

Course Code: ZOO-434

Course Title: Bioinformatics: Introduction to Biological Databases

Number of Credits: 2

Effective from AY: 2020 -21

Prerequisite for the Course:	Basic knowledge of Cell and Molecular biology, Genetics.	
Objectives:	<ol style="list-style-type: none">1. To expose students to this subject and modern technology.2. To make students understand fundamental principles of bio informatics for an in depth understanding of the related subjects.3. Application of this novel field in Zoology and use of this subject with respect to Evolutionary significance.	
Content:	<p>Module1: Bioinformatics- Introduction and definition, History and Scope, Applications of Bioinformatics in various fields.</p> <p>Nucleic Acid Sequence Databases :</p> <ul style="list-style-type: none">• Nucleic acid sequence databases (GenBank, EMBL, DDBJ), Keyword-based search at Entrez Search Engine at NCBI.• Sequence Submission tools at NCBI, EMBL etc. <p>Protein sequence database: UniProtKB (SwissPort, TrEMBL). Species specific database for mouse and zebra fish Metabolic pathway databases.</p> <p>Open Access Bibliographic Resources and Literature Databases:</p> <ul style="list-style-type: none">• PubMed,• MEDLINE,• PubMedCentral at NCBI <p>Module 2 Sequence Analysis: Various File Formats for Biomolecular Sequences:</p> <ul style="list-style-type: none">• GenBank• FASTA <p>Basic concepts of sequence analysis:</p> <ul style="list-style-type: none">• Global Pair wise Sequence Alignment• Local Pair wise Sequence Alignment• Needleman and Wunsch, Smith and Waterman algorithms for pair wise alignments, gap penalties, use of pair wise alignments for analysis of Nucleic acid and Protein sequences and interpretation of results. <p>Databases Searches :</p> <ul style="list-style-type: none">• BLAST• FASTA <p>Multiple Sequence Alignment:</p>	<p>1 hr</p> <p>8 hrs</p> <p>3 hrs</p> <p>7 hrs</p> <p>5 hrs</p>

	<ul style="list-style-type: none"> • The need for MSA • Basic concepts of various approaches for MSA (e.g. progressive, hierarchical, iterative etc.). <p>Concept of Phylogeny:</p> <ul style="list-style-type: none"> • Molecular Phylogeny • Various Methods of Phylogenetic Tree Construction <p>Scoring matrices: Basic concept of a scoring matrix, Matrices for nucleic acid and proteins sequences, PAM and BLOSSUM series.</p>	
Pedagogy:	Lectures/ tutorials/online teaching mode/self-study.	
Learning Outcome:	<p>The students will acquire the knowledge about:</p> <ol style="list-style-type: none"> 1. Various bioinformatics tools and techniques and how to use that for the analysis of the biological experimental data. 2. Concepts of various databases and various methods for the data retrieval, data storage, and data mining and use that data for the further analysis. 3. In-Silico approach for the protein modeling and drug discovery process. 4. Sequencing techniques and gene annotation as well as submission of the sequences to the various databases. 	
References /Reading:	<ol style="list-style-type: none"> 1. Mount D (2005), Bioinformatics Sequence and Genome Analysis. 2. Xiong J (2006), Essential Bioinformatics: Cambridge University Press. 3. Wood AH, Parry TK and Smith DJ (2001), Introduction to bioinformatics, Pearson education Asia. 4. Baxevanis AD & Ouellette BFF (2001), Bioinformatics: A practical guide to the analysis of genes and proteins, Wiley Interscience – New York. 5. Misener S & Stephen A (2000), Bioinformatics: Methods and Protocols Krawetz, Humana Press, New Jersey. 6. Higgins D & Taylor W (2000), Bioinformatics, Sequence, structure and databanks – Oxford University Press. 7. Bosu O and Thukral SK (2007), Bioinformatics Databases, Tools and Algorithms. 	