



Name :

Roll No. :

Invigilator's Signature :

CS/M.Sc.(GE)/SEM-3/MSGEN-302/2012-13

2012

**GENOMICS PROTEOMICS AND
BIOINFORMATICS**

Time Allotted : 3 Hours

Full Marks : 70

The figures in the margin indicate full marks.

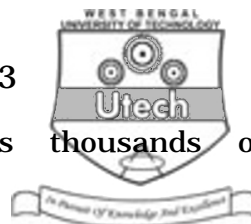
*Candidates are required to give their answers in their own words
as far as practicable.*

GROUP – A

(Multiple Choice Type Questions)

1. Choose the correct alternatives for any *ten* of the following :
10 × 1 = 10

- i) Experimental tool(s) available in proteomics is/are
 - a) UV-VIS
 - b) X-ray
 - c) SPR
 - d) all of these.
- ii) KEGG database is
 - a) Kyato Enrich of Genes and Genomes
 - b) Kyato Environment of Genes and Genomes
 - c) Kyato Energy of Genes and Genomes
 - d) Kyato Encyclopedia of Genes and Genomes.



- iii) Microarray technique encompasses thousands of interactions of the type
- a) protein-ligand
 - b) antibody-antigen
 - c) receptor-ligand
 - d) all of these.
- iv) In sex linkage the speciality is
- a) atavism
 - b) reversion
 - c) gene flow
 - d) criss-cross inheritance.
- v) EMP is database for
- a) Environment and metabolic parameter
 - b) Enzymes and metabolic pathway
 - c) Encyclopdia of metabolic pathway
 - d) Enzymes and metabolic parameter.
- vi) A widely used machine learning approach is
- a) Hidden Markov Models
 - b) Mechanical Pathways
 - c) Metal dependent Pathways
 - d) None of these.
- vii) Protein-protein interaction database is
- a) DIP
 - b) PPI
 - c) DIPTM
 - d) all of these.
- viii) MSDN stands for
- a) Microbial Strain Data Network
 - b) Microbes Sequence Diversity Nomenclature
 - c) Micro Soft Data Network
 - d) None of these.



- ix) Protein Secondary Structure contains
- Peptide bond
 - Hydrogen bond
 - Disulphide bridge
 - None of these.
- x) Genebank file ends with
- ENDML
 - > ||
 - //
 - \\.
- xi) Full form of CSD is
- Chemical Structure Database
 - Cambridge Structural Database
 - Chemical Structure Determination
 - None of these.
- xii) The Scoring matrices never reflect
- # of mutations to convert one to another
 - chemical similarity
 - observed mutation frequencies
 - GAP scores for each mismatch.
- xiii) NCBI contains
- Primary databases
 - Derivative databases
 - Both primary and derivative databases
 - none of these.



- xiv) The full form of EST is
- a) Express Sequence Test
 - b) Expression Substitution Tags
 - c) Extra Substitution Term
 - d) Express Sequence Tags.
- xv) Clustal W is tool for doing
- a) Multiple sequence alignment
 - b) Phylogenetic analysis
 - c) Multiple sequence alignment & phylogenetic analysis
 - d) Pair-wise alignment.
- xvi) Needleman and Wunsch algorithm is used for
- a) Local alignment
 - b) Global alignment
 - c) MSA
 - d) None of these.
- xvii) Example of metabolic pathway database is
- a) RESS
 - b) KEGG
 - c) OMIM
 - d) BLISS.
- xviii) The full form of BLOSUM is
- a) Block Summation Mode
 - b) Block Substitution Matrix
 - c) Block Sequence Monitoring
 - d) None of these.



GROUP - B

(Short Answer Type Questions)

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

2. a) Name two structural databases.
b) What is the difference between Accession and GI number of NCBI ? $1 + 4$
3. FISH to identify chromosome landmark. Explain.
4. Write application of proteomics for identification of disease genes.
5. a) What is the basis of database searching in bioinformatics ?
b) Explain the main features of CSD. $3 + 2$
6. Explain Genomic Variation.
7. Discuss in brief about protein array assay.
8. Write short notes on any *one* of the following :
 - a) Chou-Fasman algorithm
 - b) Fasta
 - c) INSDC.
9. Discuss the use of RAPD in DNA typing.
10. Name three tools available in Expasy proteomics server and write their application.
11. What is substitution matrix ? In "PAM 250" the 250 stands for what ? Write down the difference of PAM 250 and BIOSM 62 matrix. $1 + 1 + 3$



GROUP - C

(Long Answer Type Questions)

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

12. Align the following two sequences by dynamic programming method. Write down the score.

GAATTCAGTTA (sequence#1)

GGATCGA (sequence#2) $12 + 3$

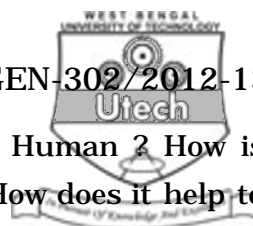
13. Write notes on any the *three* databases : 3×5

- a) Blast algorithm
- b) RFLP method of DNA typing
- c) PPI
- d) OMIM.

14. Draw a diagram showing the steps of determining of secondary structure of protein in a bioinformatics approach. Write any four publically accessible nucleotide databases. Write one family and one motif determining tool of protein in bioinformatics. What is file format ? Why is it necessary ?

$8 + 2 + 2 + 1 + 2$

15. a) What is Secondary database ?
- b) Name two Secondary databases.
- c) What information can be obtained from secondary database ?
- d) Compare the features between primary, secondary & tertiary databases.
- e) Write a short note on International Nucleotide Sequence Database Collaboration.
- f) What do you mean by SRS ? $2 + 1 + 2 + 4 + 4 + 2$



16. Explain Structural variation and types in Human ? How is Structural Variation mapped in Human ? How does it help to identify human diseases ? 5 + 5 + 5
17. Write selection way of KEGG database. What are the characteristics of KEGG ? Discuss the information available from this database. Write in short what type of information can be obtained from EMP database. Name two metabolic pathway databases other than EMP. 2 + 3 + 5 + 3 + 2
18. What is functional Genomics ? What is comparative Genomics ? How does it relate to functional Genomics ? How do Model Organism Data help in functional Genomics ? 2 + 3 + 5 + 5
19. What are NPs ? How SNPs effect in the following situations :
 a) SNPs in genes
 b) SNPs in Regulatory region
 c) SNPs in Non-Regulatory Intergenic regions. 3 + 4 + 4 + 4
20. Write the basic principle of protein array technique. Explain briefly the attachment methods available for protein array. Write the general steps for antibody array. Give its few applications. Give example of three databases and three tools of ExPASy proteomics server. 2 + 3 + 4 + 2 + 2 + 2
21. Why is the study of protein-protein interactions important in proteomics ? What are the different forces behind protein-protein interactions ? What are the differences between DIP and PPI servers ? How is Prosite search done ? Explain a detection technique in protein array. 4 + 2 + 3 + 4 + 2

=====