UPES
End Semester Examination, May 2024

Course: Computational Biology and Bioinformatics   Semester   : IV
Program: B.Sc. Microbiology   Duration   : 3 Hours
Course Code: HSMB2010   Max. Marks: 100

Instructions: Attempt all the questions

S. No.	Section A Short answer questions/ MCQ/T&F (20Qx1.5M= 30 Marks)	Marks	COs
Q 1	Application of bioinformatics include a) Data storage and management b) Drug designing c) Understand relationships between organisms d) All of the above	1.5	CO1
Q 2	What are literature databases. Given an example.	1.5	CO1
Q 3	Which of the following database is a gene expression database a) GEO b) MMDB c) DDBJ d) d) EMBL	1.5	CO1
Q 4	Which of the following is the replacement of a single amino acid in the primary structure of a protein with another single amino acid, which is accepted by the processes of natural selection. a) PAM b) BLOSUM c) Pairwise d) Multiple	1.5	CO1
Q 5	What is a FASTA file format?	1.5	CO1
Q 6	Which of the following is an example of Homology and similarity tool? a) BLAST b) RasMol c) EMBOSS d) PROSPECT	1.5	CO2
Q 7	What is the significance of scoring matrices?	1.5	CO2
Q 8	What do you mean by unrooted phylogenetic tree?	1.5	CO2
Q 9	"Both rooted and unrooted trees can be either bifurcating or multifurcating". a) True	1.5	CO2

	b) False		
Q 10	Which of the following statements is FALSE? a) In bioinformatics, the BLOSUM (BLOcks SUbstitution Matrix) matrix is a substitution matrix used for sequence alignment of proteins. b) BLOSUM matrices are used to score alignments between evolutionarily divergent protein sequences. c) All BLOSUM matrices are based on observed alignments; they are not extrapolated from comparisons of closely related proteins like the PAM Matrices. d) None of the above	**1.5**	**CO2**
Q 11	Proteomics is the study of a) Set of proteins b) Set of proteins in a specific region of the cell c) Entire set of expressed proteins in a cell d) None of these	**1.5**	**CO3**
Q 12	What is the approximate genome size of Arabidopsis?	**1.5**	**CO3**
Q 13	Which types of interactions are present in the tertiary structure of protein	**1.5**	**CO3**
Q 14	How many genes does the human genome contain?	**1.5**	**CO3**
Q 15	What is the difference between PAM and BLOSUM?	**1.5**	**CO3**
Q 16	Which of the following does not affect the stability of an α-helix? a) Electrostatic repulsion b) Bulkiness c) Interaction between R groups spaced three residues apart d) Occurrence of alanine and glycine residues	**1.5**	**CO4**
Q 17	Which of the following is not true about secondary protein structure? a) The hydrophilic/hydrophobic character of amino acid residues is important to secondary structure b) The ability of peptide bonds to form intramolecular hydrogen bonds is important to secondary structure c) The alpha helix, beta pleated sheet and beta turns are examples of protein secondary structure d) The steric influence of amino acid residues is important to secondary structure	**1.5**	**CO4**
Q 18	Which of the following is also known as "Dihedral angles"? a) Right angles b) Obtuse angles c) Acute angles d) Torsion angles	**1.5**	**CO4**
Q 19	Ramachandran plot can be used to predict which of the following structure? a) Quaternary structure	**1.5**	**CO4**

	b) Tertiary structure c) Primary structure d) Secondary structure		
Q 20	"Left-handed α-helix allowed region" is present in which of the following quadrants of Ramachandran plot? a) Fourth quadrant b) Third quadrant c) Second quadrant d) First quadrant	1.5	CO4
	Section B **(4Qx5M=20 Marks)**		
Q 1	Explain the BLAST tool at NCBI. What is the significance of the E value or expected value resulting from a blast?	5	CO1
Q 2	Compare pairwise and multiple sequence alignment	5	CO2
Q 3	Write a short note on diversity of genomes	5	CO3
Q 4	With the help of a neat and labelled diagram, explain energy minimizations and evaluation by Ramachandran plot	5	CO4
	Section C **(2Qx15M=30 Marks)**		
Q 1	Case Study: The study involves the comparative analysis of genomic data from three distinct organisms representing viral, prokaryotic, and eukaryotic domains. We will explore the organization, composition, and functional significance of their genomes, transcriptomes, and proteomes, as well as employ advanced proteomic techniques to characterize protein expression profiles. Based on your understanding of omics, answer the following questions A) Describe how the study of genomes, transcriptomes, and proteomes can be helpful in exploration of diversity across different domains of life. B) Which public databases can be used for genomics and transcriptomics data acquisition? C) What are the advanced proteomics techniques to analyze protein expression profiles? D) List major features of viral, prokaryotic, and eukaryotic genomes.	4+4+4+3	CO3
Q 2	Case Study: The study involves the structural characterization of a target protein implicated in a specific disease pathway. We will employ computational modeling techniques to predict the protein's	4+4+4+3	CO4

	structure, assess its stability through energy minimization, and evaluate its conformational quality using Ramachandran plot analysis. Finally, we will explore how knowledge of the protein structure can inform the design of small molecule inhibitors as potential therapeutics.
	
	Based on your understanding of protein structure prediction, answer the following questions
	    A) Explain hierarchical organization of protein structures, including motifs, folds, and domains.
	    B) Which bioinformatics tool and algorithms can help us predict 3D structure of protein based on its amino acid sequence
	    C) Describe the process of energy minimization in protein structure refinement.
	    D) Discuss the role of protein structure in rational drug design.		

Section D **(2Qx10M=20 Marks)**			
Q 1	A) What is MALDI TOF spectroscopy? Explain its working principle. B) Discuss the applications of MALDI-TOF mass spectrometry in clinical diagnosis	**5+5**	**CO3**
Q 2	A) What is the Ramachandran plot, and how is it used to evaluate the quality of protein structures? B) Describe the interpretation of the Ramachandran plot in terms of allowed and disallowed regions for phi (ϕ) and psi (ψ) torsion angles. C) Discuss the implications of Ramachandran plot analysis for protein structure validation.	**4+3+3**	**CO4**