1) STAR- **alignment** of reads to the transcriptome/genome

2) RSEM- **quantification** of transcripts

Genome/Transcript annotations:

http://www.gencodegenes.org/

Combined pipeline:

<https://github.com/ENCODE-DCC/long-rna-seq-pipeline/blob/master/DAC/STAR_RSEM.sh>

STAR manual:

<https://github.com/alexdobin/STAR/blob/master/doc/STARmanual.pdf>

RSEM manual:

<http://deweylab.biostat.wisc.edu/rsem/README.html>

3) generate a read counts matrix:

rsem-generate-data-matrix \

./Sample1.genes.results \

./Sample2.genes.results \

./SampleX.genes.results \

> raw\_counts.matrix.txt

4) use Rstudio to run DEseq2 for **differential expression**:

DESeq2 Vignette:

http://bioconductor.org/packages/release/bioc/vignettes/DESeq2/inst/doc/DESeq2.pdf

DESeq2 Manual: https://bioconductor.org/packages/release/bioc/manuals/DESeq2/man/DESeq2.pdf

------------------------------------------------------------------------------------------------------------

Galaxy RNA-seq tutorial:

http://birg.cs.cofc.edu/index.php/Galaxy\_RNA-seq\_Tutorial

SCG3 cluster user guide:

<https://web.stanford.edu/group/scgpm/cgi-bin/informatics/wiki/index.php/Cluster_user_guide>

https://www.encodeproject.org/rna-seq/long-rnas/