

2056

M.Sc. (Bio-Informatics) Second Semester
MBIN-8008: Sequence Bio-Informatics and Software Tools

Time allowed: 3 Hours

Max. Marks: 60

NOTE: Attempt five questions in all, including Question No. I which is compulsory and selecting atleast one question from each Unit.

x-x-x

I. Answer briefly:

- a) Consensus sequence
- b) Internal nodes of a tree
- c) Motifs
- d) Global alignment
- e) Pattern & Profiles
- f) tBLASTx

(6x2)

UNIT - I

- II. a) Discuss the construction and role of SWISSPROT as a protein sequence database. (2x6)
- b) Explain the Genbank flat file format.
- III. a) Describe the structure of a FASTA file and explain how sequence data is represented using it? (2x6)
- b) How is sequence retrieved using GQuery?

UNIT - II

- IV. a) Discuss the construction of PAM matrices. (2x6)
- b) Write a note on Needleman and Wunsch algorithm.
- V. a) What are the key features of Clustal W that make it useful for analyzing evolutionary relationships and conserved regions in DNA or protein sequences?
- b) Elaborate on the notion of Homology with reference to Orthologues and Paralogues. (2x6)

UNIT - III

- VI. a) Describe the use of PHYLIP for phylogenetic analysis. (2x6)
- b) Explain the importance of distance-based methods for phylogenetic analysis.
- VII. a) Differentiate between Phylogram and Cladogram.
- b) Elaborate on the maximum parsimony method for phylogenetic estimation. (2x6)

x-x-x