

# **BINC SYLLABUS for all Papers**

## **BINC Bioinformatics Syllabus - Basic**

### **Major Bioinformatics Resources: NCBI, EBI, ExPASy, RCSB**

The knowledge of various databases and bioinformatics tools available at these resources, organization of databases: data contents and formats, purpose and utility in Life Sciences

### **Open access bibliographic resources and literature databases:**

Open access bibliographic resources related to Life Sciences viz., PubMed, BioMed Central, Public Library of Sciences (PLOS)

### **Sequence databases**

Formats, querying and retrieval

Nucleic acid sequence databases: GenBank, EMBL, DDBJ;

Protein sequence databases: Uniprot-KB: SWISS-PROT, TrEMBL, PIR-PSD

Repositories for high throughput genomic sequences: EST, STS GSS, etc.;

Genome Databases at NCBI, EBI, TIGR, SANGER

Viral Genomes

Archeal and Bacterial Genomes;

Eukaryotic genomes with special reference to model organisms (Yeast, Drosophila, C. elegans, Rat, Mouse, Human, plants such as Arabidopsis thaliana, Rice, etc.)

### **3D Structure Database: PDB, NDB**

Chemical Structure database: Pubchem

Gene Expression database: GEO, SAGE

### **Derived Databases**

Knowledge of the following databases with respect to: basic concept of derived databases, sources of primary data and basic principles of the method for deriving the secondary data, organization of data, contents and formats of database entries, identification of patterns in given sequences and interpretation of the same

Sequence: InterPro, Prosite, Pfam, ProDom, Gene Ontology

Structure classification database: CATH, SCOP, FSSP

Protein-Protein interaction database: STRING

**Compilation of resources:** NAR Database and Web server Issues and other resources published in Bioinformatics related journals

### **Sequence Analysis**

**File formats:** Various file formats for bio-molecular sequences: GenBank, FASTA, GCG, MSF etc

**Basic concepts:** Sequence similarity, identity and homology, definitions of homologues, orthologues, paralogues

**Scoring matrices:** basic concept of a scoring matrix, Matrices for nucleic acid and proteins sequences, PAM and BLOSUM series, principles based on which these matrices are derived

**Pairwise sequence alignments:** basic concepts of sequence alignment: local and global alignments, Needleman and Wunsch, Smith and Waterman algorithms for pairwise alignments, gap penalties, use of pairwise alignments for analysis of Nucleic acid and protein sequences and interpretation of results

**Multiple sequence alignments (MSA):** the need for MSA, basic concepts of various approaches for MSA (e.g. progressive, hierarchical etc.). Algorithm of CLUSTALW and PileUp and their application for sequence analysis (including interpretation of results), concept of dendrogram and its interpretation

#### **Database Searches:**

Keyword-based searches using tools like ENTREZ and SRS

Sequence-based searches: BLAST and FASTA

**Sequence patterns and profiles:** Basic concept and definition of sequence patterns, motifs and profiles, various types of pattern representations viz. consensus, regular expression (Prosite-type) and sequence profiles; profile-based database searches using PSI-BLAST, analysis and interpretation of profile-based searches

**Taxonomy and phylogeny:** Basic concepts in systematics, taxonomy and phylogeny; molecular evolution; nature of data used in Taxonomy and Phylogeny, Definition and description of phylogenetic trees and various types of trees

**Protein and nucleic acid properties:** Computation of various parameters using proteomics tools at the ExPASy server and EMBOSS

**Comparative genomics:** Basic concepts and applications, whole genome alignments: understanding significance. Artemis as an example

#### **Structural Biology**

**3-D structure visualization and simulation:** Visualization of structures using Rasmol or SPDBV or CHIME or VMD

Basic concepts in molecular modeling: different types of computer representations of molecules. External coordinates and Internal Coordinates

Non-Covalent Interactions and their role in Biomolecular structure and function

Fundamentals of Receptor-ligand interactions.

**Proteins:** Principles of protein structure; Peptide bond, phi, psi and chi torsion angles, ramachandran map, anatomy of proteins – Hierarchical organization of protein structure – Primary. Secondary, Super secondary, Tertiary and Quaternary structure; Hydrophobicity of amino acids, Packing of protein structure, Structures of oligomeric proteins and study of interaction interfaces

**DNA and RNA:** types of base pairing – Watson-Crick and Hoogsteen; types of double helices (A, B, Z), triple and quadruple stranded DNA structures, geometrical as well as structural features; structural and geometrical parameters of each form and their comparison; various types of interactions of DNA with proteins, small molecules

RNA secondary and tertiary structures, t-RNA tertiary structure

**Carbohydrates:** The various building blocks (monosaccharides), configurations and conformations of

the building blocks; formations of polysaccharides and structural diversity due to the different types of linkages

Glyco-conjugates: various types of glycolipids and glycoproteins

### **Classification and comparison of protein 3D structures:**

Purpose of 3-D structure comparison and concepts, Algorithms : CE, VAST and DALI, concept of coordinate transformation, RMSD, Z-score for structural comparison

Databases of structure-based classification; CATH, SCOP and FSSP

**Secondary structure prediction:** Algorithms viz. Chou Fasman, GOR methods; nearest neighbor and machine learning based methods, analysis of results and measuring the accuracy of predictions.

**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/comparative Modeling, fold recognition, threading approaches, and *ab initio* structure prediction methods

# **BINC Bioinformatics Syllabus - Advanced**

## **Sequence analysis**

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

**Pairwise sequence alignments:** Needleman and Wuncsh, 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, PSI-BLAST

**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, 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 assesment 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 tools (especially clustering approaches)

### **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)

### **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 – Use of computers in simulation of cellular subsystems

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.

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

# **BINC Biology Syllabus: Basic**

## **Cell Biology and Genetics**

Basic aspects of Prokaryotic and eukaryotic cells (plant and animal cells); membranes and cellular compartments, cell organelles, structure and function

Cell motility and shape: cytoskeletal elements, cilia and flagella; motor proteins

Cell-cell interactions: Intercellular junctions

Cell cycle and its regulation, events during mitosis and meiosis

Concepts of Bioenergetics, respiration, electron transport systems.

Concepts of gene: Allele, multiple alleles, pseudoallele, complementation tests.

Mendelian principles : Inheritance, sex linked inheritance, Dominance, segregation, independent assortment.

Mutations : Types, causes and detection, Mutant types – lethal, conditional, biochemical, loss of function, gain of function, germinal versus somatic mutants, point/deletion/insertional mutations.

Basic concepts in immunology, antigen-antibody interaction, humoral and cell mediated immunity. Concepts of development and pattern formation.

## **Molecular Biology**

Prokaryotic genome organization and structure

Prokaryotic gene expression, factors involved in gene regulation

Eukaryotic genome organization and structure, mechanisms of gene expression in Eukaryotes, basic mechanisms of transcription and translation

Mechanisms of genome alterations: Recombination, mutation, inversion, duplication, transposition.

Basic concepts of replication.

## **Biochemistry**

Carbohydrates and lipids, their importance in cells

Proteins: Amino acids and their physico-chemical properties, peptide bond and peptides

Nucleic acids: Nucleosides, nucleotides, RNA and DNA. Denaturation and renaturation of DNA

Enzymes: Units of activity, coenzymes and metal cofactors, temperature and pH effects, Michaelis-Menten kinetics, inhibitors and activators, active site

Carbohydrate metabolism: Glycolysis, gluconeogenesis, glycogenolysis, glycogenesis, TCA cycle and oxidative phosphorylation

Pentose phosphate pathway; hormonal control,  $\beta$ -oxidation and biosynthesis of fatty acids

Transamination and deamination of amino acids, ketogenic and glyco-genic amino acids, urea cycle

Purine and pyrimidine biosynthesis

# **BINC Biology Syllabus – Advanced**

## **Cell Biology and Genetics**

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, 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 : 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 Principle

## **Molecular Biology**

Genome organization, initiation, elongation and termination of transcription, template and enzyme properties, promoter and regulatory sequences. Regulation of translation, Post transcriptional modifications, processing of RNA and proteins

Methods for studying gene expression and regulatory sequences, large-scale expression analysis, use of microarrays

Genetic information transfer, details of regulation in eukaryotes and prokaryotes, horizontal gene transfer

Organization of eukaryotic genomes, 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

## **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, feed back and feed forward systems, biochemical oscillations

**Suggested Books for Reading:**

1. Principles of Biochemistry by Lehninger
2. The Molecular biology of the Cell by Alberts et al
3. Genes X by B. Lewin
4. Essential Cell Biology 2<sup>nd</sup> Ed B.Alberts, D.Bray, K.Hopkin and A.Johnson
5. Biochemisty, 6<sup>th</sup> Edition Berg, Jermy M., Tymoczko, John L and Styer, Lubert
6. Biochemistry Vol1: Biomolecules, mechanisms of enzyme action and metabolism Voet, D and Voet, J.
7. Molecular Biology of the Gene, Watson, JD., Hopkins,NH., Roberts, JW and Steitz,JA

## **BINC Physical Science syllabus : Basic**

Particle dynamics, Newton's laws of motion, velocity, acceleration, momentum. Conservative forces, Conservation of Energy.

First law of thermodynamics, second law of thermodynamics, reversible and irreversible processes, Isothermal, isobaric and quasistatic processes. Concepts of Enthalpy and Entropy, Interrelation between potential energy and force. Thermodynamic, Gibbs and Helmholtz free energies.

Chemical potential. First-order phase transitions

Equation of state for ideal gases. Departures from ideality. Maxwell-Boltzman Distribution

## **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.

### **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 : Basic**

Hybridization states of atoms. Electronic structure of molecules, Chemical bonding (ionic bonds, covalent bonds, hydrogen bond, hydrophobic effects, coordinate bonds). Basic Molecular orbital theory. Valence bonds. Non-covalent bonding in protein structure.

Tautomerization, geometrical isomerism, inductive effect, stereochemistry (R/S,D/L); nucleophile, electrophile, nucleophilic substitution, electrophilic substitution, nomenclature of organic compounds. Bioisosterism.

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

Concept of pH, pK, chemical equilibrium, Henderson-Hasselbach equation, structure of water.

## **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 and Julio de Paula

Organic Chemistry , Morrison & Boyd.

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

## **BINC Mathematics and Statistics Syllabus – Basic**

### ***Mathematics***

Functions and Graphs of polynomial, logarithm, exponential and trigonometric functions.

2D co-ordinate geometry: Equation of a line, circle, ellipse, parabola, hyperbola; focal point, eccentricity and other properties.

3D geometry: Equation of sphere.

Solution of simultaneous and quadratic equations

Sequences and series.

Limits.

Differentiation and integration of the above mentioned functions.

Matrix algebra: Multiplication, inverse and solution of linear equations.

### ***Statistics***

Discrete random variables, their probability mass function, probability distribution function, mean and variance.

Binomial and Poisson random variables and their properties.

Continuous random variable, their probability density function, probability distribution function, mean and variance.

Normal random variable and its properties.

Conditional probability and Bayes' theorem.

## **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 distributions: Chi-square, t and F distributions.

Central limit theorem.

Methods of least squares and regression analysis.

Estimation: un-biased, maximum likelihood.

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 and variance.

Confidence interval for unknown mean and variance

Markov chains, their transition probability and stationary distributions.

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

Hogg, Mckean and Craig: Introduction to 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 : Basic**

Fundamentals in Computing

Types of Processing: Batch, Real-Time, Online, Offline.

Types of modern computing: Workstations, Servers, Parallel Processing Computing, Cluster computing, Grid computing

Introduction to operating systems: Operating System concept, UNIX/LINUX.

Basic Programming Concepts – sequential, conditional and loop constructs

Introduction to Database Systems

SQL Queries

Geometric transformation

Coordinate transformations

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

### A. Data Structures and Algorithm

Arrays, Link Lists, Stacks, Queues, Graphs, Trees – Programs to be implemented using either C or Java

Sorting, Searching, string comparison – programs using Perl

### B. Databases

SQL, indexing and Hashing.

## **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