Reviewer’s report

Title: Haplotype frequencies in a sub-region of chromosome 19q13.3, related to risk and prognosis of cancer, differ dramatically between ethnic groups.

Version: 2 Date: 20 November 2008

Reviewer: Fabienne Lesueur

Reviewer’s report:

Manuscript Number: 1805227278226189
Haplotype frequencies in a sub-region of chromosome 19q13.3, related to risk and prognosis of cancer, differ dramatically between ethnic groups.

The 19q13.3 region encompasses 4 genes related to DNA repair and cell survival, or cell proliferation. A number of polymorphisms within this locus have been associated with increased risk of several cancers in Caucasians and in Chinese population, although identification of the causative variants is still lacking. In this article, Schierup et al. performed haplotype-based association mapping in a Danish breast cancer case/control data set, and surveyed the locus for differences in haplotype structure and frequencies between populations of European, Asian and African descent. They conclude that the risk haplotype has increased in frequency in the European and Asian populations, since the out-of-Africa migration, which is rather unusual. The cause of this increase is unknown.

Reviewer’s comments:
This is a well-written manuscript, in which the authors present their data in a comprehensive manner. The study is clearly documented, and the methodological approach is appropriate to address the question. However, some statements and some Materials could be described in a little more details, given that this is a relatively short manuscript.

Minor Essential Revisions

1. Although results of the linkage disequilibrium mapping in the Danish breast cancer case/control data set as been previously published (Nexo et al. BMC Medical Genetics, 2008), a brief description of the samples series used in this study should be given.
2. The nature of the RAI-3’d1 polymorphism (5-bases tandem repeat polymorphism) should be described.
3. Page 4. What is the meaning of the “grandparent effect”?
4. In the introduction, it is specified that PPP1R13L and ERCC2 are located in a haplotype block with very limited recombination in Danes. Then, in the Results section, it is stated that the Danish Breast cancer case/control data set is pooled to European HapMap Caucasian data, and that untyped SNPs have been
inferred from the HapMap CEU data set. Has the limited number of recombinations of the associated haplotypes in the Danish population been also observed in the HapMap data? In other words, are the LD structure and polymorphisms frequencies of the 2 data sets comparable?

5. Position of polymorphic markers that had been previously reported to be associated with increased risk of cancer should be given, relatively to the SNPs and haplotype blocks presented on Figure 5. Which and where are the variations associated in the Caucasian population? In the Asian population? In which block are located RAI-3’d1, RAI-3’d3, RAI-3’7?

6. In the Method section, paragraph Typing strategy, it should be indicated that the complete list of markers is available and listed in the supplementary material from the previous paper of the investigators (Nexo et al. BMC Medical Genetics, 2008).


8. Page 2, Abstract: “the data does not allow us to distinguish …”[end of the sentence missing]

9. Figure 1: Although ERCC2 is also known as XPD, the authors always refer to as ERCC2 in the text of the paper whereas they use the other name of the gene on the Figure. Same name should be used along the manuscript.

**Level of interest:** An article whose findings are important to those with closely related research interests

**Quality of written English:** Acceptable

**Statistical review:** No, the manuscript does not need to be seen by a statistician.

**Declaration of competing interests:**

I declare that I have no competing interests.