B.Sc. (HONOURS) BIOINFORMATICS

REGULATIONS AND SYLLABI

(Effective from 2025-2026)

Department of Bioinformatics

SCHOOL OF LIFE SCIENCES

PONDICHERRY UNIVERSITY

PUDUCHERRY

Pondicherry University
School of Life Sciences
Department of Bioinformatics

B.Sc. (Honours) in Bioinformatics

PROGRAM OBJECTIVES

The main objective of the program is to train the students to learn an innovative and evolving field of bioinformatics with a multi-disciplinary approach. Hands-on sessions will be provided to train the students in both computational and experimental labs.

PROGRAM OUTCOMES

On completion of this program, students will be able to:

- Gain understanding of the principles and concepts of both biology along with computer science
- To use and describe bioinformatics data, information resource and also to use the software effectively from large databases
- To know how bioinformatics methods can be used to relate sequence to structure and function
- To develop problem-solving skills, new algorithms and analysis methods are learned to address a range of biological questions.

Eligibility for B.Sc. (Honours) Bioinformatics

- Successful completion of higher secondary (+2/ PUC) certificate or Equivalent stage of education in Sciences with a minimum of 50% aggregate marks

SEMESTER - I

Biological Sciences

Course Code : UBIN101

Major Course

Credits : 4 **Hours : 60**

Course Objective:

Enable Students to understand the diversity and uniqueness of the plant and animal kingdom. This module also intends to provide basic knowledge on the structural organisation of plants and animals, genetics, evolution, and the developmental biology and physiology of plants and animals.

Unit 1 **12 Hours**

Classification of the living organisms: Classification of organisms- Linnaeus and Whittaker's Five Kingdom classification. Systematic and binomial System of nomenclature. Salient features of plants (major groups; Angiosperms up to class level). Salient features of animals (Major phyla, Vertebrates up to class level).

Unit 2 **12 Hours**

Structural organisation and development of plants: Morphology and anatomy of flowering plants, with emphasis on differences between monocots and dicots. Phytohormones and their functions; Photoperiodism, vernalisation; Basic concepts of development-gametogenesis, fertilization, morphogenesis and organogenesis.

Unit 3 **12 Hours**

Structural organisation and development of animals: Coelom and metamerism, protostomes and deuterostomes, Metamorphosis in insects and amphibians and its hormonal regulation, Retrogressive metamorphosis, Spermatogenesis, oogenesis, fertilisation, blastulation, gastrulation, Menstrual and oestrous cycle, placental hormones, parturition, and lactation.

Unit 4 **12 Hours**

Comparative physiology of mammals: homeostasis, thermoregulation, blood, acid-base balance, transport function of haemoglobin, digestion and absorption, osmoregulation, conduction of nerve impulse, neurotransmission, contraction of skeletal muscles, reproduction in mammals, Endocrine system in humans.

Unit 5 **12 Hours**

Genetics and evolution: Mendel's laws of inheritance, linkage, recombination, multiple alleles, genetics of blood group, pedigree analysis, DNA fingerprinting. Modern concept of the gene, genetic code, sex chromosomes, sex determination in *Drosophila* and human. Theories of evolution, Hardy-Weinberg Law, natural selection, genetic drift, isolation and speciation, Evolution of man.

Course Outcome:

Understand the basics of the unique features of plants and animals with their cellular and system-level understanding. Comprehend the diversity of plant and animal kingdoms with an awareness of their development, physiology, genetics and evolution.

Text Books:

1. Life: The Science of Biology, 10th Edition by David E. Sadava, David M. Hillis, H. Craig Heller and May Berenbaum-2012.
2. The Science of Plants by Tom Michaels, Saint Paul, Emily Hoover, Saint Paul, Laura Irish, Saint Paul, 2022, Publisher: University of Minnesota Libraries Publishing- ISBN 13: 9781946135872
3. Hickman C, Jr., Keen S, Larson A, Eisenhour D, I'Anson H, Roberts LS (2017). Integrated Principles of Zoology, McGraw-Hill Education; 17th edition.
4. Gilbert SF, Barresi MGF (2016). Developmental biology. Sinauer Associates Inc; 11th edition.
5. Gardener EJ, Simons, Snustad DP (2006). Principles of Genetics. Wiley, India.
6. Moyes CD, and Schulte PM (2007). Principles of Animal Physiology. Pearson Press, 2007
7. Strachan T and Read A (2018). Human Molecular Genetics. Garland Science, CRC Press, Florida, 2018
8. Futuyama DJ (1986). Evolutionary Biology. Sinauer Associates; 2nd edition.

Reference Books:

1. Introduction to Plant Science by Rick Parker, Taylor & Francis, 2004
2. Nielson (Knut Schmidt), Animal Physiology Adaptation and Environment, Cambridge University Press, 1994
3. Hadely ME and Levine JE, Endocrinology, Benjamin Cummings, San Francisco. 2006

⬅ Back to Index

Statistical Methods

Course Code : UBIN102

Minor Course

Credits : 4 **Hours : 60**

Course Objective:

To train the students with essential statistical methods which are important in data analysis

Unit 1 **8 Lectures**

Univariate data analysis : Types of data – nominal, ordinal, discrete, and continuous, Data presentation- Tabular and Graphical representation, Data analysis **-** Measures of Central tendency, Measures of Dispersion- Absolute and Relative measures, Measures of Shape – Skewness and Kurtosis

Unit2 **10 Lectures**

Bivariate data analysis: Types of variables, Methods – Scatter plot, Correlation analysis – Pearson's correlation, Spearman's rank correlation, and Partial correlation with properties, Principle of least squares, Linear Regression analysis and properties

Unit 3 **10 Lectures**

Multivariate data analysis: Understanding multivariate data structures, Methods: Principal, Component Analysis (PCA), Cluster Analysis – Hierarchical clustering and K-means clustering, **Multivariate Regression – Logistic regression**

Unit 4 **10 Lectures**

Probability Theory and Distribution: Classical and Axiomatic definition of probability, Probability on union and intersection, Conditional probability, Bayes’ Theorem, Bonferroni's inequality; Concept on Theoretical Univariate Distribution – Discrete and Continuous probability distribution- Binomial, Poisson, Hypergeometric and Normal distribution

Unit 5 **10 Lectures**

Estimation and Testing of Hypothesis: Concept of Estimation- Use of Statistic and Population Parameter, Hypothesis Testing- Null hypothesis and Alternative hypothesis, Type 1 and Type 2 error, Level of significance, Specific **Statistical test -** Student’s t-test**, Chi-square test and F-test**

Course Outcome:

The students will learn how to use statistical methods to gain meaningful insights from data

Text Books:

1. Biostatistics (9th Ed.) by Wayne W. Daniel, John Wiley & Sons Publishers, 2018.
2. Biostatistical Analysis (5thedition) by Jerrold H. Zar, Pearson Publisher, 2018

Reference Books:

1. Statistical Methods (1stEd.) by N. G. Das, Tata McGraw-Hill, 2017.
2. Fundamentals of Biostatistics (6th Ed.), by Bernard Rosner, Thomson Brooks/Cole, 2015.

⬅ Back to Index

Bioinformatics - A Primer

Course Code : UBIN103

Multidisciplinary Course

Credits : 3 **Hours : 45**

Course Objective:

Understanding the basic and principles of Bioinformatics.

Unit 1

Physical and chemical properties of water: Polarity of water, Hydrophilic nature, Hydrophobic nature, Osmosis; Ionising property, pH of solution, Buffering system. **Introduction to proteins**, amino acids - building blocks, chirality of amino acids, ionisation of amino acids, different nature of amino acids. Absorption behaviour, symbols, peptide bond, conformation, trans and cis peptide, Ramachandran plot. Protein purification and analysis,bProtein sequencing techniques: Edman degradation method, Mass Spec.

Unit 2

Fundamentals of DNA and RNA: Nucleotides, Chemical structure of nucleic acids, polymers of Nucleotides, base pairing, Hydrogen bond, DNA double helix, RNA single strand. Involvement of genetic information, protein coding. Nucleic acid sequencing methods – Sanger Method, Maxum-Gilbert method. **Carbohydraes:** Monosaccharides-aldoses and ketoses, Polysaccharides, Glycoproteins. **Lipids**: Properties of Fatty acids, Triacylglycerols, Glycerophospholipids, Sphingolipids, steroids.

Unit 3

Introduction to Bioinformatics:

Aim and branches of Bioinformatics, Application of Bioinformatics, Role of internet in bioinformatics. **Primary Databases**: Protein and Nucleic acid sequences - Types of Nucleotide Sequence: Genomic DNA, Complementary DNA (cDNA), Recombinant DNA (rDNA), Expressed sequence tags (ESTs), Genomic survey sequences (GSSs). Nucleic acid sequence databases: GenBank, EMBL, DDBJ. **Protein Sequence Databases:** Swiss-Prot, TrEMBL, UniProt, UniProtKB, UniParc, UniRef, UniMES;

Unit 4

Derived Databses: Sequence motifs Databases:- Prosite, ProDom, Pfam, InterPro, Gene Ontology; Polymorphism and mutation database- BioMuta, dbSNP- Database of short Genetic Variation, DMDM- Domain Mapping of Disease Mutations. Secondary structure databases:-Structural Classification of Proteins –SCOP, Class Architecture Topology Homology –CATH, Families of Structurally Similar Proteins –FSSP, Catalytic Site Atlas –CSA; Molecular functions / Enzymatic catalysis databases:- KEGG ENZYME database; Protein-Protein interaction database:- STRING, BioGRID, MINT;

Unit 5

Structure databases: PDB, NDB, MMDB; Chemical Structure database:- Pubchem, DrugBank, ChEMBL

Course Outcome:

Students come to know the thorough understanding of Principles of Bioinformatics.

⟵ Back to Index

Good laboratory practice

Course Code : UBIN104

Skill Development Course

Credits : 3 **Hours : 45**

Course Objective:

To understand and train the students in basic laboratory handling procedures.

- Good laboratory practice
- Good manufacturing practice and National and International regulations
- Regulations for recombinant DNA research and manufacturing process
- Regulations for clinical trials, Documentation and Compliance, in India and selected countries –
- Rules for import and export of biological materials.

Course Outcome:

Students will learn how to handle the biological samples and how effectively students use the laboratory.

Click or tap here to enter text.

⟵ Back to Index

Introduction to Computers

Course Code : UBIN105

Value added Courses

Credits : 2 **Hours : 30**

Course Objective:

To understand the basic knowledge on computer, & Internet terminologies and problem-solving techniques and browsing of biological databases and Websites.

Unit I-

Introduction, Basic Applications of Computer Components of Computer, Connecting Computer Components, Computer Hardware & Software, Basic concepts– Semiconductor RAM– ROM– Speed– Size and cost– Cache memories– Improving cache performance– Virtual memory– Memory management requirements– Associative memories Secondary storage devices.

Unit II-

Introduction to Databases and Transactions Physical Database design, what is database system, purpose of database system, view of data, relational databases, the importance of data models, Database security, Introduction to SQL, Data Definition of SQL, Basic structure of SQL queries, Basic SQL operations. e.g. Biological databases

Unit III-

Problem Solving Techniques Algorithm & Flowchart for Biological problem, Overview of programming language Hierarchy Development of source code in C language– Compilation and Execution of the problem– Refinement of source code by Testing and Debugging

Unit IV-

Internet Technologies Introduction to diverse Internet terminologies- Domain Name System– Function and Services of Internet Protocol Address, Web organizations and addressing, web browsers, web servers, Security and vulnerability, Web system Architecture-URL, Client side and server. e.g. Biological Web Portals

Unit V-

Web technologies Basics of HTML tags & skill of creating web pages, designing of webpage- Document Layout, Creating a WebPage and WebSite, Cloud Service Components- Deployment Models– Application of Cloud Computing. e.g. Biological Websites

Course Outcome:

Able to work in different operating systems and to learn shell programming that helps to handle UNIX which is mandatory for Bioinformatics.

Text Books:

1. Dharmendra B. Kadia. Basic Knowledge of Computer. Shanti Prakashan 2010.

Reference Books:

1. Internet Technologies Handbook:2023 Mark A Miller
2. Narasimha Karumanchi. 2023 Elements of Computer Networking. Career Monk publications.

⬅ Back to Index

SEMESTER - II

Biological and Chemical Databases

Course Code : UBIN201

Major Course

Credits : 4 **Hours : 60**

Course Objective:

To introduce most of the effectively used Bioinformatics databases and their applications in the field of Bioinformatics.

Unit 1 **12 lectures**

Introduction to biological databases: Biological data and formulation of databases. Emergence and objectives of Bioinformatics, Components of Bioinformatics: Data (example: Sequences, structures, pathways, literature, etc.;), Databases, Search engines.

Unit 2 **12 lectures**

Types of Biological databases: Primary databases: Nucleic acid database (Genbank, ENA, DDBJ), Secondary databases (Swiss-Prot, TrEMBL, UniProt, UniProtKB,InterPro, etc). Composite databases:- OWL, NRD and Swissport+TREMBL. Literature Databases:- Open access and open sources, PubMed, PLoS, Biomed Central, NAR databases; Bioinformatic Resources:- NCBI, EBI, ExPASy, RCSB.

Unit 3 **12 lectures**

Structure databases: PDB; Sequence motifs Databases:-Prosite, ProDom, Pfam, InterPro, Secondary structure databases:-Structural Classification of Proteins –SCOP, ClassArchitecture Topology Homology –CATH, Catalytic Site Atlas –CSA; Molecular functions / Enzymatic catalysis databases:-KEGGdatabase; Chemical Structure database:- Pubchem

Unit 4 **12 lectures**

Organismal Databases –Microbial Genome Database-MBGD: C.elegans (WormBase), Yeast(SGD), Drosophila (FlyBase), PlasmodiumDB;; Viral genome database:-ICTVdb; Mammalian Database: Rat, Mouse, Human (OMIM / OMIA), plants – Arabidopsis thaliana (TAIR), Rice

Unit 5 **12 lectures**

Database search engines: Text-based search engine (Entrez), Sequence similarity based search engines (BLAST), Motif-based searchengines (ScanProsite), Structure similarity based search engines (VAST, DALI). Proteomics tools: - ExPASy server.

Course Outcome:

Students will understand the databases available and their applications in the field of Bioinformatics, Biomedical research, etc.

Text Books:

1. Bioinformatics- a Practical Guide to the Analysis of Genes and Proteins byBaxevanis, A.D. and Francis Ouellellette, B.F., Wiley India Pvt Ltd. 2009
2. Essential Bioinformatics by Jinxiong.,Combrridge University press,New York.2006

Reference Books:

1. Bioinformatics: Sequence and Genome Analysis by Mount D., Cold Spring HarborLaboratory Press, New York. 2004
2. Introduction to bioinformatics by Teresa K. Attwood, David J. Parry-Smith. PearsonEducation. 1999 Old editions

⟵ Back to Index

Introduction to Bioinformatics

Course Code : UBIN202

Minor Course

Credits : 4 **Hours : 60**

Course Objective:

The main objective of this soft course is to introduce general concepts of Bioinformatics to the allied department students of Pondicherry University.

Unit 1 **9 lectures**

Introduction: Aim and branches of Bioinformatics, Application of Bioinformatics, Role of internet and www in bioinformatics. Basic biomolecular concepts: Protein and amino acid, DNA & RNA, Sequence, structure and function. Forms of biological information, Types of Nucleotide Sequence: Genomic DNA, Complementary DNA (cDNA), Recombinant DNA (rDNA), Expressed sequence tags (ESTs), Genomic survey sequences (GSSs). DNA sequencing methods: Basic and Automated DNA sequencing, DNA sequencing by capillary array and electrophoresis, Gene expression data.

Unit 2 **9 lectures**

Bioinformatics Resources: NCBI, EBI, ExPASy, RCSB, DDBJ: The knowledge of databases and bioinformatics tools available at these resources, organization of databases: data contents, purpose and utility. **Open access bibliographic resources and literature databases:** PubMed, BioMed Central, Public Library of Sciences (PloS), CiteXplore.

Unit 3 **9 lectures**

Sequence databases: Nucleic acid sequence databases: GenBank, EMBL, DDBJ; Protein sequence databases: Uniprot-KB: SWISS-PROT, TrEMBL, UniParc; **Structure Databases**: PDB, NDB, PubChem, ChemBank. **Sequence file formats:** Various file formats for bio- molecular sequences: GenBank, FASTA, GCG, MSF etc. **Protein and nucleic acid properties:** Proteomics tools at the ExPASy server, GCG utilities and EMBOSS, Computation of various parameters.

Unit 4 **9 lectures**

Sequence Analysis: Basic concepts of sequence similarity, identity and homology, definitions of homologues, orthologues, paralogues and xenologues **Scoring matrices:** basic concept of a scoring matrix, Matrices for nucleic acid and proteins sequences, PAM and BLOSUM series, matrix derivation methods and principles.

Unit 5 **9 lectures**

Sequence alignment: Measurement of sequence similarity; Similarity and homology. **Pairwise sequence alignment:** Basic concepts of sequence alignment, Needleman and Wunsch, Smith and Waterman algorithms for pairwise alignments, gap penalties, use of pairwise alignments for analysis of Nucleic acid and protein sequences and interpretation of results.

Course Outcome:

The outcome of the paper is students from other departments gained the knowledge of how to utilize bioinformatics resources.

Text Books:

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 Ouellellette, B.F., Wiley India Pvt Ltd. 2009

Reference Books:

1. Introduction to bioinformatics by Teresa K. Attwood, David J. Parry-Smith, Pearson Education. 1999
2. Bioinformatics for Dummies by Jean-michel Claverie Cedric Notredame. Publisher: Dummies (Jan 2007)

⬅ Back to Index

Concepts in programming language

Course Code : UBIN203

Multidisciplinary Course

Credits : 3 **Hours : 45**

Course Objective:

To introduce the computer programming concepts to students. This course suits for students with non-computer background

Unit 1 **9 Lectures**

Basics of Computing principles: Concepts of Input and Output – Input devices – Output devices – Number systems – Binary, Octal, and hexadecimal systems – Boolean algebra (AND, OR, NOT and EXOR operations) – DeMorgan's theorem.

Unit 2 **9 Lectures**

Operating Systems: Concept of operating system – DOS, UNIX and LINUX – Types of processing – multiprogramming, multitasking, and multithread processing.

Unit 3 **9 Lectures**

Problem Solving Techniques: Basics of problem solving techniques – Understanding the problem – Algorithm – Flowchart – coding.

Unit **9 Lectures**

Classification of computer languages: Human and computer languages – Machine level language – Assembly level language – High level language – Assembler and compiler – compilation – basic principles of coding – Testing of a program – debugging – syntax and semantic errors.

Unit 5 **9 Lectures**

Concepts in computer programming: Role of decision making – illustration of 'if' – Concept of loops – importance of loops in programming – arrays and stacks – file handling – interface with internet.

Course Outcome:

The students will possess basic knowledge about programming languages. On successful completion of the course, the student will be able to self-learn most of the programming languages.

Reference Books:

- Fundamentals of Computers, Second Edition (2019) by Reema Thareja, Oxford University Press.
- Operating System Concept, Ninth Edition (2018) by Abraham Silberschatz and Peter B. Galvin, Wiley.

⬅ Back to Index

Analytical Techniques

Course Code : UBIN204

Skill Development Course

Credits : 3 **Hours : 45**

Course Objective:

To understand the principles of physical sciences that form the basis of the techniques and instrumentation used in research field.

Unit 1 **9 lectures**

Electrophoresis: Theory and types; moving boundary electrophoresis, zone electrophoresis, paper, cellulose acetate gel electrophoresis, Native PAGE, disc PAGE, Gradient PAGE, SDS PAGE, DNA agarose gel electrophoresis, Southern, Northern, Western blotting techniques, Isoelectric focusing, finger printing, DNA sequencing, Pulsed - field Electrophoresis, Capillary Electrophoresis.

Unit 2 **9 lectures**

Chromatography: Principles, methodology and applications of chromatography using paper, thin layer, column (gel filtration, ion exchange, and affinity), gas and types of HPLC.

Unit 3 **9 lectures**

Centrifugation : Principles, types and applications. Ultracentrifugation- types, optical methods used and applications of preparative and analytical ultracentrifuges.

Unit 4 **9 lectures**

Enzyme kinetics: Membrane potential, Active site, Cofactors, apo-enzymes, Enzyme specificity, Factor affecting enzyme activity, Michaelis-Menten, LB Plot, Km/kcat, Types of inhibition, Allosteric enzymes.

Unit 5 **9 lectures**

Macromolecular interactions: Isothermal Titration Calorimetry Optical and magnetic tweezers, Fluorescence Resonance Energy Transfer (FRET) Dual Polarisation Interferometry [DPI] CD/ORD, DLS.

Course Outcome:

Students will know the physical basis of appropriate strategies and instrumentation for analysis of different biological sample types.

Text Books:

1. Principles and Techniques of Practical Biochemistry (7th Ed) by Keith Wilson and John Walker, Cambridge University Press. 2010.
2. Principles of Biochemistry by Nelson and Cox, Lehninger. W H Freeman & Co. 2009

Reference Books:

1. Physical Biochemistry (2nd Ed) by D. Freifelder., Freeman. 1982
2. Fundamentals of Biochemical calculation (2nd Ed.) by Krish Moorthy CRC Press. 2007
3. Protein Purification - Principles & Practices (3rd Ed.) by R. Scopes., Springer Verlag. 1994
4. Biophysical Chemistry: Techniques for the study of biological structure and functions by Charles C. R. & Paul. S. R., W.H. Freeman & Co. New York. 2004

Introduction to R Language

Course Code : UBIN205

Value added Course

Credits : 2 **Hours : 30**

Course Objective:

The main goal of this course is to introduce the student to the R environment for biological data analysis using various statistical methods

1. Introduction to R (Installation of R and Rstudio, R Script, text editor, R project)

2. Understanding data types in R (Numeric, characters, string, integer, logical, complex) and functions (built-in and user-defined), Variables.

3. Understanding primary data structures in R: Vectors, Matrices, Lists

4. Understanding primary data structures in R: Arrays, data frames, factors

5. Built-in datasets in R, data exploration, data summarization, external file reading, data storing, R workspaces, file handling (working directory, reading .txt/.csv files, writing to files, checking files existence, listing files in a folder)

6. Manipulation of objects in R: Mathematical operations on R objects: Basic operations (addition, subtraction, element-wise multiplication, element-wise division), Recycling rule in R, propagation of names, dimensional attributes, handling missing values

7. Creating the matrices & basic matrix operations (matrix addition, subtraction, element-wise multiplication, matrix multiplication, outer product, transpose of matrix, eigenvalues and eigenvectors)

8. Textual operations in R (create and view strings, combine strings with and without space, changing case, substrings & string length, finding patterns, replacing substrings, splitting strings)

9. Statistical operations in R (Hypothesis testing, parametric & non-parametric tests, chi-square test, t-test, ANOVA, Correlation & regression analysis, principal component analysis PCA)

10. Basic graphics in R (high-level plotting, low-level plotting), scatter plot, barplot, boxplot, histogram, heatmap)

11. Interactive graphics (interactive click & keyboard event), advanced Parameters for creating plots, graphs, and charts in R (for bioinformatics analysis)

12. Introduction to Bioconductor and R Packages for Bioinformatics; Working with Bioinformatics Data (Sequences & Annotations)

Course Outcome:

The student will have an understanding of various statistical methods employed in biological data analysis. He/she will be able to perform statistical modelling and analysis in the R environment.

Text Books:

1. Paul Gerrard and Radia M. Johnson. Mastering Scientific Computing with R. Packt Publishing, UK, 2015.
2. P.P. Sinha. Bioinformatics with R Cookbook. Packt Publishing, UK, 2014.

Reference Books:

1. Florian Hahne, Wolfgang Huber, Robert Gentleman, Seth Falcon. Bioconductor case studies. Springer, 2008.
2. Paul D. Lewis, R for Medicine and Biology, Jones and Bartlett Series, 2010.

⬅ Back to Index

SEMESTER - III

Physical sciences

Course Code : UBIN301

Major Course

Credits : 4 **Hours : 60**

Course Objective:

The main objective of this course is to help the students to understand the basic concept of fundamental Physics and Chemistry and their applications in biology

Unit 1: **10 Lectures**

Classical Mechanics: Types of Motion**:-**Uniform, projectile, circular and relative motions, Newton’s Laws of Motion, Law of Gravitation, **Work and energy**:- work energy theorem, conservative / non-conservative forces, energy conservation, power, Linear momentum and collisions (elastic and inelastic), impulse, momentum theorem, **Rigid body rotation**:**-** angular velocity and acceleration, rotational kinetic energy, inertia, torque, dynamics of rotation, **Angular Momentum**:- conservation of angular momentum, translation and rotation, Statics Oscillatory motion

Unit 2: **10 Lectures**

Black body radiation, photoelectric effect, Bohr’s Model of Hydrogen atom, De Broglie’s Hypothesis, Harmonic wave function, wave packets, Heisenberg uncertainty principle, Eigen states and eigen values, Pauli Exclusion Principle, Schrodinger equation,

Unit 3: **10 Lectures**

Continuum Model, System (closed, isolated), State functions & variables, Adiabatic & diathermal boundary walls, Equilibrium, Process, equation of state. Heat, Zeroth Law of Thermodynamics, Heat Conduction Equation, The First Law of Thermodynamics, Work, Entropy, The Second Law of Thermodynamics:- reversibility and irreversibility, free and isothermal expansions, Heat Capacity, Isothermal and reversible-adiabatic expansion of an Ideal Gas , Enthalpy, Change of state, Latent heat and Enthalpy, Carnot cycle, Gibbs and Helmholtz free energy, Young's Modulus, The Third Law of Thermodynamics.

Unit 4 **12 Lectures**

Introduction to Inorganic Chemistry: Atomic Structure - Elements and compounds, atoms and molecules-definition, Classical atomic models - J. J. Thomson, E. Rutherford, N. Bohr. Electronic configuration- aufbau principle - Pauli exclusion principle - Hund's rule- Modern periodic table, periodicity. Chemical bonds - ionic bonding - covalent bonding - Coordinate covalent bonding. Overlap of σ and π orbitals – hybridization, resonance, Bond properties, Molecular geometry, Intermolecular forces.

Unit 5 **6 Lectures**

Symmetry and Principles: Definitions and theorems of group theory, subgroups, Classes. Molecular symmetry and symmetry groups – symmetry elements and operations. Symmetry planes, reflections, inversion centre, proper/ improper axes of rotation, symmetry elements and optical isomerism, symmetry point groups, classes of symmetry operations, classification of molecular point groups.

Unit 6 **12 Lectures**

Introduction to Organic Chemistry: Carbon and its compounds, Position of Carbon in periodic table, tetra covalency of carbon, functional groups. **Stereochemistry:** Concept of isomerism, types of isomerism, optical isomerism, elements of symmetry, molecular chirallity, enantiomers, stereogenic centres, optical activity, properties of enantiomers, chiral and achiral molecules with two stereogenic centres, distereoisomers, mesocompounds, resolution of enantiomers. Relative and absolute configurations, sequence rules, D & L R & S systems of nomenclature. **Aromatic and Heteroaromatics molecules**: Five / six membered hetero aromatics and analogues, Nucleic acid bases. **Chemistry of Biomolecules**: Proteins, enzymes, vitamins, nucleic acids. Drugs and their classification, Drug-target interaction, therapeutic action of different drugs.

Course Outcome:

Students would gain the knowledge of physics and chemistry as well as their applications in biology

Reference Books:

1. Physics for Scientists and Engineers (6th Ed.) by Raymond A. Serway, John W. Jewett, Thomson Brooks/Cole. 2004.
2. Fundamental Principles of Physical Chemistry (Prutton, Carl F.; Maron, Samuel H.),1949.
3. Organic Chemistry by Morrison and Boyd Sixth Edition. 1992

Reference Books:

1. Physics for Scientists and Engineers by Paul A. Tipler, Gene P. Mosca. Freeman Company. 2007.
2. Fundamentals of Physics by Resnick, Halliday and Walker. 200.
3. Chemistry, The Central Science, 10th edition, Theodore L. Brown; H. Eugene LeMay, Jr.; and Bruce E. Bursten. 2006.
4. Selected Topics in Inorganic Chemistry, Wahid U. Malik, G. D. Tuli and R.D. Madan.1993.
5. Chemistry3 Introducing inorganic, organic and physical chemistry, Andrew Burrows, John Holman, Andrew Parsons, Gwen Pilling, Gareth. 2013.
6. Organic Chemistry by Paula Yurkanis Bruice, Prentice Hall. 2010 Heterocyclic chemistry at a glance, John A. Joule and Keith Mills.2007

⟵ Back to Index

Mathematical Sciences

Course Code : UBIN302

Major Course

Credits : 4 **Hours : 60**

Course Objective:

To train the students with essential mathematical and statistical concepts which are vital to Bioinformatics.

Unit 1 **8 Lectures**

Matrices: Different types of matrices – Matrix operations – Transpose of a matrix – Symmetric and skew-symmetric matrices – Determinant of a matrix – Minors and Cofactors – Properties of Determinants – Adjoint matrix – Inverse of a matrix – Applications of matrices: Cramer’s rule – Matrix inverse method – Gauss-Jordan Method – Eigen values and eigen vectors.

Unit 2 **8 Lectures**

Calculus: Derivative of a function –Product rule, Chain rule, Quotient rule. Maxima and Minima of a function, Introduction to Partial Differentiation, Integral Calculus: The Idea of the Integral, The Definite Integrals, Indefinite Integrals.

Unit 3 **8 Lectures**

Differential Equation: Introduction to Ordinary Differential Equation, Equations of First order - Variables separable equation, Homogeneous equation, Linear equation (Leibniz’s and Bernoulli’s equation), Integrating factor and Exact equation

Unit 4 **8 Lectures**

Descriptive Statistics: Data types – Quantitative and Qualitative, Data presentation- Tabular and Graphical representation, Data analysis **-** Measures of Central tendency, Measures of Dispersion- Absolute and Relative measures, Measures of Shape – Skewness and Kurtosis

Unit 5 **8 Lectures**

Association of Variables and Classification: Scatter diagrams, Correlation – Pearson's correlation and Spearman's rank correlation coefficient, Correlation coefficient, Linear Regression analysis, Classification – Supervised Learning-Logistic Regression, Decision Trees, K-Nearest Neighbours, Evaluating Classification Models- Accuracy, precision, ROC Curve and AUC.

Unit 6 **8 Lectures**

Estimation and Testing of Hypothesis: Concept of estimation-Sample Statistic and Population Parameter, Null hypothesis and Alternative hypothesis, Type 1 and Type 2 error, Level of significance, Parametric Tests: t-test (paired and independent) , F-test, ANOVA, Non-parametric Tests: **Chi-square test and Mann-Whitney U test.**

Course Outcome:

The students will gain skills in solving mathematical and statistical problems which are essential to understand advanced courses in Bioinformatics.

Text Books:

1. Algebra by Serge A. Lang, Pearson Education. 2003
2. Introduction to Calculus & Analysis, Vol I and II by Richard Courant & Fritz John, Springer publisher.1999
3. Biostatistics (9th Ed.) by Wayne W. Daniel, John Wiley & Sons Publishers, 2018.
4. Biostatistical Analysis (5thedition) by Jerrold H. Zar, Pearson Publisher, 2018

Reference Books:

1. Basic Mathematics by Serge A. Lang. Springer Publisher. 1988
2. A First Course in Calculus by Serge A. Lang. Springer publisher. 1986
3. Higher Engineering Mathematics (40th Ed), by B.S. Grewal and J.S. Grewal. Khanna Publishers, New Delhi. 2007.
4. Statistical Methods (1stEd.) by N. G. Das, Tata McGraw-Hill, 2017.
5. Fundamentals of Biostatistics (6th Ed.), by Bernard Rosner, Thomson Brooks/Cole, 2015.

⬅ Back to Index

Bioethics and Research Methodology

Course Code : UBIN303

Minor Course

Credits : 4 **Hours : 60**

Course Objective:

To provide overview of how to identify research problem and conduct research ethically and communicated them to peers.

Unit 1 **10 Lectures**

Regulatory Procedures for research and manufacturing of biotechnology products: Good laboratory practice, Good manufacturing practice and National and International regulations - Regulations for recombinant DNA research and manufacturing process - Bio-safety and Bioethics - Regulations for clinical trials, Documentation and Compliance, in India and selected countries - Rules for import and export of biological materials.

Unit 2 **10 Lectures**

Research Methodology: Objectives of research and motivation; Problem Identification & Formulation – Research Question - Hypothesis and Hypothesis Testing; Types of research - Qualitative vs Quantitative Research - Applied vs. Fundamental Research; Data Collection - Data Analysis - Interpretation of results and Report writing.

Unit 3 **10 Lectures**

Scientific writing – Introduction - Types of scientific writings - Thesis or dissertation writing – Research paper writing; Types of publications - Open access and subscription based resources; Scientific paper writing - Choosing a journal- Instructions to authors - Structure and Style- Authorships –figures tables with legends - References and citations - Acknowledgements- Conflict of interest; Peer review mechanism and publication process; Scientometric Analyses of a paper/journal; Ethics in publishing and Plagiarism issues. Use of software for Reference Management – (Mendeley/endnote) and detection of Plagiarism (Turnitin).

Unit 4 **15 Lectures**

Oral presentation – Planning the oral presentations and visuals- In-class discussion (Students in small groups or individually will take up the assignments or select a research project/ topic and prepare oral presentations followed by a Q&A sessions).

Unit 5 **15 Lectures**

Poster Presentation – Elements and Significance of poster presentations- Planning and designing a poster- Individual Poster presentation (Students select a research project/topic and prepare posters followed by a Q&A sessions).

Course Outcome:

Students can understand the basics of how to design, conduct research, analyze and communicate the results to research community in an ethical manner. Also team work ethos and stress management strategies would help to cope-up with their day-to-day life in a competitive world.

Text Books:

1. Bioethics and Biosafety in Biotechnology by Sree Krishna V., New Age International
2. (P) Ltd., Publ., Mumbai. 2007.
3. Successful Scientific Writing: A step-by-step Guide for Biomedical Scientists by J.R. Matthews and R.W. Matthews, 4th Ed., Cambridge University Press. 2014
4. Writing and Presenting Scientific Papers, by Brigitta Malmfors, Phil Garnsworthy and Michel Grossman, 2nd Ed., Viva Books Pvt. Ltd. 2011.

Reference Books:

1. IPR, Biosafety and Bioethics, by Deepa Goel and Shomini Parashar 1st Ed., Pearson Education India, 2013.
2. From Research to Manuscript: A Guide to Scientific Writing by Michael Jay Katz, 2nd Ed., Springer. 2009
3. Scientific Writing: Easy When You Know How by Jennifer Peat, BMJ books. 2002

⬅ Back to Index

Bioinformatics – Databases and sequence analysis

Course Code : UBIN304

Multidisciplinary Course

Credits : 3 **Hours : 45**

Course Objective:

To introduce most of the effectively used Bioinformatics databases and their applications in the field of Bioinformatics.

Unit 1 12 lectures

Introduction to biological databases: Biological data and formulation of databases. Emergence and objectives of Bioinformatics, Components of Bioinformatics: Data (example: Sequences, structures, pathways, literature, etc.;), Databases, Search engines.

Unit 2 12 **lectures**

Types of Biological databases: Primary databases: Nucleic acid database (Genbank, ENA, DDBJ), Secondary databases (Swiss-Prot, TrEMBL, UniProt, UniProtKB,InterPro, etc). Composite databases:- OWL, NRD and Swissport+TREMBL. Literature Databases:- Open access and open sources, PubMed, PLoS, Biomed Central, NAR databases; Bioinformatic Resources:- NCBI, EBI, ExPASy, RCSB.

Unit 3 **12 lectures**

Structure databases: PDB; Sequence motifs Databases:-Prosite, ProDom, Pfam, InterPro, Secondary structure databases:-Structural Classification of Proteins –SCOP, ClassArchitecture Topology Homology –CATH, Catalytic Site Atlas –CSA; Molecular functions / Enzymatic catalysis databases:- KEGGdatabase; Chemical Structure database:- Pubchem

Unit 4 **12 lectures**

Organismal Databases –Microbial Genome Database-MBGD: C.elegans (WormBase), Yeast(SGD), Drosophila (FlyBase), PlasmodiumDB;; Viral genome database:-ICTVdb; Mammalian Database: Rat, Mouse, Human (OMIM / OMIA), plants – Arabidopsis thaliana (TAIR), Rice

Unit 5 12 lectures

Database search engines: Text-based search engine (Entrez), Sequence similarity based search engines (BLAST), Motif-based searchengines (ScanProsite), Structure similarity based search engines (VAST, DALI). Proteomics tools: - ExPASy server.

Course Outcome:

Students will understand the databases available and theirapplications in the field ofBioinformatics, Biomedical research, etc.

Text Books:

1. Bioinformatics- a Practical Guide to the Analysis of Genes and Proteins byBaxevanis, A.D. and Francis Ouellellette, B.F., Wiley India Pvt Ltd. 2009
2. Essential Bioinformatics by Jinxiong.,Combrridge University press,New York.2006

Reference Books:

1. Bioinformatics: Sequence and Genome Analysis by Mount D., Cold Spring HarborLaboratory Press, New York. 2004
2. Introduction to bioinformatics by Teresa K. Attwood, David J. Parry-Smith. PearsonEducation. 1999 Old editions

⬅ Back to Index

Application of PCR

Course Code : UBIN305

Skill Development Course

Credits : 3 **Hours : 45**

Course Objective:

Students of this course should be able to learn:

• Concept of PCR and different types of PCR. • Principles of oligonucleotide (primer) synthesis and purification. • Designing of Primers for PCR. • Hands-on setting up of PCR reaction and analysis of the amplified product. • Purification of PCR product

Unit 1 Concept of PCR and different types of PCR with some applications 10 hours

Principle of Polymerase Chain reaction and amplification process, use of thermocycler and other equipment required to perform PCR and analysis of amplified DNA, use of synthetic oligonucleotide as primers in PCR.

Practical: 1.1 Demonstration of thermocycler, setting of conditions for a PCR for different types of templates and understanding the reagents used in PCR.

1.2 Chemistry of oligonucleotide synthesis and purification techniques, designing the Primer sequence for PCR through the use of online free software. Handling and storage of Primers for long term use.

Unit 2: PCR reaction for amplification and analysis of the amplified product 10 hours

Designing and Setting up PCR, understanding about different reaction components, use of different types of polymerases for different length of amplicons, concept of fidelity and processivity of polymerases.

Practical: 2.1

Demonstration of a PCR to explain all the steps required for setting up the reaction and analysis of the amplified product.

2.2 Designing and setting up a PCR individually to amplify a 500 bp product and analysis of the amplified product using agarose gel electrophoresis.

2.3. Designing and setting up a PCR individually to amplify a 1500 bp product and analysis of the amplified product using agarose gel electrophoresis.

Unit 3: Purification of PCR-amplified DNA and downstream applications 10 hours

Purpose of purification of PCR product, methods for purification, downstream applications of PCR products.

Practical: 3.1 Purification of PCR products by spin column method

3.2 Agarose gel electrophoresis and visualization of the product after purification.

Course Outcome:

At the end of this course, students should be able to learn and perform: • PCR and its application for research. • Designing of Primers for PCR and obtaining Primers as synthetic oligonucleotides of • appropriate quality from commercial sources. • Designing conditions for setting up a PCR and perform analysis of amplified DNA. • Purification of PCR-amplified DNA using columns for downstream application.

Reference Books:

1. Sambrook J, Fritsch EF & Maniatis T. Molecular Cloning. A laboratory Manual. 3rd Edition. Cold Spring Harbor Laboratory Press. New York.
2. Ausubel FM, Brent R, Kingston RE, Moore DD, Seidman JG, Smith JA, Struhl K. Current Protocols in Molecular Biology. (eds.) John Wiley & Sons, Inc. New York.

⬅ Back to Index

Intellectual Property Rights

Course Code : UBIN306

Value added Courses

Credits : 2 **Hours : 30**

Course Objective:

Aims to provide UG students with the knowledge about IPR

Unit-I **8 lectures**

IPR Definition and implications. Forms of IPR Protection - Copyright, Trademark, Trade secrets and Patents. IPR in Indian Scenario & Indian laws for IPR protection.

Unit-II **8 lectures**

International organization and treaties for IPR protection – World Trade Organization (WTO) & Trade Related Aspects of Intellectual Property Rights (TRIPS) Agreements. World Intellectual Property organization (WIPO) & Patent Cooperation Treaty (PCT).

Unit-III **9 lectures**

Patents: National and international agencies for patenting - US Patent & Trademark office (USPTO), Controller General of Patents, Designs & Trade Marks, India (CGPDTM), Requirements for patentability, Composition of a patent, how to apply and get patents – US, Indian and PCT.

Unit-IV **5 lectures**

Case studies and Hands-on session for searching patents on US, Indian and PCT databases.

Course Outcome:

Students will have basic understanding of different forms of IPR protection, national and international agency for IPR protection and how to protect them

Reference Books:

1. Laws of Patents: Concepts and Cases Edited by A. V. Narasimha Rao, The ICFAI University Press (2005).
2. Intellectual Property Rights In India: General Issues And Implications by Prankrishna Pal. Publisher: Deep & Deep Publications Pvt. Ltd (2008).

3. Intellectual Property: The Law of Trademarks, Copyrights, Patents, and Trade Secrets, by Deborah E. Bouchoux, CENGAGE Learning Custom Publishing; 4th ed. edition (January 2013)

⟵ Back to Index

SEMESTER - IV

Database Management Systems

Course Code : UBIN401

Major Course

Credits : 3+1P **Hours : 45**

Course Objective:

- To teach the basic database concepts, applications, data models
- To familiarize Entity Relationship model for a database.
- To Describe the basics of SQL and construct queries using SQL
- To Emphasize the importance of normalization in databases.
- To Demonstrate the basic concepts of transaction processing and concurrency control.
- To familiarize the concepts of database storage structures and identify the access techniques.

Unit 1 **10 lectures**

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.

Unit 2 **9 lectures**

Data models: ER Model: Keys, Constraints, Design Issues, Extended ER features, Reductions of ER Schema to Tables. Relational Model: Structure.

Unit 3 **10 lectures**

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

Unit 4 **9 lectures**

Relational Database and Storage: Pitfalls in Relational Design Database, Functional dependencies, Decomposition Normal Forms -1NF, 2NF, 3NF, ordered indices, Hashing concepts - Security and Authorization.

Unit 5 **10 lectures**

Concurrency control techniques & Information retrieval: Transactions: Properties of transactions: Concurrency problems, Serialisability and Locking techniques, Granularity of Data Items – Database System Architecture and Information retrieval: Centralized and Client- Server Architecture.

Course Outcome:

- At the end of the course the students are able to:
- Use the basic concepts of Database Systems in Database design
- Design a Database using ER Modelling
- Apply normalization on database design to eliminate anomalies
- Analyze database transactions and can control them by applying ACID properties..

Reference Books:

1. "Database System Concepts", 6th Edition by Abraham Silberschatz, Henry F. Korth, S. Sudarshan, McGraw-Hill. Preferably the 6th or 7th edition
2. "Fundamentals of Database Systems", 7th Edition by R. Elmasri and S. Navathe, Pearson
3. "An introduction to Database Systems", C J Date, Pearson.
4. "Modern Database Management", Hoffer , Ramesh, Topi, Pearson. 5. "Principles of Database and Knowledge – Base Systems", Vol 1 by J. D. Ullman, Computer Science Press.

Lab - Database Management System

Hours : 15

COURSE OBJECTIVES:

- *To gain fundamental knowledge of SQL and to store the data and to share the data*
- *Understanding the role of SQL in Relational Database Management Systems.*
- *To practice SQL queries for database management.*

Structured Query Language

Introduction to SQL- Data Types

1. Data Definition Language (DDL):
    - **a.** Creating a Database
    - b. Creating a Table
    - c. Specifying Relational Data Types
    - d. Specifying Constraints
    - e. Creating Indexes
2. Data Manipulation Language (DML)
    - a. INSERT statement
    - b. Using SELECT and INSERT together
    - c. DELETE, UPDATE, TRUNCATE statements
    - d. DROP, ALTER statements
3. Data Retrieval Language (DRL)
    - a. Implementation of different types of function •
        - i. Number function
        - ii. Aggregate Function
        - iii. Character Function
        - iv. Conversion Function
        - v. Date Function
    - b. Implementation of different types of operators in SQL •
        - i. Arithmetic Operators
        - ii. Logical Operators
        - iii. Comparison Operator
        - iv. Special Operator
        - v. Set Operation
4. Implementation of different types of Joins
    - a. Inner Join, Outer Join, Natural Join
    - b. Study and Implementation of • Group By & having clause • Order by clause
5. An exercise using Open-Source Software like MySQL

COURSE OUTCOME: At the end of the course, the students are able to: Use the basics of SQL and construct queries using SQL in database creation and interaction.

SQL: The Complete Reference" by James R. Groff and Paul N. Weinberg

List of Open Source Software/learning website:

1. https://www.tutorialspoint.com/dbms/

2. https://www.w3schools.com/sql/

3. https://www.codecademy.com/learn/learn-sql 4. https://in.udacity.com/

⟵ Back to Index

Molecular Evolution

Course Code : 402

Major Course

Credits : 3+1P **Hours : 45**

Course Objective:

The main goal of this course is to help students in learning the basic concepts and computational methods involved in the molecular evolutionary analysis of genes and proteins.

Unit 1 **5 Lectures**

Evolution of DNA and proteins, the origin of the genetic code. Hardy-Weinberg equilibrium; Evolutionary changes by mutation, gene flow, natural selection and genetic drift.

Unit 2 **5 Lectures**

The concept of homology in molecular evolution. Role of transitions and transversions; chromosomal deletions and insertions in evolution. Role of pseudogenes, repetitive DNA, transposable elements and junk DNA in evolution.

Unit 3 **6 Lectures**

Neutral theory (Kimura) and nearly neutral theory (Ohta) of molecular evolution. Phylogenetic tree. Reconstruction of phyogenetic trees using distance matrix methods, the Maximum Parsimony method, Maximum likelihood and Bayesian inference. Estimation of selection at the molecular level.

Unit 4 **4 Lectures**

The concept of the Molecular Clock. Calibration. Limitation of molecular clock models. Human molecular clock: deducing evolutionary histories through mitochondrial DNA and Y chromosome.

Unit 5 **6 Lectures**

Evolution of the genome: Human Genome Project, ENCODE, Genome duplication (Ohno's hypothesis), Exon Shuffling, Concerted evolution. Evolutionary Medicine.

Course Outcome:

The student will be able to understand the molecular basis of the evolution of the genome. He/she will be able to analyse the genomic data using phylogenetics and infer the evolutionary explanation of a biological phenomenon.

Reference Books:

1. An Introduction to Molecular Evolution and Phylogenetics by Lindell Bromham, 2016, Oxford University Press.
2. Molecular Evolution by Wen Hsiung-Li, 1997, Sinauer Associates, Sunderland, MA. Reference Books:
3. Molecular Evolution and Phylogenetics by Masatoshi Nei and Sudhir Kumar, 2000, Oxford University Press.
4. Neutral Theory of Molecular Evolution by Motoo Kimura, 1985, Cambridge University Press.
5. Bioinformatics and Molecular Evolution by Paul G. Higgs and Teresa K. Attwood, 2013, Willey-Blackwell.

Lab - Molecular Evolution

Hours : 15

Course Objective:

The main goal of this course is to help students in learning the computational methods involved in the molecular evolutionary analysis of genes and proteins.

1. Introduction to genomic database- Ensembl
2. Retrieval of DNA and protein orthologues from Ensembl.
3. Multiple sequence alignment of DNA and protein sequences using AliView.
4. Introduction to MEGA, a software for molecular evolutionary analysis.
5. Estimation of the best model for DNA and protein using Modeltest.
6. Reconstruction of a neighbour-joining phylogenetic tree using MEGA.
7. Reconstruction of a maximum parsimony phylogenetic tree using MEGA.
8. Reconstruction of a maximum likelihood phylogenetic tree using MEGA.
9. Molecular clock tests using MEGA.
10. Reconstruction of a Bayesian inference phylogenetic tree using MrBayes.
11. Estimation of non-synonymous (dN) and synonymous (dS) substitutions on coding sequences using MEGA.
12. Gene-wide selection analysis using the BUSTED program in HyPHy.

Course Outcome:

The student will be able to understand the molecular basis of the evolution of the genome. He/she will be able to analyse the genomic data using phylogenetics and infer the evolutionary explanation of a biological phenomenon.

⟵ Back to Index

Fundamental of Biochemistry and Molecular Biology

Course Code : UBIN403

Major Course

Credits : 3+1P **Hours : 45**

Course Objective:

Make the students understand the fundamentals of biological macromolecules and their functions. Provide basic knowledge on biological concepts and processes in living system. The course includes their importance in understanding various bio molecular reactions and how they fold to their native, functional forms

Unit 1 **7 Hours**

The foundations of biochemistry: Cellular and chemical foundations of life. Water- Unique properties, weak interactions in aqueous systems, ionization of water, buffers, water as a reactant and fitness of the aqueous environment.

Unit 2 **8 Hours**

Carbohydrates: Monosaccharides - structure of aldoses and ketoses, ring structure of sugars, conformations of sugars, mutarotation, anomers, epimers and enantiomers, structure of biologically important sugar derivatives, oxidation of sugars. Formation of disaccharides, reducing and nonreducing disaccharides. Polysaccharides – homo- and heteropolysaccharides, structural and storage polysaccharides.

Unit 3 **8 Hours**

Lipids: Building blocks of lipids - fatty acids, glycerol, ceramide. Storage lipids - triacyl glycerol and waxes. Structural lipids in membranes – glycerophospholipids, galactolipids and sulpholipids, sphingolipids and sterols, structure, distribution and role of membrane lipids.

Unit 4 **8 Hours**

Amino acids and Vitamins: Structure and classification, physical, chemical and optical properties of amino acids. Structure and active forms of water soluble and fat soluble vitamins, deficiency diseases and symptoms, hyper-vitaminosis

Unit 5 **8 Hours**

Nucleic acids: Nucleotides – Chemical structure and properties. Nucleic acid structure and confirmation of nucleic acids: A-, B-, Z and Watson-Crick model of DNA. Structure of major species of RNA - mRNA, tRNA and rRNA. Nucleic acid chemistry - UV absorption, effect of acid and alkali on DNA. Other functions of nucleotides - source of energy, component of coenzymes, second messengers.

Unit 6 **6 Hours**

DNA replication, repair and recombination: Unit of replication, enzymes involved, replication origin and replication fork, fidelity of replication, Mechanism of bidirectional DNA replication, extrachromosomal replicons, DNA damage and repair mechanisms, Molecular basis of mutation, types of mutation.

Course outcome: By the end of this course, the student will be able to understand the physical, chemical and functional properties of macromolecules and fundamentals of biogenesis of macromolecules.

Reference Books:

1. Lehninger: Principles of Biochemistry (2013) 6th ed., Nelson, D.L. and Cox, M.M., W.H. Freeman and Company (New York), ISBN:13: 978-1-4641-0962-1 / ISBN:10:1-4292- 3414-8.
2. Textbook of Biochemistry with Clinical Correlations (2011) 7th ed., Devlin, T.M., John Wiley & Sons, Inc. (New York), ISBN:978-0-470-28173-4.
3. Watson, J.D., Baker, T.A., Bell, S.P., Gann, A., Levine, M., Losick, R. (2014) Molecular Biology of Gene. Cold Spring harbor, New York.
4. Voet D, Voet JG and Pratt CW (2011). Fundamentals of Biochemistry, 5 th Edition. John Wiley & Sons. New York.

Lab - Fundamentals of Biochemistry and Molecular Biology

Hours: 15

Course Objective:

Make the students understand the preparation of buffers and various assays.

1) Preparation of buffers.
2) Effect of pH on enzyme activity
3) Effect of Temperature on enzyme activity
4) Estimation of proteins using UV absorbance and Biuret method.
5) Assay of proteins using Lowry/Bradford method
6) Isolation of DNA from bacteria/ plant tissue and determination of DNA/ RNA concentration by A260nm

⟵ Back to Index

Functional Plant Pathology

Course Code : UBIN404

Minor Course

Credits : 4 **Hours : 45**

Course Objective:

To explore the core bioinformatic competencies and resources to understand and analyze the disease mechanism in plants.

Unit 1 (12 hrs)

Plant Diseases: Introduction to terminologies of plant diseases (pathogen, host, vector), types of plant diseases- bacterial (wilt, blight, rot, crown gall)-viral and viroid (mosaic- TMV & CMV)- fungal (blight, rust, smut, wilt) with their associated symptoms and examples.

Unit 2 (12 hrs)

Disease gene identification and Functional analysis: Exploring the online databases (Genome and RNA-seq data) for identifying the expression level disease genes in plants, Co-expression studies and Promoter analysis. Functional Analysis of the genes of interest and its network: Functional classification and pathway visualization.

Unit 3 (12 hrs)

Alternative splicing analysis of RNA-seq data: Introduction to alternative splicing (AS) analysis, Preparing the input data, Isoform-level analysis with Cuffdiff2, Event-level analysis with rMATS, Intron-level analysis with MntJULiP, Intron retention analysis with IRFinder, AS visualization with IGV and Jutils I & II.

Unit 4 (12 hrs)

Transcriptomic reconstruction: Introduction to transcriptomics with long RNA sequencing reads, Preparing the input sequence data, Aligning the long RNA sequencing reads, Differential gene and transcript expression (DESeq2), Differential splicing detection with long read tools (LIQA) to quantify isoform expression and detect differential alternative splicing (DAS) using Full-Length Alternative Isoform analysis of RNA (FLAIR tools- drimSeq & diffSplice), Building transcript models.

Course Outcome:

The learner will gain fundamental knowledge on the disease aspects of plant and will develop skills in relevant bioinformatics tools to identify and characterize the expression levels of the plant disease under specific condition to understand the disease mechanisms

Text Books:

1. Mehrotra, R.S. 1980. Plant Pathology. Tata McGraw-Hill Publishing Co. Ltd. New Delhi
2. Singh, R.S. 2002. Introduction to Principles of Plant Pathology. Oxford & IBH Publishing Co.Pvt. Ltd.,New Delhi.
3. Discovering Genomics, Proteomics & Bioinformatics A M Campbell & L J Heyer Pearson Education, 2007
4. Hartl & Ruvolo. Genetics: Analysis of Genes and Genomes. Jones & Bartlett Learning.
5. Mount. Bioinformatics sequence and Genome analysis. Cold Spring Harbor Laboratory Press
6. Deonier, Tavaré, & Waterman. Computational Genome Analysis: An Introduction. Springer
7. From Genes to Genomes by Dale and Schartz

Reference Books:

1. Pevsner, J. (2009). Bioinformatics and Functional Genomics. II Edition. John Wiley & Sons.
2. Bioinformatics: A Practical Guide to the Analysis of Genes and Proteins(2004). Andreas D. Baxevanis& B. F. Francis Ouellette. 3rd Edition. Wiley & Sons.

⟵ Back to Index

SEMESTER - V

Genomics & Proteomics

Course Code : UBIN501

Major Course

Credits : 4 **Hours : 60**

Course Objective:

The goal is to understand the expression pattern of genes within a genome and to explore their interaction for proteome that develops into a functional organism.

UNIT – I **12 Hours**

Introduction to genome and their projects: Overall structure and organisation of the nuclear and organellar genomes; Repetitive DNA, satellite bodies, interspersed repeated DNAs, Transposable elements, LINES, SINES, Alu family and their application in genome Mapping. Goals and outlook of various genome projects with their outcome- Human Genome Project, HapMap Project, the 1000 genome project and ENCODE; Ethical, Legal, Social Issues of HGP - IPR & Patents.

UNIT – II **12 Hours**

Genome Techniques: Methods for DNA/ RNA sequencing - Maxum Gilbert and Sanger sequencing; Next- Generation Sequencing (NGS)- Whole genome sequencing, Exome sequencing; Array techniques- DNA microarray, understanding of microarray data, normalizing microarray data; Introduction to Genome assembly and gene annotation, Gene networks.

UNIT – III **12 Hours**

Proteomic Technologies:

Protein separation technologies- 2D-PAGE for proteome analysis, Liquid chromatography. Protein detection – Protein identification and characterization: Mass spectrometry – EMI – MS, Tandem MS, TOF – MS, - HPLC: Peptide Mass Fingerprinting, calibration plots, proteomic libraries. Protein-protein and protein-DNA interactions techniques-Yeast two-hybridization and Phage display.

UNIT –IV **12 Hours**

Genomics Vs Proteomics: Fine structure of gene, split genes, pseudogenes, non-coding genes, overlapping genes and multi-gene families. Computational analysis tools for inferring gene function from relatedness to other genes, Finding Open Reading Frames, the prediction of genes, promoters, splice sites, regulatory regions; Genetic polymorphism- Mutation variability across the genome, Mutation detection-identification of SNPs, clinical and biomedical applications- metabolomics, lipidomics, metagenomics and pharmacogenomics, Ethical Consequences of Genomic Variations.

Course Outcome:

On the successful completion of the course, students will be able to learn the different methods of sequencing, microarrays, protein fingerprints and the role of bioinformatics tools applied to analyse and interpret the data in different cell types.

Text Books:

1. Introduction to Genomics Arthur M Lesk Oxford University Press 2007
2. Discovering Genomics, Proteomics & Bioinformatics A M Campbell & L J Heyer Pearson Education, 2007
3. Proteins and Proteomics Richard J Simpson IK International 2003
4. Genomics & Proteomics Sabesan Ane Books 2007
5. Purifying Proteins for Proteomics Richard J Simpson IK International 2004
6. Hartl & Ruvolo. Genetics: Analysis of Genes and Genomes. Jones & Bartlett Learning.
7. Mount. Bioinformatics sequence and Genome analysis. Cold Spring Harbor Laboratory Press
8. Deonier, Tavaré, & Waterman. Computational Genome Analysis: An Introduction. Springer

Reference Books:

1. Pevsner, J. (2009). Bioinformatics and Functional Genomics. II Edition. John Wiley & Sons.
2. Molecular Cloning: A Laboratory Manual (3rd Edition) Sambrook and Russell Vol. I to III, 1989.
3. Introduction to Proteomics by Daniel C. Liebler., Humana Press, 2002.
4. Discovering Genomics, Proteomics, & Bioinformatics (2003). Campbell &Heyer Pearson Education,
5. Bioinformatics, Methods of Biochemical Analysis (2001), Series Vol. 43, Baxevanis & Ouellette, John Wiley & Sons.
6. Bioinformatics: A Practical Guide to the Analysis of Genes and Proteins(2004). Andreas D. Baxevanis& B. F. Francis Ouellette. 3rd Edition. Wiley & Sons.

⬅ Back to Index

Bio-Sequence Analysis

Course Code : UBIN502

Major Course

Credits : 3+1P **Hours : 45**

Course Objective:

To train students to analyze biological sequences using computational tools and techniques to correlate with structure, function and evolutionary relationship.

Unit 1 **9 Lectures**

Basic biomolecular concepts: Protein and amino acid, DNA & RNA, Sequence, structure and function – relationship. **Generation of biological data**: Protein and DNA sequencing methods; Sanger method, Maxam and Gilbert method. **Protein Sequence submission**, DNA Sequence submission; Quality of Data, private and public data sources; **Different file formats** for bio-molecular sequences: GenBank, FASTA, GCG, MSF etc. Conversion of sequence formats – Seqret (EMBOSS). Protein sequence databases; Generation of large scale biological data - Genome Sequencing Method.

Unit 2 **9 Lectures**

Sequence Comparison – Basic concepts of sequence similarity, identity and homology, definitions of homologues, orthologues, paralogues and xenologues **Scoring matrices:** basic concept of a scoring matrix, Matrices for nucleic acid and proteins sequences, PAM and BLOSUM series, matrix derivation methods and principles. Repeats: Tandem and Interspersed repeat finding, Motifs, consensus, position weight matrices. **Pairwise sequence alignment** – Basic concepts of sequence alignment, gap penalties, Needleman and Wunsch, Smith and Waterman algorithms for pairwise alignments and application in Nucleic acid and protein sequences alignments.

Unit 3 **9 Lectures**

Basic Local Alignment Search Tool (BLAST) and Advanced Searching: Introduction, BLAST Search Steps, BLAST Algorithm Uses Local Alignment Search Strategy, BLAST Search Strategies, **BLAST Searching with Multidomain Protein**: HIV-1 Pol, Using Blast For Gene Discovery: Find-a-Gene, Using Blast For primer design, Pitfalls, **Advanced Database Searching**; Specialized BLAST Sites, Specialized BLAST-Related Algorithms, **Finding Distantly Related Proteins**: Position-Specific Iterated BLAST (PSI-BLAST) and DELTA-BLAST, **Profile Searches**: Hidden Markov Models, BLAST-Like Alignment Tools to Search Genomic DNA Rapidly, Aligning Next-Generation Sequence (NGS) Reads to a Reference Genome.

Unit 4 **9 Lectures**

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 application, concept of dendrogram and its interpretation, Use of HMM-based Algorithm for MSA (e.g. SAM method). **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. **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, PSI-BLAST

Unit 5 **9 Lectures**

Phylogenetic Analysis: Sequence Acquisition, Multiple Sequence Alignment, Models of DNA and Amino Acid Substitution, Tree-Building Methods - Distance-Based, Phylogenetic Inference: Maximum Parsimony, Model-Based Phylogenetic Inference: Maximum Likelihood, Tree Inference: Bayesian Methods; Evaluating Trees, Perspective, Pitfalls, Bootstrapping methods, use of tools such as Phylip, MEGA.

Course Outcome:

Students will learn to identify sequence patterns, compare sequences, predict protein structures, and understand evolutionary relationships; and apply these skills to uncover biological problems and make new discoveries

Reference Books:

1. Bioinformatics and Functional Genomics, Jonathan Pevsner, 3rd Edition, Wiley, 2022.
2. Bioinformatics: Sequence and Genome Analysis by Mount D., Cold Spring Harbor Laboratory Press, New York. 2004
3. Bioinformatics- a Practical Guide to the Analysis of Genes and Proteins by Baxevanis, A.D. and Francis Ouellellette, B.F., Wiley India Pvt Ltd. 2009

LAB - BIOSEQUENCE ANALYSIS

Hours: 15

Exercices:

- Sequence Analysis Packages: EMBOSS, NCBI ToolKit, SMS
- Pair wise alignment:
  - Search tools against Databases:
    - BLAST
    - FASTA
- Multiple sequence alignment:
  - Clustal
  - Dialign
  - Multalign
- Sequence patterns and profiles:
  - generation of sequence profiles
    - PSI-BLAST
  - derivation of and searching sequence patterns:
    - MEME/MAST
    - PHI-BLAST
    - SCanProsite
    - PRATT
- Protein motif and domain analysis:
  - MEME/MAST
  - InterproScan
  - ProSite
  - ProDom
  - Pfam
- Phylogenetic analysis – MEGA, PHYLIP

⟵ Back to Index

Programming languages C & C++

Course Code : UBIN503

Major Course

Credits : 3+1P **Hours : 45**

Course Objective:

To train the students in writing programs in C- language and to introduce them to the concepts of object-oriented language through C++.

Unit 1 **9**
Lectures

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

Unit 2 **9**
Lectures

Controls and loops: Conditional and Unconditional Control Statements – Looping Statements: while, do-while, for – Nested loops – Arrays.

Unit 3 **9**
Lectures Procedural Concept: Structured Programming – Built-in library functions – User defined functions – Pointer concept – Passing a pointer to a function – Structure – Union – File handling: Read and Write numerical and character data from/to a file.

Unit 4 **9**
Lectures

String Handling & Sorting: String declaration – String library functions - String Manipulation - Sorting: Bubble sort, Selection sort, Insertion sort – Searching algorithms : Linear search and Binary search.

Unit 5 **9**
Lectures

Object Oriented Programming and C++: Object Oriented Concept: Encapsulation, Inheritance, Polymorphism – Different forms of Constructors – Destructors – Abstract class – Virtual function. Understanding of C++ as object oriented language.

Course Outcome:

On successful completion of the course students will get familiarize with coding for bioinformatics problems in C/C++ language and with the object-oriented programming approach.

Reference Books:

1. Programming in ANSI C by E. Balagurusamy. Tata McGrawHill Publishing Company Limited. 2007
2. Object Oriented Programming using C++ (4th Edition) by Lafore, R. Galgotia Publishers. 2008
3. Sams Teach Yourself C++ in 24 hours (5th edition) by Jesse Liberty and Rogers Candenhead, Pearson Education Inc., 2012.
4. Head First C by David Griffiths and Dawn Griffiths, O’Reilly. 2013.

Lab - Programming languages C & C++

Hours : 15

Programming in C

1. Simple Input and Output statements.
2. Working with if, if else and switch constructs.
3. Working with arrays and strings.
4. Loops and nested loops.
5. Working with user defined functions.
6. Working with pointers.
7. Working with structures and Unions.
8. File handling with numerical and character data.

Programming in C++

9. Creation of a simple class and working with its objects.
10. Implementing the inheritance in C++.
11. Working with function overloading.
12. Working with operator overloading.

⬅ Back to Index

Next Generation sequencing

Course Code : UBIN504

Major Course

Credits : 4 **Hours : 60**

Course Objective:

The main goal of this course is to help students in learning the various next-generation sequencing platforms and computational methods involved in the reference-based and de novo assembly of short reads for genome mapping and gene expression analysis, especially in the R environment

Unit 1.

DNA Sequencing Technologies: Next generation sequencing- Cyclic reversible Termination – Pyrosequencing – Sequencing by ligation – Genome sequencing by Measuring pH- single Molecule sequencing with Long Read Lengths- Recent Sequencing Technique – Nanopore DNA Sequencing – MinION.

Unit 2

Experimental Design and sample preparation – Generating sequence data to File Formats – Quality assessment of Data- Genome Assembly – Competitions and critical evaluations of the performance of Genome assemblers – End of Assembly – Sequence alignment – Genome analysis Toolkit Workflow – Calculating Read Depth – Variant Calling – visualizing and Tabulating Next Generation sequence Data – Interpreting the biological significance of Variants.

Unit 3

Bioinformatic approaches to RNA – Noncoding RNA – Transfer RNA – Ribosomal RNA – Small Nuclear RNA – small Nucleolar RNA – MicroRNA – Short interfering RNA – Long Noncoding RNA – Circular RNA- Messenger RNA – Microarrays and RNA – Seq- Genome wide Measurement of Gene expression- RNA Analysis. Understanding the genetic Basis of variation in Gene Expression through combined RNA-seq and DNA-seq.

Unit -4

Genome assembly algorithms: Alignment of short-reads to reference genome using spaced seed (ELAND, SOAP), FM-index (Bowtie, BWA, SOAP2), and suffix tree (MUMmer). Sequence Alignment formats: Sequence Alignment/Map (SAM) format, Binary Alignment/Map (BAM) format, Tools for conversion (SAMtools), Alignment viewers (IGV, MGAviewer). De-novo assemblies: Overlap-layout-consensus (OLC) approach (Arachne, Phusion), de Bruijn and Euler path approach (Euler, SOAPdenovo), string graph assembler (SGA). Scaffolding: Supercontig, contig orientation, contig ordering, contig distancing and gap closing.

Unit – 5

Big Data and R language in NGS analysis: Elements of Big data and R, Introduction to Bioconductor, Reading of RNA-seq data (ShortRead, Rsamtools, GenomicRanges), annotation (biomaRt, genomeIntervals), reads coverage and assign counts (IRanges, GenomicFeatures), differential expression (DESeq). Biological applications of NGS: Whole-genome sequencing, Exome sequencing, Transcriptome sequencing, DNA-Protein Interactions (CHIP-Seq), Epigenomics and DNA methylation analysis, Metagenome analysis.

Text Books:

1. Bioinformatics and Functional Genomics by Jonathan Pevsner – An Indian Adaptation, Third Edition – 2022.
2. Next-generation DNA sequencing Informatics by Stuart M. Brown, Cold Spring Harbor Laboratory, 2013.
3. Big Data Analysis for Bioinformatics and Biomedical Discoveries by Shui Quing Ye. Chapman and Hall/CRC Press, 2016.

Reference Books:

1. Next generation sequencing: Translation to Clinical Diagnostics by Wong Lee-Jun C. (ed.), Springer, 2013.
2. Next-generation genome sequencing: Towards Personalized Medicine by Michal Janitz, 4

⬅ Back to Index

Microbial Genome Informatics

Course Code : UBIN505

Minor Course

Credits : 4 **Hours : 60**

Course Objective:

This course enable the students to Learn about key genomic technologies, including Next-Generation Sequencing (NGS) and gene expression analysis and helps to Learn bioinformatics tools and resources for microbial genome analysis and data visualization finally helps them to Explore how microbial genomics helps in public health, disease control, and antimicrobial resistance research

Unit I: Introduction to Microbial Genomics (8 hours)

Overview of microbial genomics and introduction to omics sciences. Importance of genomics in studying microorganisms (bacteria, archaea, fungi, and viruses). Genome organization: Microbial genome structure (circular vs. linear genomes), genome size, gene content, and microbial diversity. Organization and complexity of microbial genomes. Prokaryotic vs. eukaryotic genomes. Importance of genome sequencing in microbial research. Current trends and future directions in microbial genomics research.

Unit II: Microbial Genomic Databases and Resources (10 hours)

Genome databases: GenomeIndia Project (GIP), goals, and impact. Introduction to NCBI-SRA and its microbial genomics resources. Annotation of microbial genomes using databases such as KEGG, UniProt, and NCBI. PATRIC: A comprehensive database for pathogenic microbial genomics. Virulence and antimicrobial resistance databases: ResFinder, CARD, and VFDB. Metagenomic databases: Kraken, KEGG Microbe database: Ensemble genome browsing and annotation for microorganisms.

Unit III: Introduction to Genomic Technologies (10 hours)

Fundamental concepts of genomics, epigenomics, transcriptomics, and metabolomics. Evolution of genomic technologies. Key milestones in microbial genome sequencing (e.g., Human Genome Project, list of microbial genomes sequenced). Genome sequencing and a basic understanding of Next-Generation Sequencing (NGS) technologies. Overview of metagenomic sequencing. Gene expression analysis: Identification of unknown genes. Emerging technologies: Highlights of CRISPR/Cas systems and gene editing in micro-organisms, and the role of AI in microbial genomics.

Unit IV: Bioinformatics Approaches for Microbial Genomics (10 hours)

Microbial genome data analysis. Genomics file formats: FASTA, FASTQ, SAM, BAM, VCF, and GFF/GTF files. Quality control of sequencing data (using tools like FASTQC). Microbial genome assembly and annotation. Galaxy: An open-source platform for genomic analysis. Bacterial genomics pipelines (e.g., GARLIC, Prokka, Roary). Antimicrobial resistance and virulence factor prediction.

Comparative genomics and genome alignment. Visualization of microbial genomic data using tools like Geneious and iTOL. Case studies in microbial pathogenesis, evolution, and resistance.

Unit V: Applications of Bioinformatics in Microbial Genomics (10 hours)

Overview of microbial genome data science. Pathogenic gene detection: Genomes of pathogenic microbes (e.g., *E. coli*, *Salmonella*, or *Mycobacterium tuberculosis*) and their resistance gene prediction. Global Antimicrobial Resistance and Use Surveillance System (GLASS). Genomic data analysis for outbreak surveillance and public health monitoring. Metagenomics and microbiome research: The contribution of microbial communities to health, disease, and environmental processes. Current trends and future directions in microbial genomics.

Course Outcome:

The students will be able to understand Genome of an organism and its functional annotations, and demonstrate fundamental knowledge of emerging genome sequencing technologies and know the tools applied to microbial genome data analysis using open-source software and demonstrate skills for genomic visualization and understanding the application of bioinformatics in microbial genomic science

Reference Books:

1. Microbial genomes By Claire M. Fraser, Timothy D. Read, Karen E. Nelson Humana Press
2. Bacterial genomes: physical structure and analysis. By Frans J. de Bruijn, James R. Lupski, George M. Weinstock.
3. Cheba B A. Review on Microbial Bioinformatics: Novel and Promoting Trend for Microbiomics Research and Applications. Cham: Springer International Publishing, 2021: 718-729.
4. David W, 2005, Bio-informatics; sequence and Genome Analysis, 2ndEdition By Mount CB Spublishers.
5. Essentials of Genomics and Bioinformatics by C. W. Sensen (John Wiley & Sons Inc.)
6. Genome analysis: a laboratory manual, Volume 2 edited by Eric D. Green
7. Next-generation genome sequencing: towards personalized medicine, By Michal Janitz
8. Hiraoka S, Yang C, Iwasaki W. Metagenomics and bioinformatics in microbial ecology: current status and beyond. Microbes and environments, 2016, 31(3): 204-212.

⟵ Back to Index

SEMESTER - VI

Algorithms

Course Code : UBIN601

Major Course

Credits : 4 **Hours : 60**

Course Objective:

To impart knowledge about the importance of data structures in programming and to familiarize basic searching and sorting algorithms.

Unit 1 **10 lectures**

Computing Algorithms: Algorithms in Computing, analyzing algorithms, Designing algorithms, Asymptotic notation, Standard notations, Big 'O' notations, Time and space complexity of algorithms and common functions.

Unit 2 **10 lectures**

Sorting, Searching & Strings Matching: Sorting: Bubble sort, Insertion sort, Selection sort, Merge Sort, Quick Sort, External sort: K-way merge sort, balanced merge sort, Searching: Binary Search, String Matching: Naïve algorithm.

Unit 3 **9 Lectures**

Graphs: Representation of Graphs, Breadth First Search, Depth First Search, Connected Components, Minimum Spanning Tree, Single-Source Shortest Path: Dijkstra's Algorithm, All-Pairs Shortest Paths.

Unit 4 **9 Lectures**

Trees: Forests, DAGs, Ancestors, and Descendants, Binary Search Trees, Querying a Binary search tree, Insertion and Deletion, Tree Traversals, AVL-Trees, Rotations, Insertion, Deletion

Unit 5 **10 Lectures**

Algorithm Design and Analysis: The substitution method, the iteration method, Divide and Conquer, Greedy Algorithms, Dynamic Programming: Traveling Sales Person Problem Backtracking Algorithms: 8-queens Problem.

Course Outcome:

To comprehend the basics of algorithms and understand the operations performed using arrays. To realize the properties of tree data structure and its importance in searching large database. To understand graph data structure and its applications.

Text Books:

1. Introduction to Algorithms by Cormen et al., 4th Ed., 2022
2. Algorithms: Design and Analysis by Collins B., 1st Ed., 2023

Reference Books:

1. Algorithms by Sedgewick & Wayne, 4th Ed.2020
2. Introduction to Algorithms (4th Ed.) by T .H. Cormen, C. E. Leiserson, R .L. Rivest., The MIT Press. 2022

⟵ Back to Index

Systems Biology

Course Code : UBIN602

Major Course

Credits : 3+1P **Hours : 45**

Course Objective:

The main goal of this course is to help students in learning the basic concepts and computational methods involved in the computational modelling of the biological systems.

Unit 1 **5 lectures**

Introduction & Biological Networks – Systems Biology: Emergent property, Applications in health and diseases. Microarrays and its applications in systems biology. Connectivity maps (CMap) and Library of Integrated Network-based Cellular Signatures (LINCS) -definition and its uses. Biological Networks: Degree distribution, Clustering coefficient, Random networks, Scale-free networks, small-world effect.

Unit 2 **5 lectures**

Simulation of pathways – Metabolic network, Metabolic reconstruction, Flux Balance Analysis (FBA): Translating biochemical networks into linear algebra, Stoichiometric matrix, Elementary mode, Extreme pathways, Objective function, Optimization using linear programming. Genome-scale cellular models: Virtual Erythrocytes, Global human metabolic model (Recon 3D).

Unit 3 **5 lectures**

Signalling & Experimental methods in systems biology – slow and auto–regulation The coherent FFL and incoherent FFL, single-input module (SIM): LIFO and FIFO, DOR, signalling networks and neuronal circuits.

Robustness and optimality in Biological complex systems – Biological Robustness: System control, modularity, decoupling. Optimal design of gene circuits I- cost and benefit: gene circuits II- selection of regulation. Stochasticity in gene expression.

Unit 4 **4 lectures**

Databases and softwares for Systems Biology – Introduction- databases: KEGG, EMP, MetaCyc. Expression databases and other databases related to systems biology. Cytoscape, visANT & CellDesigner.

Unit 5 **5 lectures**

Synthetic Biology – Introduction, definition and Basics, Synthetic Oligonucleotide/DNA-based, RNA-based, Peptide-based Technologies and Applications, Technologies and Applications of Directed Evolution and Microbial Engineering, Potential Hazards of Synthetic Biology, iGEM

Course Outcome:

The student will have a system-level understanding of the biological systems. He/she will be able to develop and analyse the properties of in silico models of gene-gene interactions and protein-protein interactions

Text Books:

1. Introduction to Systems Biology: Design Principles of Biological Circuits by Uri Alon, Chapman & Hall/CRC, 2007.
2. Synthetic Biology: A Primer by P.S. Freemont & R.I. Kitney, Imperial College Press, 2012.

Reference Books:

1. Introduction to Systems Biology, S. Choi, Humana Press, 2007.
2. Linked – The New Science of Networks, Albert-László Barabási, Perseus Publishing, 2002.
3. Networks – an Introduction, Mark Newman, Oxford University Press, 2010.

Lab - Systems Biology

Hours: 15

1. Introduction to Transcriptome databases: SRA, GEO, ArrayExpress, TCGA.
2. Introduction to Metabolome databases: MetaboLights, Human Metabolome Database, KEGG.
3. Introduction to Visualisation tools: visANT, CellDesigner, Cytoscape.
4. Construction of gene-coexpression network using minet package in R.
5. Study of network topology using igraph package in R.
6. Visualizations of protein-protein interaction network using different algorithms in igraph package in R and Cytoscape.
7. Construction of adjacency matrix from the protein-protein interactions in R environment.
8. Kinetic modeling of *E. coli* using JSim tool.
9. Study of Flux Balance Analysis (FBA) in *E. coli* by simulating metabolic flux using the Genome-Scale Metabolic Model.
10. Reconstruction of the multi-omics network in the R environment.

⬅ Back to Index

Structural Biology

Course Code : UBIN603

Major Course

Credits : 3+1P **Hours : 45**

Course Objective:

The main objective of this course is to make the students to thorough understanding of structural biology of biological macromolecules.

Unit 1 **9 Lectures**

Macromolecules: DNA and RNA- types of base pairing – Watson-Crick and Hoogstein; types of double helices A, B, Z and their geometrical as well as structural features; structural and geometrical parameters of each form and their comparison; various types of interactions of DNA with proteins, small molecules. RNA secondary and tertiary structures, t-RNA tertiary structure. **Proteins**: Principles of protein structure; anatomy of proteins – Hierarchical organization of protein structure – Primary. Secondary, Super secondary, Tertiary and Quaternary structure; Ramachandran Map.

Unit 2 **9 Lectures**

Protein structure determination techniques: X-ray crystallography Electromagnetic radiation, origin of X-rays, diffraction basics, Bragg's Law , phase problem, different methods for solving phase problem ,**NMR techniques**, Nuclear Magnetic Resonance: Chemical Shift, Coupling constant, spin-spin relaxation, spin-lattice relaxation, COSY, NOESY and NOE, **cryo-electron microscopy.**

Unit 3 **9 Lectures**

Structural modeling- Homology modeling, Template recognition and initial alignment, Alignment correction, Backbone generation, Loop modelling, Side-chain modelling, Model optimization, Model validation, Threading, ab initio method, Protein-protein interaction ,Protein-ligand interaction, Protein-DNA interaction, Prediction of binding cavities.

Unit 4 **9 Lectures**

Structure Prediction Strategies – Secondary structure prediction: Alpha Fold, Algorithms viz. Chou Fasman, GOR methods; analysis of results and measuring the accuracy of predictions using Q3, Segment overlap, Mathew's correlation coefficient Identification/assignment of secondary structural elements from the knowledge of 3-D structure of macromolecule using DSSP and STRIDE methods.

Unit 5 **9 Lectures**

Classification and comparison of protein 3D structures – Purpose of 3-D structure comparison and concepts; Algorithms such as FSSP, CE, VAST and DALI, Fold Classes. Databases of structure-based classification: CATH and SCOP. Structures of oligomeric proteins and study of interaction interfaces.

Course Outcome:

Students gained the knowledge of structural biology of biological macromolecules.

Text Books:

1. Introduction to protein structure. By C Branden and J Tooze, New York. 2 editions (January 3, 1999).
2. Textbook of Structural Biology, by Anders Liljas, Lars Liljas, Jure Piskur, Göran Lindblom, Poul Nissen, and Morten Kjeldgaard, 2nd Edition edition (January 4, 2017.
3. Protein Structures and molecular properties: By T E Creighton, W H Freeman, New York. Second edition (August 15, 1992).

Reference Books:

1. Crystals, X-rays and Proteins: Comprehensive Protein Crystallography 1st Edition, Kindle Edition by Dennis Sherwood and Jon Cooper, 2011
2. Fundamentals of Crystallography by Giacovazzo Carmelo, Third Edition 2011.
3. Principles of Protein X-ray Crystallography by Drenth Jan, Third edition, 2007
4. Introduction to Crystallography by Donald E. Sands, 1st Edition, 1994

LAB - STRUCTURAL BIOLOGY

Hours: 15

COURSE OBJECTIVES: The main objective of the course is to train the students how to determine the small molecule structure through X-ray crystallography

1. Advanced Visualization Software and 3D representations.
2. Small Molecule Structure determination
    a. Structure Solution: SHELXS
    b. Structure Refinement: SHELXL
3. Thermal Ellipsoid Plot:
    a. ORTEP
4. Structure analysis
    a. PARST
    b. Platon
    c. Mercury
5. Comparative structure alignment and analysis using three dimensional structures of protein and nucleic acids.
    a. DALI
6. Structure Validation
    a. Procheck, WHATIF, VERIFY 3D
7. Exploration of CCP4 for macromolecular crystallography

COURSE OUTCOME: Students gained the knowledge of how to solve the small molecule structure through X-ray crystallography method.

⬅ Back to Index

JAVA programming Language

Course Code : UBIN604

Major Course

Credits : 3+1P **Hours : 45**

Course Objective:

Students will

- Learn the fundamentals of Java programming and object-oriented concepts.
- Build, compile, and execute Java applications, including bioinformatics-related programs.
- Gain skills in using Java Collections, generics, and file handling for managing data.
- Develop problem-solving abilities and apply logical thinking to design and implement Java solutions.

The course also aims to develop strong problem-solving and logical thinking skills.

UNIT 1: Introduction to Java (7 hours)

Basics of Java and OOP concepts, Features of Java, Java Development Kit (JDK) and Java Runtime Environment (JRE), Writing, compiling, and executing a Java program, Java data types, variables, and operators, Control statements: if, switch, loops

UNIT 2: Classes, Objects, and Methods (7 hours)

Defining classes and objects, Class Fundamentals, Method declaration and calling, Assigning Object Reference Variables, Constructors and overloading, this keyword, Static members and blocks.

UNIT 3: Inheritance and Polymorphism (7 hours)

Inheritance Basics, Types of inheritance in Java, Method overriding and overloading, super keyword and constructor chaining, Abstract classes and interfaces. Using final with Inheritance.

UNIT 4: Arrays, Strings, and Exception Handling (7 hours)

Single and multidimensional arrays, String handling: String and StringBuilder classes, Command-line arguments, Exception handling: try, catch, finally, throw, throws, File handling basics

UNIT 5: Java Collections and Bioinformatics Applications (8 hours)

Introduction to Collection Framework and core interfaces: Collection, List, Set, Map. Common implementations: ArrayList, LinkedList, HashSet, HashMap and their basic usage. Simple iteration using loops and basic operations like add, remove, contains, and size. Generics in Java, Java I/O and file reading, Simple GUI using Swing. Biojava and applications

Course Outcome:

Upon successful completion of this course, students will be able to

- Understand the fundamental concepts of Java programming, including OOP principles, control structures, and data handling.
- Apply core Java features such as classes, inheritance, exception handling, and collections to develop basic Java applications.
- Analyze the use of Java Collections, generics, and file I/O to efficiently manage data and build modular programs.
- Create simple code using Java, integrating GUI elements and biojava libraries to solve domain-specific problems.

Reference Books:

1. 1The Complete Reference - Java, Herbert Schildt, 10th Edition, 2017, McGraw Hill Education Publications, ISBN-10: 9789387432291, ISBN-13: 978-9387432291
2. 2Herbert Schildt, 2018, JAVA: The complete reference, 11th Edition, TATA MCGRAW
3. HILL Edition, Kindle Edition.
4. 3Introduction to Java Programming, Y Daniel Liang, 10th Edition, 2014, Comprehensive Version Pearson education, ISBN 10: 0-13-376131-2, ISBN 13: 978-0-13-376131-3
5. 4Core Java – Vol 1, Cay S. Horstmann, 10th Edition, 2016, Pearson Education, ISBN-10: 9332582718, ISBN-13: 978-9332582712 4 Object-Oriented Analysis and Design With applications, Grady Booch, Robert A Maksimchu
6. Andreas Prlic, Andrew Yates, Spencer E. Bliven, et al., BioJava: on open-source framework for bioinformatics in 2012. Bioinformatics. 28(20): 2693-2695. https://www.biojava.org

Lab - JAVA programming

Hours : 15

This lab components enables students to set up Java programming environments using popular IDEs and the command line. They will practice writing programs with variables, operators, decision-making statements, and loops. Students will gain hands-on experience creating classes, objects, and working with abstract classes and interfaces. The course also covers exception handling, file reading/writing, and using core collections like ArrayList and HashMap. Optionally, students can explore building simple GUIs with Swing to enhance their Java skills.

Practical Components

1. Java program setup in IDE (Eclipse/NetBeans, etc)
2. Write programs using variables, operators
3. decision-making statements like if-else and switch-case.
4. Demonstrate loops: for, while, and do-while.
5. Compile and run Java programs from the command line.
6. Create classes and objects.
7. Abstract classes and interfaces.
8. Work with single and multidimensional arrays.
9. Handle exceptions with try-catch blocks.
10. Read and write files
11. Use ArrayList, LinkedList, HashSet, and HashMap.
12. Create a simple GUI with Swing.

⟵ Back to Index

Spectroscopy in Biology

Course Code : UBIN605

Minor Course

Credits : 3 + 1 (Seminar) **Hours : 45**

Course Objective:

A course that is meant for teaching various spectroscopic tools available to extract useful information in respect of biological samples.

Unit 1 **9 Lectures**

UV- Visible spectroscopy: Absorption laws - calculations involving Beer - lambert's law - instrumentation – photo colorimeter and spectrophotometer - block diagrams with description of components - theory - types of electronic transitions - chromophore and auxochromes - absorption bands and intensity - factors governing absorption maximum and intensity Fluorescence spectroscopy.

Unit 2 **9 Lectures**

Infrared spectroscopy: Principle - types of stretching and bending vibrations - vibrational frequencies - instrumentation - block diagram - source - monochromator - cell sampling techniques - detector and recorders - identification of organic molecules from characteristic absorption bands. FTIR and its advantages.

Unit 3 **9 lectures**

Raman spectroscopy: Raleigh and Raman scattering - Stoke's and anti Stokes lines - instrumentation - block diagram - differences between IR and Raman spectroscopy - mutual exclusion principle - applications - structural diagnosis.

Unit 4 **9 lectures**

Nuclear Magnetic Resonance Spectroscopy: Nuclear spin magnetic moment, Interaction of nuclear magnet with external magnetic field, NMR spectrometer, relaxation and dynamic processes, chemical shift, coupling constants in ^{1}H, ^{13}C and ^{31}P NMR spectra; application of NMR spectroscopy for structure elucidation of simple biomolecules. Heteronuclear NMR experiments.

Unit 5 **9 lectures**

Electron Spin Resonance Spectroscopy: Electron spin and Magnetic moment, Resonance condition in ESR and significance of 'g' value, applications of ESR of (i) transition metal ions (ii) free radicals and (iii) spin labels in Biology. Introduction to Pulsed EPR.

Course Outcome:

The students will be introduced to the working principles and applications of various common spectroscopic techniques.

Reference Books:

1. Fundamentals of molecular spectroscopy by C. N. Banwell., McGraw-Hill.1983
2. Modern Spectroscopy (Fourth Edition) by J. Michael Hollas (2004) John Wiley & Sons.
3. Molecular spectroscopy by I. N. Levins, Wiley Interscience. 1975
4. Electron Spin Resonance : Elementary Theory and Practical Applications by John Wertz Springer (1986)
5. Nuclear Magnetic Resonance Ed. Hans-Ferdinand Linskens and F John Springer (1986)
6. Introduction to molecular spectroscopy by G. M. Barrow., McGraw-Hill.1962

Seminar (Spectroscopy in Biology)

Hrs : 15

The students are to give a seminar for 1 credit. A research paper (published in last 5 years) that used any of the spectroscopic tools discussed here is to be chosen by the student for the seminar.

⬅ Back to Index

Immuno-informatics

Course Code : UBIN606

Minor Course

Credits : 4 **Hours : 45**

Course Objective:

The main objectives of this course are to:

1. Make the students understand the immune system, its components and their functions.
2. Provide fundamental knowledge in Immunoinformatics databases and tools.
3. Learn the informatics-based approaches for prediction of Epitopes, design of vaccines and immuno-diagnostic tools.

Unit- **16 Hours**

Introduction to Immunology: Innate and Adaptive Immunity, Anatomical and Physiological Barriers, Inflammation, Soluble molecules. Active and passive immunity, natural and artificial immunity Innate Immune response and their recognition structures, Humoral and Cell mediated responses of immune system.

Unit-2 **7 Hours**

Immune system: Cells of Immune system**:** Lymphocytes, Mononuclear phagocytes, Antigen Presenting cells, polymorphs, Natural Killer cells, Granulocytes, Mast cells, Dendritic cells, Cluster designation (CD) and antigen specific receptors. **Organs of Immune system:** Primary lymphoid organs (Bone marrow and Thymus). Secondary Lymphoid organs (Lymph node, spleen and MALT).

Unit- 3 **8 Hours**

Immunoglobulins: Structure and function **-** Clonal selection theory – Ig Classes and subclasses, DR and LDR regions. Major Histochemical molecules/peptide complexes- Structure and Function. **Antigen and antibody reaction/interaction:** Precipitation, Haemagglutination, direct and indirect immunofluorescence. Complement system- activation, pathways and biological effects.

Unit-4 **8 Hours**

Antibody production Technology: Hybridoma technology for monoclonal antibody production. Chimeric antibodies, antibody engineering via computational tools, large scale manufacture of antibodies. DNA vaccine, Plant and protein based recombinant antigens as vaccines. Reverse vaccinology, Toxoid as vaccine, Conjugate vaccine.

Unit-5 **8 Hours**

Computational Immunology: Databases in Immunology, dbMHC-MHC database at NCBI. Prediction of immunogenicity - Pipeline & workflows, **Structure-based Vaccine design** - Web-Based Tools for Vaccine Design. The IMGT® Immunoinformatics page. Databases associated with Immunoglobulins (or Antibodies) (IG).

Unit-6 **8 Hours**

MHC- PREDICTION: Prediction of Cytotoxic T Cell (MHC Class I) Epitopes- Antigen Processing in the MHC Class I Pathway. MHC-II PREDICTION: Prediction of Helper T Cell (MHC Class II) Epitopes- Processing of MHC Class II Epitopes. T-cell epitope databases, B-cell epitope databases, SYFPEITHI MHC-presented epitopes. HLA Nomenclature and the IMGT/HLA Sequence Database.

Course Outcome:

At the end of the course the student will be able to: • Understand the computational aspects of immunology • Tools and databases related to immunoinformatics studies. Application of Bioinformatics tools in immunology.

Reference Books:

1. Immunological Bioinformatics” by Ole Lund Darren Flower (2006) MIT press, Springer.
2. Immunoinformatics: Predicting Immunogenicity in Silico” by Darren R Flower (2007) Humana Press
3. Computational Immunology: Basics” by Shyamasree Ghosh (2020) CRC Press
4. Kuby Immunology “(2006|) W. H. Freeman & Company

⟵ Back to Index

SEMESTER - VII

Molecular modelling, Mechanics and simulations

Course Code : UBIN701

Major Course

Credits : 3+1P **Hours : 45**

Course Objective:

Aims to provide knowledge about the concepts of modeling, mechanics and dynamics of biomolecular systems by introducing the essential concepts from Physics, Chemistry and Biology.

Unit 1 **9 Lectures**

BASIC CONCEPTS: Introduction, Coordinate System, Structure Drawing, 3D Effects, Optical Activity, Modelling and 3D-Displays.

Unit 2 **9 Lectures**

MOLECULAR MODELS: Molecular representation in Balls on Springs model, Vibrational Motion, Newtonian Law, A Simple Diatomic models, Harmonic Potentials, Morse Potential, More Advanced molecular Potentials.

Unit 3 **9 Lectures**

FORCE FIELDS: Introduction**,** Bonded Interactions (Bond-stretching, Bond-bending, Dihedral motions, inversion), Non-bonded interactions, Coulomb interactions, Modelling the Solvent, United atoms, Cut-offs, Example: AMBER Force Fields.

Unit 4 **9 Lectures**

POTENTIAL ENERGY SURFACE AND OPTIMIZATION: Introduction, PES Features (Multiple Minima, Saddle Points), Characterization, Finding Minima, Multivariate Grid Search, Derivative Methods, First-Order optimization Methods:- Steepest descent, Conjugate gradients.

Unit 5 **9 Lectures**

MOLECULAR DYNAMICS SIMULATION AND ANALYSIS: Introduction, Newtonian mechanics, Integrators- Leapfrog and Verlet algorithm, solvation and models, Periodic boundary conditions, Temperature and pressure control in molecular dynamics simulations**. Molecular Dynamics Analysis** (stability of the system, RMSD, RMSF, radius of gyration, Hydrogen bonds, Contact maps, etc).

Course Outcome:

Students could understand the theories in macromolecular simulations and perform research work in the area of computational structural biology.

Reference Books:

1. Molecular Modelling for Beginners by Alan Hinchliffe, 2003, John Wiley & Sons Ltd, England.
2. Computational Chemistry and Molecular Modeling-Principles and Applications by Ramachandran, Deepa and Namboori, 2008, Springer-Verlag.
3. Molecular Modeling Principles and Applications (2nd Ed.) by Andrew R. Leach., Prentice Hall, USA. 2001.

Reference Books:

1. Molecular Modelling for Beginners, (2nd Edition) by Alan Hinchliffe., John Wiley & Sons Ltd.2008.
2. Molecular Modeling and Simulation – An Interdisciplinary Guide by Tamar Schlick., Springer-Verlag 2000.
3. Computational Medicinal Chemistry for Drug Discovery, edited by Patrick Bultinck., Hans De Winter, Wilfried Langenaeker, Jan P. Tollenare, CRC press, 2003.
4. The art of molecular dynamics simulation, second edition by D. C. Rapaport, Cambridge University Press, 2004.

LAB - MOLECULAR MODELING, MECHANICS AND SIMULATIONS

Hours: 15

COURSE OBJECTIVES: Aims to train students with technical skills to perform molecular dynamics simulations

Exercises

- Molecular Visualization: Pymol and Chimera
  - Pdb file format and Parsing
  - Visualizing a molecule in different representations
  - Identifying interacting residues
- Homology modeling of protein 3Dstructure (Repeat)
  - Model building using Modeller
  - Model validation
- Geometry Optimization using SwissPdb Viewer
  - Energy Minimization of protein molecule
  - Determining Maxima and Minima energy points
- Molecular Dynamics Simulation
  - Protein dynamics using Gromacs
  - Protein-ligand complex MD simulation
    - Binding Site Identification (Structure Biology)

      Tools:- Cast-P, POCASA, 3D ligand site

COURSE OUTCOME: Students will be skilled to perform macromolecular simulations which will be useful for their research/project work

⬅ Back to Index

Data science & Machine Learning

Course Code : UBIN702

Major Course

Credits : 3+1P **Hours : 45**

Course Objective:

- To learn the basic concepts, terminology, models and methods in the data Science
- To apply data transformations such as aggregation and filtering for visualization
- To identify opportunities for application of data visualization in various domains

Unit 1 **10 Lectures**

Introduction to Data Science& ML

Data Science & Machine Learning: Definition & its Application, Different Types of Data - Structured, Unstructured, & Semi-Structured, insight into the data science process,: KDD Process, Kinds of Data can be Mined, Kind of Patterns can be Mined, Technologies, Applications – Issues in Data Mining, Data Objects & Attribute Types, Data Pipelines, Data Quality, Machine Learning Definition and Relation with Data Science..

Unit-2 **10 Lectures**

Data management and Analysis: Data collection and management: Introduction, Sources of data, Data collection and APIs, Exploring and fixing data, Data storage and management, using multiple data sources. **Data analysis**: Introduction, Terminology and concepts, Introduction to statistics, Central tendencies and distributions, Variance, Distribution properties and arithmetic, Samples/CLT, Basic machine learning algorithms, Linear regression, SVM, Naive Bayes

Unit-3

Basics machine learning algorithms **10 Lectures**

Supervised Learning**:** Regression & Classification: Definition, Types, and Applications, Linear Regression, Logistic Regression, Decision Trees, Support Vector Machines (SVM), Naive Bayes. ***Unsupervised Learning*****:** Clustering: Definition and Applications, Partitioning Methods: k-means, K-Medoids and Dimensionality Reduction: Principal Component Analysis.

Unit-4 **9 Lectures**

Model Evaluation: Metrics: Residual Standard Error, R^2, Mean Squared Error, Akaike Information Criterion, Bayesian Information Criterion, Confusion Matrix, Accuracy, Precision, Recall, F1-score, Cross-Validation, Grid Search and Randomized Search, Under fitting, Over fitting, Model Selection.

Unit-5 **9 Lectures**

Data visualization

Data visualization: Introduction, Types of data visualization, Data for visualization: Data types, Data encodings, Retinal variables, mapping variables to encodings, Visual encodings, various visualization techniques, application development methods of used in data science.

Course Outcome:

On the successful completion of the course, student will be able to:

- Explore the fundamental concepts of data science.
- Understand data analysis techniques for applications handling large data.
- Understand various machine learning algorithms used in data science process.

Reference Books:

1. Michael Berthold, David J. Hand, Intelligent Data Analysis, Springer, 200
2. Anand Rajaraman and Jeffrey David Ullman, Mining of Massive Datasets, Cambridge University Press, 2012
3. Jan Holler, VlasiosTsiatsis, Catherine Mulligan, Stefan Avesand, StamatisKarnouskos, David Boyle, "From Machine-to-Machine to the Internet of Things: Introduction to a New Age of Intelligence", 1st Edition, Academic Press, 2014
4. Vijay Madisetti and ArshdeepBahga, "Internet of Things (A Hands-on-Approach)", 1st Edition, VPT, 2014.
5. Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow" by Aurélien Géron

Lab - Data Science & Machine Learning

Hours : 15

Course Objective:

To obtain practical experience using data mining tool such as WEkA.

Emphasize hands-on experience working with all real data sets. tools such as WEKA

Exercises:

1. Demonstration of Data mining tools: Weka, Rapid miner, Keel, Orange
2. Introduction, Data pre-processing on dataset
3. Association rule process on dataset using apriori algorithm
4. Classification rule process on dataset using j48 algorithm
5. Classification rule process on dataset using id3 algorithm
6. Classification rule process on dataset using naïve bayes algorithm
7. Clustering rule process on dataset using simple k-means

Course Outcome:

1. Ability to add mining algorithms as a component to the exiting tools
2. Ability to apply mining techniques for realistic data.

References:

- Pang-NingTan,MichaelSteinbach, VipinKumar, IntroductiontoDataMining, 2nd edition, Pearson education, 2018.
- Jlawei Han & Micheline Kamber, Data Mining, “Concepts and Techniques”, 3 rd edition, Morgan Kaufmann Publishers, 2012.
- Margaret H Dunham, Data Mining Introductory and advanced topics, 6th edition, Pearson Education, 2009
- www.cs.waikato.ac.nz/ml/weka/downloading.html

⬅ Back to Index

Computer Aided drug design: Fundamentals & application

Course Code : UBIN703

Major Course

Credits : 3+1P **Hours : 45**

Course Objective:

Aims to provide UG Bioinformatics (Hons.) students with the knowledge and ability to use computer-assisted methods for the design of new drugs

Unit-I **8 lectures**

Introduction to Drugs: How drugs work - Drug targets, drug-target interaction and dose-response relationships; ADME & oral bioavailability of drugs – drug-drug interaction & drug toxicity. New Drug Discovery & Development: Target & Lead Discovery methods;–Preclinical & Clinical Testing of New Drugs. Drug Regulatory Agencies (US FDA, EMA and CDSCO) & Regulatory requirement for New Drug approval (cGLP, cGMP, cGCP, IND, NDA and ANDA).

Unit-II **5 lectures**

Basics of drug design: Introduction to the concept of structure-activity relationship (SAR). Effect of functional group modifications, variation of substituents and ring structures on pharmacokinetics & pharmacodynamics properties of drugs, Bioisosteric replacement and prodrug design to improve oral bioavailability. Importance of Chirality in drugs,

Unit-III **8 lectures**

Computer-aided drug design (CADD): Introduction to CADD and software/tools/webservers used. Structure based drug design: Basics of Docking (pose prediction and scoring algorithms), flexible and rigid docking, protein-ligand and protein-protein docking, De Novo Drug Design concepts (Fragment Placements, Connection Methods, Sequential Grow) and its application.

Unit-IV **8 lectures**

Ligand based drug design Pharmacophore generation (3D database searching, conformation searches, deriving and using 3D Pharmacophore, constrained systematic search, Genetic Algorithm, clique detection techniques, maximum likelihood method). Introduction to QSAR, descriptors used in QSAR study, model building (Regression Analysis, Partial Least Squares (PLS), Principle Components Analysis (PCA)) and model validation methods. 2D vs 3D QSAR. Use of machine learning in QSAR.

Unit-V **8 lectures**

Lead identification & optimization though CADD: Virtual screening for lead identification – drug-likeness concept & Lipinski's rule of 5; cheminformatics and compound library design (tanimoto similarity, drug-like, fragment-like, natural product, etc), database searching using structure-based and ligand-based methods. Chemical space docking and drug repurposing strategy. Structure-based methods for lead optimization. ADMET predictions. Case study: use of CADD for the development of kinase inhibitors for anti-cancer therapy.

Unit-VI **8 lectures**

AI/ML in drug design: Role of AI/ML in modern drug discovery pipelines and common techniques used. Supervised learning: Regression and Classification (SVM, Random Forest, XGBoost, Neural Networks); Unsupervised learning: Clustering, PCA, t-SNE; Deep learning in drug discovery: CNNs, RNNs, Transformers. Evaluation metrics: Accuracy, Precision, Recall, ROC-AUC, F1 Score. Applications of AI/ML in QSAR, Virtual screening and De novo design with case **studies.**

Course Outcome:

Students will have basic understanding of how drugs work and how they are discovered. Also students will have in-depth knowledge on the theory behind computer-aided drug design, as well as apply them for drug design.

Text Books:

1. Drugs: From discovery to approval by Rick NG. 3rd Ed., Wiley Blackwell, 2015.
2. Molecular Modeling – Principles and Applications by Andrew R. Leach Second Edition, Prentice Hall, USA, 2001.
3. Computational Drug Design: A guide for Computational & Medicinal Chemists by David C Young, John Wiley & Sons, Inc. 2009.

Reference Books:

1. Burger's Medicinal Chemistry, Drug discovery and Development. Volume 1, Methods in Drug Discovery. 8th Edition. Edited by Donald J. Abraham and Michael Myers and Kent D. Stewart, Wiley, 2021
2. Essentials of Medical Pharmacology, by Tripathi KD. 6th Ed., Jaypee Brothers (2008)
3. Molecular Modeling: Basic Principles and Applications, by Hans-Dieter Höltje and Gerd Folkers, 3rd Ed., Wiley-vch Verlag Gmbh, 2008.
4. The organic chemistry of drug design and drug action by Richard B. Silverman & Mark W. Holladay, Academic Press, 2015.

Lab - Computer Aided Drug Design

Hours : 15

Course Objective:

Aims to train students with technical skills to perform computer-assisted drug design

1. Small Molecule sketching using Marvin sketch and bond optimization in 2D & 3D format
    a. SMILES, SDF, MOL2 file formats
2. Chemoinformatics analysis of large database of compounds using DataWarrior
    a. Similarity, rule of five and sub-structure based methods
3. Molecular Visualization of proteins using Pymol and Chimera
    a. Pdb file format and Parsing
    b. Visualizing a molecule in different representations
    c. Identifying interacting residues (protein and ligand interactions)
    d. Measuring distances between atoms
    e. B-factor visualization
    f. Image tracing and preparation
4. Protein geometry optimization using SwissPdb Viewer
    a. Energy Minimization of protein molecule
    b. Determining Maxima and Minima energy points
5. Structure based Drug design
    a. Molecular docking using AutoDock
    b. Virtual Screening using AutoDock Vina
    c. Commercial software for docking (eg., Schrodinger GLIDE - Demo)
6. Ligand based Drug design
    a. QSAR studies
    b. Descriptor generation
    c. Model building and validation

Course Outcome:

Students will be skilled to perform computer-aided drug design which will be useful for their research/project work

⬅ Back to Index

Python Programming

Course Code : UBIN704

Major Course

Credits : 3+1P **Hours : 45**

Course Objective:

- This course is designed to build a solid foundation in Python programming with a specific focus on bioinformatics applications.
- It aims to develop students' understanding of programming logic and computational thinking while introducing key concepts such as data types, control structures, and functions.
- Students will learn to handle biological data efficiently using Python’s built-in data structures and file operations.

The course also emphasizes practical experience with biological data and introduces the use of essential libraries such as BioPython

UNIT 1: Basics of Programming and Python Introduction (7 hours)

Understanding computational thinking and programming logic, Importance of programming in bioinformatics, Overview of programming languages and paradigms, Introduction to Python: History, features, and environment setup. Installing Python and setting up the environment, Running Python code (IDLE, terminal, Jupyter, etc.). Writing and executing the first Python program, Python syntax, indentation, and code structure, Keywords, identifiers, and comments in Python, Basic interactive programming using input/output, Summary and practice problems

UNIT 2: Data Types, Variables, and Operators (6 hours)

Data types in Python: integer, float, string, Boolean, Type conversion and type casting, Variables and constants: declaration and usage, Rules for naming variables, String operations and formatting, Operators: Arithmetic operators and expressions, Relational and logical operators, Assignment operators and shorthand notation, Input/output operations.

UNIT 3: Control Structures and Functions (8 hours)

Decision making: if, if-else, elif statements. Looping structures: while and for loops. Loop control statements: break, continue, pass. Introduction to functions: definition and calling, Function parameters, arguments, return values, Scope of variables and recursion basics, Application: GC content calculation, DNA reverse complement.

UNIT-4: Data Structures and File Handling (8 hours)

Lists and operations: creation, indexing, slicing, Tuples and sets: characteristics and uses. Dictionaries: keys, values, operations. Iteration over data structures. Reading and writing text files in

Python. Parsing and analyzing biological data formats (FASTA). Introduction to exception handling: try, except, finally block.

UNIT 5: python packages and Bioinformatics Applications (7 hours)

Modules and Packages: Importing built-in modules (`math`, `random`, etc.), Introduction to custom modules, Using `pip` to install third-party packages. Overview of BioPython: installation and features, Handling sequence data with BioPython, parsing biological files and extracting sequences, Basic tools for sequence manipulation. Applications in Computer aided drug designing.

Course Outcome:

Upon successful completion of this course, students will be able to

- Understand the programming concepts, Python features and demonstrate Python installation, write and execute basic programs.
- Apply python functions to perform calculations and manage data in Python programs and identify Python data types, variables, and naming rules.
- Analyse decision-making statements and loops to control program flow and choose appropriate python function for solving biological data problems.
- Evaluate file handling methods and exception handling techniques while parsing and analysing biological data formats like FASTA and Test and debug Python programs effectively.
- Create small-scale Python scripts and Python-based fundamental solutions for diverse applications in bioinformatics.

Reference Books:

1. Jones, M. (2013). Python for Biologists: A Complete Programming Course for Beginners. CreateSpace Independent Publishing Platform. ISBN: 9781492346135,
2. Lutz, M. (2013). Learning Python (5th Edition). O'Reilly Media. ISBN: 9781449355739
3. Cock, P. J. A., Antao, T., Chang, J. T., Chapman, B. A., Cox, C. J., Dalke, A., ... & de Hoon, M. J. L. (2009). Biopython: Python tools for computational molecular biology and bioinformatics. Bioinformatics, 25(11), 1422-1423. DOI: 10.1093/bioinformatics/btp163
4. Charles Dierbach (2012). Introduction to Computer Science Using Python: A Computational Problem-Solving Focus. Wiley. ISBN: 9780470555155
5. SWAYAM/NPTEL Course Materials (Government of India, Ministry of Education)

Lab - Python Programming

Total Hrs: 15

Course Objective :

This lab components introduces Python programming fundamentals and bioinformatics applications, covering data types, control structures, functions, and file handling.

1. Installing Python and Running First Script- Setting up Python environment, running “Bio, World!” in IDLE, terminal, and Jupyter.
2. Data Types and Variable Manipulation-Exploring Python data types, type conversion, and variable assignment exercises.
3. Conditional Statements Practice- Writing programs using if, if-else, and elif statements
4. Loop Structures and Control Statements- Using for and while loops, and loop control statements (break, continue, pass)
5. Writing and Calling Functions- Defining functions, using parameters, return values, and simple recursion.
6. Calculating GC Content of DNA Sequences-Implementing a function to calculate GC percentage from a DNA string.
7. Working with Lists, Tuples,
8. Sets, and Dictionaries
9. Reading and Writing FASTA Files- Parsing FASTA format files, extracting sequence data, and writing output.
10. Exception Handling in File Operations- Handling errors during file reading/writing with try-except-finally blocks.
11. BioPython Basics and Sequence Analysis- Installing BioPython, loading sequences, and performing basic bioinformatics operations.

Course Outcome:

Students will learn basic concpets of python and apply python for biological data handling and use Bio Python for sequence analysis. Emphasis is placed on writing code to solve real-world bioinformatics problems

⟵ Back to Index

Microscopic Technique for Image Processing

Course Code : UBIN705

Minor Course

Credits : 3 + 1 (Seminar) **Hours : 45**

Course Objective:

To impart knowledge on instrumentation and sample handling methods for several microscopic techniques

Unit 1 **8 Lectures**

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

Unit 2 **8 Lectures**

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

Unit 3 **11 Lectures**

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 - Applications.

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 - Applications.

Unit 4 **9 Lectures**

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 5 **9 Lectures**

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, Resolution.

Course Outcome:

The students will possess theoretical knowledge in handling the several microscopic techniques for imaging biological samples.

Reference Books:

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

Seminar (Microscopic techniques for Biological samples)

Hours : 15

The students are to give a seminar for 1 credit. A research paper (published in last 5 years) that used any of the microscopic techniques discussed here is to be chosen by the student for the seminar.

⬅ Back to Index

Genetic Engineering & Applications

Course Code : UBIN706

Minor Course

Credits : 3+1P **Hours : 45**

Course Objective:

Paste your content here.

Unit 1 Introduction to recombinant DNA technology **6 Lectures**

Overview of recombinant DNA technology and scope of genetic engineering. Good laboratory practices, Regulations for recombinant DNA research, Biosafety and Bioethics. Restriction and modification systems, restriction endonucleases and other enzymes used in manipulating DNA molecules.

Unit 2 Cloning and expression vectors **8 Lectures**

Plasmids and bacteriophages as vectors for gene cloning. Cloning vectors based on E. coli plasmids, pBR322, pUC8, pGEM3Z. Vectors for expression of foreign genes in E. coli, cassettes and gene fusions. Challenges in producing recombinant protein in E. coli. Production of recombinant protein by eukaryotic cells. Fusion tags and their role in purification of recombinant proteins.

Unit 3 Ligation and Introduction of DNA into cells **8 Lectures**

Ligation of DNA molecules. DNA ligase, sticky ends, blunt ends, linkers and adapters. Uptake of DNA by cells, preparation of competent cells. Selection for transformed cells. Identification for recombinants - insertional inactivation, blue-white selection. Introduction of phage DNA into bacterial cells. Introduction of DNA into plant and animal cells, electroporation.

Unit 4 Methods for clone identification **6 Lectures**

The problem of selection, direct selection, radioactive non- radioactive probes. Gene libraries, identification of a clone from gene library, colony and plaque hybridization probing. Blotting techniques.

Unit 5 Polymerase chain reaction **8 Lectures**

Fundamentals of polymerase chain reaction, designing primers for PCR. Studying PCR products. Cloning PCR products. Real time PCR. DNA sequencing by Sanger’s method, Automated DNA sequencing. Pyrosequencing.

Unit 6 Applications of genetic engineering in Biotechnology 9 Lectures

Site–directed mutagenesis and protein engineering. Applications in medicine, production of recombinant pharmaceuticals such as insulin, human growth hormone, factor VIII. 33 Recombinant vaccines. Gene therapy. Applications in agriculture - plant genetic engineering, herbicide resistant crops, problems with genetically modified plants, safety concerns.

Course Outcome:

Paste your content here.

Reference Books:

1. Gene Cloning and DNA Analysis (2010) 6th ed., Brown, T.A., Wiley-Blackwell publishing (Oxford, UK), ISBN: 978-1-4051-8173-0.
2. Principles of Gene Manipulation and Genomics (2006) 7th ed., Primrose, S.B., and Twyman, R. M., Blackwell publishing (Oxford, UK) ISBN:13: 978-1-4051-3544-3.
3. Molecular Biotechnology: Principles and Applications of Recombinant DNA (2010) 4th ed., Glick B.R., Pasternak, J.J. and Patten, C.L., ASM Press (Washington DC), ISBN: 978-1-55581-498-4 (HC).

Lab - Genetic Engineering and Applications

Hours : 15

1. Isolation of plasmid DNA from E. coli cells.
2. Digestion of plasmid DNA with restriction enzymes.
3. Amplification of a DNA fragment by PCR.
4. Preparation of competent E. coli cells
5. Transformation of E. coli cells with plasmid DNA.
6. Molecular weight estimation of fragmented DNA using agarose gel electrophoresis

⬅ Back to Index

SEMESTER - VIII

Introduction to Cheminformatics

Course Code : 801

Major Course

Credits : 4 **Hours : 60**

Course Objective:

This course enables the students to get introduced with the field of cheminformatics by understanding the representation of molecules and chemical reactions as the chemical data information and the methodologies towards the designing of novel drug candidates.

Unit I: Introduction and Basics of Graph theory **12 hrs**

Origin and scope of cheminformatics, Definition and concepts of Cheminformatics, Basics of Graph Theory:- Matrix Representations:- Adjacency, Distance, Atom Connectivity and Bond Matrix. Connection Table.

Unit 2: Chemical representation of molecules **12 hrs**

Representation of Chemical Compounds:- Line Notations-WLN, SMILES, SMARTS, SYBYL and InCHI; Exchange formats for chemical structures (MOL, SDF and PDB) and reactions (RXN and RDF), Representation of 3D structures:-Z Matrix.

Unit 3 Chemical structure searching and Databases **12 hrs**

Molecular structure Searching techniques:- Full structure search, Substructure search, Three dimensional similarity searching methods. Chemical Databases:- PubChem, Chembank, Drugbank, ZINC, eMolecules, etc. Literature databases:- Pubmed, Medline. Other databases:- PDB, PRIDE, chemPDB, KEGG, Ligand Info, etc.

Unit 4: Molecular properties as descriptors **12 hrs**

Molecular descriptors and fingerprints:- Different types of descriptors and fingerprints. Tools and methods for their calculations: PaDEL and RDKit; Development and validation of QSAR/QSPR models. Statistical assessment and validation of models; Molecular Interaction Fields:- 3D QSAR.

Unit 5: Applications **12 hrs**

Introduction to drug design; Target Identification and Validation; Lead Finding and Optimization; Analysis of HTS data; Virtual Screening; Design of Combinatorial Libraries; Ligand-Based and Structure Based Drug design; Application of Chemoinformatics in Drug Design. Case studies.

.

Course Outcome:

Students will gain expertise in the field of cheminformatics with an understanding of the methodologies and the applications of chemical databases towards the designing of novel drug candidates.

Reference Books:

1. Johann Gasteiger (2003), Thomas Engel (ed.), Chemoinformatics: A Textbook, Wiley VCH, Weinheim.
2. Leach, A. R., & Gillet, V. J. (2017). An Introduction to Cheminformatics, Springer, Varnek, A.
3. Engel, T., & Gasteiger, J. (2018). Chemoinformatics: basic concepts and methods. John Wiley & Sons.
4. Sharma, N., Ojha, H., Raghav, P., & Goyal, R. K. (2021). Chemoinformatics and Bioinformatics in the Pharmaceutical Sciences. Elsevier.
5. Stromgaard, K., Krogsgaard-Larsen, P., & Madsen, U. (2009). Textbook of drug design and discovery. CRC press. ISBN: 9780429111242.
   Larsen (2004) Textbook of Drug Design and Discovery, 3rd edition, Taylor and Francis, London and New York

⬅ Back to Index

Drug Design, Discovery and Development

Course Code : UBIN802

Minor Course

Credits : 3+1P **Hours : 45**

Course Objective:

Aims to provide UG students (Non-Bioinformatics) with the knowledge about drugs and how to design them using computer-assisted methods.

Unit-I **8 Lectures**

Introduction to Drugs: Drug nomenclature, Routes of drug administration and dosage forms, Principles of Pharmacokinetics and Pharmacodynamics: ADME, Bioavailability of drugs. How drugs work - Drug targets, drug-target interaction and dose-response relationships.

Unit II **8 Lectures**

New Drug Discovery & Development: Overview of new drug discovery, development, cost and time lines. **Target Identification & Validation**. **Lead Discovery**: Rational and random approaches - Drug repurposing, Natural products, High-throughput screening (HTS), Combinatorial chemistry and computer aided drug design (CADD).

Unit III **8 Lectures**

Preclinical trials of New Drugs: Pharmacology - In vitro/in vivo Pharmacokinetics and Pharmacodynamics; Toxicology - Acute, chronic, carcinogenicity and reproductive toxicity; Drug formulation. **Clinical Trial of New Drugs**: Phase I, Phase II and Phase III; Good clinical practice (GCP) guidelines - Investigators brochures, Clinical trial protocols and trial design; Ethical issues in clinical trials - How are patient rights protected?

Unit IV **8 Lectures**

Drug Regulatory Agencies: US Food & Drug Administration (US FDA) and Central Drugs Standard Control Organization (CDSCO), India. **Regulatory Applications & New Drug Approval**: Investigational new drug (IND) application & New drug application (NDA); Regulatory review and approval process. **Regulatory Requirements for Drug Manufacturing**: Current Good manufacturing practice (cGMP) and GMP manufacturing facility inspection & approval.

Unit-V **13 Lectures**

Computer-aided drug design (CADD): Introduction to CADD, Bioinformatics & chemoinformatics**;** software/tools/webservers used for the design of drugs. **Ligand-based drug design (LBDD)**: QSAR/ QSPR methodology for drug design - Various Descriptors used in QSAR studies - deriving & validating QSAR equations. 3D Pharamcophore development - conformation generation, deriving and using 3D Pharmacophores. **Structure-based drug design (SBDD)**: 3D structure of drug targets, Molecular Docking and *De novo* ligand design. Virtual screening for lead identification.

Course Outcome:

Students will understand how drugs work and how they are discovered. Also students will understand the basics of computer-aided drug design, as well as apply them for drug design.

Text Books:

1. Essentials of Medical Pharmacology, by Tripathi KD. 6th Ed., Jaypee Brothers (2008)
2. Drugs: From discovery to approval by Rick NG. 3rd Ed., Wiley Blackwell, 2015.
3. Computational Drug Design: A guide for Computational & Medicinal Chemists by David C Young, John Wiley & Sons, Inc. 2009.

Reference Books:

1. Burger's Medicinal Chemistry, Drug discovery and Development. Volume 1, Methods in Drug Discovery. 8th Edition. Edited by Donald J. Abraham and Michael Myers and Kent D. Stewart, Wiley, 2021
2. The organic chemistry of drug design and drug action by Richard B. Silverman & Mark W. Holladay, 3rd Ed., Academic Press, 2014.
3. Molecular Modeling – Principles and Applications by Andrew R. Leach Second Edition, Prentice Hall, USA, 2001.

LAB - DRUG DESIGN & DISCOVERY

Hours : 15

COURSE OBJECTIVES: *Aims to train students with technical skills to perform computer-assisted drug design*

1. Small Molecule sketching using Marvin sketch and bond optimization in 2D & 3D format
    - SMILES, SDF, MOL2 file formats
    - Databases for small molecules (Drug bank, ChEMBL, etc)

2. Protein sequence retrieval
    - UniProt

3. Molecular Visualization of proteins using Pymol and Chimera
    - PDB file format and Parsing
    - Visualizing a molecule in different representations
    - Identifying interacting residues (protein and ligand interactions)
    - Measuring distances between atoms

4. Protein geometry optimization using SwissPdb Viewer
    - Energy Minimization of protein molecule
    - Determining Maxima and Minima energy points

5. Binding Site Identification
    - Different approaches for binding site identification
    - Tools - Cast-P, POCASA, 3D ligand site, Metapocket, Ghecom

6. Structure based Drug design
    - Molecular docking using AutoDock
    - Virtual Screening using AutoDock Vina

COURSE OUTCOME: *Students will be skilled to perform computer-aided drug design which will be useful for their research/project work*

⬅ Back to Index