**Algorithms in Bioinformatics:**

**Assignment 2 Report**

**Lucas M. Damler**

**Northern Illinois University**

Human comparison to chimp results in around a 98% sensitivity and specificity and standard deviation of around 1.8%. Increasing or decreasing the L value does not result in any significant change to the data while increasing the O value decreases the standard deviation to around 0.5%. This still does not significantly alter the sensitivity and specificity by more than 1-2%. Lowering the C value from 5 to 0 also has little impact on accuracy of the alignment.

Comparison of human and dog sequences results in around a 67% sensitivity and 76% specificity under an O value of 300. Increasing or decreasing the L value does not change this more than a percent. Increase of the O value to 500 increases sensitivity by around 5% hitting an average of 80%, and raises specificity to approximately 90%, an average increase of 10-14%. Dropping C from 5 to 0 lowered sensitivity and specificity by an average of 10% (Sensitivity: M = 58%; Specificity M = 68%). With C = 0 and O = 500 sensitivity was around a 61% and specificity a 79%.

Comparing human to mouse sequences at C = 5, L = 2000, and O = 300 results in a sensitivity of 31% and specificity of 62%. Once again, the L value impacts values no more than 1% while increase of the O value to 500 decreases sensitivity by around 10% but increases specificity by around 20% (Sensitivity: M = 24%; Specificity M = 83%). Decreasing C to 0 drops sensitivity and specificity by 6% and 14% respectively (Sensitivity: M = 25%; Specificity: M = 48%). Increasing O to 500 under a C value of 0 decreases sensitivity to 19% but increases specificity to 67%.

Comparing the human genome to chimp, dog, and mouse genomes produces widely different results. According to the data the L value has little effect on accuracy. Altering of the O and C values changes both accuracy and standard distribution. This program was written in C++ It took around 15 to 20 hours to program and test. Final production included computing and comparing all variations of the data. Only data that showed any significant differences were included in the final report. L and O values in between displayed values showed the same upward or downward trends in represented data.