RNA-seq pipeline (please change wording to prevent plagiarism when submitting for publication):

1. Trimmed poor quality reads using fastq-mcf (1.05) [1-2] and checked quality using fastqc (v0.11.7)[3].
2. Align reads to respective genome (hg38 or mm10) using STAR (version 2.6.0) [4]

--outFilterMultimapNmax 20 --alignSJoverhangMin 8 --alignSJDBoverhangMin 1 --outFilterMismatchNmax 999 --outFilterMismatchNoverLmax 0.04 --alignIntronMin 20 --alignIntronMax 1000000 --alignMatesGapMax 1000000

1. Ensembl gene annotation (hg38 or mm10)
2. Count reads using subread counts (1.6.2) [5].
3. May filter out genes with row means less than 5 or greater than 5000 (see R code/html output).
4. Differential expression using edgeR (3.24.3) [6]
5. Gene Ontology using TopGO (2.34.0) fisher elim method [7]. ‘

References

1. Erik Aronesty (2011). ea-utils : "Command-line tools for processing biological sequencing data"; <https://github.com/ExpressionAnalysis/ea-utils>
2. Erik Aronesty (2013). TOBioiJ : "Comparison of Sequencing Utility Programs", [DOI:10.2174/1875036201307010001](http://benthamscience.com/open/openaccess.php?tobioij/articles/V007/1TOBIOIJ.htm)
3. Andrews S. (2010). Available online at:<http://www.bioinformatics.babraham.ac.uk/projects/fastqc>
4. Dobin et al. 2013. STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29(1):15-21. [pubmed:23104886](https://www.ncbi.nlm.nih.gov/pubmed/23104886); [doi:10.1093/bioinformatics/bts635](https://doi.org/10.1093/bioinformatics/bts635)
5. Liao Y, Smyth GK and Shi W (2014). [featureCounts: an efficient general purpose program for assigning sequence reads to genomic features.](http://www.ncbi.nlm.nih.gov/pubmed/24227677) Bioinformatics, 30(7):923-30.
6. Robinson MD, McCarthy DJ and Smyth GK (2010). “edgeR: a Bioconductor package for differential expression analysis of digital gene expression data.” Bioinformatics, 26, pp. -1.
7. Alexa A, Rahnenfuhrer J (2019). topGO: Enrichment Analysis for Gene Ontology. R package version 2.36.0.