



# 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 Mus Musculus
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

### Best Matches in Database for Each Motif (Highest to Lowest)

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

Original motif

Consensus sequence: GGCGGGGC



Reverse complement motif

Consensus sequence: GCCCGGCC



### Best Matches for Motif ID 1 (Highest to Lowest)

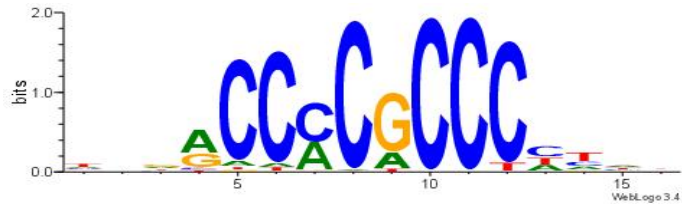
Motif ID:	<a href="#">UP00093</a>
Motif name:	Klf7_primary
Species:	Mus musculus
Matching format of first motif:	Reverse Complement
Matching format of second motif:	Original Motif
Direction:	Backward
Position number:	6
Number of overlap:	8
Similarity score:	0

Alignment:

```
HBDRCCMCGCCCHHHD  
---GCCCGGCC-----
```

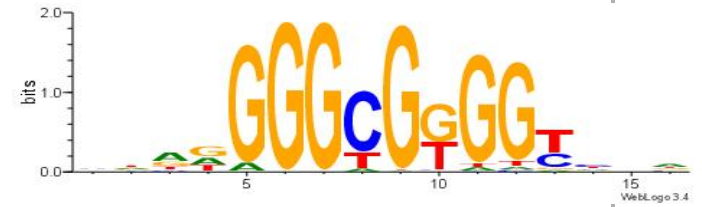
Original motif

Consensus sequence: HBDRCCMCGCCCHHHD



Reverse complement motif

Consensus sequence: DHHDGGGCGRGGKHBH



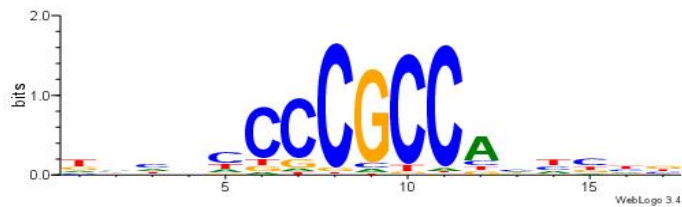
Motif ID: [UP00000](#)  
 Motif name: Smad3\_secondary  
 Species: Mus musculus  
 Matching format of first motif: Original Motif  
 Matching format of second motif: Reverse Complement  
 Direction: Forward  
 Position number: 7  
 Number of overlap: 8  
 Similarity score: 0.00617043

Alignment:

```
BHBHDTGGCGGGBDHD
-----GGCGGGGC----
```

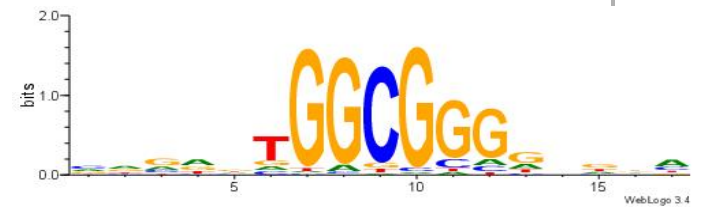
Original motif

Consensus sequence: DHHBCCCCGCCAHHBHB



Reverse complement motif

Consensus sequence: BHBHDTGGCGGGBDHD



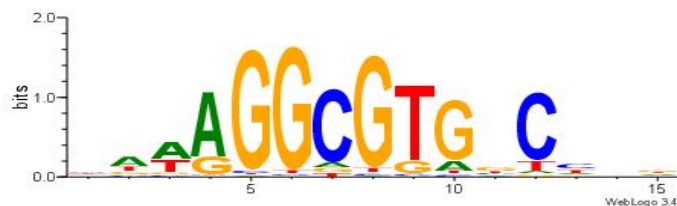
Motif ID: [UP00002](#)  
Motif name: Sp4\_secondary  
Species: Mus musculus  
Matching format of first motif: Original Motif  
Matching format of second motif: Original Motif  
Direction: Forward  
Position number: 5  
Number of overlap: 8  
Similarity score: 0.0145996

Alignment:

BDWAGGCGTGBCHHD  
----GGCGGGC---

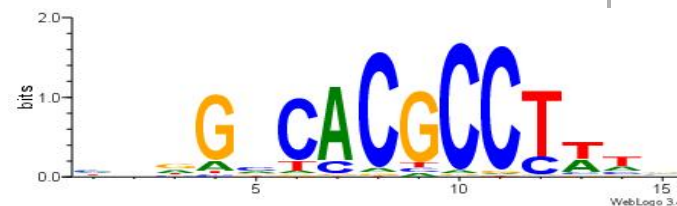
Original motif

Consensus sequence: BDWAGGCGTGBCHHD



Reverse complement motif

Consensus sequence: HHDGBCACGCCTWDB



Motif ID: [UP00002](#)  
Motif name: Sp4\_primary  
Species: Mus musculus  
Matching format of first motif: Reverse Complement  
Matching format of second motif: Original Motif  
Direction: Forward

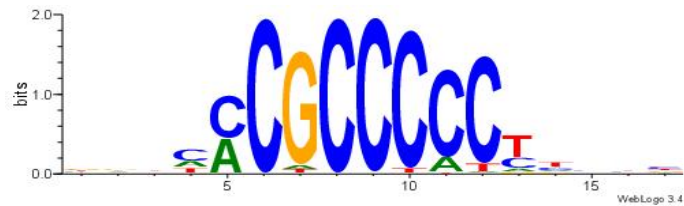
Position number: 2  
Number of overlap: 8  
Similarity score: 0.0233506

Alignment:

BDHHMCGCCCCCTHVBB  
-GCCCCGCC-----

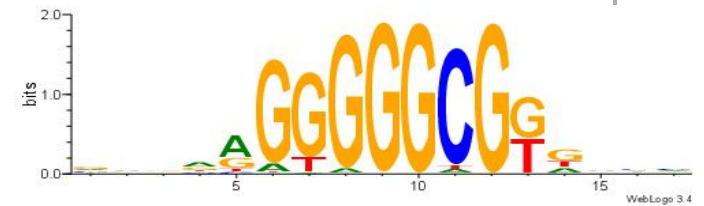
Original motif

Consensus sequence: BDHHMCGCCCCCTHVBB



Reverse complement motif

Consensus sequence: BVVHAGGGGGGCGRDHBB



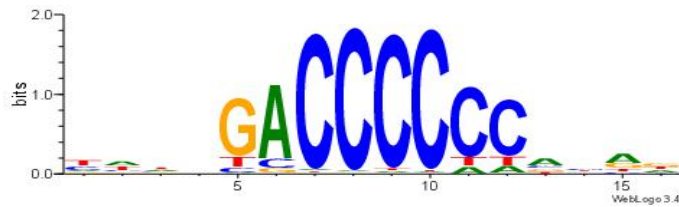
Motif ID: [UP00024](#)  
Motif name: Glis2\_primary  
Species: Mus musculus  
Matching format of first motif: Reverse Complement  
Matching format of second motif: Original Motif  
Direction: Forward  
Position number: 5  
Number of overlap: 8  
Similarity score: 0.0291232

Alignment:

HHHVGACCCCCCVBRD  
----GCCCCGCC----

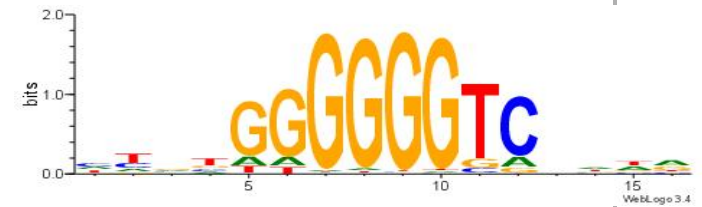
Original motif

Consensus sequence: HHHVGACCCCCVBRD



Reverse complement motif

Consensus sequence: HKBBGGGGGGTCVHHH



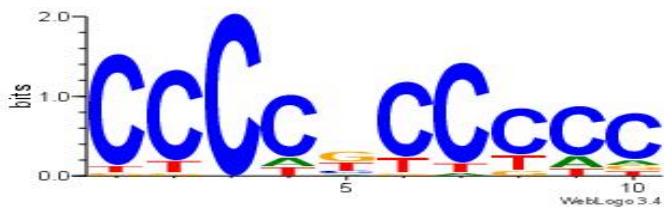
**Dataset #: 3**

**Motif ID: 24**

**Motif name: SP1**

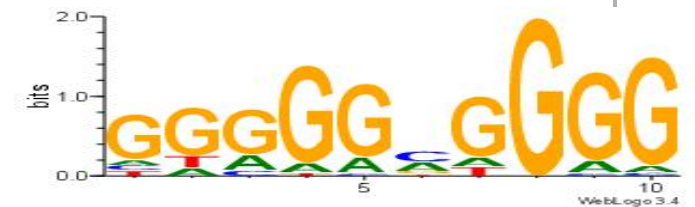
Original motif

Consensus sequence: CCCCKCCCC



Reverse complement motif

Consensus sequence: GGGGGYGGGG



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

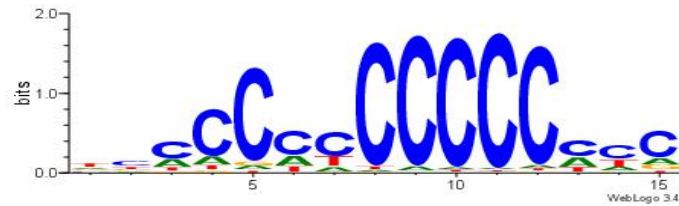
Motif ID:	<a href="#">UP00021</a>
Motif name:	Zfp281_primary
Species:	Mus musculus
Matching format of first motif:	Original Motif
Matching format of second motif:	Original Motif
Direction:	Backward
Position number:	4
Number of overlap:	10
Similarity score:	0.0231018

Alignment:

```
DBCCCCCCCCCMYC  
--CCCCKCCCCC---
```

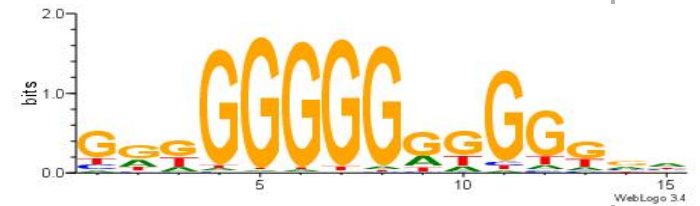
Original motif

Consensus sequence: DBCCCCCCCCCMYC



Reverse complement motif

Consensus sequence: GKRGGGGGGGGGGBD



---

Motif ID:	<a href="#">UP00093</a>
Motif name:	Klf7_primary
Species:	Mus musculus
Matching format of first motif:	Original Motif
Matching format of second motif:	Original Motif
Direction:	Backward
Position number:	3
Number of overlap:	10
Similarity score:	0.0303207

Alignment:

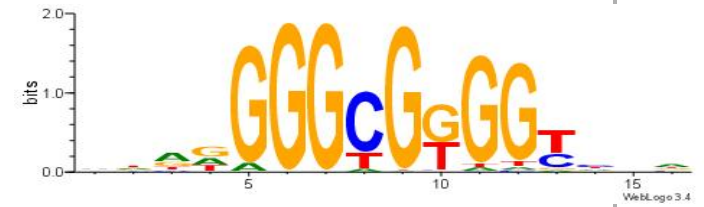
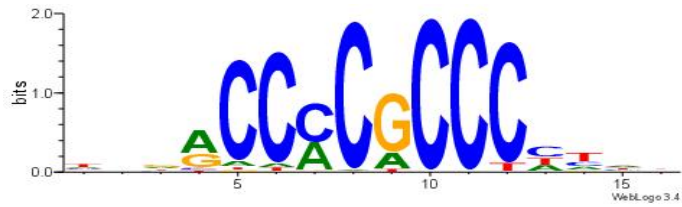
```
HBDRCCMCGCCCHHHD  
----CCCCKCCCCC--
```

Original motif

Consensus sequence: HBDRCCMCGCCCHHHD

Reverse complement motif

Consensus sequence: DHHDGGGCGRGGKHBH

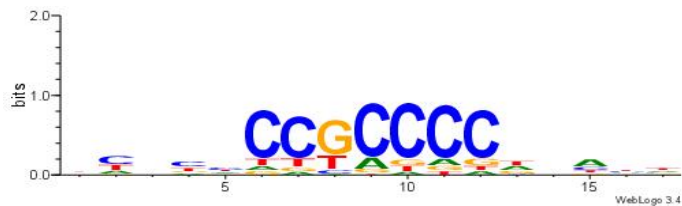


Motif ID: [UP00033](#)  
 Motif name: Zfp410\_secondary  
 Species: Mus musculus  
 Matching format of first motif: Reverse Complement  
 Matching format of second motif: Reverse Complement  
 Direction: Forward  
 Position number: 5  
 Number of overlap: 10  
 Similarity score: 0.0323007

Alignment:  
 HHHHDGGGGCGGDBHDV  
 ----GGGGYGGG----

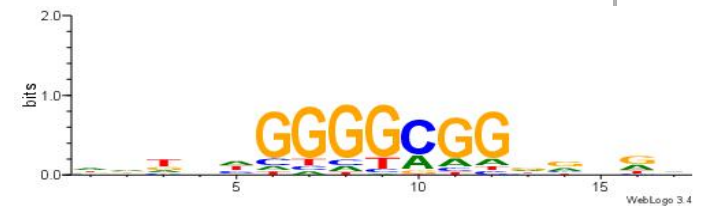
Original motif

Consensus sequence: BHHBHCCGCCCCDHHHH



Reverse complement motif

Consensus sequence: HHHHDGGGGCGGDBHDV



Motif ID: [UP00099](#)



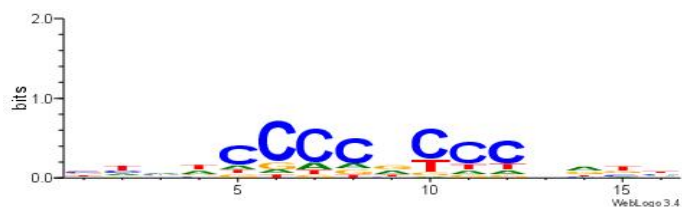
Motif name: Ascl2\_secondary  
 Species: Mus musculus  
 Matching format of first motif: Reverse Complement  
 Matching format of second motif: Reverse Complement  
 Direction: Forward  
 Position number: 3  
 Number of overlap: 10  
 Similarity score: 0.0340973

**Alignment:**

HVDVGGGHGGGGDBHB  
 --GGGGYGGGG-----

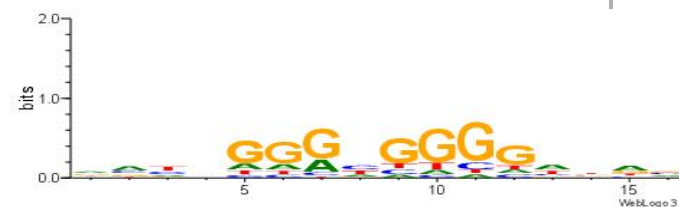
**Original motif**

Consensus sequence: BHVDCCCDCCCBDBH



**Reverse complement motif**

Consensus sequence: HVDVGGGHGGGGDBHB




---

Motif ID: [UP00043](#)  
 Motif name: Bcl6b\_secondary  
 Species: Mus musculus  
 Matching format of first motif: Reverse Complement  
 Matching format of second motif: Reverse Complement  
 Direction: Backward  
 Position number: 3  
 Number of overlap: 10

Similarity score: 0.0351051

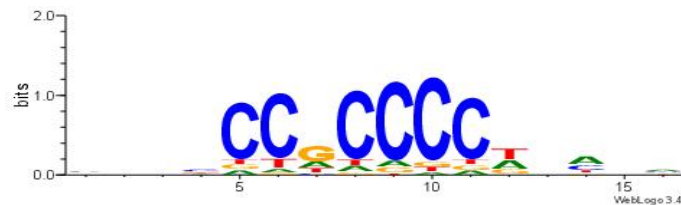
Alignment:

BHHHWGGGGCGBBVH

----GGGGYGGG--

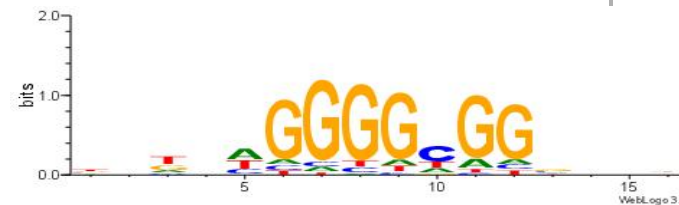
Original motif

Consensus sequence: HBBCCGCCCCWHHHV



Reverse complement motif

Consensus sequence: BHHHWGGGGCGBBVH



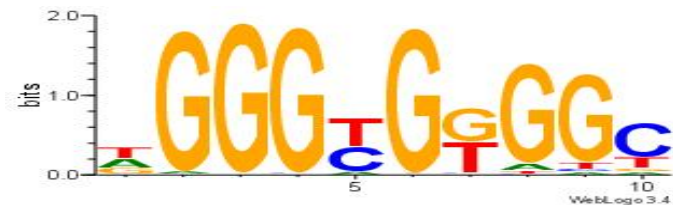
**Dataset #: 3**

**Motif ID: 27**

**Motif name: Klf4**

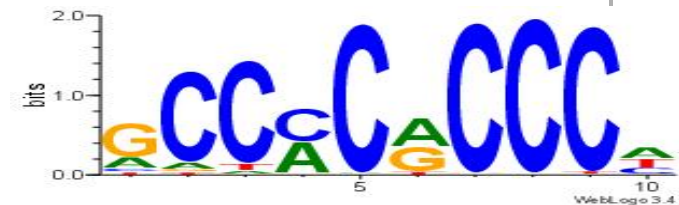
Original motif

Consensus sequence: DGGGYGKGGC



Reverse complement motif

Consensus sequence: GCCYCMCCCD



### Best Matches for Motif ID 27 (Highest to Lowest)

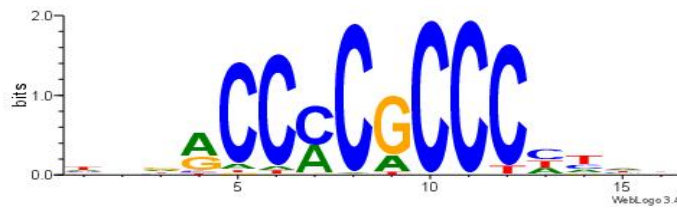
Motif ID: [UP00093](#)  
Motif name: Klf7\_primary  
Species: Mus musculus  
Matching format of first motif: Original Motif

Matching format of second motif: Reverse Complement  
 Direction: Forward  
 Position number: 4  
 Number of overlap: 10  
 Similarity score: 0

Alignment:  
 DHH DGGGCGRGGKHBH  
 ---DGGGYGKGGC---

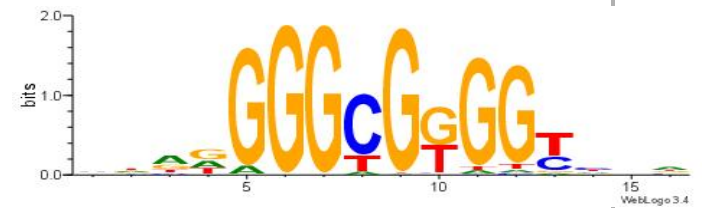
Original motif

Consensus sequence: HBDRCMCGCCCHHHD



Reverse complement motif

Consensus sequence: DHH DGGGCGRGGKHBH




---

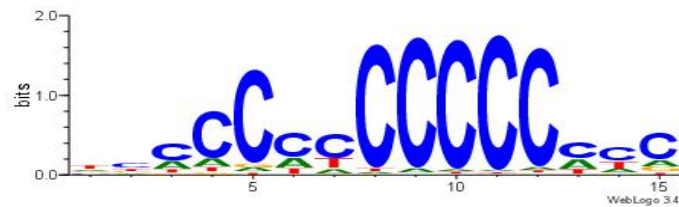
Motif ID: [UP00021](#)  
 Motif name: Zfp281\_primary  
 Species: Mus musculus  
 Matching format of first motif: Reverse Complement  
 Matching format of second motif: Original Motif  
 Direction: Forward  
 Position number: 4  
 Number of overlap: 10  
 Similarity score: 0.0153308

Alignment:

DBCCCCCCCCMYC  
---GCCYCMCCCD---

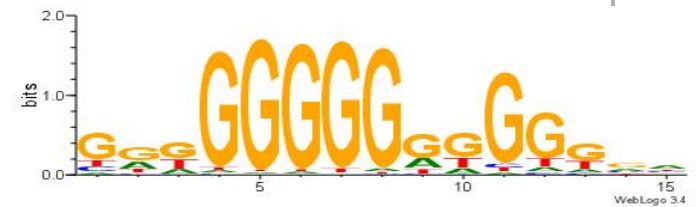
Original motif

Consensus sequence: DBCCCCCCCCMYC



Reverse complement motif

Consensus sequence: GKRGGGGGGGGGGBD



Motif ID: [UP00002](#)  
Motif name: Sp4\_secondary  
Species: Mus musculus  
Matching format of first motif: Original Motif  
Matching format of second motif: Original Motif  
Direction: Forward  
Position number: 3  
Number of overlap: 10  
Similarity score: 0.0186791

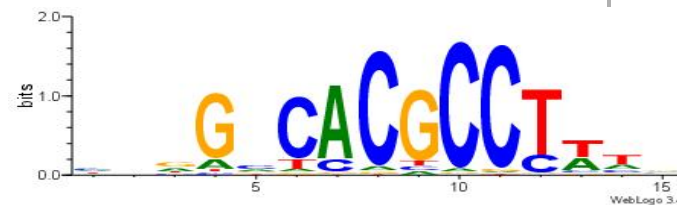
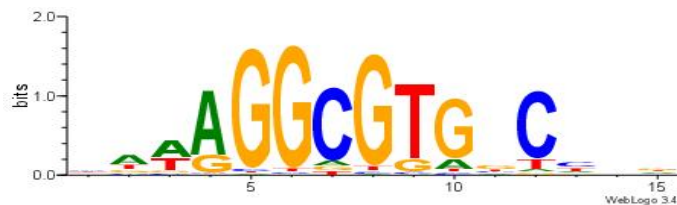
Alignment:  
BDWAGGCGTG BCHHD  
--DGGGYGKGGC---

Original motif

Consensus sequence: BDWAGGCGTG BCHHD

Reverse complement motif

Consensus sequence: HDDGBCACGCCTWDB

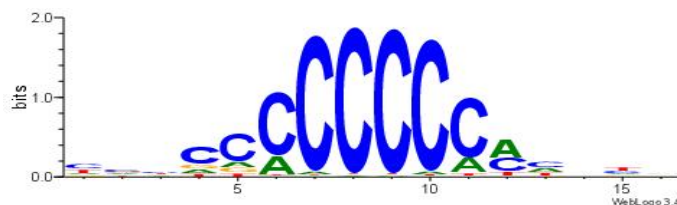


Motif ID: [UP00022](#)  
 Motif name: Zfp740\_primary  
 Species: Mus musculus  
 Matching format of first motif: Reverse Complement  
 Matching format of second motif: Original Motif  
 Direction: Forward  
 Position number: 3  
 Number of overlap: 10  
 Similarity score: 0.0218399

Alignment:  
 HVBCCCCCCCCMHHHB  
 --GCCYCMCCCD----

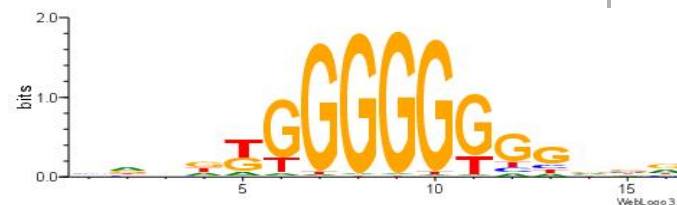
Original motif

Consensus sequence: HVBCCCCCCCCMHHHB



Reverse complement motif

Consensus sequence: BHHDYGGGGGGGBVD



Motif ID: [UP00007](#)

Motif name: Egr1\_primary  
 Species: Mus musculus  
 Matching format of first motif: Reverse Complement  
 Matching format of second motif: Original Motif  
 Direction: Forward  
 Position number: 1  
 Number of overlap: 10  
 Similarity score: 0.0249586

Alignment:  
 VHTGCGGGGGCGGH  
 GCCYCMCCCD-----

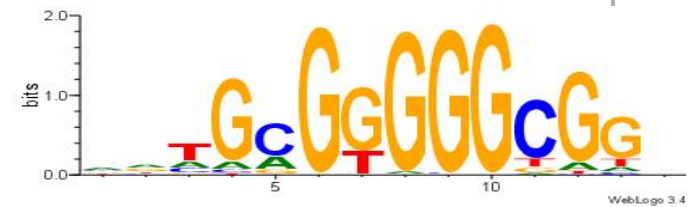
Original motif

Consensus sequence: HCCGCCCCCGCAHB



Reverse complement motif

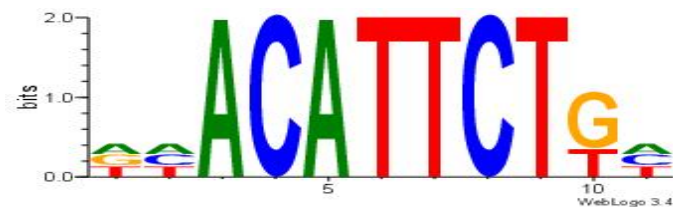
Consensus sequence: VHTGCGGGGGCGGH



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

Original motif

Consensus sequence: DHACATTCTGH



Reverse complement motif

Consensus sequence: HCAGAATGTHD



## Best Matches for Motif ID 44 (Highest to Lowest)

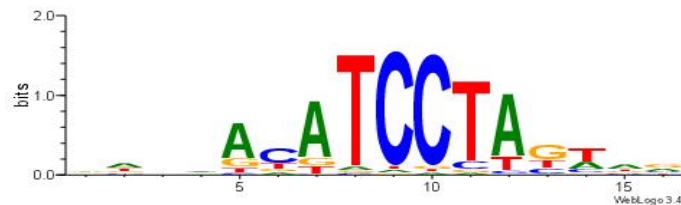
Motif ID: [UP00038](#)  
Motif name: Spdef\_secondary  
Species: Mus musculus  
Matching format of first motif: Reverse Complement  
Matching format of second motif: Reverse Complement  
Direction: Forward  
Position number: 4  
Number of overlap: 11  
Similarity score: 0.00581553

### Alignment:

```
VDWYTAGGATGTHVDH  
---HCAGAATGTHD---
```

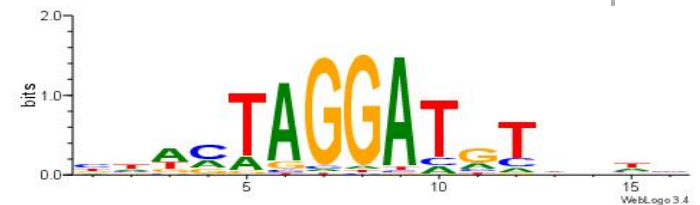
### Original motif

Consensus sequence: DDBHACATCCTAKWDV



### Reverse complement motif

Consensus sequence: VDWYTAGGATGTHVDH



---

Motif ID: [UP00528](#)  
Motif name: Foxm1\_secondary  
Species: Mus musculus  
Matching format of first motif: Original Motif

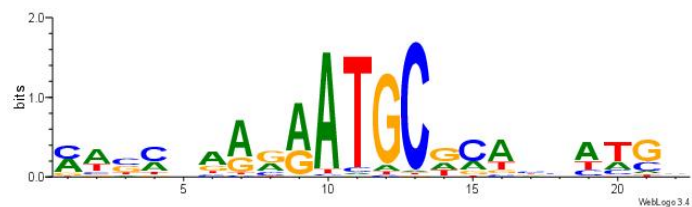
Matching format of second motif: Reverse Complement  
 Direction: Forward  
 Position number: 8  
 Number of overlap: 11  
 Similarity score: 0.00582895

Alignment:

DCATDBTGHGCATKMTTBRSWR  
 -----DHACATTCTGH-----

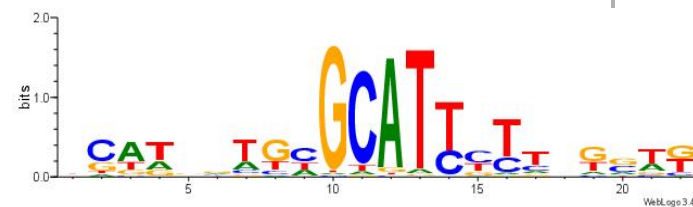
Original motif

Consensus sequence: MWSMBAARRATGCD CABHATGD



Reverse complement motif

Consensus sequence: DCATDBTGHGCATKMTTBRSWR



Motif ID: [UP00023](#)  
 Motif name: Sox30\_primary  
 Species: Mus musculus  
 Matching format of first motif: Original Motif  
 Matching format of second motif: Reverse Complement  
 Direction: Backward  
 Position number: 4  
 Number of overlap: 11  
 Similarity score: 0.00613057

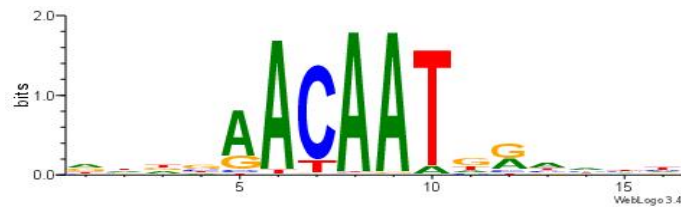
Alignment:



HHHHMHATTGTTBDHD  
--DHACATTCTGH---

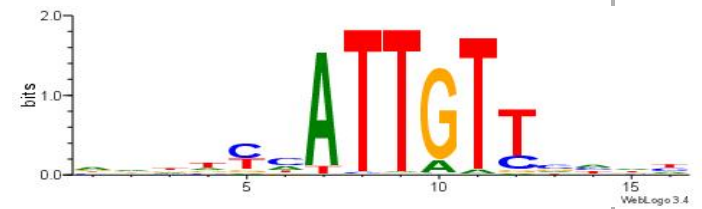
Original motif

Consensus sequence: DHDBAACAATDRHHHH



Reverse complement motif

Consensus sequence: HHHMHATTGTTBDHD



Motif ID: [UP00034](#)  
Motif name: Sox7\_primary  
Species: Mus musculus  
Matching format of first motif: Original Motif  
Matching format of second motif: Reverse Complement  
Direction: Forward  
Position number: 6  
Number of overlap: 11  
Similarity score: 0.0119275

Alignment:

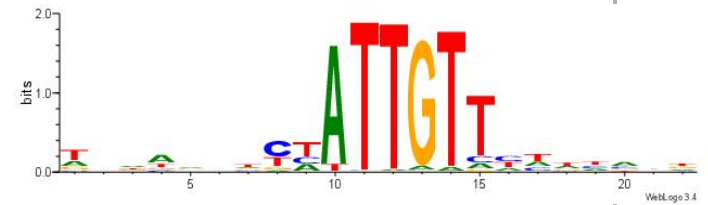
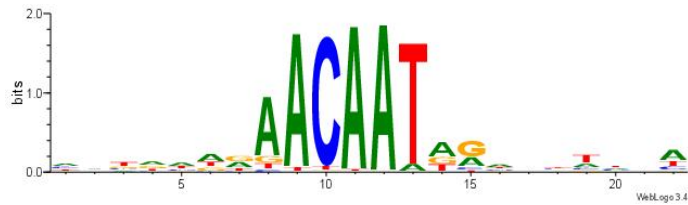
WBHHHHHMDATTGTTDHDVHH  
-----DHACATTCTGH-----

Original motif

Consensus sequence: HHBDHDDAACAATDRHHHHHBW

Reverse complement motif

Consensus sequence: WBHHHHHMDATTGTTDHDVHH

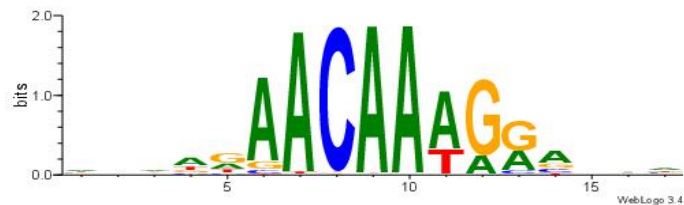


Motif ID: [UP00030](#)  
 Motif name: Sox11\_primary  
 Species: Mus musculus  
 Matching format of first motif: Original Motif  
 Matching format of second motif: Reverse Complement  
 Direction: Forward  
 Position number: 3  
 Number of overlap: 11  
 Similarity score: 0.0122803

Alignment:  
 HHBTMCTTTGTTMDDVH  
 --DHACATTCTGH----

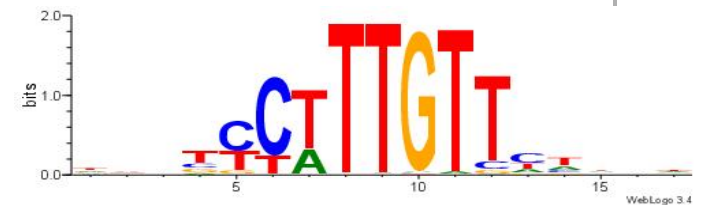
Original motif

Consensus sequence: HBDDRAACAAAGRABHH



Reverse complement motif

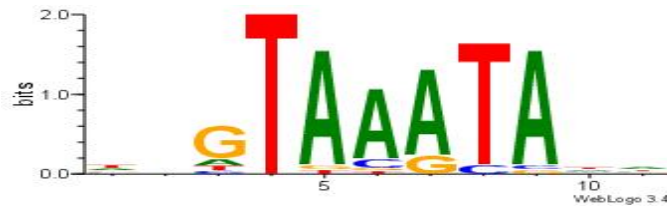
Consensus sequence: HHBTMCTTTGTTMDDVH



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

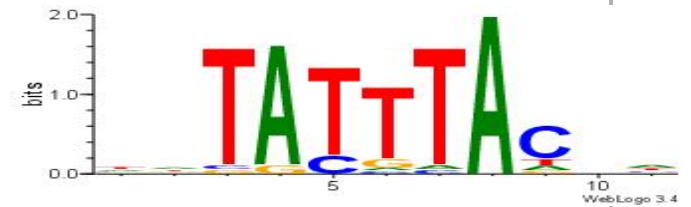
Original motif

Consensus sequence: DBGTAATAHD



Reverse complement motif

Consensus sequence: DHTATTTACBD



### Best Matches for Motif ID 45 (Highest to Lowest)

Motif ID: [UP00242](#)  
Motif name: Hoxc8  
Species: Mus musculus  
Matching format of first motif: Reverse Complement  
Matching format of second motif: Reverse Complement  
Direction: Forward  
Position number: 3  
Number of overlap: 11  
Similarity score: 0

Alignment:

AHDHTAATTACHHHHV

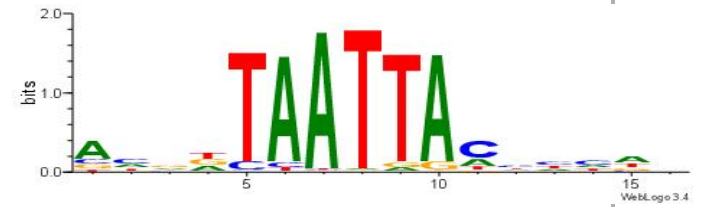
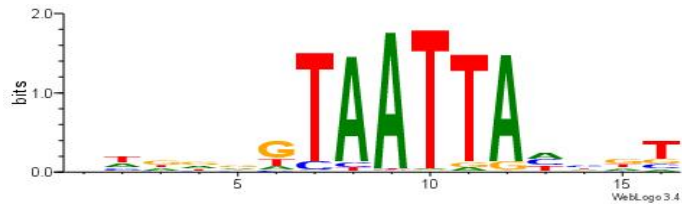
--DHTATTTACBD---

Original motif

Consensus sequence: BHDDDGTAAATTAHHDT

Reverse complement motif

Consensus sequence: AHDHTAATTACHHHHV



Motif ID: [UP00073](#)  
 Motif name: Foxa2\_primary  
 Species: Mus musculus  
 Matching format of first motif: Original Motif  
 Matching format of second motif: Original Motif  
 Direction: Backward  
 Position number: 4  
 Number of overlap: 11  
 Similarity score: 0.00453517

Alignment:  
 HHWAWGTAAAYAAAVDB  
 ---DBGTAAATAHD---

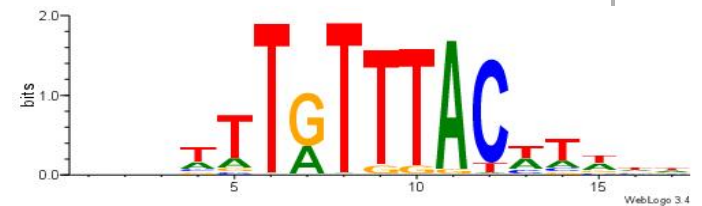
Original motif

Consensus sequence: HHWAWGTAAAYAAAVDB



Reverse complement motif

Consensus sequence: BDVTTTKTTTACWTWHH



Motif ID: [UP00039](#)

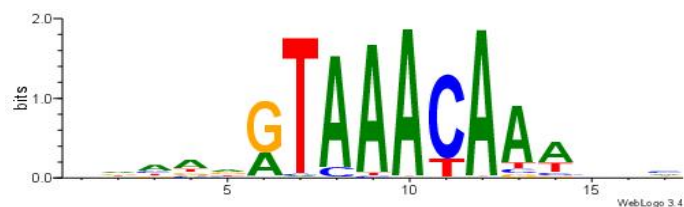
Motif name: Foxj3\_primary  
Species: Mus musculus  
Matching format of first motif: Original Motif  
Matching format of second motif: Original Motif  
Direction: Forward  
Position number: 4  
Number of overlap: 11  
Similarity score: 0.00490457

Alignment:

HDHADGTAAACAAAVVH  
---DBGTAAATAHD---

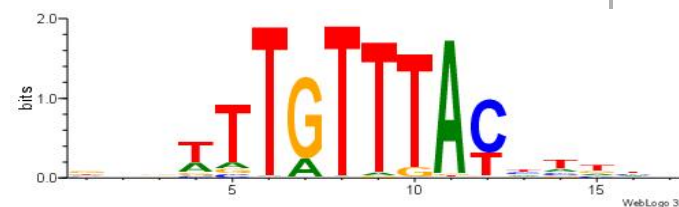
Original motif

Consensus sequence: HDHADGTAAACAAAVVH



Reverse complement motif

Consensus sequence: DVVTTTGTTCACDTHDH



---

Motif ID: [UP00041](#)  
Motif name: Foxj1\_primary  
Species: Mus musculus  
Matching format of first motif: Original Motif  
Matching format of second motif: Original Motif  
Direction: Backward  
Position number: 5  
Number of overlap: 11

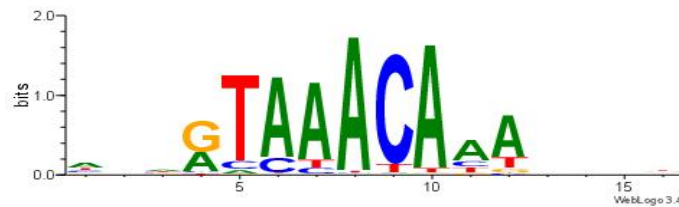
Similarity score: 0.00656547

Alignment:

HDHRTAAACAAHDDD  
-DBGTAAATAHD----

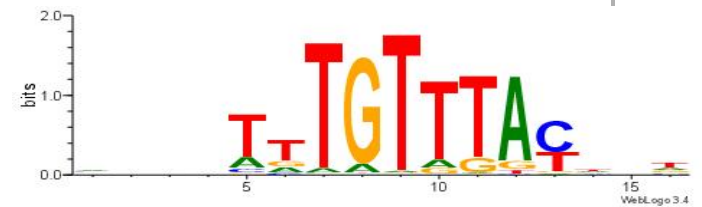
Original motif

Consensus sequence: HDHRTAAACAAHDDD



Reverse complement motif

Consensus sequence: DDDHTTTGTTTAMHDD



Motif ID: [UP00182](#)  
Motif name: Hoxa6  
Species: Mus musculus  
Matching format of first motif: Original Motif  
Matching format of second motif: Reverse Complement  
Direction: Backward  
Position number: 3  
Number of overlap: 11  
Similarity score: 0.0086354

Alignment:

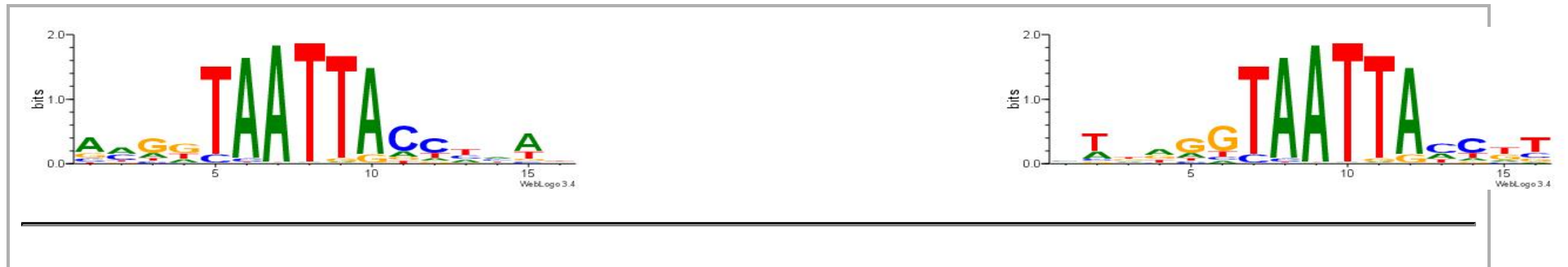
DTHHGGTAATTAYCYT  
---DBGTAAATAHD--

Original motif

Consensus sequence: AMGKTAATTACCHHAD

Reverse complement motif

Consensus sequence: DTHHGGTAATTAYCYT



Results created by MOTIFSIM on 02-03-2023 09:15:18

Runtime: 362.144 seconds

MOTIFSIM is written by Ngoc Tam L. Tran

Motif logo generated by [weblogo](#)