



Canadian Bioinformatics Workshops

www.bioinformatics.ca

bioinformaticsdotca.github.io

Supported by



Creative Commons

This page is available in the following languages:

Afrikaans Български Català Dansk Deutsch Ελληνικά English English (CA) English (GB) English (US) Esperanto
 Castellano Castellano (AR) Español (CL) Español (Ecusador) Castellano (MX) Castellano (PE)
 Euskara Suomeksi français français (CA) Galego മലബാറി ഹിന്ദി മന്ത്രി മാര്യു ഇറാഖി ജപ്പാൻ ഹംഗരി മാക്സി മലൈ നേഡൻസ് നോർഖ് സെസോ ലേബോ പോളി പോർട്ടുഗീസ് റോമാൻ സ്ലോവെനി ജെങ്കി ചൈനീസ് സ്രീ ലാൻ കാ സോ ഥോ സ്വീഡൻ
 中文 草語 (臺灣) 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](#)

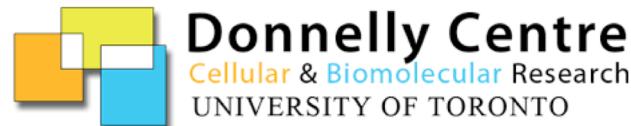
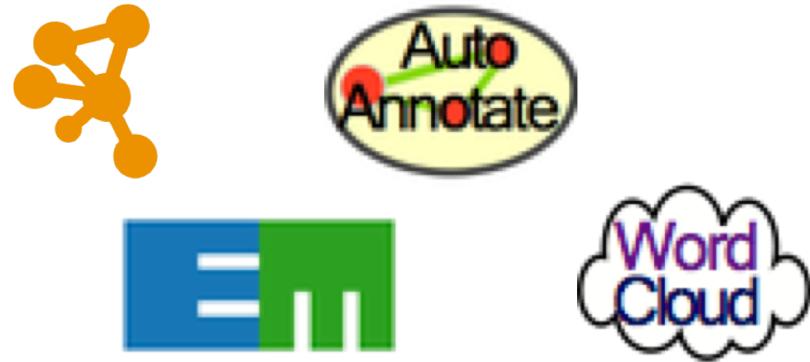
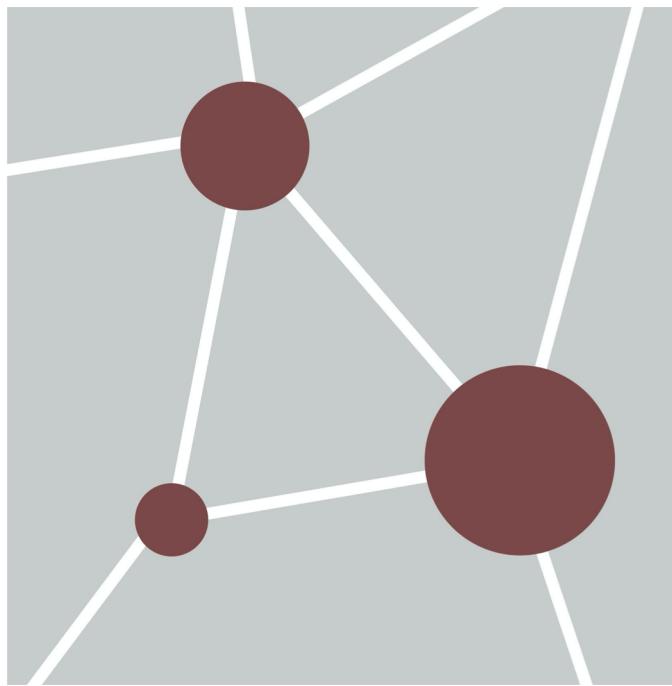
Your fair dealing and other rights are in no way affected by the above.
 This is a human-readable summary of the Legal Code (the full licence) available in the following languages:
[English](#) [French](#)

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

Module 3 Enrichment Maps



Ruth Isserlin
Pathway and Network Analysis of -omics Data
July 27-29, 2020



Learning Objectives

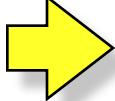
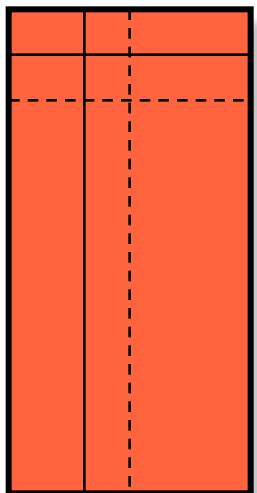
- By the end of this lecture, you will:
 - Understand how to transform your enrichment results from g:profiler, gsea or other enrichment algorithms to a Network
 - Understand the differences between a network and an Enrichment Map
 - Be able to summarize your enrichment results with annotation using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

Results of Lab 2 - Enrichment Results

- 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

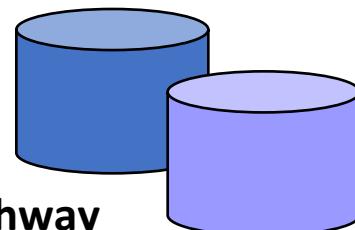
Experimental
Data



ENRICHMENT
TEST

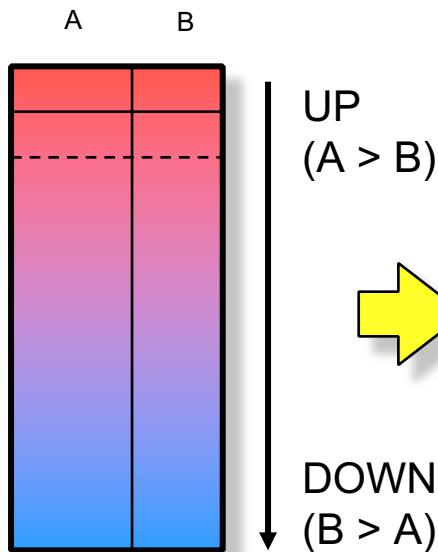
Enrichment Table

Spindle	0.00001
Apoptosis	0.00025

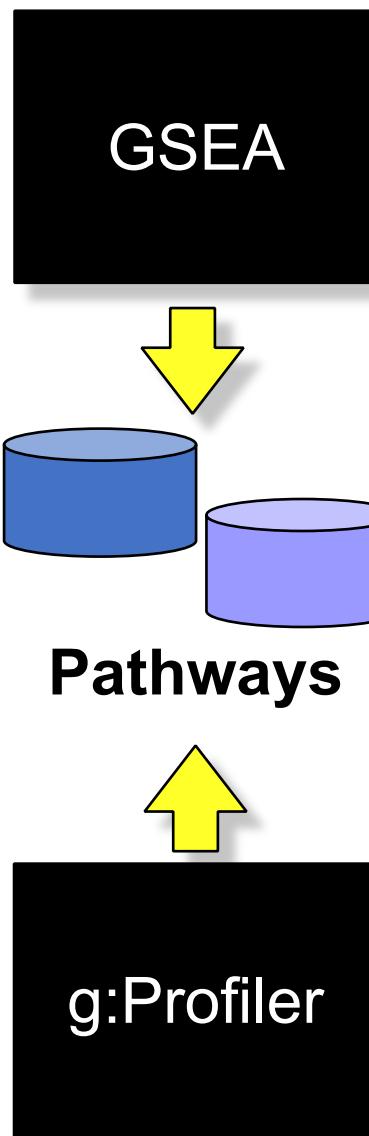


Pathway Enrichment Analysis

Ranked Gene List



Thresholded lists (selected genes)



Enrichment in Condition A vs. B	
Gene-set	Significance
Cell Cycle	0.0001
EGF Pathway	0.003
Spindle	0.007
...	...

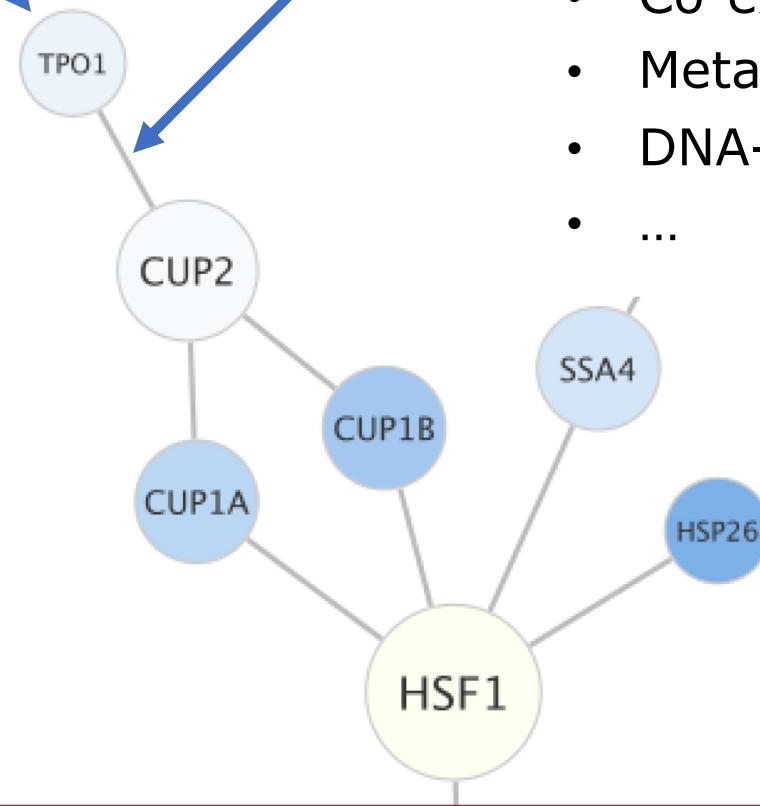
Enrichment in Condition B vs. A	
Gene-set	Significance
Proteasome	0.0002
Apoptosis	0.005
Caspase	0.009
...	...

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
- ...



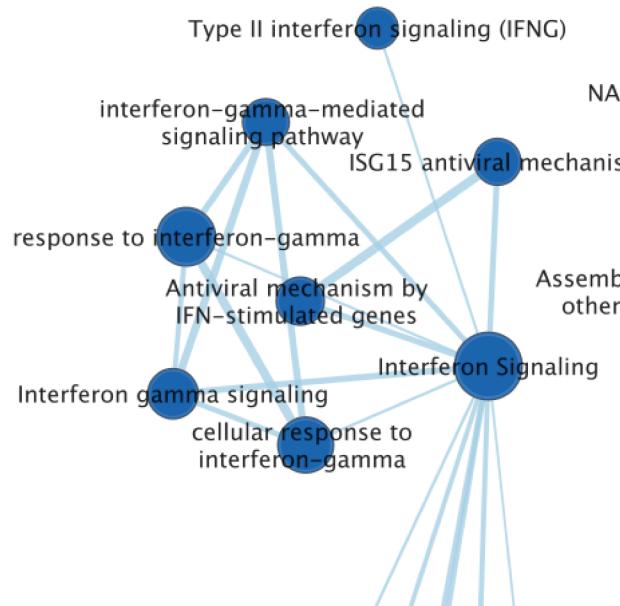
Edge (interaction/relationship)

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

Enrichment Map Basics

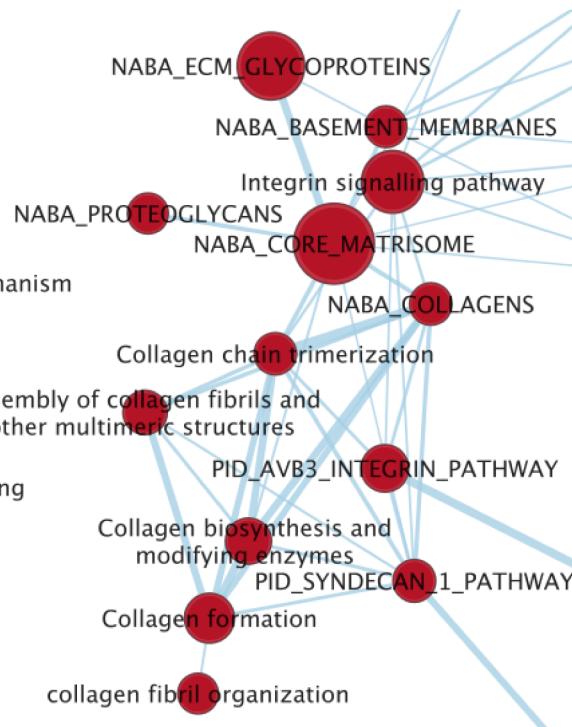
Node (molecule/entity)

- 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

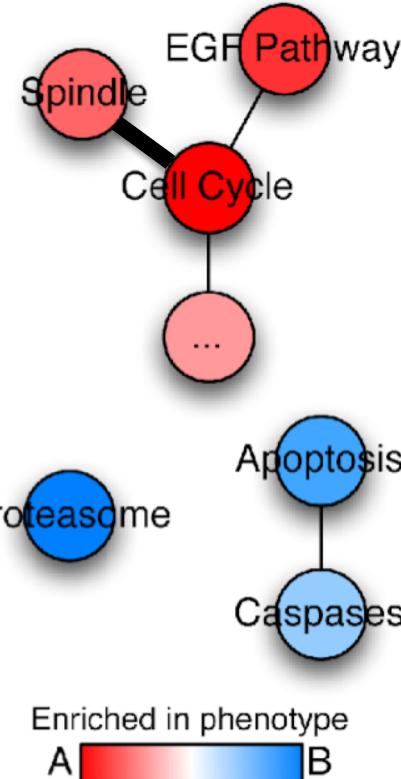
Enrichment in
Condition A vs. B

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

Enrichment in
Condition B vs. A

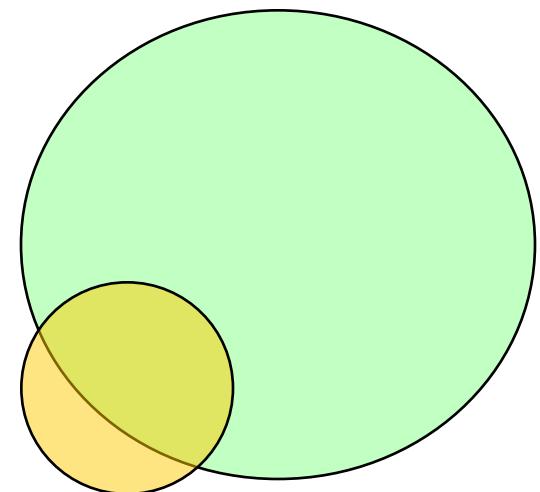
Gene-set	Significance
Proteasome	0.0002
Apoptosis	0.005
Caspase	0.009
...	...

GENE-SET LIST



ENRICHMENT MAP

Overlap



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

Typical Output

RNA HELICASE ACTIVITY%GO:GO-0003724	28	1.77	0.0041	0.0464386
MRNA SURVEILLANCE PATHWAY%KEGG%HSA03015	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%MSIGDB_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%HSA04145	147	1.76	0	0.0486164
PROTEASOME COMPLEX%GO:GO-0000502	29	1.76	0.007	0.0490215
ANTIGEN PRESENTATION: FOLDING, ASSEMBLY AND PEPTIDE LOADING OF CLASS I MHCI%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 RAD50/MERK1/ATR RESULTING IN CELL CYCLE ARREST%	67	1.75	0	0.052886
REGULATION OF MACROPHAGE ACTIVATION%GO:GO-0003930	11	1.75	0.003	0.0534709
PROTEIN FOLDING%REACTOME%REACT_16952	52	1.75	0.002	0.0537717
ENDOPLASMIC RETICULUM UNFOLDED PROTEIN RESPONSE%GO:GO-0009568	73	1.75	0	0.0546052
PROTEIN EXPORT%KEGG%HSA03060	24	1.75	9.75E-04	0.0548699
TRANSCRIPTION INITIATION FROM RNA POLYMERASE II PROMOTERS%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-00040	20	1.74	0.0059	0.0556722
ACID-AMINO ACID LIGASE ACTIVITY%GO:GO-00168	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%MSIGDB_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 D1%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%MSIGDB_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

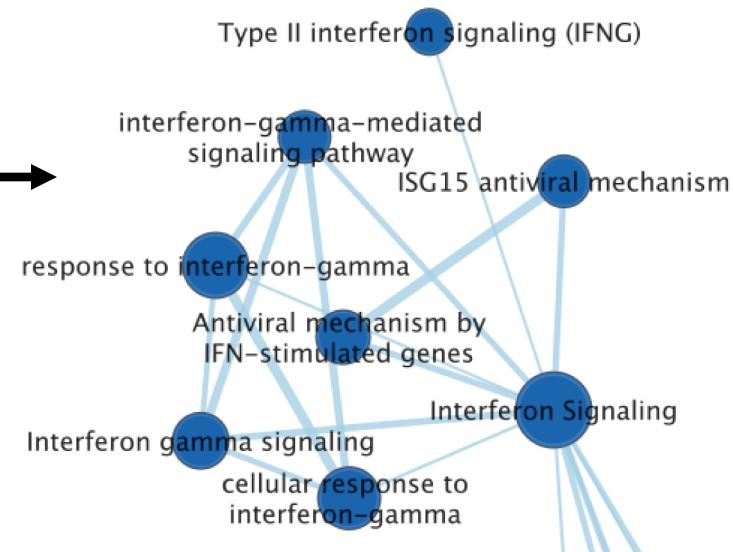
DATA
FORMAT

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.

Network Visualization



Enrichment Map: use case I

Single enrichment

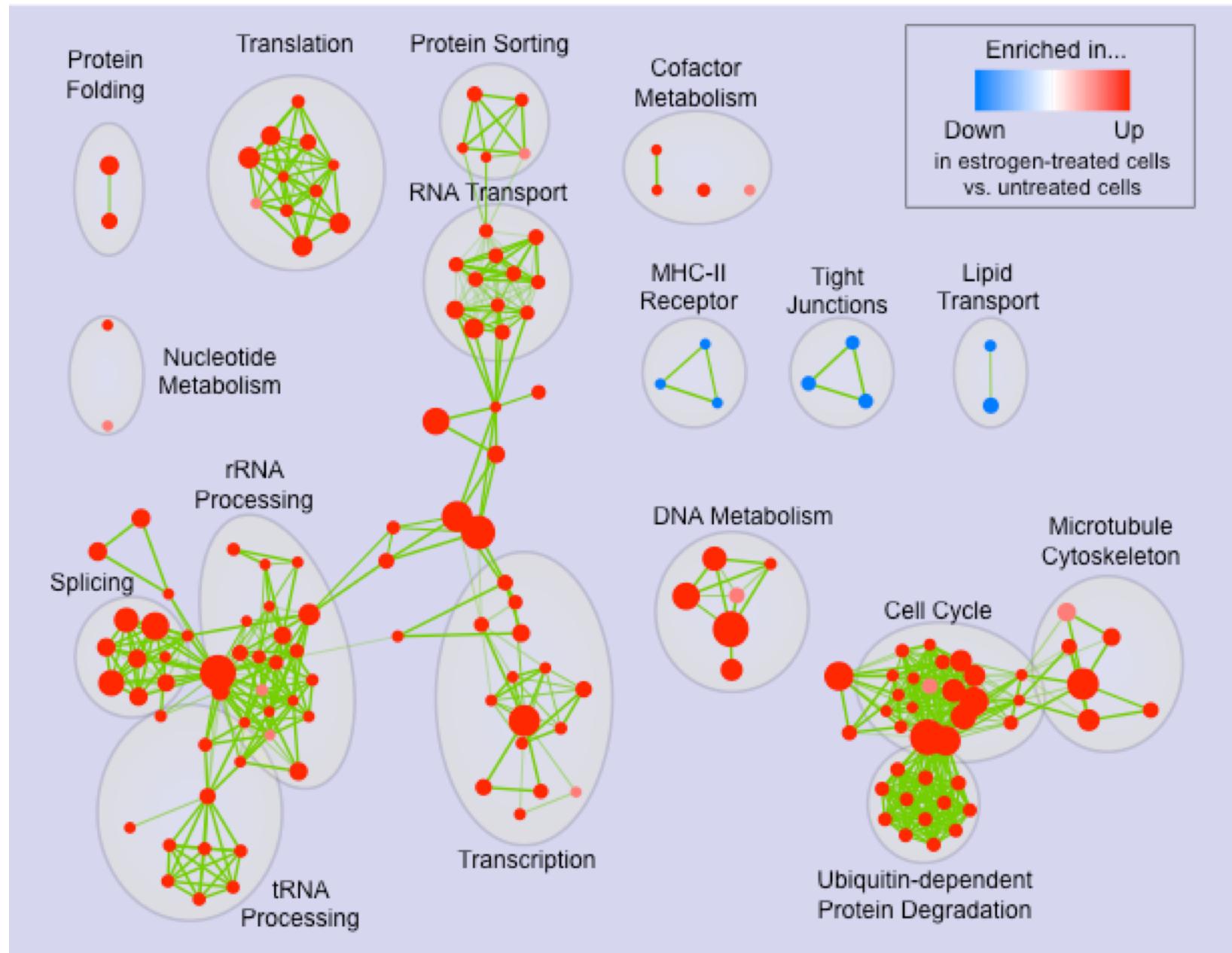
Estrogen treatment of breast cancer cells

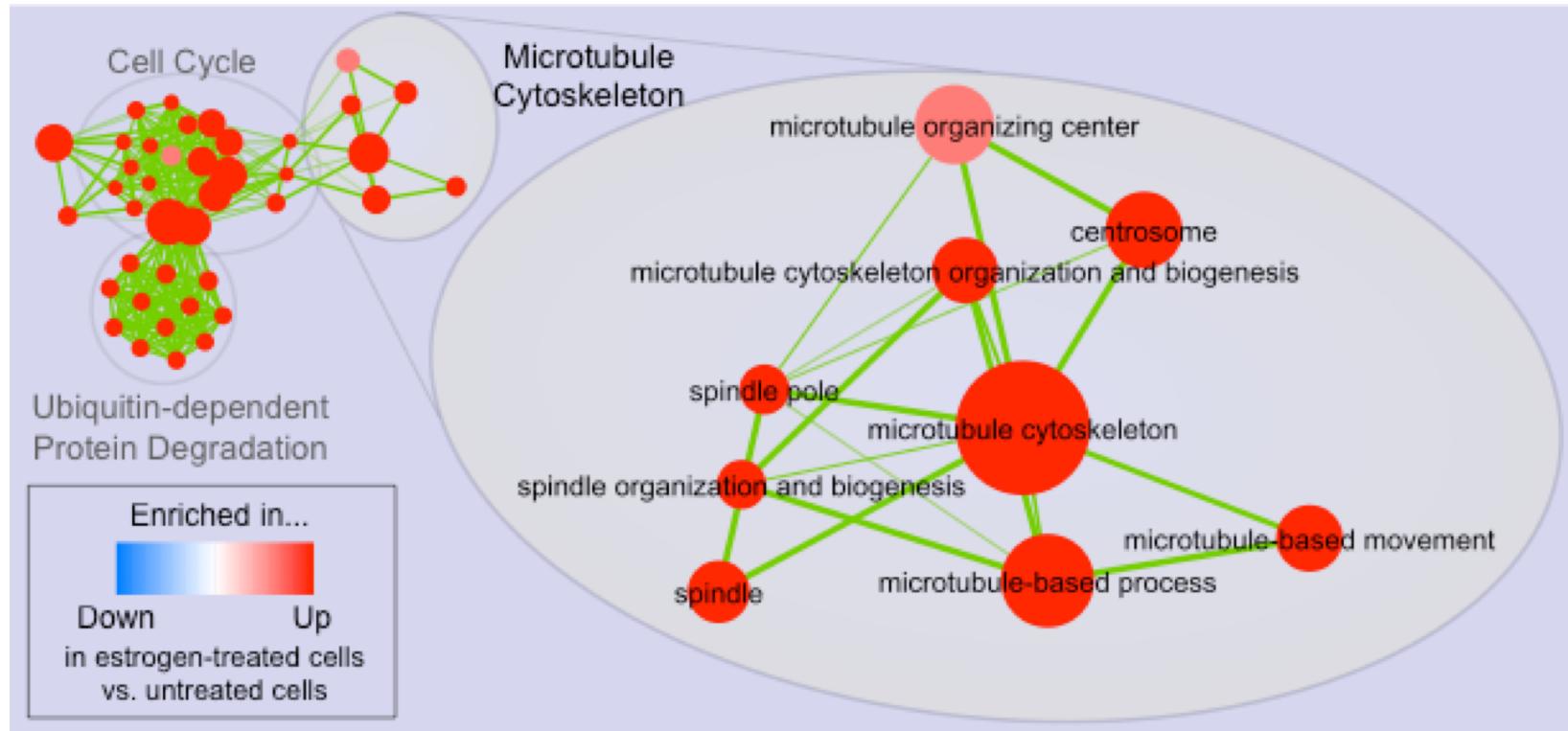
- Design:
2-time points, two-class

	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

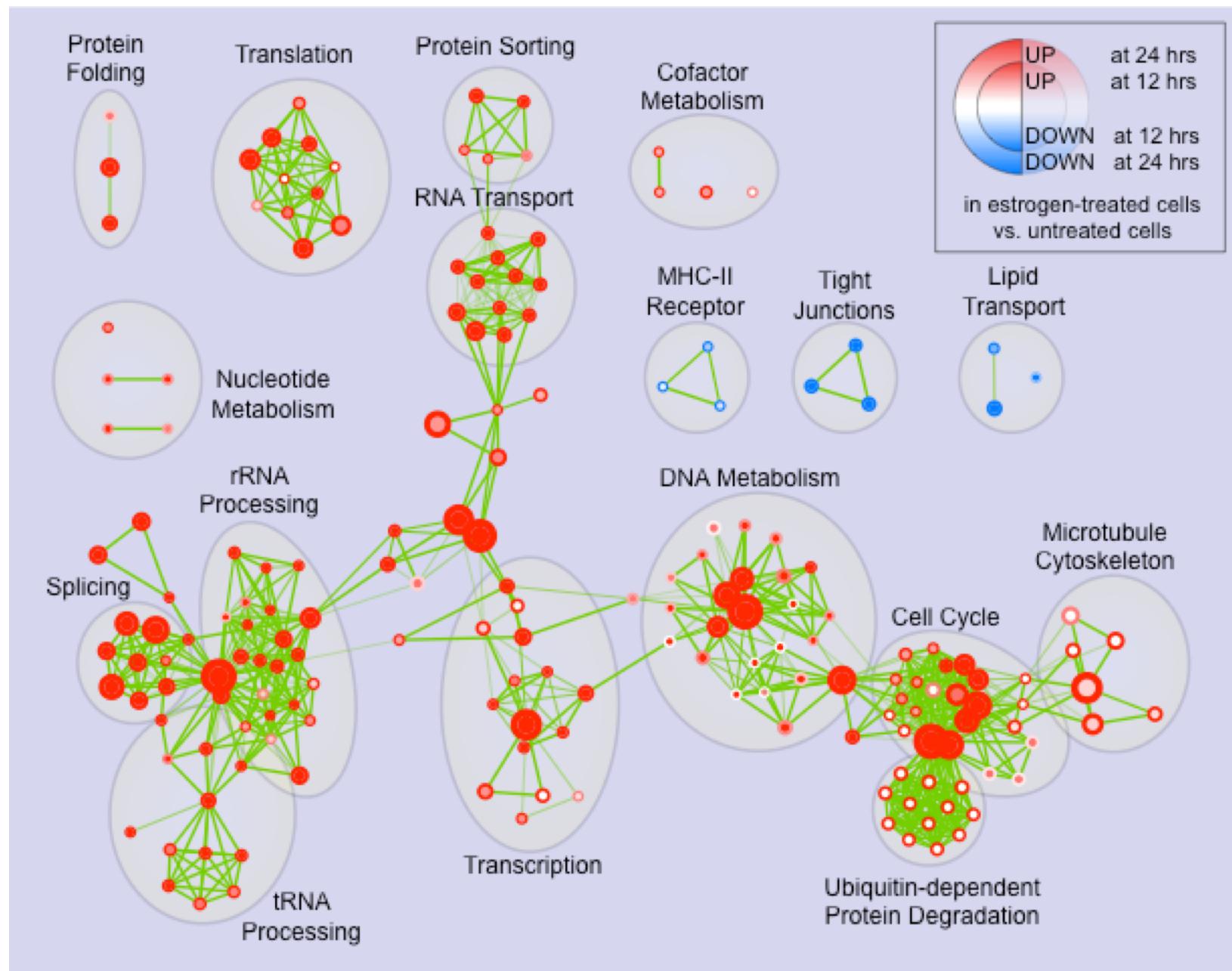
Comparison of two enrichments

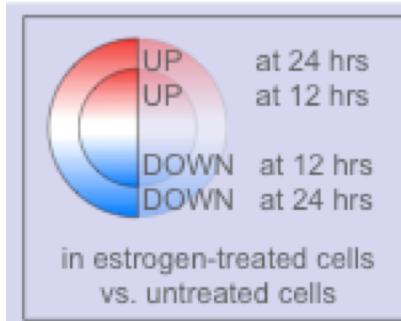
Estrogen treatment of breast cancer cells

- Design:
2-time points, two-class

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

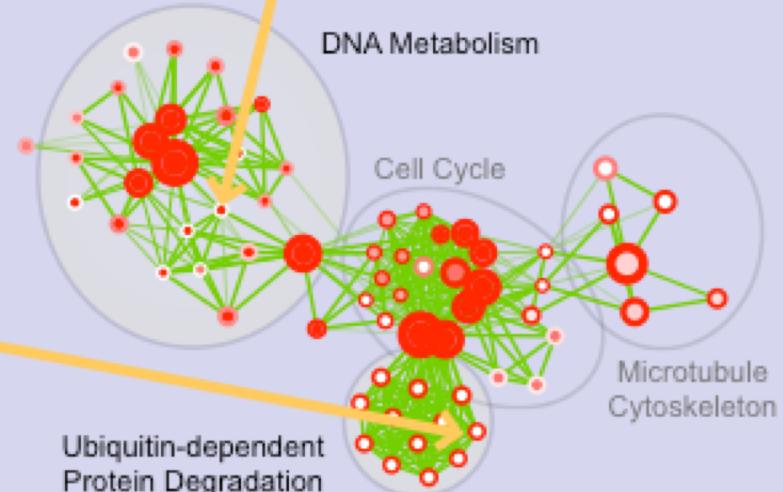
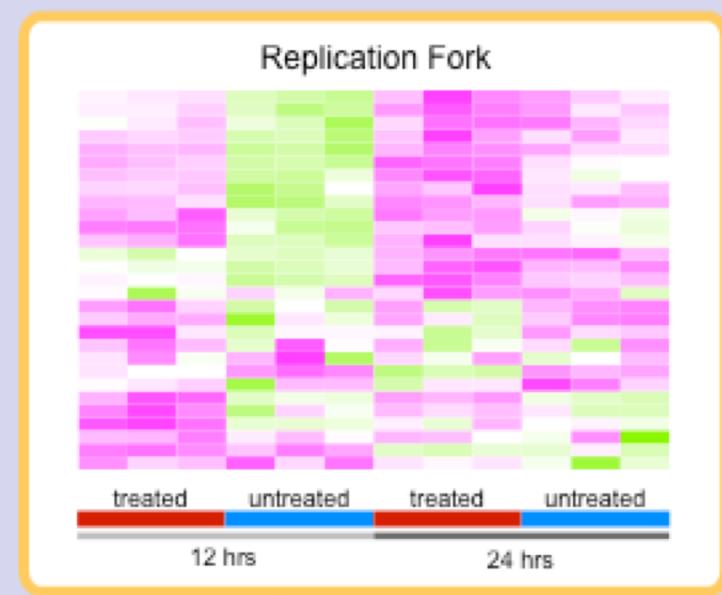
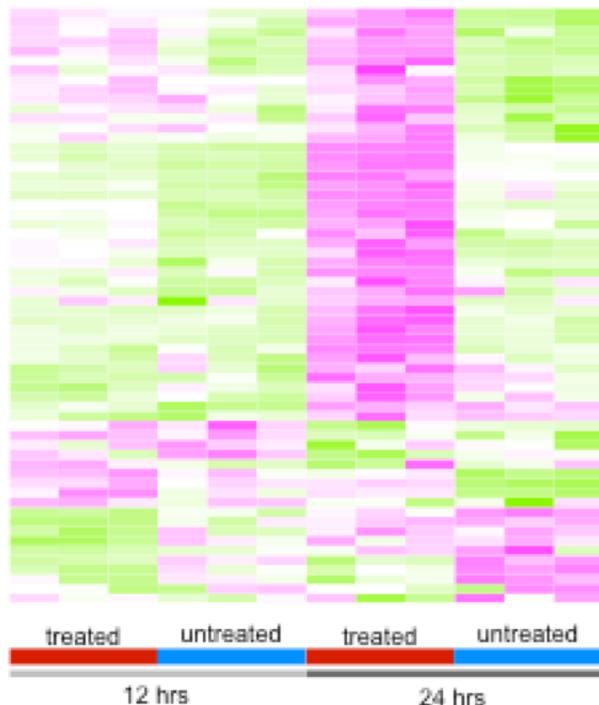
- Gene set Database:
Gene Ontology





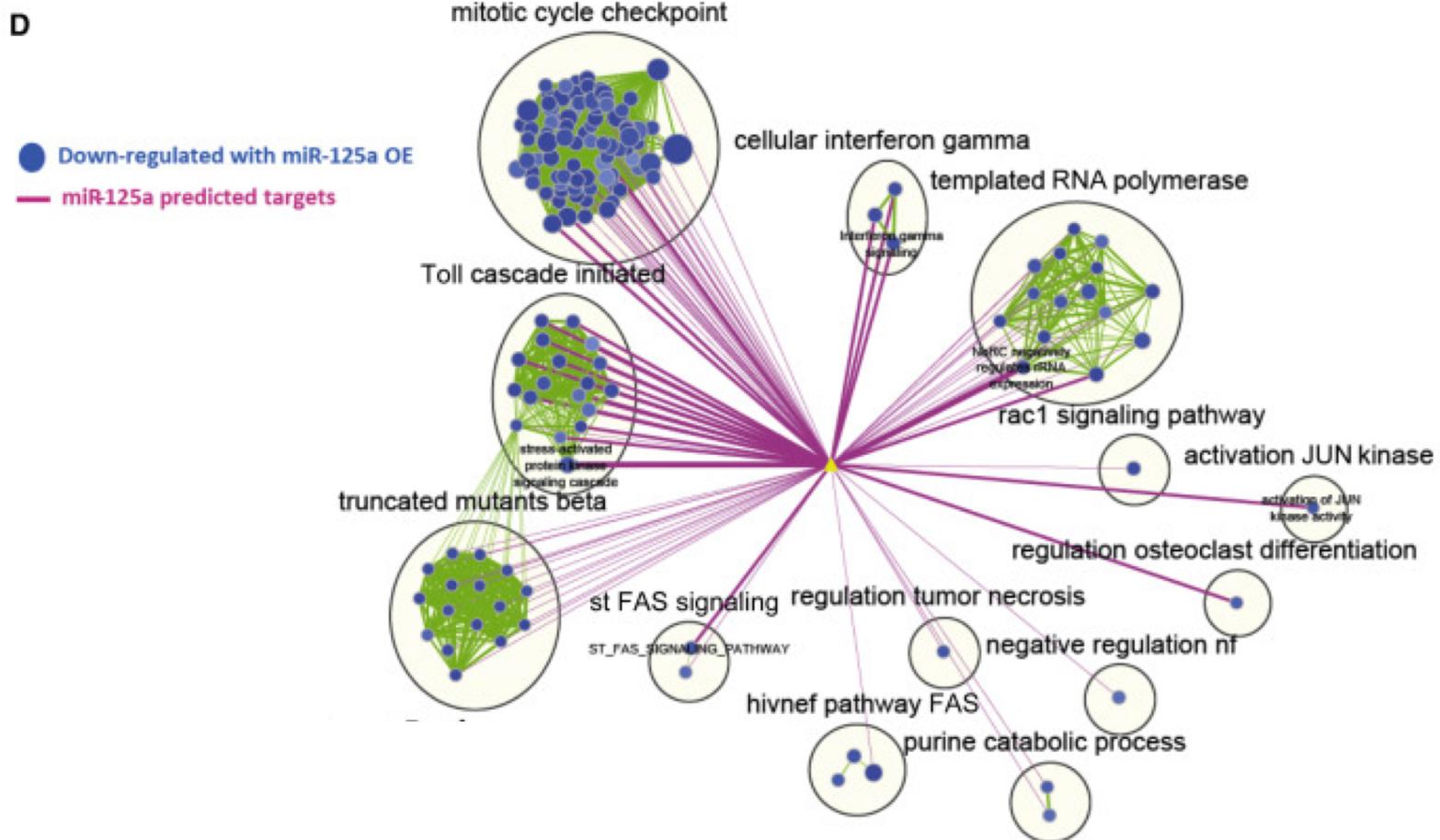
Estrogen-treated vs untreated:
█ Up
█ Down

APC-dependent Protein Degradation



Enrichment Map: use case III

Query Set Analysis

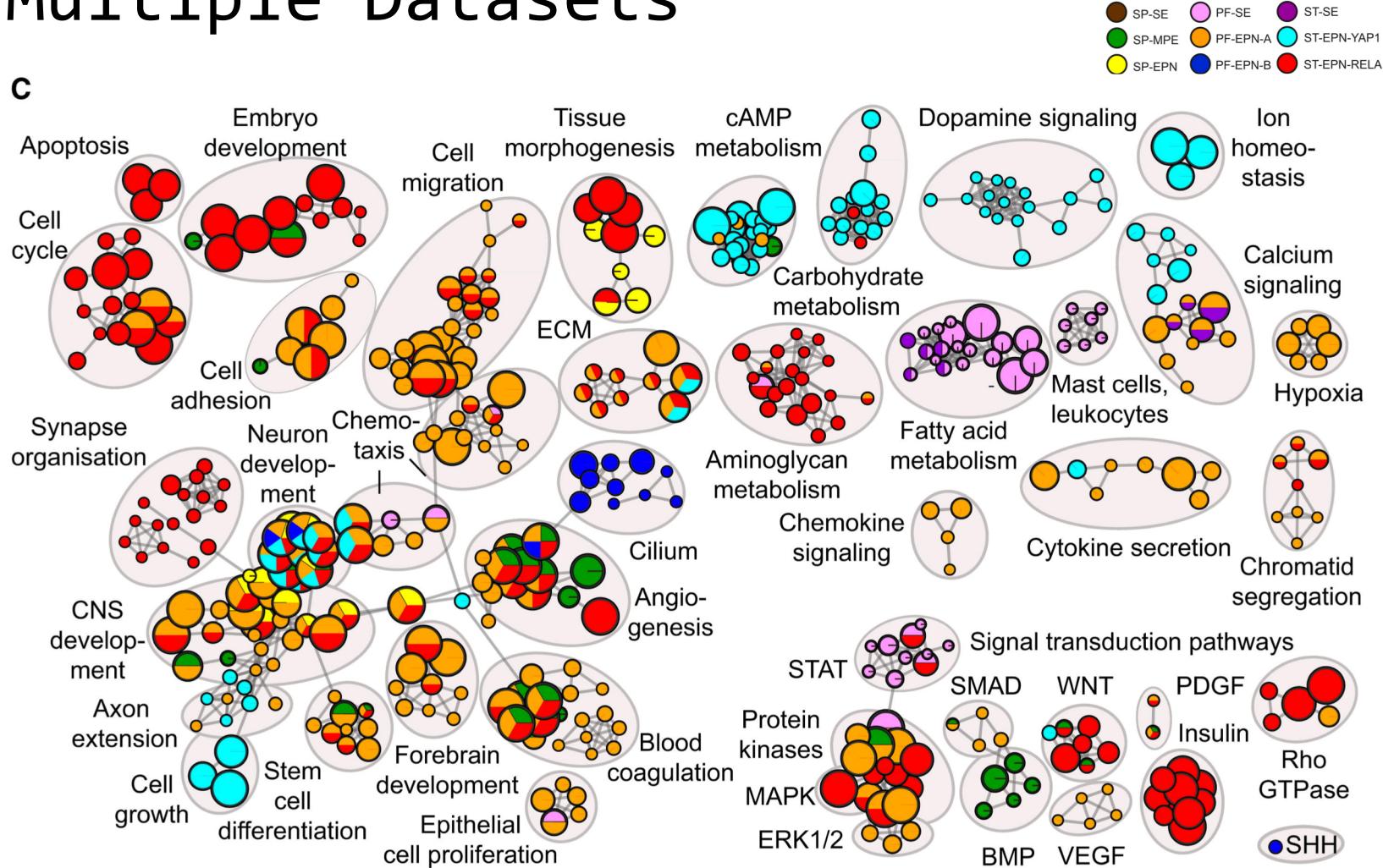


Wojtowicz EE, Lechman ER, Hermans KG, et al. Ectopic miR-125a Expression Induces Long-Term Repopulating Stem Cell Capacity in Mouse and Human Hematopoietic Progenitors. *Cell Stem Cell*. 2016;19(3):383-396

Enrichment Map: use case IV

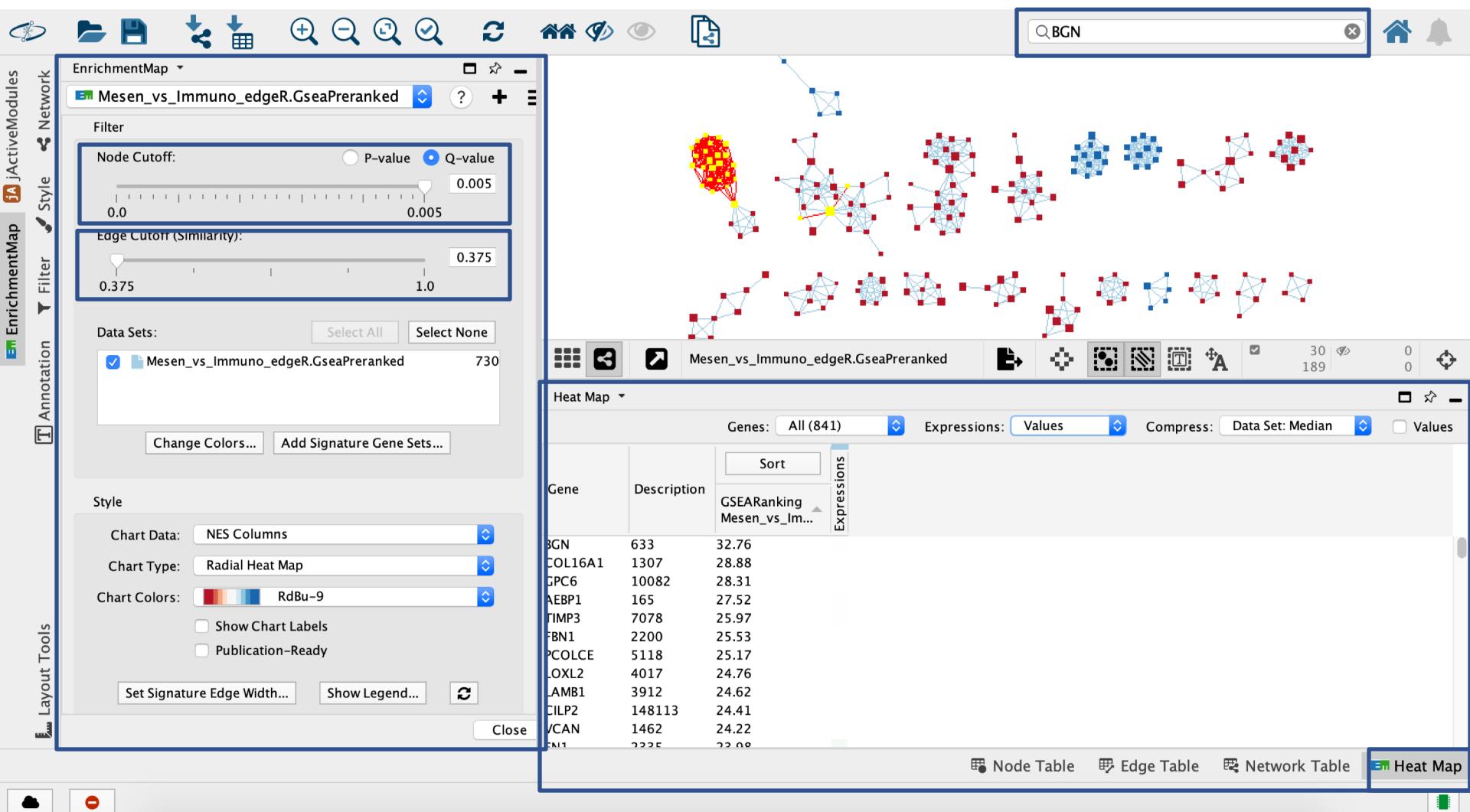
Multiple Datasets

c

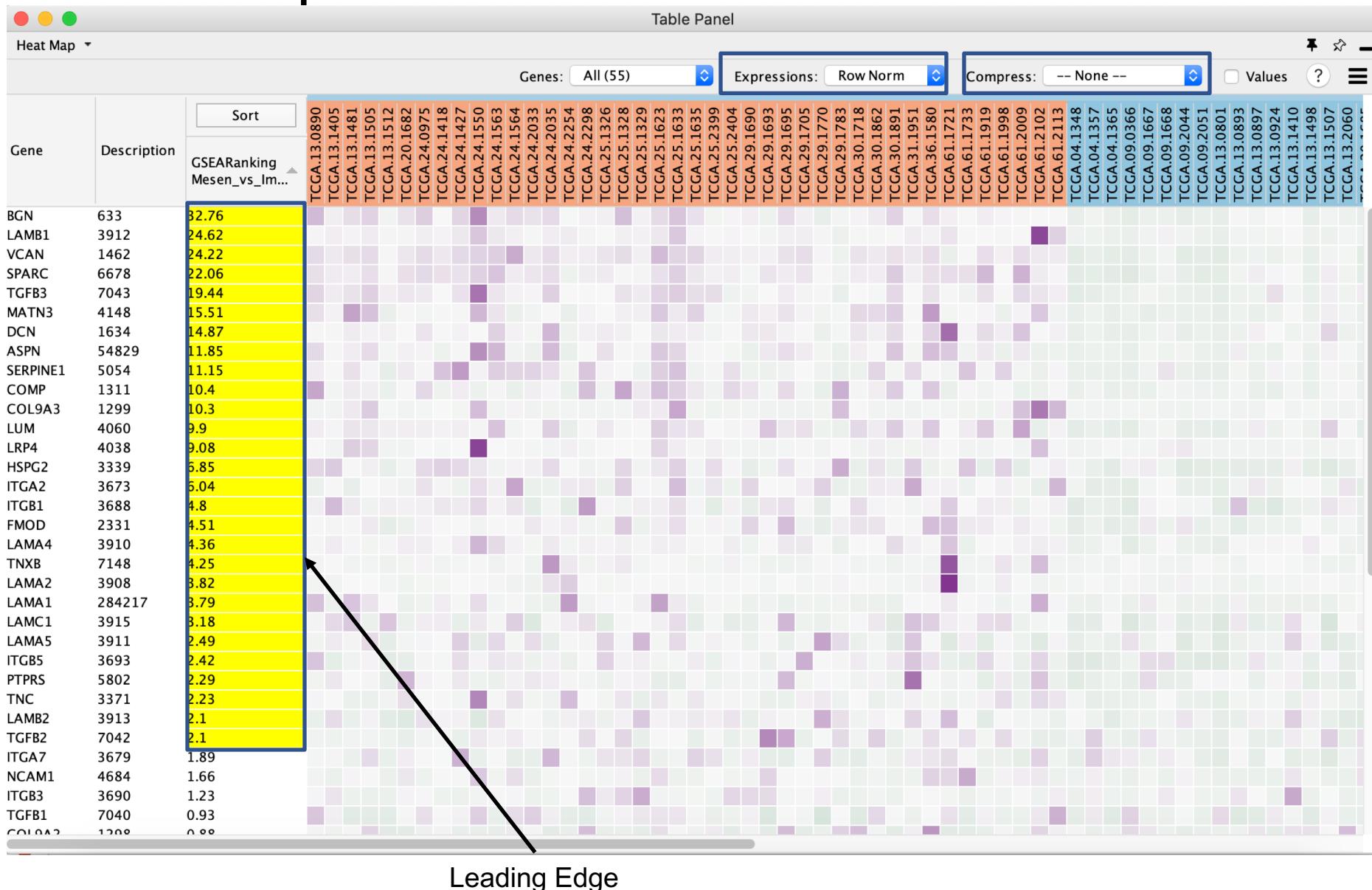


Pajtler KW, Witt H, Sill M, et al. Molecular Classification of Ependymal Tumors across All CNS Compartments, Histopathological Grades, and Age Groups. *Cancer Cell*. 2015;27(5):728-743

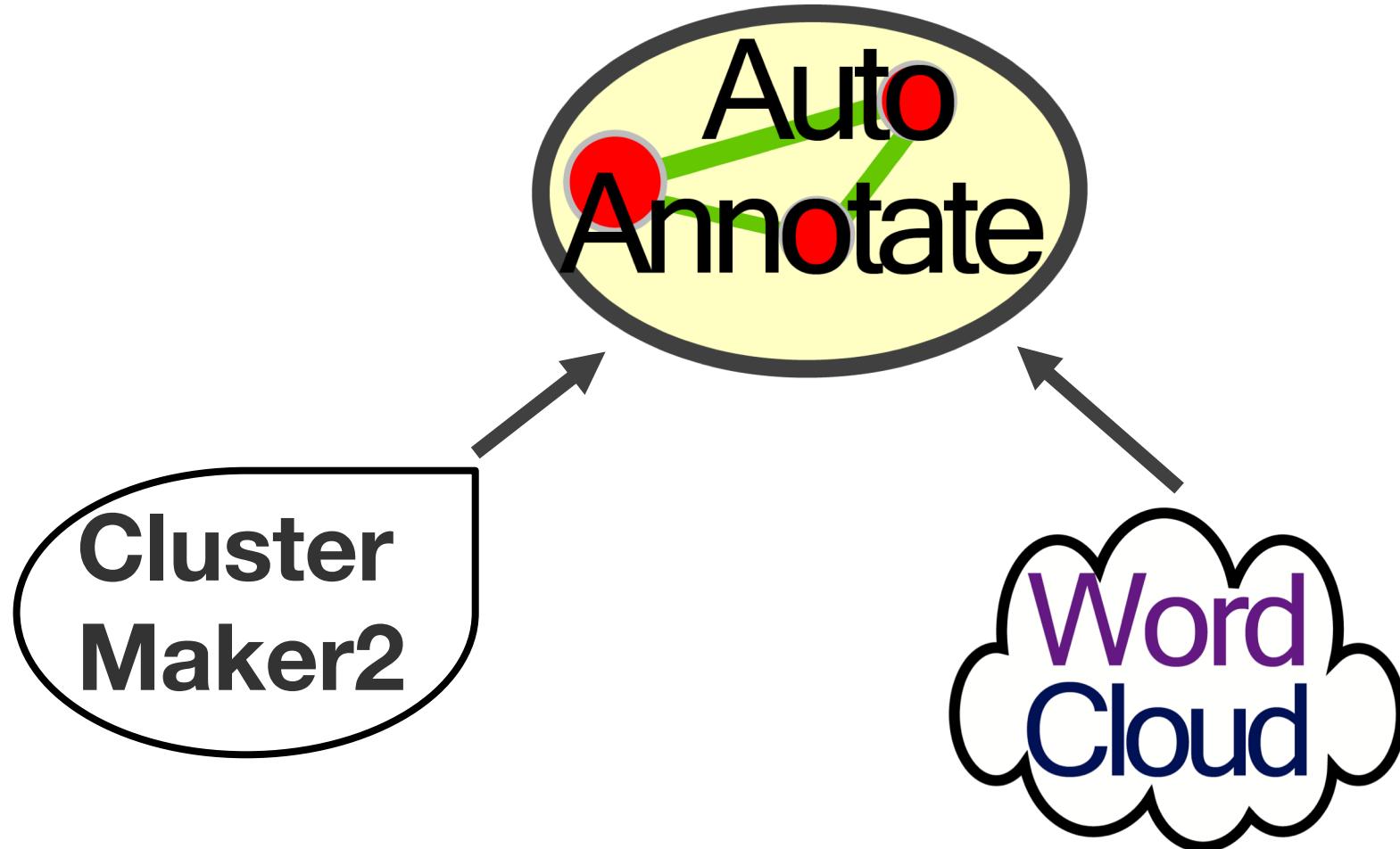
Enrichment Map Features



Heatmap Panel



Enrichment Map Features

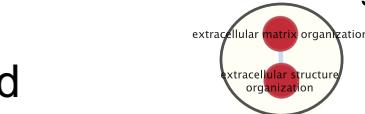


AutoAnnotate

Steps:

1. Cluster the Network

extracellular structure organization



2. For each cluster, find frequent words in node labels

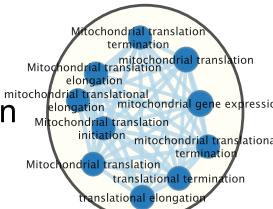
hallmark epithelial mesenchymal



3. Select top 3 words to create labels

4. Possibility to move or edit labels to make figure publication ready.

mitochondrial translational translation

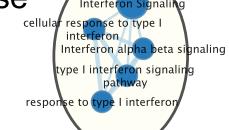


interferon gamma signaling



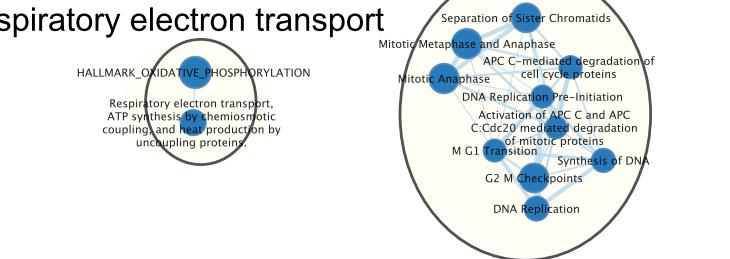
cellular response type

hallmark interferon response

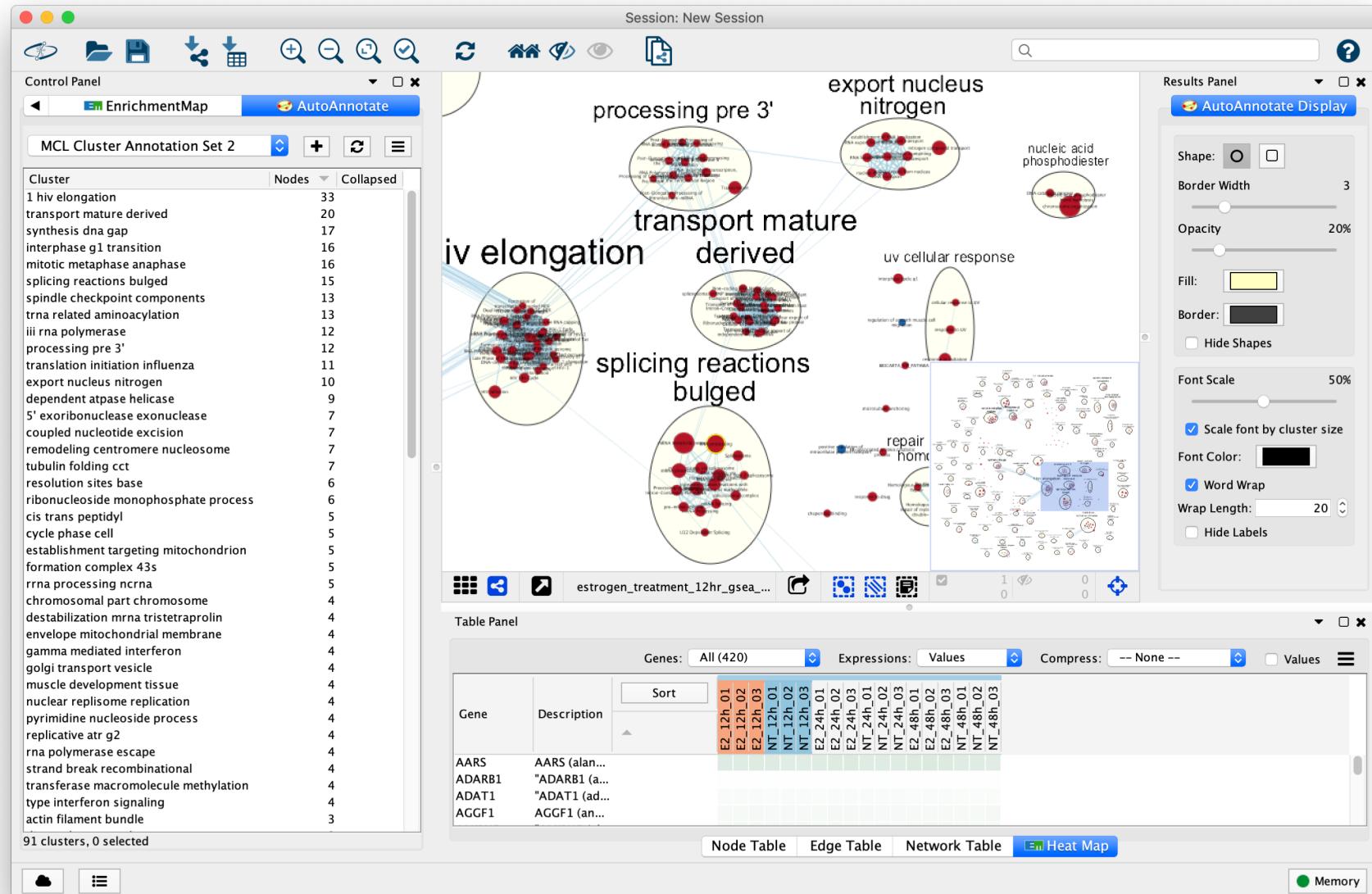


apc dna mitotic

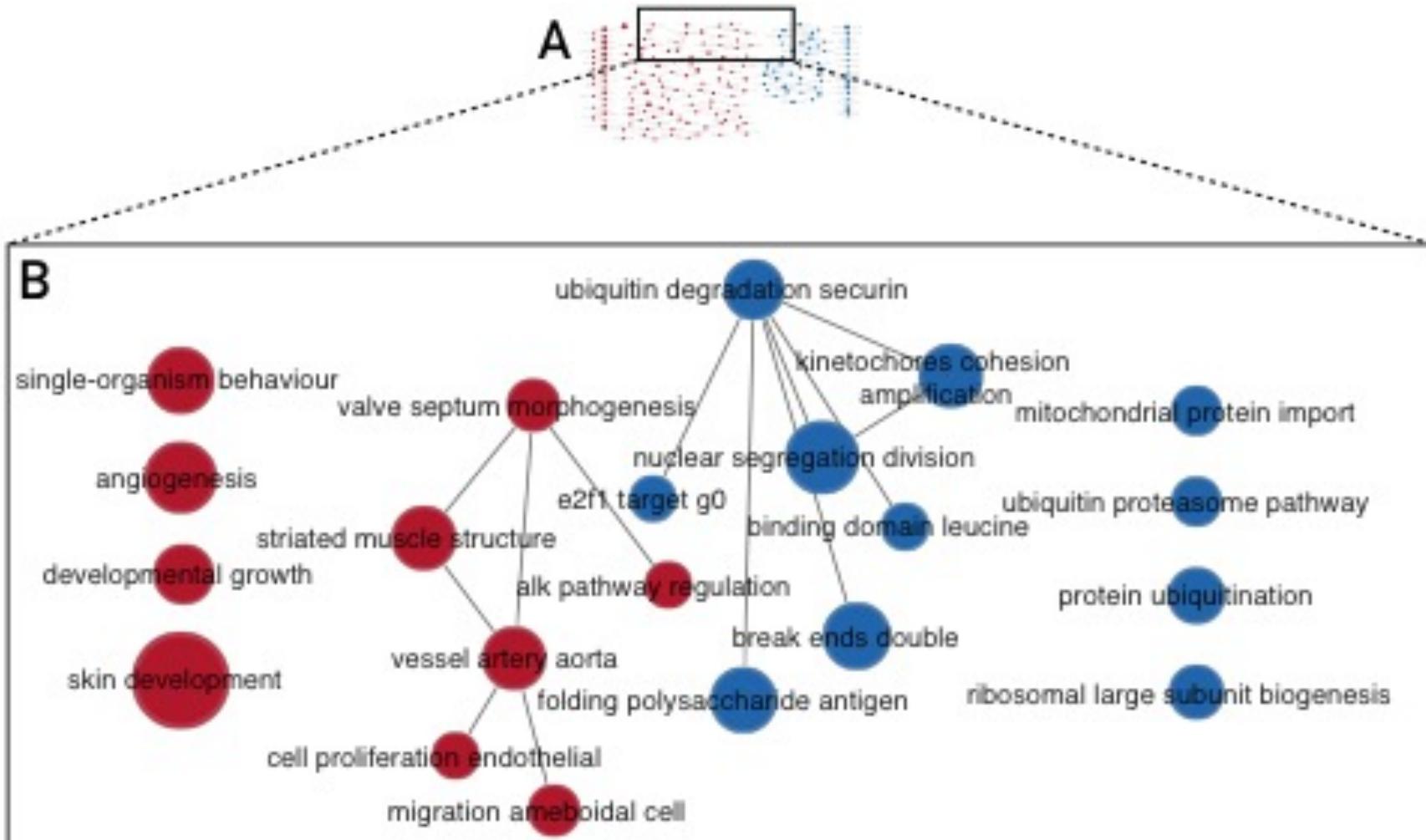
respiratory electron transport



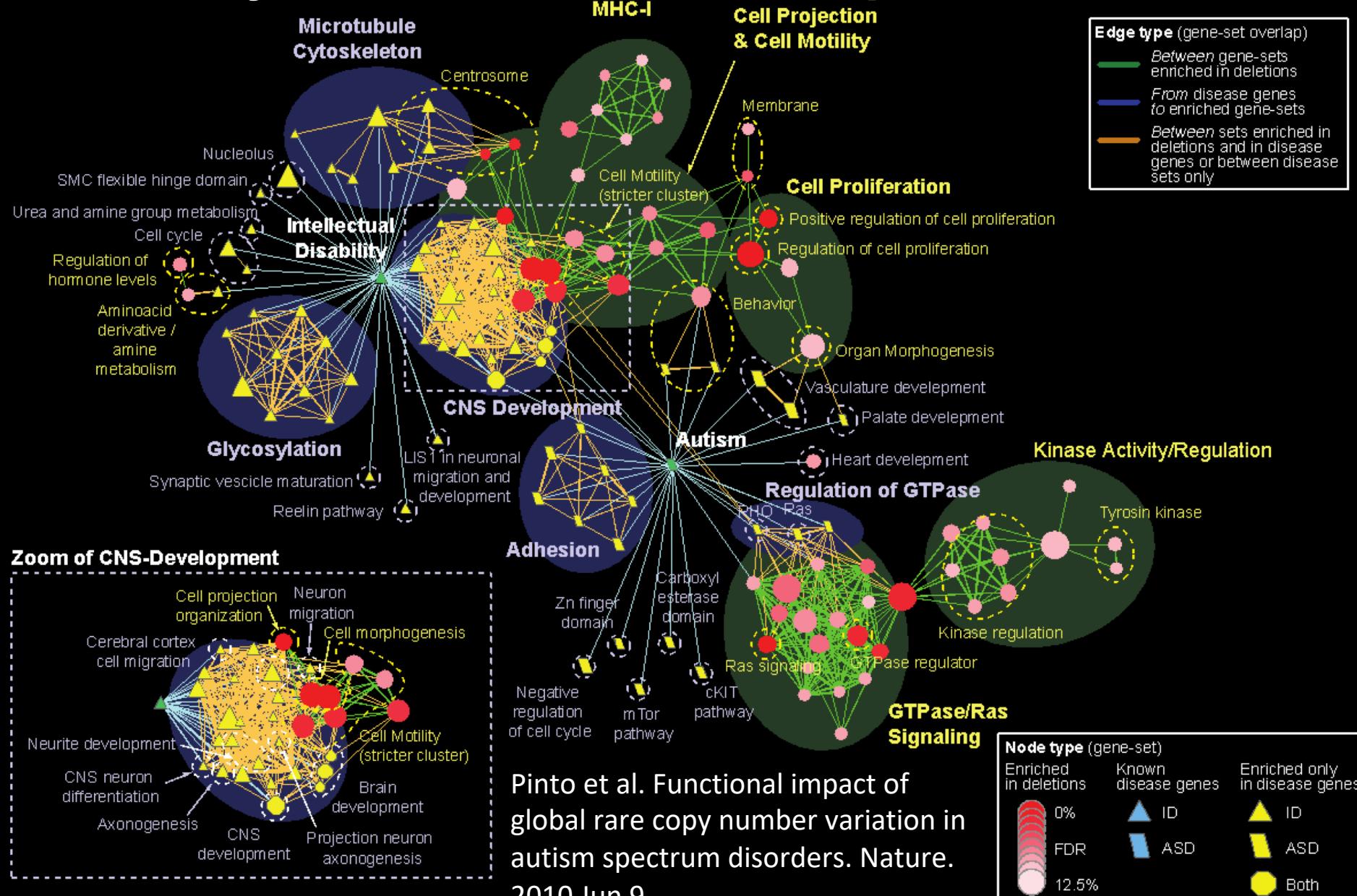
AutoAnnotate



Collapsed network



Pathways Enriched in Autism Spectrum Disorder



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

compute | calcul
canada | canada



Workshop Sponsors:

Canadian Centre for
Computational
Genomics

MicM McGill initiative in
Computational Medicine