Reviewer's report

Title: Second-generation PLINK: rising to the challenge of larger and richer datasets

Version:3**Date:**30 October 2014

Reviewer: Gad Abraham

Reviewer's report:

The authors describe PLINK 1.9, a major rewrite of the widely used PLINK 1.07 tool, enabling substantially more efficient analysis of genomic data. I have personally been using PLINK 1.9 for a while and have found it immensely useful, making analyses that would previously take hours on PLINK 1.07 much faster, sometimes taking only seconds. This tool represents a large amount of work and I thank the authors for this undertaking.

1) Major compulsory revisions:

None

2) Minor essential revisions:

2.1) Given that PLINK 1.9 is becoming widely used, and that will be relied upon by many researchers, some sort of automated code testing facility would be very useful to verify that no regressions are introduced in subsequent versions. You could start by generating some random test data and ensure that the most widely used functions return consistently correct results.

2.2) For the tables, make it more clear that timing is in seconds, in the table itself or in table captions.

3) Discretionary revisions:

3.1) You may want to consider a plot showing time as a function of dataset size for some common operation, e.g. --assoc, just to make it immediately obvious to the reader what the magnitude of the speedup is; tables are useful but plots are more intuitive.

Level of interest: An article of importance in its field

Quality of written English: Acceptable

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.