M.Tech. Computational Biology

REGULATIONS AND SYLLABI

(Effective from 2014-2015)

Centre for Bioinformatics

SCHOOL OF LIFE SCIENCES

PONDICHERRY UNIVERSITY

PUDUCHERRY

Eligibility for M. Tech. Computational Biology **

1) Minimum of 55% of marks in Master's degree in any relevant area of Bioinformatics/ Physics/Chemistry/ Mathematics/ Statistics/ Computer Science/ Biotechnology/ Biochemistry/ Microbiology/ Plant Biology/ Botany/ Animal Biology/ Zoology for Pondicherry University

2) Minimum of 55% of marks in B.Tech/B.E degree in Industrial Biotechnology, Biotechnology, Pharmaceutical Technology, Food Technology, Bioinformatics, Chemical Engineering, Leather, Bioengineering, Information Technology, Information Science, Computer Science and Engineering, Electrical and Electronics Engineering, Electronics and Communication Engineering, Mechanical Engineering, Biomedical Engineering, Electronics and Instrumentation Engineering for both Anna University and Pondicherry University

** Network teaching programme in collaboration with Anna University, Chennai. Admission to Anna University, Chennai will also be based on the entrance exam conducted by Pondicherry University

PONDICHERRY UNIVERSITY

SCHOOL OF LIFE SCIENCES

Centre for Bioinformatics

SYLLABUS FOR M. Tech. Computational Biology

(Academic Year 2014-2015 onwards)

- Students with Biology background are expected to choose “Physical Sciences for Biologist” and students with Mathematical and Physical Science background are expected to choose “Introductory Biology” as compulsory papers.

CBIO 601 BIOINFORMATICS AND SEQUENCE ANALYSES

Total Credits: 3 **Total: 35 Hrs.**

Unit I **7 Lectures**

Introduction to primary Databases: Types of Biological data- Genomic DNA, cDNA, rDNA, ESTs, GSSs; Primary Databases -Nucleotide sequence databases-GenBank, EMBL, DDBJ, Protein Sequence Databases- UniProtKB, UniProt, TrEMBL, Swiss-Prot, UniProt Archive-UniParc, UniProt Reference Clusters-UniRef, UniProt Metagenomic and Environmental Sequences-UniMES. Literature Databases-PubMed, PLos, BioMed Central.

Unit II **7 Lectures**

Introduction to Secondary or Derived Databases- PDB, CSD, MMDB, SCOP, CATH, FSSP, CSA, KEGG ENZYME, BRENDA; Sequence motifs Databases:-Prosite, ProDom, Pfam, InterPro; Composite Databases-NRDB, Genome Databases- Viral genome database (ICTV db), Bacterial Genome database (GOLD, MBGD), Organism specific database (OMIM/OMIA, SGD, WormBase, PlasmoDB, FlyBase, TAIR), Genome Browsers (Ensembl, VEGA, NCBI map viewer, UCSC Genome Browse). Bioinformatics Database search engines:-Text-based search engines (Entrez, DBGET/LinkDB).

Unit III **7 Lectures**

File formats, sequence patterns and profiles: Sequence file formats – GenBank, FASTA, ALN/ClustalW2, PIR; Basic concept and definition of sequence patterns, motifs and profiles, various types of pattern representations viz. consensus, regular expression (Prosite-type) and sequence profiles; Sequence similarity based search engines (BLAST and FASTA); Pattern based search using MeMe and PRATT); Motif-based search using ScanProsite and eMOTIF; Structure similarity based search using VAST and DALI; Profile-based database searches using PSI-BLAST and HMMer.

Unit IV **7 Lectures**

Sequence Analysis and predictions: Nucleic acid sequence analysis- Reading frames; Codon Usage analysis; Translational and transcriptional signals, Splice site identification, Gene prediction methods and RNA fold analysis; Protein sequence analysis-Compositional analysis, Hydrophobicity profiles, Amphiphilicity detection, Moment analysis, Transmembrane prediction methods, Secondary structure prediction methods.

Unit V **7 Lectures**

BIG DATA in OMICS: Big data industry standards, Data acquisition, cleaning, distribution, and best practices, Visualization and design principles of big data infrastructures, Biological databases for big data management, High Performance Computing, grid, and cloud computing for omics sciences, Real-Time Processing of Proteomics Data Using Hadoop.

Text books

1. Bioinformatics: Sequence and Genome Analysis by Mount D., Cold Spring Harbor Laboratory Press, New York. 2004
2. Bioinformatics- a Practical Guide to the Analysis of Genes and Proteins by Baxevanis, A.D. and Francis Ouellellette, B.F., Wiley India Pvt Ltd. 2009
3. Introduction to Bioinformatics by Teresa K. Attwood, David J. Parry-Smith. Pearson Education. 1999

References

1. Near real-time processing of proteomics data using HADOOP by Hillman et al., (2014) Mary ann Liebert, Inc- Big Data. 2 (1): BD44- BD49.
2. Curating Big Data Made Simple: Perspectives from Scientific Communities by SoweSulayman K. and ZettsuKoji **(2014).** Big Data. 2 (1): 23-33
3. The quantified self: Fundamental Disruption in Big Data Science and Biological Discovery by Melanie Swan (2013) Mary ann Liebert, Inc. Big data , 1(2): BD85-99.

CBIO 602 FUNDAMENTALS OF BIOSTATISTICS

Total Credits: 3 **Total: 34 Hrs.**

Unit I **6 Lectures**

Review of Basic statistical measures: Numerical description of data, Measures of central tendency, Measuring variations in data, Standard deviation and its significance, Percentile, Quartiles, Box Plots, Correlation and Regression, Application to Biology.

Unit II **6 Lectures**

Probability theory: Classical and modern definition of probability, Sample space and events, Axioms of probability, Sample space having equally likely outcomes, Conditional probability, Independent events, Bayes formula and its application to Biology, Random Variables- Types of Variables, Eexpected Value, Variance

Unit III **7 Lectures**

Discrete and Continuous Distribution: Bernoulli and Binomial, Poisson, Geometric, Uniform, Exponential, Normal, Standard normal, Student's t, Chi-squared, Fisher-Snedector distribution, their importance and usage in Biology.

Unit IV **7 Lectures**

Sampling Distributions and Estimation: Statistic, Distribution of sample mean, sample variance, central limit theory, Biased and unbiased estimator, Confidence interval, Population mean, Population variance.

Unit V **8 Lectures**

Tests of Hypotheses: Formulation of Hypothesis Simple and Composite, Type I and Type II errors, Power of a test, Significance of a test, P-value, Testing Normal, Chi-square, t-test and F-test, Non-parametric - Mann-Whitney test, Applications to Biology.

Text Books

1. Biostatistics (9th Ed.), Wayne W. Daniel, John Wiley, 2004
2. Probability and Statistics (5^{th} Ed.), J.L. Devore, Thomson Asia, 2002
3. Statistics (3^{rd} Ed.), Murray R. Spiegel and Larry J. Stephens, Tata McGraw-Hill, 2000

Reference Books

1. Statistical Methods (Volume 1 and 2) (1^{st} Ed.), N. G. Das, Tata McGraw-Hill, 2009
2. Fundamentals of Biostatistics (6th Ed.), Bernard Rosner, Thomson Brooks/Cole, 2006

CBIO 603 DESIGN AND ANALYSIS OF ALGORITHM

Total Credits: 3 **Total: 36 Hrs.**

Unit 1 **5 lectures**

Algorithm Analysis: Analyzing algorithms-Designing algorithms-Asymptotic notation-Standard notations and common functions-The substitution method-The recursion tree method-The master method -Basics of time complexity estimates, General norms for running time calculation

Unit 2 **6 lectures**

Divide and conquer: General method, applications-Binary search, Quick sort, Merge sort, Strassen's Matrix Multiplication. Greedy method: General method, applications-Job sequencing with deadlines, 0/1 knapsack problem, Minimum cost spanning trees, Single source shortest path problem.

Unit 3 **7 lectures**

Dynamic Programming: General method, applications-Matrix chain multiplication, Optimal binary search trees, 0/1 knapsack problem, All pairs shortest path problem, Traveling sales person problem, Reliability design.

Unit 4 **6 lectures**

Searching and Traversal Techniques: Efficient non-recursive Tree Traversal Algorithms, DFS, BFS of Graphs, AND/OR graphs, game trees, Bi-Connected components, Search Trees- Balanced search trees-AVL trees, representation, Operations-insertion, deletion and searching,
B-Trees-B-Tree of order m, Operations- insertion, deletion and searching.

Unit 5 **7 lectures**

Backtracking and Branch and Bound: General method (Backtracking), Applications-n-queen problem, sum of subsets problem, graph coloring, Hamiltonian cycles. General method (Branch and Bound), Applications - Traveling sales person problem.

Text Books:

1. Computer Algorithms/C++, E.Horowitz, S.Sahani and S.Rajasekharan, Galgotia Publishers pvt. Limited.
2. Data Structures and Algorithm Analysis in C++, 2nd Edition, Mark Allen Weiss, Pearson Education.
3. Introduction to Algorithms, 2nd Edition, T.H.Cormen, C.E.Leiserson, R.L.Rivest, and C.Stein, PHI Pvt.Ltd./ Pearson Education.

Reference Books:

1. Design and Analysis of algorithms, Aho, Ullman and Hopcroft, Pearson Education.
2. Introduction to the Design and Analysis of Algorithms, A.Levitin, Pearson Education.
3. Data structures, Algorithms and Applications in C++, S.Sahni, University press (India) pvt ltd, 2nd edition, Orient Longman pvt.ltd.
4. Object Oriented Programming Using C++, 2nd Edition, I.Pohl, Pearson Education.
5. Fundamentals of Sequential and Parallel Algorithms, K.A.Berman, J. L.Paul, Thomson
6. Data Structures And Algorithms in C++, 3rd Edition, Adam Drozdek, Thomson.
7. Algorithm Design: Foundations, Analysis and Internet examples, M.T.Goodrich and R.Tomassia, John Wiley and sons.

CBIO 604 - C++ IN COMPUTATIONAL BIOLOGY

Total Credits: 3 **Total: 36 Hrs.**

Unit 1 **7 lectures**

C++ programming basics: Compilation of C++ programs – Input and output statements – integer, float, and character variables – arithmetic operations and built-in library functions.

Unit 2 **9 lectures**

Procedural concept – decision making, functions and structures: Loops and decision making statements – structures and unions – arrays and strings – user defined functions.

Unit 3 **7 lectures**

Pointers and file handling: Pointer concept – pointers and arrays – pointers and functions – pointers to pointers – File handling – Reading and Writing the data from file.

Unit 4 **6 lectures**

Object Oriented Programming: Object oriented concepts – working with objects and classes in C++ – scope resolution operator – constructors – destructors – overloading of constructors and operators – string class.

Unit 5 **7 lectures**

Inheritance : Concept of inheritance – base class and derived class – overriding of member functions – abstract class – public and private inheritance – Levels of inheritance and multiple inheritance – inheritance and graphic shapes – virtual function and friend function.

Text Books:

1. Object Oriented Programming using C++ (4^{th} Ed.) by Lafore, R. Sams Publishers. 2002

CBIO 605 - CELL AND MOLECULAR BIOLOGY

Total Credits: 3 **Total: 36 Hrs.**

Unit 1 — 6 Lectures

Molecules of life - Structural organization of prokaryotic and eukaryotic cells- Concept of a composite cell and Molecular composition of cells. Biomembranes- Structural organization- Models of a plasma membrane, Membrane permeability- Transport across cell membranes- Transmembrane signals- Artificial membranes- liposome.

Unit 2 — 7 Lectures

Mitochondrial Structure and Function – Oxidative Metabolism in the Mitochondrion – The Role of Mitochondria in the formation of ATP – Translocation of Protons and the Establishment of a proton-motive force – The Machinery for ATP formation – Peroxisomes.

Unit 3 — 7 Lectures

Chloroplast structure and function – An overview of photosynthetic Metabolism – The absorption of light – Photosynthetic units and reaction centers – Photophosphorylation – Carbondioxide fixation and the synthesis of carbohydrates

Unit 4 — 7 Lectures

Cellular Components – Cytoskeleton – components of Cytoskeleton, Microtubules, Intermediate filaments – Microfilaments, Cell cycle, Endoplasmic reticulum, Golgi complex, Types of vesicles - transport and their functions, Lysosomes.

Unit 5 — 9 Lectures

DNA and Protein Synthesis - Structure of DNA - evidence for DNA as genetic material. Gene transfer in microorganisms – conjugation, transformation, transduction - DNA replication, Transcription – mRNA processing, Translation. Protein synthesis – Ribosomes, enzymes, Protein processing.

Text Book:

1. Cell and Molecular Biology – Concepts and Experiments by Gerald Karp. Wiley International Student Version. 2008

Reference Books:

1. Genes VIII (8 th Ed.) by Lewin, B. Pearson Education International. 2004.
2. Cell and Molecular Biology by De Robertis and De Robertis. Saunders College, Philadelphia, USA. 2002

CBIO 606- METABOLISM AND IMMUNOLOGY

Total Credits: 3 **Total: 36 Hrs.**

Unit 1 **6 lectures**

Basic enzymology: Enzyme nomenclature and classification of enzymes according to I.U.B. convention, General properties of enzymes, substrate specificity and active site. Enzyme kinetics, Michales-Menten equation, Lineweavar Burk, Factors effect Enzyme activity, Enzyme Inhibition, allosteric enzymes.

Unit 2 **7 lectures**

Overview of metabolism, high energy compounds, oxidation-reduction reactions, the reactions of glycolysis, fermentation, control of glycolysis. The pentose phosphate pathway, glycogen breakdown and synthesis, control of glycogen metabolism, gluconeogenesis. Citric acid cycle: enzymes of the citric acid cycle, regulation of the citric acid cycle,

Unit 3 **8 lectures**

Protein metabolism: amino acid deamination, the urea cycle, breakdown of amino acids, amino acid biosynthesis. Fatty acid metabolism Lipid digestion, adsorption and transport, fatty acid oxidation, ketone bodies, fatty acid biosynthesis, regulation of fatty acid metabolism. Nucleic acid metabolism: Synthesis of purine ribonucleotides, synthesis of pyrimidiine ribonucleotides, formation of deoxyribonucleotides. Heme biosynthesis and degradation

Unit 4 **8 Lectures**

Introduction and Antibodies: Innate and acquired immunity, active and passive immunity, natural and artificial immunity and humoral. Lymphoid system- primary or secondary organ. Cells- Lymphocytes, mononuclear, phagocytes, antigen presenting, polymorphs, mast cells, cluster designation (CD) and antigen specific receptors – Principles and its uses. **Antibody generation:** structure and function –clonal selection theory-different types of immunoglobulins, effectors, receptors and antibody diversity. Complement system- activation, pathways and biological effects. Major Histochemical molecules/peptide complexes- Structure and Function.

Unit 5 **7 Lectures**

Antigen and antibody reaction/interaction: Precipitation, Haemagglutination, direct and indirect immunofluroescence, hybridoma technology for mass production. Chimeric antibodies, antibody engineering; large scale manufacture of antibodies. **Vaccine development and Immunoinformatics**: Recombinant vaccines, combined vaccines, polyvalent vaccines. Immunoinformatics, databases in immunology, DNA, Plant and protein based recombinant antigens as vaccines.

Text book:

1. Biochemistry by Voet and Voet. Wiley. 2011
2. Text book of Immunology by Kuby, 2008

Reference Books:

1. Principles of Biochemistry by Nelson and Cox, Lehninger. W H Freeman & Co. 2009
2. Biochemistry by Berg, Tymoczko & Stryer. W.H.Freeman and Co New York. 2007
3. Text book of Immunology by Riott, 2006

CBIO 607- PHYSICAL SCIENCES FOR BIOLOGIST

Total Credits: 3 **Total: 36 Hrs**

Unit 1: **5 lectures**

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

Unit 2: Quantum Mechanics **4 lectures**

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

Unit 3: Thermodynamics **6 lectures**

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

Unit4 **6 Lectures**

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

Unit5 **8 Lectures**

Introduction to Organic chemistry: Carbon and its compounds, Position of Carbon in periodic table, tetra covalency of carbon, functional groups. **Stereochemistry**: Concept of isomerism, types of isomerism, optical isomerism, elements of symmetry, molecular chirallity, enantiomers, stereogenic centres, optical activity, properties of enantiomers, chiral and achiral molecules with two stereogenic centres, distereoisomers, mesocompounds, resolution of enantiomers. Relative and absolute configurations, sequence rules, D & L , R & S systems of nomenclature. **Heteroaromatics:** Five / six membered hetero aromatics and analogues, Nucleic acid bases, Structure, electron rich electron deficient heterocycles.

Text Books:

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

Reference Books:

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

CBIO 608 - INTRODUCTORY BIOLOGY

Total Credits: 3 **Total: 35 Hrs.**

Unit I **7 Lectures**

Diversity of Life forms: Diversity of living organisms - Classification of the living organisms (five kingdom classification, major groups and principles of classification within each kingdom). Systematics and binomial System of nomenclature - Salient features of animal (non-chordates up to phylum level and chordates up to class level) and plant (major groups; Angiosperms up to class - linnaeus) classification.

Unit II **7 Lectures**

Inheritance biology: Mendelian principles- Dominance, segregation, independent assortment, Codominance, incomplete dominance, genomic imprinting, linkage and crossing over; extra chromosomal inheritance, microbial genetics, human genetics-pedigree analysis, mutations, recombination, structural and numerical alterations of chromosomes.

Unit III **7 Lectures**

Developmental Biology: Basic concepts of development, gametogenesis, fertilization and early development, morphogenesis and organogenesis in animals and plants, programmed cell death, aging and senescence.

Unit IV **7 Lectures**

Ecology & Evolution: Habitat and niche, population growth curves, Ecosystems stability- species interactions, competition, conservation methods (both in situ and ex situ); Origin of life, theories and evidences.

Unit V **7 Lectures**

Applied Biology: Microbial fermentation and production of micro and macro molecules, Tissue and cell culture methods for plants and animals, Transgenic animals and plants, Genomics and its application to health and agriculture, Breeding in plants and animals, Bioremediation and phytoremediation, Biosensors.

Text books

1. Concepts of Biology (2nd Ed.) by Sylvia S. Mader. McGraw Hill Publishers, 2011.
2. Molecular Biology of the cell (4 th Ed.) by Bruce Alberts. Garland publishing Inc. 2002

References

1. Fundamentals of Ecology Eugene P.*ODUM* 1971 (Third Edition): W. B. Saunders, Comp. Philadelphia – London – Toronto.
2. Developmental biology by Scott F Gilbert 9th edition.
3. Principles of Gene Manipulation- An Introduction to Genetic Engineering By S.B. Primrose, university of California Press

CBIO 651-CELL BIOLOGY – LAB

Credit: 1

List of Experiments:

1. Microscopical identification of cells in permanent slides
2. Identification of microorganisms from soil by serial dilution
3. Blood cell differentiation by Giemsa staining
4. Preparation of Culture media and growth determination by sigmoid curve
5. Axenization by streaking and pour plate techniques
6. Estimation of Chlorophylls
7. Estimation of proteins
8. Estimation of DNA
9. Estimation of total carbohydrates by anthrone method
10. Estimation of Aminoacid by Ninhydrin Method
11. Estimation of Cholesterol by Zak's Method
12. Estimation of Ascorbic Acid by volumetric Method
13. Stages of mitosis by staining in allium sepa.

CBIO 652 – PROGRAMMING IN C++ - LAB

Total Credits: 1

1. Simple C++ programs to demonstrate various decision making and loop constructs.
2. Working with matrices.
3. Demonstration of switch construct.
4. User defined functions.
5. Working with pointers.
6. String handling functions.
7. Creating and working with classes.
8. Illustration of constructors and destructors.
9. Demonstration of scope resolution operator.
10. Operator and function overloading.
11. Simple and multiple inheritance.
12. Overloading a constructor.

CBIO 621 - MATHEMATICS FOR BIOSCIENCES

Total Credits: 3 **Total: 34 Hrs.**

Unit 1 **6 Lectures**

Reviewing Limits, Continuity and Differentiability: Limits of Functions, Continuity of Functions; Basics of Differentiation- Differentiability, Derivatives, Interpretations of Derivatives, General Rules of Differentiations

Unit 2 **7 Lectures**

Integration: Review of Definite Integrals, Double (Surface) Integrals - Definition, Iterated Integrals (Fubini's Theorem), Properties; Triple (Volume) Integrals- Definition, Properties, Geometric Interpretation of Double and Triple Integrals

Unit 3 **7 Lectures**

Differential Equation: Ordinary Differential Equation (ODE)- Definition, Equations of First order and degree- Homogeneous and Non-homogeneous Equations, Exact Differential Equations, Condition, Rules of Finding Solution, Partial Differential Equation (PDE)- Definition, Derivation of a PDE, PDE of First Order, Linear PDE

Unit 4 **7 Lectures**

Introduction to Laplace Transform: Definition, Some Elementary Functions and their Laplace Transform, Piecewise Continuity, Sufficient Conditions for Existence and Important Properties of Laplace Transform, Convolution of Laplace Transform, Inverse Laplace Transform

Unit 5 **7 Lectures**

Introduction to Fourier Transform: Definition, Fourier Series, Fundamental Fourier Transform Properties, Fourier Transform Application

Text Books:

1. Fundamentals of University Mathematics (3rd Ed.), Colin McGregor, Woodhead Publishing in Mathematics, 2010
2. Introduction to Mathematics for Life Scientists, Edward Batschelet, Springer

Reference Books

1. Higher Engineering Mathematics (40th Ed), B.S. Grewal and J.S. Grewal, Khanna Publishers, 2007

CBIO 622 ALGORITHMS IN COMPUTATIONAL BIOLOGY

Total Credits: 3 **Total: 36 Hrs.**

Unit 1 **9 lectures**

DNA Computing: DNA Structure, and Processing , Computational operations and Step involve in DNA computing, Bio-soft Computing Based on DNA Length, Beginnings of Molecular Computing-Adelman Experiment. **RNA secondary structure prediction**: Base pair maximisation and the Nussinov folding algorithm, Energy minimisation and the Zuker folding algorithm, Design of covariance models, Application of RNA fold.

Unit 2 **10 lectures**

Combinatorial Pattern Matching- Hash Tables, Repeat Finding, Exact Pattern Matching; **Genetic Algorithm**:- Basic Concepts, Reproduction, Cross over, Mutation, Fitness Value, Optimization using GAs; Applications of GA in bioinformatics.

Unit 3 **9 lectures**

Hidden Markov Model: Markov processes and Markov Models, Hidden Markov Models. Forward and Backward Algorithms, Most probable state path: Viterbi algorithm, Parameter Estimation for HMMs:-Baum-Welch Algorithm, EM Algorithm, Applications of profile HMMs for multiple alignment of proteins and for finding genes in the DNA.

Unit 4 **7 Lectures**

Support Vector Machines: Introduction, hyperplane separation (maximum and soft margin hyperplanes), linear classifier, Kernel functions, Large Margin Classification, Optimization problem with SVM, Applications of SVM in bioinformatics. **Bayesian network**: Bayes Theorem, Inference and learning of Bayesian network, BN and Other Probabilistic Models.

Unit 5 **8 lectures**

Artificial Neural Network: Historic evolution – Perceptron, characteristics of neural networks terminology, models of neuron Mc Culloch – Pitts model, Perceptron, Adaline model, Basic learning laws, Topology of neural network architecture, single layer ANN, multilayer perceptron, back propagation learning, input - hidden and output layer computation, back propagation algorithm, Applications of ANN

Text Books:

1. Biological sequence analysis: Probabilistic models of proteins and nucleic acids by Richard Durbin, Eddy, Anders Krogh, 1998
2. An introduction to bioinformatics algorithms by Neil C. Jones, Pavel Pevzner. MIT Press.2004
3. Algorithms for Molecular Biology by Ron Shamir Lecture, Fall Semester, 2001
4. Neural Networks: A Systematic Introduction by Raul Rojas. Springer. 1996
5. DNA Computing: New Computing Paradigms By Gheorghe Paun, Grzegorz Rozenberg, Arto Salomaa.

Reference Books:

1. Bioinformatics: the machine learning approach by Pierre Baldi, Søren Brunak. MIT Press.2001

CBIO 623 - JAVA AND BIOJAVA IN COMPUTATIONAL BIOLOGY

Total Credits: 3 **Total: 36 Hrs.**

Unit 1 **7 lectures**

Introduction to Java: Compilation of java programs – Java Development kit – virtual machine – byte code – data types (int, long, char, and Boolean) – operators (arithmetic, relational, bitwise and assignment) – arrays – operator precedence – type conversion – control statements and loops.

Unit 2 **9 lectures**

Working with java classes: Declaring classes – super and sub classes – constructors – instances of classes – inheritance (simple, multiple and multilevel) – overriding and overloading – exception handling – file handling.

Unit 3 **7 lectures**

Multi-thread programming: Life cycle of a thread – creating a thread (extension of thread class and implementing runnable) – thread priorities – synchronization – deadlock.

Unit 4 **6 lectures**

Event handling and applets: Event handling mechanisms – delegation event model – event classes – event listener interfaces – mouse and keyboard events – adapter classes and inner classes. Applet basics – passing parameters to applets – applet display methods – drawing lines, ovals, rectangles and polygons – threads and animation.

Unit 5 **7 lectures**

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

Text Books:

1. Java:The completer Reference. (7^{th} Ed.) by Herbert Schildt, TMH. 2012

CBIO 624 – DRUG DISCOVERY AND IPR

Total Credits: 3 **Total: 36 hrs**

Unit 1 8 Lectures

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

Unit 2 6 Lectures

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

Unit 3 6 Lectures

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

Unit 4 6 Lectures

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

Unit 5 10 Lectures

Intellectual Property Rights (IPR): IPR Definition and implications for discovery & development. Forms of IPR Protection - Copyright, Trademark and Patents. International organization and treaties for IPR protection – World Trade Organization (WTO) & Trade Related Aspects of Intellectual Property Rights (TRIPS) Agreements. Importance of IPR in Indian Scenario & Indian laws for IPR protection. **Patents**: National and international agencies for patenting - US Patent & Trademark office (USPTO), Controller General of Patents, Designs & Trade Marks, India (CGPDTM), World Intellectual Property organization (WIPO)-Patent Cooperation Treaty (PCT); Requirements for patentability, Composition of a patent, How to apply and get patents – US, Indian and PCT.

Text Books:

1. **Drugs: From discovery to approval** 2nd ed by Rick NG. Wiley Blackwell (2009)
2. **Intellectual Property Rights** by Deborah E. Bouchoux,. Delmar Cenage Learning. 2005

Reference books:

1. **Burger's Medicinal Chemistry and Drug discovery**. Volume 2, Drug Discovery and development.6th Edition. Ed Donald J AbrahamWiley-Interscience.
2. **Essentials of Medical Pharmacology**, 6the Edition (Hardcover) by Tripathi Kd. Publisher: Jaypee Brothers (2008)
3. **Laws of Patents: Concepts and Cases** Edited by A. V. Narasimha Rao © 2005 The ICFAI University Press
4. **Intellectual Property Rights In India: General Issues And Implications** by Prankrishna Pal. Publisher: Deep & Deep Publications Pvt.ltd (2008)

CBIO 625 – MOLECULAR EVOLUTION

Total Credits: 3 **Total: 36 Hrs.**

Unit 1 **6 Lectures**

History of evolution of life on earth: Chemical basis of evolution, Evolution of DNA, RNA and proteins, origin of the genetic code. Hardy-Weinberg equilibrium; Evolutionary changes by mutation, gene flow, genetic drift and natural selection.

Unit 2 **8 Lectures**

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

Unit 3 **10 Lectures**

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

Unit 4 **6 Lectures**

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

Unit 5 **6 Lectures**

Evolution of the genome: Human Genome Project, ENCODE, Genome 10 K, Genome duplication (Ohno's hypothesis), Gene duplication, Exon Shuffling, Concerted evolution.

Reference Books:

1) Molecular Evolution by Wen Hsiung-Li, 1997, Sinauer Associates, Sunderland, MA. ISBN 0878934634.
2) Evolution (3rd Edition) by Ridley, M., 2004, Blackwell Science. ISBN 1-4051-0345-0

CBIO 626 – GENOMICS AND PROTEOMICS

Total Credits: 3 **Total: 36 Hrs.**

Unit 1 **8 Lectures**

Genomics and Metagenomics: Large scale genome sequencing strategies. Metagenomics, basic principles, application of methods to prokaryotic and eukaryotic genomes and interpretation of results. Basic concepts on identification of disease genes, role of bioinformatics-OMIM database, reference genome sequence, integrated genomic maps, gene expression profiling; identification of SNPs, SNP database (DbSNP). Role of SNP in Pharmacogenomics, SNP arrays. Basic concepts in identification of Drought stress response genes, insect resistant genes, nutrition enhancing genes

Unit 2 **7 Lectures**

Epigenetics: DNA microarray: database and basic tools, Gene Expression Omnibus (GEO), ArrayExpress, SAGE databases DNA microarray: understanding of microarray data, normalizing microarray data, detecting differential gene expression, correlation of gene expression data to biological process and computational analysis tools (especially clustering approaches)

Unit 3 **7 Lectures**

Comparative genomics: Basic concepts and applications, whole genome alignments: understanding the significance; Artemis, BLAST2, MegaBlast algorithms, PipMaker, AVID, Vista, MUMmer, applications of suffix tree in comparative genomics, synteny and gene order comparisons Comparative genomics databases: COG, VOG

Unit 4 **7 Lectures**

Functional genomics: Application of sequence based and structure-based approaches to assignment of gene functions – e.g. sequence comparison, structure analysis (especially active sites, binding sites) and comparison, pattern identification, etc. Use of various derived databases in function assignment, use of SNPs for identification of genetic traits. Gene/Protein function prediction using Machine learning tools viz. Neural network, SVM etc

Unit 5 **7 Lectures**

Proteomics: Protein arrays: basic principles. Computational methods for identification of polypeptides from mass spectrometry. Protein arrays: bioinformatics-based tools for analysis of proteomics data (Tools available at ExPASy Proteomics server); databases (such as InterPro) and analysis tools. Protein-protein interactions: databases such as DIP, PPI server and tools for analysis of protein-protein interactions

Text Books:

1. Discovering Genomics, Proteomics and Bioinformatics 2nd edition - by A. Malcolm Campbell and Laurie J. Heyer. by Cold Spring Harbor Laboratory Press 2006.

Reference books:

1. Principles of Genome Analysis and Genomics (3rd Ed.) by Primrose, S.B. and Twyman, R.M., Blackwell Publishing Company, Oxford, UK. 2003
2. Introduction to Proteomics – Tools for the new biology (1st Ed.) by Liebler, D.C., Humana Press Inc., New Jersey, USA. 2002
3. Bioinformatics and Functional Genomics by Pevsner, J., John Wiley and Sons, New Jersey, USA. 2003
4. Bioinformatics: Sequence and Genome Analysis by Mount, D., Cold Spring Harbor Laboratory Press, New York. 2004

CBIO 627-STRUCTURAL BIOLOGY

Total Credits: 3 **Total: 36 Hrs**

Unit I **8 Lectures**

Fundamentals of protein structure- amino acids fundamental building blocks, Peptide bond, rigid planar peptide unit, *cis* and *trans* configuration. **Structural Heirarchy**: Primary, Secondary, Tertiary, Quaternary structures. **Motifs and domains**: α - domain structures, β - domain structures, α/β (alpha/beta) - structures. **Types of proteins**: globular, membrane, fibrous proteins. **Complex proteins**: Virus, ribosomes

Unit II **8 Lectures**

Principles of nucleic acid structure: Chemical structure of nucleic acids, Watson and Crick's base-pairings and their implications. Non Watson and Crick pairing schemes - base stacking interactions - DNA polymorphism - structure of ADNA, BDNA and ZDNA - helical transitions. Non-uniform helical DNA Structure. Unusual DNA structures - hairpins, bulges, cruciform, triplexes, tetraplexes.

Unit III **8 Lectures**

Elementary crystallography: Introduction: symmetry in crystals, lattices and unit cells, crystal systems, Bravais lattices, **Elements of symmetry** - rotation axis, mirror planes and center of inversion, proper/ improper axes of rotation, translational symmetry- screw axis and glide planes. **Symmetry operation:** classes of symmetry operations, classification of symmetry point groups and molecular space groups and equivalent points. X-ray diffraction - Laue equations - Bragg's law - reciprocal lattice and its application to geometrical Crystallography.

Unit IV **6 Lectures**

X-ray scattering: Atomic scattering factor - diffraction by a space lattice - structure factor equation - electron density and Fourier series - Fourier Transform and crystal diffraction - diffraction by real crystals - Lorentz and polarization factor - primary and secondary extinctions.

Unit V **6 Lectures**

Nuclear Magnetic Resonance:- Introduction**,** Nuclear spin, NMR sensitivity, shielding and deshielding effects of NMR, nuclear Overhauser effect. **Spectral parameters**: chemical shift, spin-spin splitting, coupling, non-equivalent proton. Carbon-13 NMR spectra of protein, FTNMR, spin-spin splitting, proton spin decoupling, off-resonance decoupling. 1D- NMR spectra, 2D-NMR spectroscopy.

References Books:

1. Introduction to protein structure, C. Branden and J. Tooze
2. Molecular Modeling Principles and Applications (2^{nd} Ed.) by Andrew R. Leach., Prentice Hall, USA. 2001
3. X-Ray Structure Determination: A Practical Guide, 2nd Edition, by George H. Stout , Lyle H. Jensen
4. Principles of Protein Structure by G. E. Schulz., Springer 2009
5. Lehninger Principles of Biochemistry by David L. Nelson and Michael M. Cox, W. H. Freeman.2005
6. Structural Bioinformatics, Philip E. Bourne, Helge Weissig, Wiley Publication

CBIO 655 – STRUCTURAL BIOLOGY – LAB

Total Credits: 1

1) Different types of structure representation and implications – PyMol, Chimera
2) Surface calculation and implications: Hydrophobic, charge representation
3) Secondary structure prediction
4) Structure based alignment
5) Structural Blast - DALI
6) Binding pocket prediction – Castp; Glycosilation, phosphorilation sites prediction
7) Modeller – homology modeling, threading
8) Energy Minimisation
9) Validation of models – Procheck, Whatif, Verify 3d,
10) Auto dock
11) NMR structure analysis – Demo
12) Structure analysis of Chemical compounds using IR - Demo

CBIO 656- PROGRAMMING IN JAVA -LAB

Total Credits: 1

1. Simple java programs to demonstrate decision making, and loops.
2. Handling of arrays and working with matrices.
3. Working with Classes and objects in java.
4. Use of constructors and demonstration of overloading of constructors.
5. Demonstration of simple, multiple and multilevel inheritances.
6. Exception handling.
7. Creation of multiple threads.
8. Reading and writing files.
9. Applets.
10. Animation and Threads.
11. Managing Simple Events and Interactivity.
12. Alignment of sequences (biojava)

CBIO 701 BIOMOLECULAR SIMULATIONS

Total Credits: 3 **Total: 36 Hrs.**

Unit 1 **8 Lectures**

Molecular Mechanics: Introduction, The Morse Potential, The Harmonic Oscillator Model for Molecules, Comparison of Morse and Harmonic Potential, Two atoms connected by a bond, Poly atomic Molecules, Energy due to Stretch, Bend, Stretch-Bend, Torsional strain, van der Waals and Dipole-Diploe interactions. Types of Potentials: Lennard-Jones, Truncated Lennard-jones, Exponential-6, Ionic and Polar potentials. Types of Force Fields: AMBER, CHARMM, Merck Molecular Force Field, Consistent Force Field, MM2, MM3 and MM4 force fields.

Unit 2 **5 Lectures**

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

Unit 3 **8 Lectures**

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

Unit 4 **8 Lectures**

Molecular modelling in Drug design:- Conformational analysis, lead identification, optimization and validation. Methods and Tools in Computer-aided molecular Design, Analog Based drug design:- Pharmacophores and QSAR. Structure based drug design:- Docking, De Novo Drug Design, Virtual screening.

Unit 5 **7 Lectures**

Structure Activity Relationship: Introduction to QSAR, QSPR, Various Descriptors used in QSARs, Regression Analysis, Significance and Validity of QSAR Regression Equations, Partial Least Squares (PLS) Analysis, Multi Linear Regression Analysis. Application of Genetic Algorithms, Neural Networks and Principle Components Analysis in QSAR analysis.

Text Books:

1. Computational Chemistry and Molecular Modeling-Principles and Applications by Ramachandran, Deepa and Namboori, 2008, Springer_Verlag. Reference for Unit 1 and 2.
2. Molecular Modeling Principles and Applications (2nd Ed.) by Andrew R. Leach, Prentice Hall, USA. 2001

Reference:

1. Molecular Modelling for Beginners, (2nd Edition) by Alan Hinchliffe, John Wiley & Sons Ltd. 2008
2. Molecular Modeling and Simulation – An interdisciplinary Guide by Tamar Schlick, Springerverlag. 2000
3. Computational medicinal chemistry for drug discovery edited by Patrick Bultinck, Marcel Dekker Inc. 2004

CBIO 702 BIOPHYSICAL TECHNIQUES

Total Credits: 3 **Total: 36 Hrs.**

Unit 1 **8 Lectures**

Spectroscopy: Introduction to spectroscopy: basic principles, instrumentation and applications of UV-VIS absorption, infrared, Raman, fluorescence spectroscopy

Unit 2 **8 Lectures**

Application of Spectroscopy to macromolecules: Amino acid, Protein absorption at UV spectra, DNA absorption spectrum, Protein-DNA interaction study using UV spectra. CD and ORD introduction, linear and circular Dichroism for biological molecules, secondary structure prediction using CD. NMR application to macromolecules. Mass spectroscopy and application to macromolecules.

Unit 3 **5 Lectures**

Scattering from Solutions of Macromolecules: Principles of light scattering, Rayleigh scattering, scattering from particles comparable to wavelength of radiation, static light scattering, dynamic light scattering, low angle X−ray scattering

Unit 4 **8 Lectures**

Separation techniques: Chromatography- column chromatography, TLC, paper chromatography, adsorption chromatography, partition chromatography, Gas liquid chromatography, Ion exchange chromatography, Molecular exclusion chromatography, affinity chromatography, Hydrophobic interaction chromatography. **Electrophoresis**: Moving boundary electrophoresis, zone electrophoresis, low voltage electrophoresis, high voltage electrophoresis, gel electrophoresis, SDS, Iso electric focusing, continuous flow electrophoresis, capillary electrophoresis in DNA sequencing. Centrifugation, Ultra centrifugation.

Unit 5 **7 Lectures**

Membrane Biophysics and neuro Biophysics: Membrane Constituents : Review of chemistry and biochemistry of constituents of membranes - lipids, phospholipids, lipoproteins, models of membrane structure. **Nervous System**: Organization of the nervous system - Membrane potentials - origins of membrane potential - electrochemical potentials - Donnan equilibrium - Nernst equation - Goldman equation.

Reference Books:

1. Biophysics, V. Pattabhi, N. Gautham, 2002, Narosa Publishing House
2. Guide to protein purification, 2nd edition, Methods in enzymology V463, JN. Abelson and Melvin I. Simon
3. Spectroscopy for the biological sciences, Gordon G. Hammes, John Wiley & Sons, Inc., Publication
4. Lehninger Principles of Biochemistry, Fourth Edition, David L. Nelson, Michael M. Cox
5. Introduction to biological membrane, Jain RK
6. Biomembrane structure and function, Chapman D

CBIO 703 SYSTEMS BIOLOGY

Total Credits: 3 **Total: 36 Hrs.**

Unit-I **5 Lectures**

Networks and graph theory: Basic properties of Network: Degree, average degree and degree distribution. Adjacency matrix, weighted and unweighted networks, Bipartite network, Paths and distances.

Unit-II **5 Lectures**

Random Networks: Erdos-Renyi model, Small-world effect, clustering coefficient.

Scale-free networks: Power laws, Hubs, ultra-small property, degree exponent, The Barabasi-Albert Model. Degree correlations: assortativity and disassortativity.

Unit-III **8 Lectures**

Biological networks: Complex Biological Systems, Types of Biological networks, Intra-cellular networks: Gene-regulatory network, Protein-interaction network, Metabolic networks and Signaling network; Inter-cellular networks: Neuronal networks, Network motifs, Network medicine.

Unit-IV **8 Lectures**

Modularity: Motifs and sub-graphs, Feed-forward loops, Single-input modules: LIFO, FIFO. Dense overlapping regulons (DORs). Optimal gene design circuits: fitness function and optimal expression of a protein in bacteria, Robustness.

Unit-V **10 Lectures**

Constraint-based modelling – Metabolic reconstruction, Flux Balance Analysis (FBA): Translating biochemical networks into linear algebra, Stoichiometric matrix, Elementary mode, Extreme pathways, Objective function, Optimization using linear programming. Genome-scale cellular models: Virtual Erythrocytes, Global human metabolic model (Recon 1).

Text Books:

1. Networks: An Introduction by M.E.J. Newman, Oxford University Press, 2010.
2. Introduction to Systems Biology: Design Principles of Biological Circuits by Uri Alon, Chapman & Hall/CRC, 2007.

Reference Books:

1. Introduction to Systems Biology, S. Choi, Humana Press, 2007.
2. Linked – The New Science of Networks, Albert-László Barabási, Perseus Publishing, 2002.

CBIO 704 NEXT GENERATION SEQUENCING

Total Credits: 3 **Total: 36 Hrs.**

Unit-I 6 **Lectures**

NGS Platforms: Introduction to NGS, Roche/454 FLX, Illumina/Solexa Genome Analyzer, Applied Biosystems SOLiD system, Helicos Heliscope, Pacific Biosciences/single molecule real time (SMRT) sequencing.

Unit-II 8 **Lectures**

Genome assembly algorithms: Alignment of short-reads to reference genome using spaced seed (ELAND, SOAP), index-filtering algorithm (SeqMap), quality-score (RMAP), q-filter algorithm (SHRiMP), FM-index (Bowtie, BWA,SOAP2), suffix tree (MUMmer). Sequence Alignment formats: Sequence Alignment/Map (SAM) format, Binary Alignment/Map (BAM) format, Tools for conversion (SAMtools), Alignment viewers (IGV, MGAviewer).

Unit-III 8 **Lectures**

De-novo assembly: Overlap-layout-consensus (OLC) approach (Arachne, Phusion), de Bruijn and Euler path approach (Euler, SOAPdenovo), string graph assembler (SGA).Scaffolding: Supercontig, contig orientation, contig ordering, contig distancing and gap closing using SOAPdenovo, ABySS, OPERA and RACA.

Unit-IV 8 **Lectures**

Application of R in NGS analysis: Introduction to Bioconductor, Reading of RNA-seq data (ShortRead, Rsamtools, GenomicRanges), annotation (biomaRt, genomeIntervals), reads coverage and assign counts (IRanges, GenomicFeatures), differential expression (DESeq).

Unit-V 6 **Lectures**

Biological applications of NGS: Whole-genome sequencing, Exome sequencing, Transcriptome sequencing, Epigenome sequencing, Interactome sequencing, methylome sequencing.

Text Books:

1. Next-generation DNA sequencing Informatics by Stuart M. Brown, Cold Spring Harbor Laboratory, 2013.
2. RNA-seq Data Analysis: A Practical Approach by Eija Korpelainen, Jarno Tuimala, Panu Somervuo, Mikael Huss, Garry Wong. Chapman & Hall/CRC, 2014.

Reference Books:

1. Next generation sequencing: Translation to Clinical Diagnostics by Wong Lee-Jun C. (ed.), Springer, 2013.
2. Next-generation genome sequencing: Towards Personalized Medicine by Michal Janitz, Wiley-VCH, 2008

CBIO 705 DATA MINING AND DATA WAREHOUSING

Total Credits: 3 **Total: 36 Hrs.**

Unit 1 **5 lectures**

Introduction

Need for data warehouse, definition, goals of data warehouse, Data Mart, Data warehouse architecture, extract and load process, clean and transform data, Designing fact tables, partitioning, Data warehouse and OLAP technology.

Importance of Data Mining, Relational Databases, Data Warehouses, Transactional Databases, Advance Database Systems and Applications, Data Mining Functionalities, Classification of Data Mining Systems, Major issues in Data Mining.

Unit 2 **6 lectures**

Primitives and System Architectures

Architectures of Data Mining Systems, Data Mining Primitives, Data Mining Query Language, Designing Graphical User, Interfaces Based on a Data Mining Query Language

Unit 3 **7 lectures**

Concept Description and Association Rules

Concept Description, Characterization and comparison, Data Generalization and Summarization - Analytical Characterization, Mining Class Comparisons, Association Rule Mining, Mining Association Rules in Large Databases, Mining Single-Dimensional Boolean Association Rules from Transactional Databases, Mining descriptive statistical measures in large data bases, multidimensional association rules from relational DBS and DWS, Correlation analysis, Constraint based association mining.

Unit 4 **6 lectures**

Classification and Prediction Classification and Prediction, Issues: Data preparation for classification and Prediction, Comparing classification Methods, Classification by Decision Tree Induction, Back propagation, Bayesian classification.

Unit 5 **7 lectures**

Clustering Methods

Clustering Analysis, Types data in clustering analysis: Scaled variable, Binary variables, Variables of Mixed Types, Partitioning Methods: K-means and K-Medoids, Data Mining Applications: Data mining for Biomedical and DNA Data Analysis.

Text Books:

1. Data Mining Concepts and Techniques – Jiawei Hen, Micheline Kambler, Academic Press Morgan Kaufman Publishers. 2006
2. Building Data Ware House by W.H.Inmon, John Wiley & Sons
3. Data warehousing by S. Anahory and D.Murray, Pearson Education, ASIA

CBIO 706 PERL PROGRAMMING FOR BIOLOGISTS

Total Credits: 3 **Total: 34 Hrs.**

Unit 1 **7 Lectures**

Basic Datatypes: Scalar Variables, Scalar Operations and Functions, Array Variables, Literal Representation of Array, Array Operations and Functions, Scalar and List Context, Hash Variables, Literal Representation of a Hash, Hash Functions

Unit 2 **7 Lectures**

Perl Regular Expression: Concepts on Regular Expressions, Uses of Regular Expressions in Biology, Patterns, Matching Operator, Substitutions, Split and Join functions.

Unit 3 **7 Lectures**

Modular Programming: Subroutines, Advantage of Subroutines, Scoping and Subroutines, Arguments, Passing Data to Subroutines, Modules and Libraries of Subroutines, Concept on File handle, Opening and Closing a File Handle, Opening and Closing a Directory Handle, Reading a Directory Handle, File and Directory Manipulation.

Unit 4 **7 Lectures**

Common Gateway Interface (CGI): The CGI.pm Module, CGI program in Context, Simple CGI programs, Passing Parameters via CGI, Perl and the Web

Unit 5 **8 Lectures**

Bioperl: Introduction to Bioperl, Installing Procedures, Architectures, General Bioperl Classes, Sequences -Bio::Seq Class, Sequence Manipulation, Features and Location Classes-Extracting CDS, Alignments -AlignIO, Analysis -Blast, Databases- Database Classes, Accessing a Local Database

Text Books

1. Beginning Perl for Bioinformatics (1st Ed.), J. Tisdall, O'Reilly, 2004
2. Learning Perl (5th Ed.), Randal L. Schwartz, Tom Phoenix and Brain d foy, O'Reilly, 2008
3. Programming Perl (3rd Ed), W. Wall, T. Christiansen and J. Orwant, O'Reilly, 2000

CBIO 707 PROJECT

Total Credits: 4

The course is designed to result in the satisfactory completion and defense of the Masters dissertation.

This process includes

a) the conceptualization of the independent research that will comprise the dissertation,
b) the preparation of and satisfactory defense of the dissertation proposal,
c) the collection, analysis, and interpretation of data,
d) presentation of findings in the dissertation format, and
e) oral defense of the dissertation.

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

CBIO 751 BIMOLECULAR SIMULATIONS - LAB

Total Credits : 1

Exercises

1. Molecular Visualization Softwares: Pymol and Rasmol
2. Geometry Optimization
3. Tutorial on Molecular Dynamics: Gromacs
4. Binding Site Identification
5. Structure based Drug Design:- Molecular Docking
6. Ligand based Drug Design:- QSAR

CBIO 752 DATA MINING AND DATA WAREHOUSING -LAB

Total Credits : 1

Exercises

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

CBIO 721 - BIOMEDICAL INFORMATICS AND TRANSLATIONAL RESEARCH

Total Credits: 2 **Total 24 hrs**

Unit I **4 Lectures**

Overview of Medical Informatics - Healthcare functions and information technology, Key Players in Health Information technology (HIT), Organizations involved with HIT, Barriers to HIT Adoption. Public Health Informatics - Information systems in public health - National Health Information Infrastructure (NHII). Internet based consumer health information - telehealth and telemedicine.

Unit II **6 Lectures**

Biomedical data - Their acquisition, storage and use, Electronic health records (EHR), Information Retrieval from Digital Libraries, Imaging Systems in Radiology and Picture archiving. Genomics and Proteomics data - Human Genome project, HapMap and 1000 genomes projects, Genetic profiling of individuals and large populations, Creation and use of Bioinformatics databases - gene, metabolic pathways of diseases.

Unit III **4 Lectures**

Managing Information Security and Privacy in Health Care Data. General approaches to assuring appropriate use of data, data tracking and deidentifying data. Methods and Evaluation in biomedical decision making: Sampling, appropriate use of controls, data collection, testing of statistical significance, sensitivity and specificity, ROC plots. Standards in Biomedical informatics; Ethics, legal and regulatory matters in health informatics.

Unit IV **4 Lectures**

Clinical Decision-Support Systems - The Nature of clinical decision making, types of decisions, The role of computers in decision support, Historical perspectives- Leeds abdominal pain system, MYCIN, HELP; Ilustrative examples of clinical decision-support systems-Internist-1, DXplain system. Patient monitoring system and information management in intensive care unit.

Unit V **6 Lectures**

Translational Research - Concepts and Principles. Therapeutic discovery in an academic setting, Technology Transfer and Commercialization process of a product. Bringing drugs from bench to bedside for cancer therapy - Molecular basis of cancer, strategies for developing therapeutic treatments, how imatinib and dasatinib were developed. Principles of Clinical Trials: Genetics/-Omics in Clinical Investigation, Principles of biomarker development and utility, pharmacogenomics including utilization of key knowledge from the human genome projects for personalized medicine. Regulatory and ethical issues involved in translational clinical research.

Text Books:

1. Biomedical Informatics: computer applications in Health care and Biomedicine (3rd ed), by Shortliffe EH, Ciminio JJ., 2000, New York Springer-Verlag, ISBN 0-387-28986-0.
2. Translational Research in Genetics and Genomics, Ed. Moyra Smith; 2008, Oxford University press, ISBN: 978-0-19-531376-5.

Reference Books:

1. Evaluation methods in medical Informatics by Friedman CP. Wyatt JC, 1996, New York Springer Verlag, ISBN 0-387-25899-2.
2. Biomedical Informatics in Translational Research, Ed. Hai Hu, Richard J. Mural and Michael N. Liebman, 2008 Artech House, INC, ISBN-13: 978-1-59693-038-4.

CBIO 722 – GENETIC ENGINEERING-(rDNA Technology)

Total Credits: 3 **Total 36 hrs**

UNIT I **7 Lectures**

Scope of Genetic Engineering, Milestones in Genetic Engineering, Genetic engineering guidelines, Regulatory Procedures: Good laboratory practice, Good manufacturing practice and FDA regulations - Regulations for recombinant DNA research and manufacturing process - Bio-safety and Bioethics - Regulations for clinical trials, Documentation and Compliance, in India and selected countries

UNIT II **8 Lectures**

Nucleic Acid cloning and amplification methods: Molecular Tools in genetic engineering: Restriction enzymes, Restriction Mapping of DNA Fragments and Map Construction. Ligases, S1 nuclease, terminal deoxynucleotides, transferases, polymerases, Reverse Transcriptase and Alkaline phosphatase. Gene Cloning Vectors- Plasmids, bacteriophages, phagemids, cosmids, artificial chromosomes. Ligation – transformation methods, Polymerase chain reaction, Primers, Gene amplification and applications, Real Time PCR, Microarray techniques, DNA Sequencing methods; Enzymatic sequencing; Chemical sequencing of DNA; Automated DNA sequencing; RNA sequencing.

UNIT III **8 Lectures**

cDNA Synthesis and cDNA library preparations: Cloning mRNA enrichment, reverse transcription, DNA primers, Linkers, adaptors and their chemical synthesis, Library construction and screening. Genomic libraries (complete sequencing projects). Alternative Strategies of Gene Cloning Cloning interacting genes- screening for genes of interest – site directed mutagenesis – Labeling of DNA: Nick translation, Random priming, Radioactive and non-radioactive probes, Hybridization techniques: Northern, Southern and Colony hybridization, Fluorescence in situ hybridization; Chromatin Immunoprecipitation; DNA-Protein Interactions-Electromobility shift assay; DNaseI footprinting;.

UNIT IV **6 Lectures**

Gene Regulation methods: S1 mapping, RNase protection assay, Reporter assays. Transgenic and Gene Knockout Technologies Targeted gene replacement, Strategies of gene delivery, gene regulation silencing and transcription factors. Differential gene expression and protein array.

UNIT V **7 Lectures**

Expression Strategies for Heterologous Genes, Vector engineering and codon optimization, host engineering, *In-vitro* transcription and translation, expression in bacteria, expression in Yeast, expression in insects and insect cells, expression in mammalian cells, expression in plants. Processing of Recombinant Proteins Purification, Characterization of recombinant proteins. Transgenic plants, animals, genetically modified organisms (GMO) and GM food etc. Agricultural: Hybrid and modified seeds - Bio-pesticides - Bio-fertilizers.

BOOKS

1. *Molecular Cloning, a Laboratory Manual*, J. Sambrook, E.F. Fritsch and T. Maniatis, Cold Spring Harbor Laboratory Press, New York, ,1st edition, 2000.
2. *DNA Cloning- a Practical Approach,* .M. Glover and B.D. Hames, IRL Press, Oxford, 1st edition, 1995.

REFERENCES:

1. *Molecular and Cellular Methods in Biology and Medicine*, P.B. Kaufman, W. Wu. D. Kim and L.J; Cseke, CRC Press, Florida, 6th edition,2006.
2. Lewin's GENES XI by Jocelyn E. Krebs, Elliott S. Goldstein and Stephen T. Kilpatrick (Dec 31, 2012)
3. *Methods in Enzymology* vol. 152, *Guide to Molecular Cloning Techniques*, S.L. Berger and A.R. Kimmel, Academic Press, Inc. San Diego, 2nd edition,1998,
4. *Methods in Enzymology* Vol 185, Gene Expression Technology, D.V. Goeddel, Academic Press, Inc., San Diego, 1990.
5. *DNA Science. A First Course in Recombinant Technology*, D,A. Mickloss and G.A. Froyer. Cold Spring Harbor Laboratory Press, New YorK, 1st edition, 1990.
6. *Molecular Biotechnology*, S.B. Primrose. Blackwell Scientific Publishers, Oxford, 2nd Edition, 1994.
7. *Milestones in Biotechnology. Classic papers on Genetic Engineering, J*.A. Davies and W.S. Reznikoff, Butterworth-Heinemann, Boston, 2nd edition, 1992.
8. *Route Maps in Gene Technology*, M.R. Walker and R. Rapley, Blackwell Science Ltd., Oxford, 6th edition, 2001.

CBIO 723 PROJECT

Total Credits: 8

The course is designed to result in the satisfactory completion and defense of the Masters dissertation.

This process includes

f) the conceptualization of the independent research that will comprise the dissertation,
g) the preparation of and satisfactory defense of the dissertation proposal,
h) the collection, analysis, and interpretation of data,
i) presentation of findings in the dissertation format, and
j) oral defense of the dissertation.

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