

## **BINC SYLLABUS for Paper-II**

### **BINC BioInformatics Syllabus - Advanced**

#### **Sequence analysis**

**Scoring matrices:** Detailed method of derivation of the PAM and BLOSUM matrices

**Pairwise sequence alignments:** Needleman and Wunsch, Smith and Waterman algorithms and their implementation

**Multiple sequence alignments (MSA):**

**Use of HMM-based Algorithm** for MSA (e.g. SAM method)

#### **Sequence patterns and profiles:**

**Repeats: Tandem and Interspersed repeats, repeat finding, Motifs, consensus, position weight matrices**

Algorithms for derivation of and searching sequence patterns: MEME, PHI-BLAST, SCanProsite and PRATT

Algorithms for generation of sequence profiles: Profile Analysis method of Gribskov, HMMer, PSIBLAST

**Protein and nucleic acid properties:** e.g. Proteomics tools at the ExPASy server and EMBOSS

**Taxonomy and phylogeny:** Phylogenetic analysis algorithms such as maximum Parsimony, UPGMA, Transformed Distance, Neighbors-Relation, Neighbor-Joining, Probabilistic models and associated algorithms such as Probabilistic models of evolution and maximum likelihood algorithm, Bayesian inference algorithm, Bootstrapping methods, use of tools such as PHYLIP, MEGA, PAUP

Analysis of regulatory RNA's: Databases and tools

#### **Structural Biology**

Experimental methods for Biomolecular structure determination: X-ray and NMR

Identification/assignment of secondary structural elements from the knowledge of 3-D structure of macromolecule using DSSP and STRIDE methods

Prediction of secondary structure: PHD and PSI-PRED methods

**Tertiary Structure prediction:** Fundamentals of the methods for 3D structure prediction (sequence similarity/identity of target proteins of known structure, fundamental principles of protein folding etc.) Homology Modeling, fold recognition, threading approaches, and ab -initio structure prediction methods **Structure analysis and validation:** Pdbsum, Whatcheck, Procheck, Verify3D and ProsaII

Critical Assessment of Structure Prediction(CASP)

Structures of oligomeric proteins and study of interaction interfaces

### **Molecular modeling and simulations**

Macro-molecular force fields, salvation, long-range forces

Geometry optimization algorithms: Steepest descent, conjugate gradient

Various simulation techniques: Molecular mechanics, conformational searches, Molecular Dynamics, Monte Carlo, genetic algorithm approaches, Rigid and Semi-Flexible Molecular Docking

### **Genomics**

Large scale genome sequencing strategies

Genome assembly and annotation

Genome databases of Plants, animals and pathogens

Metagenomics

Gene networks: basic concepts, computational model such as Lambda receptor and *lac* operon Prediction of genes, promoters, splice sites, regulatory regions: 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

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.

Next Generation sequencing & assembly: Elements of big data analysis, NGS Platforms based on pyrosequencing, sequencing by synthesis, emulsion PCR approach with small magnetic beads and single molecule real time (SMRT) sequencing; Genome assembly algorithms, *De-novo* assembly algorithms, Sequence Alignment formats: Sequence Alignment/Map (SAM) format, Binary Alignment/Map (BAM) format.

### **Comparative genomics:**

Basic concepts and applications, BLAST2, MegaBlast algorithms, PipMaker, AVID, Vista, MUMmer, applications of suffix tree in comparative genomics, synteny and gene order comparisons Comparative genomics databases: Clusters of Orthologous Groups (COGs), Ensembl

### **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: supervised/unsupervised learning, Neural network, SVM etc

### **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 STRINGS, DIP, PPI server and tools for analysis of protein-protein interactions

### **Modeling biological systems**

Systems biology – Topology of biological networks: Random vs Scale-Free networks. Use of computers in simulation of cellular subsystems: Simulation and analysis of biochemical networks and their dynamics using ODEs and stochastic algorithm, Flux Balance Analysis (FBA), Boolean network simulations.

Metabolic networks, or network of metabolites and enzymes, Signal transduction networks, Gene regulatory networks, Metabolic pathways: databases such as KEGG, EMP , MetaCyc, AraCyc

## **Drug design**

Drug discovery process

Role of Bioinformatics in drug design

Target identification and validation and lead optimization

Different systems for representing chemical structure of small molecules like SMILES etc

Generation of 3D coordinates of small molecules

Structure-based drug design: Identification and Analysis of Binding sites and virtual screening

Ligand based drug design: Structure Activity Relationship – QSARs and QSPRs, QSAR Methodology, Pharmacophore mapping

*In silico* prediction ADMET properties for Drug Molecules

## **Vaccine design:**

Reverse vaccinology and immunoinformatics

Databases in Immunology

Principles of B-cell and T-cell epitope prediction

## **Suggested Books for Reading:**

David W Mount, Bioinformatics: Sequence and Genome Analysis, 2nd Edition, Cold Spring Harbor Press

Durbin et al (2007) Biological Sequence Analysis: Probabilistic models of protein and Nucleic acids Cambridge University Press.

Stuart M. Brown (2013) Next-generation DNA sequencing Informatics. Cold Spring Harbor Press

M.E.J. Newman (2010) Networks: An Introduction, Oxford University Press.

Thomas E. Creighton, Proteins: structures and molecular properties

Chemoinformatics Edited by Johann Gasteiger and Thomas Engel

Structural Bioinformatics, Edited Philip E. Bourne and Helge Weissig

Lee A Segel (2008), Biological Kinetics, Cambridge University Press Cambridge

Cornish-Bowden (2012), Fundamentals of Enzyme Kinetics ,Wiley-Blackwell

Alberghina L (2005), System Biology : Definitions and Perspectives, Springer-Verlag Berlin Heidelberg.

Najarian K, Najarian S, Gharibzadeh S, Eichelberger CN (2009) Systems Biology and Bioinformatics: A Computational Approach, CRC Press.

Klipp E, Liebermeister W, Wierling C, Kowald A, Lehrach H, Herwig R (2009) Systems Biology : A Text Book, Wiley-Blackwell

Integrative approaches for finding modular structure in biological networks, NATURE REVIEWS , GENETICS, VOLUME 14, OCTOBER 2013

BIOINFORMATICS, Vol. 19 no. 2, 2003

Nucleic Acids Research (2014), Vol. 42, Database issue D199–D205 doi:10.1093/nar/gkt1076

Nucleic Acids Research (2012), Vol. 40, Database issue D109–D114 doi:10.1093/nar/gkr988

An extended bioreaction database that significantly improves reconstruction and analysis of genome-scale metabolic networks (2011), Integrative Biology, 2011.3, 1071-1086.

Computational Systems Bioinformatics — Methods and Biomedical Applications By Xiaobo Zhou (Harvard Medical School and Brigham & Women's Hospital, USA), Stephen T C Wong (Harvard Medical School and Brigham & Women's Hospital, USA).

Bioinformatics for Systems Biology (2009) by Stephen Krawetz, Published by Humana Press.

## **BINC Biology Syllabus – Advanced**

### **Cell Biology and Genetics**

Membrane structure and membrane transport, Intercellular compartment and protein sorting, Vesicular transport and protein traffic in cells

Different mechanisms of signal transduction, concepts in signal network, second messenger, molecules involved in various signaling pathways such as G-protein coupled receptors, protein kinases, calcium binding proteins

Extensions of Mendelian principles: Codominance, incomplete dominance, gene interactions, gene and environment interactions, pleiotropy, genomic imprinting, penetrance and expressivity, linkage and crossing over, sex linkage

Gene mapping methods: Linkage maps, tetrad analysis, mapping with molecular markers, mapping by using somatic cell hybrids, development of mapping population in plants.

Extra chromosomal inheritance: Inheritance of mitochondrial and chloroplast genes, maternal inheritance.

Basic microbial genetics: Bacterial and viral genetic system. Extrachromosomal DNA and resistance genes. Methods of genetic transfers – transformation, conjugation, transduction and sexduction, mapping genes by interrupted mating, fine structure analysis of genes.

Basic Human genetics: Pedigree analysis, linkage testing, karyotypes, genetic disorders. Population genetics: Hardy Weinberg Law, Effect of mutation, migration, selection and genetic drift on population.

## **Molecular Biology**

Genome organization of complex organisms, complexity of genome – C value paradox, Structure of bacterial and Eukaryotic chromosomes.

Molecular basics of prokaryotic and eukaryotic DNA replication and recombination.

Gene Expression: initiation, elongation and termination of transcription, template and enzyme properties, promoter and regulatory sequences. Genetic code and translation. Post transcriptional modifications, processing of RNA and proteins

Principles of Gene regulations: Gene regulations in bacterial cells -Inducible and repressible operons, Positive control and catabolite repression. Eukaryotic gene regulations – Histone modifications, Posttranscriptional and posttranslational regulations. Methods for studying gene expression and regulatory sequences, large-scale expression analysis, use of microarrays.

Methods for studying variation and polymorphism at genome level, PCR, northern. Southern, western blotting, RFLP, fingerprinting, RAPDs, DNA and protein sequencing methods

Epigenetic mechanisms of inheritance, regulatory RNA molecules (miRNA, siRNA), antisense RNA and their applications

Recombinant DNA technology: Restriction and ligation of DNA fragments, cloning.

## **Biochemistry**

Enzyme kinetics, Lineweaver-Burk plot, competitive and non competitive inhibition

Molecular mechanisms of interactions of small and large molecules including ions, regulation of protein pathways, mechanism of enzyme action and ribozymes

Isoenzymes, allosteric enzymes, regulation by covalent modification

Mechanism of enzyme action with special reference to lysozyme and carboxypeptidase – A.  
Functional classification of enzymes, concept of EC numbers.

Organization of metabolic systems: Enzyme chains, multi enzyme complexes and multifunctional enzymes and regulatory enzymes.

Inborn errors of metabolism. Concept of biochemical regulation, feedback and feed forward systems, biochemical oscillations

### **Suggested Books for Reading:**

1. Life, the biology of science, 10<sup>th</sup> edition, David Sadava
2. Genetics: A conceptual approach, 5<sup>th</sup> Edition, Benjamin Pierce.
3. Principles of Biochemistry by Lehninger
4. The Molecular biology of the Cell, 5<sup>th</sup> Edition, Bruce Alberts
5. Genes X by B. Lewin
6. Essential Cell Biology 2<sup>nd</sup> Ed B.Alberts, D.Bray, K.Hopkin and A.Johnson
7. Biochemistry, 6<sup>th</sup> Edition Berg, Jermy M., Tymoczko, John L and Styer, Lubert
8. Biochemistry Vol 1: Biomolecules, mechanisms of enzyme action and metabolism Voet, D and Voet, J.
9. Molecular Biology of the Gene, Watson, JD., Hopkins, NH., Roberts, JW and Steitz, JA

### **BINC Physical Science Syllabus -Advanced**

Electrostatics, Coulomb's Law, Dielectric Constant, dipole-dipole and dipole-monopole interaction Basic Properties of lasers.

Quantum mechanics: the time-independent Schrödinger equation, energy eigenvalues and eigenfunctions. Simple solvable models in Quantum Mechanics: Particle in a box, harmonic oscillator, hydrogen like atoms.

Vibrations and Waves, Properties of Schrodinger wave function

### **Suggested Books for Reading:**

1. Fundamentals of Physics, 6th Edition David Halliday & Robert Resnik
2. The Feynman Lectures on Physics: Volumes 1, 2 & 3

### **BINC Chemical Science Syllabus: Advanced**

Concept of free energy, activation energy. Polar molecules, molar refraction and polarization. Basic Molecular orbital theory. Dipole moment, potentiometric determination of pK of aminoacids.

Biophysical techniques for determining size and shape of macromolecules – ultra centrifugation, electrophoresis and chromatography. Application of spectroscopy and X-ray diffraction for determination of biomolecular secondary and tertiary structure. Mass-spectroscopy of biological molecules.

Luminescence, fluorescence, phosphorescence.

First law of thermodynamics, isothermal process, entropy and second law of thermodynamics, reversible and irreversible processes; Concepts and calculations of enthalpy, internal energy and potential energy, free energy, entropy; Inter-relation between potential energy and force, heat of formation.

Basic principles of chemical kinetics – Zero, first and second order kinetics, enzyme kinetics: Michaelis-Menten kinetics; Activation energy and the Arrhenius law.

### **Suggested Books for Reading:**

Physical Chemistry, P.W. Atkins and Julio de Paula

Organic Chemistry, Morrison & Boyd.

Biophysical chemistry vol I, Charles R Cantor & Paul Reinhard Schimmel

### **BINC Mathematics, Statistics Syllabus –Advanced**

#### **Mathematics**

Vector – addition, subtraction, multiplication and their geometric understanding.

Matrices, their eigenvalues and eigenvectors.

Differential equations: Second order linear differential equation and initial value problems.



## **Statistics**

Sampling distribution: chi-square, t and F distributions.

Central limit theorem.

Methods of least squares and regression analysis.

Estimation: un-biased, maximum likelihood, Bayesian.

Testing of hypothesis: Type I and Type II errors, power of a test, p-value.

Large sample test: one and two sample tests for mean variance.

Confidence interval for unknown mean and variance

Markov chains, their transition probability and stationary distribution.

## **Suggested Books for Reading:**

Hogg, Mckean and Craig: Introduction mathematical Statistics 6<sup>th</sup> edition Pearson, Prentice Hall, 2005.

Sheldon M. Ross: Introduction to probability models, 9<sup>th</sup> edition, Academic Press, 2007.

Gilbert Strang: Linear Algebra and its application, 4<sup>th</sup> edition, Cengage Learning, 2006.

NCERT class 12 mathematics books.

Ewens and Grant: Statistical methods in bioinformatics.

## **BINC Information Technology Syllabus: Advanced**

- A. Data Structures and Algorithm  
Arrays, Link Lists, Stacks, Queues, Graphs, Trees, Sorting, Searching, string comparison  
– Programs to be implemented using C or Python or Java or Perl
- B. Databases  
SQL, indexing and Hashing.
- C. Elements of scripting languages.
- D. Elements of NoSQL

**Suggested Books for Reading:**

1. Database Management System – Ramakrishnan and Gehrke
2. Data Structure : Andrew S Tannenbaum
3. Complete Reference to C
4. Complete Reference to Java
5. Complete reference to Perl
6. Complete Reference to Python