**Supplementary Documentation**

By: Yogindra Raghav

Supervisor: Dr. Miler Lee

**Note:** All computational tools are denoted using this font.

***Retrieving ChIP-Seq Data from NCBI Sequence Read Archive Using*** Fastq-Dump**:**

fastq-dump SRR392589 --split-files --readids --defline-qual + --gzip

fastq-dump SRR392592 --split-files --readids --defline-qual +  --gzip

***Building***Hisat2 Index**:**

hisat2-build reference\_files/danRer10\_chr.fa zebrafish\_index -p 4

***Alignment to danRer10 Genome Using***Hisat2***Genome Aligner*:**

hisat2 -p 4 --no-unal --no-spliced-alignment --phred33 -x index\_files/zebrafish\_index -U raw\_genomic\_data/SRR392589\_1.fastq.gz --summary-file my\_chip\_summary.txt | samtools view -hSb > nanog.bam

hisat2 -p 4 --no-unal --no-spliced-alignment --phred33 -x index\_files/zebrafish\_index -U raw\_genomic\_data/SRR392592\_1.fastq.gz --summary-file my\_chip\_summary.txt | samtools view -hSb > input.bam

***Calling Peaks Utilizing***Macs2 callpeak**:**

macs2 callpeak --bdg -g 6.4e8 --outdir nanog\_peaks -n nanog\_output -t bam\_files/nanog.bam -c bam\_files/input.bam

***Forming Promoter BED File with*** Bedtools**:**

bedtools flank -i reference\_files/danRer10\_refseq.bed -g reference\_files/danRer10\_chr\_sizes.txt -l 1 -r 0 -s > reference\_files/TSS.bed

bedtools slop -i reference\_files/TSS.bed -g reference\_files/danRer10\_chr\_sizes.txt -l 1000 -r 0 -s > reference\_files/all\_promoters.bed

***Seperating Peaks into Promoters and Enhancers Using*** *Bedtools Intersect***:**

bedtools intersect -a reference\_files/all\_promoters.bed -b nanog\_peaks/nanog\_output\_peaks.narrowPeak > intersected\_files/promoters\_bound.bed

bedtools intersect -a reference\_files/all\_promoters.bed -b nanog\_peaks/nanog\_output\_peaks.narrowPeak -v > intersected\_files/enhancers\_bound.bed

***Generating BigWig File from*** Macs2 callpeak***BEDGRAPH File Using*** wigToBigWig**:**

wigToBigWig nanog\_peaks/nanog\_output\_control\_lambda.bdg reference\_files/danRer10\_chr\_sizes.txt deeptools\_heatmap/nanog.bw

***Creating Combined Enhancer/Promoter Bed File for***deepTools Heatmap**:**

cat intersected\_files/promoters\_bound.bed > intersected\_files/combined\_promoters\_and\_enhancers.bed

echo "#Promoters" >> intersected\_files/combined\_promoters\_and\_enhancers.bed

cat intersected\_files/enhancers\_bound.bed >> intersected\_files/combined\_promoters\_and\_enhancers.bed

echo "#Enhancers">> intersected\_files/combined\_promoters\_and\_enhancers.bed

***Producing Matrix File for Heatmap Using***deepTools ComputeMatrix**:**

computeMatrix reference-point --referencePoint center -b 500 -a 500 -p 4 --sortRegions descend --missingDataAsZero -S deeptools\_heatmap/nanog.bw -R intersected\_files/combined\_promoters\_and\_enhancers.bed -o deeptools\_heatmap/nanog.matrix

***Making Final Heatmap with***deepTools plotHeatmap**:**

plotHeatmap -m deeptools\_heatmap/nanog.matrix -o deeptools\_heatmap/nanog\_heatmap.png --plotTitle "Nanog: Bound Promoters vs. Enhancers" -y "Pileup" -x "Gene Distance (base pairs)"

***Retrieving FASTA Sequence of Bound Promoters from*** Bedtools getfasta**:**

bedtools getfasta -name -fi reference\_files/danRer10\_chr.fa -bed intersected\_files/promoters\_bound.bed -fo motif\_finding/promoters\_bound.fa

***Retrieving FASTA Sequence of Bound Enhancers from*** Bedtools getfasta**:**

bedtools getfasta -name -fi reference\_files/danRer10\_chr.fa -bed intersected\_files/enhancers\_bound.bed -fo motif\_finding/enhancers\_bound.fa