



# Cytoscape

**Network Visualization and Analysis**

**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, NCCR 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

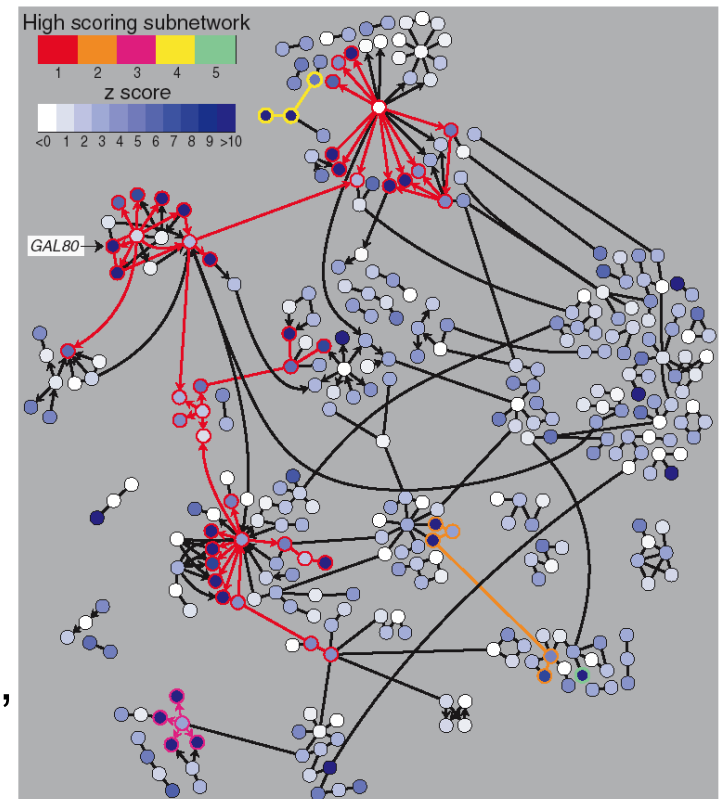
- **Network tools allow...**

Analysis

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

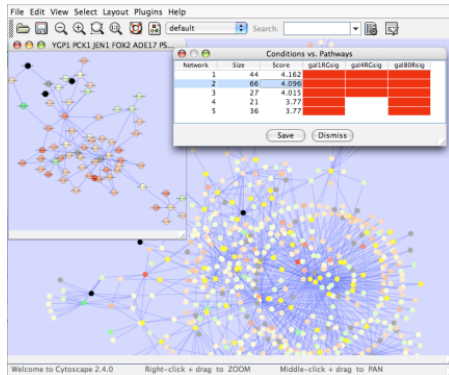
Visualization

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





# Applications of Network Biology



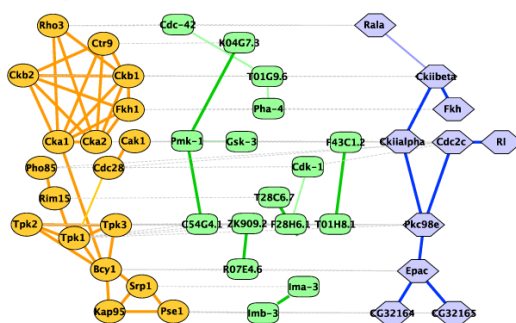
jActiveModules, UCSD

- **Gene Function Prediction**  
shows connections to sets of genes/proteins involved in same biological process
- **Detection of protein complexes/subnetworks**  
discover modularity & higher order organization (motifs, feedback loops)



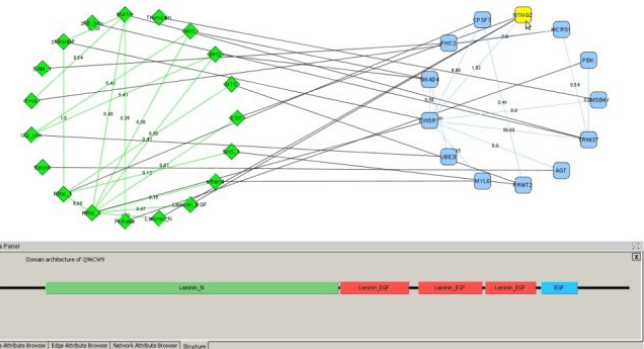
mCode, University of Toronto

[b] Phosphorus metabolism  
Complexes 32, 296, 728, 822, 894, 927



PathBlast, UCSD

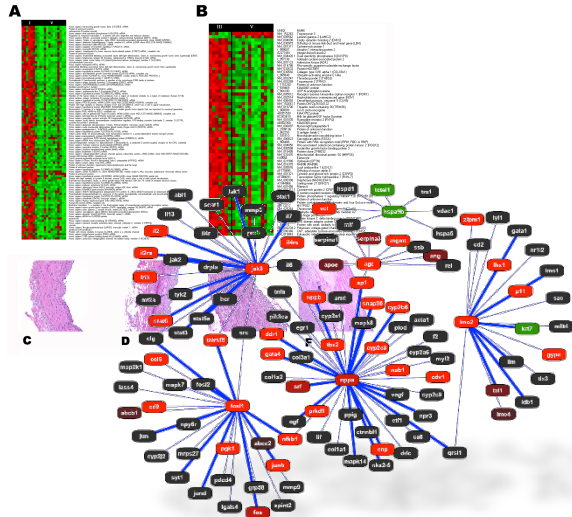
- **Network evolution**  
biological process(s)  
conservation across species
- **Prediction of interactions & functional associations**  
statistically significant domain correlations in protein interaction network to predict protein-protein or genetic interaction



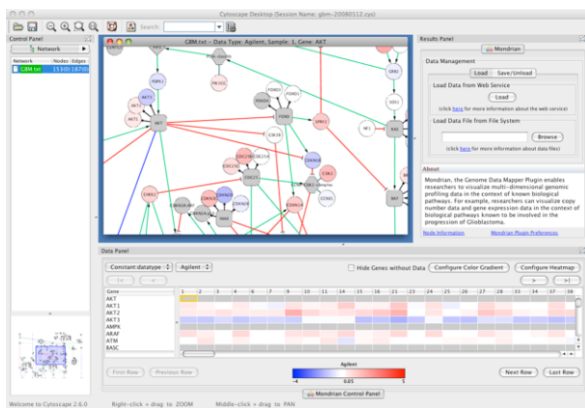
DomainGraph, Max Planck Institute



# Applications in Disease

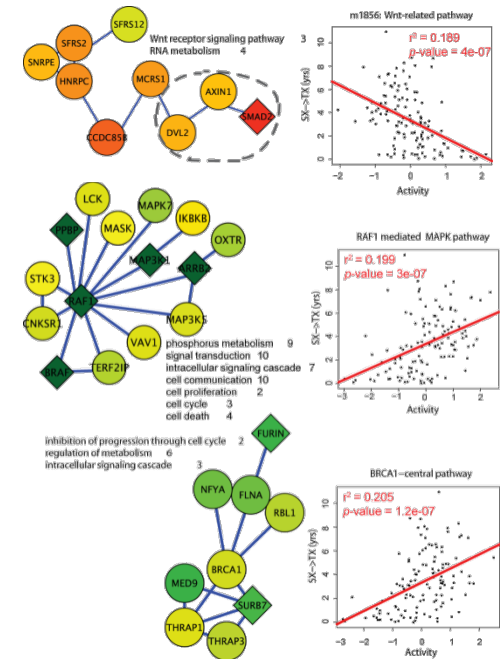


Agilent Literature Search



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

combined\_scores.txt

	A	
1	# Feature1 Sy Feat	
2	Q0010	Q00
3	Q0010	Q00
4	Q0010	Q00
5	Q0010	Q00
6	Q0010	Q00
7	Q0010	Q00
8	Q0010	Q00
9	Q0010	Q00
10	Q0010	Q00
11	Q0010	Q00
12	Q0010	Q00
13	Q0010	Q00
14	Q0010	Q00
15	Q0010	Q00
16	Q0010	Q00
17	Q0010	Q00
18	Q0010	Q00
19	Q0010	Q00
20	Q0010	Q00
21	Q0010	Q00
22	Q0010	Q00
23	Q0010	Q00
24	Q0010	Q00
25	Q0010	Q00
26	Q0010	Q00
27	Q0010	Q00
28	Q0010	Q00
29	Q0010	Q00
30	Q0010	Q00
31	Q0010	Q00
32	Q0010	Q00
33	Q0010	Q00
34	Q0010	Q00

combir

Normal View Ready

Sum=0

SCRL CAPS NUM



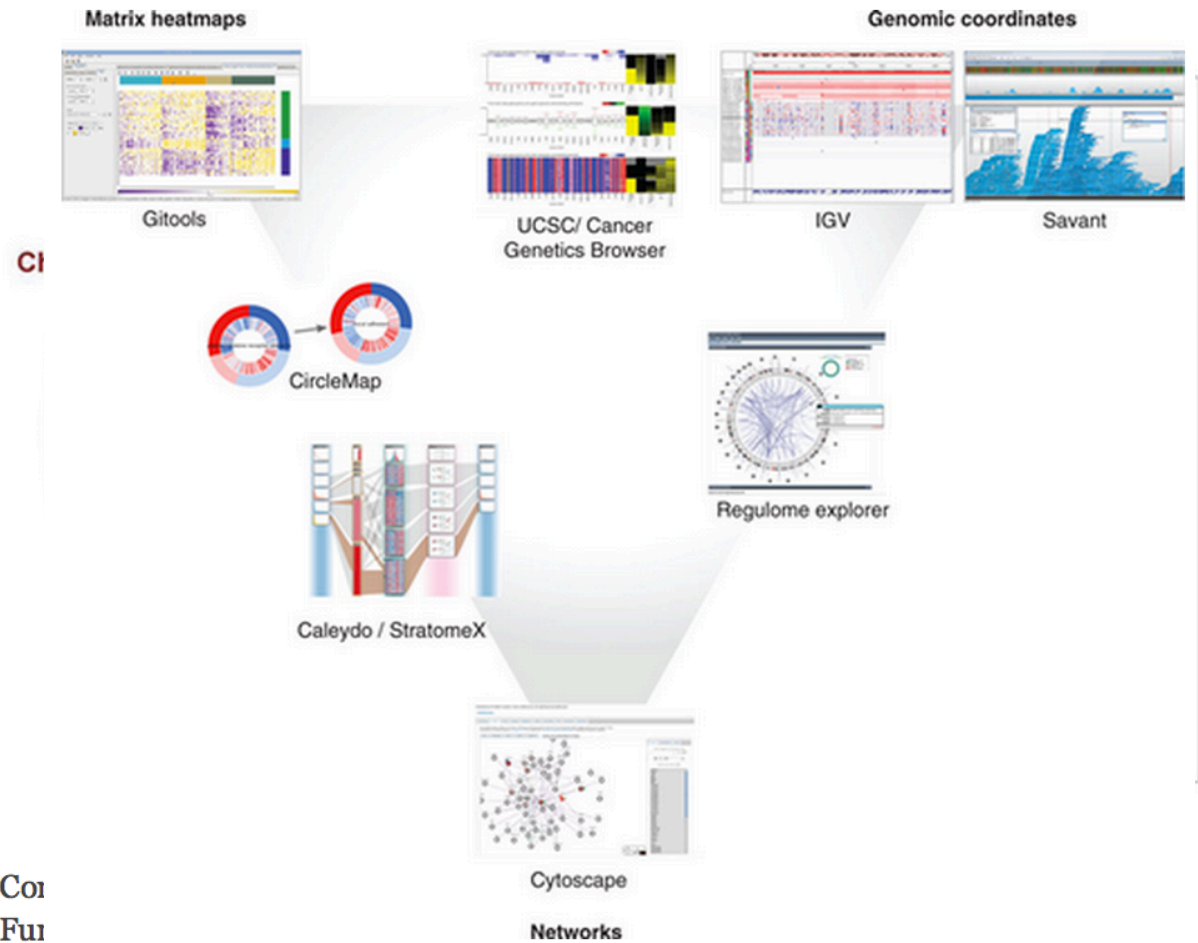


# The Challenge

- 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



Visualizing multidimensional cancer genomics data.

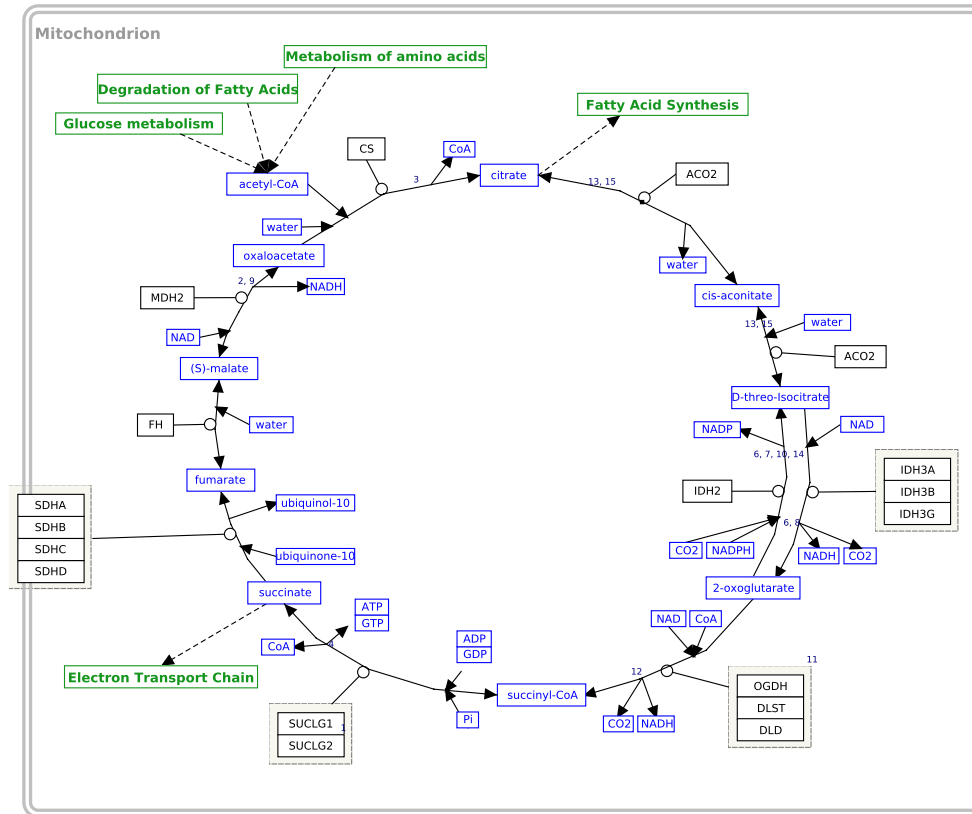
“ Schroeder et al, *Genome Med.* 2013 Jan 31;5(1):9.

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

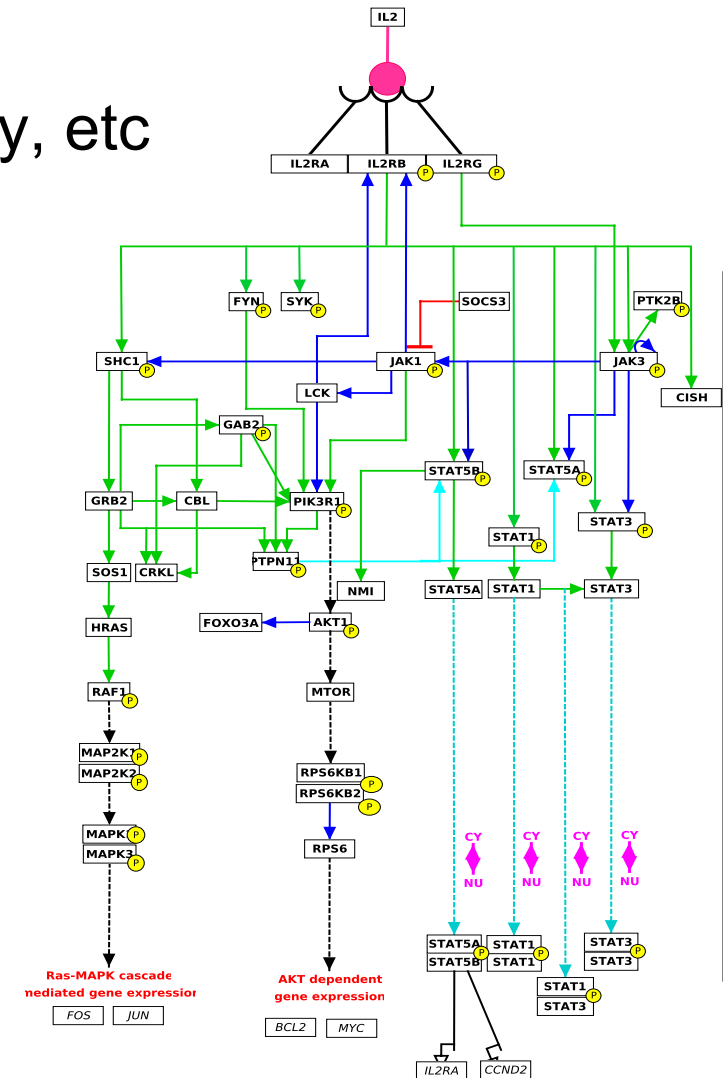


# Biological Network Taxonomy

- Pathways
  - Signaling, Metabolic, Regulatory, etc



IL-2 Signaling Pathway

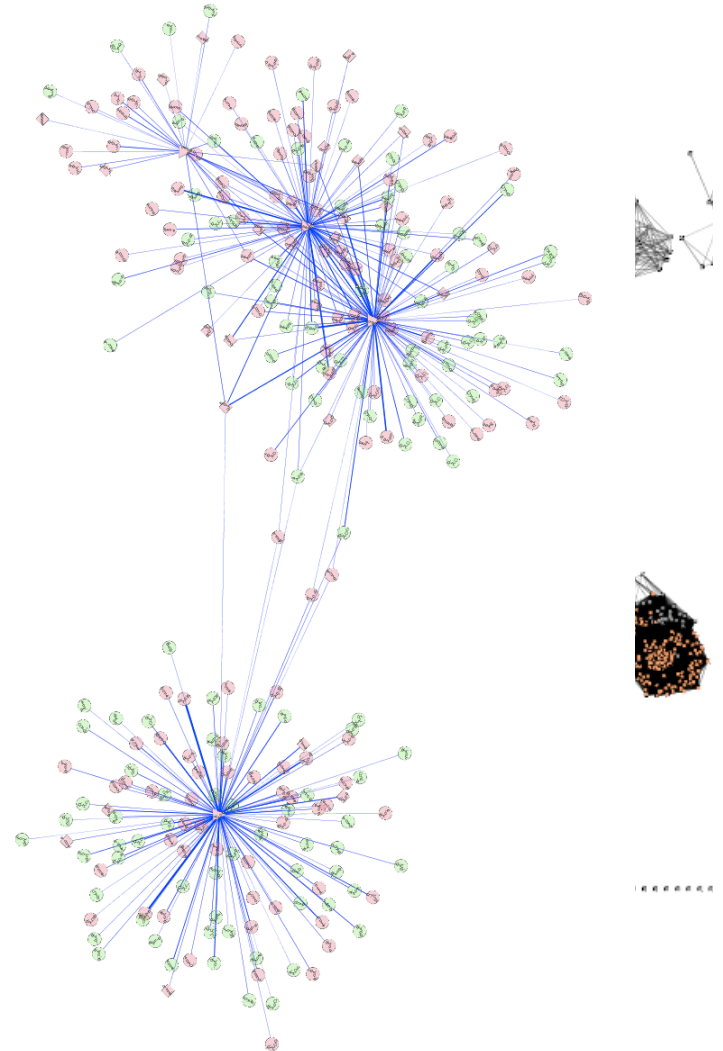






# Biological Network Taxonomy

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

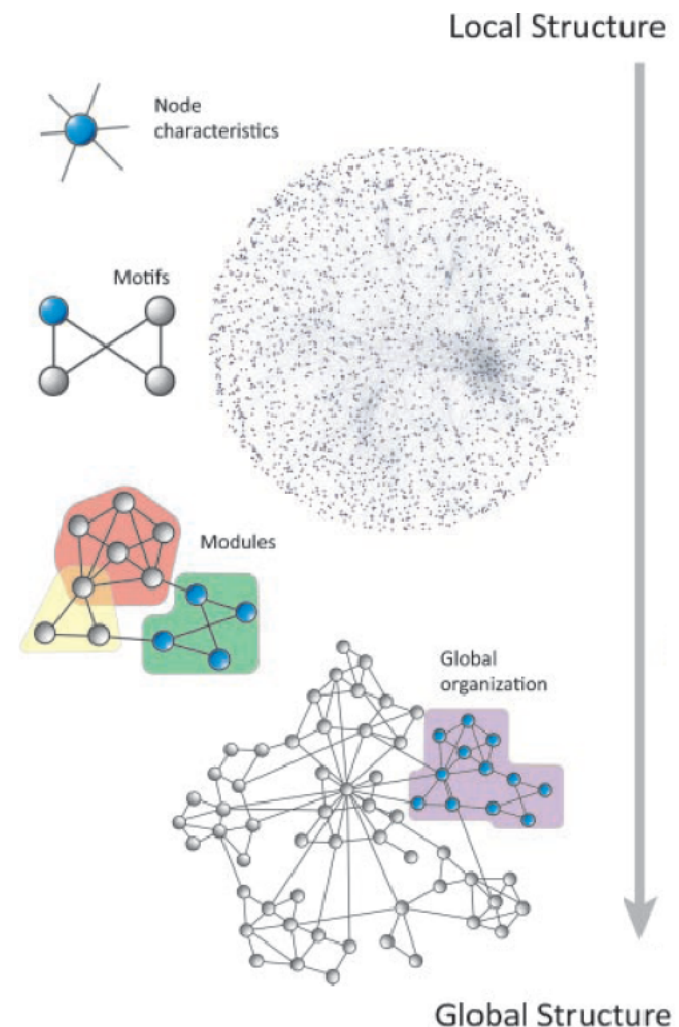




# Analytical Approaches

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, centrality, 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.

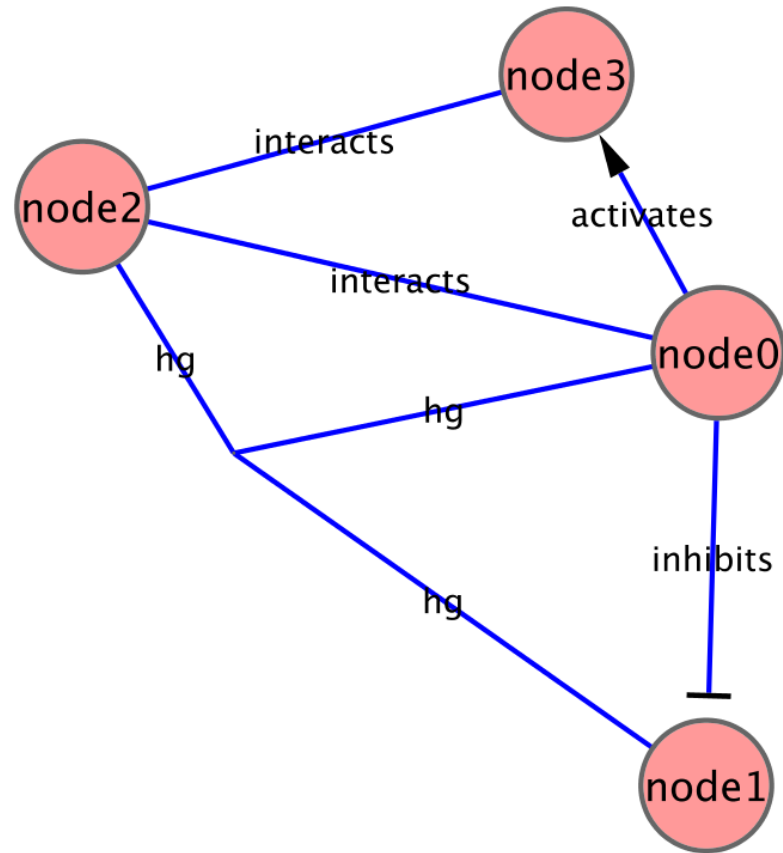
**Motif analysis** is the identification of small network patterns that are over-represented when compared with a randomized version of the same network. Regulatory elements are often composed of such motifs.

**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

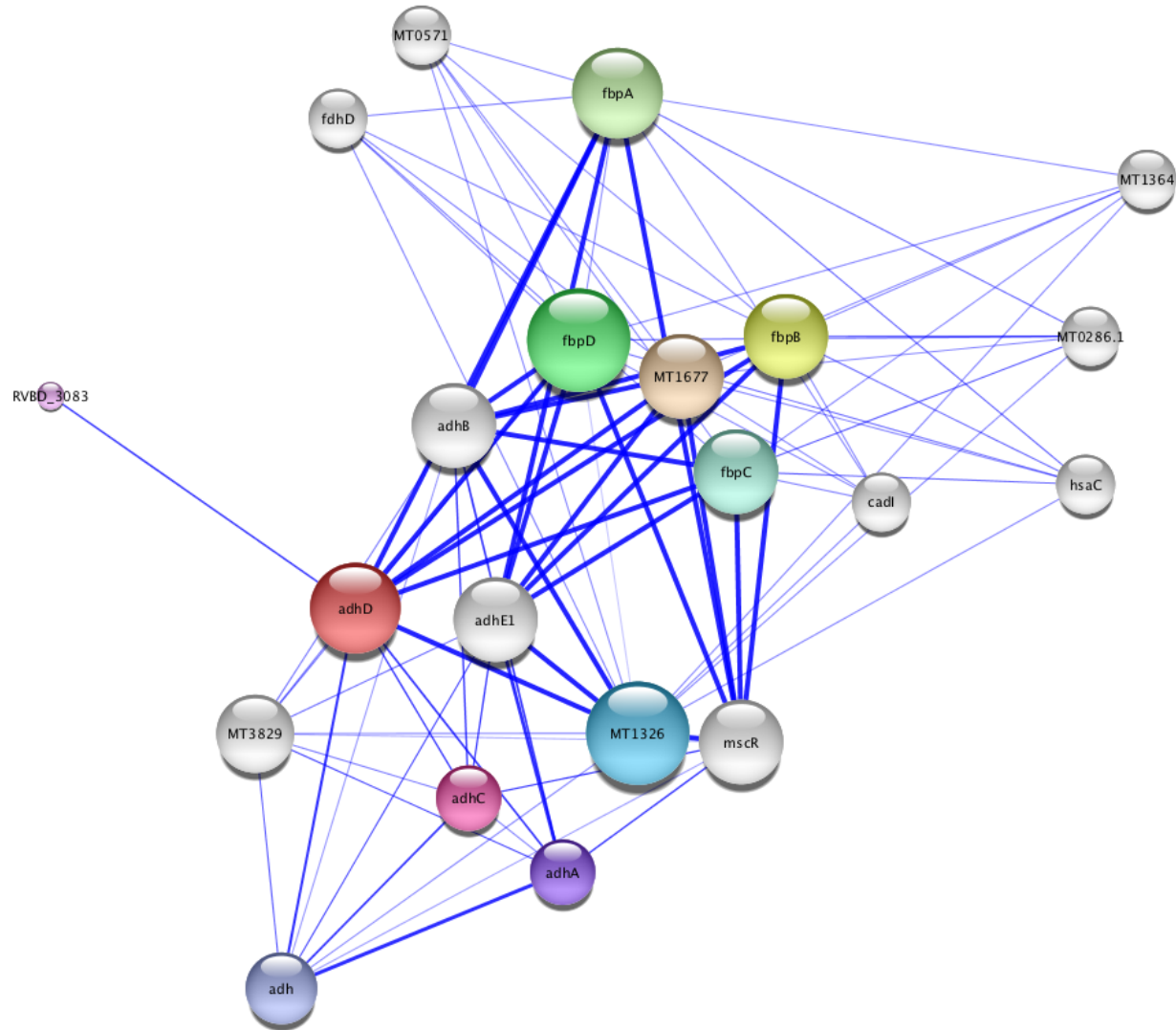
- 





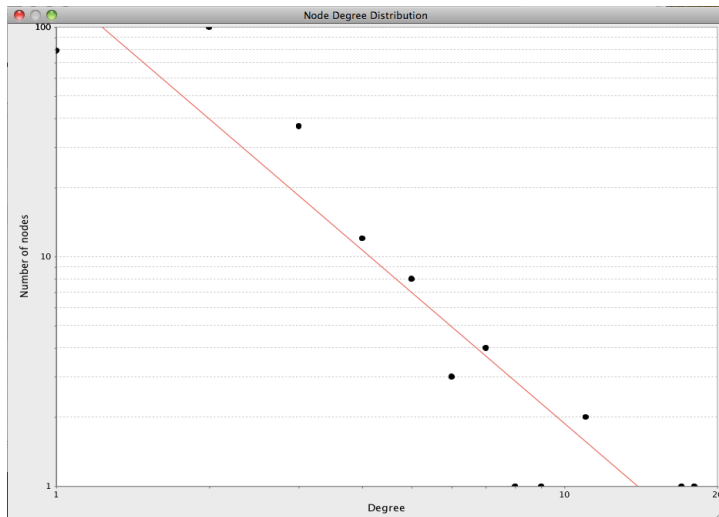


# Analytical Approaches





# Analytical Approaches



NetworkAnalyzer - Fitted Function

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

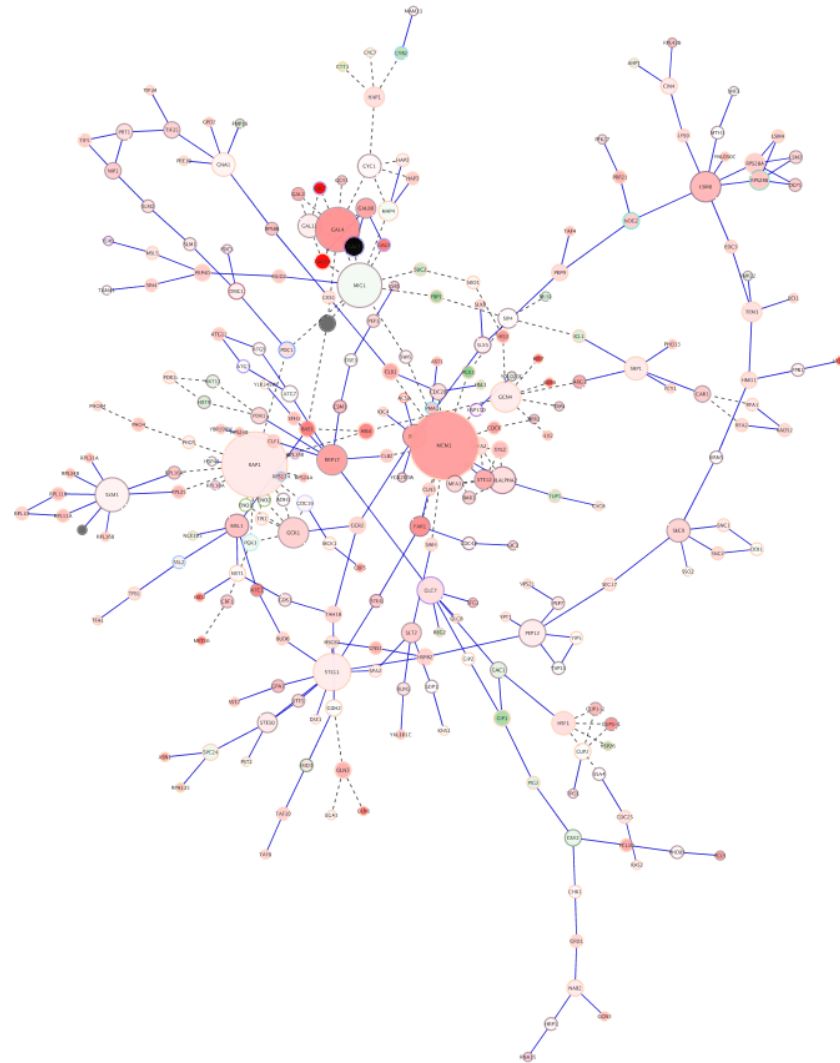
a = 148.46      b = -1.896

Correlation = 0.742

R-squared = 0.871

Note: R-squared is computed on logarithmized values.

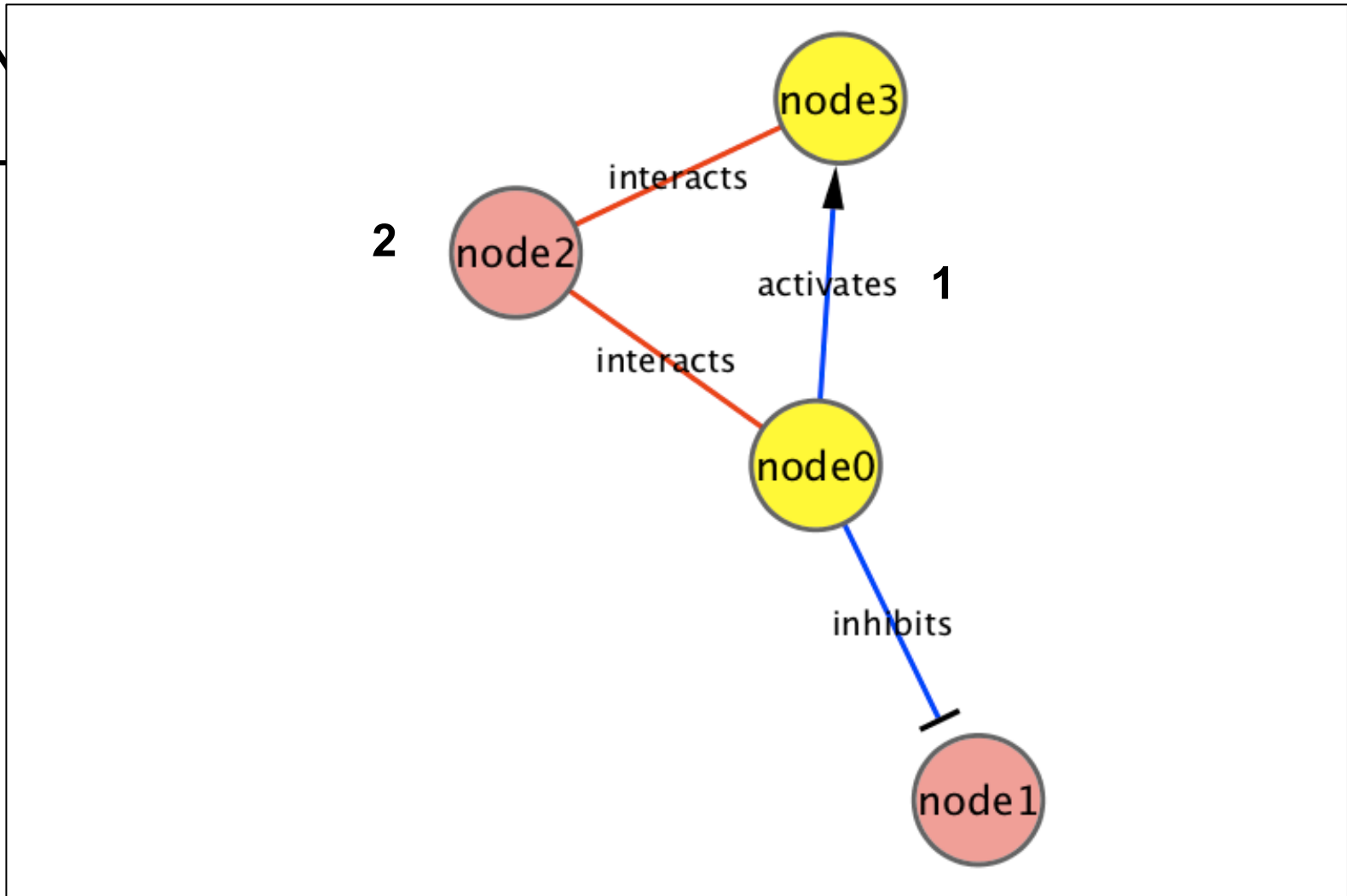
OK      Help





# Analytical Approaches

- N





# 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

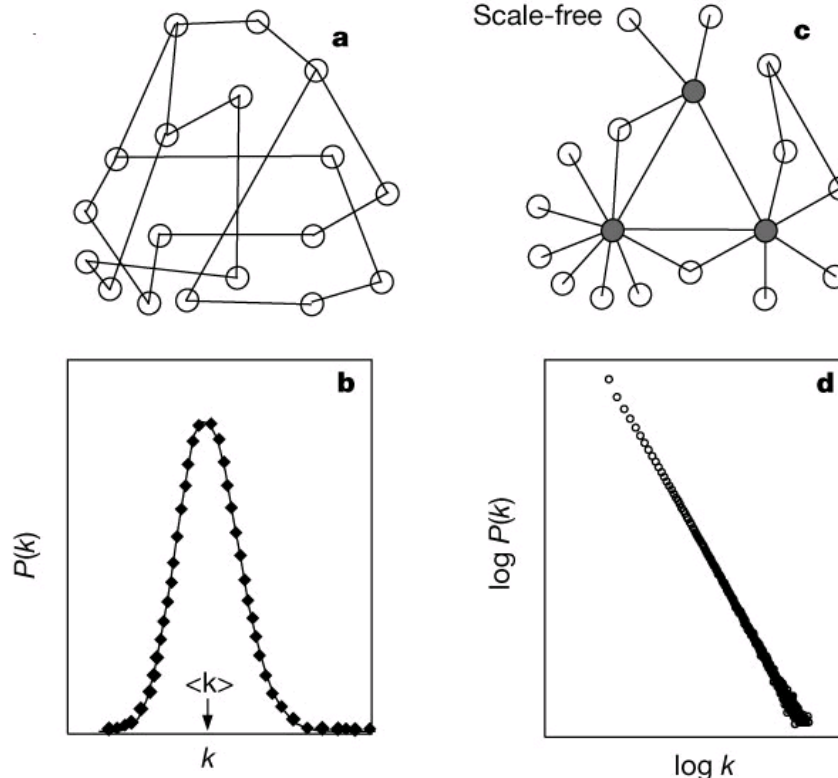
- Random networks

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

- Algorithm

- Flat random
- Scale-free
- Small-world

- Useful to network



networks

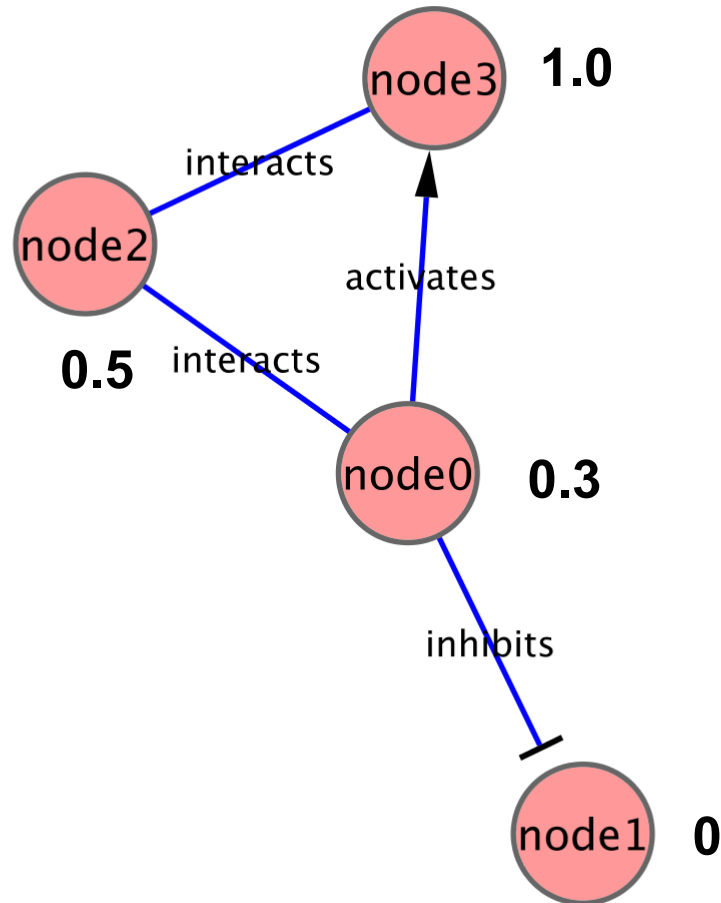
logatz

a random



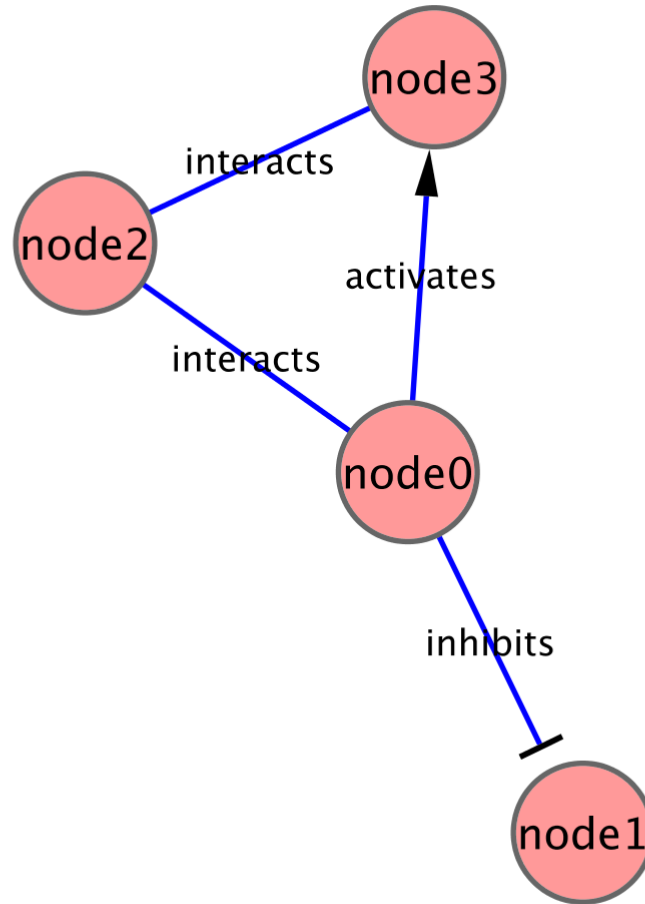
# Analytical Approaches

- 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

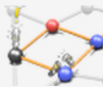
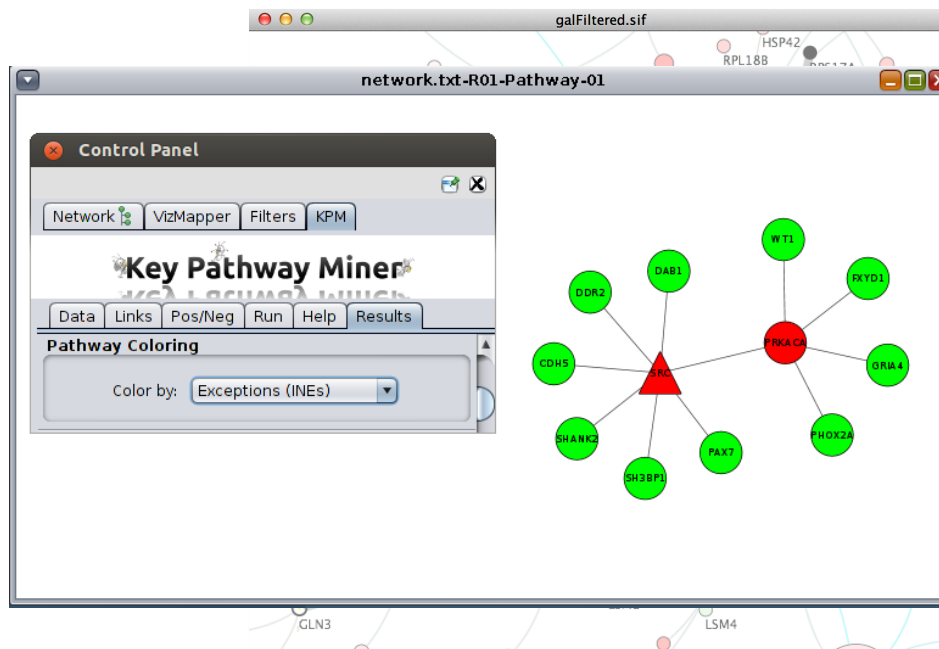
- Network Analyzer Demo...





# Analytical Approaches

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



**KeyPathwayMiner** 3.0+ 3.0+

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

★★★★☆ (10) 177 downloads

sed and loads



# 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 displays the ClueGO software interface. On the left, a network graph is visualized with nodes of various colors and sizes, representing different biological terms. Labels for some clusters include "mesoderm formation", "regulation of T", "regulation of inflammatory response", "The Co-Stimulatory Signaling", "cellular defense response", and "chemotaxis".

Three tool descriptions are overlaid on the interface:

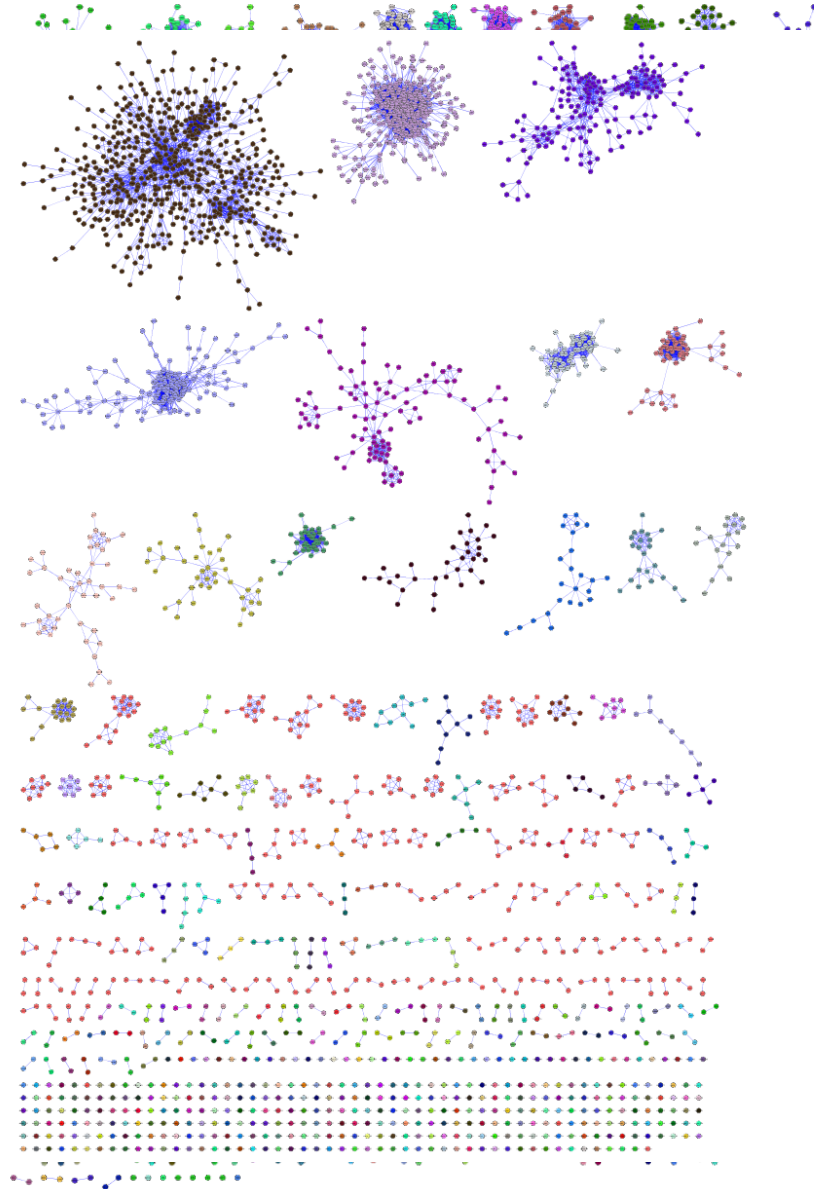
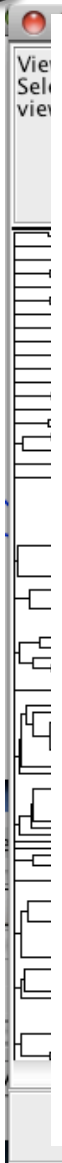
- ClueGO** (3.0+): Creates and visualizes a functionally grouped network of terms/pathways. (47) 4807 downloads.
- BiNGO** (3.0+): Calculates overrepresented GO terms in the network and display them as a network of significant GO terms. (20) 1325 downloads.
- NOA**: The Network Ontology Analysis plugin performs ontology overrepresentation analysis based on the network connections among annotate nodes. (7).

On the right side of the screenshot, a table titled "Project" shows analysis results. The table has columns for "Term PVal", "Term", "Group PVal", "Group Value", and "Associated Genes Found".

Term PVal	Term	Group PVal	Group Value	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, CYP11B1, 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, MNMAT1]	
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, MNMAT1]	
E-3 1.6 E-1	2.1 E-3	8.4 E-3	[KYNU, MNMAT1]	
E-4 5.8 E-3	2.1 E-3	8.4 E-3	[INDO, KYNU, MNMAT1]	
E-3 4.5 E-2	2.1 E-3	8.4 E-3	[KYNU, MNMAT1]	
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, UPF 3B]	
E-2 5.3 E-1	3.3 E-2	1.3 E-1	[SMG1, UPF 3B]	



# Analytical Approaches



edges

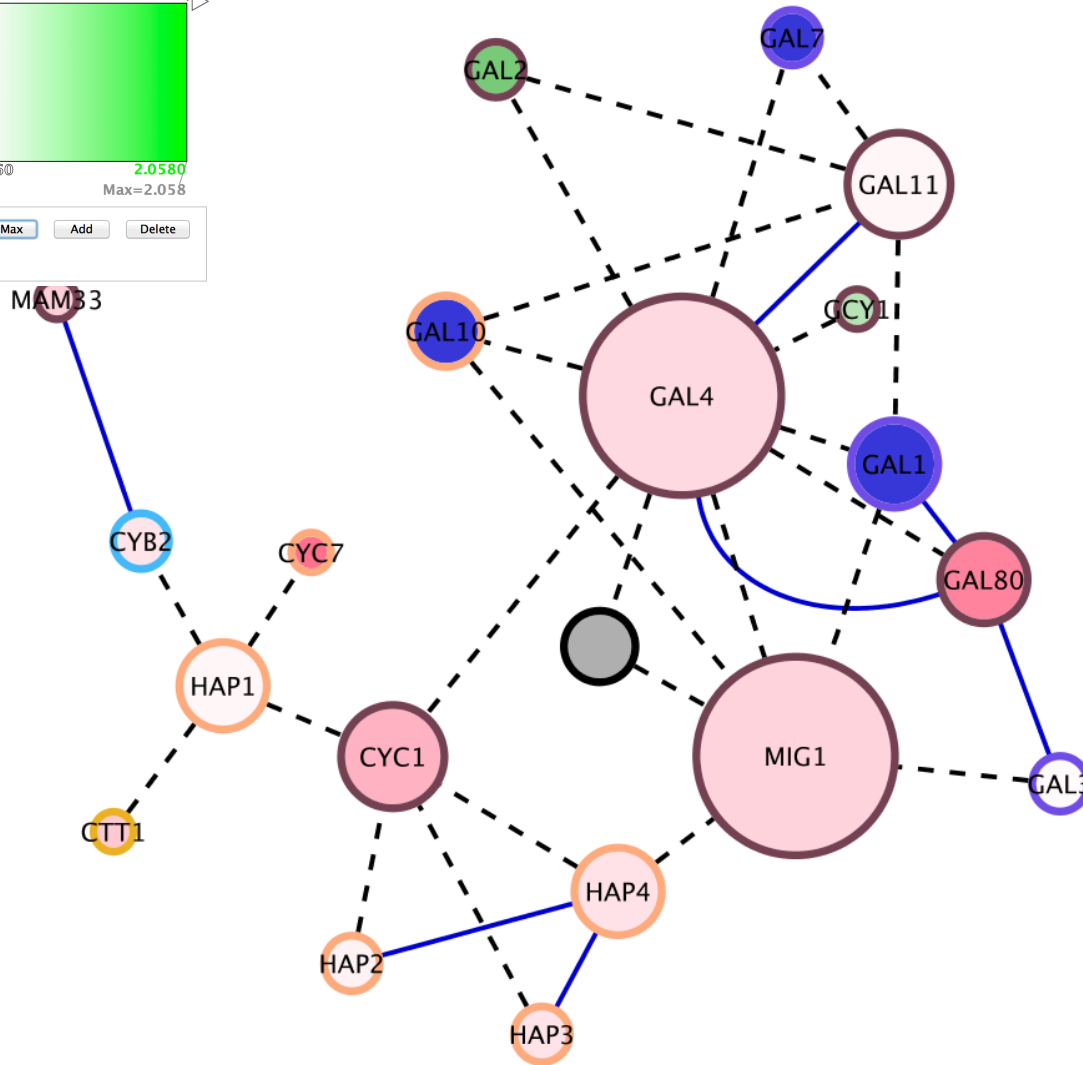
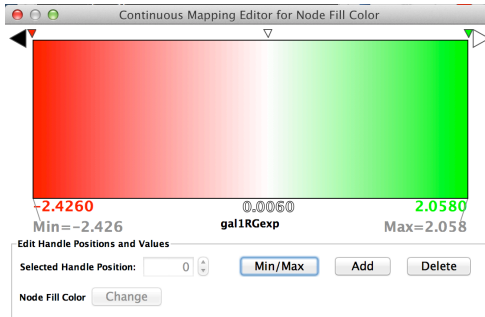


# Visualization of Biological Networks

- Data Mapping
- Layouts
- Animation



# Data Mapping



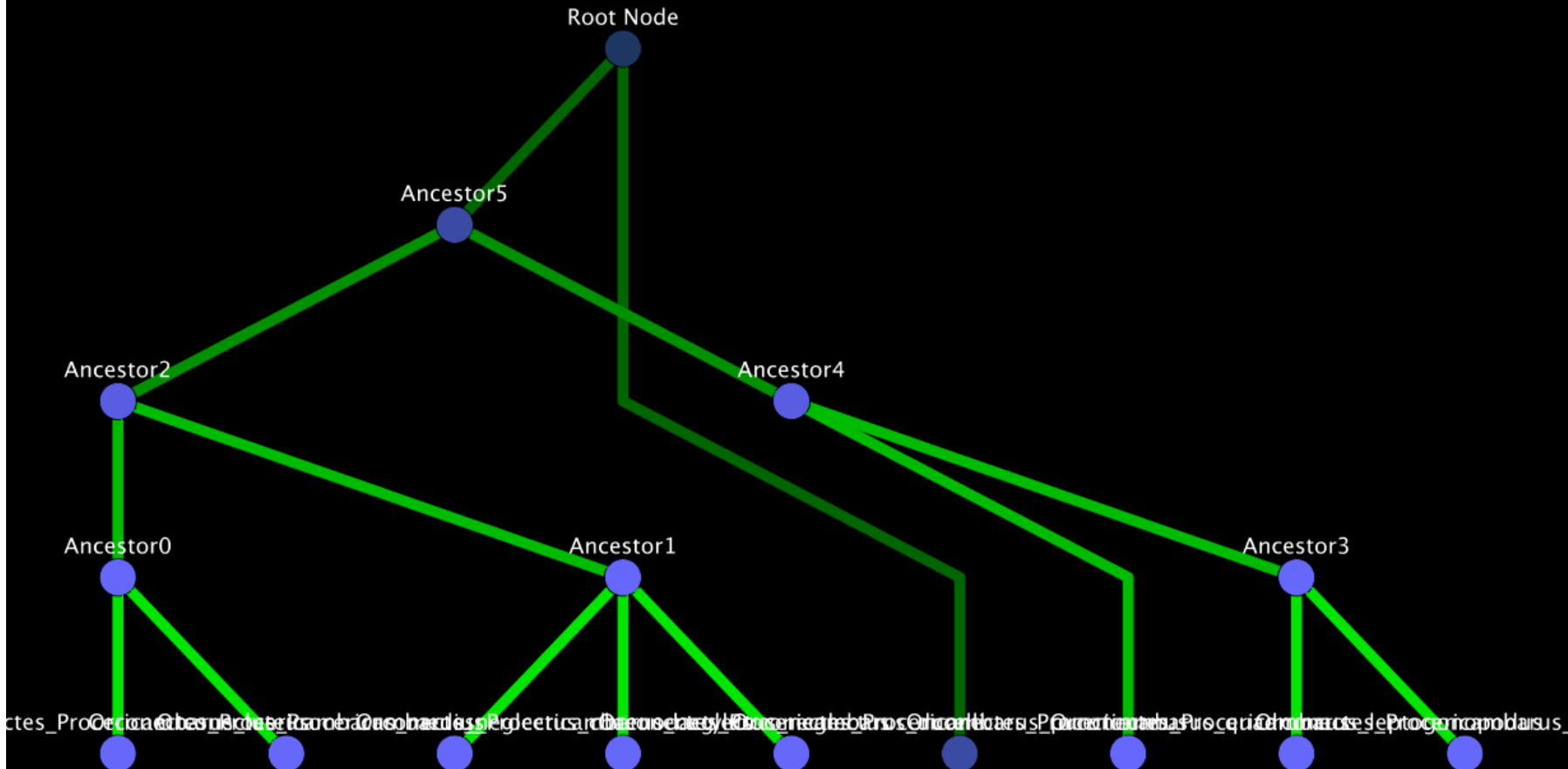


# Data Mapping

- 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

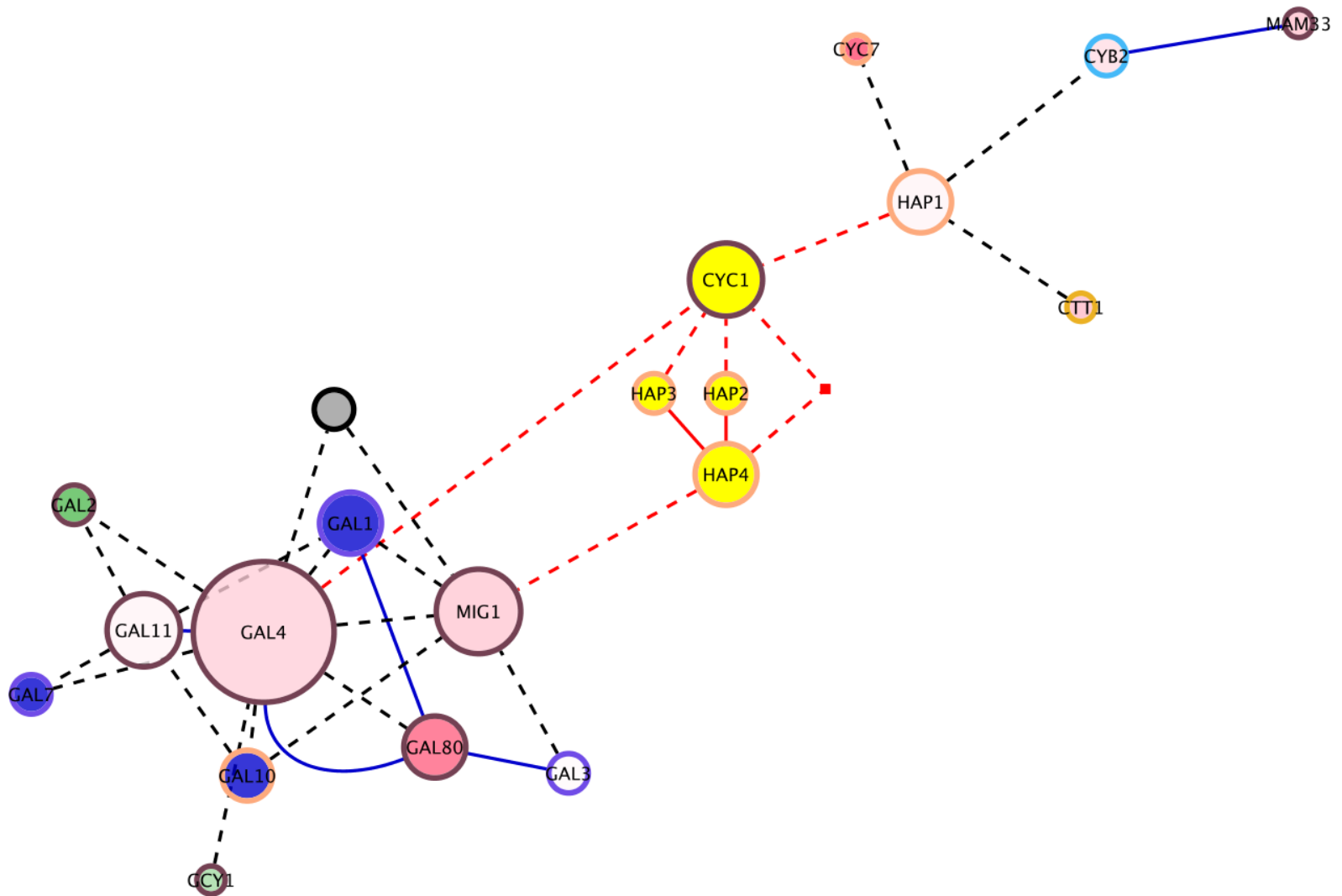


# Layouts





# Layouts



– many, many others



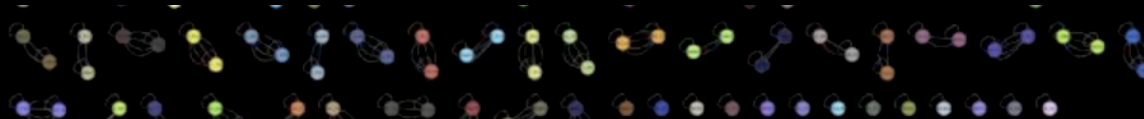
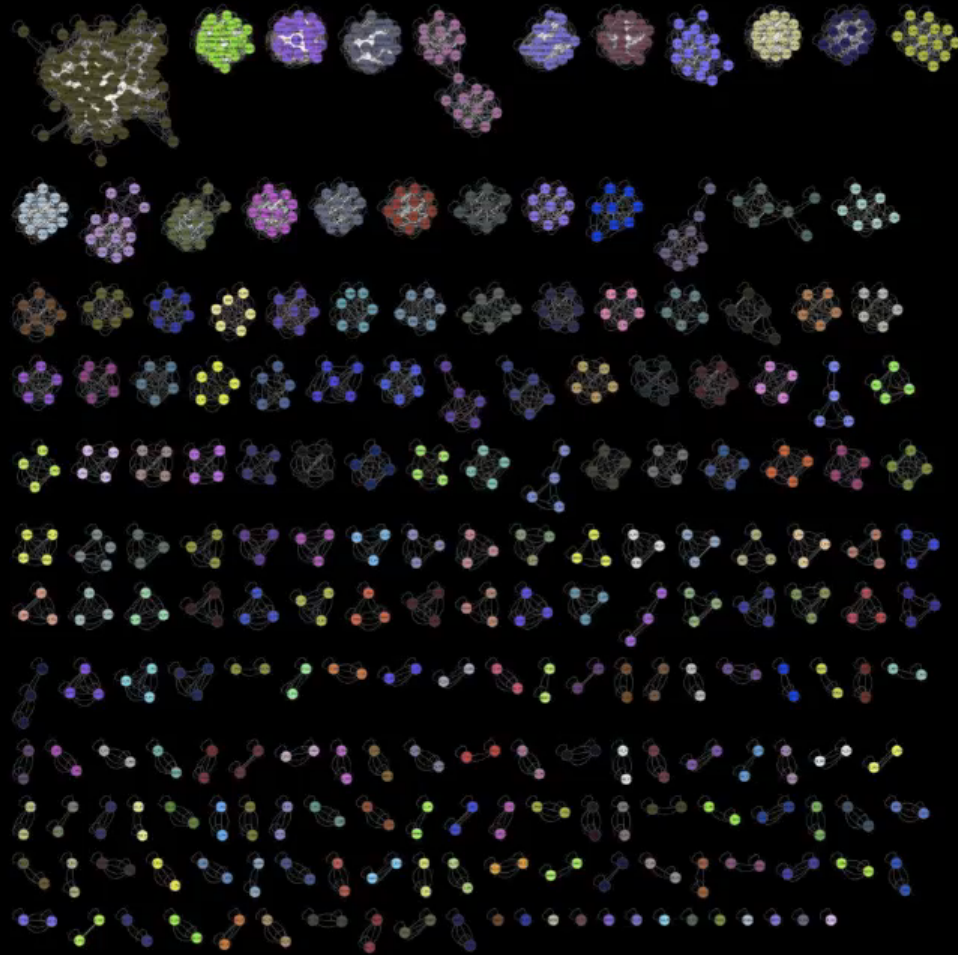


# Layouts

- 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

- Open source
- Cross platform
- Consortium



Institute for Systems Biology



University of California at San Diego



Memorial Sloan-Kettering Cancer Center



Institut Pasteur



Agilent Technologies

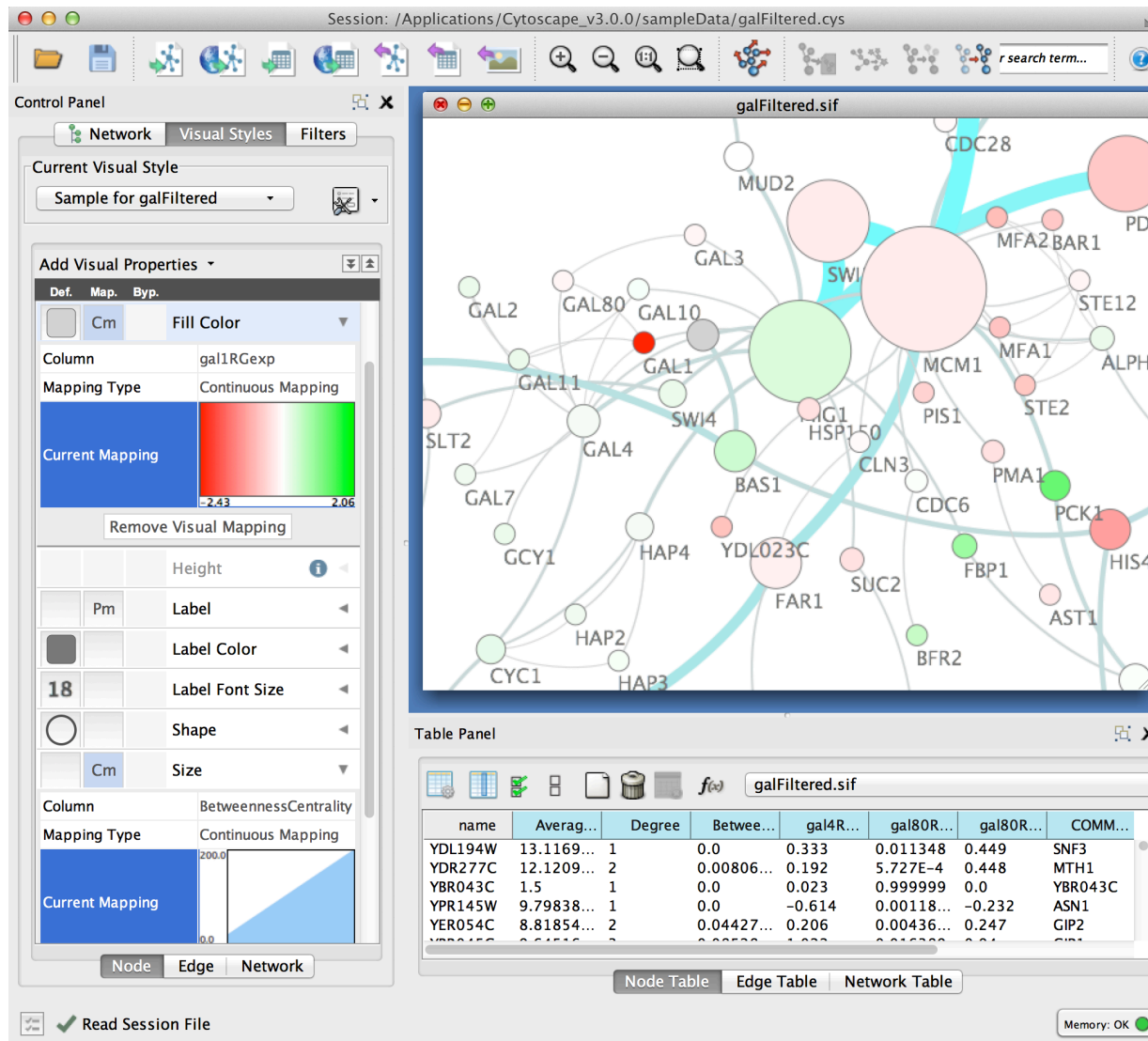


University of California at San Francisco



University of Toronto

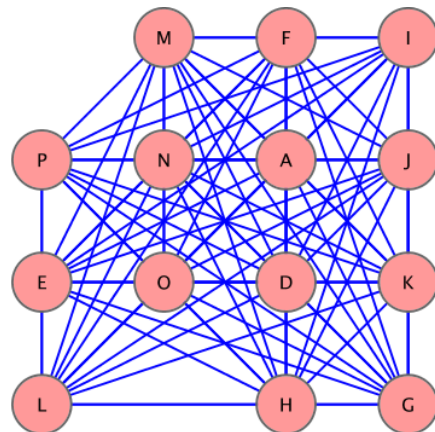
**GLADSTONE  
INSTITUTES**





# Core Concepts

- Networks and Tables



## Networks

e.g., PPIs or pathways

A screenshot of a table with columns: GI\_Number, Gene, HPRD\_ID, REFSEQ, Entrez\_ID, OMIM\_ID, SWISS56 Gene\_Description, Architecture, Arch. The table contains 26 rows of data, including gene names like SDSL, SEC23B, and SERPINB12, and their corresponding identifiers and descriptions.

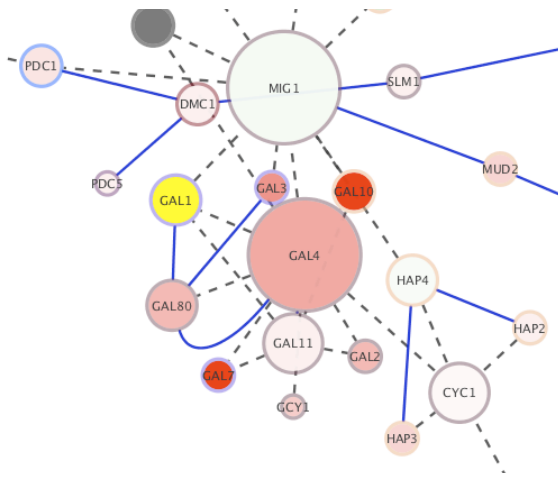
## Tables

e.g., data or annotations



# Core Concepts

- Networks and Tables



GI_Number	Gene	HPRD_ID	REFSEQ	Entrez_ID	OMIM_ID	SWISSF	Gene_Description	Architecture	Arch
gi19923959	SDSL	11542	NP_612441.1	113675			SDSL	SP	Motif
gi14591926	SEC23B	11543	NP_116781.1	10483		Q15437	Protein transport protein SEC23B	GEL	Dom
gi14602492	SCMH1	11540	AAH09752.1	22955		Q96GD3	SCMH1	SAM	Dom
gi21361625	EVOX2	11544	NP_060773.3	55770		Q96KP1	SEC5 like 1	IPP	Dom
gi17998551	SERPINB12	11548	NP_536722.1	89777		Q96P63	Serpin B12	SERPIN	Dom
gi65287717	EIF2AK4	18447	NP_00101372	440275	609280		Similar to GCN2 eIF2alpha kinase	S_T_Y_Kinase	Dom
gi4507755	TYROBP	4996	NP_003323.1	7305	604142	O43914	DAP12	ITAM	Dom
gi5803086	IL24	4995	NP_006841.1	11009	604136	Q13007	Interleukin 24	IL10	Dom
gi21265034	ADAMTS13	4994	NP_620594.1	11093	604134		ADAMTS 13	TSP1	Dom
gi1743873	PIPSK1A	4470	AAC50911.1	8394	603275	Q99755	Phosphatidylinositol-4-phosphate 5 PIPKc	PIPKc	Dom
gi11496982	SVIL	4992	NP_068506.1	6840	604126	O95425	Supervillin	VHP	Dom
gi31563386	SULT2B1	4991	NP_814444.1	6820	604125		Sulfotransferase family 2B, member 1		
gi21040399	RBBP8	4990	AAH30590.1	5932	604124	Q99708	CTIP	LZ	Motif
gi4758378	FIGF	2102	NP_004460.1	2277	300091	O43915	VEGF D	PDGF	Dom
gi11321617	DPYSL4	7463	NP_006417.1	10570	608407	O14531	Collapsin response mediator protein 3		
gi8923202	TASP1	7460	NP_060184.1	55617	608270	Q9H6P5	Threonine aspartase 1		
gi5454090	SSR4	2101	NP_006271.1	6748	300090	P51571	Signal sequence receptor delta	SP	Motif
gi5730045	SLC16A2	2106	NP_006508.1	6567	300095	P36021	X linked PEST containing transport	TM	Dom
gi11968027	FTS	7467	NP_071921.1	64400	608483	Q9H8T0	FTS	UBC	Dom
gi48255885	PRKCI	2105	NP_002731.3	5584	600539	P41743	Protein kinase C, iota type	S_T_Kinase	Dom
gi9558731	RPA4	6591	NP_037479.1	29935			RPA4	TRNA	Dom
gi13376812	PPP1R2P9	6593	NP_079486.1	80316			Type 1 protein phosphatase inhibitor		
gi15826862		6595	NP_296375.1	90060			JM1 protein	CC	Motif
gi7661844	CCDC22	6594	NP_054727.1	28952			JM1 protein	CC	Motif
gi6005794	PRAF2	6596	NP_009144.1	11230			JM4 protein	TM	Dom

Networks

Tables

Visual Styles





# 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
- 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

Control Panel

Network

Start New Session

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

From Preset Network

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

Load Preset Network

Open Recent Session

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

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

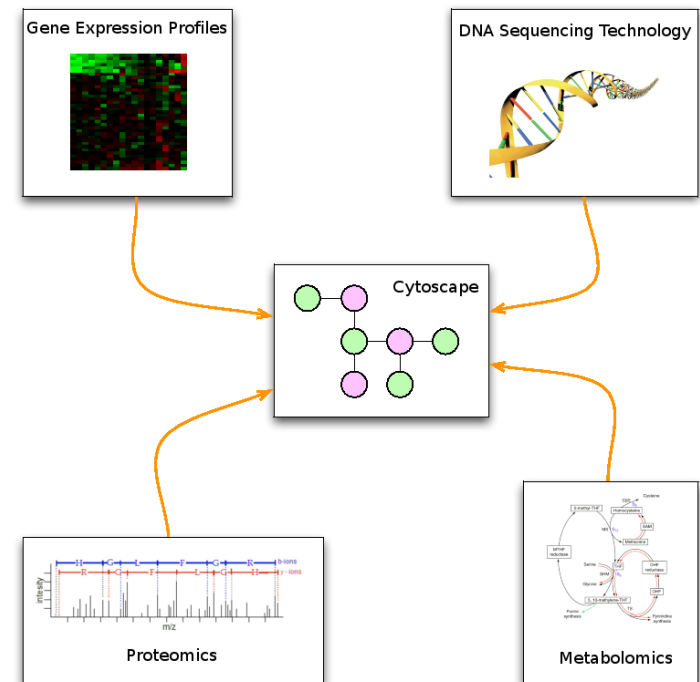
Memory: OK

?!



# 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

shared...	name	isExclu...	COMM...	gal1R...	gal4R...	gal80R...	gal1R...	gal4i
	YGL229C	false	SAP4	-0.521	0.171	-0.208	2.3655E...	0.0208
	YBR217W	false	APG12	0.088	-0.332	0.378	0.143	0.0110
	YBR160W	false	CDC28	-0.016	-0.087	-0.405	0.7432	0.3210
	YNL216W	false	RAP1	0.205	0.015	0.234	0.024692	0.9349
	YLR249W	false	YEF3	-0.39	-0.394	-0.769	2.713E-8	0.0474
	YGL202W	false	ARO8	-0.305	-0.286	-0.536	4.2677E...	4.2373
	YML123C	false	PHO84	0.283	-0.114	0.692	0.00151...	0.0194
	YBR018C	false	GAL7	0.153	-1.995	3.126	7.8855E...	3.6284
	YIL105C	false	YIL105C	0.106	0.022	0.107	6.0834E...	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 **Sample for galFiltered** Solid

Universe

Size Transparency Width

Node Edge Network

galFiltered.sif

galFiltered.sif

	name	Averag...	Cluster...	Closen...	IsSingl...	Partne...
W	YDL194W	13.1169...	0.0	0.07623...	false	0
C	YDR277C	12.1209...	0.0	0.08250...	false	0
YBR043C	YBR043C	1.5	0.0	0.66666...	false	0
YPR145W	YPR145W	9.79838...	0.0	0.10205...	false	0
YER054C	YER054C	8.81854...	0.0	0.11339...	false	0
YBR045C	YBR045C	9.64516...	0.0	0.10367...	false	0
YBL079W	YBL079W	2.0	0.0	0.5	false	0
YLR345W	YLR345W	9.11693...	0.0	0.10968...	false	0
YLR052C	YLR052C	9.11693...	0.0	0.10968...	false	0

Node Table Edge Table Network Table

Searching PSICQUIC Services

Memory: OK



# Selection Filters

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

galFiltered.sif

Table Panel

age...	Cluster...	Closen...	IsSingl...	Partner...	SelfLoo...	Eccentri...	Stress	Degree	Between...
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	20	70312	5	0.15820...

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


Control Panel

Network Style Filter Filter (New)

Network Nodes

galFiltered.sif 331(0)

galFiltered.sif



Indexing Network Data

### 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



PHO5

S24B

RPL1

Be
0.0
0.00
0.0
0.0
0.04
0.08
0.0
0.0
0.0

memory: OK

— Click OK





# Expression Data Analysis

The screenshot displays the 'clusterMaker TreeView : galFiltered.sif' application window. The interface is divided into several sections:

- View Status:** Shows 'NODEID: GROUP327X' and 'CORRELATIO'.
- Dendrogram:** A tree structure at the top center, with a red line highlighting a specific branch.
- Heatmap:** A large central area showing a heatmap of expression data with a color scale from dark blue to yellow.
- Usage Hints:** A box on the right stating 'Click to select node - use arrow keys to navigate tree'.
- Gene List:** A list of gene identifiers on the right, including 'YBR018C', 'YBR019C', and 'YBR020W'. A small legend with colored squares (black, blue, yellow) is visible above this list.
- Buttons:** A bottom toolbar with buttons for 'Settings...', 'Save Data...', 'Export Graphics...', 'Flip Tree Nodes', 'Map Colors Onto Network...', and 'Close'.

Labels 'tree' and 'network' are placed to the right of the dendrogram and heatmap areas, respectively.



# 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	72(0)	127(0)

cluster1

BiNGO output

cluster1

GO,Saccharomyces cerevisiae,default,bingo,namespace close

GO-ID	Description	p-val	corr p-val	cluster freq	total freq	genes
<input type="checkbox"/> 33499	galactose catabolic process via UDP-galactose	2.5090...	1.9069...	3/3	100.0%	3/6208 0... YBR020W YBR018C YBR019C
<input type="checkbox"/> 19388	galactose catabolic process	5.0181...	1.9069...	3/3	100.0%	6/6208 0... YBR020W YBR018C YBR019C
<input type="checkbox"/> 6012	galactose metabolic process	4.1399...	1.0488...	3/3	100.0%	11/6208 ... YBR020W YBR018C YBR019C
<input type="checkbox"/> 19320	hexose catabolic process	7.3414...	1.3726...	3/3	100.0%	57/6208 ... YBR020W YBR018C YBR019C
<input type="checkbox"/> 46365	monosaccharide catabolic process	9.0300...	1.3726...	3/3	100.0%	61/6208 ... YBR020W YBR018C YBR019C
<input type="checkbox"/> 46164	alcohol catabolic process	1.1481...	1.4543...	3/3	100.0%	66/6208 ... YBR020W YBR018C YBR019C
<input type="checkbox"/> 44275	cellular carbohydrate catabolic process	1.9088...	2.0724...	3/3	100.0%	78/6208 ... YBR020W YBR018C YBR019C
<input type="checkbox"/> 16052	carbohydrate catabolic process	2.5677...	2.4394...	3/3	100.0%	86/6208 ... YBR020W YBR018C YBR019C
<input type="checkbox"/> 19318	hexose metabolic process	6.6970...	5.6552...	3/3	100.0%	118/6208... YBR020W YBR018C YBR019C
<input type="checkbox"/> 5996	monosaccharide metabolic process	9.4003...	7.1443...	3/3	100.0%	132/6208... YBR020W YBR018C YBR019C
<input type="checkbox"/> 44282	small molecule catabolic process	1.3282...	0.1731...	3/3	100.0%	148/6208... YBR020W YBR018C YBR019C

Select All Unselect All Select nodes

Start BiNGO



# Expression Data Analysis

- Active modules
  - The jActiveModule

that show

r time-

The screenshot displays the jActiveModules software interface. At the top, a 'Control Panel' window shows the 'Target Network' set to 'galFiltered.sif'. Below it, a 'Numeric Node Attributes' table lists columns for 'Name', 'Most sig', 'Least sig', 'Reverse sig', and 'Scaling'. The main window, titled 'Module\_0\_4', shows a network diagram with nodes colored by their activity: green (e.g., MCM1, MLS1, SIP4, CAR1, ICL1, FBP1), red (e.g., ARG1, SRP1), and pink (e.g., PCK1). Two inset diagrams show zoomed-in views of other network modules. Below the network diagram is a 'Table Panel' with a table of expression data for 'Module\_0\_4'.

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

ules  
ages



# Protein Complexes

Session: /Users/scooter/Documents/collinsPlus.cys

Control Panel

Network | Style | Filter | Filter (New) | Sets

Network	Nodes	Edges
combined_scores_good.txt	2401(0)	16218...
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

Table Panel

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	-2.952	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
PUP1											

Node Table | Edge Table | Network Table

Indexing Network Data

Memory: OK



# Protein Complexes

Session: /Users/scooter/Documents/collinsPlus3.cys

GIM3

Control Panel

Network Style Filter Filter (New) Sets

Network	Nodes	Edges
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

Table Panel

combined\_scores\_good.txt--clustered

shared...	0_MCL...	canoni...	cluster	compl...	cytos...	DNA a...	Endo f...	Featur...	Featur...	GPLS1...	GI
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	TFC8	0.385	0.49
vnc8	68	vnc8	165	1506	[vnc8]	523	225	vnc8		0.166	0.20

Node Table Edge Table Network Table

Indexing Network Data

Memory: OK

ges



# Protein Complexes

**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

Memory: OK

Control Panel

Network Style Filter Filter (New)

Network Nodes

combined\_scores\_good.txt 2401

combined\_scores\_good.txt 2401

DNA and Tran 07-21-06b.csv 743(0

RNAPuberNov2+Meg6c.csv 552(0

combined\_scores\_good.txt--cluster2401

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
		0.166	0.30

Indexing Network Data



# Protein Complexes

The image displays the clusterMaker TreeView software interface, which is used for analyzing protein complexes. The main window is titled "clusterMaker TreeView : DNA and Tran 07-21-06b.csv". It features a large heatmap visualization of gene expression data, with a dendrogram on the left side. A "View Stat" panel on the left indicates "6 genes".

Below the heatmap, a network graph is shown, representing the relationships between genes. The nodes are labeled with gene names: YKE2, PAC10, GIM5, GIM4, PFD1, and GIM3. The graph shows a complex network of connections between these genes.

The interface also includes a "Control Panel" on the left, which lists the loaded files: "combined\_scores\_good.txt", "DNA and Tran 07-21-06b.csv", and "RNAPuberNov2+Meg6C.csv". A "Table Panel" at the bottom provides a detailed view of the data, with columns for "chared", "MCL", "Cannoni", "cluster", "comml", "cvtoc", "DNA", "Endo", "Feat", "Feat", "GPI.S1", "GPI.S1", and "GPI.S1".

Buttons at the bottom of the interface include "Settings...", "Save Data...", "Export Graphics...", "Flip Tree Nodes", "Map Colors Onto Network...", and "Close". A status bar at the bottom indicates "Performing Hierarchical cluster" and "Memory: OK".



# From Networks to Structures

The screenshot displays the UCSF Chimera software interface. On the left, the 'Control Panel' shows a network graph with the following data:

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

Below the network graph is the 'Cytoscape Molecular Structure Navigator' showing '1 Open Chimera Models' and 'Model #0 1EZ2 (1 chains, 332 residues) [Network {RIN 1EZ2 ligand}, No...'. The main window displays a 3D ribbon structure of a protein, colored by domain (blue, orange, yellow, green, cyan). A detailed view of the protein structure is shown on the right, with specific residues labeled: EU 211.A, HIS 201.A, DII 1403.A, TRP 131.A, ILE 106.A, and R 308.A. The interface includes standard menu bars (File, Edit, View, Select, Layout, Apps, Tools, Help) and window controls.





# 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

The screenshot displays a network visualization software interface. The main window shows a network graph with nodes represented as pink circles. A sub-network of five nodes (DCP1, LSM2, LSM4, RPS28B, and RPS28A) is highlighted with a green rounded rectangle and labeled "RPS Complex". Other nodes include CIN4, LYS9, and RPS28A. The interface includes a Control Panel on the left with various node and edge properties, and a Table Panel at the bottom with a data table.

**Control Panel**

Network Style Select

galFiltered Style

Properties

Def. Map. Byp.

Border Paint

Border Width: 4.0

Fill Color

Height

Label

Column: SGD symbol

Mapping Type: Passthrough Mapping

Label Color

Label Font Size: 14

Shape

Size: 35.0

Tooltip

Transparency: 245

Width

Lock node width and height

Node Edge Network

**Table Panel**

Prot Accession	Degree	gal1RGsig	SGD symbol	gal80Rexp	annotation.GO CELLULAR_COMPONENT	Additional Links	Ensembl Family ID	annotation.Assigned_by	Family Descri
	1	1.6598E-8	[GIC2]	-0.798	[actin cap, cellular bud neck, cellular bud tip, c...	[http://www.genome.jp...	[ENSF00000000183]	SGD	[UNKNOWN]
	2	0.0013953	[CBF1]	0.103	[chromosome, chromosome, pericentric regio...	[http://www.genome.jp...	[ENSF00000000003]	UniProt	[UNKNOWN]
	1	6.0394E-10	[YIC1]	0.597	[cytoplasm, cytosol, nucleus]	[http://www.genome.jp...	[ENSF00000000397]	SGD	[UNKNOWN]
	2	0.091906	[ASM4]	0.124	[nuclear pore, nucleus]	[http://www.genome.jp...	[ENSF00000000291]	UniProt	[AMBIGUOUS]
	2	0.74865	[MSL5]	0.101	[commitment complex, membrane coat, nucleu...	[http://www.genome.jp...	[ENSF00000000332]	UniProt	[AMBIGUOUS]
	1	4.7824E-4	[YEF1]	-0.265	[cellular_component]	[http://www.genome.jp...	[ENSF00000001025]	SGD	[AMBIGUOUS]
	17	0.024692	[RAP1]	0.234	[chromosome, chromosome, telomeric region,...	[http://www.genome.jp...	[ENSF00000000424]	UniProt	[AMBIGUOUS]
	1	4.514E-5	[YBR190W]	0.07	[ ]	[http://www.genome.jp...	[ENSF00000000260]	[ ]	[AMBIGUOUS]
	1	4.6513E-11	[FMP16]	-0.089	[mitochondrion]	[ ]	[ENSF00000000183]	SGD	[UNKNOWN]

Node Table Edge Table Network Table

Memory: OK



# 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			LARS	5859	QARS
Q9P2J5	hu-hu			LARS	9255	SCYE1
Q9P2J5	hu-hu			LARS	5917	RARS
Q9P2J5	hu-hu			LARS	4141	MARS
P14868	hu-hu			DARS	5859	QARS
P14868	hu-hu			DARS	3376	IARS
P14868	hu-hu			DARS	4141	MARS
P14868	hu-hu			DARS	5917	RARS
P14868	hu-hu			DARS	7965	JTV1
P14868	hu-hu			DARS	9255	SCYE1
P14868	hu-hu			DARS	51520	LARS
P14868	hu-hu			DARS	3735	KARS
P14868	hu-hu			DARS	9521	EEF1E1
P14868	hu-hu			DARS	2058	EPRS
P54136	hu-hu			RARS	9255	SCYE1
Q15046	hu-hu			KARS	5917	RARS
Q15046	hu-hu			KARS	3735	KARS
Q15046	hu-hu			KARS	5859	QARS
Q15046	hu-hu			KARS	9255	SCYE1
Q15046	hu-hu			KARS	4141	MARS
Q15046	hu-hu			KARS	51520	LARS
P47897	hu-hu			QARS	9255	SCYE1
P47897	hu-hu			QARS	5917	RARS
P41252	hu-hu			IARS	7965	JTV1
P41252	hu-hu			IARS	3735	KARS
P41252	hu-hu			IARS	4141	MARS
P41252	hu-hu			IARS	9255	SCYE1
P41252	hu-hu			IARS	51520	LARS
P41252	hu-hu			IARS	5917	RARS

UniprotID2

ab 1 123 1.0 y/n

[ ab ] [ 1 ] [ 123 ] [ 1.0 ] [ y/n ]

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)