

Lab__book__12__2__16

Claire Green

Monday

I have now completed my second attempt at a literature review. This will be uploaded to the repository. Now that this is finished my next step is to look at creating a statistically robust protocol for generating lists of DEGs.

Notes for meeting with Wenbin

Data: Affy HG Plus 2.0 RNA expression (except FTLN - HG U133A 2.0)

C9orf72 motor neurons: 8 patients, 3 controls CHMP2B motor neurons: 3 patients, 6 controls SALS motor neurons: 7 patients, 3 controls FTLN cortical neurons: 16 patients, 8 controls VCP muscle: 7 patients, 3 controls

R script

1. Run each data set through Wenbin's code individually
2. Take top X most DEGs -> Do I rank by fold change, pvalue or adjusted p value?
3. Find genes that are consensus across all data sets -> which genes are commonly dysregulated?
4. Validate number is more than you would expect by chance by random permutations test
 - say top 2000 from each data set yields 10 common genes
 - take 5 sets of 2000 random genes
 - calculate consensus value
 - repeat 10,000 times
 - number of times value is >10 / 10,000 should equal p value (less than 0.05 is acceptable)