Comp 50 / Bio 40 / BME 94-08: Bioinformatics Fall 2010

Course Goal: This course is designed to introduce future biologists and physicians to bioinformatics tools and analysis methods. Upon completion of the course, students should be more comfortable working with the vast amounts of biomedical and genomic data and online tools that will be relevant to their work in the coming decades.

Course strategy: This is a hands-on, project-oriented class. Most weeks will include one classroom lecture and one period of computer lab time. There will be several homework projects assigned throughout the term. Although labs are only graded pass/fail, timely completion of labs is essential for students to develop the basic skills needed to complete the deeper and more open-ended project assignments. Additional time in the lab will be offered for students who want more help with the online portion of their work. Lab work may also be completed on students' own machines at any time.

Staff: Professor Donna Slonim, slonim@cs.tufts.edu; Halligan 234.
TAs: TBA

Prerequisites: Bio 41 or BME 62 or equivalent (by consent). We will assume some background in molecular biology and genetics. There are no computer science prerequisites for this course.

Textbook: *Bioinformatics and Functional Genomics* (2nd edition) by Jonathan Pevsner, Wiley-Liss, ISBN# 978-0-470-08585-1.

Computational Resources: You will need access to a computer with an internet connection. If you are registered for the course in SIS, a computer account on the Windows machines in Halligan labs 120 and 122 should have been created for you. Your login name is your Trumpeter login or UTLN (usually something like "dsloni01"), and your password is your student ID number. If you would like to work on the labs and projects on your own machine, you will need a web browser and the ability to install your own software. You may bring your own laptop for use during lab time if you prefer.

Grading: Grades will be based on lab effort (15%), project assignments (50%), an in-class midterm (15%), and a final exam (20%).

Labs: Labs will be graded only as pass/fail. Labs will typically take place on Wednesdays, and solutions will be accepted until the following Sunday at 6pm, so that they can be discussed in class the next day. If you have not completed your lab work during the allotted time, we have the room reserved during the open block immediately following class. If you need additional time, you are welcome to work in either Halligan 120 or 122 at any time when it is not reserved for a class. (The schedule will be posted on the door.) We may also schedule additional office hours in the labs if there is interest.

Late policy: *Project* assignments handed in after the *start of the class period in which they are due* will be marked down 5% for each day they are late. (For simplicity, the "first day" is deemed to run from the beginning of the class in which the assignment is due until midnight *the following day*; subsequent days are counted from midnight to midnight.) *Labs* will be due the Sunday after they are assigned; turn in what you have by that point.

Collaboration Policy: All written work submitted should be your own unless you obtain prior permission to collaborate. You are free to discuss assignments with others in the class unless specifically asked not to, but you must write up your solutions, lab notes, and code yourself.

Tentative Schedule:

Wed. Sep. 8: Class Introduction. Bioinformatics: What and why?
Reading: Chapter 1

Mon. Sep 13: Genomic sequences. Online databases. Intro to sequence alignment.
Reading: Chapter 2
Wed. Sep 15: Lab 1 – Finding information in online databases.

Mon. Sep 20: Sequence alignment. Scoring Matrices. Pairwise alignment. Gaps.
Reading: Chapter 3, pp. 47-76 and 84-94
Wed. Sep 22: Lab 2 – Pairwise sequence alignment.

Mon. Sep 27: Database searching; BLAST. Limits of detection, significance.
Reading: Chapter 4
Wed. Sep 29: Lab 3 – BLAST.

Mon. Oct. 4: Advanced BLAST: PSI-BLAST, Genomic DNA. Find-a-gene project.
Reading: Chapter 5.
Wed. Oct 6: Lab 4 – Advanced BLAST and find-a-gene help.

Mon. Oct 11: Multiple sequence alignment. Relevance to inferences about evolution.
Reading: Chapter 6.
Wed. Oct 13: Lab 5 – Using online multiple sequence alignment tools.

Mon. Oct 18: Midterm review; molecular phylogeny introduction.
No Reading.
Wed. Oct 20: Midterm.

Mon. Oct 25: Molecular phylogeny and evolution.
Reading: Chapter 7.
Wed. Oct 27: Lab 6 – Molecular phylogeny.

Mon. Nov 1: mRNA and gene expression introduction. Unigene.
Reading: Chapter 8, pp. 300-323.
Wed. Nov 3: differential expression intro, normalization. visualization/clustering.
GenePattern.

Mon. Nov 8: Statistics for differential expression, multiple testing.
Reading: Chapter 9.
Wed. Nov 10: Lab 7 – Finding differentially expressed genes.

Mon. Nov 15: Functional interpretation of array data.
Reading: DAVID and GSEA papers.
Wed. Nov 17: Lab 8 – Interpreting expression variation.

Mon. Nov 22: Characterizing eukaryotic genomes.
Reading: Chapter 16.
Wed. Nov 24: NO CLASS: Thanksgiving vacation.

Mon. Nov 29: Human variation and disease.
Reading: Chapter 19.
Wed. Dec 1: Linking genes and disease.
Reading: Chapter 21.

Mon. Dec 6: Sequence variation, phenologs, comparative genomics.
Reading: Webb Miller paper.
Wed. Dec 8: Personalized medicine. Multiple testing, revisited.
Reading: Kohane paper.

Thu. Dec 16, 3:30pm: Final exam.