Name/student number:

Score:

Total: 20 points

Time 30 min

No aid

1. What is the difference between the SAM and BAM file format standard? Which one should we generally use and why? (3 point)
2. How many lines is in a fastq file? What does each of the four lines contain? (2 point)
3. Why is it so important to know the exact type of encoding (Sanger, Solexa etc.)? What does a Phred quality score of 20 mean? (2 point)
4. Mention two reasons for trimming reads (2 point)
5. How does one check a 16s rRNA amplicon sample for adequate sequencing depth? (1 point)
6. How does one check a shotgun metagenome sample for sequencing depth? (1 point)
7. Why should we check for sequencing depth in a metagenomic study? (2 point)
8. Name 2 of the 3 major types of sequencing read categories that one can encounter? Additionally, write down the read directions of those two (directions drawings are sufficient): (3 points)
9. Write how you would make a directory called Yucky\_samples and copy the files yucky\_biofilm\_R1.fastq & yucky\_biofilm\_R2.fastq to this directory (2 point)
10. Mention a program one could use for having a first look at raw sequencing reads (1 point)
11. Mention programs one could use to trim adapters from raw reads (1 point)