CompMoby incorporates evolutionary conservation information with the existing MobyDick algorithm to detect over-represented motifs in the upstream or 3' UTR sequences of co-expressed genes.

For example, download zip file containing all sample input files for upstream promoter or 3' UTR sequences for test run. Explanation of the results of the upstream promoter example can be found here. It is recommended that you <a href="http://genome.ucsc.edu/compmoby/">RepeatMask</a> all your input sequences before running CompMoby.

1. Please input an upstream or 3' UTR sequence file of the genes in the reference species in FASTA format. (Maximum number of sequences: 200) [sample input file].
2. Please input an upstream or 3' UTR sequence file of the orthologous genes in the orthologous species in FASTA format. (Maximum number of sequences: 200) [sample input file].
3. Please input the text file containing aligned non-coding regions for your subset of genes from the alignment wrapper script (alignment extraction script) or a FASTA format file containing pair-wise alignments from an algorithm of your choosing [sample input file in FASTA format].
4. Please input a file containing a background set of upstream or 3' UTR sequences in the reference species in FASTA format for p-value calculation [sample input file]. For files with greater than 1000 sequences, please gzip your file to reduce uploading time.
5. Type of sequence to analyze:
6. Length of upstream sequence (promoter option) or downstream sequence (3' UTR option) to analyze. For the 3' UTR option, the user can analyze the entire length of the given 3' UTR by entering 0 as the option.
7. Filter out motifs (e.g., short AT-rich repeats) with occurrences greater than (dependent on length of sequences; 200 used in references 1 and 2). 200
8. Clustering threshold parameter (0.35 used in references 1 and 2).
9. Bonferroni corrected $\log_{10} p$-value cutoff of enriched motifs.
10. Please provide a current email address. A link to your results will be emailed to this address when the CompMoby analysis is completed.

References:


(last updated: June 3, 2008)