Balan Ramesh and TJ Firneno

Methods

We demultiplexed the raw Ilumina reads and removed the barcodes using the fastx\_barcode\_splitter and the fastx\_trimmer modules of the FASTX-TOOLKIT (WEBSITE), respectively. We examined sequence quality on a per sample basis using FASTQC v0.11.3 and MULTIQC v1.8(WEBSITE). We indexed the *Rana pipiens* reference transcriptome (www.davislab.net/rana/) and then quantified the expression of our reads to the indexed reference transcriptome using the program Salmon v0.11.3. We used the DESeq2 package (CITATION) in a custom R script find the differential expression between the treated test group and the control group. In brief, this script takes the counts for each transcript for each sample and then, using the given design, it calculates the size factors for each transcript, estimates the dispersion of each transcript, and tries to fit a negative binomial GLM overall. The p-value is then calculated and adjusted using the negative binomial Wald’s test---. All of the transcripts with an adjusted p-value < 0.05 are considered differentially expressed. Then transcripts with read counts </= 10 across all twelve samples are removed from further analysis. Finally, the script normalizes the read counts using variance stabilizing transformations (VST) and removes the batch effects due to developmental stages, and creates a biplot using PC1 and PC2. We then identified and blasted differentially expressed transcripts against the Xenopus laevis and Xenopus tropicalis transcriptomes to annotate the function of our differentially expressed transcripts.

Results