



Approved in 44th BoA Meeting (24-11-2021)

Course number	: BE505
Course Name	: Computational Biology
Credit Distribution	: 2-0-2-3
Intended for	: Core for Integrated Dual Degree Bioengineering students with Specialisation in Computational Bioengineering, elective for other B.Tech and Mtech students.
Prerequisite	: Understanding Biotechnology and its applications (IC136), and Bioinformatics (BE304) or with permission of the instructor.
Mutual Exclusion	: None

1. Preamble: The course aims at providing an understanding to the students about advanced topics in Computational Biology and their in-depth applications for solving biological problems. There are two goals for this course. The first goal is to introduce the students to the foundations of the field of computational biology. Namely, introduce the fundamental biological problems of the field, and learn the algorithmic and machine learning techniques needed for tackling them. The second goal of the course is to tackle the research frontiers of computational biology. The course will include practical sessions for the students to help them master some of the advanced bioinformatics techniques from hands-on experience. The course will cover the following broad topics:

2. Course Modules with quantitative lecture hours:

A. Pattern matching: Finding regulatory sequences in DNA, Motif discovery, Exhaustive search, Greedy motif clustering, wordlets and motifs refinements, Probabilistic solutions (expectation maximisation, Gibbs sampling) (2 hours)

B. Genomic Analysis: (14 hours)

- Genome assembly with Graphs and Networks,
- Gene prediction, Hidden Markov models, Viterbi, expectation maximisation)
- Next Generation Sequencing analysis (short read mapping, ChIP-seq and RNA-seq analysis)
- Comparative genomics and genome rearrangements
- Population genomics, Medical genomics, Personal genomics, disease epigenomics, Systems approaches to disease

C. Phylogenetic Inferences: Introduction to phylogenetic inferences, Alignments to Distances, probabilistic models of divergence (Jukes Cantor, Kimura, hierarchy), Distances to trees, types of trees, algorithms for tree building (UPGMA, neighbor joining), optimality (least squared error, minimum evolution), Alignments to trees, alignment scoring given a tree, parsimony, greedy vs dynamic programming, maximum likelihood, Max-a-Posteriori, bootstrapping, Tree visualisation. (3 hours)

D. Structure prediction and Molecular modelling: RNA and protein structure prediction methods, homology modelling and ab – initio structure prediction methods, models of

proteins, discrete conformational search, binding and docking, molecular dynamics simulations. (5 hours)

F. Biological network analysis and modelling: Gene regulatory networks, Protein interaction networks, Logic modelling of Cell signalling networks, network modelling, formulating models, nonlinear dynamics and stability, steady-state problems, parameter fitting and estimation, basic overview of the modeling of metabolic networks in genome scale by Flux Balance or modeling of reaction kinetic for smaller networks/pathways (4 hours)

Lab Course content: (28 hours)

The below mentioned 10 topics will be covered over the 14 weeks:

1. Pattern matching
2. Genome assembly
3. Gene prediction (prokaryotic and eukaryotic)
4. Read mapping and NGS data-analysis
5. Comparative genomics methods
6. Population genomics methods
7. Phylogenetic and analysis molecular evolution
8. Protein structural analysis (Secondary structure prediction and homology modelling)
9. Protein structural analysis (Ab-initio molecular modelling and simulation, docking)
10. Biological networks analysis

3. Text books:

1. Computational Molecular Biology: An Introduction 1st Edition by Peter Clote (2000), ISBN-13: 978-0471872528, ISBN-10: 0471872520
2. An Introduction to Bioinformatics Algorithms (Computational Molecular Biology), by Neil C. Jones and Pavel A Pevzner, (2004), ISBN-10 : 0262101068, ISBN-13 : 978-0262101066.

4. References:

1. Bioinformatics Algorithms: Techniques and Applications: 03 (Wiley Series in Bioinformatics) by Ion Mandoiu (Author), Alexander Zelikovsky (2008), ISBN-10 : 0470097736, ISBN-13 : 978-0470097731
2. Introduction to Bioinformatics. Arthur M. Lesk (3rd Edition) Oxford University Press.
3. Biological Sequence Analysis-Probabilistic Models of Proteins and Nucleic Acids, by Richard Durbin, Sean R. Eddy, Anders Krogh, Graeme Mitchison. Cambridge University Press.
4. Bioinformatics and Functional Genomics 3rd Edition (Wiley-Blackwell) by Jonathan Pevsner. ISBN-13: 978-1118581780

5. Similarity Content declaration with existing courses:

S. No.	Course Code	Similarity Content	Approx. % of Content
01.	BY512	Gene Prediction, Phylogenetic analysis	10%

6. Justification of new course proposal if cumulative similarity content is >30%: