Department: Bioinformatics Centre, Savitribai Phule Pune University

Course: M. Sc. Bioinformatics

Duration: 2 years

Total Number of Credits: 88

Semester I		
Mandatory Courses		
Subject Code	**Subject Title**	**Number of Credits**
BIM 101 (T+P)	Biological Data & Databases	2T + 2P
BIM 102 (T+P)	Biological Sequence Analysis	2T + 2P
BIM 103 (T+P)	Biomolecular Structure & Organization	2T + 2P
BIM 104 (P)	Python Programming	2P
	Total Mandatory Credits	6T + 8P =14 credits
BIM 105 (T)	Research Methodology	4T
Elective Courses ***Courses worth of 4 credits from among the following:***		
Subject Code	**Subject Title**	**Number of Credits**
BIE 106 (T)*	Mathematics & Statistics	2T
BIE 107 (T)*	Basic Biology	2T
BIE 108 (T)	Molecular Biology	2T
BIE 109 (T)	Biochemistry and Cell Biology	2T
	Total Elective credits	2 credits
	Semester Total	**22 credits**

* Depending upon the UG training background, the students coming from Biology, Chemistry and all streams other than Mathematics and Statistics will opt for the course BIE 106 Mathematics & Statistics and the students coming from Mathematics and Statistics streams will opt for BIE 107 Basic Biology.

Semester II		
Mandatory Courses		
Subject Code	**Subject Title**	**Number of Credits**
BIM 201 (T+P)	Structural Bioinformatics	2T + 2P
BIM 202 (T+P)	Chemoinformatics	2T + 2P
BIM 203 (T+P)	Genomics	2T + 2P
BIM 204 (T+P)	Introduction to Database Systems	2P
	Total Mandatory Credits	6T + 8P = 14 credits
BIM 205 (P)	On Job Training/ Field Project	4P
Elective Courses ***Courses worth of 4 credits from among the following:***		
Subject Code	**Subject Title**	**Number of Credits**
BIE 206 (P)	Biological Data Curation and Analysis	2P
BIE 207 (P)	Advanced Techniques in Experimental Biology	4P
BIE 208 (T+P)	Molecular Phylogenetics	1T + 1P
BIE 209 (T+P)	Programming in Java	2T + 2P
BIE 210 (P)	Linux Operating System	2P
BIE 211 (T+P)	Software Testing	1T + 1P
	Total Elective credits	4 credits
	Semester Total	**22 credits**

Semester I

BIM 101 (T+P): Biological Data & Databases **(2T + 2P)**

Objectives:

This course will enable the students to:

- understand the nature of biological data and need for biological databases
- to understand and explore a few major biological databases (organization and contents); search and retrieve data from the databases using their respective search engines

Theory

Syllabus:

- Introduction to Bioinformatics (2)
- Overview of Bioinformatics resources on the web (2)
  NCBI/EBI/EXPASY etc
- Nature of biological data and formats (2)
- Biological literature databases (2)
  PubMed
- Nucleic acid sequence databases (6)
  - GenBank, EMBL, DDBJ
  - RefSeq, dbSTS, dbEST
- Protein sequence databases (6)
  - UniProtKB
  - UniRef, UniParc, Proteomes, NextProt
- Derived databases (5)
  - InterPro and constituent databases
  - Recent derived databases.
- RNA sequence databases (3)
  - miRBase, lncRNAdb, MIT/ICBP siRNA database.
- Species and Biodiversity databases / resources (2)
  - NCBI Taxonomy database, GBIF, Sahyadri

Practicals

Objectives:

This course will enable the students to:

- To understand and explore major biological sequence databases (organization and contents); search and retrieve data from the databases using their respective search engines.

Syllabus:

- Exploring the integrated database system at NCBI server and querying the PUBMED, GenBank, NCBI Taxonomy database and other databases using the ENTREZ search engine (6)
  - Use of operators (AND, OR & NOT)
  - Use of limits

    - Exporting GI list
    - Batch retrieval
- Exploring the integrated database system at EMBL-EBI server (4)
- Exploring & querying UniProtKB and other protein sequence databases. (4)
- Sequence Formats & Format converters (2)
- Exploring tools on ExPASY (2)
- Exploring and using InterPro, its constituent databases and recent derived databases. (6)
- Exploring and querying RNA sequence databases: miRBase, lncRNAdb, MIT/ICBP siRNA database. (4)
- Exploring and querying Biodiversity databases / resources: GBIF and Sahyadri (2)

References:

- Nucleic Acids Research - Database issue (Most recent issues relevant to appropriate databases)
- Baxevanis A.D., Davison D.B., Page R. D. M. & Petsko G.A. Current Protocols in Bioinformatics. New York, John Wiley & Sons Inc., Latest Edition
- Korf Ian, Yandell Mark, Bedell Joseph. BLAST: an essential guide to the basic local alignment search tool. Shroff Publishers and Distributors Pvt. Ltd., Latest Edition
- Baxevanis Andreas D. Bioinformatics: A Practical Guide to the Analysis of Genes and Proteins, Latest Edition. Publisher: New York, John Wiley & Sons, Inc.
- Teresa Attwood, Parry-Smith David J. Introduction to Bioinformatics. Publisher: Pearson Education (Singapore) Pte.Ltd., Latest Edition.
- Mount David W.. Bioinformatics: Sequence and Genome Analysis. Publisher: Cold Spring Harbor Laboratory Press; Latest Edition
- Gibas Cynthia, Jambeck Per. Developing Bioinformatics Computer Skills. Publisher: Shroff Publishers and distributors O'Reilly Media, Inc., Latest Edition

BIM 102 (T+P): Biological Sequence Analysis (2T + 2P)

Objectives:

This course will enable the students to:

- understand the concepts and methods for biomolecular sequence analysis
- study various algorithms for sequence analysis and understand underlying statistics
- understand when and why to use these algorithms to answer some of the biological questions

Syllabus: Theory

- Biomolecular sequence analysis (2)
  - Overview
  - Concepts
- Analysis of single sequence (2)
  - Nucleotide
  - Protein
- Pairwise sequence alignment algorithms (3)
  - Needleman & Wunsch
  - Smith & Waterman
- Scoring matrices for Protein and Nucleotide sequences (3)
  - MDM/PAM series
  - BLOSUM series
  - CSW
- Database Similarity Searches (6)
  - FASTA
  - BLAST
- Multiple sequence alignment algorithms (6)
  - CLUSTALW
  - MUSCLE
  - DALIGN
  - T-Coffee
- Sequence logos, consensus & patterns (3)
- Basic concept of sequence profiles, Derivation of profiles; applications (5)
  - Gribskov's Profile Analysis method
  - PSI-BLAST

References:

- Pevsner, Jonathan, Bioinformatics and functional genomics 2nd ed. Publisher: John Wiley & Sons,2009. ISBN: 978 0470085851
- Posada, David, Bioinformatics for DNA sequences analysis. Publisher: Humana Press, 2009. ISBN 978 1588299109
- Haubold, Bernhard, Introduction to computational Biology: An Evolutionary Approach. Publisher: Birkhäuser, 2006, ISBN: 9783764367008
- Xiong, Jin, Essential Bioinformatics. Publisher: Cambridge University Press, 2006. ISBN: 978 0521706100
- Buehler, Lukas K Ed., Bioinformatics Basics: Applications in Biological Science and Medicine. Publisher: CRC Press, 2005 ISBN: 0849312833
- Mount,David. Bioinformatics : sequence and genome analysis. Publisher: Cold Spring Harbor Laboratory Press, 2004. ISBN: 087967121
- Ian, Korf. Blast. Publisher: Oreilly. 2003. ISBN: 9788173665127

- Baxevanis, A.D. Bioinformatics: a practical guide to the analysis of genes and proteins. 2nd Ed.. Publisher: John Wiley & Sons, 2002, ISBN: 9814126756
- Attwood, T.K. Introduction to Bioinformatics. Publisher: Pearson Education Pte. Ltd, 2001. ISBN: 8178085070

Practicals/Tutorials:

Objectives:

This course will enable the students to:

- understand how to use various algorithms for biomolecular sequence analysis
- understand use of various parameter in respective algorithms and their effects on the outcome
- learn to write codes and practice programming skills for some of the simple tasks and to understand how bioinformatics evolved

Syllabus:

- Computing & plotting mono-, di-, tri-mer frequencies (2)
- Analysis & comparison of sequence with itself (1)
- Compute properties of sequences using tools on ExPASY (2)
- Exploring utilities in EMBOSS packages (2)
- Pairwise global alignment using Needleman-Wunsch algorithm (2)
    - DNA
    - Protein
- Pairwise local alignment using Smith-Waterman algorithm (2)
    - DNA
    - Protein
- Pairwise alignments (2)
    - Use of scoring matrices
    - Interpretation of results: dos & don'ts
- Database (homology) searches using different versions of FASTA (2)
    - DNA/Protein sewaquence as a query
    - Interpretation of results
- Database (homology) searches using different versions of BLAST (4)
    - DNA/Protein sequence as a query
    - Interpretation of results
- Multiple sequence alignment programs (5)
    - Compilation of dataset
    - Learning to use programs
    - Input/output formats & parameters
    - Analysis & interpretation of results
    - Alignment visualization and editing tools (JalView)
- Build sequence logos (1)
- Deriving consensus & Prosite patterns (2)
- Using PSI-BLAST for detection of distant homologs (3)

Reference books:

- Sung Wing-Kin, Algorithms in Bioinformatics: pratical introduction. Publisher: CRC Press, 2010. ISBN: 978 1420070330
- Gibas, Cyntbia, Developing bioinformatics computer skills. Publisher: Oreilly, 2001. ISBN: 8173662428

- Pevsner, Jonathan, Bionformatics and functional genomics 2nd ed.Publisher: John Wiley & Sons,2009. ISBN: 978 0470085851
- Posada, David, Bioinformatics for DNA sequences analysis. Publisher: Humana Press, 2009. ISBN 978 1588299109
- Haubold, Bernhard, Introduction to computational Biology: An Evalutionary Approach. Publisher: Birkhäuser, 2006, ISBN: 9783764367008
- Xiong, Jin, Essential Bioinformatics. Publisher: Cambridge University Press, 2006. ISBN: 978 0521706100
- Mount,David. Bioinformatics: sequence and genome analysis. Publisher: Cold Spring Harbor Laboratory Press, 2004. ISBN: 087967121
- Ian, Korf. Blast. Publisher: Oreilly. 2003. ISBN: 9788173665127
- Baxevanis, A.D. Bioinformatics: a practical guide to the analysis of genes and proteins. 2nd Ed.. Publisher: John Wiley & Sons, 2002, ISBN: 9814126756
- Attwood, T.K. Introduction to Bioinformatics. Publisher: Pearson Education Pte. Ltd, 2001. ISBN: 8178085070

BIM 103 (T+P): Biomolecular Structure & Organization (2T + 2P)

Objectives:

This course will introduce the students to:

- Basic physico-chemical principles involved in structural organization of macromolecules
- Organization of the structures of proteins, DNA, RNA, carbohydrates and lipids
- experimental methods of structure determination

Theory

Syllabus:

- **Physicochemical Principles of biomolecular structure organization** (1)
  - Basic concepts of atom structure, hybridization of atomic orbitals, valence, covalent bonds, atomic interactions and forces, formation of polymeric molecules
- Co-ordinate systems: Rectangular, Cylindrical and spherical coordinate systems (2)
- **Experimental Methods for determination of biomolecular structures** (4)
  - X-ray Diffraction
  - NMR Spectroscopy
- **Protein Structure** (10)
  - Internal Coordinates – Bond lengths, bond angles, torsional angles; peptide unit, Ramachandran Map; Calculation of dihedral angles; Fourth atom fixing
  - Hierarchical Organization of Protein structure - Primary, Secondary, Supersecondary, Tertiary and Quaternary structure; Membrane Protein Structures
  - Principles of protein folding and Energetics; Mechanism of Protein folding; methods to study protein folding
- **DNA and RNA Structure** (8)
  - Base pairing in DNA & RNA
  - Double Helix – Organization, types and structural features
  - Structural & Geometric parameters associated with DNA
  - Secondary structures of DNA (triple helices, quadruplex, cruciform)
  - Sequence-structure relationships in DNA
  - Secondary Structures in RNA; Representations of RNA structures
  - Energetics of RNA structure
  - tRNA structure
- **Carbohydrates** (3)
  - Conformations of mono and oligosaccharides and Correlation with properties
  - Glycoproteins, Proteoglycans and Glycolipids: Structural aspects
- **Lipids & Membranes** (3)
  - Membrane microdomains and organization

Practicals:

- Protein Data Bank (2)
- Nucleic Acid Databank (2)
- Visualization of structures (SWISS-PDB Viewer, Discovery Studio) (5)
- Calculation of structural parameters of Proteins & Ramachandran Plot (3)
- Calculation of structural parameters of DNA & RNA (3)

- Calculation of structural parameters of Carbohydrates and Lipids (3)
- Understanding Macromolecular interactions through visualisation & structure analysis (6)

    Protein – Protein
    Protein – Nucleic acids
    Protein - carbohydrates

Understanding assemblies of biomolecules through visualization: Ribosome, Nucleosome, & Viral particles (6)

BIM 104 (P): Python Programming **(2P)**

Objectives:

This course will enable the students to:

- Learn python programming concepts
- Analysis of biological data using Python codes

Syllabus:

- Introduction, Features of Python, IDLE environment (1)
- Data types (1)
- Variables (2)
- Data operators (2)
- Input and Output processing (2)
- Control statements, loops, iterations (3)
- Exception handling (1)
- Collections: list, tuple, set, dictionary (5)
- File handling (4)
- Pattern matching (2)
- Object oriented concepts – classes, attributes, methods (1)
- Functions (2)
- Python libraries: using standard modules and creating a new module (2)
- Working with Biological data (2)

References:

- Mitchell L Model, Bioinformatics programming using Python. Publisher: O'Reilly Media. ISBN: 9788184048988
- Paul Berry, Head First Python. Publisher: Shroff/ O'Reilly ISBN: 9350231883
- Wesley J Chun, Core Python Applications Programming. Publisher: Pearson Education ISBN: 8131791343
- Sebastian Bassi, Python for Bioinformatics. Publisher: CRC Press. ISBN: 9781584889298
- Jason Kinser, Python for Bioinformatics. Publisher: Jones and Bartlet. ISBN: 0763751863

BIM 105 (P): Research Methodology **(4T)**

Objectives:
This course will enable the students to gain insights into the detailed methodologies and various other aspects of research

Module 1: Foundation of Research Methodology **(15)**

1. History and Philosophy of Research: What is Research? Ancient methodologies and research in various fields of science. Ancient Indian research methodologies. Scientific Method in research. Philosophy of Research including Rationalism, Empiricism, Skepticism, Reductionism, Pseudoscience etc. Qualities of a Researcher.

2. **Problem Identification & Formulation**: definition and formulating the research problem, Necessity of defining the problem, Importance of literature review in defining a problem, Literature survey: primary and secondary; web sources; critical literature review. Design of Research Question.

3. **Approaches to Research Design:** Induction and deduction in research. Variables in research. Features of a good research design - Exploratory Research Design - Concept, Types and uses, Descriptive Research Design - concept, types and uses. Experimental Design - Concept of Independent & Dependent variables. Biased and unbiased research design.

Module 2: Data Collection, Processing and Analysis **(15)**

1. **Types of Biological data and nature of research:** Qualitative data – Quantitative data Concept of measurement, causality, generalization, replication. Basic concepts in qualitative and quantitative data analysis.

2. **Data Collection:** Execution of the research - Observation and Collection of data - Methods of data collection Databases (ENA, UniProtKB, PDB, GEO etc) searching databases, data retrieval, preparing data sets, data curation - Positive and negative data sets. Basic concepts in organization of data into databases - hypothesis-testing - Generalization and Interpretation. Observation, Surveys, Monitoring research. records to be maintained by researcher.

3. **Sampling & Measurement**: Concept of Statistical population, Sample, Sampling Frame, Sampling Error, Sample size, Non-Response. Characteristics of a good sample. Probability Sample - Simple Random Sample, Systematic Sample, Stratified Random Sample & Multi-stage sampling etc. Determining size of the sample - Practical considerations in sampling and sample size. Concept of measurement - what is measured? Problem in measurement in research - Validity and Reliability. Levels of measurement - Nominal, Ordinal, Interval, Ratio.

5. **Data Analysis and Interpretation:** Graphical representation of data, Univariate analysis (frequency tables, bar charts, pie charts, percentages), Bivariate analysis – Cross tabulations and Chi-square test including testing hypothesis of association. Correlation and regression analysis. Probability distributions: Discrete distribution (bionomial & Poisson) – Continuous distribution (Normal & Exponential) Bayesian Modeling, Estimation of accuracy (sensitivity specificity, Mathew's Correlation coefficient)

Module 3: Scientific Paper Writing and Communication **(20)**

1 **Literature Review and Writing a Review**: Planning and Writing a Research Proposal to seek funding

2. **Use of tools / techniques for referencing**: methods to search required information effectively, PubMed, effective literature search using Entrez, Google Scholar

3. **Scientific writing and communication**: Different types of publications, their structure and components, Choice of journals for publication based on various parameters such as impact factor, scope of journal etc. Types of report - Technical reports and thesis, Different steps in the preparation - Layout, Structure and Language of typical reports, Paraphrasing, Illustrations and tables, Bibliography, referencing. Oral presentation, Making presentation, Use of visual aids, Importance of effective communication. Tools/Software for paper formatting like MSOffice, software for detection of Plagiarism

Module 4: Research Ethics & Safety **(10)**

1. Ethics in Research: Ethical issues, ethical committees, Reproduction of published material - Plagiarism - citation and acknowledgement, Self-Plagiarism, Reproducibility and accountability

2. IPR: Commercialization, Copy right, Royalty, Intellectual property rights and patent law, Trade related aspects of intellectual property Rights

3. Biosafety: Good lab practices, Guidelines for Biosafety, Institutional Biosafety committee, Institutional Animal ethics committee.

References:

- 'History of the Scientific Methods' by Martin Shuttleworth, https://explorable.com/history-of-the- scientific-method.
- 'The Statistical Analysis of Experimental Data' by, John Mandel, ISBN: 0486646661, ISBN13: 9780486646664
- 'Research Methodology and Scientific Writing' by C. George Thomas, ISBN 978-3-030-64864-0
- 'Research Methods for Science' by Michael Marder, Online ISBN:9781139035118
- Research Methodology: An Introduction - Stuart Melville and Wayne
- Practical Research Methods - Catherine Dawson
- Research Methodology - C. R. Kothari Essential Bioinformatics – Jin Xiong (Cambridge University Press)

Elective Courses

Courses worth of 4 credits from among the following:

BIE 106 (T): Mathematics & Statistics (2T)

Objectives:

This course will enable the students to achieve skills in statistics and mathematics that are essential for application in bioinformatics.

Mathematics (1 credit)

- Introduction to Set Theory (1)
- Trigonometry: (2)
  Trigonometric Functions, Series Expansion, Inverse, General Values, Graphs, Taylor series
- Vector & Matrices (2)
  Vector Algebra, Vector Calculus, Basic Computations, Matrices
- Calculus: (4)
  Limits, Continuity, Analysis, Differentiation (1D & Partial), Reimann Integration, Definite Integrals.
- Ordinary & Partial Differential Equation (3)
  1st Order & 2nd Order Ordinary Differential Equations. Self-Adjoint Equations, Method of Separation of Variables
- Integral transform (3)
  Fourier Series, Fourier Transform, Laplace Transform

Statistics (1 credit)

- Overview of applications of statistics in Bioinformatics (1)
- Introduction to principles of statistical sampling from a population (1)
- Frequency Distributions and Statistical Measures: mean, mode, median, variance, standard deviation, coefficient of variation, measures of skewness and kurtosis. Computation of these measurements for the given data. (2)
- Introduction to theory of Probability, Conditional Probability, Bayesian Rules, Random variable, Distributions of random variables, Binomial, Poisson, Geometric, Normal and extreme value distribution, (5)
- Regression, correlation, fitting regression line. (2)
- Hypothesis testing: (4)
  - Test of significance viz. Z test, t test, paired t test, chi^2 test of goodness of fit

References:

- Isaev Alexander. Introduction to Mathematical Methods in Bioinformatics. Publisher: Berlin; New York: Springer, 2004. ISBN: 3540219730
- Raman K. V. & Pal Sourav. Mathematics in Chemistry. New Delhi, Vikas Publishing House Pvt Ltd., 2004. ISBN: 8125912886
- Jones D.S., Sleeman B.D.. Differential Equations and Mathematical Biology Publisher : Chapman & Hall. 2003. ISBN:1584882964
- Bracewell Ronald. The Fourier transform and it's applications 3rd edition. Publisher: New Delhi : McGraw Hill, 2000. ISBN: 0073039381.
- Stephenson G., Radmore P. M.. Advanced Mathematical Methods for Engineering and Science Students Cambridge: Cambridge University Press, 1990.
- Arfken George. Mathematical methods for physicists. Publisher: Orlando : Academic Press 1985. ISBN: 0120598205.

BIE 107 (T): Basic Biology **(2T)**

Objectives:

This course will enable the students to:

- Understand the basic principles of biology.
- Appreciate the basic nature and diversity of microbial, plant and animal life.
- Understand the classification of organisms and taxonomy.

Syllabus:

- **Origin and evolution of Life**: Living and nonliving things; theories and evidences of origin and evolution of life; chemical evolution- atom to procell and biological evolution- procell to human; cellular tree of life. (4)

- **Unit of life**: Basic unit of life; prokaryotic cell, eukaryotic cell, plant cell and animal cell; Structure and functions of cell organelles- membrane and non-membrane organelles. (4)

- **Diversity of life**: Classification of organisms- five-kingdom classification (Monera, Protista, Fungi, Plantae and Animalia); classification and characteristics of different kingdoms; Biodiversity- hotspots of biodiversity, ecological and economic role of biodiversity; threats to biodiversity. (16)

- **Ecology and nutrition**: Biotic and abiotic components; food chain; trophic levels; food webs; ecological pyramids; ecosystems- structure, types of ecosystem; Nutrition- macronutrients and micronutrients; modes of nutrition- parasitic, saprophytic, symbiotic and insectivorous. (3)

- **Viruses**: Biology of viruses; bacteriophages, plant and animal viruses. (3)

References:

- Wallace Robert A., Sanders Gerald P., Ferl Robert J. The science of life. Publisher: New York, NY : Harper Collins, 1991. ISBN: 0673380440.
- Solomon Eldra P., Berg Linda R., Martin Diana W. Biology 6th edition. Publisher: Pacific Grove, CA, Brooks/Cole Thomson Learning, 2002. ISBN: 0030335035.
- Recknagel F. 2002 Ecological Informatics: Understanding Ecology by Biologically-Inspired Computation, Springer, New York.
- Odum E.P. 1983 Basic Ecology. Saunders International Edition, Japan.
- The Cell: A Molecular Approach (5th edition) by Cooper, G.M. and Hoffman, R.E., A. S. M. Press, 2009.
- The world of the Cell (7th edition) by Becker, W.M., Kleinsmith, L.J., Hardin, J and Bertoni, G.P., Pearson / Benjamin Cummings, 2009
- Molecular Cell Biology (5th edition) by H. Lodish, A. Berk, P. Matsudaira, C.A. Kaiser, M. Krieger, M. P. Scott, S. L. Zipursky & J. Darnell, W. H. Freeman & Company, 2004.

BIE 108 (T): Molecular Biology (2T)

Objectives:

This course will enable the students to:

- understand the concepts in gene structure and expression which leads to specific functions.
- understand the concepts in vital processes such as replication, transcription, post-transcriptional modifications, translation etc.

Syllabus:

- Nucleic acid: Composition, Primary and Secondary structures, Circular DNA (2)
- Genome organization: Prokaryotic and eukaryotic genomes – C value paradox, repetitive and non-repetitive DNA., transposons and retroposons; Exons and introns – organization of interrupted genes, one gene-many proteins concept; Gene numbers – essential genes and total gene number, gene clusters, pseudogenes; Gene families – globin and rDNA gene families; Organelle genome – mitochondrial and chloroplast. (4)
- Packaging of genome – Bacterial genome as nucleoid; Eukaryotic genome – nucleosomes, chromatin, solenoids, loops, domains, scaffolds, chromosomes (2)
- DNA Replication- Details of prokaryotic and eukaryotic DNA replication: DNA polymerases, initiation, elongation and termination of replication; multiple origins of replication, Regulation of replication (4)
- DNA damage, repair and recombination- DNA damaging agents physical and chemical, types of DNA damages; DNA repair systems in prokaryotes and eukaryotes -Single step repair, Base excision repair, Nucleotide excision repair, Mismatch repair, Recombination repair; Recombination - homologous and non-homologous recombination. (2)
- Gene Expression - Transcription: Details of prokaryotic and eukaryotic transcription: RNA polymerase, promoters, initiation, elongation and termination of transcription; regulation of transcription, operons. (4)
- Processing of transcripts - 5' capping, 3' polyadenylation, splicing and editing, self-splicing (1)
- Translation: Details of prokaryotic and eukaryotic translation: Protein synthesis machinery, initiation, elongation and termination, Genetic code, accuracy of translation, regulation, Post-translational modifications of proteins (6)
- Recombinant DNA technology: Enzymes, cloning strategies, vectors, screening for recombinants (3)
- Small non-coding RNAs – micro-RNA, small nuclear RNA, small nucleolar RNA, si-RNA, piwi-RNA and Genome editing. (2)

References:

- Watson James D., Baker Tania A., Bell Stephen P., Alexander Gann, Levine, Michael Losick Richard. Molecular Biology of the Gene 6th Edition. Publisher: New York, Cold Spring Harbor Laboratory Press. 2008. ISBN: 9780321507815
- Benjamin Lewin, Jocelyn E. Krebs, Elliott S. Goldstein, Stephen T. Genes XIth Edition. Publisher: Kilpatrick Jones & Bartlett Publishers, 2014.

BIE 109 (T): Biochemistry and Cell Biology (2T)

Objectives:

This course will enable the students to:

- Understand the physicochemical properties of bio-macro-molecules with their building blocks and their interactions in an aqueous environment
- Understand the structure-function relationships of bio-macro-molecules, the principles of enzyme catalysis, regulation and inhibition.
- Study the principles of structure and function of cells, membranes and organelles.
- Study the cell division cycles with their regulation and mechanism of signal transduction.

Biochemistry (1 credit)

- **Chemistry of life**: Atoms, elements, ions, molecules and macromolecules of life. (2)
- **Water:** Water as the universal biological solvent; physicochemical properties of water; Ionization of water; concepts of pH, pI, pKa, pKb, buffer, acidosis, alkalosis and osmolarity; Henderson-Hasselbalch equation. (2)
- **Biomolecules:** Classification and physico-chemical properties of biomolecules; (2)
- **Enzymes**: Historical perspective, general characteristic of enzymes; enzyme cofactors, prosthetic groups, coenzymes, apoenzyme, holoenzyme, metalloenzymes, proenzymes, isoenzymes, enzyme activity; specificity activity; enzyme models; enzyme nomenclature and classification; enzyme catalysis; enzyme regulation- reversible and non-reversible regulations; Enzyme kinetics- basic enzyme reactions, energy levels, Factors affecting enzyme activity (temperature, pH, enzyme concentration, substrate concentration, inhibitors and activators); Michaelis-Menten kinetics- K_m, V_{max}, K_{cat} and K_{cat}/K_m; enzyme inhibition. (4)
- **Vitamins**: General characteristics of vitamins; classification- water-soluble and fat-soluble vitamins, sources, structures, dietary requirements and deficiency conditions; coenzyme forms; Role of vitamins in metabolism; therapeutic uses of vitamins; antivitamins and hypervitaminosis. (2)
- **Hormones**: General characteristics of hormones; classification, structures, functions, mechanism of action and abnormalities of the hormones (thyroid, pancreas, hypothalamus, pituitary and gonads); Plant hormones (auxin, gibberllic, abscisic acids and cytokinins); Pheromones. (2)
- **Minerals**: Structure, classification, daily requirement, sources, uptake, transport, excretion and functions of minerals (calcium, phosphorus, magnesium, iron, iodine, zinc and copper). (1)

Cell Biology (1 credit)

- **Cellular organization of Prokaryotic and eukaryotic cells**: Cell types; Cell theory; Structure and function of organelles- Double membrane bounds organelles (Nucleus, mitochondria, chloroplasts); Single membrane bounds organelles (endoplasmic reticulum,

Golgi apparatus, lysosomes, peroxisomes, vacuoles) and Non-membrane bounds organelles (Ribosomes, nucleoles and centrioles). (4)

- **Membrane structure and function**: Structure, transports, channels, carriers, receptors, membrane potentials. (3)
- **Cell motility and shape**: Structure, polymerization, organization and functions cytoskeletal elements; Structure and functions of motor proteins. (2)
- **Signal transduction**: Types of signaling; signaling molecules; receptors and major signaling pathways. (3)
- **Cell cycle and its regulation**: events during mitosis and meiosis (3)

References:

Biochemistry-

- Voet, Donald, Voe Judith, Pratt, Charlotte W. Fundamentals of Biochemistry: Life at the molecular Level 2nd Edition. Publisher: Asia, John Wiley & Sons. 2006. ISBN: 0471753416.
- Nelson David L., Cox Michale. Lehninger Principles of Biochemistry 5th Edition. Publisher: New York. W. H. Freeman. 2008. ISBN 978 0716771081.
- Berg, Jeremy M, Tymoczko, John L. Stryer, Lubert. Biochemistry 6th Edition. Publisher: New York : W.H. Freeman. 2007.ISBN: 071676766X.
- Hames David, Hooper Nigel. Instant Notes in Biochemistry 3rd Edition. Publisher . Nodia, Taylor & Francis. 2007. ISBN: 185996 2491.
- Zubay, Geoffrey.Biochemistry 4th Edition, Publisher: Boston,Wm C. Brown, 1998, ISBN 0697219003.
- Horton, Robert, Moran, Laurence A, Scrimgeour, Perry Gray Marc, Rawn.David. Principles of biochemistry. Publisher: New Jersey, Pearson Prentice Hall, 2006. ISBN: 0131453068.
- Mathews, Christopher K. van Holde, K. E., Ahern Kevin G. Biochemistry. Publisher: San Francisco, Pearson Education 2000.ISBN: 81 29702150.
- Garrett, Reginald H., Grisham, Charles M. Principles of biochemistry: with a human focus. Publisher: Australia Brooks/Cole, Thomson Learning, 1997. ISBN:0030973694.

Cell Biology

- Essential Cell Biology (3rd edition), by B. Alberts, D. Bray, K. Hopkin, A. Johnson, J. Lewis, M. Raff, K. Roberts & P. Walter, Garland Science, 2010.
- Molecular Biology of the Cell (5th edition) by B. Alberts, A. Johnson, J. Lewis, M. Raff, K. Roberts & P. Walter Garland Science, 2008.
- The Cell: A Molecular Approach (5th edition) by Cooper, G.M. and Hoffman, R.E., A. S. M. Press, 2009.
- The world of the Cell (7th edition) by Becker, W.M., Kleinsmith, L.J., Hardin, J and Bertoni, G.P., Pearson / Benjamin Cummings, 2009
- Molecular Cell Biology (5th edition) by H. Lodish, A. Berk, P. Matsudaira, C.A. Kaiser, M. Krieger, M. P. Scott, S. L. Zipursky & J. Darnell, W. H. Freeman & Company, 2004.
- Molecular Biology of the Cell: A Problems Approach, (4th edition) by Wilson, J and Hunt, T., Garland Science publishers, 2002.

Semester II

BIM 201 (T+P): Structural Bioinformatics (2T+2P)

Objectives:

- Know how three dimensional structures can be analyzed for gaining insights into functions and other biological aspects
- know the computational approaches for structure analysis
- acquire knowledge of various algorithms & methods for structure prediction
- understand the principles of macromolecular interactions

Syllabus:

Theory:

- Overview of Structural Bioinformatics (2)
- Prediction of protein structure
  - secondary structure prediction methods (4)
    - First, second and third generation methods
  - Tertiary structure prediction
    - Homology modeling (4)
    - Fold Recognition: (5)
      - 1D-3D Profile-based methods,
      - Threading methods
    - *ab initio* methods (4)
- Structural alignments of proteins (4)
  - Superimposition of structures & calculation of RMSD
  - Vector-based, distance matrix-based and combined algorithms for structural alignments
- Structure-based classification of proteins: SCOP & CATH (2)
- Prediction of binding pockets on protein structures (3)
- Structure-based function Prediction (2)
- Prediction of RNA structures (2)

Practicals:

- Write codes for Calculation of inter-atomic distances (4)
  - to find chain breaks in polypeptides
  - to detect presence of non-bonded interactions
- Write codes for computation of dihedral angles (3)
- Write code for 4^{th} atom fixing (3)
- Prediction of Secondary structures of proteins using various methods and computation of prediction accuracies (2)
- Prediction of tertiary structures of proteins using Homology Modeling Protocol/s and validation of templates as well as predicted structures (4)
- Prediction of tertiary structures of proteins using Fold Recognition methods (2)
- Prediction of tertiary structures of proteins using *ab initio* methods (2)
- Write code for construction of distance matrix using atomic coordinates (e.g. C^{α} - C^{α} distance matrix) and analysis for detection of tertiary interactions (2)
- Detection of structural similarities among proteins and finding remote homologues using protein structural alignment methods (2)
- Searching and browsing structural classification databases (2)

- Prediction of surface pockets of proteins using various methods (2)
- Prediction of RNA structures using various methods (2)

References:

- Forbes Burkowski. Structural bioinformatics: An algorithmic approach. Publisher: CRC Press, 2009. ISBN: 9781584886839.
- Drenth Jan. Principles of Protein X-Ray Crystallography. Publisher: Netherlands, Springer Science. 2007. ISBN: 9780387333342.
- Bourne Philip E., Weissig Helge. Structural Bioinformatics (Methods of Biochemical Analysis, V. 44), 2003. Publisher: Wiley-Liss. ISBN: 0471202002.
- Höltje Hans-Dieter, Sippl Wolfgang, Rognan Didier, Folkers Gerd. Molecular Modeling: Basic Principles and Applications. Publisher: New York, Wiley-VCH. 2003. ISBN: 3527305890.
- Leach, Andrew. Molecular Modelling: Principles and Applications. Publisher: Prentice Hall. 2001. ISBN: 0582239338.
- Friesner Richard A. Computational Methods for Protein Folding:advances in Chemical Physics Volume 120 Kindle Edition. Publisher: New York, John Wiley & Sons. 2002. ISBN: 0471209554.
- Heilmeyer L., Friedrich P. Protein Modules in Cellular Signalling. Publisher: Amsterdam, IOS Press. 2001. ISBN: 1586031805.
- Rhodes Gale.Crystallography Made Crystal Clear, Third Edition: A Guide for Users of Macromolecular Models. Publisher: USA, Academic Press 2000 ISBN: 0125870728.
- Branden ,Tooze John. Introduction to Protein Structure. Publisher: New York, Garland Publishing Inc. 1999. ISBN: 0815323050.
- Hill H.A.O. Sadler P.J., A.J. Ed. Metal Sites in Proteins and Models Redox Centres Publisher: New York, Springer 1999. ISBN: 3540655564.
- Sternberg Michael J. E. Protein Structure Prediction: A Practical Approach. Publisher: USA, Oxford University Press. 1997. ISBN: 0199634953.
- Fasman G.D. Prediction of Protein Structure and the Principles of Protein Conformation. Publisher: New York, Plenum Press. 1989 ISBN: 0306431319.
- Creighton T. E. Editor. Protein Structure: A Practical Approach. Publisher: IRL Press at Oxford University Press. 1989. ISBN: 0199630011.

BIM 202 (T+P): Chemoinformatics **(2T+2P)**

Objectives:
This course will enable the students to:

- understand the representation and use of chemical information on computer
- appreciate the complementary aspects of cheminformatics and bioinformatics

Theory
Syllabus:

- Representation of chemical compounds: 1D, 2D, 3D (7)
  - SMILES, InChI, Fingerprints (Daylight), Matrices, Connection Tables, mol & sdf files, Markush structures, Rotatable bonds and conformers, Molecular surfaces
- Representation of reactions: (3)
  - SMARTS, SMIRKS, Matrices
- Searching chemical structures: (6)
  - Exact (Canonicalizing SMILES: Morgan, CANGEN), Substructure, Superstructure and Similarity search (Tanimoto, Euclidean and Tversky index)
  - Clustering small molecules : Hierarchical and non-hierarchical methods
  - 3D pharmacophore based searching: Common Pharmacophore features, building hypotheses and searching databases
- Chemical Databases: CSD, Pubchem, other relevant databases (5)
  - Information content and applications
  - Data mining
- Quantitative Structure Activity Relationship (QSAR) (4)
  - Structure descriptors: topological and shape indices
  - Training and test data set
  - Applicability domain
  - 2D QSAR
- Combinatorial chemistry and Library Design (3)
  - Historical methods: Mix and Split library design and Iterative deconvolution
  - Computer based methods
  - Diversity analysis
  - Lipinski Rules
- Tools for cheminformatics algorithm development (2)
- Future perspectives (1)

Practicals
Objectives:
This course will enable the students to

- understand the representation and use of chemical information on computer
- using chemical databases
- introduction to analysis and predictive modeling

Syllabus:

- Practical sessions on structure storage and representation in various formats: SMILES, InChI, mol & sdf files (2)
- Practical sessions on understanding reaction transforms and applying them to various small molecules (2)
- Building a database of small molecules (1)

- Structure searching: Substructure, Superstructure and Similarity searches (1)
- Exploring and data mining of : CSD and Pubchem (3)
- Pharmacophore hypothesis and searching (2)
- Descriptors and 2D QSAR studies (2)
- Tools for cheminformatics algorithm development (2)

References:

Books

- Gasteiger Johann, Engel Thomas. Chemoinformatics: A Textbook. Publisher: Wiley-VCH; 1st edition. 2003. ISBN: 3527306811.
- Leach Andrew R., Valerie J. Gillet. An introduction to chemoinformatics. Publisher: Kluwer academic , 2003. ISBN: 1402013477.
- Gasteiger Johann, Handbook of Chemoinformatics: From Data to Knowledge (4 Volumes). Publisher: Wiley-VCH. 2003. ISBN:3527306803.
- Muthukumarasamy Karthikeyan, Renu Vyas. Practical Chemoinformatics. Publisher: Springer. 2014. ISBN: 9788132217794
- Bajorath Jürgen. Chemoinformatics and computational chemical biology. Publisher: Humana Press. 2011. ISBN: 9781607618386

Other learning resources:

- Indiana Cheminformatics Education Portal (http://icep.wikispaces.com/)
- Henry Stewart Talks Introducing Cheminformatics

BIM 203 (T+P): Genomics **(2T + 2P)**

Objectives:

This course will enable the students to:

- Understand concepts in genome organization, genome packaging.
- Understand concepts in recombinant DNA technology and DNA markers.
- understand genome sequencing, assembly, alignment and annotation etc. and related technologies.

Syllabus

Theory

- Genome projects – importance, objectives and strategies (1)
- Genome markers and mapping: STS, EST, RFLP, SNP (4)
- Genome sequencing – First, Second and Next generation sequencing platforms (5)
- Genome sequence assembly (5)
  - Output file formats
  - Quality assessment of sequence reads
  - Removal of adapter contamination
  - Reference-guided and *de novo* Assembly: Approaches and Tools
  - Visualization of alignments in Reference-guided Assembly
  - Removal of redundancy
  - Archival of assembled sequences in databases
- Genome Annotation (9)
  - Basic Aspects of Genome Annotation
  - Prediction of ORFs
  - Algorithms for gene prediction & gene modeling
  - Prediction of Promoters, splice sites, UTRs etc
- Structural Genomics (2)
- Functional Genomics (2)
- Gene Ontology (2)
  - Basic concepts
  - GO terms and relationships (DAGs)
  - Tools for GO data: (AmiGO, BioConductor)
  - GO Browsers and editors
  - GO visualization tools
  - Microarray-related tools
  - Protein interaction related tools
  - Statistical analysis
  - Term enrichment tools

References:

- Brown, T.A. Genomes 2 Publisher: New York, BIOS Scientific Publishers Ltd. 2002, ISBN: 1859960294
- Old R.W. and Primrose S.B. Principles of Gene Manipulation: An Introduction to Genetic Engineering Publisher: University of California Press, 1980
- Benjamin Lewin, Jocelyn E. Krebs, Elliott S. Goldstein, Stephen T. Genes XIth Edition. Publisher: Kilpatrick Jones & Bartlett Publishers, 2014.

Practicals:

Objectives:

- This course intend to introduce the students with genome browsers and databases, assembly softwares and annotation procedures.
- Hands-on exercises will help the students to understand steps involved in comphrensive genomic study of an organism

Syllabus:

- Browsing & viewing genome data (3)
  - Ensembl@EBI
  - MapViewer@NCBI
  - UCSC Genome Browser

  Visualization, Browsing and searching of
  - Protein coding and non-coding Genes, ESTs, STSs, Retrotransposons, RFLPs, SNPs
  - regions exhibiting Synteny
- Exploration of Genome Databases (4)
- Genome assembly (8)
- Genome Annotation
  - Using integrated genome annotation servers such as the server developed at IMTech, Chandigarh (http://imtech.res.in/raghava/gp.html) (2)
  - Gene Prediction and Gene Modeling
  - Prediction of genes and gene structures (gene modeling) using online (web) servers of different methods tailored for prokaryotic and eukaryotic organisms such as GLIMMER, GeneMark, Grail, GENSCAN etc. Interpretation of results and comparison with known gene models (where available). Evaluation of accuracy of the methods. (4)
  - Prediction of promoters using methods such as Neural Network Promoter Prediction (NNPP) at Berkeley Drosophila Genome Project server, Genome inspector for combined analysis of multiple signals in genomes etc. Using Promoter databases. (2)
  - Prediction of alternate splice sites using methods such as Splice Site Prediction by Neural Network (at Berkeley Drosophila Genome Project server), GenScan, NetGene2, GeneSplicer etc. Prediction of PCR primers using Primer 3, ePCR etc. (2)
- Structural Genomics: PSI Knowledgebase, VIPs (2)
- Functional Genomics (3)
  - Using sequence -based and structure-based Function Annotation Servers such as
    - ProKnow (http://www.doe-mbi.ucla.edu/Services/ProKnow/)
    - Joined Assembly of Function Annotations (JAFA) at http://jafa.burnham.org/learnMore.html etc. which are integrated services for function annotation
    - ProFunc (http://www.ebi.ac.uk/thornton-srv/databases/ProFunc)

BIM 204 (P): Introduction to Database Systems (2P)

Objectives:

This course will enable the students to:

- understand the concepts of data, data models and relationships be aware of various data representation techniques and various types of databases
- appreciate and implement relational database design create the database systems with user-friendly front-ends for fast and efficient data retrieval and storage
- acquire the skills of using MySQL, SQL and basic skills in creating front end applications

Syllabus:

- Database Management System basic concepts (1)
- Data abstraction, Data Models, Instance & Schema (2)
- Data Normalization (1)
- Overview of MySQL and its features (1)
- Installing MySQL Server and MySQL Workbench (1)
- Data Definition Language (DDL) (10)
    - Creating and altering database tables
    - Defining primary keys, foreign keys, and constraints
    - Understanding data types in MySQL
- Data Manipulation Language (DML) (10)
    - Inserting data into tables
    - Updating and deleting data
    - Retrieving data using SELECT statements
    - Arithmetic and logical operators.
    - Grouping and aggregate functions
    - Filtering grouped data with HAVING clause
    - Pattern matches
    - Performing subqueries for complex queries
- Joins (1)
- Importing and Exporting Data (2)
    - Importing data from CSV and other formats
    - Exporting query results to files
    - Backup and restore databases
- Advanced options: Defining Indices, Management using Grant and Revoke (1)

References:

- Silberschatz, A, Korth, H F & Sudarshan, S. Database system concepts; McGraw-Hill higher education, 2002. ISBN: 0071148108.
- Date C. J. An Introduction to Database Systems. 1999. Publisher: Addison Wesley. ISBN: 0201327546.
- Bayross Ivan. SQL, PL/SQL The Programming Language of Oracle. 2nd revised edition. Publisher: BPB ISBN: 8176560723.
- Ramez Elmasri, Shamkant Navathe, Fundamentals of Database Systems. Publisher: Addison Wesley. ISBN: 0136086209
- Vikram Vaswani, MySQL: The Complete Reference. Publisher: Tata Mcgraw Hill. ISBN: 0070586845

BIM 205 (P): On Job Training / Field Project **(4P)**

Student will perform Internship/ Apprenticeship with a company related to Bioinformatics / Computational Biology. Alternately the student will engage in a Field Project that addresses the use of computational biology to solve applied problems.

Elective Courses

Courses worth of 4 credits from among the following:

BIE 206 (P): Biological Data Curation and Analysis **(2P)**

Objectives:

This course will enable the students to:

- Understand the Biological data as Big Data.
- Understand the significance of data curation.
- Understand and use the web services provided by the biological databases.
- Programmatic access to biological resources.
- Write advanced codes for biological data curation and analysis.
- Design workflows for data processing.

Syllabus:

- Biological Data (1)
  - Types, complexity & big data
- Biological data curation (1)
  - Need, issues & challenges
- Using APIs (5)
  - Introduction to APIs and their importance
  - Making API requests with Python
  - Parsing data from API responses
  - Web scraping and its applications
  - Use of requests, BeautifulSoup modules
- Web services of commonly used Bioinformatics databases: (10)
  - NCBI: E-utilities
  - EMBL-ENA: Programmatic access using ENA browser REST URLs
  - Uniprot: Uniprot resources programmatically access
  - PDB: The RCSB PDB RESTful Web Service interface
  - KEGG: KEGG API
- Data Analysis with Pandas: using data frames, data cleaning, filtering (2)
- Biopython: Introduction, installation, important components like seq, seqIO, alignIO, BLAST, ClustalW, PDB, SwissProt etc., parsing output (6)
- Workflows in Bioinformatics (5)
  - Known workflow environments
  - Case studies: Existing workflows (ex: Immunoinformatics)

- Workflows design & implementation
  - Development of one server

References:

- Web services documentation of various biological databases & resources
- Review and research articles on Big data and data curation & web services
- Agostino, Michael, Practical Bioinformatics. Publisher: Garland Science, 2012. ISBN: 9780815344568
- Sung Wing-Kin, Algorithms in Bioinformatics: pratical introduction. Publisher: CRC Press, 2010. ISBN: 978 1420070330

BIE 207 (P): Advanced Techniques in Experimental Biology (4P)

Objectives:

This course will enable the students to:

- gain hands-on experience with some of the laboratory techniques necessary to understand genomics and proteomics.

Practical:

A. Microbiological techniques: (15)
   - Media preparation, Sterilization, Inoculum preparation.
   - Staining: Gram staining, endospore staining.
   - Streak, pour and spread plate methods.
   - Growth curve and kinetics
   - Screening of antibiotic producer, amylase producers, PHA producers
   - Fermentation

B. Biochemical techniques: (30)
   - Extraction and fractionation methods- Cell lysis, protein precipitations, dialysis, density gradient centrifugation/differential centrifugation, ultracentrifugation
   - Electrophoresis- Native-PAGE, SDS-PAGE, 2DE, 2D-DIGE, agarose gel
   - Chromatography- AC, IEC, TLC, HPLC, GC, gel filtration and paper chromatography
   - Isolation and Purification of enzymes (Invertase, Amylase, Alkaline Phosphatase, Acid Phosphatase etc.)
   - Enzymes kinetics- Enzyme activity, factors affecting enzyme activity (pH, Temp., Substrate, Inhibitors, Activators), Determination of pKa, pI, K_{cat}, K_m, K_I,$V_{max.}$
   - Other techniques- Western blot, In-gel digestion, metabolism assays, gel image analysis, CD/NMR/MS spectra

C. Molecular biology techniques: (15)
   - Bacterial Transformation
   - Plasmid isolation
   - Restriction digestion
   - *Polymerase Chain Reaction*
   - Gel Elution of DNA
   - TA/TOPO-TA cloning of PCR product

References:

- Microbiological Methods (Eighth edition) by C.H. Collins and P.M. Lyne, Arnold, a member of the Hodder Headline Group,338 Euston Road, London NW1 3BH
- Medical Microbiology: The practice of medical microbiology. Vol. 2 (Twelth edition) by Robert Cruickshank, Churchill Livingstone, 1975.
- Practical Handbook of Microbiology (Second Edition) by Emanuel Goldman, Lorrence H Green, CRC Press, 2008.
- Biochemical Techniques: Theory and Practice by John F. Robyt and Bernard J. White, Waveland Press, 1987.
- Introduction to Practical Biochemistry (Third edition) by Plummer Mu, David T. Plummer, Tata McGraw-Hill Education, 1988 - 332 oldal.

- Principles and Techniques of Practical Biochemistry (7th Edition) by Keith Wilson and John Walker, Cambridge University Press.
- Introductory Practical Biochemistry by S.K. Sawhney and R. Singh, Narosa Publishing House, 2000.
- Biochemical Methods (Second edition) by S Sadasivam, A. Manickam, New Age International (P) Ltd. Publisher, 1996.

BIE 208 (T+P): Molecular Phylogenetics (1T + 1P)

Objectives:

This course will enable the students to:

- understand concepts of molecular evolution and the nature of data for deriving molecular phylogeny
- learn and apply the statistical approaches and models for phylogenetic analysis and tree reconstruction

Theory

Syllabus:

- Molecular phylogenetics: Overview (2)
  - Concepts in molecular evolution
  - Nature of data
- Molecular Phylogeny (5)
  - Concept & overview
  - Distance-based methods: UPGMA & NJ
  - Character-based methods: Maximum Parsimony
- Probabilistic models and associated algorithms (3)
  - Probabilistic models of evolution
  - Maximum likelihood algorithm
- Phylogenetic analysis algorithms (5)
  - Distance-based: UPGMA, Transformed Distance, Neighbors-Relation, Neighbor-Joining
  - Character optimization; delayed and accelerated transformation
  - Maximum Parsimony
  - Reliability of trees: Bootstrap, jackknife, decay, randomization tests.
- Phylogenetic trees and their comparison: (2)
  - Definition and description, various types of trees;
  - Consensus (strict, semi-strict, Adams, majority rule, Nelson).
  - Data partitioning and combination.
  - Tree to tree distances, similarity
- Case studies (2)
- Numerical taxonomy & bacterial identification (1)

Practicals

Objectives:

This course will enable the students to:

- use computational approaches for phylogenetic analysis
- explore and use packages available for molecular phylogeny

Syllabus:

- Compilation & curation of dataset, format conversion (1)

- Survey of software programs available for phylogenetic analysis (1)
  - Installation of at least 2 public domain packages for both Windows & Unix environment: Phylip, PAUP, MEGA
- MSA using MUSCLE (1)
  - Informative and variable sites; Singleton sites etc.

- Reconstruction of phylogenetic trees using distance-based methods (1 datasets) (3)
    - Converting sequence data into distance data
    - UPGMA, Neighbor-joining, Neighbor-relations & Transformed distance
    - Difference between dendogram & phylogenetic tree

- Reconstruction of phylogenetic tress using character-based methods
    - Maximum Parsimony (1)
    - Maximum likelihood (1)

- Using bootstrapping tool to generate multiple datasets from the original input data & generation of consensus tree (1)
- Plotting, visualizing & printing phylogenetic trees: TreeView and other tools (1)
    - Various rendering
    - Formatting & labeling
    - Interpretation of trees
- Comparison of trees drawn using (2)
    - RNA, Nucleotide & protein data
    - Gene tress & species tree
- Un-rooted & rooted tree (1)
    - Rooting un-rooted tree using an out group
- Reconstruction of phylogenetic trees using whole genome data of viruses (1)

- Numerical taxonomy & Bacterial identification using matrices (1)

References:

- Bromham Lindell. Reading the Story in DNA: A Beginner's Guide to Molecular Evolution. Publisher: USA, Oxford University Press. 2008. ISBN: 9780199290918.
- Bernardi Giorgio. Structural and Evolutionary Genomics, Volume 37: Natural Selection in Genome Evolution (New Comprehensive Biochemistry). Publisher: Netherlands, Elsevier Science.2005. ISBN: 9780444521361.
- Marco Salemi, Anne-Mieke Vandamme. The phylogenetic handbook: a practical approach to DNA and protein phylogeny. Publisher: Cambridge University Press, 2003. ISBN: 052180390X.
- Patthy Laszlo. Protein Evolution. Publisher: London, Blackwell Science Ltd. 1999. ISBN: 0632047747.
- Takahata Naoyuki, Clark Andrew G.(Editor). Mechanisms of Molecular Evolution: Introduction to Molecular Paleopopulation Biology. Publisher: Japan, Japan Scientific Socities Press and Sinauer Associates, Inc. 1993. ISBN: 476226718X.
- Graur Dan, Wen-Hsiung Li. Fundamentals of molecular evolution Publisher: Sinauer Associates, 1991. ISBN: 0878932666.

BIE 209 (T+P): Programming in Java **(2T + 2P)**

***Objectives*:**
The course will enable the students to:

- Acquire skills in programming using object oriented language: Java
- Develop and implement programs to analyze biological data

***Syllabus*:**

- An introduction to Java programming
- Features of Object-Oriented programming
- Java Basics
- Working with objects
- Arrays & Control structures
- Modifiers, Access control and class design
- Working with data structures
- Packages & Interfaces
- Using Native methods & Libraries
- Exception handling in Java
- Multithreading
- File I/O
- Brief Introduction to Applets & AWT

***References*:**

- Java In a Nutshell by David Flanagan, Oreilly Publications.
- Java Examples in a Nutshell by David Flanagan, Oreilly Publications.
- Java 2: The Complete Reference by Patrick Naughton and Herbert Schildt, McGraw Hill
- Horstmann Cay S.Cornell Gary. Core Java Volume II Advanced Features. Publisher: New Delhi, Education
- Schildt Herbert. Java 2: The Complete Reference Publisher: New Delhi, Tata McGraw-Hill Publishing Company.
- Bruce Eckel. Thinking in Java. Publisher: Pearson Education Flanagan David. Java in a Nutshell: A Desktop Quick Reference. Publisher: Mumbai, Shroff Publishers & Dist
- Balagurusamy E. Programming With Java A Primer. Publisher: New Delhi, Tata McGraw Hill

BIE 210 (P): Linux Operating System **(2P)**

Objectives:

This course will enable the students to:

- Writing basic and advanced Linux/Unix commands
- Understand and write shell scripts

Syllabus:

- Introduction to Unix/Linux, Linux distributions (1)
- Using graphical and command line environments (1)
- Linux file system and file/directory manipulation commands (3)
- File permissions and commands to modify permissions (1)
- File manipulation commands, redirection, pipes (4)
- Searching and filtering files, grep command (4)
- Monitoring system resources (1)
- Text editors: basic editing and navigation (2)
- Shell Basics, writing and executing scripts (1)
- Conditional statements: If-else-elif, test, logical operators-AND, OR, NOT (3)
- Loops: while, for, until, break & continue (3)
- Command line arguments (1)
- Advanced commands: process management, package management, sed and awk commands, archiving and compressing files (5)

References:

- Richard Blum and Christine Bresnahan, Linux Command Line and Shell Scripting Bible, 3ed, Wiley publishers. ISBN: 8126554983
- Richard Petersen, Linux: The Complete Reference, Sixth Edition, McGraw Hill Education. ISBN: 0070222940
- Sumitabha Das, Unix Concepts And Applications, McGraw Hill Education. ISBN: 0070635463

BIE 211 (T+P): Software Testing **(1T + 1P)**

Objectives*:*

- Enable students to become aware of errors and inaccuracies in software programs and software routines.
- To learn the process of making error free new or existing computing software programming systems/packages.
- To optimize the performance of software programming systems/packages.

Aim:

To enhance the quality of source code, software program, web-enabled databases released in open source or charged environment.

Syllabus:

- Principles of testing- test-case, test scenarios, different methods in testing. (1)
- Principle of Automation Testing (1)
- JUnit testing (Brief introduction) (1)
- Types of testing: black box testing, white box testing. (1)
- Defect life cycle, STLC (1)
- Different methodologies for: V model, water fall, Agile, continuous delivery, test driven development, extreme programming. (3)
- Databases Testing (1)
- Fundamentals of Server-side Testing: What is server, server types, types of testing servers (2)
- Web-services testing. (1)
- Algorithmic Testing (1)
- Some existing tools used in Software Testing industry (2)

Practicals:

- Perform testing of existing well known General Purpose applications, Life Sciences/ Bioinformatics applications. Store and analyze the results.
- Testing of Application of languages/interpreters (perl, Java, etc) and other standard services / desktop programs**.** (15)