Swami Ramanand Teerth Marathwada University, Nanded

BIOINFORMATICS SYLLABUS
For Ph.D. Entrance
Section - B

Unit I- Basics of Biological Science

Origin of Life Prebiological chemical evolution, Proteinoids, Protocells. **Systematic** Species concepts, Kingdom to species, the five kingdoms –Classical, Phenetic and cladistic approaches. **Bacteria** Structure of bacterial cell, Bacterial types, Transformation, Transaction, Transduction and conjugation, Nutrition, Phylogeny. **Viruses** Biology of viruses, Bacteriophases, Plant and Animal Viruses, Replication of Virus, genome, HIV. **Fungi:** Distinguishing characters of fungi, Morphology, Classification of fungi, Yeasts, Fungal role in plant diseases, brewing and antibiotics. **Algae and Lichens** Characteristics, Classification (Outline) biological and economic importance of Algae, Lichens: Morphology, Reproduction, Symbiotic nature, Chemical Interaction.

Unit II- Introduction to Plant and Animal world

Plant Diversity: Plant world classification up to the level of genus and species, Angiosperms multicellular fungi, Bryophytes, Pteridophytes, monocot, dicot, Gymnosperms (One example each), Morphology General Organization of plant body such as aerial, underground parts. Differences of Angiosperms and gymnosperms.

Animal Life: Major animal phyla, characteristics and interrelationships; tissues, organs and organ systems; principles of nutrition, digestion, thermoregulation, Osmoregulation and excretion, muscle contraction, neural reflexes, circulation, respiration and endocrines.

Unit III - Biocomputing

Computer system Characteristics and Capability

Basic structure, ALU, Memory, CPU, I/O devices, development of computers, classification of computers (microcomputers, minicomputers, mainframe, supercomputer, PC, server, workstation) binary number system, low level and high level languages. Flow charts and programming techniques

Input/Output Devices Keyboard entry, direct entry: Card readers, scanning devices (Barcode, OMR MICR), Voice input devices, pointing devices (light pen, mouse, touch screen, digitizer, scanner), printers: Dot matrix, inkjet, laser, plotter, CRT, LCD, CD-Writer, ZIP drive

Memory RAM, ROM, PROM, EPROM, EEPROM etc., Base memory, Extended memory, expanded memory, virtual memory, cache memory, storage devices: tape, FDD, HDD, CD, DVD

Operating System Fundamentals of DOS, file and directory, booting procedure of DOS, DOS commands Features of MS-Windows, GUI, Multitasking etc., Main modules of Windows OS: program manager, control panel, networks, Elements of Windows: desktop, windows, applications, icons, group window,

Switching between applications: running MS-DOS applications, Windows help Windows Accessories: Notepad, paintbrush, study of important files of windows (e.g. DLL, INI, etc.),

Networking environment: concept of internet, internet and their uses and application, Domain Name Formats (e.g. .COM, .EDU, .MIL, .ORG, .NET, etc) (internal and external), configuration of DOS (config.sys), Batch file concept (autoexec.bat)

MS Office: MS office covering word processing- MS Word, spreadsheet- MS Excel and presentation of softwares MS Power Point.

Unit IV- Introduction to Bioinformatics

Introduction to Bioinformatics

Various definitions of bioinformatics, history of bioinformatics, bioinformatics in business
Fundamentals of internet: WWW, HTML, URL, HTTP
Browsers: - Netscape, Opera, Explorer, Google Chrome, Mozila Firefox
Search engines: - Google, SRS, Entrez, PubMed, , Unigene, PDB, Swissprot and TrEMBL
Database Searching: Nucleotide Databases -NCBI GenBank, DDBJ, EMBL, BRENDA, TrEMBL, KEGG. Microarray –MIAME, GEO Diseases -OMIM Database. Protein databases- PIR, UniProt PDB. BLAST: BLAST X, PHI-PSI- BLAST, FASTA

Unit V -Mathematics and Statistics

Real Number Classification and measurements, algebraic law, inequalities, summation, power and fraction power. **Sets and Symbolic Logic** Sets notating and symbols, variables members, complementary sets union, symbolic logic, Boolean algebra, negation and implication
Relation and Functions Definition, types of relations, equivalence relations, partial orderings, definition of function as relations, composition of function, inverse function.
Calculus Differentiation: Calculation gradients of cords first and higher order derivatives,
Application: Increasing and decreasing functions, maximum and minimum points derivatives as rates of change.
Matrix Algebra Addition, subtraction and multiplication of matrix, transpose of matrix, inverse of matrix, conjugative matrix. **Limits and Complex Numbers** Limits of sequences, series, limit of functions, the Fibonacci sequence, complex plane, algebraic operations, exponential function of complex variable, Oscillations

Statistics Definitions: Webster's and Secrist's definitions of statistics, importance of statistics, Scope of statistics: Industry, economy, computer science, social science, concept of statistical population with illustration, concept of sample with illustrations
Elementary Statistics Representation of data, discrete data, continuous data, histogram, polygram, frequency curves. Arithmetic Mean: - Definition of grouped and ungrouped data, combine mean, weighted mean, merits and demerits. Median: - definition, formula and computation for ungrouped and grouped data, graphical methods, merits and demerits Mode: - definition, formula and computing for grouped and ungrouped data, merits and demerits. Quartiles: Definition, formula for grouped and ungrouped data, merits and demerits. Numerical problems
Probability Experimental probability, probability when outcomes are equality likely subjective probabilities, probabilities law, probability rules for combined events, conditional probability and independent events, probability trees, bayes theorem
Random Variables and Distributions Discrete and continuous random variables, cumulative distribution function, probability mass function and probability density function, Bernoulli distribution, binomial distribution, poison distribution, Uniform distribution, normal distribution, normal approximation to binomial distribution, central limit theorem
Measures of Dispersion Concept of dispersion, measures of dispersion, Absolute and relative measures of dispersion. Range: - Definition for grouped and ungrouped data, combined variance for two grouped, merits and demerits Variance.

Swami Ramanand Teerth Marathwada University, Nanded

BIOINFORMATICS SYLLABUS

For Ph.D. Entrance

Unit I- Molecular Cell Biology & Genetics

Molecules of Cell. Cells and Genomes Cell Chemistry and Biosynthesis
An overview of organelles, (mitochondria, chloroplasts, ER, Golgi, lysosomes and peroxisomes; nucleus and nucleolus) and organelle genetic systems. Working of Cell

Internal Organization of Cell
Prokaryotic and eukaryotic cell membranes and cellular compartmentation.
Cellular membranes: Structure, transport, channels, carriers, receptors, endocytosis, membrane potentials.
Cell motility and shape: cytoskeletal elements, cilia and flagella; motor proteins.

Cell-cell interactions and signal transduction: Intercellular junctions, signaling by hormones and neurotransmitters; receptors, G-proteins, protein kinases and second messengers
Protein traffic in cells: Protein sorting and signal sequences; protein translocation in ER and vesicular transport to Golgi, lysosmes and plasma membrane; protein import into nuclei, mitochondria, chloroplasts and peroxisomes.

Cell cycle and its regulation; events during mitosis and meiosis; Mechanics of cell division; Programmed cell death.

Integrating Cells into Tissues
Cell-Cell and cell matrix adhesion: an overview. The extracellular matrix of epithelial sheets
The extracellular matrix of non-epithelial tissues, Plant Tissues.
Universal mechanism of animal development Cell movements and shaping of the vertebrate body. Plant development

Science of genetics – objectives, terminologies, methods Mendelian principles of inheritance, sex linked inheritance Chromosomes, linkage, linkage maps and recombination Mutations – molecular, gene/point and chromosomal Genetics of viruses and bacteria Phenotype and genotype relationships, role of environment, from gene to phenotype, gene interactions
Study of quantitative traits Genetics of populations, genetics and evolution

Unit II –Biochemistry

Water as the universal biological solvent; concept of osmolarity.

Carbohydrates: monosaccharides, oligosaccharides, polysaccharides, proteoglycans and glycoproteins.

Lipids: fatty acids, acylglycerols; phospholipids, sphingolipids, cholesterol and membranes Isoprenoids, icosanoids and their biological importance.

Proteins: amino acids and peptides; primary, secondary, tertiary and quaternary structures; structure, function and evolutionary relationships; protein – protein interactions protein folding; allosteric proteins.

Nucleic acids: bases, nucleotides, RNA and DNA; different structural forms of DNA; denaturation, renaturation and hybridization of DNA; different types of RNA; Protein-nucleic acid interaction.

Enzymes: details of enzyme nomenclature and classification; units of enzyme activity; coenzymes and metal cofactors; temperature and pH effects;

Michaelis-Menten kinetics; Inhibitors and activators; active site and catalytic mechanisms; covalent and non-covalent regulations; isoenzymes; osmolytes and intracellular modulation of enzymes.

Oxidation of glucose in cells: high energy bond, glycolysis, citric acid cycle and oxidative phosphorylation.

Unit III- Genetic Information flow and Processing

Introduction

Introduction: DNA as a genetic material.– Experiments done to prove this.
Nucleic acid structure: Single stranded, double stranded. Secondary structures in single stranded molecules. Alternative double helical structures in double stranded DNA; Closed DNA as Super coiled molecule.

Genome Organization and Evolution

Prokaryotic and eukaryotic genomes: C value paradox. Repetitive and non-repetitive DNA, transposons and retroposons. Exons and introns – organization of interrupted genes, one DNA sequence may code for multiple proteins. Gene numbers – essential genes and total gene number, gene clusters, Pseudogenes. Gene families – globin gene and immunoglobulin gene families;

Organelle genome – mitochondrial and chloroplast. Packaging of genome – Bacterial genome as nucleoid. Eukatryotic genome – nucleosomes, chromatin, solenoids, loops, domains, scaffolds, chromosomes. Evolution of Genomes.

DNA replication
Prokaryotic DNA replication – DNA polymerases, origin of replication, initiation, elongation and termination of replication, Rolling circle model of replication.
Eukaryotic DNA polymerases – multiple origins of replication, process of replication; Regulation of replication in both prokaryotes and eukaryotes.

DNA damage, repair and recombination
Different types of DNA damages. Variety of DNA repair systems in prokaryotes and eukaryotes – Base excision repair system. Nucleotide excision repair system, Mismatch repair system, Recombination repair system; Recombination – homologous and nonhomologous recombination.

Gene Expression – Transcription and Translation:
Transcription in prokaryotes – RNA polymerase, initiation, elongation and termination of transcription. Regulation of transcription – operon concept, lactose and Tryptopahn operons.
Transcription in eukaryotes: Different RNA polymerases – requirement of promoters by these RNA polymerases, Initiation, elongation and termination by these polymerases;. Processing of transcripts – 5' capping, 3' polyadenylation, splicing and editing; Regulation of transcription – Response elements, enhancers and silencers, HLH, Leucine zipper proteins., noncoding RNAs.
Translation in prokaryotes: protein synthesis machinery – mRNA, tRNA and rRNA molecules; initiation, elongation and termination of translation; Genetic code – interpreting genetic code; Accuracy of translation. Eukaryotic translation: protein synthesis – initiation, elongation and termination;

Post-transnational modifications of proteins; protein degradation; Regulation of translation - mRNA stability, 5' and 3' UTRs, mRNA Localization. Gene regulation by post-translational modifications of proteins (acetylation, methylation, ribosylation, phosphorylation etc.) and different intermediate RNAs (ribozymes, miRNAs, siRNAs etc.)

Unit IV- Chemo Informatics & Biodiversity Informatics

Introduction
Role of Chemoinformatics in pharmaceutical/chemical research.
Integrated databases; HTS analysis; Ligand based design of compounds; Structure based design of compounds. Structure representation systems, 2D and 3D structures.
Chemical Databases – Design, Storage and Retrieval methods
Introduction to database filters, property based & (drug-like)-Lipinski Rule of Five

Characterization of chemicals
Introduction to pharmocophore Identification of pharmacophore features. Building pharmacophore hypothesis Searching databases using pharmocophores. Chemoinformatics tools for drug discovery

QSAR & applications to Hit to lead optimization
Introduction to QSAR. Methods- Hansch analysis, Free-wilson QSAR, GFA. QSAR using public domain tools. *in silico* ADMET; QSAR approach, Knowledge-based approach. Design & Analysis of combinatorial libraries

Biological Diversity
Biological Information. Biodiversity of life – Species diversity: taxonomic information plants, animals, microbes and viruses; Genetic diversity; Ecological/ecosystem diversity; urban biodiversity Methods for species identification classification

Biodiversity Databases
Organizing biological species information Accessing existing databases on World Wide Web – Species 2000; Tree of Life Project; GBIF; ATCC; NBII; ICTV; AVIS; Species analyst Collaboration

Species Identification
Software for species identification. Probabilistic and deterministic identification Delta, Micro IS, AVIS, ICTV. Bio complexity issues in Biodiversity

Unit V- Immunology and Parasite Bioinformatics

Introduction
Immune systems and systems biology – Innate and adaptive immunity in vertebrates; Antigen processing and presentation.
Antibodies – Immunoglobulins, Immunoglobulin classes and subclasses CDR and LDR regions and sequence numbering; Immunogenetics Membrane receptors for antigen – The B-cell surface receptor for antigen
(BCR), The T-cell surface receptor for antigen (TCR); Antigen recognition diversity; The major histocompatibility complex (MHC) Contemporary challenges to the immune system – Infectious diseases, Clustering of infectious disease organisms, Autoimmune diseases

Epitopes and primary interaction with antigen
Epitopes – Affinity Maturation, Recognition of Antigen by B cells, Neutralizing Antibody, Prediction of epitopes The primary interaction with antigen – The nature of B-cell epitopes,
Antigen and Antibody interaction, The specificity of antigen recognition, Binding of the peptides on the MHC, Super antigens, Tools and servers available

Vaccine design
Categories of vaccines. Polytope vaccines, Therapeutic vaccines. Evolution and escape due to variations, Immunogenomics and viral bioinformatics

Parasitic Diseases to be covered
Malaria, Leishmaniasis, Trypanosoma, Entamoeba histolytica, Toxoplasma, Filariasis

Parasites and Vectors of parasites
Biology of Parasites – Life Cycle, Infectivity, Demographic distribution of strains
Parasite Genome and Proteome Databases Vectors of parasites – Biology of vectors; Genome & Proteome databases: Mosquito

Application of Bioinformatics Data Mining tools for Identification of
Parasite-specific genes/gene products (e.g. house-keeping genes, genes essential for survival), Resistant Genes, Structural genomics of parasites, Host-Parasite and Host-Vector-Parasite Interactions –Pathway databases

Multi-Drug Resistance
Mechanism of MDR: genomic, molecular, cellular Identification of genes responsible for MDR
Approaches to novel drug discovery, Challenges and opportunities in vaccine development
Plant Parasites and diseases – Disease resistance genes of plants; Plant-pathogen interactions

Unit VI- Programming in C

Introduction: Types of Programming Languages, Introduction To C, Historical Development of C Language, Structure of C Program.

C Fundamentals: The C Character Set, Identifiers and Keywords, Data Types, Constants, Operators Used In C, Variables And Types Of C Variables, Declaration, Expressions, Statements, Symbolic Constants, I/O Statements used in C

Control Statements Introduction. **Branching:** The If Statements, If-Else Statements, If-Else, Ladder, Switch Statements. **Looping:** While Loop, Do-While Loop, For Loop, Multiple, Initialization In, the For Loop, Nested Control Structures, Jumps in Loop, Break Statement, Continue Statements, Goto Statement

Arrays: What Is An Array? , Declaring and Initializing an Array, One-Dimensional Array, Multi-Dimensional Array, Passing array to function, Strings.

Storage Classes: Automatic Storage Class, Register Storage Class, Static Storage Class, External Storage Class.

User Defined Functions: Introduction, Need for User-Defined Functions, Return Types, Passing. Arguments to a Functions, Scope and Life Time of Variables in Function, Nesting Of Functions, Recursion, Functions and Array.

Character String: What Are Strings, Declaring And Initializing String Variables, Reading And Writing Of String, Standard Library String Function

Pointers Introduction to Pointers, Declaring And Initializing Pointers, Accessing A

Variable Through Its Pointer , Pointer Expression , Pointers And Arrays, Pointers And Character Strings , Array of Pointer and pointer to pointer

Structures And Union: Introduction ,Structure ,Structure Initialization , Array Of Structures , Arrays Within Structure ,Structure Within Structure ,Introduction To Union

File Management in C: Introduction, Defining And Opening A File, Closing A File, Input/ Output Operations On File, Error Handling During I/O Operations, Random Access To File.

Introduction to OOPS: Object oriented programming, basic concepts of OOPS, Benefits of OOPS.

Unit VII- Programming in Perl

Introduction: What is Perl? Why use Perl in Bioinformatics? History of Perl, Availability, Support, Basic Concepts
Scalar Data: What Is Scalar Data? Numbers, Strings, Scalar Operators, Scalar Variables, Scalar Operators and Functions
Arrays and List Data: What Is a List or Array? Literal Representation, Variables, Array Operators and Functions, Scalar and List Context
Control Structures: Statement Blocks
Hashes: What Is a Hash? Hash Variables, Literal Representation of a Hash, Hash Functions, Hash Slices. **BasicI/O**
Regular Expressions: Concepts About Regular Expressions, Simple Uses of Regular Expressions, Patterns, More on the Matching Operator, Substitutions, The split and join Functions
Subroutines: System and User Functions, The local Operator, Variable-length Parameter Lists, Notes on Lexical Variables
Miscellaneous Control Structures, File handles and File Tests: What Is a File handle? Opening and Closing a File handle, Using Pathnames and Filenames, A Slight Diversion: die, Using File handles, The -x File Tests, The stat Function
Formats: What Is a Format? Defining a Format, Invoking a Format
Directory Access: Moving Around the Directory Tree, Globbing, Directory Handles, Opening and Closing a Directory Handle, Reading a Directory Handle File and Directory Manipulation
Process Management: Using system and exec, Using Backquotes,
Other Data Transformation: Finding a Substring, Extracting and Replacing a Substring
Formatting Data: Sorting, Transliteration
System Information: Getting User and Machine Information, Packing and Unpacking Binary Data Getting Network Information
Database Manipulation: DBM Databases and DBM Hashes, Opening and Closing, DBM Hashes, Fixed-Length Random-Access Databases, Variable-Length (Text), Databases, Win32 Database Interfaces
CGI Programming: The CGI.pm Module, Your CGI Program in Context, Simplest CGI, Program, Passing Parameters via CGI, Perl and the Web

Object oriented perl: Introduction to modules, Creating Objects
Bioperl: Introduction, Installation procedures, Architecture, Uses of bioperl

Unit VIII –Taxonomy and Phylogeny

Introduction
Systematics, Taxonomy and Phylogeny: Basic concepts – Species concept; kingdom to species; the five kingdoms; classical, phenetic and cladistic approaches; taxonomic information on viruses, microbes, plants & animals; Concepts in Molecular Evolution Nature of data used in Taxonomy and Phylogeny – Morphological and molecular character data

Phylogenetic trees and their comparison
Definition and description Various types of trees Consensus (strict, semi-strict, Adams, majority rule, Nelson). Data partitioning and combination. Tree to tree distances, similarity

Phylogenetic analysis algorithms
Maximum Parsimony Distance-based – UPGMA, Transformed Distance, Neighbors-Relation, Neighbor-Joining Probabilistic models and associated algorithms – Probabilistic models of evolution; Maximum likelihood algorithm

Approaches for tree reconstruction
Character optimization Delayed and accelerated transformation. Reliability of trees. Bootstrap Jackknife Decay Randomization tests.

Applications of phylogeny analyses
Comparison of Phylogenetic Trees obtained using DNA seq. vs. protein seq. vs. Full genomes. Need for addition of other properties towards total phylogenetic analysis Comparative methods for detection of species / organism relationships Gene duplication, Horizontal transfer, Domain evolution Study of co-evolution: Plant-insect interactions. Host-parasite interactions. Viral evolution.

Unit IX - Comparative Genomics and Proteomics

Objective and Overview of Genome Comparisons

GenomeAlignments BLAST2 MUMmer PipMaker VISTA

Comparison of Gene Order
GeneOrder

Comparative Genomics
Viruses, Microbes, Pathogens, Eukaryotes

Comparative Genomics Databases
COG, VirGen CORG, HOBACGEN, Homophila, XREFdb, Gramene

Single Nucleotide Polymorphism, dbSNP and other SNP-related databases

Overview of Proteomics
Experimental Techniques, Bioinformatics Approaches

Protein-Protein Interaction Networks, databases and software
DIP (Database of Interacting Proteins), PPI Server, BIND - Biomolecular Interaction Network Database, PIM –Hybrigenics, PathCalling Yeast Interaction Database, MINT - a Molecular Interactions Database. GRID - The General Repository for Interaction Datasets
InterPreTS - protein interaction prediction through tertiary structure

Unit X- Structural Biology & Molecular Modeling

Macromolecular Structure
Protein - Primary, Secondary, Super secondary, Tertiary and Quaternary structure
Nucleic acid – DNA and RNA, Carbohydrates, 3D Viral structures

Proteins
Principles of protein folding and methods to study protein folding
Structure of Ribosome, Macromolecular interactions: Protein – Protein, Protein – Nucleic acids, Protein - carbohydrates

Macromolecular x-ray crystallography
Principles of crystallography, Methods to study 3D structure, Co-ordinate systems
Fitting and refinement, Validation, Analysis of 3D structures
Mass spectrometry and computational approaches in structural biology

Molecular modeling
An Overview, Introduction and challenges, Molecular modeling methods – Conformational searching, Potential energy, maps, Ramachandran maps, Ab-initio methods, Semi-empirical methods, Empirical methods.

Conformational analysis
Introduction and Methods – Molecular fitting, Conformations: global vs. local
Force fields: expressions for stretch, bond, torsion, etc.
Energy Minimization – Non-derivative and derivative methods, Free energy calculations
Global optimization (simulated annealing, Tabu search, genetic algorithms)
Applications of energy minimization

Methods for 3D structure prediction
Knowledge based & Fold recognition Advance techniques in Prediction of 3D Structure –
Hidden Markov Model, Neural networks, Rossetta Stone, Genetic algorithms
Designing of molecules like drug, inhibitors using – Structure based & ligand based docking methods. Different Scoring schemes