Reviewer's report

Title: Genetically protected young smokers? - A case-control study on lung cancer before the age of 50 years

Version: 1 Date: 12 September 2007

Reviewer: ADONINA Tardon

Reviewer's report:

General

****************************************************************************

Minor Essential Revisions (such as missing labels on figures, or the wrong use of a term, which the author can be trusted to correct)

General

In this study, the authors investigated the relationship between 11 SNPs in phase I and II metabolizing, tumour suppressor and DNA repair genes [CYP1A1 (Val462Ile), EPHX1 (His113Tyr and Arg139His), GSTP1 (A-193C), NAT2 (Thr114Ile and Gln197Arg), GPX1 (Pro200Leu), p53 (Arg72Pro), XRCC1 (Arg280His and Arg399Gln), and XPD (Asp312Asn)] and the risk of lung cancer in a population consisting of 246 early onset lung cancer patients (younger than 51 years at diagnosis) and 223 control subjects.

They observed genetic association for GPX1 (Pro200Leu) polymorphism and gene-smoking interactions for GPX1 (Pro200Leu) and EPHX1 (His113Tyr) polymorphisms. Leu-allele carriers of GPX1 (Pro200Leu) polymorphism showed a significant decrease of lung cancer risk in general and in heavy smokers, while His-allele carriers of EPHX1 (His113Tyr) polymorphism showed a significant decrease of lung cancer risk in moderate smokers. When they considered both variants together, they found a monotone decrease of the OR for each protective allele in smokers.

1. The study is well designed but the authors need to explain better the control selection.

2. Sample size is large considering that the study only recruits cancer patients younger than 51 years, but authors should provide proof of the adequacy of their sample size to detect associations, thus proving their study is adequately powered for the stratification.

3. This study has a reasonable size to pursue a few focused hypotheses. However, if one assumes the need to assess multiple genes and stratify by many categories in the analysis, then the sample size need power calculation.

4. The authors should clarify how they have grouped into never smokers (0-1
PY), moderate (1-19 PY) and heavy smokers (20 and more PY).

5. For about 20% of the cases at least one parent was non-German coming from other European countries or North-America. A recent paper from Campbell et al. (Nat Genet 2005, 37(8):868-872) reported that European populations may display various levels of genetic substructure which may lead to false positive associations due to population stratification. Could this affect the present study?

6. The authors only have shown results for 4 SNPs [EPHX1 (His113Tyr), GSTP1 (A-193C), NAT2 (Gln197Arg), GPX1 (Pro200Leu)], they could be better to present the results for the other [CYP1A1 (Val462Ile), EPHX1 (Arg139His), NAT2 (Thr114Ile), p53 (Arg72Pro), XRCC1 (Arg280His and Arg399Gln), and XPD (Asp312Asn)].

7. The abstract conclusion is too general

8. The reference list is completely incorrect: the reference 12 is the same to 76; 15 is 71; 49 is 80; 65; 66 and 67 is the same; 28 is 39; 70 is 74;

9. In general, tables are understood with difficulty. The variables and dates organization is not clear, especially in tables 1, 4 and 5.

10. Meaning of P and C should be clarified in Table 2.

Discretionary Revisions (which the author can choose to ignore)

1. Adenocarcinoma instead of adenoma carcinoma.

2. Abbreviation for squamous cell carcinoma (SCC) and small cell carcinoma (SCLC)

3. In page 11, line 8, perhaps it would be better write “exon 3” instead of “3rd exon”.

4. In page 11, line 9, there is a mistake: “withn” instead of “within”.

5. In page 17, line 13, there is a mistake: “EPHX13” instead of “EPHX1”.

6. In Table 2, polymorphisms in NAT2 and CYP1A1 are written in capitals.

7. Reference list showed in “Table 2” could be included in manuscript’s reference list.

What next?: Accept after minor essential revisions

Level of interest: An article of importance in its field

Quality of written English: Acceptable

Statistical review: Yes, but I do not feel adequately qualified to assess the statistics.
Declaration of competing interests:

'I declare that I have no competing interests'