Reviewer's report

Title: Association of Nrf2-encoding NFE2L2 haplotypes with Parkinson's disease

Version: 1 Date: 23 September 2009

Reviewer: Jemma Wilk

Reviewer's report:

The authors have performed a SNP and haplotype study of 2 genes (NFE2L2 and KEAP1) involved in oxidative stress to evaluate association to Parkinson Disease (PD). The association studies were performed in two independent case/control samples to evaluate replication of any observed effect. The authors identified a protective haplotype that was associated with older age of onset in the Swedish sample and decreased risk of PD in the Polish sample.

Major Compulsory Revisions

1) The haplotype analysis is a critical result for the study, but not enough information is provided in the methods. Specifically, how were haplotype probabilities handled in the logistic/linear models? One would expect an EM algorithm to provide all possible haplotype configurations with their associated probabilities. Was only the highest probability haplotype used? How were two haplotypes per person handled in the models? Was software designed for haplotype association analysis used and therefore could be referenced (such as PLINK or Haplo.stats)?

2) The main finding of the manuscript is a haplotype allele associated with onset age in one sample and risk in the other. This does not meet the "same allele, same phenotype" standard for replication. While the global haplotype p-values appear to be significant for risk in both samples, the haplotype-specific results do not replicate. The Swedish result appears to be driven by haplotypes increasing risk, which are not significant in the Polish sample, or in one analysis significantly having the opposite direction of effect. Thus, that they have demonstrated replication is not convincing and statements (bottom of page 12) that "haplotype GAGCAAAA showed strong protective association with PD in two independent European populations" is not accurate when the association was observed only to onset in one of the populations. The authors should address this issue in the discussion and provide an argument for why association to onset and risk constitutes replication.

Discretionary Revisions

There is a wealth of publicly available genome-wide association data for PD that could possibly be used (depending on the availability of these exact SNPs) to address replication of the finding.

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

Quality of written English: Acceptable

Statistical review: Yes, and I have assessed the statistics in my report.

Declaration of competing interests:
I declare that I have no competing interests