Author's response to reviews

Title: An association study on contrasting cystic fibrosis endophenotypes recognizes KRT8 but not KRT18 as a modifier of cystic fibrosis disease severity and CFTR mediated residual chloride secretion

Authors:

Frauke Stanke (mekus.frauke@mh-hannover.de)
Silke Hedtfeld (hedtfeld.silke@mh-hannover.de)
Tim Becker (becker@imbie.meb.uni-bonn.de)
Burkhard Tümmler (tuemmler.burkhard@mh-hannover.de)

Version: 2 Date: 14 February 2011

Author's response to reviews: see over
Dear Sir,

please find enclosed the revised version of MS:9170606874860806 “An association study on contrasting cystic fibrosis endophenotypes recognizes KRT8 but not KRT18 as a modifier of cystic fibrosis disease severity and CFTR mediated residual chloride secretion”. As detailed below in our point-to-point responses to the reviewers`comments, we have edited the entire manuscript and added new information to make the manuscript more self-sufficient. Complying with the editorial request, the entire text of the revised manuscript was carefully copy edited. Details on ethics/consent are now given in a separate methods paragraph “Ethics approval”. References are now formatted in the Journal’s inhouse style. As requested, all changes made in the revised version of the manuscript are highlighted in yellow.

We hope that the manuscript is now acceptable for publication in BMC Medical Genetics.

Yours sincerely,

Frauke Stanke
Point-to-point answers to reviewer`s comments:

Reviewer 1 - Major Compulsory Revisions 1. The abstract is written in unclear language and is difficult to understand.
The abstract has been thoroughly edited. Please see also our response to related Minor Essential Revisions 1-6 below for changes made in the revised version.

Reviewer 1 - Major Compulsory Revisions 2. In the background and title, the authors use the term endophenotype to describe the phenotypes they are analyzing. However there is no evidence presented that these phenotypes are indeed endophenotypes at all. And no definition for an endophenotype presented. If indeed these are endophenotypes, then this approach could be very interesting, but this indeed requires discussion and justification, or reference to literature substantiating this claim.

We concur with the reviewer that the subject on endophenotypes was not adequately presented in the submitted manuscript. In our revised version, we have added two paragraphs to the introduction to define endophenotypes and to explain why the CF basic defect, manifesting as impaired chloride conductance, is an endophenotypes and as such useful for the analysis of CF modifier genes. Text reads " As CFTR is expressed epithelial cells of many tissues throughout the body, CF is a complex multi-organ disease (Strausbaugh & Davies 2007) whereby the severity of some pathological manifestations are correlated. Patients who are pancreatic sufficient are less susceptible to chronic colonization of the airways by opportunistic pathogens such as Pseudomonas aeruginosa than pancreatic insufficient patients (Kubesch et al 2003, Green et al 2010). F508del-CFTR homozygotes who display CFTR mediated residual chloride secretion in the airways or the intestine have a milder clinical phenotype in comparison to F508del-CFTR homozygotes who do not express any chloride secretion in their tissues (Brosnved et al 2001). These two examples and other similar observations suggest that the variability of the individual disease manifestations are based on overlapping or shared causes. The corresponding concept from the field of epidemiology describes such manifestations as endophenotypes. The term endophenotypes was coined by the entomologists John and Lewis who derived the concept in the 1960s studying the geographical distributions of grasshopper populations (John & Lewis 1966). At the present time, endophenotypes are taken into account by genetic epidemiologists who analyze inherited factors that determine psychiatric diseases such as schizophrenia, depression or bipolar disorder (Gottesman & Gould 2003). Using endophenotypes is considered advantageous to the identification of causative genes because these are likely to be determined be fewer genes and less likely prone to environmental perturbations than the global disease manifestation (Gottesman & Gould 2003). Hence, as the CF basic defect is a manifestation of CFTR dysfunction at the cellular level and likely to be less complex than a global disease phenotype such as lung function, we hypothesize that measures of impaired CFTR function can be used to get insight into CF modifying genes." See also our answer to Reviewer 1 - Discretionary Revisions 3.

Reviewer 1 - Major Compulsory Revisions 3. There is reference at the end of the background to a published candidate gene study by the authors. It is difficult to assess how the current study differs from the candidate gene study referenced and why they are requiring a separate publication for this study. For example, were the same phenotypes used? Were KRT8 and 18 not candidates in that study? This is
distracting and requires further explanation as to why we should be treating this as a different study from a statistical perspective.

We have supplemented the introductory paragraph in the methods section according to the reviewer’s criticism. The text now reads: “While the entire data set of two indels, 101 SNPs and 79 microsatellites, which include the KRT8/KT18 locus as a candidate gene, has recently been published, this work gives details on KRT8 and KRT18 that have not been reported elsewhere. In particular, we now report on the results of our fine-mapping approach using a low density SNP map, assess the mode of inheritance and name the risk and the benign allele at the KRT8/KRT18 locus. These results have not been detailed in our previous work.” In addition to this information, we would like to emphasize that the detailed data given in either the figures or the tables of the present manuscript has not been displayed before in our previous publication (J Med Genet. 2011;48:24-31) which only mentioned that the locus gives valid association signals and reports Pbest values without referring to the marker combination. In summary, our recent publication in JMedGenet is focused on the global analysis of all candidate genes that were studied and does not elaborate on detailed findings for KRT8/KRT18.

Reviewer 1 - Major Compulsory Revisions 4. In the methods section, it is difficult to assess what the distribution of the phenotypes look like. Some plots and summaries would be very helpful.

In the methods section “Case-reference populations to investigate association to CF disease severity”, we have added the distribution of the clinical parameters used for ranking as suggested by the reviewer. Text now reads: “Based on the clinical disease severity of the individual siblings and the intrapair discordance, concordant mildly affected pairs (CON+; two sibs with similar and mild phenotype) and discordant severely affected pairs (CON-; two sibs with similar and severe phenotype) were defined. As described previously, severely affected sibs and mildly affected sibs differ significantly in both clinical parameters utilized by the ranking algorithm. Respectively, mean [inner quartiles; range] were as follows: for the intrapair sum of wfh% - 177 [170-187; 142-191] for concordant severely affected sib pairs and 212 [206-217; 202-229] for concordant mildly affected patient pairs; p < 0.001 (comparison carried out by Mann-Whitney rank test); for the intrapair sum of FEVPerc - 43 [18-66;9-97] for concordant severely affected sib pairs and 116 [85-143;60-190] for concordant mildly affected patient pairs; p < 0.001 (comparison carried out by Mann-Whitney rank test).” Similarly, the basic defect phenotype is now described in more detail based on the suggestion of reviewer 3 (see answer to comment 3 of reviewer 3).

Reviewer 1 - Major Compulsory Revisions 5. In the methods section we are referred to a previous publication for the details of the phenotypes. However, without understanding the phenotypes analyzed in the current study, it is difficult to follow. Therefore, a clear presentation of how the phenotypes are derived and how they are distributed is imperative to the understanding of the manuscript.

We have extended the methods section as detailed in our answer to the following comments: Reviewer 1 - Major Compulsory Revisions 4, Reviewer 1 - Major Compulsory Revisions 6, Reviewer 3 – Comment 3 and Reviewer 3 – Minor Comment 2. To avoid redundancy, kindly refer to the respective sections of this point-to-point-responses.
Reviewer 1 - Major Compulsory Revisions 6. **Methods, second paragraph, again it is difficult to follow how and why parameters needed to be combined using a ranking algorithm and how that was accomplished. Some substantial work clarifying the phenotypes is required to fully understand the implications of the results.**

We have added explanatory text on why the parameter used to describe CF disease severity was based on a composite parameter and how the composite parameter was defined. The paragraph now reads: “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 was inquired from the treating physician using a one-page questionnaire. To describe the severity of the disease with one parameter only which accounts for the severity of CF disease in the two major afflicted organ systems, i.e. the gastrointestinal and pulmonary tract, these basic clinical parameters were combined in a composite parameter using a ranking algorithm. For this purpose, weight and height were converted to weight as % of predicted weight for height (wfh%) and FEV1 values were converted to CF population centiles for FEV1 as % of predicted value (FEVPerc). Centiles of wfh% and FEVPerc were age independent in the analyzed CF patient population. Next, wfh% and all FEVPerc values for the entire study population were converted to rank numbers. The composite parameter was defined as the distance from origin in a plot of rank number for wfh% vs. rank number for FEVPerc. We have chosen this approach in order to describe overall CF disease severity assuming an equal weight of the anthropometric and the pulmonary component. In order to validate that the 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 monozygous twin pairs as long as the composite parameter was considered, while no association of intrapair concordance and twin zygosity was seen when only wfh% or only FEVPerc were examined.5 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.” These data have been published in Twin Res 2000;3:277-93, which is now included in the reference list.

Reviewer 1 - Major Compulsory Revisions 7. **The analysis section is uninformative. There is no description of what tests were employed and how exactly analysis was carried out, with the exception of a reference to the statistical analysis software used.**

The paragraph “Data evaluation” has been extended whereby the following text has been added: “Case and reference population, as specified above, were analyzed for an association of KRT8/KRT18 markers with the manifestation of CF disease severity and the CF basic defect. Genotyping data was evaluated by comparing cases and references with respect to their allele and haplotype distributions as well as genotype and diplotype distributions.” As it now refers back to two previous methods paragraphs that describe the patient populations in detail it has been moved so that the methods section of the revised manuscript now has the following headings in sequence: “Study population”, “Case-reference populations to investigate association to CF disease severity”, “Case-reference populations to investigate association to the manifestation of the CF basic defect”, “Data evaluation” and “Genotyping” (previously “Study population”, “Case-reference populations to
investigate association to CF disease severity”, “Case-reference populations to investigate association to the manifestation of the CF basic defect” , “Genotyping” and “Data evaluation”). Further information on how FAMHAP works has been added based on Reviewer 2 - Comment 6 and to Reviewer 1 - Minor Essential Revisions 11 (see below in this point-to-point responses).

Reviewer 1 - Major Compulsory Revisions 8. Page 12, top, you conclude that the causative variant responsible for the association is in KRT8, however there is not sufficient evidence to make such a conclusion. Degree of association is affected by many factors such as minor allele frequency, $r^2$ with the causal variant, etc. # 2 in discretionary revisions might help to provide additional information towards these sorts of conclusions, albeit, it is very difficult with common SNPs presumed only to be in LD with the causal variant.

Our conclusion is based on the observation at the microsatellite locus as well as on the SNP results: Typically, microsatellites detect association across longer distances than necessary to cover KRT8 and KRT18 and thus should be able to notice causative variants in both genes. However, the association is seen with KRT8 SNPs only. We have looked at the Dprime values as a measure of LD among the SNPs. Within KRT8, the range of Dprime between any two of the four SNPs is 0.901 to 1.000 (mean 0.962). SNP pairs composed of any of the four KRT8 SBNPs and any of the two KRT18 SNPs show Dprime values from 0.370 to 0.487 (mean 0.435). $R^2$ values are just as significant: Within KRT8, the range of $r^2$ between any two of the four SNPs is 0.768 to 0.986 (mean 0.864). SNP pairs composed of any of the four KRT8 SBNPs and any of the two KRT18 SNPs show $r^2$ values from 0.138 to 0.202 (mean 0.162). Thus, these two sections of the genome are on different haplotype blocks. In conclusion, we are confident that our microsatellite association signal is based on causative variants on the KRT8 haplotype block. We have changed our wording based on the reviewer’s criticism. Text now reads: “We conclude that the causative variant responsible for the allelic association with CF disease severity and the CF basic defect which we observed at KRT8Sat is likely to be found in KRT8 and not in KRT18.”

Reviewer 1 - Minor Essential Revisions 1. First sentence, abstract, “…the most frequent disease-causing lesion..,”; is lesion the correct word here? We have edited this paragraph. Text now reads: "F508del-CFTR, the most frequent disease-causing mutation among Caucasian cystic fibrosis (CF) patients...."

Reviewer 1 - Minor Essential Revisions 2. Abstract, methods” “...a low density SNP map...” you should provide information about how SNPs were chosen (i.e. as tag SNPs?) and how many were analyzed

SNPs were selected based on the polymorphism information content which is maximal for SNPs with equal allele frequencies. As all SNPs were typed using PCR-RFLP, another criterion is that the SNP must alter a recognition site of a restriction enzyme or it must be possible to create such a site using sequence modified primers for PCR. Furthermore, map position was a criterion for selection. We have edited the text in the abstract according to the reviewer’s suggestion. Text now reads: “Subsequently, a low density SNP map with four SNPs in KRT8 and two SNPs in KRT18, each selected for high polymorphism content, was used to localize the association signal.” Selection criteria beyond PIC (compatibility with PCR-RFLP and map position) are summarized in the methods section in the paragraph “Genotyping.”
Reviewer 1 - Minor Essential Revisions  3. Abstract, results section: Pbest and Pcorr were never defined  
In consistency with the comments of reviewers 2 and 3, FAMHAP is now explained in more detail in the revised manuscript (see see answer to Reviewer 2 – Minor Comment 6 for changes made in the revised manuscript). We have now defined Praw, Pbest and Pcorr in the methods section “Data evaluation”.

Reviewer 1 - Minor Essential Revisions  4. Abstract, results: no sample size reported  
As suggested, this information is now also provided in the abstract. Text reads: “We have selected contrasting F508del-CFTR homozygous patient subpopulations stratified for disease severity, comparing 13 concordant mildly affected sib pairs vs. 12 concordant severely affected sib pairs, or manifestation of the CF basic defect in intestinal epithelium, comparing 22 individuals who exhibit CFTR-mediated residual chloride secretion vs. 14 individuals who do not express any chloride secretion, for an association.”

Reviewer 1 - Minor Essential Revisions  5. Abstract, results: is it supposed to read “homozygosity of the recessive haplotype”? This statement is unclear.  
We have reworded the sentence which now reads: “Absence of chloride secretion was associated with the recessive haplotype 1122 at rs1907671, rs4300473, rs2035878 and rs2035875.”

We apologize for this non-standard terminology. As an explanation, the german terminus technicus is “dominant-rezessiver Erbgang” and obviously, it did not occur to us that this does not translate verbatim to English. We thank the reviewer for correcting this mistake (which nobody has done in the last one and a half decades). We have eliminated the non-appropriate wording and refer to the dominant allele only or the recessive allele only throughout the manuscript.

Reviewer 1 - Minor Essential Revisions  7. In the methods section, last line, it is unclear whether KRT8 and KRT18 were candidate genes in the previous published study and if so some explanation or clarification of what is different here is required.  
Text is changed according to the reviewer’s comment in the revised version as detailed in the answer to: Reviewer 1 - Major Compulsory Revisions 3.

Reviewer 1 - Minor Essential Revisions  8. Results section, first paragraph. What is Praw? I assume unadjusted for multiple hypothesis tests, but if it was one STR, why is there a multiple testing burden? How many degrees of freedom is the test?  
In consistency with the comments of reviewers 2 and 3, FAMHAP is now explained in more detail in the revised manuscript (see see answer to Reviewer 2 – Minor Comment 6 for changes made in the revised manuscript). As for Praw on KRT8Sat: all P values reported within this manuscript are derived from the permutation analysis using all weighted haplotype explanations at either the entire KRT8/KRT18 locus analysed (7-marker-haplotypes, see results in Figure 1) or restricted to KRT8 markers only (4-marker-haplotypes, see results in tables 1 and 2). After haplotype reconstruction and assignment of weighted haplotype explanations, FAMHAP reports Praw values for all marker combinations and single loci. Praw for the STR marker is derived from the analysis of the 7-marker-set. In conclusion, the reviewer is correct that testing for only the STR would not require correction for multiple testing, but as
the entire data set consists of 7 markers, we feel that Praw is the correct label for this value.

Reviewer 1 - Minor Essential Revisions 9. In results, second section, “...KRT8 markers outweigh signals...” This is not technical terminology, and it is unclear what exactly is meant by “outweigh.” We have reworded the sentence which now reads: “In both analyses, KRT8 markers provide smaller P values than KRT18 markers.”

Reviewer 1 - Minor Essential Revisions 10. Results section 2, it would be important to show that single marker analysis and two-marker haplotype analysis tagged all SNPs in the region, so the lack of significance isn’t due to lack of sufficient coverage. Please see or answer to Major Compulsory Revisions 8, which also explains why KRT18 does not need to be covered with more markers to trace the association signal of KRT8Sat.

Reviewer 1 - Minor Essential Revisions 11. Results section, page 12, first paragraph: it is unclear how many haplotypes were corrected for. Why were 4 SNP haplotypes used, or were all combinations calculated? If so how many adjustments for multiple testing are required? Some clarification on the haplotype construction would be helpful.

In consistency with comments of reviewers 2 and 3, FAMHAP is now explained in more detail in the revised manuscript (see answer to Reviewer 2 – Minor Comment 6 for changes made in the revised manuscript). All combinations using the seven markers were calculated and corrected for as FAMHAP reconstructs haplotype, using weighted haplotype explanations when non-informative phases are encountered, and does the permutation analysis based weighted haplotype explanations assigned to cases and controls. Pcorr refers to this corrected P-value.

Reviewer 1 - Minor Essential Revisions 12. Some discussion of Hardy-weinberg equilibrium would be instructive, especially at the end of page 12.

We have compared the observed diplotype frequencies in all case and reference populations analyzed with the expectancy value derived under the assumption of the Hardy-Weinberg-Law in Table 2. As can be seen from the data, two major deviations are observed for the diplotype 2211/2211 in CON+ and for the diplotype 1122/1122, both of which are overrepresented in the respective sample in comparison to the expected frequency. Hence, disturbances from HWE have been observed. As suggested by the reviewer, these observations are now stated in the manuscript text in the results section “The major contrasting KRT8 haplotypes constitute a dominant and a recessive allele which determine the manifestation of the CF basic defect”. Text now reads: “In support of these conclusions, a deviation from the expectancy frequencies derived under the condition of the Hardy Weinberg-Law are observed among mildly affected patient pairs who are more frequently homozygous for the haplotype 2211 and among patients who do not exhibit any chloride secretion who are more frequently homozygous for the allele 1122 (Table 2).”

Reviewer 1 - Minor Essential Revisions 13. Top of page 13, again the terminology of a dominant-recessive mode of inheritance is confusing and requires change.

We have edited this paragraph. Text now reads: “We have observed that the CF basic defect is determined via KRT8 whereby the phylogenetically older KRT8 haplotype 2211 is dominant for the manifestation of CFTR mediated residual chloride...”
secretion and also recognized consistently as the benign modifier allele that is associated with a milder clinical phenotype among concordant CF sibpairs (Table 2, Figure 2). Accordingly, the phylogenetically younger contrasting haplotype 1122 constitutes the recessive risk allele (Table 2, Figure 2)."

Reviewer 1 - Minor Essential Revisions 14. In the analysis of haplotype age on page 13, why were two-marker haplotypes chosen? Why were the haplotypes chosen in this way, or were more markers considered as well. The haplotypes are exclusively STRP-SNP pairs because we wanted to know which of the two SNP alleles is older and whether or not the alleles marked as older will segregate with one of the two frequent haplotypes that we have seen in our population (as it is the case). We did not consider other marker combinations and we did not analyze more markers or higher marker combinations.

Reviewer 1 - Minor Essential Revisions 15. Conclusions section, first sentence “…KRT8 haplotypes that determine the manifestation…” Determine is a very strong word and the data presented do not back up this claim. We have changed the text accordingly to “contrasting KRT8 haplotypes that influence the manifestation of the CF basic defect”.

Reviewer 1 - Discretionary Revisions 1. Genotyping section/results section it would be important to have some sort of figure depicting the degree of linkage disequilibrium. The information requested by the reviewer is now provided in the text of the results section “Intragenic SNPs in KRT8 allocate the association signal to CF basic defect and disease severity”. Text reads: “Within KRT8, linkage disequilibrium is strong with D’ between any two of the four SNPs ranging from 0.901 to 1.000 (mean 0.962). SNP pairs composed of any of the four KRT8 SBNPs and any of the two KRT18 SNPs show D’ values from 0.370 to 0.487 (mean 0.435). Thus, the two KRT genes reside on different haplotype blocks of the genome.”

Reviewer 1 - Discretionary Revisions 2. Results, top of page 12, it would be interesting to see what is in high LD with your SNPs that were not already typed to assist with localization and determine if there are any functional SNPs in LD with your associated SNPs. We fully agree with the reviewer. On other occasions, we have been able to map the modifier by the base by comparative resequencing analysis of patients who carry contrasting haplotypes and exhibit contrasting phenotypes (Method is validated in: 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). However, as stated in the discussion section, the haplotype block encompassing KRT8 is at least 13 kb in size, hence we expect that cases and reference haplotypes differ at many SNPs. The area of interest involves several kb of intergenic sequence with unknown function which will be difficult to annotate without further information on the molecular mechanism that determines the role of KRT8 as a CF modifier. We hope that, even though the genomic sequence of the entire LD block is not analyzed yet, our data will be of interest to the scientific community for a replication study at the present stage of the project.

Reviewer 1 - Discretionary Revisions 3. The last sentence of the conclusions on the “endophenotype” being a superior measure for modifier studies is an important and
interesting one. This should be played stronger throughout the manuscript, with greater emphasis as one of the theses of this paper. Of course, presuming that one can back up the fact that this is indeed an endophenotype, and if not, that it is a heritable trait with less environmental influence.

We thank the reviewer for this kind assessment. As already outlined in our answer to Reviewer 1 - Major Compulsory Revisions 2, the data supports the concept that the CF basic defect is indeed an endophenotype.

Reviewer 2 Comment 1: The selection criteria for cases and references of the association study, e.g. were 101 CF families consist of both siblings and twins?

The information on the composition of the study panel of 101 CF families is now provided as suggested. Text in the methods section “Study population” of the revised manuscript reads:” The study population and the selection criteria for cases and references of the association study has been described in detail elsewhere.5 Briefly, genotyping data from 101 CF families, 85 of which are a subgroup of the twin and sibling study panel of 466 twin and sibling pairs, was used for the association study.5 16 F508del-CFTR homozygous singletons with known basic defect and their parents were included into the analysis of the manifestation of the CF basic defect. These 16 patients were recruited from the local CF clinic at Hannover Medical School for a study on the manifestation of the basic defect in excised intestinal biopsies and subsequent CFTR protein analysis and chip-based transcriptome analysis.”

Reviewer 2 Comment 2: The composite parameter and the ranking algorithm used for disease severity phenotype : How CON+ and CON- groups are exactly defined, i.e. what is the intra-pair discordance measure? How large? Is it somewhat arbitrary? (Also see below for further comments on this point.) While the proposed statistical approach provided a positive result, I do wonder about the power as well as the validity of the method.

In accordance with the reviewer’s comment, and also based on the suggestion of reviewer 1, we have extended the paragraph on the description of the clinical phenotype. The revised manuscript now provides the information on how the composite parameter was defined and the distribution of the anthropometry and lung function values in the study population. As for the power and the validity of our method, we would like to emphasize that the composite parameter utilized by the ranking algorithm recognizes monozygous twins to be more concordant than dizygous twins in our twin sample of 29 monozygous pairs and 12 dizygous pairs. (Twin Research 2000: 277-293). This higher degree of concordance in MZT is not seen in the isolated anthropometry or lung function parameter. We conclude that our composite parameter employed to categorize the twin and sib pairs capturers an inherited component of CF disease and as such is most valuable to study modifying genes in CF. To avoid repetition, kindly refer to the answer to Reviewer 1 - Major compulsory revision 4 & 6 for the complete text that was changed in the revised manuscript.

Reviewer 2 Comment 3: Genotyping data are available for 101 CF families, yet only 13 CON+ and 12 CON- pairs are used in the analysis for CF disease severity (similar comments for the other phenotype of interest). Given the analytical approach used in the paper, what is the expected power for different range of genetic effect? Is KRT18 negative mainly due to insufficient power?

We have characterized informative patient cohorts and have conducted an association study on these patient subsamples as extreme phenotypes are generally
considered to be more informative in analyzing quantitative traits (Eaves L, Meyer J. Locating human quantitative trait loci: guidelines for the selection of sibling pairs for genotyping. Behav Genet 1994;24:443-55. Risch N, Zhang H. Extreme discordant sib pairs for mapping quantitative trait loci in humans. Science 1995;268:1584-89. 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.) Given the size of the genotyped sample, it is unlikely that we could detect rare variants with small effects on the modifying genes. However, major modifiers who have a large impact on the clinical phenotype – which we aim to find - should not escape our notice. As for KRT18: if there are variants in KRT18 that modify CF, these are a) not mirrored in KRT8Sat b) do not manifest in the impaired ion conductance of the nasal or the intestinal epithelium and c) are likely of minor importance to the clinical phenotype in CF.

Reviewer 2 Comment 4: If phenotype is defined on a pair, how do authors resolve the issue of potential genotype/haplotype inconsistency between a sib pair?
FAMHAP allows association tests of affected sib pairs vs unaffected sib pairs or a reference population of sib pairs. The software accounts for within sibship dependency of diplotypic status in regions of linkage. MC permutation is carried out for the entire sibship, i.e. affection status is permuted or not with equal probability for both siblings. As haplotypes of all individuals are reconstructed prior to association testing and , in cases of non-informative phase or haplotype uncertainty, weighted haplotype explanation lists are assigned to each individual whereby the haplotype frequencies of the entire data set are taken into account to compute the conditional likelihood weights, the haplotypes of both sibs are allowed to be different during the association test. This information is also provided in the text of the revised version (see answer to minor comment 6).

Reviewer 2 Comment 5: Discordant pairs also provide association evidence. I believe all individuals can be used in the association analysis using a regression framework but with correct covariance matrix to account for the relatedness between the samples. This also avoids the problem of requiring a somewhat arbitrary cutoff point to define concordance.
All individuals can be used in the association analysis, without the necessity for a regression framework or other similarly thorny methods. We agree with the reviewer that discordant pairs are very useful to detect modifiers and have implemented discordant pairs in our strategy as follows: 1. Intrapair comparison whereby mild and severe sib are compared (algorithm implemented in FAMHAP accounting for sibship dependency) and 2. Interpair comparison of all concordant pairs to all discordant pairs. The latter will detect gene-gene interactions and modifiers acting in trans to the analysed genomic area (see CEACAM3 / Hum Genet. 2010;127:383-394; SCNN1B / Hum Genet. 2006;119:331-343; CFTR linkage group / Hum Genet. 2003;112:1-11).

Reviewer 2 Minor Comment 6: Add some material on haplotype inference, e.g. which program was used, how to deal with haplotype uncertainty etc.
As suggested by the reviewer, we have provided more information on FAMHAP in the revised version. The text of the revised methods section “Data evaluation” now reads: “Genetic data for the association study was evaluated using the FAMHAP software package which allows family-based analysis and accepts data evaluation in association studies on unrelated individuals as well as on affected sib pairs. Case and reference population, as specified above, were analysed for an association of
KRT8/KRT18 markers with the manifestation of CF disease severity and the manifestation of impaired ion conductance. Genotyping data was evaluated by comparing cases and references with respect to allele and haplotype as well as genotype and diplotype distributions. All case-reference comparisons were carried out using 10,000 Monte-Carlo simulated data sets. The analysis of more than one marker per locus is corrected for multiple testing by haplotype permutation. For this purpose, the entire data set of cases and references is used to estimate haplotype frequencies. Haplotype, or, in cases of non-informative phase or haplotype uncertainty, weighted haplotype explanation lists are assigned to each individual whereby the haplotype frequencies of the entire data set are taken into account to compute the conditional likelihood weights. Permutation is done by randomly assigning the affection status to the individuals in each replication whereby the ratio of cases and controls is kept constant. For the comparison of case sib pairs to reference sib pairs, the affection status is permuted or not with equal chance for both siblings simultaneously. P-values for comparison of n-marker-haplotype and all marker subsets derived thereof are computed as s/n, where n is the number of permutation replicates, and where s is the number of permutation replicates leading to a test statistic higher than or equal to that of the real data. Similarly, diplotype distributions between cases and controls are compared whereby a diplotype is a haplotype pair of the individual. Reported P values are: Praw, referring to a computed P value of a single marker or a marker subset, Pbest, referring to the best observed Praw value, and Pcorr, referring to the P value of the entire marker set that is corrected for multiple testing. The adjustment for multiple testing properly accounts for LD within the Monte-Carlo simulation framework that evaluates the corrected significance minP, the smallest observed raw P-value. The computational details of the minP principle have been described in: Becker T & Knapp M A powerful strategy to account for multiple testing in the context of haplotype analysis. Am J Hum Genet 75:561-70. To allow a comparable assignment of weighted haplotype explanations in all subpopulations, the entire genotyping data of 101 CF families were provided as training set to FAMHAP for all case-reference comparisons.”

Reviewer 2 Minor Comment 7: Page 9, last line, change 10.000 to 10,000
Done as suggested.

Reviewer 2 Minor Comment 8: Page 24, for ICM no Res. vs ICM CFTR Res. why Pcorr < Praw?
Table 2 provides Praw values for the 4-marker-haplotype. The corrected P value takes all possible marker subsets into account, some of which give lower P values than the 4-marker-haplotype (see figure 1C, 4-MHap shown in yellow, 3MHap in orange and 2MHap in red). We have chosen not to interpret this further as LD is strong between the KRT8 markers and all P values are in the same order of magnitude.

Reviewer 3 – Comment 1) The authors do not fully introduce/consider other factors that have been reported to promote trafficking and partially restore del508 function at the plasma membrane such as over-expression, reduced temperature, compounds that affect the folding environment, and other mutations in cis with del508 (I539T, R553Q, and others). I expected some discussion of the other mutations and why their physiological effect may be (or may not be, as reported by members of your group previously) important in this analysis. It appears that other genotype data are
available for these samples. It would be important to know if any of the potential cis acting CFTR modifying mutations are present among this group.

The samples were not investigated for I539T or R553Q. We consider both mutations found in cis with F508del too rare to have an impact on a sample of the size of our study population. However, we have analyzed the CFTR locus as a potential modifier on the manifestation of the basic defect using intragenic markers. We did not see any influence of the CFTR intragenic background on the ICM derived phenotypes. In contrast, with three endophenotypes based on nasal potential difference measurements, we could detect an association signal as described in our previous publication (J Med Genet. 2011;48:24-31). As this manuscript is focused on ICM and not on NPD, we did not discuss these findings.

Reviewer 3 – Comment 2) Along the lines of the point above, have you looked at KRT8 and ion conductance in severe CFTR mutations other than del508? What about other Class II mutations? Do you have access to another population of patients you could attempt to replicate this result in?

During the recruitment phase, we have concentrated on F508del-CFTR homozygotes to eliminate the CFTR mutation genotype as a confounding variable when looking for modifying genes. Furthermore, we have exclusively analyzed F508del-CFTR homozygotes as this is the only CFTR mutation that is frequent enough to study a patient group who is homogeneous with respect to the disease-causing gene. Hence, we currently do not have access to another patient population with class II mutations.

Reviewer 3 – Comment 3) The introduction could benefit from a clearer explanation of functional consequences of delF508 and the variation of CFTR activity found in delF508 homozygotes. For example, do the rectal suction biopsies range from 0% function to 30% function? 50%? How much of a clinical impact does a CFTR-del508 protein functioning in the epithelial membrane have? I believe having a more detailed sense of the variation in CFTR del-508 ion conductance will add to the understanding and clinical relevance of the results.

The introduction and the methods section “Case-reference populations to investigate association to the manifestation of the CF basic defect” has been extended to provide the information as requested. In the introduction, the information on the clinical relevance of CFTR mediated residual function among F508del homozygotes is given. Text reads: "F508del-CFTR homozygotes who display CFTR mediated residual chloride secretion in the airways or the intestine have a milder clinical phenotype in comparison to F508del-CFTR homozygotes who do not express any chloride secretion in their tissues." In the methods section “Case-reference populations to investigate association to the manifestation of the CF basic defect”, the following information has been added to explain the variation of CFTR activity in F508del-CFTR homozygotes: "When analysed by ICM, non-CF controls display upon stimulation by carbachol or histamine short circuit currents (Isc) that exclusively monitor the Cl⁻ secretory response mediated by CFTR, because the obligatory K⁺ efflux response response is not visible in the presence of fully functional CFTR. For the diagnosis of CF, both the direction and the magnitude of Isc is used. Typically, CF patients display a biphasic response to carbachol and histamine as the fast K⁺ current and the slower Cl⁻ current, both visible in classical CF cases, show opposite directions. Responses range from 0% Cl⁻ response (equivalent to 100% K⁺ signal) to 100% Cl⁻ response (equivalent to 0% K⁺ signal) among F508del-CFTR homozygotes. In the absence of a K⁺ signal, a Cl⁻ secretory response below 10µA/cm² is taken as an indicator of CFTR dysfunction and is consistent with CF. The judgment of whether
or not a recording shows evidence of residual CFTR activity is based on the relative proportions of the $K^+$ and the Cl$^-$ signals in the biphasic response to carbachol and to histamine after the application of DIDS. No evidence of a chloride secretory current values in both cases is interpreted as absence of CFTR function. Residual Cl$^-$ secretion that is mediated by CFTR is assumed if both, carbachol and histamine response after the incubation with DIDS show a proportion of at least 40% of Cl$^-$ secretory response (equivalent to 60% or less of K+ response) in the presence of a cAMP-mediated Cl$^-$ secretory response."

Reviewer 3 – Comment 4) The design of this project is not clear. While the methods describe a twin/sibling sample, it is unclear how these were used for the association analyses presented. My interpretation of the description is that all concordant severe sib pairs were compared to all concordant mild sib pairs for the severity phenotype. However, there is no mention of what specific test was done or how the sibling correlations within each group were accommodated in the testing. In consistency with comments of reviewers 1 and 3, FAMHAP is now explained in more detail in the revised manuscript (see answer to Reviewer 2 – Minor Comment 6 for changes made in the revised manuscript). The revised methods paragraph “Data evaluation” now provides the information how the sibling correlations were accommodated. Patient populations for the association study are explained in the methods sections “Case-reference populations to investigate association to CF disease severity” and “Case-reference populations to investigate association to the manifestation of the CF basic defect”. In summary, the analysis of association to disease severity is based on a sib pair sample while the analysis of the basic defect is based on unrelated patients, which are derived from sib pairs using an index case strategy to ensure that cases and references are not related for this analysis.

Reviewer 3 – Comment 5) It is unclear why the results for the microsatellite analysis (first paragraph of results) are presented in allele comparisons for one phenotype but genotype comparison for the other. Were the results not significant for the genotype comparison?

The entire manuscript is based on a comprehensive data evaluation as documented in our previous publication (J Med Genet. 2011;48:24-31), whereby every locus is investigated for an association of haplotype distribution, genotype distribution or an accumulation of rare variants among cases and controls. In this manuscript, we have done a follow-up analysis on the best $P$ values at the KRT8/KRT18 locus, which are $P_{raw}=0.0409$ (allele distributions, CON+ vs CON-)and $P_{raw} = 0.0177$ (genotype distributions, ICM) for the microsatellite. After fine-mapping, we have checked both modes of evaluation for the the 4-SNP-haplotype in KRT8. For the CON+/CON- contrast, $P$ is lower by one order of magnitude for all allele distribution. For the ICM contrast, both $P$ values are in the same order of magnitude. The corresponding data is displayed in table 1 and table 2.

Reviewer 3 – Minor Comment 1) The term “CF basic defect” is used often in the text. It appears that the use is intended to refer to the mutation, but at times is appears that you specifically mean the physiologic consequence. This can be confusing at times in the presentation. I would suggest stating “impaired ion conductance” in some instances when explaining data or a process to make the idea you are trying to express more clear to the reader.

We have edited the manuscript accordingly.
Reviewer 3 – Minor Comment 2) In your methods, please also cite the Mekus paper (Twin Res. 2000 Dec;3(4):277-93) when referencing the ranking algorithm.
Done as suggested.

Reviewer 3 – Minor Comment 3) Page 12, line 4 worded like KRT8 is the cause of the CF basic defect, which it is not. I would recommend “the modification of the CF basic defect in CFTR-del508 homozygotes.”
Done as suggested.

Reviewer 3 – Minor Comment 4) The phylogenetic age of the allele section seems irrelevant to the results reported. If it is important to the interpretation of this work, this should be more clearly discussed.
We have now added a paragraph to the discussion section on the relative age of the two contrasting haplotypes. Text reads: “As the phylogenetically older and dominant allele conveys CFTR mediated chloride secretion, indicative of proper targeting of CFTR to the epithelial apical membrane, we conclude that this condition has provided a selective advantage to human beings. Consistently, the younger recessive risk allele manifests only in the homozygous state.”

Reviewer 3 – Minor Comment 5) Pg 15, last line: I would suggest the use of “may be” instead of “apparently” since there is no biological data to back up this statement. Have you thought about trying to show the role of the KRT8 haplotypes and their effect on CFTR-del508 molecules in cell culture?
The wording in the concluding paragraph of the discussion section was changed as suggested. We agree with the reviewer that functional analysis – in cell culture or by other means – would be essential to clarify the underlying mechanism. However, our analysis of KRT8 as a CF modifier at the present stage has identified two contrasting haplotypes which are likely to result in functional consequences. We cannot name the causative variant(s) in KRT8. As detailed in the discussion section of the manuscript, the haplotype block that would need to be characterized by comparative sequence analysis is larger than 13 kb and likely to yield more than one candidate SNP to analyze in any follow-up work whether it is causal or not. While this is beyond the scope of our present work, we were hoping that experts in KRT biology might be interested in genotyping their resources to unravel the mechanism from their end.

Reviewer 3 – Minor Comment 6) Figure 2: A. Is there a reason why you did not combine axis labels with the overall microsatellite graph? B. Show microsatellite distribution by haplotype (1122, 2211, and other)
Axis labels are given in the separate plot only to avoid redundancy in the figure – the bar diagrams would show nine times the same labels otherwise. As for the combination of the individual SNPs to haplotypes: the purpose of this figure is to see which allele at the individual SNP loci (allele 1 or 2) is older. That the older alleles line up to form one haplotype and the younger alleles line up to form the contrasting haplotype is not obligatory any recent recombination event could have changed that. Hence, to show the microsatellite distribution by haplotype and not by individual SNP would have drawn the (in the case of KRT8 correct) conclusion in advance.

Reviewer 3 – Minor Comment 7) Table 1: Change 0.000 to <0.001
Done as suggested.