

Introduction to Bioinformatics (Graduate Section)

For the time being, I will not require written reviews. However, unlike previous years, I will not assign papers in advance to specific reviewers. Rather I will call on one participant at random to lead the discussion of a particular paper.

When you are called, you should assume that everyone has actually read the paper so you do not have present a detailed summary of the paper. Instead, focus on the following:

- very brief synopsis of paper - not a detailed description of every point; focus on the logical argument and the following questions: What data? What analysis? What is their argument and does it support their conclusions? Why is this paper important?
- Any confusing or difficult to understand sections
- strong points of the paper
- weak points of the paper
- validity of conclusions

59500 Paper schedule		
		Tentative (Don't read yet)
Week	Lecture Topic	Papers
6 September		<ul style="list-style-type: none"> • Needleman SB, Wunsch CD., A general method applicable to the search for similarities in the amino acid sequence of two proteins. <i>J Mol Biol</i> 1970 48:443-453. PDF • Fitch WM, Smith TF. Optimal sequence alignments. <i>Proc Natl Acad Sci USA</i> 1983 80:1382-6. PDF
13 September	Sequence Alignment and Scoring	<ul style="list-style-type: none"> • Henikoff S, Henikoff JG., Amino acid substitution matrices from protein blocks. <i>Proc Natl Acad Sci USA</i> 1992, 89, 10915-10919. PDF • Chang MSS, Benner SA, Empirical Analysis of Protein Insertions and Deletions Determining Parameters for the Correct Placement of Gaps in Protein Sequence Alignments, <i>J Mol Biol</i> 2004 341:617-631. PDF
20 September	Database searching	<ul style="list-style-type: none"> • Altschul SF, Madden TL, Schaeffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ, Gapped BLAST and PSI-BLAST: a new generation of protein database search programs, 1997, <i>Nucleic Acids Res</i> 25:3389-3402. PDF • Kent WJ. BLAT--the BLAST-like alignment tool. <i>Genome Res</i> 2002 12:656-64. PDF
27 September	Motifs	<ul style="list-style-type: none"> • Nevill-Manning CG, Wu TD, Brutlag DL. Highly specific protein sequence motifs for genome analysis. <i>Proc Natl Acad Sci USA</i> 95, 5865-5871, 1998. PDF • Trachana K, Larsson TA, Powell S, Chen WH, Doerks T, Muller J, Bork P. Orthology prediction methods: a quality assessment using curated protein families. <i>Bioessays</i> 33, 769-80, 2011. PDF

59500 Paper schedule

Tentative (Don't read yet)

Week	Lecture Topic	Papers
4 October	Trees	<ul style="list-style-type: none"> • L. Vigilant, M. Stoneking, H. Harpending, K. Hawkes, A. C. Wilson, <i>Science</i> 253, 1503 (1991). <ul style="list-style-type: none"> ◦ Templeton AR, Human origins and analysis of mitochondrial DNA sequences, <i>Science</i> 255, 737 (1992) ◦ Hedges SB, Kumar, S, Tamura K, <i>Science</i> 255, 737-739 (1992) • Felsenstein J. Cases in which parsimony or compatibility methods will be positively misleading. <i>Systematic Zool</i> 27, 401-410, 1978.
11 October	Multiple Alignment	<ul style="list-style-type: none"> • Thompson JD, Higgins DG, Gibson TJ, CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice, <i>Nucleic Acids Research</i> 22, 4673-4680, 1994. • RC Edgar. MUSCLE: multiple sequence alignment with high accuracy and high throughput. <i>Nucleic Acids Res.</i> 32, 1792-1797, 2004.
18 October	Profiles & HMMs	<ul style="list-style-type: none"> • LS Johnson, SR Eddy, E Portugaly. Hidden Markov model speed heuristic and iterative HMM search procedure. <i>BMC Bioinformatics</i> 11, 431, 2010. • Kawaji H, Schönbach C, Matsuo Y, Kawai J, Okazaki Y, Hayashizaki Y, Matsuda H. Exploration of novel motifs derived from mouse cDNA sequences. <i>Genome Res.</i> 12, 367-378, 2002
25 October	Genome Sequencing and annotation	<ul style="list-style-type: none"> • AM Phillippy, MC Schatz and M Pop, Genome assembly forensics: finding the elusive mis-assembly, <i>Genome Biology</i> 9, R55, 2008. • Lomsadze A, Ter-Hovhannisyann V, Chernoff YO, Borodovsky M. Gene identification in novel eukaryotic genomes by self-training algorithm. <i>Nucleic Acids Res.</i> 33, 6494-6506, 2005.
1 November	Gene Finding and Annotation	<ul style="list-style-type: none"> • Kamoun C, Payen T, Hua-Van A, Filée J. Improving prokaryotic transposable elements identification using a combination of de novo and profile HMM methods. <i>BMC Genomics</i> 14:700, 2013. • R Overbeek, M Fonstein, M D'souza, G D. Pusch, and N Maltsev, The use of gene clusters to infer functional coupling, <i>Proc. Natl. Acad. Sci. USA</i>, 96, 2896-2901, 1999
8 November	Protein Structure Prediction	<ul style="list-style-type: none"> • Iakoucheva LM, Radivojac P, Brown CJ, O'Connor TR, Sikes JG, Obradovic Z, Dunker AK. "The importance of intrinsic disorder for protein phosphorylation" <i>Nucleic Acids Res.</i> 32, 1037-1049, 2004 • A. Kloczkowski, K.-L. Ting, R.L. Jernigan, and J. Garnier. "Combining the GORV Algorithm With Evolutionary Information for Protein Secondary Structure Prediction From Amino Acid Sequence." <i>Proteins</i> 49, 154-166. 2002.
15 November	Protein Modeling	<ul style="list-style-type: none"> • Bryant SH, Lawrence CE. An empirical energy function for threading protein sequence through the folding motif. <i>Proteins</i> 16, 92-112, 1993 • Simons KT, Kooperberg C, Huang E, Baker D. Assembly of protein tertiary structures from fragments with similar local sequences

59500 Paper schedule

Tentative (Don't read yet)

Week	Lecture Topic	Papers
		using simulated annealing and bayesian scoring functions. J.Mol.Biol 268, 209-225, 1997.
22 Nov	Thanksgiving holiday	Thanksgiving holiday - TBA - there will be a written review for this week
29 November	Protein Homology modeling and docking	<ul style="list-style-type: none"> Zhou JM, Lee E, Kanapathy-Sinnaiaha F, Park Y, Kornblatt JA, Lim Y, Ibrahim RK. Structure-function relationships of wheat flavone O-methyltransferase: Homology modeling and site-directed mutagenesis. BMC Plant Biology, 10:15, 2010. http://www.biomedcentral.com/content/pdf/1471-2229-10-156.pdf Geisler M, Wilczynska M, Karpinski S, Kleczkowski LA. Toward a blueprint for UDP-glucose pyrophosphorylase structure/function properties: homology-modeling analyses. Plant Mol Biol. 56:783-794, 2004. http://www.springerlink.com/content/g64688w486218k56/fulltext.pdf
6 December	Systems Biology	<ul style="list-style-type: none"> Barabasi A-L, Albert R. Emergence of scaling in random networks. Science 286, 509-512, 1999. Michael P. H. Stumpf, Thomas Thorne, Eric de Silva, Ronald Stewart, Hyeon Jun An, Michael Lappe and Carsten Wiuf Estimating the size of the human interactome. PNAS 105, 6959-6964, 2008.

-
-