Ruan lab CTCF abcam antibody characterization

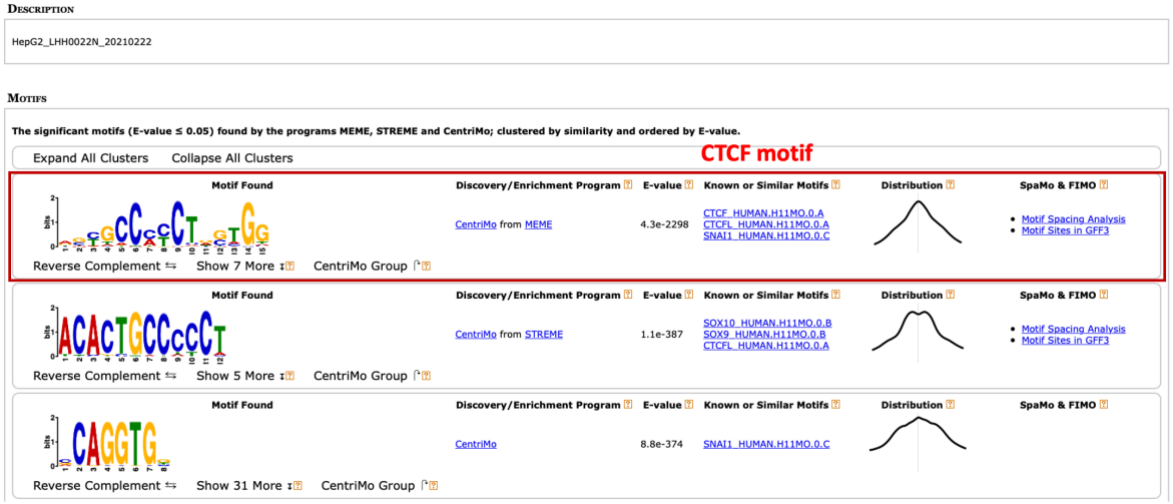
Goal: to characterize a CTCF abcam antibody (ENCAB871TEY: https://www.encodeproject.org/antibodies/ENCAB871TEY/) used by the Ruan lab for ChIA-PET experiments

2. Secondary characterization: Motif analysis

Cell line	Antibody	ChIA-PET experiment ID	Peak file	Binding coverage file
HepG2	ENCAB871TEY	ENCSR411IVB	ENCFF226GVS	ENCFF812KSY
K562	ENCAB871TEY	ENCSR597AKG	ENCFF995WHC	ENCFF803SOP

Description:
-Of the peaks listed in the Ruan lab peak file, those with maximum summit of greater than 10 reads are identified and retained (n=36,306 for HepG2, n=28,398 for K562) [`pyBedGraph` v0.5.43].
-Maximum summit location of filtered peaks is extended by 50 bps upstream and downstream, resulting in a bed file with regions of length 100 bps [`pyBedGraph` v0.5.43].
-Sequences in the 100 bps regions are extracted via bedtools getfasta function using hg38 reference genome [`bedtools` v2.27.0].
-Motif analysis on these primary sequences are performed with the following command: meme-chip -oc . -time 240 -ccut 100 -fdesc description -dna -order 2 -minw 6 -maxw 15 -db db/HUMAN/HOCOMOCOv11_core_HUMAN_mono_meme_format.meme -meme-mod zoops -meme-nmotifs 3 -meme-searchsize 100000 -streme-pvt 0.05 -streme-totallength 4000000 -centrimo-score 5.0 -centrimo-ethresh 10.0 ${input_seq_file} [`MEME-ChIP` v5.3.3]. Top 3 motifs are presented below.

a. Motif analysis on HepG2 cell line (top 3 motifs identified by `MEME-ChIP`)



b. Motif analysis on K562 cell line (top 3 motifs identified by `MEME-ChIP`)

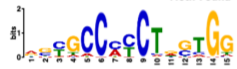
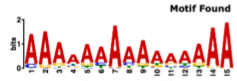
Description

K562_LHK0001N_20210222

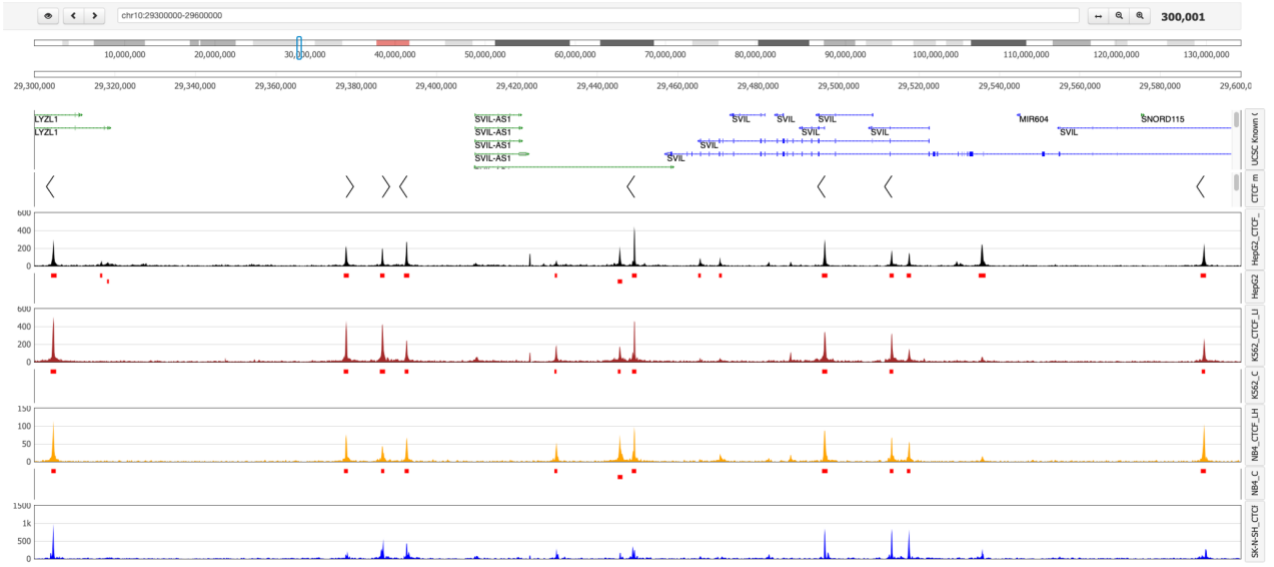
Motifs

The significant motifs (E-value ≤ 0.05) found by the programs MEME, STREME and CentriMo; clustered by similarity and ordered by E-value.

Expand All Clusters Collapse All Clusters **CTCF motif**

Motif Found	Discovery/Enrichment Program	E-value	Known or Similar Motifs	Distribution	SpaMo & FIMO
 Reverse Complement ⇆ Show 7 More ↕ CentriMo Group	MEME	2.0e-1092	CTCF_HUMAN.H11MO.0.A CTCFL_HUMAN.H11MO.0.A SNAI1_HUMAN.H11MO.0.C		• Motif Spacing Analysis • Motif Sites in GFF3
 Reverse Complement ⇆ Show 4 More ↕	MEME	6.7e-233	PRDM6_HUMAN.H11MO.0.C ANDR_HUMAN.H11MO.0.A FOXJ3_HUMAN.H11MO.0.A	Not Centrally Enriched	• Motif Spacing Analysis • Motif Sites in GFF3
Reverse Complement ⇆	CentriMo	2.0e-077	ARI5B_HUMAN.H11MO.0.C		

3. Examples

chr3:50100000-50640000 (540kb)

chr10:29300000-29600000 (300kb)