Reviewer's report

Title: Novel Variants in the PRDX6 Gene and the Risk of Acute Lung Injury Following Major Trauma

Version: 1 Date: 22 December 2010

Reviewer: Li Gao

Reviewer's report:

Dear Editor,

Doran and colleagues have performed variant discovery and case-control genetic association analysis of acute lung injury following major trauma with the peroxiredoxin 6 (PRDX6) gene in the manuscript entitled “NOVEL VARIANTS IN THE PRDX6 GENE AND THE RISK OF ACUTE LUNG INJURY FOLLOWING MAJOR TRAUMA”. Although the variation discovery and statistical analysis are quite solid, the primary weakness is no association found between these PRDX6 SNPs studied and ALI. This is a potentially interesting piece of work given the protective role of PRDX6 gene against oxidant injury.

Major Compulsory Revisions.

1. Methods, SNP genotyping. Please describe whether quality control approaches were implemented in genotyping, if so, how (e.g., percentage of random samples repeated). Given the high genotype data missing rate (up to 66% for 7 SNPs) mentioned in the discussion section, both the genotyping completion rate and data concordance rate should be reported.

2. Methods, Statistical analysis of ALI association. It’s not clear whether the association testing was performed separately for the European and African Americans. Also authors need provide more details on potential confounding factors adjusted in the multivariable genetic association analyses.

3. Results, Identification of novel polymorphisms in PRDX6, line 5. Table 2 listed 37 SNPs which are known. Providing allele frequency information in addition to number of subjects carrying the SNP (last column in Table 2), and a comparison of allele frequency information available in the public databases (e.g., dbSNP, HapMap or the 1000 Genomes Database), will be more informative.

4. Results, Identification of novel polymorphisms in PRDX6, line 6. Authors mentioned that 25 of the novel SNPs uncovered were submitted to NCBI, are dbSNP ss# available for those? If so, please list them in Table 1.

5. Results, In Silico function of novel SNPs in PRDX6. Given that no association was found between PRDX6 SNPs studied and ALI and no variants in the coding regions was identified, the major strength of this paper is the discovery of potential functional variants in either gene promoter that could regulate gene transcriptional activity or in 3’UTR that might be altering mRNA stability. Authors described the findings of TESS database search in Table 3, however, more details are needed, i.e., 1) how many SNP sites within PRDX6 gene either create
or abolish a transcription factor binding site? 2) Of them, how many are novel?

6. Results, Association of PRDX6 with ALI. Since “Blunt Mechanism (%))” only appeared in Table 5, and was not described in Methods/Patient population, please provide the description in Table 5 for “Blunt Mechanism (%))”.

7. Results, Association of PRDX6 with ALI. Is there a reason why the authors chose 3 and 10 consecutive SNPs for the sliding window haplotype analysis?

8. Was the association testing performed for severity (APACHE) and outcome variables?

9. Discussion. Given that the major finding of this study is the identification of novel SNPs within the PRDX6 gene and its 5’ and 3’ flanking regions via direct sequencing, the subscript is to justify future studies in identifying PRDX6 gene as genetic risk factors in other diseases. Therefore, I believe that more discussion directly on the involvement of Prdx6 in other diseases and conditions (e.g., lung cancer) needs to be provided in the text in order that the reader can evaluate more precisely the significance of the findings of the study. Similarly, authors briefly mentioned the important interaction between Prdx6 and GSTpi in the background section, but didn’t further explore that as possible future direction testing gene x gene interactions in the Discussion, especially given that genetic association between glutathione S-transferase (GST) gene variants and a variety of human diseases have been established.

10. In summary, this is a potentially interesting piece of work given the protective role of PRDX6 against oxidant injury. However, discussing the biochemical and mechanistic implications of the resequencing results (rather than the genetics) would enhance the value of the manuscript to the general reader.

Minor Essential Revisions.

1. Background, 2nd paragraph. Please spell out class pi glutathione-S-transferase (GSTpi) when it first appeared in the text.

2. Background, 3rd paragraph. Please provide URLs for SNPbrowser and HapMap.

3. Methods, PRDX6 resequencing. Authors referenced published literature (Kruglyak L, et al., Nature Genetics) for the power calculation method used in this study, please provide direct information on the software or program and parameters used for the power calculation.

4. Methods, Statistical analysis of ALI association. Line 7, please spell out LD when it first appeared in the text.

5. Results, Identification of novel polymorphisms in PRDX6, line 4. Please provide NCBI dbSNP build number for the SNP matching comparison and the URL for Genewindow.

6. Results, Haplotype Analysis. Authors mentioned that a total of 37 SNPs were genotyped and 7 were excluded from the haplotype analysis due to a high genotyping failure rate. However, in Figures 2 and 3 legends, authors showed the haplotype structure generated from 31 SNPs (should be 30).

Discretionary Revisions.
1. In the abstract, authors may want to clearly state the study design: trauma- ALI cases were compared to ICU sick controls.

2. Results, Identification of novel polymorphisms in PRDX6, line 10, Figure 1. For the readership, I would suggest label all (if possible) or some prioritized novel SNPs with their corresponding IDs. And more descriptions are need for Figure 1 legend to explain PRDX6 gene structure and symbols/abbreviations used for illustration.

3. Results, In Silico function of novel SNPs in PRDX6. Since haplotype analyses did not strengthen the single marker association signals, I would suggest move Table 7 to the supplementary section.

Other minor issues:
1. Methods, In silico modeling of putative function in SNP sites, first paragraph. Italic font should be used for “in silico”.
2. Methods, Statistical analysis of ALI association. line 6. Please delete extra period after “ALI =0.30”.
3. Discussion, 4th paragraph. Should “mrR-942” be replaced by “miR-942”?

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

**Quality of written English:** Needs some language corrections before being published

**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.