

# **AMB17 BIOINFORMATICS**

## **UNIT-1 THE INTERNET AND BIOLOGIST**

Internet basics, FTP, Gopher, World Wide Web.

## **UNIT-2 THE GEN BANK SEQUENCE DATABASE**

Introduction, Primary & Secondary database, Format vs content: computer vs humans, GenBank Flat File dissection, GCG, ACDEB.

## **UNIT-3 STRUCTURE DATABASES**

Introduction to structures, PDB, MMDB, Structure file formats, Visualizing structural information, Database structure viewers.

## **UNIT-4 INFORMATION RETRIEVAL FROM BIOLOGICAL DATABASES**

Retrieving database entries, integrated information retrieval: The entrez system, sequence databases beyond NCBI, Medical Databases

## **UNIT-5 THE NCBI DATABASE**

Introduction, SeqIDS, Bioseq: Sequences, Bioseqsets: Collections of sequences, Seq. Annot: Annotating the sequence, Seqdiscr: Describing the sequence

## **UNIT-6 SEQUENCE ALIGNMENT AND DATABASE SEARCHING**

Introduction, Evolutionary basis of sequence alignment, Optimal alignment methods, Substitution scores & gap penalties, Statistical significance of alignments, Database similarity searching, FASTA, BLAST, Low complexity regions, Repetitive elements

## **UNIT-7 MULTIPLE SEQUENCE ALIGNMENT**

Progressive alignment methods, Motifs and patterns, Hocks, MOST, Probe, Presentation methods, Abscript

## **UNIT-8 PHYLOGENETIC ANALYSIS**

Elements of phylogenetic models, data analysis: Alignment, substitution model building, tree building and tree evaluation, building methods, searching for trees, hooting trees, Evaluating trees and data, phylogenetic software Some simple practical consideration

## **UNIT-9 PREDICTIVE METHODS USING NUCLEOTIDE SEQUENCE**

Framework, marking repetitive DNA, Database search, Codon bias detection, Detecting function sites in the DM, Integrated gene passing, finding tRMA genes

## **UNIT-10 PREDICTIVE METHODS USING PROTEIN SEQUENCES**

Protein identity based on composition, Propsearch, Physical properties based on sequences, secondary structure and folding classes, Sspread sopma, specialized structures of features, Tertiary structure

## **UNIT-11 GENOME MAPPING**

Different types of maps: physical, genetical, etc. Synteny, Human genome project, Application of genome mapping, Chromosome maps.

## **UNIT-12 SUBMITTING DNA SEQUENCES TO THE DATABASES**

Introduction, Where to submit, What to submit, How to submit on the world wide web, How to submit with sequin.

### **Reference Books**

1. Bioinformatics: A practical guide to the analysis of genes and proteins A.D. Baxevanis and B.F.F. Ouellette (Eds). 2002 John Wiley and Sons.
2. Bioinformatics: Sequence and Genome Analysis by D.W. Mount, 2001, Cold Spring Harbor Laboratory Press.

