Priya Narain

Anna Ritz

Biology 131

14 May 2020

Project Report

I performed alignment on COVID-19(coronavirus) and MERS-CoV(Middle Eastern Respiratory Syndrome) ORF1a polyprotein sequences. MERS belongs to the betacoronavirus category, a subtype of the coronavirus family, and SARS-CoV-2 is the novel coronavirus that led to the upper-respiratory disease commonly known as COVID. Sequences from the NIH website for the ORF1a polyprotein were available, which allowed for the comparisons between the two viruses. Performing alignments on the two strings allows us to see the similarities and differences in the viral and genomic nature of the diseases. I defined the longest common subsequence and local alignment functions, and had two files for the COVID and MERS polyprotein sequence FASTA files.

I modified code from homework #10 on sequence alignment and used the longest common subsequence in order to find out the longest sequence commonalities that exist between the two strings. LCS identifies commonalities in the sequences that appear in the same relative order. When running the LCS function, the output I get shows that there are no common subsequences.

When running the LCS function, the output I get shows that there are no long common subsequences.

Next I attempted to perform local alignment on the two sequences. I think I had issues

Severe acute respiratory syndrome coronavirus 2 isolate Wuhan-Hu-1, complete genome:

https://www.ncbi.nlm.nih.gov/nuccore/NC_045512.2?report=fasta&from=266&to=21555

Middle East respiratory syndrome coronavirus, complete genome:

https://www.ncbi.nlm.nih.gov/nuccore/NC_019843.3?report=fasta&from=279&to=21514