



**FORMAT FOR COURSE CURRICULUM**

**Course Title: Introduction to Computational Biology**

**Course Code: BIOF401**

**Units: 04**

L	T	P/ S	SW/F W	TOTAL CREDIT UNITS
3	0	2	0	4

**Course Objectives:**

Theory: This course will enable the students to achieve the understanding in integrating biology and different sciences together under the umbrella term Computational biology.

**Pre-requisites:** 10+2 (PCM/B), Basics of Computers.

**Student Learning Outcomes:**

- The student will gain knowledge about biological data submission and retrieval from databases
- The student will acquire skills to analyze biological data and produce and interpret the predictions of the software.

**Course Contents/Syllabus- Theory:**

	Weightage (%)
<b>Module I</b>	<b>15</b>
<b>Descriptors/Topics: Introduction and overview</b> String view of DNA and Proteins: Basic file formats, codon-genetic code-transcription & translation in parallel to computational biology, sense/coding as anti-sense/template strands. Sequence Data Bases and their types, detailed study of GenBank of NCBI- typical Gen Bank (DDBJ+EMBL) entry and for DNA and RNA, concepts of similarity - homologous, orthologous and paralogous sequences.	
<b>Module II</b>	<b>25</b>
<b>Descriptors/Topics: Sequence alignment</b> Global and local alignments, statistical significance of alignments, Pair-wise sequence alignment, Scoring matrices for amino acid sequence alignment, PAM, BLOSUM, Needleman-Wunch algorithm, position specific scoring matrices, BLAST, FASTA, Smith – Waterman algorithm for local alignment; Multiple sequence alignment- Progressive,	

Iterative and Block based MSA. Applications of Alignments-Patterns, Profiles, PSI-BLAST.	
<b>Module III</b>	<b>20</b>
<p><b>Descriptors/Topics: Molecular Phylogeny</b></p> <p>Concept of phylogenetics –Application of Phylogenetic trees- Basic terminology-taxa, taxonomy, clade, root, leaf, node graph &amp; tree, Dendrogram, cladogram, rooted tree, unrooted tree, scaled trees- Molecular clock hypothesis, Distance based methods-UPGMA, NJ algorithm, Character based methods-Maximum parsimony and ML methods. Newick format of trees. Validating phylogenetic trees – jack knifing and bootstrapping Tools for Phylogenetic analysis- MEGA6, PHYLIP etc.</p>	
<b>Module IV</b>	<b>15</b>
<p><b>Descriptors/Topics: Structural and Functional Annotation of DNA/Proteins.</b></p> <p>Computational Gene Prediction in Prokaryotes and Eukaryotes. Tools based on different strategies. , Promoter Prediction, Transcription factor binding sites prediction, Prediction of Restriction sites on DNA/Proteins, Prediction of vector contamination.</p>	
<b>Module V</b>	<b>25</b>
<p><b>Descriptors/Topics: Molecular modeling and Basics of Tools &amp; Databases in Bioinformatics</b></p> <p>Molecular modeling – PDB and MMDB, structure file formats, visualizing structural information, advance structure modeling, Internal and external co-ordinate system, cartesian and cylindrical polar co-ordinate system. Protein secondary and tertiary structure prediction: JPred, 3DPSSM, 123D, Modeller, Procheck, ITASSER; Protein visualization tools- Swiss PDB Viewer, Pymol, Rasmol.</p>	

**Pedagogy for Course Delivery:**

Lectures: 41

Tutorial:

Presentation/ Seminar: 3

Class Test: 1

Total: 45

## Practical

Practical	: 28
Lab internal	: 2
Total	: 30

## List of Experiments

- Biological databases: NCBI, EMBL, DDBJ, iHOP, PDB, UniProt, KEGG, Ensembl, STRING; Sequence file formats: GenBank, FASTA, EMBL, PDB format
- DotPlot Analysis: DOTPLOT, DOTTER, DOTMATCHER., Pairwise Sequence Alignment programs: LALIGN, EMBOSS NEEDLE, EMBOSS Water, Clustalw, Muscle, T-Coffee, Similarity Searching: BLAST, Variants of Blast.
- Phylogenetic analysis software: MEGA, PHYLIP.
- Primer Designing: PRIMER3, Gene Identification Programs: GENSCAN, ORF finder. Fgenesh, Glimmer, Protein Identification and characterization: Protparam, Peptide cutter, Motif and Patterns program: Prosite, InterProScan, Pfam.
- Modelling software: Swiss Model workspace, ArgusLab, Model Evaluation: PROcheck. Docking-Hex

### Assessment/ Examination Scheme:

<b>Theory L/T (%)</b>	<b>Lab/Practical/Studio (%)</b>	<b>Total</b>
<b>75%</b>	<b>25%</b>	<b>100%</b>

### Theory Assessment (L&T):

<b>Continuous Assessment/Internal Assessment</b>						<b>End Term Examination</b>
<b>Components (Drop down)</b>	<b>Class Test 1</b>	<b>Viva</b>	<b>Home Assignment</b>		<b>Attendance</b>	
<b>Weightage (%)</b>	15	5	10		5	70

### Lab/ Practical/ Studio Assessment:

	<b>Continuous Assessment/Internal Assessment</b>				<b>End Term Examination</b>			
<b>Components (Drop down)</b>	<b>Performance</b>	<b>Lab record</b>	<b>viva</b>	<b>Attendance</b>	<b>Lab record</b>	<b>Performance</b>	<b>Viva</b>	<b>Total</b>

<b>Weightage (%)</b>	10	10	5	5	10	50	10	70
----------------------	----	----	---	---	----	----	----	----

### **TEXT BOOKS**

- Bioinformatics: Sequence, Structure and Databanks: A Practical Approach (The Practical Approach Series, 236), Des Higgins (Editor), Willie Taylor. 1<sup>st</sup> edition, October 2000, Oxford University Press
- Bioinformatics: Sequence and Genome Analysis, David W. Mount. 2<sup>nd</sup> edition, June 2004, Cold spring harbor laboratory press.

### **REFERENCES**

- Introduction to Bioinformatics, Teresa Attwood, David Parry-Smith, 1st edition, May 2001, Pearson Education
- Bioinformatics: A Practical Guide to the Analysis of Genes and Proteins, Second Edition, Andreas D. Baxevanis, B. F. Francis Ouellette. 3rd edition, October 2004, A John Wiley & Sons, Inc., Publication
- Introduction to Bioinformatics, Arthur M. Lesk, 2002, Oxford University Press

### **Any other Study Material:**

- Online tutorials
- Molecular Evolution and Phylogenetics Masatoshi Nei and Sudhir Kumar