Molecular Phylogenetics (EEOB 563)
Homework #4: Distance Matrix Methods

Part I. Distance methods using paper and pencil ☺.

Table 1:

Dog	ATG	ACC	AAC	ATT	CGA	AAA	ACC	CAC	CCA	CTA
Cat	ATG	ACC	AAC	ATT	CGA	AAA	TCA	CAC	CCC	CTT
Mouse	ATG	ACA	AAC	ATA	CGA	AAA	ACA	CAC	CCA	TTA
Pig	ATG	ACC	AAC	ATC	CGA	AAA	TCA	CAC	CCA	CTA
Human	ATG	ACC	CCA	ATA	CGC	AAA	ACT	AAC	CCC	CTA

1. Table 1 shows the first 30 bp of the mitochondrial cytochrome b gene for five mammals. **Find the Jukes-Cantor distance for each pair of species** (I talked briefly about different models of DNA evolution on Thursday, but here is the formula again: $D(t)=-\ln(1-(4/3)p)$, were p is uncorrected (observed) dissimilarity.)

2. Apply the algorithm we used in class to find the UPGMA tree based on JC distances. **Draw this tree and indicate all branch lengths.**

3. Apply the neighbor-joining (NJ) algorithm to find the NJ tree. **Draw this tree and indicate all branch lengths.**

Part II. Distance methods in PHYLIP.

You may want to check the PHYLIP documentation on http://evolution.genetics.washington.edu/phylip and/or lab3 tutorial before doing this part.

Create a multiple sequence alignment using complete mitochondrial cytochrome b sequences from the *cob_nt.fasta* file we used in class and save it in the PHYLIP format.

4. Calculate 4 different matrices using 4 models available in **dnadist** and perform a NJ analysis with each of them. **Do not submit these trees.** Instead, calculate a strict consensus tree**. Root your consensus trees using an appropriate outgroup and include it with the rest of the assignment.**

5. Create 200 bootsrtap replicates of your data file and build a NJ tree for each of them (you should choose one of the four models of nucleotide substitutions you used in #4).
Build a majority rule consensus tree for these trees, print it out, and submit with the rest of the assignment.

Part III (Bonus point). FastME.

8. Use one of the distance matrices you calculated in #4 to calculate NJ trees with/without the tree refinement with NNI and SPR. **Did NNI/SPR post-processing changed the resulting tree phylogeny? Which of the trees corresponds better to your understanding of mammalian evolution?** ☺