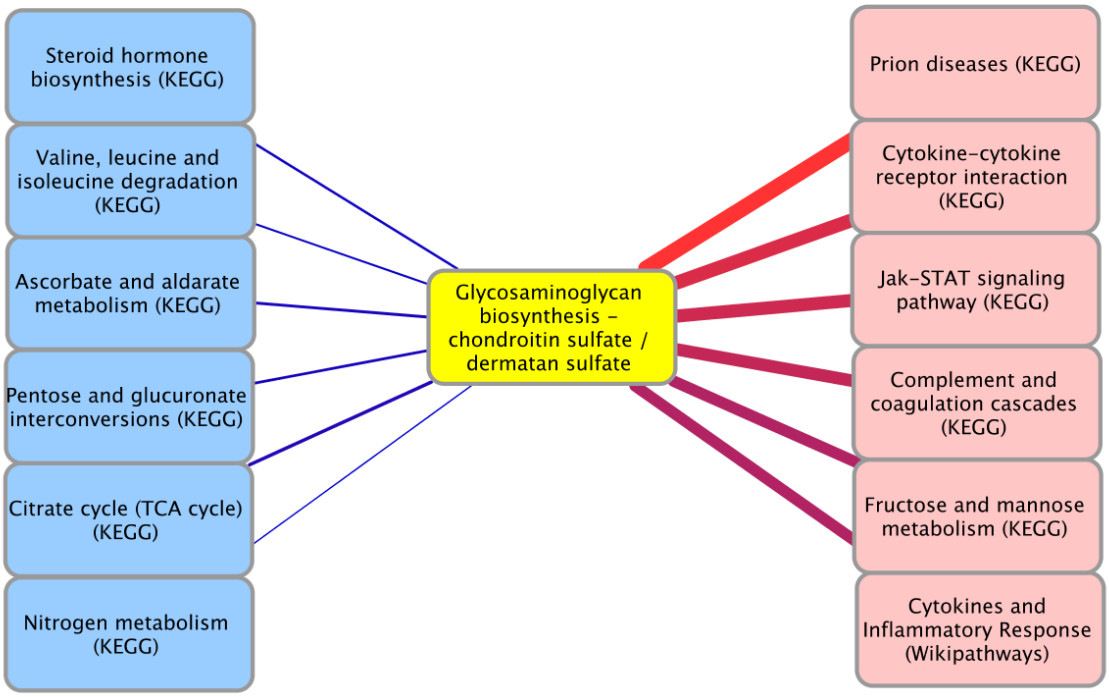
LabBook_08_04_16

Claire Green

Monday

Spent the morning tidying up what I had discovered on Friday and continued working on my powerpoint presentation for Win. I looked this afternoon at comparing the DEG enriched pathways and the pathprint pathways. This was not easy as GeneMANIA uses GO terms whereas pathprint uses KEGG, Wikipathways etc. I had to try and find equivalents of the GO pathways in the Pathprint database, which I think is super sketchy but I didn't have a choice. To be honest, I'm just trying to get somthing interesting for the presentation so it doesn't have to be groundbreaking. It was interesting though - I could only find a pathway related to the dermatan/chondroitin metabolism, but when I correlated that with the pathprint pathways I ended up with metabolic pathways being negatively correlated, and immunity/inflammation pathways positively correlated.



Tuesday

I worked a little on the presentation today, still trying to cut it down as it's far too long. In the mean time John had sent me some different GWAS lists, which tightened the thresholds of significance (and therefore lost me some genes). I also went through GWAS catalog and looked to see if there were any on there, and I found that there were some in there that had been reported. So overall:

GWAS central (ALS) CSRP1 (0.0006729)
RAB40B (0.0005151)
SERBP1 (0.0008722)
TUG1 (0.0006935)

GWAS catalog PFDN1 (Alzheimer's Disease, 3x10-7)
ZFYVE26 (ALS, 6x10-7)

Note that TUG1 is a non-protein-coding gene so is often missed. For full details on the SNPs, go to the file ALLDEGs.xlsx

Wednesday

Today I looked at extending my gene list to see if there are any SNP-containing genes associated with my seed list. I generated a list of 95 by including 50 more genes by genemania (co-expression, genetic interaction, physical interaction, weighting based on biological processes)

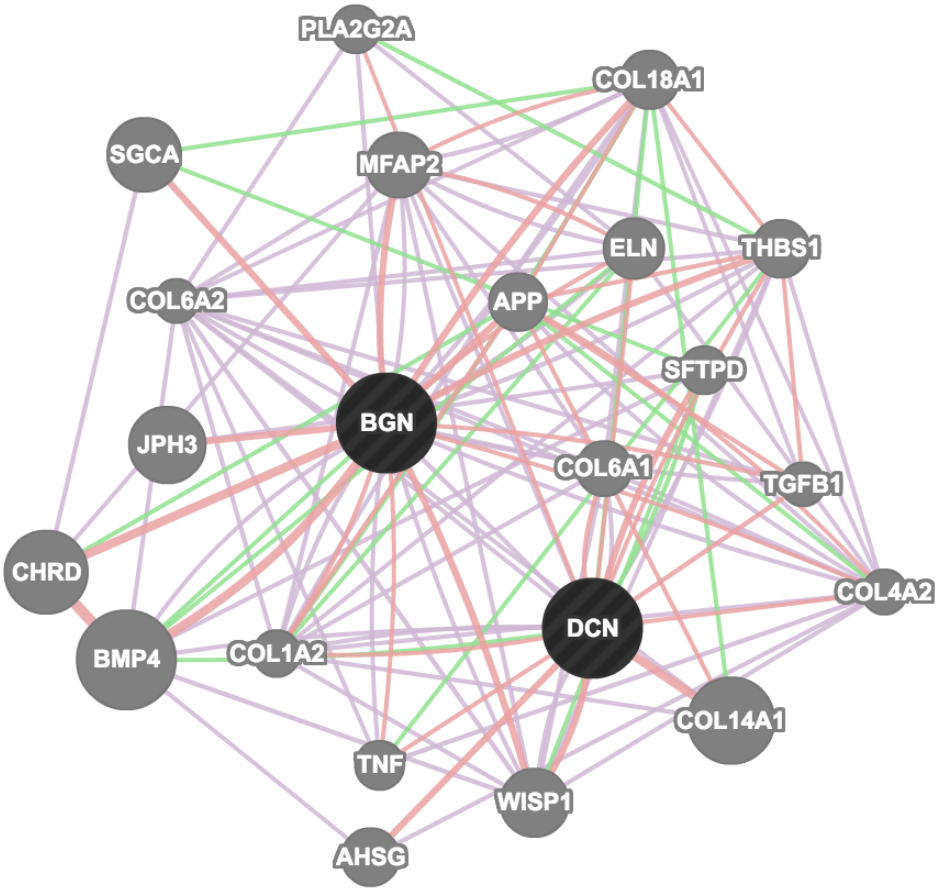
This was the output: TUG1 NDUFS5 PLEKHB1 BBIP1 DCN RPL35A RAB40B VPS13B JAG1 KPNA6 MPHOSPH9 NKTR ZNF518A ZFYVE26 CDH11 PTPN13 PPP1R7 STMN1 SERBP1 POGZ PFDN1 SCN1B CST3 BPTF CREG1 RPL37 PLOD2 OTUB1 BGN CSRP1 NUTF2 ANXA1 TCF4 MXI1 GBAS KCTD12 PPP2CB TUBB4B PRPF3 ACTN1 SPARC ZFP36 SF3B1 PPP2CA TARDBP SF3B14 COL3A1 NELFE MAZ PPP2R1A UBC MAX OTUD6B ATXN2 SH2D3A RAN AGO4 PPP1CA MACF1 NOTCH1 CALD1 RBM17 UHMK1 PLRG1 ACTN4 SCN2B DDX17 ANXA4 KPNB1 PFDN5 TMEM189-UBE2V1 C4A SMNDC1 TLX1 BMP4 MFAP1 TUBB COL14A1 THOC2 PCBP1 SET SF3A2 TUBA1B GNGT1 PPP2R1B S100A11 RNPS1 PAN2 PRKAR2A NEURL LRRC7 CTNNB1 CNOT1 FOS EIF4A3

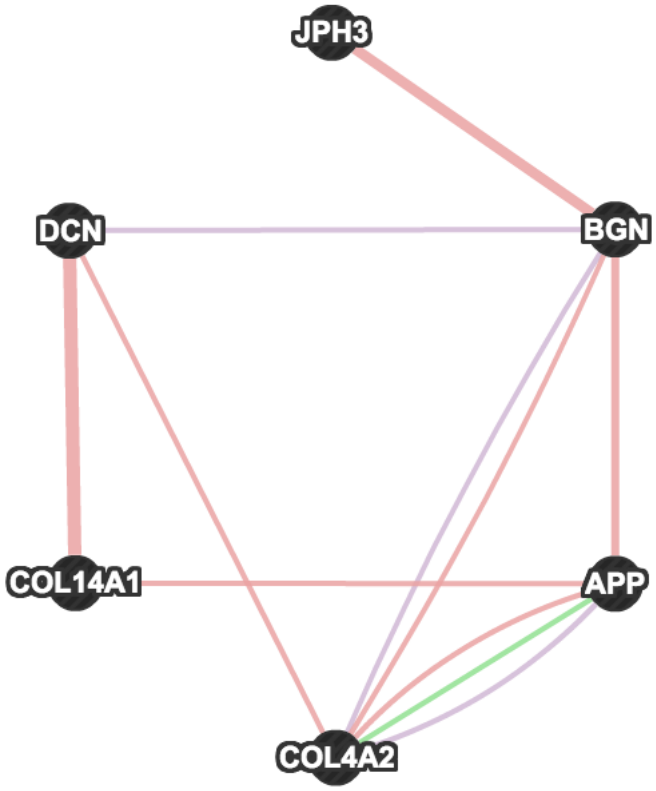
From this I identified 4 more SNPS

COL14A1 (ALS) COL3A1 (AD and Lewy body dementia) CTNNB1 (ALS) LRCC7 (ALS)

I tried using GWAS central myself but it's horrible to navigate. John says that they don't allow large amounts of information to be downloaded, and when he tried to contact them about AD SNPs they basically said no because of collaboration/authorship problems.

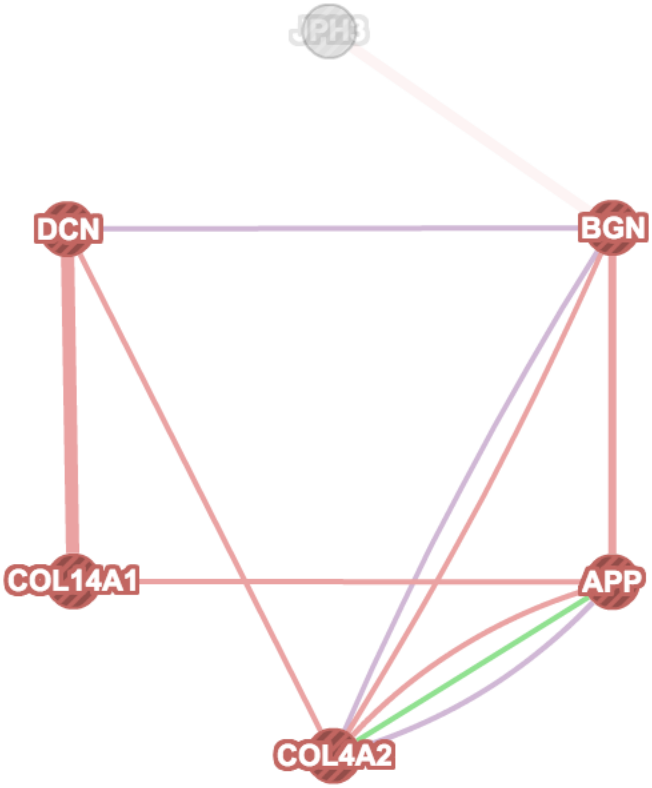
I did a little investigation of the DCN/BGN story as I had been emphasising it in my presentation. I decided to input just DCN and BGN in genemania and ask for the 20 most related genes. This contained 4 GWAScentral SNPs, including APP.


PLA2G2A
COL18A1
SGCA
MFAP2
ELN
THBS1
COL6A2
APP
SFTPD
BGN
JPH3
COL6A1
TGFB1
CHRD
COL4A2
DCN
BMP4
COL1A2
COL14A1
TNF
WISP1
AHSG



Just SNP genes

As it turned out, of these 6 genes, 5 were enriched for extracellular matrix organisation, but not dermatan/chondroitin metabolism.


PH3
DCN
BGN
COL14A1
APP
COL4A2