[USER@biowulf gz]$ mkdir zcat

for i in {1..X}; do (echo "cd /data/USER/FOLDER/gz && \zcat IDENTIFIER\_${i}\_\*\_R1\_001.fastq.gz > /data/USER/FOLDER/gz/zcat/${i}\_R1\_Combined.fastq && \

zcat IDENTIFIER\_${i}\_\*\_R2\_001.fastq.gz > /data/USER/FOLDER/gz/zcat/${i}\_R2\_Combined.fastq"); done > zcat.swarm

[USER@biowulf gz]$ less zcat.swarm (press “q” for quit seeing)

[USER@biowulf gz]$ swarm -f zcat.swarm -g 20 -t 10

[USER@biowulf gz]$ cd zcat

[USER@biowulf zcat]$ mkdir fastqc

[USER@biowulf zcat]$ mkdir bowtie2

[USER@biowulf zcat]$ mkdir Remove\_Dup

[USER@biowulf zcat]$ mkdir MACS2

[USER@biowulf zcat]$ mkdir mm10\_TSS

Copy and paste mm10.refGene.TSS.bed file to mm10\_TSS/ folder

Copy and paste Encode\_blacklist/mm10-blacklist.v2.bed

for i in {1..X}; do (echo "cd /data/USER/FOLDER/gz/zcat && \

cutadapt -a CTGTCTCTTATACACATCT -A CTGTCTCTTATACACATCT -q 30 --cores=\$SLURM\_CPUS\_PER\_TASK --minimum-length 36 -o ${i}\_R1\_trimmed.fastq \

-p ${i}\_R2\_trimmed.fastq ${i}\_R1\_Combined.fastq ${i}\_R2\_Combined.fastq && \

fastqc -o /data/USER/FOLDER/gz/zcat/fastqc -t \$SLURM\_CPUS\_PER\_TASK -noextract ${i}\_R1\_Combined.fastq ${i}\_R2\_Combined.fastq && \

export BOWTIE2\_INDEXES=/fdb/igenomes/Mus\_musculus/UCSC/mm10/Sequence/Bowtie2Index && \

bowtie2 --very-sensitive -X 2000 -p \$(( SLURM\_CPUS\_PER\_TASK - 2 )) -x genome --no-mixed --no-discordant \

-1 ${i}\_R1\_trimmed.fastq -2 ${i}\_R2\_trimmed.fastq \

| samtools view -u - | samtools sort - > /data/USER/FOLDER/gz/zcat/bowtie2/${i}.bam && \

cd /data/USER/FOLDER/gz/zcat/bowtie2 && \

samtools index ${i}.bam &&\

samtools idxstats ${i}.bam | cut -f 1 | grep -v -P '^chrM$' | xargs samtools view ${i}.bam -b > ${i}\_RM.bam && \

java -Xmx10g -XX:ParallelGCThreads=10  -jar \$PICARDJARPATH/picard.jar MarkDuplicates I=${i}\_RM.bam \

O=/data/USER/FOLDER/gz/zcat/Remove\_Dup/${i}\_RM.bam \

M=${i}\_dups\_RM.txt REMOVE\_DUPLICATES=false VALIDATION\_STRINGENCY=LENIENT && \

cd /data/USER/FOLDER/gz/zcat/Remove\_Dup && \

samtools view -h -b -F1804 -f 2 -q 30 ${i}\_RM.bam > ${i}\_RMD.bam && samtools index ${i}\_RMD.bam && \

bedtools bamtobed -i ${i}\_RMD.bam > ${i}\_RM.bed && \

awk 'BEGIN {OFS = \"\t\"} ; {if (\$6 == \"+\") print \$1, \$2 + 4, \$3 + 4, \$4, \$5, \$6; else print \$1, \$2 - 5, \$3 - 5, \$4, \$5, \$6}' \

${i}\_RM.bed > ${i}\_shifted\_RM.bed && \

bedClip ${i}\_shifted\_RM.bed /data/USER/Biowulf\_STAR\_genomeDir\_ENsembl/UCSC\_Mouse\_Chromsizes.txt  ${i}\_shifted\_cliped\_RM.bed && \

cd /data/USER/FOLDER/gz/zcat/MACS2 && \

macs2 callpeak -t /data/USER/FOLDER/gz/zcat/Remove\_Dup/${i}\_shifted\_cliped\_RM.bed -f BED -q 0.05 --nomodel \

--shift -75 --extsize 150 -g mm --keep-dup all --call-summits -n ${i} && \

bedtools intersect -v -a ${i}\_peaks.narrowPeak -b /data/USER/FOLDER/Encode\_blacklist/mm10-blacklist.v2.bed \

> ${i}\_fileted\_peaks.narrowPeak && \

cd /data/USER/FOLDER/gz/zcat/Remove\_Dup && \

samtools index ${i}\_RM.bam &&\

bamCoverage --bam ${i}\_RM.bam -o ${i}\_RM.bw -bs=1 --normalizeUsing RPKM --extendReads -p=max && \

computeMatrix reference-point -S ${i}\_RM.bw -R /data/USER/FOLDER/gz/zcat/mm10\_TSS/mm10.refGene.TSS.bed \

--referencePoint TSS  -b 2000 -a 2000 -bs=1 -p=max -out ${i}\_matrix.gz && plotHeatmap -m ${i}\_matrix.gz -out \

${i}\_matrix.pdf --outFileNameMatrix ${i}.gz --colorMap=Blues"); done > XXXX.swarm

[USER@biowulf gz]$ swarm -f XXXX.swarm -g 38 -t 40 --module cutadapt/3.4,bowtie/2-2.4.4,samtools,macs/2.2.7.1,ucsc/418,deeptools/3.5.1,bedtools/2.30.0,fastqc/0.11.9,picard/2.25.7 --time 3-00:00:00 --gres=lscratch:100