

\*\*\*\*\*  
\*\*

MEME - Motif discovery tool

\*\*\*\*\*  
\*\*

MEME version 3.0 (Release date: 2001/03/05 14:24:28)

For further information on how to interpret these results or to get  
a copy of the MEME software please access <http://meme.sdsc.edu>.

This file may be used as input to the MAST algorithm for searching  
sequence databases for matches to groups of motifs. MAST is available  
for interactive use and downloading at <http://meme.sdsc.edu>.

\*\*\*\*\*  
\*\*

\*\*\*\*\*  
\*\*

#### REFERENCE

\*\*\*\*\*  
\*\*

If you use this program in your research, please cite:

Timothy L. Bailey and Charles Elkan,  
"Fitting a mixture model by expectation maximization to discover  
motifs in biopolymers", Proceedings of the Second International  
Conference on Intelligent Systems for Molecular Biology, pp. 28-36,  
AAAI Press, Menlo Park, California, 1994.

\*\*\*\*\*  
\*\*

\*\*\*\*\*  
\*\*

#### TRAINING SET

\*\*\*\*\*  
\*\*

DATAFILE= /home/tools/hap\_upstreams.txt

ALPHABET= ACGT

Sequence name	Weight	Length	Sequence name	Weight	Length
YGL188C	1.0000	800	CYT1	1.0000	800
COX5A	1.0000	800	YGR182C	1.0000	800
COX13	1.0000	800	COX6	1.0000	800
QCR7	1.0000	800	COX7	1.0000	800
COX12	1.0000	800	COX4	1.0000	800
COX8	1.0000	800	QCR6	1.0000	800
QCR9	1.0000	800	ERG8	1.0000	800
RIP1	1.0000	800	ERG9	1.0000	800
YMR145C	1.0000	800	ACS2	1.0000	800
ERG7	1.0000	800	ERG2	1.0000	800
MVD1	1.0000	800	YEL033W	1.0000	800
ERG13	1.0000	800	CYC1	1.0000	800
ERG26	1.0000	800	ERG3	1.0000	800
ANB1	1.0000	800	HYP2	1.0000	800
ERG10	1.0000	800	ERG1	1.0000	800

ERG11	1.0000	800	YER044C	1.0000	800
ERG5	1.0000	800	CYB5	1.0000	800
YDL086W	1.0000	800	SCM4	1.0000	800
YJL048C	1.0000	800	ERG20	1.0000	800

\*\*\*\*\*  
 \*\*  
 MOTIF 3 width = 11 sites = 38 llr = 301 E-value = 5.4e+004  
 \*\*\*\*\*  
 \*\*

-----  
 --  
 Motif 3 Description  
 -----

--  
 Simplified A ::435:111:1  
 pos.-specific C a:1::11:682  
 probability G :a4:52:2115  
 matrix T ::17:877211

bits	2.4 *
	2.1 **
	1.9 **
	1.7 **
Information	1.4 ** *
content	1.2 ** *
(11.4 bits)	0.9 ** *** *
	0.7 ** ***** *
	0.5 ** *****
	0.2 *****
	0.0 -----

Multilevel CGGTGTTTCCG  
 consensus AAA C  
 sequence

-----  
 --  
 Motif 3 sites sorted by position p-value  
 -----

Sequence name	Start	P-value	Site
CYT1	297	2.20e-07	AATGGGCGGC CGGTATTTCCG GCGGCCAAAA
ERG7	295	3.55e-07	TCGATAAAGT CGATGTTTCCG GATATTGTGT
COX8	622	2.94e-06	GAGAAGTGTT CGGTGTTTACG GAAAAGCTAG
SCM4	554	3.71e-06	CGCGGTTGCG CGATATTTCCC TTGCTTAATG
ERG11	97	3.71e-06	CAGCAAAATA CGAAATTTCCG GCAAAATGTC
ERG5	230	5.85e-06	GTAAAAACTT CGTTATTTCCG CGGCGCGGTT
COX6	532	5.85e-06	GGGCCCCGCG CGTTATTTCCG CCGCTTACAG
YGL188C	379	5.85e-06	TAGAGGGTAT CGATGTTTTCG ACACCAAACC
YER044C	546	9.13e-06	TTAAACTTTA CGGTATTACCG ATAGGAAACT
COX4	231	2.54e-05	TTCGAAAAAG CGATATCTCCC ATCTTGGTTA

COX5A	205	2.54e-05	TCGGAAACAA	CGAAGTTTCCT	CGACTACGAA
ACS2	485	2.89e-05	GGGGCCGAAG	CGTTATTGCCG	ATATTCGGTA
CYC1	442	3.25e-05	ATGTTTTTAC	CGATCTTTCCG	GTCTCTTTGG
ERG13	188	4.56e-05	GTATCGTAGC	CGGTGGTTCTG	GTATCACGCC
ERG26	591	6.35e-05	ATTTTATTGT	CGGTGTTTGTG	GATTTTATAT
ERG2	541	6.35e-05	TTATTATCAA	CGAAGTATCCC	ATTATGATCT
YMR145C	352	6.35e-05	CTGGCCTCAG	CGGTATTTGCA	TTCACAACCA
YEL033W	333	6.93e-05	AAACTGCTGA	CGCTGGTTCCT	CCGCCACCTA
RIP1	604	7.74e-05	TGCGGCGACG	CGGAGGCTCCC	TTATAAGGAC
ERG1	398	9.49e-05	CGCTTTAATG	CGATACTGCCG	TAGCGGGCCT
MVD1	683	9.49e-05	TCCACGATGC	CGAAATCGCCG	AAATGGCTCG
QCR6	512	9.49e-05	GGCGGCCTGC	CGGTACATCCG	TCACCTACAG
ERG3	485	1.04e-04	TATACGAGGC	CGATGGCTGCG	ATAAACGAAA
ERG10	429	1.15e-04	ACATAATCTA	CGATATATCCT	GTAAATAGAA
YJL048C	580	1.26e-04	CTATTCTTTT	CGGAGTTAACG	AAGCCAGGAG
CYB5	463	1.26e-04	GGTAATAGCC	TGGAGTTTCCG	GATGAAAAAT
ERG8	687	1.26e-04	GGTTTTAAGG	CGGAGTTAACG	TCATGTGCGAA
QCR7	308	1.26e-04	AATGAGTCTC	CGGAGTTGACC	AAGTCATACA
HYP2	439	1.38e-04	CCATACCCCT	CGATGTATTCC	GTAGCGTTAT
ERG9	40	1.97e-04	AGCCTCAGTA	CGCTGGTACCC	GTCACAATGT
YGR182C	483	2.31e-04	TCAAACGAAA	CGCTGTTTTTG	TTGGTACTAT
COX13	545	2.72e-04	TGATATGGTG	CGGTATTGACA	CAATATAGTA
YDL086W	393	3.43e-04	GACATTCTGA	CGGTAGTTGGC	AGCTTTCTGC
ERG20	501	3.68e-04	GTATTCTAAG	CGGTATATTCA	CCGTCTCTTA
QCR9	763	4.58e-04	TCATATCCTT	CGCAATTTTCGT	AAAGCAACAA
ANB1	295	5.63e-04	TCATATTCGA	CGATGTCTGTCT	CACACGGAAA
COX12	623	7.76e-04	ATGCGTCGCG	CGTAGTTTTGC	CCGATATCCC
COX7	403	8.24e-04	TTTCTATAGA	CGCTATTTGGA	AACAAGATGT

--

--

Motif 3 block diagrams

--

SEQUENCE NAME	POSITION	P-VALUE	MOTIF DIAGRAM
CYT1		2.2e-07	296_[3]_493
ERG7		3.6e-07	294_[3]_495
COX8		2.9e-06	621_[3]_168
SCM4		3.7e-06	553_[3]_236
ERG11		3.7e-06	96_[3]_693
ERG5		5.8e-06	229_[3]_560
COX6		5.8e-06	531_[3]_258
YGL188C		5.8e-06	378_[3]_411
YER044C		9.1e-06	545_[3]_244
COX4		2.5e-05	230_[3]_559
COX5A		2.5e-05	204_[3]_585
ACS2		2.9e-05	484_[3]_305
CYC1		3.2e-05	441_[3]_348
ERG13		4.6e-05	187_[3]_602
ERG26		6.4e-05	590_[3]_199
ERG2		6.4e-05	540_[3]_249
YMR145C		6.4e-05	351_[3]_438
YEL033W		6.9e-05	332_[3]_457
RIP1		7.7e-05	603_[3]_186

ERG1	9.5e-05	397_[3]_392
MVD1	9.5e-05	682_[3]_107
QCR6	9.5e-05	511_[3]_278
ERG3	0.0001	484_[3]_305
ERG10	0.00011	428_[3]_361
YJL048C	0.00013	579_[3]_210
CYB5	0.00013	462_[3]_327
ERG8	0.00013	686_[3]_103
QCR7	0.00013	307_[3]_482
HYP2	0.00014	438_[3]_351
ERG9	0.0002	39_[3]_750
YGR182C	0.00023	482_[3]_307
COX13	0.00027	544_[3]_245
YDL086W	0.00034	392_[3]_397
ERG20	0.00037	500_[3]_289
QCR9	0.00046	762_[3]_27
ANB1	0.00056	294_[3]_495
COX12	0.00078	622_[3]_167
COX7	0.00082	402_[3]_387

--

--

Motif 3 in BLOCKS format

--

BL MOTIF 3 width=11 seqs=38

CYT1	( 297)	CGGTATTTCCG	1
ERG7	( 295)	CGATGTTTCCG	1
COX8	( 622)	CGGTGTTTACG	1
SCM4	( 554)	CGATATTTCCC	1
ERG11	( 97)	CGAAATTTCCG	1
ERG5	( 230)	CGTTATTTCCG	1
COX6	( 532)	CGTTATTTCCG	1
YGL188C	( 379)	CGATGTTTTTCG	1
YER044C	( 546)	CGGTATTACCG	1
COX4	( 231)	CGATATCTCCC	1
COX5A	( 205)	CGAAGTTTCCT	1
ACS2	( 485)	CGTTATTGCCG	1
CYC1	( 442)	CGATCTTTCCG	1
ERG13	( 188)	CGGTGGTTCTG	1
ERG26	( 591)	CGGTGTTTGTG	1
ERG2	( 541)	CGAAGTATCCC	1
YMR145C	( 352)	CGGTATTTGCA	1
YEL033W	( 333)	CGCTGGTTCCCT	1
RIP1	( 604)	CGGAGGCTCCC	1
ERG1	( 398)	CGATACTGCCG	1
MVD1	( 683)	CGAAATCGCCG	1
QCR6	( 512)	CGGTACATCCG	1
ERG3	( 485)	CGATGGCTGCG	1
ERG10	( 429)	CGATATATCCT	1
YJL048C	( 580)	CGGAGTTAACG	1
CYB5	( 463)	TGGAGTTTCCG	1
ERG8	( 687)	CGGAGTTAACG	1
QCR7	( 308)	CGGAGTTGACC	1
HYP2	( 439)	CGATGTATTCC	1

```

ERG9          (  40) CGCTGGTACCC  1
YGR182C      ( 483) CGCTGTTTTTG  1
COX13        ( 545) CGGTATTGACA  1
YDL086W      ( 393) CGGTAGTTGGC  1
ERG20        ( 501) CGGTATATTCA  1
QCR9         ( 763) CGCAATTTTCGT  1
ANB1         ( 295) CGATGTCGTCT  1
COX12        ( 623) CGTAGTTTTGC  1
COX7         ( 403) CGCTATTTGGA  1
//

```

```

-----
--
-----
--

```

Motif 3 position-specific scoring matrix

```

-----
--
log-odds matrix: alength= 4 w= 11 n= 30020 bayes= 9.62388 E= 5.4e+004
-1189   227  -1189   -347
-1189  -1189   237  -1189
   24   -62   103   -148
  -11  -1189  -1189   128
   60  -294   137  -1189
-1189  -194   -29   143
  -125  -62  -1189   133
  -157 -1189   -29   133
  -125  152   -55   -89
-1189   202   -87  -189
  -157   23   145  -116
-----
--

```

Motif 3 position-specific probability matrix

```

-----
--
letter-probability matrix: alength= 4 w= 11 n= 30020 E= 5.4e+004
0.000082  0.973481  0.000051  0.026386
0.000082  0.000053  0.999788  0.000077
0.368406  0.131597  0.394684  0.105313
0.289480  0.000053  0.000051  0.710417
0.473642  0.026362  0.499919  0.000077
0.000082  0.052671  0.157904  0.789343
0.131626  0.131597  0.000051  0.736725
0.105318  0.000053  0.157904  0.736725
0.131626  0.578848  0.131595  0.157930
0.000082  0.815628  0.105286  0.079004
0.105318  0.236833  0.526228  0.131621
-----
--

```

Time 663.58 secs.

\*\*\*\*\*  
\*\*

\*\*\*\*\*  
\*\*

SUMMARY OF MOTIFS

\*\*\*\*\*  
\*\*

-----  
--  
          Combined block diagrams: non-overlapping sites with p-value < 0.0001  
-----

-----

SEQUENCE NAME	COMBINED P-VALUE	MOTIF DIAGRAM
YGL188C 06)]_156	3.36e-04	378_[3(5.85e-06)]_243_[2(6.70e-
CYT1 05)]_224	2.23e-05	296_[3(2.20e-07)]_257_[1(1.05e-
COX5A 05)]_89	3.02e-03	204_[3(2.54e-05)]_484_[1(1.89e-
YGR182C 05)]_273	1.06e-02	206_[1(3.74e-05)]_297_[2(6.46e-
COX13	1.74e-03	379_[2(2.26e-05)]_397_[1(9.62e-06)]
COX6 06)]_258	1.31e-03	73_[2(2.54e-05)]_446_[3(5.85e-
QCR7	1.01e-02	443_[1(7.34e-05)]_333_[1(1.17e-05)]
COX7 06)]_175_[1(2.25e-06)]_174	2.89e-04	316_[2(2.82e-05)]_99_[2(4.58e-
COX12 05)]_36	2.47e-02	492_[1(7.34e-05)]_248_[2(3.57e-
COX4 05)]_234_[1(5.24e-05)]_243	2.17e-03	230_[3(2.54e-05)]_58_[2(5.42e-
COX8 06)]_168	8.99e-04	575_[2(5.42e-05)]_34_[3(2.94e-
QCR6 05)]_5	2.20e-02	511_[3(9.49e-05)]_261_[1(3.26e-
QCR9	4.52e-02	778_[1(3.74e-05)]_10
ERG8	3.05e-02	734_[1(2.39e-05)]_54
RIP1 05)]_173_[1(2.30e-05)]_1	1.61e-03	521_[2(2.82e-05)]_70_[3(7.74e-
ERG9 05)]_52	4.42e-03	602_[2(2.82e-05)]_122_[1(3.26e-
YMR145C 05)]_80_[2(5.42e-05)]_250_[1(1.89e-05)]_84	1.99e-03	272_[1(8.11e-05)]_67_[3(6.35e-
ACS2 06)]_269	6.90e-04	484_[3(2.89e-05)]_24_[2(2.08e-
ERG7 06)]_118	6.85e-06	294_[3(3.55e-07)]_365_[1(2.25e-
ERG2 06)]_62_[3(6.35e-05)]_249	4.22e-04	375_[2(7.18e-05)]_79_[1(2.25e-

MVD1	1.32e-02	371_[2(1.23e-05)]_299_[3(9.49e-05)]_107
YEL033W	8.37e-03	209_[2(1.06e-05)]_111_[3(6.93e-05)]_457
ERG13	4.06e-02	187_[3(4.56e-05)]_602
CYC1	4.69e-05	49_[1(8.11e-05)]_107_[1(2.25e-06)]_36_[1(8.11e-05)]_20_[1(2.25e-06)]_181_[3(3.25e-05)]_244_[2(1.06e-05)]_92
ERG26	6.04e-04	81_[2(4.02e-05)]_466_[2(2.08e-06)]_19_[3(6.35e-05)]_199
ERG3	4.59e-04	33_[1(2.25e-06)]_186_[2(4.85e-05)]_557
ANB1	1.10e-03	361_[2(5.42e-05)]_6_[2(2.08e-06)]_28_[2(2.08e-06)]_123_[2(1.23e-05)]_194_[1(3.26e-05)]_28
HYP2	3.91e-04	3_[1(3.45e-05)]_16_[1(6.47e-06)]_669_[2(1.06e-05)]_76
ERG10	3.13e-03	102_[1(1.17e-05)]_255_[2(8.68e-05)]_419
ERG1	4.44e-05	397_[3(9.49e-05)]_193_[2(4.54e-07)]_11_[2(6.46e-05)]_30_[2(2.08e-06)]_82_[1(1.74e-05)]_28
ERG11	2.65e-05	96_[3(3.71e-06)]_485_[2(2.08e-06)]_2_[2(5.92e-05)]_156_[1(5.24e-05)]_14
YER044C	4.82e-04	545_[3(9.13e-06)]_52_[1(6.47e-06)]_180
ERG5	2.12e-04	36_[1(1.89e-05)]_181_[3(5.85e-06)]_134_[3(9.13e-06)]_318_[2(4.02e-05)]_85
CYB5	8.42e-03	495_[1(1.89e-05)]_293
YDL086W	3.47e-03	450_[2(3.09e-06)]_338
SCM4	1.78e-05	211_[2(1.06e-05)]_94_[1(8.79e-05)]_224_[3(3.71e-06)]_59_[1(6.47e-06)]_165
YJL048C	2.54e-02	13_[2(5.92e-05)]_775
ERG20	8.46e-04	198_[2(3.12e-05)]_188_[1(2.25e-06)]_155_[2(8.68e-05)]_223

-----  
 --

\*\*\*\*\*  
 \*\*

\*\*\*\*\*  
 \*\*

Stopped because nmotifs = 3 reached.

\*\*\*\*\*  
 \*\*

CPU: rana.lbl.gov

\*\*\*\*\*  
 \*\*