7. Write short notes on **(any three)** : (12)

   (i) Significance of hypothesis testing

   (ii) Scoring matrices

   (iii) Drug discovery methods

   (iv) Machine Learning in Bioinformatics

---

6 Dec

[This question paper contains 6 printed pages.]

Your Roll No..............

Sr. No. of Question Paper : 1162

06 DEC [illegible]

Unique Paper Code : 32237905

Name of the Paper : DSE Computational Biology

Name of the Course : **B.Sc. (H) Zoology Examination, 2022-LOCF**

Semester : V

Duration : 3 Hours Maximum Marks : 75

Instructions for Candidates

1. Write your Roll No. on the top immediately on receipt of this question paper.
2. Use of simple calculator is allowed.
3. Statistical tables should be provided.
4. Attempt **five** questions in all, including Question No **1** which is compulsory.

1. (a) Expand the following abbreviations : (3)

   (i) DDBJ

   (ii) GWAS

(iii) BLAST

(iv) BLOSUM

(v) GCG

(vi) ADMET

(b) Match the terms in column A with those given in column B : (3)

A	B
Dynamic Programming	Chi square test
Composite Database	PDB
Structural Bioinformatics	MSA
Heuristic Method	Needlemann-Wunsch (NW)
Progressive Alignment	OMIM
Karl Pearson	FASTA

(c) Define the following : (5)

(i) Biological Database

(ii) Gap Penalty

(iii) Dot matrix

(iv) Functional Genomics

(v) Null Hypothesis

4. (a) Classify and explain major nucleotide databases in bioinformatics with examples.

(b) Explain similarities and differences between BLAST and FASTA tools for sequence alignment. (4+8)

5. (a) Discuss different methods of protein sequencing.

(b) What are the differences between chain termination and chemical cleavage method of gene sequencing? (8+4)

6. (a) What are confidence intervals? Explain their significance using a suitable example.

(b) A researcher was analyzing the results obtained from a heterozygous dihybrid cross and following data was obtained: dominant for both traits: 570, dominant for trait 1 and recessive for trait 2: 185, dominants for trait 2 and recessive for trait 1: 190, and recessive for both traits: 55. Perform a chi-square analysis to see if the data above agrees with the predicted outcome of this cross ($\alpha = 0.05$). (4+8)

(g) State whether the following statements are *true or false*. Give reasons for your answer. (3)

(i) BLAST and FASTA are faster than Dynamic programming.

(ii) Gap extension penalty is always smaller than gap opening penalty.

(iii) Variance is the square root of standard deviation.

2. (a) Define Bioinformatics. Discuss its aim and scope in biological sciences.

(b) Describe major applications as well as limitations of Bioinformatics. (6+6)

3. (a) What is the advantage of Needleman-Wunsch alignment over seeded alignment?

(b) What are indels and indel penalty? Explain different types of gap penalties used in scoring alignments?

(c) What would happen if gap penalties were too high or too low? Explain your answer in each case.

(4+6+2)

(d) Differentiate between the following : (8)

(i) Curated and Non-curated databases

(ii) Local and Global Alignment

(iii) Type I and Type II error

(iv) Dependent and Independent t-test

(e) Explain the following symbols : (3)

(i) p

(ii) C.V.

(iii) Σ

(iv) μ

(v) α

(vi) σ^2

(f) Write the contribution of following scientists : (2)

(i) Stephen Altschul

(ii) Christian D. Wunsch

(iii) Steven Henikoff

(iv) William Sealy Gosset