-bronchodilator (BD) $FEV_1 \leq 65\%$ predicted and a post-BD FEV_1/FVC ratio $< 70\%$. Control subjects had no physician-diagnosed COPD and normal spirometry.

All study participants completed a protocol that included a questionnaire used to collect demographics, smoking history, and respiratory health data⁹; pulmonary function testing; and collection of blood samples for DNA extraction. Spirometry was conducted with a Collins Survey Tach spirometer (Collins Medical). All subjects had to be free of respiratory illnesses for ≥ 4 weeks before spirometry and were instructed (when possible) to avoid the use of inhaled short-acting BDs for ≥ 4 h before testing. Of the 578 study participants, 543 (about 94%) performed spirometry before and 15 min after the administration of 180 μ g (two puffs) albuterol. Forced expiratory maneuvers were judged to be acceptable if they met or exceeded American Thoracic Society criteria.¹⁰ For simplicity and to maximize statistical power, pre-BD values were used for the analysis of lung function measures (FEV_1 , FVC, and FEV_1/FVC).

Written informed consent was obtained from all participants. The study was approved by the institutional review boards of the Hospital Nacional de Niños (UBIHNN-010-2003), Brigham and Women's Hospital (2001-P-001393/49), and the University of Pittsburgh (PRO10040165).

Genotyping

Whereas the first 115 participants (probands, unrelated case subjects, and unrelated control subjects) were genotyped with the Illumina Human610-Quad platform (Illumina Inc) (containing about 610,000 single-nucleotide polymorphisms [SNPs]), the remaining participants ($n = 463$) were genotyped at a later date with the Illumina HumanOmniExpress platform (containing about 730,000 SNPs). We applied stringent filters for quality control of each dataset by excluding SNPs with a minor allele frequency $< 1\%$, a completion rate $< 90\%$, or out of Hardy-Weinberg equilibrium ($P < 1 \times 10^{-6}$ in unrelated control subjects). We then merged the two SNP datasets and implemented quality control measures similar to those previously conducted for each separate dataset while also examining potential batch effects by comparing allele frequencies and performing a clustering analysis. To assess for Mendelian inconsistencies in members of families of probands, the KING (Kinship-based Inference for Gwas) program¹¹ was used to determine relationships based on the genome scan marker data.

Estimation of Racial Ancestry

To select SNPs for ancestry estimation, we first merged about 300,000 SNPs that overlapped between the two genotyping platforms used in this study with reference panels from the Human Genome Diversity Project (HGDP).¹² Such reference panels included data from 88 Europeans, 159 Africans, and 31 Native Americans (seven Colombians, 13 Mayans, and 11 Pima). After applying an SNP-pruning algorithm,¹³ we had 50,000 SNPs with $r^2 < 0.1$ for any pair within a window of 500 kb. Using this panel of SNPs, we applied a model-based program (STRUCTURE,¹⁴ version 2.3.3) to estimate the percentage of racial ancestry from each founder population (African, European, and Native American) for each subject. To avoid introducing bias, we did not explicitly label the reference populations but instead allowed the program

Manuscript received June 5, 2013; revision accepted November 10, 2013.

Affiliations: From the Division of Pulmonary Medicine, Allergy and Immunology (Drs Chen, Brehm, Boutaoui, and Celedón), Children's Hospital of Pittsburgh of UPMC, Pittsburgh, PA; Division of Pediatric Pulmonology (Drs Soto-Quiros and Avila), Hospital Nacional de Niños, San José, Costa Rica; Division of Pulmonary and Critical Care Medicine (Dr Celli), Brigham and Women's Hospital, Boston, MA; and Lovelace Respiratory Research Institute (Drs Bruse and Tesfaigzi), Albuquerque, NM. Drs Chen and Brehm contributed equally to this article.

Funding/Support: This work was supported by the National Heart, Lung, and Blood Institute of the National Institutes of Health [Grant R01-HL073373] and by Children's Hospital of Pittsburgh of UPMC.

Correspondence to: Juan C. Celedón, MD, DrPH, FCCP, Division of Pulmonary Medicine, Allergy and Immunology, Children's Hospital of Pittsburgh of UPMC, 4401 Penn Ave, Pittsburgh, PA 15224; e-mail: juan.celedon@chp.edu

© 2014 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.13-1308

to perform the clustering. We performed 10 independent Markov Chain Monte Carlo runs with 5,000 burn-in and 10,000 iterations and averaged them into the final estimate of ancestral proportions.

As a confirmatory analysis, we repeated the global ancestry estimation using Local Ancestry in Admixed Populations (LAMP),¹⁵ an efficient algorithm for global and local ancestry inference. The estimates of Native American ancestry obtained with STRUCTURE and LAMP were highly correlated ($r^2 > 0.95$). To visualize relative contributions of ancestral populations to the racial admixture of the present cohort and to detect potential outliers, we performed a principal component analysis with the program EIGENSTRAT.¹⁶ For this analysis, reference panels of Europeans, Africans, and Native Americans from the HGDP were combined with the present study samples.

Statistical Analysis

For this analysis, COPD was defined as stage II or higher according to GOLD criteria (both a post-BD $FEV_1 < 80\%$ predicted and a post-BD $FEV_1/FVC < 0.70$). Control subjects had to have a post-BD $FEV_1 \geq 80\%$ predicted and a post-BD $FEV_1/FVC \geq 0.70$. Because there are no reference FEV_1 or FVC values for Costa Ricans, those published for Mexican Americans were used instead.^{17,18} Of the 578 study participants, 94 were classified as having COPD (26 family members of 13 probands and 68 case subjects) and 412 as control subjects; 72 participants (37 family members and 35 unrelated individuals) could not be confidently classified as case or control subjects because they either had no post-BD spirometry ($n = 35$) or had equivocal findings on post-BD spirometry (eg, low FEV_1 but normal FEV_1/FVC ratio, normal FEV_1 but low FEV_1/FVC ratio) and were, thus, excluded from the analysis of COPD.

To account for within-family correlations for members of COPD pedigrees using a kinship coefficient matrix, we used linear mixed-effects models for the analysis of (quantitative) lung function measures (FEV_1 , FVC, FEV_1/FVC ratio) and generalized estimating equation models for the analysis of COPD (a binary outcome). All multivariate analyses were adjusted for age, sex, highest educational level attained by the participant (high school or greater vs noncompletion of high school), smoking status (current vs former or never), and intensity of cigarette smoking (in pack-years); the analyses of FEV_1 and FVC were also adjusted for height. All analyses were conducted using R, version 2.15.0 (www.r-project.org) with

R package GWAF (Genome-Wide Association analyses with Family),¹⁹ kinship, and in-house R functions.

To further assess whether smoking habits explain an association between Native American ancestry and COPD or lung function, we examined the relation among racial ancestry, smoking, and COPD or FEV_1/FVC using mediation analysis^{20,22} (e-Appendix 1, e-Fig 1). We aimed to compare two models relating Native American ancestry to COPD or FEV_1 as follows: one with the mediator (smoking behavior) in the model and the other without. Both analyses were also adjusted for the same variables as those included in the final multivariate analysis of Native American ancestry and COPD. We estimated the proportion of the ancestral effect on COPD or FEV_1 that is mediated by smoking, calculated as follows: indirect effect of Native American ancestry (path ab in e-Fig 1) divided by the total effect of Native American ancestry (path c in e-Fig 1).^{20,21} All estimated coefficients were rescaled before computing indirect effects. All analyses were conducted with Stata 12.0, module function KHB and SEM (StataCorp LP).

RESULTS

The lung function and COPD analyses comprised 578 and 506 subjects, respectively. The main characteristics of the study participants are summarized in Table 1. Compared with control subjects, those with COPD were older and more likely to be current smokers and to have a greater number of pack-years and a lower FEV_1/FVC ratio. Of the 316 members of families of probands with COPD, 269 were aged ≥ 21 years (85.1%) (e-Table 1). Consistent with prior reports in Costa Ricans, study participants were predominantly of European (mean, 62%; range, 28%-87%) and Native American (mean, 35%; range, 5%-72%) ancestry, with a small proportion of African ancestry (mean, 3%; range, 0%-23%) (Fig 1, e-Fig 2).

In a multivariate analysis, Native American ancestry was significantly associated with height and pack-years of cigarette smoking, which are known to influence

Table 1—Main Characteristics of Study Participants

Covariate	All Samples (N = 578)	COPD (n = 94)	Control (n = 412)	Other (n = 72)
Age, y	46.45 $\pm$ 15.51	60.08 $\pm$ 9.19	42.63 $\pm$ 14.83	50.47 $\pm$ 14.95
Male sex	291 (50.34)	58 (61.7) ^a	192 (46.6)	41 (56.9)
Height, cm	163.08 $\pm$ 9.11	161.80 $\pm$ 9.76	163.15 $\pm$ 8.81	165.18 $\pm$ 10.15
BMI, z score	26.15 $\pm$ 5.00	24.52 $\pm$ 5.15 ^a	26.40 $\pm$ 4.77	27.10 $\pm$ 5.98
Smoking				
Ever	398 (69.45)	94 (100) ^a	253 (61.86)	51 (72.9)
Current	216 (37.76)	39 (41.49)	153 (37.41)	24 (34.8)
Pack-y	27.15 $\pm$ 32.79	56.57 $\pm$ 35.01 ^a	18.78 $\pm$ 25.64	37.12 $\pm$ 41.32
Prebronchodilator FEV_1 , L	2.79 $\pm$ 1.09	1.18 $\pm$ 0.52 ^a	3.20 $\pm$ 0.83	2.34 $\pm$ 0.82
Postbronchodilator FEV_1 , L	2.90 $\pm$ 1.14	1.27 $\pm$ 0.52 ^a	3.32 $\pm$ 0.89	2.39 $\pm$ 0.77
Prebronchodilator FVC, L	3.55 $\pm$ 1.16	2.18 $\pm$ 0.81 ^a	3.91 $\pm$ 0.97	3.16 $\pm$ 1.14
Postbronchodilator FVC, L	3.63 $\pm$ 1.18	2.34 $\pm$ 0.78 ^a	3.97 $\pm$ 1.02	3.21 $\pm$ 1.20
Prebronchodilator FEV_1/FVC	0.77 $\pm$ 0.13	0.54 $\pm$ 0.11 ^a	0.82 $\pm$ 0.06	0.75 $\pm$ 0.08
Postbronchodilator FEV_1/FVC	0.78 $\pm$ 0.13	0.54 $\pm$ 0.11 ^a	0.84 $\pm$ 0.05	0.76 $\pm$ 0.09
NAA, %	34.6 $\pm$ 7.4	32.7 $\pm$ 7.9 ^a	35.3 $\pm$ 7.3	33.4 $\pm$ 7.3

Data are presented as mean $\pm$ SD and No. (%). Seventy-two subjects could not be classified into case or control status and were, thus, included in the analysis of lung function but not in the analysis of COPD. NAA = Native American ancestry.

^a $P < .05$ for comparison between case and control subjects.

FIGURE 1. PC analysis of genetic data. The study samples (Costa Rican) were combined with those of Africans, Native Americans, and Europeans from the Human Genome Diversity Project panel. PC analysis was conducted on the combined dataset, and the first two eigenvectors were plotted in pairs. PC = principal component.

lung function (Table 2). In bivariate analyses, Native American ancestry was inversely associated with COPD and FEV₁/FVC (Table 3). After adjustment for age, sex, educational level, current smoking, and case/control status, Native American ancestry was significantly and positively associated with all lung function measures (FEV₁, FVC, and FEV₁/FVC ratio) but significantly and inversely associated with COPD (e-Fig 3, Table 3). For these analyses, the β represents the mean change for each 10% of Native American ancestry. For example, the odds of COPD in a subject with 30% Native American ancestry would be, on average, 45% lower than those of an adult with 20% Native American ancestry. For FEV₁/FVC, a subject with

40% Native American ancestry would have, on average, a ratio 1.9% higher than that of a subject with 20% Native American ancestry.

We then repeated this multivariate analysis after additional adjustment for intensity (pack-years) of cigarette smoking (Table 4). In this analysis, the estimated effect of Native American ancestry on COPD or FEV₁/FVC was slightly attenuated and became non-statistically significant ($P = .06$ for both COPD and FEV₁/FVC). On the other hand, the estimated effect of Native American ancestry on FEV₁ or FVC was slightly attenuated but remained statistically significant.

We conducted three confirmatory analyses of the association between Native American ancestry and lung function or COPD. First, we repeated the multivariate analysis after excluding 227 subjects (all members of families of probands with COPD) who did not have a history of at least 10 pack-years of smoking. Findings from this restricted analysis were very similar to those of our main analysis, despite reduced sample size (e-Table 2). Second, we repeated the multivariate analysis of COPD and lung function after excluding subjects aged < 21 years, obtaining similar results. Third, we repeated the multivariate analysis of lung function measures after excluding 72 subjects who could not be classified as case or control subjects, obtaining very similar results (data not shown).

To attempt to examine whether the observed association between Native American ancestry and FEV₁ or FVC differed between case and control subjects, we conducted an analysis stratified by case/control status (e-Table 3). Among control subjects, Native American ancestry remained significantly and positively associated with FEV₁ or FVC. Among case subjects, Native American ancestry was also positively associated with FEV₁ or FVC, but this association was not statistically significant (likely because of limited statistical power due to the small number of subjects with COPD).

Table 2—Multivariate Analysis of NAA and Height or Pack-Years of Smoking

Subjects	Height, cm		Smoking, Pack-y	
	β (95% CI)	P Value	β (95% CI)	P Value
All subjects (N = 578)				
Each 10% increment in NAA	−2.08 (−2.93 to −1.24)	< .001	−5.37 (−8.75 to −1.99)	.002
Male sex	12.2 (11.2 to 13.2)	< .001	13.5 (9.2 to 17.8)	< .001
Current smoking	0.97 (−0.063 to 1.99)	.07	8.28 (3.77 to 12.8)	< .001
COPD	−2.88 (−4.3 to −1.45)	< .001	17.3 (10.9 to 23.7)	< .001
Noncase subjects (n = 484)				
Each 10% increment in NAA	−1.64 (−2.58 to −0.69)	< .001	−4.44 (−7.96 to −0.92)	.01
Male sex	11.9 (10.9 to 12.9)	< .001	13.2 (8.89 to 17.6)	< .001
Current smoking	1.48 (0.35 to 2.62)	.01	8.71 (4.03 to 13.4)	< .001
Case subjects (n = 94)				
Each 10% increment in NAA	−3.1 (−4.73 to −1.47)	< .001	−11.5 (−20.2 to −2.88)	.01
Male sex	13.7 (11.0 to 16.3)	< .001	14.9 (0.69 to 29.2)	.04
Current smoking	−1.14 (−3.68 to 1.4)	.38	6.46 (−7.59 to 20.5)	.37

All multivariate analyses were adjusted for age in addition to the variables listed in each column. See Table 1 legend for expansion of abbreviation.

Table 3—NAA, COPD, and Lung Function in Study Participants, Without Pack-Years

Predictor	COPD			FEV ₁ , mL			FVC, mL			FEV ₁ /FVC, %		
	OR (95% CI)	P Value		β (95% CI)	P Value		β (95% CI)	P Value		β (95% CI)	P Value	
Unadjusted												
NAA (each 10% increment)	0.62 (0.47 to 0.81)	<.001		778 (−549 to 2,105)	.25		166 (−1,212 to 1,544)	.81		25.8 (8.9 to 42.7)		.003
Multivariate model												
NAA (each 10% increment)	0.55 (0.41 to 0.75)	<.001		118 (42.7 to 192)	.002		119 (29.9 to 208)	.009		0.95 (0.06 to 1.85)		.04
Age (every 10 y)	2.66 (1.86 to 3.81)	<.001		−238 (−272 to −205)	<.001		−216 (−255 to −177)	<.001		−1.85 (−2.26 to −1.44)		<.001
Male sex	3.87 (1.67 to 8.97)	.002		414 (280 to 549)	<.001		571 (413 to 729)	<.001		−1.68 (−2.83 to −0.53)		.004
Height (every 10 cm)	0.61 (0.38 to 0.99)	.04		365 (286 to 444)	<.001		483 (390 to 577)	<.001		NA		...
Education (high school or greater)	0.67 (0.33 to 1.37)	.27		104 (−3.08 to 211)	.06		126 (0.32 to 252)	.05		0.09 (−1.21 to 1.38)		.9
Current smoking	1.18 (0.68 to 2.05)	.56		96.6 (−1.31 to 195)	.05		169 (54.4 to 284)	.004		0.12 (−1.09 to 1.32)		.9
Case	NA	...		−1,489 (−1,627 to −1,352)	<.001		−1,239 (−1,400 to −1,078)	<.001		−23.9 (−25.5 to −22.2)		<.001

NA = not applicable. See Table 1 legend for expansion of other abbreviation.

On the basis of the findings for Native American ancestry and COPD or lung function measures, we used nonparametric structural equation multivariate models to quantify the mediation effect of intensity (pack-years) of smoking within a causal inference network (e-Table 4). In this multivariate analysis, we estimated that about 31% of the estimated effect of Native American ancestry on COPD is mediated by intensity of smoking ($P = .001$). In addition, we estimated that about 9.6% of the estimated effect of Native American ancestry on FEV₁ is mediated by intensity of smoking; however, this was not statistically significant ($P = .09$).

DISCUSSION

The study of racial ancestry as it pertains to human health has become popular again in the scientific literature as a result of projects like HapMap,²³ the HGDP,¹² and the 1000 Genomes Project.²⁴ Data on founder human populations in these studies can be used to infer the racial ancestry of admixed individuals (using statistical methods such as LAMP¹⁵ and STRUCTURE¹⁴) far more precisely than using self-reported race or ethnicity.

In the present report, we show that Native American ancestry is inversely associated with COPD but positively associated with lung function in racially admixed adults in Costa Rica. To our knowledge, this study is the first of Native American ancestry (determined by genetic markers) and COPD or lung function in an exclusively (and homogeneous) Hispanic population as well as the first such study in a Latin American country.

We found that Native American ancestry was inversely associated with COPD but positively associated with lung function measures (FEV₁, FVC, and FEV₁/FVC ratio), even after accounting for age, sex, height, and current smoking. However, the observed association between Native American ancestry and COPD or FEV₁/FVC became nonsignificant after additional adjustment for intensity (pack-years) of cigarette smoking, suggesting that smoking behavior partly mediates the estimated effect of Native American ancestry on COPD or FEV₁/FVC. In fact, we estimated that nearly one-third of the observed association between ancestry and COPD is mediated by intensity of smoking.

In contrast to the findings for COPD or FEV₁/FVC, the estimated effect of Native American ancestry on FEV₁ or FVC remained statistically significant after excluding subjects with COPD or adjusting for intensity of smoking, suggesting that such an effect may be mediated by genetic factors on lung function either directly or indirectly (eg, by affecting body habitus [trunk-to-limb ratio]). Although we do not have data

Table 4—NAA, COPD, and Lung Function in Study Participants, After Additional Adjustment for Pack-Years of Cigarette Smoking

Predictor	COPD		FEV ₁ , mL		FVC, mL		FEV ₁ /FVC, %	
	OR (95% CI)	P Value	β (95% CI)	P Value	β (95% CI)	P Value	β (95% CI)	P Value
NAA (each 10% increment)	0.67 (0.45 to 1.01)	.06	109 (33.6 to 184)	.005	112 (22.2 to 202)	.02	0.87 (−0.027 to 1.78)	.06
Age (every 10 y)	2.33 (1.68 to 3.23)	<.001	−231 (−268 to −194)	<.001	−212 (−255 to −169)	<.001	−1.73 (−2.18 to −1.29)	<.001
Male sex	2.91 (1.2 to 7.06)	.02	435 (300 to 570)	<.001	590 (431 to 749)	<.001	−1.52 (−2.71 to −0.34)	.01
Height (every 10 cm)	0.53 (0.3 to 0.92)	.02	367 (288 to 446)	<.001	485 (392 to 578)	<.001	NA	...
Education (high school or greater)	0.63 (0.3 to 1.31)	.22	103 (−4.09 to 209)	.06	124 (−1.32 to 250)	.05	0.11 (−1.19 to 1.41)	.87
Current smoking	0.88 (0.46 to 1.66)	.69	116 (17.3 to 215)	.02	187 (71.2 to 304)	.002	0.21 (−1.02 to 1.43)	.74
Pack-y smoking	1.03 (1.02 to 1.05)	<.001	−1.65 (−3.44 to 0.13)	.07	−1.35 (−3.45 to 0.75)	.21	−0.013 (−0.035 to 0.0095)	.26
Case	NA	...	−1,438 (−1,581 to −1,296)	<.001	−1,190 (−1,358 to −1,022)	<.001	−23.7 (−25.5 to −22.0)	<.001

See Table 1 and 3 legends for expansion of abbreviations.

on body habitus, previous results in African Americans suggested that less than one-half of the estimated effect of ethnicity on lung function is explained by changes in body habitus or socioeconomic status.²⁵

Native American ancestry likely captures not only genetic variation (as some alleles may be more or less frequent in Native Americans) but also social patterns and behaviors. For example, Native American ancestry was associated with lower intensity of smoking and shorter height among Costa Ricans in this study. Determining the proportion of these results that is explained by genetics (eg, susceptibility variants for nicotine addiction, height) or lifestyle (eg, shorter height due to inadequate nutrition in childhood) is of interest but beyond the scope of this report.

This study has several strengths, including a cohort that is homogenous on the basis of ethnicity, racial ancestry (predominantly European and Native American), health care (all Costa Ricans have universal access to a nationwide health-care system), and ability to adjust for potential confounders of the relation between ancestry and COPD (age, sex, educational level, and smoking status). We also recognize several limitations. First, we had a relatively small number of case subjects, thus, reducing the statistical power for the analysis of COPD (in all subjects) or lung function (in case subjects only). Second, we used data from Native American populations in the HGDP (and not from [largely extinct] Native Americans in the founder population of Costa Rica) as reference panels for ancestry estimation. However, this likely resulted in a small degree of bias that would not substantially affect our conclusions. Third, this was a cross-sectional study, and, thus, reverse causation is possible and likely explains the positive association between current smoking and lung function measures (because subjects with lower lung function may be more symptomatic and, thus, less likely to smoke). Fourth, the mediation analysis methods currently available were not designed for family-based cohorts, which can lead to some bias. However, we obtained similar findings through two different approaches (see e-Appendix 1). Finally, the results may not be generalizable to Hispanic subgroups with a different racial admixture (eg, those with a significant proportion of African ancestry). However, the findings are generally consistent with and significantly extend those for Hispanic ethnicity and COPD or lung function in New Mexico⁷ and are, thus, relevant to Hispanic subgroups of predominantly European and Native American descent (eg, Mexicans, Mexican Americans).

In summary, Native American ancestry is inversely associated with COPD but positively associated with FEV₁ or FVC in Costa Ricans. Whereas the findings for COPD are partly mediated by ancestral effects on smoking behavior (intensity), those for FEV₁ or FVC

are not. Taken together with other findings,^{3,7} the present results may explain the Hispanic paradox: a lower prevalence of COPD in Hispanic populations with a significant proportion of Native American ancestry (eg, Mexicans) than in those with a significant proportion of African ancestry (eg, Puerto Ricans). In addition, these results highlight the importance of considering individual variability in culture and racial ancestry when assessing risk for COPD. Future studies are necessary to identify the genetic and behavioral factors underlying potential effects of Native American ancestry on lung function.

ACKNOWLEDGMENTS

Author contributions: Drs Chen and Brehm had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

Dr Chen: contributed to the data analysis and interpretation, writing of the manuscript, and review and approval of the manuscript.

Dr Brehm: contributed to the data analysis and interpretation, writing of the manuscript, and review and approval of the manuscript.

Dr Boutaoui: contributed to the data collection and review and approval of the manuscript.

Dr Soto-Quiros: contributed to the data collection and review and approval of the manuscript.

Dr Avila: contributed to the data collection and review and approval of the manuscript.

Dr Celli: contributed to the data analysis and interpretation and review and approval of the manuscript.

Dr Bruse: contributed to the data analysis and interpretation and review and approval of the manuscript.

Dr Tesfaigzi: contributed to the data analysis and interpretation and review and approval of the manuscript.

Dr Celedón: contributed to the data analysis and interpretation, writing of the manuscript, and review and approval of the manuscript.

Financial/nonfinancial disclosures: The authors have reported to CHEST the following conflicts of interest: Dr Celedón served as a single-time consultant for Genentech, Inc, in 2011 on an issue unrelated to this article. Drs Chen, Brehm, Boutaoui, Soto-Quiros, Avila, Celli, Bruse, and Tesfaigzi have reported that no potential conflicts of interest exist with any companies/organizations whose products or services may be discussed in this article.

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.

Other contributions: The authors thank the participants and their families in Costa Rica. All analyses were conducted at Children's Hospital of Pittsburgh of UPMC under local guidelines and policies.

Additional information: The e-Appendix, e-Figures, and e-Tables can be found in the "Supplemental Materials" area of the online article.

REFERENCES

- Lopez AD, Shibuya K, Rao C, et al. Chronic obstructive pulmonary disease: current burden and future projections. *Eur Respir J*. 2006;27(2):397-412.
- Silverman EK. Perspective: how can genetics help? *Nature*. 2012;489(7417):S7.
- Brehm JM, Celedón JC. Chronic obstructive pulmonary disease in Hispanics. *Am J Respir Crit Care Med*. 2008;177(5):473-478.
- Lisker R, Babinsky V. Admixture estimates in nine Mexican Indian groups and five East Coast localities. *Rev Invest Clin*. 1986;38(2):145-149.

- Sans M, Salzano FM, Chakraborty R. Historical genetics in Uruguay: estimates of biological origins and their problems. *Hum Biol*. 1997;69(2):161-170.
- Celedón JC, Palmer LJ, Litonjua AA, et al. Body mass index and asthma in adults in families of subjects with asthma in Anqing, China. *Am J Respir Crit Care Med*. 2001;164(10):1835-1840.
- Bruse S, Sood A, Petersen H, et al. New Mexican Hispanic smokers have lower odds of chronic obstructive pulmonary disease and less decline in lung function than non-Hispanic whites. *Am J Respir Crit Care Med*. 2011;184(11):1254-1260.
- Harhay MO. The Hispanic paradox and chronic obstructive pulmonary disease. *Am J Respir Crit Care Med*. 2012;185(11):1246.
- Ferris BG. Epidemiology standardization project (American Thoracic Society). *Am Rev Respir Dis*. 1978;118(6 pt 2):1-120.
- American Thoracic Society. Standardization of spirometry, 1994 update. *Am J Respir Crit Care Med*. 1995;152(3):1107-1136.
- Manichaikul A, Mychaleckyj JC, Rich SS, Daly K, Sale M, Chen WM. Robust relationship inference in genome-wide association studies. *Bioinformatics*. 2010;26(22):2867-2873.
- Li JZ, Absher DM, Tang H, et al. Worldwide human relationships inferred from genome-wide patterns of variation. *Science*. 2008;319(5866):1100-1104.
- Purcell S, Neale B, Todd-Brown K, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet*. 2007;81(3):559-575.
- Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. *Genetics*. 2000;155(2):945-959.
- Pasaniuc B, Sankararaman S, Kimmel G, Halperin E. Inference of locus-specific ancestry in closely related populations. *Bioinformatics*. 2009;25(12):i213-i221.
- Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet*. 2006;38(8):904-909.
- Hankinson JL, Odencrantz JR, Fedan KB. Spirometric reference values from a sample of the general US population. *Am J Respir Crit Care Med*. 1999;159(1):179-187.
- Wang S, Ray N, Rojas W, et al. Geographic patterns of genome admixture in Latin American Mestizos. *PLoS Genet*. 2008;4(3):e1000037.
- Chen MH, Yang Q. GWAf: an R package for genome-wide association analyses with family data. *Bioinformatics*. 2010;26(4):580-581.
- Kohler U, Karlson KB, Holm A. Comparing coefficients of nested nonlinear probability models. *Stata J*. 2011;11(3):420-438.
- MacKinnon DP. *Introduction to Statistical Mediation Analysis*. New York, NY: Lawrence Erlbaum Associates; 2008.
- Baron RM, Kenny DA. The moderator-mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. *J Pers Soc Psychol*. 1986;51(6):1173-1182.
- Thorisson GA, Smith AV, Krishnan L, Stein LD. The International HapMap Project Web site. *Genome Res*. 2005;15(11):1592-1593.
- Abecasis GR, Altshuler D, Auton A, et al; 1000 Genomes Project Consortium. A map of human genome variation from population-scale sequencing [published correction appears in *Nature*. 2011;473(7319):1061-1073]. *Nature*. 2010;467(7319):1061-1073.
- Harik-Khan RI, Muller DC, Wise RA. Racial difference in lung function in African-American and White children: effect of anthropometric, socioeconomic, nutritional, and environmental factors. *Am J Epidemiol*. 2004;160(9):893-900.