CSC 314**, Bioinformatics Lab #10: Name:\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**BLAST**

For each question, include screen shots of the relevant results.

***Can genetic differences explain the infectiousness of SARS-Cov-2?***

Here you will use BLAST, following the framework and settings of the in-class example, to identify possible proteins that may explain the difference in infectiousness between SARS-Cov-2 and SARS, based on a search of the *Reference proteins* database.

***For the top hit questions, you must include screenshots that show where you are finding the answers. The answers should also be typed or written for each question. You should also include screenshots showing the list of domain results.***

Protein #1: *nsp7*, <https://www.ncbi.nlm.nih.gov/protein/YP_009725303.1>

1. For the top hit, what is the
   1. Accession number
   2. Description
   3. Percent Identity
   4. Percent Similarity
   5. Query coverage
2. For the conserved domain results, what is the first *pfam* domain listed. Based on this domain, what is the function of this protein?
3. Based on this analysis, do you think this protein may be a reason why SARS-Cov-2 is more infectious than SARS?

Protein #2: surface (spike) glycoprotein, <https://www.ncbi.nlm.nih.gov/protein/1796318598>

1. For the top hit, what is the
   1. Accession number
   2. Description
   3. Percent Identity
   4. Percent Similarity
   5. Query coverage
2. This protein binds to receptors on human cells, which allows the coronavirus to infect human cells. Based on the conserved domain results, answer the questions below.
   1. What domain gives the coronavirus its name?
   2. What two proteins (receptors) does the surface (spike) glycoprotein bind to?
3. Based on this analysis, do you think this protein may be a reason why SARS-Cov-2 is more infectious than SARS?

Protein #3: *nsp2,* <https://www.ncbi.nlm.nih.gov/protein/1802476806>

1. For the top hit, what is the
   1. Accession number
   2. Description
   3. Percent Identity
   4. Percent Similarity
   5. Query coverage
2. According to the NY Times article[[1]](#footnote-1) describing the SARS-Cov2 genome, this protein is a “mystery protein”. What is the first *pfam* domain identified for this protein. Based on this domain, what is the protein’s function?
3. Based on this analysis, do you think this protein may be a reason why SARS-Cov-2 is more infectious than SARS?

1. <https://www.nytimes.com/interactive/2020/04/03/science/coronavirus-genome-bad-news-wrapped-in-protein.html> [↑](#footnote-ref-1)