Author's response to reviews

**Title:** Association between polymorphisms in RMI1, TOP3A, and BLM and risk of cancer: a case-control study

**Authors:**

Karin Broberg (karin.broberg@med.lu.se)
Elizabeth Huynh (liseo_82@hotmail.com)
Karin Schläwicke Engström (karin.engstrom@med.lu.se)
Jonas Björk (jonas.bjork@skane.se)
Maria Albin (maria.albin@med.lu.se)
Christian Ingvar (christian.ingvar@med.lu.se)
Håkan Olsson (hakan.olsson@med.lu.se)
Mattias Höglund (mattias.hoglund@med.lu.se)

**Version:** 2 **Date:** 12 December 2008

**Author's response to reviews:** see over
Dear Editor,
Please, find our response to each of the three reviewers, point-by-point, as well as the comment from the Biomed Central editorial team.

Best regards,
Karin Broberg

Reviewer’s report
Title: Association between polymorphisms in RM11, TOP3A, and BLM and risk of cancer: a case-control study
Version: 1 Date: 4 September 2008
Reviewer: Donghui Li
Reviewer’s report:
1. The role of BLM and BLM-related gene in cancer susceptibility is an important research area.
   Major compulsory revisions:
   The study design is very confusing and problematic. Why men are used as control for female breast cancer?

   Based on the reviewer’s comments, the men in the breast cancer control group have now been excluded from the analysis. The results have been modified accordingly, but the effect estimates were very similar and the significance levels did not change compared to previous analyses. Actually, there was a mistake in the first version. Table 2 and 3 did originally only include women as controls.

2. The four different types of human cancer investigated in this study involve quite different etiology and risk factor. None of the known risk factors was considered in the study.

   Age is a risk factor for most cancers and sex for bladder and breast cancer. Both age and sex were included as risk factors in the analysis. However, it is true that other risk factors were not evaluated. Nevertheless, since our primary aim was to analyse the impact of genetic variation of the BLM-system on different cancer forms, we did not adjust for other risk factors. This is now included on page 12 in the discussion.

3. A convenient sample approach plus a small sample size for each disease could easily lead to some spurious associations. The authors did not even acknowledge any of the limitations of their study and the conclusion is very much overstated.

   We have now modified our discussion addressing limitations beginning on page 11 with “There are several limitations with the study. One problem is that smaller case-control studies may lead to spurious association and it is important to stress that our findings need to be confirmed in other case-control studies. For bladder cancer, the drop out in genotype data/assay was fairly high and these data should thus be cautiously interpreted. The characteristics of the control groups differed; they were selected in different ways during different time frames, showed varying degree of participation, and varied in sex and age distribution. However, allelic frequencies in the different control
groups were similar, indicating that the observed frequencies are good estimates of the true frequency in this region of Sweden. On the other hand, there were different allelic distributions for the SNPs in the different cancer forms. Different repair pathways are important for different cancer forms, but based on our results on each cancer form, homologous recombination may be important in several cancer types. However, the genetic influence is not expected to be exactly the same for the different cancers and therefore, the allelic frequencies may differ. Since the separate analysis for each cancer points at influence of the BLM system, it was considered acceptable to combine all cases and controls, despite the fact that the cancer forms in this study are very different in etiology, and genetic heterogeneity is present in different tumour forms. When the cases from all cancer forms and the different control groups were combined in one analysis, the case and control groups differed somewhat in sex, but were very similar in age distribution. Still, it should be noted that apart from factors as age and sex, we have not evaluated how cancer-specific risk factors, which could influence the associations between SNPs in the BLM complex and cancer risk.”

Also, the conclusions have been modified (page 13)

4. I’m very frustrated in reading this manuscript. I could not figure out what exactly the numbers presented in Table 2 are. For example, for TOP3A rs1563634 in leukemia, the number of cases with GG, GA and AA genotype was 84, 54, and 11, respectively according to Supplemental Table 1. In the manuscript, AA seems to be used as the referent group, and AG/GG genotype showed an increased risk for cancer. What is the number 75/66 represent here?

Because data presented in this table is the main finding of the study, without clear understanding of the presented data, I really can not properly evaluate the study.

We apologize for the unclear description of the data. However, the reason for different allelic frequencies for rs1563634 between the supplementary tables and the tables in the manuscripts was as follows: The results for leukemia and malignant melanoma indicated that the variant allele of rs1563634 was protective. In order to evaluate the cancer risk genotypes, we used the variant homozygotes of rs1563634 as reference genotype in subsequent analyses of breast cancer, and in the analyses of all cases and controls combined. This reasoning resulted in risk allele frequencies of 75/66 (cases/controls) for the rs1563634. This was in the original version indicated in Tables 2-4 in the footnotes. However, we understand that this change needs to be included in the text as well. We have now added the following sentence in the result section on page 9 and the statistical analysis section on page 7 as well as in the footnotes of the Supplementary tables I-II: “For rs1563634 in TOP3A, the reference category was changed to the variant homozygous genotype in order to display cancer risk increment.”

5. Explain why the drop rate for bladder cancer is so high.

The drop rate probably reflects that the DNA has been used in other studies of ours and that there was limited amount of DNA left. We therefore carefully evaluated the results and only included those samples that demonstrated positive signals in most assays. This information has now been included in the materials and methods, page 6.
6. Explain what type of cancer of the control spouses were accompanying?

The Oncology clinic is a regional clinic and responsible for treatment of all kind of tumors, except for leukemias and pediatric tumors. This means that the control spouses were accompanying patients with malignancies, reflecting the cancer forms found in the general population.

7. Change the Table 2 format as that in supplemental Table 1. Give the number of cases and controls for each genotype.

The format has now been changed and the number of cases and controls has been included.

8. Minor revision:
Cite the reference for Haploview on page 5 in the reference list.

The reference is now cited in the reference list.

9. Give the Taqman assay numbers in the supplemental table and cut down on the text.

This has now been changed according to the reviewer’s suggestion.

10. Explain how multiplicative interaction was estimated.

Interaction is here defined as departure from the additivity of effects. As a measure of interaction, Relative Excess Risk due to Interactions (RERI), could be estimated with the formula: $\text{RERI} = \text{OR}_{11} - \text{OR}_{10} - \text{OR}_{01} + 1$, (Greenland,S. and Rothman,K.J. 1998 Modern Epidemiology, 2nd edn. Lippincott Raven, Philadelphia, pp. 329--342.). However, we did not perform any formal testing of multiplicative interaction in this study. We thought that this was not necessary, based on the figures.

11. Explain what is “therapy-related AML” on page 12

One risk factor for acute myeloid leukemia is previous history of chemo- and/or radiotherapy, and this type of leukemia is defined as therapy-related AML. This is now explained on page 13.
Minor essential revisions
1. There is no mention of the ethnicity of the control and disease populations. Was it the same in the two populations? Polymorphic allele frequencies vary greatly between different ethnic groups.

We did not ask for ethnicity of the study participants. However, based on their names, the absolute majority of the participants was of European descent (predominating Swedish). We have now included this information about the study participants on page 4.

2. It is important to indicate whether there were any differences in polymorphism distribution in males and females particularly because the proportion of males and females between control and cancer groups are so different for some of the analyses.

The ratio between men and women differed in particular for the breast cancer study. However, since we now have excluded the men from the control group to breast cancer cases (please, see comment to previous reviewer), we have not analysed sex differences in allelic distribution.

3. It is mis-leading to refer to the ‘leukemia’ group when a large proportion of this group have MDS, not leukemia.

We have now changed the term leukemia group to AML/MDS group throughout the manuscript.

Discretionary revisions
It would help the reader if the research goals were more clearly defined – see points below:

1. Why were the 3 different, unrelated, malignancies chosen? I would have expected cancers to be associated with Bloom syndrome to be chosen. Justification of the choice of cancers should be provided in the introduction.

According to German’s study on the first 100 cancer cases among individuals with Bloom’s syndrome, there is a predominance of leukemias during younger ages and epithelial cancer forms during adult life, reflecting the cancer pattern seen in the general population. Our hypothesis is that the DNA repair genes of the homologous recombination repair complex increases the cancer risk in general. Therefore different malignancies such as leukemia, epithelial cancer of the breast and bladder as well as
malignant melanoma are relevant to study in relation to genetic variants of the BLM complex. Case-control materials were available to test our hypothesis. We have tried to clarify this in the introduction (page 3): “Since mutations that alter BLM function are associated with elevated cancer susceptibility, we reasoned that genetic variants of BLM and other proteins that form a complex with BLM might affect the risk for different cancer forms. In order to test this hypothesis we analysed in this study polymorphisms in the RMI1, TOP3A and BLM, and their association with cancer risk in available case-control materials, namely AML/MDSs (acute myeloid leukemia and myelodysplastic syndromes), malignant melanoma, and bladder and breast cancer.”

2. Is there a rationale to combine all cancer cases? This needs justification in the manuscript. Firstly, the individual SNPs show different distributions between the different malignancies. Secondly, the etiology of the 3 diseases studied is diverse and each disease is associated with different DNA damage types with different DNA repair pathways expected to play an important role in the disease. Both these points raise the issue of whether the cancer cases should be combined.

It is true that there are different allelic distributions for the SNPs in the different cancer forms, i.e. for rs1563634 and rs401549. Different repair pathways are important for different cancer forms, but based on our results on each cancer form, we believe that homologous recombination may be important in several cancer types. However, we do not expect the genetic influence to be exactly the same for the different cancers and therefore, the allelic frequencies may differ. Since the separate analysis for each cancer points at influence of the BLM system, we considered it acceptable to combine all cases and controls.

We have now added a comment upon this in the discussion on page 11.

3. Whilst the genomic position of the 4 selected SNPs is referred to, some discussion of whether these SNPs are likely to be functional would be useful.

As we point out on page 11, there is not yet any functional data for these polymorphisms. These SNPs were selected, because they are tagSNPs for the region. The rs2532105 of BLM and rs12945597 of TOP3A is not in LD with any SNP that is non-synonymous. On the other hand, rs1563634 is situated in a putative CpG island, DNA sequences associated with gene regulation. We have now modified the discussion of the possible functionality of the SNPs (page 11). Also, due to the limited functional data, we present a strict prior probability for the FPRP analysis (0.01) as discussed on page 12.
**Reviewer's report**

**Title:** Association between polymorphisms in RMI1, TOP3A, and BLM and risk of cancer: a case-control study  
**Version:** 1  
**Date:** 2 October 2008  
**Reviewer:** Ulla Vogel

**Reviewer's report:**

- **Major Compulsory Revisions**

1. There is no description of the quality control of the genotyping of the leukemia, melanoma and bladder case-control studies. This should be added.

   We have now included a description of the quality control for the Sequenom analysis in the manuscript on page 5. “For the Sequenom assays, 20 percent of the samples were rerun and negative controls (water instead of DNA) were included in each run.”

   As mentioned on page 6, the quality control for the Taqman assays encompassed: “For all assays, at least 5% of the samples were reanalyzed and the concordance rate of these analyses was 100%.”

2. I really think that all four case-control studies are too small to merit publishing in relation to genetic epidemiology. Moreover, it is a poor epidemiological design to compare female breast cancer cases with a control group which also includes 30% men.

   Based on the reviewer’s comments, the men in the breast cancer control group have now been excluded from the analysis. The results have been modified accordingly, but the effect estimates were very similar and the significance levels did not change compared to previous analyses. Actually, there was a mistake in the first version. Table 2 and 3 did originally only include women as controls.

3. In order to publish this work, I think that you should reproduce your results regarding the four selected SNPs in a cohort that includes at least 400 cases and 400 controls for example for breast cancer. There are several possibilities in Scandinavia, if you want to use the same etnichic group. I think that there is some potentially very interesting findings, but the study cohorts are simply too small to merit publication in my opinion, even though you have four. This is a very interesting pilot study that should open doors in terms of starting a collaboration with larger cohorts.

   Our hypothesis that polymorphic variants in the BLM complex increase the risk for several cancer types has been substantiated in our study by looking at different cancer forms as bladder cancer, leukemia, MDS, breast cancer and melanoma. Although the number of cases for each cancer type is 100-200 cases the finding for each type supports each other. Also our data from the different control groups strengthen the finding. Further cases for replication and subgroup studies combining other risk factors with gene variant carrihship have been gathered. We however think that the present study merits publication as a preliminary report.
BMC Cancer Editorial office:
We request that you make some formatting changes to the manuscript, details of these changes are just below. Your cover letter should include details on how the requested formatting changes have been incorporated into the manuscript.

1. Informed consent must have been given, and a statement regarding this should be included in the methods section. Manuscripts may be rejected if the editorial office considers that the research has not been carried out within an ethical framework, e.g. if the severity of the experimental procedure is not justified by the value of the knowledge gained.

Informed consent has been given of each participant and we have now added this information in the Methods section, page 4.