BIO/CSE/STAT 598F
Bioinformatics I: basic analysis of DNA and protein sequences

Fall semester, 2009
Tues/Thurs 2:30-3:45, 174 Willard
3 credits
Schedule numbers: BIO 598F - 401533; CSE 598F - 396454; STAT 598F - 403315
Web site: http://www.bx.psu.edu/courses/bx-fall09/

Prerequisites: Graduate standing or consent of instructors. Familiarity with molecular biology or computer methods is not assumed (though a willingness to learn a little about them is required).

Topics: (estimated time for discussion)

- Introduction to molecular biology for the non-biologist. (0.5 week)

- Internet resources for identifying genes in a genomic sequence and for predicting a gene's function. Students will use the UCSC Genome Browser, RepeatMasker, GenScan, Blast, and PipMaker. (1 week)

- Information about potential class projects. (2 weeks). At that point, students will need to specify groups and topics for their term projects.

- Selected topics in bioinformatics, including: introduction to algorithms and complexity (1 week), exhaustive search (1 week), greedy algorithms (0.5 weeks), dynamic programming algorithms (1 week), graph algorithms (1 week), combinatorial pattern matching (0.5 weeks), and hidden Markov models (1 week).

- Phylogenetic algorithms, gene and genome history, gene family evolution, models of sequence evolution, and phylogenetic analysis as a predictive tool. (1 week)

- Students' in-class reports on their term projects. (3 weeks). If necessary, additional reports will be scheduled for the week of final exams.

The course will cover neither protein structure prediction nor "post-sequencing" bioinformatics, such as analysis of expression data (e.g., from micro-arrays), proteomics, and analysis of regulatory networks. Protein structure prediction is covered in courses offered by the Chemistry Department, and post-sequencing bioinformatics is the topic of Bioinformatics II, which is taught most Spring Semesters.

Grading will be based on (1) a term project of the student's choice (60%), (2) an exam over the lecture material (30%), and (3) a few homework assignments, e.g., using World Wide Web resources for analyzing genomic sequences (10%).

Required text:

An Introduction to Bioinformatics Algorithms
Neal C. Jones and Pavel A. Pevzner.
MIT Press, 2004
ISBN 0-262-10106-8

Academic Integrity: All University and Departmental policies regarding academic integrity/academic dishonesty apply to this course and the students enrolled in this course. Refer to the following URLs for further details on the academic integrity policies at PSU of the Eberly College of Science: http://www.science.psu.edu/academic/Integrity/Policy.htm.

For more information:
Webb Miller, 506B Wartik Lab, 865-4551, webb@bx.psu.edu