Weekly Report

Lin Yang 08/26/2012

Progress in database website construction

All the files (php scripts) and the database have been moved to the lab server machine and can be accessed through link http://rohslab.cmb.usc.edu/server. However, two problems arose: one is undesired characters appeared on the heat maps generated by GNUPLOT scripts and has been solved, the other problem remaining unsolved is the symbol Å in the title of heat maps cannot be displayed correctly by GNUPLOT, possibly due to system settings of the server’s operation system (Figure 1). A workaround for this is to generate heat maps without title using GNUPLOT and then use an external program to merge the figure without title to a blank figure that contains only the title.

Figure 1 The Å symbol in the title cannot be displayed correctly.
**Things to be done**

1. Add all JASPAR data to the database.

2. A better way of showing the shape-sequence motif logos:
   
   One idea is to keep the sequence logo as one figure to show and show the shape logo separately as the following: divide minor groove width into groups of very narrow, narrow, wide and very wide which can further be represented by four different letters and thus can be used to make shape logos just as making sequence logos. One advantage of this scheme is the sequence logo and shape logo can be combined together by choosing to use sequence letter or shape letter depending on which contains more information content for each position. And such combination of sequence and shape logos would be the target of the potential motif discovery project.


5. More searching criteria.