DNA Barcoding Lab Post-Lab Assignment

Collin Coviello and Olivia Compton

BIO110-Molecules & Cells

**Part A. Gel Electrophoresis Analysis**

Figure 1. Agarose Gel Standard Curve. This standard curve was created by measuring the distance between the base pairs of a strand of DNA of known length.

Based on the Standard Curve that was based on the size standard DNA strand, our unknown (Sample C) was approximately 636 base pairs long, and the other unkown, (Sample D) was approximately 597 base pairs long (the cm measurement of the strand from the end of the gel was plugged in for the x value of the equation on the graph, and a reverse log was applied to the end result to give us the number of base pairs.) The size standard went from 900 base pairs to 100 base pairs, and just judging by first glance, one can see that Sample C was between 700 and 600 base pairs long, so this answer is congruent with the initial expectation, and we can assume that this graph is fairly representative of the data. Not to mention, the correlation value is very close to 1, so 97.83% of how the distance migrated changes can be described by our model with the amount of base pairs in the size standard.

**Part B. Bioinformatics Analysis**

The best sequence match for Sample C from Scenario 1 is the from the CO1 gene from salvelinus fontinalis, or Brook Trout. This match had 100% coverage, and the E-value is 0.0. This gene had 652 identities. For Sample D, the gene is the CO1 gene from the mitochondria of salmo trutta trutta, or Sea Trout. However, when found in this region, it would likely have been a Brown Trout, which is a part of the same species as Sea Trout.

During the BLASTn search, the query sequence is built from the data for the nitrogenous bases that is inputted by the user, whereas the subject sequence is the consequent match from the database.

The probability that the match was random is 0%, according to NCBI, since the E value can be interpreted as the chance you have of getting a random match in the database using the sequence that was given.

Phosphodiester bonds join adjacent nucleotides in a DNA double helix.

Sample C:

CCTTTATTTAGTATTTGGTGC

GGA AAU AAA UCA UAA ACC ACG

Glycine-Asparagine-Lysine-Serine-STOP

Sample D:

TTGGTGCCTGAGCCGGAATAG

AAC CAC GGA CUC GGC CUU AUC

Asparagine-Histidine-Glycine-Leucine-Glycine-Leucine-Isoleucine

The COX1 gene codes for a protein of the mitochondrial matrix in fish- cytochrome c oxidase-and plays an important part in the electron transport chain. This is crucial in the cellular respiration for these fish. On top of that, this gene is easily mutated, so differentiating between different species using this gene is very easy compared to other genes.

**Part C. Forensic Analysis Report**

To Whom It May Concern,

My name is Collin Coviello, and I am a student at Keene State College. The samples of fish DNA from this NH Department of Fish & Game case were examined in a laboratory at the school by a team of Biology Students, and it was decided that the man in question was indeed in possession of both brown trout and brook trout.

To figure out which species of fish were contained in the cooler of the perpetrator, the process of DNA barcoding was performed.

As a part of the preliminary steps, we isolated the total DNA from the small samples of tissue, and purified it using several chemical steps. Then the DNA was centrifuged a couple of times, and the samples of purified fish DNA were run through a polymerase chain reaction, which isolates a very specific part of the gene by changing the temperature of the sample and giving way to replication of that isolated section, because of the way DNA polymerase functions.

The next step, Agarose Gel Electrophoresis, required samples of the DNA isolated in the first steps to be placed into a gel, and pushed through with electricity. The electrophoresis of these dyed samples made it easier for the size of the DNA strand to be found. In the gel, we used a size standard DNA which was marked from 900 base pairs to 100 base pairs to create a standard curve from which to extrapolate the approximate size of the unknown strands. Using this method, Sample C was found to be approximately 636 base pairs, and Sample D was found to be approximately 597 base pairs.

Then, the nitrogenous bases were analyzed in a bioinformatics database using BLAST analysis. Samples are run against the archive in order to find the best match of known DNA. The best sequence match for Sample C is from the CO1 gene from salvelinus fontinalis, or Brook Trout. This match had 100% coverage, and the E-value is 0.0, meaning that the sample strand (or query sequence) matches 100% of the subject sequence (the sequence received from the database), and that there is a chance of 0.0 of getting this result randomly with the given sample. Sample C had 652 identities. The gene in question is the CO1 gene from the mitochondria of salmo trutta trutta, or Sea Trout. However, when found in this region, it would likely have been a Brown Trout, which is a part of the same species as Sea Trout.

These results would indicate that the owner of the cooler is guilty of catching fish out of season. There were steps along the way in which the data could have been slightly wrong. For example, the size standard that was measured was actually based on a very small image of the DNA gel. There could have been an issue measuring since precision was more difficult to achieve. Not to mention, it is hard to tell from a picture where the wells in the agarose gel truly were. However, with the checks and balances of our multi-faceted approach, it is still safe to say that the subject is guilty.