MAST

Motif Alignment & Search Tool

For further information on how to interpret these results https://meme-suite.org/meme/doc/mast.html.

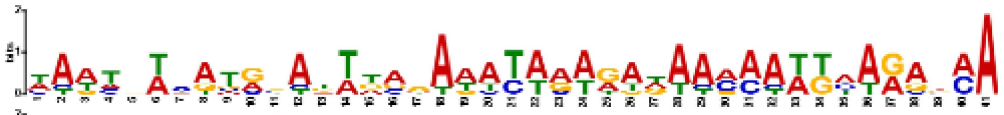
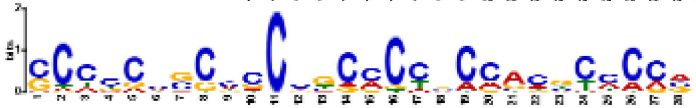
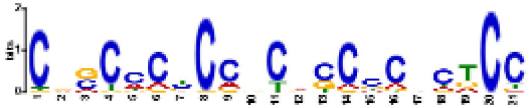
To get a copy of the MEME software please access https://meme-suite.org.

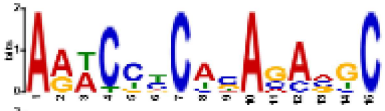
If you use MAST in your research, please cite the following paper:
Timothy L. Bailey and Michael Gribskov, "Combining evidence using p-values: application to sequence homology searches", *Bioinformatics*, **14**(1):48-54, 1998. [full text]

MOTIFS

Next Top

Motifs with a pale red background are very similar to other earlier specified motifs and may be biasing the results.
It is recommended that you re-run MAST and request it to remove redundant motifs.

	Logo	Name	Alt. Name
1.		HAWTNWBATRVAKWDAVAWAYWRWRADAAMMAWKHARADMA	MEME-1
2.		SCCCCMSCCCCCGCMCYSCCACKCMCCV	MEME-2
3.		WKTVTRTBDTCAYATTATARBTRAKGMA	MEME-3
4.		AMAYWAAATAWATRWAMTC	MEME-4
5.		TTTCWHRTYTT	MEME-5
6.		CBSCCCBCCNCHCCCCNCTCC	MEME-6

	Logo	Name	Alt. Name
7.		AAAWRTTYKTGWTWG	MEME-7
8.		TTAAGTWSAGAA	MEME-8
9.		ATAMAAAGSTCAVTTCCMAWGATBNT	MEME-9
10.		TWWTWTAGMAGWAGCAARA	MEME-10
11.		CCCGCSCCCGCCCSG	MEME-11
12.		ARWCCYCACASAVGC	MEME-12
13.		CBMCTCTCAGGAA	MEME-13
14.		TCWTCTAAGAA	MEME-14
15.		TGTGAAAT	MEME-15
16.		ATCTAGARAA	MEME-16
17.		ATAAWAT	MEME-17
18.		CTGCACCACGAYCAC	MEME-18

	Logo	Name	Alt. Name
19.		AGASACATTCA	MEME-19
20.		TTMTTCCTCTYAMA	MEME-20
21.		AAATGA	MEME-21
22.		AACTSATGAA	MEME-22
23.		TTTGCACTGCT	MEME-23
24.		AGCTATTKTT	MEME-24
25.		TTTWGAAT	MEME-25
26.		AAWCTTGAGCA	MEME-26
27.		AATAWCCTT	MEME-27
28.		ATRAMGTAAT	MEME-28
29.		TWTTAWTT	MEME-29
30.		TGAAATGT	MEME-30

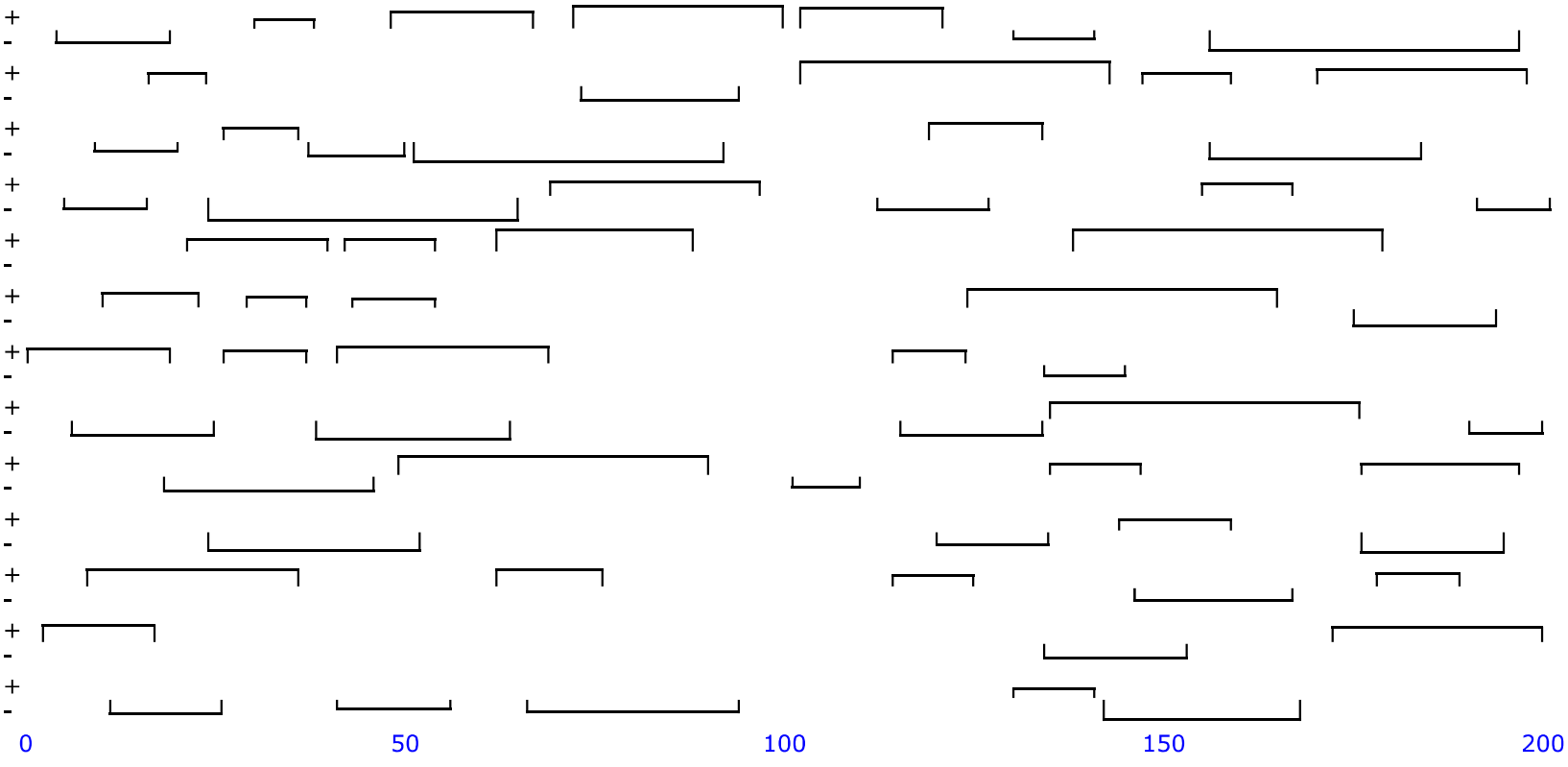
Search Results

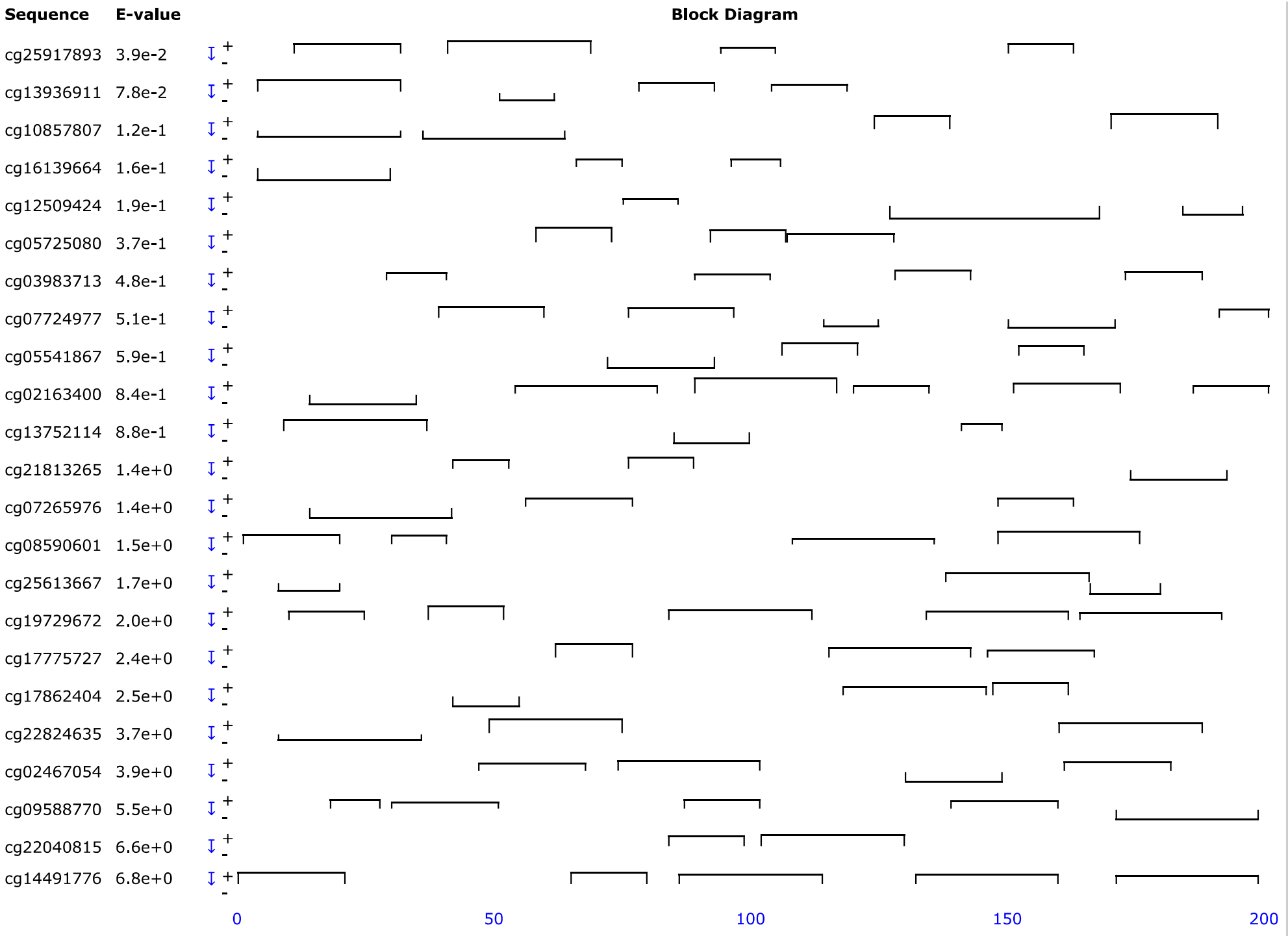
Top Scoring Sequences

Each of the following 37 sequences has an *E*-value less than 10.
The motif matches shown have a position *p*-value less than 0.0001.
Hover the cursor over the sequence name to view more information about a sequence.
Hover the cursor over a motif for more information about the match.
Click on the arrow (↧) next to the *E*-value to see the sequence surrounding each match.

☐ HAWTNWBATRVAKWDAVAWAYWRWRADAAMMAWKHARADMA ☐ SCCCCMSCCCCCGCMCYSCCACKCMCCV ☐ WKTVTRTBDTCAYATTATARBTRAKGMA
☐ AMAYWAAATAWATRWAMTC ☐ TTTCWHRTYTT ☐ CBSCCCBCCNCHCCCCNCTCC ☐ AAAWRTTYKTGWTWG ☐ TTAAGTWSAGAA
☐ ATAMAAAGSTCAVTTCCMAWGATBNT ☐ TWWTWTAGMAGWAGCAARA ☐ CCCGCSCCCGCCCSG ☐ ARWCCYCACASAVGC ☐ CBMCTCTCAGGAA
☐ TCWTCTAAGAA ☐ TGTGAAAT ☐ ATCTAGARAA ☐ ATAAWAT ☐ CTGCACCACGAYCAC ☐ AGASACATTCA ☐ TTMTTCCTCTYAMA ☐ AAATGA
☐ AACTSATGAA ☐ TTTGCACTGCT ☐ AGCTATTKTT ☐ TTTWGAAT ☐ AAWCTTGAGCA ☐ AATAWCCTT ☐ ATRAMGTAAT ☐ TWTTAWTT ☐ TGAAATGT

Sequence	E-value	Block Diagram
cg00957886	8.6e-17	
cg02632822	1.6e-13	
cg04156077	1.3e-10	
cg17040924	2.1e-10	
cg02573152	3.1e-10	
cg07747690	1.3e-6	
cg00896220	2.4e-6	
cg24872610	2.9e-6	
cg19931348	4.8e-5	
cg12500857	2.3e-4	
cg10637260	3.9e-3	
cg20667822	6.6e-3	
cg03532422	1.5e-2	




Sequence
E-value
Block Diagram
cg25917893 3.9e-2
cg13936911 7.8e-2
cg10857807 1.2e-1
cg16139664 1.6e-1
cg12509424 1.9e-1
cg05725080 3.7e-1
cg03983713 4.8e-1
cg07724977 5.1e-1
cg05541867 5.9e-1
cg02163400 8.4e-1
cg13752114 8.8e-1
cg21813265 1.4e+0
cg07265976 1.4e+0
cg08590601 1.5e+0
cg25613667 1.7e+0
cg19729672 2.0e+0
cg17775727 2.4e+0
cg17862404 2.5e+0
cg22824635 3.7e+0
cg02467054 3.9e+0
cg09588770 5.5e+0
cg22040815 6.6e+0
cg14491776 6.8e+0
0
50
100
150
200

Sequence	E-value	Block Diagram
cg09368670	7.4e+0	

HAWTNWBATRVAKWDAVAWAYWRWRADAAMMAWKHARADMA SCCCCMSCCCCCGCMCYSCCACKCMCCV WKTVTRTBDTCAYATTATARBTRAKGMA
AMAYWAAATAWATRWAMTC TTTCWHRTYTT CBSCCCBCCNCHCCCCNCTCC AAAWRTTYKTGWTWG TTAAGTWSAGAA
ATAMAAAGSTCAVTTCCMAWGATBNT TWWTWTAGMAGWAGCAARA CCCGCSCCCGCCCSG ARWCCYCACASAVGC CBMCTCTCAGGAA
TCWTCTAAGAA TGTGAAAT ATCTAGARAA ATAAWAT CTGCACCACGAYCAC AGASACATTCA TTMTTCCTCTYAMA AAATGA
AACTSATGAA TTTGCACTGCT AGCTATTKTT TTTWGAAT AAWCTTGAGCA AATAWCCTT ATRAMGTAAT TWTTAWTT TGAAATGT

Inputs & Settings

Prev Next Top

Alphabet

Background Source: an old version of the NCBI non-redundant database

Name	Bg.				Bg.	Name
Adenine	0.274262	A	~	T	0.274262	Thymine
Cytosine	0.22521	C	~	G	0.22521	Guanine

Sequences

The following sequence database was supplied to MAST.

Database	Sequence Count	Residue Count	Last Modified
dmp_cpg_sequences.fasta	75	15075	Fri Jun 20 07:54:51 2025
Total	75	15075	

Motifs

The following motif database was supplied to MAST.

Database	Last Modified
meme.xml	Fri Jun 20 08:03:00 2025

Other Settings

Strand Handling	The result of scanning both strands is **combined**. When matches overlap the non-overlapping combination with the best p-value is shown.
Max Correlation	Motifs with a correlation greater than **0.6** are marked for potential removal dependant on the `--remcorr` option.
Remove Correlated	Correlated motifs exceeding the threshold are **highlighted** and their removal is recommended.
Max Sequence *E*-value	Sequences with an *E*-value less than **10** are included in the output.
Adjust Hit *p*-value	The hit *p*-value is **not adjusted** for the length of the sequence.
Displayed Hits	The *p*-value of a hit must be less than **0.0001** to be shown in the output.
Displayed Weak Hits	Weak hits are **not displayed**.

Prev Top

MAST version 5.5.8
(Release date: Thu May 15 15:01:46 2025 -0700)

Reference
Timothy L. Bailey and Michael Gribskov, "Combining evidence using p-values: application to sequence homology searches", *Bioinformatics*, **14**(1):48-54, 1998. [full text]

Command line summary

```
mast -oc . -nostatus meme.xml dmp_cpg_sequences.fasta
```

Ran in 0.08 seconds on Fri Jun 20 08:03:00 2025 on noble-meme.grid.gs.washington.edu