**Using IDEP: Shiny App for DEseq2**

Advantages: easy to use web interface; can export/save processed data and figures.
Disadvantages: cannot save the R project online; need to repeat analysis every visit.

Help: At the bottom of the left panel of each page, **?** is a link to help/documentation.
R page (far right at top): cam download customized R code (with markdown) and more as electronic notebook.

1) Web site: http://bioinformatics.sdstate.edu/idep/

2) Load Data page: load raw counts and target files.
   (minor changes to sample names may be required; see https://idepsite.wordpress.com/data-format/)
   1 – Best matching species (EnsemblIDs are species-specific)
   2 – Read counts data
   3A – Upload file: RawCounts_IDEP.txt
   4B – Upload expt design file: target_IDEP.txt
   After loading, you should see two tables on the right: Study_design (header + group) and counts table (Ensembl IDs).
   If there are errors, correct them now!
   Delete columns from input files as needed to subset dataset and reload.

3) Pre-Process page
   Keep genes with: minCPM=1 in at least 2 libraries (as a starting point, adjust as needed)
   Transform counts: VST
   Missing values: Gene median (default)
   Option to plot or download results
   Processed data = filtered normalized counts (VST)
   Converted counts data = filtered raw counts with gene symbols

4) Heatmap, k-Means, PCA
   **Unsupervised analyses: driven by genes variable among samples, not statistically-DE genes between groups**
   Heatmap: typically genes are centered and normalized = Z-score. Samples are not.
   Click through different options, plots, and analyses

5) DEG1
   Method: DEseq2
   FDR (adj-p value) cutoff: 0.05
   Min fold change: 2 (relax this filter first to include more DE genes)
   **Important! Select factors & comparisons to include/exclude appropriate comparisons.**
   By default, all pairs and orders (AvB, BvA) are shown. For AvB, A=numerator and B=denominator in log2FC ratio.
   Option to plot overlap counts (venn diagram) and download results
   Gene lists = EnsemblIDs for up- and down-DE genes for each comparison
   FDR & fold-changes for all genes = log2FC + adj-p value for all genes, all comps + normalized counts for each sample

6) DEG2, Pathway
   **Supervised analyses: driven by DE genes**
   Click through different options, plots, and analyses
   PGSEA w/ all samples is nice, uses all comparisons at once
   Use the help page for documentation/description of options

**If you get errors, start over…**