



HAL
open science

Genes that determine immunology and inflammation modify the basic defect of impaired ion conductance in cystic fibrosis epithelia

Frauke Stanke, Tim Becker, Vinod Kumar, Silke Hedtfeld, Christian Becker, Harry Cuppens, Stephanie Tamm, Jennifer Yarden, Ulrike Laabs, Benny Siebert, et al.

► To cite this version:

Frauke Stanke, Tim Becker, Vinod Kumar, Silke Hedtfeld, Christian Becker, et al.. Genes that determine immunology and inflammation modify the basic defect of impaired ion conductance in cystic fibrosis epithelia. *Journal of Medical Genetics*, 2010, 48 (1), pp.24. 10.1136/jmg.2010.080937 . hal-00565395

HAL Id: hal-00565395

<https://hal.science/hal-00565395>

Submitted on 12 Feb 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Genes that determine immunology and inflammation modify the basic defect of impaired ion conductance in cystic fibrosis epithelia

Running Title: Host defense genes modify CF basic defect

Frauke Stanke,¹ Tim Becker,² Vinod Kumar,¹ Silke Hedtfeld,¹ Christian Becker,¹ Harry Cuppens,³ Stephanie Tamm,¹ Jennifer Yarden,³ Ulrike Laabs,¹ Benny Siebert,¹ Luis Fernandez,¹ Milan Macek Jr.,⁴ Dragica Radojkovic,^{3,5} Manfred Ballmann,¹ Joachim Greipel,⁵ Jean-Jacques Cassiman,³ Thomas F. Wienker² & Burkhard Tümmler¹

¹ *Department of Pediatrics, Hannover Medical School, Hannover, Germany*

² *Institute of Medical Biometry, Informatics and Epidemiology, University of Bonn, Bonn, Germany*

³ *Department of Human Genetics, University of Leuven, Leuven, Belgium*

⁴ *Department of Biology and Medical Genetics, Charles University Prague, Medical School and University Hospital Motol, Prague, Czech Republic*

⁵ *Institute of Molecular Genetics and Genetic Engineering, Belgrade, Serbia*

⁶ *Department of Biophysical Chemistry, Hannover Medical School, Hannover, Germany*

Corresponding author:

Dr. rer. nat. Frauke Stanke
Department of Pediatrics OE6710
Hannover Medical School
Carl-Neuberg-Strasse 1
30625 Hannover
Germany
Phone: +49-511-532-6722
Fax: +49-511-532-6723
e-mail: mekus.frauke@mh-hannover.de

Key words: cystic fibrosis, modifier genes, association study, affected sib pairs, CF basic defect

Word count: 3428 words (I,M,R,D). Abstract 243 words. The MS contains 4 figures. Supplementary material is provided attached to the MS text.

ABSTRACT

Background The CF basic defect, caused by dysfunction of the apical chloride channel CFTR in the gastrointestinal and respiratory tract epithelia, has not been employed so far to support the role of CF modifier genes.

Methods Patients were selected from 101 families with a total of 171 F508del-*CFTR* homozygous CF patients to identify CF modifying genes. We performed a candidate gene based association study of 52 genes on 16 different chromosomes with a total of 182 genetic markers. Differences in haplotype and/or diplotype distribution between case and reference CF subpopulations were analyzed.

Results Variants at immunologically relevant genes were associated with the manifestation of the CF basic defect ($0.01 < P_{\text{raw}} < 0.0001$ at *IL1B*, *TLR9*, *TNF α* , *CD95*, *STAT3* and *TNFR*). The intragenic background of F508del-*CFTR* chromosomes determined disease severity and manifestation of the basic defect ($P_{\text{raw}} = 0.0009$). Allele distributions comparing transmitted and non-transmitted alleles were distorted at several loci unlinked to *CFTR*.

Conclusions The inherited capabilities of the innate and adaptive immune system determine the manifestation of the CF basic defect. Variants on F508del-*CFTR* chromosomes contribute to the observed patient-to-patient variability among F508del-*CFTR* homozygotes. A survivor effect, manifesting as a transmission disequilibrium at many loci, is consistent with the improvement of clinical care over the last decades resulting in a depletion of risk alleles at modifier genes. We conclude that awareness of non-genetic factors such as improvement of patient care over time is crucial for the interpretation of CF modifier studies.

INTRODUCTION

Cystic fibrosis (CF) is the most common severe monogenic autosomal-recessively inherited disease affecting 1 out of 3200 newborns in the Caucasian population.¹ The manifestation of the multi-organ disease varies among individuals and even among siblings homozygous for the most frequent disease-causing lesion F508del-*CFTR*.²⁻⁴ Improvement of medical care throughout the last decades has increased the patient's life expectancy from less than 5 years in the 1940s to 37 years estimated for the current CF population,^{5,6} demonstrating convincingly that mainly non-genetic factors determine the disease outcome. Hallmarks of improved symptomatic therapy are the use of antibiotics and nutritional repletion in centers specialized for CF care starting in the mid-1950s⁵ and the development of specialized drugs such as novel pancreatic enzymes or recombinant DNase since the late 1980s.⁵ Center-to-center variations of quality-of-care⁷ and the influence of the socioeconomic status on the course of CF disease⁶ furthermore emphasize the influence of non-inherited factors on disease manifestation. Consequently, attempts to uncover clinically meaningful modifier genes require that non-genetic confounding factors are taken into account.

The basic defect in cystic fibrosis is defined as an impaired chloride conductance in epithelia that express *CFTR*, a chloride channel localized in the apical membrane of differentiated epithelial cells.¹ Nasal potential difference measurement (NPD) and intestinal current measurement (ICM) allow the assessment of the CF basic defect in the two major affected organ systems in CF, i.e. the respiratory and gastrointestinal tract.^{8,9} Even though the manifestation of the basic defect on the cellular and epithelial level is less vulnerable to environmental influences than any clinical outcome parameter such as lung function, growth and stature, substantial variability of the basic defect phenotype among sib pairs with shared *CFTR* mutation genotype is observed.^{10,11} We hypothesize that CF genetic modifiers can be recognized through a genotype-phenotype association with the basic defect.

If the influence of modulating genes on the CF phenotype is at the core of interest, the power of the study to detect a modifying gene will increase with the standardization of the *CFTR*

mutation genotype^{1,2,5} and environmental influences^{1,5-7} as the two major confounding variables of CF disease severity. While the high frequency of the F508del-*CFTR* allele among Caucasians^{1,2} allows the recruitment of a sufficient number of patients who are homozygous for the same *CFTR* disease causing lesion,^{3,12,13} standardization of the environment cannot be achieved for a human study population. Hence, the best choice are twin and sibling patient pairs to minimize many environmental effects as siblings will share their family environment, visit the same physician and be treated at the same CF center. Monozygous twins, being genetically identical, and dizygous patient pairs, sharing half of their genome, can be compared to deduce the influence of inherited factors on the phenotype.¹⁴

The *European Cystic Fibrosis Twin and Sibling Study Consortium*³ has collected clinical data on more than 500 CF twin and sibling pairs. We have previously reported how informative patient pairs with extreme phenotypes can be selected from this cohort for candidate gene analysis³ and described CF modifier genes on 6 Chromosomes.¹⁵⁻²² In this study, we follow the objectives to: 1. qualify the relative impact of the *CFTR* gene, environmental and other inherited factors on CF disease severity and 2. to identify genetic modifiers for the CF basic defect and 3. to observe whether or not these are the same genes that modify CF disease severity. We have tested a total of 182 genetic markers targeting 52 candidate gene loci on 16 different chromosomes for their association with CF disease severity and the manifestation of the *CFTR*-mediated basic defect by genotyping a cohort composed of 101 families with a total of 171 F508del-*CFTR* homozygous patients. Among these are 12 genes studied as CF modifiers by others,^{12,13,23-26} 26 additional genes which represent plausible modifier candidates such as direct interaction partners of *CFTR*, alternative ion channels and components of the innate immune system and 14 candidate genes which were selected by transcriptome analysis comparing intestinal epithelial biopsies of controls and CF patients. To our knowledge, this is the only study on CF modifiers that combines the powerful approach to analyze affected sib pairs¹³ while at the same time exploiting the phenotypic contrast between informative extreme clinical phenotypes.¹²

METHODS

Patients & Phenotypes

This work reports upon clinical data from 466 twin and sibling pairs and genetic data from 101 CF families, 85 of which are a subgroup of the twin and sibling study cohort. Enrollment of patient pairs into the association study was based on one of the following criteria: the twin or sibling pair exhibits an extreme clinical phenotype³ and/or the patient was characterized for the CF basic defect.^{10,11} 16 unrelated F508del-*CFTR* homozygotes from the CF clinic in Hannover were included who participated in an analysis of basic defect and transcriptome in the intestine to define 14 candidate genes (van Barneveld et al, unpublished).

We have selected the 12% most informative pairs from the entire sample of 318 CF twin and sibling pairs by a ranking algorithm³ that relies on two clinical parameters most sensitive for course and prognosis of CF.^{27,28} These 37 dizygous F508del-*CFTR* homozygous patient pairs, enrolled due to their contrasting extreme clinical phenotype (Figure 1), were comparable with respect to their birth cohort.¹⁸

CF basic defect phenotypes derived from intestinal current measurement (ICM) and nasal potential difference measurement (NPD) were obtained for 71 patient pairs.^{10,11} Briefly, secretagogues that activate or block ion channels, ion exchangers or components of the cellular signal transduction pathways were applied by superfusion of the surface of the lower nasal turbinate⁹ or to excised rectal suction biopsies mounted in a micro-Ussing chamber⁸ using a fixed sequence of defined pharmacologically active substances.

Based on the sib pairs' clinical phenotype and the individuals manifestation of the CF basic defect, subgroups were defined as cases and references for an association study (Figure 1). These subgroups are by definition partially (clinical phenotype vs. ICM) or nearly completely (ICM and NPD phenotypes) overlapping sets of patient pairs (see Supplement for details).

This study was approved by the ethics committee of the Medizinische Hochschule Hannover. Written informed consent was obtained from all participants or their parental guides.

Candidate Genes & Genetic Markers

Genotypes of two indels, 101 SNPs and 79 microsatellites were analyzed for this study. 73 markers were developed de novo for this study building on raw genomic sequences containing microsatellite repeats within or near the targeted candidate gene to develop polyallelic markers (40 microsatellites) or on the public SNP database (33 diallelic markers). To compensate for the loss of power inflicted upon the study by the limited sample size of subgroups stratified for extreme clinical phenotypes (being infrequent by definition) or for basic defect phenotypes such as F508del-*CFTR* homozygotes who display *CFTR*-mediated residual function (being infrequent due to the etiology of the disease), care was taken to include only informative markers in the study: between one and 13 microsatellite motifs (median: 4 motifs) were tested per candidate gene locus to obtain one marker with a suitable polymorphism information content of 0.61 – 0.91 (mean 0.76) per targeted candidate gene. SNP markers were preselected based on HAPMAP data.²⁹ Technical details on all markers and genotyping methods are provided in the Supplement.

Genetic Modelling

Genetic modeling was done with the LISREL8 software package.³⁰ Briefly, covariance matrices from standard normal distributions derived from weight for height and FEV1%pred data were used to fit a model taking genetic effects as well as random and shared environmental effects into account. Applicability of the model was judged by the Aikaike information criterion.³⁰

Association study

Genetic data for the association study was evaluated using the FAMHAP software package³¹ which allows family-based analysis^{32,33} and accepts data evaluation in association studies on unrelated individuals as well as on affected sib pairs.³¹ All case-reference comparisons were carried out using 10,000 Monte-Carlo simulated data sets.³¹⁻³³ Nuclear families were analyzed by the transmission-disequilibrium test (TDT)³⁴ extended to both nuclear families with more than one affected child and

to multi-marker haplotypes.³¹⁻³³ To allow a comparable assignment of weighted haplotype explanations, the entire genotyping data was provided as training set to FAMHAP for all case-reference comparisons.

All candidate genes were analyzed for an underrepresentation of alleles among transmitted chromosomes (TDT) as well as by case-reference comparisons for their association with disease severity (two case-reference comparisons), ICM (two case-reference comparisons) or NPD (three case-reference-comparisons). Assignment of case and reference status are outlined in Figure 1 and detailed in the Supplement. To confirm observed associations, we have employed within-data-set validation strategies using partial replicates as detailed in the Supplement. Significance within the entire data set of 182 markers was conservatively judged by Bonferroni-correction for 200 independent markers (threshold $P_{\text{raw}} = 3 \times 10^{-4}$ for $\alpha=0.05$) or 52 independent loci (threshold $P_{\text{raw}} = 9 \times 10^{-4}$ for $\alpha=0.05$).

RESULTS

Inherited, environmental and maternal factors influence CF disease manifestation

We have compared intrapair differences in weight expressed as % of predicted weight for height (wfh%) and CF-population centiles for the lung function parameter FEV1% (FEVPer_c) and conducted genetic modeling on 466 (wfh%) and 318 (FEV1%) patient pairs (Figure 2). While monozygous twins were more concordant in their FEVPer_c than either dizygous twins or sibs, all twins were more concordant than sibs in their wfh% irrespective of their zygosity status. This indicates that the shared pre- and early postnatal period which distinguishes twins from sibs is associated with lower intrapair differences in wfh% while concordance in lung function is dominated by genetic factors as only monozygous twins share their entire genetic information in contrast to dizygous twins or siblings who have half of their genetic information in common. The results of the genetic modeling analysis imply that inherited factors have a larger impact on the patient-to-patient variation in wfh% than for lung function (Figure 2B). These data are in accordance with heritability estimates from other CF twin and sibling cohorts.⁴

Association study

Candidate genes were interrogated for their association with disease severity using the phenotypic contrast between concordant/mildly affected patient pairs, concordant/severely affected patient pairs and discordant patient pairs. For this purpose, 37 dizygous pairs – representing the 12% most informative patient pairs of the entire sample – were selected based on a non-parametric ranking algorithm as described previously.³ NPD data were used firstly to inquire for an association with sodium transport via the amiloride-sensitive proportion of the potential, using the upper and lower 30% of the entire sample to define extreme phenotypes, secondly to evaluate an association with residual chloride secretion activated by chloride-free gluconate solution and isoproterenol and thirdly to analyze a contribution by the selected candidate genes on ATP-stimulated chloride

response. Basic defect assessed by ICM was evaluated using a set of patients devoid of residual chloride secretion as controls. The pleiotropic chloride-channel inhibitor DIDS, to which CFTR is not sensitive, was used to discriminate CFTR-mediated residual chloride secretion from chloride secretion mediated by alternative channels which are DIDS-sensitive.^{10,11} After correction for multiple testing, significance was retained at the following loci and for the following phenotypes (Figure 3): 7q31.1/*CFTR*: $P=0.0009$ (NPD_Na+) and $P_{\text{raw}}=0.0008$ (NPD_ATP); 7q34/D7Sat3: $P_{\text{raw}}=0.0005$ (disease severity cis); 11q13/*GSTP1* $P_{\text{raw}}=0.0006$ (NPD_ATP); 12p13/*TNFR* $P_{\text{raw}}=0.0004$ (NPD_GI/Iso) and $P_{\text{raw}}=0.0003$ (NPD_ATP); 12q13/*KRT8* $P_{\text{raw}}=0.0004$ (ICM_Res); 16p13.3/*NHERF2* $P_{\text{raw}}=0.0001$ (ICM_Res). We have previously published association with DIDS-sensitive residual chloride secretion in ICM at the *CLCA*-gene cluster¹⁵ and association with disease severity for *LEP*,¹⁶ *CD95*,¹⁷ *TNFR*,¹⁸ *SCNNIB*,¹⁸ *SCNNIG*,¹⁸ at the *CEACAM*-gene cluster,^{19,20} and at a paternally imprinted gene on 7q34.²¹ Association of these genes to other phenotypes are reported here for the first time.

Informative microsatellite markers allow the assessment of candidate genes as CF modifiers

Most candidate gene association studies in CF rely on the typing of polymorphisms which are suspected or known to alter the performance of the gene or its protein product,^{12,23} a strategy that cannot be applied as long as functional variants are unknown to date. Building on our previous experience,^{18,22} we developed informative microsatellite markers to interrogate these genes (see Supplement for details). Results of the association study are displayed in Figure 3. To confirm initial findings by microsatellite markers, a low-density SNP scan was done for *PRSS8*, *KRT8/KRT18*, *TLR4* and *IL1B*, and previously obtained genotyping data¹⁸ was evaluated for basic defect phenotypes at *SCNNIB/SCNNIG* and *TNFR/SCNNIA*. Significance after correction for multiple testing was retained for the association of the member of the CFTR network *NHERF2* with DIDS-insensitive residual chloride secretion detected by ICM ($P_{\text{raw}}=0.0001$, diplotype model; $P_{\text{raw}}=0.0007$, haplotype model; $P_{\text{raw}}<10^{-5}$, accumulation of rare alleles among patients with residual

function; Figure 3 and Figure 4). *NHERF2* was interrogated with two microsatellite motifs located in a distance of 30kb and 70kb. 7 further genes are encoded on the segment between marker and targeted candidate gene. Even though none of these can be reasonably associated with CF pathophysiology, neighbours of the targeted candidate gene have to be considered: We observed an association of the response to gluconate and isoproterenol in NPD at the *TNFR/SCNNIA* locus (Figure 3). Haplotype analysis of the 12p13 locus with four intragenic markers for each of the two neighboring genes *TNFR* and *SCNNIA* revealed that the effect is mediated by the cytokine receptor *TNFR* and not by the alpha subunit of the amiloride-sensitive epithelial sodium channel *SCNNIA* (Figure 4).

Diversity of the F508del-*CFTR* genetic background determines the clinical variability of F508del-*CFTR* homozygotes

We have included intragenic and *CFTR*-flanking markers in our association study. As the enrolled patients are F508del-*CFTR* homozygous, the major disease causing variant is identical for all study subjects. In contrast, these F508del-*CFTR* chromosomes differ at closely linked neighboring SNPs.³⁵⁻³⁷ We tested in our association study whether these different F508del-*CFTR* alleles are asymmetrically distributed when comparing mildly and severely affected patients or subgroups stratified for the manifestation of the basic defect. We observed significant association of the *CFTR* core haplotype to the response of the nasal epithelium to superfusion with amiloride ($P_{\text{raw}}=0.0009$; Figure 3 and Figure 4) and ATP ($P_{\text{raw}}=0.0008$; Figure 3 and Figure 4). Moderate association of *CFTR* was observed to CF disease severity and response to gluconate and isoproterenol in NPD ($0.001 < P_{\text{raw}} < 0.01$; Figure 3 and Figure 4). These findings demonstrate that the different F508del-*CFTR* alleles are functionally nonequivalent. In other words, sequences adjacent to the F508del-*CFTR* mutation which are phylogenetically younger than the F508del variant and/or sequences in the vicinity of the *CFTR* gene allowing for allelic diversity among F508del chromosomes through recombination have a functional impact on the manifestation of the CF basic defect and disease

severity. We would like to emphasize the point that variant F508del-*CFTR* genes ($P_{\text{raw}}=0.0009$) more than ENaC variants ($P_{\text{raw}}=0.0695$ at locus *SCNNIB/SCNNIG*; $P_{\text{raw}} = 0.0183$ at T663A in *SCNNIA*) determined the sensitivity of the CF nasal epithelium to amiloride in our study population (Figure 3).

Host defense capabilities affect the manifestation of the CF basic defect

We have selected cytokines as well as cytokine and pathogen receptors as candidate genes and tested these for an association with basic defect phenotypes. Minor ($0.01 < P_{\text{raw}} < 0.1$; *TLR5*, *CD14*, *IFNGR1*, *TLR4*, *TGFB1*), moderate ($0.001 < P_{\text{raw}} < 0.01$; *IL1B*, *TLR9*, *TNF α* , *CD95*, *STAT3*) and significant ($0.0001 < P_{\text{raw}} < 0.001$; *TNFR*) associations of host defense genotype to basic defect phenotype were observed (Figure 3). As none of these genes encode ion channels or their intermediate interaction partners, these genotype-phenotype relations were a priori unanticipated. However, crosstalk between the immune system and secretory properties of epithelial cells have been described before: for cytokines IL10, IFN γ , TNF α , IL1 β , TGF β , IL4 and IL13, alterations of CFTR and/or ENaC and/or CaCC expression and/or activity, resulting in altered ion and fluid transport, have been communicated.³⁸⁻⁴¹ This has been interpreted as an unspecific host response to pathogens as the surface liquid is increased on the epithelium upon induction of a hypersecretory state by inflammatory stimuli,³⁹ and consequently mucociliary clearance of the intruding pathogen will be alleviated. Our data shows that the effect of the host defense system on ion conductance properties of CF epithelia outweighs the effect of major ion channels which we also analyzed in our association study.

DISCUSSION

Modifier genes have now been studied in CF since Zielenski et al described CFM1 in 1999.²³ Among those studies which analyze several candidate genes at once, one modifier only is reported.^{12,26} On the other hand, negative outcomes in these studies are often contradicted by other isolated reports (reviewed by Cutting²⁴) and genetic modeling has lead to the expectation that many more modifiers need to be uncovered in CF.^{4,25} This ambiguity is likely to be understood if non-inherited factors,^{1,6} which act differently between CF centers⁷ as well as between patients from different birth cohorts,⁵ are taken into account. Improvement of therapy, quality of life and life expectancy is at the heart of medical care for all diseases. However, the increase in life span among CF patients within the last decades is outstanding in comparison to other chronic diseases, signifying that better medication and improvement of therapeutic management dramatically alter the environment with which our patients have to cope. As a consequence for the human geneticist studying a CF modifier in a cross-sectionally recruited patient sample, carriers of high-risk alleles will be less frequent, especially among early-born patient cohorts.²⁰ Consequently, case-control-comparisons might lack power to detect an allelic imbalance if the frequency of the risk alleles within the entire CF population is reduced owing to the survivor effect. If parental genotype information is available, the allelic bias introduced by a survivor effect can be visualized by comparing transmitted and non-transmitted alleles within CF families.¹⁸ In support of this, many loci reported upon within this manuscript are positive in the transmission disequilibrium test (Figure 3).

The impact of non-inherited factors on the course of CF disease, and moreover the change of these confounders over time challenges classical strategies to validate findings of association studies. Ideally, replication studies with large sample sizes are used to verify findings. However, large CF patient sample are likely to be heterogeneous and thus will not share the major environmental determinants of CF disease progression. Taking this into account, we choose to replicate the findings of our association study using confirmatory CF endophenotypes and within-

data-set validation strategies. No evidence for the role of the gene as a CF modifier was found for six loci. Isolated findings – defined as an association to only disease severity or only a basic defect phenotype were observed for 20 genes. The remaining 26 loci analyzed in this study are positive for both, an association with a clinical phenotype and a CF basic defect phenotype (see Figure 3 and Supplement - Table 8). As the basic defect in CF is correlated with disease severity¹¹ these consistent findings must be interpreted with caution with respect to the underlying causality: either the gene analyzed has a direct influence on the efficiency of F508del-CFTR processing and trafficking and thus causes a better phenotype. Alternatively, the patient subsamples analyzed for CFTR-mediated residual current in ICM or NPD can be considered as enriched with individuals who are mildly diseased due to the association of mild disease with residual function in respiratory and intestinal tissue.¹¹ In the latter case, the findings of the association study do not necessarily imply a direct causal action of the analyzed modifier on the CF basic defect but can be interpreted as a supporting finding in two independent patient subsamples. As an example, the mild *TNFR* haplotype (allele 1-10-2-2 at markers rs767455, D12S889, rs1800692 and rs1800693)¹⁸ is significantly enriched on transmitted chromosomes in entire European CF sibling population as indicated by TDT (Ref. 18 and Figure 3) and is associated with the concordant mildly affected phenotype (Ref. 18 and Figure 3) as well as with residual chloride secretion of the nasal epithelium (Figure 4). As a causal interpretation, this might reflect the interaction of the cytokine pathway and ion secretory properties of the epithelium,³⁸⁻⁴¹ promoting residual chloride secretion among carriers of the mild *TNFR* allele. Alternatively, as we observed residual chloride secretion more frequently among patients with mild disease phenotype,¹¹ the *TNFR* association with the basic defect phenotype might reflect an overrepresentation of mild modifier alleles in this patient subgroup, which is equivalent to a replication study with confirmatory outcome.

In conclusion, describing modifiers in cystic fibrosis remains challenging even though the underlying disease is monogenic by definition. In contrast, the manifestation of endophenotypes such as diabetes, the CF basic defect and ultimately disease severity is most likely polygenic by

nature, implying that when looking for CF modifiers, the same obstacles will be encountered as in complex diseases such as asthma or cardiovascular disorders. Based on our experience, we suggest that patient samples in which individuals share as many non-inherited characteristics as possible and genetic markers which are selected for a high information content to overcome the loss of power in small-sized study cohorts will facilitate mapping of clinically meaningful CF modifiers in the future.

SUPPLEMENTARY INFORMATION

Part of the candidate genes analysed in this study were extracted from transcriptome of intestinal epithelium of F508del-CFTR homozygotes and non-CF controls (van Barneveld et al., personal communication; MS in preparation co-authored by FS and BT). The corresponding transcriptome data have been deposited in the GEO database under accession no. GSE15568. Detailed information on the definition of patient cohorts, the genotyping of markers and the statistical evaluation is provided within the supplement of this manuscript.

ACKNOWLEDGEMENTS

We are indebted to all participating patients and their families, physicians, CF centers and staff for their time, co-operation and assistance. The authors also thank Prof T. Epplen (University of Münster) for his contribution of the genomic sequence for two microsatellites used to interrogate the HLA class I and III gene loci.

V.K. was a member of the International Research Training Group *Pseudomonas*: Pathogenicity and Biotechnology (IRTG653 of the Deutsche Forschungsgemeinschaft). The support of the Fritz-Thyssen-Stiftung (to F.S. and T.B.) for analysis of the SCNN1B/SCNN1G-locus is gratefully acknowledged. M.M.Jr was supported by VZFM00064203 and NE9488/3. This study was supported by a grant to B.T. from the Deutsche Forschungsgemeinschaft (SFB 621, project C7). This work was executed as part of the European Cystic Fibrosis Twin and Sibling Study and financially supported by the European Union (QL G1-CF- 2001-01005), the Deutsche Fördergesellschaft für die Mukoviszidoseforschung eV and the Mukoviszidose eV.

Role of funding organizations

The funding organizations had no influence on the study design, data analysis and interpretation or intention to publish.

Author contributions

Genotyping data which including the responsibility for the accuracy of primary data and its quality control, was provided to the study by FS, VK, SH, CB, HC, ST, JY, UL, BS, LF, DR, MMJr. Analysis and interpretation of genetic data was done by FS and TB. Analysis of clinical data was carried out by FS, MB and BT. Consultants to the European CF twin and sibling study for statistics are JG and TFW. The concept of the study was designed by BT, JJC and TFW. All authors have revised the manuscript critically for content and have approved the final version.

The Corresponding Author (FS) has the right to grant on behalf of all authors and does grant on behalf of all authors, an exclusive licence on a worldwide basis to the BMJ Publishing Group Ltd and its Licensees to permit this article (if accepted) to be published in Journal of Medical Genetics and any other BMJPGL products to exploit all subsidiary rights.

Conflict of interest statement

The authors have no conflicting financial interests.

REFERENCES

1. Strausbaugh SD, Davis PB. Cystic fibrosis: a review of epidemiology and pathobiology. *Clin Chest Med* 2007;28:279-288.
2. Koch C, Cuppens H, Rainisio M, Madessani U, Harms H, Hodson M, Mastella G, Navarro J, Strandvik B, McKenzie S, Investigators of the ERCF. European Epidemiologic Registry of Cystic Fibrosis (ERCF): comparison of major disease manifestations between patients with different classes of mutations. *Pediatr Pulmonol* 2001;31:1-12.
3. Mekus F, Ballmann M, Bronsveld I, Bijman J, Veeze H, Tümmler B. Categories of deltaF508 homozygous cystic fibrosis twin and sibling pairs with distinct phenotypic characteristics. *Twin Res* 2000;3:277-293.
4. Vanscoy LL, Blackman SM, Collaco JM, Bowers A, Lai T, Naughton K, Algire M, McWilliams R, Beck S, Hoover-Fong J, Hamosh A, Cutler D, Cutting GR. Heritability of lung disease severity in cystic fibrosis. *Am J Respir Crit Care Med* 2007;175:1036-1043.
5. Davis PB. Cystic fibrosis since 1938. *Am J Respir Crit Care Med* 2006;173:475-482.
6. Wolfenden LL, Schechter MS. Genetic and non-genetic determinants of outcomes in cystic fibrosis. *Paediatr Respir Rev* 2009;10:32-36.
7. Quinton HB, O'Connor GT and the Northern New England CV Cystic Fibrosis Consortium. Current Issues in Quality Improvement in Cystic Fibrosis. *Clin Chest Med* 2007;28:459-72.
8. De Jonge HR, Ballmann M, Veeze H, Bronsveld I, Stanke F, Tümmler B, Sinaasappel M. Ex vivo CF diagnosis by intestinal current measurements (ICM) in small aperture, circulating Ussing chambers. *J Cyst Fibros* 2004;3:159-63.
9. Schüler D, Sermet-Gaudelus I, Wilschanski M, Ballmann M, Dechaux M, Edelman A, Hug M, Leal T, Lebacq J, Lebecque P, Lenoir G, Stanke F, Wallemacq P, Tümmler B, Knowles

- MR. Basic protocol for transepithelial nasal potential difference measurements. *J Cyst Fibros* 2004;3:151-155.
10. Bronsveld I, Mekus F, Bijman J, Ballmann M, Greipel J, Hundrieser J, Halley DJ, Laabs U, Busche R, De Jonge HR, Tümmler B, Veeze HJ. Residual chloride secretion in intestinal tissue of deltaF508 homozygous twins and siblings with cystic fibrosis. The European CF Twin and Sibling Study Consortium. *Gastroenterology* 2000;119:32-40.
 11. Bronsveld I, Mekus F, Bijman J, Ballmann M, de Jonge HR, Laabs U, Halley DJ, Ellemunter H, Mastella G, Thomas S, Veeze HJ, Tümmler B. Chloride conductance and genetic background modulate the cystic fibrosis phenotype of Delta F508 homozygous twins and siblings. *J Clin Invest* 2001;108:1705-1715.
 12. Drumm ML, Konstan MW, Schluchter MD, Handler A, Pace R, Zou F, Zariwala M, Fargo D, Xu A, Dunn JM, Darrah RJ, Dorfman R, Sandford AJ, Corey M, Zielenski J, Durie P, Goddard K, Yankaskas JR, Wright FA, Knowles MR, Gene Modifier Study Group. Genetic modifiers of lung disease in cystic fibrosis. *N Engl J Med* 2005;353:1443-1453.
 13. Bremer LA, Blackman SM, Vanscoy LL, McDougal KE, Bowers A, Naughton KM, Cutler DJ, Cutting GR. Interaction between a novel TGFB1 haplotype and CFTR genotype is associated with improved lung function in cystic fibrosis. *Hum Mol Genet* 2008;17:2228-2237.
 14. Martin N, Boomsma D, Machin G. A twin-pronged attack on complex traits. *Nat Genet* 1997;17:387-392.
 15. Ritzka M, Stanke F, Jansen S, Gruber AD, Pusch L, Woelfl S, Veeze HJ, Halley DJ, Tümmler B. The CLCA gene locus as a modulator of the gastrointestinal basic defect in cystic fibrosis. *Hum Genet* 2004;115:483-491.

16. Mekus F, Laabs U, Veeze H, Tümmler B. Genes in the vicinity of CFTR modulate the cystic fibrosis phenotype in highly concordant or discordant F508del homozygous sib pairs. *Hum Genet* 2003;112:1-11.
17. Kumar V, Becker T, Jansen S, van Barneveld A, Boztug K, Wölfl S, Tümmler B, Stanke F. Expression levels of FAS are regulated through an evolutionary conserved element in intron 2, which modulates cystic fibrosis disease severity. *Genes Immun* 2008;9:689-696.
18. Stanke F, Becker T, Cuppens H, Kumar V, Cassiman JJ, Jansen S, Radojkovic D, Siebert B, Yarden J, Ussery DW, Wienker TF, Tümmler B. The TNFalpha receptor TNFRSF1A and genes encoding the amiloride-sensitive sodium channel ENaC as modulators in cystic fibrosis. *Hum Genet* 2006;119:331-343.
19. Stanke F, Tümmler B, Becker T. Genetic modifiers in cystic fibrosis. *N Engl J Med* 2006;354:88-90.
20. Becker T, Jansen S, Tamm S, Wienker TF, Tümmler B, Stanke F. Transmission ratio distortion and maternal effects confound the analysis of modulators of cystic fibrosis disease severity on 19q13. *Eur J Hum Genet* 2007;15:774-778.
21. Stanke F, Davenport C, Hedtfeld S, Tümmler B. Differential decay of parent-of-origin-specific genomic sharing in cystic fibrosis-affected sib pairs maps a paternally imprinted locus to 7q34. *Eur J Hum Genet* 2010 [Epub ahead of print] PMID: 20051989
22. Stanke F, Becker T, Hedtfeld S, Tamm S, Wienker TF, Tümmler B. Hierarchical fine mapping of the cystic fibrosis modifier locus on 19q13 identifies an association with two elements near the genes CEACAM3 and CEACAM6. *Hum Genet* 2010;127:383-394. Epub 2010 Jan 3.
23. Zielenski J, Corey M, Rozmahel R, Markiewicz D, Aznarez I, Casals T, Larriba S, Mercier B, Cutting GR, Krebsova A, Macek M Jr., Langfelder-Schwind E, Marshall BC, DeCeglie-Germana J, Claustres M, Palacio A, Bal J, Nowakowska A, Ferec C, Estivill X, Durie P,

- Tsui LC. Detection of a cystic fibrosis modifier locus for meconium ileus on human chromosome 19q13. *Nat Genet* 1999;22:128-129.
24. Cutting GR. Modifier genetics: cystic fibrosis. *Annu Rev Genomics Hum Genet* 2005;6:237-260.
 25. Collaco JM, Cutting GR. Update on gene modifiers in cystic fibrosis. *Curr Opin Pulm Med* 2008;14:559-566.
 26. Hillian AD, Londono D, Dunn JM, Goddard KA, Pace RG, Knowles MR, Drumm ML, CF Gene Modifier Study Group. Modulation of cystic fibrosis lung disease by variants in interleukin-8. *Genes Immun* 2008;9:501-508.
 27. Corey M, Farewell V. Determinants of mortality from cystic fibrosis in Canada, 1970–1989. *Am J Epidemiol* 1996;143:1007-1017.
 28. Sharma R, Florea VG, Bolger AP, Doehner W, Florea ND, Coats AJ, Hodson ME, Anker SD, Henein MY. Wasting as an independent predictor of mortality in patients with cystic fibrosis. *Thorax* 2001;56:746-750.
 29. International HapMap Consortium. A second generation human haplotype map of over 3.1 million SNPs. *Nature* 2007;449:851-861.
 30. Neale MC, Cardon LR eds 1992: *Methodology for genetic studies of twins and families*. Dordrecht: Kluwer Academic Publishers BV
 31. Herold C, Becker T. Genetic association analysis with FAMHAP: a major program update. *Bioinformatics* 2009;25:134-136.
 32. Knapp M, Becker T. Family-based association analysis with tightly linked markers. *Hum Hered* 2003;56:2-9.
 33. Becker T, Knapp M. A powerful strategy to account for multiple testing in the context of haplotype analysis. *Am J Hum Genet* 2004;75:561-570.

34. Spielman RS, McGinnis RE, Ewens WJ. Transmission test for linkage disequilibrium: the insulin gene region and insulin-dependent diabetes mellitus (IDDM). *Am J Hum Genet* 1993;52:506-516.
35. Rowntree R, Harris A. DNA polymorphisms in potential regulatory elements of the CFTR gene alter transcription factor binding. *Hum Genet* 2002;111:66-74.
36. Sereth H, Shoshani T, Bashan N, Kerem BS. Extended haplotype analysis of cystic fibrosis mutations and its implications for the selective advantage hypothesis. *Hum Genet* 1993;92:289-295.
37. Morral N, Dörk T, Llevadot R, Dziadek V, Mercier B, Férec C, Costes B, Girodon E, Zielenski J, Tsui LC, Tümmler B, Estivill X. Haplotype analysis of 94 cystic fibrosis mutations with seven polymorphic CFTR DNA markers. *Hum Mutat* 1996;8:149-159.
38. Frank J, Roux J, Kawakatsu H, Su G, Dagenais A, Berthiaume Y, Howard M, Canessa CM, Fang X, Sheppard D, Matthay MA, Pittet JF. Transforming growth factor-beta1 decreases expression of the epithelial sodium channel alphaENaC and alveolar epithelial vectorial sodium and fluid transport via an ERK1/2-dependent mechanism. *J Biol Chem* 2003;278:43939-43950.
39. Galietta LJ, Folli C, Caci E, Pedemonte N, Taddei A, Ravazzolo R, Zegarra-Moran O. Effect of inflammatory stimuli on airway ion transport. *Proc Am Thorac Soc* 2004;1:62-65.
40. Roux J, Kawakatsu H, Gartland B, Pespenti M, Sheppard D, Matthay MA, Canessa CM, Pittet JF. Interleukin-1beta decreases expression of the epithelial sodium channel alpha-subunit in alveolar epithelial cells via a p38 MAPK-dependent signaling pathway. *J Biol Chem* 2005;280:18579-18589.
41. Eisenhut M. Reduction of alveolar epithelial ion and fluid transport by inflammatory mediators. *Am J Respir Cell Mol Biol* 2007;36:388-389.

Figure 1: Assembly of patient subsets for genetic modeling and association study

Weight, height, and lung function data of at least one out of two siblings was received in 1995-1996 for 540 cystic fibrosis sib pairs from 158 CF clinics in 14 European countries. Pairs for whom clinical data was obtained for both sibs of a pair were included into genetic modeling (466 pairs with weight and height, 318 pairs with lung function).³ Subsamples of F508del homozygotes were selected for the association study based on the clinical phenotype or the basic defect phenotype which was measured by nasal potential difference measurement (NPD) and intestinal current measurement (ICM).^{10,11} In total, 171 patients from 101 families were genotyped. Two patients with CFTR mutation genotypes other than F508del/F508del were only included for the phenotype “DIDS-sensitive chloride conductance” determined by ICM. The association study compared subsets of affected patient pairs in their manifestation of the clinical phenotype, and correspondingly only dizygous pairs were included. In case of case-control-comparisons of the basic defect in subsamples of unrelated index cases one twin per monozygous twin pair was added to the cohort. For the association study on ICM phenotypes, 16 families with F508del homozygous offspring recruited from the CF clinic in Hannover, Germany were included.

Integers within this figure correspond to the number of independent families while all numbers in brackets refer to the number of individual patients. Relative proportions of samples and subsamples stratified for extreme phenotypes are displayed on the basis of the number of independent nuclear families, i.e. of sib pairs instead of individual patients, in order to represent sample sizes based on their number of independent genetic contributions which define the effective number of chromosomes considered for the genetic association study. Integers on top of diagrams report the sample size as sum of all subgroups. Integers adjacent to diagrams report sizes of subgroups which are visualized as slices. For subgroups consisting of 540, 466 and 318 pairs, the entire circle diameter corresponds to the sample size. For subgroups consisting of 46 and 55 families, the sample size is represented in proportion to the larger data sets by the circle diameter carrying the sample size label (white central circle with red rim on top of the diagram).

Composition of subgroups stratified for manifestation of the basic defect is displayed on the basis of individual patients. The numbers displayed for the contrasting phenotypes analyzed in the case-control scenarios show the effective sample size, taking into account that sib pairs who show a discordant basic defect phenotype can contribute to one out of two (NPD phenotypes) or one out of three (ICM phenotypes) subgroups only in the association study. In other words, subsets of unrelated patients were defined by an index-case strategy whereby extreme basic defect phenotypes were selected for in pairs for whom discordance ICM or NPD was observed. As NPD measurements were obtained for two siblings of a pair in most cases, two subsets for each case-control-test were evaluated to provide an internal control depending on which sibling was assigned at random to be the index case.

Candidate genes were interrogated for their association with disease severity using the phenotypic contrast between concordant/mildly affected patient pairs, concordant/severely affected patient pairs and discordant patient pairs. For this purpose, 37 dizygous pairs – representing the 12% most informative patient pairs of the entire sample – were selected based on a non-parametric ranking algorithm as described previously.³ NPD data were used firstly to inquire for an association with sodium transport via the amiloride-sensitive proportion of the potential, using the upper and lower 30% of the entire sample to define extreme phenotypes, secondly to evaluate an association with residual chloride secretion activated by chloride-free gluconate solution and isoproterenol and thirdly to analyze a contribution of the selected candidate genes on ATP-stimulated chloride response. The basic defect assessed by ICM was evaluated using a set of patients devoid of residual chloride secretion as controls. The pleiotropic chloride-channel inhibitor DIDS, to which CFTR is not sensitive at the chosen concentration, was used to discriminate CFTR-mediated residual chloride secretion from chloride secretion mediated by alternative channels which are DIDS-sensitive.^{10,11}

Figure 2: Relative impact of environmental and inherited factors on CF

A Distribution of intrapair differences in weight for height % (wfh%, left) and the CF-population centiles for FEV1%pred (FEVPerc, right) among sib pairs, dizygous twins (DZT) and monozygous twins (MZT). Overall comparisons were done by Kruskal-Wallis rank test (wfh%: $P = 5 \cdot 10^{-6}$; FEVPerc: $P = 0.067$) and individual subgroups were compared by Dunn rank test. Please note that both monozygous and dizygous twins are similarly concordant in wfh%, indicating that the shared in-utero environment determines outcome in wfh%. In contrast, monozygous twins are more concordant than either dizygous twins or sibs in FEVPerc, implying that lung function is influenced by inherited factors.

B Relative impact of inherited and environmental factors on disease manifestation as estimated by genetic modeling. Covariance matrices of weight as % of predicted weight for height (WFH) and forced expiratory volume in 1 sec as % of predicted value (LF) were analyzed by genetic modeling.³⁰ All patient pairs and the subgroup of F508del-*CFTR* homozygotes were evaluated. Models were based on a linear combination of the four following factors: random environmental effects, shared environmental effects, additive and dominant genetic effects. Goodness-of-fit was judged by χ^2 -measure, whereby in case of two or more models with a similar goodness-of-fit the least complicated model, based on the fewest out of the four linearly combined factors, was accepted as suggested by the Akaike information criterion.³⁰ Consensus diagrams displaying the impact of inherited (black bars) and environmental (white bars) factors are shown.

Figure 3: Association of candidate genes with CF disease severity and basic defect

Candidate genes are listed according to their position in the human genome. Association to genes instead of results of individual markers are displayed if several linked markers were analyzed per gene or gene cluster. For *PRSS8*, *SCNN1B/SCNN1G*, *KRT8/KRT18*, *TNFR/SCNN1A*, *TLR4*, the results of the microsatellite scan and the confirmation by low-resolution SNP typing are shown. Solid lines are used to indicate unlinked genes, adjacent loci within 10 Mb distance are separated by dotted lines and neighboring genes are not separated by lines between gene names. All candidate genes were analyzed for an underrepresentation of alleles among transmitted chromosomes (TDT) as well as by case-reference comparisons for their association with disease severity, ICM or NPD: disease severity cis — case = concordant mildly affected patient pairs, reference = concordant severely affected patient pairs;¹⁶⁻¹⁸ disease severity trans — case = discordant patient pairs, reference = concordant patient pairs;^{16,17} ICM_Res — case = DIDS-insensitive (CFTR-mediated)^{10,11} residual chloride secretion, reference = no residual chloride secretion; ICM_DIDSRes — case-DIDS-sensitive (\neq CFTR)^{10,11} residual chloride secretion, reference = no residual chloride secretion; NPD_Na+ — case = low response to amiloride, reference = high response to amiloride; NPD_Gl/Iso — case = no response to gluconate and isoproterenol, reference = residual chloride secretion response to gluconate and isoproterenol; NPD_ATP — case = no response to ATP, reference = residual chloride secretion response to ATP. Solid dots correspond to raw P-values for single chromosomes (marked H), chromosome pairs (marked D) or accumulation of rare variants (marked r). All displayed values are uncorrected P_{raw} (see also Supplement – Table 8).

Figure 4: Comparison of allele distributions among patients with different basic defect phenotypes

Allele distributions are shown for the *CFTR*-locus (A-C), *SCNNIA*, encoding the α -subunit of the epithelial sodium channel ENaC and the neighboring gene *TNFRSF1A*, encoding the 55 kDa-receptor for TNF α (*TNFR*) (D-E) and the CFTR interaction partner *NHERF2* (F). Pictograms represent the relative location of targeted candidate gene (black box) and genotyped markers (vertical lines). Markers are: XV-2c, the variant HUG16RS³⁵ and J3.11 (A,B); XV-2c, KM19 5' and IVS17bTA³⁶ (C); four markers in *SCNNIA* (nt7AG, SC3, SC4 and rs2228576; D); four markers in *TNFR* (rs767455, D12S889, rs1800692 and rs1800693; E); two microsatellites near *NHERF2* (F). Alleles at IVS17bTA are calibrated in accordance with Morral et al.³⁷ For other microsatellites, arbitrary repeat units were assigned with 10 corresponding to the most frequently observed allele. SNPs typed by PCR-RFLP are named according to the presence (allele 2) or absence (allele 1) of the diagnostic restriction site. P-values and corresponding allele distributions (A-B; D-F) or diplotypes distributions (C) are given for 6 case-control-tests. P_{best} is the best observed p-value while P_{global} was corrected for multiple testing³³ of all markers at the locus. Phenotypes are: NPD_Amil: NPD change to amiloride [mV]; NPD_ATP: NPD change to ATP; NPD_GI: NPD change to gluconate and isoproterenol [mV]; ICM no Res.: no residual response; ICM Res.: cAMP sensitive and DIDS-insensitive residual chloride secretion. The four-marker haplotype at *TNFR* correlates with CF disease severity¹⁸ whereby allele 1-10-2-2 is associated with a mild CF phenotype (E) and is also overrepresented among patients exhibiting residual chloride secretion assessed by NPD (D).

Supplementary Information

Genes that determine immunology and inflammation modify the basic defect of impaired ion conductance in cystic fibrosis epithelia

Running Title: Host defense genes modify CF basic defect

Frauke Stanke,¹ Tim Becker,² Vinod Kumar,¹ Silke Hedtfeld,¹ Christian Becker,¹ Harry Cuppens,³ Stephanie Tamm,¹ Jennifer Yarden,³ Ulrike Laabs,¹ Benny Siebert,¹ Luis Fernandez,¹ Milan Macek Jr.,⁴ Dragica Radojkovic,^{3,5} Manfred Ballmann,¹ Joachim Greipel,⁵ Jean-Jacques Cassiman,³ Thomas F. Wienker² & Burkhard Tümmler¹

¹ *Department of Pediatrics, Hannover Medical School, Hannover, Germany*

² *Institute of Medical Biometry, Informatics and Epidemiology, University of Bonn, Bonn, Germany*

³ *Department of Human Genetics, University of Leuven, Leuven, Belgium*

⁴ *Department of Biology and Medical Genetics, Charles University Prague, Medical School and University Hospital Motol, Prague, Czech Republic*

⁵ *Institute of Molecular Genetics and Genetic Engineering, Belgrade, Serbia*

⁶ *Department of Biophysical Chemistry, Hannover Medical School, Hannover, Germany*

This Supplement provides details on the patient cohorts, the genetic markers tested and the statistical analysis reported upon in this study. We have characterized informative patient cohorts (Section A of this Supplement) and have conducted an association study on these patient subsamples. Extreme phenotypes were selected as these are generally considered to be more informative in analyzing quantitative traits.^{S1-S3} We have summarized at 9 to 29 independent nuclear families as subsamples (Figure 1 of the main manuscript) whereby the sample size — for instance, for a rare phenotype such as “presence of CFTR-mediated residual chloride secretion in F508del-*CFTR* homozygotes” — was limited by the incidence of CF in the central European population and by the etiology and pathophysiology of CF disease that determines the frequencies of the endophenotype in the CF population. Consequently, appropriate strategies were undertaken when selecting genetic markers for the study (see Section B of this Supplement) and applying statistical tests to the genetic data obtained (see Section C of this Supplement) in order to maintain the study’s power to detect a modifier.

Within this entire Supplement, numbers of nuclear families rather than individual patients are considered to illustrate the effective sample size in an association study that is determined by the number of independent chromosomes. Monozygous twin pairs are, of course, accounted for by one twin per pair only.

A Patients and Phenotypes

The association study reported upon in this manuscript was done on 101 CF families with a total of 171 CF patients. As detailed below in sections A1-A5, these patients were classified based on either their clinical phenotype or their basic defect phenotype defining overlapping subsamples. Data on the clinical phenotype and/or the basic defect phenotype for the 101 families enrolled into the study was as follows: Clinical data, but no information on the manifestation of the basic defect, was known for F508del homozygous twin and sib pairs of 37 families. Both, clinical phenotype and

basic defect was characterized for 46 F508del homozygous twin and sibling pairs. 16 F508del-CFTR homozygous singletons with known basic defect were included in the study. These 16 patients were recruited for a study on the manifestation of the basic defect in excised intestinal biopsies and subsequent chip-based transcriptome analysis (van Barneveld et al., unpublished). For the analysis of non-CFTR mediated, alternative residual chloride conductance, two non-F508del homozygous patient pairs were enrolled. The overlap between the subset analyzed in the association study for extreme clinical phenotype (37 dizygous F508del-CFTR homozygous patient pairs, see section A2) and the patient subsets of which NPD and/or ICM (see section A4) is shown within Supplement - Table 1.

Supplement — Table 1: Overlap between patient subsamples

	Total no. of nuclear families	Part of subsample of 37 dizygous F508del homozygous characterized for extreme clinical phenotype?		Part of subsample analyzed for association with NPD and / or ICM phenotypes?	
		Yes	No	Yes	No
F508del homozygous siblings with known basic defect ^A and known clinical phenotype	46	20	26	46	0
F508del homozygous siblings with known clinical phenotype	37	17	20	0	37
non-F508del homozygous siblings, basic defect known ^B	2	0	2	2	0
unrelated F508del homozygotes, basic defect known	16	0	16	16	0
All	101	37	64	46	55

^A NPD and ICM was measured for patients of 39 families; only NPD data was available for 7 families

^B included only for the analysis of non-CFTR mediated residual chloride secretion

A1 Clinical Phenotype

The *European CF twin and sibling* study was initialized in 1995 as a project of the European Community Concerted Action for the Coordination of Cystic Fibrosis Research and Therapy under the auspices of the European Cystic Fibrosis Society. Twin and sibling pairs with cystic fibrosis were recruited from 158 CF clinics in 14 European countries. Basic clinical data such as actual weight, height and forced expiratory volume in one second (FEV1) was inquired from the treating physician using a one-page questionnaire. Weight and height data was converted to weight as % of predicted weight for height (wfh%) using the centiles by Prader et al.^{S4} FEV1 was expressed as %

of predicted values according to the reference data set of Knudson et al.^{S5} To account for the age-dependent decline in FEV1%pred which is characteristic for the progressive lung disease CF, FEV1%pred was transformed into its CF population centiles FEVPerc^{S6} based on the data accumulated by the European CF registry. Using wfh% and FEVPerc, the clinical phenotype within the European twin and sib pair population was described based on two parameters with age-independent centiles^{S6} which is a prerequisite to quantify disease severity and intrapair discordance in a crosssectionally recruited CF patient population.

Both, wfh% and FEVPerc were combined in a composite parameter using a ranking algorithm to describe the severity of the disease with one parameter only.^{S6} In order to validate that this composite parameter is suitable to detect the influence of inherited factors on CF disease, dizygous and monozygous twin pairs were compared in their intrapair differences in FEVPerc, wfh% and the composite parameter derived thereof. Intrapair discordance was significantly lower in 29 monozygous twin pairs compared to 12 dizygous twin pairs as long as the composite parameter was considered ($P = 0.04$, Mann-Whitney rank test; previously reported in Mekus et al. 2000^{S6}) while no association of intrapair concordance and twin zygosity was seen when only wfh% or only FEVPerc were examined. Interpreting the higher concordance of monozygous twin pairs as an indication of inherited factors that determine the phenotype, we concluded from this finding that the composite parameter was more sensitive to detect inherited factors than either of the individual clinical parameters wfh% and FEVPerc. Consequently, we relied on the composite parameter to select patient pairs with extreme clinical phenotypes for the association study.

In summary, the clinical phenotype of the CF twin and sibling pairs was described using a composite parameter derived from a linear combination of wfh% and FEVPerc to account for the severity of CF disease in the two major afflicted organ systems, the gastrointestinal and pulmonary tract. By comparing monozygous and dizygous twin pairs, we have observed that the composite parameter was more sensitive towards the influence of inherited factors in CF disease than each of the individual parameters describing nutritional and pulmonary status. As a consequence, we

assume that the selection of patient pairs for the association study based on this composite parameter has increased the power of the study to detect inherited factors in comparison to any recruitment based on a single clinical parameter.

A2 Patient pairs with extreme clinical phenotypes

Only F508del-*CFTR* homozygous sib pairs were enrolled to study the association of candidate genes with the clinical phenotype. Besides the obvious rationale to include only F508del homozygotes – as these subgroup is homogeneous with respect to their *CFTR* mutation genotype which is a major determinant of CF disease severity^{S7} – we have carried out an association study on sib pairs as they share many environmental factors such as the treating physician, family environment and the socioeconomic status all of which are known to influence outcome in CF disease.^{S8-S10} Three different patient pair categories were defined based on the clinical disease severity of the individual siblings: concordant severely affected pairs (CON-; two sibs with similar and severe phenotype), concordant mildly affected pairs (CON+; two sibs with similar and mild phenotype) and discordant pairs (DIS; two sibs with dissimilar phenotype). Based on our ranking procedure,^{S6} we have identified 19 pairs, 20 pairs and 18 pairs with the phenotypes CON-, CON+ and DIS, respectively. Candidate genes were tested for association with the clinical phenotype using sib pairs, thus three CON- and three CON+ monozygous twins were excluded from the association study. Of the remaining 16 CON-, 17 CON+ and 18 DIS sibling pairs, DNA of a reliable quality suitable for multiwell-PCR-based genotyping was available for the entire duration of the study until 2007 from 11 CON-, 12 CON+ and 14 DIS patient pairs. Parents were recruited along with the patients whereby both parents were enrolled for 8 CON-, 10 CON+ and 10 DIS families and genomic DNA of at least one parent was recruited for one further CON-, two further CON+ and one further DIS family (previously reported in Stanke et al. 2006^{S11}).

In summary, out of 318 patient pairs for whom wfh% and FEVPer was available, we have genotyped 37 sibling pairs and their parents who were recruited because of their homogeneous *CFTR* mutation genotype and their extreme clinical phenotype. This highly selected sample

corresponds to the most informative 12% of the central European patient pair population. The phenotypic contrast between these pairs was employed to interrogate candidate genes for an association with CF disease severity.

A3 Basic defect phenotype

The CF basic defect is defined as impaired or absent CFTR-mediated chloride secretion in CFTR-expressing tissues such as the sweat gland and epithelia of the respiratory and gastrointestinal tracts. In the study reported upon in this manuscript, we have used nasal potential difference measurements (NPD) and intestinal current measurements (ICM) to characterize the patients' CF basic defect in respiratory and intestinal tissue. The patients were recruited from the European CF twin and study cohort. NPD and ICM measurements on twins and sibs have been performed between 1997 and 1999 at selected CF core centers in Hannover, Innsbruck, London, Rotterdam, and Verona. All ICM and NPD results have been previously described by Bronsveld et al.^{S12,S13} Briefly, secretagogues that activate or block ion channels, ion exchangers or components of the cellular signal transduction pathways were applied by superfusion of the lower nasal turbinate^{S14} or to excised rectal suction biopsies mounted in a micro-Ussing chamber.^{S15} Mainly, we have used both techniques to discriminate between patients with and without residual chloride secretion. As ICM is an ex-vivo method applied to patient's biopsies, the toxic compound DIDS (4,4'-diisothiocyanostilbene-2,2'-disulfonic acid) which has been reported to block chloride channels other than CFTR, could be used to differentiate between CFTR-mediated residual chloride secretion and chloride secretion through alternative channels.^{S12} Based on NPD measurements, contrasting phenotypes for the response to amiloride, to chloride-free gluconate solution and isoproterenol, and to ATP were assigned to F508del homozygotes whereby amiloride blocks the epithelial sodium channel ENaC and thus is indicative of ENaC function, a response to chloride-free gluconate solution and isoproterenol is interpreted as CFTR-mediated residual function and a response to ATP taken as an indication of the activity of other ion channels.^{S13}

A4 Patients with extreme basic defect phenotypes

For the present study, families with CF twins or sibs who have participated in the basic defect study were genotyped at the candidate gene loci. While NPD data was obtained for 46 families of F508del homozygous siblings for whom DNA of a reliable quality suitable for multiwell-PCR-based genotyping was available for the entire duration of the study until 2007, ICM results were available for a subset of 39 families only. 16 unrelated F508del homozygotes for whom the ICM phenotype was known were included into the study to allow the investigation of rare ICM phenotypes such as residual chloride secretion. Inclusion criteria were extended further for the analysis of non-CFTR mediated chloride conductance as two non-F508del homozygous sib pairs were included to investigate modifiers of this CF endophenotype.

In contrast to the study on extreme clinical phenotypes for which subsets of patient pairs were compared in an association study, an index-case strategy had to be employed for the association study on basic defect phenotypes as the sample is too small to define a reasonable number of extreme concordant for each of the NPD and ICM phenotypes and discordant pairs for each of the phenotype combinations. Patients whose ICM phenotype showed a combination of DIDS-sensitive and DIDS-insensitive residual chloride secretion (6 individuals) were defined as an intermediate phenotype. The following hierarchical rules were applied to the entire data set in order to define which sibling of a pair was included as the index case of the family:

1. for all pairs in whom the phenotype was obtained for one sibling only, this sibling was enrolled irrespective of his/her phenotype
2. for all pairs in whom a discordant phenotype was observed, the sibling with the most extreme phenotype was defined as an index case by
 - a. NPD, amiloride response: if one of the sibs displays a response of $>31\text{mV}$ or $<21\text{mV}$, this sib was enrolled as index case with an extreme phenotype

- b. NPD, response to chloride-free gluconate solution and isoproterenol: if one of the sibs displays a response of $>5\text{mV}$, this sib was included as the index case with an extreme phenotype
 - c. ICM: if one sibling displays DIDS-insensitive residual chloride secretion, this sibling is enrolled as index case
 - d. ICM: if one sibling displays DIDS-sensitive residual chloride secretion, this patient was included as index case
3. For all pairs in whom a concordant phenotype was observed, one sibling was selected as index case by randomly alternating between sibling A and sibling B of a pair. For the NPD measurements, more than two concordant sib pairs were observed in most subsamples and hence the randomization was done a second time by selecting the opposite sib of all concordant pairs, providing two non-overlapping randomized subsets which were analyzed in parallel.

Applying these rules, we have enrolled patients with no residual chloride secretion (absence of chloride secretion upon stimulation with carbachol and histamine), chloride secretion mediated by alternative ion channels (presence of chloride secretion upon stimulation with carbachol but absence of chloride secretion upon stimulation with histamine after inhibition with DIDS) or chloride secretion mediated by CFTR (presence of chloride secretion upon stimulation with carbachol and presence of chloride secretion upon stimulation with histamine after inhibition with DIDS).^{S12,S13} Individuals from 17, 13 and 9 families could be assigned to these three respective ICM-based phenotypes using stringent criteria. As a result of the index-case selection strategy, the individuals of these families were not related to each other. Similarly, using stringent criteria for definition, an amiloride response exceeding the 70th centile ($> 31 \text{ mV}$) in the studied population of 86 individuals, was assigned to siblings from 16 families. An amiloride response below the 30th centile ($< 22 \text{ mV}$) was observed for siblings of 15 families. A response of more than 5 mV chloride-free gluconate solution and isoproterenol was noted for siblings in 12 families. No response to

chloride-free gluconate solution and isoproterenol was seen for siblings of 21 families. Finally, in 29 families, a response upon superfusion with ATP was observed while in sibpairs of 16 families, no reaction to ATP was detected.

A5 Ascertainment of zygosity and CFTR mutation genotype

As reported previously,^{S6,S16} zygosity of same-sex twin pairs was ascertained using the AmpFLSTR Profiler PlusTM typing kit on an IBI Prism 377 (Perkin Elmer Applied Biosystems, Foster City, CA, USA) or by oligonucleotide fingerprinting of simple repeats applying in situ gel hybridization of *MboI* or *HinfI* genomic digests. 8 pairs for whom DNA was not available and zygosity uncertain (see Figure 1) were excluded from further analysis. The *CFTR* mutation genotype was verified for all F508del-*CFTR* homozygous patient pairs and their parents enrolled in the association study by heteroduplex analysis. Non-paternity and consanguinity were excluded based on genotyping of 16 polymorphic markers located on five autosomes other than chromosome 7.

B1 Selection of candidate genes

Candidate genes were selected from reports of CF modifiers by others,^{S17-S20} based on our own previous findings,^{S11,S16,S21-S23} on a differential expression comparing F508del-*CFTR* and non-*CF* epithelial cells (van Barneveld et al, unpublished) and on an annotated function of the gene product that implies its role in CF disease pathophysiology, *CFTR* processing, *CFTR* trafficking or host defense.

The following genes were analyzed because they have been suggested or published as CF modifiers: the glutathione-S-transferases *GSTM1*,^{S24-S27} *GSTT1*^{S27} and *GSTP1*,^{S27-S29} the beta-2-adrenergic receptor *ADRB2*,^{S30-S32} the cytokines *IL8*,^{S19,S33} *IL10*,^{S33,S34} *TNF α* ^{S25,S26,S33,S36} and *TGFB1*,^{S17,S33,S37} the components of the innate immune defense mannan-binding lectin (*MBL*),^{S35-S44} the receptor for bacterial LPS *CD14*,^{S45} the alpha-1-antitrypsin gene (*AAT*),^{S46-S50} the surfactant proteins *SPA*^{S51} and *SPD* and the cystic fibrosis modifier 1 locus *CFM1*.^{S18}

We have previously identified CF modifiers at 1p31,^{S21} CFTR,^{S16} the CFTR linkage region encompassing the leptin gene,^{S16} and a genomic area on 19q13.^{S23} We also described the FAS receptor CD95,^{S22} the tumor necrosis factor receptor 1 (TNFR)^{S11} and the beta and gamma subunits of the amiloride-sensitive sodium channel ENaC (SCNN1B and SCNN1G)^{S11} as modulating genes in CF. Findings on these genes and loci are reported again firstly, as our previous analyses were restricted to only the basic defect phenotype (1p31)^{S21} or to only the disease severity phenotypes^{S11,S16,S22,S23} while this manuscript presents the comprehensive analysis of both, the clinical and the basic defect phenotypes for all genes. Secondly, this presentation allows us to jointly correct for multiple testing of all loci investigated on European CF twins and sibs so far. In the context of this manuscript, data presentation for loci 1p31 and 19q13 is restricted to the markers used for the initial scan of these genomic regions^{S11,S16,S21-S23}. Fine mapping of the modifier variants in these genomic areas was done using a multimarker-SNP map and subsequent resequencing of contrasting haplotypes, both of which is in its detail and extent beyond the scope of this work and will be communicated separately to the scientific community.

We have furthermore included the following genes among our candidates because of their reported role in ion transport, host defense or their interaction with CFTR: Two bestrophins were selected as alternative chloride channels (BEST4, BEST2).^{S52} The PDZ-binding proteins NHERF1, NHERF2, PDZK1 (alias NHERF3), and GOPC (alias CAL) were included as putative interaction partners with the PDZ-binding domain of CFTR.^{S53} Caveolin 2 (CAV2)^{S54} and paraoxonase 2 (PON2)^{S55} are located upstream of CFTR and are positional candidates. The QTL detected at D2S1788^{S56} and adiponectin (ADIPOQ)^{S57} were included due to their role in the regulation of body weight, which can be considered as a follow-up analysis to our previously reported association of the weight regulating hormone leptin with CF disease severity.^{S16} The toll-like receptors^{S58-S60} TLR2, TLR4, TLR5 and TLR9 were selected to inquire whether these receptors for bacterial and viral compounds modify host defense and consequently, CF disease severity. CXCR2^{S61} was included as it is a receptor for IL8, a proinflammatory cytokine reported to be upregulated in CF.

PRKAA1^{S62}, PRKAA2^{S62} and the syntaxin STX8^{S63} were tested as CFTR interaction partners. The gene SMPD1 encodes for the acetic sphingomyelinase which catalyses the conversion of sphingosine-1-phosphate to ceramide, a metabolite that has been reported to mediate infection susceptibility in CF.^{S64}

Finally, the following genes were selected as candidates based on transcriptome data (van Barneveld et al, unpublished). Briefly, RNA was isolated from rectal suction biopsies of CF patients or healthy volunteers and the transcriptome profile was compared using Affymetrix U133A chips. From the larger set of differentially expressed genes only those candidates which could be reasonably associated with ion transport, CFTR trafficking or host defense were enrolled into the association study. As candidates for ion transport modification, we have analyzed the sulfate transporter SLC26A2^{S65} and the serine protease PRSS8 which has been reported to regulate the activity of the sodium channel ENaC.^{S66} We targeted proteins that might facilitate the interaction of CFTR with the cytoskeleton and thus modify CFTR trafficking to the apical membrane, i.e. the calcium binding proteins S100A11 and S100P,^{S67} and the syntaxin STX12,^{S68} as well as two receptors of the membrane of the endoplasmic reticulum (KDEL2 and KDEL3)^{S69} and the raft-associated protein stomatin (STOM),^{S70} hypothesizing that these might interfere with the turnover kinetics of misfolded F508del-CFTR. Cytokeratin 8 (KRT8) and its neighboring gene KRT18 were included because an interaction of F508del-CFTR and KRT8, influencing the amount of apically located F508del-CFTR, was shown in cell culture models.^{S71} The phospholipase PLA2G2A^{S72} was analyzed due to its role in the pathway that generates inflammatory mediators from lipid components of cellular membranes. The interleukin IL1B,^{S33,S73} the cytokine receptor IFNGR1^{S74} and the signal transduction molecule STAT3^{S75} were considered as plausible candidates to modify host defense. We have included the mucin gene MUC2^{S76} as viscous secretions from submucosal glands are a hallmark of CF disease.

B2 Markers and genotyping

The marker set used for genotyping is heterogeneous with respect to the genotyping methods employed and the density of markers typed per candidate gene. As outlined in Supplement-Table 2, we have typed 182 markers, composed of 101 single nucleotide polymorphisms (SNPs), 79 microsatellite markers and two large genomic deletions. 73 of these markers (40 microsatellites and 33 SNPs) were developed de novo as many candidate genes could not be targeted based on existing genotyping protocols.

40 microsatellites were generated to target candidate genes. For marker development, the genomic sequence was retrieved from the NCBI database selecting 200,000 bp upstream and downstream of the candidate gene. The primary sequence was analyzed with the program TandemRepeatFinder^{S77} to reveal repetitive sequences. Short repetitive elements composed of two to 35 nucleotides were included in an amplicon of 100 bp up to 1200 bp in length for polymerase-chain reaction amplification using one biotinylated primer and products were visualized by direct blotting electrophoresis on a high-resolution polyacrylamide gel and chemoluminescence detection of biotinylated PCR products as described elsewhere.^{S11,S16} Utilization of the PCR amplicon for routine genotyping purposes was judged based on a set of five unrelated control samples. Markers were excluded from analysis if no length variation was observed from sample-to-sample, implying that the marker is not informative, or if multiple non-specific products obscured the interpretation of the genotype pattern. Priority was given to sequences with high numbers of motif copies in the reference sequence as these are more likely to be polymorphic than short repeats, and to sequences which are intragenic or closely located to the candidate gene as these are more likely to be in linkage disequilibrium with functional variants of the gene. On average, four repeat motifs had to be tested to develop one marker suitable to interrogate the candidate gene. For half of the genes, genotyping could be accomplished after testing one or two motifs while for three genes (PDZK1, NHERF1 and NHERF2), more than ten motifs were tested in order to establish a marker. Consequently, while all novel candidate genes were interrogated using at least one informative

microsatellite marker, these markers are not equivalent from gene to gene as they vary with respect to their repeat motif and the distance between marker and targeted candidate gene (Supplement-Table 2). For four genes, a distance of more than 100 kB between the microsatellite marker and the targeted candidate gene were accepted (BEST4, PDZK1, S100P, and BEST2). All other loci were interrogated with closely positioned microsatellite markers and/or intragenic SNPs. Intragenic microsatellite markers were analyzed for STX12, S100P, IFNGR1, CFTR, TNFRSF1A, SCNN1B and STAT3.

SNPs were typed using the SNaPshot technology^{S25} or by polymerase-chain reaction with subsequent restriction digest (PCR-RFLP) using an enzyme that discriminates between the two alleles. SNPs typed by PCR-RFLP were selected from internet resources. Informativity of the SNP was verified by genotyping unrelated control samples.

References for genetic markers are given in Table 2. For all markers developed or applied *de novo* within this work, primers and – if applicable – restriction enzymes for PCR-RFLP tests are shown in Supplement-Table 3.

Supplement - Table 2: Genotyped markers

Chr.	Gene	Reference for association study results on <i>European CF twins and sibs</i>	Functional classification of targeted candidate gene	Genetic marker	Type of polymorphism	Reference for genotyping method	Distance to targeted candidate gene	PIC ^A	Comments
1p35	PLA2G2A	this work	inflammation	PLA2SAT1	SAT	see Table 3 within Suppl.	31 kB	0.61	transcript PLA2G2A upregulated in CF ^B
1p34.1	STX12	this work	CFTR Network	STX12SAT2	SAT	see Table 3 within Suppl.	intragenic STX12	0.68	transcript STX12 upreg. in CF-no Res. ^D
1p33	BEST4	this work	ion channel	BEST4SAT1	SAT	see Table 3 within Suppl.	888 kB	0.81	
1p31	PRKAA2	this work	CFTR interaction partner	rs715405	SNP	Yarden et al. ^{S25}	intragenic PRKAA2		
1p31	genomic region	Ritzka et al. ^{S21} (ICM basic defect), this work (other phenotypes)	ion channel	PGM1	SNP	Ritzka et al. ^{S21}	not applicable ^C		
				D1S1162	SAT	Ritzka et al. ^{S21}	not applicable ^C		
				D1S2618	SAT	Ritzka et al. ^{S21}	not applicable ^C		
				D1S551	SAT	Ritzka et al. ^{S21}	not applicable ^C		
				D1S2807	SAT	Ritzka et al. ^{S21}	not applicable ^C		
				CLCA2SAT	SAT	Ritzka et al. ^{S21}	not applicable ^C		
				CLCA1SAT	SAT	Ritzka et al. ^{S21}	not applicable ^C		
				CLCA4SAT	SAT	Ritzka et al. ^{S21}	not applicable ^C		
				D1S2865	SAT	Ritzka et al. ^{S21}	not applicable ^C		
				D1S1588	SAT	Ritzka et al. ^{S21}	not applicable ^C		
				D1S1170	SAT	Ritzka et al. ^{S21}	not applicable ^C		
				D1S1163	SAT	Ritzka et al. ^{S21}	not applicable ^C		
				D1S2767	SAT	Ritzka et al. ^{S21}	not applicable ^C		
				D1S1174	SAT	Ritzka et al. ^{S21}	not applicable ^C		
				D1S2888	SAT	Ritzka et al. ^{S21}	not applicable ^C		
1p13.3	GSTM1	this work	detoxification	genomic deletion	indel	Baranov et al. ^{S24}	intragenic GSTM1		
1q21	PDZK1	this work	CFTR interaction partner	PDZK1SAT14	SAT	see Table 3 within Suppl.	200 kB	0.70	
1q21	S100A11	this work	CFTR Network	S100A11SAT1	SAT	see Table 3 within Suppl.	150 kB	0.87	transcript S100A11 upreg. in CF-no Res. ^D
1q31	IL10	this work	inflammation	-1082 g/a	SNP	Yarden et al. ^{S25}	promoter IL10		
				-819 c/t	SNP	Yarden et al. ^{S25}	promoter IL10		
				-592 c/a	SNP	Yarden et al. ^{S25}	promoter IL10		
				+19 c/t	SNP	Yarden et al. ^{S25}	intragenic IL10		
				+117 c/t	SNP	Yarden et al. ^{S25}	intragenic IL10		
				+953 t/c	SNP	Yarden et al. ^{S25}	intragenic IL10		
1q41	TLR5	this work	pathogen receptor	rs1861172	SNP	see Table 3 within Suppl.	intragenic TLR5		
2p21	D2S1788	this work	LEP/OB-network	D2S1788	SAT	see Table 3 within Suppl.	not applicable ^C	0.86	
2q14.1	IL1B	this work	inflammation	IL1BSAT7	SAT	see Table 3 within Suppl.	26 kB	0.65	transcript IL1B upregulated in CF ^B
				IL1BSNP2	SNP	see Table 3 within Suppl.	intragenic IL1B		
				IL1BSNP3	SNP	see Table 3 within Suppl.	intragenic IL1B		
				IL1BSNP4	SNP	see Table 3 within Suppl.	intragenic IL1B		
				IL1BSAT2	SAT	see Table 3 within Suppl.	23 kB	0.47	
				IL1BSAT4	SAT	see Table 3 within Suppl.	96 kB	0.80	
2q35	CXCR2	this work	inflammation	rs2230054	SNP	see Table 3 within Suppl.	intragenic CXCR2		
3q21	TLR9	this work	pathogen receptor	rs352140	SNP	see Table 3 within Suppl.	intragenic TLR9		
				rs187084	SNP	see Table 3 within Suppl.	intragenic TLR9		
3q27	ADIPOQ	this work	LEP/OB-network	ADIPSAT2	SAT	see Table 3 within Suppl.	9 kB	0.87	
4p16	S100P	this work	CFTR Network	S100PSAT2	SAT	see Table 3 within Suppl.	29 kB	0.36	transcript S100P upregulated in CF ^B
				S100PSAT4	SAT	see Table 3 within Suppl.	intragenic S100P	0.76	
4p13	IL8	this work	inflammation	-251 a/t	SNP	Yarden et al. ^{S25}	intragenic IL8		
				+396 t/g	SNP	Yarden et al. ^{S25}	intragenic IL8		
4q32	TLR2	this work	pathogen receptor	rs3804099	SNP	see Table 3 within Suppl.	intragenic TLR2		
5p12	PRKAA1	this work	CFTR interaction partner	rs1002424	SNP	see Table 3 within Suppl.	intragenic PRKAA1		
5q31.1	CD14	this work	pathogen receptor	rs2569190	SNP	see Table 3 within Suppl.	intragenic CD14		
				rs2563298	SNP	see Table 3 within Suppl.	intragenic CD14		

5q31	ADRB2	this work	CFTR Network	-1023 a/g -654 a/g -468 c/g C19 _{BP} R(-47c/t) G16R (46 g/a) Q27E (79 c/g)	SNP SNP SNP SNP	Yarden et al. ^{S25} Yarden et al. ^{S25} Yarden et al. ^{S25} see Table 3 within Suppl. see Table 3 within Suppl. see Table 3 within Suppl.	promoter ADRB2 promoter ADRB2 promoter ADRB2 intragenic ADRB2 intragenic ADRB2		
5q31	SLC26A2	this work	ion channel	SLC26A2SAT2 SLC26A2SNP4	SAT SNP	see Table 3 within Suppl. see Table 3 within Suppl.	84 kB	0.72	transcript SLC26A2 upreg. in CF-Res. ^D
6p21.3	HLA	this work	inflammation, immunology	D6S2239 HLA-F SAT DPB1SAT SATLTA LTA+250 TNF-1031 TNF-851 TNF-308 TNF-238 TNF+691 D6S273 HLA-DQB1 TAP2-Q687X LMP2-R60H	SAT SAT SAT SNP SNP SNP SNP SNP SNP SNP SAT — SNP SNP	see Table 3 within Suppl. see Table 3 within Suppl. Tamiya et al. ^{S78} Udalova et al. ^{S79} Waterer et al. ^{S80} see Table 3 within Suppl. Yarden et al. ^{S25} Wilson et al. ^{S81} Yarden et al. ^{S25} Yarden et al. ^{S25} see Table 3 within Suppl. Nomura et al. ^{S82} Fraile et al. ^{S83} Deng et al. ^{S84}	intragenic SLC26A2 not applicable ^C not applicable ^C not applicable ^C not applicable ^C intragenic LTA promoter TNF α promoter TNF α promoter TNF α promoter TNF α intragenic TNF α not applicable ^C intragenic DQB1 intragenic TAP2 intragenic LMP2		
6q21	CAL	this work	CFTR interaction partner	GOPCSAT5	SAT	see Table 3 within Suppl.	35 kB	0.66	
6q23	IFNGR1	this work	inflammation	IFNGR1SAT1	SAT	see Table 3 within Suppl.	intragenic IFNGR1	0.91	transcript IFNGR1 upregulated in CF ^B
7p22	KDELR2	this work	CFTR Network	KDELR2SAT2	SAT	see Table 3 within Suppl.	7 kB	0.87	transcript KDELR2 upregulated in CF ^B
7q21.2	PON2	Mekus et al. ^{S16} ; this work		PON2/DpnII PON2SAT1	SNP SAT	Mekus et al. ^{S16} see Table 3 within Suppl.	intragenic PON2 32 kB	0.47	
7q31	CAV2	Mekus et al. ^{S16} ; this work		D7S525	SAT	Mekus et al. ^{S16}	not applicable ^C		
7q32	CFTR	Mekus et al. ^{S16} (disease severity), this work (basic defect phenotypes)		Q130E MetH XV2c KM19 IVS17bTA TUB20 DHS1 ^F DHS2 ^F J3.11	SNP SNP SNP SNP SAT SNP SNP SNP SNP	see Table 3 within Suppl. Mekus et al. ^{S16} Mekus et al. ^{S16} Mekus et al. ^{S16} see Table 3 within Suppl. Mekus et al. ^{S16} see Table 3 within Suppl. see Table 3 within Suppl. Mekus et al. ^{S16}	intragenic CAV2 600 kB 200 kB 100 kB intragenic CFTR intragenic CFTR	0.52 ^E	
7q32-qtel	genomic region	Mekus et al. ^{S16} (disease severity), this work (basic defect phenotypes)		D7S514 LEP-2548GA LEP-1387GA LEP-188AC LEP+19AG 7qSAT2 7qSAT9 7qSAT5 D7S500 7qSAT3 D7S495 7qSAT12 7qSAT15 7qSAT14 7qSAT4 7qS1521 D7S500	SAT SNP SNP SNP SNP SAT SAT SAT SAT SAT SAT SAT SAT SAT SAT SAT SAT	Mekus et al. ^{S16} see Table 3 within Suppl. see Table 3 within Suppl. Mekus et al. ^{S16} Mekus et al. ^{S16} see Table 3 within Suppl. see Table 3 within Suppl. see Table 3 within Suppl. see Table 3 within Suppl. see Table 3 within Suppl. Mekus et al. ^{S16} see Table 3 within Suppl. see Table 3 within Suppl. see Table 3 within Suppl. see Table 3 within Suppl. see Table 3 within Suppl. see Table 3 within Suppl. see Table 3 within Suppl. see Table 3 within Suppl. see Table 3 within Suppl.	not applicable ^C intragenic LEP intragenic LEP intragenic LEP intragenic LEP not applicable ^C not applicable ^C not applicable ^C not applicable ^C not applicable ^C not applicable ^C not applicable ^C not applicable ^C not applicable ^C not applicable ^C not applicable ^C not applicable ^C not applicable ^C not applicable ^C not applicable ^C not applicable ^C	0.58 ^E	
								0.42 ^E	
								0.80 ^E	
								0.76 ^E	
								0.85 ^E	
								0.53 ^E	
								0.76 ^E	
								0.85 ^E	
								0.73 ^E	
								0.81 ^E	
								0.53 ^E	
								0.61 ^E	
								0.75 ^E	

9q32	TLR4	this work	pathogen receptor	rs10759930 rs1927914 rs1927911 rs2149356 rs11536891	SNP SNP SNP SNP SNP	see Table 3 within Suppl. see Table 3 within Suppl. see Table 3 within Suppl. see Table 3 within Suppl. see Table 3 within Suppl.	promoter TLR4 promoter TLR4 intragenic TLR4 intragenic TLR4 intragenic TLR4			
9q34	STOM	this work		TLR4SAT1 STOMSAT7	SAT SAT	see Table 3 within Suppl. see Table 3 within Suppl.	3 kB 27 kB	0.61 0.94	transcript STOM upreg. in CF-no Res. ^D	
10q11.2	MBL2	this work	antimicrobial defense	MBL2prom Y/X MBL2prom H/L MBL2exon1 A/D MBL2exon1 A/C MBL2exon1 A/B	SNP SNP SNP SNP SNP	Yarden et al. ^{S25} Yarden et al. ^{S25} Yarden et al. ^{S25} Yarden et al. ^{S25} Yarden et al. ^{S25}	promoter MBL2 promoter MBL2 intragenic MBL2 intragenic MBL2 intragenic MBL2			
10q22.3	SPA, SPD	this work	antimicrobial defense	SPA2: T9N SPA2: A91P SPA1: P62P SPD: T2M	SNP SNP SNP SNP	Yarden et al. ^{S25} Yarden et al. ^{S25} Yarden et al. ^{S25} see Table 3 within Suppl.	intragenic SPA2 intragenic SPA2 intragenic SPA1 intragenic SPD			
10q24	CD95	Kumar et al. ^{S22} (disease severity); this work (basic defect phenotypes)	host defence	rs1800682 rs1324551 rs2147420 rs2296603 rs7901656 rs1571019	SNP SNP SNP SNP SNP SNP	Kumar et al. ^{S22} Kumar et al. ^{S22} Kumar et al. ^{S22} Kumar et al. ^{S22} Kumar et al. ^{S22} Kumar et al. ^{S22}	promoter CD95 intragenic CD95 intragenic CD95 intragenic CD95 intragenic CD95 intragenic CD95			
11p15.5	MUC2	this work	antimicrobial defense	MUC2SAT1 MUC2SAT3	SAT SAT	see Table 3 within Suppl. see Table 3 within Suppl.	24 kB 305 kB	0.89 0.78	transcript MUC2 upregulated in CF ^B	
11p15.4	SMPD1	this work	antimicrobial defense	SMPD1SAT4	SAT	see Table 3 within Suppl.	53 kB	0.75		
11q13	GSTP1	this work	detoxification	I105V Val114A	SNP SNP	see Table 3 within Suppl. see Table 3 within Suppl.	intragenic GSTP1 intragenic GSTP1			
12p13	TNFR SCNN1A	Stanke et al. ^{S11} (disease severity); this work (basic defect phenotypes)	inflammation ion channel	D12S374 rs3181301 rs740842 rs4149584 rs1800693 rs1800692 D12S889 rs767455 rs2228576 SC4 SC3 nt7 A to G	SAT SNP SNP SNP SNP SNP SAT SNP SNP SAT SAT SNP	Stanke et al. ^{S11} Stanke et al. ^{S11} Stanke et al. ^{S11} Stanke et al. ^{S11} Stanke et al. ^{S11} Stanke et al. ^{S11} Stanke et al. ^{S11} Stanke et al. ^{S11} Stanke et al. ^{S11} Stanke et al. ^{S11} Stanke et al. ^{S11} Stanke et al. ^{S11}	480 kB intragenic CD9 intragenic FLJ10665 intragenic TNFR intragenic TNFR intragenic TNFR intragenic TNFR intragenic ENaC intragenic ENaC intragenic ENaC intragenic ENaC			
12q13	KRT8, KRT18	this work	CFTR Network	KRT8SNP6 KRT8SNP1 KRT8SNP2 KRT8SNP3 KRT8SAT1 KRT18SNP1 KRT18SNP2	SNP SNP SNP SNP SAT SNP SNP	see Table 3 within Suppl. see Table 3 within Suppl. see Table 3 within Suppl. see Table 3 within Suppl. see Table 3 within Suppl. see Table 3 within Suppl. see Table 3 within Suppl.	intragenic KRT8 intragenic KRT8 intragenic KRT8 intragenic KRT8 1 kB intragenic KRT18 intragenic KRT18		0.63	transcript KRT8 upreg. in CF-no Res. ^D
14q32.1	AAT	this work	antimicrobial defense	1237 g/a M/Z (7774 g/a) M/S (9628 a/t)	SNP SNP SNP	Yarden et al. ^{S25} Yarden et al. ^{S25} Yarden et al. ^{S25}	intragenic AAT intragenic AAT intragenic AAT			
16p13.3	NHERF2	this work	CFTR interaction partner	NHERF2SAT2 NHERF2SAT5	SAT SAT	see Table 3 within Suppl. see Table 3 within Suppl.	76 kB 32 kB	0.54 0.40		
16p12	SCNN1B, SCNN1G	Stanke et al. ^{S11} (disease severity); this work (basic defect phenotypes)	ion channel	HS3ST2-SAT rs5735 rs5723 rs1004749 rs238547 rs152730 rs250563 betaENaCGT D16S417	SAT SNP SNP SNP SNP SNP SNP SAT SAT	Stanke et al. ^{S11} Stanke et al. ^{S11} Stanke et al. ^{S11} Stanke et al. ^{S11} Stanke et al. ^{S11} Stanke et al. ^{S11} Stanke et al. ^{S11} Stanke et al. ^{S11} Stanke et al. ^{S11}	320 kB intragenic SCNN1G intragenic SCNN1G intragenic SCNN1B intragenic SCNN1B intragenic SCNN1B intragenic SCNN1B intragenic SCNN1B intragenic SCNN1B 400 kB			

16p11.2	PRSS8	this work	ion channel activator	PRSS8SAT4 PRSS8SAT5 PRSS8SNP1 PRSS8SNP2 PRSS8SNP3	SAT SAT SNP SNP SNP	see Table 3 within Suppl. see Table 3 within Suppl. see Table 3 within Suppl. see Table 3 within Suppl. see Table 3 within Suppl.	62 kB 14 kB	0.73 0.68	transcript PRSS8 upregulated in CF ^B
17p12	STX8	this work	CFTR Network	895 c/g	SNP	Yarden et al. ^{S25}	intragenic PRSS8		
17q21.31	STAT3	this work	transcription factor	STAT3SAT71A	SAT	see Table 3 within Suppl.	intragenic PRSS8	0.79	transcript STAT3 upreg. in CF-no Res. ^D
17q25.1	NHERF1	this work	CFTR interaction partner	NHERF1SAT11 NHERF1SAT6	SAT SAT	see Table 3 within Suppl. see Table 3 within Suppl.	14 kB 30 kB	0.69 0.82	
19p13.2	BEST2	this work	ion channel	BEST2SAT1	SAT	see Table 3 within Suppl.	826 kB	0.40	
19q13.2	TGFB1	Stanke et al. ^{S23}	inflammation	Leu10Pro	SNP	Stanke et al. ^{S23}	intragenic TGFB1		
19q13	genomic region	Stanke et al. ^{S23} (disease severity); this work (basic defect phenotypes)	contains CFM1 ^G	Arg25Pro D19S197 19qSAT1 19qSAT3 19qSAT6 PSGSAT8 D19S112	SNP SAT SAT SAT SAT SAT SAT	Stanke et al. ^{S23} Stanke et al. ^{S23} Stanke et al. ^{S23} Stanke et al. ^{S23} Stanke et al. ^{S23} Stanke et al. ^{S23} Stanke et al. ^{S23}	intragenic TGFB1 not applicable ^C not applicable ^C not applicable ^C not applicable ^C not applicable ^C not applicable ^C		
22q11	GSTT1	this work	detoxification	genomic deletion	indel	Sprenger et al. ^{S85}	intragenic GSTT1		
22q31.1	KDELR3	this work	CFTR Network	KDELR3SAT2	SAT	see Table 3 within Suppl.	10 kB	0.73	transcript KDELR3 upregulated in CF ^B

A polymorphism information content (PIC) values were calculated from 128 parental genotypes for microsatellite markers developed de novo for this study

B the gene was selected based on the comparison of transcriptome data from 14 F508del-*CFTR* homozygous CF vs 8 non-CF rectal epithelial tissue samples (van Barneveld et al, unpublished; The transcriptome data have been deposited in the GEO database under accession no. GSE15568.)

C the marker used to map modulator of unknown identity and/or interrogation of a genomic area suspected to contain a modifier with the following rationales:

on 1p31: genomic region contains a mediator of calcium-dependent chloride conductance,^{S21}

on 2p21: genomic region contains a QTL for serum leptin levels and fat mass,^{S56}

on 6p21.3: MHC class I, II and III genes were analyzed as CF modifiers;

on 7q31-7qtel: the genomic region contains a paternally imprinted CF modifier (Stanke et al, unpublished);

on 19q13: genomic region contains the two described CF modifiers *TGFB1* and CFM1^{S18}

D the gene was selected based on the comparison of transcriptome data from rectal suction biopsies from 4 F508del-*CFTR* homozygous patients without residual function , 3 F508del-*CFTR* homozygous patients with median residual function and 5 F508del-*CFTR* homozygous patients with high CFTR mediated residual function (van Barneveld et al, unpublished; The transcriptome data have been deposited in the GEO database under accession no. GSE15568.)

E PIC values of markers located between CFTR and 7qtel were estimated for F508del-CFTR chromosomes

F SNPs DHS1 and DHS2 correspond to the variants identified by Rowntree et al.^{S86} within the DNase hypersensitive site at 4572+14.6 kB.

DHS1 is the C/T exchange at pos. 72957 and DHS2 is the C/T exchange at pos. 73211, respectively.

G the cystic fibrosis modifier 1 (CFM1) was described near D19S112 by Zielenski et al^{S18}

Supplement - Table 3: Polymerase chain reaction primers for PCR-RFLP and microsatellite markers developed or applied *de novo* for this study

Chr.	Gene	Marker		Primer A	Primer B	Restriction enzyme
1p35	PLA2G2A	D1PLA2SAT1	SAT	AGC CCT CTG ATT TTT CAC CA	Bio- CGA CAC AAA CTC CTC CCA CT	—
1p34.1	STX12	D1STX12SAT2	SAT	TGA GGT CAG GGG TGA GTT TT	Bio- CAT GTA CTC CAG CCT CAG CA	—
1p33	BEST4	D1VMDL1SAT1	SAT	CGA GAT TGC ACC ACT GGA	Bio- CCA GCC CAG TCT CTT TGT CT	—
1p31	PRKAA2	rs715405	SNP	CCC ATT AAT GCT GAC ATT CAA A	TGG AAA CAG TAG GAG GAG ATC AA	<i>Mbo</i> I
1q21	PDZK1	D1PDZK1SAT14	SAT	ACC TCA GGG AAG TGC TGA TG	Bio- TGG GAC TGA CTG AAC CCT TC	—
1q21	S100A11	D1S100A11SAT1	SAT	TTT GTA TTA GAG GCA GAA AAG TGA	Bio- TGG CTT TGT TAT GCC TTT CC	—
1q41	TLR5	rs1861172	SNP	AGA TAA GAG GTG GCC CCA AA	GCT GAA AAG GTA GGT TGG TGA	<i>HpyCH4</i> III
2p23.3	POMC	D2S1788	SAT	AAT GGA TGG ACA AAT GGAT G	Bio-CCC TCC ATA ATT AGA TGA GCC	—
2q14.1	IL1B	D2IL1BSAT7	SAT	GGA GAC AAA GTG AAT TCC GAG T	Bio- GGC AAA TGA GAG CCT GAG AC	—
		D2IL1BSNP2	SNP	AAG GGC TCT TTT AAC CAT CAC A	GAC AAA GAT GCT ATG G	<i>Bsl</i> I
		D2IL1BSNP3-2	SNP	CTA GGT CAG TTG CCC AGG TT	CAG CTG CAG CCA ACA AGT TA	<i>Taq</i> I
		D2IL1BSNP3-3	SNP	AGG TCA GTT GTG CAG GTT CG	CAG CTG CAG CCA ACA AGT TA	<i>Bsl</i> I
		D2IL1BSNP4	SNP	GGG GAC ATG CAC CGT CCA A	ACA TCA GGG AAA AGC CAT TG	<i>Bsl</i> I
		D2IL1BSAT2	SAT	CCC AAA AGT AAT GAT CCA TGC	Bio- CCC TTG TAC CGA GCA GAT TG	—
D2IL1BSAT4	SAT	CAG GCT TGG CAT TAG GAT GT	Bio- GTG CCA ACT GAC CCA GAA CT	—		
2q35	CXCR2	rs2230054	SNP	GCT GTC GTC CTC ATC TTC CCG	AGT CCA TGG CGA AAC TCC TG	<i>BsrB</i> I
3q21	TLR9	rs352140	SNP	TTG GCT GTG GAT GTT GTT GT	AAG CTG GAC CTC TAC CAC GA	<i>BstU</i> I
		rs187084	SNP	TGT ACT GGA TCC TGG GGA TG	GAG CTC CTT TGC CTG GTC TA	<i>Hpy188</i> III
3q27	ADIPOQ	D3ADIPSAT2	SAT	AAA GGA GGG TGT TGA GTC CA	Bio- GGC TGC ATG ACA ACT GAA AG	—
4p16	S100P	D4S100PSAT2	SAT	GGG TGA CAA GAG CGA AAC TT	Bio- CTT CCA TCT TGC TGG CTC TT	—
		D4S100PSAT4	SAT	ATC TGC TGA TGG ACA GTT GG	Bio- CGA GAT CCC GCC ATT G	—
4q32	TLR2	rs3804099	SNP	TGG TGG ACT TAC CTT CCT TGA	CAA ACA TTC CAC GGA ACT TG	<i>HpyCH4</i> IV
5p12	PRKAA1	rs1002424	SNP	GAA GAG GGC CAC AAT CAA AG	AAA GCC CTC CCT CTT ACG AA	<i>Hpa</i> II
5q31.1	CD14	rs2569190	SNP	CCT CCC CAC CTC TCT TCC T	CAC CCA CCA GAG AAG GCT TA	<i>Ava</i> II
		rs2563298	SNP	CCG CAG TTC TTT TCT TGA GG	CGT CAG GAC GTT GAG GAC TT	<i>Sau96</i> I
5q31	ADRB2	C19 _{BUP} R(-47c/t)	SNP	CGC TGA ATG AGG CTT CCA G	CAG GCC AGT GAA GTG ATG AA	<i>MspA1</i>
		G16R (46 g/a)	SNP	CTT CTT GCT GGC ACG CAA T	GGA CGA TGA GAG ACA TGA CG	<i>BsrD1</i>
		Q27E (79 c/g)	SNP	CTT CTT GCT GGC ACG CAA T	CAG GCC AGT GAA GTG ATG AA	<i>Fnu4H</i> I
5q31	SLC26A2	D5SLC26A2SAT2	SAT	TCT TCC CTC TGC CTG TGT CT	Bio- GTG CCA CTG CCC TCC A	—
		D5SLC26A2SNP4	SNP	GGC CCT CTC TGT CTT CTC AG	CGA GAC GGC TTG GTT ATC TT	<i>Tsp509</i> I
6p21.3	HLA	D6S2239	SAT	GTT GGA AGC AAT GGA TTA GAT GTC C	Bio- CTA CCT GCC AGG AAC AAT ATA CAC	—
		HLA-F SAT	SAT	CCT GAG AAT GAA GGT CTA GAG AC	Bio- GCT CAG GTA CAA CTT TTC CAG AG	—
		DPB1SAT	SAT	ACT TGA ACC CCA GAG GCA G	Bio- TTC AGT AGC TTT AGT CTT TCC CC	—
		SATLTA	SAT	Bio- CCT CTA GAT TTC ATC CAG CCA CA	CTC TCT CCC CTG CAA CAC ACA	<i>Nco</i> I
		LTA	SNP	CCG TGC TTC GTG CTT TGG ACT A	AGA GGG GTG GAT GCT TGG GTT TC	<i>Bbs</i> I
		TNF-1031	SNP	CTG TGG GGA GAA CAA AAG GA	CTC CTA CCC ATT GCT GTG GT	<i>Nco</i> I
		TNF-308	SNP	AGG CAA TAG GTT TTG AGG GCC AT	GCT CAT CTG GAG GAA GCG GTA	—
		D6S273	SAT	GCA ACT TTT CTG TCA ATC CA	ACC AAA CTT CAA ATT TTC GG	—
		TAP2	SNP	AAG CTG CAG AAG CTT GCC CGG ATC	CCG CAC AGC TTT AGG GAA ACT C	<i>Bfa</i> I
		PSMB9	SNP	GCC AGC AAG AGC CGA AAC AAG	GTG AAC CGA GTG TTT GAC AAG C	<i>Hha</i> I

Chr.	Gene	Marker		Primer A	Primer B	Restriction enzyme
6q21	CAL	D6GOPCSAT5	SAT	GAA AAC ACA CTT GTG GTG TCT CA	Bio- AGC CAG CCG AGA TCG AG	—
6q23	IFNGR1	D6IFNGR1SAT1	SAT	AAG CAT ATA TTG TAA CTC AGT CTT TGA	Bio- TTC CTC GAA AAT ATA CTT GCA TCA	—
7p22	KDELR2	D7KDELR2SAT2	SAT	ATA GCA GCC CTA GCA AAC GA	Bio- CCC CAT CCT CTG AGT GAA AA	—
7q21.2	PON2	PON2 PON2SAT1	SNP SAT	GGC ACA AAT GAT CAC TAT TTT CTT GAC TGC AAA TAA GAT TAA AGT ACC CTT TG	GAG TAA ATC CAC TAC ATT TCA Bio- CAG CTA CTC AGG CAG GTG AA	<i>Dpn II</i> —
7q31	CAV2	D7S525 Q130E	SAT SNP	CTT GCT GTT TAA GTA CCA CAA GTT C GAG AAT GTG GAG GTC CCA GA	Bio- GTT AGC CGA GAT TGC CC CGT CCT ACG CTC GTA CAC AA	— <i>HpyCH IV</i>
7q32-qtel	genomic region	D7S514 7qSAT2 7qSAT9 7qSAT5 D7S500 7qSAT3 D7S495 7qSAT12 7qSAT15 7qSAT14 7qSAT4 D7S1521 D7S550	SAT SAT SAT SAT SAT SAT SAT SAT SAT SAT SAT SAT SAT SAT	TGC AGT TAG TGC CAA TGT CA AGG GAG GAT CTC CTG AGC AC AGT TAG TTG GCT GGG CTC TG CAC ATG CTT GCC ACC TCT TA ATT GAT TGA GGA ACT GAA CTT ACC T ATG GCG TGA ACC CGA GAG AGC ACC TGG TCC AAT TTT CT CCA GTA CTC AAT GCC CCT GT CTG ATT AGG GAA AAT GGG TGA CTG GGT AAC AGG GTG AGA CG GGG AGG TGG GGG TTA CAG CCA AGA TTG TGC CAC TGT GT TGT GAA TGC ACT ATC ATC CA	Bio- TGG GCA ACA CAG CAA A Bio- AGG CAC ATT CCA CCA TGT CT Bio- TGA GAC CCT CTC TCC ATG ACT Bio- TGC CTT CCT CTT GCA CTC TT Bio- CCA GAA TTG AAA ACT CAG CA Bio- CCA CTC AGA AGT CCC AGC AT Bio- TGG CAT TCA TTT ACA ATA GCC Bio- TTG TGC AGA GCA GAC AGG TT Bio- ATG CTA CCG CAC TCC AGT CT Bio- AGG TGT GCA GTG GTG TCT TG Bio- TTC ACC TAC TTT GCA GAC CAG A Bio- CTT GAT GAA AGC TGG TAA GTG Bio- GTG TAT CCA TTC ACC TGA TAG G	— — — — — — — — — — — — — —
9q32	TLR4	rs10759930 rs1927914 rs1927911 rs2149356 rs11536891 TLR4SAT1	SNP SNP SNP SNP SNP SAT	GGG TGC ACT CAC TCA CCT CT TGG GAT TAA ATG AAC TGG CAT T TCCATATCATTGGGGAGACTG TGA CTG GTA AAT ATC CAT TTC AGA GA GGG TGT GTT TCC ATG TCT CA Bio- TGG TTG AAA AGG CAC ATG AA	CCT TGG ACA CCC ATT ACC AG ACA AAA TGG TCC CTC ACA GC TGGGAATCCATGCACTCTAAA TTT CCA CAA AAC TCG CTC CT GCA TAA GGG ATA AGG GGA GA ACC CTG CAT TTT GGC ATA AC	<i>Dra I</i> <i>Sph I</i> <i>BsaI I</i> <i>Tsp509 I</i> <i>Hha I</i> —
9q34	STOM	D9STOMSAT7	SAT	CTG TGA CTT TAG AAC CTC CAG TT	Bio- ATT CCA GCC TGG GTG A	—
10q22.3	SPD	rs721917	SNP	GAA GAC CTA CTC CCA CAG AGC AA	TTG GGA GGA AGA AAC ACG TC	<i>BsrD I</i>
10q24	CD95	rs1800682 rs1324551 rs2147420 rs2296603 rs7901656 rs1571019	SNP SNP SNP SNP SNP SNP	ATA TAG CTG GGG CTA TGC GAT CTC TGA ATC TGC AAG CTG GCA TTT CT TTC TGT CTC TGA TGA AAT CTT GG TTT TAC AGT TTT TGG TTC CCC TA TCT GGG AAT CTC CAG TTT GTT T GAT CTT TTT AGG CAG GAG TTC TG	CTC AGA GAA AGA CTT GCG GG TGT AAG TCG CTG CCT GAG TG ACA GCG CAA TGA GAT CCT AAA TTC TCA TTT CAG AGG TGC ATG T GCT CTG CTC ACC TAT ACA GCA A ACC TGC TCA GCA TAA AGC AT	<i>ScrF I</i> <i>Hph I</i> <i>HpyCH4 IV</i> <i>Nla III</i> <i>HpyCH4 IV</i> <i>Rsa I</i>
11p15.5	MUC2	D11MUC2SAT1 D11MUC2SAT3	SAT SAT	GCC ACC GTT GTG AAG AAG AT TTC AAA AGA GCG GAA GTT CA	Bio- AGT CAG GAC AGA CAG CGT CA Bio- GCA GAG CAC GAG ACT CCA TC	— —
11p15.4	SMPD1	D11SMPD1SAT4	SAT	CCA ATT GAC AGT GGA TTT TTG A	Bio- CAT AAA AGG GCC CGA TAC AA	—
11q13	GSTP1	I105V Val114A	SNP SNP	TGT GTG GCA GTC TCT CAT CCT T TGG CAG CTG AAG TGG ACA GGA TT	TAC TTG GCT GGT TGA TGT CCC A ATG GCT CAC ACC TGT GTC CAT	<i>BsmA I</i> <i>Aci I</i>

Chr.	Gene	Marker		Primer A	Primer B	Restriction enzyme
12q13	KRT8, KRT18	D12KRT8SNP6 D12KRT8SNP1 D12KRT8SNP2 D12KRT8SNP3 D12KRT8SAT1 D12KRT18SNP1 D12KRT18SNP2	SNP SNP SNP SNP SAT SNP SNP	GTT CTG CTC ACC CCT TCC TC TGA TCT GGG CTA AGG TGG TC GAG ATC AAC TTC CTC CAG CAG TGA ATT AAG AGA AAA GAC GAA TTG C AAA TGA GTG AAT AAA CAT CAC ACG TCA CCT AAT GGT GGG GAG AG AGA ACC ACG AAG AGG CAA GC	AAC TCT TTC CTT TTG GGG AGA GGT GCT TCC TCT TCC TTT CC GCC TCT GGT TGA GTC TCA GG TCC AGC ATC TTG TTC TGC TG Bio- CCT TCA GAT GTA GAG GGA CGA CTC ACA TTC ACT GCC ACC TG AGA ATG CTC TTC ATC AGA GC	<i>Hae</i> III <i>Ava</i> II <i>Alw</i> NI <i>Xmn</i> I — <i>Hpy</i> CH4 IV <i>Aci</i> I
16p13.3	NHERF2	D16NHERF2SAT2 D16NHERF2SAT5	SAT SAT	TCA GCA GAG TTC TGG GAG GT ACA ACG GGC CTA GAG TGA AC	Bio- CCC TGG CAG TAA GTC CTG AA Bio- CTG AGC AGG GAG GAT CAC AT	— —
16p11.2	PRSS8	D16PRSS8SAT4 D16PRSS8SAT5 D16PRSS8SNP1 D16PRSS8SNP3 D16PRSS8SNP4	SAT SAT SNP SNP SNP	ACC TGG GCA GAT CAC TTG AG CAG TCT TGG AGC TCC TGG TT GCC AGG ATG GTG ACA GGT CAT TGC TTG ATT GGA AGA ACA GC AGG GAG GGA AGA AGG CCG GAG TG	Bio- AGA GGA GTG GGA GGT GAG GT Bio- GGG CGA CAG AGC ATG ATT CTG TCC ATC AAG GAC CTC AG ATT GTG AGA CTG GGC AGC AT CGC CTC AGG TCT TTC AGG T	— — <i>Hpy</i> 188 III <i>Alo</i> I <i>Bgl</i> I
17p12	STX8	895 c/g	SNP	CGA CCA ACT GAT GGC AGT AA	GGC CGG AGC TTT AAG TCT TT	<i>Scr</i> F I
17q21.31	STAT3	D17STAT3SAT71A	SAT	TTC TGC CTG GTC ACT GAC TG	Bio- GGA GGT ACG GGT CCT CAA AG	—
17q25.1	NHERF1	D17NHERF1SAT11 D17NHERF1SAT6	SAT SAT	TGG ATG ACA GAG CTC CAC CT CTA TCT TAG TCC CGG CTG CT	Bio- GCA CCT GGC CTG GTA TGC Bio- AGC CAA GAT CGT GCC ATT	— —
19p13.2	BEST2	D19VMDL2L1SAT1	SAT	CCC CAT CTC AAA AAC CAA AA	Bio- TGC TGT ATG AGG ATG CCC TA	—
19q13	genomic region	D19S197 D19SAT1 D19SAT3 D19SAT6 D19PSGSAT8 D19S112	SAT SAT SAT SAT SAT SAT	Bio- TCC TCA GCA GGT CTC CTC AT TTT GTG TAA TGC CTG TTT CTC C AAA ACC TAT TGC CCT CTT CTC A TTT GGT ATC TAC CCA GAA GAA AA GGT TCT CTT CAC CCA TGA GC Bio- GCC AGC CAT TCA GTC ATT TGA AG	CAC CAG GGA AAT GCC AAT Bio- GCC CAA ATT ATT GCA ACA GAA Bio- GGG AGA CTC CAT CTC ATC TCA Bio- GCT GAG GAT GCC ATT AAT TC Bio- GGA TTC CAC ACA GGG AAA TG CTG AAA GAC ACG TCA CAC TGG T	— — — — — —
22q31.1	KDELR3	D22KDELR3SAT1	SAT	GGC AAC AGA GCC AGA TCC TA	Bio- TGG GCC TTT GTT TCC TCA	—

C Computation of P-values

Genetic data for the association study was evaluated using the FAMHAP software package^{S86} which allows family-based analysis and accepts data evaluation in association studies on unrelated individuals as well as on affected sib pairs. FAMHAP uses an expectation-maximization algorithm to reconstruct haplotypes at loci investigated with more than one marker.^{S88,S89} Non-informative phases are handled by the program through the acknowledgement of likelihood-weighted haplotype explanations whereby all haplotypes consistent with the Mendelian law are considered and their weight is attached in accordance with the haplotype distribution observed in the data set. In contrast to many programs that restrict further analyses to the most likely haplotype only, FAMHAP allows all weighted haplotype explanations proportionally during the association study (see below). To enable a consistent assignment of weighted haplotype explanations within all subgroups, the entire genotyping data set composed of 101 CF families with a total of 171 CF patients was provided as training set to FAMHAP in all case-reference comparisons.

Comparisons of observed genotype distribution to a distribution expected for a population with similar allele frequencies under the conditions required for Hardy-Weinberg-equilibrium were not analyzed. Instead, all loci were checked for distortions due to a selection bias or survivor effect using the transmission-disequilibrium test (TDT). The TDT compares allele- and/or haplotype distributions between transmitted and non-transmitted chromosomes. An overrepresentation of alleles among the non-transmitted chromosomes indicates that this allele is a risk allele in the analyzed population, which can for instance be due to the fact that carriers of the risk-allele were not recruited into the study.^{S11} As described previously,^{S6} we have recruited sibs and twins from pairs where both siblings were alive in 1996. Even though the basic clinical data necessary for ranking might have been accessible through clinical charts, we did not recruit sib pairs where one sibling was already deceased.

Genotype data was compared between groups of patients or patient pairs with contrasting phenotypes for disease severity of manifestation of the basic defect assessed in-vivo through

potential difference measurement of the nasal epithelium or ex-vivo in intestinal biopsies mounted in an micro-Ussing chamber. FAMHAP uses a Monte Carlo simulation based association test to compare the data sets for the phenotypically defined cases and references.^{S88,S89} In order to account for the dependence of the genotypes of the individuals within each sib pair, in each permutation replicate the affection status was simultaneously permuted or not permuted with equal probability for both sibs. 10.000 permutations were carried out at all loci and for all tests. As we cannot specify a priori a genetic model, haplotype (allele) as well as diplotype (genotype) distributions were compared. In addition, the accumulation of rare alleles among cases or references was analyzed for.

Supplement – Table 4: Overview of case-reference tests

Phenotype	Case	Reference	Stringency	Model
disease severity	CON+	CON-	S	hap, dip, div
disease severity	DIS-	DIS+	E	hap, dip, div
disease severity	DIS	CONC	S,E	hap, dip, div
basic defect	ICM Res.	ICM no Res.	S,E	hap, dip, div
basic defect	ICM DIDS Res.	ICM no Res.	S,E	hap, dip, div
basic defect	NPD Amil ≤ 21	NPD Amil ≥ 31	S,E	hap, dip, div
basic defect	NPD GI = 0	NPD GI < -5	S,E	hap, dip, div
basic defect	NPD ATP = 0	NPD ATP < 0	E	hap, dip, div

CON+: concordant/mildly affected sib pair; CON-: concordant severely affected sib pair; DIS+: mildly affected sibling of a discordant sib pair; DIS-: severely affected sibling of a discordant sib pair; DIS: discordant sib pair; CONC: concordant sib pair; ICM no Res.: no residual chloride secretory response observed by intestinal current measurement (ICM); ICM Res.: cAMP activatable and DIDS-insensitive residual chloride secretion observed by ICM; ICM DIDS Res.: DIDS-sensitive residual chloride secretion observed by ICM; NPD Amil: change of nasal potential upon superfusion with amiloride [mV]; NPD GI: change of nasal potential upon superfusion with chloride-free gluconate solution and isoproterenol solution [mV]; NPD ATP: change of nasal potential upon superfusion with adenosine triphosphate solution; S: test was carried out using a stringent phenotype definition; E: test was carried out using an extended phenotype definition referring to a slightly larger sample set; hap: comparison of haplotype distribution using FAMHAP command hapcc; dip: comparison of diplotype distributions using FAMHAP-command dipcc; div: investigation of accumulation of rare haplotypes using FAMHAP-command diversityhap

D Within-data-set-validation strategies

Our association study is based on a systematic comparison of patient subsamples characterized by non-overlapping phenotypic properties for CF disease severity or CF basic defect. We have summarized at 9 to 29 independent nuclear families as subsamples (Figure 1 of the main manuscript) whereby the sample size — for instance, for a rare phenotype such as “presence of CFTR-mediated residual chloride secretion in F508del-*CFTR* homozygotes” — was limited by the incidence of CF in the central European population and by the etiology and pathophysiology of CF disease that determines the frequencies of the endophenotype in the CF population. Consequently, we cannot generate two independent subsamples of reasonable size for the same phenotypic contrast from our patient pool to do a replication study. Instead, the stability of the observed associations was assessed by testing overlapping subsamples as partial replicates.

Association tests that rely on nasal potential difference measurement (NPD) phenotypes were done using an index-case strategy. When selecting the index case from a pair with a concordant NPD phenotype, two subsets were generated which differ in the sibling recruited from each concordant pair whereby the assignment of index sibling from a concordant pair to the subset was done at random. The resulting two subsets are identical in sample size of cases and controls and they assess the same repertoire of independent parental chromosomes. Consequently, the level of significance of an association test must match if it is a true-positive finding and not a random perturbation. Furthermore, we generated partial replicates by using a stringent and an extended definition for most phenotypes (see Mekus et al.^{S6} for stringent and extended definition of disease severity). In these cases, the entire stringent data set is a subgroup of the extended data set.

Supplement – Table 5: Sample size for stringent and extended phenotype definition

Phenotype	Case	N° of chr.	N° of pat.	N° of fam.	Reference	N° of chr.	N° of pat.	N° of fam.
disease severity	S CON+	52	26	13	CON-	46	23 ^a	12
disease severity	E DIS-	38	19	19	DIS+	38	19	19
disease severity	S DIS	56	28	14	CONC	98	49	24
disease severity	E DIS	88	44	22	CONC	122	61	30
basic defect	S ICM Res.	18	9	9	ICM no Res.	28	14	14
basic defect	E ICM Res.	44	22	22	ICM no Res.	28	14	14
basic defect	S ICM DIDS Res.	20	10	10	ICM no Res.	28	14	14
basic defect	E ICM DIDS Res.	28	14	14	ICM no Res.	28	14	14
basic defect	S NPD Amil ≤ 21	26	13	13	NPD Amil ≥ 31	34	17	17
basic defect	E NPD Amil ≤ 27	30	15	15	NPD Amil ≥ 28	54	27	27
basic defect	S NPD GI = 0	40	20	20	NPD GI < -5	26	13	13
basic defect	E NPD GI = 0	40	20	20	NPD GI < 0	38	19	19
basic defect	E NPD ATP = 0	30	15	15	NPD ATP < 0	54	27	27

a: includes one sibling from a CON- pair; no DNA suitable for genotyping available for other sibling of this pair for all loci analyzed since 1998
N° of chr. : number of chromosomes; N° of pat: number of patients; N° of fam: number of independent families

Supplement– Table 6: Number of tests on partial replicates

Phenotype	Case	Reference	Phenotype definition	randomized subsampling of index sib done?	N° of tests done to investigate phenotype
disease severity	CON+	CON-	S	not applicable	1
disease severity	DIS-	DIS+	E	not applicable	1
disease severity	DIS	CONC	S,E	not applicable	2
basic defect	ICM Res.	ICM no Res.	S,E	no	2
basic defect	ICM DIDS Res.	ICM no Res.	S,E	no	2
basic defect	NPD Amil low	NPD Amil high	S,E	yes	4
basic defect	NPD GI = 0	NPD GI < 0	S,E	yes	4
basic defect	NPD ATP = 0	NPD ATP < 0	E	yes	2

CON+: concordant/mildly affected sib pair; CON-: concordant severely affected sib pair; DIS+: mildly affected sibling of a discordant sib pair; DIS-: severely affected sibling of a discordant sib pair; DIS: discordant sib pair; CONC: concordant sib pair; ICM no Res.: no residual chloride secretory response observed by intestinal current measurement (ICM); ICM Res.: cAMP activatable and DIDS-insensitive residual chloride secretion observed by ICM; ICM DIDS Res.: DIDS-sensitive residual chloride secretion observed by ICM; NPD Amil: change of nasal potential upon superfusion with amiloride [mV]; NPD GI: change of nasal potential upon superfusion with chloride-free gluconate solution and isoproterenol solution [mV]; NPD ATP: change of nasal potential upon superfusion with adenosine triphosphate solution; S: test was carried out using a stringent phenotype definition; E: test was carried out using an extended phenotype definition.

Supplement – Table 7: Examples for outcome of related tests on partial replicates

Gene/Marker	Phenotype	Case ^A	Reference ^A	Set ^B	Sub ^C	raw P-value (Model) ^D	P-value given in Tab. 8
<i>IL1B</i>	basic defect	ICM Res.	ICM no Res.	S	n.a.	0.0930 (hap)	0.0235
		ICM Res.	ICM no Res.	E	n.a.	0.0235 (hap)	
	basic defect	ICM DIDS Res.	ICM no Res.	S	n.a.	0.0294 (dip)	0.0294
		ICM DIDS Res.	ICM no Res.	E	n.a.	0.0664 (dip)	
	basic defect	NPD Amil low	NPD Amil high	S	A	0.0166 (dip)	0.0039
		NPD Amil low	NPD Amil high	S	B	0.0094 (dip)	
		NPD Amil low	NPD Amil high	E	A	0.0039 (dip)	
		NPD Amil low	NPD Amil high	E	B	0.0059 (dip)	
basic defect	NPD ATP = 0	NPD ATP < 0	E	A	0.0055 (hap)	0.0055	
	NPD ATP = 0	NPD ATP < 0	E	B	0.0293 (hap)		
<i>CLCA2Sat</i>	basic defect	NPD GI = 0	NPD GI < -5	S	A	0.0138 (dip)	0.0138
		NPD GI = 0	NPD GI < -5	S	B	0.0508 (dip)	
		NPD GI = 0	NPD GI < 0	E	A	0.0324 (dip)	
		NPD GI = 0	NPD GI < 0	E	B	0.0590 (dip)	
<i>GSTM1</i>	basic defect	NPD GI = 0	NPD GI < -5	S	A	0.1090	0.0260
		NPD GI = 0	NPD GI < -5	S	B	0.0567	
		NPD GI = 0	NPD GI < 0	E	A	0.0840	
		NPD GI = 0	NPD GI < 0	E	B	0.0260	
<i>IL8</i>	basic defect	NPD GI = 0	NPD GI < -5	S	A	0.0980 (dip)	0.0215
		NPD GI = 0	NPD GI < -5	S	B	0.0215 (dip)	
		NPD GI = 0	NPD GI < 0	E	A	0.1010 (dip)	
		NPD GI = 0	NPD GI < 0	E	B	0.0259 (dip)	
<i>ADRB2</i>	disease severity	DIS	CONC	S	n.a.	0.0039 (div) 0.0081 (hap)	0.0039
		DIS	CONC	E	n.a.	0.0901 (div) 0.0419 (hap)	
	basic defect	NPD GI = 0	NPD GI < -5	S	A	0.0100 (hap)	0.0044
		NPD GI = 0	NPD GI < -5	S	B	0.0359 (hap)	
		NPD GI = 0	NPD GI < 0	E	A	0.0044 (hap)	
		NPD GI = 0	NPD GI < 0	E	B	0.0082 (hap)	
<i>TNFα</i>	disease severity	CON+	CON-	S	n.a.	0.0081 (dip) 0.0699 (hap)	0.0081
	basic defect	NPD Amil low	NPD Amil high	E	A	0.0069 (hap)	0.0069
		NPD Amil low	NPD Amil high	E	B	0.0263 (hap)	
	basic defect	NPD GI = 0	NPD GI < 0	E	A	0.0735 (hap)	0.0326
		NPD GI = 0	NPD GI < 0	E	B	0.0326 (hap)	
	basic defect	NPD ATP = 0	NPD ATP < 0	E	A	0.0039 (dip)	0.0039
NPD ATP = 0		NPD ATP < 0	E	B	0.0170 (dip)		
<i>IFNGR1Sat</i>	disease severity	DIS	CONC	S	n.a.	0.0021 (hap) 0.0260 (dip) 0.0068 (div)	0.0021
		DIS	CONC	E	n.a.	0.0024 (hap) 0.0056 (dip) 0.0207 (div)	
<i>CAV2</i>	basic defect	NPD GI = 0	NPD GI < -5	S	A	0.0403 (dip)	0.0064
		NPD GI = 0	NPD GI < -5	S	B	0.0445 (dip)	
		NPD GI = 0	NPD GI < 0	E	A	0.0064 (dip)	
		NPD GI = 0	NPD GI < 0	E	B	0.0064 (dip)	
<i>CFTR</i>	disease severity	DIS	CONC	S	n.a.	0.0678 (hap) 0.0580 (dip)	0.0039
		DIS	CONC	E	n.a.	0.0168 (hap) 0.0039 (dip)	
	basic defect	NPD GI = 0	NPD GI < -5	S	A	0.0067 (dip)	0.0067
		NPD GI = 0	NPD GI < -5	S	B	0.0116 (dip)	
		NPD GI = 0	NPD GI < 0	E	A	0.0268 (dip)	
		NPD GI = 0	NPD GI < 0	E	B	0.0956 (dip)	
	basic defect	NPD ATP = 0	NPD ATP < 0	E	A	0.0009 (hap)	0.0008
		NPD ATP = 0	NPD ATP < 0	E	B	0.0008 (hap)	
<i>LEP</i>	disease severity	DIS	CONC	S	n.a.	0.0438 (hap) 0.0063 (dip)	0.0063
		DIS	CONC	E	n.a.	0.0494 (hap) 0.0485 (dip)	
	basic defect	NPD Amil low	NPD Amil high	S	A	0.0162 (hap)	0.0103
		NPD Amil low	NPD Amil high	S	B	0.0103 (hap)	
	basic defect	NPD GI = 0	NPD GI < -5	S	A	0.0076 (hap)	0.0070
		NPD GI = 0	NPD GI < -5	S	B	0.0076 (hap)	
		NPD GI = 0	NPD GI < 0	E	A	0.0070 (hap)	
		NPD GI = 0	NPD GI < 0	E	B	0.0070 (hap)	
	basic defect	NPD ATP = 0	NPD ATP < 0	E	A	0.0125 (hap)	0.0015
		NPD ATP = 0	NPD ATP < 0	E	B	0.0015 (hap)	

Gene/Marker	Phenotype	Case ^A	Reference ^A	Set ^B	Sub ^C	raw P-value (Model) ^D	P-value given in Tab. 8
<i>CD95</i>	basic defect	ICM DIDS Res.	ICM no Res.	S	n.a.	0.0430 (div)	0.0035
		ICM DIDS Res.	ICM no Res.	E	n.a.	0.0035 (div)	
	basic defect	NPD GI = 0	NPD GI < -5	S	A	0.0047 (dip)	0.0047
		NPD GI = 0	NPD GI < -5	S	B	0.0422 (dip)	
		NPD GI = 0	NPD GI < 0	E	A	0.0063 (dip)	
		NPD GI = 0	NPD GI < 0	E	B	0.0236 (dip)	
<i>MUC2</i>	disease severity	DIS	CONC	S	n.a.	0.0021 (hap) 0.0906 (dip)	0.0021
		DIS	CONC	E	n.a.	0.0043 (hap) 0.0748 (dip)	
<i>GSTP1</i>	basic defect	NPD ATP = 0	NPD ATP < 0	E	A	0.0017 (hap) 0.0021 (dip) 0.0027 (div)	0.0006
		NPD ATP = 0	NPD ATP < 0	E	B	0.0006 (hap) 0.0008 (dip) 0.0014 (div)	
<i>TNFR</i>	basic defect	NPD GI = 0	NPD GI < -5	S	A	0.0031 (hap)	0.0004
		NPD GI = 0	NPD GI < -5	S	B	0.0004 (hap)	
		NPD GI = 0	NPD GI < 0	E	A	0.0051 (hap)	
		NPD GI = 0	NPD GI < 0	E	B	0.0025 (hap)	
	basic defect	NPD ATP = 0	NPD ATP < 0	E	A	0.0035 (hap)	0.0003
		NPD ATP = 0	NPD ATP < 0	E	B	0.0003 (hap)	
basic defect	ICM Res.	ICM no Res.	S	n.a.	0.0379 (hap) 0.0061 (dip)	0.0012	
	ICM Res.	ICM no Res.	E	n.a.	0.0348 (hap) 0.0012 (dip)		
<i>KRT8</i>	basic defect	ICM Res.	ICM no Res.	S	n.a.	0.0085 (hap) 0.0106 (dip)	0.0004
		ICM Res.	ICM no Res.	E	n.a.	0.0018 (hap) 0.0004 (dip)	
<i>NHERF2</i>	basic defect	ICM Res.	ICM no Res.	S	n.a.	0.0339 (hap) 0.0044 (dip) 0.0297 (div)	0.0001
		ICM Res.	ICM no Res.	E	n.a.	0.0007 (hap) 0.0001 (dip) 0.0001 (div)	
<i>SCNN1B/ SCNN1G</i>	basic defect	NPD GI = 0	NPD GI < -5	S	A	0.0106 (hap)	0.0024
		NPD GI = 0	NPD GI < -5	S	B	0.0135 (hap)	
		NPD GI = 0	NPD GI < 0	E	A	0.0024 (hap)	
		NPD GI = 0	NPD GI < 0	E	B	0.0075 (hap)	

A DIS+: mildly affected sibling of a discordant sib pair; DIS-: severely affected sibling of a discordant sib pair; DIS: discordant sib pair; CONC: concordant sib pair; ICM no Res.: no residual chloride secretory response observed by intestinal current measurement (ICM); ICM Res.: cAMP activatable and DIDS-insensitive residual chloride secretion observed by ICM; ICM DIDS Res.: DIDS-sensitive residual chloride secretion observed by ICM; NPD Amil: change of nasal potential upon superfusion with amiloride [mV]; NPD GI: change of nasal potential upon superfusion with chloride-free gluconate solution and isoproterenol solution [mV]; NPD ATP: change of nasal potential upon superfusion with adenosine triphosphate solution; S: test was carried out using a stringent phenotype definition; E: test was carried out using an extended phenotype definition.

B S = stringent, E = extended data set as defined in Supplement-Table 5

C Sub: n.a. – not applicable; A,B: when selecting the index case from a pair with a concordant NPD phenotype, two subsets (A,B) were generated which differ in the sibling recruited from each concordant pair whereby the assignment of the index sibling to the subset was done at random

D raw P-values were calculated by 10000 MC simulations. Model: hap = comparison of haplotype distributions; dip = comparison of diplotypes distributions; div = accumulation of rare haplotypes among cases or references

Supplement - Table 8: Results of the association study (P_{raw} -values)

	Locus	Gene	Phenotype	P	Test	Case	Reference	Model	A
hypothesis rejected	1p31	PRKAA2		> 0.12	all				
	3q27	ADIPOQ		> 0.14	all				
	4p16	S100P		> 0.25	all				
	5p12	PRKAA1		> 0.37	all				
	9q34	STOM		> 0.15	all				
	14q32.1	AAT ^B		> 0.14	all				
isolated finding ^C	1p34.1	STX12	basic defect	0.0295	CC	ICM Res.	ICM DIDS Res.	E	dip
			basic defect	0.0402	CC	NPD Amil ≤ 27	NPD Amil ≥ 28	E	hap
	1p33	BEST4	disease severity	0.0506	CC	CON+	CON-	S	hap
			disease severity	0.0866	CC	DIS	CONC	E	dip
	1q21	S100A11	disease severity	0.0460	CC	CON+	CON-	S	dip
	1q31	IL10	disease severity	0.0038	CC	DIS-	DIS+	E	dip
			disease severity	0.0491	CC	DIS	CONC	S	dip
	1q41	TLR5	basic defect	0.0472	CC	ICM DIDS Res.	ICM no Res.	S	hap
			basic defect	0.0920	CC	NPD GI =0	NPD GI < 0	E	hap
	2p23.3	POMC	disease severity	0.0232	CC	DIS	CONC	E	hap
	2q35	CXCR2	disease severity	0.0922	TDT	transmitted	non-transmitted	E	hap
	4q32	TLR2	disease severity	0.0709	CC	CON+	CON-	S	hap
	5q31	SLC26A2	disease severity	0.0110	CC	CON+	CON-	S	div
			disease severity	0.0785	CC	DIS	CONC	S	hap
	6q21	GOPC	basic defect	0.0066	CC	ICM Res.	ICM DIDS Res.	E	hap
			basic defect	0.0669	CC	ICM DIDS Res.	ICM no Res.	E	hap
			basic defect	0.0102	CC	NPD GI =0	NPD GI < 0	E	dip
	7q21.1	PON2	disease severity	0.0515	CC	CON+	CON-	S	hap
	7q31.1	CAV2	basic defect	0.0064	CC	NPD GI =0	NPD GI < 0	E	dip
	10q11.2	MBL2	disease severity	0.0333	CC	DIS-	DIS+	E	hap
			disease severity	0.0426	CC	DIS	CONC	E	dip
	17p12	STX8	basic defect	0.0410	CC	ICM Res.	ICM no Res.	E	hap
	17q21.31	STAT3	basic defect	0.0035	CC	ICM Res.	ICM no Res.	E	div
			basic defect	0.0171	CC	NPD Amil ≤ 22	NPD Amil ≥ 31	S	dip
	19p13.2	BEST2	disease severity	0.0833	CC	DIS	CONC	S	hap
	19q13.2	TGFB1	basic defect	0.0354	CC	ICM DIDS Res.	ICM no Res.	S	hap
	19q13	D19S112 (CFM1)	basic defect	0.0859	CC	ICM Res.	ICM no Res.	E	hap
	22q11	GSTT1	disease severity	0.0198	TDT	transmitted	non-transmitted	E	hap
			disease severity	0.0409	CC	DIS-	DIS+	E	hap
			disease severity	0.0699	CC	DIS	CONC	S	dip
22q31.1	KDEL3	disease severity	0.0184	TDT	transmitted	non-transmitted	E	hap	

	Locus	Gene	Phenotype	P	Test	Case	Reference	Model	A		
supported finding ^D	1p35	PLA2G2A	disease severity	0.0792	CC	DIS	CONC	S	dip		
			basic defect	0.0689	CC	ICM Res.	ICM no Res.	E	hap		
			basic defect	0.0619	CC	ICM DIDS Res.	ICM no Res.	E	hap		
	1p31	CLCA2Sat	basic defect	0.0138	CC	NPD GI =0	NPD GI ≤ -5	S	dip		
			CLCA1Sat	basic defect	0.0419	CC	ICM DIDS Res.	ICM no Res.	S	hap	
				basic defect	0.0301	CC	NPD Amil ≤ 22	NPD Amil ≥ 31	S	hap	
		CLCA4Sat	basic defect	0.0302	CC	NPD GI =0	NPD GI ≤ -5	S	div		
			disease severity	0.0124	CC	DIS	CONC	S	hap		
			basic defect	0.0533	CC	ICM DIDS Res.	ICM no Res.	E	hap		
	1p31.1	GSTM1	disease severity	0.0514	TDT	transmitted	non-transmitted	E	hap		
			disease severity	0.0727	CC	CON+	CON-	S	hap		
			basic defect	0.0260	CC	NPD GI =0	NPD GI < 0	E	hap		
	1q21	PDZK1	disease severity	0.0983	CC	CON+	CON-	S	dip		
			basic defect	0.1070	CC	NPD ATP = 0	NPD ATP < 0	E	dip		
	2q14.1	IL1B	disease severity	0.0263	TDT	transmitted	non-transmitted	E	hap		
			disease severity	0.0034	CC	CON+	CON-	S	hap		
			disease severity	0.0541	CC	DIS	CONC	E	hap		
			basic defect	0.0235	CC	ICM Res.	ICM no Res.	E	dip		
			basic defect	0.0294	CC	ICM DIDS Res.	ICM no Res.	S	dip		
			basic defect	0.0039	CC	NPD Amil ≤ 27	NPD Amil ≥ 28	E	dip		
			basic defect	0.0628	CC	NPD GI =0	NPD GI ≤ -5	E	dip		
			basic defect	0.0055	CC	NPD ATP = 0	NPD ATP < 0	E	hap		
			3q21	TLR9	disease severity	0.0655	TDT	transmitted	non-transmitted	E	hap
					basic defect	0.0022	CC	ICM Res.	ICM no Res.	S	hap
	basic defect	0.0088			CC	ICM DIDS Res.	ICM no Res.	E	div		
	4p13	IL8	disease severity	0.0296	CC	DIS-	DIS+	E	div		
			basic defect	0.0215	CC	NPD GI =0	NPD GI ≤ -5	S	dip		
	5q31.1	CD14	disease severity	0.0530	TDT	transmitted	non-transmitted	E	hap		
			disease severity	0.0386	CC	DIS-	DIS+	E	div		
			disease severity	0.0873	CC	DIS	CONC	S	hap		
			basic defect	0.0864	CC	ICM Res.	ICM no Res.	S	dip		
			basic defect	0.0307	CC	ICM DIDS Res.	ICM no Res.	S	div		
			basic defect	0.0144	CC	NPD ATP = 0	NPD ATP < 0	E	hap		
5q31			ADRB2	disease severity	0.0765	TDT	transmitted	non-transmitted	E	hap	
	disease severity	0.0935		CC	DIS-	DIS+	E	dip			
	disease severity	0.0039		CC	DIS	CONC	S	div			
	basic defect	0.0597		CC	NPD Amil ≤ 27	NPD Amil ≥ 28	E	hap			
	basic defect	0.0044		CC	NPD GI =0	NPD GI < 0	E	hap			

Locus	Gene	Phenotype	P	Test	Case	Reference	Model	A	
6p21.3	HLA TNF α - LT α	disease severity	0.0636	TDT	transmitted	non-transmitted	E	hap	
		disease severity	0.0081	CC	CON+	CON-	S	dip	
		basic defect	0.0010	CC	ICM Res.	ICM DIDS Res.	E	dip	
		basic defect	0.0326	CC	NPD GI =0	NPD GI < 0	E	dip	
		basic defect	0.0039	CC	NPD ATP = 0	NPD ATP < 0	E	hap	
6p21.3	HLA DQ - TAP2	disease severity	0.0317	TDT	transmitted	non-transmitted	E	hap	
		disease severity	0.0034	CC	CON+	CON-	S	dip	
		disease severity	0.0190	CC	DIS	CONC	S	dip	
		basic defect	0.0072	CC	NPD GI =0	NPD GI \leq -5	S	dip	
6q23	IFNGR1	disease severity	0.0021	CC	DIS	CONC	S	hap	
7p22	KDELR2	basic defect	0.0694	CC	NPD GI =0	NPD GI \leq -5	S	div	
		disease severity	0.0137	CC	DIS	CONC	S	div	
7q31.2	CFTR ^E	basic defect	0.0382	CC	NPD Amil \leq 22	NPD Amil \geq 31	S	hap	
		disease severity	0.0107	CC	CON+	CON-	S	hap	
		disease severity	0.0093	CC	DIS	CONC	E	dip	
		basic defect	0.0009	CC	NPD Amil \leq 22	NPD Amil \geq 31	S	hap	*
7q31.3	LEP	basic defect	0.0067	CC	NPD GI =0	NPD GI \leq -5	S	dip	
		basic defect	0.0008	CC	NPD ATP = 0	NPD ATP < 0	E	hap	*
		disease severity	0.0514	CC	CON+	CON-	S	hap	
		disease severity	0.0063	CC	DIS	CONC	S	dip	
		basic defect	0.0489	CC	ICM Res.	ICM no Res.	E	hap	
7q34 ^F	D7Sat5(2)	basic defect	0.0103	CC	NPD Amil \leq 22	NPD Amil \geq 31	S	hap	
		basic defect	0.0076	CC	NPD GI =0	NPD GI \leq -5	S	hap	
	D7Sat3	basic defect	0.0015	CC	NPD ATP = 0	NPD ATP < 0	E	hap	
		basic defect	0.0020	CC	ICM DIDS Res.	ICM no Res.	S	hap	*
	D7S495	basic defect	0.0012	CC	NPD ATP = 0	NPD ATP < 0	E	hap	
		disease severity	0.0367	CC	CON+	CON-	S	hap	
		disease severity	0.0087	CC	DIS	CONC	E	dip	
		basic defect	0.0155	CC	ICM Res.	ICM no Res.	E	div	
		basic defect	0.0316	CC	ICM DIDS Res.	ICM no Res.	S	div	
		basic defect	0.0087	CC	DIS	CONC	E	dip	
9q32	TLR4	disease severity	0.0378	TDT	transmitted	non-transmitted	E	hap	
		disease severity	0.0033	CC	DIS-	DIS+	E	dip	
		basic defect	0.0625	CC	ICM Res.	ICM no Res.	E	dip	
		basic defect	0.0581	CC	NPD Amil \leq 27	NPD Amil \geq 28	E	hap	
		basic defect	0.0140	CC	NPD GI =0	NPD GI < 0	E	dip	

Locus	Gene	Phenotype	P	Test	Case	Reference	Model	A	
10q22.3	SPA - SPD	disease severity	0.0633	TDT	transmitted	non-transmitted	E	hap	
		disease severity	0.0197	CC	DIS-	DIS+	E	div	
10q24	CD95	basic defect	0.0357	CC	NPD ATP = 0	NPD ATP < 0	E	dip	
		disease severity	0.0039 ^G	CC	CON+	CON-	S	dip	
		basic defect	0.0740	CC	ICM Res.	ICM no Res.	E	div	
		basic defect	0.0035	CC	ICM DIDS Res.	ICM no Res.	E	div	
		basic defect	0.0309	CC	NPD Amil ≤ 22	NPD Amil ≥ 31	S	hap	
		basic defect	0.0047	CC	NPD GI =0	NPD GI ≤ -5	S	dip	
		basic defect	0.0776	CC	NPD ATP = 0	NPD ATP < 0	E	hap	
11p15.5	MUC2	disease severity	0.0021	CC	DIS	CONC	S	hap	
11p15.4	SMPD1	basic defect	0.0040	CC	ICM Res.	ICM no Res.	E	dip	
		disease severity	0.0949	CC	DIS-	DIS+	E	dip	
11q13	GSTP1	basic defect	0.0786	CC	ICM Res.	ICM no Res.	E	dip	
		disease severity	0.0614	TDT	transmitted	non-transmitted	E	hap	
		disease severity	0.0433	CC	DIS	CONC	S	dip	
		basic defect	0.0938	CC	NPD GI =0	NPD GI < 0	E	dip	
12p13	TNFR	basic defect	0.0006	CC	NPD ATP = 0	NPD ATP < 0	E	hap	*
		disease severity	0.0067	TDT	transmitted	non-transmitted	E	hap	
		disease severity	0.0044	CC	CON+	CON-	S	hap	
		disease severity	0.0253	CC	DIS	CONC	S	hap	
		basic defect	0.0012	CC	ICM Res.	ICM no Res.	E	dip	
		basic defect	0.0247	CC	ICM DIDS Res.	ICM no Res.	E	hap	
		basic defect	0.0217	CC	NPD Amil ≤ 27	NPD Amil ≥ 28	E	hap	
		basic defect	0.0004	CC	NPD GI =0	NPD GI ≤ -5	S	hap	*
		basic defect	0.0003	CC	NPD ATP = 0	NPD ATP < 0	E	hap	*
		disease severity	0.0019	CC	CON+	CON-	S	hap	
12q13	KRT8	disease severity	0.0133	CC	DIS	CONC	S	dip	
		basic defect	0.0004	CC	ICM Res.	ICM no Res.	E	dip	*
		basic defect	0.0212	CC	ICM Res.	ICM DIDS Res.	S	hap	
		basic defect	0.0366	CC	ICM DIDS Res.	ICM no Res.	E	hap	
		basic defect	0.0529	CC	NPD Amil ≤ 22	NPD Amil ≥ 31	S	dip	
		basic defect	0.0181	CC	NPD GI =0	NPD GI ≤ -5	S	dip	
		basic defect	0.0888	CC	NPD ATP = 0	NPD ATP < 0	E	hap	
		disease severity	0.0959	TDT	transmitted	non-transmitted	E	hap	
16p13.3	NHERF2	disease severity	0.0589	CC	DIS	CONC	E	dip	
		basic defect	0.0001	CC	ICM Res.	ICM no Res.	E	dip	*
		basic defect	0.0123	CC	ICM DIDS Res.	ICM no Res.	S	dip	
		basic defect	0.0597	CC	NPD Amil ≤ 22	NPD Amil ≥ 31	S	hap	
		basic defect	0.0758	CC	NPD GI =0	NPD GI ≤ -5	S	dip	

Locus	Gene	Phenotype	P	Test	Case	Reference	Model	A
16p12	SCNN1B/G ^H	disease severity	0.0034	TDT	transmitted	non-transmitted	E	hap
		disease severity	0.0125	CC	DIS-	DIS+	E	hap
		disease severity	0.0046	CC	DIS	CONC	S	hap
		basic defect	0.0079	CC	ICM Res.	ICM no Res.	S	dip
		basic defect	0.0164	CC	ICM DIDS Res.	ICM no Res.	E	dip
		basic defect	0.0695	CC	NPD Amil ≤ 27	NPD Amil ≥ 28	E	hap
		basic defect	0.0024	CC	NPD GI = 0	NPD GI < 0	E	hap
		basic defect	0.0113	CC	NPD ATP = 0	NPD ATP < 0	E	hap
16p11.2	PRSS8	disease severity	0.0168	TDT	transmitted	non-transmitted	E	hap
		disease severity	0.0073	CC	CON+	CON-	S	dip
		disease severity	0.0596	CC	DIS	CONC	S	dip
		basic defect	0.1000	CC	NPD GI = 0	NPD GI ≤ -5	S	hap
17q25.1	NHERF1	basic defect	0.0080	CC	NPD ATP = 0	NPD ATP < 0	E	dip
		disease severity	0.0169	CC	CON+	CON-	S	dip
		basic defect	0.0941	CC	ICM DIDS Res.	ICM no Res.	E	hap
19q13	D19S197	basic defect	0.0810	CC	NPD GI = 0	NPD GI ≤ -5	S	hap
		disease severity	0.0746	CC	CON+	CON-	S	dip
		disease severity	0.0011	CC	DIS	CONC	S	dip
		basic defect	0.0849	CC	NPD GI = 0	NPD GI ≤ -5	S	hap

Abbreviations within above table are: CON+: concordant/mildly affected sib pair; CON-: concordant severely affected sib pair; DIS+: mildly affected sibling of a discordant sib pair; DIS-: severely affected sibling of a discordant sib pair; DIS: discordant sib pair; CONC: concordant sib pair; ICM no Res.: no residual chloride secretory response observed by intestinal current measurement (ICM); ICM Res.: cAMP activatable and DIDS-insensitive residual chloride secretion observed by ICM; ICM DIDS Res.: DIDS-sensitive residual chloride secretion observed by ICM; NPD Amil: change of nasal potential upon superfusion with amiloride [mV]; NPD GI: change of nasal potential upon superfusion with chloride-free gluconate solution and isoproterenol solution [mV]; NPD ATP: change of nasal potential upon superfusion with adenosine triphosphate solution; S: test was carried out using a stringent phenotype definition; E: test was carried out using an extended phenotype definition.

A significance retained after correction of multiple testing for 200 markers indicated by * in this column

B AAT was previously described as modifier^{S46-S50}; however the effect of these rare functional SNPs cannot be detected in this cohort (MAF < 0.05).

C isolated finding refers to an observed association with only disease severity phenotype or only basic defect phenotypes

D supported finding refers to observed associations with both, disease severity and basic defect phenotypes

E The strongest effect within this data set, represented by the trivial finding of a TDT $P < 10^{-6}$ at CFTR-spanning markers, is due to the transmission of the F508del-CFTR allele at this disease-causing locus. However, all case-reference tests are done on subgroups of F508del homozygotes and thus reflect heterogeneity of F508del alleles at this locus.

F The 7q34 region contains a paternally imprinted gene^{S16} which maps to a 2 Mb segment encompassing markers D7Sat3 and D7S495 (Stanke et al., unpublished)

G Previously published data^{S22} gives the haplotype *P* for CD95 intronic variant. In this manuscript, we show the association of diplotypes distributions at CD95 (best *P*: 0.0039; corrected for six SNPs typed at the CD95 locus $P_{corr} = 0.0347$).

H This locus contains two distinct haplotype blocks whereby the TDT is seen at SCNN1G while the association to disease severity is mediated by the neighboring gene SCNN1B.^{S11}

References within the Supplement

- S1 Eaves L, Meyer J. Locating human quantitative trait loci: guidelines for the selection of sibling pairs for genotyping. *Behav Genet* 1994;24:443-55.
- S2 Risch N, Zhang H. Extreme discordant sib pairs for mapping quantitative trait loci in humans. *Science* 1995;268:1584-89.
- S3 Dolan CV, Boomsma DI. Optimal selection of sib pairs from random samples for linkage analysis of a QTL using the EDAC test. *Behav Genet* 1998;28:197-206.
- S4 Prader A, Largo RH, Molinari L, Issler C. Physical growth of Swiss children from birth to 20 years of age. First Zurich longitudinal study of growth and development. *Helv Pediatr Acta* 1998; 52:1-125.
- S5 Knudson RJ, Lebowitz MD, Holberg CJ, Burrows B. Changes in the normal maximal expiratory flow volume curve with growth and ageing. *Am Rev Respir Dis* 1983;127:752-34.
- S6 Mekus F, Ballmann M, Bronsveld I, Bijman J, Veeze H, Tümmler B. Categories of deltaF508 homozygous cystic fibrosis twin and sibling pairs with distinct phenotypic characteristics. *Twin Res* 2000;3:277-93.
- S7 Koch C, Cuppens H, Rainisio M, et al. European Epidemiologic Registry of Cystic Fibrosis (ERCF): comparison of major disease manifestations between patients with different classes of mutations. *Pediatr Pulmonol* 2001;31:1-12.
- S8 Davis PB. Cystic fibrosis since 1938. *Am J Respir Crit Care Med* 2006;173:475-82.
- S9 Wolfenden LL, Schechter MS. Genetic and non-genetic determinants of outcomes in cystic fibrosis. *Paediatr Respir Rev* 2009;10:32-6.
- S10 Quinton HB, O'Connor GT and the Northern New England CV Cystic Fibrosis Consortium. Current Issues in Quality Improvement in Cystic Fibrosis. *Clin Chest Med* 2007;28:459-72.
- S11 Stanke F, Becker T, Cuppens H, et al. The TNFalpha receptor TNFRSF1A and genes encoding the amiloride-sensitive sodium channel ENaC as modulators in cystic fibrosis. *Hum Genet* 2006;119:331-43.
- S12 Bronsveld I, Mekus F, Bijman J, et al. Residual chloride secretion in intestinal tissue of deltaF508 homozygous twins and siblings with cystic fibrosis. The European CF Twin and Sibling Study Consortium. *Gastroenterology* 2000;119:32-40.
- S13 Bronsveld I, Mekus F, Bijman J, et al. Chloride conductance and genetic background modulate the cystic fibrosis phenotype of Delta F508 homozygous twins and siblings. *J Clin Invest* 2001;108:1705-15.

- S14 Schüler D, Sermet-Gaudelus I, Wilschanski M, et al. Basic protocol for transepithelial nasal potential difference measurements. *J Cyst Fibros* 2004;3:151-5.
- S15 De Jonge HR, Ballmann M, Veeze H, et al. Ex vivo CF diagnosis by intestinal current measurements (ICM) in small aperture, circulating Ussing chambers. *J Cyst Fibros* 2004;3:159-63.
- S16 Mekus F, Laabs U, Veeze H, Tümmler B. Genes in the vicinity of CFTR modulate the cystic fibrosis phenotype in highly concordant or discordant F508del homozygous sib pairs. *Hum Genet* 2003;112:1-11.
- S17 Drumm ML, Konstan MW, Schluchter MD, et al. Genetic modifiers of lung disease in cystic fibrosis. *N Engl J Med* 2005;353:1443-53.
- S18 Zielenski J, Corey M, Rozmahel R, et al. Detection of a cystic fibrosis modifier locus for meconium ileus on human chromosome 19q13. *Nat Genet* 1999;22:128-9.
- S19 Hillian AD, Londono D, Dunn JM, et al. Modulation of cystic fibrosis lung disease by variants in interleukin-8. *Genes Immun* 2008;9:501-8.
- S20 Cutting GR. Modifier genetics: cystic fibrosis. *Annu Rev Genomics Hum Genet* 2005;6:237-60.
- S21 Ritzka M, Stanke F, Jansen S, et al. The CLCA gene locus as a modulator of the gastrointestinal basic defect in cystic fibrosis. *Hum Genet* 2004;115:483-91.
- S22 Kumar V, Becker T, Jansen S, et al. Expression levels of FAS are regulated through an evolutionary conserved element in intron 2, which modulates cystic fibrosis disease severity. *Genes Immun* 2008;9:689-96.
- S23 Stanke F, Tümmler B, Becker T. Genetic modifiers in cystic fibrosis. *N Engl J Med* 2006;354:88-90.
- S24 Baranov VS, Ivaschenko T, Bakay B, et al. Proportion of the GSTM1 0/0 genotype in some Slavic populations and its correlation with cystic fibrosis and some multifactorial diseases. *Hum Genet* 1996;97:516-20.
- S25 Yarden J, Radojkovic D, De Boeck K, et al. Association of tumour necrosis factor alpha variants with the CF pulmonary phenotype. *Thorax* 2005;60:320-5.
- S26 Hull J, Thomson AH. Contribution of genetic factors other than CFTR to disease severity in cystic fibrosis. *Thorax* 1998;53:1018-21.
- S27 Flamant C, Henrion-Caude A, Boëlle PY et al. Glutathione-S-transferase M1, M3, P1 and T1 polymorphisms and severity of lung disease in children with cystic fibrosis. *Pharmacogenetics* 2004;14:295-301.
- S28 Henrion-Caude A, Flamant C, Roussey M et al. Liver disease in pediatric patients with cystic fibrosis is associated with glutathione S-transferase P1 polymorphism. *Hepatology* 2002;36:913-7.
- S29 Feuillet-Fieux MN, Nguyen-Khoa T, Lorient MA, et al. Glutathione S-transferases related to *P. aeruginosa* lung infection in cystic fibrosis children: preliminary study. *Clin Biochem* 2009;42:57-63.

- S30 Büscher R, Eilmes KJ, Grasemann H, et al. Beta2 adrenoceptor gene polymorphisms in cystic fibrosis lung disease. *Pharmacogenetics* 2002;12:347-53.
- S31 Hart MA, Konstan MW, Darrah RJ, et al. Beta 2 adrenergic receptor polymorphisms in cystic fibrosis. *Pediatr Pulmonol* 2005;39:544-50.
- S32 Steagall WK, Barrow BJ, Glasgow CG, et al. Beta-2-adrenergic receptor polymorphisms in cystic fibrosis. *Pharmacogenet Genomics* 2007;17:425-30.
- S33 Corvol H, Boelle PY, Brouard J, et al. Genetic variations in inflammatory mediators influence lung disease progression in cystic fibrosis. *Pediatr Pulmonol* 2008;43:1224-32.
- S34 Tesse R, Cardinale F, Santostasi T, et al. Association of interleukin-10 gene haplotypes with *Pseudomonas aeruginosa* airway colonization in cystic fibrosis. *J Cyst Fibros* 2008;7:329-32.
- S35 Yarden J, Radojkovic D, De Boeck K, et al. Polymorphisms in the mannose binding lectin gene affect the cystic fibrosis pulmonary phenotype. *J Med Genet* 2004;41:629-33.
- S36 Buranawuti K, Boyle MP, Cheng S, et al. Variants in mannose-binding lectin and tumour necrosis factor alpha affect survival in cystic fibrosis. *J Med Genet* 2007;44:209-14.
- S37 Dorfman R, Sandford A, Taylor C, et al. Complex two-gene modulation of lung disease severity in children with cystic fibrosis. *J Clin Invest* 2008;118:1040-9.
- S39 Garred P, Pressler T, Madsen HO, et al. Association of mannose-binding lectin gene heterogeneity with severity of lung disease and survival in cystic fibrosis. *J Clin Invest* 1999;104:431-7.
- S40 Gabolde M, Guilloud-Bataille M, Feingold J, Besmond C. Association of variant alleles of mannose binding lectin with severity of pulmonary disease in cystic fibrosis: cohort study. *BMJ*. 1999;319:1166-7.
- S41 Gabolde M, Hubert D, Guilloud-Bataille M, Lenaerts C, Feingold J, Besmond C. The mannose binding lectin gene influences the severity of chronic liver disease in cystic fibrosis. *J Med Genet* 2001;38:310-1.
- S42 Garred P, Pressler T, Lannig S, et al. Mannose-binding lectin (MBL) therapy in an MBL-deficient patient with severe cystic fibrosis lung disease. *Pediatr Pulmonol* 2002;33:201-7.
- S43 Davies JC, Turner MW, Klein N; London MBL CF Study Group. Impaired pulmonary status in cystic fibrosis adults with two mutated MBL-2 alleles. *Eur Respir J* 2004;24:798-804.
- S44 Olesen HV, Jensenius JC, Steffensen R, Thiel S, Schiøtz PO. The mannan-binding lectin pathway and lung disease in cystic fibrosis--disfunction of mannan-binding lectin-associated serine protease 2 (MASP-2) may be a major modifier. *Clin Immunol* 2006;121:324-31.

- S45 Martin AC, Laing IA, Zhang G, et al. CD14 C-159T and early infection with *Pseudomonas aeruginosa* in children with cystic fibrosis. *Respir Res* 2005;6:63.
- S46 Döring G, Krogh-Johansen H, Weidinger S, Høiby N. Allotypes of alpha 1-antitrypsin in patients with cystic fibrosis, homozygous and heterozygous for deltaF508. *Pediatr Pulmonol* 1994;18:3-7.
- S48 Henry MT, Cave S, Rendall J, O'Connor CM, Morgan K, FitzGerald MX, Kalsheker N. An alpha1-antitrypsin enhancer polymorphism is a genetic modifier of pulmonary outcome in cystic fibrosis. *Eur J Hum Genet* 2001;9:273-8.
- S49 Frangolias DD, Ruan J, Wilcox PJ, et al. Alpha 1-antitrypsin deficiency alleles in cystic fibrosis lung disease. *Am J Respir Cell Mol Biol* 2003;29:390-6.
- S50 Courtney JM, Plant BJ, Morgan K, et al. Association of improved pulmonary phenotype in Irish cystic fibrosis patients with a 3' enhancer polymorphism in alpha-1-antitrypsin. *Pediatr Pulmonol* 2006;41:584-91.
- S51 Choi EH, Ehrmantraut M, Foster CB, Moss J, Chanock SJ. Association of common haplotypes of surfactant protein A1 and A2 (SFTPA1 and SFTPA2) genes with severity of lung disease in cystic fibrosis. *Pediatr Pulmonol* 2006;41:255-62.
- S52 Stöhr H, Marquardt A, Nanda I, Schmid M, Weber BH. Three novel human VMD2-like genes are members of the evolutionary highly conserved RFP-TM family. *Eur J Hum Genet* 2002;10:281-4.
- S53 Shenolikar S, Weinman EJ. NHERF: targeting and trafficking membrane proteins. *Am J Physiol Renal Physiol* 2001;280:F389-95.
- S54 Razani B, Wang XB, Engelman JA, et al. Caveolin-2-deficient mice show evidence of severe pulmonary dysfunction without disruption of caveolae. *Mol Cell Biol* 2002;22:2329-44.
- S55 Teiber JF, Horke S, Haines DC, et al. Dominant role of paraoxonases in inactivation of the *Pseudomonas aeruginosa* quorum-sensing signal N-(3-oxododecanoyl)-L-homoserine lactone. *Infect Immun*. 2008;76:2512-9.
- S56 Comuzzie AG, Hixson JE, Almasy L, et al. A major quantitative trait locus determining serum leptin levels and fat mass is located on human chromosome 2. *Nat Genet* 1997;15:273-6.
- S57 Vasseur F, Meyre D, Froguel P. Adiponectin, type 2 diabetes and the metabolic syndrome: lessons from human genetic studies. *Expert Rev Mol Med* 2006;8:1-12.
- S58 Iwasaki A, Medzhitov R. Toll-like receptor control of the adaptive immune responses. *Nat Immunol* 2004;5:987-95.

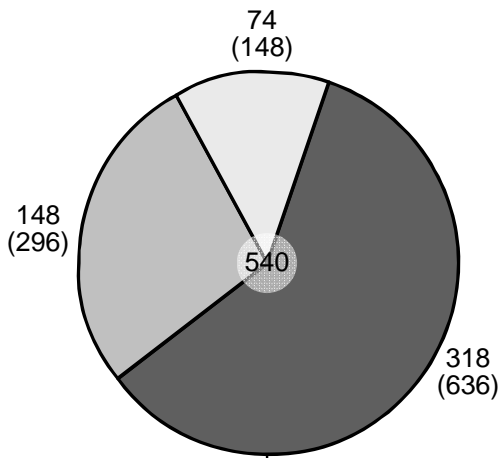
- S59 Basu S, Fenton MJ. Toll-like receptors: function and roles in lung disease. *Am J Physiol Lung Cell Mol Physiol* 2004;286:L887-92.
- S60 Urquhart DS, Allen J, Elrayess M, Fidler K, Klein N, Jaffé A. Modifier effect of the Toll-like receptor 4 D299G polymorphism in children with cystic fibrosis. *Arch Immunol Ther Exp (Warsz)* 2006;54:271-6.
- S61 Matheson MC, Ellis JA, Raven J, Walters EH, Abramson MJ. Association of IL8, CXCR2 and TNF-alpha polymorphisms and airway disease. *J Hum Genet* 2006;51:196-203.
- S62 Vastiau A, Cao L, Jaspers M, et al. Interaction of the protein phosphatase 2A with the regulatory domain of the cystic fibrosis transmembrane conductance regulator channel. *FEBS Lett* 2005;579:3392-6.
- S63 Thoreau V, Bergès T, Callebaut I, et al. Molecular cloning, expression analysis, and chromosomal localization of human syntaxin 8 (STX8). *Biochem Biophys Res Commun* 1999;257:577-83.
- S64 Teichgräber V, Ulrich M, Endlich N, et al. Ceramide accumulation mediates inflammation, cell death and infection susceptibility in cystic fibrosis. *Nat Med* 2008;14:382-91.
- S65 Ohana E, Yang D, Shcheynikov N, Muallem S. Diverse transport modes by the Solute Carrier 26 family of anion transporters. *J Physiol* 2008 Nov 17.
- S66 Vallet V, Chraïbi A, Gaeggeler HP, Horisberger JD, Rossier BC. An epithelial serine protease activates the amiloride-sensitive sodium channel. *Nature* 1997;389:607-10.
- S67 Heizmann CW, Ackermann GE, Galichet A. Pathologies involving the S100 proteins and RAGE. *Subcell Biochem* 2007;45:93-138.
- S68 Bared SM, Buechler C, Boettcher A, et al. Association of ABCA1 with syntaxin 13 and flotillin-1 and enhanced phagocytosis in tangier cells. *Mol Biol Cell* 2004;15:5399-407.
- S69 Pfeffer SR. Unsolved mysteries in membrane traffic. *Annu Rev Biochem* 2007;76:629-45.
- S70 Price MP, Thompson RJ, Eshcol JO, Wemmie JA, Benson CJ. Stomatin modulates gating of acid-sensing ion channels. *J Biol Chem* 2004;279:53886-91.
- S71 Davezac N, Tondelier D, Lipecka J, et al. Global proteomic approach unmasks involvement of keratins 8 and 18 in the delivery of cystic fibrosis transmembrane conductance regulator (CFTR)/deltaF508-CFTR to the plasma membrane. *Proteomics* 2004;4:3833-44.
- S72 Dennis EA. Phospholipase A2 in eicosanoid generation. *Am J Respir Crit Care Med* 2000;161:S32-5.
- S73 Muselet-Charlier C, Roque T, Boncoeur E, Chadelat K, Clement A, Jacquot J, Tabary O. Enhanced IL-1beta-induced IL-8 production in cystic fibrosis lung epithelial cells is dependent of both mitogen-activated protein kinases and NF-kappaB signaling. *Biochem Biophys Res Commun* 2007;357:402-7.

- S74 Stein CM, Zalwango S, Chiunda AB, et al. Linkage and association analysis of candidate genes for TB and TNFalpha cytokine expression: evidence for association with IFNGR1, IL-10, and TNF receptor 1 genes. *Hum Genet* 2007;121:663-73.
- S75 Stepkowski SM, Chen W, Ross JA, Nagy ZS, Kirken RA. STAT3: an important regulator of multiple cytokine functions. *Transplantation* 2008;85:1372-7.
- S76 Voynow JA, Gendler SJ, Rose MC. Regulation of mucin genes in chronic inflammatory airway diseases. *Am J Respir Cell Mol Biol* 2006;34:661-5.
- S77 Benson G. Tandem repeats finder: a program to analyze DNA sequences. *Nucl Acids Res* 1999;27: 573-80.
- S78 Tamiya G, Ota M, Katsuyama Y, et al. Twenty-six new polymorphic microsatellite markers around the HLA-B, -C and -E loci in the human MHC class I region. *Tissue Antigens* 1998;51:337-46.
- S79 Udalova IA, Nedospasov SA, Webb GC, Chaplin DD, Turetskaya RL. Highly informative typing of the human TNF locus using six adjacent polymorphic markers. *Genomics* 1993;16:180-6.
- S80 Waterer GW, Quasney MW, Cantor RM, Wunderink RG. Septic shock and respiratory failure in community-acquired pneumonia have different TNF polymorphism associations. *Am J Respir Crit Care Med* 2001;163:1599-604.
- S81 Wilson AG, di Giovine FS, Blakemore AI, Duff GW. Single base polymorphism in the human tumour necrosis factor alpha (TNF alpha) gene detectable by NcoI restriction of PCR product. *Hum Mol Genet* 1992;1:353.
- S82 Nomura N, Ota M, Tsuji K, Inoko H. HLA-DQB1 genotyping by a modified PCR-RFLP method combined with group-specific primers. *Tissue Antigens* 1991;38:53-9.
- S83 Fraile A, Vinasco J, Nieto A, Pareja E, Martín J. Typing for TAP alleles using PCR-RFLP with amplification-created restriction sites. *Tissue Antigens* 1996;47:346-52.
- S84 Deng GY, Muir A, Maclaren NK, She JX. Association of LMP2 and LMP7 genes within the major histocompatibility complex with insulin-dependent diabetes mellitus: population and family studies. *Am J Hum Genet* 1995;56:528-34.
- S85 Sprenger R, Schlagenhauer R, Kerb R, Bruhn C, Brockmüller J, Roots I, Brinkmann U. Characterization of the glutathione S-transferase GSTT1 deletion: discrimination of all genotypes by polymerase chain reaction indicates a trimodular genotype-phenotype correlation. *Pharmacogenetics* 2000;10:557-65.
- S86 Rowntree R, Harris A. DNA polymorphisms in potential regulatory elements of the CFTR gene alter transcription factor binding. *Hum Genet* 2002;111:66-74.

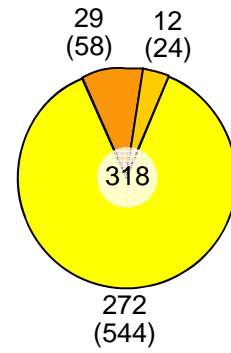
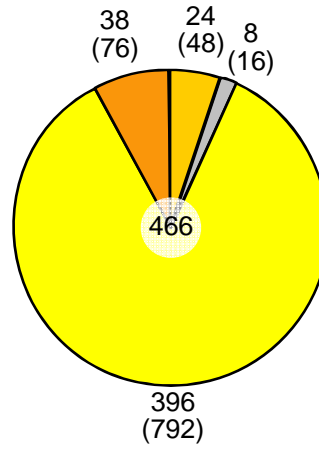
- S87 Herold C, Becker T. Genetic association analysis with FAMHAP: a major program update. *Bioinformatics* 2009;25:134-6.
- S88 Knapp M, Becker T. Family-based association analysis with tightly linked markers. *Hum Hered* 2003;56:2-9.
- S89 Becker T, Knapp M. A powerful strategy to account for multiple testing in the context of haplotype analysis. *Am J Hum Genet* 2004;75:561-70.

Statement to grant a Licence for Publication

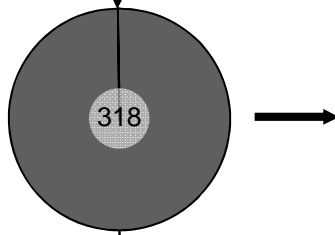
The Corresponding Author has the right to grant on behalf of all authors and does grant on behalf of all authors, an exclusive licence(or non exclusive for government employees) on a worldwide basis to the BMJ Publishing Group Ltd to permit this article (if accepted) to be published in Journal of Medical Genetics and any other BMJPGL products and sublicences such use and exploit all subsidiary rights, as set out in <http://group.bmj.com/products/journals/instructions-for-authors/licence-forms>.



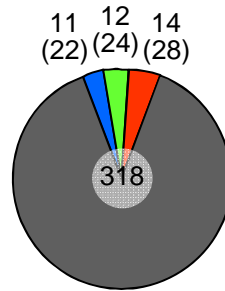
Genetic Modelling



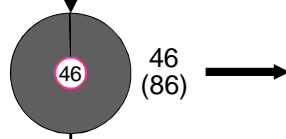
- Sibling pairs
- Monozygous twins
- Dizygous twins



Clinical Phenotype

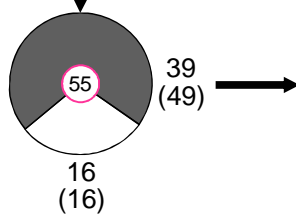
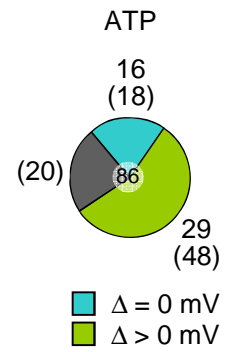
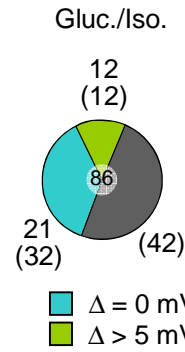
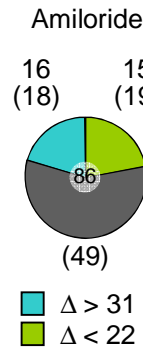


- dizygous, F508del-homozygous sibs with extreme phenotype:
- concordant / severely affected
 - concordant / mildly affected
 - discordant

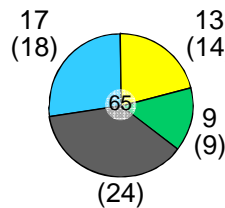


Association study

Basic Defect Phenotype

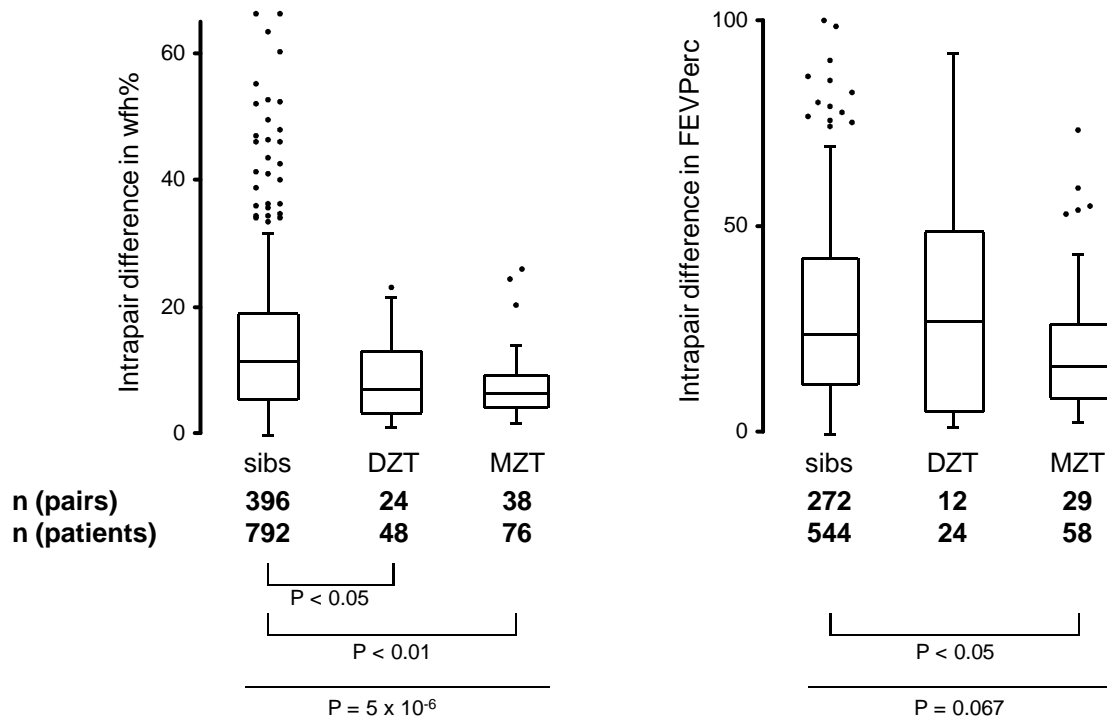


ICM



- Residual chloride secretion:
- none
 - DIDS-sensitive
 - DIDS-insensitive

A



B

