



bioinformatics.ca

Canadian Bioinformatics Workshops

www.bioinformatics.ca

This page is available in the following languages:

Afrikaans Български Català Dansk Deutsch Ελληνικά English English (CA) English (GB) English (US) Esperanto
Castellano Castellano (AR) Español (CL) Castellano (CO) Español (Ecuador) Castellano (MX) Castellano (PE)
Euskara Suomiksi français français (CA) Galego ජ්‍යෙනිස් hrvatski Magyar Italiano 日本語 한국어 Macedonian Melayu
Nederlands Norsk Sesotho sa Leboa polski Português română slovenský jezik srpski (latinica) Sotho svenska
中文 華語 (台灣) isiZulu

 creative commons

Attribution-Share Alike 2.5 Canada

You are free:

 to Share — to copy, distribute and transmit the work

 to Remix — to adapt the work





Under the following conditions:

 **Attribution.** You must attribute the work in the manner specified by the author or licensor (but not in any way that suggests that they endorse you or your use of the work).

 **Share Alike.** If you alter, transform, or build upon this work, you may distribute the resulting work only under the same or similar licence to this one.

- For any reuse or distribution, you must make clear to others the licence terms of this work.
- Any of the above conditions can be waived if you get permission from the copyright holder.
- The author's moral rights are retained in this licence.

[Disclaimer](#)

Your fair dealing and other rights are in no way affected by the above.
This is a human-readable summary of the Legal Code (the full licence) available in the following languages:
[English](#) [French](#)

Module 5

Gene Function Prediction

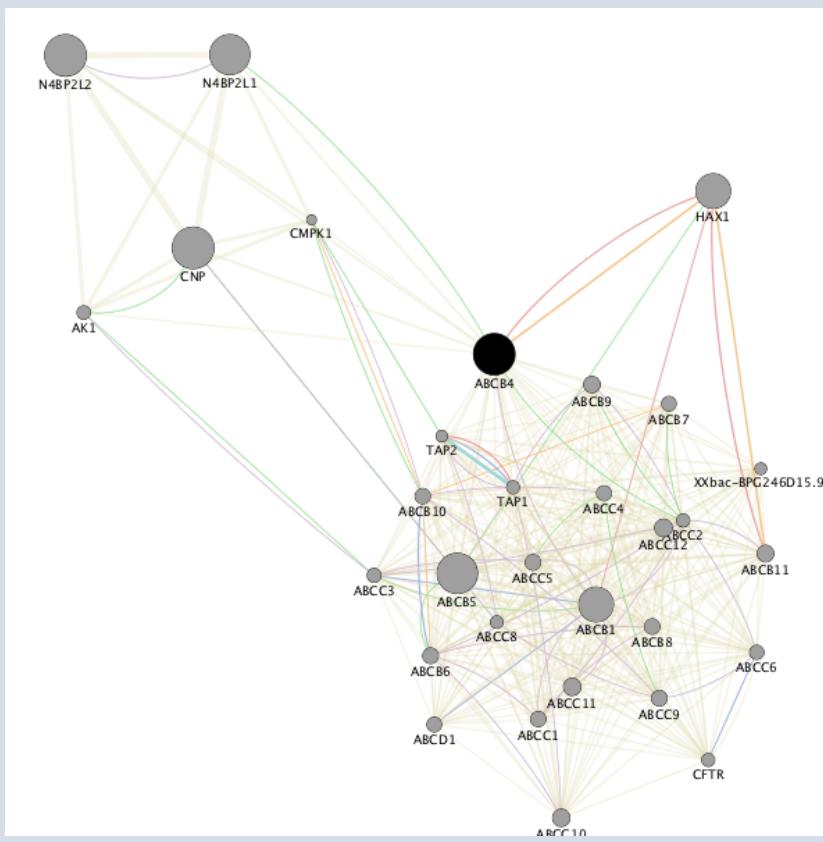
Quaid Morris

Pathway and network analysis of –omic data

June 1-3, 2015



GNAQ
GNAS
DGKZ
GUCY1A3
PDE4B
PDE4D
ATP2A2
ATP2A3
NOS1
CNN1
GSTO1
NOS3
CNN2
MYLK2
CALD1
ACTA1
MYL2



Learning Objectives of Module 5

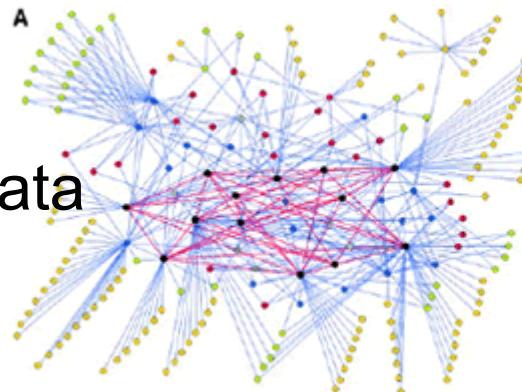
- **Understand** the concepts: *functional interaction network*, *guilt-by-association*, *gene recommender systems*.
- **Understand** the concept of context-specific network weighting schemes.
- **Understand** the difference between *direct interaction* and *label propagation* methods for predicting gene function.
- **Be able** to use gene recommender systems (e.g. GeneMANIA) to answer two types of questions about gene function: “what does my gene do?” and “give me more genes like these”
- **Be able** to select the appropriate network weighting scheme to answer your questions about gene function.

Outline

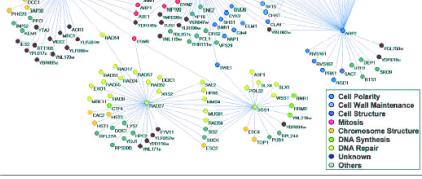
- Functional interaction networks
- Concepts in gene function prediction:
 - Guilt-by-association
 - Gene recommender systems
- Scoring interactions by guilt-by-association
- GeneMANIA
- GeneMANIA demo
- Explanation of network weighting schemes
- STRING

Using genome-wide data in the lab

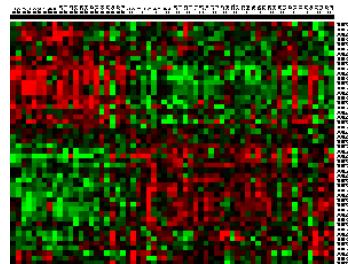
Protein domain similarity network



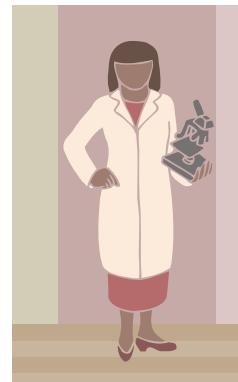
Genetic interaction data



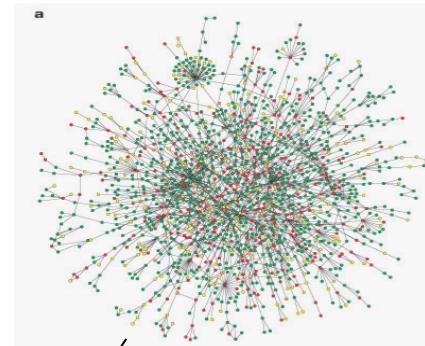
Microarray expression data



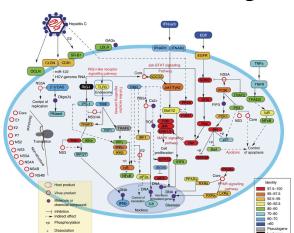
?!



Protein-protein interaction data

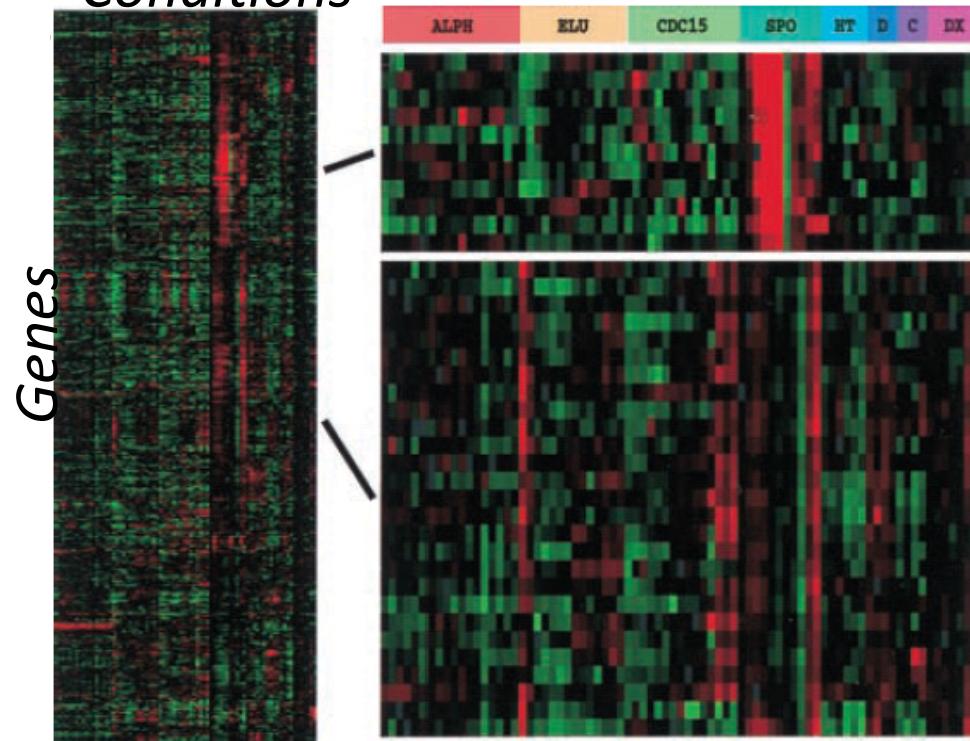


Pathways



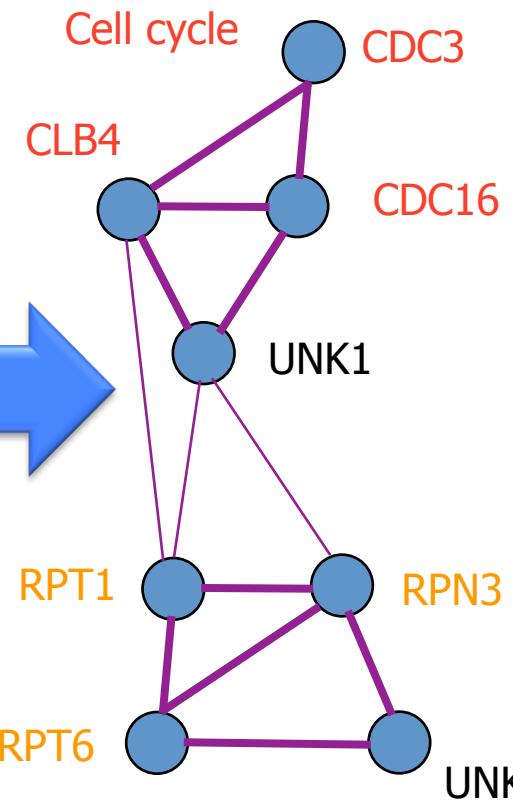
Functional interaction networks

Microarray expression data
Conditions



Eisen et al (PNAS 1998)

Co-expression network



B	STU2	CYTOSKELETON
	DHS1	DNA REPAIR
	BRM1	CYTOSKELETON
	SPC42	CYTOSKELETON
	CIN67	CYTOSKELETON
	CLB4	CELL CYCLE
	CDC10	CYTOKINESIS
	CDC3	CYTOKINESIS
	CLB3	CELL CYCLE
	APC4	CELL CYCLE
	CDC16	CELL CYCLE
C	RPN11	PROTEIN DEGRADATION
	UPF1	PROTEIN DEGRADATION
	RPM9	PROTEIN DEGRADATION
	RPT1	PROTEIN DEGRADATION
	RPM6	PROTEIN DEGRADATION
	PFE4	PROTEIN DEGRADATION
	RPM6	PROTEIN DEGRADATION
	RPT4	PROTEIN DEGRADATION
	RPT7	PROTEIN DEGRADATION
	RPM3	PROTEIN DEGRADATION
	FUP2	PROTEIN DEGRADATION
	SCL1	PROTEIN DEGRADATION
	PFE5	PROTEIN DEGRADATION
	PFE9	PROTEIN DEGRADATION
	PFE1	PROTEIN DEGRADATION
	PFE2	PROTEIN DEGRADATION
	PFE3	PROTEIN DEGRADATION
	PFE10	PROTEIN DEGRADATION
	FUP1	PROTEIN DEGRADATION
	PFE6	PROTEIN DEGRADATION
	PFE7	PROTEIN DEGRADATION
	RPM10	PROTEIN DEGRADATION
	RPT3	PROTEIN DEGRADATION
	RPT5	PROTEIN DEGRADATION
	RPM12	PROTEIN DEGRADATION
	RPM5	PROTEIN DEGRADATION
	RPM8	PROTEIN DEGRADATION

Protein degradation

A useful reference: Fraser AG, Marcotte EM - A probabilistic view of gene function - Nat Genet. 2004 Jun;36(6):559-64

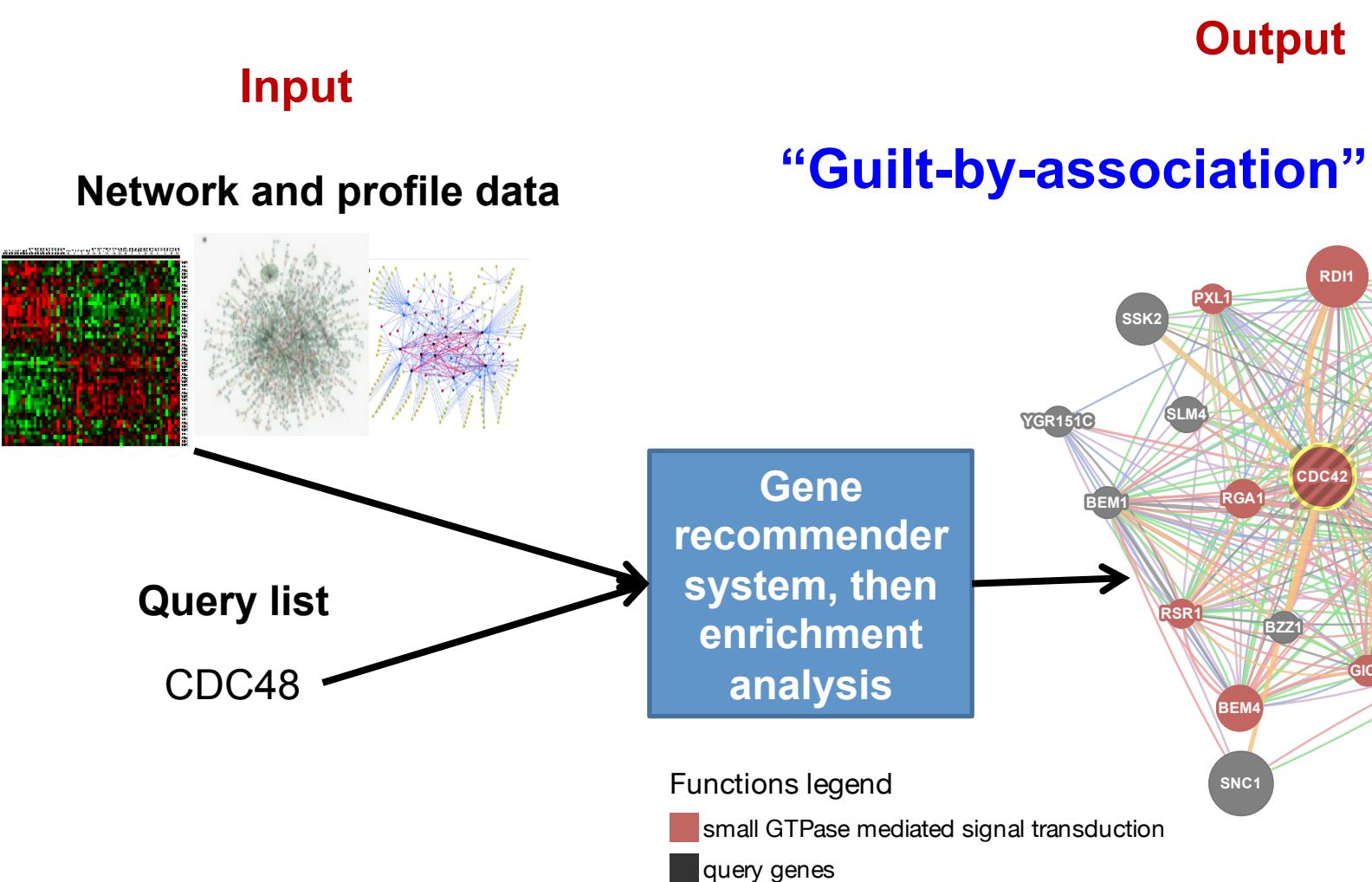
Varieties of functional interaction networks

- **Directly measured interactions, e.g.:**
 - protein interaction networks
 - genetic interaction networks
- **Inferred interactions from a single data source, e.g.:**
 - co-expression networks computed from gene expression profiling studies (whether microarray, RNA-seq, proteomics)
- **Inferred interactions from multiple data sources, e.g.:**
 - *Context-independent*: FI network, STRING, {Human, Worm, etc} Net, bioPIXIE
 - *Context-dependent*: GeneMANIA, HEFaIMp

Two types of function prediction

- “**What does my gene do?**”
 - Goal: determine a gene’s function based on who it interacts with: “guilt-by-association”
- “**Give me more genes like these**”
 - e.g. find more genes in the Wnt signaling pathway, find more kinases, find more members of a protein complex

“What does my gene do?”



Network types used

- **Directly measured interactions, e.g.:**
 - protein interaction networks
 - genetic interaction networks
- **Inferred interactions from a single data source, e.g.:**
 - co-expression networks computed from gene expression profiling studies (whether microarray, RNA-seq, proteomics)
- **Inferred interactions from multiple data sources, e.g.:**
 - *Context-independent*: FI network, STRING, {Human, Worm, etc} Net, bioPIXIE
 - *Context-dependent*: GeneMANIA, HEFaMp

“What does p53 do?”

- Question could be about its
 - biological process,
 - biochemical/molecular function,
 - subcellular/Cellular localization,
 - regulatory targets,
 - temporal expression pattern,
 - phenotypic effect of deletion,
 - role in disease.

Some networks may be better for some types of gene function than others

Network types needed

- **Directly measured interactions, e.g.:**
 - protein interaction networks
 - genetic interaction networks
- **Inferred interactions from a single data source, e.g.:**
 - co-expression networks computed from gene expression profiling studies (whether microarray, RNA-seq, proteomics)
- **Inferred interactions from multiple data sources, e.g.:**
 - *Context-independent*: FI network, STRING, {Human, Worm, etc} Net, bioPIXIE
 - *Context-dependent*: GeneMANIA, HEFalMp

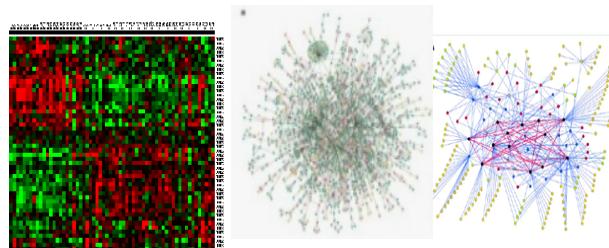
Defining queries by providing context

- **Memphis**, Knoxville, Nashville...
 - Chattanooga, Morristown
- **Memphis**, Alexandria, Cairo...
 - Luxor, Giza, Aswan

“Give me more genes like these”

Input

Network and profile data



Query list

GNAQ
GNAS
DGKZ
GUCY1A3
PDE4B
PDE4D
ATP2A2
ATP2A3
NOS1
CNN1
GSTO1
NOS3
CNN2
MYLK2
CALD1
ACTA1
MYL2

Gene
recommender
system

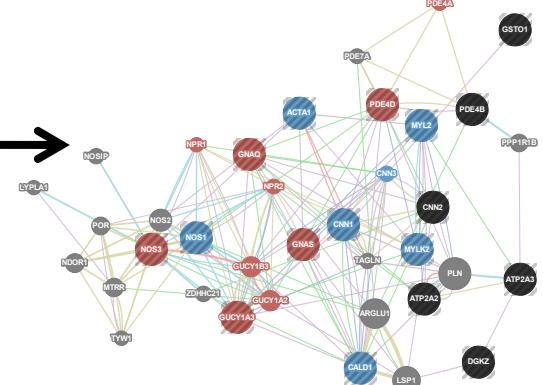
Output

Functions legend

- muscle contraction
- cyclic nucleotide metabolic process
- query genes

Networks legend

- Co-expression
- Co-localization
- Genetic interactions
- Pathway
- Physical interactions
- Shared protein domains



GENEMANIA

Find genes in **H. sapiens (human)**

related to **gnaq; gnas; dgkz; gucy1a3; pde4b; pde4d; atp2a2; atp**

Go

Showing 20 related genes with 37 total genes, 10 attributes, and 239 total links

Show advanced options

File ▾ View ▾ Query ▾

Networks legend

Functions legend

Networks Genes Functions

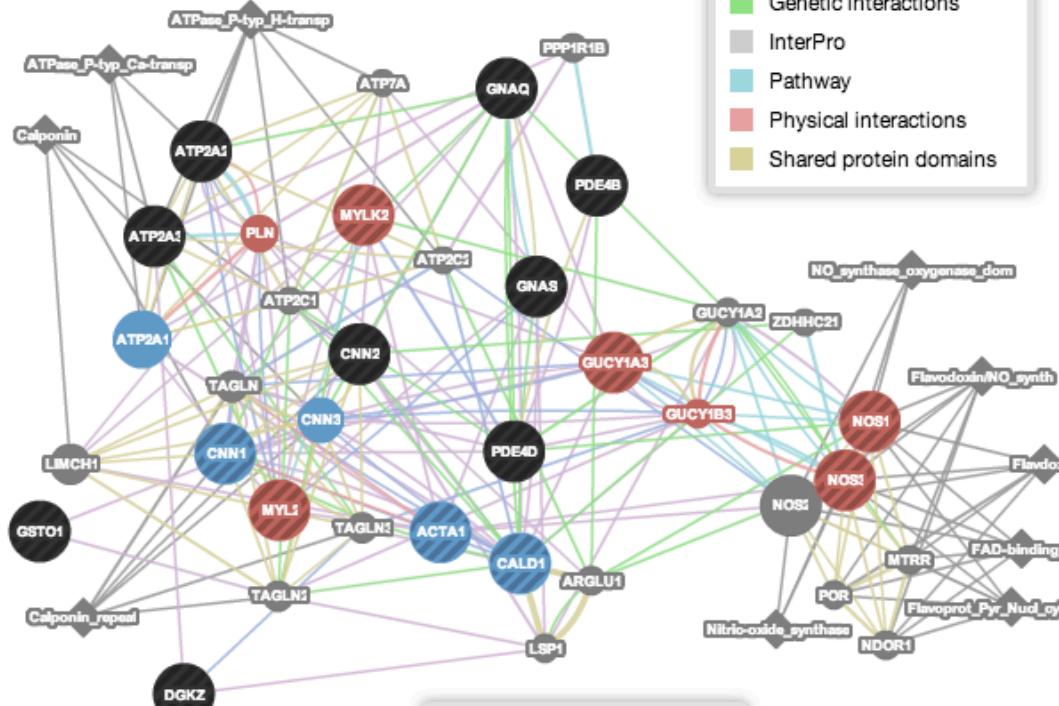


Sort by: [name, percent weight](#)

Expand: [all](#), [only top level](#), [none](#)

Enable: [all](#), [none](#)

<input checked="" type="checkbox"/> InterPro	66.96 %
<input checked="" type="checkbox"/> Co-expression	20.60 %
<input checked="" type="checkbox"/> Pathway	5.92 %
<input checked="" type="checkbox"/> Shared protein domains	4.33 %
<input checked="" type="checkbox"/> Co-localization	1.79 %
<input checked="" type="checkbox"/> Physical interactions	0.22 %
<input checked="" type="checkbox"/> Genetic interactions	0.17 %



Functions legend

- blood circulation
- muscle system process
- query genes



GeneMANIA: Selecting networks I

Click links to select all, zero or a pre-defined (default) set of networks

The screenshot shows the GeneMANIA web interface. At the top, there is a search bar with 'Find genes in S. cerevisiae (baker's yeast)' and a dropdown menu 'related to CDC27; APC11; APC4; XRS2; RAD54; APC2; RAD'. Below the search bar is a 'Help' button and a navigation menu with 'Video tutorials', 'Blog', 'Contact us', and 'About'. A large green arrow points to the left side of the interface, which contains a list of network types. To the right of this list is another green arrow pointing to a 'Hide advanced options' button.

Networks
Enable: all, none, default (209 of 321 currently enabled)
Sort by: first author, last author, publication date, size

Network Type	Status
Attributes	0/1
Co-expression	20/129
Co-localization	1/1
Genetic interactions	63/63
Physical interactions	76/76
Predicted	44/46
Shared protein domains	2/2
Other	3/3
Uploaded	0/0

Upload help Upload network...

Network weighting

Query-dependent weighting
 Automatically selected weighting method
 Assigned based on query genes

Gene Ontology (GO)-based weighting
 Biological process based
 Molecular function based
 Cellular component based

Equal weighting
 Equal by network
 Equal by data type

Number of gene results
In the results generated by GeneMANIA, 20 related genes and at most 10 related attributes will be displayed.

Click phrase to open or close the advanced options panel

GeneMANIA: Selecting networks II

Click check boxes to select all (or no) networks or attributes of that type.

The screenshot shows the GeneMANIA web interface. At the top, there is a search bar with 'Find genes in S. cerevisiae (baker's yeast)' and 'related to CDC27; APC11; APC4; XRS2; RAD54; APC2; RAD'. Below the search bar is a 'Networks' section with a green header. The header contains the text 'Enable: all, none, default (209 of 321 currently enabled)' and 'Sort by: first author, last author, publication date, size'. It also has 'Upload help' and 'Upload network...' buttons. A large black arrow points to the left side of the 'Networks' section. Inside this section, there is a table of network types with their counts:

Network Type	Count
Attributes	0/1
Co-expression	20/129
Co-localization	1/1
Genetic interactions	63/63
Physical interactions	76/76
Predicted	44/46
Shared protein domains	2/2
Other	0/0
Uploaded	0/0

A small black arrow points upwards from the bottom of the table towards the fraction values. To the right of the table, there is a 'weighting' section with 'Equal weighting' selected, and options for 'Equal by network' and 'Equal by data type'. At the bottom of the 'Networks' section, there is a note: 'In the results generated by GeneMANIA, 20 related genes and at most 10 related attributes will be displayed.'

Fraction indicates # of networks selected out of total available (for this organism).

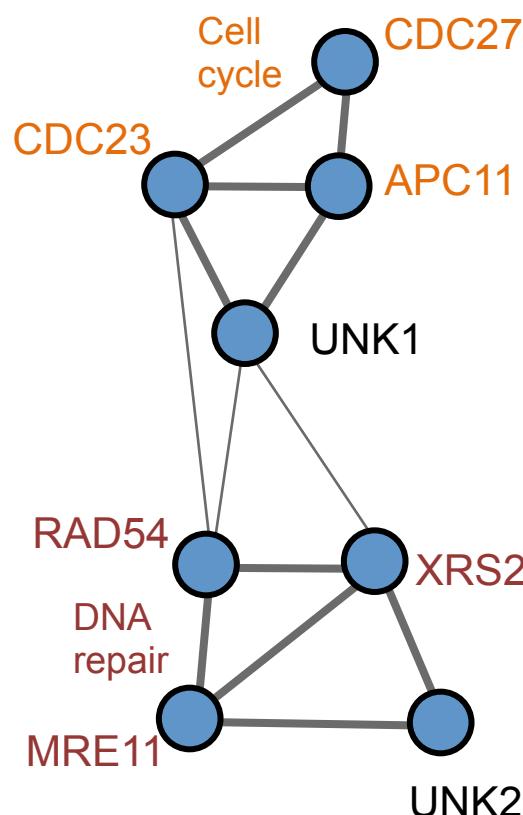
GeneMANIA: Selecting networks III

Click on network type to view list of networks (of that type) in right panel

The screenshot shows the GeneMANIA interface for selecting networks. At the top, there is a search bar with 'Find genes in S. cerevisiae (baker's yeast)' and a dropdown menu 'related to cdc27; cdc23; apc11; rad54; mre11; xrs2'. Below the search bar, it says 'Showing 20 related genes' and 'Hide advanced options'. A large section titled 'Networks' lists various network types with their counts: Co-expression (20/95), Co-localization (0/1), Genetic interactions (63/63), Physical interactions (74/74), Predicted (0/53), Shared protein domains (0/2), Other (0/3), and Uploaded (0/0). A yellow box highlights the 'Huh-O'Shea-2003' entry, which includes a link to 'Global analysis of protein localization in budding yeast' by Huh et al. (2003) in Nature, and a note that the source is 904,057 interactions from supplementary material. To the left of the network list, a black arrow points to the 'Co-localization' entry. To the right of the expanded network entry, another black arrow points to the 'Upload...' button. A callout box in the center-right area says 'Click on check box to select (or deselect) network'. At the bottom, there are sections for 'Network weighting' (with radio buttons for 'Automatically selected weighting method' and 'Assigned based on query genes'), 'Gene Ontology (GO)-based weighting' (with radio buttons for 'Biological process based', 'Molecular function based', and 'Cellular component based'), and 'Equal weighting' (with radio buttons for 'Equal by network' and 'Equal by data type').

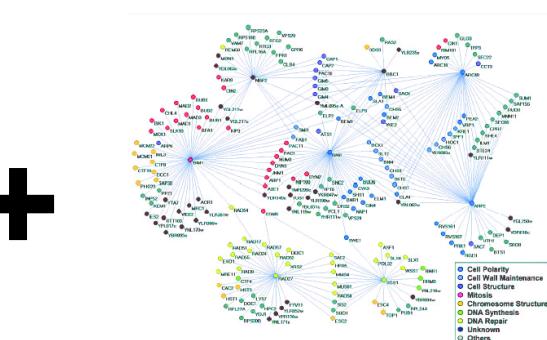
Click on network name to expand entry to get more information on network. HTML link points to Pubmed abstract

Context-independent networks



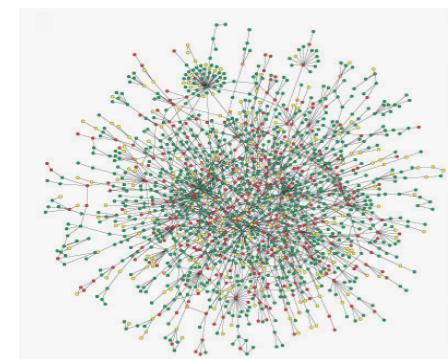
Co-expression

Pre-combine networks e.g. by simple addition or by pre-determined weights



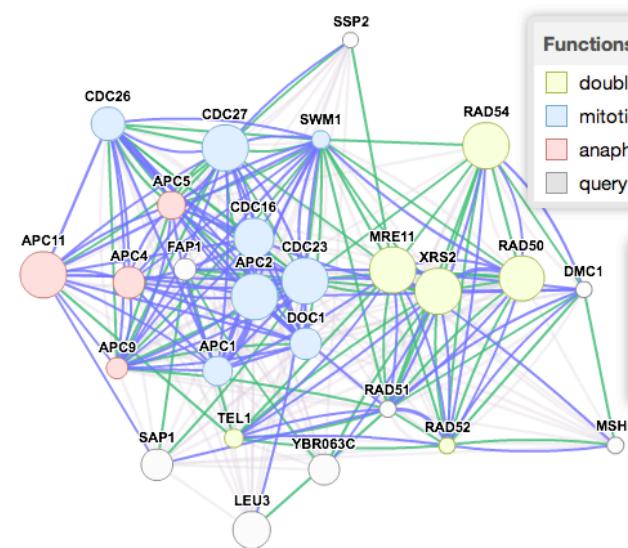
Genetic

e.g. Tong et al. 2001



Co-complexed

e.g. Jeong et al 2002



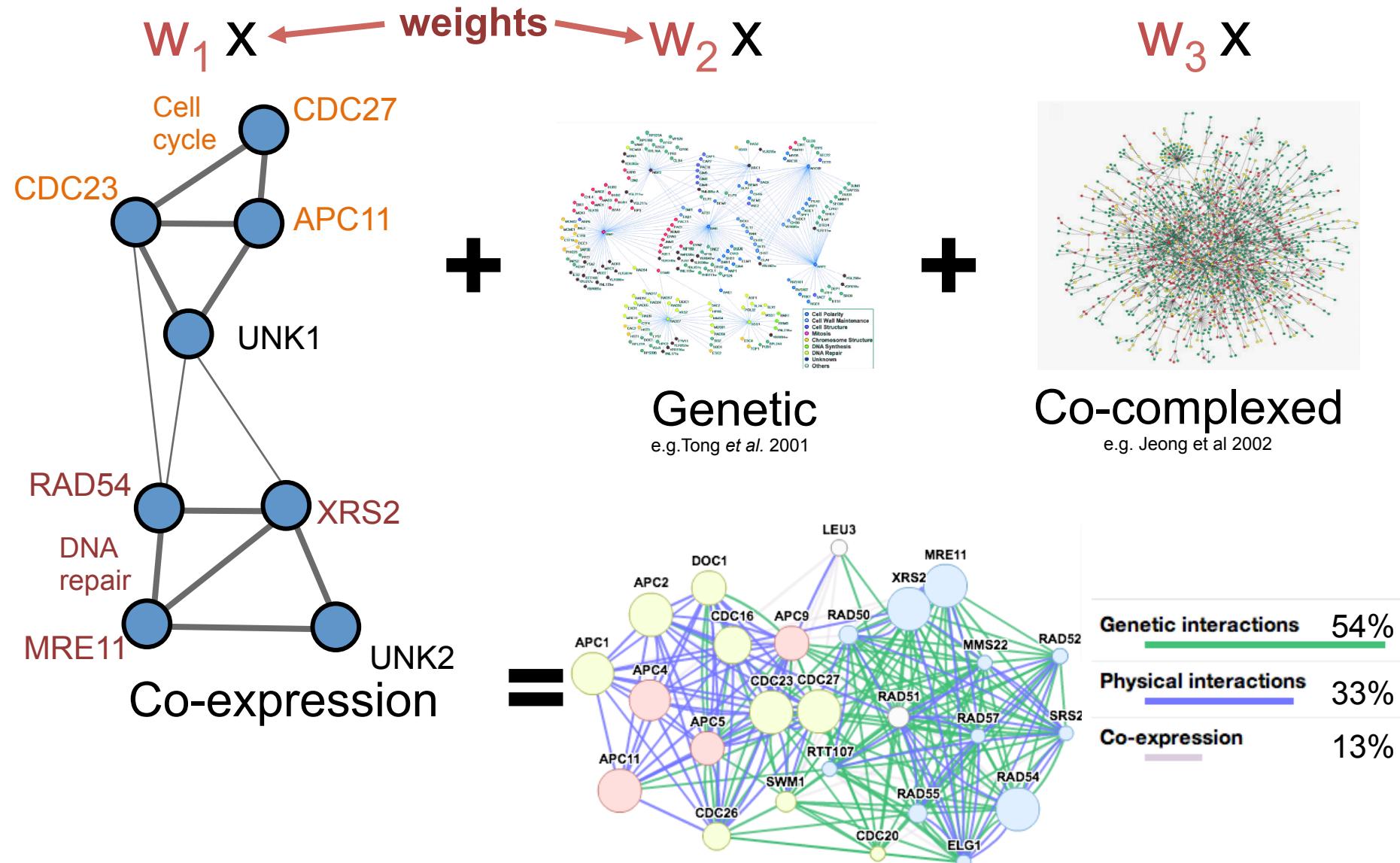
Functions legend

- double-strand break repair
- mitotic metaphase/anaphase transition
- anaphase-promoting complex
- query genes

Networks legend

- Co-expression
- Genetic interactions
- Physical interactions

Context-dependent networks



Two rules for network weighting

Relevance

The network should be relevant to predicting the function of interest

- **Test:** Are the genes in the query list more often connected to one another than to other genes?

Redundancy

The network should not be redundant with other datasets – particularly a problem for co-expression

- **Test:** Do the two networks share many interactions?
- *Caveat:* Shared interactions also provide more confidence that the interaction is real.

Network weighting schemes I

Click radio button
to change the
network weight
scheme

The screenshot shows the GeneMANIA web interface. At the top, there are search fields: "Find genes in S. cerevisiae (baker's yeast)" and "related to cdc27; cdc23; apc11; rad54; mre11; xrs2". Below the search fields, it says "Showing 20 related genes" and "Hide advanced options".

In the "Networks" section, there is a list of network types with checkboxes: Co-, Co-, Ger, Phy, Pre, Sha, Oth, and Up. The "Ger" and "Phy" checkboxes are checked. Below this, a large text box contains the following message:

By default, GeneMANIA decides between GO-dependent and query-specific weighting scheme based on the size of your list. We recommend using the default scheme in most cases

Below the message, there is a list of publications:

- ▶ □ Brauer-Botstein-2008
- ▶ □ Busti-Vanoni-2011
- ▶ □ Caba-Aubrecht-2005
- ▶ □ Chattopadhyay-Tabor-2009

The "Network weighting" section is highlighted with a green background. It contains three radio button groups:

- Query-dependent weighting**:
 - Automatically selected weighting method
 - Assigned based on query genes
- Gene Ontology (GO)-based weighting**:
 - Biological process based
 - Molecular function based
 - Cellular component based
- Equal weighting**:
 - Equal by network
 - Equal by data type

At the bottom, the "Number of gene results" section states: "In the results generated by GeneMANIA, 20 related genes will be displayed."

Network weighting schemes II

The screenshot shows the GeneMANIA interface for yeast genes. At the top, there are search fields for 'Find genes in' (S. cerevisiae (baker's yeast)) and 'related to' (cdc27; cdc23; apc11; rad54; mre11; xrs2). Below this, a sidebar lists network types: Co-expression, Co-localization, Gene ontology, Phylogenetic, Preprint, Shared interaction, Other, and Upload. Under 'Network weighting', the 'Gene Ontology (GO)-based weighting' section is highlighted with a green box. It contains three radio button options: 'Automatically selected weighting method' (selected), 'Assigned based on query genes', and 'Equal weighting'. The 'Equal weighting' section has two options: 'Equal by network' and 'Equal by data type'. At the bottom, it says 'Number of gene results' and 'In the results generated by GeneMANIA, 20 related genes will be displayed.'

- GO-based weighting assigns network weights based on how well the networks reproduce patterns of GO co-annotations (“Are genes that interact in the network more likely to have the same annotation?”),
- Can choose any of the three hierarchies,
- Ignores query list when assigning network weight.

Network weighting schemes III

Can force query
list based
weighting by
selecting this
option

The screenshot shows the GeneMANIA interface for querying related genes in *S. cerevisiae* (baker's yeast) related to *cdc27; cdc23; apc11; rad54; mre11; xrs2*. It displays a list of networks and their counts, and a section for network weighting.

Networks:

Network Type	Count
Co-expression	20/95
Co-localization	0/1
Genetic interactions	63/63
Physical interactions	74/74
Predicted	0/53
Shared protein domains	0/2
Other	0/3
Uploaded	0/0

Network weighting:

Query-dependent weighting:

- Automatically selected weighting method
- Assigned based on query genes

Gene Ontology (GO)-based weighting:

- Biological process based
- Molecular function based
- Cellular component based

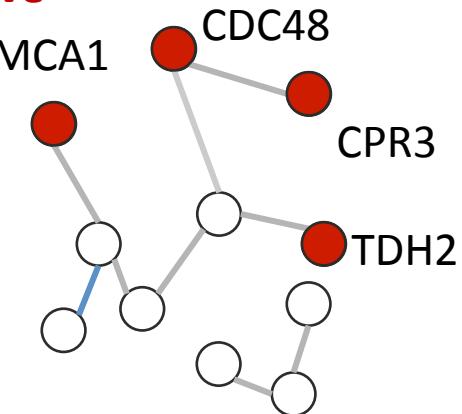
Equal weighting:

- Equal by network
- Equal by data type

Select these and either all networks or all data types get the same weight

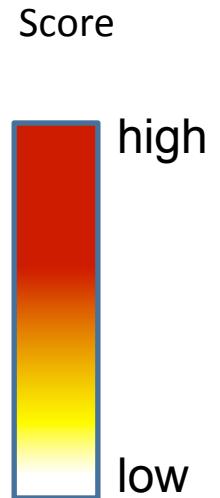
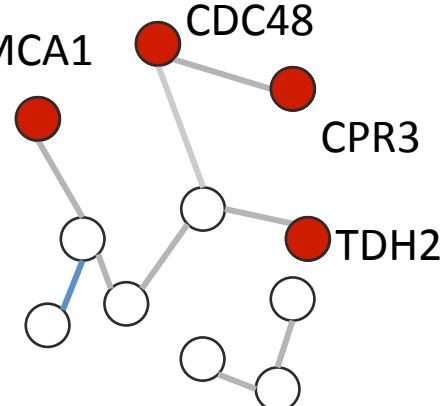
Predicting gene function by finding “guilty associates”

Query list: “positive examples”

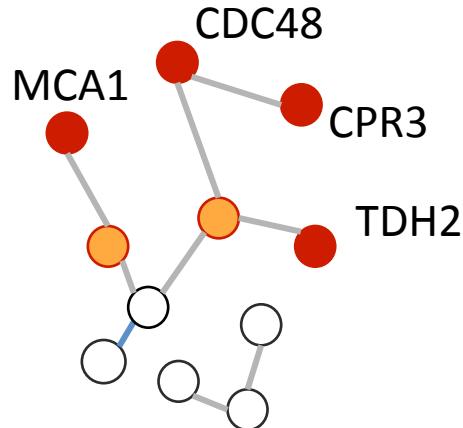


Predicting gene function by finding “guilty associates”

Query list: “positive examples”

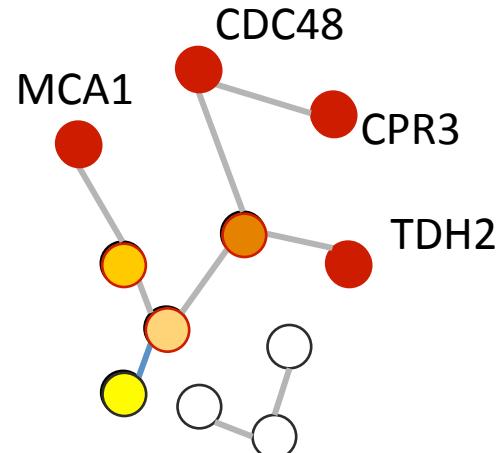


Direct interaction



Two main algorithms

Label propagation

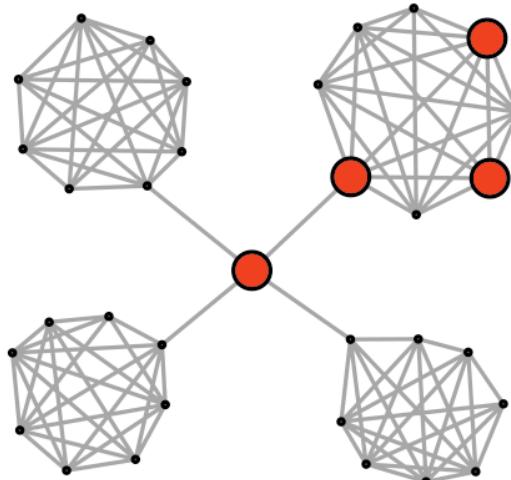


Association scoring algorithm details

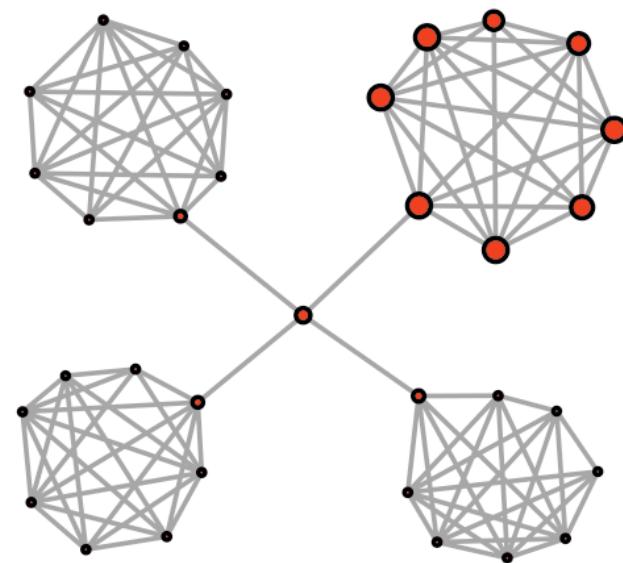
- **Direct interaction** scoring depends on:
 - Strength of links to query genes,
 - # of query gene neighbors,
 - Example algorithm: Naïve Bayes
- **Label propagation** scoring depends on:
 - Iteratively propagating ‘direct neighbour score’ allowing indirect links to impact scores,
 - Whether or not a gene is in a connected cluster of genes with query gene(s)
 - Example algorithm: GeneMANIA

Label propagation example

Before



After



Three parts of GeneMANIA:

- A large, automatically updated collection of interactions networks.
- A query algorithm to find genes and networks that are functionally associated to your query gene list.
- An interactive, client-side network browser with extensive link-outs

GeneMANIA data sources

Various sources, largely
mSigDB, compiled by Bader lab



Interologs

Network

Enable: none, default (140 of 400 currently enabled)
Sort by: first author, last author, publication date, size

Upload help Upload network...

Attributes	Count
<input checked="" type="checkbox"/> Attributes	5/5
<input checked="" type="checkbox"/> Co-expression	20/279
<input checked="" type="checkbox"/> Co-localization	2/2
<input checked="" type="checkbox"/> Genetic interactions	3/3
<input checked="" type="checkbox"/> Pathway	6/6
<input checked="" type="checkbox"/> Physical interactions	61/61
<input checked="" type="checkbox"/> Predicted	41/42
<input checked="" type="checkbox"/> Shared protein domains	2/2
<input type="checkbox"/> Uploaded	0/0

Consolidated-Pathways-2013

Source: Pathways consolidated by Bader lab. from [Enrichment map: a network-based method for gene-set enrichment visualization and interpretation.](#)

- Drug-interactions-2013
- InterPro
- miRNA-target-predictions-20
- Transcriptional-factor-targets



+ some organism-specific datasets
(click around to see what's available)

-Gene ID mappings from
Ensembl and Ensembl Plant

-Network/gene descriptors
from Entrez-Gene and
Pubmed

-Gene annotations from
Gene Ontology, GOA, and
model org. databases

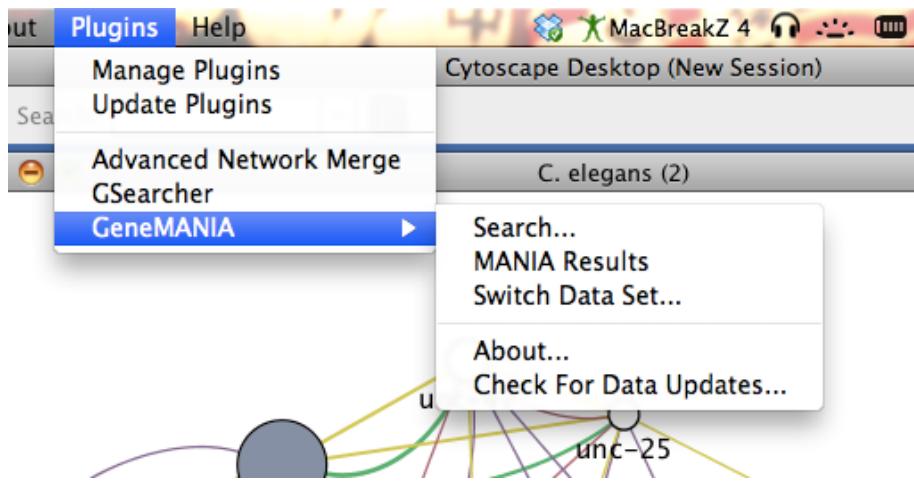
Gene identifiers

- All unique identifiers within the selected organism: e.g.
 - Entrez-Gene ID
 - Gene symbol
 - Ensembl ID
 - Uniprot (primary)
 - also, some synonyms & organism-specific names
- We use Ensembl database for gene mappings (but we mirror it once / 3 months, so sometimes we are out of date)

Current status

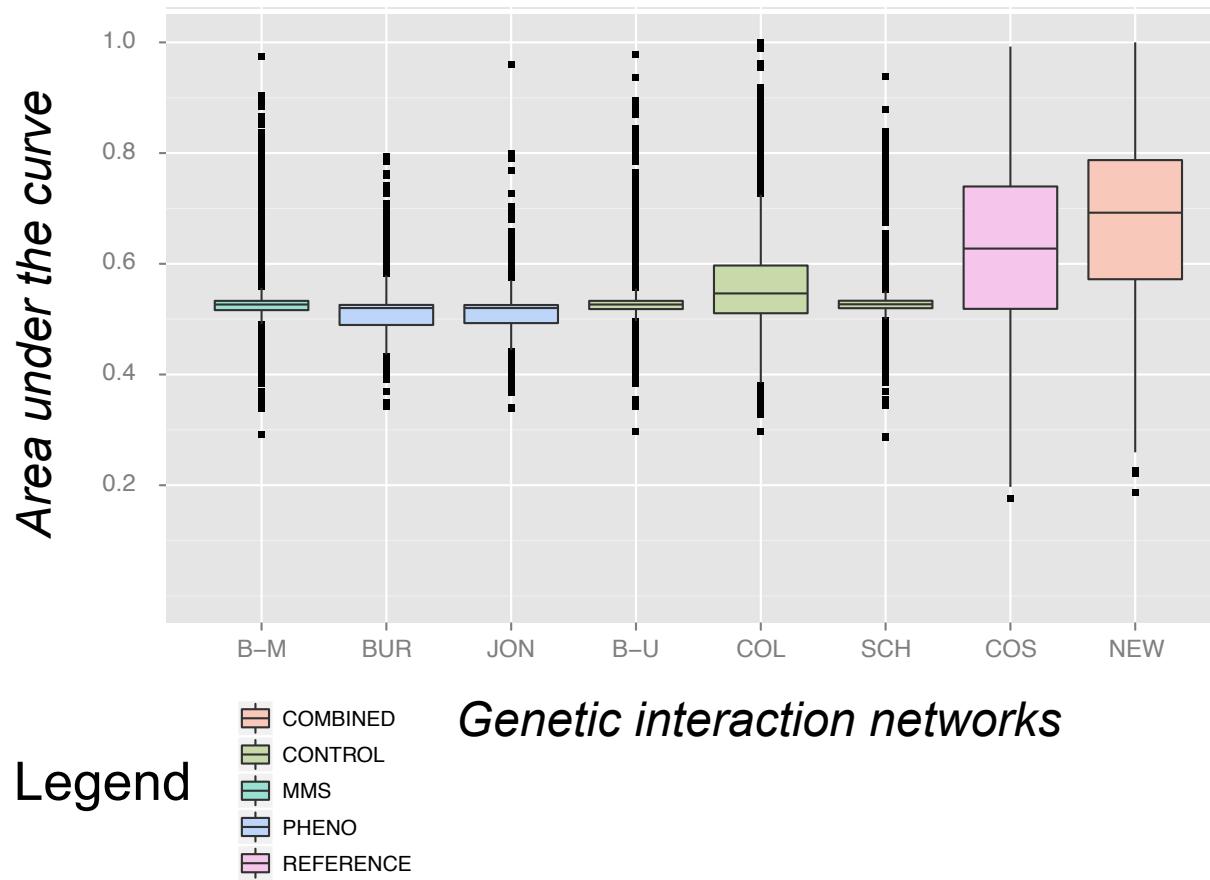
- Eight organisms:
 - Human, Mouse, Rat, Zebrafish, worm, fly, A Thaliana, yeast, E Coli
- >2,000 networks (many co-expression and physical interaction)
- Web network browser

Cytoscape plug-in



- Has all GeneMANIA functionally,
- Can use it to access older GeneMANIA data releases,
- Can add new organisms,
- Can integrate GeneMANIA networks with other Cytoscape analyses,
- Supports longer query lists.

QueryRunner



- Runs GO function prediction from the command line.
- Does cross-validation to assess predictive performance of a set of networks
- Can assess “added predictive value of new data”

(Michaut et al, in press)

STRING: <http://string-db.org/>

Home · Download · Help/Info  STRING 8.3

STRING - Known and Predicted Protein-Protein Interactions

search by name

protein name:
(examples: #1 #2 #3)

(STRING understands a variety of protein names and accessions; you can also try a [random entry](#))

organism: 

interactors wanted: COGs Proteins

please enter your protein of interest...

What it does ...

STRING is a database of known and predicted protein interactions. The interactions include direct (physical) and indirect (functional) associations; they are derived from four sources:

Genomic Context

High-throughput Experiments

(Conserved) Coexpression

Previous Knowledge

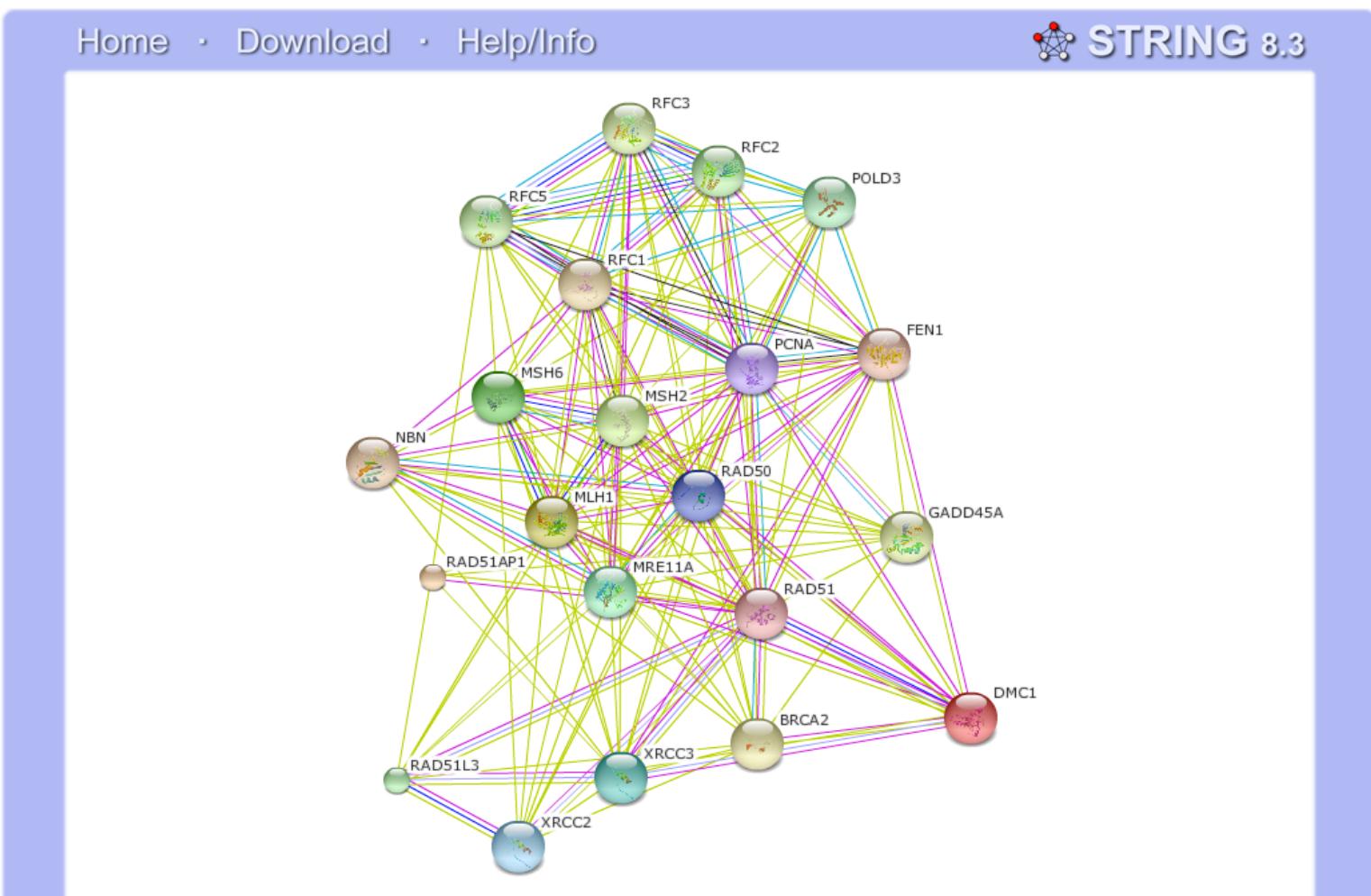
STRING quantitatively integrates interaction data from these sources for a large number of organisms, and transfers information between these organisms where applicable. The database currently covers 2,590,259 proteins from 630 organisms.

[More Info](#) [Funding / Support](#) [Acknowledgements](#) [Use Scenarios](#)

STRING (*Search Tool for the Retrieval of Interacting Genes/Proteins*) is being developed at [CPR](#), [EMBL](#), [SIB](#), [KU](#), [TUD](#) and [UZH](#).
STRING references: [Szklarczyk et al. 2011](#) / [2009](#) / [2007](#) / [2005](#) / [2003](#) / [Snel et al. 2000](#).
Miscellaneous: [Access Statistics](#), [Robot Access Guide](#), [STRING/STITCH Blog](#), [Supported Browsers](#).

What's New? This is version 8.3 of STRING - January 2011:[payload](#) mechanism introduced.
Sister Projects: check out [STITCH](#) and [eggNOG](#) - two sister projects built on STRING data!
Previous Releases: Trying to reproduce an earlier finding? Confused? Refer to our [old releases](#).

STRING results

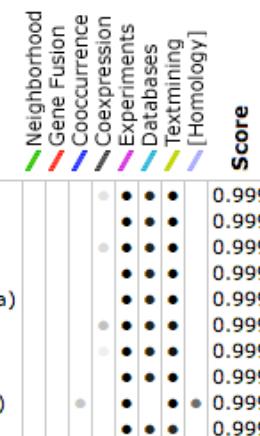


STRING results

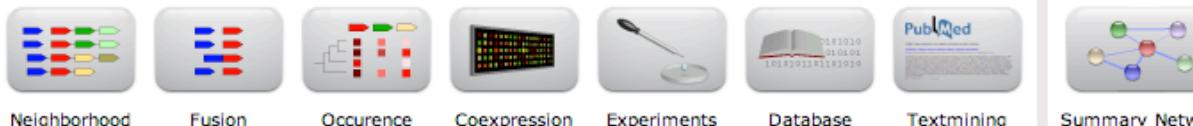
● RAD51 DNA repair protein RAD51 homolog 1 (hRAD51) (HsRAD51); May participate in a common DNA damage response pathway associated with the activation of homologous recombination and double-strand break repair. Binds to single and double stranded DNA and exhibits DNA-dependent ATPase activity. Underwinds duplex DNA and forms helical nucleoprotein filaments (340 aa)
(*Homo sapiens*)

Predicted Functional Partners:

● FEN1	Flap endonuclease 1 (EC 3.1.-.-) (Flap structure-specific endonuclease 1) (FEN-1) (Maturation f [...] (380 aa)							0.999
● NBN	Nibrin (Nijmegen breakage syndrome protein 1) (Cell cycle regulatory protein p95); Component of [...] (754 aa)							0.999
● RFC1	Replication factor C subunit 1 (Replication factor C large subunit) (RF-C 140 kDa subunit) (Act [...] (1148 aa)							0.999
● BRCA2	Breast cancer type 2 susceptibility protein (Fanconi anemia group D1 protein); Involved in doub [...] (3418 aa)							0.999
● GADD45A	Growth arrest and DNA-damage-inducible protein GADD45 alpha (DNA- damage-inducible transcript 1 [...] (165 aa)							0.999
● RFC3	Replication factor C subunit 3 (Replication factor C 38 kDa subunit) (RFC38) (Activator 1 38 kD [...] (356 aa)							0.999
● RFC5	Replication factor C subunit 5 (Replication factor C 36 kDa subunit) (RF-C 36 kDa subunit) (RFC [...] (340 aa)							0.999
● RFC2	Replication factor C subunit 2 (Replication factor C 40 kDa subunit) (RF-C 40 kDa subunit) (RFC [...] (354 aa)							0.999
● RAD51L3	DNA repair protein RAD51 homolog 4 (R51H3) (RAD51-like protein 3) (TRAD); Involved in the homol [...] (328 aa)							0.999
● POLD3	DNA polymerase subunit delta 3 (DNA polymerase subunit delta p66); Required for optimal DNA pol [...] (466 aa)							0.999



Views:



GeneMANIA/STRING comparison

- **STRING (2003-present)**
 - Large organism coverage
 - Protein focused, nodes link to protein structures
 - Very good information links, integration with Uniprot
 - Uses eight pre-computed networks
 - Heavy use of phylogeny to infer functional interactions, also contains text mining derived interactions
 - Uses “direct interaction” to score nodes
 - Link weights are “Probability of functional interaction”
- **GeneMANIA webserver (2010-present)**
 - Covers nine major model organisms (but can add more with plugin)
 - Gene focused
 - Thousands of networks, weights are not pre-computed, **can upload your own network**
 - Relies heavily on functional genomic data: so has genetic interactions, phenotypic info, chemical interactions
 - Allows enrichment analysis
 - Uses “label propagation” to score nodes

Learning Objectives of Module 5

- **Understand** the concepts: *functional interaction network*, *guilt-by-association*, *gene recommender systems*.
- **Understand** the concept of context-specific network weighting schemes.
- **Understand** the difference between *direct interaction* and *label propagation* methods for predicting gene function.
- **Be able** to use gene recommender systems (e.g. GeneMANIA) to answer two types of questions about gene function: “what does my gene do?” and “give me more genes like these”
- **Be able** to select the appropriate network weighting scheme to answer your questions about gene function.

We are on a Coffee Break &
Networking Session

