**Phylogenetics Homework Assignment**

Name:

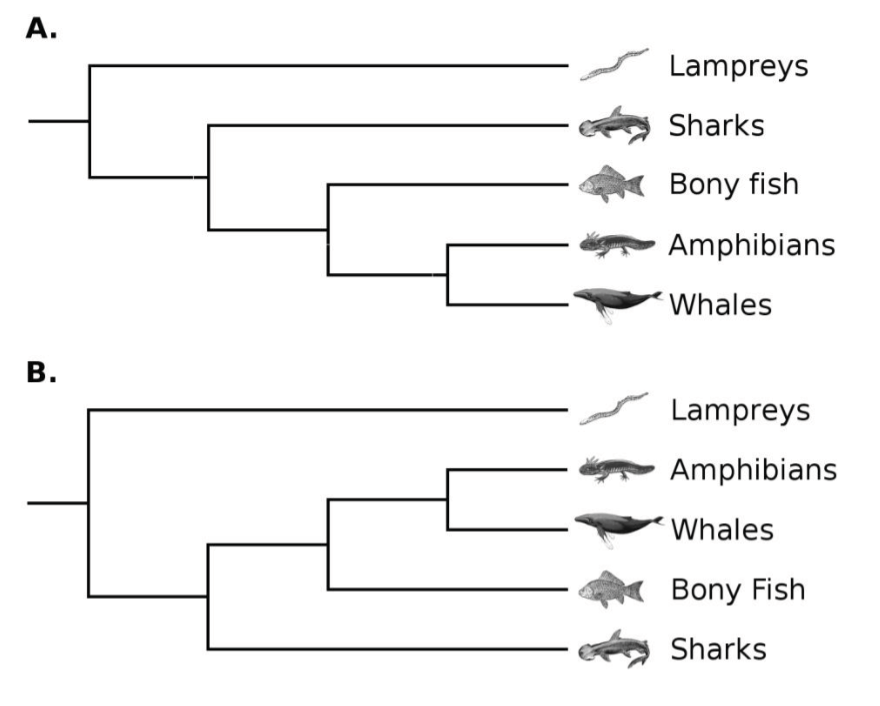
**Please turn in this assignment via Slack to Jessica by 5:00pm on Thursday, November 9th. Points will be deducted for late assignments.**

**Part 1: Concepts of Phylogenetics (12 points)**

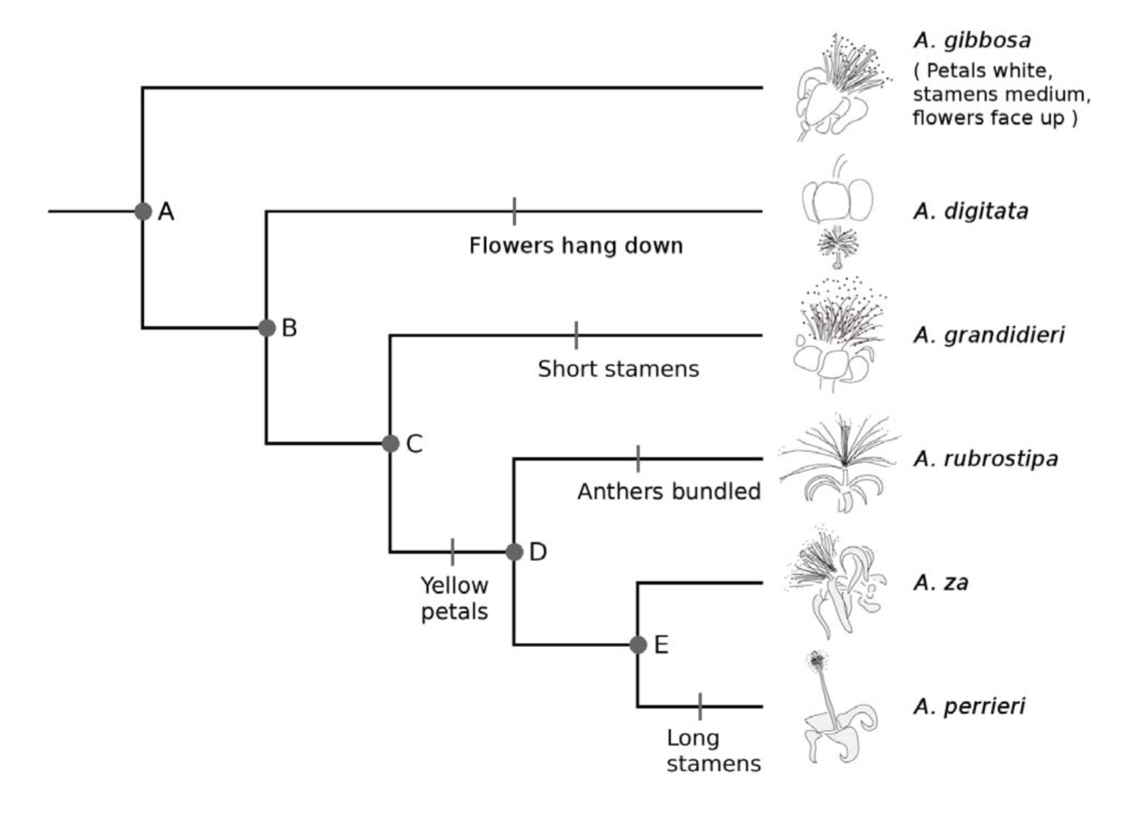
*Answer the following questions regarding the interpretation of phylogenetic trees and what relationships they depict.*

1. Explain whether the two trees below show the same or different phylogenetic relationships among the taxa. (2 points)

Different relationships: the sister species (most closely related / with most recent common ancestor) in tree A are Amphibians and Sharks, while the sister species in tree B are Amphibians and Whales. Can also state that in A, Whales share a recent common ancestor with the group including Bony fish, Amphibians, and Sharks; in B, Sharks share a recent common ancestor with the group including Bony fish, Amphibians, and Whales.



1. The diagram below shows the possible evolutionary relationships among six living species of baobab trees. The drawings show the flowers of the living baobab species. The letters A- E represent ancestral species.



For each of the following three traits, choose the option which describes what the flowers of species D would have looked like according to this phylogenetic tree (3 points):

Species D is the common ancestor of A. rubrostipa, A. za, and A. perrieri. Since the branch leading to D has the trait change to yellow petals, D should have yellow petals. The tree also informs you (under the description for A. gibbosa) that the outgroup of this tree has medium stamens. Since this outgroup shares a common ancestor with the rest of the tree, and the branches leading to D have no changes in stamen length, we can infer that species D had medium stamens. The same logic applies to the orientation of the flowers, which should be face up for species D.

1. Petals: a. White b. Yellow c. Can’t tell from diagram
2. Stamens: a. Short b. Medium c. Long d. Can’t tell from diagram
3. Flowers: a. Face down b. Face up c. Can’t tell from diagram
4. The following image depicts an unrooted tree with 6 species represented. How many locations are possible for the root of the tree? (1 point)

9 locations : Based on your notes from class, the equation is 2n-3, where n is the number of taxa represented on the tree (6). So the result of the equation is 2(6)-3 = 9. You can also count the branches (don’t forget to count the three branches directly in the middle) and you get 9.

**E**

**D**

**A**

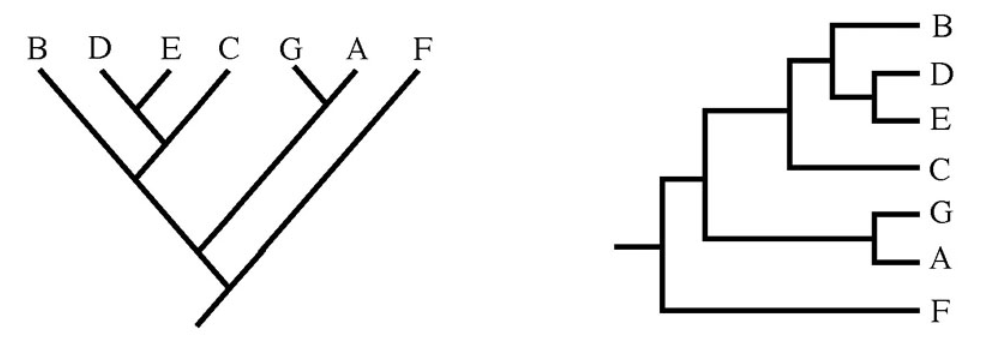
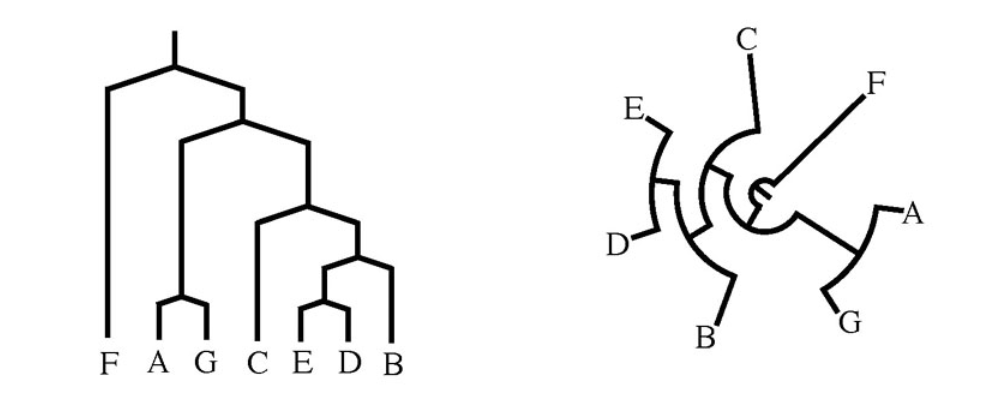
**F**

**C**

**B**

1. Which of these depicts a different relationship from the others? (1 point)

Tree C is different from the others. The sister species relationships are the same for all trees, but in C, the most recent common ancestor of D and E is C, whereas in the other trees, the most recent common ancestor of D and E is B.

****

A

B

D

C

1. Based on the tree below, which of the following is an accurate statement of relationships? (1 point)

A is correct.

A

B

C

D

1. Species A is more closely related to species C than species D.
2. Species A is more closely related to species D than species C.
3. Species A is equally related to species C and D.
4. Species A is more closely related to species C than species B.
5. Explain why you chose the answer you did from part 5. (1 point)

Species A shares a more recent common ancestor with species C than it does with species D. Species A shares a more recent common ancestor with species B than species C, so (d.) is not correct.

1. Does the tree below depict the same relationships as the tree from 5? Why? (2 points)

D

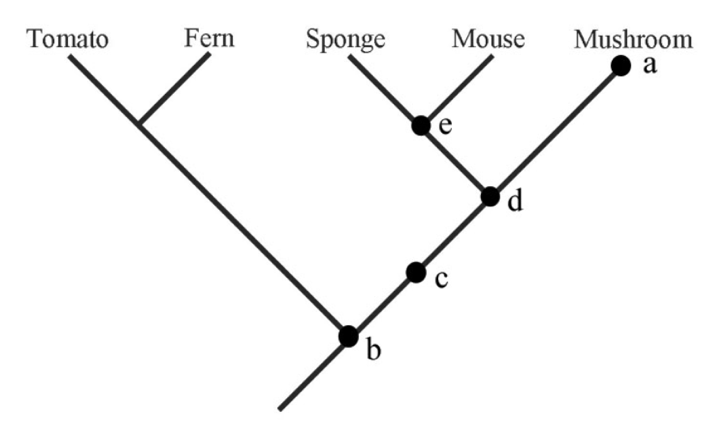
A

B

C

Yes, it depicts the same relationship. This is because in this tree, A and B are most closely related, which is the same as the previous tree, and A and B share a more recent common ancestor with C than D, which is the same as the previous tree. Students can also mention that if you just rotate the first tree around the point of the recent common ancestor of A, B, and C, the second tree is recreated (as discussed in lecture, you can rotate trees at the ancestral nodes, and the relationships stay the same).

1. Which point marks the most recent common ancestor between a fern and a sponge? (1 point)



The most recent common ancestor between fern and sponge is point b.

**Part 2: Creating a Tree Using Sequences from GenBank (9 points)**

Using MEGA, search for and include in your alignment the following 9 sequences:

* Canis lupus haplotype GW01 cytochrome b (cytb) gene, partial cds; mitochondrial
* Canis lupus familiaris cytochrome b gene, complete cds; mitochondrial
* Canis rufus cytochrome b (cytb) gene, mitochondrial gene encoding mitochondrial protein, partial cds
* Canis latrans isolate CLA\_RKW3591 cytochrome b (CYB) gene, complete cds; mitochondrial
* Canis simensis isolate CSI\_RKW2631 cytochrome b (CYB) gene, complete cds; mitochondrial
* Vulpes zerda haplotype ZMo cytochrome b (cytb) gene, partial cds; mitochondrial
* Chrysocyon brachyurus isolate SP104 cytochrome b (cytb) gene, complete cds; mitochondrial
* Urocyon cinereoargenteus cytochrome b (cytb) gene, complete cds; mitochondrial
* Crocuta crocuta cytochrome b (CYTB) gene, complete cds; mitochondrial

1. Form a hypothesis regarding relationships of the species represented in your sequences. Which species do you expect to be the outgroup? Which should be more closely related to each other? *Remember, a hypothesis is a well-educated guess based on your previous knowledge; it is not expected to be perfectly correct, but it should be well-reasoned.* (2 points)

For your hypothesis, you should include:

* which species you expect to be the outgroup (answers may vary, but make sure there is some justification; the outgroup in this case should be the hyena (Crocuta crocuta) because it does not belong in the canid group) 1 point
* which species you expect to be most closely related (answers may vary, but make sure there is some justification; the most closely related in this case should be the wolf (Canis lupus) and dog (Canis lupus familiaris) because the dog was domesticated from the wolf within the last 50,000 years). Other closely related species would be the foxes. 1 point

Align the sequences using MEGA, and save and export the alignment in MEGA format.

1. How many variables sites are in the alignment? (1 point)

387

1. How many sites are parsimony-informative? (1 point)

181

1. Explain how you could test your hypothesis you wrote above. Be explicit about your method and what you would expect regarding the support for your method (how would you test tree support based on our in-class activity?) (2 points)

You can test your hypothesis by building a tree. You should mention the type of tree you would use, which should be maximum likelihood, as it is the best method available for MEGA and accounts for multiple evolutionary options. To test the tree hypothesis, bootstrapping will give you the support for your tree (how often will you rebuild the relationships in your tree when you subset your data in different ways).

1. Test your hypothesis and paste the result below. \*\**Make sure you change the setting for Gaps/Missing Data Treatment to “Use All Sites”\*\** Does your result support your original hypothesis? Explain. (3 points)

Paste Tree: 1 point

Answers vary for the explanation of whether it supports their original hypothesis. They should explain whether their outgroup choice was supported (1 point) and whether the sister species they predicted were supported by the tree (1 point).



1. You should have used a method you are relatively confident in to build your tree for question 13. Now try a method that I recommended NOT be used in class. Paste the tree below, explain how this tree differs from the one in question 13 and why it might differ from your tree in question 13. (3 points)



Answers may vary depending on the tree they ran. Students should mention the differences between the trees, and explain that distance methods are less desirable because they summarize information instead of directly using sequence. Additionally, UPGMA assumes constant substitution rates, which is problematic because the molecular clock is not the same across lineages or genes.

1. Explain the concept of a molecular clock. (1 point)

The relationship between the number of mutations in a sequence and the amount of time of divergence of that sequence. We can time the divergence between two sequences by knowing the mutation rate of the sequences.