Cystic Fibrosis Gene Modifier SLC26A9 Modulates Airway Response to CFTR-directed Therapeutics

Lisa J Strug, Tanja Gonska, Gengming He, Katherine Keenan, Wan Ip, Pierre-Yves Boelle, Fan Lin, Naim Panjwani, Jiafen Gong, Weili Li, et al.

Cystic Fibrosis Gene Modifier *SLC26A9* Modulates Airway Response to CFTR-directed Therapeutics

Lisa J. Strug[1,2,3]*, Tanja Gonska[4,5,6], Gengming He[1], Katherine Keenan[6], Wan Ip[6], Pierre-Yves Boelle[7,8], Fan Lin[1], Naim Panjwani[1], Jiafen Gong[1], Weili Li[1,3], David Soave[1,2], Bowei Xiao[1], Elizabeth Tullis[9], Harvey Rabin[10,11], Michael D. Parkins[10,11], April Price[12], Peter C. Zuberbuhler[13], Harriet Corvol[7,14], Felix Ratjen[6,15], Lei Sun [2,16], Christine E. Bear[17,18,19], Johanna M. Rommens[1,20]

Affiliations:

[1] Program in Genetics and Genome Biology, Research Institute, The Hospital for Sick Children, Toronto, Ontario, M5G 1X8, Canada

[2] Division of Biostatistics, Dalla Lana School of Public Health, University of Toronto, Ontario, M5T 3M7, Canada

[3] The Centre for Applied Genomics, The Hospital for Sick Children, Toronto, Ontario, M5G 1X8, Canada

[4] Department of Pediatrics, University of Toronto, Toronto, Ontario, M5G 1E2, Canada

[5] Division of Gastroenterology, Hepatology and Nutrition, The Hospital for Sick Children, Toronto, Ontario, M5G 1X8, Canada

[6] Program in Physiology and Experimental Medicine, The Hospital for Sick Children, Toronto, Ontario, M5G 1X8, Canada

[7] Sorbonne Universités, Université Pierre et Marie Curie (UPMC) Paris 06, Paris 75005, France

[8] Assistance Publique-Hôpitaux de Paris (AP-HP), Hôpital St. Antoine, Biostatistics Department; Inserm U1136, Paris 75012, France

[9] Department of Respiratory Medicine and Li Ka Shing Knowledge Institute, St. Michael's Hospital, University of Toronto, Toronto, Ontario, M5B 1W8, Canada

[10] Department of Medicine, University of Calgary, Calgary, Alberta, T2N 4N1, Canada

[11] The Department of Microbiology, Immunology and Infectious Disease, University of Calgary, Calgary, Alberta, T2N 4N1, Canada

[12] Division of Paediatric Respirology, Department of Paediatrics, Children's Hospital at London Health Sciences Centre, London, Ontario, N6A 5W9, Canada

[13] Department of Pediatrics, University of Alberta, Edmonton, Alberta, T6G1C9, Canada

[14] AP-HP, Hôpital Trousseau, Pediatric Pulmonary Department; Institut National de la Santé et al Recherche Medicale (INSERM) U938, Paris 75012, France

[15] Division of Respiratory Medicine, The Hospital for Sick Children, Toronto, Ontario, M5G 1X8, Canada

[16] Department of Statistical Sciences, University of Toronto, Toronto, Ontario, M5S 3G3, Canada

[17] Program in Molecular Structure and Function, The Hospital for Sick Children, Toronto, Ontario, M5G 1X8, Canada

Departments of [18]Biochemistry, [19]Physiology and [20]Molecular Genetics, University of Toronto, Toronto, Ontario, M5S 1A8, Canada

*To whom correspondence should be directed: lisa.strug@utoronto.ca,

Tel: 1-416-813-7654 ext. 301762.

Fax: 1-416-813-8421

Abstract

Cystic fibrosis is realizing the promise of personalized medicine. Recent advances in drug development that target the causal *CFTR* directly result in lung function improvement, but variability in response is demanding better prediction of outcomes to improve management decisions. The genetic modifier *SLC26A9* contributes to disease severity in the CF pancreas and intestine at birth and here we assess its relationship with disease severity and therapeutic response in the CF lungs.

SLC26A9 association with lung disease was assessed in individuals from the Canadian and French CF Gene Modifier consortia with CFTR-gating mutations and in those homozygous for the common Phe508del mutation. Variability in response to a CFTR-directed therapy attributed to *SLC26A9* genotype was assessed in Canadian patients with gating mutations. A primary airway model system determined if *SLC26A9* shows modification of Phe508del CFTR function upon treatment with a CFTR *corrector.*

In those with gating mutations that retain cell-surface localized CFTR we show that *SLC26A9* modifies lung function while this is not the case in individuals homozygous for Phe508del where cell surface expression is lacking. Treatment response to ivacaftor, which aims to improve CFTR-channel opening probability in patients with gating mutations, shows substantial variability in response, 28% of which can be explained by rs7512462 in *SLC26A9* (p=0.0006).

When homozygous Phe508del primary bronchial cells are treated to restore surface CFTR,

SLC26A9 likewise modifies treatment response (p=0.02).

Our findings indicate that SLC26A9 airway modification requires CFTR at the cell surface, and

that a common variant in *SLC26A9* may predict response to CFTR-directed therapeutics.

Introduction

Cystic fibrosis (CF; MIM:219700) is a common life-limiting genetic disease caused by mutations in the cystic fibrosis transmembrane conductance regulator (*CFTR*; MIM:602421) gene. CF affects multiple organs including the lungs, pancreas, sweat glands, intestine and male reproductive tract. Meconium ileus (intestinal obstruction at birth; ~ 15% of patients) and exocrine pancreatic dysfunction occur early in life, while the majority of CF-associated mortality results from repeated pulmonary exacerbations and reduced lung function later in the course of disease. Individuals with the same *CFTR* mutations have variable disease outcomes across CF-affected organs. Mutation heterogeneity and several environmental exposures (such as infection history, secondhand smoke, and ambient air pollution) (1-3) along with other genes beyond *CFTR* (referred to as modifier genes) (4-6) also influence lung disease severity.

CFTR is an anion channel present in select epithelial tissues. CF-causing variants affect the amount of protein and/or the function of the CFTR channel at the apical membrane surface (7). For example, the most common CF-causing allele is a three base pair deletion resulting in the loss of a phenylalanine residue at position 508 in the CFTR protein, Phe508del (c.1521_1523delCTT; p.Phe508del). Phe508del represents ~70% of all CF-causing alleles (8) and results in only nominal CFTR levels at the apical membrane due to processing deficiencies (9, 10); any residual mutant channel that does reach the cell surface exhibits reduced open probability (11) and appears subject to rapid elimination (12). In contrast, the missense mutation G551D (c.1652G>A; p.Gly551Asp), occurring on at least one allele in ~4-5% of individuals with CF (with varying geographical frequency; (13)), results in cell-surface channel localization but exhibits defective gating (14, 15).

Until recently, treatment of CF has focused on lessening the manifestations of CF disease in the affected organs (7). The development of ivacaftor led to the first prescribed drug that targets CFTR. Ivacaftor increases the opening probability of CFTR at the apical surface to aid chloride and bicarbonate ion transport in CF-epithelia (16). Ivacaftor was shown to significantly improve lung function as measured by percent-predicted forced expiratory volume in 1 second (FEV_1pp) in individuals with at least one G551D-CFTR allele (17), and is now approved for use in CF individuals with gating mutations (18). More recently, ivacaftor has been combined with lumacaftor (VX-809), which is a CFTR corrector that improves the Phe508del CFTR processing to increase the amount of cell surface-localized protein, demonstrating improved outcomes in a clinical trial (19). Although important, the average improvements appear more modest than were observed in the ivacaftor trial (17, 19).

Individuals treated with ivacaftor, or ivacaftor and lumacaftor combinations responded to their respective treatments with marked variability in both mutation groups (17, 19, 20). The factors that explain this variability remain unknown, although the extent and nature of compromised tissue integrity due to lung disease or infection status, bioavailability and aspects of drug metabolism (including interaction), compliance, and effects of modifier genes all may play a role. Identifying the genetic factors that contribute to this variability could lead to improved regimens and personalized approaches to treatment.

SLC26A9 has been shown to modify CF phenotypes, but its functional role is not well understood. An inhibitory relation between CFTR and SLC26A9 had initially been suggested by *in vitro* studies of individual regulatory (R) and Sulfate Transporter and Anti-Sigma factor antagonist domains, respectively (21). However, biochemical and electrophysiological studies support the hypothesis that SLC26A9 is an anion channel (22, 23) that may physically interact

6

and enhance the functional expression of CFTR (24-26). The latter studies are further supported by *in vivo* studies, wherein the disruption of *Slc26a9* exacerbates the gut CF-related morbidities in *Cftr*-deficient mice (27).

Genome-wide studies of CF patient populations consisting of individuals with any two severe loss-of-function *CFTR* mutations (including Phe508del) have shown *SLC26A9* to be associated with early-onset CF phenotypes in the gastrointestinal tract (28, 29), possibly by providing alternate anion transport to CFTR. In fact, one intronic SNP, rs7512462 in *SLC26A9*, accounts for almost 12% of the variability in a biomarker of CF exocrine pancreatic disease at birth (30). However, no association of *SLC26A9* was seen with lung function in CF patients (31), even with stratification by age (32) or in a recent meta-analysis of the largest CF population studied for genetic modification to date (6). This finding is inconsistent with observations that SLC26A9 is expressed in the human (33, 34) and mouse (27, 35)) adult lung (and stomach), with only low expression in other organs such as the pancreas or the intestinal tract.

Here we attempt to reconcile why *SLC26A9* did not appear to influence lung disease across the CF population with severe *CFTR* mutations. Given our current understanding of SLC26A9 and the failings of different *CFTR* mutations (7), we hypothesized that modification by *SLC26A9* may be evident in individuals with CF-causing variants such as G551D that co-localize at the surface of airway epithelial cells. Such specificity could then imply that SLC26A9 may modulate response to CFTR-directed therapies aimed at increasing cell-surface localized CFTR protein, particularly in the airway (7). Given the recent introduction of CFTR-directed therapies and the low frequency of the G551D mutation, we provide convergent evidence from CF population studies (Canada and France), small cohorts of patients with G551D undergoing

7

treatment, as well as studies of primary tissue explants homozygous for Phe508del, to support

that CFTR response is influenced by *SLC26A9* genotype.

Results

SLC26A9 rs7512462 is a modifier of CF lung disease in individuals with a G551D allele

Of the individuals enrolled in the Canadian and French CF Gene Modifier studies with genome-

wide genotyping (29) and previously analyzed for lung function associations with *SLC26A9* (6,

36, 37), 70 individuals carried at least one G551D allele and 1,759 carried two Phe508del alleles.

None of the patients were being treated with ivacaftor at the time of lung function assessment,

which involved multiple longitudinal measures of forced expiratory volume in 1 second (FEV_1)

over a three-year period. For genetic association analysis a **S**urvival-adjusted **a**veraged CF-

specific **K**ulich (38) FEV_1 percentile (calculated as a function of age, height and sex) that was

normalized (Saknorm) was used, as has been previously described (37). The minor allele C of

rs7512462 (sample MAF=0.38) occurring in intron 5 of *SLC26A9* was identified as a significant

and protective modifier of CF phenotypes in the pancreas and intestine in previous studies (29,

30, 39), with other variants in high linkage disequilibrium (LD) generally aggregating toward the

transcription start region of *SLC26A9*.

The distribution of lung function in individuals homozygous for Phe508del (n=1,759) is

similar across the rs7512462 genotypes (Figure 1A, Table 1, top panel, p=0.20), whereas for

those with at least one G551D CFTR allele (or other gating mutation), the number of C alleles is

positively associated with improved lung function (Figure 1B, Table 1, bottom panel, p=0.04); in

those with a G551D allele each additional C allele at rs7512462 is associated with a ~0.31 unit

increase in Saknorm, which corresponds to a change in FEV_1 of approximately 8.5 percentage

8

points for an 18-year-old at the 50th percentile for height. Genotypic analysis suggests individuals with the CC genotype display a 0.69 (s.e.=0.32, p=0.04) unit improvement in Saknorm over those with the TT genotype, while those who are CT have a 0.21 (s.e.=0.22, p=0.34) unit improvement over those with TT. To determine whether the effect size of rs7512462 on lung function differs significantly between individuals homozygous for Phe508del and those who carry a G551D-CFTR allele, we assessed statistical interaction in an analysis of the combined samples (all individuals with a G551D allele and those homozygous for Phe508del, n=1,829) which indicated that even with the modest number of individuals with a G551D allele there is evidence of CFTR-by-SLC26A9 interaction (p=0.055). This is consistent with an hypothesis that the SLC26A9 channel modifies lung disease through interaction with CFTR.

[Figure 1 here]

Response to ivacaftor in patients with a gating mutation is dependent on SLC26A9 rs7512462 genotype

The association with lung function in individuals with a G551D or other gating mutation together with the earlier protein interaction studies suggested that *SLC26A9* may influence response to CFTR-directed therapies that have the objective of increasing the activity of surface localized mutant CFTR. We therefore tested whether rs7512462 could explain variability in FEV$_1$pp response to treatment in a total of 24 individuals prescribed the CFTR-directed therapy ivacaftor, with recruitment in two phases. Baseline and post-treatment measures of FEV$_1$pp with ivacaftor were first obtained from 11 individuals with a G551D allele or other gating mutation

enrolled in a prospective observational study of ivacaftor response at The Hospital for Sick Children and St. Michael's Hospital CF clinics. Most had at least 3 post-treatment measurements that were averaged (see below; Suppl. Table 1). The primary treatment response outcome was defined as the difference between the average FEV_1pp post treatment and the baseline FEV_1pp taken prior to treatment.

Individuals displayed variable response to treatment with ivacaftor, with observed treatment differences ranging from -6.25 to 29.33 FEV_1pp. Although almost all individuals on ivacaftor exhibited improvement in their FEV_1pp from baseline, each rs7512462 C allele was associated with an improvement of 9.56 (se=2.98) FEV_1pp after adjustment for baseline FEV_1pp. The magnitude of the improvement associated with the rs7512462 C allele was substantial enough to achieve statistical significance (p=0.01) despite the small sample size.

We then obtained an independent sample of 13 individuals with a G551D allele on ivacaftor from across Canada. Remarkably, in this replication sample where there were no individuals with the CC genotype, having the CT genotype at rs7512462 was associated with an improvement of 9.93 (se=3.98) in FEV_1pp (p=0.03) over individuals with the TT genotype after adjusting for baseline FEV_1pp, very similar to the discovery sample.

The response to treatment as a function of rs7512462 genotype (coded additively) in the combined sample of n=24 is shown in Figure 2; using the average of all recorded FEV_1pp post-treatment measurements (15 to 400 days; Figure 2A, 9.8 increase in FEV_1pp per additional C allele, p=0.0006, n=24; Table 2; Suppl. Table 1), or with only the first FEV_1pp post-treatment measurement (up to 55 days; Figure 2B, 8.5 increase in FEV_1pp per additional C allele, p=0.004, n=21; Table 2; Suppl. Table 1). Therefore, analyses of treatment response defined by the difference between baseline and the average or the first post-treatment measurements yielded the

10

same qualitative conclusion, although the variability in response by rs7512462 genotype appears

reduced when using the single post-treatment measure within the first 2 months. The number of

post-treatment measurements per individual varied, but was not associated with baseline FEV_1pp

(p=0.51), sex (p=0.34), age (p=0.41), or treatment response (p=0.73). Whether the individual

was from the discovery or replication sample had no effect on the lung function response

(p=0.96), and therefore results are only adjusted for baseline FEV_1pp and are not unadjusted for

study sample. The variance explained in the treatment response by rs7512462 (n=24) is 28%, and

inclusion of baseline FEV_1pp results in a model that explains 34% of the variation. The TC

genotype results in an improved treatment response of 8.98 and 6.00 percentage points over

individuals with the TT genotype (p=0.01, p=0.05; Table 2) when treatment response is defined

by averaging the measurements or by using the first post-treatment measurement, respectively.

[Figure 2]

Although we include baseline FEV_1pp in all reported models to adjust for inclusion of

participants with a broad range of initial lung function values (i.e. 30-96% FEV_1pp), it did not

reach statistical significance with treatment response alone (p=0.35) or in the multivariate model

that included the rs7512462 genotype (p=0.18), and rs7512462 was not associated with baseline

FEV_1pp in a univariate analysis (p=0.52). Analysis of participants specifically with the G551D

variant (removing the three participants with non-G551D-gating mutations) resulted in a larger

estimated effect size (11.66 increase in FEV_1pp per additional C allele, p<0.0001).

It is important to determine how modification of the treatment response by *SLC26A9* may

translate more generally to other therapeutics and other CFTR genotypes. However large

numbers of individuals on other CFTR-directed therapies such as the lumacaftor (VX809) and

ivacaftor combination therapy are not yet available. Therefore we evaluated the effects of

SLC26A9 genotypes in a cell culture system devised to directly measure CFTR activity with

corrector treatment of the Phe508del mutation.

CFTR-mediated current in Phe508del CF-monolayers with VX-809 treatment is dependent on

SLC26A9 rs7512462 genotype

Given that CF lung disease severity was influenced by *SLC26A9* genotype in individuals where

mutant CFTR (G551D) is known to successfully traffic to the apical surface, where effects were

enhanced upon potentiation, we asked whether *SLC26A9* would also influence responses to

CFTR-directed therapies aimed at improved folding and trafficking of mutant CFTR to the apical

surface of epithelial cells.

Primary human bronchial epithelial monolayers of 11 individuals with CF, homozygous

for Phe508del, were assessed in an Ussing chamber following 24 hours exposure to vehicle alone

(control) or the corrector drug, VX-809 (Figure 3A). Forskolin-stimulated currents mediated by

CFTR (measured as change in current after application of forskolin, ΔI_{eq}-forskolin) were

increased following treatment with corrector VX-809 (Figure 3, more negative values correspond

to more CFTR activity). Differences in residual CFTR-mediated current across the rs7512462

genotypes in the cultures without drug treatment (DMSO, n=10) did not reach statistical

significance (p=0.09, -0.17 $\mu A/cm^2$ ΔI_{eq}-forskolin per C protective allele, Figure 3B, red).

However, CFTR-medicated currents generated with corrector VX-809 treatment (n=11, Figure

3), indicated that each additional C allele of rs7512462 was associated with improved function

(p=0.02, -0.49 $\mu A/cm^2$ ΔI_{eq}-forskolin per C protective allele, Figure 3B, blue). Taken together

with the G551D findings above, the rs7512462-*SLC26A9* allele effect upon treatment with corrector drug would be consistent with modification of surface-localized Phe508del CFTR.

[Figure 3 here]

SLC26A9 genome landscape and expression

The rs7512462 SNP (C/T) occurs within intron 5 of the two known human RefSeq *SLC26A9* transcripts (Figure 4). Associations with common variation in the protein coding portions of *SLC26A9* have not been reported in CF population samples. Data obtained from the Roadmap Epigenomics Mapping Consortium (REMC) (40, 41) and ENCODE projects (42) indicate that rs7512462 itself does not overlap with known functional annotations (Figure 4 and Suppl. Figure 1). The association evidence (on the -$\log_{10}$ p-value scale) of rs7512462 and other SNPs in the gene region with Saknorm (n=70) and with FEV_1pp ivacaftor treatment difference in the combined discovery and replication set (n=24, Suppl. Table 1) are given in Figure 4A. Linkage disequilibrium estimation using D' (Figure 4A, upper panel) indicates strong LD between rs7512462 and SNPs that extend upstream of the gene, with association evidence for both Saknorm and FEV_1pp difference also extending in this direction. This 5' region of association falls within non-coding regulatory elements. Notably one of these FEV_1pp-associated 5' variants, rs1342063 (sample MAF=0.29, p=0.002, n=24), which is also in high LD with variants reported to be associated with meconium ileus and CF-related diabetes (eg. rs4077468 (29, 39)), is only 271 base pairs from the *SLC26A9* mRNA start site occurring within and immediately adjacent to enhancer and transcription start 'state' segments, respectively, in both gastric and adult lung tissues (Figure 4B). These data suggest that rs7512462 may be marking the rs1342063 variant, which could contribute to variability in SLC26A9 expression level, although relationships

13

between the noted and other nearby variants and gene expression remain to be established.

[Figure 4 here]

The observation that modification of lung function is dependent on *CFTR* mutation type is in contrast to investigations of gene modifiers in other CF-affected organs. Previous studies have indicated that *SLC26A9* common variants are associated with CF exocrine pancreatic disease and other phenotypes close to birth, and that these associations exist in patient populations with any two severe *CFTR* mutations, the majority being Phe508del (30). The specific early timing of the CFTR-independent associated phenotypes as well as the distinct regulatory state patterns in fetal versus adult human lung tissues (Figure 4B) prompted us to consider expression of SLC26A9 in mouse tissues (Figure 4C), where notable difference in mRNA levels are apparent between late gestation and adult stages for SLC26A9 in the pancreas and the lung, and in opposing directions.

Discussion

CFTR-directed therapeutics have tremendous potential to improve outcomes in individuals with CF. However, not all individuals respond uniformly to a given therapy as both genetic background and environmental exposures vary. The rs7512462 C allele in *SLC26A9* is associated with better lung function in untreated population-based studies of Canadian and French CF individuals with at least one G551D allele or other gating mutation, and leads to notably improved response after treatment with ivacaftor in an observational study of treated patients from Canada. Although rs7512462 does not show association with lung function in CF populations homozygous for Phe508del, they do appear to influence forskolin-stimulated CFTR currents in primary cell monolayers from CF patients upon treatment with a recently developed drug to restore Phe508del CFTR maturation and cell surface delivery. Together, these findings

14

support that SLC26A9 can influence the lung function in CF, especially in the context of CFTR-directed therapies.

That the SLC26A9 benefit is realized when CFTR is present at the apical membrane is insightful with regard to how modification may be occurring. In the airways, the presence of CFTR at the membrane appears necessary, possibly to enable interaction with SLC26A9 that is then further accommodated when CFTR is functioning. These notions are consistent with the rationale of ivacaftor treatment for G551D (wherein this potentiator drug is used to obviate the gating limitation imposed by the missense change), and with earlier studies showing that SLC26A9 can enhance CFTR channel function (24-26). Further, CF individuals with the homozygous Phe508del genotype would not show benefit in lung function from their respective *SLC26A9* genotype (as we observed), without facilitation of mutant protein delivery and restoration of some CFTR function as should occur with the introduction of corrector drugs such as VX-809. Together these findings predict that SLC26A9 may influence airway outcomes in any therapeutic situation with increased apical-surface localized CFTR protein and therefore would be generalizable to pharmacological, gene therapy-driven or other efforts to increase or restore working CFTR. Of immediate interest are how the *SLC26A9* alleles are enabling benefit, and more precisely, how this is being achieved.

The CFTR-independent SLC26A9 association with the gastrointestinal phenotypes in contrast to the lungs, argue that there are unique aspects in the airway and future work must address these interesting cross-organ differences. Expression differences, as noted here using mouse tissues, may contribute to the mode of modification. In the pancreas, SLC26A9 could simply provide an alternate anion channel whereas its role in the lung may involve a more complex relationship where direct or indirect interaction with CFTR may be needed. Detailed

15

delineation of where and how SLC26A9 interacts with mutant CFTR in CF-affected tissues, as well as a better understanding of SLC26A9 function is needed.

Future development of models based on *SLC26A9* genotype that can accurately predict individual response to treatment could be used to guide companion or alternative therapeutic strategies, and patient monitoring. Since most individuals on ivacaftor benefit, we anticipate that predictive models involving the *SLC26A9* modifier will have their greatest utility in settings where CFTR-directed therapies result in more modest effects on lung function, such as for the ivacaftor and lumacaftor (VX-809) combination therapy.

Limitations of the study are important to note. We present data on a limited number of subjects in each analysis, due to the rarity of the gating mutations (~4% of individuals with CF), as well as poorly coordinated efforts to capture baseline FEV_1 measurements at ivacaftor treatment initiation. Although larger sample sizes were not available for the individual association analyses, converging evidence show consistent benefit of the C allele of rs7512462, with notably large effect sizes. Our results prompt further evaluation of *SLC26A9* as a modifier of treatment response in larger samples and in individuals on alternative therapeutics with different *CFTR* genotypes; both would be important next steps in extending the current findings. Specifically, we maintain for future studies there should be careful consideration of inclusion criteria such as availability of a baseline measure within 30 days (ideally <7 days) of treatment initiation, and restriction to younger individuals (<~40) with baseline FEV_1pp within 40-90 (as in (17)). While statistical association evidence suggests rs7512462 is a robust marker for therapeutic response, understanding the impact of *SLC26A9* alleles on *in vivo* gene expression is essential and not addressed here.

The discovery that rs7512462 is a marker of treatment response with presumed altered SLC26A9 expression level has clinical relevance in multiple settings. Knowledge of the rs7512462 genotype in an individual with CF can assist in guiding CFTR-directed therapeutic strategies, especially for those treatments with variable effects across patients. Such personalized approaches to therapy can then indicate additional monitoring or complementary considerations for patients with alleles that predict less response to treatment. For future clinical trials of new (or revisited) CFTR-directed therapies or gene therapy, stratification by rs7512462 genotype should be considered to optimize outcomes.

Together, our findings highlight the impact of genetic associations beyond *CFTR*, and their potential to personalize therapeutic approaches to improve outcomes, possibly, for all CF patients. Recognition of the clinical utility of common genetic variants found to be associated with disease has been challenging, largely due to the complexity of modeling the pertinent interactions in the correct context. Despite the earlier experimental evidence of interactions with CFTR and SLC26A9 in lung tissue models (24, 25), *SLC26A9* was identified as a modifier gene in genome-wide association studies of meconium ileus in CF (29) and of CF-related diabetes (39), but not initially for the most concerning CF feature of progressive lung disease (6, 31, 32).

That SLC26A9 may provide an alternative therapeutic target in CF has already been suggested (43, 44). In fact, our findings argue that targeting both SLC26A9 and CFTR may have benefit, but also highlight that any SLC26A9-directed therapies should also carefully consider the nature of the mutant CFTR, especially in the airways. Finally, SLC26A9 has been reported to be a target for lung diseases beyond CF, including asthma and idiopathic bronchiectasis (45, 46). Given that a common *SLC26A9* variant can discriminate lung function with pharmacological rescue of CFTR, it would appear pertinent that *CFTR* and *SLC26A9* genotype combinations be

carefully re-considered in these sino-pulmonary conditions, especially where CF carrier frequencies are known to be elevated (47-50).

Materials and Methods

Human Subjects, Phenotypes and Genotypes

The majority of participant genotyping was performed as part of the Canadian and French CF gene modifier studies using Illumina platforms or by Taqman genotyping (as described in (29, 31)). For the replication subset of individuals on ivacaftor (n=13), DNA genotyping was performed for 10 individuals using the Illumina Omni2.5 BeadChip at The Centre for Applied Genomics, The Hospital For Sick Children. Informed consent was obtained for all participants.

Data from the Canadian and French CF Gene Modifier studies were used to assess whether *SLC26A9* is a modifier of lung function using a phenotype of **S**urvival-**a**djusted averaged CF-specific FEV_1 percentiles (38) that is then **norm**alized (Saknorm) (37). The percentiles are calculated by sex, height and age since FEV1 in liters is associated with lung size. These adjustments provide for the direct comparison of patients of various ages and their differential cohort-specific mortalities. We analyzed n=56 participants from the Canadian CF Gene Modifier study (Suppl. Table 2) and n=14 from the French CF Gene Modifier Study with one G551D allele and n=1,036 and 723 individuals homozygous for Phe508del from the Canadian and French studies respectively, for association with Saknorm.

For the ivacaftor response study, only participants that consented to have their clinical data linked to their genetic data were included. There were eleven individuals with a *G551D* or other gating mutation that were enrolled in a prospective observational study of ivacaftor at the CF clinics of The Hospital for Sick Children and St. Michael's Hospital from April 2013 to

18

August 2015 and were already enrolled in the Canadian CF Gene Modifier study with rs7512462 genotype data available. These eleven comprised the discovery sample to assess whether *SLC26A9* rs7512462 was associated with response to treatment with ivacaftor. A replication sample was subsequently assembled using a retrospective observational study design. Thirteen individuals with a G551D or other gating mutation from CF clinics at University of Calgary, London Health Sciences Centre, University of Alberta, The Hospital for Sick Children and St. Michael's Hospital who were on ivacaftor and had baseline and post-treatment FEV_1 measurements were included. FEV_1pp (51) was used as the lung function measure to mirror the original clinical trial (17). All participants were taking a constant dose of 150mg of ivacaftor every 12 hours, over the entire study period.

For inclusion in either the discovery or replication ivacaftor treated sample, all individuals had to have at least one lung function measurement post-treatment and a baseline measurement between 30 and 96 FEV_1pp. The majority of participants had their baseline measurement taken on the day of treatment initiation, although patients with a baseline measurement as early as 30 days prior to initiating treatment were included (Suppl. Table 1). The number of follow-up measurements ranged from 1-6, with the majority having 3-4 post-treatment measurements; essentially aligning with routine clinic visit schedules in Canada. The first post-treatment measurement included in the study was taken as early as 15 days (Suppl. Table 1). Individuals on compassionate use of ivacaftor due to severe disease, or who were previously enrolled in the Phase II or III ivacaftor clinical trials were excluded from the study. The primary treatment response was defined as the difference between the average of their post-treatment measures (taken up to 400 days post treatment for some participants, Suppl Table 1) and their baseline measure, given an early treatment response within the first few weeks

19

generally persists, with some measure-to-measure variability (Suppl. Figure 2). For comparison, treatment response was also defined as the difference between the first post-treatment measure and their baseline measure (this analysis included only n=21 individuals who had a first measurement within 55 days of initiating treatment). Individual treatment responses (defined both ways) were then assessed for association with SLC26A9 rs7512462 genotype adjusting for their baseline FEV_1pp measurement. Adjustment for baseline FEV_1pp measurement had little affect on the rs7512462 association, but is included in all model results presented.

Ussing chamber studies with primary human bronchial epithelial cells

The primary human bronchial epithelial cells were obtained from 11 CF lung explants homozygous for the Phe508del *CFTR* mutation (as described in (52)). SLC26A9 rs7512462 genotypes included n=3 with CC, n=5 with CT and n=3 with TT. Cell monolayers were mounted in a circulating Ussing chamber (Physiological Instruments Inc), maintained at 37°C and continuously perfused with buffer at pH7.4 (with 126 mM NaCl, 24 mM $NaHCO_3$, 2.13 mM K_2HPO_4, 0.38 mM KH_2PO_4, 1 mM $MgSO_4$, 1 mM $CaCl_2$, 10 mM glucose) with symmetrical chloride concentrations in apical and basolateral chambers, and constant gassing with 5% CO_2 and 95% O_2. Transepithelial voltage was recorded and resistance was measured, following brief 1 μA current pulses every 30 seconds (53) to obtain calculated equivalent short-circuit currents (I_{eq}). CFTR function was assessed following inhibition of the epithelial Na^+ channel with amiloride (100μM, Spectrum Chemical Mfg Corp) as forskolin (10μM, Sigma Alrich)-activated currents (ΔI_{eq}-forskolin; μA/cm^2) in monolayers treated with lumacaftor (VX-809; 3 μM) or with DMSO (vehicle), for 48h (53, 54).

Statistical Analyses

All association analyses used linear regression implemented in the R statistical package (version 3.1.3, http://www.R-project.org/). Linear regression analysis of SLC26A9-rs7512462 to predict lung function severity (Saknorm) adjusted for whether the individual was from the Canadian or French CF Gene modifier study. Primary analysis coded rs7512462 additively with respect to the number of C alleles. The C allele was presumed to be the *protective* allele of rs7512462 since it was previously shown to associate with better pancreatic function at birth in CF (30), lower risk of meconium ileus at birth (29), and lower risk of CF-related diabetes (39). For comparison, genotypic models comparing CC and TC to the risk TT genotype are provided by including two indicator variables for the CC and TC genotypes in the multiple linear regression model.

Likewise, linear regression was used to determine the association between rs7512462 and FEV_1pp treatment difference in the discovery, replication and combined analyses, as well as in the functional studies. To account for the presence of related (one set of siblings) individuals in the treatment studies as well as for the possibility of heteroscedastic errors we used a robust sandwich variance estimator (version 4.4-2, https://cran.r-project.org/web/packages/rms/index.html). All analyses of treatment response were adjusted for baseline FEV_1pp measurement. Despite the specific hypothesis throughout that the C allele of rs7512462 was associated with improved CFTR function, all reported p-values are two-sided to be conservative. The variability in treatment response explained by baseline and rs7512462 was estimated by the coefficient of determination (R-squared). The treatment response for TC was compared to the TT genotype using a genotypic model as described above with the CC excluded. Univariate linear regression was used to assess the association between the numbers of post-treatment measurements and baseline FEV_1pp, sex, age or treatment response.

21

Expression analysis

Total RNA was isolated from RNAlater (QIAGEN) stabilized pancreas and lung tissues of wild type mice with C57BL/6J background using the RNeasy Mini Plus Kit (QIAGEN) according to manufacturer's instructions. Quality control, preparation of complementary DNA (cDNA) from RNA followed by quantitative real-time PCR using SYBR-green (BioRad Labs) labeling, was carried out as previously described (55). For the pancreas at E17.5, n=4; P5, n=4 and P40, n=3. For the lung at E18.5, n=4 and P55, n=3. Results are presented relative to the expression of the control gene, *Gusb*. Oligonucleotide primers for *Slc26a9* were 5'-GCTTTGTGGATTTGATGGGC and 5'-ATTTGCTTGGGCAAAGAGGA (product of 203bp); for *Cftr* were 5'-CGGAGTGATAACACAGAAAGT and 5'-CAGGAAACTGCTCTATTACAGAC (product of 140bp) and for *Gusb* were 5'-CCGATTATCCAGAGCGAGTATG and 5'-CTCAGCGGTGACTGGTTCG (product of 197bp).

Bioinformatics Analyses

To assess the genomic regulatory landscape of the *SLC26A9* region, the University of Santa Cruz (UCSC) genome browser (56) was used to query and download integrated and consolidated track data from the Roadmap Epigenomics Mapping Consortium (REMC) (40, 41) regarding predicted chromatin states and RNASeq expression evidence from CF-relevant human primary tissues. Brain hippocampus information was included as a non-CF tissue for comparison. Transcription factor (15) chromatin immunoprecipitation followed by sequencing (ChIP-seq) experimental data from multiple cell line investigations were also accessed from the

22

ReMap database, which includes data from cell lines in the ENCODE consortium (42, 57) and other public data (58). The findings were displayed in context of the association evidence for Saknorm and FEV$_1$pp treatment difference at SNPs across the region. Linkage disequilibrium patterns and haplotypes were estimated using Haploview (59), and the functional annotations were interpreted as relevant for SNPs that have a D'>0.85 with rs7512462 and provided association evidence across the *SLC26A9* locus in the population studies (*i.e* Saknorm and FEV$_1$pp with ivacaftor), or had been previously reported in other relevant studies (29, 32, 39).

Supplementary Materials

Supplementary Material is available in a separate submitted file.

Conflict of Interest statement

References

1   Sanders, D.B., Li, Z., Laxova, A., Rock, M.J., Levy, H., Collins, J., Ferec, C. and Farrell, P.M. (2014) Risk factors for the progression of cystic fibrosis lung disease throughout childhood. *Ann. Am. Thorac. Soc.*, **11**, 63-72.

2   Collaco, J.M., Vanscoy, L., Bremer, L., McDougal, K., Blackman, S.M., Bowers, A., Naughton, K., Jennings, J., Ellen, J. and Cutting, G.R. (2008) Interactions between secondhand smoke and genes that affect cystic fibrosis lung disease. *JAMA*, **299**, 417-424.

3   Goss, C.H., Newsom, S.A., Schildcrout, J.S., Sheppard, L. and Kaufman, J.D. (2004) Effect of ambient air pollution on pulmonary exacerbations and lung function in cystic fibrosis. *Am. J. Respir. Crit. Care Med.*, **169**, 816-821.

4   Cutting, G.R. (2015) Cystic fibrosis genetics: from molecular understanding to clinical application. *Nat. Rev. Genet.*, **16**, 45-56.

5   Soave, D., Corvol, H., Panjwani, N., Gong, J., Li, W., Boelle, P.Y., Durie, P.R., Paterson, A.D., Rommens, J.M., Strug, L.J. *et al.* (2015) A Joint Location-Scale Test Improves Power to Detect Associated SNPs, Gene Sets, and Pathways. *Am. J. Hum. Genet.*, **97**, 125-138.

6   Corvol, H., Blackman, S., Boelle, P., Gallins, P., Pace, R., Stonebraker, J., Accurso, F., Clement, A., Collaco, J., Dang, H. *et al.* (2015) Genome-wide association meta-analysis identifies five modifier loci of lung disease severity in cystic fibrosis. *Nature Commun.*, **6**:8382.

7   Ratjen, F., Bell, S.C., Rowe, S.M., Goss, C.H., Quittner, A.L. and Bush, A. (2015) Cystic fibrosis. *Nat. Rev. Dis. Primers*, **1**, 1-19.

8   Sosnay, P.R., Siklosi, K.R., Van Goor, F., Kaniecki, K., Yu, H., Sharma, N., Ramalho, A.S., Amaral, M.D., Dorfman, R., Zielenski, J. *et al.* (2013) Defining the disease liability of variants in the cystic fibrosis transmembrane conductance regulator gene. *Nat. Genet.*, **45**, 1160-1167.

9   Rabeh, W.M., Bossard, F., Xu, H., Okiyoneda, T., Bagdany, M., Mulvihill, C.M., Du, K., di Bernardo, S., Liu, Y., Konermann, L. *et al.* (2012) Correction of both NBD1 energetics and domain interface is required to restore DeltaF508 CFTR folding and function. *Cell*, **148**, 150-163.

10   Thibodeau, P.H., Richardson, J.M., 3rd, Wang, W., Millen, L., Watson, J., Mendoza, J.L., Du, K., Fischman, S., Senderowitz, H., Lukacs, G.L. *et al.* (2010) The cystic fibrosis-causing mutation deltaF508 affects multiple steps in cystic fibrosis transmembrane conductance regulator biogenesis. *J. Biol. Chem.*, **285**, 35825-35835.

11   Dalemans, W., Barbry, P., Champigny, G., Jallat, S., Dott, K., Dreyer, D., Crystal, R.G., Pavirani, A., Lecocq, J.P. and Lazdunski, M. (1991) Altered chloride ion channel kinetics associated with the delta F508 cystic fibrosis mutation. *Nature*, **354**, 526-528.

12   Sharma, M., Pampinella, F., Nemes, C., Benharouga, M., So, J., Du, K., Bache, K.G., Papsin, B., Zerangue, N., Stenmark, H. *et al.* (2004) Misfolding diverts CFTR from recycling to degradation: quality control at early endosomes. *J. Cell. Biol.*, **164**, 923-933.

13   Bobadilla, J.L., Macek, M., Jr., Fine, J.P. and Farrell, P.M. (2002) Cystic fibrosis: a worldwide analysis of CFTR mutations--correlation with incidence data and application to screening. *Hum. Mutat.*, **19**, 575-606.

14   Bompadre, S.G., Sohma, Y., Li, M. and Hwang, T.C. (2007) G551D and G1349D, two CF-associated mutations in the signature sequences of CFTR, exhibit distinct gating defects. *J. Gen. Physiol.*, **129**, 285-298.

15   Ramjeesingh, M., Ugwu, F., Stratford, F.L., Huan, L.J., Li, C. and Bear, C.E. (2008) The intact CFTR protein mediates ATPase rather than adenylate kinase activity. *Biochem. J.*, **412**, 315-321.

16   Van Goor, F., Hadida, S., Grootenhuis, P.D., Burton, B., Cao, D., Neuberger, T., Turnbull, A., Singh, A., Joubran, J., Hazlewood, A. *et al.* (2009) Rescue of CF airway epithelial cell function in vitro by a CFTR potentiator, VX-770. *Proc. Natl. Acad. Sci. USA*, **106**, 18825-18830.

17   Ramsey, B.W., Davies, J., McElvaney, N.G., Tullis, E., Bell, S.C., Drevinek, P., Griese, M., McKone, E.F., Wainwright, C.E., Konstan, M.W. *et al.* (2011) A CFTR potentiator in patients with cystic fibrosis and the G551D mutation. *N. Engl. J. Med.*, **365**, 1663-1672.

18   Yu, H., Burton, B., Huang, C.J., Worley, J., Cao, D., Johnson, J.P., Jr., Urrutia, A., Joubran, J., Seepersaud, S., Sussky, K. *et al.* (2012) Ivacaftor potentiation of multiple CFTR channels with gating mutations. *J. Cyst. Fibros.*, **11**, 237-245.

19   Wainwright, C.E., Elborn, J.S., Ramsey, B.W., Marigowda, G., Huang, X., Cipolli, M., Colombo, C., Davies, J.C., De Boeck, K., Flume, P.A. *et al.* (2015) Lumacaftor-Ivacaftor in Patients with Cystic Fibrosis Homozygous for Phe508del CFTR. *N. Engl. J. Med.*, **373**, 220-231.

20   Durmowicz, A.G., Witzmann, K.A., Rosebraugh, C.J. and Chowdhury, B.A. (2013) Change in sweat chloride as a clinical end point in cystic fibrosis clinical trials: the ivacaftor experience. *Chest*, **143**, 14-18.

21   Chang, M.-H., Plata, C., Sindic, A., Ranatunga, W.K., Chen, A.-P., Zandi-Nejad, K., Chan, K.W., Thompson, J., Mount, D.B. and Romero, M.F. (2009) Slc26a9 Is Inhibited by the R-region of the Cystic Fibrosis Transmembrane Conductance Regulator via the STAS Domain. *J. Biol. Chem.*, **284**, 28306-28318.

22   Loriol, C., Dulong, S., Avella, M., Gabillat, N., Boulukos, K., Borgese, F. and Ehrenfeld, J. (2008) Characterization of SLC26A9, facilitation of Cl(-) transport by bicarbonate. *Cell Physiol. Biochem.*, **22**, 15-30.

23   Ohana, E., Yang, D., Shcheynikov, N. and Muallem, S. (2009) Diverse transport modes by the solute carrier 26 family of anion transporters. *J. Physiol.*, **587**, 2179-2185.

24   Bertrand, C.A., Zhang, R., Pilewski, J.M. and Frizzell, R.A. (2009) SLC26A9 is a constitutively active, CFTR-regulated anion conductance in human bronchial epithelia. *J. Gen. Physiol.*, **133**, 421-438.

25   Avella, M., Loriol, C., Boulukos, K., Borgese, F. and Ehrenfeld, J. (2011) SLC26A9

stimulates CFTR expression and function in human bronchial cell lines. *J. Cell Physiol.*, **226**,

212-223.

26   Ousingsawat, J., Schreiber, R. and Kunzelmann, K. (2011) Differential Contribution of

SLC26A9 to Cl- conductance in polarized and non-polarized epithelial cells. *Cell Physiol.*

Biochem., **227**, 2323-2329.

27   Liu, X., Li, T., Riederer, B., Lenzen, H., Ludolph, L., Yeruva, S., Tuo, B., Soleimani, M.

and Seidler, U. (2015) Loss of Slc26a9 anion transporter alters intestinal electrolyte and

HCO3(-) transport and reduces survival in CFTR-deficient mice. *Pflugers Arch*, **467**, 1261-

1275.

28   Soave, D., Miller, M.R., Keenan, K., Li, W., Gong, J., Ip, W., Accurso, F., Sun, L.,

Rommens, J.M., Sontag, M. *et al.* (2014) Evidence for a causal relationship between early

exocrine pancreatic disease and cystic fibrosis-related diabetes: a Mendelian

randomization study. *Diabetes*, **63**, 2114-2119.

29   Sun, L., Rommens, J.M., Corvol, H., Li, W., Li, X., Chiang, T.A., Lin, F., Dorfman, R.,

Busson, P.F., Parekh, R.V. *et al.* (2012) Multiple apical plasma membrane constituents are

associated with susceptibility to meconium ileus in individuals with cystic fibrosis. *Nat.*

Genet., **44**, 562-569.

30   Miller, M.R., Soave, D., Li, W., Gong, J., Pace, R.G., Boelle, P.Y., Cutting, G.R., Drumm,

M.L., Knowles, M.R., Sun, L. *et al.* (2015) Variants in Solute Carrier SLC26A9 Modify Prenatal

Exocrine Pancreatic Damage in Cystic Fibrosis. *J. Pediatr.*, **166**, 1152-1157 e1156.

31   Wright, F.A., Strug, L.J., Doshi, V.K., Commander, C.W., Blackman, S.M., Sun, L.,

Berthiaume, Y., Cutler, D., Cojocaru, A., Collaco, J.M. *et al.* (2011) Genome-wide association

and linkage identify modifier loci of lung disease severity in cystic fibrosis at 11p13 and

20q13.2. *Nat. Genet.*, **43**, 539-546.

32   Li, W., Soave, D., Miller, M.R., Keenan, K., Lin, F., Gong, J., Chiang, T., Stephenson, A.L.,

Durie, P., Rommens, J. *et al.* (2014) Unraveling the complex genetic model for cystic

fibrosis: pleiotropic effects of modifier genes on early cystic fibrosis-related morbidities.

Hum. Genet., **133**, 151-161.

33   GTEx Consortium. (2015) The Genotype-Tissue Expression (GTEx) pilot analysis:

multitissue gene regulation in humans. *Science*, **348**, 648-660.

34   Lonsdale, J., Thomas, J. et al. (2013) The Genotype-Tissue Expression (GTEx) project.

Nat. Genet., **45**, 580-585.

35   Braun, J., Mundhenk, L., Range, F. and Gruber, A.D. (2010) Quantitative expression

analyses of candidates for alternative anion conductance in cystic fibrosis mouse models. *J.*

Cyst. Fibros., **9**, 351-364.

36   Dorfman, R., Sandford, A., Taylor, C., Huang, B., Frangolias, D., Wang, Y., Sang, R.,

Pereira, L., Sun, L., Berthiaume, Y. *et al.* (2008) Complex two-gene modulation of lung

disease severity in children with cystic fibrosis. *J. Clin. Invest.*, **118**, 1040-1049.

37   Taylor, C., Commander, C.W., Collaco, J.M., Strug, L.J., Li, W., Wright, F.A., Webel, A.D.,

Pace, R.G., Stonebraker, J.R., Naughton, K. *et al.* (2011) A novel lung disease phenotype

adjusted for mortality attrition for cystic fibrosis Genetic modifier studies. *Pediatr.*

Pulmonol., **46**, 857-869.

38   Kulich, M., Rosenfeld, M., Campbell, J., Kronmal, R., Gibson, R.L., Goss, C.H. and

Ramsey, B. (2005) Disease-specific reference equations for lung function in patients with

cystic fibrosis. *Am. J. Respir. Crit. Care Med.*, **172**, 885-891.

39   Blackman, S., Commander, C., Watson, C., Arcara, K., Strug, L., Stonebraker, J., Wright, F., Rommens, J., Sun, L., Pace, R. *et al.* (2013) Genetic modifiers of cystic fibrosis-related diabetes. *Diabetes*, **62**, 3627-3635.

40   Roadmap Epigenomics Consortium, Kundaje, A., Meuleman, W., Ernst, J., Bilenky, M., Yen, A., Heravi-Moussavi, A., Kheradpour, P., Zhang, Z., Wang, J. *et al.* (2015) Integrative analysis of 111 reference human epigenomes. *Nature*, **518**, 317-330.

41   Ernst, J. and Kellis, M. (2012) ChromHMM: automating chromatin-state discovery and characterization. *Nat. Methods*, **9**, 215-216.

42   Encode Project Consortium. (2012) An integrated encyclopedia of DNA elements in the human genome. *Nature*, **489**, 57-74.

43   Mall, M.A. and Galietta, L.J. (2015) Targeting ion channels in cystic fibrosis. *J. Cyst. Fibros.*, **14**, 561-570.

44   El Khouri, E. and Toure, A. (2014) Functional interaction of the cystic fibrosis transmembrane conductance regulator with members of the SLC26 family of anion transporters (SLC26A8 and SLC26A9): physiological and pathophysiological relevance. *Int. J. Biochem. Cell Biol.*, **52**, 58-67.

45   Anagnostopoulou, P., Riederer, B., Duerr, J., Michel, S., Binia, A., Agrawal, R., Liu, X., Kalitzki, K., Xiao, F., Chen, M. *et al.* (2012) SLC26A9-mediated chloride secretion prevents mucus obstruction in airway inflammation. *J. Clin. Invest.*, **122**, 3629-3634.

46   Bakouh, N., Bienvenu, T., Thomas, A., Ehrenfeld, J., Liote, H., Roussel, D., Duquesnoy, P., Farman, N., Viel, M., Cherif-Zahar, B. *et al.* (2013) Characterization of SLC26A9 in patients with CF-like lung disease. *Hum. Mutat.*, **34**, 1404-1414.

47   Bienvenu, T., Sermet-Gaudelus, I., Burgel, P.R., Hubert, D., Crestani, B., Bassinet, L., Dusser, D. and Fajac, I. (2010) Cystic fibrosis transmembrane conductance regulator channel dysfunction in non-cystic fibrosis bronchiectasis. *Am. J. Respir. Crit. Care. Med.*, **181**, 1078-1084.

48   Casals, T., De-Gracia, J., Gallego, M., Dorca, J., Rodriguez-Sanchon, B., Ramos, M.D., Gimenez, J., Cistero-Bahima, A., Olveira, C. and Estivill, X. (2004) Bronchiectasis in adult patients: an expression of heterozygosity for CFTR gene mutations? *Clin. Genet.*, **65**, 490-495.

49   Girodon, E., Cazeneuve, C., Lebargy, F., Chinet, T., Costes, B., Ghanem, N., Martin, J., Lemay, S., Scheid, P., Housset, B. *et al.* (1997) CFTR gene mutations in adults with disseminated bronchiectasis. *Eur. J. Hum. Genet.*, **5**, 149-155.

50   Bergougnoux, A., Viart, V., Miro, J., Bommart, S., Molinari, N., des Georges, M., Claustres, M., Chiron, R. and Taulan-Cadars, M. (2015) Should diffuse bronchiectasis still be considered a CFTR-related disorder? *J. Cyst. Fibros.*, **14**, 646-653.

51   Stanojevic, S., Wade, A., Stocks, J., Hankinson, J., Coates, A.L., Pan, H., Rosenthal, M., Corey, M., Lebecque, P. and Cole, T.J. (2008) Reference ranges for spirometry across all ages: a new approach. *Am. J. Respir. Crit. Care Med.*, **177**, 253-260.

52   Eckford, P.D., Ramjeesingh, M., Molinski, S., Pasyk, S., Dekkers, J.F., Li, C., Ahmadi, S., Ip, W., Chung, T.E., Du, K. *et al.* (2014) VX-809 and related corrector compounds exhibit secondary activity stabilizing active F508del-CFTR after its partial rescue to the cell surface. *Chem. Biol.*, **21**, 666-678.

53   Cao, H., Ouyang, H., Ip, W., Du, K., Duan, W., Avolio, J., Wu, J., Duan, C., Yeger, H., Bear, C. *et al.* (2015) Testing gene therapy vectors in human primary nasal epithelial cultures. *Mol. Ther. Methods. Clin. Dev.*, in press.

54   Yang, H., Shelat, A.A., Guy, R.K., Gopinath, V.S., Ma, T., Du, K., Lukacs, G.L., Taddei, A., Folli, C., Pedemonte, N. *et al.* (2003) Nanomolar affinity small molecule correctors of defective Delta F508-CFTR chloride channel gating. *J. Biol. Chem.*, **278**, 35079-35085.

55   Tourlakis, M.E., Zhong, J., Gandhi, R., Zhang, S., Chen, L., Durie, P.R. and Rommens, J.M. (2012) Deficiency of Sbds in the mouse pancreas leads to features of Shwachman-Diamond syndrome, with loss of zymogen granules. *Gastroenterology*, **143**, 481-492.

56   Speir, M.L., Zweig, A.S., Rosenbloom, K.R., Raney, B.J., Paten, B., Nejad, P., Lee, B.T., Learned, K., Karolchik, D., Hinrichs, A.S. *et al.* (2015) The UCSC Genome Browser database: 2016 update. *Nucleic Acids Res.*, in press.

57   Encode Project Consortium. (2011) A user's guide to the encyclopedia of DNA elements (ENCODE). *PLoS Biol.*, **9**, e1001046.

58   Griffon, A., Barbier, Q., Dalino, J., van Helden, J., Spicuglia, S. and Ballester, B. (2015) Integrative analysis of public ChIP-seq experiments reveals a complex multi-cell regulatory landscape. *Nucleic Acids Res.*, **43**, e27.

59   Barrett, J.C., Fry, B., Maller, J. and Daly, M.J. (2005) Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics*, **21**, 263-265.

60   Pollard, K.S., Hubisz, M.J., Rosenbloom, K.R. and Siepel, A. (2010) Detection of nonneutral substitution rates on mammalian phylogenies. *Genome Res.*, **20**, 110-121.

Legends to Figures

Figure 1. *SLC26A9* rs7512462 modifies lung function in CF patients with G551D**.** Boxplots of

Saknorm measures grouped by rs7512462 genotype; A, individuals homozygous for

Phe508del (n=1,759) and B, individuals with a G551D allele (n=70).

Figure 2. Response to ivacaftor is dependent on *SLC26A9* rs7512462 genotype. Boxplots of

FEV_1pp treatment difference for the combined discovery and replication samples. A,

treatment response defined as the difference between the averaged FEV_1pp measurements

within 400 days of treatment initiation and baseline measure prior to onset of treatment

(n=24); B, treatment response defined as the difference between the first FEV_1pp

measurement within 55 days of treatment initiation and baseline measure (n=21).

Differences are presented as a function of *SLC26A9* rs7512462 genotype.

Figure 3. CFTR-mediated current in primary CF bronchial monolayers treated with VX-809 is a

function of SLC26A9 rs7512462 genotype. Forskolin-stimualted currents were assessed

in cultured primary human bronchial cells from n=11 individuals with homozygous

Phe508del CF alleles treated with VX-809 or vehicle only (DMSO, n=10). CFTR

function (ΔIeq-forskolin) was measured as the current difference obtained with and

without forskolin stimulation. More negative ΔIeq-forskolin values are associated with

greater CFTR activity. A, Individual trajectories of ΔIeq-forskolin measures with vehicle

followed by VX-809 treatments are shown color-coded by rs7512462 genotype; one

33

sample treated with VX-809 had a missing baseline measurement. B, Boxplots of the

ΔIeq-forskolin with VX-809 treatment (blue) and DMSO vehicle measurements (red)

stratified by rs7512462 genotype.

Figure 4. Regional association evidence, genomic landscape and expression of SLC26A9.

SLC26A9 with at least two major RefSeq transcript isoforms (in human) is shown

aligned with genetic, physical and regulatory landmarks. SNPs with p-values <0.05 for

both Saknorm and FEV_1pp difference, and other variants of interest are indicated with

light blue shaded vertical lines. A, Regional association evidence. Linkage disequilibrium

(LD, upper panel; D') was estimated using Canadian CF Gene Modifier Consortium

participants (n=1,659; (29)), revealing strong D' between rs7512462 and previously

reported CF-associated variants including rs4077468 upstream of the gene (29, 39), and

weaker D' with an asthma-associated region in the 3' untranslated region, including

rs12733647 (45). Pairwise relationships with no numeric value indicate complete LD

(D'=1). The stronger LD between rs7512462 and variants upstream of the gene lead us to

favor the 5' region as contributing to the phenotypic variability. P-values for association

evidence of Saknorm (n=70) and combined ivacaftor studies (difference in FEV_1pp,

n=24) are shown as separate tracks across the interval with black vertical lines. B,

Regulatory features of the SLC26A9 locus. Chromatin state prediction tracks are shown

aligned to SLC26A9 transcripts based on chromHMM modeling (41) using the integrated

and consolidated Roadmap Epigenomics Mapping Consortium (REMC) data, which

consists of histone methylation and DNase hypersensitivity assays in primary human

tissues (40). A cluster of 'active' features including enhancers, active transcription start

sites and/or transcription are indicated in adult gastric, lung and pancreas tissues. Features

corresponding to active repression or no expression are evident in other tissues shown.

Mammalian conservation (60) and transcription evidence based on RNASeq studies from

REMC (40) are shown aligned below the chromHMM tracks. C, High expression of

Slc26a9 in young murine pancreas. Transcript expression of Slc26a9 and Cftr in mouse

pancreas (left panel) and lung (right panel) tissues. Late gestation (E) and post birth (P)

time points are indicated in days. Expression levels are normalized to the control gene,

Gusb.

Tables

Table 1. Association between Saknorm and rs7512462 under both additive and genotypic models.

Phe508del[♭] (n = 1,759)				
Additive Model	Estimate	Standard Error	T statistic	p-value
rs7512462[♮]	0.03434	0.0267	1.286	0.1986
Cohort*	0.08185	0.0384	2.131	0.0332
Genotypic Model				
TC	0.0216	0.04202	0.514	0.6073
CC	0.07384	0.05502	1.342	0.1797
Cohort*	0.08174	0.03841	2.128	0.0335

G551D[♯] (n = 70)				
Additive Model	Estimate	Standard Error	T statistic	p-value
rs7512462[♮]	0.3084	0.1497	2.061	0.0432
Cohort*	-0.2895	0.2485	-1.165	0.2481
Genotypic Model				
TC	0.2091	0.2162	0.967	0.3370
CC	0.6924	0.3231	2.143	0.0358
Cohort*	-0.2882	0.2496	-1.155	0.2524

[♭]Association analysis in individuals homozygous for Phe508del CF-causing variants

[♮]rs7512462 was coded additively as the number of C alleles

[♯]Association analysis in individuals with at least one G551D CF-causing variant.

*Cohort refers to French or Canadian sample

[+]There are no additional covariates included in the model, because the phenotype Saknorm is a normalized function of FEV1 adjusted for age, sex, and cohort-specific survival.

Table 2. Association between ivacaftor treatment response and rs7512462 for additive and genotypic models

Treatment Response defined as averaged FEV1pp measures on Ivacaftor				
Additive Model (n=24)	Estimate	Standard Error[#]	T statistic	p-value
Baseline	-0.1477	0.1055	-1.4	0.1761
rs7512462	9.8059	2.4104	4.07	0.0006
Genotypic Model* (n=24)				
Baseline	-0.1526	0.105	-1.45	0.1615
TC	8.9809	3.266	2.75	0.0123
CC^	22.5572	2.331	9.68	<0.0001

Treatment Response defined as first FEV1pp measure on Ivacaftor				
Additive Model (n=21)	Estimate	Standard Error[#]	T statistic	p-value
Baseline	-0.1124	0.0844	-1.33	0.1994
rs7512462	8.4982	2.583	3.29	0.0041
Genotypic Model* (n=20)				
Baseline	-0.1284	0.0823	-1.56	0.1371
TC	6.0027	2.8141	2.13	0.0478
CC^	24.4784	2.0671	11.84	<0.0001

*Genotypic model compares the improvement in treatment response for TC over the TT genotype.

^There is only one individual with the CC genotype.

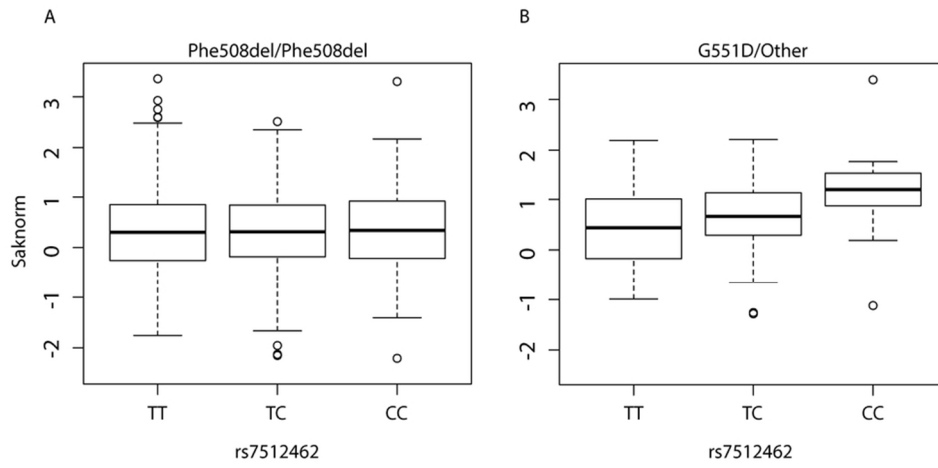# Robust standard errors calculated using the robcov function in R.

Figure 1. SLC26A9 rs7512462 modifies lung function in CF patients with G551D. Boxplots of Saknorm measures grouped by rs7512462 genotype; A, individuals homozygous for Phe508del (n=1,759) and B, individuals with a G551D allele (n=70).
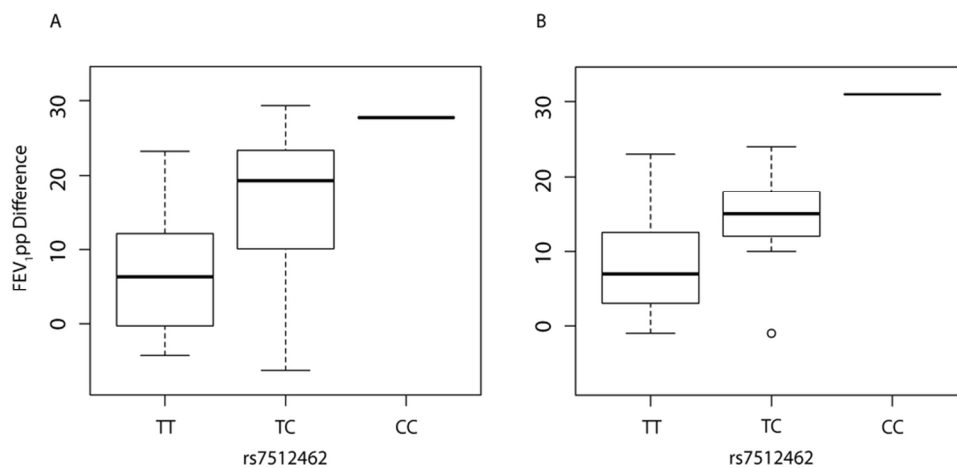
87x42mm (300 x 300 DPI)

Figure 2. Response to ivacaftor is dependent on SLC26A9 rs7512462 genotype. Boxplots of FEV1pp treatment difference for the combined discovery and replication samples. A, treatment response defined as the difference between the averaged FEV1pp measurements within 400 days of treatment initiation and baseline measure prior to onset of treatment (n=24); B, treatment response defined as the difference between the first FEV1pp measurement within 55 days of treatment initiation and baseline measure (n=21). Differences are presented as a function of SLC26A9 rs7512462 genotype.
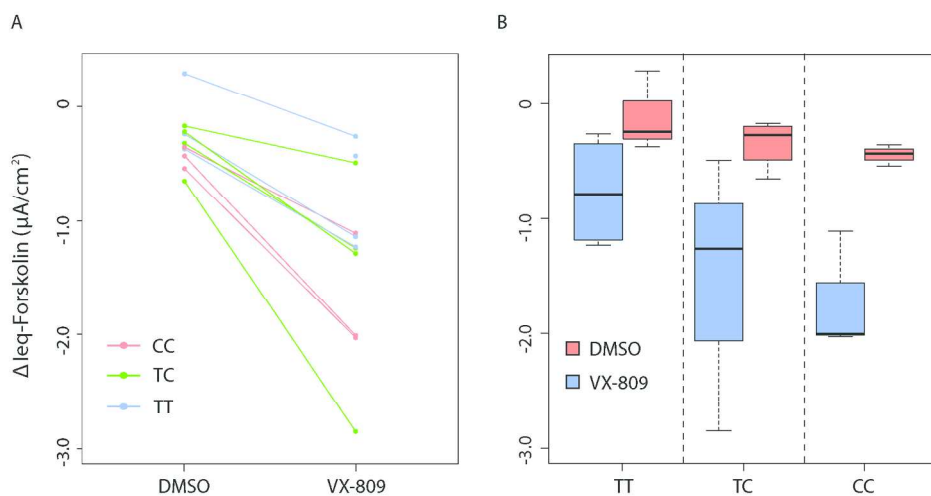
87x42mm (300 x 300 DPI)

Figure 3. CFTR-mediated current in primary CF bronchial monolayers treated with VX-809 is a function of SLC26A9 rs7512462 genotype. Forskolin-stimualted currents were assessed in cultured primary human bronchial cells from n=11 individuals with homozygous Phe508del CF alleles treated with VX-809 or vehicle only (DMSO, n=10). CFTR function (ΔIeq-forskolin) was measured as the current difference obtained with and without forskolin stimulation. More negative ΔIeq-forskolin values are associated with greater CFTR activity. A, Individual trajectories of ΔIeq-forskolin measures with vehicle followed by VX-809 treatments are shown color-coded by rs7512462 genotype; one sample treated with VX-809 had a missing baseline measurement. B, Boxplots of the ΔIeq-forskolin with VX-809 treatment (blue) and DMSO vehicle measurements (red) stratified by rs7512462 genotype.
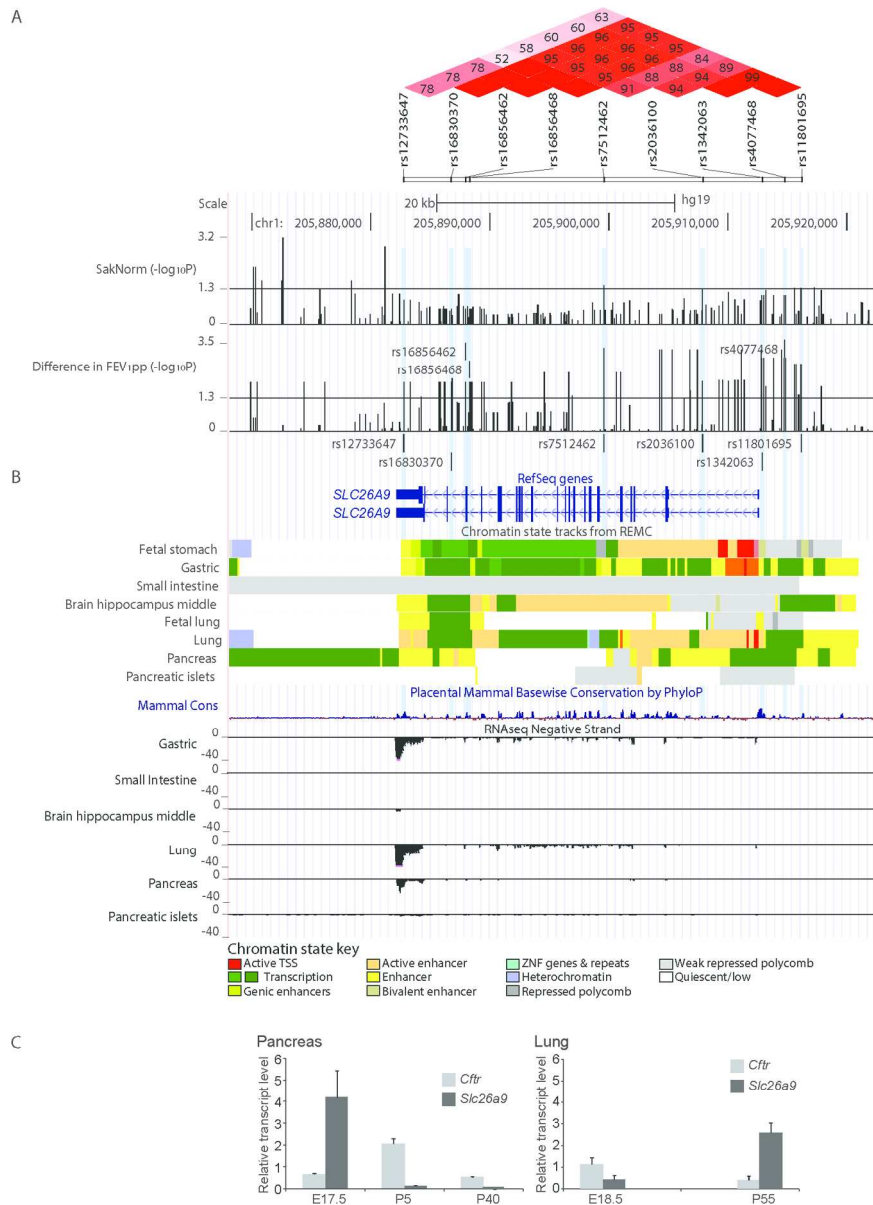
180x92mm (300 x 300 DPI)

Figure 4. Regional association evidence, genomic landscape and expression of SLC26A9. SLC26A9 with at least two major RefSeq transcript isoforms (in human) is shown aligned with genetic, physical and regulatory landmarks. SNPs with p-values <0.05 for both Saknorm and FEV1pp difference, and other variants of interest are indicated with light blue shaded vertical lines. A, Regional association evidence. Linkage disequilibrium (LD, upper panel; D') was estimated using Canadian CF Gene Modifier Consortium participants (n=1,659; (29)), revealing strong D' between rs7512462 and previously reported CF-associated variants including rs4077468 upstream of the gene (29, 39), and weaker D' with an asthma-associated region in the 3' untranslated region, including rs12733647 (45). Pairwise relationships with no numeric value indicate complete LD (D'=1). The stronger LD between rs7512462 and variants upstream of the gene lead us to favor the 5' region as contributing to the phenotypic variability. P-values for association evidence of Saknorm (n=70) and combined ivacaftor studies (difference in FEV1pp, n=24) are shown as separate tracks across the interval with black vertical lines. B, Regulatory features of the SLC26A9 locus. Chromatin state prediction tracks are shown aligned to SLC26A9 transcripts based on chromHMM modeling (41) using the

integrated and consolidated Roadmap Epigenomics Mapping Consortium (REMC) data, which consists of histone methylation and DNase hypersensitivity assays in primary human tissues (40). A cluster of 'active' features including enhancers, active transcription start sites and/or transcription are indicated in adult gastric, lung and pancreas tissues. Features corresponding to active repression or no expression are evident in other tissues shown. Mammalian conservation (60) and transcription evidence based on RNASeq studies from REMC (40) are shown aligned below the chromHMM tracks. C, High expression of Slc26a9 in young murine pancreas. Transcript expression of Slc26a9 and Cftr in mouse pancreas (left panel) and lung (right panel) tissues. Late gestation (E) and post birth (P) time points are indicated in days. Expression levels are normalized to the control gene, Gusb.

141x195mm (300 x 300 DPI)