

POST GRADUATE (P.G.) CERTIFICATE COURSE IN BIOINFORMATICS

(Syllabus for the session 2014-15 onwards)

CBI 101: Basics of Bioinformatics

MM: Th 80 + IA:20

Time: 3 Hours

Note: Nine questions are to be set in all and the candidates are required to attempt five questions including compulsory question. Question no. I is compulsory covering the whole syllabus. Out of remaining eight questions, two questions are to be set from each unit. Candidate is required to attempt four questions, selecting one question from each unit.

UNIT I

Overview of Bioinformatics and Information Technology

History, Scope and application, Internet and World Wide Web; Generation of computers; Concept of Networking; Introduction to Data Mining; Application of data mining in Bioinformatics.

UNIT II

Bioinformatics Resources

Biological databases, Basic classification – Sequence & Structure; Generalized & Specialized; Primary & Secondary, with example databases (Genbank, EMBL, DDBJ, INSDC, Swiss Prot, PIR, PDB, NDB, BLOCKS, Pfam, ProSITE, etc.); Literature databases.

UNIT III

Bioinformatics techniques

Sequence comparison and alignment; Local and Global Alignment – (Smith Waterman Algorithm; Needleman Wunsch Algorithm); Concept of Gap, Gap Penalty & Scoring Matrices (PAM, BLOSSUM); Dot Plot Analysis.

UNIT IV

Bioinformatics tools

Information retrieval system (Entrez, SRS); Sequence alignment tools (BLAST, FASTA, CLUSTAL-W/X, MUSCLE, TCOFFEE), Variants of BLAST (BLASTn, BLASTp, PSI-BLAST, PHI-BLAST, etc)

POST GRADUATE (P.G.) CERTIFICATE COURSE IN BIOINFORMATICS

(Syllabus for the session 2014-15 onwards)

CBI 102: Genomics and Proteomics MM: Th 80 + IA:20

Time: 3 Hours

Note: Nine questions are to be set in all and the candidates are required to attempt five questions including compulsory question. Question no. I is compulsory covering the whole syllabus. Out of remaining eight questions, two questions are to be set from each unit. Candidate is required to attempt four questions, selecting one question from each unit.

UNIT I

Introduction to Genomics

Introduction to the concept of genome, gene networks: basic concepts, Prediction of genes, promoters, splice sites, regulatory regions: basic principles, application of methods genome projects. Tools for genome analysis-RFLP, DNA fingerprinting, RAPD, PCR, Linkage and Pedigree analysis-physical and genetic mapping.

UNIT II

Genome sequencing projects

Microbes, plants and animals; Accessing and retrieving genome project information from web; Comparative genomics, Identification and classification using molecular markers-16S rRNA typing/sequencing, ESTs and SNPs. Human Genome Project. Large scale genome sequencing strategies, Genome assembly and annotation. Genome databases of Plants, animals and pathogens, Metagenomics: Concept and applications.

UNIT III

Proteomics

Concept of proteome, Protein analysis (includes measurement of concentration, amino-acid composition, N-terminal sequencing); 2-D electrophoresis of proteins; Microscale solution isoelectricfocusing; Peptide fingerprinting; LC/MS-MS for identification of proteins and modified proteins; MALDI-TOF; SAGE and Differential display proteomics, Protein-protein interactions, Yeast two hybrid system.

UNIT IV

Functional genomics and proteomics

Analysis of microarray data; Protein and peptide microarray-based technology; PCR-directed protein in situ arrays; Structural proteomics. Deriving function from sequence, Proteomics in drug discovery and toxicology.

POST GRADUATE (P.G.) CERTIFICATE COURSE IN BIOINFORMATICS

(Syllabus for the session 2014-15 onwards)

CBI 103: Structural Bioinformatics

MM: Th 80 + IA:20

Time: 3 Hours

Note: Nine questions are to be set in all and the candidates are required to attempt five questions including compulsory question. Question no. I is compulsory covering the whole syllabus. Out of remaining eight questions, two questions are to be set from each unit. Candidate is required to attempt four questions, selecting one question from each unit.

UNIT I

Macromolecular Structure

Conformational Analysis of protein, Forces that determine protein structure, polypeptide chain geometries, Ramachandran Map, potential energy calculations, observed values for rotation angles, structure comparison and alignment. Conformational analysis of nucleic acids and carbohydrates, general characteristics of nucleic acid structure – geometries, glycosidic bond, rotational isomers and ribose puckering - forces stabilizing ordered forms – base pairing, base stacking.

UNIT II

Computational structural biology

Overview of molecular modelling - Introduction and challenges; Molecular modelling methods; Conformational searching, Potential energy maps, Ramachandran maps, *Ab-initio* methods, Semi-empirical methods; Empirical methods, Conformational analysis-Introduction and Methods: Molecular fitting, Energy Minimisation. Molecular dynamics simulations of Bio-macromolecules.

UNIT III

Molecular visualization tools

Visualization of tertiary structures, quaternary structures, architectures and topologies of proteins and DNA using molecular visualization softwares such as RasMol, Cn3D, SPDBV, Chime, Mol4D, etc.

UNIT IV

Structure prediction tools and homology modelling

Prediction of secondary structures of proteins using at least 5 different methods with analysis and interpretation of the results. Comparison of the performance of the different methods for various classes of proteins. Internet based modeling tools. Prediction of tertiary structures of proteins using Homology Modeling approach: SWISSMODEL, SWISS-PDB Viewer Prediction of tertiary structures of proteins using at least 3 methods for fold recognition along with analysis and interpretation of results. Structure Databases – PDB, NDB, CCD - Structural Classification – SCOP, CATH, FSSP.

POST GRADUATE (P.G.) CERTIFICATE COURSE IN BIOINFORMATICS

(Syllabus for the session 2014-15 onwards)

CBI 104: Biostatistics

MM: Th 80 + IA:20

Time: 3 Hours

Note: Nine questions are to be set in all and the candidates are required to attempt five questions including compulsory question. Question no. I is compulsory covering the whole syllabus. Out of remaining eight questions, two questions are to be set from each unit. Candidate is required to attempt four questions, selecting one question from each unit.

UNIT I

Probability Concepts and Distributions for Analyzing Large Biological Data, Basic Concepts, Conditional Probability and Independence, Random Variables, Expected Value and Variance, Distributions of Random Variables, Joint and Marginal Distribution, Multivariate Distribution, Sampling Distribution.

UNIT II

Clustering: Unsupervised Learning In Large Biological Data, Measures of Similarity, Clustering, Assessment of Cluster Quality, Statistical Network Analysis For Biological Systems And Pathways, Boolean Network Modeling, Bayesian Belief Network, Modeling of Metabolic Networks.

UNIT III

Multidimensional Analysis and Visualization on Large Biomedical Data, Classical Multidimensional Visualization Techniques, Two-Dimensional Projections, Issues and Challenges, Systematic Exploration of Low-Dimensional Projections, One-Dimensional Histogram Ordering, Two-Dimensional Scatterplot Ordering.

UNIT IV

Samples and Sampling Distribution, Standard Error, significance level, Degrees of freedom, Tests of significance, tests for proportion, t and F tests Confidence Intervals, Contingency tables of χ^2 (Chi square) tests of goodness of fit and homogeneity, Correlation: Simple, Partial and Multiple Correlation, Methods of averages and least squares, polynomial fitting, Regression Analysis. Analysis of variance for one and two way classification.