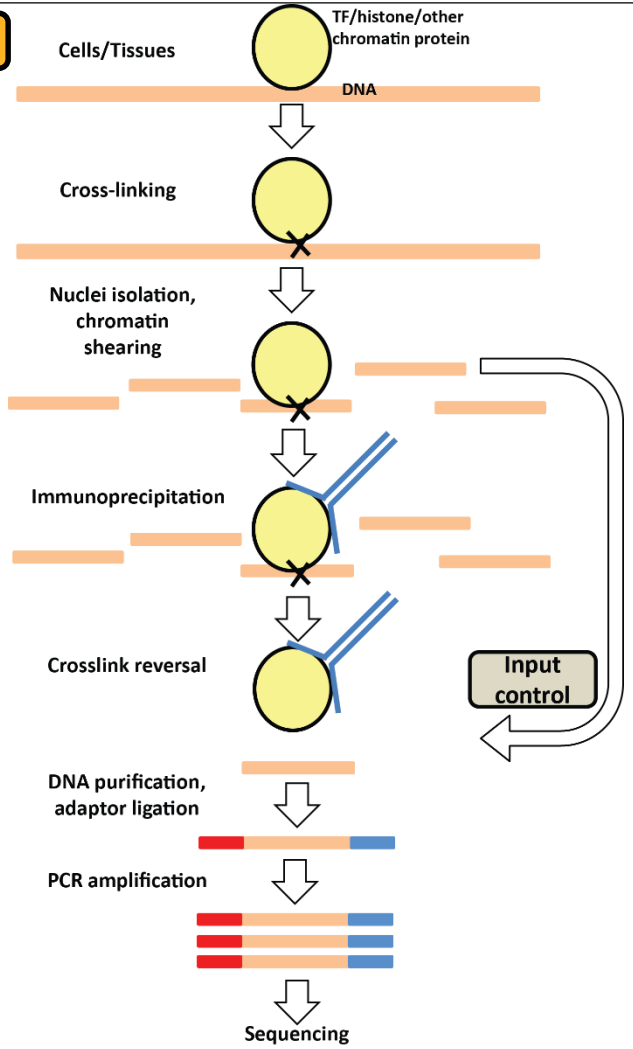
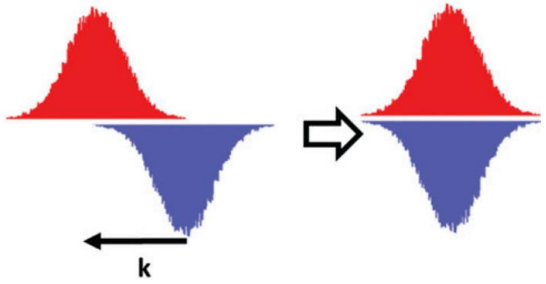
CHIP-SEQ/CHIP-EXO/CUT&RUN COMPARISON

GEORGI K. MARINOV

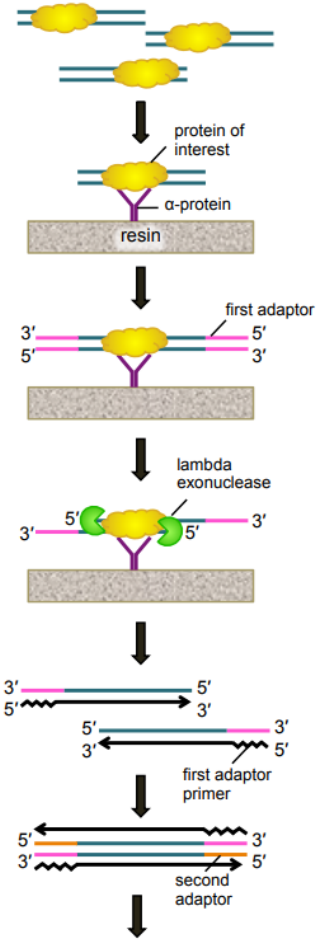
July 7th 2019

CHIP-SEQ


Cells/Tissues
TF/histone/other chromatin protein
DNA
Cross-linking
Nuclei isolation, chromatin shearing
Immunoprecipitation
Crosslink reversal
Input control
DNA purification, adaptor ligation
PCR amplification
Sequencing


k

CHIP-EXO



Crosslink proteins to DNA, shear DNA

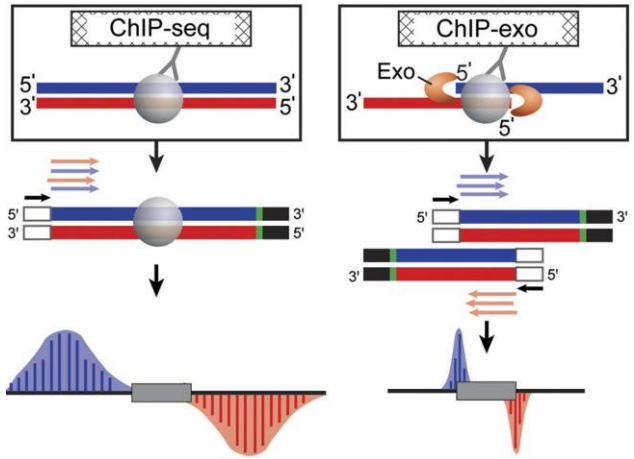
Chromatin-immunoprecipitation

Ligated first adaptor, fill in ends

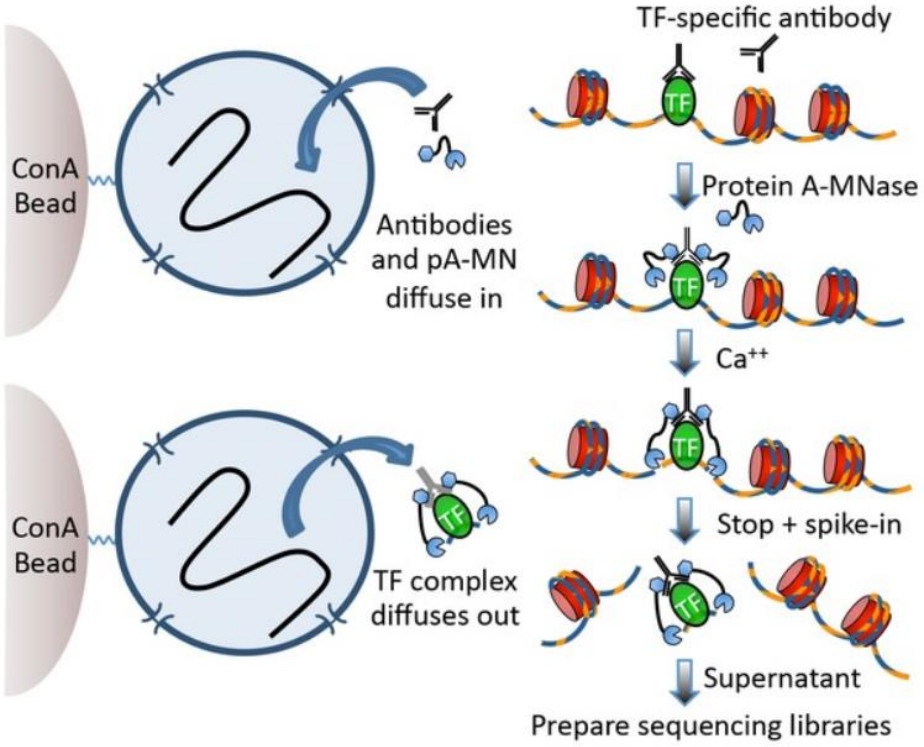
Exonuclease digestion

LM-PCR with primers to first adaptor

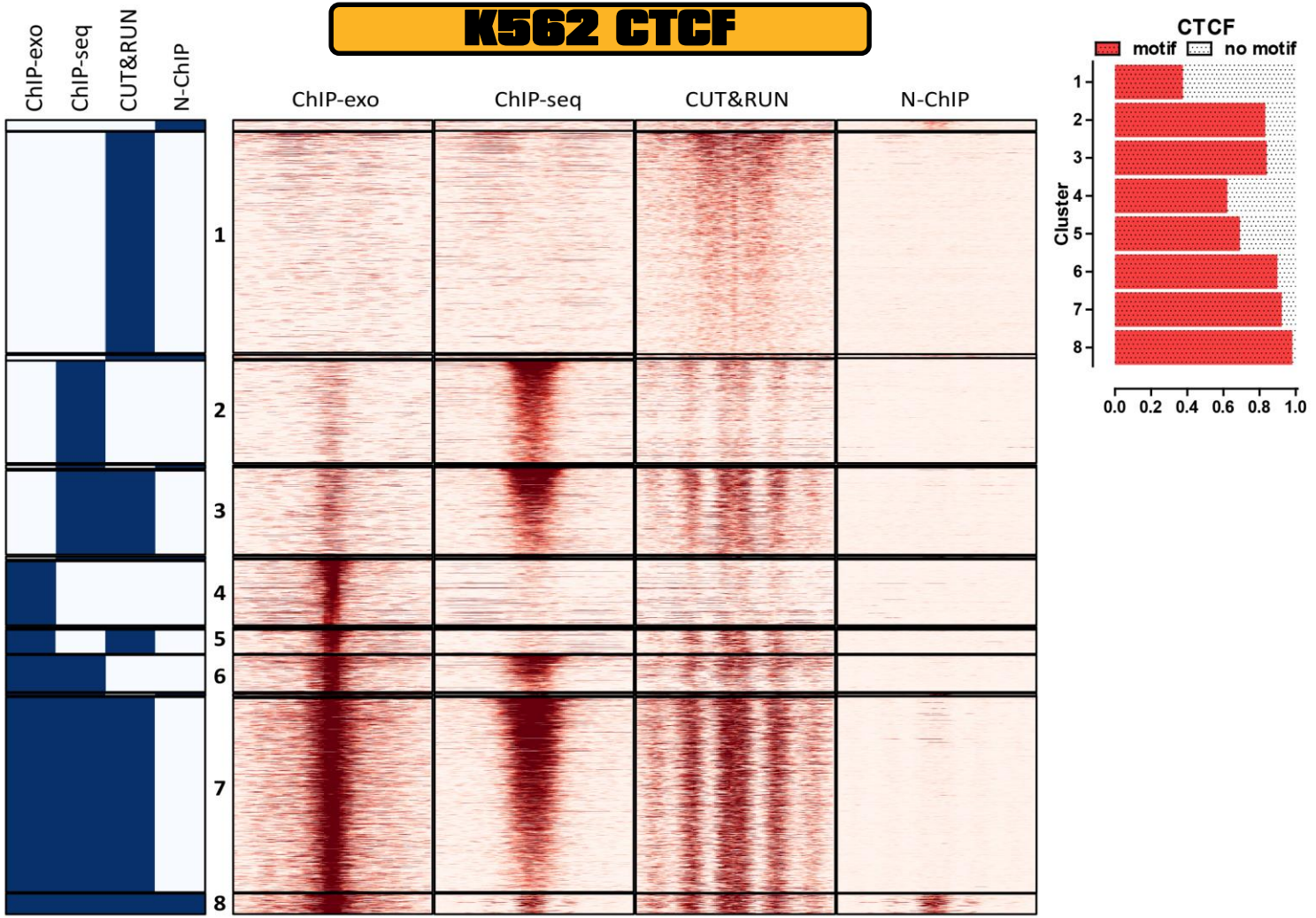
Ligate second adaptor, amplify DNA with LM-PCR



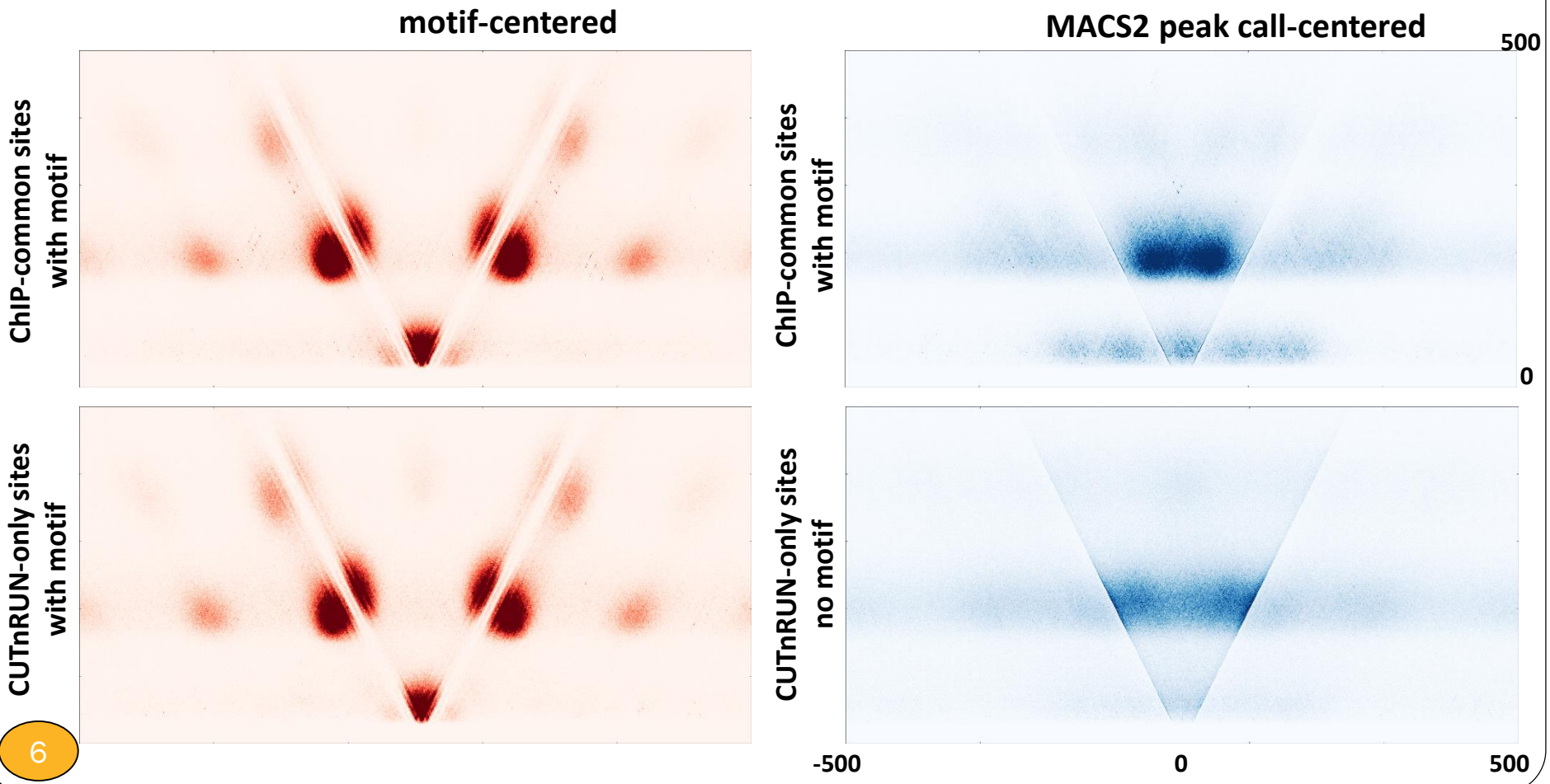
CHIP-EXO


TF-specific antibody
ConA Bead
Antibodies and pA-MN diffuse in
Protein A-MNase
Ca++
ConA Bead
TF complex diffuses out
Stop + spike-in
Supernatant
Prepare sequencing libraries

K562 CTCF


ChIP-exo
ChIP-seq
CUT&RUN
N-ChIP
1
2
3
4
5
6
7
8
CTCF
motif
no motif
Cluster
0.0
0.2
0.4
0.6
0.8
1.0

CTCF


motif-centered
MACS2 peak call-centered
ChIP-common sites with motif
ChIP-common sites with motif
CUTnRUN-only sites with motif
CUTnRUN-only sites no motif
500
0
-500
0
500

GATA1

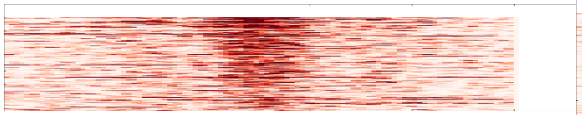
number peaks

	ChIP peak		no ChIP peak	
	motif	no motif	motif	no motif
ENCODE K562 GATA1	6,328	8,277		
GATA1_1_100_250k_neb	511	558	117	525
GATA1_1_2500_250k_neb	213	261	22	167
GATA1_1_2500_50k_neb	595	568	101	476
GATA1_1_500_250k_neb	1,678	1,613	788	2,064
GATA1_1_500_50k_neb	65	111	9	82

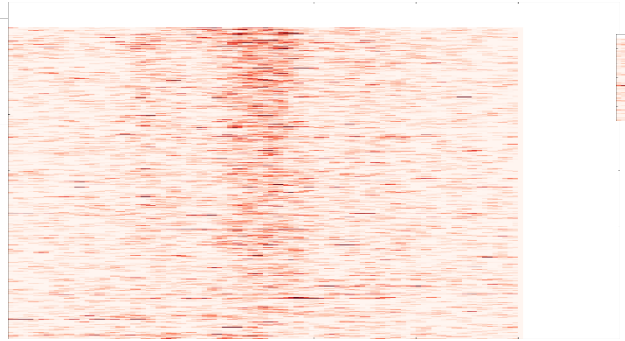
fraction

	ChIP peak	no ChIP peak	ChIP peak		no ChIP peak	
			motif	no motif	motif	no motif
ENCODE K562 GATA1	1.00	0.00	0.43	0.57	n/a	n/a
GATA1_1_100_250k_neb	0.62	0.38	0.48	0.52	0.18	0.82
GATA1_1_2500_250k_neb	0.71	0.29	0.45	0.55	0.12	0.88
GATA1_1_2500_50k_neb	0.67	0.33	0.51	0.49	0.18	0.82
GATA1_1_500_250k_neb	0.54	0.46	0.51	0.49	0.28	0.72
GATA1_1_500_50k_neb	0.66	0.34	0.37	0.63	0.10	0.90

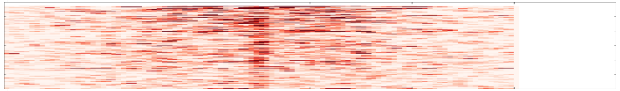
common sites



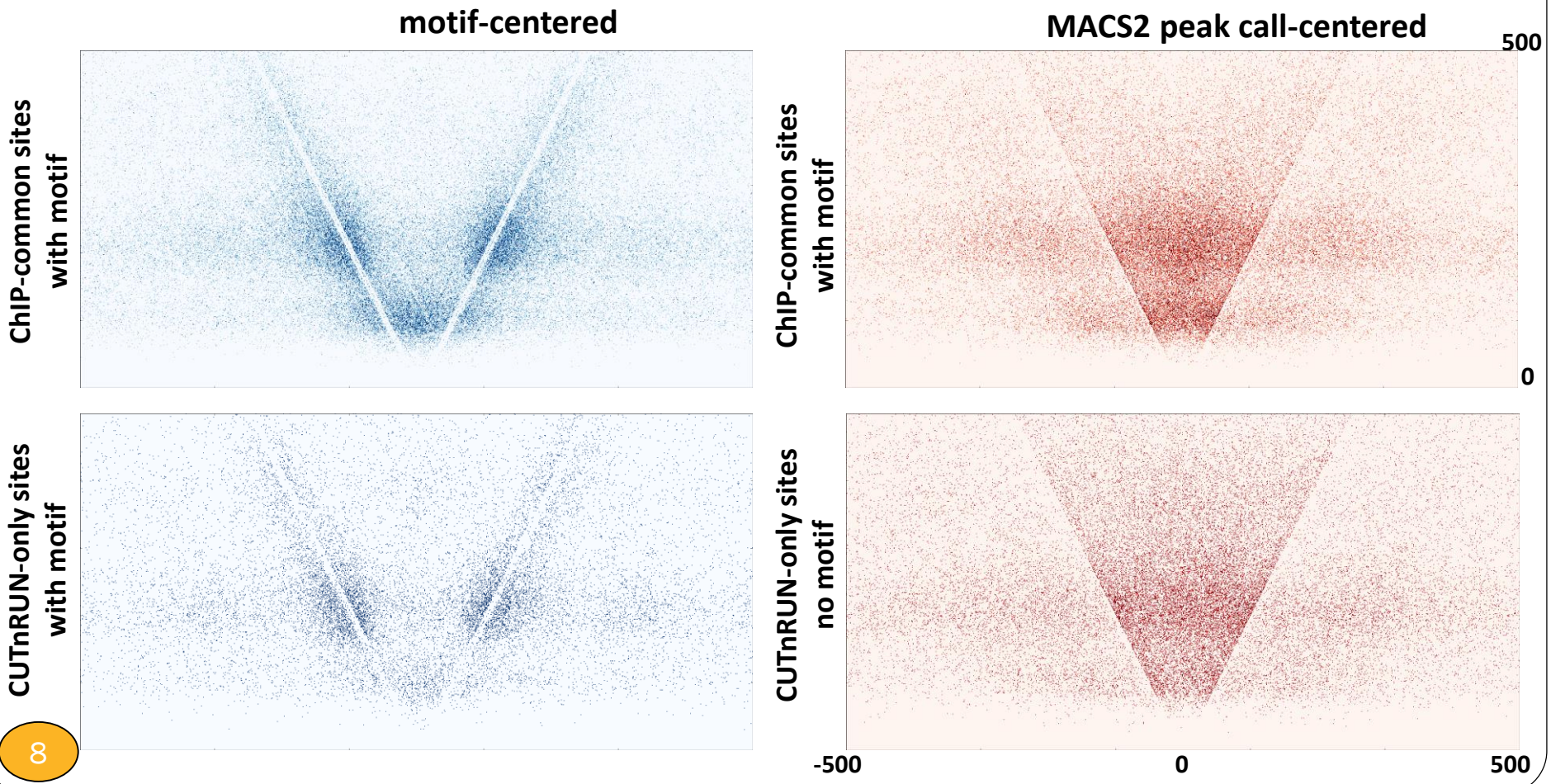
ChIP-only sites



CUT&RUN only sites



GATA1


motif-centered
MACS2 peak call-centered
ChIP-common sites with motif
ChIP-common sites with motif
CUTnRUN-only sites with motif
CUTnRUN-only sites no motif
500
0
-500
0
500

(SOME) GOALS FOR WORKSHOP:

1. Can we develop predictive models that relate sequence to the signal profile of each assay, especially CUT&RUN?

2. Can we apply these models to understand the differences between the occupancy sets derived from the different assays at the sequence level?

3. Can we identify epigenetic properties that distinguish assay-specific sites based on other ENCODE assays?

DATASETS:

1. ChIP-seq/ChIP-exo/CUT&RUN for CTCF in K562 cells
2. ChIP-seq/CUT&RUN for Myc/Max/GATA1 in K562 cells
3. ChIP-seq/CUT&RUN for FOSL2 in A549 cells
4. ChIP-seq/CUT&RUN for CTCF/Nanog/Oct4 in H1-hESC cells