Problem Set 1

Zachary Foster

March 31, 2014

1

Since there are 4 possible nucleotides and 3 bases per codon, there are $4^3 = 64$ codon sequences possible. There are three stop codons and one start codon. If each codon is equally likely to occur, than for every 64 codons, 3 are going to be stop codons on average. The remaining 61 are broken into 3 ORFs on average (since real sequence is not in 64bp chunks, ORFs on the ends of adjacent theoretical 64bp chunks are shared). Therefore, if start codons are not considered, ORFs of random sequence should be $\frac{61}{3} = \mathbf{20.\bar{3}bp}$ long. If start codons are considered, then for every 64bp, there will be 60 normal codons, one start codon, and three stop codons. The one start and three stop codons will delineate 4 segments per 64bp on average, one of which will begin with a start codon. Therefore, the average length of ORFs in a random sequence is $\frac{60}{4} + 1 = \mathbf{16bp}$ long.

2

A 99% sequence accuracy rate implies a 1% error rate. On average, $0.01 \cdot 100 = \mathbf{1bp}$ of every 100bp read would be incorrect. If a particular region is sequenced 10 times the probability that all 10 reads will produce an error at the same position is $0.01^{10} = \mathbf{1 \cdot 10^{-20}}$. The probability that all 10 reads will produce the *same* error at the same position is $0.01 \cdot \left(\frac{0.01}{3}\right)^9 \approx \mathbf{5.08 \cdot 10^{-25}}$.

3

There are $4^{20} \approx 1.10 \cdot 10^{12}$ possible 20bp oligomers and $4.7 \cdot 10^6 - 19 \approx 4.7 \cdot 10^6$ possible places a 20bp oligomer could match to a random sequence the size of the *E. coli* genome. Therefore, the probability that a random oligomer of 20 bases will hybridize accidentally by a perfect match is:

$$1 - \left(1 - \frac{1}{4^{20}}\right)^{4.7 \cdot 10^6 - 19} \approx \mathbf{4.27 \cdot 10^{-6}}$$

4

There are 770 (1540/2) variable positions, each of which has four options. Therefore, the absurdly large number of permutations is:

$$4^{770} \approx \mathbf{3.86 \cdot 10^{463}}$$