STREME

Sensitive, Thorough, Rapid, Enriched Motif Elicitation

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

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

If you use STREME in your research, please cite the following paper:
Timothy L. Bailey, "STREME: accurate and versatile sequence motif discovery", *Bioinformatics*, Mar. 24, 2021. [full text]

DISCOVERED MOTIFS | INPUTS & SETTINGS | PROGRAM INFORMATION | MOTIFS IN MEME TEXT FORMAT | MATCHING SEQUENCES | MATCHING SITES | RESULTS IN XML FORMAT

DESCRIPTION

Compared to background sequences from homer

DISCOVERED MOTIFS

Next Top

Motif	Logo	RC Logo	P-
1–CACHAGRKGGCRSTAKWGS			1
2–RTTGCAGTTCTM			1.
3–AACGCAGTWTTTCHNNGG			1.

Stopped because 3 consecutive motifs exceeded the p-value threshold (0.05).
STREME ran for 160.66 seconds.

Motif	Logo	RC Logo	P-
4-AGAGGGAGCDATW			1.

Stopped because 3 consecutive motifs exceeded the p-value threshold (0.05).
STREME ran for 160.66 seconds.

INPUTS & SETTINGS

Sequences

Role	Source	Alphabet	Sequence Count	Total Size
Positive (primary) Sequences	150_top_insertion_30bp_flanks.fasta	DNA	150	11850
Negative (control) Sequences	output.bg.sequences.fasta	DNA	50632	3999928

Background Model

Source: built from the negative (control) sequences

Order: 2 (only order-0 shown)

Name	Freq.	Bg.			Bg.	Freq.	Name
Adenine	0.259	0.268	A	~ T	0.268	0.259	Thymine
Cytosine	0.241	0.232	C	~ G	0.232	0.241	Guanine

Other Settings

Strand Handling	Both the given and reverse complement strands are processed.
Objective Function	Differential Enrichment
Statistical Test	Fisher Exact Test
Minimum Motif Width	8
Maximum Motif Width	19
Test Set	10% of the input sequences were randomly assigned to the test set.
Word Evaluation	Up to 25 words of each width from 8 to 19 were evaluated to find seeds.
Seed Refinement	Up to 4 seeds of each width from 8 to 19 were further refined.
Refinement Iterations	Up to 20 iterations were allowed when refining a seed.
Random Number Seed	0
Trimming of Control Sequences	Trimming of control sequences was allowed.
Total Length	The total length of each sequence set was limited to 4.00e+6.
Maximum Motif p-value	Stop when the p-value is greater than 0.05 for 3 consecutive motifs.
Maximum Motifs to Find	No maximum number of motifs.
Maximum Run Time	14400 seconds.

STREME version
5.5.5 (Release date: Thu Sep 14 08:48:04 2023 +1000)

Reference
Timothy L. Bailey, "STREME: accurate and versatile sequence motif discovery", *Bioinformatics*, Mar. 24, 2021. [full text]

Command line
```
streme --verbosity 1 --oc . --dna --totallength 4000000 --time 14400 --minw 8 --maxw 19
--thresh 0.05 --align center --dfile description --p 150_top_insertion_30bp_flanks.fasta
--n output.bg.sequences.fasta
```