**GUIDELINES FOR BIONINFORMATICS PROJECT**

For this project, I prefer that you work in a group of up to three or four students. If you would prefer to work individually, that is OK as well.

Steps of project: (1) Identify a group of organisms that interests you and select a specific evolutionary question that can be investigated via phylogenetic/bioinformatic methods; (2) Carry out an in depth literature review; (3) construct a novel data matrix (DNA or protein sequence data from at least two gene/loci for a minimum of 30 taxa); (4) analyze the data – align data and reconstruct phylogenetic trees; (5) interpret trees in light of your question - map question onto tree; (6) go through a peer review and instructor review process of a poster prior to producing the final version; and (7) present (stand by your poster, explain the research, and answer questions) at the end of the semester as if at a scientific meeting. I am willing to help groups out with additional computational analyses if appropriate.

1. Choose a group of organisms and a question
   1. Your group will choose a group of organisms and decide upon a question that can be answered from insights from a phylogenetic tree. For example the question could be “Are the big cats monophyletic?” or “How many times did big catness evolve?” After reconstructing a tree from sequence data (note the genes does not have to have anything to do with your question, it is simply just data to produce a tree), you then map your question on the tree. In other words, look at how the evolution of “big catness” evolved on the tree. Did it just happen once or multiple times? Did “big catness” ever revert back to “small catness”?, and so forth. Coming up with a group of organisms and good question may be difficult. Maybe start with google searches like “the evolution and phylogeny of cats” or whatever you are thinking about. It may be useful for you to search biological abstracts using terms that are of interests to your group. Most likely your question will be one that has been asked before and that is OK (remember science is repeatable).
   2. Two groups cannot do the same group and question – there is a whole tree of life out there – let’s enjoy the diversity.
2. Literature Review
   1. Carrying out a robust literature is an important skill you should attain. You will need to research your particular group of organisms from primary literature and secondary sources. You must have at least 10 peer reviewed papers cited.
3. Dataset
   1. Depending upon the question that you want to address, your group will need to think about the appropriate taxon sampling (including outgroups) and gene(s) that will be used.
   2. You might use NCBI databases to get your data. I suggest using the Nucleotide, Protein, Taxonomy, and PopSet databases to find your sequences. Come see me for additional help – don’t be shy, this is sometimes the hardest step of the process.
   3. Your dataset must consist of at least 30 terminals, including outgroups (species, organisms, lineages).
   4. You will submit your dataset via canvas.
4. Data Analysis
   1. Multiple Sequence Alignment – if necessary. Remember to keep track of your parameters and maybe try parameters other than the defaults.
   2. Phylogenetic tree reconstruction
      1. Parsimony
      2. Distance
      3. Maximum Likelihood
      4. Bayesian???
   3. Node support values
      1. Bootstrap
      2. Consensus
      3. Bayesian posterior probability
      4. Sensitivity analysis
5. Interpret tree in light of your question
   1. Now that you have trees showing hypotheses of proposed relationships, you can map your question onto the trees. Realize that your trees may be incorrect and that this could lead you to incorrect conclusions about your question. So, think about this. If you want, you may want to tweak your dataset, start over, change parameters, etc.
6. Peer reviews and instructor reviews will occur before the final submission.
7. Posters will be printed (at a size of 36 X 48 inches) and you will present (stand by your poster, explain the research, and answer questions) at the end of the semester, as if at a scientific meeting.