LabBook_15_04_16

Claire Green

11 April 2016

Monday

I started with calculating the enrichment of my DEGs in the pathprint pathway gene list. I did so using this script:

```
pathprint <- read.table(file = "Pathprintgenes.txt")
pathprint <- pathprint$V1

pathprintunique <- pathprint[!duplicated(pathprint)]

overlap <- Reduce(intersect, list(x, pathprint))
print(overlap)


x <- read.table(file = "DEGs.txt")
x <- x$V1

library(hgu133plus2.db)
sym <- hgu133plus2SYMBOL
sym1 <- mappedkeys(sym)
sym2 <- as.list (sym[c(sym1)])
sym3 <- data.frame (sym2)
sym.probes <- names (sym2)
sym.genes <- sym3[1,]

x.in <- length (which(x %in% pathprintunique))
x.out <- length(x) - x.in
tot.in <- length (pathprintunique)
tot.out <- length (sym.genes)

counts <- matrix (nrow=2, ncol=2)
counts [1,] <- c(x.in, tot.in)
counts [2,] <- c(x.out, tot.out)

a5 <-fisher.test (counts)
enrich <- a5$p
```

Overlapping genes were “KPNA6” “NUTF2” “PLOD2” “PPP2CA” “PPP2CB”

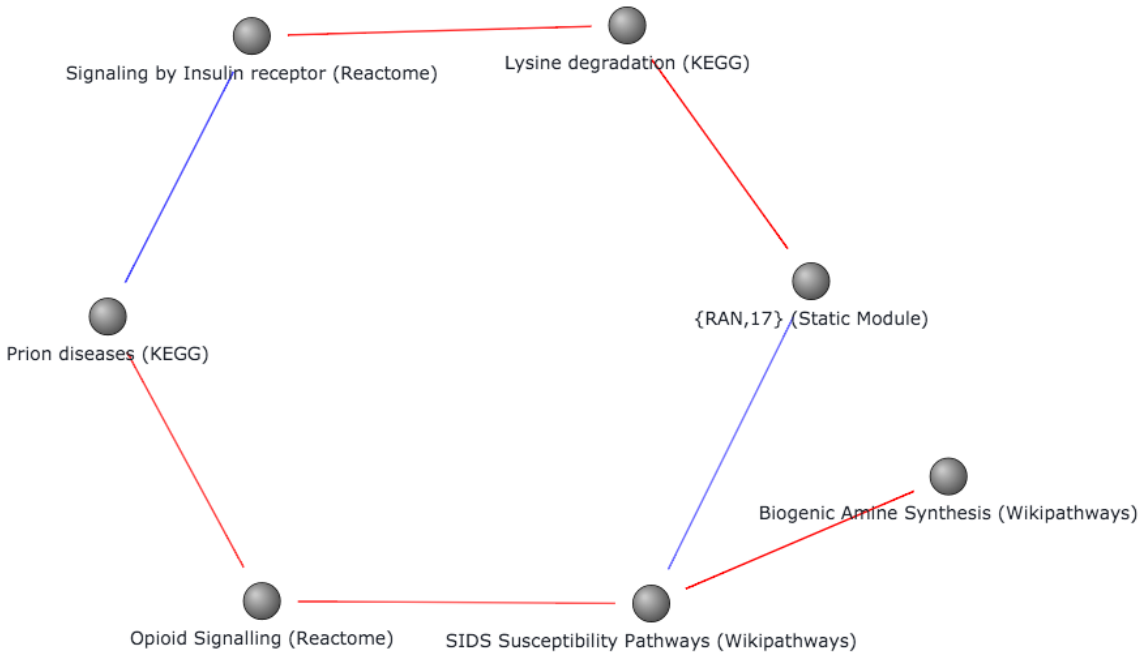
The result was that when the duplicates were not removed, significance was 0.004. With duplicates removed, significance was p = 0.002. This means that my 5 genes are significantly enriched in the pp gene list as compared to the proportion of all genes represented by the list.

I was interested to see if the Pathprint genes were enriched with SNPs. I used the no-duplicates pathprint list and the GWAS central list of SNPs where p<.0001 (“signif.snp.GWAScentral.p0.0001.1.txt”). There were 4 genes containing SNPS (“KCNQ1” “PPARGC1A” “GNG7” “ITPR2”) but no significant enrichment.

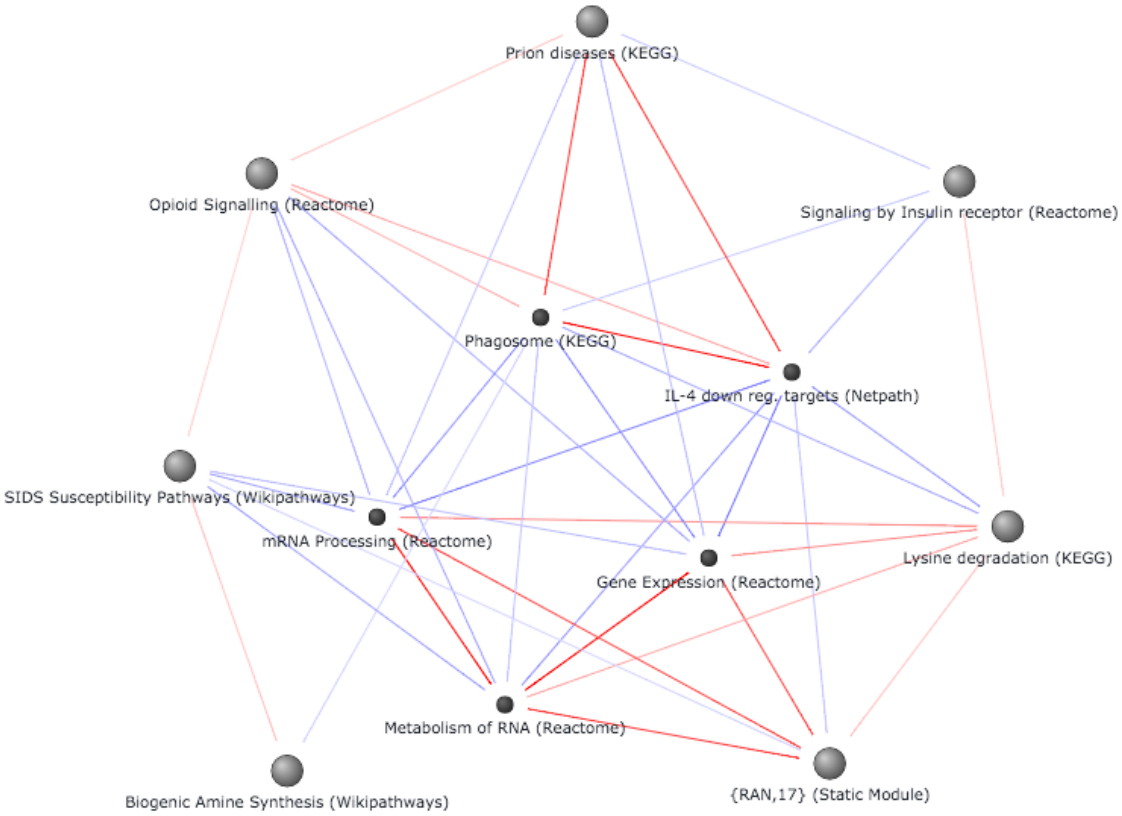
Next, I looked at the NeuroX list. There were 7 genes overlapping these lists ("NOTCH1" "SOD1" "COMT" "CHRNA4" "FGFR3" "STK11" "TSC2") and this enrichment was significant (p = 0.0002)

When I look at the pathways in which these are enriched, Prion diseases (KEGG) contains 2 NeuroX genes (SOD1, NOTCH1), Biogenic Amine Synthesis (KEGG) contains 1 (COMT), SIDS susceptibility pathway (Wikipathways) contains 1 neuroX (CHRNA4) and two GWAS central (KCNQ1, PPARGC1A), Signalling by insulin receptor (Reactome) contains 3 neuroX genes (FGFR3, STK11, TSC2), Opioid signalling (Reactome) contains 2 GWAS central genes (GNG7, ITPR2) and 2 DEGS (PPP2CA, PPP2CB).

The first thing I did was put these pathways in PCxN, including RAN,17 which contains 2 DEGs.

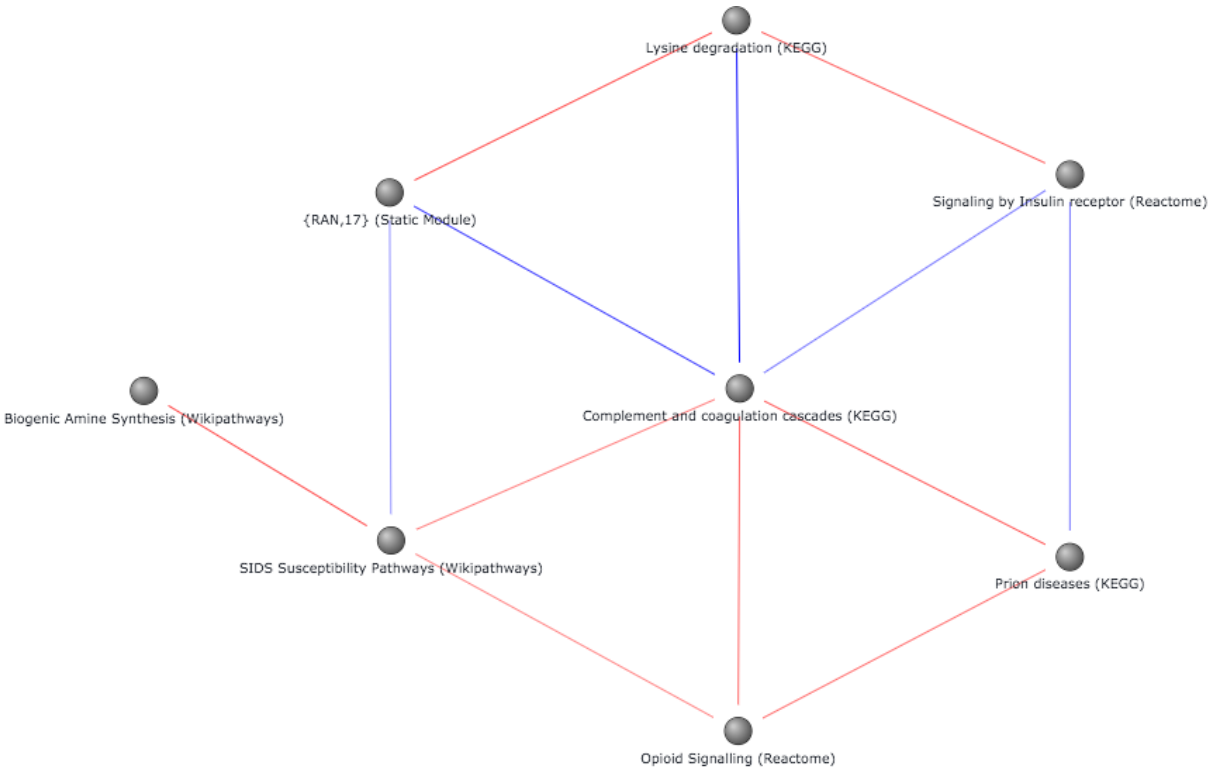


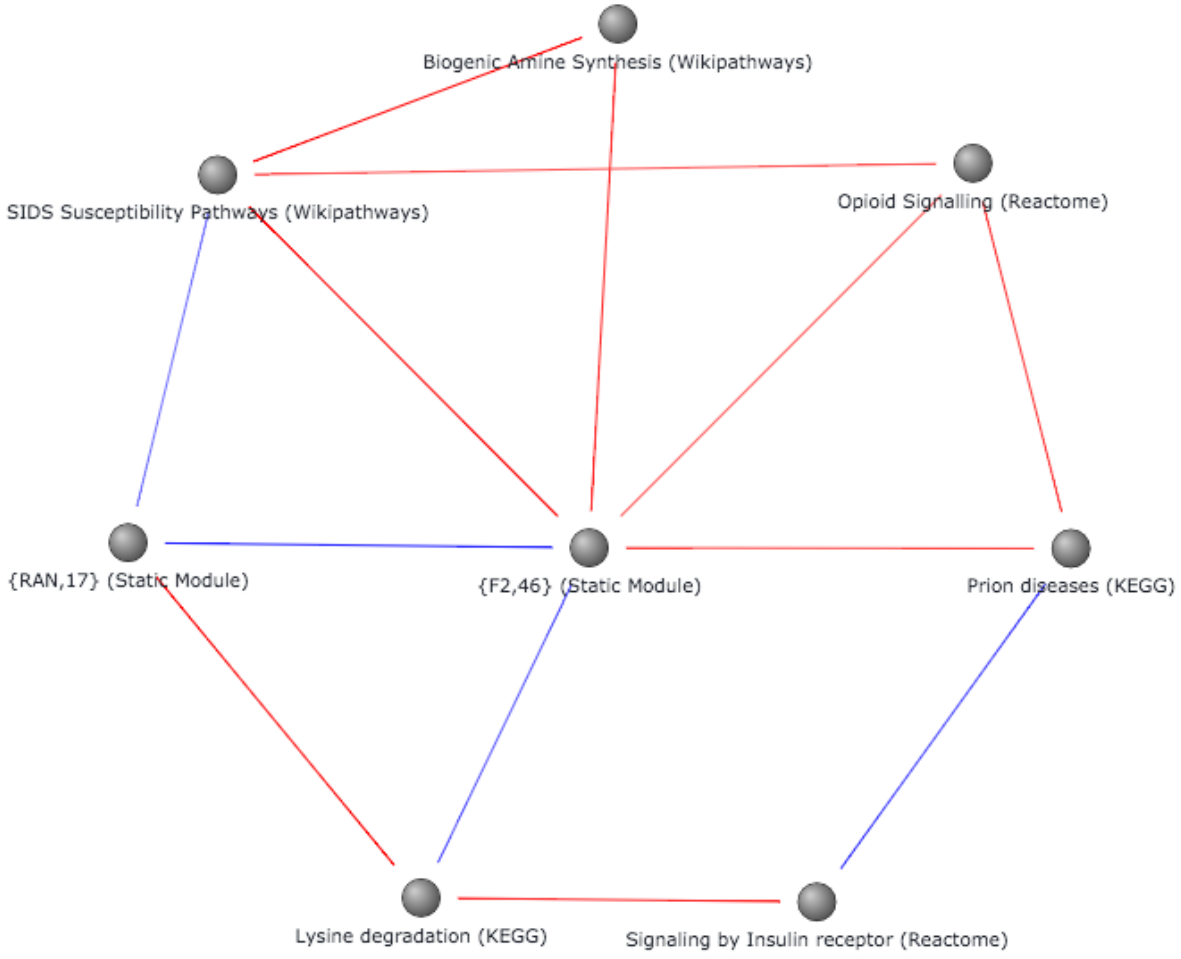
It's not particularly connected, so I tried adding the 5 most correlated gene sets

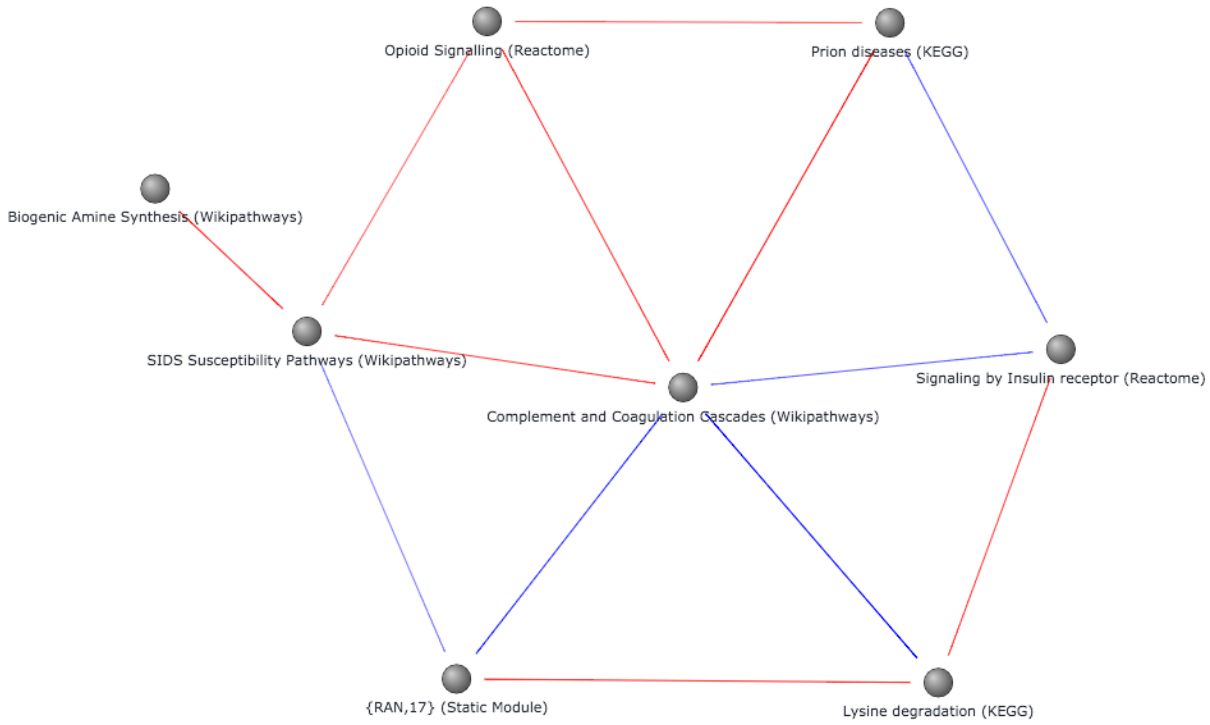


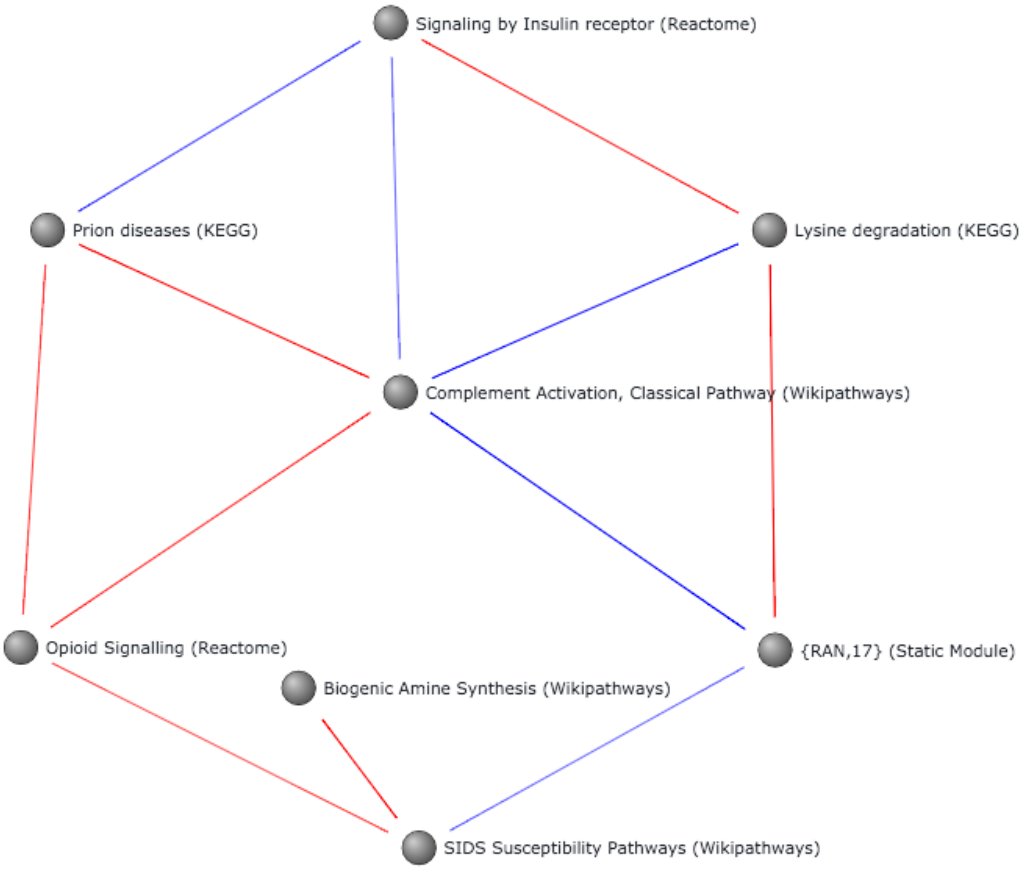
What we get out is Phagosome, IL-4 down reg. targets, mRNA processing, metabolism of RNA and gene expression. The last three can particularly be linked to TDP-43 activity.

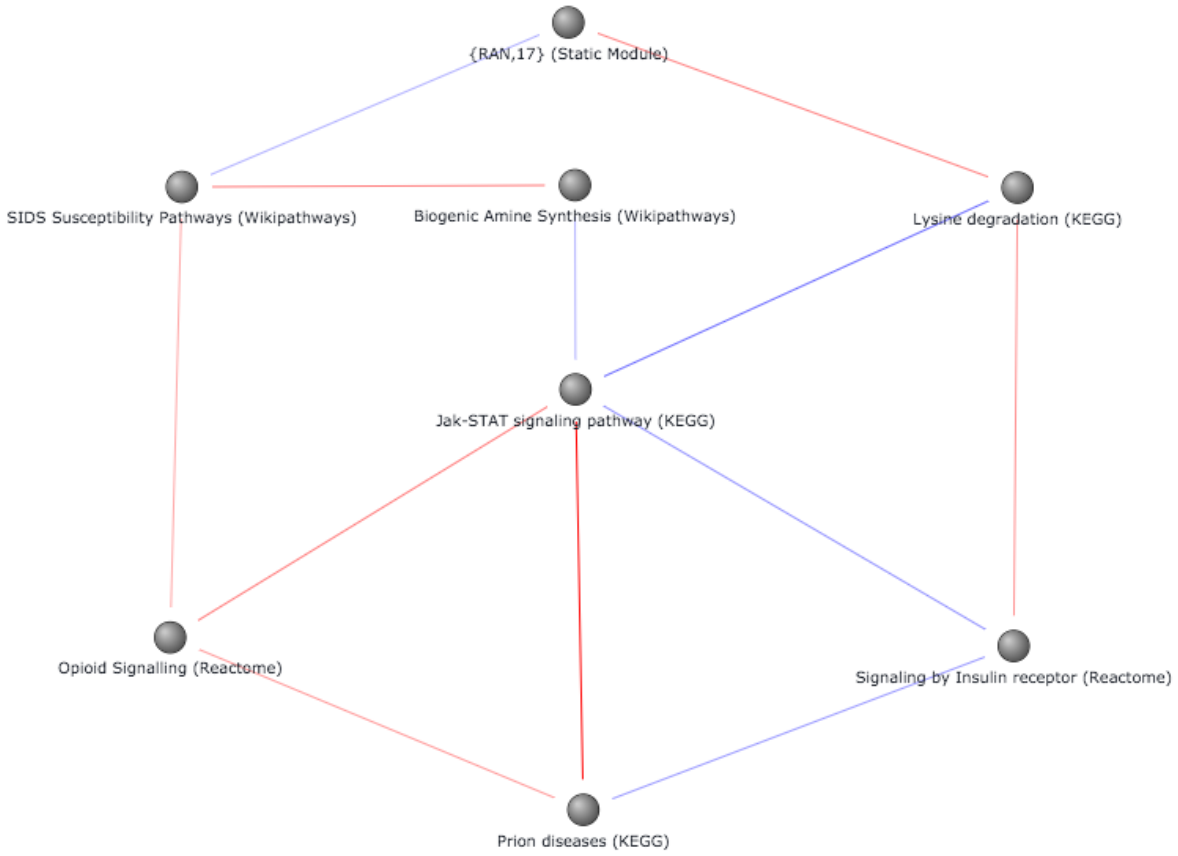
What I then tried to do is discover if any of the other pathprint pathways were able to connect the enriched pathways better. I identified 5 pathways:


Lysine degradation (KEGG)
{RAN,17} (Static Module)
Signaling by Insulin receptor (Reactome)
Biogenic Amine Synthesis (Wikipathways)
Complement and coagulation cascades (KEGG)
SIDS Susceptibility Pathways (Wikipathways)
Prion diseases (KEGG)
Opioid Signalling (Reactome)


Biogenic Amine Synthesis (Wikipathways)
SIDS Susceptibility Pathways (Wikipathways)
Opioid Signalling (Reactome)
{RAN,17} (Static Module)
{F2,46} (Static Module)
Prion diseases (KEGG)
Lysine degradation (KEGG)
Signaling by Insulin receptor (Reactome)


Opioid Signalling (Reactome)
Prion diseases (KEGG)
Biogenic Amine Synthesis (Wikipathways)
SIDS Susceptibility Pathways (Wikipathways)
Signaling by Insulin receptor (Reactome)
Complement and Coagulation Cascades (Wikipathways)
{RAN,17} (Static Module)
Lysine degradation (KEGG)


Signaling by Insulin receptor (Reactome)
Prion diseases (KEGG)
Lysine degradation (KEGG)
Complement Activation, Classical Pathway (Wikipathways)
Opioid Signalling (Reactome)
{RAN,17} (Static Module)
Biogenic Amine Synthesis (Wikipathways)
SIDS Susceptibility Pathways (Wikipathways)



It appears that complemet and coagulation cascades seems to be the pathway that links these enriched pathways together.

Next, I downloaded all the SNPs identified by GWAS catalog and added them in. Now the pathways are this:

Pathways	GWAS Central	NeuroX	GWAS catalog	DEGS
Lysine degradation (KEGG)	x	x	HADH	PLOD2
ABC transporters (KEGG)	x	x	ABCG1 ABCC12	x
Complement and coagulation cascades (KEGG)	x	x	MASP1	x
Prion diseases (KEGG)	x	NOTCH1 SOD1	x	x
Biogenic Amine Synthesis (Wikipathways)	x	COMT	x	x
SIDS Susceptibility Pathways (Wikipathways)	CHRNA4	KCNQ1 PPARGC1A	x	x
Signaling by Insulin receptor (Reactome)	x	FGFR3 STK11 TSC2	x	x
Opioid Signalling (Reactome)	GNG7 ITPR2			PPP2CA PPP2CB
{RAN,17} (Static Module)	x	x	x	NUTF2 KPNA6

Now we can confirm the inclusion of complement and coagulation cascades.