Name/student number:

Score:

Total: 20 points

Time 30 min

No aid

1. What does multiplex sequencing mean? Why do people use multiplex sequencing? (2 point)
2. Give at least two reasons as to why we cannot just do exact string matching when aligning our reads to reference sequences (2 points)
3. The Burrows-Wheeler Transform creates an alfabetically sorted suffix array and a Full-text index in Minute space (FM-index). Mention some important effects of having the data sorted and compressed (2 points)
4. BWA is a program that uses the Burrows-Wheeler Transform in two different implementations: aln and mem. How does each method work? drawing is encouraged (3 points)
5. Why does the average sequencing depth we observe when mapping reads to a reference genome not always reflect the actual coverage of the reference genome? (1 point)
6. Why is the 16s rRNA gene used for microbiome phylogenetic analysis? (1 point)
7. Why does different primers for 16s rRNA amplicon analysis matter so much? (1 point)
8. What is the difference between an Operational Taxonomic Unit (OTU) and Amplicon Sequence Variant (ASV) (2 points)
9. In the following figure. Are we looking at alpha or beta diversity? Have we sequenced deep enough to describe the different samples? ( 2 points)
10. Shannon index incorporates the species ... and ... of a sample (2 points)
11. When accounting for different sample sizes. What does it mean to rarefy samples? What effect does sample rarefication not take into account (2 points)