

2012

**COMPUTATIONAL BIOLOGY &  
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 any *ten* of the following :

10 × 1 = 10

- i) Each Gen Bank entry includes
  - a) a concise description of the sequence
  - b) the scientific name
  - c) taxonomy of the source organism
  - d) all of these.
- ii) Methods to perform DNA or RNA sequencing include
  - a) chromosome walking
  - b) shotgun sequencing
  - c) chain termination method
  - d) all of these.
- iii) Brookhaven ..... is a database of 3D structures.
  - a) DDBJ b) PDB
  - c) Swiss Prot d) UniGene.
- iv) Phylogenetic tree depends on
  - a) MSA b) BLAST

- c) Global alignment d) none of these.
- v) Which of the following is not true about Swiss Prot ?
- a) Highly curated
  - b) Highly cross-referenced
  - c) Non-redundant database
  - d) redundancy.
- vi) The low complexity region among the following is
- a) 'aaggtcctagtagtcga' b) 'aaaaaaaa'
  - c) 'attattaaagcgctgcat' d) 'nnnnnna'.
- vii) Dynamic programming is used to
- a) develop the phylogenetic tree
  - b) detect the most suitable alignment
  - c) detect the conserved domain of different protein
  - d) none of these.
- viii) Pairwise Sequence alignment is associated with
- a) the Baum-Welch Algorithm
  - b) the Viterbi Algorithm
  - c) Needleman-Wunuch Algorithm
  - d) None of these.
- ix) Blastx is a tool of
- a) search protein database using a translated nucleotide query
  - b) search nucleotide database using a translated nucleotide query
  - c) search protein database using a protein query
  - d) search translated nucleotide database using a protein query.
- x) The E-value of an alignment

a) is the expected number of sequences that give the same Z-score or better if the database is probed with a random sequence

b) is expertise value of sequences that give the same P-value

c) is probability value of an alignment having more than these similarities

d) none of these.

xi) Fasta algorithm uses

a) *k*-tuple values b) LSP values

c) complex values d) none of these.

xii) Gap penalty in sequence alignment is

a) the penalty score for umatch between later

b) the penalty score for umatch between two gaps

c) The penalty score for umatch between same residues

d) the penalty score for giving a gap to get maximum match interpreting as insertion, deletion and mutation.

## **GROUP – B**

### **( Short Answer Type Questions )**

Answer any *three* of the following.  $3 \times 5 = 15$

2. What is Ramachandran plot ? Explain the utility of Ramachandran plot to predict secondary structure.

3. How would you get the conserved domain and get the phylogenetic relation of the prion protein of horse, human, macaca mulatta and chimpanzee. What is Pubmed ? What is the name of the database where you get the inheritable

disease of human ?

4. Explain the term 'GIR'. What is central dogma in biology ?

Give example of experimental and theoretical approaches for proteomics.

5. Write short note on any *one* of the following :

a) MSA

b) The experimental approaches to genome sequence data.

6. What is substitution matrix ? In PAM-250 the 250 stands for what ? Write down the difference of PAM250 and BLOSUM62 matrix.

### **GROUP – C**

#### **( Long Answer Type Questions )**

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

7. What are *E*-value and *Z*-value of an alignment ? Write down the steps. Find a nucleotide from NCBI namely

Topoisomerase. Convert it to protein. Find the homologue of

the protein from the databases. Then find the similarity of

the corresponding protein with other species. What is pairwise alignment and *k*-tuple ? Write the use of HSP in Blast.

$2 + 2 + 6 + 2 + 3$

8. Explain the main features of CSD. Write the main features of

SCOP and CATH. Explain the information obtained from

PROSEARCH method.  $5 + 3 + 3 + 4$

9. Align the following two sequences by dynamic programming

method : 15

G A A T T C A G T T A ( sequence # 1 )

G G A T C G A ( sequence # 2 ).

10. Write down the name of the tools of NCBI where local

alignment is done. What is the algorithm behind the program ? What are the applications of doing an MSA ? You have been given a protein along with the sequence of it. How would you identify the protein ? If not possible, then how would you get an idea of the functionality of the protein ? Give proper explanation. Define domain of a protein. Write a short note on Blast's working procedure.

1 + 1 + 2 + 3 + 1 + 2 + 2 + 1 + 2

11. State the differences between homology and similarity.

Write down six emerging fields in bioinformatics. You have given a sequence of African elephant and Indian elephant.

How would you draw the phylogenetic relation between them ? What result would you expect ? What is relative mutability ? 3 + 3 + 4 + 2 + 3

12. What is E-value and Z-value of an alignment ? Write down the steps only. How will you show through bioinformatics that the insulin sequence of HORSE, MONKEY, HUMAN and RAT are very near. What is Pairwise alignment and  $k$ -tuple ? Write the use of HSSP in Blast. 2 + 2 + 6 + 2 + 3

=====