| MST 304: BIOINFORMATICS | |
|---|---|
| **Teaching Scheme**<br>Lectures: 3 hrs/Week<br>Tutorials: 1 hr/Week<br>Credits: 4 | **Examination Scheme**<br>Class Test - 12Marks<br>Teachers Assessment - 6Marks<br>Attendance – 12 Marks<br>End Semester Exam – 70 marks |

Prerequisite: - Computer fundamentals, Computer Applications & Biostatistics, Concepts on biomolecules and function, Molecular Biology.

**Course Objectives:**

1. To give an overview on computing methods and the bioinformatics tools commonly used for analyzing the sequencing data.
2. To provide basics knowledge on unix and the fundamentals in networking.
3. To describe the importance of phylogenetic analysis and the mathematical models as a prerequisite to calculate the evolutionary linkages.
4. To explain the computing models and concepts to understand the computational techniques
5. To explain the annotation to the study proteins, protein coding genes and DNA and genomes.
6. To understand the structure prediction methods for the proteins and nucleic acids.

**Detailed Syllabus**

**Unit-1**

**Introduction to computers and bioinformatics-** Types of operating systems, concepts of networking and remote login, basic fundamentals of working with unix. **Biological databases-** Introduction to NCBI, NCBI data bases, BLAST, BLASTn, BLASTp, PSI-BLAST,modes of database search, mode of data storage (Flat file format, db-tables), flatfile formats of GenBank, EMBL, DDBJ, PDB. Sequence alignment –Concept of local and global sequence alignment, Pairwise sequence alignment, scoring analignment, substitution matrices, multiple sequence alignment, Primer designing.

**Unit-2**

**Phylogenetic analysis-** Basic concepts of phylogenetic analysis, rooted/uprooted trees, approaches for phylogenetictree construction (UPGMA, Neighbor joining, Maximum parsimony, Maximum likelihood). Cluster analysis; Phylogenetic clustering by simple matching coefficients; Sequence Comparison; Sequence pattern; Regular expression based pattern; Theory of profiles and their use in sequence analysis; Markov models; Concept of HMMS; Baum-Welch algorithm; Use of profile HMM for protein family classification; Pattern recognition methods.

**Unit-3**

**Generation and analysis of high throughput sequence data-** Assembly pipeline for clustering of HTGS data, format of ".ace" file, quality assessment of genomic assemblies, International norms for sequence data quality, Clustering of EST sequences, concept of Unigene.

**Annotation procedures for high through-put sequence data-** Identification of various genomic elements (protein coding genes, repeat elements, strategies for annotation of whole genome, functional annotation of EST clusters, gene ontology (GO) consortium

Head
Department of Biotechnology
Invertis University, Bareilly (U.P.)

Dean
Faculty of Science
Invertis University, Bareilly (U.P.)

Registrar
Invertis University
Bareilly

**Unit-4**

**Structure predictions for nucleic acids and proteins**- Approaches for the prediction of RNA secondary and tertiary predictions, energy minimization and base covariance models, Basic approaches for protein structure predictions, comparative modeling, fold recognition/threading and ab-initio prediction.

**Text and Reference Books**

1. *Bioinformatics: A Practical Guide to the Analysis of Genes and Proteins by Baxevanis A.D. and Ouellette, Third Edition. John Wiley and Son Inc., 2005.*
2. *Bioinformatics Sequence and Genome Analysis by Mount D.W., CSHL Press, 2004.*
3. *Introduction to Bioinformatics by Tramontano A., Chapman & Hall/CRC, 2007.*
4. *Understanding Bioinformatics by Zvelebil, M. and Baum, Chapman & Hall/CRC, 2008.*

**Course Outcomes:**

After completing the course, students will be able to:

1. Understand the importance of bioinformatics and the computational techniques.
2. Analyze the sequencing data generated and available in the databases and to interpret these Results.
3. Identify the important mathematical models and techniques for biological data analysis
4. Understand importance of techniques for structure and function prediction of proteins and genes.
5. Understand the nucleic acid and protein structure prediction tools
6. Understand the genome annotation methods and some of the techniques.

Head
Department of Biotechnology
Invertis University, Bareilly (U.P.)

Dean
Faculty of Science
Invertis University, Bareilly (U.P.)

Registrar
Invertis University
Bareilly