

Genomics and Applied Bioinformatics Biol 4150/6150, Fall 2012

Jung H. Choi

jung.choi@biology.gatech.edu; Twitter @jung_gt

Recommended (Optional) Textbooks:

8803B: Zvelebil and Baum, *Understanding Bioinformatics*, Garland Science, 2008 (Z&B)

4150: Jonathan Pevsner, *Bioinformatics and Functional Genomics*, 2nd ed. Wiley-Blackwell, 2009 (JP)

Required readings: Journal articles, blog posts will be posted as PDF files on the course T-square site, or emailed as hyperlinks.

Recommended equipment: If you do not have your own laptop, you will need a USB drive to save your work from the computer lab sessions. I recommend 4 Gb minimum. If you have not already done so, I also recommend that your laptop be configured as a dual-boot linux/Windows configuration, or be able to run the 2nd OS as a virtual machine. There are many “flavors” or “distros” of linux to choose from; most are available free.

Learning objectives: You should be able to:

1. Locate and access a variety of public databases and software tools for bioinformatics.
2. Use/apply the common tools appropriately, with knowledge of their limits and pitfalls.
3. Evaluate the statistical significance of results of bioinformatics searches and analyses, where appropriate.
4. Evaluate the biological significance of, or formulate new questions based on the results of, bioinformatics searches and analyses.

Grading:

10%	Web assignments – blog, wiki, etc.
30%	Two midterm exams (take-home)
40%	Computer projects
10%	Group paper presentations
10%	In-class group work

Projects for graduate students will be more complex, involve writing scripts, and require a higher level of analysis.

Policy on collaboration and the Honor Code: Discussion with peers is encouraged, both during class group discussion & work periods and during computer exercises. However, all work turned in must be of your own original efforts. Using the work of your peers (copying and pasting, even with alterations) or from any other sources, on-line or in print, without attribution is plagiarism. Plagiarism will result in zero credit for the assignment, and possible further penalties as an Honor Code violation.

Tentative Schedule of Topics - first iteration: The table below is the first iteration of the schedule of topics. As the course progresses, this schedule may be further adjusted to fit the pace of learning.

<i>Week</i>	<i>Topic</i>	<i>Computer Lab/Homework</i>
1:	Intro, public databases Next-generation DNA sequencing technologies	Genbank tutorial Configure laptop Set up blog
2:	Pairwise sequence alignment – local and global alignments	Needleman-Wunch, Smith-Waterman
3:	BLAST & Advanced BLAST	BLAST, PSI-BLAST
4:	Multiple sequence alignment & molecular phylogeny	Take-home exam 1
5:	High-throughput approaches to gene expression	microarray data normalization
6:	Microarray data analysis – normalization, statistical analysis	Affymetrix data analysis project
7:	Microarray data analysis – clustering, pathways	
8:	Microarray data analysis	
9:	Other types of microarrays	Exam 2
10:	Protein structures - classification and prediction	Homology modeling project
11:	Protein structures – homology modeling	
12:	Genome sequencing and annotation	Final Group Project
13:	Bacterial genomes and Paper Presentations	
14:	Fungal genomes & Paper Presentations	
15:	Vertebrate genomes & Paper Presentations	
16:	Medical genetics & Paper Presentations	