Name : ....

Roll No. : ....

Invigilator's Signature : ....

CS/B. Tech (CSE) /SEM-7/CS-704B/2011-12

2011
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 × 1 = 10

    i) Threading refers to which of the following ?

        a) Sequence alignment

        b) Fold recognition

        c) Gap in structure

        d) Sequencing method of DNA.

    ii) Which one of the following does not match with the rest *three* ?

        a) EMBL   b) DDBJ

        c) Gene bank   d) MIPS.

    iii) Orthologs are defined as

        a) homologous sequences in different species that share an ancestral gene

        b) homologous sequences that share little amino acid identity but share great structural similarity

        c) homologous sequences in the same species that arose through gene duplication

        d) homologous sequences in the same species which have similar and often redundant function.

7402   [Turn over

CS/B. Tech (CSE) /SEM-7/CS-704B/2011-12

iv) Data archival means

    a)   storage of data   b)   analyzing data

    c)   retrieving data   d)   removal of data.

v) Only four different bases are used in RNA molecules. These are

    a)   A, B, C and D   b)   A, T, G and C

    c)   A, U, G and C   d)   all of these.

vi) Total number of amino acid is

    a)   20   b)   23

    c)   21   d)   all of these

vii) Hamming or edit distance between the following sequences is

```
H   I   T
    |   |
I   I   T
```

    a)   1   b)   2

    c)   3   d)   4.

viii) Calculate the score, if the scoring scheme used was + 5 for a match, – 3 for a mismatch and – 4 for a gap.

```
G A A T T C
  | | | |
G G A – T C
```

    a)   10   b)   11

    c)   12   d)   13.

ix) UPGMA stands for

    a)   Unpair group method with Arithmetic mean

    b)   Unweighted pair group method with Arithmetic mean

    c)   Unweighted pair group method with mean Arithmetic

    d)   none of these.

CS/B. Tech (CSE) /SEM-7/CS-704B/2011-12

x) Which of the following is not an example of neighbour algorthms ?

a) SOPM   b) GORIV

c) SOPMA   d) PREDATOR.

GROUP – B

(Short Answer Type Questions)

Write short notes on any *three* of the following.

$3 \times 5 = 15$

2. Describe Central Dogma   5

3. a) Explain the differences between Bioinformatics and Computational Biology.

   b) Writ down the complementary and reverse DNA sequences for

     GGGATTTCGGAATCGTACGCCTTGGAATTG   3 + 2

4. Write down the distinction between Global Alignment and Local Alignment for two sequences.

5. a) What do you mean by Blosum 62 in scoring protein sequence alignment ?

   b) How disease predication is done in Bioinformatics ?

     2 + 3

6. What is PAM ? Explain with an example   2 + 3

GROUP – C

(Long Answer Type Questions)

Answer any *three* of the following. $3 \times 15 = 45$

7. a) Write down the methods of sequence alignment.

   b) Define EMBL.

   c) What are the applications of multiple alignment, sub-optimal alignment and parametric alignment.   5 + 5 + 5

7402   3   [Turn over

CS/B. Tech (CSE) /SEM-7/CS-704B/2011-12

8. a) Using a gap penalty of 2 (i.e., deduction 2 from the score for each gap), find the best alignment of AWAP and APP using a Needleman-Wunsch dynamic programming approach. Show your working (i.e., draw and fill in the S-W matrix).

   b) Find a simple dot plot for the words CORRELATIONS and RELATIONSHIP.   10 + 5

9. a) What is Microarray ?

   b) Write down some features of microarray.

   c) Explain Fold recognition in Bioinformatics.   5 + 5 + 5

10. a) What are the applications of Neural network in Bioinformatics ? Explain with an example.

    b) What are the modern usages of phylogenies ? (5 + 5) + 5

11. a) What are the parts of FASTA ?

    b) Explain Chao-Fasman algorithm with an example.

                                                          5 + 10

12. Write short notes on any *two* of the following : $2 \times 7\frac{1}{2}$

    a) Metabolic pathway

    b) HMM

    c) Nearest Neighbour method for secondary structure prediction.