Heatmaps and average profiles

**Fig. 2d and Supplementary Fig. 5b**

extract 2 groups:

promoter

non promoter: intragenic (exon, intron, 5UTR, 3UTR) + intergenic

4 groups of ATAC CpGs based on methylation and coverage (cov <5 reads removed)

make one group of each for each group of CpGs

file1.txt

#1-meth\_00to20-cov\_50andabove

file2.txt

#2-meth\_00to20-cov\_5to50

file3.xt

#3-meth\_20to80

file4.txt

#4-meth\_80to100

/ifs/home/lhoump01/Method/deeptools\_heatmaps/genomic-annotations

## BW files

BW\_1="/ifs/home/lhoump01/Method/deeptools\_heatmaps/ATAC-merged.deduplicated.sort.Q30.bw"

BW\_2="/ifs/home/lhoump01/mESC\_ChIP-BS/Public/K4me1\_GSM1000121\_q30\_rmdup\_sorted.bw"

BW\_2="/ifs/home/lhoump01/mESC\_ChIP-BS/Public/K27ac\_GSM1000126\_q30\_rmdup\_sorted.bw"

BW\_2="/ifs/home/lhoump01/mESC\_ChIP-BS/Public/K4me3\_GSM1000124\_q30\_rmdup\_sorted.bw"

################## for promoter ################

cat 1-promoter.bed file1.txt 2-promoter.bed file2.txt 3-promoter.bed file3.txt 4-promoter.bed file4.txt > 4groups\_Promoter.bed

qlogin -q all.q

module unload python

module unload deeptools

module load deeptools/2.3.3

BED="/ifs/home/lhoump01/Method/deeptools\_heatmaps/genomic-annotations/4groups\_Promoter.bed"

################### Try using sort by ATAC signal

#then run in one bloc:

group\_plot(){

computeMatrix reference-point --referencePoint center --scoreFileName ${BW\_1} --regionsFileName ${BED} --outFileName ${OUTPUT\_FILE\_PREFIX}\_matrix.gz --sortRegions **descend --sortUsingSamples 1** --missingDataAsZero --beforeRegionStartLength ${Midpoint} --afterRegionStartLength ${Midpoint}

plotHeatmap --matrixFile ${OUTPUT\_FILE\_PREFIX}\_matrix.gz --outFileName ${OUTPUT\_FILE\_PREFIX}\_heatmap.pdf --colorMap ${color} --sortRegions **descend** **--sortUsingSamples 1** --heatmapHeight 50 --legendLocation none --whatToShow 'heatmap and colorbar' --zMin ${zMin} --zMax ${zMax}

plotProfile --matrixFile ${OUTPUT\_FILE\_PREFIX}\_matrix.gz --outFileName ${OUTPUT\_FILE\_PREFIX}\_profilePlot.pdf --colors red blue green purple --legendLocation 'upper-center' --yMin ${y\_scale\_min} --yMax ${y\_scale\_max}

plotProfile --matrixFile ${OUTPUT\_FILE\_PREFIX}\_matrix.gz --outFileName ${OUTPUT\_FILE\_PREFIX}\_pergroupprofilePlot.pdf --colors red --legendLocation 'upper-center' --yMin ${y\_scale\_min} --yMax ${y\_scale\_max} --perGroup

}

#then run in one bloc for ATAC:

RUNDIR=/ifs/home/lhoump01/Method/deeptools\_heatmaps

Midpoint=1000

y\_scale\_min=0

y\_scale\_max=150

color=Oranges

zMin=0

zMax=75

BW\_1="/ifs/home/lhoump01/Method/deeptools\_heatmaps/ATAC-merged.deduplicated.sort.Q30.bw"

cd ${RUNDIR}

OUTPUT\_FILE\_PREFIX="ATAC\_meth-cov\_4groups\_ATAC\_promoter"

group\_plot

#### for histone marks

group\_plot(){

computeMatrix reference-point --referencePoint center --scoreFileName ${BW\_1} ${BW\_2} --regionsFileName ${BED} --outFileName ${OUTPUT\_FILE\_PREFIX}\_matrix.gz --sortRegions **descend --sortUsingSamples 1** --missingDataAsZero --beforeRegionStartLength ${Midpoint} --afterRegionStartLength ${Midpoint}

plotHeatmap --matrixFile ${OUTPUT\_FILE\_PREFIX}\_matrix.gz --outFileName ${OUTPUT\_FILE\_PREFIX}\_heatmap.pdf --colorMap ${color} --sortRegions **descend** **--sortUsingSamples 1** --heatmapHeight 50 --legendLocation none --whatToShow 'heatmap and colorbar' --zMin ${zMin} --zMax ${zMax}

plotProfile --matrixFile ${OUTPUT\_FILE\_PREFIX}\_matrix.gz --outFileName ${OUTPUT\_FILE\_PREFIX}\_profilePlot.pdf --colors red blue green purple --legendLocation 'upper-center' --yMin ${y\_scale\_min} --yMax ${y\_scale\_max}

plotProfile --matrixFile ${OUTPUT\_FILE\_PREFIX}\_matrix.gz --outFileName ${OUTPUT\_FILE\_PREFIX}\_pergroupprofilePlot.pdf --colors red blue --legendLocation 'upper-center' --yMin ${y\_scale\_min} --yMax ${y\_scale\_max} --perGroup

}

#then run in one bloc for K4me1:

RUNDIR=/ifs/home/lhoump01/Method/deeptools\_heatmaps

Midpoint=1000

y\_scale\_min=0

y\_scale\_max=0.20

color=Greens

zMin=0

zMax=0.25

BW\_1="/ifs/home/lhoump01/Method/deeptools\_heatmaps/ATAC-merged.deduplicated.sort.Q30.bw"

BW\_2="/ifs/home/lhoump01/mESC\_ChIP-BS/Public/K4me1\_GSM1000121\_q30\_rmdup\_sorted.bw"

cd ${RUNDIR}

OUTPUT\_FILE\_PREFIX="ATAC\_meth-cov\_4groups\_K4me1\_promoter"

group\_plot

#then run in one bloc for K27ac:

RUNDIR=/ifs/home/lhoump01/Method/deeptools\_heatmaps

Midpoint=1000

y\_scale\_min=0

y\_scale\_max=0.4

color=Greens

zMin=0

zMax=0.5

BW\_1="/ifs/home/lhoump01/Method/deeptools\_heatmaps/ATAC-merged.deduplicated.sort.Q30.bw"

BW\_2="/ifs/home/lhoump01/mESC\_ChIP-BS/Public/K27ac\_GSM1000126\_q30\_rmdup\_sorted.bw"

cd ${RUNDIR}

OUTPUT\_FILE\_PREFIX="ATAC\_meth-cov\_4groups\_K27ac\_promoter"

group\_plot

#then run in one bloc for K4me3:

RUNDIR=/ifs/home/lhoump01/Method/deeptools\_heatmaps

Midpoint=1000

y\_scale\_min=0

y\_scale\_max=2

color=Greens

zMin=0

zMax=1

BW\_1="/ifs/home/lhoump01/Method/deeptools\_heatmaps/ATAC-merged.deduplicated.sort.Q30.bw"

BW\_2="/ifs/home/lhoump01/mESC\_ChIP-BS/Public/K4me3\_GSM1000124\_q30\_rmdup\_sorted.bw"

cd ${RUNDIR}

OUTPUT\_FILE\_PREFIX="ATAC\_meth-cov\_4groups\_K4me3\_promoter"

group\_plot

#

 ################## for inter\_intra ################

### combine intergenic and intragenic together

/ifs/home/lhoump01/Method/deeptools\_heatmaps/genomic-annotations

cat 1-Intragenic.bed 1-Intergenic.bed file1.txt 2-Intragenic.bed 2-Intergenic.bed file2.txt 3-Intragenic.bed 3-Intergenic.bed file3.txt 4-Intragenic.bed 4-Intergenic.bed file4.txt > 4groups\_InterIntra.bed

qlogin -q all.q

module unload python

module unload deeptools

module load deeptools/2.3.3

BED="/ifs/home/lhoump01/Method/deeptools\_heatmaps/genomic-annotations/4groups\_InterIntra.bed"

################### sort by ATAC signal

#then run in one bloc:

group\_plot(){

computeMatrix reference-point --referencePoint center --scoreFileName ${BW\_1} --regionsFileName ${BED} --outFileName ${OUTPUT\_FILE\_PREFIX}\_matrix.gz --sortRegions **descend --sortUsingSamples 1** --missingDataAsZero --beforeRegionStartLength ${Midpoint} --afterRegionStartLength ${Midpoint}

plotHeatmap --matrixFile ${OUTPUT\_FILE\_PREFIX}\_matrix.gz --outFileName ${OUTPUT\_FILE\_PREFIX}\_heatmap.pdf --colorMap ${color} --sortRegions **descend** **--sortUsingSamples 1** --heatmapHeight 50 --legendLocation none --whatToShow 'heatmap and colorbar' --zMin ${zMin} --zMax ${zMax}

plotProfile --matrixFile ${OUTPUT\_FILE\_PREFIX}\_matrix.gz --outFileName ${OUTPUT\_FILE\_PREFIX}\_profilePlot.pdf --colors red blue green purple --legendLocation 'upper-center' --yMin ${y\_scale\_min} --yMax ${y\_scale\_max}

#plotProfile --matrixFile ${OUTPUT\_FILE\_PREFIX}\_matrix.gz --outFileName ${OUTPUT\_FILE\_PREFIX}\_pergroupprofilePlot.pdf --colors red --legendLocation 'upper-center' --yMin ${y\_scale\_min} --yMax ${y\_scale\_max} --perGroup

}

#then run in one bloc for ATAC:

RUNDIR=/ifs/home/lhoump01/Method/deeptools\_heatmaps

Midpoint=1000

y\_scale\_min=0

y\_scale\_max=150

color=Oranges

zMin=0

zMax=75

BW\_1="/ifs/home/lhoump01/Method/deeptools\_heatmaps/ATAC-merged.deduplicated.sort.Q30.bw"

cd ${RUNDIR}

OUTPUT\_FILE\_PREFIX="ATAC\_meth-cov\_4groups\_ATAC\_InterIntra"

group\_plot

#### for histone marks

group\_plot(){

computeMatrix reference-point --referencePoint center --scoreFileName ${BW\_1} ${BW\_2} --regionsFileName ${BED} --outFileName ${OUTPUT\_FILE\_PREFIX}\_matrix.gz --sortRegions **descend --sortUsingSamples 1** --missingDataAsZero --beforeRegionStartLength ${Midpoint} --afterRegionStartLength ${Midpoint}

plotHeatmap --matrixFile ${OUTPUT\_FILE\_PREFIX}\_matrix.gz --outFileName ${OUTPUT\_FILE\_PREFIX}\_heatmap.pdf --colorMap ${color} --sortRegions **descend** **--sortUsingSamples 1** --heatmapHeight 50 --legendLocation none --whatToShow 'heatmap and colorbar' --zMin ${zMin} --zMax ${zMax}

plotProfile --matrixFile ${OUTPUT\_FILE\_PREFIX}\_matrix.gz --outFileName ${OUTPUT\_FILE\_PREFIX}\_profilePlot.pdf --colors red blue green purple --legendLocation 'upper-center' --yMin ${y\_scale\_min} --yMax ${y\_scale\_max}

#plotProfile --matrixFile ${OUTPUT\_FILE\_PREFIX}\_matrix.gz --outFileName ${OUTPUT\_FILE\_PREFIX}\_pergroupprofilePlot.pdf --colors red blue --legendLocation 'upper-center' --yMin ${y\_scale\_min} --yMax ${y\_scale\_max} --perGroup

}

#then run in one bloc for K4me1:

RUNDIR=/ifs/home/lhoump01/Method/deeptools\_heatmaps

Midpoint=1000

y\_scale\_min=0

y\_scale\_max=0.15

color=Greens

zMin=0

zMax=0.25

BW\_1="/ifs/home/lhoump01/Method/deeptools\_heatmaps/ATAC-merged.deduplicated.sort.Q30.bw"

BW\_2="/ifs/home/lhoump01/mESC\_ChIP-BS/Public/K4me1\_GSM1000121\_q30\_rmdup\_sorted.bw"

cd ${RUNDIR}

OUTPUT\_FILE\_PREFIX="ATAC\_meth-cov\_4groups\_K4me1\_InterIntra"

group\_plot

#then run in one bloc for K27ac:

RUNDIR=/ifs/home/lhoump01/Method/deeptools\_heatmaps

Midpoint=1000

y\_scale\_min=0

y\_scale\_max=0.4

color=Greens

zMin=0

zMax=0.5

BW\_1="/ifs/home/lhoump01/Method/deeptools\_heatmaps/ATAC-merged.deduplicated.sort.Q30.bw"

BW\_2="/ifs/home/lhoump01/mESC\_ChIP-BS/Public/K27ac\_GSM1000126\_q30\_rmdup\_sorted.bw"

cd ${RUNDIR}

OUTPUT\_FILE\_PREFIX="ATAC\_meth-cov\_4groups\_K27ac\_InterIntra"

group\_plot

#then run in one bloc for K4me3:

RUNDIR=/ifs/home/lhoump01/Method/deeptools\_heatmaps

Midpoint=1000

y\_scale\_min=0

y\_scale\_max=0.5

color=Greens

zMin=0

zMax=0.5

BW\_1="/ifs/home/lhoump01/Method/deeptools\_heatmaps/ATAC-merged.deduplicated.sort.Q30.bw"

BW\_2="/ifs/home/lhoump01/mESC\_ChIP-BS/Public/K4me3\_GSM1000124\_q30\_rmdup\_sorted.bw"

cd ${RUNDIR}

OUTPUT\_FILE\_PREFIX="ATAC\_meth-cov\_4groups\_K4me3\_InterIntra"

group\_plot

#

**Fig. 3c**

Of note: location of CTCF motif (MA0139.1) into each cytosine group (as per the fig. 2a) were obtained using FIMO on MEME suite.

#### merge CTCF motifs from the 4 groups

cat 1-CpGs0to20\_cov50andabove\_intoCTCFmotif\_uniq.bed file1.txt 2-CpGs0to20\_cov5to50\_intoCTCFmotif\_uniq.bed file2.txt 3-CpGs20to80intoCTCFmotif\_uniq.bed file3.txt 4-CpGs80to100intoCTCFmotif\_uniq.bed file4.txt > fimoCTCF\_4groups.bed

module load deeptools/2.3.3

BED="/ifs/home/lhoump01/Method/deeptools\_heatmaps/fimoCTCF\_4groups.bed"

group\_plot(){

  computeMatrix reference-point --referencePoint center --scoreFileName ${BW\_1} ${BW\_2} ${BW\_3} --regionsFileName ${BED} --outFileName ${OUTPUT\_FILE\_PREFIX}\_matrix.gz --sortRegions **descend --sortUsingSamples 1**--missingDataAsZero --beforeRegionStartLength ${Midpoint} --afterRegionStartLength ${Midpoint}

  plotHeatmap --matrixFile ${OUTPUT\_FILE\_PREFIX}\_matrix.gz --outFileName ${OUTPUT\_FILE\_PREFIX}\_heatmap.pdf --colorMap ${color} --sortRegions **descend** **--sortUsingSamples 1**--heatmapHeight 50 --legendLocation none --whatToShow 'heatmap and colorbar' --zMin ${zMin} --zMax ${zMax}

  plotProfile --matrixFile ${OUTPUT\_FILE\_PREFIX}\_matrix.gz --outFileName ${OUTPUT\_FILE\_PREFIX}\_profilePlot.pdf --colors red blue green purple --legendLocation 'upper-center' --yMin ${y\_scale\_min} --yMax ${y\_scale\_max}

  plotProfile --matrixFile ${OUTPUT\_FILE\_PREFIX}\_matrix.gz --outFileName ${OUTPUT\_FILE\_PREFIX}\_pergroupprofilePlot.pdf --colors red blue green --legendLocation 'upper-center' --yMin ${y\_scale\_min} --yMax ${y\_scale\_max}  --perGroup

}

############# Plot CTCF

#then run in one bloc:

RUNDIR=/ifs/home/lhoump01/Method/deeptools\_heatmaps

Midpoint=500

y\_scale\_min=0

y\_scale\_max=15

color=Blues

zMin=0

zMax=15

BW\_1="/ifs/home/lhoump01/Method/deeptools\_heatmaps/ATAC-merged.deduplicated.sort.Q30.bw"

BW\_2="/ifs/home/lhoump01/Method/deeptools\_heatmaps/CTCF-merged.deduplicated.sort.Q30.bw"

BW\_3="/ifs/home/lhoump01/Method/deeptools\_heatmaps/mES\_CTCF.deduplicated.sort.Q30.bw"

cd ${RUNDIR}

OUTPUT\_FILE\_PREFIX="**ATACpeaks\_4groups\_CTCFmotifs\_CTCF**"

group\_plot

### plot ATAC

RUNDIR=/ifs/home/lhoump01/Method/deeptools\_heatmaps

Midpoint=500

y\_scale\_min=0

y\_scale\_max=200

color=Oranges

zMin=0

zMax=75

BW\_1="/ifs/home/lhoump01/Method/deeptools\_heatmaps/ATAC-merged.deduplicated.sort.Q30.bw"

BW\_2="/ifs/home/lhoump01/Method/deeptools\_heatmaps/CTCF-merged.deduplicated.sort.Q30.bw"

BW\_3="/ifs/home/lhoump01/Method/deeptools\_heatmaps/mES\_CTCF.deduplicated.sort.Q30.bw"

cd ${RUNDIR}

OUTPUT\_FILE\_PREFIX="**ATACpeaks\_4groups\_CTCFmotifs\_ATAC**"

group\_plot