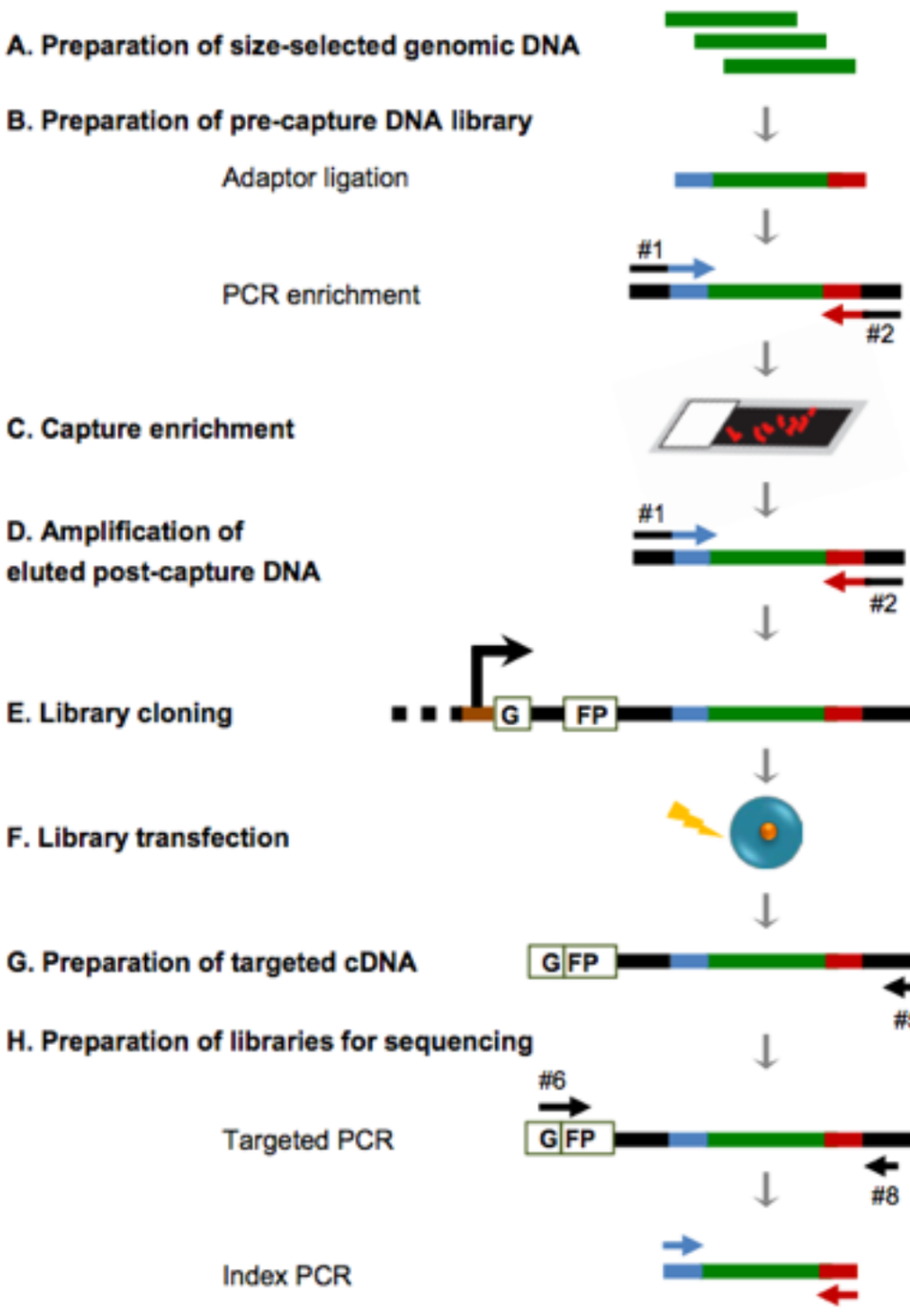
MCF-7 EnhancerSeq

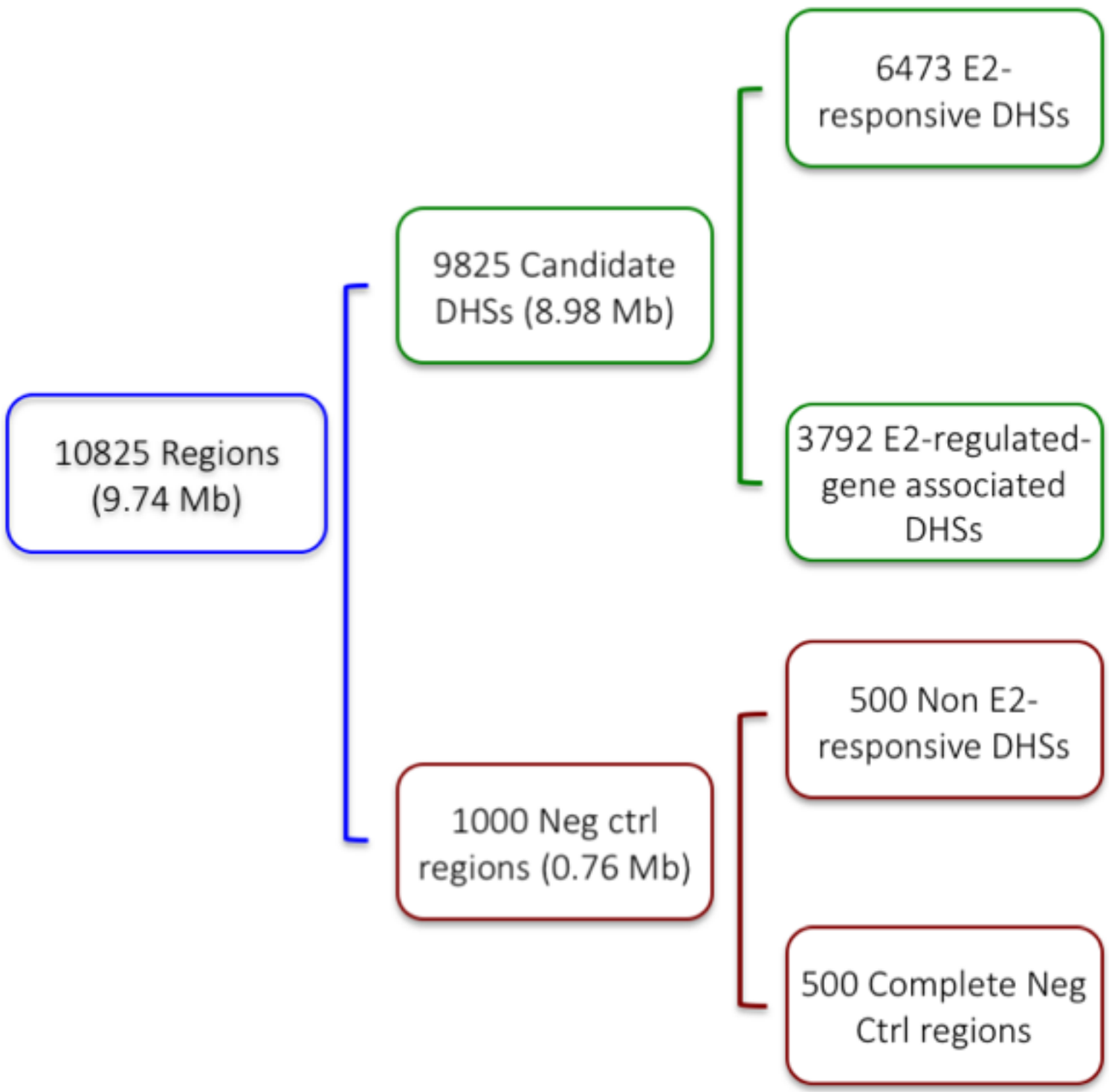
Donghoon Lee
Gerstein Lab

Experiment Design



Nimblegen capture
~10 Mb total target regions

Enhancer-seq Captured Regions



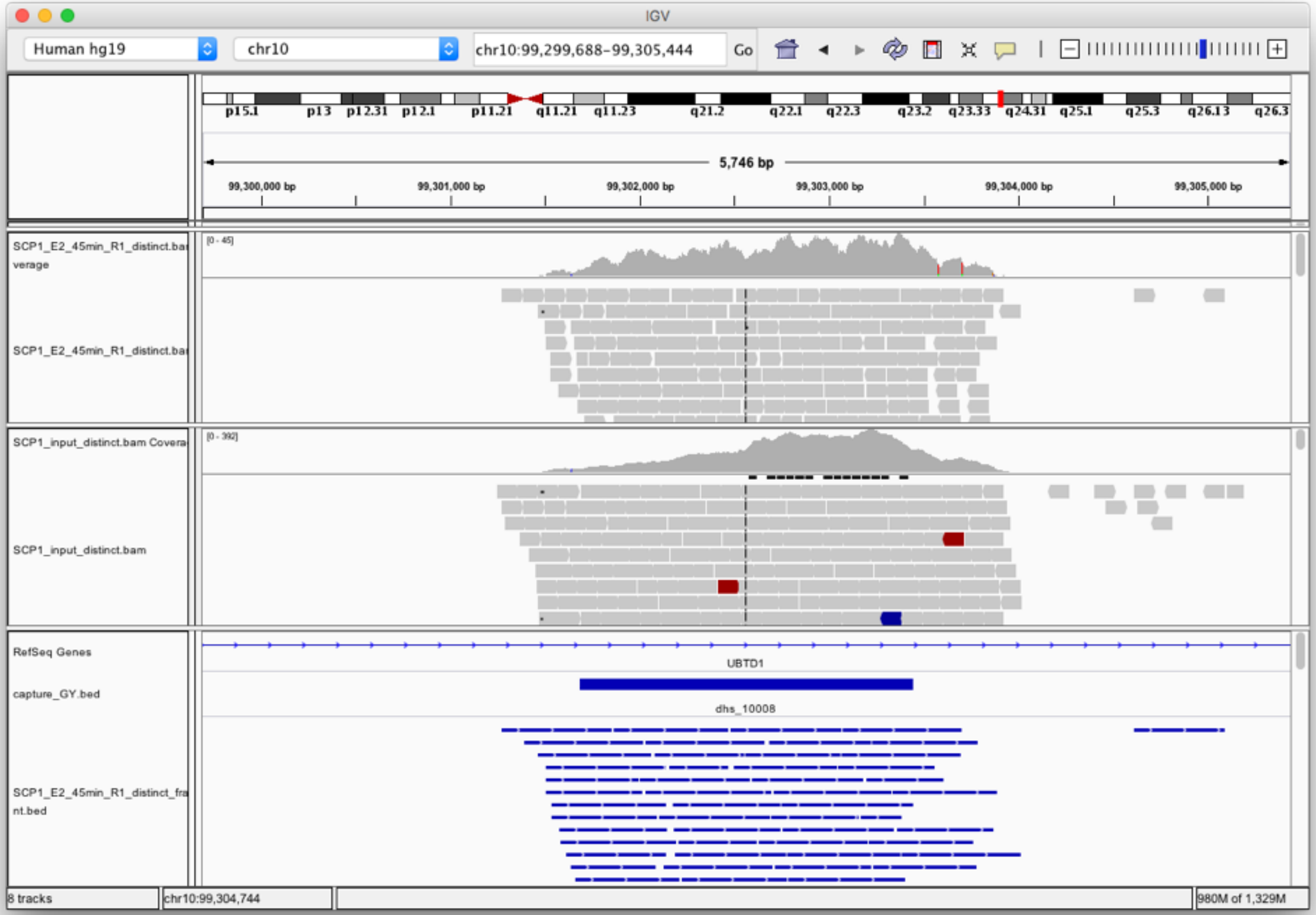
- 2000 E2 constantly induced DHSs
- 2000 E2 constantly repressed DHSs
- Dynamically induced DHSs (top 1000 induced DHSs at each time point)

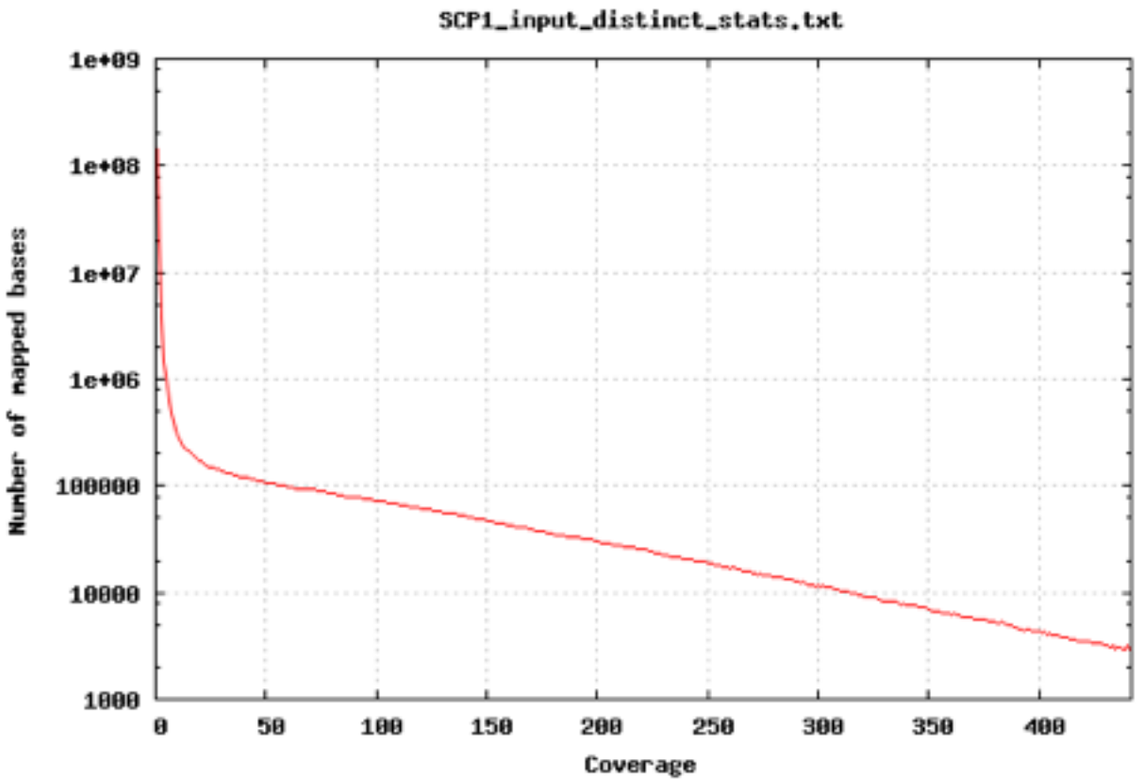
- DHSs within 50 kb upstream of 834 high confident E2-regulated genes
- Exclude +/- 1kb of TSS

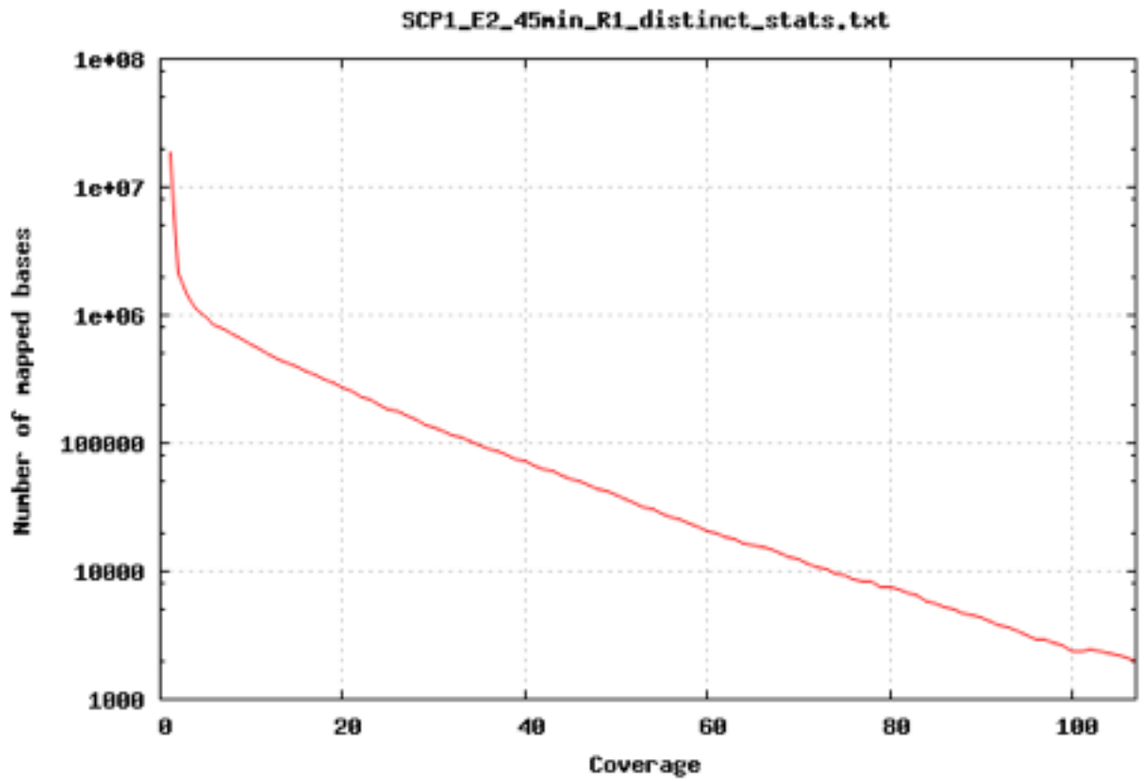
- DHSs w/ lowest |logFC| across all time points

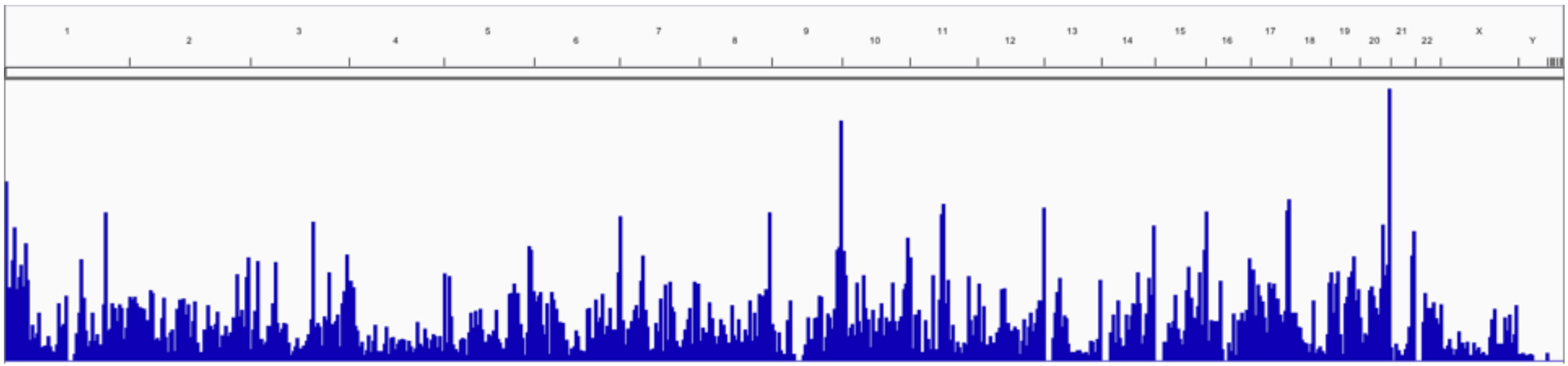
- Regions in closed chromatin
- No H3K27ac signal(data from GSM946850)
- Not NR sites
- Not nested or simple repeats

Slide adopted from Grace Yu


IGV
Human hg19
chr10
chr10:99,299,688-99,305,444
Go
p15.1 p13 p12.31 p12.1 p11.21 q11.21 q11.23 q21.2 q22.1 q22.3 q23.2 q23.33 q24.31 q25.1 q25.3 q26.13 q26.3
5,746 bp
99,300,000 bp
99,301,000 bp
99,302,000 bp
99,303,000 bp
99,304,000 bp
99,305,000 bp
SCP1_E2_45min_R1_distinct.bam Coverage
[0 - 45]
SCP1_E2_45min_R1_distinct.bam
SCP1_input_distinct.bam Coverage
[0 - 392]
SCP1_input_distinct.bam
RefSeq Genes
UBTD1
capture_GY.bed
dhs_10008
SCP1_E2_45min_R1_distinct_fragment.bed
8 tracks
chr10:99,304,744
980M of 1,329M

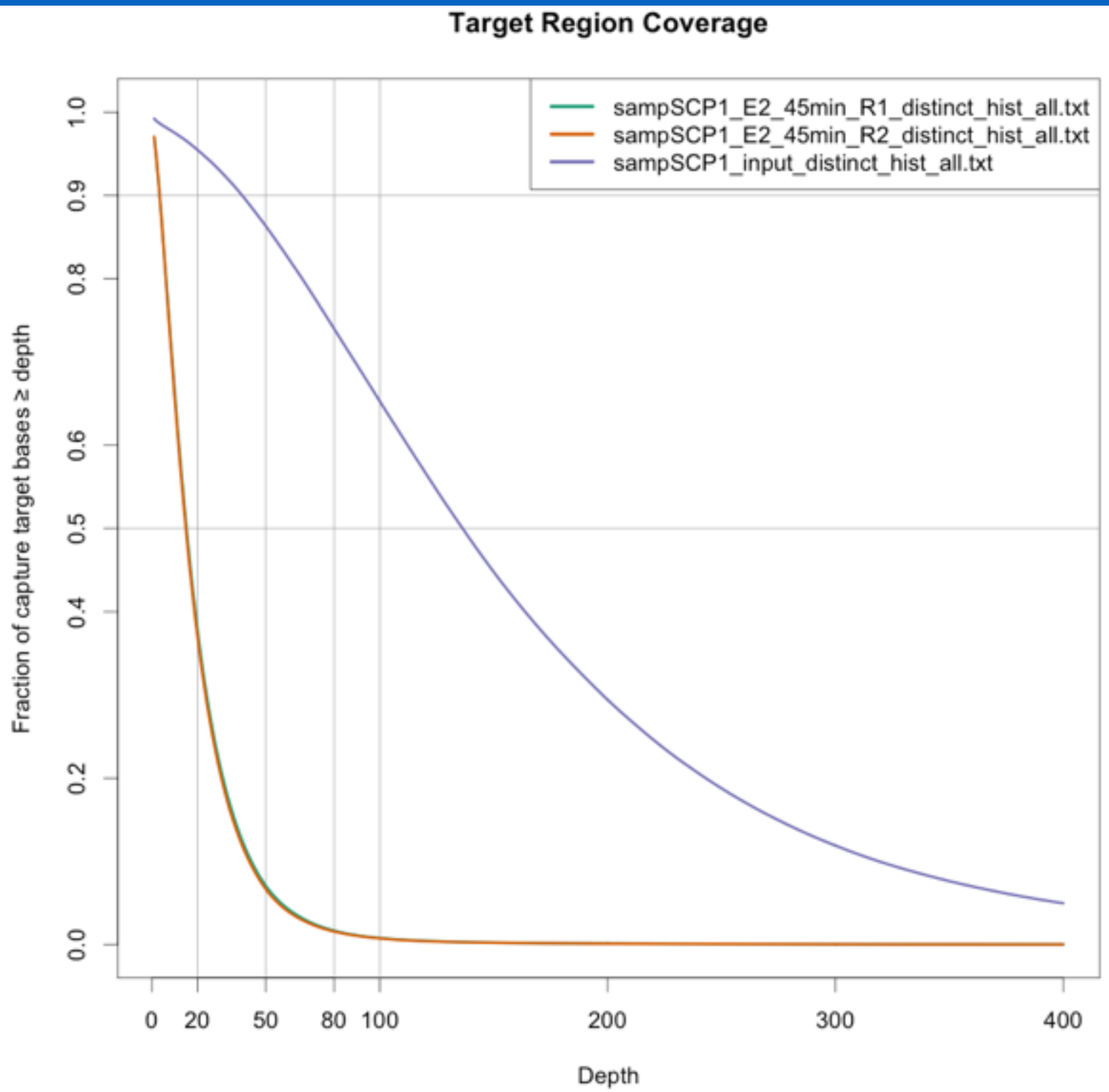

SCP1_input_distinct_stats.txt
Number of mapped bases
1e+09
1e+08
1e+07
1e+06
100000
10000
1000
0
50
100
150
200
250
300
350
400
Coverage


SCP1_E2_45min_R1_distinct_stats.txt
Number of mapped bases
1e+08
1e+07
1e+06
100000
10000
1000
0
20
40
60
80
100
Coverage


1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
X
Y

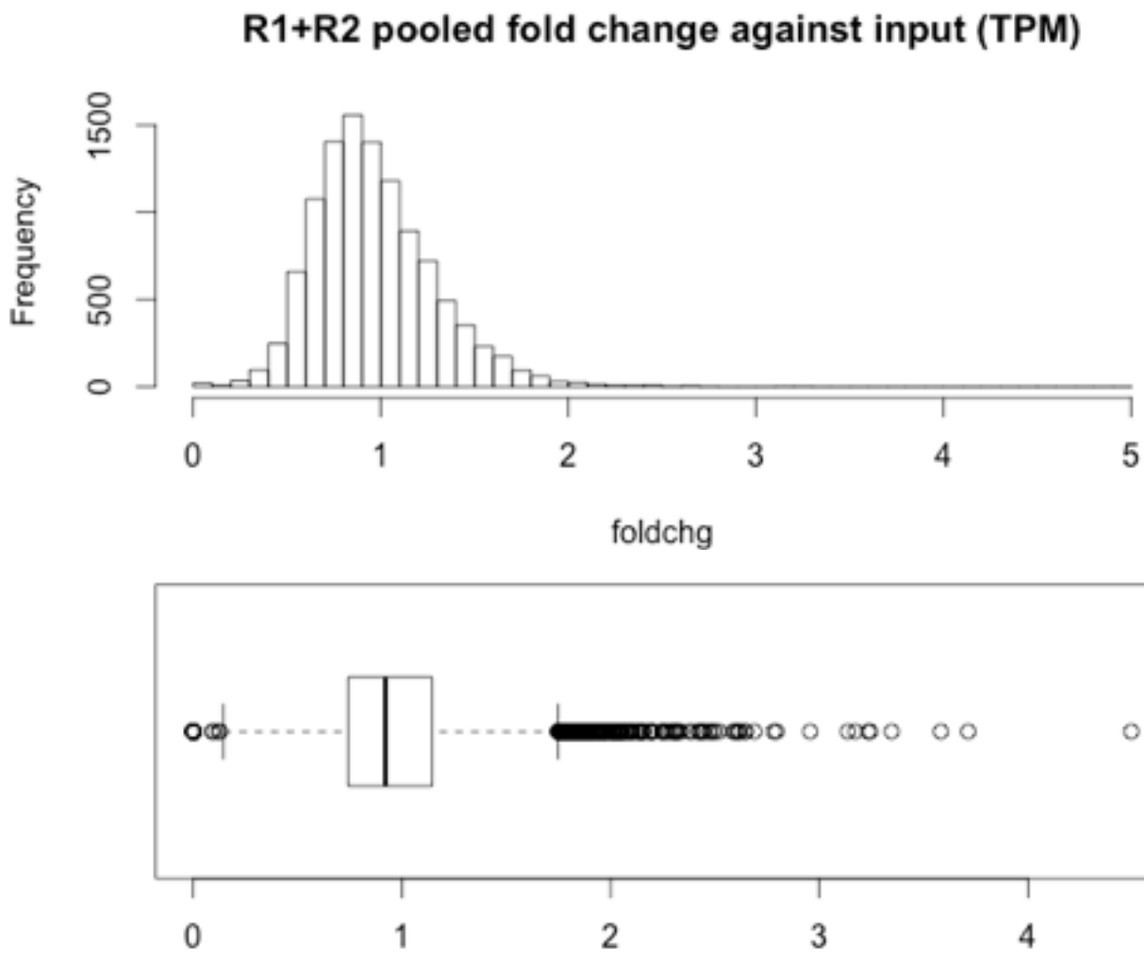
Alignment Statistics

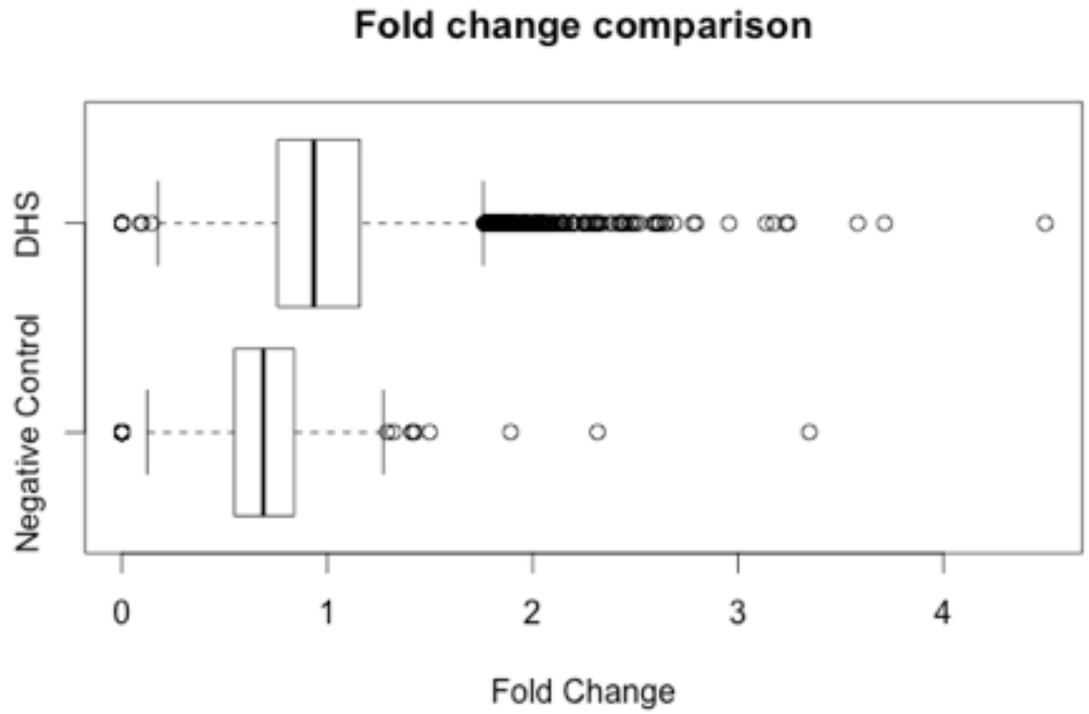
	Mapped paired reads	Uniquely mapped reads	Uniquely mapped fragments
Input	24,153,972	24,035,885	12,015,055
E2_45min_R1	2,758,706	2,745,755	1,372,364
E2_45min_R2	2,706,016	2,693,177	1,346,136
E2_4h_R1	3,224,178	3,208,671	1,603,741
E2_4h_R2	2,543,772	2,531,799	1,265,457
E2_12h_R1	2,536,432	2,523,851	1,261,447
E2_12h_R2	2,566,568	2,551,780	1,275,364
E2_24h_R1	1,832,258	1,823,256	911,240
E2_24h_R2	2,016,094	2,005,778	1,002,443



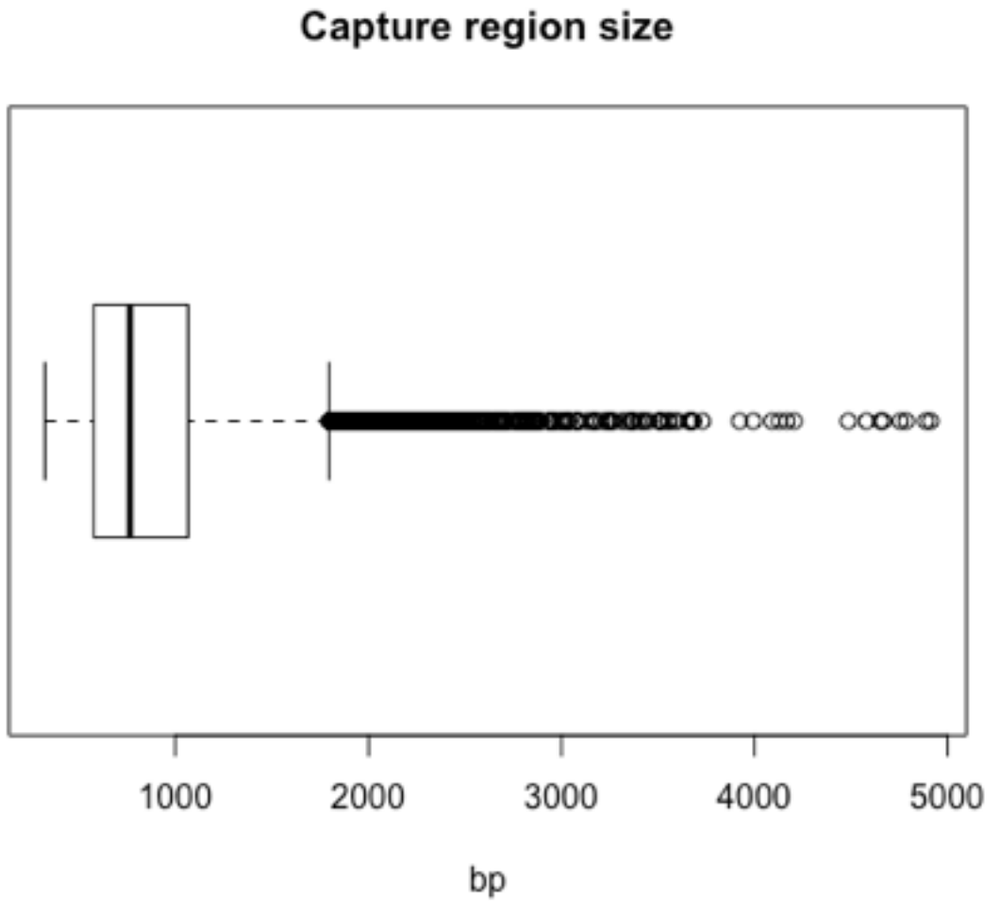
	Mapped paired reads	Reads WITHIN capture	Reads OUTSIDE capture	% capture
Input	24,153,972	17,194,372	6,959,600	**71.19%**
E2_45min_R1	2,758,706	2,085,933	672,773	**75.61%**
E2_45min_R2	2,706,016	2,038,102	667,914	**75.32%**

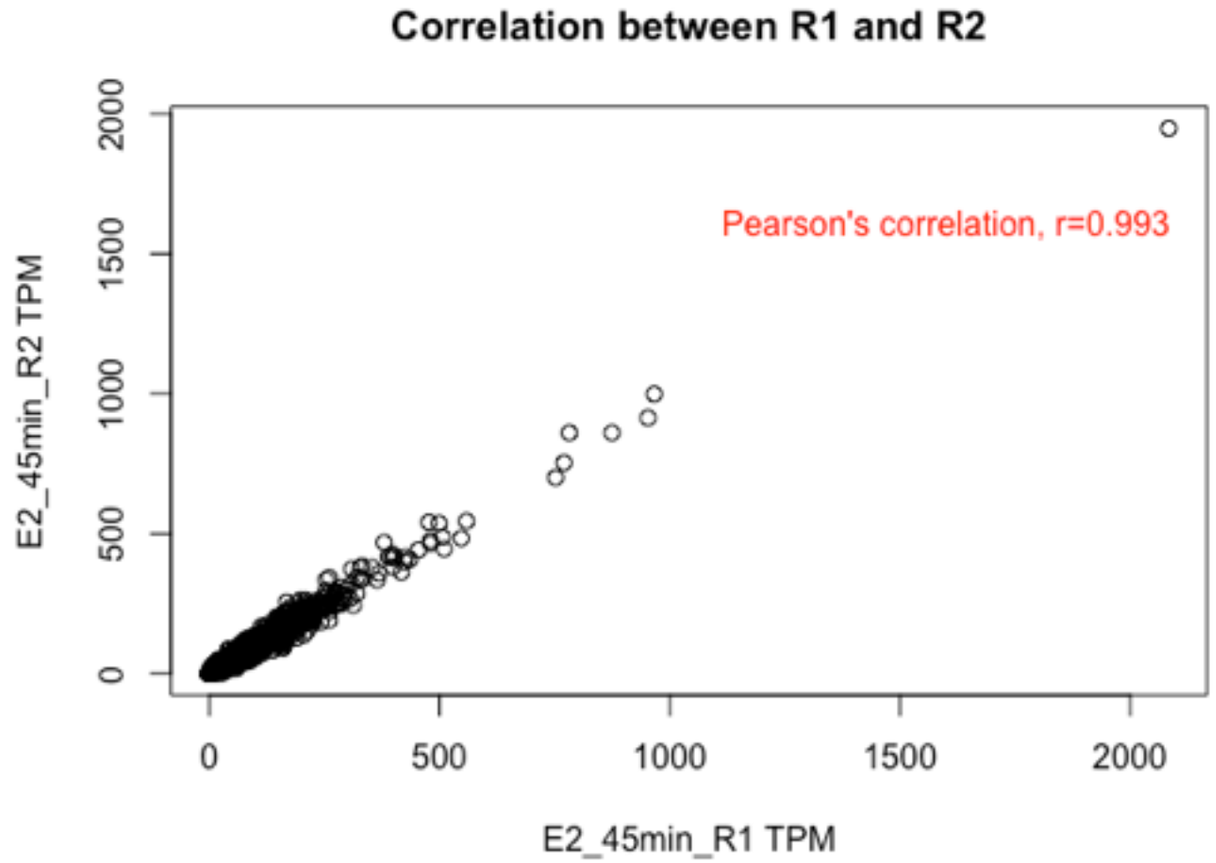
TPM over whole capture regions





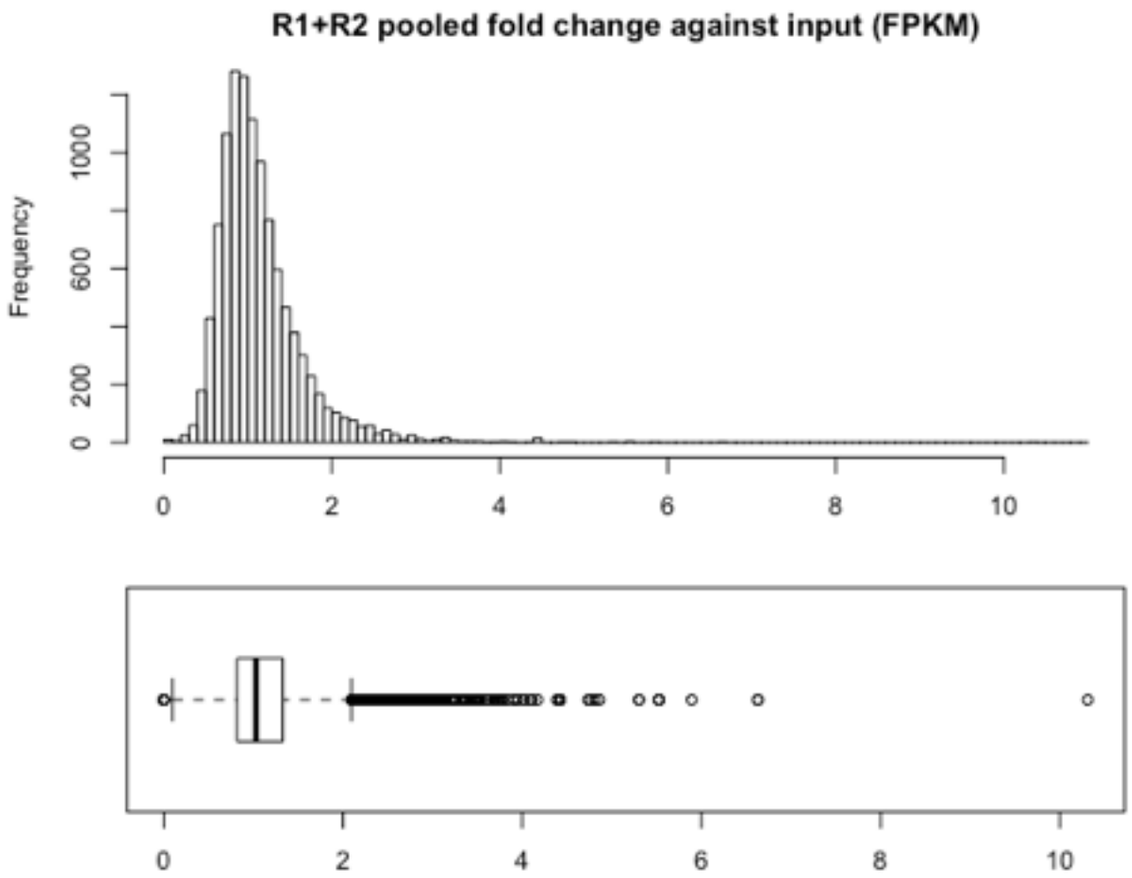
TPM over whole capture regions

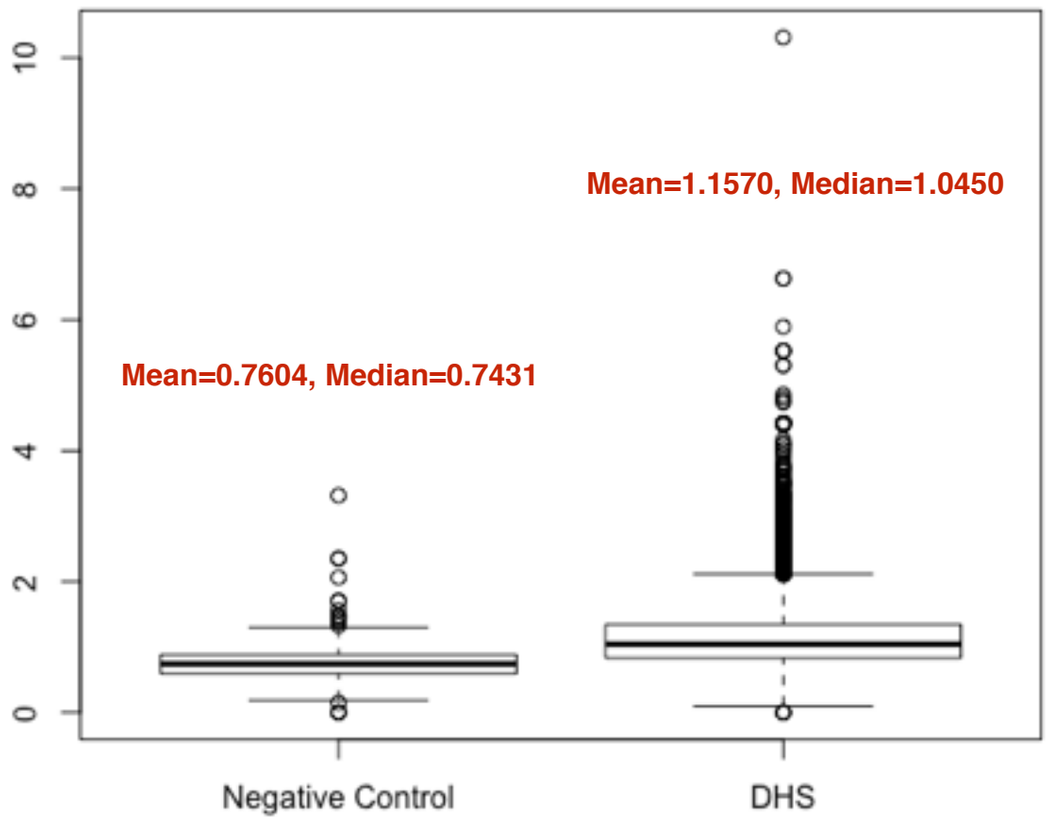




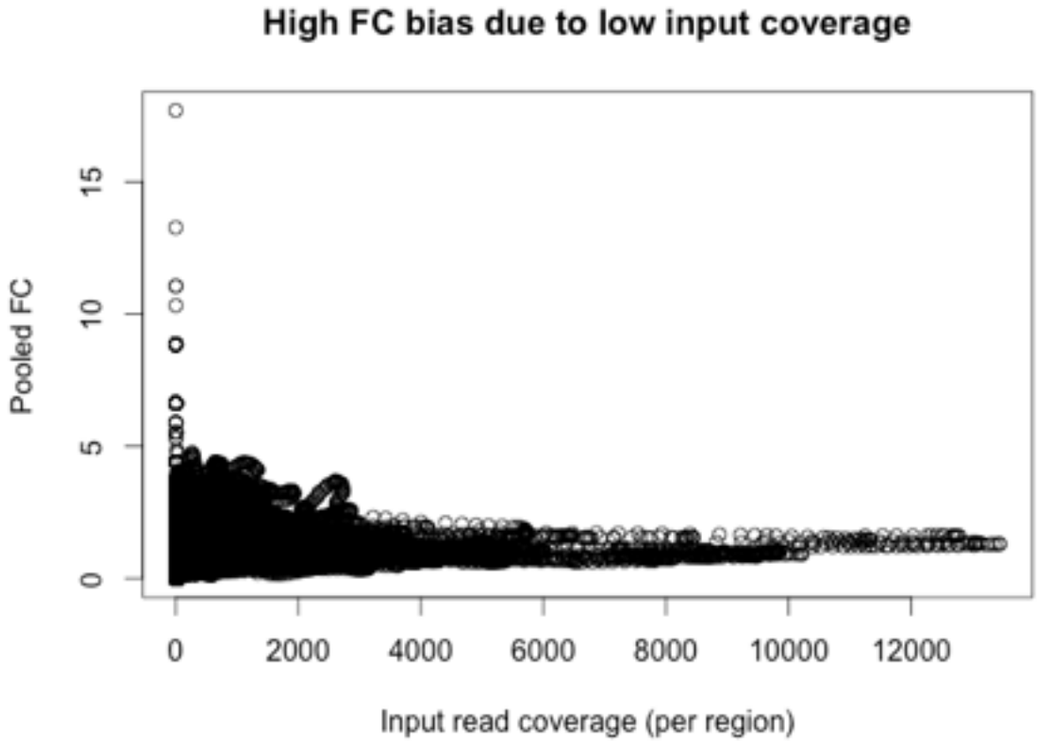
- One captured regions may contain multiple enhancers
- Real enhancers may only be small portions (up to 1kb?) within the captured region
- By directly calculating the fold change over the whole captured region, you may “dilute” the signal

FPKM over sliding 200 bp window (20bp step)

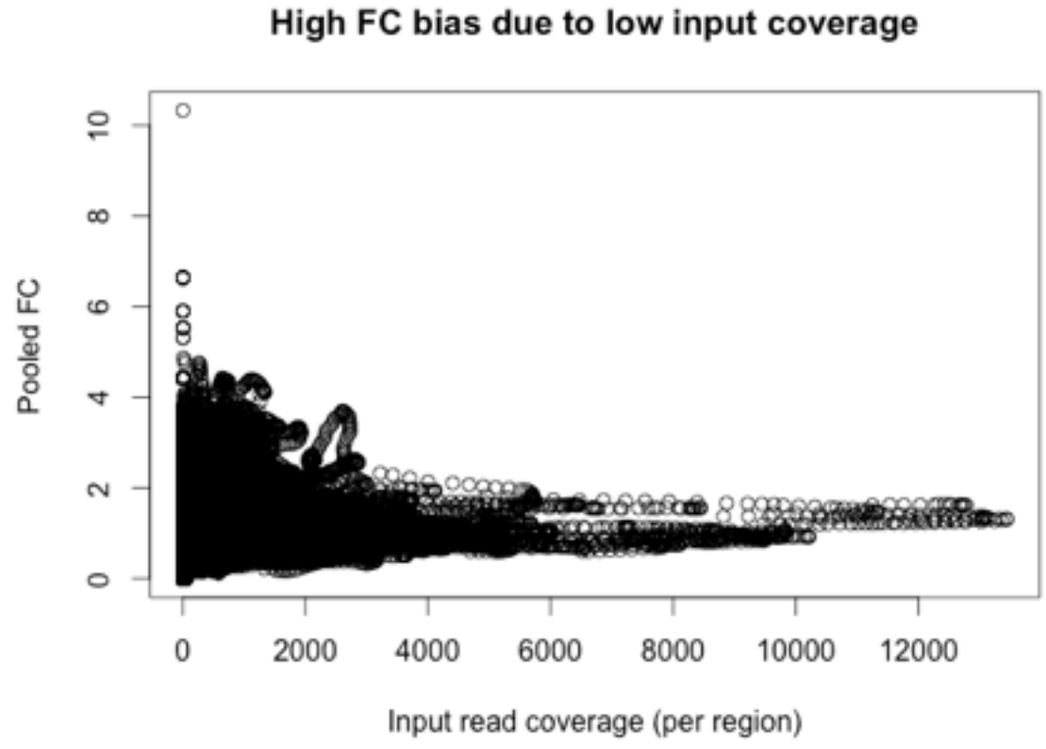




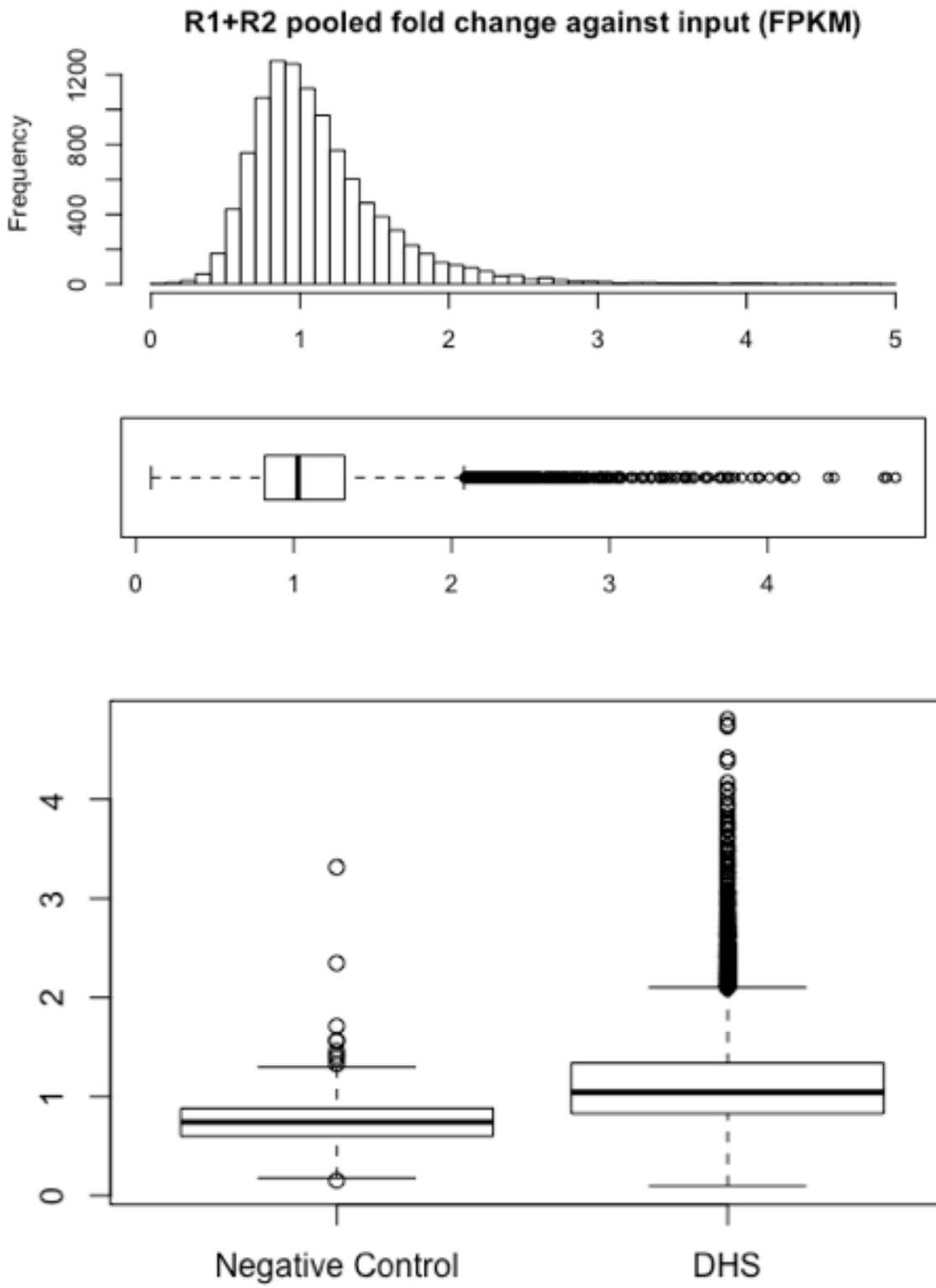
FPKM over sliding 200 bp window (20bp step)

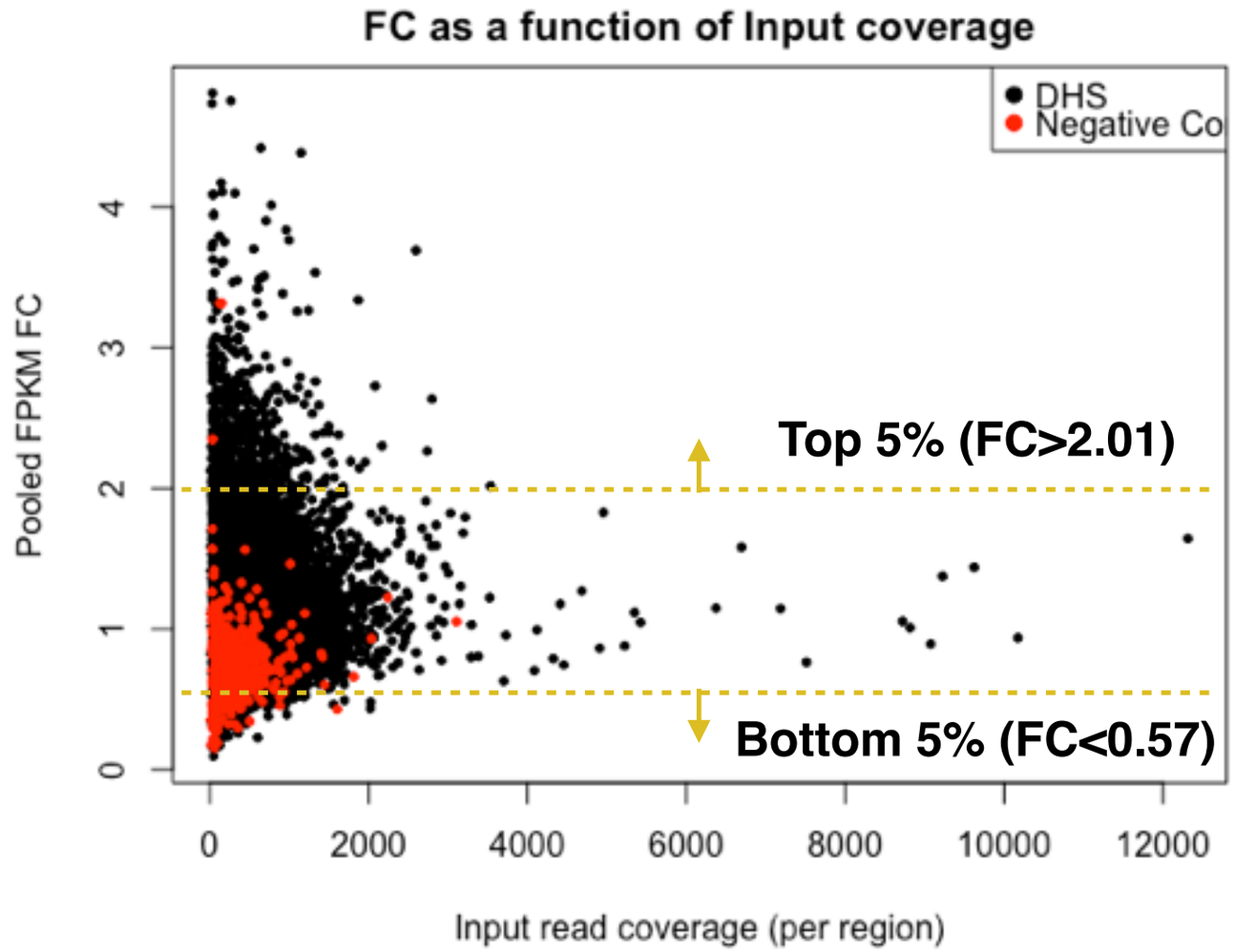


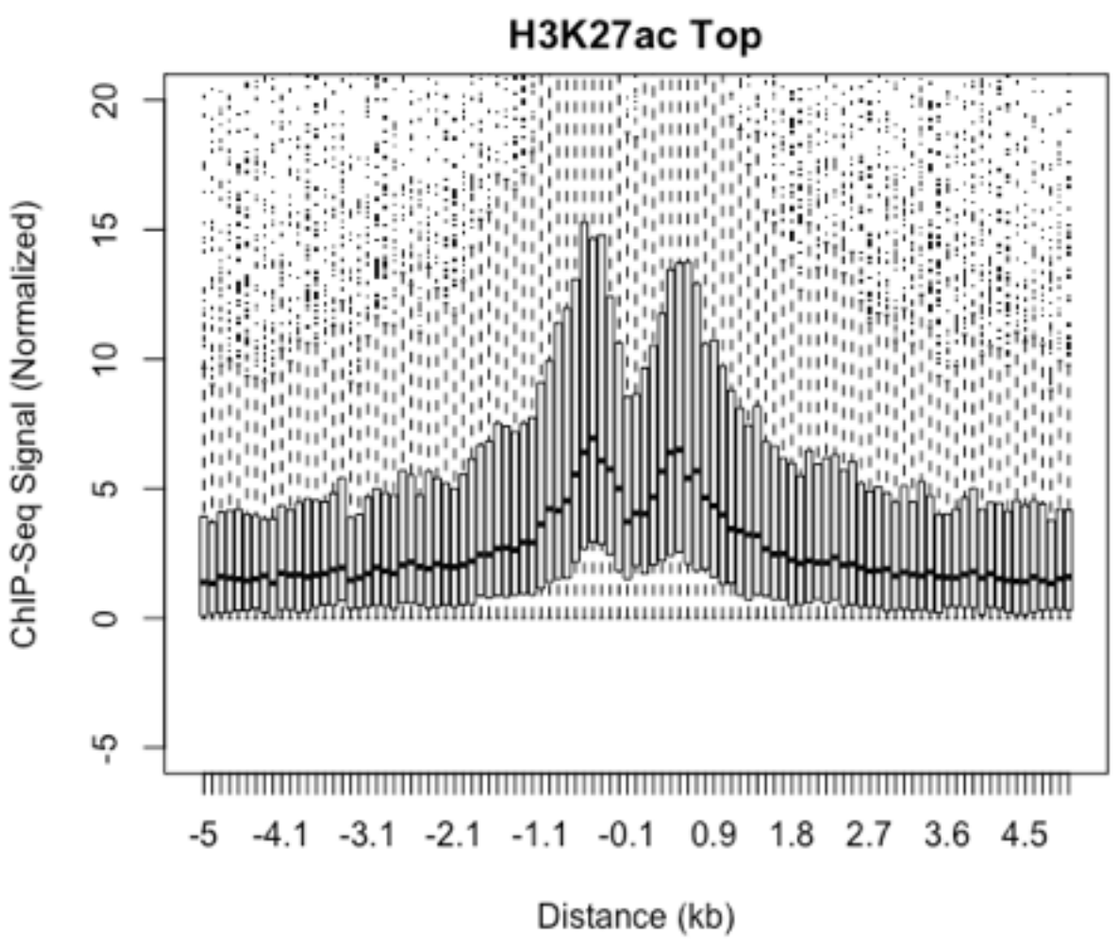
Remove input FPKM ≤ 1

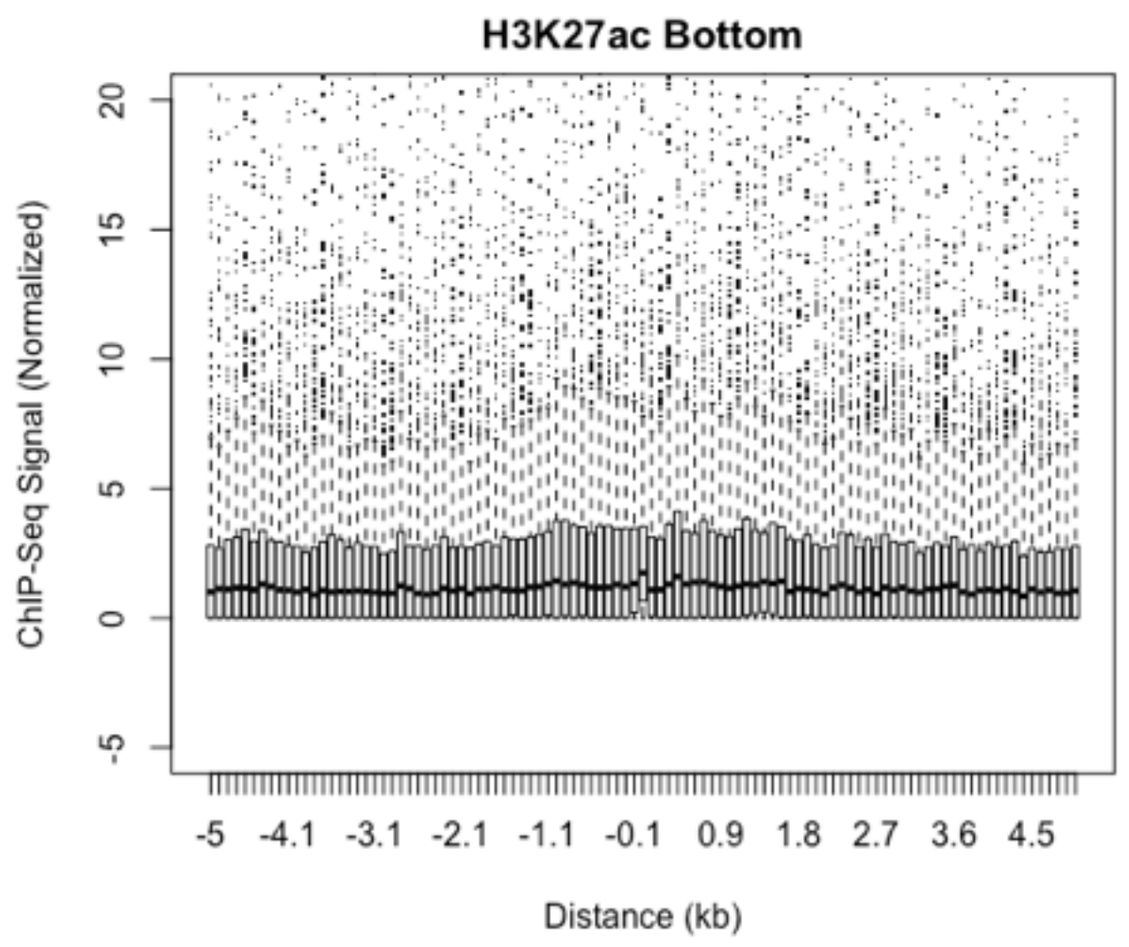


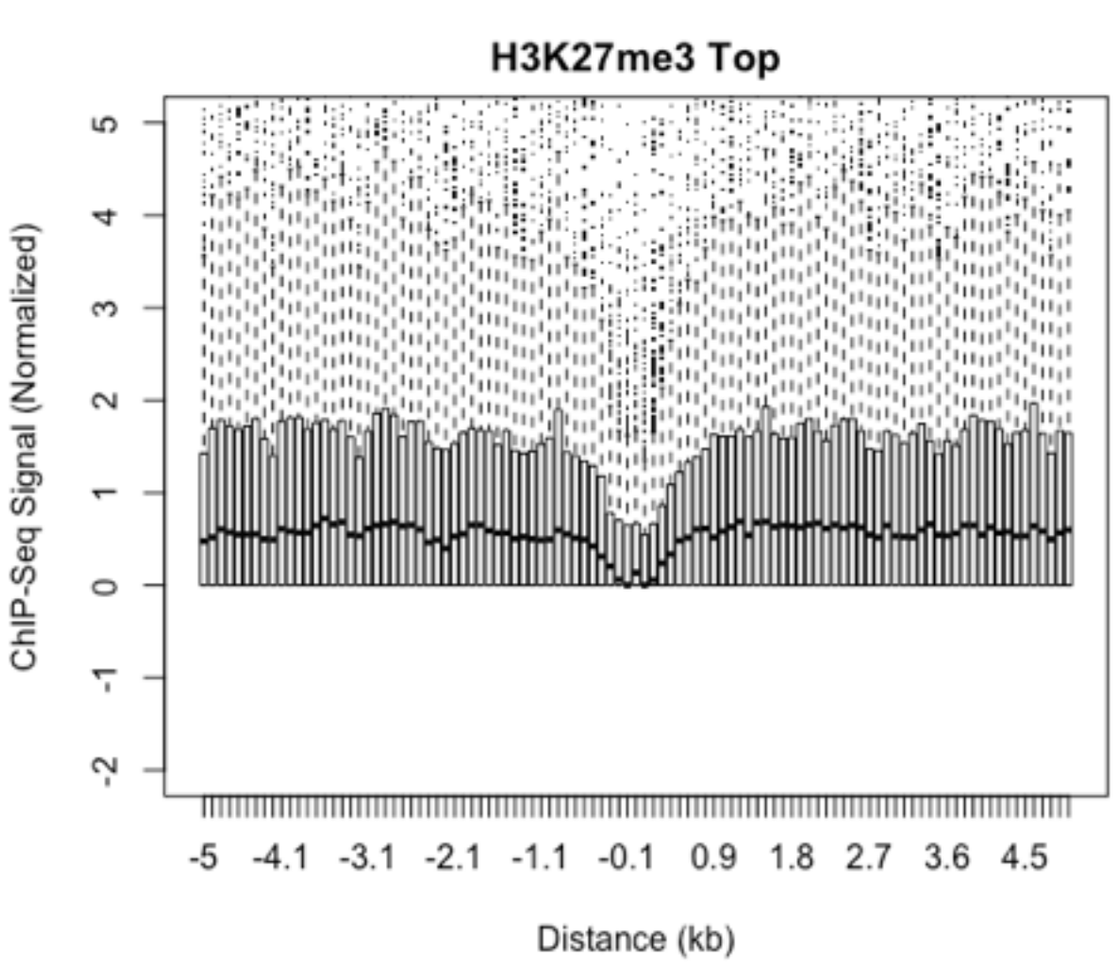
Stringent QC (input FPKM ≥ 10)

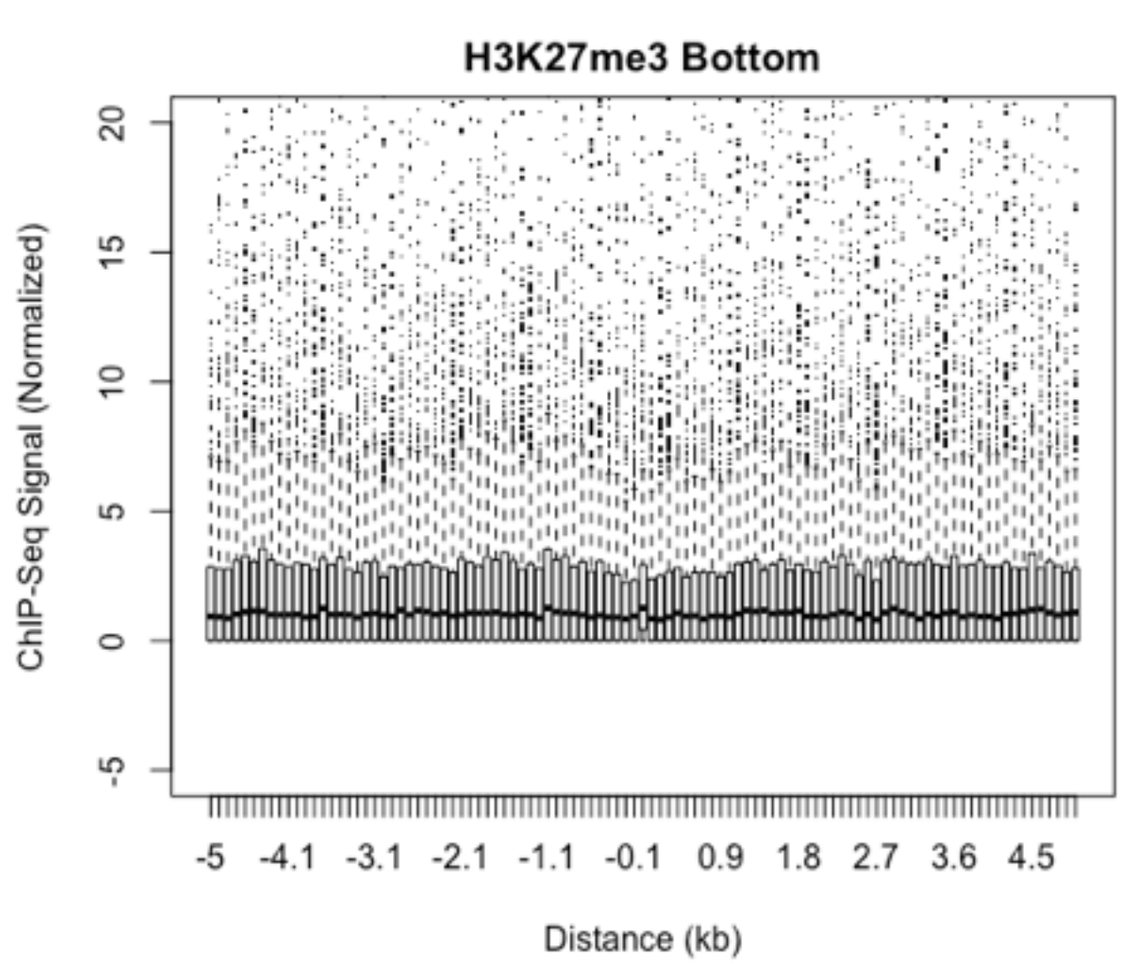


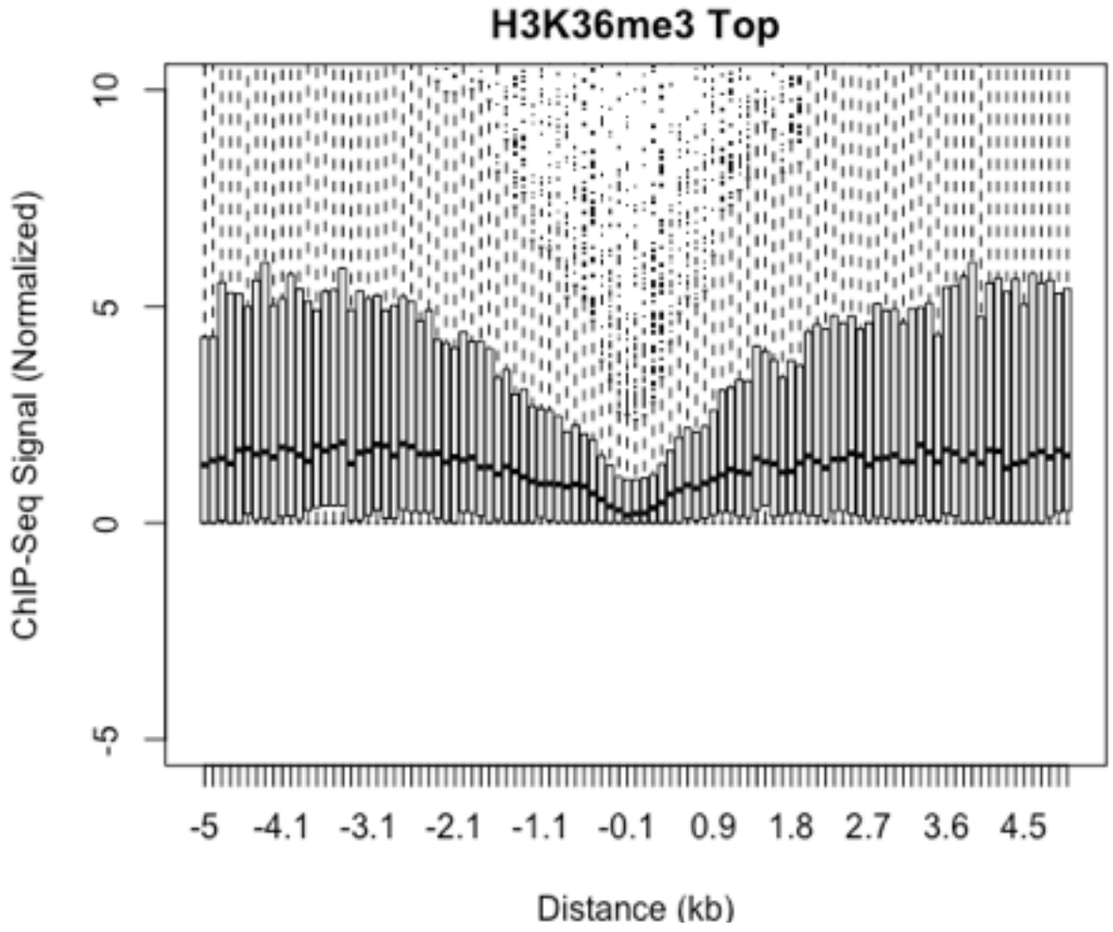


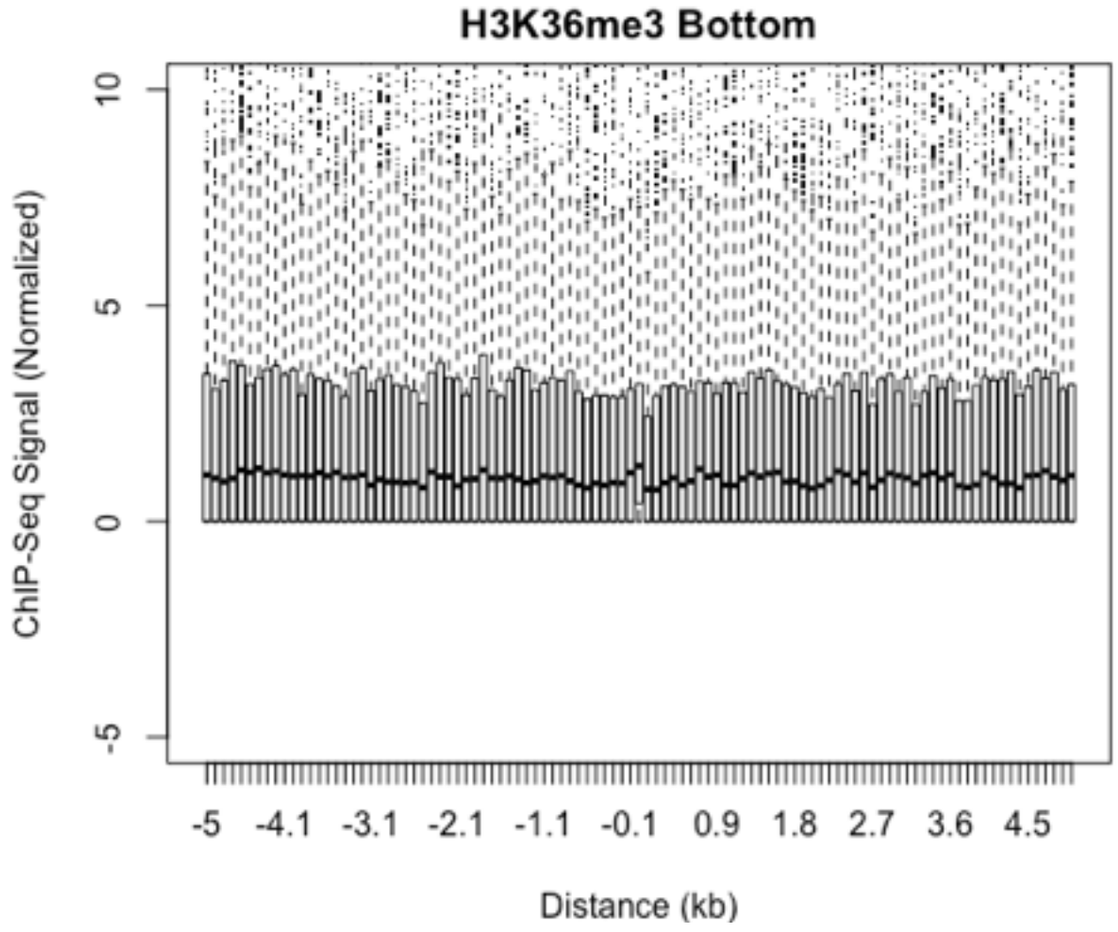

H3K27ac Top
ChIP-Seq Signal (Normalized)
-5
0
5
10
15
20
-5
-4.1
-3.1
-2.1
-1.1
-0.1
0.9
1.8
2.7
3.6
4.5
Distance (kb)

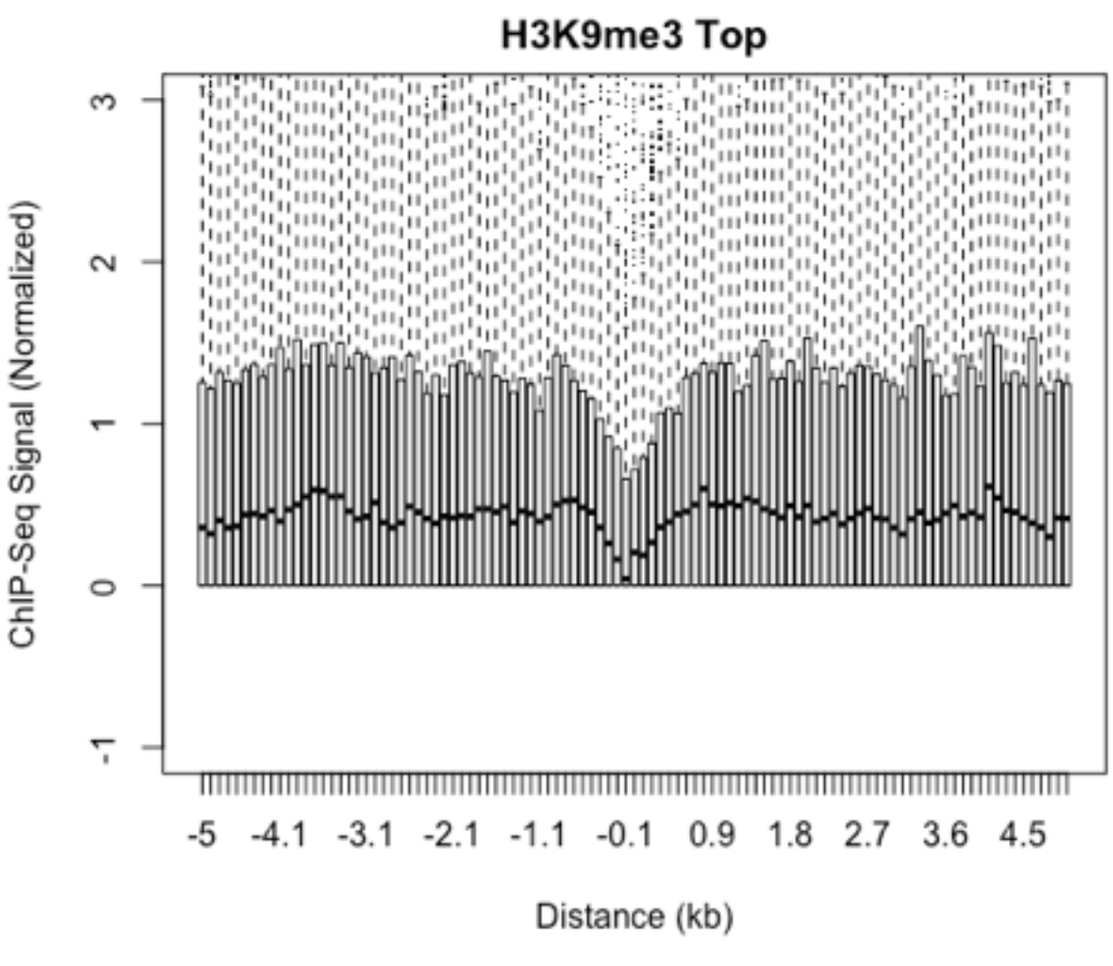

H3K27ac Bottom
ChIP-Seq Signal (Normalized)
-5
0
5
10
15
20
-5
-4.1
-3.1
-2.1
-1.1
-0.1
0.9
1.8
2.7
3.6
4.5
Distance (kb)

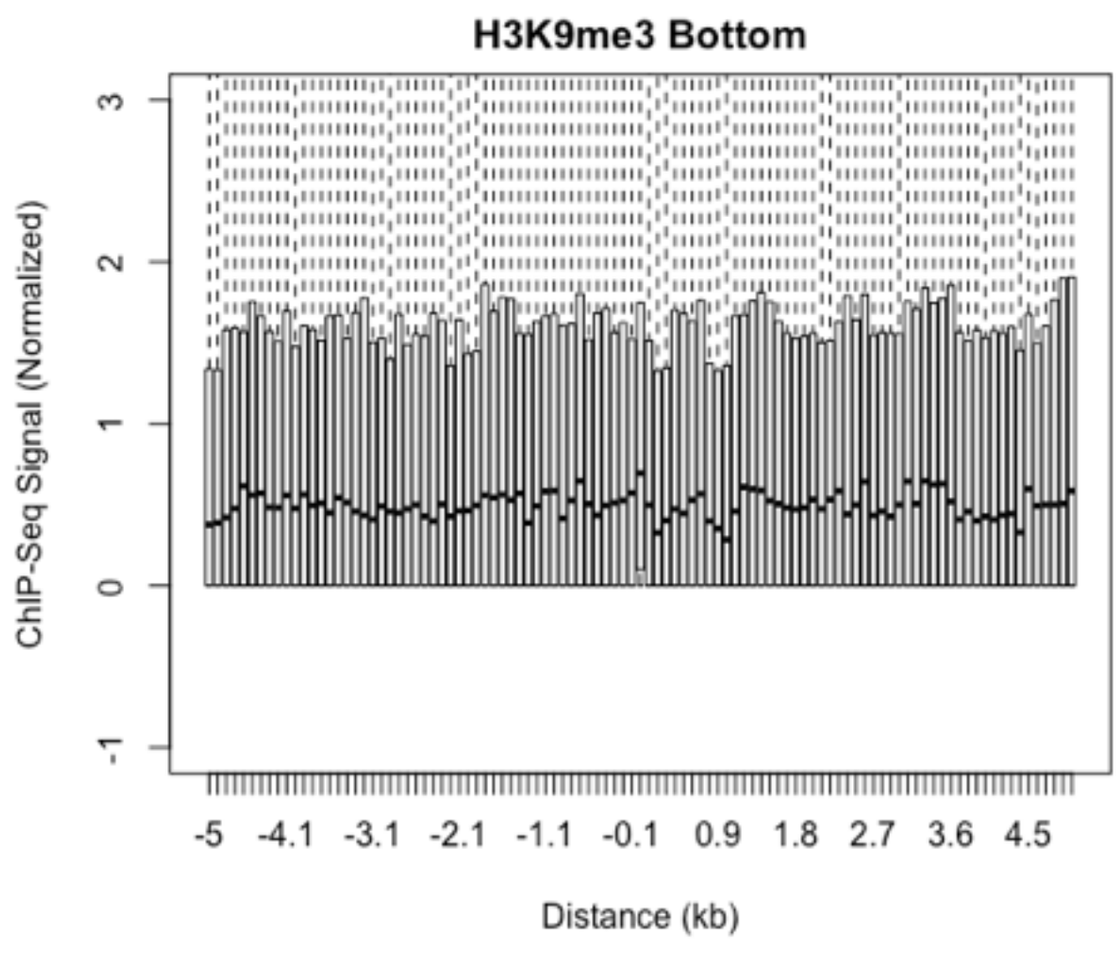

H3K27me3 Top
ChIP-Seq Signal (Normalized)
-2
-1
0
1
2
3
4
5
-5
-4.1
-3.1
-2.1
-1.1
-0.1
0.9
1.8
2.7
3.6
4.5
Distance (kb)

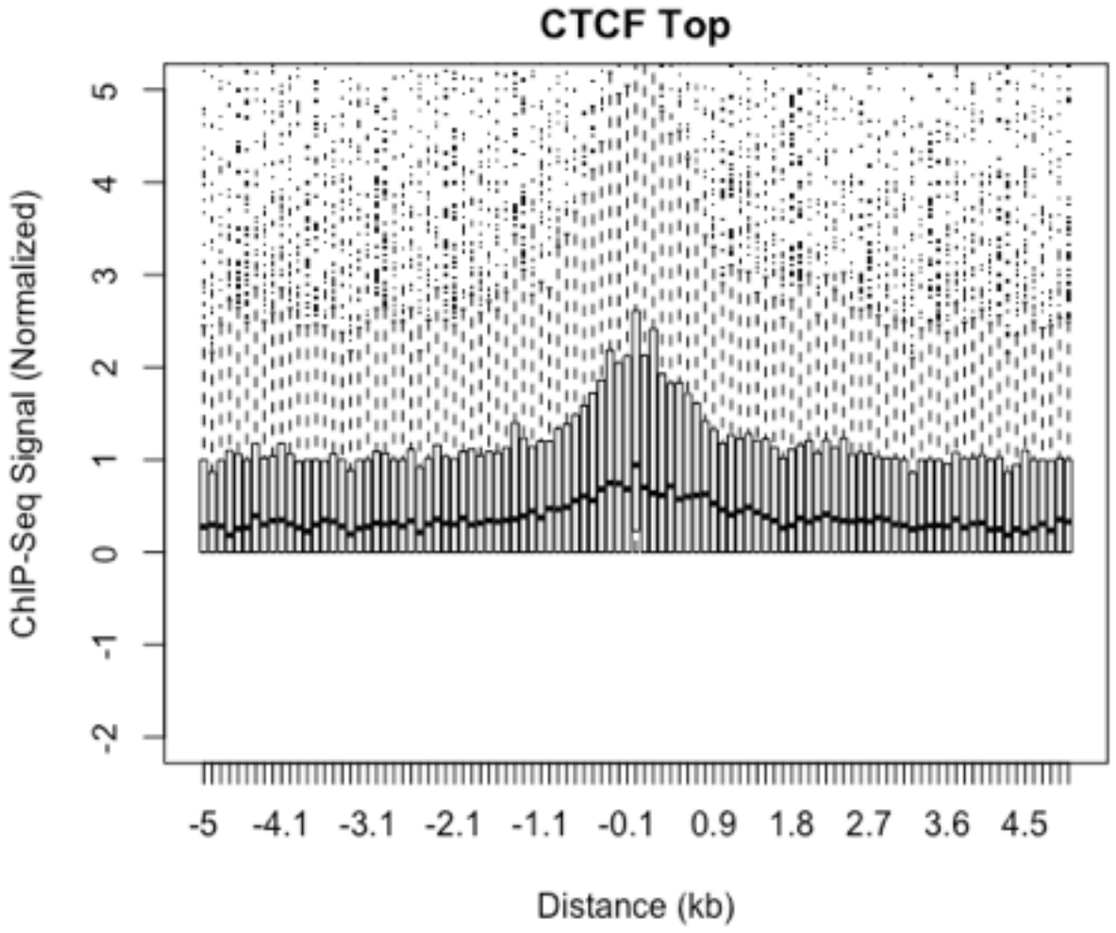

H3K27me3 Bottom
ChIP-Seq Signal (Normalized)
-5
0
5
10
15
20
-5
-4.1
-3.1
-2.1
-1.1
-0.1
0.9
1.8
2.7
3.6
4.5
Distance (kb)

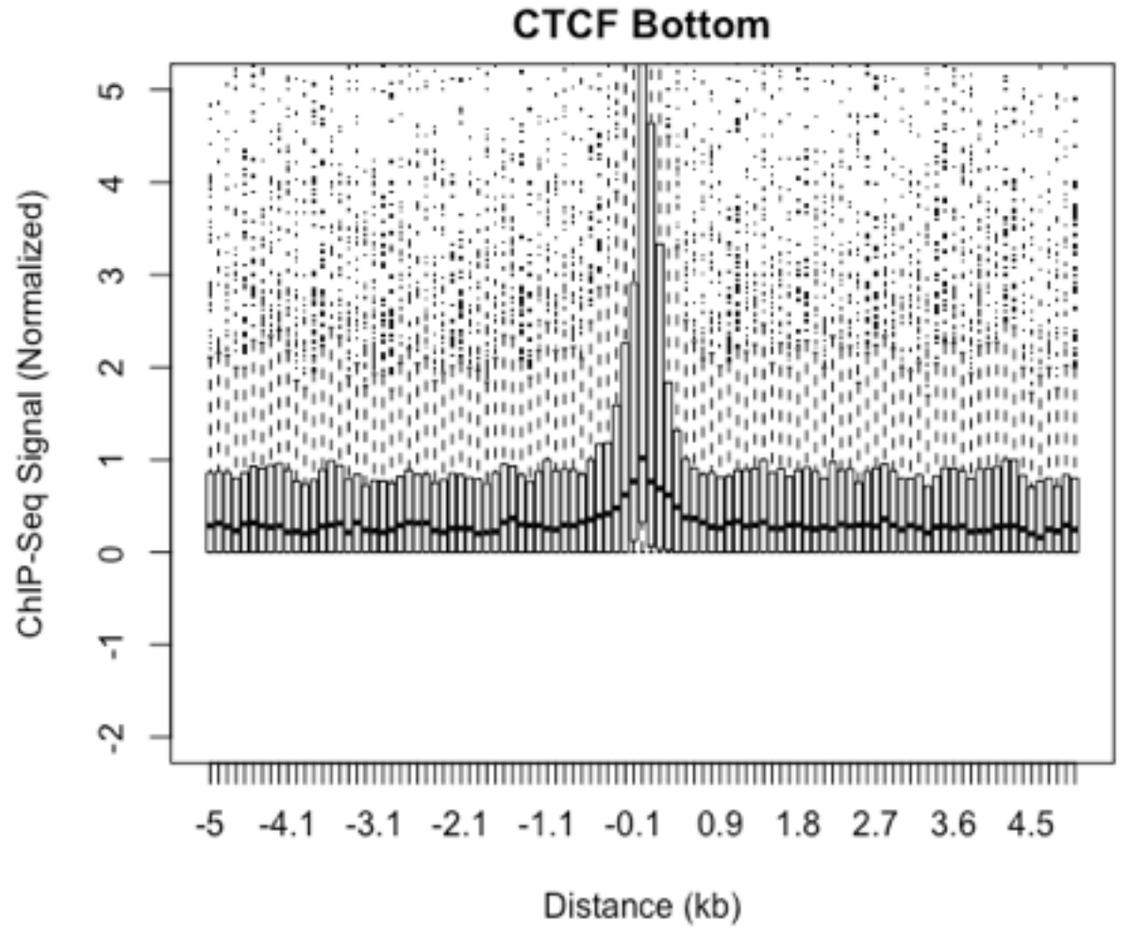

H3K36me3 Top
ChIP-Seq Signal (Normalized)
10
5
0
-5
-5 -4.1 -3.1 -2.1 -1.1 -0.1 0.9 1.8 2.7 3.6 4.5
Distance (kb)

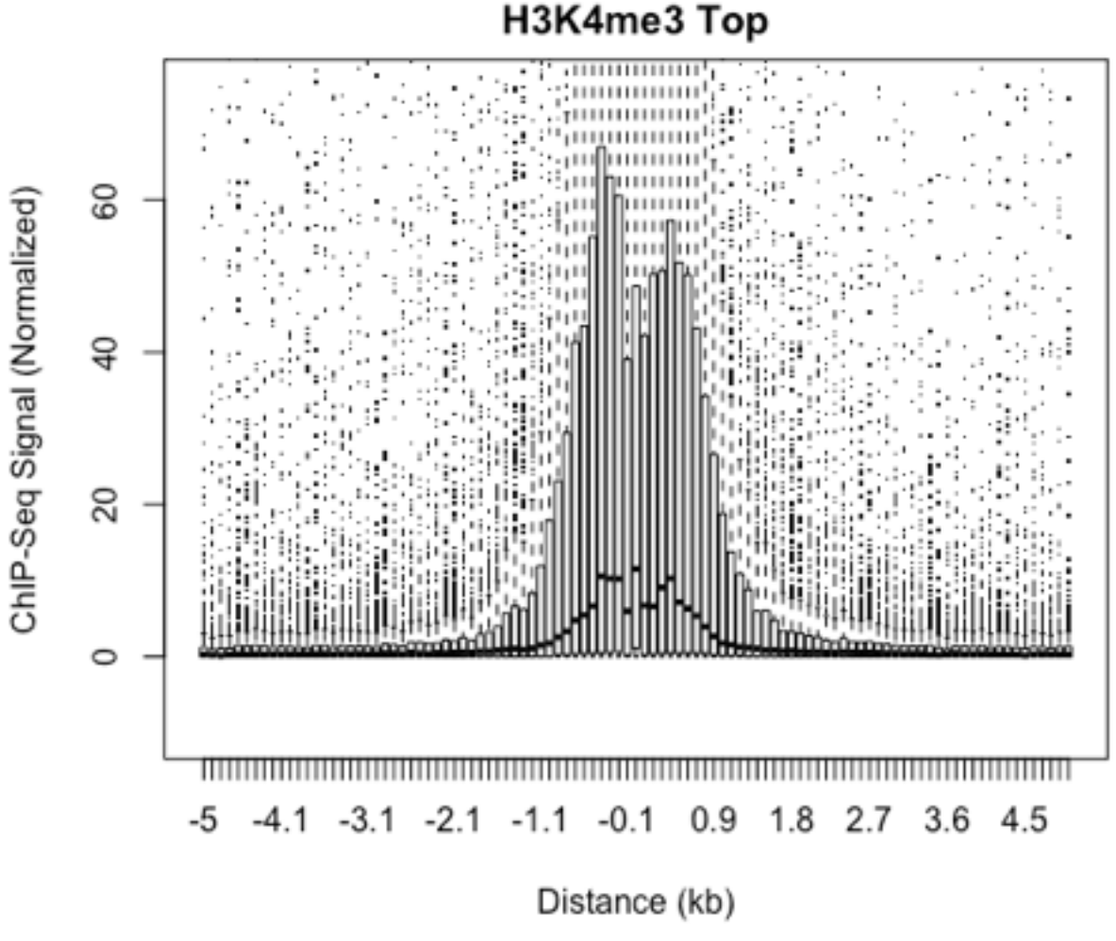

H3K36me3 Bottom
ChIP-Seq Signal (Normalized)
10
5
0
-5
-5 -4.1 -3.1 -2.1 -1.1 -0.1 0.9 1.8 2.7 3.6 4.5
Distance (kb)

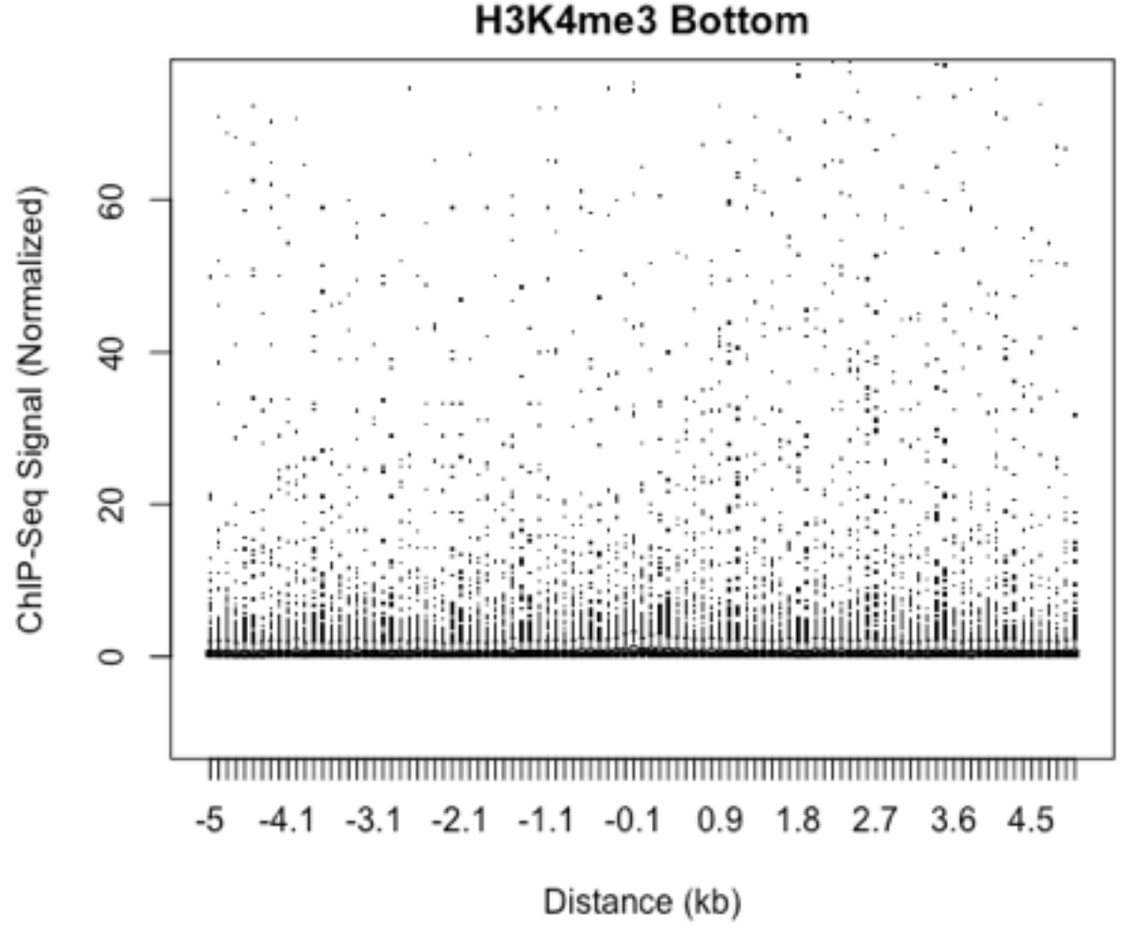

H3K9me3 Top
ChIP-Seq Signal (Normalized)
3
2
1
0
-1
-5 -4.1 -3.1 -2.1 -1.1 -0.1 0.9 1.8 2.7 3.6 4.5
Distance (kb)

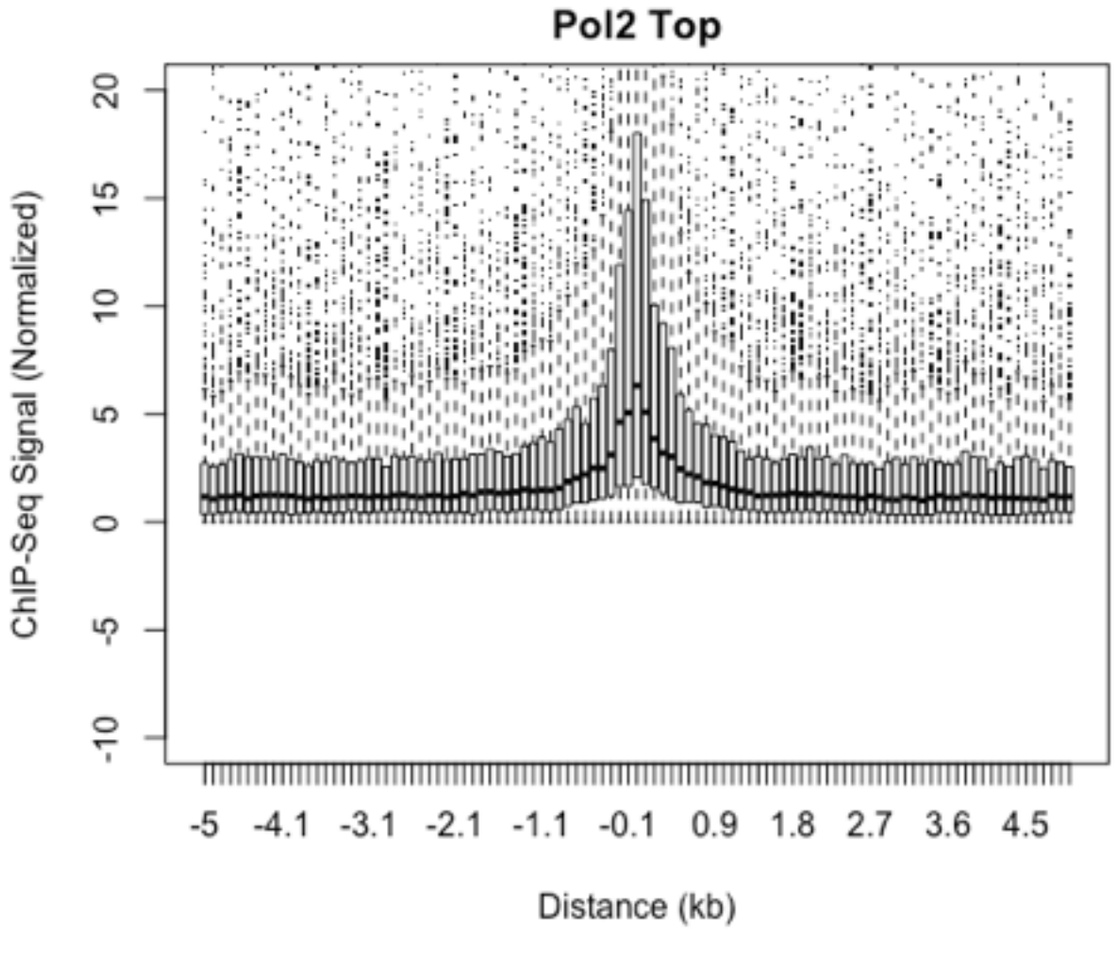

H3K9me3 Bottom
ChIP-Seq Signal (Normalized)
3
2
1
0
-1
-5 -4.1 -3.1 -2.1 -1.1 -0.1 0.9 1.8 2.7 3.6 4.5
Distance (kb)

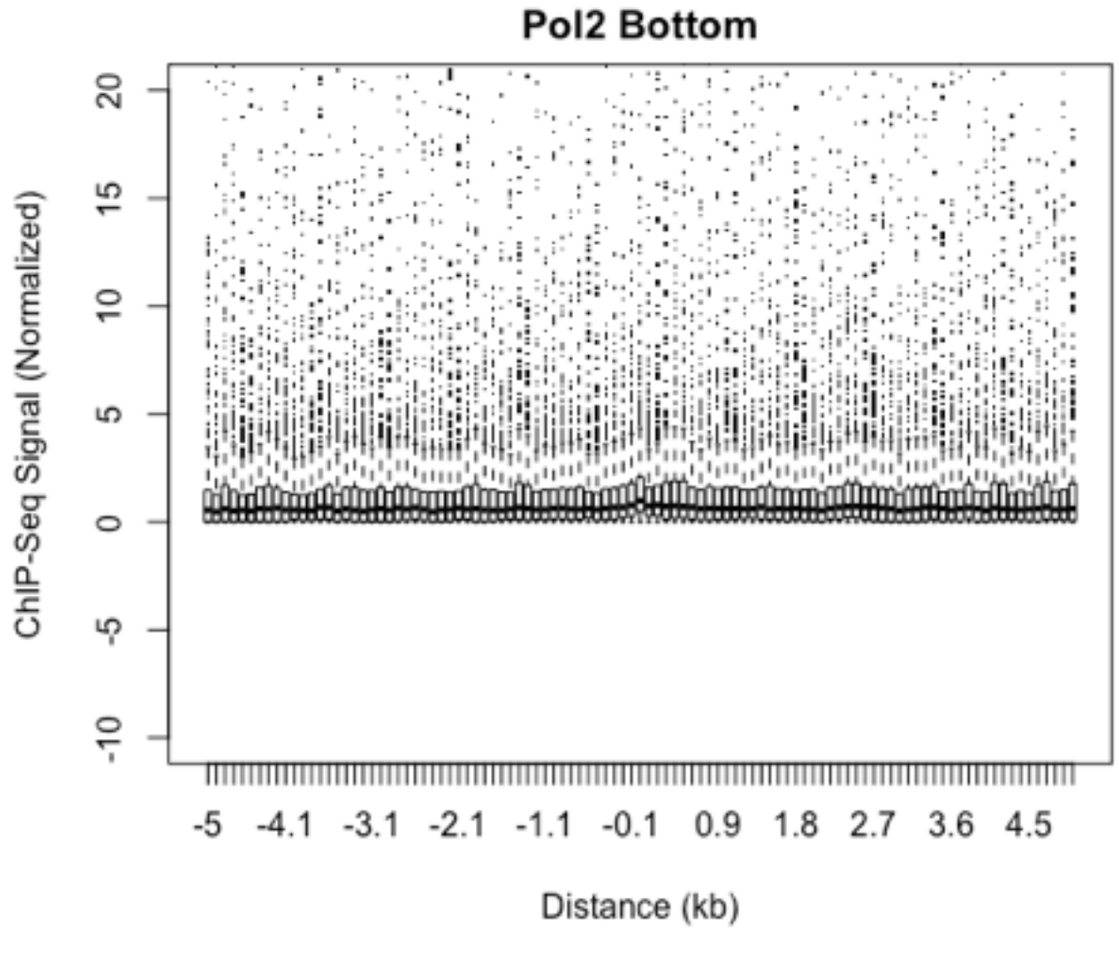

CTCF Top
ChIP-Seq Signal (Normalized)
-2
-1
0
1
2
3
4
5
-5
-4.1
-3.1
-2.1
-1.1
-0.1
0.9
1.8
2.7
3.6
4.5
Distance (kb)

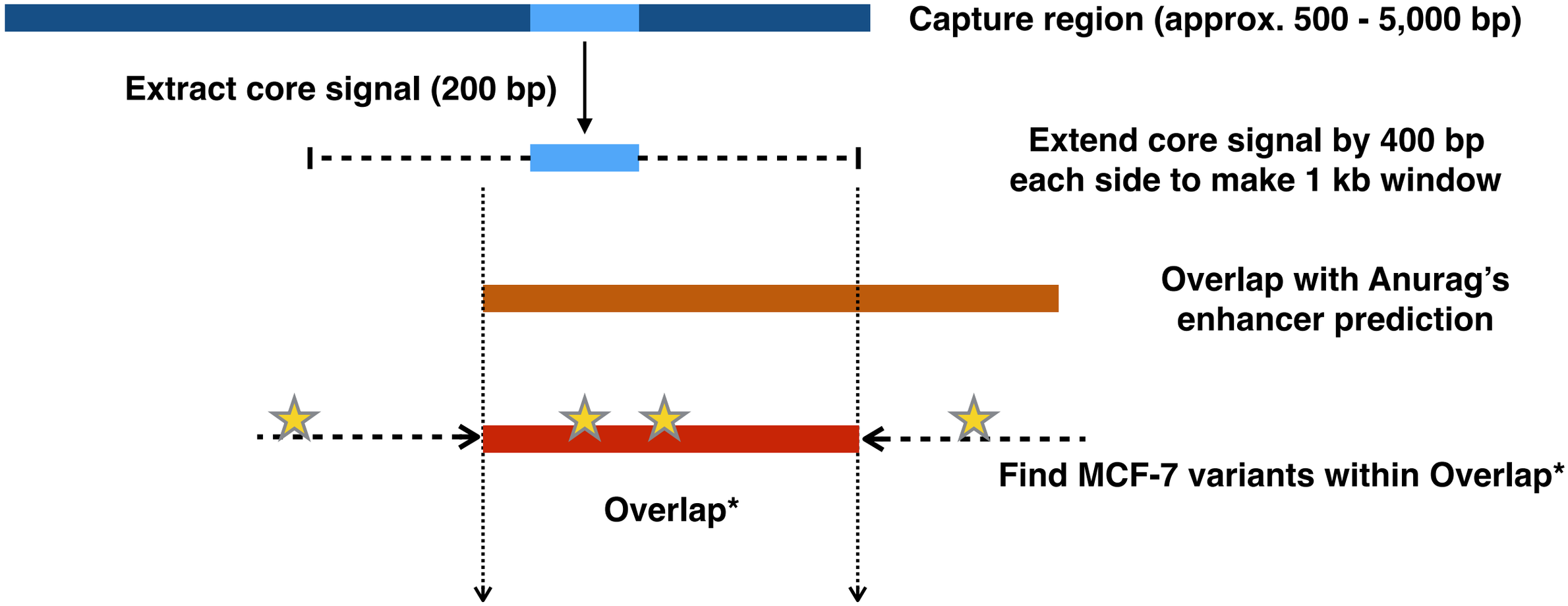

CTCF Bottom
ChIP-Seq Signal (Normalized)
-2
-1
0
1
2
3
4
5
-5
-4.1
-3.1
-2.1
-1.1
-0.1
0.9
1.8
2.7
3.6
4.5
Distance (kb)


H3K4me3 Top
ChIP-Seq Signal (Normalized)
0
20
40
60
-5
-4.1
-3.1
-2.1
-1.1
-0.1
0.9
1.8
2.7
3.6
4.5
Distance (kb)


H3K4me3 Bottom
ChIP-Seq Signal (Normalized)
0
20
40
60
-5
-4.1
-3.1
-2.1
-1.1
-0.1
0.9
1.8
2.7
3.6
4.5
Distance (kb)


Pol2 Top
ChIP-Seq Signal (Normalized)
20
15
10
5
0
-5
-10
-5
-4.1
-3.1
-2.1
-1.1
-0.1
0.9
1.8
2.7
3.6
4.5
Distance (kb)


Pol2 Bottom
ChIP-Seq Signal (Normalized)
20
15
10
5
0
-5
-10
-5
-4.1
-3.1
-2.1
-1.1
-0.1
0.9
1.8
2.7
3.6
4.5
Distance (kb)



	MCF-7 variants within Overlap*		Enhancer Overlap*	
	Top 5% FC	Bottom 5% FC	Top 5% FC	Bottom 5% FC
Using input FPKM>1	80	4	325 / 541 (60.07%)	28 / 541 (5.18%)
Using input FPKM≥10	97	4	356 / 538 (66.17%)	28 / 538 (5.20%)

Acknowledgement

- **Mark Gerstein**
  - Jing Zhang
  - Anurag Sethi
  - Joel Rozowsky
  - Arif Harmanci
  - Michael R. Schoenberg
- **Kevin White**
  - Grace Yu