

Total number of printed pages-4

**3 (Sem-5/CBCS) ZOO HE 1**

**2022**

**ZOOLOGY**

(Honours Elective)

Paper : ZOO-HE-5016

( DSE(H)-1)

**(Computational Biology and Biostatistics)**

Full Marks : 60

Time : Three hours

**The figures in the margin indicate full marks for the questions.**

1. Fill in the blanks : **(any seven)**  $1 \times 7 = 7$
- (a) SWISPROT protein sequence database began in the year \_\_\_\_\_ .
- (b) Bioinformatics database was first created by \_\_\_\_\_ .
- (c) The human genome contains approximately \_\_\_\_\_ base pairs.

*Contd.*

- (d) COPIA is used for identification of \_\_\_\_\_ .
- (e) STAG is maintained by \_\_\_\_\_ .
- (f) The information retrieval tool of NCBI Gene Bank is \_\_\_\_\_ .
- (g) Proteomics refers to the study of \_\_\_\_\_ .
- (h) EMBL is a \_\_\_\_\_ sequence database.
- (i) Phylogenetic relation can be shown by \_\_\_\_\_ .
- (j) The first molecular biology server expasy was used in the year \_\_\_\_\_ .
- (k) \_\_\_\_\_ is the first completed and published gene sequence.
- (l) The identification of drugs through the genome study is called \_\_\_\_\_ .

2. Answer the following : **(any four)**

2×4=8

- (a) EMBL
- (b) Global alignment
- (c) Standard error

(d) Phylogram

(e) DNA database of Japan

(f) Goodness-of-fit

(g) OMIM

(h) Genetic Code

3. Answer the following questions : **(any three)**

5×3=15

- (a) Explain the hierarchical method of multiple sequence alignment.
- (b) Discuss the database that contain the information of protein sequences.
- (c) Explain the secondary database.
- (d) What is dendrogram ?
- (e) What are biological databases based on data type and database design ?
- (f) Explain co-efficient of variance and its importance.
- (g) How standard deviation is differed from standard error ?
- (h) Application of Chi-square tests.

4. Answer the following questions : **(any three)**  
10×3=30

- (a) What is phylogeny ? Discuss the various method for phylogenetic analysis.
- (b) Elaborate different file formats in biological resources with the help of an arbitrary example each.
- (c) Give various types of BLAST. Explain briefly the algorithm of BLAST.
- (d) Discuss STRING database. Describe its various features.
- (e) How can you predict the structure of a protein sequence ? Explain in detail one of the knowledge-based method to predict the protein structure.
- (f) Explain the various apparatus used in pairwise sequence alignment.
- (g) What is Biostatistics ? How biostatistics help in biological sequence analysis ?
- (h) Calculate the standard deviation from the following data :

X :	20-30	30-40	40-50	50-60	60-70	70-80	80-90	90-100
Y :	30	58	62	85	112	70	57	26