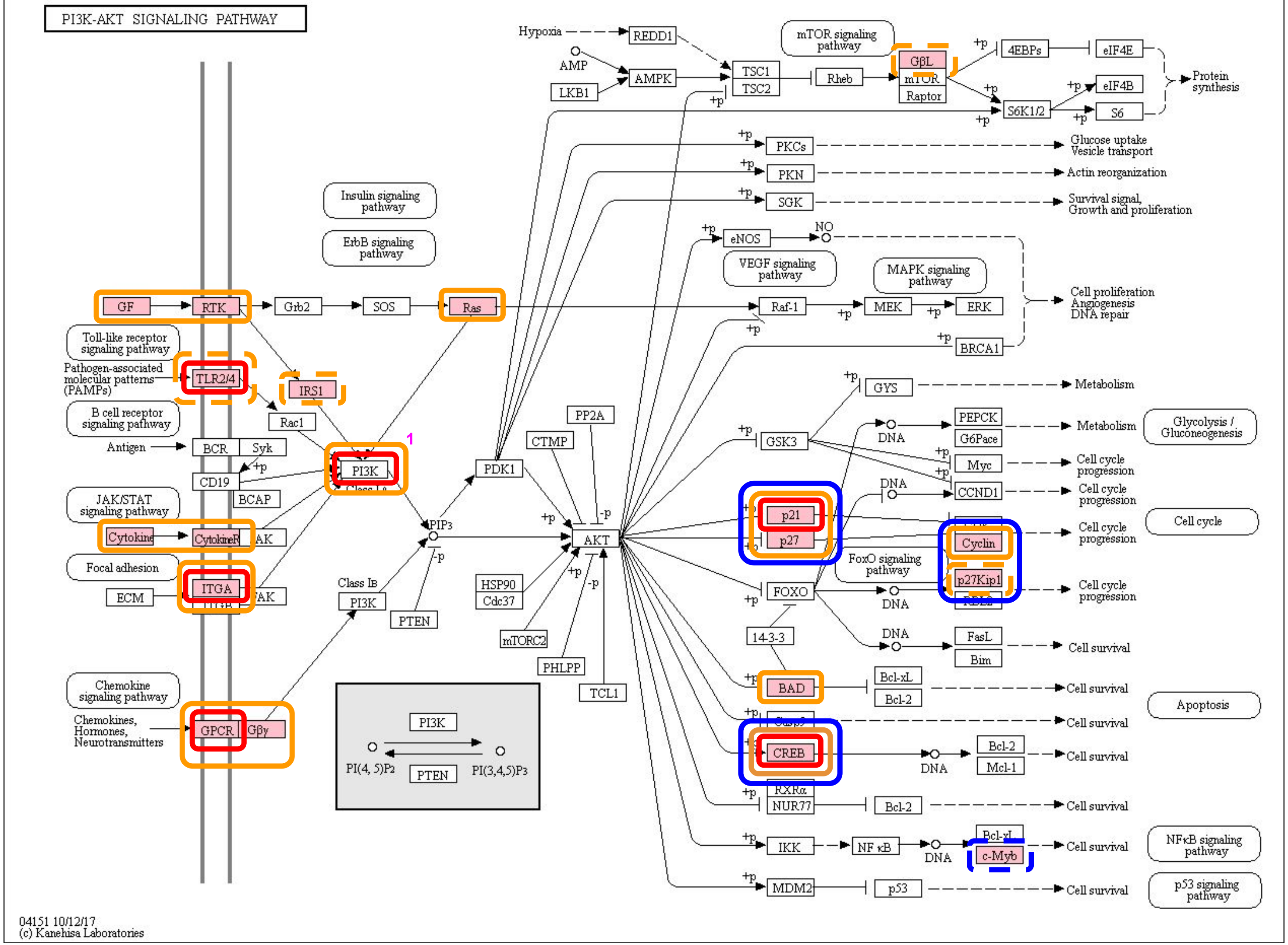


Differentially expressed genes within the PI3K-AKT signaling pathway:

FDR <=0.05, abs(LFC)>=1

Genes that decrease in expression with DMSO treatment

earlyG1

lateG1

SG2M

Genes that increase in expression with DMSO treatment

earlyG1

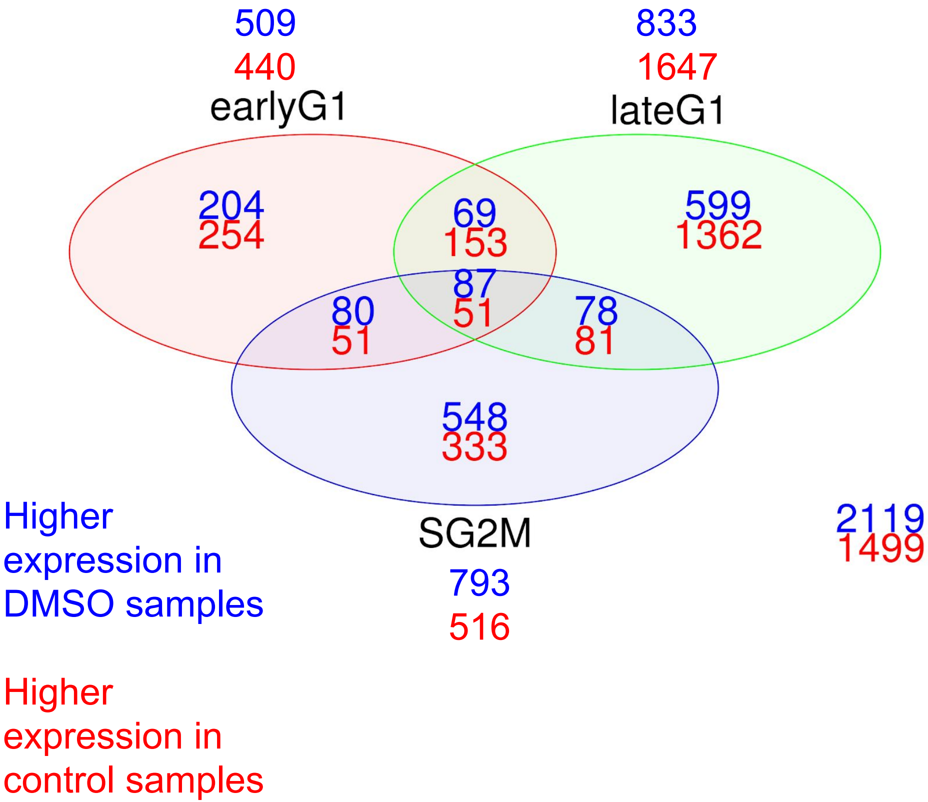
lateG1

SG2M

Number of differentially expressed genes in DMSO treated vs control samples

FDR <=0.05, abs(LFC)>=1

(RNA-seq assay)



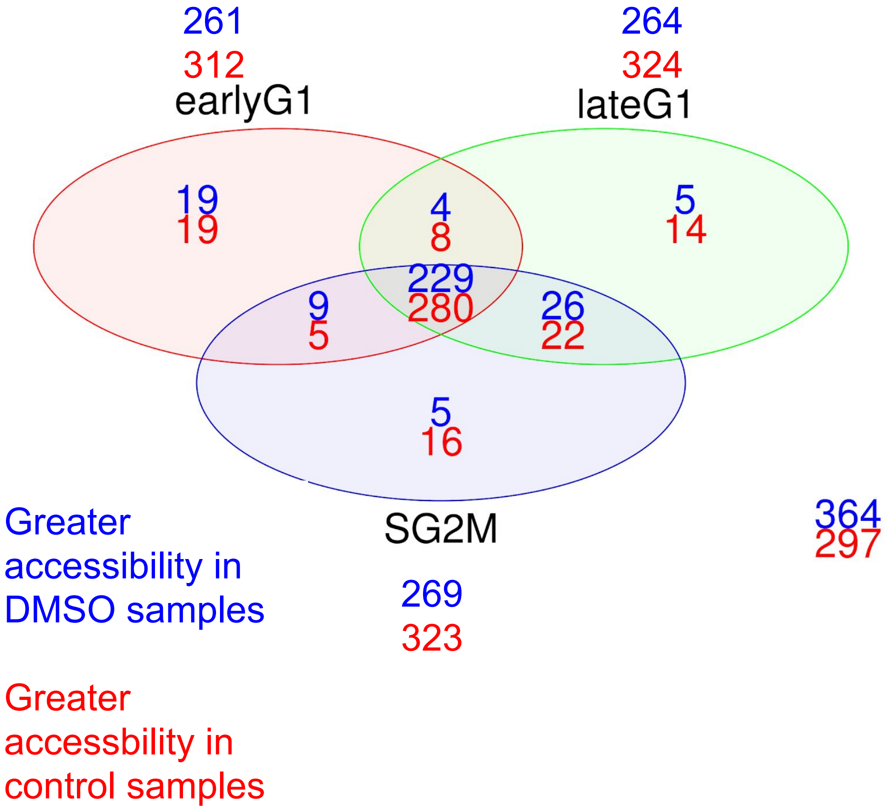
Higher expression in DMSO samples

Higher expression in control samples

Number of differentially accessible genome regions in DMSO treated vs control samples

FDR <=0.05, abs(LFC)>=1

(ATAC-seq assay)



Greater accessibility in DMSO samples

Greater accessbility in control samples

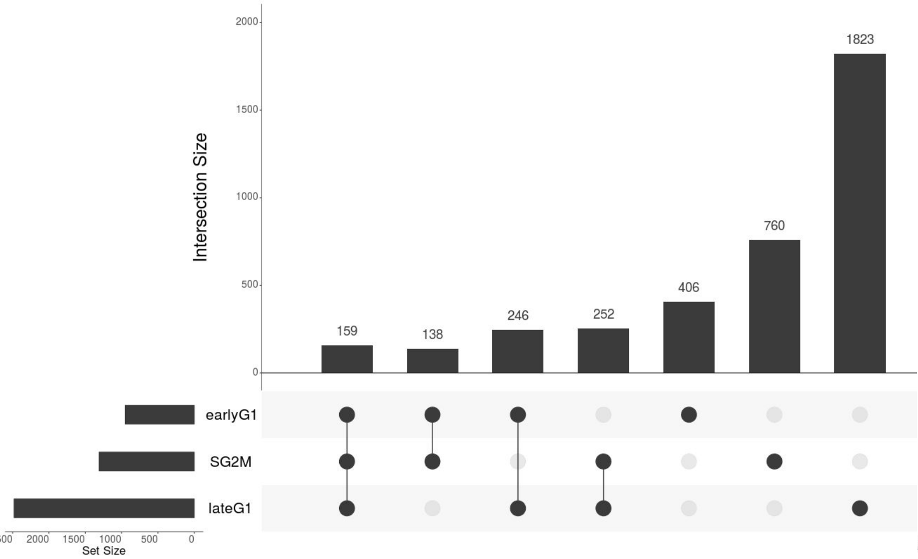
RNA-seq assay

	Upregulated in earlyG1	Upregulated in lateG1	Upregulated in SG2M
Control # of differentially expressed genes	440	1647	516
DMSO # of differentially expressed genes	509	833	793

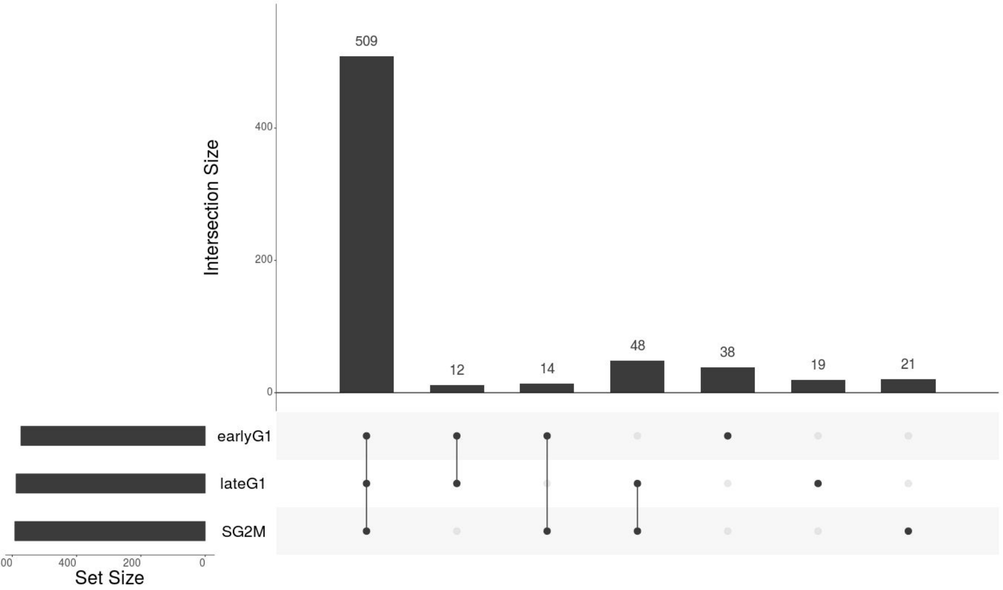
ATC-seq assay

	Upregulated in earlyG1	Upregulated in lateG1	Upregulated in SG2M
Control # of differentially accessible genomic regions	312	324	323
DMSO # of differentially accessible genomic regions	261	264	269

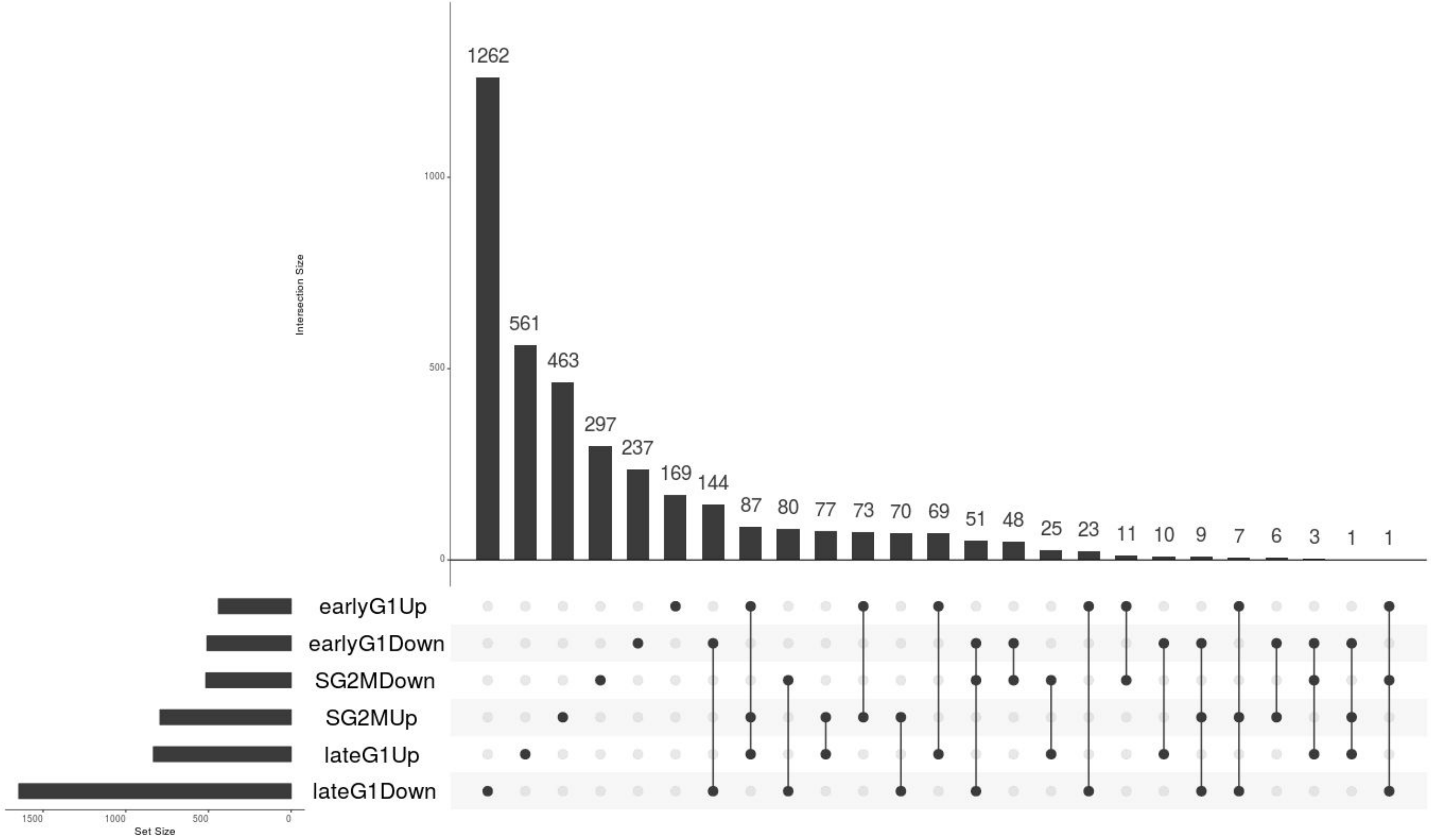
Differentially expressed genes



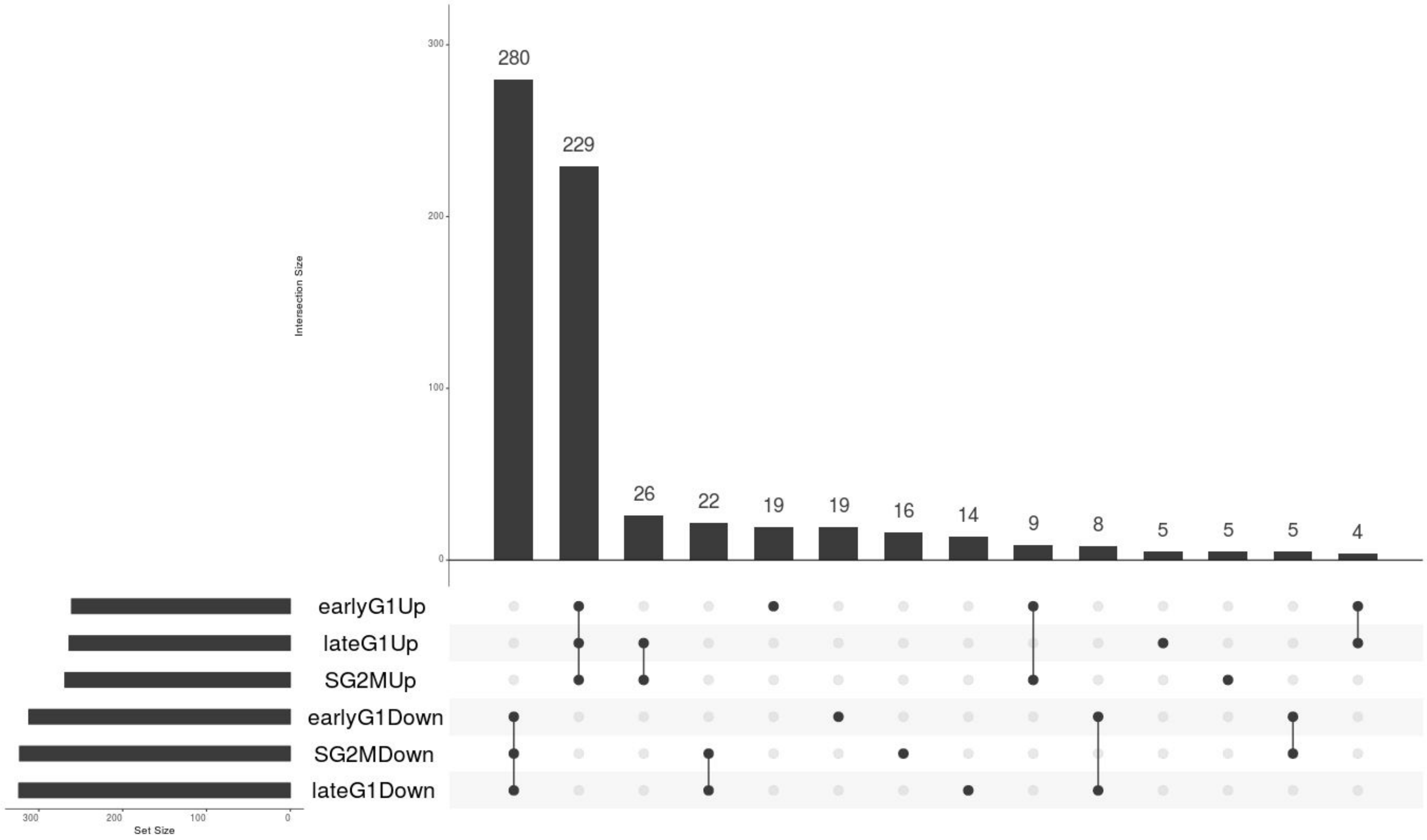
Differentially accessible chromatin regions



Differentially expressed genes

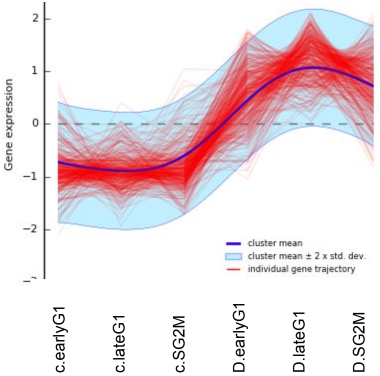


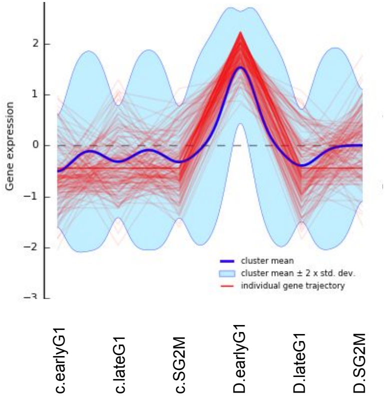
Differentially accessible chromatin regions

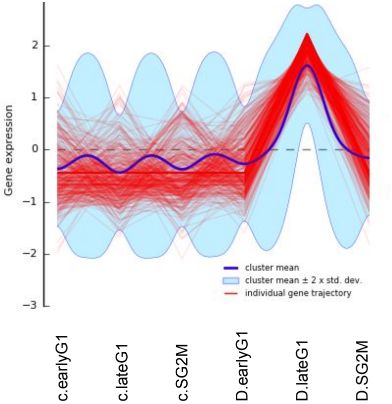


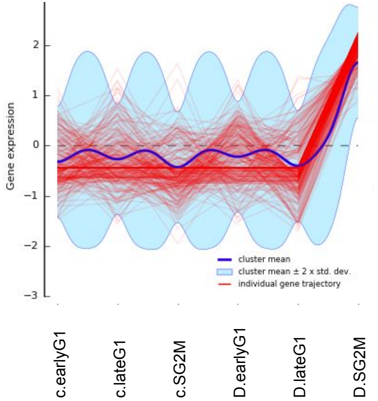
Clustering of differential genes via Dirichlet process Gaussian process mixture model[1]

Expression up in DMSO samples

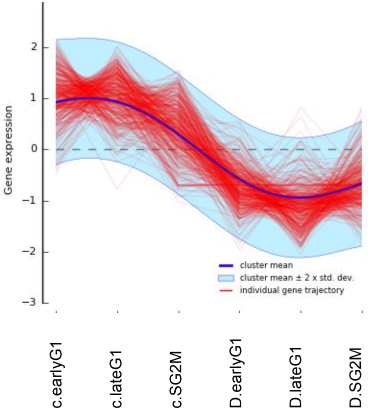


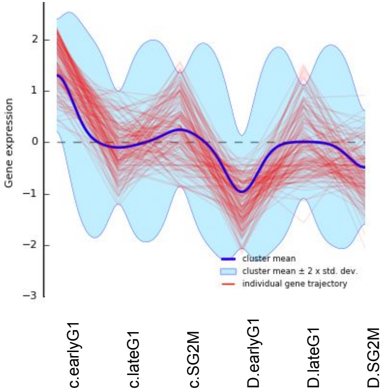


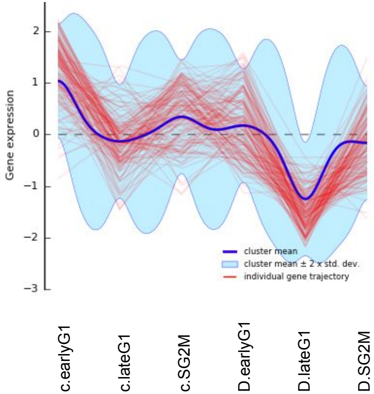




Expression down in DMSO samples



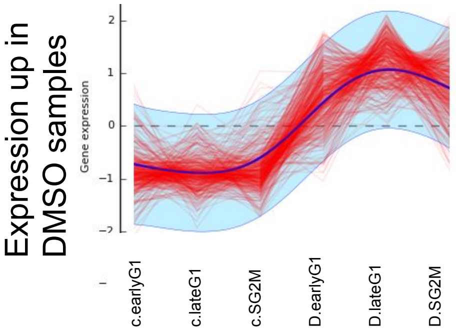




---

[1] https://www.biorxiv.org/content/early/2017/04/26/131151

Clustering of differential genes via Dirichlet process Gaussian process mixture model[1]



Genes upregulated with DMSO treatment at each phase of the cell cycle

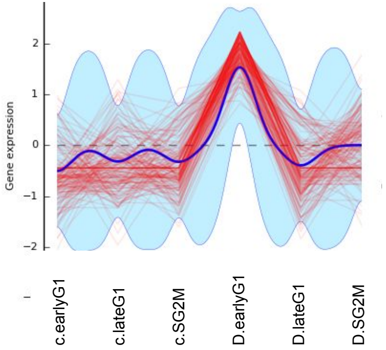
GO Term	**FDR q-value**
Phosphoprotein	3.90e-8
Integral component of plasma membrane	9.85e-4
Cell surface	1.53e-3
Extracellular matrix	1.49e-2
Sodium ion binding	1.69e-2
I band	3.96e-2

---

1. https://www.biorxiv.org/content/early/2017/04/26/131151

Clustering of differential genes via Dirichlet process Gaussian process mixture model[1]

Expression up in DMSO samples



Genes upregulated with DMSO treatment at earlyG1 phase of the cell cycle

GO Term	**FDR q-value**
Plasma membrane	7.9e-3

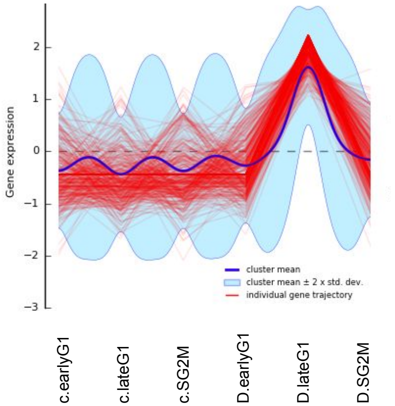
---

[1] https://www.biorxiv.org/content/early/2017/04/26/131151

Clustering of differential genes via Dirichlet process Gaussian process mixture model[1]

Expression up in DMSO samples

GO Term	FDR q-value
Glycoprotein	1.5e-5
Signal peptide	1.2e-4
Extracellular topological domain	1.6e-4
Neuroactive ligand-receptor interaction	2.0e-3
Transmembrane signaling receptor activity	8.7e-3
Signaling receptor activity	2.54e-2
G-protein coupled receptor activity	1.65e-2



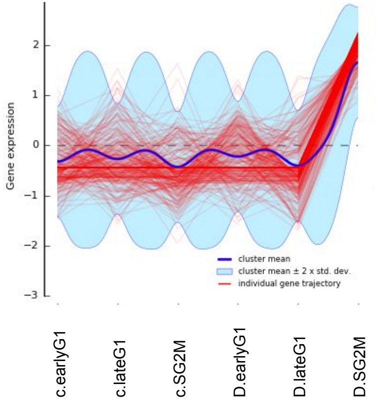
Genes upregulated with DMSO treatment at lateG1 phase of the cell cycle

1. https://www.biorxiv.org/content/early/2017/04/26/131151

Clustering of differential genes via Dirichlet process Gaussian process mixture model[1]

Expression up in DMSO samples

GO Term	FDR q-value
Glycosylation sites	1.40e-7
Signal peptide	3.3e-6
Disulfide bond	3.0e-4
Detection of chemical stimulus involved in sensory perception	4.17e-3
Signaling	3.6e-3
Transmembrane region	4.9e-3
G-protein coupled receptor signaling pathway	3.41e-2



Genes upregulated with DMSO treatment at SG2M phase of the cell cycle

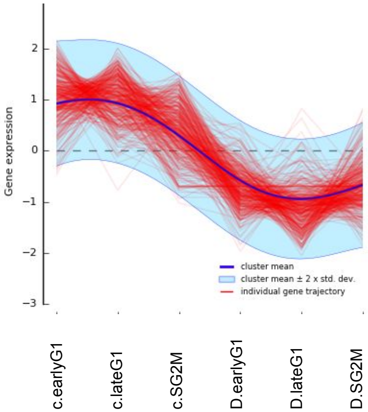
---

[1] https://www.biorxiv.org/content/early/2017/04/26/131151

Clustering of differential genes via Dirichlet process Gaussian process mixture model[1]

Genes downregulated with DMSO treatment at each phase of the cell cycle

Expression down in DMSO samples



GO Term	FDR q-value
Alternative splicing	4.2e-5
Cytoplasm	1.1e-2
Protein binding	1.6e-2
Centrosome	1.8e-2
Phosphoprotein	4.6e-2

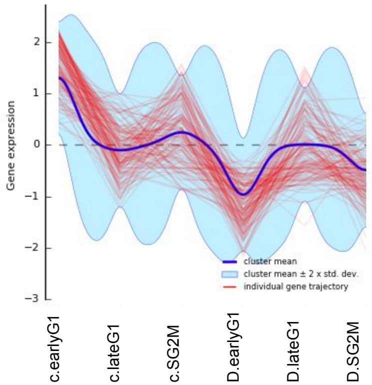
---

1. https://www.biorxiv.org/content/early/2017/04/26/131151

Clustering of differential genes via Dirichlet process Gaussian process mixture model[1]

Expression down in
DMSO samples

Genes downregulated with DMSO treatment in earlyG1 phase of the cell cycle



GO Term	FDR q-value
Cytosol	5.2e-4
Acetylation	8.7e-3
Alternative initiation	7.0e-2

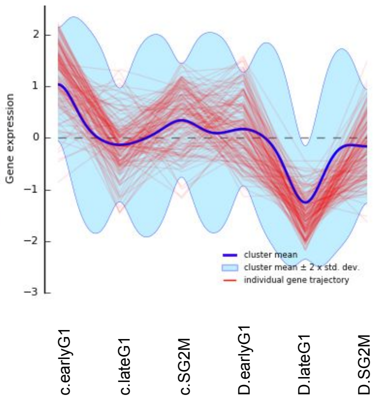
1. https://www.biorxiv.org/content/early/2017/04/26/131151

Clustering of differential genes via Dirichlet process Gaussian process mixture model[1]

Expression down in DMSO samples

GO Term	FDR q-value
Transcription regulator SCAN	2.9e-3
Alternative splicing	7.4e-3
Metabolic pathways	8.5e-3
SCAN box domain	2.4e-2

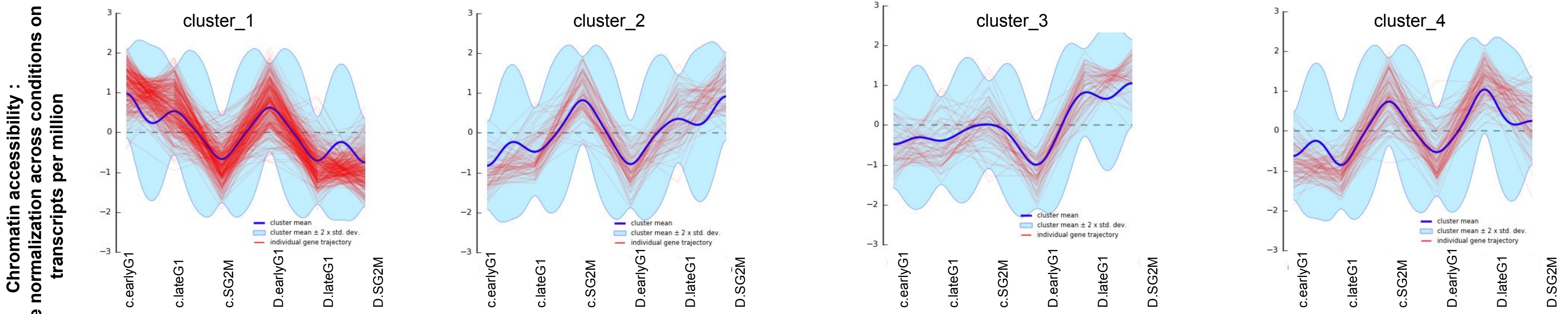
Genes downregulated with DMSO treatment in the lateG1 phase of the cell cycle



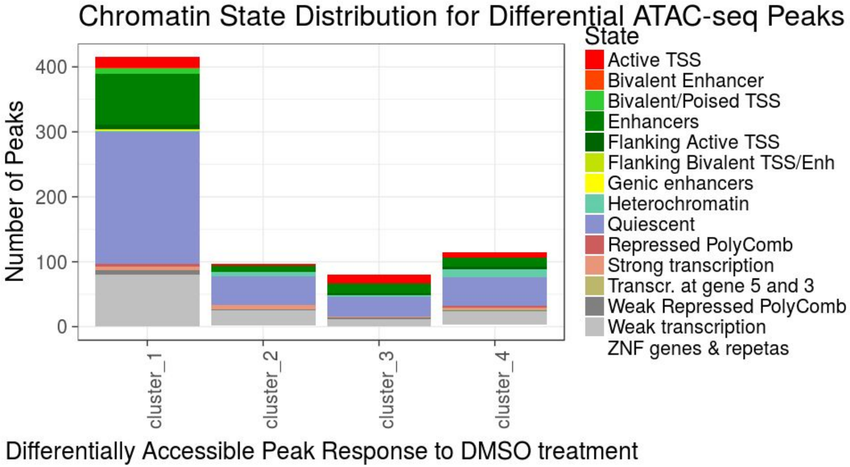
---

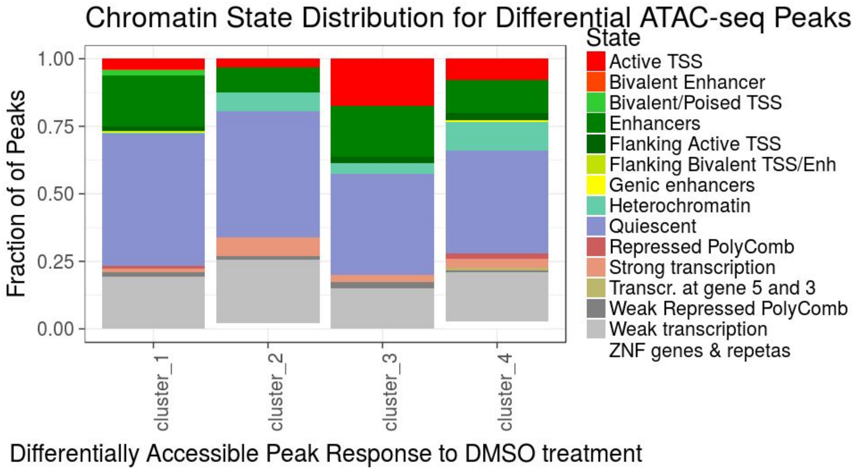
1. https://www.biorxiv.org/content/early/2017/04/26/131151

Clustering of differentially accessible genomic regions via Dirichlet process Gaussian process mixture model[1]



Per-cluster chromatin state distribution from 15-state ChromHMM Model

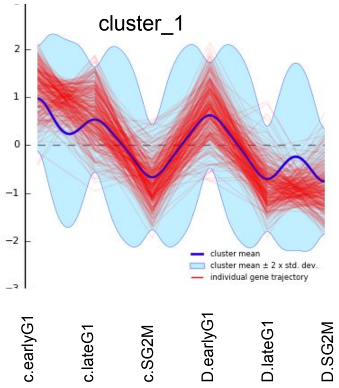




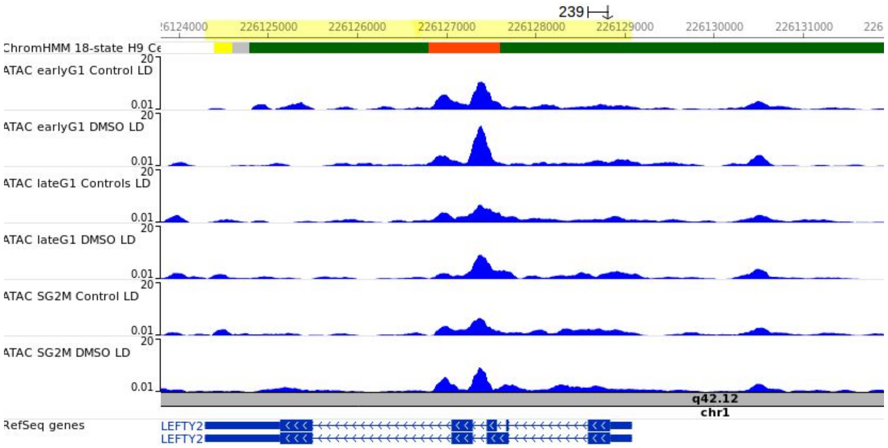
---

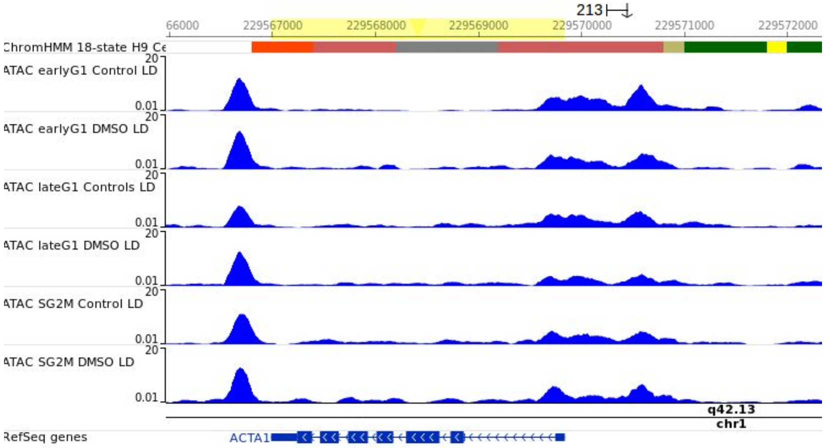
1. https://www.biorxiv.org/content/early/2017/04/26/131151

Peak-gene associations within ATAC-seq data clusters: Differential genes with strong peak associations

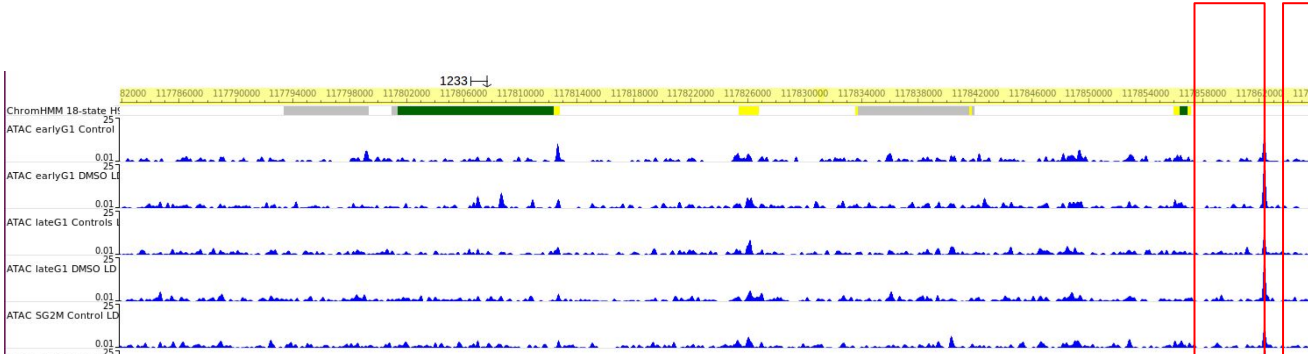


Pathway	Differentially expressed genes on pathway with peaks associations in cluster
Cilium Assembly 1, C	IFT22, IFT80
GPCR ligand binding	OR2A7, UTS2, LPAR3, EDN1
PI3K-AKT signaling pathway	CSF1, ITGA9, MYB, PCK2, TNC, EGF2
Wnt signaling pathway	ACTA1, PCDH7, SIAH2, GNA14, GNG2
VEGFR signaling pathway	CAV1, FGF2

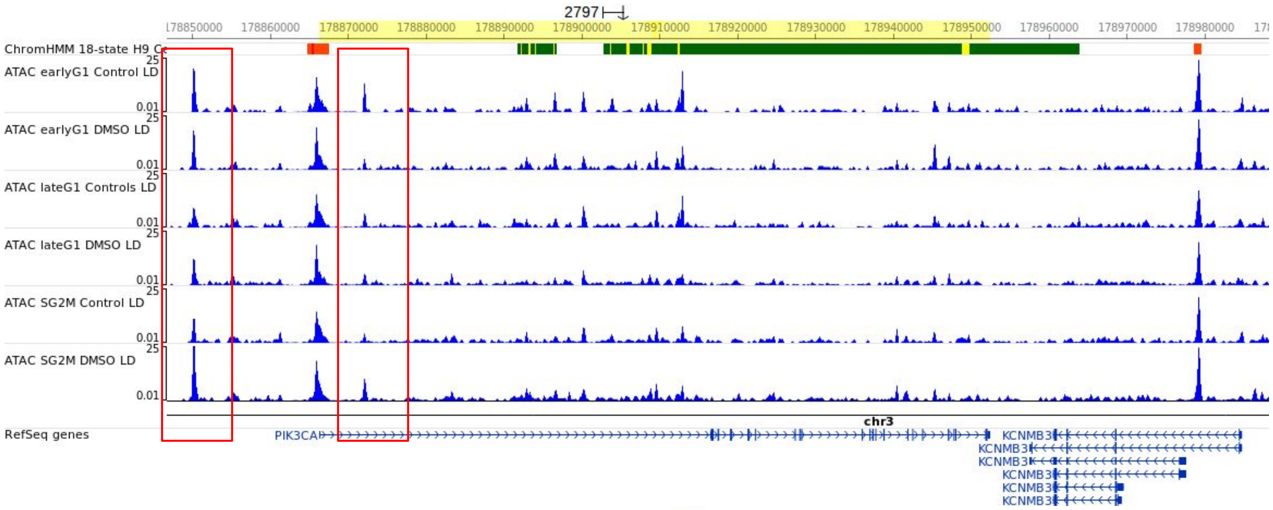


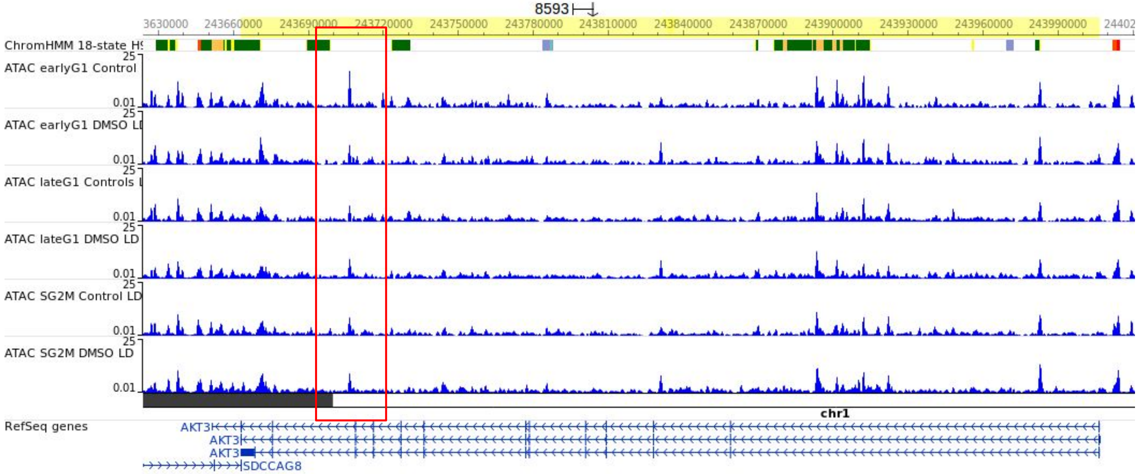


ATAC-seq browser tracks for differentially expressed genes along the PI3K-AKT signaing pathway

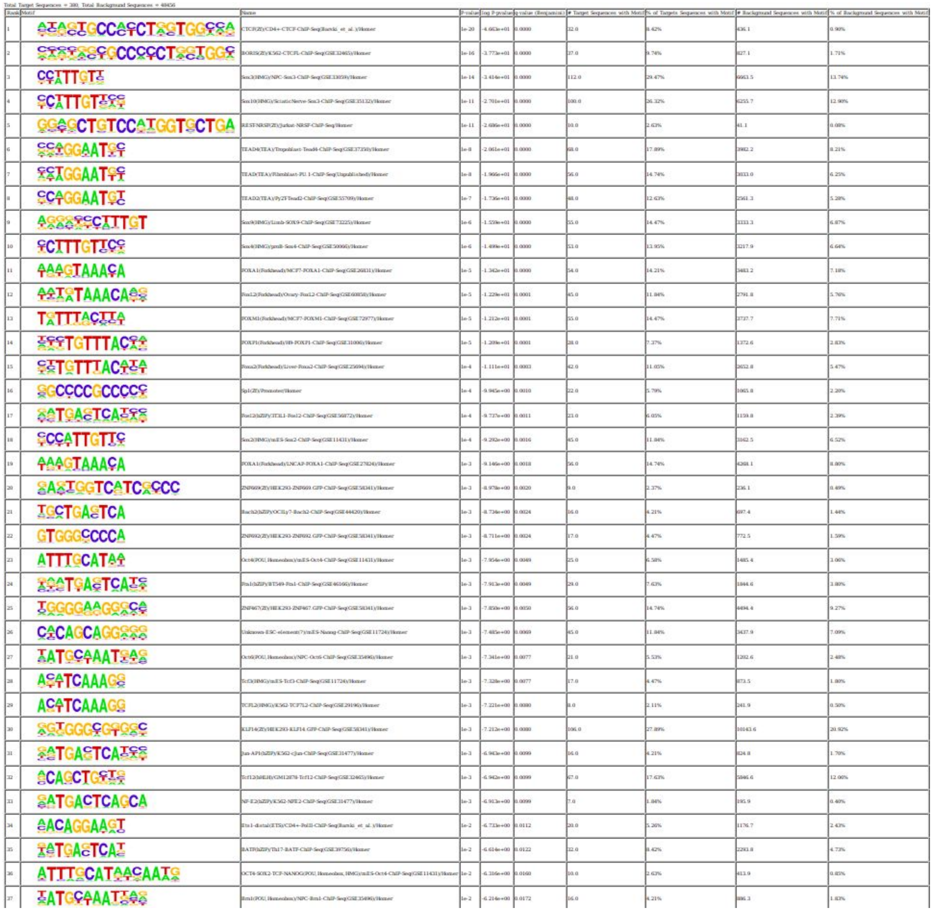


ATAC-seq peaks for PI3K and AKT genes





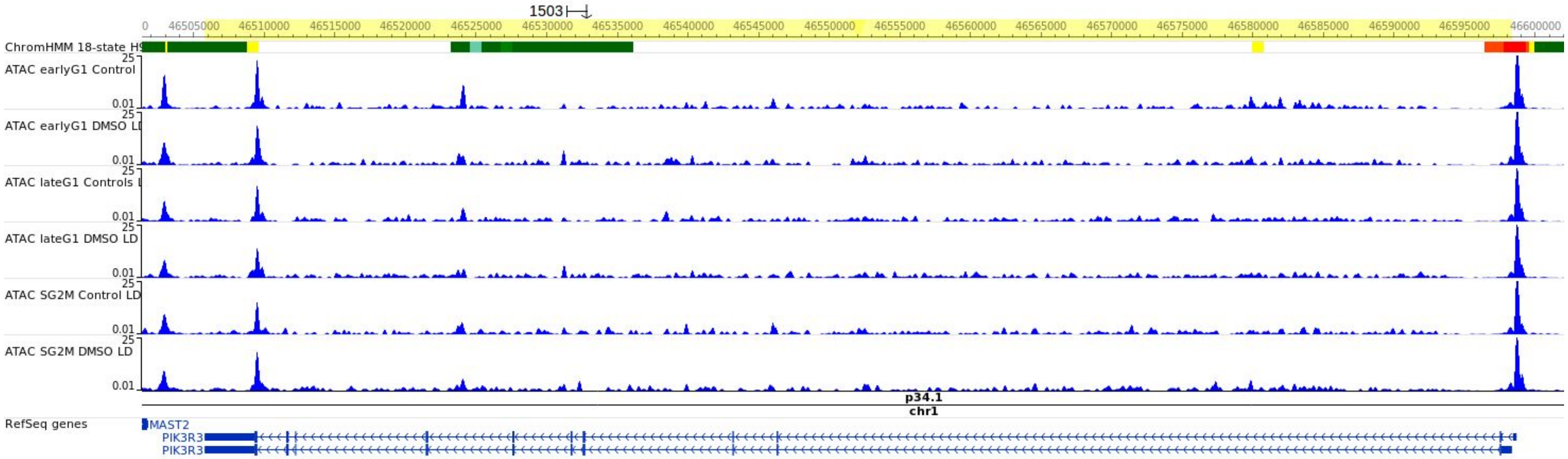
Quiescent ?????? why so many?

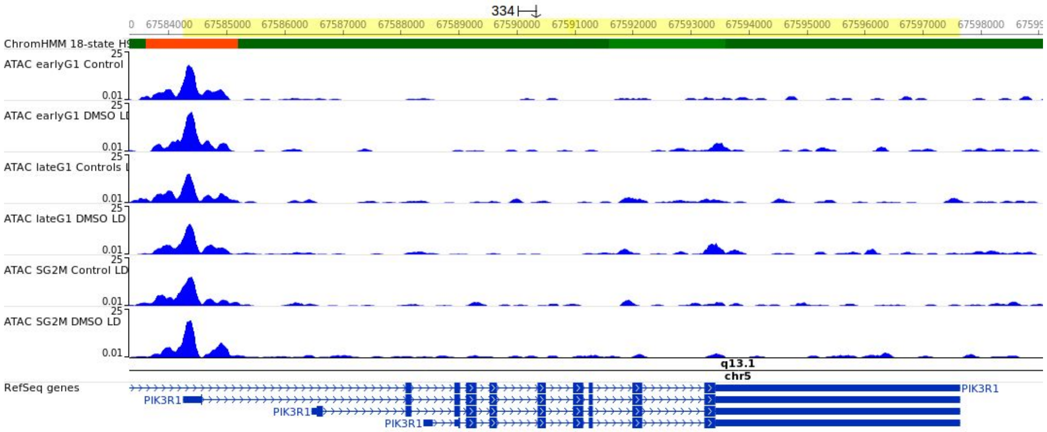


Total Target Sequences = 380, Total Background Sequences = 48456

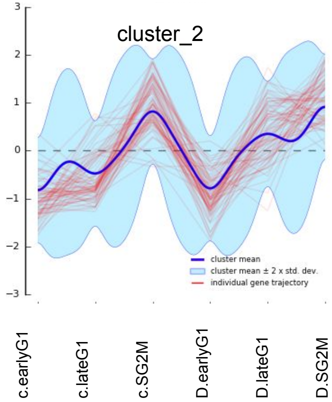
Rank	Motif	Name	P-value	log P-pvalue	q-value (Benjamini)	# Target Sequences with Motif	% of Targets Sequences with Motif	# Background Sequences with Motif	% of Background Sequences with Motif
1		CTCF(Zf)/CD4+-CTCF-ChIP-Seq(Barski_et_al.)/Homer	1e-20	-4.663e+01	0.0000	32.0	8.42%	436.1	0.90%
2		BORIS(Zf)/K562-CTCFL-ChIP-Seq(GSE32465)/Homer	1e-16	-3.773e+01	0.0000	37.0	9.74%	827.1	1.71%
3		Sox3(HMG)/NPC-Sox3-ChIP-Seq(GSE33059)/Homer	1e-14	-3.414e+01	0.0000	112.0	29.47%	6663.5	13.74%
4		Sox10(HMG)/SciaticNerve-Sox3-ChIP-Seq(GSE35132)/Homer	1e-11	-2.701e+01	0.0000	100.0	26.32%	6255.7	12.90%
5		REST-NRSF(Zf)/Jurkat-NRSF-ChIP-Seq/Homer	1e-11	-2.686e+01	0.0000	10.0	2.63%	41.1	0.08%
6		TEAD4(TEA)/Tropoblast-Tead4-ChIP-Seq(GSE37350)/Homer	1e-8	-2.061e+01	0.0000	68.0	17.89%	3982.2	8.21%
7		TEAD(TEA)/Fibroblast-PU.1-ChIP-Seq(Unpublished)/Homer	1e-8	-1.966e+01	0.0000	56.0	14.74%	3033.0	6.25%
8		TEAD2(TEA)/Py2T-Tead2-ChIP-Seq(GSE55709)/Homer	1e-7	-1.736e+01	0.0000	48.0	12.63%	2561.3	5.28%
9		Sox9(HMG)/Limb-SOX9-ChIP-Seq(GSE73225)/Homer	1e-6	-1.559e+01	0.0000	55.0	14.47%	3333.3	6.87%
10		Sox4(HMG)/proB-Sox4-ChIP-Seq(GSE50066)/Homer	1e-6	-1.499e+01	0.0000	53.0	13.95%	3217.9	6.64%
11		FOXA1(Forkhead)/MCF7-FOXA1-ChIP-Seq(GSE26831)/Homer	1e-5	-1.342e+01	0.0000	54.0	14.21%	3483.2	7.18%
12		Foxl2(Forkhead)/Ovary-Foxl2-ChIP-Seq(GSE60858)/Homer	1e-5	-1.229e+01	0.0001	45.0	11.84%	2791.8	5.76%
13		FOXM1(Forkhead)/MCF7-FOXM1-ChIP-Seq(GSE72977)/Homer	1e-5	-1.212e+01	0.0001	55.0	14.47%	3737.7	7.71%
14		FOXP1(Forkhead)/H9-FOXP1-ChIP-Seq(GSE31006)/Homer	1e-5	-1.209e+01	0.0001	28.0	7.37%	1372.6	2.83%
15		Foxa2(Forkhead)/Liver-Foxa2-ChIP-Seq(GSE25694)/Homer	1e-4	-1.111e+01	0.0003	42.0	11.05%	2652.8	5.47%
16		Sp1(Zf)/Promoter/Homer	1e-4	-9.945e+00	0.0010	22.0	5.79%	1065.8	2.20%
17		Fosl2(bZIP)/3T3L1-Fosl2-ChIP-Seq(GSE56872)/Homer	1e-4	-9.737e+00	0.0011	23.0	6.05%	1159.8	2.39%
18		Sox2(HMG)/mES-Sox2-ChIP-Seq(GSE11431)/Homer	1e-4	-9.292e+00	0.0016	45.0	11.84%	3162.5	6.52%
19		FOXA1(Forkhead)/LNCAP-FOXA1-ChIP-Seq(GSE27824)/Homer	1e-3	-9.146e+00	0.0018	56.0	14.74%	4268.1	8.80%
20		ZNF669(Zf)/HEK293-ZNF669.GFP-ChIP-Seq(GSE58341)/Homer	1e-3	-8.978e+00	0.0020	9.0	2.37%	236.1	0.49%
21		Bach2(bZIP)/OCILy7-Bach2-ChIP-Seq(GSE44420)/Homer	1e-3	-8.734e+00	0.0024	16.0	4.21%	697.4	1.44%
22		ZNF692(Zf)/HEK293-ZNF692.GFP-ChIP-Seq(GSE58341)/Homer	1e-3	-8.711e+00	0.0024	17.0	4.47%	772.5	1.59%
23		Oct4(POU,Homeobox)/mES-Oct4-ChIP-Seq(GSE11431)/Homer	1e-3	-7.954e+00	0.0049	25.0	6.58%	1485.4	3.06%
24		Fra1(bZIP)/BT549-Fra1-ChIP-Seq(GSE46166)/Homer	1e-3	-7.913e+00	0.0049	29.0	7.63%	1844.6	3.80%
25		ZNF467(Zf)/HEK293-ZNF467.GFP-ChIP-Seq(GSE58341)/Homer	1e-3	-7.850e+00	0.0050	56.0	14.74%	4494.4	9.27%
26		Unknown-ESC-element(?)/mES-Nanog-ChIP-Seq(GSE11724)/Homer	1e-3	-7.485e+00	0.0069	45.0	11.84%	3437.9	7.09%
27		Oct6(POU,Homeobox)/NPC-Oct6-ChIP-Seq(GSE35496)/Homer	1e-3	-7.341e+00	0.0077	21.0	5.53%	1202.6	2.48%
28		Tcf3(HMG)/mES-Tcf3-ChIP-Seq(GSE11724)/Homer	1e-3	-7.328e+00	0.0077	17.0	4.47%	873.5	1.80%
29		TCFL2(HMG)/K562-TCF7L2-ChIP-Seq(GSE29196)/Homer	1e-3	-7.221e+00	0.0080	8.0	2.11%	241.9	0.50%
30		KLF14(Zf)/HEK293-KLF14.GFP-ChIP-Seq(GSE58341)/Homer	1e-3	-7.212e+00	0.0080	106.0	27.89%	10143.6	20.92%
31		Jun-AP1(bZIP)/K562-cJun-ChIP-Seq(GSE31477)/Homer	1e-3	-6.943e+00	0.0099	16.0	4.21%	824.8	1.70%
32		Tcf12(bHLH)/GM12878-Tcf12-ChIP-Seq(GSE32465)/Homer	1e-3	-6.942e+00	0.0099	67.0	17.63%	5846.6	12.06%
33		NF-E2(bZIP)/K562-NFE2-ChIP-Seq(GSE31477)/Homer	1e-3	-6.913e+00	0.0099	7.0	1.84%	195.9	0.40%
34		Ets1-distal(ETS)/CD4+-PolII-ChIP-Seq(Barski_et_al.)/Homer	1e-2	-6.733e+00	0.0112	20.0	5.26%	1176.7	2.43%
35		BATF(bZIP)/Th17-BATF-ChIP-Seq(GSE39756)/Homer	1e-2	-6.614e+00	0.0122	32.0	8.42%	2293.8	4.73%
36		OCT4-SOX2-TCF-NANOG(POU,Homeobox,HMG)/mES-Oct4-ChIP-Seq(GSE11431)/Homer	1e-2	-6.316e+00	0.0160	10.0	2.63%	413.9	0.85%
37		Brn1(POU,Homeobox)/NPC-Brn1-ChIP-Seq(GSE35496)/Homer	1e-2	-6.214e+00	0.0172	16.0	4.21%	886.3	1.83%

ATAC-seq peaks for PI3K receptors: PI3KR1, PI3KR3

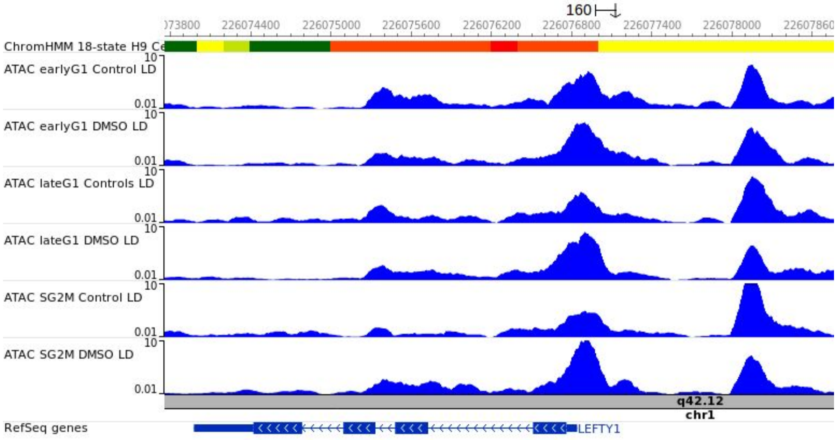


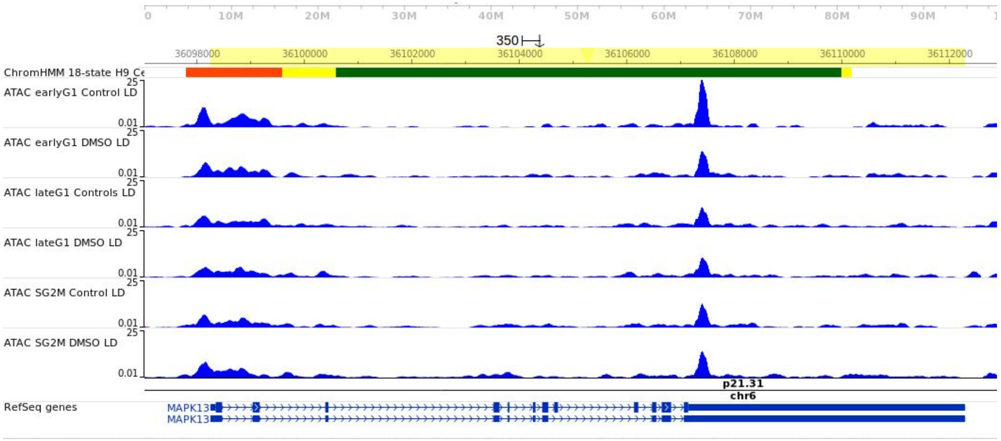


Peak-gene associations within ATAC-seq data clusters: Differential genes with strong peak associations

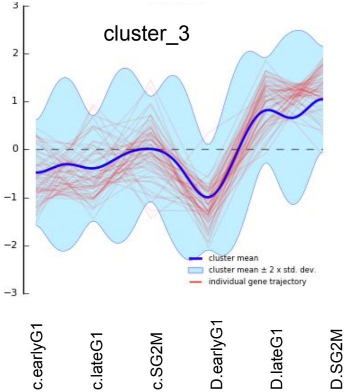


Pathway	**Differentially expressed genes on pathway with peaks associations in cluster**
Cilium assembly	SFI1
Circadian entrainment	RASD1
Gastrin-CREB signaling pathway	NTS
VEGFR signaling	KITLG, PSMB9
Signaling mediated by p38-gamma and p38-delta(N)	MAPK13

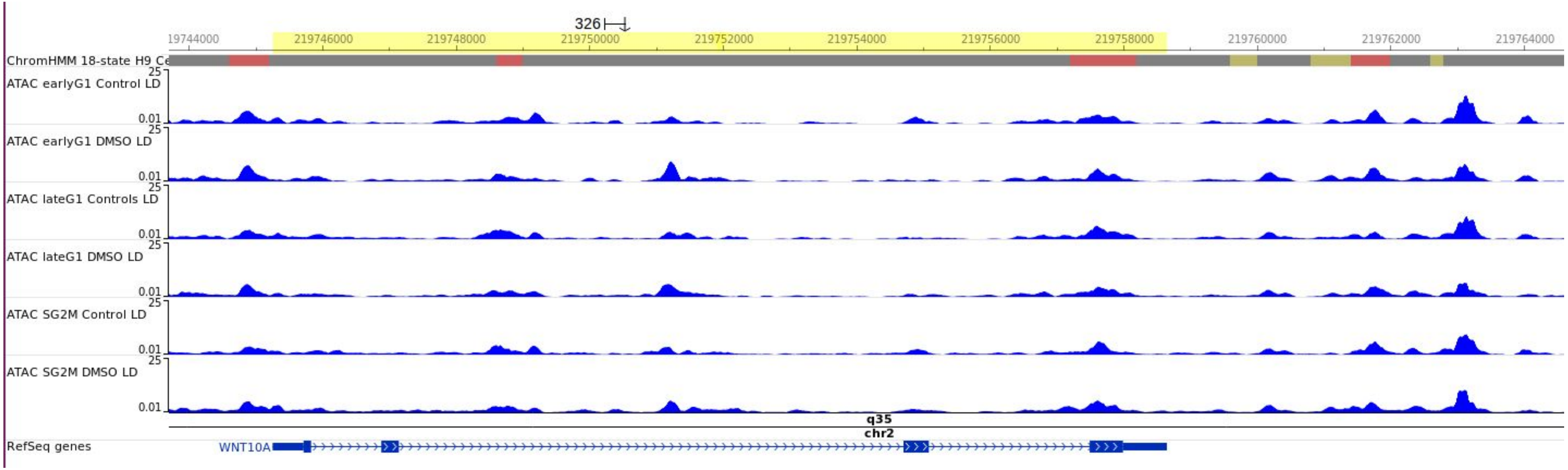




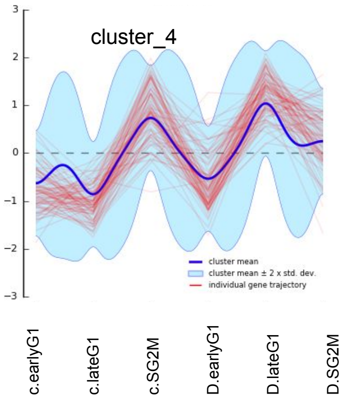
Peak-gene associations within ATAC-seq data clusters: Differential genes with strong peak associations



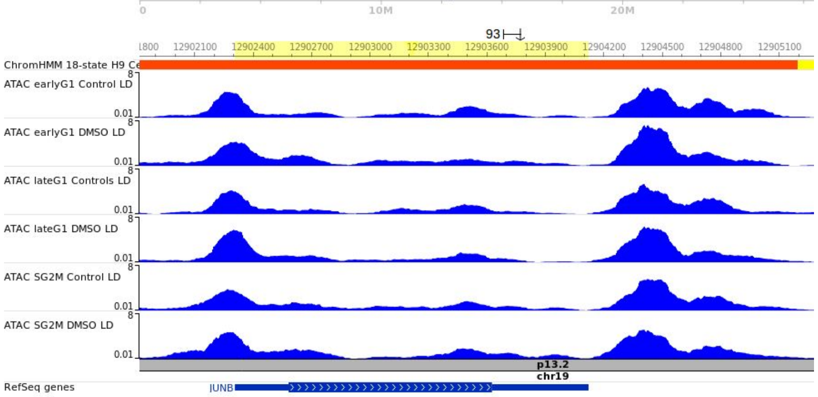
Pathway	**Differentially expressed genes on pathway with peaks associations in cluster**
Wnt signaling pathway	WNT10A,ACTA1
PI3K-AKT signaling pathway	ITGA9
Circadian entrainment	RASD1
Cilium assembly	DYNC2LI1

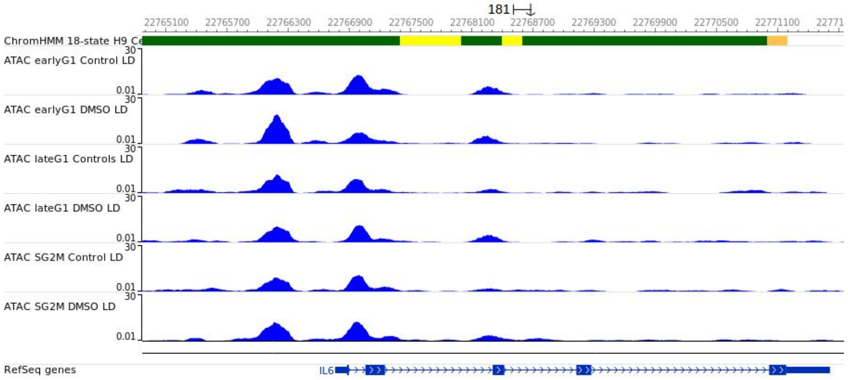


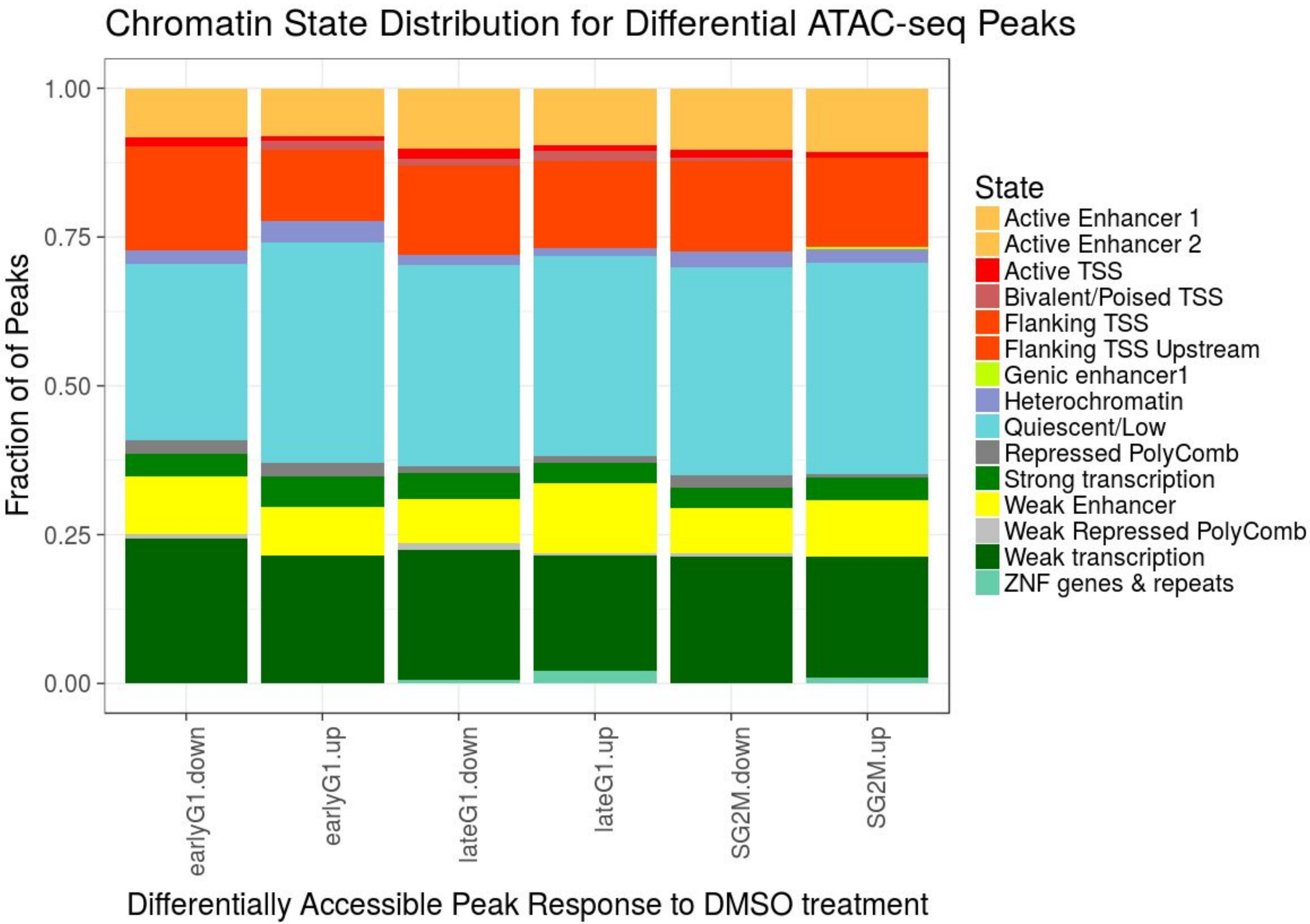
Peak-gene associations within ATAC-seq data clusters: Differential genes with strong peak associations



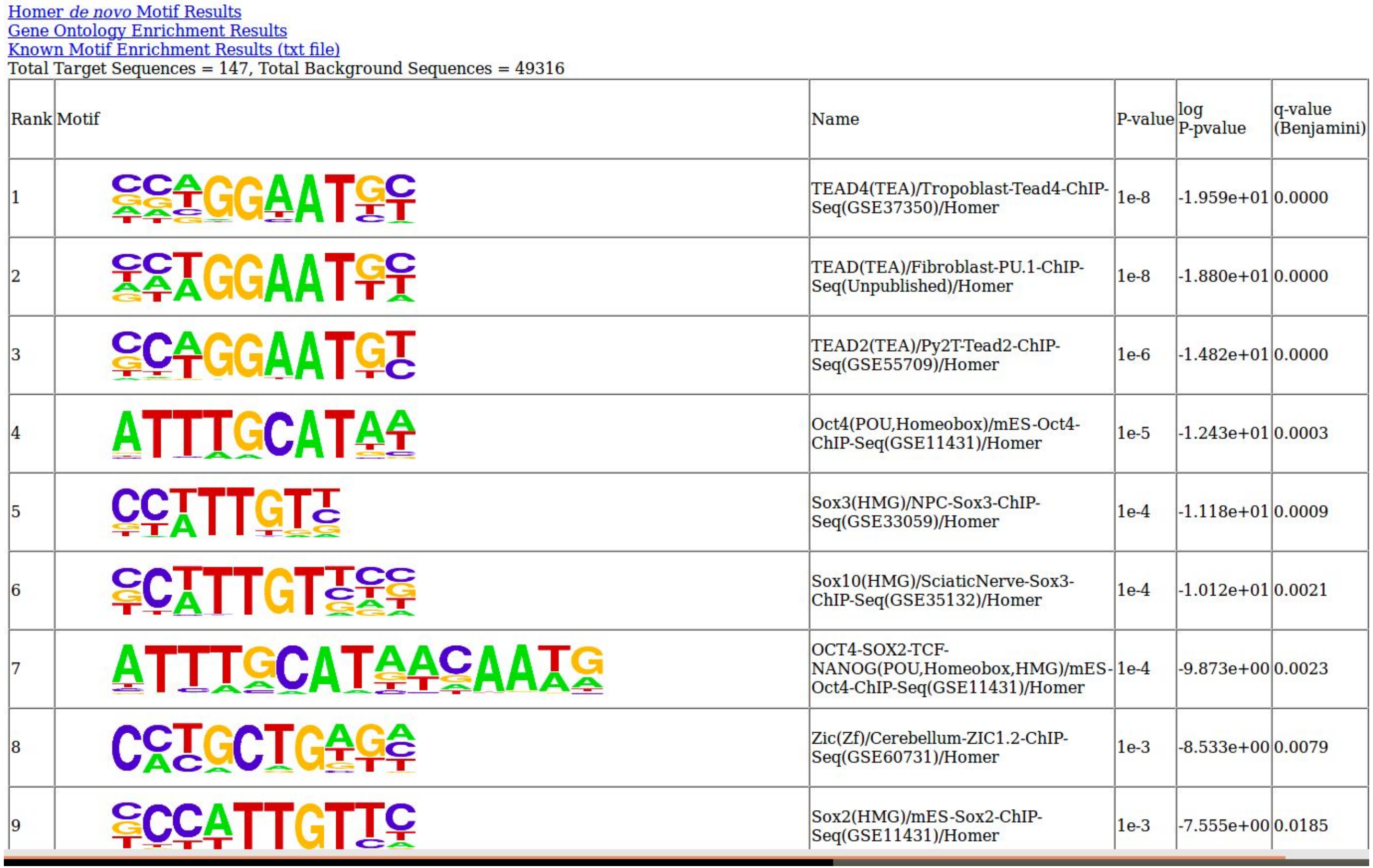
Pathway	Differentially expressed genes on pathway with peaks associations in cluster
Cilium assembly	ODF2
ER to golgi anterograde transport	GOLGA2
GPCR ligand binding	OR2A7, ACKR3, ANXA1, FZD5
Interleukin signaling pathway	IL6
VEGFR signaling pathway	CAV1
TNF signaling pathway	JUNB
WNT signaling pathway	CER1, FZD5, PCDH10, GNA14






Chromatin State Distribution for Differential ATAC-seq Peaks
Fraction of of Peaks
1.00
0.75
0.50
0.25
0.00
earlyG1.down
earlyG1.up
lateG1.down
lateG1.up
SG2M.down
SG2M.up
Differentially Accessible Peak Response to DMSO treatment
State
Active Enhancer 1
Active Enhancer 2
Active TSS
Bivalent/Poised TSS
Flanking TSS
Flanking TSS Upstream
Genic enhancer1
Heterochromatin
Quiescent/Low
Repressed PolyComb
Strong transcription
Weak Enhancer
Weak Repressed PolyComb
Weak transcription
ZNF genes & repeats

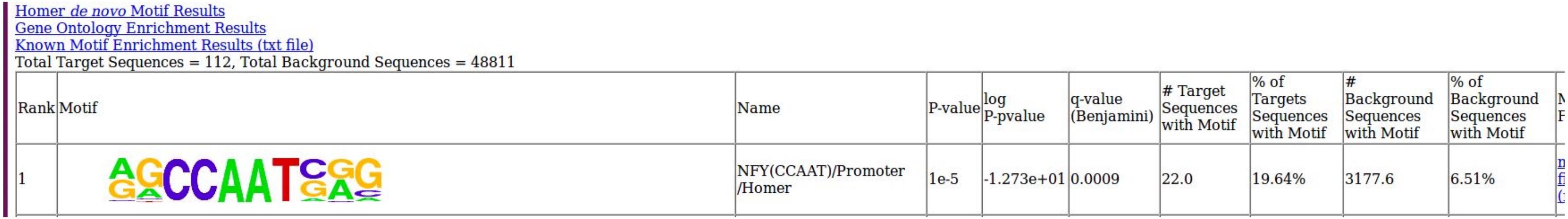
Motifs enriched in enhancer peaks



Homer *de novo* Motif Results
Gene Ontology Enrichment Results
Known Motif Enrichment Results (txt file)
Total Target Sequences = 147, Total Background Sequences = 49316

Rank	Motif	Name	P-value	log P-pvalue	q-value (Benjamini)
1	CCAGGAATGC	TEAD4(TEA)/Tropoblast-Tead4-ChIP-Seq(GSE37350)/Homer	1e-8	-1.959e+01	0.0000
2	CCTGGAATGC	TEAD(TEA)/Fibroblast-PU.1-ChIP-Seq(Unpublished)/Homer	1e-8	-1.880e+01	0.0000
3	CCAGGAATGT	TEAD2(TEA)/Py2T-Tead2-ChIP-Seq(GSE55709)/Homer	1e-6	-1.482e+01	0.0000
4	ATTTGCATAA	Oct4(POU,Homeobox)/mES-Oct4-ChIP-Seq(GSE11431)/Homer	1e-5	-1.243e+01	0.0003
5	CCTTTGTT	Sox3(HMG)/NPC-Sox3-ChIP-Seq(GSE33059)/Homer	1e-4	-1.118e+01	0.0009
6	CCTTTGTTCC	Sox10(HMG)/SciaticNerve-Sox3-ChIP-Seq(GSE35132)/Homer	1e-4	-1.012e+01	0.0021
7	ATTTGCATAACAATG	OCT4-SOX2-TCF-NANOG(POU,Homeobox,HMG)/mES-Oct4-ChIP-Seq(GSE11431)/Homer	1e-4	-9.873e+00	0.0023
8	CCTGCTGAGA	Zic(Zf)/Cerebellum-ZIC1.2-ChIP-Seq(GSE60731)/Homer	1e-3	-8.533e+00	0.0079
9	CCCATTGTTC	Sox2(HMG)/mES-Sox2-ChIP-Seq(GSE11431)/Homer	1e-3	-7.555e+00	0.0185

Motifs enriched in active TSS



Homer *de novo* Motif Results
Gene Ontology Enrichment Results
Known Motif Enrichment Results (txt file)
Total Target Sequences = 112, Total Background Sequences = 48811

Rank	Motif	Name	P-value	log P-pvalue	q-value (Benjamini)	# Target Sequences with Motif	% of Targets Sequences with Motif	# Background Sequences with Motif	% of Background Sequences with Motif	M F
1		NFY(CCAAT)/Promoter /Homer	1e-5	-1.273e+01	0.0009	22.0	19.64%	3177.6	6.51%	m fi (

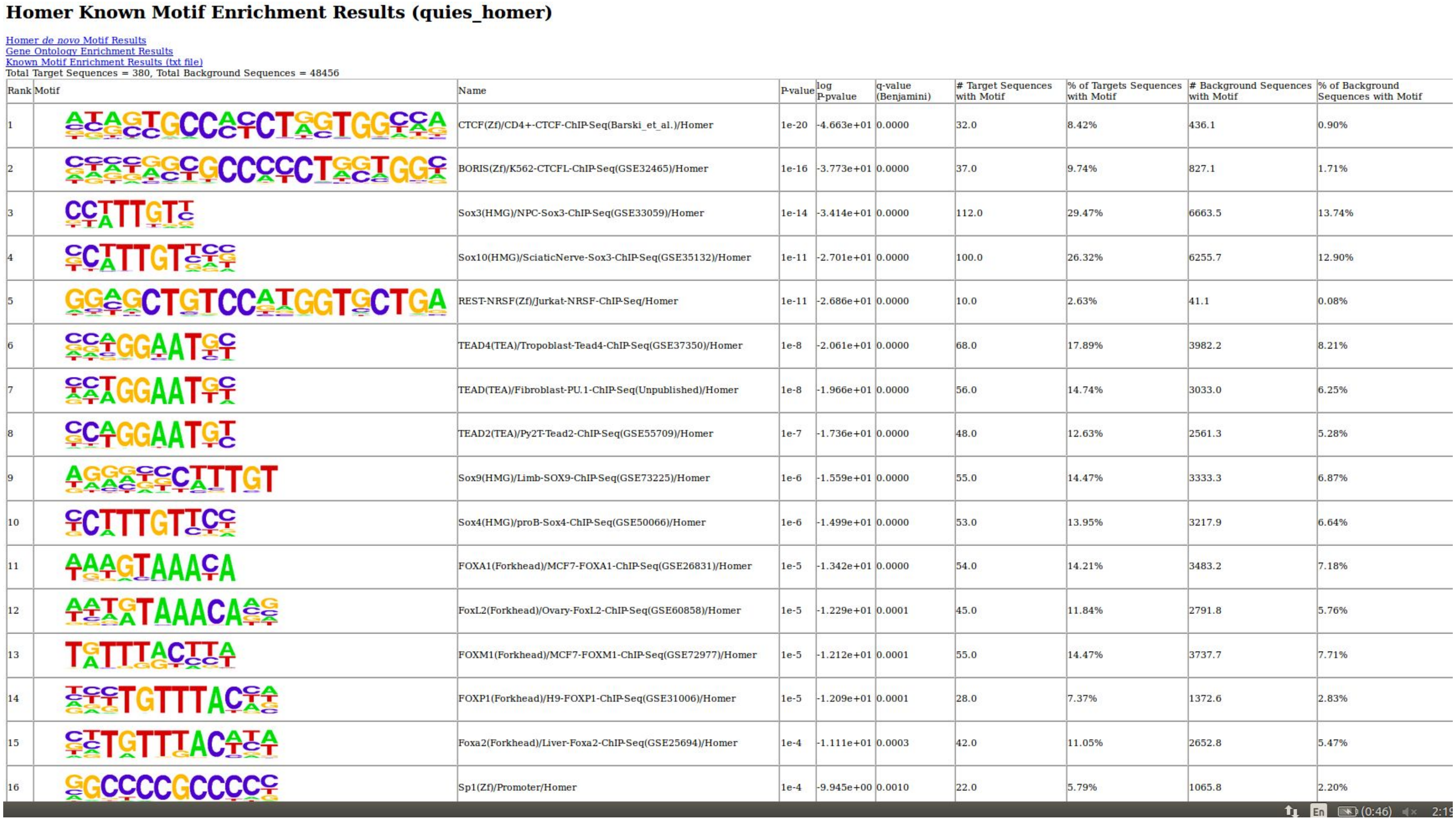
Motifs enriched in Transcription peaks

Total Target Sequences = 168, Total Background Sequences = 49216

Rank	Motif	Name	P-value	log P-pvalue	q-value (Benjamini)	# Target Sequences with Motif	% of Targets Sequences with Motif	# Background Sequences with Motif	% of Background Sequences with Motif
1		TEAD(TEA)/Fibroblast-PU.1-ChIP-Seq(Unpublished)/Homer	1e-7	-1.661e+01	0.0000	30.0	17.86%	2913.6	5.92%
2		BORIS(Zf)/K562-CTCFL-ChIP-Seq(GSE32465)/Homer	1e-5	-1.366e+01	0.0002	15.0	8.93%	951.7	1.93%
3		TEAD4(TEA)/Tropoblast-Tead4-ChIP-Seq(GSE37350)/Homer	1e-5	-1.299e+01	0.0002	33.0	19.64%	4037.9	8.20%
4		Fosl2(bZIP)/3T3L1-Fosl2-ChIP-Seq(GSE56872)/Homer	1e-5	-1.241e+01	0.0003	16.0	9.52%	1195.6	2.43%
5		TEAD2(TEA)/Py2T-Tead2-ChIP-Seq(GSE55709)/Homer	1e-5	-1.159e+01	0.0006	24.0	14.29%	2594.5	5.27%
6		CTCF(Zf)/CD4+-CTCF-ChIP-Seq(Barski_et_al.)/Homer	1e-5	-1.156e+01	0.0006	10.0	5.95%	496.9	1.01%
7		Sox3(HMG)/NPC-Sox3-ChIP-Seq(GSE33059)/Homer	1e-4	-1.140e+01	0.0006	45.0	26.79%	6923.8	14.06%
8		Jun-AP1(bZIP)/K562-cJun-ChIP-Seq(GSE31477)/Homer	1e-4	-1.134e+01	0.0006	13.0	7.74%	877.3	1.78%
9		Bach2(bZIP)/OCILy7-Bach2-ChIP-Seq(GSE44420)/Homer	1e-4	-9.456e+00	0.0028	11.0	6.55%	770.2	1.56%
10		OCT4-SOX2-TCF-NANOG(POU,Homeobox,HMG)/mES-Oct4-ChIP-Seq(GSE11431)/Homer	1e-3	-9.183e+00	0.0033	8.0	4.76%	415.6	0.84%
11		BATF(bZIP)/Th17-BATF-ChIP-	1e-3	-8.993e+00	0.0036	20.0	11.90%	2297.1	4.67%

Chromatin State Distribution QC

-- Why are most peaks in the quiescent state, and why are we seeing so many motif hits in these peaks?



Homer Known Motif Enrichment Results (quies_homer)

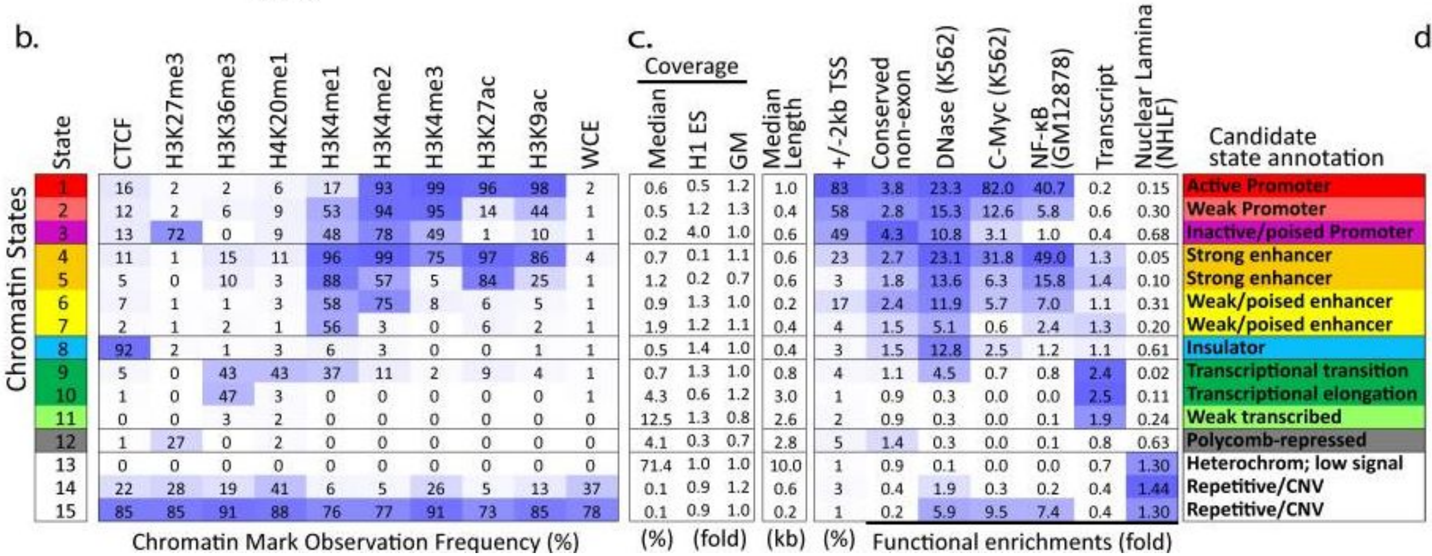
Homer *de novo* Motif Results
Gene Ontology Enrichment Results
Known Motif Enrichment Results (txt file)
Total Target Sequences = 380, Total Background Sequences = 48456

Rank	Motif	Name	P-value	log P-pvalue	q-value (Benjamini)	# Target Sequences with Motif	% of Targets Sequences with Motif	# Background Sequences with Motif	% of Background Sequences with Motif
1		CTCF(Zf)/CD4+-CTCF-ChIP-Seq(Barski_et_al.)/Homer	1e-20	-4.663e+01	0.0000	32.0	8.42%	436.1	0.90%
2		BORIS(Zf)/K562-CTCFL-ChIP-Seq(GSE32465)/Homer	1e-16	-3.773e+01	0.0000	37.0	9.74%	827.1	1.71%
3		Sox3(HMG)/NPC-Sox3-ChIP-Seq(GSE33059)/Homer	1e-14	-3.414e+01	0.0000	112.0	29.47%	6663.5	13.74%
4		Sox10(HMG)/SciaticNerve-Sox3-ChIP-Seq(GSE35132)/Homer	1e-11	-2.701e+01	0.0000	100.0	26.32%	6255.7	12.90%
5		REST-NRSF(Zf)/Jurkat-NRSF-ChIP-Seq/Homer	1e-11	-2.686e+01	0.0000	10.0	2.63%	41.1	0.08%
6		TEAD4(TEA)/Tropoblast-Tead4-ChIP-Seq(GSE37350)/Homer	1e-8	-2.061e+01	0.0000	68.0	17.89%	3982.2	8.21%
7		TEAD(TEA)/Fibroblast-PU.1-ChIP-Seq(Unpublished)/Homer	1e-8	-1.966e+01	0.0000	56.0	14.74%	3033.0	6.25%
8		TEAD2(TEA)/Py2T-Tead2-ChIP-Seq(GSE55709)/Homer	1e-7	-1.736e+01	0.0000	48.0	12.63%	2561.3	5.28%
9		Sox9(HMG)/Limb-SOX9-ChIP-Seq(GSE73225)/Homer	1e-6	-1.559e+01	0.0000	55.0	14.47%	3333.3	6.87%
10		Sox4(HMG)/proB-Sox4-ChIP-Seq(GSE50066)/Homer	1e-6	-1.499e+01	0.0000	53.0	13.95%	3217.9	6.64%
11		FOXA1(Forkhead)/MCF7-FOXA1-ChIP-Seq(GSE26831)/Homer	1e-5	-1.342e+01	0.0000	54.0	14.21%	3483.2	7.18%
12		FoxL2(Forkhead)/Ovary-FoxL2-ChIP-Seq(GSE60858)/Homer	1e-5	-1.229e+01	0.0001	45.0	11.84%	2791.8	5.76%
13		FOXM1(Forkhead)/MCF7-FOXM1-ChIP-Seq(GSE72977)/Homer	1e-5	-1.212e+01	0.0001	55.0	14.47%	3737.7	7.71%
14		FOXP1(Forkhead)/H9-FOXP1-ChIP-Seq(GSE31006)/Homer	1e-5	-1.209e+01	0.0001	28.0	7.37%	1372.6	2.83%
15		Foxa2(Forkhead)/Liver-Foxa2-ChIP-Seq(GSE25694)/Homer	1e-4	-1.111e+01	0.0003	42.0	11.05%	2652.8	5.47%
16		Sp1(Zf)/Promoter/Homer	1e-4	-9.945e+00	0.0010	22.0	5.79%	1065.8	2.20%

10-mark model

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3088773/figure/F1/

Model	Input marks	Cell/Tissue types	Assembly	No. of states	Ref.	URLs
Ernst *et al.*, 2011 *Nature* model	10-marks-observed human data (H3K4me1, H3K4me3, H3K27me3, H3K36me3, H3K27ac, H3K9ac, H4K20me1, H3K4me2, CTCF, and Input)	9 ENCODE cell types	hg18 and hg19 liftover	15	1	https://genome.ucsc.edu/cgi-bin/hgTrackUi?g=wgEncodeBroadHmm&db=hg19 https://www.encodeproject.org

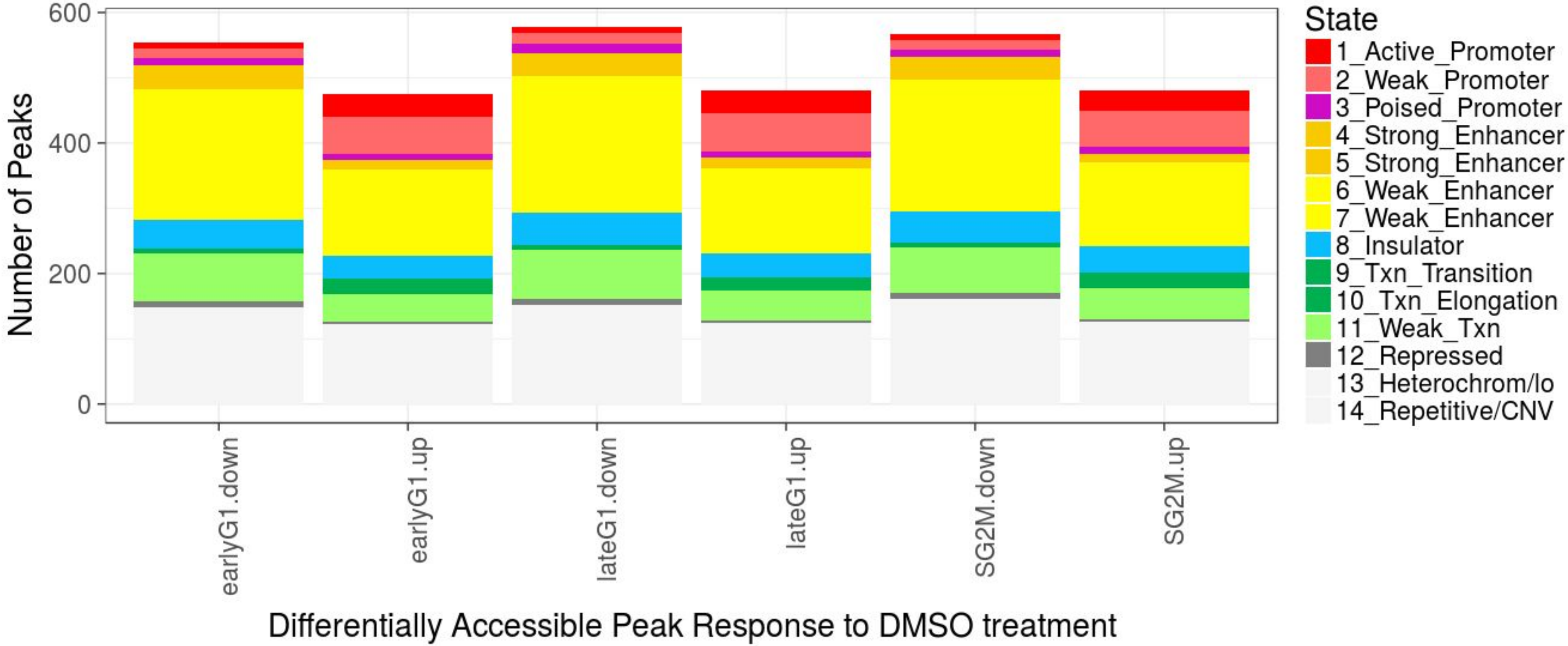


State	CTCF	H3K27me3	H3K36me3	H4K20me1	H3K4me1	H3K4me2	H3K4me3	H3K27ac	H3K9ac	WCE	Coverage Median (%)	Coverage H1 ES (fold)	Coverage GM (fold)	Median Length (kb)	+/-2kb TSS (%)	Conserved non-exon	DNase (K562)	C-Myc (K562)	NF-κB (GM12878)	Transcript	Nuclear Lamina (NHLF)	Candidate state annotation
1	16	2	2	6	17	93	99	96	98	2	0.6	0.5	1.2	1.0	83	3.8	23.3	82.0	40.7	0.2	0.15	Active Promoter
2	12	2	6	9	53	94	95	14	44	1	0.5	1.2	1.3	0.4	58	2.8	15.3	12.6	5.8	0.6	0.30	Weak Promoter
3	13	72	0	9	48	78	49	1	10	1	0.2	4.0	1.0	0.6	49	4.3	10.8	3.1	1.0	0.4	0.68	Inactive/poised Promoter
4	11	1	15	11	96	99	75	97	86	4	0.7	0.1	1.1	0.6	23	2.7	23.1	31.8	49.0	1.3	0.05	Strong enhancer
5	5	0	10	3	88	57	5	84	25	1	1.2	0.2	0.7	0.6	3	1.8	13.6	6.3	15.8	1.4	0.10	Strong enhancer
6	7	1	1	3	58	75	8	6	5	1	0.9	1.3	1.0	0.2	17	2.4	11.9	5.7	7.0	1.1	0.31	Weak/poised enhancer
7	2	1	2	1	56	3	0	6	2	1	1.9	1.2	1.1	0.4	4	1.5	5.1	0.6	2.4	1.3	0.20	Weak/poised enhancer
8	92	2	1	3	6	3	0	0	1	1	0.5	1.4	1.0	0.4	3	1.5	12.8	2.5	1.2	1.1	0.61	Insulator
9	5	0	43	43	37	11	2	9	4	1	0.7	1.3	1.0	0.8	4	1.1	4.5	0.7	0.8	2.4	0.02	Transcriptional transition
10	1	0	47	3	0	0	0	0	0	1	4.3	0.6	1.2	3.0	1	0.9	0.3	0.0	0.0	2.5	0.11	Transcriptional elongation
11	0	0	3	2	0	0	0	0	0	0	12.5	1.3	0.8	2.6	2	0.9	0.3	0.0	0.1	1.9	0.24	Weak transcribed
12	1	27	0	2	0	0	0	0	0	0	4.1	0.3	0.7	2.8	5	1.4	0.3	0.0	0.1	0.8	0.63	Polycomb-repressed
13	0	0	0	0	0	0	0	0	0	0	71.4	1.0	1.0	10.0	1	0.9	0.1	0.0	0.0	0.7	1.30	Heterochrom; low signal
14	22	28	19	41	6	5	26	5	13	37	0.1	0.9	1.2	0.6	3	0.4	1.9	0.3	0.2	0.4	1.44	Repetitive/CNV
15	85	85	91	88	76	77	91	73	85	78	0.1	0.9	1.0	0.2	1	0.2	5.9	9.5	7.4	0.4	1.30	Repetitive/CNV

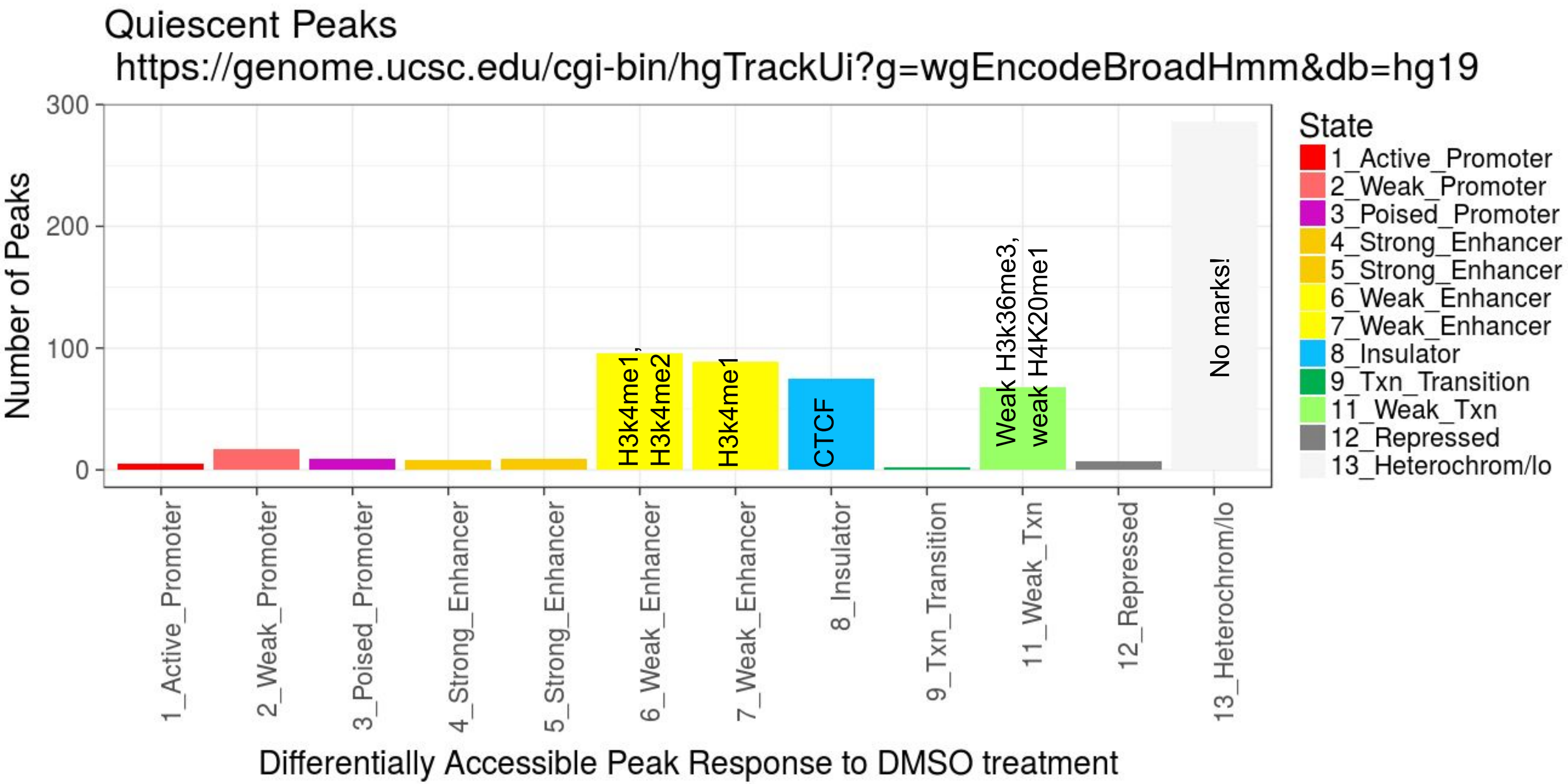
Chromatin Mark Observation Frequency (%)

Functional enrichments (fold)

https://genome.ucsc.edu/cgi-bin/hgTrackUi?g=wgEncodeBroadHmm&db=hg19



Peaks from 15-state Quiescent state, as determined by 10-mark model



Quiescent Peaks

https://genome.ucsc.edu/cgi-bin/hgTrackUi?g=wgEncodeBroadHmm&db=hg19

H3k4me1, H3k4me2

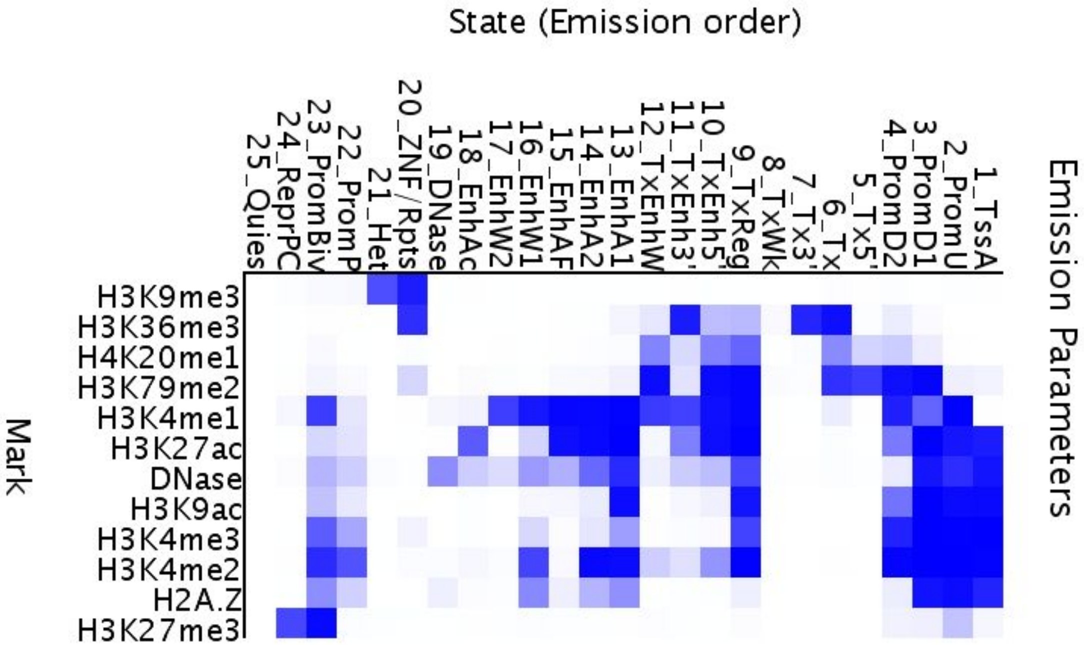
H3k4me1

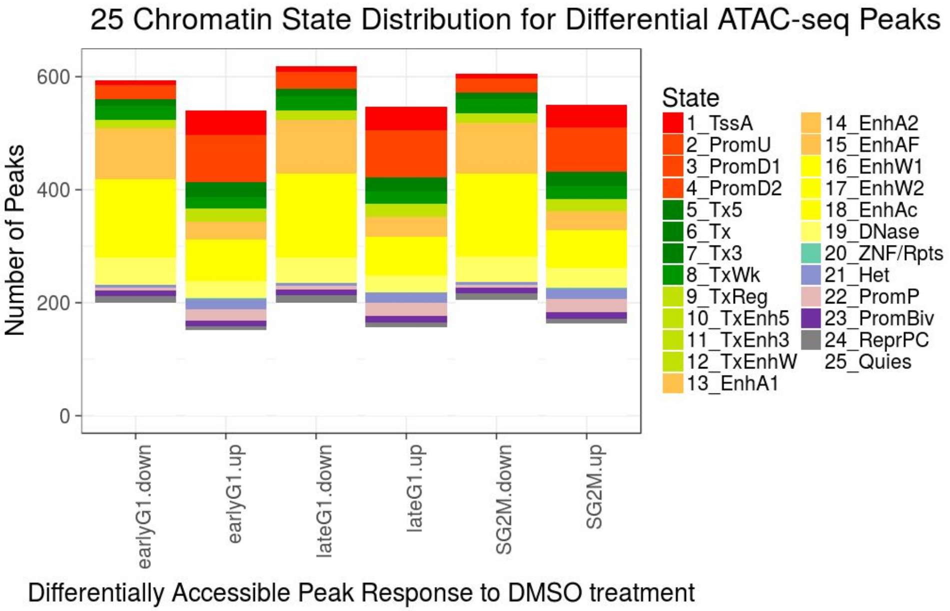
CTCF

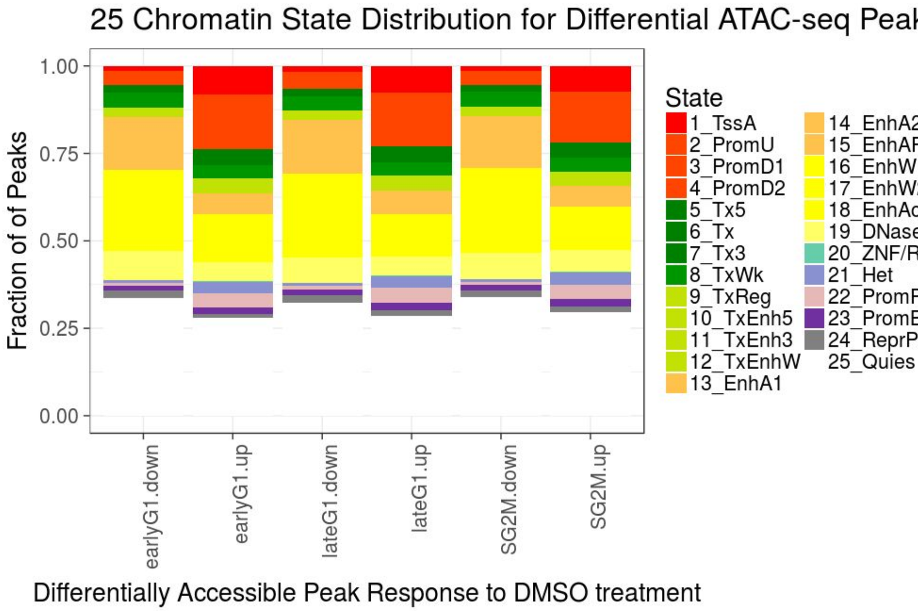
Weak H3k36me3, weak H4K20me1

No marks!

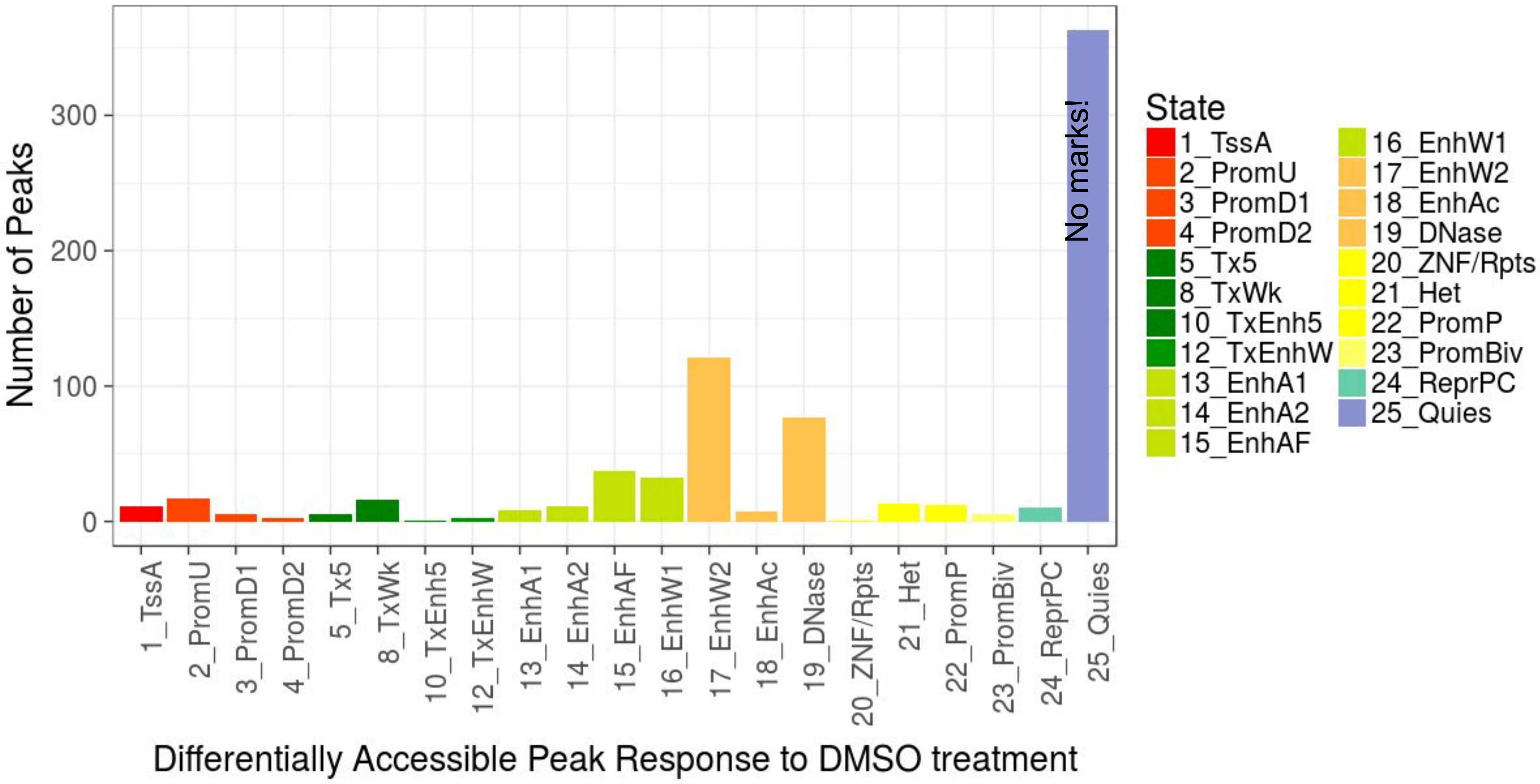
25 state model based on 12 marks



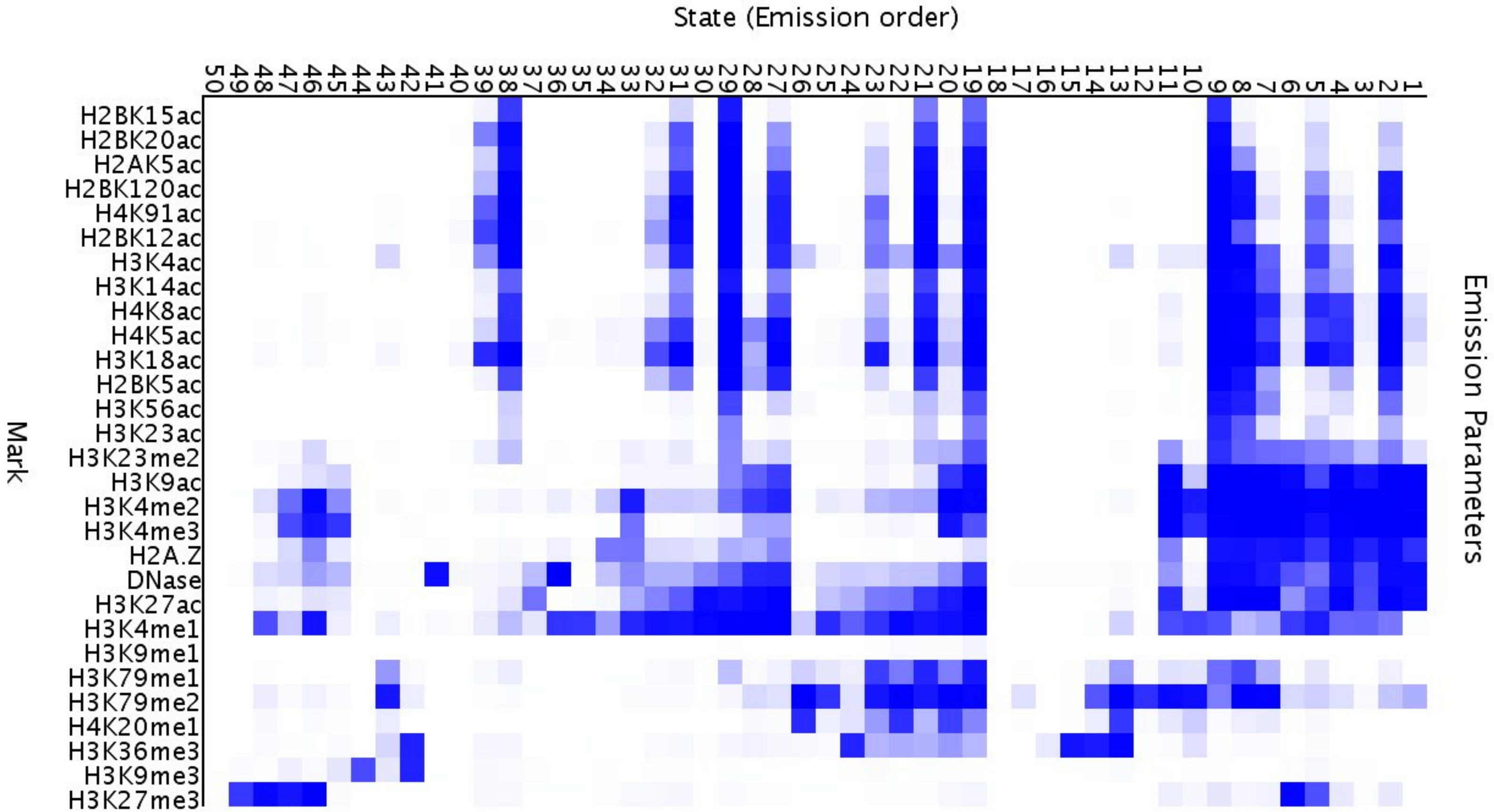

25 Chromatin State Distribution for Differential ATAC-seq Peaks
Number of Peaks
0
200
400
600
earlyG1.down
earlyG1.up
lateG1.down
lateG1.up
SG2M.down
SG2M.up
Differentially Accessible Peak Response to DMSO treatment
State
1_TssA
2_PromU
3_PromD1
4_PromD2
5_Tx5
6_Tx
7_Tx3
8_TxWk
9_TxReg
10_TxEnh5
11_TxEnh3
12_TxEnhW
13_EnhA1
14_EnhA2
15_EnhAF
16_EnhW1
17_EnhW2
18_EnhAc
19_DNase
20_ZNF/Rpts
21_Het
22_PromP
23_PromBiv
24_ReprPC
25_Quies


25 Chromatin State Distribution for Differential ATAC-seq Peak
Fraction of of Peaks
0.00
0.25
0.50
0.75
1.00
earlyG1.down
earlyG1.up
lateG1.down
lateG1.up
SG2M.down
SG2M.up
Differentially Accessible Peak Response to DMSO treatment
State
1_TssA
2_PromU
3_PromD1
4_PromD2
5_Tx5
6_Tx
7_Tx3
8_TxWk
9_TxReg
10_TxEnh5
11_TxEnh3
12_TxEnhW
13_EnhA1
14_EnhA2
15_EnhAF
16_EnhW
17_EnhW
18_EnhAc
19_DNase
20_ZNF/R
21_Het
22_PromP
23_PromB
24_ReprP
25_Quies

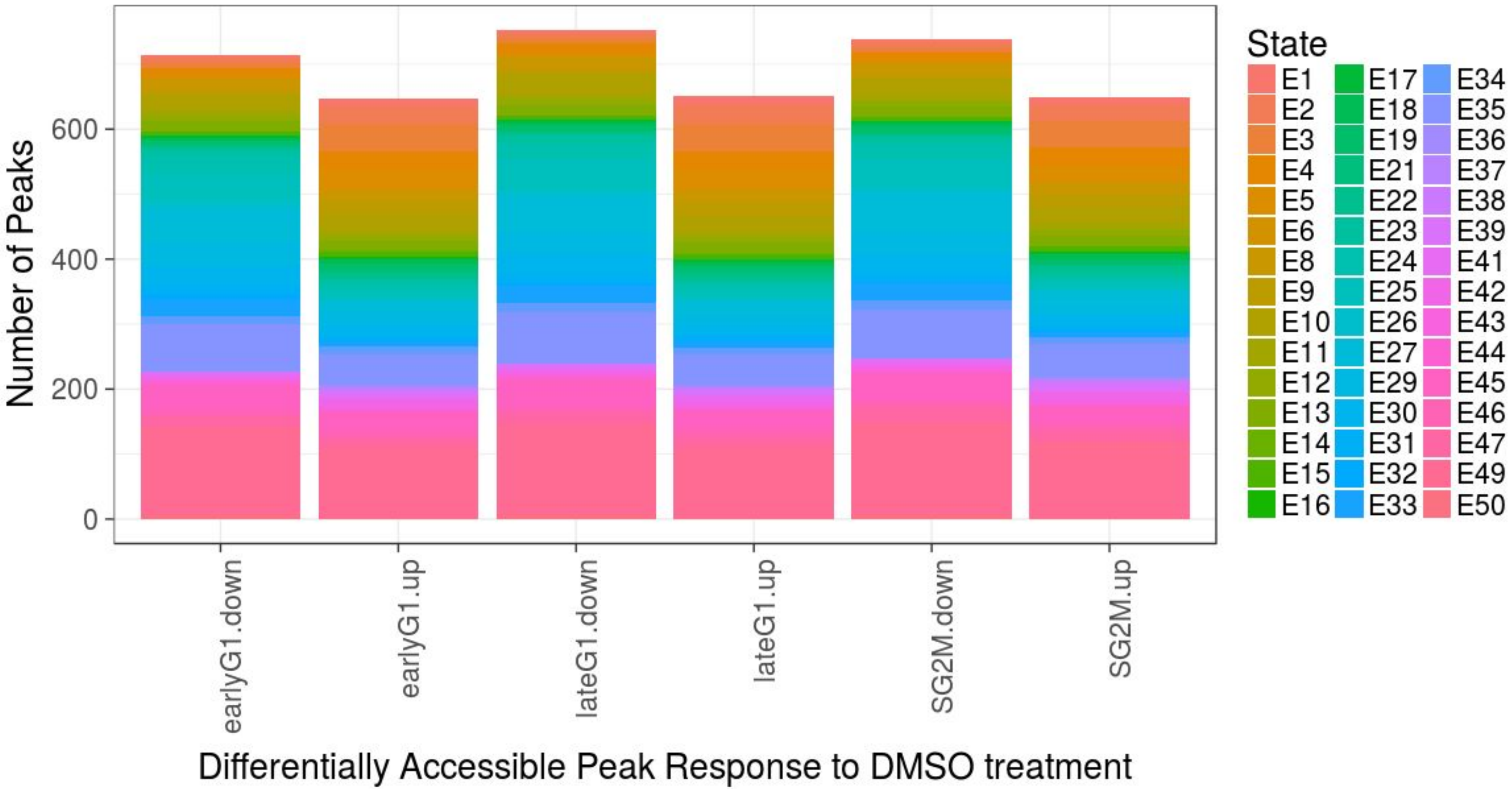
Quies. Peak 25 Chromatin State Distribution for Differential ATAC-seq Peaks

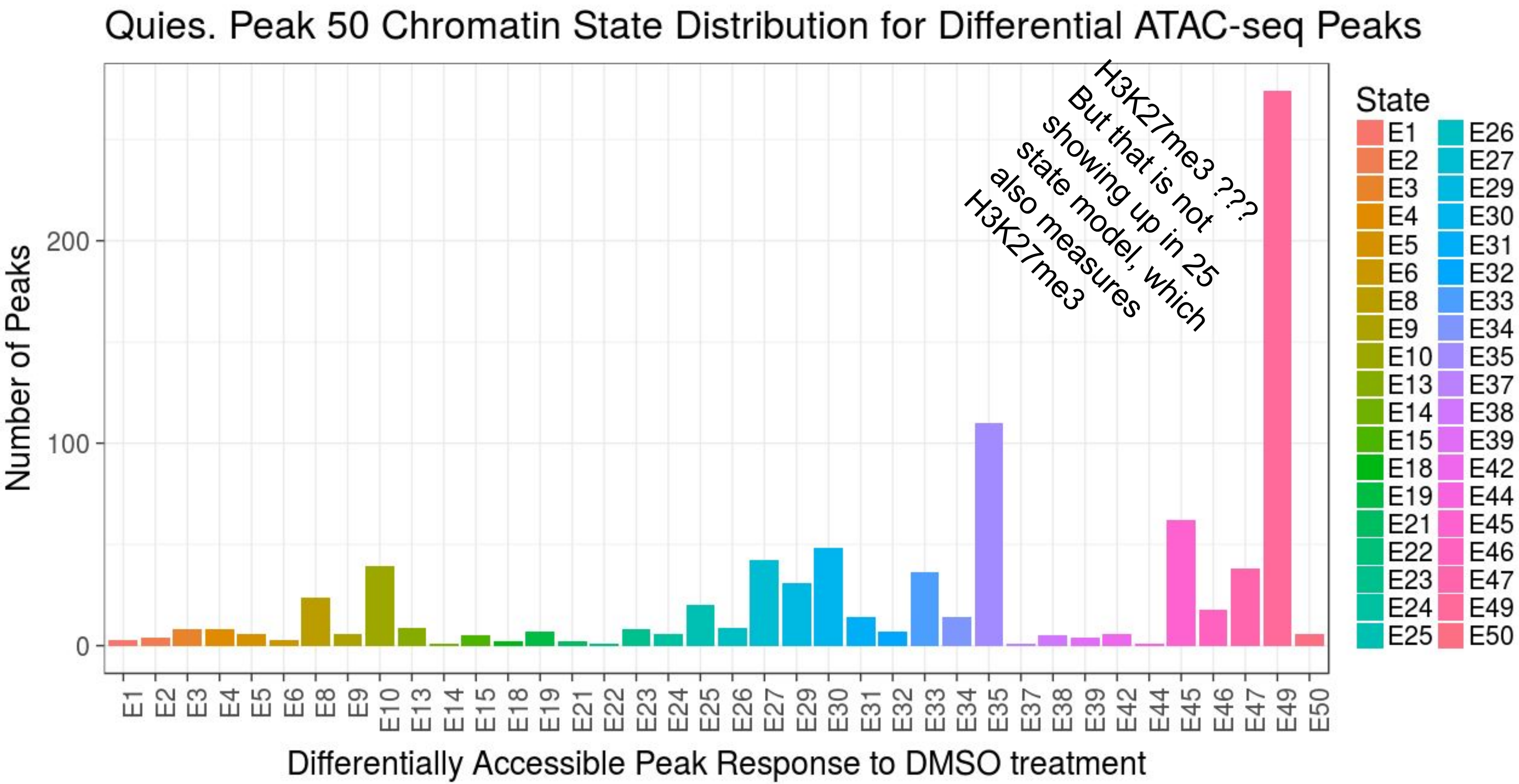


50 state model based on 29 marks



50 Chromatin State Distribution for Differential ATAC-seq Peaks

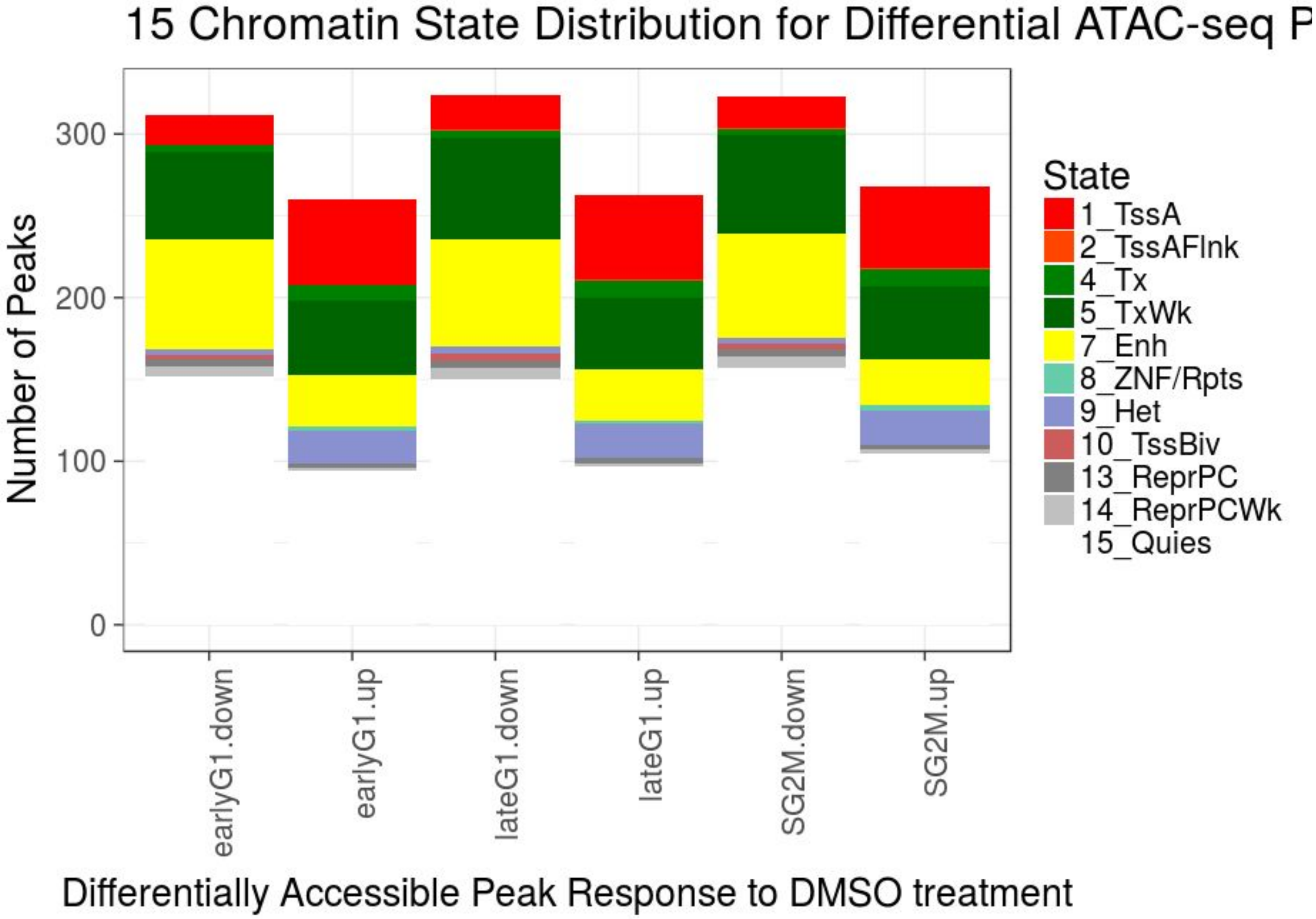


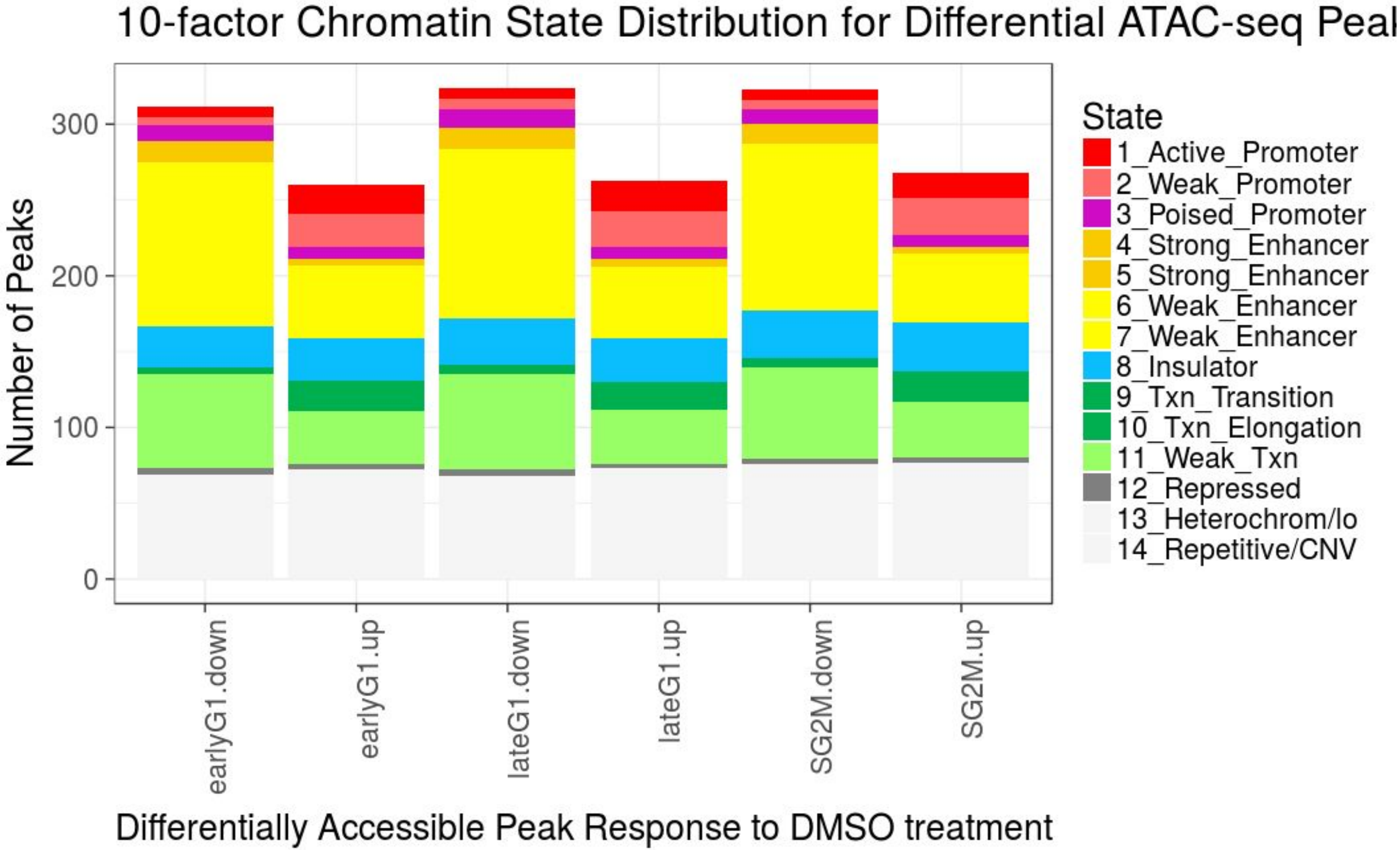


Quies. Peak 50 Chromatin State Distribution for Differential ATAC-seq Peaks

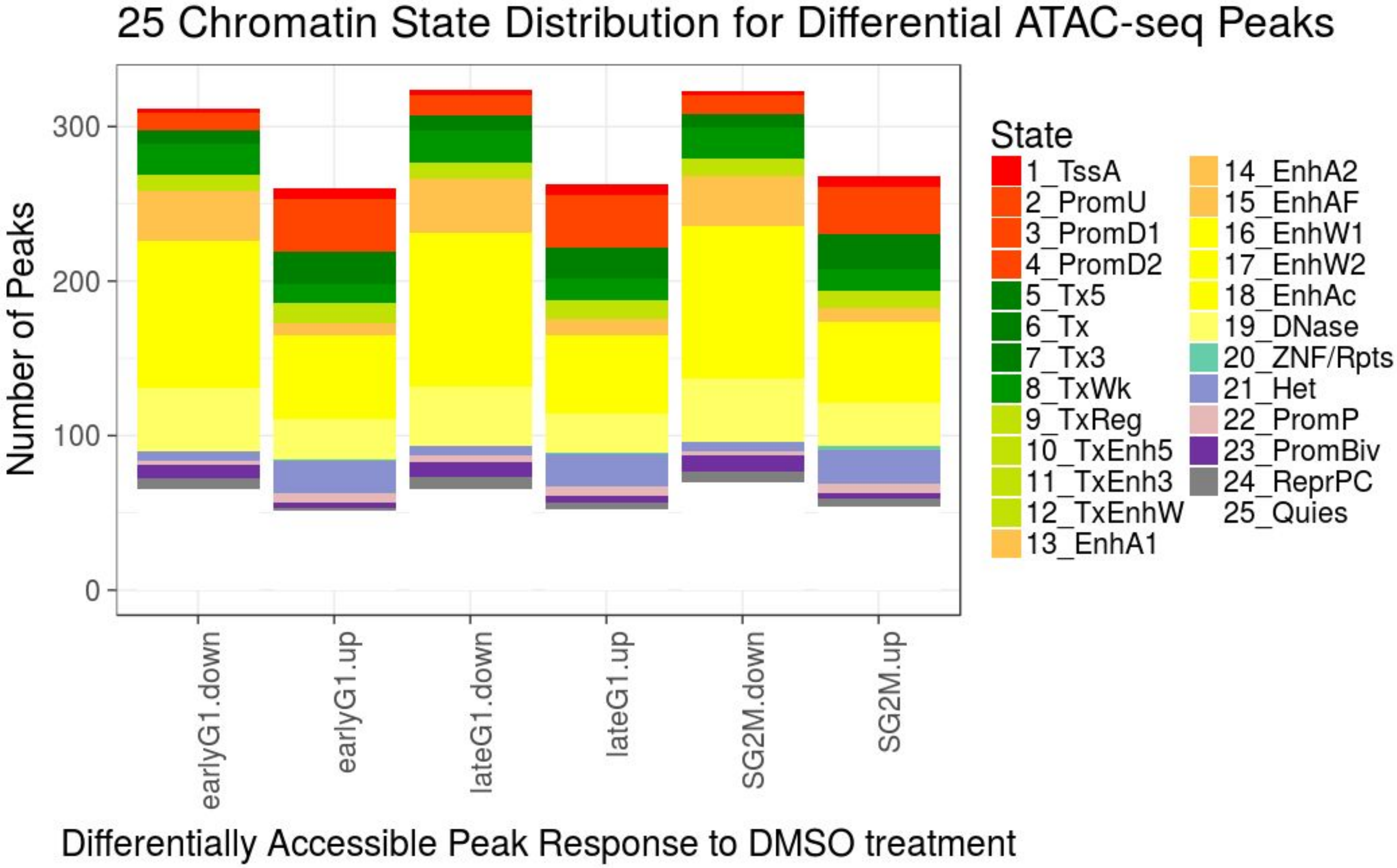
H3K27me3 ???
But that is not
showing up in 25
state model, which
also measures
H3K27me3

Expanding peaks by a flank of 400 bp and assigning peaks that overlap both enhancer & quiescent state to enhancer state


15 Chromatin State Distribution for Differential ATAC-seq P
Number of Peaks
300
200
100
0
earlyG1.down
earlyG1.up
lateG1.down
lateG1.up
SG2M.down
SG2M.up
Differentially Accessible Peak Response to DMSO treatment
State
1_TssA
2_TssAFlnk
4_Tx
5_TxWk
7_Enh
8_ZNF/Rpts
9_Het
10_TssBiv
13_ReprPC
14_ReprPCWk
15_Quies



10-factor Chromatin State Distribution for Differential ATAC-seq Peal



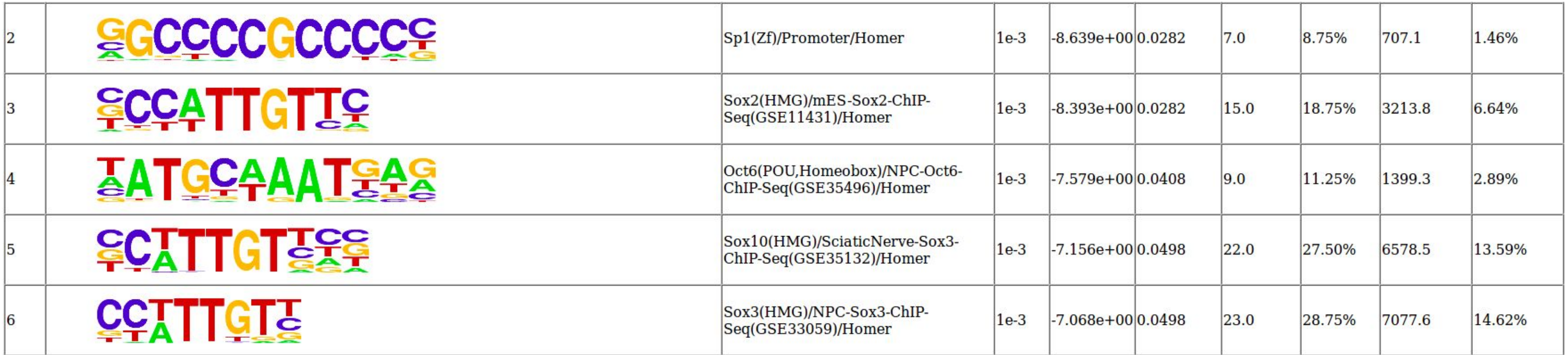
25 Chromatin State Distribution for Differential ATAC-seq Peaks

75 peaks appear completely in quiescent state in 50-state model, 25-state model, 10-factor model, and 15-state model.

- earlyG1 up: 31
- earlyG1 down: 38
- lateG1 up: 31
- lateG1 down: 38
- SG2M up: 34
- SG2M down: 42

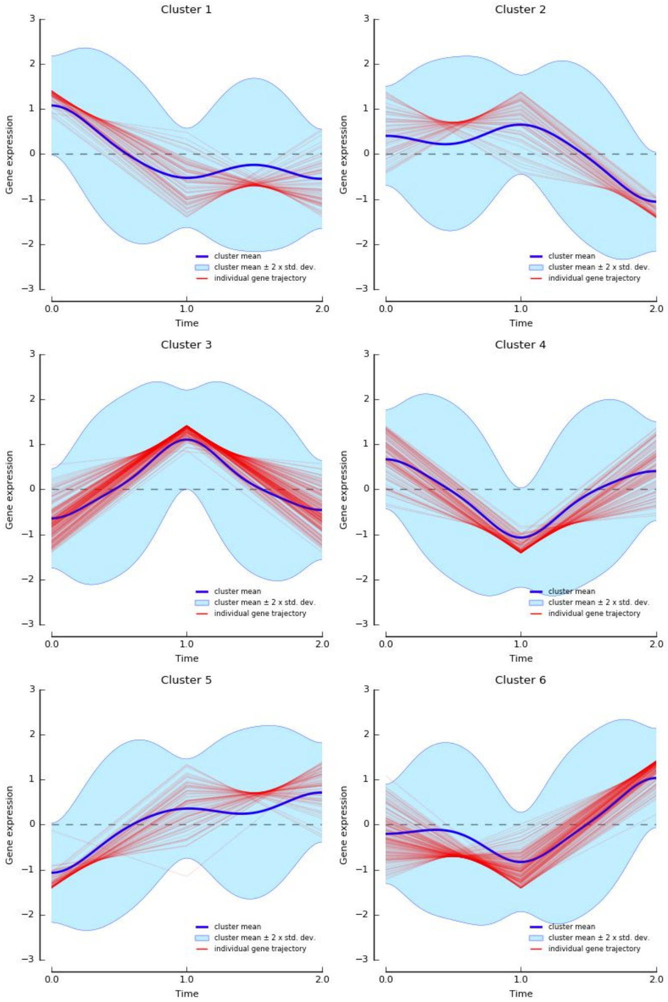
These seem to be associated with a drop in chromatin accessibility with DMSO treatment.
There appears to be a weak increase in H3K27me3 binding with DMSO treatment in these regions.
Homer results for the 75 peaks suggest that most have a Sox motif -- no Chip-seq data is available to verify though.

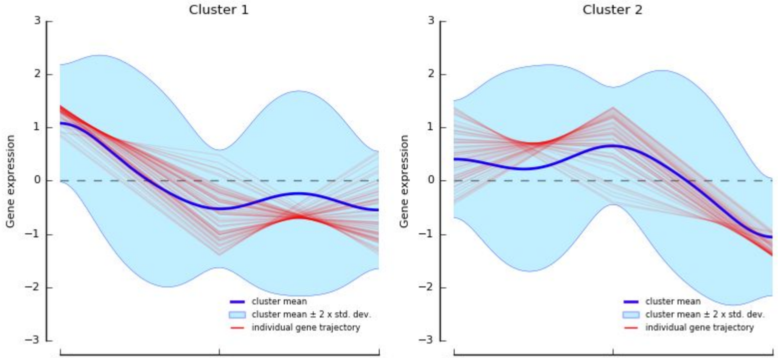
Relevant article about role of Sox2,Sox3,Sox11 in neuron differentiation:
https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3243056/



2		Sp1(Zf)/Promoter/Homer	1e-3	-8.639e+00	0.0282	7.0	8.75%	707.1	1.46%
3		Sox2(HMG)/mES-Sox2-ChIP-Seq(GSE11431)/Homer	1e-3	-8.393e+00	0.0282	15.0	18.75%	3213.8	6.64%
4		Oct6(POU,Homeobox)/NPC-Oct6-ChIP-Seq(GSE35496)/Homer	1e-3	-7.579e+00	0.0408	9.0	11.25%	1399.3	2.89%
5		Sox10(HMG)/SciaticNerve-Sox3-ChIP-Seq(GSE35132)/Homer	1e-3	-7.156e+00	0.0498	22.0	27.50%	6578.5	13.59%
6		Sox3(HMG)/NPC-Sox3-ChIP-Seq(GSE33059)/Homer	1e-3	-7.068e+00	0.0498	23.0	28.75%	7077.6	14.62%

Clustering Version 2


Cluster 1
Gene expression
3
2
1
0
−1
−2
−3
0.0
1.0
2.0
Time
cluster mean
cluster mean ± 2 x std. dev.
individual gene trajectory
Cluster 2
Gene expression
3
2
1
0
−1
−2
−3
0.0
1.0
2.0
Time
cluster mean
cluster mean ± 2 x std. dev.
individual gene trajectory
Cluster 3
Gene expression
3
2
1
0
−1
−2
−3
0.0
1.0
2.0
Time
cluster mean
cluster mean ± 2 x std. dev.
individual gene trajectory
Cluster 4
Gene expression
3
2
1
0
−1
−2
−3
0.0
1.0
2.0
Time
cluster mean
cluster mean ± 2 x std. dev.
individual gene trajectory
Cluster 5
Gene expression
3
2
1
0
−1
−2
−3
0.0
1.0
2.0
Time
cluster mean
cluster mean ± 2 x std. dev.
individual gene trajectory
Cluster 6
Gene expression
3
2
1
0
−1
−2
−3
0.0
1.0
2.0
Time
cluster mean
cluster mean ± 2 x std. dev.
individual gene trajectory



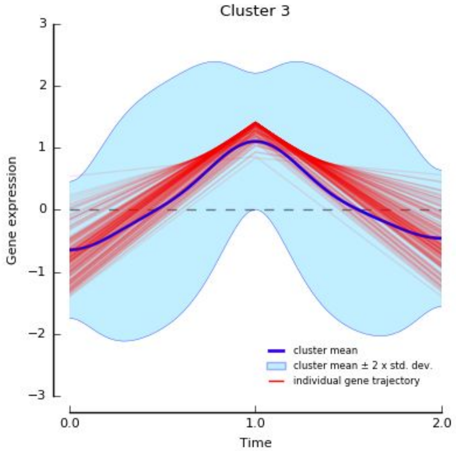
Homer *de novo* Motif Results
Gene Ontology Enrichment Results
Known Motif Enrichment Results (txt file)
Total Target Sequences = 72, Total Background Sequences = 49000

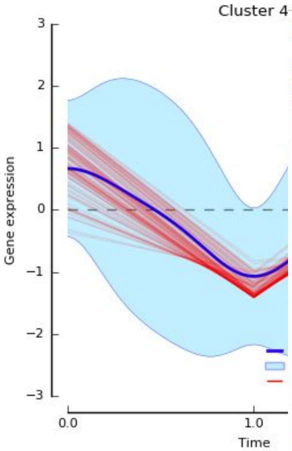
Rank	Motif	Name	P-value	log P-pvalue	q-value (Benjamini)	# Target Sequences with Motif	% of Targets Sequences with Motif	# Background Sequences with Motif	% of Background Sequences with Motif	Motif File	PDF
1		CTCF(Zf)/CD4+-CTCF-ChIP-Seq(Barski_et_al.)/Homer	1e-4	-9.518e+00	0.0235	6.0	8.33%	475.3	0.97%	motif file (matrix)	pdf

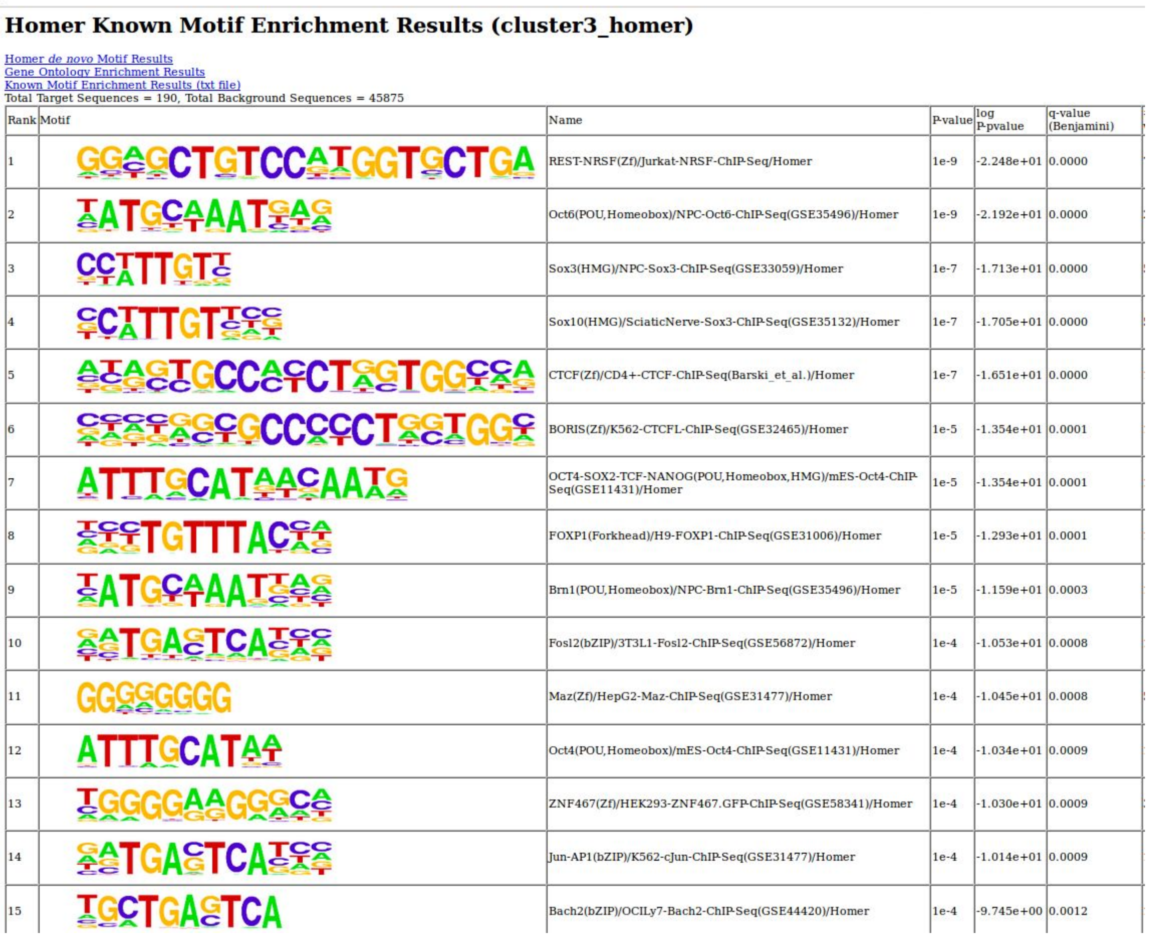
Homer Known Motif Enrichment Results (cluster2_homer)

Homer *de novo* Motif Results
Gene Ontology Enrichment Results
Known Motif Enrichment Results (txt file)
Total Target Sequences = 59, Total Background Sequences = 46000

Rank	Motif	Name	P-value	log P-pvalue	q-value (Benjamini)	# Target Sequences with Motif	% of Targets Sequences with Motif	# Background Sequences with Motif	% of Background Sequences with Motif
1		REST-NRSF(Zf)/Jurkat-NRSF-ChIP-Seq/Homer	1e-7	-1.633e+01	0.0000	4.0	6.78%	32.2	0.07%
2		CTCF(Zf)/CD4+-CTCF-ChIP-Seq(Barski_et_al.)/Homer	1e-3	-8.647e+00	0.0280	5.0	8.47%	429.2	0.88%
3		Sox10(HMG)/SciaticNerve-Sox3-ChIP-Seq(GSE35132)/Homer	1e-3	-7.836e+00	0.0420	18.0	30.51%	6417.1	13.13%
4		Sox4(HMG)/proB-Sox4-ChIP-Seq(GSE50066)/Homer	1e-3	-7.802e+00	0.0420	12.0	20.34%	3225.6	6.60%







Homer Known Motif Enrichment Results (cluster3_homer)

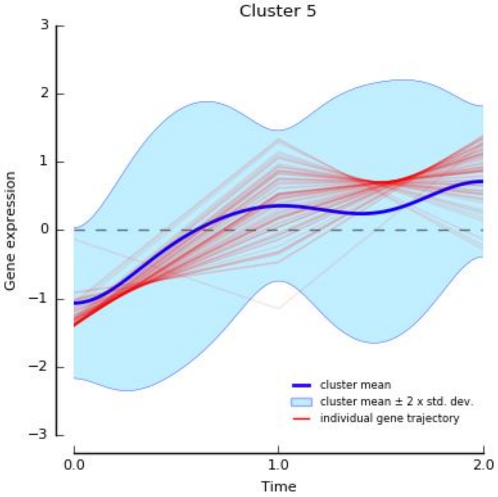
Homer *de novo* Motif Results
Gene Ontology Enrichment Results
Known Motif Enrichment Results (txt file)
Total Target Sequences = 190, Total Background Sequences = 45875

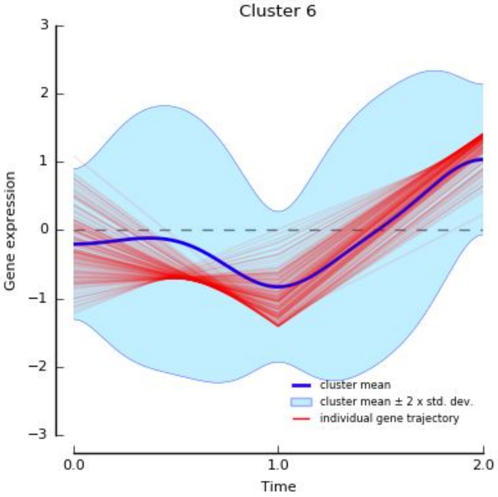
Rank	Motif	Name	P-value	log P-pvalue	q-value (Benjamini)
1		REST-NRSF(Zf)/Jurkat-NRSF-ChIP-Seq/Homer	1e-9	-2.248e+01	0.0000
2		Oct6(POU,Homeobox)/NPC-Oct6-ChIP-Seq(GSE35496)/Homer	1e-9	-2.192e+01	0.0000
3		Sox3(HMG)/NPC-Sox3-ChIP-Seq(GSE33059)/Homer	1e-7	-1.713e+01	0.0000
4		Sox10(HMG)/SciaticNerve-Sox3-ChIP-Seq(GSE35132)/Homer	1e-7	-1.705e+01	0.0000
5		CTCF(Zf)/CD4+-CTCF-ChIP-Seq(Barski_et_al.)/Homer	1e-7	-1.651e+01	0.0000
6		BORIS(Zf)/K562-CTCFL-ChIP-Seq(GSE32465)/Homer	1e-5	-1.354e+01	0.0001
7		OCT4-SOX2-TCF-NANOG(POU,Homeobox,HMG)/mES-Oct4-ChIP-Seq(GSE11431)/Homer	1e-5	-1.354e+01	0.0001
8		FOXP1(Forkhead)/H9-FOXP1-ChIP-Seq(GSE31006)/Homer	1e-5	-1.293e+01	0.0001
9		Brn1(POU,Homeobox)/NPC-Brn1-ChIP-Seq(GSE35496)/Homer	1e-5	-1.159e+01	0.0003
10		Fosl2(bZIP)/3T3L1-Fosl2-ChIP-Seq(GSE56872)/Homer	1e-4	-1.053e+01	0.0008
11		Maz(Zf)/HepG2-Maz-ChIP-Seq(GSE31477)/Homer	1e-4	-1.045e+01	0.0008
12		Oct4(POU,Homeobox)/mES-Oct4-ChIP-Seq(GSE11431)/Homer	1e-4	-1.034e+01	0.0009
13		ZNF467(Zf)/HEK293-ZNF467.GFP-ChIP-Seq(GSE58341)/Homer	1e-4	-1.030e+01	0.0009
14		Jun-AP1(bZIP)/K562-cJun-ChIP-Seq(GSE31477)/Homer	1e-4	-1.014e+01	0.0009
15		Bach2(bZIP)/OCILy7-Bach2-ChIP-Seq(GSE44420)/Homer	1e-4	-9.745e+00	0.0012

Homer Known Motif Enrichment Results (cluster4_homer)

Homer *de novo* Motif Results
Gene Ontology Enrichment Results
Known Motif Enrichment Results (txt file)
Total Target Sequences = 107, Total Background Sequences = 48432

Rank	Motif	Name	P-value	log P-pvalue	q-value (Benjamini)	# Target Sequences with Motif	% of Targets Sequences with Motif	# Background Sequences with Motif	% of Background Sequences with Motif
1		CTCF(Zf)/CD4+-CTCF-ChIP-Seq(Barski_et_al.)/Homer	1e-7	-1.621e+01	0.0000	10.0	9.35%	460.1	0.95%
2		TEAD(TEA)/Fibroblast-PU.1-ChIP-Seq(Unpublished)/Homer	1e-6	-1.468e+01	0.0001	23.0	21.50%	3190.3	6.59%
3		TEAD4(TEA)/Tropoblast-Tead4-ChIP-Seq(GSE37350)/Homer	1e-5	-1.346e+01	0.0002	26.0	24.30%	4245.3	8.77%
4		BORIS(Zf)/K562-CTCFL-ChIP-Seq(GSE32465)/Homer	1e-4	-9.417e+00	0.0065	9.0	8.41%	806.4	1.67%
5		TEAD2(TEA)/Py2T-Tead2-ChIP-Seq(GSE55709)/Homer	1e-4	-9.383e+00	0.0065	17.0	15.89%	2682.0	5.54%


Cluster 5


Cluster 6

Homer *de novo* Motif Results
Gene Ontology Enrichment Results
Known Motif Enrichment Results (txt file)
Total Target Sequences = 74, Total Background Sequences = 48718

Rank	Motif	Name	P-value	log P-pvalue	q-value (Benjamini)	# Target Sequences with Motif
1		CTCF(Zf)/CD4+-CTCF-ChIP-Seq(Barski_et_al.)/Homer	1e-9	-2.117e+01	0.0000	11.0
2		BORIS(Zf)/K562-CTCFL-ChIP-Seq(GSE32465)/Homer	1e-6	-1.570e+01	0.0000	13.0
3		GABPA(ETS)/Jurkat-GABPa-ChIP-Seq(GSE17954)/Homer	1e-3	-6.928e+00	0.1042	16.0

Homer *de novo* Motif Results
Gene Ontology Enrichment Results
Known Motif Enrichment Results (txt file)
Total Target Sequences = 158, Total Background Sequences = 48975

Rank	Motif	Name	P-value	log P-pvalue	q-value (Benjamini)	# Target Sequences with Motif	% of Targets Sequences with Motif
1		TEAD4(TEA)/Tropoblast-Tead4-ChIP-Seq(GSE37350)/Homer	1e-10	-2.532e+01	0.0000	43.0	27.22%
2		TEAD(TEA)/Fibroblast-PU.1-ChIP-Seq(Unpublished)/Homer	1e-9	-2.174e+01	0.0000	35.0	22.15%
3		TEAD2(TEA)/Py2T-Tead2-ChIP-Seq(GSE55709)/Homer	1e-7	-1.764e+01	0.0000	29.0	18.35%
4		Sox2(HMG)/mES-Sox2-ChIP-Seq(GSE11431)/Homer	1e-7	-1.652e+01	0.0000	32.0	20.25%
5		Sox3(HMG)/NPC-Sox3-ChIP-Seq(GSE33059)/Homer	1e-6	-1.584e+01	0.0000	49.0	31.01%
6		Sox4(HMG)/proB-Sox4-ChIP-Seq(GSE50066)/Homer	1e-4	-1.000e+01	0.0024	26.0	16.46%
7		Sox10(HMG)/SciaticNerve-Sox3-ChIP-Seq(GSE35132)/Homer	1e-4	-9.763e+00	0.0026	40.0	25.32%
8		Sox6(HMG)/Myotubes-Sox6-ChIP-Seq(GSE32627)/Homer	1e-3	-9.001e+00	0.0049	38.0	24.05%
9		Ets1-distal(ETS)/CD4+-PolII-ChIP-Seq(Barski_et_al.)/Homer	1e-3	-8.330e+00	0.0085	13.0	8.23%
10		Sp1(Zf)/Promoter/Homer	1e-2	-6.907e+00	0.0319	12.0	7.59%
11		Sox15(HMG)/CPA-Sox15-ChIP-Seq(GSE62909)/Homer	1e-2	-6.522e+00	0.0426	28.0	17.72%
12		VDR(NR),DR3/GM10855-VDR+vitD-ChIP-Seq(GSE22484)/Homer	1e-2	-6.473e+00	0.0426	10.0	6.33%

108 differential ZNF genes