



# MOTIFSIM - MOTIF SIMilarity Detection Tool

Version 2.2

## INPUT

### Input Parameters

Number of files:	5
Number of top significant motifs:	5
Number of best matches:	5
Similarity cutoff $\geq$	0.75
Matching motif database:	Jaspar Urochordates
Motif tree:	Yes
Combined similar motifs:	Yes
Output file type:	All
Output file format:	All

### Input files and motif counts

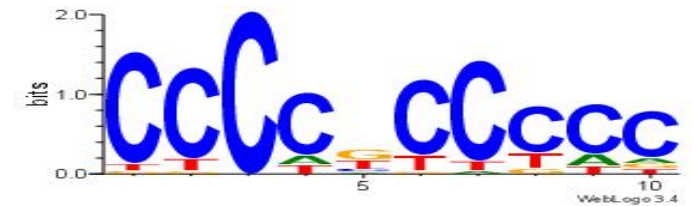
File name	Count of motifs	Dataset number
DREME_DM230.txt	1	1
MEME_DM230.txt	20	2
PScanChIP_DM230.txt	14	3
RSAT_peak-motifs_DM230.txt	10	4
W-ChIPMotifs_DM230.txt	11	5

## RESULTS

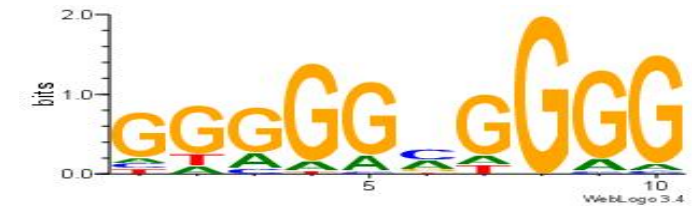
### Top 5 Significant Motifs - Global Matching (Highest to Lowest)

Dataset #: 3      Motif ID: 24      Motif name: SP1

Original motif      Consensus sequence: CCCCKCCCC



Reverse complement motif      Consensus sequence: GGGGGYGGGG



### Best Matches for Top Significant Motif ID 24 (Highest to Lowest)

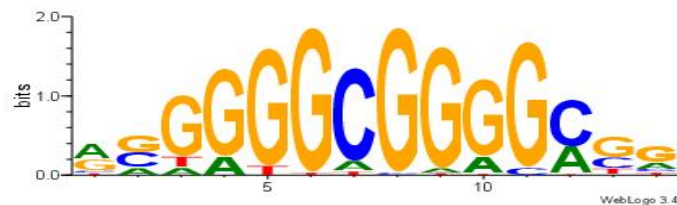
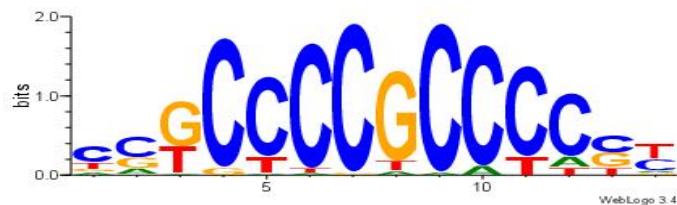
Dataset #:	2
Motif ID:	7
Motif name:	Motif 7
Matching format of first motif:	Original Motif
Matching format of second motif:	Original Motif
Direction:	Backward
Position number:	2
Number of overlap:	10
Similarity score:	0.00451594

Alignment:

```
CSKCCCCGCCCSY  
---CCCCKCCCC--
```

Original motif      Consensus sequence: CSKCCCCGCCCSY

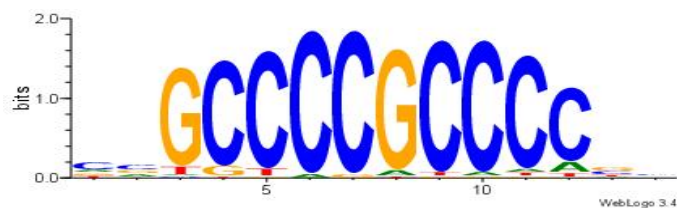
Reverse complement motif      Consensus sequence: MSGGGCGGGGY



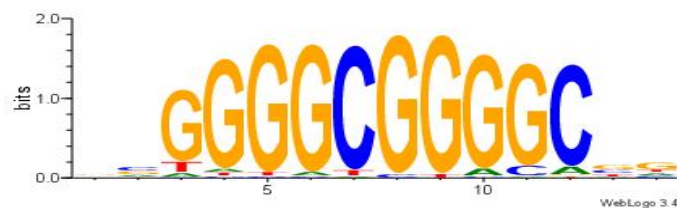
Dataset #: 4  
 Motif ID: 36  
 Motif name: csGCCCCGCCCCsc  
 Matching format of first motif: Original Motif  
 Matching format of second motif: Original Motif  
 Direction: Forward  
 Position number: 4  
 Number of overlap: 10  
 Similarity score: 0.00965796

Alignment:  
 HVGCCCCGCCCCBB  
 ---CCCCKCCCC-

Original motif    Consensus sequence: HVGCCCCGCCCCBB



Reverse complement motif    Consensus sequence: BBGGGCGGGGc

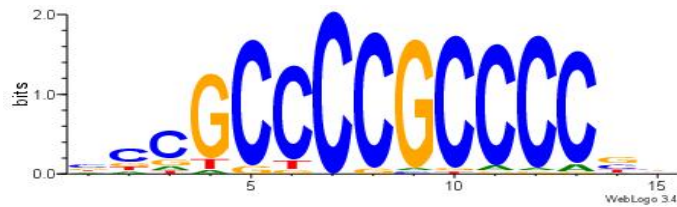


Dataset #: 4  
 Motif ID: 38

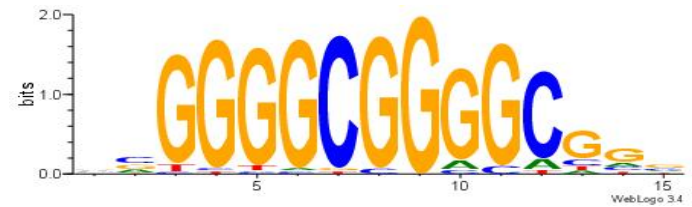
Motif name: cccGCCCCGCCCCsb  
 Matching format of first motif: Original Motif  
 Matching format of second motif: Original Motif  
 Direction: Forward  
 Position number: 5  
 Number of overlap: 10  
 Similarity score: 0.0127292

Alignment:  
 BCCGCCCCGCCCCBB  
 ----CCCCKCCCC--

Original motif    Consensus sequence: BCCGCCCCGCCCCBB



Reverse complement motif    Consensus sequence: BBGGGGCGGGGCGGB

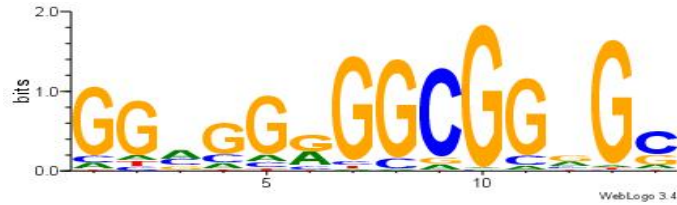


Dataset #: 5  
 Motif ID: 50  
 Motif name: TFF11  
 Matching format of first motif: Reverse Complement  
 Matching format of second motif: Original Motif  
 Direction: Forward  
 Position number: 4  
 Number of overlap: 10  
 Similarity score: 0.0174405

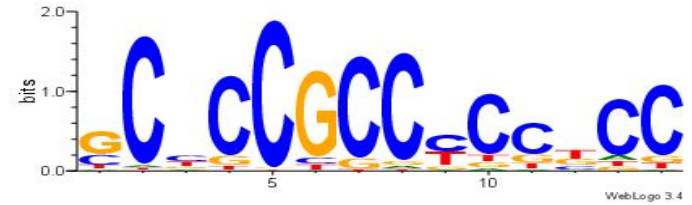
Alignment:

GGMGGRGGCGGVGC  
---GGGGYGGG---

Original motif    Consensus sequence: GGMGGRGGCGGVGC



Reverse complement motif    Consensus sequence: GCVCCGCCMCCYC



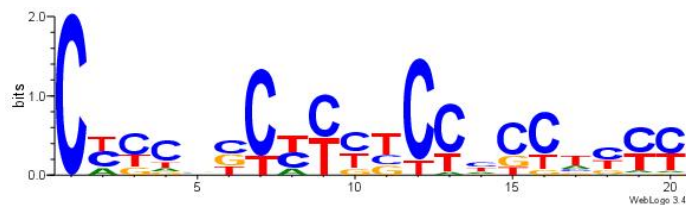
Dataset #: 5  
Motif ID: 54  
Motif name: TFM12  
Matching format of first motif: Original Motif  
Matching format of second motif: Original Motif  
Direction: Forward  
Position number: 7  
Number of overlap: 10  
Similarity score: 0.0301548

Alignment:

CYCBBCYYYTCCHCCTYYY  
-----CCCCKCCCC-----

Original motif    Consensus sequence: CYCBBCYYYTCCHCCTYYY

Reverse complement motif    Consensus sequence:  
KKKAGGDGGAKKMGBBGKMG

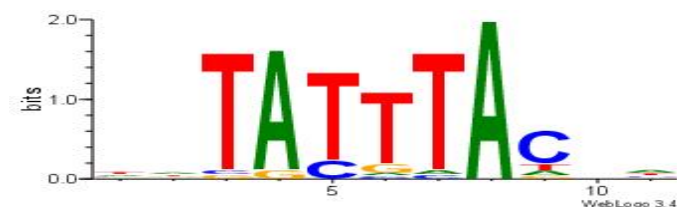


**Dataset #: 4      Motif ID: 45      Motif name: wbgTAAATAww**

Original motif      Consensus sequence: DBGTAATAHD



Reverse complement motif      Consensus sequence: DHTATTTACBD



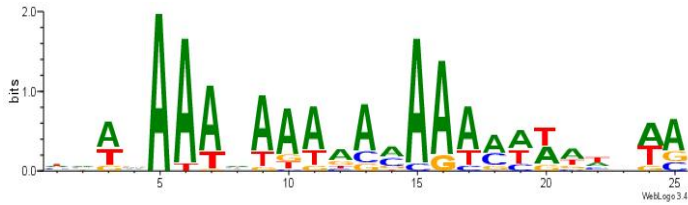
**Best Matches for Top Significant Motif ID 45 (Highest to Lowest)**

Dataset #:	5
Motif ID:	56
Motif name:	TFM11
Matching format of first motif:	Reverse Complement
Matching format of second motif:	Reverse Complement
Direction:	Forward
Position number:	15
Number of overlap:	11
Similarity score:	0.0192768

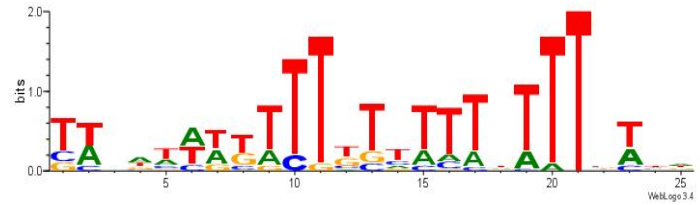
Alignment:

```
TWVHWWYTTTYTTTTHTTTVWBH
-----DHTATTTACBD
```

Original motif    Consensus sequence:  
 HDWVAAAHAAAAAAMAAAMWWWHBWA



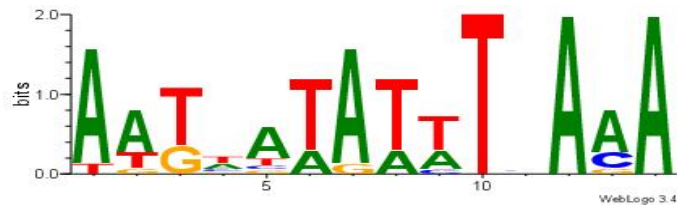
Reverse complement motif    Consensus sequence:  
 TWVHWWWYTTTTYTTTTTHTTTTVWBH



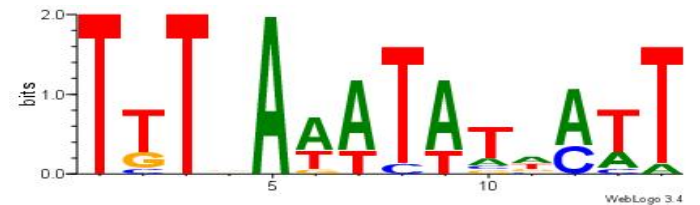
Dataset #: 2  
 Motif ID: 9  
 Motif name: Motif 9  
 Matching format of first motif: Reverse Complement  
 Matching format of second motif: Original Motif  
 Direction: Backward  
 Position number: 4  
 Number of overlap: 11  
 Similarity score: 0.0250541

Alignment:  
 AATHATATWTHAAA  
 DHTATTTACBD---

Original motif    Consensus sequence: AATHATATWTHAAA



Reverse complement motif    Consensus sequence: TTTDAWATATHAT



Dataset #: 2  
 Motif ID: 6  
 Motif name: Motif 6  
 Matching format of first motif: Original Motif  
 Matching format of second motif: Original Motif  
 Direction: Backward  
 Position number: 1  
 Number of overlap: 11  
 Similarity score: 0.027895

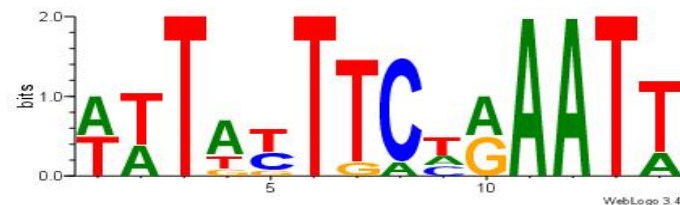
Alignment:

AATTYDGAARTAWW  
 ---DBGTAAATAHD

Original motif      Consensus sequence: AATTYDGAARTAWW



Reverse complement motif      Consensus sequence: WWTAKTTCDKAAT



Dataset #: 5  
 Motif ID: 52  
 Motif name: TFM1  
 Matching format of first motif: Original Motif  
 Matching format of second motif: Reverse Complement  
 Direction: Forward  
 Position number: 1  
 Number of overlap: 11

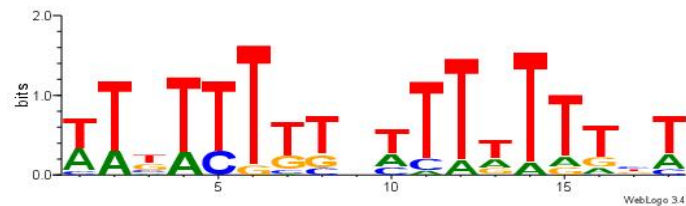


Similarity score: 0.0283217

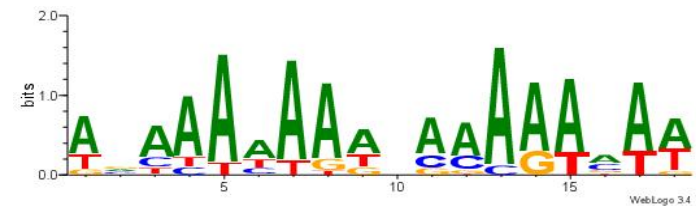
Alignment:

ABAAAAAWhAAAAARAW  
DBGTAATAHD-----

Original motif      Consensus sequence: WTKTTTTTHWTTTTTBT



Reverse complement motif      Consensus sequence:  
ABAAAAAWhAAAAARAW



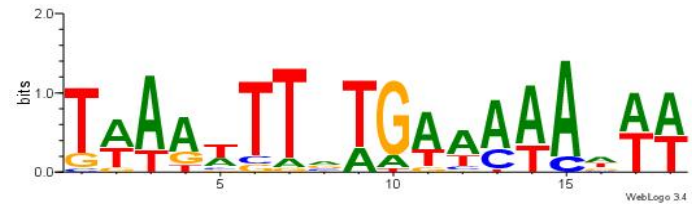
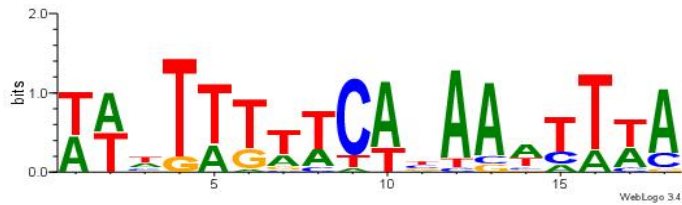
Dataset #: 5  
Motif ID: 53  
Motif name: TFM3  
Matching format of first motif: Original Motif  
Matching format of second motif: Reverse Complement  
Direction: Backward  
Position number: 2  
Number of overlap: 11  
Similarity score: 0.0308097

Alignment:

TWAAWTTVTGAAAAHWW  
-----DBGTAATAHD-

Original motif      Consensus sequence: WWHTTTTTTCABAAWTTWA

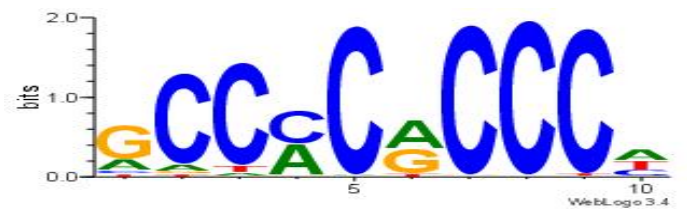
Reverse complement motif      Consensus sequence:  
TWAAWTTVTGAAAAHWW



**Dataset #: 3      Motif ID: 27      Motif name: Klf4**

Original motif      Consensus sequence: DGGYGKGGC

Reverse complement motif      Consensus sequence: GCCYCMCCCD

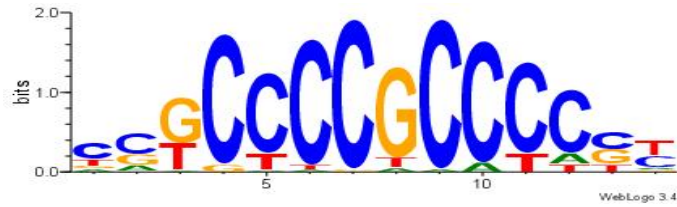


**Best Matches for Top Significant Motif ID 27 (Highest to Lowest)**

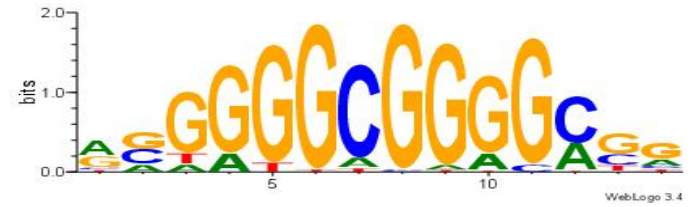
Dataset #:	2
Motif ID:	7
Motif name:	Motif 7
Matching format of first motif:	Reverse Complement
Matching format of second motif:	Original Motif
Direction:	Forward
Position number:	3
Number of overlap:	10
Similarity score:	0.0152457

Alignment:  
 CSKCCCCGCCCSY  
 --GCCYCMCCCD--

Original motif    Consensus sequence: CSKCCCCGCCCSY



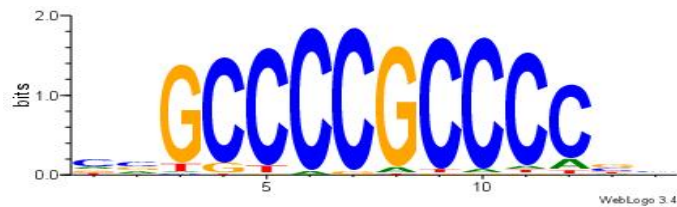
Reverse complement motif    Consensus sequence: MSGGGCGGGGY



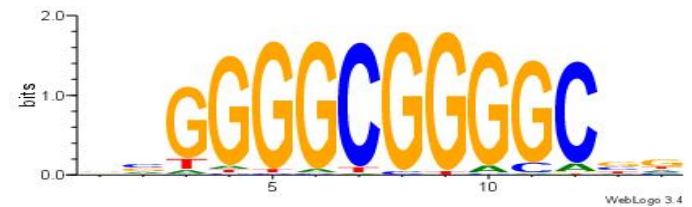
Dataset #: 4  
Motif ID: 36  
Motif name: csGCCCCGCCCSy  
Matching format of first motif: Original Motif  
Matching format of second motif: Reverse Complement  
Direction: Backward  
Position number: 3  
Number of overlap: 10  
Similarity score: 0.0160185

Alignment:  
BBGGGGCGGGGCVD  
--DGGGYGKGGC--

Original motif    Consensus sequence: HVGCCCCGCCCB



Reverse complement motif    Consensus sequence: BBGGGGCGGGGC

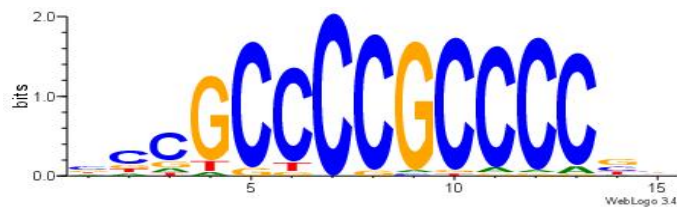


Dataset #: 4  
 Motif ID: 38  
 Motif name: cccGCCCCGCCCCsb  
 Matching format of first motif: Original Motif  
 Matching format of second motif: Reverse Complement  
 Direction: Forward  
 Position number: 3  
 Number of overlap: 10  
 Similarity score: 0.0194893

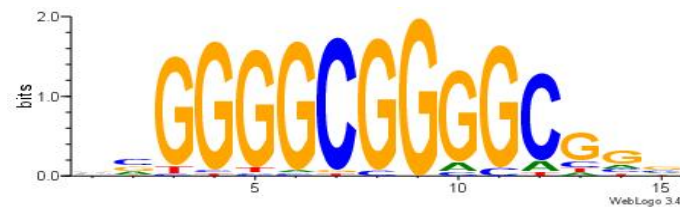
Alignment:

BBGGGGCGGGGCGGB  
 --DGGGYGKGGC---

Original motif      Consensus sequence: BCCGCCCCGCCCCBB



Reverse complement motif      Consensus sequence: BBGGGGCGGGGCGGB

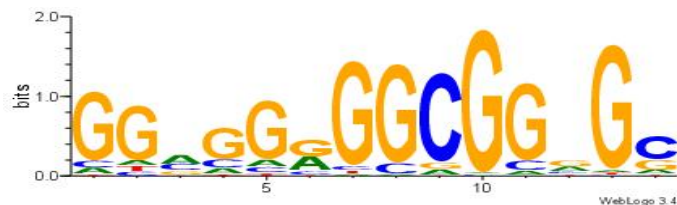


Dataset #: 5  
 Motif ID: 50  
 Motif name: TFF11  
 Matching format of first motif: Original Motif  
 Matching format of second motif: Original Motif  
 Direction: Forward  
 Position number: 5

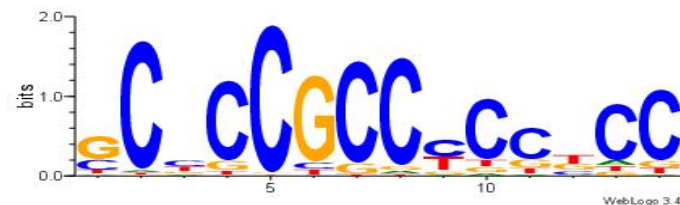
Number of overlap: 10  
Similarity score: 0.0402906

Alignment:  
GGMGGRGGCGGVGC  
----DGGGYGKGGC

Original motif      Consensus sequence: GGMGGRGGCGGVGC



Reverse complement motif      Consensus sequence: GCVCCGCCMCCYC

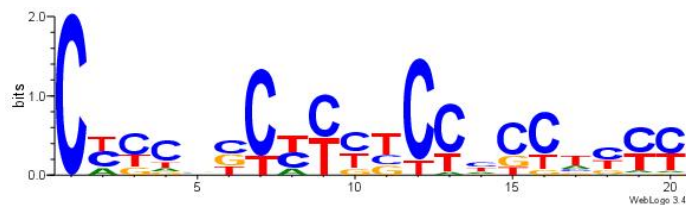


Dataset #: 5  
Motif ID: 54  
Motif name: TFM12  
Matching format of first motif: Reverse Complement  
Matching format of second motif: Original Motif  
Direction: Backward  
Position number: 7  
Number of overlap: 10  
Similarity score: 0.0556427

Alignment:  
CYCBBCYYYTCCHCCTYYY  
----GCCYCMCCD-----

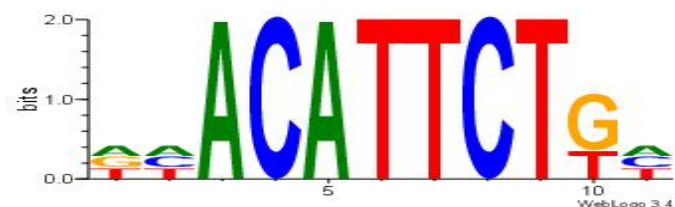
Original motif      Consensus sequence: CYCBBCYYYTCCHCCTYYY

Reverse complement motif      Consensus sequence:  
KKKAGGDGGAKKMGBBGKMG



**Dataset #: 4      Motif ID: 44      Motif name: dhACATTCTkh**

Original motif      Consensus sequence: DHACATTCTGH



Reverse complement motif      Consensus sequence: HCAGAATGTHD



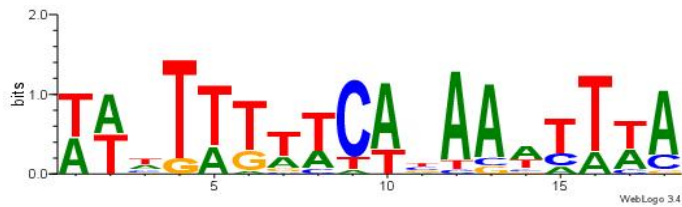
**Best Matches for Top Significant Motif ID 44 (Highest to Lowest)**

Dataset #:	5
Motif ID:	53
Motif name:	TFM3
Matching format of first motif:	Original Motif
Matching format of second motif:	Reverse Complement
Direction:	Backward
Position number:	8
Number of overlap:	11
Similarity score:	0.0217516

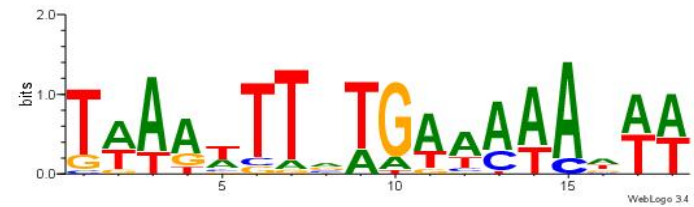
Alignment:

```
TWAAWTTVTGAAAAHWW
DHACATTCTGH-----
```

Original motif      Consensus sequence: WWHTTTTTTCABAAWTTWA



Reverse complement motif      Consensus sequence: TWAAWTTVTGAAAAHWW

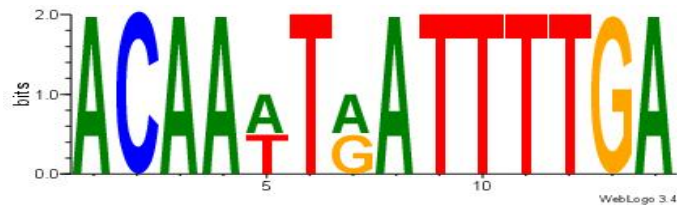


Dataset #: 2  
Motif ID: 16  
Motif name: Motif 16  
Matching format of first motif: Reverse Complement  
Matching format of second motif: Reverse Complement  
Direction: Backward  
Position number: 4  
Number of overlap: 11  
Similarity score: 0.0258838

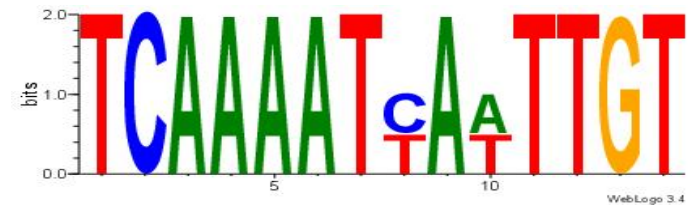
Alignment:

TCAAATKAWTTGT  
HCAGAATGTHD---

Original motif      Consensus sequence: ACAAWTRATTTTGA



Reverse complement motif      Consensus sequence: TCAAATKAWTTG

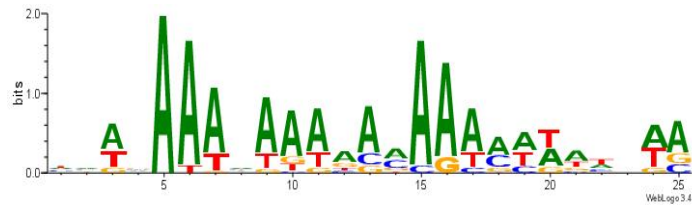


Dataset #: 5  
 Motif ID: 56  
 Motif name: TFM11  
 Matching format of first motif: Original Motif  
 Matching format of second motif: Reverse Complement  
 Direction: Backward  
 Position number: 11  
 Number of overlap: 11  
 Similarity score: 0.0307858

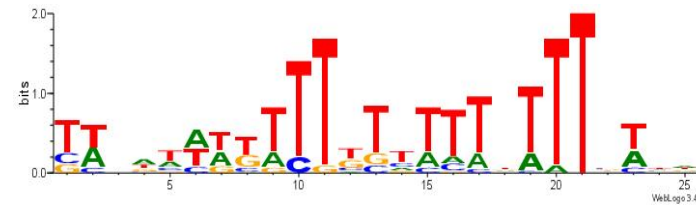
Alignment:

TWVHWWWYTTTYTTTTHTTTVWBH  
 ----DHACATTCTGH-----

Original motif    Consensus sequence:  
 HDWVAAAHAAAAAMAAAMWWWHBWA



Reverse complement motif    Consensus sequence:  
 TWVHWWWYTTTYTTTTHTTTVWBH



Dataset #: 5  
 Motif ID: 55  
 Motif name: TFM13  
 Matching format of first motif: Original Motif  
 Matching format of second motif: Original Motif  
 Direction: Backward  
 Position number: 9

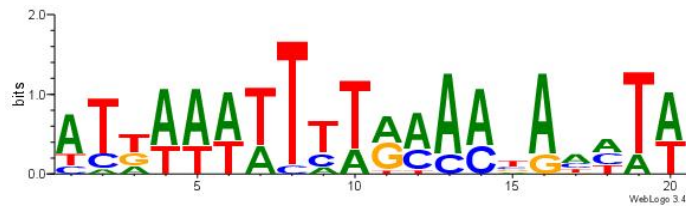


Number of overlap: 11  
Similarity score: 0.0357323

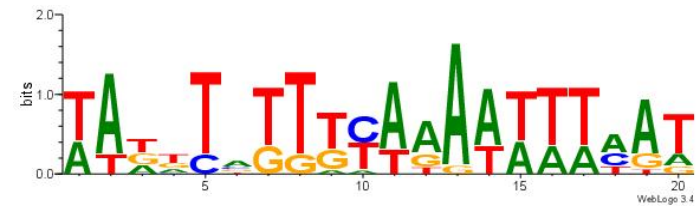
Alignment:

ATKAAWTTTTTRMAABAHTW  
-DHACATTCTGH-----

Original motif      Consensus sequence: ATKAAWTTTTTRMAABAHTW



Reverse complement motif      Consensus sequence:  
WAHHTVTTYKAAAATTRAT



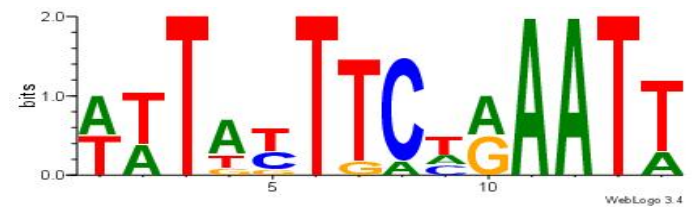
Dataset #: 2  
Motif ID: 6  
Motif name: Motif 6  
Matching format of first motif: Reverse Complement  
Matching format of second motif: Original Motif  
Direction: Backward  
Position number: 4  
Number of overlap: 11  
Similarity score: 0.0400884

Alignment:

AATTYDGAARTAWW  
HCAGAATGTHD---

Original motif      Consensus sequence: AATTYDGAARTAWW

Reverse complement motif      Consensus sequence: WWTAKTTCDKAAT



**Dataset #: 1      Motif ID: 1      Motif name: Motif 1**

Original motif      Consensus sequence: GCGGGGC



Reverse complement motif      Consensus sequence: GCCCGCC

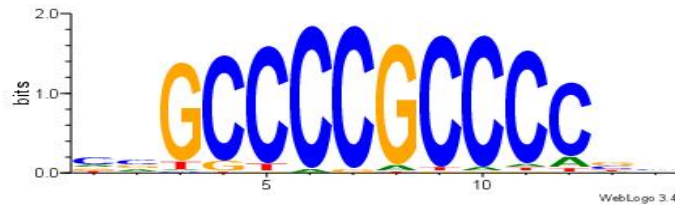


**Best Matches for Top Significant Motif ID 1 (Highest to Lowest)**

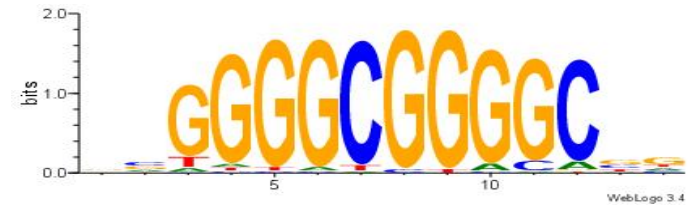
Dataset #:	4
Motif ID:	36
Motif name:	csGCCCCGCCCCsc
Matching format of first motif:	Reverse Complement
Matching format of second motif:	Original Motif
Direction:	Forward
Position number:	3
Number of overlap:	8
Similarity score:	0

Alignment:  
HVGCCCCGCCCCBB  
--GCCCGCC----

Original motif    Consensus sequence: HVGCCCCGCCCCBB



Reverse complement motif    Consensus sequence: BBGGGGCGGGGC

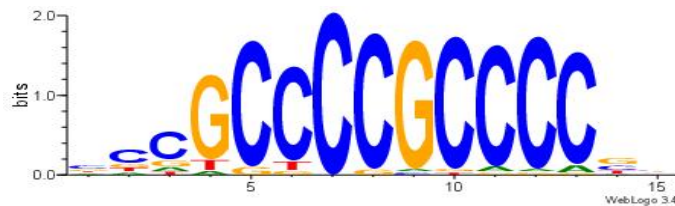


Dataset #: 4  
Motif ID: 38  
Motif name: cccGCCCCGCCCCsb  
Matching format of first motif: Reverse Complement  
Matching format of second motif: Original Motif  
Direction: Forward  
Position number: 4  
Number of overlap: 8  
Similarity score: 0.00250079

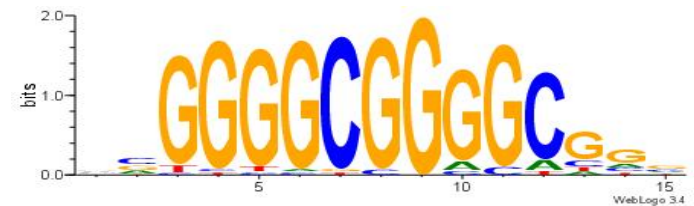
Alignment:

BCCGCCCCGCCCCBB  
---GCCCCGCC----

Original motif    Consensus sequence: BCCGCCCCGCCCCBB



Reverse complement motif    Consensus sequence: BBGGGGCGGGGCGB

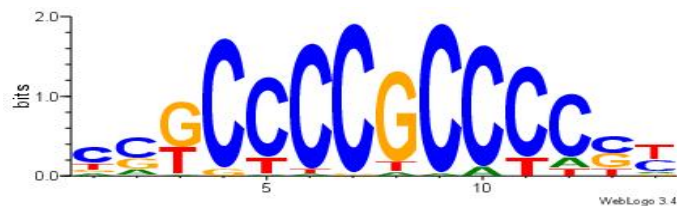


Dataset #: 2  
 Motif ID: 7  
 Motif name: Motif 7  
 Matching format of first motif: Reverse Complement  
 Matching format of second motif: Original Motif  
 Direction: Backward  
 Position number: 5  
 Number of overlap: 8  
 Similarity score: 0.0149958

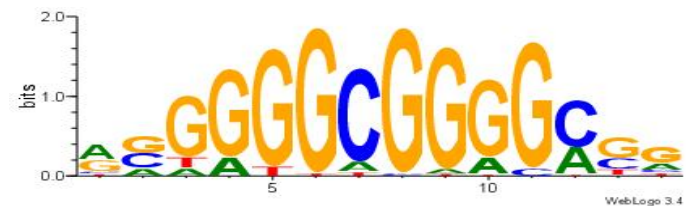
Alignment:

CSKCCCCGCCCSY  
 --GCCCGCC----

Original motif      Consensus sequence: CSKCCCCGCCCSY



Reverse complement motif      Consensus sequence: MSGGGGCGGGY



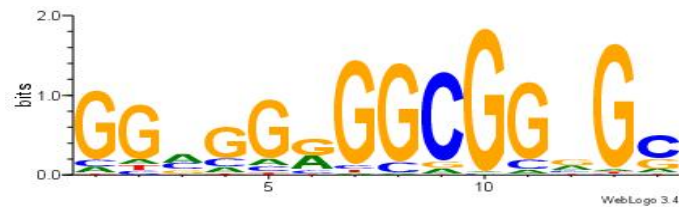
Dataset #: 5  
 Motif ID: 50  
 Motif name: TFF11  
 Matching format of first motif: Original Motif  
 Matching format of second motif: Original Motif  
 Direction: Backward  
 Position number: 1  
 Number of overlap: 8

Similarity score: 0.0348969

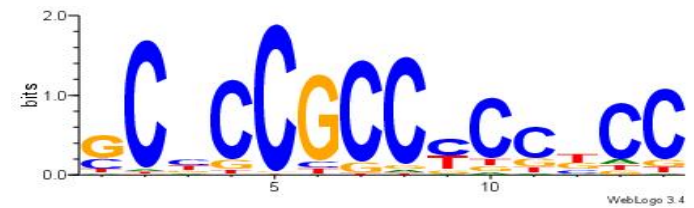
Alignment:

GGMGGRGGCGGVGC  
-----GGCGGGGC

Original motif    Consensus sequence: GGMGGRGGCGGVGC



Reverse complement motif    Consensus sequence: GCVCCGCCMCCYC



---

Dataset #: 3  
Motif ID: 27  
Motif name: Klf4  
Matching format of first motif: Original Motif  
Matching format of second motif: Original Motif  
Direction: Forward  
Position number: 3  
Number of overlap: 8  
Similarity score: 0.0349983

Alignment:

DGGGYGKGGC  
--GGCGGGGC

Original motif    Consensus sequence: DGGGYGKGGC

Reverse complement motif    Consensus sequence: GCCYCMCCCD



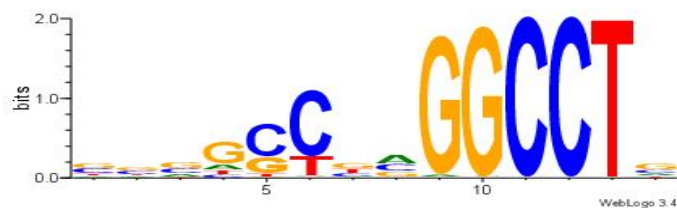
Dataset #: 3  
 Motif ID: 22  
 Motif name: Zfx  
 Matching format of first motif: Reverse Complement  
 Matching format of second motif: Reverse Complement  
 Direction: Forward  
 Position number: 4  
 Number of overlap: 8  
 Similarity score: 0.0708224

**Alignment:**

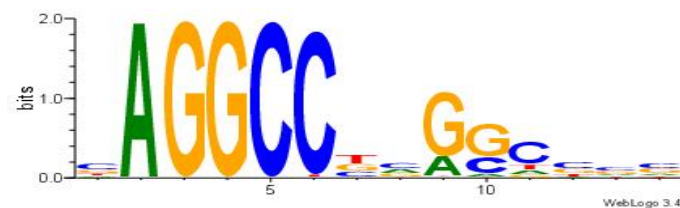
```

VAGGCCBBGGCVBB
---GCCCCGCC---
  
```

Original motif      Consensus sequence: BBVGCCBVGCCCTV



Reverse complement motif      Consensus sequence: VAGGCCBBGGCVBB



Results created by MOTIFSIM on 11-21-2023 10:01:18  
Runtime: 65.8356 seconds

MOTIFSIM is written by Ngoc Tam L. Tran  
Motif logo generated by [weblago](#)