1. What kind of search does BlastX perform? Why would you use BlastX?
1. Describe a scoring matrix (i.e. PAM250).
2. A blast score has an E-value of e^{-3}, is this statistically significant? Is it biologically significant?
3. I want to search for EST sequences, should I search the DNA nr database?
4. BlastP replaced part of my sequence with X's, why did it do that?
5. Name of few databases you would search using BlastN.
6. Describe the type of search that is performed by BlastN, BlastP, and BlastX?
7. What does an E-value from blast mean?
8. Define homologous, orthologous, and paralogous. Draw phylogenic trees to show the distinction of orthologous and paralogous.
9. You have a protein coding sequence and want to compare it to the database. Should you search the protein or DNA sequence? Why?
10. You are told a sequence alignment has a raw score of 1000. Is this a significant match?
11. What does the filter do in a Blast search?
12. The bit score includes the parameters k and lamda (λ). Why are these parameters in the equation to calculate bit scores?
13. The E-value score includes parameters m and n. Why are these parameters in the equation to calculate E-values?
14. What is Markov Chain Monte Carlo, and how is it used in maximum likelihood calculations?
15. What is bootstrap analysis?
16. Does a heuristic search guarantee that you will find the most parsimonious tree?
17. Does a branch and bound search search all possible trees?
18. What is an exhaustive search?
19. What is one method to test the significance of branching order in a phylogenetic analysis?
20. The current size of the NR protein database is 680,984,053 and the PDB is 3,816,875. Which should you use in a Blast search to have the most sensitivity in detecting homologs with known 3D structures? Why?
21. BlastX translates a DNA sequence into the 6 reading frames and searches each one against the protein NR database. BlastP takes a known translation and searches against the protein NR database. Do you expect both searches to have the identical E-value? Why or why not?
22. Describe the type of search done by BlastX, BlastP and BlastN. What type of data is the query and the database?
23. Describe DNA sequencing.
1. What is the difference between a BAC and Plasmid?
2. What are two methods for sequencing genomes?
3. What is hierarchical sequencing?
4. What is whole genome shotgun sequencing?
5. Describing a tilling path.
6. What is a phred score?
7. What is the difference between a scaffold and contig sequence?
8. What is the difference between dNTP and ddNTP?

1. My gene is the fastest evolving gene between two organisms I am studying, it is therefore been subjected to positive Darwinian selection. Describe why this statement is true or false.
2. Does the neutral model of molecular evolution allow for deleterious mutations?
3. Define two fold and four fold degenerate sites.
4. Define a synonymous substitution.
5. Define a nonsynonymous substitution
6. What is codon bias?
7. Do two fold degenerate sites evolve at the same rate as pseudogenes?
8. Describe one way to test for positive Darwinian selection.
9. Describe the difference between the neutralist model of molecular evolution and the selectionist model of molecular evovlution.
10. Between two closely related species, my gene has 2 nonsynonymous and 1 synonymous difference, it is therefore been subjected to adaptive evolution.
11. What are three possible fates for duplicated genes?
12. Calculate the dn/ds ratio for the following alignment (you will be given a genetic code);

```
Sequence 1        GAC   ACA   GCG   GTT
Sequence 2        GCC   ACT   TCG   GTT
```

13. Describe how the modeltest program selects the best model of DNA evolution for a set of aligned DNA sequences.
14. For a set of aligned sequences with a NJ tree, the likelihood of the Jukes-Cantor model is –100 while the likelihood of the F81 model is –90. Which model should you use in a likelihood phylogenetic analysis (you will be given the parameters for both models) Show any calculations use to make your decision.
15. What is a likelihood ratio test?