

1-5: Multiple choice. Choose the single best answer. (2 pts. each)

1. In a 1978 paper, "Cases in which parsimony or compatibility methods will be positively misleading", Felsenstein described analytical conditions under which maximum parsimony would find the incorrect tree with increasingly strong support as the number of characters increased, a phenomenon called "inconsistency". Which of the following types of characters you expect to be *least* prone to inconsistency in parsimony analysis?

- a. Morphological characters with two character states
- b. DNA sequence data coded with four character states
- c. DNA ambiguity codes with a total of eight character states.
- d. Amino acid data coded with 20 character states

2. Which of the following adequately defines *homologous* characters?

- a. Characters are homologous if they are more similar to each other than to any other identifiable character.
- b. Characters are homologous if they are modified versions of the same character in a common ancestor.
- c. Characters are homologous if it can be shown that there are no intermediate stages or forms found in nature.
- d. Characters are homologous if they were independently acquired via organic evolution under similar conditions of natural selection.

3. Zuckerkandl and Pauling distinguish between semantic, episemantic, and asemantic molecules. What is the critical distinction between these three classes of molecules?

- a. The degree to which they are able to convey information about the evolutionary history of the organism.
- b. The degree to which they are involved in the reproduction and fitness of the organism.
- c. The degree to which they are subject to natural selection.
- d. None of the above.

4. Which of the following is *not* an assumption shared by maximum parsimony, distance, and maximum likelihood methods?

- a. Descendants inherit their character-states from their ancestors.
- b. Characters cannot assume character-states that they hold elsewhere on the tree.
- c. Character-state changes occur independently of each other.
- d. The evolutionary process can be represented by a dichotomously branching tree.

5. In which of the following cases would you have greatest confidence that you have identified the correct phylogeny (i.e., the phylogeny that reflects the true evolutionary history of the group)?

- a. Parsimony analysis of a combined *rbcL*, and SSU rRNA dataset does not show any very long branches, and at least two of the ten clades found in the tree are almost certainly natural groups.
- b. Maximum parsimony, minimum evolution, and maximum likelihood analyses of a SSU rRNA dataset all show the same tree topology.
- c. Bootstrap analysis of an *rbcL* dataset shows at least 100% bootstrap support at every node.
- d. A variety of phylogenetic analyses of *rbcL*, SSU rRNA, *coxIII* and a morphological data set all show exactly the same tree topology.

6-10: Fill in the Blank (2 pts. each)

6. _____ is the gradual change of a population over time in the absence of speciation.
7. In cladistic (parsimony) terminology, a character that shows a unique character state in one taxon, and has another character state in all other taxa is considered to be _____, and is considered to be one type of uninformative character.
8. An _____ is a defined series of steps that can be used to solve an analytical problem.
9. _____ is a "hill climbing" algorithm that uses branch swapping to test variations on the topology of an initial starting tree.
10. The _____ uses synthetic oligonucleotide primers, a thermostable DNA polymerase, and temperature cycling to generate usable quantities of a selected region of the genome from small quantities of genomic template.

11-16: Answer these questions accurately, but briefly, in standard written English or with diagrams and calculations as appropriate. In most cases a single sentence should be adequate, and brief, but thorough, answers will be favored over longer answers.

11. For the following tree, diagram the trees that would result from nearest-neighbor-interchange on the indicated branch (5 points).

12. Compare and contrast NNI and TBR branch swapping. (5 pts.)

13. Describe a branch-and-bound search, and discuss its relationship to an exhaustive search. Include in your description a diagram of a hypothetical partial search path that illustrates the branch and bound algorithm. (10 points).

14. Imagine that you are sitting on an NSF panel considering dissertation improvement proposals. Summaries of several proposals are given below. For each of these, indicate whether or not you feel that the student has an adequate understanding of phylogenetic methods to justify funding their proposal, and indicate why you are of this opinion. (5 pts each)

a. Tony has determined 300 base pairs of the mitochondrial gene *coxII* from thirty individuals in seven isolated populations of an endangered plant in the mustard family, and would like to collect the same segment of the genome from another 200 individuals to perform paternity analysis and determine the degree of inbreeding in each population.

b. Chia-Wen has a collection of forty species of nudibranchs from Australia, and she would like to determine nuclear SSU rRNA and *coxII* sequences from each of these. She proposes to use these data in parsimony and maximum likelihood analyses to examine whether the several species that retain captured chloroplasts in their mantles constitute a monophyletic group.

c. François has DNA sequences from thirty species of the fungal component of an arctic lichen, and proposes to collect additional sequences from the algal component of the lichen. He then plans to combine these sequences into a single dataset, and perform 'total evidence' parsimony analyses with the combined data.

d. Mark has 8000 sequences of the gene *rbcL* from angiosperms, and proposes to use these sequences to perform a parsimony analysis using a branch and bound search because it is guaranteed to find the shortest possible tree.

15. Consider the article “An empirical test of bootstrapping as a method for assessing confidence in phylogenetic analysis”, by D.M. Hillis and J.J. Bull [Syst. Biol. 42:182-192 (1993)], which refers to nonparametric bootstrapping.

a. What empirical tests of bootstrapping did they perform? (5 pts.)

b. In general, what relationship did they find between bootstrap proportions and the probability that the corresponding clade is correct? (5 pts.)

c. Based on the observations of Hillis and Bull, what recommendations would you make concerning the interpretation of bootstrap proportions? (5 pts.)

16. Consider the following sequence alignment:

Critter-a	AAAAAAACCC	CCCCCGGGG	GGGGTTTTTT
Critter-b	ATAAGCACCC	CGGCTGGGAG	GGTGTTTTAT
Critter-c	AAAAGCTCAC	CGGCCGCGGG	GTTGATTTAT
Critter-d	AAACGCTCCC	CGCGCGCGAG	GGTGCTCTAT

a. Show the divergence matrix (F_{xy}) for Critter-a vs Critter-b (5 pts):

b. Calculate the p-distance and Jukes Cantor distance for Critter-b to Critter-c and Critter-c to Critter-d. Be sure to show what formula you are using for each of these distance measures, and show your calculations! (10 pts)

c. Determine the length and topology of the globally optimum tree for these four taxa under the maximum-parsimony optimality criterion. Explicitly work the tree-length algorithm we learned in class for at least two instructive characters (10 pts)

-end-