

## **Elementary Bioinformatics**

### **Unit I**

Introduction to bioinformatics and data generation: What is bioinformatics and its relation with molecular biology. Examples of related tools (FASTA, BLAST, BLAT, RASMOL), databases (GENBANK, Pubmed, PDB) and software (RASMOL, Ligand Explorer). Applications of Bioinformatics.

### **Unit II**

Biological Database and its Types Introduction to data types and Source. Population and sample, Classification and Presentation of Data. Quality of data, private and public data sources. General Introduction of Biological Databases; Nucleic acid databases (NCBI, DDBJ, and EMBL). Protein databases (Primary, Composite, and Secondary). Specialized Genome databases: (SGD, TIGR, and ACeDB). Structure databases (CATH, SCOP, and PDBsum)

### **Unit III**

Data storage and retrieval and Interoperability: Flat files, relational, object oriented databases and controlled vocabularies. File Format (Genbank, DDBJ, FASTA, PDB, SwissProt). Introduction to Metadata and search; Indices, Boolean, Fuzzy, Neighboring search. The challenges of data exchange and integration. Ontologies, interchange languages and standardization efforts. General Introduction to XML, UMLS, CORBA, PYTHON and OMG/LIFESCIENCE.

### **Unit IV**

Sequence Alignments and Visualization: Introduction to Sequences, alignments and Dynamic Programming; Local alignment and Global alignment (algorithm and example), Pairwise alignment (BLAST and FASTA Algorithm) and multiple sequence alignment (Clustal W algorithm). Methods for presenting large quantities of biological data: sequence viewers (Artemis, SeqVISTA), 3D structure viewers (Rasmol, SPDBv, Chime, Cn3D, PyMol), Anatomical visualization.

### **Unit V**

Gene Expression and Representation of patterns and relationship: General introduction to Gene expression in prokaryotes and eukaryotes, transcription factors binding sites. SNP, EST, STS. Introduction to Regular Expression, Hierarchies, and Graphical models (including Markov chain and Bayes notes). Genetic variability and connections to clinical data.