**PrediXcan**

1. The authors train using the DGN whole-blood data and test on the lymphoblastoid cell lines from GEUVADIS. It is interesting that the expected R2 is low (as one would hope, for the expression levels of the genes significantly disregulated in the test samples versus the controls). But the actual result is a very high R2 indicating high concordance in expression values. Why would this be? The authors dwell on the ‘prediction R2 values’ but their explanation for the divergent R2 scores is very unsatisfactory ([the substantial departure from expected values…suggests that models developed for whole blood are still useful for understanding diseases that affect other primary tissues]). Can it be that the use of cell-line expression data, in this case, is causing the high concordance in expression? But even then, won’t you expect cell-lines to be under less ‘regulatory control’ than, say, blood cells in a human bone marrow, and so have a lesser expected R2 in the first place?
2. The authors suggest that their method is ‘likely to identify causal genes’ in the Discussion. How would this happen? Their method gives a direction for changes in gene expression in an set of genes (be it genomewide or a panel), but I do not understand at all how that can be directly deconvoluted into causal genes and effector/carrier genes?

**LUMPY**

1. I do not understand the motivation behind the coverages that the authors chose to create their simulated genome for testing in this study. So to simulate a 5% SV allele frequency genome, the abnormal genome was sequenced at 0.5X and the normal at 9.5X, pooled to get 10X. Why would they even do this? A 0.5X coverage implies less than 1 read on average at each spanning position. Why not just increase the max coverage to something like 50X, so even a 5% SV allele frequency will mean atleast 2.5X coverage from the abnormal genome? This approach does seem to favour LUMPY over DELLY/Pindel in the FDR rate, but the authors also do not seem to use any parameter tuning for the latter, whereas they do do so for LUMPY in low coverage samples.
2. The way the probability distribution merging around possible breakpoint reads work is that the max point of the sum of distributions is identified, and any distributions that don’t intersect that point are removed. Does this mean that any neighbouring breakpoints will be ignored and only the majority one will be called by this tool? Is that alright, or can another maximization approach be used to separate out the summed probability distributions into ‘subsets’ of regions (ex. A bimodal distribution -> 2 separate regions with one significant overlap)?