**BIOM262: ChIP-Seq workshop – Tuesday Feb 13th 2018**

Here we will explore IGVtools, use HOMER for studying NDRs (nucleosome depleted regions), and follow a tutorial for bedtools.

**IGVtools:**

Creating tdf files:

We will need to sort the bam files as well as index them.

We will use data from Javasky & Shamir *et al*. The bam files for chromosome 12 are located at

/home/ucsd-train36/biom262-2018/class-iii/data/

To sort the files, we will run:

***for f in <path>/data/\*bam; do fname=`basename $f -chr12.bam`; samtools sort -o*** *<path****>/sorted/$fname.sorted.bam $f; done***

***for f in*** *<path>/****sorted/\*bam; do samtools index $f; done***

we will need interactive shell:

***qsub -I -l nodes=1:ppn=2 -l walltime=1:00:00 -q hotel***

and

***for f in*** *<path>/* ***sorted /\*bam; do fname=`basename $f -chr12.sorted.bam`;***

**igvtools count -e 164 $f** *<path>****/tdfs/$fname.tdf hg19; done***

copy the tdfs to your computer and view on IGV.

**Plotting the NDR signals**

The data from Javasky & Shamir *et al*. is derived from human cells, so we first should add to HOMER the needed genome:

**perl** <path; mine was at ~/software/homer/>/configureHomer.pl **-install hg19**

The bam files for chromosome 12 are located at

/home/ucsd-train36/biom262-2018/class-iii/data/

Now, let’s make a tag-directories for all files (as we did before):

***for f in*** *<path to the bam files>****/\*.bam; do fname=`basename $f -chr12.bam`; makeTagDirectory <****path to where your tag directories will be****>/$fname -genome hg19***

***-checkGC $f; done***

Once we have the tag-directories, we can use the annotatePeaks.pl command

***annotatePeaks.pl tss hg19 -size 4000 -hist 10 -d*** *<*path to the tagdirs>/\*k27\* > <path>/h3k27ac-hist.txt

and

***annotatePeaks.pl tss hg19 -size 4000 -hist 10 -d*** *<*path to the tagdirs>/\*k4\* > <path>/h3k4me3-hist.txt

Open the files using a spreadsheet program (R, Excel or similar). As we discussed, the first column gives the distance offsets from the TSS followed by columns corresponding to the ‘coverage’ (calculated by extending tags by their estimated ChIP-fragment length; analogous to the profiles made for the Genome Browser), ‘+ Tags’, and ‘- Tags’ (density of 5’ and 3’ aligned tags; these are independent of fragment length) for each experiment.

Plot the coverage of mitosis and interphase H3K27ac and H3K4me3 – do you see a difference in the way they are organized around the TSS?

**BEDtools:**

We will follow the tutorial by Aaron Quinlan

<http://quinlanlab.org/tutorials/bedtools/bedtools.html>