Assigment Report: Dewald Eygelaar

Report to: Roche

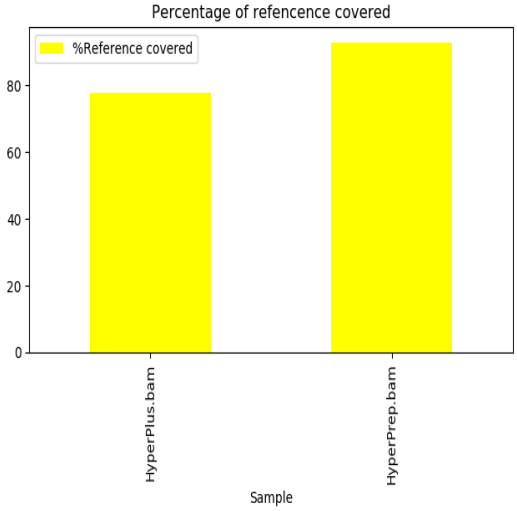
FastQC reports:

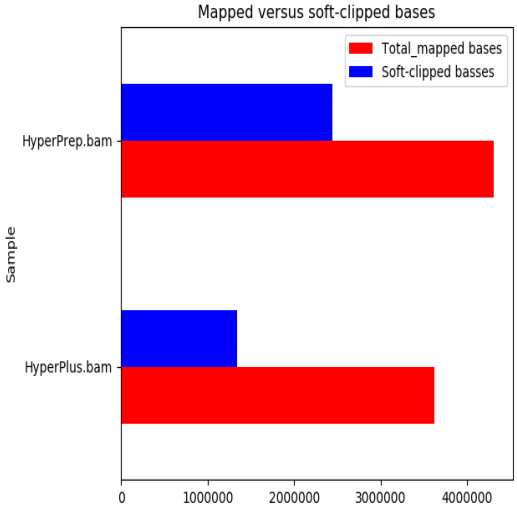
HyperPlus_fastqc > HyperPrep_fastqc >

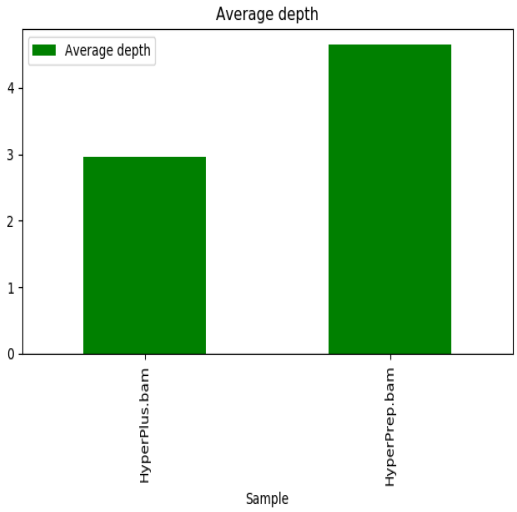
Bam basic stats:

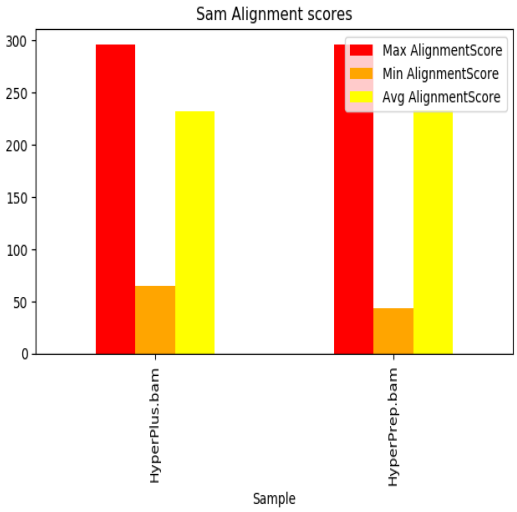
Bam statistics saved to metrics/{sample}_stats.txt

Soft-clip rate was calculated as soft-clipped bases / total bases * 100 = 10% for both samples









After an initial check using a QC program, the reads needed trimming as well as adpater removal. According to QC it the universal Illumina adapter that was used. The HyperPrep sample has roughly twice as many soft-clipped. Soft-clipping may force reads to align at the wrong location, mostly seen at repetitive regions. This may the explain the double amount of mismatches in HyperPrep vs HyperPlus.

The HyperPlus sample contains overrepresentation that may point to some form of contaminant, this corresponds with the sharper per sequence GC content. Even after adapter trimming, fastQC still failed on Adapter content, which again point to over-representation of sequences maybe contamination. The reduced coverage in HyperPlus may be explained by the PCR step not having enough cycles or the PCR was prematurely stop or ran out of dNTPS. Geneious inspection of the alignments as well as the reference fasta, showed that the reference contains some ambigious N's and reads will not map to these parts.