Reviewer's report

Title: Admixture Mapping of End Stage Kidney Disease Genetic Susceptibility Using Estimated Mutual Information Ancestry Informative Markers.

Version: 1 Date: 7 July 2010

Reviewer: Elad Ziv

Reviewer's report:

This paper represents a replication of an admixture mapping analysis of end stage kidney disease (ESKD) which the authors map to a region in chromosome 22. This mapping has been published before, although as the authors state, replication is important. The authors also go through a series of exercises to determine the power of their marker set and then increase power to enhance the signal on chromosome 22.

However, there are some confusing aspects of the paper that require clarification and/or revision:

Major revisions/clarifications:

1. The neg log p values reported from admixmap seem too high compared with the results from ancestrymap, esp in figure 4. Assuming the authors use base 10 negative log p values then values of 15-20 suggest a very strong association result that is more significant than a genome wide association. I am wondering whether the authors are reporting negative LN(p) or negative Log base10(P). They do not clarify this in their methods section. Based on the fact that a neg log p of ~20 vs. an LGS of ~4-5 is seen in figure 4 it seems more likely that this is a natural log. If so, please convert to log base 10. If not, please clarify the discordance in statistical significance.

2. The fact that the authors are running a case only analysis (as opposed to case-control) analysis should be noted in the introduction or methods and the ramifications should be addressed in the discussion. In particular, the fact that false positive results may occur due to mis-estimation of locus specific ancestry should be explained. Since there are already previous publications on this locus, this is unlikely to be the case, but it should still be addressed as a limitation.

3. The inclusion of the X chromosome is confusing. In fact, as the authors note, the X is often skewed in terms of ancestry due to differential proportions of sexes in the founding of the admixed population. Therefore, it seems simpler to exclude the X-chromosome, or at least account for it differently (use the X mean ancestry to adjust for the locus specific ancestry at each locus on the X.) Currently, table 2 makes it seem like there is a signal on the X which is not fully addressed by the authors.

Minor revisions:
1. The use of the term "MALD" should probably be avoided. The term "admixture mapping" is more appropriate for this paper. While the 2 terms are often used interchangeably, they historically refer to different approaches. MALD, as originally proposed in AJHG 1994, used long range LD in admixed populations to do association mapping (ie marker and trait direct association). But there was no estimation of ancestral states of the chromosomes (locus specific ancestry.) In contrast, admixture mapping, as proposed by McKeigue in AJHG 1997 explicitly infers ancestral states of chromosomes and associates these with the trait. Thus, MALD and admixture mapping are qualitatively different if these original definitions are used. This paper uses admixture mapping and should refer to the procedure by that term.

2. For the presentation of the results, it may be instructive to present in either figure or table form the excess ancestry or the risk associated with locus specific ancestry at chromosome 22. That is essentially the key parameter. The discussion should also comment about whether this excess ancestry explains only part or most/all of the excess non-diabetes ESKD seen in the population.

**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 I have no competing interests