Name:

Bioinformatics Take Home Test #6

Due Date Wednesday 11/16/2016 before class

(This is an open book exam based on the honors system -- you can use notes, lecture notes, online manuals, and text books. ***Teamwork is not allowed*** *on the exams*, write down your own answers, do not cut and paste from webpages. If your answer uses a citation, give the source of the quoted text.)

MORE THAN ONE ANSWER MAY BE CORRECT!

1. Which of the following can be used to root the tree of life?
A. Rock
B. Virus
C. Random sequence
D. Unrelated sequence
E. Paralogs that diverged Pre-LUCA
F. Composition of reconstructed sequences that reflect a more primitive state (symplesiomorphy) closer to the early expansion of the genetic code.

2. How many rooted and unrooted trees with different topology are possible for trees with 4, 5, and 6 OTUs?

3. In a phylogenetic tree, OTU can be synonymous with which of the following term(s)?
A. Leaf
B. Taxa
C. Terminal Node
D. Species
E. All of the above

4. Given a specific topology of one unrooted tree with 6 OTUs, how many rooted phylogenies are possible that conform to the given topology.
A) 6,   B) 7,   C) 9,   D) 105,   E) 945

    5. How many rooted trees with 4 taxa can be collapsed down into a single unrooted tree?
    A) 1,   B) 2,   C) 3,   D) 4,   E) 5,   F) None of the above

    6. What happens when a dataset is not first aligned using a Multiple Sequence Alignment program?
    A. An error is produced and the program refuses to calculate the tree
    B. A horrible tree is produced, sometimes with 100% bootstrap support
    C. The program alerts you to the problem and suggests that you do an alignment
    D. The program calculates a tree, but it is a random tree with 0% bootstrap support

    7. Which of the following are approaches to build trees from of alignments?
    A. Neighbor Joining
    B. UPGMA (Although this one should NEVER be used)
    C. Maximum Likelihood
    D. Bayesian Inference
    E. Parsimony

    8. True/False All tree building programs (excepting UPGMA which is awful) have their advantages and disadvantages over other programs, which is why more than one should always be used to analyze any dataset, to verify that the findings are independent of the program used.

    9. What does an internal node on a phylogenetic tree represent?
    A. Time
    B. The common ancestor
    C. Rate of evolution

D. Total evolution

E. A species or sequence present today

10. Which of the following is NOT a proper taxonomic groups (i.e., they do NOT form a clade in the "traditional*" version of tree of life)?

A. Prokaryotes

B. Fish

C. Protists

D. Algae

11. Which of the follow term(s) is often used as a synonym for the splits of a tree?

A. Interior Node

B. Bifurcation

C. Branch

D. Bipartition

12. Which of the following is true of the gap penalty in alignments?

A. With a high penalty, there are too few gaps, making the sequences clump together and poorly aligned.

B. If the penalty is too low, there are too many gaps, creating homoplasies

C. The default parameters are set to produce the ideal number of gaps for the majority of sequences, so that the individual alignment columns are likely to contain homologous positions.

13. In alignment programs, insertions are best dealt with by doing which of the following?

A. Adding gaps

B. Removing gaps

C. Correcting for multiple substitutions

D. Bootstrapping the dataset

14. The place of the root in the tree of life was first determined using which molecule?

A. Inteins

B. The signal recognition particle and receptor

C. ATPsynthase catalytic and non-catalytic subunits

D. rRNA

E. None of the above

15. A. Is the following tree rooted or unrooted? ((((G1E, G1S), G1T), ((G2E, G2T), G2S)), (GA, GB))

B. Draw it:

16. Draw all possible unrooted 4 taxa trees with the OTUs 1, 2, 3, and 4 (we are only interested in different topologies).

17. True/False Clustal and many other alignment programs treats gaps inserted at the beginning and end of a sequence differently from gaps inserted into the middle of a sequence, because doing so creates better alignments.

18. How did the sequencing of rRNA genes revolutionize biology?
A. It was the first biomolecule to place microorganisms onto the tree of life
B. Led to the discovery that Archaea are not a type of Bacteria
C. Provided molecular datasets with a large number of characters to analyze
D. Rooted the tree of life

19. According to the currently favored version of the tree of life, which are the closest relative of the Archaea?
A. Bacteria
B. Viruses
C. Eukarya
D. Archaeabacteria

20. Which comes first in phylogenetic analysis?
A. Tree evaluation
B. Determination of substitution model
C. Alignment
D. Tree building
E. Compilation of sequence dataset

21. True/False: Bootstrap analysis can be used for more than just neighbor joining tree calculations.

22. How are evolutionary relationships between organisms represented?
A) web-like diagrams
B) tree-like diagrams
C) Venn diagrams
D) Klenow diagrams

23 True/False. Phylogenetic analysis is an inference of evolutionary relationships between organisms.

24. Which tree is identical (only with respect to topology) to the tree depicted bellow?

```
--+--+----------A
  |   `--+--+-----B
  |      |    `--+--C
  |      |       `--D
  |      `--------E
  `-------------F
```

a) --+--F
     `--+--A
        `--+--E
           `--+--B
              `--+--D
                 `--C

b) --+--F
     `--+--A
        `--+--B
           `--+--E
              `--+--D
                 `--C

c) --+--F
     `--+--E
        `--+--B
           `--+--A
              `--+--D
                 `--C

d) --+--E
     `--+--A
        `--+--F
           `--+--B
              `--+--D
                 `--C
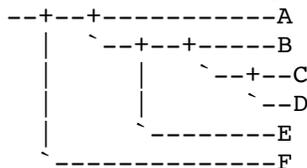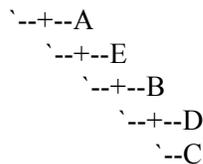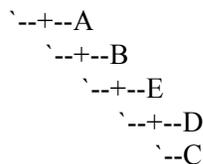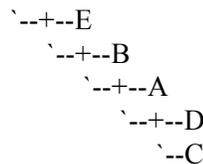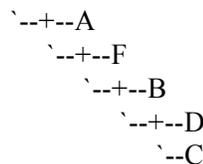
25. Which approach finds the "true" tree?
A. Parsimony
B. Maximum Likelihood
C. Bayesian Inference
D. Neighbor Joining
E. All of the above
F. None of the above

26. True/False Rotating branches of a phylogenetic tree around a node changes the meaning of that tree.

27. Bootstrap values belong to which of the following?
A. Taxa, Species, or OTUs
B. Leaves
C. Splits
D. The ancestral sequence
F. None of the above

Use the following figure to answer the following questions:
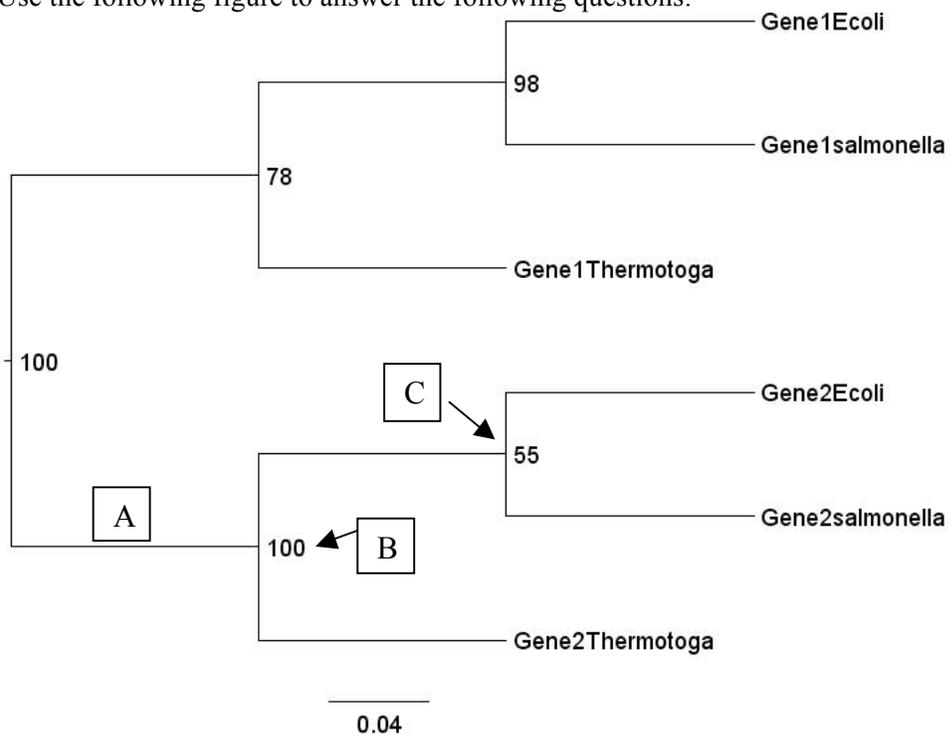
Numbers give bootstrap support values.

28. What are the lettings A, B, and C pointing to?
A.
B.
C.

29. Gene1 and Gene2 are?
A. Orthologs
B. Paralogs
C. Xenologs
D. Can't tell, other than some type of homolog

30. True/False The number 98 refers only to the confidence of Gene1Ecoli Gene1salmonella going together

31. True/False Given the above tree, Gene2Ecoli could group with Gene2Thermotoga in 45% of the trees calculated from bootstrap samples.

Extra credit:
1. Draw all possible rooted versions of the tree given below. (Note: this is a different question than question 16, here we ask only about trees that have the same topology as the one depicted below)

A   C

B
   D