

MEME workshop Wits May 2013

Philip Machanick
Department of Computer Science, Rhodes University



RUBi
Rhodes University Bioinformatics
Research Group



RHODES UNIVERSITY
Where leaders learn

Who I am

- Rhodes Computer Science associate professor
- previously worked with MEME author Tim Bailey, Institute for Molecular Bioscience, University of Queensland, Australia
- worked on MEME extensions and new web services including MEME-ChIP and CentriMo

IMB



Dr Tim Bailey
8 group members

400 research staff and students
A\$80-million annual budget
<http://imb.uq.edu.au/>

Part I: Overview

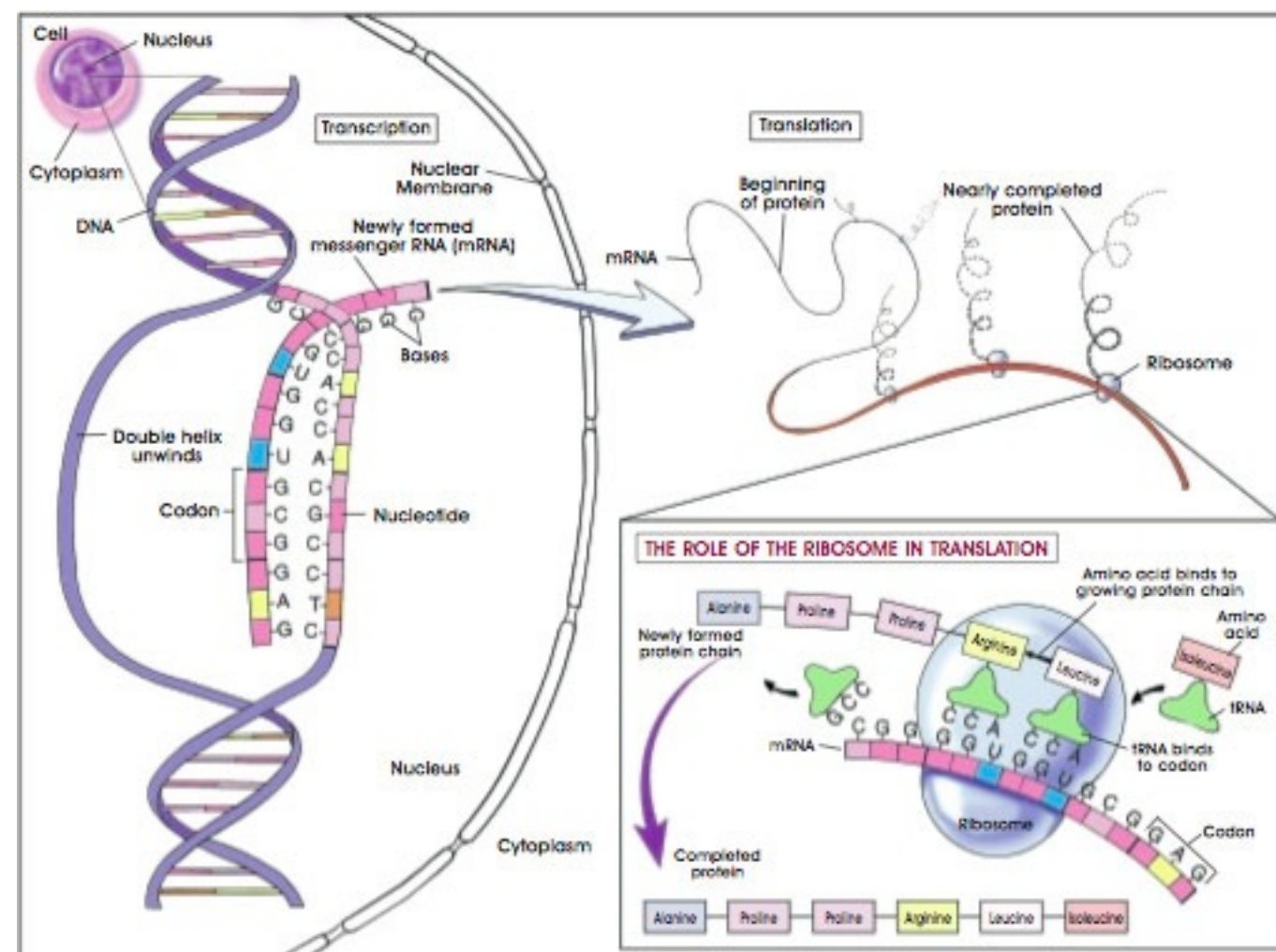
Focus of MEME tools

- motif analysis
 - mostly transcription factors
 - protein motifs also supported

TF binding can get quite complex

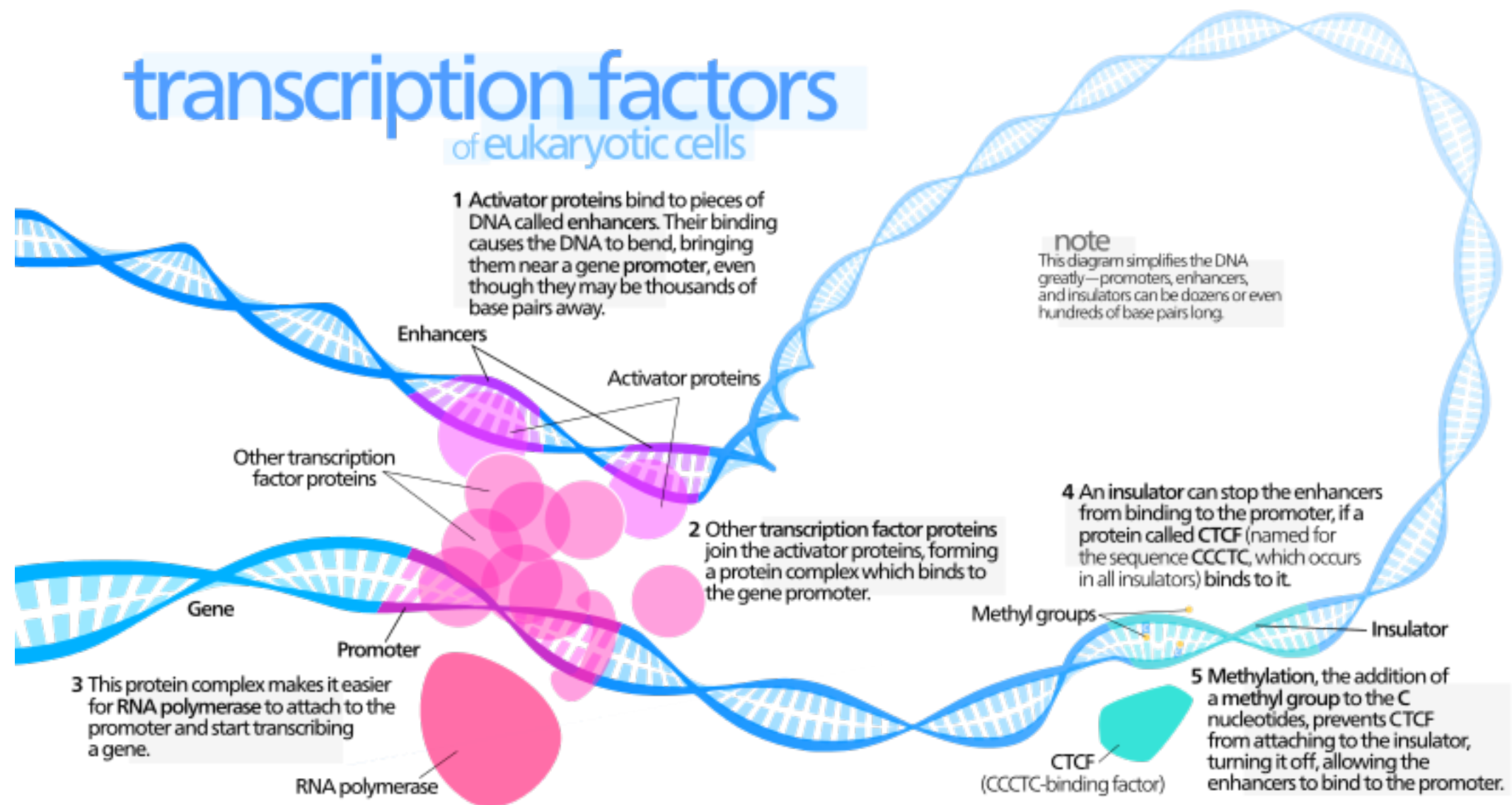
- focus here on simpler cases
 - single TF easily identified
 - pairs of cofactors
 - but also consider indirect binding

Transcription overview



Stem Cell Information, Appendix A: Early Development, NIH
<http://stemcells.nih.gov/info/2001report/pages/appendixA.aspx>

Transcription complexity



http://en.wikipedia.org/wiki/Transcription_factor

Tasks

- characterise putative binding sites
 - a *motif* is a model of binding sites
- given a specific TF, identify genomic regions where it is up or down regulated
- determine what other TFs are involved in binding

What is a motif?

- numeric representation of the probability of a given base at each position in a binding site

Motif by example

Motif by example

aligned sequences

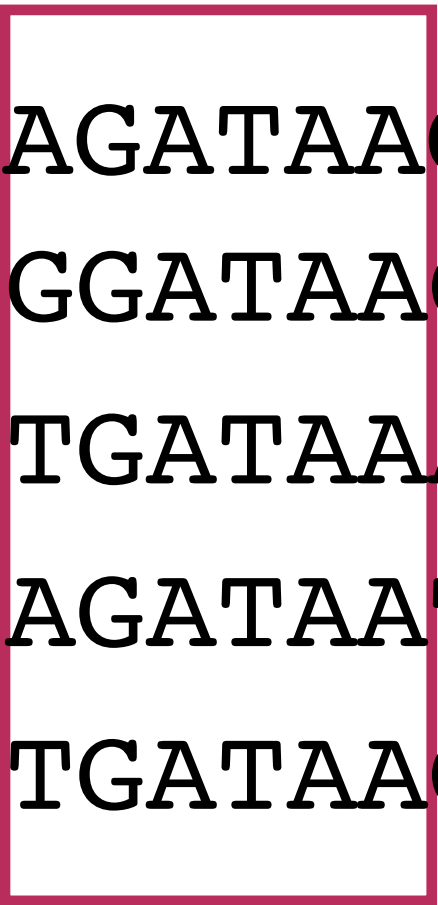
Motif by example

aligned sequences

```
GCAGATAAGCCGCA  
CCGGATAACCTGAT  
TCTGATAAAATCTG  
CCAGATAATATACA  
ACTGATAACAGATG
```

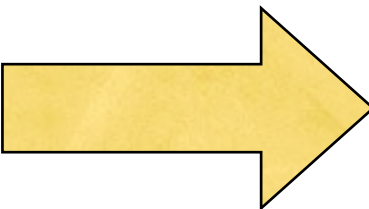

Motif by example

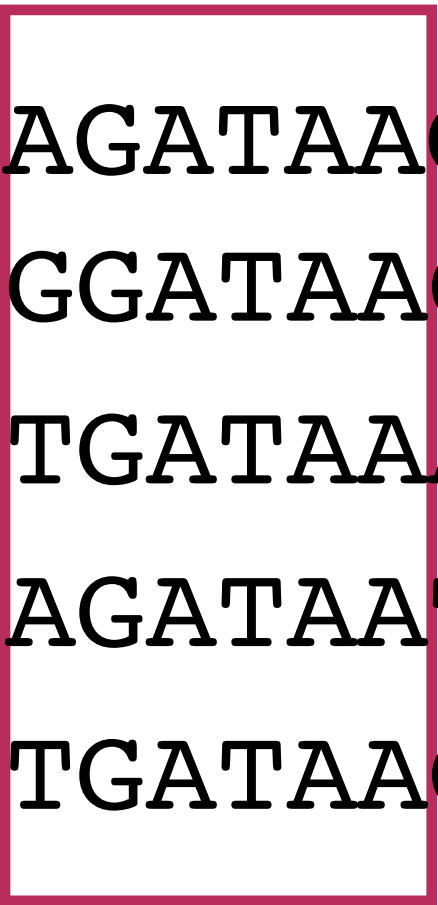
aligned sequences



GCAGATAAGCCGCA
CCGGATAACCTGAT
TCTGATAAAATCTG
CCAGATAATATACA
ACTGATAACAGATG

Motif by example

aligned sequences 



```
GCAGATAAGCCGCA
CCGGATAACCTGAT
TCTGATAAAATCTG
CCAGATAATATACA
ACTGATAACAGATG
```

Motif by example

aligned sequences  counts

G	C	A	G	A	T	A	A	G	C	C	G	C	A
C	C	G	G	A	T	A	A	C	C	T	G	A	T
T	C	T	G	A	T	A	A	A	A	T	C	T	G
C	C	A	G	A	T	A	A	T	A	T	A	C	A
A	C	T	G	A	T	A	A	C	A	G	A	T	G

Motif by example

aligned sequences  counts

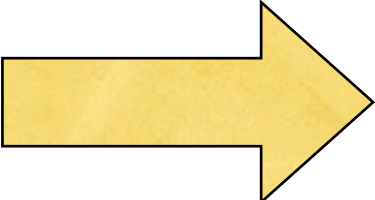
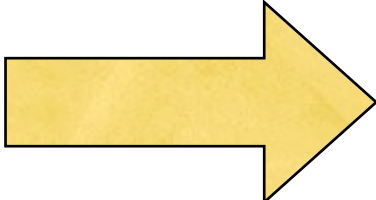
GCAGATAAGCCGCA	A	2	0	5	0	5	5
CCGGATAACCTGAT	C	0	0	0	0	0	0
TCTGATAAAATCTG	G	1	5	0	0	0	0
CCAGATAATATACA	T	2	0	0	5	0	0
ACTGATAACAGATG							

Motif by example

aligned sequences  counts 

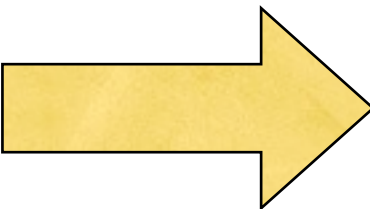
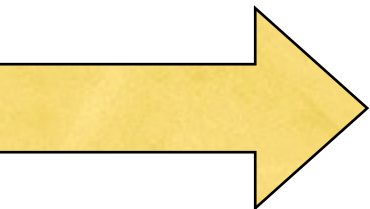
GCAGATAAGCCGCA	A	2	0	5	0	5	5
CCGGATAACCTGAT	C	0	0	0	0	0	0
TCTGATAAAATCTG	G	1	5	0	0	0	0
CCAGATAATATACA	T	2	0	0	5	0	0
ACTGATAACAGATG							

Motif by example

aligned sequences  counts  probabilities

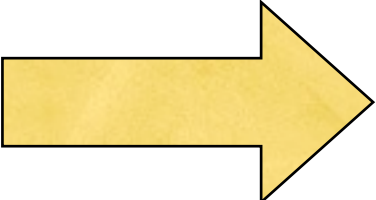
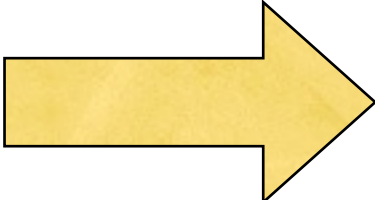
GCAGATAAGCCGCA	A	2	0	5	0	5	5
CCGGATAACCTGAT	C	0	0	0	0	0	0
TCTGATAAAATCTG	G	1	5	0	0	0	0
CCAGATAATATACA	T	2	0	0	5	0	0
ACTGATAACAGATG							

Motif by example

aligned sequences  counts  probabilities

GCAGATAAGCCGCA	A	2	0	5	0	5	5	0.4	0	1	0	1	1
CCGGATAACCTGAT	C	0	0	0	0	0	0	0	0	0	0	0	0
TCTGATAAAATCTG	G	1	5	0	0	0	0	0.2	1	0	0	0	0
CCAGATAATATACA	T	2	0	0	5	0	0	0.4	0	0	1	0	0
ACTGATAACAGATG													

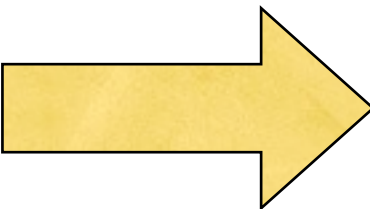
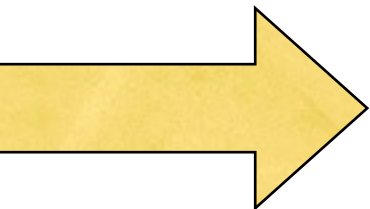
Motif by example

aligned sequences  counts  probabilities

GCAGATAAGCCGCA	A	2	0	5	0	5	5	0.4	0	1	0	1	1
CCGGATAACCTGAT	C	0	0	0	0	0	0	0	0	0	0	0	0
TCTGATAAAATCTG	G	1	5	0	0	0	0	0.2	1	0	0	0	0
CCAGATAATATACA	T	2	0	0	5	0	0	0.4	0	0	1	0	0
ACTGATAACAGATG													

think of either as a
model or ***motif***

Motif by example

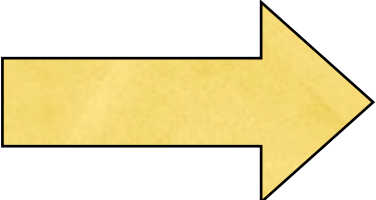
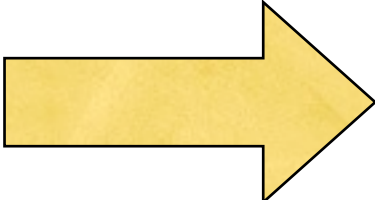
aligned sequences  counts  probabilities

GCAGATAAGCCGCA	A	2	0	5	0	5	5	0.4	0	1	0	1	1
CCGGATAACCTGAT	C	0	0	0	0	0	0	0	0	0	0	0	0
TCTGATAAAATCTG	G	1	5	0	0	0	0	0.2	1	0	0	0	0
CCAGATAATATACA	T	2	0	0	5	0	0	0.4	0	0	1	0	0
ACTGATAACAGATG													

think of either as a
model or **motif**

logo

Motif by example

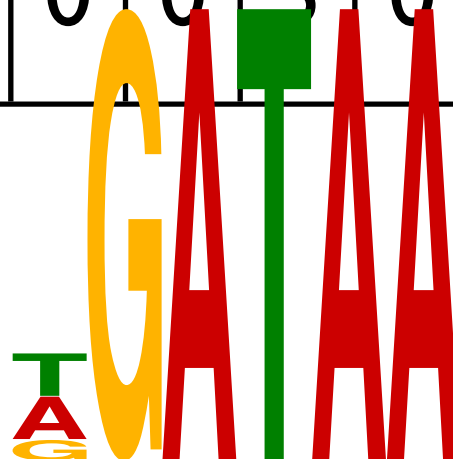
aligned sequences  counts  probabilities

GCAGATAAGCCGCA
CCGGATAACCTGAT
TCTGATAAAATCTG
CCAGATAATATACA
ACTGATAACAGATG

A	2	0	5	0	5	5
C	0	0	0	0	0	0
G	1	5	0	0	0	0
T	2	0	0	5	0	0

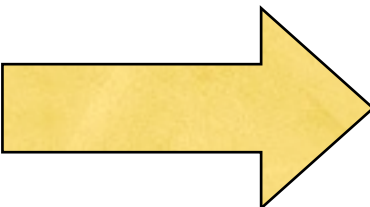
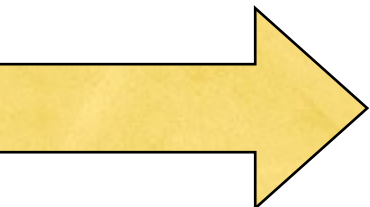
0.4	0	1	0	1	1
0	0	0	0	0	0
0.2	1	0	0	0	0
0.4	0	0	1	0	0

logo



think of either as a
model or ***motif***

Motif by example

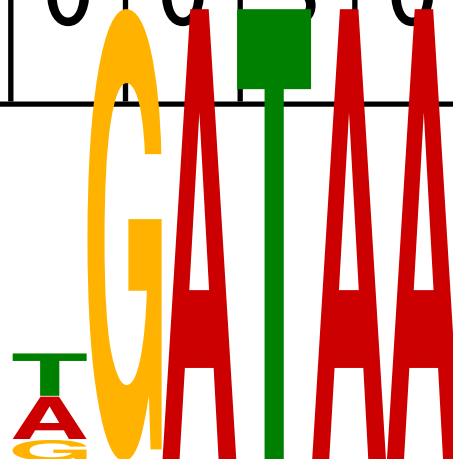
aligned sequences  counts  probabilities

GCAGATAAGCCGCA
CCGGATAACCTGAT
TCTGATAAAATCTG
CCAGATAATATACA
ACTGATAACAGATG

A	2	0	5	0	5	5
C	0	0	0	0	0	0
G	1	5	0	0	0	0
T	2	0	0	5	0	0

0.4	0	1	0	1	1
0	0	0	0	0	0
0.2	1	0	0	0	0
0.4	0	0	1	0	0

logo



height = $2 + \sum \log_2 f_{b,j} \times f_{b,j}$
information content

think of either as a
model or **motif**

What motifs can't do

- cannot model a binding site with variable *spacing*
- cannot model *indirect* binding

Data sources

- TF binding sites
 - older: ChIP-chip
 - more recent: ChIP-seq
 - many public data sets, e.g. ENCODE
- motif databases
- sites of interest in genome

Focus here

- given a sequence set representing binding sites, find the binding motif
- rank a set of motifs for enrichment in a sequence set
- attempt to identify binding partners
- compare motifs for similarity

MEME web service

- wide range of available tools – a sample:
 - MEME, DREME – *de novo* motif finders
 - MAST, FIMO – match motifs to sequences
 - SpaMo, CentriMo – motif enrichment analysis
 - TOMTOM – measure motif similarity

What we'll do

- brief summary of each tool
 - then use a compendium: MEME-ChIP web service
- examine how to solve a specific problem in more detail

Motif discovery

- a set of sequences representing binding regions
- if a TF binds directly, we expect a section of each sequence to stand out from the background

Sources of binding regions

- currently popular method: ChIP-seq (chromatin immunoprecipitation followed by deep sequencing)
- many public data sets
- unfortunately a lot of data needs work to use

ChIP-seq workflow

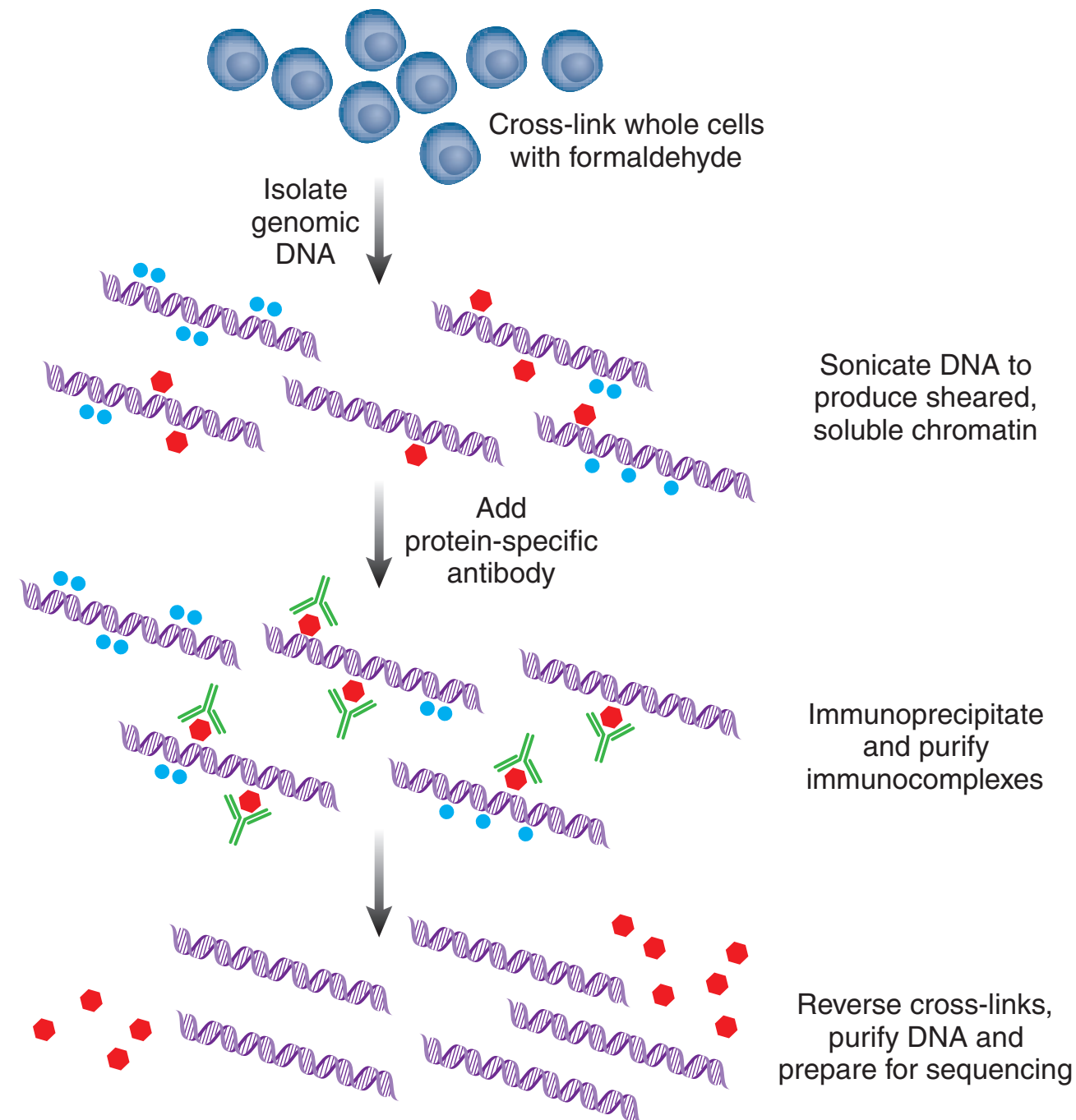


Figure 1 | Workflow of Chip-seq. DNA and proteins are cross-linked and purified; then bound DNA is analyzed by massively parallel short-read sequencing.

For now assume we have binding regions

- each region a short sequence of DNA
- a set of such regions is needed for enough statistical power to find the binding sites

MEME

- finds an alignment of sequences in which there is a site of high statistical similarity across as many sequences as possible
- creates a numeric model of that site
- tests that model (motif) for likelihood it's a valid motif
- refines the model

a largely statistical approach

DREME

- identifies *words* (*k*-mers) with statistically significant over-representation
- combines similar words using *regular expressions*
- keeps going until no significant *k*-mers left

MEME vs DREME

- MEME is much more computation-intensive
 - can find longer motifs
 - but a lot slower
- DREME faster
 - can find many shorter *core motifs*

Examples

MEME

<http://meme.nbcr.net>

a toy example to illustrate principles

Examples

MEME

<http://meme.nbcr.net>

MEME Suite Menu

- + Submit A Job
- + Documentation
- + Downloads
- + User Support
- + Alternate Servers
- Authors
- Citing

The MEME Suite

Motif-based sequence analysis tools

a toy example to illustrate principles

Examples

MEME

<http://meme.nbcr.net>

MEME Suite Menu

- + Submit A Job
- + Documentation
- + Downloads
- + User Support
- + Alternate Servers
- Authors
- Citing

The MEME Suite

Motif-based sequence analysis tools

a toy example to illustrate principles

Examples

MEME

<http://meme.nbcr.net>

MEME Suite Menu

- + Submit A Job
- Documentation
 - Detailed Documentation**
 - Release Notes
 - Installation Guide
 - Command Line Documentation :
 - FAQ
 - Publications

The MEME Suite

Motif-based sequence analysis tools

a toy example to illustrate principles

Examples

MEME

<http://meme.nbcr.net>



The MEME Suite

Motif-based sequence analysis tools

a toy example to illustrate principles

Examples

MEME

<http://meme.nbcr.net>



The MEME Suite

Motif-based sequence analysis tools

Motif Discovery

DREME	Manual	Tutorial	Sample Output
GLAM2	Manual	Tutorial	Sample Output
MEME	Manual	Tutorial	<u>Sample Output</u>

a toy example to illustrate principles

Examples

MEME

<http://meme.nbcr.net>



The MEME Suite

Motif-based sequence analysis tools

Motif Discovery

DREME	Manual	Tutorial	Sample Output
GLAM2	Manual	Tutorial	Sample Output
MEME	Manual	Tutorial	<u>Sample Output</u>

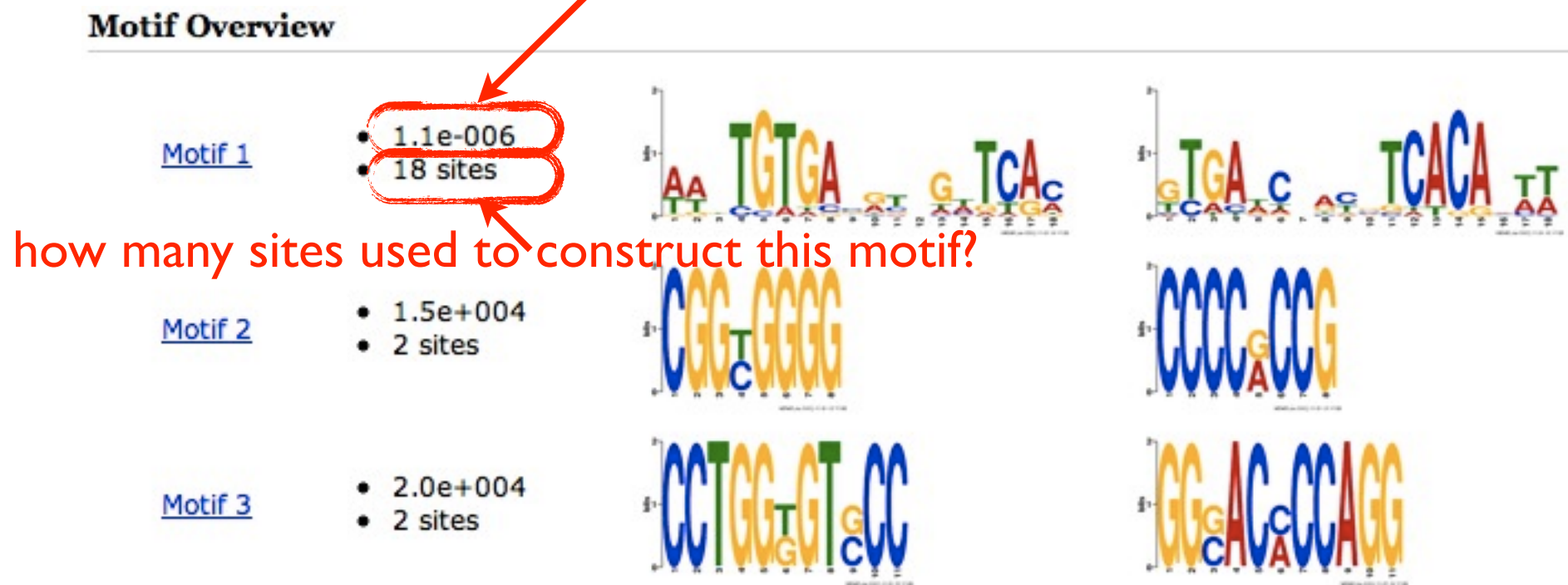
a toy example to illustrate principles

MEME detail

scroll through the example

things to note

E-value: how many like this motif in a random sequence set?



toy example: only 18 sequences

MEME more detail

for each motif




Sites [?](#)

Click on any row to highlight sequence in all motifs.

Name	Strand	Start	p-value	Sites ?
ara	-	58	2.51e-07	TGGCATAGCA AAGTGTGACGCCGTGCAA ATAATCAATG
lac	+	8	5.35e-07	AACGCAAT TAATGTGAGTTAGCTCAC TCATTAGGCA
malt	+	40	8.61e-07	AAAGATTTGG AATTGTGACACAGTGCAA ATTCAGACAC

Block Diagrams [?](#)

The height of the motif "block" is proportional to $-\log(p\text{-value})$, truncated at the height for a motif with a p-value of $1e-10$. Click on any row to highlight sequence in all motifs. Mouse over the center of the motif blocks to see more information.

Name	Lowest p-value	Motif Location
celcg	7.54e-06 \pm	
ara	2.51e-07 \pm	
bglr1	8.30e-05 \pm	
crp	9.89e-06 \pm	
cya	1.81e-05 \pm	

plus more detail, options to invoke other tools

MEME outputs

- motifs based on sequences
 - stats on those motifs
 - details on where they are in the sequences

Examples

DREME

<http://meme.nbcr.net>



a smallish example to illustrate principles

Examples

DREME

<http://meme.nbcr.net>

MEME Suite Menu

- + Submit A Job
- + Documentation
- + Downloads
- + User Support
- + Alternate Servers
- Authors
- Citing

The MEME Suite

Motif-based sequence analysis tools



a smallish example to illustrate principles

Examples

DREME

<http://meme.nbcr.net>

MEME Suite Menu
+ Submit A Job
+ Documentation
+ Downloads
+ User Support
+ Alternate Servers
+ Authors
+ Citing

The MEME Suite
Motif-based sequence analysis tools

a smallish example to illustrate principles

Examples

DREME

<http://meme.nbcr.net>

MEME Suite Menu

- + Submit A Job
- Documentation
 - Detailed Documentation
 - Release Notes
 - Installation Guide
 - Command Line Documentation :
 - FAQ
 - Publications

The MEME Suite

Motif-based sequence analysis tools



a smallish example to illustrate principles

Examples

DREME

<http://meme.nbcr.net>



The MEME Suite
Motif-based sequence analysis tools

a smallish example to illustrate principles

Examples

DREME

<http://meme.nbcrc.net>

MEME Suite Menu

- + Submit A Job
- Documentation
 - Detailed Documentation**
 - Release Notes
 - Installation Guide
 - Command Line Documentation
 - FAQ
 - Publications

The MEME Suite

Motif-based sequence analysis tools

Motif Discovery

DREME	Manual	Tutorial	Sample Output
GLAM2	Manual	Tutorial	Sample Output
MEME	Manual	Tutorial	<u>Sample Output</u>













a smallish example to illustrate principles

DREME detail

scroll through the example

small sample of motifs found:

E-value: Fisher exact test p -value vs. shuffled version of sequences scaled by number of motifs

	Motif ?	Logo ?	RC Logo ?	E-value
1.	CCMCRCCC			3.7e-072
2.	TTATCW			1.3e-029
3.	CACRCMC			7.7e-021
4.	RARGAAA			8.5e-013
5.	TTTTAWW			1.1e-009
6.	CTCCDCCC			3.2e-009

bigger example: 904 sequences

DREME outputs

- motifs based on sequences
- stats on those motifs

less detail than MEME

Analysis of found motifs

- are they similar to known motifs?
- could they have binding partners?

Comparing motifs

- TOMTOM compares motifs to motifs
 - query set: one or more motifs
 - target set: one or more motif databases

Motif Enrichment Analysis (MEA)

- how likely is it that a given motif represents direct binding?
- what other motifs could be involved?

Two options available

- *central* enrichment
 - ChIP-seq binding regions tend to be centred on the binding site
- enriched *spacing*
 - binding partners likely to have a preferred spacing from each other

CentriMo: central motif enrichment analysis

- given
 - binding regions as equal-length sequences
 - database(s) of motifs
- rank the motifs by how strongly their *best site* is centred

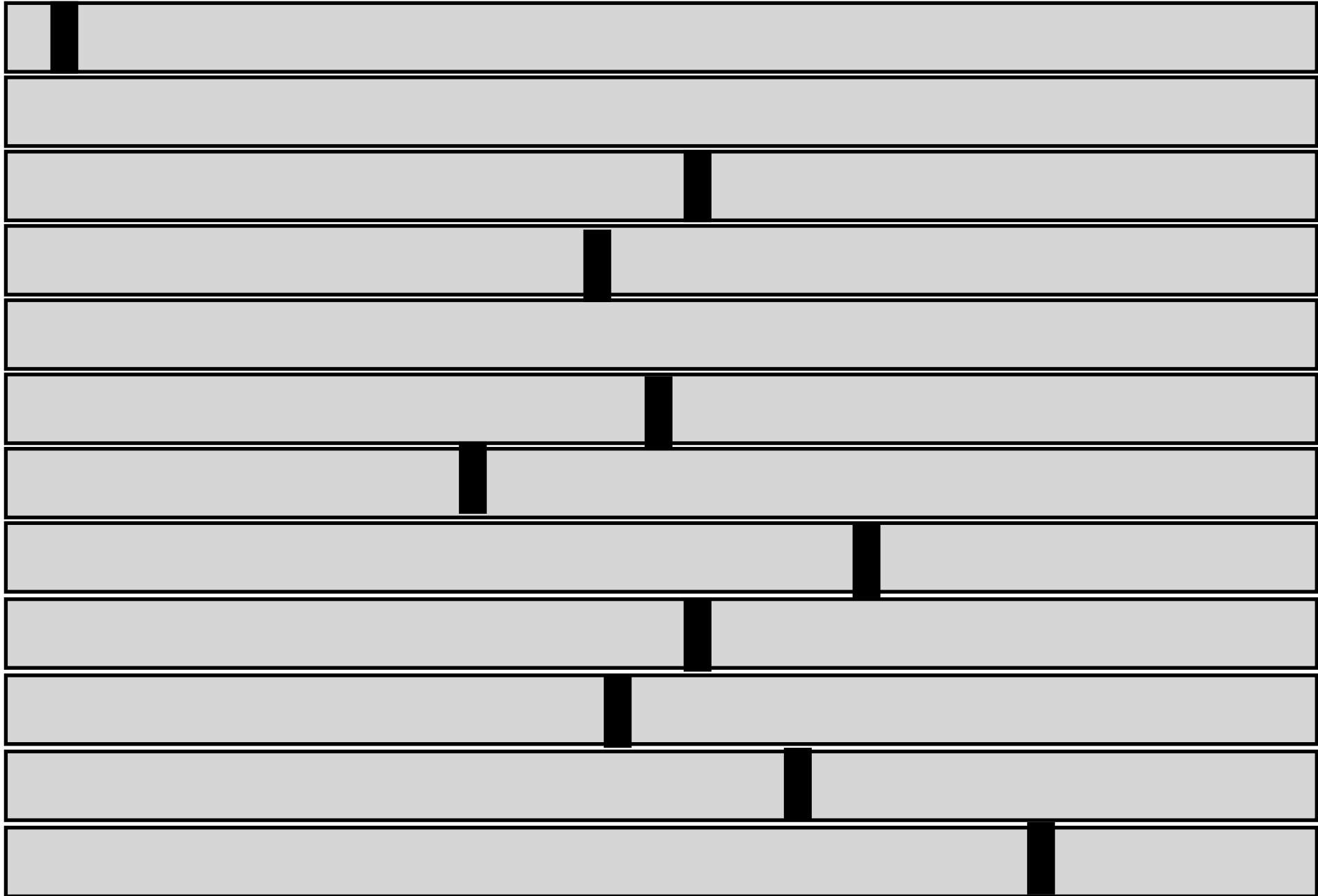
CentriMo statistical approach

- for each motif find the site that best matches it and discard any with *log likelihood ratio* $<$ cutoff (usually 5)
- calculate the binomial *p*-value for the fraction of best sites within a central window

best site



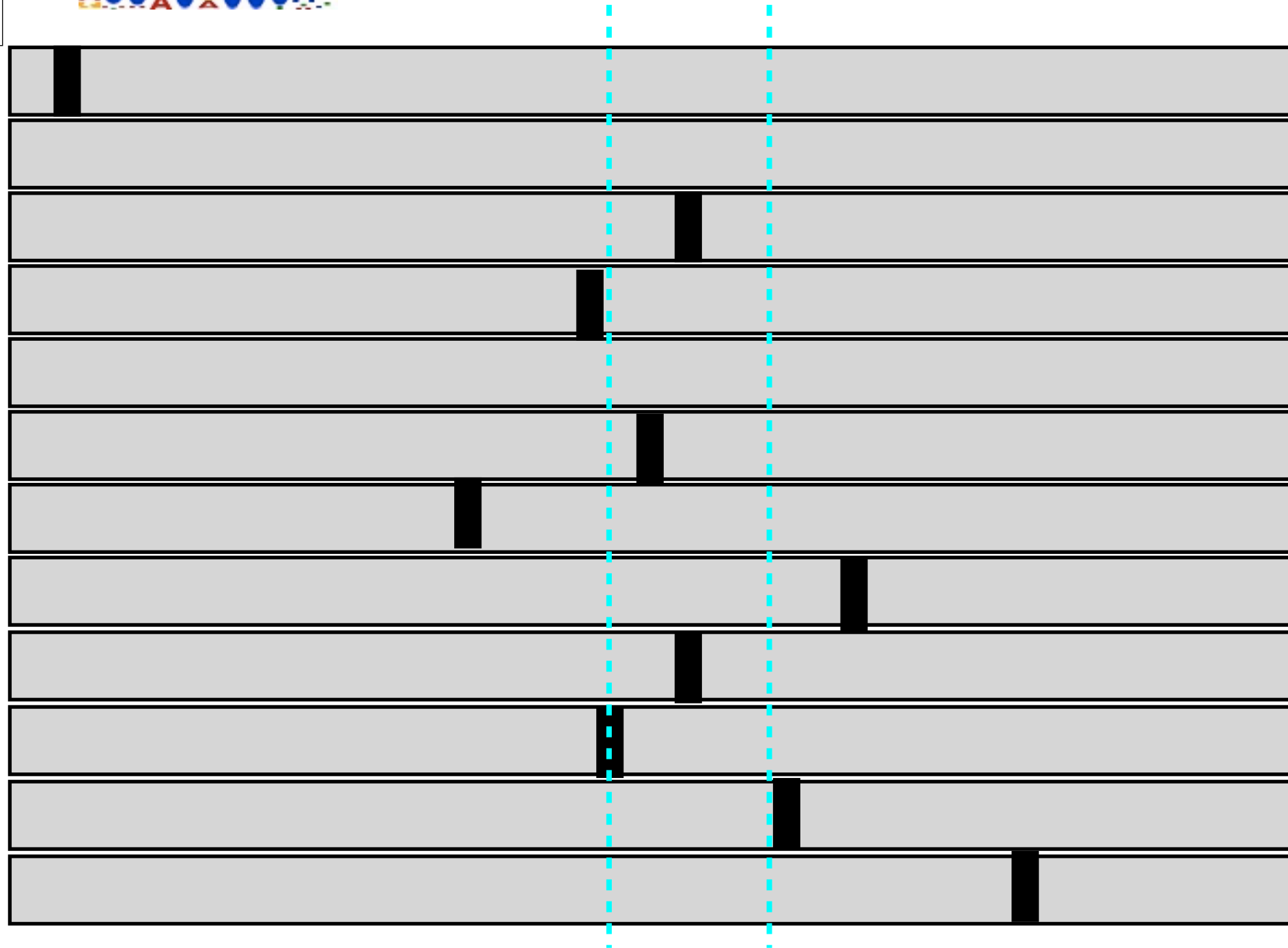
sequences



best site ■



sequences

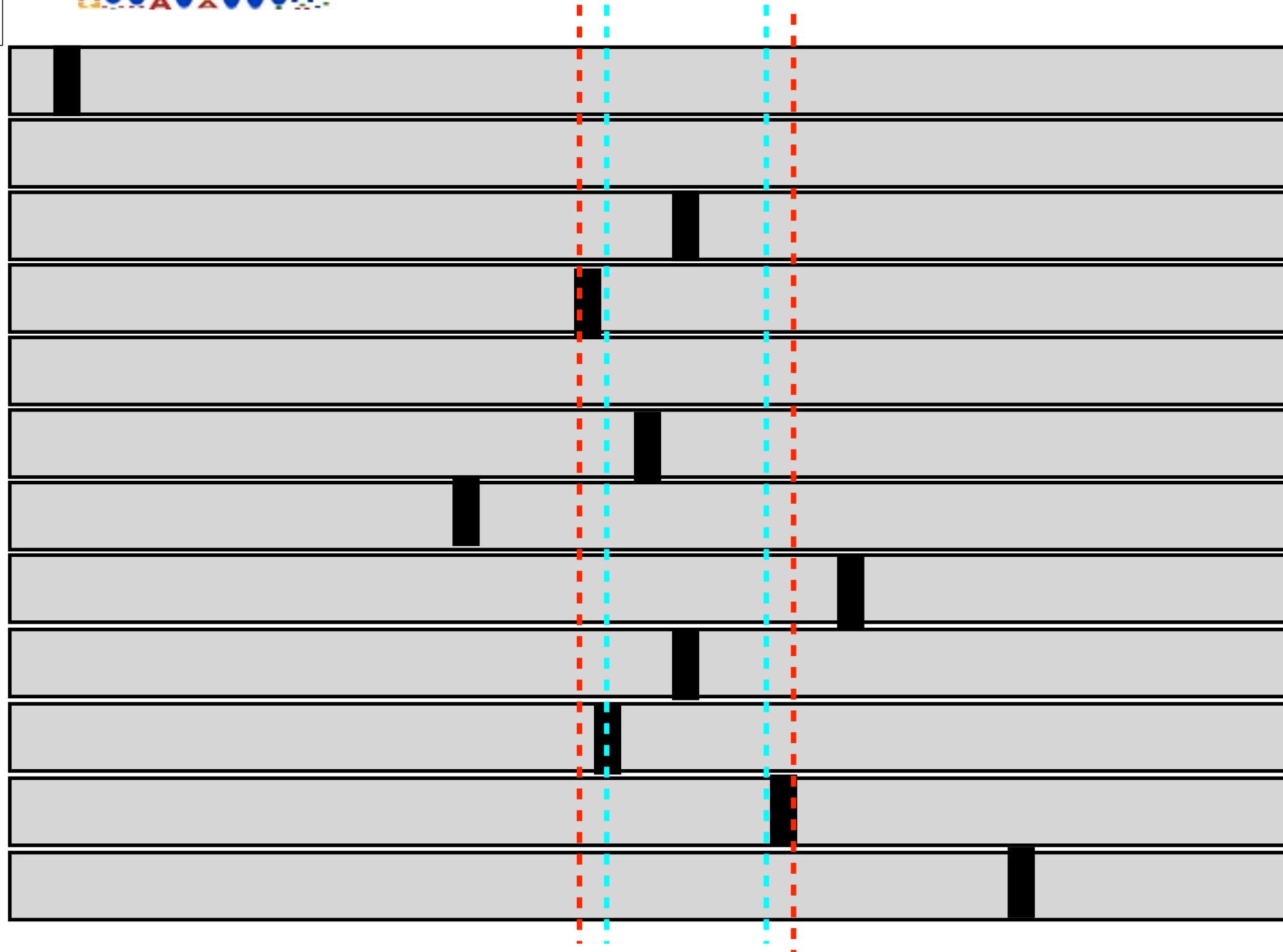


1/8 of sequences in window containing 0.3 of best sites, $p=0.12$

best site ■

CCcCgCCc

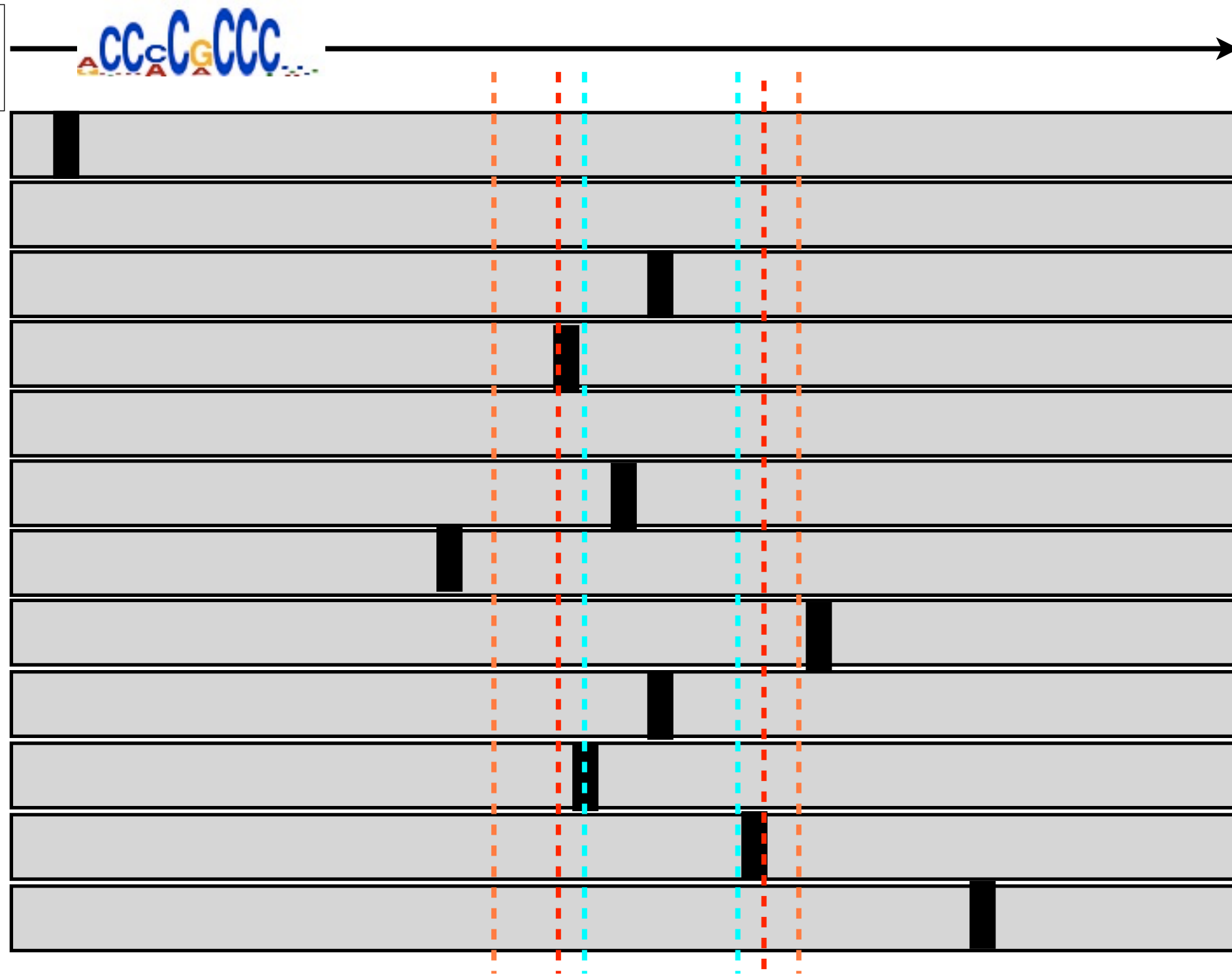
sequences



1/8 of sequences in window containing 0.3 of best sites, $p=0.12$
1/6 of sequences in window containing 0.6 of best sites, $p=0.00244$

best site ■

sequences



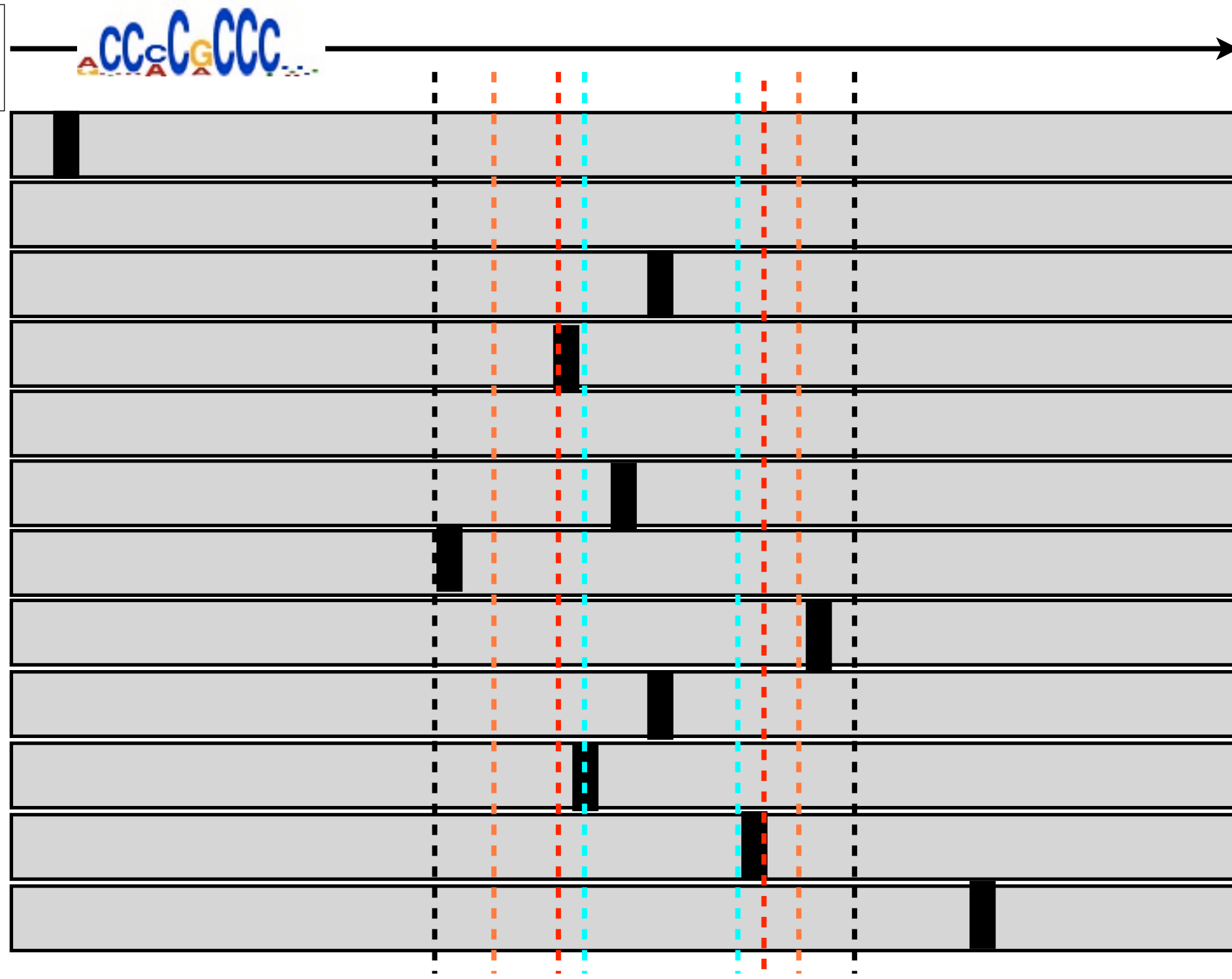
1/8 of sequences in window containing 0.3 of best sites, $p=0.12$

1/6 of sequences in window containing 0.6 of best sites, $p=0.00244$

1/4 of sequences in window containing 0.6 of best sites, $p=0.0197$

best site ■

sequences



1/8 of sequences in window containing 0.3 of best sites, $p=0.12$

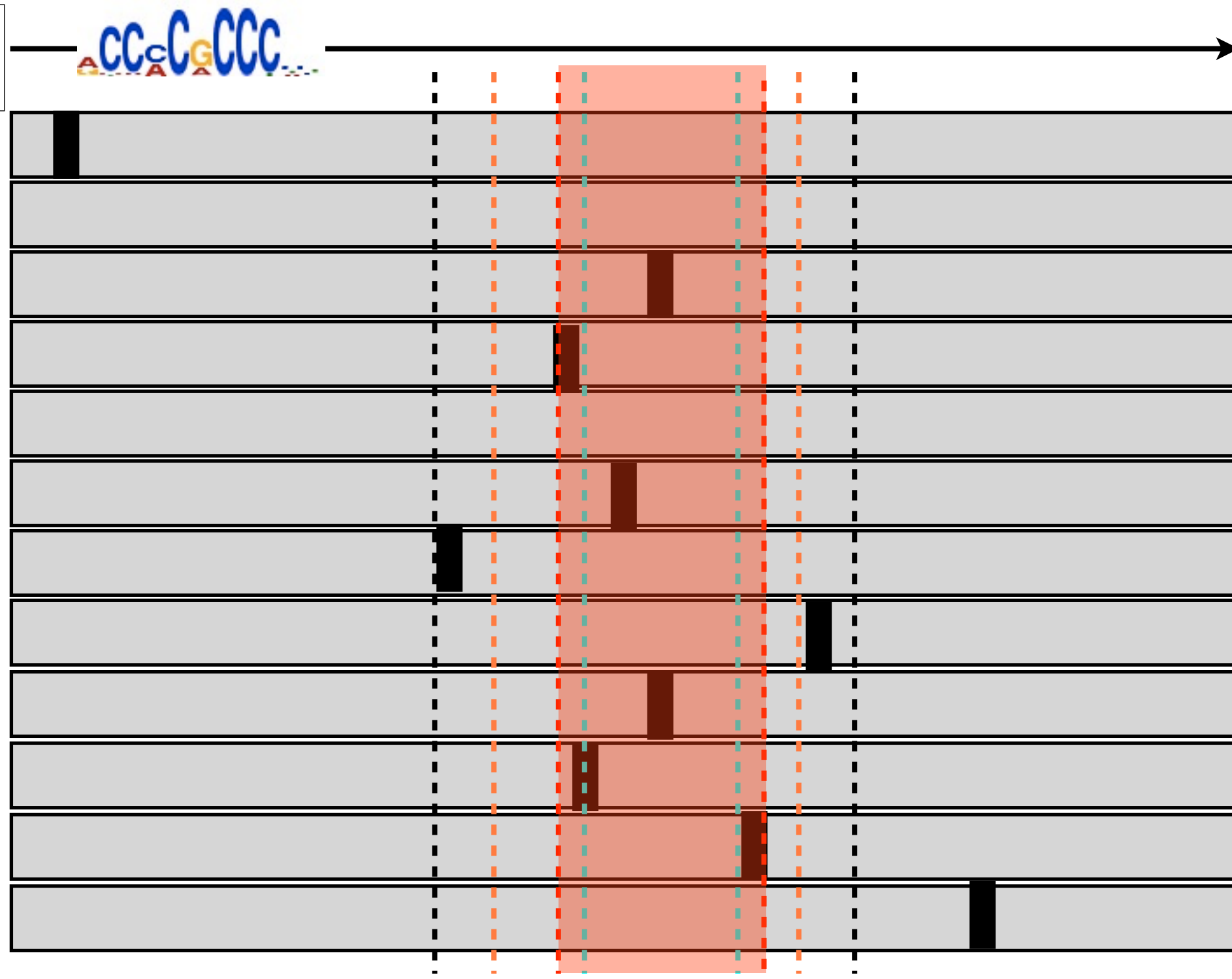
1/6 of sequences in window containing 0.6 of best sites, $p=0.00244$

1/4 of sequences in window containing 0.6 of best sites, $p=0.0197$

1/3 of sequences in window containing 0.8 of best sites, $p=0.0034$

best site ■

sequences



1/8 of sequences in window containing 0.3 of best sites, $p=0.12$

1/6 of sequences in window containing 0.6 of best sites, $p=0.00244$

1/4 of sequences in window containing 0.6 of best sites, $p=0.0197$

1/3 of sequences in window containing 0.8 of best sites, $p=0.0034$

CentriMo outputs

- central enrichment of motifs
 - binomial p -value used to calculate central tendency
 - also look at how many sequences have a significant match to each motif
 - narrowness of lowest p -value window

SpaMo: spaced motif analysis

- given a *primary* motif, a motif database and sequence set
- determine which motifs have a statistically significant preferred spacing from the primary motif

Examples

CentriMo

<http://meme.nbcr.net>



a smallish example to illustrate principles

Examples

CentriMo

<http://meme.nbcr.net>

MEME Suite Menu

- + Submit A Job
- + Documentation
- + Downloads
- + User Support
- + Alternate Servers
- Authors
- Citing

The MEME Suite

Motif-based sequence analysis tools

a smallish example to illustrate principles

Examples

CentriMo

<http://meme.nbcr.net>

MEME Suite Menu

- + Submit A Job
- + Documentation
- + Downloads
- + User Support
- + Alternate Servers
- Authors
- Citing

The MEME Suite

Motif-based sequence analysis tools

a smallish example to illustrate principles

Examples

CentriMo

<http://meme.nbcr.net>

MEME Suite Menu

- + Submit A Job
- Documentation
 - Detailed Documentation
 - Release Notes
 - Installation Guide
 - Command Line Documentation :
 - FAQ
 - Publications

The MEME Suite

Motif-based sequence analysis tools

a smallish example to illustrate principles

Examples

CentriMo

<http://meme.nbcr.net>

MEME Suite Menu

- + Submit A Job
- Documentation
 - Detailed Documentation**
 - Release Notes
 - Installation Guide
 - Command Line Documentation
 - FAQ
 - Publications

The MEME Suite

Motif-based sequence analysis tools

Motif Enrichment Analysis

SpaMo	Manual	Tutorial	Sample Output
CentriMo	Manual	Tutorial	Sample Output

a smallish example to illustrate principles

Examples

CentriMo

<http://meme.nbcr.net>



The MEME Suite

Motif-based sequence analysis tools

Motif Enrichment Analysis

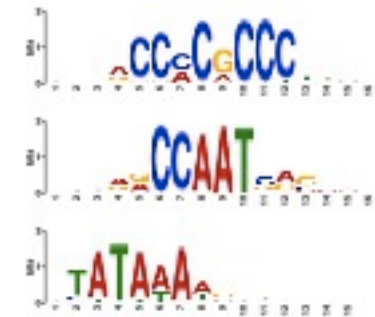
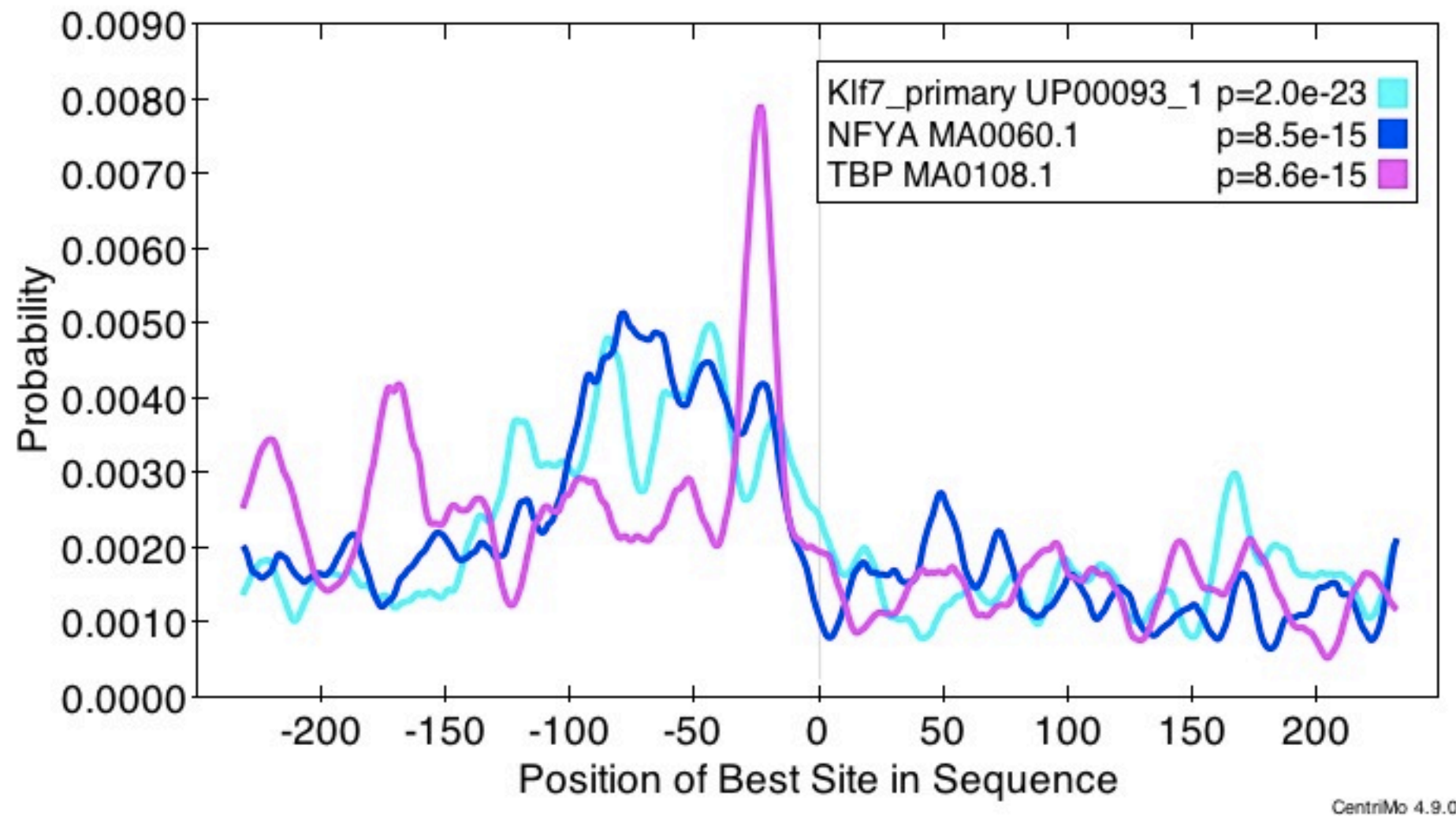
SpaMo	Manual	Tutorial	Sample Output
CentriMo	Manual	Tutorial	Sample Output

a smallish example to illustrate principles

CentriMo detail

scroll through the example

Motif Probability Graph (score ≥ 5 bits) ?



Enriched motifs ($E\text{-value} \leq 1$ using the Binomial test)

this time 1000 sequences

Examples

SpaMo

<http://meme.nbcr.net>

a smallish example to illustrate principles

Examples

SpaMo

<http://meme.nbcr.net>

MEME Suite Menu

- + Submit A Job
- + Documentation
- + Downloads
- + User Support
- + Alternate Servers
- Authors
- Citing

The MEME Suite

Motif-based sequence analysis tools

a smallish example to illustrate principles

Examples

SpaMo

<http://meme.nbcr.net>

MEME Suite Menu

- + Submit A Job
- + Documentation
- + Downloads
- + User Support
- + Alternate Servers
- Authors
- Citing

The MEME Suite

Motif-based sequence analysis tools

a smallish example to illustrate principles

Examples

SpaMo

<http://meme.nbcr.net>

MEME Suite Menu

- + Submit A Job
- Documentation
 - Detailed Documentation
 - Release Notes
 - Installation Guide
 - Command Line Documentation :
 - FAQ
 - Publications

The MEME Suite

Motif-based sequence analysis tools

a smallish example to illustrate principles

Examples

SpaMo

<http://meme.nbcr.net>

MEME Suite Menu

- + Submit A Job
- Documentation
 - Detailed Documentation**
 - Release Notes
 - Installation Guide
 - Command Line Documentation
 - FAQ
 - Publications

The MEME Suite

Motif-based sequence analysis tools

Motif Enrichment Analysis

SpaMo	Manual	Tutorial	Sample Output
CentriMo	Manual	Tutorial	Sample Output

a smallish example to illustrate principles

Examples

SpaMo

<http://meme.nbcr.net>



The MEME Suite

Motif-based sequence analysis tools

Motif Enrichment Analysis

SpaMo	Manual	Tutorial	Sample Output
CentriMo	Manual	Tutorial	Sample Output

a smallish example to illustrate principles

Examples

SpaMo

<http://meme.nbcr.net>



The MEME Suite

Motif-based sequence analysis tools

Motif Enrichment Analysis

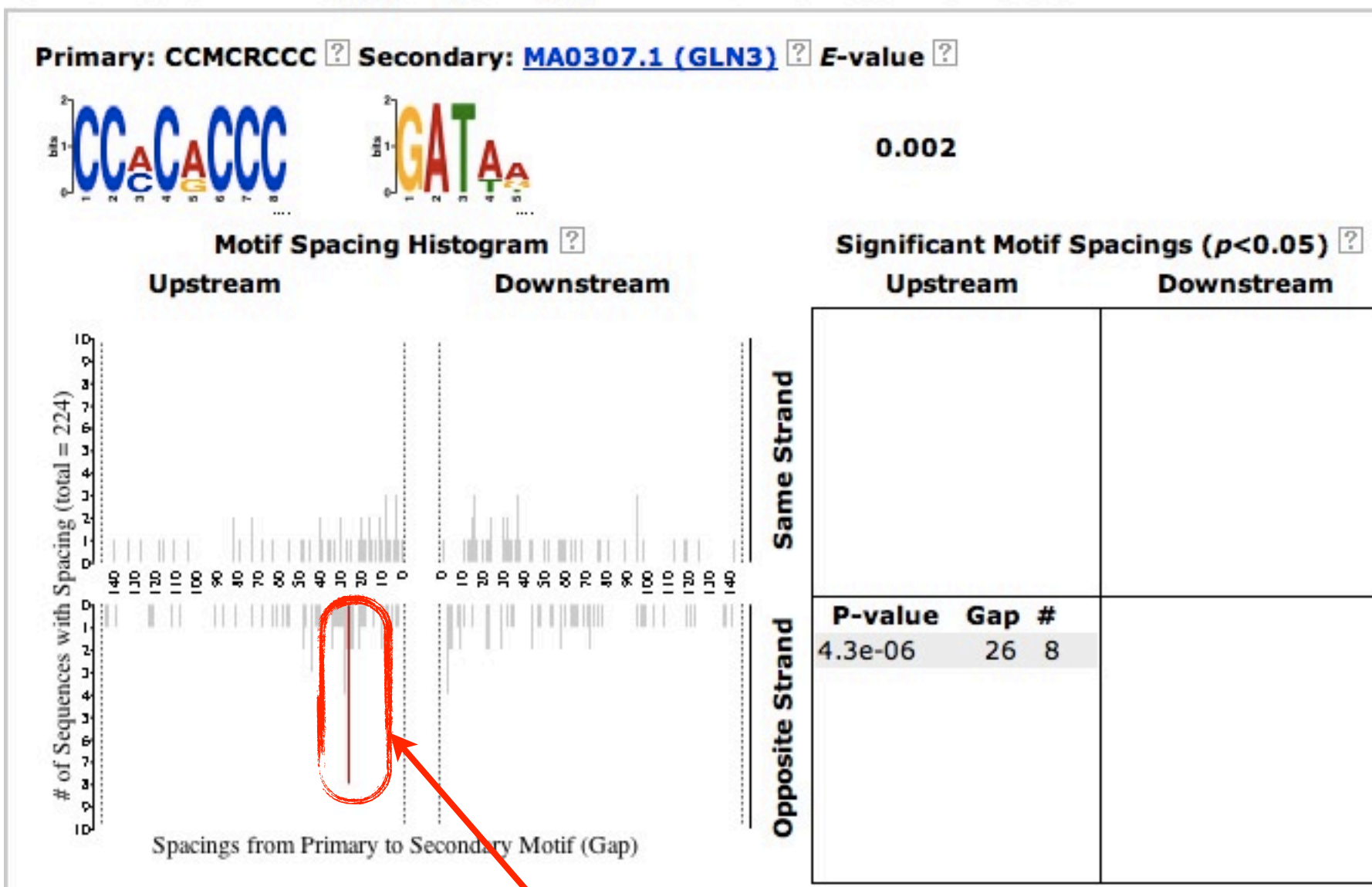
SpaMo	Manual	Tutorial	Sample Output
CentriMo	Manual	Tutorial	Sample Output

a smallish example to illustrate principles

SpaMo detail

scroll through the example

SPACINGS OF "MA0307.1 (GLN3)" RELATIVE TO "CCMCRCRCCC"



TOMTOM: motif comparison

- given a motif and a motif database
 - rank the motif database by match to the given motif (at all possible alignments)

Examples

TOMTOM

<http://meme.nbcr.net>

a smallish example to illustrate principles

Examples

TOMTOM

<http://meme.nbcr.net>

MEME Suite Menu

- + Submit A Job
- + Documentation
- + Downloads
- + User Support
- + Alternate Servers
- Authors
- Citing

The MEME Suite

Motif-based sequence analysis tools

a smallish example to illustrate principles

Examples

TOMTOM

<http://meme.nbcr.net>

MEME Suite Menu

- + Submit A Job
- + Documentation
- + Downloads
- + User Support
- + Alternate Servers
- Authors
- Citing

The MEME Suite

Motif-based sequence analysis tools

a smallish example to illustrate principles

Examples

TOMTOM

<http://meme.nbcr.net>

MEME Suite Menu

- + Submit A Job
- Documentation
 - Detailed Documentation
 - Release Notes
 - Installation Guide
 - Command Line Documentation :
 - FAQ
 - Publications

The MEME Suite

Motif-based sequence analysis tools

a smallish example to illustrate principles

Examples

TOMTOM

<http://meme.nbcr.net>

MEME Suite Menu

- + Submit A Job
- Documentation
 - Detailed Documentation**
 - Release Notes
 - Installation Guide
 - Command Line Documentation :
 - FAQ
 - Publications

The MEME Suite

Motif-based sequence analysis tools

Motif Comparison

TOMTOM

Manual

Tutorial [Sample Output](#)

a smallish example to illustrate principles

Examples

TOMTOM

<http://meme.nbcrc.net>



The MEME Suite

Motif-based sequence analysis tools

Motif Comparison

TOMTOM

Manual

Tutorial [Sample Output](#)

a smallish example to illustrate principles

Examples

TOMTOM

<http://meme.nbcrc.net>



The MEME Suite

Motif-based sequence analysis tools

Motif Comparison

TOMTOM

Manual

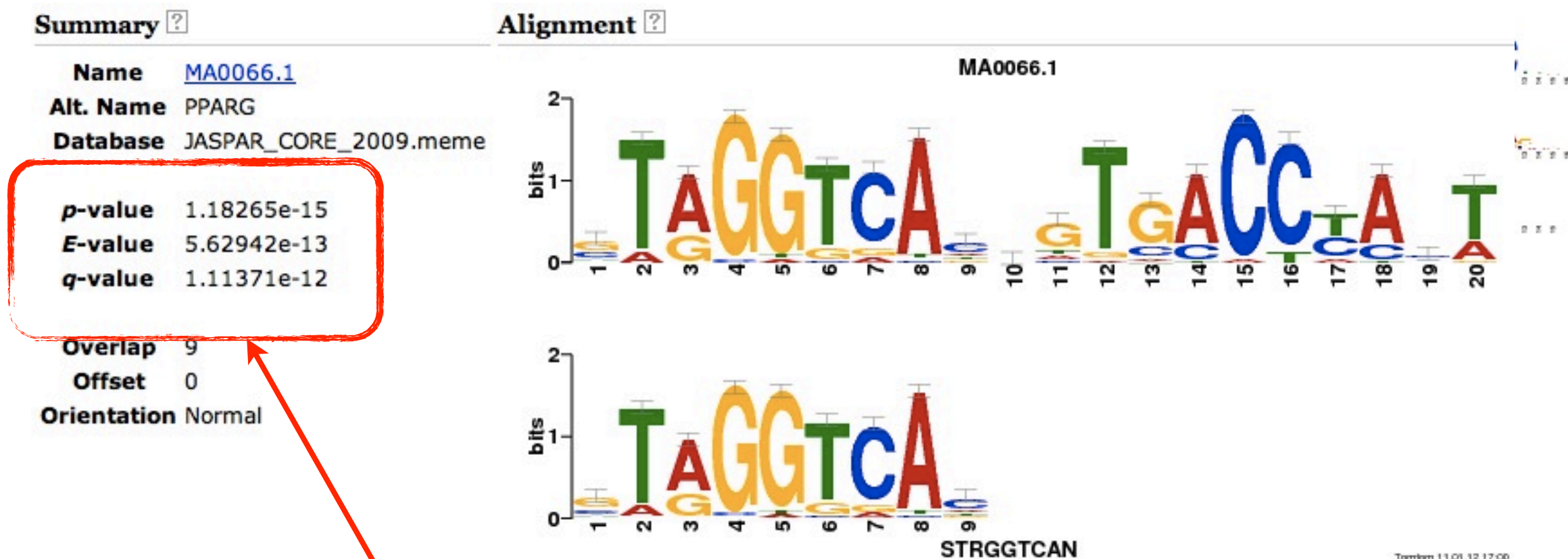
Tutorial

Sample Output

a smallish example to illustrate principles

TOMTOM detail

scroll through the example



statistics on the quality of match (E-value is p-value scaled for DB size; q minimum false discovery rate to include the match)

Finally a quick and easy combo

- MEME-ChiP runs a combination of tools
 - trims large sequence sets so MEME can run in a reasonable amount of time

MEME-ChIP outputs

- MEME and DREME
- CentriMo
 - MEME and DREME motifs combined with JASPAR CORE 2009 database
- TOMTOM
 - separately: MEME, DREME vs. JASPAR CORE 2009

Examples

MEME-ChIP

<http://meme.nbcr.net>



a smallish example to illustrate principles

Examples

MEME-ChIP

<http://meme.nbcr.net>

MEME Suite Menu

- + Submit A Job
- + Documentation
- + Downloads
- + User Support
- + Alternate Servers
- Authors
- Citing

The MEME Suite

Motif-based sequence analysis tools

a smallish example to illustrate principles

Examples

MEME-ChIP

<http://meme.nbcr.net>

MEME Suite Menu

- + Submit A Job
- + Documentation
- + Downloads
- + User Support
- + Alternate Servers
- Authors
- Citing

The MEME Suite

Motif-based sequence analysis tools

a smallish example to illustrate principles

Examples

MEME-ChIP

<http://meme.nbcr.net>

MEME Suite Menu

- + Submit A Job
- Documentation
 - Detailed Documentation
 - Release Notes
 - Installation Guide
 - Command Line Documentation :
 - FAQ
 - Publications

The MEME Suite

Motif-based sequence analysis tools

a smallish example to illustrate principles

Examples

MEME-ChIP

<http://meme.nbcr.net>

MEME Suite Menu

- + Submit A Job
- Documentation
 - Detailed Documentation**
 - Release Notes
 - Installation Guide
 - Command Line Documentation
 - FAQ
 - Publications

The MEME Suite

Motif-based sequence analysis tools

DREME	Manual	Tutorial	Sample Output
GLAM2	Manual	Tutorial	Sample Output
MEME	Manual	Tutorial	Sample Output
MEME-ChIP	Manual		<u>Sample Output</u>

a smallish example to illustrate principles

Examples

MEME-ChIP

<http://meme.nbcr.net>



The MEME Suite

Motif-based sequence analysis tools

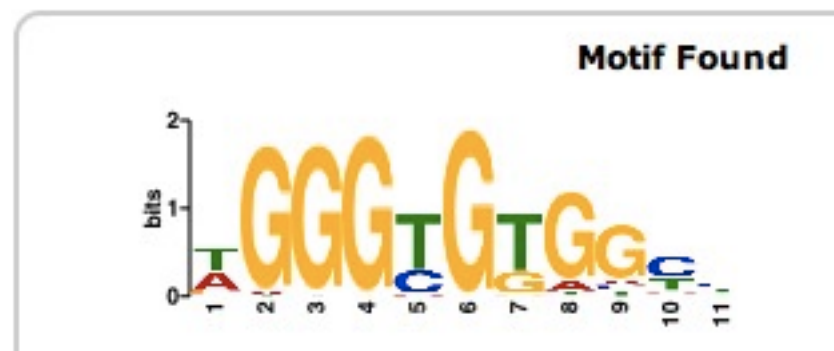
DREME	Manual	Tutorial	Sample Output
GLAM2	Manual	Tutorial	Sample Output
MEME	Manual	Tutorial	Sample Output
MEME-ChIP	Manual		<u>Sample Output</u>

a smallish example to illustrate principles

MEME-ChIP summary

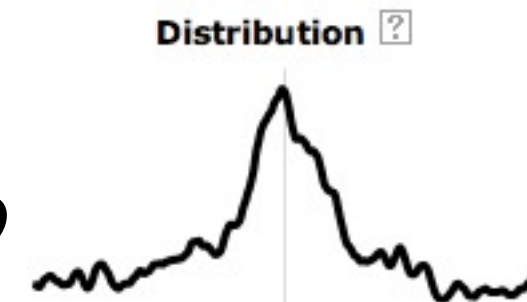
- motifs ranked by *E*-value no matter what the source – e.g., a CentriMo motif could be from a motif DB or from one of our motif finders
- CentriMo distribution summary shown for each

MEME-ChIP output examples



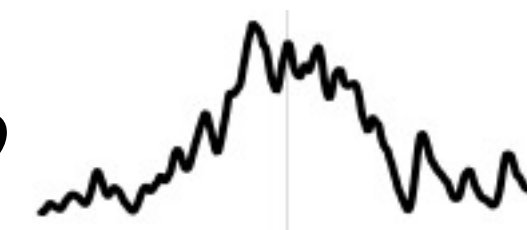
MEME

$$E=8.1 \times 10^{-229}$$



CentriMo

$$E=8.1 \times 10^{-229}$$



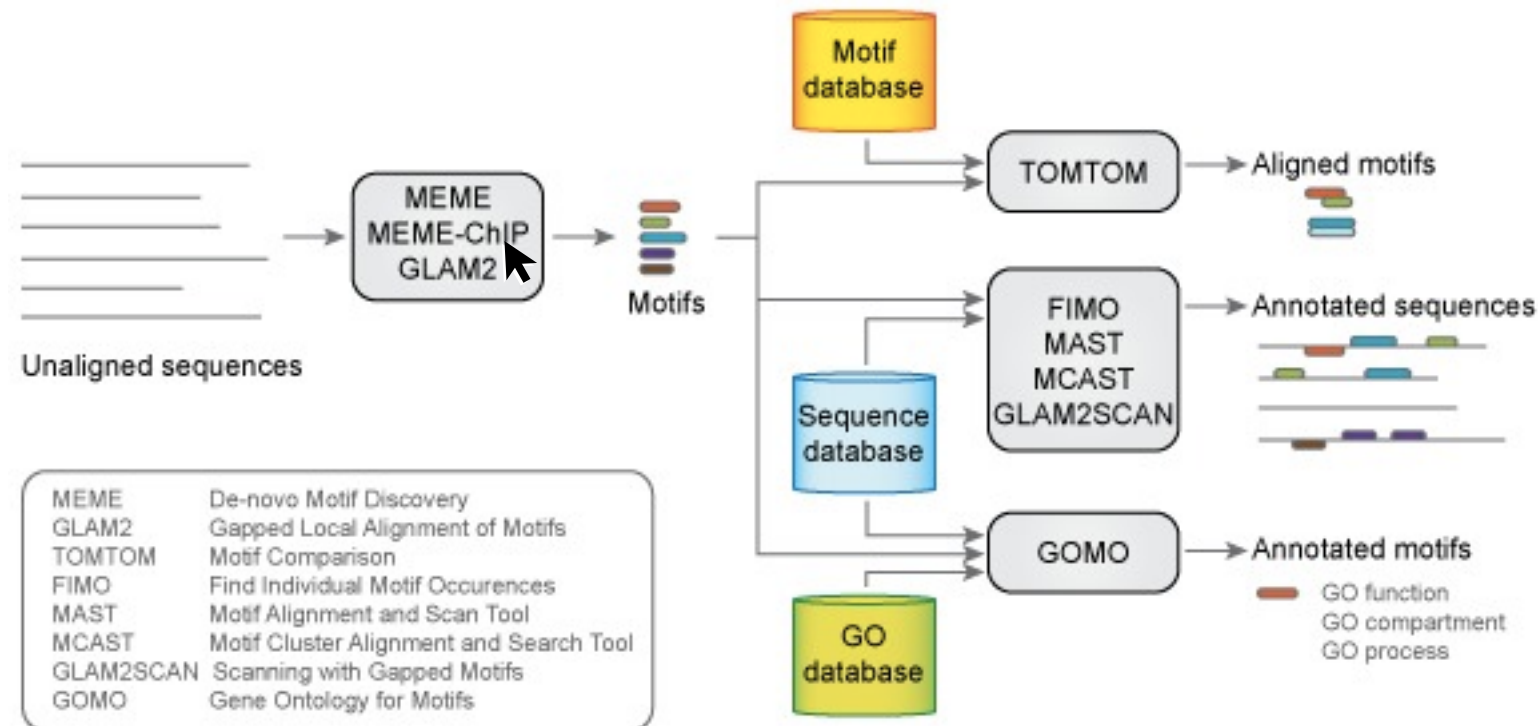
JASPAR:TaII::GataI (MA0140.1)

check out the other outputs

Part 2 application

Run our own example

<http://meme.nbcrc.net>

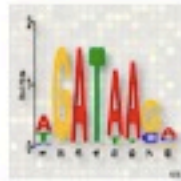


SCL is an interesting example

- motif is low in information, an E-box:
CANNTG
- binds in cooperation with other more distinctive motifs

Files on memory stick

- **Wits workshop** – contains
 - **SCL.fa** – sequences, you will need them now
 - **memeWits2013.ppt** and **memeWits2013.pdf** – copy of slides
 - **MEMECHIP-run** – same outputs as the example you are running now



MEME-ChIP

Motif Analysis of Large DNA Datasets

Version 4.9.0

Use this form to submit DNA sequences to MEME-ChIP. MEME-ChIP is designed especially for discovering motifs in **LARGE** (50MB maximum) sets of short (around 500bp) DNA sequences centered on locations of interest such as those produced by ChIP-seq experiments.

Data Submission Form

Perform motif discovery and enrichment on large DNA datasets.

Input the sequences

Enter DNA sequences in which you want to find motifs [?](#)

or ☐ paste the sequences

Input the motif database

[?](#)

Input job queue details

Enter your email address. [?](#)

Re-enter your email address.

Optionally enter a job description. [?](#)

Your job id is: **appMEMECHIP_4.9.01369051150859-1073137303**

You can view your job results at: http://nbc-222.ucsd.edu/meme_4.9.0/cgi-bin/querystatus.cgi?jobid=appMEMECHIP_4.9.01369051150859-1073137303

You can view server activity [here](#).

Description

SCL Wits workshop 2013

Settings

Sequences	Scl.fa
Motif Database	JASPAR and UniPROBE
Use given strand only	No

MEME Specific Settings

Distribution of motif occurrences	Zero or one per sequence
Number of different motifs	3
Minimum motif width	6
Maximum motif width	30

DREME Specific Settings

Motif E-value Threshold	0.05
-------------------------	------

CentriMo Specific Settings

Minimum Site Score	5
E-value Threshold	10
Allow Uncentered Regions	Disabled
Store Sequence IDs	Enabled

Sequences Details

this will take a while

Break

then work through the uploaded example

My version

http://is.gd/MEME_Wits_2013

Support

- this trip funded by Rhodes University
- other support: NRF