Course title: Bioinformatics and Computational Biology- Part I			
Course code: BBP174	No. of credits: 2	L-T-P: 20-8-0	Learning hours: 28
Pre-requisite course code and title (if any): None			
Department: Department of Biotechnology			
Course coordinator(s):		Course instructor(s): Prof Mukesh Jain / Dr. Rohini Garg	
Contact details:			
Course type: Core		Course offered in: Semester II	

Course description: This course is designed to introduce students of various academic backgrounds to the interdisciplinary knowledge that underlies Bioinformatics and Computational Biology. The students will get acquainted with fundamentals of computers, operating systems and basic concepts in computing and networking.
The students will be trained to use computational tools and approaches to extract information from different types of bioinformatics data (gene, protein, disease, etc.) and to analyse them in their area of research work.

Course objectives:
1. To educate the interdisciplinary nature of advances in bioinformatics and computational biology.
2. To provide basic understanding of how biological data is stored and retrieved from various biological databases.
3. To develop an understanding of algorithms of sequence alignment (pair-wise and multiple) and scoring algorithms.

Course contents

S.No	Topic	L	T	P
Module1	**Introduction to Information Technology**			
1	**Concepts in information processing and peripherals** Information technology, an overview of current IT applications, Difference between data and information, Information system, Programming languages, Input and output devices, port introduction.	2		0
2	**Internet technology and World Wide Web** Data compression, entropy of information, networking in computers, operating system, Intranet and internet, file transfer protocols, world wide web, internet requirements, Internet-a global network, host& terminal, TCP/IP, common protocols used in internet, web browsers, internet addresses, domain names, basic concepts of HTML, web search engines, electronic mail.	3	2	0
3	**Information search and data retrieval** Introduction, tools for web search, data retrieval tools, data mining of biological databases, biological databases, difference between primary, secondary and tertiary databases, types of databases. Probabilistic information retrieval, language models for information retrieval, managing bioinformatics tools, command line sequence extraction and analysis.	5	2	0
Module2	**Sequence Analysis and alignment**			
4	**Sequence analysis** Sequence analysis, various file formats for bio-molecular sequences: Genbank, FASTA, GCG, MSF, NBRF-PIR etc. Basic concepts of sequence similarity, identity and homology, definitions of homologues, orthologues, paralogues. Scoring matrices: basic concept of a scoring matrix, PAM and BLOSUM series. Applications of substitution matrices in protein sequence alignment and evolution. Sequence-based Database Searches: what are sequence-based database searches, BLAST and FASTA algorithms, various versions of basic BLAST and FASTA,	6	2	0

		filtering and gapped BLAST.			
5		**Pairwise and Multiple sequence alignments** Basic concepts of sequence alignment, Needleman & Wunsch, Smith & Waterman algorithms for pairwise alignments, Progressive and hierarchical algorithms for MSA. Use of pairwise alignments and Multiple sequence alignment for analysis of Nucleic acid and protein sequences and interpretation of results.	4	2	0
		Total	**20**	**8**	**0**

Evaluation criteria:
1. 2 minor tests   30% (each)
2. 1 major test (end semester)   40%

Learning outcomes:
1. On the completion of this course students shall have knowledge to identify, adapt and develop *in silico* models appropriate to the specific study of different biological projects.
2. The students will be familiar with the use of bioinformatics software, tools in their area of research.

Pedagogical Approach:
Classroom lectures and tutorials, with a major emphasis on the detailed discussion of uses of software in original research work in the class.

Skill Set:
1. Data analysis: intelligent bioinformatics methods.

Employability:
1. Bioinformatic software development companies.
2. Academic Institutions and Industries involving R&D.

Materials:
Required text
Suggested readings
1. David W Mount**,** Bioinformatics: Sequence And Genome Analysis, 2nd Edition, cold Spring Harbor Press
2. Durbin et al (2007) Biological Sequence Analysis: Probabilistic models of protein and Nucleic acids .Cambridge University Press.
3. Korf Ian, Yandell Mark, Bedell Joseph. BLAST: an essential guide to the basic local alignment search tool. Shroff Publishers and Distributors Pvt. Ltd., 2003. ISBN: 8173665125.
4. Teresa Attwood, Parry-Smith David J. Introduction to Bioinformatics. Publisher: Pearson Education (Singapore) Pte.Ltd. 2001. ISBN:8178085070.

Case studies
Websites: http://www.ncbi.nlm.nih.gov/
                  www.embl.org

                  www.ddbj.nig.ac.jp

                  www.uniprot.org

                  www.rcsb.org/
Journals
Other readings

Additional information (if any):

Student responsibilities:
1. Class attendance.
2. Study of course materials as specified by the instructor.

Reviewed by:

Dr. Dinesh Gupta, ICGEB, New Delhi

Prof. B.N. Mishra, Ins of Eng & Tech., Lucknow