**Computer Lab 3 – Sequence Alignment**

**Conservation Genetics (BIOL 4174 / 5174)**

Due Date: **2/12/2018 by Midnight.**

Submit via Blackboard.

Questions – Part 1

1. The sequences you edited during part I of lab 2 belong to Bowfin (*Amia calva*). Does this match the top BLAST hit for this sequence?
2. What percentage of bases matched between your sequence and the top BLAST result?
3. Provide the scientific names of the four species that had the next closest matching sequences.
4. Recall that last week I told you this 842bp sequence contained two mitochondrial genes that overlap one another. Use your BLAST results to determine the names of these two genes. (Hint: What are the most common gene names you see in the output?)
5. What might cause you NOT to get a top BLAST hit matching *Amia calva*?
6. Spend some additional time exploring Genbank. Search the nucleotide database for sequences from your favorite animal, or an organism that you study. Use the scientific name of the organism when you search. In the unlikely event that your search results return absolutely nothing, try searching for another organism.
   1. What organism did you search for, and how many sequences were returned by your search?
   2. Do all of them belong to your organism? (skipping ahead to the last page of search results will usually show if your search returned sequences from other organisms).
   3. If your search returned results from other organisms, click on one of these sequences. Can you figure out why your search results included this sequence?
   4. How do the number of sequences found for your organism compare to the number found for some common model organisms (examples: *Drosophila*, *Xenopus*, *Gasterosteus*, *Danio*)?

Questions – Part 2

1. Were there any gaps inserted into the sequences by the program?
2. How did the addition of the python and boa sequences affect your alignment? If ClustalX has inserted gaps into the alignment, report which taxa have gaps and at which columns of the alignment they occur. If gaps have been inserted into all *Heloderma* sequences, you can list these as a group rather than individually.
3. Given what you know about how mutations will affect protein-coding sequences, do you think this alignment is accurate (i.e., has the program accurately identified homologous bases; would a functional protein be produced from a sequence with these gaps)? Support your answer.
4. Report your observations here after modifying the default settings in Clustal.

Questions – Part 3,4 and 5

1. Compare the alignments made by ClustalX, Clustal Omega, and Muscle. Given what you know about how mutations will affect protein-coding sequences, which of these alignments would you consider to be the best? Which is the worst? Do you think any of them are accurate?