

M.Sc. (Biotechnology) Part – II Exam, 2017

Computer and Bioinformatics

Paper – 2/3

Group –A

Answer Q1 and any three from rest

1. Answer any **five** Questions: 5X4
- a) What is compiler? How it differs from interpreter?
 - b) Assume 8-bits are used to store an integer. Perform the following using 2's complement representation.
 - i) $(-42)_{10} - (25)_{10}$
 - ii) $(-42)_{10} + (25)_{10}$
 - c) Represent $(-0.02)_{10}$ using IEEE single precession representation.
 - d) Explain for and while loop in C language.
 - e) Differentiate between Call by value and Call by reference.
 - f) Perform the following BCD arithmetic:
 - i) $73 + 37$
 - ii) $73 - 37$
2. a) Draw a flowchart to find GCD of two positive integers.
b) Write an algorithm to find minimum value among N numbers given by user. 5+5
3. a) Explain the operation performed by the following string functions:
i) strlen() ii) strcat() iii) strcmp()
- b) Write a C program that evaluates the series: $1 + x + x^2 + x^3 + \dots + x^N$ 5+5
4. a) Write a C program that accept a string as input and finds number of words in the string.
b) Differentiate between gets() and puts() function in C language. 7+3
5. a) Write a C program that takes an integer and check that it is prime or not.
b) A word that reads same backwards as forwards is known as palindrome, e.g. madam. Write a C program that takes a word as input and check whether it is palindrome or not. 5+5

[Turn over

GROUP - B

Answer all the questions.

The figures in the margins indicate full marks.

6. Answer the following questions: $1 \times 10 = 10$

- What is primary and secondary database?
- What makes BLAST more useful than FASTA for database search?
- What are the basic structure of a GenBank and an EMBL entry?
- Write the name of scoring model used to build similarity between nucleotide sequences?
- For what purpose BLOSUM matrices are used?
- What is FASTA format?
- Which one is used for search all NCBI associated databases?
 - BLAST
 - CLUSAL W
 - ENTREZ
 - PDB
- If you want to BLAST the non-redundant database using a new protein sequence as query, which is the BEST search program to use?
 - blastp
 - blastn
 - tblastx
 - blastx
- Submission to GenBank include
 - Sequin and BankIt
 - Sequin
 - BankIt
 - BankIt and sequen
- Name two different methods used in pair wise sequence alignment.

7. Answer **any five** of the following: $2 \times 5 = 10$

- Write the two most commonly used scoring matrix.
- Why the alignment between two sequences is performed?
- Write name of two web based tools used for restriction mapping.
- What do you mean by Express Sequence Tags (EST) database?
- What is bioinformatics?
- What do you mean by local and global alignment?
- What do you mean by identity and similarity between two protein sequences?

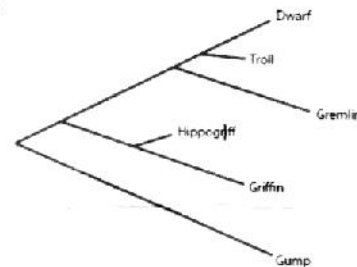
Answer **any four** of the following: $5 \times 4 = 20$

- What is DNA sequencing chromatogram? Explain how to interpret it? $3+2$
- What is Homologs, Paralogs, Orthologs. Explain them with examples. $3+2$
- What do you mean by similarity matrix? Explain with example.

- Define the Total Score, Max Score, Query coverage and E-value, Identity in the context of sequence alignment 5
- What's the three major databases with their respective operative countries. Name the two structure databases. $3+2$
- What are differences between PAM and BLOSUM matrix? 5

Answer **any one** of the following: $10 \times 1 = 10$

- What are Databanks in the context of Bioinformatics? How do you classify various types of databanks pertaining to bioinformatics $5+5=10$
- Answer the following questions about this tree:



- Which are the two most closely related species?
 - Which of these is probably the outgroup?
 - Which is most distantly related to the Gremlin other than the outgroup species?
 - Write down the few basic character of outgroup.
- c) What is out-group sequence in phylogenetic tree and why it is so important in phylogenetic analysis? Define the phylogenetic tree construction methods that directly based on sequences. $5+2+3$