

# Biological Network Analysis:

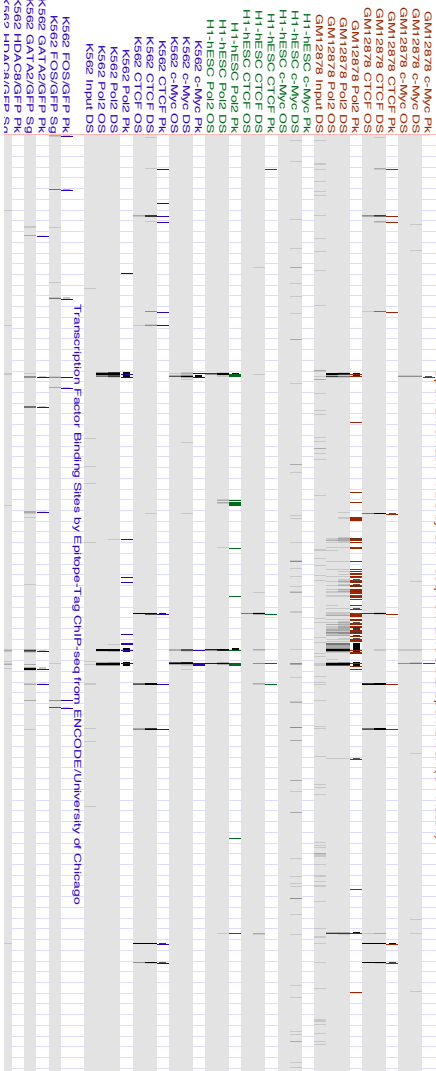
Hierarchies, Mutational  
Constraints & Logical  
Circuits in Regulation

Mark Gerstein  
Yale

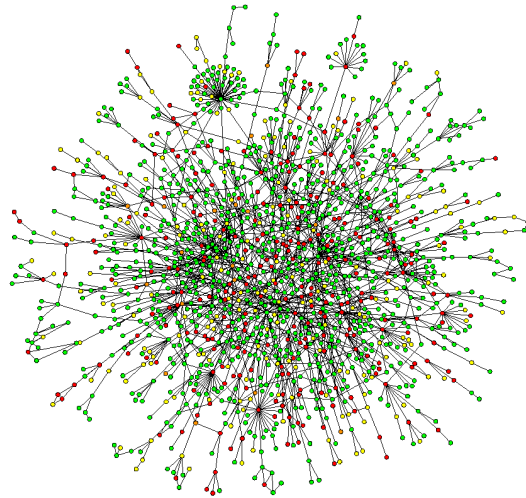
Slides freely downloadable from  
**Lectures.GersteinLab.org**  
& “tweetable” (via @markgerstein).  
See last slide  
for references & more info.



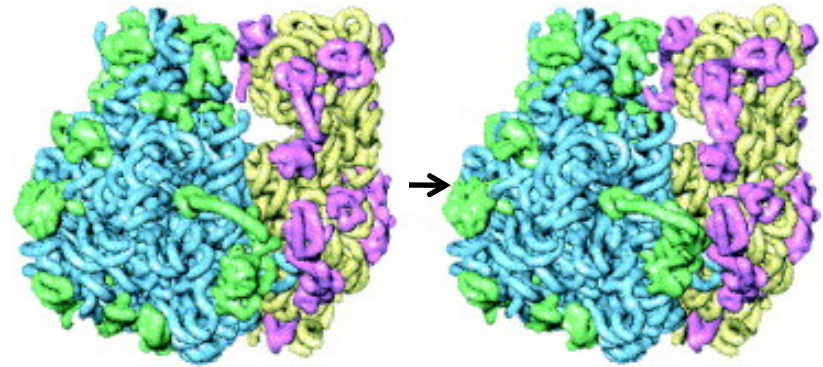
# Networks for Genome Annotation: midway point in terms of level of understanding



1D: Complete  
Partlist  
("Elements" in  
genomic tracks)



~2D: Network  
Wiring Diagram of a  
Molecular System



3D & 4D:  
Detailed structural  
understanding of cellular  
machinery

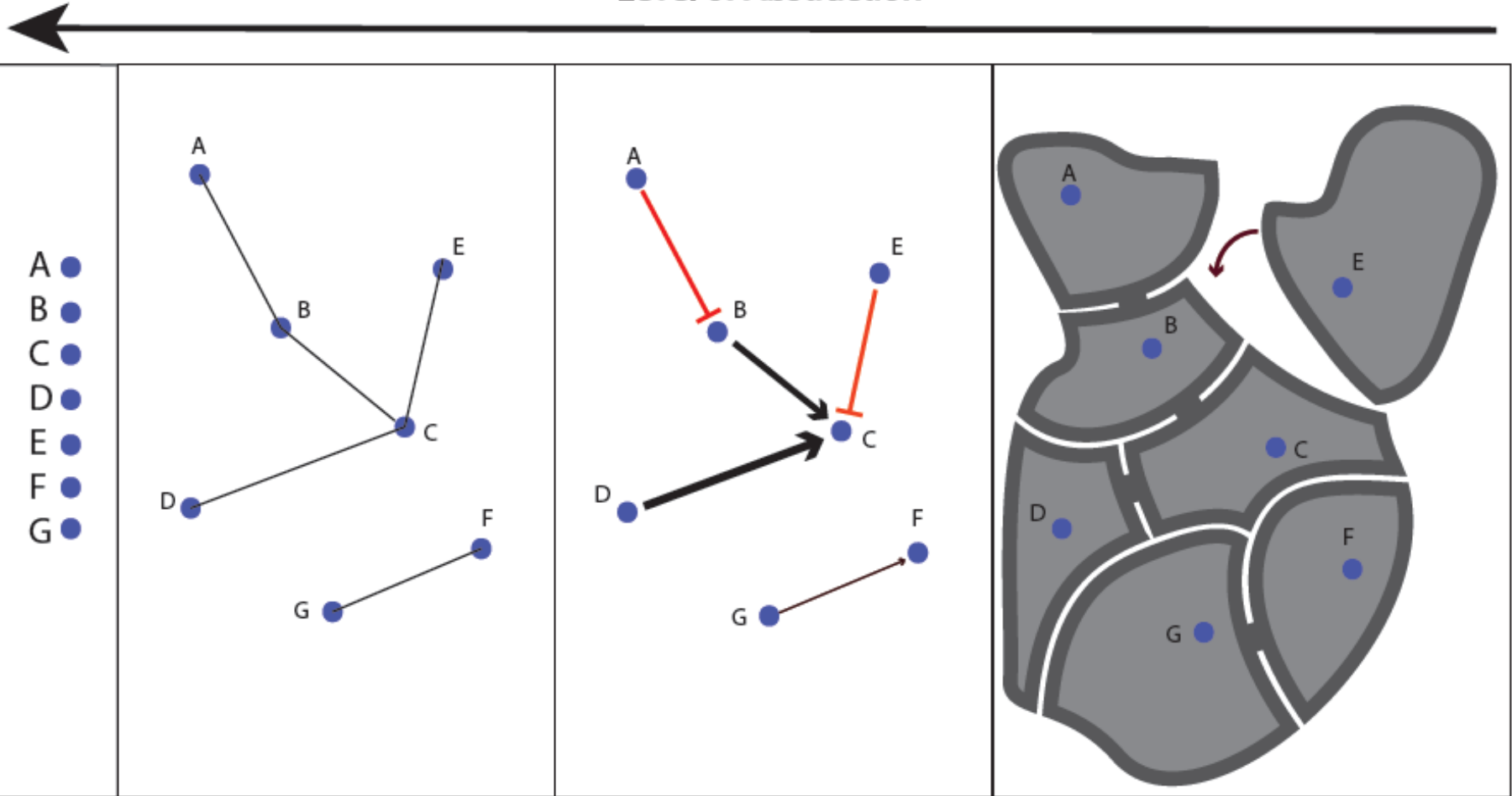
[UCSC genome browser]

[Jeong et al. Nature, 41:411]

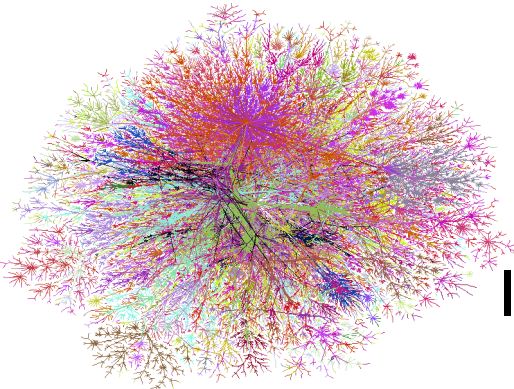
[Chiu et al. Trends in Cell Biol, 16:144]

Connecting systems biology to  
genomics

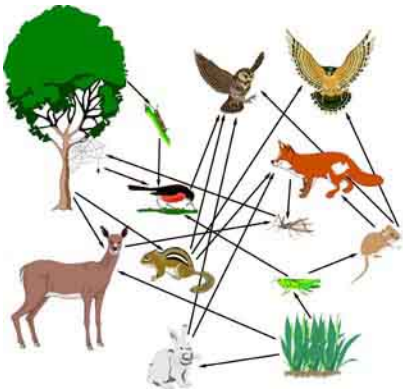
# Level of Abstraction



# Networks as a common language in many scientific contexts



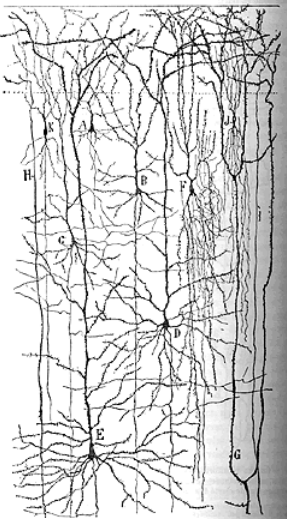
Internet



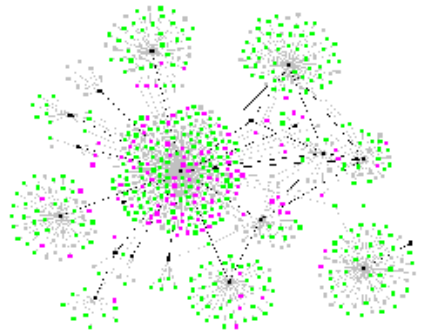
Food Web



Electronic Circuit



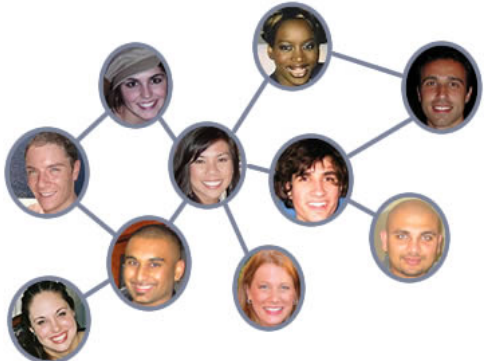
Neural Network



Disease Spread

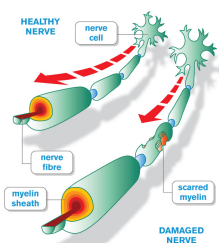
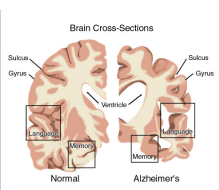
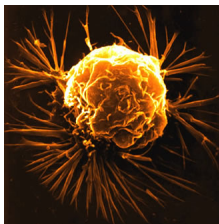


Protein Interactions



Social Network

# Network pathology & pharmacology

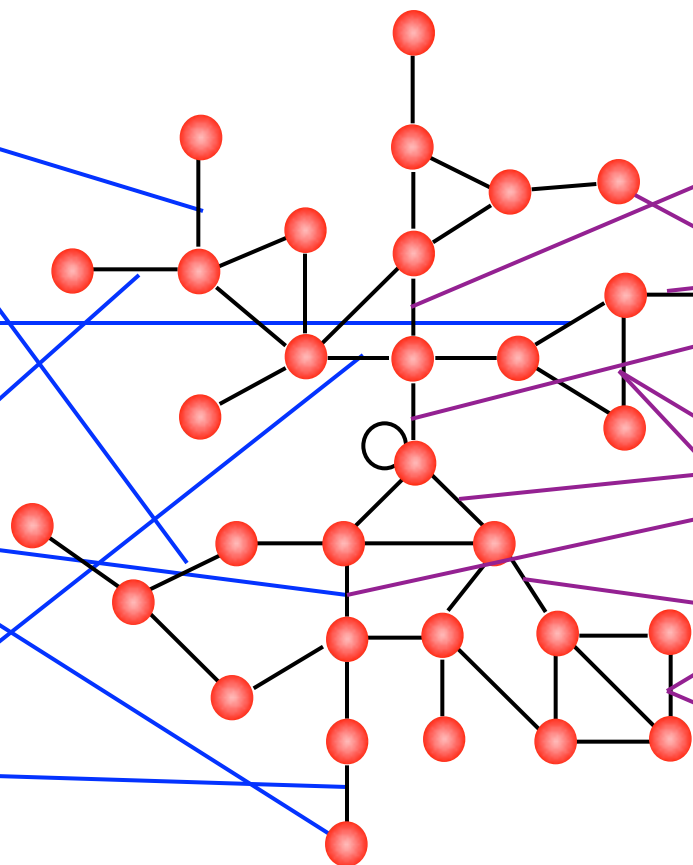


**Breast Cancer**

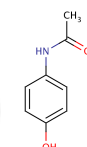
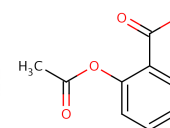
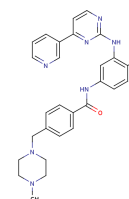
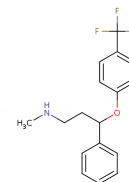
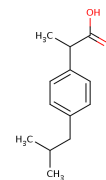
**Alzheimer's Disease**

**Parkinson's Disease**

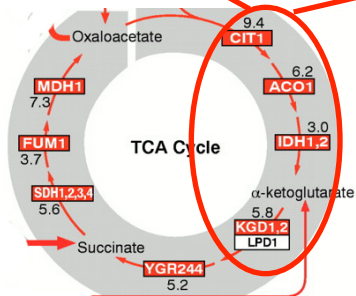
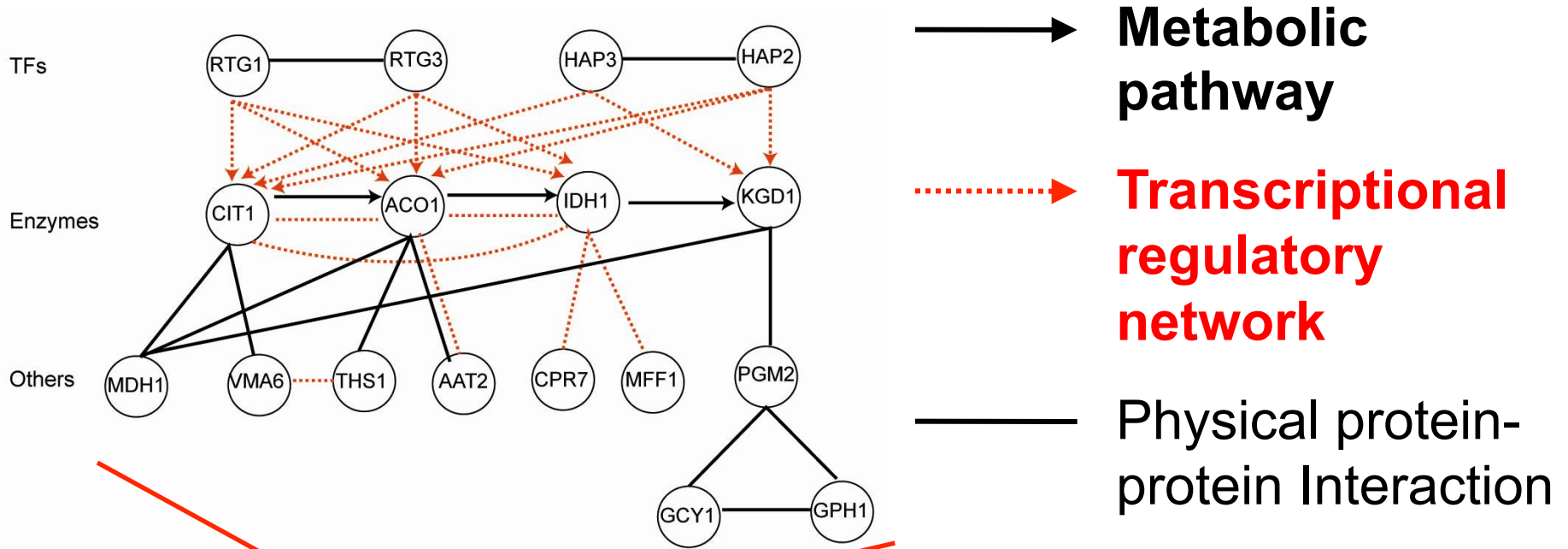
**Multiple Sclerosis**



Interactome networks



# Combining networks forms an ideal way of integrating diverse information

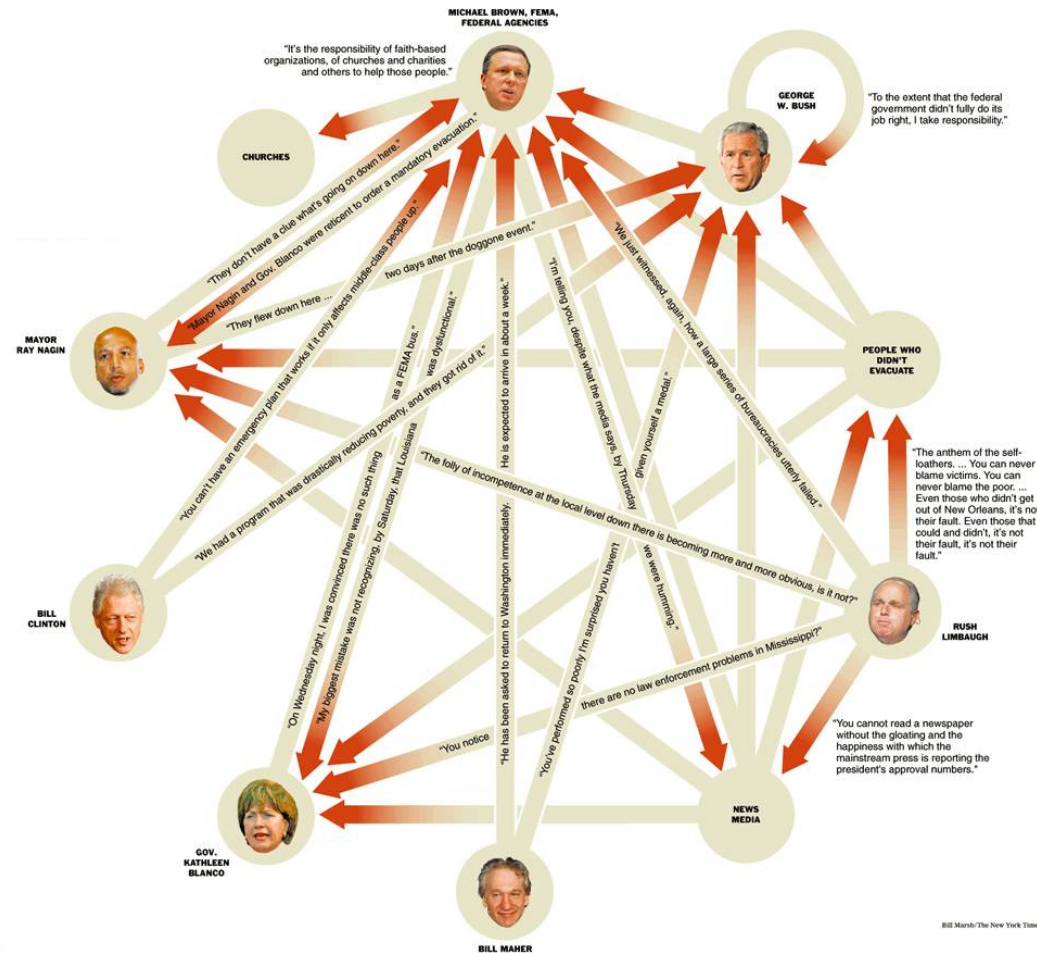
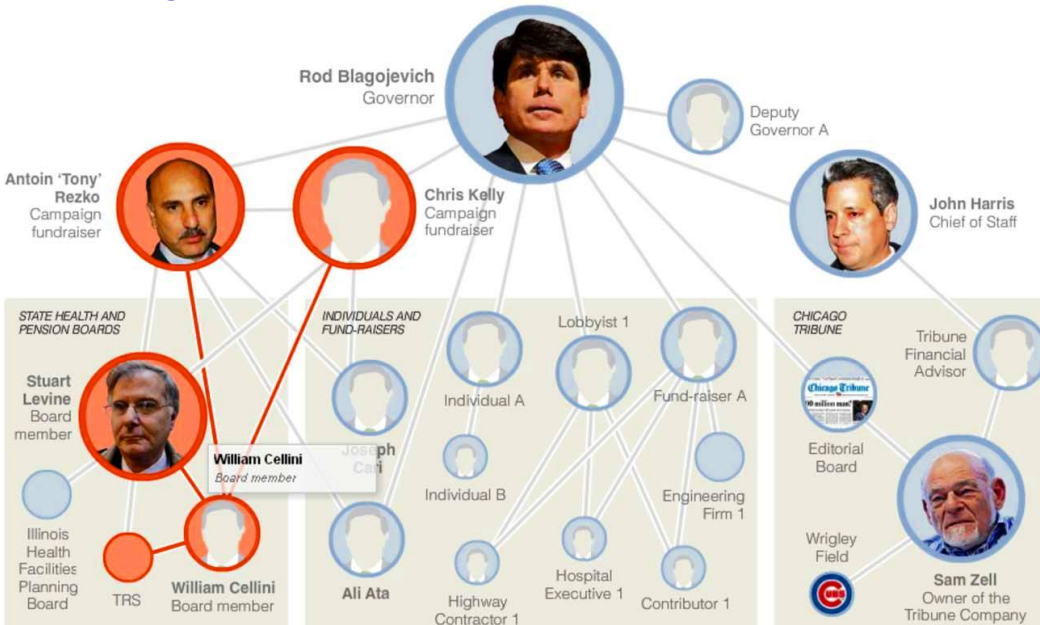


**Part of the TCA cycle**

Genetic interaction (synthetic lethal)  
Signaling pathways

# Exploiting Network Analogies to Gain Intuition

## Guilt by association



## Finding the causal regulator (the "Blame Game")

[NY Times, 2-Oct-05, 9-Dec-08]

# Biological Network Analysis: Hierarchies, Mutational Constraints & Logical Circuits in Regulation

- **Why Networks ?**

- Representational sweet spot + Intuition from cross-disciplinary comparisons

- **Organizing the Regulatory Network into a Hierarchy**

- Construction: local BFS v global simulated annealing
- Information flow bottlenecks in the middle & greater mid-level **collaboration** to ease them, esp. in more complex organisms
- Differences between kinase & TF hierarchy

- **Analyzing the Impact of Variation on the Network**

- Node Variation: **more connectivity = more constraint**
- Useful analogies to **designed systems**

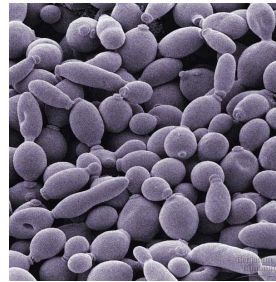
- **Going from Regulatory Networks to Logical Circuits**

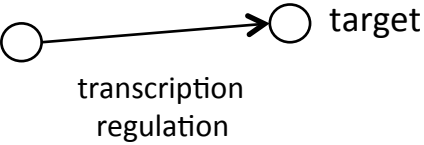
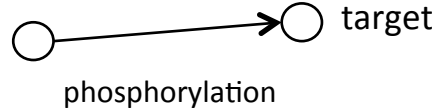
- Preponderance of OR gates in the human network v yeast
- Relation to cancer (myc)
- More logical structure at top of hierarchy



# Regulatory networks: from yeast to human

Also: Regulatory networks from E. coli, mouse, rat, &c for comparison



<p>Transcriptional regulatory network</p> <p>transcription factor  target</p>	<p>150 TFs, 6000 genes</p>	<p>1300 TFs, 20000 genes</p>
<p>Kinase network (regulation in translational level)</p> <p>kinase  target</p>	<p>130 kinases, 1000 targets</p>	<p>500 kinases, up to 6000 targets</p>

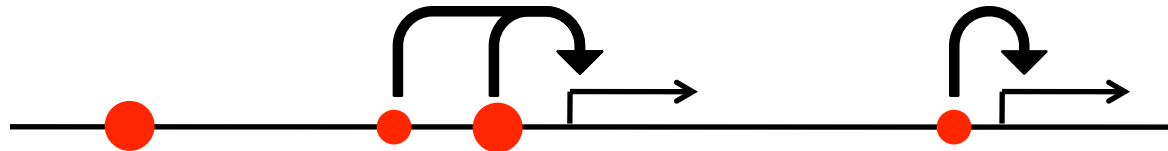
# Data Flow: peaks to proximal & distal networks

[ Cheng et al., *Bioinfo.* ('11);  
Gerstein et al. *Nature* (in press, '12) ;  
Yip et al., *GenomeBiology* (in press, '12)]

Peak Calling

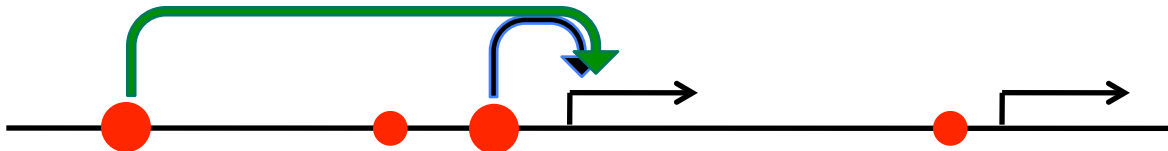


Assigning TF binding sites to targets

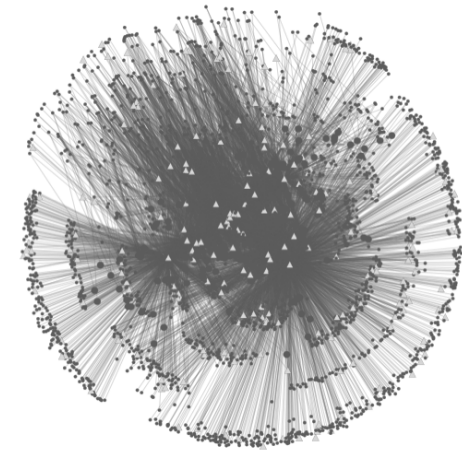


Filtering high confidence edges & distal regulation

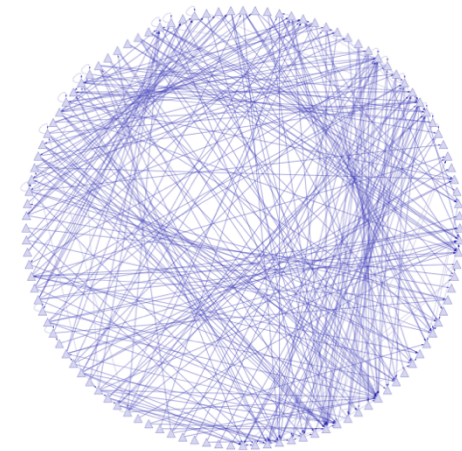
Based on stat. model combining  
signal strength & location relative to typical binding



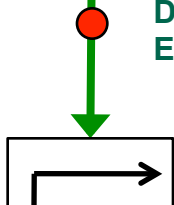
~500K  
Edges



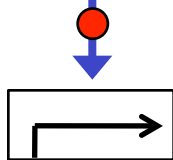
~26K  
Edges

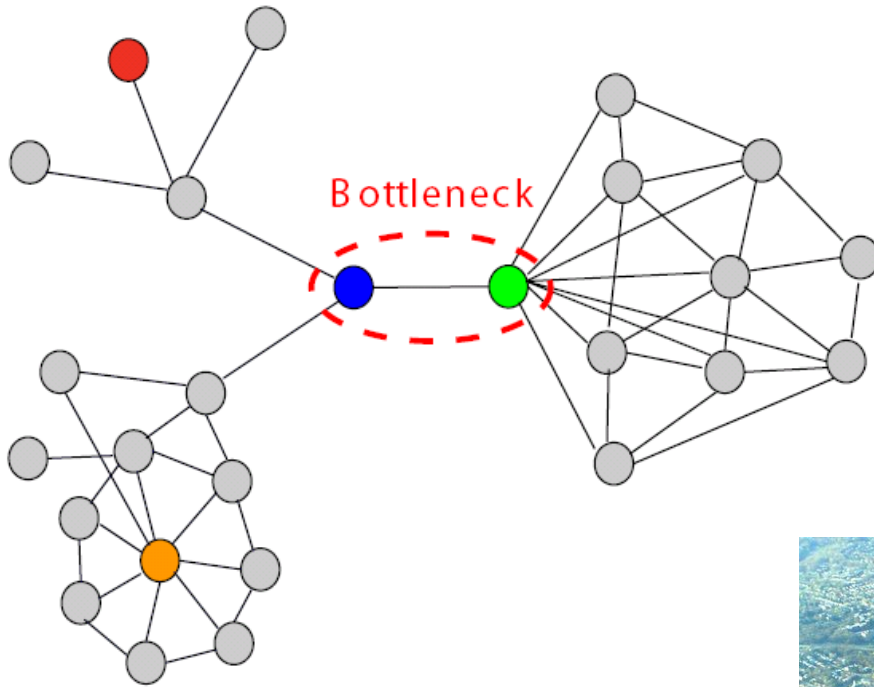


Potential  
Distal  
Edge



Strong  
Proximal  
Edge



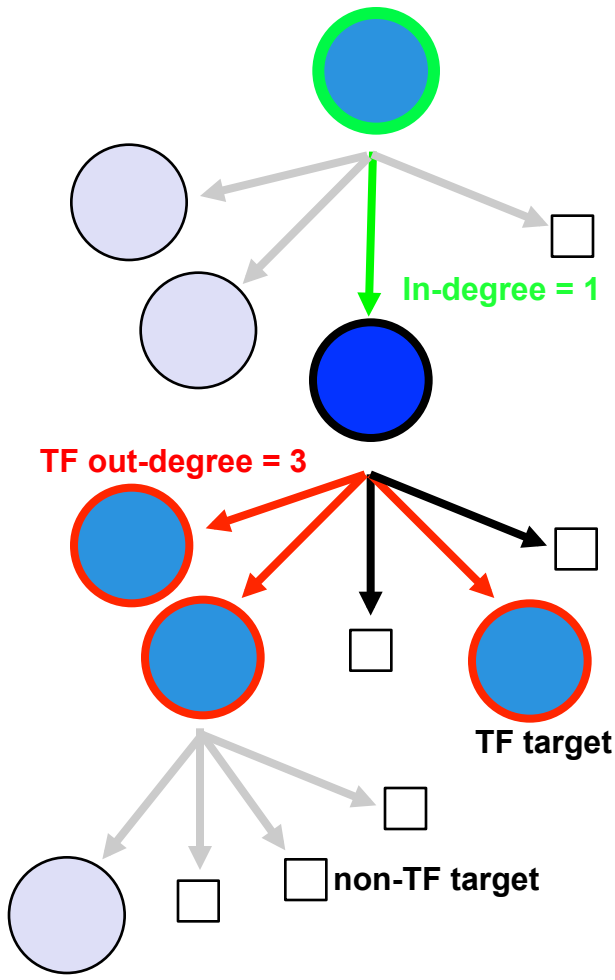


- Hub-bottleneck node
- Non-hub-bottleneck node
- Hub-non-bottleneck node
- Non-hub-non-bottleneck node

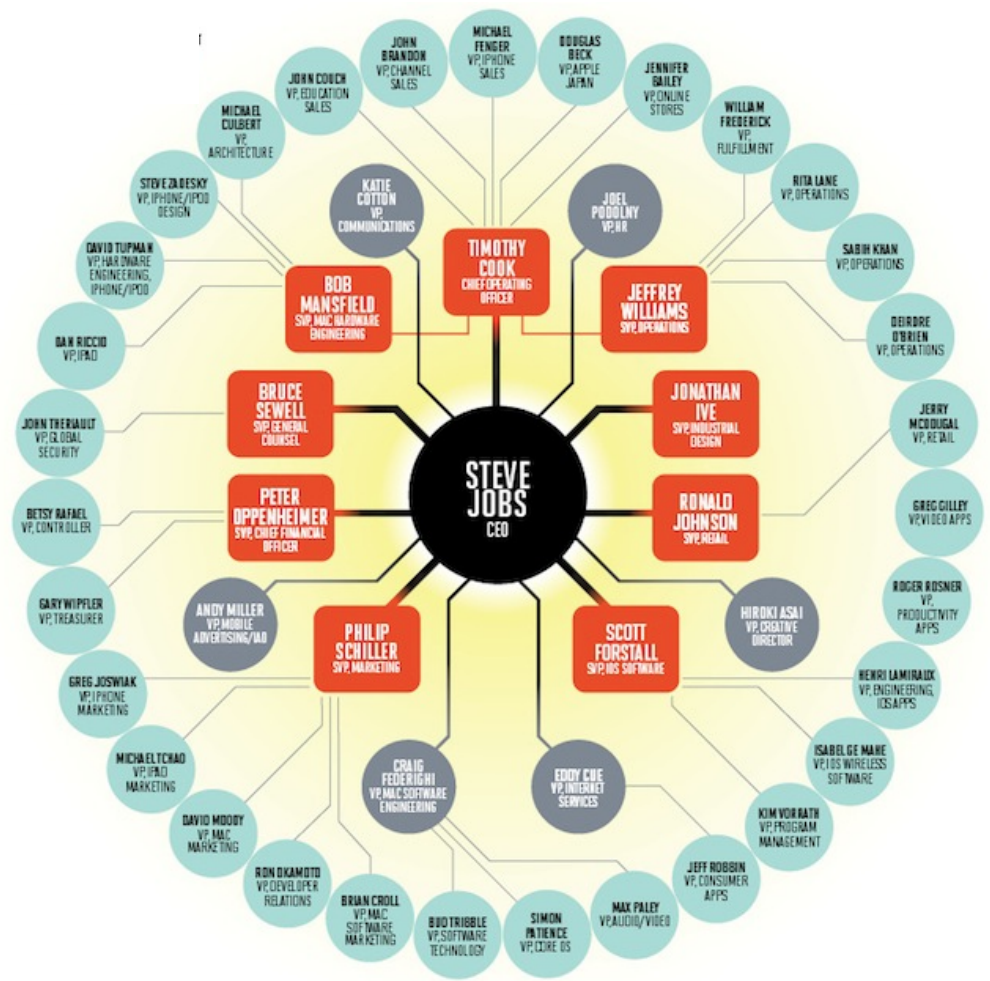


## Network Stats to Identify Bottlenecks & Hubs

[Yu et al., PLOS CB (2007)]



**Hierarchy Height Statistic = (normalized TF Out deg. – In deg.)**



## Network Stats to Identify Hierarchy

# Biological Network Analysis: Hierarchies, Mutational Constraints & Logical Circuits in Regulation

- **Why Networks ?**

- Representational sweet spot + Intuition from cross-disciplinary comparisons

- **Organizing the Regulatory Network into a Hierarchy**

- Construction: local BFS v global simulated annealing
- Information flow bottlenecks in the middle & greater mid-level **collaboration** to ease them, esp. in more complex organisms
- Differences between kinase & TF hierarchy

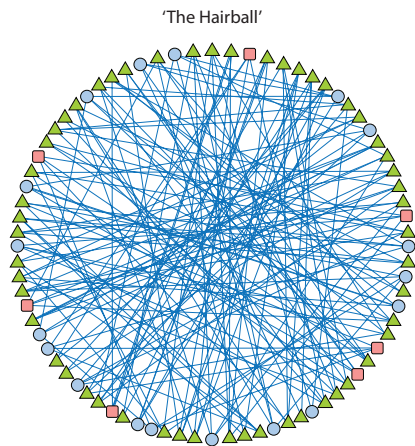
- **Analyzing the Impact of Variation on the Network**

- Node Variation: **more connectivity = more constraint**
- Useful analogies to **designed systems**

- **Going from Regulatory Networks to Logical Circuits**

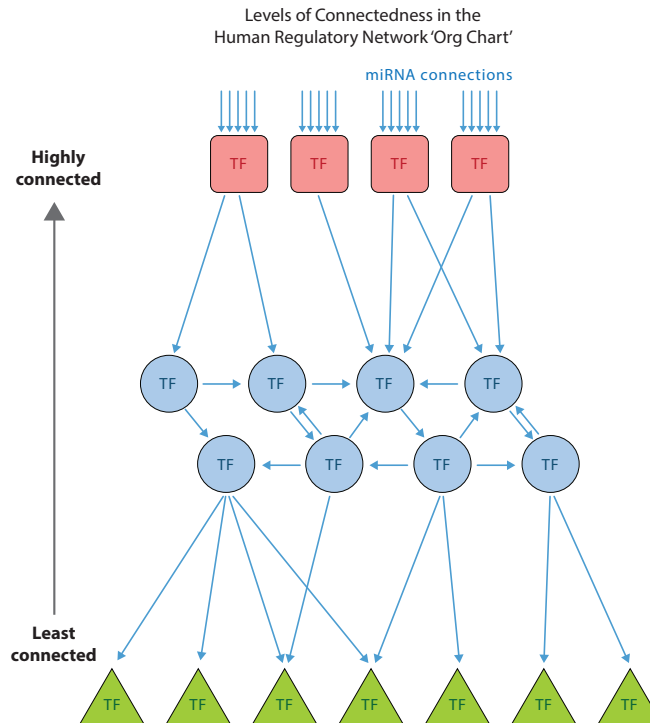
- Preponderance of OR gates in the human network v yeast
- Relation to cancer (myc)
- More logical structure at top of hierarchy

# Hierarchy algorithm: from Hairball to Hierarchy



There are an enormous number of connections and dependencies among DNA transcription factors.

However, bioinformatics techniques of network analysis can help visualize patterns in the seeming chaos of interconnections.



## Most influential

This level of highly connected transcription factors is the most conservative and persistent over time, and the least affected by immediate selection pressures and mutations.

## Middle management

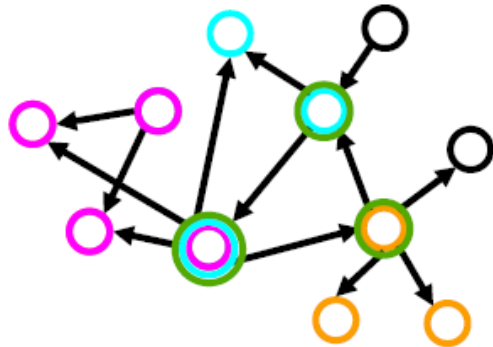
Potential connection bottlenecks between TF 'middle managers' are ameliorated by many 'cooperative' interconnections at this level, gaining robustness through great redundancy.

## Least influential

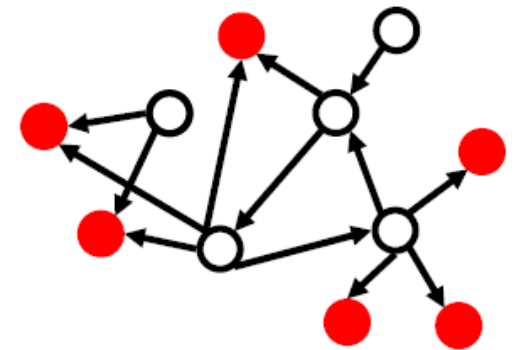
The least connected transcription factors are the most affected by selection pressures and mutations, but fewer connections and interdependencies make them less likely to cause fatal mutations.

# Determination of "Level" in Regulatory Network Hierarchy with Breadth-first Search

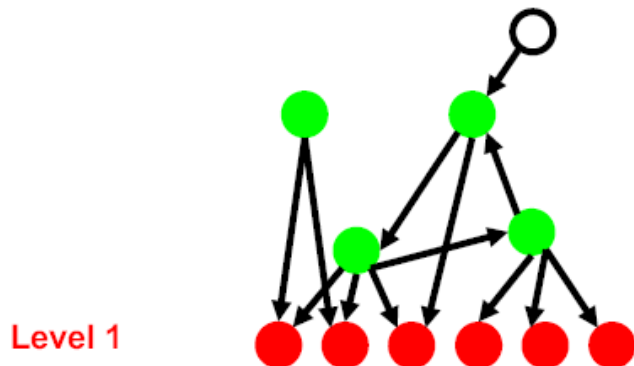
I. Example network with all 4 motifs



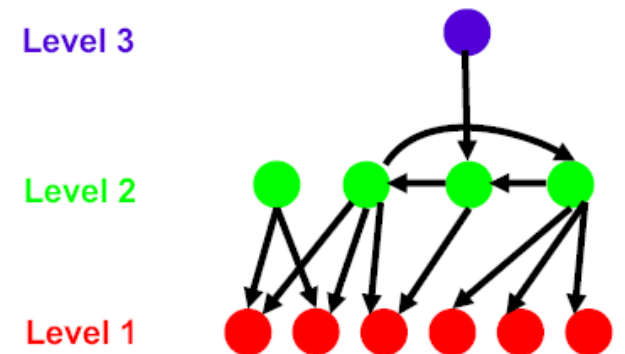
II. Finding terminal nodes (Red)



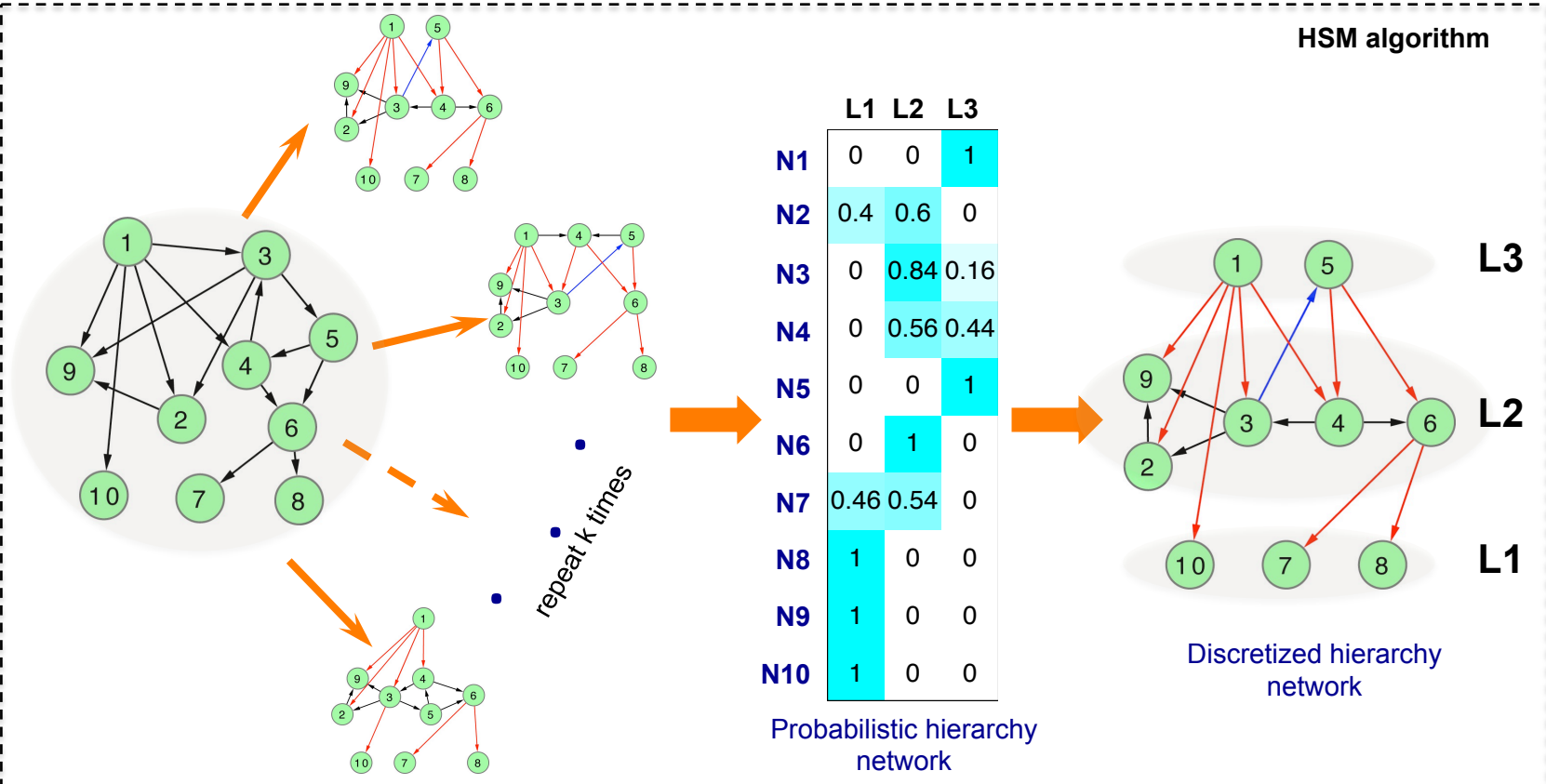
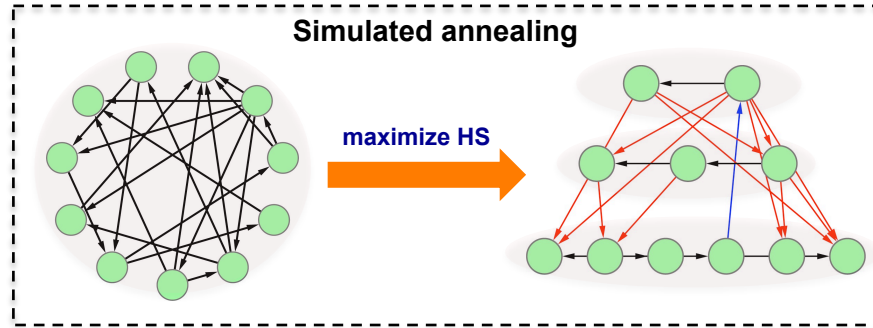
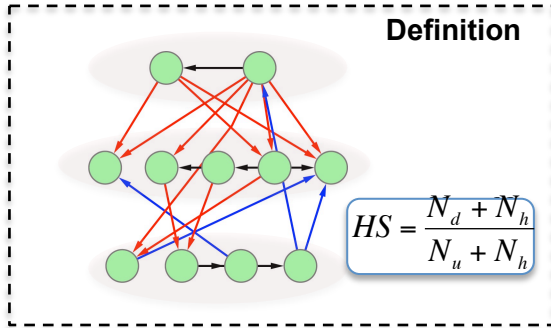
III. Finding mid-level nodes (Green)



IV. Finding top-most nodes (Blue)

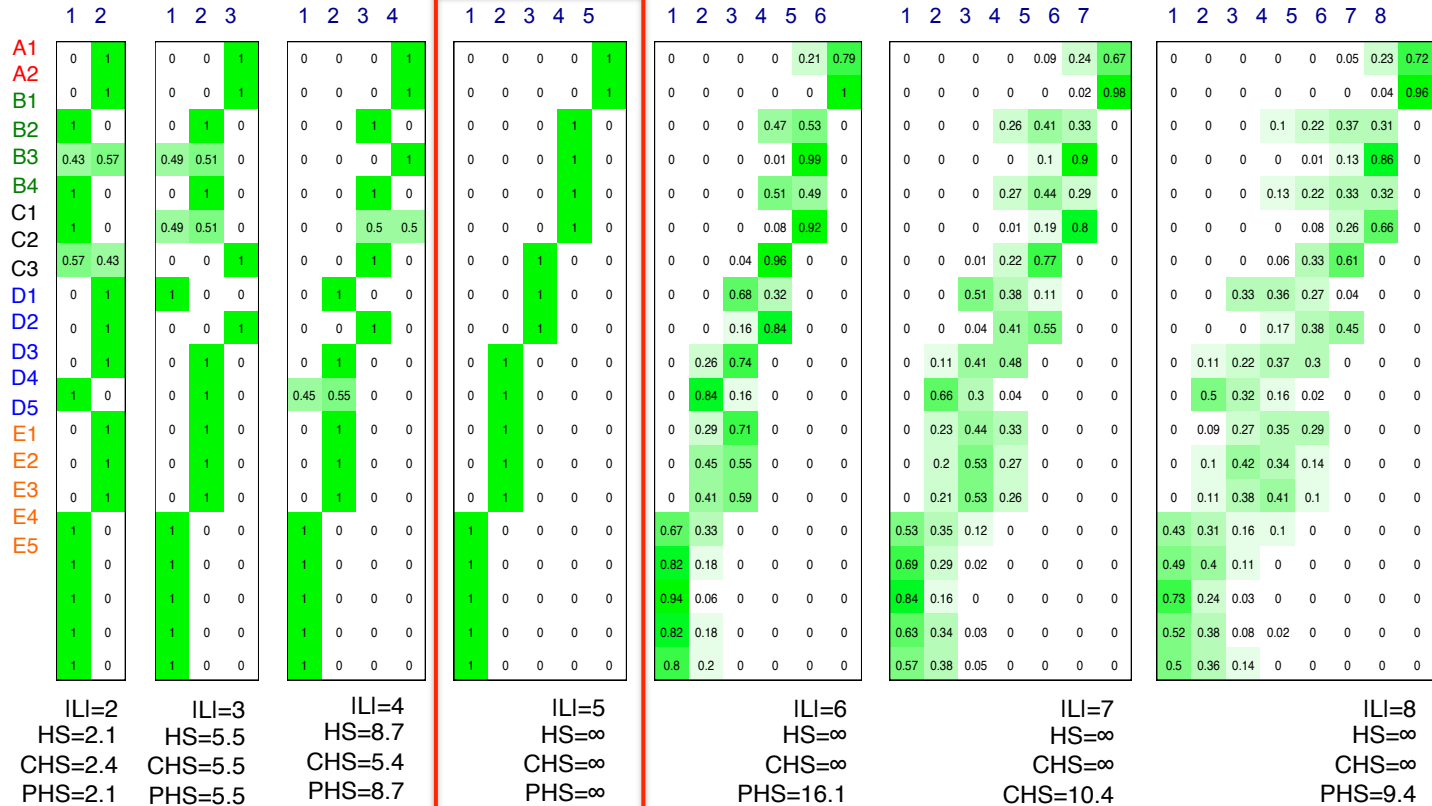
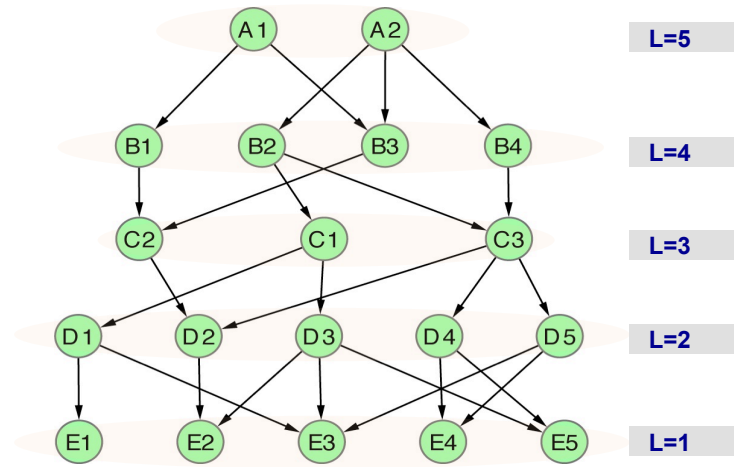


# Hierarchy Score Maximization Algorithm





# Apply HSM to a toy example



# Biological Network Analysis: Hierarchies, Mutational Constraints & Logical Circuits in Regulation

- **Why Networks ?**

- Representational sweet spot + Intuition from cross-disciplinary comparisons

- **Organizing the Regulatory Network into a Hierarchy**

- Construction: local BFS v global simulated annealing
- Information flow bottlenecks in the middle & greater mid-level **collaboration** to ease them, esp. in more complex organisms
- Differences between kinase & TF hierarchy

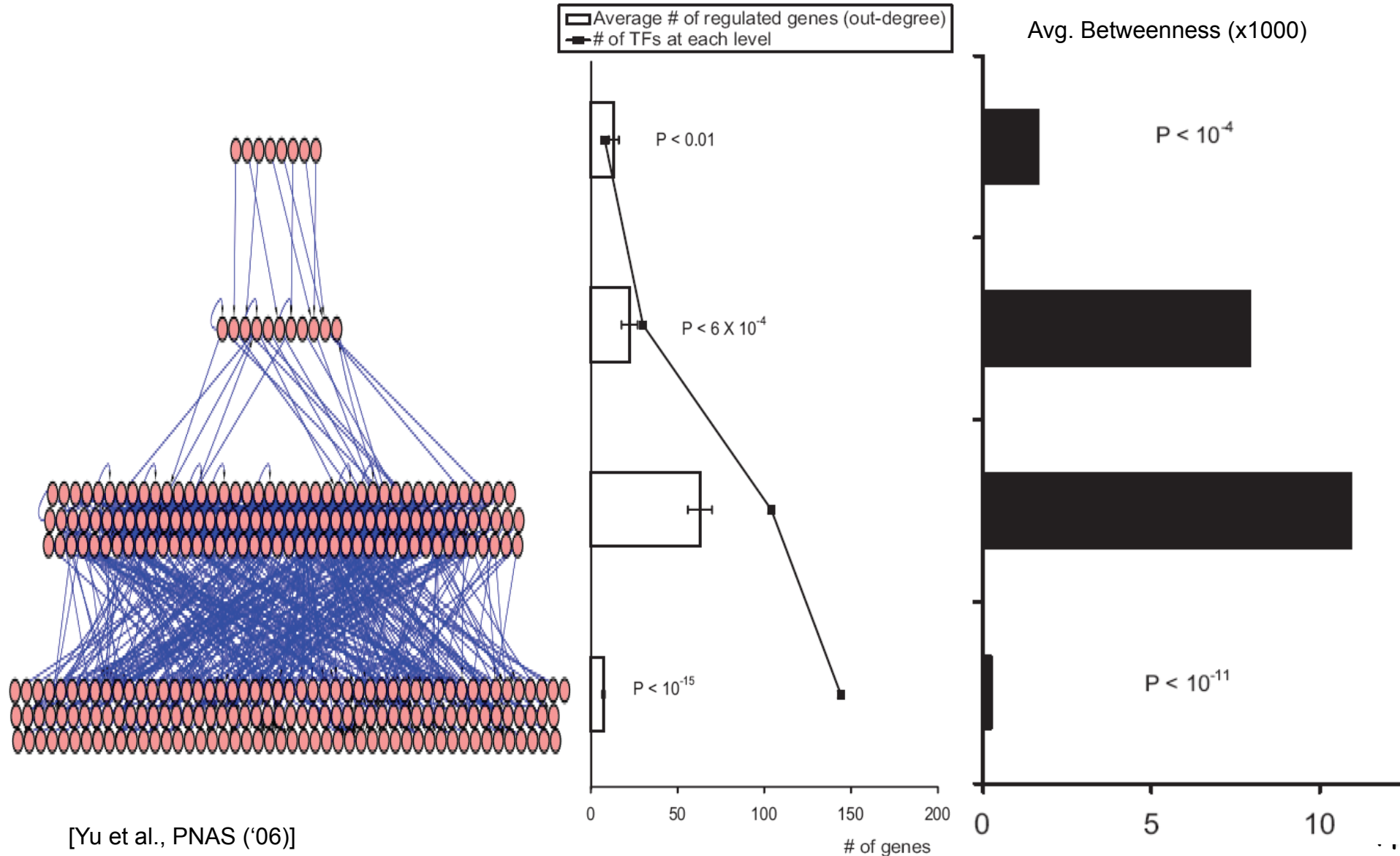
- **Analyzing the Impact of Variation on the Network**

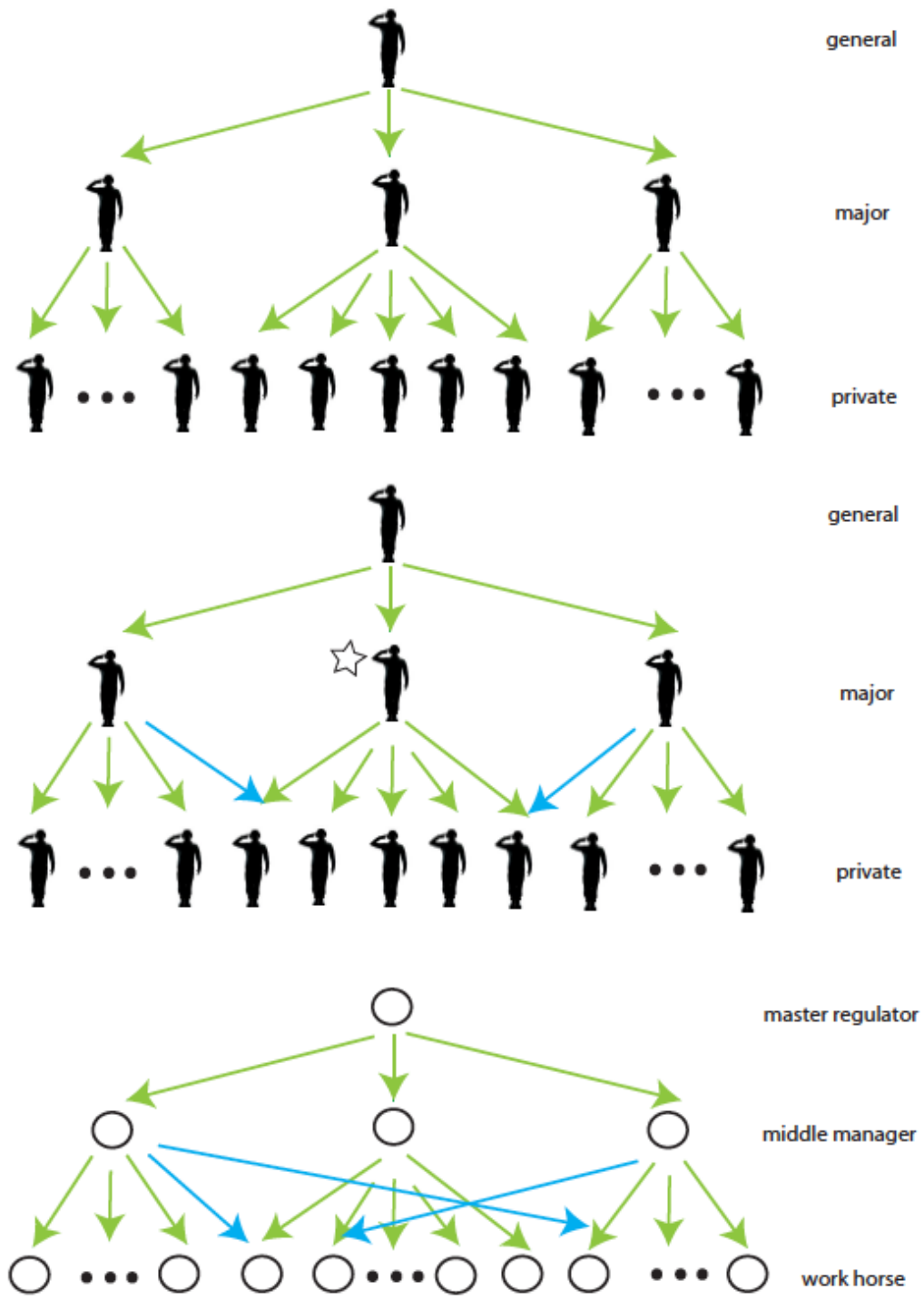
- Node Variation: **more connectivity = more constraint**
- Useful analogies to **designed systems**

- **Going from Regulatory Networks to Logical Circuits**

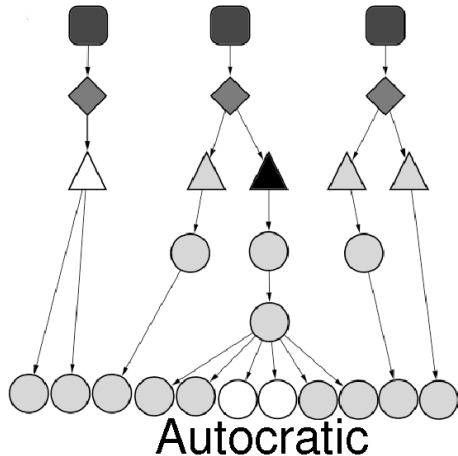
- Preponderance of OR gates in the human network v yeast
- Relation to cancer (myc)
- More logical structure at top of hierarchy

# Yeast Regulatory Hierarchy: Middle-managers Rule

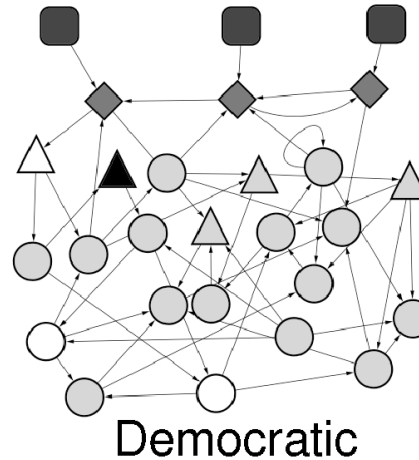




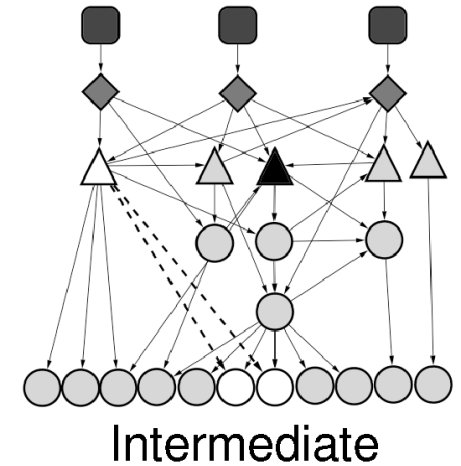
# Different kinds of Hierarchies



- Well-defined levels and a clear chain of command
- A military hierarchy



- Without well-defined levels & with more co-regulatory partnerships
- A club or a scientific collaboration network

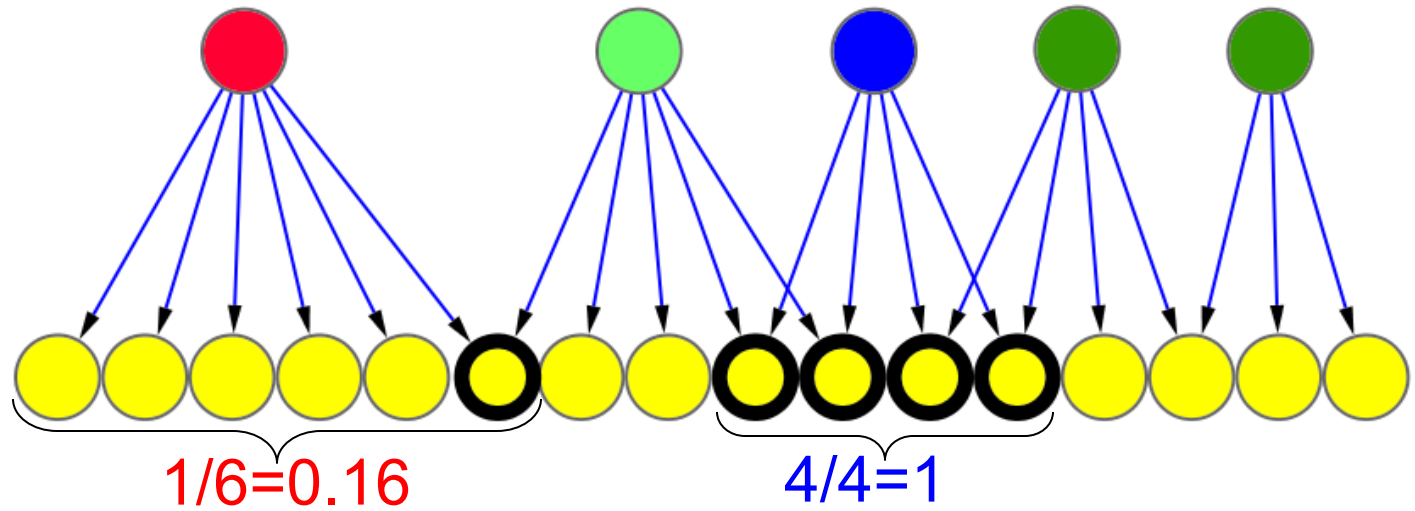


- High degree of co-regulation and can be organized into hierarchies
- A law firm

# Collaborative Nature of the Nodes

Autonomous

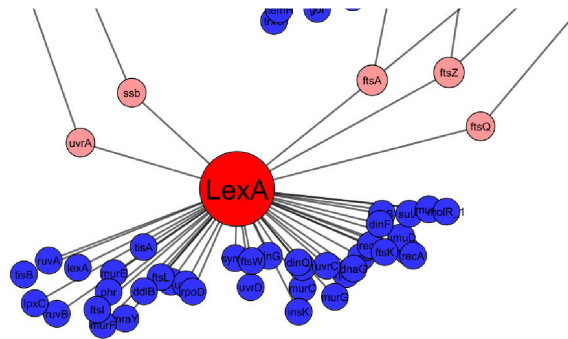
Collaborative



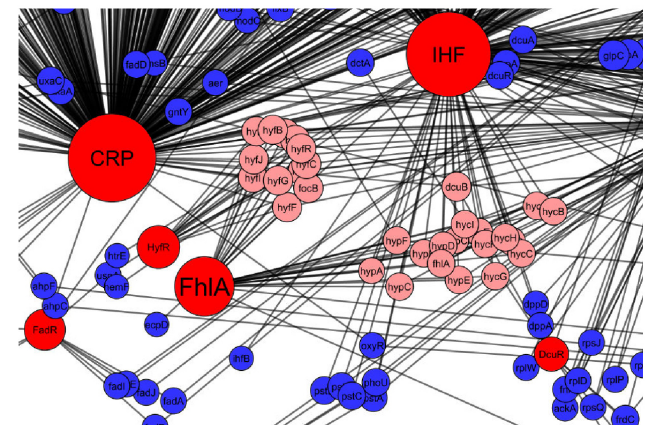
More Collaborative:  
Democratic

More Autonomous:  
Autocratic

Autonomous

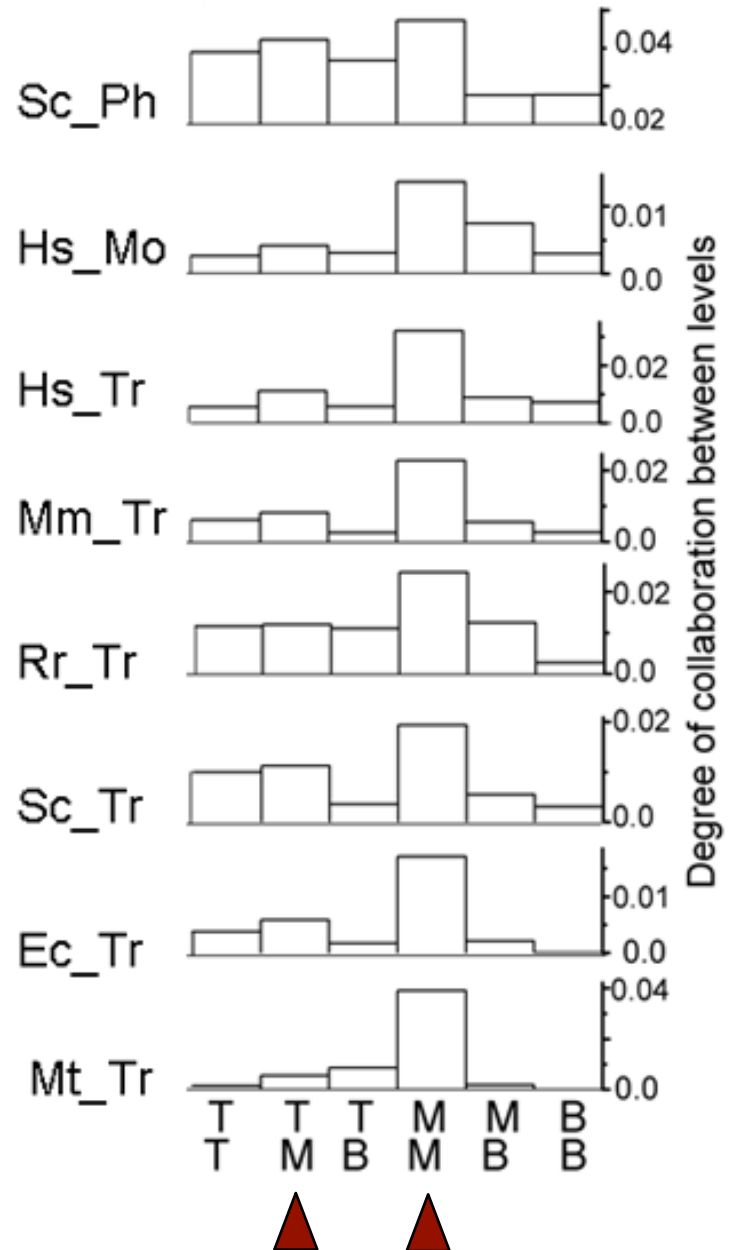
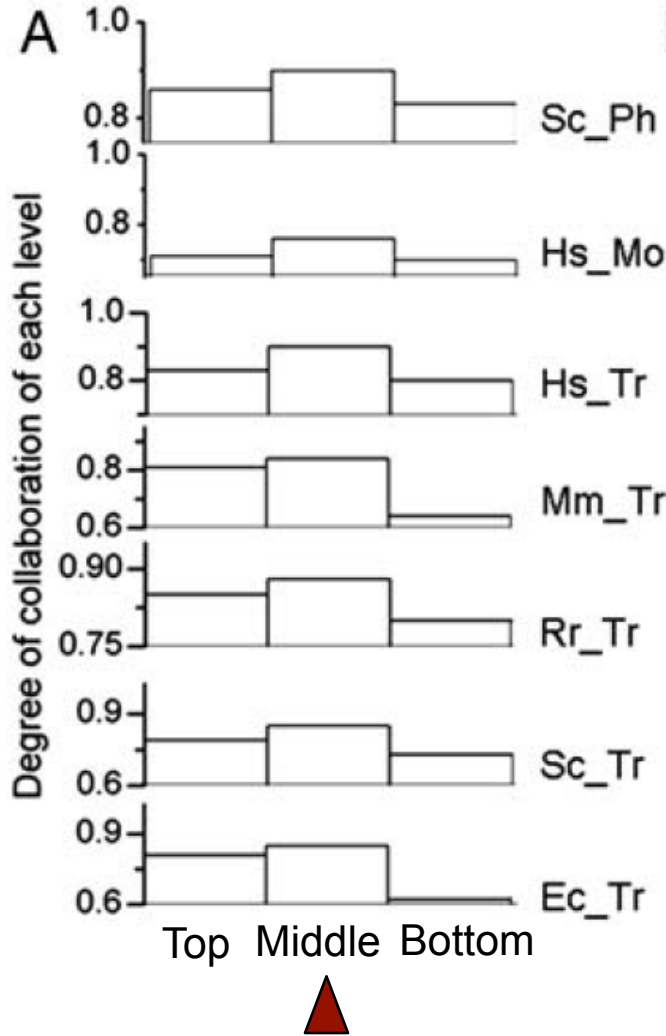


Collaborative



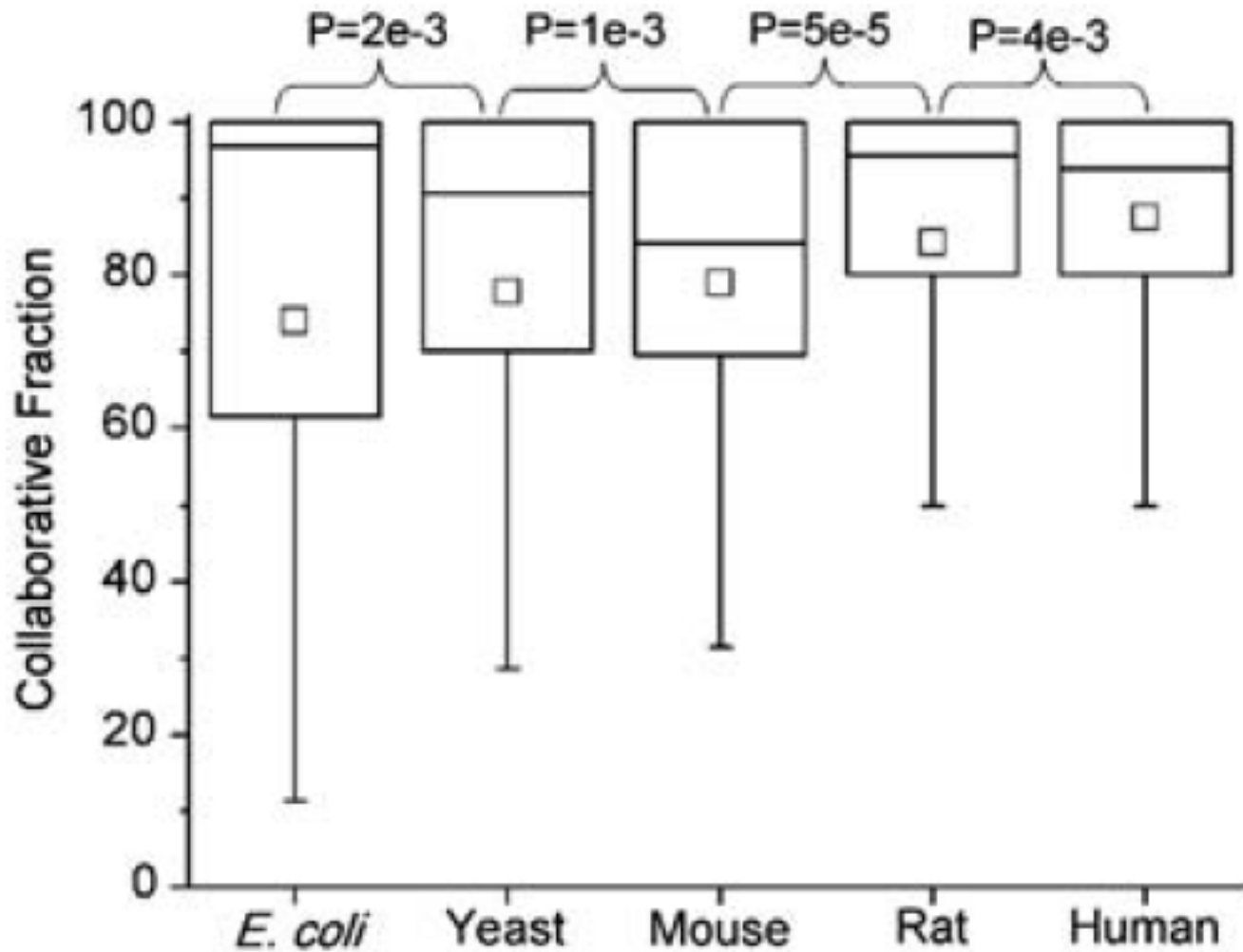
[Bhardwaj et al., PNAS (2010), in press]

# Most collaboration involves middle level



[Bhardwaj et al., PNAS (2010), in press]

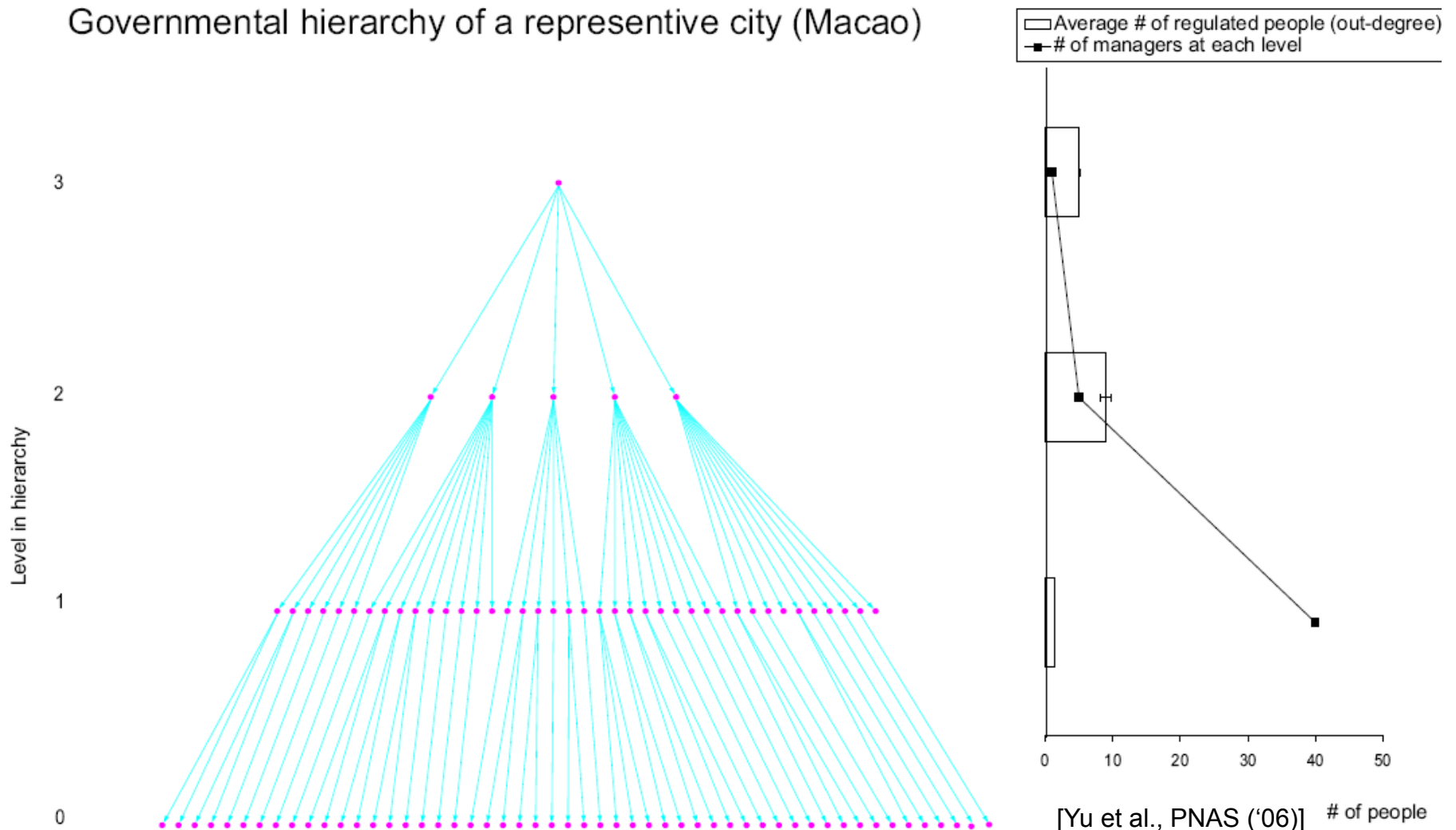
# Higher species have more collaborative nodes





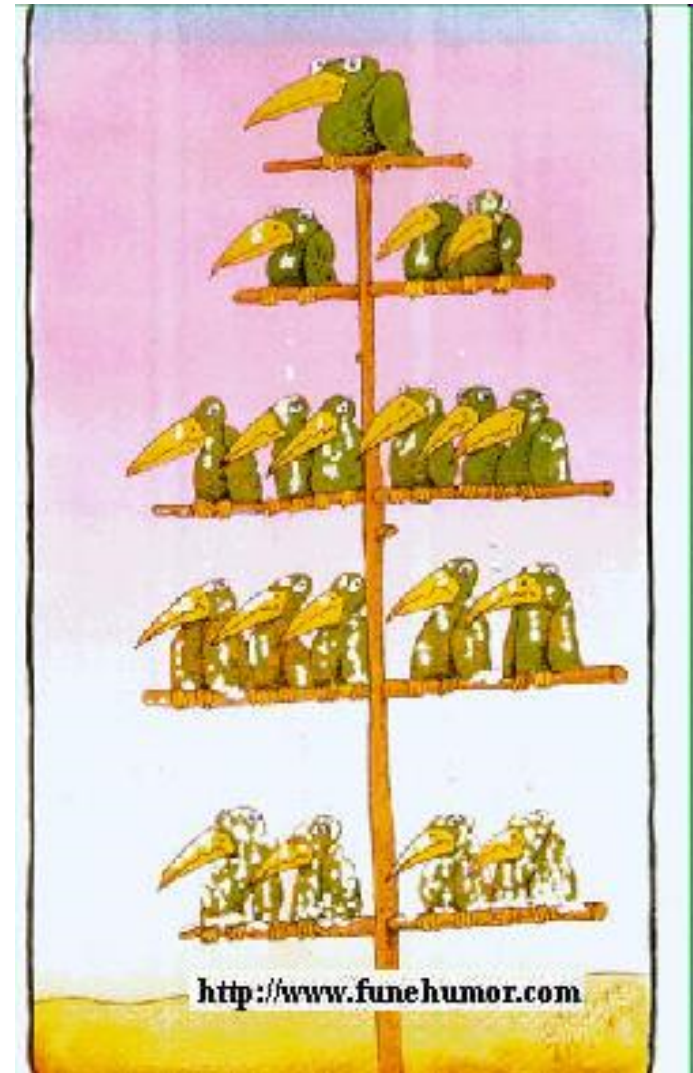
# Yeast Network Similar in Structure to Government Hierarchy with Respect to Middle-managers

Governmental hierarchy of a representative city (Macao)



# Middle Managers Interact the Most in Efficient Corporate Settings

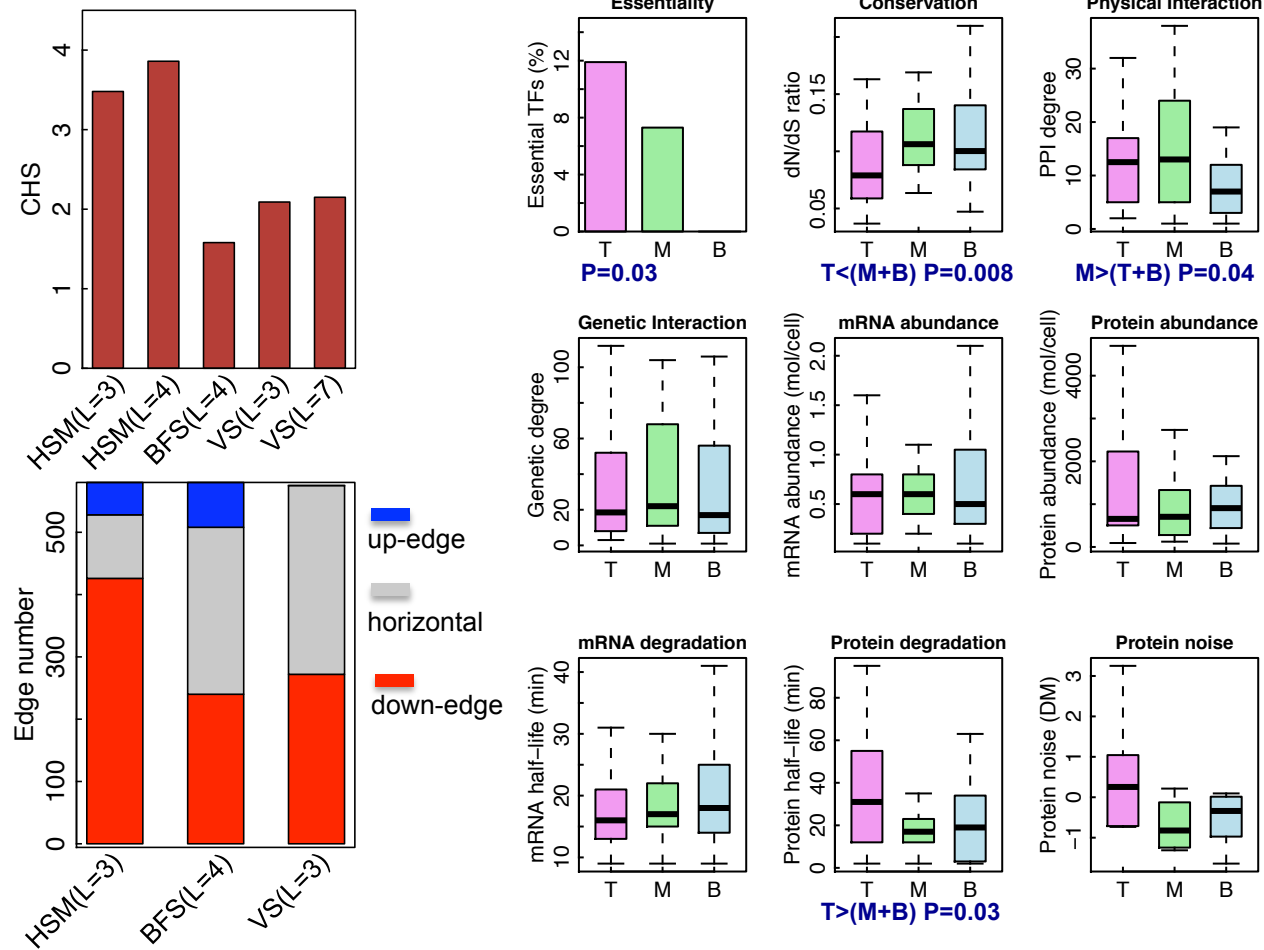
- Floyd, S. W. et al (1992)  
**Middle management involvement in strategy and its association with strategic type**  
*Strategic Management Journal* 13, 153-167.
- Woodward, J. (1982) *Industrial Organization: Theory and Practice* (Oxford University Press, Oxford).
- Floyd, S. W. et al (1993)  
**Dinosaurs or Dynamos?  
Recognizing Middle Management's Strategic Role**  
*The Academy of Management Executive* 8, 47-57.
- Floyd, S. W. et al (1997)  
**Middle management's strategic influence and organizational performance**  
*Journal of Management Studies* 34, 465-485.



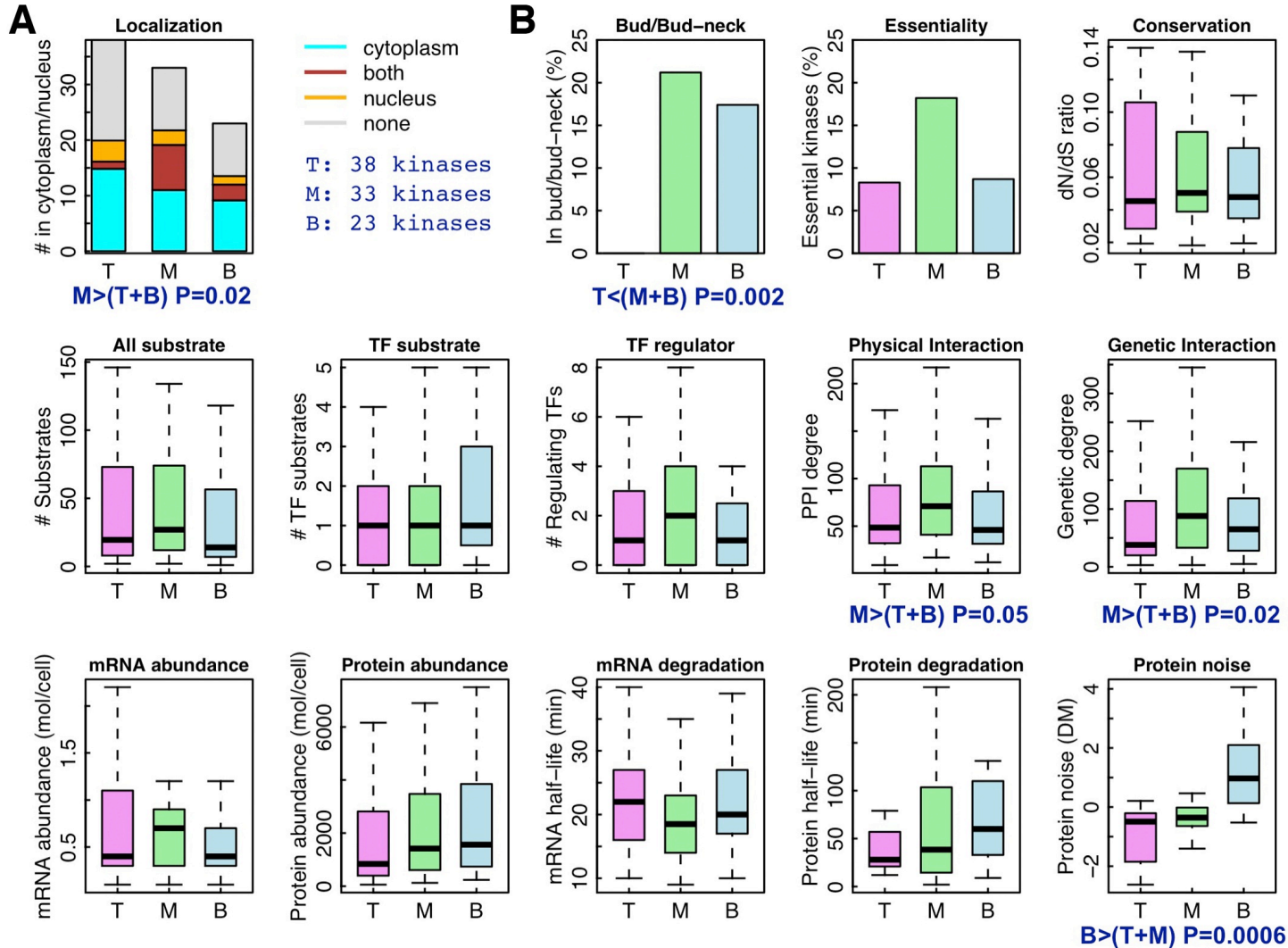
# Kinase network is more hierarchical than the TF reg. network

	#nodes	#edges	#levels	Reciprocity	Krackhardt	HS	CHS	PHS	Z-s
Worm neural	297	2359	4	0.184	0.186	2.80	2.364	2.707	
Political blogs	1224	19087	5	0.487	0.514	3.04	3.177	2.972	
Yeast TF	149	580	4	0.559	0.611	4.67	3.869	4.330	
Human TF	112	513	4	0.631	0.718	7.08	5.608	5.848	
P2P file sharing	6301	20777	4	0.486	0.772	4.34	5.878	2.401	
Foodweb	63	612	3	0.259	0.261	5.78	6.407	5.788	
Human Kinase	373	2171	7	0.492	0.798	14.08	13.396	12.874	
Yeast Kinase	94	200	5	0.645	0.775	17.45	13.982	11.777	

# Application: yeast TF regulatory network



# Application: yeast kinase network



# Biological Network Analysis: Hierarchies, Mutational Constraints & Logical Circuits in Regulation

- **Why Networks ?**

- Representational sweet spot + Intuition from cross-disciplinary comparisons

- **Organizing the Regulatory Network into a Hierarchy**

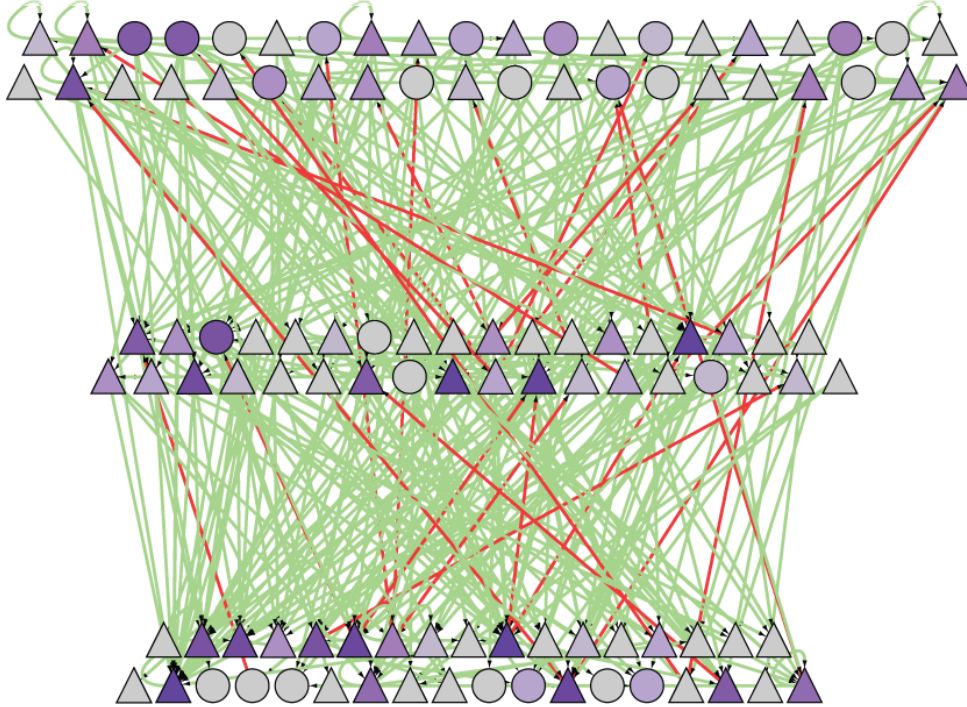
- Construction: local BFS v global simulated annealing
- Information flow bottlenecks in the middle & greater mid-level **collaboration** to ease them, esp. in more complex organisms
- Differences between kinase & TF hierarchy

- **Analyzing the Impact of Variation on the Network**

- Node Variation: **more connectivity = more constraint**
- Useful analogies to **designed systems**

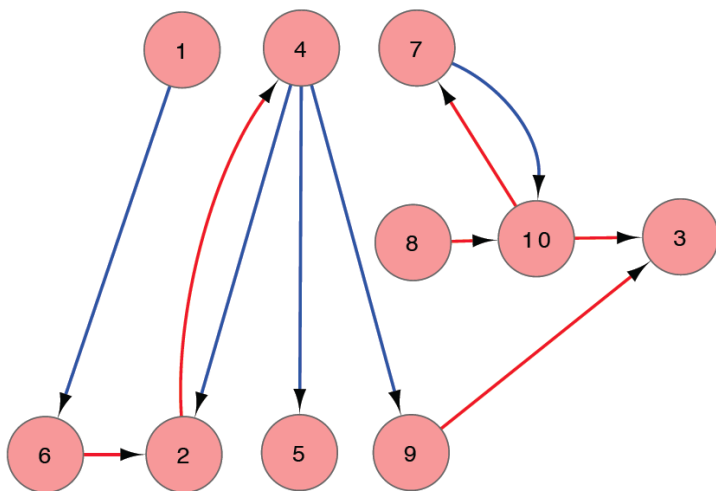
- **Going from Regulatory Networks to Logical Circuits**

- Preponderance of OR gates in the human network v yeast
- Relation to cancer (myc)
- More logical structure at top of hierarchy

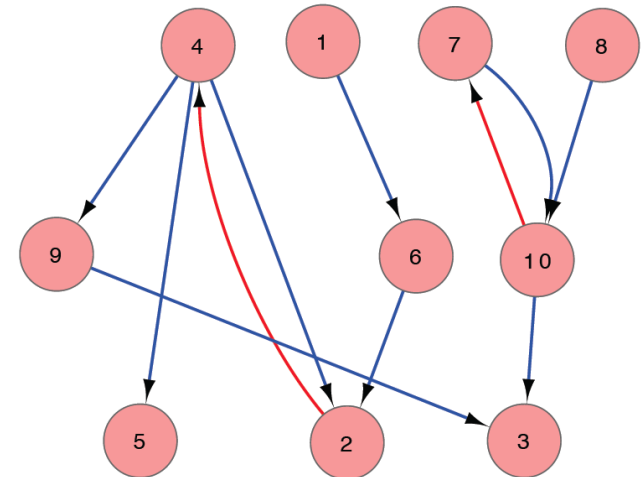


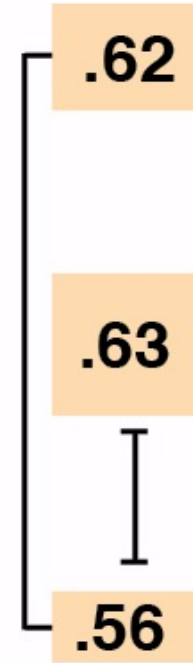
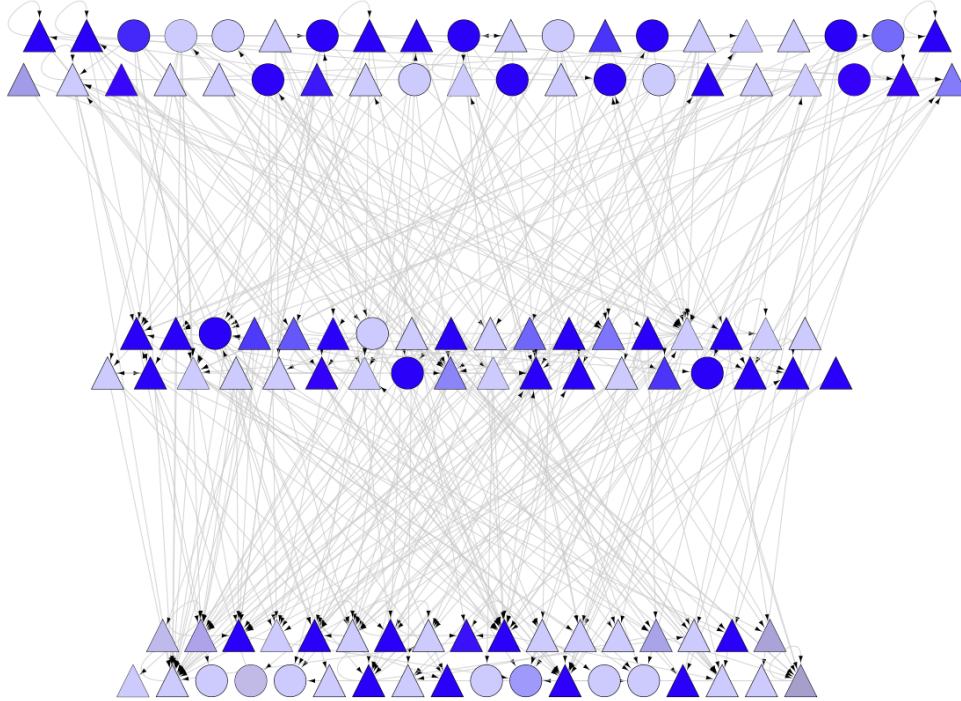
**Optimally arrange TFs into 3 levels by simulated annealing, maximizing downward-pointing edges**

Hierarchy height distribution approximated by 3 levels  
 Probing direction framed as an optimization problem

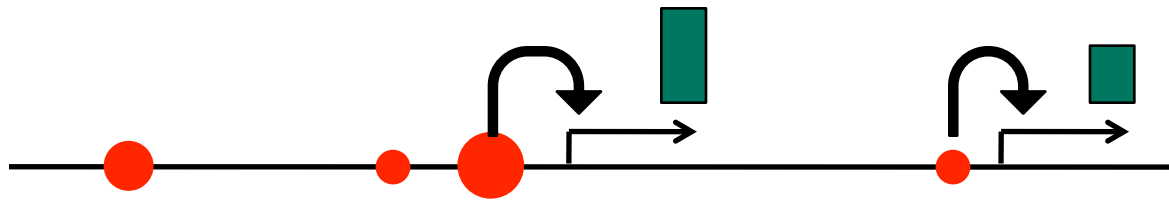


simulated annealing



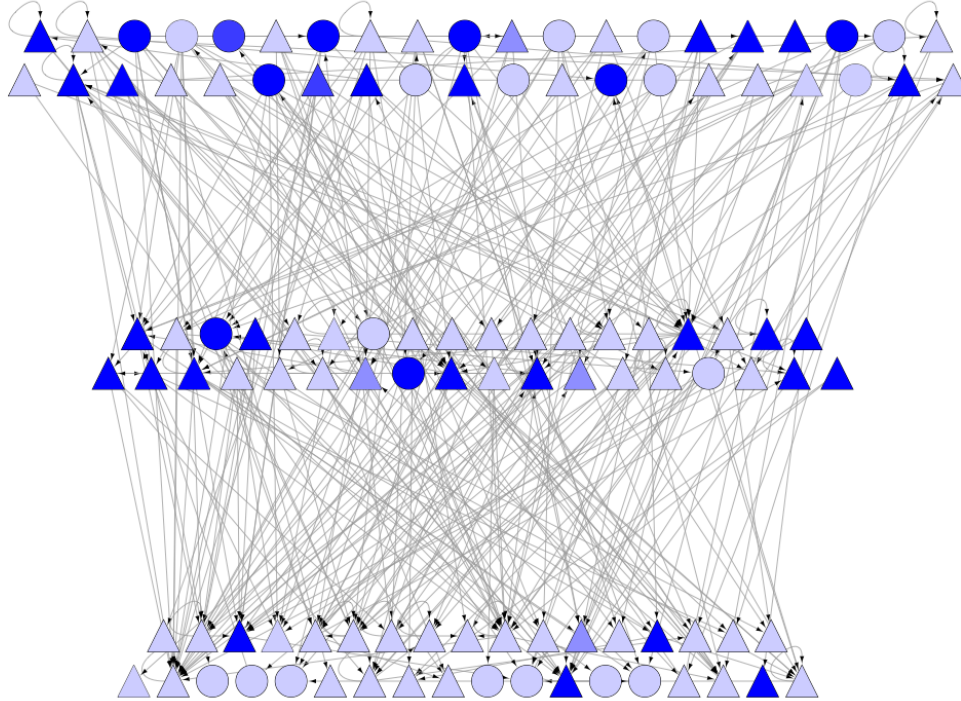


**Avg. correlation  
betw. binding  
signal of TF &  
gene expr. of its  
target**



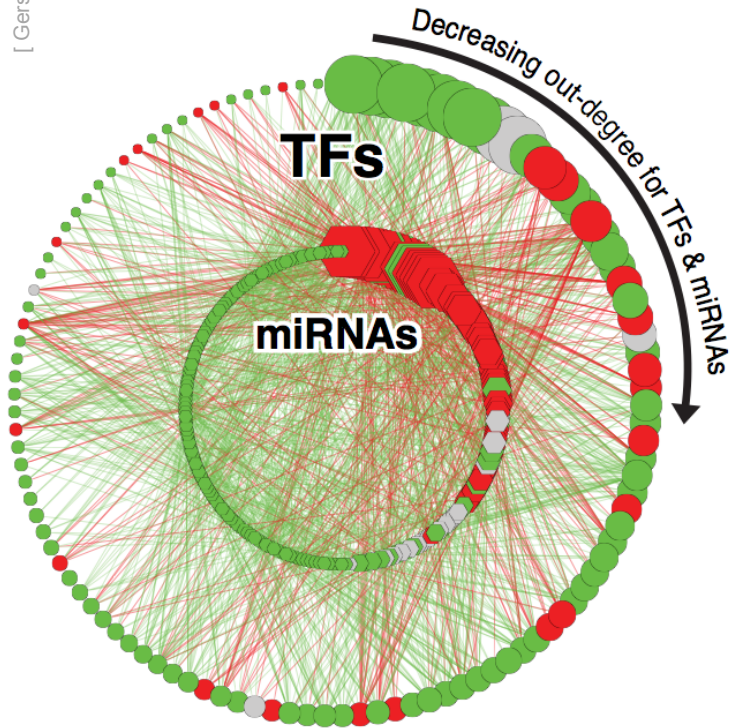
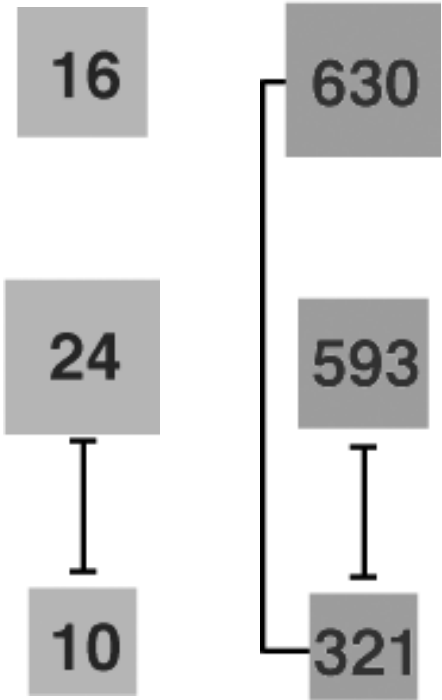
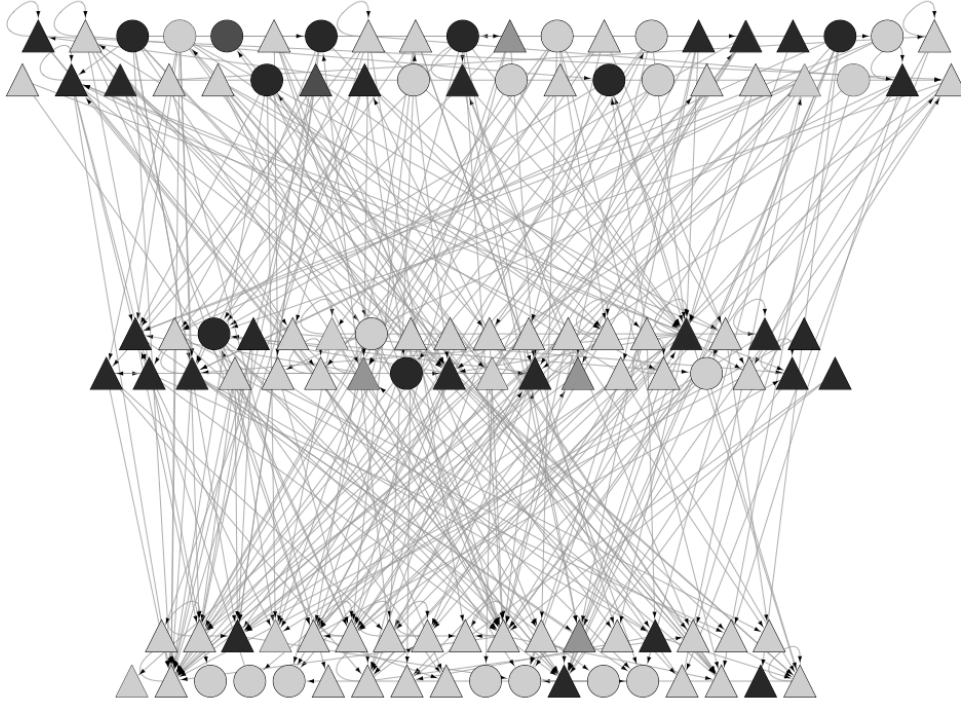
**Integration of TF hierarchy  
with other 'omic information :  
more influential & connected TFs on the top**





Avg. #  
of PPI for  
each TF

**Integration of TF hierarchy  
with other 'omic information :  
more influential & connected TFs on the top**

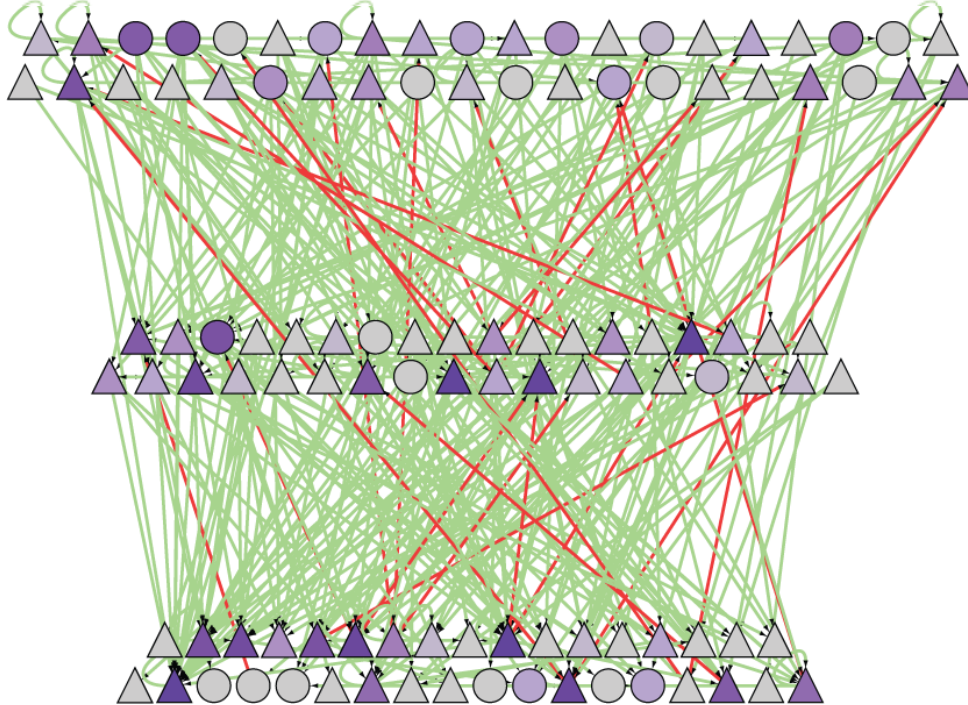


**Sig. corr. w/ TF  
hubbiness**  
(.24 & .62)

**Avg. values**

**# regulating miRNAs & # regulated miRNAs**

**Integration of TF hierarchy  
with other 'omic information :**  
more influential & connected TFs on the top



49



115



107

betweenness

**Strongest  
Proximal  
Regulatory  
Edges Can  
be Arranged  
into a  
Hierarchy**

Global wiring pattern of TFs

Middle level has highest betweenness, creating info. flow bottlenecks

# Biological Network Analysis: Hierarchies, Mutational Constraints & Logical Circuits in Regulation

- **Why Networks ?**

- Representational sweet spot + Intuition from cross-disciplinary comparisons

- **Organizing the Regulatory Network into a Hierarchy**

- Construction: local BFS v global simulated annealing
- Information flow bottlenecks in the middle & greater mid-level **collaboration** to ease them, esp. in more complex organisms
- Differences between kinase & TF hierarchy

- **Analyzing the Impact of Variation on the Network**

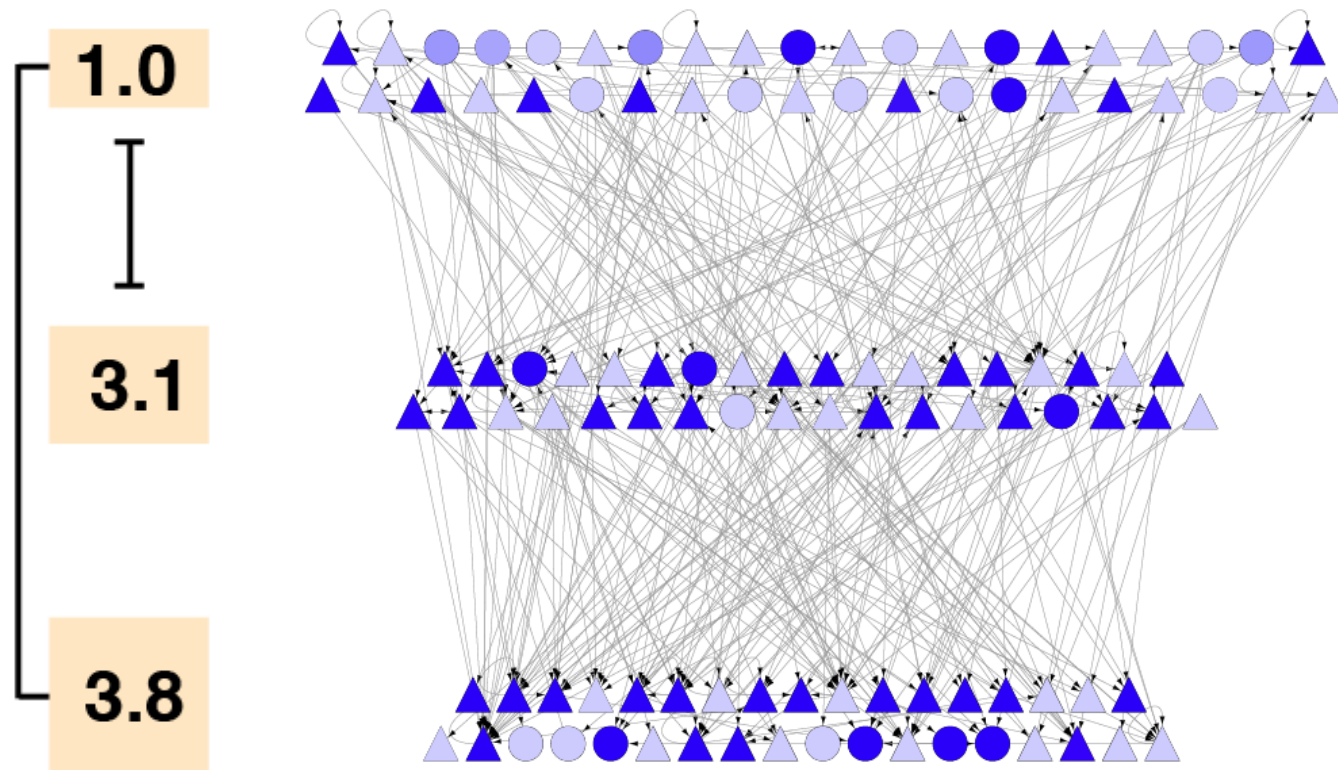
- Node Variation: **more connectivity = more constraint**
- Useful analogies to **designed systems**

- **Going from Regulatory Networks to Logical Circuits**

- Preponderance of OR gates in the human network v yeast
- Relation to cancer (myc)
- More logical structure at top of hierarchy

# TFs at the Top Under Stronger Negative Selection

SNP dens.  
 $\times 10^{-3}$



[ Gerstein et al. *Nature* (in press, '12) ]

# More connected components (“hubs”) have less variation

Integrate TFs & their binding sites with 1000G variation data & primate alignments (GERP score).

This shows:

TF **target in-degree**

**Neg. corr.** with

(SCC=-.2,  $P < 0.5$ )

dN/dS

(from chimp alignments)

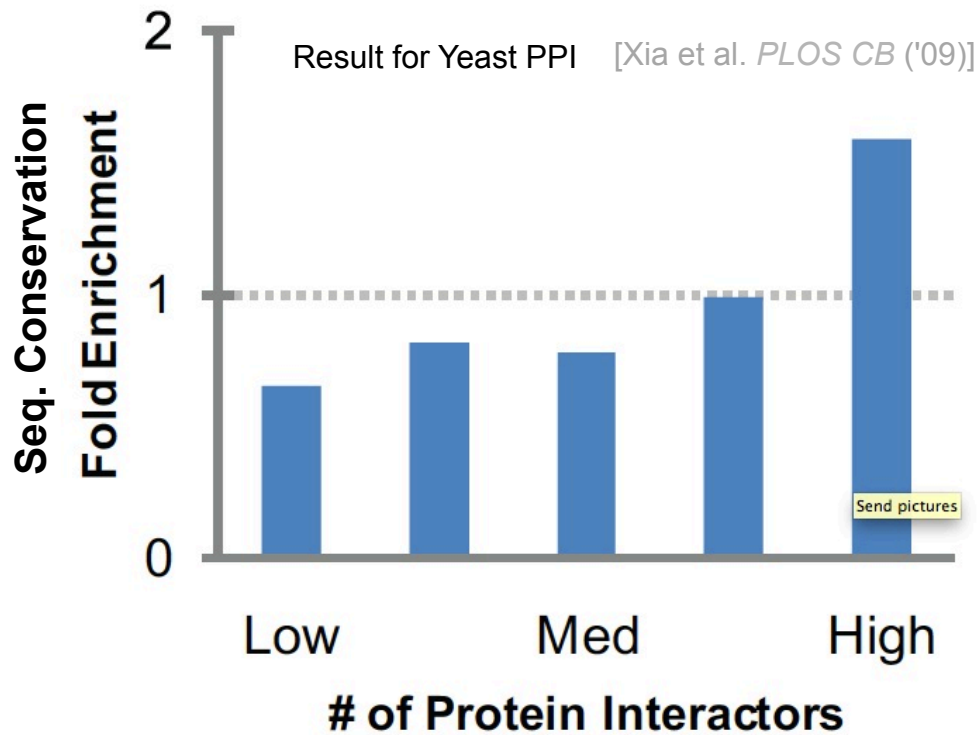
TF **target in-degree**

&

**TF out-degree**

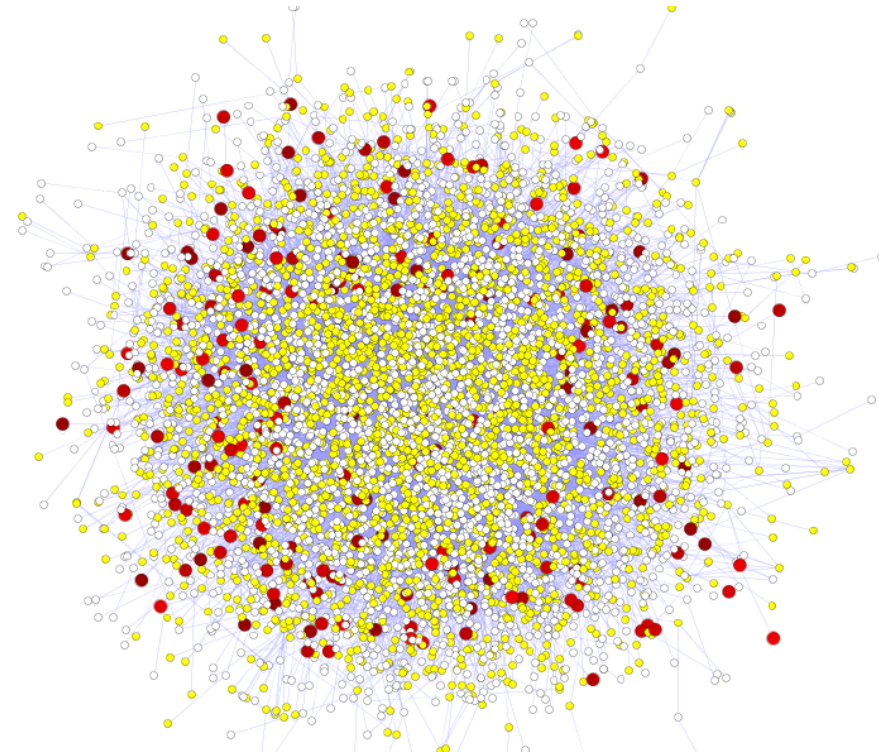
**Neg. corr.** with

ns SNP density, pN/pS, avg. DAF, frac rare SNPs



**More Connectivity,  
More Constraint : A  
theme borne out in  
many studies**

- High likelihood of positive selection
- Lower likelihood of positive selection
- Not under positive selection
- No data about positive selection



• **Sequence variation v. centrality**

- Nonsyn / synonymous SNPs v. deg. centrality
- $\rho = -.1, P < 4.0e-4$
- $\rho = -.3, P < 2.2.0e-16$
- (updated to 1000G phase I)

Proteins that have a more central position evolve more slowly and are more likely to be essential. This phenomenon is observed in many organisms and different kind of networks: Fraser et al. ('02) *Science*, ('03) *BMC Evo. Bio.* [yeast PPI]; Butland et al. ('04) *Nature* [*E. coli*. PPI]; Hahn et al. ('05), *MBE* [worm, fly PPI]; Cheng et al. ('09), *BMC Genomics* [miRNA nets]

# Biological Network Analysis: Hierarchies, Mutational Constraints & Logical Circuits in Regulation

- **Why Networks ?**

- Representational sweet spot + Intuition from cross-disciplinary comparisons

- **Organizing the Regulatory Network into a Hierarchy**

- Construction: local BFS v global simulated annealing
- Information flow bottlenecks in the middle & greater mid-level **collaboration** to ease them, esp. in more complex organisms
- Differences between kinase & TF hierarchy

- **Analyzing the Impact of Variation on the Network**

- Node Variation: **more connectivity = more constraint**
- Useful analogies to **designed systems**

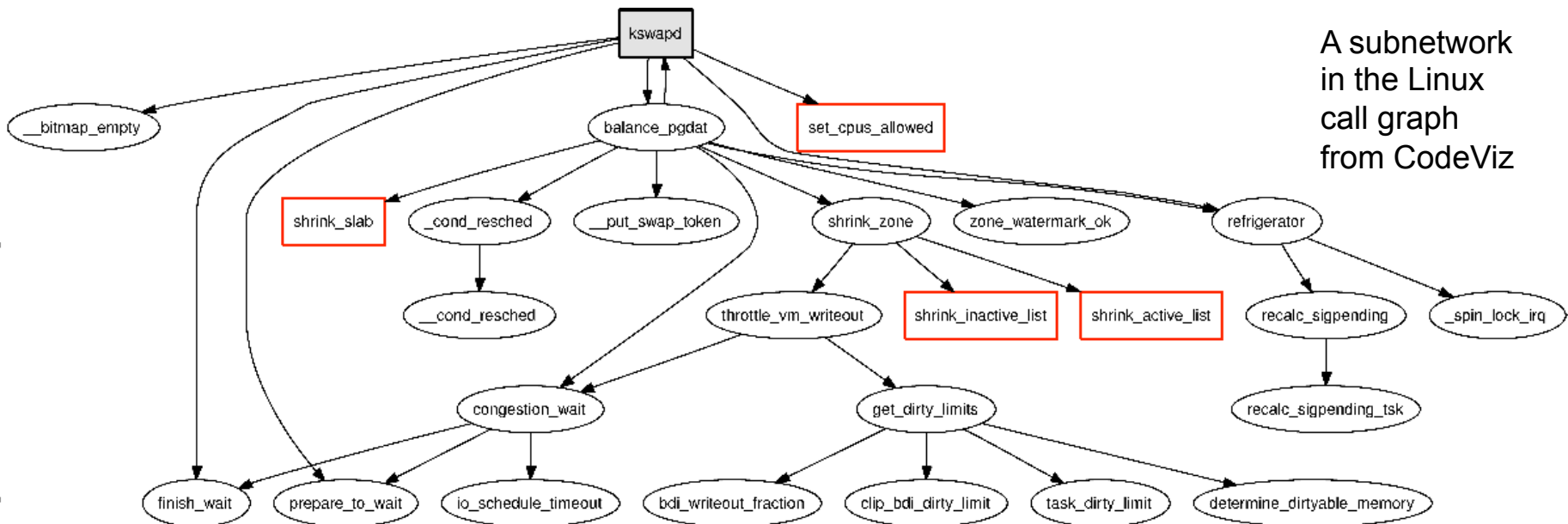
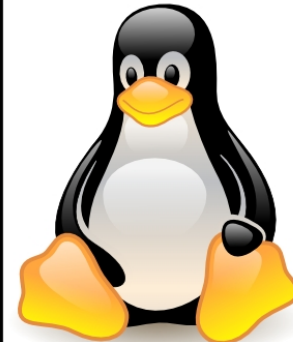
- **Going from Regulatory Networks to Logical Circuits**

- Preponderance of OR gates in the human network v yeast
- Relation to cancer (myc)
- More logical structure at top of hierarchy



# E. Coli Transcriptional regulatory network vs Linux call graph

		<i>E. coli</i> transcriptional regulatory network	Linux call graph
<b>Basic properties of systems</b>	Nodes	Genes (TFs & targets)	Functions (subroutines)
	Edges	Transcriptional regulation	Function calls
	External constraints	Natural environment	Hardware architecture, customer requirements
	Origin of evolutionary changes	Random mutation & natural selection	Designers' fine tuning



A subnetwork in the Linux call graph from CodeViz

E. coli transcriptional regulatory network

Linux call graph

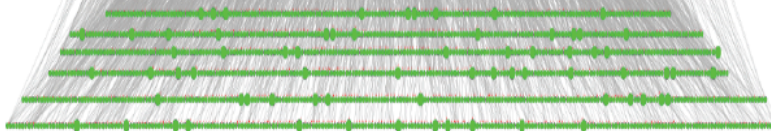
master regulator



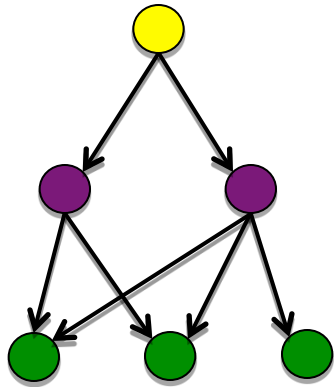
middle manager



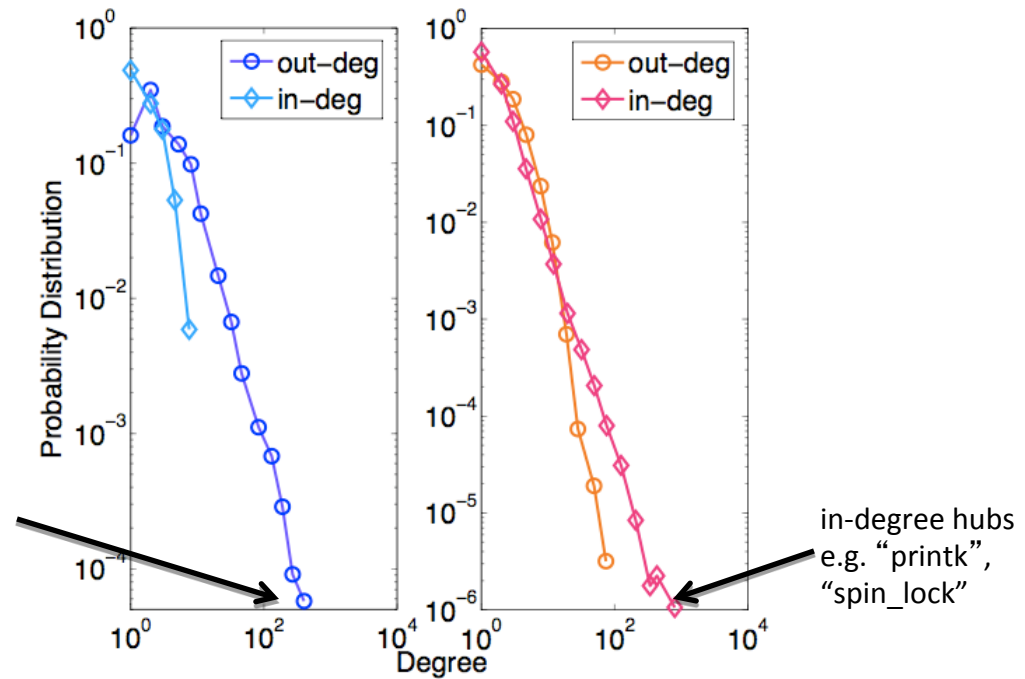
workhorse

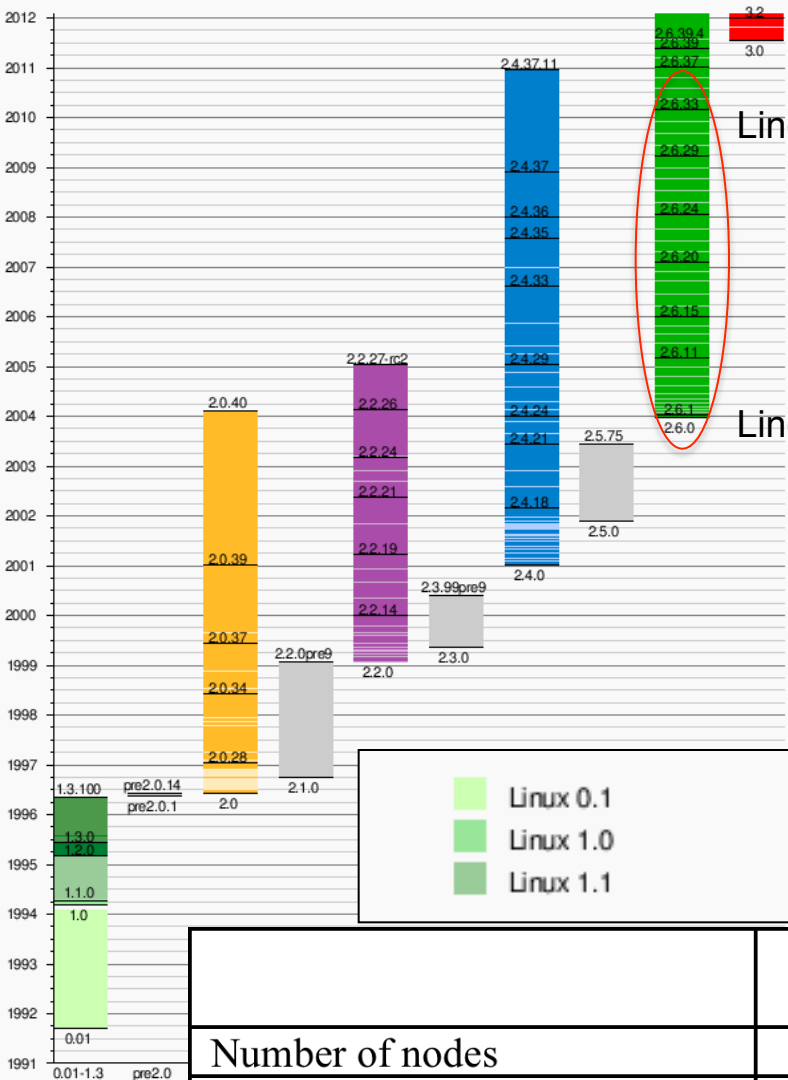


# Comparison: hierarchical organization



	% in <i>E. coli</i> regulatory network	% in Linux call graph
master regulator	4.6	29.6
middle manager	5.1	58.2
workhorse	90.2	12.3





## The Linux Kernel Evolves!

We can track the evolution of a function as the

“rate of evolution of a function”

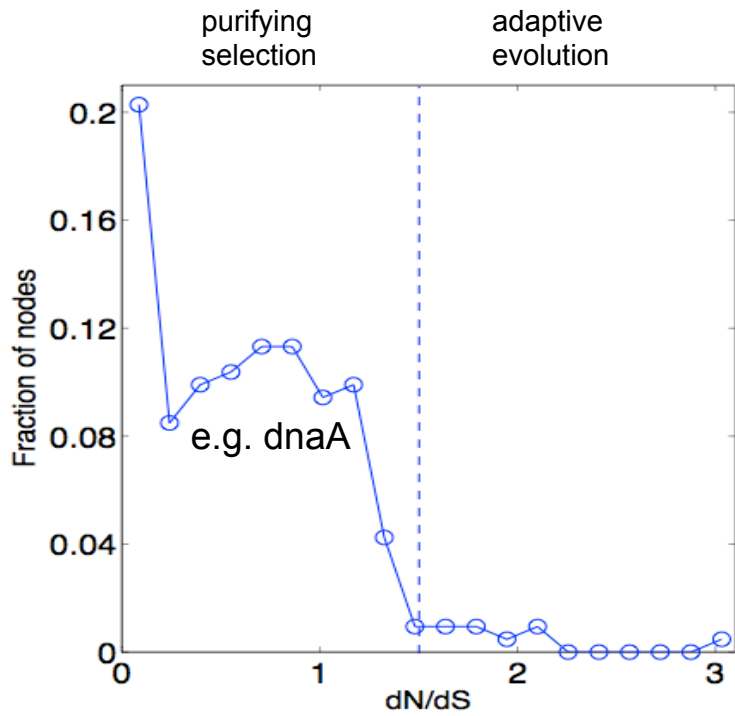
~  
the number of times it got revised



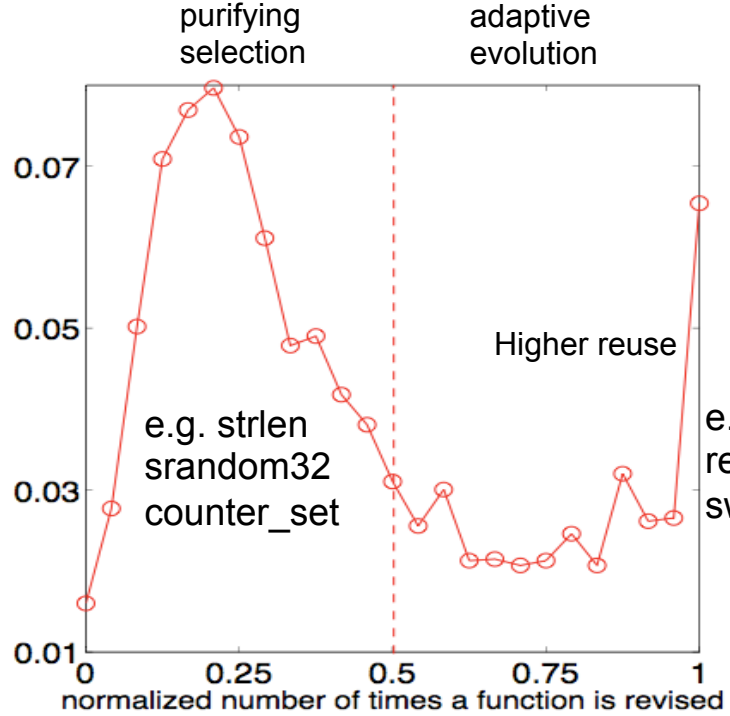
	<i>E. coli</i> transcriptional regulatory network	Linux call graph
Number of nodes	1378	12391
Number of persistent nodes	72* (5%)	5120 (41%)
Number of edges	2967	33553
Number of modules	64	3665
Number of comparative references	200 bacterial genomes	24 versions of kernels
Years of evolution	Billions years	20 years

# Distribution of Evolutionary Rates of components in E. coli vs Linux

*E. coli*



Persistent genes evolve slowly



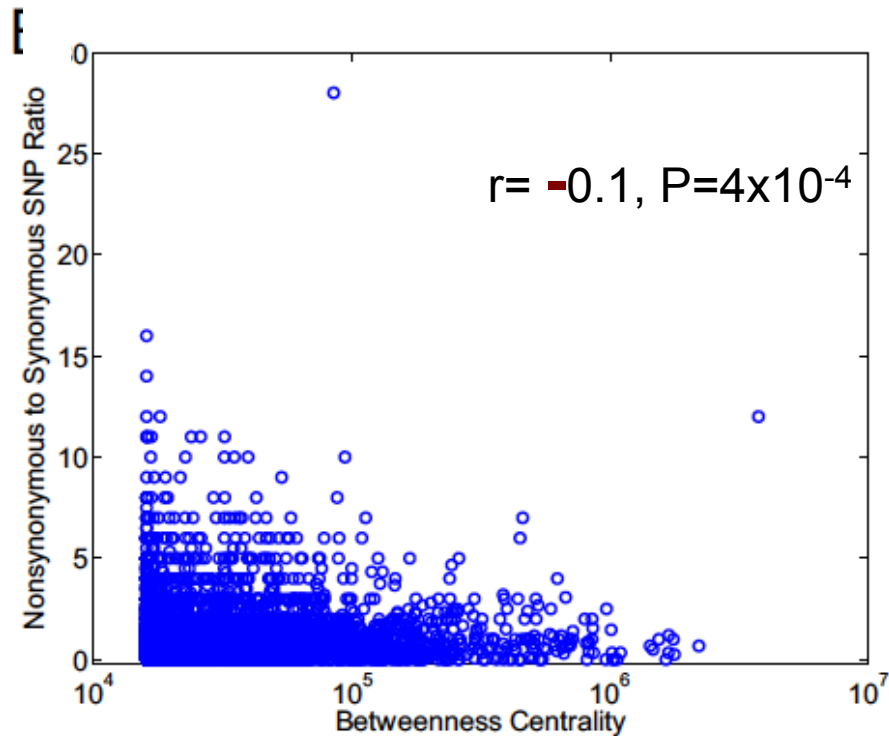
Linux

Two classes of persistent functions

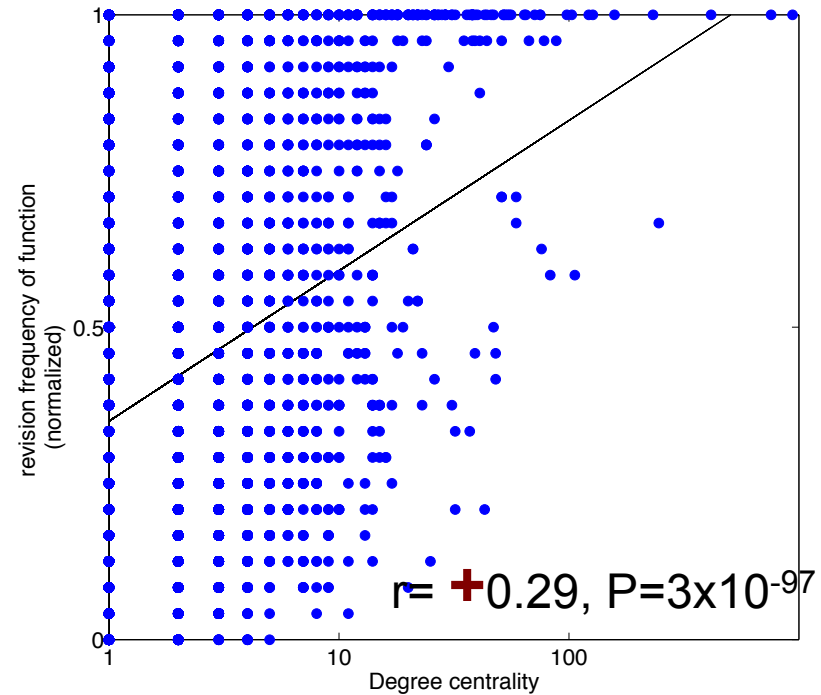
# Tinkering versus Design: Connectivity and Constraint

More Connectivity, More Constraint : A theme borne out in many evolutionary studies of biological network

Centrality is correlated with variation in technological systems



==Rate of Mutation==>



==Centrality==>

# Network of R package dependencies



## **mixtools: Tools for analyzing finite mixture models**

A collection of R functions for analyzing finite mixture models. This package is based upon under Grant No. SES-0518772.

Version: 1.0.2  
Depends: R ( $\geq 2.10.0$ ), [boot](#), [MASS](#), [segmented](#)  
Published: 2014-05-14  
Author: Derek Young [aut, cre], Tatiana Benaglia [aut], Didier Chauveau [aut], Hettmansperger [ctb], Hoben Thomas [ctb], Fengjuan Xuan [ctb]  
Maintainer: Derek Young <dsy109 at stat.psu.edu>  
License: [GPL-2](#) | [GPL-3](#) [expanded from: GPL ( $\geq 2$ )]  
NeedsCompilation: yes  
Citation: [mixtools citation info](#)  
Materials: [NEWS](#)  
In views: [Cluster](#), [Distributions](#)  
CRAN checks: [mixtools results](#)

### Downloads:

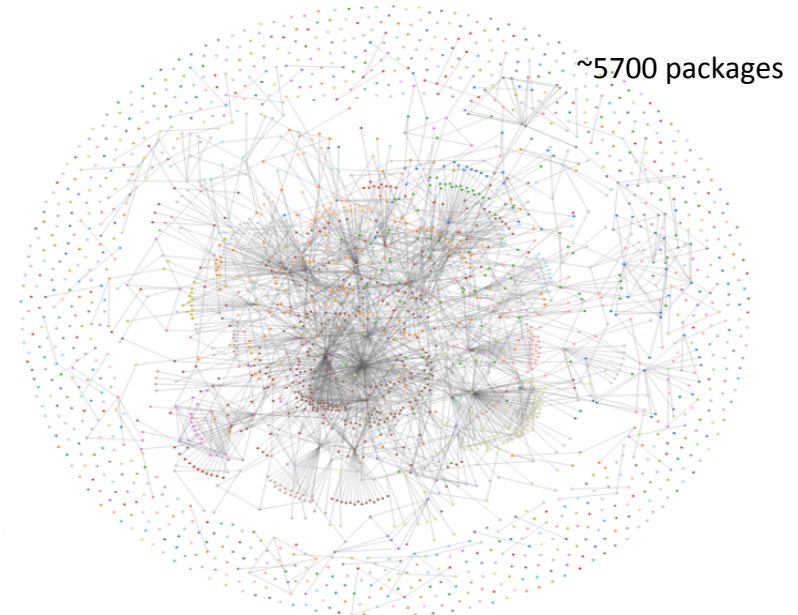
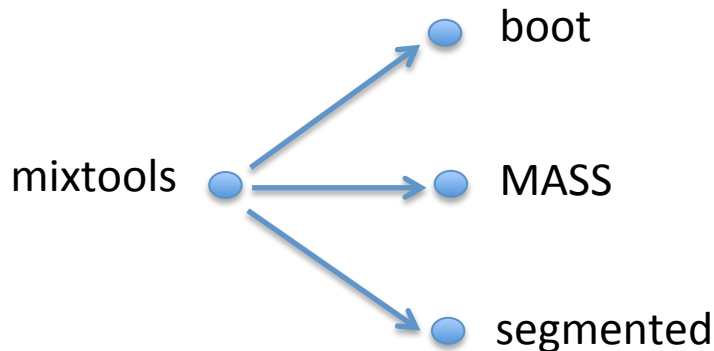


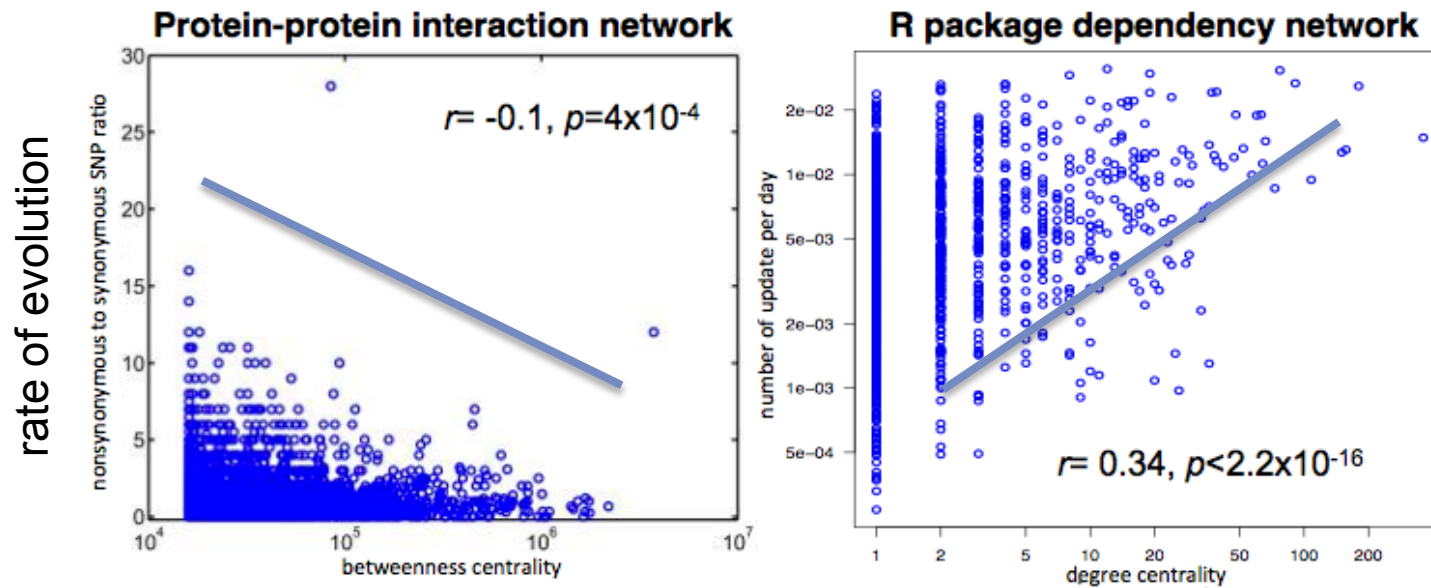
image: tlfvincent R Bloggers

[CRAN](#)  
[Mirrors](#)  
[What's new?](#)  
[Task Views](#)  
[Search](#)

[About R](#)  
[R Homepage](#)  
[The R Journal](#)

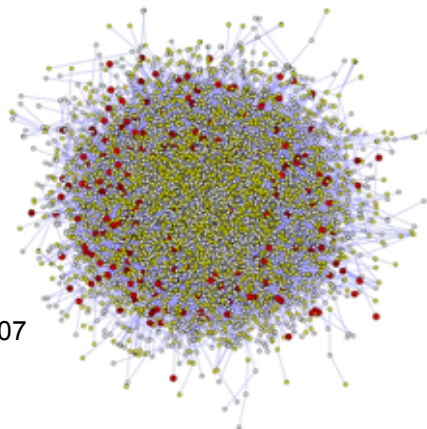
[Software](#)  
[R Sources](#)  
[R Binaries](#)  
[Packages](#)  
[Other](#)

# Tinkering versus Design: Connectivity and Constraint

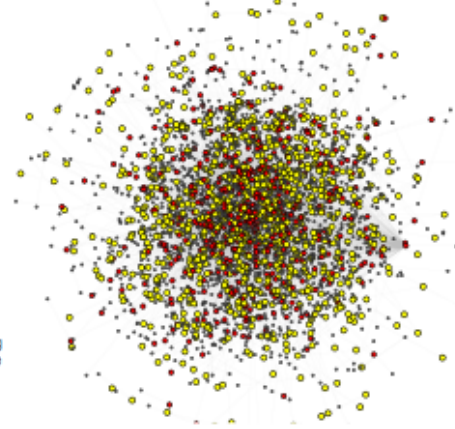


centrality

Kim et al. PNAS 2007



evolving protein { ● Fast  
● Slow  
○ Medial } updating package

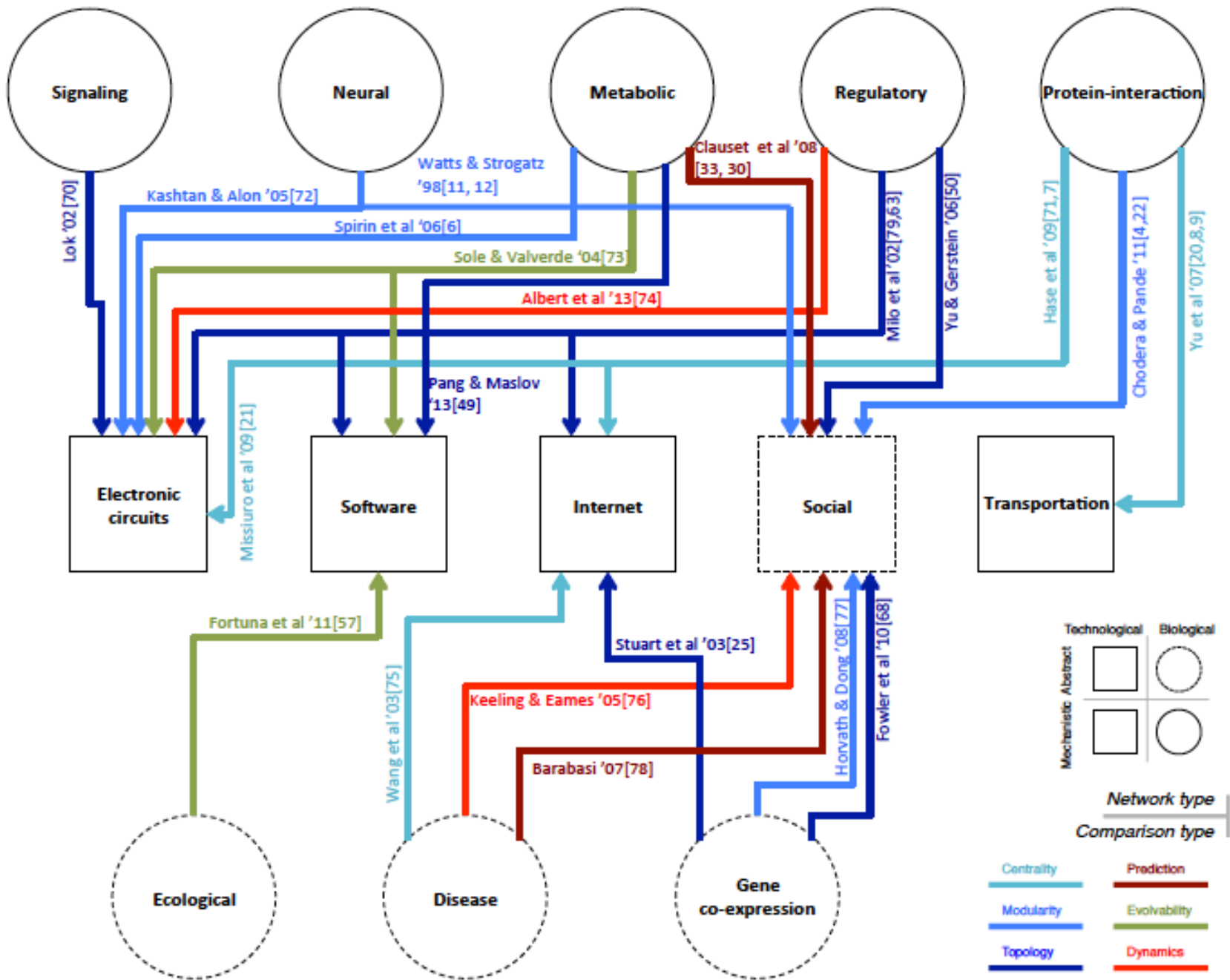




# Perspectives on Random Change v Intelligent Design

- Central points = hubs & bottlenecks
- If changes random, best not to put them in central pts.
- If changes made rationally, can put them into central pts.
  - Moreover, good to do this, as these more often used
    - i.e more efficient
  - Why there's so much GWB construction





[Yan et al., PLOS Biol. ('15, in revision)]

# Biological Network Analysis: Hierarchies, Mutational Constraints & Logical Circuits in Regulation

- **Why Networks ?**

- Representational sweet spot + Intuition from cross-disciplinary comparisons

- **Organizing the Regulatory Network into a Hierarchy**

- Construction: local BFS v global simulated annealing
- Information flow bottlenecks in the middle & greater mid-level **collaboration** to ease them, esp. in more complex organisms
- Differences between kinase & TF hierarchy

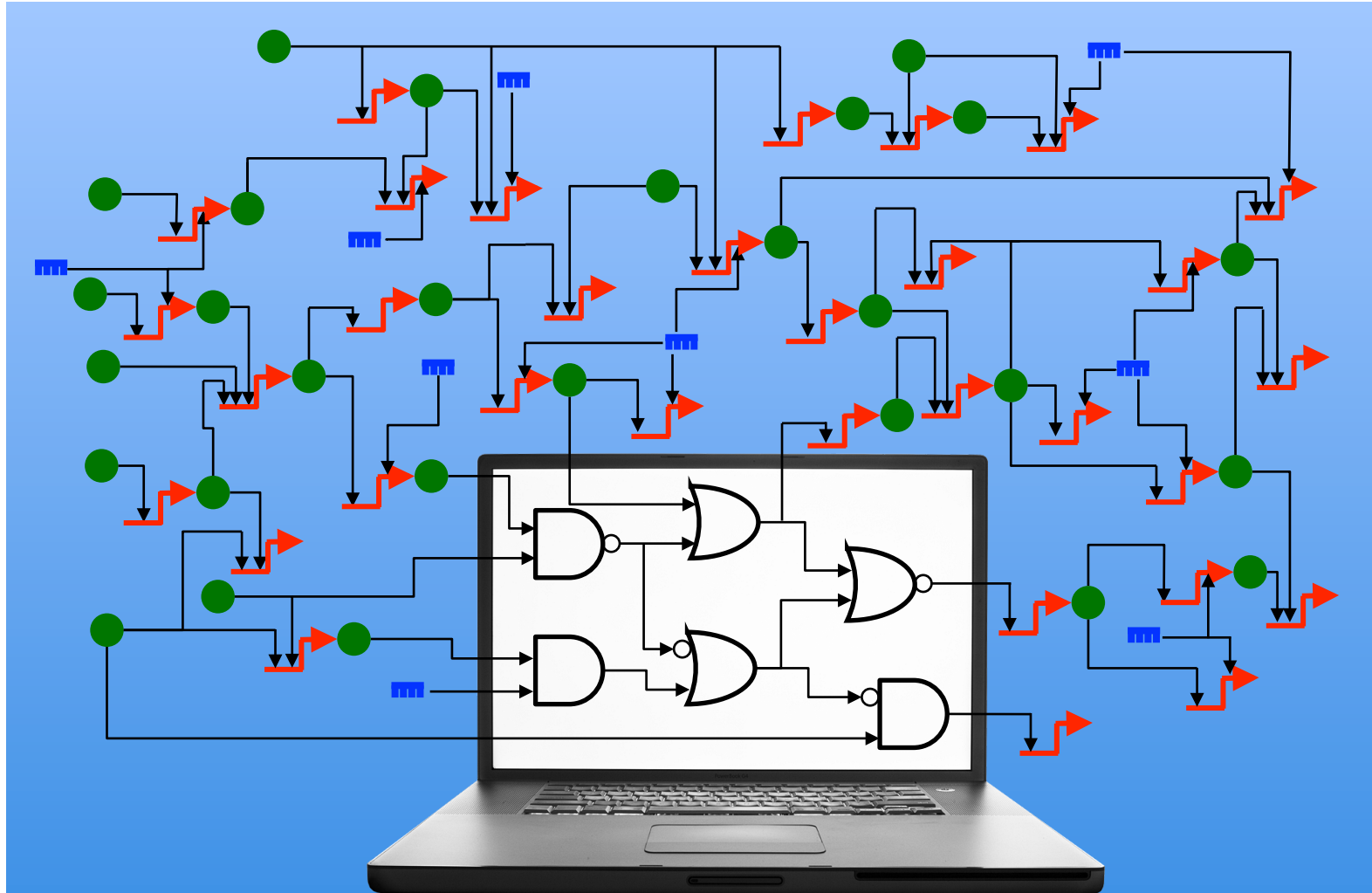
- **Analyzing the Impact of Variation on the Network**

- Node Variation: **more connectivity = more constraint**
- Useful analogies to **designed systems**

- **Going from Regulatory Networks to Logical Circuits**

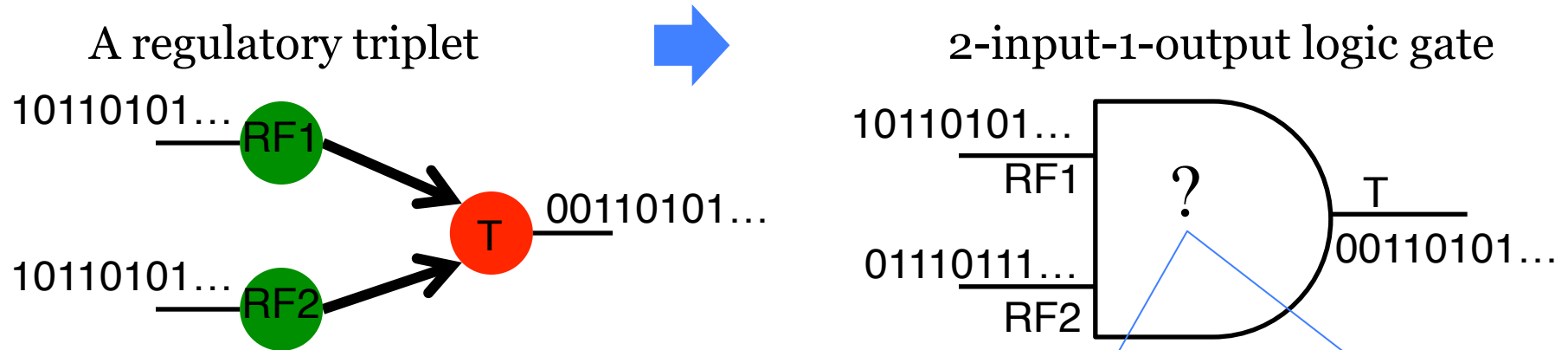
- Preponderance of OR gates in the human network v yeast
- Relation to cancer (myc)
- More logical structure at top of hierarchy

# Logic: A method to characterize the cooperative logic of regulatory factors



Wang, et al., *PLoS Computational Biology*, in press, 2015

# Modeling cooperativity between RFs to target gene using logic gates



0 – gene off  
 1 – gene on  
 after binarizing gene expression data\*

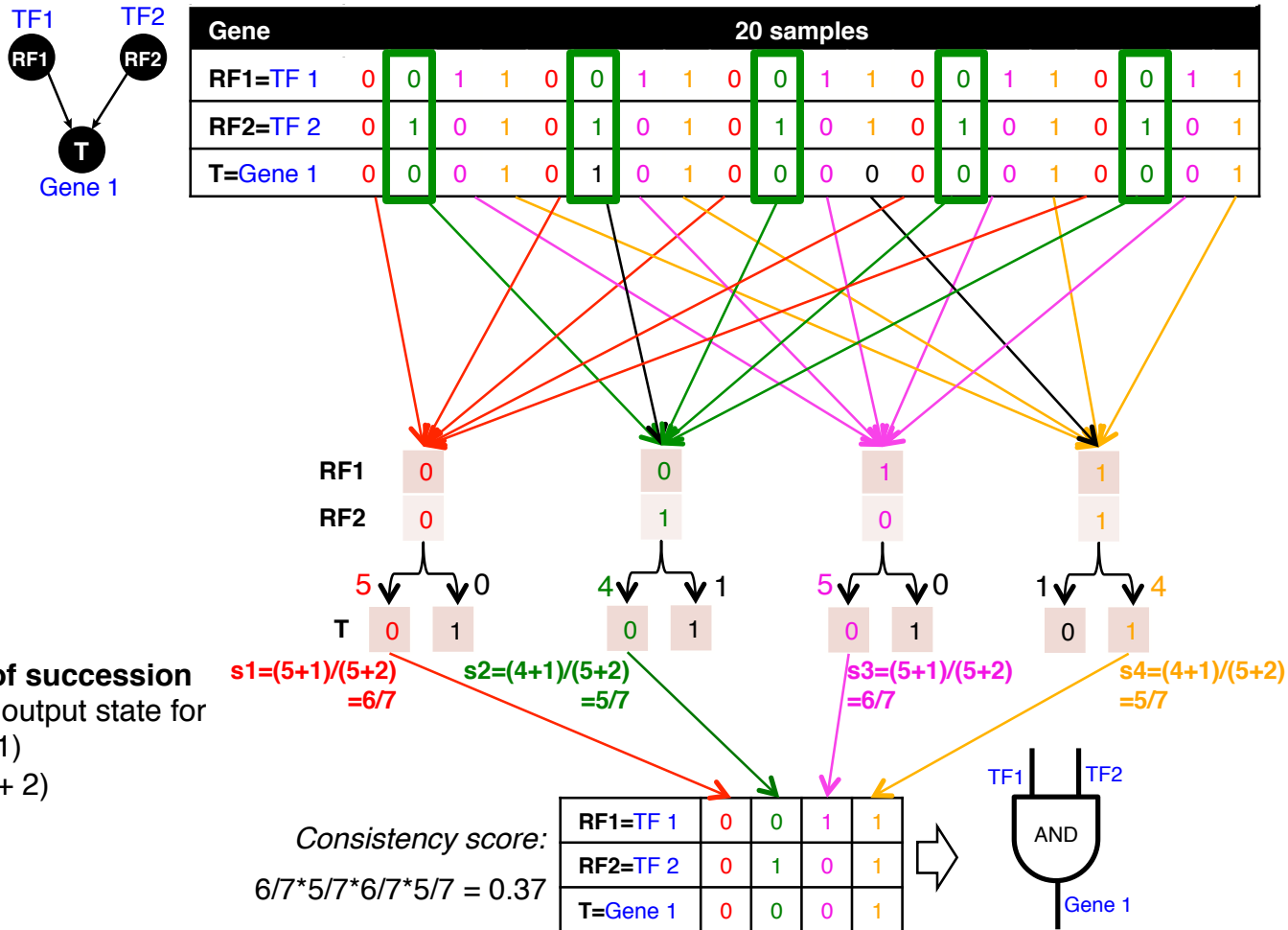
Input type (RF1, RF2)	RF1	0	0	1	1	} Binarized expression
	RF2	0	1	0	1	
Output	T	X	X	X	X	

X can be 0 or 1, so there are  $2^4=16$  possible output combinations, each of which corresponds to a unique 2-input-1-output logic gate

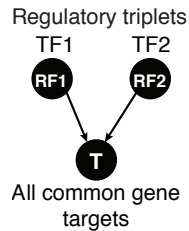


\*BoolNet, R package

# An example: selection of the best-matched logic gate



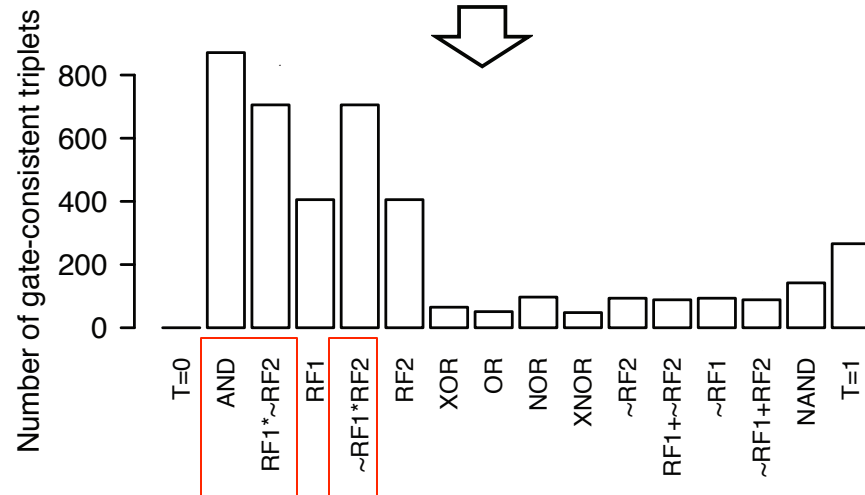
# Application 1 – transcription factor cooperativity in Yeast cell cycle



Target gene	2464
TF	176
Triplet	39,011
Time point	59

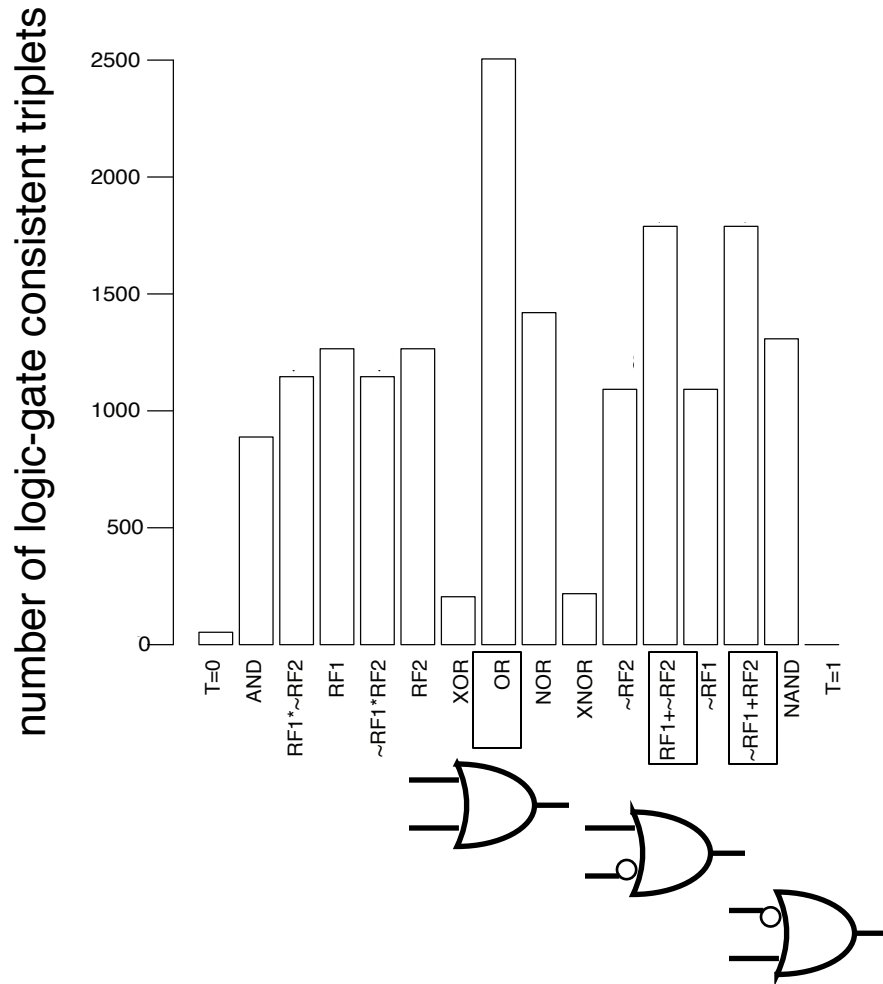
Yeast Cell Cycle

Triplet ID	RF1	RF2	Common Target Gene (T)	Matched logic gate
1	YHR084W	YBR083W	YBR082C	AND
2	YKL112W	YIL131C	YMR198W	OR
...	...	...	...	...
39011	YOR113W	YBL103C	YDR042C	XOR



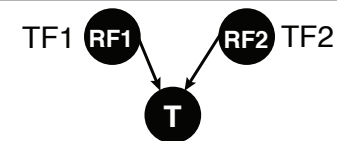
Wang, et al., *PLoS Computational Biology*, 2015

# Application – transcription factor cooperativity in Acute Myeloid Leukemia (AML)



## Human TF-TF-target

RF1	RF2	Common Target Gene (T)	Matched logic gate
ATF3	BDP1	YPEL1	AND
MYC	BCL3	BCR	T=RF1
ATF3	BRF2	AIF1L	AND
...	...	...	...



All common gene targets

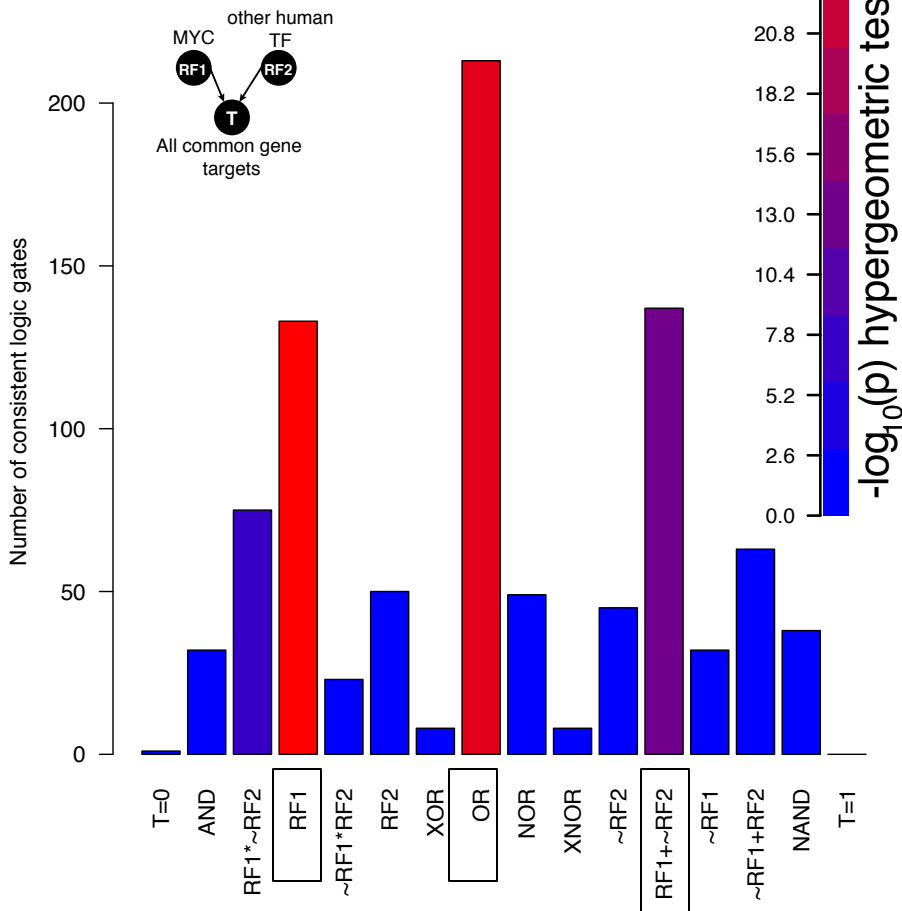
Target gene	1824
TF	70
Triplet	50,865
Patient	197

Wang, et al., *PLoS Computational Biology*, in press, 2015



# Cancer-related TF, MYC universally amplifies target expression

2,153 (RF1=MYC, RF2=other TFs, T=all common targets) triplets



- RF1
- **OR**(RF1, RF2)
- **OR**(RF1, **NOT** RF2)



High expression of MYC is sufficient for high target gene expression

**c-Myc Is a Universal Amplifier of Expressed Genes in Lymphocytes and Embryonic Stem Cells**



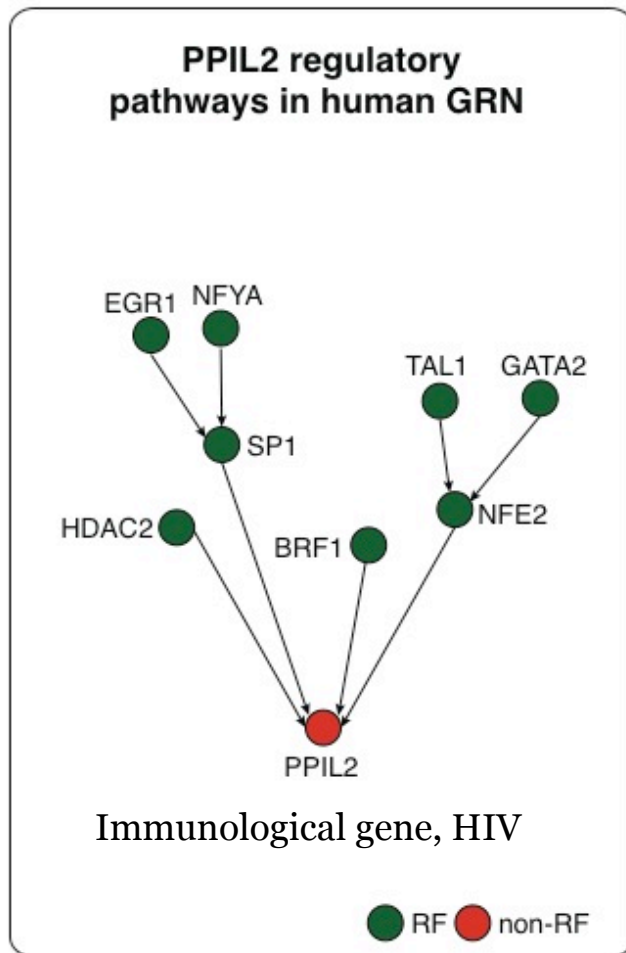
Zuqin Nie,<sup>1,6</sup> Gangqing Hu,<sup>2,6</sup> Gang Wei,<sup>2</sup> Kairong Cui,<sup>2</sup> Arito Yamane,<sup>3</sup> Wolfgang Resch,<sup>3</sup> Ruoning Wang,<sup>4</sup> Douglas R. Green,<sup>4</sup> Lino Tessarollo,<sup>5</sup> Rafael Casellas,<sup>3</sup> Keji Zhao,<sup>2,\*</sup> and David Levens<sup>1,\*</sup>



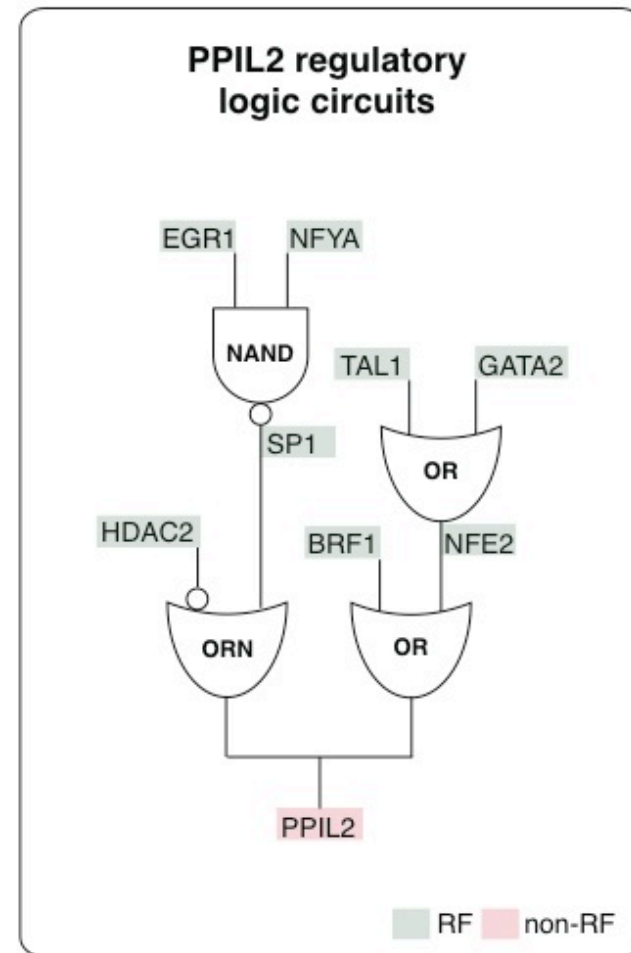
Predict outcomes of genome engineering in leukemia

Wang, et al., *PLoS Computational Biology*, in press, 2015

# Gene regulatory pathways have logic-circuit behaviors



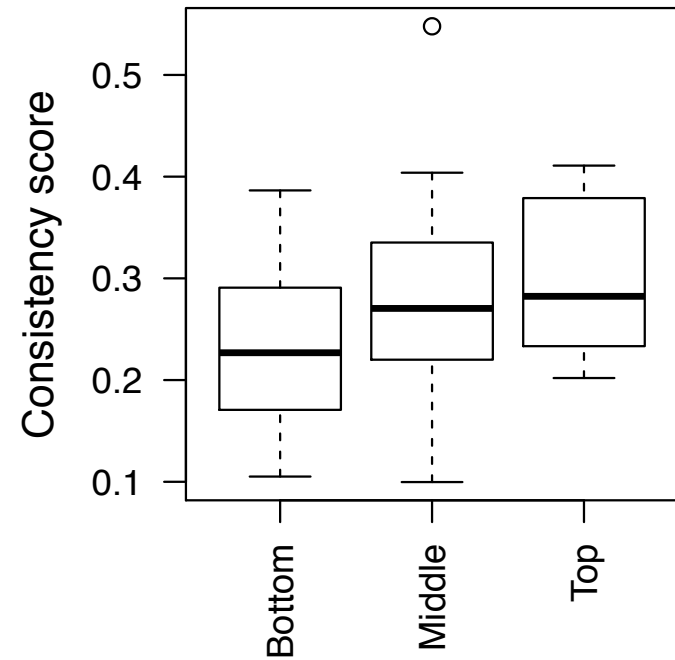
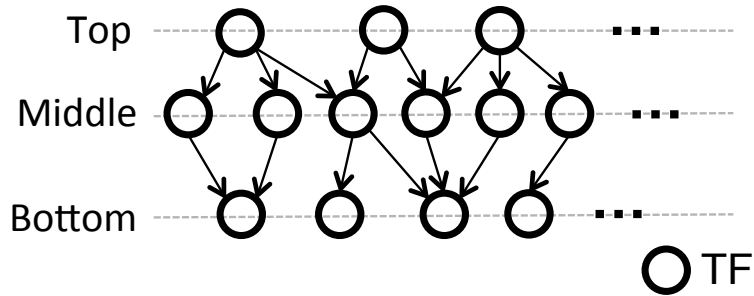
**Logic** →



Wang, et al., *PLoS Computational Biology*, in press, 2015

# Logical cooperativity across hierarchical layers in gene regulatory network

Hierarchical gene regulatory network



The regulations of middle and top TFs more likely follow logical operations than the bottom TFs.

# Biological Network Analysis: Hierarchies, Mutational Constraints & Logical Circuits in Regulation

- **Why Networks ?**

- Representational sweet spot + Intuition from cross-disciplinary comparisons

- **Organizing the Regulatory Network into a Hierarchy**

- Construction: local BFS v global simulated annealing
- Information flow bottlenecks in the middle & greater mid-level **collaboration** to ease them, esp. in more complex organisms
- Differences between kinase & TF hierarchy

- **Analyzing the Impact of Variation on the Network**

- Node Variation: **more connectivity = more constraint**
- Useful analogies to **designed systems**

- **Going from Regulatory Networks to Logical Circuits**

- Preponderance of OR gates in the human network v yeast
- Relation to cancer (myc)
- More logical structure at top of hierarchy



- an automated web tool

**tYNA**

(vers. 2 :

"TopNet-like

Yale Network Analyzer")

ID	Name	Creator	Creation date	Delete
14	Uetz 2000 yeast two hybrid	kevin	21-Feb-06	Delete
15	Ito 2001 yeast two hybrid	kevin	21-Feb-06	Delete
16	Ho 2002 pull down	kevin	21-Feb-06	Delete
17	Gavin 2002 pull down	kevin	21-Feb-06	Delete
18	Jansen 2003 PIT	kevin	21-Feb-06	Delete
19	MIPS yeast PPI	kevin	21-Feb-06	Delete
21	BIND yeast data	kevin	21-Feb-06	Delete
22	DIP yeast data	kevin	21-Feb-06	Delete
23	Kim 2006 structural interaction	kevin	21-Feb-06	Delete
24	Han 2004 FYI data	kevin	21-Feb-06	Delete
25	Luscombe 2004 regulatory	kevin	21-Feb-06	Delete

FILES

- enes24.figure54b.xgml
- enes25.figure2a\_proximal.xgml
- enes25.figure2a\_distal.xgml
- enes26.figure54c.xgml
- enes27.figure5b.xgml
- enes28.figure54d.xgml
- enes29.figure5a.xgml

Info  
Click a node or edge.  
Download file

**Encodenets.gersteinlab.org, networks.gersteinlab.org**

(tYNA: Normal website + Downloaded code (JAVA) + Web service (SOAP) with Cytoscape plugin)

[Yu et al., NAR (2004); Yip et al. Bioinfo. (2006);

Similar tools include Cytoscape.org, Idekar, Sander et al]



(11 Main Projects, ~50 labs, >700 substantial contributors + NHGRI)

## Network Acknowledgements

Networks/Elements (~60 participants):



**A Kundaje, M Hariharan, S Landt, K Yan, C Cheng, X Mu, E Khurana, J Rozowsky, R Alexander, R Min, P Alves, A Abyzov, N Addleman, N Bhardwaj, A Boyle, P Cayting, A Charos, D Chen, Y Cheng, D Clarke, C Eastman, G Euskirchen, S Fietze, Y Fu, J Gertz, F Grubert, A Harmanci, P Jain, M Kasowski, P Lacroute, J Leng, J Lian, H Monahan, H O'Geen, Z Ouyang, E Partridge, D Patacsil, F Pauli, D Raha, L Ramirez, T Reddy, B Reed, M Shi, T Slifer, J Wang, L Wu, X Yang, K Yip, G Zilberman-Schapira, S Batzoglou, A Sidow, P Farnham, R Myers, S Weissman, M Snyder**

Hiring Postdocs. See [gersteinlab.org/jobs](http://gersteinlab.org/jobs) !

- Hierarchy Construction –  
Chao **Cheng**, Erik Andrews,  
Koon-Kiu Yan, Matthew Ung, Daifeng  
Wang

[papers.gersteinlab.org/papers/hinet coming soon]

- Loregic –  
Daifeng **Wang**, Koon-Kiu Yan,  
Cristina Sisu, Chao Cheng, Joel  
Rozowsky, William Meyerson

[papers.gersteinlab.org/papers/loregic coming soon]

- Net hierarchy analysis –  
N **Bhardwaj**, KK Yan

[papers.gersteinlab.org/netheir]

- Call graph analysis –  
KK **Yan**, G Fang, N Bhardwaj, RP  
Alexander

[papers.gersteinlab.org/callgraph]

- Network Sci. Analogies –  
Koon-Kiu **Yan**, Daifeng Wang,  
Anurag Sethi, Paul Muir, Robert  
Kitchen, Chao Cheng

[papers.gersteinlab.org/papers/netessay coming soon]

# Acknowledgements



Hiring Postdocs. See [gersteinlab.org/jobs](http://gersteinlab.org/jobs) !

# Info about content in this slide pack

- General PERMISSIONS
  - This Presentation is copyright Mark Gerstein, Yale University, 2014.
  - Please read permissions statement at <http://www.gersteinlab.org/misc/permissions.html> .
  - Feel free to use slides & images in the talk with PROPER acknowledgement (via citation to relevant papers or link to gersteinlab.org).
  - Paper references in the talk were mostly from Papers.GersteinLab.org.
- For SeqUniverse slide, please contact Heidi Sofia, NHGRI
- PHOTOS & IMAGES. For thoughts on the source and permissions of many of the photos and clipped images in this presentation see <http://streams.gerstein.info> .
  - In particular, many of the images have particular EXIF tags, such as kwpotppt , that can be easily queried from flickr, viz: <http://www.flickr.com/photos/mbgmbg/tags/kwpotppt>