



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:	UniProbe Homo Sapiens
Motif tree:	Yes
Combined similar motifs:	Yes
Output file type:	All
Output file format:	PDF

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_DM05.txt	11	5

RESULTS

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

Dataset #: 5 **Motif ID: 46** **Motif name: TFW3**

Original motif Consensus sequence: GCACTG



Reverse complement motif Consensus sequence: CAGTGC



Best Matches for Top Significant Motif ID 46 (Highest to Lowest)

Dataset #: 4
Motif ID: 43
Motif name: wSTACwGTAsw
Matching format of first motif: Reverse Complement
Matching format of second motif: Reverse Complement
Direction: Backward
Position number: 2
Number of overlap: 6
Similarity score: 0.00773193

Alignment:
DBTACWGTA VH
----CAGTGC--

Original motif Consensus sequence: HVTACWG TABD

Reverse complement motif Consensus sequence: DBTACWGTA VH

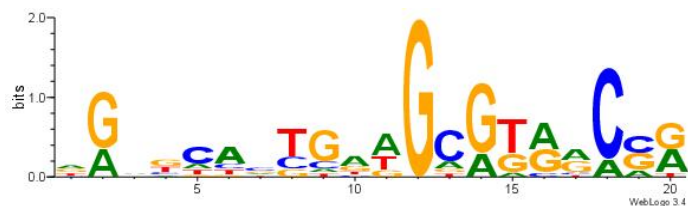


Dataset #: 3
 Motif ID: 31
 Motif name: Pax5
 Matching format of first motif: Original Motif
 Matching format of second motif: Original Motif
 Direction: Forward
 Position number: 4
 Number of overlap: 6
 Similarity score: 0.0111111

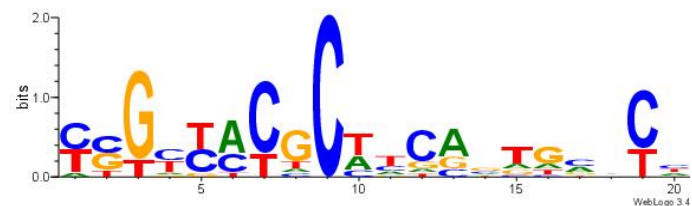
Alignment:

DGVCBABCATGDWCGKRRCSR
 ---GCACTG-----

Original motif Consensus sequence: DGVCBABCATGDWCGKRRCSR



Reverse complement motif Consensus sequence: MSGKKRCGCWDCABTGBBCD



Dataset #: 3

Motif ID: 34
 Motif name: Myc
 Matching format of first motif: Reverse Complement
 Matching format of second motif: Reverse Complement
 Direction: Backward
 Position number: 2
 Number of overlap: 6
 Similarity score: 0.0148473

Alignment:
 DCCACGTGCV
 ---CAGTGC-

Original motif Consensus sequence: VGCACGTGGH



Reverse complement motif Consensus sequence: DCCACGTGCV



Dataset #: 3
 Motif ID: 35
 Motif name: GABPA
 Matching format of first motif: Original Motif
 Matching format of second motif: Reverse Complement
 Direction: Backward
 Position number: 2
 Number of overlap: 6
 Similarity score: 0.0156785

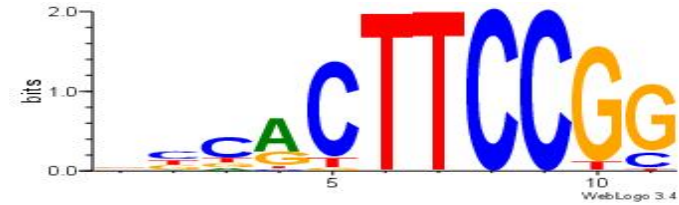
Alignment:

VVCACTTCCGG
----GCACTG--

Original motif Consensus sequence: CCGGAAGTGVV



Reverse complement motif Consensus sequence: VVCACTTCCGG



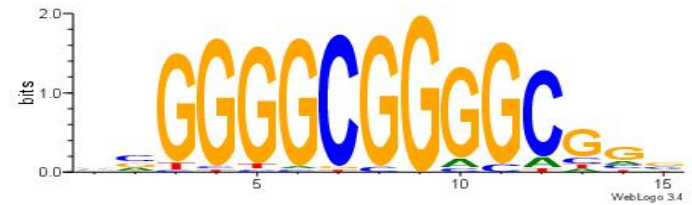
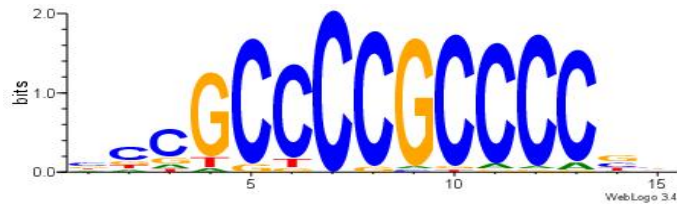
Dataset #: 4
Motif ID: 38
Motif name: cccGCCCCGCCCCsb
Matching format of first motif: Original Motif
Matching format of second motif: Original Motif
Direction: Forward
Position number: 4
Number of overlap: 6
Similarity score: 0.0173006

Alignment:

BCCGCCCCGCCCCBB
---GCACTG-----

Original motif Consensus sequence: BCCGCCCCGCCCCBB

Reverse complement motif Consensus sequence:
BBGGGGCGGGGCGGB

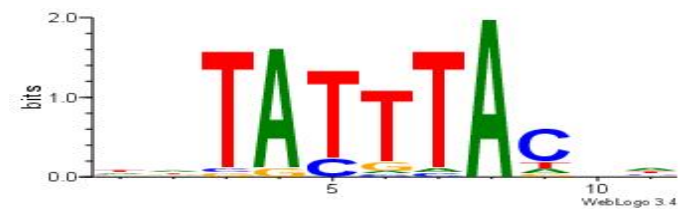


Dataset #: 4 Motif ID: 45 Motif name: wbgTAAATAww

Original motif Consensus sequence: DBGTAAATAHD



Reverse complement motif Consensus sequence: DHTATTTACBD

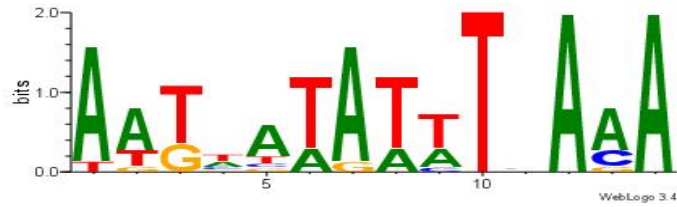


Best Matches for Top Significant Motif ID 45 (Highest to Lowest)

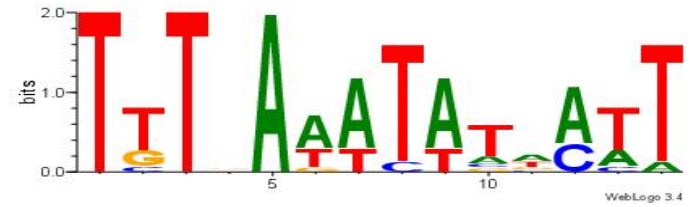
Dataset #:	2
Motif ID:	9
Motif name:	Motif 9
Matching format of first motif:	Reverse Complement
Matching format of second motif:	Original Motif
Direction:	Backward
Position number:	4
Number of overlap:	11
Similarity score:	0.0177412

Alignment:
 AATHATATWTHAAA
 DHTATTTACBD---

Original motif Consensus sequence: AATHATATWTHAAA



Reverse complement motif Consensus sequence: TTTDAWATATHAT

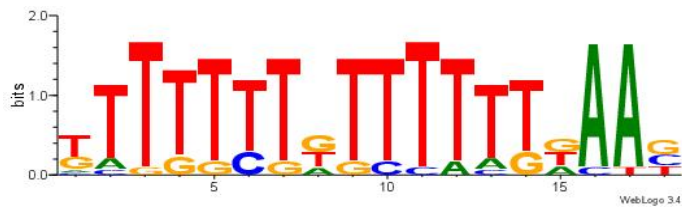


Dataset #: 5
 Motif ID: 53
 Motif name: TFM3
 Matching format of first motif: Original Motif
 Matching format of second motif: Reverse Complement
 Direction: Forward
 Position number: 4
 Number of overlap: 11
 Similarity score: 0.0203792

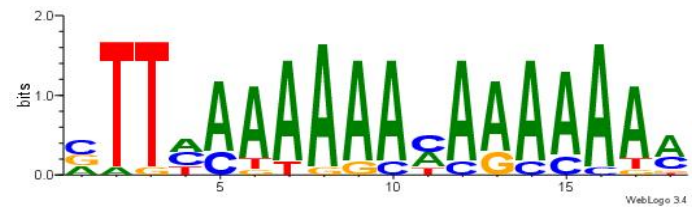
Alignment:

```
BTTHAAAAAYAAAAAR
---DBGTAATAHD----
```

Original motif Consensus sequence: KTTTTTKTTTTTTDAAB



Reverse complement motif Consensus sequence: BTTHAAAAAYAAAAAR



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: 0.0205821

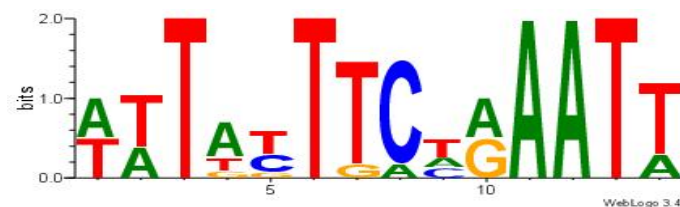
Alignment:

AATTYDGAARTAWW
 ---DBGTAAATAHD

Original motif Consensus sequence: AATTYDGAARTAWW



Reverse complement motif Consensus sequence: WWTAKTTCDKAAT



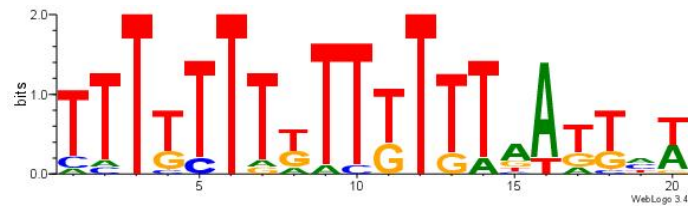
Dataset #: 5
 Motif ID: 56
 Motif name: TFM13
 Matching format of first motif: Reverse Complement
 Matching format of second motif: Original Motif
 Direction: Backward
 Position number: 3
 Number of overlap: 11

Similarity score: 0.0258581

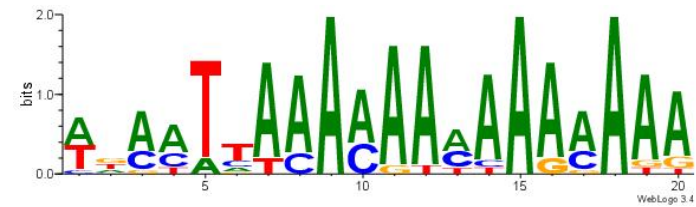
Alignment:

```
TTTTTTTKTTKTTTAATTHW  
-----DHTATTTACBD--
```

Original motif Consensus sequence: TTTTTTKTTKTTTAATTHW



Reverse complement motif Consensus sequence: WHAATTAARAARAAAAAA



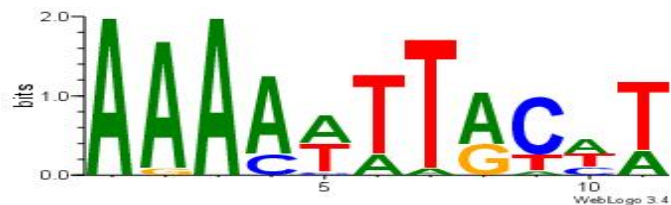
Dataset #:	2
Motif ID:	4
Motif name:	Motif 4
Matching format of first motif:	Original Motif
Matching format of second motif:	Reverse Complement
Direction:	Backward
Position number:	1
Number of overlap:	11
Similarity score:	0.0288058

Alignment:

```
AWGKAAWTTTT  
DBGTAAATAHD
```

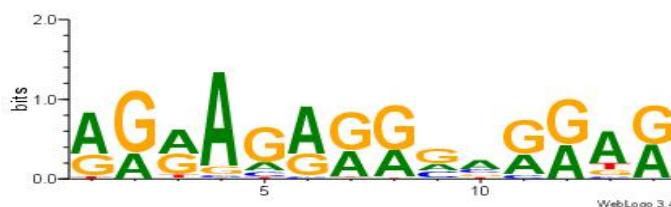
Original motif Consensus sequence: AAAAWTTRCWT

Reverse complement motif Consensus sequence: AWGKAAWTTTT

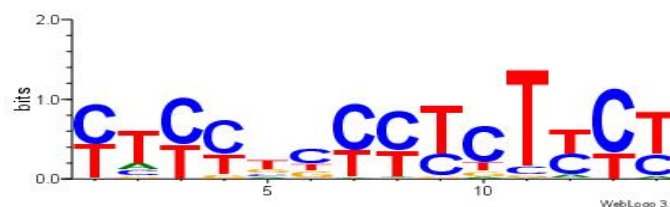


Dataset #: 2 Motif ID: 2 Motif name: Motif 2

Original motif Consensus sequence: RGRAGARRGARRAR



Reverse complement motif Consensus sequence: MTMMTCMMTCTK

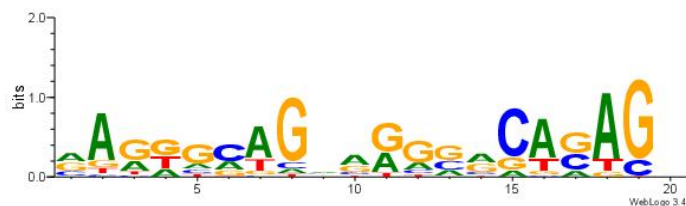


Best Matches for Top Significant Motif ID 2 (Highest to Lowest)

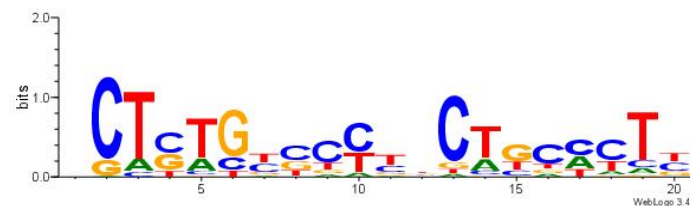
Dataset #: 5
 Motif ID: 55
 Motif name: TFM12
 Matching format of first motif: Original Motif
 Matching format of second motif: Original Motif
 Direction: Forward
 Position number: 4
 Number of overlap: 14
 Similarity score: 0.00215466

Alignment:
 RAGKGMAGVRRGSRCASAGV
 ---RGRAGARRGARRAR---

Original motif Consensus sequence: RAGKGMAGVRRGSRCASAGV



Reverse complement motif Consensus sequence: VCTSTGKSCMKBCTRCYCTK

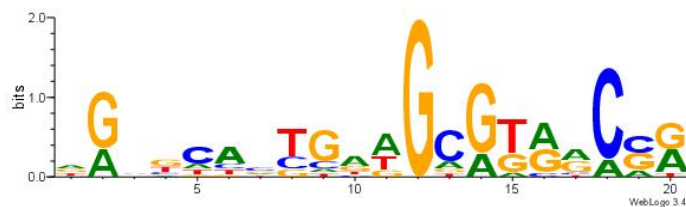


Dataset #: 3
Motif ID: 31
Motif name: Pax5
Matching format of first motif: Reverse Complement
Matching format of second motif: Reverse Complement
Direction: Forward
Position number: 7
Number of overlap: 14
Similarity score: 0.0213256

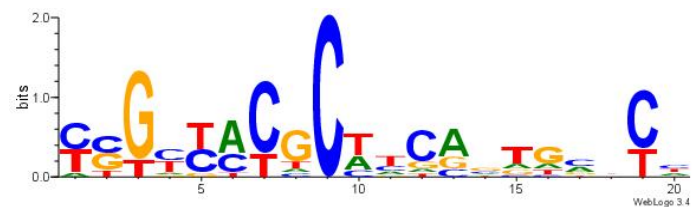
Alignment:

```
MSGKKRCGCWDCABTG B B C D  
-----MTMMTCMMTCTKCK
```

Original motif Consensus sequence: DGVBCABTGDWGCGRRCRCSR



Reverse complement motif Consensus sequence: MSGKKRCGCWDCABTG B B C D

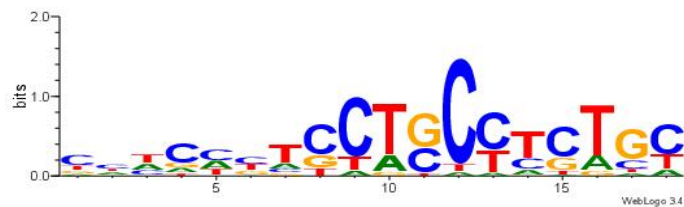


Dataset #: 5
 Motif ID: 52
 Motif name: TFM2
 Matching format of first motif: Reverse Complement
 Matching format of second motif: Original Motif
 Direction: Forward
 Position number: 3
 Number of overlap: 14
 Similarity score: 0.0283463

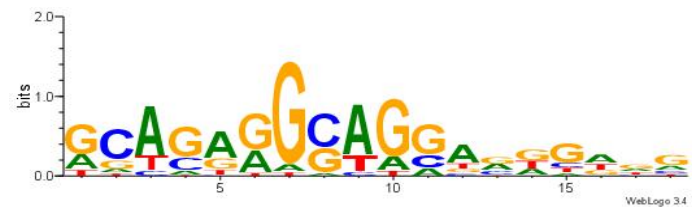
Alignment:

CHHCHBTCCTSCYTCTGC
 --MTMMTCMMTCTKCK--

Original motif Consensus sequence: CHHCHBTCCTSCYTCTGC



Reverse complement motif Consensus sequence: GCAGAKGSAGGABDGH DG

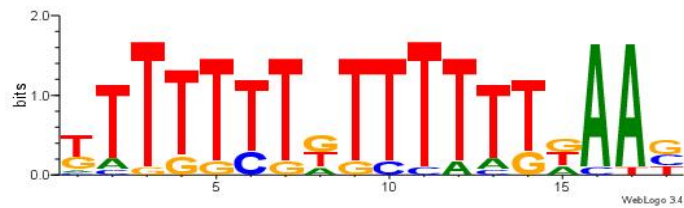


Dataset #: 5
 Motif ID: 53
 Motif name: TFM3
 Matching format of first motif: Original Motif
 Matching format of second motif: Reverse Complement
 Direction: Forward
 Position number: 5

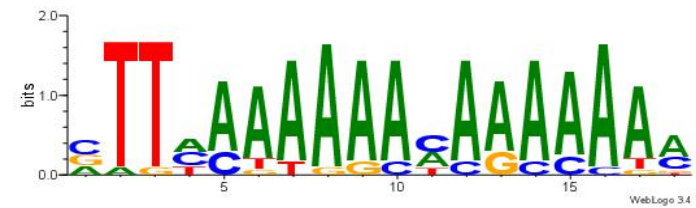
Number of overlap: 14
 Similarity score: 0.0316147

Alignment:
 BTTHAAAAAYAAAAAR
 ----RGRAGARRGARRAR

Original motif Consensus sequence: KTTTTTKTTTTTDAAB



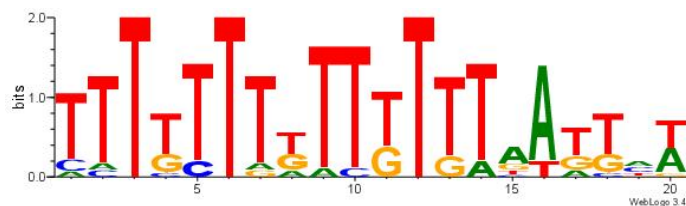
Reverse complement motif Consensus sequence: BTTHAAAAAYAAAAAR



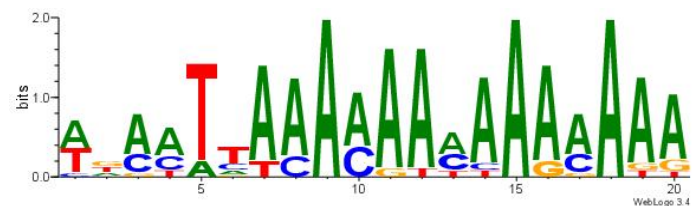
Dataset #: 5
 Motif ID: 56
 Motif name: TFM13
 Matching format of first motif: Original Motif
 Matching format of second motif: Reverse Complement
 Direction: Backward
 Position number: 4
 Number of overlap: 14
 Similarity score: 0.0368869

Alignment:
 WHAATTAAARAARAAAAA
 ----RGRAGARRGARRAR----

Original motif Consensus sequence: TTTTTTKTTKTTTAATTHW

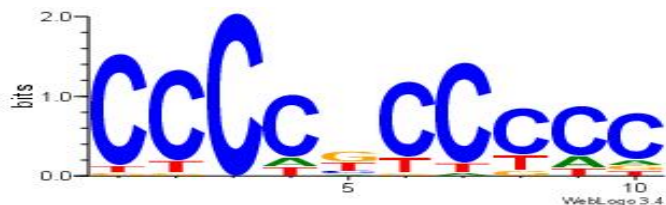


Reverse complement motif Consensus sequence: WHAATTAARAARAAAAA

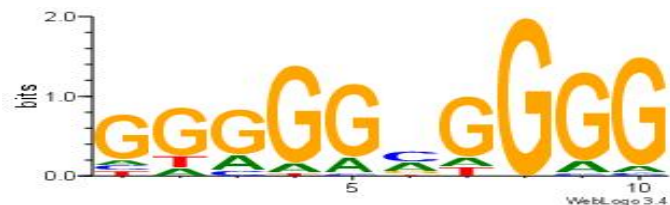


Dataset #: 3 Motif ID: 24 Motif name: SP1

Original motif Consensus sequence: CCCCKCCCC



Reverse complement motif Consensus sequence: GGGGYGGGG



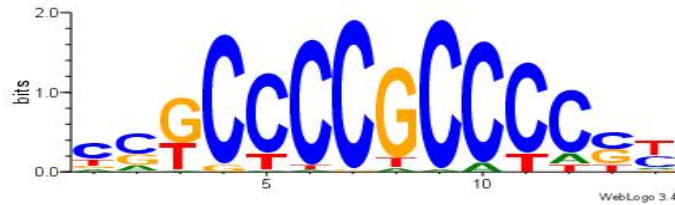
Best Matches for Top Significant Motif ID 24 (Highest to Lowest)

Dataset #:	2
Motif ID:	7
Motif name:	Motif 7
Matching format of first motif:	Reverse Complement
Matching format of second motif:	Reverse Complement
Direction:	Backward
Position number:	4
Number of overlap:	10
Similarity score:	0.00451594

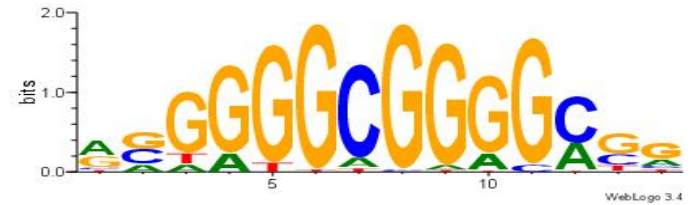
Alignment:

MSGGGGCGGGYSG
-GGGGYGGGG---

Original motif Consensus sequence: CSKCCCCGCCCSY



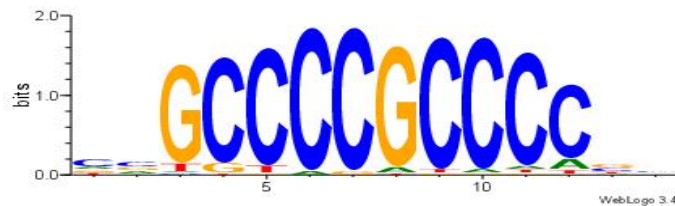
Reverse complement motif Consensus sequence: MSGGGGCGGGGY



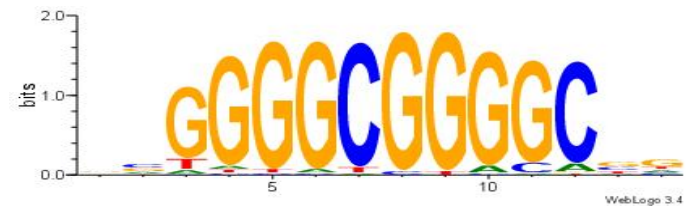
Dataset #: 4
Motif ID: 36
Motif name: csGCCCCGCCCSy
Matching format of first motif: Reverse Complement
Matching format of second motif: Reverse Complement
Direction: Backward
Position number: 4
Number of overlap: 10
Similarity score: 0.00965796

Alignment:
BBGGGGCGGGCVD
-GGGGYGGGG---

Original motif Consensus sequence: HVGCCCCGCCCB



Reverse complement motif Consensus sequence: BBGGGGCGGGC

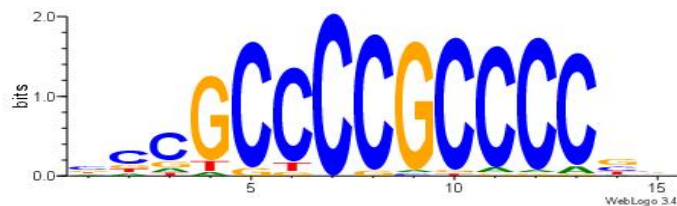


Dataset #: 4
 Motif ID: 38
 Motif name: cccGCCCCGCCCCsb
 Matching format of first motif: Original Motif
 Matching format of second motif: Original Motif
 Direction: Backward
 Position number: 2
 Number of overlap: 10
 Similarity score: 0.0127292

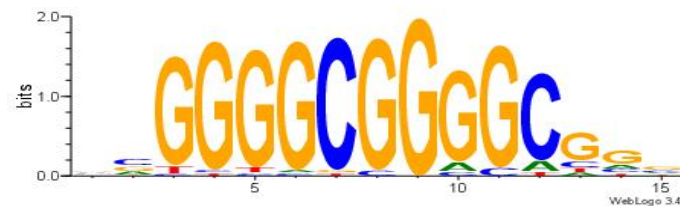
Alignment:

BCCGCCCCGCCCCBB
 ----CCCCKCCCC-

Original motif Consensus sequence: BCCGCCCCGCCCCBB



Reverse complement motif Consensus sequence: BBGGGGCGGGGCGGB



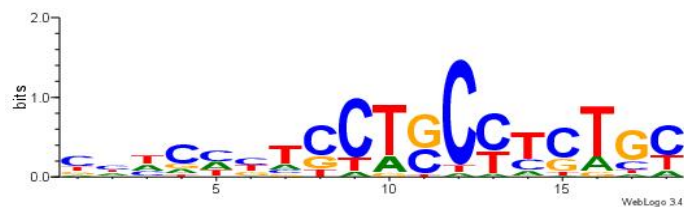
Dataset #: 5
 Motif ID: 52
 Motif name: TFM2
 Matching format of first motif: Original Motif
 Matching format of second motif: Original Motif
 Direction: Forward

Position number: 6
Number of overlap: 10
Similarity score: 0.0523047

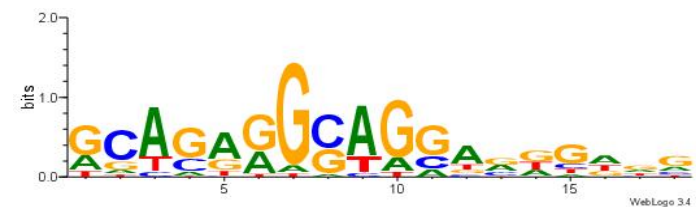
Alignment:

CHHCHBTCCTSCYTCTGC
-----CCCCKCCCC-----

Original motif Consensus sequence: CHHCHBTCCTSCYTCTGC



Reverse complement motif Consensus sequence:
GCAGAKGSAGGABDGHG

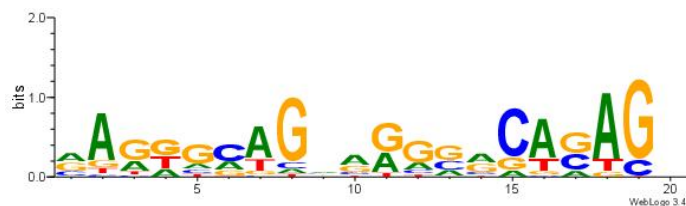


Dataset #: 5
Motif ID: 55
Motif name: TFM12
Matching format of first motif: Original Motif
Matching format of second motif: Reverse Complement
Direction: Backward
Position number: 5
Number of overlap: 10
Similarity score: 0.0639596

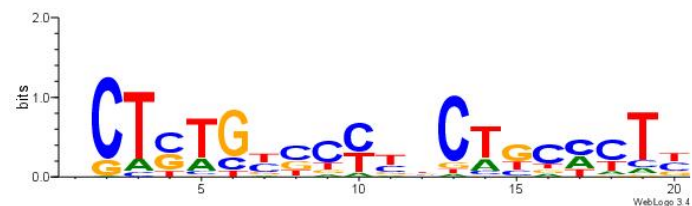
Alignment:

VCTSTGKSCMKBCTRCYCTK
-----CCCCKCCCC-----

Original motif Consensus sequence: RAGKGMAGVRRGSRCASAGV



Reverse complement motif Consensus sequence: VCTSTGKSCMKBCTRCYCTK

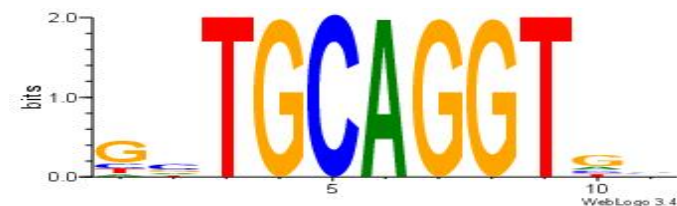


Dataset #: 4 Motif ID: 40 Motif name: kcACCTGCAGc

Original motif Consensus sequence: BCACCTGCABC



Reverse complement motif Consensus sequence: GBTGCAGGTGB



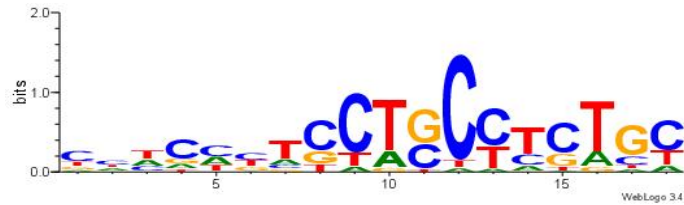
Best Matches for Top Significant Motif ID 40 (Highest to Lowest)

Dataset #:	5
Motif ID:	52
Motif name:	TFM2
Matching format of first motif:	Reverse Complement
Matching format of second motif:	Reverse Complement
Direction:	Backward
Position number:	5
Number of overlap:	11
Similarity score:	0.0194012

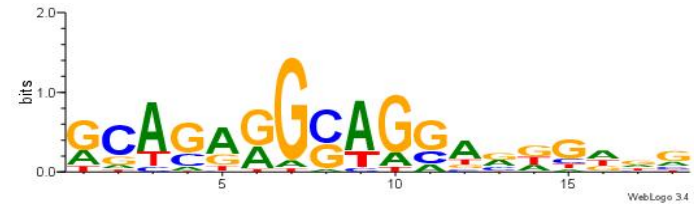
Alignment:

GCAGAKGSAGGABDGHDG
---GBTGCAGGTGB----

Original motif Consensus sequence: CHHCHBTCCTSCYTCTGC



Reverse complement motif Consensus sequence:
GCAGAKGSAGGABDGHDG

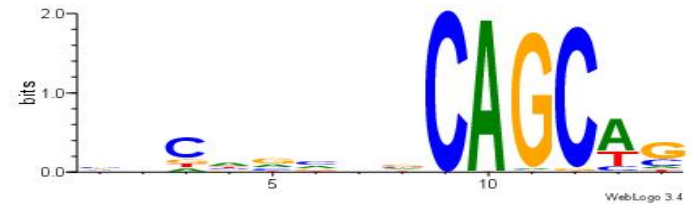
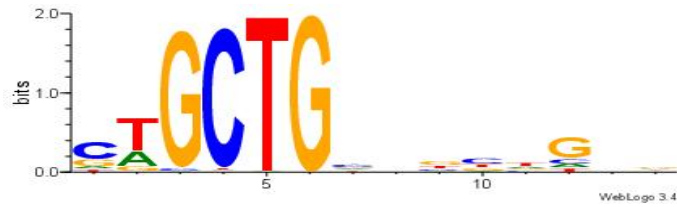


Dataset #: 5
Motif ID: 50
Motif name: TFF11
Matching format of first motif: Reverse Complement
Matching format of second motif: Original Motif
Direction: Backward
Position number: 2
Number of overlap: 11
Similarity score: 0.0222944

Alignment:
CTGCTGBDBBDGBD
--GBTGCAGGTGB-

Original motif Consensus sequence: CTGCTGBDBBDGBD

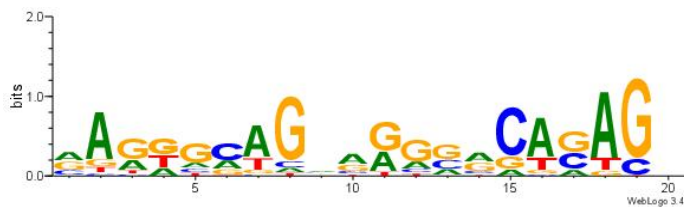
Reverse complement motif Consensus sequence: HVCDBBDBCAGCA



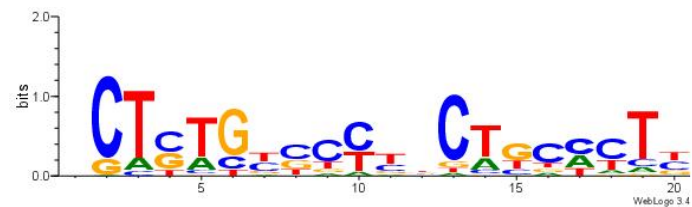
Dataset #: 5
 Motif ID: 55
 Motif name: TFM12
 Matching format of first motif: Reverse Complement
 Matching format of second motif: Original Motif
 Direction: Backward
 Position number: 6
 Number of overlap: 11
 Similarity score: 0.0374254

Alignment:
 RAGKGMAGVRRGSRCASAGV
 ----GBTGCAGGTGB-----

Original motif Consensus sequence: RAGKGMAGVRRGSRCASAGV



Reverse complement motif Consensus sequence: VCTSTGKSCMKBCTRCYCTK



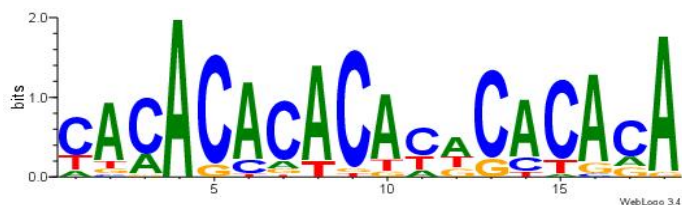
Dataset #: 5

Motif ID: 51
 Motif name: TFM1
 Matching format of first motif: Reverse Complement
 Matching format of second motif: Reverse Complement
 Direction: Forward
 Position number: 7
 Number of overlap: 11
 Similarity score: 0.0402527

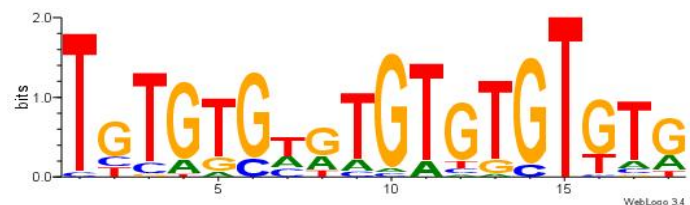
Alignment:

TGTGTGWGTGTGTGTGTG
 -----GBTGCAGGTGB-

Original motif Consensus sequence: CACACACACWCACACA



Reverse complement motif Consensus sequence: TGTGTGWGTGTGTGTGTG



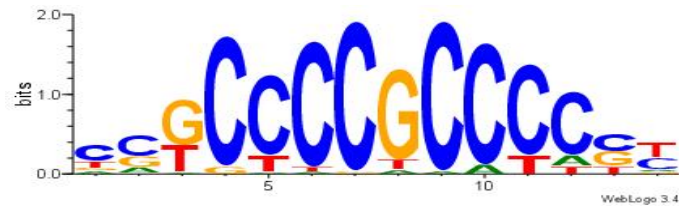
Dataset #: 2
 Motif ID: 7
 Motif name: Motif 7
 Matching format of first motif: Original Motif
 Matching format of second motif: Original Motif
 Direction: Forward
 Position number: 2
 Number of overlap: 11

Similarity score: 0.0411257

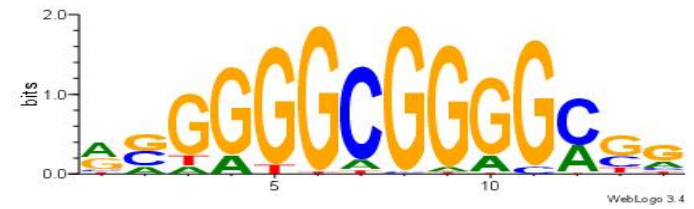
Alignment:

CSKCCCCGCCCSY
-BCACCTGCABC--

Original motif Consensus sequence: CSKCCCCGCCCSY



Reverse complement motif Consensus sequence: MSGGGCGGGGY



Results created by MOTIFSIM on 06-24-2023 16:08:01
Runtime: 118.467 seconds

MOTIFSIM is written by Ngoc Tam L. Tran
Motif logo generated by [weblogo](http://weblogo.threebio.com)