1. I was able to assemble all contigs into two super-contigs.
2. I am confident in the ordering and orientation of contigs within my super-contigs. However, it would be good to find a link between the two. To do this we would have to sequence a mate pair that maps to a location in both super-contigs. Increasing sequencing coverage would be the best way to do this – by increasing coverage, you increase the probability that a read will span the gap between the two super-contigs (provided they are close enough to be captured by a mate pair, increasing coverage won’t help if there’s a centromere between the two or something).
3. If mate pairs were generated from a diploid genome it adds an extra layer of complexity. Instead of being from either the forward or reversed strand, the read could come from either strand on either chromosome copy. It’s fair to assume most of the genome is homozygous, so this wouldn’t be a problem in most cases. Heterozygosities like SNPs or copy number variations would complicate things, though. A scaffold could have ‘bubbles’ similar to contig assembly (see picture below).

CNV

1. I used blastn against the nr database to look for similarity. These contigs come from ebolavirus, one of the most deadly viruses in the world. This project interestingly correlates with the outbreak happening in Africa right now…