

## BIOL595A/CS590B : Protein Bioinformatics

**Course Content:** Accumulation of biological data, such as genome sequences, protein structures and sequences, metabolic pathways, opened up a new way of research in biology – bioinformatics. Through the survey of the various active research topics in bioinformatics, in this course we will learn bioinformatics databases, tools, and algorithms behind these tools. Special emphasis is placed on protein sequence and structure analyses. Covered topics will include methods for protein sequence comparison, protein structure comparison, protein structure prediction/modeling, protein docking prediction, protein function prediction, and protein network analysis.

**Grading:** Grades in the course will be based on one mid-term and one final exam, homework assignments, and projects. The course will be concluded by presentations of group projects. Semester grades will be awarded based on the following minima of performance: 90% = A, 80% = B, 70% = C and 60% = D; < 60% = F. These thresholds will not be raised but may be lowered at the instructor's discretion.

**Projects:** Considering that this course is cross-listed between Biology and Computer Science, students will choose either of programming-type or mini-review-type (+ using existing software) project by the consent of instructor.

**Instructor:** Dr. Daisuke Kihara ([dkihara@purdue.edu](mailto:dkihara@purdue.edu))

**Time & Place:** MWF, 12:30-1:20pm, LILY G424.

### Spring 2006: Provisional Schedule

Week	Lec. #	Date	Topic
1	1.	Jan. 9	Introduction
	2	Jan. 11	Sequence alignment
	3.	Jan. 13	Suboptimal alignments <b>Homework #1 (sequence alignment) out</b>
2	-	Jan. 16	<i>(Official University Holiday)</i>
	4.	Jan. 18	Profiles, multiple sequence alignments
	5.	Jan. 20	homology search, FASTA, BLAST, PSI-BLAST <b>Homework #1 turn in</b>
3	6.	Jan. 23	Scoring matrices
	7.	Jan. 25	Hidden Markov models (protein family)

	8.	Jan. 27	Protein motif search
4	9.	Jan 30	<i>several databases tools for motif, domain search</i> <b>Homework #2 (using motif DB, HMM) out</b>
	10.	Feb. 1	Protein localization prediction (PSORT)
	11.	Feb. 3	Analysis of protein-protein interaction data
5	12.	Feb. 6	Protein-protein interaction cont.
	13.	Feb. 8	Protein function: GO, EC number and function prediction <b>Homework #2 in</b>
	14.	Feb. 10	Function prediction through comparative genomics, <i>tools</i>
6	15.	Feb. 13	neural network: Protein secondary structure prediction
	15.	Feb. 15	Protein Secondary Str. prediction cont. <b>Project #1 out</b>
	16.	Feb. 17	Support vector machine
7	17.	Feb. 20	Transmembrane domain prediction
	18.	Feb. 22	Coiled-coil region, disordered region
	19.	Feb. 24	Homology modeling
8	20.	Feb. 27	Midterm exam review <b>Midterm</b>
	21.	Mar. 1	Threading
	22.	Mar. 3	Threading cont. <b>Homework #3 out</b>
9	23.	Mar. 6	<b>Project middle report</b>
	24.	Mar. 8	Ab initio prediction
	25.	Mar. 10	Ab initio prediction cont. <b>Homework #3 in</b>
11	-	Mar. 13	(Spring Break)
	-	Mar. 15	
	-	Mar. 17	
12	26.	Mar. 20	Meta-server
	27.	Mar. 22	CASP
	28.	Mar. 24	<b>Project #1 presentation</b>
13	29.	Mar. 27	<b>Project #1 presentation cont.</b>
	30.	Mar. 29	Geometric hashing: Protein-protein docking prediction
	31.	Mar. 31	Fast Fourier Transform and the others: Protein-protein docking prediction cont.

14	32.	Apr. 3	Protein docking cont. <b>Project #2 out</b>
	33.	Apr. 5	Prediction of protein docking interface
	34.	Apr. 7	Protein ligand docking
15	35.	Apr. 10	Protein function prediction from 3D structure <b>Homework #4 out</b>
	36.	Apr. 12	Protein structure comparison
	37.	Apr. 14	Protein structure comparison cont.
16	38.	Apr. 17	RNA secondary structure prediction
	39.	Apr. 19	Suffix tree
	40.	Apr. 21	Phylogenetic tree <b>Homework #4 in</b>
17	41.	Apr. 24	<b>Project #2 presentation</b>
	42.	Apr. 26	<b>Project #2 presentation cont.</b>
	43.	Apr. 28	Course Evaluation
18		May1-5.	<b>Final Exam Week</b>