Lab #8

By the beginning of the next lab (March 26th), send what you have to [afodor@uncc.edu](mailto:afodor@uncc.edu) with “Lab #8” in the subject line.

1. We return to our RNA seq dataset of E. Coli genes from mice. The URL is here:

http://afodor.github.io/classes/stats2015/longitdunalRNASeqData.zip

1. Read and normalize the counts table ( “nc101\_scaff\_dataCounts.txt “ into R). For example:

setwd("C:\\somewhere")

myT<-read.table("nc101\_scaff\_dataCounts.txt",sep="\t",header=TRUE)

# remove rare genes

myT <- myT[ apply( myT,1, median)> 5,]

myTNorm <- myT

for ( i in 1:ncol(myT))

{

colSum = sum(myT[,i])

myTNorm[,i] =myTNorm[,i]/colSum

}

(The first 3 columns are “day 2”, the next 3 columns are “week 12” and the last 5 are “week 18”). Remember, that day 2 is before the mice have symptoms, week 12 is associated with inflammation and week 18 is associated with cancer.

1. For every row in the normalized spreadsheet, run three t-tests ( “day 2” vs. “week 12”, “day 2” vs. “week 18” and “week 12” vs. “week 18”. At a p < .05 threshold fill in the following table:

|  |  |  |  |
| --- | --- | --- | --- |
|  | # of genes significant at p <0.05 uncorrected | # genes significant at p <0.05 BH FDR corrected | # genes significant at p <0.05 Bonferroni corrected |
| “day 2” vs. “week 12” |  |  |  |
| “day 2” vs. “week 18” |  |  |  |
| “week 12” vs. “week 18” |  |  |  |

1. Based on this table, when is the biggest shift in the transcriptome? Which samples are most different from one another?