EUROPEAN RESPIRATORY *journal*

FLAGSHIP SCIENTIFIC JOURNAL OF ERS

Early View

Original research article

CFTR variants are associated with chronic bronchitis in smokers

Aabida Saferali, Dandi Qiao, Wonji Kim, Karen Raraigh, Hara Levy, Alejandro A. Diaz, Garry R. Cutting, Michael H. Cho, Craig P. Hersh

Please 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; in press (https://doi.org/10.1183/13993003.01994-2021).

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], Craig P. Hersh[1,2,5*] NHLBI TransOmics in Precision Medicine (TOPMed)

[1]Channing Division of Network Medicine, Brigham and Women's Hospital, Boston, Massachusetts, USA

[2]Harvard Medical School, Boston, Massachusetts, USA

[3]Johns Hopkins University School of Medicine

[4]Division of Pulmonary Medicine, Department of Pediatrics, National Jewish Health, Denver, Colorado, USA

[5] Division of Pulmonary and Critical Care Medicine, Brigham and Women's Hospital

* Corresponding Author

E-mail: craig.hersh@channing.harvard.edu

Funding: R01HL133137, R01HL149861, R01DK044003, R01HL130512, R01HL149861, R01HL135142, R01HL137927, R01 HL089856, R01HL147148, U01HL089897, U01HL089856, T32HL007427, K01HL157613,K01 HL129039.

COPDGene is also supported by the COPD Foundation through contributions made to an Industry Advisory Board comprised of AstraZeneca, Boehringer-Ingelheim, Genentech, GlaxoSmithKline, Novartis, Pfizer, Siemens, and Sunovion.

Disclosure of potential conflict of interest: CPH has received grants from NHLBI, Alpha-1 Foundation, Bayer, Boehringer-Ingelheim, Novartis and Vertex, and consulting fees from Takeda. AAD has received grants from NHLBI. GRC has received grants from the NIDDK and U.S. CF Foundation. MHC has received grant support from Bayer and GSK, and consulting or speaking fees from Genentech, Astrazeneca, and Illumina. HL has received grants from NHLBI and NIH Office of the Director, and consulting fees as part of the Chan Zuckerberg Rare Disease Consortium. AS, DQ, WK, and KR do not have any conflicts of interest to disclose.

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 hypothesize that *CFTR* variants in individuals with a history of smoking are associated with COPD and related phenotypes.

Methods: Whole genome sequencing was performed through the NHLBI TOPMed program in 8597 subjects from the COPDGene 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 2,118 subjects from the Evaluation of COPD Longitudinally to Identify Predictive Surrogate Endpoints (ECLIPSE) 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-value=0.0025, OR =1.53) and in meta-analysis of COPDGene and ECLIPSE (one sided p-value=0.0060, OR =1.52). Single variant testing revealed that the F508del variant was associated with chronic bronchitis in COPDGene (one sided p-value=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 pathologic 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 thirty-five Americans is a carrier of a loss of function variant in the Cystic Fibrosis Transmembrane Conductance Regulator (*CFTR*). 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 FEV_1, as well as reduce systemic inflammation and sputum colonization 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 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 non-smokers 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 Genetic Epidemiology of COPD (COPDGene) and Evaluation of COPD Longitudinally to Identify Predictive Surrogate Endpoints (ECLIPSE) 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 multicenter multinational 3-year longitudinal study that enrolled 3,291 subjects of GOLD stage 2-4. In COPDGene, COPD was defined by a postbronchodilator ratio of forced expiratory volume in 1 second (FEV_1) to forced vital capacity (FVC) < 0.7 (Global Initiative for Chronic Obstructive Lung Disease (GOLD) 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 two years. Bronchodilator response was defined as the % change in pre/post bronchodilator FEV_1. Visual scoring of bronchiectasis was performed using CT scans for 1,372 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

Whole genome sequencing data was generated through the NHLBI TOPMed consortium to a mean depth of 30X using DNA from blood, PCR-free library construction and Illumina HiSeq X technology(28). For COPDGene, Freeze 5b WGS data was used which includes 8,598 subjects

including 5,773 non-Hispanic white (NHW) and 2825 African American (AA). For replication in ECLIPSE, Freeze 8 WGS data was used which included 2345 subjects, and a subset of 2212 were included in this analysis. Reads were mapped to human genome assembly version GRCh38 and computational phasing was performed using Eagle 2.4 (Dec 13, 2017).

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 characterize 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 Variant Effect Predictor (VEP) consequence. Annotation of known CF-causing variants was downloaded from the CFTR2 consortium website (https://cftr2.org/) (accessed on May 4, 2021). These variants are categorized as CF-causing, varying-clinical significance, non-CF causing and unknown significance. For variants that were not reported in the CFTR2 database, 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 was 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 hypothesized *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 severe COPD, increased risk of severe exacerbations, decreased BMI, decreased FEV_1 percent predicted, 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 percent 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 count ≥ 10 were included. Analyses were adjusted for age, sex, pack-years of smoking, current smoking status, and principal components of genetic ancestry (PC). Calculation of PCs has been previously described (31, 32). Analyses in COPDGene were performed in NHW and AA individuals combined, using 3 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/non-variant carrier status among all subjects 20,000 times, then computing the p-value using the number of permutations in which the test statistic is more extreme than the observed test statistic. As described above, 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 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 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 COPDGene cases and 4691 smoking controls were available for analysis (**Table 1**). In these subjects, we identified 11,567 variants within the gene boundary of CTFR 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 (SNV) and 990 are insertion-deletion polymorphisms (indel). Of these, there were 301 variants that are located within the coding region of the RefSeq Select transcript (NM_000492.3) (**Supplementary Table 1**). 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 chr7:117595001:T:G), while rs213950 (lV470M) is a missense variant known to be non-CF causing. After these, the most frequent variants were chr7:117509093:G:A (R75Q) with 459 counts and chr7:117559655:G:A (1716G/A) with 290 counts, both of which are non-CF causing missense variants. We additionally identified 177 subjects that are 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 1 stop-gain and 89 missense variants which are predicted to have moderate to high impact on CFTR through SNPEff functional impact prediction.

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 CFTR2 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 subjects out of 248 with CF-causing variants have chronic bronchitis (27.4%) while 1,597 out of 8,345 subjects without CF-causing variants have chronic bronchitis (19.1%)($p=0.0025$, OR=1.53)(**Table 2**). We hypothesized 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 % 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 2**). Since we hypothesized 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 vs former smokers, where we found that 39.5% of currently smoking subjects with CF-causing variants had chronic bronchitis, compared to 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 3**). In contrast, in former smokers we found that 17.2% of subjects with CF-causing variants had chronic bronchitis, compared to 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 to subjects without CFTR variants (25.5%)(p=0.0022, OR=1.72)(**Supplementary Table 4**). 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 1,372 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 of at least 10. This resulted in one CF-causing variant (F508del) tested for association with chronic bronchitis (**Table 3**), and 36 variants tested for association with severe COPD (**Supplementary Table 5**). We found that F508del was significantly associated with chronic bronchitis (one sided p-value=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* variant or were heterozygous for two *CFTR* variants that have varying clinical consequence (**Supplementary Table 6**). 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 non-compound 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. Whole genome sequencing and phenotyping data were available for 2212 subjects including 1953 cases and 165 controls. 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 CFTR2 or that have unknown significance (**Supplementary Table 8**). 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 of greater than 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-value=0.055, OR=1.67)(**Table 5**), and in meta-analysis of COPDGene and ECLIPSE (one sided p-value=0.081, OR=1.52).

DISCUSSION:

This study is the largest to date characterizing 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.

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 dyspnea (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, 20, 22) (18). Most

strikingly, a recent study including 108,035 Danish individuals identified 2858 F508del individuals and found that these individuals had an increased risk of bronchiectasis with an odds ratio of 1.31, as well as an increased risk of bronchiectasis with a hazard ratio of 1.88 (23). In addition, Miller *et al*. reported that CFTR variants were associated with an increase of chronic bronchitis with and odds ratio of 1.24 (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, though 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 etiology of COPD has been unclear, possibly due to the small sample size of studies to date, and 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 chronic bronchitis with statistical significance. The OR for the association in COPDGene was 1.53, and the OR in the meta analysis of COPDGene and ECLIPSE was 1.52. Similarly, the OR for the association of F508del with chronic bronchitis was 1.47 in COPDGene and 1.52 in the

meta-analysis of COPDGene and ECLIPSE. This indicates that smokers with CF-causing variants are approximately 1.5 times more likely to have chronic bronchitis than subjects without CFTR variants, and the consistency of the OR across the two studies is an indicator of the validity of our findings. The finding that the OR is slightly higher in our study of only current or former smokers, compared to what has been reported in the literature (OR ranges 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 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 to non-compound heterozygotes. It is not possible to definitively conclude that these compound heterozygous subjects do 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), suggesting that these subjects did not have 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 characterize *CFTR* variants using whole genome sequencing 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 8 subjects that were heterozygous for this variant. The functional impact of most of the variants identified in our study are not known, and combining functional and non-functional variants reduces power for association studies. In addition, almost all subjects in both COPDGene and ECLIPSE have 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) program 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 (3R01HL089856-08S1). 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 centralized genomic read mapping and genotype calling, along with variant quality metrics and filtering were provided by the TOPMed Informatics Research Center (3R01HL-117626-02S1; contract HHSN268201800002I). Core support including phenotype harmonization, data management, sample-identity QC, and general program coordination were provided by the TOPMed Data Coordinating Center (R01HL-120393; U01HL-120393; 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, MD (PI); Edwin K. Silverman, MD, PhD (PI); Barry J. Make, MD; Elizabeth A. Regan, MD, PhD

Genetic Analysis Center: Terri Beaty, PhD; Ferdouse Begum, PhD; Peter J. Castaldi, MD, MSc; Michael Cho, MD; Dawn L. DeMeo, MD, MPH; Adel R. Boueiz, MD; Marilyn G. Foreman, MD, MS; Eitan Halper-Stromberg; Lystra P. Hayden, MD, MMSc; Craig P. Hersh, MD, MPH; Jacqueline Hetmanski, MS, MPH; Brian D. Hobbs, MD; John E. Hokanson, MPH, PhD; Nan Laird, PhD; Christoph Lange, PhD; Sharon M. Lutz, PhD; Merry-Lynn McDonald, PhD; Margaret M. Parker, PhD; Dmitry Prokopenko, Ph.D; Dandi Qiao, PhD; Elizabeth A. Regan,

MD, PhD; Phuwanat Sakornsakolpat, MD; Edwin K. Silverman, MD, PhD; Emily S. Wan, MD; Sungho Won, PhD

Imaging Center: Juan Pablo Centeno; Jean-Paul Charbonnier, PhD; Harvey O. Coxson, PhD; Craig J. Galban, PhD; MeiLan K. Han, MD, MS; Eric A. Hoffman, Stephen Humphries, PhD; Francine L. Jacobson, MD, MPH; Philip F. Judy, PhD; Ella A. Kazerooni, MD; Alex Kluiber; David A. Lynch, MB; Pietro Nardelli, PhD; John D. Newell, Jr., MD; Aleena Notary; Andrea Oh, MD; Elizabeth A. Regan, MD, PhD; James C. Ross, PhD; Raul San Jose Estepar, PhD; Joyce Schroeder, MD; Jered Sieren; Berend C. Stoel, PhD; Juerg Tschirren, PhD; Edwin Van Beek, MD, PhD; Bram van Ginneken, PhD; Eva van Rikxoort, PhD; Gonzalo Vegas Sanchez-Ferrero, PhD; Lucas Veitel; George R. Washko, MD; Carla G. Wilson, MS;

PFT QA Center, Salt Lake City, UT: Robert Jensen, PhD

Data Coordinating Center and Biostatistics, National Jewish Health, Denver, CO: Douglas Everett, PhD; Jim Crooks, PhD; Katherine Pratte, PhD; Matt Strand, PhD; Carla G. Wilson, MS

Epidemiology Core, University of Colorado Anschutz Medical Campus, Aurora, CO: John E. Hokanson, MPH, PhD; Gregory Kinney, MPH, PhD; Sharon M. Lutz, PhD; Kendra A. Young, PhD

Mortality Adjudication Core: Surya P. Bhatt, MD; Jessica Bon, MD; Alejandro A. Diaz, MD, MPH; MeiLan K. Han, MD, MS; Barry Make, MD; Susan Murray, ScD; Elizabeth Regan, MD; Xavier Soler, MD; Carla G. Wilson, MS

Biomarker Core: Russell P. Bowler, MD, PhD; Katerina Kechris, PhD; Farnoush Banaei-Kashani, Ph.D

COPDGene® Investigators – Clinical Centers

Ann Arbor VA: Jeffrey L. Curtis, MD; Perry G. Pernicano, MD

Baylor College of Medicine, Houston, TX: Nicola Hanania, MD, MS; Mustafa Atik, MD; Aladin Boriek, PhD; Kalpatha Guntupalli, MD; Elizabeth Guy, MD; Amit Parulekar, MD;

Brigham and Women's Hospital, Boston, MA: Dawn L. DeMeo, MD, MPH; Alejandro A. Diaz, MD, MPH; Lystra P. Hayden, MD; Brian D. Hobbs, MD; Craig Hersh, MD, MPH; Francine L. Jacobson, MD, MPH; George Washko, MD

Columbia University, New York, NY: R. Graham Barr, MD, DrPH; John Austin, MD; Belinda D'Souza, MD; Byron Thomashow, MD

Duke University Medical Center, Durham, NC: Neil MacIntyre, Jr., MD; H. Page McAdams, MD; Lacey Washington, MD

Grady Memorial Hospital, Atlanta, GA: Eric Flenaugh, MD; Silanth Terpenning, MD

HealthPartners Research Institute, Minneapolis, MN: Charlene McEvoy, MD, MPH; Joseph Tashjian, MD

Johns Hopkins University, Baltimore, MD: Robert Wise, MD; Robert Brown, MD; Nadia N. Hansel, MD, MPH; Karen Horton, MD; Allison Lambert, MD, MHS; Nirupama Putcha, MD, MHS

Lundquist Institute for Biomedical Innovationat Harbor UCLA Medical Center, Torrance, CA: Richard Casaburi, PhD, MD; Alessandra Adami, PhD; Matthew Budoff, MD; Hans Fischer, MD; Janos Porszasz, MD, PhD; Harry Rossiter, PhD; William Stringer, MD

Michael E. DeBakey VAMC, Houston, TX: Amir Sharafkhaneh, MD, PhD; Charlie Lan, DO

Minneapolis VA: Christine Wendt, MD; Brian Bell, MD; Ken M. Kunisaki, MD, MS

National Jewish Health, Denver, CO: Russell Bowler, MD, PhD; David A. Lynch, MB

Reliant Medical Group, Worcester, MA: Richard Rosiello, MD; David Pace, MD

Temple University, Philadelphia, PA: Gerard Criner, MD; David Ciccolella, MD; Francis Cordova, MD; Chandra Dass, MD; Gilbert D'Alonzo, DO; Parag Desai, MD; Michael Jacobs, PharmD; Steven Kelsen, MD, PhD; Victor Kim, MD; A. James Mamary, MD; Nathaniel Marchetti, DO; Aditi Satti, MD; Kartik Shenoy, MD; Robert M. Steiner, MD; Alex Swift, MD; Irene Swift, MD; Maria Elena Vega-Sanchez, MD

University of Alabama, Birmingham, AL: Mark Dransfield, MD; William Bailey, MD; Surya P. Bhatt, MD; Anand Iyer, MD; Hrudaya Nath, MD; J. Michael Wells, MD

University of California, San Diego, CA: Douglas Conrad, MD; Xavier Soler, MD, PhD; Andrew Yen, MD

University of Iowa, Iowa City, IA: Alejandro P. Comellas, MD; Karin F. Hoth, PhD; John Newell, Jr., MD; Brad Thompson, MD

University of Michigan, Ann Arbor, MI: MeiLan K. Han, MD MS; Ella Kazerooni, MD MS; Wassim Labaki, MD MS; Craig Galban, PhD; Dharshan Vummidi, MD

University of Minnesota, Minneapolis, MN: Joanne Billings, MD; Abbie Begnaud, MD; Tadashi Allen, MD

University of Pittsburgh, Pittsburgh, PA: Frank Sciurba, MD; Jessica Bon, MD; Divay Chandra, MD, MSc; Carl Fuhrman, MD; Joel Weissfeld, MD, MPH

University of Texas Health, San Antonio, San Antonio, TX: Antonio Anzueto, MD; Sandra Adams, MD; Diego Maselli-Caceres, MD; Mario E. Ruiz, MD; Harjinder Singh

Table 1: Description of study subjects in COPDGene

	COPDGene		ECLIPSE	
	COPD Cases	**Smoking Controls**	**COPD Cases**	**Smoking Controls**
Number of subjects	3848	4691	1953	165
% Male	56.31	51.18	34.46	43.64
Age	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 Smokers	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)

Table 2: Burden testing in COPDGene. P-values and effect sizes for variant-set testing of *CFTR* variants with COPD and related phenotypes.

		CF-causing	**CF-causing + Varying clinical consequence**	**CF-causing + Varying clinical consequence + Predicted functional**	**All coding variants**	**Controls**
# of Variants		36	61	206	297[1]	
# of Subjects[3]		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 (OR)	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 (OR)	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 (OR)	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 (OR)	0.19	0.20	0.19	0.26	
BMI	*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	
FEV1 percent predicted	*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 % FEV1[2]	*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 (*beta=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	

1. Four variants with allele frequency > 5% (881 counts) were excluded from analysis
2. All -p-values are one-sided except for BDR which is two sided
3. Chronic bronchitis, severe exacerbation, and BMI data were unavailable for 2 subjects; COPD and severe COPD data were unavailable for 58 subjects; FEV1 percent predicted data was unavailable for 58 subjects; percent emphysema data was unavailable for 618 subjects; airway wall thickness data was unavailable for 619 subjects; bronchodilator response data was unavailable for 169 subjects; and bronchiectasis data was unavailable for 7209 subjects.

*Indicates p-values that are significant after correction for multiple comparisons ($p<0.05/10$ or 0.005). Odds ratios or beta coefficients are shown for all nominally significant associations ($p<0.05$).

Table 3: Single Variant testing of F508del for association with chronic bronchitis in COPDGene and ECLIPSE.

	COPDGene	ECLIPSE	Meta-analysis
Allele Counts	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

Table 4: Compound heterozygotes in COPDGene. Numbers of subjects identified who are compound heterozygotes for *CFTR* variants.

	Clinically significant or predicted to be functional[1]	Clinically significant or predicted to be functional + Varying clinical consequence[2]	All Compound Heterozygotes	Controls	One sided p-value for all compound heterozygotes[3]
Total number of 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	

[1]These 8 subjects all carry one copy of the F508del variant and one variant of unknown function according to CFTR2 that is predicted to have moderate effect according to SNPeff

[2]This group includes the 8 subjects from the first group, one subject that carries one F508del variant and one variant of varying clinical consequence, and 5 subjects that carry two variants of varying clinical consequence

[3] p-values were computed using Fishers exact test to test whether COPD, chronic bronchitis and bronchiectasis cases were enriched in all compound heterozygotes compared to controls. Statistical testing was not performed for the other two groups due to the small sample sizes.

* indicates p-values that are significant after correction for multiple comparisons ($p<0.05/3$)

Table 5: Burden testing of association between CF-causing variants and chronic bronchitis in ECLIPSE. P-values and effect sizes for association between CF-causing variants and chronic bronchitis in ECLIPSE and meta-analysis between ECLIPSE and COPDGene

	Number of variants	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

Figure 1: Workflow for 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) 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) greater than 10.

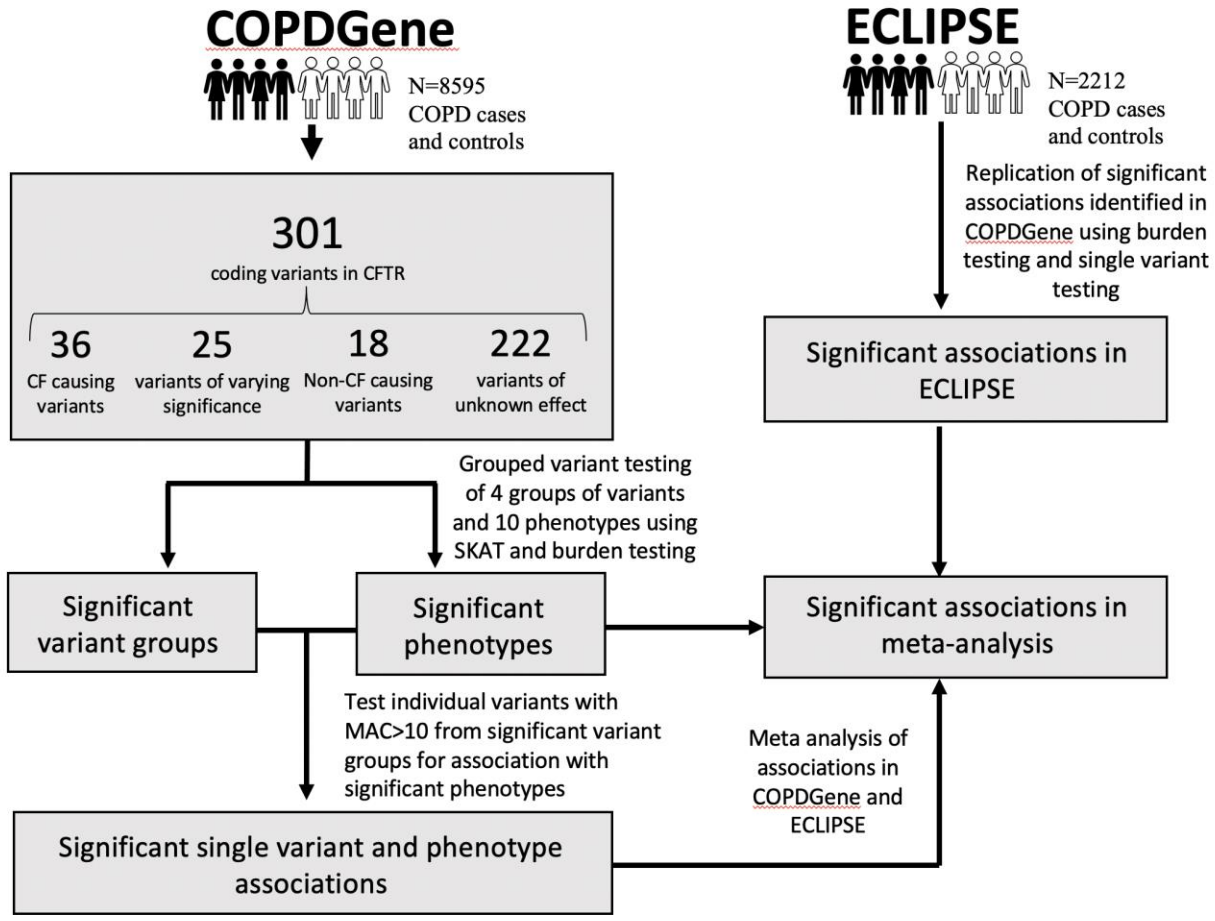
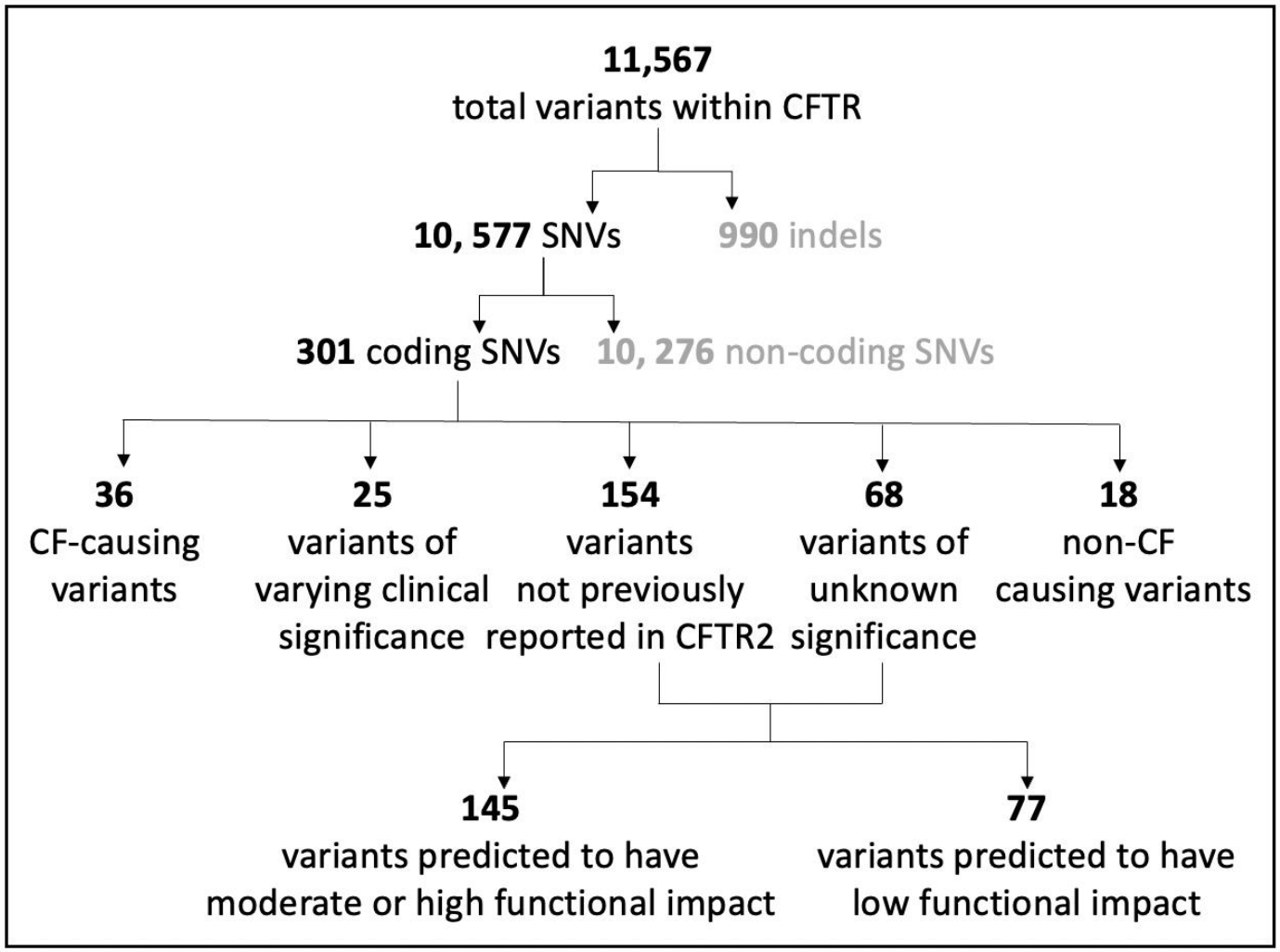


Figure 2: Breakdown of CFTR variants. A total of 301 coding SNVs were included in analysis.



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, Make BJ, Han MK, Martinez F, Criner GJ, National Emphysema Treatment Trial Research G. Severe chronic bronchitis in advanced emphysema increases mortality and hospitalizations. *COPD* 2013; 10: 667-678.
3. Sharer N, Schwarz M, Malone G, Howarth A, Painter J, Super M, Braganza J. 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, Knowles MR, Silverman LM, Jowell PS. 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, Bassas L, Lissens W, Silber S, Romey MC, Ruiz-Romero J, Verlingue C, Claustres M, 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., Greenberger PA, MacLean J, Walden SM, Slavin RG, Cutting GR. 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, Courville C, Tang LP, Backer K, Levin E, Raju SV, Li Y, Mazur M, Byan-Parker S, Grizzle W, Sorscher EJ, Dransfield MT, Rowe SM. 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, Courville C, Tidwell SL, Raju SV, Gaggar A, Steele C, Tang LP, Liu B, Rowe SM. 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, Ellis L, Dupuis A, Liao J, Zielenski J, Durie P. 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, Aleksandrov AA, Henderson AG, Zeman KL, Worthington EN, Gentzsch M, Kreda SM, Cholon D, Bennett WD, Riordan JR, Boucher RC, Tarran R. 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, MacWilliam L, Cole PJ. Effect of cigarette smoking on nasal mucociliary clearance and ciliary beat frequency. *Thorax* 1986; 41: 519-523.
12. Raju SV, Lin VY, Liu L, McNicholas CM, Karki S, Sloane PA, Tang L, Jackson PL, Wang W, Wilson L, Macon KJ, Mazur M, Kappes JC, DeLucas LJ, Barnes S, Kirk K, Tearney GJ, Rowe SM. 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, Raju SV, Reeves G, Acosta EP, Dransfield MT, Rowe SM. Pilot evaluation of ivacaftor for chronic bronchitis. *Lancet Respir Med* 2016; 4: e32-33.

14. Rowe SM, Jones I, Dransfield MT, Haque N, Gleason S, Hayes KA, Kulmatycki K, Yates DP, Danahay H, Gosling M, Rowlands DJ, Grant SS. 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, Tybjærg-Hansen A. 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, Nordestgaard BG. ΔF508 heterozygosity in cystic fibrosis and susceptibility to asthma. *Lancet* 1998; 351: 1911-1913.
17. Divac A, Nikolic A, Mitic-Milikic M, Nagorni-Obradovic L, Petrovic-Stanojevic N, Dopudja-Pantic V, Nadaskic R, Savic A, Radojkovic D. 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, Debaecker D, Kesteloot M, Lalau G, Roussel P. Abnormal distribution of cystic fibrosis DF508 allele in adults with chronic bronchial hypersecretion. *Lancet* 1990; 335: 1340.
19. Gervais R, Lafitte JJ, Dumur V, Kesteloot M, Lalau G, Houdret N, Roussel P. Sweat chloride and delta F508 mutation in chronic bronchitis or bronchiectasis. *Lancet* 1993; 342: 997.
20. Kostuch M, Semczuk A, Szarewicz-Adamczyk W, Gasowska-Giszczak U, Wojcierowski J, Kulczycki L. 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, Fang P, Oster RA, Dransfield MT, Rowe SM. Impact of heterozygote CFTR mutations in COPD patients with chronic bronchitis. *Respir Res* 2014; 15: 18.
22. Stankovic M, Nikolic A, Divac A, Tomovic A, Petrovic-Stanojevic N, Andjelic M, Dopudja-Pantic V, Surlan M, Vujicic I, Ponomarev D, Mitic-Milikic M, Kusic J, Radojkovic D. 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.
24. Miller AC, Comellas AP, Hornick DB, Stoltz DA, Cavanaugh JE, Gerke AK, Welsh MJ, Zabner J, Polgreen PM. Cystic fibrosis carriers are at increased risk for a wide range of cystic fibrosis-related conditions. *Proc Natl Acad Sci U S A* 2020; 117: 1621-1627.
25. Regan EA, Hokanson JE, Murphy JR, Make B, Lynch DA, Beaty TH, Curran-Everett D, Silverman EK, Crapo JD. Genetic epidemiology of COPD (COPDGene) study design. *COPD* 2010; 7: 32-43.
26. Vestbo J, Anderson W, Coxson HO, Crim C, Dawber F, Edwards L, Hagan G, Knobil K, Lomas DA, MacNee W, Silverman EK, Tal-Singer R, investigators E. Evaluation of COPD Longitudinally to Identify Predictive Surrogate End-points (ECLIPSE). *Eur Respir J* 2008; 31: 869-873.
27. Diaz AA, Maselli DJ, Rahaghi F, Come CE, Yen A, Maclean ES, Okajima Y, Martinez CH, Yamashiro T, Lynch DA, Wang W, Kinney GL, Washko GR, San Jose Estepar R. Pulmonary vascular pruning in smokers with bronchiectasis. *ERJ Open Res* 2018; 4.
28. Taliun D, Harris DN, Kessler MD, Carlson J, Szpiech ZA, Torres R, Taliun SAG, Corvelo A, Gogarten SM, Kang HM, Pitsillides AN, LeFaive J, Lee SB, Tian X, Browning BL, Das S, Emde AK, Clarke WE, Loesch DP, Shetty AC, Blackwell TW, Smith AV, Wong Q,

Liu X, Conomos MP, Bobo DM, Aguet F, Albert C, Alonso A, Ardlie KG, Arking DE, Aslibekyan S, Auer PL, Barnard J, Barr RG, Barwick L, Becker LC, Beer RL, Benjamin EJ, Bielak LF, Blangero J, Boehnke M, Bowden DW, Brody JA, Burchard EG, Cade BE, Casella JF, Chalazan B, Chasman DI, Chen YI, Cho MH, Choi SH, Chung MK, Clish CB, Correa A, Curran JE, Custer B, Darbar D, Daya M, de Andrade M, DeMeo DL, Dutcher SK, Ellinor PT, Emery LS, Eng C, Fatkin D, Fingerlin T, Forer L, Fornage M, Franceschini N, Fuchsberger C, Fullerton SM, Germer S, Gladwin MT, Gottlieb DJ, Guo X, Hall ME, He J, Heard-Costa NL, Heckbert SR, Irvin MR, Johnsen JM, Johnson AD, Kaplan R, Kardia SLR, Kelly T, Kelly S, Kenny EE, Kiel DP, Klemmer R, Konkle BA, Kooperberg C, Kottgen A, Lange LA, Lasky-Su J, Levy D, Lin X, Lin KH, Liu C, Loos RJF, Garman L, Gerszten R, Lubitz SA, Lunetta KL, Mak ACY, Manichaikul A, Manning AK, Mathias RA, McManus DD, McGarvey ST, Meigs JB, Meyers DA, Mikulla JL, Minear MA, Mitchell BD, Mohanty S, Montasser ME, Montgomery C, Morrison AC, Murabito JM, Natale A, Natarajan P, Nelson SC, North KE, O'Connell JR, Palmer ND, Pankratz N, Peloso GM, Peyser PA, Pleiness J, Post WS, Psaty BM, Rao DC, Redline S, Reiner AP, Roden D, Rotter JI, Ruczinski I, Sarnowski C, Schoenherr S, Schwartz DA, Seo JS, Seshadri S, Sheehan VA, Sheu WH, Shoemaker MB, Smith NL, Smith JA, Sotoodehnia N, Stilp AM, Tang W, Taylor KD, Telen M, Thornton TA, Tracy RP, Van Den Berg DJ, Vasan RS, Viaud-Martinez KA, Vrieze S, Weeks DE, Weir BS, Weiss ST, Weng LC, Willer CJ, Zhang Y, Zhao X, Arnett DK, Ashley-Koch AE, Barnes KC, Boerwinkle E, Gabriel S, Gibbs R, Rice KM, Rich SS, Silverman EK, Qasba P, Gan W, Consortium NT-OfPM, Papanicolaou GJ, Nickerson DA, Browning SR, Zody MC, Zollner S, Wilson JG, Cupples LA, Laurie CC, Jaquish CE, Hernandez RD, O'Connor TD, Abecasis GR. 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, Johnson AD, Brody JA, Li AH, Huang Z, Carroll A, Wei P, Gibbs R, Klein RJ, Boerwinkle E. WGSA: an annotation pipeline for human genome sequencing studies. *J Med Genet* 2016; 53: 111-112.
31. Prokopenko D, Hecker J, Silverman EK, Pagano M, Nothen MM, Dina C, Lange C, Fier HL. 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, Qiao D, Parker MM, McDonald MN, Manichaikul A, Rich SS, Barr RG, Williams CJ, Brantly ML, Lange C, Beaty TH, Crapo JD, Silverman EK, Cho MH. 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, Barnes KC, Rieder MJ, Nickerson DA, Team NGESP-ELP, Christiani DC, Wurfel MM, Lin X. 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, Hubert D, Crestani B, Bassinet L, Dusser D, Fajac I. 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, Nguyen AT, Bessou G, Lenoir G, Descamps-Latscha B. Disturbed myeloperoxidase-dependent activity of neutrophils in cystic fibrosis homozygotes and heterozygotes, and its correction by amiloride. *Journal of immunology* 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. *Journal of innate immunity* 2010; 2: 260-266.
37. Schneider A, Larusch J, Sun X, Aloe A, Lamb J, Hawes R, Cotton P, Brand RE, Anderson MA, Money ME, Banks PA, Lewis MD, Baillie J, Sherman S, Disario J, Burton FR, Gardner TB, Amann ST, Gelrud A, George R, Rockacy MJ, Kassabian S, Martinson J, Slivka A, Yadav D, Oruc N, Barmada MM, Frizzell R, Whitcomb DC. 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, Barnes PJ, Buist SA, Calverley P, Fukuchi Y, Jenkins C, Rodriguez-Roisin R, van Weel C, Zielinski J, Global Initiative for Chronic Obstructive Lung D. 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.

HGVS ID	Variant ID	Legacy Name	CFTR2 Determination	SnpEff predicted effect	SnpEff impact	Pancreatic Insufficiency	Number of alleles in COPDGene
c.1408G>A	rs213950	V470M	Non CF-causing	missense_variant	MODERATE		9581
c.2562T>G	chr7:117595001:T:G		Not reported in CFTR2	synonymous_variant	LOW		7315
c.4389G>A	rs1800136	4521G/A	Not Evaluated	synonymous_variant	LOW		4037
c.3870A>G	rs1800130	P1290P	Not Evaluated	synonymous_variant	LOW		1181
c.224G>A	chr7:117509093:G:A	R75Q	Non CF-causing	missense_variant	MODERATE	0.28	459
c.1584G>A	chr7:117559655:G:A	1716G/A	Non CF-causing	splice_region_variant&synonyn	LOW	0.67	290
c.1521_1523delCTT	rs199826652	rs1800136	CF-causing	disruptive_inframe_deletion	MODERATE	0.98	177
c.2898G>A	rs1800109	3030G/A	Not Evaluated	synonymous_variant	LOW		166
c.4272C>T	rs1800135	4404C/T	Not Evaluated	synonymous_variant	LOW		136
c.3705T>G	rs34911792	S1235R	Non CF-causing	missense_variant	MODERATE	0.38	131
c.3285A>T	rs1800118	3417A/T	Not Evaluated	synonymous_variant	LOW		109
c.2002C>T	rs1800100	R668C	Non CF-causing	missense_variant	MODERATE	0.44	105
c.3808G>A	chr7:117642528:G:A	D1270N	Varying clinical consequence	missense_variant	MODERATE	0.17	86
c.1727G>C	chr7:117590400:G:C	G576A	Non CF-causing	missense_variant	MODERATE	0.33	85
c.220C>T	rs115545701	R74W	Varying clinical consequence	missense_variant	MODERATE	0.15	82
c.1581A>G	rs1800094	1713A/G	Not reported in CFTR2	synonymous_variant	LOW		78
c.2900T>C	rs1800110	L967S	Varying clinical consequence	missense_variant	MODERATE	0	29
c.2991G>C	rs1800111	L997F	Non CF-causing	missense_variant&splice_regio	MODERATE	0.32	29
c.91C>T	chr7:117504290:C:T	R31C	Non CF-causing	missense_variant	MODERATE	0.3	28
c.2260G>A	rs150157202	V754M	Non CF-causing	missense_variant	MODERATE	0.73	28
c.350G>A	chr7:117530975:G:A	R117H	Varying clinical consequence	missense_variant	MODERATE	0.23	27
c.2820T>G	rs60887846		Not reported in CFTR2	synonymous_variant	LOW		25
c.4243-5C>T	chr7:117666903:C:T	4375-5C->T	Not Evaluated	splice_region_variant&intron_v	LOW		23
c.443T>C	chr7:117531068:T:C	I148T	Non CF-causing	missense_variant	MODERATE	0.88	22
c.853A>T	rs151073129	I285F	Not Evaluated	missense_variant	MODERATE		21
c.1365G>A	chr7:117548796:G:A		Not reported in CFTR2	synonymous_variant	LOW		21
c.1523T>G	chr7:117559594:T:G	F508C	Non CF-causing	missense_variant	MODERATE	0.6	19
c.3897A>G	rs1800131	4029A/G	Not Evaluated	synonymous_variant	LOW		17
c.3485G>T	chr7:117627538:G:T	R1162L	Non CF-causing	missense_variant	MODERATE	0.38	17
c.890G>A	rs143486492	R297Q	Not Evaluated	missense_variant	MODERATE		17
c.2735C>T	chr7:117603609:C:T	S912L	Unknown significance	missense_variant	MODERATE	0.6	16
c.2245C>T	rs151235408	2377C/T	Not Evaluated	synonymous_variant	LOW		14
c.3454G>C	rs75541969	D1152H	Varying clinical consequence	missense_variant	MODERATE	0.24	14
c.2249C>T	rs140455771	P750L	Varying clinical consequence	missense_variant	MODERATE	0.33	14
c.274-6T>C	rs371315549	406-6T->C	Not Evaluated	splice_region_variant&intron_v	LOW		13
c.509G>A	rs1800079	R170H	Non CF-causing	missense_variant	MODERATE	0.33	12
c.4333G>A	rs148783445	D1445N	Not Evaluated	missense_variant	MODERATE		11
c.3080T>C	rs1800112	I1027T	Non CF-causing	missense_variant	MODERATE	0.7	11
c.221G>A	rs142540482	R74Q	Not Evaluated	missense_variant	MODERATE		11
c.2421A>G	rs1800103	I807M	Non CF-causing	missense_variant	MODERATE	0	10
c.2988+1G>A	rs75096551	3120+1G->A	CF-causing	splice_donor_variant&intron_v	HIGH	0.98	8
c.2079T>G	rs145540754	F693L(TTG)	Not Evaluated	missense_variant	MODERATE		8
c.1652G>A	rs75527207	G551D	CF-causing	missense_variant	MODERATE	0.96	8
c.1865G>A	rs121908759	G622D	Varying clinical consequence	missense_variant	MODERATE	0	8
c.418C>T	rs145900055	P140S	Not Evaluated	missense_variant	MODERATE		8
c.601G>A	rs138338446	V201M	Unknown significance	missense_variant	MODERATE	0.25	8
c.1046C>T	rs121909021	A349V	Unknown significance	missense_variant	MODERATE	0.33	7
c.3558A>G	rs1800121	Q1186Q (3690A/G)	Not reported in CFTR2	synonymous_variant	LOW		7
c.1624G>T	rs113993959	G542X	CF-causing	stop_gained	HIGH	0.98	6
c.92G>T	chr7:117504291:G:T	R31L	Unknown significance	missense_variant	MODERATE	0	6
c.31G>A	chr7:117480125:G:A	V11I	Not Evaluated	missense_variant	MODERATE		6
c.650A>G	rs121909046	E217G	Not Evaluated	missense_variant	MODERATE		5
c.3205G>A	rs200321110	G1069R	Varying clinical consequence	missense_variant	MODERATE	0.67	5
c.617T>G	rs121908752	L206W	CF-causing	missense_variant	MODERATE	0.2	5
c.2855T>C	rs142773283	M952T	Unknown significance	missense_variant	MODERATE	0	5
c.1052C>G	chr7:117540282:C:G	T351S	Not Evaluated	missense_variant	MODERATE		5
c.806T>C	rs201016820		Not reported in CFTR2	missense_variant	MODERATE		5
c.-4282G>T	rs530099256		Not reported in CFTR2	intergenic_region	MODIFIER		5
c.3204C>T	rs1800116	3336C/T	Not reported in CFTR2	synonymous_variant	LOW		4
c.489+3A>G	rs377729736	621+3A->G	Varying clinical consequence	splice_region_variant&intron_v	LOW	0.25	4
c.1364C>A	chr7:117548795:C:A	A455E	CF-causing	missense_variant	MODERATE	0.34	4
c.2506G>T	rs201386642	D836Y	Non CF-causing	missense_variant	MODERATE	0.5	4
c.3154T>G	rs150212784	F1052V	Varying clinical consequence	missense_variant	MODERATE	0.15	4
c.3151A>G	chr7:117611592:A:G	I1051V	Not reported in CFTR2	missense_variant	MODERATE		4
c.3909C>G	rs80034486	N1303K	CF-causing	missense_variant	MODERATE	0.98	4
c.3485G>A	chr7:117627538:G:A	R1162Q	Not Evaluated	missense_variant	MODERATE		4
c.1657C>T	chr7:117587811:C:T	R553X	CF-causing	stop_gained	HIGH	0.97	4
c.1558G>A	chr7:117559629:G:A	V520I	Not Evaluated	missense_variant	MODERATE		4
c.1684G>A	chr7:117590357:G:A	V562I	Non CF-causing	missense_variant	MODERATE	0.43	4
c.902A>G	rs150691494	Y301C	Not Evaluated	missense_variant	MODERATE		4
c.1584+5130C>A	chr7:117564785:C:A		Not reported in CFTR2	intron_variant	MODIFIER		4
c.2153C>G	rs142432539		Not reported in CFTR2	missense_variant	MODERATE		4
c.4092G>A	rs148878126		Not reported in CFTR2	synonymous_variant	LOW		4
c.663G>A	rs758147990	795G/A	Not reported in CFTR2	synonymous_variant	LOW		3
c.137C>T	chr7:117504336:C:T	A46V	Not reported in CFTR2	missense_variant	MODERATE		3
c.1516A>G	chr7:117559587:A:G	I506V	Not Evaluated	missense_variant	MODERATE		3
c.1519_1521delATC	rs763199062	I507del	CF-causing	inframe_deletion	MODERATE	0.98	3
c.3209G>A	chr7:117611650:G:A	R1070Q	Varying clinical consequence	missense_variant	MODERATE	0.86	3
c.3484C>T	rs74767530	R1162X	CF-causing	stop_gained	HIGH	0.97	3
c.1001G>A	chr7:117540231:G:A	R334Q	Varying clinical consequence	missense_variant	MODERATE	0	3
c.1454G>C	rs143980575	S485T	Not reported in CFTR2	missense_variant	MODERATE		3
c.3041A>G	rs149279509	Y1014C	Unknown significance	missense_variant	MODERATE	0	3
c.1680-6T>G	chr7:117590347:T:G		Not reported in CFTR2	splice_region_variant&intron_v	LOW		3
c.4197C>G	chr7:117665519:C:G		Not reported in CFTR2	synonymous_variant	LOW		3
c.2450G>T	rs148604667		Not reported in CFTR2	missense_variant	MODERATE		3
c.3429G>A	rs375845215		Not reported in CFTR2	synonymous_variant	LOW		3
c.4108G>C	rs760336091		Not reported in CFTR2	missense_variant	MODERATE		3
c.3033A>G	rs773752573		Not reported in CFTR2	synonymous_variant	LOW		3
c.1519A>G	rs1801178	1651A/G	Not reported in CFTR2	missense_variant	MODERATE		2
c.2052delA	rs777301769	2184delA	CF-causing	frameshift_variant	HIGH	0.98	2
c.2620-6T>C	rs371315682	2752-6T->C	Not Evaluated	splice_region_variant&intron_v	LOW		2
c.313delA	rs779091180	444delA	CF-causing	frameshift_variant	HIGH	1	2
c.489+1G>T	rs78756941	621+1G->T	CF-causing	splice_donor_variant&intron_v	HIGH	0.99	2
c.3200C>T	chr7:117611641:C:T	A1067V	Not Evaluated	missense_variant	MODERATE		2
c.3503A>G	rs150326506	D1168G	Not reported in CFTR2	missense_variant	MODERATE		2

c.2770G>A	rs201759207	D924N	Unknown significance	missense_variant	MODERATE	0	2
c.846A>T	rs142864834	E282D	Not Evaluated	missense_variant	MODERATE		2
c.948T>G	rs78742051	F316L	Not reported in CFTR2	missense_variant	MODERATE		2
c.254G>A	chr7:117509123:G:A	G85E	CF-causing	missense_variant	MODERATE	0.85	2
c.3415A>G	rs397508556	I1139V	Not Evaluated	missense_variant	MODERATE		2
c.1853T>C	rs139468767	I618T	Varying clinical consequence	missense_variant	MODERATE	0	2
c.202A>G	rs397508332	K68E	Not Evaluated	missense_variant	MODERATE		2
c.1247A>G	rs777850419	N416S	Not reported in CFTR2	missense_variant	MODERATE		2
c.772A>G	rs191456345	R258G	Varying clinical consequence	missense_variant	MODERATE	0.5	2
c.889C>T	chr7:117540119:C:T	R297W	Not reported in CFTR2	missense_variant	MODERATE		2
c.1163C>T	rs143860237	T388M	Not Evaluated	missense_variant	MODERATE		2
c.3322G>C	rs397508542	V1108L	Not reported in CFTR2	missense_variant	MODERATE		2
c.2758G>A	chr7:117603632:G:A	V920M	Not Evaluated	missense_variant	MODERATE		2
c.2813T>G	rs193922511	V938G	Not Evaluated	missense_variant	MODERATE		2
c.3846G>A	chr7:117642566:G:A	W1282X	CF-causing	stop_gained	HIGH	0.99	2
c.2739T>A	rs149790377	Y913X	CF-causing	stop_gained	HIGH	0.88	2
c.1480T>G	chr7:117559551:T:G		Not reported in CFTR2	missense_variant	MODERATE|HIGH		2
c.2557A>G	chr7:117594996:A:G		Not reported in CFTR2	missense_variant	MODERATE		2
c.1920T>C	rs145877746		Not reported in CFTR2	synonymous_variant	LOW		2
c.4232A>C	rs150177304		Not reported in CFTR2	missense_variant	MODERATE		2
c.1210G>C	rs200899224		Not reported in CFTR2	missense_variant&splice_regio	MODERATE		2
c.1734A>G	rs201025424		Not reported in CFTR2	synonymous_variant	LOW		2
c.1691A>G	rs375325315		Not reported in CFTR2	missense_variant	MODERATE		2
c.2665C>T	rs61738523		Not reported in CFTR2	missense_variant	MODERATE		2
c.2917C>T	rs747139295		Not reported in CFTR2	missense_variant	MODERATE		2
c.2916T>A	rs773273576		Not reported in CFTR2	synonymous_variant	LOW		2
c.837A>T	rs773509355		Not reported in CFTR2	missense_variant	MODERATE		2
c.1116+1G>C	chr7:117540347:G:C	1248+1G->C	Not Evaluated	splice_donor_variant&intron_v	HIGH		1
c.1209+6A>G	rs749054857	1341+6A->G	Not reported in CFTR2	splice_region_variant&intron_v	LOW		1
c.1585-1G>A	rs76713772	1717-1G->A	CF-causing	splice_acceptor_variant&intror	HIGH	0.97	1
c.1680-1G>A	rs121908794	1812-1G->A	CF-causing	splice_acceptor_variant&intror	HIGH	1	1
c.2052dupA	rs746460279	2184insA	CF-causing	frameshift_variant	HIGH	0.85	1
c.2657+5G>A	rs80224560	2789+5G->A	CF-causing	splice_region_variant&intron_v	LOW	0.43	1
c.147A>G	chr7:117504346:A:G	279A/G	Not reported in CFTR2	synonymous_variant	LOW		1
c.165-3C>T	chr7:117509031:C:T	297-3C->T	Unknown significance	splice_region_variant&intron_v	LOW	1	1
c.2988G>A	rs121908797	3120G->A	CF-causing	splice_region_variant&synonyn	LOW	0.55	1
c.3067_3072delATAGTG	rs397508492	3199del6	CF-causing	inframe_deletion	MODERATE		1
c.233dupT	rs397508360	365-366insT	CF-causing	frameshift_variant	HIGH	1	1
c.3564G>A	rs146804928	3696G/A	Not reported in CFTR2	synonymous_variant	LOW		1
c.3807C>T	chr7:117642527:C:T	3939C/T	Not reported in CFTR2	synonymous_variant	LOW		1
c.262_263delTT	rs754147777	394delTT	CF-causing	frameshift_variant	HIGH	0.97	1
c.3874-4A>G	rs201381687	4006-4A->G	Not reported in CFTR2	splice_region_variant&intron_v	LOW		1
c.273+4A>G	rs387906374	405+4A->G	Not Evaluated	splice_region_variant&intron_v	LOW		1
c.3964-6C>T	chr7:117664682:C:T	4096-6C>T	Not reported in CFTR2	splice_region_variant&intron_v	LOW		1
c.489+8T>G	chr7:117531122:T:G	621+8T->G	Not Evaluated	splice_region_variant&intron_v	LOW		1
c.3017C>A	rs397508480	A1006E	CF-causing	missense_variant	MODERATE	0.2	1
c.3025G>A	rs184724618	A1009T	Not reported in CFTR2	missense_variant	MODERATE		1
c.358G>A	chr7:117530983:G:A	A120T	Varying clinical consequence	missense_variant	MODERATE	0.33	1
c.925G>A	rs148013312	A309T	Not reported in CFTR2	missense_variant	MODERATE		1
c.1675G>A	rs75549581	A559T	CF-causing	missense_variant	MODERATE	0.98	1
c.26C>T	chr7:117480120:C:T	A9V	Not Evaluated	missense_variant	MODERATE		1
c.384C>A	chr7:117531009:C:A	C128X	Not Evaluated	stop_gained	HIGH		1
c.2597G>A	rs193922506	C866Y	Not Evaluated	missense_variant	MODERATE		1
c.328G>C	chr7:117530953:G:C	D110H	CF-causing	missense_variant	MODERATE	0.17	1
c.4297G>A	rs750559671	E1433K	Not Evaluated	missense_variant	MODERATE		1
c.3297C>A	rs747754623	F1099L	Varying clinical consequence	missense_variant	MODERATE	0.25	1
c.571T>G	rs141482808	F191V	CF-causing	missense_variant	MODERATE	0.25	1
c.3517G>A	rs368393738	G1173S	Not Evaluated	missense_variant	MODERATE		1
c.3893G>T	chr7:117652861:G:T	G1298V	Not Evaluated	missense_variant	MODERATE		1
c.638G>A	rs775701644	G213E	Not Evaluated	missense_variant	MODERATE		1
c.715G>A	rs397508788	G239R	Not Evaluated	missense_variant	MODERATE		1
c.4123C>A	rs146947665	H1375N	Not Evaluated	missense_variant	MODERATE		1
c.355A>G	rs193922518	I119V	Not reported in CFTR2	missense_variant	MODERATE		1
c.3238A>C	rs766126240	K1080Q	Not reported in CFTR2	missense_variant	MODERATE		1
c.2851A>G	rs181137679	K951E	Not Evaluated	missense_variant	MODERATE		1
c.3177A>G	rs1800113	L1059L (3309A/G)	Not reported in CFTR2	synonymous_variant	LOW		1
c.3230T>C	rs139304906	L1077P	CF-causing	missense_variant	MODERATE	0.93	1
c.4003C>T	rs145545286	L1335F	Not reported in CFTR2	missense_variant	MODERATE		1
c.958T>G	rs144476686	L320V	Non CF-causing	missense_variant	MODERATE	0	1
c.1125A>C	rs73215912	L375F	Not reported in CFTR2	missense_variant	MODERATE		1
c.1399C>T	rs1800089	L467F	Not Evaluated	missense_variant	MODERATE		1
c.3409A>G	rs397508553	M1137V	Not Evaluated	missense_variant	MODERATE		1
c.454A>T	chr7:117531079:A:T	M152L	Not Evaluated	missense_variant	MODERATE		1
c.794T>G	rs148519623	M265R	Varying clinical consequence	missense_variant	MODERATE	0.83	1
c.2856G>C	chr7:117603730:G:C	M952I	Not Evaluated	missense_variant	MODERATE		1
c.1253A>G	rs397508185	N418S	Not Evaluated	missense_variant	MODERATE		1
c.3038C>T	chr7:117610568:C:T	P1013L	Not Evaluated	missense_variant	MODERATE		1
c.332C>T	rs140502196	P111L	Not Evaluated	missense_variant	MODERATE		1
c.14C>T	rs193922501	P5L	Varying clinical consequence	missense_variant	MODERATE	0.1	1
c.3713A>G	rs397508594	Q1238R	Not reported in CFTR2	missense_variant	MODERATE		1
c.4054C>G	chr7:117664778:C:G	Q1352E	Not reported in CFTR2	missense_variant	MODERATE		1
c.4426C>T	rs374705585	Q1476X	Varying clinical consequence	stop_gained	HIGH	0.18	1
c.451C>A	chr7:117531076:C:A	Q151K	Not Evaluated	missense_variant	MODERATE		1
c.535C>A	chr7:117534321:C:A	Q179K	Not Evaluated	missense_variant	MODERATE		1
c.3197G>A	chr7:117611638:G:A	R1066H	CF-causing	missense_variant	MODERATE	0.33	1
c.3208C>T	rs202179988	R1070W	Varying clinical consequence	missense_variant	MODERATE	0.34	1
c.349C>T	chr7:117530974:C:T	R117C	CF-causing	missense_variant	MODERATE	0.24	1
c.349C>G	chr7:117530974:C:G	R117G	Varying clinical consequence	missense_variant	MODERATE	0.13	1
c.350G>T	chr7:117530975:G:T	R117L	Varying clinical consequence	missense_variant	MODERATE	0.13	1
c.4357C>T	chr7:117667022:C:T	R1453W	Not reported in CFTR2	missense_variant	MODERATE		1
c.1040G>A	chr7:117540270:G:A	R347H	CF-causing	missense_variant	MODERATE	0.24	1
c.224G>T	chr7:117509093:G:T	R75L	Not Evaluated	missense_variant	MODERATE		1
c.2353C>T	rs374946172	R785X	CF-causing	stop_gained	HIGH	1	1
c.2374C>G	chr7:117592541:C:G	R792G	Not Evaluated	missense_variant	MODERATE		1
c.2375G>A	rs369040061	R792Q	Not Evaluated	missense_variant	MODERATE		1
c.2428A>G	rs377447726	R810G	Not Evaluated	missense_variant	MODERATE		1
c.3353C>G	chr7:117611794:C:G	S1118C	Not Evaluated	missense_variant	MODERATE		1
c.4276T>C	rs397508708	S1426P	Not reported in CFTR2	missense_variant	MODERATE		1

c.4364C>G	rs121909043	S1455X	Varying clinical consequence	stop_gained	HIGH	0.1	1
c.125C>T	rs143456784	S42F	Not Evaluated	missense_variant	MODERATE		1
c.149C>A	rs397508220	S50Y	Not reported in CFTR2	missense_variant	MODERATE		1
c.1647T>G	rs121909005	S549R(T->G)	CF-causing	missense_variant	MODERATE		1
c.2930C>T	rs141033578	S977F	Varying clinical consequence	missense_variant	MODERATE	0.25	1
c.3737C>T	rs397508600	T1246I	Varying clinical consequence	missense_variant	MODERATE	0.45	1
c.3896C>T	rs397508634	T1299I	Not Evaluated	missense_variant	MODERATE		1
c.490A>G	rs200885306	T164A	Not reported in CFTR2	missense_variant&splice_regio	MODERATE		1
c.1811C>G	chr7:117591978:C:G	T604S	Not reported in CFTR2	missense_variant	MODERATE		1
c.2723C>A	rs369521395	T908N	Not Evaluated	missense_variant	MODERATE		1
c.695T>A	rs397508783	V232D	CF-causing	missense_variant	MODERATE	0.13	1
c.2537G>A	rs397508393	W846X	CF-causing	stop_gained	HIGH		1
c.1731C>T	rs55928397	Y577Y (1863C/T)	Not reported in CFTR2	synonymous_variant	LOW		1
c.-4287G>A	chr7:117475808:G:A		Not reported in CFTR2	intergenic_region	MODIFIER		1
c.-4280G>A	chr7:117475815:G:A		Not reported in CFTR2	intergenic_region	MODIFIER		1
c.40A>G	chr7:117480134:A:G		Not reported in CFTR2	missense_variant	MODERATE		1
c.70T>G	chr7:117504269:T:G		Not reported in CFTR2	missense_variant	MODERATE		1
c.127G>A	chr7:117504326:G:A		Not reported in CFTR2	missense_variant	MODERATE		1
c.336T>C	chr7:117530961:T:C		Not reported in CFTR2	synonymous_variant	LOW		1
c.377G>C	chr7:117531002:G:C		Not reported in CFTR2	missense_variant	MODERATE		1
c.393T>C	chr7:117531018:T:C		Not reported in CFTR2	synonymous_variant	LOW		1
c.472A>G	chr7:117531097:A:G		Not reported in CFTR2	missense_variant	MODERATE		1
c.492T>A	chr7:117534278:T:A		Not reported in CFTR2	splice_region_variant&synonyn	LOW		1
c.529A>T	chr7:117534315:A:T		Not reported in CFTR2	missense_variant	MODERATE		1
c.603G>A	chr7:117535271:G:A		Not reported in CFTR2	synonymous_variant	LOW		1
c.640C>T	chr7:117535308:C:T		Not reported in CFTR2	synonymous_variant	LOW		1
c.912C>T	chr7:117540142:C:T		Not reported in CFTR2	synonymous_variant	LOW		1
c.1079C>T	chr7:117540309:C:T		Not reported in CFTR2	missense_variant	MODERATE		1
c.1164G>A	chr7:117542063:G:A		Not reported in CFTR2	synonymous_variant	LOW		1
c.1232A>G	chr7:117548663:A:G		Not reported in CFTR2	missense_variant	MODERATE		1
c.1301C>T	chr7:117548732:C:T		Not reported in CFTR2	missense_variant	MODERATE		1
c.1366G>C	chr7:117548797:G:C		Not reported in CFTR2	missense_variant	MODERATE		1
c.1392+11delT	chr7:117548830:CT:C		Not reported in CFTR2	intron_variant	MODIFIER		1
c.1393-6T>C	chr7:117559458:T:C		Not reported in CFTR2	splice_region_variant&intron_v	LOW		1
c.1446T>A	chr7:117559517:T:A		Not reported in CFTR2	synonymous_variant	LOW		1
c.1484C>T	chr7:117559555:C:T		Not reported in CFTR2	missense_variant	MODERATE		1
c.1522T>G	chr7:117559593:T:G		Not reported in CFTR2	missense_variant	MODERATE		1
c.1584+5016T>C	chr7:117564671:T:C		Not reported in CFTR2	intron_variant	MODIFIER		1
c.1614T>A	chr7:117587768:T:A		Not reported in CFTR2	missense_variant	MODERATE		1
c.1785G>T	chr7:117591952:G:T		Not reported in CFTR2	missense_variant	MODERATE		1
c.1804T>C	chr7:117591971:T:C		Not reported in CFTR2	synonymous_variant	LOW		1
c.1830A>T	chr7:117591997:A:T		Not reported in CFTR2	missense_variant	MODERATE		1
c.2078T>C	chr7:117592245:T:C		Not reported in CFTR2	missense_variant	MODERATE		1
c.2118C>T	chr7:117592285:C:T		Not reported in CFTR2	synonymous_variant	LOW		1
c.2229G>A	chr7:117592396:G:A		Not reported in CFTR2	synonymous_variant	LOW		1
c.2305G>A	chr7:117592472:G:A		Not reported in CFTR2	missense_variant	MODERATE		1
c.2354G>A	chr7:117592521:G:A		Not reported in CFTR2	missense_variant	MODERATE		1
c.2757C>T	chr7:117603631:C:T		Not reported in CFTR2	synonymous_variant	LOW		1
c.2780T>A	chr7:117603654:T:A		Not reported in CFTR2	missense_variant	MODERATE		1
c.2931C>T	chr7:117606696:C:T		Not reported in CFTR2	synonymous_variant	LOW		1
c.2974T>G	chr7:117606739:T:G		Not reported in CFTR2	missense_variant	MODERATE		1
c.3125A>G	chr7:117610655:A:G		Not reported in CFTR2	missense_variant	MODERATE		1
c.3140-4A>T	chr7:117611577:A:T		Not reported in CFTR2	splice_region_variant&intron_v	LOW		1
c.3308T>C	chr7:117611749:T:C		Not reported in CFTR2	missense_variant	MODERATE		1
c.3543A>G	chr7:117627596:A:G		Not reported in CFTR2	synonymous_variant	LOW		1
c.3588A>T	chr7:117627641:A:T		Not reported in CFTR2	synonymous_variant	LOW		1
c.3592G>A	chr7:117627645:G:A		Not reported in CFTR2	missense_variant	MODERATE		1
c.3637A>C	chr7:117627690:A:C		Not reported in CFTR2	missense_variant	MODERATE		1
c.3762A>G	chr7:117642482:A:G		Not reported in CFTR2	synonymous_variant	LOW		1
c.3933T>G	chr7:117652901:T:G		Not reported in CFTR2	missense_variant	MODERATE		1
c.4028G>T	chr7:117664752:G:T		Not reported in CFTR2	missense_variant	MODERATE		1
c.4142A>G	chr7:117665464:A:G		Not reported in CFTR2	missense_variant	MODERATE		1
c.4144C>A	chr7:117665466:C:A		Not reported in CFTR2	missense_variant	MODERATE		1
c.4181A>G	chr7:117665503:A:G		Not reported in CFTR2	missense_variant	MODERATE		1
c.4243-3T>C	chr7:117666905:T:C		Not reported in CFTR2	splice_region_variant&intron_v	LOW		1
c.4313G>A	chr7:117666978:G:A		Not reported in CFTR2	missense_variant	MODERATE		1
c.4320C>T	chr7:117666985:C:T		Not reported in CFTR2	synonymous_variant	LOW		1
c.4336A>C	chr7:117667001:A:C		Not reported in CFTR2	synonymous_variant	LOW		1
c.4340T>A	chr7:117667005:T:A		Not reported in CFTR2	missense_variant	MODERATE		1
c.4367G>A	chr7:117667032:G:A		Not reported in CFTR2	missense_variant	MODERATE		1
c.3521A>C	rs137875514		Not reported in CFTR2	missense_variant	MODERATE		1
c.3180A>G	rs142526976		Not reported in CFTR2	synonymous_variant	LOW		1
c.2424T>C	rs143954792		Not reported in CFTR2	synonymous_variant	LOW		1
c.3064G>A	rs144441835		Not reported in CFTR2	missense_variant	MODERATE		1
c.2708A>G	rs147297080		Not reported in CFTR2	missense_variant	MODERATE		1
c.264A>C	rs149662778		Not reported in CFTR2	missense_variant	MODERATE		1
c.2003G>A	rs199623561		Not reported in CFTR2	missense_variant	MODERATE		1
c.2502T>G	rs200735475		Not reported in CFTR2	missense_variant	MODERATE		1
c.1301C>A	rs367934560		Not reported in CFTR2	stop_gained	HIGH		1
c.4206C>T	rs368044495		Not reported in CFTR2	synonymous_variant	LOW		1
c.310A>G	rs369715785		Not reported in CFTR2	missense_variant	MODERATE		1
c.1690A>G	rs371291116		Not reported in CFTR2	missense_variant	MODERATE		1
c.3256A>T	rs373043500		Not reported in CFTR2	missense_variant	MODERATE		1
c.538C>G	rs374163420		Not reported in CFTR2	missense_variant	MODERATE		1
c.1340A>G	rs746941790		Not reported in CFTR2	missense_variant	MODERATE		1
c.2192C>T	rs748634753		Not reported in CFTR2	missense_variant	MODERATE		1
c.4363T>G	rs748845320		Not reported in CFTR2	missense_variant	MODERATE		1
c.66A>G	rs748899228		Not reported in CFTR2	synonymous_variant	LOW		1
c.3650C>T	rs749662161		Not reported in CFTR2	missense_variant	MODERATE		1
c.1629A>G	rs751730446		Not reported in CFTR2	synonymous_variant	LOW		1
c.4064G>T	rs755028771		Not reported in CFTR2	missense_variant	MODERATE		1
c.589T>C	rs755619078		Not reported in CFTR2	synonymous_variant	LOW		1
c.3984A>G	rs755917129		Not reported in CFTR2	missense_variant	MODERATE		1
c.2505T>C	rs761043298		Not reported in CFTR2	synonymous_variant	LOW		1
c.29G>A	rs762241850		Not reported in CFTR2	missense_variant	MODERATE		1
c.3340G>C	rs762831873		Not reported in CFTR2	missense_variant	MODERATE		1
c.4053G>A	rs763602969		Not reported in CFTR2	synonymous_variant	LOW		1
c.727A>G	rs763914313		Not reported in CFTR2	missense_variant	MODERATE		1

c.2490+5G>T	rs764466147	Not reported in CFTR2	splice_region_variant&intron_v	LOW	1
c.810A>G	rs767312350	Not reported in CFTR2	synonymous_variant	LOW	1
c.2882T>C	rs769377991	Not reported in CFTR2	missense_variant	MODERATE	1
c.251A>G	rs769754499	Not reported in CFTR2	missense_variant	MODERATE	1
c.2742T>C	rs769879940	Not reported in CFTR2	synonymous_variant	LOW	1
c.2743G>C	rs770502501	Not reported in CFTR2	missense_variant	MODERATE	1
c.3783G>T	rs771812900	Not reported in CFTR2	synonymous_variant	LOW	1
c.62G>T	rs777520137	Not reported in CFTR2	missense_variant	MODERATE	1
c.1227T>C	rs778548877	Not reported in CFTR2	synonymous_variant	LOW	1
c.510T>A	rs780772620	Not reported in CFTR2	synonymous_variant	LOW	1

Supplementary Table 2: Grouped variant testing in COPDGene using SKAT-O

	CF-causing		**CF-causing + Varying clinical consequence**		**CF-causing + Varying clinical consequence + Predicted functional**	**All coding variants**
# of Variants	36 (33 variants in weighted analysis)		61 (58 variants in weighted analysis)		206	297[1]
# of Alleles	254		548		866	2820
	p-value	**p-value with weighting[2]**	**p-value**	**p-value with weighting[2]**	**p-value**	**p-value**
Chronic Bronchitis	0.016	0.015	0.055	0.015	0.056	0.15
COPD	0.049	0.044	0.14	0.045	0.12	0.088
Severe COPD	0.37	0.34	0.42	0.35	0.20	0.029
Severe Exacerbations	0.18	0.17	0.38	0.18	0.31	0.30
BMI	0.22	0.20	0.40	0.21	0.42	0.30
FEV1 percent predicted	0.25	0.22	0.18	0.22	0.15	0.23
Percent Emphysema	0.37	0.38	0.42	0.38	0.44	0.20
Airway Wall Thickness	0.11	0.11	0.019	0.098	0.024	0.028
Bronchodilator Response % FEV1[1]	0.023	0.020	0.025	0.020	0.020	0.21

1. All -p-values are one-sided except for BDR which is two sided
2. SKAT was performed with weighting for %pancreatic insufficiency as a measure of disease severity. Three variants did not have pancreatic insufficiency data and were therefore excluded from the weighted analysis

Supplementary Table 3: Burden testing in COPDGene stratified by current smoking status. P-values and effect sizes for variant-set testing of *CFTR* variants with chronic bronchitis.

	Current Smokers			**Former Smokers**		
	Chronic Bronchitis		p-value/ OR*	Chronic Bronchitis		p-value/ OR*
	Yes	No		Yes	No	
CF-causing	45 (39.5%)	69	0.0082 *(OR=1.62)*	23 (17.2%)	111	0.082
CF-causing + Varying clinical consequence	76 (31.7%)	164	0.037 *(OR=1.22)*	33 (15.3%)	182	0.31
CF-causing + Varying clinical consequence + Predicted functional	115 (29.8%)	271	0.011 *(OR=1.24)*	54 (15.6%)	292	0.23
All coding variants	313 (26.8%)	856	0.45	150 (13.2%)	988	0.46
Controls (no CFTR variants)	813 (23.9%)	2579	-	389 (13.5%)	2500	-

*All -p-values are one-sided. Odds ratios or beta coefficients are shown for all nominally significant associations ($p<0.05$).

Supplementary Table 4: Burden testing in COPDGene stratified by COPD case control status. P-values and effect sizes for variant-set testing of *CFTR* variants with chronic bronchitis.

	COPD Cases			**Smoking Controls**		
	Chronic Bronchitis		p-value/ OR*	Chronic Bronchitis		p-value/ OR*
	Yes	No		Yes	No	
CF-causing	51 (38.1%)	83	0.0022 *(OR=1.72)*	17 (14.9%)	97	0.37
CF-causing + Varying clinical consequence	72 (33.8%)	141	0.038 *(OR=1.27)*	37 (15.4%)	203	0.23
CF-causing + Varying clinical consequence + Predicted functional	106 (31.5%)	231	0.022 *(OR=1.24)*	62 (15.9%)	327	0.13
All coding variants	308 (28.1%)	787	0.16	150 (12.5%)	1047	0.11
Controls (no CFTR variants)	701 (25.5%)	2050	-	493 (14.1%)	2996	-

*All -p-values are one-sided. Odds ratios or beta coefficients are shown for all nominally significant associations ($p<0.05$).

Supplementary Table 5: Single variant testing of all variants in COPDGene with severe COPD. Variants with minor allele count > 10 were included in this analysis.

Variant ID	HGVS ID (legacy name)	Allele counts	Severe COPD		Effect	CFTR2 Determination	SNFEff function
			One sided p-value (Effect Size)	One sided p-value with permutation			
rs142540482	c.221G>A (R74Q)	11	0.0038 *(OR=7.30)*	0.16	Missense variant	Reported in CFTR2; not yet annotated	MODERATE
chr7:117509093:G:A	c.224G>A (R75Q)	459	0.0068 *(OR=1.38)*	0.02	Missense variant	Non CF-causing	MODERATE
rs150157202	c.2260G>A (V754M)	28	0.015 *(OR=2.76)*	0.09	Missense variant	Non CF-causing	MODERATE
rs143486492	c.890G>A (R297Q)	17	0.050	0.17	Missense variant	Reported in CFTR2; not yet annotated	MODERATE
chr7:117530975:G:A	c.350G>A (R117H)	27	0.062	0.16	Missense variant	Varying clinical consequence	MODERATE
rs151073129	c.853A>T (I285F)	21	0.081	0.14	Missense variant	Reported in CFTR2; not yet annotated	MODERATE
chr7:117548796:G:A	c.1365G>A	21	0.081	0.15	Synonymous variant	Not reported in CFTR2	LOW
chr7:117666903:C:T	c.4243-5C>T (4375-5C->T)	23	0.11	0.16	Splice region variant & intron variant	Reported in CFTR2; not yet annotated	LOW
chr7:117559594:T:G	c.1523T>G (F508C)	19	0.11	0.23	Missense variant	Non CF-causing	MODERATE
rs34911792	c.3705T>G (S1235R)	131	0.13	0.13	Missense variant	Non CF-causing	MODERATE
rs1800094	c.1581A>G	78	0.15	0.14	Synonymous variant	Not reported in CFTR2	LOW
rs60887846	c.2820T>G	25	0.26	0.07	Synonymous variant	Not reported in CFTR2	LOW

rs1800118	c.3285A>T	109	0.30	0.35	Synonymous variant	Reported in CFTR2; not yet annotated	LOW
rs1800103	c.2421A>G (I807M)	10	0.31	0.21	Missense variant	Non CF-causing	MODERATE
rs371315549	c.274-6T>C (406-6T->C)	13	0.42	0.28	Splice region variant & intron variant	Reported in CFTR2; not yet annotated	LOW
rs1800111	c.2991G>C (L997F)	29	0.45	0.45	Missense variant & splice region variant	Non CF-causing	MODERATE
chr7:117531068:T:C	c.443T>C (I148T)	22	0.50	0.63	Missense variant	Non CF-causing	MODERATE
chr7:117559655:G:A	c.1584G>A (1716G/A)	290	0.51	0.54	Splice region variant & synonymous variant	Non CF-causing	LOW
rs1800112	c.3080T>C (I1027T)	11	0.57	0.53	Missense variant	Non CF-causing	MODERATE
rs199826652	c.1521_1523 delCTT (F508del)	177	0.62	0.68	Disruptive inframe deletion	CF-causing	MODERATE
chr7:117642528:G:A	c.3808G>A (D1270N)	86	0.65	0.57	Missense variant	Varying clinical consequence	MODERATE
rs75541969	c.3454G>C (D1152H)	14	0.68	0.66	Missense variant	Varying clinical consequence	MODERATE
chr7:117504290:C:T	c.91C>T (R31C)	28	0.69	0.70	Missense variant	Non CF-causing	MODERATE
rs140455771	c.2249C>T (P750L)	14	0.74	0.73	Missense variant	Varying clinical consequence	MODERATE
rs1800110	c.2900T>C (L967S)	29	0.76	0.69	Missense variant	Varying clinical consequence	MODERATE
rs1800109	c.2898G>A (3030G/A)	166	0.78	0.76	Synonymous variant	Reported in CFTR2; not yet annotated	LOW
rs1800100	c.2002C>T (R668C)	105	0.84	0.82	Missense variant	Non CF-causing	MODERATE
rs115545701	c.220C>T (R74W)	82	0.85	0.81	Missense variant	Varying clinical consequence	MODERATE

rs151235408	c.2245C>T (2377C/T)	14	0.86	0.84	Synonymous variant	Reported in CFTR2; not yet annotated	LOW
chr7:117590400:G:C	c.1727G>C (G576A)	85	0.88	0.93	Missense variant	Non CF-causing	MODERATE
chr7:117627538:G:T	c.3485G>T (R1162L)	17	0.93	0.05	Missense variant	Non CF-causing	MODERATE
rs1800079	c.509G>A (R170H)	12	0.94	0.13	Missense variant	Non CF-causing	MODERATE
rs148783445	c.4333G>A (D1445N)	11	0.95	0.12	Missense variant	Reported in CFTR2; not yet annotated	MODERATE
rs1800135	c.4272C>T (4404C/T)	136	0.97	0.98	Synonymous variant	Reported in CFTR2; not yet annotated	LOW
rs1800131	c.3897A>G (4029A/G)	17	0.98	0.97	Synonymous variant	Reported in CFTR2; not yet annotated	LOW
chr7:117603609:C:T	c.2735C>T (S912L)	16	1.00	0.94	Missense variant	Unknown significance	MODERATE

Supplementary Table 6: CFTR Compound Heterozygous Subjects in COPDGene

Variant 1/ Variant 2 HGVSc	Variant Determination	Case/ Control	Current Smoker Status	Chronic Bronchitis	Bronchiectasis
c.1521_1523delCTT c.1584G>A	CF-causing/ Non CF causing	Control	Current	No	
c.1521_1523delCTT c.1584G>A	CF-causing/ Non CF-causing	Control	Former	No	
c.1521_1523delCTT c.1584G>A	CF-causing/ Non CF-causing	Case	Former	Chronic Bronchitis	Bronchiectasis score=2; 2 lobes with bronchiectasis
c.1521_1523delCTT c.1584G>A	CF-causing/ Non CF-causing	Case	Former	No	
c.1521_1523delCTT c.2735C>T	CF-causing/ MODERATE	Control	Former	No	
c.1521_1523delCTT c.3485G>T	CF-causing/ Non CF-causing	Case	Former	No	
c.1521_1523delCTT c.650A>G	CF-causing/ MODERATE	Case	Former	No	
c.1521_1523delCTT / c.221G>A and c.890G>A[1]	CF-causing/ MODERATE	Case	Current	Chronic Bronchitis	
c.1521_1523delCTT c.221G>A and c.890G>A	CF-causing/ MODERATE	Case	Current	Chronic Bronchitis	
c.1521_1523delCTT c.2855T>C	CF-causing/ MODERATE	Control	Former	No	
c.1521_1523delCTT c.2245C>T	CF-causing/ LOW	Case	Current	Chronic Bronchitis	
c.1521_1523delCTT c.2002C>T	CF-causing/ MODERATE	Case	Former	No	
c.1521_1523delCTT c.2002C>T	CF-causing/ MODERATE	Case	Former	Chronic Bronchitis	No bronchiectasis
c.1521_1523delCTT	CF-causing/	Case	Former	No	

c.2900T>C	Varying clinical consequence				
c.1521_1523delCTT c.3285A>T	CF-causing/ LOW	Case	Current	No	
c.1521_1523delCTT c.4272C>T	CF-causing/ LOW	Case	Former	No	
c.1521_1523delCTT c.4272C>T	CF-causing/ LOW	Control	Current	Chronic Bronchitis	
c.1521_1523delCTT c.4272C>T	CF-causing/ LOW	Case	Current	Chronic Bronchitis	
c.1521_1523delCTT c.3705T>G	CF-causing/ Non CF-causing	Control	Former	No	Bronchiectasis score=5; 3 lobes with bronchiectasis
c.1521_1523delCTT c.3705T>G	CF-causing/ Non CF-causing	Case	Current	No	
c.1521_1523delCTT c.589T>C	CF-causing/ LOW	Case	Current	Chronic Bronchitis	
c.1521_1523delCTT c.224G>A	CF-causing/ Non CF-causing	Case	Former	No	
c.1521_1523delCTT c.224G>A	CF-causing/ Non CF-causing	Case	Former	Chronic Bronchitis	No bronchiectasis
c.1521_1523delCTT c.224G>A	CF-causing/ Non CF-causing	Case	Current	No	
c.1521_1523delCTT c.224G>A	CF-causing/ Non CF-causing	Case	Former	No	
c.1521_1523delCTT c.224G>A	CF-causing/ Non CF-causing	Control	Current	No	
c.1521_1523delCTT c.727A>G	CF-causing/ MODERATE	Control	Current	No	
c.3209G>A/ c.3808G>A	Varying clinical consequence/ Varying clinical consequence	Control	Current	No	
c.350G>A/ c.2900T>C and	Varying clinical consequence/	Case	Former	No	

c.3808G>A[2]	Varying clinical consequence				
c.1865G>A/ c.220C>T and c.3808G>A[3]	Varying clinical consequence/ Varying clinical consequence	Case	Former	No	
c.220C>T and c.3808G>A/ c.220C>T and c.3808G>A[3]	Varying clinical consequence/ Varying clinical consequence	Control	Current	No	
c.220C>T and c.3808G>A/ c.220C>T and c.3808G>A[3]	Varying clinical consequence/ Varying clinical consequence	Control	Former	No	Bronchiectasis score=6; 3 lobes with bronchiectasis

[1] c.890G>A is often in cis with c.221G>A. Both variants are missense variants predicted to have moderate impact on CFTR protein

[2] c.2900T>C was found in cis with c.3808G>A in one subject. Both variants have varying clinical consequence.

[3] c.2900T>C is often in cis with c.3808G>A. Both variants have varying clinical consequence

Supplementary Table 7: Alternative variants IDs for compound heterozygous subjects in COPDGene

Variant 1/ Variant 2 Variant ID	**Variant 1/ Variant 2 HGVSc ID**	**Variant 1/ Variant 2 Legacy Name**
rs199826652/ chr7:117559655:G:A	c.1521_1523delCTT/ c.1584G>A	F508del/ 1716G/A
rs199826652/ chr7:117559655:G:A	c.1521_1523delCTT/ c.1584G>A	F508del/ 1716G/A
rs199826652/ chr7:117559655:G:A	c.1521_1523delCTT/ c.1584G>A	F508del/ 1716G/A
rs199826652/ chr7:117559655:G:A	c.1521_1523delCTT/ c.1584G>A	F508del/ 1716G/A
rs199826652/ chr7:117603609:C:T	c.1521_1523delCTT/ c.2735C>T	F508del/ S912L
rs199826652/ chr7:117627538:G:T	c.1521_1523delCTT/ c.3485G>T	F508del
rs199826652/ rs121909046	c.1521_1523delCTT/c.650A>G	F508del/ E217G
rs199826652/ rs142540482 and rs143486492	c.1521_1523delCTT/ c.221G>A and c.890G>A	F508del/ R74Q and R297Q
rs199826652/ rs142540482 and rs143486492	c.1521_1523delCTT/ c.221G>A and c.890G>A	F508del/R74Q and R297Q
rs199826652/ rs142773283	c.1521_1523delCTT/ c.2855T>C	F508del/M952T
rs199826652/ rs151235408	c.1521_1523delCTT/ c.2245C>T	F508del/2377C/T
rs199826652/ rs1800100	c.1521_1523delCTT/ c.2002C>T	F508del
rs199826652/ rs1800100	c.1521_1523delCTT/ c.2002C>T	F508del
rs199826652/ rs1800110	c.1521_1523delCTT/ c.2900T>C	F508del/L967S
rs199826652/ rs1800118	c.1521_1523delCTT/ c.3285A>T	F508del
rs199826652/ rs1800135	c.1521_1523delCTT/ c.4272C>T	F508del/4404C/T
rs199826652/ rs1800135	c.1521_1523delCTT/ c.4272C>T	F508del/4404C/T
rs199826652/ rs1800135	c.1521_1523delCTT/ c.4272C>T	F508del/4404C/T
rs199826652/ rs34911792	c.1521_1523delCTT/ c.3705T>G	F508del/S1235R
rs199826652/ rs34911792	c.1521_1523delCTT/ c.3705T>G	F508del/S1235R
rs199826652/ rs755619078	c.1521_1523delCTT/ c.589T>C	F508del
rs199826652/ chr7:117509093:G:A	c.1521_1523delCTT/ c.224G>A	F508del/R75Q

rs199826652/ chr7:117509093:G:A	c.1521_1523delCTT/ c.224G>A	F508del/R75Q
rs199826652/ chr7:117509093:G:A	c.1521_1523delCTT/ c.224G>A	F508del/R75Q
rs199826652/ chr7:117509093:G:A	c.1521_1523delCTT/ c.224G>A	F508del/R75Q
rs199826652/ chr7:117509093:G:A	c.1521_1523delCTT/ c.224G>A	F508del/R75Q
rs199826652/ rs763914313	c.1521_1523delCTT/ c.727A>G	F508del
chr7:117611650:G:A/ chr7:117642528:G:A	c.3209G>A/ c.3808G>A	R1070Q / D1270N
chr7:117530975:G:A/ rs115545701 and chr7:117642528:G:A	c.350G>A/ c.2900T>C and c.3808G>A	R117H/ R74W and D1270N
rs121908759/ rs115545701 and chr7:117642528:G:A	c.1865G>A/ c.220C>T and c.3808G>A	G622D/ R74W and D1270N
rs115545701 and chr7:117642528:G:A/ rs115545701 and chr7:117642528:G:A	c.220C>T and c.3808G>A/ c.220C>T and c.3808G>A	R74W and D1270N/ R74W and D1270N
rs115545701 and chr7:117642528:G:A/ rs115545701 and chr7:117642528:G:A	c.220C>T and c.3808G>A/ c.220C>T and c.3808G>A	R74W and D1270N/ R74W and D1270N

HGVS ID	Variant ID	Legacy Name	CFTR2 Determination	SnpEff predicted effect	Pancreatic Insufficiency	Number of alleles in ECLIPSE
c.1408G>A	rs213950	V470M	Non CF-causing	missense_variant		1963
c.2562T>G	chr7:117595001:T:G		Not reported in CFTR2	synonymous_variant		1543
c.4389G>A	rs1800136	4521G/A	Not evaluated	synonymous_variant		1128
c.744-9_744-6delGATT	rs1432807327		Not reported in CFTR2	splice_region_variant&intron_variant		882
c.3870A>G	rs1800130	P1290P	Not evaluated	synonymous_variant		179
c.224G>A	chr7:117509093:G:A	R75Q	Non CF-causing	missense_variant	0.28	160
c.1210-7_1210-6dupTT	chr7:117548628:G:GTT		Not reported in CFTR2	splice_region_variant&intron_variant		137
c.1210-7_1210-6delTT	rs1491448762		Not reported in CFTR2	splice_region_variant&intron_variant		136
c.1251C>A	chr7:117548682:C:A		Not reported in CFTR2	missense_variant		119
c.1584G>A	chr7:117559655:G:A	1716G/A	Non CF-causing	splice_region_variant&synonymous_varian	0.67	114
c.1521_1523delCTT	rs1297060838	F508del	CF-causing	disruptive_inframe_deletion	0.98	57
c.3285A>T	rs1800118	3417A/T	Not Evaluated	synonymous_variant		49
c.3705T>G	rs34911792	S1235R	Non CF-causing	missense_variant	0.38	41
c.4272C>T	chr7:117666937:C:T	4404C/T	Not evaluated	synonymous_variant		29
c.2898G>A	rs1800109	3030G/A	Not evaluated	synonymous_variant		27
c.350G>A	chr7:117530975:G:A	R117H	Varying clinical consequence	missense_variant	0.23	14
c.744-9_744-6dupGATT	rs386134231		Not reported in CFTR2	splice_region_variant&intron_variant		13
c.2900T>C	rs1800110	L967S	Varying clinical consequence	missense_variant	0.00	11
c.443T>C	chr7:117531068:T:C	I148T	Non CF-causing	missense_variant	0.88	11
c.91C>T	chr7:117504290:C:T	R31C	Non CF-causing	missense_variant	0.30	10
c.2991G>C	rs1800111	L997F	Non CF-causing	missense_variant&splice_region_variant	0.32	10
c.2855T>C	rs142773283	M952T	Unknown significance	missense_variant	0.00	8
c.3485G>T	chr7:117627538:G:T	R1162L	Non CF-causing	missense_variant	0.38	8
c.2260G>A	rs150157202	V754M	Non CF-causing	missense_variant	0.73	8
c.1523T>G	chr7:117559594:T:G	F508C	Non CF-causing	missense_variant	0.60	7
c.2245C>T	rs151235408	2377C/T	Not evaluated	synonymous_variant		7
c.3897A>G	rs1800131	4029A/G	Not evaluated	synonymous_variant		7
c.890G>A	rs143486492	R297Q	Not evaluated	missense_variant		5
c.2421A>G	rs1800103	I807M	Non CF-causing	missense_variant	0.00	4
c.3154T>G	chr7:117611595:T:G	F1052V	Varying clinical consequence	missense_variant	0.15	4
c.3873+2T>C	rs146795445	4005+2T->C	CF-causing	splice_donor_variant&intron_variant	0.27	4
c.3909C>G	rs80034486	N1303K	CF-causing	missense_variant	0.98	4
c.650A>G	rs121909046	E217G	Not evaluated	missense_variant		4
c.3415A>G	rs397508556	I1139V	Not Evaluated	missense_variant		4
c.2249C>T	rs140455771	P750L	Varying clinical consequence	missense_variant	0.33	3
c.1624G>T	chr7:117587778:G:T	G542X	CF-causing	stop_gained	0.98	3
c.3558A>G	chr7:117627611:A:G	Q1186Q (3690A/G)	Not reported in CFTR2	synonymous_variant		3
c.1043T>A	rs142920240	M348K	Not evaluated	missense_variant		3
c.1052C>G	chr7:117540282:C:G	T351S	Not evaluated	missense_variant		3
c.1079C>T	chr7:117540309:C:T		Not reported in CFTR2	missense_variant		3
c.2770G>A	chr7:117603644:G:A	D924N	Unknown significance	missense_variant	0.00	2
c.1001G>A	chr7:117540231:G:A	R334Q	Varying clinical consequence	missense_variant	0.00	2
c.220C>T	rs115545701	R74W	Varying clinical consequence	missense_variant	0.15	2
c.349C>T	chr7:117530974:C:T	R117C	CF-causing	missense_variant	0.24	2
c.2657+2_2657+3insA	rs397508414	2789+2insA	Unknown significance	splice_region_variant&intron_variant	0.31	2
c.3080T>C	rs1800112	I1027T	Non CF-causing	missense_variant	0.70	2
c.262_263delTT	rs754147777	394delTT	CF-causing	frameshift_variant	0.97	2
c.3846G>A	chr7:117642566:G:A	W1282X	CF-causing	stop_gained	0.99	2
c.274-6T>C	rs371315549	406-6T->C	Not evaluated	splice_region_variant&intron_variant		2
c.997C>T	rs193922533	L333F	Not evaluated	missense_variant		2
c.1734A>G	rs201025424		Not reported in CFTR2	synonymous_variant		2
c.31G>A	chr7:117480125:G:A	V11I	Not evaluated	missense_variant		2
c.2758G>A	chr7:117603632:G:A	V920M	Not evaluated	missense_variant		2
c.92G>T	chr7:117504291:G:T	R31L	Unknown significance	missense_variant	0.00	1
c.14C>T	rs193922501	P5L	Varying clinical consequence	missense_variant	0.10	1
c.3808G>A	chr7:117642528:G:A	D1270N	Varying clinical consequence	missense_variant	0.17	1
c.328G>C	chr7:117530953:G:C	D110H	CF-causing	missense_variant	0.17	1
c.4426C>T	rs374705585	Q1476X	Varying clinical consequence	stop_gained	0.18	1
c.617T>G	rs121908752	L206W	CF-causing	missense_variant	0.20	1
c.579+3A>G	chr7:117534368:A:G	711+3A->G	CF-causing	splice_region_variant&intron_variant	0.21	1
c.377G>A	chr7:117531002:G:A	G126D	CF-causing	missense_variant	0.22	1
c.489+3A>G	rs377729736	621+3A->G	Varying clinical consequence	splice_region_variant&intron_variant	0.25	1
c.1055G>A	rs121908753	R352Q	CF-causing	missense_variant	0.40	1
c.2657+5G>A	rs80224560	2789+5G->A	CF-causing	splice_region_variant&intron_variant	0.43	1
c.2506G>T	rs201386642	D836Y	Non CF-causing	missense_variant	0.50	1
c.2735C>T	chr7:117603609:C:T	S912L	Unknown significance	missense_variant	0.60	1
c.3205G>A	rs200321110	G1069R	Varying clinical consequence	missense_variant	0.67	1
c.3484C>T	rs74767530	R1162X	CF-causing	stop_gained	0.97	1
c.1585-1G>A	rs76713772	1717-1G->A	CF-causing	splice_acceptor_variant&intron_variant	0.97	1
c.2052delA	rs1164974840	2184delA	CF-causing	frameshift_variant	0.98	1
c.2988+1G>A	rs75096551	3120+1G->A	CF-causing	splice_donor_variant&intron_variant	0.98	1
c.579+1G>T	rs77188391	711+1G->T	CF-causing	splice_donor_variant&intron_variant	0.98	1
c.580-1G>T	chr7:117535247:G:T	712-1G->T	CF-causing	splice_acceptor_variant&intron_variant	1.00	1
c.715G>A	rs397508788	G239R	Not evaluated	missense_variant		1
c.853A>T	rs151073129	I285F	Not evaluated	missense_variant		1
c.1516A>G	chr7:117559587:A:G	I506V	Not evaluated	missense_variant		1
c.41A>T	chr7:117480135:A:T		Not reported in CFTR2	missense_variant		1
c.92G>A	chr7:117504291:G:A		Not reported in CFTR2	missense_variant		1
c.672_674delCTG	chr7:117535339:TCTG:T		Not reported in CFTR2	disruptive_inframe_deletion		1
c.1429C>T	chr7:117559500:C:T		Not reported in CFTR2	missense_variant		1
c.1516A>C	chr7:117559587:A:C	I506L	Not evaluated	missense_variant		1
c.1886C>T	chr7:117592053:C:T		Not reported in CFTR2	missense_variant		1
c.2110C>A	chr7:117592277:C:A		Not reported in CFTR2	missense_variant		1
c.2255T>G	chr7:117592422:T:G	I752S	Not evaluated	missense_variant		1
c.2493G>A	chr7:117594932:G:A		Not reported in CFTR2	splice_region_variant&synonymous_variant		1
c.2657+6T>C	chr7:117602869:T:C		Not reported in CFTR2	splice_region_variant&intron_variant		1
c.2684G>C	chr7:117603558:G:C		Not reported in CFTR2	missense_variant		1

c.2758G>T	chr7:117603632:G:T	V920L	Not evaluated	missense_variant	1
c.2824A>G	chr7:117603698:A:G		Not reported in CFTR2	missense_variant	1
c.2937T>C	chr7:117606702:T:C		Not reported in CFTR2	synonymous_variant	1
c.3005T>C	chr7:117610535:T:C		Not reported in CFTR2	missense_variant	1
c.3219C>T	chr7:117611660:C:T		Not reported in CFTR2	synonymous_variant	1
c.3590A>G	chr7:117627643:A:G		Not reported in CFTR2	missense_variant	1
c.3592G>A	chr7:117627645:G:A		Not reported in CFTR2	missense_variant	1
c.3743C>G	chr7:117642463:C:G		Not reported in CFTR2	stop_gained	1
c.3780A>C	chr7:117642500:A:C		Not reported in CFTR2	synonymous_variant	1
c.3918C>G	chr7:117652886:C:G		Not reported in CFTR2	synonymous_variant	1
c.3945A>G	chr7:117652913:A:G		Not reported in CFTR2	missense_variant	1
c.4409A>T	chr7:117667074:A:T		Not reported in CFTR2	missense_variant	1
c.82T>C	rs1012752433		Not reported in CFTR2	missense_variant	1
c.4243-7delT	rs1170705810	4375-7delT	Not evaluated	splice_region_variant&intron_variant	1
c.2262G>C	rs1201012182		Not reported in CFTR2	synonymous_variant	1
c.1585-7G>A	rs1367201083		Not reported in CFTR2	splice_region_variant&intron_variant	1
c.2831T>C	rs141747560	V944A	Not evaluated	missense_variant	1
c.964G>A	rs1800085	V322M	Not evaluated	missense_variant	1
c.2559T>C	rs1800104		Not reported in CFTR2	synonymous_variant	1
c.580-2A>G	rs193922730	712-2A->G	Not evaluated	splice_acceptor_variant&intron_variant	1
c.2173G>A	rs199791061		Not reported in CFTR2	missense_variant	1
c.1270G>A	rs371107552	G424S	Not evaluated	missense_variant	1
c.365A>G	rs377295859	Y122C	Not evaluated	missense_variant	1
c.1601C>A	rs387906368	A534E	Not evaluated	missense_variant	1
c.2417A>G	rs397508375	D806G	Not evaluated	missense_variant	1
c.2563G>A	rs397508397	V855I	Not evaluated	missense_variant	1
c.2756A>G	rs397508430	Y919C	Not evaluated	missense_variant	1
c.3409A>G	rs397508553	M1137V	Not evaluated	missense_variant	1
c.3680T>C	rs397508593	L1227S	Not evaluated	missense_variant	1
c.4091C>T	rs397508670		Not reported in CFTR2	missense_variant	1
c.4312C>T	rs397508711	R1438W	Not evaluated	missense_variant	1
c.2475C>T	rs746961486		Not reported in CFTR2	synonymous_variant	1
c.2052A>G	rs750642366		Not reported in CFTR2	synonymous_variant	1
c.54-8T>A	rs778197563		Not reported in CFTR2	splice_region_variant&intron_variant	1
c.2933A>G	rs943473311		Not reported in CFTR2	missense_variant	1
c.2395C>A	rs984281283		Not reported in CFTR2	missense_variant	1
c.332C>T	chr7:117530957:C:T	P111L	Not evaluated	missense_variant	1
c.221G>A	rs142540482	R74Q	Not evaluated	missense_variant	1
c.3588A>T	chr7:117627641:A:T		Not reported in CFTR2	synonymous_variant	1
c.2153C>G	rs142432539		Not reported in CFTR2	missense_variant	1
c.2424T>C	rs143954792		Not reported in CFTR2	synonymous_variant	1
c.925G>A	rs148013312	A309T	Not reported in CFTR2	missense_variant	1
c.3874-4A>G	rs201381687	4006-4A->G	Not reported in CFTR2	splice_region_variant&intron_variant	1
c.3713A>G	rs397508594	Q1238R	Not reported in CFTR2	missense_variant	1
c.589T>C	rs755619078		Not reported in CFTR2	synonymous_variant	1
c.948T>G	rs78742051	F316L	Not reported in CFTR2	missense_variant	1