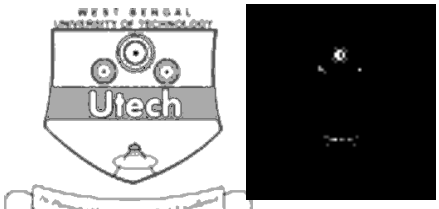
CS/M.Tech(BT)/SEM-2/MBT-201/09
BIOINFORMATICS (SEMESTER - 2)



1. .....
Signature of Invigilator

2. .....
Signature of the Officer-in-Charge

Reg. No.

Roll No. of the Candidate

CS/M.Tech(BT)/SEM-2/MBT-201/09

ENGINEERING & MANAGEMENT EXAMINATIONS, JUNE – 2009

BIOINFORMATICS (SEMESTER - 2)

Time : 3 Hours] [Full Marks : 70

INSTRUCTIONS TO THE CANDIDATES :

1. This Booklet is a Question-cum-Answer Booklet. The Booklet consists of **32 pages**. The questions of this concerned subject commence from Page No. 3.
2. a) In **Group – A**, Questions are of Objective type. You have to write the answer in the space provided **marked 'Answer Sheet'.**
   b) For **Groups – B** & **C** you have to answer the questions in the space provided marked 'Answer Sheet'. Questions of **Group – B** are Short answer type. Questions of **Group – C** are Long answer type. Write on both sides of the paper.
3. **Fill in your Roll No. in the box** provided as in your Admit Card before answering the questions.
4. Read the instructions given inside carefully before answering.
5. You should not forget to write the corresponding question numbers while answering.
6. Do not write your name or put any special mark in the booklet that may disclose your identity, which will render you liable to disqualification. Any candidate found copying will be subject to Disciplinary Action under the relevant rules.
7. **Use of Mobile Phone and Programmable Calculator is totally prohibited in the examination hall.**
8. You should return the booklet to the invigilator at the end of the examination and should not take any page of this booklet with you outside the examination hall, **which will lead to disqualification**.
9. Rough work, if necessary is to be done in this booklet only and cross it through.

No additional sheets are to be used and no loose paper will be provided

---

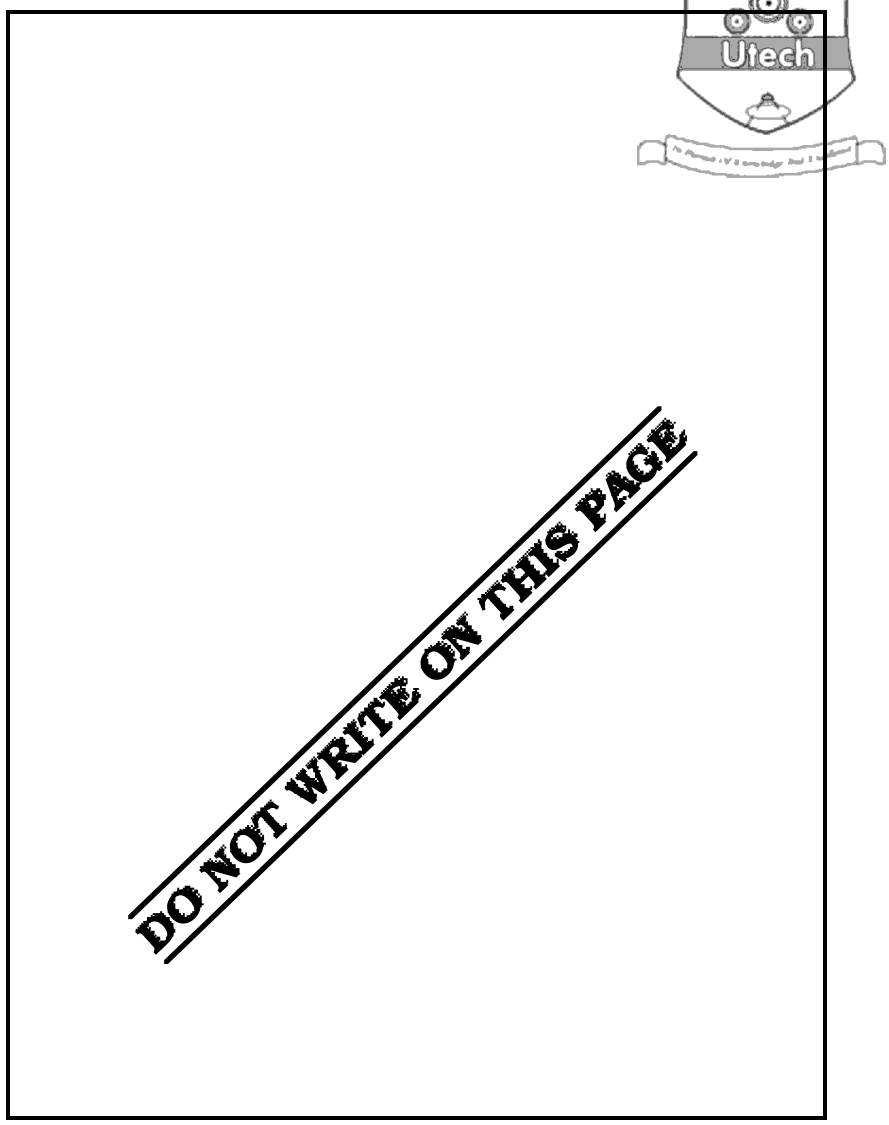
FOR OFFICE USE / EVALUATION ONLY

Marks Obtained

	Group – A										Group – B			Group – C			***Total Marks***	***Examiner's Signature***
Question Number																		
Marks Obtained																		

.....
Head-Examiner/Co-Ordinator/Scrutineer

42010 (30/06)


DO NOT WRITE ON THIS PAGE

CS/M.Tech(BT)/SEM-2/MBT-201/09

BIOINFORMATICS

SEMESTER - 2

Time : 3 Hours] [Full Marks : 70

GROUP – A

(Objective Type Questions)

1. Answer *all* questions : 10 × 1 = 10

A) Answer very briefly the following questions :

i) What does the number 62 signify in the substitution matrix BLOSUM62 ?

ii) What is the difference between homology and homoplasy ?

iii) What is the difference between a cladogram and a phylogram ?

iv) What is KEGG ?

v) What is a Specialist or Boutique database ?

B) Choose the correct alternative for the following :

vi) What is PROSITE ?

a) A database of protein structures

b) A database of interacting proteins

c) A database of protein motifs

d) A search tool.

vii) You have two distantly related proteins. Which BLOSUM or PAM matrix would you choose to compare them ?

a) BLOSUM45 or PAM250

b) BLOSUM60 or PAM1

c) BLOSUM80 or PAM120

d) BLOSUM80 or PAM1.

viii) The two main features of any phylogenetic tree are the

a) clades and the nodes

b) topology and the branch lengths

c) clades and the root

d) alignment and the bootstrap.

ix) Multiple sequence alignment is used for

a) searches for sequences for an unknown genome

b) prediction of structural similarities in unknown proteins based on known proteins in the alignment

c) phylogenetic analysis

d) all of these.

x) Which one is not a Data format ?

a) GenBank b) FASTA

c) dbGap d) None of these.

GROUP – B

(Short Answer Type Questions)

Answer any *three* of the following. 3 × 5 = 15

2. Present the aim of Blast software. Describe in words how the algorithm works.

3. Define multiple sequence alignment ? What is the goal of multiple sequence alignment ?

4. Define 'FASTA' format. Write the full form of NCBI. Classify databases on data source. Give one example of each database.

5. What is bioinformatics ? Explain the importance of bioinformatics ? What is scoring matrix ?

6. Compare the following :

a) Unigene and locuslink

b) Global and local alignment

c) Similarity and identity.

GROUP – C

(Long Answer Type Questions)

Answer any *three* of the following. 3 × 15 = 45

7. Describe with one example the difference between Hamming and Edit distances. Discuss the Smith-Waterman algorithm. If you had 1,500 base pair pieces of random DNA and you wanted to know how many of them had homology to known genes, what would you do to determine that ? What is *E* value and Bit Score ? What is paralogues ? 2 + 5 + 3 + 4 + 1

8. Discuss the complexity of the Neighbour-Joining algorithm. Discuss the properties and assumptions of the Jukes-Cantor one parameter model. Compare between Parsimony, Distance (NJ) and Likelihood-based algorithms. 2 + 10 + 3

9. What is gap penalty ? Compare the use of the affine gap penalty with the constant gap penalty. Do pairwise sequence alignment of the following two strings, *A* = ggaatgg and *B* = atg, where match = 0, mismatch = 20 and deletion = – 25. 2 + 3 + 10

10. Build the tree from the following distance matrix between species *A*, *B*, *C*, *D* using the UPGMA (Unweighted Pair Group Method using Arithmetic Averages) method.

Table

Describe the inputs in MP method. State the advantages and disadvantages of this method. 10 + 2 + 3

11. Write short notes on the following : 5 × 3

    a) Psi-Blast

    b) Dot-matrix

    c) Pfam

    d) SCOP

    e) Molecular clock.

END