

## **BIOL 4540, Bioinformatics I, Sequence Analysis. FALL 2004**

This course covers the theory and practice of biological sequence analysis, including alignment, database searching and algorithms.

**Place:** Jonsson-Rowland Science Center Rm 2C13

**Time:** Monday, Thursday 2:00-3:50

**Instructors:** Chris Bystroff

John Salerno

**office:** MRC 303

SC 3C38

**phone:** 518-276-3185

518-276-6392

**email:** [bystrc@rpi.edu](mailto:bystrc@rpi.edu)

[salerj@rpi.edu](mailto:salerj@rpi.edu)

**TA:** None.

**Office Hours:** Tuesdays, 12-1pm, 2C13

**Web sites:** <http://www.bioinfo.rpi.edu/~bystrc/courses/biol4540.html>

<http://www.bioinformaticsonline.org>

### **REQUIRED TEXT:**

David W. Mount “Bioinformatics. Sequence and Genome Analysis. Second Edition” Cold Spring Harbor Laboratory Press, 2004.

Other required readings will be handouts, reserves, or online material.

### **GRADING:**

Homework 30%

Midterm exam 20%

Term project 20% (Due on last day of class)

Final exam: 20% (date and time TBA)

In-class exercises and participation: 10%

Please turn in homework on paper at the beginning of class on the day due unless otherwise specified. Both exams have a written and a practical (computer) part.

## BIOL 4540, Bioinformatics I, Sequence Analysis. FALL 2004

	<i>Date</i>	<i>Topic</i>	<i>Reading [Mount 2<sup>nd</sup> ed.]</i>	<i>HW, exercises</i>
1	Mon 8/30	Basics. Sequencing . Edit distance.	Ch.1, Ch 2:29-59	Unix, vi
2	Thu 9/2	Databases. Dot plots.	Ch 3:65-79	SeqLab
3	Thu 9/9	Pairwise sequence alignment without gaps	Ch 9:366-370	HW 1 due
4	Mon 9/13	Gapped alignment. Dynamic programming	Ch 3:79-90	DP on paper
5	Thu 9/16	Amino acid substitution matrices	Ch 3:94-113	affine DP
6	Mon 9/20	Global versus local alignment	Ch 3:91-94	HW 2 due local DP
7	Thu 9/23	Database searching using FASTA and BLAST	Ch 6:227-259	NCBI blast
8	Mon 9/27	Statistical Significance.	Ch 4:127-148	e-value calculation
9	Thu 9/30	Multiple sequence alignment	Ch 5: 164-189	HW 3 due Distance calc.
10	Mon 10/4	Editing MSAs by hand.	Ch 5:215-217	SeqLab editing
11	Thu 10/7	Phylogenetic tree construction.	Ch 7:281-295,301-306	distance tree
12	Tues 10/12	Parsimony. Orthologs and paralogs.	Ch 7:296-301, 321-322, Ch 11:519-521	HW 4 due parsimony tree
13	Thu 10/14	Weights, Odds, Profiles and Logos	Ch 5:189-194, 211-215, Ch 6:261-271	Drawing logos
14	Mon 10/18	<i>review session</i>	Ch 2-7	
15	Thu 10/21	<b>Midterm Exam</b>		
16	Mon 10/25	Markov chains and hidden Markov models	handouts	Assign projects. Viterbi algor.
17	Thu 10/28	Profile HMMs	Ch 5:204-210	Profile HMM
18	Mon 11/1	Clustering, classification	handouts	K-means
19	Thu 11/4	Bayesian adaptive alignment	Ch 6:259-261	HW 5 due work on project
20	Mon 11/8	Single Nucleotide Polymorphisms (SNPs)	<a href="http://www.ncbi.nlm.nih.gov/About/primer/snps.html">www.ncbi.nlm.nih.gov/About/primer/snps.html</a>	bio-war scenario
21	Thu 11/11	Footprints	Ch 9:385-400	MEME
22	Mon 11/15	RNA secondary structure	Ch 8	work on project
23	Thu 11/18	PCR primer design	handouts	chimera
24	Mon 11/22	Microrepeats, Satellites and transposons	Ch 11: 506-509 handouts	HW 6 due, Finding Alu
25	Mon 11/29	Microarrays	Ch 13	work on project
26	Thu 12/2	Gene finding, splicing.	Ch 9:366-385	work on project
27	Mon 12/6	<b>Student Presentations</b>		Term papers due
28	Thu 12/9	<i>review session</i>		
29	(Thu12/16)	<b>Final Exam</b>		