Mckenna Brown, John Zhang, and Anupam Pokharel

Dr. Joshua Kangas

02-261, Spring 2019

5th February 2019

Lab 1 Writeup

**Approach for Task 1:**

Our CalculatePrimerFeatures() function returned a list containing the number of base pairs of type A/T, the number of base pairs of type C/G, and the length of the sequence. Though rather short, this list turned out to be concise yet thorough, achieving 95% or greater prediction accuracy on multiple runs. We briefly dabbled with also including the ΔH (Enthalpy) of the strand, but we realized that, since the Enthaply is derived from a linear calculation involving data already in the list, its inclusion did not yield appreciable improvements in the accuracy of the model.

**Approach for Task 3:**

The goal of the task was to generate one forward-primer compatible with all three organisms’ (pork, chicken, and beef) DNA and then three unique reverse-primers, so we bifurcated our approach accordingly. To find this forward primer, we…

1. …searched the entire length of the chicken strand for an 18-base segment (this was our chosen primer length) that aligned well[[1]](#footnote-1) with the beef DNA.
2. …assessed if that strand also achieved the same standard of alignment with the pork DNA
3. …verified that the melting point was acceptable (within three degrees of 60 Celsius) by generating features on the potential primer using the function composed for Task 1 and extracting the first feature.

If any of the three components of the search was unsuccessful, then the search would proceed from the next index on the chicken’s DNA strand.

Upon obtaining the forward primer, we chose three arbitrary locations higher on the strands (50, 150, and 250 base pairs away for pork, chicken, and beef, respectively to be the starting location of the three unique reverse primers. Then, we generated the reverse-complements of the segments of the DNA starting at those locations, again for a length of 18 base pairs. In a fashion similar to that for finding the forward primer, we successively shifted forward by one the segments of DNA designated as our reverse-primers, if they did not generate PCR products on the given strands when paired with the universal forward-primer (according to the PredictPCRProduct() function composed for an earlier task). Once all three sequences underwent the search for reverse-primers and were observably viable for making PCR products, we performed one final check to determine if the strand lengths were different for all three organisms.

1. For our purposes, “aligning well” in quantitative terms meant to meet a standard such that the first return value of alignment.local\_align() was a score value equal to at least 90% of 10 times 18, our chosen primer length [↑](#footnote-ref-1)