We thank Wan et al. (2011) for their useful comments on our recently published paper in Bioinformatics (Wang et al., 2011). We are grateful also to the editor for providing us this opportunity to clarify the issues.

As mentioned by Wan et al. (2011), the two common procedures for testing statistical epistasis are as follows: (i) ‘two-locus interaction’ as used in BOOST (Wan et al., 2010a); and (ii) ‘two-locus association allowing for interaction’ as used in SNPRuler (Wan et al., 2010b), SNPHarvester (Yang et al., 2009), TEAM (Zhang et al., 2010) and Screen and Clean (Wu et al., 2010). In our paper, we tested BOOST, TEAM, SNPRuler, SNPHarvester, and Screen and Clean using both types of epistatic interactions. We believe this is a more realistic and more comprehensive treatment because, in practice, researchers may have no idea whether the underlying epistatic interactions in their data have main effects or not. Our results suggest that researchers should apply both TEAM and BOOST to their data. Given the same nominal threshold, if BOOST outputs more significant pairs, this hints that pure epistatic interaction effects hidden in the data dominate over marginal effects. A similar logic follows for TEAM.

Wan et al. (2011) queried our experiments on the completeness of BOOST. Upon a careful review of our experiments, we discovered a bug in our program script. In particular, we missed out in our program script a step to sort the output of BOOST. This caused the top interacting SNP pair to be not always chosen in each test. Consequently, we misreported in Section 6.5 of our paper that BOOST wrongly pruned the most significant SNP pairs in 4195 datasets without main effect and 756 datasets with main effect. After correcting the program script, BOOST was verified to be complete and did not mis-prune any most significant SNP pairs. This point is corrected in our published paper. Thus, Figure 5 of our original paper should be revised as per Figure 1 here.

Wan et al. (2011) also commented on our experiments on ‘type-1 error rate’ of BOOST. The type-1 error rate we observed is relatively high compared with the theoretical nominal significance level (0.05). However, this is the empirical result obtained from our simulated datasets. We double-checked our experiments and results; and we also investigated those false positives identified by BOOST. We think there are two reasons for this result. First, with the nominal significance threshold set to 0.05, the test value of the likelihood ratio test with 4 degrees of freedom is 38.24 (after bonferroni correction), while the test value of the chi-square test with 8 degrees of freedom is 47.87 (after bonferroni correction). For example, BOOST reports the top SNP pair (SNP669 and SNP723) with likelihood ratio test statistic value 39.173 845 in the 19th dataset of our simulated datasets. SNPHarvester can identify 22 297 (20 320 + 1977) of them. Among the 28 000 top SNP pairs, 20 320 of them can be identified by all three methods. TEAM’s outputs represent the 28 000 datasets. SNPHarvester can identify 22 297 (20 320 + 1977) of them. Among the 28 000 top SNP pairs, 20 320 of them can be identified by all three methods.

In (a), all three methods—TEAM, SNPRuler and SNPHarvester—use chi square test. TEAM’s outputs represent the 28 000 datasets. SNPHarvester can identify 22 297 (20 320 + 1977) of them. Among the 28 000 top SNP pairs, 20 320 of them can be identified by all three methods. (b) and (d) follow similar explanations.

Fig. 1. The completeness space for the four methods. As there are two types of datasets and two types of test statistics, four venn diagrams are drawn respectively. In (a), all three methods—TEAM, SNPRuler and SNPHarvester—use chi square test. TEAM’s outputs represent the 28 000 datasets. SNPHarvester can identify 22 297 (20 320 + 1977) of them. Among the 28 000 top SNP pairs, 20 320 of them can be identified by all three methods.
simulations from 1000 to 10,000, the type-1 error we obtained is 0.0545, which is close to the nominal significance level (0.05).

**Funding:** This research is funded in part by a National University of Singapore NGS scholarship (to Y.W.) and by a Singapore Agency for Science Technology & Research grant SERC 102 101 0030 (to L.W. and G.L.).

**Conflict of Interest:** none declared.

**REFERENCES**


