


VIT

Vellore Institute of Technology

 Course: BBIT205L - Bioinformatics
 Class NBR(s): 1283 / 1286
 Time: Three Hours

Final Assessment Test - April 2025

Slot: A1

Max. Marks: 100

- KEEPING MOBILE PHONE/ANY ELECTRONIC GADGETS, EVEN IN 'OFF' POSITION IS TREATED AS EXAM MALPRACTICE
- DON'T WRITE ANYTHING ON THE QUESTION PAPER

 Answer ALL Questions
 (10 X 10 = 100 Marks)

- How are the sequences stored in computer? Mention the single and three letter codes of nucleic acids and amino acids.
- List out the different sequence file formats. Explain in detail about any four file formats.
- Describe the three database layers of UniProt.
- List out the difference between Global and Local alignment. [5]
 - Calculate the score $S(a,b)$ for the given sequence alignment by Needleman-Wunsch algorithm. [5]

GATATCTTGAAGC

GAAT - - - GAACCA

Use the similarity matrix below, with a gap penalty of -2.

	A	G	C	T
A	10	-1	-3	-4
G	-1	7	-5	-3
C	-3	-5	9	0
T	-4	-3	0	8

- What is dot plot? Construct a dot plot for the given set of sequences.
 S1 = GCAGGAAGTCTA
 S2 = GCAGGAAGTCTA
- Write any eight protein sequences of your choice; perform alignment and interpret at least 6 salient features from your alignment.
- If you are given a novel sequence and tasked to identify similar sequences – What will be your approach? Which algorithm will you use? Reason out.
- Interpret the relationship of phylogenetic analysis to sequence alignment. [4]

9.1) b) Construct a rooted phylogenetic tree for the given set of sequence by distance method. [6]

Seq 1: ASTYISTPQSTYIK

Seq 2: ASTRISTRQSTIIK

Seq 3: ATTRITTPQSYIYK

Seq 4: TSTYKTTRQSTIIK

9.a) List the different algorithms used for 2-D structure prediction and write about Chou-Fasman algorithm in detail.

OR

9.b) With the help of a neat flow chart describe the steps involved in protein modelling by homology-based methods.

10.a) Explain the methods employed for gene prediction. Add a note on gene prediction tools.

OR

10.b) Discuss *Insilico* drug designing.

⇔⇔⇔ U/D/TY ⇔⇔⇔