Reviewer's report

Title: Polymorphisms in the glutathione pathway modulate cystic fibrosis severity: a cross-sectional study

Version: 5 Date: 11 June 2013

Reviewer: Frauke Stanke

Reviewer's report:

MS „Polymorphisms in the glutathione pathway modulate cystic fibrosis severity: a cross-sectional study“ by de Lima Marson and colleagues

For clarification: The reviewer understands that this is a resubmission of the data previously submitted to The Journal in a manuscript entitled “Genetic interaction of GSH metabolic pathway genes in Cystic Fibrosis”.

In their current manuscript “Polymorphisms in the glutathione pathway modulate cystic fibrosis severity: across-sectional study” de Lima Marson and colleagues report on the role of polymorphisms in the glutathione metabolism genes GCLC and GSTP1 as well as large genomic deletions in GSTM1 and GSTT1 on the manifestation of cystic fibrosis disease. The authors report on 180 CF patients who have been characterized for these candidate modifier genes and the CFTR mutation genotype. Association to 28 clinical phenotypes was tested for each of these clinical phenotypes in this cohort.

The manuscript is well organized and the results are presented very clearly, summarising the findings on each genetic variation considered for each of the clinical phenotypes studied in a separate supplementary table. The data presented by de Lima Marson and colleagues is highly relevant for those who analyse cystic fibrosis modifiers as it shows convincingly that the GST metabolic pathway influences the CF phenotype on several levels. However, at the first glance the data is confusing as the association signals do not appear consistently in all CF subgroups and the reader does not get help in understanding this in the present manuscript. Moreover, some association results substantiate each other and the reader does not get help to realize this in the present manuscript. Hence, the findings (which are currently listed comprehensively and discussed in context with other association studies that target the same candidate genes) need to be discussed more extensively in context with each other and the population specifics.

Major compulsory revisions:

1. The authors group their patients prior to analysis with respect to the number of resolved CFTR mutations. Does the frequency of pancreatic sufficient patients vary between these patient subsets? This might help to explain the findings of an
association signal in one patient subgroups only based on an underlying survivor effect and/or selection bias. Example: Table 6, GCLC-129C>T / association signal with phenotype P aeruginosa mucoid, observed only in the subgroup “one CFTR mutation identified”. The reader might be puzzled: If this GCLC polymorphism modifies CF disease, why is this signal absent among those patients with two identified mutations who are typical CF patients? What makes these subgroups heterogeneous? There are currently three possible explanations:

1A. About the only phenotype that can be reliably predicted based on the CFTR mutation genotype is the pancreatic status, given rise to the “rule” that those patients who do carry one mild CFTR mutation (class IV and V) are pancreatic sufficient. Table 5 states that 80% of patients in this study are PI, consequently 20% are PS. This is a high prevalence of PS in comparison to the Caucasian CF populations (incidence of PS genotypes 5 – 10%) and might reflect the high incidence of mild CFTR mutations that cause pancreatic sufficiency in the Brazilian CF population. Alternatively, it might reflect that some patients are carriers of CF mutations who suffer from another disease but cystic fibrosis and not CF patients. These misdiagnosed individuals would then be pooled in the groups “one CFTR mutation identified” or “no CFTR mutation identified” and be absent from the group “both CFTR mutations identified”. Consequently, the groups defined by the authors would be highly heterogeneous and the lack of the association signal in the group “both CFTR mutations identified” would point to the fact that this modifier has no impact on classical CF, but might play a role on atypical CF / CF-like disease and other misclassified phenotypes.

1B. If CFTR transports GSH the finding (-129C>T / association signal with phenotype P aeruginosa mucoid, observed only in the subgroup “one CFTR mutation identified”) might reflect that the GSH pathway modifies CF in a subgroup of patients who display CFTR mediated residual function (class IV and V mutations on one CF chromosome, phenotype PS), presuming that these individuals are more prevalent in the group “one CFTR mutation identified” than “two CFTR mutations identified”. It would also point out that CFTR can indeed mediate GSH transport, at least in some tissues as residual CFTR activity is likely too lead to higher GSH transport capabilities of the mutant CFTR protein.

1C. The patients were recruited cross-sectionally and hence, mild CFTR mutations might be overrepresented in this patient panel (indicating that patients who carry a severe CFTR mutation genotype have died with a higher likelihood than patients with mild CFTR mutations prior to being recruited for this study). This would also explain convincingly why the prevalence of PS patients is high in this patient panel. Again, this would then lead to an asymmetric distribution of PS patients in the subgroups stratified for CFTR mutation genotype.

The authors have access to the data to discriminate between 1A, 1B and 1C:

1A – review clinical charts for patients without identified CFTR mutations. Misdiagnosis with a sweat test value > 60 mval/l occurs, but hardly if the sweat test is taken repeatedly. Causes for fals-positive sweat tests are: atopic
dermatitis and malnutrition and several other conditions (e.g. listed in: O'Sullivan BP, Freedman SD. Cystic fibrosis. Lancet. 2009 May 30;373(9678):1891-904. doi: 10.1016/S0140-6736(09)60327-5.)

1B – count prevalence of identified class IV / class V mutations in subgroups “both CFTR mutations identified” versus “only one CFTR mutation identified” versus “no CFTR mutation identified”. Likewise, assuming that some PS CFTR mutations have not been identified, count the prevalence of PS phenotype in subgroups “both CFTR mutations identified” versus “only one CFTR mutation identified” versus “no CFTR mutation identified”

1C – rank patients according to date of birth. If PS patients are more frequently among early-born individuals, this indicates a survivor effect in the population. Likewise, if patients with two unresolved CFTR mutations and/or patients with one resolved CFTR mutation and / or patients with two identified CFTR mutations cluster by birth cohort, this indicates a survivor effect in the population.

To clarify the reviewer’s assessment: the data is, according to the reviewer’s opinion, valid and true, i.e. there is no flaw in the analysis. However, the reviewer urges the authors to evaluate their valuable data pool comprehensively to elucidate why the association signals are observed in one patient subgroup only. This is likely best done in a separate discussion paragraph. Apart from the aforementioned finding (GCLC-129C>T / association signal with phenotype P aeruginosa mucoid, observed only in the subgroup “one CFTR mutation identified”), the other findings for which a similar discussion of elevated frequency of PS cystic fibrosis is necessary are:

GCLC-3506A>G / P. aeruginosa no mucoid / One identified CFTR mutation
GSTM1 deletion polymorphism / SpO2 / No CFTR mutations identified
GSTT1 deletion polymorphism / P. aeruginosa no mucoid / No CFTR mutations identified
GSTM1/GSTT1 deletion polymorphism / Bhalla score / No CFTR mutations identified
And complementary for
GSTP1+313A>G / Osteoporosis / Two identified CFTR mutations
for which the prevalence of osteoporosis in the other two patient subgroups (One identified CFTR mutation, No CFTR mutations identified) might differ and this difference in prevalence might explain the result.

2. The interrelationship of the association findings is not clearly discussed in the present manuscript. For instance, finding of an association for two different phenotypes that describe the same pathology substantiates such a result considerably: Two different GCLC variants are associated with two different P. aeruginosa phenotypes (albeit in the same patient subgroup “one identified CFTR mutation only”). This appears to be confirmatory of the interrelationship of GCLC and P. aeruginosa colonization in CF? Can the authors construct haplotypes of these two SNPs and deduce which haplotype one is protective against P.
aeruginosa and which one promotes chronic colonization?

Minor points:

3. Page 8, text reads: “adjusted by the Bonferroni correction (# corrected = 0.05/number of tests” – the tables imply that alpha observed was divided by 4 to display the corrected P-value? Please also state the rationale for the number 4 (as some readers might expect correction for 28 clinical variables, or worse, for 4 X 28 tests).

4. Page 8, text reads “Data distribution showing a high standard deviation was analyzed by the median value.” This is not interpretable- was the primary data transformed? If so, how? Please explain differently.

5. page 10, text reads “The COPD pathophysiology is similar to CF in that it involves cellular responses, inflammatory mediators, and oxidative stress” There are clearly many observations that point to differences in COPD and CF. Please reword to avoid that readers who do not have a profound knowledge of the pathophysiology of CF and COPD (such as pure geneticists with limited clinical expertise) can confuse these two different diseases in the future.

6. page 10, text reads: “little is known about GST genes and CF severity.” “Little” is unspecific – of course, the GST genes deserve a more thorough look as a modifier, but the analysis of nearly 2000 CF patients for the role of the candidate modifier genes GSTM1 does not qualify as “little”. Please reword.

7. page 11: text reads “previously investigated in 1,940 children (aged 8–11 years).” Please quote the population & reference 37 already here. Also: as this is a very large study, do the results presented within this manuscript identify agree with this previously published work in an entirely different study population with respect to the assignment of the risk allele at GSTM1, GSTT1 and GSTP1?

8. page 11 and following: PAM and PAMN are defined in the manuscript, but cryptic to read within the text. Please avoid these abbreviations and use full-text instead.

9. page 11, text reads “higher frequency of the PAM to CC genotype” please specify the SNP.

10. page 12, text reads “the GSTP1+313A>G polymorphism was associated with a low risk of osteoporosis (p=0.036; with two CFTR mutations identified) as a protective factor and with young age # 154 months (p=0.044; without taking the CFTR gene into account) as a risk factor.” This needs to be edited – the A allele is protective against osteoporosis? The A allele is increased among young patients with unresolved CFTR mutation genotype?

11. page 12, text reads “The presence of osteoporosis is influenced by several different factors, including mutations in the CFTR gene, the environment, modifier genes, and increased life expectancy [39].” Please add a comment here with respect to the observations of the osteoporosis incidence in your patient subgroups (see above, comment 1: “ ……. GSTP1+313A>G / Osteoporosis / Two identified CFTR mutations
for which the prevalence of osteoporosis in the other two patient subgroups (One identified CFTR mutation, No CFTR mutations identified) might differ and this difference in prevalence might explain the result.” Any survivor bias in this population will emphasise the effect of a modifier on a condition associated with longevity.

12. page 12, text reads: “recent discoveries indicate that CFTR modulates the transport of GSH, creating a dysfunction in the antioxidant defense.” Please give references to substantiate this claim.

Level of interest: An article of importance in its field

Quality of written English: Acceptable

Statistical review: Yes, and I have assessed the statistics in my report.

Declaration of competing interests:

I declare that I have no competing interests.