BIOL 526H: Computational Genetics
University of North Carolina at Chapel Hill
Fall 2008

Class: T-Th 11:00am-12:15pm, Hamilton 150
Lab: Th 1:00pm-1:50pm, Hamilton 150

Instructor: Todd Vision
Office: Coker 102
E-mail: tjv@bio.unc.edu
Office Phone: 843.4507
Office Hours: Tues 1:00-3:00pm and by appt.

TA: Toby Clarke
E-mail: thclarke@email.unc.edu
Office: Coker 122
Office Phone: 843.4508
Office Hours: Wed 1:00-3:00pm and by appt.

Description
A study of the concepts underlying the bioinformatic tools used in genetics. Topics include alignment, gene finding, expression analysis, mapping, phylogenetics and measuring sequence divergence and polymorphism. Course includes a computer lab session. 4 credit hours.

Prerequisites
This course is intended for honors undergraduates and beginning graduate students in the life sciences. Some familiarity with concepts in statistics and computer science will be helpful. Ideally students should have taken the following UNC courses or their equivalents, or receive a waiver from the instructor:

- BIOL 202 - Molecular biology and genetics
- COMP 101 - Introduction to computer programming
- STOR 155 - Probability and statistics

Reading
The textbook is *Introduction to Computational Genomics: A Case Studies Approach*, by Nello Cristianini and Matthew W. Hahn, Cambridge University Press (ISBN-13 978-0-521-67191-0).

Supplementary readings will be made available via class handouts and through the Blackboard site. Several advanced books are on reserve at the House Undergraduate Library, including

- Deonier R, Tavare S, Waterman M (2005) Computational genome analysis: An Introduction, Springer- Verlag (DTW05).
- Durbin R, Eddy S, Krogh A, Mitchison G (1998) Biological sequence analysis (DEKM98)
- Gibson G, Muse S (2004) A primer of genome science. Sinauer (GM04)

- Jones NC, Pevzner PA (2004) An introduction to bioinformatics algorithms. MIT Press (JP04)
- Mount D (2004) Bioinformatics, 2nd ed. Cold Spring Harbor Press (M04)

Computer labs

The course includes a 1-hour computer lab each week. In the first two-thirds of the semester, the lab sessions will be designed to help you complete the weekly problem-set assignments. There will be five problem set assignments, each to be completed as homework and handed in the following Thursday at the beginning of class. The assignments should require no more than four hours to complete (and usually less). If it is taking you longer than this, contact the instructor it is likely that you are not approaching the assignment as intended. You are encouraged to work together, but you must be able to show your own work and be able to defend your answers if asked. In the final third of the semester, the lab sessions will provide an opportunity to work on your independent project in the presence of the instructor and TA.

Grading

The distribution of letter grades will not be curved but letter grade divisions will be drawn at the instructors discretion. The grading scale will differ for graduate students and undergraduates. Points obtained on assignments are halved if they are late, unless permission is granted at least the day before. You are expected to be aware of and observe the UNC honor code (http://honor.unc.edu). Graded components include:

- Problem sets: five assignments worth 5% each (25% of the final grade in total).
- A midterm exam: worth 25% of the final grade. Questions will be similar to those on the problem sets, with the addition of some simple-answer questions (e.g. multiple choice, true/false).
- A class presentation on a primary-research paper: 10% of the final grade.
- An independent project, composed of both a paper and an oral presentation: 30% of the final grade.
- Attendance and participation is critical to the success of the seminar: 10% of the final grade.

Month	Date	Topic	
Aug	19	Introduction to computational genetics	
	21	Basic sequence analysis	Ch. 1
		Lab: Introduction to R	
	26	Gene finding	Ch. 2
	28	Gene finding	
	28	Lab	
Sept	2	Alignment	Ch. 3
	4	Alignment	Assignment 1 due
		Lab	
	9	Hidden Markov models	Ch. 4
	11	Hidden Markov models	
		Lab	
	16	Population genetics	Ch. 5
	18	Population genetics	Assignment 2 due
		Lab	
	23	Coding sequence evolution	Ch. 6
	25	Coding sequence evolution	
		Lab	
	30	Phylogenetics	Ch. 7
Oct	2	Phylogenetics	Assignment 3 due
		Lab	
	7	Genome evolution	Ch. 8
	9	Genome evolution	
		Lab	
	14	Reading the primary literature	TBA
	16	Fall Break - no class	
	21	Gene expression	Ch. 9
	23	Gene expression	Assignment 4 due
		Lab	
	28	Motif finding	Ch. 10
	30	Motif finding	
		Lab	
Nov	4	Paper discussions	
	6	Paper discussions	Assignment 5 due
		Lab: Independent project discussions	
	11	Review session	
	13	Midterm	
		No lab	
	18	Special topic: Collaborative software development in R	
	20	Special topic: Writing and presenting science	
		Lab: Independent projects	
	25	Special topic: Databases	
	27	Thanksgiving Break - no class	
Dec	2	Independent project presentations	
	8	Independent project papers due	