

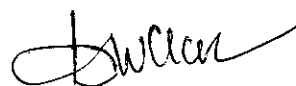
DEPARTMENT OF MICROBIOLOGY AND BIOINFORMATICS

**POST GRADUATE DIPLOMA IN COMPUTATIONAL BIOLOGY
(PGDCB)**

SYLLABUS
(w.e.f admitted batch 2015 -16)

BILASPUR UNIVERSITY

**Old High Court Bhavan, Near Gandhi Chowk
Bilaspur (C.G) - 495001**



PG Diploma in Computational Biology

Course Code	Semester I	Credits	Marks 100 (External/ Internal)
PGDCB 101	Basics in Biology	03	75/25
PGDCB 102	Programming in C and C++	03	75/25
PGDCB 103	BioPERL and BioJava	03	75/25
PGDCB 104	Bioinformatics	03	75/25
PGDCB 105	Minor Project/ INFORMATICS LAB	08	100
Semester II			
PGDCB 201	Machine learning	03	75/25
PGDCB 202	Genomics and Proteomics	03	75/25
			75/25
PGDCB 203	Molecular modeling and Drug Designing	03	75/25
PGDCB 204	Systems Biology	03	75/25
PGDCB 205	Dissertation/ Viva-Voce	08	100
Total		40	1000

[Handwritten Signature]

PGDCB 101

BASICS IN BIOLOGY

Unit I

Structure and properties of purine and pyrimidine, two-dimensional structure of DNA, nucleosides, nucleotides, chemical synthesis of oligonucleotides, physicochemical properties of nucleic acids, Models and strategies of DNA replication, DNA damage and repair.

Unit II

Principles of transcription- prokaryotic and eukaryotic, General features of genetic code, structural components of prokaryotic and eukaryotic ribosomes, protein synthesis in prokaryotes and eukaryotes, inhibitors of protein synthesis, post translational modifications, protein targeting.

Unit III

Principles of bioenergetics- free energy, enthalpy and entropy. Free energy changes in biological transformations in living systems. Redox potential. Phosphate group transfer potential and ATP, high energy compounds, oxidation and reduction reactions, Mitochondrial electron transport system - organization of components and importance. Substrate level phosphorylation, oxidative phosphorylation.

Unit IV

Carbohydrates and lipids - Classification, chemical properties, structural and functional roles. Glycolysis, TCA cycle, β -oxidation of fatty acids. Biosignaling- Molecular mechanism of signal transduction, secondary messenger - cAMP, cGMP, IP₃, diacylglycerol and nitric oxide.

Unit V

Enzymes - Classification, Assay, Nature and Mapping of active site, Factors affecting enzyme activity, Kinetics of single substrate enzyme catalyzed reactions- Michaelis and Menten equation, Line weaver - Burk, Edaie -Hofstee and Hanes plots, V_{max}, K_m, K_{cat}, specificity constant (K_{cat}/K_m) and their significance, Kinetics of multisubstrate reactions - classification with examples.

Recommended Text Books

1. Palmer, T, Understanding Enzymes, First Edition, Harvard publishing.
2. Stryer, L., Biochemistry. Fourth Edition, Freeman W.H. and CO.
3. West and Todd, Text book of Biochemistry, Fourth Edition. Oxford and IBH.
4. Nelson Cox, Principles of Biochemistry, Fourth Edition. Freeman.
5. Voet and Voet, Biochemistry, Third Edition, John Wiley and Sons.
6. Watson, J.D., Molecular Biology of the gene, Fifth Edition, Pearson Education International.
7. Lodish, H., Scott, M.P., Matsudaira, P., Darnell, J., Zipursky, L., Kaiser, C.A., Berk, A. and Krieger, M., Molecular Cell Biology, Fifth Edition, W.H. Freeman and Co.
8. Lewin, B. Genes VIII, Eighth Edition, Pearson Education International.



PGDCB 102

PROGRAMMING IN C AND C++

Unit I

Introduction to computers: Anatomy of digital computer; Memory units: RAM, ROM, types of RAM, introduction to auxiliary memory; Definition and functions of operating systems. Algorithms and flowcharts.

Unit II

Basic of C: Introduction to C, Features of C, Basic I/O functions, Control structures, Functions-Passing values between functions. Calling function by value and by reference, Recursion. Arrays, Pointers

Unit III

Structures and Unions, File handling.

Unit IV

Introduction to C++: Stream Bases I/O, Class Members, Member Functions, Static Data Members, Static Member Functions, Friend Function. Base Class, Derived Class,

Unit V

Inheritance - Single, Multiple, Multilevel, Hierarchical, Hybrid. Virtual Functions, Virtual Base Classes, Pure Virtual Functions, Virtual Destructors. Generic Classes, Function Template, Class Template.

Recommended Text Books

1. Alexis Leon and Mathews Leon, Fundamentals of Information technology, First Edition, Vikas Publishing House Pvt. Ltd.
2. Norton, P., Introduction to Computers, Sixth Edition, Tata McGraw Hill.
3. Yashwant K., Let us C, Fourth Edition, BPB publications.
4. Balagurusamy, E., Object Oriented Programming with C++, Third Edition, Tata McGraw-Hill.



PGDCB 103

BioPERL and BioJava

UNIT - I

Unix/Linux- Basics, Filesystems, Essential Commands and Text Editors.

UNIT - II

Perl - Introduction, Data Types -Lists/Arrays and Hash, Scalar Functions, Quoting Basics, Functions/Subroutines, Operators and Control Structures, Data Types: Boolean, Inputs, Pattern, Matching and Regular Expressions, File Handling and File Manipulation, Error Handling, Variable Scope, Mathematical Functions, Special Variables, References, Aggregate data structures, Using Modules.

UNIT - III

BioPerl Modules- Sequences and Strings, Motifs and Loops, flow control, String operators and Writing files, Subroutines-Scoping, Arguments, Command line arguments, Passing data to subroutines, Modules and Libraries, Debugging, CPAN – Comprehensive perl archive network) Data Structures and Algorithms for Biology Restriction Maps and Restriction Enzyme Data, Working and Analyzing with GenBank and BLAST data Subroutines, Text and String Processing.

UNIT - IV

Introduction-java, applications, hello world two ways, Variables and Operators, Control Structures, Objects, Arrays. Applet Class, basic HTML, the applet HTML tag, TCP/IP, UDP, servers, clients.

UNIT - V

Setup, Alphabets and Symbols, Basic Sequence Manipulation, Translation, Proteomics, Sequence I/O, Annotations, Locations and Features, BLAST and FASTA, Counts and Distributions, Weight Matrices and Dynamic Programming, User Interfaces, Genetic Algorithms.

Recommended Text books:

- Beginning Perl for Bioinformatics, Tisdall, J.D., First Edition O'Reilly Publishers.
- Programming Perl, Wall, W., Christiansen, T. and Orwant, J., Third Edition, O'Reilly Publishers.
- Java-2: The complete Reference, Naughton, P. and Schildt, H., Third Edition, McGraw Hill Publishers.



BIOINFORMATICS

Unit I

Types of databases. Data abstraction. Data models (ER, Hierarchical and Network). SQL: Data definition, data manipulation and control statements. Database objects, tables, views, sequences and synonyms. Application development using Visual Basic as front end: variables, procedures and controlling program executors, forms, standard controls.

Unit II

Concepts regarding nature and type of data available in biological databases: Biological literature databases, biodiversity information databases, microbiological databases – (eukaryotic, bacterial and viral), cell type databases, toxicology databases, spectroscopic databases, protein and nucleic acid sequence databases, genome databases.

Unit III

Dynamic programming – global alignment and local alignment. Substitution Matrices – PAM and BLOSUM, Tools for searching databases - BLAST & FASTA, different types of BLAST & FASTA, Significance of BLAST & FASTA.

Unit IV

Multiple sequence alignment, Phylogenetic trees and Sequence graphs, Phylogenetic tree construction methods - Distance based methods - Clustering based methods (UPGMA and Neighbor Joining), optimal criterion method (FM and Minimum Evolution methods) and character based methods. Traditional parsimony methods - Exhaustive, Branch and Bound, and heuristic. Maximum Likelihood method, Tree evaluation methods - Bootstrap.

Unit V

Probabilistic models of evolution – Jukes cantor model and Kimura model. Combined multiple sequence alignment and phylogeny – Sankoff and Cedergren method. Prediction of RNA secondary structure: Nussinov folding algorithm, energy minimization and Zuker folding algorithm, covariance models.

Recommended Text Books

1. Silberschatz, A., Korth, H.F. and Sudarshan, S., Database system Concepts, Fourth Edition, McGraw Hill Publishers.
2. Date, C.J., An introduction to Database systems, Seventh Edition, Addison Wesley Publishers.
3. Baxevanis, A.D. and Francis Ouellette, B.F., Bioinformatics – a practical guide to the analysis of Genes and Proteins, First Edition, John Wiley & Sons.
4. Attwood, T.K. and Parry, D.J., Introduction to Bioinformatics, Second Edition, Addison Wesley Longam Limited.
5. Mount, D., Bioinformatics: Sequence and Genome Analysis, First Edition, Cold Spring Harbor Laboratory Press.
6. Lesk, A.M., Introduction to Bioinformatics, First Edition, Oxford University Press.



PGDCB 111

Mini Project/ INFORMATICS LAB

1. Practicals on Microsoft Word, Microsoft Excel (Spreadsheet Application).
2. Downloading and installing software (MS-Office, Acrobat Reader, etc).
3. Internet surfing and searching information,
4. Practical sessions on Microsoft Access (Database related applications).
5. Designing of computer presentation with MS Power Point.
6. Program to find the factorial of a given number using Recursion in C.
7. Program to count number of bases in given DNA sequence in C++
8. Use PERL/ BioPERL to convert DNA sequence information to amino acid sequence.
9. Transcription program using BioJava
10. Translation program using BioJava

A handwritten signature in black ink, appearing to read 'K. S. Srinivasan', is located to the right of the list of tasks.

PGDCB 201

MACHINE LEARNING

UNIT – I

Definition, goals and applications of machine learning. Training data, concept representation, function approximation. Representing concepts as decision trees. Recursive induction of decision trees. Searching for simple trees and computational complexity. Active learning with ensembles

UNIT – II

Experimental Evaluation of Learning Algorithms

Measuring the accuracy of learned hypotheses. Comparing learning algorithms: cross-validation, learning curves, and statistical hypothesis testing. Artificial Neural Networks.

UNIT – III

Bayesian Learning and Instance-Based Learning

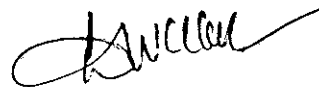
Probability theory and Bayes rule. Naive Bayes learning algorithm. Parameter smoothing. Generative vs. discriminative training. Logistic regression. Bayes nets and Markov nets for representing dependencies. Constructing explicit generalizations versus comparing to past specific examples. k-Nearest-neighbor algorithm. Case-based learning.

UNIT – IV

Graph Algorithms: Graphs and digraphs. Shortest path problems. Complexity. Bellman's optimality principle. Dijkstra's algorithm. Shortest spanning trees. Kruskal's greedy algorithm. Prim's algorithm for shortest spanning trees. Networks. Flow augmenting paths. Ford-fulkerson algorithm for maximum flow. assignment problems. bipartite matching.

UNIT – V

Text Processing: Pattern matching algorithms- Brute force, the Boyer-Moore algorithm, the KMP algorithm, tries- standard tries, compressed tries, suffix tries, text compression- Huffman coding algorithm.



PGDCB 202

GENOMICS AND PROTEOMICS

Unit I

Genome organization, DNA sequencing – Overlap, layout and consensus. Combinatorial approaches to genome sequencing – shot gun method, Craig Venter method, hierarchal method, High throughput sequencing method. Fragment and map assembly, Research tools for genome assembly - Phred, Phrap, Consed.

Unit II

Structural Genomics: Gene and genome annotation Functional Genomics: EST, STS and RNA expression. Predicting function of gene, gene ontology, DNA microarrays, Comparative genomics: homology, paralogy, synteny – gene order, genome alignment, genome rearrangements. Tools in comparative genomics (MuMmer, BLAST2).

Unit III

Basic methods for studying proteomics – Proteome, Isolation of protein, Purification of proteins - Principles of chromatography - Gel chromatography, ion exchange chromatography, affinity chromatography, Protein digestion techniques, Principles of electrophoresis - Capillary electrophoresis, SDS PAGE,

Unit IV

Proteome techniques: X-ray crystallography technique, Nuclear Magnetic Resonance (NMR), Mass Spectroscopy (MS), 2D gel and Microarray, Acquisition and analysis of data, Tools for correlation of protein expression data to biological processes.

Unit V

Structure prediction Tools: Wet bench and *in silico* tools for protein-protein interactions. Predicting function of the protein - Motifs, Domains, Profiles; Protein sequence-structure-function relationship; Comparative proteomics. Applications of proteomics.

Recommended Text Books

1. Andrzej Polanski and Marek Kimmel, Bioinformatics, First Edition, Springer Publications.
2. Pevsner, J., Bioinformatics and Functional Genomics, John Wiley and Sons.
3. Primrose, S.B. and Twyman, R.M., Principles of genome analysis and. Genomics, Third Edition, Blackwell Publishing Company.
4. I. Howe, D.H.Williams, and R.D. Bowen, Mass spectrometry, principles and applications. 2nd edition, McGraw Hill, London, 1981.
5. Daniel .C.Liebler Introduction to Proteomics – Tools for the new biology. Humana press 2002.



PGDCB 203

MOLECULAR MODELING AND DRUG DESIGNING

Unit I

Concepts of computational chemistry. Molecular mechanics: general features of force fields, Quantum mechanics: Hamiltonian operators.

Unit II

Energy minimization: statement of the problem, line minimization and energy minimization, Simulation methods: Time averages, ensemble averages, Molecular dynamics methods, Monte Carlo methods, Differences between MD and MC, Simulated Annealing procedure.

Unit III

Comparative modeling of proteins: Homology modeling – Superimposition, Template selection, Backbone modeling, loop building (Search/generation), Side chain generation, and Model evaluation and validation. Fold recognition methods - Threading-Bowie-Luthy-Eisenberg, Profile-Profile alignment, combined modeling approaches. *Ab initio* method - Anfinsen Thermodynamic Hypothesis.

Unit IV

First and Second generation of drug designing: QSAR, Fundamentals of QSAR-objectives, expressions of biological activity, QSAR parameters related to chemical structure, Analysis of results - correlative methods, and Regression analysis., MFA, 3-D QSAR, QSAR descriptors, Finding new drug targets to treat diseases.

Unit V

Computer aided drug design, Pharmacopore properties, Ligand and structure based drug design, Receptors - Structure, function, and pharmacology, Virtual screening, Limitations of CADD. Molecular Docking – Principles and methods for Docking, Docking problem, 3-D database search approaches, Rigid body and flexible docking; Hydrophathy, Denovo ligand design.

Recommended Text Books

1. Andrew R. L., Molecular modeling: principles and applications, Second Edition, Prentice Hall.
2. Bourne, P.E. and Weisig, H., Structural bioinformatics, Second Edition.
3. Creighton, T.E., Proteins: structure and molecular properties, Second Edition, Freeman W.H. and Company.
4. Penelope W Coddig. Structure based drug design,



PGDCB 204
SYSTEMS BIOLOGY

UNIT - I

Introduction, basics, applications of systems biology. Systems Microbiology: System Structure Identification, System Behavior Analysis, System Control, System Design, Measurement for Systems Biology

The cells as a well stirred bioreactor, Michaelis Menten kinetics, a genetic switch in Lambda phage, synthetic genetic switches, stability analysis, Modeling *E.coli* chemotaxis, Genetic oscillators.

UNIT - II

Cell systems biology: computational cell biology, computer simulation of the cell, modeling approaches: differential equation, dynamic systems theory, dealing with uncertainty, computational modeling.

Importance of diffusion and gradients for cellular regulation, local excitation, global inhibition model, Rapid Pole-to-pole oscillations in *E. coli*, Models for Eukaryotic gradient sensing.

UNIT - III

Developmental Systems Biology: System structure identification, Network Structure Identification: Bottom-Up Approach, Top-Down Approach, Parameter Identification

Building and organism starting from a single cell, Quorum sensing, *Drosophila* development, Whole cell simulation, Virtual erythrocytes, pathological analysis, fomentation analysis, flux balance analysis, minimal gene complement.

UNIT - IV

Metabolic systems biology: system behavior analysis, Simulation, Analysis Methods: bifurcation analysis, metabolic control analysis, and sensitivity analysis. Robustness of biological systems, Complex Engineering Systems

UNIT - V

Classification of metabolic pathways, metabolic pathway databases – KEGG, EMP, Malarial parasite metabolic pathways, EcoCyc and MetaCyc, Boehringer-Mannheim biochemical pathways.

UNIT – V

References books:

- Noise-based switches and amplifiers for gene expression, Jeff Hasty, Joel Pradines, Milos Dolnik, and Collins, J. J. (2000). PNAS, USA 97: 2075-2080.
- Molecular Cell Biology of the Cell, Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K and Walter, P, Fourth Edition, Taylor and Francis Group.



PGDCB 211

PROJECT WORK (Students have to carry out a project work in the II semester)/

Viva-voce

Students will have a Viva-voce in the subjects concerned to I and II semesters.

A handwritten signature in black ink, appearing to be 'R. Mani', with a long horizontal stroke extending to the right.