Annexure – A

Ph.D. Entrance Test
College of Biotechnology

Bioinformatics

UNIT I

Bioinformatics: introduction, origin and history, Applications of Bioinformatics in agriculture, Analysis of protein and nucleotide sequences, Basic concepts of sequence similarity, Introduction to sequences alignments: local and global alignment, pairwise and multiple sequence alignment, Dynamic programming-Needleman and Wunsch, Smith-Waterman algorithms, Scoring matrices: PAM & BLOSUM, Motifs and Patterns, Phylogenetic analyses, Types of phylogenetic trees, Tree-Building Methods, Character-based and Distance-based methods, Maximum-likelihood, Maximum parsimony, Unweighted Pair Group Method with Arithmetic Mean (UPGMA), Neighbor-Joining (NJ), Fitch-Margoliash (FM), Minimum Evolution (ME), Tree Evaluation, Bootstrapping, Introduction to systems and network biology. Concepts of Molecular Modelling, Molecular mechanics, Force Fields, Local and global energy minima, Techniques in MD and Monte Carlo Simulation for conformational analysis, DFT and semi-empirical methods, Simulated annealing, RNA Secondary Structure prediction techniques, Algorithmic perspective of RNA folding, Protein structure prediction: *ab initio*, homology modeling and fold recognition methods, Receptor-based and ligand-based drug design, Design of ligands, docking, Classical SAR/QSAR, COMFA & COMSIA, Molecular descriptors, Pharmacophore mapping and applications

UNIT II

Fundamentals of Computing, Introduction to Operating Systems, WINDOWS, UNIX/Linux operating systems, Programming in C, Arrays and Pointers, Variables, Procedures and functions, Standard Controls, Flow control and loops, string operators, Declaration and definition of user defined functions, Call by value, Call by references, File handling in C, Overview of Object Oriented Programming, Class and Objects, function and operator overloading, inheritance. Definition, purpose of database system, Advantages of Database System, Components of Database System, Data Models-Relational, Network, Hierarchical, Three level Architecture for Database System ; internal, conceptual and external levels, Data independence, Data Abstraction, Mapping, Data Definition Language, Data Manipulation Language, Role of Schemas, Client/Server architecture, Relational Databases- Relational data models (binary, ternary, quaternary & n-ary relations), Important terms in relational database system, Primary and secondary keys, Structured Query Language, Primary, secondary and derived biological databases, submitting sequence to the Database and retrieval, Data mining & Knowledge discovery in Biological databases, Supervised and unsupervised learning, machine learning techniques, Artificial Neural Network, Support Vector Machine and genetic algorithms

UNIT III

Introduction to Statistical Bioinformatics, Random sampling. Frequency distributions: Graphical representations and Descriptive measures; Standard Probability Distributions; Correlation and regression analysis. Hypothesis testing; Markov Models, Cluster Analysis: Hierarchical and Non-Hierarchical methods. Coordinate geometry with basic concepts of 2D and 3D geometry, Vector algebra Addition and subtraction of vectors, Dot and cross product, Scalar triple product. Matrix algebra: basic definitions, matrix operations, transpose of a matrix, inverse of matrix, eigen values, Boolean algebra, Geometric and Arithmetic Progression. Solution of equation by bisection method, Iteration method, Newton Raphson method, numerical differentiation. Numerical integration; Trapezoidal rule, Simpson's 1/3 and 3/8 rules, Runge-Kutta method of nth order, Fast Fourier transformation

UNIT IV

Whole genome analysis and comparative genomics, classical ways of genome analysis, large fragment genomic libraries, Physical mapping of genomes, Genome sequencing strategies, Sequence assembly and annotation, exome sequencing, Genome structural and functional annotation, Functional genomics and proteomics, Metabolomics for elucidating metabolic pathways, Linkage analysis, genotyping analysis, Applications of genomics and proteomics in agriculture, Evolution of sequencing technology, Microarrays analysis and applications, Next generation sequencing technologies, Sequencing by synthesis, ligation, single molecular sequencing, emerging NGS technologies