

Fall 2009 Honors Core Course (BME/CSE/MCB 1401)

Computational Molecular Biology

Lecture: E2322, Monday/Wednesday 11-11:50 am.

Lab: E2305, Friday 10-10:50 (section 001), 11-11:50 (section 002)

Instructors:

Craig Nelson, MCB Department (craig.nelson@uconn.edu), OH: TBD

Yufeng Wu, CSE Department (ywu@engr.uconn.edu), OH: T/Th: 2-3 pm (ITE 235).

Teaching Assistant: Jin Jun, phone: 486-1576, (jinjun@engr.uconn.edu), OH: Tue 5-6 pm (E2305)

Asav Dharia, phone: 486-1576, (asav.dharia@uconn.edu), OH: Wed 5-6 pm (E2305)

Textbook: N. Cristianini and M.W. Hahn, *Introduction to computational genomics: a case studies approach*, Cambridge University Press, 2007. Textbook website: <http://www.computational-genomics.net/>.

Course outline: This course is an introduction to computational genomics through lectures, computer lab exercises, and mentored research projects. Started in 1995 by the completion of the first genome sequence of a free-living organism, *H. influenzae*, the genomic era has already led to hundreds of complete genome sequences deposited in public databases and many more genome projects at various stages of completion. The huge amounts of available genome data are revolutionizing biomedical research, but fully exploiting them requires powerful computational and statistical methods. The main objective of the course is to provide students with a general understanding of the field of computational genomics, including current problems and research. Students will become familiar with fundamental molecular biology concepts and computational techniques, and will learn how to use the Matlab bioinformatics toolbox for solving problem in genomics.

Grading and course policies: Grading will be based on in-class quizzes, homeworks, computer lab assignments, and team final projects. Computer assignments will be assigned on Fridays and will be due by midnight the following Wednesday. Assignments must be submitted electronically via HuskyCT (see below). *No late assignments and make-up quizzes will be allowed.* The lowest quiz/homework and computer assignment grade will be omitted from the computation of the final grade. The last five weeks of the class will be devoted to a final project done in teams of three students. For the project you will pick a computational genomics topic not discussed in lectures and research it in more depth. You will be required to give weekly progress reports, submit a written final report, and give a 15-minute presentation at the end of the semester. The final project component of the grade will include participation in discussions of progress reports and final presentations of other teams.

HuskyCT: We have a HuskyCT site for the class; you can access it by logging in with your NetID and password at <https://huskyct.uconn.edu/>. You must use HuskyCT for submitting assignments and check it regularly for class materials, grades, problem clarifications, changes in class schedule, and other class announcements.

Academic honesty: You are expected to adhere to the highest standards of academic honesty. All submitted solutions must be your own work. You may discuss ideas and concepts with other people, but *must not share written solutions or computer code.* Use of published materials is allowed, but the sources

should be explicitly stated in your submission. Violations will be reviewed and sanctioned according to the University Policy on Academic Integrity.

Students with disabilities: If you have a documented disability for which you are or may be requesting an accommodation, you are encouraged to contact the instructor and the Center for Students with Disabilities or the University Program for College Students with Learning Disabilities as soon as possible to better ensure that such accommodations are implemented in a timely fashion.

Tentative Lecture Schedule

Date	Text	Topics	Computer Lab
8/31	Ch.1	DNA sequences, replication, mutations and evolution (CEN)	
9/2	Ch.1	Introduction to computer algorithms and probabilistic models (YW)	
9/4			Lab1: Introduction to Matlab and databases
9/7		No class (Labor Day)	
9/9	Ch.2	Transcription, translation, genetic code and structure of genes (CEN)	
9/11			Lab2: Sequence analysis, Exercise 1.1-1.3
9/14	Ch.3	Gene finding (YW)	
9/16	Ch.3	Homology, orthology, and paralogy, gene duplication and deletion (CEN)	
9/18			Lab3: ORF finding, Exercise 2.1-2.3
9/21	Ch.3	Sequence Alignment (YW)	
9/23	Ch.3	BLAST (YW)	
9/25			Lab4: Sequence alignment & BLAST Exercise 3.1-3.4
9/28	Ch.5	Mutations, Mitochondrial DNA and human evolution (CEN)	
9/30	Ch.5	genetic distances (YW)	
10/2			Lab5: Multiple alignment & genetic distance Exercise 5.1-5.3
10/5	Ch.6	Evolution and selection (CEN)	
10/7	Ch.6	Estimating Ka/Ks (YW)	
10/9			Lab6: Ka/Ks Exercise 6.1-6.3
10/12	Ch.7	SARs, virus and virus-host interaction (CEN)	
10/14	Ch.7	Neighbor joining (YW)	
10/16			Lab7: Phylogeny Exercise 7.1-7.3
10/19	Ch.8	Whole-genome comparison, synteny (YW)	
10/21	Ch.8	Sorting by reversal and phylogeny (YW)	
10/23			Lab8: Reversal distance in

			phylogeny Exercise 8.1-8.4
10/26	Ch.10	Motif finding (YW)	
10/28	Notes	Application of genomics in forensics and identification (YW)	
10/30			Lab9: Motif finding, Exercise 10.1-10.3
11/2		Project: topic selection	
11/4			
11/6			Topic selection and specific aim due
11/9		Progress report (10' presentation): Literature review and specific aims	
11/11			
11/13			Literature review and hypothesis due
11/16		Progress report (10' presentation): define hypothesis, analysis strategy and data source	
11/18			
11/20			Analysis strategy due
11/23-11/27		Thanksgiving recess	
11/30		Progress report (10' presentation): Preliminary results	
12/2			
12/4			Preliminary results due
12/7		Final project presentation	
12/9			
12/11			