

JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY, WAKNAGHAT

TEST -3 EXAMINATION- 2023

M.Sc. (Biotechnology/Microbiology)

COURSE CODE(CREDITS): 20MS1BT213(2)

MAX. MARKS: 35

COURSE NAME: Bioinformatics

COURSE INSTRUCTORS: Dr. Shikha Mittal

MAX. TIME: 2 Hours

Note: (a) All questions are compulsory.

(b) Marks are indicated against each question in square brackets.

(c) The candidate is allowed to make suitable numeric assumptions wherever required for solving problems

Q1. Explain the concept of scoring matrices for aligning amino acid sequences. Briefly explain how PAM is derived? [3 marks] (CO-2)

Q2. Phylogenetic tree construction is based on some assumptions. What are the assumptions that are used for construction of phylogenetic tree? Define root node and clade with the help of a diagram. [4 marks] (CO-4)

Q3. Let, S1= AATTCGCGTA [4 marks] (CO-1 & CO-2)

S2 = TATCGCTACA

(a) Build the complete dynamic programming table for these strings

(b) List all optimal global alignment between S1 & S2.

Q4. If you have a gene sequence of an organism whose genome is yet to be annotated, how will you proceed forward via *in silico* analysis to trace a specific region with a known function? Which tool will be useful? Explain the concept and different categories of that tool in detail. [4 marks] (CO-3)

Q5. Define the UPGMA algorithm and state and justify its complexity. What is the output of the algorithm given the distance matrix of the species X1, X2, X3, X4 below? [5 marks] (CO-4)

$$\begin{pmatrix} \text{species} & X_1 & X_2 & X_3 \\ X_2 & 2 & & \\ X_3 & 4 & 4 & \\ X_4 & 6 & 6 & 6 \end{pmatrix}$$

1.

Q6. In a phylogenetic tree, the number of topologies varies with the number of taxa. If I have 5 taxa, then what will be the number of rooted and unrooted topologies? [4 marks] (CO-3 & CO-4)

Q7. What are the common structural motifs in RNA? What base-pairing are used in RNA? [3 marks] (CO-3)

Q8. Explain the following – [8 marks] (CO-1, CO-2, CO-3 & CO-4)

- a. Difference between cladogram and phylogram
- b. Affine gap penalty
- c. Sum of Pair method (SOP) in MSA
- d. What is FASTA format? Give an example of nucleotide sequence in FASTA format