

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 2014
BIOINFORMATICS
SECOND SEMESTER

COURSE : MAJOR – CORE
PAPER : BASIC BIOINFORMATICS
TIME : 3 HOURS

MAX. MARKS: 100

SECTION – A

ANSWER ALL THE QUESTIONS

(30x1=30)

1. Linear branched genomes are found in
 - A. archael genomes
 - B. prokaryotic nuclear DNA
 - C. organelle genomes
 - D. eukaryotic nuclear DNA
2. The largest known human gene is
 - A. myoglobin
 - B. dystrophin
 - C. kerat in
 - D. rennin
3. The homologs resulting from whole genome duplication are
 - A. orthologs
 - B. xenologs
 - C. paralogs
 - D. ohnologs
4. Functional genomics involves the study of
 - A. genes
 - B. genic and non-genic elements
 - C. proteins
 - D. SNPs
5. Messenger RNA
 - A. is formed in the 3' to 5' direction
 - B. is found only in prokaryotes
 - C. contains uracil instead of thymine
 - D. is synthesized using the enzyme ligase
6. The penalty score for gap opening and extension in BLAST is
 - A. the same for both
 - B. higher for gap extension than opening
 - C. always zero
 - D. higher for gap opening than extension

7. Regions of genes that do not code for a protein are called
- A. primers
 - B. introns
 - C. okazaki fragments
 - D. exon
8. The default scoring matrix used in Blastp is
- A. BLOSUM85
 - B. BLOSUM62
 - C. PAM70
 - D. PAM30
9. Identify the derived database in NCBI
- A. GenBank
 - B. PubMed
 - C. RefSeq
 - D. EMBL
10. Higher PAM or lower BLOSUM matrices are used to obtain
- A. More closely related sequences
 - B. hydrophilic amino acids
 - C. More divergent sequences
 - D. substitutions of rare amino acids
11. Word size options available in NCBI nucleotide blast is
- A. 2, 3
 - B. 6, 12, 21
 - C. 1, 2
 - D. 7, 11, 15
12. The tool to search EBI databases is
- A. Entrez
 - B. Fasta
 - C. SRS
 - D. Getentry
13. HSP's in blast is used to denote
- A. Heat Shock Proteins
 - B. High-scoring Segment Pairs
 - C. Heat Shock Pairs
 - D. High Scoring Proteins
14. Identify the structure database in NCBI
- A. PubMed
 - B. OMIM
 - C. MeSH
 - D. MMDB
15. Number of matches as good as the observed one that would be expected to appear by chance in a database of the size probed is called
- A. Bit Score (S')
 - B. K-value
 - C. E-value
 - D. P-value
16. Algorithm developed by dynamic programming in 1970 to completely align pairs of sequence using matrix was given by
- A. Needleman and Wunsch
 - B. Smith and Waterman
 - C. Henikoff and Henikoff
 - D. Pearson and Lipman

17. The second step in hierarchical method of multiple sequence alignment
- A. Create a guide tree
 - B. Progressive alignment
 - C. Pairwise alignment
 - D. Amino acid weight matrices
18. Branch length in evolutionary trees indicate
- A. Number of changes that have occurred in the branch
 - B. Relationship between taxa
 - C. Common ancestor and all of their descendents
 - D. Common ancestor of all taxa
19. Phylogenetic method employed when there is a strong sequence similarity
- A. Maximum Parsimony
 - B. Distance
 - C. Bootstrapping
 - D. Maximum likelihood
20. BLOCKS & PRINTS is linked with
- A. Multiple Alignment
 - B. Structure Prediction
 - C. Ligand Binding
 - D. E -Score DBs
22. BOOT Strap :
- A. Data Curation
 - B. Data Modeling
 - C. Data Duplication
 - D. Data Compliance
23. Most commonly used algorithm in bioinformatics sequence alignment is
- A. Greedy algorithm
 - B. Brute-force algorithm
 - C. Dynamic programming algorithm
 - D. Branch and bound algorithm
24. The process of extracting patterns from data is called
- A. Dynamic programming
 - B. Artificial neural network
 - C. Gene extraction
 - D. Data mining
25. Global sequence alignment algorithm was proposed in 1970 by
- A. Temple Smith and Michael Waterman
 - B. Saul Needleman and Christian Wunsch
 - C. Henikoff and Henikoff
 - D. Thompson and Pearson

26. Sequence alignment tool FASTA works on
- A. Dynamic programming method
 - B. Monte Carlo
 - C. Greedy
 - D. Heuristic method
27. The default word size in Blastn is
- A. 3
 - B. 7
 - C. 11
 - D. 2
28. PAM250 is used to detect
- A. Divergent sequences
 - B. Closely related sequences
 - C. Substitutions involving hydrophobic amino acids
 - D. Substitutions involving hydrophilic amino acids
29. Lower E-value and high score indicates
- A. Most similar sequence
 - B. Highly divergent sequence
 - C. Sequence with maximum substitutions
 - D. Sequence length is short
30. Artificial neural network technique is an inspiration from the examination of
- A. Computer chips
 - B. Central nervous system
 - C. Road traffic network
 - D. Mobile phone network system

SECTION - B

Answer any THREE of the following; each answers not exceeding 500 words. Draw diagram wherever necessary.

(3 X 10 =30)

- 31. Significance of E-value and word size
- 32. PSI Blast
- 33. BLOSUM matrices
- 34. Markov models
- 35. Hidden Markov model
- 36. Protocols in Phylogenetic Tree Construction

SECTION – C

Answer any TWO of the following, each answer not exceeding 1200 words. Draw diagram wherever necessary.

(2 X 20 = 40)

- 37. What is BLAST and explain its variants and uses
- 38. Explain the methods employed for multiple sequence alignment with its pros & cons
- 39. Write an essay on Sequence Submission tools and their advantages.
- 40. Elaborate : Phylogenetic Analysis and its implications
