Brad Olinger

Quant Bio

Week 5 Assignment

Mapping Reads

Command:

bowtie2 -p 4 -x mm10 CTCF\_ER4.fastq > CTCFER4\_mapped

bowtie2 -p 4 -x mm10 CTCF\_G1E.fastq > CTCFG1E\_mapped

bowtie2 -p 4 -x mm10 input\_ER4.fastq > inputER4\_mapped

bowtie2 -p 4 -x mm10 input\_G1E.fastq > inputG1E\_mapped

Calling Peaks

Command: macs2 callpeak -t CTCFER4\_mapped -c inputER4\_mapped --name=ER4 --gsize=61000000 –bdg

macs2 callpeak -t CTCFG1E\_mapped -c inputG1E\_mapped --name=G1E --gsize=61000000 --bdg

Differential Binding

Command: bedtools subtract -A -f 0.90 -a G1E\_peaks.narrowPeak -b ER4\_peaks.narrowPeak > old\_peaks.bam

bedtools subtract -A -f 0.90 -a ER4\_peaks.narrowPeak -b G1E\_peaks.narrowPeak > new\_peaks.bam

Feature Overlapping

Command: bedtools intersect -a Mus\_musculus.GRCm38.94\_features.bed -b ER4\_peaks.narrowPeak > ER4\_overlap

bedtools intersect -a Mus\_musculus.GRCm38.94\_features.bed -b G1E\_peaks.narrowPeak > G1E\_overlap

counting features

command: wc -l ER4\_overlap (output 415)

wc -l G1E\_overlap (output 324)

determining number of lost/gained peaks upon differentiation

command: wc -l old\_peaks.bam

wc -l new\_peaks.bam

results: lost = 207

gained = 343

Part 2: Motif Discovery

Save the 100 strongest ChIP-seq peaks

Command: sort -k 9 -n -r ER4\_peaks.narrowPeak | head -n 100 > ER4\_100.narrowPeak

Extract sequences using bedtools

Command: bedtools getfasta -fi /Users/cmdb/data/genomes/mm10.fa -bed ER4\_100.narrowPeak > ER4\_seq\_100

Using memechip for motif finding

Command: meme-chip -maxw 20 ER4\_seq\_100

Download jaspar database and unzip

Command: wget <https://bx.bio.jhu.edu/data/msauria/cmdb-lab/JASPAR2018_CORE_non-redundant_pfms_meme.zip>

Copy meme motif database

Command: cp /Users/cmdb/data/meme\_db/motifs/motif\_databases.12.21.tgz .

gunzip -c motif\_databases.12.21.tgz | tar xvf –

Using Tomtom

Command: tomtom MA\* memechip\_out/combined.meme

**Logo:**

