

# UCSF

## UC San Francisco Previously Published Works

### Title

The Effects of Rare SERPINA1 Variants on Lung Function and Emphysema in SPIROMICS

### Permalink

<https://escholarship.org/uc/item/3h9537nb>

### Journal

American Journal of Respiratory and Critical Care Medicine, 201(5)

### ISSN

1073-449X

### Authors

Ortega, Victor E  
Li, Xingnan  
O'Neal, Wanda K  
et al.

### Publication Date

2020-03-01

### DOI

10.1164/rccm.201904-0769oc

Peer reviewed

# The Effects of Rare *SERPINA1* Variants on Lung Function and Emphysema in SPIROMICS

Victor E. Ortega<sup>1</sup>, Xingnan Li<sup>2</sup>, Wanda K. O'Neal<sup>3</sup>, Lela Lackey<sup>4</sup>, Elizabeth Ampleford<sup>1</sup>, Gregory A. Hawkins<sup>1</sup>, Philip J. Grayeski<sup>3</sup>, Alain Laederach<sup>4</sup>, Igor Barjaktarevic<sup>5</sup>, R. Graham Barr<sup>6</sup>, Christopher Cooper<sup>5</sup>, David Couper<sup>3</sup>, MeiLan K. Han<sup>7\*</sup>, Richard E. Kanner<sup>8</sup>, Eric C. Kleerup<sup>5</sup>, Fernando J. Martinez<sup>9\*</sup>, Robert Paine III<sup>8</sup>, Stephen P. Peters<sup>1</sup>, Cheryl Pirozzi<sup>8</sup>, Stephen I. Rennard<sup>10,11</sup>, Prescott G. Woodruff<sup>12\*</sup>, Eric A. Hoffman<sup>13,14,15</sup>, Deborah A. Meyers<sup>2</sup>, and Eugene R. Bleecker<sup>2</sup>; for the SPIROMICS Investigators

<sup>1</sup>Center for Precision Medicine, Department of Internal Medicine, Wake Forest School of Medicine, Winston-Salem, North Carolina; <sup>2</sup>Department of Medicine, University of Arizona, Tucson, Arizona; <sup>3</sup>University of North Carolina at Chapel Hill School of Medicine, Chapel Hill, North Carolina; <sup>4</sup>Department of Biology, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina; <sup>5</sup>Department of Medicine, David Geffen School of Medicine, Los Angeles, California; <sup>6</sup>Columbia University Medical Center, New York City, New York; <sup>7</sup>Division of Pulmonary and Critical Care Medicine, Michigan Medicine, University of Michigan, Ann Arbor, Michigan; <sup>8</sup>Division of Respiratory, Critical Care, and Occupational Pulmonary Medicine, Department of Medicine, University of Utah Health Sciences Center, Salt Lake City, Utah; <sup>9</sup>Division of Pulmonary and Critical Care Medicine, Department of Medicine, Weill Cornell Medical College of Cornell University, New York City, New York; <sup>10</sup>Division of Pulmonary, Critical Care, Sleep, and Allergy, Department of Medicine, University of Nebraska, Omaha, Nebraska; <sup>11</sup>Innovative Medicines and Early Development (IMED) Biotech Unit, AstraZeneca, Cambridge, United Kingdom; <sup>12</sup>Division of Pulmonary and Critical Care Medicine, Department of Medicine, Cardiovascular Research Institute, University of California, San Francisco, California; and <sup>13</sup>Department of Radiology, <sup>14</sup>Department of Medicine, and <sup>15</sup>Department of Biomedical Engineering, University of Iowa Carver College of Medicine, Iowa City, Iowa

## Abstract

**Rationale:** The role of PI (protease inhibitor) type Z heterozygotes and additional rare variant genotypes in the gene encoding alpha-1 antitrypsin, *SERPINA1* (serpin peptidase inhibitor, clade A, member 1), in determining chronic obstructive pulmonary disease risk and severity is controversial.

**Objectives:** To comprehensively evaluate the effects of rare *SERPINA1* variants on lung function and emphysema phenotypes in subjects with significant tobacco smoke exposure using deep gene resequencing and alpha-1 antitrypsin concentrations.

**Methods:** DNA samples from 1,693 non-Hispanic white individuals, 385 African Americans, and 90 Hispanics with  $\geq 20$  pack-years smoking were resequenced for the identification of rare variants (allele frequency  $< 0.05$ ) in 16.9 kB of *SERPINA1*.

**Measurements and Main Results:** White PI Z heterozygotes confirmed by sequencing (MZ;  $n = 74$ ) had lower post-bronchodilator FEV<sub>1</sub> ( $P = 0.007$ ), FEV<sub>1</sub>/FVC ( $P = 0.003$ ), and greater computed tomography–based emphysema ( $P = 0.02$ ) compared with 1,411 white individuals without PI Z, S, or additional rare variants

denoted as V<sub>R</sub>. PI Z–containing compound heterozygotes (ZS/ZV<sub>R</sub>;  $n = 7$ ) had lower FEV<sub>1</sub>/FVC ( $P = 0.02$ ) and forced expiratory flow, midexpiratory phase ( $P = 0.009$ ). Nineteen white heterozygotes for five non-S/Z coding variants associated with lower alpha-1 antitrypsin had greater computed tomography–based emphysema compared with those without rare variants. In African Americans, a 5' untranslated region insertion (rs568223361) was associated with lower alpha-1 antitrypsin and functional small airway disease ( $P = 0.007$ ).

**Conclusions:** In this integrative deep sequencing study of *SERPINA1* with alpha-1 antitrypsin concentrations in a heavy smoker and chronic obstructive pulmonary disease cohort, we confirmed the effects of PI Z heterozygote and compound heterozygote genotypes. We demonstrate the cumulative effects of multiple *SERPINA1* variants on alpha-1 antitrypsin deficiency, lung function, and emphysema, thus significantly increasing the frequency of *SERPINA1* variation associated with respiratory disease in at-risk smokers.

**Keywords:** chronic obstructive pulmonary disease; alpha-1 antitrypsin; *SERPINA1*; rare variant; emphysema

(Received in original form April 9, 2019; accepted in final form October 24, 2019)

\*M.K.H. and P.G.W. are Associate Editors and F.J.M. is Deputy Editor of *AJRCCM*. Their participation complies with American Thoracic Society requirements for recusal from review and decisions for authored works.

Correspondence and requests for reprints should be addressed to Victor E. Ortega, M.D., Ph.D., Center for Precision Medicine, Wake Forest School of Medicine, Medical Center Boulevard, Winston-Salem, NC 27157. E-mail: vortega@wakehealth.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](http://www.atsjournals.org).

More information about the study and how to access SPIROMICS data is at [www.spiromics.org](http://www.spiromics.org).

Am J Respir Crit Care Med Vol 201, Iss 5, pp 540–554, Mar 1, 2020

Copyright © 2020 by the American Thoracic Society

Originally Published in Press as DOI: 10.1164/rccm.201904-0769OC on October 29, 2019

Internet address: [www.atsjournals.org](http://www.atsjournals.org)

The strongest genetic risk factor for chronic obstructive pulmonary disease (COPD) and emphysema susceptibility has been the locus coding for *SERPINA1* (alpha-1 antitrypsin, serpin peptidase inhibitor, clade A, member 1) on chromosome 14q32. Alpha-1 antitrypsin deficiency is caused by genetic variants in *SERPINA1*, resulting in increased risk for early-onset obstructive lung disease and emphysema, specifically in cigarette smokers. *SERPINA1* variation has been estimated to account for <1% of all cases of COPD and emphysema (1–3).

Traditionally, alpha-1 antitrypsin deficiency is diagnosed based on serum alpha-1 antitrypsin levels and classified with protein phenotyping (PI [protease inhibitor] typing) through protein isoelectric focusing (PIEF). Based on PIEF, PI type M is the major wild-type phenotype associated with normal alpha-1 antitrypsin levels and homozygotes for this phenotype (MM) have a similar COPD risk as the general population.

In individuals of European white descent, the most frequent PI types associated with alpha-1 antitrypsin deficiency are PI type S due to low-frequency coding variants at amino acid position 288 (Glu<sup>288</sup>Val, rs17580) and PI type Z at position 366 (Glu<sup>366</sup>Lys, rs28929474) (4–6). Homozygotes for the Lys<sup>366</sup> allele (ZZ), Val<sup>288</sup> allele (SS), and compound heterozygotes (ZS) have been associated with lower serum alpha-1 antitrypsin levels and an increased risk for COPD and emphysema (2, 7, 8). The influence of the PI Z heterozygote genotype on COPD and emphysema risk and severity is still controversial with multiple studies

showing contradictory results (9, 10). Earlier PIEF-based studies suggested that MZ heterozygotes with a history of smoking have an increased risk for COPD and lower lung function but were based on cohorts with a minimal smoking history (without the necessary gene–environment interaction) or not ascertained for COPD (9, 11, 12). More recent studies with targeted genotyping have shown that PI type Z heterozygote genotype was associated with lower lung function, COPD, and computed tomography scan measures of emphysema; however, these analyses did not account for additional rare *SERPINA1* variation or alpha-1 antitrypsin concentrations (13–15). A genome-wide association study for emphysema and two for Global Initiative for Chronic Obstructive Lung Disease (GOLD) grade 2 to 4 COPD have also identified loci adjacent to *SERPINA1*, suggesting the importance of rare variation at this locus (13, 16–18). None of these studies performed DNA sequencing to exclude the effects of background rare variation in MZ heterozygotes or for the identification of novel pathogenic rare genotypes.

Sequencing of *SERPINA1* is necessary to confirm whether one PI Z allele is sufficient to affect disease risk and severity, and is essential for identifying additional rare variants that could influence pathogenesis (19). Altogether, MZ heterozygotes and novel pathogenic rare genotypes could increase the influence of *SERPINA1* on COPD risk and severity in ever-smokers. We hypothesize that PI type Z and additional rare variants cumulatively

affect lung function and emphysema, and, therefore, COPD risk and severity in subjects with significant tobacco smoke exposure (3). Thus, we performed a deep resequencing study of *SERPINA1* in a comprehensively characterized (including alpha-1 antitrypsin concentration measurements in a large subset), multiethnic cohort from the NHLBI SPIROMICS (Subpopulations and Intermediate Outcomes Measures in COPD Study). The results of this study have been previously reported in abstract form (20, 21).

## Methods

### Study Population

SPIROMICS participants were 40 to 80 years of age with a smoking history  $\geq 20$  pack-years with COPD (GOLD spirometric grades 1–4) and without COPD (22, 23). Participants were comprehensively characterized with annual pre- and post-bronchodilator lung function measures for up to 3 years, computed tomography scans, and standardized questionnaires (22). Alpha-1 antitrypsin concentrations were obtained in 1,391 subjects using Myriad RBM's CustomMAP using Luminex xMAP technology at the initial baseline visit.

### Resequencing of *SERPINA1*

DNA samples from 1,693 non-Hispanic white individuals (1,097 with alpha-1 antitrypsin concentrations [see baseline characteristics in Table E1 in the online supplement]), 385 African Americans

---

SPIROMICS was supported by contracts from the NIH/NHLBI (HHSN268200900013C, HHSN268200900014C, HHSN268200900015C, HHSN268200900016C, HHSN268200900017C, HHSN268200900018C, HHSN268200900019C, and HHSN268200900020C) and supplemented by contributions made through the Foundation for the NIH and the COPD Foundation from AstraZeneca/MedImmune; Bayer; Bellerophon Therapeutics; Boehringer-Ingelheim Pharmaceuticals, Inc.; Chiesi Farmaceutici S.p.A.; Forest Research Institute, Inc.; GlaxoSmithKline; Grifols Therapeutics, Inc.; Ikaria, Inc.; Novartis Pharmaceuticals Corporation; Nycomed GmbH; ProterixBio; Regeneron Pharmaceuticals, Inc.; Sanofi; Sunovion; Takeda Pharmaceutical Co.; and Theravance Biopharma. Additional NIH grant funding included a career development award from the NHLBI (K08 HL118128, R01HL142992, and R01HL111527). The project officers from the Lung Division of the NHLBI were Lisa Postow and Lisa Viviano. Sequencing services were provided through the Resequencing and Genotyping Service by the Northwest Genomics Center at the University of Washington, Department of Genome Sciences, under U.S. federal government contract number HHSN268201100037C from the NHLBI.

Author Contributions: V.E.O.: study design, data quality control and harmonization, data analysis, and manuscript preparation. X.L.: data quality control and harmonization, data analysis, and critical review of manuscript. W.K.O'N.: measurement of alpha-1 antitrypsin concentrations, data quality control and harmonization, data analysis, and critical review of manuscript. L.L.: noncoding variant analysis, reporter assay, and manuscript preparation. E.A.: study design, data collection, and critical review of manuscript. G.A.H.: study design, data collection, and critical review of manuscript. P.J.G.: noncoding variant analysis, reporter assay, and manuscript preparation. A.L.: noncoding variant analysis and manuscript preparation. I.B.: study design, data collection, and critical review of manuscript. R.G.B.: study design, data collection, and critical review of manuscript. C.C.: study design, data collection, and critical review of manuscript. D.C.: study design, data collection, and critical review of manuscript. M.K.H.: study design, data collection, and critical review of manuscript. R.E.K.: study design, data collection, and critical review of manuscript. E.C.K.: study design, data collection, and critical review of manuscript. F.J.M.: study design, data collection, and critical review of manuscript. R.P.: study design, data collection, and critical review of manuscript. S.P.P.: study design, data collection, and critical review of manuscript. C.P.: study design, data collection, and critical review of manuscript. S.I.R.: study design, data collection, and critical review of manuscript. P.G.W.: study design, data collection, and critical review of manuscript. E.A.H.: study design, data collection, and critical review of manuscript. D.A.M.: data quality control and harmonization, data analysis, and critical review of manuscript. E.R.B.: data quality control and harmonization, data analysis, and critical review of manuscript.

## At a Glance Commentary

### Scientific Knowledge on the

**Subject:** The influence of the MZ heterozygote genotype on chronic obstructive pulmonary disease phenotypes is controversial, with multiple studies showing contradictory results. Over the past decade, protein isoelectric focusing and genotyping-based studies in European white individuals, including one large genotyping study of African Americans and white individuals, have demonstrated that protease inhibitor Z heterozygotes have lower lung function and more emphysema when appropriately evaluating at-risk individuals with sufficient pack-years smoking history. These studies did not evaluate for the effects of other rare variants with deep sequencing, were not complemented by alpha-1 antitrypsin measurements, and most cohorts had lower smoking histories (<20 pack-years) (8, 11, 12, 14, 15, 17, 18).

### What This Study Adds to the Field:

This study is the most comprehensive analysis of rare *SERPINA1* variants in a multiethnic cohort with significant cigarette-smoking exposure ascertained for a broad range of chronic obstructive pulmonary disease severity, not alpha-1 antitrypsin deficiency. This study is the first to perform deep gene sequencing in combination with alpha-1 antitrypsin concentrations to detect the cumulative effects of protease inhibitors Z, S, and additional rare *SERPINA1* variants, combined and independently (34).

(231 with alpha-1 antitrypsin concentrations), and 90 Hispanics (63 with alpha-1 antitrypsin concentrations) were sequenced with an average depth of 61.8 reads per target region ( $\times$ ; median =  $60.5 \times$ ) for the identification of less common and rare variants with minor allele frequencies less than or equal to 0.05 in a 16.9 kB region (chromosome 14, nucleotide position 94841102–94857987 based on human genome assembly GRCh37 [hg19]) of *SERPINA1* consisting of the five coding exons, four introns, and the 5' and 3' untranslated regions (UTRs) (23). Resequencing services were provided

through the NHLBI Resequencing and Genotyping Service. Twelve variants with minor allele frequency less than or equal to 0.05 identified with sequencing, including PI Z, were also genotyped and confirmed with the Illumina OmniExpress HumanExome BeadChip (Illumina, Inc.) and PI S with the MassARRAY system (Sequenom, Inc.). A flow diagram of the analytical methods is shown as Figure 1.

### Statistical Methods

*SERPINA1* polymorphisms that had a call rate less than or equal to 0.98 and common variants that did not meet Hardy–Weinberg expectations ( $P < 0.05$ ) or showed nonrandom missingness were excluded with the exception of PI type Z, which did not meet Hardy–Weinberg expectations ( $P = 6.29 \times 10^{-6}$ ) owing to a strong influence on COPD risk in white individuals (24). We annotated variants identified with resequencing based on GRCh37 using National Center for Biotechnology Information ClinVar, the NHLBI Grand Opportunity Exome Sequencing Program, Exome Aggregation Consortium, International Genome Sample Resource based on the 1000 Genomes Project, and Online Mendelian Inheritance in Man databases to determine PI type and novelty (25–29).

Regression-based burden tests of PI Z, S, and other rare variants were stratified by self-reported ethnic group to detect gene-level effects on the strongly correlated co-primary outcomes of baseline post-bronchodilator FEV<sub>1</sub>% predicted, FEV<sub>1</sub>/FVC ratio, and computed tomography scan evidence of emphysema based on percentage low attenuation area of bilateral lungs  $\leq -950$  Hounsfield units (HFU) ( $\% \leq -950$  HFU) at TLC (see Table E2) (30). We subsequently analyzed individual variants with regression-based allelic association analyses. Based on the significant co-primary outcome associations found with burden tests and for PI Z (C/T alleles) (15), we evaluated six collapsed genotype groups based on the presence or absence of PI Z, PI S, or other rare variant minor alleles for which little or no *a priori* knowledge of pathogenicity was available ( $V_R$  designates all non-Z and non-S rare variants). Shown as Figure 1:

1. Subjects with two *SERPINA1* variants were grouped as 1) PI Z homozygotes (ZZ), 2) PI Z compound heterozygotes

(ZS/ZV<sub>R</sub>), and 3) compound heterozygotes without PI Z (SS/V<sub>R</sub>V<sub>R</sub>/SV<sub>R</sub>).

2. Subjects with one *SERPINA1* variant were grouped as 1) MZ heterozygotes and 2) MS/MV<sub>R</sub>.
3. Subjects with no identifiable rare variant, PI S, or Z were a single group (No Z/S/V<sub>R</sub>).

Secondary, related outcomes included computed tomography scan evidence of air trapping based on percentage of low attenuation area of bilateral lungs less than or equal to  $-856$  HFU ( $\% \leq -856$  HFU) at residual volume, small airway disease as measured by computed tomography scan-based parametric response mapping of functional small airway disease (PRM<sup>ISAD</sup>), and clinical outcome measures (see Figure 1) (31). Analyses were performed using JMPGenomics 6.0 (SAS Institute Inc.) and PLINK v1.07 adjusted for age, sex, study site, body mass index, pack-years cigarette smoking, FVC% predicted for forced expiratory flow, midexpiratory phase (FEF<sub>25-75</sub>), height for computed tomography scan measures, and C-reactive protein for alpha-1 antitrypsin concentrations (24). To account for ancestry in African Americans and Hispanics, whole-genome European, African, and Native American percentage ancestry was estimated using ADMIXTURE for regression-based models (24, 32). Haplotype analyses and phasing of compound heterozygotes was performed with 16 common tagging single nucleotide polymorphisms (SNPs;  $r^2 \leq 0.5$ ), PI Z, and S.

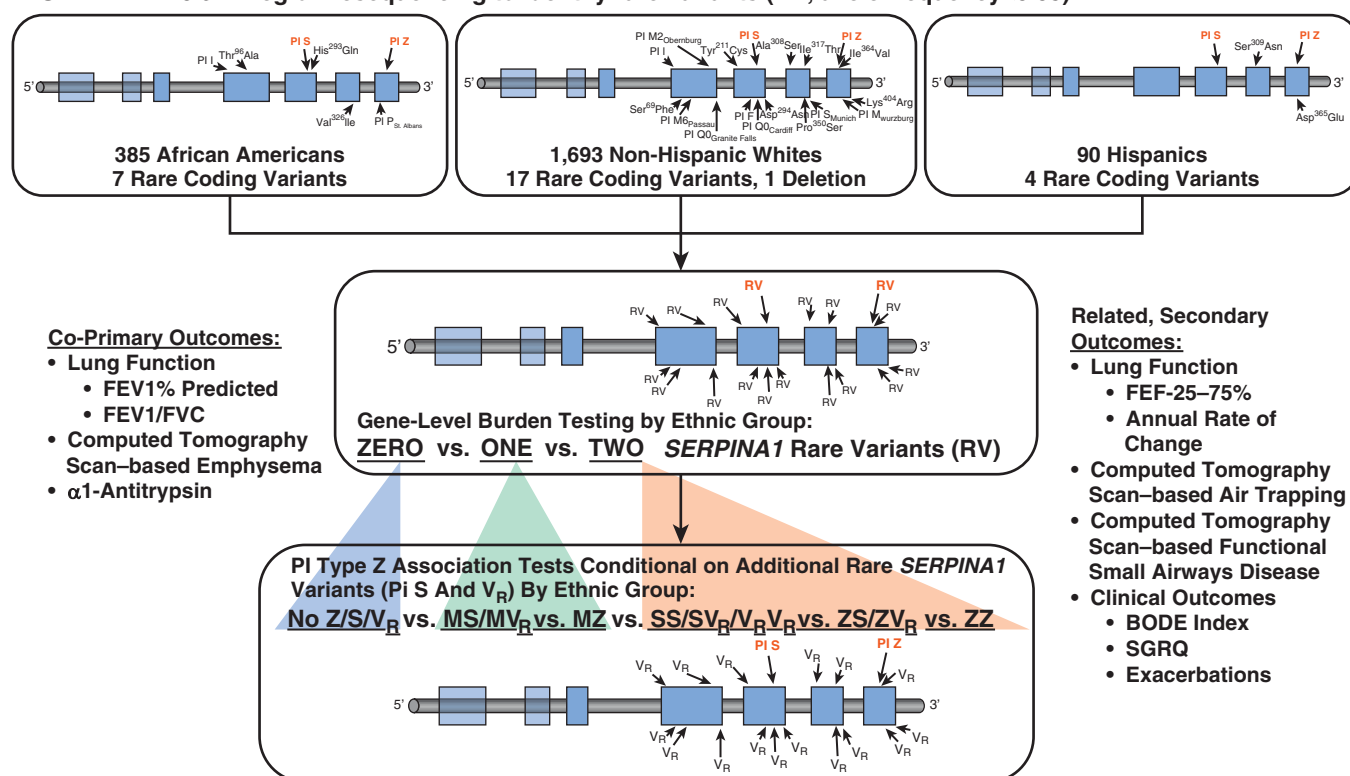
### Luciferase Reporter Assay

To evaluate a rare, 5' UTR G-insertion associated with alpha-1 antitrypsin concentrations and PRM<sup>ISAD</sup> in African Americans, the 5' UTR of *SERPINA1* NM\_000295 was synthesized and homology inserted into the pNL3.2 Nano luciferase plasmid. All experiments were performed in triplicate; experimental details, including PCR primers and genomic coordinates, are provided in Table E3 and the online supplement methods.

## Results

### *SERPINA1* Variants in the SPIROMICS Multiethnic Study Cohort

The baseline characteristics are shown by ethnic group in Table 1. Non-Hispanic

**SERPINA1 16.9kB region resequencing to identify rare variants (RV, allele frequency <0.05)**

**Figure 1.** *SERPINA1* genetic study flow diagram summarizes the sequential analytical steps for characterizing rare *SERPINA1* variants in SPIROMICS. First, burden tests for PIs (protease inhibitors) Z, S, and 21 additional rare exonic coding variants and 1 frameshift insertion, all with a minor allele frequency of less than 0.05, were performed by self-reported ethnic group to evaluate for gene-level associations with the co-primary outcomes (30). Second, individual rare variants were analyzed with regression-based allelic association tests to further evaluate significant associations. Based on the strong effect of PI type Z variant on the co-primary outcomes in white individuals, we performed association testing of this variant conditional on additional less common (PI S) and rare *SERPINA1* variants. We compared white individuals without a PI Z, S, or additional rare variant allele (no Z/S/V<sub>R</sub>, where V<sub>R</sub> designates all non-Z and non-S rare variants), heterozygotes for non-PI Z rare variation (MS/MV<sub>R</sub>), PI Z heterozygotes without another rare variant (MZ), compound heterozygotes with non-PI Z rare variation (SS/SV<sub>R</sub>/V<sub>R</sub>V<sub>R</sub>), PI Z-containing compound heterozygotes with PI S or additional rare variants (ZS/ZV<sub>R</sub>), and PI Z homozygotes (ZZ) to evaluate the effects of PI Z heterozygotes with the confirmed absence of other rare variants and PI Z in compound heterozygotes with another rare variant. We performed similar analyses for secondary related outcomes. BODE = body mass index, airflow obstruction, dyspnea, and exercise; FEF<sub>25-75</sub> = forced expiratory flow, midexpiratory phase; SGRQ = St. George's Respiratory Questionnaire; SPIROMICS = Subpopulations and Intermediate Outcomes Measures in Chronic Obstructive Pulmonary Disease Study.

white individuals had a higher smoking history (51.8 pack-years) than African Americans (41.0 pack-years) or Hispanics (46.8 pack-years) ( $P = 5.60 \times 10^{-5}$ ). Post-bronchodilator FEV<sub>1</sub>% and FEV<sub>1</sub>/FVC ratio was lower in white individuals (71.8% predicted and 0.59, respectively) compared with African Americans (76.6% and 0.64) and Hispanics (80% and 0.64;  $P = 0.0004$  and  $P = 0.0001$ ). Computed tomography scan-based emphysema based on percentage  $\leq -950$  HFU and air trapping based on percentage  $\geq -856$  HFU was highest in white individuals (8.19% and 25.9%, respectively) compared with African Americans (7.70% and 22.2%) and Hispanics (4.89% and 20.0%;  $P = 0.0006$  and  $P = 3.80 \times 10^{-5}$ ). White individuals ( $n = 1,097$ ; 1.98 mg/ml), African

Americans ( $n = 231$ ; 1.95 mg/ml), and Hispanics ( $n = 63$ ; 1.80 mg/ml) had similar alpha-1 antitrypsin concentrations ( $P = 0.18$ ).

The allele frequencies for all missense and frameshift exonic *SERPINA1* variants are shown by ethnic group in Table 2, including 16 coding variants confirmed with array genotyping. PI Z, S, 21 additional rare coding variants, and one 5' UTR nucleotide insertion variant are shown as Figure 2. One variant in white individuals (Ile<sup>317</sup>Thr) had not been previously reported. White individuals had 17 coding variants with an allele frequency less than 0.05 and one frameshift insertion resulting in a null variant (Tyr<sup>184</sup>Terfs, PI type Q0<sub>Granite Falls</sub>), of which 15 were unique to white individuals. African Americans had eight variants with a frequency less than 0.05

(including the benign PI M1/M4), four unique to African Americans; whereas, in Hispanics, four, two unique, rare variants were identified. We identified 18 of 23 exonic coding SNPs identified in the NHLBI Exome Sequencing Project (see Table E4), plus seven additional rare coding variants and a frameshift insertion (28).

Seven PI Z-containing compound heterozygotes were identified among white individuals: three with PI type S (ZS genotype), two Ala<sup>308</sup>Ser, one PI type M6<sub>Passau</sub>, and one PI type I. Two PI SS homozygotes and three compound heterozygotes without PI Z were identified in white individuals: one PI S/Ile<sup>317</sup>Thr, one PI S/PI F, and one PI F/Ala<sup>308</sup>Ser. The baseline characteristics of compound heterozygotes and nine PI Z homozygotes

**Table 1.** Baseline Characteristics of the SPIROMICS Cohort

SPIROMICS Cohort Baseline Characteristics	Ethnic Group		
	Non-Hispanic White Individuals	African Americans	Hispanics
<b>Demographics</b>			
Subjects, <i>n</i>	1,693	385	90
Age, yr, mean (SD)*	65.2 (8.30)	58.2 (8.96)	60.8 (10.0)
Sex, F, <i>n</i> (%)*	751 (44.4)	207 (53.8)	34 (37.8)
BMI, mean (SD)	27.8 (5.06)	28.0 (6.06)	28.6 (5.29)
Pack-years, mean (SD)*	51.8 (26.4)	41.0 (17.4)	46.8 (25.8)
COPD diagnosis, <i>n</i> (%)*	1,191 (70.3)	206 (53.5)	52 (57.8)
COPD grade 2–4 diagnosis, <i>n</i> (%)*	921 (54.4)	167 (43.5)	36 (40.0)
<b>Post-bronchodilator pulmonary function</b>			
Baseline FEV <sub>1</sub> , % predicted, mean (SD)*	71.8 (25.8)	76.6 (27.1)	80.0 (28.0)
Baseline FVC, % predicted, mean (SD)*	90.3 (17.7)	92.5 (20.1)	94.8 (19.3)
FEV <sub>1</sub> /FVC ratio, %, mean (SD)*	0.59 (0.16)	0.64 (0.17)	0.64 (0.17)
<b>Clinical outcome measures</b>			
Hospitalization or ED visit last 12 mo, <i>n</i> (%)	163 (9.75)	80 (21.2)	16 (18.0)
St. George's Respiratory Questionnaire, mean (SD)	32.1 (20.1)	37.8 (20.9)	34.5 (21.9)
BODE, mean (SD)	1.51 (1.94)	1.57 (1.91)	1.08 (1.68)
6-minute-walk distance, m, mean (SD)*	408.5 (118.1)	410.6 (135.5)	408.9 (90.2)
COPD assessment test, mean (SD)	13.5 (8.10)	15.9 (8.88)	15.4 (8.82)
<b>Additional measures</b>			
Computed tomography evidence of emphysema, TLC % area < –950 HFU, mean (SD)*	8.19 (10.2)	7.70 (11.1)	4.89 (7.73)
Computed tomography evidence of air trapping, residual volume % area < –856 HFU, mean (SD)*	25.9 (21.0)	22.2 (23.2)	20.0 (21.0)
Computed tomography–based functional small airways disease, PRM <sup>fSAD</sup> , %, mean (SD)*	21.9 (15.0)	17.1 (15.7)	16.8 (16.3)
Alpha-1 antitrypsin concentration, mg/ml, mean (SD) <sup>†</sup>	1.98 (0.48)	1.95 (0.48)	1.80 (0.35)

*Definition of abbreviations:* BMI = body mass index; BODE = body mass index, airflow obstruction, dyspnea, and exercise; COPD = chronic obstructive pulmonary disease; ED = emergency department; HFU = Hounsfield units; PRM<sup>fSAD</sup> = parametric response mapping of functional small airway disease; SPIROMICS = Subpopulations and Intermediate Outcomes Measures in COPD Study.

\* $P < 0.05$  for the trend between ethnic groups.

<sup>†</sup>Serum alpha-1 antitrypsin levels were measured in a subgroup of 1,097 non-Hispanic white individuals, 231 African Americans, and 63 Hispanics.

are shown as Table E5 (haplotype phases are shown as Table E6). We identified 14 synonymous SNPs, 268 5' UTR variants, 103 intronic variants, and 122 3' UTR variants, including 25 insertion-deletions (see Table E7).

#### Effects of PI Type Z, S, and Additional Rare *SERPINA1* Variants on Lung Function and Emphysema Phenotypes

Based on gene-level burden testing, white individuals with any two rare exonic *SERPINA1* variants, including PI Z and S (two rare variants [2RVs];  $n = 21$ ), had lower FEV<sub>1</sub> (ORVs = 72.6%; 2RVs = 50.2% predicted;  $P = 3.58 \times 10^{-5}$ ), FEV<sub>1</sub>/FVC (ORVs = 0.59; 2RVs = 0.44;  $P = 3.52 \times 10^{-6}$ ), and FEF<sub>25-75</sub> (ORVs = 55.7%; 2RVs = 27.1% predicted;  $P = 1.69 \times 10^{-5}$ ), as well as greater computed tomography scan–based emphysema (ORVs = 7.83%; 2RVs = 17.2%  $\leq$  –950 HFU;  $P = 1.47 \times 10^{-5}$ ), air trapping (ORVs = 25.4%; 2RVs = 46.5%  $\leq$  –856 HFU;  $P = 2.89 \times 10^{-6}$ ),

and functional small airway disease (ORVs = 21.6%; 2RVs = 34.4% PRM<sup>fSAD</sup>;  $P = 2.34 \times 10^{-5}$ ; Table 3) compared with white individuals with ORVs ( $n = 1,411$ ). White individuals with 1RV ( $n = 259$ ) had a lower FEV<sub>1</sub>/FVC (1RV = 0.57;  $P = 0.005$ ), FEF<sub>25-75</sub> (1RV = 50.4% predicted;  $P = 0.004$ ), and greater emphysema (1RV = 9.43%  $\leq$  –950 HFU;  $P = 0.049$ ) compared with those with ORVs (see Table 3). Rare variant associations were not found in African Americans or Hispanics (see Table E8).

Allelic association testing by ethnic group is shown as Tables E9 to E14. As previously demonstrated in white individuals, the PI Z SNP (C/T alleles; Table 4) was associated with lower post-bronchodilator lung function and GOLD grade 2 or higher COPD when comparing major allele homozygotes (CC;  $n = 1,601$ ) and heterozygotes (CT;  $n = 81$ ) (15). Compared with major allele homozygotes, PI Z SNP heterozygotes also had lower FEF<sub>25-75</sub> (CC = 55.1%; CT = 47.3%;  $P = 0.01$ ), greater

computed tomography scan–based emphysema (CC = 7.89%; CT = 12.3%;  $P = 0.01$ ), air trapping (CC = 25.4%; CT = 33.6%;  $P = 0.01$ ), and functional small airway disease (CC = 21.6%; CT = 26.1%;  $P = 0.02$ ; see Table 4). Haplotype analyses excluded background haplotype effects (see Table E15). No significant lung function or computed tomography scan measure associations were found for any other exonic variant (see Tables E9–E14) or PI Z in African Americans or Hispanics (see Table E16).

#### Effects of PI Type Z, S, and Specific Rare *SERPINA1* Variant Genotypes on Lung Function and Emphysema Phenotypes

Owing to the significant co–primary outcome associations between white individuals with ORVs versus 1RV or two *SERPINA1* variants (see Table 3) and PI Z genotypes (see Table 4) but not PI S or other rare variants ( $V_R$ ; see Tables E9 and E12), we grouped white individuals into six genotype groups by

**Table 2.** *SERPINA1* Polymorphisms Identified in the Coding Exons in SPIROMICS by Ethnic Group

rs Number <sup>†</sup>	<i>SERPINA1</i> Variants*			Minor Allele Frequencies		
	cDNA Nucleotide Change	Coding Change	PI Type	Non-Hispanic White Individuals (n = 1,693)	African Americans (n = 385)	Hispanics (n = 90)
rs28931570 <sup>‡</sup>	c.187C>T	Arg <sup>63</sup> Cys	I	0.003	0.001	0
rs199687431 <sup>‡</sup>	c.206C>T	Ser <sup>69</sup> Phe	—	0.0003	0	0
rs111850950 <sup>‡</sup>	c.250G>A	Ala <sup>84</sup> Thr	M6 <sup>Passau</sup>	0.0003	0	0
rs376616935	c.286A>G	Thr <sup>96</sup> Ala	—	0	0.001	0
rs709932 <sup>‡</sup>	c.374G>A	Arg <sup>125</sup> His	M2/M4	0.16	0.04	0.12
rs112030253 <sup>‡</sup>	c.514G>T	Gly <sup>172</sup> Trp	M2 <sup>Obernburg</sup>	0.002	0	0
rs267606950	c.552delC	Tyr <sup>184</sup> Terfs	Q0 <sup>Granite Falls</sup>	0.0003	0	0
rs1206520317	c.632A>G	Tyr <sup>211</sup> Cys	—	0.0003	0	0
rs6647 <sup>‡</sup>	c.710T>C	Val <sup>237</sup> Ala	M1Ala/M1Val	0.22	0.56	0.24
rs28929470 <sup>‡</sup>	c.739C>T	Arg <sup>247</sup> Cys	F	0.004	0	0
rs121912714	c.839A>T	Asp <sup>280</sup> Val	P <sup>Lowell</sup> , Q0 <sup>Cardiff</sup>	0.0009	0	0
rs17580 <sup>‡</sup>	c.863A>T	Glu <sup>288</sup> Val	S	0.04	0.008	0.05
rs141095970 <sup>‡</sup>	c.879C>A	His <sup>293</sup> Gln	—	0	0.001	0
rs772436715	c.880C>T	Asp <sup>294</sup> Asn	—	0.0003	0	0
rs141620200 <sup>‡</sup>	c.922G>T	Ala <sup>308</sup> Ser	—	0.004	0	0
rs773222881	c.926G>A	Ser <sup>309</sup> Asn	—	0	0	0.006
pos 94845916	c.950A>G	Ile <sup>317</sup> Thr	—	0.0003	0	0
rs139964603 <sup>‡</sup>	c.976G>A	Val <sup>326</sup> Ile	—	0	0.003	0
rs377739083	c.1048C>T	Pro <sup>350</sup> Ser	—	0.0003	0	0
rs201788603 <sup>‡</sup>	c.1061C>T	Ser <sup>354</sup> Phe	S <sup>Munich</sup>	0.0009	0	0
rs201318727	c.1090A>G	Ile <sup>364</sup> Val	—	0.0003	0	0
rs143370956 <sup>‡</sup>	c.1093G>A	Asp <sup>365</sup> Asn	P <sup>St. Albans</sup>	0	0.001	0
rs201774333	c.1095C>G	Asp <sup>365</sup> Glu	—	0	0	0.006
rs28929474 <sup>‡</sup>	c.1096G>A	Glu <sup>366</sup> Lys	Z	0.03	0.005	0.03
rs61761869 <sup>‡</sup>	c.1177C>T	Pro <sup>393</sup> Ser	M <sup>Wurzburg</sup>	0.0009	0	0
rs1303 <sup>‡</sup>	c.1200A>C	Glu <sup>400</sup> Asp	M3	0.25	0.09	0.22
rs754885222	c.1211A>G	Lys <sup>404</sup> Arg	—	0.0003	0	0

*Definition of abbreviations:* cDNA = complementary DNA; PI = protease inhibitor; pos = position; SPIROMICS = Subpopulations and Intermediate Outcomes Measures in Chronic Obstructive Pulmonary Disease Study.

Minor allele frequencies of missense and frameshift polymorphisms identified through resequencing of a 16.9 kB region of *SERPINA1* are shown by ethnic group.

\*Variants are described by rs number, PI type based on protein isoelectric focusing, and amino acid coding change when indicated.

<sup>†</sup>Novel variation or those without an rs number are shown by nucleotide pos on chromosome 14.

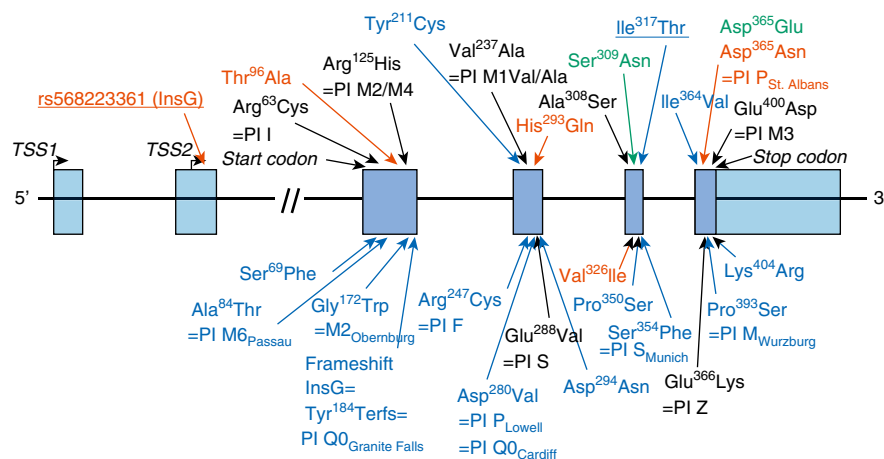
<sup>‡</sup>Twelve less common and rare coding variants (minor allele frequency < 0.05) and 3 common variants were genotyped with the Illumina OmniExpress HumanExome BeadChip (Illumina, Inc.) and PI S (not covered by the Chip) using a Sequenom Array.

collapsing PI Z-containing and non-Z variant (PI S or V<sub>R</sub>) genotypes (Table 5 and see Figure 1). These six genotype groups included 1) those without PI Z, S, or other, rare variants (“no Z/S/V<sub>R</sub>” or “those without rare variants”; n = 1,411); 2) heterozygotes for non-PI Z rare variants (“MS/MV<sub>R</sub>”; n = 132 MS; n = 53 MV<sub>R</sub>); 3) PI Z heterozygotes (CT SNP genotype) without another rare variant confirmed by sequencing (designated as “MZ”; n = 74); 4) PI S homozygotes, PI S and V<sub>R</sub> (non-PI Z) compound heterozygotes (“SS/SV<sub>R</sub>/V<sub>R</sub>V<sub>R</sub>”; n = 5); 5) PI Z-containing compound heterozygotes (“ZS/ZV<sub>R</sub>”; n = 7); and 6) PI Z homozygotes (TT SNP genotype designated as “ZZ”; n = 9) are shown as Figure 1 and Table 5. Stratifying between these genotype groups characterized the associations between ORVs and 1RV (see Table 3) by distinguishing effects of true MZ heterozygotes from Z-containing compound heterozygotes (ZS/ZV<sub>R</sub>) and non-Z

heterozygotes (MS/MV<sub>R</sub>; see Figure 1). For the associations between ORVs and 2RVs (see Table 3), this genotype stratification distinguished the known effects of ZZ homozygotes from the effects of additional rare variants in compound heterozygotes with (ZS/ZV<sub>R</sub>) and without PI Z (SS/SV<sub>R</sub>/V<sub>R</sub>V<sub>R</sub>; see Figure 1).

White MZ heterozygotes had a lower FEV<sub>1</sub> (64.3%; P = 0.007; Figure 3A) and FEV<sub>1</sub>/FVC (0.54; P = 0.003; see Figure 3B), as well as increased computed tomography scan-based emphysema (12.3%; P = 0.02; see Figure 3C) and air trapping (32.1%; P = 0.03), compared with those without rare variants (FEV<sub>1</sub> = 72.6%; FEV<sub>1</sub>/FVC = 0.60; emphysema = 7.83%; air trapping = 25.4%; see Table 5). Z-containing compound heterozygotes also had a lower FEV<sub>1</sub>/FVC (0.42; P = 0.02; see Figure 3B) and FEF<sub>25-75</sub> (20.1%; P = 0.009) compared with subjects without rare variants (see Table 5). Hence,

the association between PI Z SNP heterozygote (CT) genotype and lung function is due to both MZ and Z-containing compound heterozygotes (15). Non-PI Z rare variant heterozygotes (MS/MV<sub>R</sub>) also had a lower FEF<sub>25-75</sub> (50.6%) compared with subjects without rare variants (55.7%; P = 0.03; see Table 5). White ZZ homozygotes had lower lung function measures (FEV<sub>1</sub> = 37.9% [P = 2.11 × 10<sup>-6</sup>]; FEV<sub>1</sub>/FVC = 0.40 [P = 2.52 × 10<sup>-6</sup>]; FEF<sub>25-75</sub> = 16.1% [P = 0.0002]; see Figures 3A and 3B and Table 5) and greater computed tomography scan measures of emphysema (24.3%; P = 8.52 × 10<sup>-8</sup>; see Figure 3C), air trapping (54.2%; P = 1.15 × 10<sup>-8</sup>), and functional small airway disease (35.8%; P = 6.83 × 10<sup>-7</sup>) compared with subjects without rare variants (see Table 5). Genotype effects on emphysema and air trapping were determined by the lower lung zone emphysema in ZZ and MZ; however, upper



**Figure 2.** *SERPINA1* variants identified in the promoter and coding exons in the SPIROMICS multiethnic cohort. Missense and frameshift polymorphisms identified through resequencing of a 16.9 kB region of *SERPINA1* containing the untranslated promoter and five coding exons of *SERPINA1* in the SPIROMICS multiethnic cohort are shown by amino acid position, coding change, and protease inhibitor type where available. Of the 26 exonic coding variants identified with resequencing, three were the common and benign protease inhibitor types M1Ala/Val (Val<sup>237</sup>Ala, rs6647), M2/M4 (Arg<sup>125</sup>His, rs709932), and M3 (Glu<sup>400</sup>Asp, rs1303). A newly identified missense variant is underlined. Variants only found in white individuals are in blue text, those only in African Americans in red text (including an ethnic-specific G-nucleotide 5' untranslated region insertion [underlined] evaluated with functional studies), and those only in Hispanics in green text. SPIROMICS = Subpopulations and Intermediate Outcomes Measures in Chronic Obstructive Pulmonary Disease Study.

lung zone emphysema was also increased in ZZ (see Table E17). White SS/SV<sub>R</sub>/V<sub>R</sub>V<sub>R</sub> showed a spectrum of baseline disease risk and severity (see Table 5); however, tobacco smoke exposure varied across these five individuals (21.5–42.8 pack-years for two SS homozygotes, 63.0–80.0 pack-years for two SV<sub>R</sub>/one V<sub>R</sub>V<sub>R</sub>; see Table E5).

### Effects of PI Type Z, S, and Additional Rare *SERPINA1* Variants on COPD Risk

All white PI Z-containing compound heterozygotes ( $n=7$ ) and ZZ homozygotes ( $n=9$ ) had GOLD grade 1 to 4 COPD. Of these, six of seven Z-containing compound heterozygotes and 100% of ZZ homozygotes had grade 2 to 4 COPD. Grade 2 to 4 COPD was more frequent in MZ ( $n=49$  [66.2%]) compared with white individuals without rare variants ( $n=748$  [53.0%];  $P=0.02$ ; see Table 5). Rare *SERPINA1* genotypes were not significantly associated with COPD in African Americans or Hispanics (see Table E18).

### Effects of PI Type Z, S, and Additional Rare *SERPINA1* Variants on Longitudinal Lung Function

Although there was a significant difference in longitudinal decline of FEV<sub>1</sub> and FVC over 3 years in those with 2RVs compared

with those without rare variants (see Table 3), there were no strong genotype-specific associations (additive  $P>0.05$ ; see Tables 4 and 5). Based on burden testing, white individuals with any two of the rare variants, including PI Z and S ( $n=16$ ), had a higher rate of decline (FEV<sub>1</sub> =  $-0.14$  L/yr [ $P=0.001$ ]; FVC =  $-0.16$  L/yr [ $P=0.009$ ]; FEV<sub>1</sub>/FVC =  $-0.02$ /yr [ $P=0.02$ ]) compared with white individuals without rare variants ( $n=1,214$ ; FEV<sub>1</sub> =  $-0.04$  L/yr; FVC =  $-0.04$  L/yr; FEV<sub>1</sub>/FVC =  $-0.006$ /yr; see Table 3).

### Effects of PI Type Z, S, and Additional Rare *SERPINA1* Variants on Exploratory Clinical Outcomes

Body mass index, airflow obstruction, dyspnea, and exercise (BODE) index was higher in MZ (2.07;  $P=0.006$ ) and ZZ subjects (5.00;  $P=6.67 \times 10^{-7}$ ) compared with those without rare variants (1.46), whereas St. George's Respiratory Questionnaire scores were only higher in ZZ subjects (total score = 57.3) compared with those without rare variants (31.6;  $P=0.002$ ; see Table E19). COPD Assessment Test scores were higher in ZZ (22.1;  $P=0.007$ ) and Z-containing compound heterozygotes (20.6;  $P=0.04$ ) compared with those without rare variants (13.4; see Table E19). In white individuals, COPD exacerbations requiring emergency

department visits or hospitalization were more frequent in ZZ ( $n=4$  [44.4%]) compared with those without rare variants ( $n=131$  [9.38%];  $P=0.02$ ; see Table E19).

### Effects of Rare *SERPINA1* Variants on Alpha-1 Antitrypsin Concentrations

PI Z SNP genotypes (C/T) were associated with alpha-1 antitrypsin concentrations in white individuals ( $P=5.84 \times 10^{-46}$ ; see Table 4). Alpha-1 antitrypsin concentrations were lower in MS/MV<sub>R</sub> ( $n=113$ ; 1.79 mg/ml;  $P=1.23 \times 10^{-11}$ ), MZ ( $n=49$ ; 1.22 mg/ml;  $P=2.92 \times 10^{-37}$ ), compound heterozygotes without PI Z ( $n=2$ ; 1.20 mg/ml;  $P=0.005$ ), PI Z-containing compound heterozygotes ( $n=5$ ; 0.87 mg/ml;  $P=2.73 \times 10^{-10}$ ), and ZZ homozygotes ( $n=6$ ; 0.88 mg/ml;  $P=7.13 \times 10^{-10}$ ) compared with those without PI Z/S/V<sub>R</sub> ( $n=922$ ; 2.05 mg/ml; see Table 5). Detailed genotype analyses based on PI types Z, S, and additional rare variants demonstrate effects on alpha-1 antitrypsin concentrations, which mirror effects on lung function and emphysema (Figures 4A–4C). Rare variant burden testing excluding white subjects with either PI Z and S demonstrated lower alpha-1 antitrypsin concentrations in those with 1 of 14 different rare variants ( $n=38$ ; 1.76 mg/ml) compared with those without variants ( $n=922$ ; 2.05 mg/ml;  $P=8.09 \times 10^{-6}$ ; see Figure E1). African American ( $n=4$ ; 1.30 mg/ml) and Hispanic MZ heterozygotes ( $n=5$ ; 0.91 mg/ml) had lower alpha-1 antitrypsin concentrations compared with African Americans ( $n=368$ ; 1.96 mg/ml;  $P=0.03$ ) and Hispanics ( $n=74$ ; 1.82 mg/ml;  $P=0.02$ ) without rare variants (see Table E18).

### Alpha-1 Antitrypsin Quantitative Trait Loci Analyses

Association testing, 16.9 kB region-wide, for alpha-1 antitrypsin concentrations of *SERPINA1* conditional on PI Z did not identify loci within the 5' or 3' UTR reaching region-wide Bonferroni-adjusted significance ( $P=1.0 \times 10^{-4}$  based on 493 variants tested; see Figures E2 and E3). Five rare non-Z coding variants were at least nominally associated with alpha-1 antitrypsin concentrations in white individuals (Ile<sup>317</sup>Thr [coinherited with PI S in an individual compound heterozygote;  $P=0.05$ ; see Figures 4A–4C], PI I [ $P=0.0002$ ], PI Q<sub>0</sub>Granite Falls [ $P=0.01$ ], PI S [ $P=5.93 \times 10^{-9}$ ], PI M<sub>Wurzburg</sub> [ $P=0.046$ ]; see Figure E2 and Table E20), whereas no



**Table 3.** Effects of PI Types Z and S and Additional Rare *SERPINA1* Variants on Lung Function, Emphysema, and Related Phenotypes in Non-Hispanic White Individuals

	Number of RVs			ORVs vs. 1RV		1RV vs. 2RVs		ORVs vs. 2RVs	
	ORVs	1RV	2RVs	$\beta$ (SE)/OR (95% CI)	P Value	$\beta$ (SE)/OR (95% CI)	P Value	$\beta$ (SE)/OR (95% CI)	P Value
Total non-Hispanic white individuals, <i>n</i>	1,411	259	21	—	—	—	—	—	—
Post-BD lung function									
Post-BD, FEV <sub>1</sub> % predicted	72.6 (25.6)	69.1 (25.9)	50.2 (25.7)	-1.42 (0.82)	0.08	-8.71 (2.84)	0.002	-11.0 (2.66)	$3.58 \times 10^{-5}$
Post-BD, FVC % predicted	90.5 (17.7)	90.0 (17.8)	84.8 (20.4)	0.27 (0.55)	0.62	-3.14 (1.97)	0.11	-2.98 (1.80)	0.1
Post-BD, FEV <sub>1</sub> /FVC ratio	0.59 (0.16)	0.57 (0.16)	0.44 (0.16)	-0.01 (0.01)	0.005	-0.06 (0.02)	0.001	-0.08 (0.02)	$3.52 \times 10^{-6}$
Post-BD, FEF <sub>25-75</sub> % predicted	55.7 (40.2)	50.4 (38.8)	27.1 (31.2)	-0.03 (0.01)	0.004	-0.10 (0.03)	0.002	-0.14 (0.03)	$1.69 \times 10^{-5}$
GOLD grade 2 or higher, <i>n</i> (%)	748 (53.0%)	155 (59.8)	18 (85.7)	1.30 (0.98-1.74)	0.07	6.13 (1.99-26.8)	0.007	4.71 (1.48-20.9)	0.0009
3-yr longitudinal lung function, <i>n</i>	1,214	223	16	—	—	—	—	—	—
Annual rate of change, FEV <sub>1</sub> , L/yr	-0.04 (0.13)	-0.05 (0.15)	-0.14 (0.16)	-0.004 (0.005)	0.42	-0.05 (0.02)	0.01	-0.05 (0.02)	0.001
Annual rate of change, FVC, L/yr	-0.04 (0.23)	-0.04 (0.22)	-0.16 (0.34)	-0.001 (0.008)	0.93	-0.07 (0.03)	0.01	-0.07 (0.03)	0.009
Annual rate of change, FEV <sub>1</sub> /FVC per year	-0.006 (0.03)	-0.007 (0.03)	-0.02 (0.02)	-0.0005 (0.001)	0.6	-0.007 (0.004)	0.09	-0.008 (0.004)	0.02
Computed tomography scan-based measures									
Emphysema, TLC % area $\leq$ -950 HFU	7.83 (9.86)	9.43 (11.0)	17.2 (13.2)	0.04 (0.01)	0.049	0.19 (0.06)	0.002	0.26 (0.06)	$1.47 \times 10^{-5}$
Air trapping, residual volume % area $\leq$ -856 HFU	25.4 (20.8)	27.3 (21.3)	46.5 (21.7)	0.03 (0.01)	0.08	0.17 (0.04)	0.0005	0.23 (0.05)	$2.89 \times 10^{-6}$
Small airway disease, PRM <sup>SAD</sup> , %	21.6 (15.1)	22.2 (14.4)	34.4 (13.3)	0.02 (0.01)	0.16	0.13 (0.04)	0.002	0.17 (0.04)	$2.34 \times 10^{-5}$
Alpha-1 antitrypsin concentration, <i>n</i>	922	162	13	—	—	—	—	—	—
Alpha-1 antitrypsin concentration, mg/ml*	2.05 (0.45)	1.62 (0.42)	0.92 (0.52)	-0.22 (0.02)	$5.47 \times 10^{-33}$	-0.36 (0.06)	$1.73 \times 10^{-7}$	-0.56 (0.06)	$1.22 \times 10^{-19}$

*Definition of abbreviations:* BD = bronchodilator; CI = confidence interval; FEF<sub>25-75</sub> = forced expiratory flow, midexpiratory phase; GOLD = Global Initiative for Chronic Obstructive Lung Disease; HFU = Hounsfield units; OR = odds ratio; PI = protease inhibitor; PRM<sup>SAD</sup> = parametric response mapping of functional small airway disease; RV = rare variant. Subjects with ORVs, 1RV, and 2RVs, including PI types Z and S, were compared. Data are shown as mean (SD) unless otherwise indicated. Regression-based analyses of lung function are adjusted by study site, age, sex, body mass index, and pack-years smoking history. Log-transformed percentage predicted FEF<sub>25-75</sub> also adjusted for post-BD percentage predicted FVC and longitudinal, annualized rate of change for baseline lung function as a percentage of predicted. Computed tomography scan-based measures of emphysema ( $\leq$  -950 HFU or less [% bilateral lung area  $\leq$  -950]), air trapping ( $\leq$  -856 HFU or less [% bilateral lung area  $\leq$  -856]), and PRM<sup>SAD</sup> measures log-transformed for analysis and also adjusted for height. \*Alpha-1 antitrypsin concentrations also adjusted for C-reactive protein.

**Table 4.** Effects of *SERPINA1* PI Z Genotype on Lung Function, Emphysema, and Related Phenotypes in Non-Hispanic White Individuals

	PI Z (C/T) Genotype			CC vs. CT		CT vs. TT		CC vs. TT	
	CC	CT	TT	Additive P Value	$\beta$ (SE/OR) (95% CI)	P Value	$\beta$ (SE/OR) (95% CI)	P Value	$\beta$ (SE/OR) (95% CI)
Total non-Hispanic white individuals, <i>n</i>	1,601	81	9	—	—	—	—	—	—
Post-BD lung function	72.4 (25.5)	63.1 (27.3)	37.9 (15.3)	$2.50 \times 10^{-7}$	-4.25 (1.38)	0.002	-8.37 (5.42)	0.13	-19.1 (4.05)
Post-BD, FEV <sub>1</sub> % predicted	90.6 (17.7)	87.2 (17.3)	73.6 (20.7)	0.006	-1.20 (0.93)	0.2	-5.87 (3.65)	0.11	-8.13 (2.76)
Post-BD, FVC % predicted	0.59 (0.16)	0.53 (0.18)	0.40 (0.12)	$1.02 \times 10^{-7}$	-0.03 (0.01)	0.0006	-0.04 (0.04)	0.26	-0.12 (0.02)
Post-BD FEV <sub>1</sub> /FVC, ratio	55.1 (39.9)	47.3 (43.4)	16.0 (10.1)	$5.05 \times 10^{-5}$	-0.04 (0.02)	0.01	-0.06 (0.06)	0.38	-0.18 (0.05)
Post-BD, FEF <sub>25-75</sub> % predicted	857 (53.5)	55 (67.9)	9 (100)	$2.45 \times 10^{-5}$	1.86 (1.14-3.11)	0.01	9.07 (0.51-162)	0.05*	16.5 (0.96-284)
GOLD grade 2 or higher, <i>n</i> (%)	1,376	69	8	—	—	—	—	—	—
3-yr longitudinal lung function, <i>n</i>	—	—	—	—	—	—	—	—	—
Annual rate of change FEV <sub>1</sub> , L/yr	-0.04 (0.13)	-0.05 (0.13)	-0.11 (0.21)	0.15	-0.002 (0.008)	0.78	-0.05 (0.03)	0.11	-0.05 (0.02)
Annual rate of change FVC, L/yr	-0.04 (0.22)	-0.08 (0.26)	-0.07 (0.43)	0.13	-0.02 (0.01)	0.15	-0.09 (0.06)	0.14	-0.06 (0.04)
Rate of change FEV <sub>1</sub> /FVC per yr	-0.007 (0.03)	-0.003 (0.02)	-0.02 (0.01)	0.3	-0.001 (0.002)	0.46	-0.003 (0.005)	0.52	-0.007 (0.005)
Computed tomography scan-based measures	—	—	—	—	—	—	—	—	—
Emphysema, TLC % area $\leq$ -950 HFU	7.89 (9.87)	12.3 (12.7)	24.3 (13.8)	$6.34 \times 10^{-8}$	0.08 (0.03)	0.01	0.34 (0.12)	0.01	0.48 (0.09)
Air trapping, residual volume % area $\leq$ -856 HFU	25.4 (20.7)	33.6 (23.9)	54.2 (16.9)	$1.00 \times 10^{-8}$	0.06 (0.03)	0.01	0.28 (0.09)	0.004	0.42 (0.07)
Small airway disease, PRM <sup>tsAD</sup> , %	21.6 (14.9)	26.1 (15.9)	35.8 (8.3)	$5.84 \times 10^{-7}$	0.05 (0.02)	0.02	0.21 (0.07)	0.006	0.30 (0.06)
Alpha-1 antitrypsin concentration, <i>n</i>	1,037	54	6	—	—	—	—	—	—
Alpha-1 antitrypsin concentration, mg/ml <sup>†</sup>	2.02 (0.45)	1.19 (0.27)	0.88 (0.70)	$5.84 \times 10^{-46}$	-0.42 (0.03)	$3.84 \times 10^{-41}$	-0.20 (0.09)	0.03	-0.54 (0.09)

*Definition of abbreviations:* BD = bronchodilator; CI = confidence interval; FEF<sub>25-75</sub> = forced expiratory flow, midexpiratory phase; GOLD = Global Initiative for Chronic Obstructive Lung Disease; HFU = Hounsfield units; OR = odds ratio; PI = protease inhibitor; PRM<sup>tsAD</sup> = parametric response mapping of functional small airway disease. Homozygotes for the major allele (CC), heterozygotes for one minor or rare variant allele (CT), and homozygotes for the minor allele (TT) of PI type Z (Glu<sup>366</sup>Lys) were compared. Data are shown as mean (SD) unless otherwise indicated. Regression-based analyses were adjusted by study site, age, sex, body mass index, and pack-years smoking history for lung function. Log-transformed percentage predicted FEF<sub>25-75</sub> was also adjusted for post-BD percentage predicted FVC and longitudinal, annualized rate of change for baseline lung function as a percentage of predicted. Models also included height for log-transformed computed tomography scan-based emphysema based on percentage of the area in bilateral lungs with density of -950 Hounsfield units or less at TLC (TLC % lung area  $\leq$  -950), log-transformed computed tomography scan air trapping based on percentage of the area in bilateral lungs with density of -856 Hounsfield units or less at residual volume (residual volume % lung area  $\leq$  -856), and log-transformed functional small airway disease (PRM<sup>tsAD</sup>). Analysis of African Americans and Hispanics also adjusted for global African and Native American genetic ancestry.

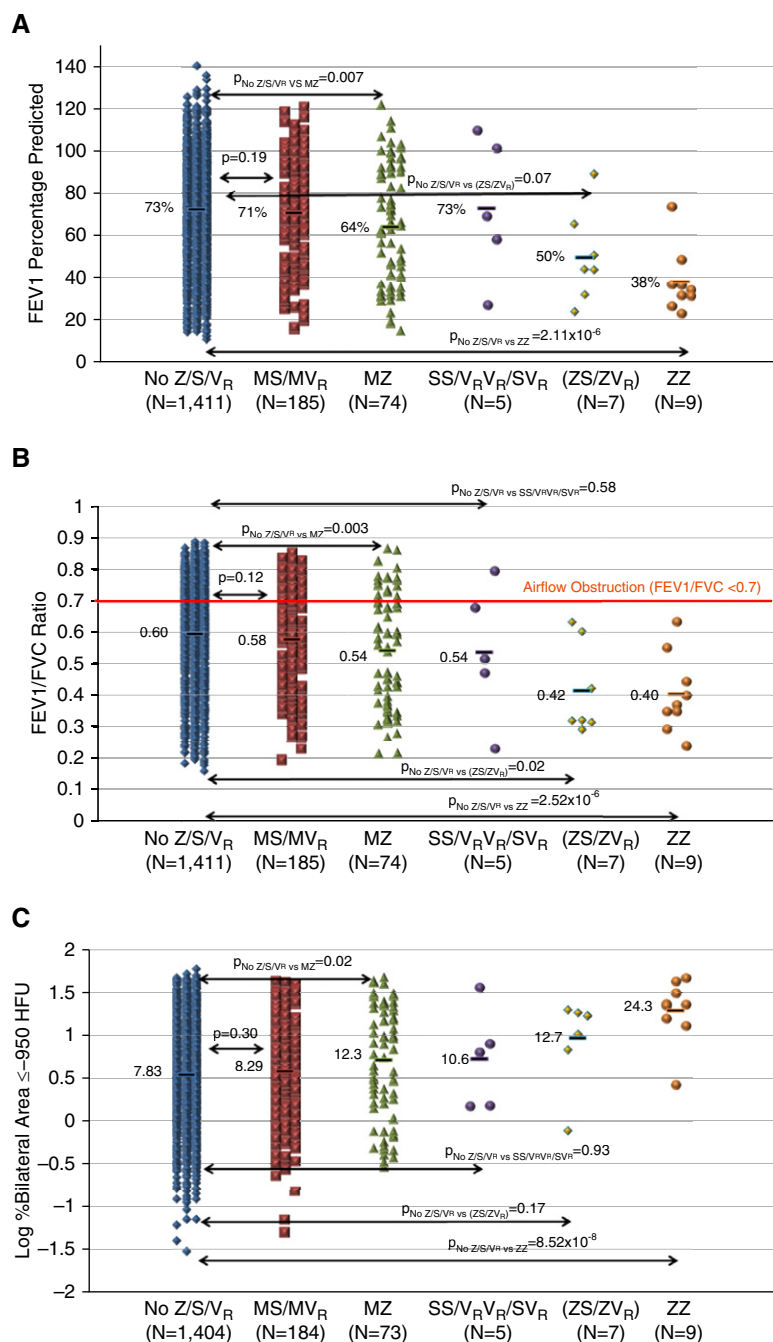
\*Owing to 100% of TT homozygotes with chronic obstructive pulmonary disease, the OR was calculated using the Haldane-Anscombe method, 95% CI using Woolf's method, and P value using the two-tail Fisher exact test.

<sup>†</sup>Alpha-1 antitrypsin concentrations also adjusted for C-reactive protein.

**Table 5.** Effects of Rare *SERPINA1* Variant Genotypes on Lung Function, Emphysema, and Related Phenotypes in Non-Hispanic White Individuals

	Rare <i>SERPINA1</i> Variant-based Genotypes							P Values				
	No Z/S <sub>R</sub>	MS/M <sub>V</sub>	MZ	SSV <sub>R</sub> V <sub>R</sub> /SV <sub>R</sub>	ZS/ZV <sub>R</sub>	ZZ	Additive trend	No Z/S <sub>R</sub> vs. MS/M <sub>V</sub>	No Z/S <sub>R</sub> vs. MZ	No Z/S <sub>R</sub> vs. SSV <sub>R</sub> V <sub>R</sub> /SV <sub>R</sub>	No Z/S <sub>R</sub> vs. ZS/ZV <sub>R</sub>	No Z/S <sub>R</sub> vs. ZZ
Total white individuals, <i>n</i>	1,411	185	74	5	7	9						
Post-BD lung function												
FEV <sub>1</sub> , % predicted	72.6 (25.6)	70.9 (25.0)	64.3 (27.6)	73.1 (33.6)	49.8 (21.9)	37.9 (15.3)	7.82 × 10 <sup>-6</sup>	0.64	0.007	0.97	0.07	2.11 × 10 <sup>-6</sup>
FVC, % predicted	90.5 (17.7)	91.2 (17.9)	87.0 (17.3)	98.6 (12.4)	89.2 (18.7)	73.6 (20.7)	0.03	0.19	0.22	0.47	0.92	0.003
FEV <sub>1</sub> /FVC, ratio	0.60 (0.16)	0.58 (0.16)	0.54 (0.18)	0.54 (0.21)	0.42 (0.14)	0.40 (0.12)	8.49 × 10 <sup>-7</sup>	0.12	0.003	0.58	0.02	2.52 × 10 <sup>-6</sup>
FEF <sub>25-75</sub> , % predicted	55.7 (40.2)	50.6 (36.3)	49.9 (44.5)	56.8 (54.3)	20.1 (13.1)	16.1 (10.1)	2.88 × 10 <sup>-5</sup>	0.03	0.05	0.49	0.009	0.0002
GOLD grade ≥ 2, <i>n</i> %	748 (53)	106 (57)	49 (66)	3 (60)	6 (86)	9 (100)	4.0 × 10 <sup>-4</sup>	0.41	0.02*	0.72	0.21	0.004†
3-yr longitudinal, <i>n</i>	1,214	158	65	4	4	8						
Rate of change, FEV <sub>1</sub> , L/yr	-0.04 (0.13)	-0.05 (0.15)	-0.04 (0.14)	-0.20 (0.10)	-0.12 (0.10)	-0.11 (0.21)	0.06	0.36	0.98	0.04	0.15	0.03
Rate of change, FVC, L/yr	-0.04 (0.22)	-0.03 (0.19)	-0.07 (0.26)	-0.24 (0.21)	-0.26 (0.22)	-0.07 (0.43)	0.14	0.65	0.42	0.22	0.04	0.18
Rate of change, FEV <sub>1</sub> /FVC	-0.006 (0.03)	-0.009 (0.03)	-0.003 (0.02)	-0.03 (0.03)	-0.008 (0.02)	-0.02 (0.01)	0.12	0.28	0.45	0.03	0.65	0.15
Computed tomography scan-based measures												
Emphysema, TLC% ≤ -950 HFU	7.83 (9.86)	8.29 (9.83)	12.3 (13.2)	10.6 (14.3)	12.7 (7.03)	24.3 (13.8)	1.84 × 10 <sup>-6</sup>	0.30	0.02	0.93	0.17	8.52 × 10 <sup>-8</sup>
Air trapping, residual volume area% ≤ -856 HFU	25.4 (20.8)	25.4 (19.9)	32.1 (23.4)	29.6 (27.1)	48.7 (19.1)	54.2 (16.9)	3.56 × 10 <sup>-7</sup>	0.41	0.03	0.9	0.12	1.15 × 10 <sup>-8</sup>
Small airways disease, PRM <sup>ISAD</sup>	21.6 (15.1)	21.2 (13.9)	24.6 (15.4)	22.4 (14.5)	41.1 (13.6)	35.8 (8.29)	1.30 × 10 <sup>-5</sup>	0.59	0.05	0.88	0.07	6.83 × 10 <sup>-7</sup>
Alpha-1 antitrypsin, <i>n</i>	922	113	49	2	5	6						
Alpha-1 antitrypsin, mg/ml <sup>†</sup>	2.05 (0.45)	1.79 (0.37)	1.22 (0.24)	1.20 (0)	0.87 (0.39)	0.88 (0.70)	7.56 × 10 <sup>-56</sup>	1.23 × 10 <sup>-11</sup>	2.92 × 10 <sup>-37</sup>	0.005	2.73 × 10 <sup>-10</sup>	7.13 × 10 <sup>-10</sup>

*Definition of abbreviations:* BD = bronchodilator; FEF<sub>25-75</sub> = forced expiratory flow, midexpiratory phase; GOLD = Global Initiative for Chronic Obstructive Lung Disease; HFU = Hounsfield units; PRM<sup>ISAD</sup> = parametric response mapping of functional small airway disease. Regression-based association testing compared white individuals without PIs (protease inhibitors) Z, S, or other rare variants (No Z/S<sub>R</sub>), heterozygotes for non-PI Z rare variants (MS and M<sub>V</sub><sub>R</sub> heterozygotes), PI Z heterozygotes without another rare variant (MZ), PI S and V<sub>R</sub> compound heterozygotes (SS, SV<sub>R</sub>, and V<sub>R</sub>V<sub>R</sub>), PI Z-containing compound heterozygotes with PI S or additional rare variants (ZS/ZV<sub>R</sub>), and PI Z homozygotes (ZZ). White compound heterozygotes with PI type Z and S or additional rare variants (ZS/ZV<sub>R</sub>) include three PI type ZS, one PI type Z/PI M<sub>V</sub><sub>R</sub>SS<sub>U</sub>, one PI type Z/PI I, and two PI type Z/Ala<sup>308</sup>Ser. White compound heterozygotes without PI type Z (SS, SV<sub>R</sub>, and V<sub>R</sub>V<sub>R</sub>) included two PI type S homozygotes (SS), one PI type S/Ile<sup>171</sup>Thr, one PI type S/PI type F, and one PI type F/Ala<sup>308</sup>Ser. Regression-based analyses were adjusted by study site, age, sex, body mass index, and pack-years smoking history. Log-transformed percentage predicted FEF<sub>25-75</sub> also adjusted for post-BD FVC% predicted and longitudinal, annualized rate of change for baseline lung function as a percentage of predicted. Models also included height for log-transformed computed tomography scan-based emphysema based on percentage of the area in bilateral lungs with density of -950 HFU or less at TLC (TLC % lung area ≤ -950 HFU), log-transformed computed tomography scan air trapping based on percentage of the area in bilateral lungs with density of -856 HFU or less at residual volume (residual volume % lung area ≤ -856 HFU), and log-transformed functional small airway disease (PRM<sup>ISAD</sup>). Data are shown as mean (SD) unless otherwise indicated. \*Owing to 100% of ZZ homozygotes having chronic obstructive pulmonary disease, the odds ratio was calculated using the Haldane-Anscombe method, 95% confidence interval using Woolf's method, and *P* value using the two-tailed Fisher Exact test; odds ratio = 16.82; 95% confidence interval = 0.98-298.5; *P* = 0.004. †Alpha-1 antitrypsin concentrations also adjusted for C-reactive protein.



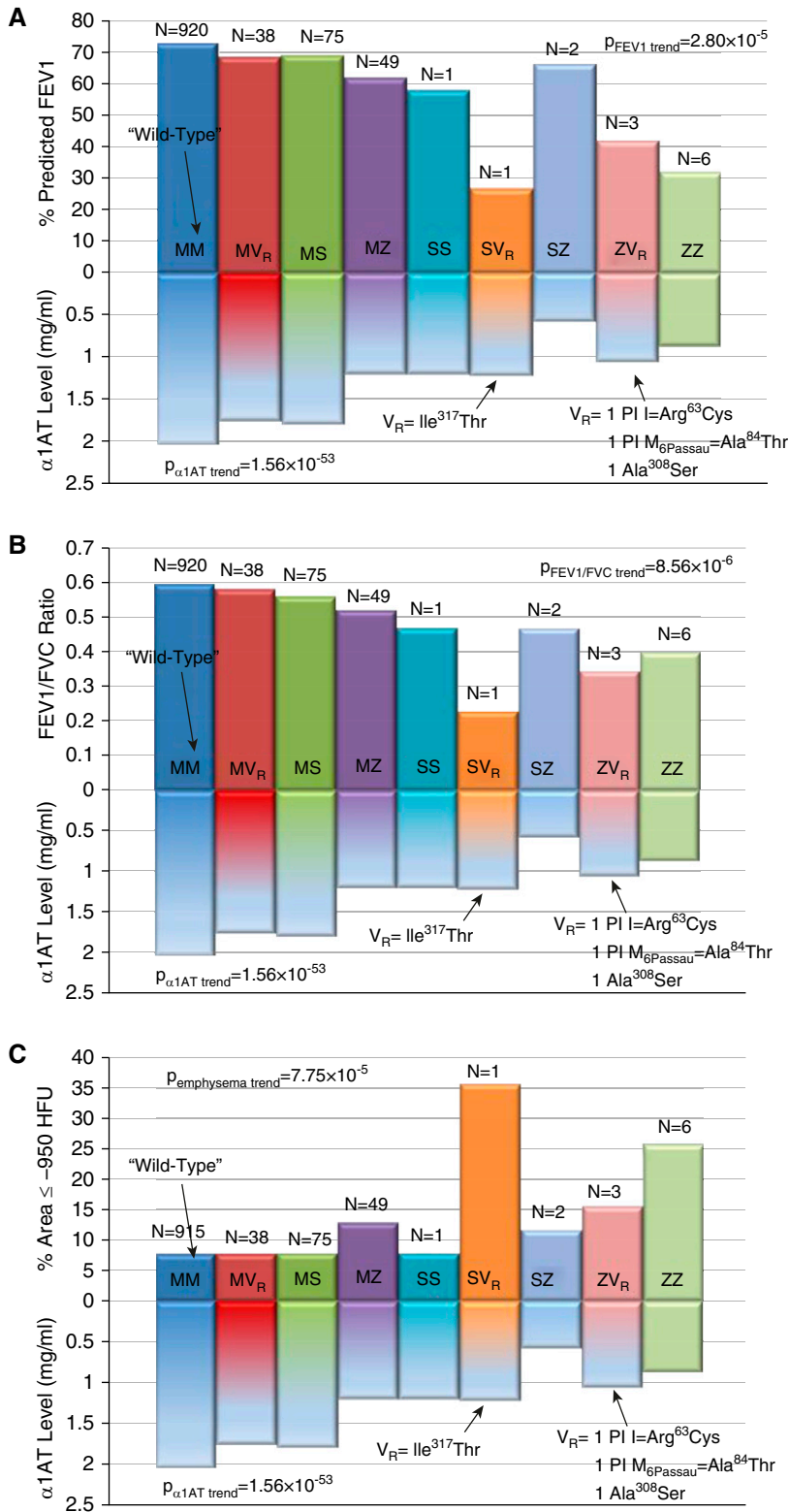
**Figure 3.** Effects of rare *SERPINA1* variant-containing genotypes on lung function and computed tomography scan-based emphysema. Regression-based association testing compared white individuals without PIs (protease inhibitors) Z, S, or other rare variants (No Z/S/V<sub>R</sub>); heterozygotes for non-PI Z rare variants (MS [*n* = 132] and MV<sub>R</sub> [*n* = 53] heterozygotes); PI Z heterozygotes without another rare variant (MZ); PI S and V<sub>R</sub> compound heterozygotes (SS, SV<sub>R</sub>, and V<sub>R</sub>V<sub>R</sub>); PI Z-containing compound heterozygotes with PI S or additional rare variants (ZS/ZV<sub>R</sub>); and PI Z homozygotes (ZZ). Compound heterozygotes with PI type Z and S or additional rare variants include three PI type ZS, one PI type Z/PI M<sub>6</sub>P<sub>Passau</sub>, one PI type Z/PI I, and two PI type Z/Ala<sup>308</sup>Ser. Compound heterozygotes without PI type Z included two PI type S homozygotes (SS), one PI type S/Ile<sup>317</sup>Thr, one PI type S/PI type F, and one PI type F/Ala<sup>308</sup>Ser. Bars represent the means. Regression-based analyses adjusted by study site, age, sex, body mass index, and pack-years smoking history for lung function ([A] FEV<sub>1</sub>% predicted and [B] FEV<sub>1</sub>/FVC ratio). Models also included height for log-transformed computed tomography scan-based emphysema based on percentage of the area in bilateral lungs with density of -950 Hounsfield units (HFU) or less ([C] log % bilateral area ≤ -950 HFU).

exonic rare variants were associated in African Americans (see Table E21) or Hispanics (see Table E22). In the entire white cohort with sequencing data, we collapsed associated variants (PI I, Q<sub>0</sub>Granite Falls M<sub>W</sub>urzburg) with rare variants showing with mean concentrations ≤ 1.5 mg/ml in carriers, the lowest 10th percentile, (PI S<sub>Munich</sub> [*n* = 3; 1.5 mg/ml], PI Q<sub>0</sub>Cardiff or P<sub>Lowell</sub> [*n* = 1; 1.3 mg/ml]; see Table E20) to test for associations in the absence of PI S and PI Z genotypes. White heterozygotes with any one of these five rare variants had increased computed tomography scan-based emphysema (*n* = 19; 9.92%) compared with those without rare variants (*n* = 1,409; 7.83%; *P* = 0.02; Figure 5 and see Table E23).

Alpha-1 antitrypsin quantitative trait loci analysis in African Americans conditional on PI Z identified a G-insertion in the 5' UTR (-5,368 bps relative to start codon, rs568223361; see Figure 2) associated with lower alpha-1 antitrypsin concentrations with similar significance as PI Z (*P* = 0.02; see Figure E3). Five African Americans with this insertion had greater functional small airway disease (PRM<sup>SAD</sup> = 43%) compared with those without (17%; *P* = 0.007; see Figure E4 and Table E24). Luciferase reporter experiments indicate that this insertion lowers alpha-1 antitrypsin expression (see Figures E5A and E5B) through either altered transcriptional or translational regulation (33).

## Discussion

This study is the largest, deep targeted resequencing analysis of *SERPINA1* variation integrated with alpha-1 antitrypsin measurements in a comprehensively characterized multiethnic cohort of subjects with COPD and unaffected cigarette smokers not ascertained for alpha-1 antitrypsin deficiency. Targeted DNA sequencing provides deeper coverage compared with whole-genome or whole-exome sequencing for the identification of rare variants. In addition, the heavy minimal smoking history of ≥ 20 pack-years in the SPIROMICS cohort provides an opportunity to detect pathogenic variation not possible in cohorts without sufficient exposure because gene-environment interactions are a necessary component for the penetrance of the emphysema and COPD induced by alpha-1 antitrypsin deficiency (34).

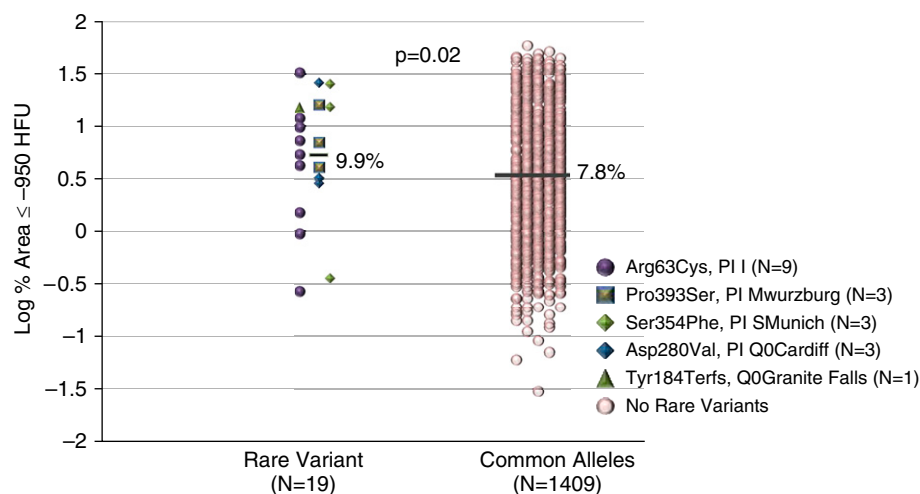


**Figure 4.** Rare *SERPINA1* variant genotypes and alpha-1 antitrypsin levels in non-Hispanic white individuals and effects on lung function and computed tomography scan-based emphysema. Detailed genotypes in white individuals are shown as those without a rare *SERPINA1* variant (MM), those with one rare variant ( $V_R$ ) that is not PI (protease inhibitor) type S or Z (MV<sub>R</sub>), those with one PI type S allele and no other variant (MS), PI type Z heterozygotes (MZ), PI type S homozygotes (SS),

Thus, this study addresses long-standing questions regarding the role of PI Z genotypes in the pathogenesis of COPD and emphysema while providing insight into novel rare *SERPINA1* genotype associations in ever-smokers. First, this study confirms that PI Z heterozygotes with a significant cigarette smoking history have lower lung function, greater airflow obstruction, and greater computed tomography scan-based quantitative measures of emphysema, air trapping, and functional small airway disease. The question of the effects of MZ heterozygosity on COPD risk and severity is an important issue that has been evaluated over the past 40 years in mostly general populations primarily limited by small sample sizes, physician-based diagnoses, and a failure to establish sufficient exposure to tobacco smoke resulting in variable results (9, 10, 12, 35). Irish MZ heterozygotes based on PIEF with  $\geq 20$  pack-years smoking history had a lower lung function and an increased COPD risk compared with MM ever-smokers, whereas no differences were observed between those with  $< 20$  pack-years (11). In contrast to our study, these subjects were not ascertained for COPD and, on average, had normal lung function.

Recent genotyping studies of COPD and controls with  $\geq 10$  pack-years smoking history also found lower lung function and higher risk for COPD, computed tomography scan-based emphysema, and air trapping in PI Z heterozygotes compared with those without PI S or Z, including what we described in a genome-wide association study of the same population (13–15). In conjunction with prior studies, we confirm the effects of the MZ heterozygote genotype on COPD and emphysema phenotypes in current and ex-smokers. In contrast to prior studies, we assessed background variation with deep sequencing of *SERPINA1* (10–15). Hence, we were able to demonstrate that the effects of PI Z heterozygote CT genotype (see Table 4) are determined by the combined effects of true MZ and PI Z-containing compound heterozygotes on alpha-1 antitrypsin concentrations, lung function, measures of small airway disease, and emphysema (see Table 5) (13–15, 36).

Second, we demonstrated that non-PI Z rare *SERPINA1* variants and PI S are significantly associated with COPD phenotypes in Z-containing compound heterozygotes. Prior studies have shown that ZZ homozygotes and ZS compound



**Figure 5.** Effects on non-PI (protease inhibitor) type S or Z rare *SERPINA1* variants associated with alpha-1 antitrypsin concentrations on computed tomography scan-based emphysema in non-Hispanic white individuals. Regression-based analyses were adjusted by study site, age, sex, pack-years smoking history, body mass index, and height. Bars represent the means, red dots on the right column represents subjects without rare *SERPINA1* variants (common, major allele homozygotes), and the dots of multiple colors on the right column represent different heterozygotes for one of five different rare *SERPINA1* variants associated with lower alpha-1 antitrypsin concentrations ( $P=0.05$ ) or with mean concentrations  $\leq 1.5$  mg/ml, excluding PI type S and Z. HFU = Hounsfield units.

heterozygotes are at an increased risk for airflow obstruction and emphysema; however, few studies have evaluated Z-containing compound heterozygotes with additional variation (7, 8). In our study, white Z-containing compound heterozygotes with four different rare variants had increased airflow obstruction, impaired small airway function, and lower alpha-1 antitrypsin concentrations (see Table 5) (10, 37). Pathogenic Z-containing compound heterozygotes with a diversity of rare variants, including Q0<sub>Granite Falls</sub>, PI I, and P<sub>Lowell</sub>, have been associated with alpha-1 antitrypsin deficiency and early-onset severe emphysema in case reports (38–41).

Third, our study is the first to show that multiple rare *SERPINA1* variant genotypes, independent of PI Z, impact alpha-1 antitrypsin deficiency and could serve as a predictor of emphysema even among heavy smokers with a single variant. We found a cumulative effect of rare variation on

alpha-1 antitrypsin deficiency, independent of PI Z and S (see Figure E1), similar to that described in a general population (42). Of these, we identified five non-S/Z variants associated with lower alpha-1 antitrypsin concentrations and cumulatively associated with computed tomography scan-based emphysema in white heterozygotes (see Figure 5) (10). White individuals with two rare *SERPINA1* variants, including PI S homozygotes and compound heterozygotes without PI Z, also had a higher rate of lung function decline (see Tables 3 and 5), demonstrating the diversity of genotypes that influence disease progression beyond PI Z genotypes (43, 44). The largest number of rare *SERPINA1* variants and phenotype associations were found in white individuals because of the smaller sample size of African Americans and Hispanics, which also resulted in a potential lack of power to detect associations in these minority racial groups. Despite this

limitation, this is the largest *SERPINA1* sequencing study in a minority cohort of ex-smokers and COPD with alpha-1 antitrypsin concentrations that identified a rare 5' UTR variant unique to African Americans associated with functional small airway disease, a putative precursor to emphysema, and alterations in alpha-1 antitrypsin expression (37). The clinical impact of this G-insertion needs to be confirmed in larger racially diverse cohorts.

We tested for associations between *SERPINA1* variation and multiple phenotypes, resulting in multiple tests. The strong effects of *SERPINA1* variation in rare compound heterozygotes demonstrated gene-level variant interactions that justified a tiered approach starting with burden testing (see Figure 1) to minimize multiple comparisons. In addition, the three co-primary outcomes are interrelated (see Table E2), resulting in association tests that are not independent. Another limitation is that sequencing does not distinguish compound heterozygotes, including singleton variants versus two variants occurring on the same chromosome. The low alpha-1 antitrypsin concentrations in compound heterozygotes provide supportive evidence that we identified true compound heterozygotes (see Tables 3 and 5, Table E5, and Figures 4A–4C). Finally, the analysis of small samples sizes of rare genotype subgroups is a major statistical challenge inherent to the analysis of rare variants, which could have weak effects, no effects, or opposing beneficial effects not detected with traditional association tests or collapsing methods. For instance, whereas Z-containing compound heterozygotes showed statistical trends similar to ZZ homozygotes, non-Z compound heterozygotes and SS homozygotes showed a spectrum of baseline disease risk and severity (see Tables 5 and E5), which could relate to differences in individual variant effect sizes or tobacco smoke exposure, and requires study in larger cohorts. The

**Figure 4.** (Continued). compound heterozygotes with PI type S except PI type Z (SV<sub>R</sub>), PI type S and Z-containing compound heterozygotes (SZ), Z-containing compound heterozygotes with additional rare variants except PI type S (ZV<sub>R</sub>), and PI type Z homozygotes (ZZ), as shown. MV<sub>R</sub> heterozygotes included those with Lys<sup>404</sup>Arg ( $n=1$ ), Pro<sup>393</sup>Ser (PI M<sub>wurzburg</sub>;  $n=3$ ), Ile<sup>364</sup>Val ( $n=1$ ), Ser<sup>354</sup>Phe (PI S<sub>Munich</sub>;  $n=2$ ), Pro<sup>350</sup>Ser ( $n=1$ ), Ala<sup>308</sup>Ser ( $n=7$ ), Asp<sup>280</sup>Val (PI P<sub>Lowell</sub>, Q0<sub>Cardiff</sub>;  $n=1$ ), Arg<sup>247</sup>Cys (PI F;  $n=8$ ), Tyr<sup>211</sup>Cys ( $n=1$ ), Tyr<sup>184</sup>Terfs (PI Q0<sub>Granite Falls</sub>;  $n=1$ ), Gly<sup>172</sup>Trp (PI M2<sub>Obernburg</sub>;  $n=5$ ), Ser<sup>69</sup>Phe ( $n=1$ ), and Arg<sup>63</sup>Ser (PI I;  $n=6$ ). ZV<sub>R</sub> compound heterozygotes included one PI Z/PI M<sub>6Passau</sub>, one PI Z/PI I, and one PI Z/Ala<sup>308</sup>Ser. SV<sub>R</sub> compound heterozygotes included one PI S/Ile<sup>317</sup>Thr. Regression-based analyses were adjusted by study site, age, sex, body mass index, and pack-years smoking history for lung function ([A] FEV<sub>1</sub>% and [B] FEV<sub>1</sub>/FVC ratio). Models also included height for (C) log-transformed computed tomography scan-based emphysema based on percentage of the area in bilateral lungs with density of  $\leq -950$  Hounsfield units (HFU) or less (% area  $\leq -950$  HFU).  $\alpha$ 1AT = alpha-1 antitrypsin.

challenge of variable effects was also important for non-Z/S rare variation for which we leveraged alpha-1 antitrypsin quantitative trait loci analyses to identify five pathogenic variants cumulatively associated with emphysema in white individuals and a rare insertion associated with functional small airway disease in African Americans.

Using targeted, deep DNA sequencing and comprehensive phenotyping, we have demonstrated the potential for sequencing as a complementary test to the measurement of alpha-1 antitrypsin concentrations for the identification of novel variant determinants for COPD and emphysema risk and progression (19). We used next-generation sequencing to detect novel and known rare variation, including insertions or frameshift variation, and ethnic-specific variants that would not have been found in SNP-based genotyping platforms primarily based on European descent populations. Sequencing distinguishes null variants (Q0<sub>Granite Falls</sub>, Q0<sub>Cardiff</sub>) that falsely appear as homozygote MM or ZZ genotypes on PIEF, as well as similar PI types that have varying pathogenic effects (normal M, M<sub>Wurzburg</sub>, M6<sub>Passau</sub>, P<sub>St. Albans</sub>, P<sub>Lowell</sub>, S, S<sub>Munich</sub>). DNA sequencing definitively demonstrated that MZ heterozygotes with the confirmed absence of background rare variation and Z-containing compound heterozygotes associated with COPD and emphysema. We also used alpha-1 antitrypsin concentrations in conjunction with computed tomography scan-based

measures and *in vitro* functional studies to identify novel, at-risk non-Z heterozygote subgroups. These findings need to be evaluated in larger cohorts with sequencing but justifies the sequencing of *SERPINA1* in a subgroup of at-risk smokers with COPD, emphysema, and low alpha-1 antitrypsin concentrations in the absence of two pathogenic SNPs covered by chip genotyping panels. Commercial testing kits already perform resequencing in those with alpha-1 antitrypsin deficiency if *SERPINA1* panel genotyping is negative (45). A similar, step-wise approach for *CFTR* resequencing is standard for individuals with clinical features of cystic fibrosis and elevated sweat chloride concentrations despite negative *CFTR* panel genotyping (46). *CFTR* sequencing has identified over 2,000 variants of which 346 variants are known to cause cystic fibrosis and are critical to the selection of targeted modulator therapies (47, 48).

Based on the frequency of Z-containing compound heterozygotes, MZ heterozygotes, and additional rare variant heterozygote genotypes associated with emphysema (see Table E23 and Figure 5) in white individuals from SPIROMICS (6.4%), the impact of *SERPINA1* variation in the smoking general population could be more than ninefold greater when compared with the prevailing hypothesis (ZZ/ZS combined frequency = 0.7%). Based on the cumulative deleterious effects of multiple rare *SERPINA1* genotypes, we should no longer

consider *SERPINA1* variation a rare cause of COPD and emphysema but as a common and important pathogenic determinant of disease risk and severity. As the costs of next-generation sequencing and high-throughput genotyping decline, the identification and study of diverse *SERPINA1* variation could lead to the development of comprehensive genotyping panels for the identification of a larger number of at-risk genotype subgroups. This would allow early targeted smoking cessation or other targeted, precision approaches for the individualized prevention and treatment of COPD and emphysema related to alpha-1 antitrypsin deficiency. ■

**Author disclosures** are available with the text of this article at [www.atsjournals.org](http://www.atsjournals.org).

**Acknowledgment:** The authors thank the SPIROMICS participants and participating physicians, investigators, and staff for making this research possible. They also thank the following current and former investigators of the SPIROMICS sites and reading centers: Neil E. Alexis, Wayne H. Anderson, Mehrdad Arjomandi, Lori A. Bateman, Surya P. Bhatt, Richard C. Boucher, Russell P. Bowler, Stephanie A. Christenson, Alejandro P. Comellas, Gerard J. Criner, Ronald G. Crystal, Jeffrey L. Curtis, Claire M. Doerschuk, Mark T. Dransfield, Christine M. Freeman, Craig Galban, Nadia N. Hansel, Annette T. Hastie, Yvonne Huang, Robert J. Kaner, Jerry A. Krishnan, Lisa M. LaVange, Stephen C. Lazarus, Wendy C. Moore, Laura Paulin, Nirupama Putcha, Elizabeth C. Oelsner, Sanjeev Raman, Donald P. Tashkin, J. Michael Wells, and Robert A. Wise.

## References

- Zaimidou S, van Baal S, Smith TD, Mitropoulos K, Ljujic M, Radojkovic D, et al. A1ATVar: a relational database of human SERPINA1 gene variants leading to alpha1-antitrypsin deficiency and application of the VariVis software. *Hum Mutat* 2009;30:308–313.
- Silverman EK, Sandhaus RA. Clinical practice: alpha1-antitrypsin deficiency. *N Engl J Med* 2009;360:2749–2757.
- Stoller JK, Aboussouan LS. Alpha1-antitrypsin deficiency. *Lancet* 2005;365:2225–2236.
- Yoshida A, Lieberman J, Gaidulis L, Ewing C. Molecular abnormality of human alpha1-antitrypsin variant (Pi-ZZ) associated with plasma activity deficiency. *Proc Natl Acad Sci USA* 1976;73:1324–1328.
- Brantly M, Nukiwa T, Crystal RG. Molecular basis of alpha-1-antitrypsin deficiency. *Am J Med* 1988;84:13–31.
- Eriksson S. Pulmonary emphysema and alpha1-antitrypsin deficiency. *Acta Med Scand* 1964;175:197–205.
- Green CE, Vayalappa S, Hampson JA, Mukherjee D, Stockley RA, Turner AM. PiSZ alpha-1 antitrypsin deficiency (AATD): pulmonary phenotype and prognosis relative to PiZZ AATD and PiMM COPD. *Thorax* 2015;70:939–945.
- Turino GM, Barker AF, Brantly ML, Cohen AB, Connelly RP, Crystal RG, et al. Clinical features of individuals with PiSZ phenotype of alpha 1-antitrypsin deficiency: alpha 1-antitrypsin deficiency registry study group. *Am J Respir Crit Care Med* 1996;154:1718–1725.
- Hersh CP, Dahl M, Ly NP, Berkey CS, Nordestgaard BG, Silverman EK. Chronic obstructive pulmonary disease in alpha1-antitrypsin Pi MZ heterozygotes: a meta-analysis. *Thorax* 2004;59:843–849.
- Thun GA, Ferrarotti I, Imboden M, Rochat T, Gerbase M, Kronenberg F, et al. SERPINA1 PiZ and PiS heterozygotes and lung function decline in the SAPALDIA cohort. *PLoS One* 2012;7:e42728.
- Molloy K, Hersh CP, Morris VB, Carroll TP, O'Connor CA, Lasky-Su JA, et al. Clarification of the risk of chronic obstructive pulmonary disease in alpha1-antitrypsin deficiency PiMZ heterozygotes. *Am J Respir Crit Care Med* 2014;189:419–427.
- Sørheim IC, Bakke P, Gulsvik A, Pillai SG, Johannessen A, Gaarder PI, et al. alpha1-Antitrypsin protease inhibitor MZ heterozygosity is associated with airflow obstruction in two large cohorts. *Chest* 2010;138:1125–1132.
- Hobbs BD, Parker MM, Chen H, Lao T, Hardin M, Qiao D, et al.; NETT Genetics Investigators; ECLIPSE Investigators; COPDGene Investigators; International COPD Genetics Network Investigators. Exome array analysis identifies a common variant in IL27 associated with chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2016;194:48–57.
- Foreman MG, Wilson C, DeMeo DL, Hersh CP, Beaty TH, Cho MH, et al.; Genetic Epidemiology of COPD (COPDGene) Investigators. Alpha-1 antitrypsin PiMZ genotype is associated with chronic obstructive pulmonary disease in two racial groups. *Ann Am Thorac Soc* 2017;14:1280–1287.

15. Li X, Ortega VE, Ampleford EJ, Graham Barr R, Christenson SA, Cooper CB, *et al.*; SPIROMICS Research Group. Genome-wide association study of lung function and clinical implication in heavy smokers. *BMC Med Genet* 2018;19:134.
16. Busch R, Hobbs BD, Zhou J, Castaldi PJ, McGeachie MJ, Hardin ME, *et al.*; National Emphysema Treatment Trial Genetics; Evaluation of COPD Longitudinally to Identify Predictive Surrogate End-Points; International COPD Genetics Network; COPDGene Investigators. Genetic association and risk scores in a chronic obstructive pulmonary disease meta-analysis of 16,707 subjects. *Am J Respir Cell Mol Biol* 2017;57:35–46.
17. Cho MH, McDonald ML, Zhou X, Mattheisen M, Castaldi PJ, Hersh CP, *et al.*; NETT Genetics, ICGN, ECLIPSE and COPDGene Investigators. Risk loci for chronic obstructive pulmonary disease: a genome-wide association study and meta-analysis. *Lancet Respir Med* 2014;2: 214–225.
18. Kong X, Cho MH, Anderson W, Coxson HO, Muller N, Washko G, *et al.*; ECLIPSE Study NETT Investigators. Genome-wide association study identifies BICD1 as a susceptibility gene for emphysema. *Am J Respir Crit Care Med* 2011;183:43–49.
19. Hersh CP, Campbell EJ, Scott LR, Raby BA. Alpha-1 antitrypsin deficiency as an incidental finding in clinical genetic testing. *Am J Respir Crit Care Med* 2019;199:246–248.
20. Ortega VE, Li X, O'Neal WK, Hawkins GA, Pennington E, Barr RG, *et al.* Comprehensive gene resequencing of SERPINA1 in the SPIROMICS multi-ethnic population [abstract]. *Am J Respir Crit Care Med* 2017; 195:A7244.
21. Ortega VE, Hawkins GA, Li X, O'Neal WK, Manichaikul A, Barr RG, *et al.* SPIROMICS. Comprehensive gene resequencing of SERPINA1 in SPIROMICS reveals novel rare loci for  $\alpha$ 1-antitrypsin deficiency and emphysema [abstract]. *Am J Respir Crit Care Med* 2018;197:A4184.
22. Couper D, LaVange LM, Han M, Barr RG, Bleecker E, Hoffman EA, *et al.*; SPIROMICS Research Group. Design of the subpopulations and intermediate outcomes in COPD study (SPIROMICS). *Thorax* 2014;69:491–494.
23. Vogelmeier CF, Criner GJ, Martinez FJ, Anzueto A, Barnes PJ, Bourbeau J, *et al.* Global strategy for the diagnosis, management, and prevention of chronic obstructive lung disease 2017 report: GOLD executive summary. *Am J Respir Crit Care Med* 2017;195:557–582.
24. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 2007;81: 559–575.
25. Landrum MJ, Lee JM, Benson M, Brown GR, Chao C, Chitipiralla S, *et al.* ClinVar: improving access to variant interpretations and supporting evidence. *Nucleic Acids Res* 2018;46:D1062–D1067.
26. Sudmant PH, Rausch T, Gardner EJ, Handsaker RE, Abyzov A, Huddleston J, *et al.*; 1000 Genomes Project Consortium. An integrated map of structural variation in 2,504 human genomes. *Nature* 2015;526:75–81.
27. Lek M, Karczewski KJ, Minikel EV, Samocha KE, Banks E, Fennell T, *et al.*; Exome Aggregation Consortium. Analysis of protein-coding genetic variation in 60,706 humans. *Nature* 2016;536:285–291.
28. Exome Variant Server. NHLBI GO Exome Sequencing Project (ESP) [accessed 2019 Jan 1]. Available from: <http://evs.gs.washington.edu/EVS/>.
29. Online Mendelian Inheritance in Man. OMIM [accessed 2019 Jun 1]. Available from: <https://omim.org/>.
30. Morris AP, Zeggini E. An evaluation of statistical approaches to rare variant analysis in genetic association studies. *Genet Epidemiol* 2010;34:188–193.
31. Galbán CJ, Han MK, Boes JL, Chughtai KA, Meyer CR, Johnson TD, *et al.* Computed tomography-based biomarker provides unique signature for diagnosis of COPD phenotypes and disease progression. *Nat Med* 2012;18:1711–1715.
32. Alexander DH, Novembre J, Lange K. Fast model-based estimation of ancestry in unrelated individuals. *Genome Res* 2009;19:1655–1664.
33. Corley M, Solem A, Phillips G, Lackey L, Ziehr B, Vincent HA, *et al.* An RNA structure-mediated, posttranscriptional model of human  $\alpha$ -1-antitrypsin expression. *Proc Natl Acad Sci USA* 2017;114: E10244–E10253.
34. Castaldi PJ, DeMeo DL, Kent DM, Campbell EJ, Barker AF, Brantly ML, *et al.* Development of predictive models for airflow obstruction in alpha-1-antitrypsin deficiency. *Am J Epidemiol* 2009;170:1005–1013.
35. Seersholm N, Wilcke JT, Kok-Jensen A, Dirksen A. Risk of hospital admission for obstructive pulmonary disease in alpha(1)-antitrypsin heterozygotes of phenotype PiMz. *Am J Respir Crit Care Med* 2000; 161:81–84.
36. Boes JL, Hoff BA, Bule M, Johnson TD, Rehemtulla A, Chamberlain R, *et al.* Parametric response mapping monitors temporal changes on lung CT scans in the subpopulations and intermediate outcome measures in COPD Study (SPIROMICS). *Acad Radiol* 2015;22:186–194.
37. Stockley JA, Ismail AM, Hughes SM, Edgar R, Stockley RA, Sapey E. Maximal mid-expiratory flow detects early lung disease in  $\alpha$ 1-antitrypsin deficiency. *Eur Respir J* 2017;49:1602055.
38. Faber JP, Weidinger S, Goedde HW, Ole K. The deficient alpha-1-antitrypsin phenotype Pi P is associated with an A-to-T transversion in exon III of the gene. *Am J Hum Genet* 1989;45:161–163.
39. Crystal RG. The alpha 1-antitrypsin gene and its deficiency states. *Trends Genet* 1989;5:411–417.
40. Seri M, Magi B, Cellesi C, Olia PM, Renieri A, De Marchi M. Molecular characterization of the P and I variants of alpha 1-antitrypsin. *Int J Clin Lab Res* 1992;22:119–121.
41. Nukiwa T, Takahashi H, Brantly M, Courtney M, Crystal RG. alpha 1-Antitrypsin null Granite Falls, a nonexpressing alpha 1-antitrypsin gene associated with a frameshift to stop mutation in a coding exon. *J Biol Chem* 1987;262:11999–12004.
42. Thun GA, Imboden M, Ferrarotti I, Kumar A, Obeidat M, Zorzetto M, *et al.* Causal and synthetic associations of variants in the SERPINA gene cluster with alpha1-antitrypsin serum levels. *PLoS Genet* 2013; 9:e1003585.
43. Dahl M, Tybjaerg-Hansen A, Lange P, Vestbo J, Nordestgaard BG. Change in lung function and morbidity from chronic obstructive pulmonary disease in alpha1-antitrypsin MZ heterozygotes: a longitudinal study of the general population. *Ann Intern Med* 2002; 136:270–279.
44. Sandford AJ, Chagani T, Weir TD, Connett JE, Anthonisen NR, Paré PD. Susceptibility genes for rapid decline of lung function in the lung health study. *Am J Respir Crit Care Med* 2001;163:469–473.
45. Foil KE, Blanton MG, Sanders C, Kim J, Al Ashry HS, Kumbhare S, *et al.* Sequencing alpha-1 MZ individuals shows frequent biallelic mutations. *Pulm Med* 2018;2018:2836389.
46. Dequeker E, Stuhmann M, Morris MA, Casals T, Castellani C, Claustres M, *et al.* Best practice guidelines for molecular genetic diagnosis of cystic fibrosis and CFTR-related disorders--updated European recommendations. *Eur J Hum Genet* 2009;17:51–65.
47. Clinical and Functional Translation of CFTR. CFTR2 [2019 Oct 10]. Available from: <https://cftr2.org/>.
48. Keating D, Marigowda G, Burr L, Daines C, Mall MA, McKone EF, *et al.*; VX16-445-001 Study Group. VX-445-tezacaftor-ivacaftor in patients with cystic fibrosis and one or two Phe508del alleles. *N Engl J Med* 2018;379:1612–1620.