Questions on Delaneau et al (2013) *Nature Methods* 29:84-91.

Improved whole-chromosome phasing for disease and population genetic studies.

Make sure you read through section 9 of the supplementary methods.

1. In the background section, the authors summarize some previous strategies for generating haplotype estimation algorithms that do not scale quadratically with sample size. Summarize the algorithms implemented in MACH and IMPUTE2.

2. The SHAPEIT algorithm also avoids scaling quadratically with sample size. How?

3. The present paper tries to improve upon the original SHAPEIT algorithm by including some features of the IMPUTE2 approach. What prompted the authors to consider this?

4. The algorithm divides chromosomes into small segments, each with a number of heterozygous sites. What are the advantages of small segments? What are the advantages of long segments?

5. How do the segment boundaries change as the algorithm proceeds?

6. Do you think the proposed algorithm will be sensitive to missing data? If so, why?

7. Do you think the proposed algorithm could be adapted to low pass sequencing data? If so, why? If not, why not?

8. What is the danger of including closely related individuals in the sample? How can this be avoided?

9. What struck you most about the paper?