



**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

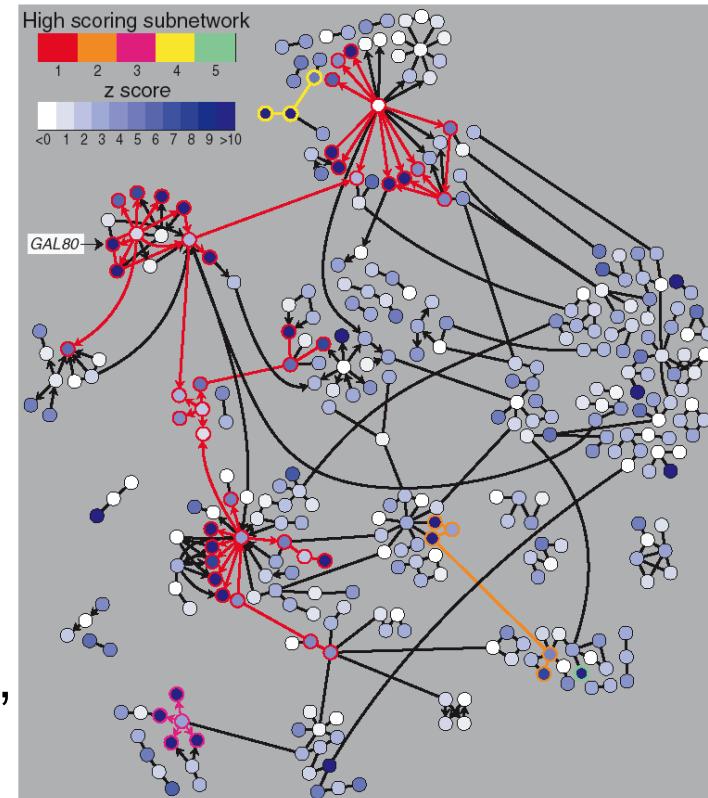
- **Network tools allow...**

## Analysis

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

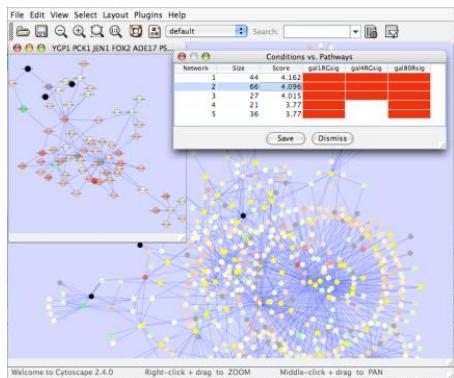
## Visualization

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

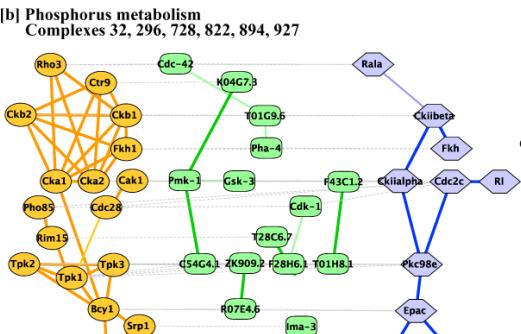




# Applications of Network Biology

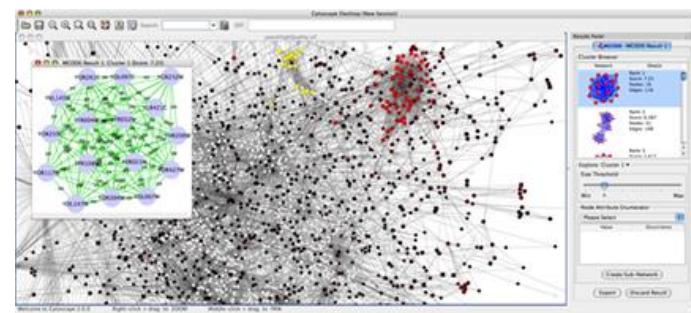


jActiveModules, UCSD

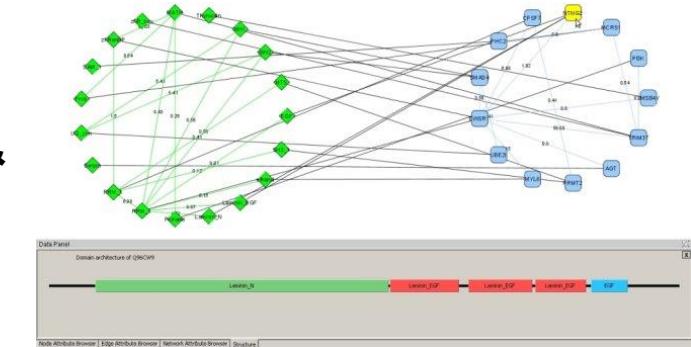


PathBlast, UCSD

- 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



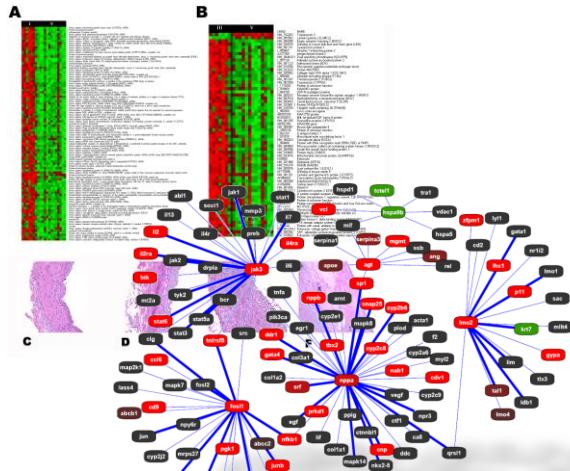
mCode, University of Toronto



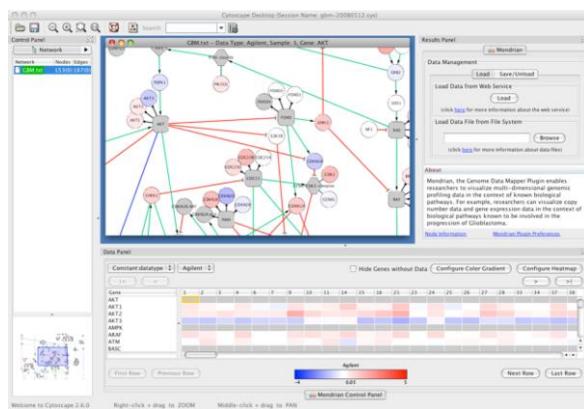
DomainGraph, Max Planck Institute



# Applications in Disease



Agilent Literature Search

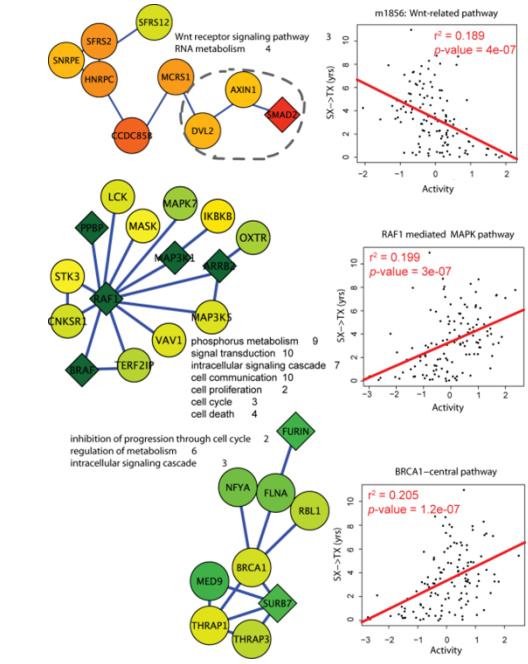


Mondrian, MSKCC

- **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

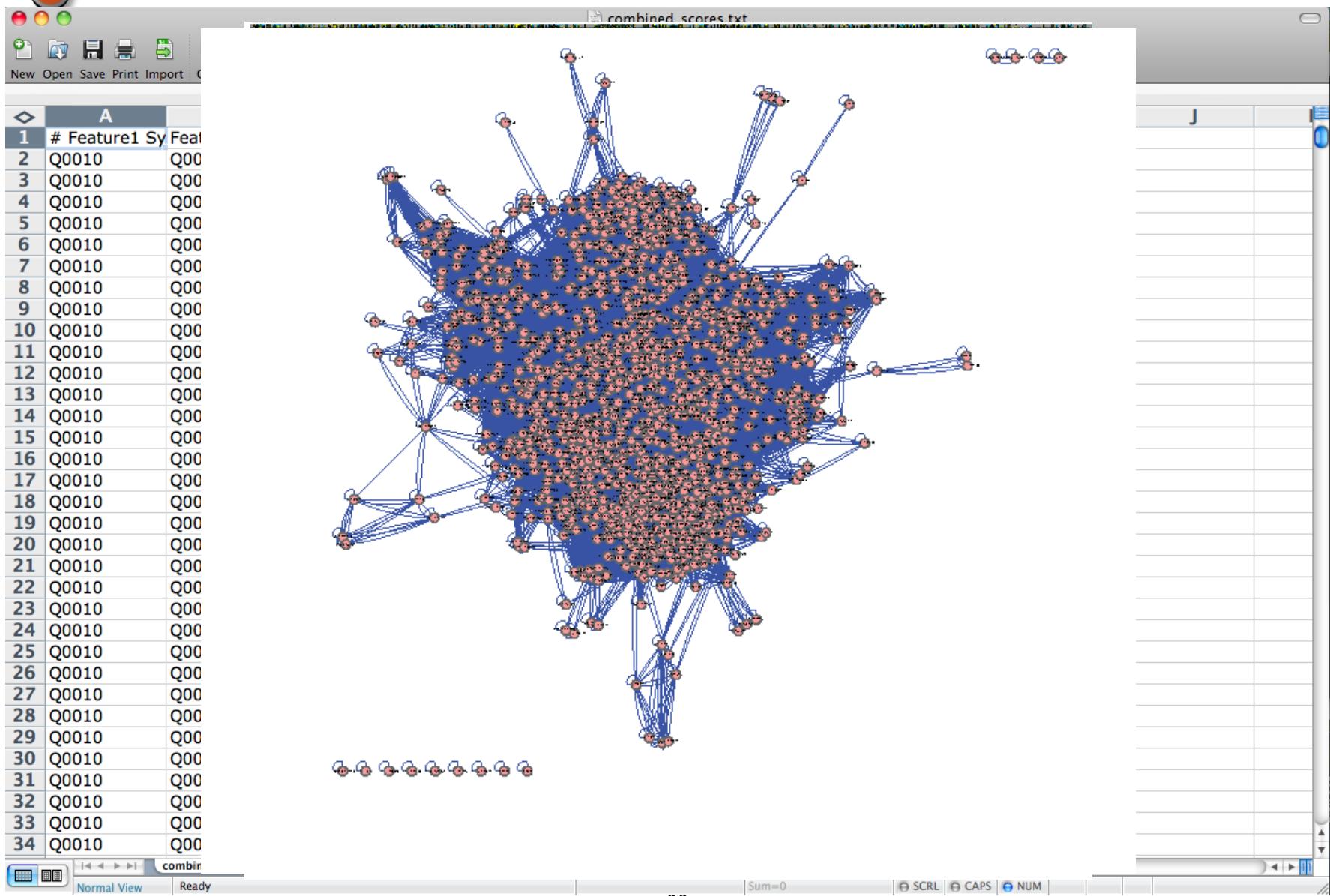
- **Subnetwork-based gene association** – map common pathway mechanisms affected by collection of genotypes (SNP, CNV)



PinnacleZ, UCSD



# The Challenge



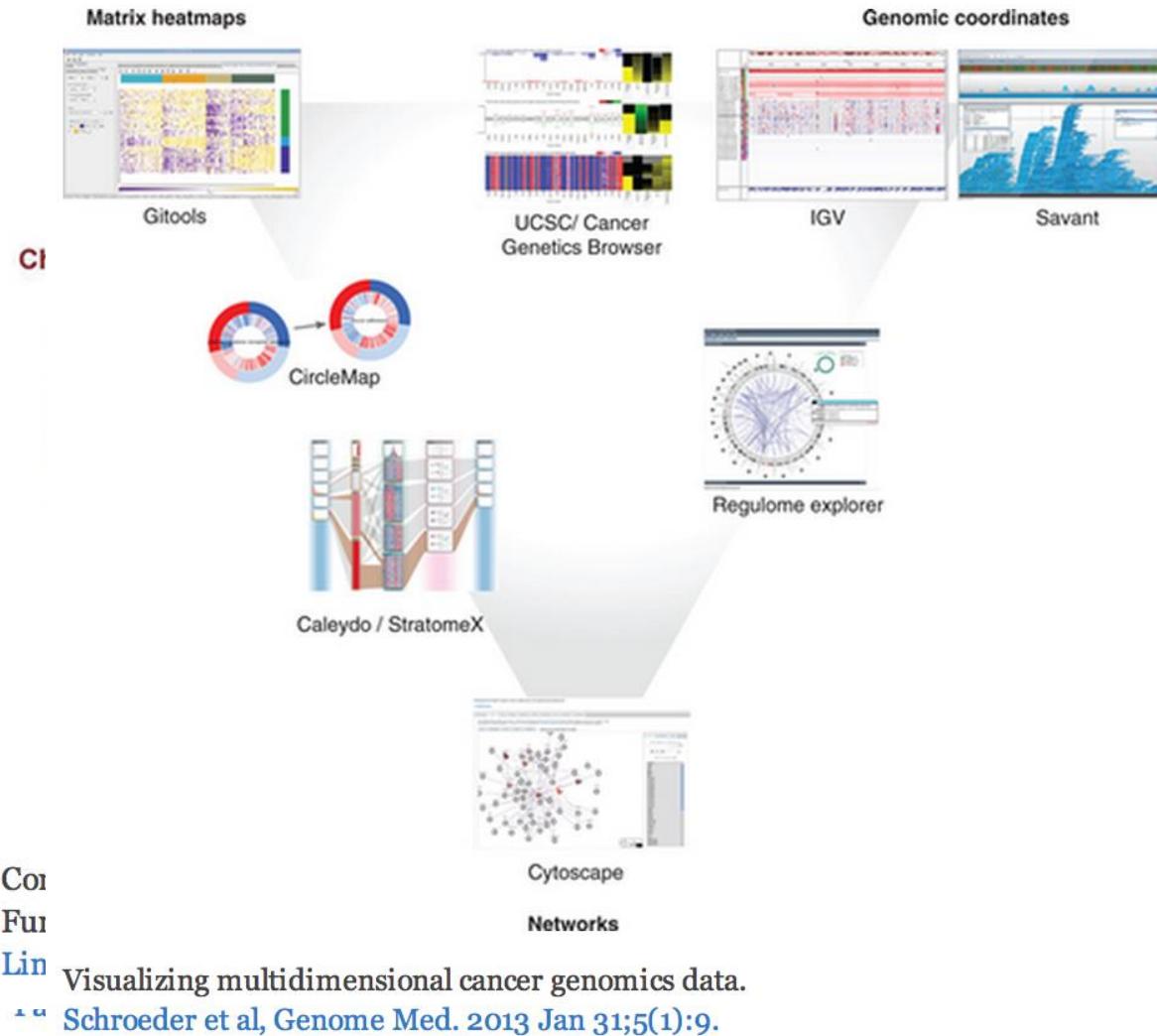


# The Challenge

- 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

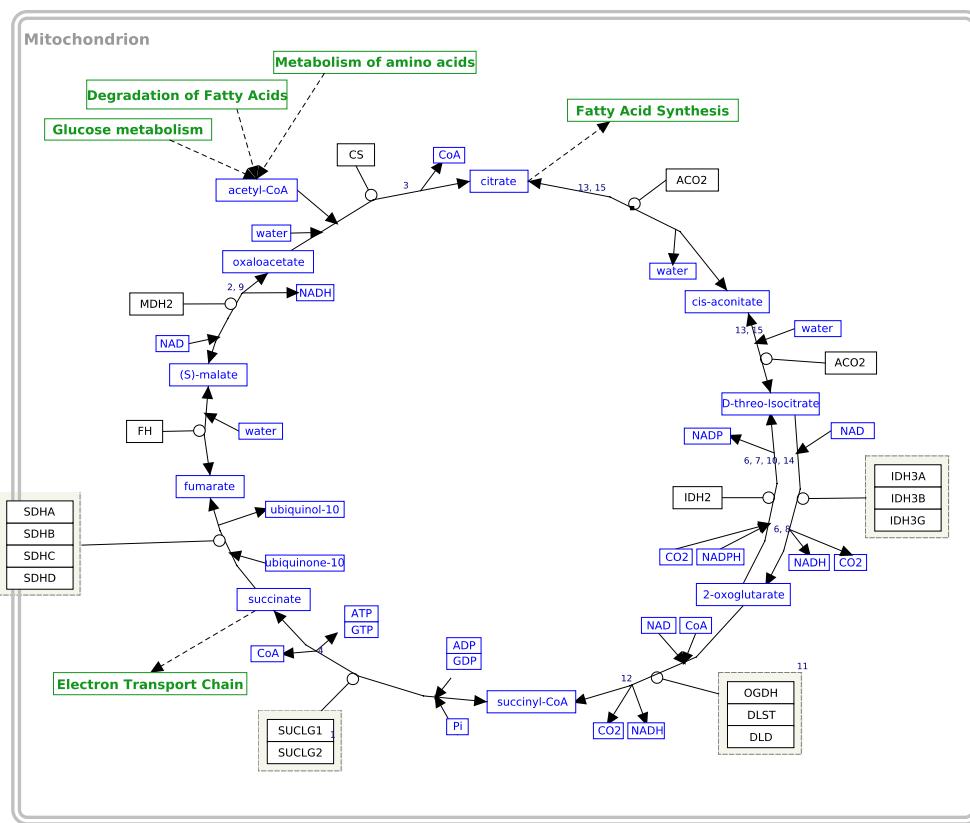


<http://cytoscape-publications.tumblr.com/archive>

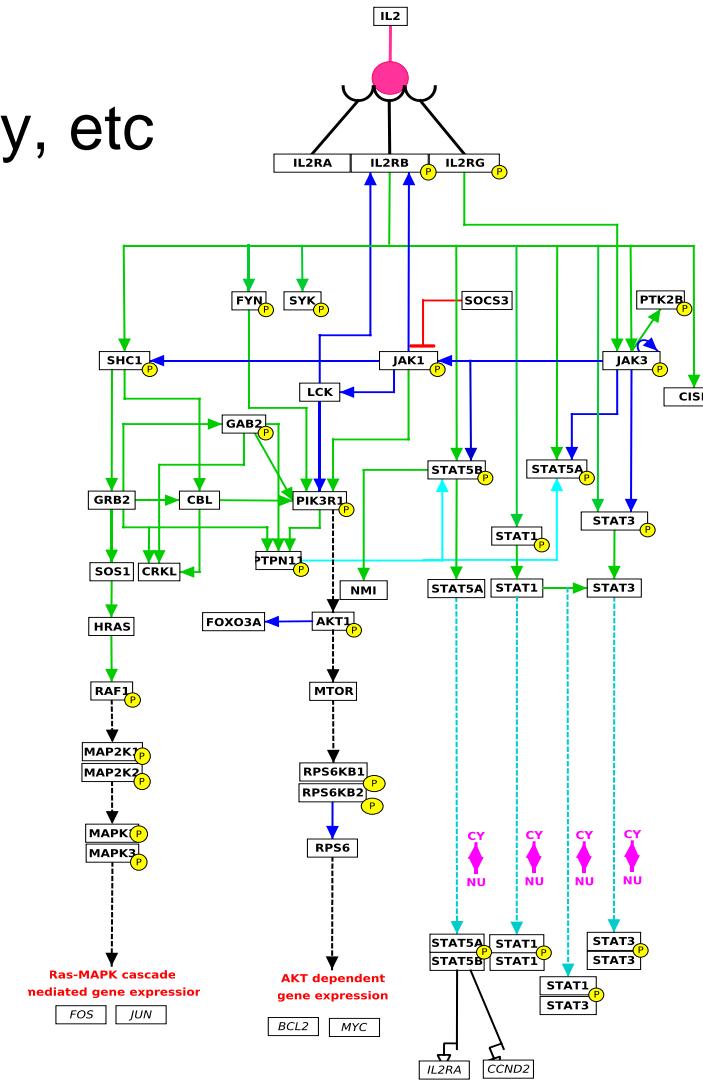


# Biological Network Taxonomy

- Pathways
  - Signaling, Metabolic, Regulatory, etc



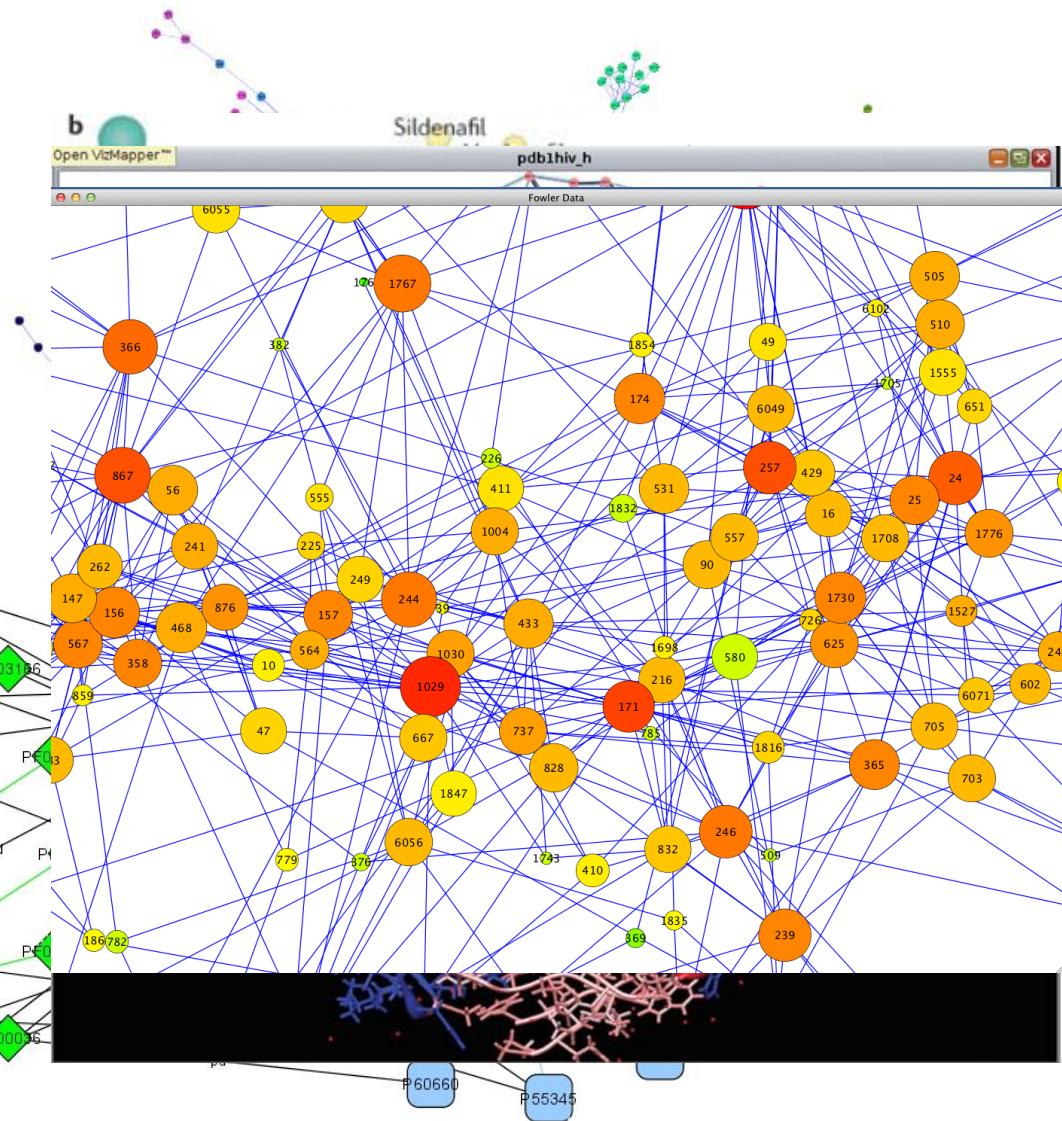
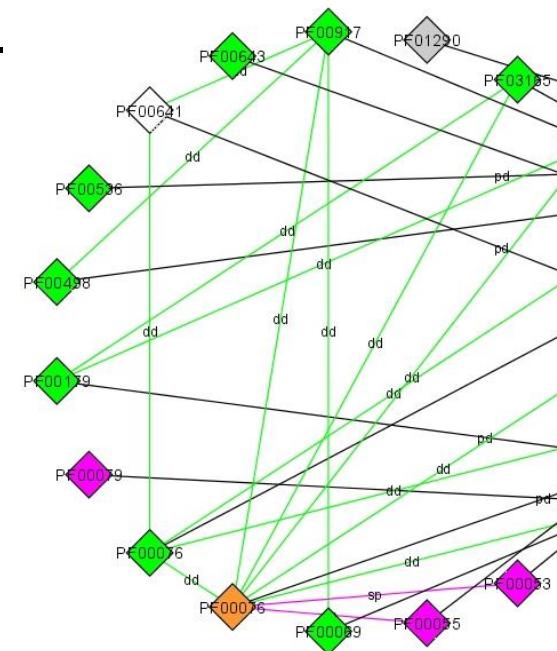
IL-2 Signaling Pathway





# Biological Network Taxonomy

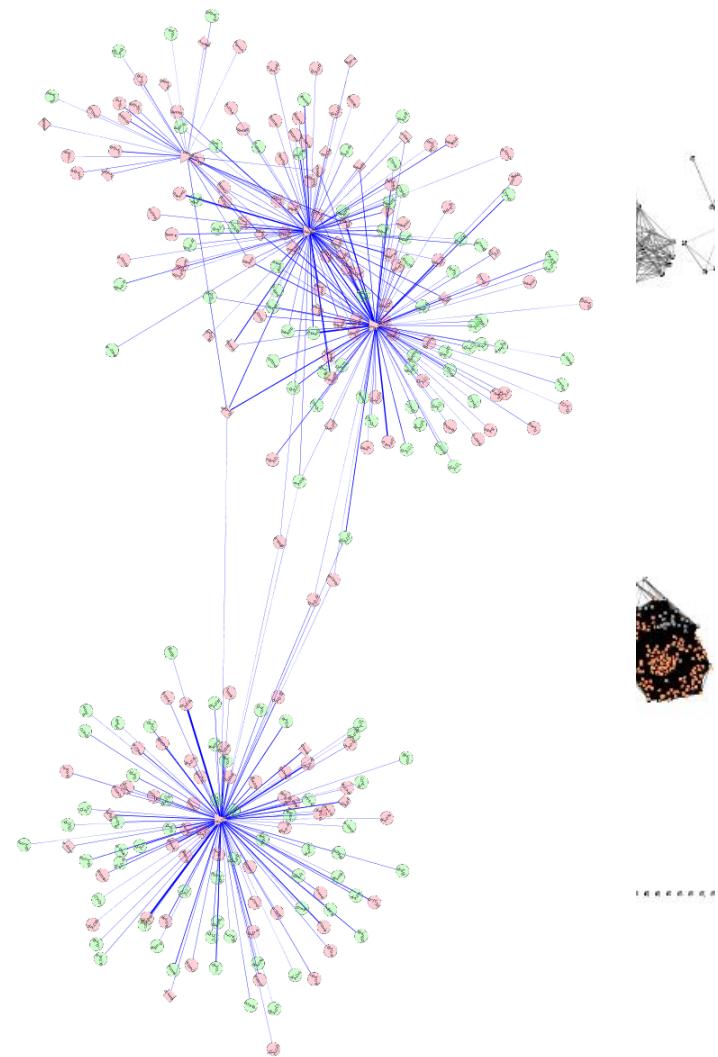
- Interactions
  - Protein-Protein
  - Protein-Ligand
  - Domain-Domain





# Biological Network Taxonomy

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

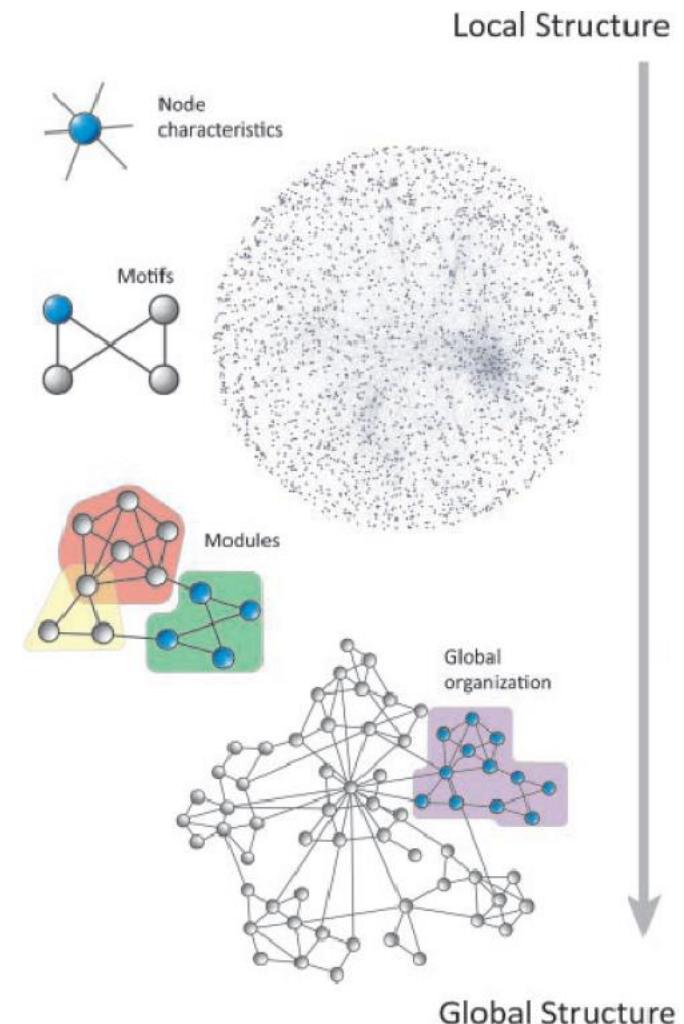




# Analytical Approaches

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.

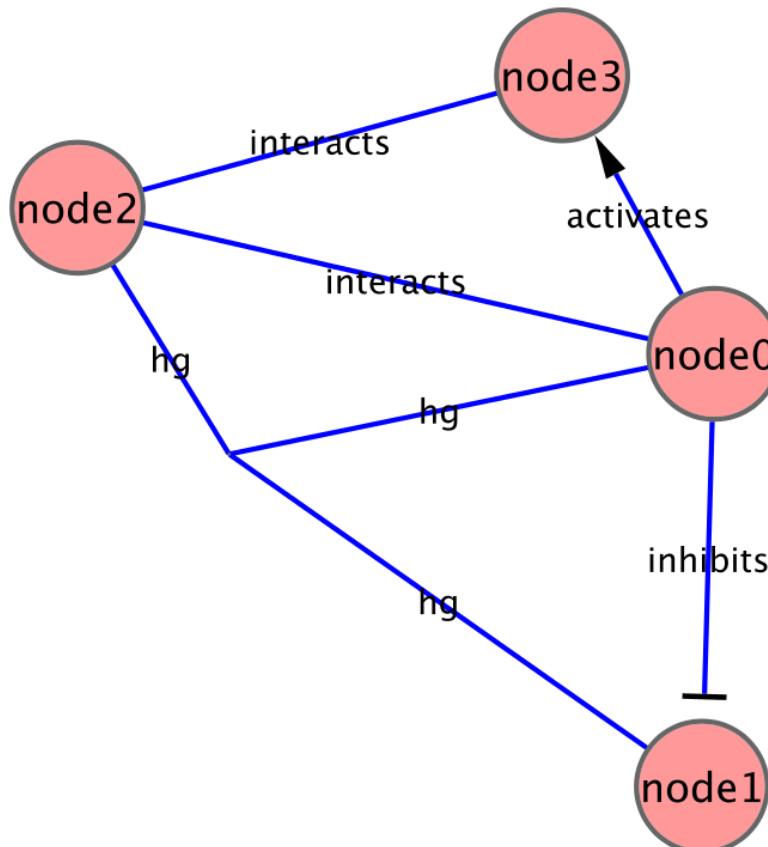
**Motif analysis** is the identification of small network patterns that are over-represented when compared with a randomized version of the same network. Regulatory elements are often composed of such motifs.

**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

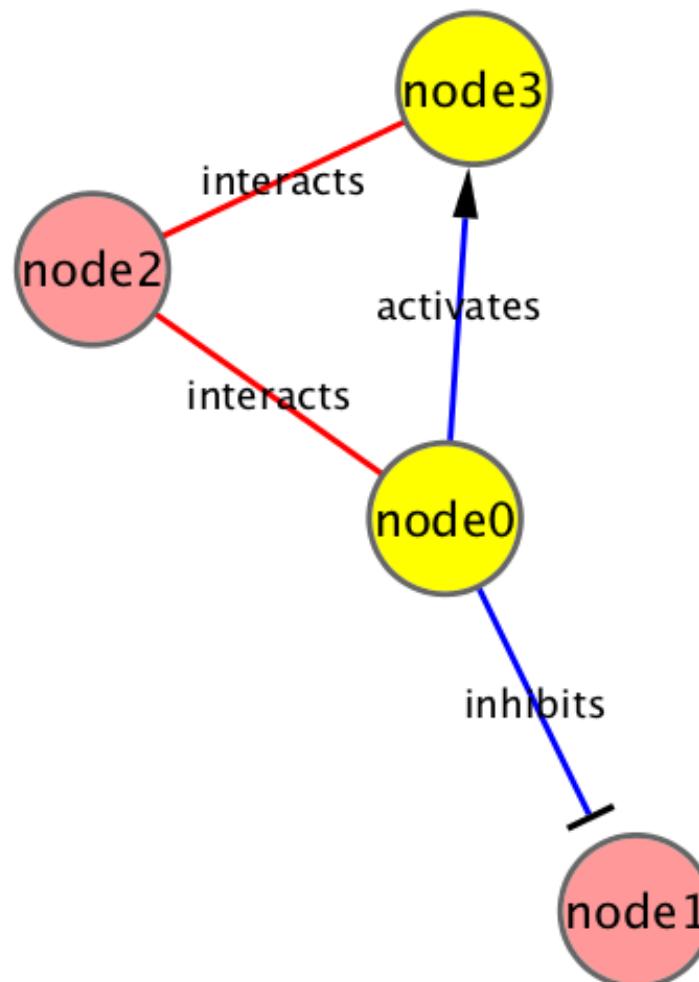
- 





# Analytical Approaches

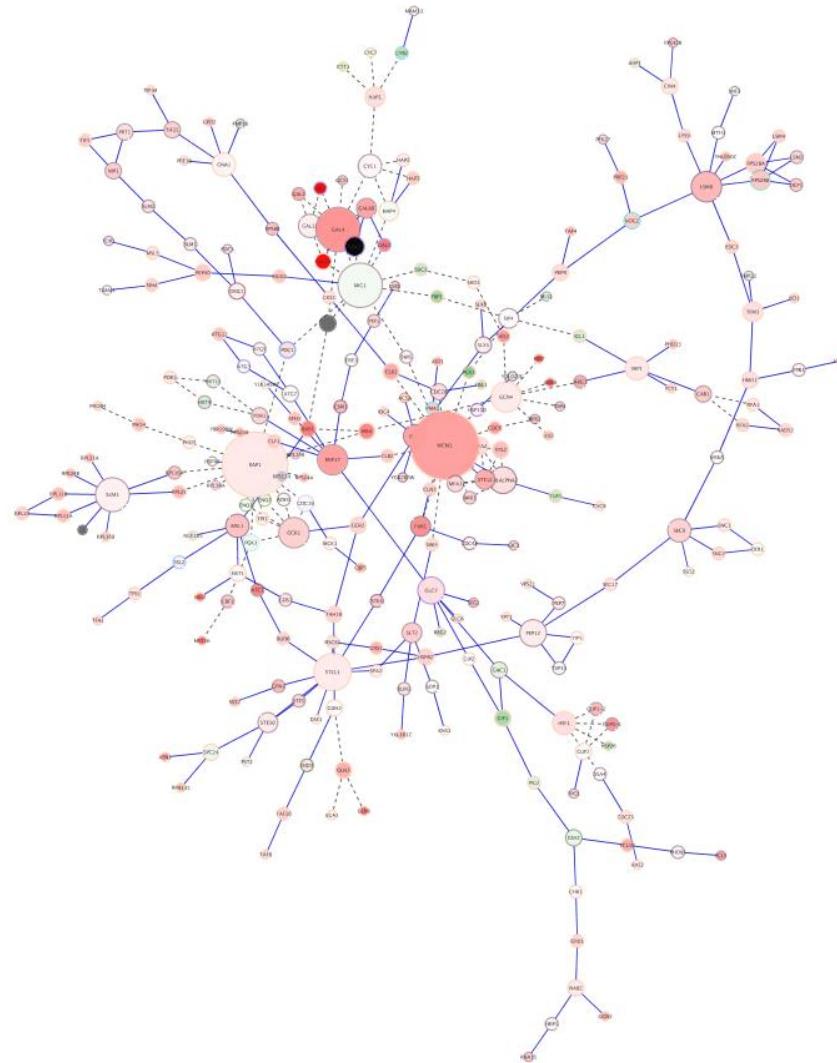
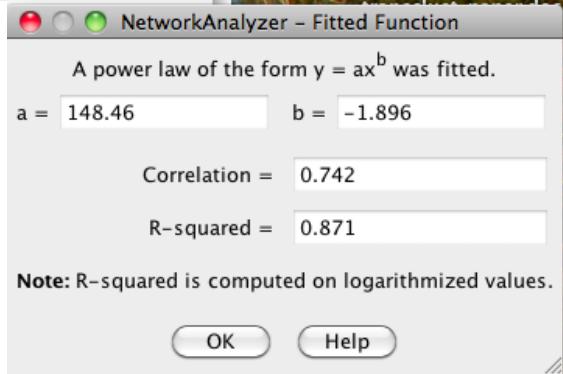
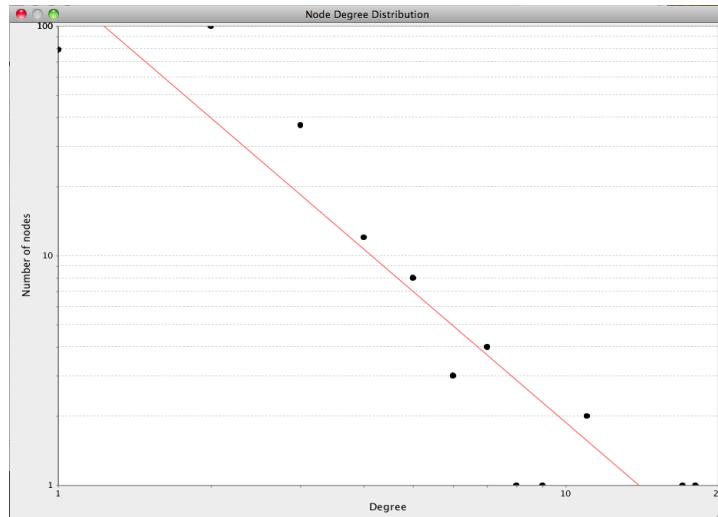
- N



not



# Analytical Approaches





# Analytical Approaches

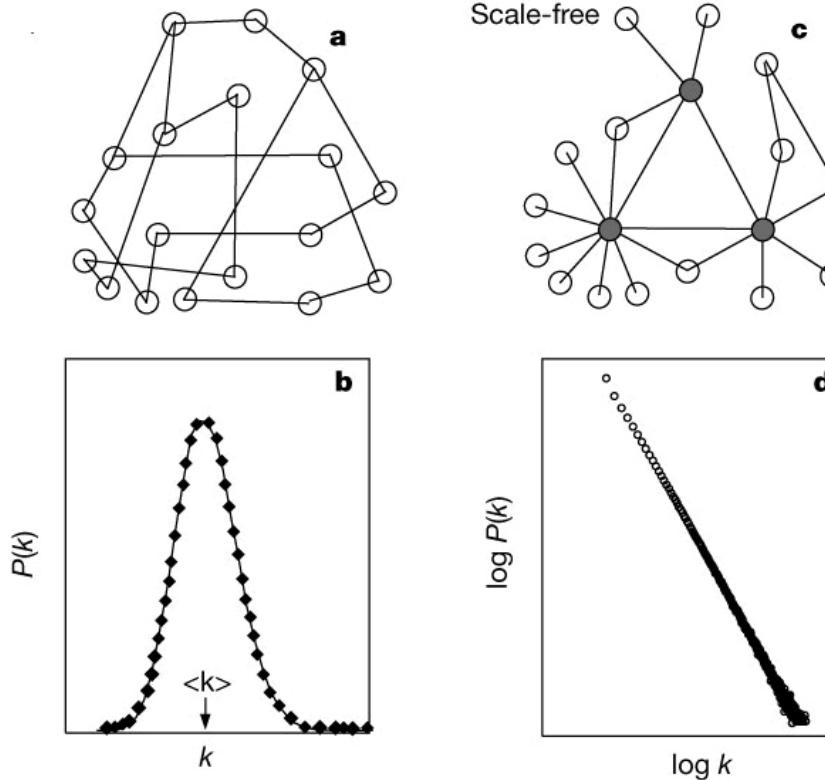
- Small-world networks
  - any two arbitrary nodes are connected by a small number of intermediate edges
  - the network has an average shortest path length much smaller than the number of nodes in the network  
(Watts, Nature, 1998).
  - Interaction networks have been shown to be small-world networks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





# Analytical Approaches

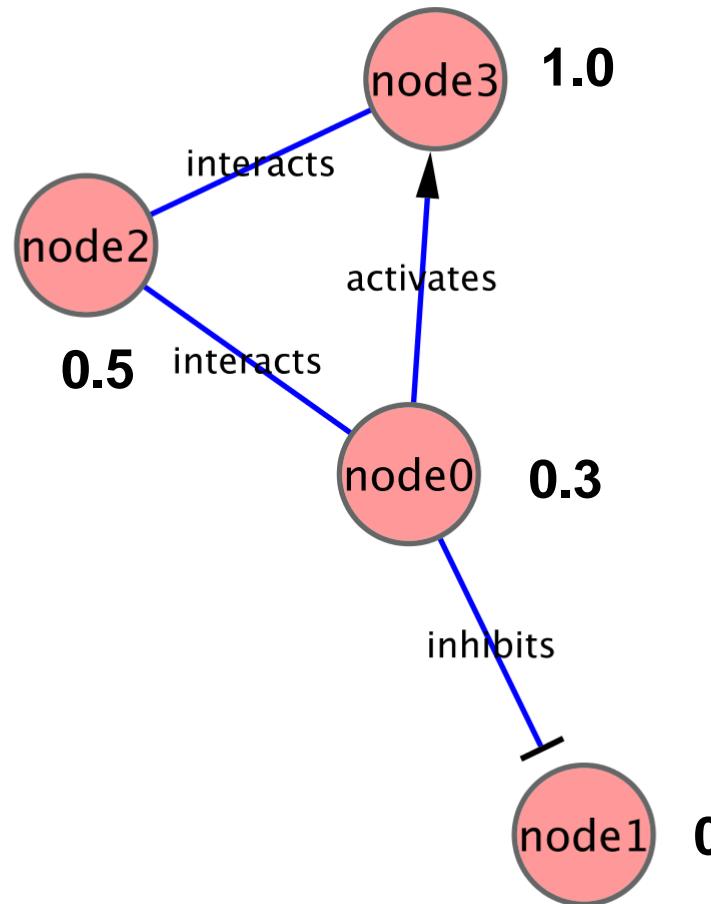
- Random networks
  - homogeneous, nodes have similar degrees, and not robust to arbitrary node failure
  - Algorithms
    - Flat random
    - Scale-free
    - Small-world
  - Useful to understand network





# Analytical Approaches

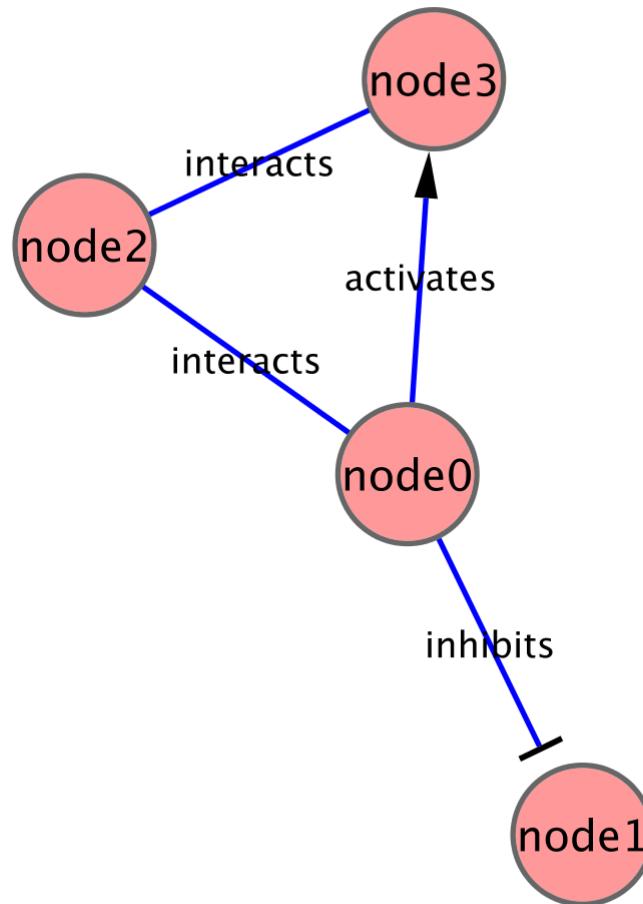
- N





# 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      |



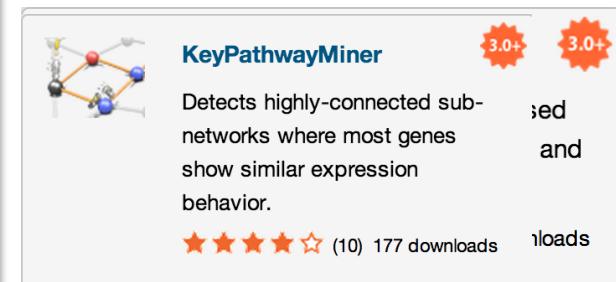
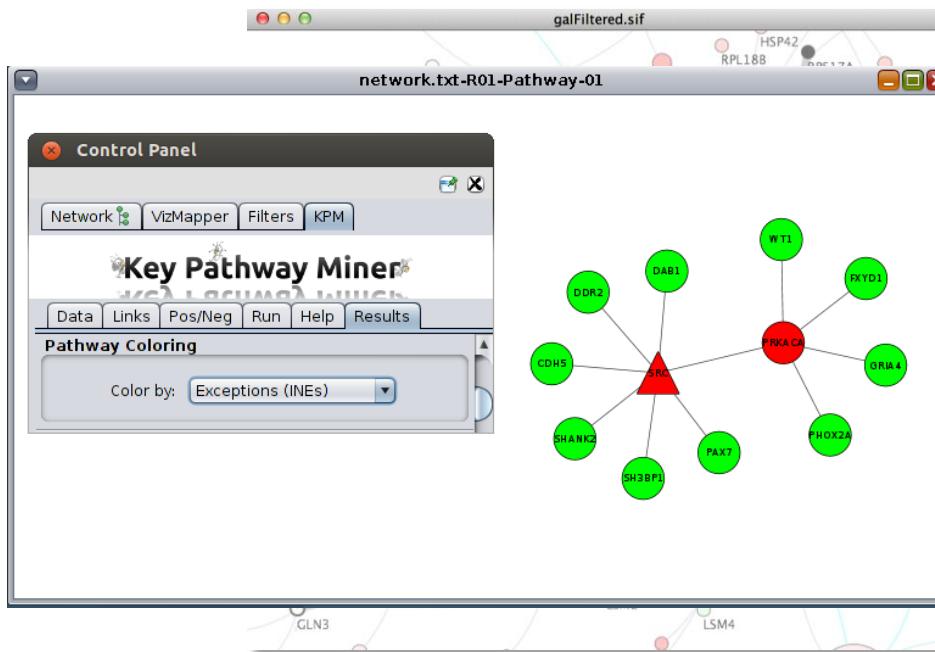
# Analytical Approaches

- Network Analyzer Demo...



# Analytical Approaches

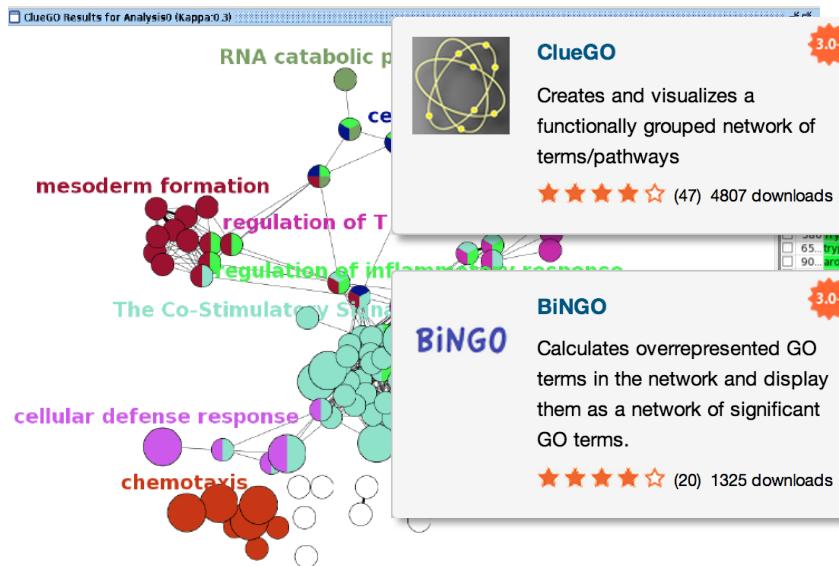
- Motif finding
  - Search directed networks for network motifs (feed-forward loops, feedback loops, etc.)





# Analytical Approaches

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

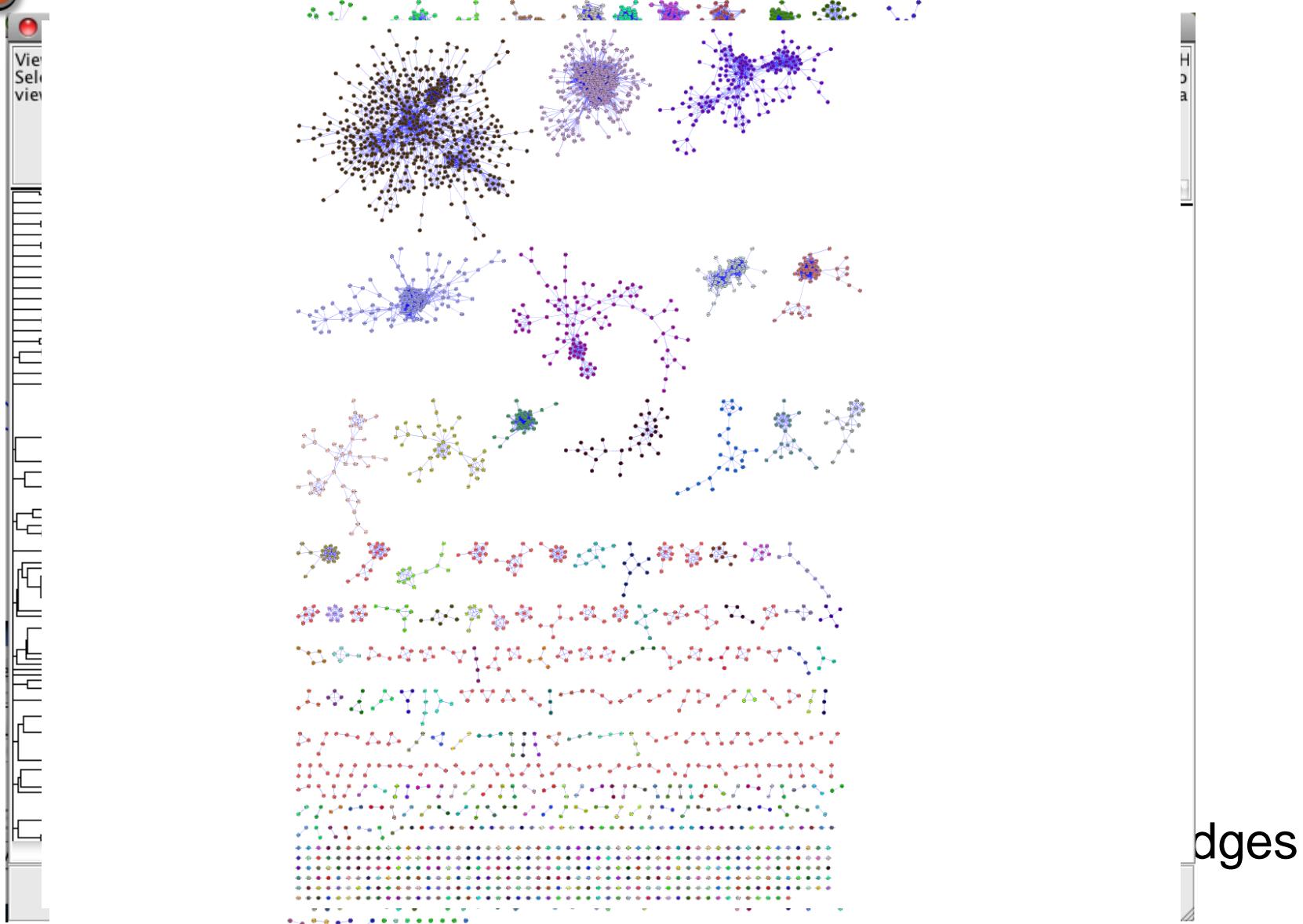


The Network Ontology Analysis (NOA) plugin performs ontology overrepresentation analysis based on the network connections among annotated nodes.

| Term PVal. | Term PValue | Group PVal. | 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]              |
| E-2        | 3.0 E-1     | 3.3 E-2     | 1.3 E-1      | [SMG1, UPF3B]              |



# Analytical Approaches



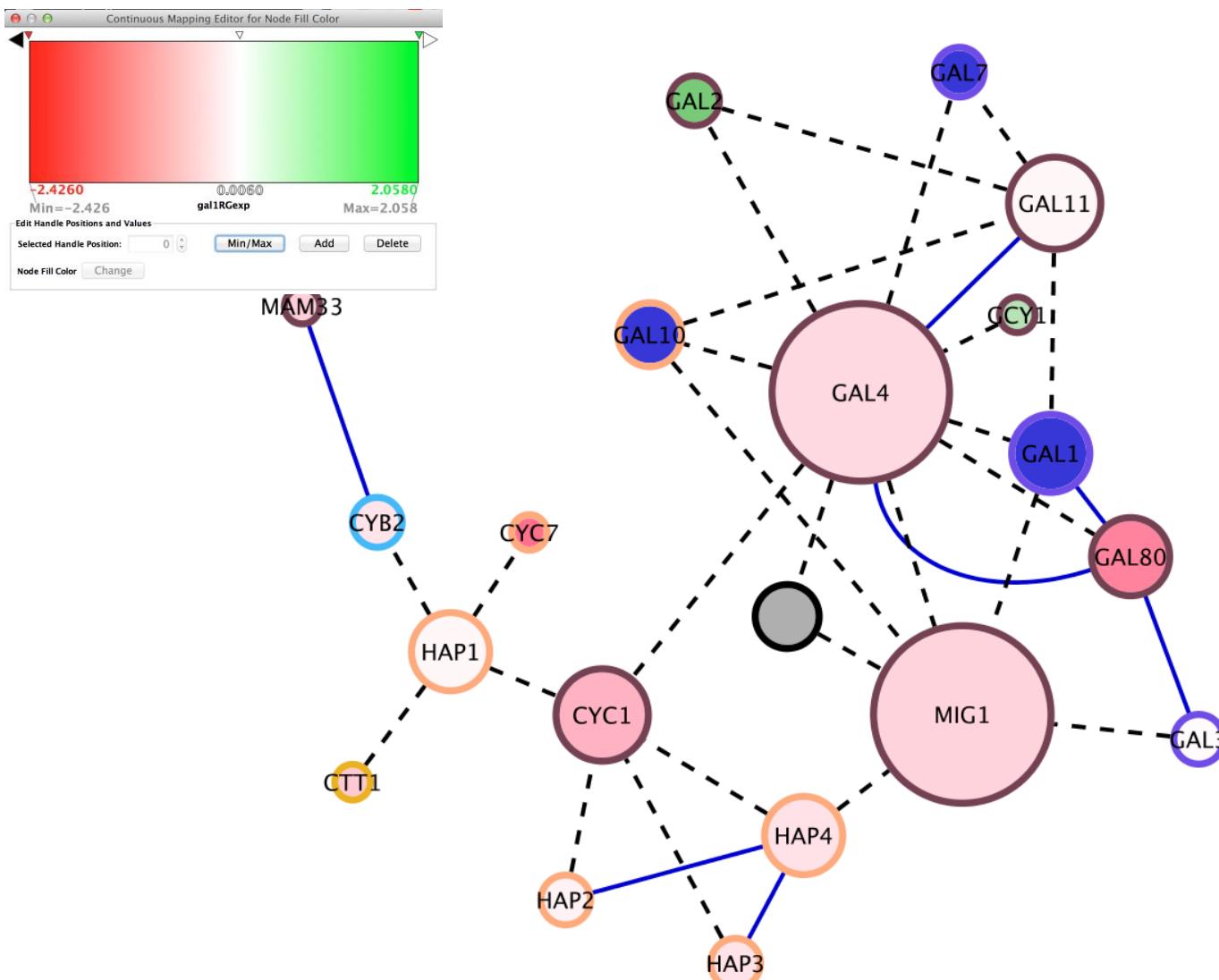


# Visualization of Biological Networks

- Data Mapping
- Layouts
- Animation



# Data Mapping



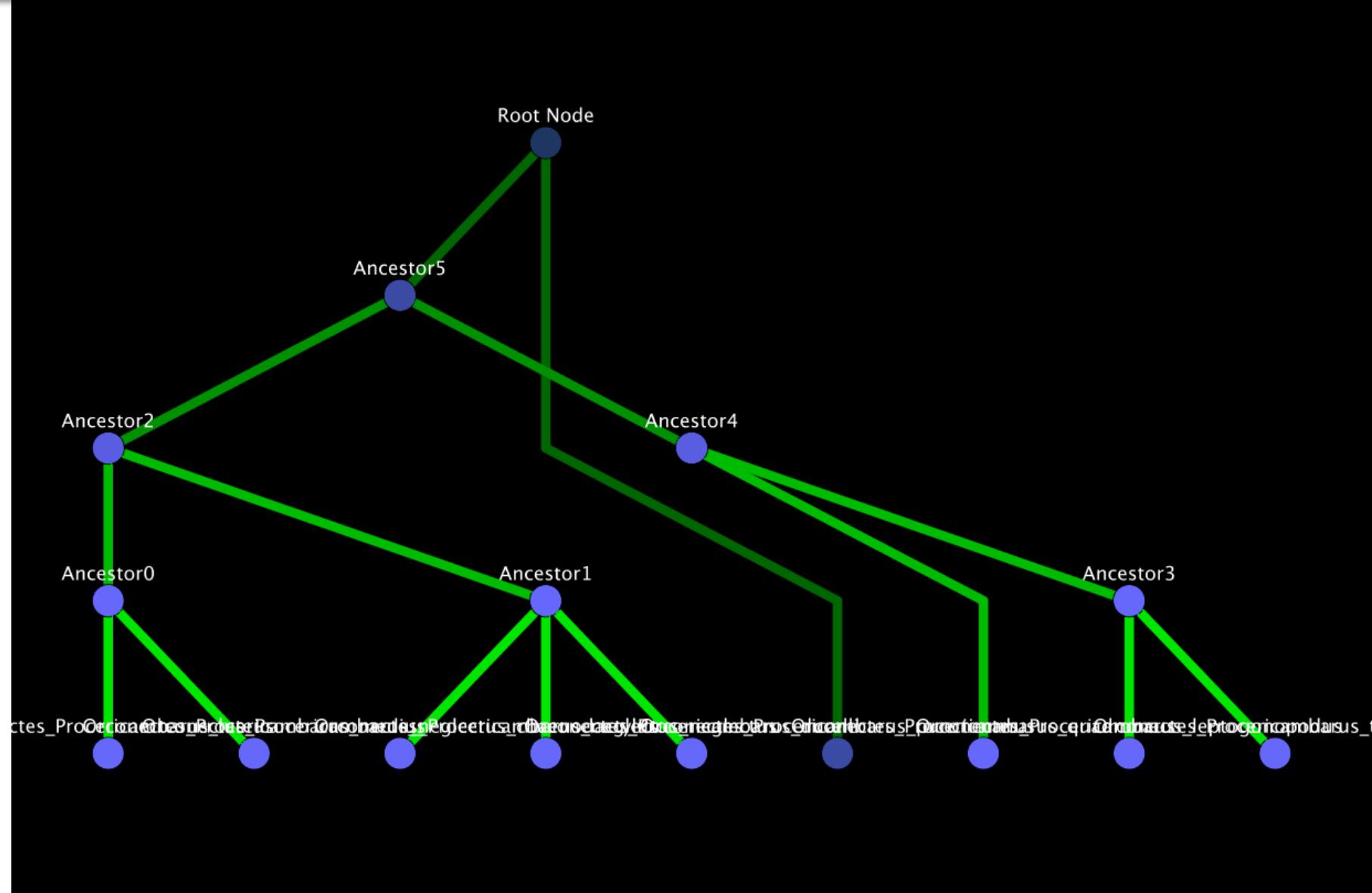


# Data Mapping

- 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

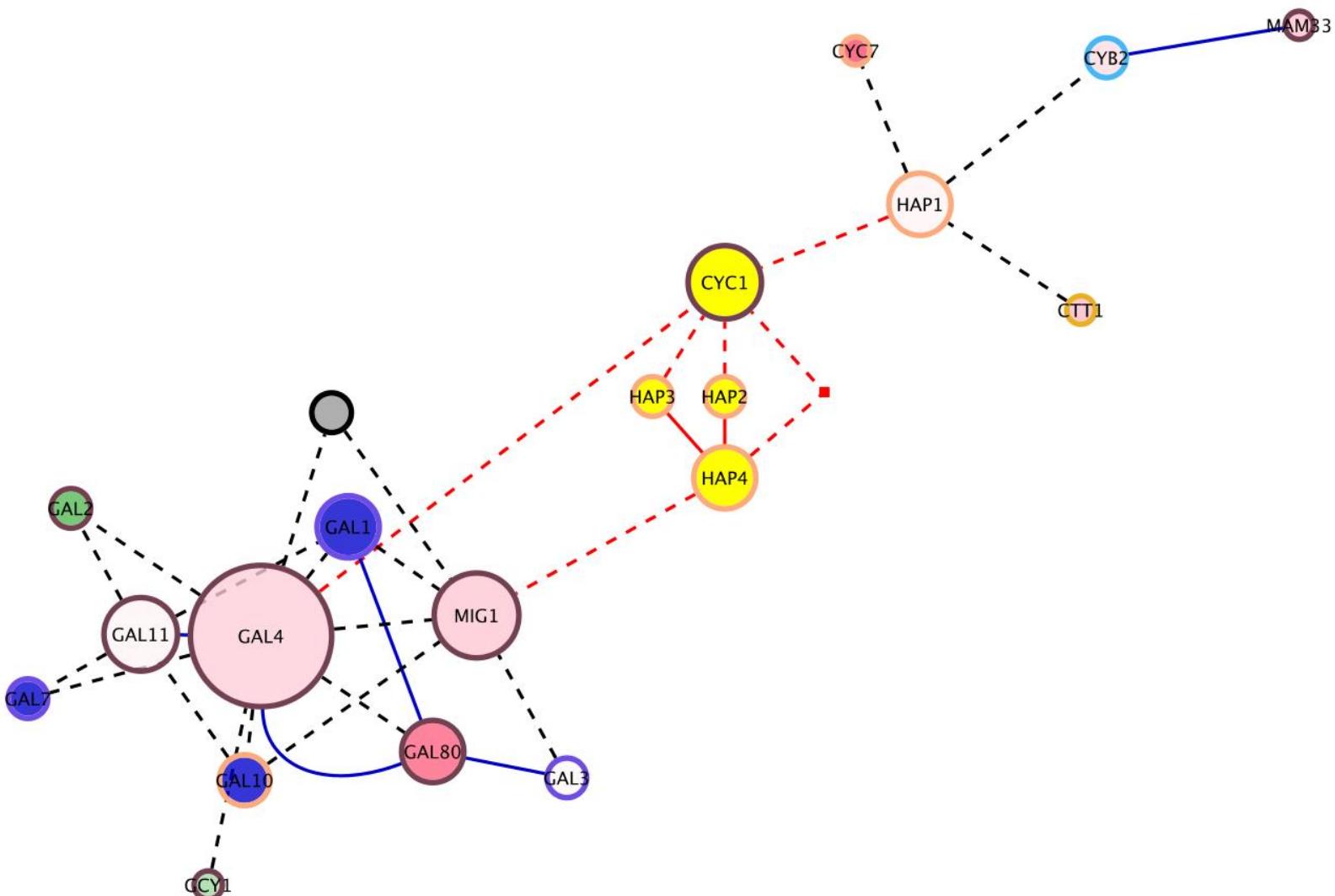


# Layouts





# Layouts



— Many, many owners

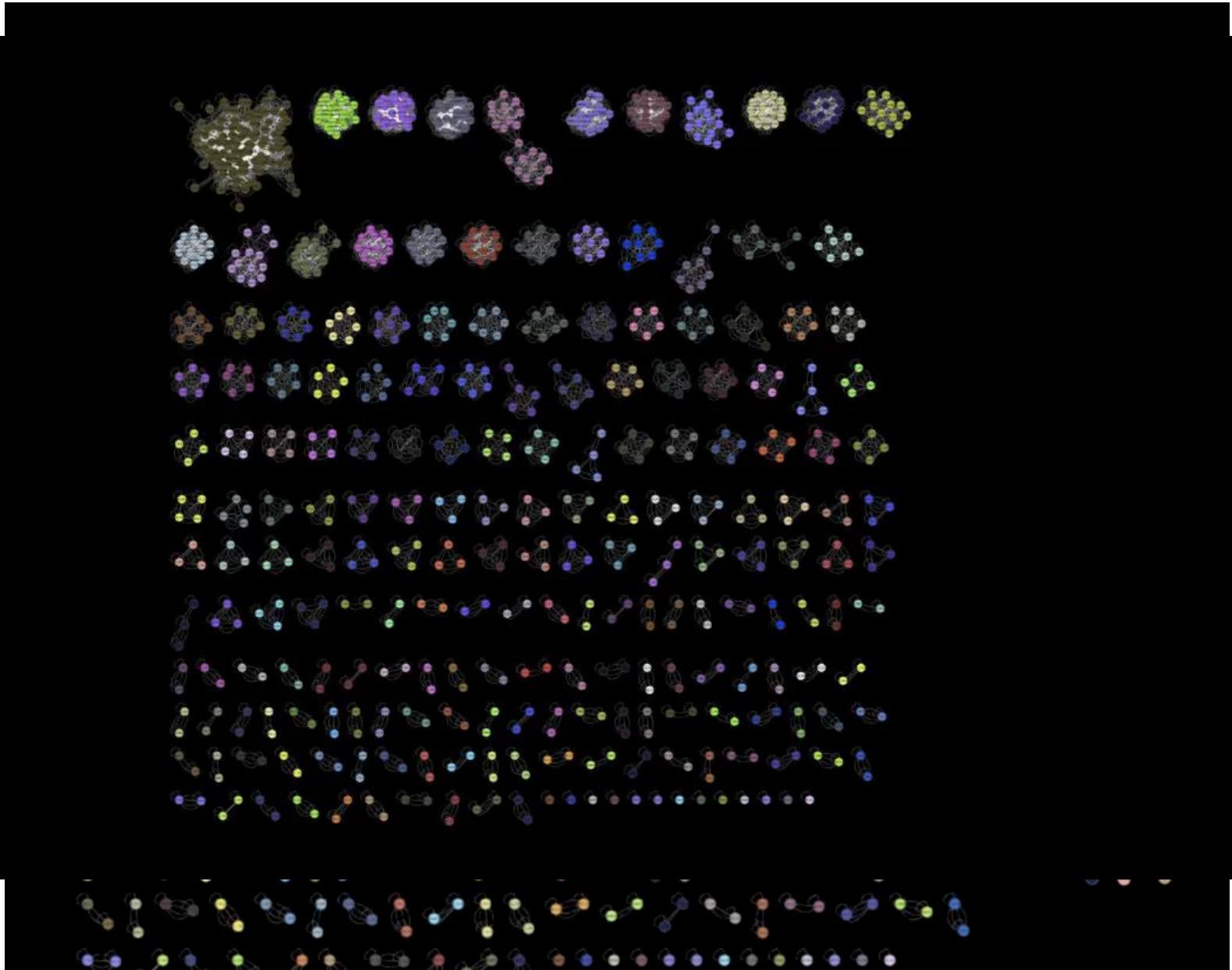


# Layouts

- 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: A color scale from blue (-2.43) to red (2.06)

Remove Visual Mapping

Height: 18

Pm Label

Label Color: 18

Label Font Size: 18

Shape: Circle

Size: Cm

Column: BetweennessCentrality Mapping Type: Continuous Mapping

Current Mapping: A color scale from blue (0.0) to red (200.0)

Node Edge Network

Table Panel

galFiltered.sif

| name    | Average... | Degree | Between... | gal4R...   | gal80R...  | gal80R... | COMM... |
|---------|------------|--------|------------|------------|------------|-----------|---------|
| YDL194W | 13.1169... | 1      | 0.0        | 0.333      | 0.011348   | 0.449     | SNF3    |
| YDR277C | 12.1209... | 2      | 0.00806... | 0.192      | 5.727E-4   | 0.448     | MTH1    |
| YBR043C | 1.5        | 1      | 0.0        | 0.023      | 0.999999   | 0.0       | YBR043C |
| YPR145W | 9.79838... | 1      | 0.0        | -0.614     | 0.00118... | -0.232    | ASN1    |
| YER054C | 8.81854... | 2      | 0.04427... | 0.206      | 0.00436... | 0.247     | GIP2    |
| YDR045C | 0.00116... | 1      | 0.00020... | 0.00020... | 0.00020... | 0.001     | GIP2    |

Node Table Edge Table Network Table

Read Session File

Memory: OK

- 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

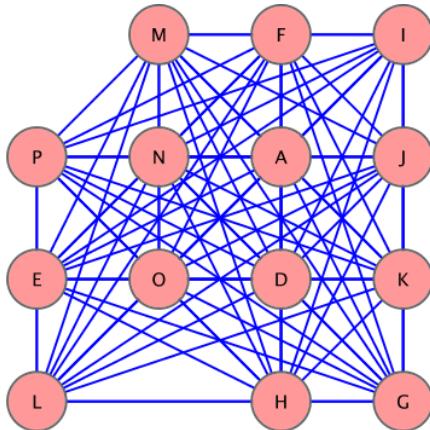
GLADSTONE  
INSTITUTES

NRNB



# Core Concepts

- Networks and Tables



|    | A           | B         | C       | D           | E         | F       | G                                | H  | I            | J     |
|----|-------------|-----------|---------|-------------|-----------|---------|----------------------------------|--|--------------|-------|
|    | GI_Number   | Gene      | HPRD_ID | REFSEQ      | Entrez_ID | OMIM_ID | SWISS_Gene_Description           | Architecture                             | Arch         | Motif |
| 1  | gi 19923959 | SDSL      | 11542   | NP_614411.1 | 113675    |         | SDSL                             | SP                                       | SP           |       |
| 2  | gi 14602926 | SEC23B    | 11543   | NP_617914.1 | 103493    | Q15437  | Secain transport protein SEC23B  | CEL                                      | CEL          | Doma  |
| 3  | gi 14602492 | SCMH1     | 11540   | AAH09752.1  | 22955     | Q96CD1  | SCMH1                            | SAM                                      | SAM          | Doma  |
| 4  | gi 14602492 | EXOC2     | 11544   | NP_060773.3 | 55770     | Q96KP1  | SECS like 1                      | IPT                                      | IPT          | Doma  |
| 5  | gi 1361625  | SERPINB12 | 11548   | NP_536722.1 | 89777     | Q96PF3  | Serpin B12                       | SERPIN                                   | SERPIN       | Doma  |
| 6  | gi 1798551  | EIF2AK4   | 18447   | NP_00101372 | 440275    | 609280  | Similar to GCN2 eIF2alpha kinase | S_T_Y_Kinase                             | S_T_Y_Kinase | Doma  |
| 7  | gi 5287717  | TYROB     | 4996    | NP_003323.1 | 7305      | 604142  | DAP12                            | ITAM                                     | ITAM         | Doma  |
| 8  | gi 4507755  | IL24      | 4995    | NP_006841.1 | 11009     | 604136  | Q13007                           | Interleukin 24                           | IL10         | Doma  |
| 9  | gi 5803086  | ADAMTS13  | 4994    | NP_620594.1 | 11093     | 604134  | ADAMTS 13                        | TSP1                                     | TSP1         | Doma  |
| 10 | gi 12165034 | PIP5K1A   | 4470    | AAC50911.1  | 8394      | 603275  | Q99755                           | Phosphatidylinositol-4-phosphate 5 PIPKc | PIP5Kc       | Doma  |
| 11 | gi 1743873  | SLC16A2   | 4992    | NP_006508.1 | 6600      | 603276  | Q95425                           | Surveillance                             | VMP          | Doma  |
| 12 | gi 1463386  | SUJ1T2B1  | 4991    | NP_144444.1 | 6820      | 604125  | Suppressor of                    | LZ                                       | LZ           | Motif |
| 13 | gi 1463386  | RBBP8     | 4990    | AAH30590.1  | 5932      | 604124  | Q99708                           | VEGF D                                   | PDGF         | Doma  |
| 14 | gi 21040399 | FIGF      | 2102    | NP_004460.1 | 2277      | 300091  | Q43915                           | Collapsin response mediator protein 3    | CRMP3        | Doma  |
| 15 | gi 758378   | DPYSL4    | 7463    | NP_006417.1 | 10570     | 608407  | Q14531                           | Threonine aspartate 1                    | TM           | Doma  |
| 16 | gi 1321617  | TASP1     | 7460    | NP_060184.1 | 55617     | 608270  | Q9H6P5                           | Signal sequence receptor delta           | SP           | Motif |
| 17 | gi 8923202  | SSR4      | 2101    | NP_006271.1 | 6748      | 300090  | P51571                           | X linked PEST containing transport       | UBC          | Doma  |
| 18 | gi 5454090  | SLC16A2   | 2106    | NP_065083.1 | 6567      | 300095  | P36021                           | Protein kinase C, iota type              | S_T_kinase   | Doma  |
| 19 | gi 5730045  | FTS       | 7467    | NP_071921.1 | 64400     | 608483  | Q9H870                           | TM                                       | TRNA         | Doma  |
| 20 | gi 1968027  | PRKCI     | 2105    | NP_002731.3 | 5584      | 600539  | P41743                           | Protein kinase C, iota type              | TRNA         | Doma  |
| 21 | gi 48255885 | PPP1R2P9  | 6591    | NP_079482.1 | 29925     |         |                                  |  |              |       |
| 22 | gi 55576812 | CDCC22    | 6593    | NP_079483.1 | 80316     |         |                                  |  |              |       |
| 23 | gi 176812   | PRAF2     | 6595    | NP_062375.1 | 90066     |         |                                  |  |              |       |
| 24 | gi 50626862 | CCDC22    | 6594    | NP_054727.1 | 28952     |         |                                  |  |              |       |
| 25 | gi 7661844  | PRAF2     | 6596    | NP_009144.1 | 11230     |         |                                  |  |              |       |
| 26 | gi 6005794  |           |         |             |           |         |                                  |  |              |       |

## Networks

e.g., PPIs or pathways

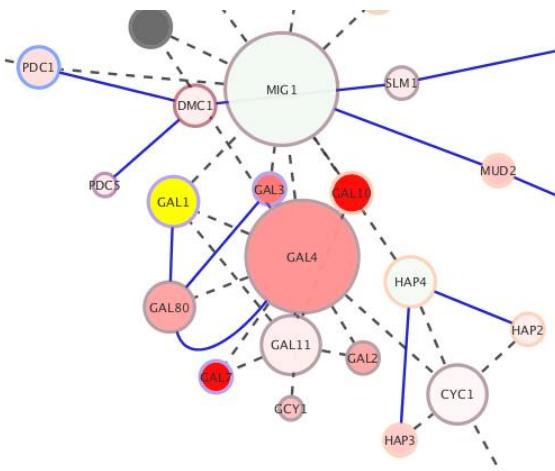
## Tables

e.g., data or annotations



# Core Concepts

- Networks and Tables



|    | A          | B                                 | C       | D           | E         | F       | G                                   | H  | I          | J     |  |
|----|------------|-----------------------------------|---------|-------------|-----------|---------|-------------------------------------|--|------------|-------|--|
|    | GI_Number  | Gene                              | HPRD_ID | REF_ID      | Entrez_ID | OMIM_ID | SWISS_Gene_Description              | Architecture                             | Arch       |       |  |
| 1  | g19923959  | SDSL                              | 11542   | NP_604411.1 | 113675    |         | SDSL                                | SP                                       | Motif      |       |  |
| 2  | g11462926  | SEC23B                            | 11543   | NP_604791.1 | 103493    | Q15437  | Serine transport protein SEC23B     | CEL                                      | Dom        |       |  |
| 3  | g14602492  | SCMH1                             | 11540   | AAH09752.1  | 22955     | Q96GDI  | SCMH1                               | SAM                                      | Dom        |       |  |
| 4  | g14602492  | EXOC2                             | 11544   | NP_060773.3 | 55770     | Q96KPL  | SEC61 like 1                        | IPT                                      | Dom        |       |  |
| 5  | g121361625 | SERPINB12                         | 11548   | NP_536722.1 | 89777     | Q96PF3  | Serpin B12                          | SERPIN                                   | Dom        |       |  |
| 6  | g17998551  | EIF2AK4                           | 18447   | NP_00101372 | 440275    | 609280  | Similar to GCN2 eIF2alpha kinase    | S_T_Y_Kinase                             | Dom        |       |  |
| 7  | g165287717 | TYROBP                            | 4996    | NP_003323.1 | 7305      | 604142  | DAP12                               | ITAM                                     | Dom        |       |  |
| 8  | g1507755   | IL24                              | 4995    | NP_006841.1 | 11009     | 604136  | Q13007                              | Interleukin 24                           | IL10       | Dom   |  |
| 9  | g15803086  | ADAMTS13                          | 4994    | NP_620594.1 | 11093     | 604134  | ADAMTS 13                           | TSP1                                     | Dom        |       |  |
| 10 | g121265034 | PIP5K1A                           | 4470    | AAC50911.1  | 8394      | 603275  | Q99755                              | Phosphatidylinositol-4-phosphate 5 PIPKc | Dom        |       |  |
| 11 | g1743873   | PIP5K1B                           | 4992    | NP_00101373 | 6000      | 603276  | Q95425                              | Survivin                                 | VIM        | Dom   |  |
| 12 | g11462926  | RBBP8                             | 4991    | NP_144444.1 | 6820      | 604125  | Subtransferease family 2B, member 1 | LZ                                       | Motif      |       |  |
| 13 | g163386    | SUJ1T2B1                          | 4990    | AAH30590.1  | 5932      | 604124  | Q99708                              | DPGf                                     | Dom        |       |  |
| 14 | g121040399 | FIGF                              | 2102    | NP_004460.1 | 2277      | 300091  | Q43915                              | VEGF D                                   | ITAM       | Dom   |  |
| 15 | g14758378  | DPYSL4                            | 7463    | NP_006417.1 | 10570     | 608407  | Q14531                              | Collapsin response mediator protein 3    | Dom        |       |  |
| 16 | g11321617  | TASP1                             | 7460    | NP_601184.1 | 55617     | 608270  | Q9H6P5                              | Threonine aspartate 1                    | PDGF       |       |  |
| 17 | g18923202  | SSR4                              | 2101    | NP_006271.1 | 6748      | 300090  | P51571                              | Signal sequence receptor delta           | SP         | Motif |  |
| 18 | g15454090  | SLC16A2                           | 2106    | NP_006508.1 | 6567      | 300095  | P36021                              | X linked PEST containing transport       | TM         | Dom   |  |
| 19 | g15730045  | FTS                               | 7467    | NP_071921.1 | 64400     | 608483  | Q9H8T0                              | FTS                                      | UBC        | Dom   |  |
| 20 | g11968027  | PRKC1                             | 2105    | NP_002731.3 | 5584      | 600539  | P41743                              | Protein kinase C, iota type              | S_T_kinase | Dom   |  |
| 21 | g18255885  | RPA4                              | 6591    | NP_002731.1 | 29925     |         | TPA                                 | TRNA                                     | Dom        |       |  |
| 22 | g155576812 | PPP1R2P9                          | 6593    | NP_079486.1 | 80316     |         |                                     |  |            |       |  |
| 23 | g155576812 | TM1 protein phosphatase inhibitor | 6595    | NP_206375.1 | 90060     |         | JM11 protein                        | CC                                       | Motif      |       |  |
| 24 | g158268682 | CCDC22                            | 6594    | NP_054727.1 | 28952     |         | JM1 protein                         | CC                                       | Motif      |       |  |
| 25 | g17661844  | PRAF2                             | 6596    | NP_009144.1 | 11230     |         | JM4 protein                         | TM                                       | Dom        |       |  |
| 26 | g16005794  |                                   |         |             |           |         |                                     |  |            |       |  |

Networks

Tables

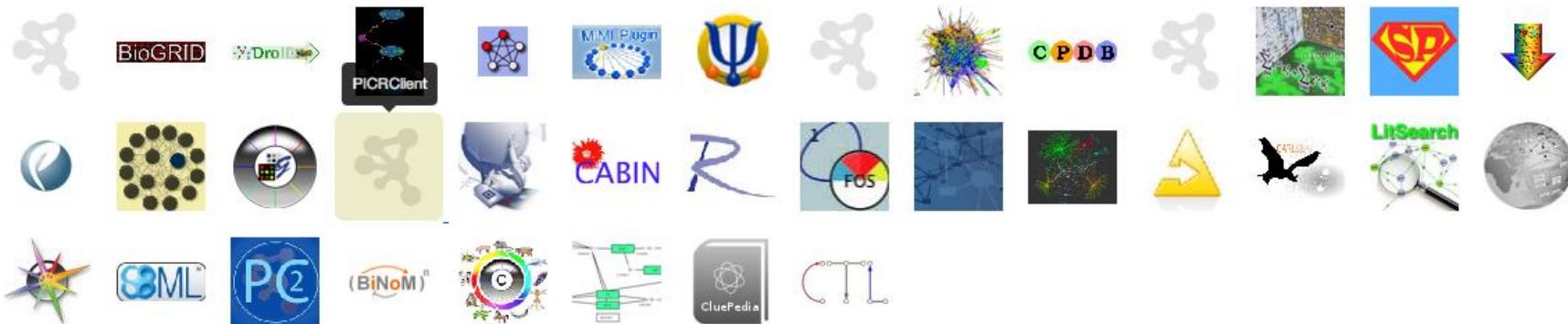
Visual Styles

## Wall of Apps 173 total

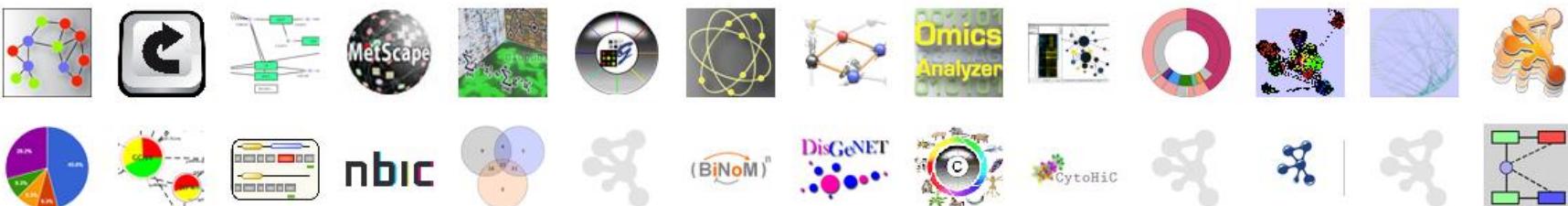
## network generation



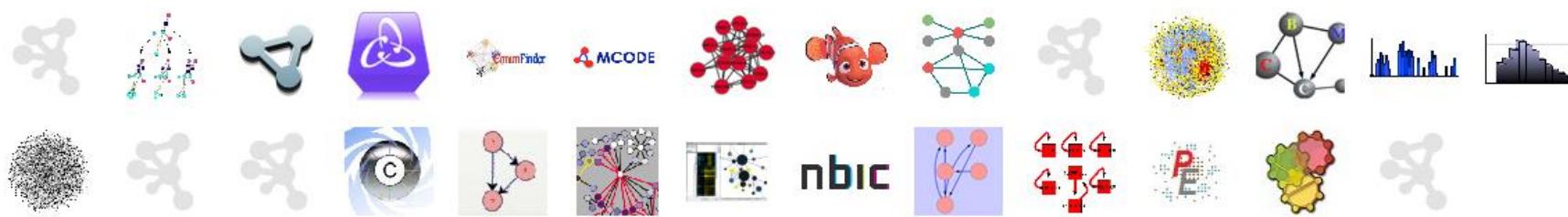
online data  
import



## data visualization



## graph analysis





# 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

The screenshot shows the Cytoscape application interface. At the top, there's a menu bar with 'Cytoscape', 'File', 'Edit', 'View', 'Select', 'Layout', 'Apps', 'Tools', and 'Help'. Below the menu is a toolbar with various icons for file operations like opening, saving, and zooming. A search bar labeled 'search term...' is also present.

The main workspace contains three panels:

- Control Panel:** Shows a tree view of networks. Under 'Network', it lists 'Yeast Network Sheet 1' and 'Yeast Network Sheet 33... 36...'. Below this, it says 'Yeast Network Sheet 1' again.
- Table Panel:** Displays a table titled 'Yeast Network Sheet 1' with columns 'shared...', 'name', and 'isExclu...'. The table is currently empty.
- Network View:** A graph visualization titled 'Yeast Network Sheet 1' showing nodes represented by teal rounded rectangles and a dense web of gray edges connecting them. The nodes are arranged in five horizontal rows, each containing approximately 10 nodes with labels like L240C, YPR062W, YIL061C, etc.

At the bottom left, a status bar indicates '✓ Loading network from table'. On the bottom right, there's a memory usage indicator showing '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...

From Preset Network

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

Load Preset Network

Don't show again  Close

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.

Node Table Edge Table Network Table

Searching PSICQUIC Services

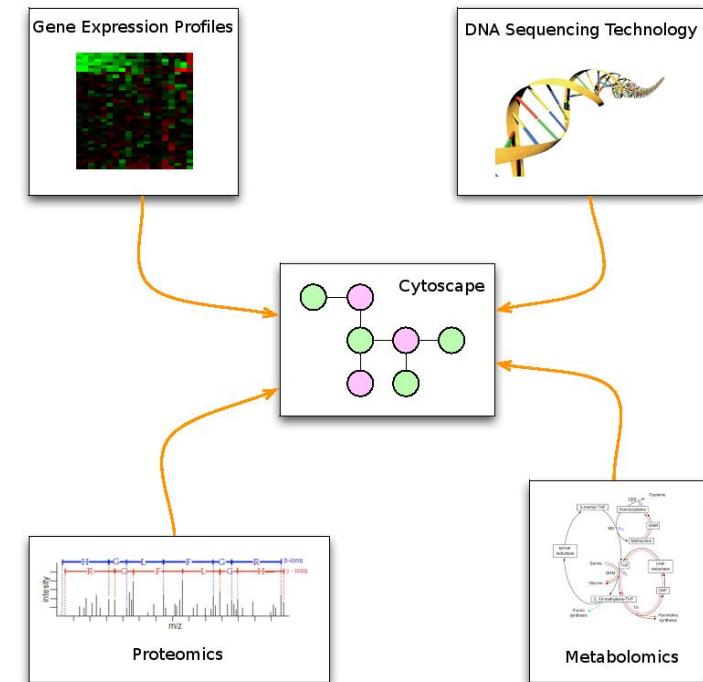
Memory: OK

?



# 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 1 33... 36...

Table Panel

Change Table Mode

Yeast Network Sheet 1

| shared... | name    | isExclu... | COMM... | gal1R... | gal4R... | gal80R... | gal1R...   | gal4R... |
|-----------|---------|------------|---------|----------|----------|-----------|------------|----------|
| 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     | 6.0234E... | 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

Sample for galFiltered

galFiltered.sif

MUD2  
GAL3  
AL80  
GAL10  
GAL1  
GAL4  
SWI5  
SWI4  
BAS1  
MIG1  
HSP150  
CLN3  
CDC6  
PIS1  
STE12  
MFA1  
STE2  
PMA1  
TUFF  
MFA2  
BAR1  
ALPHA2

Node Table Edge Table Network Table

Memory: OK

YDL194W YDR277C YBR043C YPR145W YER054C YBR045C YBL079W YLR345W YIL052C

13.1169... 12.1209... 1.5 9.79838... 8.81854... 9.64516... 2.0 9.11693... 9.14019...

0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0

0.07623... 0.08250... 0.66666... 0.10205... 0.11339... 0.10367... 0.5 0.10968... 0.10020...

false false false false false false false false false

0 0 0 0 0 0 0 0 0

0 0 0 0 0 0 0 0 0

|         | name    | Average... | Cluster... | Closeness... | IsSingular... | Partne... |   |
|---------|---------|------------|------------|--------------|---------------|-----------|---|
| YDL194W | YDL194W | 13.1169... | 0.0        | 0.07623...   | false         | 0         | 0 |
| YDR277C | YDR277C | 12.1209... | 0.0        | 0.08250...   | false         | 0         | 0 |
| YBR043C | YBR043C | 1.5        | 0.0        | 0.66666...   | false         | 0         | 0 |
| YPR145W | YPR145W | 9.79838... | 0.0        | 0.10205...   | false         | 0         | 0 |
| YER054C | YER054C | 8.81854... | 0.0        | 0.11339...   | false         | 0         | 0 |
| YBR045C | YBR045C | 9.64516... | 0.0        | 0.10367...   | false         | 0         | 0 |
| YBL079W | YBL079W | 2.0        | 0.0        | 0.5          | false         | 0         | 0 |
| YLR345W | YLR345W | 9.11693... | 0.0        | 0.10968...   | false         | 0         | 0 |
| YIL052C | YIL052C | 9.14019... | 0.0        | 0.10020...   | false         | 0         | 0 |



Searching PSICQUIC Services

Memory: OK



# 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

Selected 22 nodes and 0 edges in 9ms

Filter Chain

galFiltered.sif

Table Panel

f(x) galFiltered.sif

| Page... | Clusteri... | Closeness... | IsSingl... | Partner... | SelfLoop... | Eccentrici... | 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          | 2           | 20            | 5      | 0.5    |            |
| 58...   | 0.0         | 0.11665...   | false      | n          | n           | 2n            | 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>**

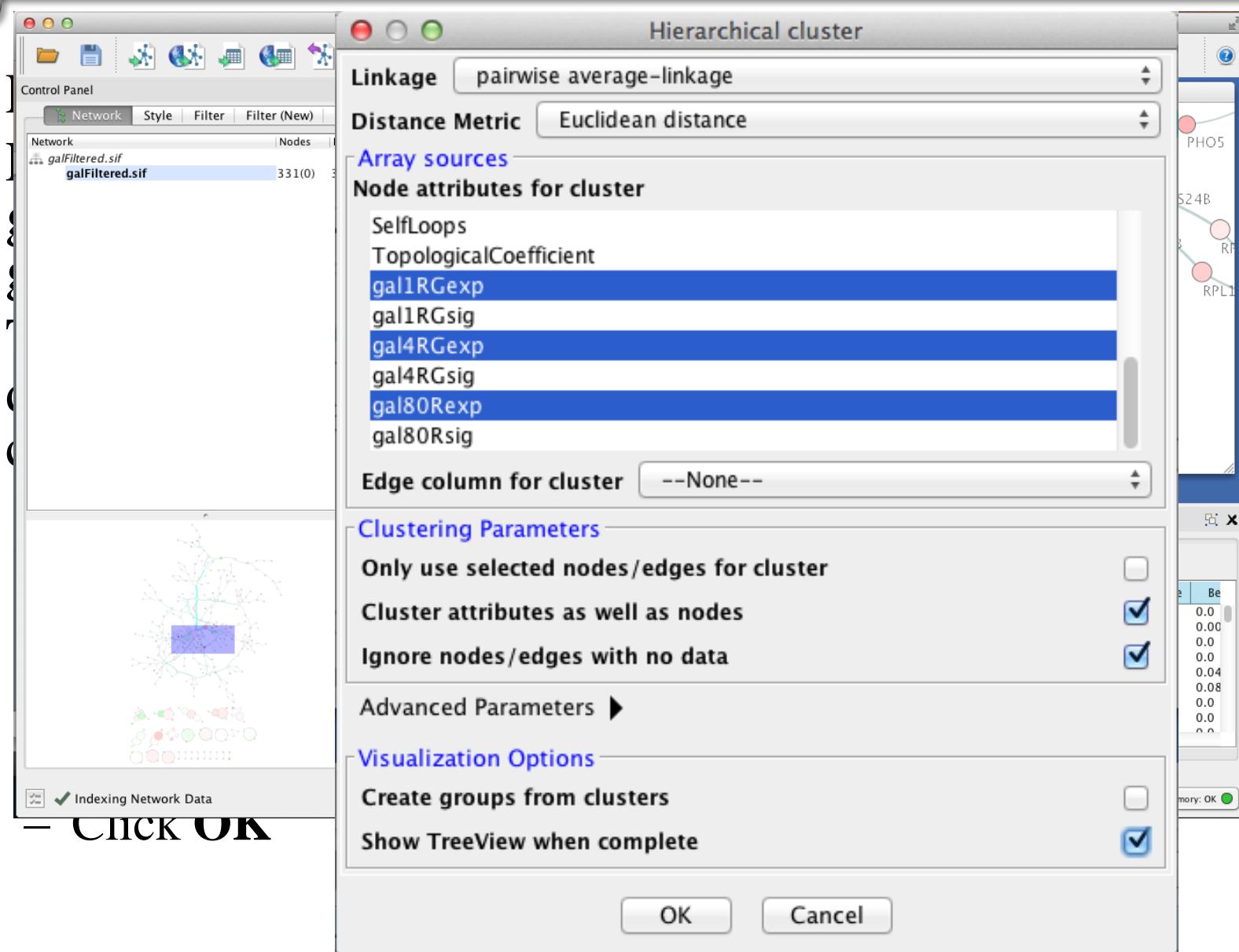


# Examples/Demos

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

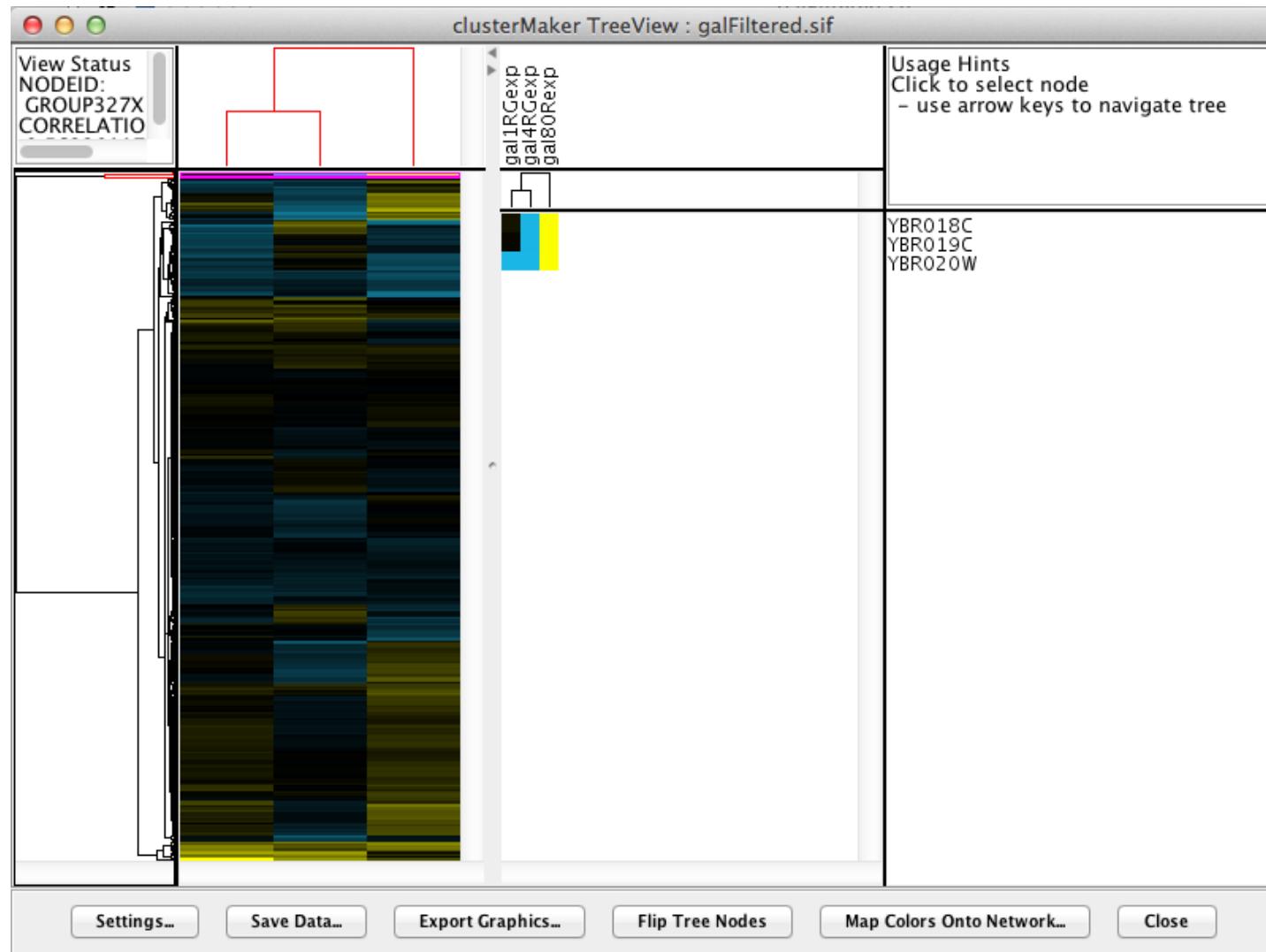


# Expression Data Analysis





# Expression Data Analysis





# Expression Data Analysis

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

Control Panel

Network Style Filter Filter (New) Sets

Nodes Edges

Network

- galFiltered.sif
- galFiltered.sif
- cluster1
- cluster1**

331(3) 362(0)

72(0) 127(0)

cluster1

monosaccharide catabolic process

hexose catabolic process

hexose metabolic process

galactose catabolic process via UDP-galactose

galactose metabolic process

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 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 |
| 44277 | small molecule catabolic process              | 1.2282... | 0.1771...  | 2/2          | 100.0%     | 148/6208... YBR020W YBR018C YBR019C |

Select All Unselect All Select nodes

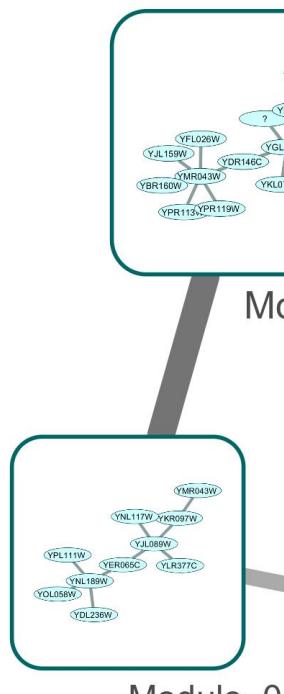
Start BiNGO



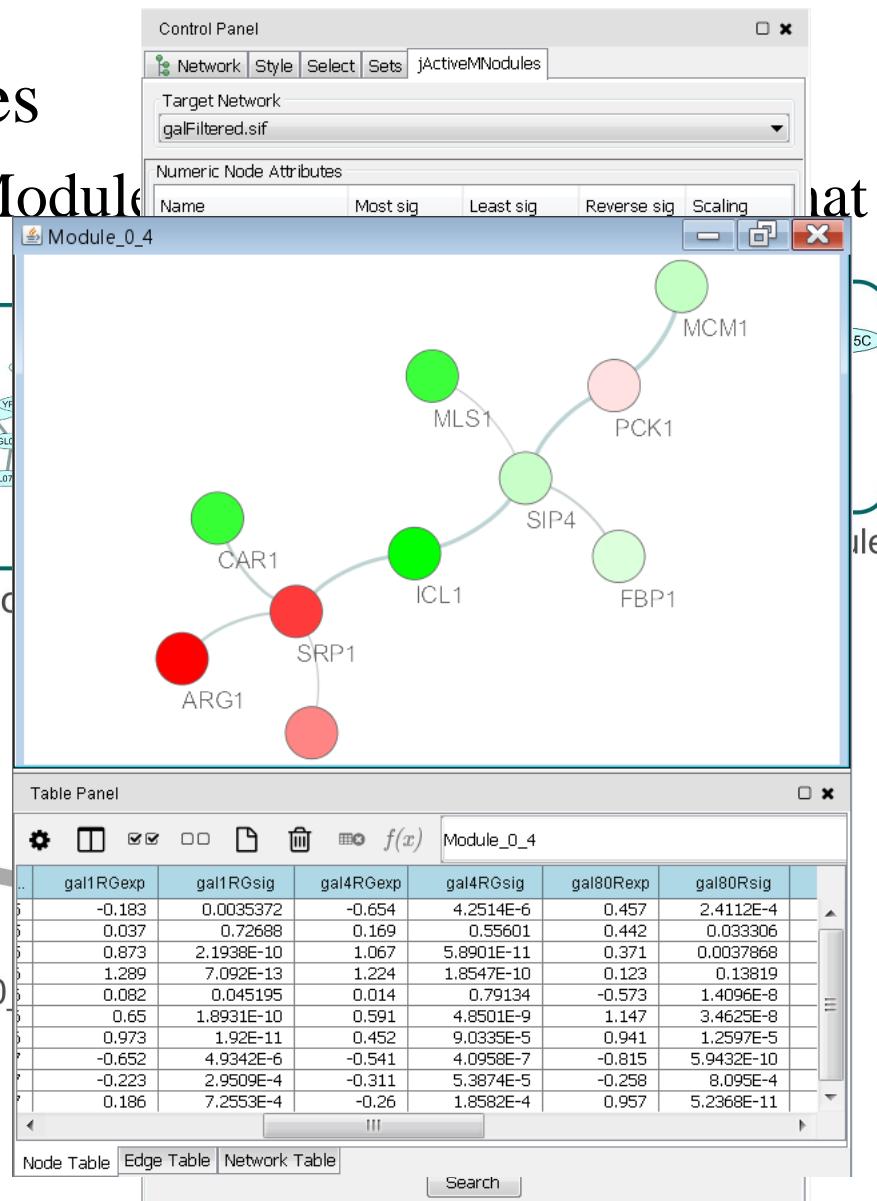
# Expression Data Analysis

- Active modules

## – The jActiveModule



Module 0



r time-

# Rules Pages



# 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... |
| DNA and Tran 07-21-06b.csv | 743(0)  | 18372... |
| RNAUberNov2+Meg6c.csv      | 552(0)  | 10784... |

combined\_scores\_good.txt

Table Panel

f(x) combined\_scores\_good.txt

| shared... | 0_MCL... | DNA a... | Endo f... | Featur... | GPL51... | GPL51... | GPL51... | GPL51... | GPL51... | Gl...  |      |
|-----------|----------|----------|-----------|-----------|----------|----------|----------|----------|----------|--------|------|
| 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 |
| PUP1      |          |          |           |           |          |          |          |          |          |        |      |

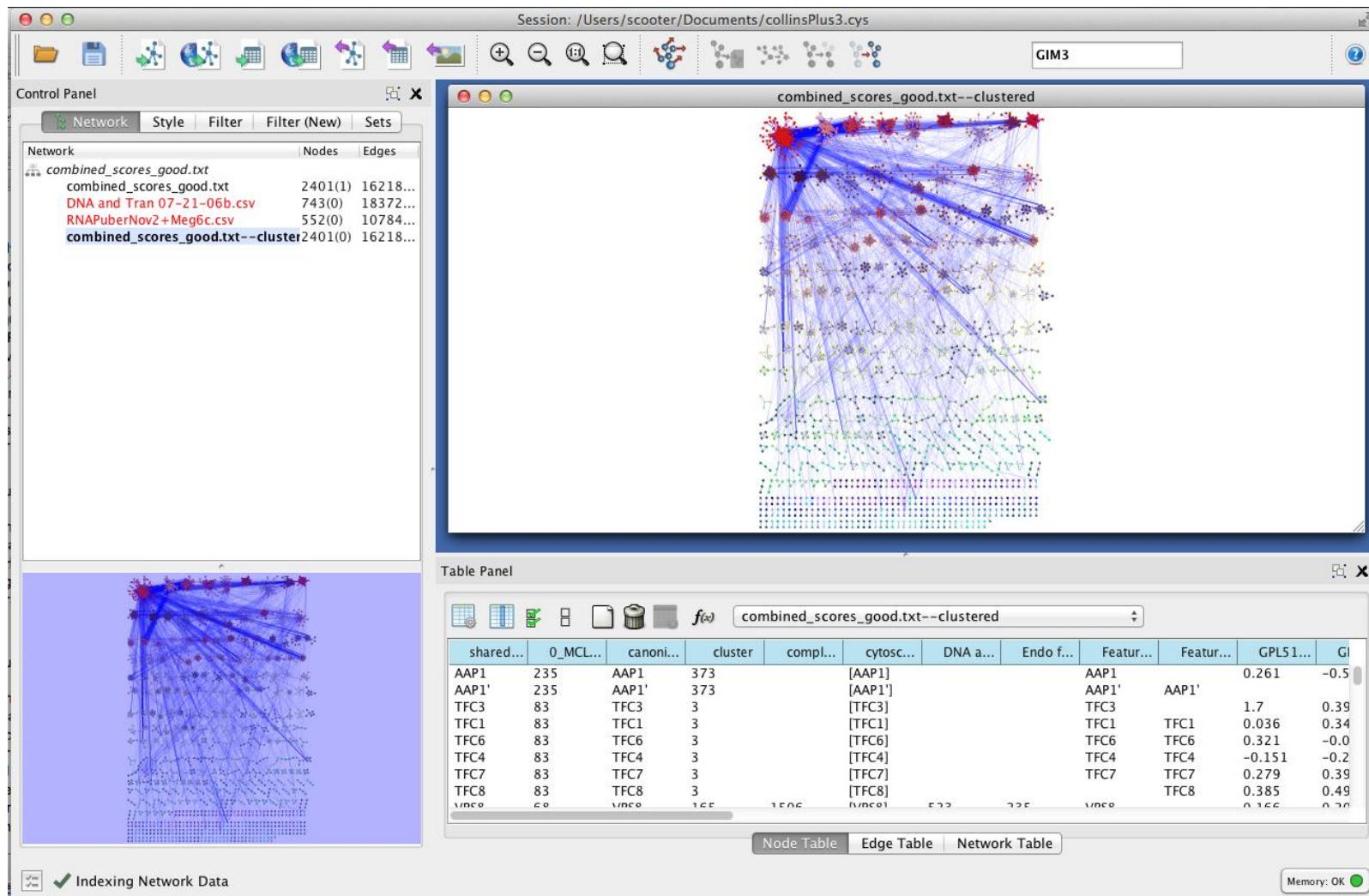
Node Table Edge Table Network Table

Indexing Network Data

Memory: OK



# Protein Complexes



ges



# Protein Complexes

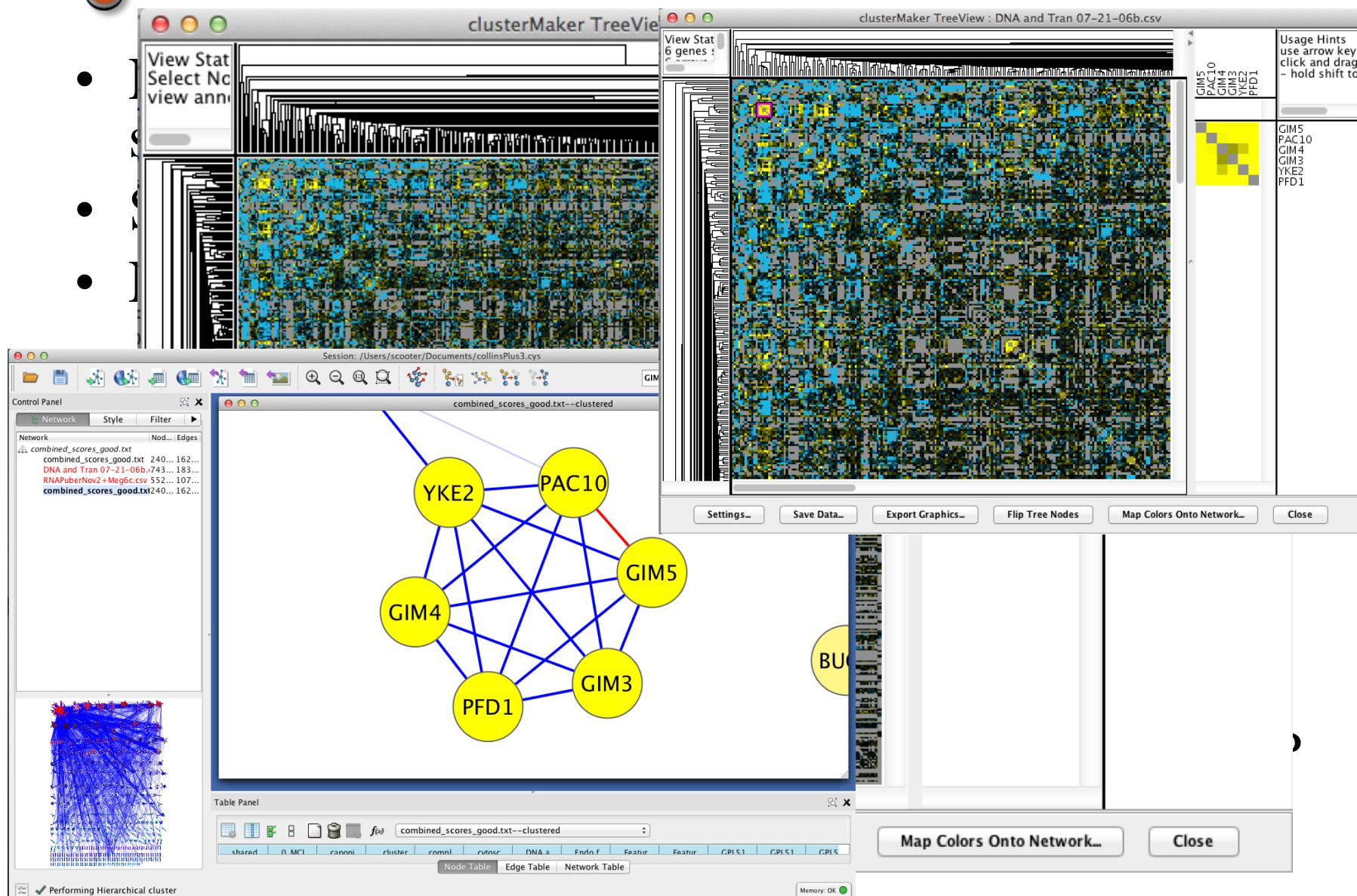
The screenshot shows the Cytoscape software interface with a 'Hierarchical cluster' dialog box open. The dialog box contains several configuration options:

- Linkage:** pairwise average-linkage
- Distance Metric:** Uncentered correlation
- Array sources:** Node attributes for cluster (selected)
- Node attributes for cluster:** --None--
- 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:

At the bottom of the dialog box are 'OK' and 'Cancel' buttons. Below the dialog box are three tabs: 'Node Table', 'Edge Table', and 'Network Table'. In the bottom right corner, there is a status bar with 'Memory: OK' and a green indicator.

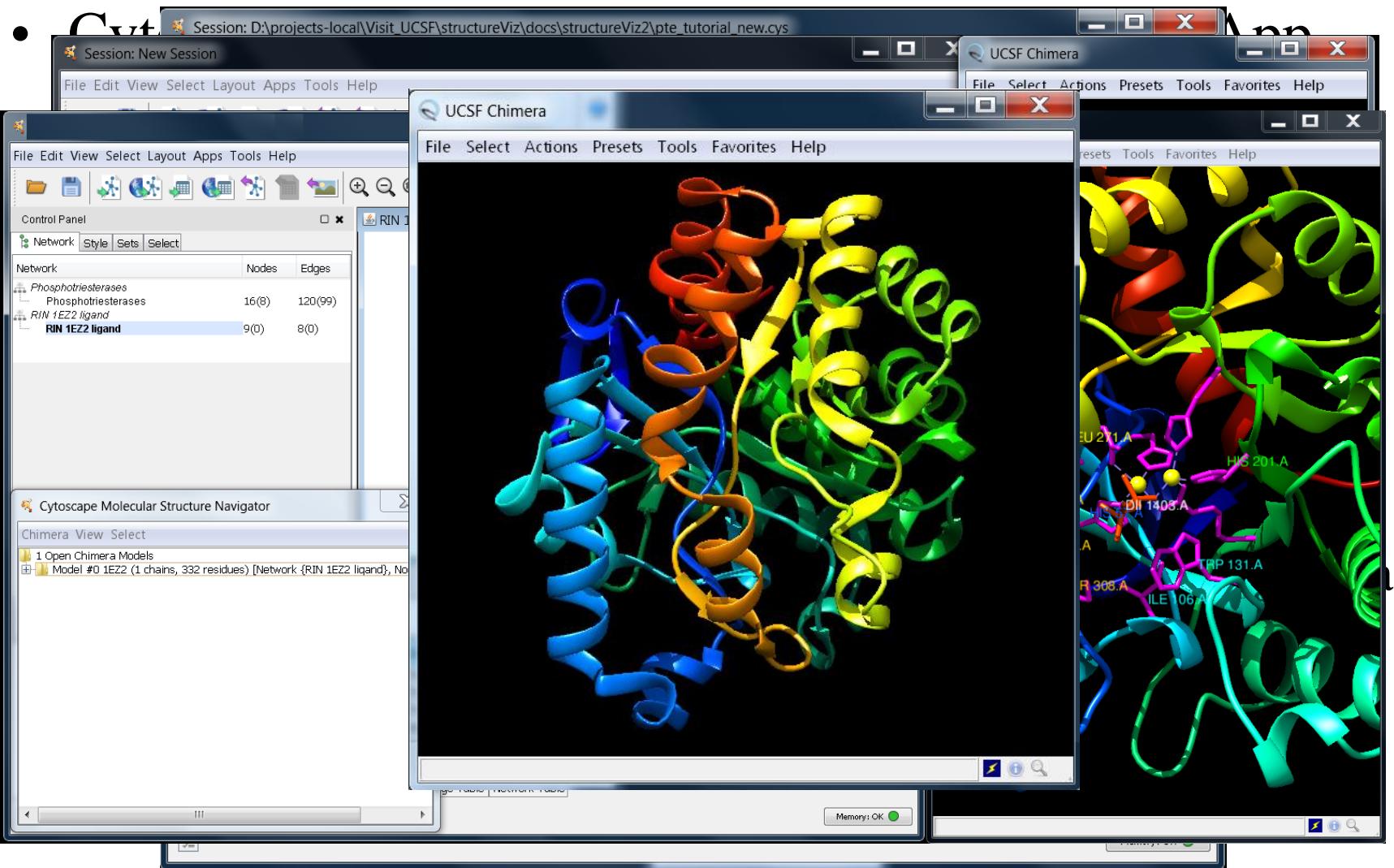


# Protein Complexes





# 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

Control Panel

Network Style Select

Properties

Def. Map. Byp.

galFiltered Style

Column SGD symbol  
Mapping Type Passthrough Mapping

Label Color  
14 Label Font Size  
Shape  
35.0 Size  
Tooltip  
245 Transparency  
Width

Lock node width and height

Table Panel

Prot Accession Degree gal1RGsig SGD symbol gal80Rexp annotation.GO\_CELLULAR\_COMPONENT Additional Links Ensembl Family ID annotation.Assigned\_by Family Descrip

66] 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 [YIG1] 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, nucle... [http://www.genome.jp... [ENSF00000000332] UniProt [AMBIGUOUS]  
1 4.7824E-4 [YEF1] -0.265 [cellular\_component] [http://www.genome.jp... [ENSF000000001025] 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] UniProt [AMBIGUOUS]  
1 4.6513E-11 [FMP16] -0.089 [mitochondrion] [http://www.genome.jp... [ENSF00000000183] SGD [UNKNOWN]

Node Edge Network

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               | UniprotID2 | 0         | LARS        | 5859      | QARS        |
| Q9P2J5     | hu-hu               |            | 0         | LARS        | 9255      | SCYE1       |
| Q9P2J5     | hu-hu               | Q15046     | 0         | LARS        | 5917      | RARS        |
| Q9P2J5     | hu-hu               | O43324     | 0         | LARS        | 4141      | MARS        |
| P14868     | hu-hu               | P07814     | 5         | DARS        | 5859      | QARS        |
| P14868     | hu-hu               | Q12904     | 5         | DARS        | 3376      | IARS        |
| P14868     | hu-hu               | P54136     | 5         | DARS        | 4141      | MARS        |
| P14868     | hu-hu               | Q15046     | 1615      | DARS        | 5917      | RARS        |
| P14868     | hu-hu               | O43324     | 1615      | DARS        | 5917      | RARS        |
| P14868     | hu-hu               | P47897     | 1615      | DARS        | 7965      | JTV1        |
| P14868     | hu-hu               | Q12904     | 5917      | RARS        | 9255      | SCYE1       |
| P14868     | hu-hu               | P54136     | 3735      | KARS        | 51520     | LARS        |
| P14868     | hu-hu               | Q15046     | 3735      | KARS        | 3735      | KARS        |
| P14868     | hu-hu               | P47897     | 3735      | KARS        | 9521      | EEF1E1      |
| P14868     | hu-hu               | Q12904     | 3735      | KARS        | 2058      | EPRS        |
| P54136     | hu-hu               | P56192     | 3735      | KARS        | 9255      | SCYE1       |
| Q15046     | hu-hu               | Q12904     | 5917      | RARS        | 5917      | RARS        |
| Q15046     | hu-hu               | P54136     | 3735      | KARS        | 5859      | QARS        |
| Q15046     | hu-hu               | Q15046     | 3735      | KARS        | 9255      | SCYE1       |
| Q15046     | hu-hu               | P47897     | 3735      | KARS        | 4141      | MARS        |
| Q15046     | hu-hu               | Q12904     | 3735      | KARS        | 51520     | LARS        |
| Q15046     | hu-hu               | P56192     | 3735      | KARS        | 9255      | SCYE1       |
| P47897     | hu-hu               | Q12904     | 5859      | QARS        | 7965      | JTV1        |
| P47897     | hu-hu               | P54136     | 5859      | QARS        | 3735      | KARS        |
| P41252     | hu-hu               | Q13155     | 3376      | IARS        | 4141      | MARS        |
| P41252     | hu-hu               | Q15046     | 3376      | IARS        | 9255      | SCYE1       |
| P41252     | hu-hu               | P56192     | 3376      | IARS        | 51520     | LARS        |
| P41252     | hu-hu               | Q12904     | 3376      | IARS        | 5917      | RARS        |
| P41252     | hu-hu               | Q9P2J5     | 3376      | IARS        | 3376      | IARS        |
| P41252     | hu-hu               | P54136     | 3376      | IARS        | 3376      | IARS        |

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)