

Reviewer's report

Title: LD-Spline: Mapping SNPs on Genotyping Platforms to Genomic Regions using Patterns of Linkage Disequilibrium

Version: 1 **Date:** 25 August 2009

Reviewer: David Reif

Reviewer's report:

- Minor Essential Revisions

1. The figure legends given in the supplemental file do not match the enumeration of the actual figures. This seems to be a simple mislabeling, as all figures are described.
2. Figures 4-5 (according to the PDF labels) are very difficult to read. Label fonts and plotting characters should be adjusted for readability.
3. (Page 4): There are more dependencies between LD measures than those cited by the authors. They should provide a reference with a more complete discussion (at least another reference to the Devlin paper previously cited).
4. (Page 5): Given the discussion of regulatory regions relating to gene boundaries, the authors should expand on the consequences of unusual genic structure for their algorithm (e.g. long-range enhancers). This is probably best discussed along with the comparisons to the global block partitioning methods.
5. (Page 8): Does the choice of threshold=1 have an effect on the balance of false/true positives/negatives? The authors should discuss consequences of the threshold choice as they relate to the goals of the particular analysis.
6. (Page 9): The increases in gene coverage using LD-Spline on the Affymetrix and Illumina data seem remarkably high. More explanation as to why this marked increase was observed (or why it was inspected given the Ensembl baseline comparison) is necessary.

- Discretionary Revisions

1. (Page 5): The adjectival form of Yoruba (Yoruban) should be used in this sentence describing the HapMap populations.
2. A more detailed breakdown/characterization of the blocks identified by the authors' algorithm (especially versus the comparison algorithms) would provide potential users a better idea of its utility. For example, how does the SNP-centric algorithm handle large genes with distant exonic segments?
3. (Page 10): The authors should check the SQL statement to "map the entire table of SNPs" for typos (e.g. Where is the specification of D' threshold?).

Level of interest: An article of importance in its field

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.