

CS 681 Advanced Topics in Computational Biology

Spring 2013

Instructor: Can Alkan

Office hours: Tuesdays 10:30-12:00 in EA-509

This course is intended for graduate students and advanced undergraduates who would like to learn more about computational methods for solving fundamental problems related to molecular biology and genomics with a special focus on the analysis of high throughput sequence data. We will talk about genome structure, discovery and evolutionary analysis of genomic repeats and duplications, genome sequence analysis and motif discovery, non-coding RNA identification and search, RNA-RNA interaction prediction and search. Grading will be based on class participation, presentations and a final exam.

Prerequisites: Discrete mathematics, algorithms and data structures, and programming proficiency with functional and/or object oriented programming languages, e.g., C/C++ will be expected. Knowledge of combinatorics and probability, fundamental algorithms for sorting, hashing, dynamic programming, and elementary graph algorithms are required.

Textbook: Not required.

Recommended reference books:

- PLOS Computational Biology: Translational Bioinformatics. <http://www.ploscollections.org/translationalbioinformatics>.
- Bioinformatics Algorithms by Neil Jones and Pavel Pevzner (MIT Press, 2004).
- Algorithms on Strings, Trees and Sequences: Computer Science and Computational Biology by Dan Gusfield (Cambridge University Press, 1997)
- Bioinformatics: Sequence and Genome Analysis by David W. Mount (Cold Spring Harbor Lab, 2001)
- Biological Sequence Analysis by Durbin, Eddy, Krogh and Mitchison (Cambridge University Press, 1998)
- Scientific papers published in Nature Reviews Genetics, Nature, Science, Genome Research, Nature Genetics, Nature Methods, Bioinformatics, etc.

Grading: There will be one paper presentation and summary report (2-3 pages; constituting to 40% of the final grade), class participation (10%) and a project (2-3 person teams are accepted; 50%).