This supplemental file will provide an assessment of students ability to

1. Find protein sequences through NCBI
2. Perform BLASTp searches to find related proteins
3. Create a FASTA file of sequences
4. Perform multiple sequence alignment of the protein sequences
5. Analyze the resulting alignment for differences between the ACE-2 (the host receptor protein for coronavirus) from different animals

Using your previous worksheet experiences, perform the following steps to analyze the human ACE-2 protein. This is the protein that the SARS and SARS-CoV-2 viruses attach to for infection. Follow the instructions and complete the assessment.

1. Using the full human ACE-2 protein sequence (NP\_001358344), perform a [BLASTp](https://blast.ncbi.nlm.nih.gov/Blast.cgi) search.
   1. Exclude from your search “Homo sapiens (taxid:9606)”.
      1. Select two sequences from the results, download the FASTA aligned sequences and add to your created FASTA file.
   2. Repeat your BLASTp search and now select “bats (taxid:9397)”. Choose at least one entry that is from a *Rhinolophus* species (this bat genus is believed to have hosted the coronavirus that gave rise to SARS-CoV-2). *Make sure to click off the ‘exclude’ button.*
   3. Repeat your BLASTp search to find sequences (take any one) from another species to include in your FASTA file. You might consider
      1. “pangolins (taxid:9971)”
      2. Or “snakes (taxid:8570)” (both considered a vector of the current SARS-CoV-2 pandemic
      3. Or “civets (taxid:9673)” (the vector of SARS-CoV).
   4. Have at least six sequences in your FASTA file.
2. The FASTA sequence file generated in question 1 should now be used to generate a multiple sequence alignment using [ClustalOmega](https://www.ebi.ac.uk/Tools/msa/clustalo/).
   1. Recommendation: For finding the sequences of interest in the next step, set the output format to be based on the input sequence order (you did this in Worksheet 4).
3. Using Table 2 from Worksheet 4, identify the amino acids in the ACE-2 protein determined by Yan *et al*. (2020) to be important for viral attachment.

**Questions**

1. Based on the multiple sequence alignment, which amino acids from Worksheet 4 Table 2 might be:
   1. common in all of the species ACE-2 sequences?
   2. might be more unique to species closely related to humans?
2. A species of *Rhinolophus* is believed to have been the host species for the SARS-CoV-2 virus. At these particular amino acids, how similar are the human and *Rhinolophus* species?
3. Predict the outcome of a ACE-2 mutation in which the D30 amino acid is altered to a lysine (D30K). Explain your answer with an illustration (hand drawn or you can use iCn3D) to support your reasoning.
4. A number of human ACE-2 mutations have been identified that in the laboratory that impact the binding of SARS S protein (no data on SARS-CoV-2 yet) to the ACE-2 protein. Using the data summarized on [Uniprot](https://www.uniprot.org/uniprot/Q9BYF1), select two mutations generated by Zhang *et al*. (2005) that abolish or strongly inhibit the SARS S protein interacting with the human ACE-2.
   1. Explain your answer with an illustration (hand drawn or you can use iCn3D) to support your reasoning.
   2. Using the mutations that are described at Uniprot, pick two mutations listed and explain why these mutations may not have had an effect on SARS S protein and ACE-2 protein interaction.