GENOMICS AND PROTEOMICS
(BIOT 5232)

Time Allotted : 2½ hrs **Full Marks : 60**

Figures out of the right margin indicate full marks.

Candidates are required to answer Group A and any 4 (four) from Group B to E, taking one from each group.

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

Group – A

1. Answer any twelve: **12 × 1 = 12**

Choose the correct alternative for the following

(i) In mass spectrometric method, which ionization method will be the best to study peptides, proteins and DNA up to 500 kD?
(a) ESI (b) MALDI (c) FAB (d) Electron impact ionization.

(ii) What is sensitivity of staining per spot by Coomassie Brilliant Blue R-250 in protein visualization?
(a) 1 mg (b) 1 ng and up (c) 100 ng to 10 μg (d) 1 ng to 1 μg

(iii) To determine K_d of protein-protein interaction, you will use which of the following assay techniques?
(a) GST-Pull down assay (b) Phage display
(c) SPR (d) Y-2H assay

(iv) You need to use a first-generation sequencing method for de novo sequencing, which template should give optimum results for this project?
(a) Genomic DNA (b) PCR product
(c) BAC (d) Plasmid DNA

(v) Name the sequences which are present in more than one copy in a haploid genome?
(a) Nonrepetitive DNA (b) Highly repetitive DNA
(c) Repetitive DNA (d) Minisatellite.

(vi) Which of the following is untrue regarding EST Index Construction?
(a) The goal of the EST databases is to improve the quality of the sequence information so the data can be used to extract full-length cDNAs
(b) The goal of the EST databases is to improve the quality of the sequence information so the data can be used to extract full-length cDNAs
(c) The process includes a preprocessing step that removes masks repeats
(d) The goal of the EST databases is to organize and consolidate the largely redundant EST data.

(vii) Which of the following is untrue about Whole Genome Alignment?
(a) This helps to reveal the presence of conserved functional elements
(b) It doesn't help to understand sequence conservation between genomes
(c) It be accomplished through direct genome comparison or genome alignment
(d) The alignment at the genome level is fundamentally no different from the basic sequence alignment.

(viii) A protein undergoes post-translational modifications. In a following experiment to identify the nature of modifications, following experimental results were found:
(1) Proteins moved more slowly in an SDS-PAGE. (2) Isoelectric focusing (IEF) showed that there was no change in the pI. (3) Mass spectrometric analysis showed that the modification was on serine. The modification that the protein undergoes is likely to be?
(a) phosphorylation (b) glycosylation
(c) ADP-ribosylation (d) All of these.

(ix) Imagine you are designing an experiment to measure gene expression in the blood of patients who have been treated with a new drug. You want to measure expression of all genes in the genome and may want to use the data to identify novel transcripts (e.g. miRNAs) in the future. Which method would you choose?
(a) ChiP-seq (b) Real-time PCR
(c) RNA-seq (d) Microarray analysis.

(x) For performing 2-Dimensional gel electrophoresis several steps are involved. Which of the following is the correct order of the steps involved?
(a) Sample solubilisation > Equilibration > Isoelectric focusing > SDS-PAGE> Staining and Image analysis > Spot picking
(b) Sample solubilisation > Isoelectric focusing > Equilibration > SDS-PAGE> Staining and Image analysis > Spot picking
(c) Equilibration > Sample solubilisation > SDS-PAGE> Isoelectric focusing > Staining and Image analysis > Spot picking
(d) Sample solubilisation > Isoelectric focusing > Spot picking> Equilibration > SDS- PAGE> Staining and Image analysis.

Fill in the blanks with the correct word

(xi) In 2-D gel electrophoresis the 2nd dimension is _____.

(xii) In mass spectrometry quadruple is used for _____ analysis.

(xiii) The _____ chromosome contains the fewest number of genes.

(xiv) _____ is an example of a repeated sequence with an open reading frame for reverse transcriptase.

(xv) The pI of a molecule is the _____ at which the net charge of the molecule is zero and the molecule don't migrate in an electrical field.

Group - B

2. (a) Human Genome Project has revealed that there are about 1.4 million locations where single-base DNA differences (SNPs) occur in humans. Justify with scientific logic how his information promises to revolutionize the processes of finding chromosomal locations for disease-associated sequences and tracing human history. *[(CO2,4)(Justify and Criticize/HOCQ)]*

   (b) What are Haplotypes? Give an overview of the 'HapMap Project' with its goals and findings. *[(CO2,4)(Understand/LOCQ)]*

6 + (2 + 4) = 12

3. (a) What are clone contigs? Give a brief description of the clone-by-clone hierarchical shotgun sequencing technique. *[(CO1)(Remember/LOCQ)]*

   (b) The 2022 Nobel Prize in Physiology & Medicine was awarded to Svante Pääbo for his contribution in the field of Paleogenomics. In this connection, what do you think are the difficulties in extracting genomic data from pre-historic samples? Give a critical appreciation of the implications of Svante Pääbo's discoveries in the area of genome research. *[(CO1)(Analyze & Criticize/HOCQ)]*

(1 + 4) + (3 + 4) = 12

Group - C

4. (a) For describing gene function using natural language creates imprecise expression analysis scientists overcome this problem by development of protein functional descriptions- explain how this is done with a key word Cytochrome C Oxidase. *[(CO2)(Evaluate/IOCQ)]*

   (b) The genetic complexity is manifested at the protein expression level where it reveals more expressed proteins than genes available to code for them- explain the phenomenon in respect to eukaryotic organisms. *[(CO2)(Explain/IOCQ)]*

6 + 6 = 12

5. (a) "Microarray based method is followed to study the global gene expression profiling". Outline the steps followed for this microarray based method *[(CO2,4) (Analyse/IOCQ)]*

   (b) To design an optimal oligonucleotide probe certain criteria are needed to be followed. Discuss those criteria citing the reasons behind it. *[(CO2,4)(Describe/LOCQ)]*

6 + 6 = 12

Group - D

6. (a) Write the names of three techniques for determination of molecular mass of a protein. Explain the principle of determination of mass of a pure protein by the technique you mentioned which does not use either electrophoresis or chromatography, with diagram. *[(CO3)(Analyse/IOCQ)]*

(b) Write names of three techniques to study protein-protein interaction (PPI). Explain anyone technique you mentioned with diagram for the study of PPI? *[(CO3)(Remember/IOCQ)]*

(1 + 5) + (1 + 5) = 12

7. (a) A protein was isolated from human tissue and subjected to a variety of purification steps. After purification, protein sample was subjected to electrospray ionization mass spectrometry and the following data obtained.

m/z	773.9	825.5	884.3	952.3	1031.3
Abundance (%)	59	88	100	66	37

Assuming that the only ions in the mixture arise by protonation, deduce an average molecular mass for the protein by this method. *[(CO6)(Analysis/IOCQ)]*

(b) Explain the basic principles of determination of mass of protein by MALDI-TOF MS with labelled diagram? *[(CO3)(Understand/LOCQ)]*

(c) Write three advantages of MALDI-TOF over ESI-TOF. *[(CO3)(Remember/IOCQ)]*

5 + 3 + 4 = 12

Group - E

8. (a) Cite two technical reasons that make structure determination of proteins by NMR difficult. Cite three steps that have been undertaken to improve this technique to make it more amenable for structure determination of large molecular weight proteins. *[(CO5)(Apply/HOCQ)]*

(b) Explain how the use of a metal substituted amino derivative can improve the xray diffraction pattern in protein crystallography. *[(CO5)(Understand/LOCQ)]*

(3 + 4) + 5 = 12

9. How has phosphoproteomics been used to develop therapeutic drugs against diseases. Cite two examples. How can proteomics technologies be used to assess drug toxicity during clinical development. *[(CO5)(Analyse/IOCQ)]*

(6 + 6) = 12

Cognition Level	LOCQ	IOCQ	HOCQ
Percentage distribution	24	56	20

Course Outcome (CO):

After completing this course, students should be able to

1. Describe recent advances in genomics, transcriptomics, metabolomics and proteomics.
2. Explain basic and high throughput techniques in Genomics and their applications.
3. Explain basic and high throughput techniques in Proteomics and their applications.
4. List and discuss the use of genomics and proteomics in human health.
5. Propose appropriate methods for analysis of given sample type with respect to purpose of analysis.
6. Suggest and outline solution to theoretical and experimental problems in Genomics and Proteomics fields.

**LOCQ: Lower Order Cognitive Question; IOCQ: Intermediate Order Cognitive Question; HOCQ: Higher Order Cognitive Question.*