

**Supplementary Table 5: Potential regulatory promoter motifs**

<b>Motif name</b>	<b>Sequence pattern<sup>a</sup></b>	<b>Associated cluster (<i>P</i> value)/gene list<sup>b</sup></b>	<b>Window size/ gene list size<sup>c</sup></b>	<b>Highest number in window/gene list</b>	<b>Total number in genome</b>	<b>Ratio<sup>d</sup></b>	<b>Binomial probability</b>	<b>Significance threshold</b>
Forkhead	TGTTTAC <sup>e</sup>	1 (<1e-40)	60	31	813	3.2	3.4e-10	1.2e-07
New 1	GTTGNCATG	1 (6.1e-07)	40	6	27	27.6	8.5e-08	8.5e-07
New 2	TTGCATTNC	1 (2.0e-05)	40	8	72	13.8	9.9e-08	8.5e-07
MCB 1	AACGCG	2 (1.6e-27)	60	19	229	6.9	1.4e-11	1.2e-07
MCB 2	CGCGNCGCG	2 (2.0e-19)	60	9	25	44.7	5.0e-13	1.2e-07
MCB 2	CGCGNCGCG	MBF-dep.	32	12	25	74.4	5.5e-20	4.8e-07
Ace2	CCAGCC <sup>e</sup>	2 (3.5e-18)	60	17	124	11.3	8.0e-14	1.2e-07
Ace2	ACCAGCCNT <sup>e</sup>	Ace2p-dep.	42	9	22	48.3	2.6e-13	6.8e-07
Histone	AACNCTAAC <sup>e</sup>	3 (4.5e-15)	40	7	50	24.2	1.5e-07	8.5e-07
New 3	ACCNCGC	4 (5.2e-11)	40	9	104	10.7	1.2e-07	8.5e-07

<sup>a</sup> The most significant of related sequence patterns is shown. N in the sequence pattern means either A, C, G, or T.

<sup>b</sup> The cluster most highly associated with a sequence pattern is indicated together with the significance of the overlap between genes in the cluster and genes with a given motif. Data are also shown for motifs in lists of MBF- and Ace2p-dependent genes.

<sup>c</sup> The sliding window size producing the highest frequency of motif occurrence, or the number of MBF- and Ace2p-dependent genes.

<sup>d</sup> Ratio between highest motif frequency in window or gene list and expected frequency, estimated from frequency in genome (see Methods).

<sup>e</sup> These motifs were also found in reverse orientation, although with lower significance. The presented data refer only to the orientation indicated.