ATACseq Analysis

Step 1 – QC & Alignment

Shell Environment:

The original “.bam” files were first converted into fastq files for fastQC analysis:

* + Converting from bam to fastq files
  + Running Fastqc
  + Trimming using Trimmomatic   
    (bash Trimmomatic\_Script.sh > ./Trimmed\_Fastqs/Trimmomatic\_Commands.txt;

bash Trimmomatic\_Commands.txt)

* + Fastqc on the trimmed fastq files (P files).
  + Fixed pairing using repair.sh command from BBMap (repair\_fastqs.sh script)
  + Mapped the fixed fastqs to hg38 genome (/home/kfirinbal/ATAC-seq\_CovVac/Reference\_Annotations/) using bwa-mem (bwa\_mem\_map.sh)
  + After that mapping, no duplicates were found, so removing duplicates step was skipped.
  + Removed Mitochondrial Chromosomes using Samtools (subset\_markdupl\_index\_stat.sh).
  + Removed blacklist regions (RemoveBlackListRegions.sh).
  + Applied shifting using alignment sieve (AlignmentSieve.sh).

Step 2 – Peak Calling

Peak calling was implemented using macs2 (Macs2\_Generic\_Script.sh)

Then downloaded the peak files (.xls, .bed, .narrowPeak) and the corresponding NoMT\_subset.blacklist-filtered.shifted.sorted bam files to a folder on my local computer. Then, created a R script to further analyze the data using DiffBind. Also, created a metadata file to split the files into groups of: CD4\_Good, CD4\_Poor, CD8\_Good and CD8\_Poor.

Step 3 - TOBIAS

* ATACorrect - nohup bash TOBIAS\_FootPrint\_Docker.sh > TOBIAS\_Docker\_output.log 2>&1 &
* ScoreBigWig - nohup bash ScoreBigwig.sh > ScoreBigwig.log 2>&1 &
* BINDetect - ?