Cytoscape

Network Visualization and Analysis

Scooter Morris

Nadezhda Doncheva

May 31, 2016

EMBO Practical Course

Budapest, Hungary

Outline

- Biological Networks
  - Why Networks?
  - Biological Network Taxonomy
  - Analytical Approaches
  - Visualization
- Introduction to Cytoscape
- Hands on Tutorial
  - Data import
  - Layout and apps

Introductions

- John “Scooter” Morris
  - 2010-Current
    - Adjunct Assistant Professor, Pharmaceutical Chemistry
  - 2004-Current
    - Director, NCRR Resource for Biocomputing, Visualization, and Informatics (RBVI) @ UCSF
  - 1985-2004
    - Principal Systems Architect: Genentech, Inc.
  - Cytoscape core team since 2006
  - Author of several Cytoscape plugins
    - SFLDLoader, *structureViz, clusterMaker, chemViz*, metanodePlugin, groupTool, commandTool, bioCycPlugin

Introductions

- Nadezhda T. Doncheva
  - 2016-Current
    - Postdoc researcher at the Novo Nordisk Foundation Center for Protein Research & Center for non-coding RNA in Technology and Health in Copenhagen
  - 2010-2016:
    - Research scientist and PhD student at the Max Planck Institute for Informatics in Saarbrücken
  - Author or co-developer of several Cytoscape plugins:
    - RINalyzer, structureViz2, setsApp, NetworkAnalyzer, NetworkPrioritizer

Introductions

- Barry Demchak
  - 2012-Current
    - Chief Software Architect, Project Manager for National Resource for Network Biology (NRNB, Ideker Lab)
  - 2005-2012
    - PhD Computer Science and Engineering, UC San Diego
  - 1987-current
    - President, Torrey Pines Software, Inc
  - Cytoscape core team since 2012
  - Architect of Cytoscape Cyberinfrastructure

Introductions

- Christian Zmasek
  - 2015-Current
    - Cytoscape core team (Ideker Lab)
  - 2006-2015
    - Postdoc (Sanford-Burnham) Comparative functional genomics
- Rintaro Saito, PhD
  - 2014-Current
    - Associate Project Scientist (Kumar Lab)
  - 2011-2014
    - Visiting Assistant Professor (Ideker Lab)
  - 2002-2011
    - Assistant Professor (Keio University)

Introductions

- You?
    - Clinician
    - Biologist
    - Bioinformatician
    - Computer Scientist
    - Chemist
    - Other?

Installation

- How many have installed:
  - Cytoscape 3.4.0
  - Apps:
    - Omics Analysis Collection

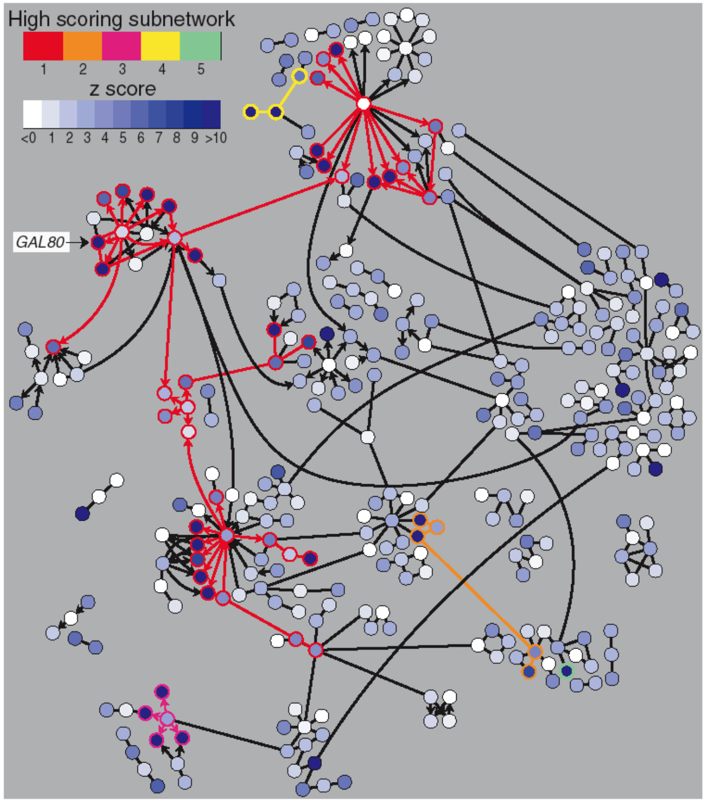
Why Networks?

- **Networks are…**
  - Commonly understood
  - Structured to reduce complexity
  - More efficient than tables

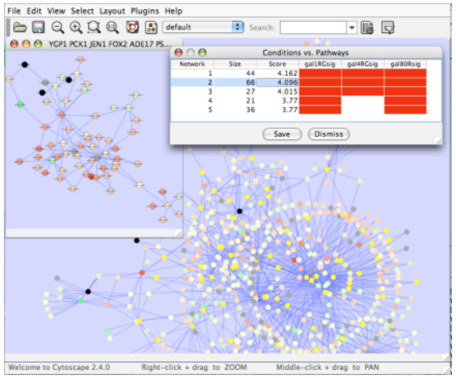
- **Network tools allow…**

  Analysis
    - Characterize network properties
    - Identify hubs and subnets
    - Classify, quantify and correlate, e.g., cluster nodes by associated data

  Visualization
    - Explore data overlays
    - Interpret mechanisms, e.g., how a process is modulated or attenuated by a stimulus



Applications of Network Biology



Network	Size	Score	gal1RGsig	gal4RGsig	gal80Rsig
1	44	4.162	🟥	🟥	🟥
2	66	4.096	🟥	🟥	🟥
3	27	4.015	🟥	🟥	🟥
4	21	3.77	🟥		🟥
5	36	3.77	🟥		🟥

jActiveModules, UCSD

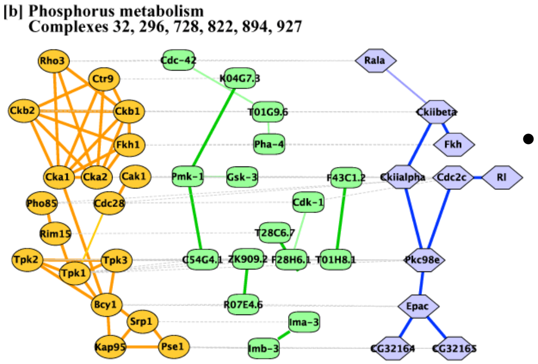
- **Gene Function Prediction**
  shows connections to sets of genes/proteins involved in same biological process

- **Detection of protein complexes/subnetworks**
  discover modularity & higher order organization (motifs, feedback loops)

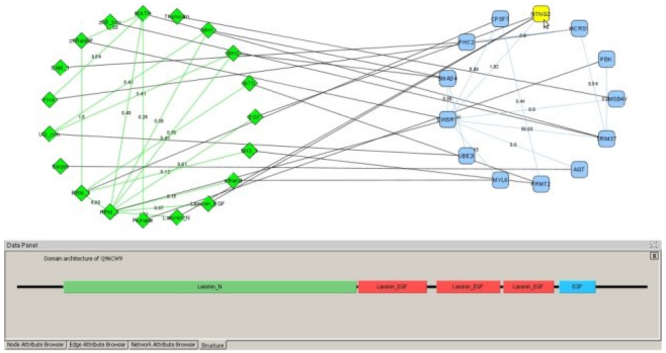
- **Network evolution**
  biological process(s) conservation across species

- **Prediction of interactions & functional associations**
  statistically significant domain-domain correlations in protein interaction network to predict protein-protein or genetic interaction

mCode, University of Toronto

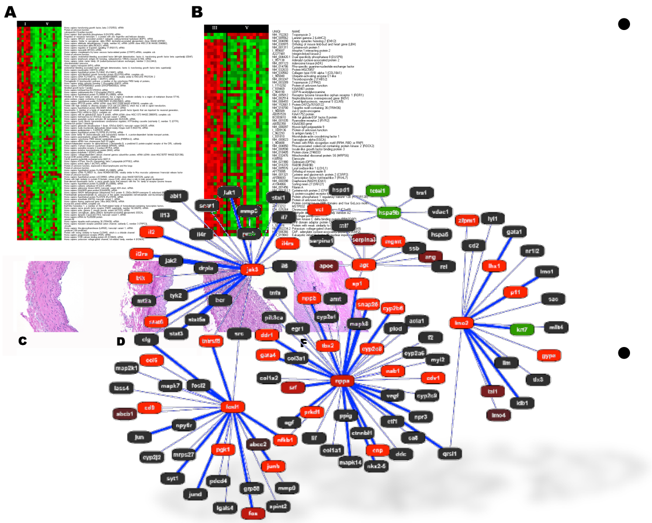


PathBlast, UCSD

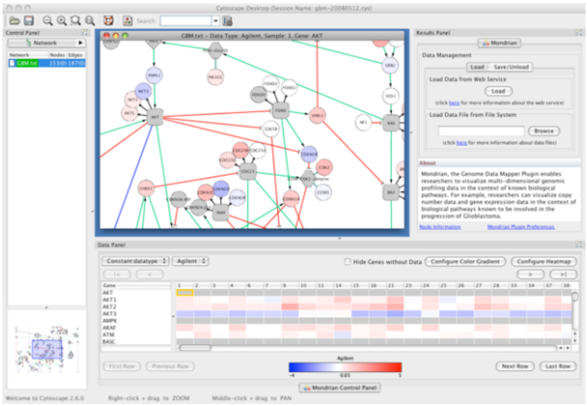


DomainGraph, Max Planck Institute

Applications in Disease



Agilent Literature Search

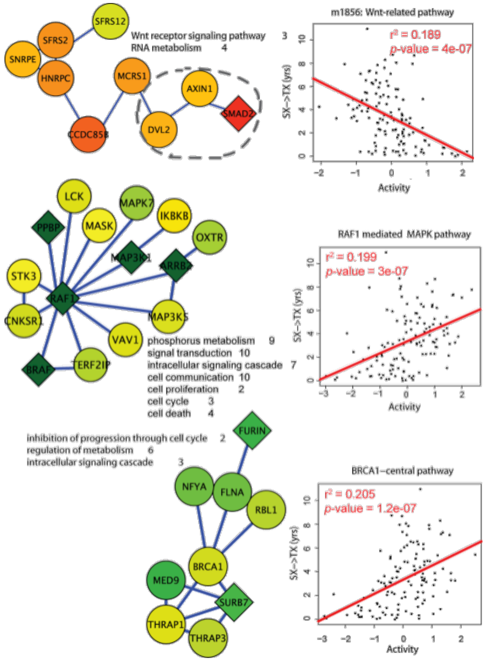


Mondrian, MSKCC

- **Identification of disease subnetworks** – identification of disease network subnetworks that are transcriptionally active in disease.

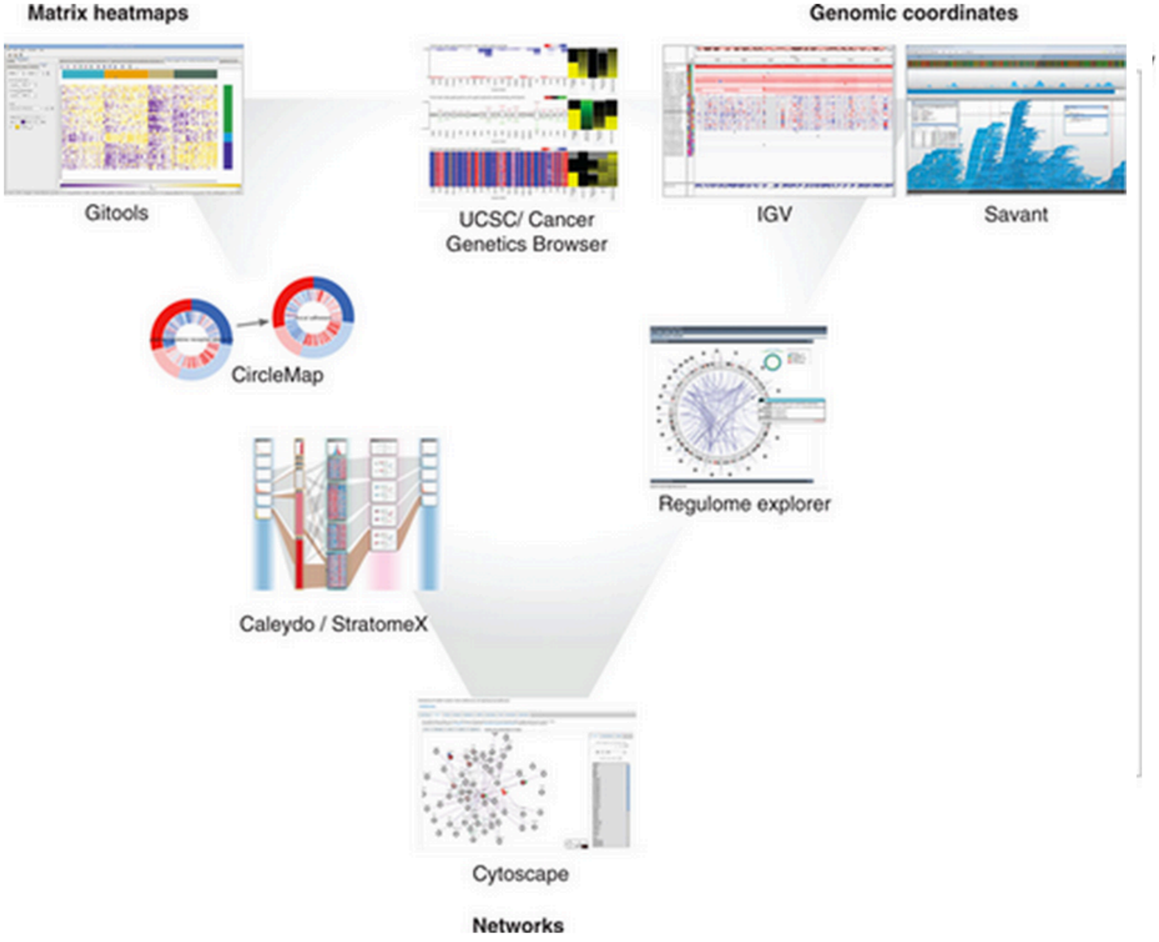
- **Subnetwork-based diagnosis** – source of biomarkers for disease classification, identify interconnected genes whose aggregate expression levels are predictive of disease state

- **Subnetwork-based gene association** – map common pathway mechanisms affected by collection of genotypes (SNP, CNV)



PinnacleZ, UCSD

The Challenge

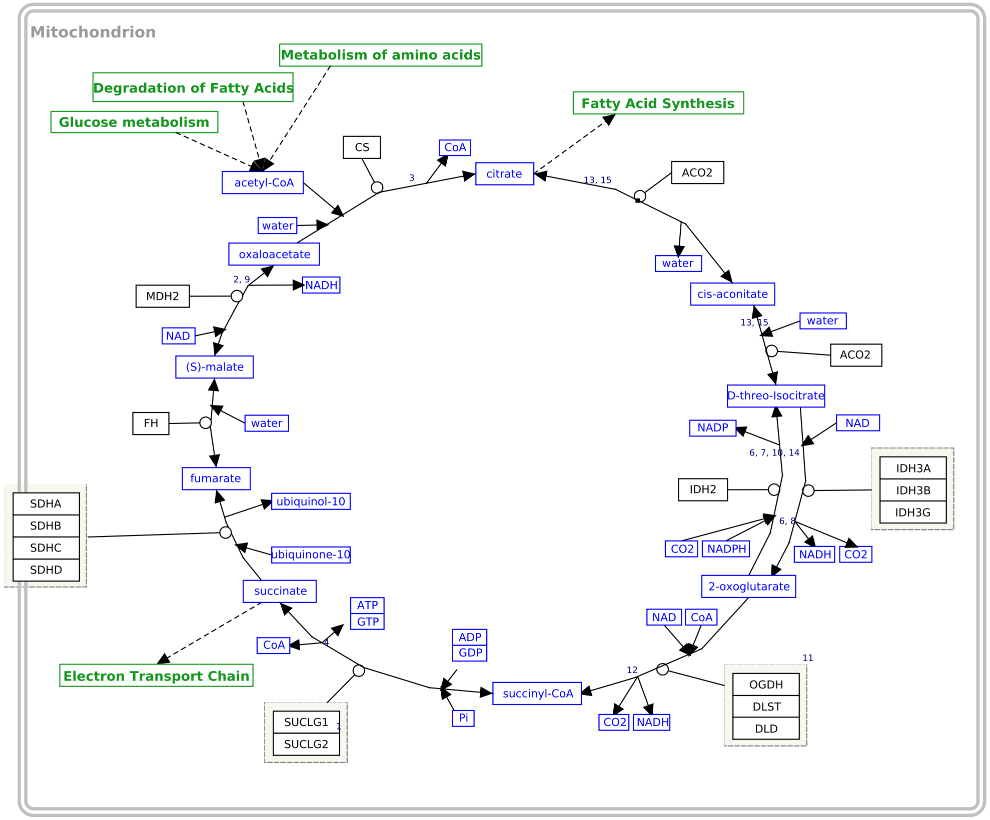


Visualizing multidimensional cancer genomics data.
Schroeder et al, Genome Med. 2013 Jan 31;5(1):9.

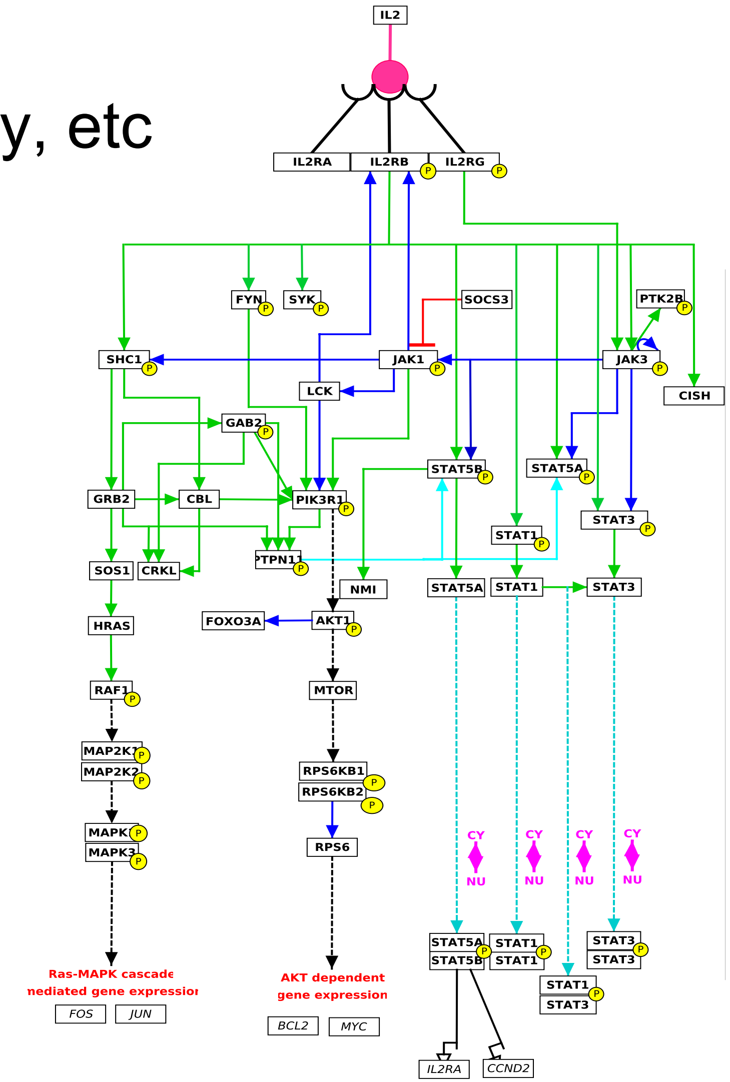
http://cytoscape-publications.tumblr.com/archive

Biological Network Taxonomy

- Pathways
  - – Signaling, Metabolic, Regulatory, etc

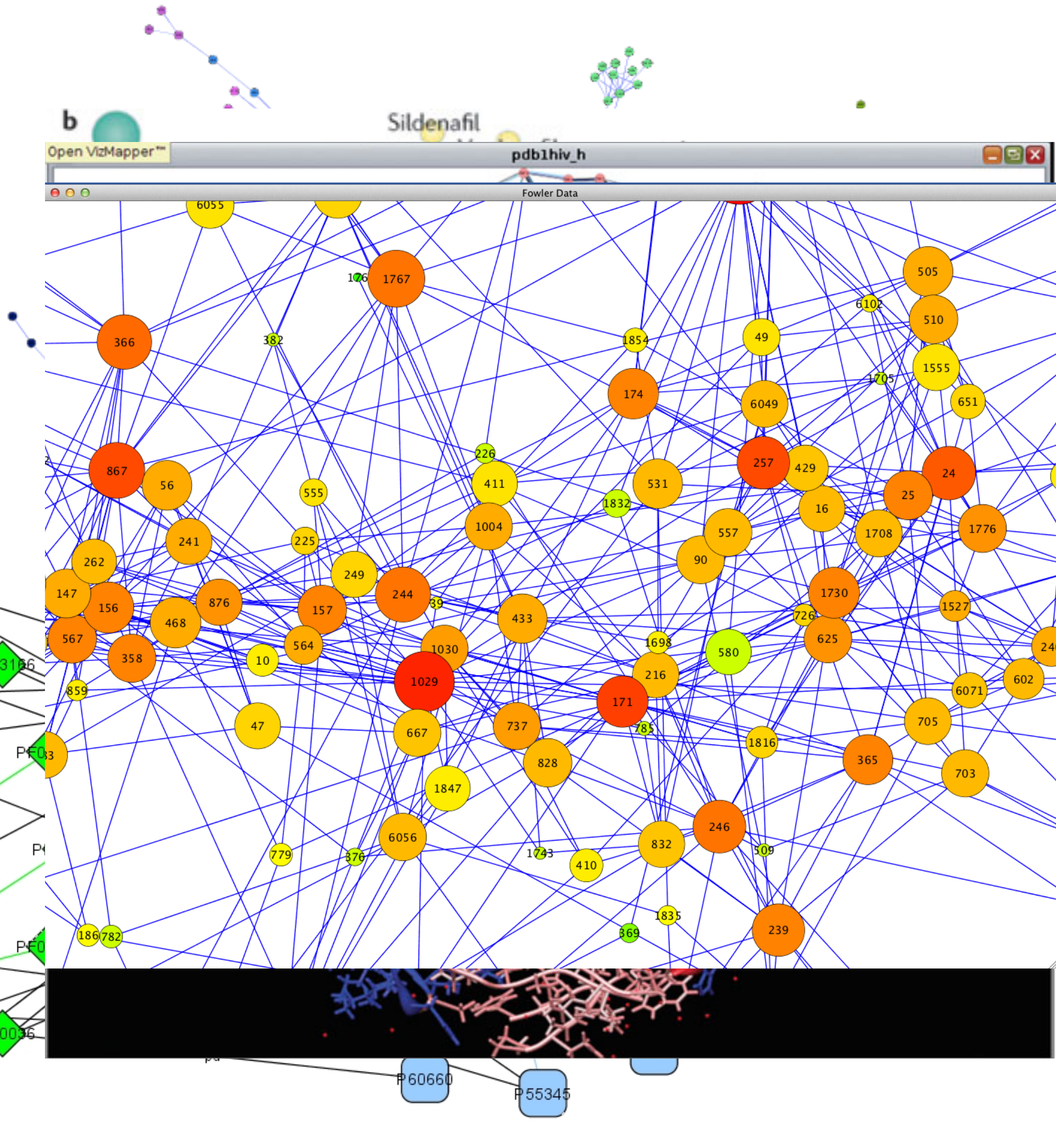


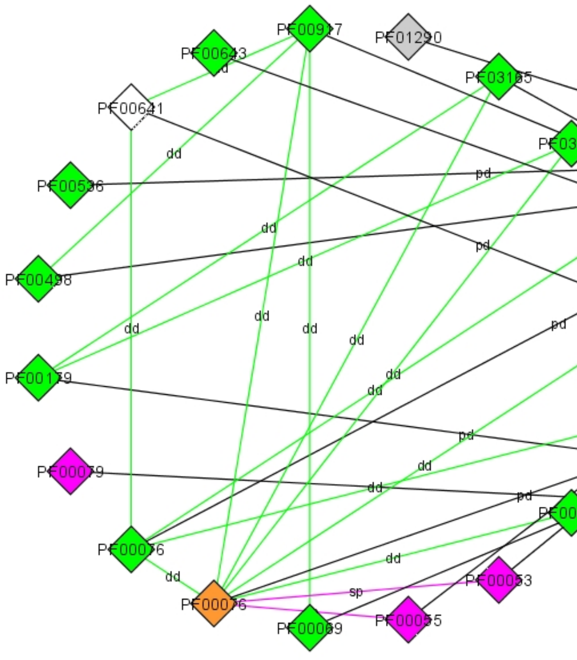
IL-2 Signaling Pathway



Biological Network Taxonomy

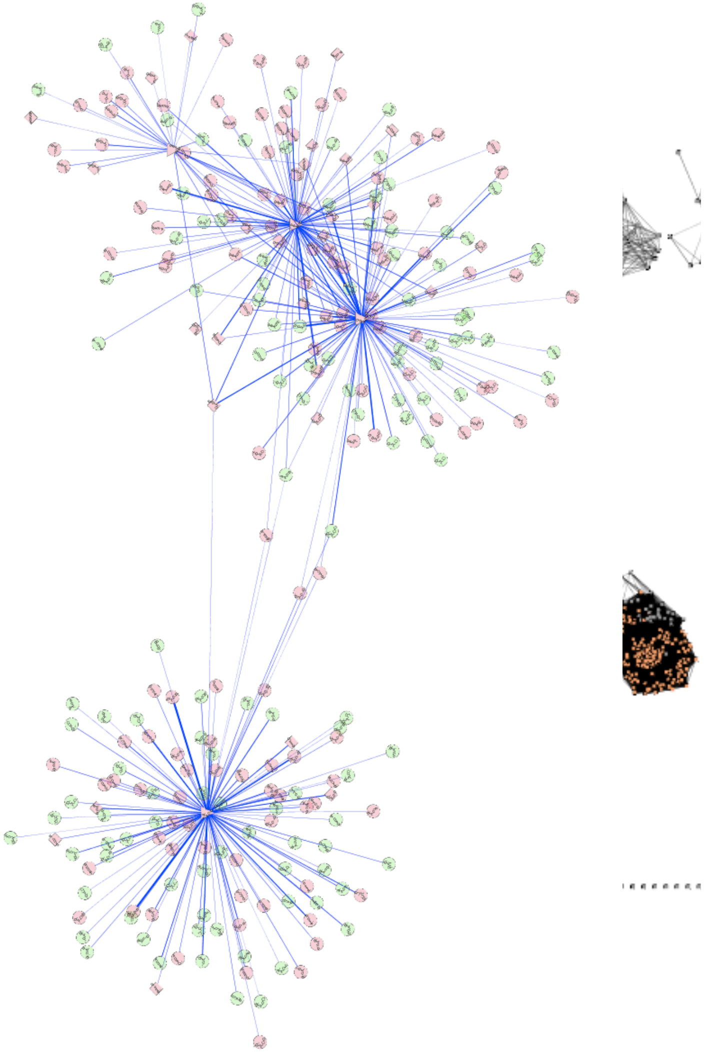
- Interactions
  - Protein-Protein
  - Protein-Ligand
  - Domain-Domain
  -





Biological Network Taxonomy

- Similarity
  - Protein-Protein
  - Chemical similarity
  - Ligand similarity (SEA)
  - Others
    - Tag clouds
    - Topic maps



Biological Network Taxonomy

Where do I get *the* network?

There is no such thing!

550 different interaction databases!

… *in 2013*

It depends on your biological question and your analysis plan.

Biological Network Taxonomy

Where do I get *the* network?

There is no such thing!

… *in 2016*

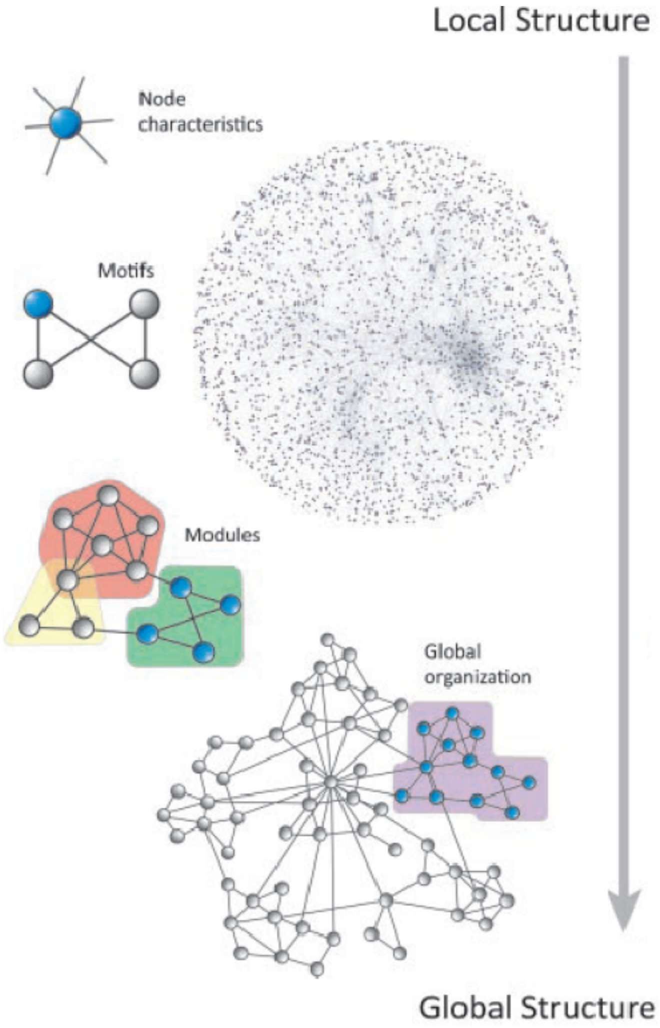
NDEx

http://ndexbio.org

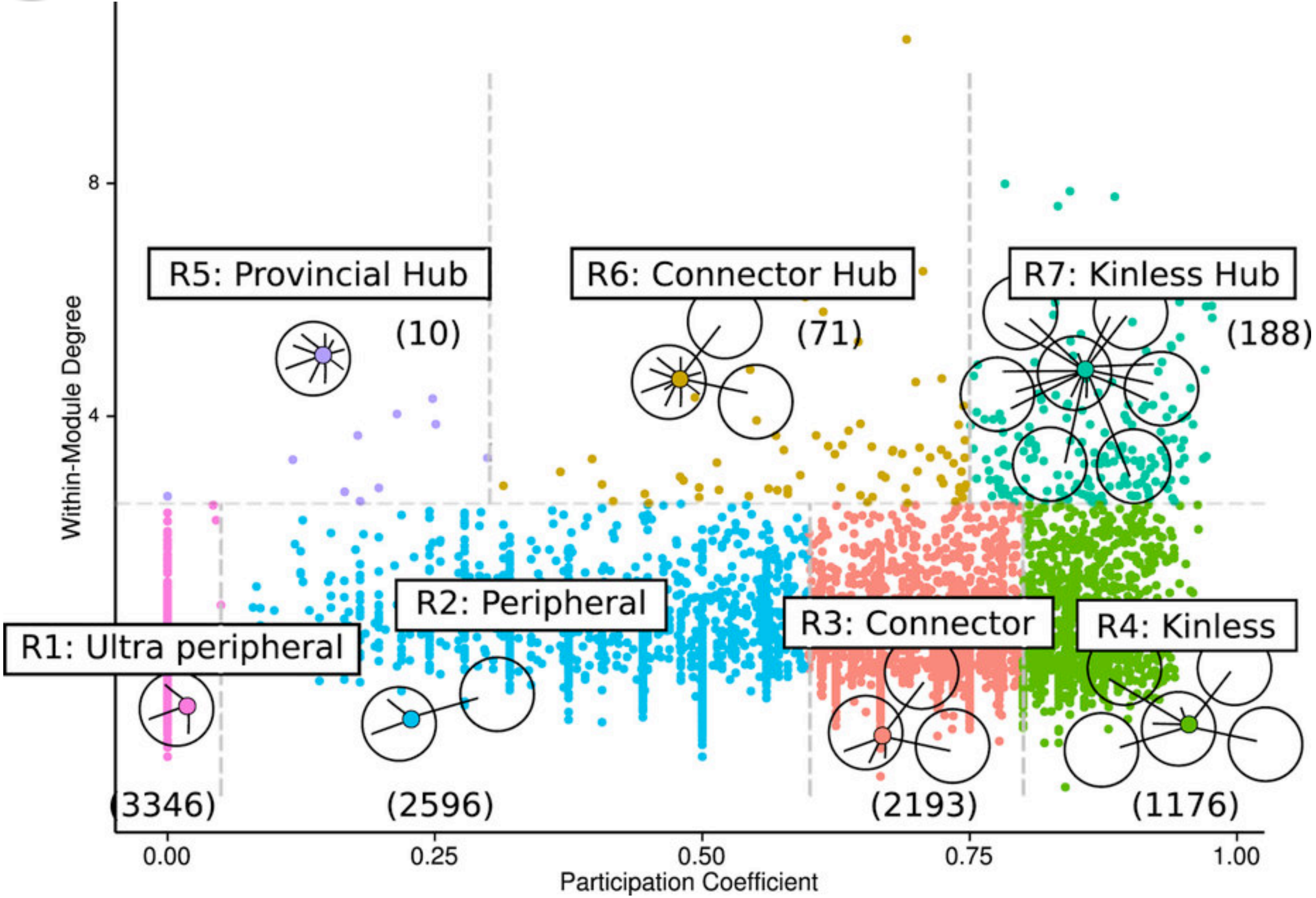
Analytical Approaches

The levels of organization of complex networks:

- Node **degree** provides information about single nodes
- Three or more nodes represent a **motif**
- Larger groups of nodes are called **modules** or **communities**
- **Hierarchy** describes how the various structural elements are combined



Analytical Approaches



Piñero, *Sci Rep,* 15 April 2016

Analytical Approaches

Network topology statistics such as node degree, degree distribution, centralitiy, clustering coefficient, shortest paths, and robustness of the network to the random removal of single nodes are important network characteristics.

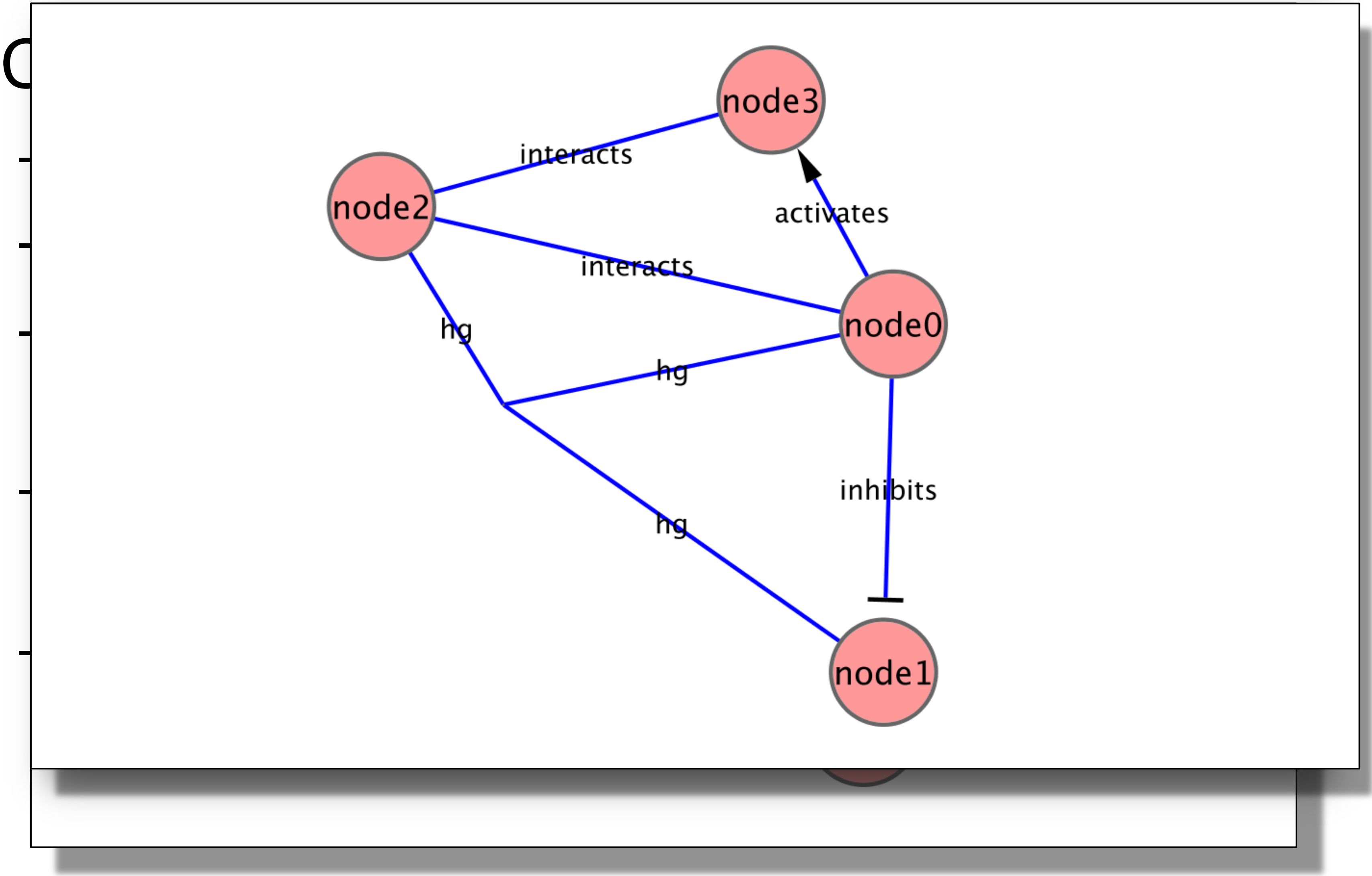
Modularity refers to the identification of sub-networks of interconnected nodes that might represent molecules physically or functionally linked that work coordinately to achieve a specific function.

Motif analysis is the identification of small network patterns that are over-represented when compared with a randomized version of the same network. Regulatory elements are often composed of such motifs.

Network alignment and comparison tools can identify similarities between networks and have been used to study evolutionary relationships between protein networks of organisms.

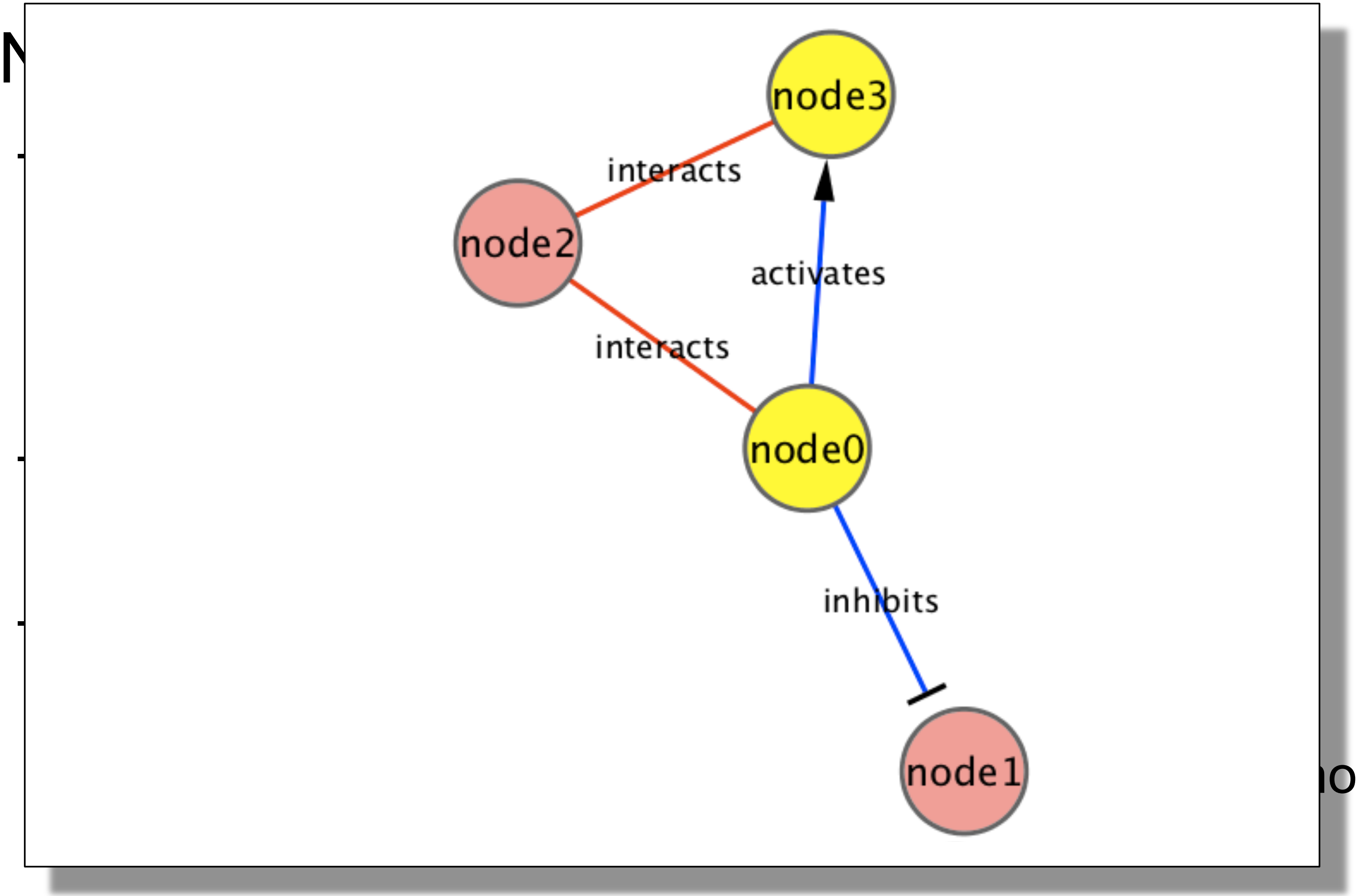
Analytical Approaches

- C



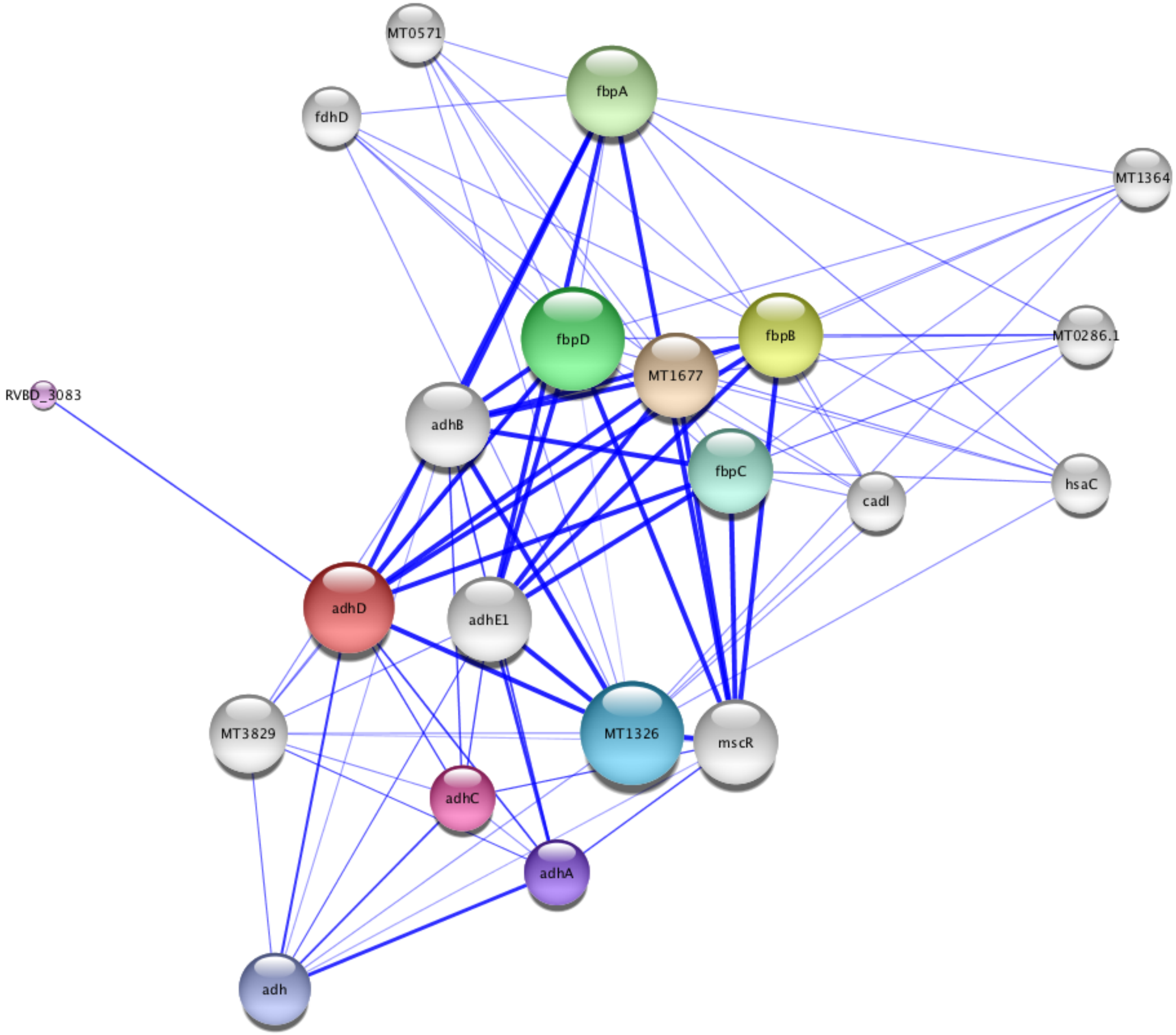
Analytical Approaches

- N

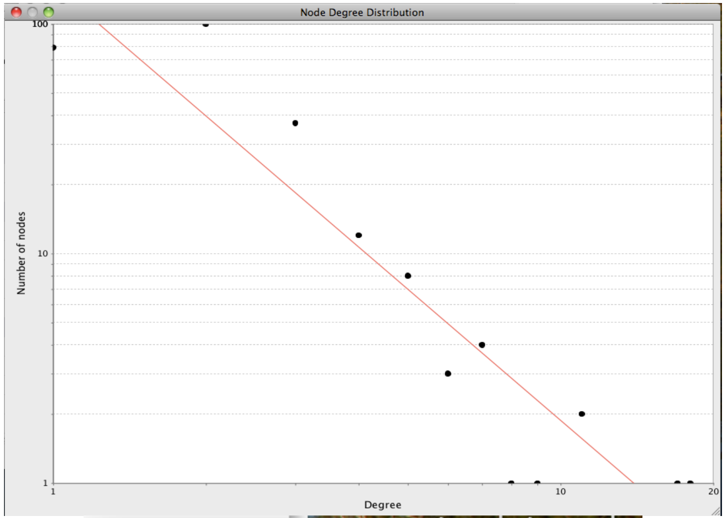


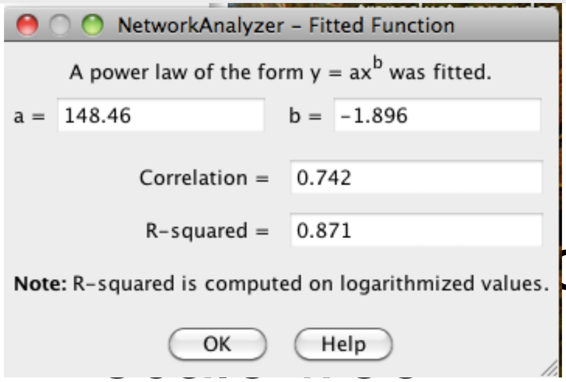
ot

Analytical Approaches



Analytical Approaches





A power law of the form $y = ax^b$ was fitted.

a = 148.46 b = −1.896

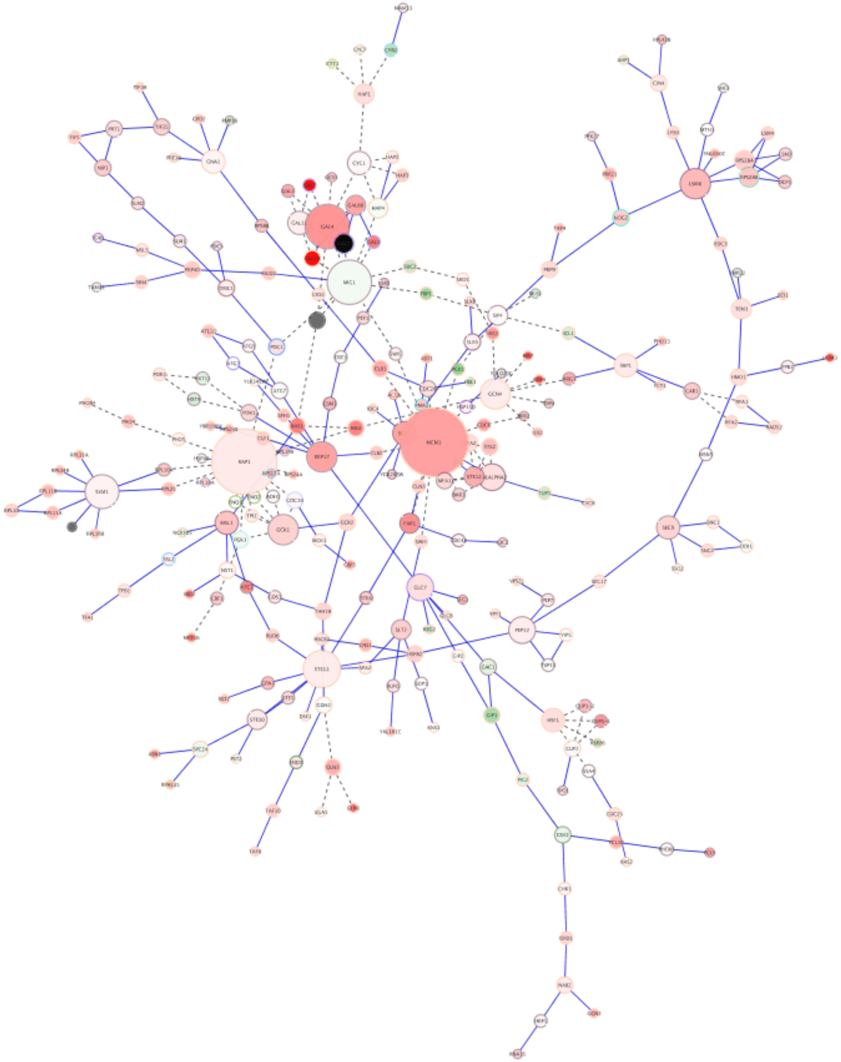
Correlation = 0.742

R-squared = 0.871

Note: R-squared is computed on logarithmized values.

h,

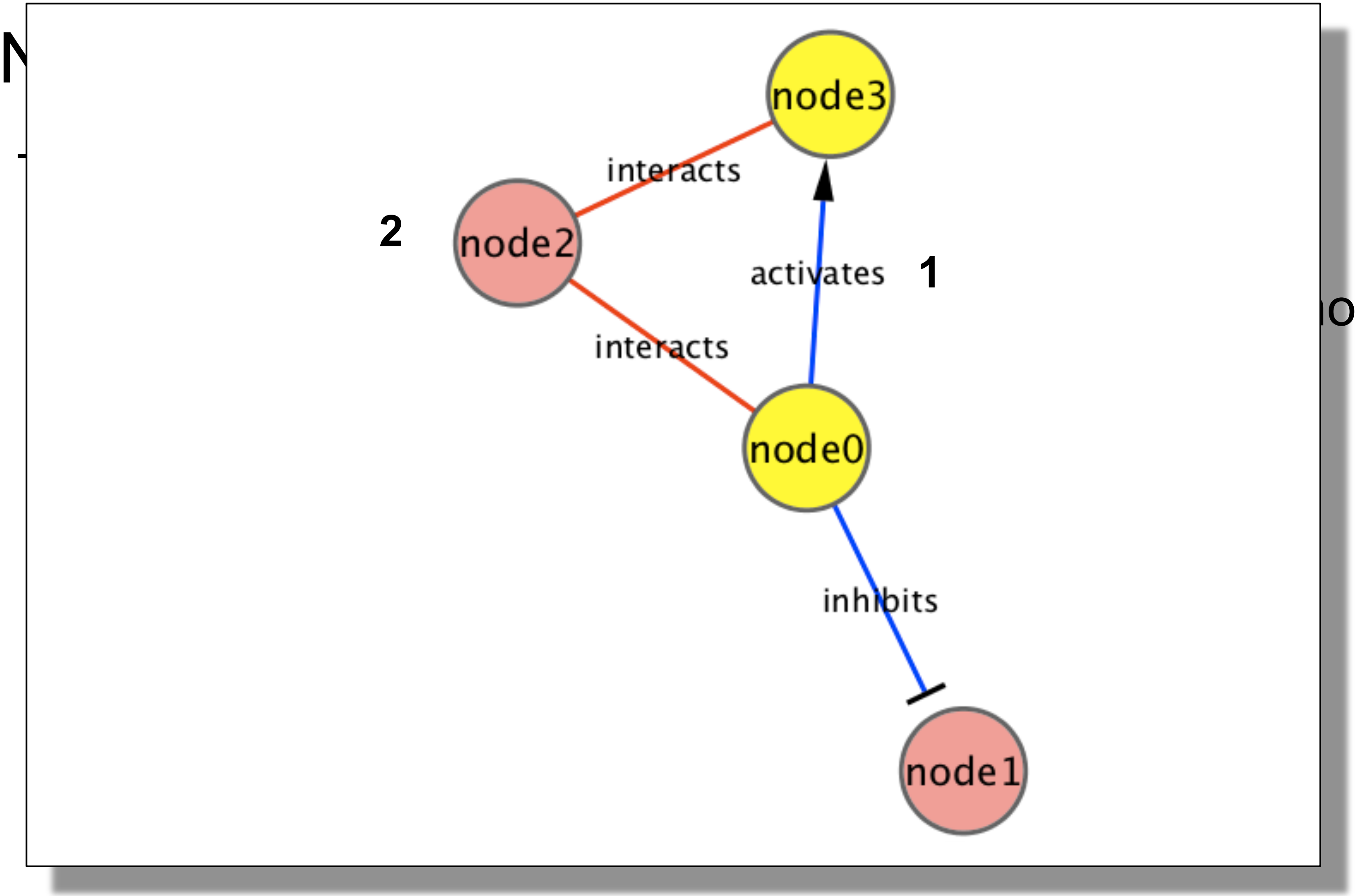
d



Analytical Approaches

- N

ot

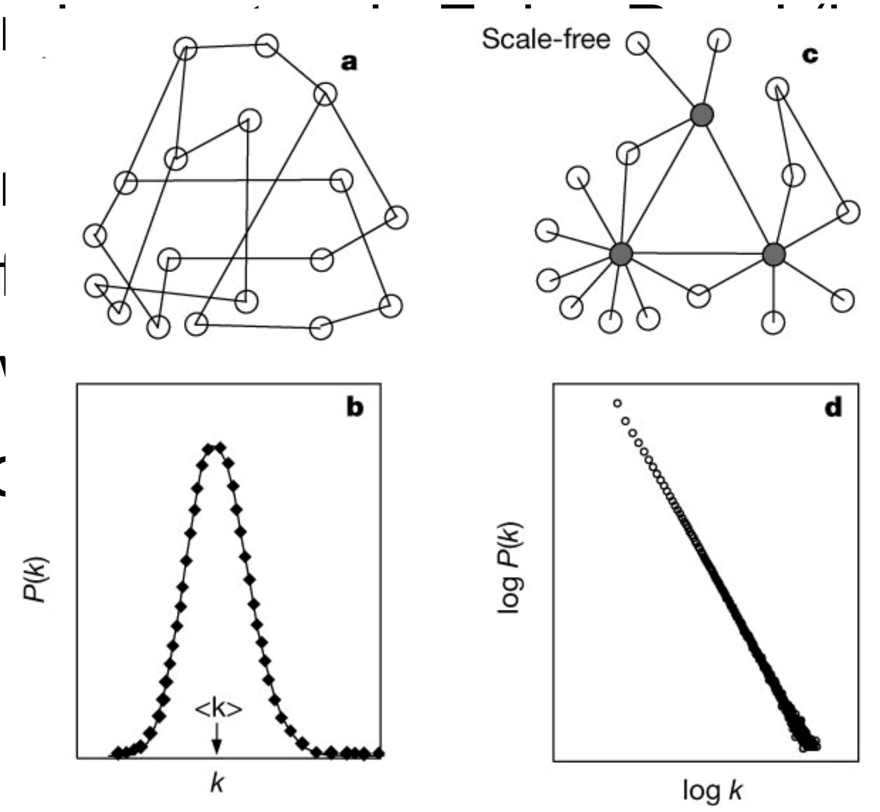


Analytical Approaches

- Small-world networks
  - any two arbitrary nodes are connected by a small number of intermediate edges
  - the network has an average shortest path length much smaller than the number of nodes in the network (Watts, Nature, 1998).
  - Interaction networks have been shown to be small-world networks (Barabási, Nature Reviews in Genetics, 2004)

Analytical Approaches

- Random networks
  - Algorithms exist to create random networks



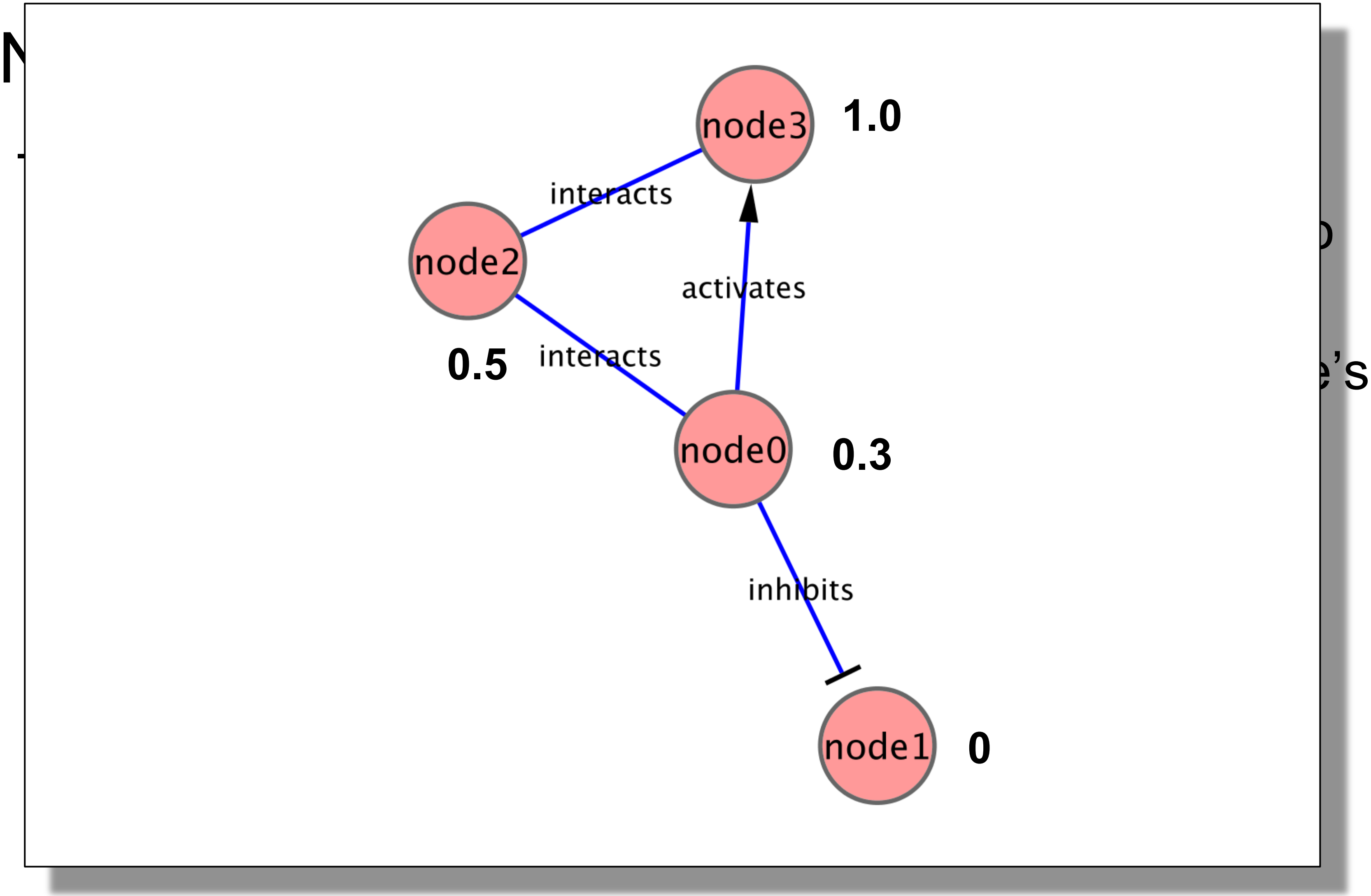
    - Flat ra ogeneous,
      nodes bust to
      arbitra
    - Scale-
    - Small- ogatz
  - Useful t a random
    network

Analytical Approaches

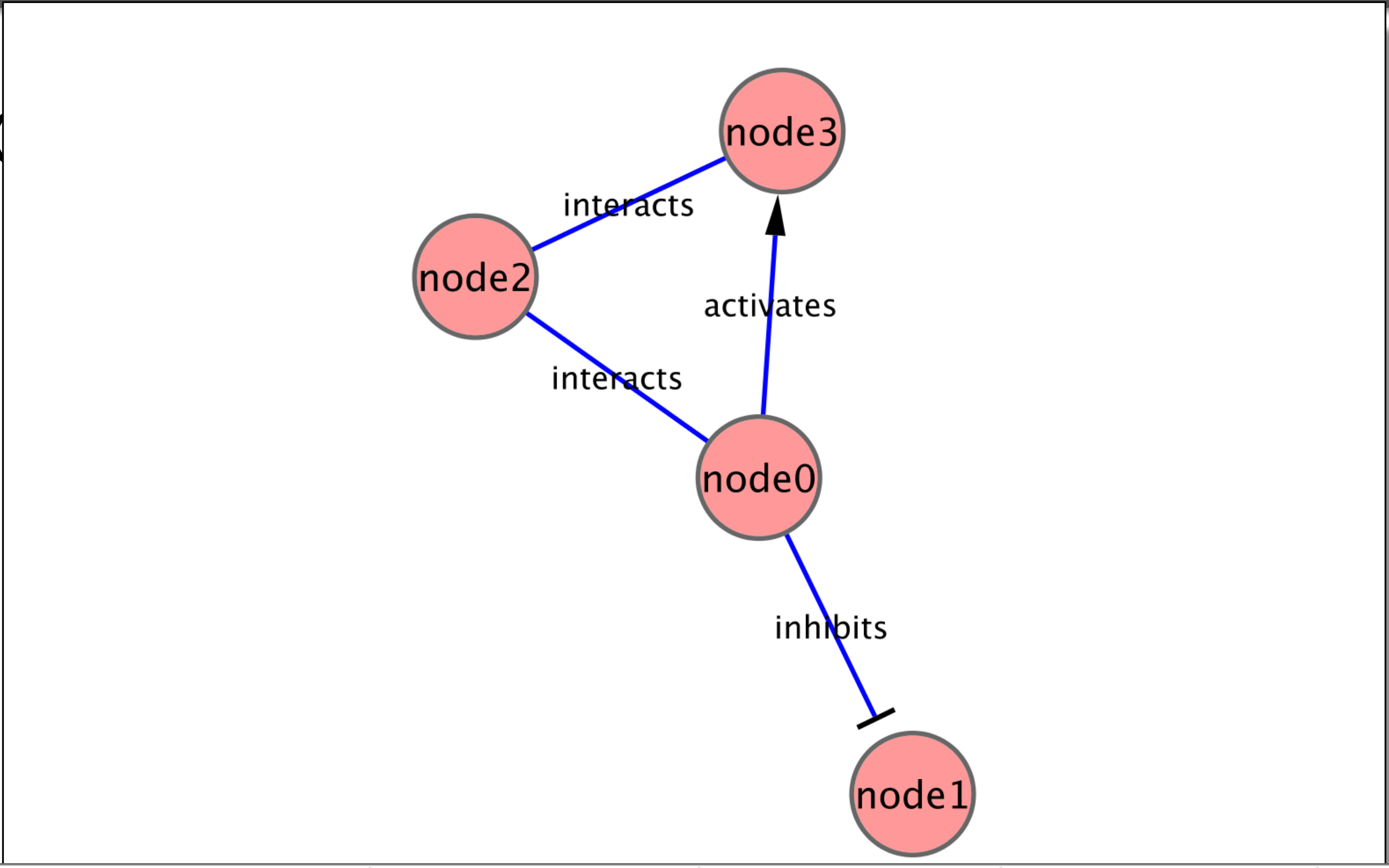
- N
  -

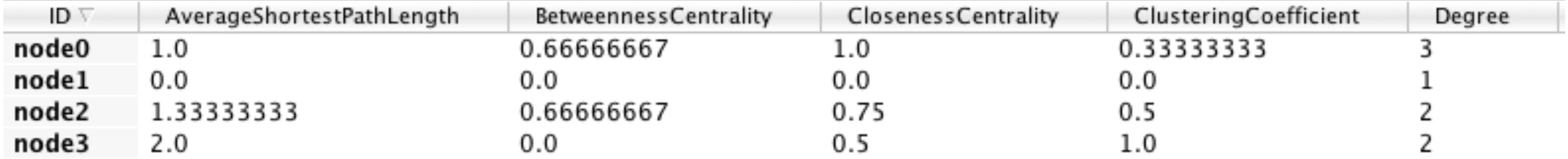
o

e’s



Analytical Approaches





ID ▽	AverageShortestPathLength	BetweennessCentrality	ClosenessCentrality	ClusteringCoefficient	Degree
node0	1.0	0.66666667	1.0	0.33333333	3
node1	0.0	0.0	0.0	0.0	1
node2	1.33333333	0.66666667	0.75	0.5	2
node3	2.0	0.0	0.5	1.0	2

Analytical Approaches

- Network Analyzer Demo…

Analytical Approaches

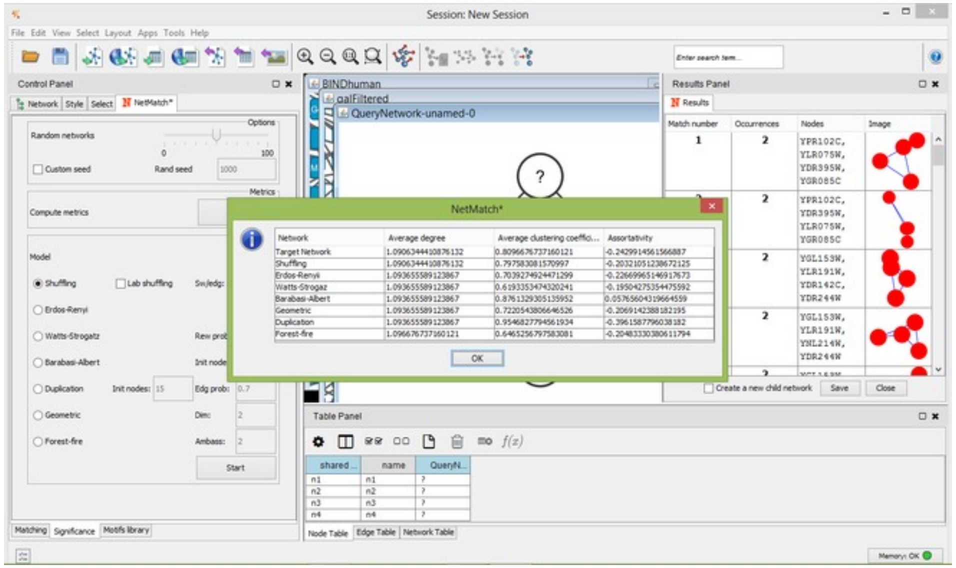
- Guilt by association
  - Combine weak signals to get a stronger one.
  - Two main applications:
    - Suggest network or functional relationship based on related data (e.g. co-expression).
    - Infer function or role (i.e. in disease) based on related data
  - Some algorithmic approaches
    - Random walk with restarts
    - Semi-supervised learning
    - Belief Propagation
  - Example: GeneMANIA

Analytical Approaches

- GeneMANIA Demo…

Analytical Approaches

- Motif finding
  - Search directed networks for network motifs (feed-forward loops, feedback loops, etc.)



Random networks 0 100

☐ Custom seed Rand seed: 1000

● Shuffling ☐ Lab shuffling

○ Erdos-Renyi

○ Watts-Strogatz

○ Barabasi-Albert

○ Duplication Init nodes: 15 Edg prob: 0.7

○ Geometric Dim: 2

○ Forest-fire Ambass: 2

NetMatch*

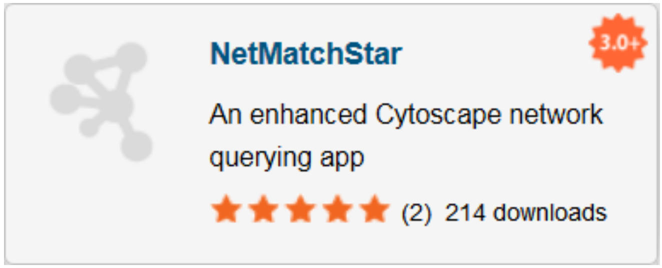
Network	Average degree	Average clustering coeffici...	Assortativity
Target Network	1.0906344410876132	0.8096676737160121	-0.242991456156687
Shuffling	1.0906344410876132	0.797583081570997	-0.2032105123867212 5
Erdos-Renyi	1.093655589123867	0.7039274924471299	-0.2266996514691767 3
Watts-Strogaz	1.093655589123867	0.6193353474320241	-0.1950427535447559 2
Barabasi-Albert	1.093655589123867	0.8761329305135952	0.0576560431966455 9
Geometric	1.093655589123867	0.7220543806646526	-0.2069142388182195
Duplication	1.093655589123867	0.9546827794561934	-0.396158779603818 2
Forest-fire	1.0966767371601 21	0.6465256797583081	-0.2048333038061179 4

OK

Match number	Occurrences	Nodes	Image
1	2	YPR102C, YLR075W, YDR395W, YGR085C	
	2	YPR102C, YDR395W, YLR075W, YGR085C	
	2	YGL153W, YLR191W, YDR142C, YDR244W	
	2	YGL153W, YLR191W, YNL214W, YDR244W	

☐ Create a new child network

shared ...	name	QueryN...
n1	n1	?
n2	n2	?
n3	n3	?
n4	n4	?



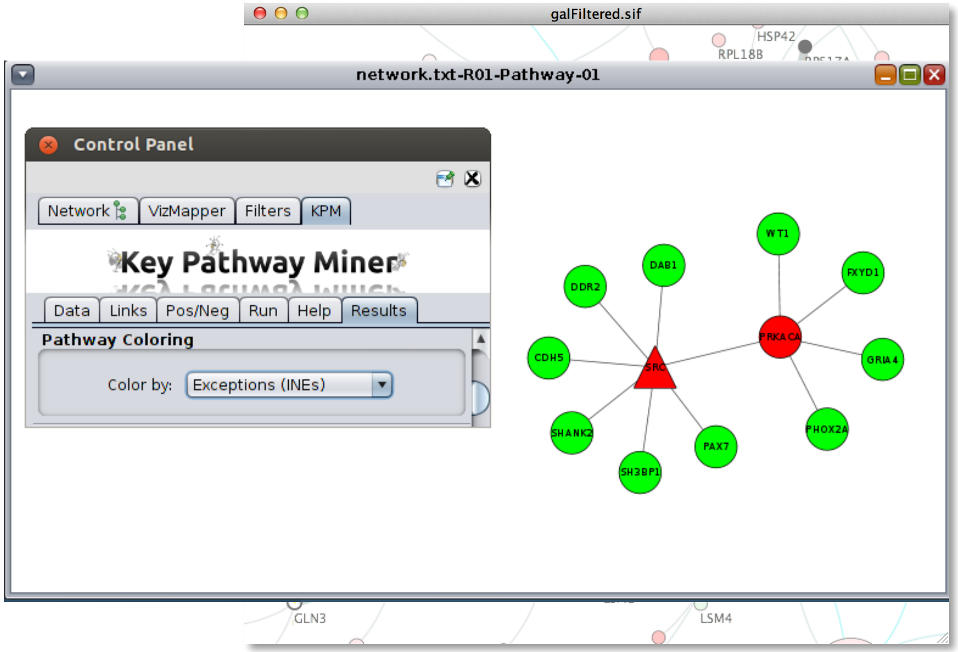
NetMatchStar

An enhanced Cytoscape network querying app

★★★★★ (2) 214 downloads

Analytical Approaches

- Finding subnetworks of similarly expressed genes
- Finding the shortest path between nodes



Pathway Coloring

Color by: Exceptions (INEs)

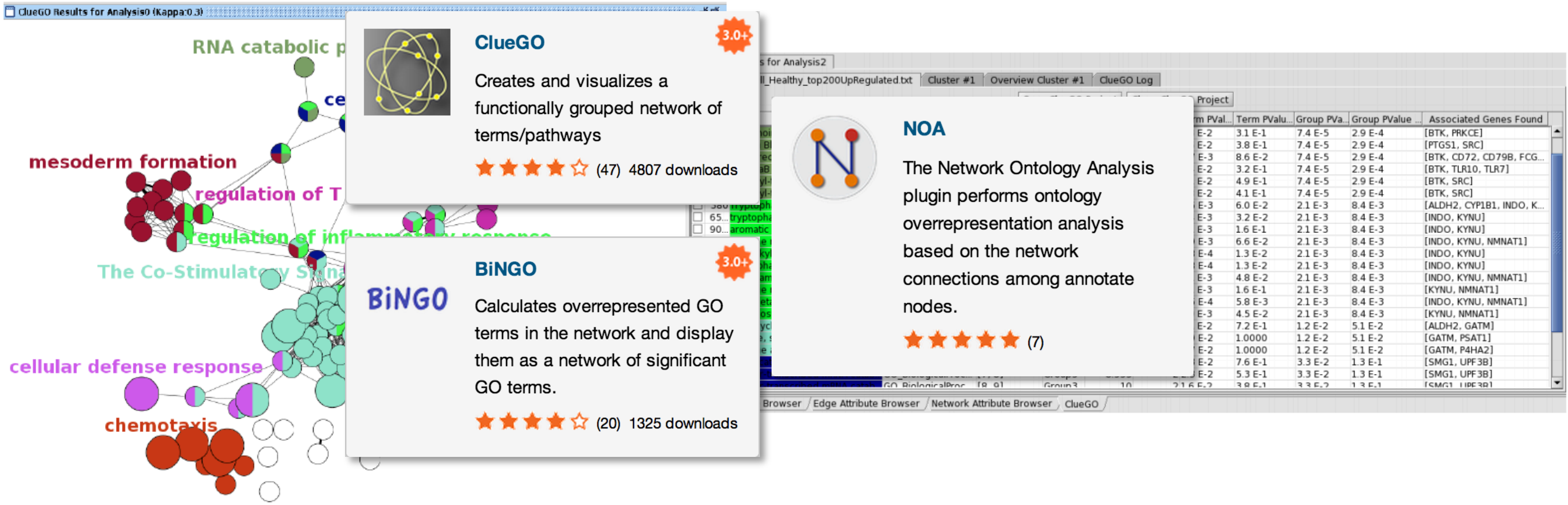
KeyPathwayMiner

Detects highly-connected sub-networks where most genes show similar expression behavior.

★★★★☆ (10) 177 downloads

Analytical Approaches

- Overrepresentation analysis
  - Find terms (GO) that are statistically overrepresented in a network
  - Not really a network analysis technique
  - Very useful for visualization



ClueGO

Creates and visualizes a functionally grouped network of terms/pathways

★★★★☆ (47) 4807 downloads

BiNGO

Calculates overrepresented GO terms in the network and display them as a network of significant GO terms.

★★★★☆ (20) 1325 downloads

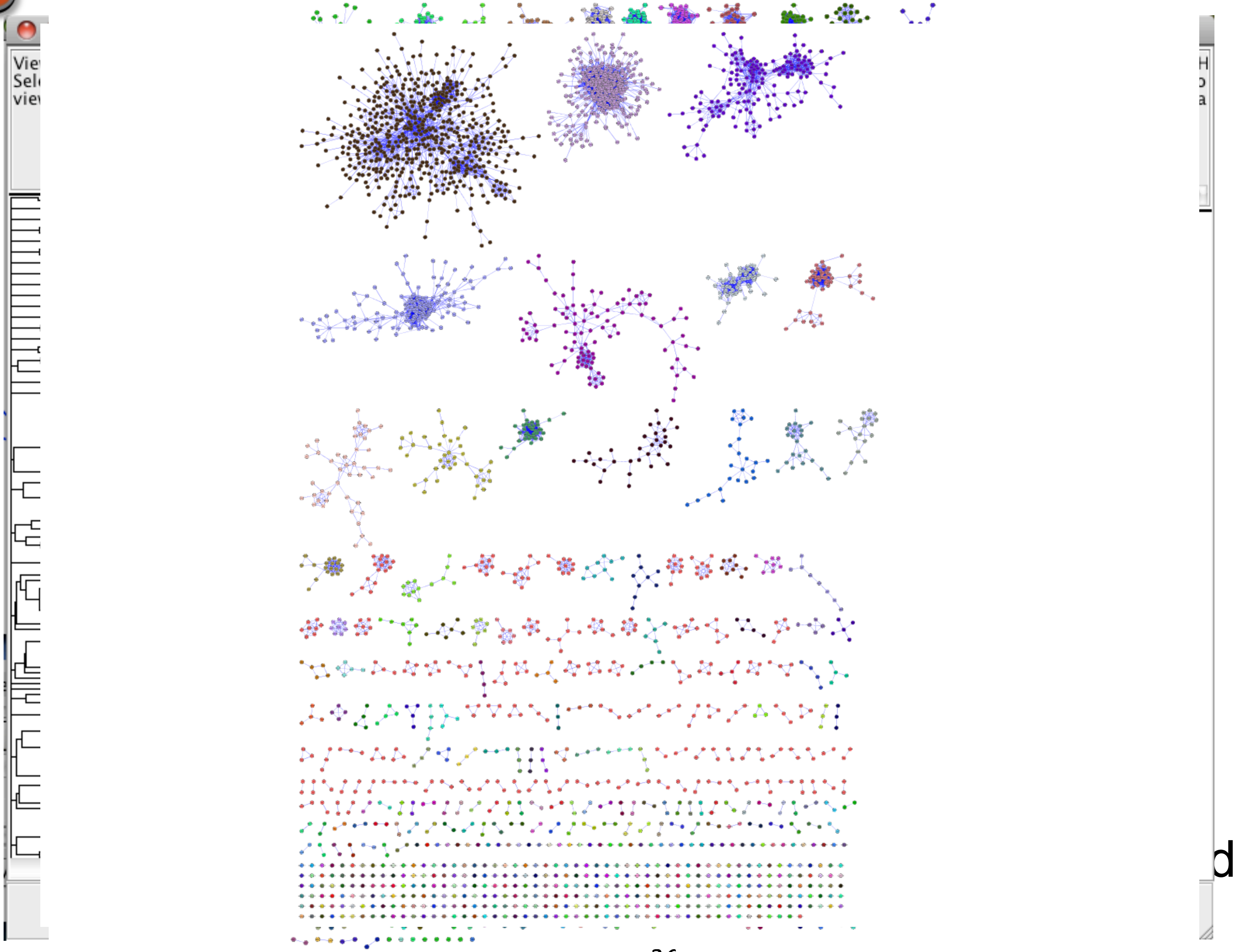
NOA

The Network Ontology Analysis plugin performs ontology overrepresentation analysis based on the network connections among annotate nodes.

★★★★★ (7)

m PVal...	Term PValu...	Group PVa...	Group PValue ...	Associated Genes Found
E-2	3.1 E-1	7.4 E-5	2.9 E-4	[BTK, PRKCE]
E-2	3.8 E-1	7.4 E-5	2.9 E-4	[PTGS1, SRC]
E-3	8.6 E-2	7.4 E-5	2.9 E-4	[BTK, CD72, CD79B, FCG...
E-2	3.2 E-1	7.4 E-5	2.9 E-4	[BTK, TLR10, TLR7]
E-2	4.9 E-1	7.4 E-5	2.9 E-4	[BTK, SRC]
E-2	4.1 E-1	7.4 E-5	2.9 E-4	[BTK, SRC]
E-3	6.0 E-2	2.1 E-3	8.4 E-3	[ALDH2, CYP1B1, INDO, K...
E-3	3.2 E-2	2.1 E-3	8.4 E-3	[INDO, KYNU]
E-3	1.6 E-1	2.1 E-3	8.4 E-3	[INDO, KYNU]
E-3	6.6 E-2	2.1 E-3	8.4 E-3	[INDO, KYNU, NMNAT1]
E-4	1.3 E-2	2.1 E-3	8.4 E-3	[INDO, KYNU]
E-4	1.3 E-2	2.1 E-3	8.4 E-3	[INDO, KYNU]
E-3	4.8 E-2	2.1 E-3	8.4 E-3	[INDO, KYNU, NMNAT1]
E-3	1.6 E-1	2.1 E-3	8.4 E-3	[KYNU, NMNAT1]
E-4	5.8 E-3	2.1 E-3	8.4 E-3	[INDO, KYNU, NMNAT1]
E-3	4.5 E-2	2.1 E-3	8.4 E-3	[KYNU, NMNAT1]
E-2	7.2 E-1	1.2 E-2	5.1 E-2	[ALDH2, GATM]
E-2	1.0000	1.2 E-2	5.1 E-2	[GATM, PSAT1]
E-2	1.0000	1.2 E-2	5.1 E-2	[GATM, P4HA2]
E-2	7.6 E-1	3.3 E-2	1.3 E-1	[SMG1, UPF3B]
E-2	5.3 E-1	3.3 E-2	1.3 E-1	[SMG1, UPF3B]

Analytical Approaches

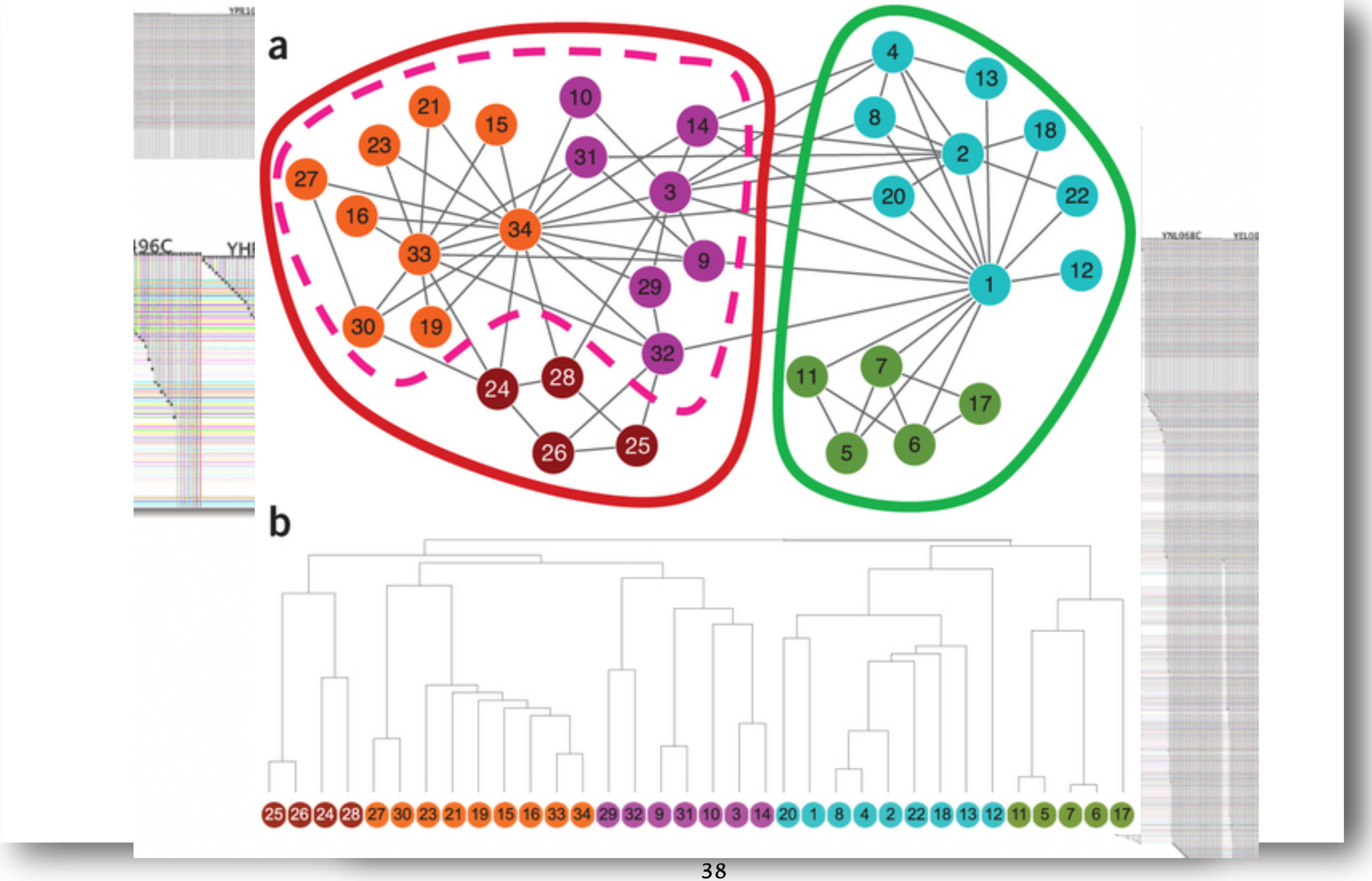


dges

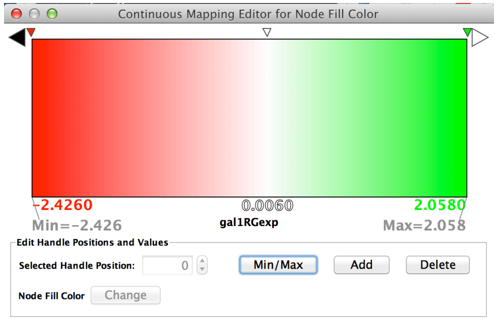
Visualization of Biological Networks

- Data Mapping
- Layouts
- Animation

Depiction


a
b
25 26 24 28 27 30 23 21 19 15 16 33 34 29 32 9 31 10 3 14 20 1 8 4 2 22 18 13 12 11 5 7 6 17

Data Mapping

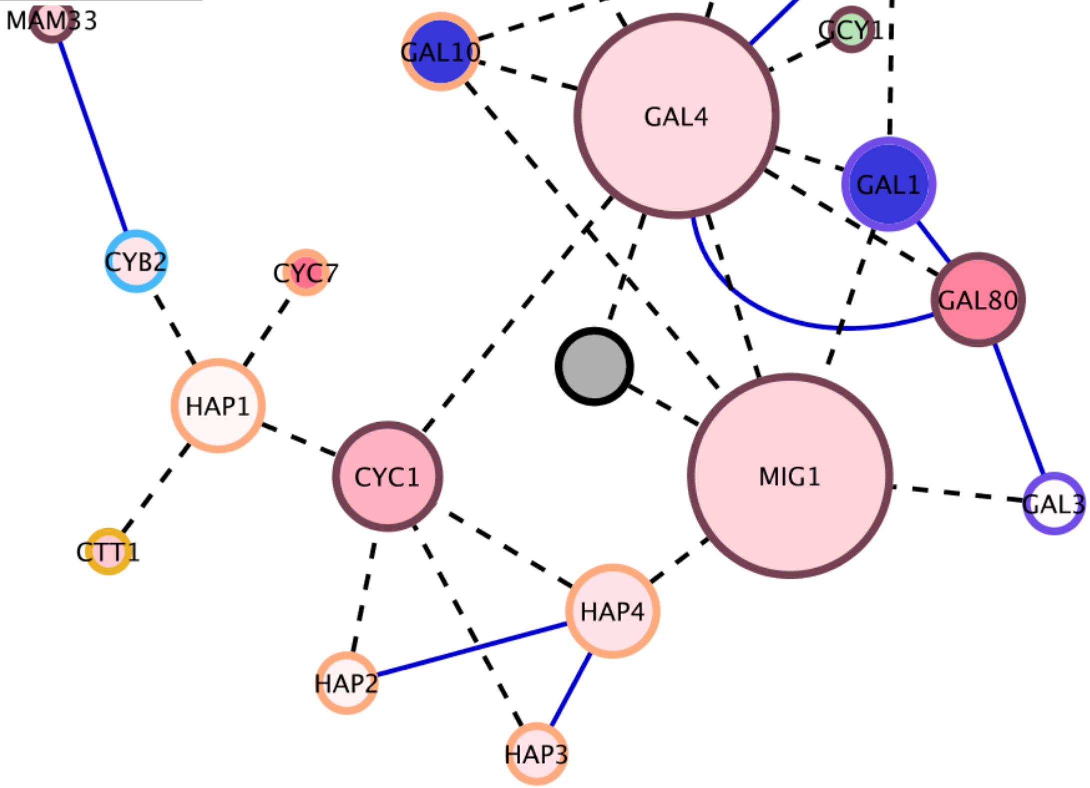

Continuous Mapping Editor for Node Fill Color

-2.4260
0.0060
2.0580
Min=-2.426
gal1RGexp
Max=2.058

Edit Handle Positions and Values

Selected Handle Position: 0

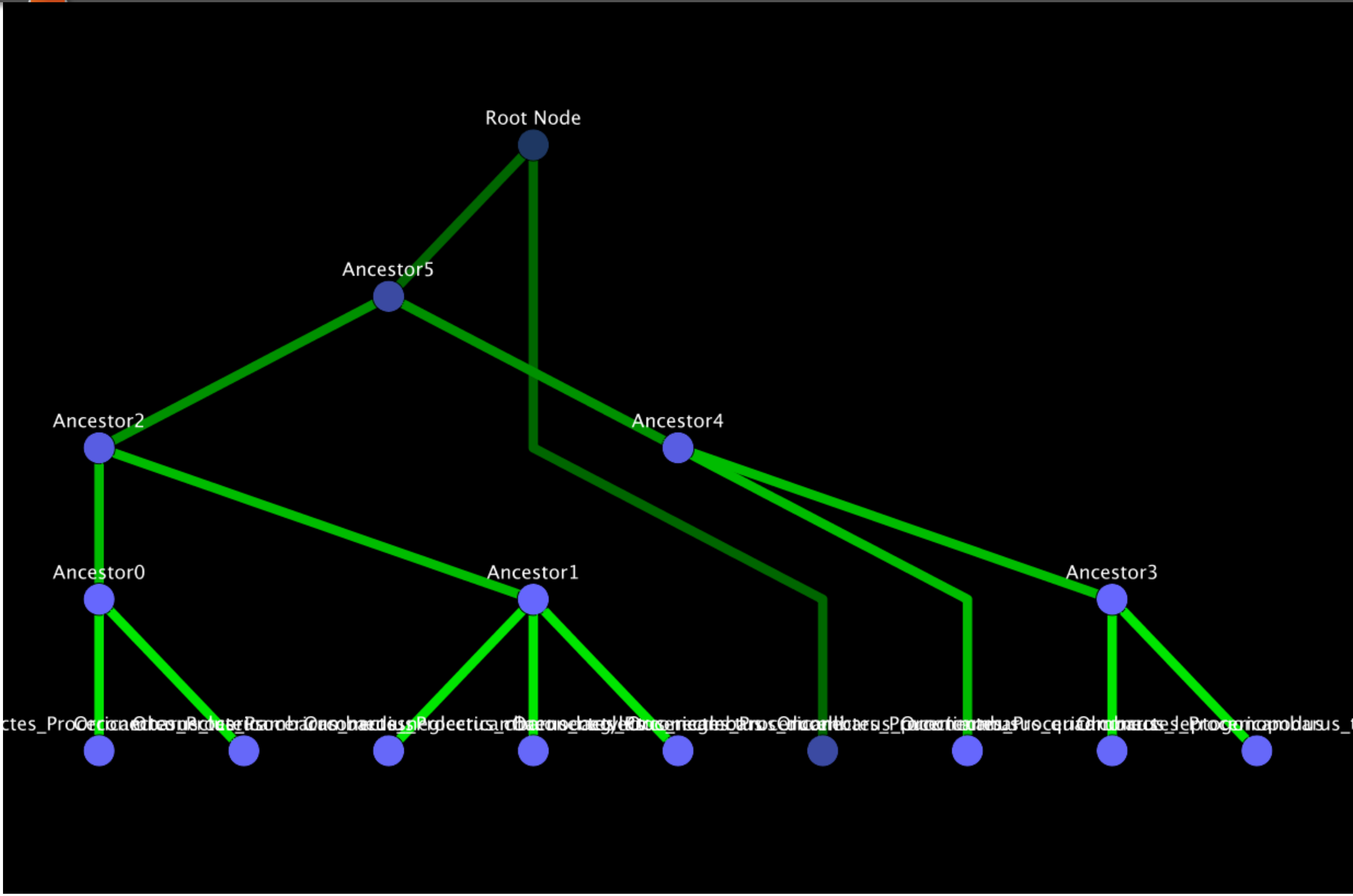
Node Fill Color



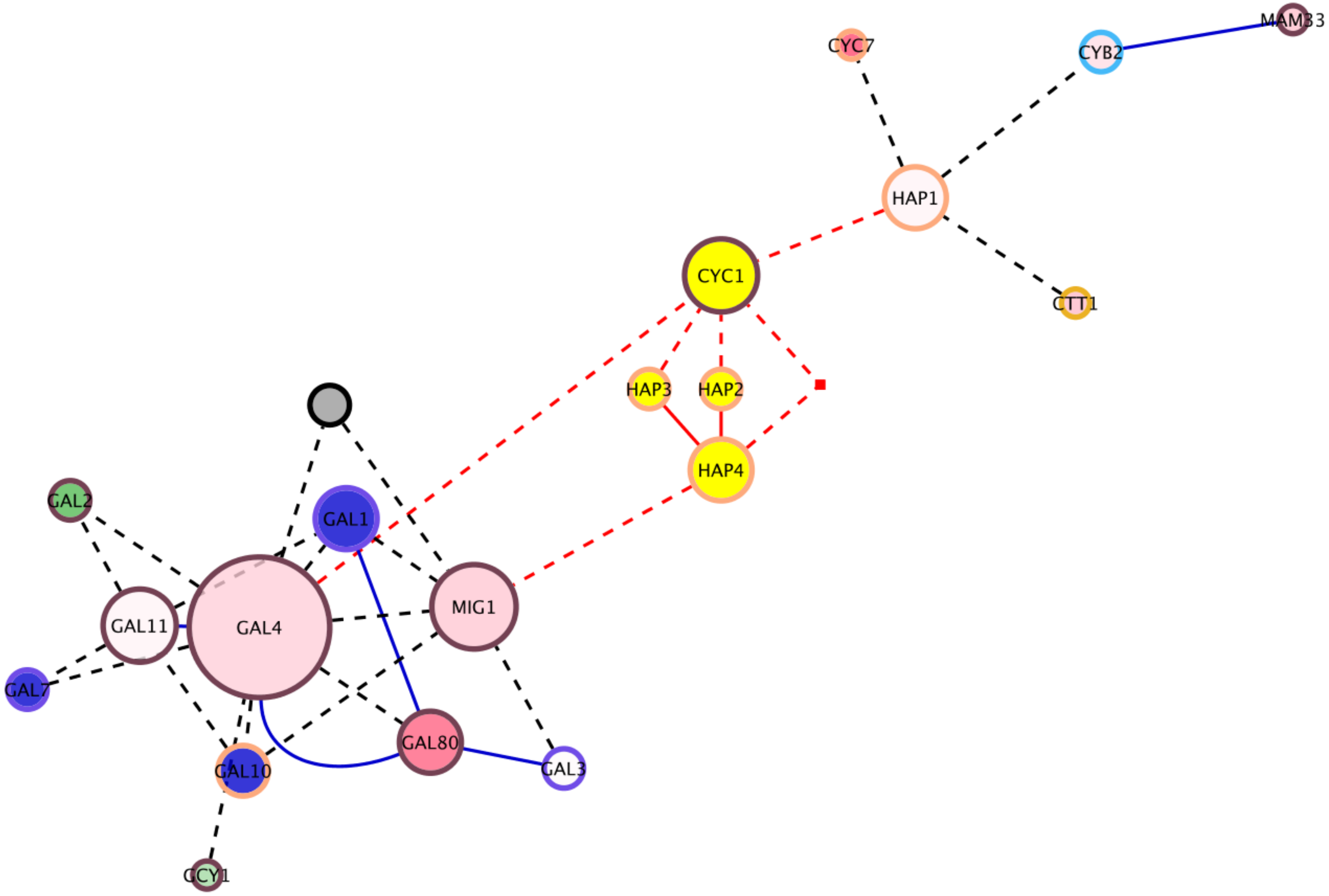
Data Mapping

- Avoid cluttering your visualization with too much data
  - Highlight meaningful differences
  - Avoid confusing the viewer
  - Consider creating multiple network images

Layouts


Root Node
Ancestor5
Ancestor2
Ancestor4
Ancestor0
Ancestor1
Ancestor3

Layouts

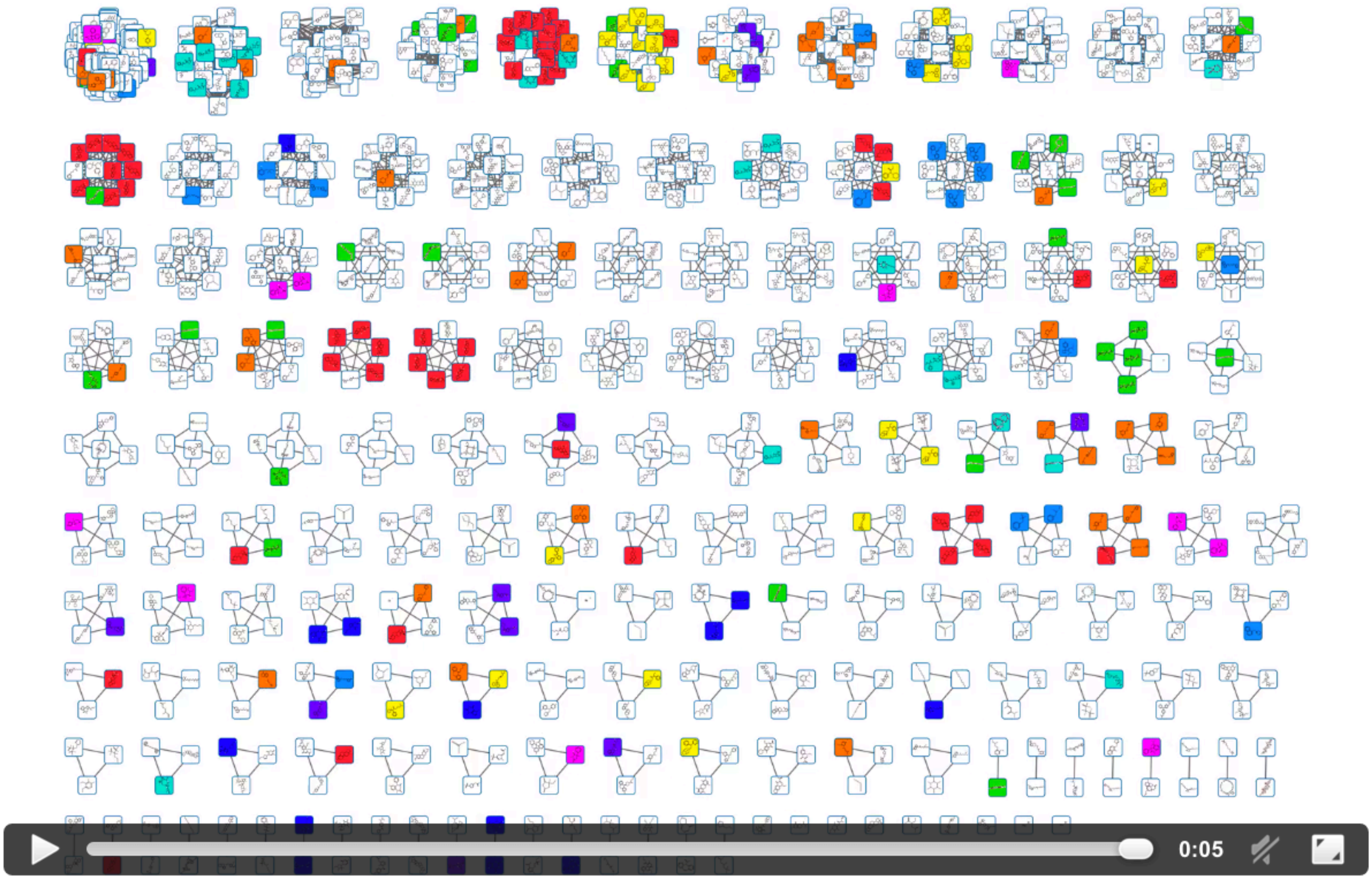


– Many, many others

Layouts

- Use layouts to convey the relationships between the nodes.
- Layout algorithms may need to be "tuned" to fit your network.
- There is not one *correct* layout. Try different things.

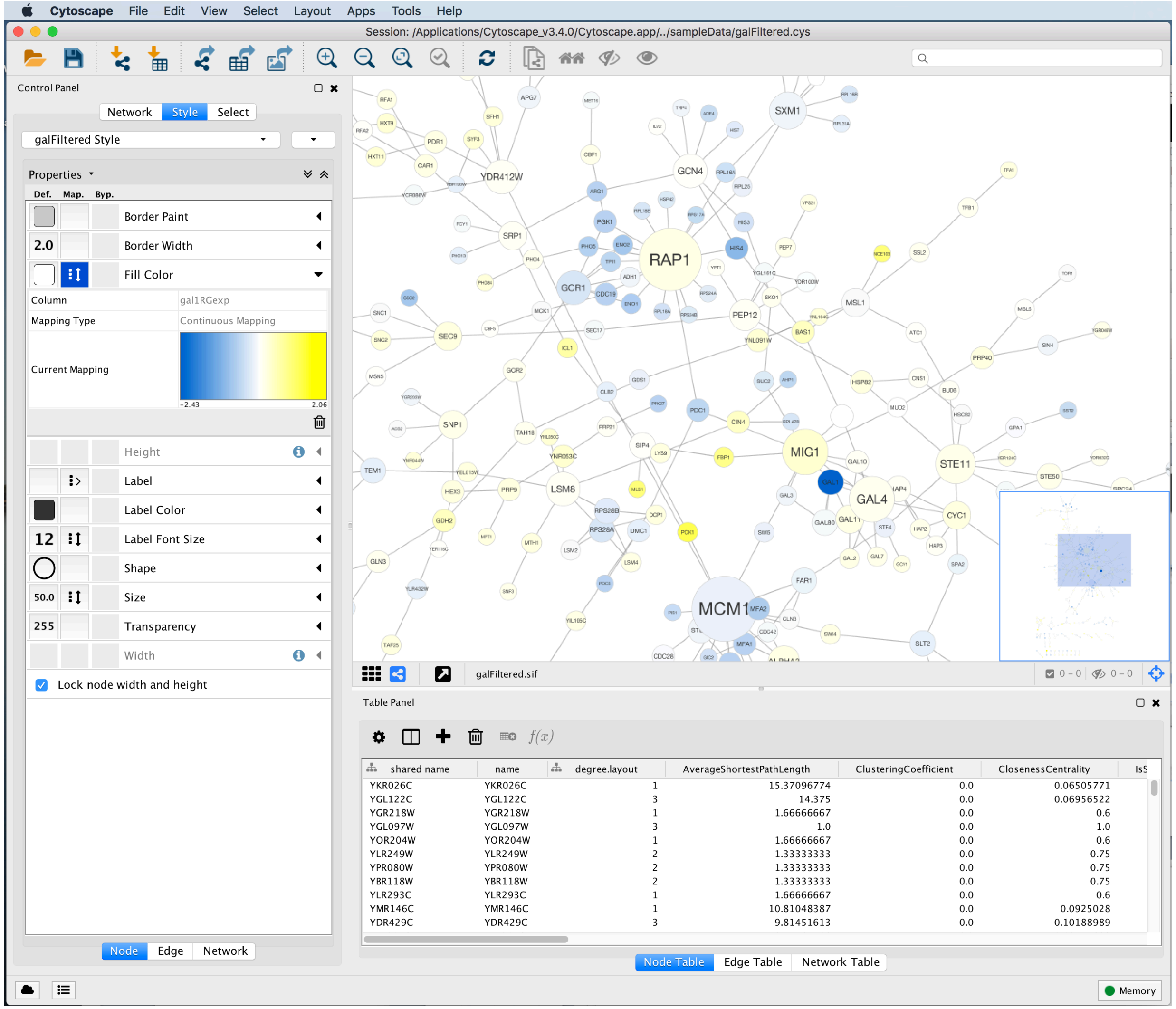
Animation



Introduction to Cytoscape

- Overview
- Core Concepts
  - Networks and Tables
  - Visual Properties
  - Cytoscape Apps
- Working with Data
  - Loading networks from files and online databases
  - Loading data tables from CSV or Excel files
  - The Table Panel

Cytoscape



Session: /Applications/Cytoscape_v3.4.0/Cytoscape.app/../sampleData/galFiltered.cys

galFiltered Style

Def.	Property
	Border Paint
2.0	Border Width
	Fill Color
	Height
	Label
	Label Color
12	Label Font Size
○	Shape
50.0	Size
255	Transparency
	Width

Column: gal1RGexp
Mapping Type: Continuous Mapping
Current Mapping: -2.43 … 2.06

☑ Lock node width and height

shared name	name	degree.layout	AverageShortestPathLength	ClusteringCoefficient	ClosenessCentrality	IsS
YKR026C	YKR026C	1	15.37096774	0.0	0.06505771	
YGL122C	YGL122C	3	14.375	0.0	0.06956522	
YGR218W	YGR218W	1	1.66666667	0.0	0.6	
YGL097W	YGL097W	3	1.0	0.0	1.0	
YOR204W	YOR204W	1	1.66666667	0.0	0.6	
YLR249W	YLR249W	2	1.33333333	0.0	0.75	
YPR080W	YPR080W	2	1.33333333	0.0	0.75	
YBR118W	YBR118W	2	1.33333333	0.0	0.75	
YLR293C	YLR293C	1	1.66666667	0.0	0.6	
YMR146C	YMR146C	1	10.81048387	0.0	0.0925028	
YDR429C	YDR429C	3	9.81451613	0.0	0.10188989	

- Open source
- Cross platform
- Consortium

Institute for Systems Biology

University of California at San Diego

Memorial Sloan-Kettering Cancer Center

Institut Pasteur

Agilent Technologies

University of California at San Francisco

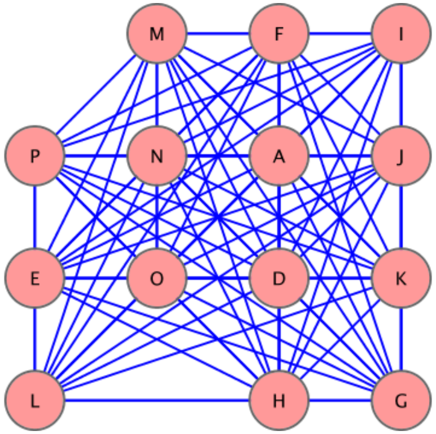
University of Toronto

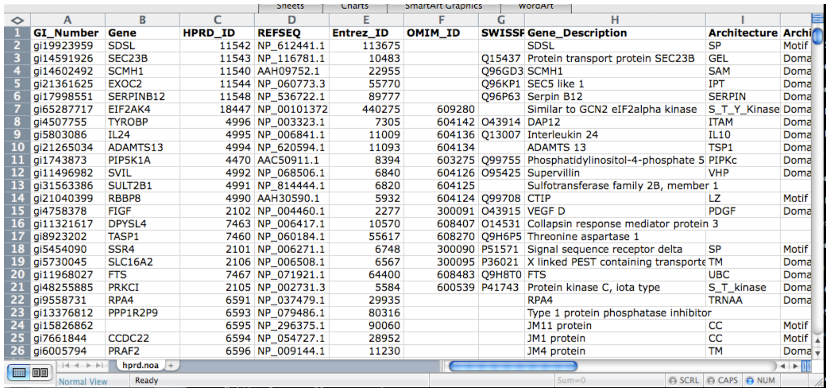
GLADSTONE INSTITUTES

NRNB

Core Concepts

- Networks and Tables





	A	B	C	D	E	F	G	H	I	
1	GI_Number	Gene	HPRD_ID	REFSEQ	Entrez_ID	OMIM_ID	SWISSF	Gene_Description	Architecture	Archi
2	gi19923959	SDSL	11542	NP_612441.1	113675			SDSL	SP	Motif
3	gi14591926	SEC23B	11543	NP_116781.1	10483		Q15437	Protein transport protein SEC23B	GEL	Doma
4	gi14602492	SCMH1	11540	AAH09752.1	22955		Q96GD3	SCMH1	SAM	Doma
5	gi21361625	EXOC2	11544	NP_060773.3	55770		Q96KP1	SEC5 like 1	IPT	Doma
6	gi17998551	SERPINB12	11548	NP_536722.1	89777		Q96P63	Serpin B12	SERPIN	Doma
7	gi65287717	EIF2AK4	18447	NP_00101372	440275	609280		Similar to GCN2 eIF2alpha kinase	S_T_Y_Kinase	Doma
8	gi4507755	TYROBP	4996	NP_003323.1	7305	604142	O43914	DAP12	ITAM	Doma
9	gi5803086	IL24	4995	NP_006841.1	11009	604136	Q13007	Interleukin 24	IL10	Doma
10	gi21265034	ADAMTS13	4994	NP_620594.1	11093	604134		ADAMTS 13	TSP1	Doma
11	gi1743873	PIP5K1A	4470	AAC50911.1	8394	603275	Q99755	Phosphatidylinositol-4-phosphate 5	PIPKc	Doma
12	gi11496982	SVIL	4992	NP_068506.1	6840	604126	O95425	Supervillin	VHP	Doma
13	gi31563386	SULT2B1	4991	NP_814444.1	6820	604125		Sulfotransferase family 2B, member 1		
14	gi21040399	RBBP8	4990	AAH30590.1	5932	604124	Q99708	CTIP	LZ	Motif
15	gi4758378	FIGF	2102	NP_004460.1	2277	300091	O43915	VEGF D	PDGF	Doma
16	gi11321617	DPYSL4	7463	NP_006417.1	10570	608407	O14531	Collapsin response mediator protein 3		
17	gi8923202	TASP1	7460	NP_060184.1	55617	608270	Q9H6P5	Threonine aspartase 1		
18	gi5454090	SSR4	2101	NP_006271.1	6748	300090	P51571	Signal sequence receptor delta	SP	Motif
19	gi5730045	SLC16A2	2106	NP_006508.1	6567	300095	P36021	X linked PEST containing transporte	TM	Doma
20	gi11968027	FTS	7467	NP_071921.1	64400	608483	Q9H8T0	FTS	UBC	Doma
21	gi48255885	PRKCI	2105	NP_002731.3	5584	600539	P41743	Protein kinase C, iota type	S_T_kinase	Doma
22	gi9558731	RPA4	6591	NP_037479.1	29935			RPA4	TRNAA	Doma
23	gi13376812	PPP1R2P9	6593	NP_079486.1	80316			Type 1 protein phosphatase inhibitor		
24	gi15826862		6595	NP_296375.1	90060			JM11 protein	CC	Motif
25	gi7661844	CCDC22	6594	NP_054727.1	28952			JM1 protein	CC	Motif
26	gi6005794	PRAF2	6596	NP_009144.1	11230			JM4 protein	TM	Doma

Networks

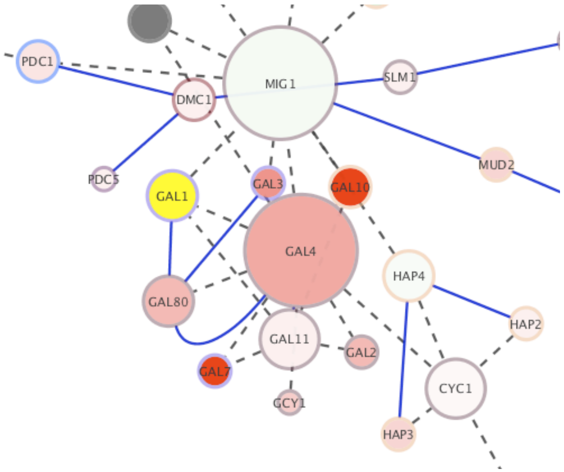
e.g., PPIs or pathways

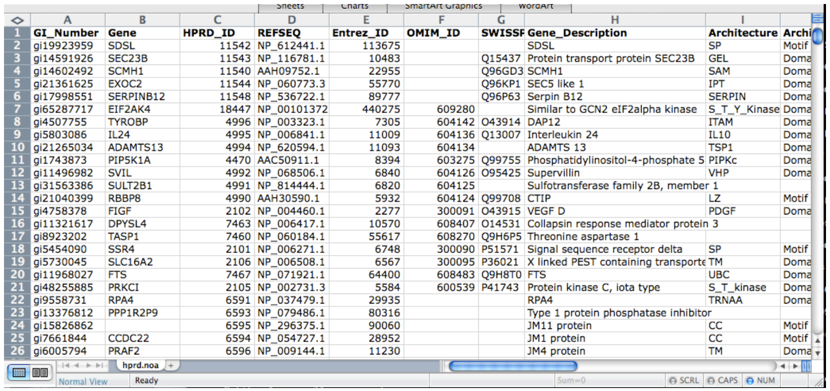
Tables

e.g., data or annotations

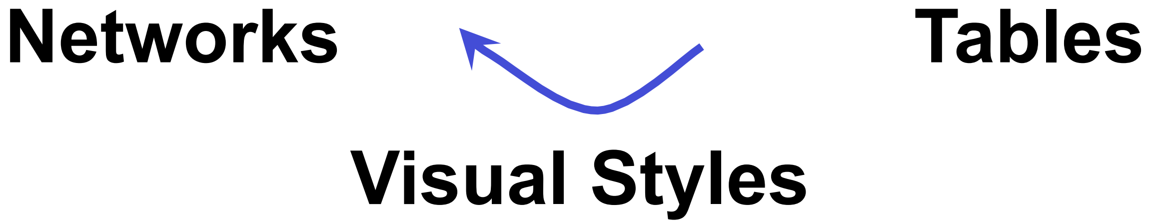
Core Concepts

- Networks and Tables





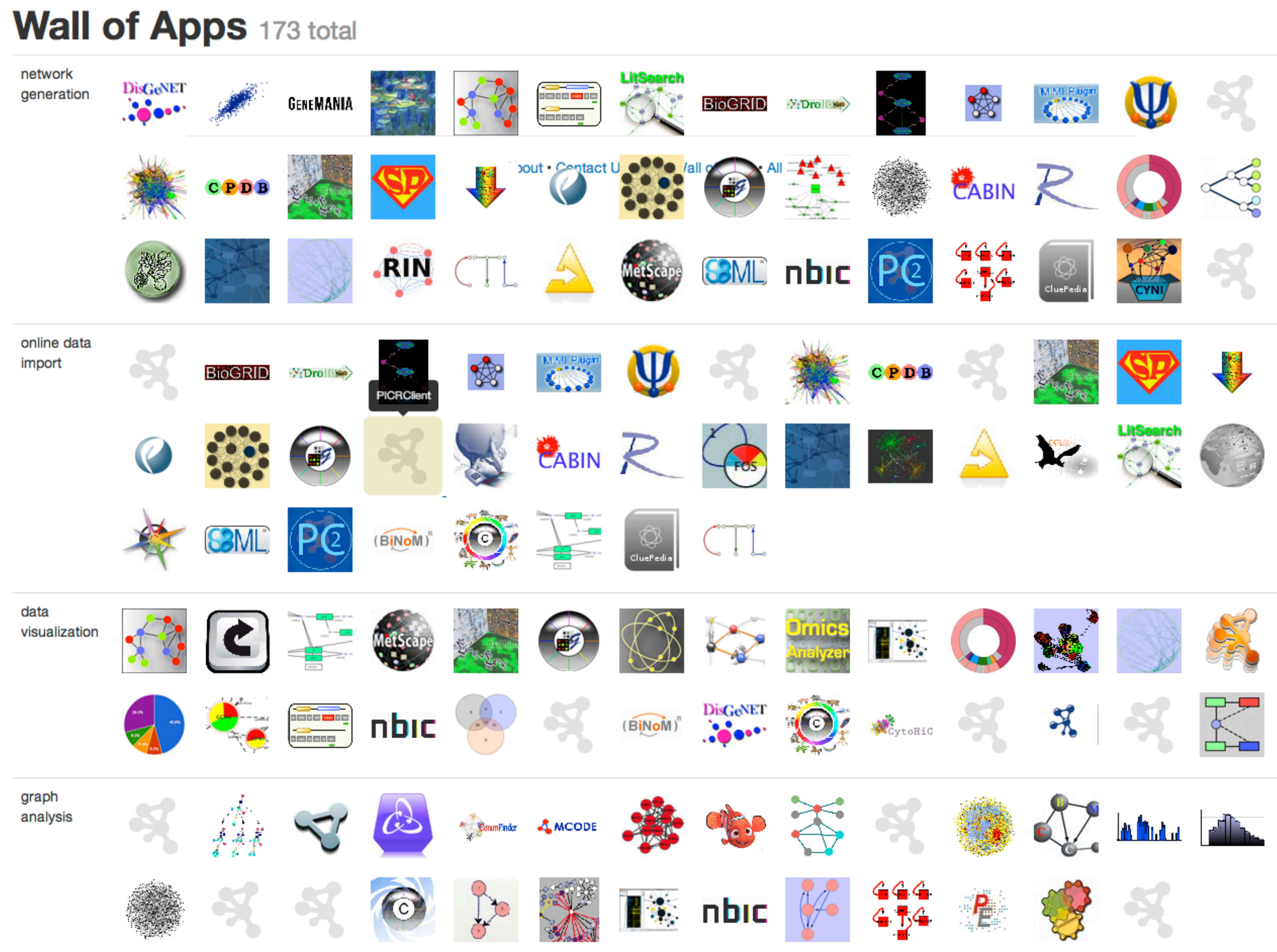
	A	B	C	D	E	F	G	H	I	
1	**GI_Number**	**Gene**	**HPRD_ID**	**REFSEQ**	**Entrez_ID**	**OMIM_ID**	**SWISSF**	**Gene_Description**	**Architecture**	**Archi**
2	gi19923959	SDSL	11542	NP_612441.1	113675			SDSL	SP	Motif
3	gi14591926	SEC23B	11543	NP_116781.1	10483		Q15437	Protein transport protein SEC23B	GEL	Doma
4	gi14602492	SCMH1	11540	AAH09752.1	22955		Q96GD3	SCMH1	SAM	Doma
5	gi21361625	EXOC2	11544	NP_060773.3	55770		Q96KP1	SEC5 like 1	IPT	Doma
6	gi17998551	SERPINB12	11548	NP_536722.1	89777		Q96P63	Serpin B12	SERPIN	Doma
7	gi65287717	EIF2AK4	18447	NP_00101372	440275	609280		Similar to GCN2 eIF2alpha kinase	S_T_Y_Kinase	Doma
8	gi4507755	TYROBP	4996	NP_003323.1	7305	604142	O43914	DAP12	ITAM	Doma
9	gi5803086	IL24	4995	NP_006841.1	11009	604136	Q13007	Interleukin 24	IL10	Doma
10	gi21265034	ADAMTS13	4994	NP_620594.1	11093	604134		ADAMTS 13	TSP1	Doma
11	gi1743873	PIP5K1A	4470	AAC50911.1	8394	603275	Q99755	Phosphatidylinositol-4-phosphate 5	PIPKc	Doma
12	gi11496982	SVIL	4992	NP_068506.1	6840	604126	O95425	Supervillin	VHP	Doma
13	gi31563386	SULT2B1	4991	NP_814444.1	6820	604125		Sulfotransferase family 2B, member 1		
14	gi21040399	RBBP8	4990	AAH30590.1	5932	604124	Q99708	CTIP	LZ	Motif
15	gi4758378	FIGF	2102	NP_004460.1	2277	300091	O43915	VEGF D	PDGF	Doma
16	gi11321617	DPYSL4	7463	NP_006417.1	10570	608407	O14531	Collapsin response mediator protein 3		
17	gi8923202	TASP1	7460	NP_060184.1	55617	608270	Q9H6P5	Threonine aspartase 1		
18	gi5454090	SSR4	2101	NP_006271.1	6748	300090	P51571	Signal sequence receptor delta	SP	Motif
19	gi5730045	SLC16A2	2106	NP_006508.1	6567	300095	P36021	X linked PEST containing transporte	TM	Doma
20	gi11968027	FTS	7467	NP_071921.1	64400	608483	Q9H8T0	FTS	UBC	Doma
21	gi48255885	PRKCI	2105	NP_002731.3	5584	600539	P41743	Protein kinase C, iota type	S_T_kinase	Doma
22	gi9558731	RPA4	6591	NP_037479.1	29935			RPA4	TRNAA	Doma
23	gi13376812	PPP1R2P9	6593	NP_079486.1	80316			Type 1 protein phosphatase inhibitor		
24	gi15826862		6595	NP_296375.1	90060			JM11 protein	CC	Motif
25	gi7661844	CCDC22	6594	NP_054727.1	28952			JM1 protein	CC	Motif
26	gi6005794	PRAF2	6596	NP_009144.1	11230			JM4 protein	TM	Doma



Networks

Tables

Visual Styles



Wall of Apps 173 total

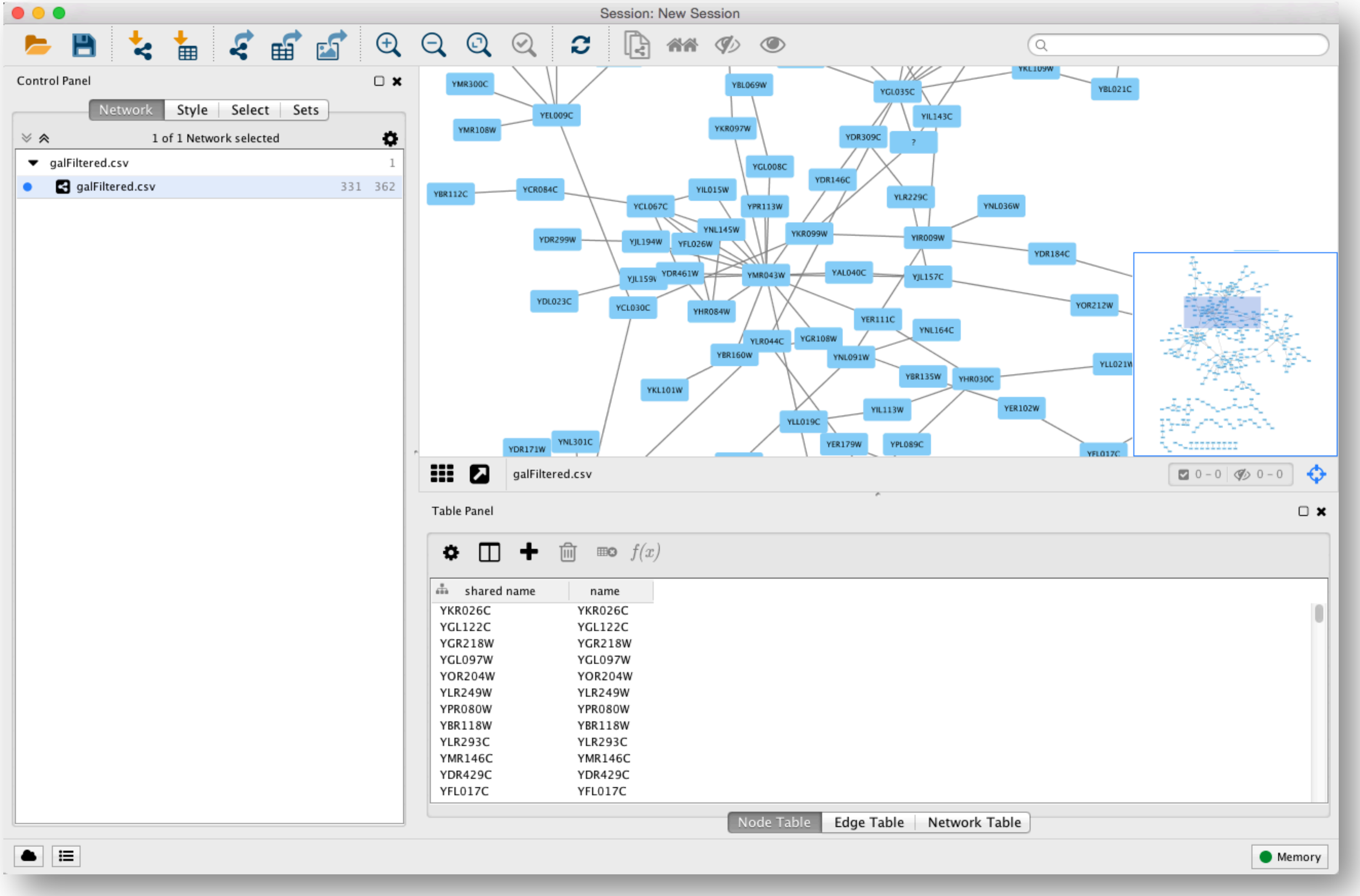
network generation

online data import

data visualization

graph analysis

Loading Networks

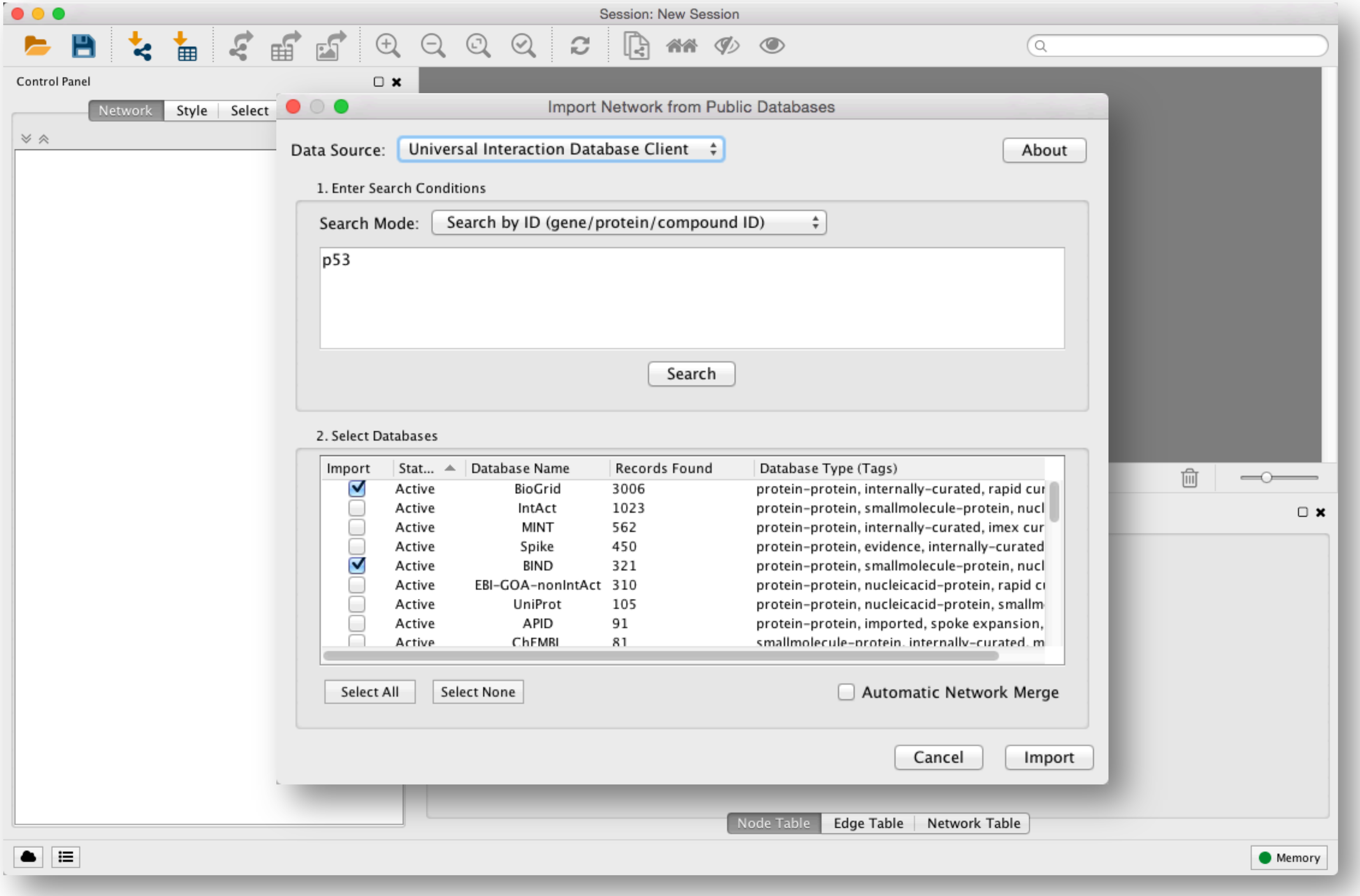


galFiltered.csv		1
galFiltered.csv	331	362

galFiltered.csv

shared name	name
YKR026C	YKR026C
YGL122C	YGL122C
YGR218W	YGR218W
YGL097W	YGL097W
YOR204W	YOR204W
YLR249W	YLR249W
YPR080W	YPR080W
YBR118W	YBR118W
YLR293C	YLR293C
YMR146C	YMR146C
YDR429C	YDR429C
YFL017C	YFL017C

Loading Networks



Import Network from Public Databases

Data Source: Universal Interaction Database Client

About

1. Enter Search Conditions

Search Mode: Search by ID (gene/protein/compound ID)

p53

Search

2. Select Databases

Import	Stat...	Database Name	Records Found	Database Type (Tags)
☑	Active	BioGrid	3006	protein-protein, internally-curated, rapid cur
☐	Active	IntAct	1023	protein-protein, smallmolecule-protein, nucl
☐	Active	MINT	562	protein-protein, internally-curated, imex cur
☐	Active	Spike	450	protein-protein, evidence, internally-curated
☑	Active	BIND	321	protein-protein, smallmolecule-protein, nucl
☐	Active	EBI-GOA-nonIntAct	310	protein-protein, nucleicacid-protein, rapid cu
☐	Active	UniProt	105	protein-protein, nucleicacid-protein, smallm
☐	Active	APID	91	protein-protein, imported, spoke expansion,
☐	Active	ChEMBL	81	smallmolecule-protein, internally-curated, m

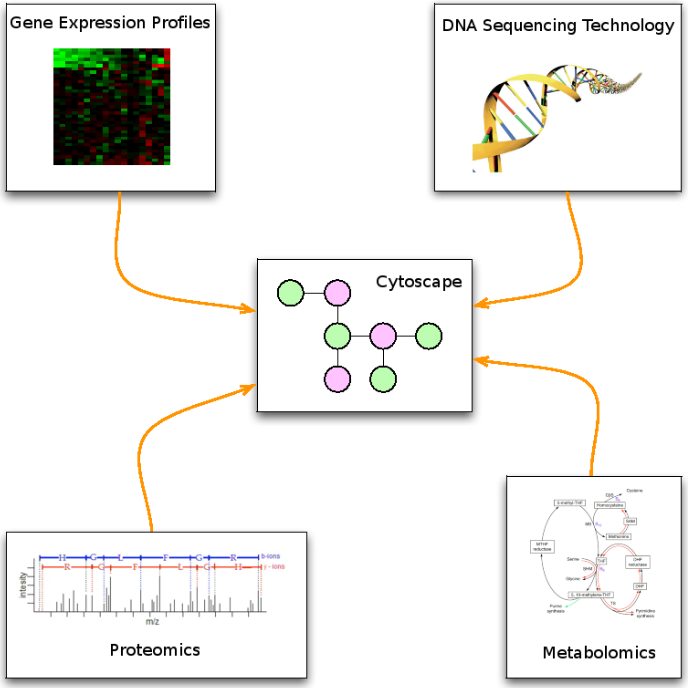
Select All · Select None

☐ Automatic Network Merge

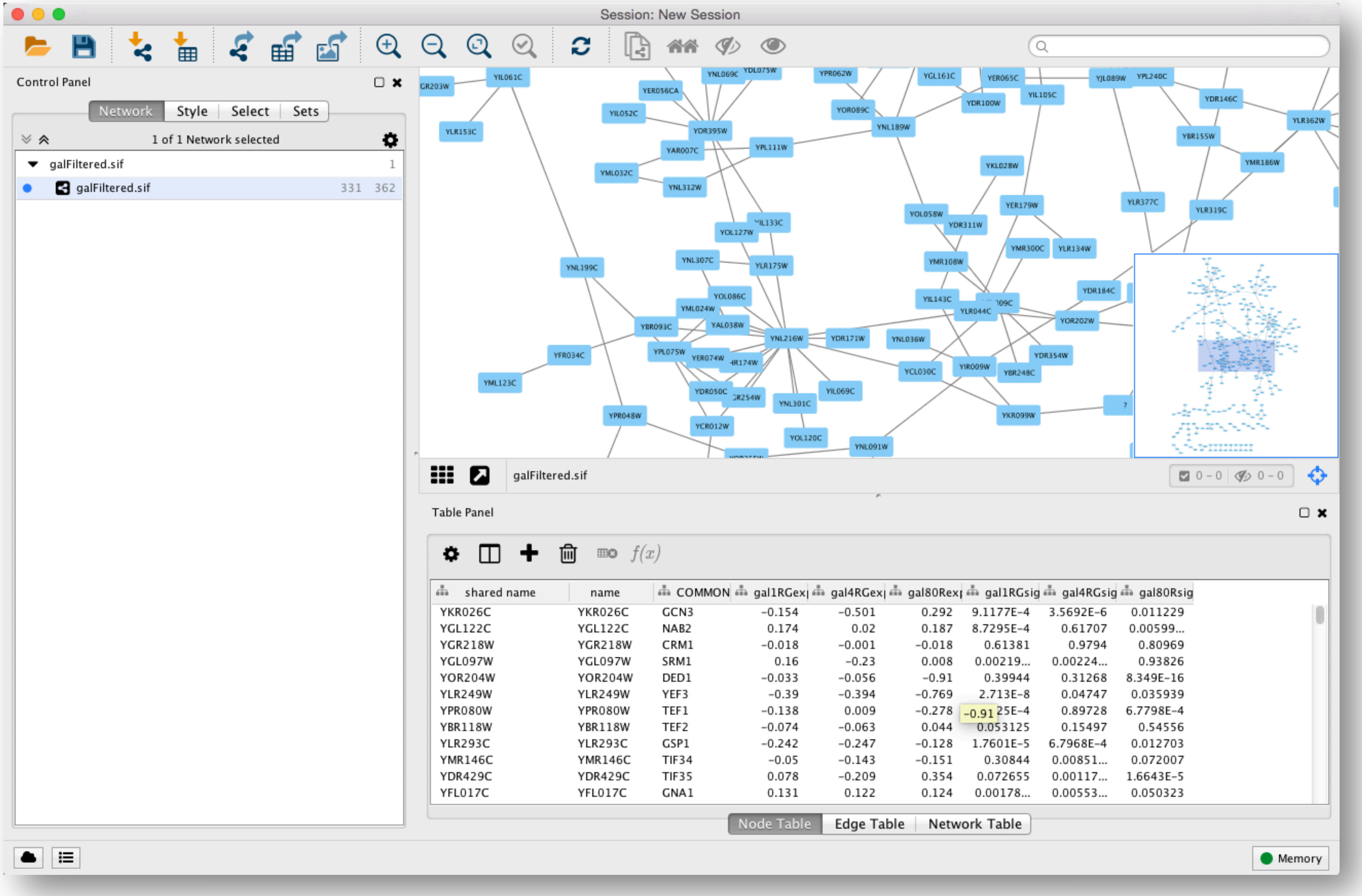
Cancel · Import

Loading Tables

- Nodes and edges can have data associated with them
  - Gene expression data
  - Mass spectrometry data
  - Protein structure information
  - Gene Ontology terms, etc.
- Cytoscape supports multiple data types: Numbers, Text, Boolean, Lists...

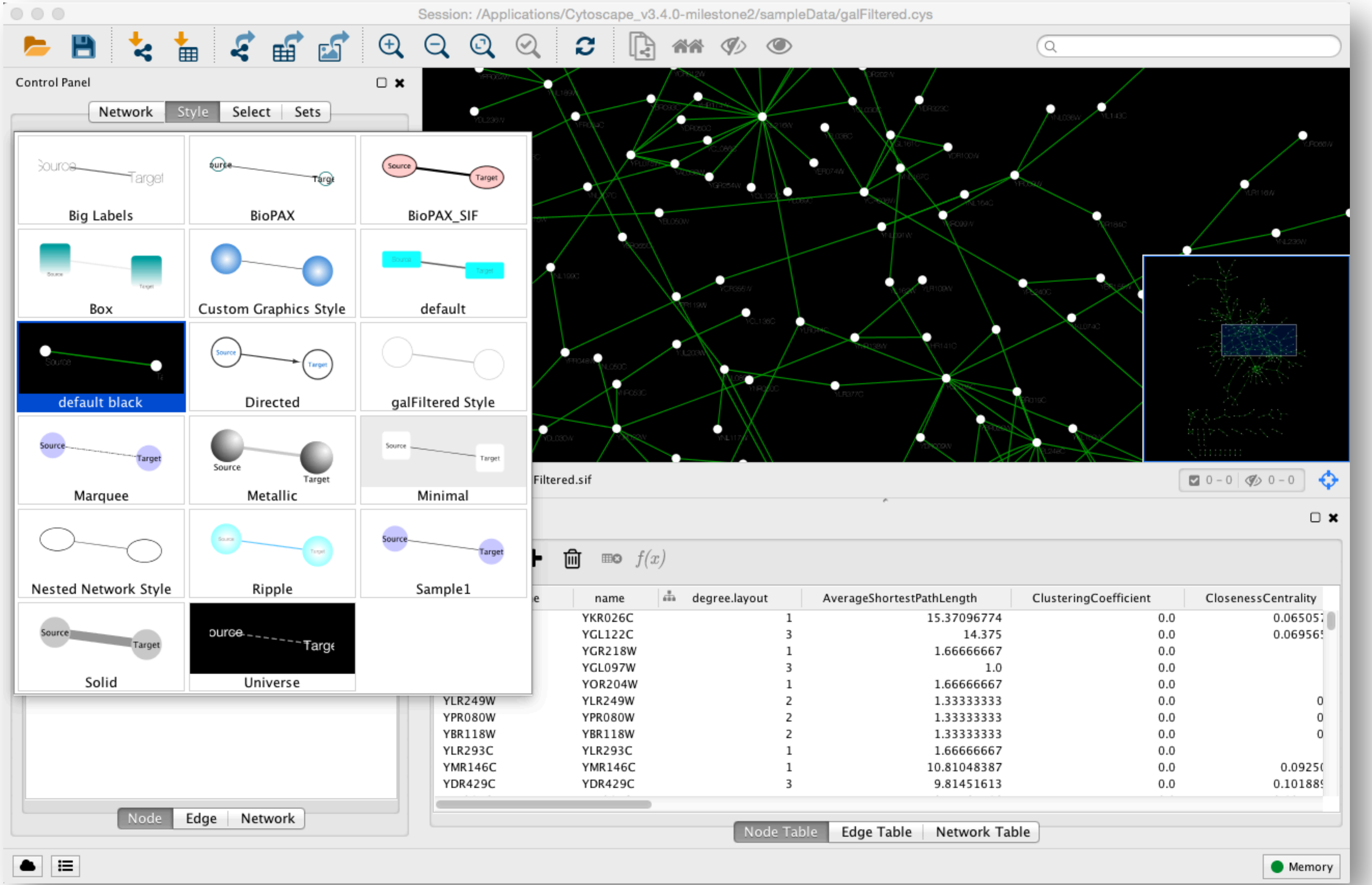


Loading Tables



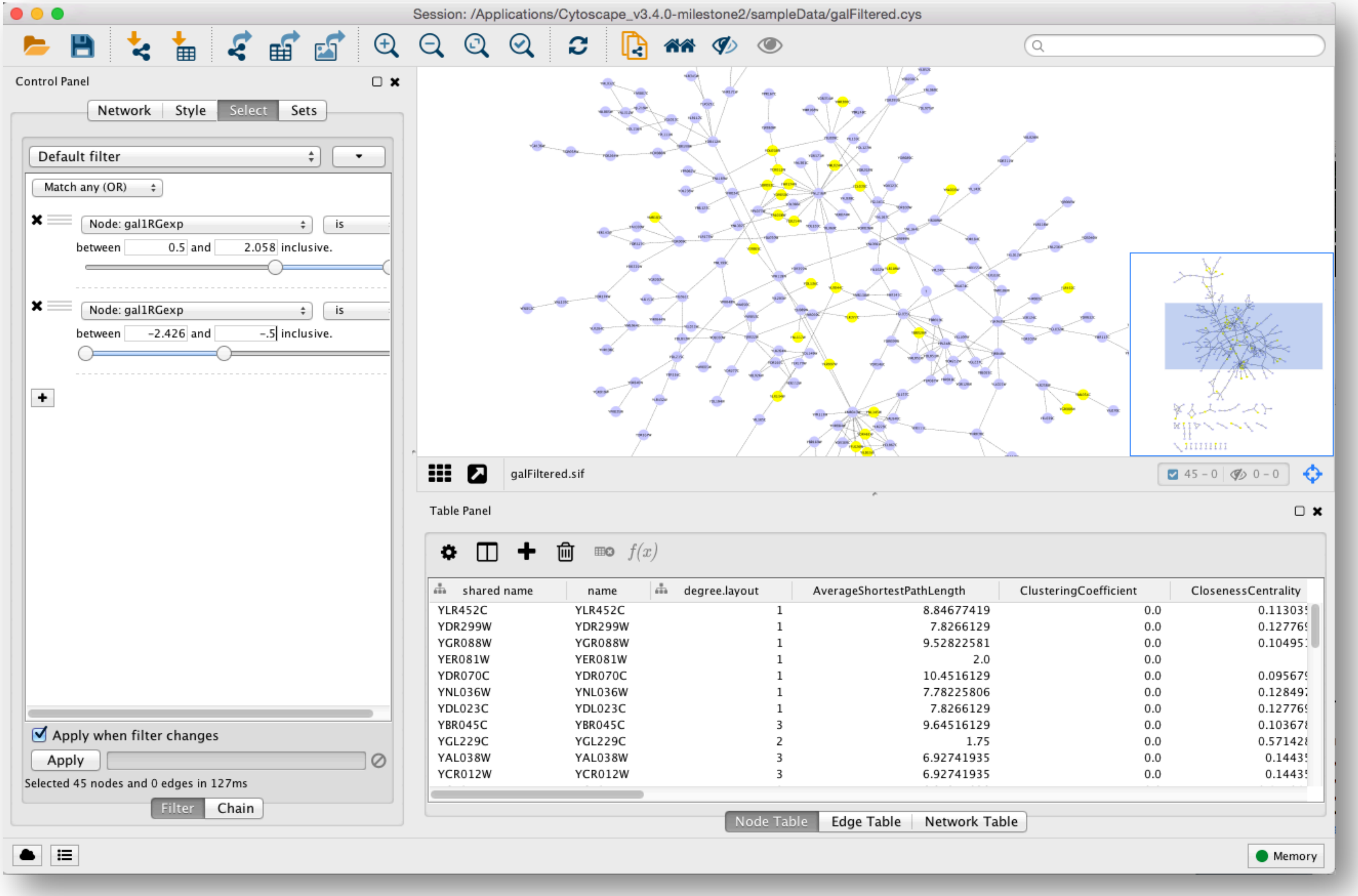
shared name	name	COMMON	gal1RGex	gal4RGex	gal80Rex	gal1RGsig	gal4RGsig	gal80Rsig
YKR026C	YKR026C	GCN3	-0.154	-0.501	0.292	9.1177E-4	3.5692E-6	0.011229
YGL122C	YGL122C	NAB2	0.174	0.02	0.187	8.7295E-4	0.61707	0.00599...
YGR218W	YGR218W	CRM1	-0.018	-0.001	-0.018	0.61381	0.9794	0.80969
YGL097W	YGL097W	SRM1	0.16	-0.23	0.008	0.00219...	0.00224...	0.93826
YOR204W	YOR204W	DED1	-0.033	-0.056	-0.91	0.39944	0.31268	8.349E-16
YLR249W	YLR249W	YEF3	-0.39	-0.394	-0.769	2.713E-8	0.04747	0.035939
YPR080W	YPR080W	TEF1	-0.138	0.009	-0.278	-0.91 25E-4	0.89728	6.7798E-4
YBR118W	YBR118W	TEF2	-0.074	-0.063	0.044	0.053125	0.15497	0.54556
YLR293C	YLR293C	GSP1	-0.242	-0.247	-0.128	1.7601E-5	6.7968E-4	0.012703
YMR146C	YMR146C	TIF34	-0.05	-0.143	-0.151	0.30844	0.00851...	0.072007
YDR429C	YDR429C	TIF35	0.078	-0.209	0.354	0.072655	0.00117...	1.6643E-5
YFL017C	YFL017C	GNA1	0.131	0.122	0.124	0.00178...	0.00553...	0.050323

Visual Style Manager



e	name	degree.layout	AverageShortestPathLength	ClusteringCoefficient	ClosenessCentrality
	YKR026C	1	15.37096774	0.0	0.06505
	YGL122C	3	14.375	0.0	0.06956
	YGR218W	1	1.66666667	0.0	
	YGL097W	3	1.0	0.0	
	YOR204W	1	1.66666667	0.0	
YLR249W	YLR249W	2	1.33333333	0.0	0
YPR080W	YPR080W	2	1.33333333	0.0	0
YBR118W	YBR118W	2	1.33333333	0.0	0
YLR293C	YLR293C	1	1.66666667	0.0	
YMR146C	YMR146C	1	10.81048387	0.0	0.0925
YDR429C	YDR429C	3	9.81451613	0.0	0.10188

Selection Filters



Session: /Applications/Cytoscape_v3.4.0-milestone2/sampleData/galFiltered.cys

Default filter

Match any (OR)

Node: gal1RGexp is between 0.5 and 2.058 inclusive.

Node: gal1RGexp is between -2.426 and -.5 inclusive.

☑ Apply when filter changes

Selected 45 nodes and 0 edges in 127ms

galFiltered.sif

shared name	name	degree.layout	AverageShortestPathLength	ClusteringCoefficient	ClosenessCentrality
YLR452C	YLR452C	1	8.84677419	0.0	0.11303
YDR299W	YDR299W	1	7.8266129	0.0	0.12776
YGR088W	YGR088W	1	9.52822581	0.0	0.10495
YER081W	YER081W	1	2.0	0.0	
YDR070C	YDR070C	1	10.4516129	0.0	0.09567
YNL036W	YNL036W	1	7.78225806	0.0	0.12849
YDL023C	YDL023C	1	7.8266129	0.0	0.12776
YBR045C	YBR045C	3	9.64516129	0.0	0.10367
YGL229C	YGL229C	2	1.75	0.0	0.57142
YAL038W	YAL038W	3	6.92741935	0.0	0.1443
YCR012W	YCR012W	3	6.92741935	0.0	0.1443

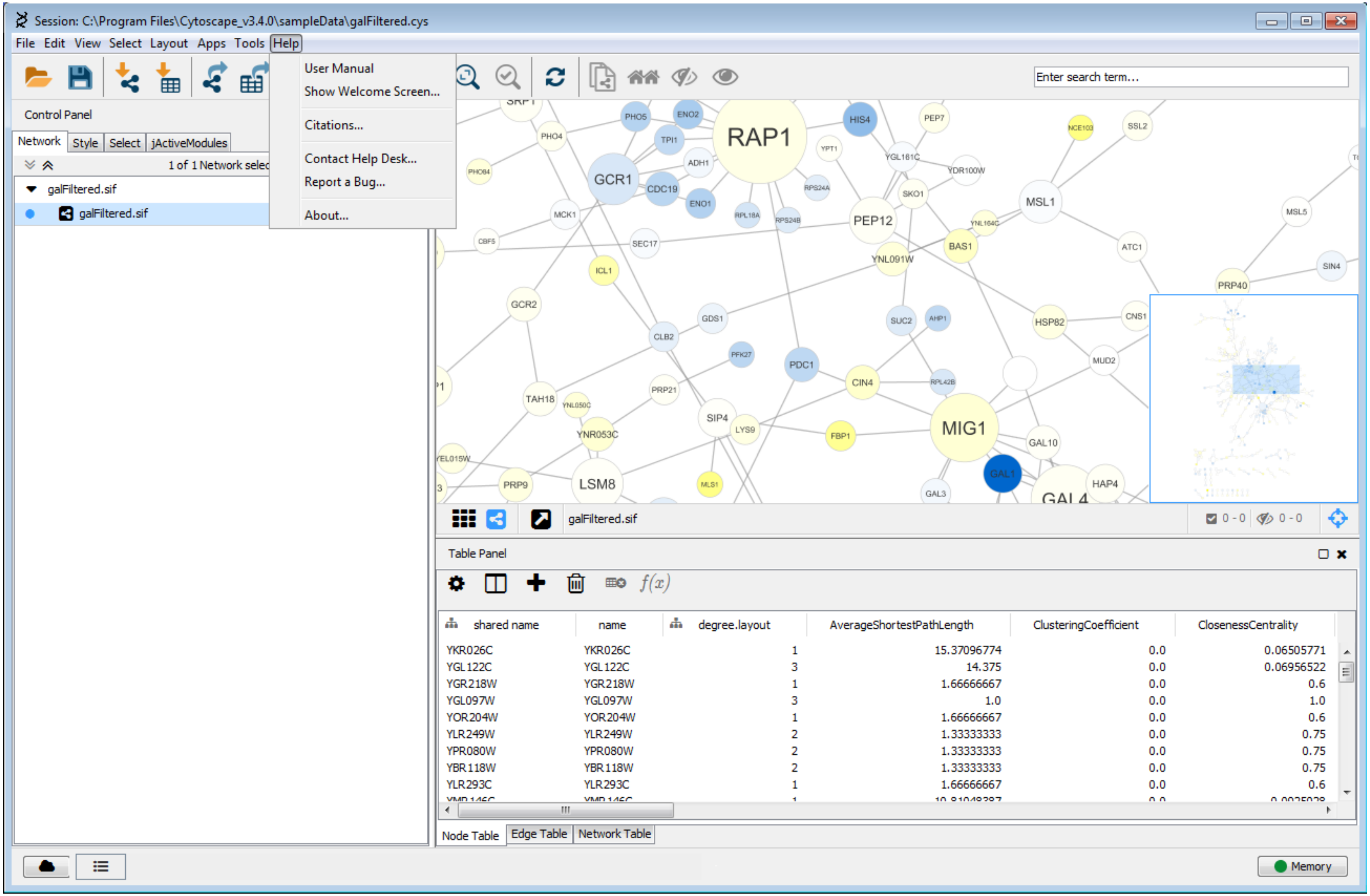
Saving and Exporting

- Sessions save everything as .cys files:
  Networks, Tables, Styles, Screen sizes, etc

- Export networks in different formats:
  SIF, GML, XGMML, BioPAX, PSI-MI 1 & 2.5

- Publication quality graphics in several formats:
  PDF, EPS, SVG, PNG, JPEG, and BMP

Getting Help



shared name	name	degree.layout	AverageShortestPathLength	ClusteringCoefficient	ClosenessCentrality
YKR026C	YKR026C	1	15.37096774	0.0	0.06505771
YGL122C	YGL122C	3	14.375	0.0	0.06956522
YGR218W	YGR218W	1	1.66666667	0.0	0.6
YGL097W	YGL097W	3	1.0	0.0	1.0
YOR204W	YOR204W	1	1.66666667	0.0	0.6
YLR249W	YLR249W	2	1.33333333	0.0	0.75
YPR080W	YPR080W	2	1.33333333	0.0	0.75
YBR118W	YBR118W	2	1.33333333	0.0	0.75
YLR293C	YLR293C	1	1.66666667	0.0	0.6

Tips & Tricks

- Network Collections
  - Each collection has a “root” network
  - Changing the attribute for a node in one network *will* also change that attribute for a node with the same SUID in all other networks within the collection
  - You can clone a network into a new collection to “decouple” it and start a new root

Tips & Tricks

- Network views
  - When you open a large network, you will not get a view by default
  - To improve interactive performance, Cytoscape has the concept of “Levels of Detail”
    - Some visual attributes will only be apparent when you zoom in
    - The level of detail for various attributes can be changed in the preferences
    - To see what things will look like at full detail:
      - View→Show Graphics Details

Tips & Tricks

- Sessions
  - – Sessions save pretty much everything:
    - Networks
    - Properties
    - Visual styles
    - Screen sizes
  - – Saving a session on a large screen may require some resizing when opened on your laptop

Tips & Tricks

- Task monitor
  - Current task displayed in status bar (lower left)
  - Icon opens complete task history
- Memory
  - Current status (lower right)
  - Toggle open for details and “Free Unused Memory” button

Tips & Tricks

- CytoscapeConfiguration directory
  - Your defaults and any apps downloaded from the App Store will go here
- App Manager
  - This is where you search/install/update/uninstall apps
  - You now have the option of disabling vs. uninstalling…
  - Can also install and update apps directly from the App Store website, if you have Cytoscape 3 up and running

Cytoscape: Platform

- Cytoscape as a platform
  - App architecture
    - http://apps.cytoscape.org
  - Use cases
    - Expression data analysis
    - Protein complexes
    - Agilent literature search

Hands-on Tutorial

Introduction to Cytoscape:

Networks, Data, Styles, Layouts and App Manager

tutorials.cytoscape.org

Hands-on Tutorial

App Tutorials:

http://tutorials.cytoscape.org

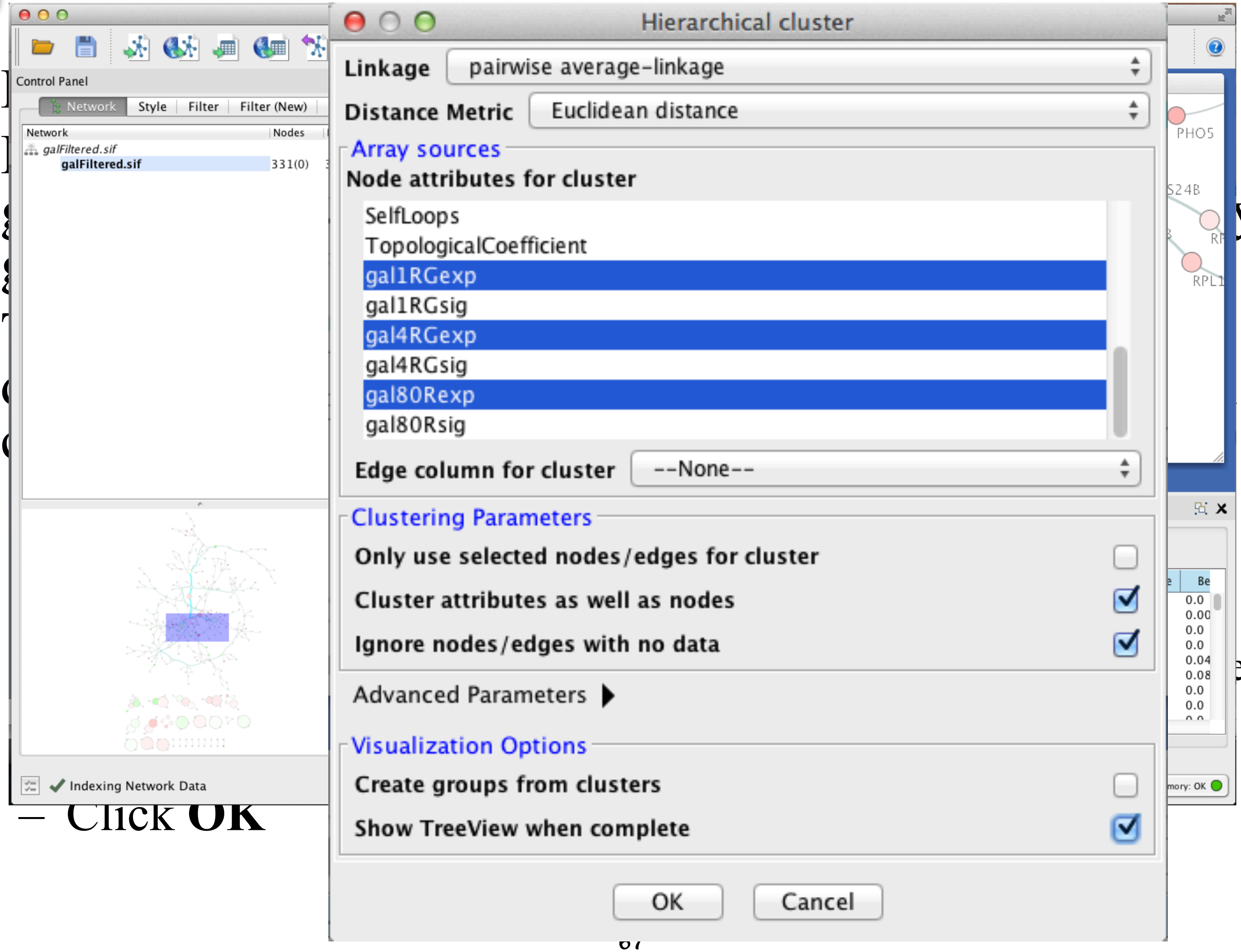
http://rbvi.ucsf.edu/cytoscape/structureViz2/tutorial.html

http://rinalyzer.de/tutorials.php

Examples/Demos

- clusterMaker
  - Clustering and cluster visualizations
- Agilent LitSearch Tool
  - Extracting networks from abstracts
- WikiPathways
  - Search and load pathway diagrams

Expression Data Analysis



Hierarchical cluster

Linkage: pairwise average-linkage

Distance Metric: Euclidean distance

Array sources

Node attributes for cluster

- SelfLoops
- TopologicalCoefficient
- gal1RGexp (selected)
- gal1RGsig
- gal4RGexp (selected)
- gal4RGsig
- gal80Rexp (selected)
- gal80Rsig

Edge column for cluster: --None--

Clustering Parameters

- **Only use selected nodes/edges for cluster** ☐
- **Cluster attributes as well as nodes** ☑
- **Ignore nodes/edges with no data** ☑

Advanced Parameters ▶

Visualization Options

- **Create groups from clusters** ☐
- **Show TreeView when complete** ☑

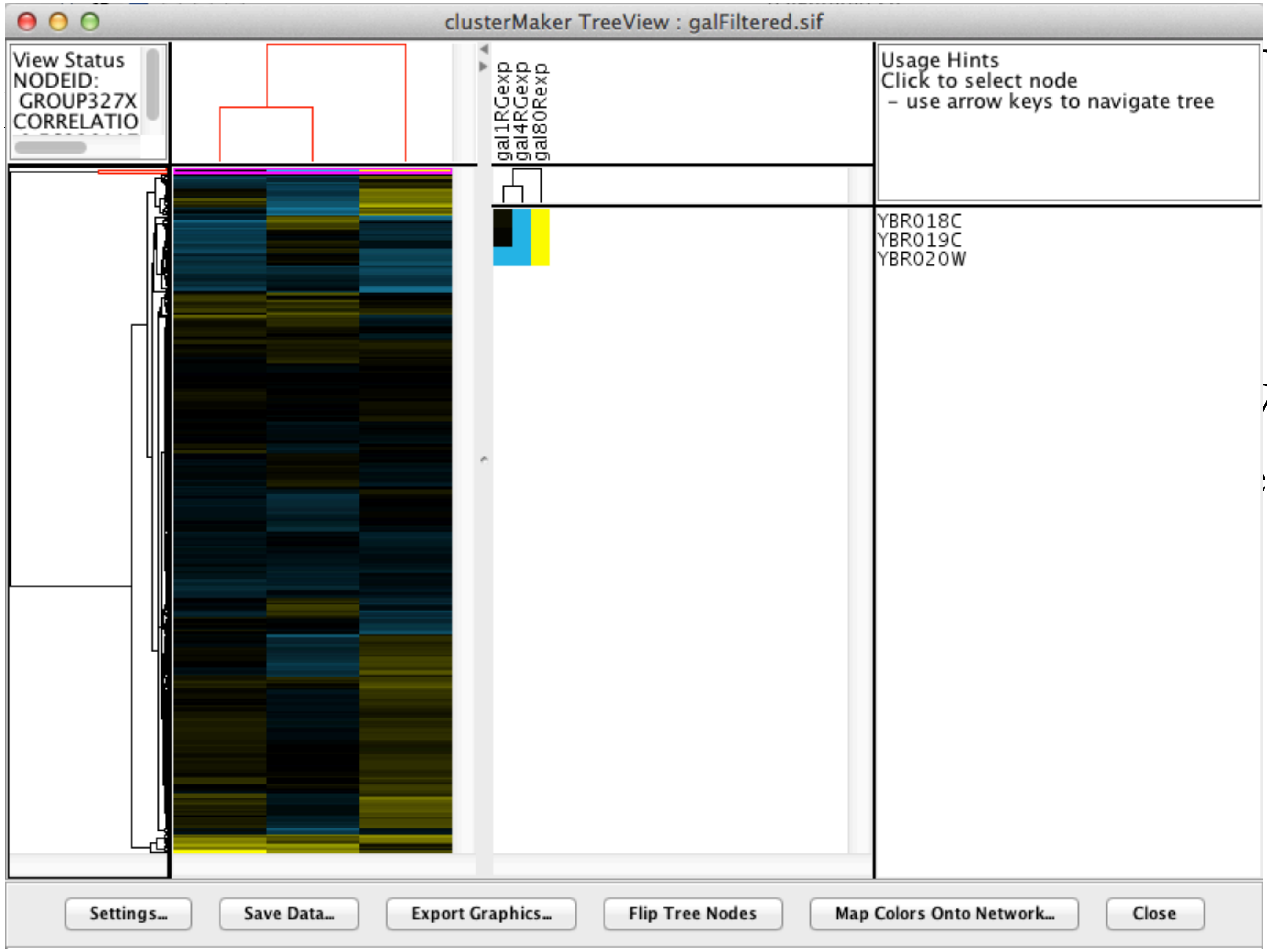
– Click **OK**

Expression Data Analysis

– ree

– wn

twork

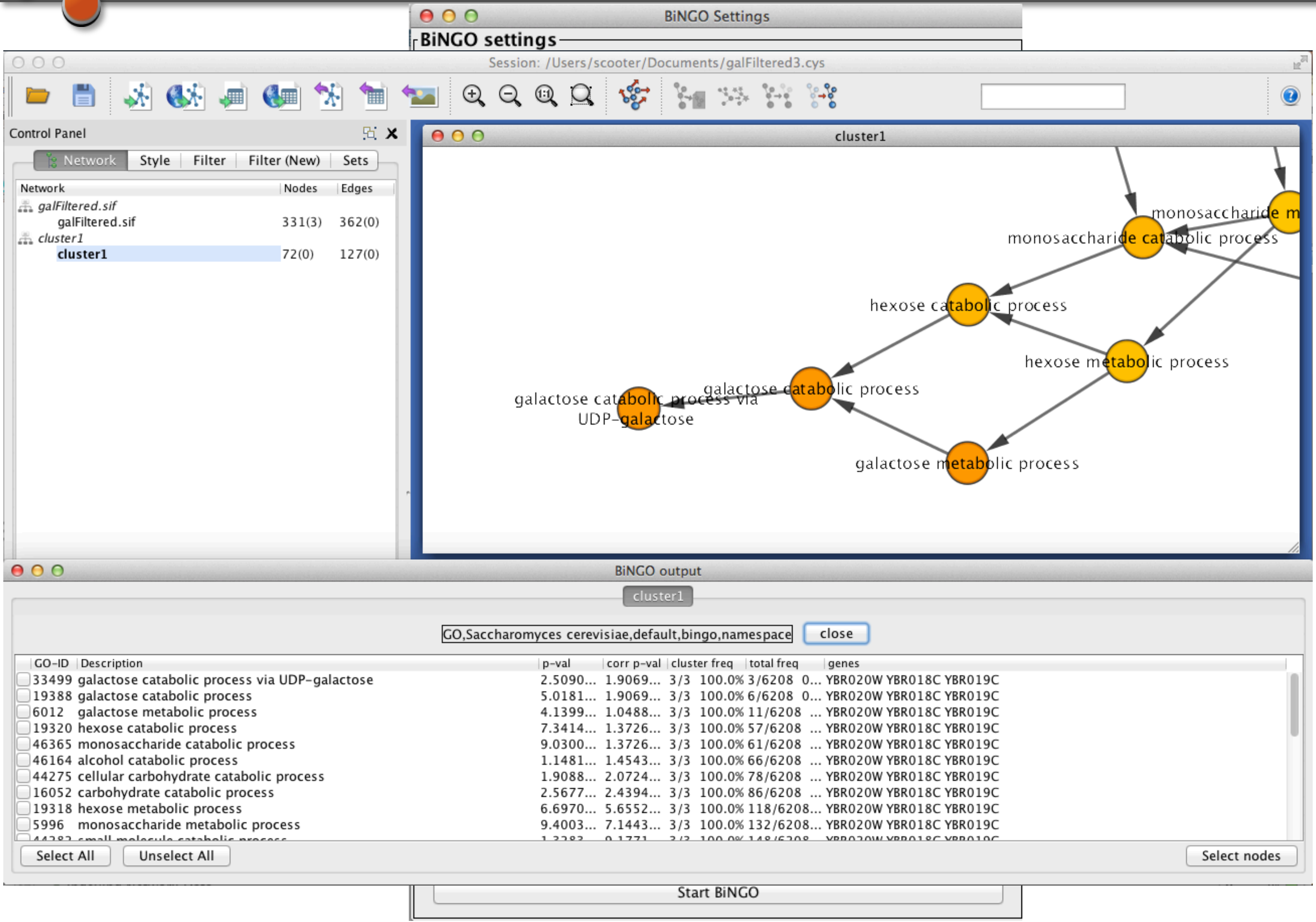


clusterMaker TreeView : galFiltered.sif

View Status
NODEID:
GROUP327X
CORRELATIO

Usage Hints
Click to select node
– use arrow keys to navigate tree

Expression Data Analysis



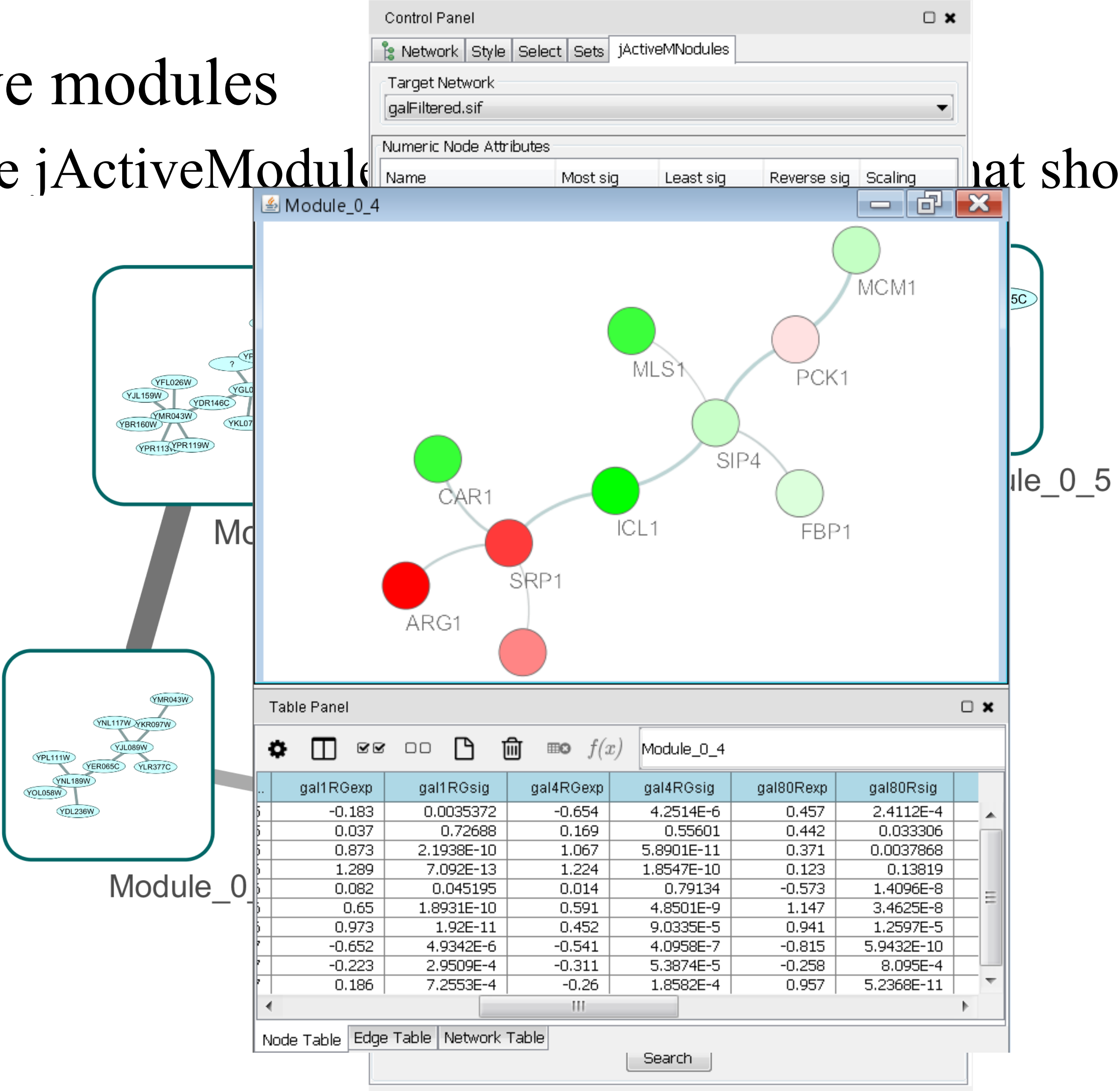
Network	Nodes	Edges
galFiltered.sif		
galFiltered.sif	331(3)	362(0)
cluster1		
cluster1	72(0)	127(0)

GO,Saccharomyces cerevisiae,default,bingo,namespace

GO-ID	Description	p-val	corr p-val	cluster freq	total freq	genes
☐ 33499	galactose catabolic process via UDP-galactose	2.5090...	1.9069...	3/3 100.0%	3/6208 0...	YBR020W YBR018C YBR019C
☐ 19388	galactose catabolic process	5.0181...	1.9069...	3/3 100.0%	6/6208 0...	YBR020W YBR018C YBR019C
☐ 6012	galactose metabolic process	4.1399...	1.0488...	3/3 100.0%	11/6208 ...	YBR020W YBR018C YBR019C
☐ 19320	hexose catabolic process	7.3414...	1.3726...	3/3 100.0%	57/6208 ...	YBR020W YBR018C YBR019C
☐ 46365	monosaccharide catabolic process	9.0300...	1.3726...	3/3 100.0%	61/6208 ...	YBR020W YBR018C YBR019C
☐ 46164	alcohol catabolic process	1.1481...	1.4543...	3/3 100.0%	66/6208 ...	YBR020W YBR018C YBR019C
☐ 44275	cellular carbohydrate catabolic process	1.9088...	2.0724...	3/3 100.0%	78/6208 ...	YBR020W YBR018C YBR019C
☐ 16052	carbohydrate catabolic process	2.5677...	2.4394...	3/3 100.0%	86/6208 ...	YBR020W YBR018C YBR019C
☐ 19318	hexose metabolic process	6.6970...	5.6552...	3/3 100.0%	118/6208...	YBR020W YBR018C YBR019C
☐ 5996	monosaccharide metabolic process	9.4003...	7.1443...	3/3 100.0%	132/6208...	YBR020W YBR018C YBR019C
☐ 44282	small molecule catabolic process	1.3283...	9.1771...	3/3 100.0%	148/6208...	YBR020W YBR018C YBR019C

Expression Data Analysis

- Active modules
  - The jActiveModule … hat show … r time- … ules … nges



Target Network: galFiltered.sif

Numeric Node Attributes

Name	Most sig	Least sig	Reverse sig	Scaling

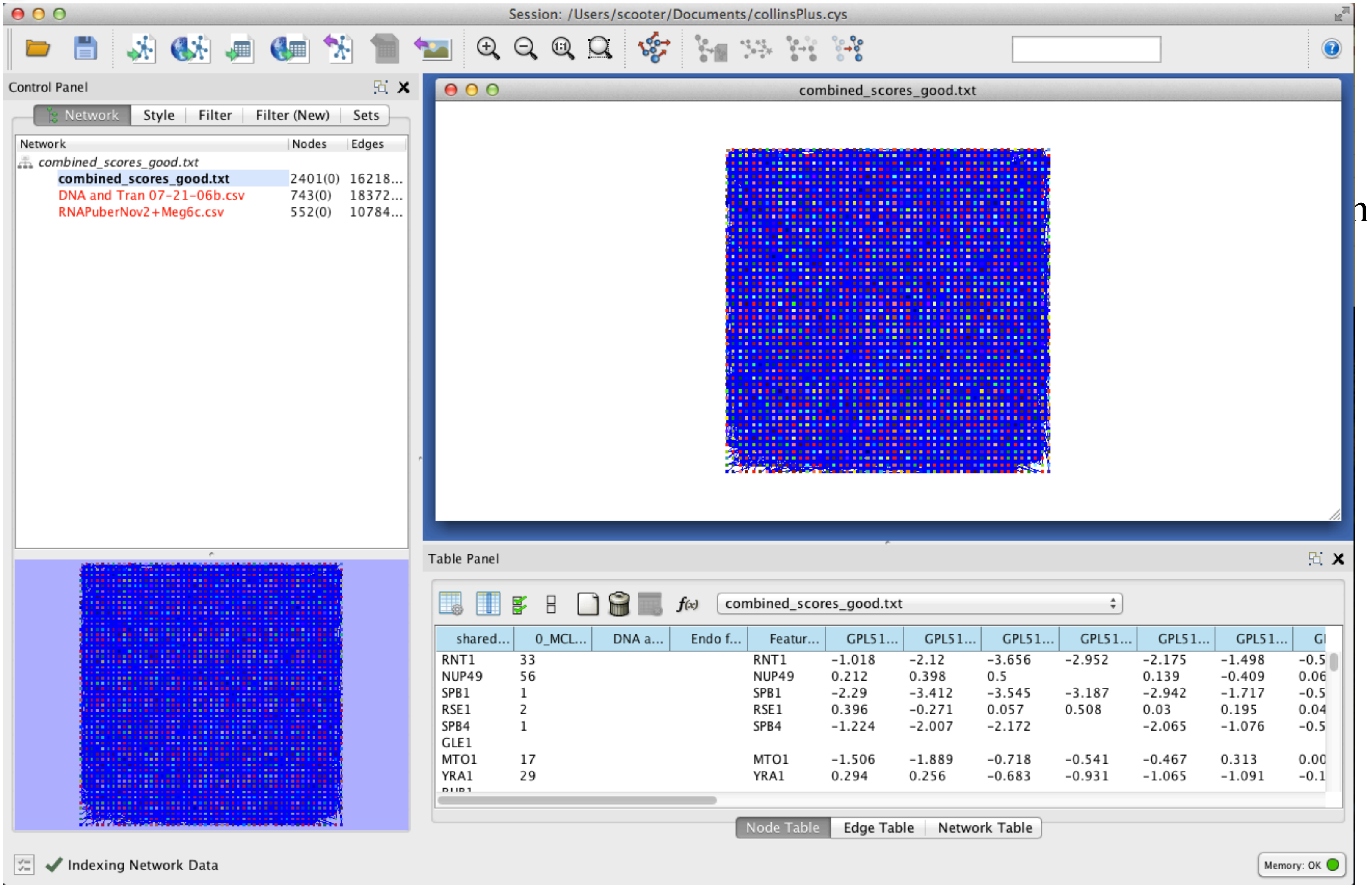
Table Panel: Module_0_4

gal1RGexp	gal1RGsig	gal4RGexp	gal4RGsig	gal80Rexp	gal80Rsig
-0.183	0.0035372	-0.654	4.2514E-6	0.457	2.4112E-4
0.037	0.72688	0.169	0.55601	0.442	0.033306
0.873	2.1938E-10	1.067	5.8901E-11	0.371	0.0037868
1.289	7.092E-13	1.224	1.8547E-10	0.123	0.13819
0.082	0.045195	0.014	0.79134	-0.573	1.4096E-8
0.65	1.8931E-10	0.591	4.8501E-9	1.147	3.4625E-8
0.973	1.92E-11	0.452	9.0335E-5	0.941	1.2597E-5
-0.652	4.9342E-6	-0.541	4.0958E-7	-0.815	5.9432E-10
-0.223	2.9509E-4	-0.311	5.3874E-5	-0.258	8.095E-4
0.186	7.2553E-4	-0.26	1.8582E-4	0.957	5.2368E-11

Protein Complexes

•

n



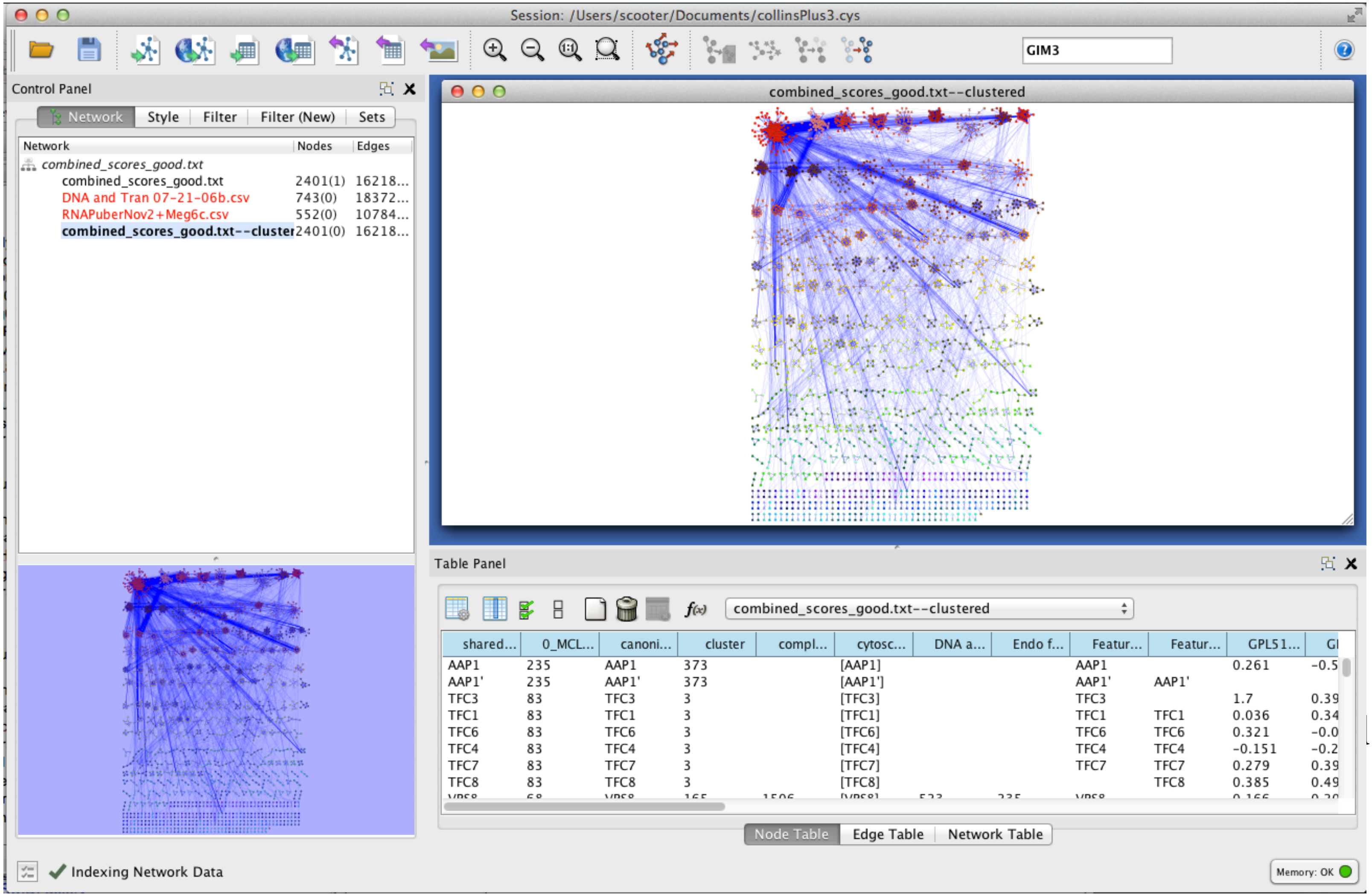
Network	Nodes	Edges
combined_scores_good.txt		
combined_scores_good.txt	2401(0)	16218...
DNA and Tran 07-21-06b.csv	743(0)	18372...
RNAPuberNov2+Meg6c.csv	552(0)	10784...

combined_scores_good.txt

shared...	0_MCL...	DNA a...	Endo f...	Featur...	GPL51...	GPL51...	GPL51...	GPL51...	GPL51...	GPL51...	GI
RNT1	33			RNT1	-1.018	-2.12	-3.656	-2.952	-2.175	-1.498	-0.5
NUP49	56			NUP49	0.212	0.398	0.5		0.139	-0.409	0.06
SPB1	1			SPB1	-2.29	-3.412	-3.545	-3.187	-2.942	-1.717	-0.5
RSE1	2			RSE1	0.396	-0.271	0.057	0.508	0.03	0.195	0.04
SPB4	1			SPB4	-1.224	-2.007	-2.172		-2.065	-1.076	-0.5
GLE1											
MTO1	17			MTO1	-1.506	-1.889	-0.718	-0.541	-0.467	0.313	0.00
YRA1	29			YRA1	0.294	0.256	-0.683	-0.931	-1.065	-1.091	-0.1

✓ Indexing Network Data

Protein Complexes



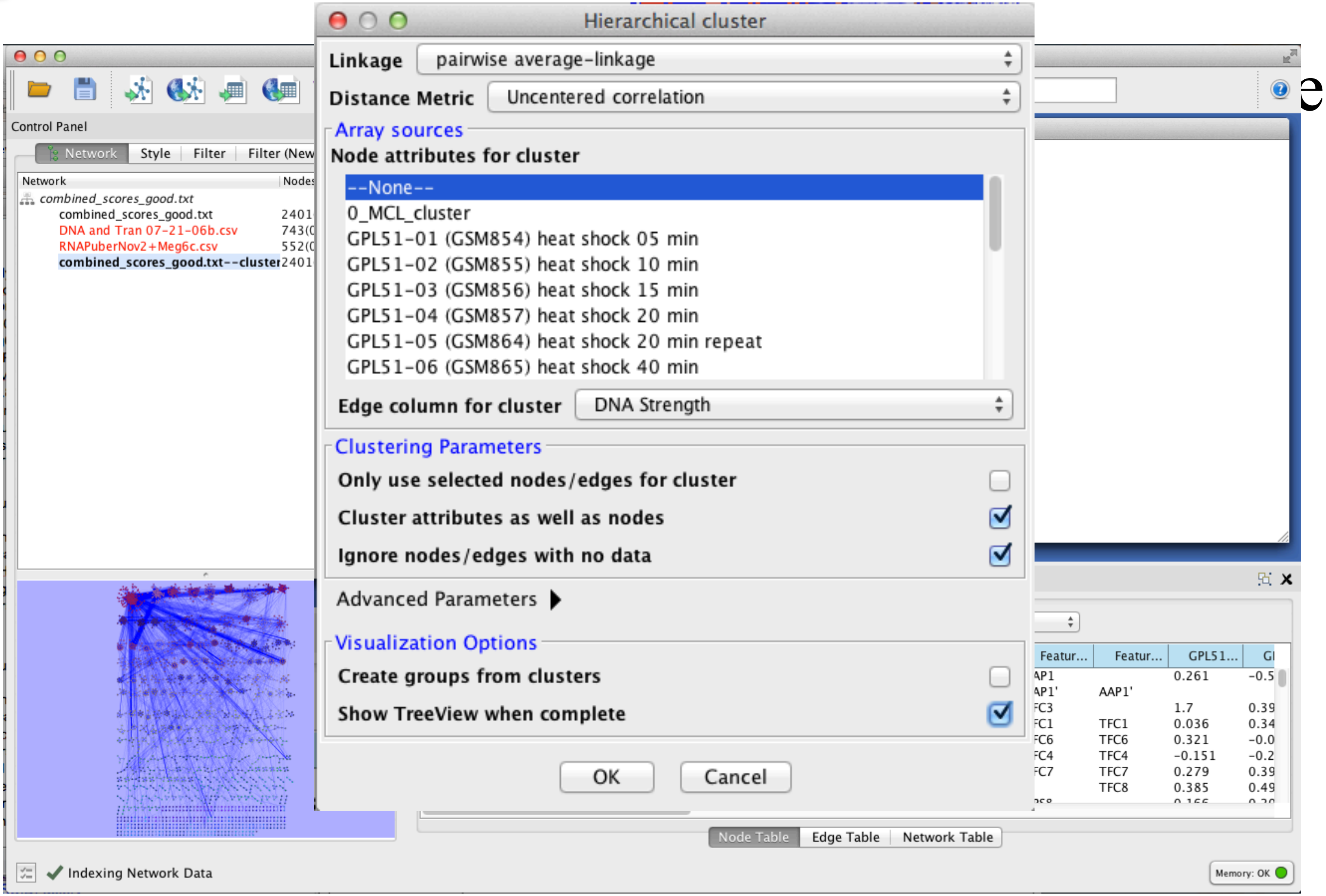
Network	Nodes	Edges
combined_scores_good.txt		
combined_scores_good.txt	2401(1)	16218...
DNA and Tran 07-21-06b.csv	743(0)	18372...
RNAPuberNov2+Meg6c.csv	552(0)	10784...
combined_scores_good.txt--cluster	2401(0)	16218...

combined_scores_good.txt--clustered

shared...	0_MCL...	canoni...	cluster	compl...	cytosc...	DNA a...	Endo f...	Featur...	Featur...	GPL51...	Gl
AAP1	235	AAP1	373		[AAP1]			AAP1		0.261	-0.5
AAP1'	235	AAP1'	373		[AAP1']			AAP1'	AAP1'		
TFC3	83	TFC3	3		[TFC3]			TFC3		1.7	0.39
TFC1	83	TFC1	3		[TFC1]			TFC1	TFC1	0.036	0.34
TFC6	83	TFC6	3		[TFC6]			TFC6	TFC6	0.321	-0.0
TFC4	83	TFC4	3		[TFC4]			TFC4	TFC4	-0.151	-0.2
TFC7	83	TFC7	3		[TFC7]			TFC7	TFC7	0.279	0.39
TFC8	83	TFC8	3		[TFC8]				TFC8	0.385	0.49

ges

Protein Complexes



Control Panel

Network	Nodes
combined_scores_good.txt	
combined_scores_good.txt	2401
DNA and Tran 07-21-06b.csv	743(
RNAPuberNov2+Meg6c.csv	552(
combined_scores_good.txt--cluster	2401

Hierarchical cluster

Linkage pairwise average-linkage

Distance Metric Uncentered correlation

Array sources

Node attributes for cluster

- --None--
- 0_MCL_cluster
- GPL51-01 (GSM854) heat shock 05 min
- GPL51-02 (GSM855) heat shock 10 min
- GPL51-03 (GSM856) heat shock 15 min
- GPL51-04 (GSM857) heat shock 20 min
- GPL51-05 (GSM864) heat shock 20 min repeat
- GPL51-06 (GSM865) heat shock 40 min

Edge column for cluster DNA Strength

Clustering Parameters

Only use selected nodes/edges for cluster ☐

Cluster attributes as well as nodes ☑

Ignore nodes/edges with no data ☑

Advanced Parameters ▶

Visualization Options

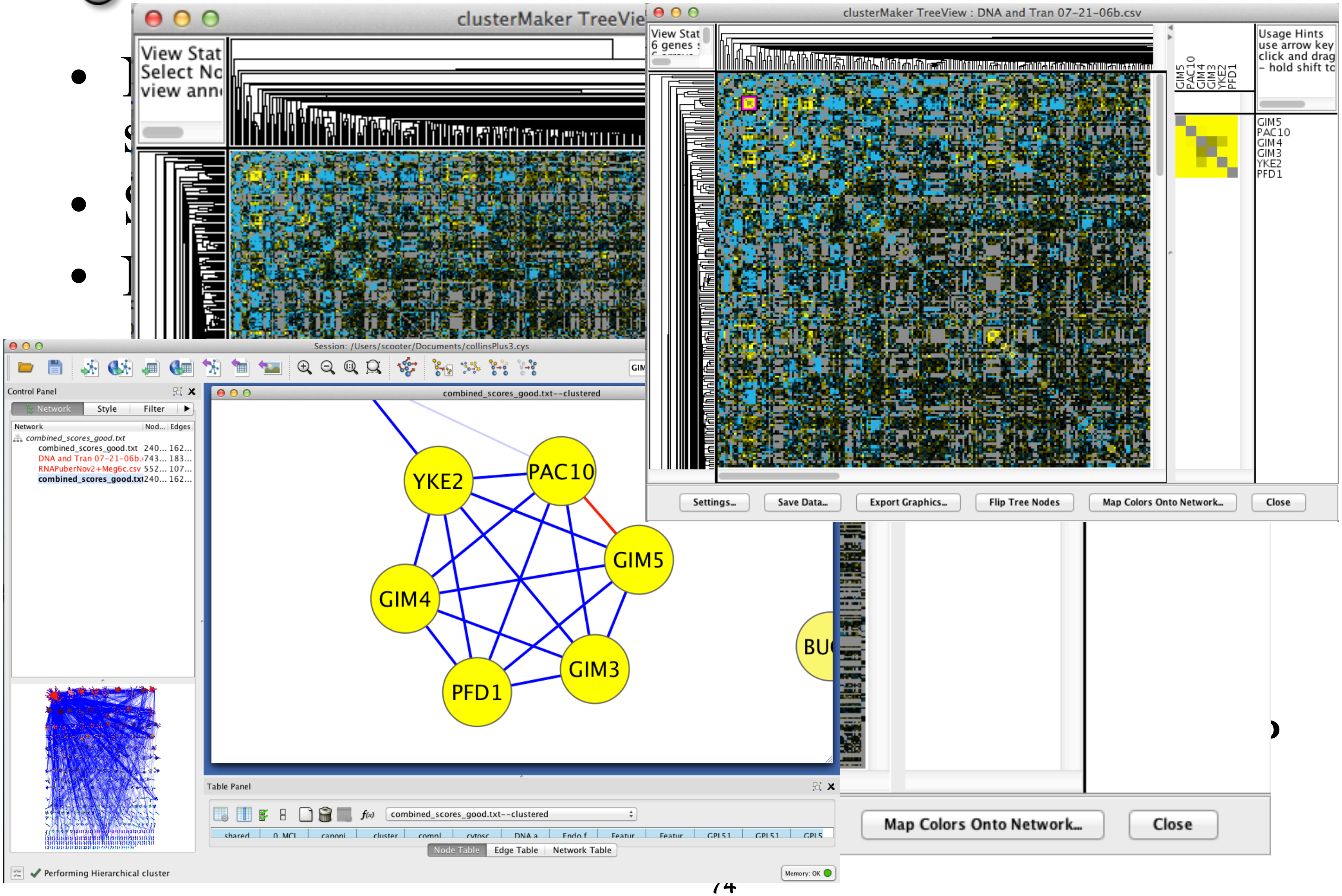
Create groups from clusters ☐

Show TreeView when complete ☑

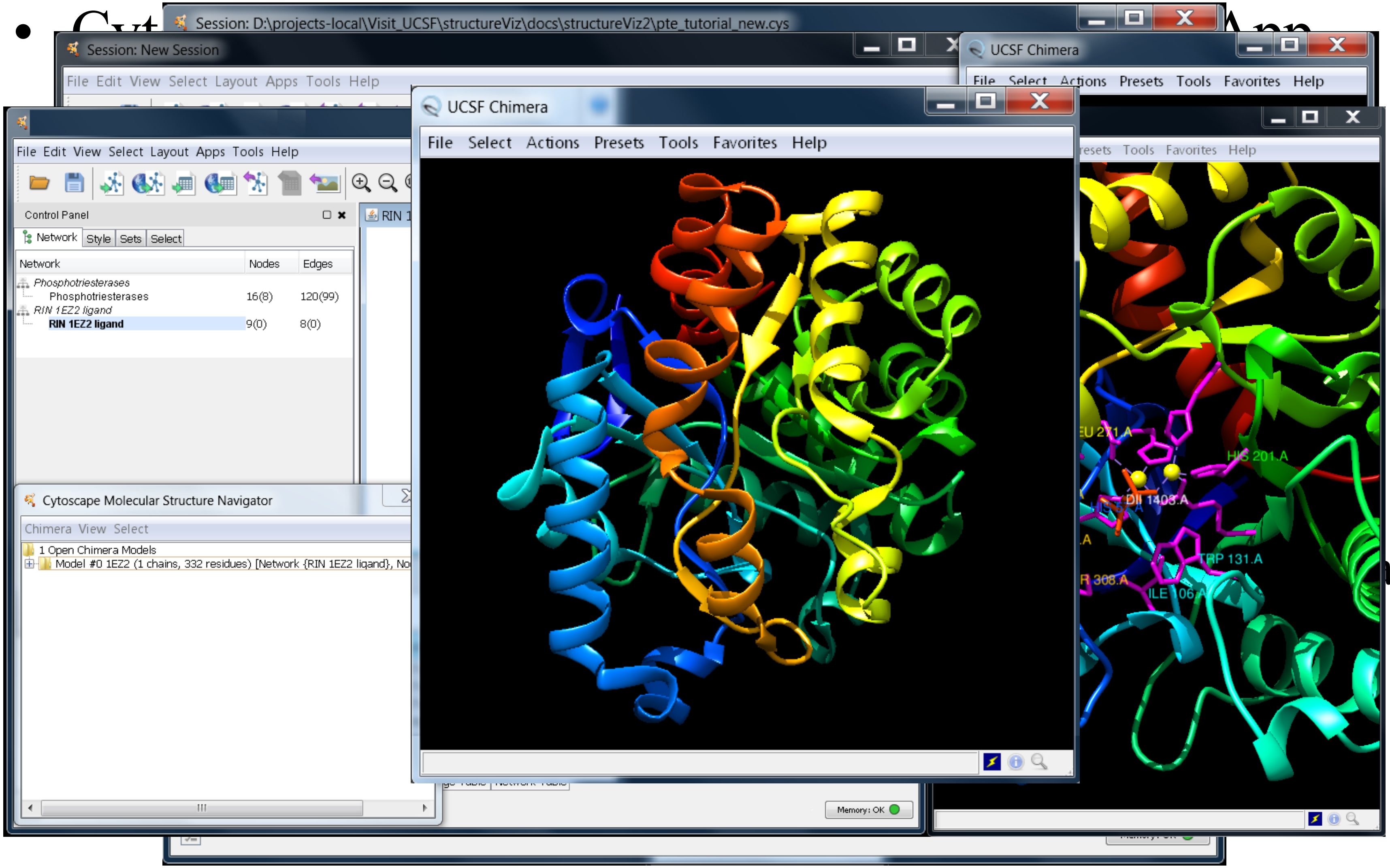
Featur...	Featur...	GPL51...	GI
AP1		0.261	-0.5
AP1'	AAP1'		
FC3		1.7	0.39
FC1	TFC1	0.036	0.34
FC6	TFC6	0.321	-0.0
FC4	TFC4	-0.151	-0.2
FC7	TFC7	0.279	0.39
	TFC8	0.385	0.49

Indexing Network Data

Protein Complexes



From Networks to Structures



Session: D:\projects-local\Visit_UCSF\structureViz\docs\structureViz2\pte_tutorial_new.cys

Session: New Session

Network	Nodes	Edges
Phosphotriesterases		
Phosphotriesterases	16(8)	120(99)
RIN 1EZ2 ligand		
RIN 1EZ2 ligand	9(0)	8(0)

Cytoscape Molecular Structure Navigator

1 Open Chimera Models

Model #0 1EZ2 (1 chains, 332 residues) [Network {RIN 1EZ2 ligand}, No

Questions?

- scooter@cgl.ucsf.edu
- cytoscape-helpdesk@googlegroups.com