Joseph Wu

112801489

CMSC423

Project 4

Question 1:

Yes, from the results of the experiment, the reads seem to correspond to a gene sequence with a mutation at codon 275. The output of the code from PileUp.py gave me a list of nucleotide positions where there are mismatching positions. Since 274 (accounting for 0-based indexing) is the codon position and each codon is made up from 3 nucleotides, the target nucleotide position would be at 822. The output from the code in that area is:

816 A C:1

818 T G:1

822 C T:10 <- target

831 G C:1

837 T G:1

Nucleotide position 822 mismatches from ‘C’ to ‘T’ frequently in the reference. At the position of 822 in the reference sequence it has codon CAC which codes for histidine and switching position 822 from ‘C’ to ‘T’ gives us TAC which codes for the amino acid Tyrosine. This shows that the reads correspond to a gene sequence with a mutation that changes an amino acid from histidine to tyrosine.