

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)

Tuesday

I discussed with Wenbin about the optimal way of ordering DEGs, and he essentially said that different people have different methods. I decided then to compare my output of DEGs ordered by adjusted p value as compared to fold change. Expression was calculated by WenBin DE Gene.R and sorted using the following lines of code:

```
###Write results to CSV files for consensus analysis
setwd("/Users/clairegreen/Documents/PhD/TDP-43/TDP-43_Data/DE Genes/TopGenes_2016-02-09/")
# dir.create(paste("TopGenes", Sys.Date(), sep = "_"))
#Take results, remove duplicate rows for genes, order by adjusted p value and take top X number of genes
uniqueresult <- result[!duplicated(result[,15]),]

#For ordering by adjusted p value

genesort <- uniqueresult[order(uniqueresult$adj.P.Val),]
topgene <- genesort[1:1000,]
write.csv(x = topgene, file = "VCP_ap_1000")
topgene <- genesort[1:2000,]
```

```

write.csv(x = topgene, file = "VCP_ap_2000")
topgene <- genesort[1:3000,]
write.csv(x = topgene, file = "VCP_ap_3000")
topgene <- genesort[1:4000,]
write.csv(x = topgene, file = "VCP_ap_4000")
topgene <- genesort[1:5000,]
write.csv(x = topgene, file = "VCP_ap_5000")

#For ordering by fold change
genesort <- uniqueresult[order(uniqueresult$`Fold Change`),]
topgene <- genesort[1:500,]

genesort <- uniqueresult[order(-uniqueresult$`Fold Change`),]
botgene <- genesort[1:500,]

topFC <- rbind(topgene,botgene)

write.csv(x = topgene, file = "VCP_fc_1000")

```

I evaluated consensus using the file DE consensus.R. It is as follows:

```

#####DIFFERENTIAL GENE EXPRESSION INTERSECT
#takes csv files of top X DE genes and identifies any consensus genes

setwd("/Users/clairegreen/Documents/PhD/TDP-43/TDP-43_Data/GeneExpressionAnalysis/TopGenes_2016-02-09/")
A <- read.csv(file = "VCP_ap_1000")

B <- read.csv(file = "VCP_fc_1000")

# sALS<- read.csv(file = "sALS_anno_5000")
#
# FTLD<- read.csv(file = "FTLD_anno_5000")
#
# VCP<- read.csv(file = "VCP_anno_5000")

C9_DE<- A$Gene.Symbol
CHMP2B_DE <- B$Gene.Symbol
# sALS_DE <- sALS$Gene.Symbol
# FTLD_DE <- FTLD$Gene.Symbol
# VCP_DE <- VCP$Gene.Symbol
overlap <- Reduce(intersect, list(C9_DE, CHMP2B_DE))
print(overlap)

```

According to this example, 250 genes were shared by the two methods. This indicates there could be a big difference in results by ranking by adjusted p value rather than fold change.

My next task is to identify whether one method produces more consensus than the other. To do this I will generate top 1000 DEGs for all 5 data sets using the two methods, and compare number of consensus results. If 1000 is not enough, I will try 2000.

UPDATE When I tried to conduct the experiment again, I got completely different results. In december, I generated:

1000 Genes

0

2000 Genes

CSRP1 RNF130

3000 Genes

CSRP1 RNF13 TUBB3 PSAP DDX39A NKTR NAGA NUTF2 /// NUTF2P4 RPL6 CCT2 DCN RNF130
TARS PFDN1 HSD17B4

4000 Genes

UPF3A DYNLT1 CSRP1 EEF1A1 RNF13 MAP4K4 CDK5R1 CST3 ICMT TUBB3 BRD3 DMD PSAP
DDX39A NKTR TFRC NAGA WBSCR22 NUTF2 /// NUTF2P4 PAICS RPL6 CCT2 DCN RPS6 GSTO1
TARDBP PRKD1 RNF130 TARS PFDN1 BPTF HSD17B4 PSMD1 GNPAT C14orf1 TMEM59 TRO
ATP6V1G2-DDX39B /// DDX39B /// SNORD84 ABCD3 PTEN ZFYVE16 DDX5 /// MIR3064 ///
MIR5047 GTF2I /// GTF2IP1 /// LOC100093631 MYL12A /// MYL12B LBR MTMR9 LSM5 WFS1
TANK PPP4R1 C18orf32 /// RPL17 /// RPL17-C18orf32 /// SNORD58A /// SNORD58B /// SNORD58C
RAB40B

When I ran the same results (generated Dec 4th 2015) again using the DE consensus.R script, I ended up
generating slightly different results:

1000

0

2000

CSRP1

3000

CSRP1 RNF13 FN1 TUBB3 PSAP RPL6 CCT2 NKTR NUTF2 /// NUTF2P4 NAGA PFDN1 TARS
RNF130 HSD17B4 DDX5 /// MIR3064 /// MIR5047 GTF2I /// GTF2IP1 /// LOC100093631

4000

UPF3A DYNLT1 CSRP1 EEF1A1 ETS2 RNF13 FN1 CST3 MAP4K4 BRD3 TUBB3 PSAP RPL6 CCT2
DMD ICMT NKTR NUTF2 /// NUTF2P4 RPS6 CDK5R1 BPTF PRKD1 NAGA GSTO1 PFDN1
ATP6V1G2-DDX39B /// DDX39B /// SNORD84 TARDBP TARS PTEN PAICS RNF130 HSD17B4
TMEM59 TRO DDX5 /// MIR3064 /// MIR5047 GNPAT GTF2I /// GTF2IP1 /// LOC100093631 WB-
SCR22 C14orf1 MTMR9 TCF4 WDR78 LBR DDX39A C18orf32 /// RPL17 /// RPL17-C18orf32 ///
SNORD58A /// SNORD58B /// SNORD58C DCN RPLP2 /// SNORA52 PSMD1 RECQL MPHOSPH9

This was generated using the following DE consensus.R script:

```
#####DIFFERENTIAL GENE EXPRESSION INTERSECT
#takes csv files of top X DE genes and identifies any consensus genes

setwd("/Users/clairegreen/Documents/PhD/TDP-43/TDP-43_Data/DE Genes/TopGenes_2016-02-09/")
C9orf72 <- read.csv(file = "C9_ap_5000")

CHMP2B <- read.csv(file = "CH_ap_5000")

sALS<- read.csv(file = "sALS_ap_5000")

FTLD<- read.csv(file = "FTLD_ap_5000")

VCP<- read.csv(file = "VCP_ap_5000")

C9_DE<- C9orf72$Gene.Symbol
CHMP2B_DE <- CHMP2B$Gene.Symbol
sALS_DE <- sALS$Gene.Symbol
FTLD_DE <- FTLD$Gene.Symbol
VCP_DE <- VCP$Gene.Symbol
overlap <- Reduce(intersect, list(C9_DE, CHMP2B_DE, sALS_DE, FTLD_DE, VCP_DE))
print(overlap)
```

I re ran the original data again in WenBin DE Gene.R and got completely different results. This is the version of the script I used: `##Differential Expression of Genes##`

```
setwd("/Users/clairegreen/Documents/PhD/TDP-43/TDP-43_Data/DE Genes/FTLD/")
library(affy)
library(Biobase)
library(tkWidgets)

#run program to choose .CEL files from directory
celfiles <- fileBrowser(textToShow = "Choose CEL files", testFun = hasSuffix("[cC][eE][lL]"))
#celfiles<-basename(celfiles)
Data<-ReadAffy(filename=celfiles) #read in files
rmaEset<-rma(Data) #normalise using RMA
#analysis.name<-"CHMP2B" #Label analysis
dataMatrixAll<-exprs(rmaEset) #takes expression from normalised expression set

#mas5call generates presence/absence calls for each probeset
mas5call<-mas5calls(Data)
callMatrixAll<-exprs(mas5call)
colnames(callMatrixAll)<-sub(".CEL", ".mas5-Detection", colnames(callMatrixAll),fixed=TRUE)
colnames(callMatrixAll)<-sub(".cel", ".mas5-Detection", colnames(callMatrixAll),fixed=TRUE)
callMatrixAll<-as.data.frame(callMatrixAll)
callMatrixAll$ProbeSetID<-rownames(callMatrixAll)
countPf<-function(x){
  sum(x=="P")
}

#count how many samples have presence calls
countPl<-apply(callMatrixAll, 1, countPf)
```

```

callMatrixAll$ProbeSetID<-rownames(callMatrixAll)
countPdf<-data.frame(ProbeSetID=names(countP1), countP=countP1)

#read annotation file

###USING BIOMART
# library (biomaRt)
# mart <- useMart("ENSEMBL_MART_ENSEMBL",dataset="hsapiens_gene_ensembl", host="www.ensembl.org")
# x <- rownames(dataMatrixAll) #create vector containing probe IDs
# mart_attribute <- listAttributes(mart)
# annotation <- getBM(attributes=c("affy_hg_u133a_2", "hgnc_symbol", "description"),
#                       filters = "affy_hg_u133a_2", values = x, mart = mart)
# annotation<-subset(annotation, subset=(hgnc_symbol !="")) #if no gene symbol, discount

#USING ANNOTATION FILE (if .csv, convert to .txt using excel)
#annotation.file<-"/Users/clairegreen/Documents/PhD/TDP-43/TDP-43_Data/HG-U133_Plus_2.na35.annot.csv/HG-U133A_2.na35.annot.csv"
annotation.file<-"/Users/clairegreen/Documents/PhD/TDP-43/TDP-43_Data/HG-U133A_2.na35.annot.csv/HG-U133A_2.na35.annot.txt"

annotation<-read.table(annotation.file, header = TRUE, row.names=NULL, sep="\t", skip=0, stringsAsFactors=FALSE)
dim(annotation)
nrow(annotation)
#[1] 39699
annotation<-subset( annotation, subset=(Gene.Symbol !="---")) #if no gene symbol, discount
nrow(annotation)

expressionMatrix<-exprs(rmaEset)
colnames(expressionMatrix)

#this is for matched samples
#tonsil<-factor(c("T101","T102","T103","T104","T105","T106"))
Treat<-factor(rep(c("Control", "Patient"),c(8,16)), levels=c("Control", "Patient"))
design<-model.matrix(~Treat)
rownames(design)<-colnames(expressionMatrix)
design

#Conduct statistical analysis of expression
library(limma)
fit<-lmFit(expressionMatrix, design) #linear model fit
fit<-eBayes(fit)
result<-topTable(fit, coef="TreatPatient", adjust="BH", number=nrow(expressionMatrix)) #BH adjust for multiple testing
#topTable normally takes top number but this takes all

result$"ProbeSetID"<-rownames(result)
head(result$"ProbeSetID")
result$"Fold Change"<-2^result$logFC
result$"Fold Change"[result$"Fold Change"<1]<-(-1)/result$"Fold Change"[result$"Fold Change"<1] #convert log2 to log10
expressionLinear<-as.data.frame(2^expressionMatrix)
expressionLinear$ProbeSetID<-rownames(expressionLinear)
result<-merge(result, expressionLinear, by.x="ProbeSetID", by.y="ProbeSetID") #merge values into one array
result<-merge(annotation, result, by.x="Probe.Set.ID", by.y="ProbeSetID")
#result<-merge(result, callMatrixAll, by.x="Probe_Set_ID", by.y="ProbeSetID")
result<-merge(result, countPdf, by.x="Probe.Set.ID", by.y="ProbeSetID")

```

```

# # write.table(result, file=paste(analysis.name, "anno RMA limma na35.txt", sep=""), sep="\t", row.names=rownames(result))

# result<-subset(result, Gene.Symbol!="")
# result<-subset(result, subset=(countP>2))
# nrow(result)
# foldchange<-1.5
# pvalue<-0.05
# #adj_P_Val<-0.05
# siggenes<-subset(result, subset=(P.Value < pvalue) & abs(logFC) > log2(foldchange))
# #siggenes<-subset(result, subset=(adj.P.Val < 0.05))
# nrow(siggenes)
# siggenesup<-subset(siggenes, subset= logFC > 0)
# siggenesdown<-subset(siggenes, subset=logFC < 0)
# colnames(siggenesup)
# nrow(siggenesup)
# nrow(siggenesdown)
# UpandDown<-intersect(siggenesup$"Gene.Symbol", siggenesdown$"Gene.Symbol")
# length(UpandDown)
#
# UporDown<-subset(siggenes, subset=(!siggenes$"Gene.Symbol"%in% UpandDown))
# upsiggenes<-subset(siggenesup, subset=(!siggenesup$"Gene.Symbol"%in% UpandDown))
# downsiggenes<-subset(siggenesdown, subset=(!siggenesdown$"Gene.Symbol"%in% UpandDown))
# length(unique(siggenes$"Gene.Symbol"))
# uniquesiggenes <- unique(siggenes$Gene.Symbol)
# length(unique(upsiggenes$"Gene.Symbol"))
# length(unique(downsiggenes$"Gene.Symbol"))
# length(unique(UporDown$"Gene.Symbol"))

###Write results to CSV files for consensus analysis
setwd("/Users/clairegreen/Documents/PhD/TDP-43/TDP-43_Data/DE Genes/TopGenesb_2016-02-09")
#dir.create(paste("TopGenesb", Sys.Date(), sep = "_")) #create directory using the day's date
#Take results, remove duplicate rows for genes, order by adjusted p value and take top X number of genes
uniquerresult <- result[!duplicated(result[,15]),]

#For ordering by adjusted p value

genesort <- uniquerresult[order(uniquerresult$adj.P.Val),]
topgene <- genesort[1:1000,]
write.csv(x = topgene, file = "FTLD_ap_1000")
topgene <- genesort[1:2000,]
write.csv(x = topgene, file = "FTLD_ap_2000")
topgene <- genesort[1:3000,]
write.csv(x = topgene, file = "FTLD_ap_3000")
topgene <- genesort[1:4000,]
write.csv(x = topgene, file = "FTLD_ap_4000")
topgene <- genesort[1:5000,]
write.csv(x = topgene, file = "FTLD_ap_5000")

#For ordering by fold change
# genesort <- uniquerresult[order(uniquerresult$`Fold Change`),]
# topgene <- genesort[1:500,]
#

```

```
# genesort <- uniqueresult[order(-uniqueresult$`Fold Change`),]
# botgene <- genesort[1:500,]
#
# topFC <- rbind(topgene,botgene)
#
# write.csv(x = topgene, file = "VCP_fc_1000")

# dir.create(paste("/Users/clairegreen/Documents/PhD/TDP-43/TDP-43 Data Sets/DE Genes/CHMP2B/unique"))
# setwd("/Users/clairegreen/Documents/PhD/TDP-43/TDP-43 Data Sets/DE Genes/CHMP2B/unique")
#
# write.table(siggenes, file=paste(analysis.name," RMA limma siggenes biomart p_", pvalue, "_fold chang
# write.table(upsiggenes, file=paste(analysis.name," RMA limma upsiggenes biomart p_", pvalue, "_fold c
# write.table(downsiggenes, file=paste(analysis.name," RMA limma downsiggenes biomart p_", pvalue, "_fo
# write.table(uniquesiggenes, file=paste(analysis.name,"unique RMA limma siggenes biomart p_", pvalue,
```

From this I generated the following results

1000

0

2000

0

3000

0

4000

MAP4K4

5000

MRPL23 MAP4K4 METTL17

Wednesday

As a result, I went back to the closest backup of the file I had (Dec 16th 2015) and the only discernible difference in Wenbin DE Gene.R was that I had commented out two lines that may be important.

```
result<-subset(result, Gene.Symbol!="") #removes any probes for which there are no gene symbols
result<-subset(result, subset=(countP>2)) #only takes results that have at least 2 samples with a prese
```

There were no differences to the consensus script. I then re-calculated expression and consensus was identical to that generated in Febuary from the December results.

Talking to Wenbin

In my discussion with Wenbin, we talked first about me removing duplicated genes before taking the top X DEGs. He at first was sceptical about removing duplicates because you couldn't be sure that you weren't left with 'bad' probes, but I explained that the only probes accepted at this point were those with 3 or more samples with a presence call (which suggests good quality).

Next, we discussed again the difference between ranking by p value or fold change, but we came to the conclusion that it's up to the individual and the project. I'm going to continue my experiment if only out of interest.

When it came to validating the results of the consensus DEGs, Wenbin said that both random permutations test and Fisher's exact test could be ways to approach it. For the RPT, I wanted his opinion on what my population pool should be for selecting random genes. We settled on using the annotation file, removing blanks, removing duplicates and also removing any probes noted to bind to negative strand matching probes. NSMPs are expressed sequence tags that have been designed in the wrong direction to the gene and so are antisense.

The random permutation test including these changes are below. Note that I have used the annotation files from both the U133 plus 2.0 and U133A to accurately represent the fact we have 4 data sets from the former and 1 from the latter.

```
## Save annotation file locations to variable
annotation.U133plus2 <- "/Users/clairegreen/Documents/PhD/TDP-43/TDP-43_Data/HG-U133_Plus_2.na35.annot.csv"
annotation.U133A2 <- "/Users/clairegreen/Documents/PhD/TDP-43/TDP-43_Data/HG-U133A_2.na35.annot.csv/HG-U133A_2.na35.annot.csv"

# Read in annotation files
annotationU133plus2 <- read.table(annotation.U133plus2, header = TRUE,
  row.names = NULL, sep = "\t", skip = 0, stringsAsFactors = F,
  quote = "", comment.char = "!", fill = TRUE)
annotationU133A2 <- read.table(annotation.U133A2, header = TRUE,
  row.names = NULL, sep = "\t", skip = 0, stringsAsFactors = F,
  quote = "", comment.char = "!", fill = TRUE)

# Remove rows in which gene symbol is absent
annotationU133plus2 <- subset(annotationU133plus2, subset = (Gene.Symbol !=
  "---")) #if no gene symbol, discount
annotationU133A2 <- subset(annotationU133A2, subset = (Gene.Symbol !=
  "---")) #if no gene symbol, discount

# Remove rows in which gene symbol is duplicated
annotationU133plus2 <- annotationU133plus2[!duplicated(annotationU133plus2[,
  15]), ]
annotationU133A2 <- annotationU133A2[!duplicated(annotationU133A2[,
  15]), ]

# Remove rows in which genes are noted to have negative
# strand matching probes
idxNegativeStrand <- grep("Negative Strand Matching Probes",
  annotationU133plus2$Annotation.Notes)
if (length(idxNegativeStrand) > 0) {
  annotationU133plus2 <- annotationU133plus2[-idxNegativeStrand,
  ]
}
nrow(annotationU133plus2)
```



```

idxNegativeStrand <- grep("Negative Strand Matching Probes",
  annotationU133A2$Annotation.Notes)
if (length(idxNegativeStrand) > 0) {
  annotationU133A2 <- annotationU133A2[-idxNegativeStrand,
    ]
}
nrow(annotationU133A2)

# indicate the number of overlapping genes identified by DE
# analysis
test <- 50

m = 10000 #number of repetitions
r <- c(1:m) #store repetition numbers in vector 'r'

for (j in 1:m) {
  random1 <- sample(annotationU133plus2$Gene.Symbol, size = 4000,
    replace = F)
  random2 <- sample(annotationU133plus2$Gene.Symbol, size = 4000,
    replace = F)
  random3 <- sample(annotationU133plus2$Gene.Symbol, size = 4000,
    replace = F)
  random4 <- sample(annotationU133plus2$Gene.Symbol, size = 4000,
    replace = F)
  random5 <- sample(annotationU133A2$Gene.Symbol, size = 4000,
    replace = F)
  random <- Reduce(intersect, list(random1, random2, random3,
    random4, random5))
  r[j] <- length(random)
}

test1 <- which(r > test) # count number of times r is larger than test value
result <- (length(test1)/m) # calculate P value

```

For both 16 (consensus value for top 3000) and 50 (consensus value for top 4000), $p < 0.01$

Next, I conducted the same analysis but using the most extreme fold change values. FC was sorted by number and the top and bottom genes were selected to create 1000, 2000, 3000, and 4000 gene sets. These were analysed for common genes.

1000

NOV

2000

MCTP1 TUBB4B NOV HN1 NAP1L2 TSPAN13 TMEM255A DMD RGS2 CDK5R1 HBA1 /// HBA2 ZIC1

3000

MCTP1 TUBB4B NOV SPP1 HN1 UPF3A NAP1L2 TSPAN13 TMEM255A DMD ESRRG RGS2 CDK5R1
ATP6V1A PVALB RTN1 NPTX2 COL4A1 HBA1 /// HBA2 ITM2A TJP2 RND3 ZIC1 MAN2A1 DCN
ATP2A2 ADAMTS1 MAP1A

4000

MCTP1 EIF1AY TUBB3 TUBB4B NOV RNF13 SPP1 YWHAH WDR47 MLF2 NEK7 KIF3A HN1 NMRK1
UPF3A GSTO1 CCT2 NAP1L2 TMEM97 PHACTR2 TSPAN13 TMEM255A DMD ESRRG RGS2 CDK5R1
ATP6V1A STAT3 BTN3A2 /// BTN3A3 PVALB ZFPM2 P4HTM VDAC3 SLC25A12 RAB15 RTN1 NQO2
IL6ST PEG3 LSM5 UCHL1 KIFAP3 ZNF271 NUPL1 CEBPD NPTX2 IFIT3 COL4A1 HBA1 /// HBA2
RGCC DDIT4 PLEKHB1 ITM2A CNKSR2 ORAI2 ADARB1 MAN2A1 MAP1A CSTA DCN SLC14A1
ATP2A2 EFS TJP2 ADAMTS1 SCN1B RND3 KATNB1 ETNK1 ZIC1 TRIM22 WSB2

The first thing to note is that when using fold change to rank genes, there are more common genes. The second thing to note is that the genes TUBB3, RNF13, UPF3A, GSTO1, CCT2, DMD, and CDK5R1 are common to both lists (top 3000 ap, top 4000 fc). However, one has to bear in mind that taking equal quantities of up and downregulated genes does not necessarily give the full picture. For example, a diseased condition may have proportionally more upregulation of genes than down, but by taking 500 from each end you are treating them as equal. By using adjusted p value, you are taking the most DE regardless of direction, which may be more useful.

Thursday

Now I have a list of 50 DEGs associated with TDP-43 pathology, I wanted to find out a little bit more about them. I first used WebGestalt to get some information on the genes themselves.

4000	Full Name
DDX39B	DEAD (Asp-Glu-Ala-Asp) box polypeptide 39B
BPTF	bromodomain PHD finger transcription factor
BRD3	bromodomain containing 3
C14orf1	chromosome 14 open reading frame 1
C18orf32	chromosome 18 open reading frame 32
CCT2	chaperonin containing TCP1, subunit 2 (beta)
CDK5R1	cyclin-dependent kinase 5, regulatory subunit 1 (p35)
CSRP1	cysteine and glycine-rich protein 1
CST3	cystatin C
DCN	decorin
DDX39A	DEAD (Asp-Glu-Ala-Asp) box polypeptide 39A
DDX5	DEAD (Asp-Glu-Ala-Asp) box helicase 5
DMD	dystrophin
DYNLT1	dynein, light chain, Tctex-type 1
EEF1A1	eukaryotic translation elongation factor 1 alpha 1
ETS2	v-ets erythroblastosis virus E26 oncogene homolog 2 (avian)
FN1	fibronectin 1
GNPAT	glyceronephosphate O-acyltransferase
GSTO1	glutathione S-transferase omega 1
GTF2I	general transcription factor Iii
HSD17B4	hydroxysteroid (17-beta) dehydrogenase 4
ICMT	isoprenylcysteine carboxyl methyltransferase
LBR	lamin B receptor

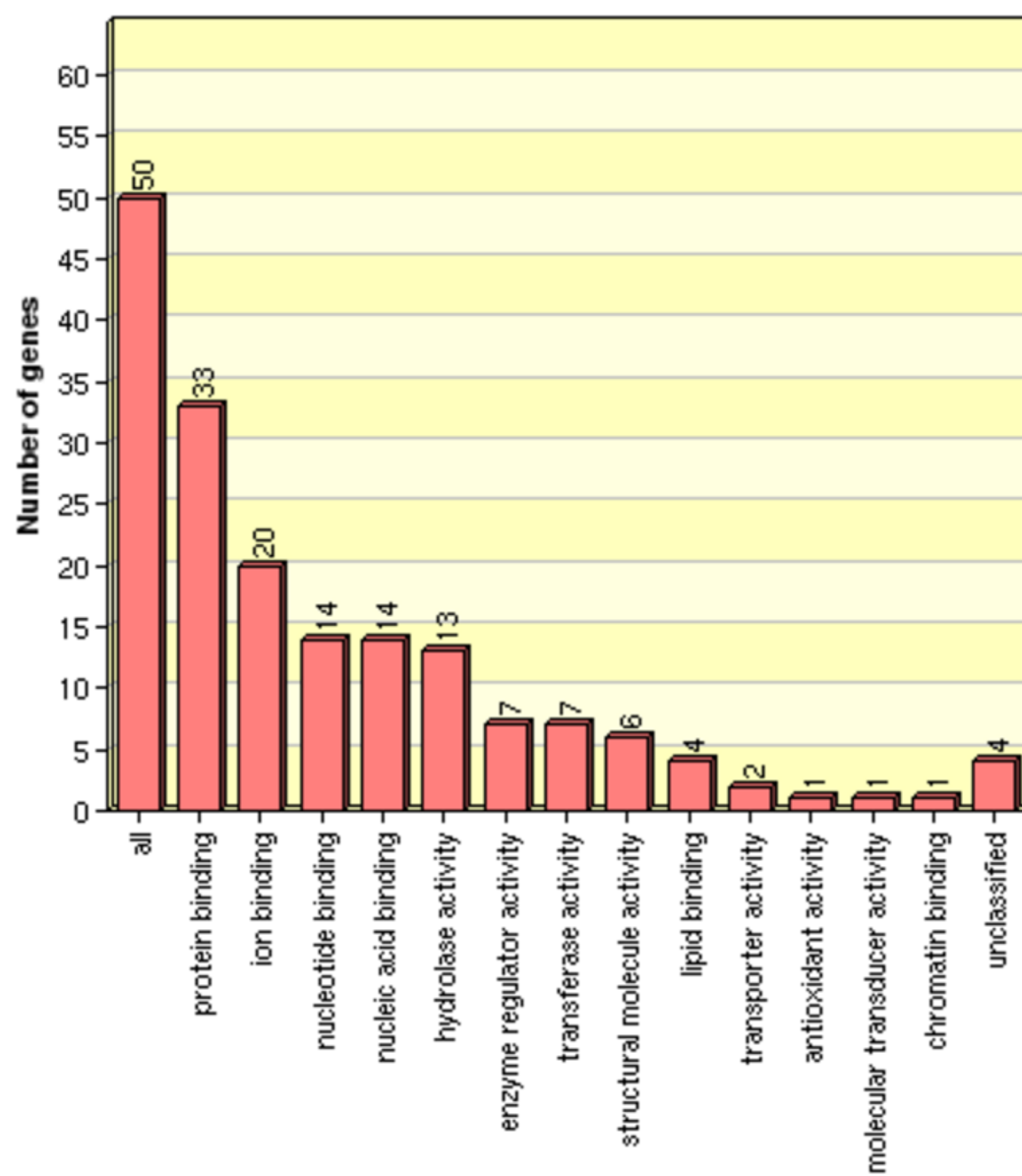
4000	Full Name
MAP4K4	mitogen-activated protein kinase kinase kinase kinase 4
MPHOSPH9	M-phase phosphoprotein 9
MTMR9	myotubularin related protein 9
NAGA	N-acetylgalactosaminidase, alpha-
NKTR	natural killer-tumor recognition sequence
NUTF2	nuclear transport factor 2
PAICS	phosphoribosylaminoimidazole carboxylase
PFDN1	prefoldin subunit 1
PRKD1	protein kinase D1
PSAP	prosaposin
PSMD1	proteasome (prosome, macropain) 26S subunit, non-ATPase, 1
PTEN	phosphatase and tensin homolog
RECQL	RecQ protein-like (DNA helicase Q1-like)
RNF13	ring finger protein 13
RNF130	ring finger protein 130
RPL6	ribosomal protein L6
RPLP2	ribosomal protein, large, P2
RPS6	ribosomal protein S6
TARDBP	TAR DNA binding protein
TARS	threonyl-tRNA synthetase
TCF4	transcription factor 4
TMEM59	transmembrane protein 59
TRO	trophinin
TUBB3	tubulin, beta 3 class III
UPF3A	UPF3 regulator of nonsense transcripts homolog A
WBSR22	Williams Beuren syndrome chromosome region 22
WDR78	WD repeat domain 78

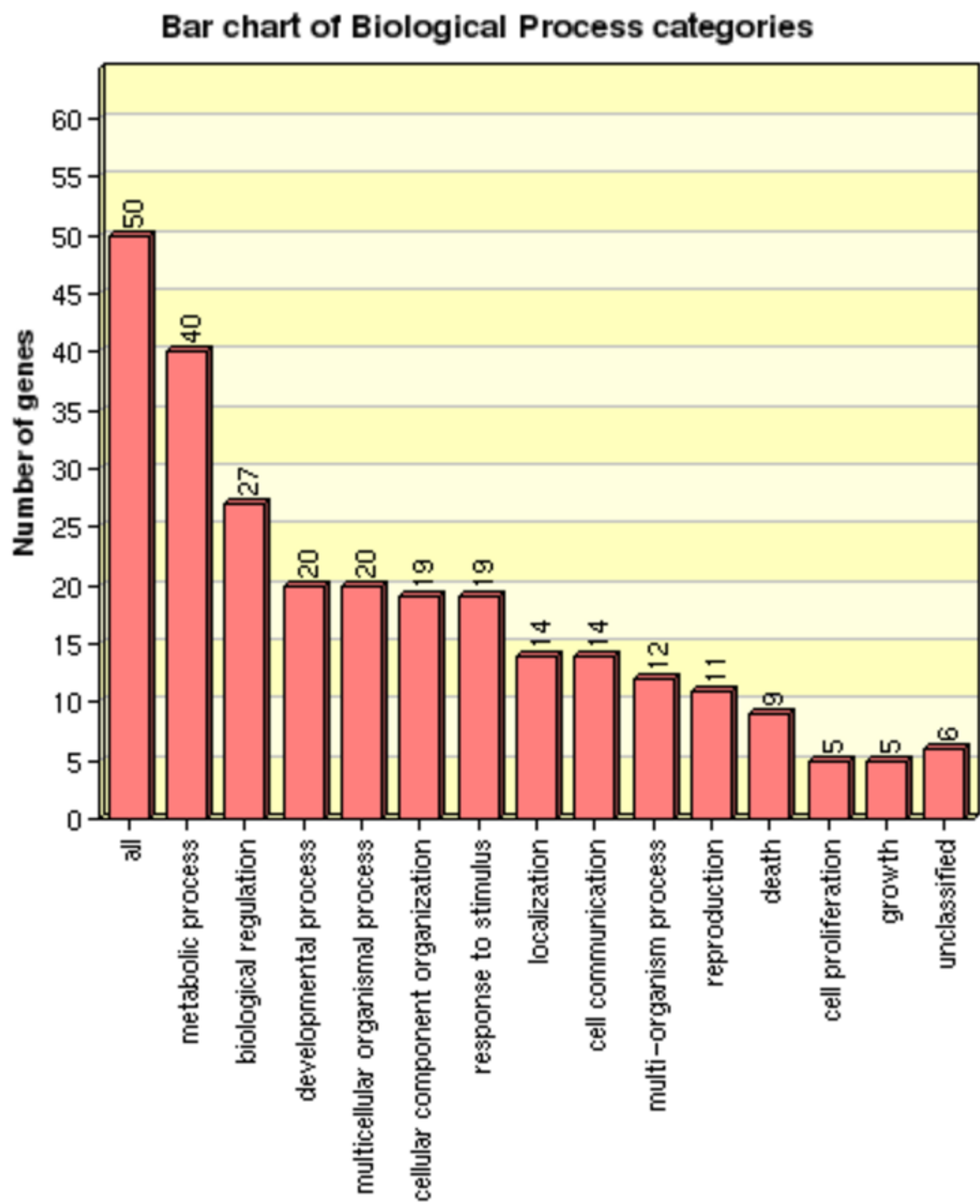
4000	Associated Diseases (WebGestalt)	Associated Drugs (WebGestalt)
DDX39B		Dactinomycin, Adenosine
BPTF	Dementia, Alzheimer's Disease	Adenosine
BRD3		
C14orf1		
C18orf32		
CCT2		
CDK5R1	Dementia, Alzheimer's Disease, Brain Diseases, Tauopathies	
CSRP1		
CST3	Dementia, Alzheimer's Disease, Brain Diseases, Tauopathies	
DCN		
DDX39A		Dactinomycin
DDX5		Dactinomycin, Adenosine
DMD	Mental Retardation, Monosomy	
DYNLT1		
EEF1A1		
ETS2	Chordoma, Mental Retardation	
FN1		
GNPAT	Mental Retardation	
GSTO1	Dementia, Alzheimer's Disease, Brain Diseases, Tauopathies	
GTF2I	Breast/Ovarian Cancer Syndrome, Mental Retardation, Monosomy	
HSD17B4	Brain Diseases	
ICMT		Adenosine

4000	Associated Diseases (WebGestalt)	Associated Drugs (WebGestalt)
LBR		
MAP4K4		
MPHOSPH9		
MTMR9		
NAGA	Brain Diseases, Sandhoff Disease	
NKTR		
NUTF2		
PAICS		
PFDN1		
PRKD1		
PSAP	Brain Diseases, Sandhoff Disease	
PSMD1		Adenosine
PTEN	Chordoma, Breast/Ovarian Cancer Syndrome	
RECQL		Adenosine
RNF13		
RNF130		
RPL6		Dactinomycin
RPLP2		
RPS6		
TARDBP	Dementia, Alzheimer's Disease, Brain Diseases, Tauopathies	
TARS		
TCF4	Mental Retardation, Monosomy	
TMEM59		
TRO		
TUBB3		
UPF3A		
WBSCR22	Mental Retardation, Monosomy	
WDR78		

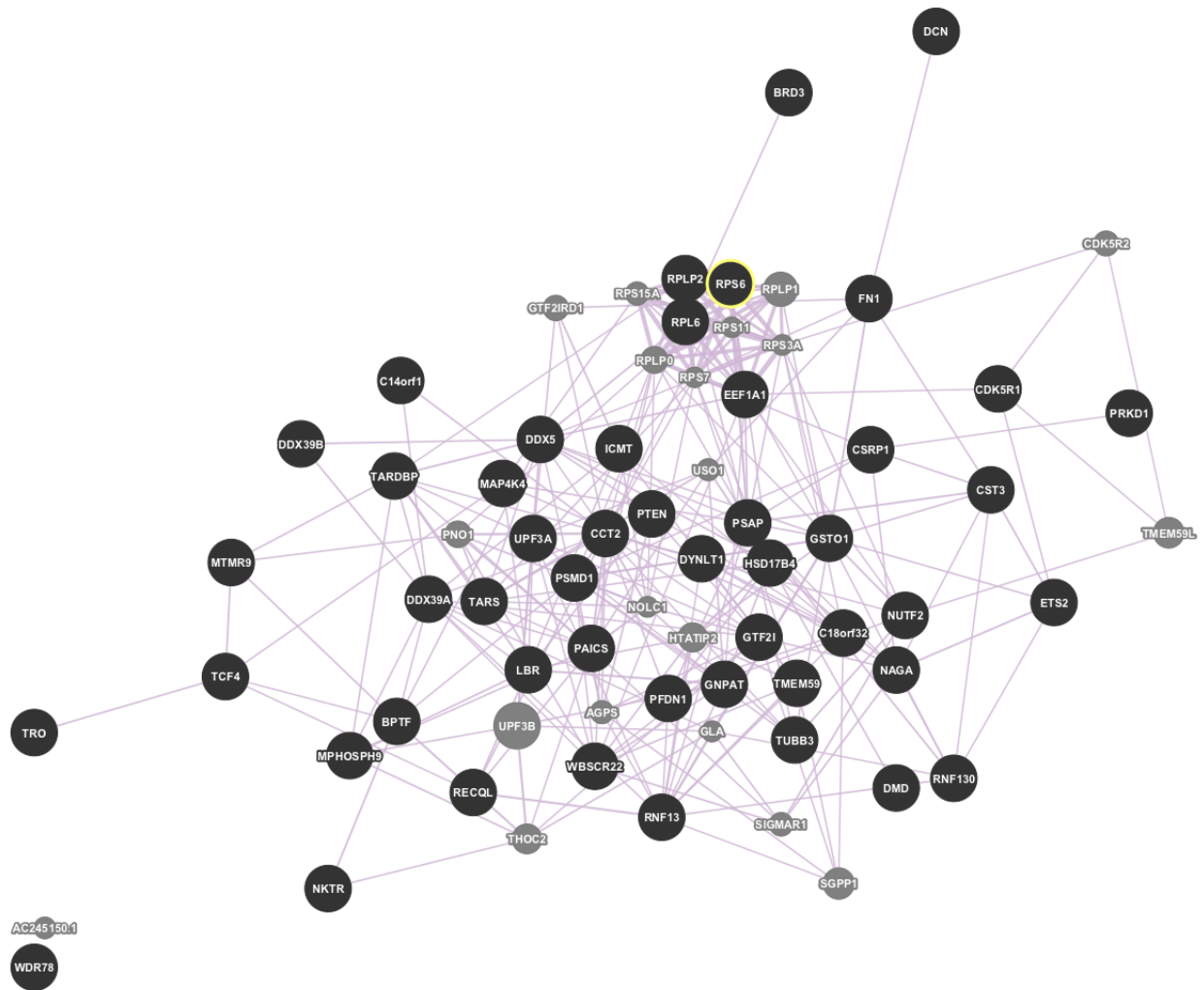
I also generated graphs showing GO biological processes and molecular functions

Bar chart of Molecular Function categories





The next thing I did was input the genes into GeneMANIA to look at the co-expression network.



There is a group of 9 genes (RPS15A, RPLP2, RPS6, RPLP1, RPL6, RPS11, RPS3A, RPLP0 and RPS7) that is very tightly connected. These genes are found in the top 29 functions listed in the table below. In one of these categories, viral transcription, these RPLP2 is significantly co-expressed with TARDBP.

Feature	FDR	Genes in network	Genes in genome
viral life cycle	4.32E-09	12	159
nuclear-transcribed mRNA catabolic process, nonsense-mediated decay	4.32E-09	11	116
translational elongation	2.10E-08	10	102
cytosolic ribosome	1.37E-07	9	90
translational termination	1.37E-07	9	89
nuclear-transcribed mRNA catabolic process	1.37E-07	11	179
viral gene expression	1.45E-07	9	94
mRNA catabolic process	1.96E-07	11	190
multi-organism metabolic process	2.84E-07	9	104
SRP-dependent cotranslational protein targeting to membrane	2.97E-07	9	107
protein targeting to membrane	2.97E-07	10	157
protein targeting to ER	2.97E-07	9	108
cotranslational protein targeting to membrane	2.97E-07	9	108
establishment of protein localization to endoplasmic reticulum	2.97E-07	9	109
single-organism cellular localization	2.97E-07	12	272
single-organism localization	2.97E-07	12	272
viral transcription	3.16E-07	10	159
RNA catabolic process	3.28E-07	11	215
protein localization to endoplasmic reticulum	4.88E-07	9	120
ribosomal subunit	5.00E-07	9	121
cellular protein complex disassembly	1.27E-06	9	135
establishment of protein localization to membrane	1.50E-06	11	253
translational initiation	2.94E-06	9	150
cytosolic part	2.99E-06	9	151
ribosome	3.61E-06	9	155
protein complex disassembly	3.67E-06	9	156
macromolecular complex disassembly	5.20E-06	9	163
protein localization to membrane	6.08E-06	10	227
protein targeting	7.89E-05	10	299
cytosolic small ribosomal subunit	1.20E-04	5	36
structural constituent of ribosome	8.64E-04	6	97
small ribosomal subunit	9.85E-04	5	55
cytosolic large ribosomal subunit	1.97E-02	4	51
sphingolipid metabolic process	1.97E-02	5	102
nuclear export	3.09E-02	5	113
nucleocytoplasmic transport	3.51E-02	7	278
nuclear transport	3.66E-02	7	281
large ribosomal subunit	5.18E-02	4	67
mRNA export from nucleus	5.35E-02	4	68
mRNA transport	5.52E-02	4	69
membrane lipid metabolic process	7.27E-02	5	140
RNA export from nucleus	7.66E-02	4	76
regulation of neurogenesis	7.75E-02	6	227
multi-organism intracellular transport	8.61E-02	3	32
multi-organism transport	8.61E-02	3	32
intracellular transport of virus	8.61E-02	3	32
transport of virus	8.61E-02	3	32
microbody lumen	9.06E-02	3	33
peroxisomal matrix	9.06E-02	3	33
establishment of RNA localization	9.35E-02	4	85
nucleic acid transport	9.35E-02	4	85
RNA transport	9.35E-02	4	85
RNA localization	9.35E-02	4	84

List of

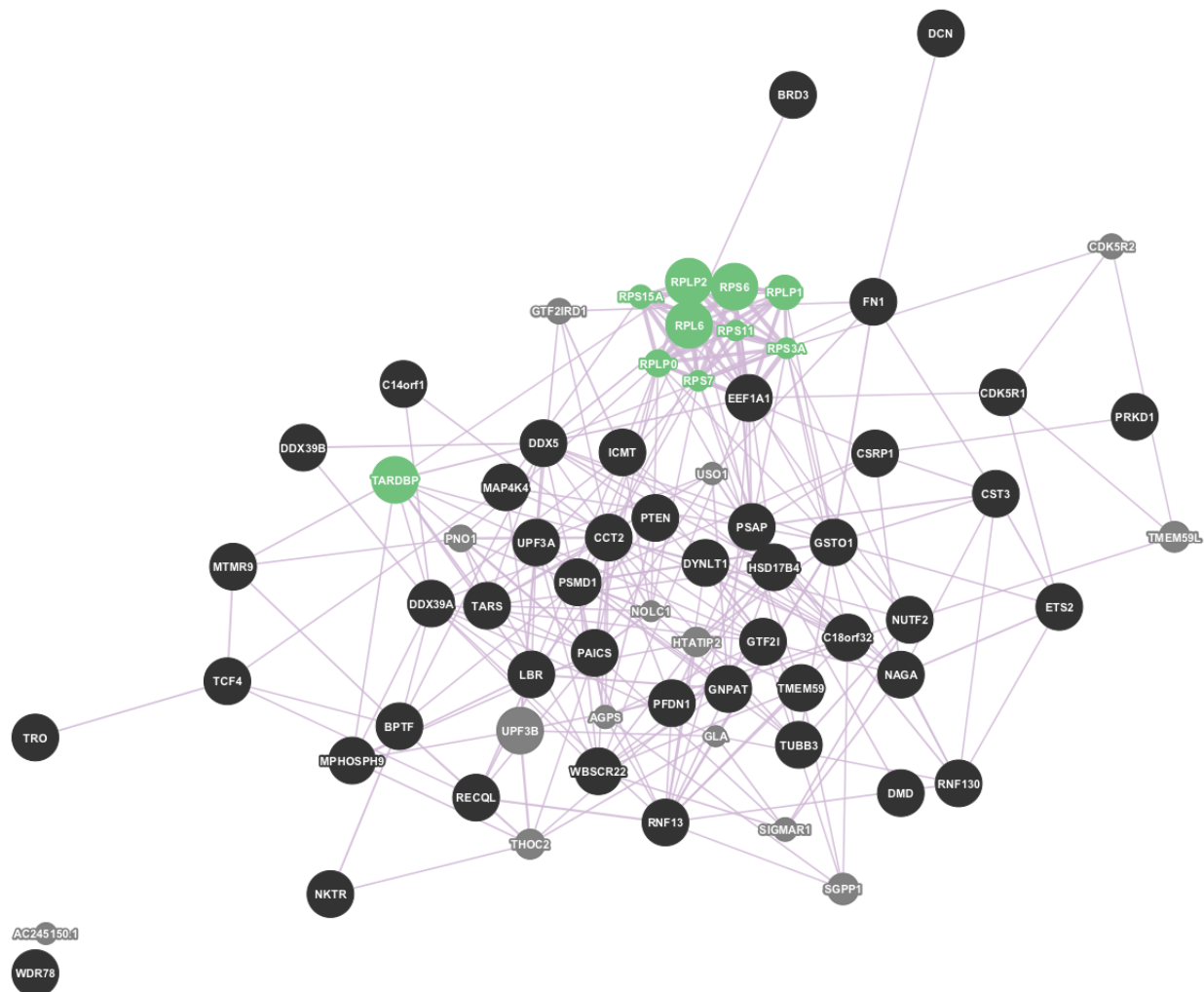
functions

4000	Gene Mania Functions
BPTF	
BRD3	
C14orf1	
C18orf32	
CCT2	
CDK5R1	Regulation of neurogenesis
CSR1	
CST3	
DCN	
DDX39A	Nuclear export, nucleocytoplasmic transport, nuclear transport, mRNA transport from nucleus, mRNA transport, RNA transport from nucleus, establishment of RNA localization, nucleic acid transport, RNA transport, RNA localization
DDX39B	Viral life cycle, nuclear export, nucleocytoplasmic transport, nuclear transport, mRNA transport from nucleus, mRNA transport, multi-organism intracellular transport, multi-organism transport, intracellular transport of virus, transport of virus, establishment of RNA localisation, nucleic acid transport, RNA transport, RNA localisation
DDX5	
DMD	Regulation of neurogenesis
DYNLT1	Viral life cycle, single-organism cellular localization, single-organism localization, regulation of neurogenesis, multi-organism intracellular transport, multi-organism transport, intracellular transport of virus, transport of virus
EEF1A1	Translational elongation
ETS2	
FN1	
GNPAT	Microbody lumen, peroxisomal matrix
GSTO1	
GTF2I	
HSD17B4	Microbody lumen, peroxisomal matrix
ICMT	Protein targeting to membrane, single-organism cellular localization, single-organism localization, establishment of protein localization to membrane, protein localization to membrane, protein targeting
LBR	
MAP4K4	
MPHOSPH9	
MTMR9	
NAGA	Sphingoloid metabolic process, membrane lipid metabolic process
NKTR	
NUTF2	Nuclear export, nucleocytoplasmic transport, nuclear transport
PAICS	
PFDN1	
PRKD1	Sphingoloid metabolic process, membrane lipid metabolic process, regulation of neurogenesis
PSAP	Sphingoloid metabolic process, membrane lipid metabolic process
PSMD1	
PTEN	Single-organism cellular localization, single-organism localization, regulation of neurogenesis
RECQL	
RNF13	
RNF130	
RPL6	Viral life cycle, nuclear-transcribed mRNA catabolic process-nonsense-mediated delay, translational elongation, cytosolic ribosome, translational termination, nuclear transcribed mRNA catabolic process, viral gene expression, mRNA catabolic process, multi-organism metabolic process, SRP-dependent cotranslational protein targeting to membrane, protein targeting to membrane, protein targeting to ER, cotranslational protein targeting to membrane, establishment of protein localization to endoplasmic reticulum, single-organism cellular localization, single organism localisation, viral transcription, RNA catabolic process, protein localisation to endoplasmic reticulum, ribosomal subunit, establishment of protein localisation to membrane, translational initiation, cytosolic part, ribosome, protein complex disassembly, macromolecular complex disassembly, protein localization to membrane, protein targeting, structural constituent of ribosome, cytosolic large ribosomal subunit, large ribosomal subunit
RPLP2	Viral life cycle, nuclear-transcribed mRNA catabolic process-nonsense-mediated delay, translational elongation, cytosolic ribosome, translational termination, nuclear transcribed mRNA catabolic process, viral gene expression, mRNA catabolic process, multi-organism metabolic process, SRP-dependent cotranslational protein targeting to membrane, protein targeting to membrane, protein targeting to ER, cotranslational protein targeting to membrane, establishment of protein localization to endoplasmic reticulum, single-organism cellular localization, single organism localisation, viral transcription, RNA catabolic process, protein localisation to endoplasmic reticulum, ribosomal subunit, establishment of protein localisation to membrane, translational initiation, cytosolic part, ribosome, protein complex disassembly, macromolecular complex disassembly, protein localization to membrane, protein targeting, structural constituent of ribosome, cytosolic large ribosomal subunit, large ribosomal subunit
RPS6	Viral life cycle, nuclear-transcribed mRNA catabolic process-nonsense-mediated delay, translational elongation, cytosolic ribosome, translational termination, nuclear transcribed mRNA catabolic process, viral gene expression, mRNA catabolic process, multi-organism metabolic process, SRP-dependent cotranslational protein targeting to membrane, protein targeting to membrane, protein targeting to ER, cotranslational protein targeting to membrane, establishment of protein localization to endoplasmic reticulum, single-organism cellular localization, single organism localisation, viral transcription, RNA catabolic process, protein localisation to endoplasmic reticulum, ribosomal subunit, establishment of protein localisation to membrane, translational initiation, cytosolic part, ribosome, protein complex disassembly, macromolecular complex disassembly, protein localization to membrane, protein targeting, cytosolic small ribosomal subunit, small ribosomal subunit
TARDBP	Viral transcription
TARS	
TCF4	Regulation of neurogenesis
TMEM59	Establishment of protein localization to membrane
TRO	
TUBB3	
UPF3A	nuclear-transcribed mRNA catabolic process-nonsense-mediated delay, nuclear transcribed mRNA catabolic process, mRNA catabolic process, RNA catabolic process, nucleocytoplasmic transport, nuclear transport.
WBSR22	
WDR78	

genes

With





Viral transcription

Next I looked at G:profiler. This produced a protein-protein interaction network using BioGRID. The high quality PDF can be found on basecamp or at [/Users/clairegreen/Documents/PhD/TDP-43/TDP-43_Data/GeneExpressionAnalysis/G.Profiler output](#). G:Profiler also produced results from multiple places including GO biological process and cellular component, as well as Human protein atlas and Reactome biological pathways.

source	term name Gene Ontology (Biological process)	term ID	n. of term genes	n. of query genes	n. of common genes	n. of corrected p-value	BRD3	BRD4	CCT2	CCT3	CCT4	CCT5	CCT6	CCT7	CCT8	CCT9	CCT10	CCT11	CCT12	CCT13	CCT14	CCT15	CCT16	CCT17	CCT18	CCT19	CCT20	CCT21	CCT22	CCT23	CCT24	CCT25	CCT26	CCT27	CCT28	CCT29	CCT30	CCT31	CCT32	CCT33	CCT34	CCT35	CCT36	CCT37	CCT38	CCT39	CCT40	CCT41	CCT42	CCT43	CCT44	CCT45	CCT46	CCT47	CCT48	CCT49	CCT50	CCT51	CCT52	CCT53	CCT54	CCT55	CCT56	CCT57	CCT58	CCT59	CCT60	CCT61	CCT62	CCT63	CCT64	CCT65	CCT66	CCT67	CCT68	CCT69	CCT70	CCT71	CCT72	CCT73	CCT74	CCT75	CCT76	CCT77	CCT78	CCT79	CCT80	CCT81	CCT82	CCT83	CCT84	CCT85	CCT86	CCT87	CCT88	CCT89	CCT90	CCT91	CCT92	CCT93	CCT94	CCT95	CCT96	CCT97	CCT98	CCT99	CCT100	CCT101	CCT102	CCT103	CCT104	CCT105	CCT106	CCT107	CCT108	CCT109	CCT110	CCT111	CCT112	CCT113	CCT114	CCT115	CCT116	CCT117	CCT118	CCT119	CCT120	CCT121	CCT122	CCT123	CCT124	CCT125	CCT126	CCT127	CCT128	CCT129	CCT130	CCT131	CCT132	CCT133	CCT134	CCT135	CCT136	CCT137	CCT138	CCT139	CCT140	CCT141	CCT142	CCT143	CCT144	CCT145	CCT146	CCT147	CCT148	CCT149	CCT150	CCT151	CCT152	CCT153	CCT154	CCT155	CCT156	CCT157	CCT158	CCT159	CCT160	CCT161	CCT162	CCT163	CCT164	CCT165	CCT166	CCT167	CCT168	CCT169	CCT170	CCT171	CCT172	CCT173	CCT174	CCT175	CCT176	CCT177	CCT178	CCT179	CCT180	CCT181	CCT182	CCT183	CCT184	CCT185	CCT186	CCT187	CCT188	CCT189	CCT190	CCT191	CCT192	CCT193	CCT194	CCT195	CCT196	CCT197	CCT198	CCT199	CCT200	CCT201	CCT202	CCT203	CCT204	CCT205	CCT206	CCT207	CCT208	CCT209	CCT210	CCT211	CCT212	CCT213	CCT214	CCT215	CCT216	CCT217	CCT218	CCT219	CCT220	CCT221	CCT222	CCT223	CCT224	CCT225	CCT226	CCT227	CCT228	CCT229	CCT230	CCT231	CCT232	CCT233	CCT234	CCT235	CCT236	CCT237	CCT238	CCT239	CCT240	CCT241	CCT242	CCT243	CCT244	CCT245	CCT246	CCT247	CCT248	CCT249	CCT250	CCT251	CCT252	CCT253	CCT254	CCT255	CCT256	CCT257	CCT258	CCT259	CCT260	CCT261	CCT262	CCT263	CCT264	CCT265	CCT266	CCT267	CCT268	CCT269	CCT270	CCT271	CCT272	CCT273	CCT274	CCT275	CCT276	CCT277	CCT278	CCT279	CCT280	CCT281	CCT282	CCT283	CCT284	CCT285	CCT286	CCT287	CCT288	CCT289	CCT290	CCT291	CCT292	CCT293	CCT294	CCT295	CCT296	CCT297	CCT298	CCT299	CCT300	CCT301	CCT302	CCT303	CCT304	CCT305	CCT306	CCT307	CCT308	CCT309	CCT310	CCT311	CCT312	CCT313	CCT314	CCT315	CCT316	CCT317	CCT318	CCT319	CCT320	CCT321	CCT322	CCT323	CCT324	CCT325	CCT326	CCT327	CCT328	CCT329	CCT330	CCT331	CCT332	CCT333	CCT334	CCT335	CCT336	CCT337	CCT338	CCT339	CCT340	CCT341	CCT342	CCT343	CCT344	CCT345	CCT346	CCT347	CCT348	CCT349	CCT350	CCT351	CCT352	CCT353	CCT354	CCT355	CCT356	CCT357	CCT358	CCT359	CCT360	CCT361	CCT362	CCT363	CCT364	CCT365	CCT366	CCT367	CCT368	CCT369	CCT370	CCT371	CCT372	CCT373	CCT374	CCT375	CCT376	CCT377	CCT378	CCT379	CCT380	CCT381	CCT382	CCT383	CCT384	CCT385	CCT386	CCT387	CCT388	CCT389	CCT390	CCT391	CCT392	CCT393	CCT394	CCT395	CCT396	CCT397	CCT398	CCT399	CCT400	CCT401	CCT402	CCT403	CCT404	CCT405	CCT406	CCT407	CCT408	CCT409	CCT410	CCT411	CCT412	CCT413	CCT414	CCT415	CCT416	CCT417	CCT418	CCT419	CCT420	CCT421	CCT422	CCT423	CCT424	CCT425	CCT426	CCT427	CCT428	CCT429	CCT430	CCT431	CCT432	CCT433	CCT434	CCT435	CCT436	CCT437	CCT438	CCT439	CCT440	CCT441	CCT442	CCT443	CCT444	CCT445	CCT446	CCT447	CCT448	CCT449	CCT450	CCT451	CCT452	CCT453	CCT454	CCT455	CCT456	CCT457	CCT458	CCT459	CCT460	CCT461	CCT462	CCT463	CCT464	CCT465	CCT466	CCT467	CCT468	CCT469	CCT470	CCT471	CCT472	CCT473	CCT474	CCT475	CCT476	CCT477	CCT478	CCT479	CCT480	CCT481	CCT482	CCT483	CCT484	CCT485	CCT486	CCT487	CCT488	CCT489	CCT490	CCT491	CCT492	CCT493	CCT494	CCT495	CCT496	CCT497	CCT498	CCT499	CCT500	CCT501	CCT502	CCT503	CCT504	CCT505	CCT506	CCT507	CCT508	CCT509	CCT510	CCT511	CCT512	CCT513	CCT514	CCT515	CCT516	CCT517	CCT518	CCT519	CCT520	CCT521	CCT522	CCT523	CCT524	CCT525	CCT526	CCT527	CCT528	CCT529	CCT530	CCT531	CCT532	CCT533	CCT534	CCT535	CCT536	CCT537	CCT538	CCT539	CCT540	CCT541	CCT542	CCT543	CCT544	CCT545	CCT546	CCT547	CCT548	CCT549	CCT550	CCT551	CCT552	CCT553	CCT554	CCT555	CCT556	CCT557	CCT558	CCT559	CCT560	CCT561	CCT562	CCT563	CCT564	CCT565	CCT566	CCT567	CCT568	CCT569	CCT570	CCT571	CCT572	CCT573	CCT574	CCT575	CCT576	CCT577	CCT578	CCT579	CCT580	CCT581	CCT582	CCT583	CCT584	CCT585	CCT586	CCT587	CCT588	CCT589	CCT590	CCT591	CCT592	CCT593	CCT594	CCT595	CCT596	CCT597	CCT598	CCT599	CCT600	CCT601	CCT602	CCT603	CCT604	CCT605	CCT606	CCT607	CCT608	CCT609	CCT610	CCT611	CCT612	CCT613	CCT614	CCT615	CCT616	CCT617	CCT618	CCT619	CCT620	CCT621	CCT622	CCT623	CCT624	CCT625	CCT626	CCT627	CCT628	CCT629	CCT630	CCT631	CCT632	CCT633	CCT634	CCT635	CCT636	CCT637	CCT638	CCT639	CCT640	CCT641	CCT642	CCT643	CCT644	CCT645	CCT646	CCT647	CCT648	CCT649	CCT650	CCT651	CCT652	CCT653	CCT654	CCT655	CCT656	CCT657	CCT658	CCT659	CCT660	CCT661	CCT662	CCT663	CCT664	CCT665	CCT666	CCT667	CCT668	CCT669	CCT670	CCT671	CCT672	CCT673	CCT674	CCT675	CCT676	CCT677	CCT678	CCT679	CCT680	CCT681	CCT682	CCT683	CCT684	CCT685	CCT686	CCT687	CCT688	CCT689	CCT690	CCT691	CCT692	CCT693	CCT694	CCT695	CCT696	CCT697	CCT698	CCT699	CCT700	CCT701	CCT702	CCT703	CCT704	CCT705	CCT706	CCT707	CCT708	CCT709	CCT710	CCT711	CCT712	CCT713	CCT714	CCT715	CCT716	CCT717	CCT718	CCT719	CCT720	CCT721	CCT722	CCT723	CCT724	CCT725	CCT726	CCT727	CCT728	CCT729	CCT730	CCT731	CCT732	CCT733	CCT734	CCT735	CCT736	CCT737	CCT738	CCT739	CCT740	CCT741	CCT742	CCT743	CCT744	CCT745	CCT746	CCT747	CCT748	CCT749	CCT750	CCT751	CCT752	CCT753	CCT754	CCT755	CCT756	CCT757	CCT758	CCT759	CCT760	CCT761	CCT762	CCT763	CCT764	CCT765	CCT766	CCT767	CCT768	CCT769	CCT770	CCT771	CCT772	CCT773	CCT774	CCT775	CCT776	CCT777	CCT778	CCT779	CCT780	CCT781	CCT782	CCT783	CCT784	CCT785	CCT786	CCT787	CCT788	CCT789	CCT790	CCT791	CCT792	CCT793	CCT794	CCT795	CCT796	CCT797	CCT798	CCT799	CCT800	CCT801	CCT802	CCT803	CCT804	CCT805	CCT806	CCT807	CCT808	CCT809	CCT810	CCT811	CCT812	CCT813	CCT814	CCT815	CCT816	CCT817	CCT818	CCT819	CCT820	CCT821	CCT822	CCT823	CCT824	CCT825	CCT826	CCT827	CCT828	CCT829	CCT830	CCT831	CCT832	CCT833	CCT834	CCT835	CCT836	CCT837	CCT838	CCT839	CCT840	CCT841	CCT842	CCT843	CCT844	CCT845	CCT846	CCT847	CCT848	CCT849	CCT850	CCT851	CCT852	CCT853	CCT854	CCT855	CCT856	CCT857	CCT858	CCT859	CCT860	CCT861	CCT862	CCT863	CCT864	CCT865	CCT866	CCT867	CCT868	CCT869	CCT870	CCT871	CCT872	CCT873	CCT874	CCT875	CCT876	CCT877	CCT878	CCT879	CCT880	CCT881	CCT882	CCT883	CCT884	CCT885	CCT886	CCT887	CCT888	CCT889	CCT890	CCT891	CCT892	CCT893	CCT894	CCT895	CCT896	CCT897	CCT898	CCT899	CCT900	CCT901	CCT902	CCT903	CCT904	CCT905	CCT906	CCT907	CCT908	CCT909	CCT910	CCT911	CCT912	CCT913	CCT914	CCT915	CCT916	CCT917	CCT918	CCT919	CCT920	CCT921	CCT922	CCT923	CCT924	CCT925	CCT926	CCT927	CCT928	CCT929	CCT930	CCT931	CCT932	CCT933	CCT934	CCT935	CCT936	CCT937	CCT938	CCT939	CCT940	CCT941	CCT942	CCT943	CCT944	CCT945	CCT946	CCT947	CCT948	CCT949	CCT950	CCT951	CCT952	CCT953	CCT954	CCT955	CCT956	CCT957	CCT958	CCT959	CCT960	CCT961	CCT962	CCT963	CCT964	CCT965	CCT966	CCT967	CCT968	CCT969	CCT970	CCT971	CCT972	CCT973	CCT974	CCT975	CCT976	CCT977	CCT978	CCT979	CCT980	CCT981	CCT982	CCT983	CCT984	CCT985	CCT986	CCT987	CCT988	CCT989	CCT990	CCT991	CCT992	CCT993	CCT994	CCT995	CCT996	CCT997	CCT998	CCT999	CCT1000	CCT1001	CCT1002	CCT1003	CCT1004	CCT1005	CCT1006	CCT1007	CCT1008	CCT1009	CCT1010	CCT1011	CCT1012	CCT1013	CCT1014	CCT1015	CCT1016	CCT1017	CCT1018	CCT1019	CCT1020	CCT1021	CCT1022	CCT1023	CCT1024	CCT1025	CCT1026	CCT1027	CCT1028	CCT1029	CCT1030	CCT1031	CCT1032	CCT1033	CCT1034	CCT1035	CCT1036	CCT1037	CCT1038	CCT1039	CCT1040	CCT1041	CCT1042	CCT1043	CCT1044	CCT1045	CCT1046	CCT1047	CCT1048	CCT1049	CCT1050	CCT1051	CCT1052	CCT1053	CCT1054	CCT1055	CCT1056	CCT1057	CCT1058	CCT1059	CCT1060	CCT1061	CCT1062	CCT1063	CCT1064	CCT1065	CCT1066	CCT1067	CCT1068	CCT1069	CCT1070	CCT1071	CCT1072	CCT1073	CCT1074	CCT1075	CCT1076	CCT1077	CCT1078	CCT1079	CCT1080	CCT1081	CCT1082	CCT1083	CCT1084	CCT1085	CCT1086	CCT1087	CCT1088	CCT1089	CCT1090	CCT1091	CCT1092	CCT1093	CCT1094	CCT1095	CCT1096	CCT1097	CCT1098	CCT1099	CCT1100	CCT1101	CCT1102	CCT1103	CCT1104	CCT1105	CCT1106	CCT1107	CCT1108	CCT1109	CCT1110	CCT1111	CCT1112	CCT1113	CCT1114	CCT1115	CCT1116	CCT1117	CCT1118	CCT1119	CCT1120	CCT1121	CCT1122	CCT1123	CCT1124	CCT1125	CCT1126	CCT1127	CCT1128	CCT1129	CCT1130	CCT1131	CCT1132	CCT1133	CCT1134	CCT1135	CCT1136	CCT1137	CCT1138	CCT1139	CCT1140	CCT1141	CCT1142	CCT1143	CCT1144	CCT1145	CCT1146	CCT1147	CCT1148	CCT1149	CCT1150	CCT1151	CCT1152	CCT1153	CCT1154	CCT1155	CCT1156	CCT1157	CCT1158	CCT1159	CCT1160	CCT1161	CCT1162	CCT1163	CCT1164	CCT1165	CCT1166	CCT1167	CCT1168	CCT1169	CCT1170	CCT1171	CCT1172	CCT1173	CCT1174	CCT1175	CCT1176	CCT1177	CCT1178	CCT1179	CCT1180	CCT1181	CCT1182	CCT1183	CCT1184	CCT1185	CCT1186	CCT1187	CCT1188	CCT1189	CCT1190	CCT1191	CCT1192	CCT1193	CCT1194	CCT1195	CCT1196	CCT1197	CCT1198	CCT1199	CCT1200	CCT1201	CCT1202	CCT1203	CCT1204	CCT1205	CCT1206	CCT1207	CCT1208	CCT1209	CCT1210	CCT1211	CCT1212	CCT1213	CCT1214	CCT1215	CCT1216	CCT1217	CCT1218	CCT1219	CCT1220	CCT1221	CCT1222	CCT1223	CCT1224	CCT1225	CCT1226	CCT1227	CCT1228	CCT1229	CCT1230	CCT1231	CCT1232	CCT1233	CCT1234	CCT1235	CCT1236	CCT1237	CCT1238	CCT1239	CCT1240	CCT1241	CCT1242	CCT1243	CCT1244	CCT1245	CCT1246	CCT1247	CCT1248	CCT1249	CCT1250	CCT1251	CCT1252	CCT1253	CCT1254	CCT1255	CCT1256	CCT1257	CCT1258	CCT1259	CCT1260	CCT1261	CCT1262	CCT1263	CCT1264	CCT1265	CCT1266	CCT1267	CCT1268	CCT1269	CCT1270	CCT1271	CCT1272	CCT1273	CCT1274	CCT1275	CCT1276	CCT1277	CCT1278	CCT1279	CCT1280	CCT1281	CCT1282	CCT1283	CCT1284	CCT1285	CCT1286	CCT1287	CCT1288	CCT1289	CCT1290	CCT1291	CCT1292	CCT1293	CCT1294	CCT1295	CCT1296	CCT1297	CCT1298	CCT1299	CCT1300	CCT1301	CCT1302	CCT1303	CCT1304	CCT1305	CCT1306	CCT1307	CCT1308	CCT1309	CCT1310	CCT1311	CCT1312	CCT1313	CCT1314	CCT1315	CCT1316	CCT1317	CCT1318	CCT1319	CCT1320	CCT1321	CCT1322	CCT1323	CCT1324	CCT1325	CCT1326	CCT1327	CCT1328	CCT1329	CCT1330	CCT1331	CCT1332	CCT1333	CCT1334	CCT1335	CCT1336	CCT1337	CCT1338	CCT1339	CCT13
--------	-------------------------------------------------	---------	------------------------	-------------------------	--------------------------	-------------------------------	------	------	------	------	------	------	------	------	------	------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	-------