Reviewer's report

Title: Gene-environment Interactions on Lung Function- A Genome-Wide Association Study in the Framingham Heart Study

Version: 1 Date: 27 September 2013

Reviewer: Lisa Maier

Reviewer’s report:

This manuscript is evaluating the interaction of an assessment of occupational exposure and genetics on lung function on two previously collected cohorts of the Framingham study, in the hopes that by combining the two that exposure and genetic factors predicting lung function will become apparent. This is a new question that has not been addressed by previous genome wide association studies, assessing lung function predictors. The exposure is classified using a population specific Job exposure matrix (JEM), which by necessity is simplified, and the authors do not refine it but characterize these exposures into low and high likelihood of dust exposure. It is not clear if this JEM would be utilized similarly or be able to be replicated by other investigators. The genetic analysis is determined using a genome wide SNP array, and an interesting approach of determining network associations of SNPs. The analysis of GWAS data is appropriate. The lung function variables chosen represent just one point in time for each individual and are not used to classify the individuals as having disease or not. This is a limitation to the study as there is no inclusion of information regarding health status into this analysis. The authors should be commended for including exposure into the genetic analysis as it is an often missed part of genetic studies.

Major Compulsory Revisions:

1) While the authors provide sound rationale for evaluating lung function predictors, including genetics and exposure, they do not characterize the lung function abnormalities into pathologic changes, such as those associated with COPD, which can be caused by the environment. In addition, it would be advantageous to explain the exams used to obtain the spirometry measurements for this study as there are many measurements to be used in these cohorts. Since the two cohorts differed in age, would it be beneficial to pick spirometry measurements from exam dates of the Offspring cohort that would match in age to the Third Generation Cohort.

2) Also the authors should note somewhere that the Framingham study only included Caucasians, as this may impact genetic results.

3) If possible, they should also report information about health, medications, and if available reported diagnosis of lung disease, like COPD in the demographic table, even though it is understood that that this is not the primary focus of the paper. It would also be helpful to compare the attributes of the two generations as they appear to have differences from a demographic standpoint- lung function,
smoking, that could impact the ability to find similar results in these groups.

4) When explaining the genotyping the authors mention that SNP analysis was performed on 9,321 subjects from the three generations of participants. However, the analyses focus on just 2 generations of participants. This should be clarified.

5) Obtaining a JEM on a population like this one is difficult and raises a concern regarding the classification of occupations into categories of “high and low likelihood for dust exposure”. Is there more information as to how the JEM was constructed? It would be helpful if there were references or an explanation why certain jobs were considered dusty or not dusty. Who classified the jobs- is the classification from the Framingham study? For instance it seems that firefighters would have dustier jobs than a mailman or delivery person. It appears that that the authors only used the occupational information from a questionnaire in exam 8 (last examination visit). Was time spent in multiple occupations recorded and if so, could this information have been included? There was no explanation as to whether this represented the longest held job or if occupation changed over time. For instance, if someone worked in a “likely dusty” occupation earlier in the follow-up period and on exam 8 reported becoming a ‘production manager’ or having an ‘administrative job’ is that previous ‘likely dusty’ occupation used in analysis? If only the occupation reported in Exam 8 was used then possible confounding effects should be addressed or at least mentioned in the discussion.

6) The use of network analysis for the GWAS study was novel and intriguing. It would benefit from greater explanation as to why the authors chose the gene list from the Huang study. Would other studies be appropriate to be used in addition? It would also be helpful if the authors provided the p values of the networks studied as this information may add strength to their use of network analysis and actually may help indicate whether certain genetic pathways are important in the prediction of lung function and interact with exposure. This would be helpful in Figure 1.

7) The discussion is quite long but and It would help if it was more connected and included some integration between findings of networks, function and or similar findings from other studies of exposure or genetics.

8) The paper needs major editing focusing on grammar, tense and typographical errors.

Minor Essential Revisions: Nothing to address

Discretionary Revisions:

1) An overall description of the study design would enhance the methods section and readability of the paper.

2) With significant data available to indicate changes in spirometry over time, the authors would have had an advantage to evaluate this measurement and determine if genetics and exposure are predictors in decline or stability of lung function. It would be helpful to discuss why they did not do this, since this would have been a more novel approach as there have been a number of studies evaluating genetic predictors of cross sectional lung function, but few evaluating changes in lung function over time. This assessment might have also been able
to help determine the impact of exposure as others have found.

3) The authors found higher lung function in those with dusty exposures. This is counterintuitive but could be related to illness or other health with non-dusty jobs or because of change in jobs due to health exposure over time. It would be helpful if the discussion addressed this issue and how it might have impacted the authors’ findings.

4) It would be helpful in the discussion if some additional discussion of the function of some of the genes was provided, such as ZNF804A and OPRM1

**Level of interest:** An article of importance in its field

**Quality of written English:** Not suitable for publication unless extensively edited

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