D 9080 (Pages : 2) Name..................................

Reg. No..............................

THIRD SEMESTER M.Sc. DEGREE EXAMINATION, NOVEMBER 2010

General Biotechnology

GBT 211—BIOSTATISTICS AND BIOINFORMATICS

Time : Three Hours Maximum : 80 Marks

Section A

Answer any **two** *questions.*
Each question carries 10 *marks.*

1. Following data is GC content (%) of fifteen nucleotide sequences of two organisms A and B are tested comparatively :

   A(x) : 3.31 2.41 2.72 2.41 2.11 2.11 3.01 2.13 2.41 2.10 2.41 2.09 3.00 2.08 2.11
   B(y) : 4.09 3.84 3.65 3.20 2.97 3.22 3.96 2.76 3.42 3.38 3.28 2.93 3.54 3.14 2.76

   Find the regression equation describing the relationship between the two variabels and infer the result.

2. Explain various types of data model in database systems.

3. Describe in detail about the BLAST algorithm in pairwise sequence alignment.

$(2 \times 10 = 20$ marks)

Section B

Answer any **ten** *questions.*
Each question carries 5 *marks.*

4. How do you represent the collected data in a table ? How does the tabulation is important in statitsical analysis ?

5 Use the following data to prepare a histogram and write its significance in biological data analysis :-

Length of the protein sequence (A) :	10-30	31-50	51-70	71-90	91-110
No. of proteins (B) :	60	85	92	70	95

6. Explain in detail the Internet and its applications in Bioinformatics study.

7. The following are the genome size (Mb) of 10 organisms. Compute standard deviation and infer the result :

   4.07, 2.71, 3.64, 3.37, 3.84, 3.83, 3.82, 4.21, 4.04, 4.50.

8. What is Harward graphics ? Explain its applications.

9. What do you mean by data structure ? Describe any one data structure with suitable example.

10. Explain how t test is applied to find the significance of a given hypothesis.

11. Explain the syntax of IF statement in QBASIC.

12. Write a C program to find median for any given n values.

Turn over

13. Explain, conversion procedure of a binary number to a decimal number with an example.

14. Describe various hardware devices in a computer.

15. Explain, bubble sort and its significance.

(10 x 5 = 50 marks)

Section C

Answer **all** *questions.*
Each question carries 2 marks.

16. What is the purpose of the coefficient of variation ?

17. Prepare a scatter diagram for the following data and discuss the result :-

X: 0.4 6.9 -0.1 12.4 -2.8 7.5 20.3 2.5 12.4 10.1 -2.7 -3.8
Y: 2.1 3.3 4.4   4.9   2.1   1.0 12.6 0.8 9.7 9.1   0.5 -3.6

18. Why C language is called as middle level language ?

19. Mention few applications of Haward graphics.

20. What are the scoring matrices used in pairwise alignment ?

(5 x 2 = 10 marks)