Monocyte analysis

Data description

Using a FDR threshold of 0.05 for hotspot association, we get $p = 381$ genes for each of the four conditions. We have $n_1 = 366$ samples for IFN-GAMMA, $n_2 = 260$ for LPS-2h, $n_3 = 321$ for LPS-24h, and $n_4 = 413$ for unstimulated cells. The number of genes in each condition mediated by the top hotspot were 294, 88, 16 and 215 respectively.

Resulting network description

The jointGHS network estimates had sparsities 0.18, 0.20, 0.21 and 0.16 for IFN-GAMMA, LPS-24, LPS-24h and unstimulated cells respectively. Their corresponding chosen global shrinkage parameters were 8 6.4, 9 and 3 respectively.

Compared to the single-network estimates of the same sparsity, the log likelihood is either increased or approximately the same for all conditions.

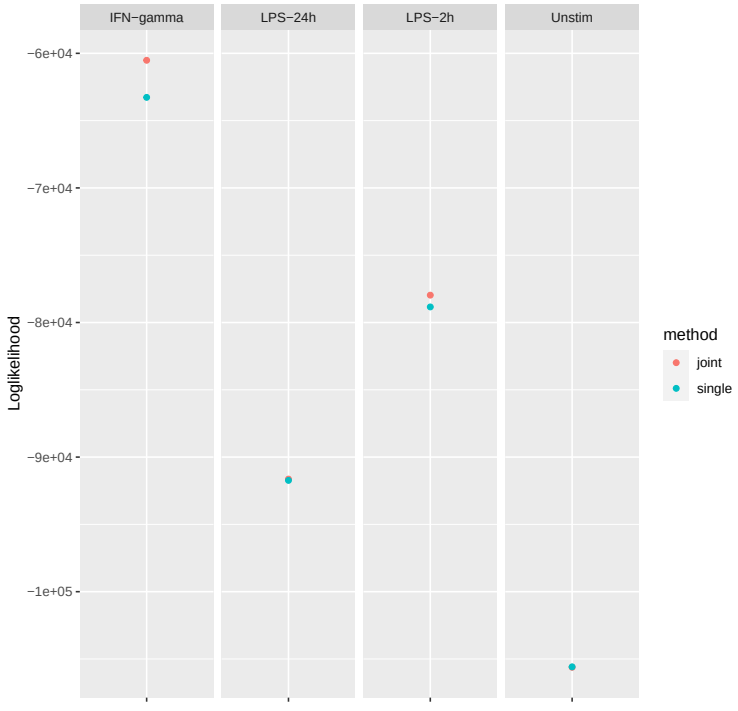


Figure 1: Log likelihood of jointGHS vs fastGHS networks

Similarity of networks across conditions

Using Matthews' correlation coefficient for networks, we see that the networks of IFN-gamma and LPS24h are the most different to each other.

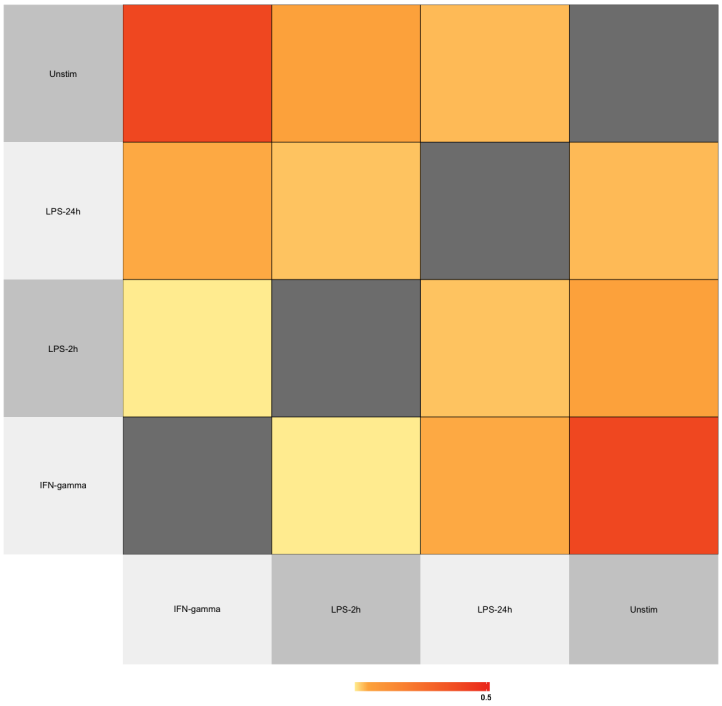


Figure 2: MCC of jointGHS networks

Degree distributions

LYZ

LYZ has a relatively high degree in all conditions, though it is not among the top hubs in any of the conditions. It does not have any edges to other cis genes, except for in LPS-24 where there is an edge between LYZ and YEATS4. In fact, this is the only cis-cis edge across all conditions. This implies that the cis genes work independently, and do not regulate each other.

YEATS4

YEATS4 also has a high edge degree in all networks, and is clearly a hub.

CREB1

The mediator trans gene CREB1 also has a fairly high degree in all networks.

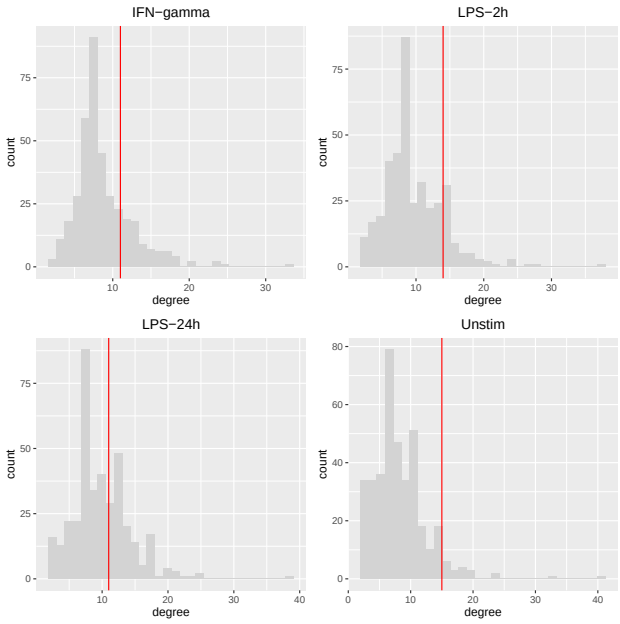


Figure 3: Degree distributions of jointGHS networks, with edge degree of LYZ marked in red

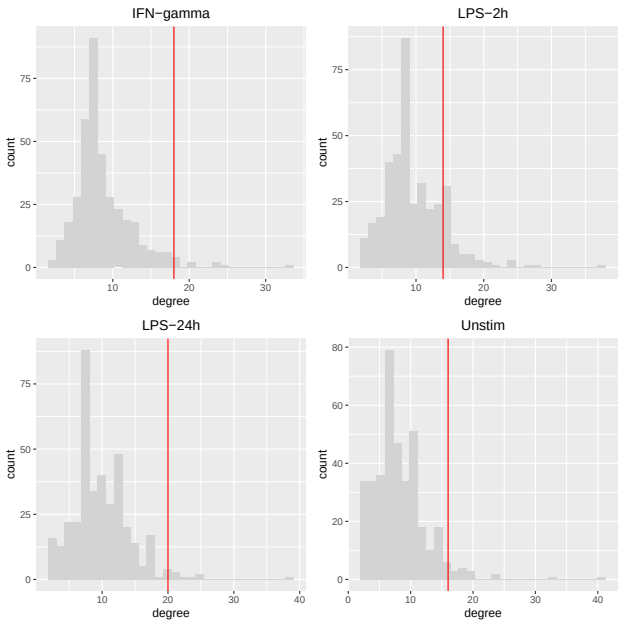


Figure 4: Degree distributions of jointGHS networks, with edge degree of YEATS4 marked in red

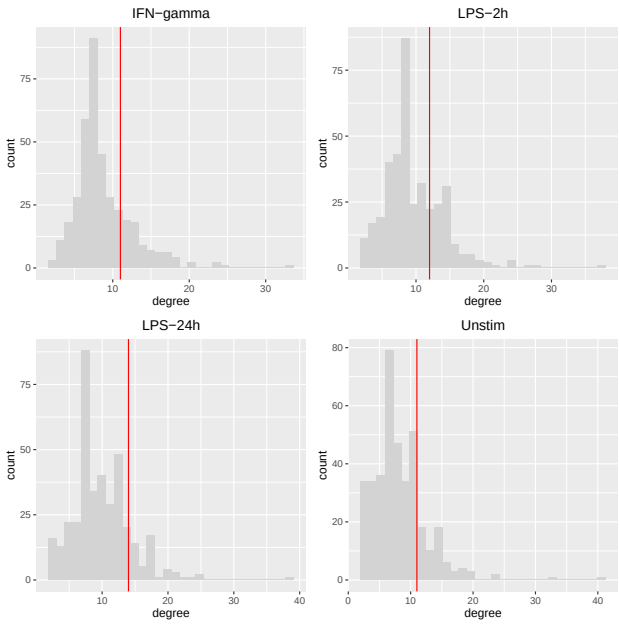


Figure 5: Degree distributions of jointGHS networks, with edge degree of CREB1 marked in red

Density plots

It apprears that the degree distribution of the unstimulated network has more density on lower degrees. Indeed, the unstimulated network is the sparsest among the four.

Density plots of top hotspot-controlled genes

The degree distribution of the genes controlled by the top hotspot clearly has the most mass on larger values in LPS-24h, also when compared to the distribution in the joint network estimate, implying more activity among these genes in this group.

Density plots of single network version

As we see, the single network estimates does not have nodes with degree larger than 25, compared to 40 for the joint estimate. Thus, the joint graphical horseshoe estimator identifies more high-degree hubs.

Hubs

It can be relevant to investigate whether the conditions have common hubs. For this purpose, we consider hubs to be the nodes with edge degree larger than the 90th percentile. This resulted in four lists with 56, 46, 49 and 48 genes for IFN-GAMMA, LPS-2h, LPS-24h and the unstimulated network respectively. The top hubs common to all four networks were: ACBD5, AFMID, AFTPH, AIRE, COX6A1, GIMAP1, IMPDH1, KIAA0101, LGALS3, RELB, SLC3A2, SORL1 and STAG3L3.

LYZ and YEATS4 had many edges with the top hubs in all conditions. YEATS4 in particular is highly connected to the top hubs in each network, in addition to being a top hub in all networks but that of LPS-2h.

Looking at the location of the top hubs, we see that they are fairly well spread out.

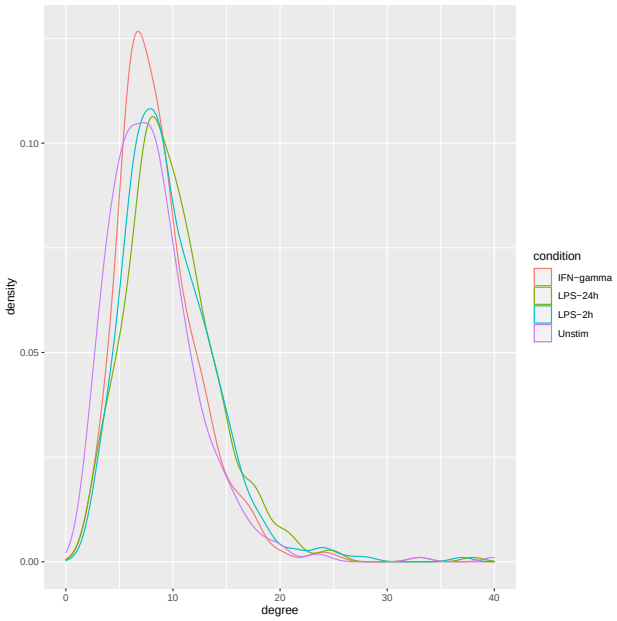


Figure 6: Degree distributions density of jointGHS networks

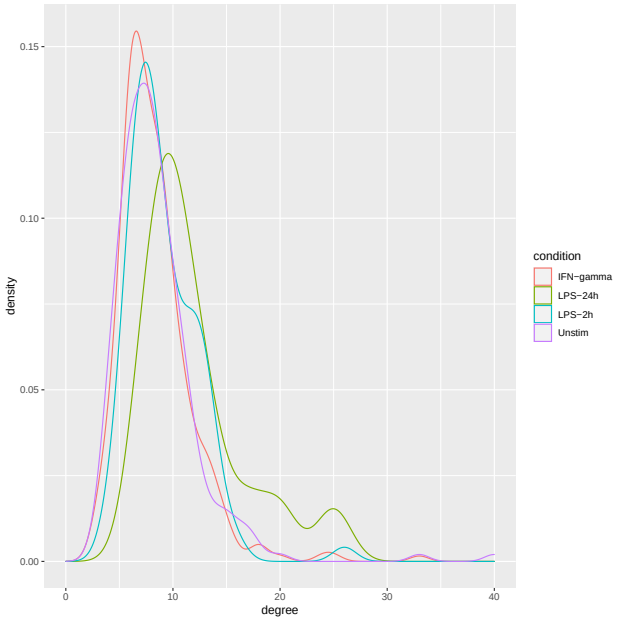


Figure 7: Degree distributions density of genes controlled by the top hotspot in the jointGHS networks

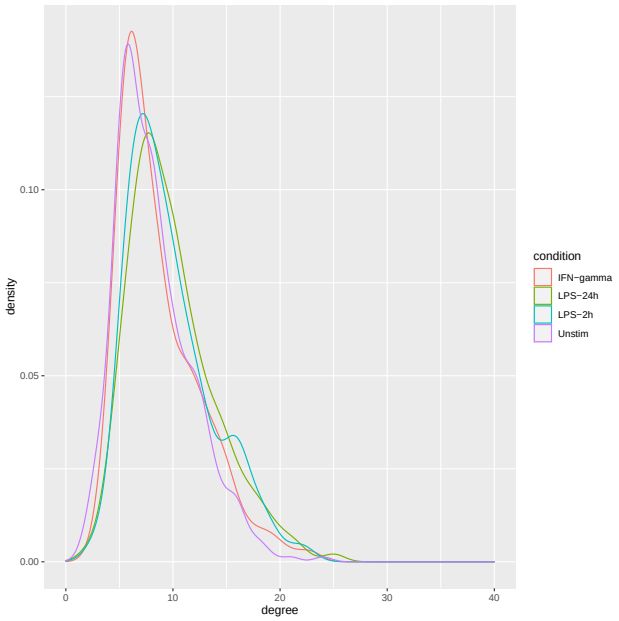


Figure 8: Degree distributions density of fastGHS networks

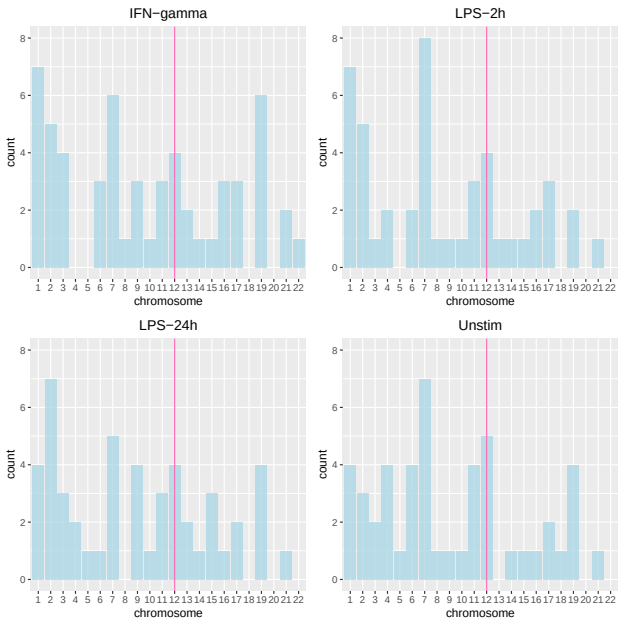


Figure 9: The location of the top hubs in the different conditions, with the 12th chromosome marked in pink

As we see from the heatmap, the unstimulated group and LPS-24h have the fewest common hubs.

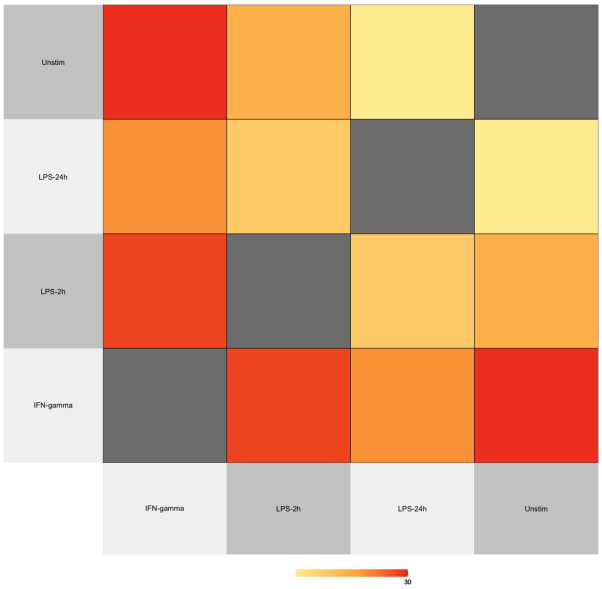


Figure 10: The number of top hubs in one condition that is present in another condition, divided by the total number of top hubs in both conditions

Resulting networks

Edges of LYZ gene

In the figure showing the subgraph with the edges of LYZ in the different conditions, a red edge indicates that the edge is unique to that condition. A blue edge is common across all networks, and a grey edge is present in two or three networks. Solid lines represent positive partial correlation while a dashed line represents negative partial correlation. Nodes are sized according to their edge degree, with larger size representing a higher degree.

As we see, LYZ does not have a striking number of neighbors in IFN-GAMMA. On the other hand, it has more edges in the others, in particular LPS-2h which also has the most unique edges. We see that a lot of the edges of LYZ represent negative partial correlations, implying it could have a role in trans down-regulation in addition to up-regulation.

The neighbors of LYZ common in the four networks are AFMID, LOC100128098, KLHL28, MAFF and SNRNP48. The partial correlations between LYZ and the the latter genes are negative in all conditions, and positive for the two former.

Edges of YEATS4 gene

YEATS4 has a high degree in all networks, and is among the top 10 hubs in all networks but that of LPS-2h. Evidently, it has many edges, corresponding to both positive and negative partial correlations. It is the cis gene with the overall highest degree in all conditions.

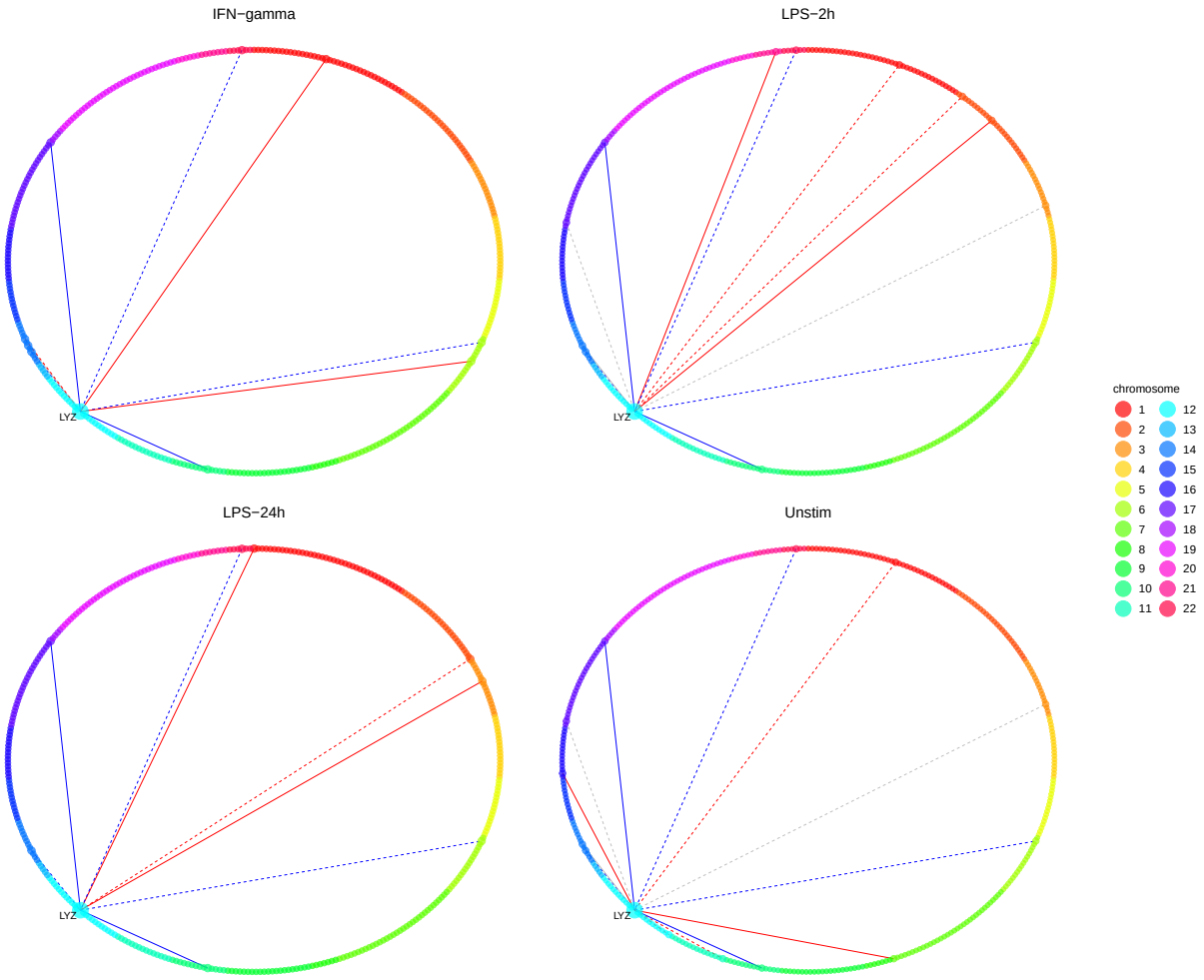


Figure 11: Subnetwork showing the edges from LYZ in the joint estimate

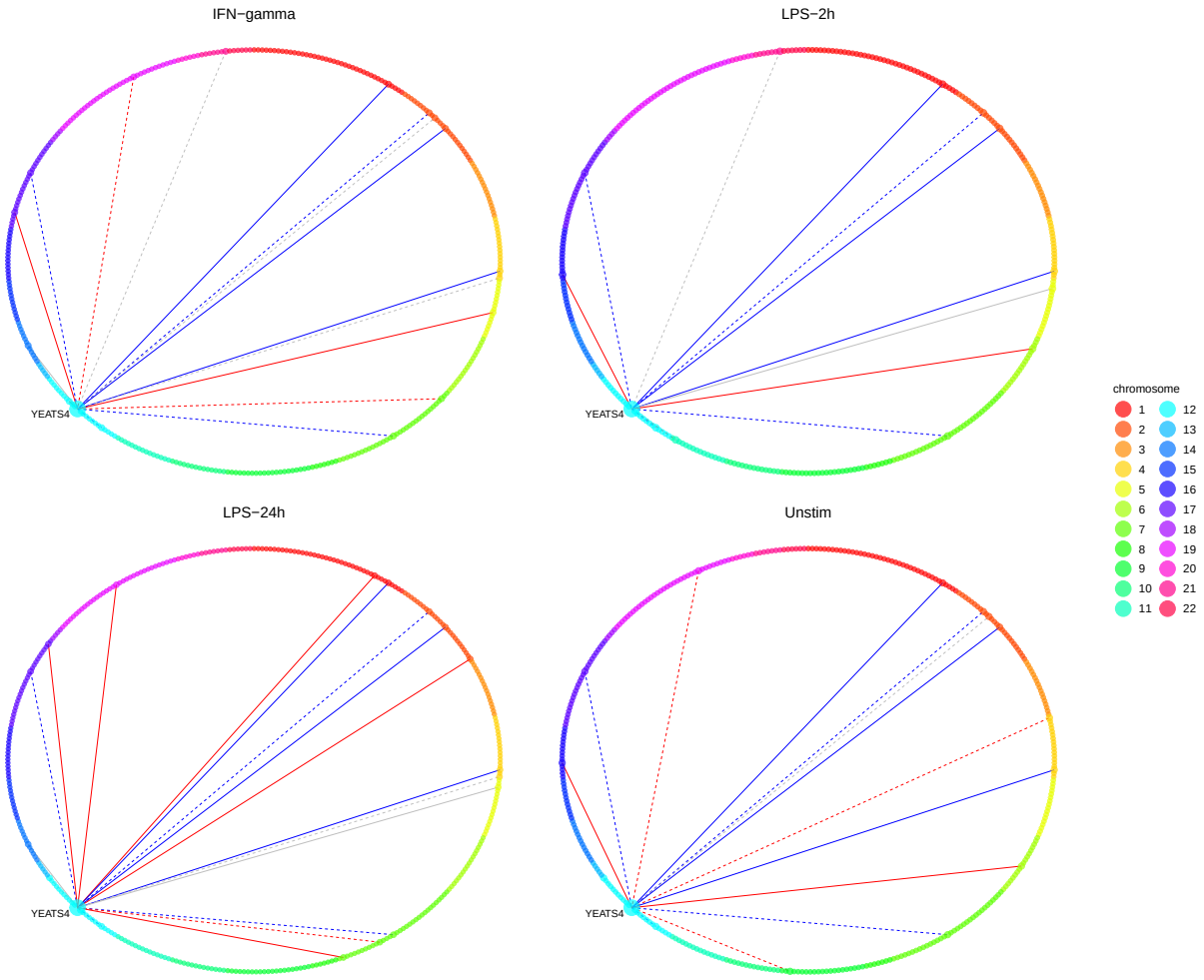


Figure 12: Subnetwork showing the edges from YEATS4 in the joint estimate

Edges of CREB1 gene

We also look at the edges of the trans gene CREB1, as it is found to have a mediating effect. CREB1 does not have any edges to the cis genes.

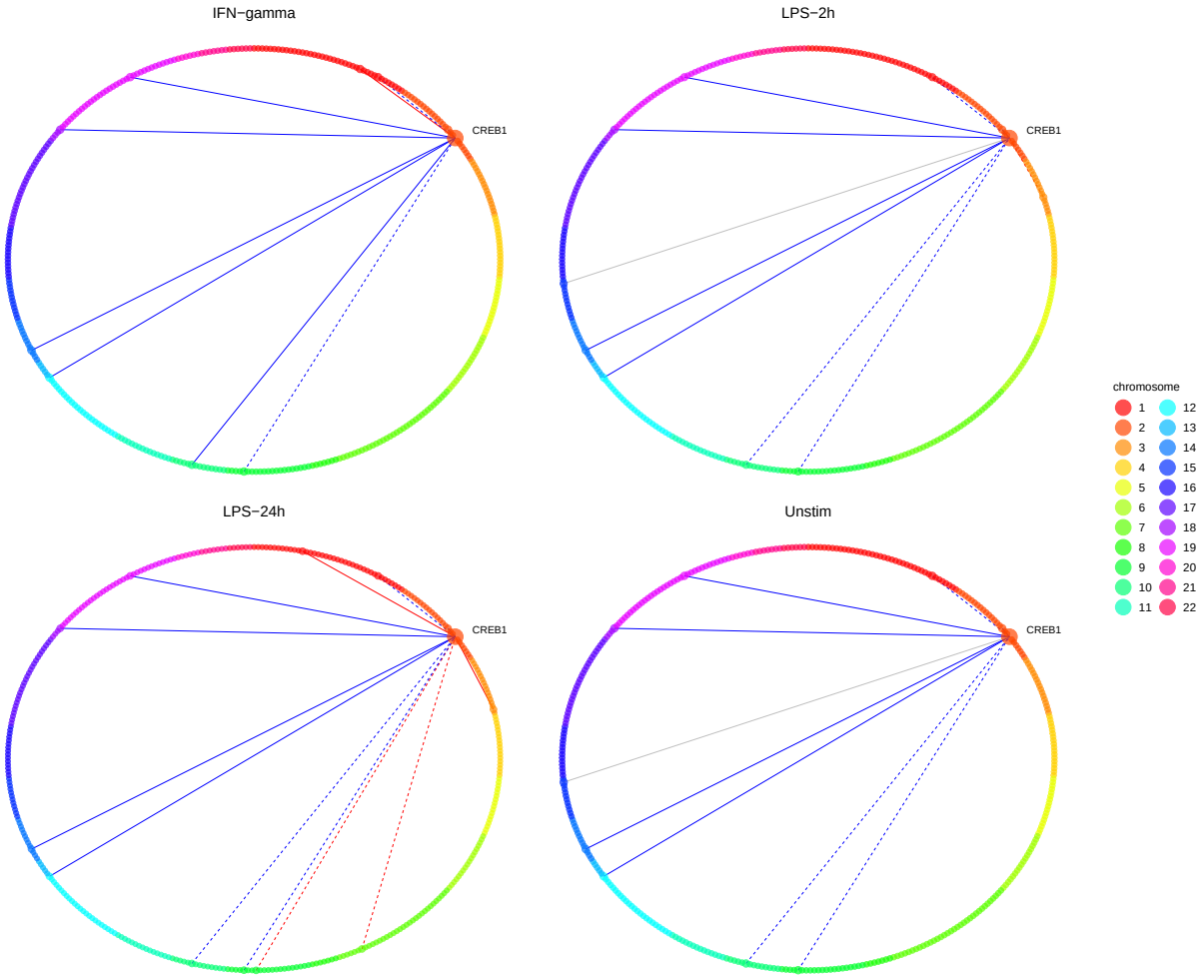


Figure 13: Subnetwork showing the edges from CREB1 in the joint estimate

Edges of both LYZ and YEATS4 genes

We also show the edges of the two cis genes controlled by the top hotspot together.

Full network

We also include the full network with all edges, though little can be concluded from this plot.

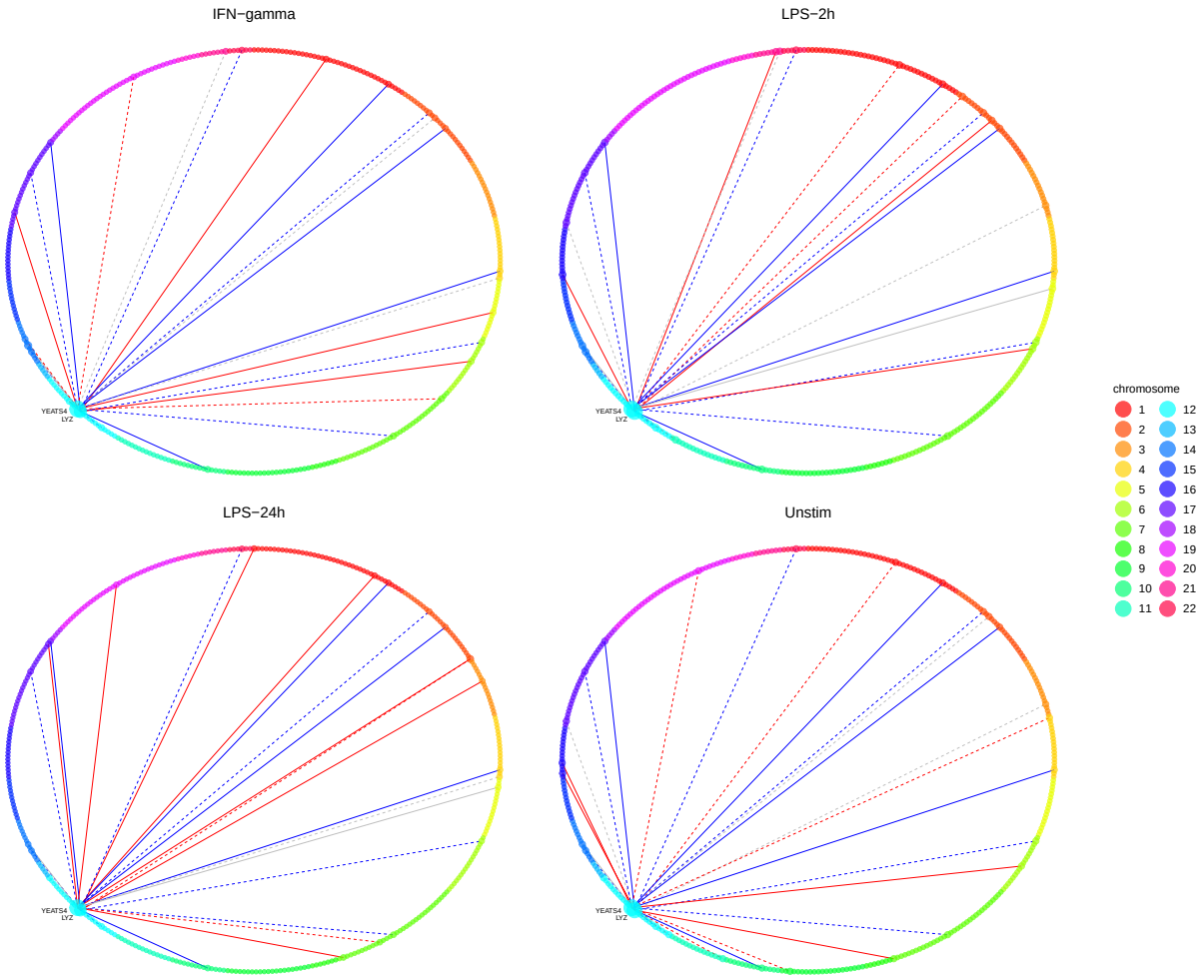


Figure 14: Subnetwork showing the edges from all genes in cis in the joint estimate

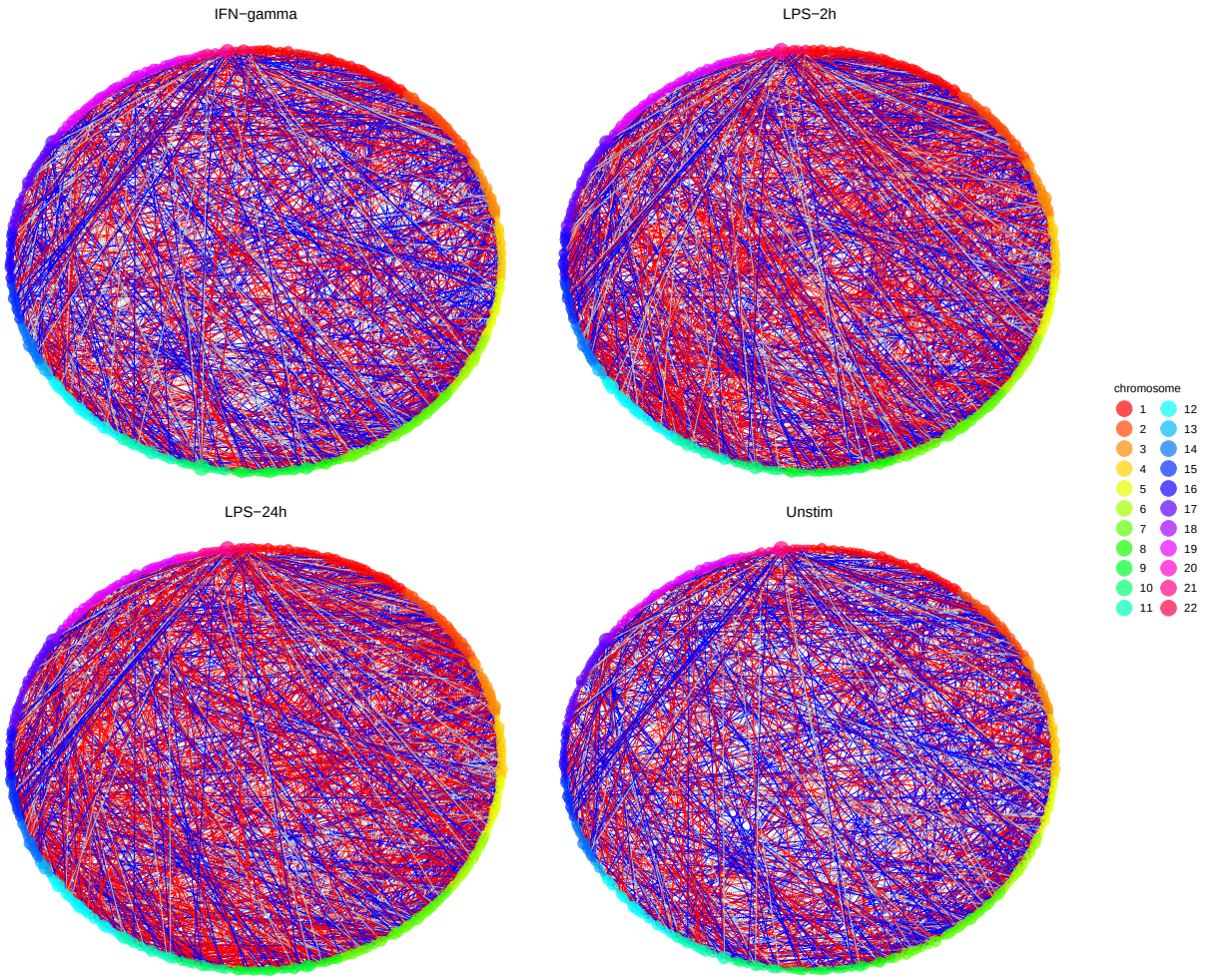


Figure 15: Subnetwork showing all edges of the joint estimate

Network intersection

All edges

As we see from the intersection plot, the networks have a fair share of edges in common. After the intersection of all conditions, the intersection with the largest edge set is LPS-2h and LPS-24h. This agrees with what we have seen previously.

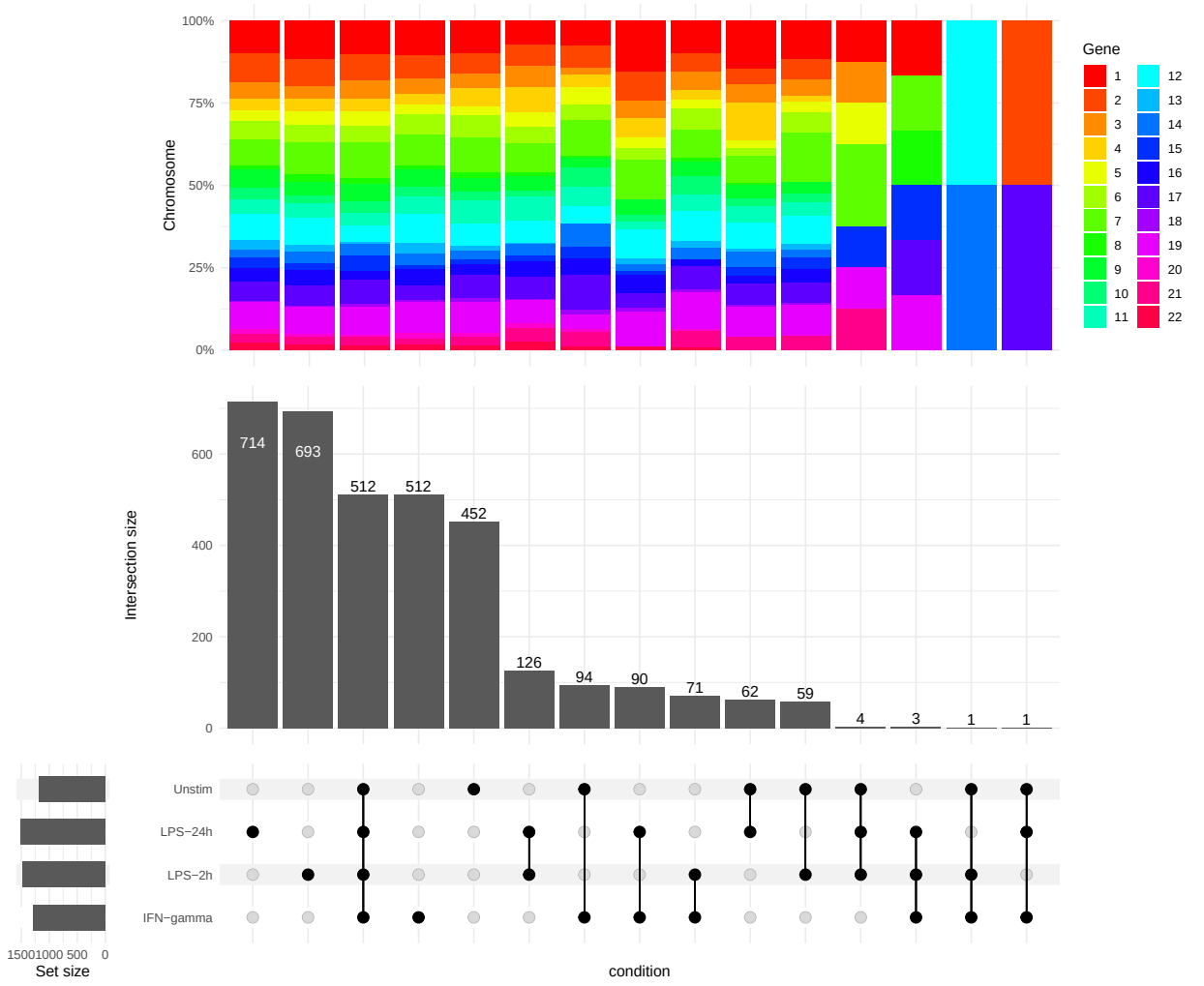


Figure 16: Intersection size of edges of different condition network combinations. The distribution of chromosome location of the genes corresponding to the common edges is shown as well.

Marked by hotspot contol

Hubs

joint vs single GHS

Comparing to the single GHS networks, we see that the joint networks have more in common. This is as expected, that we with the joint method are able to identify what is common and borrow information. From

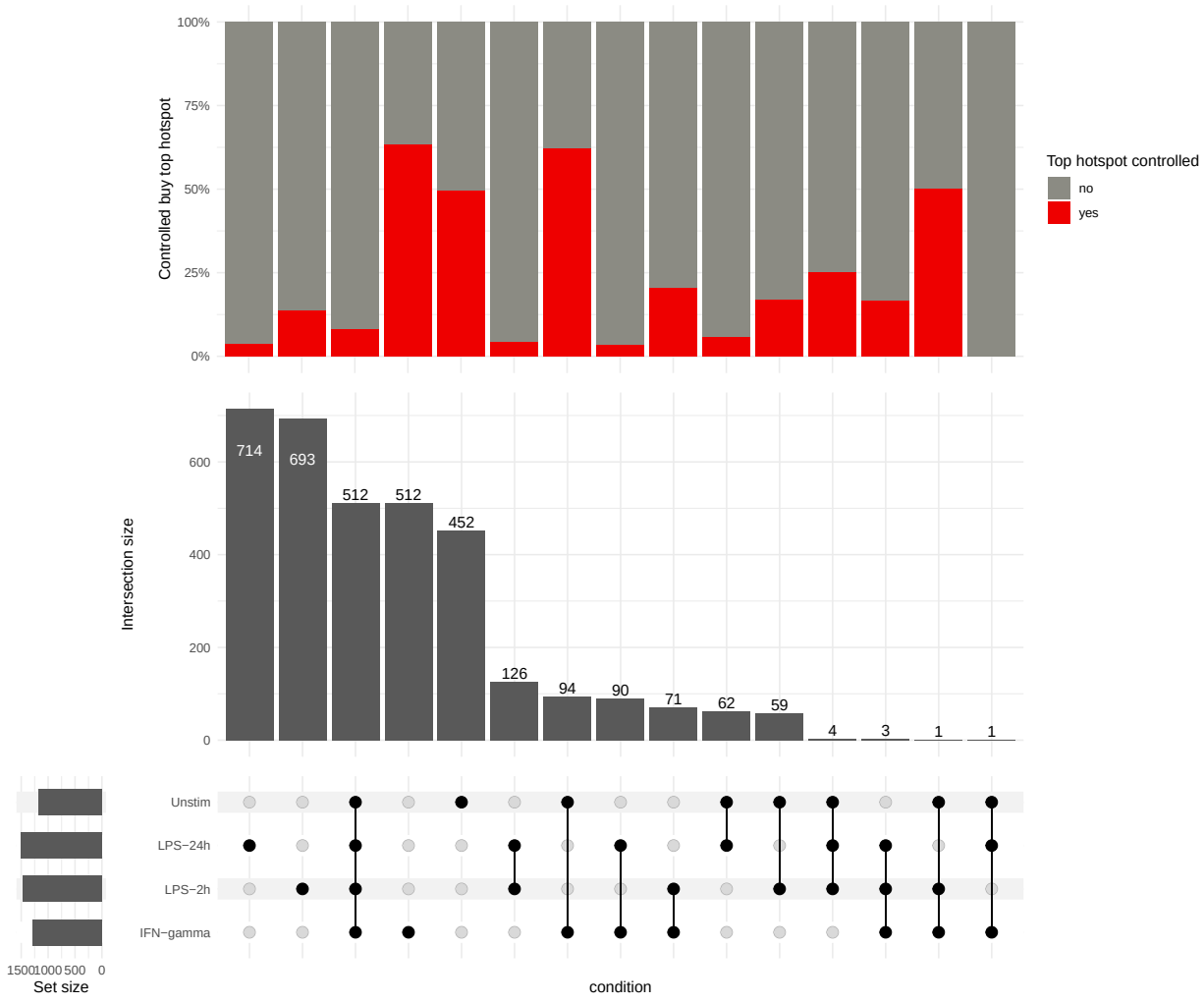


Figure 17: Intersection size of edges of different condition network combinations. Whether the corresponding hubs (genes) are controlled by the top hotspot in all conditions in the intersection is shown as well.

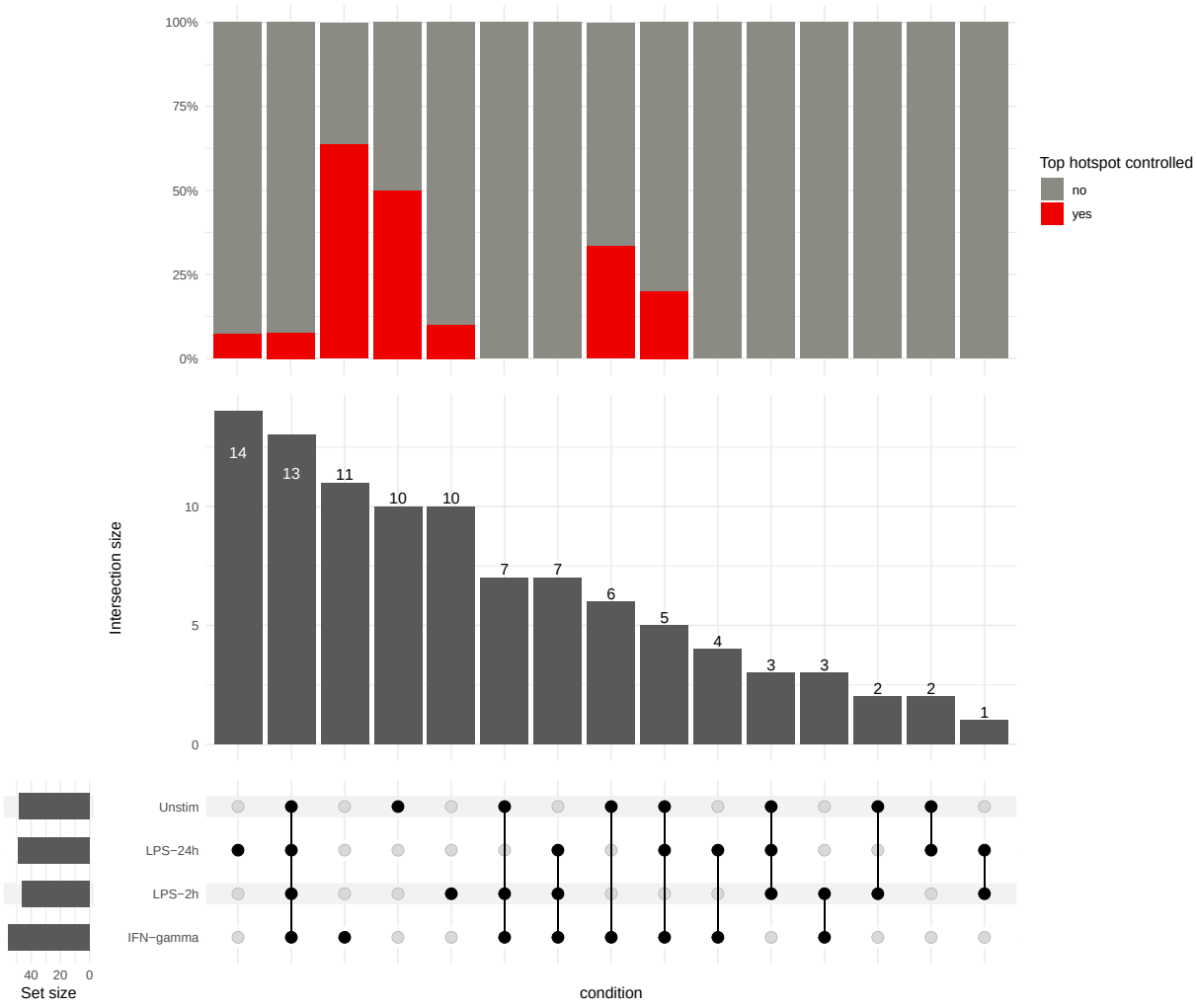


Figure 18: Intersection size of the hubs of the different condition network combinations. Whether the corresponding hubs (genes) are controlled by the top hotspot within each condition in the intersection is shown as well.

the upset plots, we see that much individuality is captured as well. The log likelihood improvement supports that the estimates have improved compared to the single-network method.

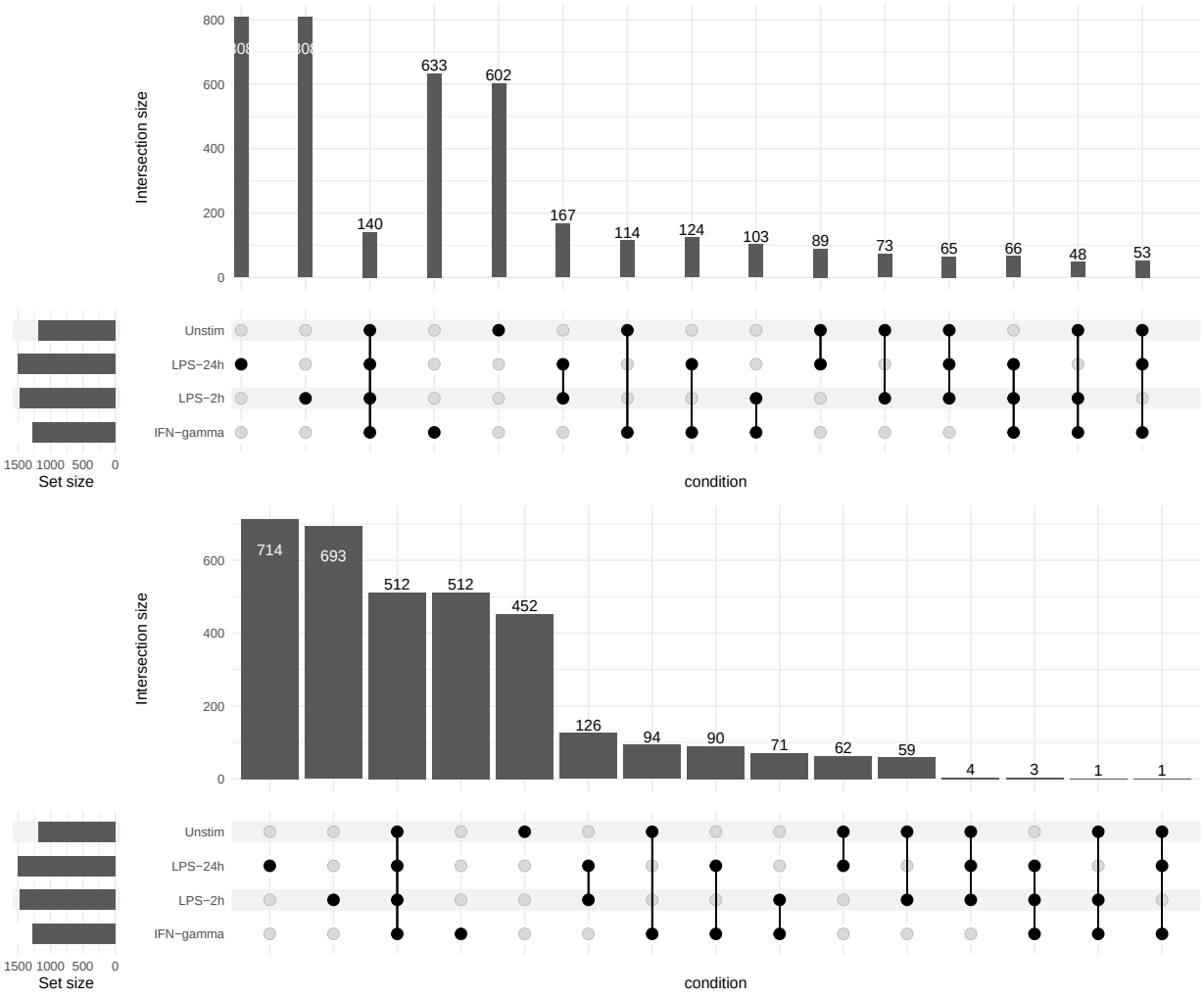


Figure 19: Intersection size of edges of different condition network combinations in single vs joint GHS.

Order of neighbourhood to cis genes

Hubs

All genes

We see that in the unstimulated group, more genes are further than three links from the cis genes compared to the other conditions. Many genes do not have any paths to the cis genes at all.

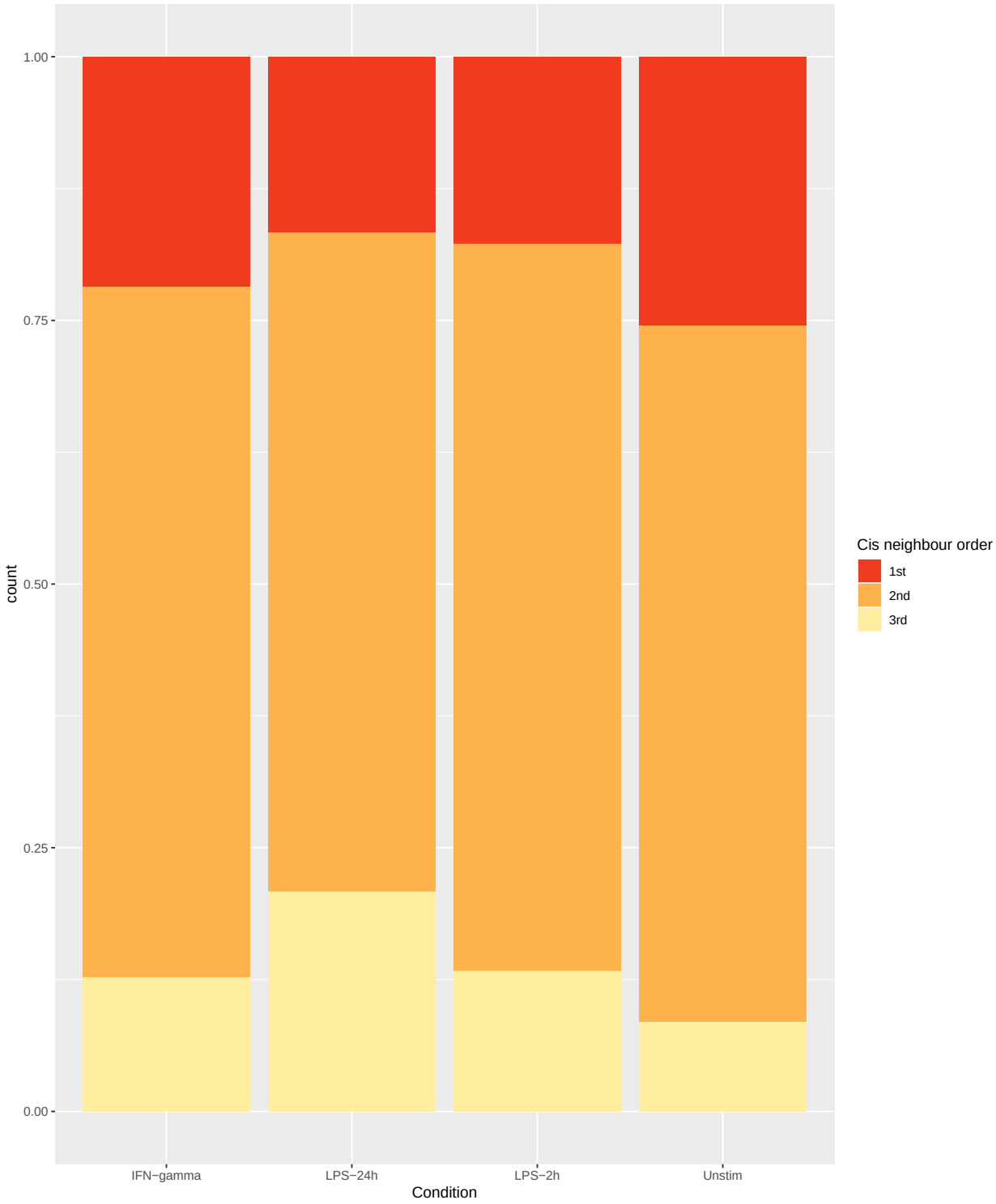


Figure 20: The order of the top hubs in each condition. By order, we mean the order of neighbourhood to either of the top hotspot mediated cis genes LYZ and YEATS4.

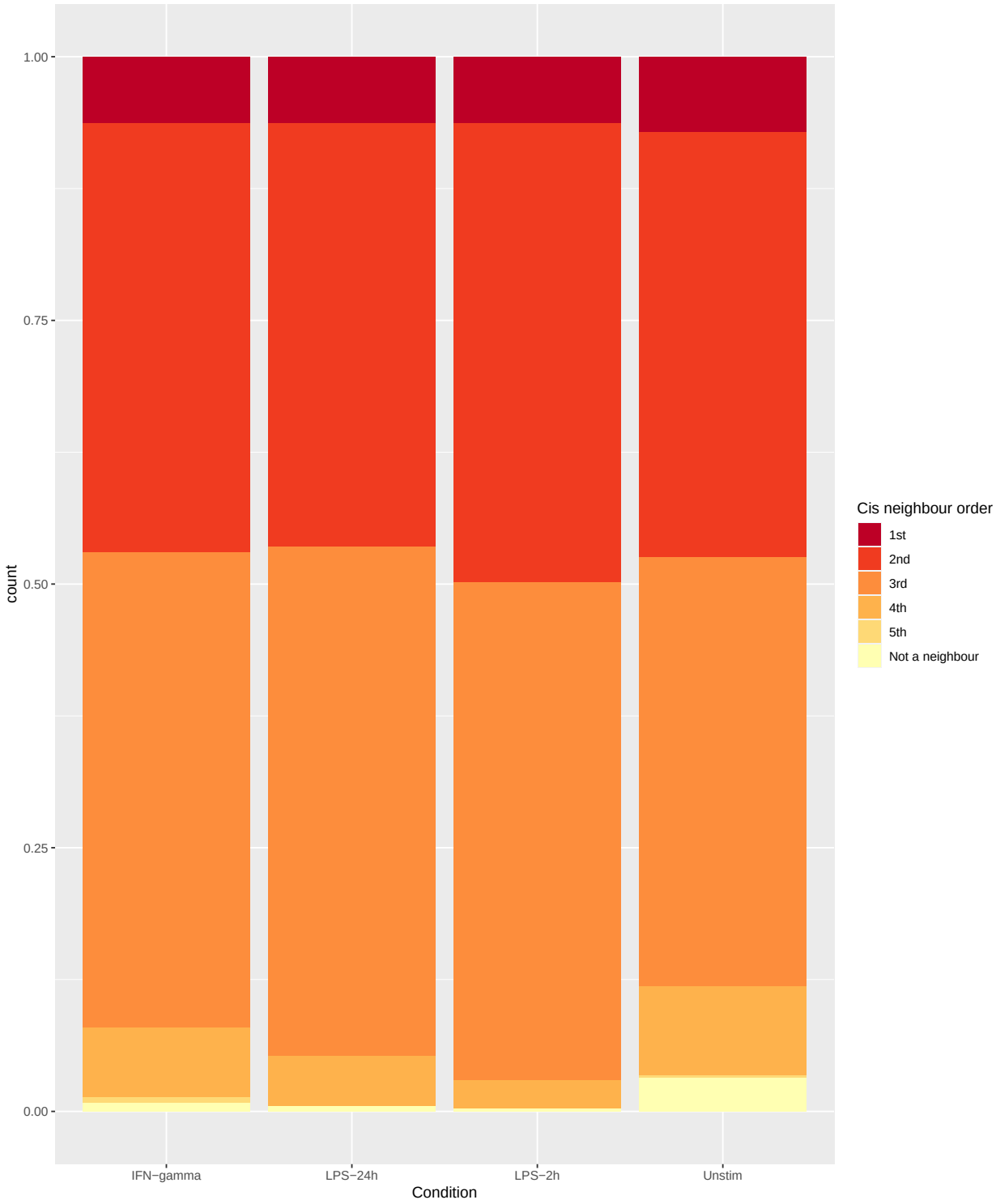


Figure 21: The order of all genes in each condition. By order, we mean the order of neighbourhood to either of the top hotspot mediated cis genes LYZ and YEATS4.

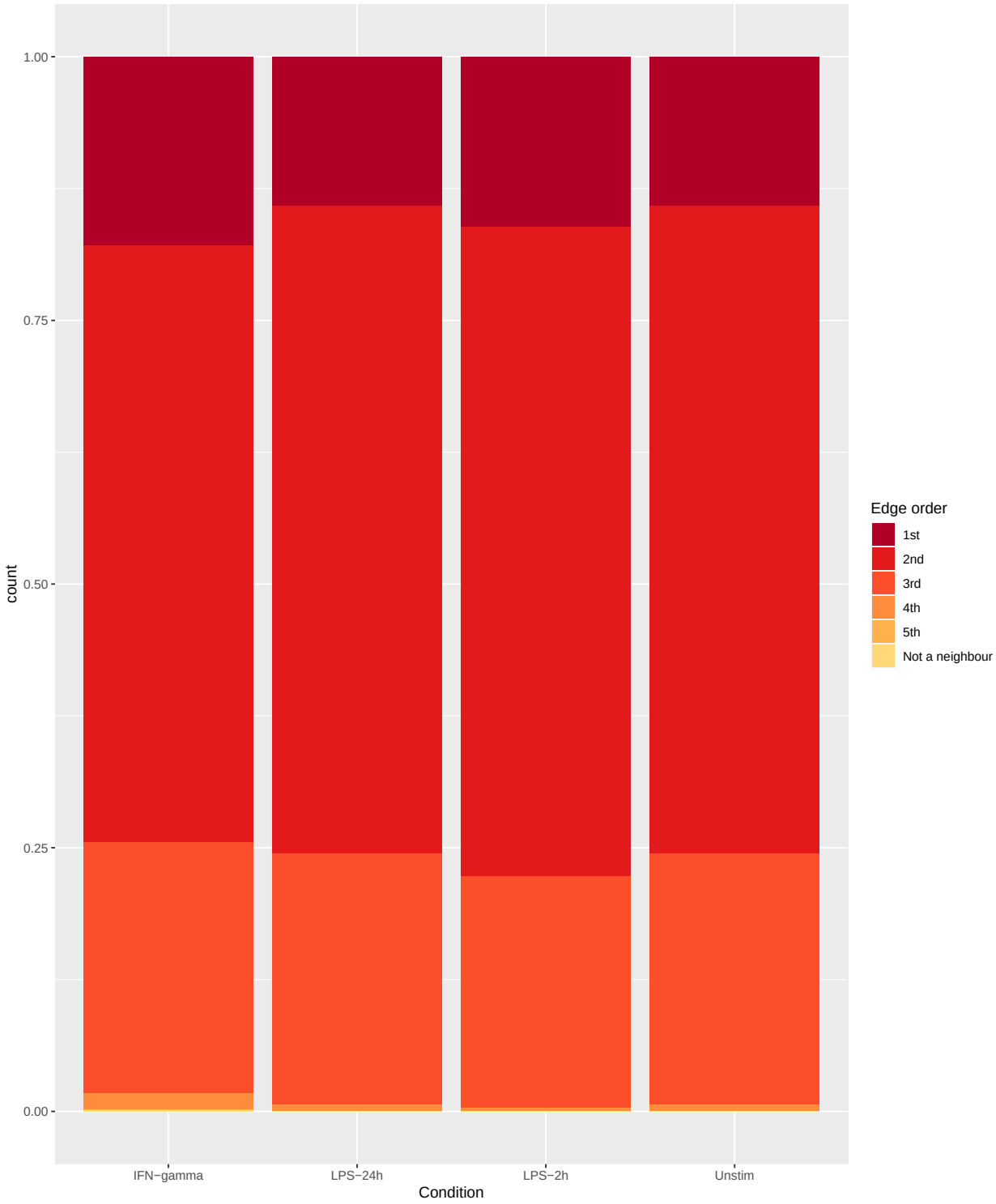


Figure 22: The order of all edges in each condition. By order, we mean the number of links to either the top hotspot mediated cis genes LYZ and YEATS4.

Edges

Common parameter estimates

As we see, while many edges are found to be in common so that information is shared, there are still many network-specific edges that have been captured. Thanks to the heavy horseshoe tail, we were able to capture these edges even though no common information about them was found between the conditions.

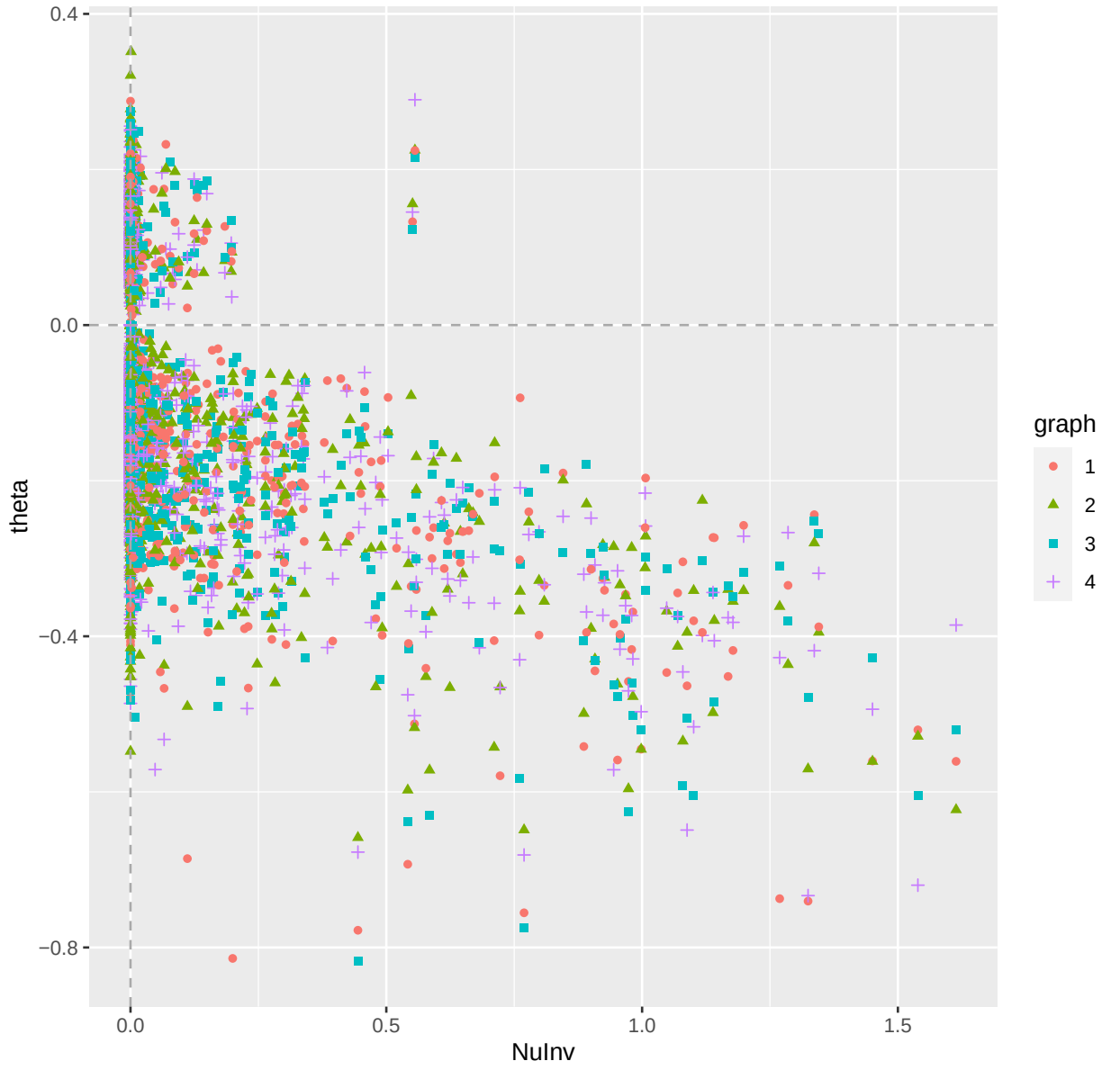


Figure 23: Estimated precision matrix elements of each condition plotted against their common hyperparameter

Are the top hotspot mediated gene subnetworks denser?

Subnetworks of top hotspot mediated genes

The subnetworks of the top hotspot mediated genes were found to be sigificantly denser than the overall networks in all conditions, with empirical p-values being 0.0001 in all conditions except IFN-GAMMA, which had p-value 0.08.

Neighbour genes of cis genes

We see that YEATS4 does not really have more mediated neighbours than the overall fraction, except in LPS-24h. CREB1 has more mediated neighbours in all conditions, and in fact *all* its neighbours are mediated in IFN-GAMMA and the unstimulated cells. LYZ also has more mediated neighbours in all conditions, though not very much more in IFN-GAMMA.

In LPS-24h, all genes have more mediated neighbours. The number of top hotspot mediated genes is lower in this group.

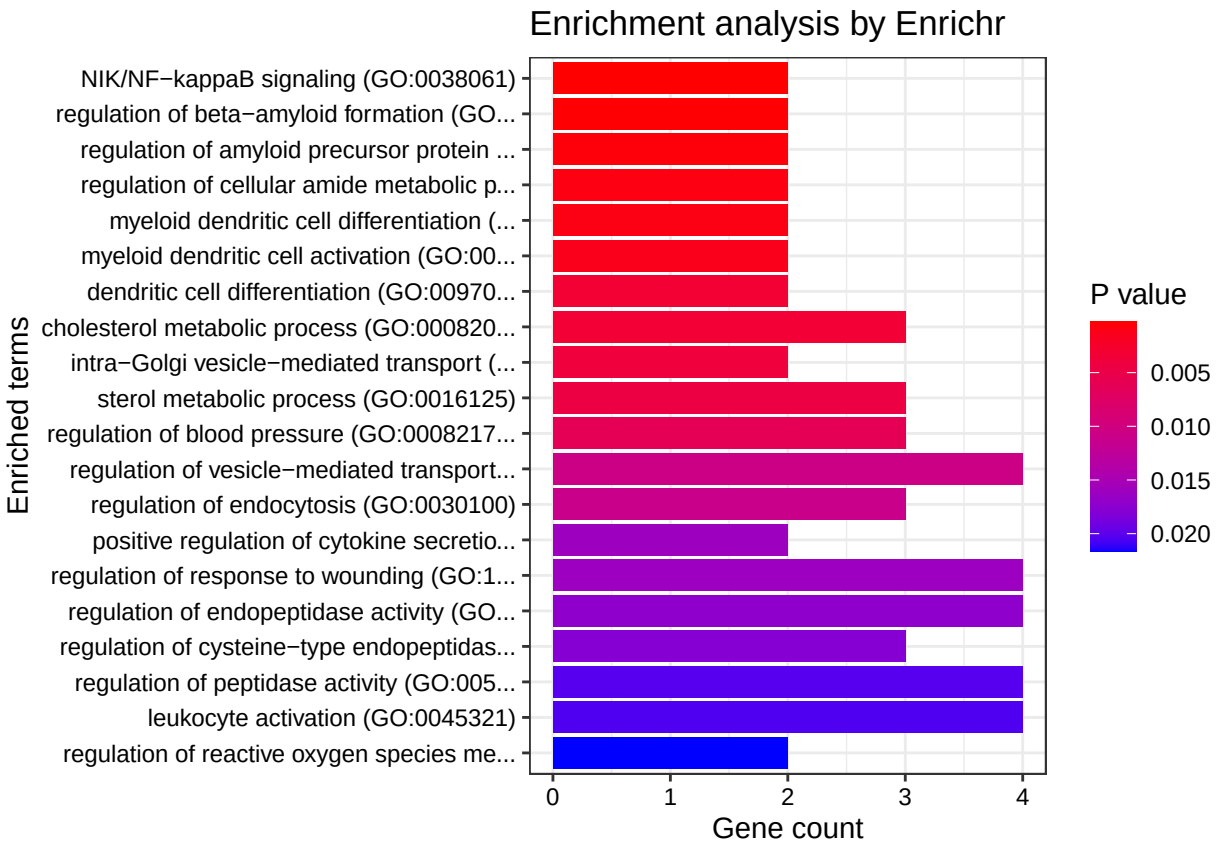
Table 1: Fraction of neighbour genes of LYZ, YEATS4 and CREB1 that are mediated by the top hotspot, as well as the overall fraction, in different conditions

	IFN-gamma	LPS-2h	LPS-24h	Unstim
Overall	0.772	0.231	0.042	0.564
LYZ neighbourhood	0.778	0.417	0.333	0.769
YEATS4 neighbourhood	0.625	0.167	0.222	0.429
CREB1 neighbourhood	1.000	0.600	0.333	1.000

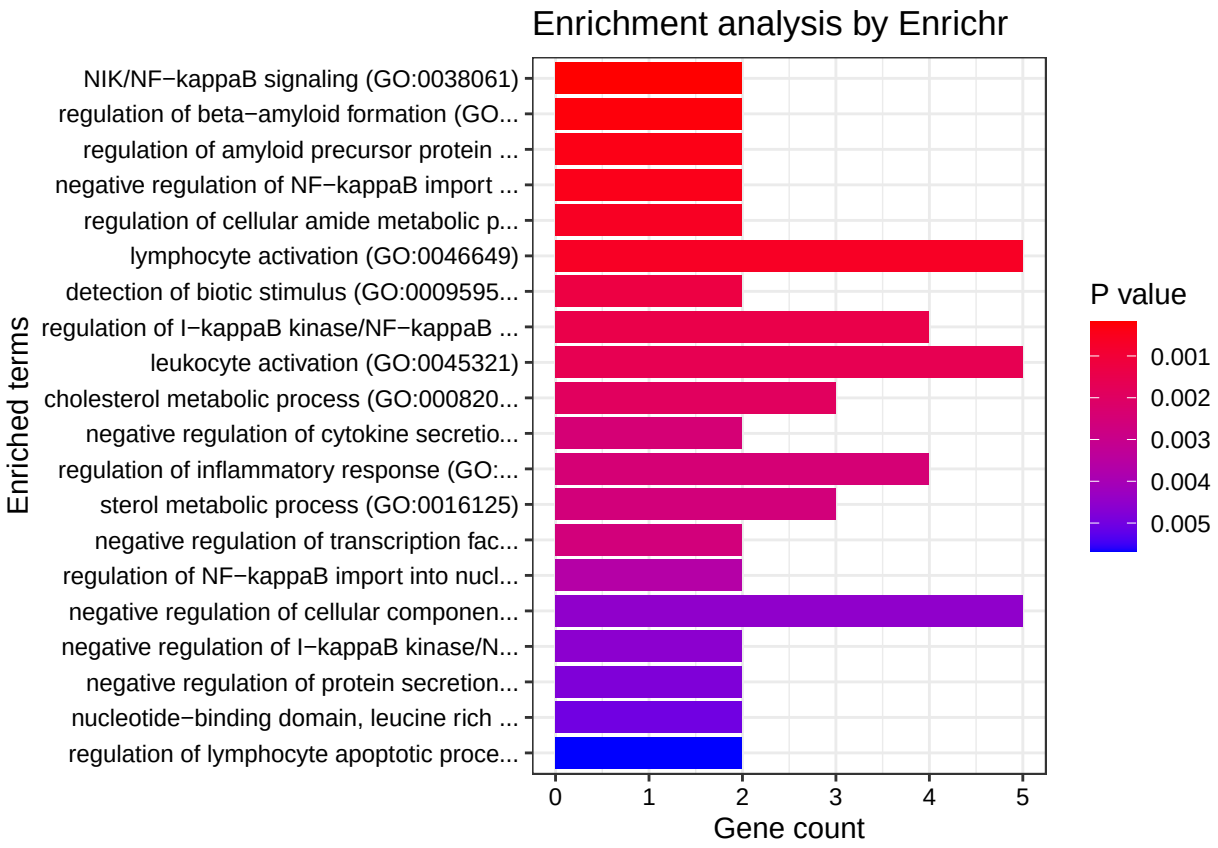
Enrichment analysis of hubs with enrichR

Exploring the top hubs of each network using enrichR. We use the data base "GO_Biological_Process_2015".

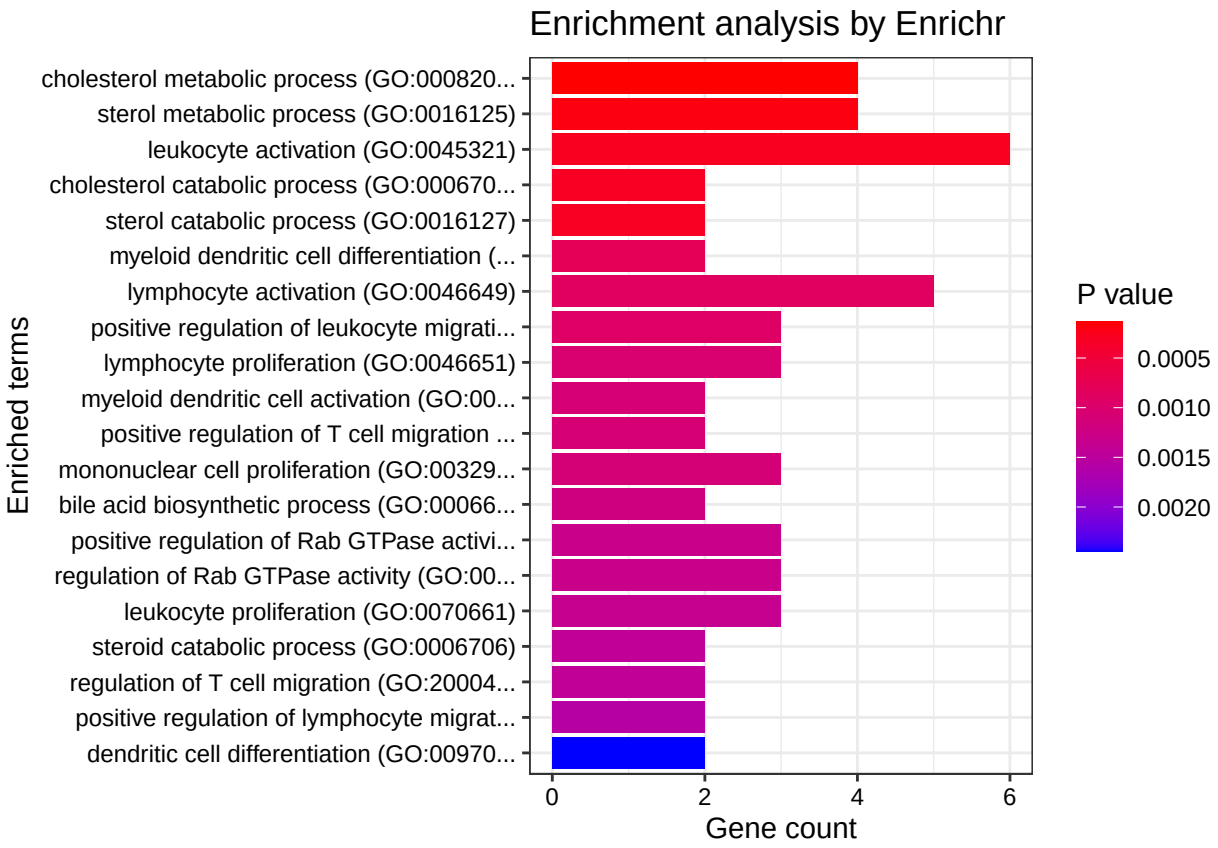
IFN-GAMMA


Enrichment analysis by Enrichr
NIK/NF−kappaB signaling (GO:0038061)
regulation of beta−amyloid formation (GO...
regulation of amyloid precursor protein ...
regulation of cellular amide metabolic p...
myeloid dendritic cell differentiation (...
myeloid dendritic cell activation (GO:00...
dendritic cell differentiation (GO:00970...
cholesterol metabolic process (GO:000820...
intra−Golgi vesicle−mediated transport (...
sterol metabolic process (GO:0016125)
regulation of blood pressure (GO:0008217...
regulation of vesicle−mediated transport...
regulation of endocytosis (GO:0030100)
positive regulation of cytokine secretio...
regulation of response to wounding (GO:1...
regulation of endopeptidase activity (GO...
regulation of cysteine−type endopeptidas...
regulation of peptidase activity (GO:005...
leukocyte activation (GO:0045321)
regulation of reactive oxygen species me...
Enriched terms
0
1
2
3
4
Gene count
P value
0.005
0.010
0.015
0.020

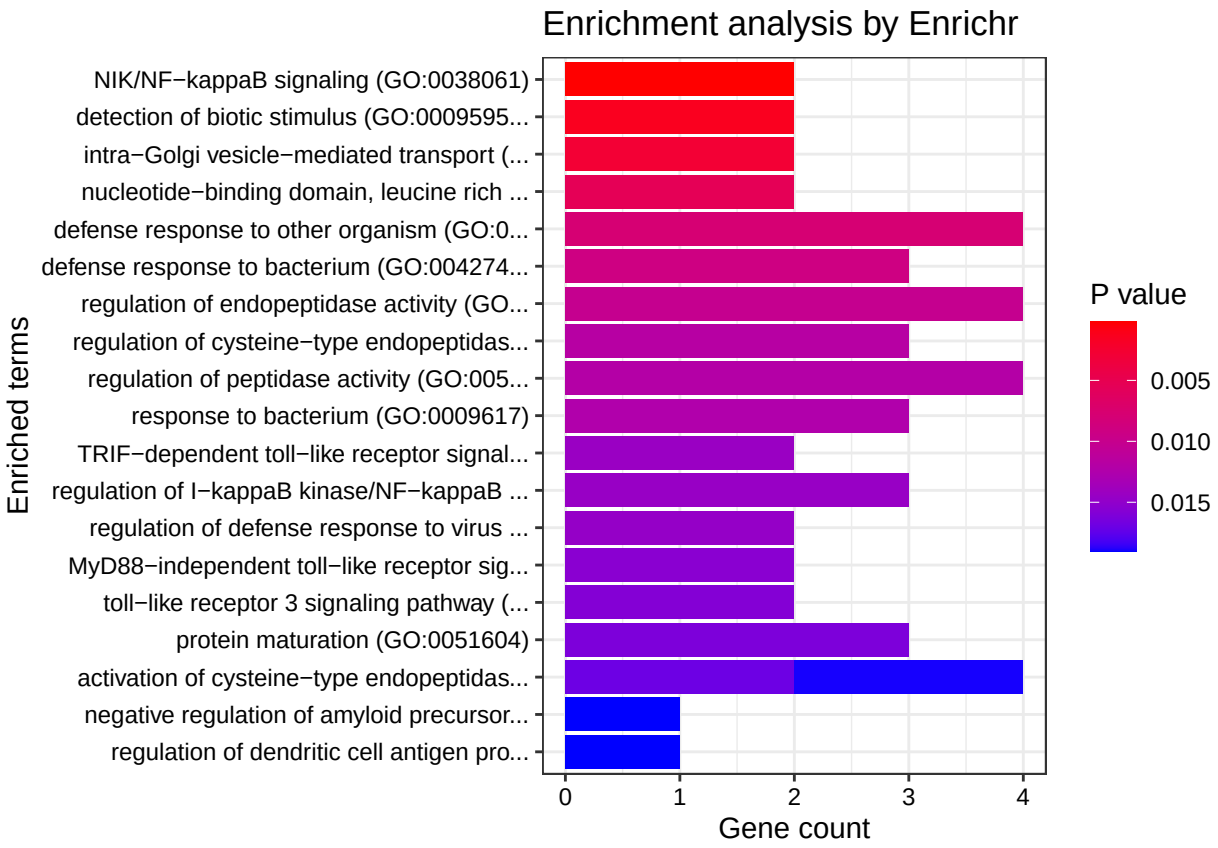
LPS-2h


Enrichment analysis by Enrichr
NIK/NF−kappaB signaling (GO:0038061)
regulation of beta−amyloid formation (GO...
regulation of amyloid precursor protein ...
negative regulation of NF−kappaB import ...
regulation of cellular amide metabolic p...
lymphocyte activation (GO:0046649)
detection of biotic stimulus (GO:0009595...
regulation of I−kappaB kinase/NF−kappaB ...
leukocyte activation (GO:0045321)
cholesterol metabolic process (GO:000820...
negative regulation of cytokine secretio...
regulation of inflammatory response (GO:...
sterol metabolic process (GO:0016125)
negative regulation of transcription fac...
regulation of NF−kappaB import into nucl...
negative regulation of cellular componen...
negative regulation of I−kappaB kinase/N...
negative regulation of protein secretion...
nucleotide−binding domain, leucine rich ...
regulation of lymphocyte apoptotic proce...
Enriched terms
0
1
2
3
4
5
Gene count
P value
0.001
0.002
0.003
0.004
0.005

LPS-24h


Enrichment analysis by Enrichr
cholesterol metabolic process (GO:000820...
sterol metabolic process (GO:0016125)
leukocyte activation (GO:0045321)
cholesterol catabolic process (GO:000670...
sterol catabolic process (GO:0016127)
myeloid dendritic cell differentiation (...
lymphocyte activation (GO:0046649)
positive regulation of leukocyte migrati...
lymphocyte proliferation (GO:0046651)
myeloid dendritic cell activation (GO:00...
positive regulation of T cell migration ...
mononuclear cell proliferation (GO:00329...
bile acid biosynthetic process (GO:00066...
positive regulation of Rab GTPase activi...
regulation of Rab GTPase activity (GO:00...
leukocyte proliferation (GO:0070661)
steroid catabolic process (GO:0006706)
regulation of T cell migration (GO:20004...
positive regulation of lymphocyte migrat...
dendritic cell differentiation (GO:00970...
Enriched terms
0
2
4
6
Gene count
P value
0.0005
0.0010
0.0015
0.0020

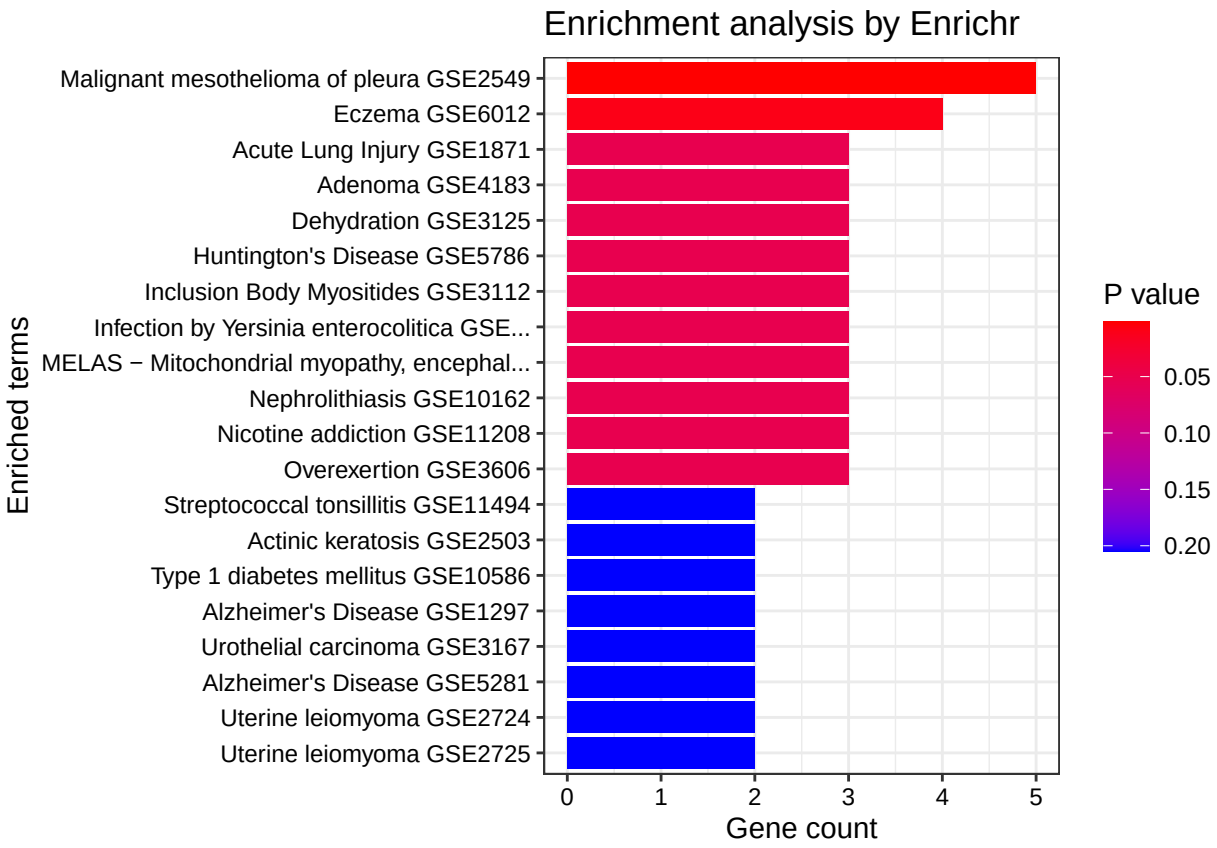
Unstimulated



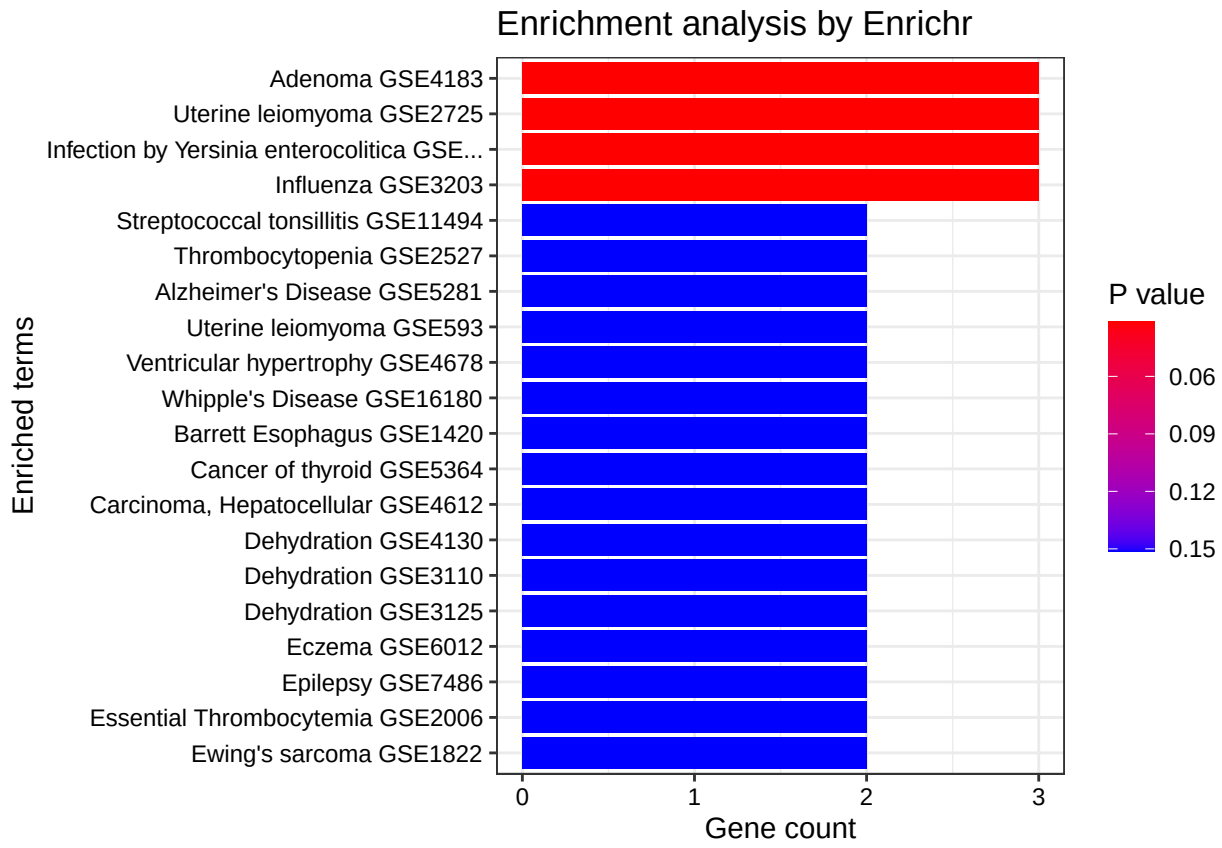
Enrichment of up-regulated genes in disease signatures

Exploring the top hubs of each network using enrichR. We use the data base "Disease_Signatures_from_GEO_up_2014".

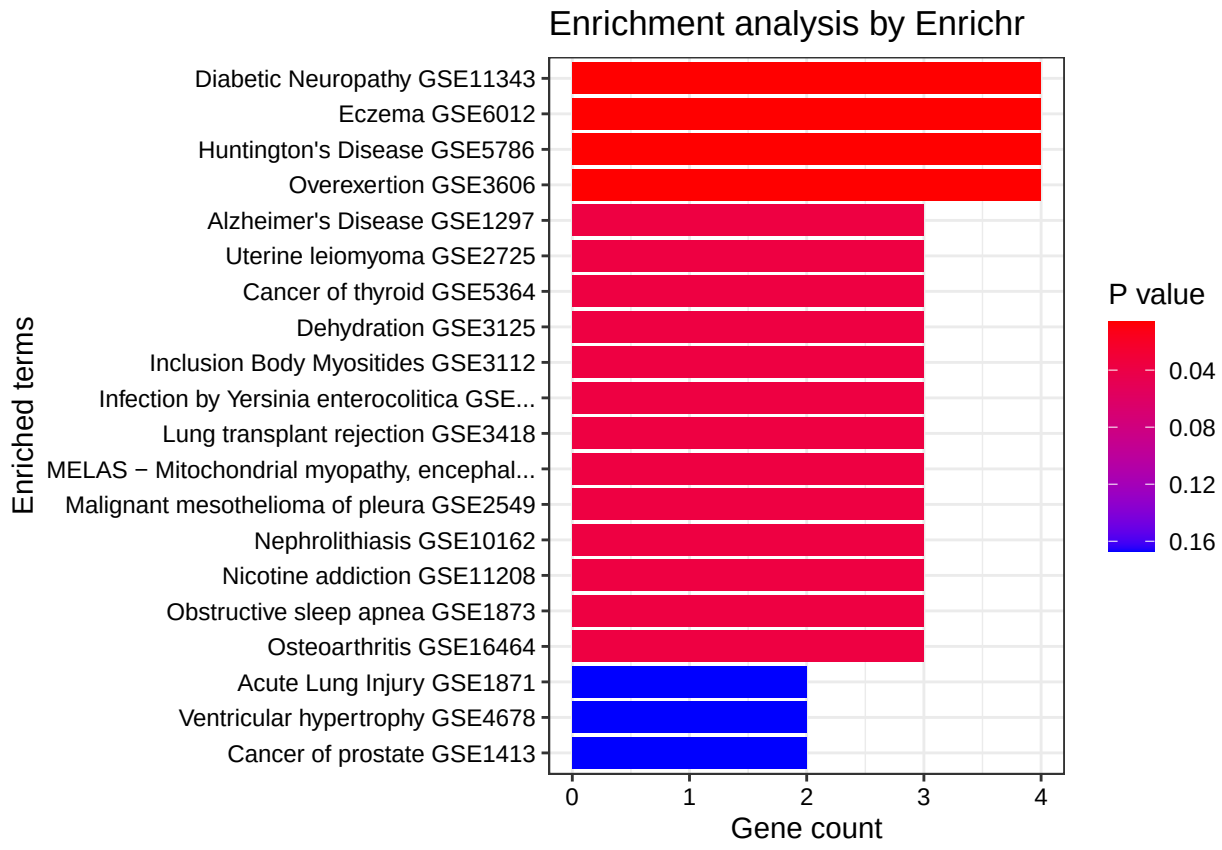
IFN-GAMMA


Enrichment analysis by Enrichr
Malignant mesothelioma of pleura GSE2549
Eczema GSE6012
Acute Lung Injury GSE1871
Adenoma GSE4183
Dehydration GSE3125
Huntington's Disease GSE5786
Inclusion Body Myositides GSE3112
Infection by Yersinia enterocolitica GSE...
MELAS − Mitochondrial myopathy, encephal...
Nephrolithiasis GSE10162
Nicotine addiction GSE11208
Overexertion GSE3606
Streptococcal tonsillitis GSE11494
Actinic keratosis GSE2503
Type 1 diabetes mellitus GSE10586
Alzheimer's Disease GSE1297
Urothelial carcinoma GSE3167
Alzheimer's Disease GSE5281
Uterine leiomyoma GSE2724
Uterine leiomyoma GSE2725
Enriched terms
0
1
2
3
4
5
Gene count
P value
0.05
0.10
0.15
0.20

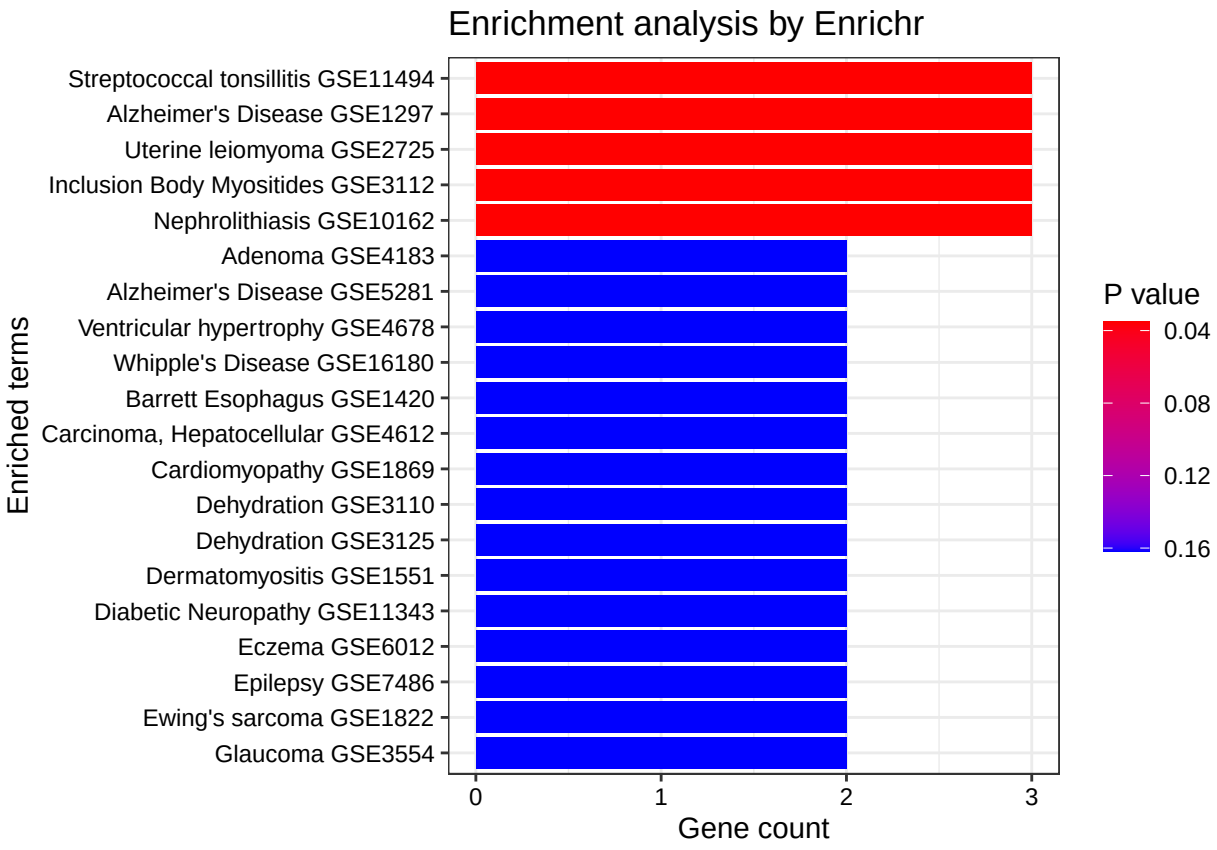
LPS-2h


Enrichment analysis by Enrichr
Adenoma GSE4183
Uterine leiomyoma GSE2725
Infection by Yersinia enterocolitica GSE...
Influenza GSE3203
Streptococcal tonsillitis GSE11494
Thrombocytopenia GSE2527
Alzheimer's Disease GSE5281
Uterine leiomyoma GSE593
Ventricular hypertrophy GSE4678
Whipple's Disease GSE16180
Barrett Esophagus GSE1420
Cancer of thyroid GSE5364
Carcinoma, Hepatocellular GSE4612
Dehydration GSE4130
Dehydration GSE3110
Dehydration GSE3125
Eczema GSE6012
Epilepsy GSE7486
Essential Thrombocytemia GSE2006
Ewing's sarcoma GSE1822
Enriched terms
0
1
2
3
Gene count
P value
0.06
0.09
0.12
0.15

LPS-24h


Enrichment analysis by Enrichr
Diabetic Neuropathy GSE11343
Eczema GSE6012
Huntington's Disease GSE5786
Overexertion GSE3606
Alzheimer's Disease GSE1297
Uterine leiomyoma GSE2725
Cancer of thyroid GSE5364
Dehydration GSE3125
Inclusion Body Myositides GSE3112
Infection by Yersinia enterocolitica GSE...
Lung transplant rejection GSE3418
MELAS − Mitochondrial myopathy, encephal...
Malignant mesothelioma of pleura GSE2549
Nephrolithiasis GSE10162
Nicotine addiction GSE11208
Obstructive sleep apnea GSE1873
Osteoarthritis GSE16464
Acute Lung Injury GSE1871
Ventricular hypertrophy GSE4678
Cancer of prostate GSE1413
Enriched terms
0
1
2
3
4
Gene count
P value
0.04
0.08
0.12
0.16

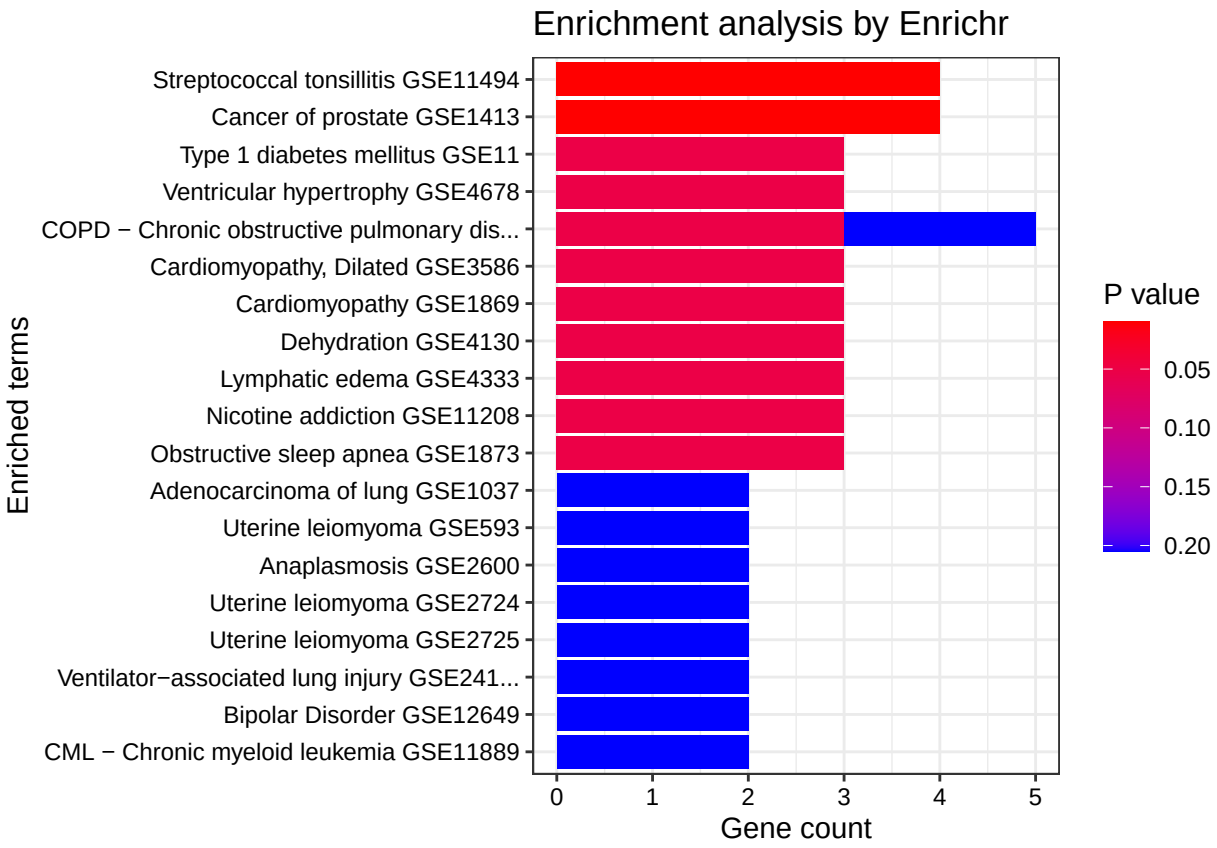
Unstimulated



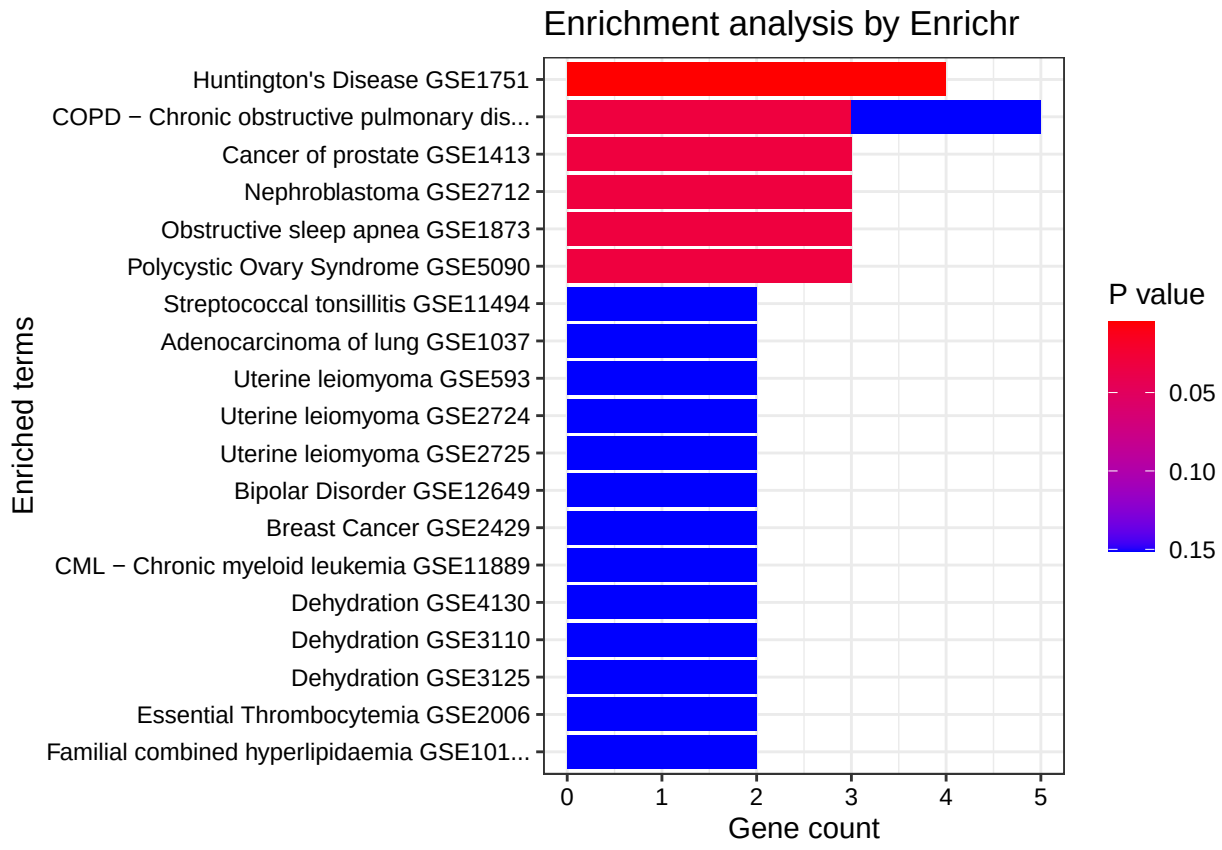
Enrichment of down-regulated genes in disease signatures

Exploring the top hubs of each network using enrichR. We use the data base "Disease_Signatures_from_GEO_down_2014".

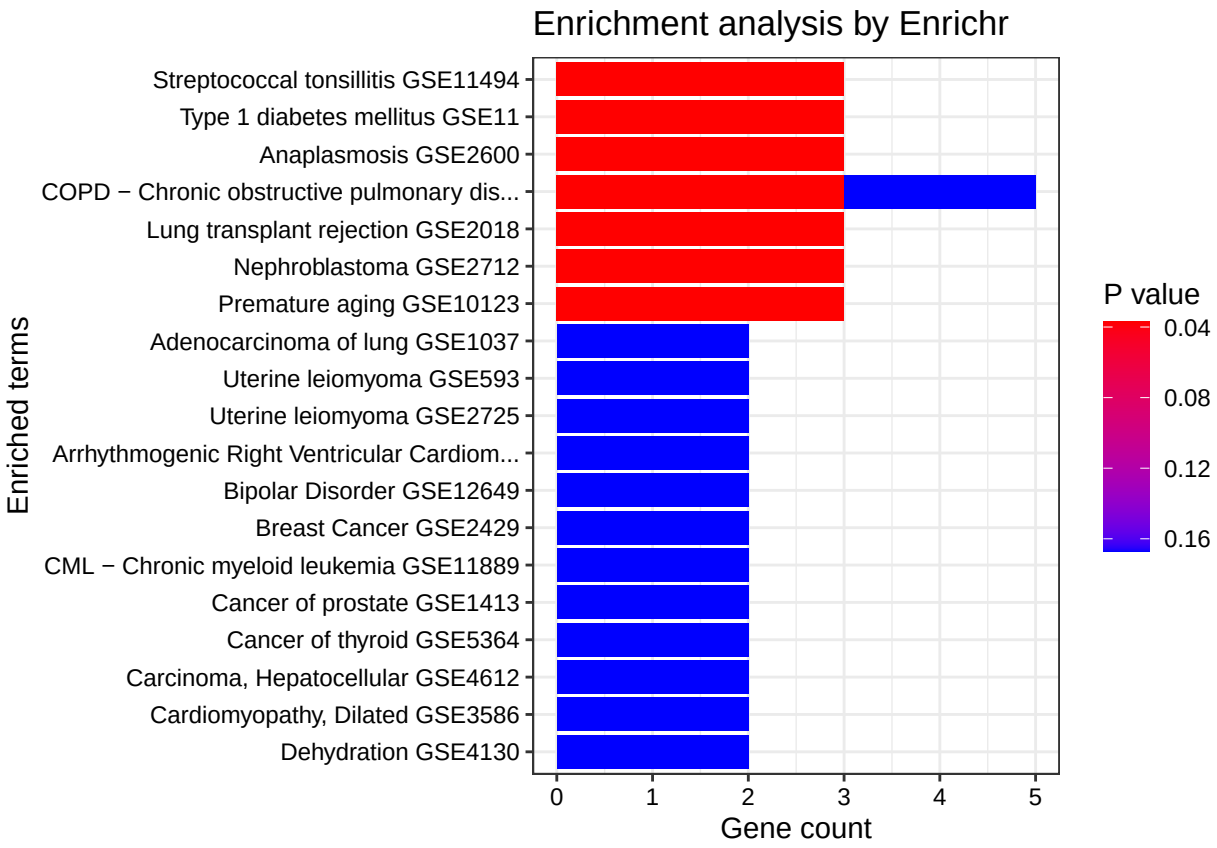
IFN-GAMMA


Enrichment analysis by Enrichr
Streptococcal tonsillitis GSE11494
Cancer of prostate GSE1413
Type 1 diabetes mellitus GSE11
Ventricular hypertrophy GSE4678
COPD − Chronic obstructive pulmonary dis...
Cardiomyopathy, Dilated GSE3586
Cardiomyopathy GSE1869
Dehydration GSE4130
Lymphatic edema GSE4333
Nicotine addiction GSE11208
Obstructive sleep apnea GSE1873
Adenocarcinoma of lung GSE1037
Uterine leiomyoma GSE593
Anaplasmosis GSE2600
Uterine leiomyoma GSE2724
Uterine leiomyoma GSE2725
Ventilator−associated lung injury GSE241...
Bipolar Disorder GSE12649
CML − Chronic myeloid leukemia GSE11889
Enriched terms
0
1
2
3
4
5
Gene count
P value
0.05
0.10
0.15
0.20

LPS-2h


Enrichment analysis by Enrichr
Huntington's Disease GSE1751
COPD − Chronic obstructive pulmonary dis...
Cancer of prostate GSE1413
Nephroblastoma GSE2712
Obstructive sleep apnea GSE1873
Polycystic Ovary Syndrome GSE5090
Streptococcal tonsillitis GSE11494
Adenocarcinoma of lung GSE1037
Uterine leiomyoma GSE593
Uterine leiomyoma GSE2724
Uterine leiomyoma GSE2725
Bipolar Disorder GSE12649
Breast Cancer GSE2429
CML − Chronic myeloid leukemia GSE11889
Dehydration GSE4130
Dehydration GSE3110
Dehydration GSE3125
Essential Thrombocytemia GSE2006
Familial combined hyperlipidaemia GSE101...
Enriched terms
0
1
2
3
4
5
Gene count
P value
0.05
0.10
0.15

LPS-24h


Enrichment analysis by Enrichr
Streptococcal tonsillitis GSE11494
Type 1 diabetes mellitus GSE11
Anaplasmosis GSE2600
COPD − Chronic obstructive pulmonary dis...
Lung transplant rejection GSE2018
Nephroblastoma GSE2712
Premature aging GSE10123
Adenocarcinoma of lung GSE1037
Uterine leiomyoma GSE593
Uterine leiomyoma GSE2725
Arrhythmogenic Right Ventricular Cardiom...
Bipolar Disorder GSE12649
Breast Cancer GSE2429
CML − Chronic myeloid leukemia GSE11889
Cancer of prostate GSE1413
Cancer of thyroid GSE5364
Carcinoma, Hepatocellular GSE4612
Cardiomyopathy, Dilated GSE3586
Dehydration GSE4130
Enriched terms
0
1
2
3
4
5
Gene count
P value
0.04
0.08
0.12
0.16

Unstimulated