



**John “Scooter” Morris  
Nadezhda T. Doncheva**

**October 3, 2015  
EMBO Practical Course  
TGAC, Norwich, UK**



# Outline

- Biological Networks
  - Why Networks?
  - Biological Network Taxonomy
  - Analytical Approaches
  - Visualization
- *Coffee Break*
- Introduction to Cytoscape
- Hands on Tutorial
  - Data import
  - Layout and apps
- *Lunch Break*
- Hands on: Using Cytoscape to explore YOUR data



# 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
  - 2010-current:
    - Research scientist at the Max Planck Institute for Informatics under the supervision of Mario Albrecht
  - Author or co-developer of several Cytoscape plugins:
    - RINalyzer, structureViz2, setsApp, NetworkAnalyzer, NetworkPrioritizer



# Why Networks?

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

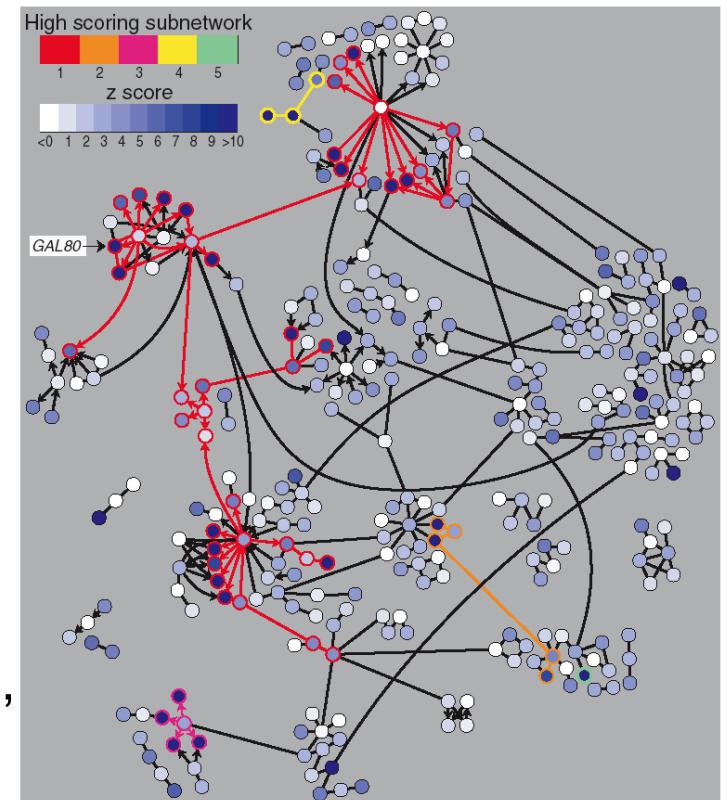
- **Network tools allow...**

## Analysis

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

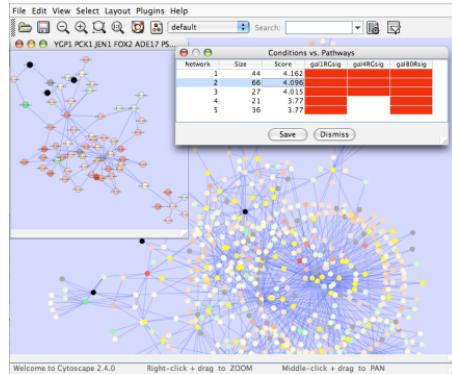
## Visualization

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





# Applications of Network Biology



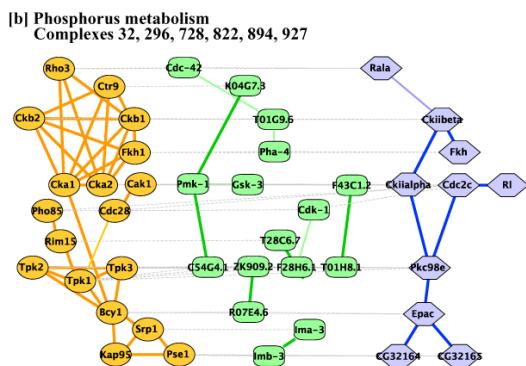
jActiveModules, UCSD

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



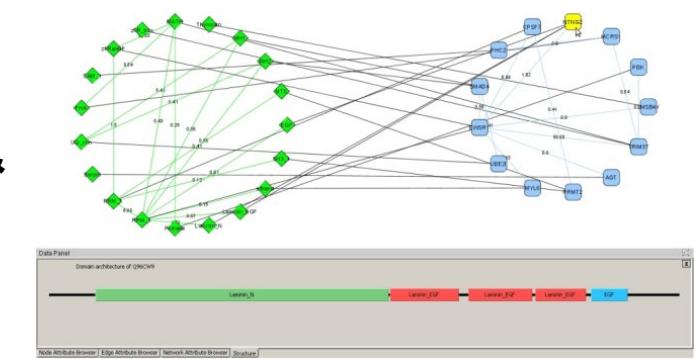
mCode, University of Toronto

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



PathBlast, UCSD

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

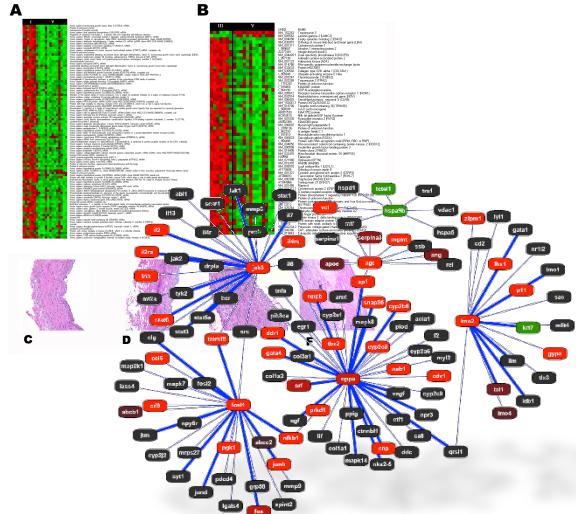


DomainGraph, Max Planck Institute

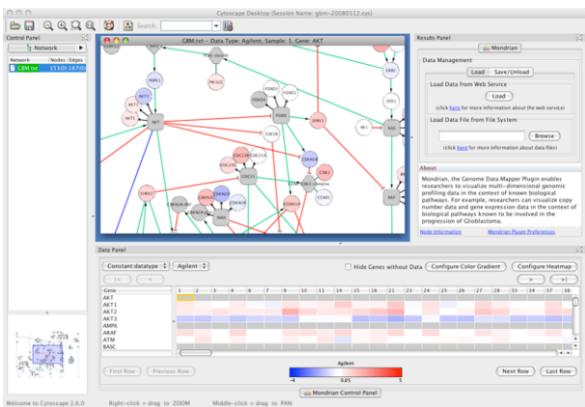
- **Prediction of interactions & functional associations** statistically significant domain-domain correlations in protein interaction network to predict protein-protein or genetic interaction



# Applications in Disease

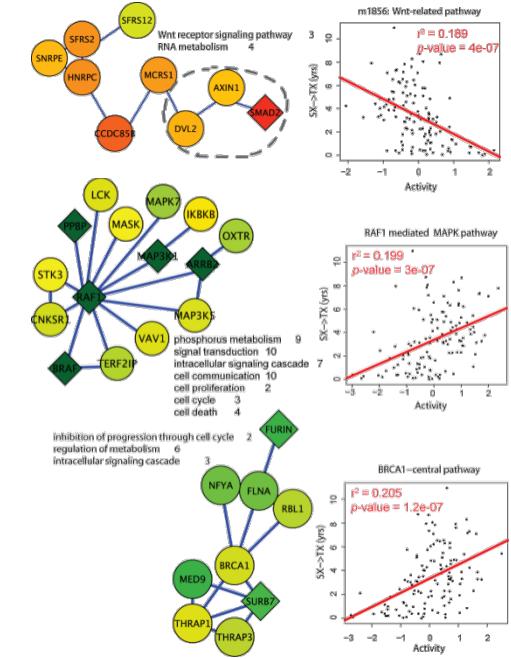


Agilent Literature Search



Mondrian, MSKCC

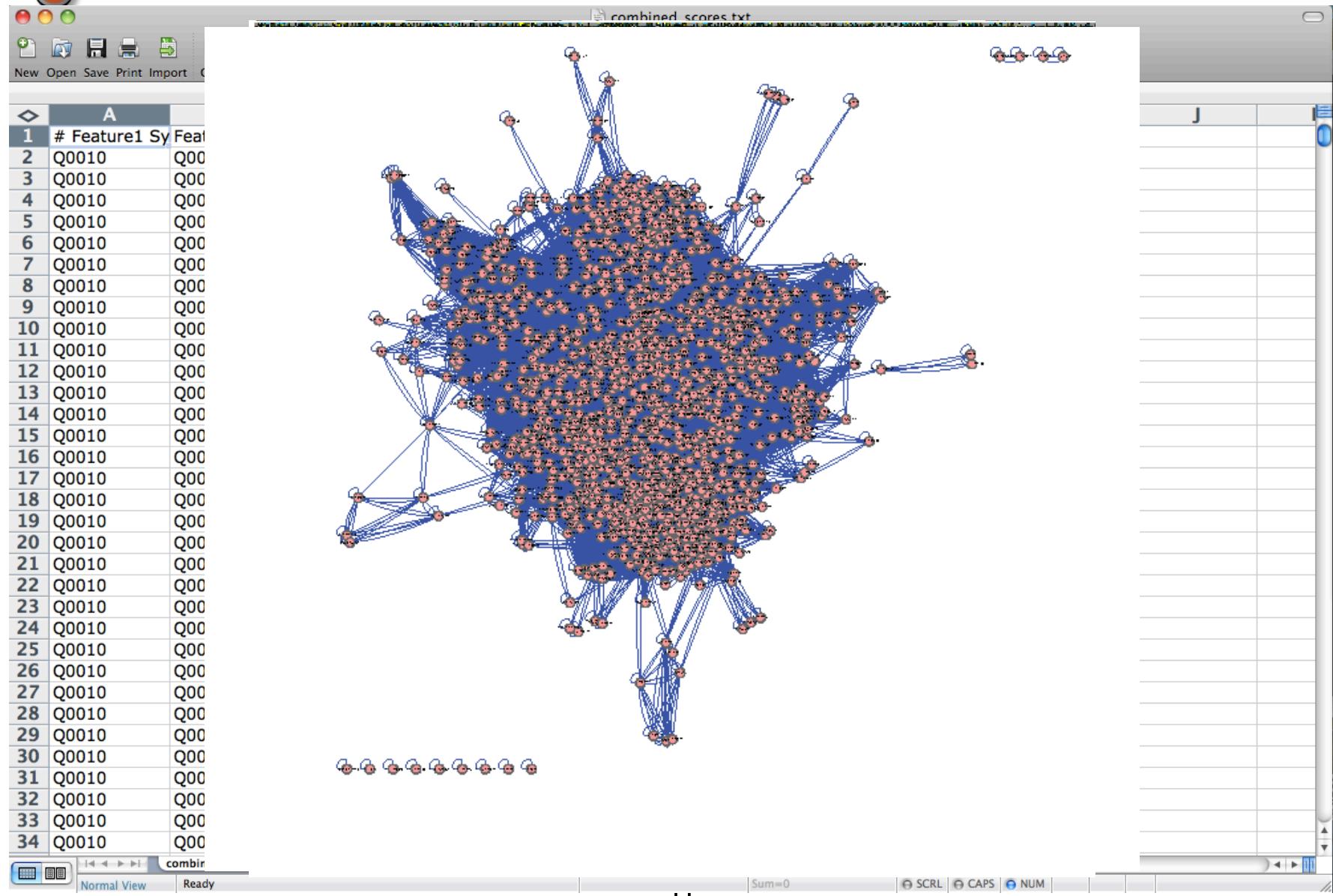
- **Identification of disease subnetworks** – identification of disease network subnetworks that are transcriptionally active in disease.
- **Subnetwork-based diagnosis** – source of biomarkers for disease classification, identify interconnected genes whose aggregate expression levels are predictive of disease state
- **Subnetwork-based gene association** – map common pathway mechanisms affected by collection of genotypes (SNP, CNV)



PinnacleZ, UCSD



# The Challenge



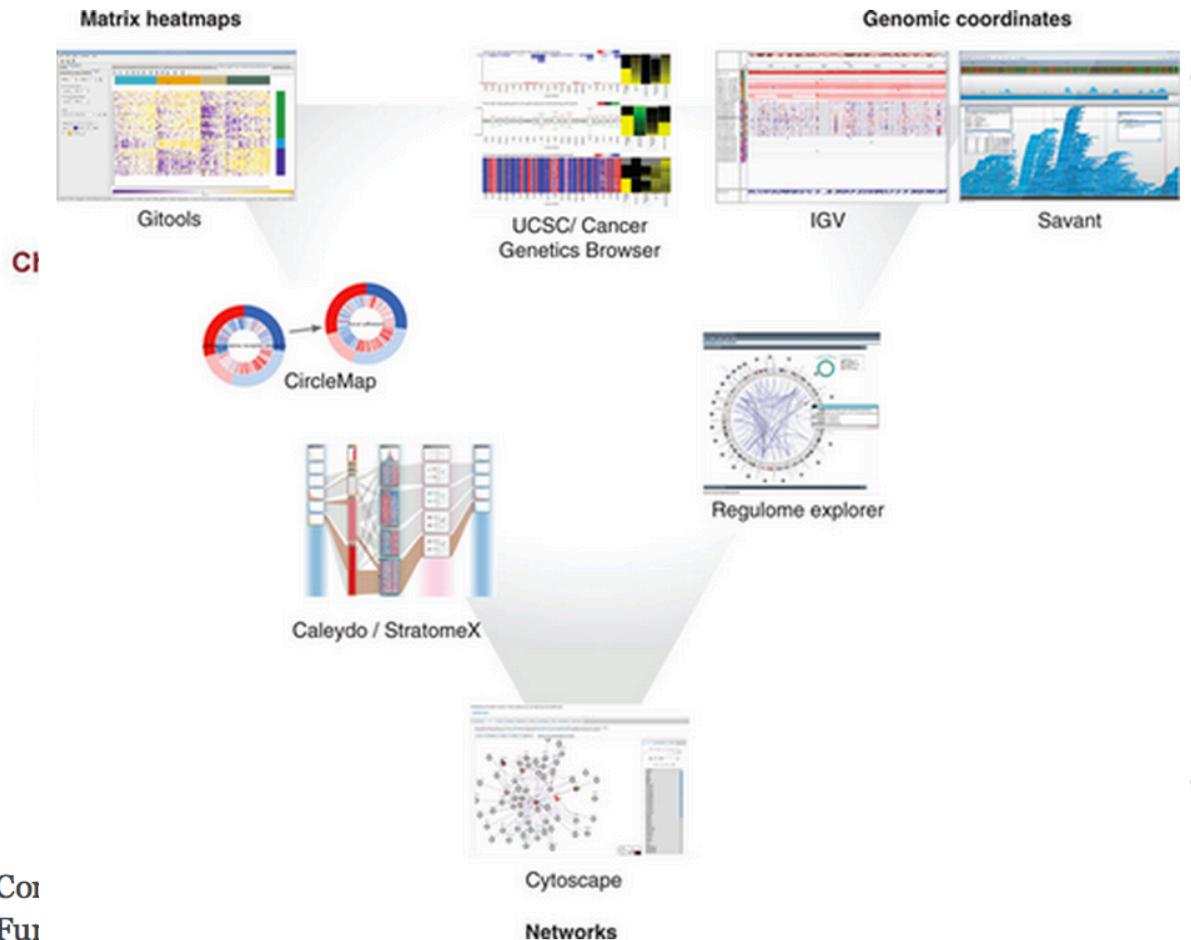


# The Challenge

- Biological networks
  - Seldom tell us anything by themselves
  - **Analysis** involves:
    - Understanding the characteristics of the network
      - Modularity
      - Comparison with other networks (i.e., random networks)
  - **Visualization** involves:
    - Placing nodes in a meaningful way (layouts)
    - Mapping biologically relevant data to the network
      - Node size, node color, edge weights, etc
    - *...which then allowing for more analysis!*



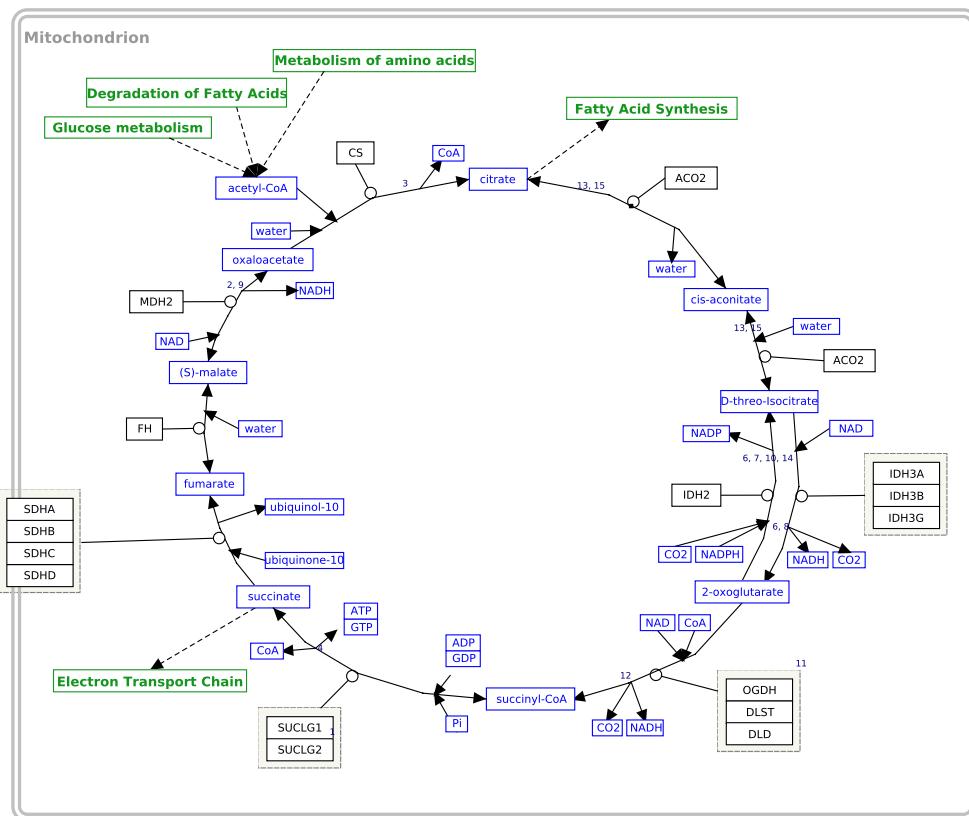
# The Challenge



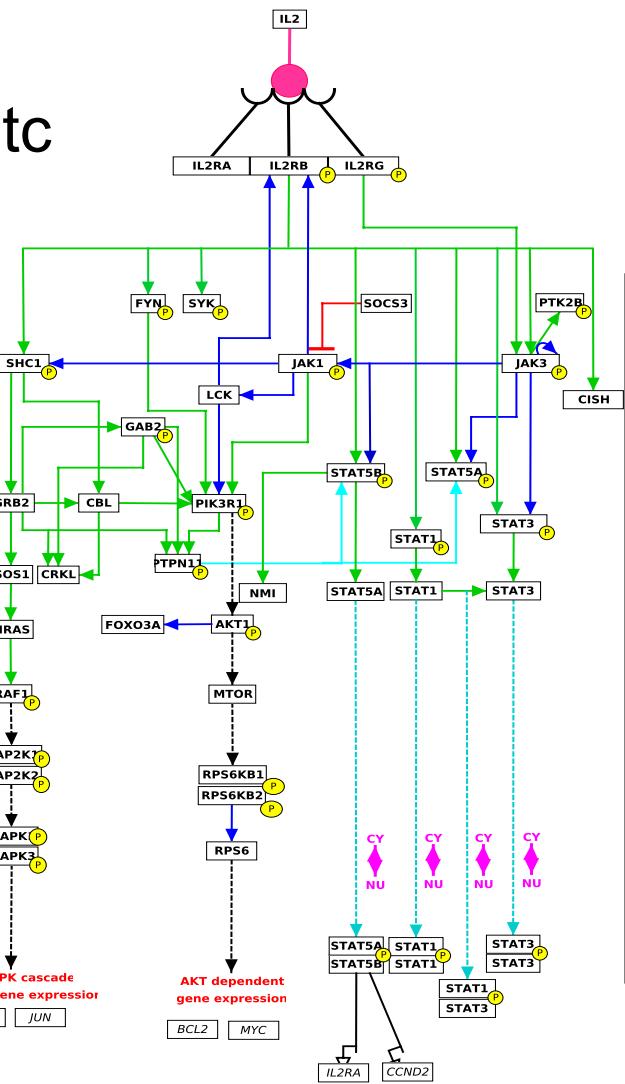


# Biological Network Taxonomy

- Pathways
  - Signaling, Metabolic, Regulatory, etc



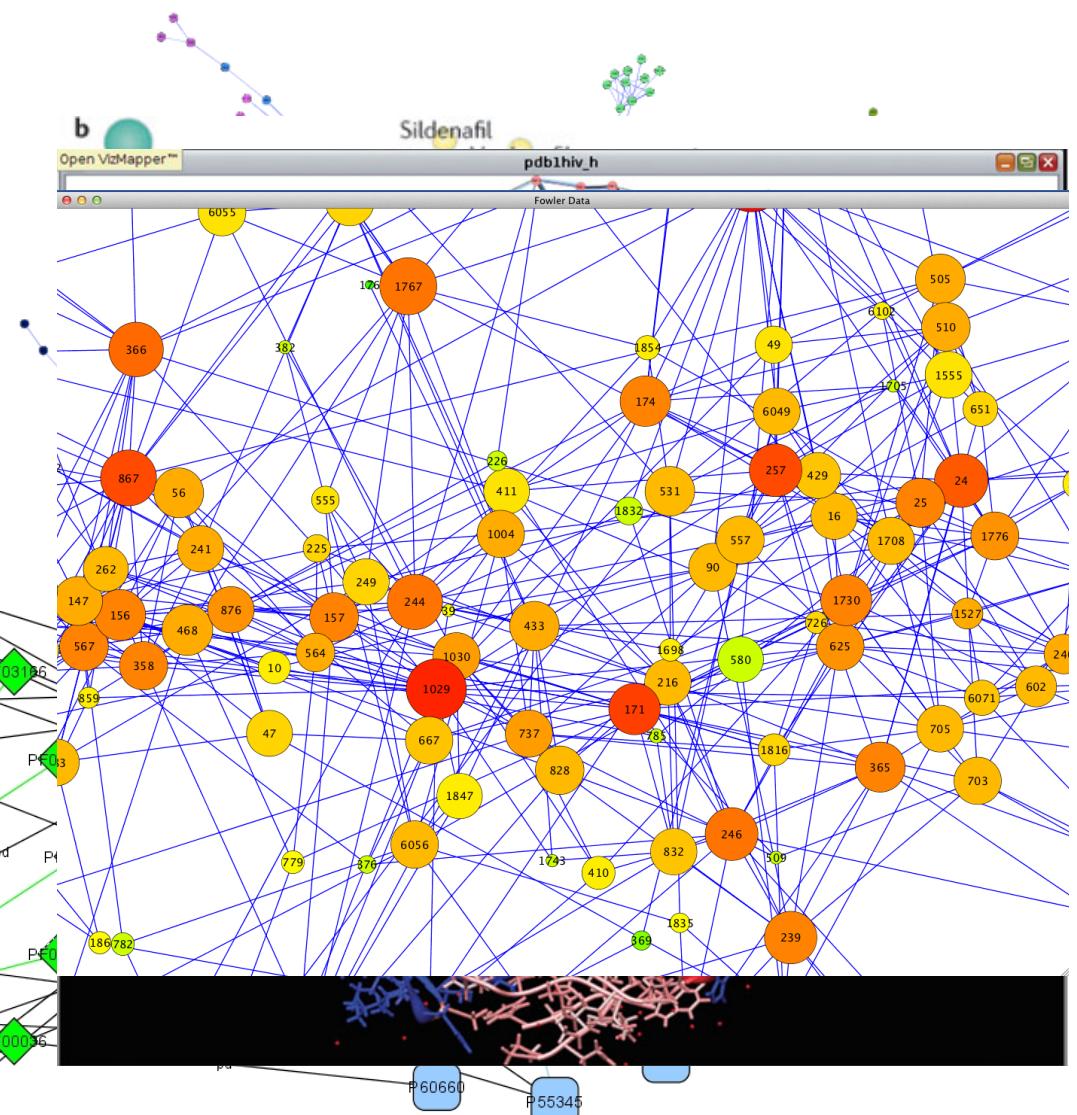
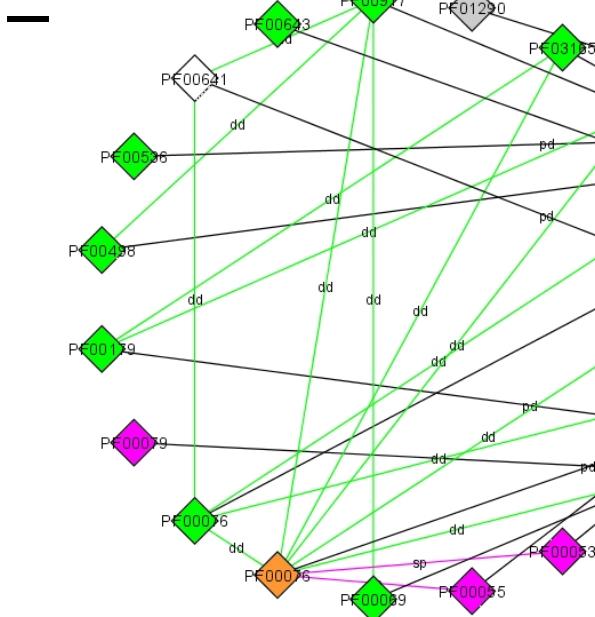
IL-2 Signaling Pathway





# Biological Network Taxonomy

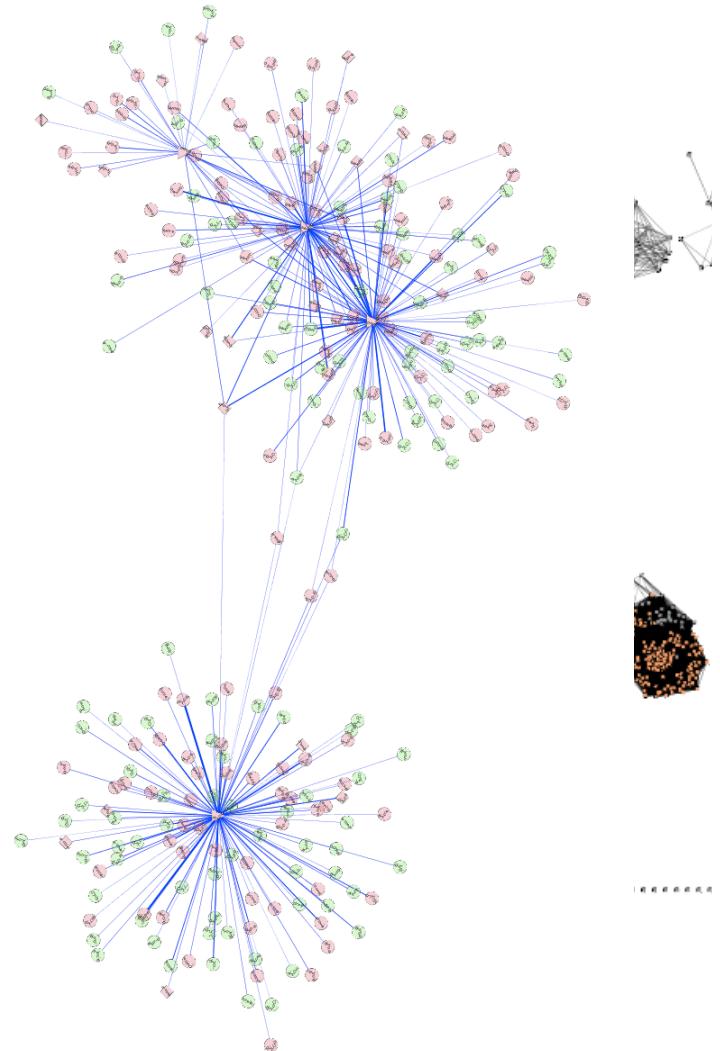
- Interactions
  - Protein-Protein
  - Protein-Ligand
  - Domain-Domain





# Biological Network Taxonomy

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

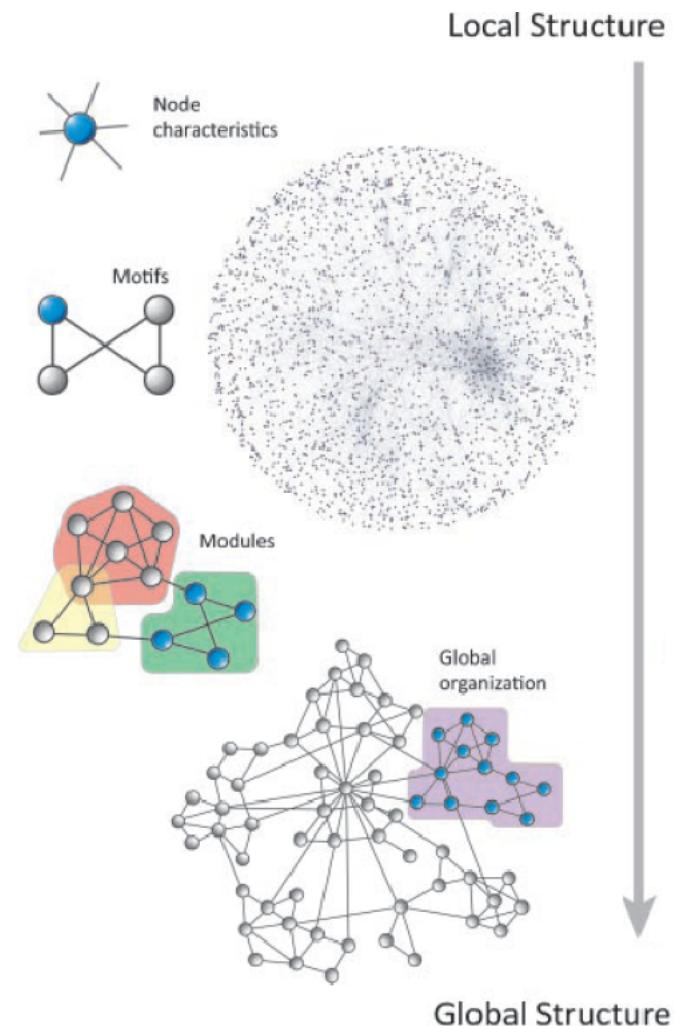




# Analytical Approaches

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

**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.

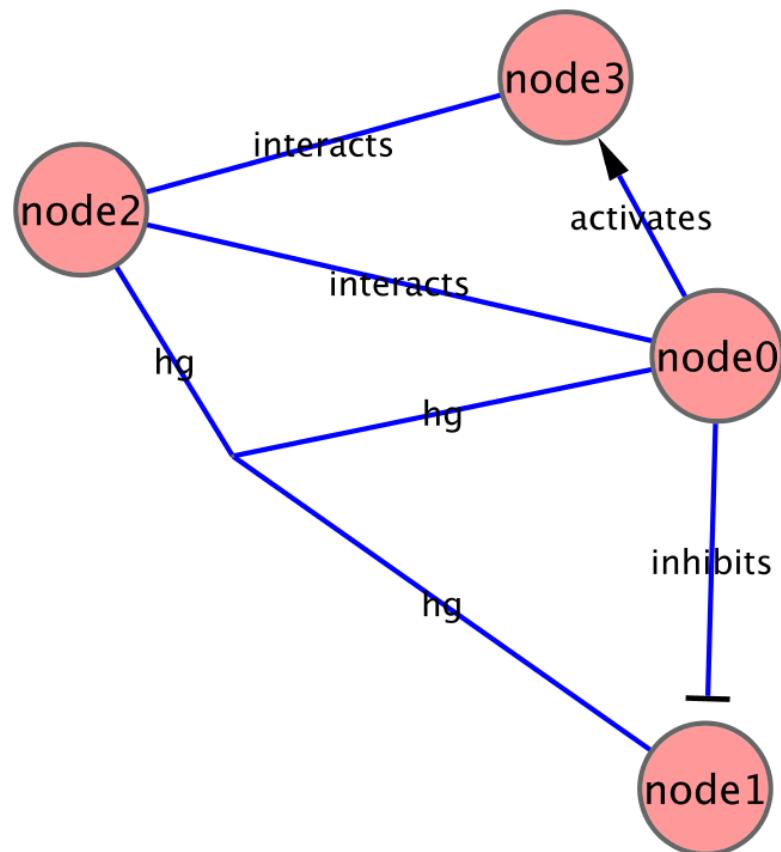
**Motif analysis** is the identification of small network patterns that are over-represented when compared with a randomized version of the same network. Regulatory elements are often composed of such motifs.

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



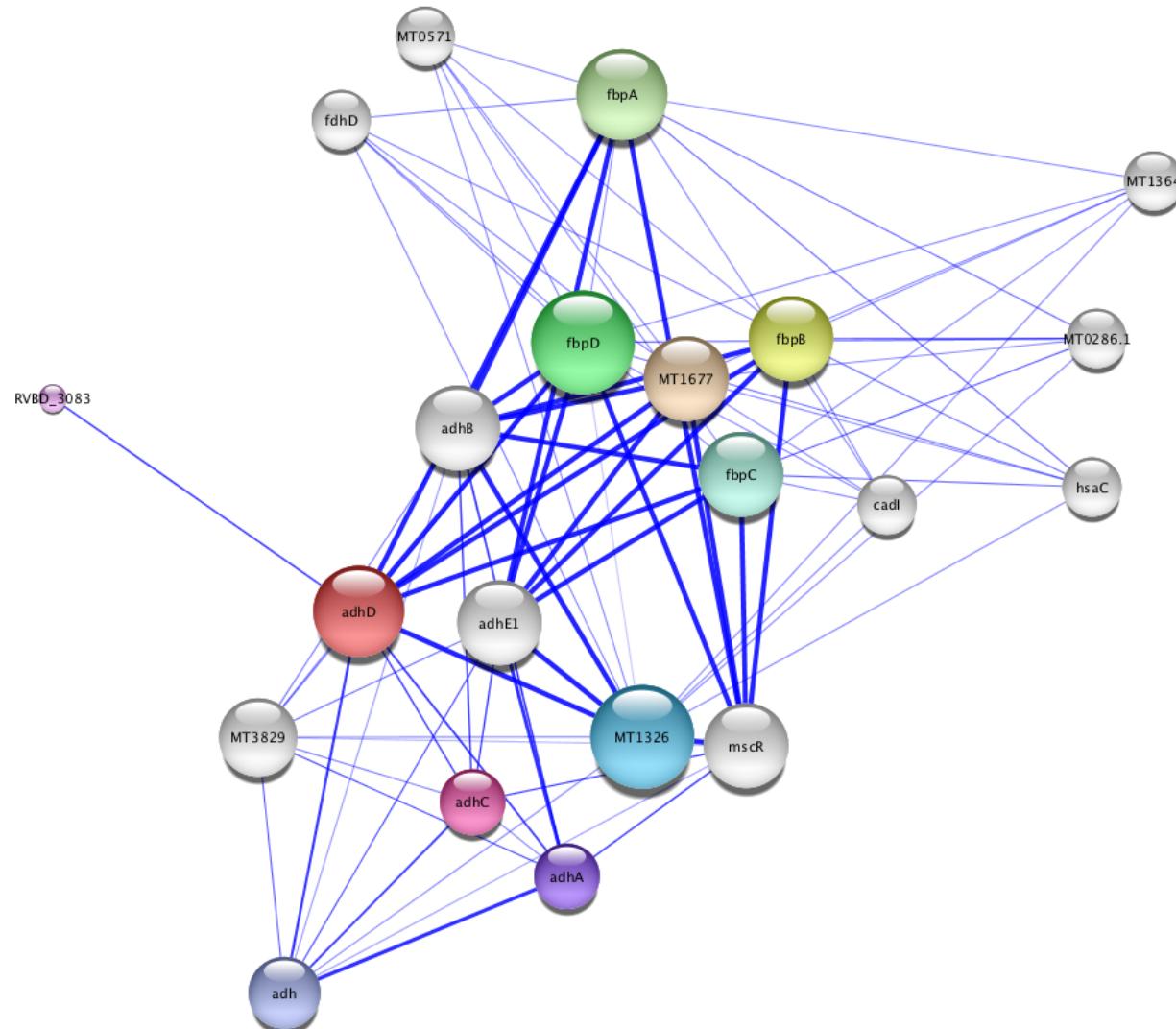
# Analytical Approaches

- 



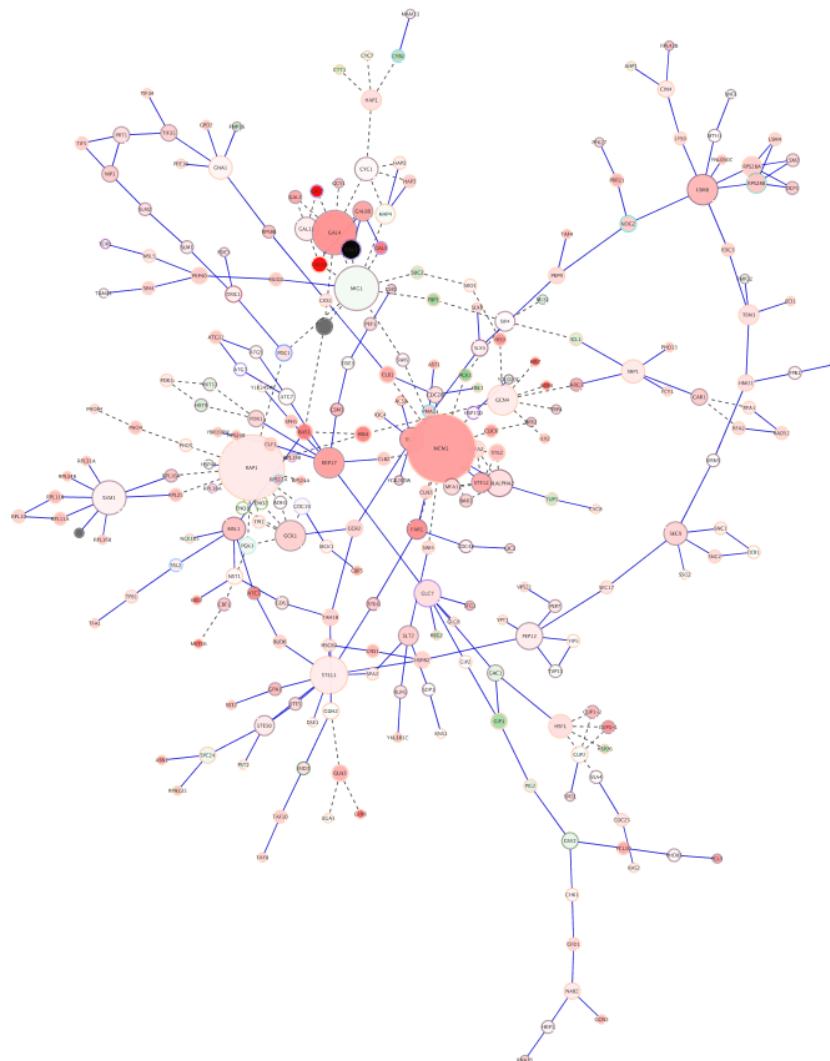
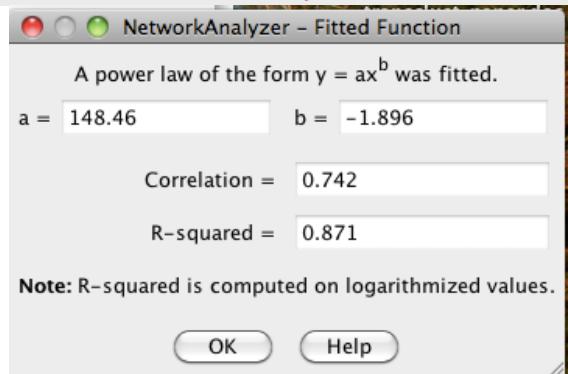
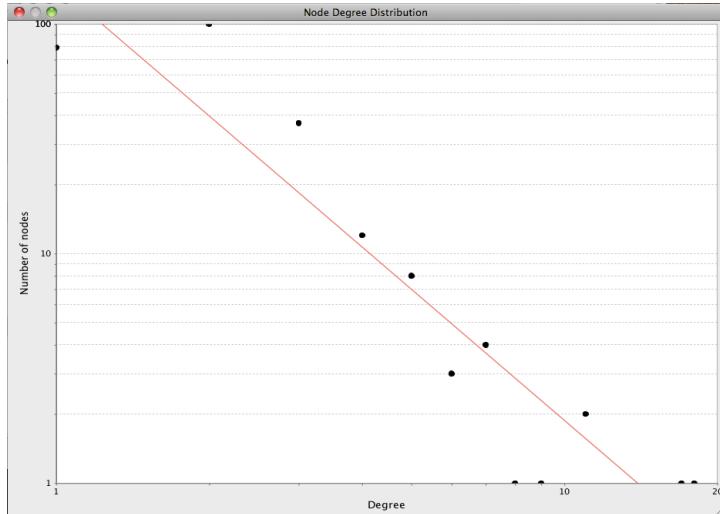


# Analytical Approaches





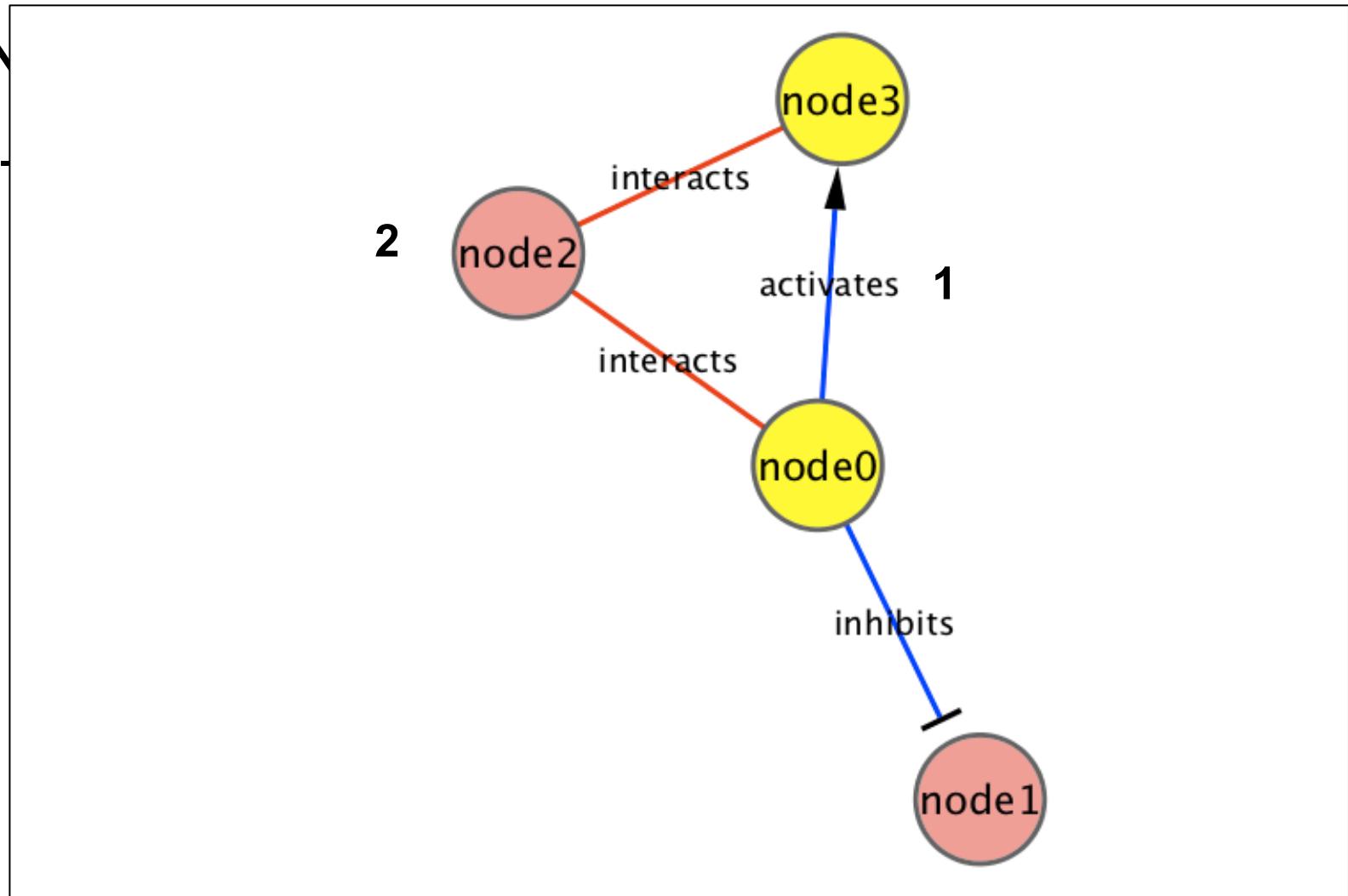
# Analytical Approaches





# Analytical Approaches

- 





# 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)





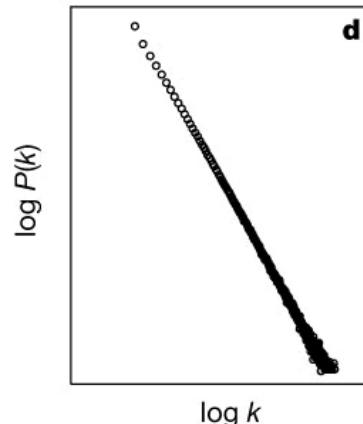
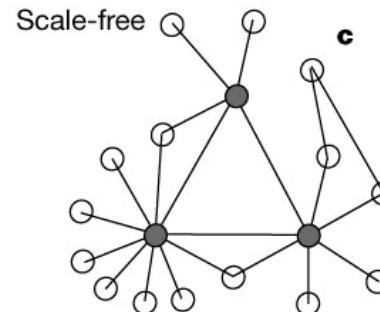
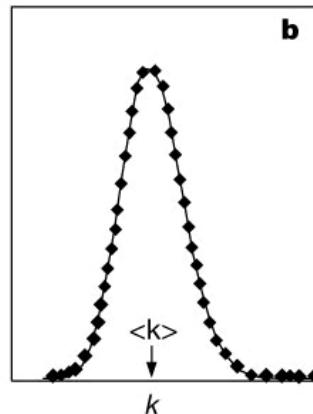
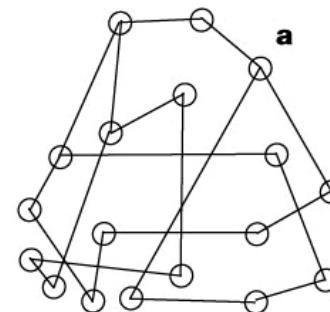
# Analytical Approaches

- Random networks
  - homogeneous, nodes have similar degrees, and not robust to arbitrary node failure

– Algorithms

- Flat random
- Scale-free
- Small-world

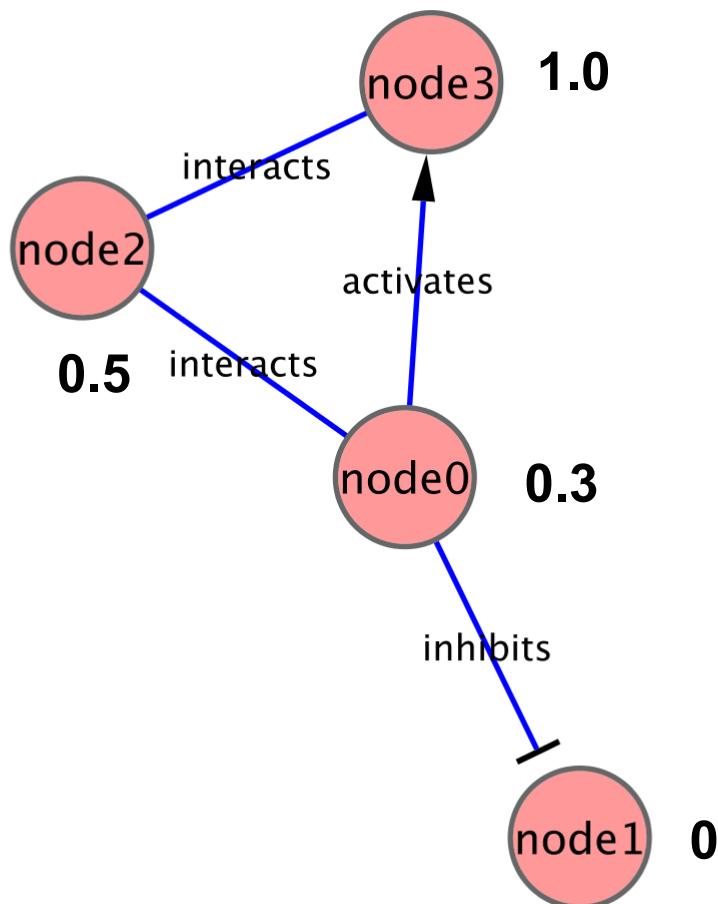
– Useful to understand a network





# Analytical Approaches

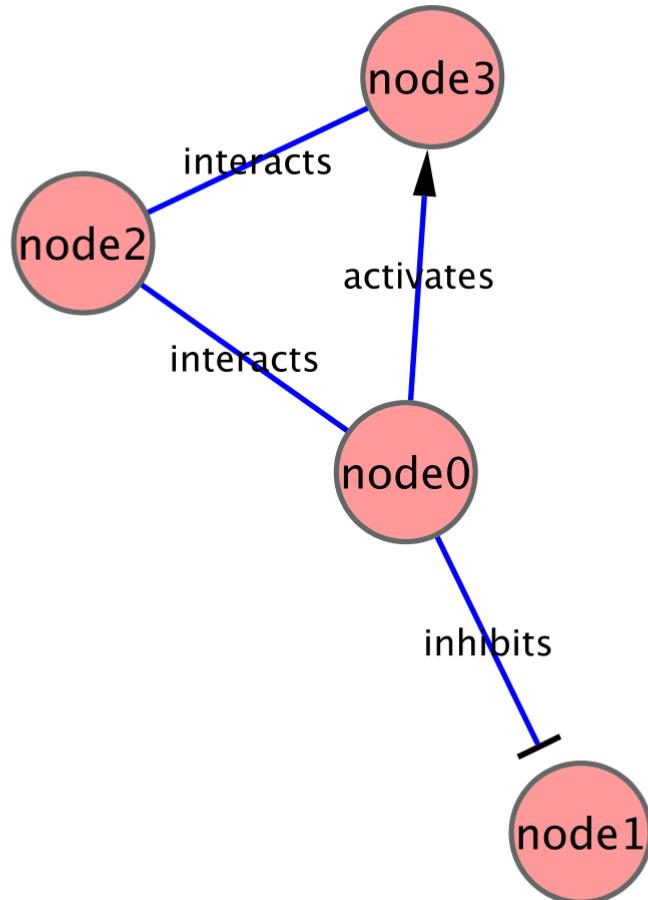
- N





# Analytical Approaches

- 



ID ▼	AverageShortestPathLength	BetweennessCentrality	ClosenessCentrality	ClusteringCoefficient	Degree
<b>node0</b>	1.0	0.66666667	1.0	0.33333333	3
<b>node1</b>	0.0	0.0	0.0	0.0	1
<b>node2</b>	1.33333333	0.66666667	0.75	0.5	2
<b>node3</b>	2.0	0.0	0.5	1.0	2



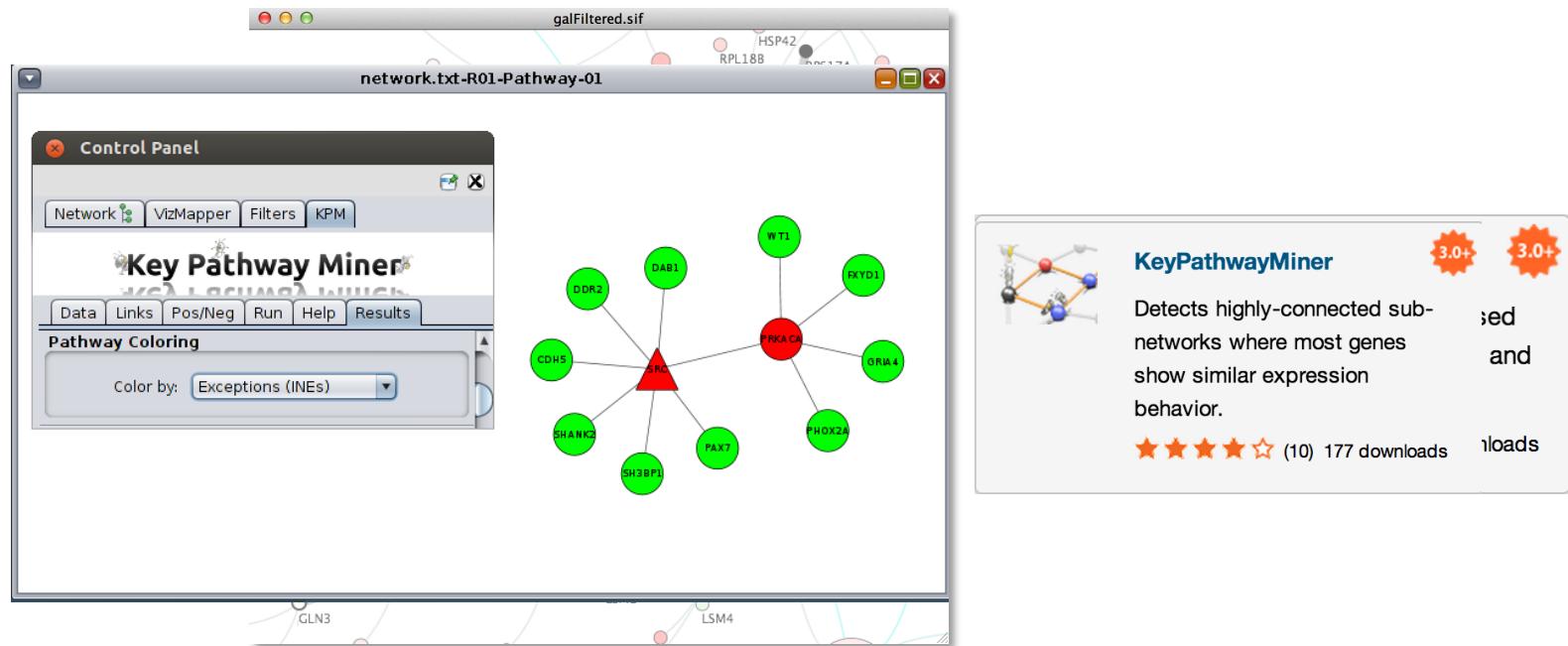
# Analytical Approaches

- Network Analyzer Demo...



# Analytical Approaches

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





# Analytical Approaches

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

The screenshot shows the ClueGO interface with a network graph. Nodes represent GO terms, and edges represent functional connections. Some prominent nodes include "RNA catabolic p", "mesoderm formation", "regulation of T", "The Co-Stimulatory Sign", "cellular defense response", and "chemotaxis". A tooltip for "RNA catabolic p" indicates it is a central node in the network.

**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

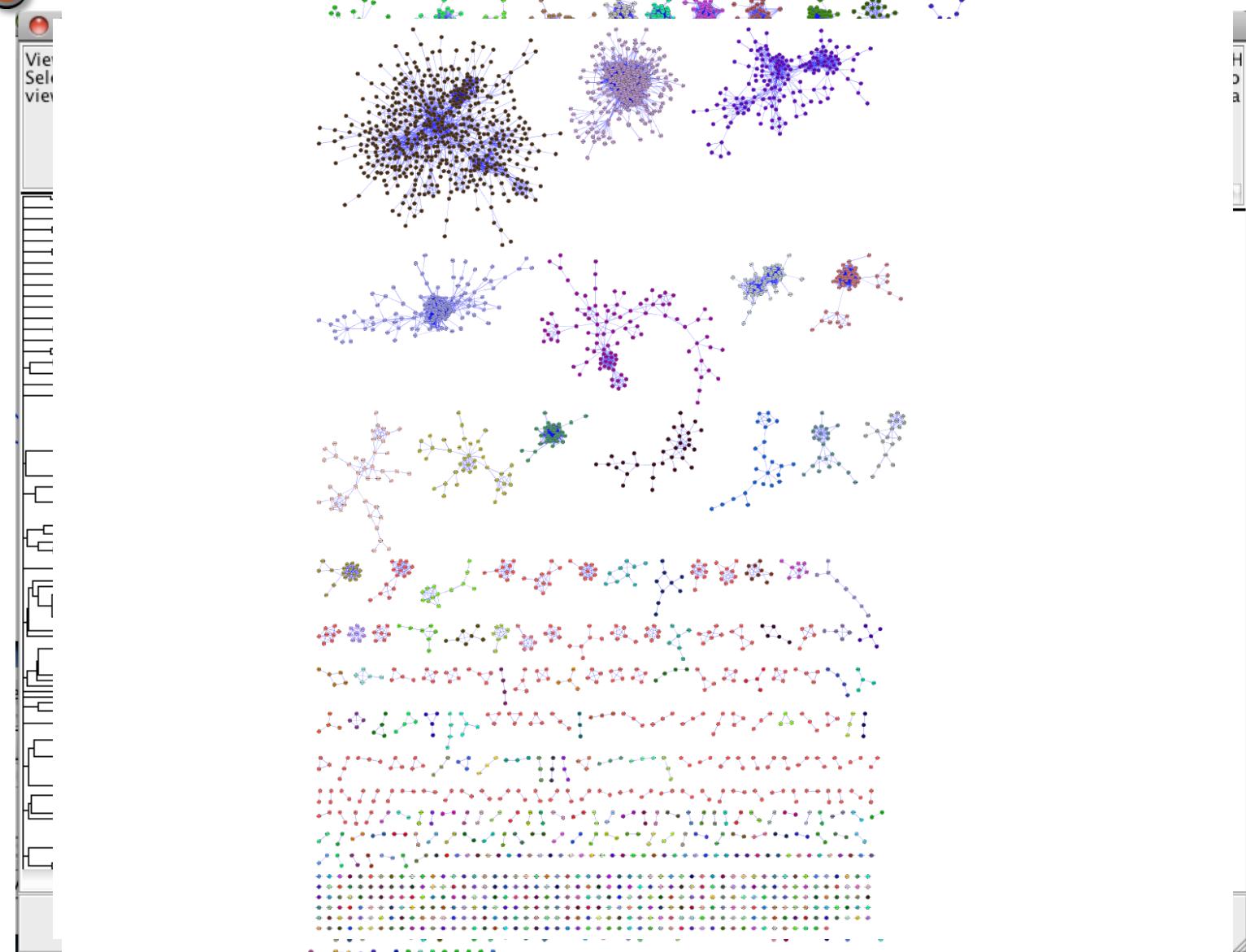
**NOA**  
The Network Ontology Analysis plugin performs ontology overrepresentation analysis based on the network connections among annotated nodes.

★★★★★ (7)

Project	Term PVal.	Term PVal.	Group PVal.	Group PValue	Associated Genes Found
Healthy_top200UpRegulated.txt	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, TLR1, 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-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]
	E-2	3.0 E-1	3.3 E-2	1.3 E-1	[SMG1, UPF3B]



# Analytical Approaches



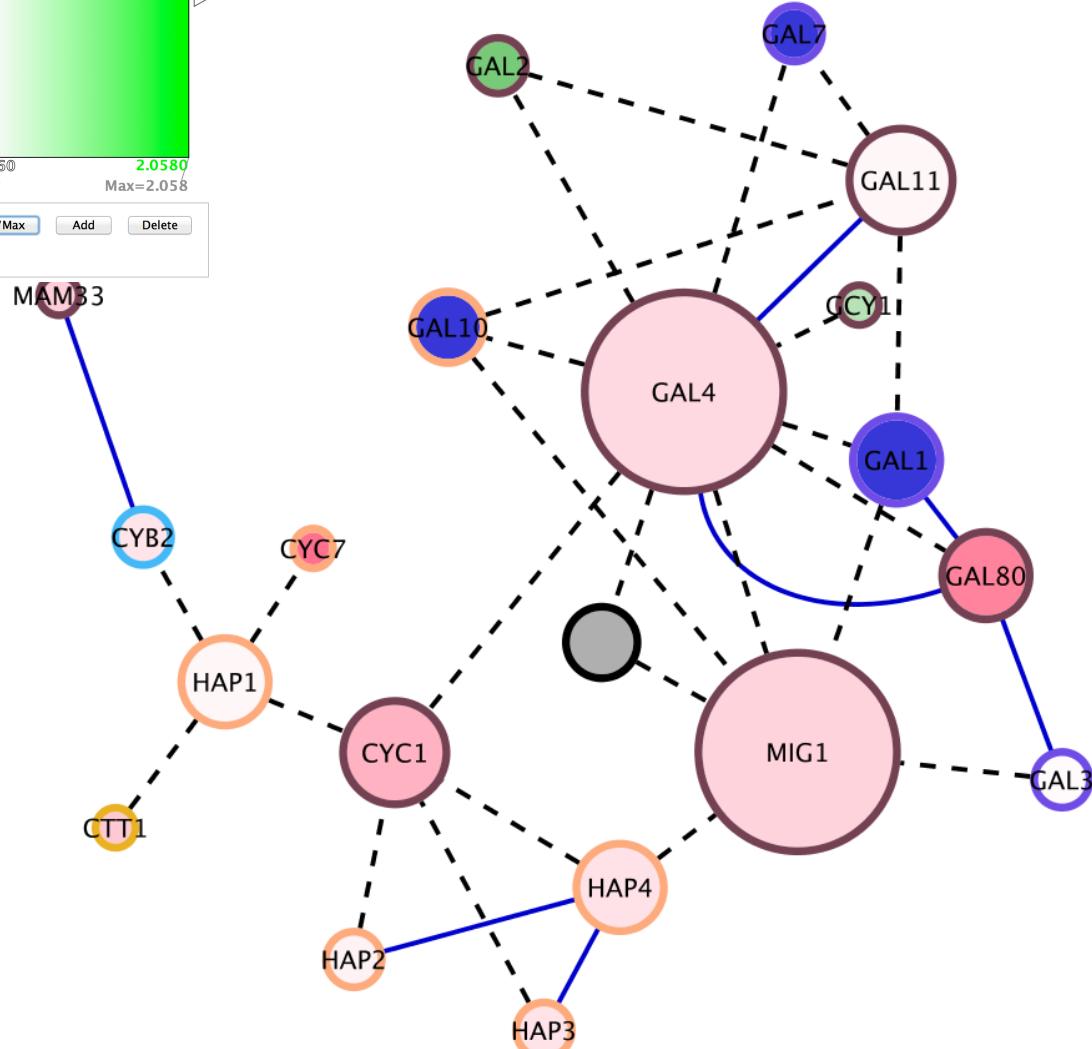
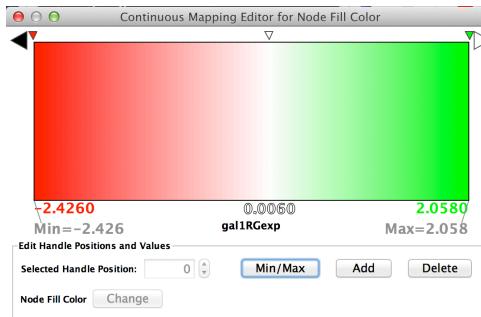


# Visualization of Biological Networks

- Data Mapping
- Layouts
- Animation



# Data Mapping



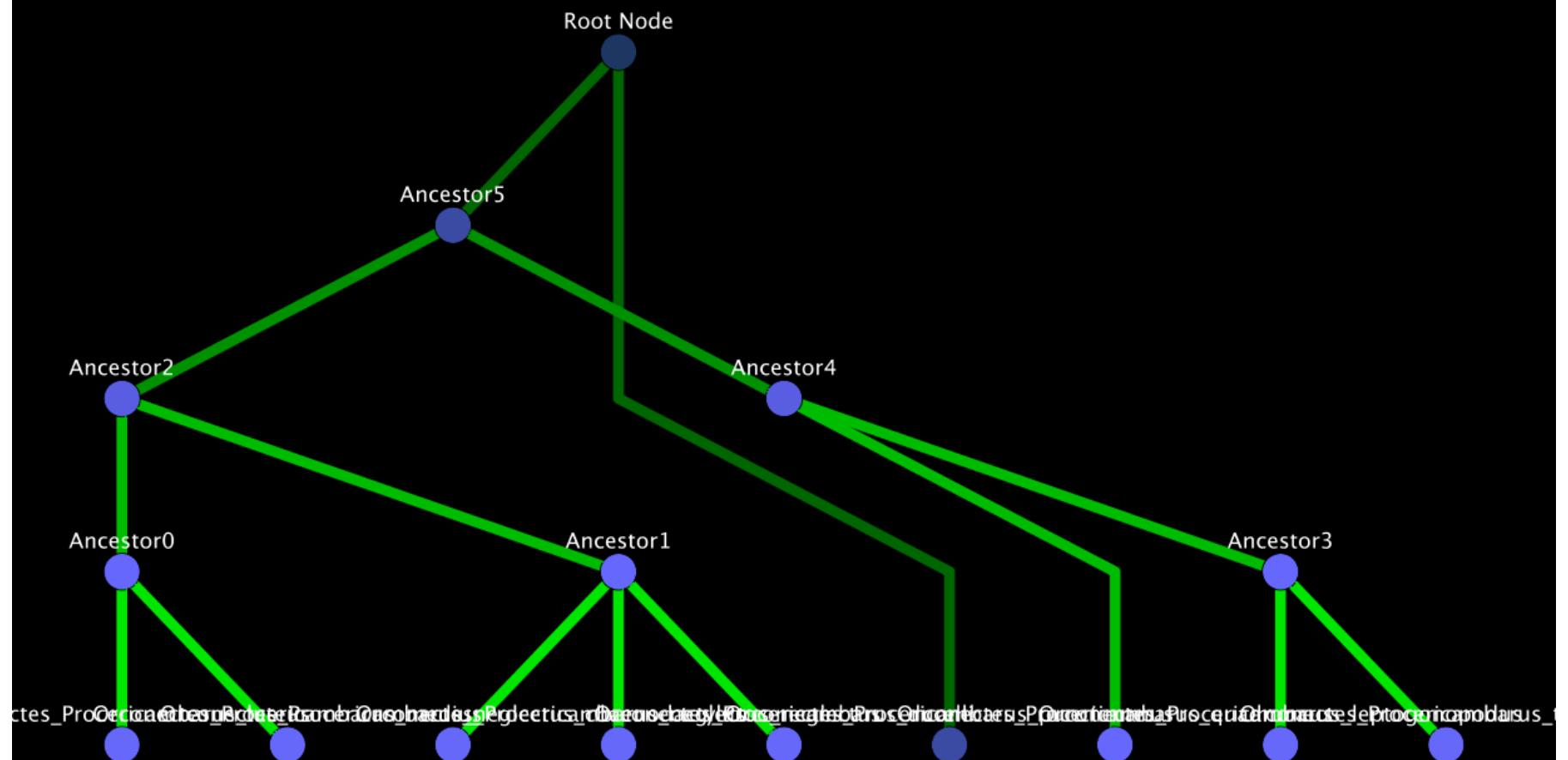


# Data Mapping

- Avoid cluttering your visualization with too much data
  - Map the data you are specifically interested in to call out meaningful differences
  - Mapping too much data to visual attributes may just confuse the viewer
  - Can always create multiple networks and map different values

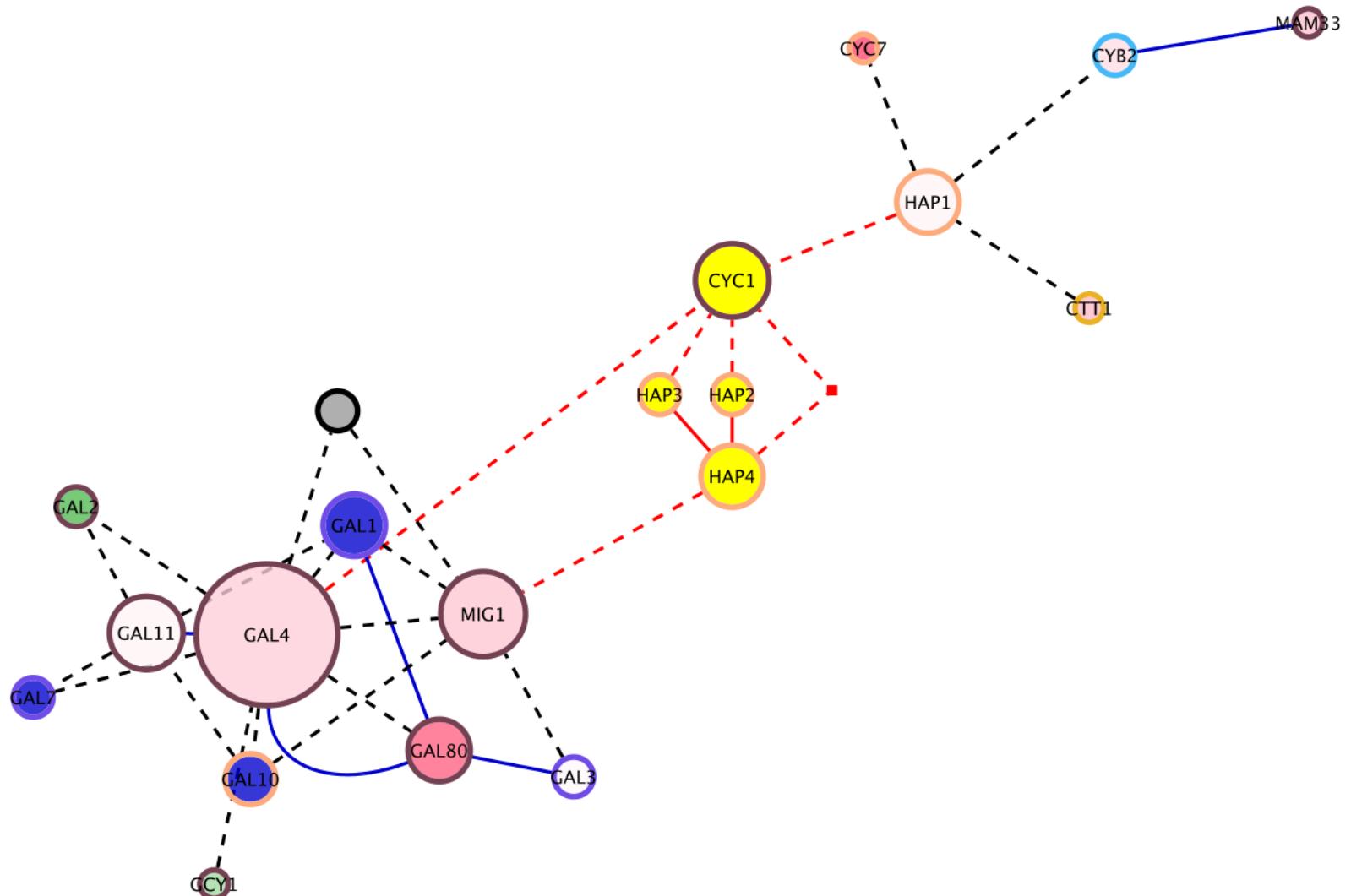


# Layouts





# Layouts



— Many, many others

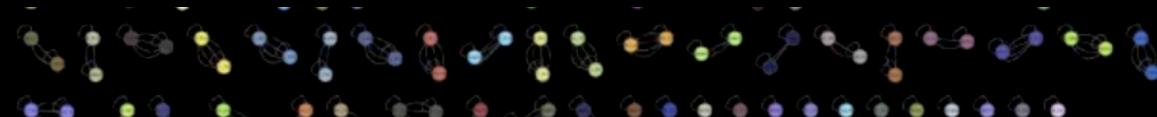
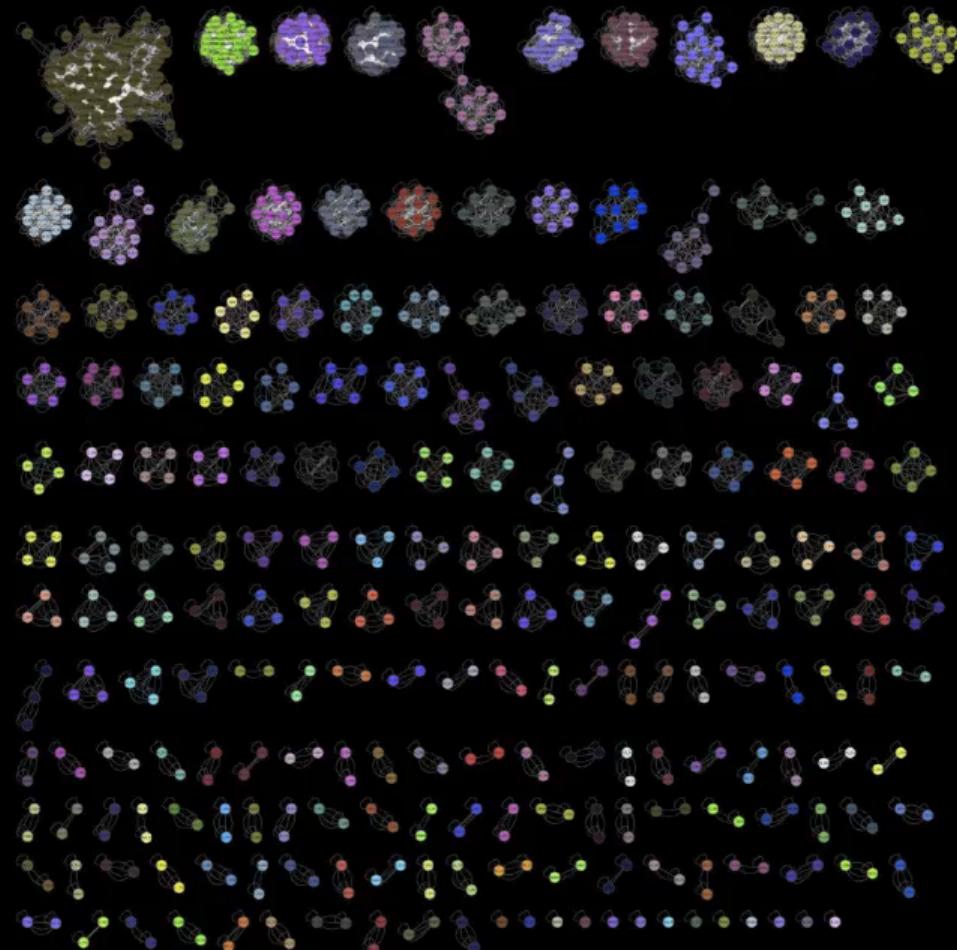


# Layouts

- Use layouts to convey the relationships between the nodes
- Layout algorithms may need to be “tuned” to fit your network
  - Layouts→Settings... menu
- Lots of parameters to change layout algorithm behavior
- Can also consider laying out portions of your network



# 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\_v3.0.0/sampleData/galFiltered.cys

Control Panel

Network Visual Styles Filters

Current Visual Style Sample for galFiltered

Add Visual Properties

Def. Map. Byp.

Column gal1RGexp Mapping Type Continuous Mapping

Current Mapping

Remove Visual Mapping

Height

Pm Label

Label Color

Label Font Size

Shape

Size

Column BetweennessCentrality Mapping Type Continuous Mapping

Current Mapping

Node Edge Network

Read Session File

galFiltered.sif

- 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

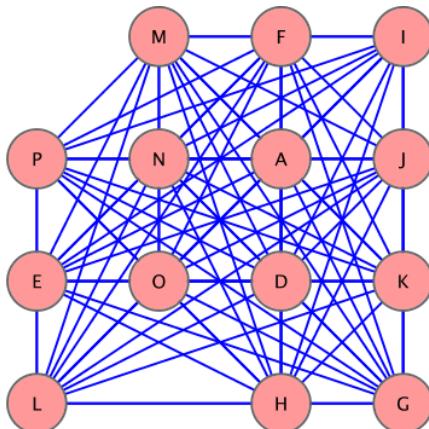
**GLADSTONE  
INSTITUTES**





# Core Concepts

- Networks and Tables



	A	B	C	D	E	F	G	H	I	J
1	GI_Number	Gene	HPRD_ID	REFSEQ	Entrez_ID	OMIM_ID	SWISSPROT	Gene_Description	Architecture	Arch
2	gi10923959	SDSL	11542	NP_012441.1	113675		SDSL		SP	Motif
3	gi14591926	SEC23B	11543	NP_116781.1	10483	Q15437	Protein transport protein SEC23B	GEL	Dom	
4	gi14602492	SCMH1	11540	AAH0752.1	22955	Q96GD3	SCMH1	SAM	Dom	
5	gi21361625	EXOC2	11544	NP_060773.3	55770	Q96KP1	SEC5 like 1	IPT	Dom	
6	gi17998551	SERPINB12	11548	NP_536722.1	89777	Q96P63	Serpin B12	SERPIN	Dom	
7	gi65287717	EIF2AK4	18447	NP_00101372	440275	609280	Similar to GCN2 eIF2alpha kinase	S_T_Y_Kinase	Dom	
8	gi4507755	TYROBP	4996	NP_003323.1	7305	604142	043914	DAP12	ITAM	Dom
9	gi5803086	IL24	4995	NP_006841.1	11009	604136	Q13007	Interleukin 24	IL10	Dom
10	gi21265034	ADAMTS13	4994	NP_620594.1	11093	604134	ADAMTS 13	TSP1	Dom	
11	gi1743873	PIP5K1A	4470	AAC50911.1	8394	603275	Q99755	Phosphatidylinositol-4-phosphate 5 PIPKc	Dom	
12	gi1459382	SVIL	4992	NP_00101006.1	6840	604135	Q99425	Supeverillin	VHP	Dom
13	gi1633955	CHLT2B1	4991	NP_014401.1	6002	604132	Q99426	Chlorotranferase family 2B, member 1	Dom	
14	gi21040399	RBBP8	4990	AAH30590.1	5932	604124	Q99708	CTIP	LZ	Motif
15	gi4758378	FIGF	2102	NP_004460.1	2277	300091	Q43915	VEGF D	PDGF	Dom
16	gi11321617	DPYSL4	7463	NP_006417.1	10570	608407	Q14531	Collapsin response mediator protein 3		
17	gi8923202	TASP1	7460	NP_060184.1	55617	608270	Q9H6P5	Threonine aseptase 1		
18	gi5454090	SSR4	2101	NP_006271.1	6748	300090	P51571	Signal sequence receptor delta	SP	Motif
19	gi5730045	SLC16A2	2108	NP_006508.1	6567	300095	P36021	X linked PEST containing transport	TM	Dom
20	gi11968027	FTS	7467	NP_071921.1	64400	608483	Q9H8T0	FTS	UBC	Dom
21	gi48255885	PRKC1	2109	NP_002731.3	5584	600539	P41743	Protein kinase C, iota type	S_T_Kinase	Dom
22	gi9104000	RPA4	6593	NP_074791.1	29335		RPA4		TRNA	
23	gi3376312	PPP1R2P9	6593	NP_074791.1	80316			Type 1 protein phosphatase inhibitor		
24	gi15826862		6595	NP_296379.1	9000			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	Dom

## Networks

e.g., PPIs or pathways

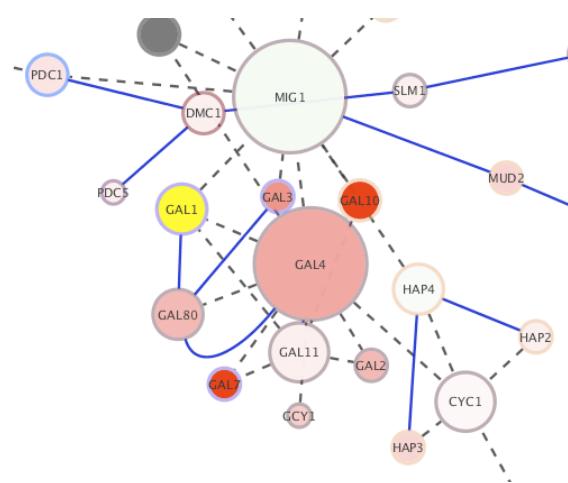
## Tables

e.g., data or annotations



# Core Concepts

- Networks and Tables



	A	B	C	D	E	F	G	H	I
	GI_Number	Gene	HPRD_ID	REFSEQ	Entrez_ID	OMIM_ID	SWISSPROT_Gene_Description	Architecture	Arch
1	gi10923959	SDSL	11542	NP_012441.1	113675		Protein transport protein SEC23B	SP	Motif
2	gi1459126	SEC23B	11543	NP_116781.1	10483	Q15437	Protein transport protein SEC23B	GEL	Dom
3	gi14620492	SCMH1	11540	AAH09752.1	22955	Q96GD3	SCMH1	SAM	Dom
4	gi21361625	EXOC2	11544	NP_060773.3	55770	Q96KP1	SEC5 like 1	IPT	Dom
5	gi17998551	SERPINB12	11548	NP_536722.1	89777	Q96P63	Serpin B12	SERPIN	Dom
6	gi65287717	EIF2AK4	18447	NP_00101372	440275	609280	Similar to GCN2 eIF2alpha kinase	S_T_Y_Kinase	Dom
7	gi4507755	TYROBP	4996	NP_003323.1	7305	604142	043914 DAP12	ITAM	Dom
8	gi5803086	IL24	4995	NP_006841.1	11009	604136	Q13007 Interleukin 24	IL10	Dom
9	gi21265034	ADAMTS13	4994	NP_620594.1	11093	604134	ADAMTS 13	TSP1	Dom
10	gi1743873	PIP5K1A	4470	AAC50911.1	8394	603275	Q99755 Phosphatidylinositol-4-phosphate 5 PIPKc	VHP	Dom
11	gi145982	SVD	4992	NP_001006.1	68440	604149	Q95425 Superfamily	VHP	Dom
12	gi163395	MLT2B1	4991	NP_014440.1	6002	604123	MLT2B1	MLT2B1	Dom
13	gi21040399	RBBP8	4990	AAH30590.1	5932	604124	Q99708 CTIP	LZ	Motif
14	gi4758378	FIGF	2102	NP_004460.1	2277	300091	043915 VEGF D	PDGF	Dom
15	gi11321617	DPYSL4	7463	NP_006417.1	10570	608407	Q14531 Collapsin response mediator protein 3		
16	gi8923202	TASP1	7460	NP_060184.1	55617	608270	Q9H6P5 Threonine aminopeptidase 1		
17	gi5454090	SSR4	2101	NP_006271.1	6748	300090	P51571 Signal sequence receptor delta	SP	Motif
18	gi5730045	SLC16A2	2108	NP_006508.1	6567	300095	P36021 X linked PEST containing transport	TM	Dom
19	gi1968027	FTS	7467	NP_071921.1	64400	608483	Q9H8T0 FTS	UBC	Dom
20	gi48255885	PRKC1	2109	NP_002731.3	5584	600539	P41743 Protein kinase C, iota type	S_T_Y_Kinase	Dom
21	gi919740	RPA4	6592	NP_0747479.1	29335	604125	Q9H8T1 RPA4	TRNA	Dom
22	gi3376312	PPP1R2P9	6593	NP_0747480.1	80316		TP53BP1 protein phosphatase inhibitor		
23	gi15826862		6595	NP_296379.1	9000		JM11 protein	CC	Motif
24	gi7661844	CCDC22	6594	NP_294727.1	28952		JM1 protein	CC	Motif
25	gi6005794	PRAF2	6596	NP_009144.1	11230		JM4 protein	TM	Dom
26									

Networks

Tables

Visual Styles

# Wall of Apps 173 total





# Cytoscape

- Common use cases
  - Visualizing:
    - Protein-protein interactions
    - Pathways
  - Integrating:
    - Expression profiles
    - Other state data
  - Analyzing:
    - Network properties
    - Data mapped onto network



# Loading Networks

Cytoscape File Edit View Select Layout Apps Tools Help

Session: New Session

Control Panel

Network

Network No... Ed...

Yeast Network Sheet 1 Yeast Network Sheet 33... 36...

Yeast Network Sheet 1

Table Panel

Yeast Network Sheet 1

shared...	name	isExclu...

Node Table Edge Table Network Table

Loading network from table

Memory: OK



# Loading Networks

Cytoscape

Session: New Session

Control Panel

Network

Welcome to Cytoscape

Start New Session

- New/Empty Network
- From Network File...
- From Network Database...

Open Recent Session

- file:/Applications/Cytoscape\_.../galFiltered.cys
- file:/Users/apico/Work/Projects/.../apr2013....
- Open file...

From Preset Network

- C. Elegans
- Human
- Arabidopsis
- Mouse
- Fly
- Yeast

Load Preset Network

News and Links

Web Links

- [About Cytoscape](#)
- [Documentation](#)
- [Tutorials](#)
- [Report a bug](#)

Latest News

- Cytoscape 3.1.0-SNAPSHOT is up to date.

Don't show again

**Close**

Node Table | Edge Table | Network Table

Searching PSICQUIC Services

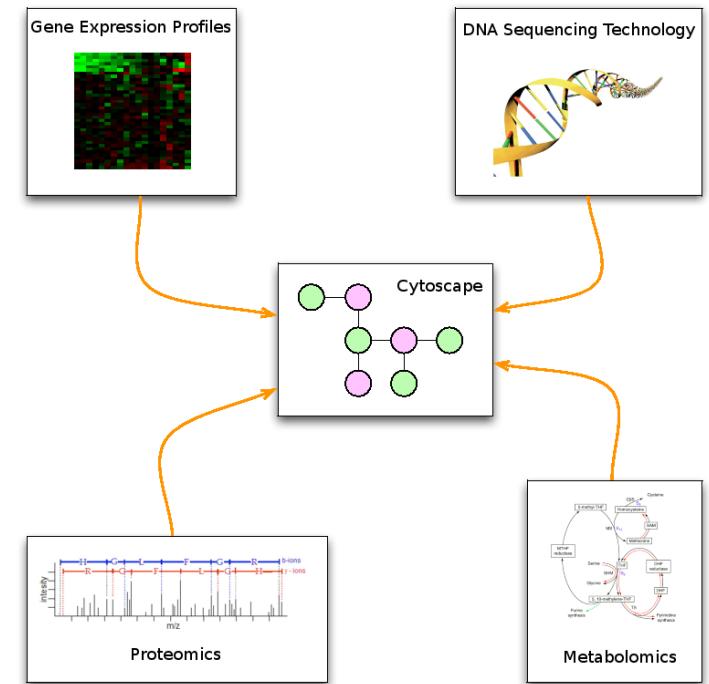
Memory: OK

A screenshot of the Cytoscape application interface. The main window title is "Cytoscape" and the session is "New Session". The toolbar has various icons for file operations like Open, Save, and Import/Export. The left sidebar is the "Control Panel" with a "Network" tab selected. A central dialog box titled "Welcome to Cytoscape" is open, containing three main sections: "Start New Session" (with "New/Empty Network", "From Network File...", and "From Network Database..." options), "Open Recent Session" (listing recent files), and "From Preset Network" (with radio buttons for C. Elegans, Human, Arabidopsis, Mouse, Fly, and Yeast). Below these is a "News and Links" section with links to "About Cytoscape", "Documentation", "Tutorials", and "Report a bug", and a "Latest News" section indicating "Cytoscape 3.1.0-SNAPSHOT is up to date". At the bottom of the dialog are "Don't show again" and "Close" buttons. The footer of the application shows tabs for "Node Table", "Edge Table", and "Network Table", and a status bar at the bottom with "Searching PSICQUIC Services" and "Memory: OK". A large question mark icon is overlaid on the right side of the window.



# 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, Logical,  
Lists...





# Loading Tables

Cytoscape File Edit View Select Layout Apps Tools Help Session: New Session

Control Panel Network Yeast Network Sheet 1 Yeast Network Sheet 33... 36...

Yeast Network Sheet 1

Table Panel

Change Table Mode

shared...	name	isExclu...	COMM...	gal1R...	gal4R...	gal80R...	gal1R...	gal41...
YGL229C	YGL229C	false	SAP4	-0.521	0.171	-0.208	2.3655E...	0.0208
YBR217W	YBR217W	false	APG12	0.088	-0.332	0.378	0.143	0.0110
YBR160W	YBR160W	false	CDC28	-0.016	-0.087	-0.405	0.7432	0.3210
YNL216W	YNL216W	false	RAP1	0.205	0.015	0.234	0.024692	0.9349
YLR249W	YLR249W	false	YEF3	-0.39	-0.394	-0.769	2.713E-8	0.0474
YGL202W	YGL202W	false	ARO8	-0.305	-0.286	-0.536	4.2677E...	4.2373
YML123C	YML123C	false	PHO84	0.283	-0.114	0.692	0.00151...	0.0194
YBR018C	YBR018C	false	GAL7	0.153	-1.995	3.126	7.8855E...	3.6284
YIL105C	YIL105C	false	YIL105C	0.106	0.022	0.107	6.0824E...	0.6670

Node Table Edge Table Network Table

Loading table data Memory: OK



# Visual Style Manager

Cytoscape File Edit View Select Layout Apps Tools Help

Session: /Applications/Cytoscape\_v3.0.0/sampleData/galFiltered.cys

Control Panel

Network Visual Styles Filters

Current Visual Style

Source Target

Big Labels

default

Directed

Minimal

Nested Network Style

Ripple

Sample1

Source Target

Universe

Node Edge Network

Search term...

galFiltered.sif



# Selection Filters

- C

Session: C:\Users\Administrator\Desktop\cytoscape-windows-3.1.1\sampleData\galFiltered.cys

File Edit View Select Layout Apps Tools Help

Control Panel

Network Sets jActiveMNodes Style Select

Default filter

Nodes with at least 5 neighbours  
within distance 1

+ [ ]

Apply Automatically

Apply [ ]

Selected 22 nodes and 0 edges in 9ms

Filter Chain

Table Panel

f(x) galFiltered.sif

age...	Clusteri...	Closen...	IsSingl...	Partner...	SelfLoo...	Eccentri...	Stress	Degree	Betwee...
30...	0.0	0.08982...	false	0	0	24	47386	7	0.12017...
29...	0.0	0.14735...	false	0	0	17	12160	5	0.12134...
74...	0.05882...	0.17115...	false	0	0	15	86646	18	0.52645...
09...	0.17777...	0.15066...	false	1	0	15	10716	11	0.05907...
66...	0.3	0.85714...	false	0	0	2	20	5	0.5
58...	0.0	0.11665...	false	0	0	21	71812	5	0.15821...

Node Table Edge Table Network Table

Memory: OK [ ]



# Saving and Exporting

- Sessions save pretty much everything:  
Networks, Properties, Visual styles, Screen sizes
- 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

**[cytoscape-helpdesk@googlegroups.com](mailto:cytoscape-helpdesk@googlegroups.com)**



# 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



# Hands-on Tutorial

Introduction to Cytoscape:  
Networks, Data, Styles, Layouts and App Manager

<http://tutorials.cytoscape.org>



# Examples/Demos

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



# Expression Data Analysis

– CLICK OK

**Hierarchical cluster**

**Linkage**: pairwise average-linkage  
**Distance Metric**: Euclidean distance

**Array sources**

**Node attributes for cluster**

- SelfLoops
- TopologicalCoefficient
- gal1RGexp
- gal1RGsig
- gal4RGexp
- gal4RGsig
- gal80Rexp
- 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** ►

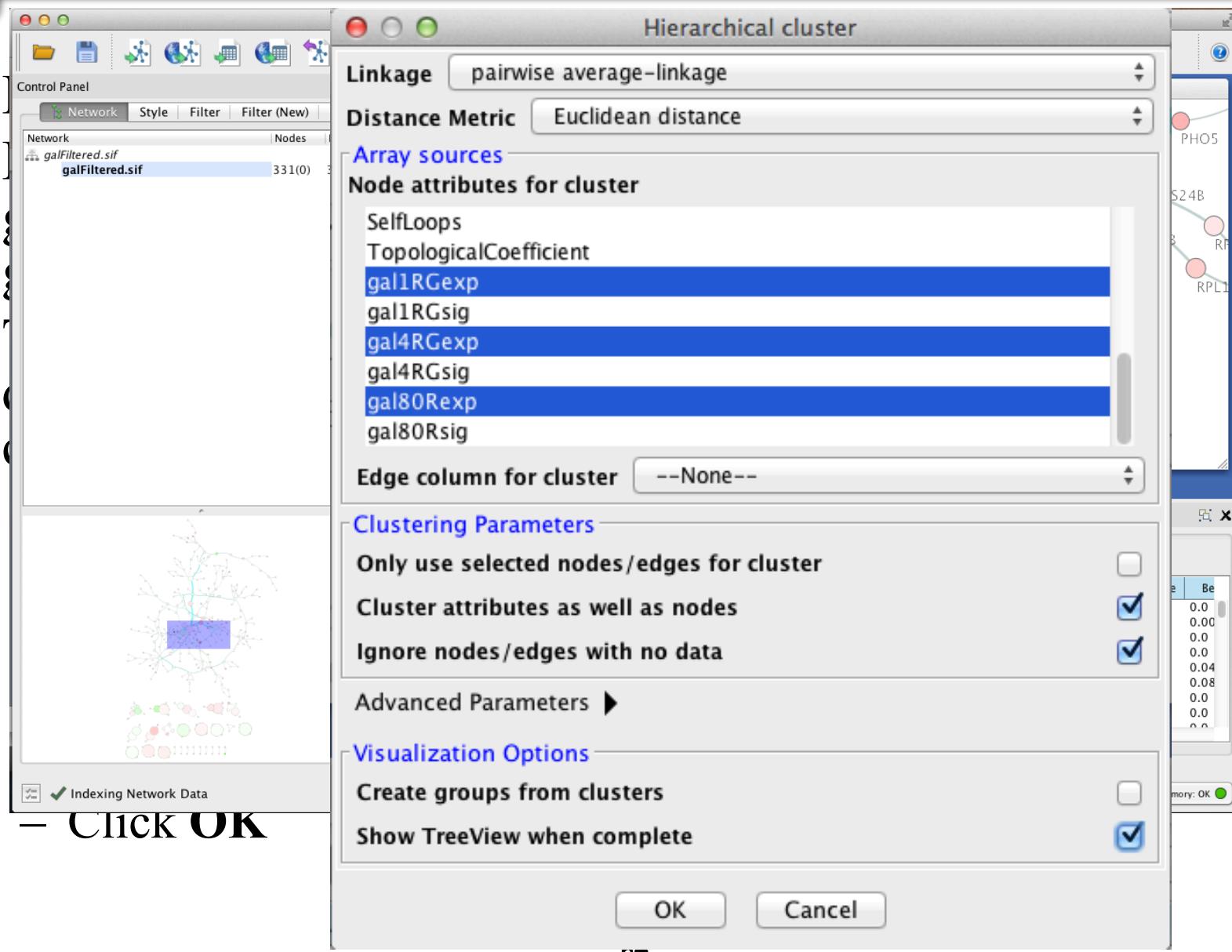
**Visualization Options**

**Create groups from clusters**:

**Show TreeView when complete**:

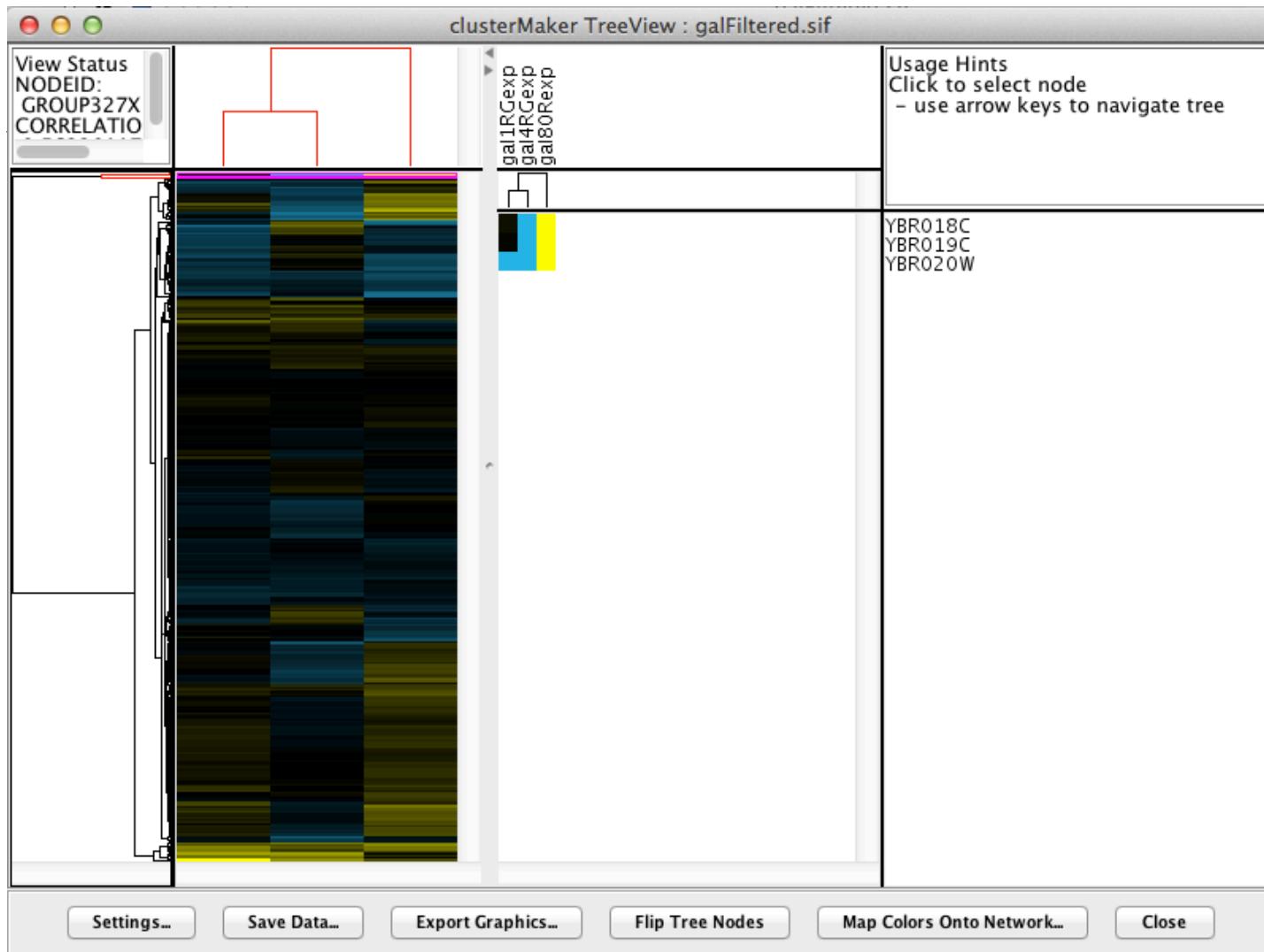
**OK**   **Cancel**

Memory: OK





# Expression Data Analysis





# Expression Data Analysis

BiNGO Settings  
Session: /Users/scooter/Documents/galFiltered3.cys

Control Panel

Network Style Filter Filter (New) Sets

Network Nodes Edges

galFiltered.sif 331(3) 362(0)

cluster1 cluster1 72(0) 127(0)

BiNGO output

cluster1

GO\_Saccharomyces cerevisiae,default,bingo,namespace close

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 ... YBR020W YBR018C YBR019C
19388	galactose catabolic process	5.0181...	1.9069...	3/3	100.0%	6/6208 ... 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
11787	small molecule catabolic process	1.2782...	0.1771...	2/2	100.0%	148/6208 ... YBR020W YBR018C YBR019C

Select All Unselect All Select nodes

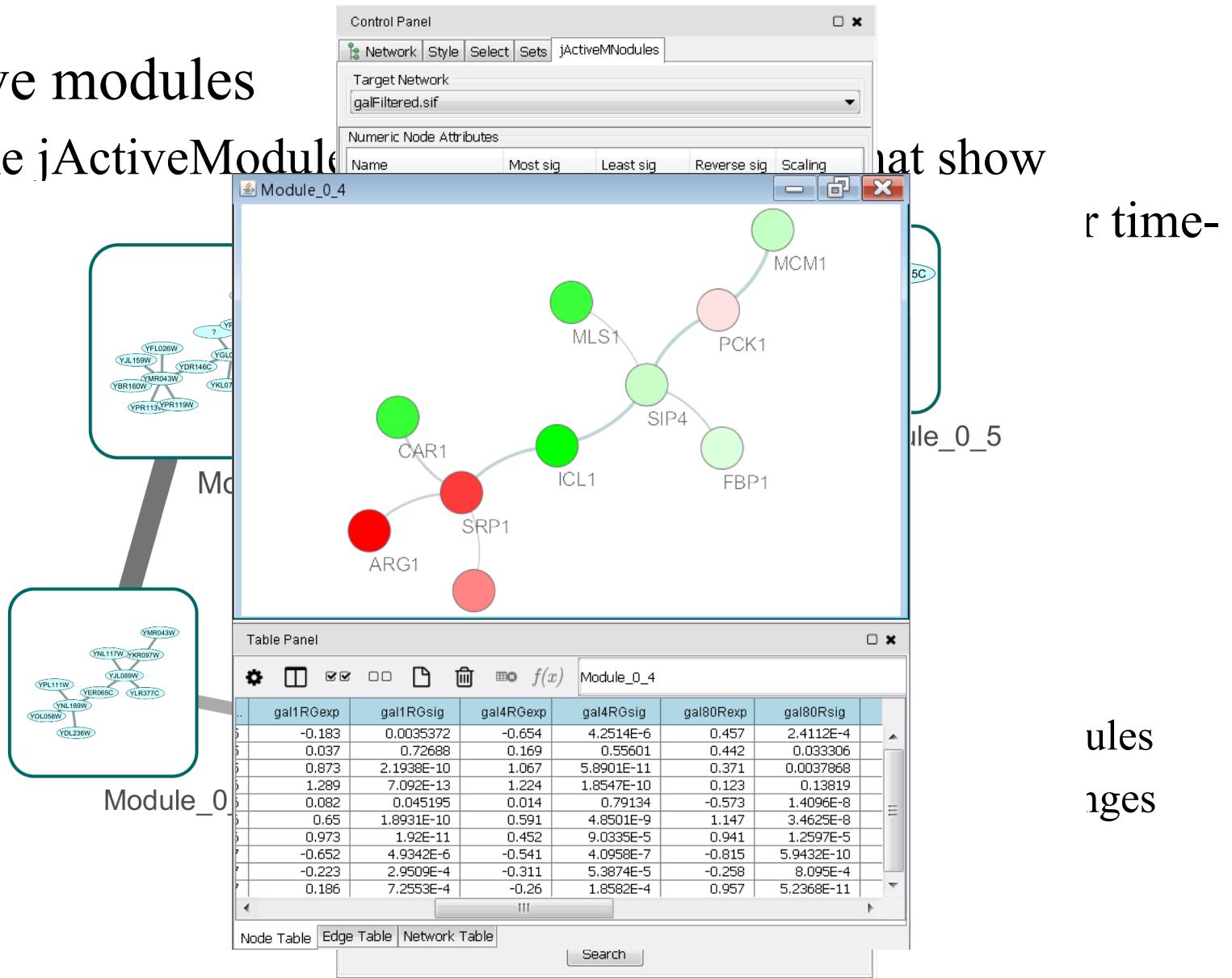
Start BiNGO



# Expression Data Analysis

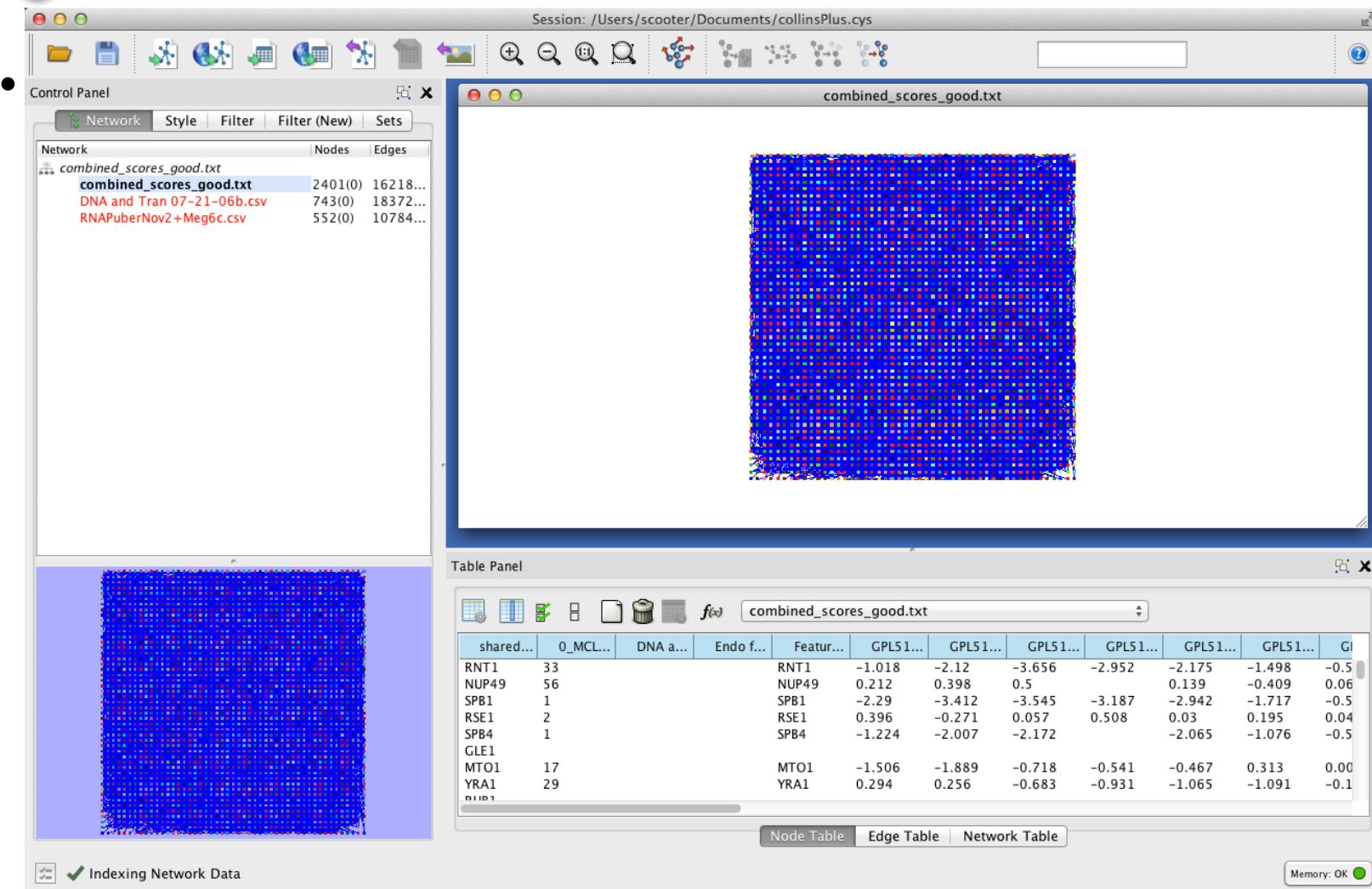
- Active modules

- The jActiveModule tool



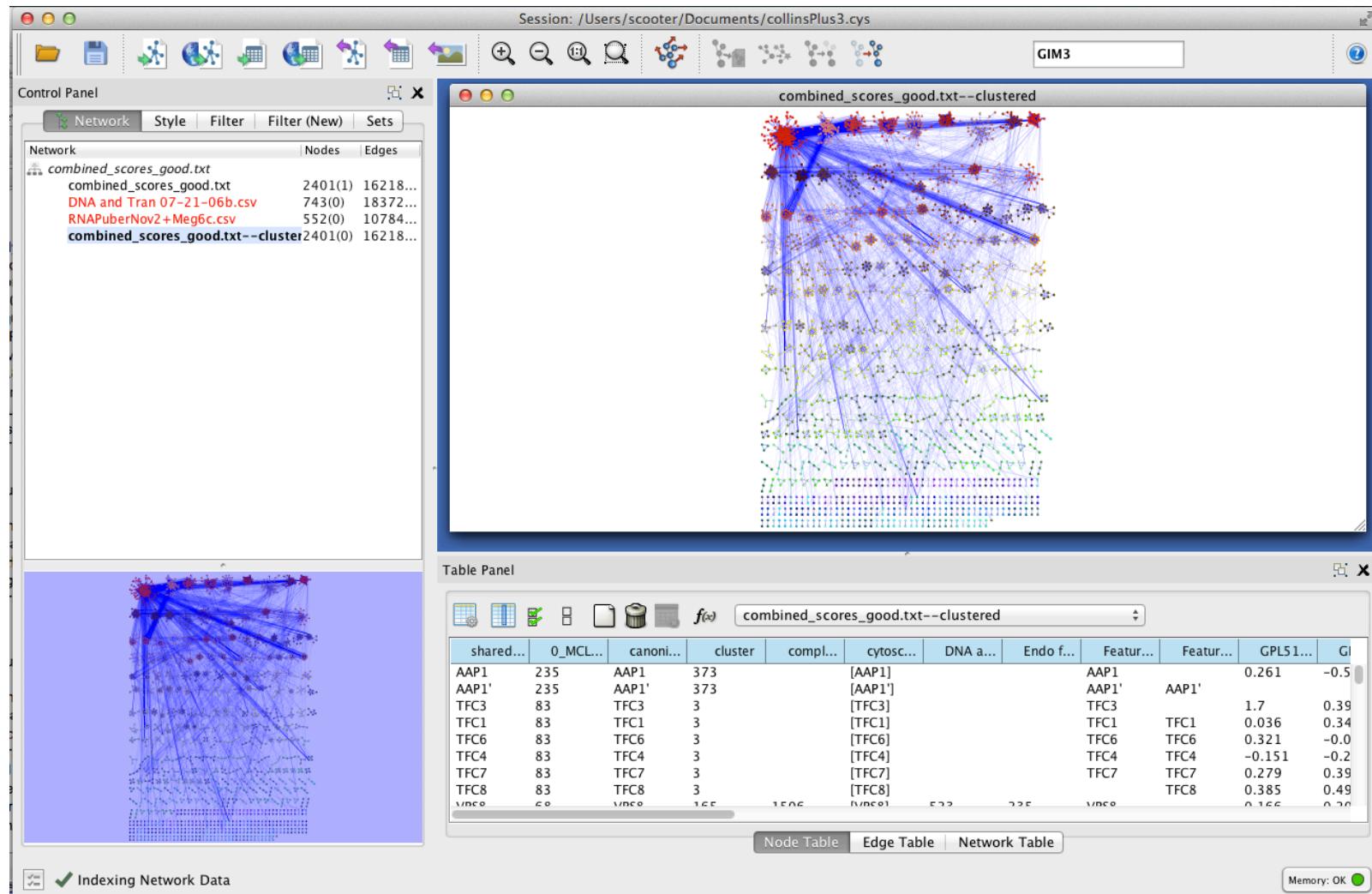


# Protein Complexes





# Protein Complexes



ges

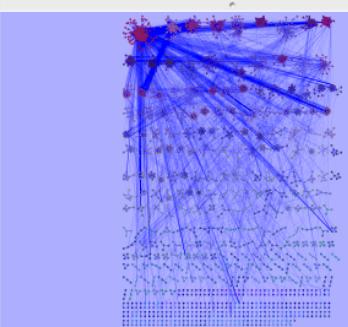


# Protein Complexes

Control Panel

Network Nodes

- combined\_scores\_good.txt
- combined\_scores\_good.txt 2401
- DNA and Tran 07-21-06b.csv 743(0)
- RNAPuberNov2+Meg6c.csv 552(0)
- combined\_scores\_good.txt--cluster2401



**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

OK Cancel

Node Table Edge Table Network Table

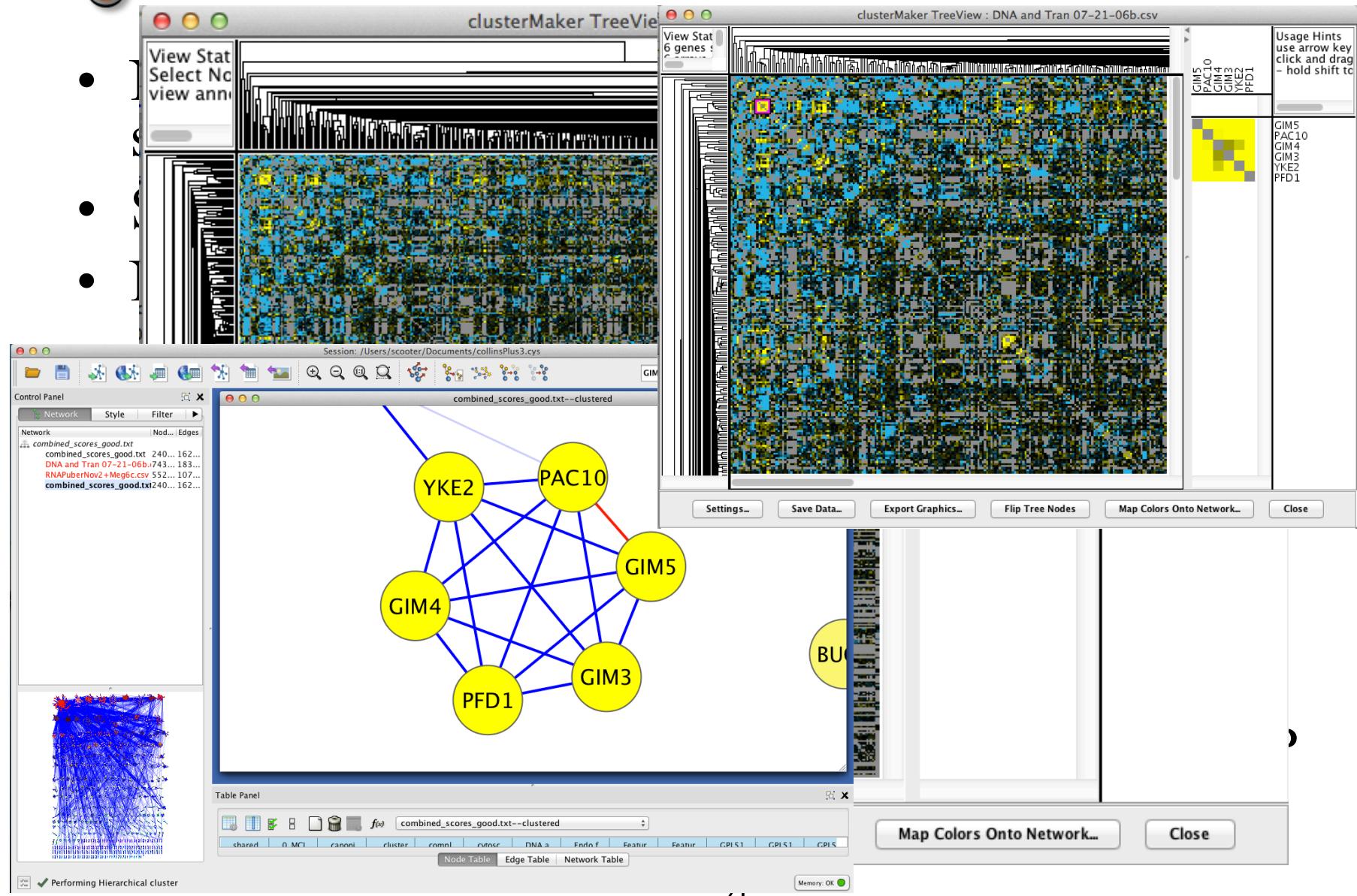
Indexing Network Data

Memory: OK

Feature	Feature	GPL51...	GI
AP1	AAP1'	0.261	-0.5
AP1'	AAP1'	1.7	0.39
FC3	TFC1	0.036	0.34
FC1	TFC6	0.321	-0.0
FC6	TFC4	-0.151	-0.2
FC4	TFC7	0.279	0.39
FC7	TFC8	0.385	0.49
FC8	TFC9	0.166	0.20

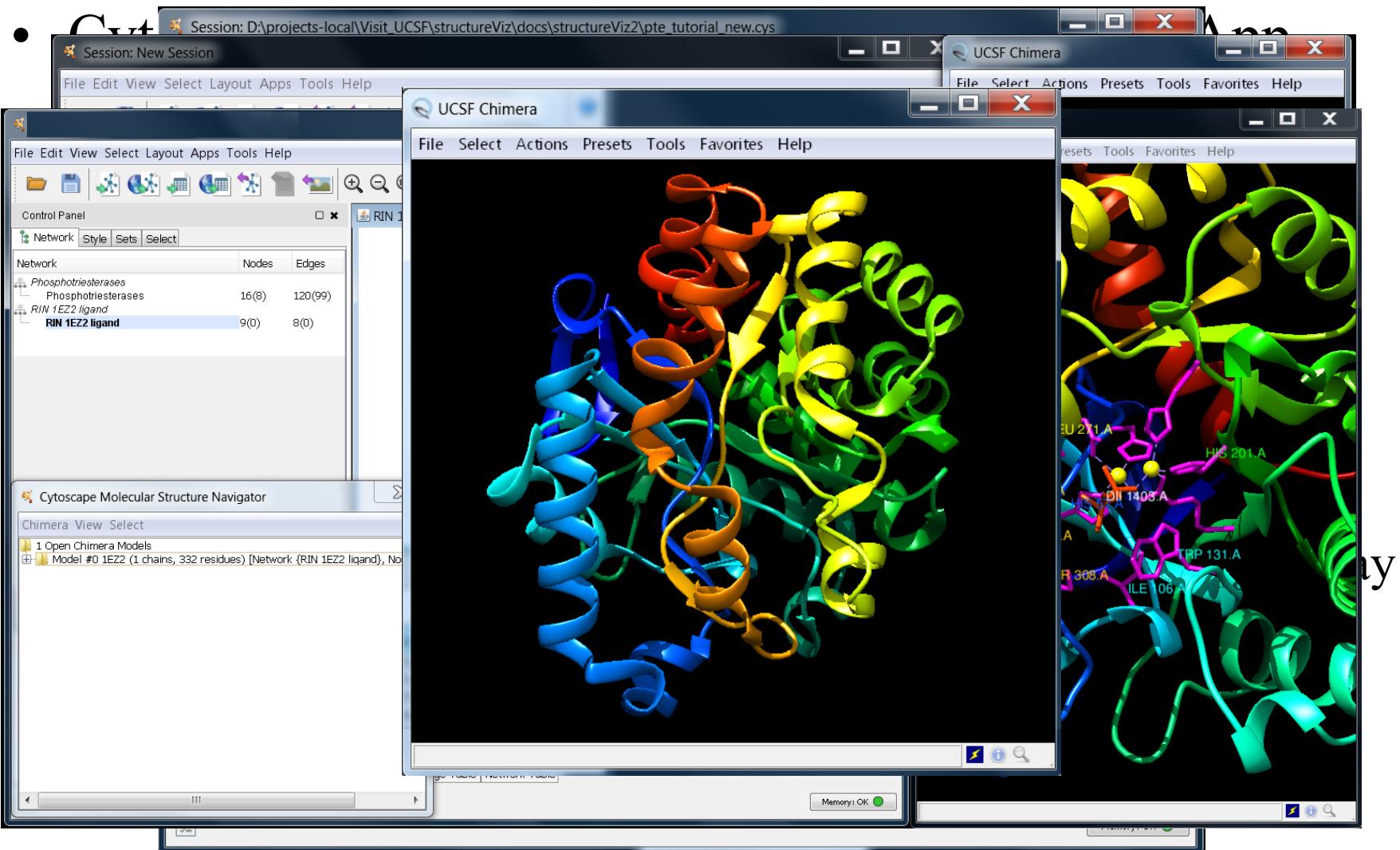


# Protein Complexes





# From Networks to Structures





# Cytoscape Roadmap

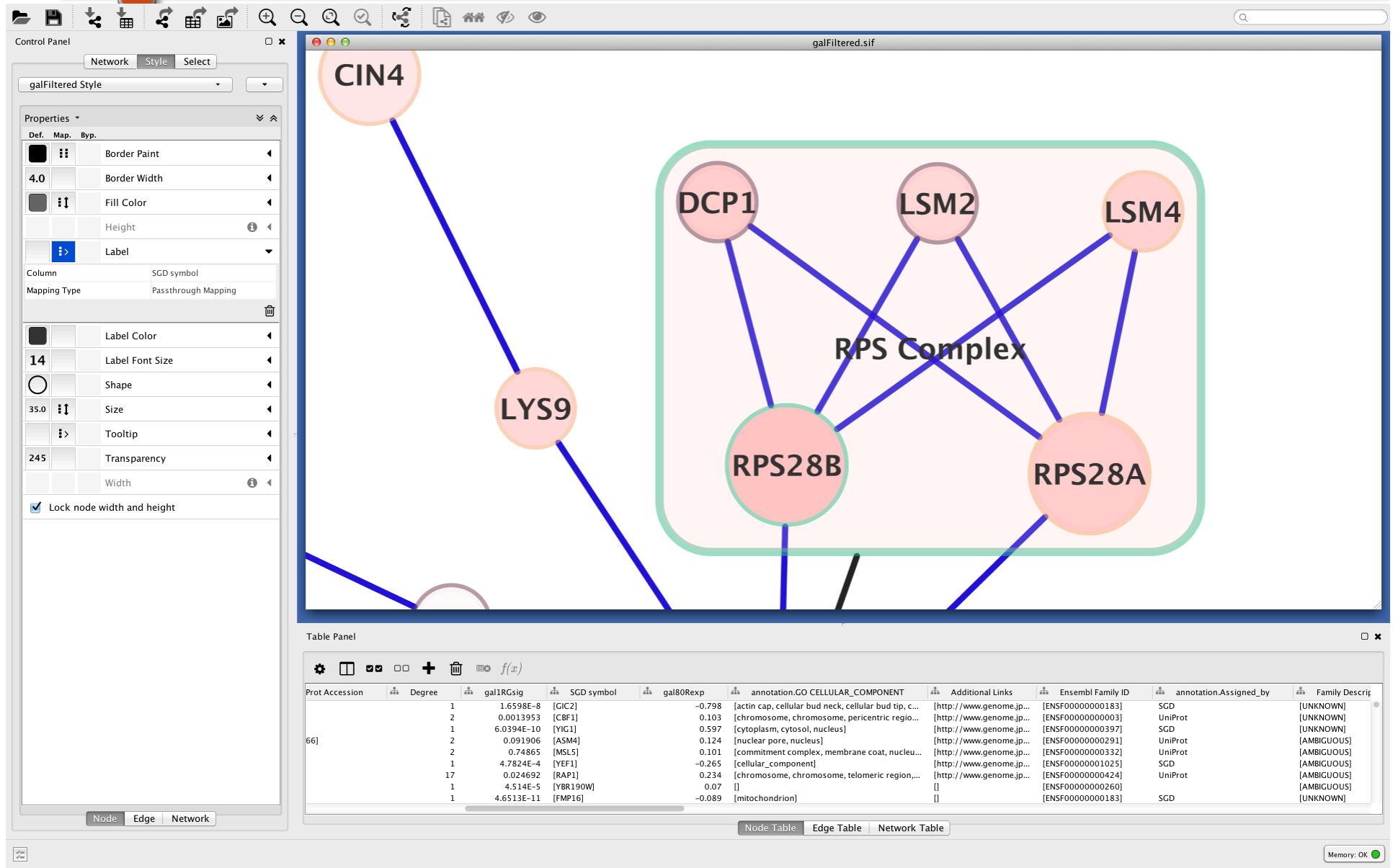


# Cytoscape 3.3

- Release target: November 2015
- Major features:
  - Compound Node Visualization
  - “Core” Apps
  - New table loader
- Other features:
  - Visual redesign
  - OpenCL-based layout
  - Marquee edges
  - Merging of REST into CyREST
  - Core-wide user logging
  - Java 8 required



# Compound Nodes





# Core Apps

- Remove “app-like” functionality from Cytoscape core
  - NetworkAnalyzer
  - Layouts
  - REST interface
- Recreate them as apps
  - Install a version as part of Cytoscape install
  - Can release newer versions via App Store



# New Table Loader

Import Network From Table

Preview

Click on a column to edit it.

Select All    Select None

UniprotID1	Type of interaction	UniprotID2	EntrezID1	GeneSymbol1	EntrezID2	GeneSymbol2
Q9P2J5	hu-hu	UniprotID2	D	LARS	5859	QARS
Q9P2J5	hu-hu		D	LARS	9255	SCYE1
Q9P2J5	hu-hu		D	LARS	5917	RARS
Q9P2J5	hu-hu		D	LARS	4141	MARS
P14868	hu-hu		5	DARS	5859	QARS
P14868	hu-hu		5	DARS	3376	IARS
P14868	hu-hu		5	DARS	4141	MARS
P14868	hu-hu		5	DARS	5917	RARS
P14868	hu-hu		5	DARS	7965	JTV1
P14868	hu-hu		5	DARS	9255	SCYE1
P14868	hu-hu		5	DARS	51520	LARS
P14868	hu-hu	Q15046	1615	DARS	3735	KARS
P14868	hu-hu	O43324	1615	DARS	9521	EEF1E1
P14868	hu-hu	P07814	1615	DARS	2058	EPRS
P54136	hu-hu	Q12904	5917	RARS	9255	SCYE1
Q15046	hu-hu	P54136	3735	KARS	5917	RARS
Q15046	hu-hu	P47897	3735	KARS	5859	QARS
Q15046	hu-hu	Q12904	3735	KARS	9255	SCYE1
Q15046	hu-hu	P56192	3735	KARS	4141	MARS
Q15046	hu-hu	Q9P2J5	3735	KARS	51520	LARS
P47897	hu-hu	Q12904	5859	QARS	9255	SCYE1
P47897	hu-hu	P54136	5859	QARS	5917	RARS
P41252	hu-hu	Q13155	3376	IARS	7965	JTV1
P41252	hu-hu	Q15046	3376	IARS	3735	KARS
P41252	hu-hu	P56192	3376	IARS	4141	MARS
P41252	hu-hu	Q12904	3376	IARS	9255	SCYE1
P41252	hu-hu	Q9P2J5	3376	IARS	51520	LARS
P41252	hu-hu	P54136	3376	IARS	5917	RARS

List Delimiter: |

Advanced Options...

Cancel    OK



# Cytoscape 3.4

- Release target: May-June 2016
- Major features:
  - Headless version
  - Jobs interface
    - Long running tasks
    - Web services integration
  - Themes
    - Groups of apps for particular purposes
  - Relative annotations
- Other features:
  - Redesigned network panel (tree structure)
  - Consistent preferences UI
  - JSON representation for all Cytoscape objects



# Questions?

- [scooter@cgl.ucsf.edu](mailto:scooter@cgl.ucsf.edu)
- [cytoscape-helpdesk@googlegroups.com](mailto:cytoscape-helpdesk@googlegroups.com)