

M. Sc. DEGREE EXAMINATION, APRIL 2011
BIOINFORMATICS
FOURTH SEMESTER

COURSE : CORE
PAPER : MOLECULAR EVOLUTION
TIME : 3 HOURS

MAX. MARKS: 100

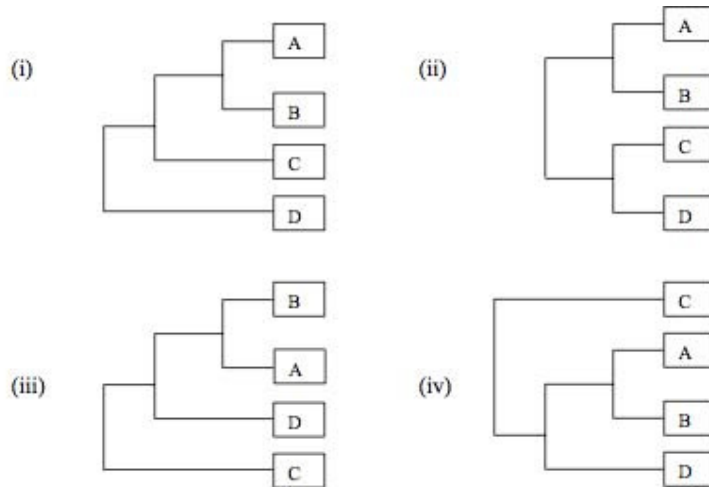
SECTION – A

ANSWER ALL QUESTIONS

(20 X 1=20)

1. Molecular evolution does not reflect
 - a. Species divergence
 - b. convergent evolution
 - c. natural selection
 - d. Neutral mutation
2. A group of organisms that is defined by a shared primitive character is
 - a) monophyletic
 - b) paraphyletic
 - c) polyphyletic
 - d. ecophylaetic
3. Molecular phylogenies are usually calculated as unrooted trees. Correct/incorrect
4. Which of the following is a good approach to create diversity on which selection can act?
 - a) puromycin linked to the 3'end of the mRNA
 - b) synonymous substitutions
 - c) recombination between divergent homologs
 - d. none of the above
5. Molecular clock concept was first proposed by
 - a. Peterson et al
 - b. Emile Zuckerkandl and Linus Pauling
 - c. Peterson and Linus Pauling
 - d. Drummond and Ayala
6. Likelihood based phylogenetic reconstruction aims to find the tree under which the data set (e.g., aligned sequences) are most probable. To calculate this probability, a model describing the evolutionary process is used. Which model assumption(s) would correspond to a normal parsimony analysis?
 - a) Gaps (-) in the sequence alignment are treated as a 5th nucleotide or as the 21st amino acid.
 - b) All types of substitutions occur with the same probability.
 - c) The ratio between probability for transition and transversion is described by a parameter kappa that is estimated from the data
 - d) All answers are incorrect
7. A new mutant allele has just arisen in a population. Which statement is true?
 - a. If the mutant is neutral with respect to the original allele, there is a 50% probability that the mutant allele will replace the original allele.
 - b. It is very likely to disappear in a few generations due to random drift.
 - c. It will only become fixed in the population if there is a strong selective advantage.
 - d. If the mutant allele reaches a frequency of 50%, it will almost always go on to fixation

8. The following four trees are unrooted and branch lengths are not drawn to scale. Which statement is correct?



- a. All four trees are non-equivalent.
 b. All four trees are equivalent.
 c. Only trees (i) and (iii) are equivalent.
 d. Only trees (iii) and (iv) are equivalent.
9. The methods adopted to evolutionary relationship among the unrelated sequences is
- a. Maximum parsimony b. Distance method
 c. Maximum likelihood method d. none of the above
10. Homoplasy refers to
- a. occurrence of the same sequence change in more than one branch of the tree.
 b. occurrence of the different sequence change in more than one branch of the tree
 c. occurrence of the same sequence change in one branch of the tree
 d. All the above
11. The programs in PHYLIP can be used to:
- a) Estimate phylogenies using the Parsimony method
 b) Bootstrap resampling of your data
 c) Estimate phylogenies using the Fitch-Margoliash distance matrix method
 d) All are correct
12. The Unweighted Pair Group Method with Arithmetic Mean (UPGMA) is
- a. Maximum parsimony b. Distance method
 c. Maximum likelihood method d. The Neighbor-joining Method

13. You do a databank search using FASTA with an amino acid sequence as query. The only reported match has an E-value of 0.5. What does this mean for the homology of the two sequences?
- this proves (beyond reasonable doubt) that the target sequence is not homologous to the query
 - the target sequence is a candidate for a homologous sequence, but an E-value of 0.5 does not prove homology
 - This proves (beyond reasonable doubt) that the two sequences are homologs.
 - none of the above
14. Pick the best answer that describes Maximum Likelihood analysis.
- Finds the tree that explains sequence data with minimum number of substitutions.
 - Given a model for sequence evolution, find the tree under which the dataset has the highest probability.
 - Measures how much support for a partition is in your data given a method of analysis.
 - All of the above.
15. Clustalx is a program that allows calculating multiple sequence alignments. What other things can you do with this program?
- Calculate distance matrices (with or without correction for multiple substitutions), neighbor joining trees, and perform bootstrap analyses
 - Calculate distance matrices, parsimony trees, and perform bootstrap analyses
 - Calculate distance matrices, neighbor joining trees, do bootstrap analyses and evaluate user trees.
 - none of the above
16. C-value refers to the amount of DNA contained within a
- diploid nucleus
 - haploid nucleus
 - both above
 - none of the above
17. Variations in GC ratio within the genomes made a mosaic-like formation with islet regions called
- chromophore
 - nucleophore
 - Isochores
 - none of the above
18. "C-value enigma is
- variation in the amount of coding DNA
 - variation in the amount of haploid DNA
 - variation in the amount of non-coding DNA
 - both a and b correct
19. The "metabolic-rate" hypothesis is
- DNA repair in organelle DNA
 - DNA damage due to oxidative metabolism
 - both a and b
 - none of the above

20. Concerted evolution is a process that may explain the observation that
- DNA of one species are closely related to same type of DNA family in other species
 - DNA of one species are closely related to different type of DNA family in other species
 - DNA of one species are closely related to same type of DNA family in same species
 - both b and c correct

SECTION – B

ANSWER ANY FOUR QUESTIONS. EACH ANSWER SHOULD NOT EXCEED 500 WORDS. ALL ANSWERS CARRY EQUAL MARKS. DRAW DIAGRAMS WHEREVER NECESSARY (4 X 10 = 40)

21. Explain the Markov model for nucleic acid substitution
22. Briefly explain the substitution rate of RNA viruses.
23. Briefly draw and explain the Universal Phylogeny tree.
24. Explain about the UPGMA method
25. What do you mean by Exon shuffling and what way is it related to protein evolution?
26. Explain in details about Intron –early and intron – late hypothesis
27. Write an account of advantage and disadvantage of non –universal genetic codes on translation

SECTION – C

ANSWER ANY TWO QUESTIONS. EACH ANSWER SHOULD NOT EXCEED 1200 WORDS. ALL ANSWERS CARRY EQUAL MARKS. DRAW DIAGRAMS WHEREVER NECESSARY (2 X 20 = 40)

28. What is nucleotide substitution? How do you calculate the rate of substitution?
What events are due to nucleotide substitution?
29. Describe the different methods involved in Distance matrix method
30. Define gene duplication. Describe in detail types of gene duplication that occur in the chromosome and its importance in evolution.
31. Describe in detail about the compositional organization of the Human genome.
