



**PERIYAR UNIVERSITY
SALEM – 11**

**M.Sc., (BIOINFORMATICS)
Semester System
(Effective from the academic year 2021-2022)**

M.Sc. BIOINFORMATICS
CHOICE BASED CREDIT SYSTEM

(For the students admitted from the year 2021-22 onwards)

Part	Course Code	Course	Hrs.	Credit	Marks		
					CIA	EA	Total
Semester I							
Core I	21PBI01	Introductory Bioinformatics	4	4	25	75	100
Core II	21PBI02	Cell and Molecular Biology	4	4	25	75	100
Core III	21PBI03	Biomolecules	4	4	25	75	100
Core IV	21PBI04	Programming in C and C++	4	4	25	75	100
Elective-I	21PBIE01	Mathematics and Biostatistics	4	4	25	75	100
Practical-I	21PBIP01	Lab Course – I Cell and Molecular Biology Practical	6	4	40	60	100
Semester II							
Core V	21PBI05	Genomics and Proteomics	4	4	25	75	100
Core VI	21PBI06	Relational Database Management systems	4	4	25	75	100
Core VII	21PBI07	Structural Biology	4	4	25	75	100
Elective-II	21PBIE02	Genetic and Genetic engineering	4	4	25	75	100
Value Edu		Human Rights	2	1	25	75	100
Practical-II	21PBIP02	Lab Course- II Programming Lab	6	4	40	60	100
Practical-III	21PBIP03	Lab Course- III Sequence and Structural Analysis Lab	6	4	40	60	100
Semester III							
Core VIII	21PBI08	Molecular Modelling and Drug Design	4	4	25	75	100
Core VIX	21PBI09	Omics and Systems Biology	4	4	25	75	100
Core X	21PBI10	Perl, Python and Web Programming	4	4	25	75	100
Core XI	21PBI11	Big Data Analysis and Next Generation Sequencing	4	4	25	75	100
Elective-III	21PBIE03	Data Warehousing and Data Mining	4	4	25	75	100
Practical-IV	21PBIP04	Lab Course-IV Programming in Perl, Python, and Web	6	4	40	60	100
Semester IV							
Core XII	21PBI12	Research Methodology	4	4	25	75	100
Add-On		PHP Scripting Language	2	-	-	-	-
Project	21PBI13	Project Work		14			200

PROJECT WORK

Dissertation submission: 120

CIA: 30

Viva-voce: 50

Total: 200

Total Credits: 91

Total Marks: 2200

SEMESTER - I

INTRODUCTORY BIOINFORMATICS

UNIT I

Introduction to Bioinformatics; Computers in Biology to understand biological system; Basic commands of Windows, Unix and Linux operating systems; Concept of open resources in Bioinformatics.

UNIT II

Sequence analysis: Biological background for sequence analysis; Sequence alignment; Global, Local, Pairwise and Multiple sequence analysis; Algorithm for alignments; Database searching; Tools for sequence alignment.

UNIT III

Biological Databases: Database concepts; Introduction to data types and source; Protein sequence and structural databases; Nucleic acid database; Genome databases; Specialized databases; Carbohydrate databases; clinically relevant drug-drug interaction databases; Information retrieval from Biological databases; Entrez system.

UNIT IV

Cheminformatics: Introduction; Chemoinformatics tools; Chemical structure representation (SMILES and SMARTS); Chemical databases: CSD, ACD, WDI, ChEMBL, PUBCHEM, Chemical structure file formats; Structural isomers; Structure visualization.

UNIT V

Medical and Pharmacy Informatics: Introduction to Pharmacy Informatics, Medical transcription; Role of informatics to enhance the services provided by pharmaceutical care givers. Health Information Systems Architecture, Health data management, Medical coding, Telemedicine and Telehealth, Ethics in medical informatics, Pharmacy systems and automation, Informatics applications in pharmacy, survey and evaluation of on-line resources.

References:

- Mount, D. 2004. "Bioinformatics: Sequence and Genome Analysis"; Cold Spring Harbor Laboratory Press, New York.
- Baxevanis, A.D. and Francis Ouellette, B.F. 1998. "Bioinformatics – a practical guide to the analysis of Genes and Proteins"; John Wiley & Sons, UK.

- Lesk, A.M. 2002. "Introduction to Bioinformatics", First edition, Oxford University Press, UK.
- Jean-Michel, C. Notredame, C. 2003. "Bioinformatics for Dummies"; John Wiley & Sons.
- Ole Lund, Nielsen, M., Lundegaard, C. Kesmir, C. and Brnak, S. 2005. "Immunological Bioinformatics"; The MIT press.
- R. Durbin, S. Eddy, A. Krogh and G. Hitchison. 2003. Biological Sequence Analysis, Cambridge University Press, Eighth edition.
- Andrew R. Leach & Valerie J. Gillet. 2007. An Introduction to Chemoinformatics, Springer, Revised Edition.

CELL AND MOLECULAR BIOLOGY

Unit-I

Cell and Sub cellular Organelles: Organization of prokaryotic and eukaryotic cells. Composition of membranes, endocytosis and exocytosis, membrane transport types and function. Structure and functions of mitochondria and chloroplast (Mitochondria – structure, biogenesis; Chloroplast – structure, biogenesis; Molecular events of electron transport chain, ATP synthesis, photosynthesis and photorespiration), endoplasmic reticulum, ribosomes, Golgi, lysosomes, vacuoles, peroxisomes, glyoxisomes, cytoskeleton.

Unit-II

Cell cycle regulation and signal transduction: The cell cycle: phases, regulation by cyclins and cyclin dependent kinases. Checkpoints. Cell cycle and regulation, cell - cell interaction, ECM, Cell junctions (Adhesion, Gap and tight), plasmodesmata, desmosomes. Biochemistry and molecular biology of Cancer – malignant growth, tumour suppressor genes and oncogenes; chemical carcinogenesis, hormonal imbalances. Mechanism of signal transduction: cellular response to environment signals in animals and plants.

Unit-III

Chromatin, replication and repair: Nucleus: Structure and organization of prokaryotes and eukaryotic chromatin, the central dogma of molecular biology, Concept of the gene - gene, cistron, coding region (ORF), transcription unit, untranslated region (UTR), pseudogenes, euchromatin and heterochromatin. Typical structure of protein-coding genes in prokaryotes and eukaryotes. Split genes- exons and introns. DNA replication in bacteria and eukaryotes: initiation, elongation, termination. DNA repair- photoreactivation, excision repair, mismatch repair, SOS response, double strand break repair.

Unit-IV

Transcription, translation and mutation: Transcriptional processing Transcription in E. coli: Transcription eukaryotes: subunit structure, transcription factors, promoters, inhibitors. Post-transcriptional processing of prokaryotic and eukaryotic rRNA, and tRNA and eukaryotic mRNA. Mutations: point mutations and frame shift mutations. Suppressor mutations - nonsense and missense suppression. Mechanism of protein synthesis in bacteria and eukaryotes: amino acid activation, initiation, elongation and termination. Post-translational modifications and molecular chaperones. Protein localization: synthesis of secretory and membrane proteins, import into nucleus, mitochondria, chloroplast and peroxisome.

Unit-V

Cellular development: Gametogenesis-spermatogenesis and oogenesis, fertilization- artificial insemination-in vitro fertilization-types of egg and pattern of cleavage, blastulation, gastrulation, Fate of germ layer-microsporogenesis and megasporogenesis, pollen development-Amphimixis and parthenocarpy. Organogenesis. Development of Drosophila and Arabidopsis – spatial and temporal regulation of gene expression; life cycle and molecular biology of some pathogens – AIDS virus, tuberculosis, malarial parasite, hepatitis virus, filarial parasite and kala-azar parasite.

Reference Books

- Geoffrey Cooper, 2000, The Cell: A molecular approach. 2nd Eds. Sinauer Associates Inc.
- Gerald Karp, 2008. Cell and Molecular Biology: Concepts and Experiments, 5th Edition. Wiley
- Wilson and Walker, 2012. Principles and techniques of Biochemistry and Molecular biology. 7th ed. Cambridge University Press.
- Sambrook, 2012. Molecular Cloning. Cold Spring Harbor Laboratory, 4th ed.
- Lodish et al Molecular Cell biology 8th ed. Freeman, 2016
- David Clark, Nanette Pazdernik and Michelle McGehee, 2018. Molecular Biology. 3rd Edition. Academic Cell.

BIOMOLECULES

UNIT I

Introduction to Biomolecules – Structure and function of DNA, RNA, protein, carbohydrate and lipid. Folding pathways: Principles of protein folding, role of chaperons,

UNIT II

Molecular interactions: protein-protein, protein-DNA, DNA-Drug, Protein-Lipid, Protein-Ligand, Protein-Carbohydrate interaction, Metalloprotein. Pi ... Pi interactions, C-H...Pi interactions.

UNIT III

Structure and Stability of Proteins: Peptide bond; Ramachandran Plot; Calculation of conformation of different states of proteins; Properties of α -helix and β -sheets; Secondary structure prediction and determination; connectivities. Tertiary structure: Determination of state of tertiary structure; characteristic balance in rigidity and flexibility; Domain concept (α -, β -, α/β - and $\alpha+\beta$ -domains) and interacting motifs. Quaternary structure: Geometry, Symmetry and intermolecular interfaces.

UNIT IV

Structure and Stability of Carbohydrates: configuration and Conformation; Step Up and Step down reactions; Conversion of Aldos to Ketos and vice versa. Conformational stability ratings, Anomeric effect, Reverse anomeric effect; Structure and stability of polysaccharides, glycoproteins, glycolipids and proteoglycans. Structure and stability of Nucleic acids (DNA and RNA), topological structure, fine structure of DNA and its organization in genome.

UNIT V

Stereochemistry: Configuration and conformation and stability; Elements of symmetry, Chirality; RS-, EZ-, DL- and dl- system of nomenclature; Stereo-specific and stereo-selective reactions; Determination of relative configuration by asymmetric synthesis. Stereochemistry of proteins and nucleic acids.

Recommended Books:

- Principles of Biochemistry, Lehninger C Rs. Publ. (1982).
- Biochemistry, L. Stryer, W.H. Freeman, San Francisco.
- Schaum's Outline Series of Theory and Problems of Biochemistry, Philip W. Kuchel and G.B. Ralston. Int. Ed., McGraw-Hill Book Co.
- Problem Approaches in Biochemistry. Wood and Hood.
- Biochemistry by Voet and voet

- Biochemistry by Zubay
- Albert cotton, F. 1971. Chemical Application of Group Theory. John Wiley and Sons, Inc. New York. 386 pp.
- Spice, J. E. 1964. Chemical Bonding and Structure. Pergamon Press Ltd., Headington Hill Hall, Oxford. 395 pp.
- Winter, m. j. 1996. Chemical Bonding. Oxford University Press, Inc., New York. 91 pp.
- Ernest Eliel, 1996. Stereochemistry of carbon compounds, Prentice Hall

PROGRAMMING IN C AND C++

UNIT I

C language Introduction – Tokens – Keywords, Identifier, Variables, Constants, Operators – Expression – Data types –Operator precedence – Statement: Input statement, Output statement,

UNIT II

Control Statements - Conditional and Unconditional Control Statement – Looping Statement: while, do-while, for –nested loop – Arrays

UNIT III

C – Procedural Concepts: Structured Programming – Built-in library function – User defined functions – Pointer introduction – Passing pointer in a function – Structure – Union – File handle: Read and Write from a file.

UNIT IV

Introduction to C++- Basic concepts of OOPS-Introduction to C++,C vs C++-data types, variables, constants, operators and statements in C++ - Conditional and looping statements

UNIT V

C++ classes - Classes & Objects – Functions in C++ – function prototype-definition– Different forms of Constructor – Destructor – Copy constructor – Inheritance –Single, Multiple and Multi level inheritance – Function & operator overloading –inline functions – Friend and virtual functions – Overloaded functions.

Recommended Books:

- B.W.Kernighan and D.M. Ritchie The C Programming Language, II Edition. Prentice Hall of India.
- Larry Wall, Tom Christiansen & John Orwant Programming Perl –3 ed 2000- O’ Reilly.
- Mark Lutz, Programming Python, II Edition., O’ Reilly.
- E. Balagurusamy , Programming in C++ - Tata Mc. Graw Hill Edition.
- Byron Gottfried.1998. Programming with C (Schaum's Outline Series) - Tata McGrawHill Publishing Company.
- Robert Laffore , Object oriented programming with c++ -Waite series.
- Larry. Wall, Programming Perl - Tom Christiansen, Orielly Publications.

MATHEMATICS AND BIOSTATISTICS

UNIT I

Trigonometry: Trigonometric Functions, Series Expansion, Inverse, General Values, Graphs, Calculus: Limits, Continuity, Analysis, Differentiation (1D & Partial), Riemann Integration, Definite Integrals, Recursion Theorems, Ordinary and Partial Differential Equation: 1st Order and 2nd Order Ordinary Differential Equations, Self Adjoint Equations, Special Functions, Nature of Partial Differential Equations, Method of Separation of Variables.

UNIT II

Vector and Matrices: Vector Algebra, Vector Calculus, Basic Computations, Matrices. Measure Theory: Introductory Concepts, Borel Sets, Lebesgue Integration, Complex Variable: Complex Functions, DeMoivre's Theorem, Conformal Map, Complex Integration, Numerical Techniques: Basic Formalism, Methods for Solving Equations, Finding Eigenvalues & Eigenvectors, Solving ODE & PDE, Differentiation and integration.

UNIT III

Data Representation: Types of Numerical data, Tables and Graphs. Measures of central tendency: Arithmetic Mean, Weighted arithmetic mean, Median and Mode - Geometric mean and Harmonic mean. Measures of Dispersion: Range, Interquartile range, Average deviation, Standard deviation and Coefficient of variation, Lorenz curve. F ratio prediction and analysis of variance ANOVA.

UNIT IV

Linear Correlation and Regression: Types of Correlation, Methods of studying Correlation: Scatter diagram, Karl Pearson's Coefficient of Correlation, Spearman's Rank Correlation. Regression lines, Equations and Coefficients.

UNIT V

Probability: The concept of probability, Sample space, Independent events, Mutually exclusive events, Addition law of probability, Conditional probability, Bayes theorem.

Probability Distributions: Expected value and Variance, Binomial distribution, Poisson distribution, Normal distribution, Analysis of variance, Chi squared distribution, F- Test, Student's t distribution.

Theory of Sampling: The purpose of sampling, Principles of sampling, Methods of samplings, Techniques of non-probability sampling, Size of Sample, Sampling and Non-Sampling errors.

Recommended Books:

- Isaev, Berlin, A. 2004. "Introduction to mathematical methods in bioinformatics"; Springer.
- Raman, K.V. and Pal Sourav, P. 2005. "Mathematics in chemistry"; Vikas publishing house Pvt. Ltd., New Delhi.
- Stephenson, F.H. 2003. "Calculations in molecular biology and biotechnology: a guide to mathematics in the laboratory"; Amsterdam, Academic Press.
- Stephenson, G. and Radmore, P. M. 1990. "Advanced mathematical methods for engineering & science students"
- Viergever, M.A. and Todd- Pokropek, Andrew 1988. "Mathematics and computer science in medical imaging"
- Arfken, G. 1970. "Mathematical Methods for Physicists"; Academic Press, NY.
- Roman P., Pergamon. 1975. "Some Modern Mathematics for Physicists and Other Outsiders"; New York, Vol. 2, p. 660.
- Murray, J. D. 1989. "Mathematical Biology"; Springer Verlag.

- Segal, L. 1980. "Mathematical Models in Molecular and Cellular Biology"; Cambridge: Cambridge University Press.
- Balaguruswamy. "Numerical Methods"; Tata Mcgraw Hill.
- Haykin, S. 1999. "Neural Networks: A Comprehensive Foundation"; Prentice Hall.
- Zar, J.H. 1984. "Bio Statistical Methods"; Prentice Hall International Edition, USA.
- Warren, J., Gregory, E. and Grant, R. 2004. "Statistical methods in Bioinformatics"; First edition, Springer-Verlag, Berlin.
- Milton, J.S. 1992. "Statistical methods in the Biological and Health Sciences" Second Edition, McGraw Hill Publishers.
- Rosner, B. 2005. "Fundamentals of Biostatistics"; Duxbury Press.

LAB COURSE – I
CELL AND MOLECULAR BIOLOGY – PRACTICAL

1. Isolation of lymphocytes.
2. Karyotype analysis: Onion and human.
3. Microscopic examination of blood cells, epithelial cells, plant cells, tissue types.
4. Cell counting and viability
5. Estimation of DNA by diphenylamine method.
6. Estimation of RNA by orcinol method.
7. DNA electrophoresis in agarose gel, determination of molecular weight.
8. Restriction enzyme digestion and ligation.
9. SDS-PAGE of proteins and Western hybridization (demonstration)
10. Real-time qPCR (demonstration).

SEMESTER – II

GENOMICS AND PROTEOMICS

UNIT I

Genomes, transcriptomes and proteomes: Viral genome- bacteriophage and eukaryotic viruses. Prokaryotic genome- chromosome, linear or multipartite genome. Eukaryotic genome - yeast and human. Repetitive DNA sequences- tandem repeats of centromeres, mini satellites and micro satellites, interspersed repeats. Mobile genetic elements - Transposition and DNA transposons.

UNIT II

Restriction mapping, FISH, Sequence tagged site. Sequencing genomes-chain termination, chemical degradation, pyro sequencing. Sequence assembly - shot gun, clone contig methods. Human genome project- mapping phase- sequencing.

UNIT III

Genetic mapping-map repositories: Entrez genome map viewer- GDB, NCBI, OMIM, MGI/MGD Linkage map resources. Gene expression analysis by Micro arrays, Serial analysis of gene expression (SAGE). Applications of structural, functional and metagenomics. Comparative genomics and databases: PEDANT, COG, NCBI- gene plot, KEGG, MPG

UNIT IV

3D structure of protein- overview-protein secondary, tertiary, quaternary structure- protein Denaturation-protein folding- reverse turns- Ramachandran plot- Exspasy tools. Sequencing of protein-N and C terminal sequencing- proteomics- tools and application of proteomics.

UNIT V

Analytical proteomics-analytical protein and protein separation techniques- 1D SDS-PAGE, isoelectric focusing, 2D SDS-PAGE, image analysis of 2D gels-MALDI-TOF, HPLC-protein digestion techniques. Protein identification and analysis- Mass spectrometry-tandem mass spectrometry-peptide mass finger printing- SALSA algorithm- protein arrays.

Recommended Books:

- Daniel C Liebler, 2000. Introduction to Proteomics- tools for the new biology. Humana Press.
- Pennington, SR., Dunn, MJ. 2001. Proteomics- from protein sequence to function. Viva books Pvt. Ltd.
- David, M. 2004. Bioinformatics sequence and genome analysis. Cold spring harbour laboratory Press, NY.
- Benjamin Lewin. Genes VIII.2003. Benjamin-Cummings Pub Co.
- Sandy B. Primrose , Richard M. Twyman , Robert W. Old, 2002. Principles of Gene Manipulation and genomics. 7th Edition. Blackwell Science.
- Campbell A and Laurie J. Heyer. 2006. Discovering Genomics, Proteomics, and Bioinformatics. 2nd Edition. Pearson Publishers.
- Desmond S. T. Nicholl. 2002. Introduction to Genetic Engineering. 2nd Edition. Cambridge University Press.
- Richard J. Reece. 2003. Analysis of Genes and Genomes. John Wiley and Sons. High Education.

RELATIONAL DATABASE MANAGEMENT SYSTEMS

Unit-I

Introduction to Database Systems: Objectives of a Database System - Clarification on Data Independence - Advantages of a Database System - Approaches to Database Design - Conventional Files - Database Approach - Desirable Features of a DBS - Database Development Life Cycle -The Database System Environment: Levels of Architecture - Database Personnel Components of DBMS Suite - Database Management System Classifications

Unit-II

The Relational Model :Basic Concepts - Domains - Relations -Relational Database System - Identifying Relationships - Representing Relationships - Implementing Relationships - Innovation: Relation-Attributes List and Relationship List - Database Naming Convention -Database Modeling and Design: - The Database Model - Database Design - The E-R Model Revisited - Database Design via the E-R Model-The Extended Relational Model

Unit-III

The Structured Query Language: Important Facts- Advantages of SQL- Overview of Oracle's SQL Environment - Basic Concepts in a Typical Oracle Database Environment -Database Creation - Database Management - Tablespace Creation - Tablespace Management - Table Creation - Dropping or Modifying a Table- SQL Data Manipulation Statements: Insertion of Data - Update Operations - Deletion of Data - Commit and Rollback Operations -Basic Syntax for Queries- Simple Queries

Unit-IV

Introduction to MySQL: Main Features of MySQL - Main Components of MySQL - Alternate Storage Engines - Shortcomings of MySQL - Overview of Microsoft SQL Server : Introduction - Brief History -- Operating Environment - Microsoft SQL Server Editions - Main Features of Microsoft SQL Server -Main Components of Microsoft SQL Server Suite - Server Components - Management Tools - Development Tools - Client Connectivity - Code Samples - Shortcomings of Microsoft SQL Server

Unit-V

Database Administration - Database Installation, Creation, and Configuration - Database Security-Database Management - Database Backup and Recovery-Database Tuning- Distributed Database Systems: Preliminaries- Advantages of Distributed Database Systems- 12 Rules for Distributed Database Systems- Challenges to Distributed Database Systems - The Future of Distributed Database Systems- Overview of Object-Oriented Database Management Systems- Challenges for Object-Oriented Database Management Systems - Hybrid Approach

Recommended Books:

1. Elvis C. Foster, Database Systems A Pragmatic Approach Second Edition, Apress, 2016 (<http://file.allitebooks.com/20161113/Database%20Systems,%202nd%20Edition.pdf>)
2. Abraham Silberschatz , Henry F. Korth , S. Sudarshan ,Database System Concepts, McGraw-Hill, Sixth Edition
3. Raghu Ramakrishnan, Johannes Gehrke, Database Management Systems, McGraw-Hill
- 4 Abraham Silberchatz, Henry F. Korth, S. Sudharshan, Database System Concepts (5th Edition), McGraw Hill, 2002.
- 5 James Martin, Computer database organization, Prentice Hall of India, 1977.
- 6 James Martin, Principles of Database management, Prentice Hall of India, 1976.
- 7 Bipin C. Desai, An introduction to database systems, Galgotia publications pvt. Ltd.,New Delhi, 2003.
- 8 Peter Rob Carlos Coronel, Database systems, design, implementation & management, Course technology, 2000.
- 9 Database systems – A practical approach to design, implementation and management, Thomas cannolly and Carolyn begg, Pearson Education, 2002.
10. Vikram Vaswani, The Complete Reference MySQL, Tata McGraw-Hill, New Delhi, 2002.
- 11 Zoe Lacroix and Ternce Critchlow, Bioinformatics - Managing Scientific Data, Morgan Kaufmann publishers, 2003.

STRUCTURAL BIOLOGY

UNIT I

Nucleic acids, Amino acids, Adenylates, Carbohydrates, Lipids, Cofactors, Vitamins, and Hormones. General characteristics of nucleic acid structures (A, T, G, C, U), forces and stabilizing geometries, glycosidic bond, rotational isomers. Stabilizing ordered forms of DNA (A, B and Z), base pairing types, base stacking, tertiary structure of DNA (Supercoiled DNA), Melting of the DNA double helix (Hyperchromicity), Interaction with small ions and small molecules.

Unit II

Composition and levels of protein structure, Conformational analysis and forces that determine protein structures, geometries, phi, psi, omega angles, Ramachandran or steric contour diagram, allowed chi angles of side chains in proteins, hydrogen bonding, disulphide bonds, hydrophobic interactions, vanderwaals forces, potential energy calculations, alpha helices, beta sheets, helix to coil transition.

Unit III

General features and thermodynamic aspects of protein folding, folding kinetics, protein-ligand interactions, Scatchard plot, Relationship between the primary, secondary, and tertiary structure of proteins. Structural families and classes, multifunctional domains. Overview and Challenges in structural bioinformatics, structural genomics: High – throughput structural genomics using x-ray crystallography method

Unit IV

Rayleigh scattering, ultra-centrifugation, viscometry. Electron microscopy (SEM-TEM, AFM), luminescence (fluorescence & phosphorescence), Calorimetry, DSC, Mass spectrometry, LCMS, MALDI-TOF - X-ray diffraction: structure determination via single crystal diffraction, fibre diffraction; Neutron diffraction. XAFS. NMR spectroscopy (structure determination). ORD/CD, UV, IR, Laser Raman, ESR/EPR.

Unit V

Association of macromolecules, molecular conjugates, supramolecular interactions, protein-protein interactions, protein-nucleic acid interactions, lipid/membrane-protein interactions. Molecular mechanics and dynamics (Newtonian and Monte Carlo simulations), theoretical

principles and its importance towards *insilico* simulations, results of molecular dynamics calculations and their implications to biological function.

Recommended Books:

- Biophysics – An Introduction by Cotterill, Wiley Student Edition.
- Foundations of Biophysics by A.L. Stanford, Academic Press.
- Principles of protein structure by G Schulz and R H Schirmer, Springer Verlag.
- Principles of nucleic acid structure by Sanger, Springer Verlag.
- Introduction to Protein Science by Arthur M Lesk, Oxford University Press.
- Biological Spectroscopy by J. D. Campbell and R. A.Dwek, Plenum Press.
- A Textbook of Biochemistry and Biophysics by S.M Gopinath, Archers & Elevators International Publishing House, India. 1st Edition, 2014.
- Structural Bioinformatics by P.E. Bourne and H. Weissig (Eds.), John-Wiley and Sons, 2003.
- Biophysics by Vasantha Pattabhi and N.Gautham, Kluwer Academic Publishers, New York, (Narosa Publishing House, Chennai) 2002.

GENETICS AND GENETIC ENGINEERING

Unit-I

Basic principles of Genetics and Chromosomes: Definitions of genetics- phenotype, genotype, heterozygous, homozygous, allele, character, gene, gene locus, pure line, hybrid. Mendel's laws. Monohybrid cross, multiple alleles, dihybrid cross, test cross, backcross, epistasis. Genetics Chromosome structure. Polytene and lampbrush chromosomes. Types of chromosomes on the basis of centromere position. Karyotyping.

Unit-II

Population genetics, Linkage and Inheritance Linkage: Population genetics - The Hardy-Weinberg law. Linkage -definition, measurement, three factor crosses. Polygenic traits. Essential features of autosomal dominant, autosomal recessive, autosomal codominant, X-linked recessive, X-linked codominant and Y-linked inheritance. Brief outline of allosomal (Klinefelter syndrome), autosomal (Down syndrome) disorders. Brief outline of Non-Mendelian inheritance.

Unit-III

Restriction endonucleases, cloning vectors, and Gene transfer methods: Basic steps in gene cloning. Type II Restriction endonucleases nomenclature and types of cleavage. Cloning vectors: plasmids (pBR322 and pUC), phage vectors (λ), cosmids, BACs and YACs. Methods of ligation of insert and vector DNA molecules: cohesive end method, homopolymeric tailing, blunt-end ligation, linkers and adapters. calcium phosphate coprecipitation, electroporation, lipofection, viruses, microinjection. Choice of host organisms for cloning.

Unit-IV

Cloning strategies and Expression systems: Cloning strategies- genomic cloning, cDNA cloning. Differences between genomic and cDNA libraries. Screening of recombinants: marker inactivation (antibiotic resistance, blue-white selection), colony hybridization, immunoscreening, screening for protein activity. Factors affecting expression of cloned genes. Expression of eukaryotic genes in bacteria- expression vector, promoters, industrial protein production. Fusion proteins, strategies to enhance protein stability, secretion and metabolic load. Expression in eukaryotic cells: Expression in yeast- yeast vectors.

Unit-V

Molecular Mapping and Techniques in Gene Manipulation: Molecular mapping of genome: genetic and physical maps. Techniques in Gene Manipulation- Extraction and purification of nucleic acids. Probes: radioactive and nonradioactive. Blotting techniques: Southern, northern, and western. Principle and applications of DNA fingerprinting, in situ hybridization, PCR, RT-PCR, real-time qPCR. DNA Sequencing: Automated sequencing. Next-generation sequencing. RFLP, RAPD and AFLP analysis. Hazards and safety aspects of genetic engineering.

Recommended Books

- Lesk A, 2013. Introduction to Genomics. 4 th ed. Oxford Univ Press.
- Primrose, 2002. Principles of genome analysis. 3 rd ed. Wiley.
- T. A. Brown, 2007. Genomes. 4 th ed Garland Science.
- Dale and von Schantz, 2011. From Genes to Genomes: Concepts and applications of DNA technology. 3rd ed. Wiley-Interscience.
- Nicholls DTS, 2008. An Introduction to Genetic Engineering. 3rd ed. Cambridge Univ Press.
- Glick and Pasternak, 2009. Molecular Biotechnology. 4 th ed. ASM Press.
- Primrose, Twyman and Old, 2016. Principles of gene manipulation. 8 th ed. Wiley-Blackwell.

LAB COURSE II

PROGRAMMING IN BIOINFORMATICS

MS-Office and HTML.

- Working with MS-Office Packages –One Exercise each in Word, Excel, Power Point and Access.
- Working with HTML Tags and HTML Forms. Creating HTML Pages (At least five different pages to be created using all tags learnt).
- Basic commands in MS-DOS and command line execution in LINUX.

C Programming

- To print the size of char, int, float and double data types.
- To convert any given temperature in a) Centigrade scale to Fahrenheit scale & b) Fahrenheit to Centigrade scale.
- To find the area and circumference of a circle with any given radius.
- To find the volume of a cylinder and cone of any given radius and height.
- To compute the pH of a given solution for a given a) H^+ ion concentration and b) OH^- ion concentration.
- To find the average molecular weight of a double-stranded DNA of length n.
- To calculate the Body Mass Index (BMI) value.

C++ programming

- Write a C++ program to find the sum for the given variables using function with default arguments.
- Write a C++ program to swap the values of two variables and demonstrates a function using call by value.
- Write a C++ program to swap the values of two variables and demonstrates a function using Call by reference using reference type (&).
- Write a C++ program to swap the values of two variables and demonstrates a function using Call by reference using pointer (*).
- Write a C++ program to swap the values of two dynamically allocated variables and release the memory after swapping. (use new & delete operators)
- Write a program to find the largest, smallest & second largest of three numbers. (use inline function MAX and MIN to find largest & smallest of 2 numbers)

LAB COURSE III
LAB IN SEQUENCE AND STRUCTURAL ANALYSIS

- 1) Data retrieval using ENTREZ (sequence, structure and specialized databases)
- 2) Database file formats
- 3) Gene finding (Genscan)
- 4) Protein prediction
- 5) Sequence search
- 6) Sequence alignment
- 7) Phylogenetic tree construction
- 8) Sequence comparison
- 9) Proteome analysis using tools
- 10) Exon finding
- 11) Genome homology
- 12) Molecular visualization using Rasmol
- 13) Model building using space filling/ball and stick models
- 14) Find the molecular weight of a DNA with n base pairs in length
- 15) Find the molecular weight of a given dephosphorylated oligonucleotide sequence
- 16) Find the molecular weight of a given DNA sequence, after checking for phosphorylation
- 17) Computing amino acid composition of a given protein sequence.
- 18) Computing base composition of a given nucleotide sequence.
- 19) Small molecule building, (ISIS DRAW and CHEMSKETCH)
- 20) Homology Modeling (SWISS-MODEL, SPDBV)
- 21) Model structure refinement (SPDBV)
- 22) Model validation (What Check and ProCheck)
- 23) Docking using (AUTODOCK or AMBER)
- 24) 2D and 3D QSAR along with CoMFA and CoMSIA

SEMESTER-III

MOLECULAR MODELING AND DRUG DESIGN

UNIT I

Quantum mechanics & concepts in molecular modeling: Introduction – Coordinate systems and potential energy surfaces; introduction to quantum mechanics –Schrodinger wave equation – hydrogen molecule – Born-Oppenheimer approximation.

Unit II

Computational quantum mechanics: one electron atom, poly electronic atoms and molecules, Hartree Fock equations; calculating molecular properties using ab initio and semi empirical methods. Molecular mechanics: general features of molecular mechanics force field, bond stretching, angle bending, torsional terms, non-bonded interactions; force field parametrisation and transferability.

Unit III

Energy Minimisation and Computer Simulation: Energy Minimisation - Non-Derivative method, 1st and 2nd order minimisation methods. Computer Simulation Methods. Simple Thermodynamic Properties and Phase Space. Boundaries. Analyzing the Results of a Simulation and Estimating Errors. GROMACS and CNS.

Unit IV

Molecular Dynamics and Monte Carlo simulation: Introduction – Use of single Model time steps, multiple steps; Setting up MD, energy conservation in MD Simulation; Monte Carlo Simulation. BIOSUITE.

Unit V

Drug design: General approach to discovery of new drugs – Structure based drug design, Ligand -based drug design; Lead discovery- Lead optimization– lead modification – Physiochemical principles of drug action – Drug stereo chemistry –Drug action; Molecular Docking, AUTODOCK and HEX; SAR studies; Pharmacophore identification – QSAR, CoMFA.

Recommended Books:

- A.R.Leach. 1996. Molecular Modeling Principles and Application, 2nd edition, Longman Publications.
- D.Baxivanis and Foulette. 2001. Bioinformatics: A Practical Guide to the Analysis of Genes and Proteins, Wiley Indian Edition.

- T K Attwood, D J parry-Smith. 2005. Introduction to Bioinformatics, Pearson Education, 1st Edition, 11th Reprint.
- J.M.Haile. 1997. Molecular Dynamics Simulation Elementary Methods, John Wiley and Sons.
- Satya Prakash Gupta. 2008. QSAR and Molecular Modeling, Springer - Anamaya Publishers.

PERL, PYTHON AND WEB PROGRAMMING

UNIT I

Programming in PERL: Introduction, Basic Operators and Control Structures, Scalars, Lists, Hashes, Pattern Matching and Regular Expressions, Subroutines, Text and String Processing

UNIT II

Advance Programming in PERL: File handling; Object oriented programming:- Classes, Objects, Methods.

UNIT III

BioPERL Programming: General Bioperl classes, Sequences (Bio::Seq Class, Sequence Manipulation), Features and Location Classes (Extracting CDS), Alignments (AlignIO), Analysis (Blast, Genscan), Databases (Database Classes, Accessing a local Database)

UNIT IV

Python Programming: Overview, Data structures, Control Flow, Modules, Basic I/O, Exception Handling, Regular Expressions, File Manipulation, Classes, Standard library

UNIT V

WEB Programming: Common Gateway Interface:- HTML form elements, GET, POST & HEAD Method, CGI Environment Variables, Handling forms, Passing Parameters via CGI, Debugging CGI programs.

Recommended Books:

- D. Tisdall, Beginning Perl for Bioinformatics, O'Reilly, 2001
- Mark Lutz , Programming Python - 2nd Ed., O' Reilly, 2003.
- Eric Ladd, J.O'Donnell, Using HTML 4, XML and JAVA, Prentice Hall of India QUE, 1999.
- Brown, The complete reference - Perl, Tata Mc Graw Hill, New Delhi, 2004.

OMICS AND SYSTEMS BIOLOGY

UNIT I

Introduction and scope of proteomics: Components of a complex mixture and protein sequencing: MALDI TOF MS, QTrap, MS/MS, 2D Gel electrophoresis and protein microarrays. qRT PCR and proteomics. Proteomic approach for clinical studies; Protein biomarker discovery and validation- Body fluid profiles, blood disease profiles, diabetes profiles, infectious diseases.

Unit II

Protein arrays: Basic principles, computational methods for identification of polypeptides, Bioinformatics-based tools for analysis of proteomics data (Tools available at ExPASy proteomics server); databases (such as Inter Pro) Protein-protein interactions; database such as STEINGS and DIP; PPI Modeling in biological systems.

Unit III

Protein complexes and aggregates: Protein interaction networks, Regulatory networks, Structures of regulatory networks, Chou-Fasman and GOR methods for protein structure prediction, Neutral network models.

Unit IV

Glycomics: The challenge and Promise of Glycomics, Identification of carbohydrates, Glycolipids, Glycoproteins, Glycan microarray and Glycan Determinants, Metaglycomes, Glycan recognition molecules, Lipidomics, Fluxomics, Biomics: systems analysis of the biome.

Unit V

Systems biology: Introduction, Integrating Networks. Computer Simulation of the whole Cell. Computer Simulation of the Cell: Human Erythrocyte Model and its Application. Software for Modeling, E-CELL, V-CELL and GROMOS. Simulation of cellular subsystems, network of metabolites and enzymes, signal transduction networks, Gene 5 regulatory networks, metabolic pathways: databases such as KEGG, EMP, MetaCyc, AraCyc.

Recommended Books:

- Hiroaki Kitano (editor). 2001. “*Foundations of Systems Biology*”; MIT Press: ISBN 0-262-11266-3.
- Fall, CP., Marland, E. Wagner, J and Tyson, JJ (Editors). 2002. “*Computational Cell Biology*”; Springer Verlag: ISBN 0-387-95369-8
- Bock, G and Goode, JA (eds). 2002. “*In Silico Simulation of Biological Processes*”; Novartis Foundation Symposium 247. John Wiley & Sons: ISBN 0-470-84480-9
- Klipp, E., Herwig, R. Kowald, A. Wierling, C and Lehrach, H. 2005. “*Systems Biology in Practice*”; Wiley-VCH: ISBN 3-527-31078-9
- Alberghina, L. and Westerhoff, H(Editors). 2005. “*Systems Biology: Definitions and Perspectives, Topics in Current Genetics 13*”; Springer Verlag, ISBN 978-3540229681
- Szallasi, Z., Stelling, J. and V.Periwal (eds). 2006. “*System Modeling in Cellular Biology: From Concepts to Nuts and Bolts (Hardcover)*”; MIT Press., ISBN 0-262-19548-8
- James M. Bower, Hamid Bolouri, 2000. *Computational Modeling of Genetic and Biochemical Networks*, MIT Press.
- Julio Collado-Vides (Editor), Ralf Hofstadt (Editor) 2002. *Gene Regulation and Metabolism: Postgenomic Computational Approaches*, MIT Press.
- Sandy B. Primrose & Richard M. Twyman, 2004. *GENOMICS: Applications in Human Biology*, Blackwell Publishing Ltd.,
- C.W. Sensen, 2002, *Essentials of Genomics & Bioinformatics*, Wiley-VCH
- Malcolm Campbell & Laurie J. Heyer Laurie J. Heyer, 2006. *Discovering Genomics, Proteomics and bioinformatics*
- Rodney F. Boyer 2000. *Modern Experimental Biochemistry*, 3rd Edition, Publisher: Prentice Hall.

BIG DATA ANALYSIS AND NEXT GENERATION SEQUENCING

UNIT I

R-Statistical package: Essentials of R-package and libraries, mathematical operations, string operations, Data structures: vectors, data framers, lists, matrices, control loop: if else, while for loops. File Input/Outpur operations. R plots and the graphics library. Overview of statistical packages and Bioconductor libraries in R. Data representation: qualitative and quantitative data types, Tabulation and visual display of data, plotting line plot, scatter plot, frequency histograms, pie-chart, heat map and 3D plots.

UNIT II

Epigenomics: Different platforms of NGS, Overview of metagenomics principles, methylation of DNA and genetics, histone modifications, ChIP-chip ChIP-seq- techniques. Impact of transcriptome on biology, volume of data produced and imported repositories.

UNIT III

Microarray data analysis: Gene expression analysis, statistical method, relative merits of various platforms, Mapping algorithm such as Burro-wheeler, Measuring gene, IncRNA, siRNA, for RNA-seq NGS data. Sequence assembly concepts and challenges in assembling short reads. Hamiltonian cycle and de Brjin. Gene prediction and annotation.

UNIT IV

Identification of genetic variants from genome sequence: SNPs, SNVs, translocation, copy number variation. Concept behind genome-aide association studies. Introduction to various applications. Concepts and algorithms to measure transcriptional regulation; methylation and alternative splicing; small RNA analysis, validation of whole-genome database.

UNIT-V

Data analysis interpretation: gene expression analysis, differential expression analysis, Allele-specific expression, prioritizing genetic variants, non-synonym variants (SIFT, Polyphen), Synonyms variants, regulatory variants, statistical methods on rare variants, statistical considerations, Hidden Markov model annotating histone markers.

Recommended Books:

- R Development Core Team 2011. R: A Language and Environment for Statistical Computing. Vienna, Austria.
- Wang J. 2012. Next Generation Microarray Bioinformatics: Methods and Protocols. Hamana publishers. NY.
- Anna Maria Paganoni. 2014. Advances in Complex Data Modeling and Computational Methods in Statistics. Springer.
- Norman Maltoff-UC Davis. 2009. The Art of R Programming
- Hadley Wickham.2013. Advanced R Programming
- Venables, WN., Smith, DM., 2008. Network Theory: An Introduction to R.

DATA WAREHOUSING AND DATA MINING

Unit-I

Overview and Concepts: Need for data warehousing, Basic elements of data warehousing, Planning and Requirements: Project planning and management, Collecting the requirements. Architecture And Infrastructure: Architectural components, Infrastructure and metadata.

Unit-II

Data Design And Data Representation: Principles of dimensional modeling, Dimensional modeling advanced topics, data extraction, transformation and loading, data quality. Information Access and Delivery: Matching information to classes of users, OLAP in data warehouse, Data warehousing and the web. Implementation and Maintenance: Physical design process, data warehouse deployment, growth and maintenance.

Unit-III

Introduction: Basics of data mining, related concepts, Data mining techniques. Data Mining Algorithms: Classification, Clustering, Association rules. Knowledge Discovery: KDD Process. Web Mining: Web Content Mining, Web Structure Mining, Web Usage mining.

Unit-IV

Advanced Topics: Spatial mining, temporal mining. Visualization : Data generalization and summarization-based characterization, Analytical characterization: analysis of attribute relevance, Mining class comparisons: Discriminating between different classes, Mining descriptive statistical measures in large databases Data Mining Primitives, Languages, and System Architectures: Data mining primitives, Query language, Designing GUI based on a data mining query language.

Unit-V

DBMS- Introduction, overview and types. Relational and transactional Database. Relational database-Introduction to relational DB, Data Definition-Manipulationcontrol-Objects, Views, sequences and Synonyms. Data Abstraction; Data Models; Instances & Schemes; E-R Model - Entity and entity sets; Relations and relationship sets; E-R diagrams; Reducing E-R Diagrams to tables. Network Data Model: Basic concepts; Hierarchical Data Model: Basic Concepts; Multimedia Databases - Basic Concepts and Applications; Indexing and Hashing; Text Databases; Introduction to Distributed Database Processing, Data Security. ORACLE and SQL- introduction and functions in DBMS.

Recommended Books:

- Dunham, M.H. "Data Mining Introductory and Advanced Topics", Pearson Education.
- Kamber, H., Kaufmann, M. "Data Mining Concepts and Techniques".
- Adriaans, P., Zantinge, D. "Data Mining", Pearson Education Asia.
- Berry, M and Linoff, G. "Mastering Data Mining", John Wiley.
- Inmon, W.H. "Building the Data Warehouses", Wiley Dreamtech.
- Kimball, R. "The Data Warehouse Toolkit", John Wiley.
- Silberschatz, A., Korth, H.F. and Sudarshan, S. 2002. "Database system Concepts"; Fourth Edition, McGraw Hill Publishers.
- Date, C.J. 2000. "An introduction to Database systems"; Seventh Edition, Addison Wesley Publishers.
- Xiaohua Hu and Yi Pan 2007. Knowledge Discovery In Bioinformatics, John Wiley & Sons, Inc
- Yan-Qing Zhang 2009. Machine Learning Approach In Bioinformatics, John Wiley & Sons, Inc
- Feldman, R and Sanger, J. 2007. "The Text Mining Handbook: Advanced approaches in analyzing unstructured data"; Cambridge University Press.

- Pierre Baldi & Søren Brunak 2001. *Bioinformatics: The Machine Learning Approach*, The MIT Press, London.

LAB-IV
LAB IN PERL, PYTHON AND WEB
PROGRAMMING

Implement the following concepts using Perl:

1. Program to perform control statements and until, for, for each, while operations.
 - a) To find greatest of 3 numbers
 - b) To check whether a number is prime or not
 - c) Generating fibonacci series.
 - d) Palindrome.
2. To perform matrix arithmetic operations
3. Functions and subroutines
4. String Operations
 - a) Concatenation & String length.
 - b) Pattern recognition and searching
 - c) Translate DNA to RNA.
 - d) DNA to protein by using Perl module.

Implement the following concepts using Python:

1. Program to perform control statements and until, for, for each, while operations.
2. String Operations
3. Load, Extract and Retrieve features from GenBank and FASTA Files
4. Split a Sequence File into a Set of smaller files
5. Sequence file formats conversion
6. Manipulations of sequence files

Implement the following concepts Web programming:

1. Write Perl / Python programs to demonstrate Write Perl / PHP programs to display
2. Write Perl / Python programs to display Server, Cookie, Session information and Digital Clock

SEMESTER- IV

RESEARCH METHODOLOGY

UNIT I

Research: Definition, types, approaches, significance, research methods, criteria of good research, Literature collection and citation, research design: Basic principles of experimental design. Result analysis and interpretation, Report writing and manuscript preparation.

UNIT II

Bio-instrumentation: Microscopy-Light, Phase Contrast, Fluorescence, Confocal, Scanning and Transmission Electron Microscopy, Atomic Force Microscope, Scanning Tunneling Microscope - principles and applications.

UNIT III

Spectroscopy and spectrometry: Fluorescence, UV-Visible spectrophotometer, NMR and ESR Spectroscopy. Mass Spectrometry, GC-MS. Blotting Techniques-Principles and techniques of Southern, Northern and Western blotting techniques and hybridization.

UNIT IV

Algorithms: Protein and Nucleic acid sequence Algorithms: Sequence Databases, Use of the algorithms BLAST, Multiple sequence alignments and Clustering algorithms. Phylogeny: Evolutionary trees; Biological networks: Pathway analysis. Protein structure analysis: Protein structure databases; Protein structure comparison; Fold recognition; 3D-1D Profiles; Threading and Comparative structure modeling.

UNIT V

Computational applications: Introduction to internet – firebox, flock, mozilla, netscape., Clustal W, Wingene, oligo primer analysis, SPDBV, Swiss-prot, PDB, restriction analysis, primer design, data mining methods for sequence analysis, web-based tools for sequence searches, motif analysis and presentation.

Recommended Books:

1. Andrews D. Baxevanis and BF. Francis Ouellette, 2001. Bioinformatics-A Practical guide to the Analysis of Genes and Proteins, A John Wiley & Sons Publishers
2. David W.Mount, 2003. Bioinformatics-Sequence and Genome Analysis, CBS Publishers.

3. Ian Korf, Mark Xandell & Joseph Bedell, 2003. BLAST. O'REILLY Publisher.
4. Kothari C.R, 2004. Research Methodology - Methods and Techniques, New Age International (P) Ltd.
5. Gupta S.P, 2008. Statistical Methods. Sultan and Sons Company, New Delhi.

SEMESTER- IV
Add-On (Internship Course)
PHP SCRIPTING LANGUAGE

UNIT - I

Basic of Coding in PHP: Mixing PHP and HTML - Introducing Variables and Operators - PHP Variables.

UNIT - II

Displaying Dynamic Content - Sending E-Mail - Using File System - Uploading Files to Website.

UNIT - III

Establishing a connection - Creating a Database Table - Inserting Data into the Table - Selecting and Displaying Data.

UNIT - IV

System Planning - Adding Contacts Modifying Contacts - Deleting Contacts - Working with ontacts.

UNIT - V

Managing a Simple Mailing List: Mailing List Software - Developing Subscription Mechanism, Mailing Mechanism. Creating Custom Logs and Reports.

RECOMMEDED TBOOKS

1. Julie Meloni and Matt Telles, "PHP 6", 3rd Edition, Cengage Learning India Edition, 2009.
2. Julie Meloni and Matt Telles, "PHP 6 Fast and Easy Web Development", 1st Edition, Cengage Learning India Edition, 2008.
3. Steven Holzner, "The Complete Reference PHP", Indian Edition, McGrawHill Education, 2007.

SEMESTER- IV

PROJECT WORK & DISSERTATION

Dissertation submission: 120

CIA: 30

Viva-voce: 50

Total: 200

Question Paper Pattern

THEORY - Question Paper Pattern [EA] (Total Marks: 75)

Time: Three Hours

Maximum Marks: 75

Part - A (15 X 1 = 15 Marks)

Answer **ALL** Questions Multiple Choice

Part - B (2 X 5 = 10 Marks)

Answer **ANY TWO** Questions out of **Five**

Part - C (5 X 10 = 50 Marks)

Answer **ALL** Questions

Either (or) Type Five questions

(One question from Each Unit)