C 3094

B.E./B.Tech. DEGREE EXAMINATION, MAY/JUNE 2007.

Sixth Semester

Biotechnology

BT 1351 — BIOINFORMATICS

(Regulation 2004)

Time : Three hours

Maximum : 100 marks

Answer ALL questions.

PART A — (10 × 2 = 20 marks)

1. Mention the various data mining tools, their description and classification.

2. Describe two methods/tools of sequence submission to databases.

3. Give examples of macromolecular structure database and point out the information present in it.

4. Explain what is an operating system with couple of examples.

5. What type of alignment would result when the gap penalty is two times higher the mismatch score explain?

6. Mention two other type of substitution matrix than PAM and BLOSUM.

7. Define genetic distance and state two distance based phylogenetic tree prediction methods.

8. In a parsimony method of tree prediction what does informative site refer to?

9. What is gene synteny in genome analysis?

10. What type of information is obtained in a cluster analysis of microarray data? Explain with a suitable example.

PART B — (5 × 16 = 80 marks)

11. (a) Write the basic properties of the Unix operating systems such as Unix shell, file system. full and relative path names and give few frequently used Unix commands.

Or

(b) Explain the following Internet basics terms.

(i) Search engines.

(ii) Internet directories.

(iii) Meta search engines.

(iv) Boolean operators and

(v) Field searching

12. (a) How biological databases are classified, write the information in the most used databases for molecular biology such as.

Or

(b) Explain How the FASTA database search algorithm works and finds and reports the matching sequences.

13. (a) Elaborate on Dynamic programming method of pairwise sequence alignment and align the following sequence by Needleman and Wunch algorithm with an identity score + 1, mismatch score – 1 and gap penalty of –2. seq 1. TCGCA. seq2. TCCA

Or

(b) Give an example of a machine learning algorithm and explain a tool developed using it to do a multiple sequence alignment.

14. (a) Expand UPGMA. Explain how it performs tree building. Use the following data to draw a UPGMA tree

	Turtle	Man	Tuna	Chicken	Moth
Turtle	–	–	–	–	–
Man	19	–	–	–	–
Tuna	27	31	–	–	–
Chicken	8	18	26	–	–
Moth	33	36	41	31	–

Or

(b) Explain the maximum parsimony and maximum likelihood method of tree prediction in detail.

15. (a) Explain proteome analysis by 2D gel and elaborate on Image analysis of 2D gels.

Or

(b) What is a DNA microarray Technology? Gene expression data can be analyzed by hierarchical cluster method describe the process.