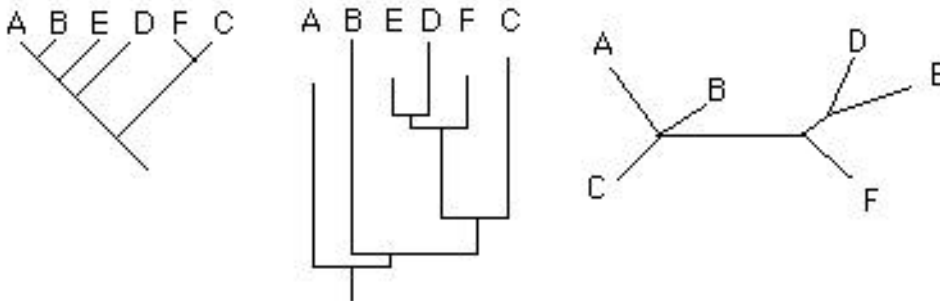


MCB/GENE/ECOL/BIOC/INSC 545
FALL 2002
FINAL EXAMINATION

- PRINT YOUR NAME LEGIBLY AT THE TOP OF EACH PAGE AND SIGN IT ON THE FIRST PAGE.
- Write your answers and show your work only in the space provided below each question. When there is a line labelled "Answer", write the answer on the line and show calculations above the line. I will not grade anything written anywhere else, and I will not grade anything that is written illegibly.
- When calculations are necessary to get an answer, you must show the calculations to get credit.
- All quantitative answers must have units (e.g. 27 substitutions/site X year).
- Approximate point values for each question are given in parentheses; these are subject to change. Partial credit will be given when appropriate.

1. (10) In each of the phylogenetic trees below in which it is possible to do so, (a) circle the coalescent (the most recent common ancestor) of genes B and C, and (b) draw a square around the root node (coalescent of all taxa).



2. (10) Evolutionary models.

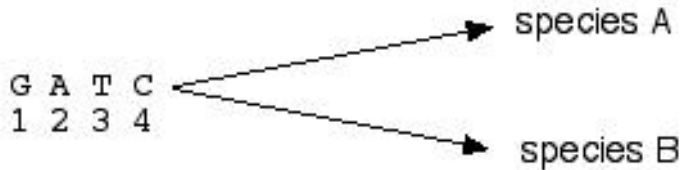
(a) The Jukes-Cantor model has only one parameter: what is it? (A couple of words or a phrase, please, not just a symbol.)

Answer _____

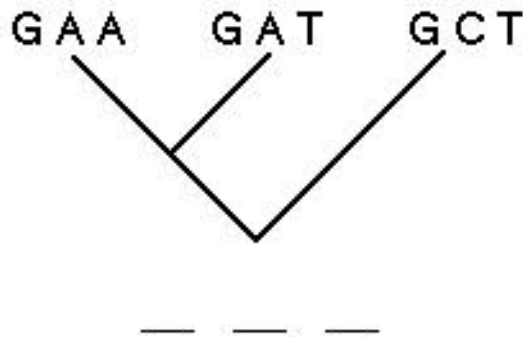
(b) Name one additional parameter or kind of parameter that is found in more complex models.

Answer _____

3. (10) On the diagram below, indicate two different kinds of multiple hits. Indicate the number of the site at which each occurs.



4. (9) Below is the sequence of the second codon in a gene in three organisms, and the cladogram describing their relationships. Write the most parsimonious reconstruction of the ancestral sequence at the base of the tree. If a site is ambiguous, write both possibilities separated by /.



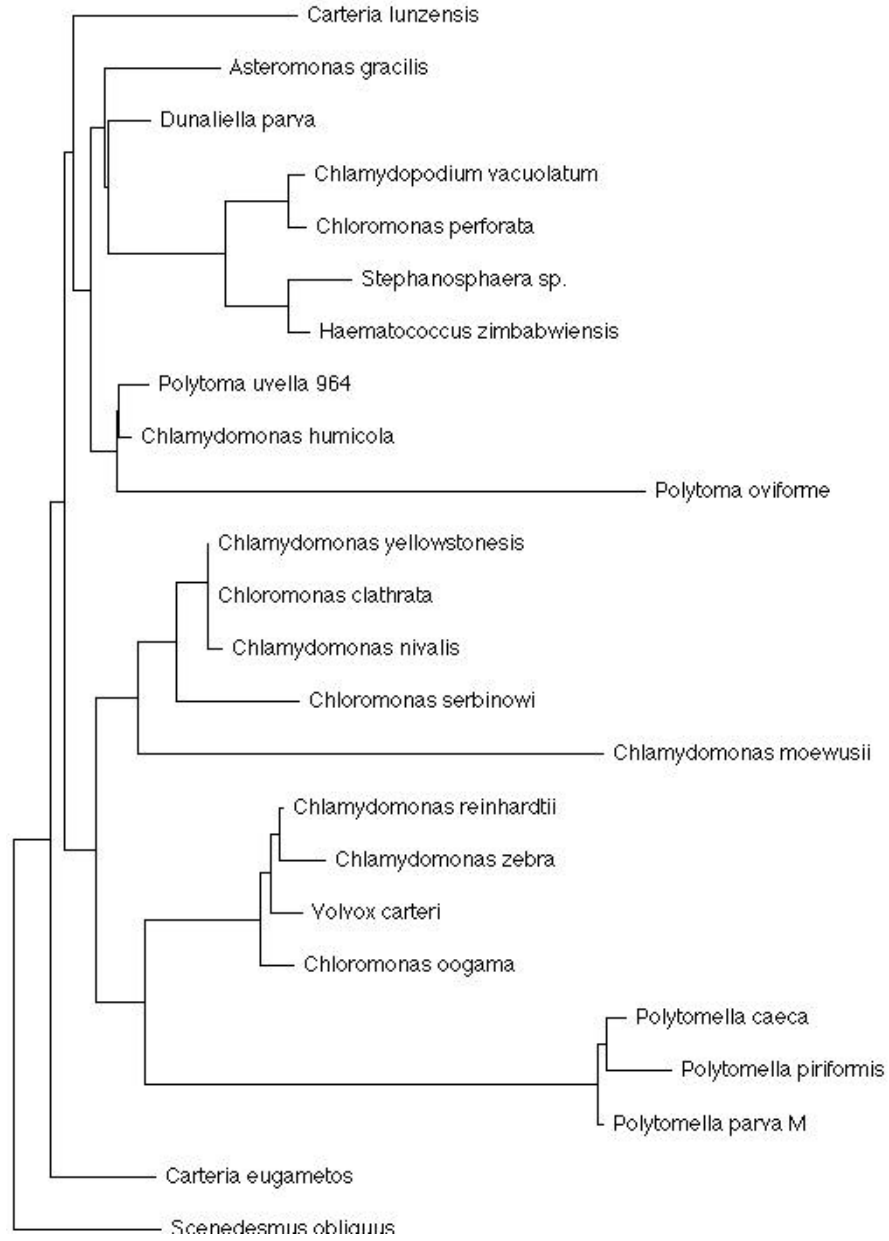
5. (16) The tree on the next page was made with *Rrn18* sequences by neighbor-joining with a maximum likelihood evolutionary model. *Scenedesmus* is the outgroup.

(a) Is the genus *Chlamydomonas* a clade? Circle your answer: yes no

(b) *Polytoma* and *Polytomella* are non-photosynthetic algae. How many times was photosynthesis lost in this group of algae? _____ Put an X on each branch where photosynthesis was lost.

(c) Is the rate of evolution of this gene constant in this group of algae? yes no

Explain your answer in one sentence.



6. (5) Phillipe calculated the nucleotide diversity of the protist *Paramecium aurelia* as = 1.14. How can you tell that his calculation is wrong? (One sentence, please.)

7. (20) George sequenced the alpha amylase gene from a human and a robin. He found that there are 31 synonymous differences and 4 nonsynonymous differences between the genes of the two species. For each of the following, briefly explain why it can, or cannot, explain the larger number of synonymous differences:

- (a) differences in mutation rate between synonymous and nonsynonymous sites
- (b) differences in population size
- (c) differences in effective population size
- (d) differences in directional selection
- (e) differences in balancing selection (I didn't discuss this, but let's see what you can do with your intuition)

8. (20) Consider a population of plants in which there are two alleles of *mfg* (my favorite gene); *mfg1* and *mfg2* have frequencies 0.9 and 0.1 respectively.

- (a) What is the expected heterozygosity of this gene?
- (b) What would be the observed heterozygosity if the population is random mating?
- (c) What can you say about the observed heterozygosity if the population is inbred?

(part d is on next page)

(d) What would you have predicted the expected heterozygosity would be if all mutations are selectively neutral, the effective population size is 10^5 , and the mutation rate is 10^{-7} ?

ALL THE EQUATIONS THAT YOU WILL NEED
(AND MAYBE SOME YOU WON'T)

$$h = 2pq = 1 - (p^2 + q^2)$$

$$h = 1 - \sum_{i=1}^m x_i^2 = 1 - (x_1^2 + x_2^2 + \dots + x_m^2)$$

$$= \sum_{ij} x_i x_j p_{ij}$$

$$K = -\frac{3}{4} \ln\left(1 - \frac{4}{3} d\right)$$

$$K = -\frac{19}{20} \ln\left(1 - \frac{20}{19} d\right) - \ln(1 - d)$$

$$E = K/2T$$

$$H = \frac{4N_e u}{1 + 4N_e u} \quad 4N_e u =$$

$$\frac{4N_e u}{1 + (4/3)4N_e u} \quad 4N_e u =$$

$$F = \frac{1 - e^{-2N_e s / N}}{1 - e^{-4N_e s}}$$

$$E = 2NuF$$