

**STELLA MARIS COLLEGE (AUTONOMOUS) CHENNAI –600 086**  
**(For candidates admitted from the academic year 2011 – 12 & thereafter)**

**SUBJECT CODE: 11BI/PC/BI24**

**M. Sc. DEGREE EXAMINATION, APRIL 2015**  
**BIOINFORMATICS**  
**SECOND SEMESTER**

**COURSE : MAJOR – CORE**  
**PAPER : BASIC BIOINFORMATICS**  
**TIME : 3 HOURS** **MAX. MARKS: 100**

**SECTION – A**

**ANSWER ALL THE QUESTIONS (30x1=30)**

1. UPGMA stands for \_\_\_\_\_.
2. Multiple sequence alignment tools are \_\_\_\_\_ and \_\_\_\_\_.
3. A statistical technique that tests the sampling errors of a phylogenetic tree is known as \_\_\_\_\_.
4. The study of evolution of a genetically related group of organisms is known as \_\_\_\_\_.
5. A tree is composed of \_\_\_\_\_ and \_\_\_\_\_.
6. The presence of same genes in different species is known as \_\_\_\_\_.
7. Margaret Dayhoff developed the first protein sequence database called
  - a) SWISS PROT
  - b) PDB
  - c) Atlas of protein sequence and structure
  - d) Protein sequence databank
8. Literature databases include
  - a) MEDLINE and PubMed
  - b) MEDLINE and PDB
  - c) PubMed and PDB
  - d) MEDLINE and PDS
9. Which of the following is a protein sequence database?
  - a) DDBJ
  - b) EMBL
  - c) GenBank
  - d) PIR
10. GenBank, the nucleic acid sequence database is maintained by
  - a) Brookhaven laboratory
  - b) DNA database of Japan (DDBJ)
  - c) European Molecular Biology laboratory (EMBL)
  - d) National Centre for Biotechnology Information (NCBI)

11. Submission to GenBank are made using
  - a) BankIt and Sequin
  - b) BankIt and BankIn
  - c) Sequin and BankIn
  - d) Entrez
12. The information retrieval tool of NCBI GenBank is
  - a) Entrez
  - b) STAG
  - c) SeqIn
  - d) text search
13. Which of the following is a sequence alignment tool
  - a) BLAST
  - b) PRINT
  - c) PROSITE
  - d) PIR
14. When performing a database search, what is the definition of an E-value?
  - a) The chance that a random sequence could achieve a better score than the query.
  - b) The chance that a homologous sequence could achieve a similar score to the query.
  - c) The chance that a random sequence could achieve a worse score than the query.
  - d) The chance that a homologous sequence could achieve a better score than the query.
15. Well-conserved regions in multiple sequence alignments:
  - a) Reflect areas of structural importance.
  - b) Reflect areas of functional importance.
  - c) Reflect areas of both functional and structural importance.
  - d) Reflect areas likely to be of functional and/or structural importance.
16. Why are colour schemes important in creating and analysing sequence alignments?
  - a) They look pretty
  - b) To make clearer printouts and presentations
  - c) To allow you to distinguish conserved residues and residue groups more easily
  - d) To allow you to detect active sites of proteins
17. Is consensus important in secondary structure prediction?
  - a) Yes, because consensus improves one's confidence in the results
  - b) No, because some algorithms predict better than others
  - c) No, because these algorithms are all built from the same datasets and thus make the same errors
  - d) Yes, because if one obtains a consensus then the results are correct
18. Bootstrapping repeats the phylogenetic analysis several times, each time reshuffling the columns of the initial alignment, in order to:
  - a) Build a consensus tree, where the number of times each branch reforms is counted and used to estimate its probability
  - b) Generate a random model from which to benchmark phylogenetic data
  - c) Produce a graphical representation of the tree
  - d) Assess the probability that the sequences in the alignment are unrelated

- 19 SRS is:
- a) Website for sequence similarity searches
  - b) Website specialising in mapping mutations related to human disease
  - c) Website that indexes many biological databases and is searchable by keywords
  - d) Website for protein family database searches
20. A Single Piece of Information in a database is called
- a) File
  - b) Field
  - c) Record
  - d) Data set
21. Which of the following is a nucleotide sequence database?
- a) EMBL
  - b) SwissProt
  - c) TrEMBL
  - d) PROSITE
22. BLOSUM matrices are used for
- a) Multiple sequence Alignment
  - b) Pairwise sequence Alignment
  - c) Phylogenetic Analysis
  - d) All the above
23. Clustal W
- a) Multiple sequence alignment tool
  - b) Protein secondary structure predicting tool
  - c) Data retrieving tool
  - d) Nucleic acid sequence analysis tool
24. Phylogenetic relationship can be shown by
- a) Dendrogram
  - b) Gene Bank
  - c) Data retrieving tool
  - d) Data search tool
25. PRINTS are software used for
- a) detection of genes from genome sequence
  - b) detection of tRNA genes
  - c) prediction of function of a new gene
  - d) Identification of functional domains/motifs of proteins
26. Rooted tree
27. BLAST
28. FASTA
29. Local alignment
30. NCBI

**SECTION - B**

**Answer any THREE of the following; each answers not exceeding 500 words. Draw diagram wherever necessary. (3 X 10 =30)**

31. What is sequence retrieval system?
32. Explain pairwise alignment in detail?
33. Detail on advanced BLAST search.
34. Explain Integrated MSA Resources.
35. Give a detailed account on evolutionary analysis.

**SECTION – C**

**Answer any TWO of the following, each answer not exceeding 1200 words. Draw diagram wherever necessary. (2 X 20 = 40)**

36. Discuss on the different types of biological databases.
37. Explain Needleman-Wunsch algorithm & Smith-Waterman algorithm.
38. Explain BLAST algorithm. Discuss the relationship between E & P values and the evaluation of output of BLAST.
39. Explain the different methods of tree construction.

**\*\*\*\*\***