COURSE TITLE	**BIOLOGICAL DATABASES**
COURSE CODE	**01CB0603**
COURSE CREDITS	**3**

Objective:

1 Develop basic knowledge of the available online biological databases.

2 Experiment with all kinds of nucleotide and protein databases and make the best use of them.

Course Outcomes: After completion of this course, student will be able to:

1 Build an extensive knowledge of model organisms and browse genome databases to retrieve useful information that will be helpful for their research work.

2 Distinguish the intersection of life and information sciences, information theory, gene expression, and database queries

3 Apply existing software effectively to extract information from large databases and use this information in computer modeling.

4 Analyze nucleotide and protein sequences from various databases.

5 Evaluate sequence, structural, and functional analysis of biomolecule

Pre-requisite of course: Basic knowledge of Bioinformatics.

Teaching and Examination Scheme

Theory Hours	Tutorial Hours	Practical Hours	ESE	IA	CSE	Viva	Term Work
2	0	2	50	30	20	25	25

Contents : Unit	Topics	Contact Hours
1	**Introduction** Relational database- Motivation of biological database, RDBMS, DBMS, File formats of different databases, Submission of Data to Databases	6
2	**Understanding Molecular Biology Databases** European Molecular Biology Laboratory (EMBL) - NCBI GenBank – DNA Data Bank of Japan (DDBJ), Genes and genetic disorders : COSMIC, Clinvar, HUMSAVAR- SNP database (DbSNP).	5
3	**Protein Databases** UniProt Knowledgebase: SwissProt/TrEMBL - Protein Information Resource (PIR), PRIDE Database	5
4	**History of structural biology** Protein Data Bank (PDB), contents of a PDB file- SCOP: SCOP: Structural Classification of Proteins - CATH : Protein Structure Classification database	5

DR. SEJAL SAHILKRISHNA SHAH
Digitally signed by (Name of HOD)

DR. RAJENDRASINH BAHADURSINH JADEJA
Digitally signed by (Name of Dean/ Principal)

Contents : Unit	Topics	Contact Hours
5	**Protein family databases** Pfam-protein family database (pfam) - GO-gene ontology, PROSITE-protein function pattern and profile, ENZYME-Enzyme commission, KEGG Pathway database	7
Total Hours		**28**

Suggested List of Experiments:

Contents : Unit	Topics	Contact Hours
1	**Module1:** To compare different types of biological database input formats	2
2	**Module1:** To compare methods for submitting sequences to databases like EMBL, GenBank, and DDBJ	2
3	**Module2:** To identify nucleotide sequence and convert sequences from one format to another using bioinformatics tools	2
4	**Module2:** To analyze genetic variants and disorders in sequences	2
5	**Module3:** To compare genes from UniProt Knowledgebase, including SwissProt and TrEMBL	2
6	**Module3:** To perform retrieval of protein-related data and annotations from databases	2
7	**Module3:** To analyze Protein Data Bank (PDB) contents for five classes of proteins	2
8	**Module3:** To analyze proteins from SCOP (Structural Classification of Proteins)	2
9	**Module4:** To compare proteins from CATH (Protein Structure Classification database)	2
10	**Module4:** To perform comparison based on structure calculate RMSD	2
11	**Module4:** To compare information on protein families and domains using MEME	2
12	**Module4:** To Analyse PROSITE for protein function pattern and profile information	2
13	**Module5:** To analyze ENZYME for Enzyme Commission information	2

Suggested List of Experiments:

Contents : Unit	Topics	Contact Hours
14	**Module5:** To compare and annotate proteins based on their functions	2
Total Hours		**28**

Textbook :

1 Bioinformatics: Sequence and Genome Analysis, Mount D, Cold Spring Harbor, 2014

References:

1 Introduction to bioinformatics, Introduction to bioinformatics, Attwood TK and Parry-Smith DJ, Pearson Education, 2014

2 Bioinformatics: a practical guide to the analysis of genes and proteins, Bioinformatics: a practical guide to the analysis of genes and proteins, Baxevanis A., Ouellette F.B.F., John Wiley and Sons, New York, 1998

3 Introduction to Biological Databases, Essential Bioinformatics, Introduction to Biological Databases, Essential Bioinformatics, Jin Xiong, Cambridge University Press, 2012

4 Bioinformatics Database Systems, Bioinformatics Database Systems, Kevin Byron, Katherine G. Herbert, Jason T. L. Wang, CRC Press, 2020

Suggested Theory Distribution:

The suggested theory distribution as per Bloom's taxonomy is as follows. This distribution serves as guidelines for teachers and students to achieve effective teaching-learning process

Distribution of Theory for course delivery

Remember / Knowledge	Understand	Apply	Analyze	Evaluate	Higher order Thinking / Creative
20.00	30.00	20.00	10.00	10.00	10.00

Instructional Method:

1 The course delivery method will depend upon the requirement of content and need of students. The teacher in addition to conventional teaching method by black board, may also use any of tools such as demonstration, role play, Quiz, brainstorming, MOOCs etc.

2 The internal evaluation will be done on the basis of continuous evaluation of students in the laboratory and class-room.

3 Practical examination will be conducted at the end of semester for evaluation of performance of students in laboratory.

Supplementary Resources:

1 https://microbenotes.com/biological-databases-types-and-importance/

2 https://training.galaxyproject.org/videos/watch.html?v=data-science/tutorials/online-resources-gene/slides

DR. SEJAL SAHILKRISHNA SHAH
Digitally signed by (Name of HOD)

DR. RAJENDRASINH BAHADURSINH JADEJA
Digitally signed by (Name of Dean/ Principal)