First Course Handout for the Course BSE633A

BSE633A: Bioinformatics and Computational Biology

Objective: Biological data in the form of protein, DNA and RNA sequences, their three-dimensional structures, gene expressions are expanding exponentially with the advent of new technologies. Hence, computational tools to organize and analyze the data to investigate basic biological problems have become integral part of modern biology. In this course, we will discuss algorithms to align two (pairwise) or more (multiple) sequences, scoring matrices and finding sequence patterns in protein and DNA. Algorithms describing the most popular tools such as BLAST and Clustal will be explained. Different types of phylogenetic algorithms will be discussed and the software packages like MEGA will be demonstrated. Other topics in this course include gene prediction algorithms and comparative genomics.

Instructor: Dr. R. Sankararamakrishnan

Course Email ID: bse633a@iitk.ac.in

Class Timings: 12:00 to 1:20 pm (Tuesdays and Fridays)

Venue: BSBE Seminar Hall

Course Website: http://home.iitk.ac.in/~rsankar/courses

Course Evaluation:

Quiz I (February First Week)	5%
Mid-semester Exam	30%
Quiz II (April First Week)	5%
Assigment/Exercise	10%
Presentation	5%
End-Semester Exam	40%
Attendance	5%
Total	100%

Course contents:

Introduction to bioinformatics: Distinction between bioinformatics and computational biology, biological databases and their growth, concept of homology and definition of associated terms

Pairwise sequence alignment: Dotmatrix plot, dynamic programming algorithm, global (Needleman-Wunsch) and local (Smith-Waterman) alignments, BLAST

Scoring matrices: PAM and BLOSUM families, gap penalty, statistical significance of alignment

Multiple sequence alignment: Sum-of-pairs method, CLUSTAL W, Genetic Algorithm, Pattern finding in protein and DNA sequencing, Gibbs Sampler, Hidden Markov Model, Profile construction and searching, PSI-BLAST

Phylogenetic algorithms: Introduction to phylogeny, maximum parsimony method, distance method (neighbor-joining), maximum-likelihood method

Gene Prediction and Comparative Genomics: Gene prediction in prokaryotes and eukaryotes, homology and ab-initio methods, genome analysis and annotation, comparative genomics

References:

1. Bioinformatics: Sequence and Genome Analysis, David W. Mount, Cold Spring Harbor Laboratory Press (2001)

2. Advances in Protein Chemistry, Vol 54 (2000)

3. Developing Bioinformatics Computer Skills. C. Gibas and P. Jambeck, O' Reilly (2001)

4. Biological Sequence Analysis: Probabilistic models of proteins and nucleic acids. R. Durbin, S. Eddy, A. Krogh and G. Mitchison, Cambridge University Press (1998)

5. Journals: Bioinformatics, BMC Bioinformatics, Nucleic Acid Research, ISMB, J. Comp. Biol., PLoS Computational Biology