582653 Computational methods of systems biology, Autumn 2009
Homework 2
Group 1: Thursday Nov. 19th 14-16 B119
Group 2: Thursday Nov. 19th 16-18 B119

General instructions

Problems for each exercise session will be distributed approximately one week before the session. You are expected to be prepared to present your solutions in the exercise session.

Assignments

1. Read the paper

   > B. Schuster-Böckler, J. Schultz, and S. Rahmann: HMM Logos for visualization of protein families. BMC Bioinformatics 2004, 5:7

   and explain what are the HMM Logos and how such a logo can be constructed for a given HMM.

2. Familiarize yourself with the PFAM database at http://pfam.sanger.ac.uk/. PFAM visualizes protein families as multiple alignments and HMM Logos. Compare the visualizations of the Piwi-domain at http://pfam.sanger.ac.uk//family/Piwi (use the buttons 'Alignments' and 'HMM logo' on the left panel to get the visualizations). Which one is better? Explain.

3. Try the HMMER tool (http://hmmer.janelia.org/). Produce a profile-HMM from the alignment given in Figure 5.3 of Durbin. Is the resulting HMM any good? (More guidance of using HMMER available http://www.cs.helsinki.fi/u/lmsalmel/cmsb09/exercises/hmmer.shtml)

4. Estimate for the HMM in Fig 5.4. of Durbin the transition probabilities from state M3 to state I3 and from state M3 to state D4 using the alignment of Fig 5.3. of Durbin.

5. Try CLUSTAL program for constructing multiple alignments (more guidance http://www.cs.helsinki.fi/u/lmsalmel/cmsb09/exercises/clustal.shtml).

6. T-Coffee is a recent multiple alignment software. Explain the main principles of T-Coffee. How does it differ from CLUSTAL? (Hint: Find answers from the WWW). Try T-Coffee (the server is here: http://www.tcoffee.org/).