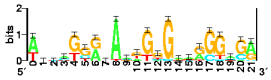
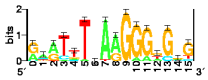
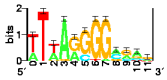
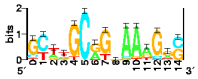
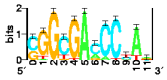
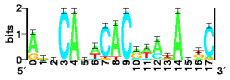
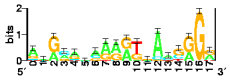
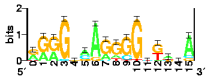
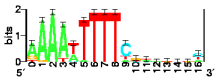
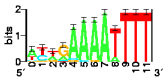
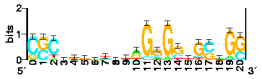
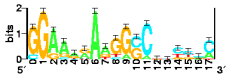
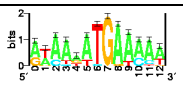
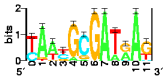
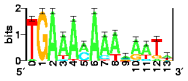
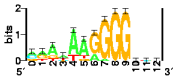
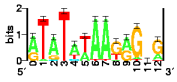
Supplemental Information

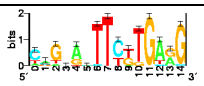
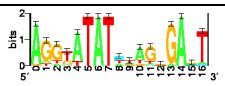
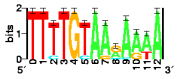
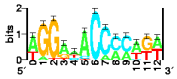
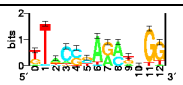
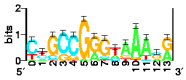
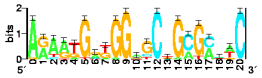
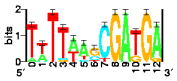
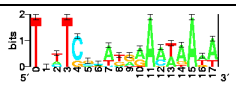
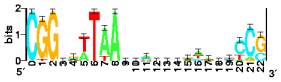
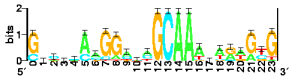
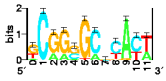
Includes: Supplemental Table I, Supplemental References, Supplemental Figure 1, Supplemental Figure 2A-D, Supplemental Figure 3.

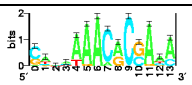
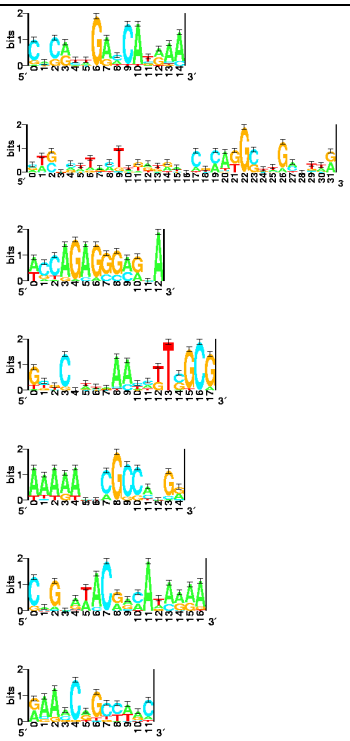
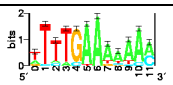
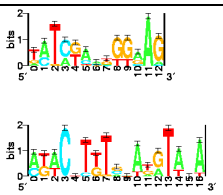
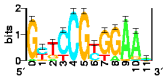
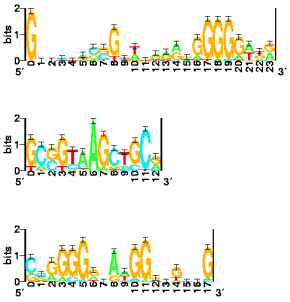
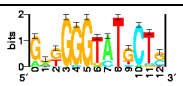
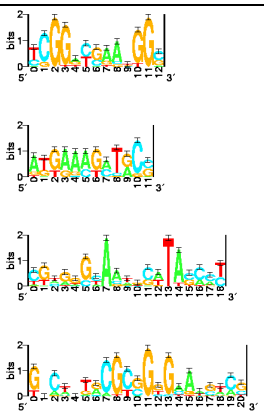
Prevalence of sequence features chosen by the Bayesian network. The types of constraints learned by the network indicate the prevalence of various modes of combinatorial regulation in *S. cerevisiae.* It is important to note that the degree of combinatorial regulation uncovered here represents a lower limit, and a more diverse sampling of physiological conditions is likely to yield a higher average number of regulators per gene, and more complex combinatorial rules. Of 6,906 parents selected on 5 cross-validations and 10 re-samplings on 49 clusters, 6,498 constraints (94%) were for the presence or absence of a motif. Constraining the distance to ATG was chosen 5.1% of the time, for subset of motifs, foremost among which are PAC, RRPE, RPN4, STRE, and RAP1, indicating potential mechanistic constraints for their function. Orientation was constrained 0.6% of the time, and the vast majority of these were for RAP1. So our analysis indicates that transcriptional regulation by RAP1 depends on its orientation, but that most other factors, including PAC and RRPE, work independently of orientation. All other constraints were chosen at even lower frequencies. While the learned sequence features are predictive, some functional constraints may not be detected if they are regularly satisfied in a set of genes selected by a previous constraint.

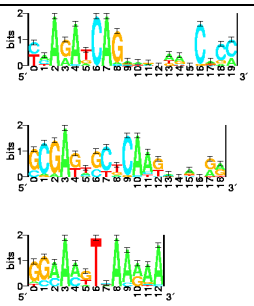
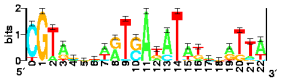
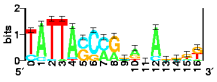
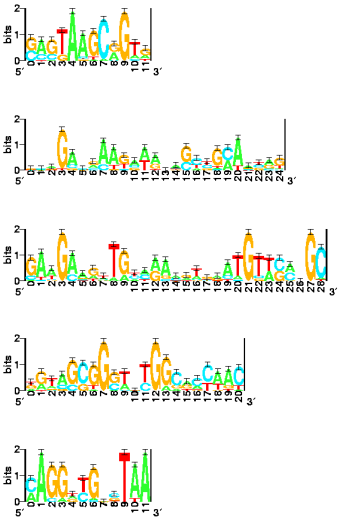
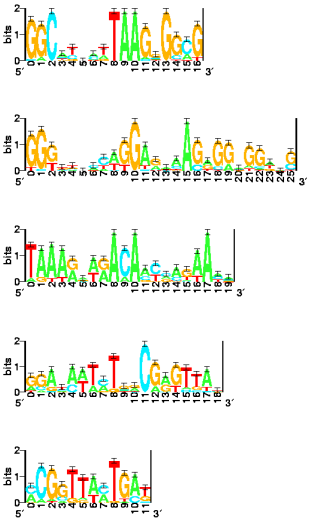
Supplemental Table I. High confidence motifs learned from the Bayesian Network for each cluster, found in at least 25% of the re-samplings.

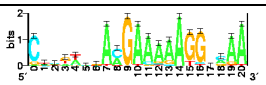
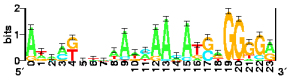
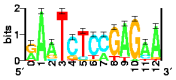
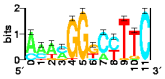
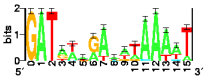
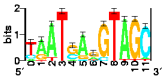
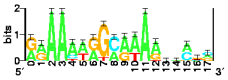
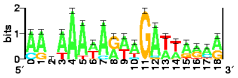
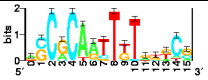
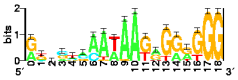
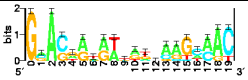
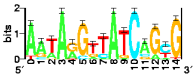
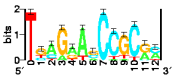
Cluster	Known motifs	New motifs
1	 RAP1  RAP1-var	
2		  STRE-like  STRE-like  
3		      STRE-like

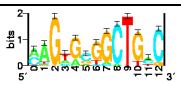
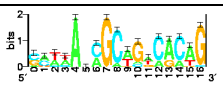
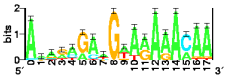
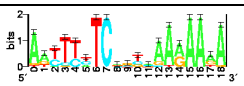
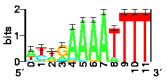
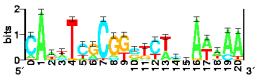
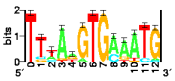
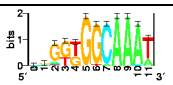
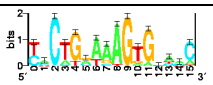
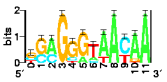
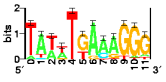
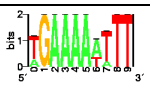
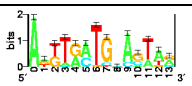
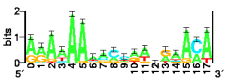
4		PAC	
		RRPE	
		RRPE	
5		HSF1-var	   
6		STE12-like	
		PAC-var	
		STRE	
		STE12	
		RRPE	
7		STRE	 TATA-var
		STRE-like	
		STRE	

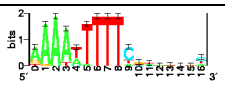
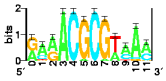
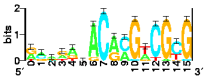
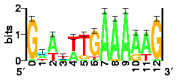
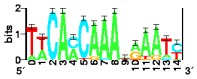
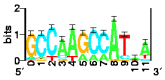
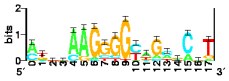
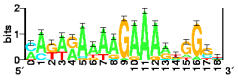
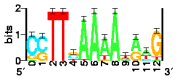
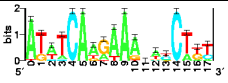
8		HSF1-like	   
9	 	MCM1-like REB1-like	  
10	  	PAC PAC-var PAC	
11	 	RPN4 OAF1	    

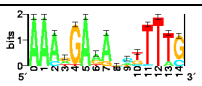
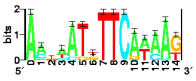
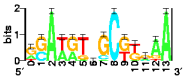
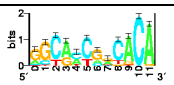
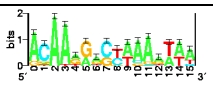
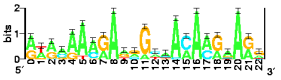
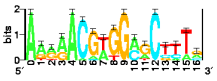
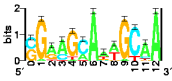
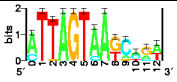
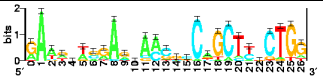
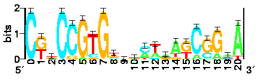
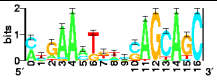
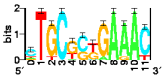
12		SWI4	
13		STE12-like	
14		PDR3	
15		MIG1	

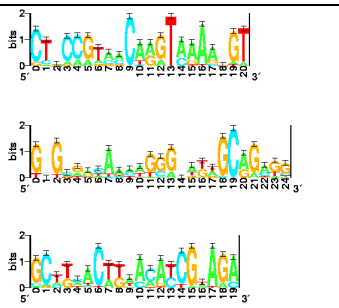
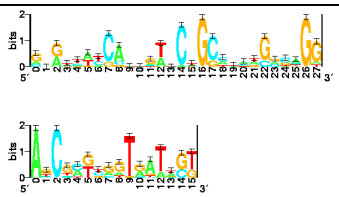
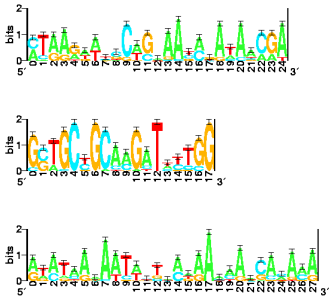
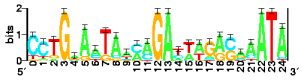
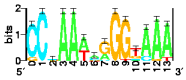
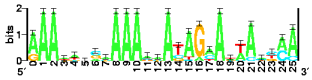
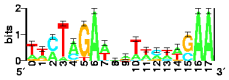
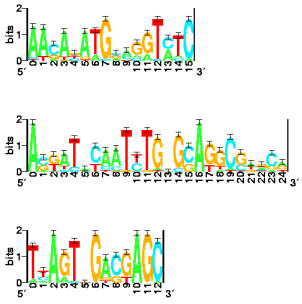
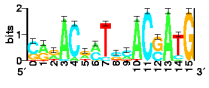
16				
17		ABF1-like		REB1-like
		RRPE-var		
		RRPE		
18		RRPE		
19		STRE		

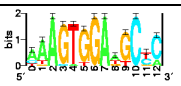
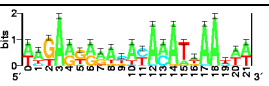
			
20		HAP4	  
21	 	HSF1-like RPN4-like	 
22			 
23			  
24	 	MET4 STRE	    

			
25		PAC	 
26		PAC	
		RRPE-var	
		RRPE	
		PAC-var	
27			    
28		RPN4	
		REB1-like	
29		RRPE	 

		RRPE-var	
30		MBP1	
		MBP1-var	
31		INO4-like	
		SWI4-like	
32		PAC-like	   
33		STRE-var	  
34			
35			 

		    
36	 RAP1-like	  
37		  
38	 YAP1	 
39		  

40		
41		
42		
43		
44		
45	 MCM1	
46	 HSF1	
47		
48		

49		GCR1	

Supplemental References for motifs learned by the network which have been previously documented in the literature:

RAP1: Planta, R.J., et al, Biochem. Cell Biol. (1995); Warner, J. Trends Biochem. Sci. (1999); Lieb, J.D., et al, Nat. Genetics (2001).

RRPE (M3a): Tavazoie, S., et al, Nat. Genetics (1999).

PAC (M3b): Tavazoie, S., et al, Nat. Genetics (1999).

HSF1: Morimoto, R.I., Science (1993).

STRE (MSN2/4): Martinez-Pastor, M.T., et al, EMBO J (1996).

UME6: Strich, R., et al, Genes Dev. (1994).

MET4: Thomas, D. & Surdin-Kerjan,Y. *Microbiol. Mol. Biol. Rev.* (1997).

MCM1: McIrney, C. J., et al, Genes Dev. (1997)

REB1: Morrow, B. E., et al, JBC (1989); Liaw, P.C.Y., et al., Yeast (1994).

RPN4: Mannhaupt, G., et al, FEBS Lett. (1999); Jelinsky, S.A., et al, MCB (2000).

OAF1: Rottensteiner, H., et al, Eur. J. Biochem. (2003).

SWI4: Koch, C, and Naysmith, K. Curr. Opin. Cell Biol. (1994).

PDR3: Delahodde, A., et al, MCB (1995).

MIG1: Lundin, M., et al, MCB (1994).

ABF1: Gray, W.M. and Fassler, J.S, Gene Expr. (1993); Planta, R.J., et al, Biochem. Cell Biol. (1995).

HAP4: Forsburg, S. L., and Guarente, L., Genes Dev. (1989).

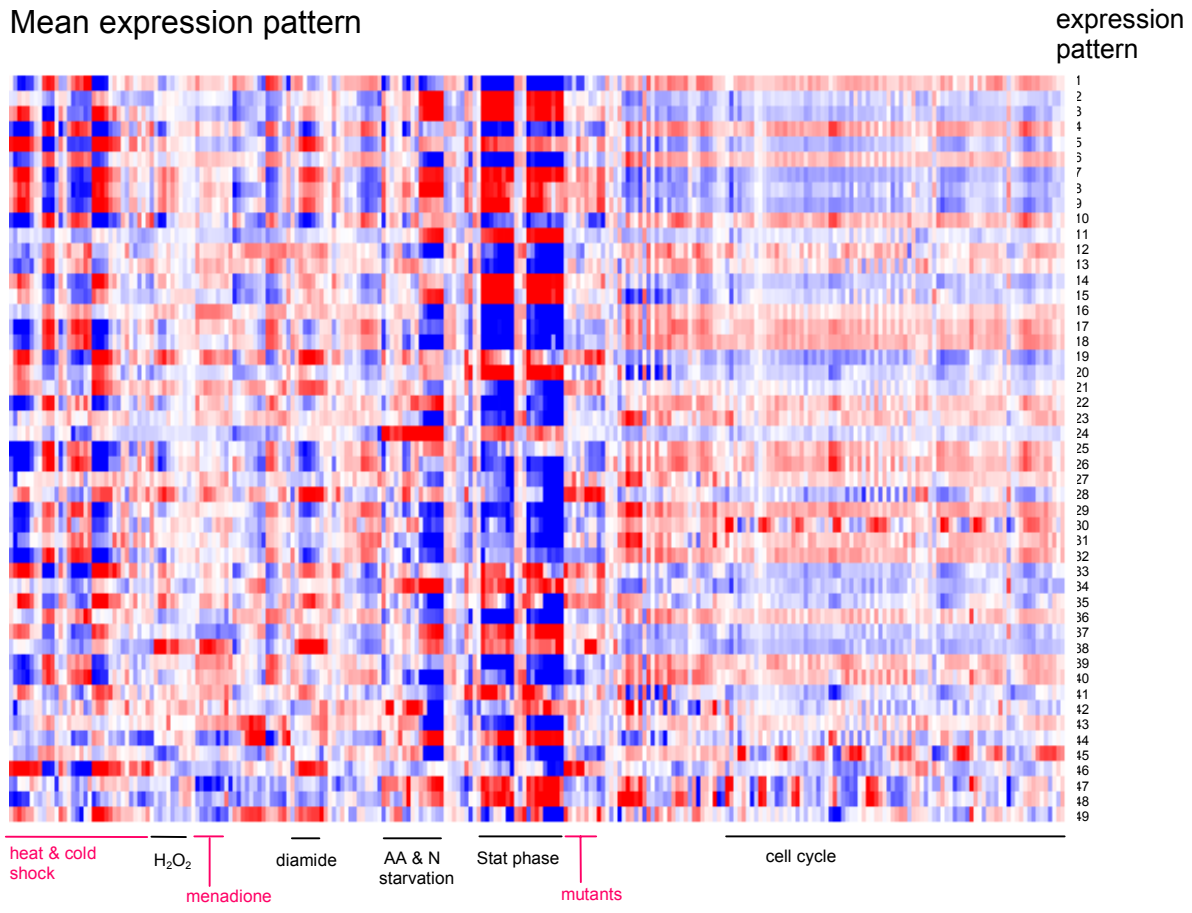
CBF1: Thomas, D. & Surdin-Kerjan,Y. *Microbiol. Mol. Biol. Rev.* (1997).

MBP1: Macintosh, E.M., et al, Curr. Gen. (1993); Koch, C, and Naysmith, K. Curr. Opin. Cell Biol. (1994).

STE12: Baur, M., et al, MCB (1997).

YAP1: Fernandes, L., et al, MCB (1997).

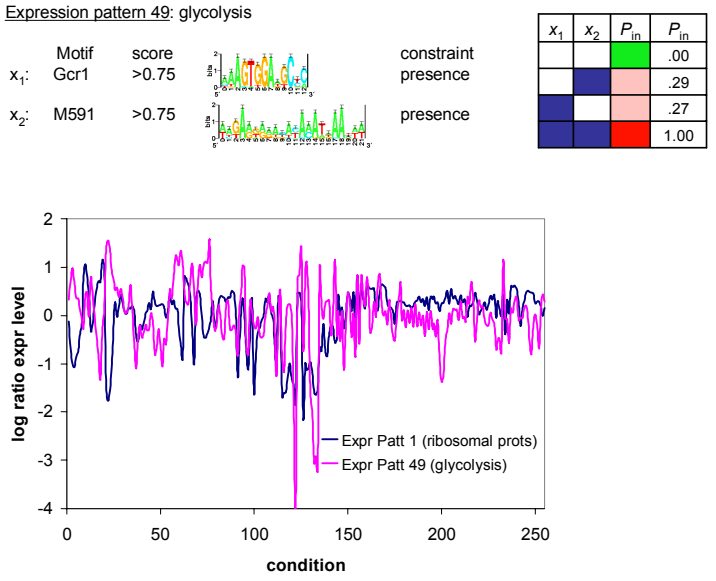
GCR1: Deminoff, S.J., et al, Genetics (2001).



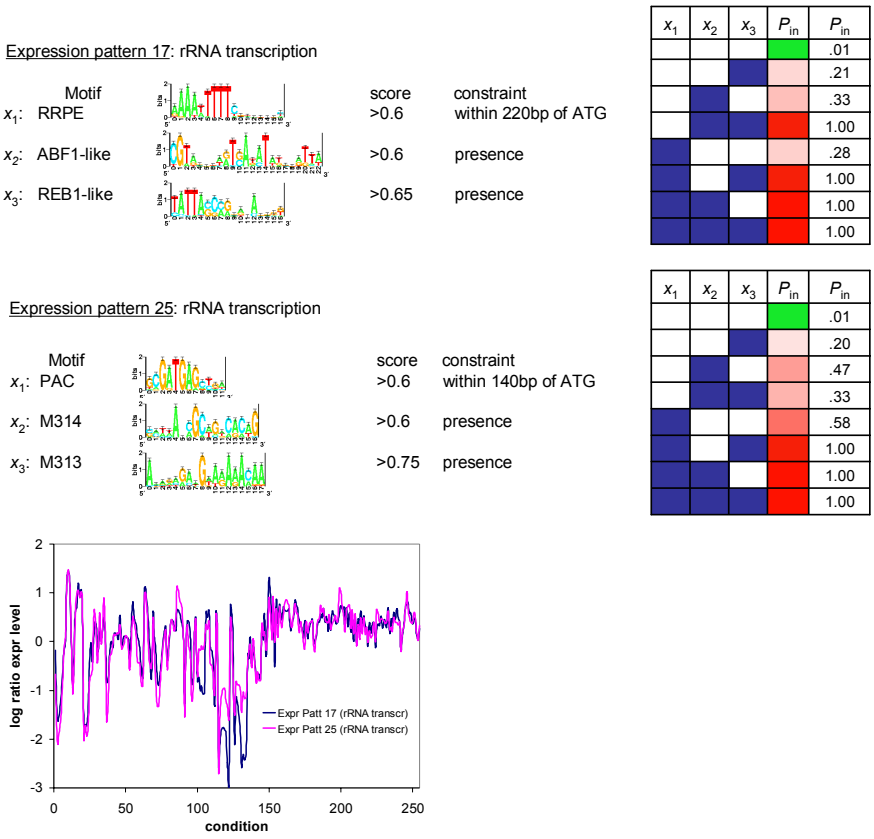
Supplemental Figure 1. Mean expression of all 49 expression patterns. Particularly informative conditions are highlighted. While the two main responses are stress induced and stress repressed, there are subtle differences between expression patterns which can distinguish functionally distinct but similarly expressed sets of genes. For example, expression pattern (1) is almost exclusively ribosomal protein genes, while the similar expression pattern (4) is highly enriched for genes involved in ribosomal RNA transcription and processing. Expression patterns (30), (45), (47) and

(48) are cell cycle regulated genes, but they remain significantly co-regulated across the environmental stress conditions.

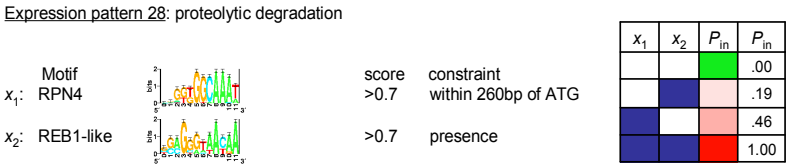
Supplemental Figure 2. Examples of expression patterns explained by the network. Here we select subsets of the constraints found by the full network, to highlight prevalent modes of combinatorial regulation and biologically interesting examples.



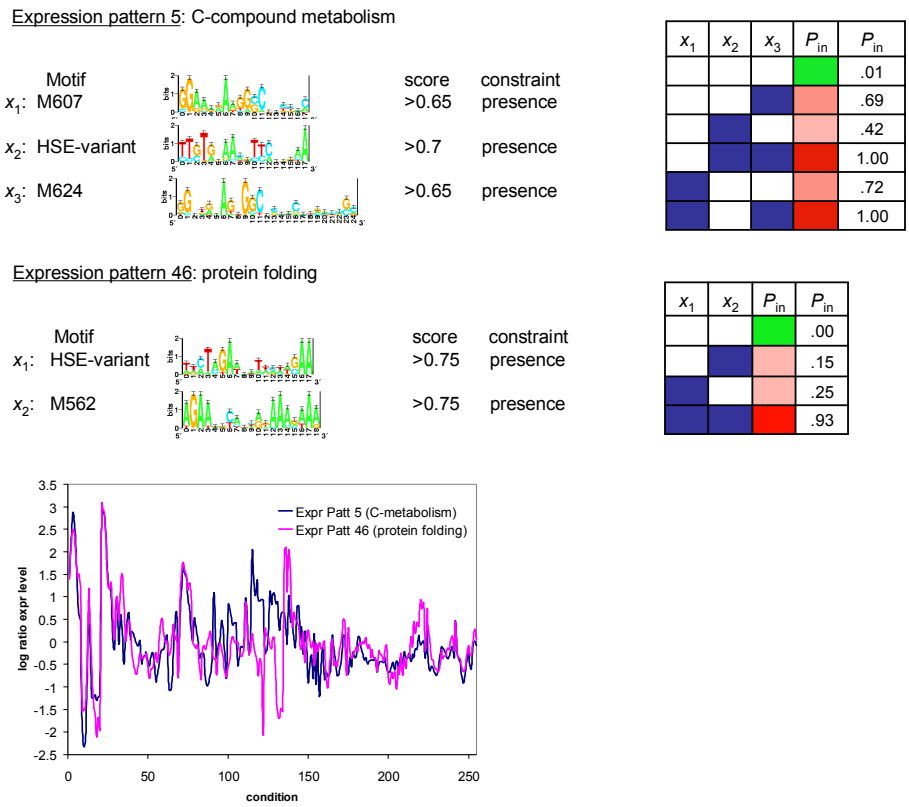
Supplemental Fig 2A. Glycolysis, expression pattern 49. Top: The network for expression pattern 49 learns Gcr1 and a new motif M591. Genes with only one are in the expression pattern 29% and 27% of the time, but genes with both motifs are in the expression pattern 100% of the time. Bottom: Expression pattern of this set compared to the ribosomal protein expression pattern (1). These GCR1 regulated genes are anti-correlated with the ribosomal proteins over many conditions.



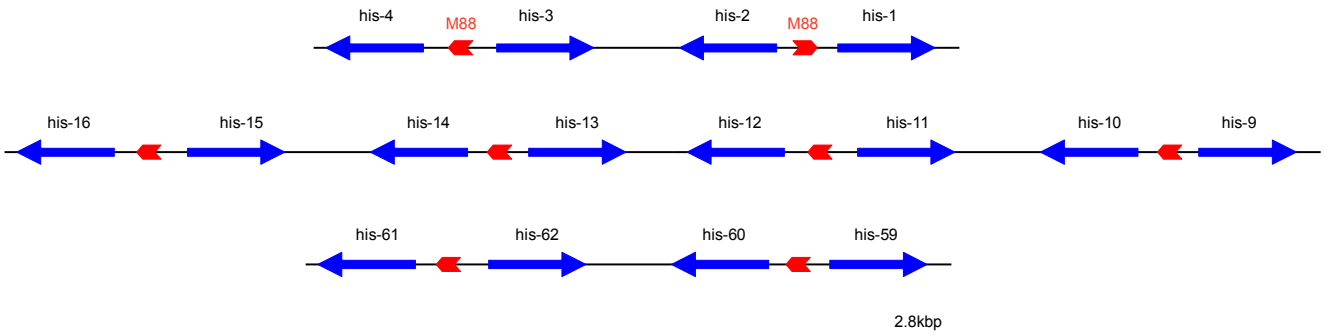
Supplemental Fig 2B. rRNA transcription, expression patterns 17 and 25. These expression patterns are similar to expression pattern 4, the largest rRNA transcription expression pattern (regulated by PAC and RRPE, Fig 3A in main text). But here RRPE and PAC are learned in combination with other motifs. Top: In expression pattern 17, the network learns RRPE with new motifs similar to the known binding sites for ABF1 and REB1. Expression pattern 25 learns PAC in combination with the new motifs M314 or M313. Bottom: Expression pattern 17 and 25 are similar, with subtle differences in the degree of repression under amino acid starvation and stationary phase. Expression pattern 4 falls between these.



Supplemental Fig. 2C. Proteolytic degradation, expression pattern 28. When RPN4 is constrained to be within 260bp of ATG, genes are in the expression pattern 46% of the time. With the additional presence of a REB1-like motif, the fraction is 100%.



Supplemental Fig. 2D. C-compound metabolism and protein folding, expression patterns 5 and 46. These expression patterns are both induced under heat shock but are anti-correlated during stationary phase, where the protein folding genes are repressed but C-compound metabolism genes are induced. Networks for both expression patterns learn variants of the heat shock element (HSE), but with different regulatory partners.



Supplemental Fig 3. Examples of *C. elegans* histone genes selected by the network for expression pattern 15. All histones are selected by motif M88.