



# Cytoscape

*Pathway Analysis Made Easy*

*Melissa Cline  
UC Santa Cruz*

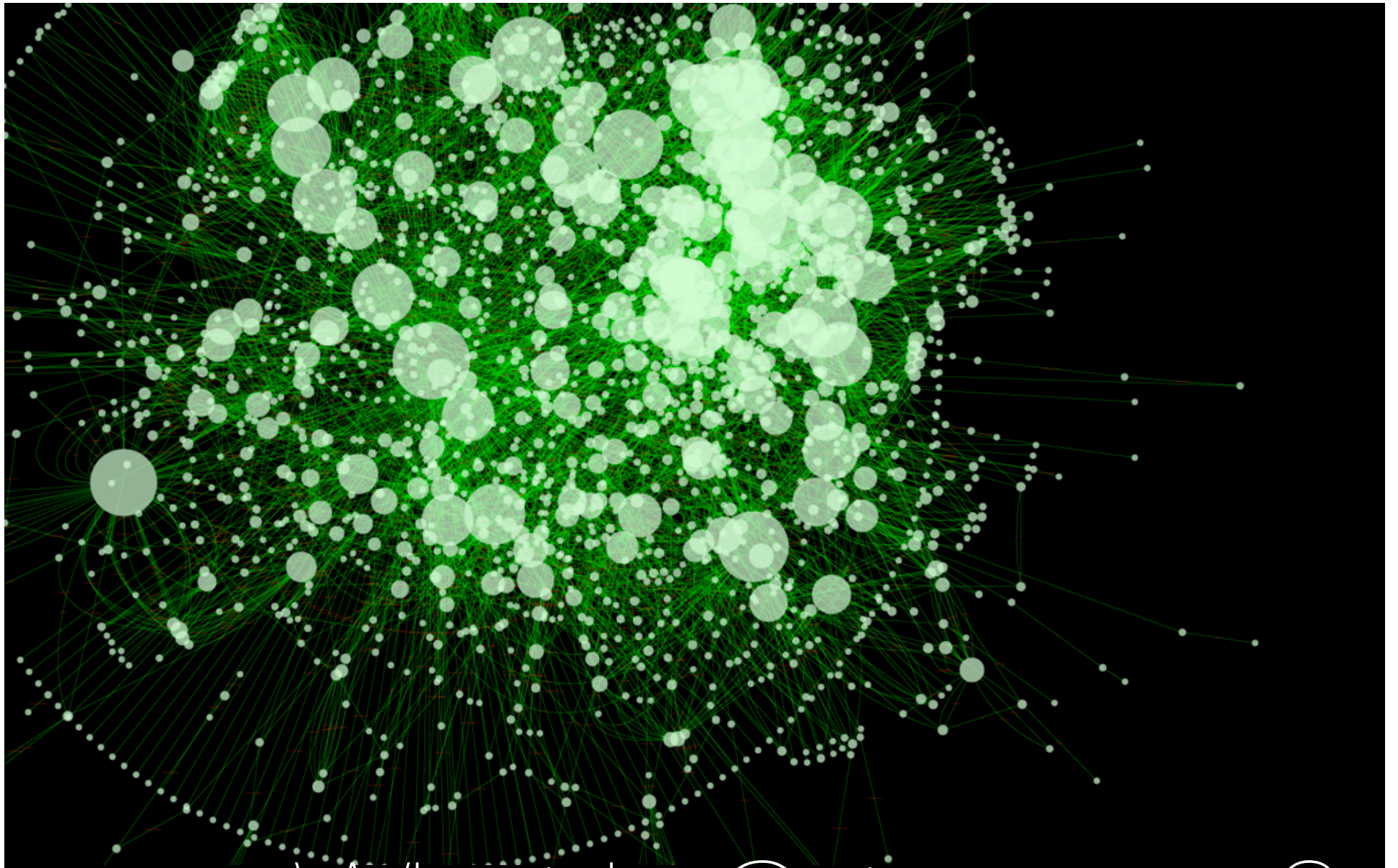
*Special thanks to Keichiro Ono*

# Thanks to the whole Cytoscape consortium





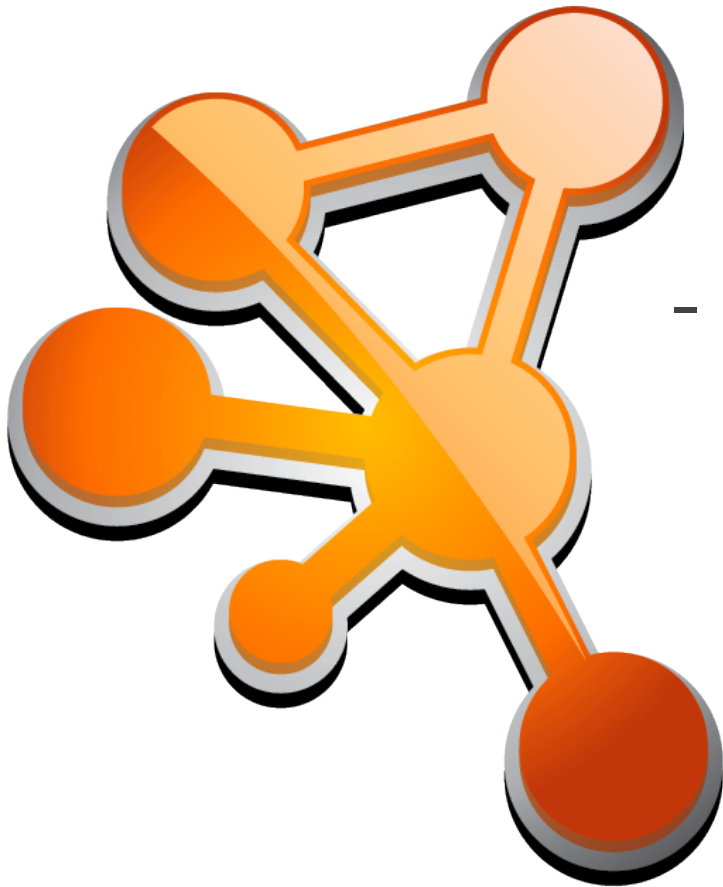




What is Cytoscape?



# Cytoscape



- Open source **platform** for biological network data integration, analysis, and visualization

Cytoscape Version 2.7.0 File Edit View Select Layout Plugins Help

Cytoscape Desktop (Session: vizbl.cys) 11月2日(月) 14:34:29

Control Panel

Network	Nodes	Edges
DNA damage response	177(15)	99(17)
MAPK signaling pathway	295(0)	150(0)
TP53 human	1081(0)	2794(0)
HPRD.sif	6112(54)	20126(0)
MCODE Result 1	54(0)	705(0)

HPRD.sif

MCODE Result 1: Cluster 1 (Score: 13.037)

MAPK signaling pathway

TP53 human

DNA damage response

Results Panel

MCODE Result 1

Cluster Browser

Network	Details
	Rank: 1 Score: 13.037 Nodes: 54 Edges: 705
	Rank: 2 Score: 8.062 Nodes: 32 Edges: 259
	Rank: 3 Score: 7.824 Nodes: 17 Edges: 133
	Rank: 4 Score: 5.893 Nodes: 103 Edges: 615
	Rank: 5 Score: 4.5 Nodes: 10 Edges: 46
	Rank: 6

Explore: Cluster 1

Size Threshold

Node Attribute Enumerator

Value Occurrence

Create Sub-Network

Export Discard Result

Data Panel

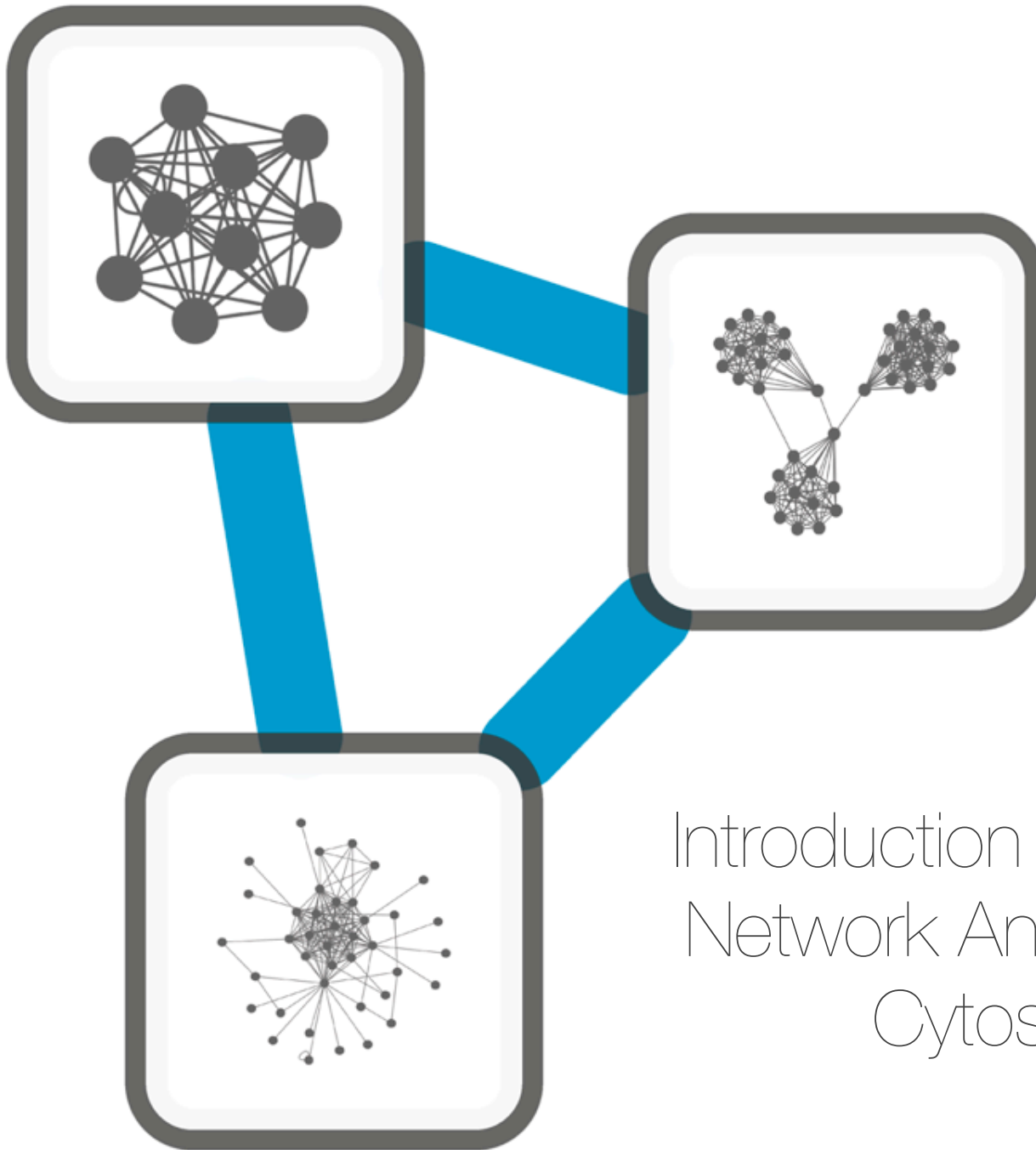
ID	GeneID
CCND1	ENSG00000110092
CCND2	ENSG00000118971

Node Attribute Browser Edge Attribute Browser Network Attribute Browser VistaClara

Welcome to Cytoscape 2.7.0 Right-click + drag to ZOOM Middle-click + drag to PAN



- Free! (LGPL)
- Developed and maintained by universities, companies, and research institutions
- De-facto standard software in biological network research community
- Expandable by Apps

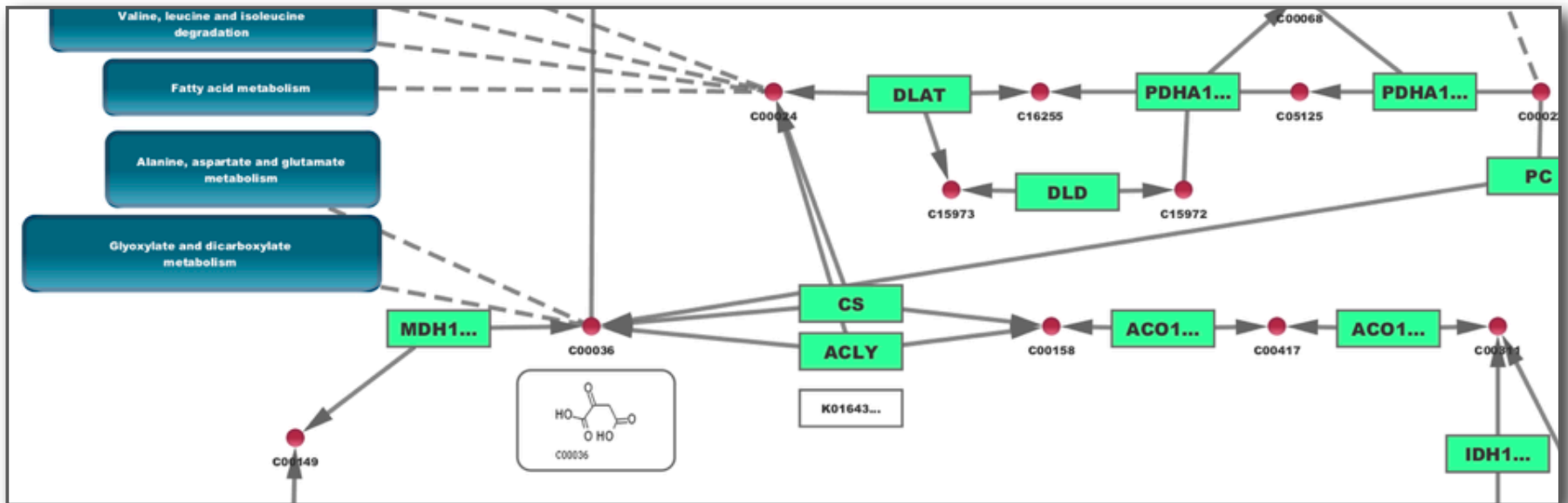


# Introduction to Biological Network Analysis Using Cytoscape



Network?

# Human-Curated Pathways





A complex undirected network graph with numerous nodes and edges. The nodes are represented by circles of varying sizes, with the largest nodes acting as hubs. The edges are thin lines connecting the nodes, forming a dense, interconnected web. The graph is centered on a large hub node, with several other significant hubs branching out from it. The overall structure is highly connected and non-linear.

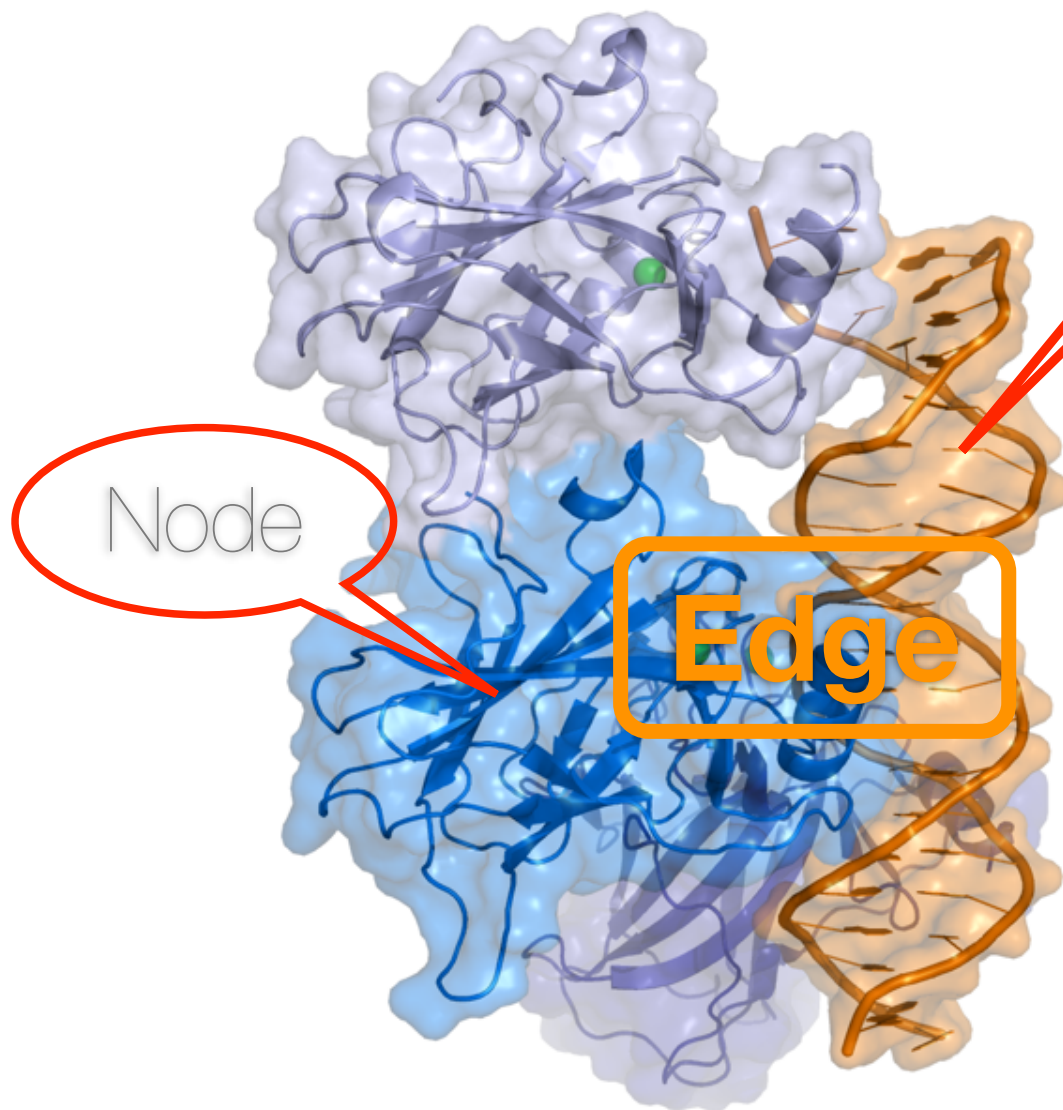
# Undirected Network

- Protein - Protein Interaction

*Network = Nodes + Edges*



# Nodes and Edges in Biology



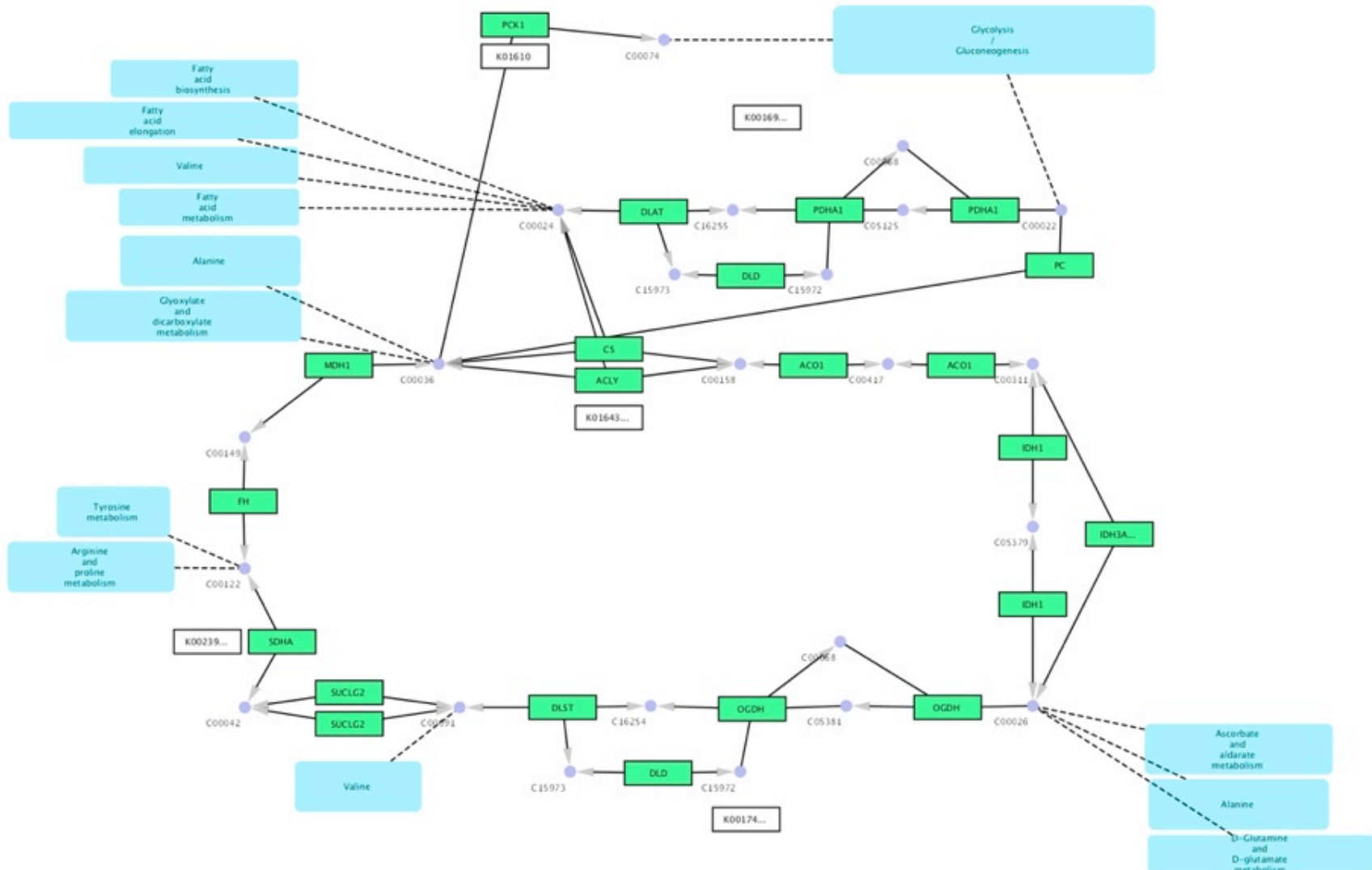
Node

Node

Edge

- Protein - Protein
- Protein - DNA
- Genetic (Epistasis)
- **Synthetic lethality**
- Biochemical Reactions
- **Compound - Enzyme - Compound**

# Directed Network



# Cytoscape Workflow

1. Load Networks (Get network data)
2. Load Attributes (Get data about networks)
3. Analyze and Visualize Networks
4. Prepare for Publication
  - A specific example of this workflow:
    - Cline, et al. "Integration of biological networks and gene expression data using Cytoscape", Nature Protocols, 2, 2366-2382 (2007).

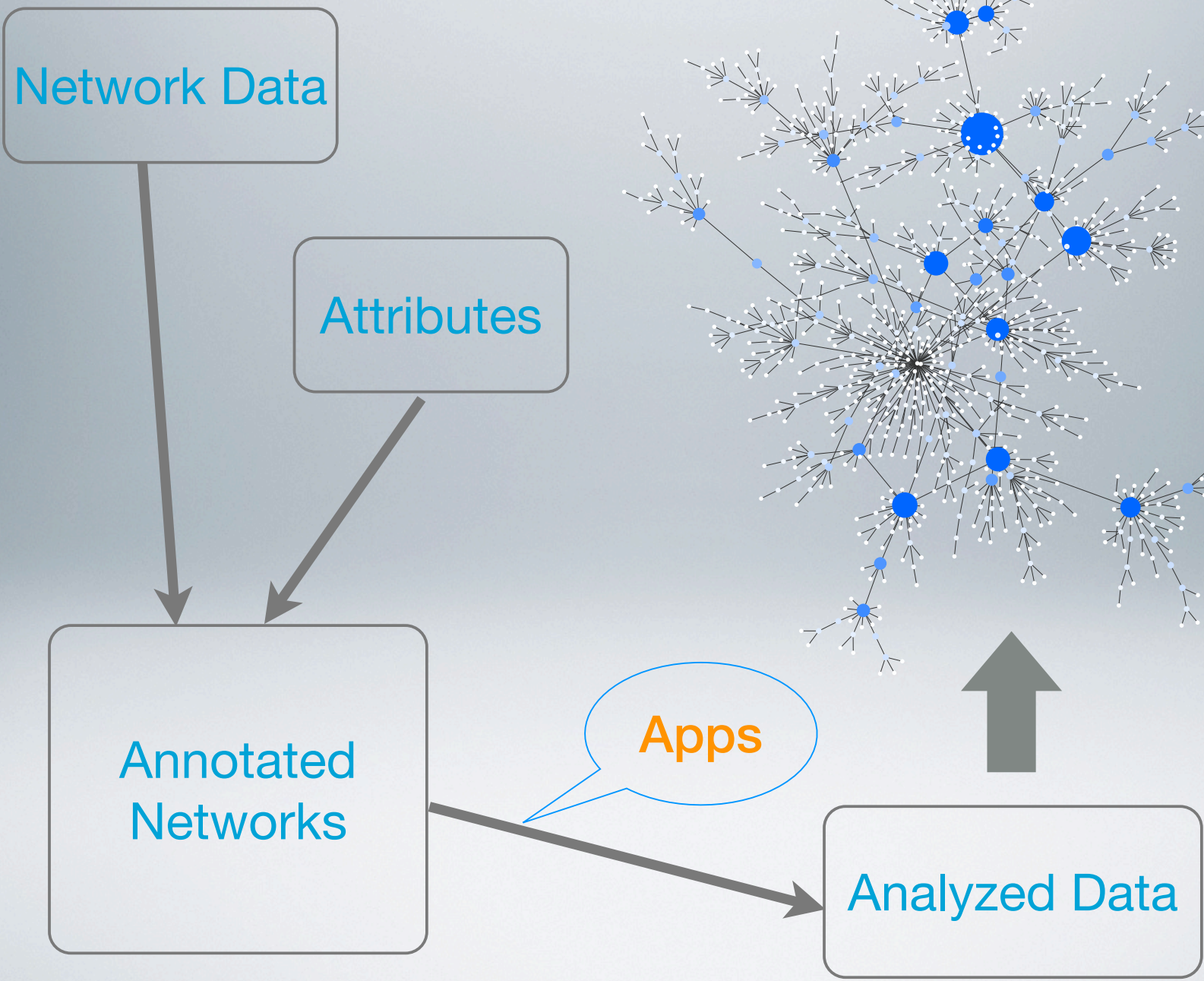
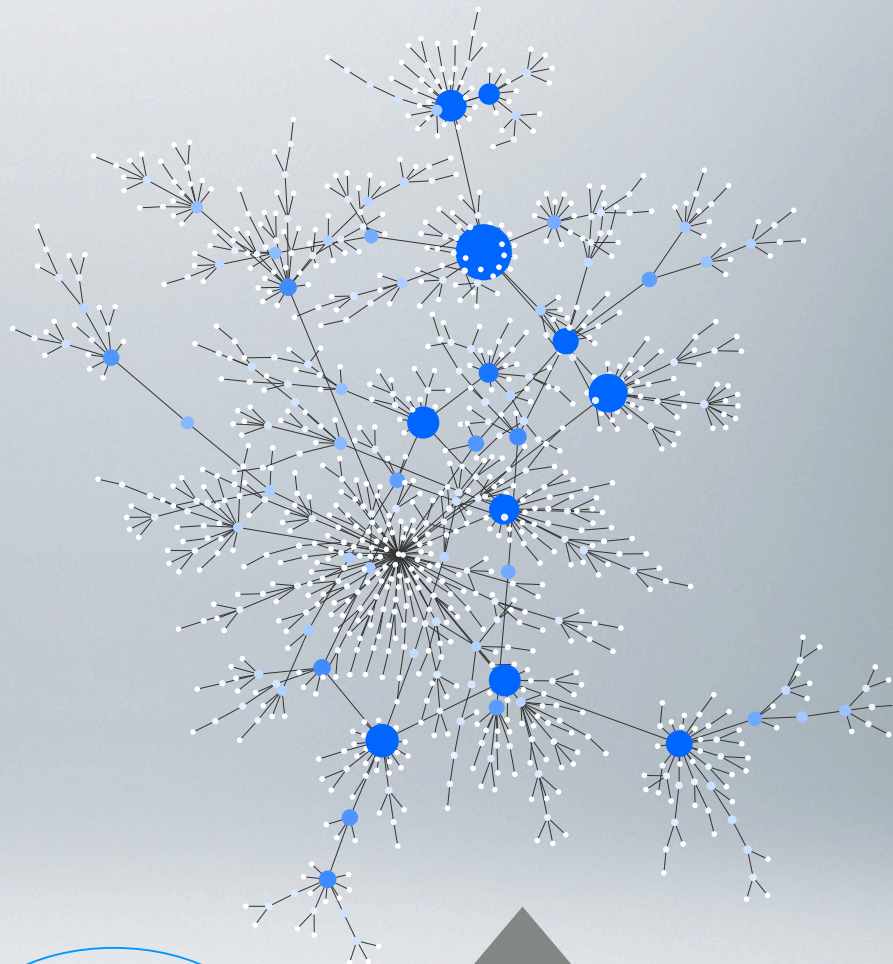
Network Data

Attributes

Annotated  
Networks

Apps

Analyzed Data

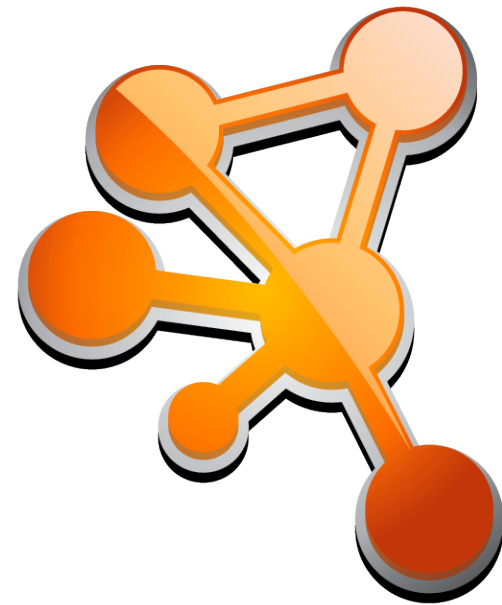
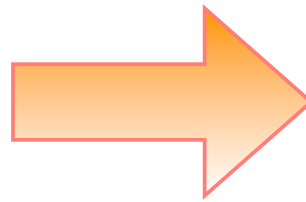




```

<?xml version="1.0" encoding="UTF-8"?>
<graphml xmlns="http://graphml.graphdrawing.org/xmlns"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="http://graphml.graphdrawing.org/xmlns
    http://graphml.graphdrawing.org/xmlns/1.0/graphml.xsd">
<!-- Created by igraph -->
<key id="degree" for="node" attr.name="degree" attr.type="double"/>
<key id="betweenness" for="node" attr.name="betweenness"
  attr.type="double"/>
<graph id="G" edgedefault="directed">
  <node id="n0">
    <data key="degree">79</data>
    <data key="betweenness">0</data>
  </node>
  <node id="n1">
    <data key="degree">9</data>
    <data key="betweenness">167</data>
  </node>
  <node id="n2">
    <data key="degree">18</data>
    <data key="betweenness">75</data>
  </node>
  <node id="n3">
    <data key="degree">8</data>
    <data key="betweenness">12</data>
  </node>
  <node id="n4">
    <data key="degree">26</data>
    <data key="betweenness">210</data>
  </node>
  <node id="n5">
    <data key="degree">29</data>
    <data key="betweenness">320</data>
  </node>

```



# Import Networks

# Network Data Formats

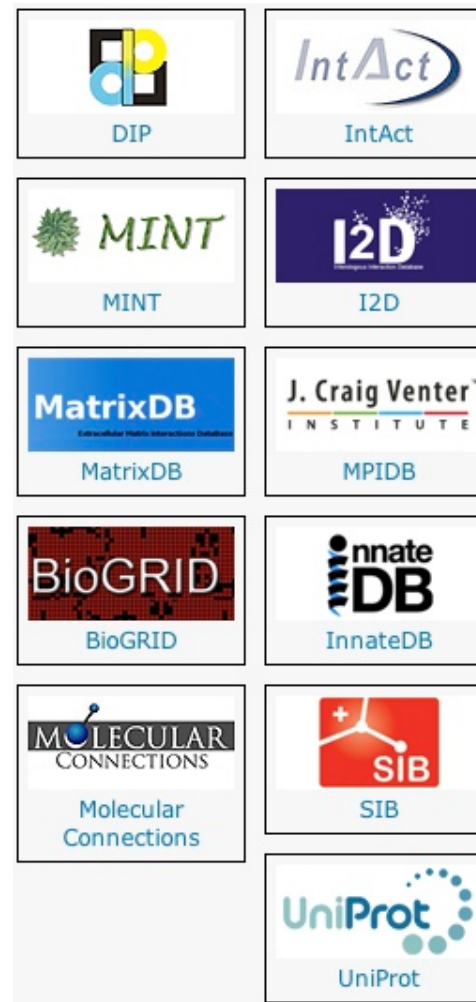
- **SIF**
- **GML**
- **XGMML**
- **GraphML**
- **BioPAX**
- **PSI-MI**
- **SBML**
- **KGML (KEGG)**
- **Excel**
- **Delimited Text Table**
- **CSV**
- **Tab**

# *I Don't Have a Network!*

- **Don't worry - There are tons of public interaction data sets**
- **Cytoscape can import those public data sets directly via Internet.**

# Which Database?

- Protein - Protein
  - STRING, IntAct
- Genetic
  - BioGRID
- Protein - Compound
  - ChEMBL
- Human-Curated Pathways
  - KEGG, Reactome, PathwayCommons



galFiltered.nodeAttrTable.xlsx

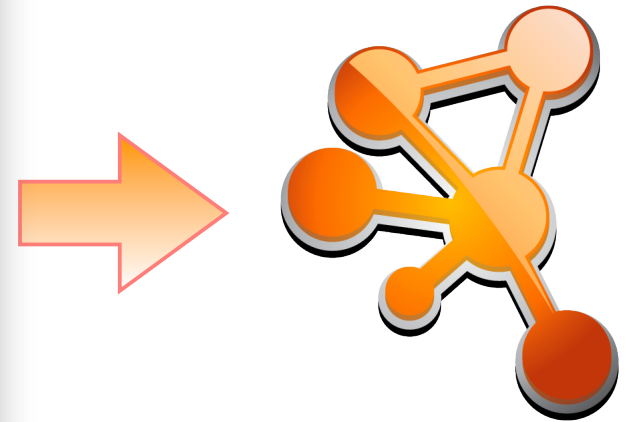
Search in Sheet

Home Layout Tables Charts SmartArt Formulas Data Review

ID	alias	annotation.GO BIOLOGICAL PROCESS	notation.GO CELLULAR COMPONED MO
YAL003W	translation elongation	translational elongation	eukaryotic translation elongation factor trans
YAL030W	S000000028, SNC1	Golgi to plasma membrane transport, endoc	Golgi trans face, SNARE complex, end v-SN
YAL038W	CDC19, PYK1, S0000	glycolysis, pyruvate metabolic process	cytosol
YAL040C	CLN3, DAF1, FUN10,	G1/S transition of mitotic cell cycle, regulati	nucleus
YAR007C	heterotrimeric RPA (R	DNA recombination, DNA replication, synthe	DNA replication factor A complex, chro dama
YBL005W	AMY2, PDR3, S0000	positive regulation of transcription, DNA-def	cytoplasm, nucleus
YBL021C	HAP3, S000000117,	tregulation of carbohydrate metabolic proces	CCAAT-binding factor complex, nucleu trans
YBL026W	Sm class, LSM2, S00	mRNA catabolic process, nuclear mRNA sp	U4/U6 x U5 tri-snRNP complex, nuclec RNA
YBL050W	S000000146, SEC17	ER to Golgi vesicle-mediated transport, vac	peripheral to membrane of membrane
YBL069W	AST1, S000000165	protein targeting to membrane	extrinsic to membrane
YBL079W	NLE3, NUP170, S000	NLS-bearing substrate import into nucleus,	nuclear pore
YBR018C	GAL7, S000000222,	galactose catabolic process	cytoplasm
YBR019C	GAL10, S000000223,	galactose catabolic process	soluble fraction
YBR020W	GAL1, S000000224,	galactose catabolic process, galactose tran	cytoplasm
YBR043C	AQR2, QDR3, S0000	multidrug transport	integral to plasma membrane, plasma
YBR045C	GIP1, Glc7p regulat	regulation of phosphoprotein phosphatase	prospore membrane, protein phosphat
YBR050C	Glc7p regulatory sub	negative regulation of transcription from RN	protein phosphatase type 1 complex
YBR072W	HSP26, S000000276	protein folding, response to stress	cytoplasm, nucleus
YBR093C	PHO5, S000000297,	cellular response to phosphate starvation, p	cell wall (sensu the Fungi research cor acid
YBR109C	CMD1, S000000313,	cell budding, cytoskeleton organization and	bud neck, bud tip, central plaque of spi calci
YBR112C	CRT8, CYC8, S0000	chromatin remodeling, negative regulation c	nucleus
YBR118W	translational elongati	translational elongation	eukaryotic translation elongation factor trans
YBR135W	CKS1, Cdc28 protein	regulation of progression through cell cycle,	cytoplasm, nucleus
YBR155W	CNS1, S000000359	protein folding	cytoplasm
YBR160W	CDC28, CDK1, HSL5	G1/S transition of mitotic cell cycle, G2/M tr	cytoplasm, nucleus
YBR170C	HRD4, NPL4, S0000	ER-associated protein catabolic process	endoplasmic reticulum, nuclear envelo
YBR190W			
YBR217W	APG12, ATG12, S000	autophagy, protein targeting to vacuole	membrane fraction

galFiltered.nodeAttr

Normal View Ready Sum = 0

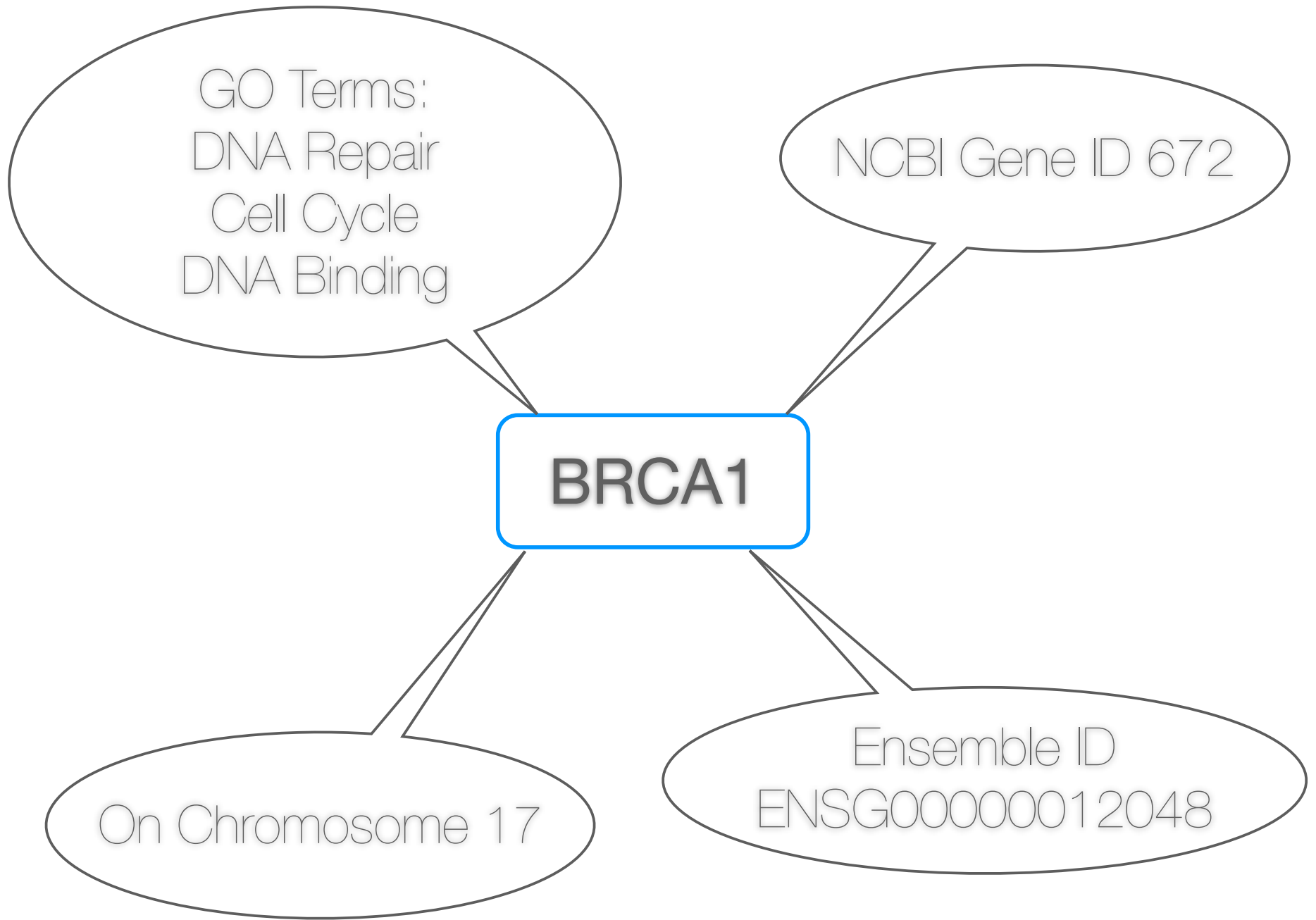


# Import Attributes



# What are Attributes?

- **Any data that describes or provides details about nodes, edges, and networks.**



# Why we need attributes?

- Analysis
  - It is hard to get biologically meaningful analysis result only from network topology (= structure)
- Meaningful Visualization
  - Expression values - Node Color
  - Gene Function - Node Shape
  - etc.

# Node Attributes

- **Gene Expression Data**
- **Human-readable gene names**
- **Gene Ontology Terms**



# Edge Attributes

- **Interaction Detection Methods**
  - **Y2H, NMR, affinity chromatography, etc.**
- **Interaction Type**
  - **Physical, genetic, predicted**
- **Publication ID**

# Network Attributes

- **Experiment details**
- **Pathway Metadata**
- **Description**
- **Publication ID**

# Public Data Source



- Access directly from Cytoscape, or



- Import as Table (Text/Excel)



### Import Attribute from Table

Data Sources: **Attributes**  Node  Edge  Network

Input File: file:/Applications/cytoscape-2.8.3/sampleData/ga...

Advanced:  Show Mapping Options  Show Text File Import Options  Import everything (Key is always ID)  Case Sensitive

Preview: **Text File** Left Click: Enable/Disable Column, Right Click: Edit Column

galExpData.pvals

✓ GENE	✓ COMMON	✓ gal1RG	✓ gal4RG	✓ gal8OR	✓ gal1RG	✓ gal4RG
YHR051W	COX6	-0.034	0.111	-0.304	3.75720e-01	1.56240e
YHR124W	NDT80	-0.090	0.007	-0.348	2.71460e-01	9.64330e
YKL181W	PRS1	-0.167	-0.233	0.112	6.27120e-03	7.89400e
YGR072W	UPF3	0.245	-0.471	0.787	4.10450e-04	7.51780e
YHL020C	OPI1	0.174	-0.015	0.151	1.40160e-04	7.19120e
YGR145W	YGR145W	0.387	-0.577	-0.088	5.37920e-03	8.27330e
YGL041C	YGL041C	0.285	-0.086	0.103	4.46050e-04	4.50790e
YGR218W	CRM1	-0.018	-0.001	-0.018	6.13810e-01	9.79400e
YOR202W	HIS3	-0.432	-0.710	0.239	1.09790e-02	1.79790e
YCR005C	CIT2	0.085	0.392	0.464	4.18980e-02	1.53050e
YER187W	KHS1	0.159	0.139	-0.045	8.51260e-04	4.17830e
YBR026C	YBR026C	0.276	0.189	0.291	3.63320e-05	6.15230e
YMR244W	YMR244W	0.078	-0.239	-0.072	5.76050e-01	3.55240e
YMR317W	YMR317W	-0.181	0.086	-0.453	5.94980e-02	3.03060e
YAR047C	YAR047C	0.234			2.46640e-02	4.96380e
YIR031C	DAL7					8.90570e
YDL177C	YDL177C					1.42900e
YLR338W	YLR338W					5.11510e
YGR073C	YGR073C					1.71320e
YGR146C	YGR146C					2.88770e
YOR130C	ORT1					6.13670e

Legend: Key Alias

Key Attributes

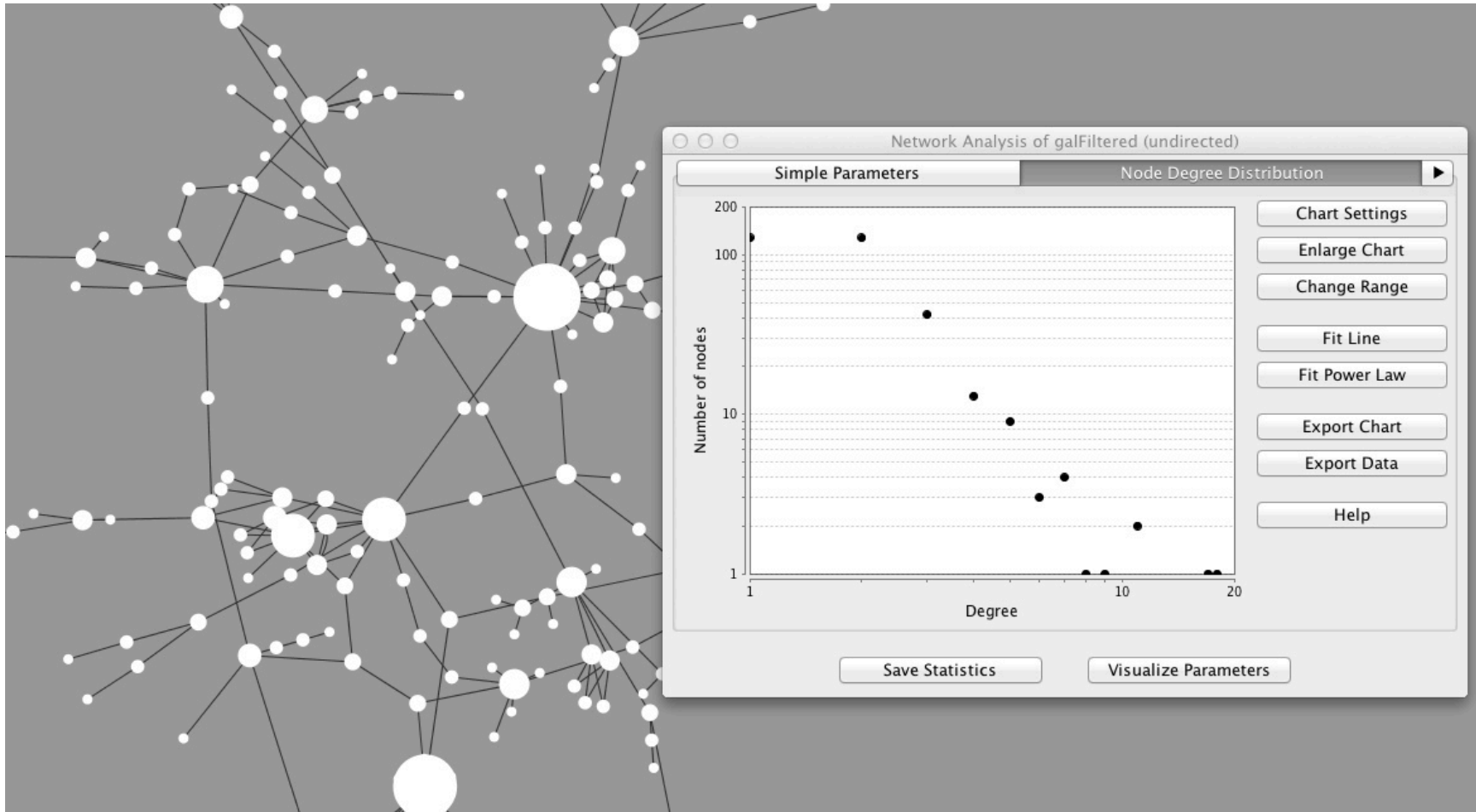
- YAL003W
- YAL030W
- YAL038W
- YAL040C
- YAR007C
- YBL005W
- YBL021C
- YBL026W
- YBL050W
- YBL069W
- YBL079W
- YBR018C
- YBR019C
- YBR020W
- YBR043C
- YBR045C
- YBR050C
- YBR072W
- YBR093C

Key-Value Matched: 9 | First 100 entries are loaded for preview. | File Size: 415 KBytes | Import Cancel

Mapping Key in the Network

Mapping Key in the Table





# Analysis

# Network Analysis


- Analysis is a huge topic in Cytoscape, so I'll show you only the very basic features only.

- **Filtering**
- Calculate network statistics by **Network Analyzer**
- Degree distribution, centrality, etc.
- Advanced analysis by **Apps**


# Filtering

- Find nodes and edges with specific conditions
  - Pick nodes with degree  $> 5$
  - Select edges extracted from publication X
  - Find nodes annotated by GO term ID Y

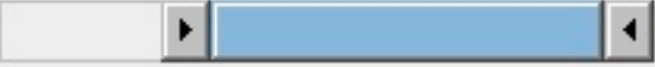

**Current Filter**

session: Default  Option ▾


**Filter Definition**

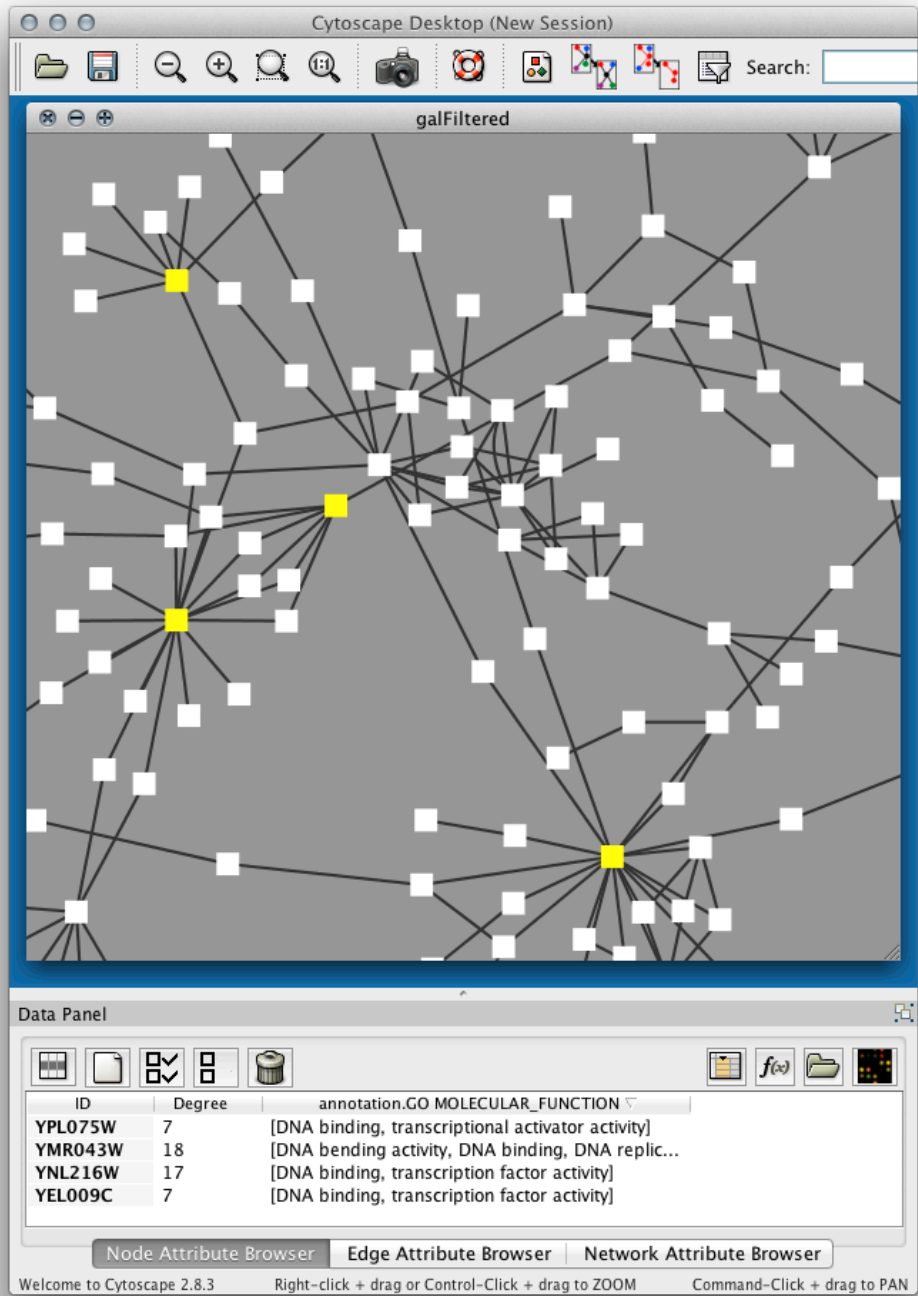
-- Attributes --  Add

**Advanced**

Degree  Not  

AND

annotation.GO MOLECULAR\_FUNCTION  Not **dna binding** 



Current Filter

session: Default Option

Filter Definition

-- Attributes -- Add

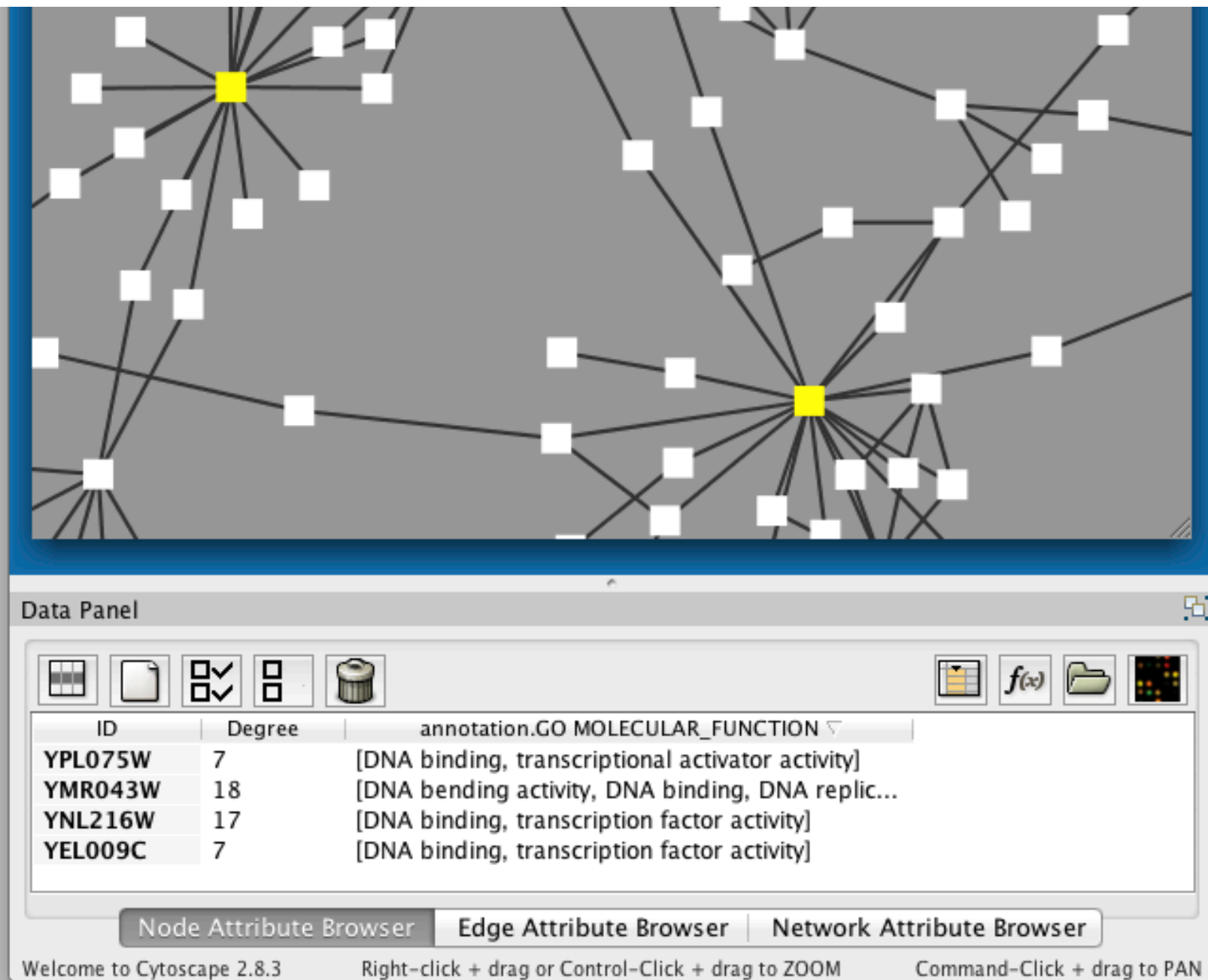
Advanced

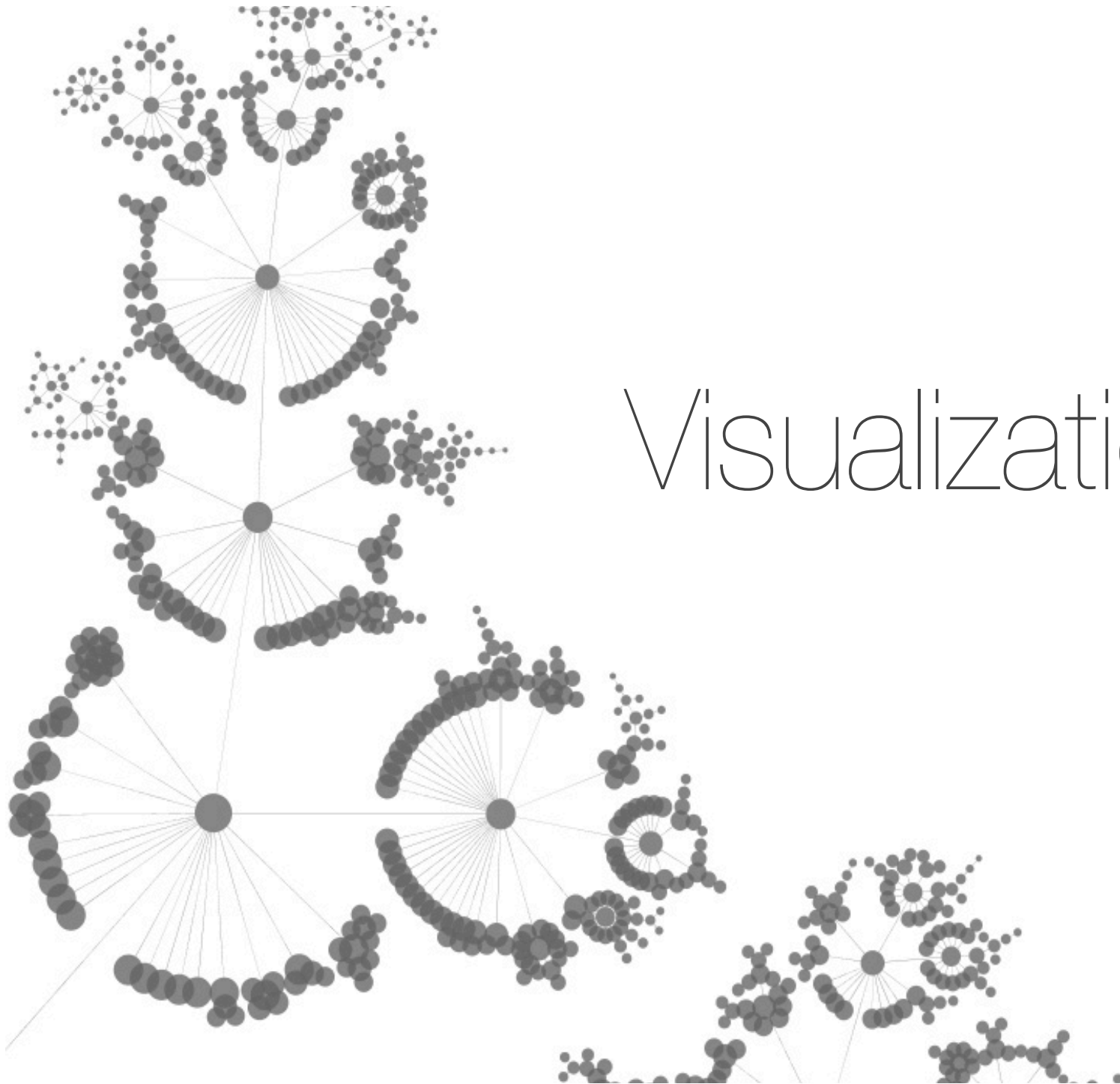
Degree  Not AND

annotation.GO MOLECULAR\_FUNCTION  Not dna binding

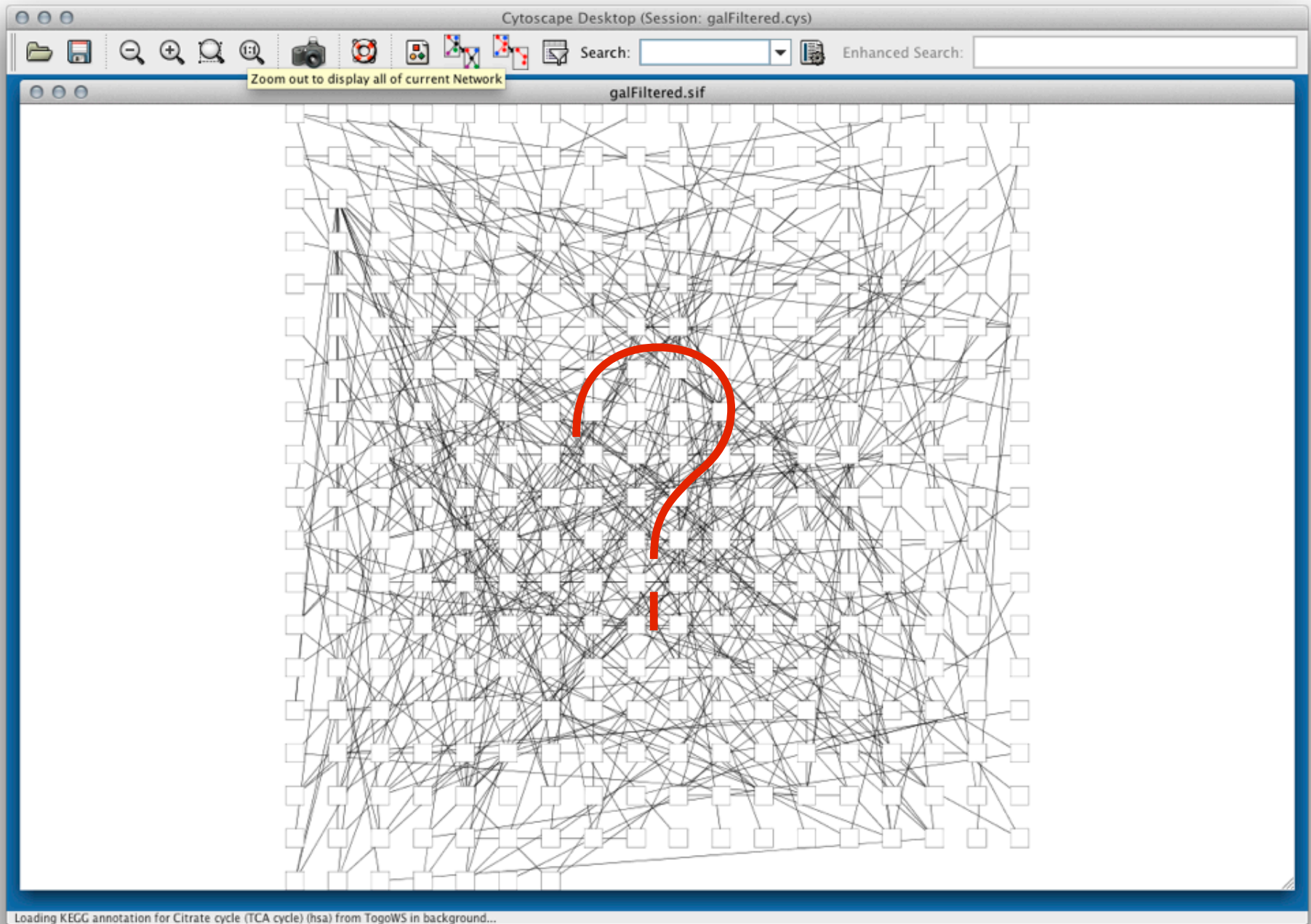
Apply Filter    Select All    Deselect All

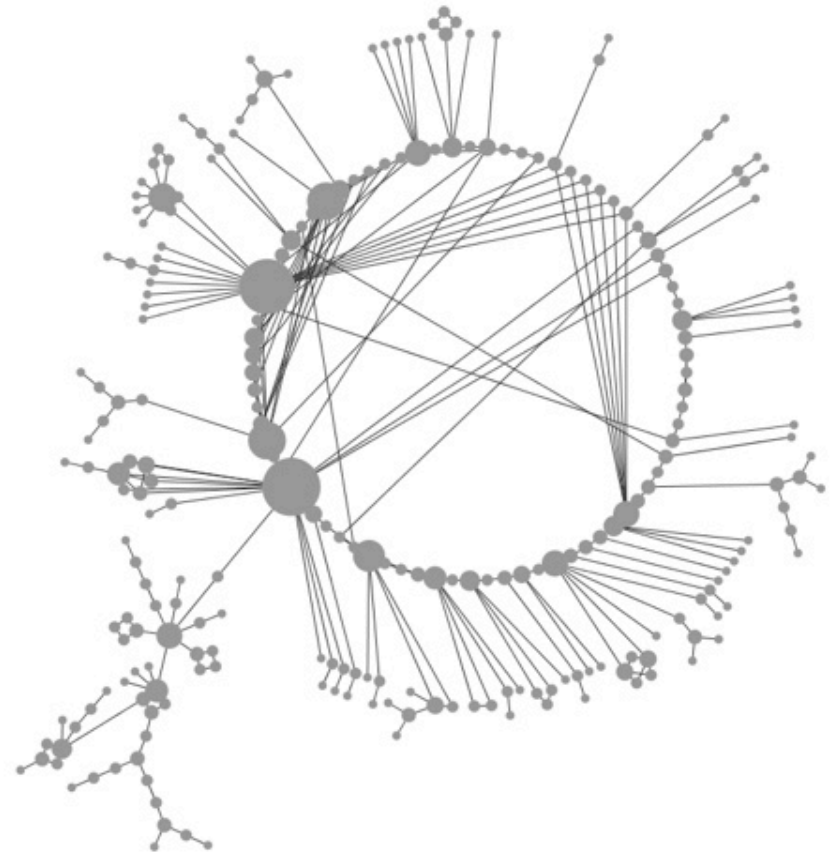
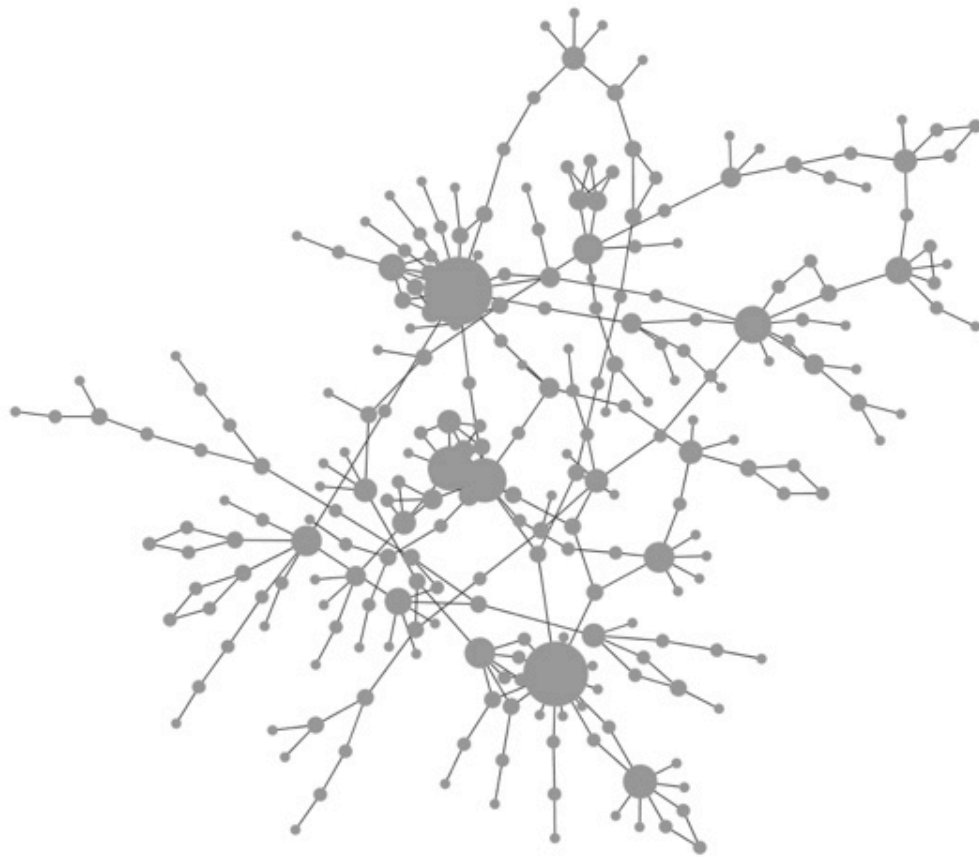




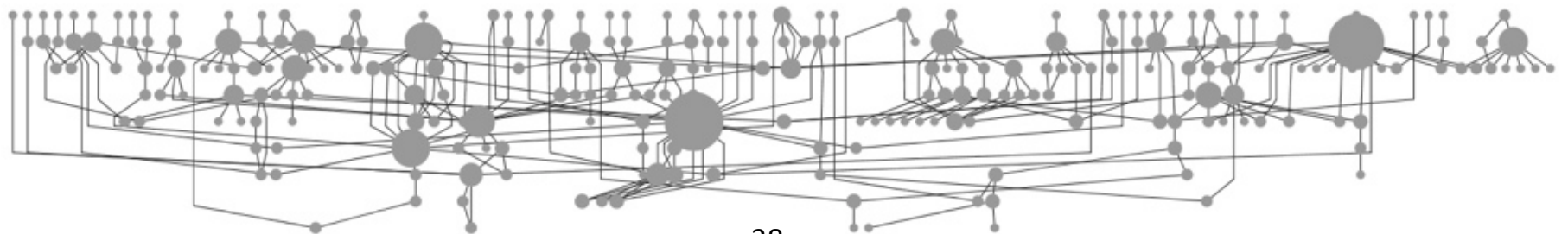


# Visualization



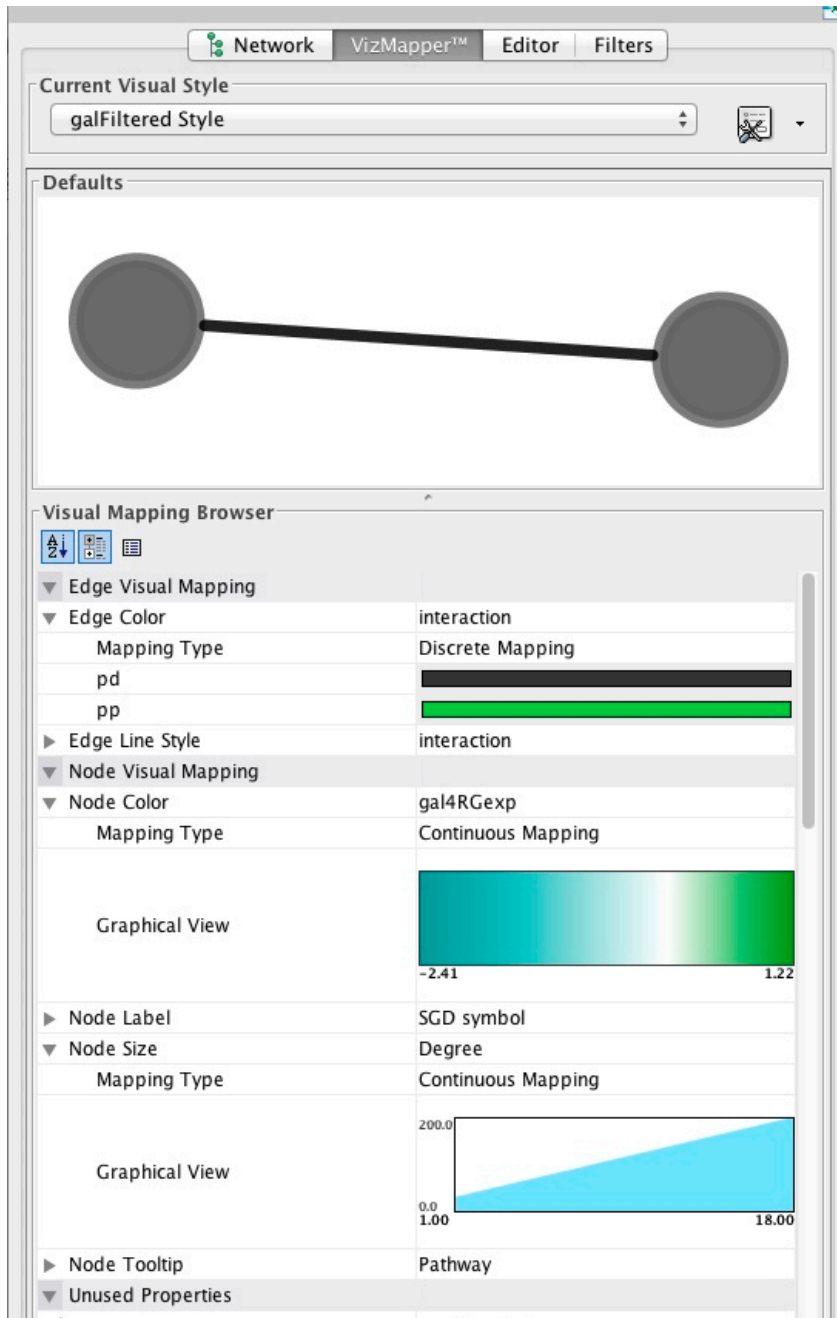


# Layouts



# Biological Data Visualization

- Help others to understand your data
- Emphasize what you want to *tell* by the image
  - Use **color**, **SHAPE**, **Size** of objects effectively!
- **Tamara Munzner Web Site:**  
<http://www.cs.ubc.ca/~tmm/>

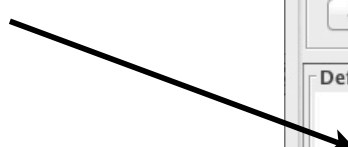


# Visual Style

- Collection of mappings from **Attributes** to **Visual Properties**



Default View Editor

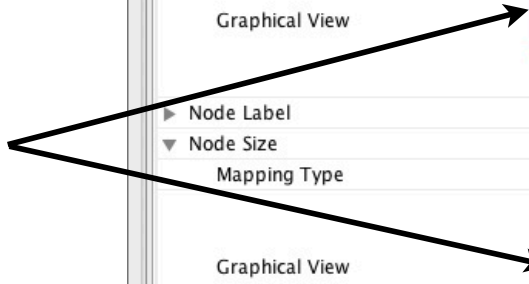


The screenshot shows the VizMapper™ Editor interface. At the top, there are tabs for 'Network', 'VizMapper™', 'Editor', and 'Filters'. Below this, the 'Current Visual Style' is set to 'galFiltered Style'. The main visualization area shows a network with two grey circular nodes connected by a thick black line. Below the visualization is the 'Visual Mapping Browser' panel, which is divided into sections for 'Edge Visual Mapping' and 'Node Visual Mapping'. The 'Edge Visual Mapping' section includes 'Edge Color' (set to 'interaction' with 'Discrete Mapping') and 'Edge Line Style' (set to 'interaction'). The 'Node Visual Mapping' section includes 'Node Color' (set to 'gal4RGexp' with 'Continuous Mapping') and 'Node Size' (set to 'Degree' with 'Continuous Mapping'). There are two 'Graphical View' sections: one for 'Node Color' showing a color gradient from cyan to green with values -2.41 and 1.22, and one for 'Node Size' showing a size gradient from 0.0 to 200.0 with values 1.00 and 18.00. Other properties like 'Node Label' (SGD symbol) and 'Node Tooltip' (Pathway) are also visible.

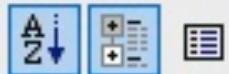
Discrete Mapping Editor



Continuous Mapping Editor



## Visual Mapping Browser



### ▼ Edge Visual Mapping

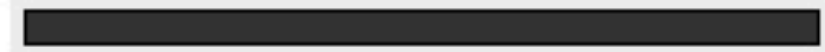
#### ▼ Edge Color

Mapping Type

interaction

Discrete Mapping

pd



pp



#### ▶ Edge Line Style

interaction

### ▼ Node Visual Mapping

#### ▼ Node Color

gal4RGexp

Mapping Type




Continuous Mapping

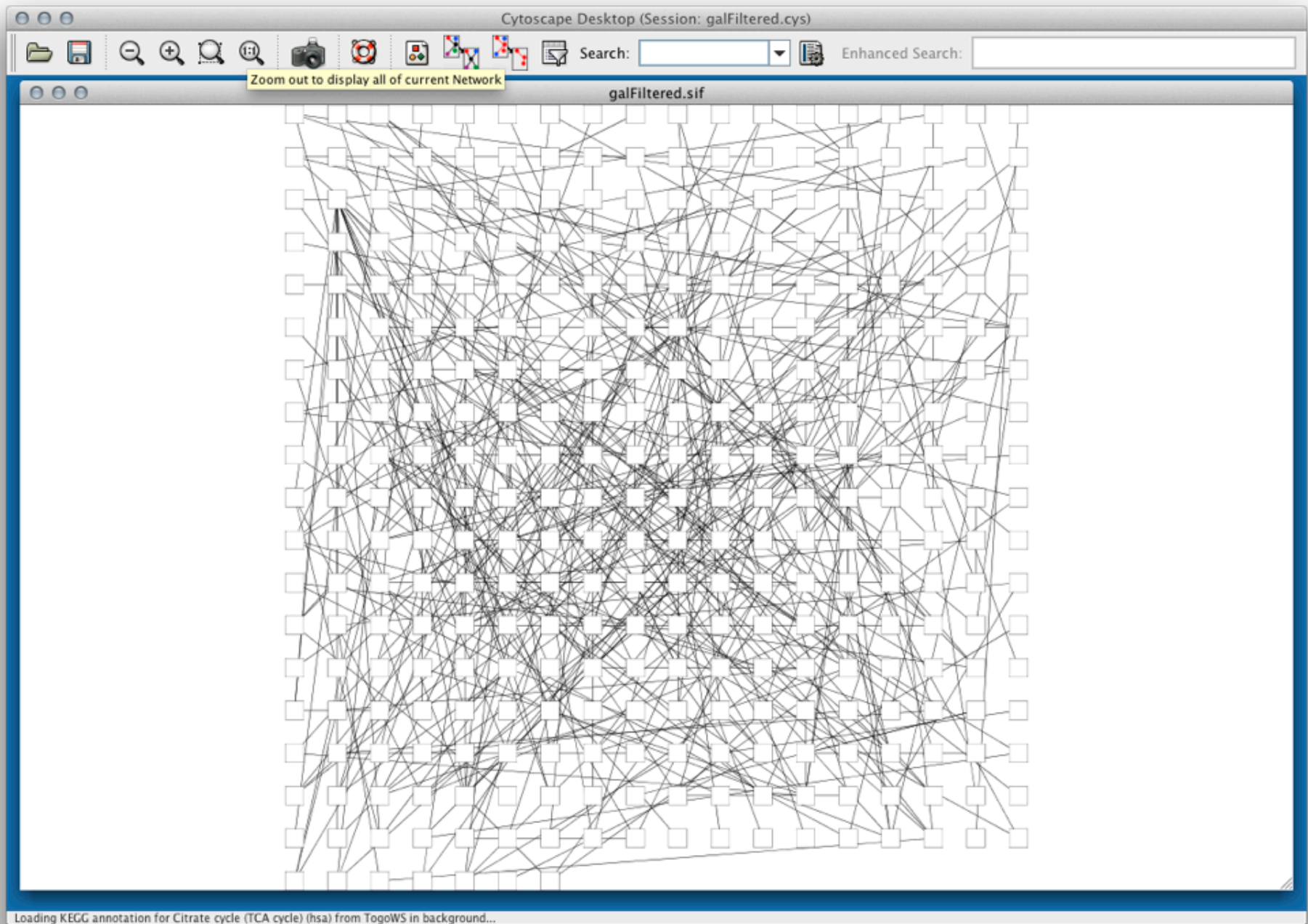
Graphical View



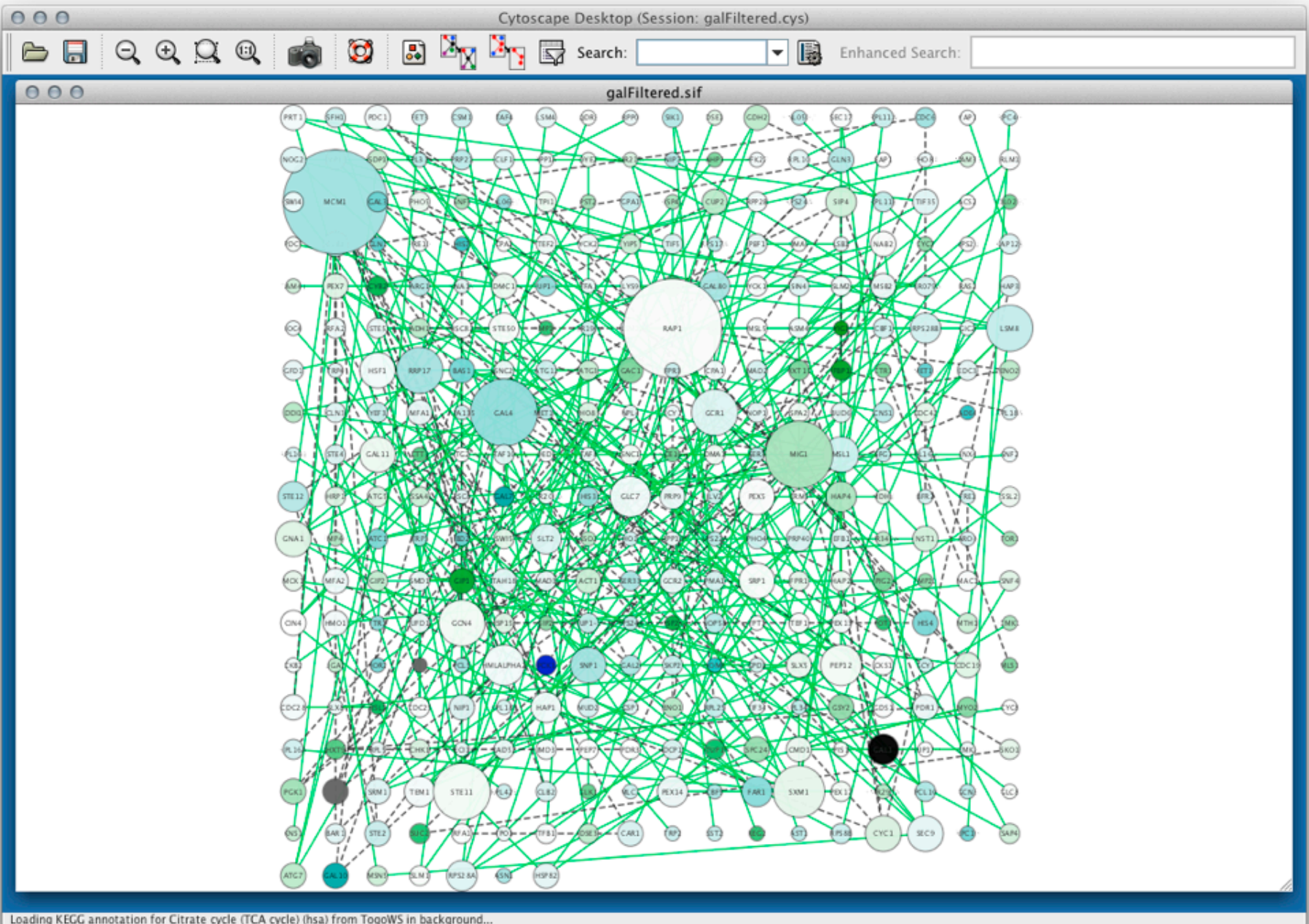
-2.41

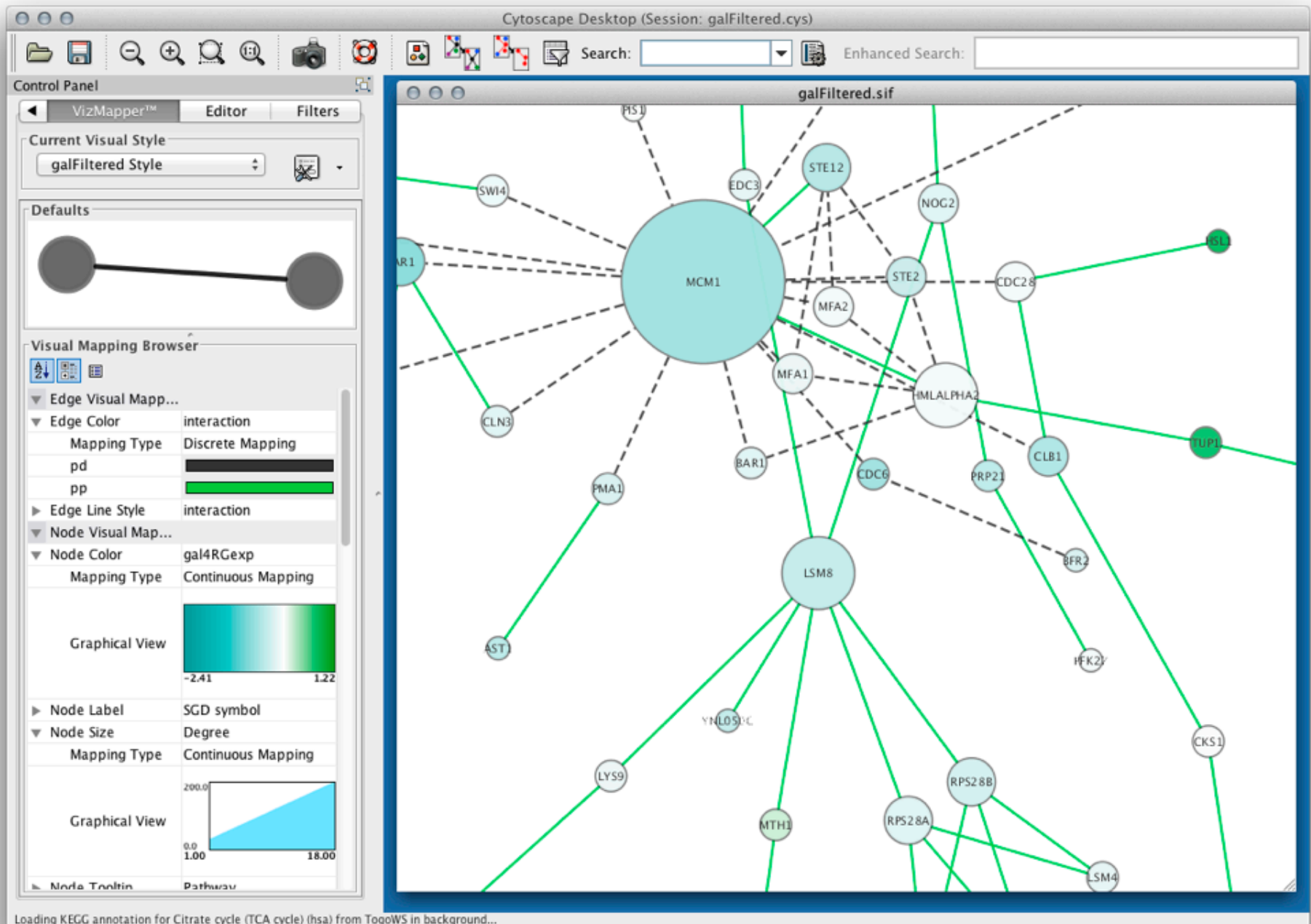
1.22

pp		
▶ Edge Line Style		interaction
▼ Node Visual Mapping		
▼ Node Color		gal4RGexp
Mapping Type		Continuous Mapping
Graphical View		 -2.41 1.22
▶ Node Label		SGD symbol
▼ Node Size		Degree
Mapping Type		Continuous Mapping
Graphical View		 200.0 0.0 1.00 18.00
▶ Node Tooltip		Pathway
▼ Unused Properties		





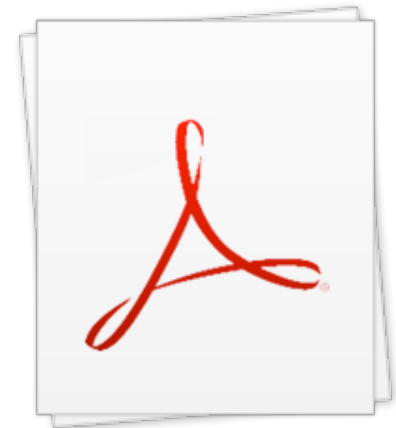






# Prepare for Publication

- Network images can be exported as PDF/PS/PNG/JPG.
- Use PDF for your publications

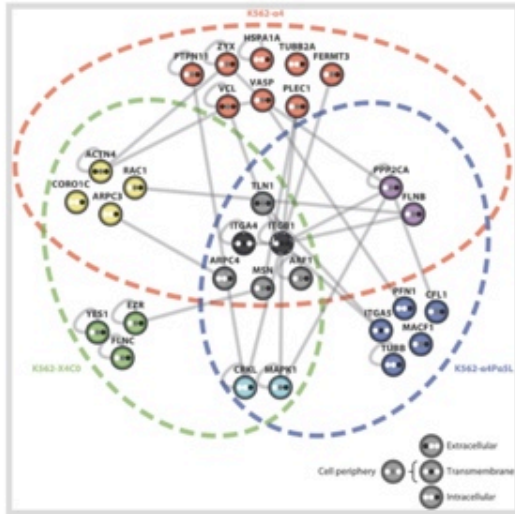


# Cytoscape Publications

Showcase for publications using Cytoscape or describing Cytoscape plugins.

[Cytoscape Website](#)

AUG  
14

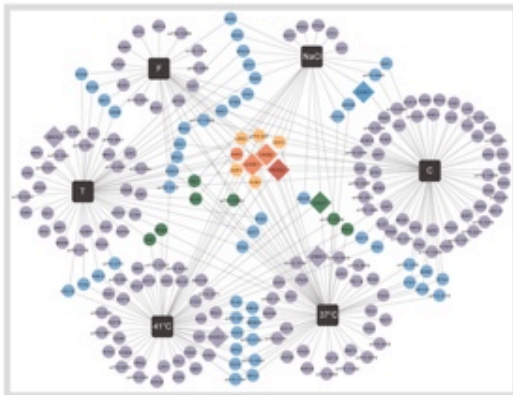


Proteomic analysis of  $\alpha_4\beta_1$  integrin adhesion complexes reveals  $\alpha$ -subunit-dependent protein recruitment.

Byron et al., *Proteomics*. 2012 Jul;12(13):2107-14.

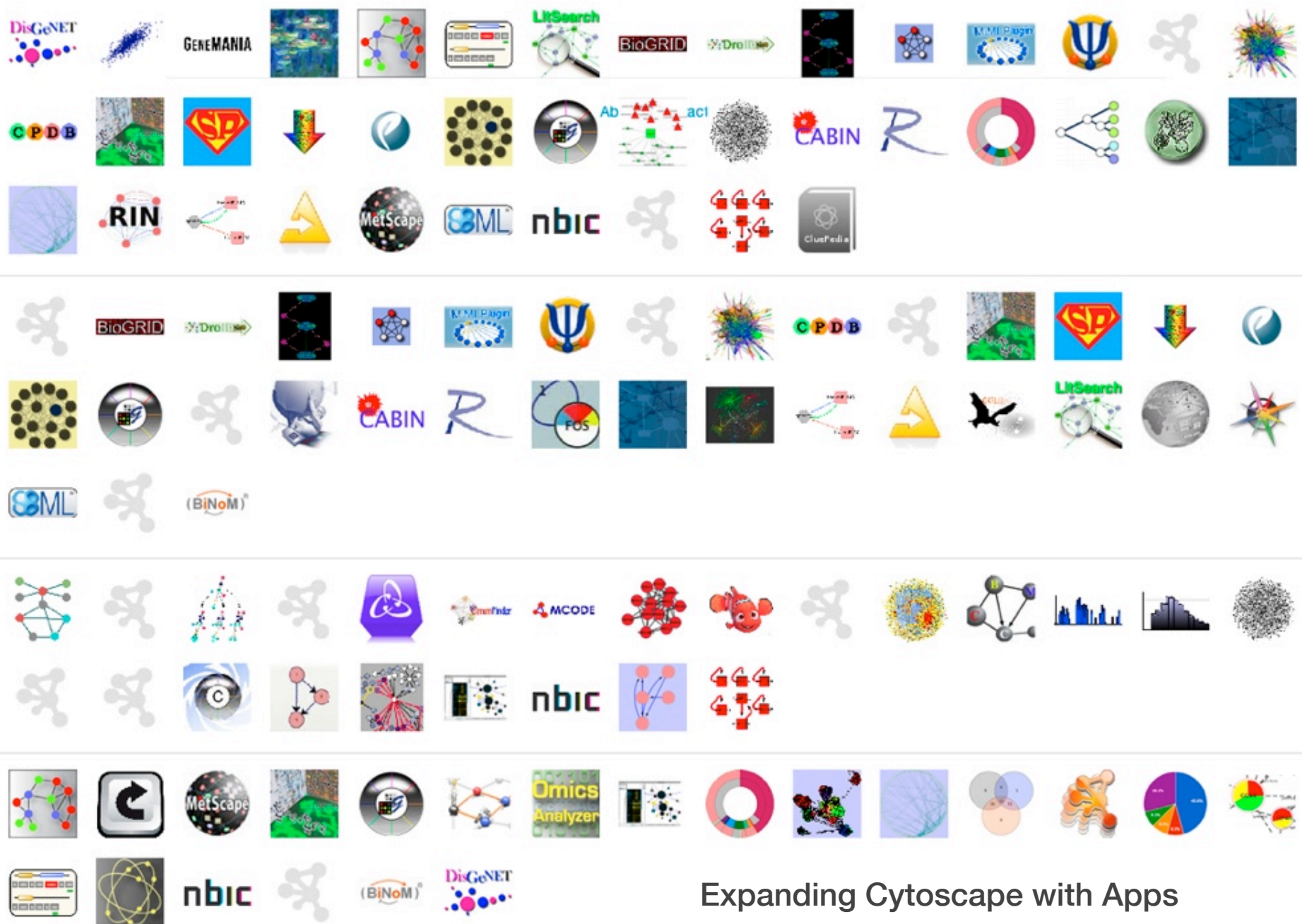
[Permalink](#) Posted at 6:36 PM

AUG  
01



# Real World Examples

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



## Expanding Cytoscape with Apps

# Cytoscape Apps

- Were called Plugins
- Add new features to Cytoscape
- Large app developer/user community
  - This is the reason why Cytoscape is so popular!

**Categories**

- network generation
- online data import
- graph analysis
- data visualization

**Featured Apps**



**PSICQUICUniversalClient**  
PSICQUIC Web Service Client for importing interactions from public



**BiNGO**  
Calculates overrepresented GO terms in the network and display them as a

<http://apps.cytoscape.org>



**jActiveModules**  
Finds clusters where member nodes show significant changes in



**BiNGO**  
Calculates overrepresented GO terms in the network and display them as a



**GeneMANIA**  
Imports interaction networks from public databases from a list of genes



**MCODE**  
Clusters a given network based on topology to find densely connected

AllegroMCODE APCluster APID2NET BioQualiPlugin BLAST2SimilarityGraph BNMatch CABIN  
CalculateNodeDegree CentiScaPe ChemViz clusterExplorerPlugin clusterMaker ClusterONE  
ClusterViz COMA CommFinder CyClus3D CyOog cytoHubba Cytomcl DualLayout  
dynamicXpr EnhancedSearch EnrichmentMap ExprEssence GraMoFoNe GraphletCounter  
HiderSlider jActiveModules mcl-new MCODE MINE NeMo NetAtlas NetCirChro netMatch  
NetworkAnalyzer NetworkEvolution OmicsAnalyzer OmicsViz OrthoNets PanGIA  
PerturbationAnalyzer PinnacleZ RandomNetworks RDFScsape Reactome Fls  
RemainingDegreeDistribution ReOrientPlugin ShortestPath Plugin SimTrek structureViz  
TransClust VennDiagramGenerator VistoClusPlugin WordCloud BiNeM bioCycPlugin  
BiogridPlugin BioGRIDPlugin BioGRIDPlugin BioGRIDPlugin BioGRIDPlugin BioGRIDPlugin  
CoryneRegNet CytoscapePlugin CytoscapePlugin CytoscapePlugin CytoscapePlugin CytoscapePlugin  
GraphMLReader GraphMLReader GraphMLReader GraphMLReader GraphMLReader GraphMLReader  
NCBIEntrezGateway NCBIEntrezGateway NCBIEntrezGateway NCBIEntrezGateway NCBIEntrezGateway  
PICRClient PICRClient PICRClient PICRClient PICRClient PICRClient  
SuperpathwaysPlugin SDBScape AgentLiteratureSearch Cytoprophet DisGeNET  
DomainGraph ExpressionCorrelation GeneMANIA MetaNetter MONET BiNGO BubbleRouter  
ClueGO CommonAttributes DisplayNetworkFromFlatFile FluxViz FunNetViz HyperEdgeEditor  
PiNGO CyGoose CytoscapeRPC GroovyScriptingEngine JavaScriptEngine MiSink  
PythonScriptingEngine RubyScriptingEngine ScriptEngineManager addParentNeighbors  
AdvancedNetworkMerge batchTool BiLayout commandTool coreCommands CyAnimator  
edgeLengthPlugin edgeLister EpiTrace FERN FM3 GoogleChartFunctions GroupTool  
MetaNodePlugin2 MultilevelLayoutPlugin NamedSelection NatureProtocolsWorkflow  
NeighborHighlight NetLink nodeCharts PhyloTree VennDiagrams

**140+ Apps**



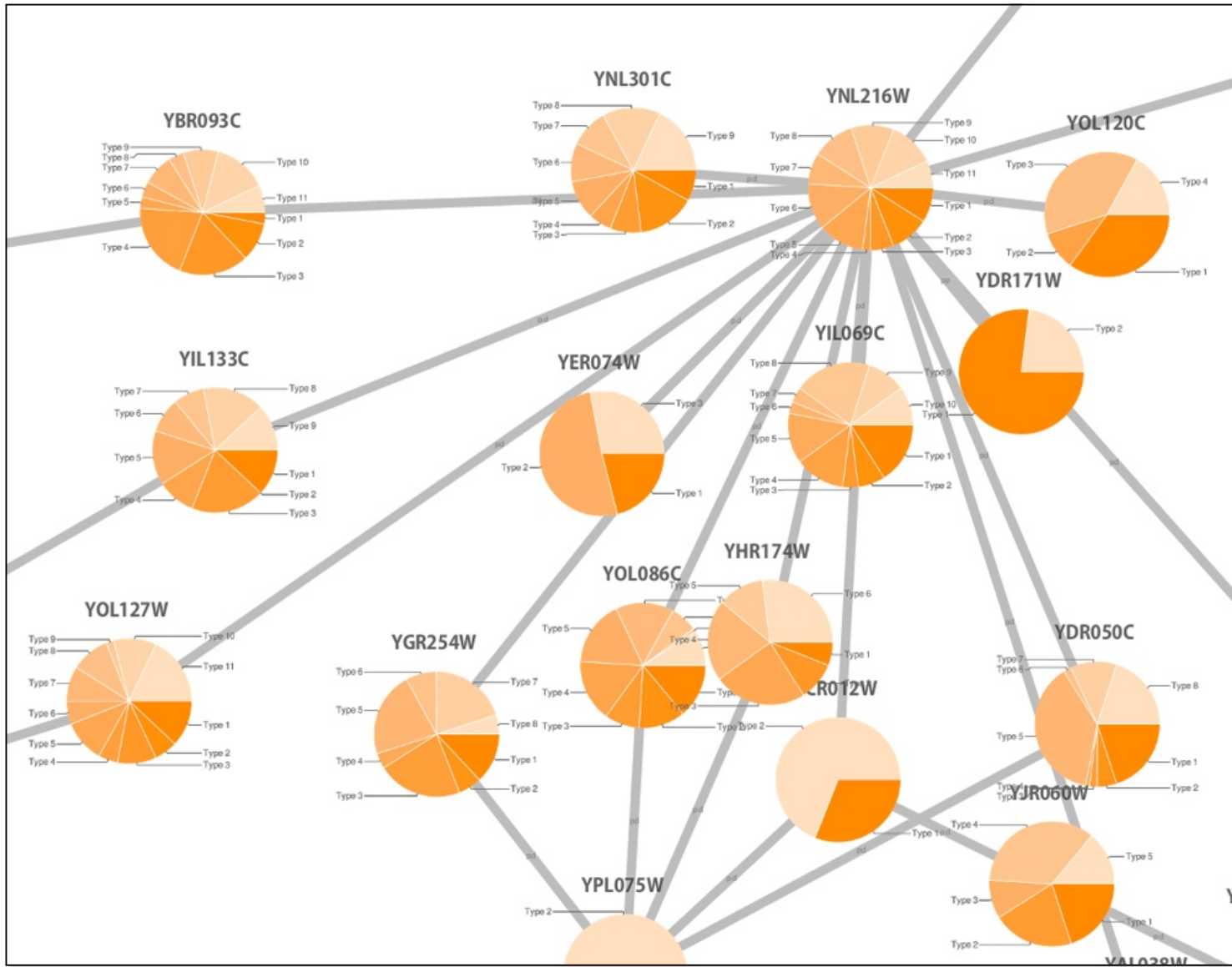
# A Must Read

## *A travel guide to Cytoscape plugins*

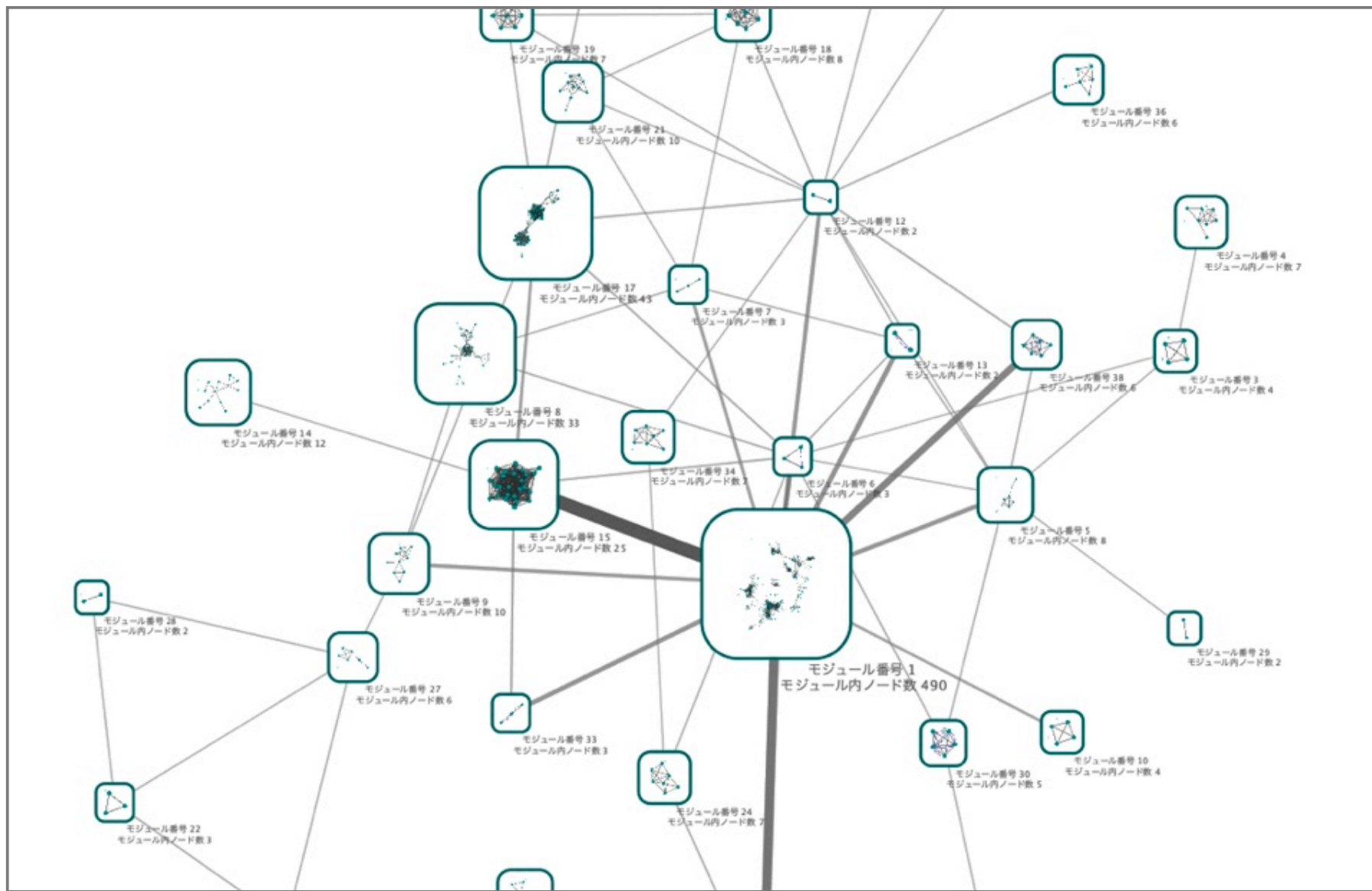
Rintaro Saito, Michael E Smoot, Keiichiro Ono, Johannes Ruscheinski, Peng-Liang Wang, Samad Lotia, Alexander R Pico, Gary D Bader, Trey Ideker  
(2012)

Nature Methods 9 (11) p. 1069-1076

# Advanced Topics



# Custom Graphics



# Nested Networks

# Getting Help

- Two Google Groups
  - cytoscape-**discuss**@googlegroups.com
  - cytoscape-**helpdesk**@googlegroups.com
- ANY question is OK!



University of California  
San Francisco

P41 RR031228 (NRNB)



**Agilent Technologies**



Memorial Sloan-Kettering  
Cancer Center

NIH GM070743-01



**UCSD**



*Revolutionizing science. Enhancing life.*

P01 HG005062



UNIVERSITY OF  
**TORONTO**

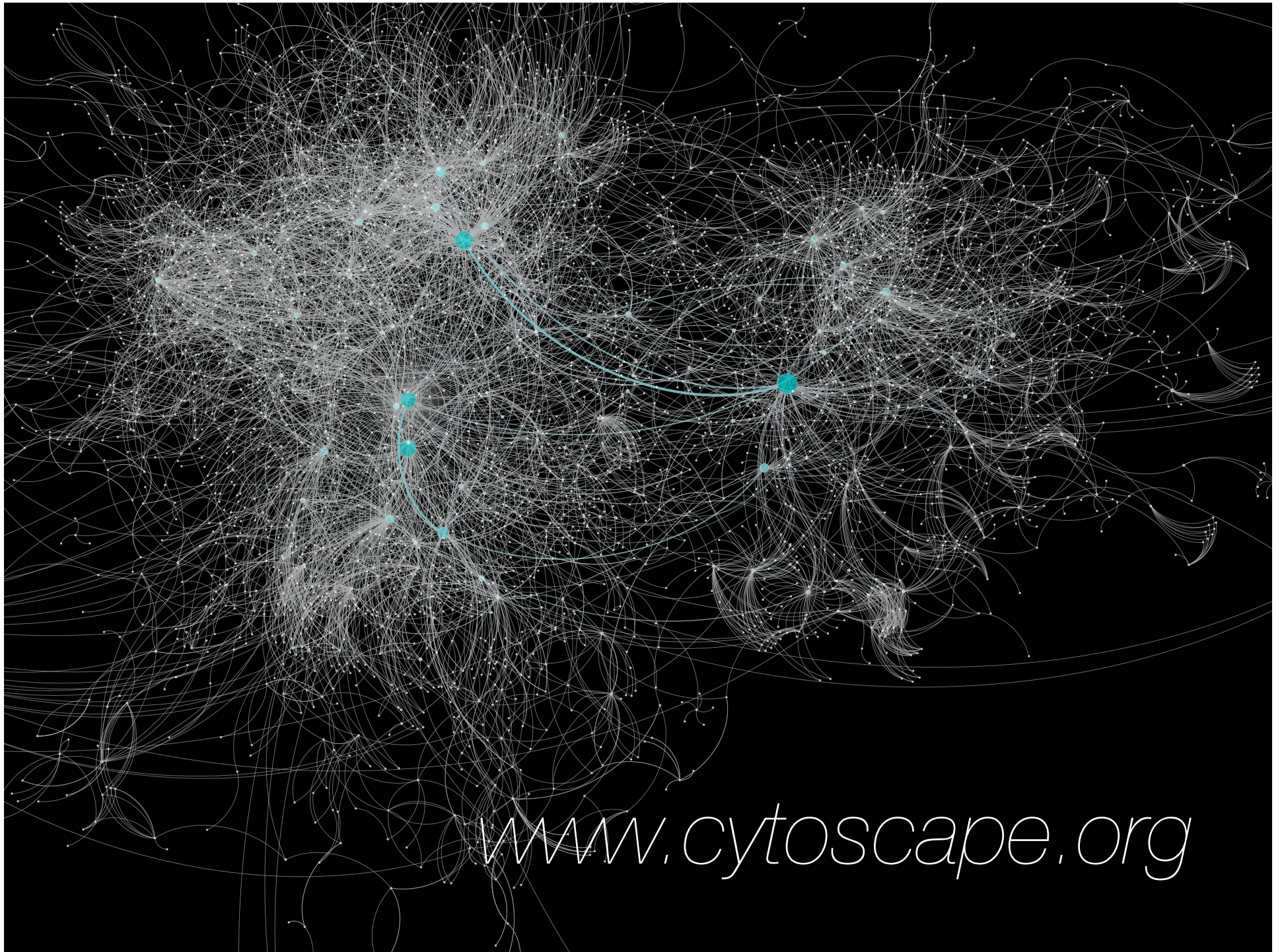


Unilever



Institut Pasteur





[www.cytoscape.org](http://www.cytoscape.org)