Results for ChIPSeq Pipeline

Output folders:

metrics:
PICARD metrics from mapping phase

alignments:
RAW unprocessed BAM files. Note before using the BAM's we post process them according to the ENCODE recipies which does the following:

- remove unmapped, mate unmapped, non-proper paired reads
- remove reads with MAPQ < 10
- remove duplicates
- remove Failed QC reads

chipSeq/macs: output of running `macs2 callpeaks` on the post processed bams. For IP's with focal transcription factors we use the following macs arguments:

- `-q 0.01`

for those with non-focal/broad binding we use:

- `--broad --broad-cutoff 0.1`

See the MACS2 website [https://github.com/taoliu/MACS] for more information on the output.

There is also an aggregate peak file there which combines the peaks from all samples
- macs/macsPeaksMerged.saf
and a file that counts the coverage for each sample within these combined peaks
- macs/peaks_raw_fcCounts.txt

chipSeq/qc: QC report/plots for ChIP related QC.

- qcChIPSeq_PROJECT-NUM_.xlsx - is a report for the total number of peaks found and the number of significant peaks. Low number of peaks could indicate an issue with the ChIP.
- qcChIPSeq_PROJECT-NUM_.pdf - plots of the number/percentage of mapped reads that fall in MACS peaks and a PCA plot of the aggregate peak counts for the samples.

chipSeq/bw:
Normalized (to 10million reads) bigwig files for loading IP profiles into IGV [http://software.broadinstitute.org/software/igv/].

chipSeq/annote:
Annotation of MACS2 peak file using HOMER `annotatePeaks.pl`. See the HOMER website [http://homer.ucsd.edu/homer/ngs/annotation.html] for more details on the output.

chipSeq/diff:
Differential analysis of peaks using `edgeR`