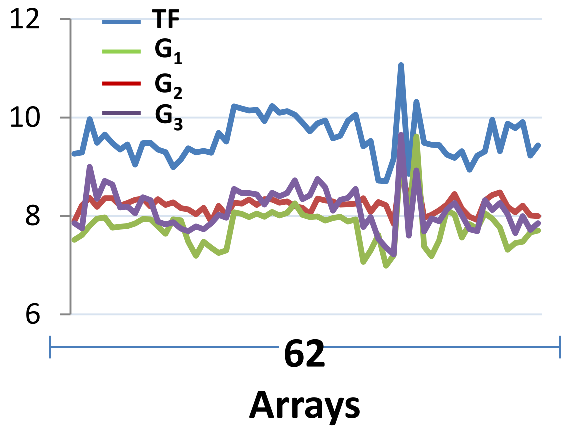
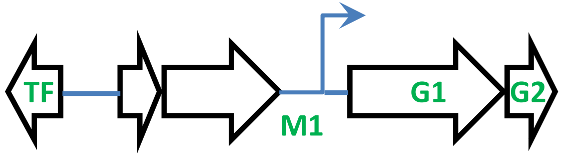
A. Correlated expression profile



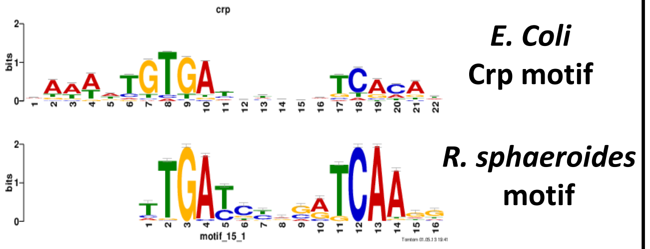
If a transcription factor (TF) is highly correlated to other genes (G) within a cluster, the TF is given a high score for that cluster.

B. TF proximity to cluster motifs



If a transcription factor's (TF) genomic location is close to one (or more) of the motifs (M1) identified in the cluster, the TF is given a high score.

C. Structure of DNA-binding domain (DBD)



The DBD family of each *R. sphaeroides* TF was determined by Pfam analysis. Then for each identified phylogenetically conserved motif, potential TFs are scored by comparing this motif to the *E. coli* motif of the DBD family to which a given TF belongs. TFs with high similarities get high scores.

D. Phylogenetic correlation

Organisms	Motif occurrence correlation	*TF* occurrence
R. sphaeroides 2.4.1	1	1
R. capsulatus	0.5	1
R. sphaeroides ATCC17025	0.9	1
R. denitrificans	0.6	1
B. japonicum	0	0
P. denitrificans	0	0
R. palustris	0.5	0
D. shibae	0.5	1
Phylogenetic correlation		**0.76**

If there is a strong correlation between the occurrence of a TF and a motif across the selected genomes, then the TF gets a high score.