

## Course Contents of M.Sc. and Ph.D., Bioinformatics

### I TRIMESTER

<b>BI 501/ MBB 509/ GP 540</b>	INTRODUCTION TO BIOINFORMATICS	2	1
<b>BI 505/ CA 561</b>	PRINCIPLES OF COMPUTER PROGRAMMING	2	1
<b>BI509/BIO601</b>	NUCLEIC ACIDS	2	1
<b>BI 510/ MBB 501</b>	PRINCIPLES OF BIOTECHNOLOGY	4	0
<b>BI 511 /BIO 501</b>	BASIC BIOCHEMISTRY	4	1
<b>BI 512</b>	ADVANCED PROGRAMMING IN BIOINFORMATICS	2	2
<b>BI 513/ GP 500</b>	ELEMENTS OF GENETICS	3	2
<b>BI 514</b>	STATISTICAL TECHNIQUES IN BIOINFORMATICS	3	1
<b>BI 601</b>	GENOME ASSEMBLY AND ANNOTATION	1	2
<b>BI 602</b>	BIOMOLECULAR MODELLING AND SIMULATION	2	1
<b>BI 611</b>	METAGENOMICS DATA ANALYSIS	2	1
<b>BI 622</b>	MOLECULAR DYNAMICS	2	1
<b>BI 631</b>	PARALLEL PROGRAMMING AND ALGORITHM DEVELOPMENT	2	1
<b>BI 641</b>	BIOLOGICAL NETWORK MODELLING AND ANALYSIS	2	1
<b>BI 651</b>	RECENT ADVANCES IN BIOINFORMATICS	1	0
<b>BI 691</b>	SEMINAR	1	0

### II TRIMESTER

<b>BI 506</b>	COMPUTATIONAL GENOMICS	3	1
<b>BI 507 /CA 566</b>	DATABASE MANAGEMENT SYSTEM	2	2
<b>BI 508</b>	COMPUTER APPLICATIONS IN BIOINFORMATICS	2	1
<b>BI 603</b>	MACHINE LEARNING TECHNIQUES IN BIOINFORMATICS	2	1
<b>BI 604</b>	COMPUTATIONAL TECHNIQUES OF TRANSCRIPTOMICS AND METABOLOMICS	1	1
<b>BI 623</b>	PEPTIDE DESIGN, SYNTHESIS AND APPLICATIONS	2	1
<b>BI 632</b>	OPTIMIZATION TECHNIQUES IN BIOINFORMATICS	2	1
<b>BI 642</b>	GENOME WIDE ASSOCIATION STUDY	2	1
<b>BI 651</b>	RECENT ADVANCES IN BIOINFORMATICS	1	0
<b>BI 691</b>	SEMINAR	1	0

### III TRIMESTER

<b>BI 502</b>	PROTEIN STRUCTURE ANALYSIS	2	1
<b>BI 503</b>	COMPUTATIONAL BIOLOGY	2	1
<b>BI 504</b>	EVOLUTIONARY BIOLOGY	2	1
<b>BI 621</b>	QUANTUM THEORY AND APPLICATIONS IN BIOLOGY	2	1
<b>BI 633</b>	BIOLOGICAL DATA INTEGRATION AND QUALITY CONTROL	2	1
<b>BI 634</b>	GRAPHICS AND VISUALIZATION OF BIOLOGICAL DATA	2	1
<b>BI 635</b>	BIOLOGICAL DATABASES AND DATA ANALYSIS	2	1
<b>BI 651</b>	RECENT ADVANCES IN BIOINFORMATICS	1	0
<b>BI 691</b>	SEMINAR	1	0

## **BI 501/MBB 509/ GP 540: Introduction to Bioinformatics (2+1)**

### **Objective**

To introduce the basics of biological data resources, tool and techniques commonly used in bioinformatics.

#### **UNIT I**

Basic molecular biology; introduction to the basic principles of structure/function analysis of biological molecules; genome analysis; different types and classification of genome databases (e.g. HTGS, DNA, Protein, EST, STS, SNPs, Unigenes etc.).

#### **UNIT II**

Role of bioinformatics in genomics; Nature of genomic data; Overview of available genomic resources on the web; NCBI/ EBI/ EXPASY etc; Nucleic acid sequence databases; GenBank/EMBL/ DDBJ; Database search engines: Entrez, SRS

#### **UNIT III**

Overview/concepts in sequence analysis; Pairwise sequence alignment algorithms: Needleman & Wunsch, Smith & Waterman ; BLAST, FASTA; Scoring matrices for Nucleic acids and proteins: PAM, BLOSUM, Dynamic Programming Algorithm, Multiple sequence alignment: PRAS, CLUSTALW

#### **UNIT IV**

Sequence based gene prediction and its function identification, Use of various derived databases in function assignment, use of SSR, SNPs and various markers for identification of genetic traits, Gene Expression

### **Practical**

Exposure of different types of databases, database search and retrieval, DNA and sequence analysis, Working with nucleic acid sequence databases, Protein sequence databases, Database search engines, Database Similarity Searches, Multiple sequence alignment, Genome databases, Structural databases, Derived databases, Gene annotation, Gene prediction software.

### **Suggested Readings**

Andreas Baxevanis and B.F. Francis Ouellette. 2004. Bioinformatics approach Guide to the analysis of genes and proteins. John Wiley.

Campbell, A.M. & Heyer, L.J. 2002 Discovering Genomics, Proteomics and Bioinformatics. Benjamin/Cummings.

David Mount. 2004. Bioinformatics: sequence and genome analysis. Cold Spring Harbour Press.

Sankoff, D. & Nadeau, J.H. 2000. Comparative genomics: empirical and analytical approaches to gene order dynamics, map alignment and the evolution of gene families. Netherlands, Kluwer Academic Publishers.

Jonathan Pevsner, 2009. Bioinformatics and Functional Genomics. Wiley Blackwell

Yang, M.C.C. 2000. Introduction to Statistical Methods in Modern Genetics. Taylor and Francis.

## **BI 502: Protein Structure Analysis (2+1)**

### **Objective**

To impart knowledge about the online protein data resources and various tools and techniques used in the analysis of protein structure and functions.

#### **UNIT I**

Nature of proteomic data; Overview of protein data bases; SWISSPROT, UniProtKB; PIR-PSD, PDB, Prosite, BLOCKS, Pfam/Prodom etc.; Structure analysis: Exploring the Database searches on PDB and CSD, WHATIF Molecular visualization tools; Visualization of tertiary structures, quaternary structures, architectures and topologies of proteins using molecular visualization softwares such as RasMol, Cn3D, SPDBV, Chime, Mol4D etc.

#### **UNIT II**

Structure prediction tools and homology modeling: Prediction of secondary structures of proteins using different methods with analysis and interpretation of the results; Comparison of the performance of the different methods for various classes of proteins. (Fasman method, Garnier Osguthorpe Robson (GOR), Neural Network based; methods); NLP approach for secondary structure prediction of RNA; Introduction to mfold and Vienna packages; Prediction of tertiary structures of proteins using Homology Modeling approach: SWISSMODEL, SWISS-PDB Viewer; along with analysis and interpretation of results. Molecular dynamics simulation and docking

### **Practicals**

Protein database search and Protein sequences retrieval from databases, Structural data, databases and structure analysis, Molecular visualization tools, Protein structure prediction with different tools and server, Molecular docking, Protein-protein interaction analysis.

### **Suggested Readings**

Baxevanis, A.D. and Francis Ouellette, B.F. 2004. Bioinformatics: A Practical Guide to the Analysis of Genes and Proteins. Wiley.

Gimona, G. Cesareni. & Yaffe, M. Sudol (EDS.) Aug 2004. Modular protein Domains. Wiley-vch verlag gmbh & co., 3-527-30813-X.

Graur, D. and Li, W-H. 2000. Fundamentals of Molecular Evolution. Sinauer Ass., USA.

Hans Dieter & Didier Rognan. 2003. Molecular Modeling: Basic Principles and Application. Wiley VeH Gmbh and Co. KGA.

Holtje, H.D. & Folkers, G., Weinheim. 1997. Molecular modeling: Basic Principles and Applications. VCH.

Webster, D. M. Ed. 2000. Protein Structure Prediction: Methods and Protocols. Totowa Humana Press.

Wilkins, M.R., Williams, K.L., Appel, R.D., Hochstrasser, D.F. (Editors) 1997 Proteome Research: New Frontiers in Functional Genomics. Springer Verlag Berlin Heidelberg

## **BI 503: Computational Biology (2+1)**

### **Objective**

To provide theoretical and practical knowledge about handling and processing of genomic data, optimization and data mining techniques used in bioinformatics.

#### UNIT I

Preprocessing of gene expression data; Data Normalization techniques, Data quality control: Modelling of errors, Imputation etc; High-throughput screening;

#### UNIT II

Optimization Techniques: concept and applications, Simulated Annealing, Genetic Algorithms: Ab initio methods for structure prediction; Information theory, entropy and relative entropy

#### UNIT III

Foundations for Machine learning Techniques: Unsupervised and Supervised Learning, Cross Validation Techniques, Markov Model, Hidden Markov Model and Application, Bayesian Inference: concepts and applications , Introduction to WEKA package;

### **Practicals**

Analysis of DNA microarray experiments, Expression profiling by microarray/gene chip, Proteomics, Pattern recognition, Hidden Markov Models, Neural Networks, Genetic algorithms, Bayesian techniques and use of Gibbs Sampling, Analysis of single and multiple DNA or protein sequences, Computationally intensive methods.

### **Suggested Readings**

- Baldi, P. and Brunak, S. 2001. Bioinformatics: The Machine Learning Approach. MIT Press.  
Baxevanis, A.D. and Francis, B.F. 2004. Bioinformatics: A Practical Guide to the Analysis of Genes and Proteins. John Wiley  
Wang, J.T.L., Zaki, M.J., Toivonen, H.T.T. and Shasha, D. 2004. Data Mining in Bioinformatics. Springer  
Amaratunga, D. & Cabrera, J. 2004. Exploration and Analysis of DNA Microarray and Protein Array. John Wiley.  
Gupta, G. K. 2006. Introduction to Data Mining with Case Studies. Prentice Hall of India, New Delhi.  
Han, J. and Kamber, M. 2006. Data Mining: Concepts and Techniques. Morgan Kaufman.  
Hand, D., H. Mannila, P. Smyth. 2001. Principles of Data Mining. Prentice Hall of India, New Delhi

## **BI 504: Evolutionary Biology (2+1)**

### **Objective**

To find out the evolutionary relationship among various species by using different phylogenetic techniques and algorithms.

#### UNIT I

Phylogenetic trees and their comparison: Definition and description, various types of trees; Consensus (strict, semi-strict, Adams, majority rule, Nelson); Data partitioning and

combination Tree to tree distances, similarity; Phylogenetic analysis algorithms: Maximum Parsimony, Distance based: UPGMA, Transformed Distance, Neighbors-Relation, Neighbor-Joining.

#### UNIT II

Probabilistic models of evolution, Maximum likelihood algorithm; Approaches for tree reconstruction: Character optimization; delayed and accelerated transformation, Reliability of trees, Bootstrap, jackknife, decay, randomization tests; Applications of phylogeny analyses: Comparison of Phylogenetic Trees obtained using DNA seq. vs. protein seq. vs. Full genomes. Need for addition of other properties towards total phylogenetic analysis, Comparative methods for detection of species / organism relationships, Gene duplication, Horizontal transfer, Domain evolution, Study of co-evolution: Plant-insect interactions. Host-parasite interactions, viral evolution.

#### Practicals

Different software for phylogenetic tree construction and evolution of tree such as EMBOSS, MrBayes, PAUP, PHYLIP, PAML, TREE puzzle, Dandogram, cladogram analysis.

#### Suggested Readings

Hall, B. G. 2001. Phylogenetic Trees Made Easy: A How to Manual for Molecular Biologists. Sinauer Ass.,USA.

Nei, M. and Kumar, S. 2000. Molecular Evolution and Phylogenetics. Oxford University Press.

Sankoff, D. & Nadeau, J. H. 2000. Comparative genomics: empirical and analytical approaches to gene order dynamics, map alignment and the evolution of gene families. Netherlands, Kluwer Academic Publisher

Gustavo Caetano. 2010. Evolutionary Genomics and Systems Biology. Wiley-blackwell.

#### BI 505/ CA 561: Principles of Computer Programming (2+1)

#### Objective

The course is aimed to develop problem-solving strategies, techniques and skills, to help students develop the logic, ability to solve the problems efficiently using object oriented programming.

#### UNIT I

Techniques of problem solving, Algorithm development, Flowcharting, Stepwise refinement.

#### UNIT II

Structured programming; Object oriented programming, classes, objects, Abstract data types, Data types, Operators (Arithmetic, Logical and Comparison) and expressions.

#### UNIT III

Branching and iteration, Arrays, Object/Message paradigm.

#### UNIT IV

Data encapsulation- modules and interfaces; Polymorphism - Static and dynamic binding, Inheritance: class and object inheritance.

## UNIT V

Object oriented software design; Generic and reusable classes, Debugging and testing of programs.

### Practical

Programming constructs, control statements: branching and looping, file operations, Creation of classes with features - overloading, inheritance, data abstraction, polymorphism and a case study using and Object oriented language.

### Suggested Readings

- Arnold, Ken and Gosling, James 1996. The Java Programming Language. The Java Series. Addison Wesley.
- Balaguruswamy, E. 1998. Programming with ANSI C. Tata McGraw Hill, New Delhi.
- Balaguruswamy, E. 2001. Programming with Object Oriented Programming using C++. Tata McGraw Hill, New Delhi.
- Bergin, J. 1994. Data Abstraction: The Object-Oriented Approach Using C++. McGraw Hill.
- Sethi, R. 1996. Programming Language Concepts. Addison Wesley.
- Stroustrup, B. 1997. The C++ Programming Language. Addison Wesley.

## BI 506: Computational Genomics (3+1)

### Objective

This course builds the basic understanding of statistical methods used in genetics and genomics.

#### UNIT I

Fundamentals of Population genetics: Hardy –Weinberg law, Effect of systematic forces on changes in gene frequency; Principles of Quantitative genetics: Values, Means and Variances, Detection and Estimation of Linkage, Inbreeding, Selection, Genetic Parameter Estimation, Variance component estimation, BLUP, G x E interaction, Path Analysis

#### UNIT II

Molecular Marker based classification: similarity measures, clustering methods, bootstrapping; QTL mapping: Detection and Estimation of QTL, Single Marker Analysis, Interval Mapping and MQM;

#### UNIT III

Design and Analysis of Expression Data; Genome Selection; Genome Prediction, Association Mapping; Genome Wide Association Analysis

### Practicals

Population genetics: Hardy-Weinberg law, Estimation of linkage, Inbreeding, Selection, Genetic parameter estimation, Variance component estimation, BLUP, Path analysis,

Molecular marker based classification, Estimation of QTL, Single marker analysis, MQM, Analysis of gene expression data, Genome selection and Genome prediction.

### **Suggested Readings**

Xu, Shizhong, 2013. Principles of Statistical Genomics. Springer

Ben Hui Liu, 1997. Statistical Genomics: Linkage, Mapping, and QTL Analysis

Sorensen, D. and Gianola, D. 2002. Likelihood, Bayesian and MCMC Methods in Genetics. Springer.

Ben Hui Liu, Leming M Shi, 2013. Statistical Genomics and Bioinformatics, Second Edition

### **BI 507/CA 566: Data Base Management Systems (2+2)**

#### **Objective**

Database systems are backbone of any information system, enterprise resource planning, research activities and other activity that require permanence of data storage. This course provides the basic introduction to database system technologies; design, concurrency, security and backup/recovery issues of database management systems. The major focus in this course is the relational database model.

#### **UNIT I**

Database system - Operational Data, Characteristics of database approach, architecture.

#### **UNIT II**

Overview of DBMS; Data associations - Entities, Attributes and Associations, Relationship among Entities, Representation of Associations and Relationship, Data Model classification.

#### **UNIT III**

Entity Relationship model; Relational Data Structure- Relations, Domains and Attributes, Relational Algebra and Operations, Retrieval Operations.

#### **UNIT IV**

Relational Database Design - Anomalies in a Database, Normalization Theory, and Normal forms; Query processing and optimization; Security, backup and recovery.

#### **UNIT V**

Distributed Databases- concepts, architecture, design; Object Oriented databases; Structured Query Language (SQL) - Data Definition Language (DDL), Data Manipulation Language (DML), Query by example.

#### **UNIT VI**

PL/SQL - Stored procedure, Database triggers; Relational Data Base Management Package.

## **Practical**

E-R diagram construction; SQL - Command Syntax, Data types, DDL Statements, DML Statements, integrity constraints; Triggers, creating stored procedures/ functions; Normalization of database and Case study on a database design and implementation.

## **Suggested Readings**

Date, C. J. 2000. Introduction to Database System. Addison Wesley.  
Desai, B. C. 2000. Introduction to Database Systems. Galgotia Publications, New Delhi.  
Elmasri and Navathe. 2006. Fundamentals of Database Systems. Addison Wesley.  
Garcia-Molina, H., Ullman, J. D. and Widom J. 2002. Database Systems: The Complete Book. Prentice Hall.  
Rob, P. and Coronel, C. 2006. Database Systems: Design, Implementation and Management. Thomson Learning.  
Silberschartz, A., Korth, H. F. and Sudarshan, S. 1997. Database Systems Concepts. Tata McGraw Hill, India.

## **BI 508: Computer Application in Bioinformatics (2+1)**

### **Objective**

To understand the basics of Linux, Windows, Web servers, Networking & protocols and programming languages used in bioinformatics.

#### Unit I

Basics of operating systems (Linux and Windows), Basics of linux commands, file system hierarchy, installation of packages, overview of system administration, web servers - IIS, Apache, Tomcat; basics of LAMP/ WAMP/ XAMPP.

#### Unit – II

Cocepts of networks, protocols (http, ftp, TCP/IP) and applications (ssh, email, chat), Internet basics, TCP/IP: addressing and routing. Internet applications: FTP, Telnet, Email, Chat. World Wide Web: HTTP protocol, overview of HTML (tags and forms), Javascript, PHP and python.

#### Unit-III

Perl: Introduction, Scalar, Arrays and List Data, Control Structures, Hashes, String Handling, Regular Expressions; Subroutines, File handling, Directory Access and Formatting, CGI Programming: CGI Module, Passing Parameters via CGI and Perl, Object oriented, Creating Objects.

### **Practicals**

Basics of Linux commands, installation of packages, overview of system administration, web servers - IIS, Apache, Tomcat; basics of LAMP/ WAMP/ XAMPP, overview of protocols (http, ftp, TCP/IP) and applications (ssh, email, chat), overview of HTML (tags and forms), Javascript, PHP and python, Perl data types, Control Structures, Hashes, String Handling, Regular Expressions, Subroutines, File handling, CGI Programming.



## Suggested Readings

Petersen, R. 2007. *Linux: The Complete Reference, Sixth Edition*. McGraw-Hill Education.  
Lewis, J. R. *Linux Utilities Cookbook*. Packt Publishing, O'Reilly  
Andrew S. Tanenbaum, *Computer Networks*, Prentice Hall.  
Douglas Comer, *Internetworking with TCP/IP, Volume 1*, Prentice Hall of India.  
W. Richard Stevens, *TCP/IP Illustrated, Volume 1*, Addison-Wesley.  
James Tisdall. 2001. *Beginning Perl for Bioinformatics*. O'Reilly.  
Randal L. Schwartz, Tom Phoenix, brian d foy. 2008. *Learning Perl*. O'Reilly.

## BI 509/ BIO 601: Nucleic Acids (2+1)

### Objective

To provide knowledge about structure and function of nucleic acids and the role of modification enzymes

#### UNIT I

History of nucleic acids; DNA as genetic material.

#### UNIT II

Chemistry of nucleic acids; Chromatin structure and function; Structure and conformation of DNA and RNAs.

#### UNIT III

DNA topoisomerases, nucleases, endonucleases and related enzymes.

#### UNIT IV

Biosynthesis of nitrogenous bases, nucleotides and their regulation.

### Practicals

Isolation and purification of RNA and DNA from plant tissue; Spectrophotometric analysis and agarose gel electrophoresis; Restriction analysis of DNA; Isolation of total RNA and Ribosomal RNA; Spectrophotometric quantification; agarose gel electrophoresis; Isolation of messenger RNA.

## Suggested Readings

Adams, R.L.P., Knowler, J.T. and Leader, D.P. 1992. *The Biochemistry of the Nucleic Acids*. 11th Ed. Chapman & Hall.  
Alberts, B., Bray, D., Lewis, J., Raff, M., Roberts, K. and Watson, J.D. 2008. *Molecular Biology of the Cell*. 6th Ed. Garland Publ.  
Blackburn, G.M. and Gait, M.J. (1996). *Nucleic Acids in Chemistry and Biology* 2nd Ed. Oxford University Press.  
Freifelder, D. and Malacinski, G.M. 1996. *Essentials of Molecular Biology*, 3rd Ed. Panima.

## **BI 510/MBB 501: Principles of Biotechnology (4+0)**

### **Objective**

To provide insight into basics and application of general biotechnology

#### UNIT I

The structure of DNA; Function of genes and genomes; Restriction enzymes and vectors; Methods of recombinant DNA technology; Nucleic acid hybridization; PCR and its applications.

#### UNIT II

Genomics, transcriptomics and proteomics.

#### UNIT III

Applications of gene cloning, Molecular markers in basic and applied research;

#### UNIT IV

Genetic engineering and transgenics;

#### UNIT V

General application of biotechnology in agriculture, Medicine, Animal husbandry, Environmental remediation, Energy production and Forensics

### **Suggested Readings**

Molecular biology (2005) by David P. Clark.

Molecular biology of the Cell (2008) by Bruce Alberts.

Molecular biology and Biotechnology (2009) by John M. Walker, Ralph Rapley

Biotechnology: Expanding Horizons (2010) by B D Singh.

## **BI 511/BIO 501: Basic Biochemistry (4+1)**

### **Objective**

To provide basic knowledge/overview of structure and functional and metabolism of biomolecules.

#### UNIT I

Scope and importance of Biochemistry in Agriculture; Fundamental principles governing life; Structure and properties of water; Acid base concepts, pH and buffers; Intra- & inter-molecular forces in biomolecules; General introduction to physical techniques for determination of structure of biopolymers.

## UNIT II

Classification, structure and function of carbohydrates, lipids, amino acids, proteins, nucleic acids and vitamins.

## UNIT III

Fundamentals of thermodynamic principles applicable to biological process, bioenergetics; respiration and oxidative phosphorylation.

## UNIT IV

Classification of enzymes and their mechanism of action, regulation and kinetics.

## UNIT V

Plant and animal hormones; Metabolism of carbohydrates, lipids & proteins, DNA replication, transcription and translation.

## Practicals

Preparation of reagents and buffers, Preparation of standard acids and alkali, Estimation of protein, free amino acids, estimation of amylolytic activity, Assay of proteolytic activity, Estimation of total sugars, Reducing sugars, Non reducing sugars, starch, Extraction and estimation of oil, Iodine value, Acid value, Fatty acid by GLC, Estimation of Vitamin C, Estimation of DNA, RNA.

## Suggested Readings

Conn, E.E. and Stumpf, P.K. 1987. Outlines of Biochemistry. John Wiley.  
Metzler, D.E. 2006. Biochemistry. Vols. I, II. Wiley International.  
Nelson, D.L. and Cox, M.M. 2004. Lehninger principles of Biochemistry. 4th Ed. MacMillan.  
Voet, D., Voet, J.G. and Pratt, C.W. 2007. Fundamentals of Biochemistry. John Wiley.

## BI 512: Advanced Programming in Bioinformatics (2+2)

### Objective

To learn programming skills for parsing biological data, parallel computing, database connectivity and web-interface.

### UNIT I

Bioperl: Introduction, Modules: SeqIO, SearchIO, Seq Feature, Finding introns, Alignments, LiveSeq and Tree

### UNIT II

Overview of Parallel Computing, Concepts and Terminology, Parallel Computer Memory Architectures, Parallel Programming Models: parallelizing compilers, parallel languages, message-passing, virtual shared memory, object-oriented programming, and programming skeletons

### UNIT III

Methodical Design of Parallel Algorithms: partitioning, communication, agglomeration and mapping, Parallel Programming Paradigms: Task-Farming (or Master/Slave), Single Program Multiple Data (SPMD), Data Pipelining, Divide and Conquer, Speculative Parallelism.

#### UNIT IV

OpenMP: Clauses, Worksharing constructs, Synchronization constructs, Environment variables, Global Data, Runtime functions, Message Passing Interface (MPI): Introduction and programming, Point to point communications, Collective communications, Advanced MPI1 concepts, MPI2 introduction, Hybrid (openMP + MPI) programming

#### UNIT V

Compute Unified Device Architecture (CUDA): Introduction and Programming

#### **Suggested Readings**

James Tisdall. 2001. Beginning Perl for Bioinformatics. O-Reilly.  
Randal L. Schwartz, Tom Phoenix, brian d foy. 2008 .Learning Perl. O-Reilly.  
Robert Orfali and Dan Harkey. 1999 .Client/Server Programming with JAVA and CORBA. John Wiley.  
Sriram Srinivasan.1997. Advanced Perl Programming. O-Reilly.  
Tim Bunce and Alligator Descartes. 2000. Programming the Perl DBI. O-Reilly.

#### **BI 513/GP 500: Principles of Genetics (3+2)**

#### **Objective**

The aim of this course is to understand basic concepts of genetics and to develop analytical, quantitative and problem-solving skills in classical and molecular genetics.

#### UNIT I

History of Genetics; Mitosis & Meiosis, Pre-Mendelian concepts of inheritance, Mendel's laws; Discussion of Mendel's paper; Probability, Chromosomal theory of inheritance. Multiple alleles, Sex-linkage, Linkage Detection, Linkage estimation by various methods in test crosses, intercrosses; recombination and genetic mapping in eukaryotes -classical to modern, Somatic cell genetics.

#### UNIT II

Structural and numerical changes in chromosomes; Nature, structure and replication of the genetic material; Organization of DNA in chromosomes, Epigenetics. Genetic code; Protein biosynthesis, Genetic fine structure analysis, Allelic complementation, Split genes, Transposable genetic elements, Overlapping genes, Pseudogenes, Gene families and clusters.

#### UNIT III

Regulation of gene activity in prokaryotes; Molecular mechanisms of mutation, repair and suppression; Bacterial plasmids, insertion (IS) and transposable (Tn) elements; Gene expression & regulation in eukaryotes.

#### UNIT IV

DNA sequencing Gene cloning, genomic and cDNA libraries, PCR-based cloning, Nucleic acid hybridization and immuno-chemical detection; DNA restriction and modification, Anti-sense RNA, Gene silencing and ribozymes; Micro-RNAs (miRNAs). Genomics: Functional, structural & comparative, proteomics, metagenomics

#### UNIT V

Methods of studying polymorphism; Transgenic bacteria and bioethics; genetics of mitochondria and chloroplasts, Extra chromosomal inheritance. Eugenics, Genetic Disorders and Behavioural Genetics

#### UNIT VI

Population - Mendelian population – Random mating population- Frequencies of genes and genotypes-Causes of change: Hardy-Weinberg equilibrium.

#### **Practicals**

Laboratory exercises in probability and chi-square; Demonstration of genetic principles using laboratory organisms; Gene mapping using three point test cross; Tetrad analysis; Induction and detection of mutations, complementation. Study of chromosome aberrations, (deletions, inversion, translocations ); DNA extraction and PCR amplification - Electrophoresis – basic principles separation of DNA; Visit to transgenic glasshouse.

#### **Suggested Readings**

Gardner, E.J. and Snustad, D.P. 1991. Principles of Genetics. John Wiley & Sons.  
Klug, W.S. and Cummings, M.R. 2003. Concepts of Genetics. Peterson Education.  
Lewin, B. 2008. Genes IX. Jones & Bartlett Publ.  
Russell, P.J. 1998. Genetics. The Benjamin/Cummings Publ. Co.  
Strickberger, M.W. 2008. Genetics. Pearson Education.  
Tamarin, R.H. 1999. Principles of Genetics. Wm. C. Brown Pubs.  
Snustad, D.P. and Simmons, M.J. 2006. Genetics, 4th Ed. John Wiley & Sons

#### **BI 514: Statistical Techniques in Bioinformatics (3+1)**

*(Pre-requisite\*: PGS 504)*

#### **Objective**

To acquaint the students with advanced statistical methods applied in bioinformatics. The course would help students in applying advanced statistical methods in biological data.

#### UNIT I

Probability Theory - Probabilities of Events, Conditional Probabilities, Independence of Events, Entropy and Related Concepts, Transformations; Many Random Variables: Covariance and Correlation, Multinomial Distribution, Multivariate Normal Distribution, Indicator Random Variables; Principal component and correspondence analysis.

## UNIT II

Statistical Inference: Properties of estimation theory, Methods of estimation; Maximum likelihood estimation, Ordinary least square, Confidence Intervals, Classical Hypothesis Testing, P-Values, Testing for the Parameters in a Multinomial Distribution, Association tests, Likelihood Ratios, Information, Nonparametric Alternatives to the One-Sample and Two-Sample t-Tests, Differential Expression – Multiple Genes: The False Discovery Rate (FDR), Bootstrap Methods: Estimation and Confidence Intervals, Bayesian inference: Monte Carlo Markov Chain, The Hastings–Metropolis Algorithm, Gibbs Sampling; The Analysis of Variance, The 2n Design, Confounding, Repeated Measures.

## UNIT III

Basics of Stochastic Processes. Classification according to state space and time domain, Finite and countable state Markov chains, Poisson processes, Transition Probabilities, Stationary Distributions; The Analysis of One DNA Sequence: Modeling DNA, Modeling Signals in DNA, Weight Matrices: Independence, Markov Dependencies, Long Repeats, r-Scans, The Analysis of Patterns, Overlaps Not Counted, Motifs; Random Walks: The Simple Random Walk, The Difference Equation Approach, General Walks. Hidden Markov Models: Multiple Sequence Alignments, Gene Finding, Modeling Protein Families, Pfam.

### **Practical**

Principal Components, Correspondence Analysis, MLE, OLS, Bootstrap Methods, Nonparametric one sample and two sample test, Association test, Differential Expression Test, Gibbs Sampling, ANOVA, Modelling DNA, Modelling Signals in DNA, Multiple Sequence Alignments, Random Walk, Gene Finding, Modelling Protein Families

### **Suggested Readings**

Sunil Mathur, 2010. Statistical Bioinformatics with R, Elsevier

W. Warren John Ewens, Gregory Robert Grant 2001. Statistical Methods in Bioinformatics: An Introduction. Springer

Koski T, 2002. Hidden Markov Models for Bioinformatics, Springer

*\* exempted for the students with major/minor in Agricultural Statistics*

## **BI 601: Genome Assembly and Annotation (1+2)**

### **Objective**

The primary objective of this course is to develop practical understanding of techniques and tools used in genome assembly with emphasis on issues and challenges of its structural and functional annotation.

## UNIT I

Types and methods of genome sequence data generation; Shot gun sequencing method; Problems of genome assembly, Approaches of genome assembly: Comparative Assembly, DE novo Assembly; Read coverages; Sequencing errors, Sequence Quality Matrix, Assembly Evaluation; Challenges in Genome Assembly

## UNIT II

Various tools and related methods of genome assembly: MIRA, Velvet, ABySS, ALLPATHS-LG, Bambus2, Celera Assembler, SGA, SOAPdenovo etc.

## UNIT III

Basic concepts of genome annotation; Structural and Functional Annotation; Identification of open reading frame (ORF) and their regularization, Identification of gene structure, coding regions and location of regulatory motifs

### **Practicals**

Genome assembly methods for data from various sequencing platform, Sequencing error determination, Sequence quality matrix; Various tools for genome assembly: MIRA, Velvet, ABySS, ALLPATHS-LG, Bambus2, Celera Assembler, SGA, SOAPdenovo etc. Structural and functional Genome annotation.

### **Suggested Readings**

Prof. Dmitrij Frishman, Alfonso Valencia, Modern Genome Annotation springer

Jung Soh, Paul M.K. Gordon, Christoph W. Sensen. 2012. Genome Annotation. Chapman and Hall/CRC

J. Craig Venter, 2000. Annotation of the Celera Human Genome Assembly. Celera.

Mark Menor (2007). Multi-genome Annotation of Genome Fragments Using Hidden Markov Model Profiles

Carson Hinton Holt. 2012. Tools and Techniques for Genome Annotation and Analysis

## **BI 602: Bio-molecular Modelling and Simulation (2+1)**

### **Objective**

The course is aimed to develop understanding of bio molecular modelling techniques and simulation.

#### UNIT I

Basic principles of modeling, modeling by energy minimization technique, concept of rotation about bonds, energy minimization by basic technique for small molecules, Ramachandran plot, torsional space minimization, energy minimization in cartesian space, molecular mechanics-basic principle

#### UNIT II

Basic concepts of Simulation Modelling: Units and derivatives, Force field and energy landscape, Truncation of non-bonded interactions, Introduction to solvation, Periodic boundary condition, Wald summation, implicit solvent model and continuum electrostatics, Monte Carlo simulation on parallel computers.

### UNIT III

Replica-exchange simulations, Restraint potentials, Free energy calculations, Membrane simulations

#### Practicals

Molecular modeling and energy minimization techniques; Ramachandran plot; Simulation dynamics, Monte carlo simulation on parallel computers. Replica exchange simulation, free energy calculation.

#### Suggested Readings

Tamar Schlick. 2010. Molecular Modeling and Simulation: An Interdisciplinary Guide. Science.

W.F. van Gunsteren, P.K. Weiner, A.J. Wilkinson. 1997. Computer Simulation of Biomolecular Systems: Theoretical and experimental application. Springer.

Martin J. Field. A Practical Introduction to the Simulation of Molecular Systems. Cambridge University Press.

### BI 603: Machine Learning Techniques in Bioinformatics (2+1)

#### Objective

The purpose of the course is to explain various machine learning techniques and its applications on biological data.

#### UNIT I

Introduction to statistical learning theory, Empirical Risk Minimization, Structural Risk Minimization; Classification: Decision tree, Bayesian, Rule based classification, ANN, SVM, KNN; Case based reasoning and Applications in Bioinformatics.

#### UNIT II

Clustering: Partition Methods, Heirarchical methods, Density based methods, Grid based clustering, Model based clustering, clustering of high dimensional data, constraints based clustering, Analysis of MD trajectories, Protein Array data Analysis.

#### UNIT III

Dimensional Reduction Techniques, Methods of Feature Selection, Resampling Techniques, Elements of Text Mining and Web Mining, Soft Computing and Fuzzy logic system & application in bioinformatics;

#### Practicals

Decision tree, classification techniques: ANN, SVM, KNN, Case based reasoning and its applications on biological data. Clustering techniques; Clustering of high dimensional data; Dimensional reduction techniques; Resampling techniques; Text mining and Web mining. Soft Computing and Fuzzy logic system & application in bioinformatics.



### **Suggested Readings:**

Witten, H. I., Frank, E. and Hall, M. A. 2011. Data Mining: Practical Machine Learning Tools and Techniques.

Hastie, T., Tibshirani, R., Friedman, J. H. 2009. The Elements of Statistical Learning: Data Mining Interface and Prediction.

Clarke, S. B., Fokoue, E. and Zhang, H. H. 2009 Principles and Theory for Data Mining and Machine Learning.

### **BI 604: Computational Techniques of Transcriptomics and Metabolomics (1+1)**

#### **Objective**

The main objective of this course is to introduce computational techniques and various tools used in transcriptomics and metabolomics.

#### UNIT I

Microarrays, RNA-seq, Chip-Seq, EST-clustering, differential expression analysis

#### UNIT II

Tools used for analysis of metabolites: XCMS, MZmine, MetAlign, MathDAMP, LCMStats

#### **Practicals**

Microarray data analysis; RNA-seq, chip-seq, EST-clustering. Tools for analysis for metabolites: XCMS, MZmine, MetAlign, MathDAMP, LCMStats

### **Suggested Readings:**

Daub, C. O. 2004. Analysis of Integrated Transcriptomics and Metabolomics.

Lindon, J. C. Nicholson, J.K. and Elaine Holmes. 2011. The Handbook of Metabonomics and Metabolomics.

Weckwerth, W. 2007. Metabolomics: Methods and Protocols.

### **BI 611: Metagenomics Data Analysis (2+1)**

#### **Objective**

The course aims is to teach basic concepts of metagenomics and various techniques used in the analysis of metagenomic data

#### UNIT I

Taxonomic and genetic annotation of high throughput sequence data, microbial diversity analyses, analyses of microbial community composition and change and metabolic reconstruction analyses

## UNIT II

Comparison between Metagenomics and AL, EC, Comparison between LCS and Metagenomics, Symbiotic Evaluations: SANE, Comparison between SANE and Metagenomics, Horizontal Gene Transfer: Microbial GA

## UNIT III

Metagenome Sequencing, Single Cell Analysis, Host-Pathogen Interaction; Shotgun metagenomics; High-throughput sequencing; Comparative metagenomics; Community metabolism; Metatranscriptomics

### Practicals

Meta genome annotation, Analyses of microbial community composition and change and metabolic reconstruction analyses; Metatranscriptomics; Comparative metagenomics.

### Suggested Readings

Diana marco, 2010. Metagenomics: Theory, Methods and Applications. Ceister academic press  
Wolfgang R. Streit, Rolf Daniel. 2010. Metagenomics: Methods and Protocols. *Springer protocols*.

Wu-Kuang Yeh, Hsiu-Chiung Yang, James R. McCarthy, 2010. Enzyme Technologies: Metagenomics, Evolution, Biocatalysis and Biosynthesis. wiley

The New Science of Metagenomics: Revealing the Secrets of Our Microbial Planet National Academies Press.

Vijay Muthukumar. 2003. Metagenomics for the Identification of Plant Viruses. ProQuest.

## BI 621: Quantum Theory and Applications in Biology (2+1)

### Objective

This course introduces the concepts of quantum theory with application in molecular biology

#### UNIT I

Classical mechanics, Newton, Lagrange and Hamilton's equations, Schrodinger's equation and its complete solution for S.H.O, central force and angular momentum.

#### UNIT II

Atomic orbital models, the wave equation, molecular orbitals, the LCAO method, the overlap method, coulomb and resonance integrals, the hydrogen molecule, charge distributions, approximate methods.

#### UNIT III

Absorbance of frequency-specific radiation (photosynthesis), Conversion of chemical energy into motion, Magneto reception in animals, DNA mutation and Brownian motors in many cellular processes

## Practicals

Classical mechanics, Central force and angular momentum; Atomic orbital model, Wave equation, Resonance integers. DNA mutation and Brownian motors in many cellular processes.

## Suggested Readings:

Heisenberg, W. 1949. The Physical Principles of the Quantum Theory.  
Bohm, D. 1951. Quantum Theory.  
Ghatak A. K. and Lokanathan, S. 2004. Quantum Mechanics: Theory and Applications.  
Bittner, E. R. 2009. Quantum dynamics: applications in biological and materials systems.  
Blinder, S. M. 2004. Introduction to Quantum Mechanics: In Chemistry, Materials Science, and Biology.

## BI 622: Molecular Dynamics (2+1)

### Objective

Basic objective of this course is to teach the theory and algorithms of molecular dynamics and its simulation.

#### UNIT I

Introduction to Molecular dynamics, Newton's equation of motion, equilibrium point, water models, thermodynamic ensembles, equilibration, radial distribution function, pair correlation functions.

#### UNIT II

MD methodology, periodic box (PBC), algorithm for time dependence. Leapfrog algorithm, Verlet algorithm, Boltzman velocity, time steps.

#### UNIT III

Basic steps in molecular dynamics simulation. Starting structures, duration of the MD run, final MD simulation structure. Visualization and analysis of trajectories

## Practicals

Molecular dynamics; MD methodology; periodic box; Starting structures, duration of the MD run, final MD simulation structure visualization and analysis of trajectories.

## Suggested Readings

J. M. Haile. 1997. Molecular Dynamics Simulation: Elementary Methods. Wiley Professional.  
D. C. Rapaport. 2004. The Art of Molecular Dynamics Simulation. Cambridge University Press,  
Perla Balbuena, Jorge M. Seminario. 1999. Molecular Dynamics: From Classical to Quantum Methods.  
Julia M. Goodfellow. 1990. Molecular Dynamics: Applications in Molecular Biology. Science.

## **BI 623: Peptide Design, Synthesis and Applications (2+1)**

### **Objective**

To teach various approaches meant for peptide design, its synthesis, docking with target and applications in agricultural sciences.

#### UNIT I

Introduction to peptides, peptide design, synthesis of peptides (solution phase and solid phase), protection and deprotection of amino and carboxyl group

#### UNIT II

Unnatural amino acids, conformation of peptides, purification and crystallization of peptides, determination of structure of small molecules

#### UNIT III

Pharmacodynamics and pharmacokinetics, Drug potency and Efficacy, Docking, Active site, Absorption, Distribution, Development of a drug: Classical steps, Chemical Parameters in drug design, Structure based drug discovery, Quantitative Structure, Activity Relationships

### **Practicals**

Peptide design; protection and deprotection of amino and carboxyl group; conformation of peptides, purification and crystallization of peptides, determination of structure of small molecules; Pharmacodynamics and pharmacokinetics, Drug potency and Efficacy, Docking, Active site, Absorption, Distribution; Structure based drug discovery, Quantitative Structure, Activity Relationships

### **Suggested Readings**

John Howl.2005. Peptide Synthesis and Applications. Springer

Norbert Sewald, Hans-Dieter Jakubke. 2009. Peptides: Chemistry and Biology. Wiley VCH

Knud J. Jensen. 2009. Peptide and Protein Design for Biopharmaceutical Applications. John Wiley & Sons

## **BI 631: Parallel Programming and Algorithm Development (2+1)**

### **Objective**

To learn the concepts of parallel computing, parallel programming for handling biological data and development of algorithms.

#### UNIT I

Parallel Programming- Introduction, Design Pattern, Pattern Languages, concurrency in parallel programs vs. operating systems; Parallel computer architecture, Flynn's Taxonomy of parallel architectures.

## UNIT II

Memory organization of parallel computers, Thread -level parallelism, Interconnection Networks, Parallel Programming Models, Parallel-Matrix vector product, Approaches for new parallel languages, Performance analysis of parallel programs.

## UNIT III

Algorithm Development- Choosing an Algorithm Structure Pattern, target platform, major organizing principle, the algorithm structure decision tree, re-evaluation

### Practicals

Parallel Programming with MPI, Parallel Programming with OpenMP, Laboratory works for estimating the parallel method efficiency, Laboratory works for developing the parallel algorithms and programs, Laboratory works for parallel solving partial differential equations, Laboratory works for studying the parallel method libraries, Laboratory works for parallel solving the problem of multidimensional multiextremal optimization

### Suggested Readings

Gebali, F. 2011. *Algorithms and Parallel Computing*. Wiley Series on Parallel and Distributed Computing. John Wiley & Sons.

Lin, Y. C. and Snyder, L. 2008. *Principles of parallel programming*. Pearson/Addison Wesley  
Zomaya, A.Y. *Parallel computing: paradigms and applications*. International Thomson Computer Press

## BI 632: Optimization Techniques in Bioinformatics (2+1)

### Objective

This course is meant for exposing the students to numerical methods of optimization, linear and nonlinear programming techniques with exposure to practical applications of these techniques for bioinformatics.

#### UNIT I

Optimization in bioinformatics, Local optimisation techniques: Simplex methods, steepest descent method, conjugate gradient and others.

#### UNIT II

Sample Subset Optimization, NP hard optimization problem, integer linear programming, integer quadratic programming, Stochastic optimization techniques: Monte Carlo, simulated annealing, particle swarm optimization.

#### UNIT III

Global optimization techniques: genetic algorithm, dynamic programming, Expectation maximization, Ant colony optimization, Applying Multi-Objective Optimisation: Mixed and hybrid techniques, Application to Genetic networks

## **Practicals**

Local optimisation techniques: Simplex methods, Steepest descent method, Conjugate gradient; Sample Subset Optimization, NP hard optimization problem, Integer linear programming, Integer quadratic programming, Stochastic optimization techniques: Monte Carlo, Simulated annealing, Particle swarm optimization. Global optimization techniques: genetic algorithm, Dynamic programming, EM, Ant colony optimization; Mixed and hybrid techniques; Genetic networks

## **Suggested Readings:**

Polanski, A. 2007. Bioinformatics.  
Xie, Wei. 2007. Optimization Algorithms for Protein Bioinformatics.  
Leondes, T. C. 1998. Optimization Techniques.  
Taha, H. A. 2011. Operation Research An Introduction, Prentice Hall

## **BI 633: Biological Data Integration and Quality Control (2+1)**

### **Objective**

To familiarize the techniques of data sources, data curation and integration of data sources

#### UNIT I

Curation of genomics, genetic, proteomics, High-throughput screening, array, qPCR data sets; Quality management of data: tools and techniques.

#### UNIT II

Biological data sources, Data granularity, Schema modelling, architecture, query design, extraction, transformation and loading, Long term data management, storage and security.

#### UNIT III

Bio-chip information system, visualization and reporting, Risk factors for data quality management.

## **Practicals**

Understanding the data sources, Data granularity, Data modeling and architecture, development of database, Storage, Security, Visualization and reporting.

## **Suggested Readings**

Kozak, K. Large scale data handling in biology. 2010. Ventus Publishing ApS. ISBN 978-87-7681-555-4.  
Harold, E. and Means W.S. XML in a Nutshell, Third Ed. O'Reilly, Sebastopol, CA, 2004.  
Witten, I.H. and Frank E. Data Mining: Practical Machine Learning Tools and Techniques (WEKA), 2nd Ed. San Francisco, Morgan Kaufmann, 2005.  
Lodish, H. et al. Molecular cell biology. New York: Freeman & Co. 2000.  
Kaneko, K. Life: An Introduction to Complex Systems Biology. Springer. 2006.

## **BI 634: Graphics and Visualization of Biological Data (2+1)**

### **Objective**

To familiarize the students with the graphical data formats and visualization of data and results

#### UNIT I

Concept of data visualization, Raster image data processing and analysis, Basic raster graphics algorithm for drawing 2D primitives, 3D object representation, Geometrical transformation.

#### UNIT II

Integrative data analysis and visualization; Concepts of data and predictive model integration.

#### UNIT III

Visualization of protein interaction, Network Analysis, Comparing protein structure, visualization of structure motifs

### **Practicals**

Understanding the image formats, Data Visualization and Browser, Network Analysis and Structure comparison.

### **Suggested Readings**

Murray, S. Interactive Data Visualization for the Web. O'Reilly Media. 2013.

McKinney, W. Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython. O'Reilly Media. 2012.

Janert, P.K. Data Analysis with Open Source Tools. O'Reilly Media; 1 edition. 2010.

Azuaje, F. and Dopazo, J. Data Analysis and Visualization in Genomics and Proteomics. Wiley; 1 edition. 2005.

Hartvigsen, G. A Primer in Biological Data Analysis and Visualization Using R. Columbia University Press. 2014.

## **BI 635: Biological Databases and Data Analysis (2+1)**

### **Objective**

To know about different biological databases existing in the public domain and perform analysis on the data available in them.

#### UNIT I

Nature of biological data; Overview of available Bioinformatics resources on the web; NCBI/EBI/EXPASY etc; Biological Databases: Nucleic acid sequence databases; GenBank/EMBL/DBJ; Biological Databases: Protein sequence databases; PIR-PSD; SwissProt, UniProtKB; Database search engines: Entrez, SRS

## UNIT II

Overview/concepts in sequence analysis; Pairwise sequence alignment algorithms: Needleman & Wunsch, Smith & Waterman; Scoring matrices for Nucleic acids and proteins: MDM, BLOSUM, CSW; Database Similarity Searches: BLAST, FASTA; Multiple sequence alignment: PRAS, CLUSTALW; Biological databases: Genome & genetic disorders.

## UNIT III

Genome databases: Human, model organisms, microbes & viral: OMIM; Biological databases: structural databases: PDB, NDB, CCSD; Derived databases: Prosite, BLOCKS, Pfam/Prodom

### Practicals

Nucleic acid sequence databases, Protein sequence databases, Database search engines, Database Similarity Searches, Multiple sequence alignment, Genome databases, Structural databases, Derived databases

### Suggested Readings

Baxevanis, A. D. & Ouellette, B., F. F. 2002. *Bioinformatics: A Practical Guide to the analysis of Genes and Proteins* (2nd Ed.). New York, John Wiley & Sons, Inc. Publications.  
Attwood, T. K. & Parry-Smith, D. J. 2001. *Introduction to Bioinformatics*. Delhi Pearson Education (Singapore) Pvt. Ltd.  
Mount, David. 2004. *Bioinformatics: Sequence and Genome Analysis*. Cold Spring Harbor Laboratory Press, New York.  
Baxevanis, A.D., Davison, D.B., Page, R. D. M. & Petsko, G.A. 2004. *Current Protocols in Bioinformatics*. John Wiley & Sons Inc. New York.

### BI 641: Biological Network Modelling and Analysis (2+1)

#### Objective

This course aims to develop basic understanding of system biology through biological network modelling and its analysis.

#### UNIT I

Introduction to biological networks, Graph theoretic modelling and analysis of biological networks, Discrete Dynamic modelling (Boolean networks, Petri nets), Continuous dynamic modelling (ODEs, stochastic simulation, etc.)

#### UNIT II

Probabilistic modelling (Probabilistic Boolean networks, Bayesian networks, Mutual Information), Network inference from experimental data, Genome-scale modelling and network integration.

#### UNIT III

Evolution of molecular networks, Network-guided GWAS studies, FBA and epistasis detection, protein function prediction



## **Practicals**

Biological networks, Graph theoretic modelling and analysis of biological networks, Discrete Dynamic modeling; Continuous dynamic modeling; Probabilistic modeling; Genome-scale modelling and network integration; Evolution of molecular networks, Network-guided GWAS studies, FBA and epistasis detection, protein function prediction.

## **Suggested Readings:**

Junker, B. H. 2008. Analysis of Biological Networks.

Koch, I. Reisig, W. Schreiber F. 2010. Modeling in Systems Biology: The Petri Net Approach.

Ramadan, E.Y. 2008. Biological Networks: Modeling and Structural Analysis.

Laubenbacher, R. 2007. Modeling and Simulation of Biological Networks.

## **BI 642: Genome Wide Association Study (2+1)**

### **Objective**

To introduce the concepts, principles, various designs and techniques of genome wide association study.

#### **UNIT I**

Definition, Allelic spectra of common diseases, Allele frequencies for susceptibility loci, Risks associated with disease-susceptibility variants, Applications of linkage-disequilibrium metrics, SNP map, Genome resequencing for full coverage in genome-wide association studies, Transmission Disequilibrium Test, common variant hypothesis, rare allele hypothesis, Genome-wide graph theory algorithms.

#### **UNIT II**

Case-Control design, Trio design, Cohort design, Cross-sectional designs for GWAS Selection of Study Participants, Environmental confounders in GWAS, Confounding by population stratification, Genotyping and Quality Control in GWA Studies, Analysis of association between SNP and traits.

#### **UNIT III**

Uses of GWAS: gene-gene interaction, detection of candidate haplotypes, association between SNPs and gene expression.

## **Practicals**

Allelic spectra of common diseases, Allele frequencies for susceptibility loci, linkage-disequilibrium metrics, SNP map, Genome resequencing for full coverage in GWAS; Case-Control design, Trio design, Cohort design, Cross-sectional designs for GWAS Selection; Genotyping and Quality Control in GWA Studies; Analysis of association between SNP and traits.

## **Suggested Readings:**

Qin, H. 2008. Statistical Approaches for Genome-wide Association Study and Microarray Analysis.

Yang, C. 2011. SNP Data Analysis in Genome-wide Association Studies.

Kraft, J. S. 2010. Genome-wide Association Study of Persistent Developmental Stuttering.

### **BI 651: Recent Advances in Bioinformatics (1+0)**

#### **Objective**

To develop proficiency of the student in his/her area of specialization in bioinformatics.

#### **UNIT I**

Recent advances in various concepts, techniques, tools, algorithms and their applications in the area of bioinformatics.

#### **Suggested Readings**

Selected topics from recent articles, reviews, books and journals.

### **BI 691: Seminar (1+0)**