Author's response to reviews

Title: Polymorphisms in the mitochondrial oxidative phosphorylation chain genes as prognostic markers for colorectal cancer

Authors:

Jesus Lascorz (j.lascorz@dkfz.de)
Melanie Bevier (m.bevier@dkfz.de)
Witigo v. Schönfels (witigo.v.schoenfels@uksh-kiel.de)
Holger Kalthoff (hkalthoff@email.uni-kiel.de)
Heiko Aselmann (heiko.aselmann@uksh-kiel.de)
Jan Beckmann (jan.beckmann@uksh-kiel.de)
Jan Egberts (jegberts@chirurgie-sh.de)
Stephan Buch (s.buch@ikmb.uni-kiel.de)
Thomas Becker (thomas.becker@uksh-kiel.de)
Stefan Schreiber (s.schreiber@mucosa.de)
Jochen Hampe (jhampe@1med.uni-kiel.de)
Kari Hemminki (k.hemminki@dkfz.de)
Asta Försti (a.foersti@dkfz.de)
Clemens Schafmayer (clemens.schafmayer@uksh-kiel.de)

Version: 2 Date: 5 February 2012

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

Thank you for reviewing our manuscript *Polymorphisms in the mitochondrial oxidative phosphorylation chain genes as prognostic markers for colorectal cancer* (MS:3409353296218273). We have revised all points indicated by referees 1 and 2, and made changes accordingly.

We hope that this version of the manuscript is acceptable for publication and look forward to hearing from you.

Yours sincerely,

Jesús Lascorz
Referee 1 (Sakari K Vanharanta)

Comments to the authors:

Major points

1) The gene selection procedure is arbitrary, and it is not clear why only 7 genes were selected. Why not study the whole pathway? Also, many genetic variants associated with cancer have been identified in regions distant from the proximal promoter and coding regions. Why did the authors only focused on the promoter, 3´, 5´and coding regions? It seems that a more unbiased effort with larger number of SNPs would be appropriate for this kind of work.

Selection of the oxidative phosphorylation (OXPHOS) pathway is based on the recently published systematic pathway-based enrichment analysis of 23 independent gene expression profiling (GEP) studies on prognosis of CRC (Lascorz et al., PLoS One 2011). This meta-analysis indicated the mitochondrial OXPHOS chain as a significantly and consistently overrepresented prognostic category for CRC. Collecting data from the 23 published independent GEP studies on CRC prognosis allowed us to overcome the lack of reproducibility observed in genes reported in individual expression studies.

Within the OXPHOS chain, the genes and SNPs for this association study were selected based on two criteria:

a) First, only those genes reported in at least one of the 23 individual GEP studies on CRC prognosis included in the pathway-based meta-analysis were considered. This way, only genes which have been reported to be differentially expressed based on CRC prognosis were selected.

b) Second, only those genes with SNPs in putative regulatory regions (5’ and 3’-UTR, promoter or non-synonymous coding polymorphisms) and MAF > 0.05 in HapMap CEPH population were considered. This way, we selected only those SNPs located in regions with a putative regulatory function on gene expression.

The approach used, based on stringent selection criteria for both the genes and the polymorphisms, allowed us to select those polymorphisms more likely to influence the expression and function of the protein.

The section Gene and SNP selection in the Methods part summarizes the selection process.
2) No correction for multiple hypothesis testing has been applied.
Taking into account the possible biological effects of the selected genes/variants on cancer progression and a possible correlation between the variants, correction for multiple comparisons was excluded. However, we used multivariate analysis to evaluate the prognostic significance of the variants.

3) Is there an explanation for the fact that variants in different genes are associated with death, metastasis and tumour stage? Are these attributes unrelated in the dataset used?
In colorectal cancer, the TNM stage includes information of primary tumor stage (T), lymph node involvement (N) and distant metastasis (M) at the time of diagnosis. Thus, these clinical parameters are related to each other. A higher tumour stage is associated with the presence of lymph node or distal metastasis and both of them with a shorter survival after diagnosis. We used multivariate analysis to identify genetic factors which may give additional prognostic information independently of clinical stage. If a genetic marker is associated with either lymph node or distant metastasis, its effect on survival would be lost in multivariate analysis.

4) No model is provided that even attempt to explain the observed associations between the genetic variants and tumour phenotypes. How do the T alleles in UQCRB 3′-UTR confer protection from CRC death? How do the COX6B1 promoter/5′-UTR variants promote metastasis? Are these SNPs the causative variants? These are difficult questions to answer, but especially when the genetic evidence is as weak as it is in this study, additional data would be essential before any conclusion can be drawn. An alternative approach would be to solidify the genetic findings.
The reviewer is right when mentioning the lack of a biological model for the associations observed in the manuscript. We hypothesize that, due to the nature of the selection process both for the genes and the polymorphisms (explained in 1), the SNPs might regulate expression. However, a biological explanation for the reported associations is out of the scope of this explorative study. In the conclusions, we remark the necessity of confirmatory studies in independent cohorts to validate the presented results. If this
would be the case, further complex functional studies could attempt to find out the exact biological consequence of the associations, as now pointed out in the discussion.

**Minor point:**
1) On page 7, why do the authors state that rs7836698 is significantly associated with death due to CRC, when in fact the effect, if any, seems to be protective? Since the hazard ratio (HR) of the association is < 1, the reviewer is right in remarking the protective effect of the minor allele of the SNP. The sentence has been modified in the revised version of the manuscript to avoid misunderstandings.
Referee 2 (Peter Söderkvist)

Comments to the authors:

SNPs representing genes involved in all five protein complexes of the oxidative respiration chain, except complex II, are selected. Why are these genes not considered in this study, since they are argued (Introduction) to be important for the Warburg hypothesis.

Selection of the oxidative phosphorylation (OXPHOS) pathway is based in the recently published systematic pathway-based enrichment analysis of 23 independent gene expression profiling (GEP) studies on prognosis of CRC (Lascorz et al., PLoS One 2011), which indicated the mitochondrial OXPHOS chain as a significantly and consistently overrepresented prognostic category for CRC. Within the OXPHOS chain, genes for this association study were selected based on two criteria: first, those reported in at least one of the 23 GEP studies on CRC prognosis included in the pathway-based meta-analysis, all these genes belong to the four complexes of the OXPHOS chain which allow flowing of protons between the mitochondrial matrix and the intermembrane space, resulting into energy generation in form of ATP; second, those with SNPs in putative regulatory regions and MAF > 0.05 in HapMap CEPH population. This way, we selected only those genes reported to have a different expression based on prognosis of CRC and, added to that, with polymorphisms in regions related to gene expression.

None of the genes in complex II fulfilled both criteria, and for that reason none of the genes from this complex were included in our study.

The section Gene and SNP selection in the Methods part summarizes the selection process.

The difference of risk between colon and rectum are not discussed in mechanistic terms (table 4). Please provide an explanation and omit table 4. The results can be included in the text.

Colon and rectal cancer are two different cancer entities, which may have different aetiologies. A short discussion about a higher familial risk for colon than rectal cancer and about a better survival of rectal than colon cancer has been added. Table 4 has been omitted.
In table 6, a significant gene dose dependent effect is expected for rs6510502 and rs10420252 and tumour stage/lymph node/metastasis, but not found. How come? A chance finding?

Both SNPs in COX6B1, rs6510502 (promoter) and rs10420262 (5´-UTR) are in relatively high linkage disequilibrium in the investigated group of patients ($r^2 = 0.71$). This might be the reason for both polymorphisms showing the same effect on the presence of affected lymph nodes (now table 5). As the SNPs were relatively rare (13% and 10%, respectively), we calculated associations only for a dominant model.

The Kaplan-Meier curves: it is not stated in the figure or legend, whether the length (X-axis) on the follow-up period represents days, weeks or months. Please clarify!

If they represent weeks, it is too short follow-up, since many patients are still alive. This figure can be omitted, since most information is already available in the tables.

The follow-up, which was up to 13 years, is represented in months, and it has now been indicated in the figure. We would like to keep the figure, as HRs and Kaplan-Meier curves give somewhat different information about survival. Kaplan-Meier curves show how the survival probabilities change during time, however, no adjustments can be applied. On the other hand, HRs can be adjusted for additional prognostic parameters and they can give information about the significance of the genetic factors as independent prognostic markers.