Register Number:……..

KUMARAGURU
college of technology
character is life

B.TECH DEGREE EXAMINATIONS: NOV/DEC 2022

(Regulation 2018)

Seventh Semester

BIOTECHNOLOGY

U18BTI7203: Bioinformatics

COURSE OUTCOMES

CO1: To introduce the concept of biological sequence alignment and various genome sequence protocols.
CO2: To familiarize with various biological database searches, parameters and algorithm.
CO3: To apply, interpret and analyze multiple sequence alignments.
CO4: To construct, interpret and access molecular phylogenetic tree prediction .
CO5: To apply, interpret and analyze protein structures prediction algorithms
CO6: To introduce the concept of computer-aided drug designing (CADD).

Time: Three Hours **Maximum Marks: 100**

Answer all the Questions:-
PART A (10 x 1 = 10 Marks)

1. Matching type item with multiple choice code CO1 [K_1]

List I	List II
A. Primary database	i.PDB
B. Secondary database	ii. genbank
C. Composite database	iii.OWL
D. Protein structure database	iv.PRINTS

	A	B	C	D
a)	ii	i	iii	iv
b)	iii	iv	ii	i
c)	ii	iv	iii	i
d)	iii	i	ii	iv

2. Lower score in scoring matrices indicates the high probability that the two aminoacids aligned CO2 [K_2]

a) By chance b) Evolutionary unrelated

c) Both A and B d) Evolutionary related

3. Which of the following statements are true regarding BLAST/FASTA? CO2 [K_2]

1.FASTA uses the dynamic programming algorithm to compute original alignment.

2. FASTA miss cases of significant homology

3. BLAST focuses on no gap alignment of any length

4. BLAST focuses on high scoring alignment between query and database using heuristic approach

a) 1,3 b) 1,4

c) 1,2 d) 2,3

4. According to Jukes-cantor model assumption, the frequency of each base in DNA sequence is CO4 [K_1]

a) 0.35 b) 0.25

c) 0.75 d) 1

5. Assertion (A):UPGMA produces ultrametric tree CO4 [K_2]

Reason (R):Sequence alignment is converted into evolutionary distance in UPGMA method

a) Both A and R are Individually true and R is the correct explanation of A b) Both A and R are Individually true but R is not the correct explanation of A

c) A is true but R is false d) A is false but R is true

6. Which of the following is not true for sum of pair method? CO4 [K_1]

a) It takes all pair-wise information into account b) It is used in multiple sequence alignment

c) It will be biased if sequence from same family are over represented d) It is used in pairwise alignment

7. Following is the correct orders of steps in homology modeling CO5 [K_1]

1. loop modeling 2.target identification 3. sequence alignment 4.backbone modeling

a) 2-3-4-1 b) 1-3-2-4

c) 3-4-2-1 d) 4-1-3-2

8. Maximum parsimony method is more suitable for CO4 [K_2]

a) Similar sequences b) Smaller variations

c) Both a and b d) Larger variations

9. Assertion (A): In iterative methods in multiple sequence alignment, guide tree and MSA are rebuild iteratively CO3 [K_2]

Reason (R): Repeated Iterations increase the accuracy of multiple sequence alignment

a) Both A and R are Individually true and R is the correct explanation of A

b) Both A and R are Individually true but R is not the correct explanation of A

c) A is true but R is false

d) A is false but R is true

10. Which of the following algorithm is used in flexible docking — CO6 [K_1]

a) Smith watermen algorithm

b) DOCK algorithm

c) Conolly's algorithm

d) Needle man algorithm

PART B (10 x 2 = 20 Marks)
(Answer not more than 40 words)

11. Calculate Levenstein and Hamming's distance for the sequences — CO1 [K_2]
A: TTCGATCCATTG
B: ATCAATCGATCG

12. Mention the application of dot plot analysis — CO1 [K_1]

13. What is PSI BLAST? Give its application — CO2 [K_1]

14. Classify different types of gaps with example. — CO2 [K_2]

15. Mention the drawbacks of iterative methods for MSA — CO3 [K_2]

16. Tell whether PILE UP is a method based on progressive or iterative alignment? Illustrate the steps involved — CO3 [K_3]

17. Illustrate Newick format with an example — CO4 [K_1]

18. Mention the significance of bootstrapping in phylogenetic analysis. — CO4 [K2]

19. List different methods used for protein secondary structure prediction. Give its significance — CO5 [K_1]

20. Discuss how virtual screening is useful in Computer Aided Drug Design? — CO6 [K_2]

Answer any FIVE Questions:-
PART C (5 x 14 = 70 Marks)
(Answer not more than 350 words)

21. a) Explain the various components of the biological sequence formats –FASTA, FASTQ, Genbank with examples. — 7 — CO1 [K_2]

b) Describe the file format used in EMBL and PDB database. — 7 — CO1 [K_2]

22. a) Explain how the scoring matrices PAM and BLOSUM are constructed? Give its significance in alignment. — 7 — CO2 [K_2]

b) Align the given sequences using Needleman Wunsch algorithm — 7 — CO2 [K_3]
A: GCATGCG
B:GATTACA

match =1, mismatch = -1,gap = -1

23.	a)	Illustrate the steps involved in progressive alignment with CLUSTALW as an example.	7	CO3	[K_2]
	b)	What is PSSM? Explain how the values in position specific scoring matrix(PSSM) are calculated and state its application	7	CO3	[K_2]
24.	a)	Illustrate the steps involved in Maximum Likelihood method. Give its significance	7	CO4	[K_2]
	b)	Construct phylogenetic tree for the given distance matrix using UPGMA method.	7	CO4	[K_2]

	B	C	D
A	14	11	7
B		11	13
C	-	-	8

25.	a)	What is homology modeling? Explain the steps involved in it for tertiary structure prediction.	10	CO5	[K_2]
	b)	Describe the tools by which the protein tertiary structures can be compared.	4	CO5	[K_2]
26.	a)	Brief the applications of computer aided case study with a case study.	7	CO6	[K_2]
	b)	Elaborate the principle of rigid and flexible docking. Compare and contrast rigid and flexible docking.	7	CO6	[K_2]

****