CFTR variants are associated with chronic bronchitis in smokers

Aabida Saferali[1,2], Dandi Qiao[1,2], Wonji Kim[1,2], Karen Raraigh[3], Hara Levy[4], Alejandro A. Diaz[2,5], Garry R. Cutting[3], Michael H. Cho[1,2,5] and Craig P. Hersh[1,2,5] on behalf of NHLBI TransOmics in Precision Medicine (TOPMed)

[1]Channing Division of Network Medicine, Brigham and Women's Hospital, Boston, MA, USA. [2]Harvard Medical School, Boston, MA, USA. [3]Johns Hopkins University School of Medicine, Baltimore, MD, USA. [4]Division of Pulmonary Medicine, Dept of Pediatrics, National Jewish Health, Denver, CO, USA. [5]Division of Pulmonary and Critical Care Medicine, Brigham and Women's Hospital, Boston, MA, USA.

Corresponding author: Craig P. Hersh (craig.hersh@channing.harvard.edu)

Shareable abstract (@ERSpublications)
Cigarette smokers who carry one deleterious *CFTR* variant have higher rates of chronic bronchitis, while presence of two *CFTR* variants associates with COPD. These results indicate that genetically mediated reduction in CFTR function contributes to COPD. https://bit.ly/3GSWUXw

Cite this article as: Saferali A, Qiao D, Kim W, *et al*. *CFTR* variants are associated with chronic bronchitis in smokers. *Eur Respir J* 2022; 60: 2101994 [DOI: 10.1183/13993003.01994-2021].

This article has an editorial commentary: https://doi.org/10.1183/13993003.00898-2022

Received: 3 Aug 2021
Accepted: 14 Dec 2021

Abstract

Introduction Loss-of-function variants in both copies of the cystic fibrosis transmembrane conductance regulator (*CFTR*) gene cause cystic fibrosis (CF); however, there is evidence that reduction in CFTR function due to the presence of one deleterious variant can have clinical consequences. Here, we hypothesise that *CFTR* variants in individuals with a history of smoking are associated with chronic obstructive pulmonary disease (COPD) and related phenotypes.

Methods Whole-genome sequencing was performed through the National Heart, Lung, and Blood Institute TOPMed (TransOmics in Precision Medicine) programme in 8597 subjects from the COPDGene (Genetic Epidemiology of COPD) study, an observational study of current and former smokers. We extracted clinically annotated *CFTR* variants and performed single-variant and variant-set testing for COPD and related phenotypes. Replication was performed in 2118 subjects from the ECLIPSE (Evaluation of COPD Longitudinally to Identify Predictive Surrogate Endpoints) study.

Results We identified 301 coding variants within the *CFTR* gene boundary: 147 of these have been reported in individuals with CF, including 36 CF-causing variants. We found that CF-causing variants were associated with chronic bronchitis in variant-set testing in COPDGene (one-sided p=0.0025; OR 1.53) and in meta-analysis of COPDGene and ECLIPSE (one-sided p=0.0060; OR 1.52). Single-variant testing revealed that the F508del variant was associated with chronic bronchitis in COPDGene (one-sided p=0.015; OR 1.47). In addition, we identified 32 subjects with two or more *CFTR* variants on separate alleles and these subjects were enriched for COPD cases (p=0.010).

Conclusions Cigarette smokers who carry one deleterious *CFTR* variant have higher rates of chronic bronchitis, while presence of two *CFTR* variants may be associated with COPD. These results indicate that genetically mediated reduction in CFTR function contributes to COPD related phenotypes, in particular chronic bronchitis.

Introduction

Chronic obstructive pulmonary disease (COPD) is a complex disease typically caused by cigarette smoke and influenced by genetic factors. COPD is phenotypically heterogeneous with varying manifestations of emphysema, chronic bronchitis, airway wall thickening and bronchiectasis despite similar degrees of lung function impairment. This variability likely reflects the contribution of multiple pathological mechanisms. Chronic bronchitis is a particularly problematic phenotype in COPD as it is associated with pulmonary exacerbations and has few treatment options [1, 2]. Since chronic bronchitis shares some clinical and pathological features with cystic fibrosis (CF), it has been proposed that there may be common mechanisms involved.

CF is the most common lethal autosomal recessive disorder in populations of European descent and one in 35 Americans is a carrier of a loss-of-function variant in the CF transmembrane conductance regulator

(*CFTR*) gene. In addition to CF, several disorders have been associated with variants in *CFTR*, such as idiopathic pancreatitis [3, 4], congenital bilateral absence of the vas deferens [5] and allergic bronchopulmonary aspergillosis [6]. Furthermore, there is evidence that cigarette smoking can lead to acquired CFTR dysfunction [7–10]. Cigarette smokers and COPD patients have reduced function of CFTR in the upper and lower airways in addition to chronic bronchitis. CFTR dysfunction has been shown to reduce airway surface liquid and decrease mucociliary transport [7, 10, 11]. Therefore, it is possible that acquired CFTR dysfunction through cigarette smoking may contribute to COPD and this effect may be compounded by genetic variation in *CFTR*.

CFTR potentiators are a new class of CF medications which function by directly correcting underlying gating defects in mutant CFTR [7]. *In vitro* studies have demonstrated that the CFTR potentiator ivacaftor can improve CFTR protein function in epithelial cells exposed to cigarette smoke and this is reflected in measures of epithelial function, including mucociliary transport, airway surface liquid depth and ciliary beating [7, 12]. In addition, a pilot study of ivacaftor in patients with COPD and chronic bronchitis demonstrated the potential for increased CFTR activity and respiratory symptoms [13]. Furthermore, there is evidence that the CFTR potentiator icenticaftor can increase forced expiratory volume in 1 s (FEV_1) as well as reduce systemic inflammation and sputum colonisation in COPD patients [14]. Collectively, these data indicate that improvement of CFTR function using existing drugs could improve lung function in COPD patients. However, the question remains as to which patients would most benefit from this treatment.

While several small studies have investigated the association of *CFTR* variants with the deleterious effects of cigarette smoke on CFTR function, results have been mixed [15–22]. Other larger studies have been limited by including nonsmokers in addition to smokers [23, 24]. To address this question with greater power, a large sample size of smokers with and without COPD along with *CFTR* gene sequencing data is required to ascertain whether *CFTR* variants, together with cigarette smoke, contribute to reduced lung function in smokers with COPD. Here, we perform the largest investigation of *CFTR* variants in COPD to date, including subjects with whole-genome sequencing (WGS) data from two large cohorts, to test the hypothesis that deleterious variants in *CFTR* are associated with COPD and related phenotypes.

Methods

Study populations

The COPDGene (Genetic Epidemiology of COPD) and ECLIPSE (Evaluation of COPD Longitudinally to Identify Predictive Surrogate Endpoints) studies have been described previously [25, 26]. Briefly, COPDGene enrolled 10 192 non-Hispanic White and African American subjects with a minimum of 10 pack-years lifetime smoking history. Subjects with diagnosed lung diseases other than COPD or asthma were excluded. The ECLIPSE study is a multicentre multinational 3-year longitudinal study that enrolled 3291 subjects of Global Initiative for Chronic Obstructive Lung Disease (GOLD) stage 2–4. In COPDGene, COPD was defined by a post-bronchodilator ratio of FEV_1 to forced vital capacity <0.7 (GOLD stage 1–4); severe COPD was defined as GOLD stages 3–4. In ECLIPSE, only subjects with GOLD stage 2–4 were included. Chronic bronchitis was defined using the classical definition of self-reported chronic cough and phlegm for ⩾3 months per year over the past 2 years. Bronchodilator response was defined as the percentage change in pre-/post-bronchodilator FEV_1. Visual scoring of bronchiectasis was performed using computed tomography (CT) scans for 1372 COPDGene subjects with WGS data [27]. Subjects who were found to have diffuse bronchiectasis on chest CT scan were excluded from COPDGene.

Institutional review boards approved the studies at all participating institutions and all participants provided written informed consent per study protocols.

Whole-genome sequencing

WGS data were generated through the National Heart, Lung, and Blood Institute TOPMed (TransOmics in Precision Medicine) consortium to a mean depth of 30× using DNA from blood, PCR-free library construction and Illumina HiSeq X technology [28]. For COPDGene, Freeze 5b WGS data were used which included 8598 subjects including 5773 non-Hispanic White and 2825 African American. For replication in ECLIPSE, Freeze 8 WGS data were used which included 2345 subjects; a subset of 2212 was included in this analysis. Reads were mapped to human genome assembly version GRCh38 and computational phasing was performed using Eagle 2.4 (13 December 2017; https://github.com/poruloh/Eagle).

Identification and annotation of CFTR *variants*

All variants within the *CFTR* gene boundary (chr7:117 465 784–117 715 971; GRCh38) were extracted from WGS data using bcftools [29]. The WGS annotator pipeline [30] was used to characterise all

variants. Coding variants were identified as variants classified as in-frame deletion, frameshift, missense, splice acceptor, splice donor, splice region, stop gained or synonymous variants according to the Ensembl VEP (Variant Effect Predictor) consequence (www.ensembl.org). Annotation of known CF-causing variants was downloaded from the CFTR2 consortium website (https://cftr2.org) (accessed on 4 May 2021). These variants are categorised as CF-causing, varying clinical significance, non-CF-causing and unknown significance. For variants that were not reported in the CFTR2 database, SnpEff (https://pcingola.github.io/SnpEff) functional effect predictions were used to identify variants with likely functional impact. Phased sequencing data from subjects with two or more known CF-causing variants were visually inspected to determine whether these subjects are compound heterozygotes with pathogenic variants on both chromosomes. These subjects are of interest as loss of function of both copies of *CFTR* would likely have a greater clinical consequence. We hypothesised *a priori* that heterozygous *CFTR* variants would have a deleterious effect in smokers due to a decrease in CFTR function; therefore, we expected that the minor allele (*i.e.* less common allele) of *CFTR* variants would be associated with increased chronic bronchitis, increased risk of severe COPD, increased risk of severe exacerbations, decreased body mass index, decreased FEV_1 % pred, decreased percent emphysema and increased airway wall thickness. We used one-sided p-values for these tests, while we used two-sided p-values for associations with bronchodilator response (as a percentage of predicted FEV_1) as we did not have a prediction regarding direction of effect.

Single-variant association testing

The workflow for genetic variant testing is described in figure 1. Testing of each individual variant for phenotype association was performed using linear regression for quantitative traits and logistic regression for binary outcomes using R base functions. For single-variant testing, only variants with a minor allele

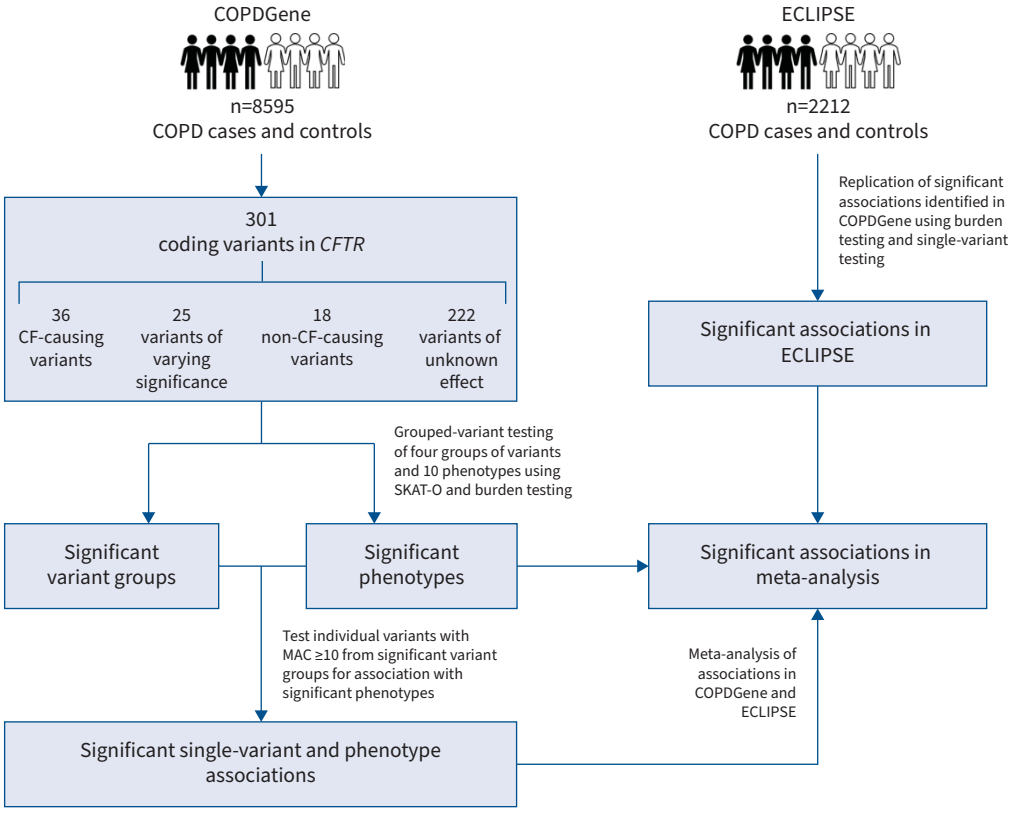

FIGURE 1 Workflow for cystic fibrosis transmembrane conductance regulator (*CFTR*) genetic variant testing. COPDGene served as the discovery cohort and significant findings were replicated in ECLIPSE. Four groups of variants were tested for association with 10 phenotypes in COPDGene using the (Sequence) Kernel Association Test (SKAT)-O and burden testing. Only variant groups and phenotypes with significant associations in grouped-variant testing were included in single-variant testing. Furthermore, single-variant testing was only performed for variants with a minor allele count (MAC) ⩾10. COPD: chronic obstructive pulmonary disease; CF: cystic fibrosis.

count ⩾10 were included. Analyses were adjusted for age, sex, pack-years of smoking, current smoking status and principal components (PCs) of genetic ancestry. Calculation of PCs has been previously described [31, 32]. Analyses in COPDGene were performed in non-Hispanic White and African American individuals combined, using three PCs of genetic ancestry. For ECLIPSE, 10 PCs of genetic ancestry were used. For each single-variant analysis, we also performed permutation analysis by permuting the variant/nonvariant carrier status among all subjects 20 000 times, then computing the p-value using the number of permutations in which the test statistic was more extreme than the observed test statistic. As described earlier, we used one-sided p-values for association testing with all phenotypes except bronchodilator response.

Gene-based association testing

Gene-based testing of rare variants (<5% minor allele frequency) was performed using burden tests in which we collapsed rare *CFTR* variants into a single burden variable and tested for association with phenotype using linear and logistic regression. In addition, we used the single nucleotide polymorphism (SNP)-set (Sequence) Kernel Association Test (SKAT)-O [33] as an additional method for gene-based association testing. All *CFTR* variants were tested in a combined analysis, in addition to testing subsets of variants grouped according to known pathogenicity using annotations from the CFTR2 database. SKAT-O tests were performed both with weighting by percent pancreatic insufficiency (obtained from the CFTR2 consortium website) as a measure of variant severity and with no weighting.

Results

Identification of CFTR *variants in COPDGene participants*

After quality control measures [28], a total 8595 subjects including 3848 COPD cases and 4691 smoking controls were available for analysis (table 1). In these subjects, we identified 11 567 variants within the gene boundary of *CFTR* (figure 2) as defined by Ensembl (chr7:117 465 784–117 715 971), which includes 14 241 bp upstream and 47 306 bp downstream of the coding region of transcript NM_000492.3. Of these variants, 10 577 are single nucleotide variants and 990 are insertion/deletion (indel) polymorphisms. Of these, there were 301 variants that are located within the coding region of the RefSeq Select transcript (NM_000492.3) (supplementary table S1). Using the CFTR2 database, we found that 147 variants have been reported in CF patients; 36 are CF-causing variants, 25 are variants of varying clinical consequence (may cause CF in some individuals but not others), 18 are non-CF-causing variants (may cause CFTR dysfunction but not sufficient to cause CF) and 68 variants have not been evaluated or are of unknown significance. Four variants with high minor allele frequency (>0.05) were excluded from further analysis; three of these are synonymous variants (rs1800136 (legacy 4521G/A), rs1800130 (P1290P) and chr7:117 595 001:T:G), while rs213950 (lV470M) is a missense variant known to be non-CF-causing. After these, the most frequent variants were chr7:117 509 093:G:A (R75Q) with 459 counts and chr7:117 559 655:G:A (1716G/A) with 290 counts, both of which are non-CF-causing missense variants. We additionally identified 177 subjects heterozygous for the common p.Phe508del (legacy F508del) variant (rs199826652). We discovered 154 variants that have not been previously described in the CFTR2 database, including one stop-gain and 89 missense variants which are predicted to have moderate-to-high impact on CFTR through SnpEff functional impact prediction.

TABLE 1 Description of study subjects in COPDGene and ECLIPSE

	COPDGene		ECLIPSE	
	COPD cases	Smoking controls	COPD cases	Smoking controls
Subjects, n[#]	3848	4691	1953	165
Male	56.31	51.18	34.46	43.64
Age, years	62.91±8.68	56.72±8.42	63.36±7.12	56.30±9.63
Race				
Non-Hispanic White	77.36	59.24	98.16	96.36
African American	22.64	40.76		
Current smoker	55.85	39.91	61.90	63.03
Smoking history, pack-years	51.62±27.40	38.44±21.29	48.94±27.44	30.02±20.30

Data are presented as % or mean±SD, unless otherwise stated. COPD: chronic obstructive pulmonary disease. [#]: in COPDGene, 56 subjects were not classified as COPD cases or controls, including 55 subjects that failed spirometry and one never-smoker; in ECLIPSE, 94 subjects were not classified as COPD cases or controls including 49 subjects that failed spirometry, 16 subjects that are GOLD stage 1 and 29 subjects with preserved ratio impaired spirometry.

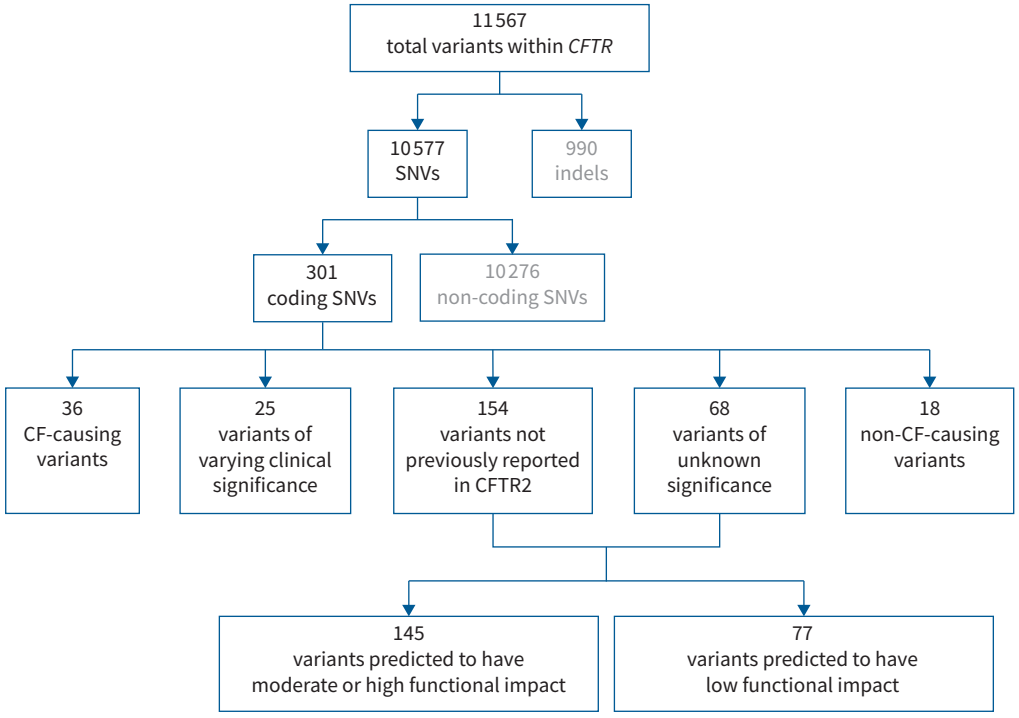

FIGURE 2 Breakdown of cystic fibrosis transmembrane conductance regulator (*CFTR*) variants. A total of 301 coding single nucleotide variants (SNVs) were included in the analysis. CF: cystic fibrosis.

Variant-set testing for association with COPD and related phenotypes

Variants were grouped according to pathogenicity. Four groupings were tested: 1) CF-causing variants, 2) CF-causing variants and variants of varying clinical consequence, 3) CF-causing, varying clinical consequence and variants that have not been reported that in the CFTR2 database that may have a functional effect (moderate or high impact in SnpEff), and 4) all coding variants. The only association that reached the threshold for significance after correction for multiple comparison ($p<0.05/10$ or 0.005) was the association of CF-causing variants with chronic bronchitis: 68 out of 248 subjects with CF-causing variants had chronic bronchitis (27.4%) while 1597 out of 8345 subjects without CF-causing variants had chronic bronchitis (19.1%) ($p=0.0025$; OR 1.53) (table 2). We hypothesised that variants associated with a larger percentage of patients having pancreatic insufficiency reflected a greater impact of the variant on CFTR function. Therefore, SKAT-O variant-set testing was performed with and without weighting for percentage pancreatic insufficiency as a measure of variant severity. This analysis confirmed that CF-causing variants are associated with chronic bronchitis, although there was no difference in the weighted and unweighted analysis, and the associations were not significant after correction for multiple comparisons (supplementary table S2). Since we hypothesised that the combination of cigarette smoke and heterozygous *CFTR* variants would result in greater reduction of CFTR function, we performed a stratified analysis of current *versus* former smokers, where we found that 39.5% of currently smoking subjects with CF-causing variants had chronic bronchitis compared with 23.9% of currently smoking subjects without *CFTR* variants; however, the p-value did not reach the stringent threshold for significance after correction for multiple comparison ($p=0.0082$; OR 1.62) (supplementary table S3). In contrast, in former smokers we found that 17.2% of subjects with CF-causing variants had chronic bronchitis compared with 13.5% of subjects without *CFTR* variants ($p=0.082$). We additionally found that in an analysis of COPD cases alone, there was a significant enrichment of chronic bronchitis in subjects with CF-causing variants (38.1%) compared with subjects without *CFTR* variants (25.5%) ($p=0.0022$; OR 1.72) (supplementary table S4). Finally, we found an association of borderline significance between all coding variants and severe COPD ($p=0.0063$; OR 1.14) (table 2). Bronchiectasis was visually scored using CT scans for 1372 subjects; however, there was no association between the presence of bronchiectasis and *CFTR* variants (table 2).

Single-variant testing for association with COPD and related phenotypes

For phenotypes in which there was a significant association using variant-set testing, we performed single-variant testing for all variants within the group with minor allele count $\geqslant 10$. This resulted in one

TABLE 2 Burden testing in COPDGene

	CF-causing	CF-causing+varying clinical consequence	CF-causing+varying clinical consequence+predicted functional	All coding variants	Controls
Variants	36	61	206	297[#]	
Subjects[¶]	248	455	732	2309	6281
Chronic bronchitis					
Cases	68 (27.4)	109 (24.0)	169 (23.1)	463 (20.1)	1202 (19.1)
Controls	180	346	563	1844	5079
p-value	0.0025* (OR 1.53)	0.033 (OR 1.19)	0.0089 (OR 1.20)	0.41	
COPD					
Cases	134 (54.0)	213 (47.0)	337 (46.4)	1095 (47.8)	2751 (44.1)
Controls	114	240	389	1197	3489
p-value	0.039* (OR 1.28)	0.28	0.084	0.021 (OR 1.09)	
Severe COPD					
Cases	50 (20.2)	81 (17.9)	131 (18.0)	437 (19.1)	1031 (16.5)
Controls	198	372	595	1855	5209
p-value	0.26	0.29	0.072	0.0063 (OR 1.14)	
Severe exacerbations					
Yes	24 (9.7)	45 (9.9)	79 (10.8)	264 (11.4)	761 (12.1)
No	224	410	653	2043	5520
p-value	0.19	0.20	0.19	0.26	
BMI, kg·m^{-2}					
Mean±SD	29.0±6.1	28.8±6.0	28.8±6.1	28.8±6.1	28.9±6.3
p-value	0.36	0.48	0.24	0.27	
FEV_1 % pred					
Mean±SD	73.4±25.5	76.0±25.4	75.8±25.4	75.4±25.4	76.5±25.2
p-value	0.16	0.49	0.20	0.16	
Percent emphysema					
Mean±SD	7.0±9.9	6.3±10.0	6.4±10.0	6.5±10.0	6.1±9.4
p-value	0.39	0.39	0.11	0.090	
Airway wall thickness					
Mean±SD	1.1±0.2	1.1±0.2	1.1±0.2	1.1±0.2	1.1±0.2
p-value	0.22	0.17	0.35	0.33	
Bronchodilator response, % FEV_1					
Mean±SD	7.7±9.4	6.3±10.2	5.9±9.4	6.2±9.4	5.7±10.4
p-value	0.021 (β=1.54)	0.85	0.85	0.27	
Bronchiectasis					
Yes	16 (30.2)	23 (28.8)	36 (28.1)	117 (30.6)	312 (31.5)
No	37	57	92	265	678
p-value	0.31	0.33	0.34	0.28	

Data are presented as n or n (%), unless otherwise stated. p-values and effect sizes for variant-set testing of cystic fibrosis transmembrane conductance regulator (*CFTR*) variants with chronic obstructive pulmonary disease (COPD) and related phenotypes. CF: cystic fibrosis; BMI: body mass index; FEV_1: forced expiratory volume in 1 s. [#]: four variants with allele frequency >5% (881 counts) were excluded from analysis; [¶]: chronic bronchitis, severe exacerbation and BMI data were unavailable for n=2 subjects, COPD and severe COPD data were unavailable for n=58 subjects, FEV_1 % pred data were unavailable for n=58 subjects, percent emphysema data were unavailable for n=618 subjects, airway wall thickness data were unavailable for n=619 subjects, bronchodilator response data were unavailable for n=169 subjects, and bronchiectasis data were unavailable for n=7209 subjects. All p-values are one-sided, except for bronchodilator response which is two-sided. *: p-values that are significant after correction for multiple comparisons (p<0.05/10 or 0.005). Odds ratios or β coefficients are shown for all nominally significant associations (p<0.05).

CF-causing variant (F508del) tested for association with chronic bronchitis (table 3) and 36 variants tested for association with severe COPD (supplementary table S5). We found that F508del was significantly associated with chronic bronchitis (one-sided p=0.016; OR 1.47). While R75Q was nominally associated (p<0.05) with severe COPD after performing permutation analysis (p=0.02), no associations with severe COPD met the threshold for significance after correction for multiple comparisons (p<0.05/36 or 0.0014).

Compound heterozygotes in COPDGene

We next searched for subjects who may be compound heterozygotes, meaning that these subjects have two different *CFTR* variants on opposite chromosomes. There were no subjects with two CF-causing variants. We identified 32 subjects that were either heterozygous for F508del in addition to carrying another *CFTR*

TABLE 3 Single-variant testing of F508del for association with chronic bronchitis in COPDGene and ECLIPSE

	COPDGene	ECLIPSE	Meta-analysis
Allele count, n	177	57	
One-sided p-value from logistic regression	0.016	0.055	0.081
One-sided p-value from Firth regression	0.016		
One-sided p-value with permutation	0.028	0.061	
Odds ratio	1.47	1.67	1.52

variant or were heterozygous for two *CFTR* variants that have varying clinical consequence (supplementary table S6). We found that compound heterozygous subjects were enriched for COPD: out of the 32 compound heterozygotes, 21 were COPD cases while 11 were controls, whereas in noncompound heterozygous individuals there were 3827 COPD cases and 4680 controls (p=0.010) (table 4). There was no enrichment of chronic bronchitis or bronchiectasis in compound heterozygotes (table 4).

Replication in ECLIPSE

To attempt to replicate the results from COPDGene, we searched for *CFTR* variants in ECLIPSE. WGS and phenotyping data were available for 2212 subjects including 1953 cases and 165 controls (table 1). We identified 133 variants within the *CFTR* gene boundary including 19 CF-causing variants, 11 variants with varying clinical consequence, 13 variants that are not CF-causing and 32 variants that were not reported in the CFTR2 database or that have unknown significance (supplementary table S8). While the association of the 19 CF-causing variants with chronic bronchitis using burden testing did not reach statistical significance in ECLIPSE alone (one-sided p=0.057), we found a significant association in meta-analysis of ECLIPSE and COPDGene (p=0.0060; OR 1.52) (table 5). The only CF-causing variant in ECLIPSE with a minor allele count ⩾10 was the F508del variant, which was present in 57 subjects. Single-variant testing revealed a suggestive association between F508del and chronic bronchitis in ECLIPSE (one-sided p=0.055; OR 1.67) (table 5) and in meta-analysis of COPDGene and ECLIPSE (one-sided p=0.081; OR 1.52).

Discussion

This study is the largest to date characterising the effect of *CFTR* variants in smokers with and without COPD. We found that CF-causing variants are associated with chronic bronchitis and this is primarily driven by the most common CF-causing variant, F508del. We also found a suggestive association between all coding *CFTR* variants and severe COPD in the COPDGene study. Furthermore, we found that subjects that are compound heterozygotes for *CFTR* variants are at increased risk for COPD.

TABLE 4 Compound heterozygotes in COPDGene

	Clinically significant or predicted to be functional[#]	Clinically significant or predicted functional+varying clinical consequence[¶]	All compound heterozygotes	Controls	One-sided p-value for all compound heterozygotes
Subjects	8	14	32	8565	
COPD					0.010*
Cases	5	8	21	3827	
Controls	3	6	11	4680	
Chronic bronchitis					0.13
Yes	3	3	9	1656	
No	5	11	23	6907	
Bronchiectasis					0.090
Yes	0	1	3	426	
No	1	1	2	941	

Data are presented as n subjects identified who are compound heterozygotes for cystic fibrosis transmembrane conductance regulator (*CFTR*) variants, unless otherwise stated. COPD: chronic obstructive pulmonary disease. [#]: n=8 subjects all carry one copy of the F508del variant and one variant of unknown function according to the CFTR2 database that is predicted to have moderate effect according to SnpEff; [¶]: includes the n=8 subjects from the first group, n=1 subject that carries one F508del variant and one variant of varying clinical consequence, and n=5 subjects that carry two variants of varying clinical consequence. p-values were computed using Fisher's exact test to test whether COPD, chronic bronchitis and bronchiectasis cases were enriched in all compound heterozygotes compared with controls. Statistical testing was not performed for the other two groups due to the small sample sizes. *: p-value significant after correction for multiple comparisons (p<0.05/3).

TABLE 5 Burden testing of association between cystic fibrosis (CF)-causing variants and chronic bronchitis in ECLIPSE

	Variants, n	ECLIPSE p-value	ECLIPSE+COPDGene meta-analysis p-value
All CF-causing variants in ECLIPSE	19	0.057	0.0060 (OR 1.52)
CF-causing variants in ECLIPSE also found in COPDGene	13	0.12	0.064

p-values and effect sizes for association between CF-causing variants and chronic bronchitis in ECLIPSE and meta-analysis between ECLIPSE and COPDGene.

Several previous studies have shown that heterozygous *CFTR* variants can have a functional effect. For example, *CFTR* heterozygous variants are associated with idiopathic pancreatitis [3, 4], congenital bilateral absence of the vas deferens [5], bronchiectasis [34] and allergic bronchopulmonary aspergillosis [6]. CF carriers may have an increased risk for developing airway obstruction, and have been shown to have abnormalities in neutrophil function [35] and apoptosis [36] that may lead to a prolonged inflammatory state that could predispose to accelerated lung function decline. Furthermore, cigarette smoke is associated with decreased CFTR function in the upper and lower airways of both healthy smokers and smokers with COPD, and defective CFTR has been associated with symptoms of chronic bronchitis and dyspnoea [7, 8]. Therefore, it is possible that the presence of heterozygous genetic variants may increase the prevalence of chronic bronchitis or COPD in smokers. While several small studies have been conducted to test this hypothesis, results to date have been mixed. One study found that F508del variants were present at an increased frequency in subjects with chronic bronchitis and elevated sweat chloride levels [19]. Several small studies have found modestly elevated *CFTR* variant frequencies in subjects with COPD or chronic bronchitis [17, 18, 20, 22]. Most strikingly, a recent study including 108 035 Danish individuals identified 2858 F508del individuals and found that these individuals had an increased risk of chronic bronchitis (OR 1.31), as well as an increased risk of bronchiectasis (hazard ratio 1.88) [23]. In addition, Miller *et al*. [24] reported that *CFTR* variants were associated with an increase in chronic bronchitis (OR 1.24). However, other studies have failed to find that *CFTR* heterozygous variants have a functional effect. A study exposing *CFTR* heterozygous mice and cell lines to cigarette smoke found that *CFTR* heterozygosity did not have an impact on residual CFTR activity [21]. In a study of obstructive pulmonary disease that included 250 F508del heterozygotes, COPD was not found to be increased, and measures of lung function were only lower in F508del heterozygotes who also had asthma [15, 16]. Furthermore, genome-wide association studies (GWAS) of lung function, COPD and emphysema have not identified *CFTR* as a susceptibility gene, although GWAS chips do not genotype the F508del variant and this variant is typically not well imputed. Thus, the contribution of heterozygosity for CF variants to the aetiology of COPD has been unclear, possibly due to the small sample size of studies to date, the use of heterogeneous groups of patients and the lack of gene sequencing to fully assess *CFTR* variants.

In this study, we sought to increase the power to detect the effect of rare *CFTR* variants by performing variant-set testing followed by individual testing of specific categories of variants. This allowed us to include ultra-rare variants, including variants only present in one subject in the dataset (singletons). We found that the combination of CF-causing variants was associated with chronic bronchitis with statistical significance. The OR for association in COPDGene was 1.53 and the OR in meta-analysis of COPDGene and ECLIPSE was 1.52. Similarly, the OR for association of F508del with chronic bronchitis was 1.47 in COPDGene and 1.52 in meta-analysis of COPDGene and ECLIPSE. This indicates that smokers with CF-causing variants are ∼1.5 times more likely to have chronic bronchitis than subjects without *CFTR* variants and the consistency of the odds ratios across the two studies is an indicator of the validity of our findings. The finding that the odds ratio is slightly higher in our study of only current or former smokers, compared with what has been reported in the literature (OR range 1.24–1.31), is consistent with the hypothesis that a history of cigarette smoking would result in a greater effect of *CFTR* variants. We also found suggestive evidence that variants with less established function (such as variants of varying clinical severity or predicted moderate impact) may be associated with chronic bronchitis. In addition, we found that the combination of all *CFTR* variants was nominally associated with severe COPD. This is of particular interest as it suggests that there could be a large number of COPD patients carrying *CFTR* variants that contribute to their disease severity and who could potentially benefit from treatment with CFTR modulators. Single-variant testing of the association of all *CFTR* variants did not identify any associated variants that were significant after correction for multiple comparison; however, the non-CF-causing variant R75Q was nominally associated with severe COPD. R75Q is a relatively common

missense variant which is not CF-causing but has been associated with pancreatitis [37] and an increased frequency of R75Q has previously been found in patients with COPD [17].

We found that the only variant that was significantly associated with either chronic bronchitis or bronchodilator response using single-variant testing was F508del. This was unsurprising given that F508del is the most common CF-causing variant identified in both COPDGene and ECLIPSE, as well as in the general population. Furthermore, F508del is a relatively severe class II variant, which produces a misfolded protein with little functional capacity. Therefore, it was one of the few variants for which we had sufficient power to detect associations with single-variant testing. We identified 32 subjects that were compound heterozygotes for *CFTR* variants, meaning that they carry two copies of *CFTR* variants on separate chromosomes, and found that these subjects were enriched for COPD cases compared with noncompound heterozygotes. It is not possible to definitively conclude that these compound heterozygous subjects did not in fact have CF, due to the lack of CF diagnostic tests such as sweat chloride measurements in the COPDGene study. However, subjects with lung disease other than COPD or asthma, or with diffuse bronchiectasis on chest CT scans, were excluded. In the 32 compound heterozygotes identified here, only one subject reported a history of pneumonia, chronic bronchitis, or chronic cough or phlegm in early life (prior to age 15 years), suggesting that these subjects did not have a history of early respiratory disease consistent with typical CF. We conclude that decreased CFTR activity due to two *CFTR* variants can result in COPD, based on the accepted GOLD definition [38].

While this study has several strengths, including being the largest study to characterise *CFTR* variants using WGS in smokers with and without COPD and having replication in an independent cohort, there are also several limitations. Despite the large sample size, there were still small numbers of subjects with the less common *CFTR* variants and therefore we are not able to determine whether these variants contribute to COPD. For example, the G551D variant is of particular interest since it can be corrected with ivacaftor; however, we only identified eight subjects that were heterozygous for this variant. The functional impact of most of the variants identified in our study is not known, and combining functional and nonfunctional variants reduces power for association studies. In addition, almost all subjects in both COPDGene and ECLIPSE had a history of smoking, and therefore we were not able to test if heterozygous *CFTR* variants have a function consequence in the absence of cigarette smoke. In summary, using unique analyses of *CFTR* variants in a cohort of smokers we found that *CFTR* variants, and particularly F508del, are associated with chronic bronchitis.

Acknowledgements: Molecular data for the Trans-Omics in Precision Medicine (TOPMed) programme was supported by the National Heart, Lung, and Blood Institute (NHLBI). Whole-genome sequencing for “NHLBI TOPMed: Genetic Epidemiology of COPD (COPDGene)” (phs000951) was performed at the Broad Institute Genomics Platform (HHSN268201500014C) and the Northwest Genomics Center (3R01HL08985608S1). Whole-genome sequencing for “NHLBI TOPMed: Evaluation of COPD Longitudinally to Identify Predictive Surrogate Endpoints (ECLIPSE)” (phs001472) was performed at the McDonnell Genome Institute (HHSN268201600037I). Core support including centralised genomic read mapping and genotype calling, along with variant quality metrics and filtering, was provided by the TOPMed Informatics Research Center (3R01HL11762602S1; contract HHSN268201800002I). Core support including phenotype harmonisation, data management, sample-identity quality control and general programme coordination was provided by the TOPMed Data Coordinating Center (R01HL120393; U01HL120393; contract HHSN268201800001I). We gratefully acknowledge the studies and participants who provided biological samples and data for TOPMed.

COPDGene Investigators – Core Units: Administrative Center: James D. Crapo (PI), Edwin K. Silverman (PI), Barry J. Make, Elizabeth A. Regan. Genetic Analysis Center: Terri Beaty, Ferdouse Begum, Peter J. Castaldi, Michael Cho, Dawn L. DeMeo, Adel R. Boueiz, Marilyn G. Foreman, Eitan Halper-Stromberg, Lystra P. Hayden, Craig P. Hersh, Jacqueline Hetmanski, Brian D. Hobbs, John E. Hokanson, Nan Laird, Christoph Lange, Sharon M. Lutz, Merry-Lynn McDonald, Margaret M. Parker, Dmitry Prokopenko, Dandi Qiao, Elizabeth A. Regan, Phuwanat Sakornsakolpat, Edwin K. Silverman, Emily S. Wan, Sungho Won. Imaging Center: Juan Pablo Centeno, Jean-Paul Charbonnier, Harvey O. Coxson, Craig J. Galban, MeiLan K. Han, Eric A. Hoffman, Stephen Humphries, Francine L. Jacobson, Philip F. Judy, Ella A. Kazerooni, Alex Kluiber, David A. Lynch, Pietro Nardelli, John D. Newell Jr, Aleena Notary, Andrea Oh, Elizabeth A. Regan, James C. Ross, Raul San Jose Estepar, Joyce Schroeder, Jered Sieren, Berend C. Stoel, Juerg Tschirren, Edwin Van Beek, Bram van Ginneken, Eva van Rikxoort, Gonzalo Vegas Sanchez-Ferrero, Lucas Veitel, George R. Washko, Carla G. Wilson. PFT QA Center, Salt Lake City, UT: Robert Jensen. Data Coordinating Center and Biostatistics, National Jewish Health, Denver, CO: Douglas Everett, Jim Crooks, Katherine Pratte, Matt Strand, Carla G. Wilson. Epidemiology Core, University of Colorado Anschutz Medical Campus, Aurora, CO: John E. Hokanson, Gregory Kinney, Sharon M. Lutz, Kendra A. Young. Mortality Adjudication Core: Surya P. Bhatt, Jessica Bon, Alejandro A. Diaz, MeiLan K. Han, Barry Make, Susan Murray, Elizabeth Regan, Xavier Soler, Carla G. Wilson. Biomarker Core: Russell P. Bowler, Katerina Kechris, Farnoush Banaei-Kashani.

COPDGene Investigators – Clinical Centers: Ann Arbor VA: Jeffrey L. Curtis, Perry G. Pernicano. Baylor College of Medicine, Houston, TX: Nicola Hanania, Mustafa Atik, Aladin Boriek, Kalpatha Guntupalli, Elizabeth Guy, Amit Parulekar. Brigham and Women's Hospital, Boston, MA: Dawn L. DeMeo, Alejandro A. Diaz, Lystra P. Hayden, Brian D. Hobbs, Craig Hersh, Francine L. Jacobson, George Washko. Columbia University, New York, NY: R. Graham Barr, John Austin, Belinda D'Souza, Byron Thomashow. Duke University Medical Center, Durham, NC: Neil MacIntyre Jr, H. Page McAdams, Lacey Washington. Grady Memorial Hospital, Atlanta, GA: Eric Flenaugh, Silanth Terpenning. HealthPartners Research Institute, Minneapolis, MN: Charlene McEvoy, Joseph Tashjian. Johns Hopkins University, Baltimore, MD: Robert Wise, Robert Brown, Nadia N. Hansel, Karen Horton, Allison Lambert, Nirupama Putcha. Lundquist Institute for Biomedical Innovation at Harbor UCLA Medical Center, Torrance, CA: Richard Casaburi, Alessandra Adami, Matthew Budoff, Hans Fischer, Janos Porszasz, Harry Rossiter, William Stringer. Michael E. DeBakey VAMC, Houston, TX: Amir Sharafkhaneh, Charlie Lan. Minneapolis VA: Christine Wendt, Brian Bell, Ken M. Kunisaki. National Jewish Health, Denver, CO: Russell Bowler, David A. Lynch. Reliant Medical Group, Worcester, MA: Richard Rosiello, David Pace. Temple University, Philadelphia, PA: Gerard Criner, David Ciccolella, Francis Cordova, Chandra Dass, Gilbert D'Alonzo, Parag Desai, Michael Jacobs, Steven Kelsen, Victor Kim, A. James Mamary, Nathaniel Marchetti, Aditi Satti, Kartik Shenoy, Robert M. Steiner, Alex Swift, Irene Swift, Maria Elena Vega-Sanchez. University of Alabama, Birmingham, AL: Mark Dransfield, William Bailey, Surya P. Bhatt, Anand Iyer, Hrudaya Nath, J. Michael Wells. University of California, San Diego, CA: Douglas Conrad, Xavier Soler, Andrew Yen. University of Iowa, Iowa City, IA: Alejandro P. Comellas, Karin F. Hoth, John Newell Jr, Brad Thompson. University of Michigan, Ann Arbor, MI: MeiLan K. Han, Ella Kazerooni, Wassim Labaki, Craig Galban, Dharshan Vummidi. University of Minnesota, Minneapolis, MN: Joanne Billings, Abbie Begnaud, Tadashi Allen. University of Pittsburgh, Pittsburgh, PA: Frank Sciurba, Jessica Bon, Divay Chandra, Carl Fuhrman, Joel Weissfeld. University of Texas Health, San Antonio, San Antonio, TX: Antonio Anzueto, Sandra Adams, Diego Maselli-Caceres, Mario E. Ruiz, Harjinder Singh.

Conflict of interest: C.P. Hersh has received grants from the NHLBI, Alpha-1 Foundation, Bayer, Boehringer Ingelheim, Novartis and Vertex, and consulting fees from Takeda. A.A. Diaz has received grants from the NHLBI. G.R. Cutting has received grants from the NIDDK and US CF Foundation. M.H. Co has received grant support from Bayer and GSK, and consulting or speaking fees from Genentech, AstraZeneca and Illumina. H. Levy has received grants from the NHLBI and NIH Office of the Director, and consulting fees as part of the Chan Zuckerberg Rare Disease Consortium. A. Saferali, D. Qiao, W. Kim and K. Raraigh do not have any conflicts of interest to disclose.

Support statement: This work was supported by National Institutes of Health grants R01HL133137, R01HL149861, R01DK044003, R01HL130512, R01HL149861, R01HL135142, R01HL137927, R01HL089856, R01HL147148, U01HL089897, U01HL089856, T32HL007427, K01HL157613 and K01HL129039. COPDGene is also supported by the COPD Foundation through contributions made to an industry advisory board comprising AstraZeneca, Boehringer Ingelheim, Genentech, GlaxoSmithKline, Novartis, Pfizer, Siemens and Sunovion. Funding information for this article has been deposited with the Crossref Funder Registry.

References

1 Kim V, Criner GJ. Chronic bronchitis and chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2013; 187: 228–237.
2 Kim V, Sternberg AL, Washko G, *et al.* Severe chronic bronchitis in advanced emphysema increases mortality and hospitalizations. *COPD* 2013; 10: 667–678.
3 Sharer N, Schwarz M, Malone G, *et al.* Mutations of the cystic fibrosis gene in patients with chronic pancreatitis. *N Engl J Med* 1998; 339: 645–652.
4 Cohn JA, Friedman KJ, Noone PG, *et al.* Relation between mutations of the cystic fibrosis gene and idiopathic pancreatitis. *N Engl J Med* 1998; 339: 653–658.
5 Chillon M, Casals T, Mercier B, *et al.* Mutations in the cystic fibrosis gene in patients with congenital absence of the vas deferens. *N Engl J Med* 1995; 332: 1475–1480.
6 Miller PW, Hamosh A, Macek M Jr, *et al.* Cystic fibrosis transmembrane conductance regulator (CFTR) gene mutations in allergic bronchopulmonary aspergillosis. *Am J Hum Genet* 1996; 59: 45–51.
7 Sloane PA, Shastry S, Wilhelm A, *et al.* A pharmacologic approach to acquired cystic fibrosis transmembrane conductance regulator dysfunction in smoking related lung disease. *PLoS One* 2012; 7: e39809.
8 Dransfield MT, Wilhelm AM, Flanagan B, *et al.* Acquired cystic fibrosis transmembrane conductance regulator dysfunction in the lower airways in COPD. *Chest* 2013; 144: 498–506.
9 Cantin AM, Hanrahan JW, Bilodeau G, *et al.* Cystic fibrosis transmembrane conductance regulator function is suppressed in cigarette smokers. *Am J Respir Crit Care Med* 2006; 173: 1139–1144.
10 Clunes LA, Davies CM, Coakley RD, *et al.* Cigarette smoke exposure induces CFTR internalization and insolubility, leading to airway surface liquid dehydration. *FASEB J* 2012; 26: 533–545.
11 Stanley PJ, Wilson R, Greenstone MA, *et al.* Effect of cigarette smoking on nasal mucociliary clearance and ciliary beat frequency. *Thorax* 1986; 41: 519–523.

12 Raju SV, Lin VY, Liu L, *et al.* The cystic fibrosis transmembrane conductance regulator potentiator ivacaftor augments mucociliary clearance abrogating cystic fibrosis transmembrane conductance regulator inhibition by cigarette smoke. *Am J Respir Cell Mol Biol* 2017; 56: 99–108.
13 Solomon GM, Hathorne H, Liu B, *et al.* Pilot evaluation of ivacaftor for chronic bronchitis. *Lancet Respir Med* 2016; 4: e32–e33.
14 Rowe SM, Jones I, Dransfield MT, *et al.* Efficacy and safety of the CFTR potentiator icenticaftor (QBW251) in COPD: results from a phase 2 randomized trial. *Int J Chron Obstruct Pulmon Dis* 2020; 15: 2399–2409.
15 Dahl M, Nordestgaard BG, Lange P, *et al.* Fifteen-year follow-up of pulmonary function in individuals heterozygous for the cystic fibrosis phenylalanine-508 deletion. *J Allergy Clin Immunol* 2001; 107: 818–823.
16 Dahl M, Tybjærg-Hansen A, Lange P, *et al.* ΔF508 heterozygosity in cystic fibrosis and susceptibility to asthma. *Lancet* 1998; 351: 1911–1913.
17 Divac A, Nikolic A, Mitic-Milikic M, *et al.* High frequency of the R75Q CFTR variation in patients with chronic obstructive pulmonary disease. *J Cyst Fibros* 2004; 3: 189–191.
18 Dumur V, Lafitte JJ, Gervais R, *et al.* Abnormal distribution of cystic fibrosis delta F508 allele in adults with chronic bronchial hypersecretion. *Lancet* 1990; 335: 1340.
19 Gervais R, Lafitte JJ, Dumur V, *et al.* Sweat chloride and delta F508 mutation in chronic bronchitis or bronchiectasis. *Lancet* 1993; 342: 997.
20 Kostuch M, Semczuk A, Szarewicz-Adamczyk W, *et al.* Detection of CFTR gene mutations in patients suffering from chronic bronchitis. *Arch Med Res* 2000; 31: 97–100.
21 Raju SV, Tate JH, Peacock SK, *et al.* Impact of heterozygote CFTR mutations in COPD patients with chronic bronchitis. *Respir Res* 2014; 15: 18.
22 Stankovic M, Nikolic A, Divac A, *et al.* The *CFTR* M470V gene variant as a potential modifier of COPD severity: study of Serbian population. *Genet Test* 2008; 12: 357–362.
23 Colak Y, Nordestgaard BG, Afzal S. Morbidity and mortality in carriers of the cystic fibrosis mutation *CFTR* Phe508del in the general population. *Eur Respir J* 2020; 56: 2000558.
24 Miller AC, Comellas AP, Hornick DB, *et al.* Cystic fibrosis carriers are at increased risk for a wide range of cystic fibrosis-related conditions. *Proc Natl Acad Sci USA* 2020; 117: 1621–1627.
25 Regan EA, Hokanson JE, Murphy JR, *et al.* Genetic Epidemiology of COPD (COPDGene) study design. *COPD* 2010; 7: 32–43.
26 Vestbo J, Anderson W, Coxson HO, *et al.* Evaluation of COPD Longitudinally to Identify Predictive Surrogate Endpoints (ECLIPSE). *Eur Respir J* 2008; 31: 869–873.
27 Diaz AA, Maselli DJ, Rahaghi F, *et al.* Pulmonary vascular pruning in smokers with bronchiectasis. *ERJ Open Res* 2018; 4: 00044-2018.
28 Taliun D, Harris DN, Kessler MD, *et al.* Sequencing of 53,831 diverse genomes from the NHLBI TOPMed Program. *Nature* 2021; 590: 290–299.
29 Li H. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics* 2011; 27: 2987–2993.
30 Liu X, White S, Peng B, *et al.* WGSA: an annotation pipeline for human genome sequencing studies. *J Med Genet* 2016; 53: 111–112.
31 Prokopenko D, Hecker J, Silverman EK, *et al.* Utilizing the Jaccard index to reveal population stratification in sequencing data: a simulation study and an application to the 1000 Genomes Project. *Bioinformatics* 2016; 32: 1366–1372.
32 Prokopenko D, Sakornsakolpat P, Fier HL, *et al.* Whole-genome sequencing in severe chronic obstructive pulmonary disease. *Am J Respir Cell Mol Biol* 2018; 59: 614–622.
33 Lee S, Emond MJ, Bamshad MJ, *et al.* Optimal unified approach for rare-variant association testing with application to small-sample case-control whole-exome sequencing studies. *Am J Hum Genet* 2012; 91: 224–237.
34 Bienvenu T, Sermet-Gaudelus I, Burgel PR, *et al.* Cystic fibrosis transmembrane conductance regulator channel dysfunction in non-cystic fibrosis bronchiectasis. *Am J Respir Crit Care Med* 2010; 181: 1078–1084.
35 Witko-Sarsat V, Allen RC, Paulais M, *et al.* Disturbed myeloperoxidase-dependent activity of neutrophils in cystic fibrosis homozygotes and heterozygotes, and its correction by amiloride. *J Immunol* 1996; 157: 2728–2735.
36 Moriceau S, Lenoir G, Witko-Sarsat V. In cystic fibrosis homozygotes and heterozygotes, neutrophil apoptosis is delayed and modulated by diamide or roscovitine: evidence for an innate neutrophil disturbance. *J Innate Immun* 2010; 2: 260–266.
37 Schneider A, Larusch J, Sun X, *et al.* Combined bicarbonate conductance-impairing variants in *CFTR* and *SPINK1* variants are associated with chronic pancreatitis in patients without cystic fibrosis. *Gastroenterology* 2011; 140: 162–171.
38 Rabe KF, Hurd S, Anzueto A, *et al.* Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease: GOLD executive summary. *Am J Respir Crit Care Med* 2007; 176: 532–555.