Homework #2 Due by 3/27. Please submit it via email.

I - Identify functional domains and motifs in your protein: Choose a protein sequence of your interest, search for motifs/domains at Pfam and Prosite. Submit the output and summarize your findings.

II - Identifying shared motifs: Compile or download a set (>3) of related protein (or DNA) sequences in FASTA format in a single text file. Load these sequences to the MEME server to identify shared motifs among these sequences. Turn in the MEME results (as .pdf or .html) or the link to the search result page. Explain briefly in writing the rationale of your input sequence set (why do you choose those?) and what you learned from the MEME analysis.

III - Phylogenetic analysis: Compile a set (>8) of protein (or DNA) sequences in FASTA format in a single text file. Perform an alignment and construct an N-J tree using ClustalOmega or other packages. Turn in the phylogenetic tree as a .pdf file.

IV – **Finish the RNA-Seq analysis**. If you did this in the gms6014/share folder, you don’t need to submit anything; I can see the results in that folder. If you did this on your own folder, list folders and files in your RNA-Seq folder with “ls -aLR”, then print the screen to show you have generated the files. Submit the screen lift image file for this assignment.

V - Gene Ontology and functional information: Search and analyze the GO entry (entries) of a biological process related to your research interest. Identify genes involved in a particular biological process of interest to you. How many genes in your model are annotated as involved in this process?

Compress all of the result files into an Hmwk2_Lastname.zip file and submit it via email to leizhou@ufl.edu.