Jacqueline Zeng

2/9/2020

SoftDes Mini Project 1 Write-Up

Results *[~1 paragraphs] Present what you by using Protein BLAST or BLAT genome search to interpret and analyze your protein results. Write up a short analysis of the candidate genes [2-3] you found. Include if the accomplished candidates were actual genes, and if so, what organism did they likely come from? What is their possible function within that organism? Is it pathogenic?*

The gene finder code I wrote returned five amino acid sequences. I fed both into Protein BLAST and all of them match proteins of Salmonella. This result was as anticipated as the program found the longest gene sequences, which were likely to be real gene encoding material. The proteins I got were: SPI-1 type III secretion system export apparatus protein SpaR, SPI-1 type III secretion system protein SpaN, FliI/YscN family ATPase, SPI-1 type III secretion system protein SpaO, SPI-1 type III secretion system export apparatus protein SpaS.

Reflection *[~1-2 paragraphs] Aside from understanding how your program works, it is also important to understand the impact it would have on others in the real world. Some questions to consider when writing the reflection:*

This project was really interesting because it got me thinking about how computer science can be used to develop fields that may seem separate from computer science. For instance, biology. Gene finding technologies can be used to aid early detection of diseases. If we are able to find out whether a person has a gene that codes for an illness, we can help them through early treatment and prevention. Although this program was successful for predicting salmonella DNA. it may not work for other DNA types. For instance, this gene finder returns the longest strands of protein. In reality, the DNA that actually encodes information may be actually shorter. Hence, we can easily miss over some DNA.