

# Canadian Bioinformatics Workshops

[www.bioinformatics.ca](http://www.bioinformatics.ca)

[bioinformaticsdotca.github.io](https://bioinformaticsdotca.github.io)

Creative Commons

This page is available in the following languages:

Afrikaans Azərbaycanca Català Dansk Deutsch Ελληνικά English English (CA) English (GB) English (US) Esperanto  
Español Castellano (AR) Español (CL) Castellano (CO) Español (Ecuador) Castellano (MX) Castellano (PE)  
Euskara Suomi français français (CA) Galego ગુજરાતી hrvatski Magyar Italiano 日本語 한국어 Macdonian Melayu  
Nederlands Norsk Sesotho sa Leboa polski Português română slovenski jezik српски (latinica) Sotho svenska  
中文 華語 (台灣) isiZulu



### Attribution-Share Alike 2.5 Canada

#### You are free:



to Share — to copy, distribute and transmit the work



to Remix — to adapt the work



#### Under the following conditions:



**Attribution.** You must attribute the work in the manner specified by the author or licensor (but not in any way that suggests that they endorse you or your use of the work).



**Share Alike.** If you alter, transform, or build upon this work, you may distribute the resulting work only under the same or similar licence to this one.

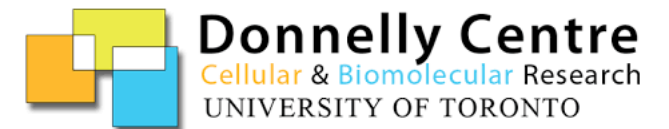
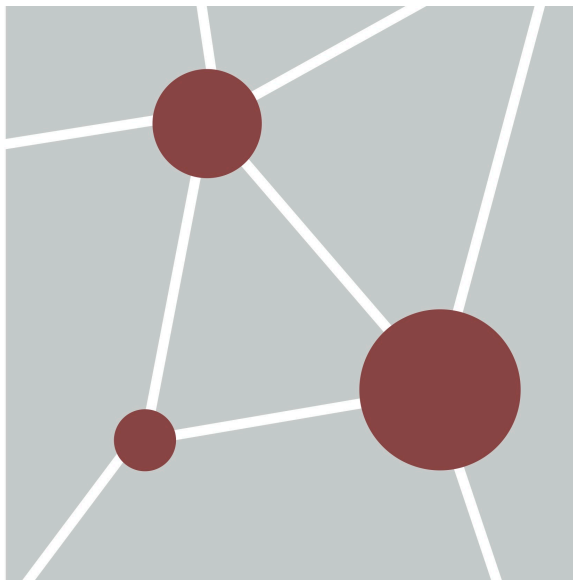
- For any reuse or distribution, you must make clear to others the licence terms of this work.
- Any of the above conditions can be waived if you get permission from the copyright holder.
- The author's moral rights are retained in this licence.

[Disclaimer](#)

Your fair dealing and other rights are in no way affected by the above.  
This is a human-readable summary of the Legal Code (the full licence) available in the following languages:  
[English](#) [French](#)

[Learn how to distribute your work using this licence](#)

# Module 3 Enrichment Maps



# Learning Objectives

- By the end of this lecture, you will:
  - Understand...
  - Be able to define...
  - Know...

# Learning Objectives

- By the end of this lecture, you will:
  - Understand how to transform your enrichment results from g:profiler, gsea or other enrichment algorithms to a Network
  - Understand the differences between a network and an Enrichment Map
  - Be able to summarize your enrichment results with annotation using the Autoannotate App.

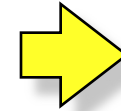
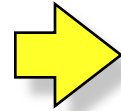
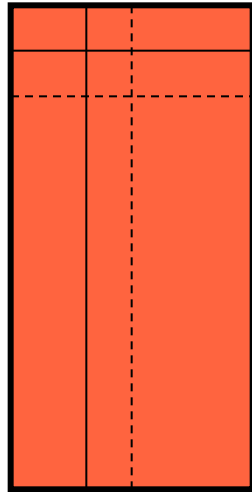
# Results of Lab 2 – Enrichment Results

GO.ID	Description	p.Val	FDR	Phenotype	Genes									
GO:0060255	regulation of macromolecule metabolic proces	2.34E-30	2.34E-30		1 TP53,PTEN,APC,VHL,KRAS,ARID1A,EGFR,NF1,PIK3R1,CDKN2A,GATA3,RB1,NOTCH									
GO:0051171	regulation of nitrogen compound metabolic pr	6.44E-30	6.44E-30		1 TP53,PTEN,APC,VHL,KRAS,ARID1A,EGFR,NF1,PIK3R1,CDKN2A,GATA3,RB1,NOTCH									
GO:0031323	regulation of cellular metabolic process	2.80E-29	2.80E-29		1 TP53,PIK3CA,PTEN,APC,VHL,KRAS,ARID1A,EGFR,NF1,PIK3R1,CDKN2A,GATA3,RB1,									
GO:0019222	regulation of metabolic process	3.50E-29	3.50E-29		1 TP53,PIK3CA,PTEN,APC,VHL,KRAS,ARID1A,EGFR,NF1,PIK3R1,CDKN2A,GATA3,RB1,									
GO:0080090	regulation of primary metabolic process	4.34E-29	4.34E-29		1 TP53,PTEN,APC,VHL,KRAS,ARID1A,EGFR,NF1,PIK3R1,CDKN2A,GATA3,RB1,NOTCH									
GO:0051173	positive regulation of nitrogen compound met	1.35E-28	1.35E-28		1 TP53,PTEN,APC,VHL,KRAS,ARID1A,EGFR,PIK3R1,CDKN2A,GATA3,RB1,NOTCH1,FB.									
GO:0010604	positive regulation of macromolecule metabol	1.51E-28	1.51E-28		1 TP53,PTEN,APC,VHL,KRAS,ARID1A,EGFR,PIK3R1,CDKN2A,GATA3,RB1,NOTCH1,FB.									
GO:1903506	regulation of nucleic acid-templated transcript	5.40E-27	5.40E-27		1 TP53,PTEN,VHL,ARID1A,EGFR,PIK3R1,CDKN2A,GATA3,RB1,NOTCH1,FBXW7,CTNN									
GO:2001141	regulation of RNA biosynthetic process	5.40E-27	5.40E-27		1 TP53,PTEN,VHL,ARID1A,EGFR,PIK3R1,CDKN2A,GATA3,RB1,NOTCH1,FBXW7,CTNN									
GO:0009893	positive regulation of metabolic process					NAME	GS  follo	GS DETAILS	SIZE	ES	NES	NOM p-val	FDR q-val	
GO:0097659	nucleic acid-templated transcription					HALLMARK_INTERFERON_GAMMA_RESPONSE%MSIGDB_C2% HALLMARK_INTERFERON_GAMMA_RESPONSE	HALLMARK_I	Details ...	195	-0.8124175	-3.1241286	0	0	
GO:0032774	RNA biosynthetic process					HALLMARK_INTERFERON_ALPHA_RESPONSE%MSIGDB_C2% HALLMARK_INTERFERON_ALPHA_RESPONSE	HALLMARK_I	Details ...	94	-0.871453	-2.969929	0	0	
GO:0019219	regulation of nucleobase-containing com					RESPONSE TO INTERFERON-GAMMA%GOBP%GO:0034341	RESPONSE T	Details ...	119	-0.7410603	-2.8620038	0	0	
GO:0031325	positive regulation of cellular metabolic p					DEFENSE RESPONSE TO VIRUS%GOBP%GO:0051607	DEFENSE RE	Details ...	102	-0.7547711	-2.8492434	0	0	
GO:0006355	regulation of transcription, DNA-templat					INTERFERON SIGNALING%REACTOME%R-HSA-913531.2	INTERFERON	Details ...	169	-0.7541074	-2.7944484	0	0	
GO:0051252	regulation of RNA metabolic process					HALLMARK_ALLOGRAFT_REJECTION%MSIGDB_C2% HALLMARK_ALLOGRAFT_REJECTION	HALLMARK_	Details ...	181	-0.7311862	-2.784118	0	0	
GO:0006351	transcription, DNA-templated					CELLULAR RESPONSE TO TYPE I INTERFERON%GOBP%GO:0071357	CELLULAR RE	Details ...	50	-0.8616462	-2.7440908	0	0	
GO:0010468	regulation of gene expression					INTERFERON GAMMA SIGNALING%REACTOME%R-HSA-877300.2	INTERFERON	Details ...	81	-0.8194266	-2.7424092	0	0	
GO:2000112	regulation of cellular macromolecule bios					ANTIGEN PROCESSING AND PRESENTATION OF EXOGENOUS PEPTIDE ANTIGEN VIA MHC CLASS I, TAP-DEPEN	ANTIGEN PR	Details ...	72	-0.7555878	-2.7347164	0	0	
						INTERFERON-GAMMA-MEDIATED SIGNALING PATHWAY%GOBP%GO:0060333	INTERFERON	Details ...	64	-0.8510788	-2.702983	0	0	
						TYPE I INTERFERON SIGNALING PATHWAY%GOBP%GO:0060337	TYPE I INTER	Details ...	50	-0.8616462	-2.6922164	0	0	
						RESPONSE TO TYPE I INTERFERON%GOBP%GO:0034340	RESPONSE T	Details ...	54	-0.8517665	-2.6604805	0	0	
						ANTIGEN PROCESSING AND PRESENTATION OF EXOGENOUS PEPTIDE ANTIGEN VIA MHC CLASS I%GOBP%GC	ANTIGEN PR	Details ...	77	-0.7585589	-2.5932803	0	0	
						ER-PHAGOSOME PATHWAY%REACTOME%R-HSA-1236974.4	ER-PHAGOSI	Details ...	81	-0.7237594	-2.589577	0	0	
						COSTIMULATION BY THE CD28 FAMILY%REACTOME%R-HSA-388841.3	COSTIMULAT	Details ...	61	-0.780563	-2.5822923	0	0	
						PROTEASOME DEGRADATION%WIKIPATHWAYS_20200610%WP183% HOMO SAPIENS	PROTEASOM	Details ...	60	-0.7425067	-2.5736508	0	0	
						REGULATION OF LEUKOCYTE PROLIFERATION%GOBP%GO:0070663	REGULATION	Details ...	142	-0.639763	-2.567436	0	0	
						ANTIGEN PROCESSING AND PRESENTATION OF PEPTIDE ANTIGEN VIA MHC CLASS I%GOBP%GO:0002474	ANTIGEN PR	Details ...	94	-0.7129676	-2.560434	0	0	
						ANTIGEN RECEPTOR-MEDIATED SIGNALING PATHWAY%GOBP%GO:0050851	ANTIGEN RE	Details ...	165	-0.6937604	-2.5514965	0	0	

- Over the years thousands of papers have been translated into digital representations as pathways or functional annotation.
- Enrichment analysis allows us to interpret data from these papers
- But .... Major cognitive burden relating overlapping information

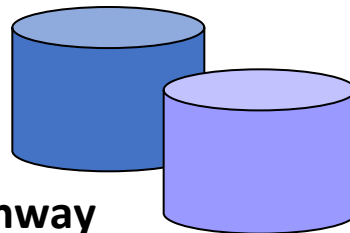
# Pathway Enrichment Test: General Framework

Experimental  
Data



Enrichment Table

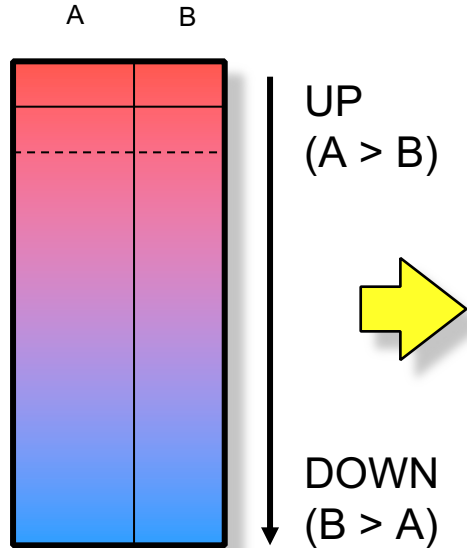
Spindle	0.00001
Apoptosis	0.00025



Pathway  
Database

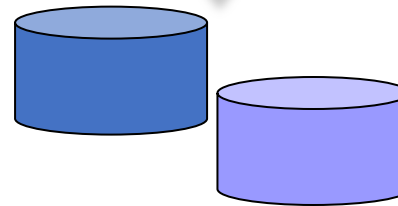
# Pathway Enrichment Analysis

## Ranked Gene List



## Thresholded lists (selected genes)

GSEA



Pathways

g:Profiler

Enrichment in  
Condition A vs. B

Gene-set	Significance
Cell Cycle	0.0001
EGF Pathway	0.003
Spindle	0.007
...	...

Enrichment in  
Condition B vs. A

Gene-set	Significance
Proteasome	0.0002
Apoptosis	0.005
Caspase	0.009
...	...

Merico D, Isserlin R, Stueker O, Emili A, Bader GD  
Enrichment map: a network-based method for  
gene-set enrichment visualization and  
interpretation PLoS One. 2010 Nov 15;5(11):e13984



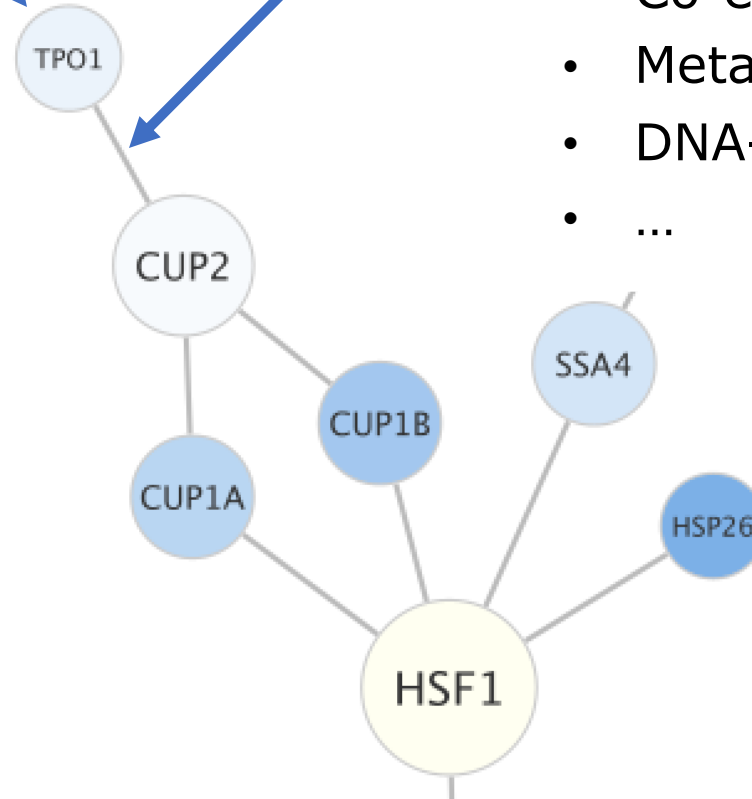
# Network Basics

## Node (molecule/entity)

- Gene
- Protein
- Transcript
- Drug
- MicroRNA
- ...

## Edge (interaction/relationship)

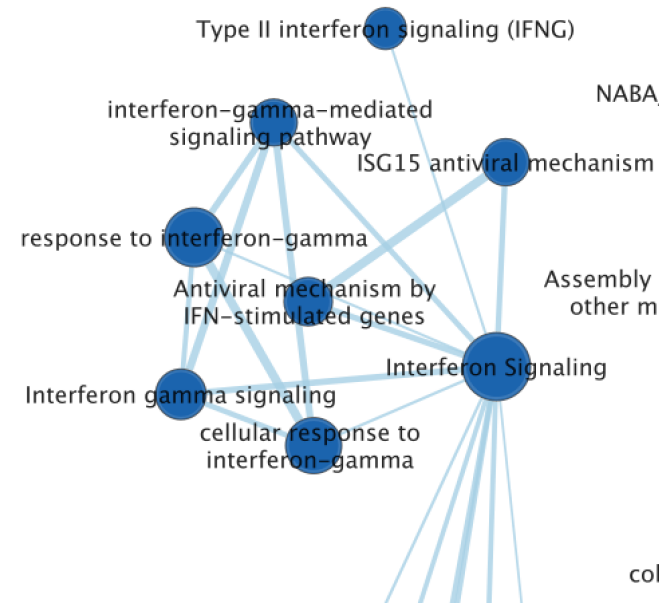
- Genetic interaction
- Physical protein interaction
- Co-expression
- Metabolic reaction
- DNA-binding
- ...



# Enrichment Map Basics

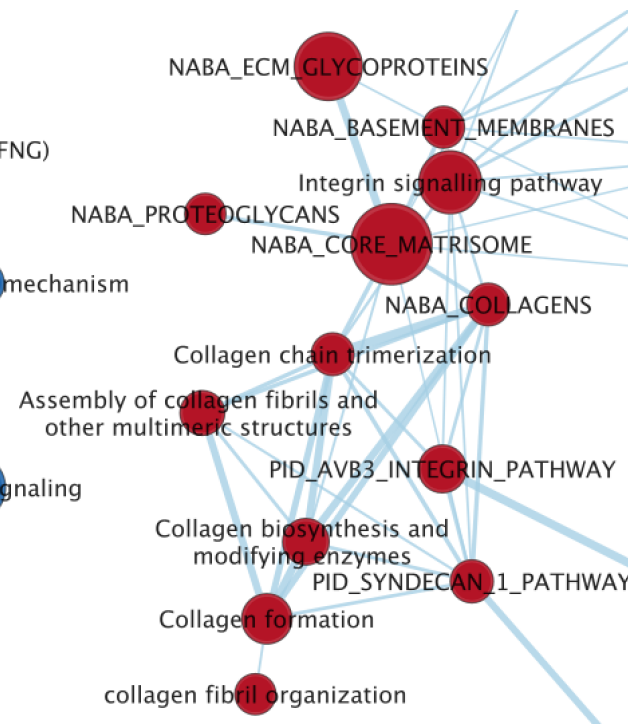
## Node (molecule/entity)

- Pathway or geneset
- Size is correlated to number of genes in set
- Color indicates class in below example (for example Up/Down, classA/classB)



## Edge (interaction/relationship)

- Degree of overlap between two genesets
- The more genes two pathways have in common the thicker/stronger the connection



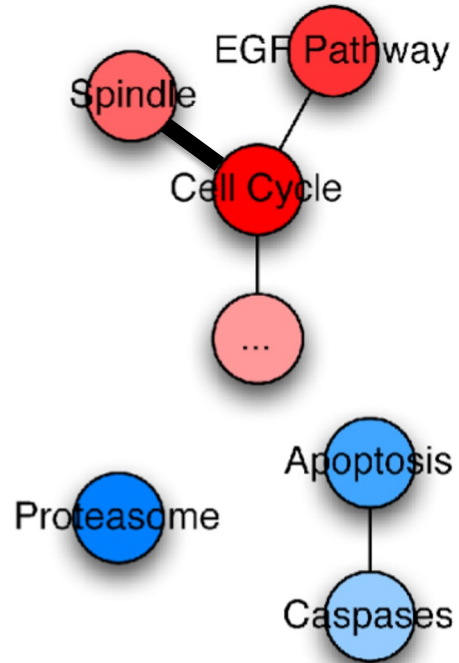
# Enrichment Map

Enrichment in  
Condition A vs. B

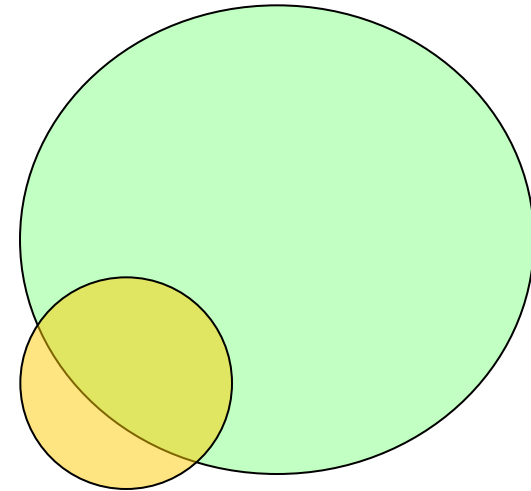
Gene-set	Significance
Cell Cycle	0.0001
EGF Pathway	0.003
Spindle	0.007
...	...


Enrichment in  
Condition B vs. A

Gene-set	Significance
Proteasome	0.0002
Apoptosis	0.005
Caspase	0.009
...	...



## Overlap



Enriched in phenotype  
A  B

GENE-SET LIST



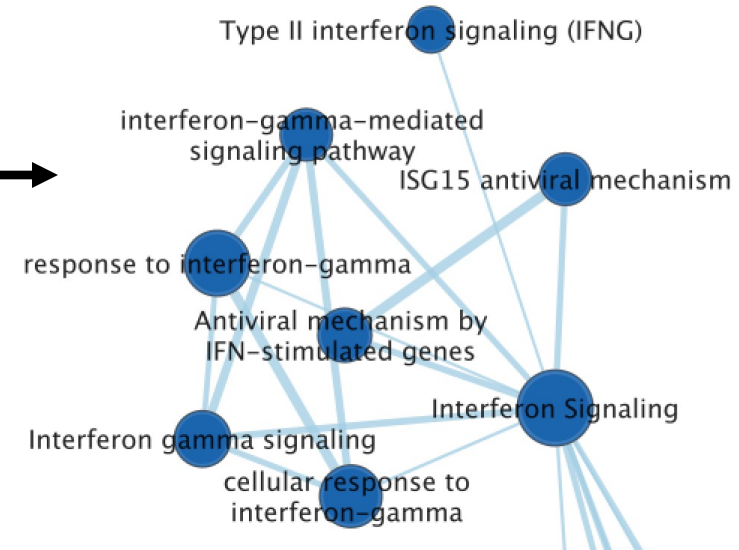
ENRICHMENT MAP

$$\frac{|A \cap B|}{\min(|A|, |B|)}$$

# Typical Output

RNA HELICASE ACTIVITY%GO%GO:0003724	28	1.77	0.0041	0.0464386
MRNA SURVEILLANCE PATHWAY%KEGG%HS03015	82	1.77	0	0.0466167
UBIQUITIN-DEPENDENT DEGRADATION OF CYCLIN D1%REACTOME%REACT_4.1	50	1.77	0.0021	0.0486015
BIOCARTA_CD40_PATHWAY%MSGDB_C2%BIOCARTA_CD40_PATHWAY	15	1.77	0.0048	0.0483781
IGF1 PATHWAY%PATHWAY INTERACTION DATABASE NCI-NATURE CURATED DATA%IGF1 PATHWAY	29	1.76	0.003	0.0489742
UBIQUITIN-DEPENDENT PROTEIN CATABOLIC PROCESS%GO%GO:0006511	204	1.76	0	0.0488442
PHAGOSOME%KEGG%HS04145	147	1.76	0	0.0486164
PROTEASOME COMPLEX%GO%GO:0000502	29	1.76	0.007	0.0490215
ANTIGEN PRESENTATION: FOLDING, ASSEMBLY AND PEPTIDE BINDING OF CLASS I MHC%REACTOME%REACT_7	24	1.76	0.0041	0.0505599
ABORTIVE ELONGATION OF HIV-1 TRANSCRIPT IN THE PRESENCE OF TAT%REACTOME%REACT_6261.3	23	1.75	0	0.0529242
DNA DAMAGE RESPONSE, SIGNAL TRANSDUCTION BY P53%KEGG%HS05030	67	1.75	0	0.052886
REGULATION OF MACROPHAGE ACTIVATION%GO%GO:000930	11	1.75	0.003	0.0534709
PROTEIN FOLDING%REACTOME%REACT_1695.2	52	1.75	0.002	0.0537717
ENDOPLASMIC RETICULUM UNFOLDED PROTEIN RESPONSE%GO%GO:0000068	73	1.75	0	0.0546052
PROTEIN EXPORT%KEGG%HS03060	24	1.75	9.75E-04	0.0548699
TRANSCRIPTION INITIATION FROM RNA POLYMERASE II PROMOTER%GO%GO:0006367	64	1.75	0.001	0.0545783
S PHASE%REACTOME%REACT_899.4	110	1.75	0	0.0546003
PROTEASOMAL PROTEIN CATABOLIC PROCESS%GO%GO:0006511	163	1.75	0	0.0550066
ATP-DEPENDENT RNA HELICASE ACTIVITY%GO%GO:0003724	20	1.74	0.0059	0.0556722
ACID-AMINO ACID LIGASE ACTIVITY%KEGG%HS03015	217	1.74	0	0.0560217
GO%GO:0072474	67	1.74	0.002	0.0565978
GO%GO:0035966	107	1.74	0	0.0562957
GO%GO:0072413	67	1.74	9.81E-04	0.05761
BIOCARTA_IL4_PATHWAY%MSGDB_C2%BIOCARTA_IL4_PATHWAY	11	1.74	0.0082	0.0581508
ASSOCIATION OF TRIC CCT WITH TARGET PROTEINS DURING BIOSYNTHESIS%REACTOME%REACT_16907.2	28	1.74	0.0039	0.0581298
UBIQUITIN-DEPENDENT DEGRADATION OF CYCLIN D%REACTOME%REACT_938.4	50	1.74	0.0029	0.057876
MODIFICATION-DEPENDENT PROTEIN CATABOLIC PROCESS%GO%GO:0019941	207	1.74	0	0.0576579
TRANSLATION INITIATION COMPLEX FORMATION%REACTOME%REACT_1979.1	55	1.74	0.0021	0.0575181
GO%GO:0001906	13	1.74	0.0117	0.0572877
G1 S TRANSITION%REACTOME%REACT_1783.2	107	1.74	0	0.0572618
GO%GO:0034620	73	1.73	0.0021	0.0576606
SIGNALING BY NOTCH%REACTOME%REACT_299.2	19	1.73	0.0069	0.0578565
RESPONSE TO UNFOLDED PROTEIN%GO%GO:0006986	102	1.73	0	0.0583864
SIGNAL TRANSDUCTION INVOLVED IN G1 S TRANSITION CHECKPOINT%GO%GO:0072404	68	1.73	0.002	0.0582213
GO%GO:0072431	67	1.73	0	0.058551
BIOCARTA_PROTEASOME_PATHWAY%MSGDB_C2%BIOCARTA_PROTEASOME_PATHWAY	19	1.73	0.0099	0.0586655
HOST INTERACTIONS OF HIV FACTORS%REACTOME%REACT_6288.4	117	1.73	0	0.0586888
AUTOPHAGIC VACUOLE ASSEMBLY%GO%GO:0000045	13	1.73	0.0122	0.0588271
CYCLIN A-CDK2-ASSOCIATED EVENTS AT S PHASE ENTRY%REACTOME%REACT_9029.2	66	1.73	0	0.0610099

## Network Visualization



Each row is a gene-set (pathway).

It displays:

- a score associated with the magnitude of overlap between gene-set and gene list.
- a pvalue that estimates the significance of the enrichment (by chance or not).
- a corrected pvalue (FDR) that corrects for multiple hypothesis testing.

# Enrichment Map: use case I

## *Single enrichment*

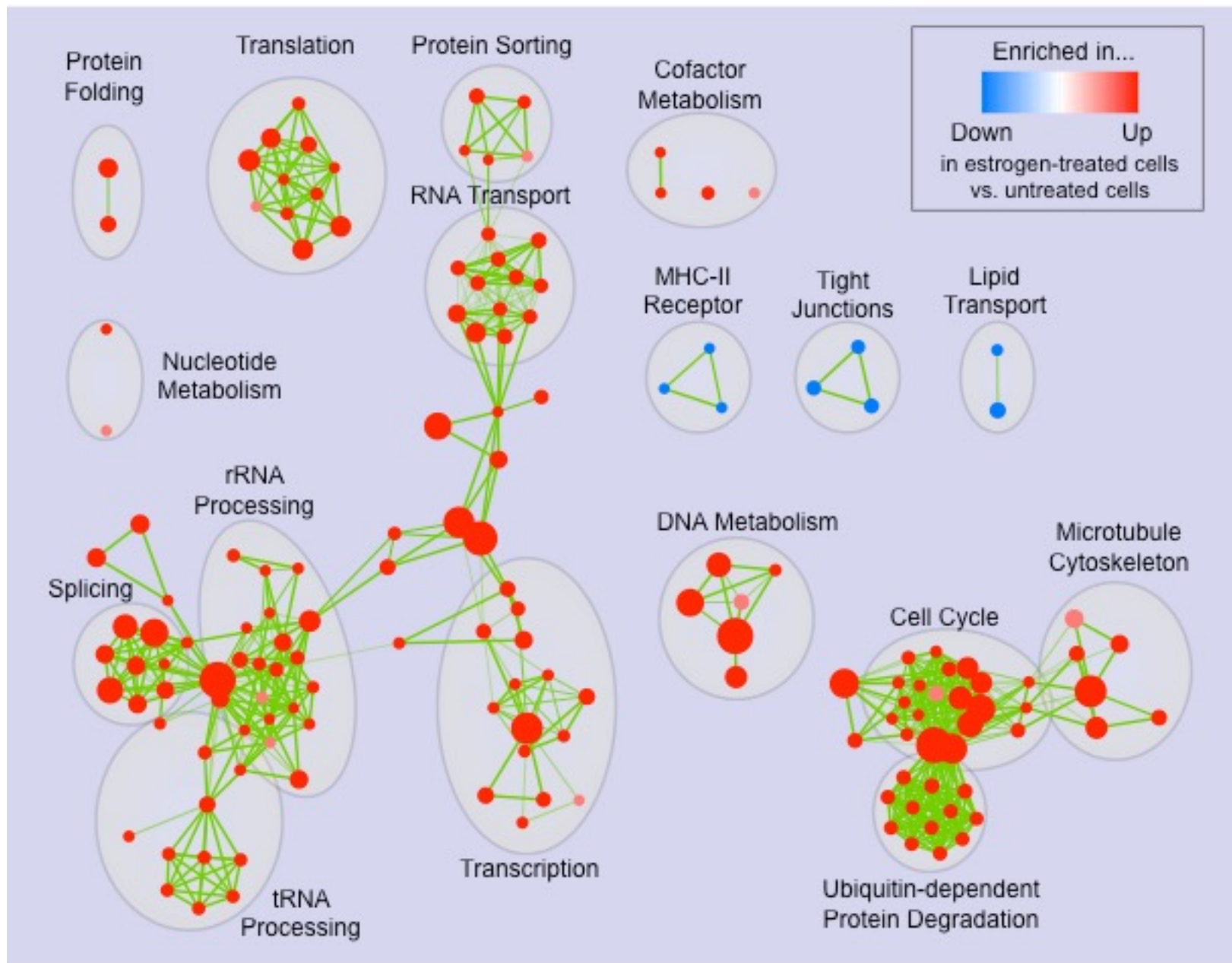
Estrogen treatment of breast cancer cells

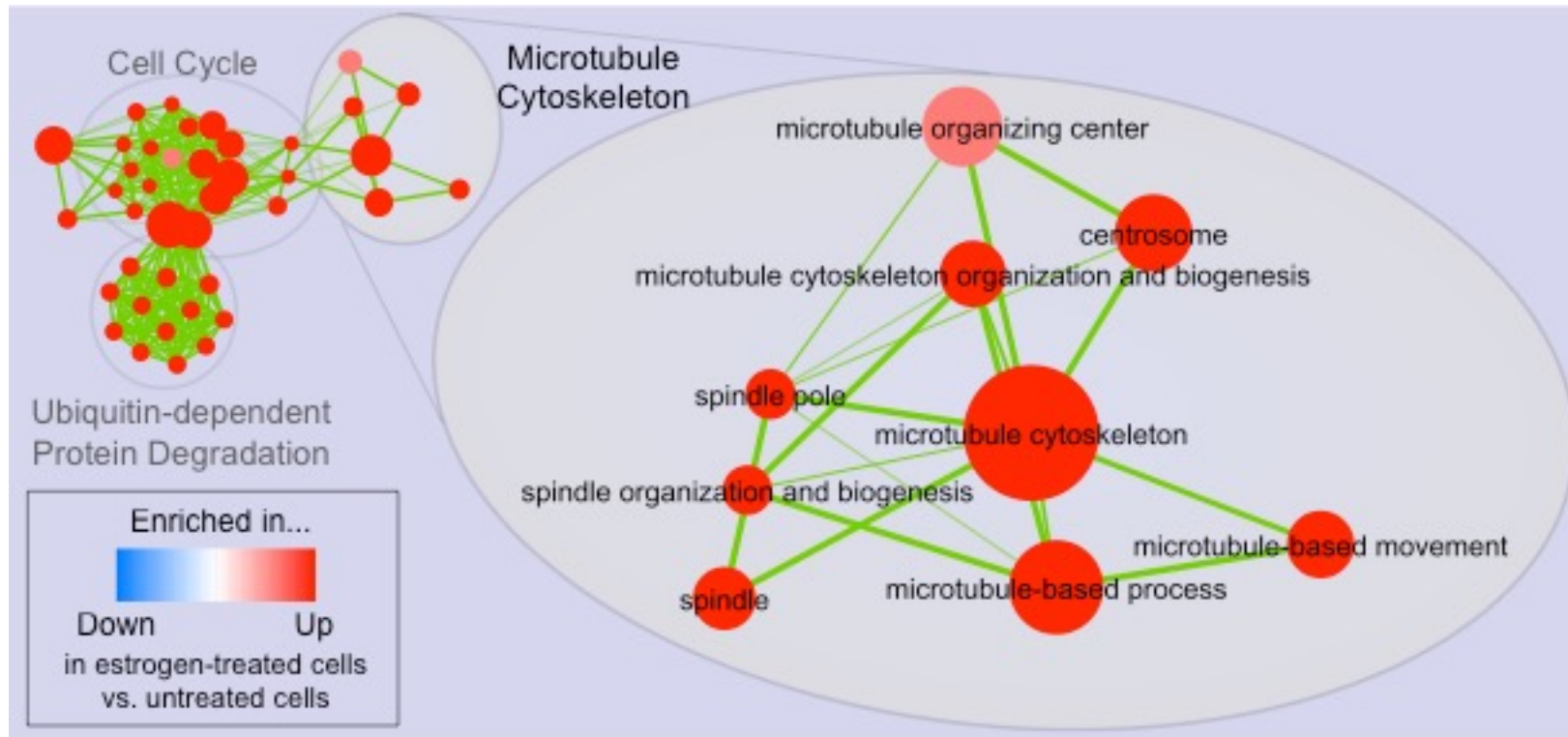
- Design:  
**2-time points, two-class**

	12 hrs	24 hrs
Estrogen-treated	3	3
Untreated	3	3

- Gene set Database:  
**Gene Ontology**

Lin C-Y, Vega VB, Thomsen JS, Zhang T, Kong SL, et al. (2007) Whole-genome cartography of estrogen receptor alpha binding sites. PLoS Genetics 3:e87





# Enrichment Map: use case II

## *Comparison of two enrichments*

Estrogen treatment of breast cancer cells

- Design:

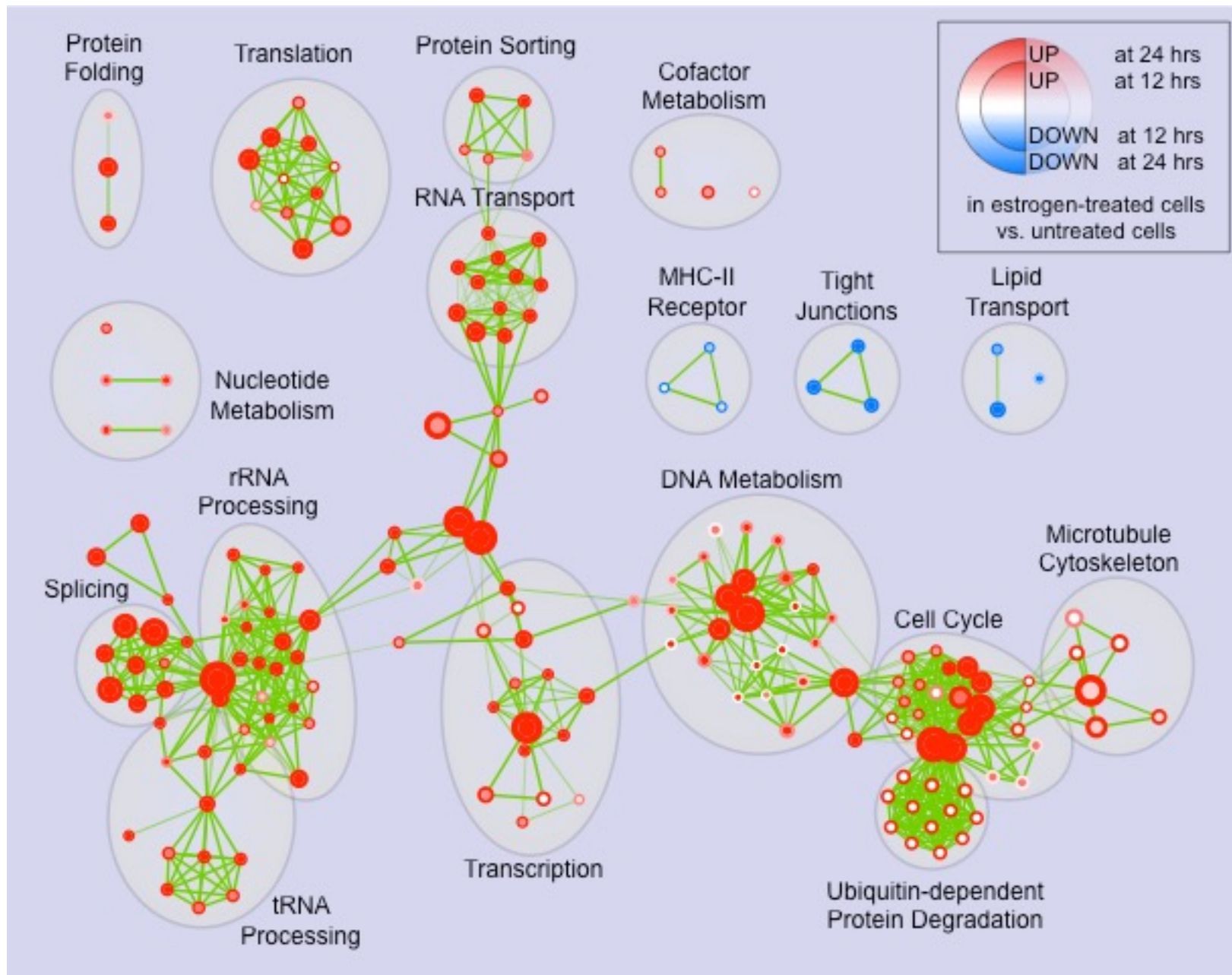
**2-time points, two-class**

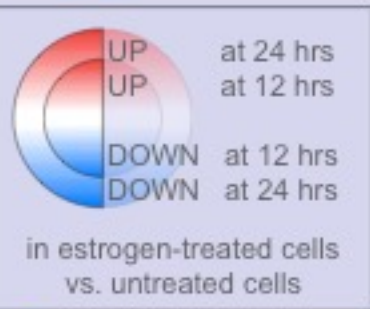
	12 hrs	24 hrs
Estrogen-treated	3	3
Untreated	3	3

- Gene set Database:

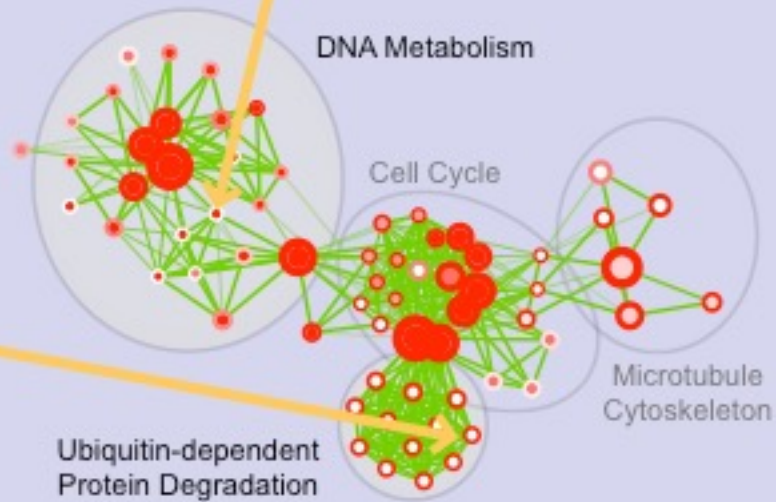
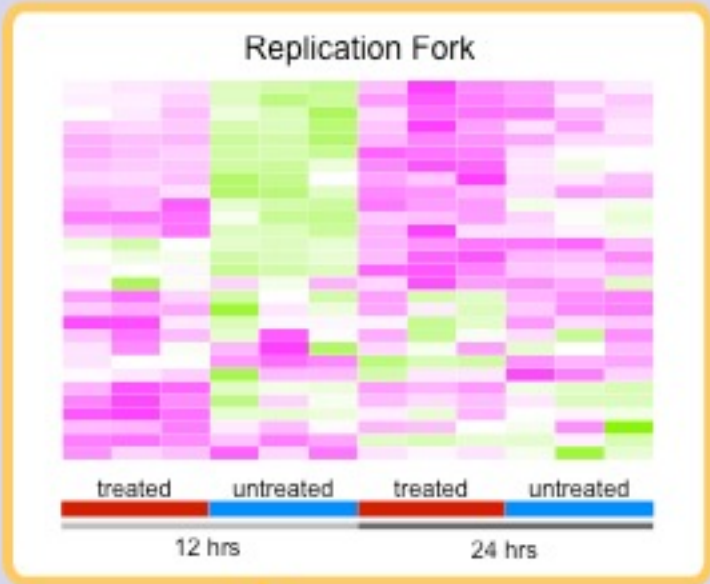
**Gene Ontology**





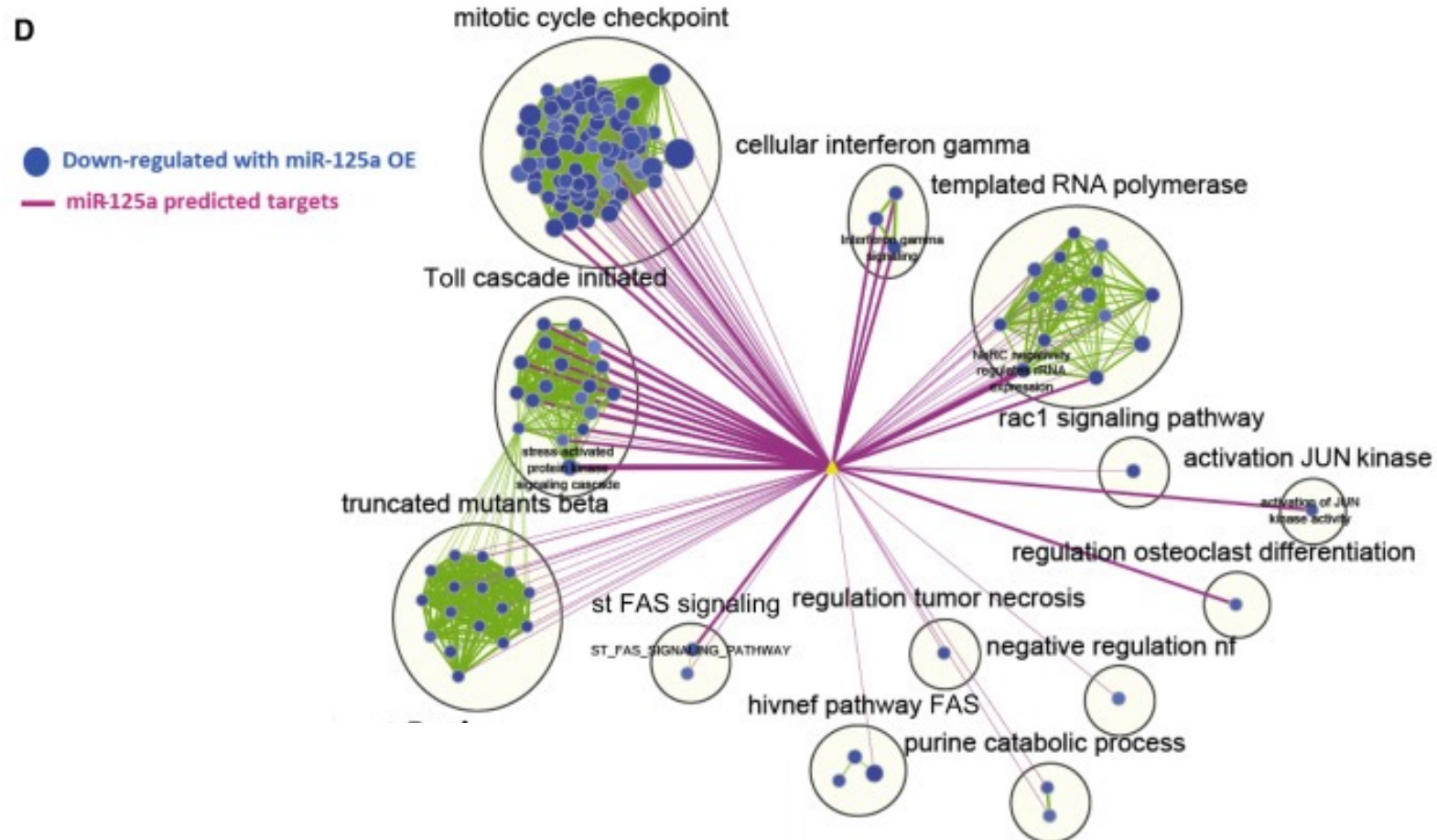


Estrogen-treated vs untreated:  
 ■ Up  
 ■ Down



# Enrichment Map: use case III

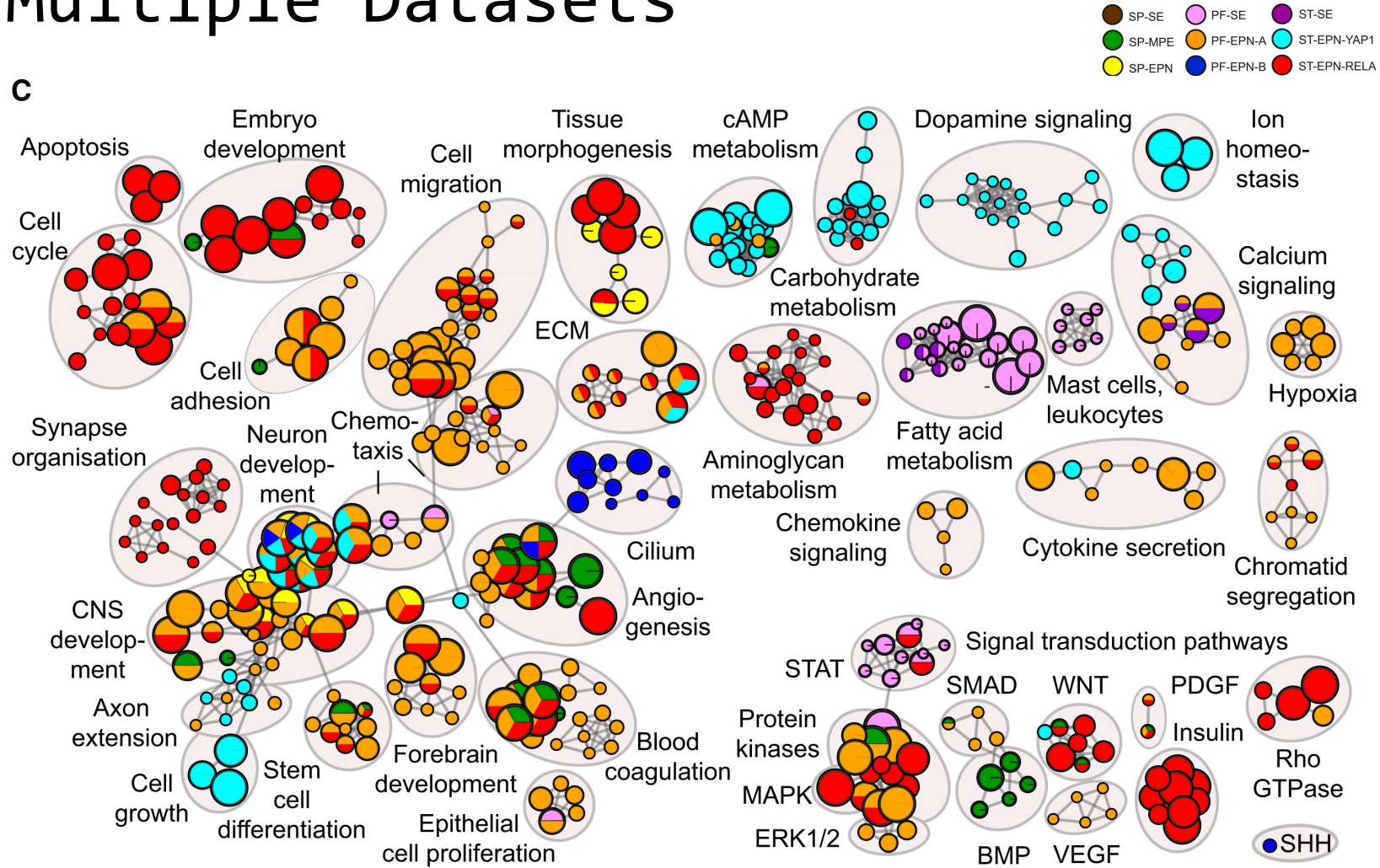
## Query Set Analysis



Wojtowicz EE, Lechman ER, Hermans KG, et al. Ectopic miR-125a Expression Induces Long-Term Repopulating Stem Cell Capacity in Mouse and Human Hematopoietic Progenitors. *Cell Stem Cell*. 2016;19(3):383-396

# Enrichment Map: use case IV

## Multiple Datasets



Pajtler KW, Witt H, Sill M, et al. Molecular Classification of Ependymal Tumors across All CNS Compartments, Histopathological Grades, and Age Groups. *Cancer Cell*. 2015;27(5):728-743

# Enrichment Map Features

The screenshot displays the EnrichmentMap software interface. The main window shows a network visualization of gene-gene interactions, with nodes colored in red and blue. A search bar at the top right contains the text "BGN".

The left sidebar contains several toolbars: "jActiveModules", "EnrichmentMap", "Annotation", "Layout Tools", "Filter", "Style", and "Network".

The "EnrichmentMap" panel on the left includes the following settings:

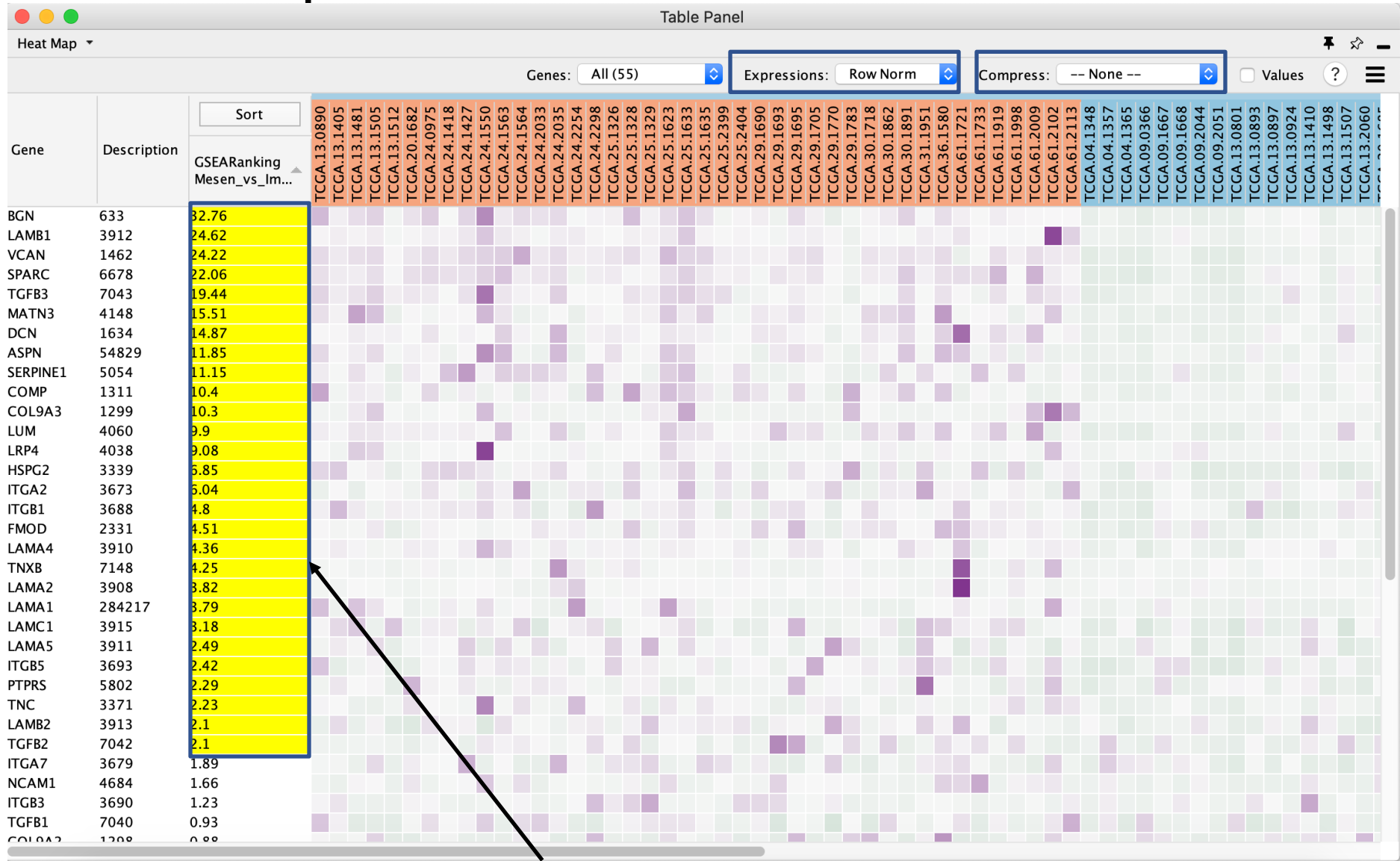
- Filter:**
  - Node Cutoff:  P-value  Q-value (0.005)
  - Edge Cutoff (Similarity): 0.375
- Data Sets:**
  - Mesen\_vs\_Immuno\_edgeR.GseaPreranked (730)
- Style:**
  - Chart Data: NES Columns
  - Chart Type: Radial Heat Map
  - Chart Colors: RdBu-9
  - Show Chart Labels
  - Publication-Ready

The "Heat Map" panel at the bottom right displays a table of gene expression data:

Gene	Description	GSEARanking	Mesen_vs_Im...	Expressions
BGN	633	32.76		
COL16A1	1307	28.88		
GPC6	10082	28.31		
AEBP1	165	27.52		
TIMP3	7078	25.97		
EBF1	2200	25.53		
PCOLCE	5118	25.17		
LOXL2	4017	24.76		
LAMB1	3912	24.62		
CILP2	148113	24.41		
VCAN	1462	24.22		
FN1	2225	22.08		

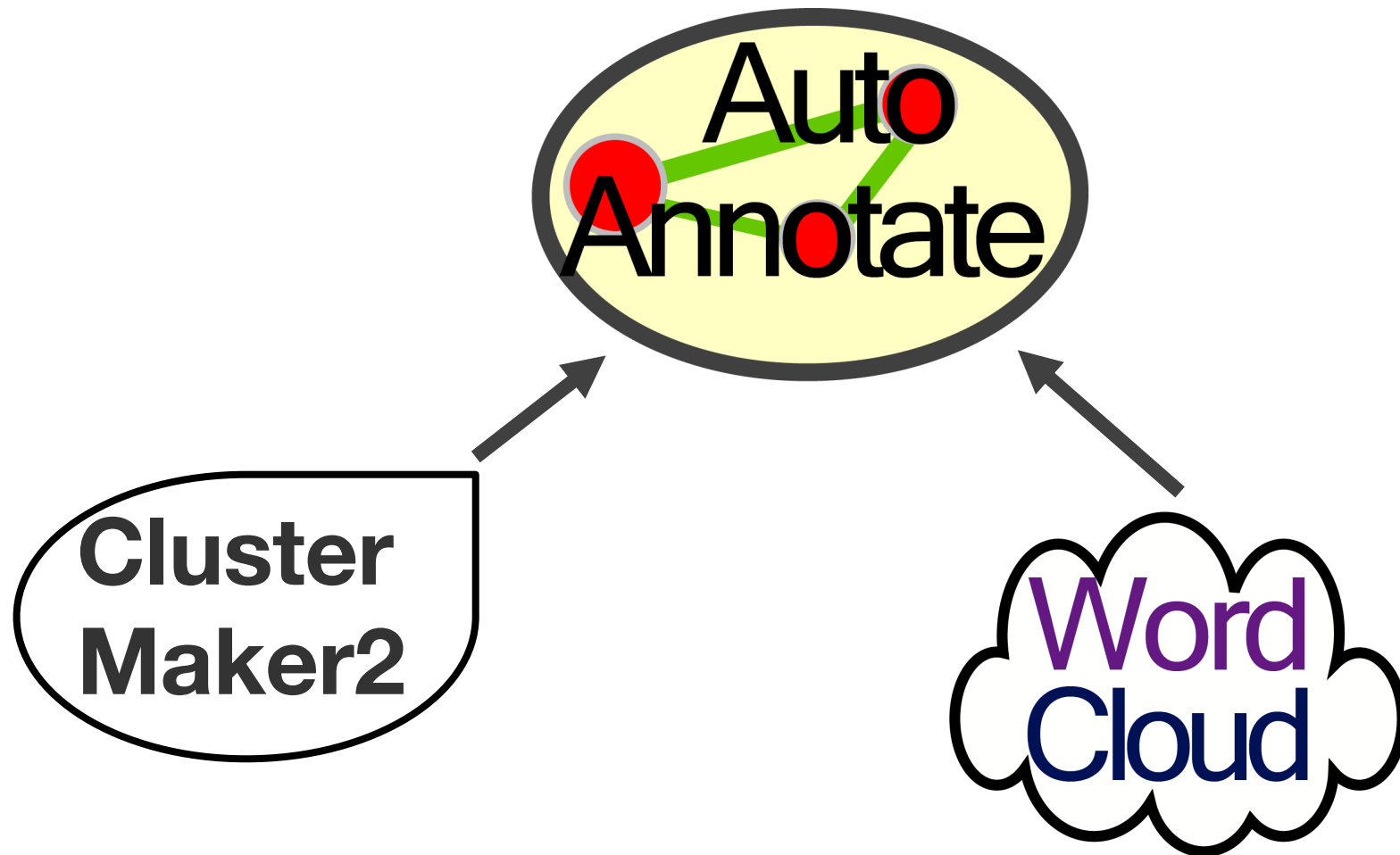
At the bottom of the interface, there are tabs for "Node Table", "Edge Table", "Network Table", and "Heat Map".

# Heatmap Panel



Leading Edge

# Enrichment Map Features

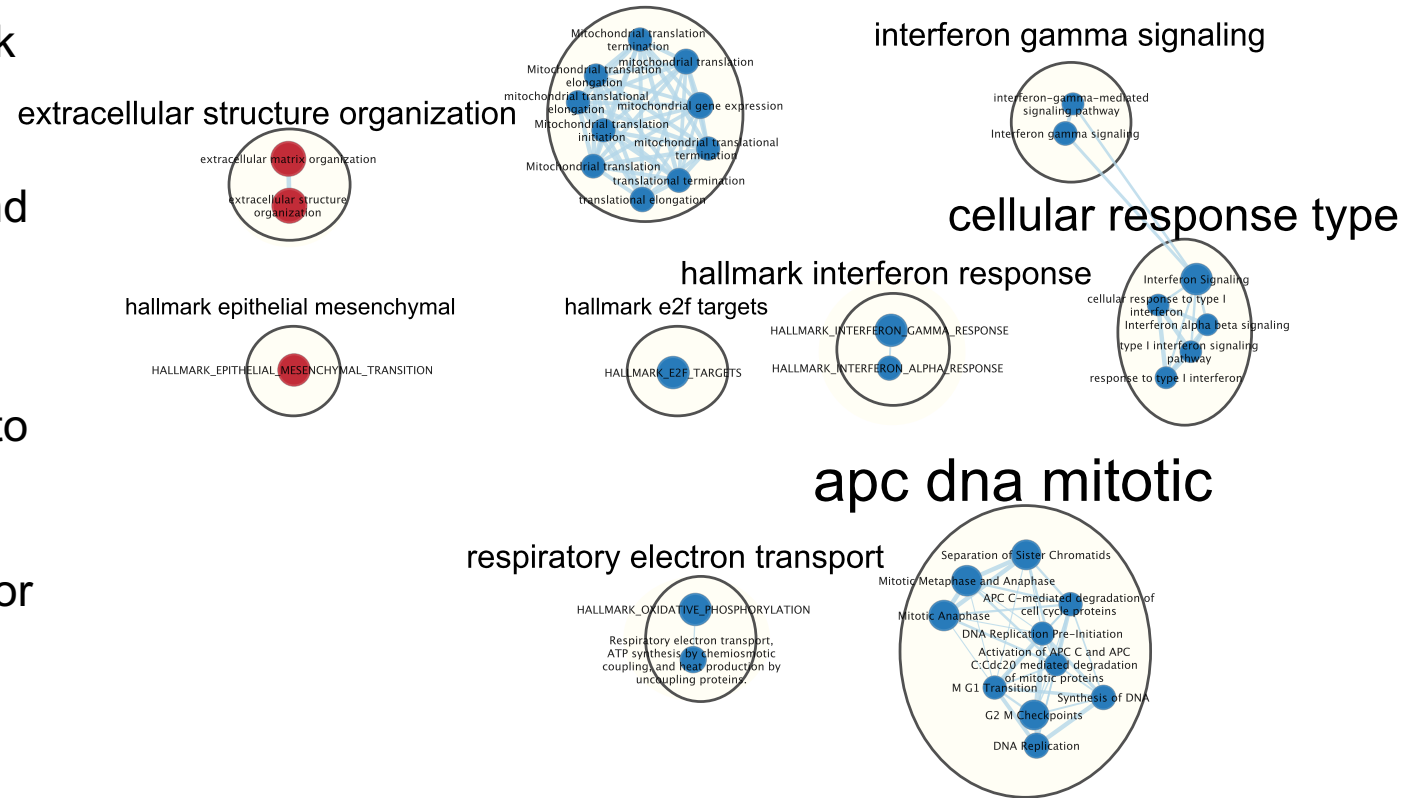


# AutoAnnotate

## Steps:

1. Cluster the Network
2. For each cluster, find frequent words in node labels
3. Select top 3 words to create labels
4. Possibility to move or edit labels to make figure publication ready.

## mitochondrial translational translation





# AutoAnnotate

Session: New Session

Control Panel

EnrichmentMap AutoAnnotate

MCL Cluster Annotation Set 2

Cluster	Nodes	Collapsed
1 hiv elongation	33	
transport mature derived	20	
synthesis dna gap	17	
interphase g1 transition	16	
mitotic metaphase anaphase	16	
splicing reactions bulged	15	
spindle checkpoint components	13	
trna related aminoacylation	13	
iii rna polymerase	12	
processing pre 3'	12	
translation initiation influenza	11	
export nucleus nitrogen	10	
dependent atpase helicase	9	
5' exoribonuclease exonuclease	7	
coupled nucleotide excision	7	
remodeling centromere nucleosome	7	
tubulin folding cct	7	
resolution sites base	6	
ribonucleoside monophosphate process	6	
cis trans peptidyl	5	
cycle phase cell	5	
establishment targeting mitochondrion	5	
formation complex 43s	5	
rrna processing ncna	5	
chromosomal part chromosome	4	
destabilization mrna tristetrprolin	4	
envelope mitochondrial membrane	4	
gamma mediated interferon	4	
golgi transport vesicle	4	
muscle development tissue	4	
nuclear replisome replication	4	
pyrimidine nucleoside process	4	
replicative atr g2	4	
rna polymerase escape	4	
strand break recombinational	4	
transferase macromolecule methylation	4	
type interferon signaling	4	
actin filament bundle	3	

91 clusters, 0 selected

export nucleus nitrogen

processing pre 3'

transport mature derived

iv elongation

uv cellular response

splicing reactions bulged

repair hom

Results Panel

AutoAnnotate Display

Shape:

Border Width: 3

Opacity: 20%

Fill:

Border:

Hide Shapes

Font Scale: 50%

Scale font by cluster size

Font Color:

Word Wrap

Wrap Length: 20

Hide Labels

Table Panel

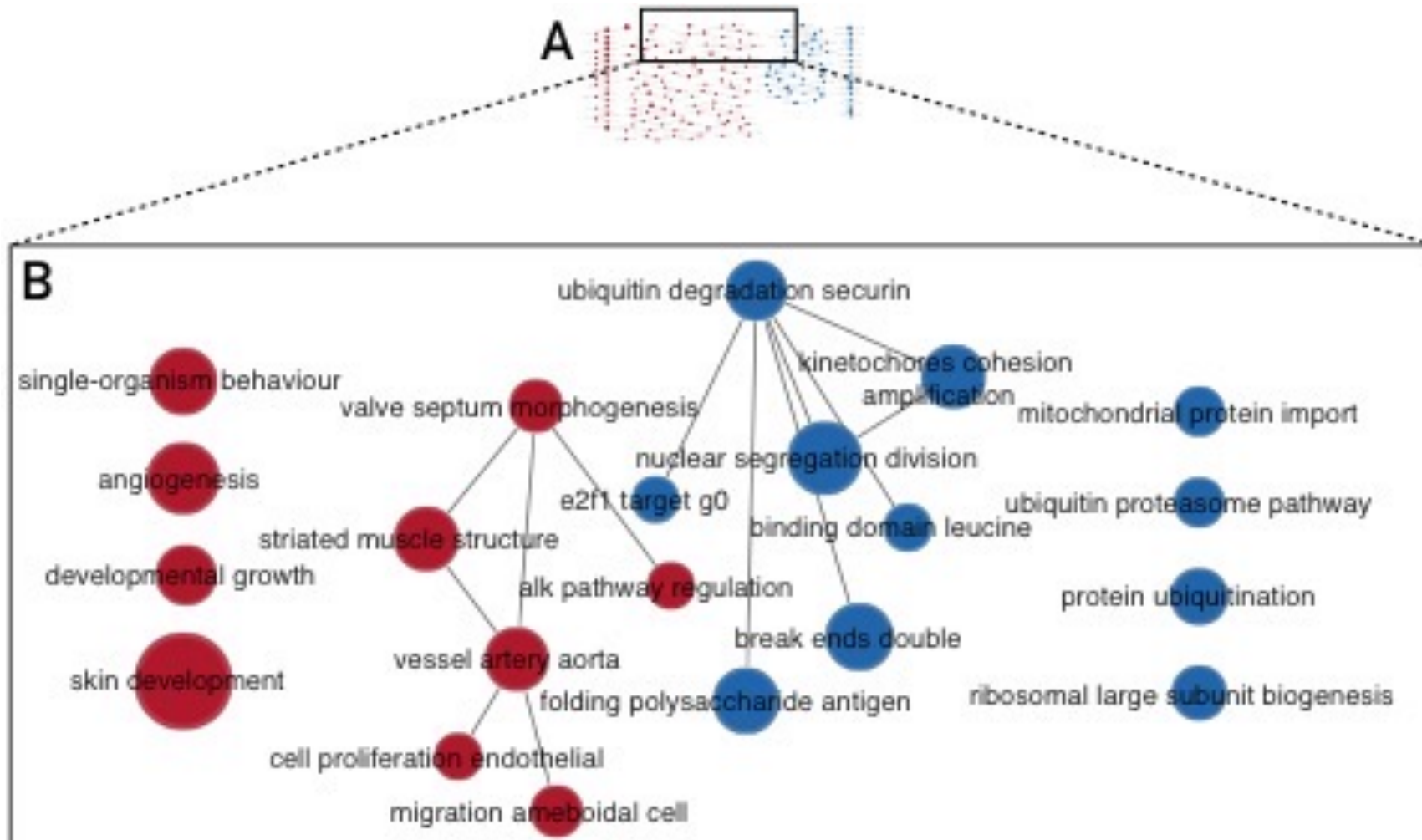
Genes: All (420) Expressions: Values Compress: -- None --  Values

Gene	Description	Sort	E2_12h_01	E2_12h_02	E2_12h_03	NT_12h_01	NT_12h_02	NT_12h_03	E2_24h_01	E2_24h_02	E2_24h_03	NT_24h_01	NT_24h_02	NT_24h_03	E2_48h_01	E2_48h_02	E2_48h_03	NT_48h_01	NT_48h_02	NT_48h_03	
AARS	AARS (alan...																				
ADARB1	"ADARB1 (a...																				
ADAT1	"ADAT1 (ad...																				
AGGF1	AGGF1 (an...																				

Node Table Edge Table Network Table **Heat Map**

Memory

# Collapsed network



# Practical Lab

- Experiment with Cytoscape
- Try out Enrichment Map
- Load your g:Profiler results – and - load the GSEA enrichment analysis files that you generated in Module 2 lab
- Try out all the different Enrichment Map features
- Try AutoAnnotate to annotate your networks.

# We are on a Coffee Break & Networking Session

Workshop Sponsors:

---



Canadian Centre for  
Computational  
Genomics



HPC4Health



OICR  
Ontario Institute  
for Cancer Research



GenomeCanada