Motif search in clusters

We looked for upstream DNA sequence motifs common to members of each cluster found by Adap_Cluster in the data from Cho et al. (1998) (See Figure 3 in our accompanying paper (De Smet et al.) - Pre-processing: removal of 90 and 100-min measurements, selection of the 3000 most variable genes using σ/μ as a metric of variation and normalisation).
We scanned for the motifs described in the paper of Tavazoie et al. (1999). For each motif mentioned in Tavazoie et al. (1999) we attempted to find the consensus sequence in the literature. An overview is given in the table below:

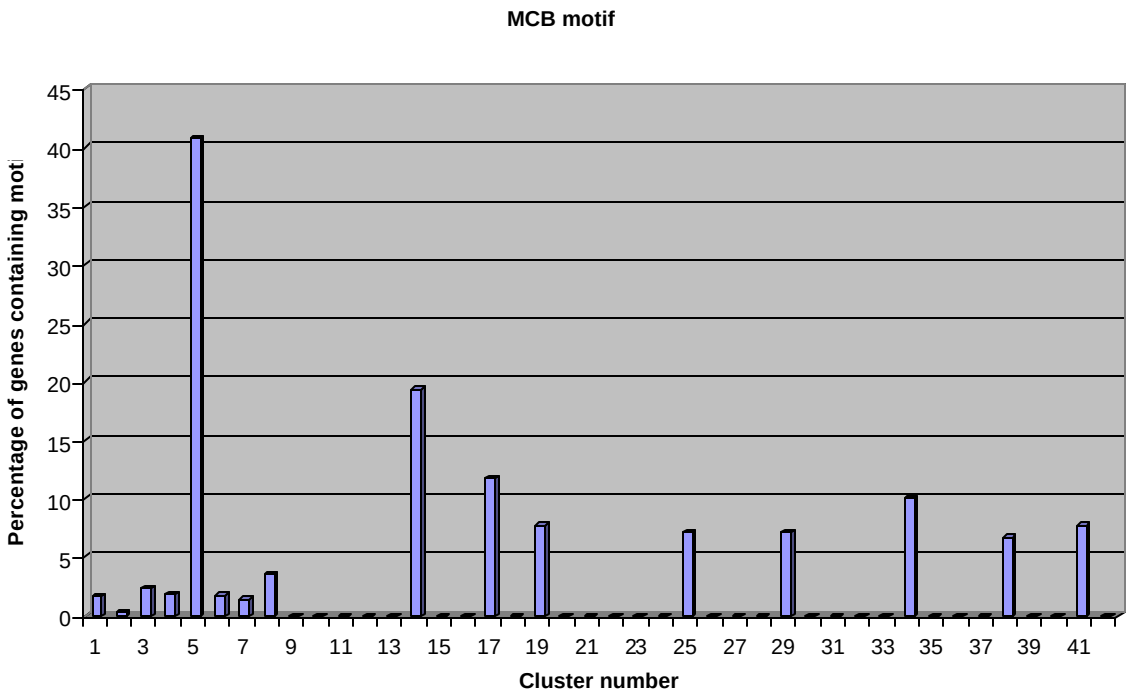
Motif name	Consensus sequence	Reference
MCB	ACGCGTNA	Machado et al. (1997); Dirick et al. (1992)
SCB	CACGAAAA	Dirick et al. (1992)
ECB	TTWCCCNNWAGGAAA	MacKay et al. (2001)
Rap1	ACACCCATACATTT	Nieuwint et al. (1989)
Cbf1p	RTCACRTG	Thomas and Surdin-Kerjan (1997)
STRE	CCCCT	Zahringer et al. (2000)
M3a	CGATGAG	Bussemaker et al. (2001)
Met31/32p	AAACTGTGG	Blaiseau et al. (1997)
M3b	TGAAAA	Bussemaker et al. (2001)

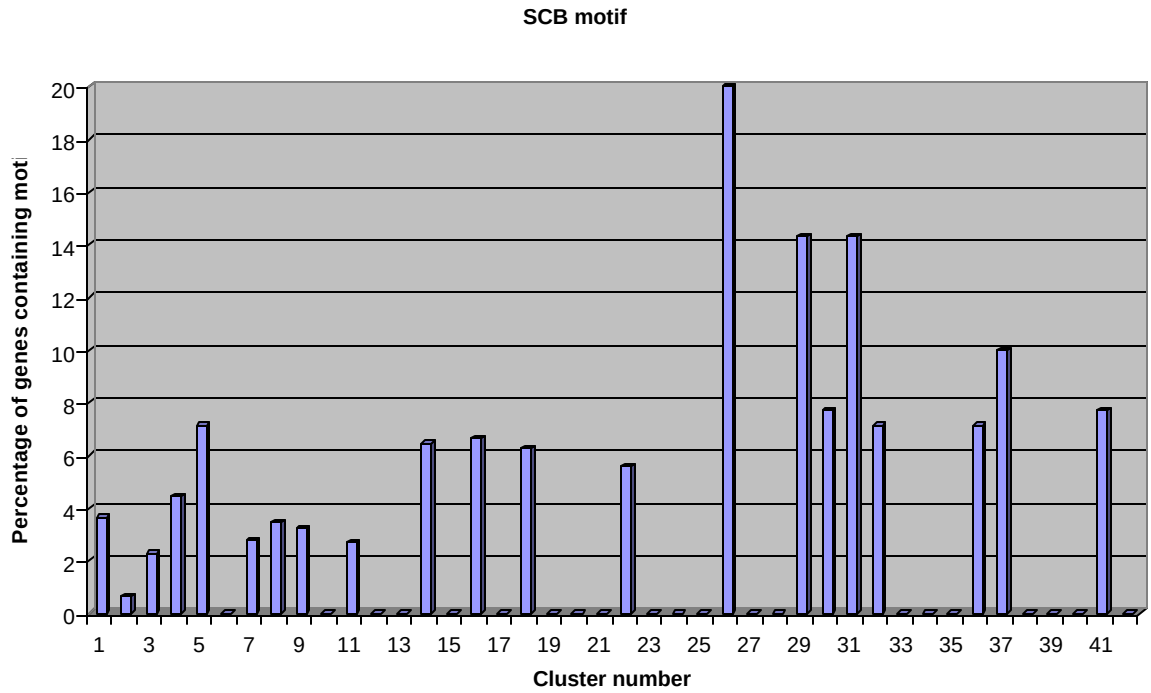
For the other motifs we used the consensus sequence reported in Tavazoie et al. (1999). An overview is given in the table below:

Motif name	Consensus sequence
M1a	TTCYCWYYG
M4	GAARKYTTT
M5	YWKNTCCYNTTT
M13	WWWTTTTTCA
M14a	TTTSGCKTY
M14b	TGTTTRKT
M26	GTKTNNGTGNG
M27	AWATTTTTC

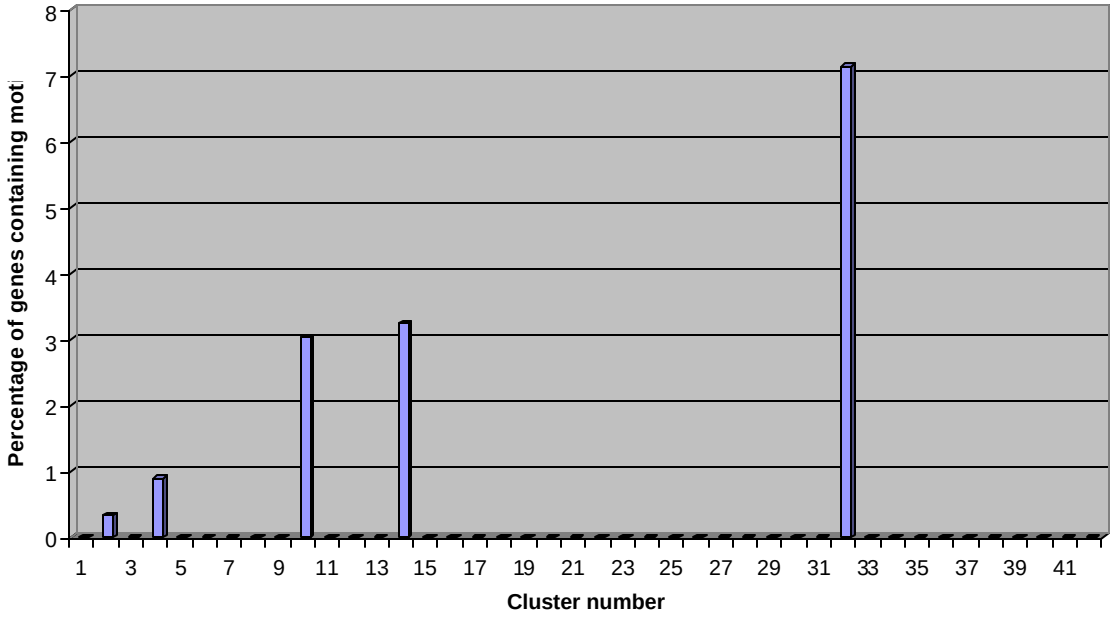
In each cluster we calculated the percentage of genes containing these motifs in their upstream region (maximum 800 bp long). For this purpose we used the Regulatory Sequence Analysis Tools available on the web (http://www.ucmb.ulb.ac.be/bioinformatics/rsa-tools/RSA_home.cgi).
These percentages can be found in the accompanying file. The results for some motifs are represented graphically in Figure 1 (see below).
Note that these results cannot be compared with the results reported by Tavazoie et al. (1999) (motif search in 30 clusters found with the K-means algorithm) because the exact position weight matrices they found are not known.

Fig 1. Percentage of genes in a cluster containing a certain motif in its upstream region.

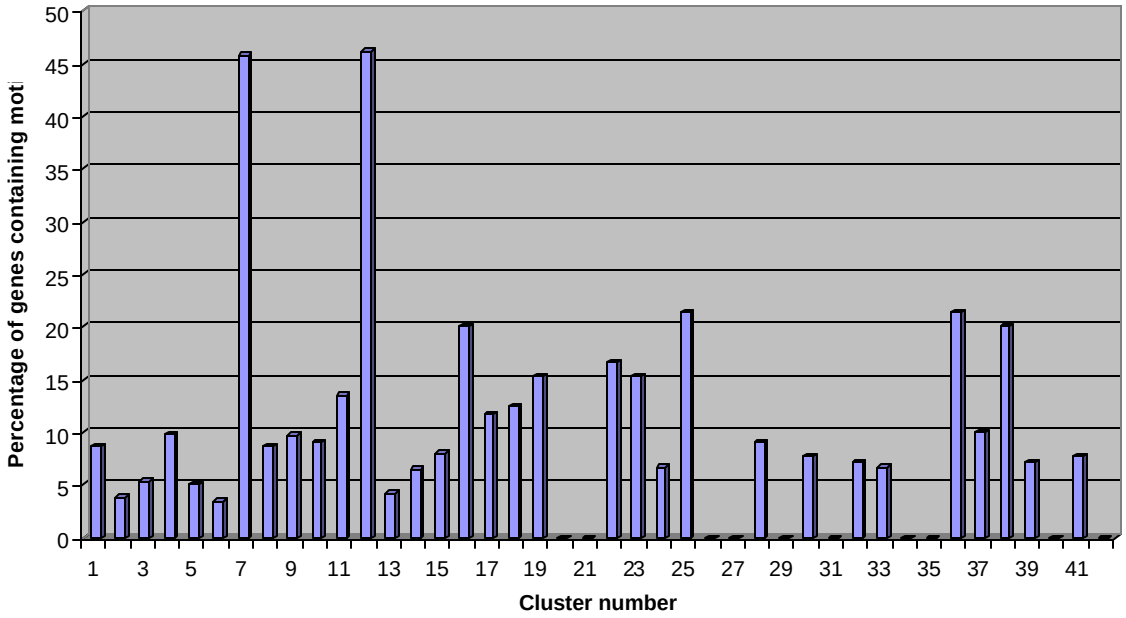




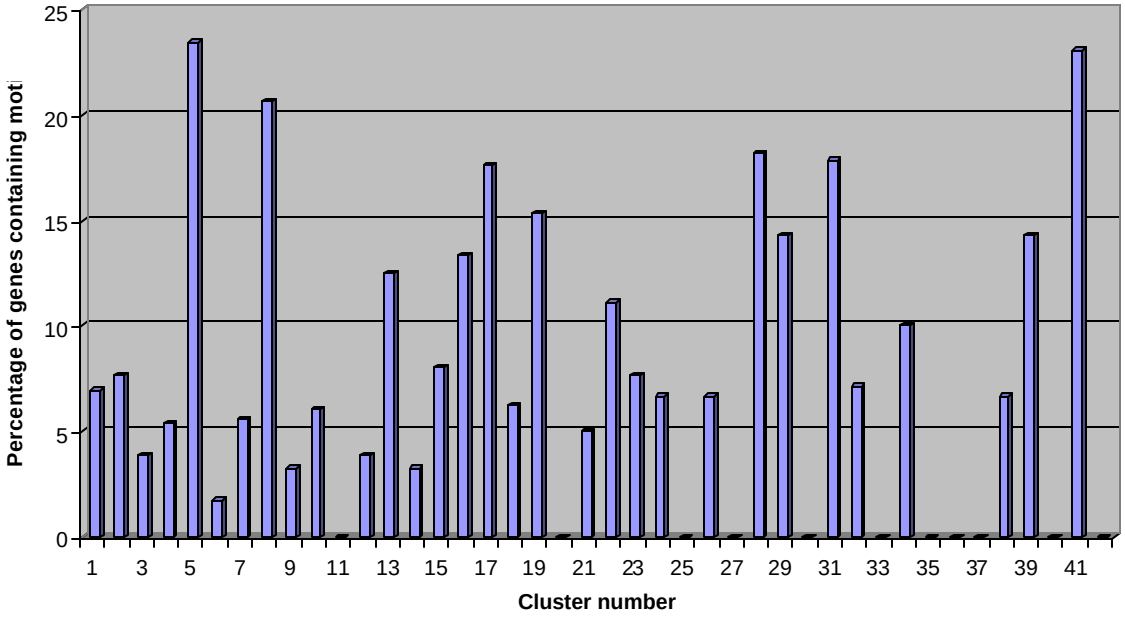
ECB motif



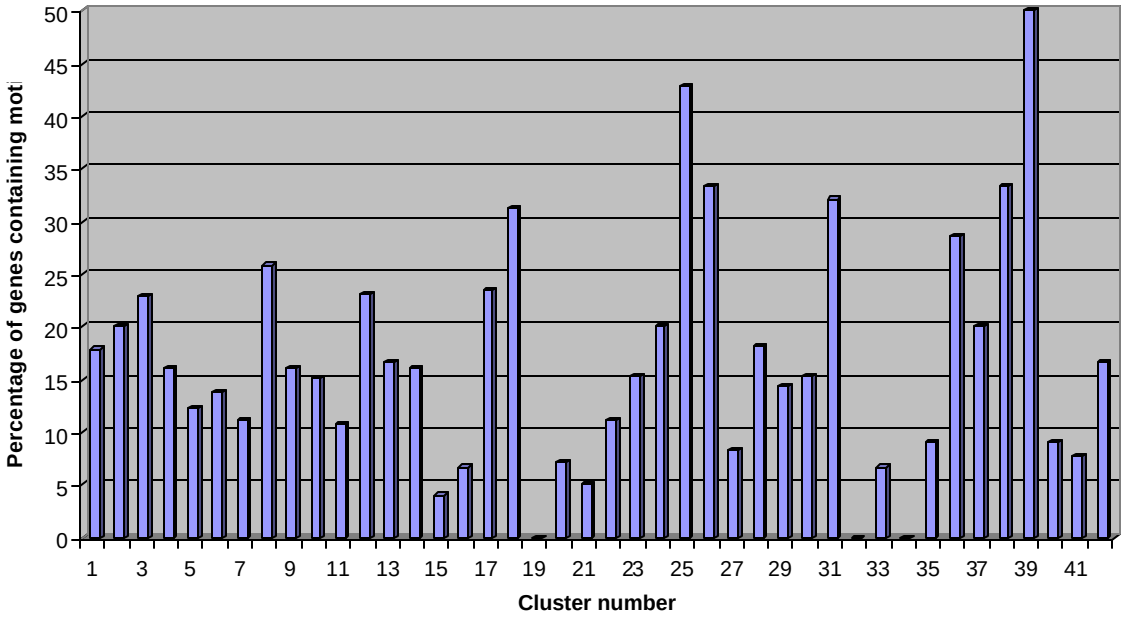
M3a motif



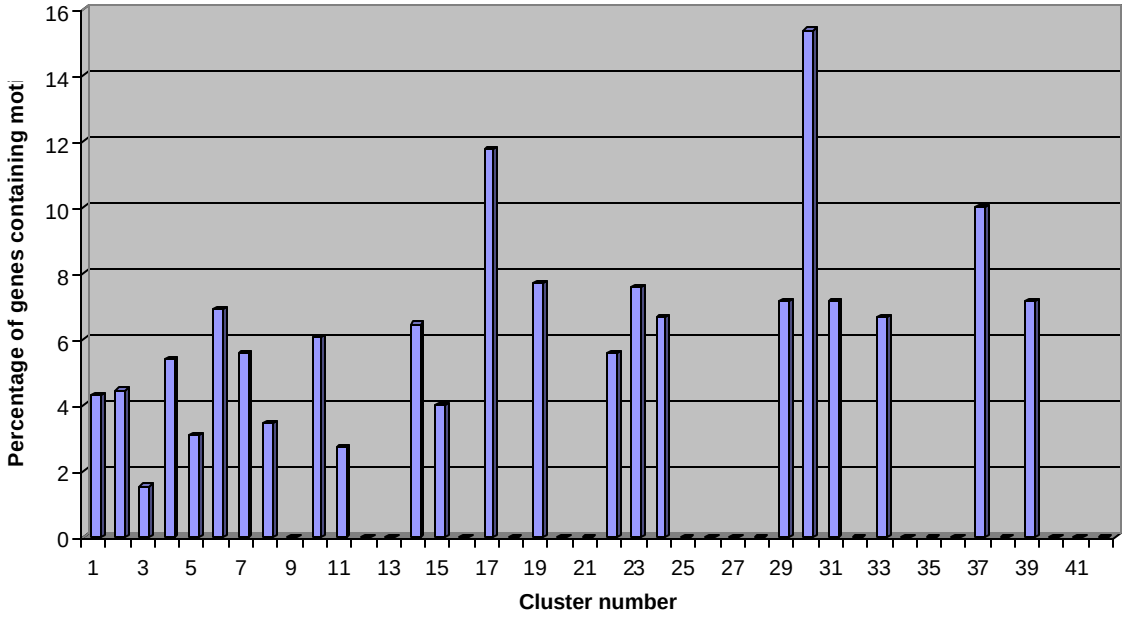
M14a motif



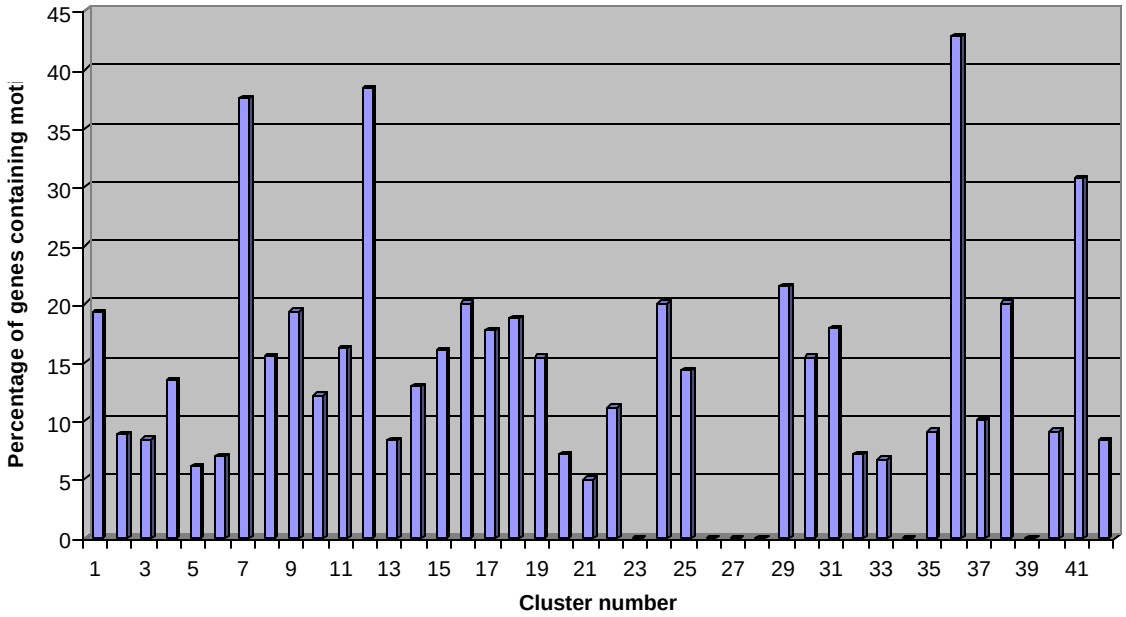
M14b motif



M26 motif



M27 motif



References:

Blaiseau, P.L., Isnard, A.D., Surdin-Kerjan, Y., and Thomas, D. (1997) Met31p and Met32p, two related zinc finger proteins, are involved in transcriptional regulation of yeast sulfur amino acid metabolism. *Mol Cell Biol.*, **17**, 3640-3648.

Bussemaker, H.J., Li, H., and Siggia, E.D. (2001) Regulatory element detection using correlation with expression. *Nat Genet*., **27**, 167-71.

Cho, R.J., Campbell, M.J., Winzeler, E.A., Steinmetz, L., Conway, A., Wodicka, L., Wolfsberg, T.G., Gabrielian, A.E., Landsman, D., Lockhart, D.J. and Davis, R.W. (1998) A genome wide transcriptional analysis of the mitotic cell cycle. *Mol. Cell*, **2**, 65-73.

Dirick, L., Moll, T., Auer, H., and Nasmyth, K. (1992) A central role for SWI6 in modulating cell cycle Start-specific transcription in yeast. *Nature*. **357**, 508-513.

Machado, A.K., Morgan, B.A., and Merrill, G.F. (1997) Thioredoxin reductase-dependent inhibition of MCB cell cycle box activity in Saccharomyces cerevisiae. *J Biol Chem*., **272**, 17045-17054.

MacKay, V.L., Mai, B., Waters, L., and Breeden, L.L. (2001) Early cell cycle box-mediated transcription of CLN3 and SWI4 contributes to the proper timing of the G(1)-to-S transition in budding yeast. *Mol Cell Biol*., **21**, 4140-4148.

Nieuwint, R.T., Mager, W.H., Maurer, K.C., and Planta, R.J. (1989) Mutational analysis of the upstream activation site of yeast ribosomal protein genes. *Curr Genet*., **15**, 247-251.

Tavazoie, S., Hughes, J.D., Campbell, M.J., Cho, R.J., and Church, G.M. (1999) Systematic determination of genetic network architecture. *Nat. Genet*., **22**, 281-285.

Thomas, D. and Surdin-Kerjan, Y. (1997) Metabolism of sulfur amino acids in Saccharomyces cerevisiae. *Microbiol Mol Biol Rev*., **61**, 503-32.

Zahringer, H., Thevelein, J.M., and Nwaka, S.(2000) Induction of neutral trehalase Nth1 by heat and osmotic stress is controlled by STRE elements and Msn2/Msn4 transcription factors: variations of PKA effect during stress and growth. *Mol Microbiol.*, **35**, 397-406.