

NDExR and Cytoscape: Interactive and automated visualization of biological networks using R

by Florian Auer

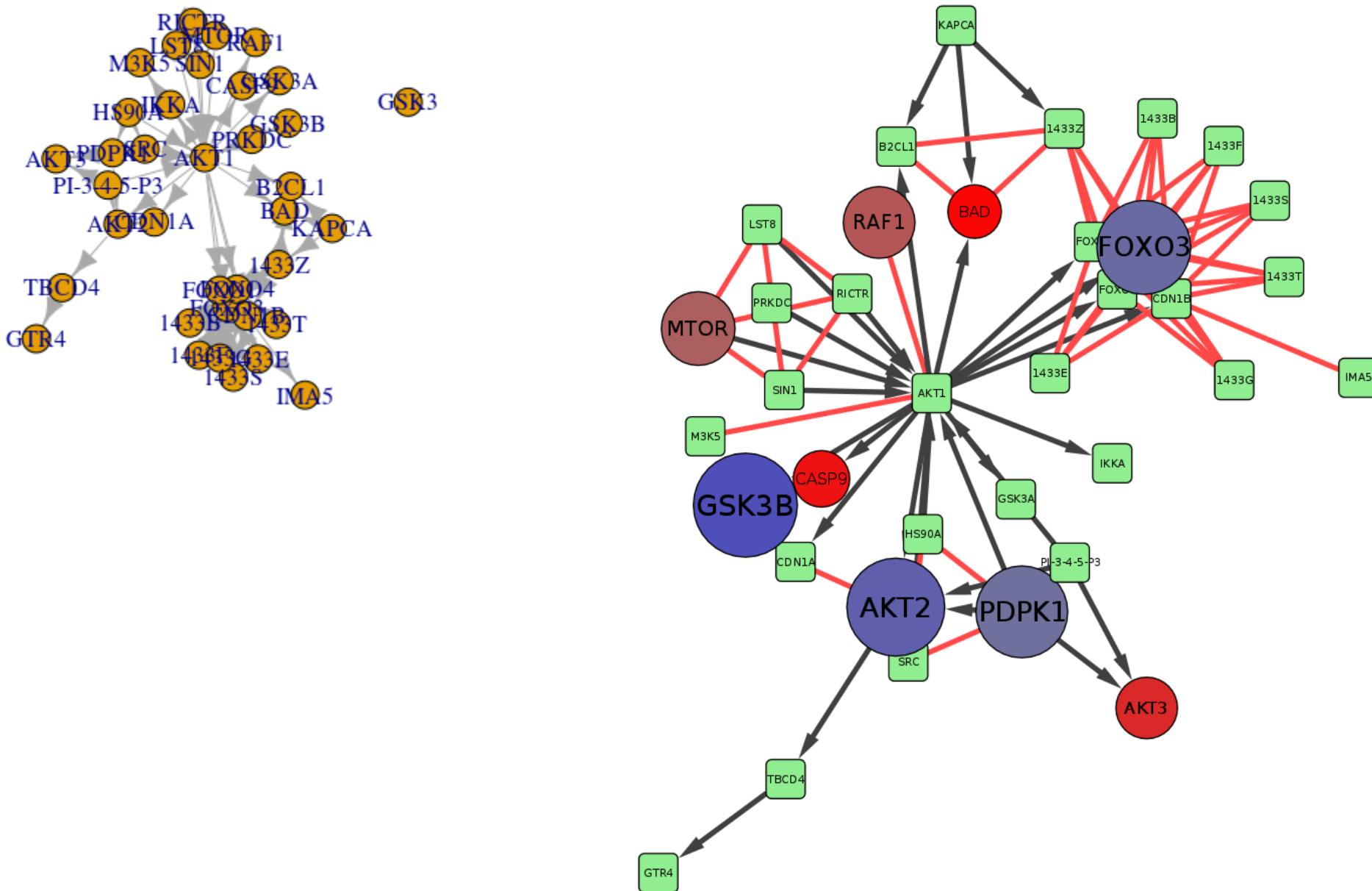
Frank Kramer, Tim Beißbarth

Statistical Bioinformatics,
University Medical Center Göttingen

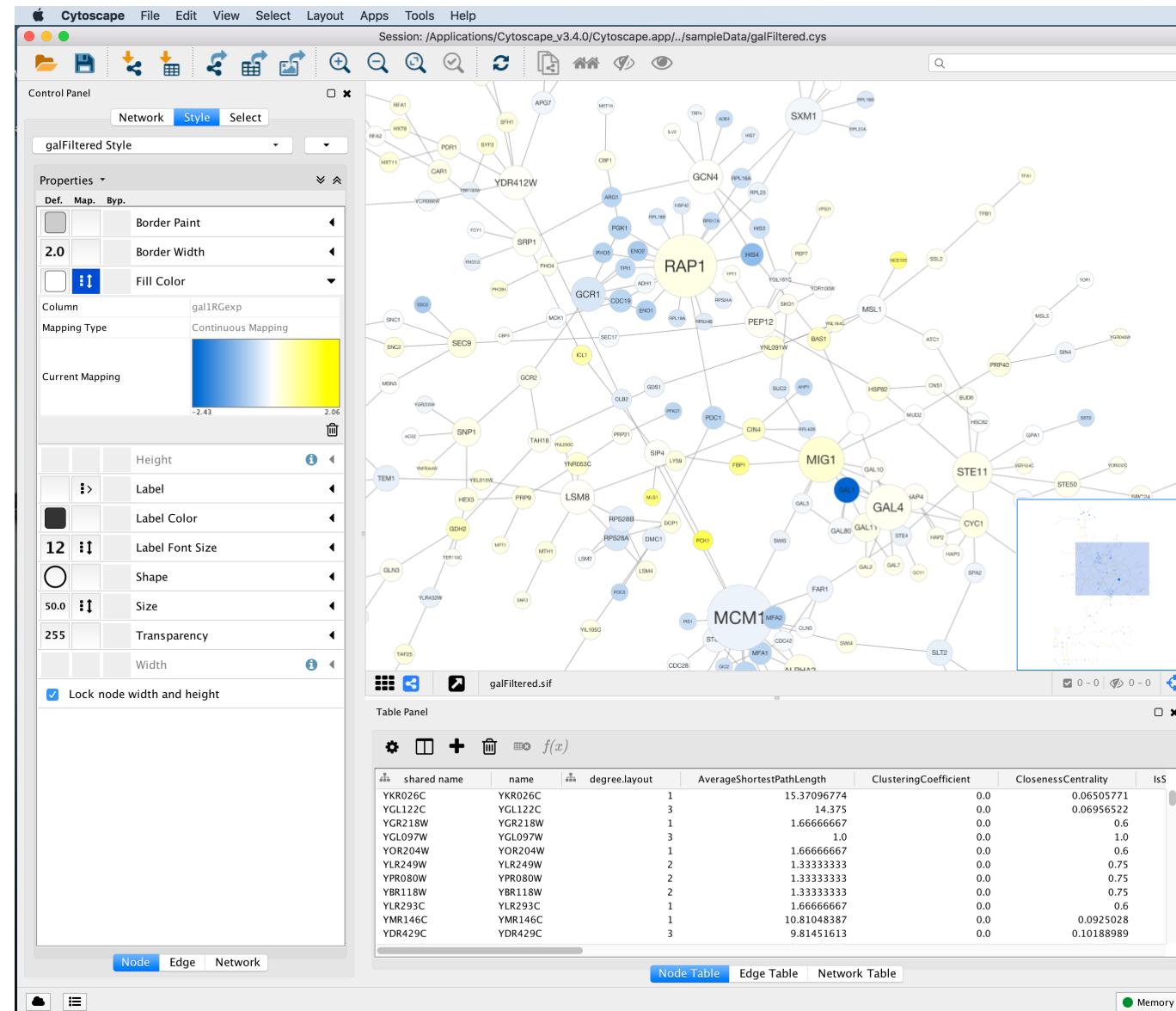
Biological Networks

- Gene regulatory network
- Signaling network
- Protein-Protein interaction
- Metabolic network
- Drug-disease associations
- ...

Network Visualization



Cytoscape



- Open source
- Cross platform
- Consortium



Institute for Systems Biology



University of California at San Diego



Institut Pasteur



UCSF University of California at San Francisco

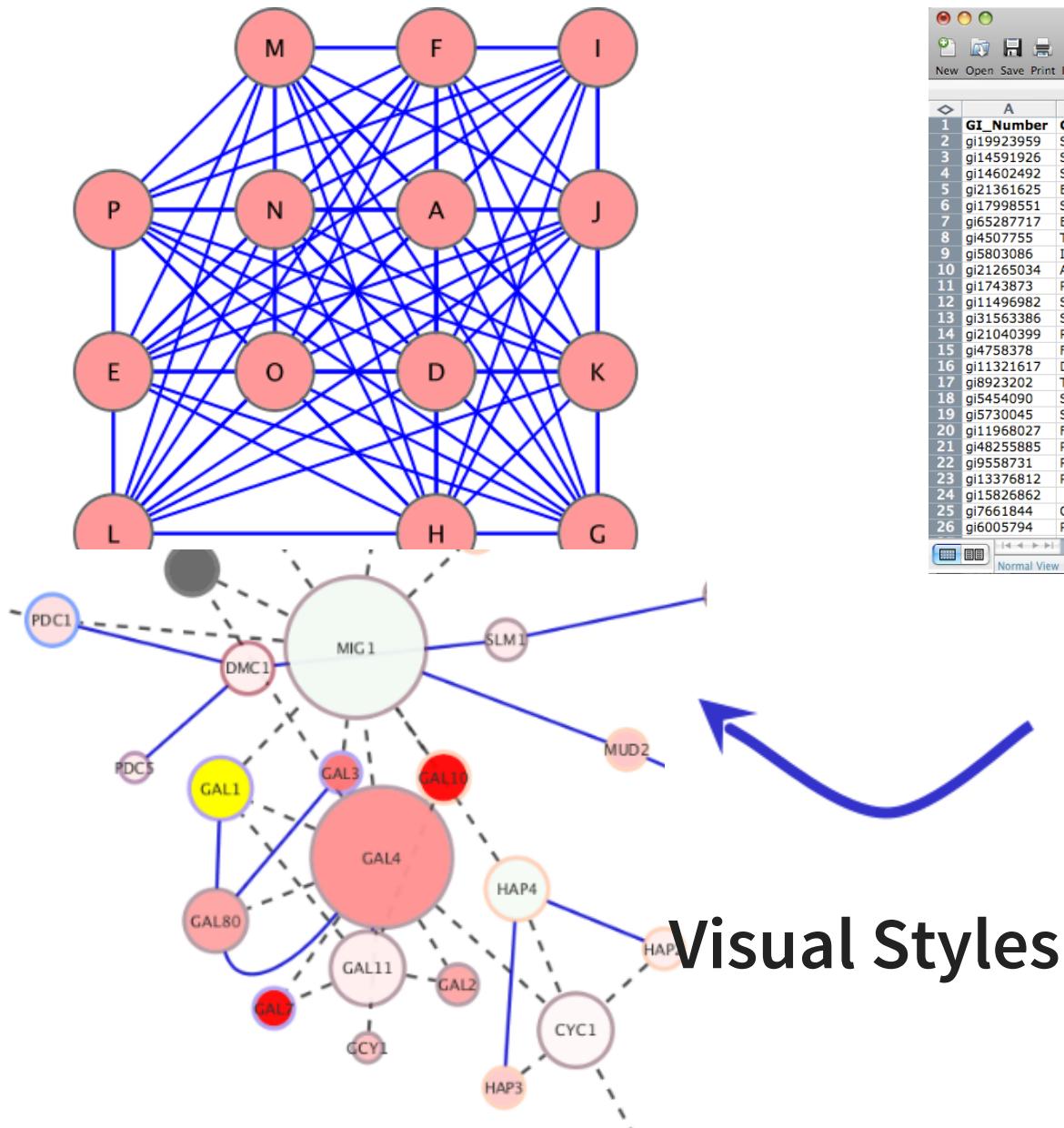


University of Toronto

**GLADSTONE
INSTITUTES**



Core Concepts: Networks and Tables



	A	B	C	D	E	F	G	H	I	J
1	GI_Number	Gene	HPRD_ID	REFSEQ	Entrez_ID	OMIM_ID	SWISSF_Gene_Description		Architecture	Arch Motif
2	gi19923959	SDSL	11542	NP_612441.1	113675		SDSL		SP	Motif
3	gi14591926	SEC23B	11543	NP_116781.1	10483		Q15437 Protein transport protein SEC23B		GEL	
4	gi14602492	SCMH1	11540	AAH09752.1	22955		Q96GD3 SCMH1		SAM	
5	gi21361625	EXOC2	11544	NP_060773.3	55770		Q96KP1 SEC5 like 1		PT	
6	gi17998551	SERPINB12	11548	NP_536722.1	89777		Q96P63 Serpin B12		SERPIN	
7	gi65287717	EIF2AK4	18447	NP_00101372	440275	609280	Similar to GCN2 eIF2alpha kinase	S_T_Y_Kinase		
8	gi4507755	TYROBP	4996	NP_003323.1	7305	604142	O43914 DAP12	ITAM		
9	gi5803086	IL24	4995	NP_006841.1	11009	604136	Q13007 Interleukin 24	IL10		
10	gi21265034	ADAMTS13	4994	NP_620594.1	11093	604134	ADAMTS 13	TSP1		
11	gi1743873	PIP5K1A	4470	AAC50911.1	8394	603275	Q99755 Phosphatidylinositol-4-phosphate 5 PIPk			
12	gi11496982	SVIL	4992	NP_068506.1	6840	604126	Q95425 Supervillin	VHP		
13	gi31563386	SULT2B1	4991	NP_814444.1	6820	604125	Sulfotransferase family 2B, member 1			
14	gi21040399	RBBP8	4990	AAH30590.1	5932	604124	Q99708 CTIP	LZ	Motif	
15	gi4758378	FIGF	2102	NP_004460.1	2277	300091	O43915 VEGF D	PDGF		
16	gi11321617	DPYSL4	7463	NP_006417.1	10570	608407	O14531 Collapsin response mediator protein 3			
17	gi8923202	TASP1	7460	NP_060184.1	55617	608270	Q9H6P5 Threonine aspartase 1			
18	gi5454090	SSR4	2101	NP_006271.1	6748	300090	P51571 Signal sequence receptor delta	SP	Motif	
19	gi5730045	SLC16A2	2106	NP_006508.1	6567	300095	P36021 X linked PEST containing transport	TM		
20	gi11968027	FTS	7467	NP_071921.1	64400	608483	Q9H8T0 FTS	UBC		
21	gi48255885	PRKCI	2105	NP_002731.3	5584	600539	P41743 Protein kinase C, iota type	S_T_kinase		
22	gi9558731	RPA4	6591	NP_037479.1	29935		RPA4	TRNAA		
23	gi13376812	PPP1R2P9	6593	NP_079486.1	80316		Type 1 protein phosphatase inhibitor			
24	gi15826862		6595	NP_296375.1	90060		JM11 protein	CC	Motif	
25	gi7661844	CCDC22	6594	NP_054727.1	28952		JM1 protein	CC		
26	gi6005794	PRAF2	6596	NP_009144.1	11230		JM4 protein	TM		

Tables

Visual Styles

Loading Network Data

- Cytoscape can import network data from:
 - Files (or URLs)
 - Excel, TSV, CSV
 - XGMML: eXtensible Graph Markup and Modelling Language
 - SBML: Systems Biology Markup Language
 - BioPAX
 - PSI-MI
 - SIF: Simple Interaction Format
 - GML: Graph Markup Language
 - ... and others depending on loaded Apps

Core Concepts: Cytoscape Apps

Over 300 Apps!

Wall of Apps 295 total



apps.cytoscape.org

Saving and Exporting

- Sessions save everything as **.cys** files:
Networks, Tables, Styles, Screen sizes, etc.
- Export networks in different formats:
SIF, GML, XGMML, BioPAX, JSON
- Export tables as CSV files
- Publication quality graphics in several formats:
PDF, SVG, PNG, and JPEG

Loading Network Data

Interactive Graph Editing & Layout « AllegroLayout | Cytoscape 3 App



Loading Network Data

- Cytoscape can import network data from:
 - Public repositories:
 - **EBI/PSICQUIC**
 - **STRING** (via the stringApp)
 - **Reactome** (via the ReactomeFI app)
 - **WikiPathways** (via the WikiPathways app)
 - **Pathway Commons** (via the CyPath2 app)
 - Automation:
 - **Command line scripts**
 - **CyREST via R, Python, etc**



The Network Data Exchange

ndexbio.org

CyNDEX App

"The NDEX Project provides an open-source framework where scientists and organizations can share, store, manipulate, and publish biological network knowledge."

Featured collections include:

- Pathway Interaction Database (NCI-PID)
- Cancer Cell Maps Initiative (CCMI)
- The NDEX Butler
- NetPath

Search on NDEX



News

About

Docs

Report Bug

Contact Us

Cite NDEX

Sign In

Search Examples ▾

Browse

All ▾

nci-pid PI3K Akt signaling

 Perform Search Term Expansion (Genes and Proteins only)[Networks \(429\)](#) [Users \(6\)](#) [Groups \(1\)](#)

	Network Name		Ref.	Disease	Tissue	Nodes	Edges	Visibility	Owner	Last Modified	
	Class I PI3K signaling events mediated by Akt					37	113	PUBLIC	nci-pid	7/6/17 9:58 PM	
	NCI Pathway Similarity Map			Cancer		212	492	PUBLIC	nci-pid	11/2/17 6:05 PM	
	Trk receptor signaling mediated by PI3K and PLC-gamma					43	188	PUBLIC	nci-pid	7/6/17 9:58 PM	
	IL2 signaling events mediated by PI3K					48	293	PUBLIC	nci-pid	7/6/17 9:58 PM	
	Class I PI3K signaling events					58	478	PUBLIC	nci-pid	7/6/17 9:58 PM	
	Glioblastoma Multiforme					23	130	PUBLIC	signor	3/7/18 2:25 AM	
	Prostate Cancer					25	70	PUBLIC	signor	3/7/18 2:26 AM	
	Thyroid cancer					29	55	PUBLIC	signor	3/7/18 2:27 AM	
	Insulin Receptor					35	114	PUBLIC	signor	3/7/18 2:26 AM	
	Class IB PI3K non-lipid kinase events					5	6	PUBLIC	nci-pid	7/6/17 9:58 PM	
	Luminal Breast Cancer					32	155	PUBLIC	signor	3/7/18 2:26 AM	
	FAP: insulin-mediated adipogenesis					12	66	PUBLIC	signor	3/7/18 2:25 AM	
	TSLP Signalling Pathway					7	7	PUBLIC	netpath	7/19/17 12:09 AM	
	PI3K/AKT					22	70	PUBLIC	signor	3/7/18 2:27 AM	
	Fibroblast Growth Factor 1 Signalling Pathway					46	80	PRIVATE	netpath	7/19/17 12:09 AM	

Total Items: 429

Explore Network on



News

About

Docs

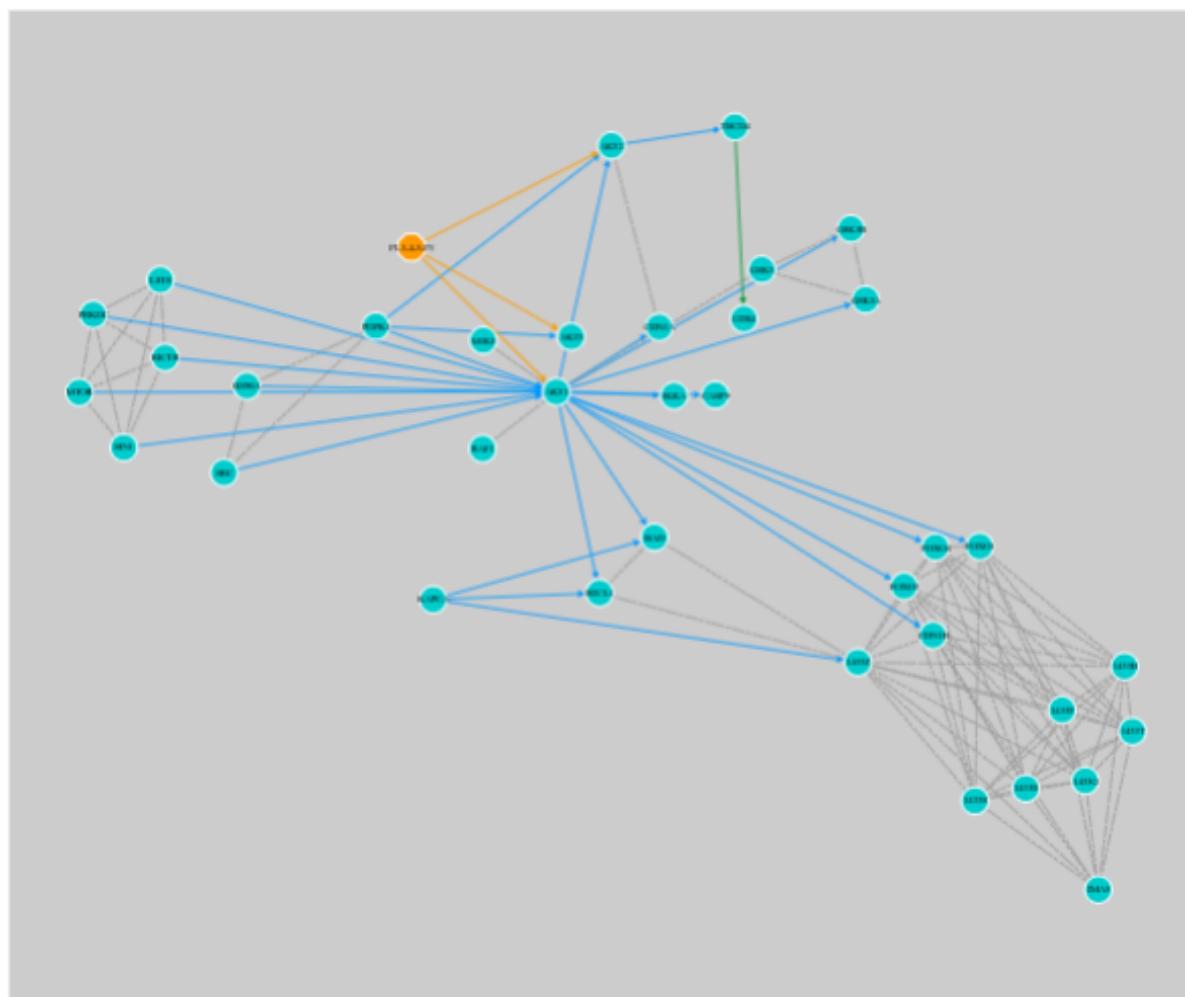
Report Bug

Contact Us

Cite NDEx



Sign In



Network Info

Nodes/Edges

Provenance

Class I PI3K signaling events mediated by Akt

Nodes: 37 Edges: 113

PUBLIC Read Only [Copy URL](#)@context: [view namespaces](#)

Owner: NCI - Pathway Interaction Database (PID)

Created: Sep 23, 2015 3:18:37 AM

Last Modified: Jul 6, 2017 9:58:37 PM

UUID: 046a84e6-6191-11e5-8ac5-06603eb7f303

Format: Unknown

Description: *Class I PI3K signaling events mediated by Akt* was derived from the latest BioPAX3 version of the Pathway Interaction Database (PID) curated by NCI/Nature. The BioPAX was first converted to Extended Binary SIF (EBS) by the PAXTools v5 utility. It was then processed to remove redundant edges, to add a 'directed flow' layout, and to add a graphic style using Cytoscape Visual Properties. This network can be found in searches using its original PID accession id, present in the 'labels' property.

Version: APR-2017**Properties:**

author Kira Anthony

labels PI3KClaktpathway

Network Terms

Depth:

1-step

Run Query

[Advanced Query](#)

Table

Log in

Search NDEx using CyNDEx App



Enter search term...

Control Panel

Network Style Select



Enter search terms for NDEx...



Sample Sessions



Find Networks <@41024c70a463>

Current Source: PUBLIC (<http://public.ndexbio.org/v2>)

Authenticated As: Not Authenticated

nci-pid PI3K Akt signaling|

Results

Network Title	Owned By	Number of Nodes	Number of Edges
Class I PI3K signaling event...	nci-pid	37	113
NCI Pathway Similarity Map	nci-pid	212	492
Trk receptor signaling med...	nci-pid	43	188
IL2 signaling events mediat...	nci-pid	48	293
Class I PI3K signaling events	nci-pid	58	478
Glioblastoma Multiforme	signor	23	130
Prostate Cancer	signor	25	70
Thyroid cancer	signor	29	55
Insulin Receptor	signor	35	114
Class IB PI3K non-lipid kina...	nci-pid	5	6
Luminal Breast Cancer	signor	32	155
FAP: insulin-mediated adip...	signor	12	66
TSLP Signalling Pathway	netpath	7	7
PI3K/AKT	signor	22	70
Fibroblast Growth Factor-1 ...	netpath	46	80
Cholin Signalling Pathway	netpath	20	62

Explore Network in Cytoscape



Enter search term...



Control Panel ▾ □ ×

Network Style Select



1 of 1 Network selected



▼ Class I PI3K signaling events mediated by Akt

● Class I PI3K signaling ev... 37 113

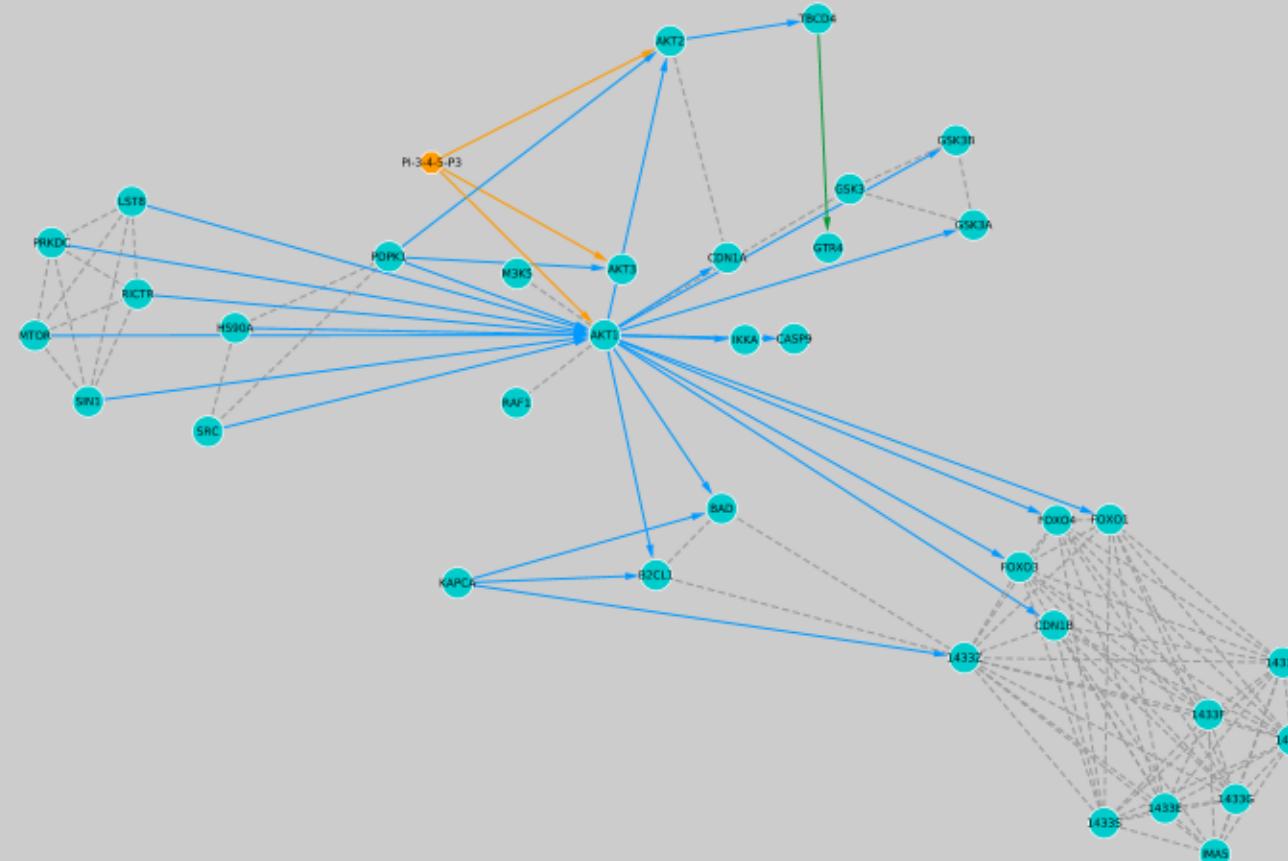

grid undo redo Class I PI3K signaling events mediated by Akt refresh 0 - 0 0 - 0 new

Table Panel

gear list plus trash filter f(x) refresh

shared name	name	alias	type
RICTR	RICTR	[uniprot kn...]	Protein
PRKDC	PRKDC	[uniprot kn...]	Protein
BAD	BAD	[uniprot kn...]	Protein
LST8	LST8	[uniprot kn...]	Protein
KAPCA	KAPCA	[uniprot kn...]	Protein
1433E	1433E	[uniprot kn...]	Protein

Node Table Edge Table Network Table


Unsilencing event source: Class I PI3K signaling events mediated by Akt default node

cloud green dot Memory

The PI3K/AKT signaling pathway: Associations of miRNAs with dysregulated gene expression in colorectal cancer.

Slattery ML, Mullany LE, Sakoda LC, Wolff RK, Stevens JR, Samowitz WS, Herrick JS.

Mol Carcinog. 2018 Feb;57(2):243-261.
doi: 10.1002/mc.22752. Epub 2017 Nov 19.

"The Kyoto Encyclopedia of Genes and Genomes (KEGG) [...] pathway map for PI3K-AKT-signaling was used to identify genes associated with this pathway."

Differential Gene Expression

	Gene <chr>	TumorMean <dbl>	NormalMean <dbl>	FoldChange <dbl>	PValue <dbl>	AdjustedPValue <dbl>
1	BAD	13.13	18.42	0.71	1.38e-11	3.42e-11
2	ITGA3	129.43	172.36	0.75	4.02e-12	1.03e-11
3	ITGA2B	0.54	0.86	0.62	9.90e-03	1.26e-02
4	BRCA1	87.46	35.35	2.47	2.77e-41	3.83e-40
5	IGF1	13.85	38.80	0.36	5.43e-28	3.59e-27

Import Gene Table

File → Import → Table → File... → "PMC5760356.csv"

Import Columns From Table <@41024c70a463>

Target Table Data

Where to Import Table Data: To a Network Collection

Select a Network Collection

Network Collection: Class I PI3K signaling events mediated by ...

Import Data as: Node Table Columns

Key Column for Network: shared na...

Case Sensitive Key Values:

Preview

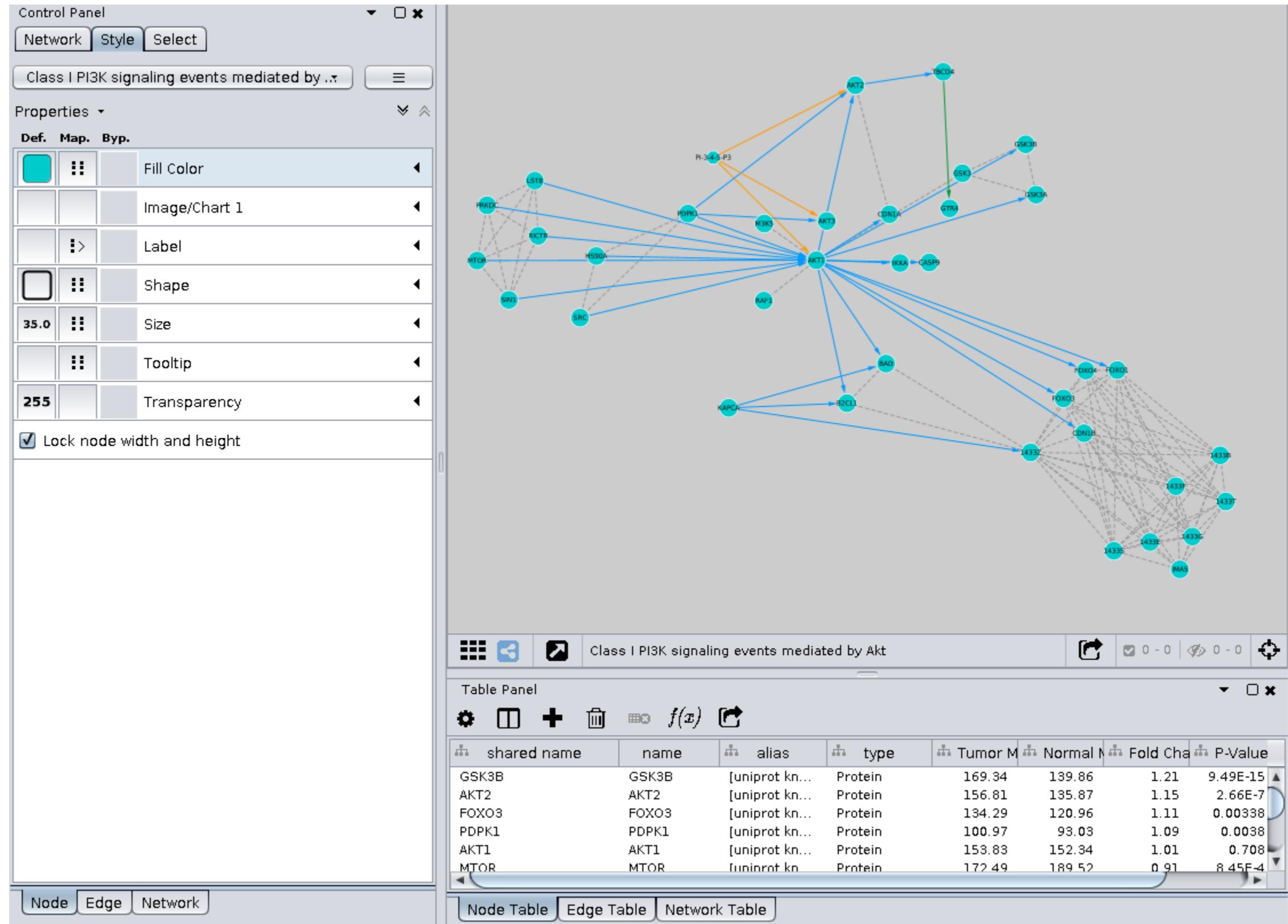
Click on a column to edit it.

Select All Select None

Gene	Tumor Mean	Normal Mean	Fold Change	P-Val
BAD	13.13	18.42	0.71	
ITGA3	129.43	172.36	0.75	
ITGA2B	0.54	0.86	0.62	
BRCA1	87.46	35.35	2.47	
IGF1	13.85	38.80	0.36	
HGF	10.18	14.00	0.73	
FLT4	12.44	14.18	0.88	

Advanced Options...

OK Cancel

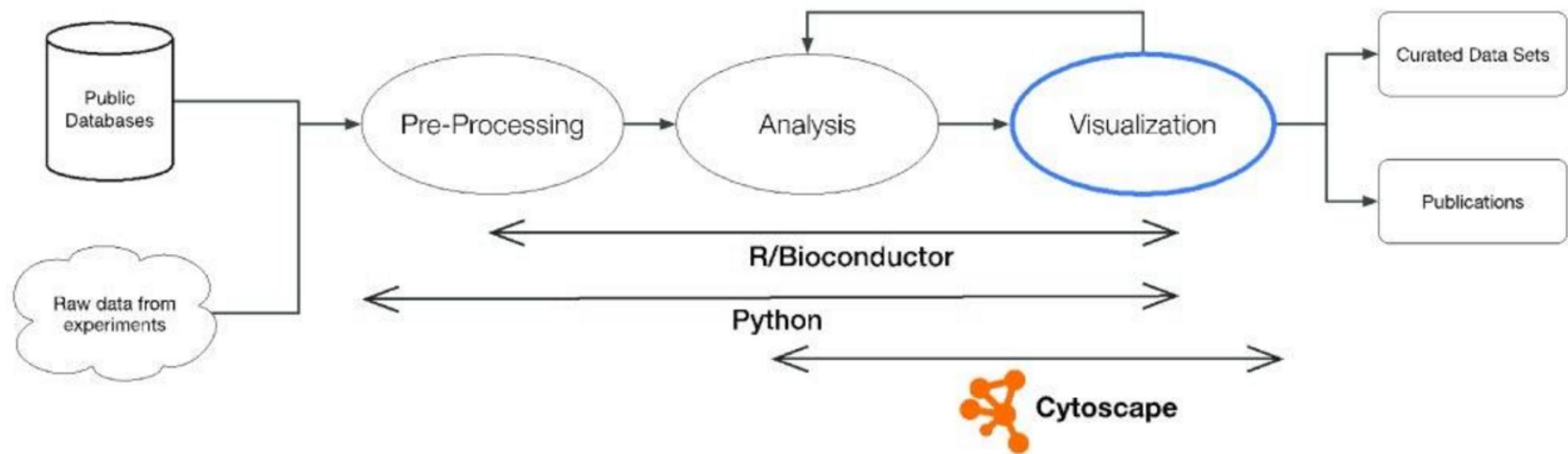


Automate!

Why run a script when I could just use Cytoscape directly?

- For things you want to do multiple times, e.g., loops
- For things you want to repeat in the future
- For things you want to share with colleagues or publish
- For things you are already working on in R or Python, etc

Bioinformatics Workflow



NDExR and Cytoscape

- Download networks from NDEx
- Connect to Cytoscape
- Submit a network to Cytoscape
- Do some fancy stuff
- ALL WITH R!

NDExR - an R package to interface with the Network Data Exchange.

Auer F, Hammoud Z, Ishkin A, Pratt D, Ideker T, Kramer F

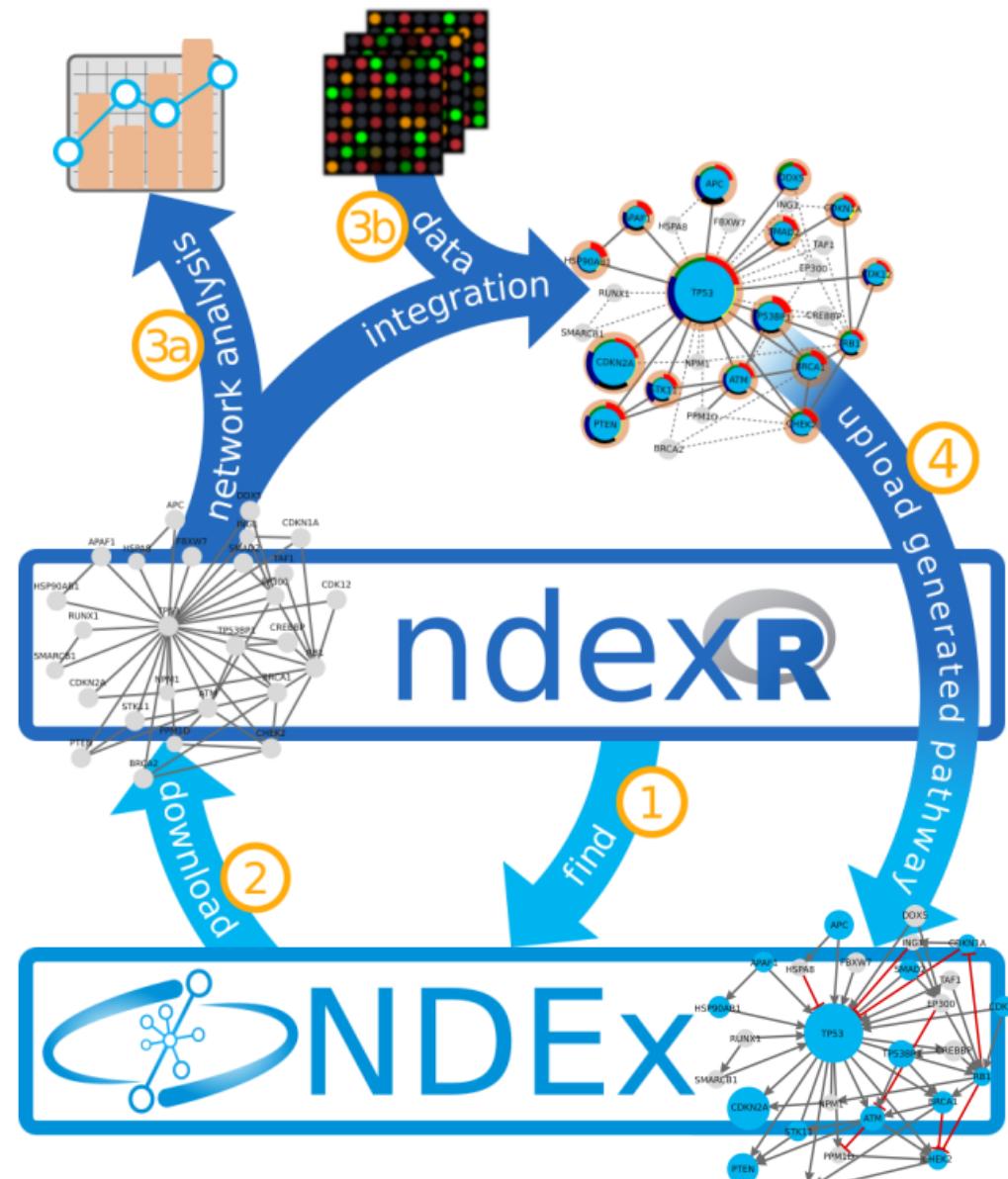
PMID: 29087446 DOI: [10.1093/bioinformatics/btx683](https://doi.org/10.1093/bioinformatics/btx683)

a software package that allows users to programmatically connect to
and interface with NDEx servers from within R

Available on [Bioconductor](#)  or from
[GitHub](#)  using [devtools](#)

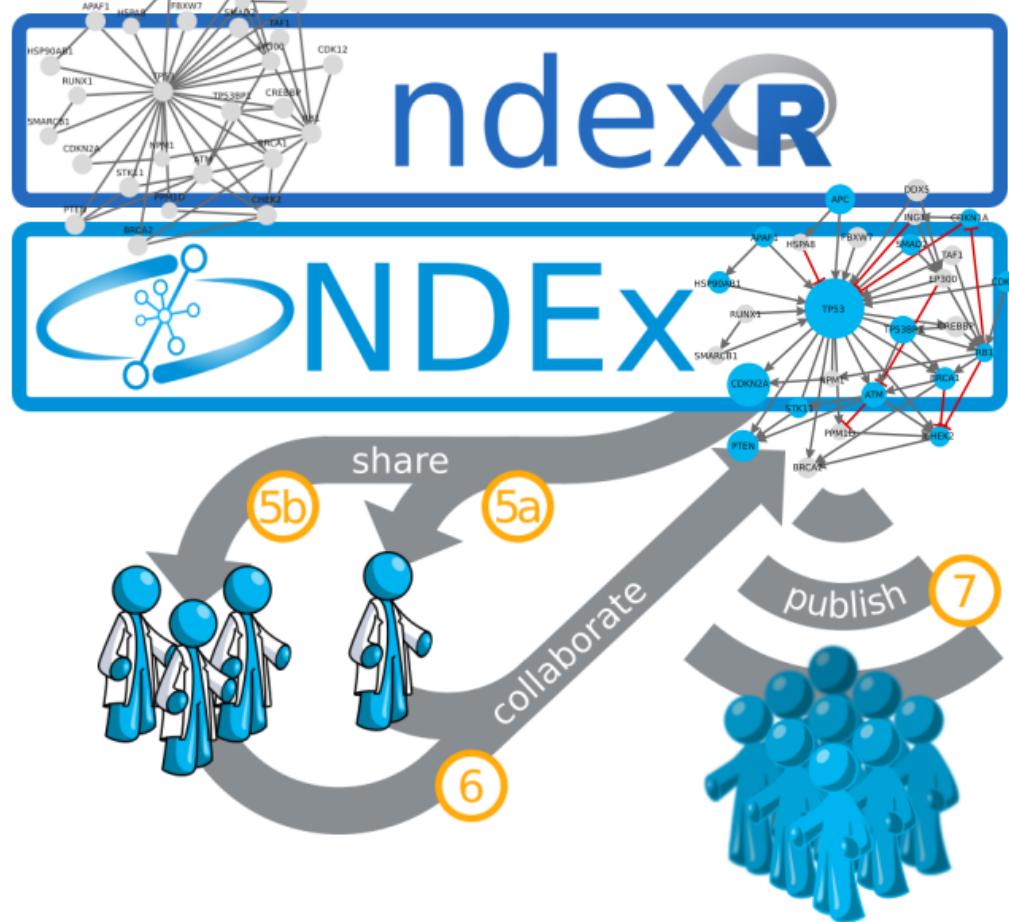
NDEx & NDExR

- 1 Find networks
- 2 Download networks
- 3 Work with the networks
- 4 Upload to NDEx server



NDEx & NDExR

- 5 Share networks
- 6 Collaborate on networks
- 7 Publish networks



Search and download networks

```
## connect to the NDEX server
ndexcon = ndex_connect()

## search the networks
networks <- ndex_find_networks(ndexcon,
  accountName = "nci-pid",
  searchString = "PI3K Akt signaling")

## download network with name matching the query
networkId = networks$externalId[1]
network = ndex_get_network(ndexcon, networkId)
```

Network as RCX

RCX object containing the following aspects:

	name	consistencyGroup	elementCount	idCounter	
1	ndexStatus	1	1	NA	1
2	provenanceHistory	1	1	NA	1
3	@context	1	1	NA	
4	nodes	1	37	37	
5	edges	1	113	207	
6	networkAttributes	1	8	NA	
7	nodeAttributes	1	73	NA	
8	edgeAttributes	1	168	NA	
9	citations	1	38	38	
10	cartesianLayout	1	37	NA	
11	visualProperties	1	3	NA	
	</na></na>				

Loading Gene Regulation

"Of the 304 genes evaluated, 254 (83.6%) were statistically significantly dysregulated after adjustment for multiple comparisons."

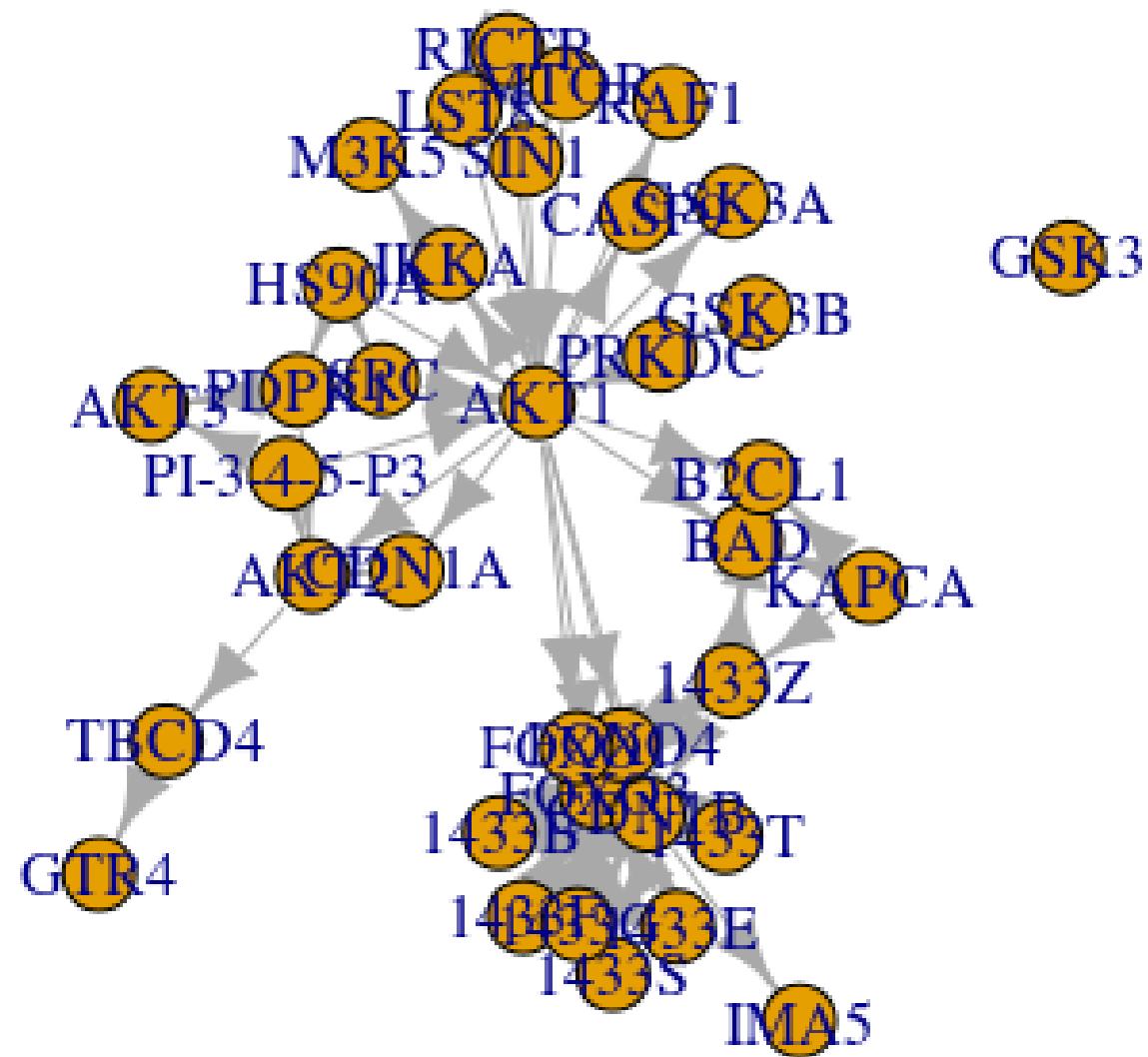
```
## Read from CSV file
geneRegulation = read.csv("PMC5760356.csv")

## filter for significant genes
geneRegulation = geneRegulation[
  geneRegulation$AdjustedPValue<=0.05, ]
```

	Gene <chr>	TumorMean <dbl>	NormalMean <dbl>	FoldChange <dbl>	PValue <dbl>	AdjustedPValue <dbl>
1	BAD	13.13	18.42	0.71	1.38e-11	3.42e-11
2	ITGA3	129.43	172.36	0.75	4.02e-12	1.03e-11
3	ITGA2B	0.54	0.86	0.62	9.90e-03	1.26e-02
4	BRCA1	87.46	35.35	2.47	2.77e-41	3.83e-40
5	IGF1	13.85	38.80	0.36	5.43e-28	3.59e-27

Visualization with iGraph

```
## using ndexr to convert an RCX to an igraph object  
graph = rcx_toNGraph(network)  
plot(graph)
```



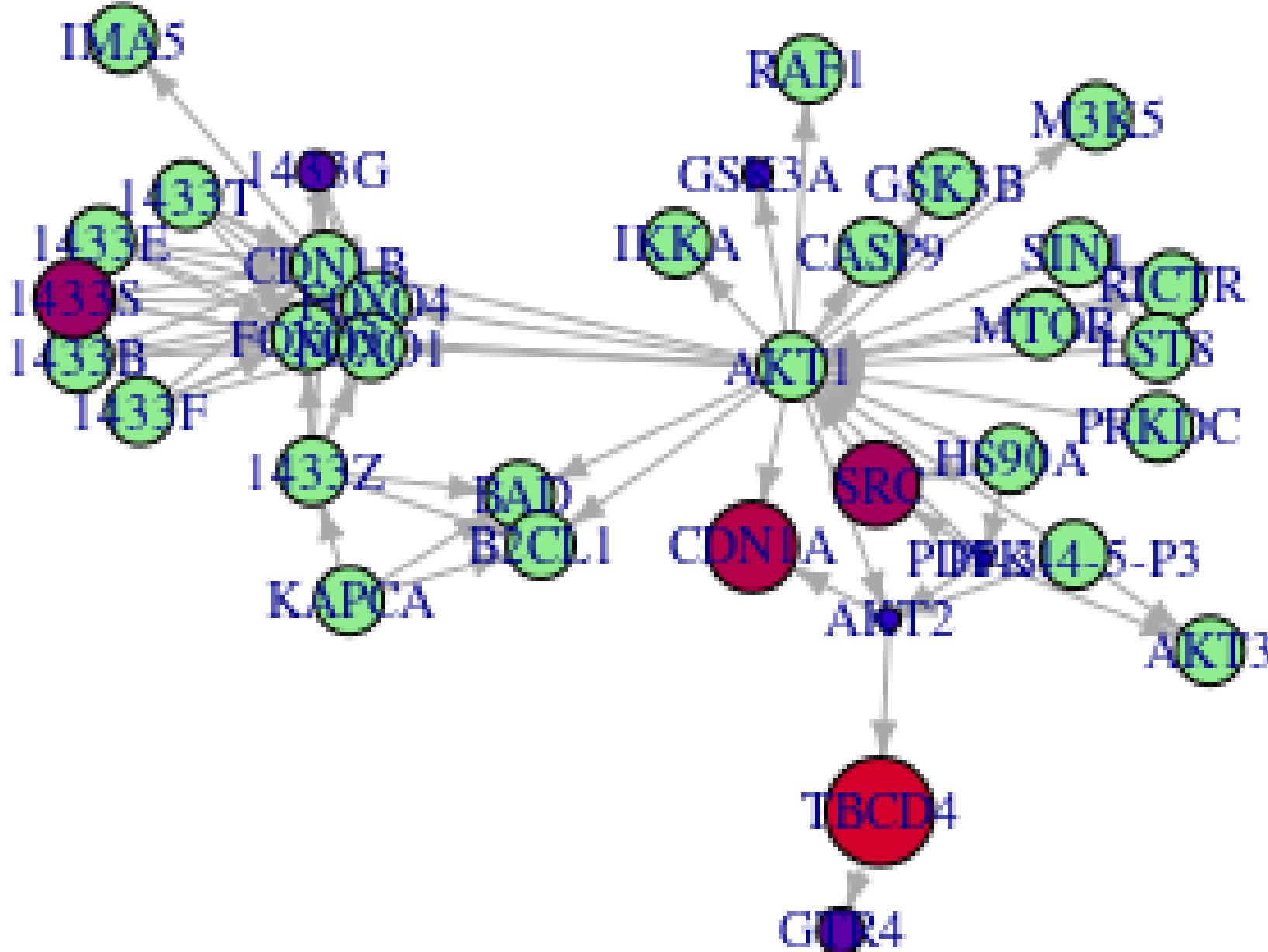
```
## map gene expression to vertices
nodeSize = merge(data.frame(name=V(graph)$name),
                  geneRegulation[,c("Gene", "FoldChange")],
                  by.x = "name", by.y = "Gene", all.x = T)

## assign color to vertices
nodeSize$color = ifelse(is.na(nodeSize$FoldChange), 0,
                        as.integer((nodeSize$FoldChange^3)*120))
V(graph)$color = ifelse(is.na(nodeSize$FoldChange), "lightgreen",
                        rgb(nodeSize$color, 0, 255-nodeSize$color,
                            maxColorValue = 255))

## adjust edge weight for force directed layout
E(graph)$weight = ifelse(E(graph)$i=="in-complex-with", 1, 2)

## visualize the igraph!
plot(graph,
      vertex.size = ifelse(is.na(nodeSize$FoldChange), 12,
                           as.integer(nodeSize$FoldChange^3*11)),
      vertex.label.cex = .7,
      edge.arrow.size=.4)
```

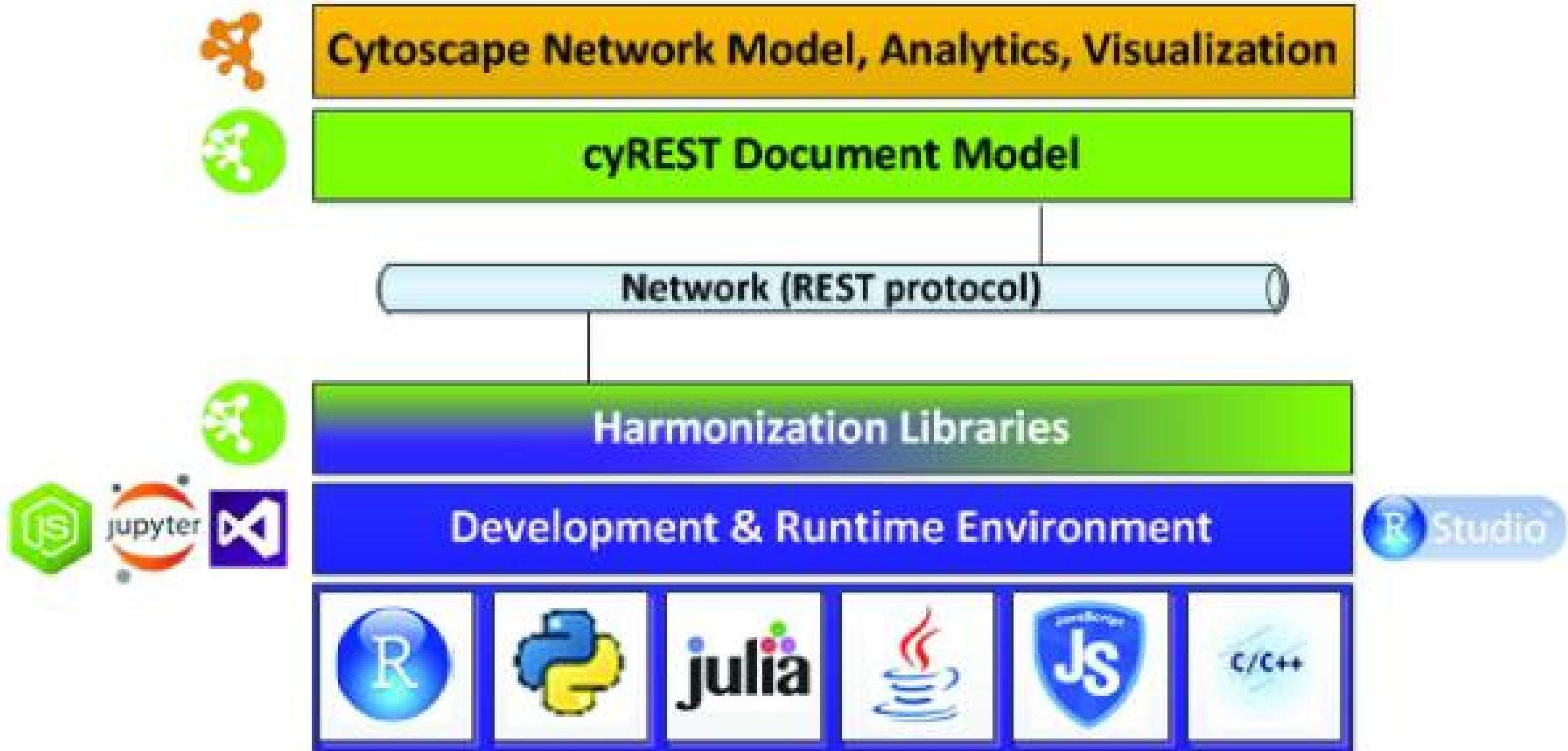
Looks better..

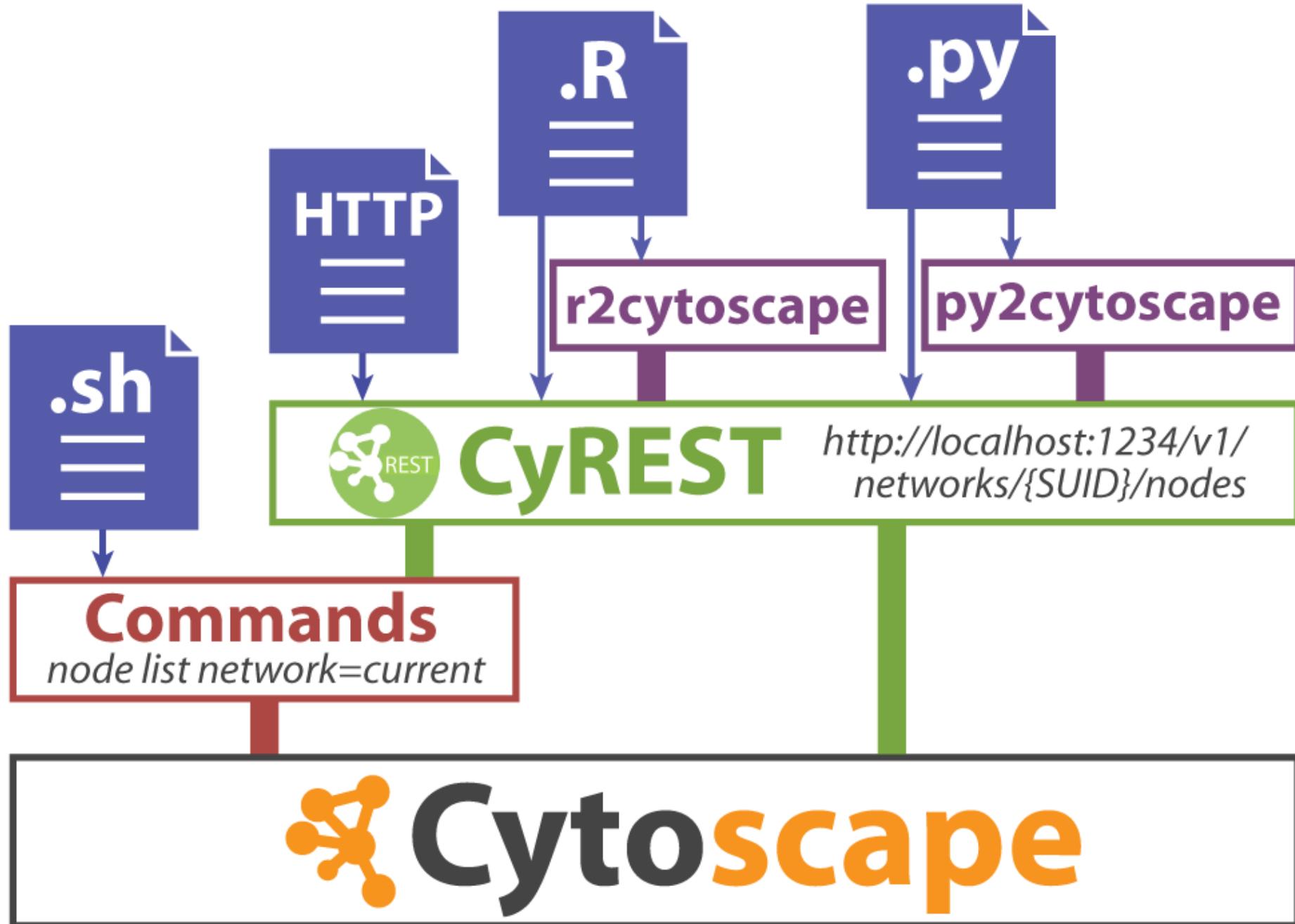


Cytoscape Script

```
#Import network
network import file indexColumnTargetInteraction=1 ...
    indexColumnSourceInteraction=2 file="input.xgmml"
#Import and set style
vizmap load file file="style.xml"
vizmap apply styles=style-name
#Set layout
layout attribute-circle
#Set view to fit display
view fit content
#Save
view export outputFile="output-name" options=PDF
```

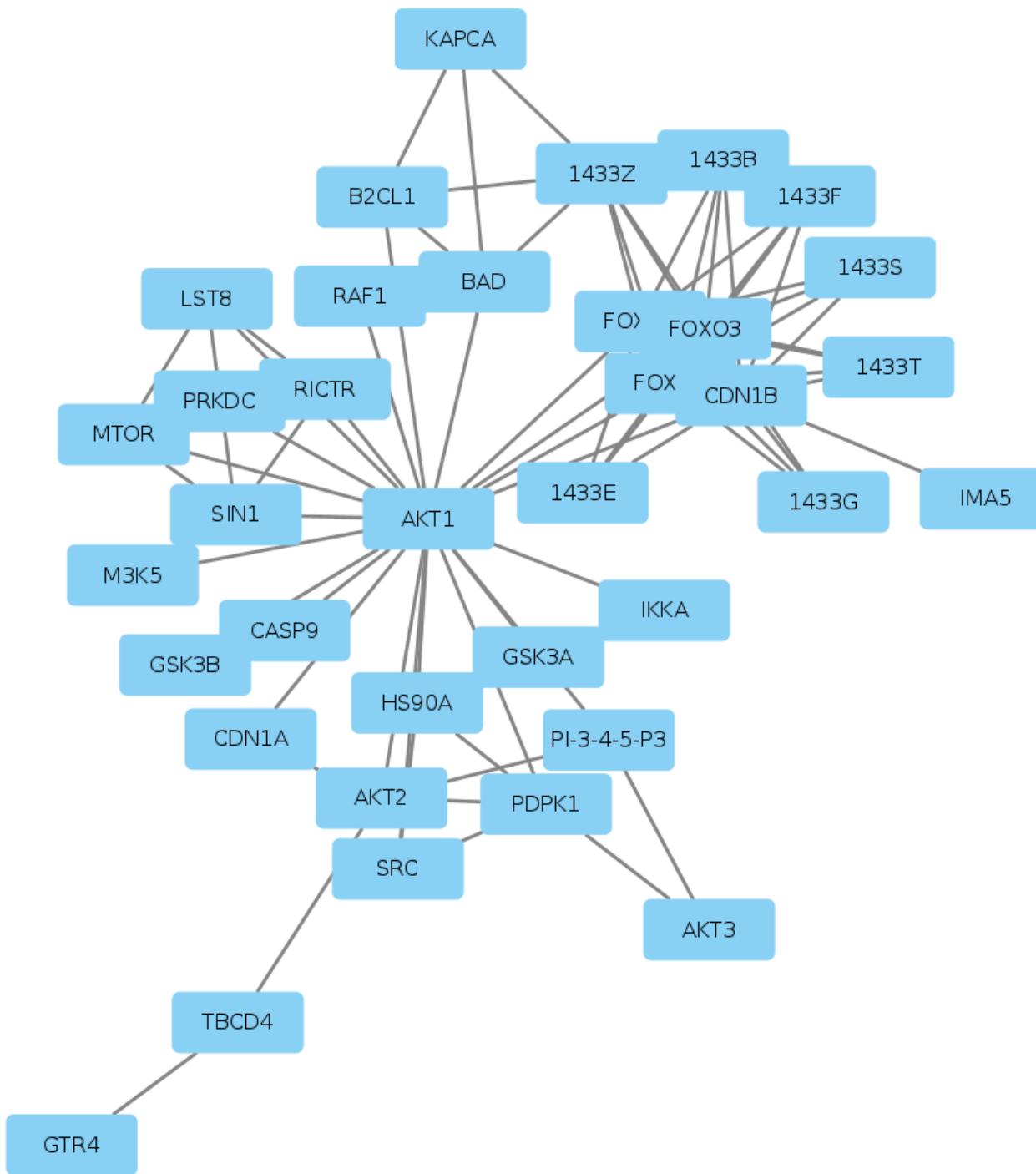
cyREST App => Core





Visualization with Cytoscape

Plain network in Cytoscape



Import Gene Table

```
## add the gene regulation as table
loadTableData(geneRegulation,
              data.key.column = "Gene",
              table = "node",
              table.key.column = "name")
```

Defining default Style

```
style.name = "FancyStyle"

defaults <- list(NODE_SHAPE="ROUND_RECTANGLE",
                 NODE_FILL_COLOR="lightgreen",
                 NODE_SIZE=30,
                 NODE_LABEL_FONT_SIZE=8,
                 EDGE_WIDTH=4,
                 EDGE_TARGET_ARROW_UNSELECTED_PAINT="#404040",
                 EDGE_TARGET_ARROW_SHAPE="delta")

## node-table:name is used as default node label
nodeLabels <- mapVisualProperty('node label', 'name', 'p')
```

Adding a custom Node Style

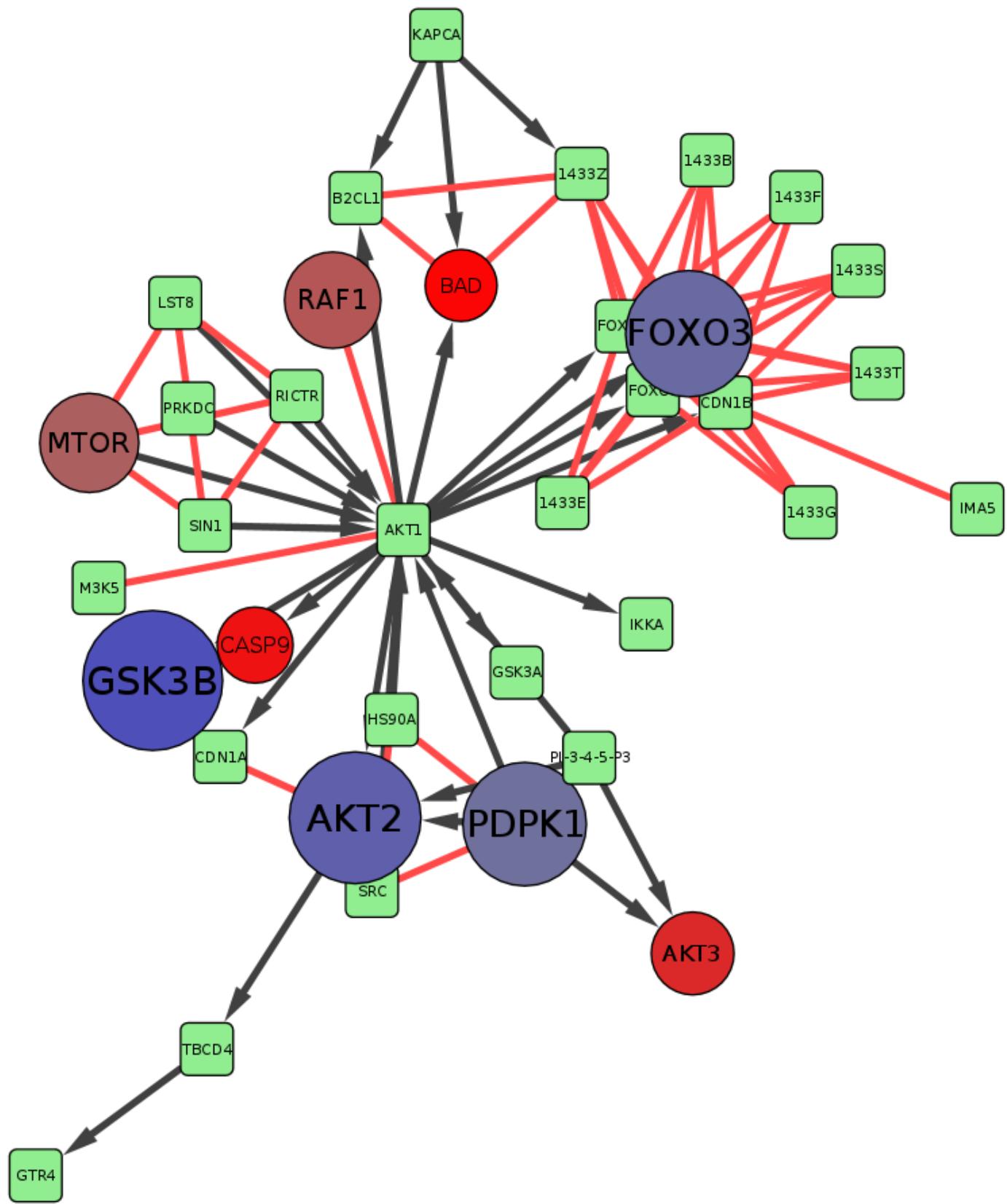
```
nodeShape <- mapVisualProperty('node shape',
  'group', 'd', c("changed"), c('ellipse'))  
  
## node size and color, and label size is based on fold change  
nodeSize <- mapVisualProperty('node size',
  'FoldChange', 'c', c(0.7,1.5), c(40,100))  
nodeFills <- mapVisualProperty('node fill color',
  'FoldChange', 'c', c(0.7,1.0,1.5),
  c('#FF0000', '#888888', '#0000FF'))  
nodeLabelFontSize <- mapVisualProperty('node label font size',
  'FoldChange', 'c', c(0.7,1.5), c(11,29))
```

...and Edge Style

```
edgeArrowShape <- mapVisualProperty('edge target arrow shape',  
    'interaction', 'd', c("in-complex-with"), c('none'))  
edgeColor <- mapVisualProperty('edge color',  
    'interaction', 'd', c("in-complex-with"), c('#FF4949'))
```

Create and apply the custom Style

```
createStyle(style.name, defaults, list(nodeLabels,  
    nodeSize,  
    nodeFills,  
    nodeLabelFontSize,  
    nodeShape,  
    edgeArrowShape,  
    edgeColor))  
  
applyStyle(style.name)
```



Export network

```
## export as png  
exportImage("cytoscape.png", "png")
```

Send Network to NDEx

```
## get the new network
newNetwork <- createIgraphFromNetwork("FancyNetwork")

## convert to RCX
newRCX = ngraph_fromRCX(newNetwork)

## create a new network on NDEx
ndex_create_network(ndexcon, newRCX, )
```

CyREST API

 **swagger**

<http://localhost:1234/v1/commands/swagger.json>

[Explore](#)

CyREST Command API

An API to offer access to Cytoscape command line commands through a REST-like service.

Cytoscape

<http://cytoscape.org/>

command

Show/Hide | List Operations | Expand Operations

Commands

Show/Hide | List Operations | Expand Operations

cybrowser

Show/Hide | List Operations | Expand Operations

diffusion

Show/Hide | List Operations | Expand Operations

edge

Show/Hide | List Operations | Expand Operations

group

Show/Hide | List Operations | Expand Operations

layout

Show/Hide | List Operations | Expand Operations

network

Show/Hide | List Operations | Expand Operations

Layouts

- Grid
- Hierarchical
- Circular (Radial)
- Partitions
- Force-Directed
- Combining layouts
- Multi-layer layouts
- Many, many others

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.

Further Reading

- Tutorials: cytoscape.github.io/cytoscape-tutorials
- Docker with Cytoscape and RStudio:
hub.docker.com/r/flockowak/cytoscape/
- Slides/RMarkdown florian-j-auer.github.io/NDExR-and-Cytoscape

