Sam Kitzerow, kitze012

Homework 1

Problem 7

The Stop (Stp) amino acid displays the largest discrepancy in amino acid counts between coding sequences and the whole genome sequence. This amino acid corresponds to the codons TAA, TAG, and TGA which match the displayed discrepancy for codons. Only 12 Stp amino acids were counted with a correct frameshift while 774 Stp amino acids were counted for a random frameshift. This is because with random frame shifts any variation of codons can be identified whereas only the proper ones can be identified with a correct frameshift. For example, one random frame shift can produce 774 Stp amino acids while another shift could produce more or less. In the separate gene fna file, there are only 13 separate genes identified, so we expect to see 12-13 Stp amino acids while the whole genome could produce any number of Stp amino acids.