

BCB716: Sequence analysis and comparison
Fall 2013
Syllabus

1 Overview

The comparison of DNA and protein sequences is a classic playground for computational and statistical methods in biology, there continue to be interesting and challenging problems in this domain, and sequence comparison algorithms such as BLAST are among the most widely-used bioinformatic tools. This module covers the fundamental concepts and methods in sequence comparison using computational techniques such as dynamic programming, hidden Markov models and Markov Chain Monte Carlo.

2 Instructors

- Todd Vision, Dept. of Biology, Coker Hall Rm 102, tjv@bio.unc.edu, 843-4507, office hours Tues 1-3.
- Zefeng Wang, Dept. of Pharmacology, Genetic Medicine Bldg Rm 4097, zefeng@med.unc.edu, 966-0131, office hours TBA

3 Prerequisite knowledge

- Molecular Genetics. The major structural features of genes and proteins and genomes, both prokaryotic and eukaryotic.
- Probability and Statistics. Gaussian, binomial, Poisson and Chi-squared distributions, Bayes theorem, tail probabilities, means and variances, and the distinctions between point estimation, model fitting and hypothesis testing.
- Computer Science. Iteration, recursion, basic complexity analysis, elemental data structures, and the ability to read and write simple pseudocode.

4 Schedule

Mondays and Fridays 1:30-3:00pm in Bondurant Hall. Note that the specific room varies from class to class (see Table 1).

5 Assignments and grading

After the first session, reading assignments are expected to be completed before class.

Table 1: Class schedule

Date	Room	Topic	Instructor
Fri, 1-Nov	2020	Pairwise sequence alignment	Vision
Mon, 4-Nov	2035	Sequence database search	Vision
Fri, 8-Nov	2020	Multiple sequence alignment	Vision
Mon, 11-Nov	2035	Phylogenetics	Vision
Fri, 15-Nov	2020	Sequence assembly	Wang
Mon, 18-Nov	2035	Profile HMMs	Wang
Fri, 22-Nov	2030	RNA structural analysis	Wang
Mon, 25-Nov	2035	Motif discovery	Wang
Mon, 2-Dec	2035	Student presentations	

There will be two problem sets, each worth 40%. These will be a combination of pencil-paper and computer exercises. Late problem sets will receive half-credit unless permission is obtained in advance.

The last session of the class consist of short presentations on selected papers by students. Paper discussions and overall class participation and preparedness are worth 20%.

All grades will be H/P/L/F.