

Dr. V. S. KRISHNA GOVERNMENT DEGREE AND PG COLLEGE (A), VSKP
DEPARTMENT OF BIOTECHNOLOGY
B.Sc BIOTECHNOLOGY SYLLABUS
SEMESTER V

Course VI –BIOINFORMATICS AND BIOSTATISTICS

(Total Hours of Teaching 60 @ 04 hrs/week)

Credits - 04

Course Outcomes:

On successful completion of this course, the students will have the knowledge and skill to:

- To become familiar with a variety of currently available genomic and proteomic databases.
- To be able to search and retrieve information from genomic and proteomic databases (e.g. GenBank, Swiss-Prot), and to analyze their search results using software available on the internet (e.g. BLAST, ClustalW).
- To compare and analyze biological sequences and how to interpret the results of their analyses.

MODULE I: Introduction to Bioinformatics

12

Hours

- 1.1. What is Bioinformatics and its relation with molecular biology Examples of related tools (FASTA, BLAST, RASMOL)
- 1.2. Examples of databases (GENBANK, Pubmed, PDB) and Examples of software (RASMOL, Ligand Explorer),
- 1.3. Data generation; Generation of large scale molecular biology data. (Through Genome sequencing, Protein sequencing, Gel electrophoresis)
- 1.4. Applications of Bioinformatics.

MODULE II: Biological Database

12

Hours

- 2.1. Introduction to data types and Source. Population and sample
- 2.2. Biological Database and its Types.
- 2.3. Classification and Presentation of Data. Quality of data, private and public data sources.
- 2.4. 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).

MODULE III: Data Storage and Retrieval

12

Hours

- 3.1 Data storage and retrieval and Interoperability
- 3.2 Flat files, relational, object oriented databases and controlled vocabularies
- 3.3 File Format (Genbank, DDBJ, FASTA, PDB, SwissProt).
- 3.4 Introduction to Metadata and search; Indices, Boolean, Fuzzy, Neighboring search.

MODULE IV: Biostatistics

12 Hours

- 5.1. Measures of central tendency: Mean, median, mode, merits, and demerits.
- 5.2. Measures of dispersion: Range, Variance, Standard deviation.
- 5.3. Tabulation of data and graphical representation
- 5.4. ANOVA, Student's t-test

MODULE V: Probability and distribution

12

Hours

- 5.1. Concepts of Probability- Additive Law and Multiplicative Law
- 5.2. Conditional probability, Dependent and Independent events
- 5.3. Normal distribution
- 5.4. Binomial distribution and Poisson distribution

PRACTICALS SEMESTER - V

Course VI– BIOINFORMATICS AND BIOSTATISTICS (Total hours of Laboratory Exercises – 30 hours @ 03 hrs/week) Credits - 02

Course Outcomes:

On successful completion of this course, the students will have the knowledge and skill to:

- To locate consensus sequences, genes, and open reading frames within biological sequences.
- To become familiar with the principles and applications of microarrays.
- To be able to perform elementary predictions of protein structure and function.
- To be able to perform elementary comparative genomic analysis.
- To use the scientific method of inquiry, through the acquisition of scientific knowledge.
- To perform various statistical tests and problem solving.

Practical Syllabus

1. NCBI, DDBJ, and EMBL.
2. Specialized Genome databases: (SGD, TIGR, and ACeDB).
3. Protein databases (Primary, Composite, and Secondary).
4. Structure databases (CATH, SCOP, and PDBsum).
5. Genbank, DDBJ, FASTA, PDB, SwissProt.
6. BLAST,FASTA.
7. Artemis, SeqVISTA
8. Rasmol, SPDBv, Chime, Cn3D, PyMol.
9. Problems on mean, median, mode, probability and standard deviation
10. Problems based on ANOVA
11. Problems based on t Test