# GENOMICS AND PROTEOMICS
## (BIOT 5232)

**Time Allotted : 3 hrs** **Full Marks : 70**

***Figures out of the right margin indicate full marks.***

***Candidates are required to answer Group A and any 5 (five) from Group B to E, taking at least one from each group.***

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

## Group – A
## (Multiple Choice Type Questions)

1. Choose the correct alternative for the following: **10 × 1 = 10**

(i) SNP which is pronounced as "snips" stands for
(a) Small nuclear protein (b) Single nucleotide polymorphism
(c) Single nucleotide particle (d) Small nicking points.

(ii) 1. DNA polymorphism plays important role in evolution and speciation.
2. 99.9% nucleotide bases are exactly same in all people.
3. Chromosome 1 has the fewest genes and Y-chromosome has most genes.
4. BAC and YAC have been used in Human Genome Project.
(a) All are correct (b) All are incorrect
(c) 1,2,4 are correct (d) 2,3,4 are correct.

(iii) Physical maps are maps of locations of identifiable landmarks on a genomic DNA _____ inheritance patterns.
(a) remotely related to (b) related to
(c) regardless of (d) associated with

(iv) 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

(v) One centiMorgan is defined as _____ percentage of the total recombination events.
(a) One (b) Ten (c) 0.1 (d) 0.01

(vi) Denaturation of a highly helical protein having disulfide bridges and two phenylalanine's can be monitored as a function of temperature by which one of the following techniques?
(a) Monitoring the ratio of absorbance at 214 nm and 250 nm at various temperatures
(b) Monitoring the absorbance at 214 nm at various temperatures
(c) Estimating the -SH content heat denaturation
(d) Recording circular dichroism spectra at various temperatures.

(vii) Which one of the following enzyme properties are typically improved in non-medical industry applications?
(a) Thermostability (b) Specificity
(c) Catalytic efficiency (d) All of (a), (b) & (c).

(viii) Which of the following techniques have been used as part of a pathogenic biomarker system?
(a) Circular dichroism (b) X-ray crystallography
(c) MALDI-TOF (d) Fluorescence spectroscopy.

(ix) The active site of a protein is at most 10% of the protein's volume. The rest of the protein is necessary partially for which ONE of the following choices
(a) It is required to bring the active site residues into their correct spatial conformation.
(b) It provides the space for the levers and fulcra for conformational change in the protein to occur
(c) Both (a) and (b)
(d) None of the above.

(x) Co-immunoprecipitation is used for which purpose?
(a) To determine if a protein-of-interest binds to a specific DNA sequence
(b) To examine protein-protein interaction in the nucleus and in the cytoplasm
(c) To examine protein-protein interactions in the cytoplasm instead of the nucleus
(d) To allow protein to be expressed in mammalian cell culture.

## Group- B

2. (a) Why was it important to construct DNA libraries before the sequencing during HGP? *[(CO1)(Analyze/IOCQ)]*

(b) Considering that the genotypes at two loci are linked with each other, calculate the coefficient of linkage of linkage disequilibrium (LD) between the alleles. *[(CO1)(Calculate/HOCQ)]*

(c) Give an overview of the '1000 Genomes Project' with its goals and findings. *[(CO1)(Remember/LOCQ)]*

**4 + 4 + 4 = 12**

3. (a) Illustrate the main features of the anatomy of a mitochondrial genome. *[(CO4)(Remember/LOCQ)]*

(b) Analyze any one techniques of SNP genotyping with an illutration. *[(CO4)(Analyze/IOCQ)]*

(c) Write a brief note on 454 sequencing mentioning the novelty of the technique. *[(CO4)(Understand/LOCQ)]*

**4 + 4 + 4 = 12**

## Group – C

4. (a) Mention the approaches briefly how raw sequence data is assembled for genome assembly. *[(CO1)(Remember-Understand/LOCQ)]*

(b) Mention the role Phred and Phrap programs in genome sequencing projects. *[(CO1)(Understand/LOCQ)]*

**6 + (2 × 3) = 12**

5. (a) in order to study gene expression, microarray serves as an important tool where data analysis comes in a digitized fashion which needs a lot of processing-briefly evaluate the process along with suitable diagrammatic representation. *[(CO2)(Evaluate/HOCQ]*

(b) Justify the utility of unsupervised analysis in microarray data analysis and briefly discuss the bottom-up and top-bottom approach. *[(CO2)(Analyze/IOCQ]*

**6 + (4 + 2) = 12**

## Group - D

6. (a) "Sample preparation is the most vital step in 2-D gel electrophoresis" justify the statement. *[(CO3)(Justify/HOCQ)]*

(b) Explain the steps of 2D-DIGE in proteomics. *[(CO3)(Understand/IOCQ)]*

(c) Explain the basic principle of protein chips. What is application of protein chips in functional proteomics? *[(CO3)(Explain/LOCQ)]*

**3 + 3 + (4 + 2) = 12**

7. (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)(Understand/LOCQ)]*

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

(c) A protein was isolated from mouse tissue and subjected to a variety of investigations. After purification, a sample was subjected to MALDI-TOF mass spectrometry and the following data obtained.

| **m/z** | **673.9** | **725.5** | **784.3** | **852.3** | **991.3** |
|---|---|---|---|---|---|
| **Abundance (%)** | **55** | **82** | **95** | **61** | **32** |

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

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

## Group - E

8. (a) "The result of a protein NMR structure is an ensemble of structures, all of which are consistent with the experimentally determined restraints, but converge to the same protein fold." Explain this statement in the context of NMR derived structural parameters for proteins and the latter's use as restraints in a computation to arrive at the most stable native fold/structure.

*[(CO3)(Analyze/IOCQ)]*

(b) Explain how single wavelength anomalous diffraction is used in structure determination of proteins. Use the example of a sulphur containing protein.

*[(CO3)(Understand/LOCQ)]*

(c) How can the temperature factor in each atom of the structural model of a protein built from an x-ray diffraction pattern be used as a measure of "the dynamics of a protein crystal"? Justify your answer. *[(CO3)(Analyse/IOCQ)]*

**5 + 3 + 4 = 12**

9. (a) Use a diagram to represent the Insulin receptor. Use the same diagram to represent the binding of insulin to this receptor. What are the underlying molecular events? How can you use this example as an example of pharmaceutical proteomics representing protein-protein interactions?

*[(CO4)(Understand-explain/IOCQ)]*

(b) Itemize the two main methods of glycoprotein enrichment. Why are they considered complementary? The high throughput identification and characterization of glycoproteins has necessitated methodological and instrumental improvements in mass spectrometry and allied methods. Use a table to represent these improvements. *[(CO3)(Understand-apply/IOCQ)]*

(c) Hen egg white lysozyme (HEWL) has a relative molecular mass of approximately 14, 300. If mass spectroscopy can measure mass to a resolution within 0.01%, could the following be confidently distinguished from the unmodified protein? (i) N-terminal acetylation (ii) phosphorylation of a single serine residue.

*[(CO6)(Calculate-Analyse/IOCQ)]*

**4 + 4 + 4 = 12**

---

| *Cognition Level* | *LOCQ* | *IOCQ* | *HOCQ* |
|---|---|---|---|
| *Percentage distribution* | *39.58* | *42.71* | *17.71* |

**Course Outcome (CO):**

After completing this course, students should be able to

CO1. Describe recent advances in genomics, transcriptomics, metabolomics and proteomics.
CO2. Explain basic and high throughput techniques in Genomics and their applications.
CO3. Explain basic and high throughput techniques in Proteomics and their applications.
CO4. List and discuss the use of genomics and proteomics in human health.
CO5. Propose appropriate methods for analysis of given sample type with respect to purpose of analysis.
CO6. 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.*