

LabBook_11_08_2017

Claire Green

Tuesday

I was thinking about the mutation differences and how to visualise that on the network. Previously what I did was build an individual network for the most dysregulated genes for each mutation and show the overlap. Yesterday I realised that there was probably a much easier way of visualising that, so I changed my approach.

What I thought was that each mutation has a list of the genes in the network, and a measure of how dysregulated those genes are (p value). What I figured I could do is show that each mutation has higher levels of dysregulation in different areas of the network. This was achieved by colouring the network nodes based on the pvalue number.

The issue came when I had to use the pvalues from multiple datasets, because the p values were quite different ranges. Instead I decided to rank the genes so that the colouring would be more “normalised” (for want of a better word). The rank was then used to colour the nodes - and this was the result:

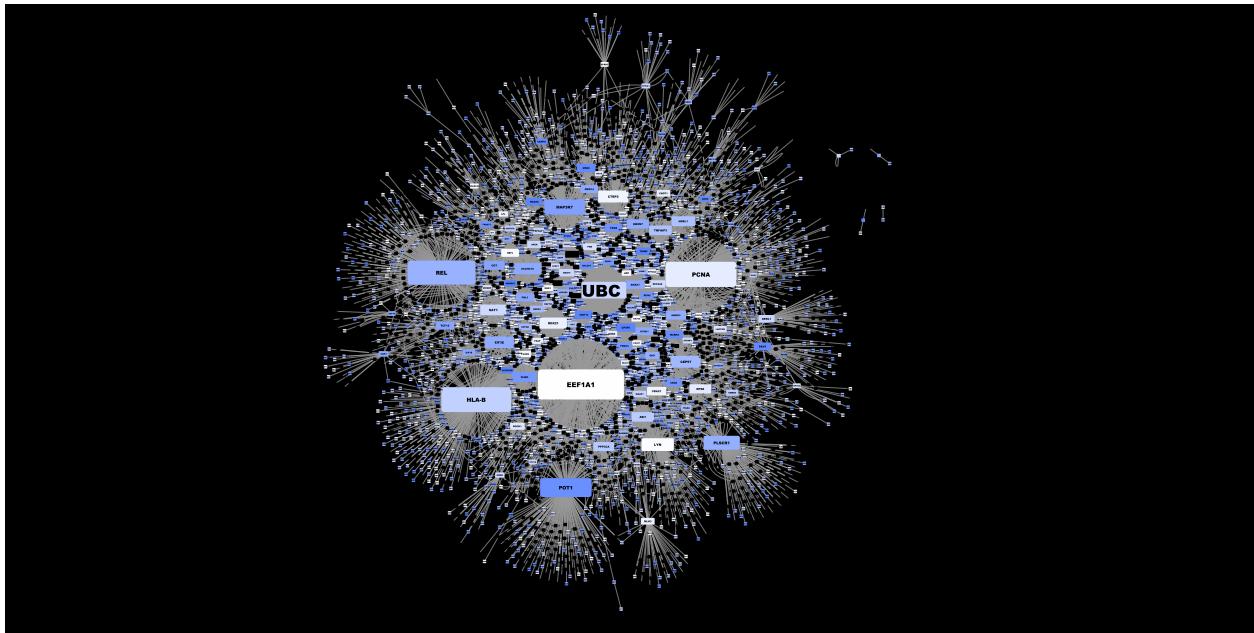


Figure 1: C9orf72

This is just a visual representation of the differences, it would have to be quantified somehow

Code can be found in file “GeneRank4PPI.R”

```
setwd("~/Documents/PhD/TDP-43/TDP-43_Code/Results/GeneExpression/noMedian/")

C9 <- read.csv("C9_unique.csv")
GRN <- read.csv("GRN_FTLRankeduniqueresult.csv")
VCP <- read.csv("vcp_unique.csv")

setwd("/Users/clairegreen/Documents/PhD/TDP-43/TDP-43_Code/Results/PPI_Network/")
```

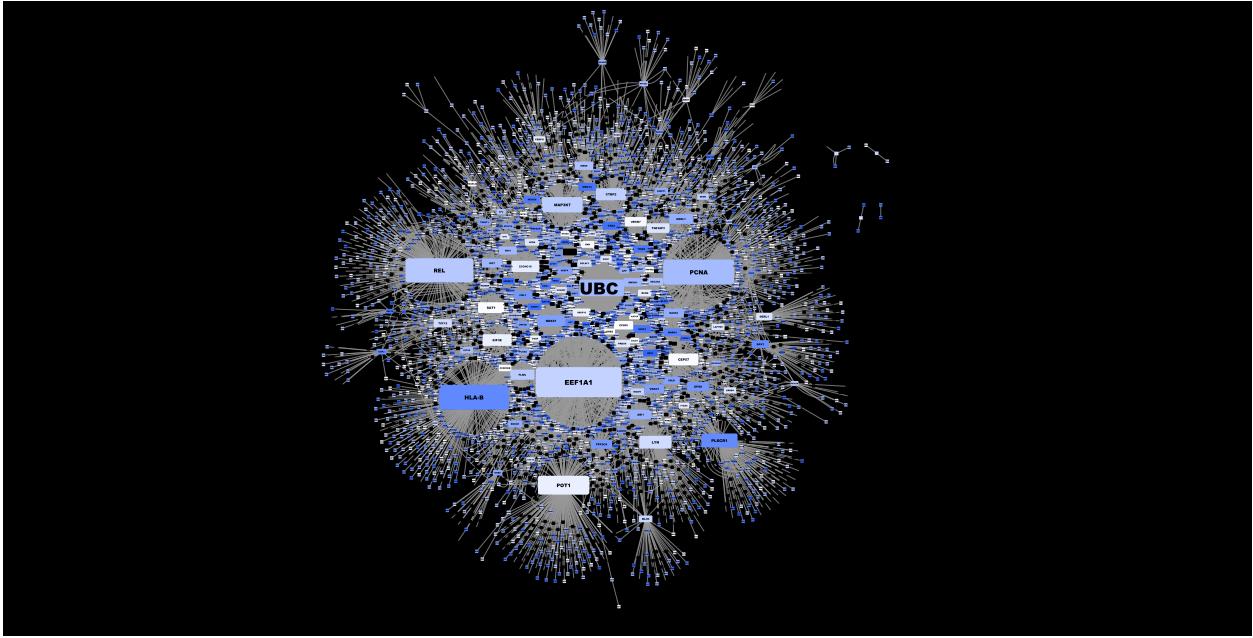


Figure 2: VCP

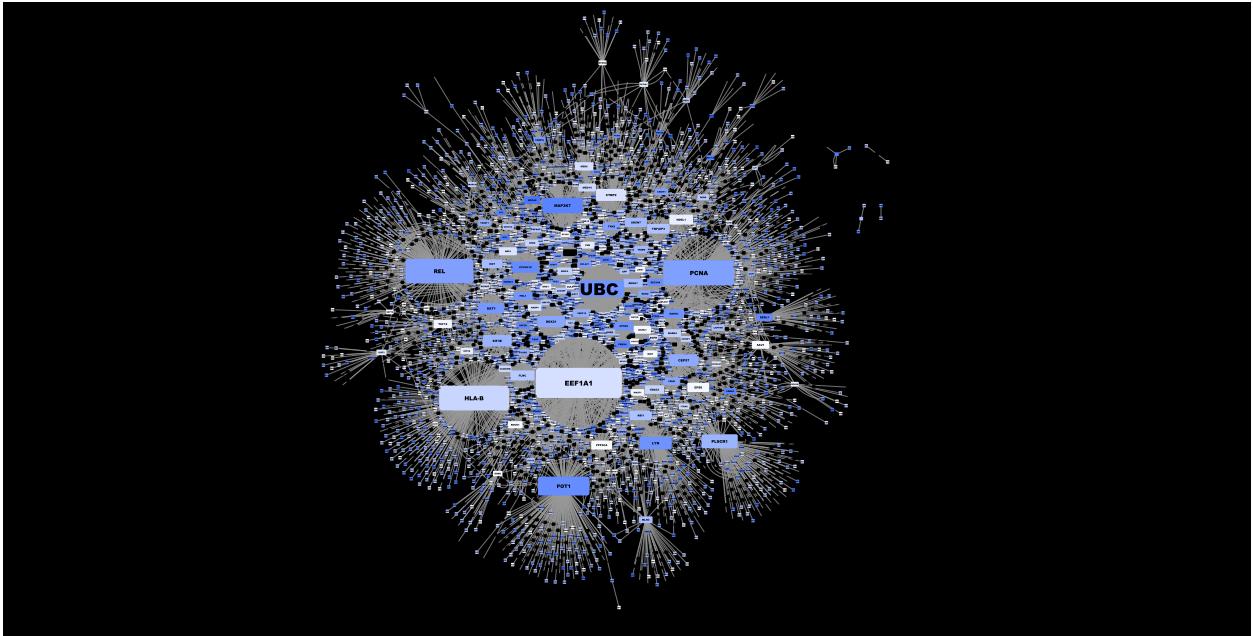


Figure 3: GRN

```
DEGPPI <- read.table("DEG_PPI_Genes.txt")
DEGPPI <- DEGPPI$V1

mut_in_DEG <- subset(GRN, GRN$Gene.Symbol %in% DEGPPI)
mut_w_var <- data.frame(mut_in_DEG$Gene.Symbol, mut_in_DEG$adj.P.Val)
colnames(mut_w_var) <- c("GeneSymbol", "GRN_adjpval")
mut_w_var$GRNrank <- 1:length(mut_w_var$GeneSymbol)

setwd("/Users/clairegreen/Documents/PhD/TDP-43/TDP-43_Code/Results/PPI_Network/heatnode/")
write.csv(mut_w_var, "GRN_overlap_adjpval.csv", row.names = F)
```