2021 Short Read Workshop

# Day 7 homework:

1. In class we ran DESeq2 with the contrast c("ifn\_treat", "ifn", "ctrl"). Now answer the following questions and perform the following analysis
   1. What did we confound by running the above contrast in our analysis?
   2. How should we interpret these results?
   3. Read the documentation for your design variable at the official GitHub repo, here: <https://github.com/mikelove/DESeq2/blob/master/R/results.R>
   4. Create a new group that allows us to perform DEA on the two genotypes, independently (HINT: this was covered in today’s slides)
   5. Compare the top 10 significant results for DEA on the two genotypes individually. Are they the same set? If not, why might that be? How do these results compare with the original analysis done in class?
2. In the following directory: /scratch/Shares/public/sread2021/homework\_data\_files/day6\_7/mousebrain/ you will find files containing a counts table file and metadata file similar to those we used in class.
   1. Check the formats of these files and how they compare to those we used previously (NOTE: this means look at the contents, not just the file extension)
   2. Reformat these files so that you can run your DESeq2 script on them in the same fashion as we did in class
3. In the following directory: /scratch/Shares/public/sread2021/homework\_data\_files/day6\_7/mousefibro/ you will find a directory containing sorted bam files, as well as a metadata file (runinfo.txt) containing the experiment info that pertains to the bam files.
   1. The first column of the metadata file corresponds to the files in the sortedbams directory, however it is not the exact path names to the files. Change this column so that it will reference the bam files explicitly (NOTE: the addinfotometa.R script will be useful for doing this)
   2. Once the metadata file is formatted correctly, run featureCounts and DESeq2 to analyze these samples.