**GRASP (Nucleic Acid Research)**

1. I wonder how effective the suffix array approach is in treating low complexity sequences (like repetitive sequences, or things like AAAATATTAAAAAAT). Won’t an SA approach lead to a higher false positive rate because these sequences can be non-specifically mapped apparently ‘homologous’ regions at a higher frequency than a less-frequent (more complex) sequence?
2. It is interesting that all four programs do not do all that well with the DS3 dataset (3C). Could this be attributed to low coverage (10x), or is it probably because of this being simulated data? I am not sure what tests can be applied to test if this is indeed because of the low coverage. Also, it is weird that all the tools do so well with the real human metagenome sample (DS4) whereas the cases where specific Pfam families are chosen (DS1, DS2) have clear differences in sensitivity between the tools. I am not sure I fully understand why this would happen?

**Corset (Genome Biology)**

1. The authors initially use the relative expression between two contigs to determine the distance between the contigs (i.e. if not constant expression, then these two contigs belong to different genes/groups). Then they also compare de novo differential gene expression using the same contig sets. Since they are relying on read counts for the first classification, and then using the first classification alongwith read counts again (they also seem to be equating read counts to expression - a bit of an odd assumption when you have overlapping contigs between genes), in order to generated the DGE profiles between samples. Perhaps I am not understanding this correctly but how does this approach avoid false negative calls? I also see a persistently low (< 0.6) recall rate in the datasets that they apply Corset to. Can this be the reason why there is a low recall rate?
2. The authors merge Trinity and Oases assemblies together and apparently get 200 additional genes ‘not detected using either constituent transcriptome alone’. How does this happen – i.e., how would you get 200 new genes that you have not seen in either transcriptome assembly alone? What’s the guarantee that these are not false negatives, with some highly abundant FP contig mappings from either tool being merged, leading to a seemingly well-covered region?
3. Also, extremely confused by what exactly is their definition of differential gene expression. All I got out of the paper was the number of different contigs mapping to different genes – there does not seem to be any comparison or validation of the expression changes in specific genes between samples, from the algorithmic side or the results.

**LayerCake (Bioinformatics)**

1. The way that LayerCake generates a consensus sequence for reference is very unclear. Which MSA algorithm do they use, and won’t that influence what particular ‘regions of high variation’ are observed relative to this population consensus? It seems like this approach can easily result in misalignments and false positives, and there seems to be no correction/validation that the authors do to address this.
2. Why do the authors divide the l1-norm for the base composition difference between two locations by 2? I do not understand the double-counting swap explanation fully. Won’t one avoid the extra oerlap counts from a shared base because of the Manhattan distance anyways?
3. In terms of visualization, how is this tool any different from IGV at a panned out view?