



School of Biotechnology
Devi Ahilya University, Indore

Syllabus

M. Sc. Bioinformatics

Revised Course Structure

Choice Based Credit System (CBCS)

The School of Biotechnology has choice based credit system (CBCS) in M.Sc. Biotechnology, students has to earn 87 actual credits and 16 virtual credits in total 04 semesters (two year duration). Maximum duration for completion of the course may be up to 04 years as per ordinance no. 31(revised).

If the student desires, credits for interdisciplinary/elective papers can be earned in any school/department.

Out of actual credits, 33 credits must be accrued from core papers, 06 credits from Discipline centric elective papers, 06 credits from Generic Electives, 05 from soft skill Enhancements courses, 22 credits from Practical's and 12 credits from project/ dissertation work. The 16 Virtual Credits have to be earned through Comprehensive Viva Voce examination conducted at the end of every semester (each of 04 credits). From these 103 credits, the credit for each subhead is as under:

S. No.	Type of Subject/Activity	Number of Subjects	Credit/Subject	Total Credit
01.	Core	11	03	33
02.	Discipline Centric Electives	04	1.5	06
03.	Generic Electives	02	03	06
04.	Soft Skill	04	(01*01 + 02*02)	05
05.	Skill Enhancement Courses	01	03	03
06.	Practical	03	(01*06+02*08)	22
07.	Project Work	01	12	12
08.	Comprehensive Viva Voce	1/ Semester	04	16
Total				103

S. No.	Core Subjects (09*03 = 27 Credit)	Elective (Generic) (03*02 = 06 Credit)	Elective (Discipline Centric) (06 Credit)	Soft Skills and Skill Enhancement Courses (30 Credit)	Project (12 Credit)
01.	Basic Mathematics	Design and Analysis of algorithms	Enzyme Technology	Internet & Web Based Programming (CGI PERL & HTML) (03 Credits)	Project Work of 12 Credits.
02.	Computer fundamentals and Biostatistics	Database management System	Immunoinformatics	Seminar (Soft Skills) (01*01 + 02*02 = 05 Credit)	
03.	Bio-molecules		Genomics & Proteomics	Practical's (In every Semester except last semester)	
04.	Cell and Developmental Biology		Metabolic Engineering & System Biology		
05.	Programming in C/C++		Pharmacogenomics		
06.	Molecular Biology		Microscopic Techniques For Image Processing		
07.	Biological Databases and Data Analysis (Bioinformatics-I)				
08.	Recombinant DNA Technology				
09.	Machine Learning Techniques & CADD(Bioinformatics II)				
10.	Structural Biology and Bioinformatics (Bioinformatics III)				
11.	Java Programming				
Total : 103 Credits					

M.Sc. Bioinformatics Syllabus

CONTENTS

Semester I

Course code	Title	Credits
BT BI 501	Basic Mathematics (Core)	03
BT BI 511	Computer fundamentals and Biostatistics (Core)	03
BT BI 521	Bio-molecules (Core)	03
BT BI 531	Cell and Developmental Biology (Core)	03
BT BI 541	Programming in C/C++(Core)	03
BT BI 551	Molecular Biology (Core)	03
BT BI 561	Practical	06
BT BI 571	Seminar & Communication Skills (Soft Skill Development)	01
	Comprehensive Viva Voce	04
	Total credits	29

Semester II

Course code	Title	Credits
BT BI 502	Biological Databases and Data Analysis (Bioinformatics-I): (Core)	03
BT BI 512	Recombinant DNA Technology (Core)	03
BT BI 522	Design and Analysis of algorithms (Generic Elective)	03
BT BI 532	Internet & Web Based Programming (CGI PERL & HTML) (Skill Enhancement Course)	03
BT BI 542	Immunoinformatics (Discipline Centric Elective)	1.5
BT BI 552	Genomics & Proteomics (Discipline Centric Elective)	1.5
BT BI 562	Enzyme & Enzyme Technology (Discipline Centric Elective)	03
BT BI 572	Practical	08
BT BI 582	Seminars (Soft Skill Development)	02
	Comprehensive Viva Voce	04
	Total Credits	29

03 Credits has to be earned from Discipline Centric electives by the students from SBT. One Generic Elective of 03 credits can be opted by the students from SBT itself or from any other department (UTD).

Semester III

Course code	Title	Credits
BT BI 601	Machine Learning Techniques & CADD(Bioinformatics II) (Core)	03
BT BI 611	Structural Biology and Bioinformatics (Bioinformatics III) (Core)	03
BT BI 621	Database management System (Generic Elective)	03
BT BI 631	Java Programming (Core)	03
BT BI 641	Metabolic Engineering & System Biology (Discipline Centric Elective)	1.5
BT BI 651	Pharmacogenomics (Discipline Centric Elective)	1.5
BT BI 661	Microscopic Techniques For Image Processing (Discipline Centric Elective)	1.5
BT BI 671	Assignments/ Practical	08
BT BI 681	Seminar (Soft Skill Development)	02
	Comprehensive Viva Voce	04
	Total Credits	29

Any 02 out of 03 Discipline Centric electives can be opted by the students from SBT. One Generic Elective of 03 credits can be opted by the students from SBT itself or from any other department (UTD).

Semester IV

Course code	Title	Credits
BT BI 602	Project Work	12
BT BI 612	Comprehensive Viva Voce	04
	Total Credits	16

Total credits for all semester in two years	103
--	------------

M.Sc. Bioinformatics

Semester – I

Course code	Title	Credits
BT BI 501	Basic Mathematics (Core)	03
BT BI 511	Computer fundamentals and Biostatistics (Core)	03
BT BI 521	Bio-molecules (Core)	03
BT BI 531	Cell and Developmental Biology (Core)	03
BT BI 541	Programming in C/C++(Core)	03
BT BI 551	Molecular Biology (Core)	03
BT BI 561	Practical	06
BT BI 571	Seminar & Communication Skills (Soft Skill Development)	01
	Comprehensive Viva Voce	04
	Total credits	29

Basic Mathematics: 3 Credits

UNIT 1

SET THEORY: Introduction, Examples of Sets, Representation of a set (Roaster form and Set builder form), Notation, Different types of sets, Set operations, , Complement of a set, Set Difference, Venn diagram De Morgan's Law.

MATRICES:

Matrices- Properties of Determinants, Minors and Cofactors, Multiplication of Determinants, Adjoint, Reciprocal, Symmetric Determinants, Cramer's rule; Different types of matrices, Matrix Operations, Transpose of a matrix, Adjoint of a square matrix, Inverse of a matrix, Eigen values and eigen vector.

UNIT 2

Vector Analysis: The concept of a Vector, Vector addition and subtraction, Products of two vectors-Dot product and Cross product, Products of three vectors- scalar triple product and vector triple product, Gradient, Divergence and Curl.

Limits and Continuity: Constants, Types of constants, Variables, Types of Variables, Function, Types of function, Right hand and left hand limits, Working rule for finding out the limit, Continuity: Define, point out discontinuity, Method of finding the continuity, Continuity from right and from left.

UNIT 3

The Binomial Theorem: Define, Binomial theorem for a positive integral index, Binomial Expansion, Finding middle term, general term, Binomial theorem for any index.

UNIT 4

Probability: Introduction, Events and types of events, Probability of events, Mutually exclusive events, favorable events, exhaustive events, independent events, addition theorem on probability, conditional probability, Multiplication theorem, Problem based on probability theorem, Bayes theorem.

UNIT 5

Differentiation and Integration: Introduction, Basic concepts and problems related to differentiation and integration.

Text references:

1. Introduction to Calculus & Analysis, Vol I and II by Richard Courant & Fritz John, Springer publisher.1999
2. Basic Mathematics by Serge A. Lang. Springer Publisher. 1988
3. Higher Engineering Mathematics (40th Ed), by B.S. Grewal and J.S. Grewal. Khanna Publishers, New Delhi. 2007

Computer Fundamentals & Biostatistics: 3

Credits

Unit I

Computer Organization: Block diagram of computer, Memory devices; Advantages and Limitations of Computers; Comparison of different operating systems DOS, Windows, Linux.

Number System: Binary, Hexadecimal, Octadecimal.

Internet Technologies: Web Services – WWW; URL; Servers: Client/ Server essentials - Domain Name Server; FTP server; E-mail server; WEB servers; Web publishing-Browsers-IP Addressing.

Database: Database concept; Database management system; Database browsing and Data retrieval; Data structures and Databases.

Sequence and Genome Databases: Databases such as GenBank; EMBL; DDBJ; Swissprot; PIR; MIPS; TIGR; TAIR; PlasmoDB; ECDC, Human Genome Project

Sequence file formats: GenBank, FASTA, PIR, ALN/ClustalW2, GCG/MSF, and PDB.

Unit II

Probability: Fundamental concepts of probability; sample space and events; independent events; mutually exclusive events; axioms of probability; conditional probability; additional and multiplication theorem of probability. Probability and analysis of one & two way samples;

Statistics: Central Limit theorem; Inference; Hypothesis; Critical region and error probabilities; Tests for proportion; Equality of proportions; Equality of means of normal populations (Variance known, Variance unknown)

Unit III

Measure of Central tendency and Dispersion; P-Value of the statistic; Confidence limit; T-Square Test; Chi-square test for independence; Introduction to one way & Two way ANOVA; Regression and Correlation coefficient; Use of statistical tools; preparation of graphs; histograms; charts and diagrams; Data Transformation.

Unit IV

Pair-wise Sequence Alignment: BLAST and its variants; FASTA.

Multiple sequence alignment: introduction

Phylogenetic Analysis: Introduction; Molecular Evolution; Cluster Analysis; Phylogenetic clustering by simple matching coefficients; Sequence comparison; Sequence pattern; : Tools used; Phylip and MEGA.

Unit V

Microarray: Goals of a Microarray experiment; Normalization of Microarray data; Detecting differential gene expression; Principal component analysis; Clustering of microarray data;

Structure Determination by X-ray crystallography; NMR spectroscopy

Structure Databases: The primary structure databases (PDB, NDB, and MMDB); secondary structure databases (SCOP, CATH, and Families of Structurally Similar Proteins). File formats for storage and dissemination of molecular structure.

In-silico Structure Prediction: Methods for modeling; Homology modeling; Threading and protein structure prediction; Structure-Structure comparison of macromolecules with reference to proteins.

Texts/References

1. **Wayne W. Daniel, *Biostatistics: A foundation for Analysis in the Health Sciences*, 8th Edition, Wiley, 2004.**
2. **Prem S. Mann, *Introductory Statistics*, 6th Edition, Wiley, 2006.**
3. **John A. Rice, *Mathematical Statistics and Data Analysis*, 3rd Edition, John A. Rice, Duxbury Press, 2006.**
4. **Campbell and Heyer, *Discovering Genomics, Proteomics, & Bioinformatics*, 2nd Edition, Benjamin Cummings, 2002.**
5. **Cynthia Gibas and Per Jambeck, *Developing Bioinformatics Computer Skill*, 1st Edition, O'Reilly Publication, 2001.**
6. **Mount D., *Bioinformatics: Sequence and Genome Analysis*, Cold Spring Harbor Laboratory Press, New York. 2004**
7. **C.R. Kothari, *Research Methodology: Methods and Techniques***

Bio-molecules - 3 Credits

Unit - I

Amino acids:

Structure and functional group properties; Peptides and covalent structure of proteins; Elucidation of primary and higher order structures; Evolution of protein structure; Structure-function relationships in model proteins like ribonuclease A, myoglobin, hemoglobin, chymotrypsin etc.; Tools to characterize expressed proteins.

Proteins- classification and separation, purification and criteria of homogeneity, end group analysis, hierarchy in structure, Ramachandran maps.

Unit – II

Sugars (Carbohydrates):

Mono, di, and polysaccharides; Suitability in the context of their different functions- cellular structure, energy storage, signaling; Glycosylation of other biomolecules - glycoproteins and glycolipids; Lipids - structure and properties of important members of storage and membrane lipids; lipoproteins.

Unit – III

Lipids –Structure and Classification of fatty acids; Structure of triglycerides and phospholipids, Chemical Reactions; structure and properties of important members of storage and membrane lipids; lipoproteins, Glycolipids, Sphingolipids, terpenes and steroids.

Unit - IV

Heterocyclic compounds and secondary metabolites in living systems - nucleotides, pigments, isoprenoids; classifications; functions and their properties in the body.

Principles of thermodynamics:

Classes of organic compounds and functional groups - atomic and molecular dimensions, space filling and ball and stick models.

Unit - V

Bioenergetics:

Basic principles; Equilibria and concept of free energy; Coupled processes; Glycolytic pathway; Kreb's cycle; Oxidative phosphorylation; Photosynthesis.

Texts/References

1. V.Voet and J.G.Voet, Biochemistry, 3rd edition, John Wiley, New York, 2004.
2. A.L. Lehninger, Principles of Biochemistry, 4th edition, W.H Freeman and Company, 2004.
3. L. Stryer, Biochemistry, 5th edition, W.H. Freeman and Company, 2002.

Cell and Developmental Biology- 3 Credits

Unit I

Cell Theory & Methods of Study:

Microscope and its modifications – Light, phase contrast and interference, Fluorescence, Confocal, Electron (TEM and SEM), Electron tunneling and Atomic Force Microscopy, etc.

Membrane Structure and Function

Structural models; Composition and dynamics; Transport of ions and macromolecules; Pumps, carriers and channels; Endo- and Exocytosis; Membrane carbohydrates and their significance in cellular recognition; Cellular junctions and adhesions; Structure and functional significance of plasmodesmata.

Unit - II

Organelles

Nucleus – Structure and function of nuclear envelope, lamina and nucleolus; Macromolecular trafficking; Chromatin organization and packaging; Cell cycle and control mechanisms; Mitochondria – structure, organization of respiratory chain complexes, ATP synthase, Structure- function relationship; Mitochondrial DNA and male sterility; Origin and evolution; Chloroplast– Structure-function relationship; Chloroplast DNA and its significance; Chloroplast biogenesis; Origin and evolution.

Endo-membrane System and Cellular Motility: Organization and role of microtubules and microfilaments; Cell shape and motility; Actin-binding proteins and their significance; Muscle organization and function; Molecular motors; Intermediate filaments; Extracellular matrix in plants and animals.

Unit – III

Cellular Movements and Pattern Formation

Laying of body axis planes; Differentiation of germ layers; Cellular polarity; Model plants like Fucus and Volvox; Maternal gene effects; Zygotic gene effects; Homeotic gene effects in Drosophila; Embryogenesis and early pattern formation in plants; Cell lineages and developmental control genes in Caenorhabditis.

Unit – IV

Cell cycle – Molecular events and model systems; Control mechanism; Apoptosis.

Cellular basis of differentiation and development - mitosis, gametogenesis and fertilization, development Arabidopsis; Spatial and temporal regulation of Gene Expression

Differentiation of Specialized Cells

Stem cell differentiation; Blood cell formation; Fibroblasts and their differentiation; Cellular basis of immunity; Differentiation of cancerous cells and role of proto-oncogenes; Phase changes in Salmonella; Mating cell types in yeast; Surface antigen changes in Trypanosomes; Heterocyst differentiation in Anabaena; Sex determination in Drosophila.

Unit - V

Biology of cancer; properties and features of cancer cells; oncogenes; tumor suppresser genes; mechanism of cancer; metagenesis; types of cancer

Genes, Mutation and Mutagenesis: UV and chemical mutagens; Types of mutation; Ames test for mutagenesis; Methods of genetic analysis.

Genetic Systems of Yeast and Neurospora.

Extra-Chromosomal Inheritance.

Texts/References

1. Lodish et al., Molecular cell Biology, 4th Edition, W.H. Freeman & Company, 2000.
2. Smith & Wood, Cell Biology, 2nd Edition, Chapman & Hall, London, 1996.
3. Watson et al., Molecular Biology of the gene, 5th Edition, Pearson Prentice Hall. USA, 2003.
4. B. M. Turner, Chromatin & Gene regulation, 1st Edition, Wiley-Blackwell, 2002.
5. Benjamin Lewin, Gene IX, 9th Edition, Jones and Barlett Publishers, 2007.

Programming in C/C++: 3 Credits

UNIT-I

C language: Introduction ; Tokens; Keywords; Identifier ; Variables; Constants; Operators ; Expression; Data types; Operator precedence

Statement: Input statement, Output statement, Conditional and Unconditional Control Statement – Looping Statement: while, do-while, for – nested loop – Arrays.

Overview of C++: Object oriented programming, Introducing C++ classes, Concepts of object oriented programming. Classes & Objects : Classes, Structure & classes, Union & Classes, Friend function, Friend classes, Inline function, Scope resolution operator, Static class members: Static data member, Static member function, Passing objects to function, Returning objects, Object assignment.

UNIT-II

Array and Pointers references: Array of objects, Pointers to object, Type checking C++ pointers, The This pointer, Pointer to derived types, Pointer to class members, References: Reference parameter, Passing references to objects, Returning reference, Independent reference, C++ 's dynamic allocation operators, Initializing allocated memory, Allocating Array, Allocating objects.

Constructor & Destructor: Introduction, Constructor, Parameterized constructor, Multiple constructor in a class, Constructor with default argument, Copy constructor, Default Argument, Constructing two dimensional Array, Destructor.

UNIT-III

Function & operator overloading : Function overloading, Overloading constructor function finding the address of an overloaded function, Operator Overloading: Creating a member operator function, Creating Prefix & Postfix forms of the increment & decrement operation, Overloading the shorthand operation (i.e. +=, -= etc), Operator overloading restrictions, Operator overloading using friend function, Overloading New & Delete, Overloading some special operators, Overloading [], (), -, comma operator, Overloading << .

UNIT-IV

Inheritance : Base class Access control, Inheritance & protected members, Protected base class inheritance, Inheriting multiple base classes, Constructors, destructors & Inheritance, When constructor & destructor function are executed, Passing parameters to base class constructors, Granting access.

Virtual functions & Polymorphism: Virtual base classes; Virtual function, Pure Virtual functions, early Vs. late binding

UNIT-V

String Handling: String declaration; String library functions; String Manipulation; Creating string objects, manipulating string objects, relational operators, string characteristics, Comparing and swapping

Sorting: Bubble sort, Selection sort, Insertion sort

Searching: Linear search, Binary search

Text References:

1. Programming in ANSI C by E. Balagurusamy. Tata McGrawHill Publishing Company Limited. 2007
2. Object Oriented Programming using C++ by Lafore, R. Galgotia Publishers. 2006
3. C: The Complete Reference by By **Herbert Schildt**, 4th Edition; McGraw-Hill/Osborne Media Publications

4. C++: The Complete Reference by By **Herbert Schildt**, 3th Edition; McGraw-Hill/Osborne Media Publications

Molecular Biology - 3 Credits

Unit I

Nucleotides: Structure; classification; Biosynthesis of purine and pyrimidine nucleotides from ribose including regulation, salvage pathways.

Genome organization

Organization of bacterial genome; Structure of eukaryotic chromosomes; Role of nuclear matrix in chromosome organization and function; Matrix binding proteins; Heterochromatin and Euchromatin; DNA re-association kinetics (Cot curve analysis); Repetitive and unique sequences; Satellite DNA; DNA melting and buoyant density; Nucleosome phasing; DNase I hypersensitive regions; DNA methylation & Imprinting.

Unit II

DNA Structure; Replication; Repair & Recombination

Structure of DNA - A-,B-, Z- and triplex DNA; Measurement of properties-Spectrophotometric, CD, AFM and Electron microscope analysis of DNA structure; Replication initiation, elongation and termination in prokaryotes and eukaryotes; Enzymes and accessory proteins; Fidelity; Replication of single strandedcircular DNA; Gene stability and DNA repair- enzymes; Photoreactivation; Nucleotide excision repair; Mismatch correction; SOS repair; Recombination: Homologous and non-homologous; Site specific recombination; Chi sequences in prokaryotes; Gene targeting; Gene disruption; FLP/FRT and Cre/Lox recombination.

Unit III

Prokaryotic & Eukaryotic Transcription

Prokaryotic Transcription; Transcription unit; Promoters- Constitutive and Inducible; Operators; Regulatory elements; Initiation; Attenuation; Termination-Rho-dependent and independent; Anti-termination; Transcriptional regulation-Positive and negative; Operon concept-lac, trp, ara, his, and gal operons; Transcriptional control in lambda phage; Transcript processing; Processing of tRNA and rRNA Eukaryotic transcription and regulation; RNA polymerase structure and assembly; RNA polymerase I, II, III; Eukaryotic promoters and enhancers; General Transcription factors; TATA binding proteins (TBP) and TBP associated factors (TAF); Activators and repressors; Transcriptional and post-transcriptional gene silencing

Post Transcriptional Modifications

Processing of hnRNA, tRNA, rRNA; 5'-Cap formation; 3'-end processing and polyadenylation; RNA editing; Nuclear export of mRNA; mRNA stability; Catalytic RNA.

RNA splicing: Nuclear splicing, spliceosome and small nuclear RNAs, group I and group II introns, Cis- and Trans-splicing reactions, tRNA splicing, alternate splicing.

Unit IV

Translation & Transport

Translation machinery; Ribosomes; Composition and assembly; Universal genetic code; Degeneracy of codons; Termination codons; Isoaccepting tRNA; Wobble hypothesis; Mechanism of initiation, elongation and termination; Co- and post-translational modifications; Genetic code in mitochondria; Transport of proteins and molecular chaperones; Protein stability; Protein turnover and degradation

Unit V

Antisense and Ribozyme Technology Molecular mechanism of antisense molecules, inhibition of splicing, polyadenylation and translation, disruption of RNA structure and capping, Biochemistry of ribozyme; hammer- head, hairpin and other ribozymes, strategies for designing ribozymes, Applications of antisense and ribozyme technologies. RNA interference.

Molecular Mapping of Genome Genetic and physical maps, physical mapping and map-based cloning, choice of mapping population, Simple sequence repeat loci, Southern and fluorescence in situ hybridization for genome analysis, Chromosome microdissection and microcloning, Molecular markers in genome analysis: RFLP, RAPD and AFLP analysis, Molecular markers linked to disease resistance genes.

Text/references

1. Benjamin Lewin, Gene IX, 9th Edition, Jones and Barlett Publishers, 2007.
2. J.D. Watson, N.H. Hopkins, J.W Roberts, J. A. Seitz & A.M. Weiner; Molecular Biology of the Gene, 6th Edition, Benjamin Cummings Publishing Company Inc, 2007.
3. Alberts et al; Molecular Biology of the Cell, 4th edition, Garland, 2002.

PRACTICAL [Total 06 Credits]

Lab on Molecular Biology

1. Plasmid DNA isolation and DNA quantization: Plasmid minipreps
2. Restriction digestion

3. Preparation of competent cells.
4. Agarose gel electrophoresis
3. Restriction Enzyme digestion of DNA
4. Purification of DNA from an agarose gel
5. DNA Ligation
6. Transformation of E.coli with standard plasmids, Calculation of transformation efficiency
7. Cloning of genomic DNA in standard plasmid vectors
8. Confirmation of the insert, Miniprep of recombinant plasmid DNA, Restriction mapping
9. Polymerase Chain reaction, using standard 16srRNA eubacterial primers
10. RFLP analysis of the PCR product
11. Transformation of yeast *Saccharomyces cerevisiae*

Lab on Computer Fundamentals and Biostatistics

1. Introduction to MS EXCEL-Use of worksheet to enter data, edit data, copy data, move data. Use of in-built statistical functions for computations of Mean, S.D., Correlation, regression coefficients etc.
2. Use of bar diagram, histogram, scatter plots, etc. graphical tools in EXCEL for presentation of data. Introduction to SYSTAT package.
3. Searching PubMed,
4. Bioinformatics Resources: NCBI, EBI, DDBJ, RCSB, ExPASy
5. BLAST BLASTn, BLASTp, PSI-BLAST,
6. Sequence manipulation Suite: BioEDIT
7. Multiple sequence alignment: Clustal W/ Clustal X and T-Coffee
8. Primer designing by primer 3
9. Phylogentic Analysis.
10. Protein Modeling,
11. Protein structure Analysis, Docking,
12. Sequence file formats: GenBank, FASTA, GCG, MSF

Lab on Biochemistry and Analytical Techniques

1. To prepare an Acetic-Na Acetate Buffer system and validate the Henderson-Hasselbach equation.
 2. To determine an unknown protein concentration by plotting a standard graph of BSA using UV-Vis Spectrophotometer and validating the Beer- Lambert's Law.
 3. Titration of Amino Acids and separation of aliphatic, aromatic and polar amino acids by TLC.
 4. AN ENZYME PURIFICATION THEME (such as E.coli Alkaline phosphatase or any enzyme of the institutions choice).
- (a) Preparation of cell-free lysates

- (b) Ammonium Sulfate precipitation
- (c) Ion-exchange Chromatography
- (d) Gel Filtration

Lab on Object Oriented Programming through 'C++'

1. Display a protein details using escape sequence
2. Calculate rotations per minute [$\text{rpm} = 1000 \sqrt{\text{RCF}} / 11.17r$]
3. Create amino acid dictionary using switch construct
4. Identify the glucose level in a blood using if - else if construct [The glucose level is identified by <70 – hypoglycemia, 70-180 hyperglycemia, > 180 diabetics]
5. Identify the type of two peptides using nested if [peptide length is < 8 small, poly otherwise]
6. Create a class which shows the various forms of constructors
7. Inheritance implementation
8. Function overloading example
9. Operator overloading example
10. Dynamic polymorphism implementation
11. Writing C programs for Bioinformatics applications: (10 assignments)
 - a) Extract a protein or nucleic acid sequence from any of the databank files (GenBank entry, Swiss-Prot, EMBL entry etc.) Syllabus
 - b) Interconverting the sequence from one databank format to the other eg. GenBank format to FASTA format, FASTA to PIR format etc.
 - c) Determining the base composition in a nucleic acid sequence and amino acid composition in a protein sequence.
 - d) Generating the complimentary sequence of a DNA sequence Pattern search algorithms
 - e) Search for a specific oligonucleotide pattern (eg. GAACATCC) in a given DNA sequence.
 - f) Find the position where a specific sequence say “GGTCCCGAC” will hybridize a given DNA sequence.
 - g) Find the restriction enzyme cleavage sites eg. where PVUZ, ECORI etc. will cut the DNA.
 - h) Locate palindromic sequence stretches in a DNA sequence.
 - i) Count the number of Open Reading frames (ORF's) in a DNA sequence.
 - j) Calculate the codon usage in a nucleic acid sequence.
 - k) Translate a DNA sequence into protein sequence in the forward and reverse frames.
 - l) Implementation of the Needleman-Wunsch algorithm for pair wise alignment and testing alignment score with randomized pairs of sequences also.

M.Sc. Bioinformatics

SEMESTER II

Course code	Title	Credits
BT BI 502	Biological Databases and Data Analysis (Bioinformatics-I): (Core)	03
BT BI 512	Recombinant DNA Technology (Core)	03
BT BI 522	Design and Analysis of algorithms (Generic Elective)	03
BT BI 532	Internet & Web Based Programming (CGI PERL & HTML) (Skill Enhancement Course)	03
BT BI 542	Immunoinformatics (Discipline Centric Elective)	1.5
BT BI 552	Genomics & Proteomics (Discipline Centric Elective)	1.5
BT BI 562	Enzyme & Enzyme Technology (Discipline Centric Elective)	03
BT BI 572	Practical	08
BT BI 582	Seminars (Soft Skill Development)	02
	Comprehensive Viva Voce	04
	Total Credits	29

03 Credits has to be earned from Discipline Centric electives by the students from SBT. One Generic Elective of 03 credits can be opted by the students from SBT itself or from any other department (UTD).

Biological Databases and Data Analysis (Bioinformatics-I): 3 Credits

UNIT I

Introduction and Bioinformatics Resources: Knowledge of various databases and bioinformatics tools available at these resources, the major content of the databases, Literature databases

Sequence databases (EMBL, GenBank, DDBJ, SWISS-PROT, PIR, TrEMBL)

Protein family/domain databases (PROSITE, PRINTS, Pfam, SMART)

Cluster databases (Prodom, Systems)

Specialist databases (Flybase, Kegg)

Search engines (SRS, Entrez)

Sequence analysis: Basic concepts of sequence similarity, identity and homology, definitions of homologues, orthologues, paralogues.

Scoring matrices: basic concept of a scoring matrix, PAM and BLOSUM series.

UNIT II

Pair Wise Sequence Alignment:

- 1) Dynamic Programming** – Smith and Waterman & Needleman and Wunsch Algorithm; Use of Pair-wise alignments for analysis of Nucleic acid and protein sequences and interpretation of results.
- 2) Heuristic Based Methods:** BLAST and FASTA algorithms, various versions of basic BLAST and FASTA.

Multiple sequence alignments (MSA): The need for MSA, basic concepts of various approaches for MSA (e.g. progressive, hierarchical etc.). Algorithm of CLUSTALW and PileUp and their application for sequence analysis (including interpretation of results), concept of dendrogram and its interpretation, Use of HMM-based Algorithm for MSA (e.g. SAM method)

UNIT III

Phylogeny: Phylogenetic analysis, Definition and description of phylogenetic trees and various types of trees, Method of construction of Phylogenetic trees [distance based method (UPGMA, NJ), Maximum Parsimony and Maximum Likelihood method]. The Fitch/Margoliash method; Character-based Methods – maximum parsimony, maximum likelihood; Phylogenetic softwares – PAUP, PHYLIP, MacClade

UNIT IV

Sequence patterns and profiles: Basic concept and definition of sequence patterns, motifs and profiles, various types of pattern representations viz. consensus, regular expression (Prosite-type) and sequence profiles; profile-based database searches using PSI-BLAST, analysis and interpretation of profile-based searches. Biological motifs (consensus, regular expressions, profiles, PSSMs, HMMs)

Algorithms for derivation and searching sequence patterns: MeMe, PHI-BLAST, SCanProsite and PRATT. Algorithms for generation of sequence profiles: Profile Analysis method of Gribskov, HMMer.

UNIT V

Data production and data flow (mapping, DNA sequencing, generation of scaffolds & contigs)

Gene prediction (ab initio & similarity based)

Genome annotation (pipelines, databases)

Text References:

1. Bioinformatics: Sequence and Genome Analysis by Mount D., Cold Spring Harbor Laboratory Press, New York. 2004
2. Bioinformatics- a Practical Guide to the Analysis of Genes and Proteins by Baxevanis, A.D. and Francis Ouellette, B.F., Wiley India Pvt Ltd. 2009
3. Bioinformatics for Dummies by Jean-michel Claverie Cedric Notredame. Publisher: Dummies
4. Introduction to Bioinformatics by Aurther M lesk
5. Developing Bioinformatics Computer Skills By: Cynthia Gibas, Per Jambeck
6. Structural Bioinformatics

RECOMBINANT DNA TECHNOLOGY : 3 Credits

Unit I

Basics Concepts : DNA Structure and properties; Restriction Enzymes; DNA ligase, Klenow enzyme, T4 DNA polymerase, Polynucleotide kinase, Alkaline phosphatase; Cohesive and blunt end ligation; Linkers; Adaptors; Homopolymeric tailing; Labeling of DNA: Nick translation, Random priming, Radioactive and non-radioactive probes, Hybridization techniques: Northern, Southern and Colony hybridization, Fluorescence in situ hybridization; Chromatin Immunoprecipitation; DNA-Protein Interactions- Electromobility shift assay; DNase footprinting; Methyl interference assay

Unit II

Cloning Vectors Plasmids: Bacteriophages; M13 mp vectors; PUC19 and Bluescript vectors, Phagemids; Lambda vectors; Insertion and Replacement vectors; Cosmids; Artificial chromosome vectors (YACs; BACs); Animal Virus derived vectors-SV-40; vaccinia/baculo& retroviral vectors; Expression vectors; pMal; GST; pET-based vectors; Protein purification; His-tag; GST -tag; MBP-tag etc.; Intein- based vectors; Inclusion bodies; Methodologies to reduce formation of inclusion bodies; Baculovirus and pichia vectors system, Plant based vectors, Ti and Ri as vectors, Yeast vectors, Shuttle vectors

Unit III

Cloning Methodologies: Insertion of Foreign DNA into Host Cells; Transformation; Construction of libraries; Isolation of mRNA and total RNA; cDNA and genomic libraries; cDNA and genomic cloning; Expression cloning; Jumping and hopping libraries; Southwestern and Far-western cloning; Protein-protein interactive cloning and Yeast two hybrid system; Phage display; Principles in maximizing gene expression

Unit IV

PCR and Its Applications: Primer design; Fidelity of thermostable enzymes; DNA polymerases; Types of PCR – multiplex, nested, reverse transcriptase, real time PCR, touchdown PCR, hot start PCR, colony PCR, cloning of PCR products; T-vectors; Proof reading enzymes; PCR in gene recombination; Deletion; addition; Overlap extension; and SOEing; Site specific mutagenesis; PCR in molecular diagnostics; Viral and bacterial detection; PCR based mutagenesis, Mutation detection: SSCP, DGGE, RFLP, Oligo Ligation Assay (OLA), MCC (Mismatch Chemical Cleavage, ASA (Allele-Specific Amplification), PTT (Protein Truncation Test)

Unit V

Sequencing methods: Enzymatic DNA sequencing; Chemical sequencing of DNA; Automated DNA sequencing; RNA sequencing; Chemical Synthesis of oligonucleotides; Introduction of DNA into mammalian cells; Transfection techniques; Gene silencing techniques; Introduction to siRNA; siRNA technology; Micro RNA; Construction of siRNA vectors; Principle and application of gene silencing; Gene knockouts and Gene Therapy; Creation of knock out mice; Disease model; Somatic and germ-line therapy- in vivo and ex-vivo; Suicide gene therapy; Gene replacement; Gene targeting; Transgenics; cDNA and intragenic arrays; Differential gene expression and protein array.

Text/references

1. S.B. Primrose, R.M. Twyman and R.W. Old; Principles of Gene Manipulation. 6th Edition, S.B. University Press, 2001.
2. J. Sambrook and D.W. Russel; Molecular Cloning: A Laboratory Manual, Vols 1- 3, CSHL, 2001.
3. Brown TA, Genomes, 3rd ed. Garland Science 2006
4. Selected papers from scientific journals.
5. Technical Literature from Stratagene, Promega, Novagen, New England Biolab etc.

Design & Analysis of Algorithm: 3 Credits

UNIT I

Basic Concepts of Algorithms: Introduction; Notion of Algorithm; Fundamentals of Algorithmic Solving; Important Problem types; Fundamentals of the Analysis Framework – Asymptotic Notations and Basic Efficiency Classes. Big 'O' notations, Time and space complexity of algorithms and common functions; String Matching: Naïve algorithm, Boyer Moore algorithm.

UNIT II

Mathematical Aspects and Analysis of Algorithms: Mathematical Analysis of Non-recursive Algorithm – Mathematical Analysis of Recursive Algorithm – Example: Fibonacci Numbers – Empirical Analysis of Algorithms – Algorithm Visualization.

UNIT III

Analysis of Sorting and Searching Algorithms: Brute Force; Selection Sort and Bubble Sort; Sequential Search and Brute-force string matching; Divide and conquer; Merge sort; Quick Sort; Binary Search; Binary tree-Traversal and Related Properties; Decrease and Conquer; Insertion Sort; Depth first Search and Breadth First Search.

UNIT IV

Algorithmic Techniques: Transform and conquer; Presorting; Balanced Search trees; AVL Trees; Heaps and Heap sort; Dynamic Programming; Warshall's and Floyd's Algorithm; Optimal Binary Search trees; Greedy Techniques; Prim's Algorithm; Kruskal's Algorithm; Single-Source Shortest Path: Dijkstra's Algorithm; Huffman trees; Minimum Spanning Tree.

UNIT V

Algorithm Design Methods: Backtracking; n-Queen's Problem; Hamiltonian Circuit problem; Subset-Sum problem; Branch and bound; Assignment problem; Knapsack problem; Traveling salesman problem.

Text References:-

1. Design and analysis of algorithms by Horrowick
2. Design and analysis of algorithms by Corman
3. Data structure by Baluja
4. Fundamentals of Algorithms by E. Horowitz and S. Sahani., Galgotia Book source Pvt. Ltd. 1999
5. Data Structures by Seymour Lipschutz., Tata Mc-Graw-Hill publication. 2007

Internet & Web Based Programming (CGI, Perl & HTML):

3 Credits

UNIT I

Internet Basics: The Basics of the Internet, Concepts of a Domain, Networking concepts, IP Addressing, Resolving Domain Names, Structure of an IP address, Overview of TCP/IP and its services, The World Wide Web, FTP and Telnet.

UNIT II

Hyper text markup language (HTML): How a Web Browser communicates with a web server, what is HTML and various HTML tags, Commonly used HTML commands, Lists, Adding Graphics to HTML documents, to create and use Tables, the concept of Hyperlink, Types of Hyperlinks, Introduction to Frames, Using the <Frameset> and the <Frame> tag. Other tags and versions of HTML such as DHTML and XML.

UNIT III

Common Gateway Interface (CGI): The concept of CGI, Why CGI is used, How CGI works, The two methods of Data submissions, the differences in the two methods of submissions, the importance of Environment variables in a CGI program, the basic steps required to process from information in a CGI program, Why Perl is the language of choice for programming in CGI.

UNIT IV

Perl Language: The basics of the Perl Language, the concept of Perl Strings and their types, the values that can be stored in scalar variables, Arrays, how to extract information from both types of Arrays, the importance of the special Hash Array, Performing operations & Controlling program Flow, Perl Functions, File Handling. Literal Representation of a Hash, Hash Functions, Using Hashes for the Genetic Code, Gene Expression Data Using Hashes

UNIT IV

Perl applications for biological data:

Bioperl: Introduction to Bioperl, Installing procedures, Architectures, 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), Implementing REBASE

Text Reference:

1. HTML/PERL/JAVA – Ivan Bayross
2. Beginning Perl for Bioinformatics (1 st Edition) by Tisdall, J., O’Reilly Publishers. 2004
3. Learning Perl (5th Edition) by Randal L. Schwartz, Tom Phoenix and Brain d Foy, O’Reilly Publishers, 2008
4. Advance Perl Programming by Sriram Srinivasan
5. Writing CGI application with Perl by Kevin Meltzer and Brent Michalski

Immunoinformatics :1.5 Credits

Unit I

Immune systems: Innate and adaptive immunity in vertebrates,

Antibodies: Immunoglobulins, Immunoglobulin classes and subclasses, Immunogenetics, Antibody numbering

The major histocompatibility complex (MHC): MHC polymorphism: Causes of MHC polymorphism, MHC supertypes. HLA nomenclature

Unit II

Antigen processing and presentation

Contemporary challenges to the immune system: Infectious diseases, clustering of infectious disease organisms, autoimmune diseases.

Membrane receptors for antigen: The B-cell surface receptor for antigen (BCR), CDR and LDR regions, The T-cell surface receptor for antigen (TCR), Antigen recognition diversity

Epitopes: Affinity Maturation, Recognition of Antigen by B cells, Neutralizing Antibody, Prediction of epitopes.

Unit III

Vaccine design: Rational Vaccine design, Categories of vaccines, Polytype vaccines, Therapeutic vaccines, Evolution and escape due to variations.

Mathematical models the immune system: Reverse immunology and approaches in computer aided vaccine design,

Viral bioinformatics: computational views of host and pathogen.

Unit IV

Browsing and searching immunological databases: Databases of epitopes, IMGT/HLA sequence database, BciPep, Epitome, CED, AgAbDb, Antibody:, Sequence analysis in immunology: Alignments, Molecular evolution and phylogeny, Prediction of functional features of biological sequences.

Methods applied in immunological bioinformatics: Sequence weighing methods, Pseudocount Correction methods, and Performance measures for prediction methods

Prediction of cytotoxic and helper T cell epitopes: MHC Class I epitopes, MHC Class II epitopes.

Web-based tools for vaccine design: Databases of MHC ligands, Prediction servers.

Text/Reference:

1. Immunological Bioinformatics (2005) by Ole Lund, Morten Nielsen, Claus Lundegaard, Can Kesmir, and Soren Brnak, The MIT press.
2. Immunoinformatics: Bioinformatics Strategies for Better Understanding of Immune Function (2003), Wiley; 1 edition.
3. Essential Immunology. Blackwell Science by Ivan M. Roitt and Peter J. Delves.

Genomics and Proteomics -1.5 Credits

Unit I

Introduction

Structural organization of genome in Prokaryotes and Eukaryotes; Organelle DNA-mitochondrial, chloroplast; DNA sequencing-principles and translation to large scale projects; Recognition of coding and non-coding sequences and gene annotation; 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.

Unit III

Proteomics

Protein analysis (includes measurement of concentration, amino-acid composition, N-terminal sequencing); 2-D electrophoresis of proteins; Microscale solution isoelectric focusing; 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

Pharmacogenetics

High throughput screening in genome for drug discovery- identification of gene targets, Pharmacogenetics and drug development

Unit V

Functional genomics and proteomics

Analysis of microarray data; Protein and peptide microarray-based technology; PCR-directed protein in situ arrays; Structural proteomics

Texts/References

1. Voet D, Voet JG & Pratt CW, *Fundamentals of Biochemistry*, 2nd Edition. Wiley 2006

2. **Brown TA**, *Genomes*, 3rd Edition. Garland Science 2006
3. **Campbell AM & Heyer LJ**, *Discovering Genomics, Proteomics and Bioinformatics*, 2nd Edition. Benjamin Cummings 2007
4. **Primrose S & Twyman R**, *Principles of Gene Manipulation and Genomics*, 7th Edition, Blackwell, 2006.
5. **Glick BR & Pasternak JJ**, *Molecular Biotechnology*, 3rd Edition, ASM Press, 1998.

Enzyme & Enzyme Technology - 3 Credits

UNIT I

Discovery, classifications and nomenclature of enzymes;

Techniques of enzyme isolation; Techniques of enzyme assay

UNIT II

Intracellular localization of enzymes; Techniques used in the purification of enzymes. Criteria of enzyme homogeneity Techniques used for determination of native and sub-unit molecular weight of enzymes; Isoenzymes; Multienzyme complexes and multifunctional enzymes

UNIT III

Physico-chemical characterization of enzymes;

Enzyme kinetics: Enzyme catalysis in solution - kinetics and thermodynamic analysis, effects of organic solvents on enzyme catalysis and structural consequences. Kinetics of enzyme inhibition;

Allosterism including half of the site activity phenomena

UNIT IV

Enzyme memory and mnemonic enzymes;

Structure and activity of the enzymes;

Mechanism of action of chymotrypsin, glyceraldehyde 3 Phosphate dehydrogenase, lysoenzyme, carboxy peptidase, ribonuclease, aldolase etc.

UNIT V

Various techniques used for the immobilization of enzymes, Applications of immobilized enzyme in Biotechnology; Ribozyme and catalytic antibodies- Functional proteins- structure and drug targets (enzymes and receptors)

Text/References:

1. **Martin F. Chaplin and Christopher Bucke; Enzyme Technology**, Cambridge, Univ Press
2. **Anil Kumar and Sarika Garg; Enzymes and Enzyme Technology**, Anshan Publishing; 1st edition

PRACTICAL [Total 08 Credits]

Lab on Bioinformatics:

1. Sequence Analysis Packages: EMBOSS, NCBI ToolKit, SMS
2. Database search engines: Entrez, SRS, DBGET
3. Pair wise alignment:
 - (i) Search tools against Databases:
 - (ii) BLAST
 - (iii) FASTA
4. Multiple sequence alignment:
 - a. Clustal
 - b. Dialign
 - c. Multalign
5. Protein motif and domain analysis:
 - a. MEME/MAST
 - b. eMotif
 - c. InterproScan
 - d. ProSite
 - e. ProDom
 - f. Pfam
6. Visualization Tool
 - a. RasMol
 - b. Cn3D
 - c. MolMol
 - d. PyMol
7. Phylogenetic analysis – MEGA, PAUP, PHYLIP
8. Protein and nucleic acid properties: Proteomics tools at the ExPASy server, GCG utilities and EMBOSS

Lab on Internet & Web Based Programming (CGI, Perl & HTML)

1. LINUX Operating System: Overview of Linux Architecture, Installation, Booting and Shutdown Process, System Processes(an overview), User Management- Types of users, Creating Users, Granting Rights, File System management
2. Uses of Scalar and Array Variables to manipulate DNA/RNA/Protein sequence data
3. Concatenation DNA fragments, transcribing DNA into RNA
4. Calculating the Reverse complement of a DNA strand
5. Uses of common Array Operators
6. Uses of Do-Until Loops
7. Uses of 'substr' function to look into the string
8. Reading a sequence data from a file and writing the results to a file
9. Opening and closing a Directory Handle, Reading a Directory and other directory manipulation functions.
10. Uses of Subroutines
11. Uses of Hashes for the genetic code: translating codons into amino acids
12. Uses of subroutine to read FASTA files
13. Translate a DNA sequence in all six reading frames
14. Uses of Regular Expressions
15. Extract annotation and sequence from GenBank file
16. Parsing GenBank annotation using arrays
17. Extract sequence chains from PDB file
18. Uses of CGI.pm Module and Passing Parameters via CGI, Debugging CGI programs
19. Installing Bioperl, Uses of Bioperl modules for sequence manipulation, accessing local database

Lab on Immunoinformatics

Browsing and searching sequence & structure databases to retrieve data of Immunoglobulin: sequence and structure

1. Study of Antibody sequence and structure
 - a) Antibody numbering: Kabat & Chothia methods o Equivalence between different numbering schemes & PDB numbering o Visualisation of structure & mapping CDRs on structure
 - b) Study of variable and constant domains o Drawing topology diagram
2. Study of Ag-Ab interactions
 - a) Searching for Ag-Ab complexes
 - b) Use of molsurfer, explorer etc
 - c) Characterisation of binding sites
 - ♣ Ab-protein complex
 - ♣ Ab-peptide complex
 - ♣ Geometric and molecular interaction
3. Sequence analysis in immunology
 - a) Alignments: antibody sequences
 - b) Alignments: antigen sequences, molecular evolution, variability analysis and phylogeny
4. Prediction of sequential epitopes
5. Prediction of conformational epitopes

6. Estimation of accuracy of predicted epitopes: use of experimental data
7. Prediction of cytotoxic and helper T cell epitopes
 - a) MHC Class I epitopes o MHC Class II epitopes
8. Web-based tools for vaccine design
9. Predicting immunogenicity
 - a) Combination of MHC and Proteasome predictions o Combination of MHC, TAP, and Proteasome predictions

M.Sc. Bioinformatics

SEMESTER III

Course code	Title	Credits
BT BI 601	Machine Learning Techniques & CADD(Bioinformatics II) (Core)	03
BT BI 611	Structural Biology and Bioinformatics (Bioinformatics III) (Core)	03
BT BI 621	Database management System (Generic Elective)	03
BT BI 631	Java Programming (Core)	03
BT BI 641	Metabolic Engineering & System Biology (Discipline Centric Elective)	1.5
BT BI 651	Pharmacogenomics (Discipline Centric Elective)	1.5
BT BI 661	Microscopic Techniques For Image Processing (Discipline Centric Elective)	1.5
BT BI 671	Assignments/ Practical	08
BT BI 681	Seminar(Soft Skill Development)	02
	Comprehensive Viva Voce	04
	Total Credits	29

Any 02 out of 03 Discipline Centric electives can be opted by the students from SBT. One Generic Elective of 03 credits can be opted by the students from SBT itself or from any other department (UTD).

Machine Learning Techniques & CADD (Bioinformatics – II): 3 Credits

UNIT I

Markov Chains & Hidden Markov Models: Introduction to Markov chains and HMM using Markov chains for discrimination of biological sequences. Forward and backward algorithms; Parameters estimation for HMMs. HMMs for pairwise and multiple sequence alignments. Profile HMMs.

UNIT II

Machine Learning and Bioinformatics: Introduction to various Machine Learning techniques and their applications in Bioinformatics. Support Vector Machine, Artificial Neural Network; Neural Networks and their practical applications towards the development of new models, methods and tools for Bioinformatics.

UNIT III

Machine Learning Algorithms

- a) Dynamic Programming
- b) Gradient Descent
- c) EM/GEM Algorithms
- d) Markov-Chain Monte-Carlo Methods
- e) Simulated Annealing
- f) Evolutionary and Genetic Algorithms

UNIT IV

CADD and Molecular Docking: Introduction, Basic Procedure; Constant and Flexible Docking.

Drug design: Drug discovery process; Target identification and validation; lead optimization and validation; Methods and Tools in Computer-aided molecular Design,

Analog Based drug design:-Pharmacophores (3D database searching, conformation searches, deriving and using 3D Pharmacophore, constrained systematic search, Genetic Algorithm, clique detection techniques, maximum likelihood method).

Structure based drug design:- Docking, De Novo Drug Design (Fragment Placements, Connection Methods, Sequential Grow), Virtual screening.

UNIT V

Structure Activity Relationship: Introduction to QSAR, QSPR, Various Descriptors used in QSARs: Electronics; Topology; Quantum Chemical based Descriptors. Regression Analysis, The Significance and Validity of QSAR Regression Equations, Partial Least Squares (PLS) Analysis, Multi Linear Regression Analysis. Use of Genetic Algorithms, Neural Networks and Principle Components Analysis in the QSAR equations.

Text References:

1. Chemoinformatics: A Textbook by Johann Gasteiger.
2. Bioinformatics second edition by David M Mount
3. Bioinformatics: Methods and Applications Genomics, Proteomics And Drug Discovery

4. Bioinformatics: the Machine Learning Approach by Pierre Baldi and Soren Brunak; Second Edition; The MIT Press

Structural Biology and Bioinformatics (Bioinformatics-III): 3 Credits

UNIT I

Prediction of protein structure: Secondary structure: Algorithms of Chou Fasman, GOR methods. Tertiary Structure: basic principles and protocols, Methods to study 3D structure.

Protein structure comparison and classification: classes, folds; the concepts in 3D structure comparison, purpose of structure comparison, algorithms such as FSSP, VAST and DALI.

Protein Folding: Principles of protein folding and methods to study protein folding.

UNIT II

Molecular Mechanics: Introduction, The Morse Potential, The Harmonic Oscillator Model for Molecules, Comparison of Morse and Harmonic Potential, Two atoms connected by a bond, Poly atomic Molecules, Energy due to Stretch, Bend, Stretch-Bend, Torsional strain, van der Waals and Dipole-Dipole interactions.

Types of Potentials: Lennard-Jones, Truncated Lennard-jones, Exponential-6, Ionic and Polar potentials. Types of Force Fields: AMBER, CHARMM, Merck Molecular Force Field, Consistent Force Field, MM2, MM3 and MM4 force fields.

UNIT III

Potential Energy Surface: Convergence Criteria, Characterizing Stationary Points, Search for Transition States. Optimization: multivariable Optimization Algorithms, level Sets, Level Curves, Gradients, Optimization Criteria, Unidirectional Search, Finding Minimum Point, Gradient based Methods-Steepest Descent and Conjugate Gradient Methods.

UNIT IV

Molecular Dynamics Simulation: Introduction, Radial distribution functions, Pair Correlation function, Newtonian dynamics, Integrators- Leapfrog and Verlet algorithm, Potential truncation and shifted-force potentials, Implicit and explicit Solvation models, Periodic boundary conditions, Temperature and pressure control in molecular dynamics simulations.

UNIT V

Computational RNA Structure analysis: Secondary and tertiary structure of RNA. Various algorithms of RNA folding and their analysis. Energy minimization in RNA folding. RNA sequence alignment based on secondary structure and its applications in functional genomics and phylogeny.

Text References:

1. Computational Chemistry and Molecular Modeling-Principles and Applications by Ramachandran, Deepa and Namboori., 2008, Springer_Verlag.
2. Molecular Modeling Principles and Applications (2nd Ed.) by Andrew R. Leach., Prentice Hall, USA. 2001
3. Molecular Modelling for Beginners, (2nd Edition) by Alan Hinchliffe., John Wiley & Sons Ltd.2008
4. Molecular Modeling and Simulation – An Interdisciplinary Guide by Tamar Schlick., Springer-verlag 2000
5. Computational Medicinal Chemistry for Drug Discovery, edited by Patrick Bultinck., Marcel Dekker Inc. 2004

Database Management System: 3 Credits

UNIT I

Introduction, Database System Versus File Systems, Characteristics of Database, Database Concepts, Schemas & Instances, DBMS architecture and Data Independence, Data Models, Database Languages & Interfaces, View of Data, Database users and Administrators, Database System Structure, Database System Applications data models- Entity Relationship Model, Relational Model. ER Model: Keys, Constraints, Design Issues, Extended ER features, Reductions of ER Schema to Tables.

UNIT II

Structured Query Language – Basic Structure, Set Operations, Aggregate Functions, Null Values, Nested Sub queries, Views, Integrity: Domain constraints, Joined Relations, Data-Definition Language

Relational Databases: Integrity and Security, Relational – Database designs; Pitfalls in Relational Design Database, Functional dependencies, Decomposition Normal Forms – 1NF, 2NF, 3NF & Boyce-Codd NF, Data Storage – Ordered indices

UNIT III

Data Storage and Querying: Storage and File Structure, Indexing and Hashing, Query Processing, Query Optimization

UNIT IV

Transaction Management: Transactions, Concurrency Control, Recovery System; Properties of transactions: Concurrency problems, Serialisability and Locking techniques, Granularity of Data Items – Database System Architecture and Information retrieval: Centralized and Client-Server Architecture

UNIT V

Database System Architectures: Distributed Databases, Parallel Databases, Data Warehousing and Data Mining

Text References:

1. Database System Concepts (4th Ed.) by Silberschatz, A., Korth, H.F. and Sudarshan, S., 2002, McGraw Hill Publishers.
2. An Introduction to Database Systems (7th Ed.) by Date, C.J., Addison Wesley Publishers. 2000
3. Fundamentals of Database Systems (4th Ed.) by Elmasri and Navathe, Addison Wesley Publishers. 2004
4. Principles of Database Systems (2nd Ed.) by Ullman, J. D., Galgotia Publications. 2001

Java Programming: 3 Credits

UNIT I

C++ Vs JAVA, JAVA and Internet and WWW, JAVA support systems, JAVA environment., JAVA program structure, Tokens, Statements, JAVA virtual machine, Constant & Variables, Data Types, Declaration of Variables, Scope of Variables, Symbolic Constants, Type Casting.

Operators: Arithmetic, Relational, Logical Assignments, Increment and Decrement, Conditional, Bitwise, Special, Expressions & its evaluation. If statement, if...else... statement, Nesting of if...else... statements,,else...if Ladder, Switch, ? Operators, Loops – While, Do, For, Jumps in Loops, Labelled Loops.

UNIT II

Defining a Class, Adding Variables and Methods, Creating Objects, Accessing Class Members, Constructors, methods Overloading, Static Members, Nesting of Methods. Inheritance: Extending a Class, Overriding Methods, Final Variables and Methods, Final Classes, Finalize Methods, Abstract methods and Classes, Visibility Control.

UNIT III

Arrays: One Dimensional & two Dimensional, strings, Vectors, wrapper Classes, Defining Interface Extending Interface, Implementing Interface, Accessing Interface Variable, System Packages, Using System Package, Adding a Class to a Packages, Hiding Classes.

UNIT IV

Creating Threads, Extending the Threads Class, Stopping and Blocking a Thread, Life Cycle of a Thread, Using Thread Methods, Thread Exceptions, Thread Priority, Synchronization, Implementing the Runnable Interface.

UNIT V

Local and Remote Applets Vs Applications, Writing Applets, Applets Life Cycle, Creating an Executable Applet, Designing a Web Page, Applet Tag, Adding Applet to HTML File, Running the Applet, Passing Parameters to Applets, Aligning the Display, HTML Tags & Applets, Getting Input from the User.

BioJava - Installing BioJava, Symbols, Basic Sequence Manipulation (DNA to RNA, Reverse Complement, motif as regular expression), Translation (DNA to Protein, Codon to amino acid, Six frame translation), Proteomics (Calculate the mass and pI of a peptide), Sequence I/O (File Formats conversions), Locations and Features (PointLocation, RangeLocation, Feature modifications), BLAST and FASTA (Blast and FastA Parser, extract information from parsed results)

Text/references

1. E. Balaguruswamy, "Programming in Java", 2nd Edition, TMH Publications Peter Norton, "Peter Norton
2. Herbert Schildt, Java- A Beginners Guide (4th Ed.), Tata Mc-Graw-Hill publication. 2007
3. Computing Concepts with Java 2 Essentials (2nd Ed.) by Horstmann, C.S., John Wiley Publishers.2000
4. Object Oriented Design and Applications (2nd Ed.) by Benjamin, Cummings and Booch, G., Addison Wesley Publishers. 1994

Metabolic Engineering & System Biology – 1.5

Credits

UNIT I

Major Metabolic Pathways: Gluconeogenesis, Pentose phosphate pathway, Glycogen synthesis and degradation, Fatty acid oxidation and synthesis, Amino acid catabolism, Purine and pyrimidine nucleotide synthesis etc.

Metabolic Pathways databases: KEGG, EcoCyc and MetaCyc; EMP, Malaria Parasite Metabolic Pathways, Boehringer Mannheim -Biochemical Pathways

UNIT-II

Enzymes, Compounds and Reactions databases

- LIGAND - Biochemical Compounds and Reactions
- ENZYME - Enzymes
- BRENDA - Comprehensive Enzyme Information System

Metabolic Engineering- Basics

Rational Metabolic engineering design

UNIT III

Systems Biology Networks - basics of computer networks, Biological uses and Integration; Self-organizing maps and Connectivity maps - definition and its uses.

Networks and Pathways – Types and methods; Metabolic networks.

Unit IV

Simulation of pathways: Whole cell: Principle and levels of simulation – E-cell and v-cell, Virtual Erythrocytes; Pathological analysis. Flux Balance Analysis. Biochemical metabolic pathways, Metabolomics and enzymes. Interconnection of pathways, metabolic regulation. Networks and Motifs: Gene Networks: basic concepts, computational models. Lambda receptor and lac operon as an example.

Text/references

KlippEdda, Liebermeister Wolfram, WierlingChristoph, Kowald Axel, Lehrach Hans, Herwig Ralf. Systems Biology: A Textbook. Publisher: Germany, Wiley-VCH. 2009. ISBN: 9783527318742.

Alterovitz Gil, Ramoni Marco F. Systems Bioinformatics: An Engineering Case-Based Approach. Publisher: Boston, Artech House Publishers. 2007. ISBN: 9781596931244

Pharmacogenomics: 1.5 Credits

1. Pharmacogenomics, benefits, practical applications, the promise of Pharmacogenomics today leading to personalized medicines, human genetic variation-example of CYP gene variation leading to variable metabolism of drugs, distribution of variation, mutation and its kinds, natural selection, variation in ethnic groups races.

2. Pharmacology, clinical pharmacology, drugs, drugs legislation and safety, types of drugs-example of latest drugs, drug potency and efficacy and toxicity, ADME of drug-drug absorption, drug distribution, drug metabolism and drug excretion, drug therapeutic levels, therapeutic index, drug abuse, drug response in patients by correlating gene expression, regulation of gene expression, polymorphism, alleles, single nucleotide polymorphism, genotyping.

3. Genetic biomarkers- biomarkers on drug development, biomarkers in clinical development, biomarkers for molecular diagnostics-example of cancer biomarkers, pharmacogenetics and drug development.

Texts/References

1. Wu R and Lin M, *Statistical & Computational Pharmacogenomics*, CRC Press, 2008
2. Yan Q, *Pharmacogenomics in Drug Discovery and Development*, Springer-Verlag New York, LLC, 2008
3. Meyer UA and Tyndale RF, *Pharmacogenomics*, 2nd Edition, CRC Press, 2005.
4. Innocenti F, *Pharmacogenomics: Methods and Applications*, Springer-verlag New York, LLC, 2005
5. Rothstein MA and Collins FS, *Pharmacogenomics: Social, Ethical and Clinical Dimensions*, Wiley John & Sons, Inc., 2003

Microscopic Techniques For Image Processing - 1.5 Credits

Unit I

Transmission electron microscopy: Wave nature of electrons – Electromagnetic lenses – Basic components of Transmission Electron Microscope – Alignment of TEM – Major operational modes of TEM.

Unit II

Scanning electron microscopy: Basic systems of the SEM – Contrast and three-dimensionality of the SEM image – Stereo imaging with the SEM

Unit III

Specimen preparation for EM: TEM : Specimen preparation for TEM – Fixation – Washing – Dehydration – Embedding – Specimen staining for TEM – Positive staining and negative staining – Metal shadowing techniques – CryoEM. Ultramicrotomy: Shaping the specimen block – Types of ultramicrotome knives – EM grids – Support films for grids – Ultramicrotome and section processing.

SEM: Surface cleaning – Rinsing and dehydration – Specimen drying techniques – Specimen fracture procedures – Replication procedures – Specimen mounting – Specimen coating for conductivity.

Unit IV

Image processing and image analysis by computer: Capturing the image – Conventional vs. digital – Image processing – Controlling contrast, brightness and gamma – Removing noise – Fast Fourier Transform – images for publication and presentation – Three dimensional imaging.

Unit V

Atomic Force microscopy and Confocal Microscopy: Atomic force microscopy (AFM) including contact-mode, tapping-mode and lateral-force AFM Confocal Microscopy: Basics of Confocal Microscopy, Sample Preparation, Confocal Optics, And Resolution.

Text/Reference:

1. Electron Microscopy: Principles and techniques for biologists by John J Bozzola, and Lonnie Dee Russell., Jones & Bartlett Learning. 1999
2. Principles and Techniques of Electron Microscopy: Biological Applications by M.A.Hayat., Cambridge University Press. 2000
3. Handbook of Biological Confocal Microscopy, by Pawley, J.B., Springer-verlag. 2006

PRACTICAL [Total 08 Credits]

Lab on Java Programming:

1. Working with Objects, Arrays, Conditionals and Loops.
2. Creating Classes and Applications in Java.
3. Java Exception handling
4. Streams and I/O, Using Native Methods and Libraries
5. Simple Animation and Threads, Advanced Animation, Images and Sound.
6. Managing Simple Events and Interactivity.
7. Local and global alignment of sequences
8. Creating User Interfaces with AWT, Modifiers.
9. Multithreading example
10. Java Programming Tools, Working with Data Structures.

Lab on Database management System:

Data Definition Language (DDL) statements:

1. Creating database, Selecting database, Deleting database, Creating table, Modifying Table, Deleting table
2. Data Manipulation statements: Inserting, updating and deleting records
3. Retrieving Records: Retrieving specific rows and columns
4. Use of MySQL operators – Arithmetic operators, Comparison Operators, Logical operators
Math functions, Aggregate functions
5. String operations
6. Limiting, Sorting and grouping query results
7. Handling null values
8. Renaming or aliasing table and column names
9. Using subqueries

10. Using Joins – joining a table to itself, joining multiple tables
11. Use of Indexes
12. Security Management
13. Granting and Revoking rights on tables

Lab on Bioinformatics II/III:

1. Structure Prediction: Advanced Modelling,
2. Modeller Interface: Chimera
3. Structural Validation: Procheck, WHATIF, VERIFY 3D
4. Energy Minimization
5. Tutorial on Molecular Dynamics: Gromacs
6. Binding Site Identification: Pocket- Finder
7. Structure based Drug Design:- Molecular Docking
8. Ligand Designing: Chemaxon
9. Ligand based Drug Design:- QSAR
10. Protein –protein Interaction and Visualization: Osprey, ViSant, Cytoscape, and STRING
11. RNA Structure Prediction

M.Sc. Bioinformatics

SEMESTER IV

Course code	Title	Credits
BT BI 602	Project Work	12
BT BI 612	Comprehensive Viva Voce	04
	Total Credits	16

Project Work (Credit: 25)

The course is required satisfactory completion and defense of the Masters dissertation.

This process includes:

- a) Conceptualization of the independent research
- b) Collection, analysis, and interpretation of data,
- c) Thesis writing
- d) Oral presentation of findings
- e) Viva-Voce.

NOTE: Dissertation activity must be completed within prescribed time frame for the semester.