RNA Sequencing Analysis Pipeline v1.0.3

**RNA-Seq
Illlumina run**
FastQ files

Create directory structure
+ <run>
 + <sample>
  + fastqc
  + jobs
  + logs
  + mapping
  + read_counts
 + read_counts
 + jobs
 + logs
 + bamMetrics

Genomes:
Human: GRCh37
Rat: Rnor5.0
Mouse: GRCm38
Zebrafish: Zv9
Dog: CanFam3.1

Perform quality control on FastQ files
FastQC v0.10.1

For each sample:
FastQC report

Align reads from FastQ files against reference genome
STAR v2.4.2a

For each sample:
-BAM file
-Chimeric junctions table
-Chimeric junctions SAM
-Splice junctions table
-Logfiles

Splice junctions analysis

Fusion genes analysis

GTF:
Human: GRCh37 v74
Rat: Rnor5.0 v71
Mouse: GRCm38 v70
Zebrafish: Zv9 v75
Dog: CanFam3.1 v75

Add read groups to BAM (Picard v1.98)
Index BAM (Sambamba v0.5.4)

Count reads in features (genes, transcripts and/or exons)
HTSeq-count 0.6.1p1

For each sample:
-Sorted BAM + BAI file

For each sample:
Txt file with number of mapped reads per feature

Collect metrics about the alignment of RNA to various functional classes
bamMetrics.pl (using Picard v1.135)

Merge
Create table from all sample count files

PDF report containing quality statistics of sorted BAM

Example design file:
sample1   test
sample2   control
sample3   test
sample4   test
sample5   control

Raw read counts table

Compute reads per kilobase per million (RPKM) values
EdgeR v3.8.6

Normalize read counts by their size factor
DESeq v1.18.0

Differential expression analysis
DESeq2 v1.6.3

RPKMs table

Normalized read counts table

- Result table of DE analysis
- MA plot
- PCA plot
- Heatmap sample-to-sample correlation