Author's response to reviews

Title: The role of IREB2 and transforming growth factor beta-1 genetic variants in COPD: a replication case-control study

Authors:

Sally L Chappell (sally.chappell@nottingham.ac.uk)
Leslie Daly (Leslie.daly@ucd.ie)
Juzer Lotya (Juzer.lotya@ucd.ie)
Aiman Alsaegh (mrxaa30@nottingham.ac.uk)
Tamar Guetta-Baranes (Tamar.guetta-baranes@nottingham.ac.uk)
Josep Roca (JROCA@clinic.ub.es)
Roberto Rabinovich (roberto.rabinovich@ed.ac.uk)
Kevin Morgan (kevin.morgan@nottingham.ac.uk)
Ann B Millar (Ann.Millar@bristol.ac.uk)
Seamus C Donnelly (seamas.donnelly@ucd.ie)
Vera Keatings (Vera.Keatings@nwhb.ie)
William MacNee (wmacnee@ed.ac.uk)
Jan Stolk (j.stolk.long@lumc.nl)
Pieter S Hiemstra (phiemstra@lumc.nl)
Massimo Miniati (miniati@ifc.cnr.it)
Simonetta Monti (monti@ifc.cnr.it)
Clare M O'Connor (Clare.oconnor@ucd.ie)
Noor Kalsheker (noor.kalsheker@nottingham.ac.uk)

Version: 2 Date: 1 December 2010

Author's response to reviews: see over
Dear Dr Pare,

Please find attached the revised version of our manuscript, which has been modified to take into account the comments of the reviewers as detailed below. All changes have been marked using “track changes” in Word.

We look forward to receiving your decision on the revised manuscript.

Yours sincerely,

Prof. Noor Kalsheker
Response to comments from Yohan Bosse:

Major revisions:

1. The authors should assess severity with lung function measurements and compare with the recent GWAS on lung function

We have now reported the results of the association testing between disease severity class and each variant, as well as the associations with FEV1 as a quantitative trait. There is no evidence for an association of IREB2 variants with lung function from the recent GWAS on lung function, and a comment has been added in the discussion to reflect this.

2. How well do the genotyped SNPs cover the common genetic variants within these two genes?

For IREB2, using 3 tag SNPs as presented here captures 64% (33 of 51) SNPs in HapMap. For TGFβ1 these calculations are not as straightforward to carry out, as only one SNP (rs1800469) is present in HapMap with any frequency. There are 8 SNPs in the TGFβ1 gene in HapMap, and using one tagSNP only will capture 2 of these 8 variants. It is likely that we have captured more variation than this as we have included two further tagSNPs which are not in HapMap. A comment has been added to the results section to reflect this information.

3. Correction for multiple testing

This has now been included, and the IREB2 associations remain significant after this correction.

Minor revisions:

1. Table 3 should be replaced by LD plots

This has now been added instead of Table 3.

2. The characteristics of all SNPs should be provided

This has been included in the new Table 3.

3. Reference to be included in introduction

The appropriate citation has now been included in the relevant part of the introduction.
4. **Does the inclusion criteria about smoking history apply to control subjects?**

   Yes, all subjects had a smoking history of at least 20 pack-years. Whilst we attempted to match for smoking between the cases and controls, the COPD patients had greater smoking history than the controls, and this was taken into account in the analysis.

5. **Methods – details about ethnicity**

   Only Caucasian individuals were recruited, so the mention of matching for ethnicity has been removed from the earlier paragraph for clarity.

6. **Clarification of linkage disequilibrium section in Methods**

   Pairwise LD coefficients were calculated using HaploView, and this has now been added to the methods section.

7. **Does any SNP fail HWE testing?**

   Testing of HWE for all SNPs (in all controls, and by each centre separately) showed that there were no deviations from HWE (all p values greater than 0.01). This is now also shown in Table 3.

8. **Minor typing error for LD values**

   This has now been corrected.

9. **Do all the controls have a FEV1/FVC ratio greater than 0.7?**

   All controls had an FEV/FVC ratio ≥0.7. The FEV1/FVC ratio declines with age. At age 50 the normal value is around 80%, and this declines to 72% at age 80. Given the age profile of the control group, the mean value of 77.8 is well within expected norms.

Discretionary revisions:

1. **Which GWAS studies were considered?**

   The two SNPs that are significant in both the COPD and lung cancer genome-wide studies are rs8034191 and rs1051730. This section of the discussion has now been modified to take
into account linkage disequilibrium with other SNPs in the general area identified during these GWAS.
Response to comments from Benjamin Raby:

1. Controls for population stratification:
   a. HWE estimates in all controls and then by study centre
      The results for all controls are shown in Table 3. The analysis was also done for each
      centre separately, and no deviation from HWE was observed. This is commented on in
      the text of the manuscript.
   b. Reports of allele frequencies in each centre, with tests for differences
      This analysis was carried out, and is mentioned in the text.
   c. Sensitivity analysis by repeating analysis by centre
      This has now been included.

2. IREB2 vs genic region
   The discussion has now been modified to be more conservative in the conclusions made.

3. Corrections for multiple comparisons
   This has now been included, and the IREB2 associations remain significant after this
   correction.

4. The discussions regarding TGFB1 and MMP12 are speculative
   The discussions are somewhat speculative, but pull together existing hypotheses about the
   interaction between MMP12 and TGFB with respect to the emphysema phenotype.
   However, as mentioned in the discussion, we could not assess this in our sample collection
   as we had no CT data available. Whilst we appreciate that TGFB1 has multiple functional
   roles, we chose not to focus on these in view of the non-significant associations seen in the
   current study.

5. Limitations
   This was a replication study which aimed to confirm or refute previous reports, which is how
   the specific SNPs were chosen for this work. A section on coverage of the genetic variation
   has now been added to account for the fact that there are other SNPs which were not
   included in the current work. We have also included a comment on future work required,
   which includes targeted resequencing combined with functional analysis. There is now an
   additional section in the discussion about the difficulties of identifying these functional
   variants/genes in areas with high levels of LD.