	Utech
Name :	
Roll No.:	To the same of the same to the same to the same of the
Invigilator's Signature :	

CS/B.TECH/CSE/SEM-7/CS-704B/2012-13 2012 PIOUNEODMATICS

BIOINFORMATICS

Time Allotted: 3 Hours Full Marks: 70

The figures in the margin indicate full marks.

Candidates are required to give their answers in their own words as far as practicable.

GROUP - A

(Multiple Choice Type Questions)

- 1. Choose the correct alternatives for the following : $10 \times 1 = 10$
 - i) GOR method is used for
 - a) primary structure prediction
 - b) pairwise sequence alignment
 - c) primary structure prediction
 - d) all of these.
 - ii) TATA box is found in
 - a) Sequence 1 : A A A T A G A G A T A T A
 - b) Sequence 2: TAAAATAGAGATAT
 - c) Sequence 3: AATTAGAGATAAA
 - d) Sequence 4 : A T A T A G A G A T A A T

7402 [Turn over

CS/B.TECH/CSE/SEM-7/CS-704B/2012-13

iii) What is the score for following two sequences (assum match = +2, mismatch = -1, indel = -2)

ATCGGATCT

A - C G G - A C T

a) 7

b) 6

c) 10

- d) none of these.
- iv) Data archival means
 - a) Storage of data
- b) Analyzing of data
- c) Retrieving of data
- d) Removal of data.
- v) Which one of the following does not match with the rest three
 - a) PDB

b) NCBI

c) EMBL

- d) ORF.
- vi) Similar sequences in two different organisms that have been derived from a common ancestor sequence.
 - a) Homologs
- b) Orthologs
- c) Paralogs
- d) Xenologs.
- vii) UPGMA stands for
 - a) Unweighted pair group method with arithmetic mean
 - b) Unpair group method with arithmetic mean
 - c) Unweighted pair group method with mean arithmetic
 - d) None of these.
- viii) Which one is the appropriate tool for a description of a sequence line with the symbol?
 - a) FASTA

b) BLAST

c) MMD

d) None of these.

7402

CS/B.TECH/CSE/SEM-7/CS-704B/2012-13

- ix) Which one of the following is not example of secondary structure prediction algorithm
 - a) SOPM

- b) PREDATOR
- c) SOPMA
- d) All of these.
- x) Purines means
 - a) A and G
- b) A and T
- c) C and G
- d) none of these.

GROUP - B

(Short Answer Type Questions)

Answer any *three* of the following.

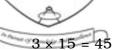
 $3 \times 5 = 15$

- 2. Suppose the Blast search returned 100 hits. Of these, 17 were false positives and we knew that there were 165 sequences in the database which should have returned a hit with our sequence. How many false negatives were there, and what is the sensitivity and selectivity of Blast in this instance?
- 3. Find a list of official and unofficial mirror sites of protein data bank. Which is closest to you? 4+1
- 4. Draw a dot plot of the following sequence from wheat dwarf virus genome: ttttcgtagagtgcgcggaggctttt against itself. In what respects is it not a perfect palindrome?
- 5. a) What is hamming distance between the words DECLENSION and RECREATION?
 - b) Explain Levenshtein distance in order to measure of sequence similarity. 3 + 2

7402 3 [Turn over

CS/B.TECH/CSE/SEM-7/CS-704B/2012-13

GROUP – C (Long Answer Type Questions) Answer any *three* of the following.



- 6. a) Why is Swiss Port important in bioinformatics?
 - b) Define Multiple sequence alignment? What is the goal of Multiple sequence alignment?
 - c) What are the differences between local and global sequence alignment. 5+7+3
- 7. a) Describe an algorithm for global alignment of strings.
 - b) Align the two strings LLNCDR and LRCDN globally using the Needleman and Wunsch algorithm. 5 + 10
- 8. a) Write down the differences between PAM matrix and BLOSUM Matrix with examples. Which one is better?
 - b) Write down steps of the FASTA. 8 + 7
- 9. a) Why HMM is required for symbol identification? How it is used.
 - b) Describe Vitervi algorithm in HMM. 10 + 5

7402 4