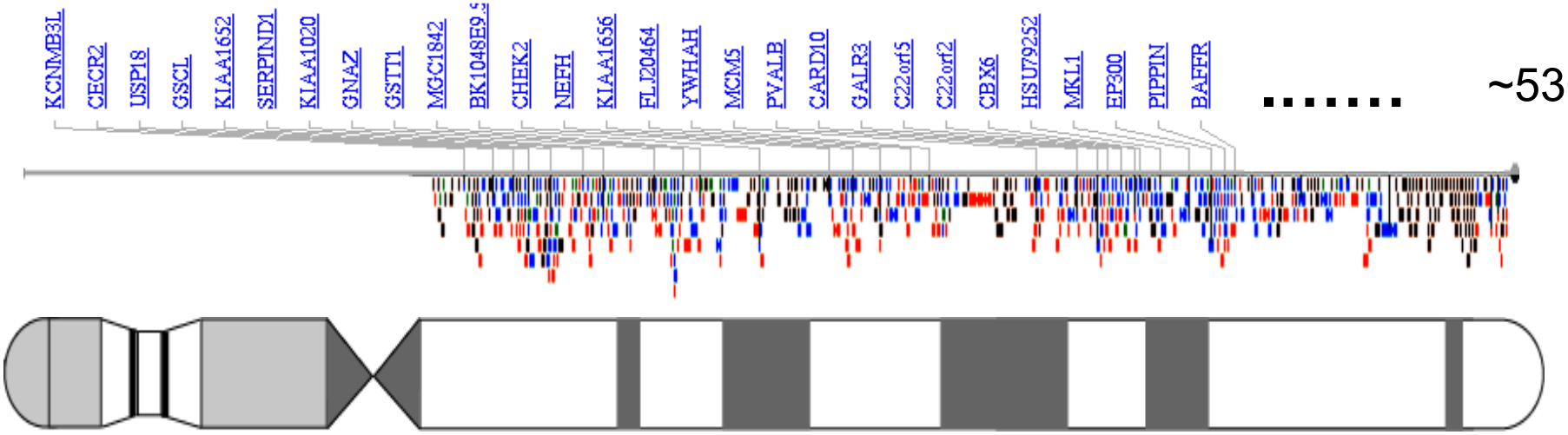
Understanding Protein Function on a Genome-scale through the Analysis of Molecular Networks

Mark B Gerstein
Yale

slides at
Lectures.GersteinLab.org

(See Last Slide for References & More Info.)

The problem: Grappling with Function on a Genome Scale?



- 250 of ~530
  originally characterized on chr. 22
  [Dunham et al. Nature (1999)]

- >25K Proteins in Entire Human Genome
  (with alt. splicing)

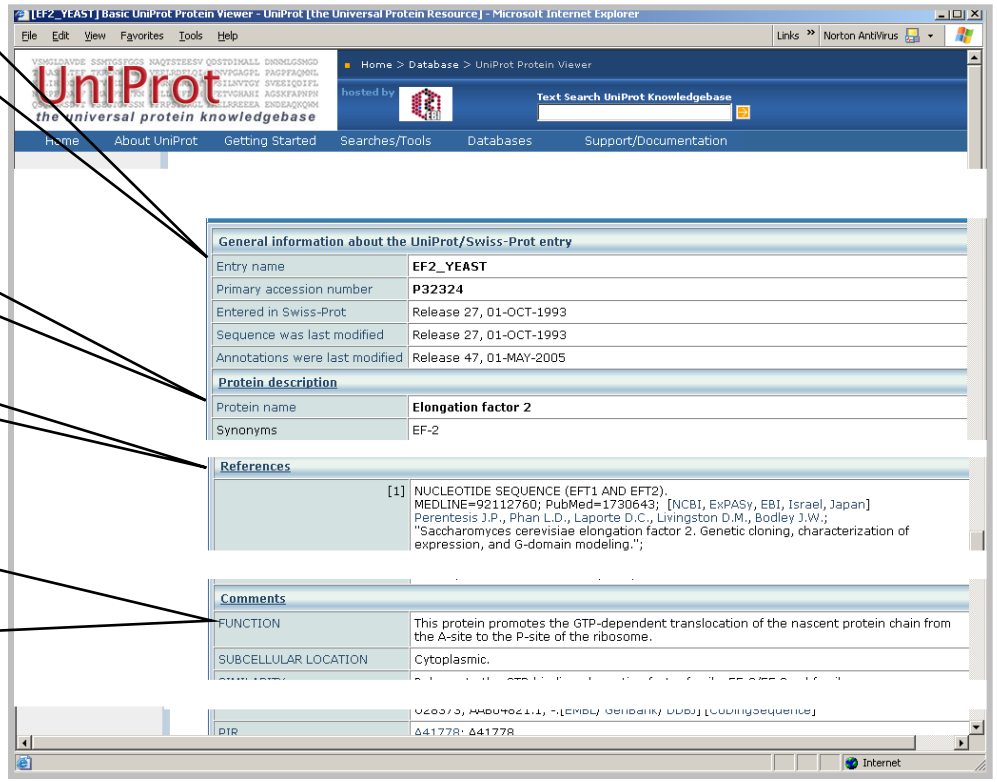
Traditional single molecule way to integrate evidence & describe function

EF2_YEAST

Descriptive Name:
Elongation Factor 2

Lots of references
to papers

Summary sentence describing function:
This protein promotes the GTP-dependent translocation of the nascent protein chain from the A-site to the P-site of the ribosome.

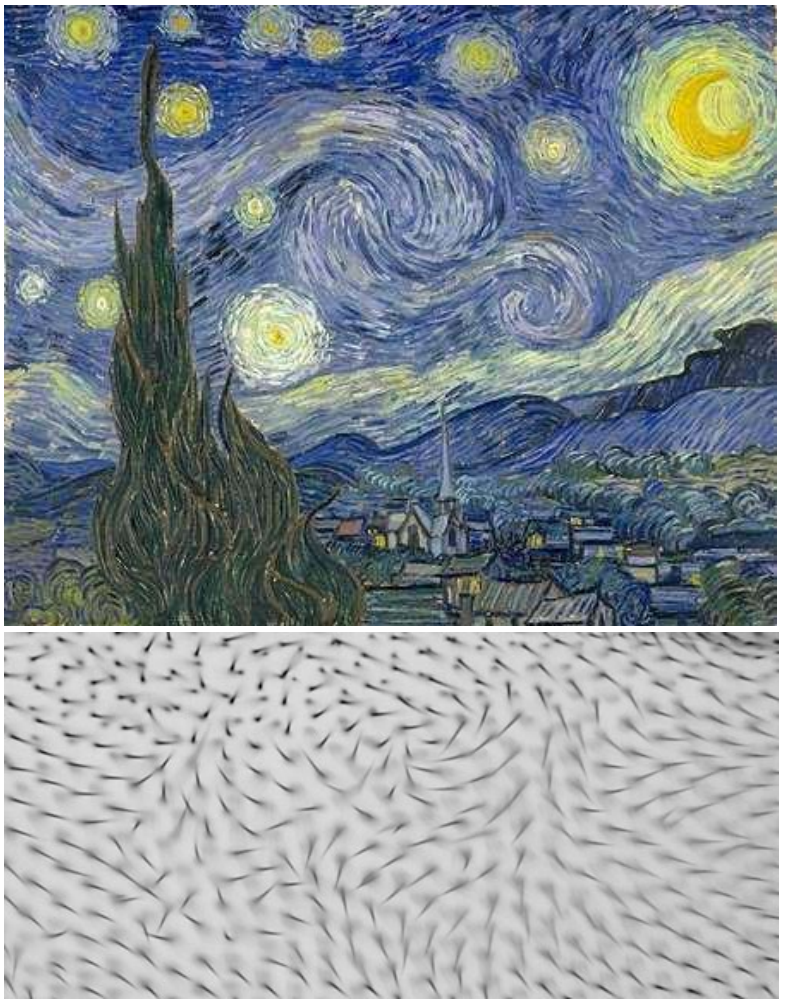


Some obvious issues in scaling single molecule definition to a genomic scale

- Fundamental complexities
  - ◊ Often >2 proteins/function
  - ◊ Multi-functionality:
    2 functions/protein
  - ◊ Role Conflation:
    molecular, cellular, phenotypic

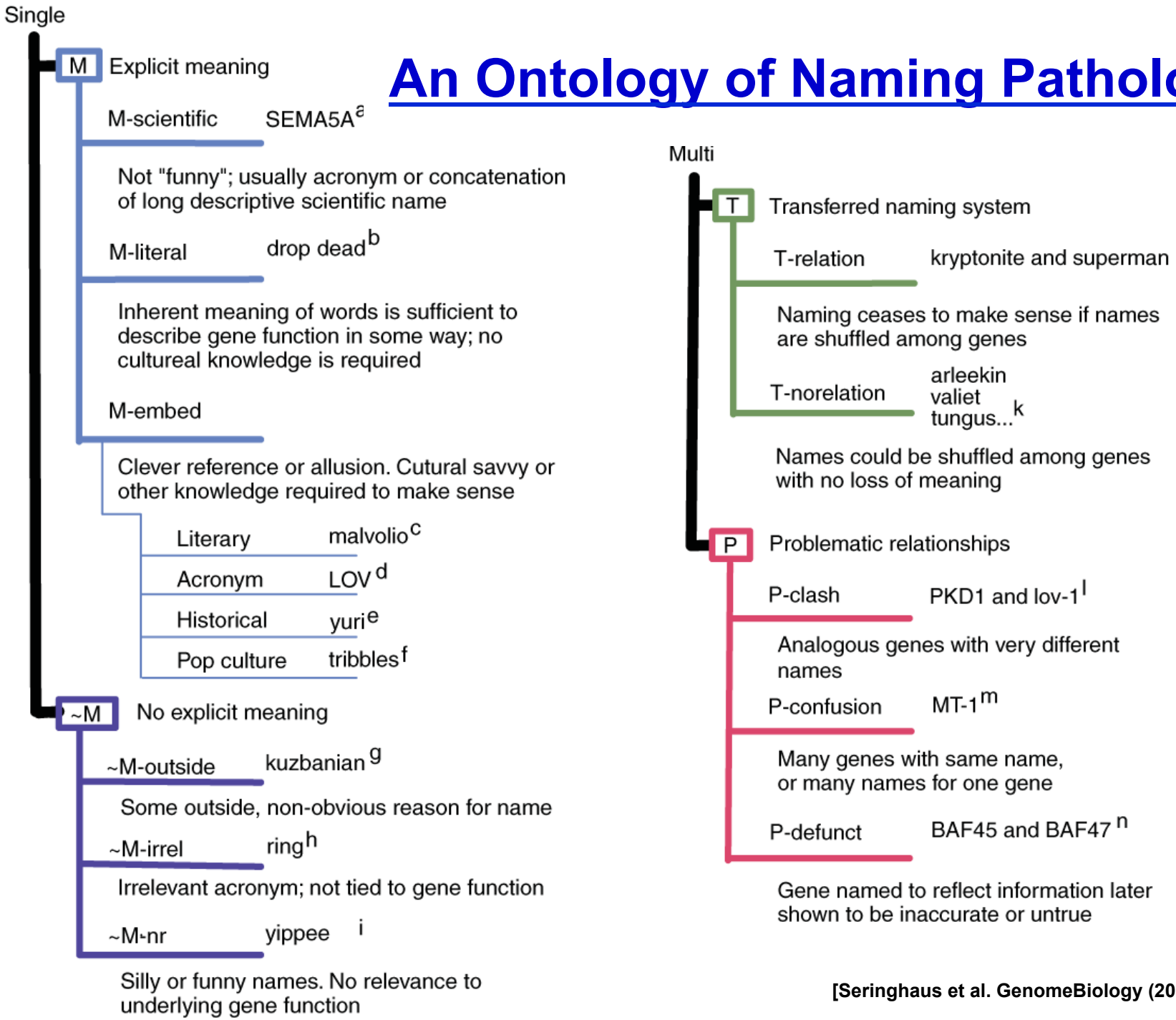
Some obvious issues in scaling single molecule definition to a genomic scale

- Fundamental complexities
  - ◊ Often >2 proteins/function
  - ◊ Multi-functionality:
    2 functions/protein
  - ◊ Role Conflation:
    molecular, cellular, phenotypic
- Fun terms… but do they scale?....
  - ◊ **Starry night** (P Adler, '94)



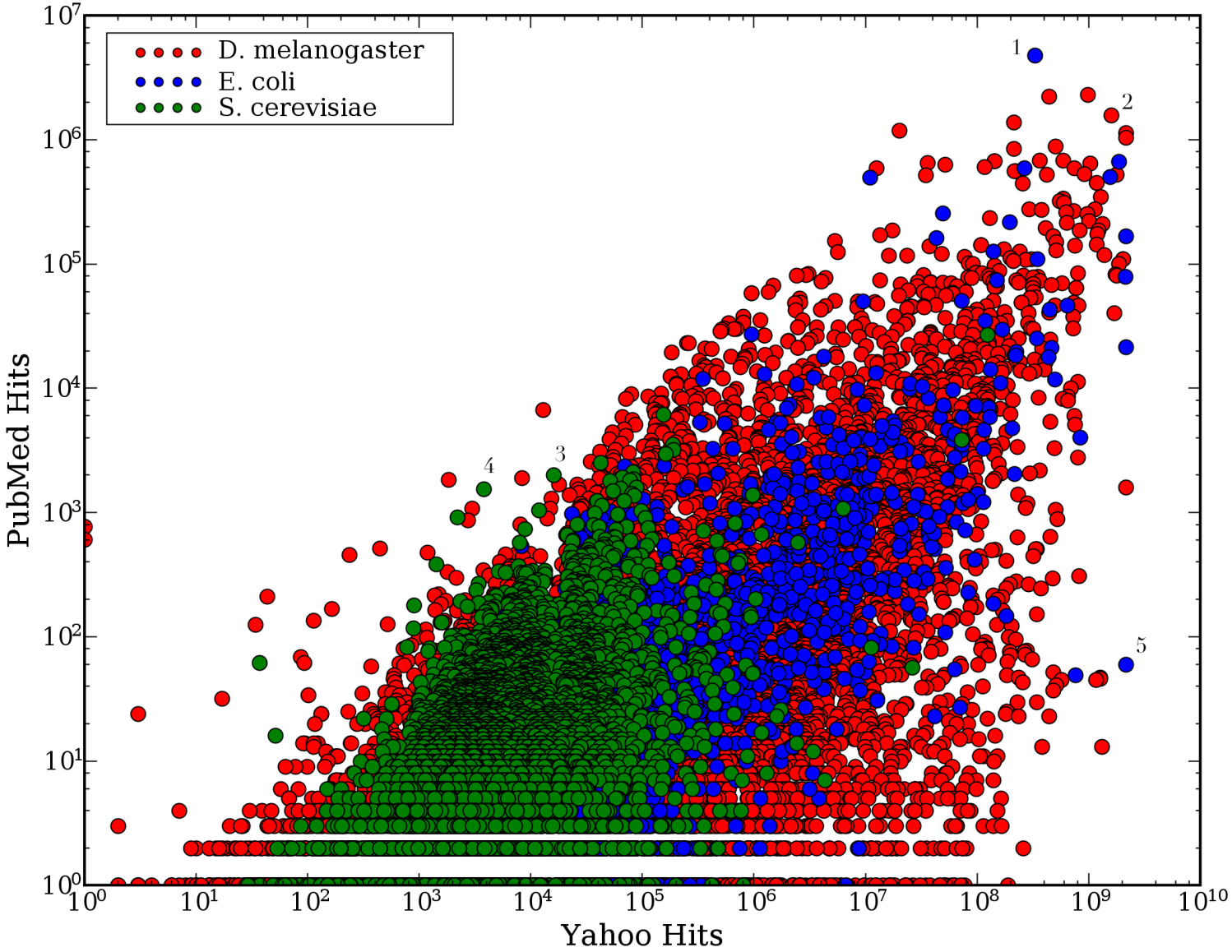
[Seringhaus et al. GenomeBiology (2008)]

An Ontology of Naming Pathologies



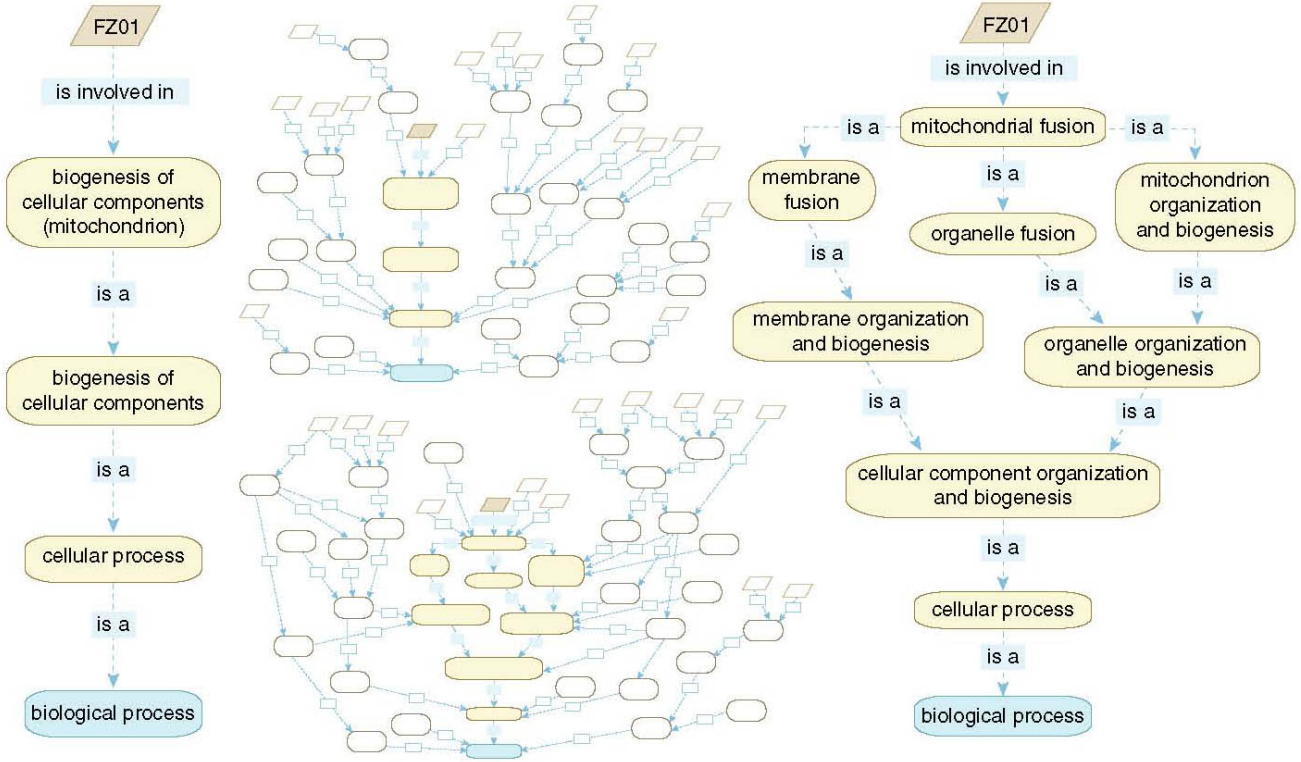
[Seringhaus et al. GenomeBiology (2008)]

Gene Name Skew



[Seringhaus et al. GenomeBiology (2008)]

Hierarchies & DAGs of controlled-vocab terms but still have issues...



MIPS (Mewes et al.)

GO (Ashburner et al.)

[Seringhaus & Gerstein, Am. Sci. '08]

Towards Developing Standardized Descriptions of Function

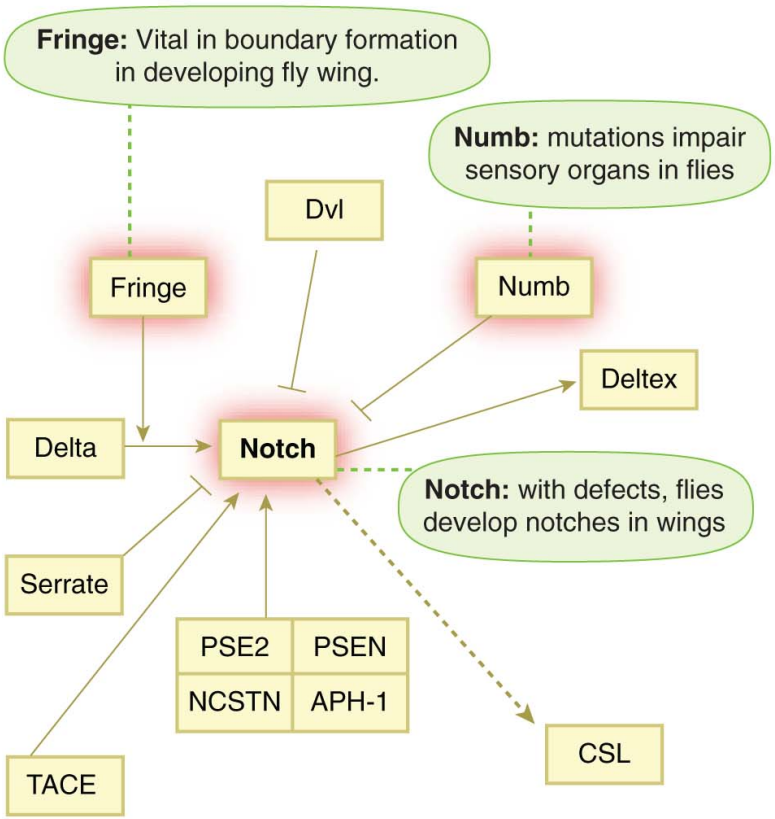
- Subjecting each gene to standardized expt. and cataloging effect
  - ◊ KOs of each gene in a variety of std. conditions => phenotypes
  - ◊ Std. binding expts for each gene (e.g. prot. chip)

- Function as a vector

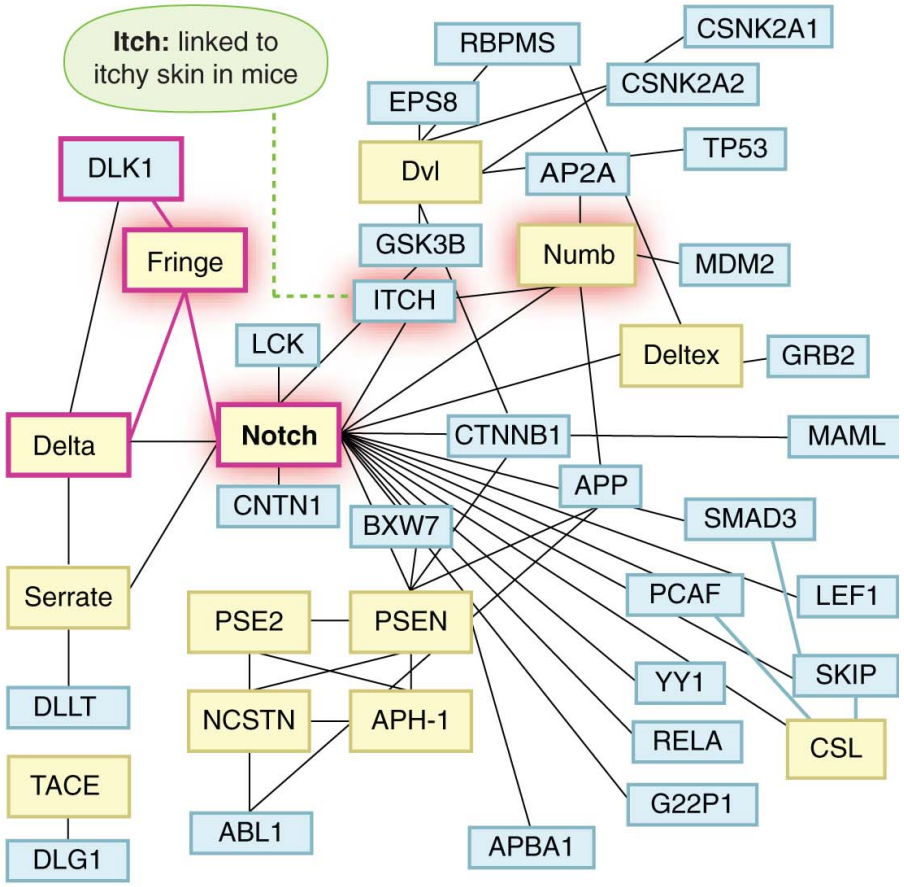
	nucleic acids		small molecules					proteins			
	DNA	RNA	ATP	Metal	CoA	NAD	.....	G protein	CDC28	Calmodulin	.....
protein 1	1.0	0	0	0	0	0	.....	0	0	0	.....
protein 2	0	0.9	0	0	0	0	.....	0	0	0	.....
protein 3	1.0	0	1.0	0	0	0	.....	0	0	0	.....
protein 4	0	0	0	0	0.8	0	.....	0	0	1.0	.....
protein 5	1.0	0	0	0	0	0	.....	0	0.9	0	.....
protein 6	0.9	0					.....				.....
protein 7	0	0.8					.....				.....
.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....

Interaction Vectors [Lan et al, IEEE 90:1848]

Networks (Old & New)



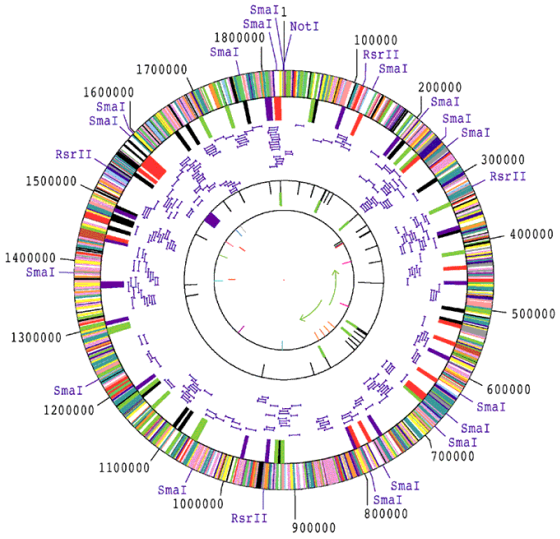
Classical KEGG pathway

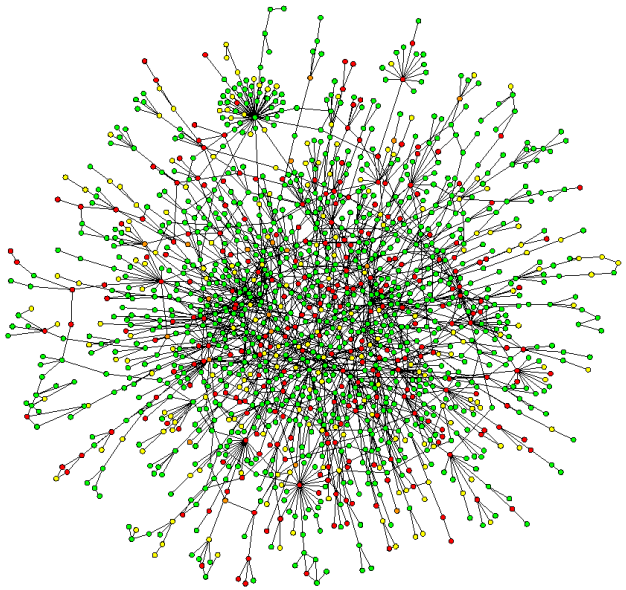


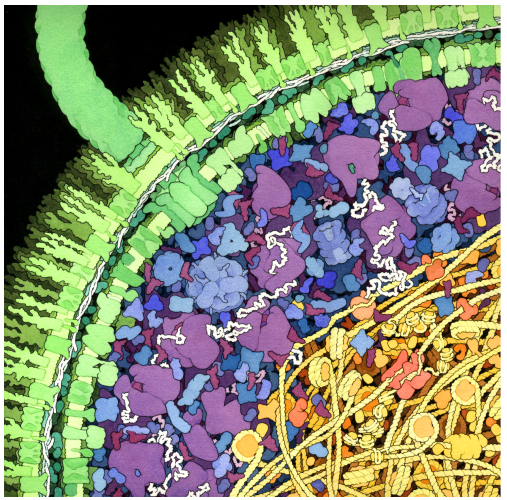
Same Genes in High-throughput Network

[Seringhaus & Gerstein, Am. Sci. '08]

Networks occupy a midway point in terms of level of understanding







1D: Complete Genetic Partslist

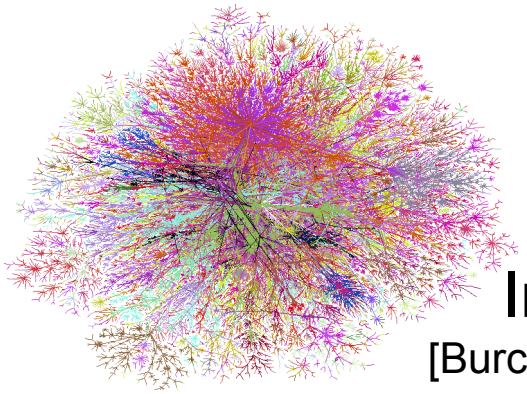
~2D: Bio-molecular Network Wiring Diagram

3D: Detailed structural understanding of cellular machinery

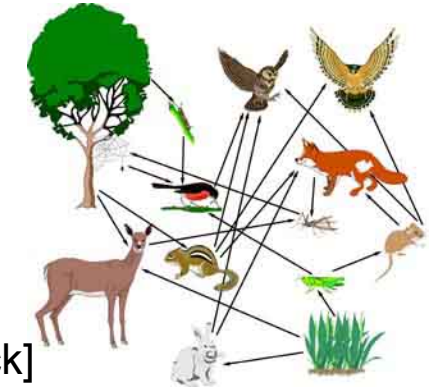
[Fleischmann et al., Science, 269 :496]

[Jeong et al. Nature, 41:411]

Networks as a universal language

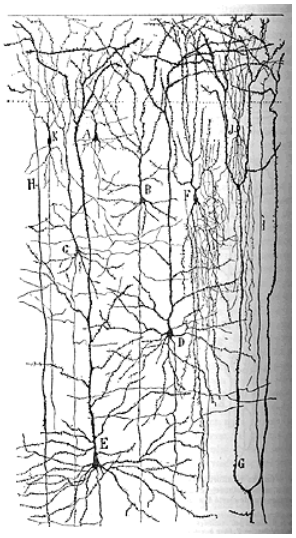


Internet
[Burch & Cheswick]

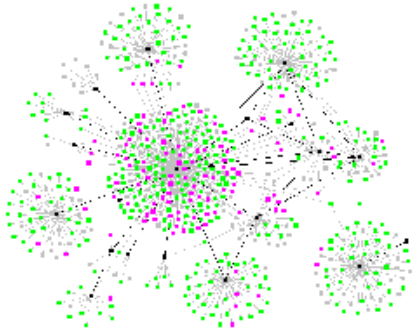


Food Web

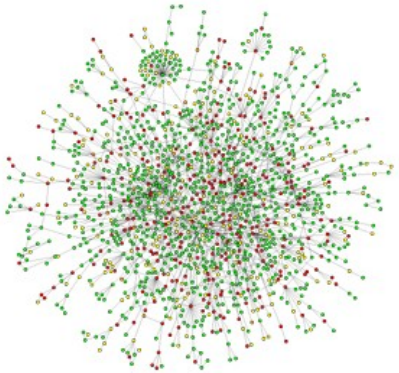
Electronic Circuit



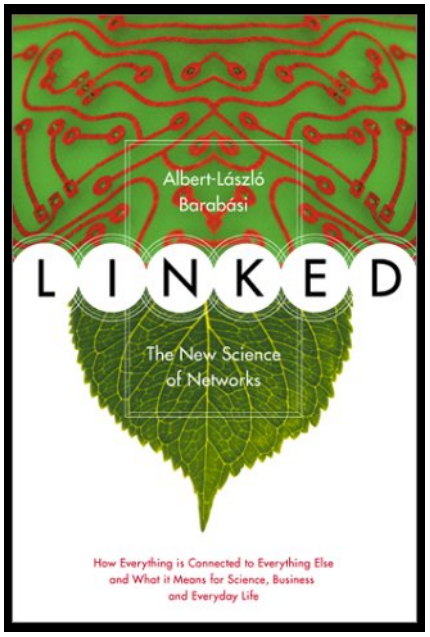
Neural Network
[Cajal]

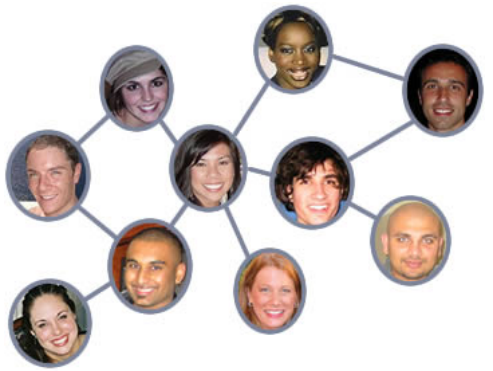


Disease Spread
[Krebs]



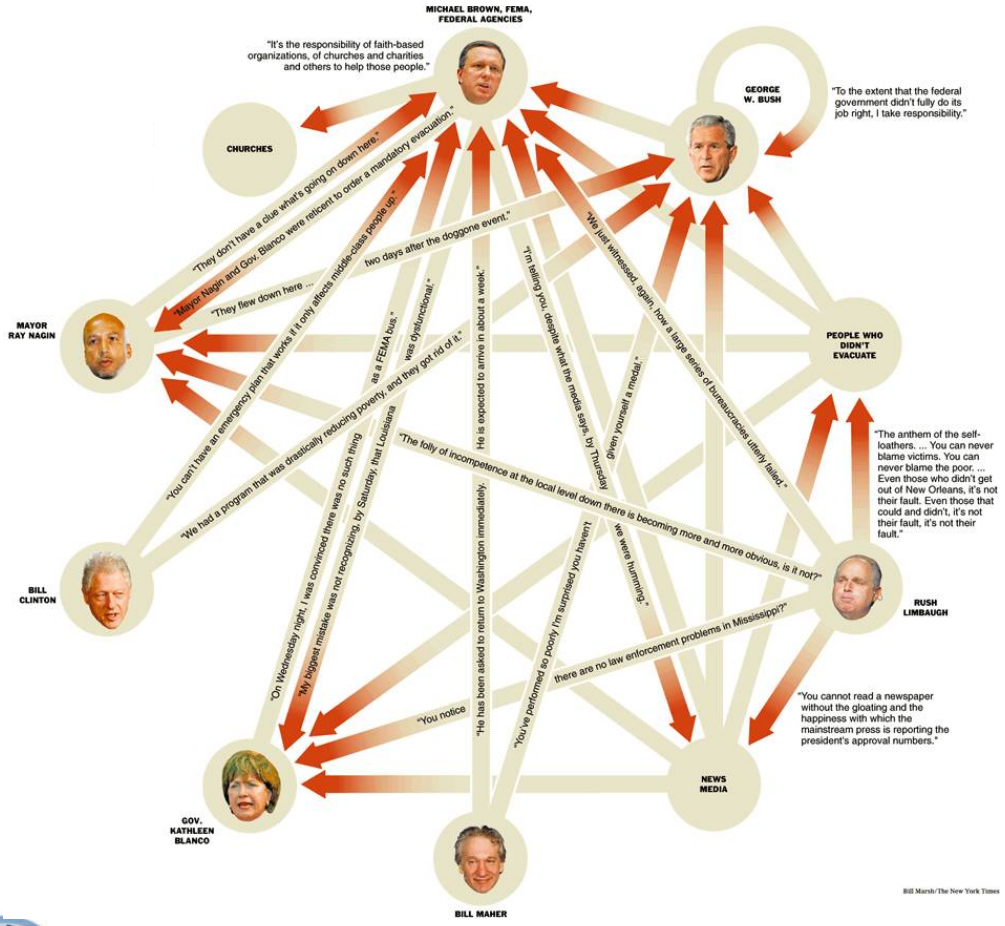
Protein Interactions
[Barabasi]



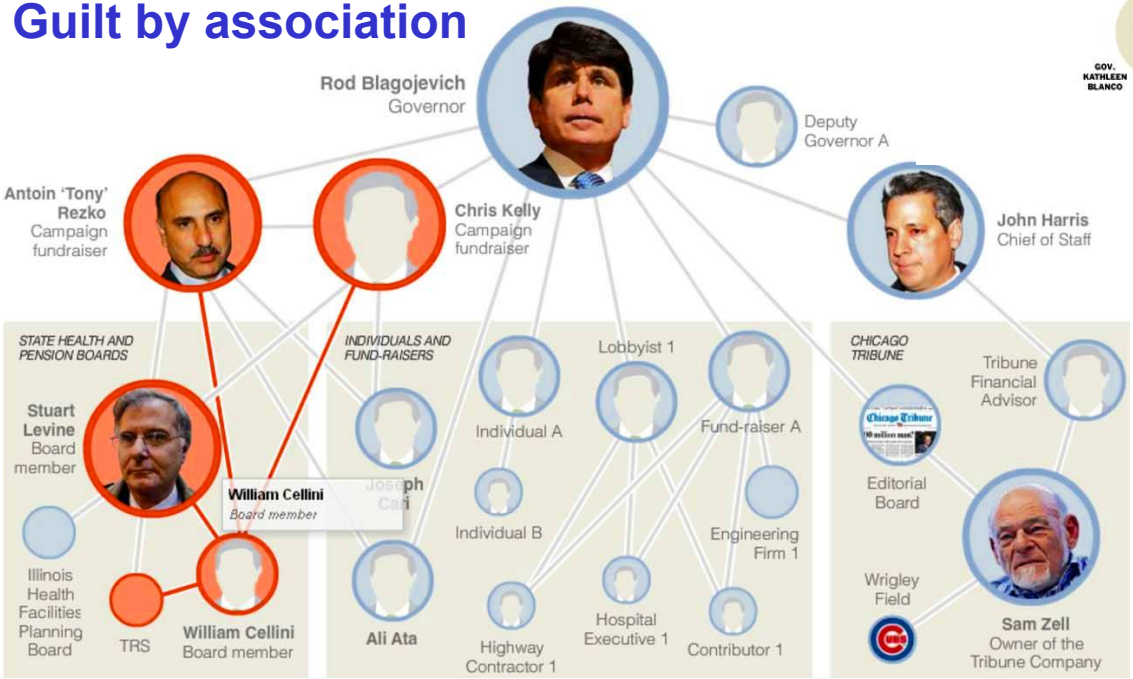


Social Network

Using the position in networks to describe function



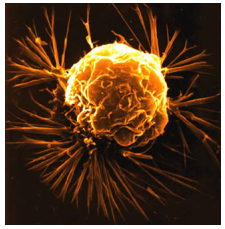
Guilt by association



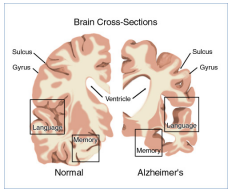
Finding the causal regulator (the "Blame Game")

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

Network pathology & pharmacology

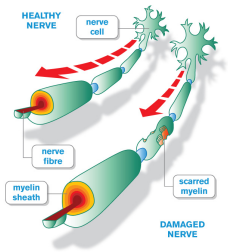


Breast Cancer

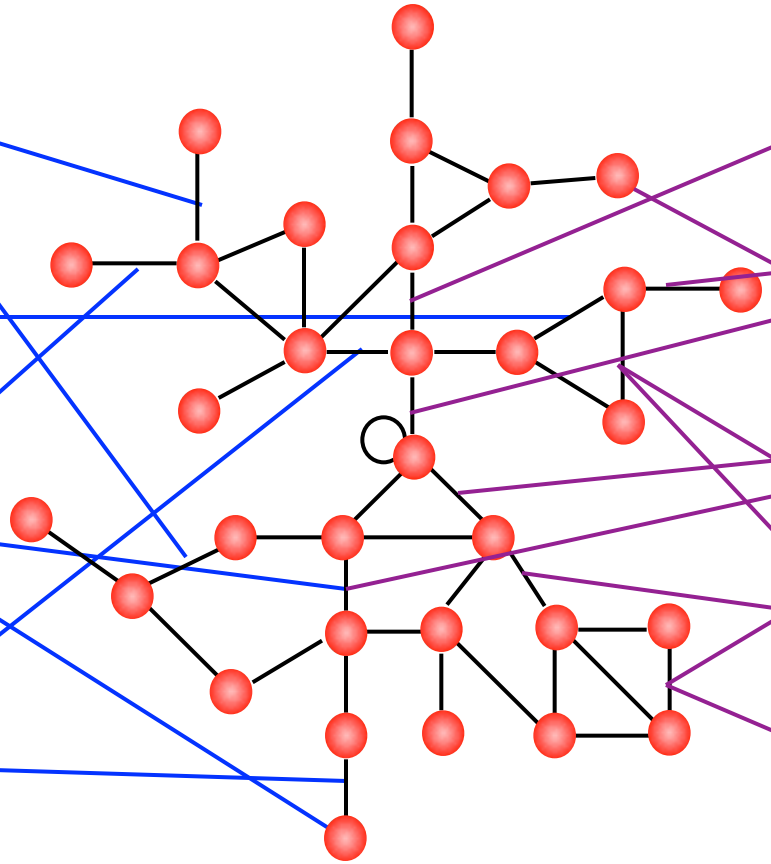


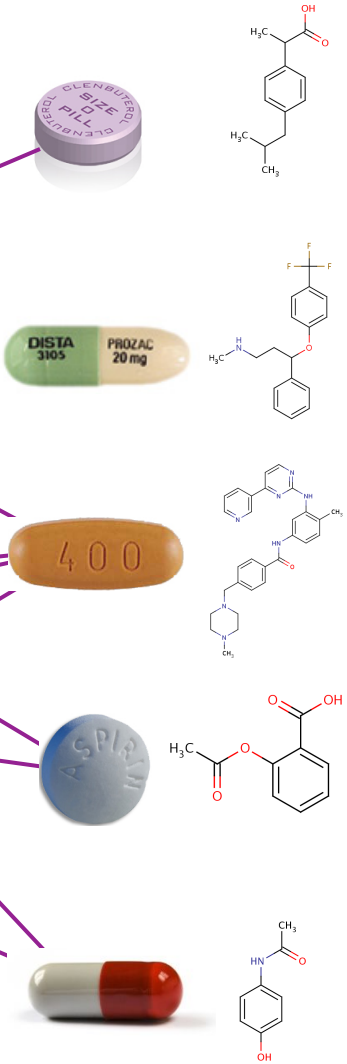
Alzheimer's Disease

Parkinson's Disease



Multiple Sclerosis





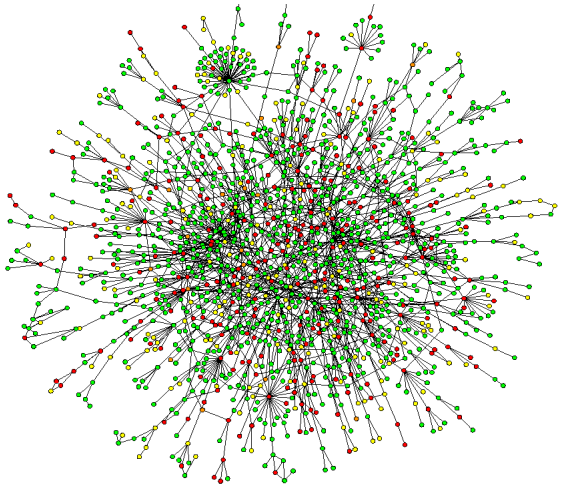
Interactome networks

[Adapted from H Yu]

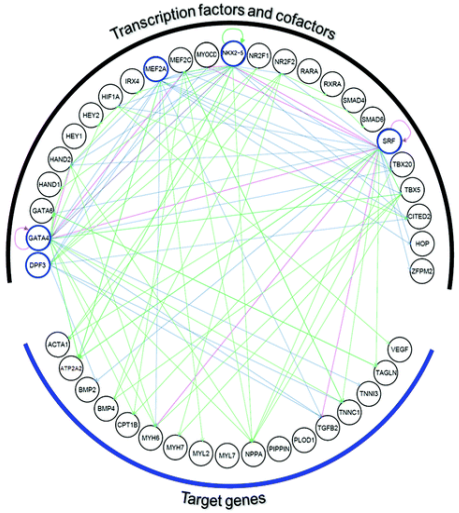
Outline: Molecular Networks

- Why Networks?
- Central Points in Networks
  - ◊ Hubs & Bottlenecks
    (yeast ppi & reg. net)
  - ◊ Tops of Heirarchies
    (yeast reg. net)
  - ◊ Identified by score
    (human miRNA-targ. net)
- Dynamics of Networks
  - ◊ Across environments
    (in prokaryote metab. pathways)
- Protein Networks & Variation
  (human ppi & miRNA-targ. net)

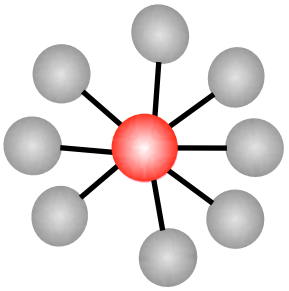
Different Types of Molecular Networks



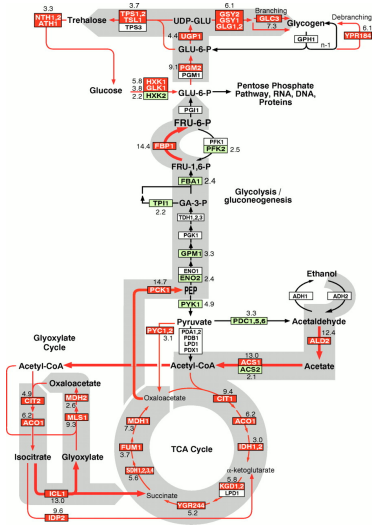
Protein-protein Interaction networks



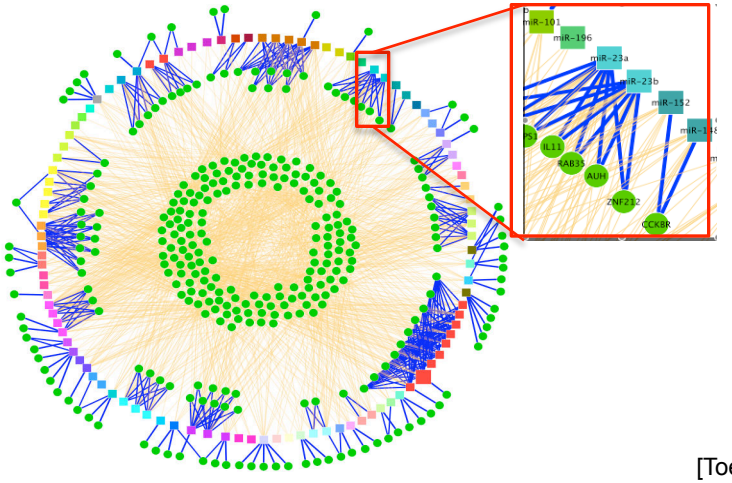
TF-target-gene Regulatory networks



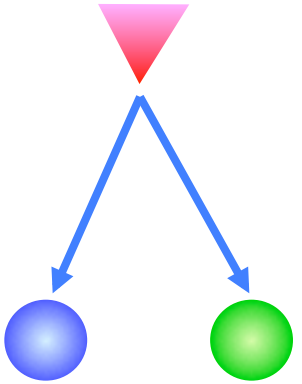
Undirected



Metabolic pathway networks



miRNA-target networks



Directed

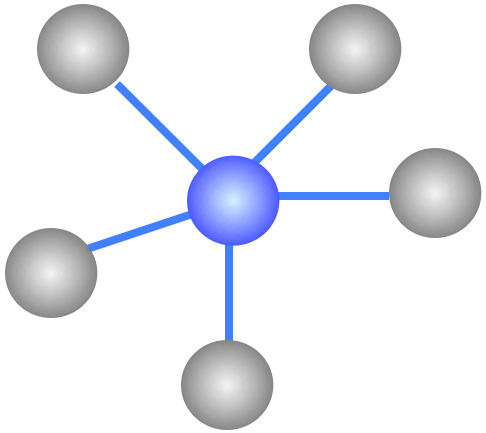
[Toenjes, *et al*, *Mol. BioSyst.* (2008); Jeong *et al*, *Nature* (2001); [Horak, et al, Genes & Development, 16:3017-3033; DeRisi, Iyer, and Brown, Science, 278:680-686]

Finding Central Points in Networks: Hubs & Bottlenecks

Where are key points networks ? How do we locate them ?

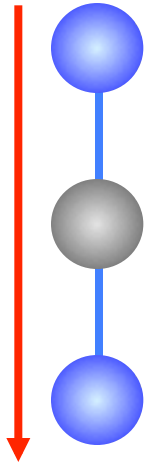
Global topological measures

Indicate the gross topological structure of the network



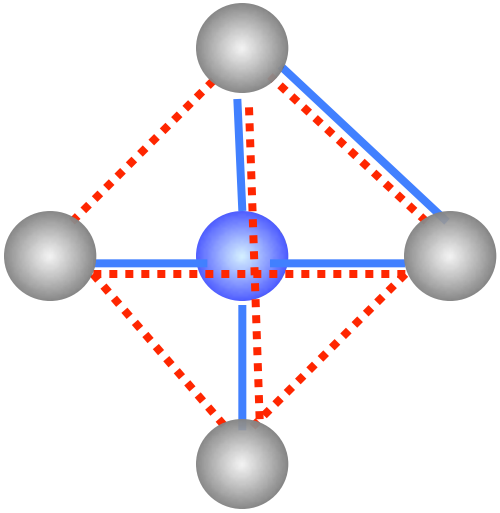
Degree (K)

5



Path length (L)

2



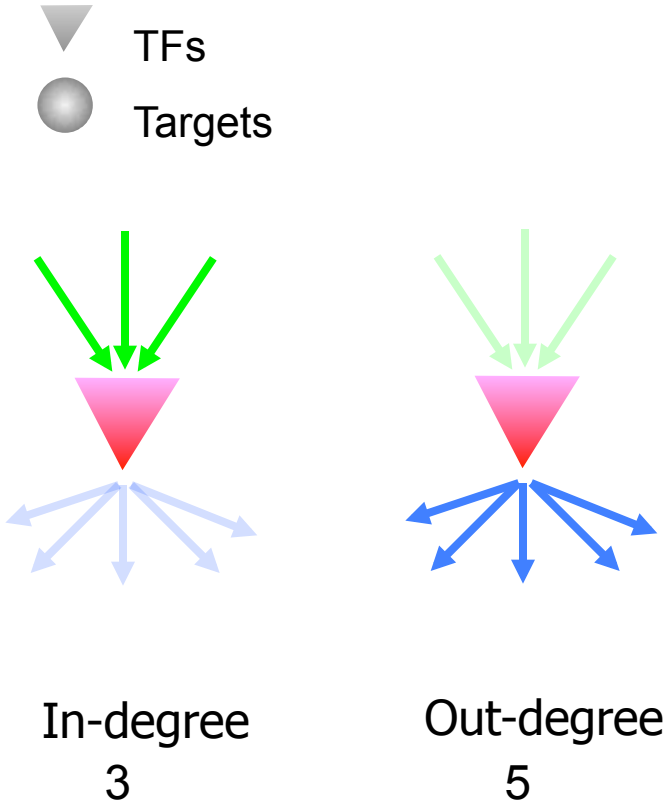
Clustering coefficient (C)

1/6

Interaction and expression networks are ***undirected***

[Barabasi]

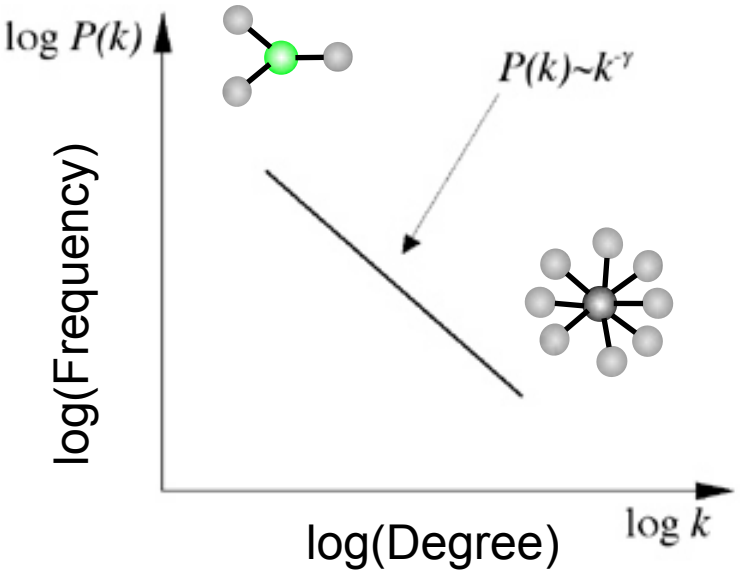
Global topological measures for directed networks

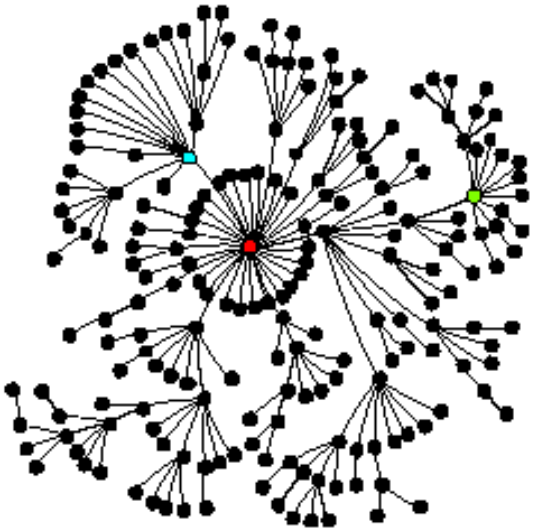


Regulatory and metabolic networks are ***directed***

Scale-free networks

Power-law distribution





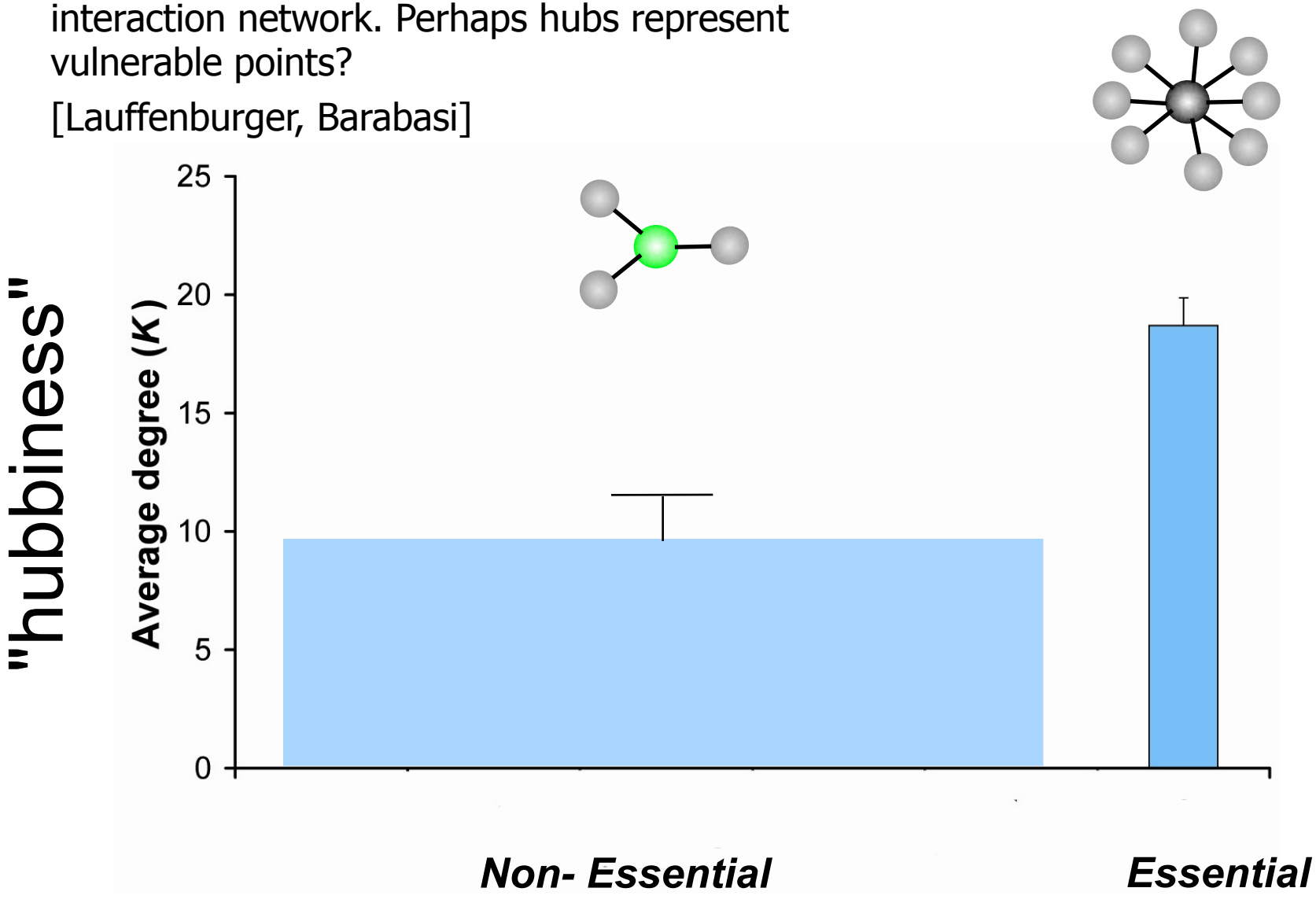
Hubs dictate the structure of the network

[Barabasi]

Hubs tend to be Essential

Integrate gene essentiality data with protein interaction network. Perhaps hubs represent vulnerable points?

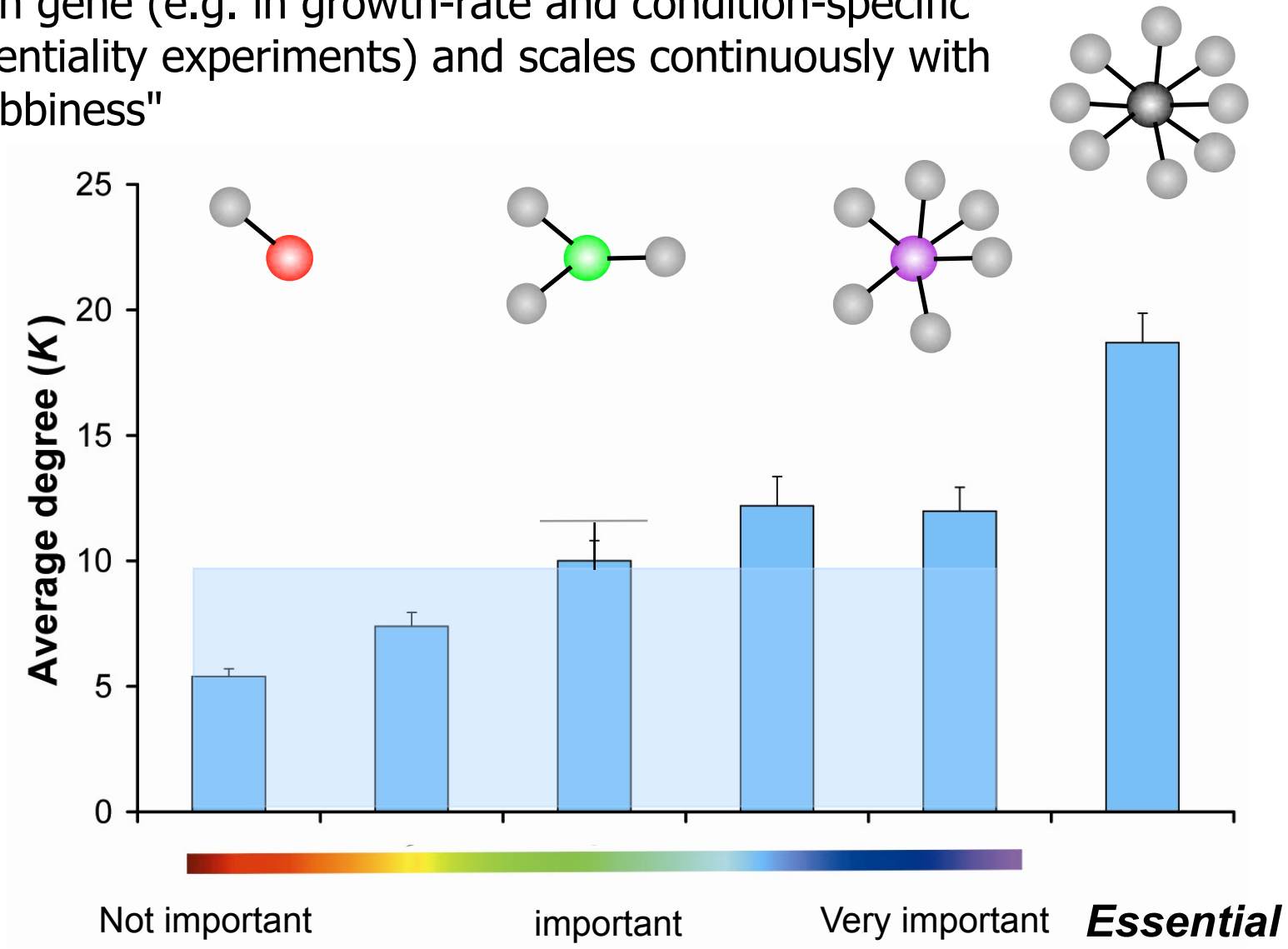
[Lauffenburger, Barabasi]



[Yu et al., 2003, TIG]

Relationships extends to "Marginal Essentiality"

Marginal essentiality measures relative importance of each gene (e.g. in growth-rate and condition-specific essentiality experiments) and scales continuously with "hubbiness"



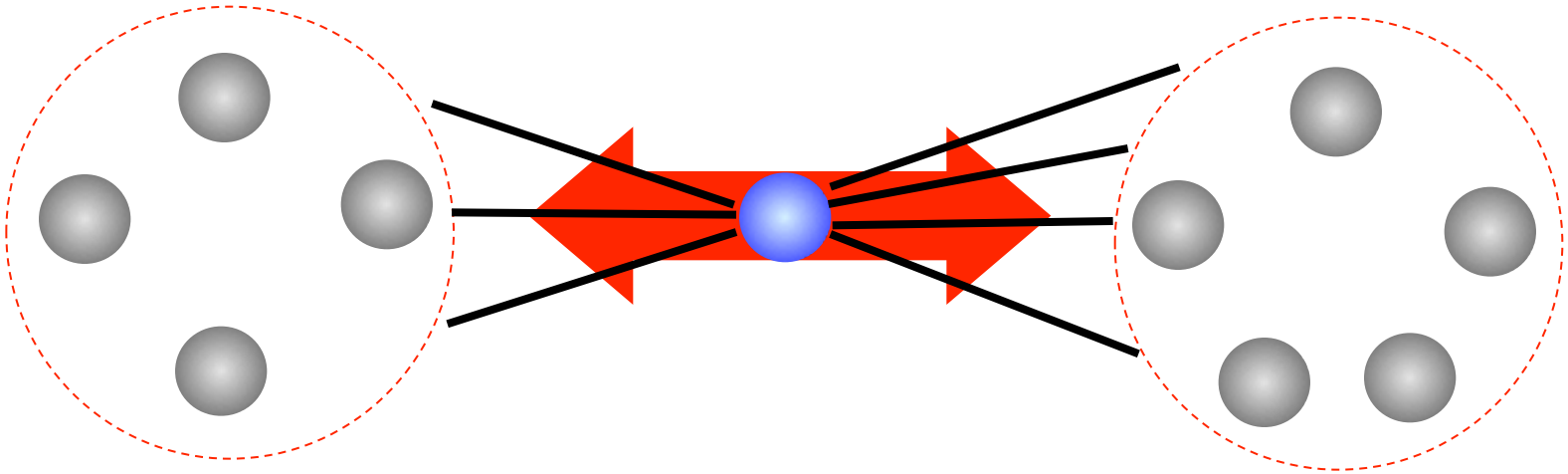
[Yu et al., 2003, TIG]

Another measure of Centrality: Betweenness centrality

Betweenness of a node is the number of shortest paths of pairs of vertices that run through it -- a measure of information flow.

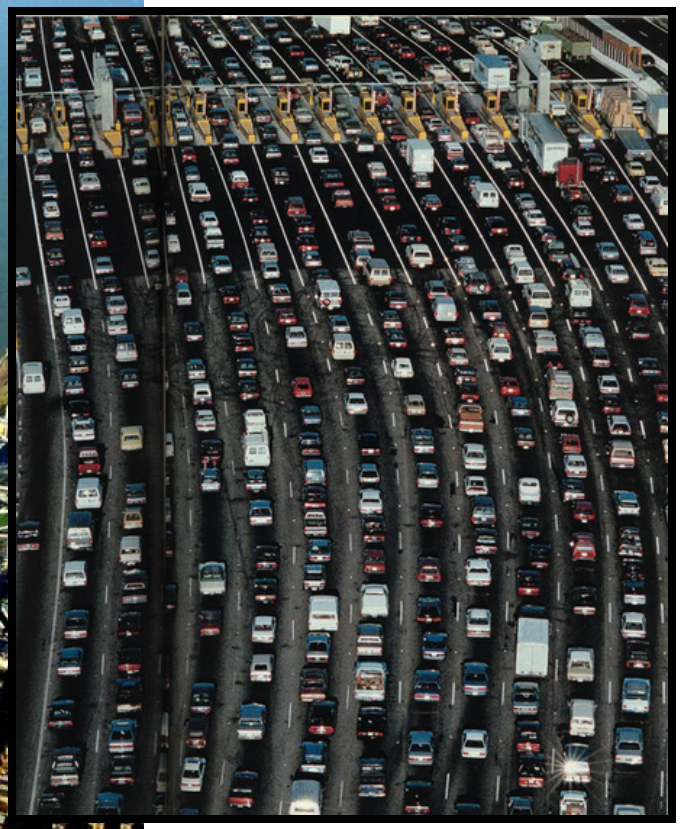
Freeman LC (1977) Set of measures of centrality based on betweenness. Sociometry 40: 35–41.

Girvan & Newman (2002) PNAS 99: 7821.

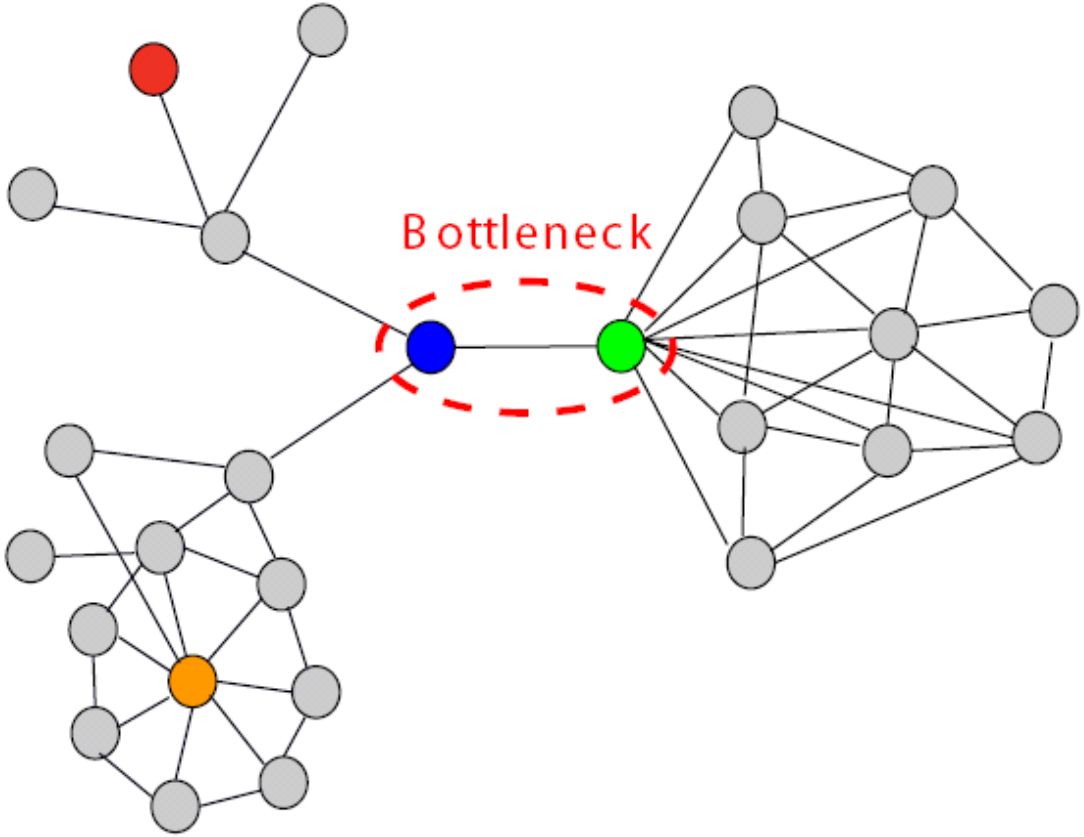


Betweenness centrality -- Bottlenecks

Proteins with high betweenness are defined as *Bottlenecks* (top 20%), in analogy to the traffic system



Bottlenecks & Hubs



Hub-bottleneck **node**

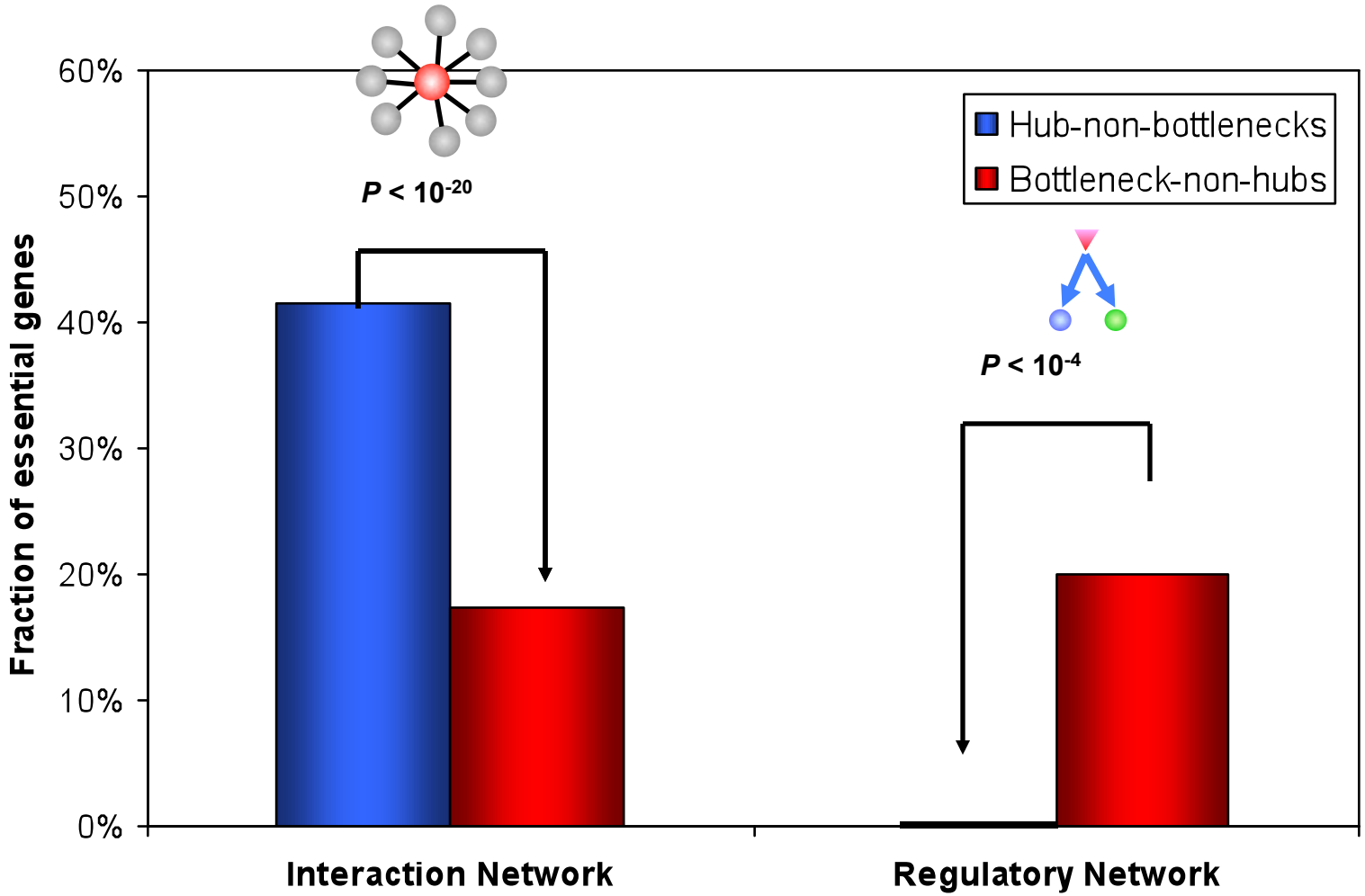
Non-hub-bottleneck **node**

Hub-non-bottleneck **node**

Non-hub-non-bottleneck **node**

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

Bottlenecks are what matters in regulatory networks

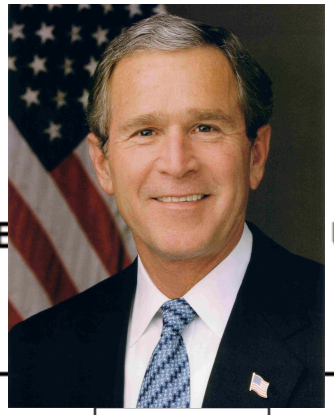


[Yu *et al.*, *PLoS Comput Biol* (2007)]

Finding Central Points in Networks #2: Tops of the Hierarchy

Where are key points networks ? How do we locate them ?

Social Hierarchy



THE GOVE UNITED STATES

LEGISLATIVE BRANCH

THE CONGRESS

SENATE HOUSE

ARCHITECT OF THE CAPITOL
UNITED STATES BOTANIC GARDEN
GENERAL ACCOUNTING OFFICE
GOVERNMENT PRINTING OFFICE
LIBRARY OF CONGRESS
CONGRESSIONAL BUDGET OFFICE

EX SIDENT

WHITE HOUSE OFFICE
OFFICE OF THE VICE PRESIDE
COUNCIL OF ECONOMIC ADVIS
COUNCIL ON ENVIRONMENTAL
NATIONAL SECURITY COUNCIL
OFFICE OF ADMINISTRATION

NAGEMENT AND BUDGET
ATIONAL DRUG CONTROL POLICY
LICY DEVELOPMENT
IENCE AND TECHNOLOGY POLICY
E U.S. TRADE REPRESENTATIVE

JUDICIAL BRANCH

THE SUPREME COURT OF THE UNITED STATES

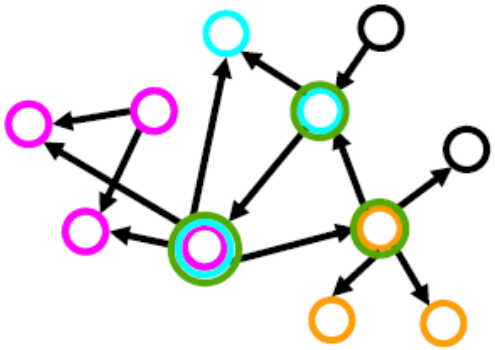
UNITED STATES COURTS OF APPEALS
UNITED STATES DISTRICT COURTS
TERRITORIAL COURTS
UNITED STATES COURT OF INTERNATIONAL TRADE
UNITED STATES COURT OF FEDERAL CLAIMS
UNITED STATES COURT OF APPEALS FOR THE ARMED FORCES
UNITED STATES TAX COURT
UNITED STATES COURT OF APPEALS FOR VETERANS CLAIMS
ADMINISTRATIVE OFFICE OF THE UNITED STATES COURTS
FEDERAL JUDICIAL CENTER
UNITED STATES SENTENCING COMMISSION

INDEPENDENT ESTABLISHMENTS AND GOVERNMENT CORPORATIONS

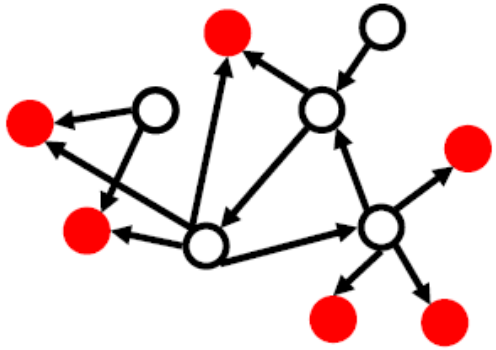
AFRICAN DEVELOPMENT FOUNDATION
CENTRAL INTELLIGENCE AGENCY
COMMODITY FUTURES TRADING COMMISSION
CONSUMER PRODUCT SAFETY COMMISSION
CORPORATION FOR NATIONAL AND COMMUNITY SERVICE
DEFENSE NUCLEAR FACILITIES SAFETY BOARD
ENVIRONMENTAL PROTECTION AGENCY
EQUAL EMPLOYMENT OPPORTUNITY COMMISSION
EXPORT-IMPORT BANK OF THE U.S.
FARM CREDIT ADMINISTRATION
FEDERAL COMMUNICATIONS COMMISSION
FEDERAL DEPOSIT INSURANCE CORPORATION
FEDERAL ELECTION COMMISSION
FEDERAL HOUSING FINANCE BOARD
FEDERAL LABOR RELATIONS AUTHORITY
FEDERAL MARITIME COMMISSION
FEDERAL MEDIATION AND CONCILIATION SERVICE
FEDERAL MINE SAFETY AND HEALTH REVIEW COMMISSION
FEDERAL RESERVE SYSTEM
FEDERAL RETIREMENT THRIFT INVESTMENT BOARD
FEDERAL TRADE COMMISSION
GENERAL SERVICES ADMINISTRATION
INTER-AMERICAN FOUNDATION
MERIT SYSTEMS PROTECTION BOARD
NATIONAL AERONAUTICS AND SPACE ADMINISTRATION
NATIONAL ARCHIVES AND RECORDS ADMINISTRATION
NATIONAL CAPITAL PLANNING COMMISSION
NATIONAL CREDIT UNION ADMINISTRATION
NATIONAL FOUNDATION ON THE ARTS AND THE HUMANITIES
NATIONAL LABOR RELATIONS BOARD
NATIONAL MEDIATION BOARD
NATIONAL RAILROAD PASSENGER CORPORATION (AMTRAK)
NATIONAL SCIENCE FOUNDATION
NATIONAL TRANSPORTATION SAFETY BOARD
NUCLEAR REGULATORY COMMISSION
OCCUPATIONAL SAFETY AND HEALTH REVIEW COMMISSION
OFFICE OF GOVERNMENT ETHICS
OFFICE OF PERSONNEL MANAGEMENT
OFFICE OF SPECIAL COUNSEL
OVERSEAS PRIVATE INVESTMENT CORPORATION
PEACE CORPS
PENSION BENEFIT GUARANTY CORPORATION
POSTAL RATE COMMISSION
RAILROAD RETIREMENT BOARD
SECURITIES AND EXCHANGE COMMISSION
SELECTIVE SERVICE SYSTEM
SMALL BUSINESS ADMINISTRATION
SOCIAL SECURITY ADMINISTRATION
TENNESSEE VALLEY AUTHORITY
TRADE AND DEVELOPMENT AGENCY
U.S. AGENCY FOR INTERNATIONAL DEVELOPMENT
U.S. COMMISSION ON CIVIL RIGHTS
U.S. INTERNATIONAL TRADE COMMISSION
U.S. POSTAL SERVICE

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

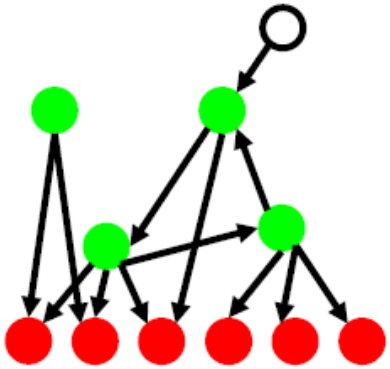
I. Example network with all 4 motifs



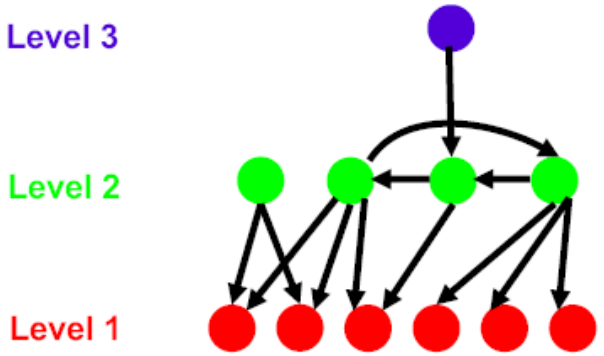
II. Finding terminal nodes (Red)



III. Finding mid-level nodes (Green)

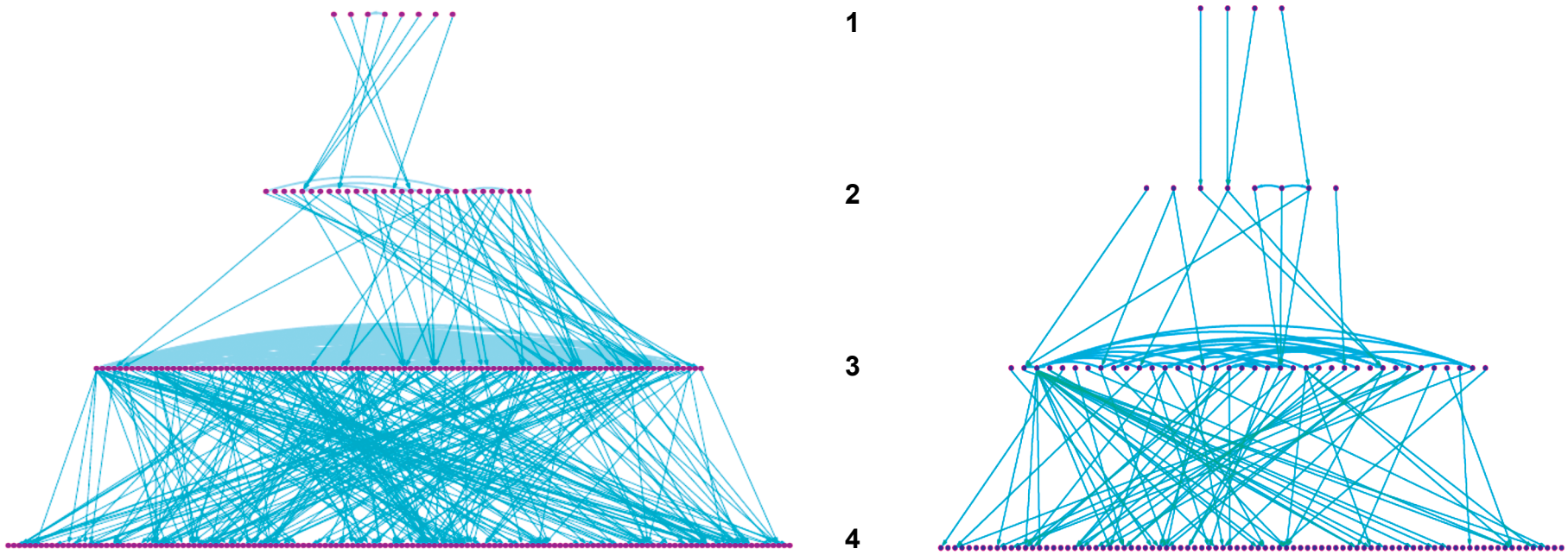


IV. Finding top-most nodes (Blue)



[Yu et al., PNAS (2006)]

Regulatory Networks have similar hierarchical structures

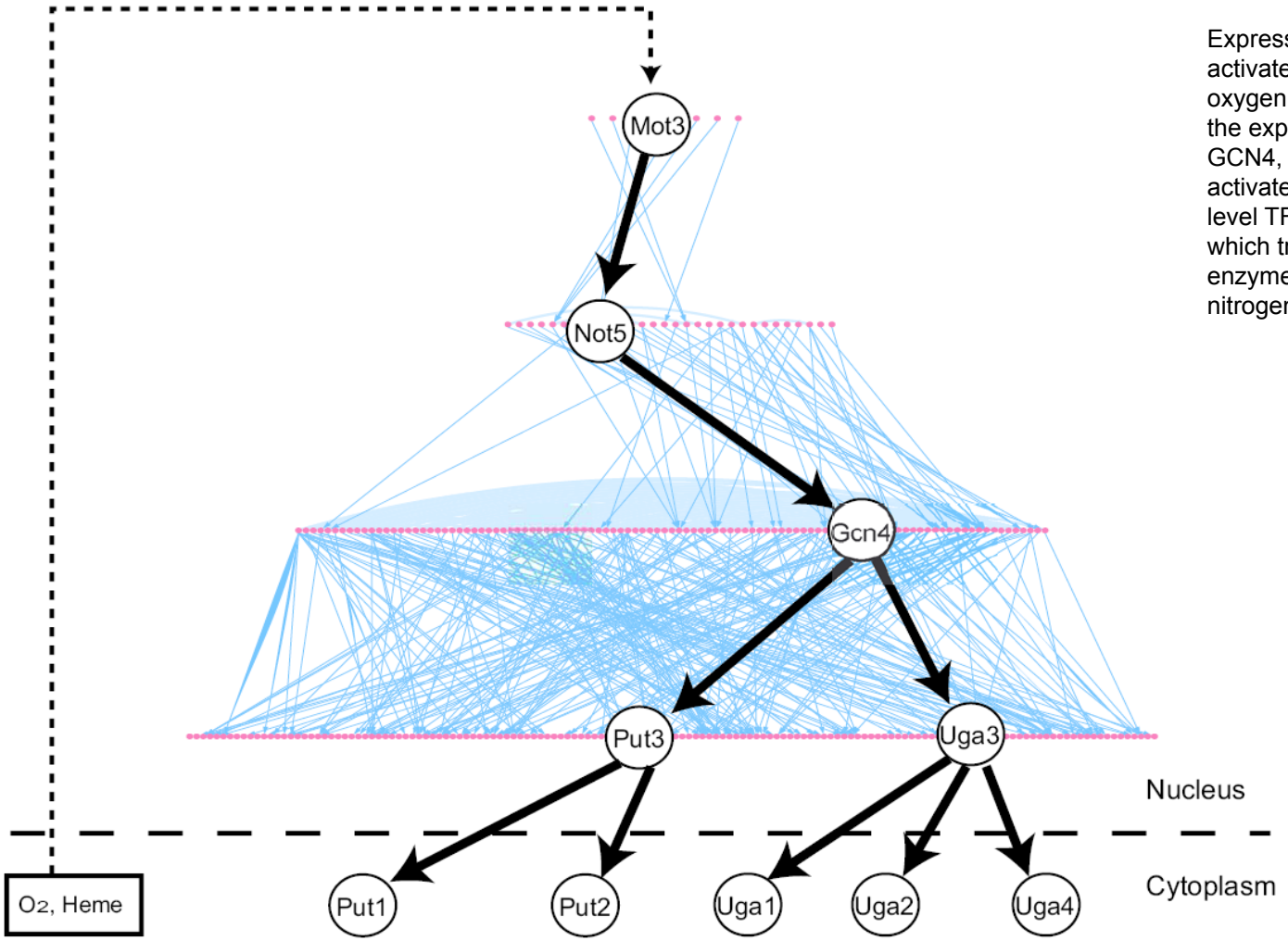


S. cerevisiae

E. coli

[Yu *et al.*, *Proc Natl Acad Sci U S A* (2006)]

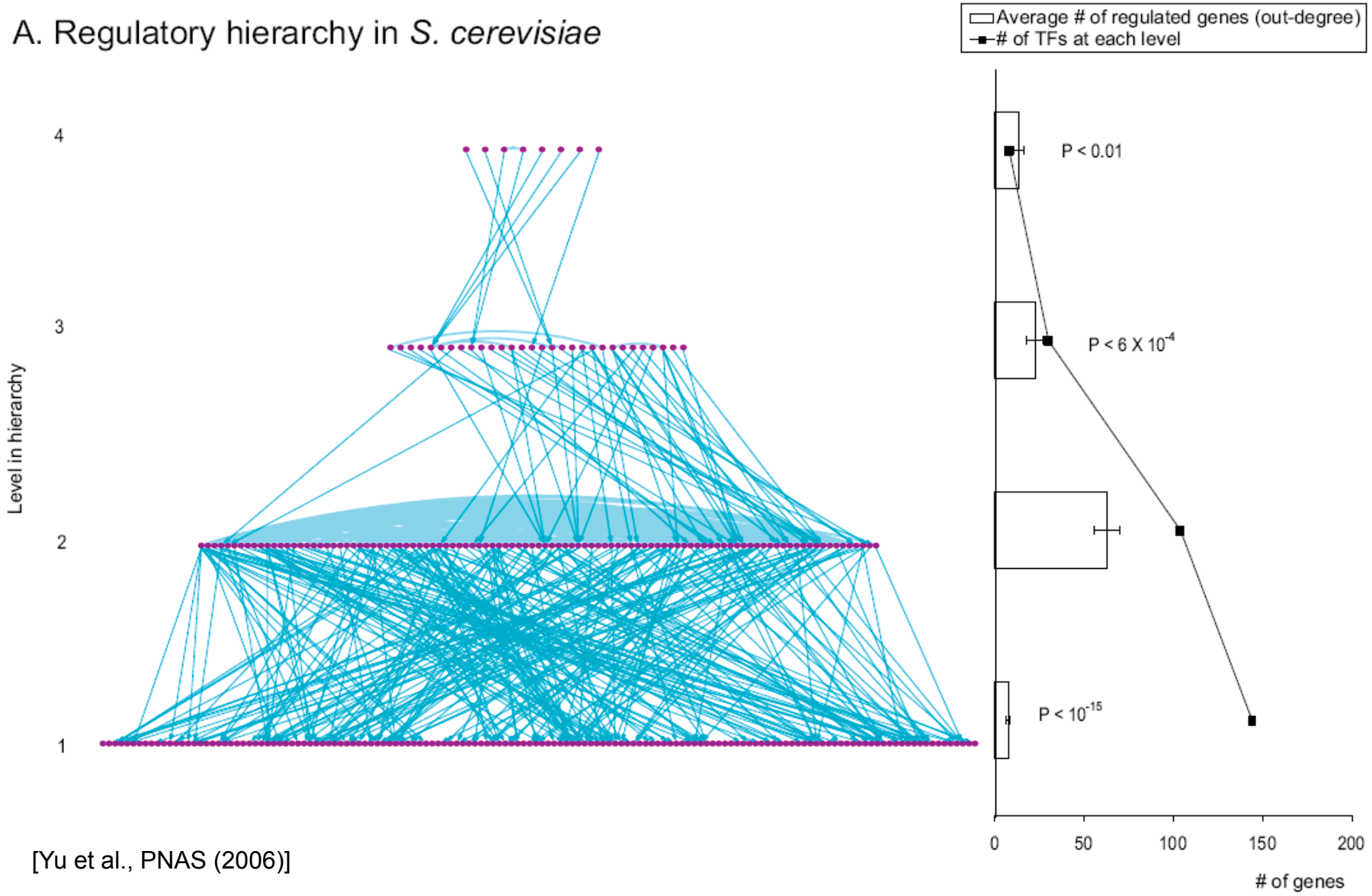
Example of Path Through Regulatory Network



Expression of MOT3 is activated by heme and oxygen. Mot3 in turn activates the expression of NOT5 and GCN4, mid-level hubs. GCN4 activates two specific bottom-level TFs, Put3 and Uga3, which trigger the expression of enzymes in proline and nitrogen utilization.

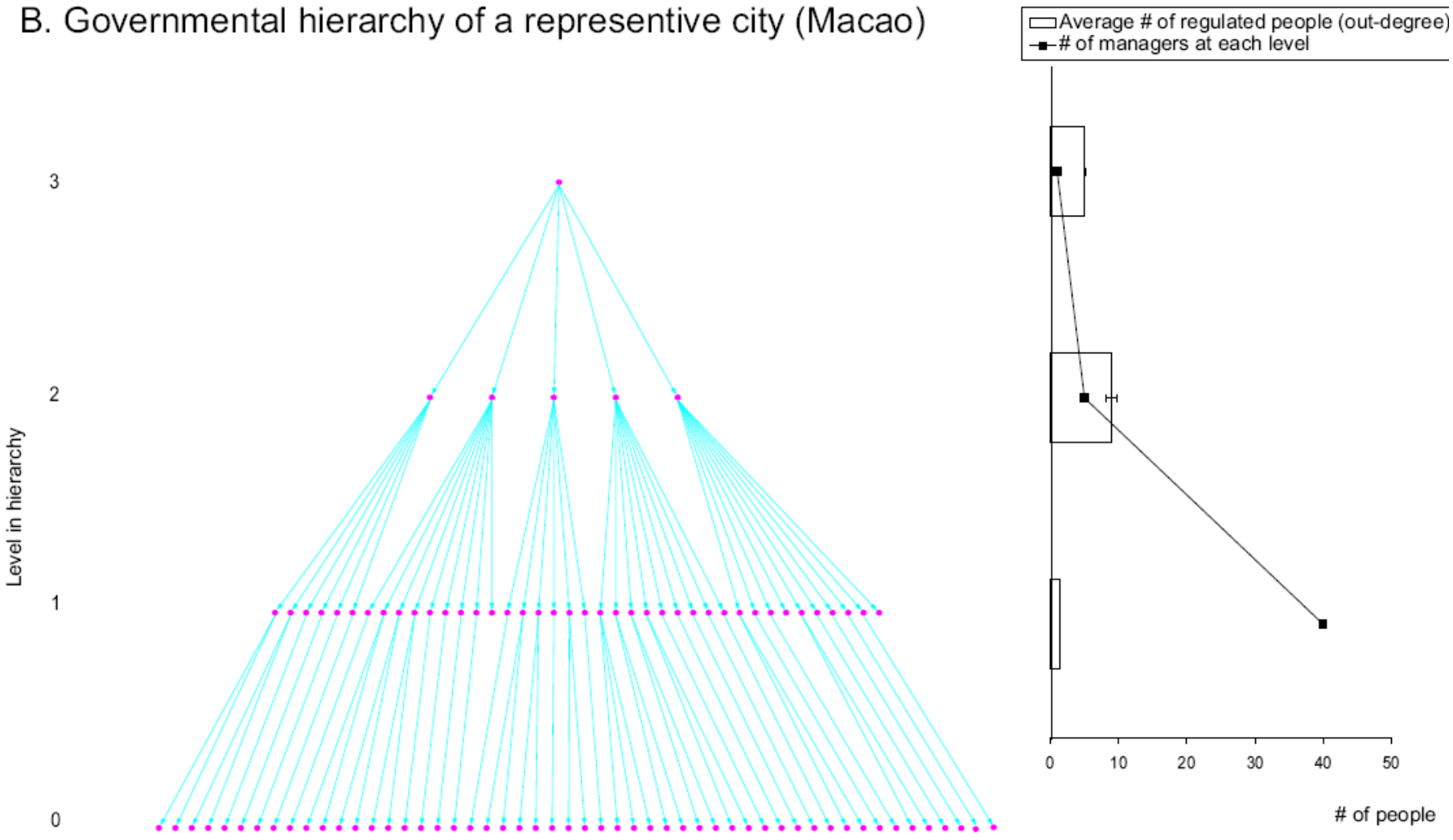
[Yu et al., PNAS (2006)]

Yeast Regulatory Hierarchy: the Middle-managers Rule

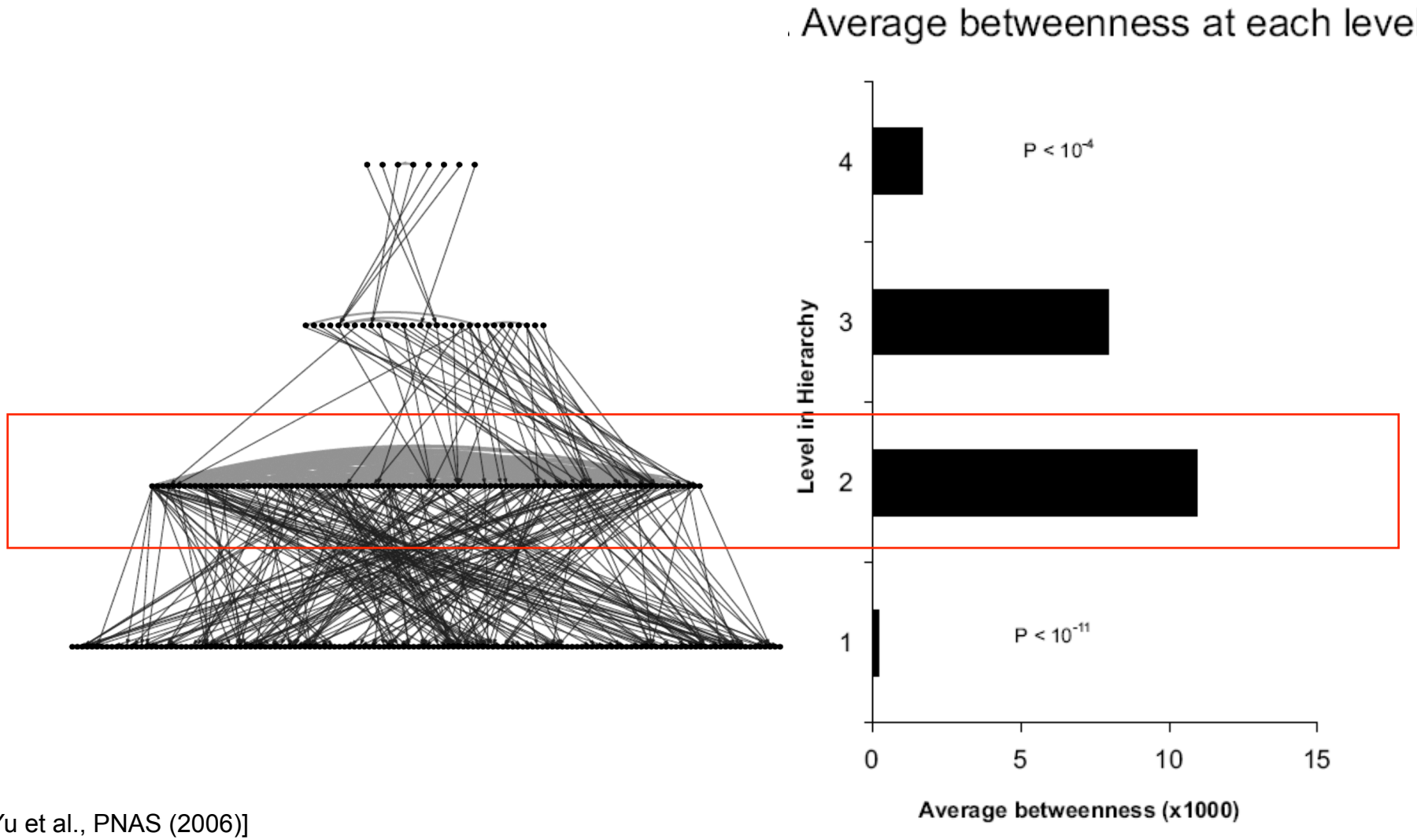


[Yu et al., PNAS (2006)]

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

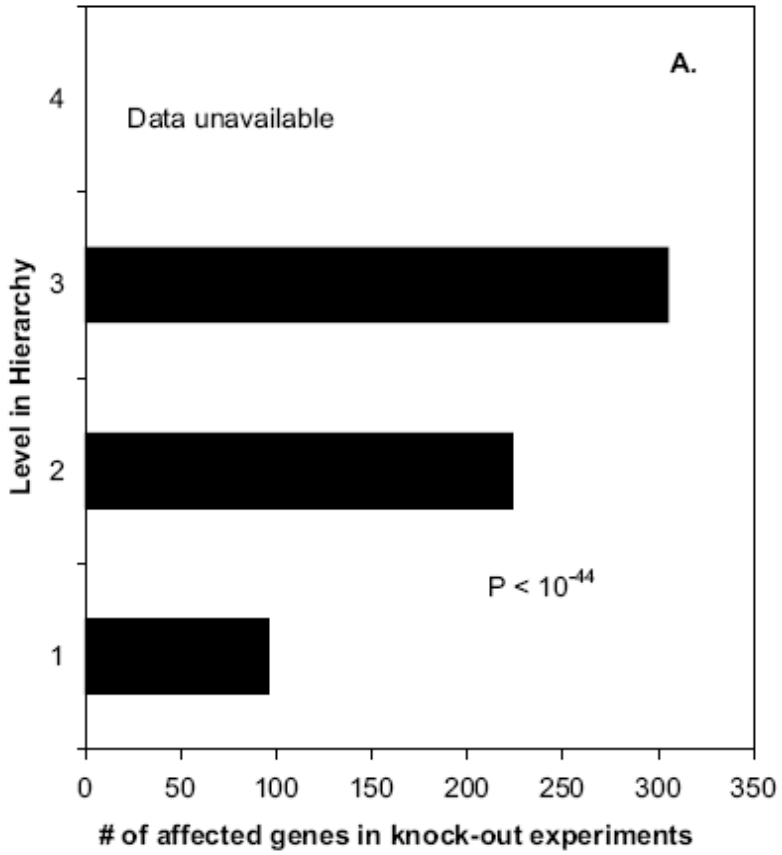


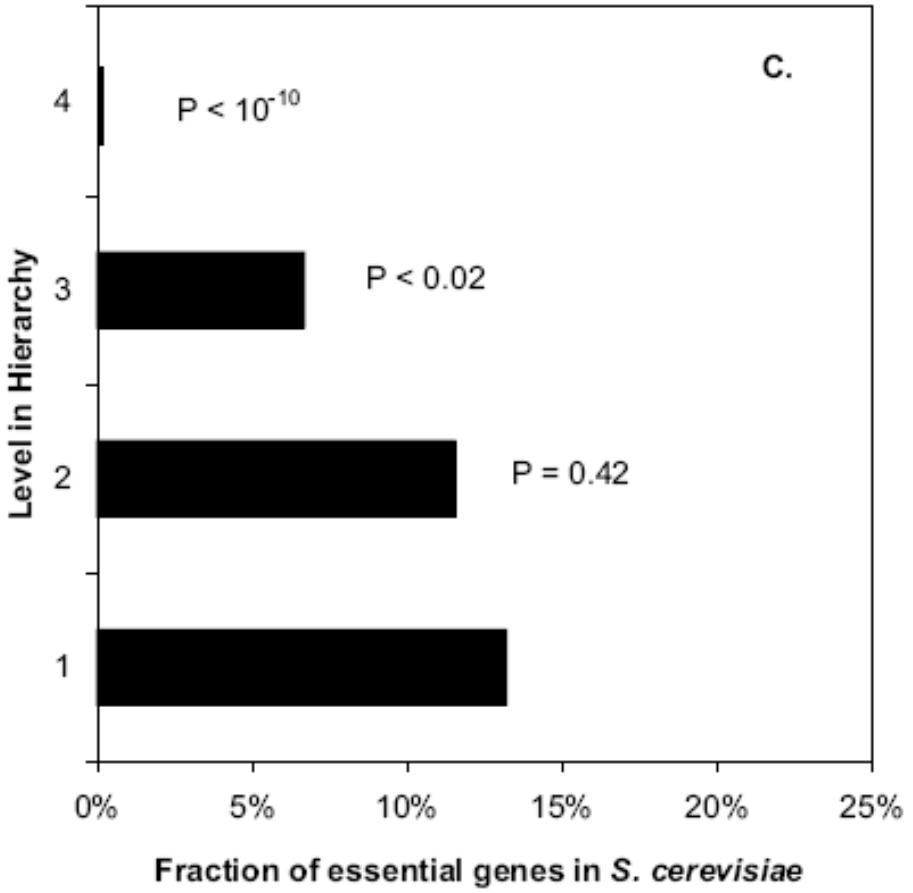
Characteristics of Regulatory Hierarchy: Middle Managers are Information Flow Bottlenecks



[Yu et al., PNAS (2006)]

Characteristics of Regulatory Hierarchy: The Paradox of Influence and Essentiality



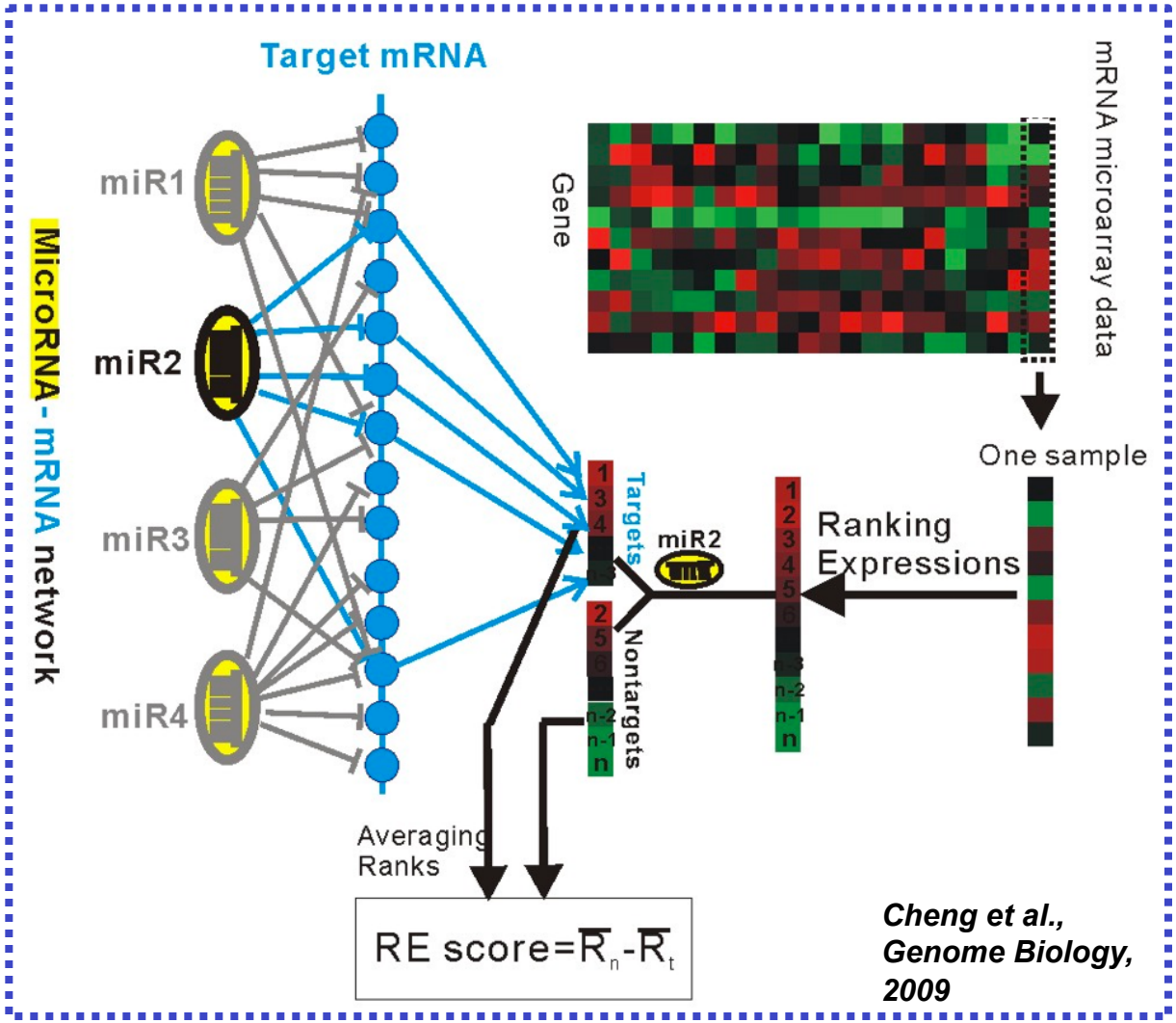


[Yu et al., PNAS (2006)]

Finding Central Points in Networks #3: Points of Maximal Regulatory Effect

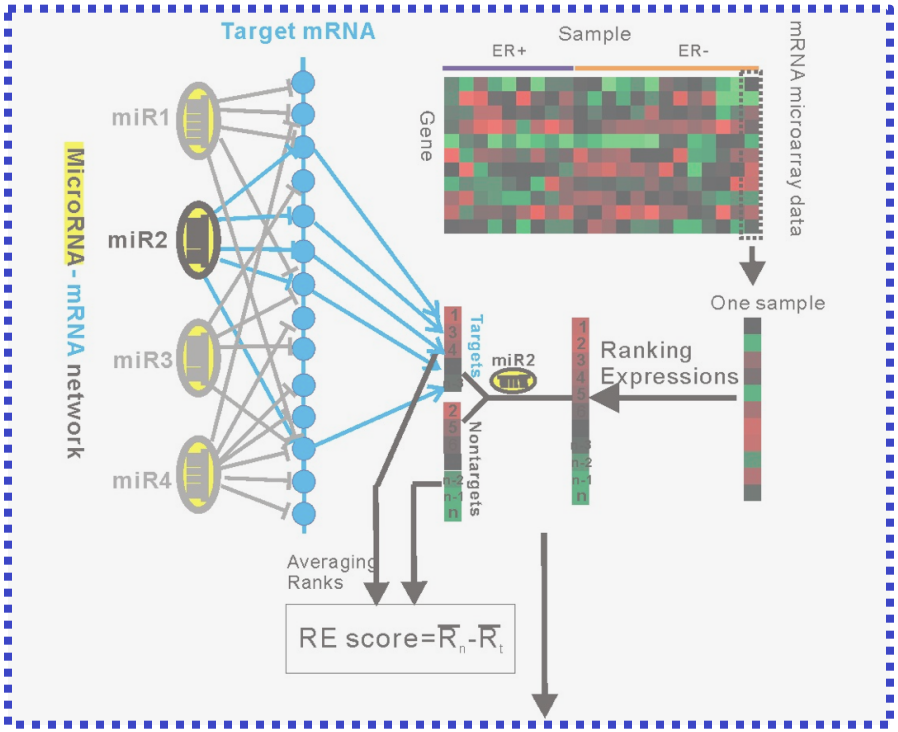
- How much does a regulator influence its targets?
- For miRNA-target networks easy to calculate, as all influence is down-regulation
  - ◊ target prediction via: TargetScan, PITA, PicTar, miRanda, …
- Look at down-reg. genes in a sample & compare with targets of a specific micro-RNA
  - ◊ more down-reg genes => stronger regulatory effect

RE-score: Another way to identify "important" network nodes

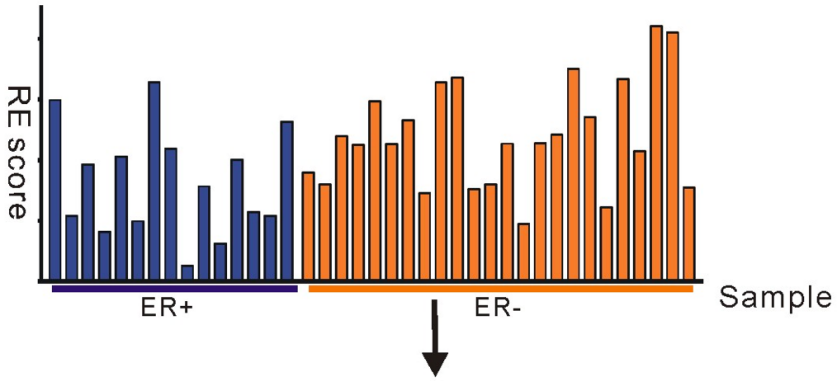


Application of RE-score to measure changing miRNA effect in different conditions

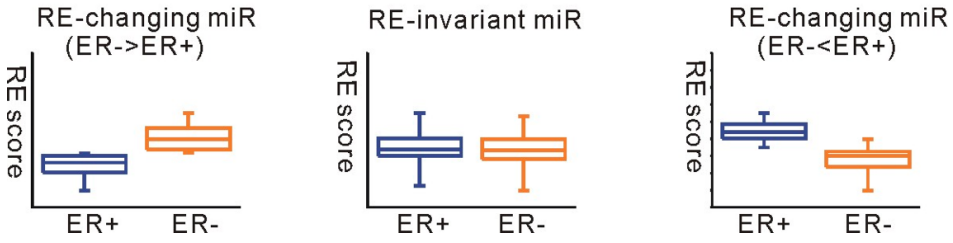
(ER- and ER+ breast cancer)



Calculating RE scores of a miRNA in each sample



Comparing the RE scores between ER+ and ER-

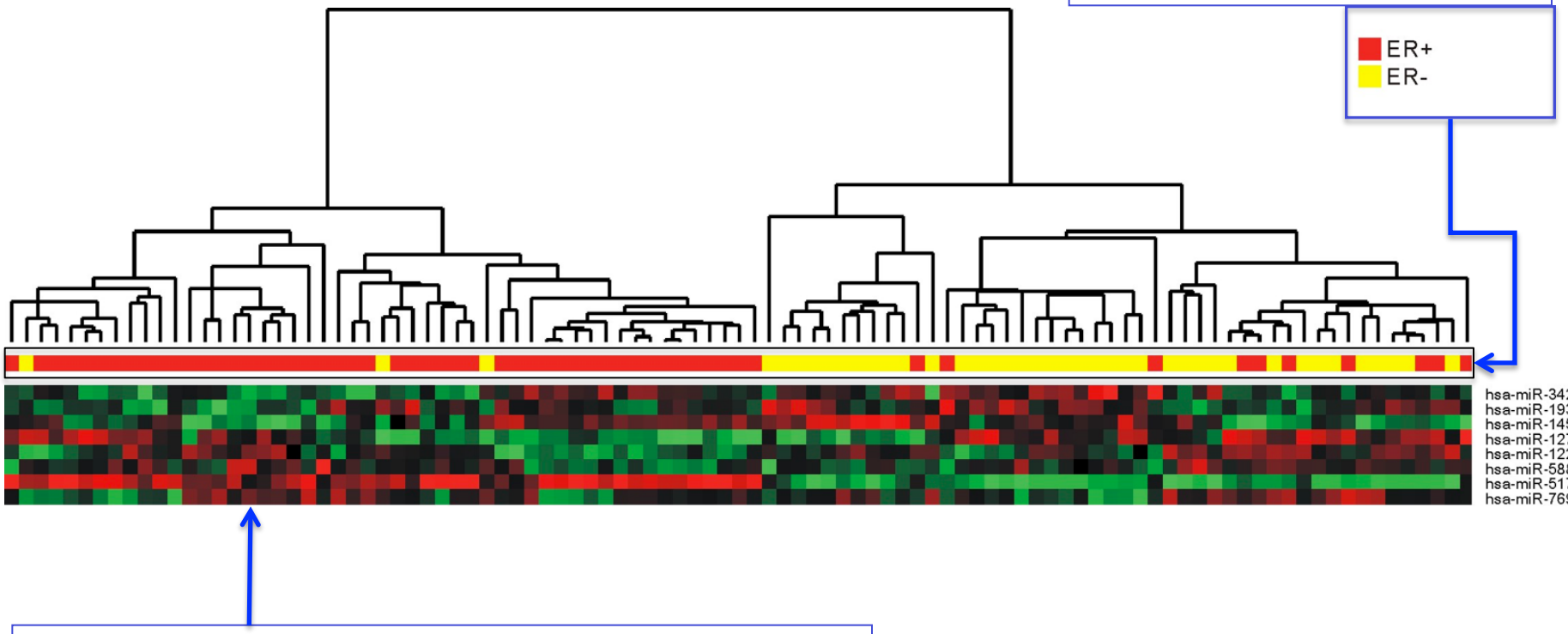


Cheng et al., Genome Biology, 2009

RE-score can be used to classify cancers

(3) Clustering based on RE score divides samples into 2 main types of cancer

(4) Clustering better than based on indiv. gene expression levels

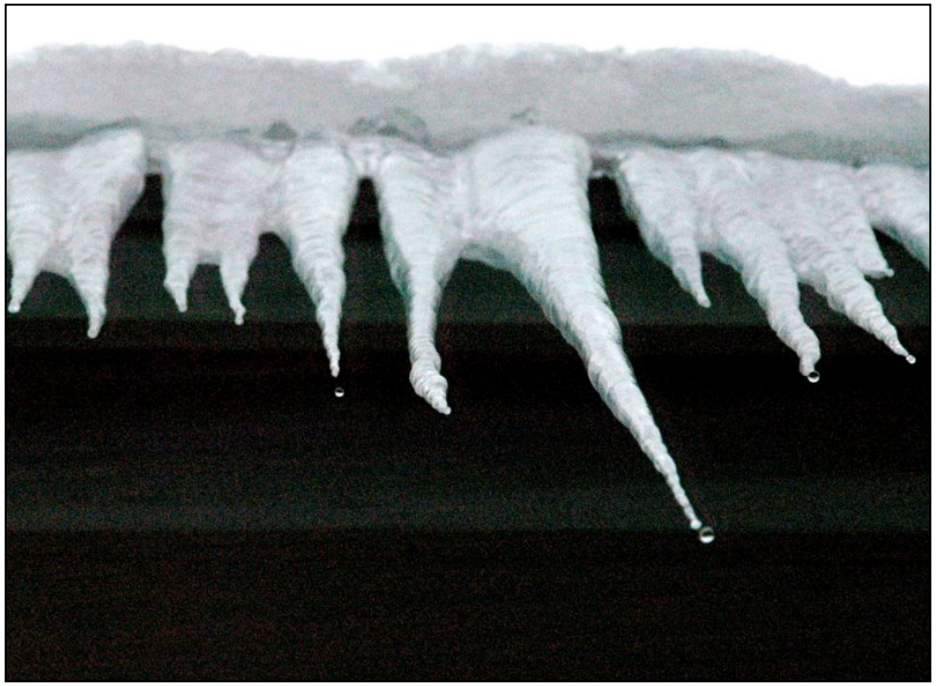


(1) RE-score profile for diff. miRNA in 1 cancer sample.
(2) Tabulate over many different breast cancer samples

Cheng et al., Genome Biology, 2009

Network Dynamics: Environments

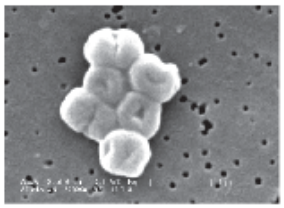
How do molecular networks change across environments?
What pathways are used more ?
Used as a biosensor ?



What is metagenomics?

Genomics Approach

Culture Microbes

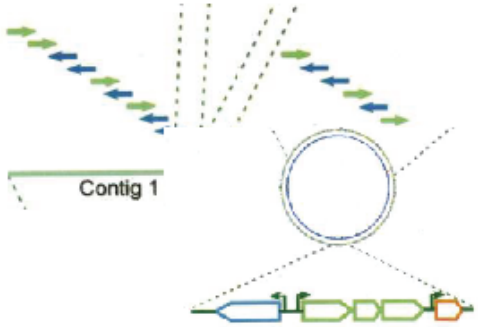


Extract DNA

Sequence

ATCGTATA
CGCGAAG
ACGTCTGA
AGTGCTGCT

Assemble and Annotate



PROBLEM: Estimated that less than 1% can be cultured in the lab

Metagenomics Approach

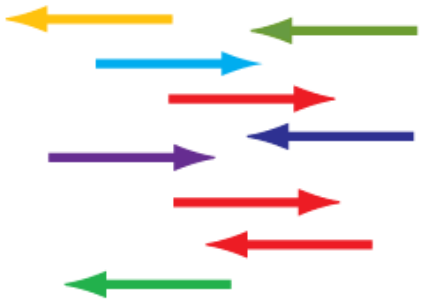
Collect Sample

Extract DNA

Sequence

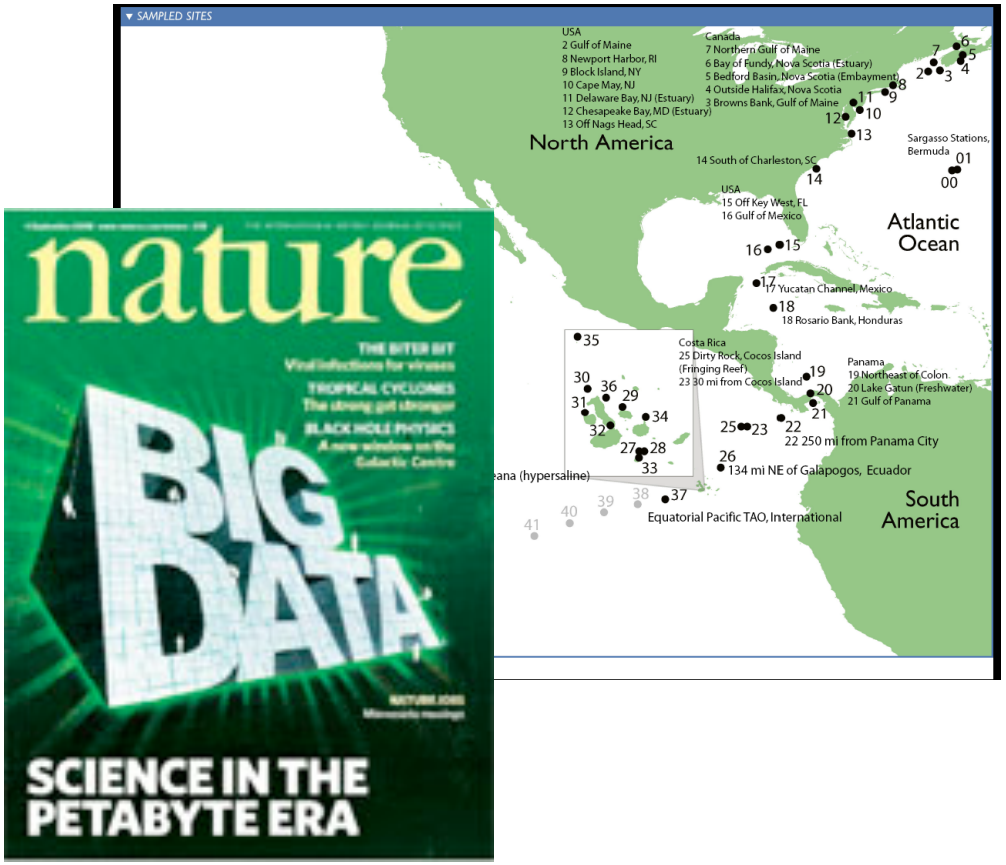
ATCGTGATAGATGATAGTAGA
ATGCTGCATGCATCTAGCACT
ACAGTAGCTAGCTACGTACTA
CAGCTGACTAGCTAGCTAGCT
ACGTAGCATGCTAGCTAGCAG
ACGTACGTAGCTAGCTAGCTAG
ACGTACGTACGTAGCTAGCATC
AGTCGACTGAGCCAGTGATGAT
ACGATGCATGAGCAGATGCTAC
AGATCGTAGCATGCTAGCATGCT
ACGTACGTAGCTAGCTAGCTAAG
AGCTAGCATGCTAGTAGCATGAG
ACGATGCTAGCTAGCTAGCTGATA
TCGATCAGCATGCTACGATGCAAG
ACGATCGATGCTAGCTAGCTAGCAT
AGCTAGCTAGTCAGCTAGCTAGATG

Partially Assemble and Annotate



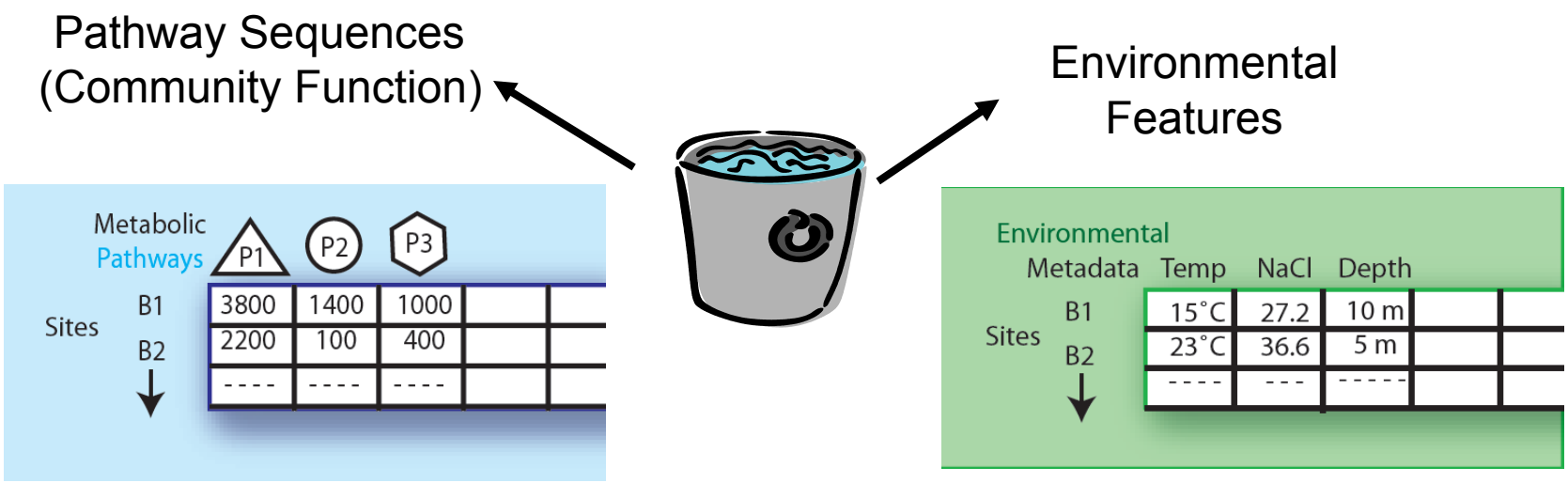
PROBLEM: Lose information about which gene belongs to which microbe.

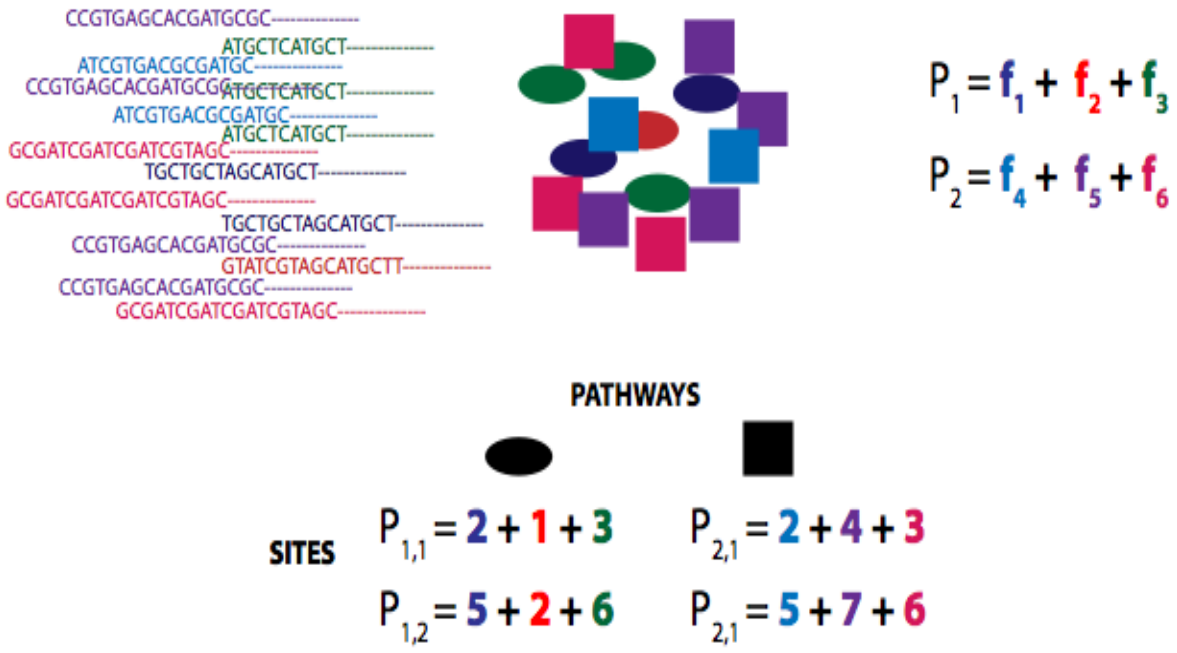
Global Ocean Survey Statistics (GOS)



6.25 GB of data
7.7M Reads
1 million CPU hours
to process

Rusch, et al., PLOS Biology 2007



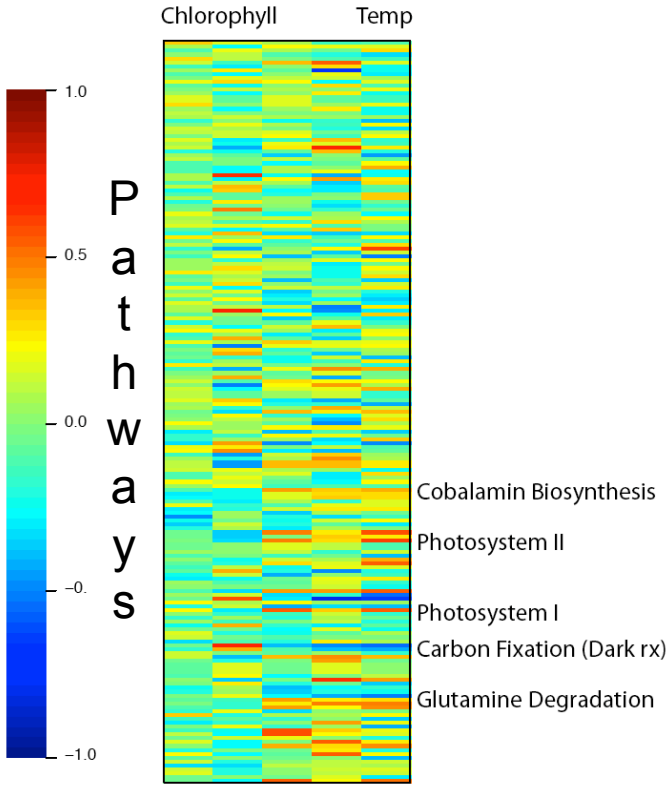


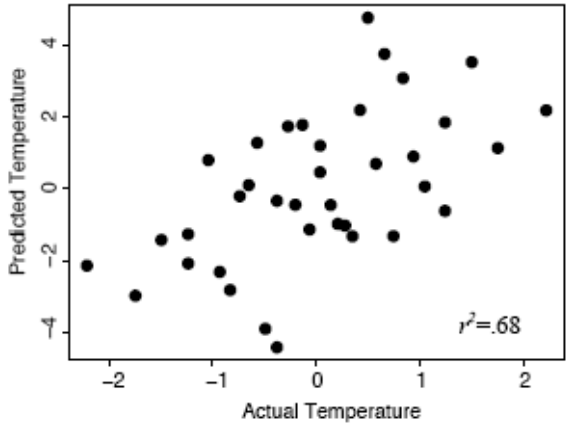
Expressing data as matrices indexed by site, env. var., and pathway usage

[Rusch et. al., (2007) PLOS Biology; Gianoulis et al., PNAS (in press, 2009]

Simple Relationships: Pairwise Correlations

Environmental Features





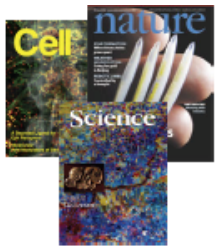
[Gianoulis et al., PNAS (in press, 2009)]

Canonical Correlation Analysis: Simultaneous weighting

Score	# of papers published
GRE	

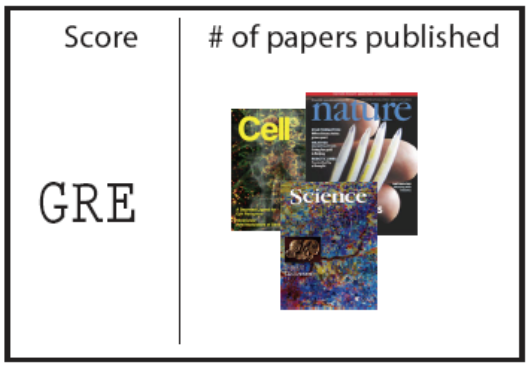
Undergraduate Performance Index (UPI)	Graduate School Performance Index (GPI)
GRE GPA	

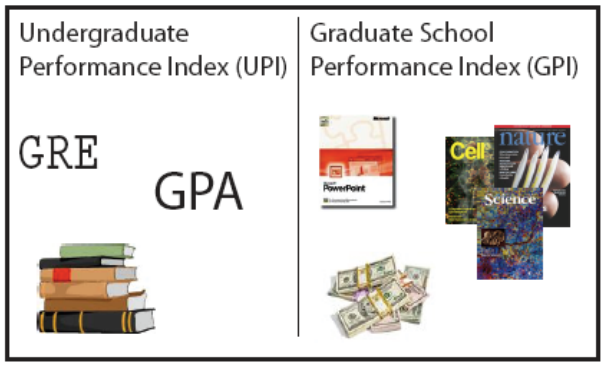
UPI = a GRE + b + c GPA

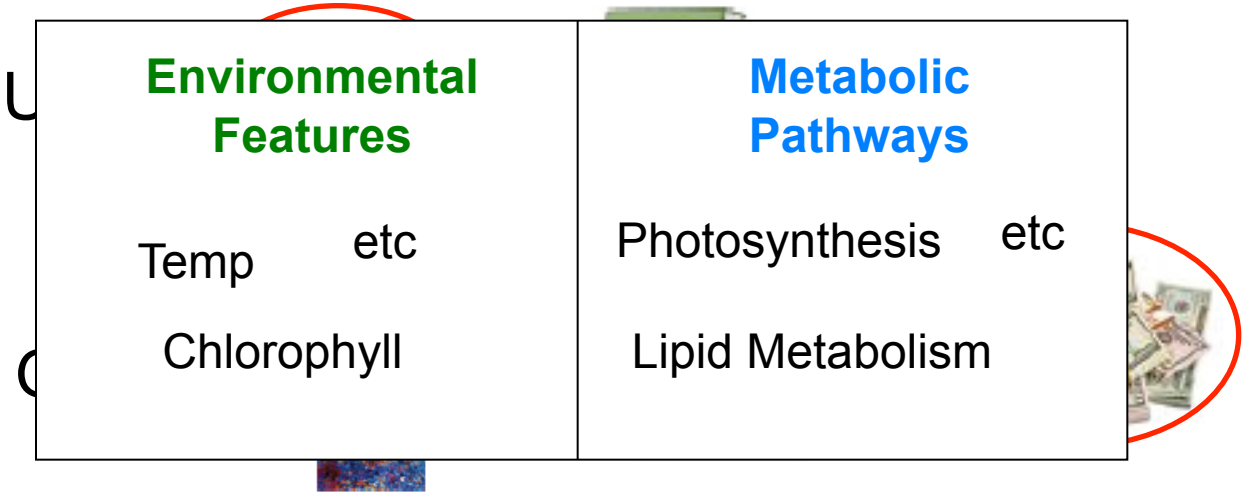
GPI = a'  + b' + c'

[Gianoulis et al., PNAS (in press, 2009)]

Canonical Correlation Analysis: Simultaneous weighting

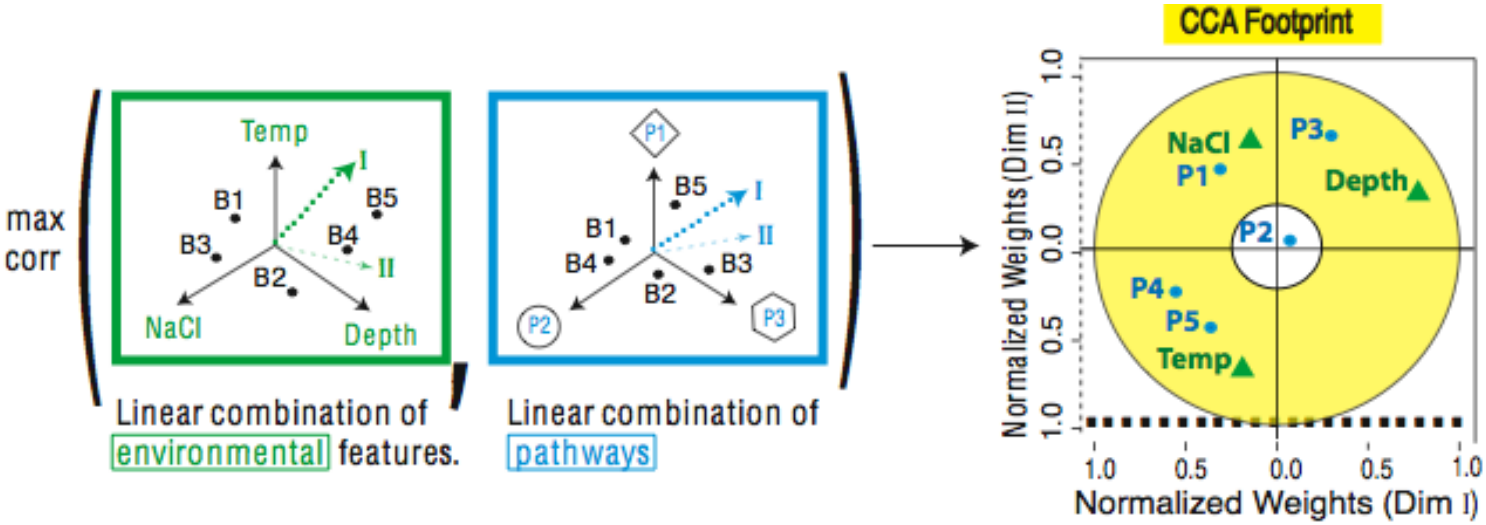






[Gianoulis et al., PNAS (in press, 2009)]

Environmental-Metabolic Space



The goal of this technique is to interpret cross-variance matrices
We do this by defining a change of basis.

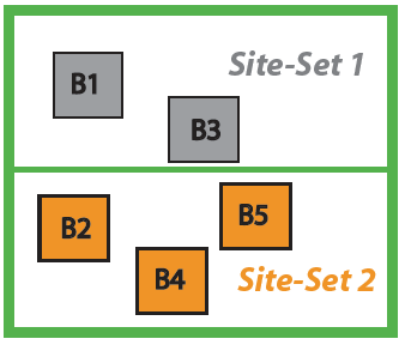
$$\text{Given } X=\{x_1,x_2,\ldots,x_n\} \text{ and } Y=\{y_1,y_2,\ldots,y_m\}$$

$$C=\begin{matrix}\Sigma_X & \Sigma_{X,Y}\\ \Sigma_Y & \Sigma_{Y,X}\end{matrix} \qquad \max_{a,b} Corr(U,V)=\frac{a'\Sigma_{12}b}{\sqrt{a'\Sigma_{11}a}\sqrt{b'\Sigma_{22}b}}$$

[Gianoulis et al., PNAS (in press, 2009)]

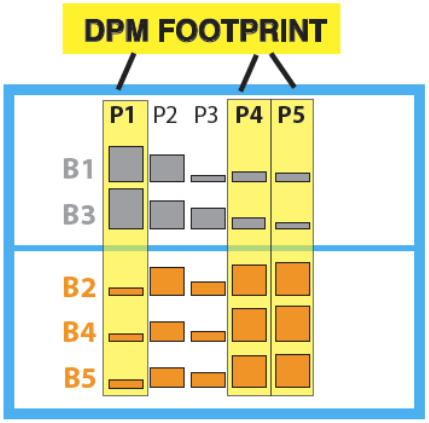
DPM: Discriminative Partition Matching

Environment



Cluster (Partition)

Metabolism

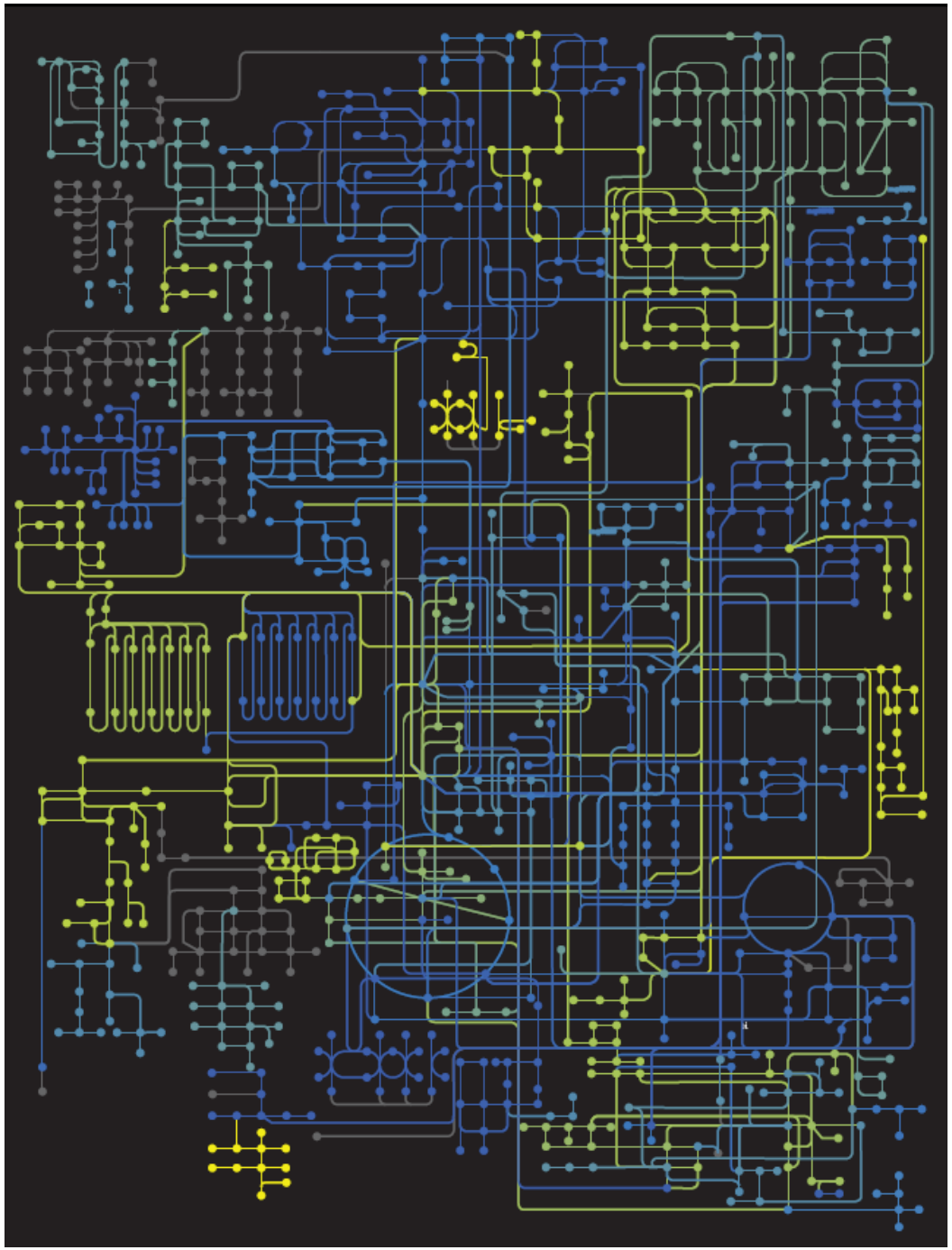


Test

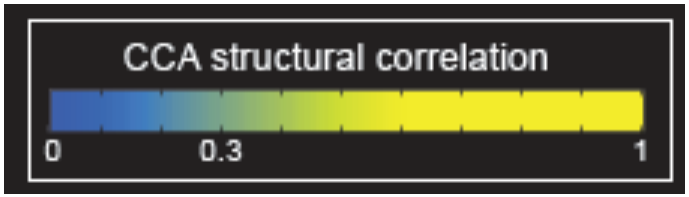
Taurine biosynthesis
Heme biosynthesis
Asparagine degradation
Nitrogen fixation
Acylglycerol degradation
Asparagine biosynthesis
Cysteine Metabolism

Functional class	pval
InfoStorage & Processing	.07
Cellular Process	.08
Metabolism	4x10-14

[Gianoulis et al., PNAS (in press, 2009)]

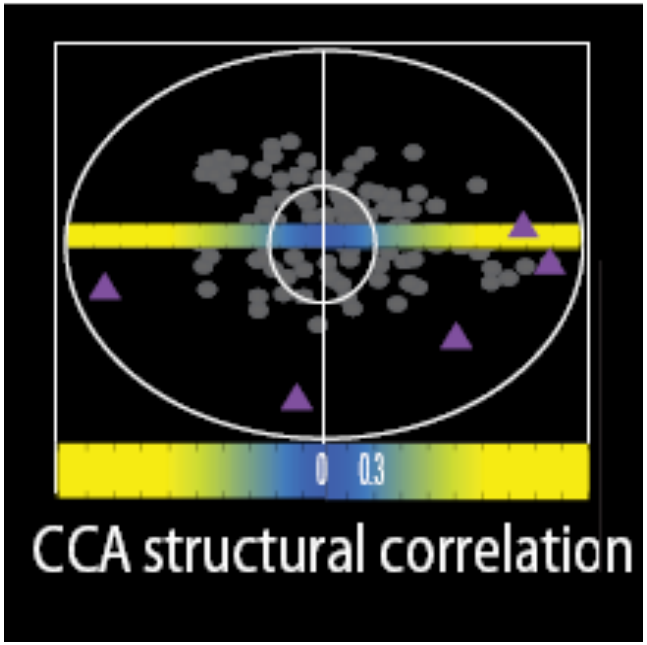


Strength of Pathway co-variation with environment



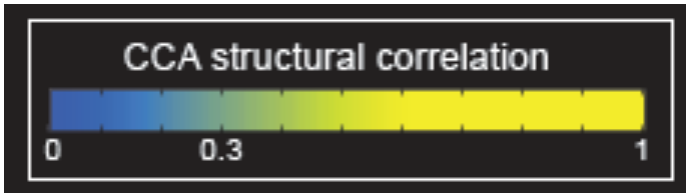
Environmentally invariant

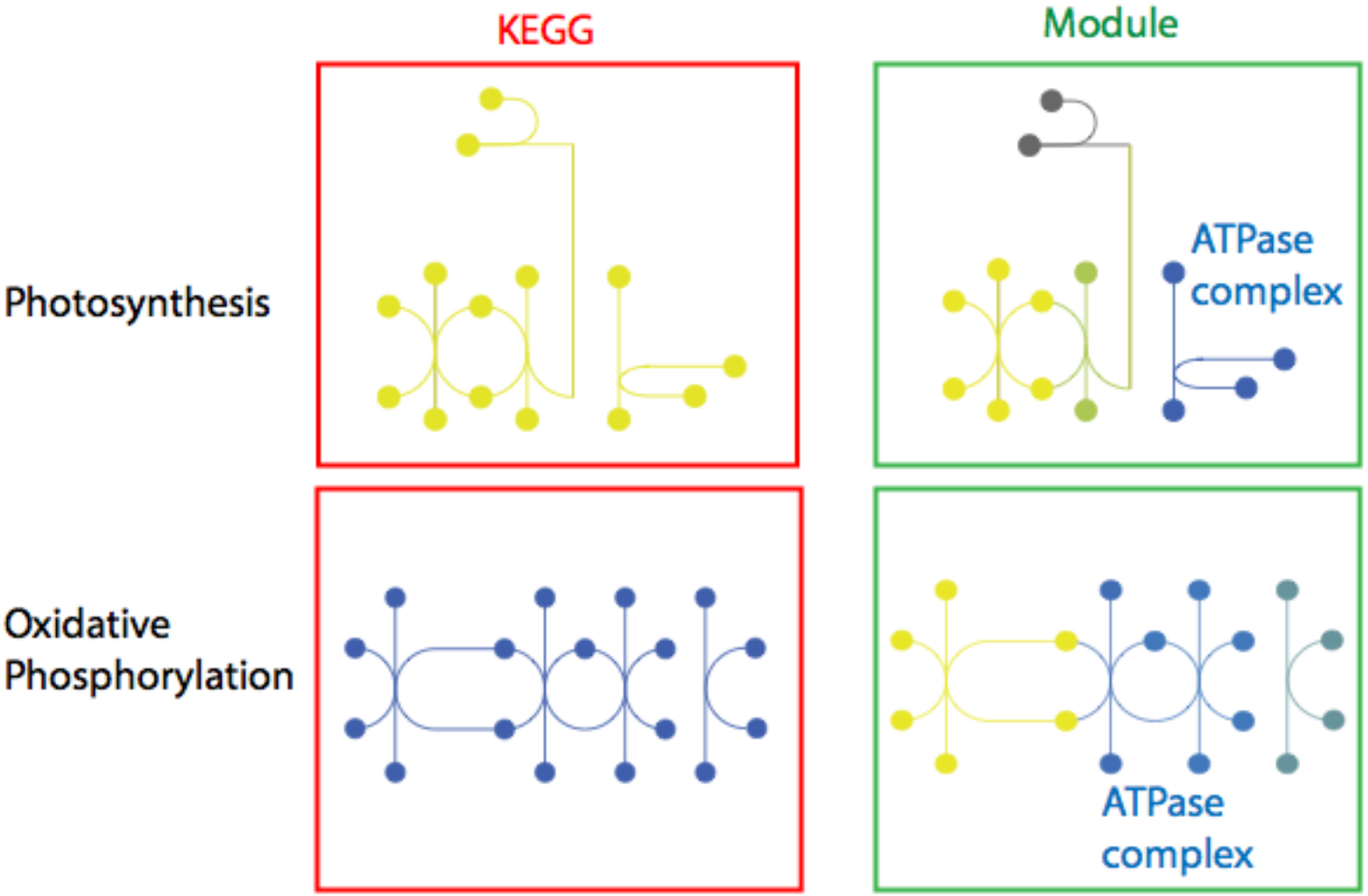
Environmentally variant



[Gianoulis et al., PNAS (in press, 2009)]

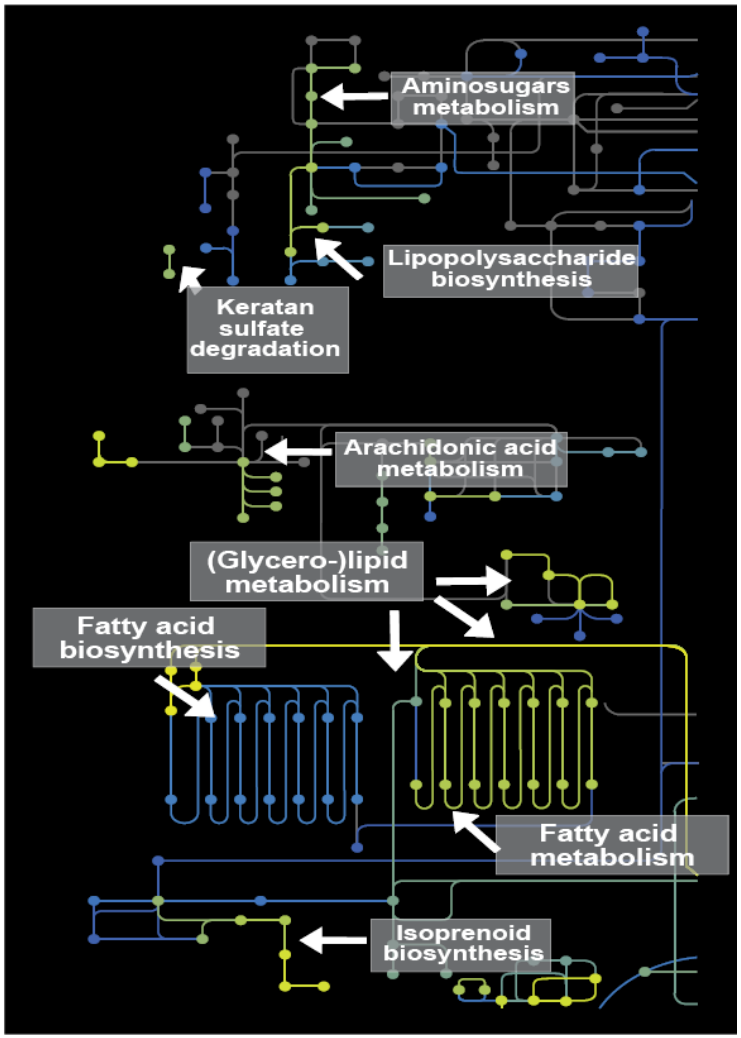
Conclusion #1: energy conversion strategy, temp and depth

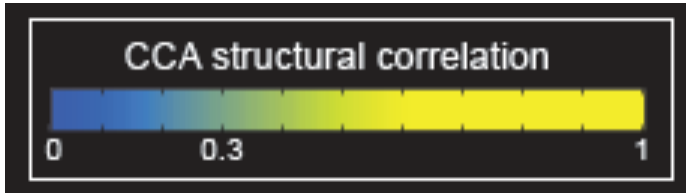




[Gianoulis et al., PNAS (in press, 2009)]

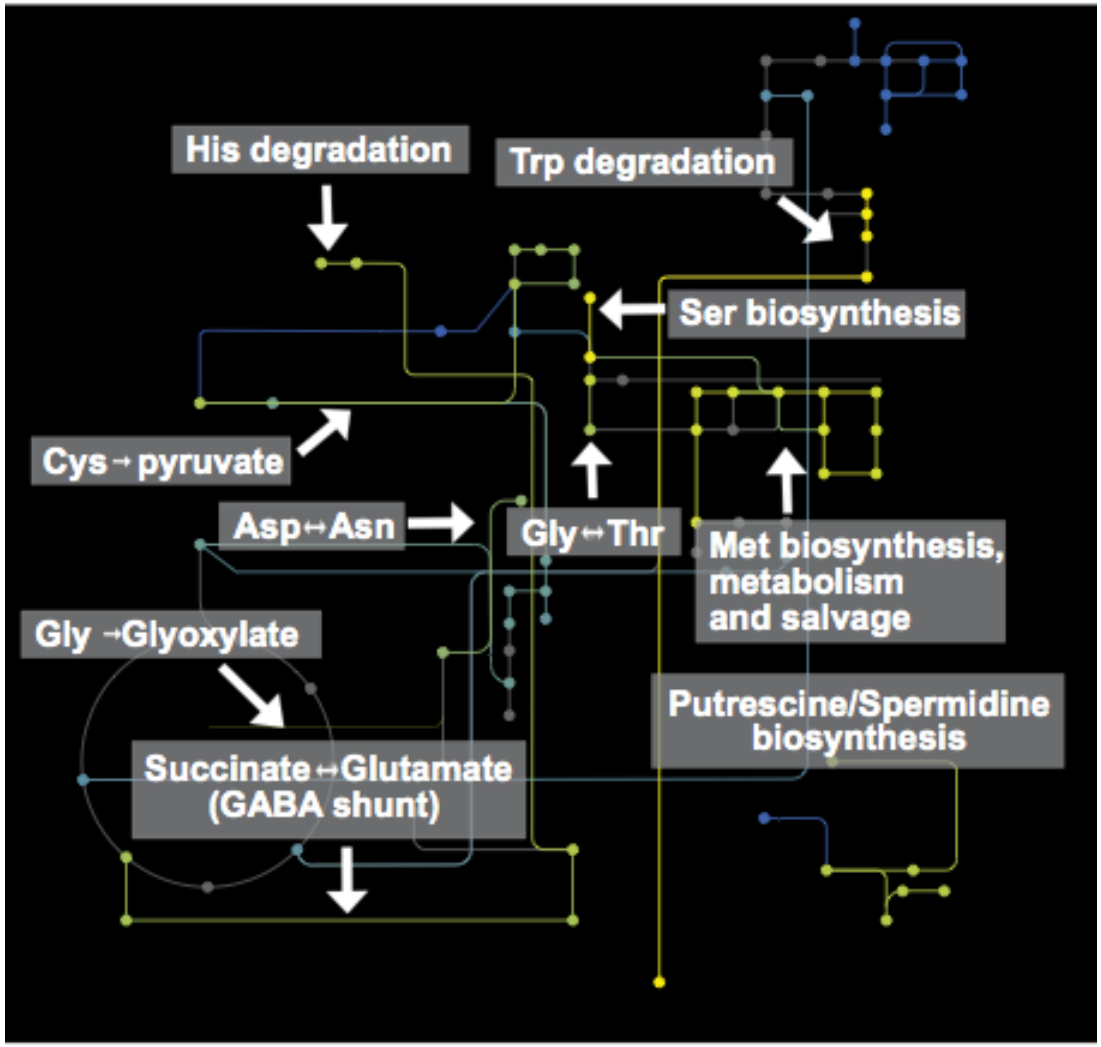
Conclusion #2: Outer Membrane components vary the environment





[Gianoulis et al., PNAS (in press, 2009)]

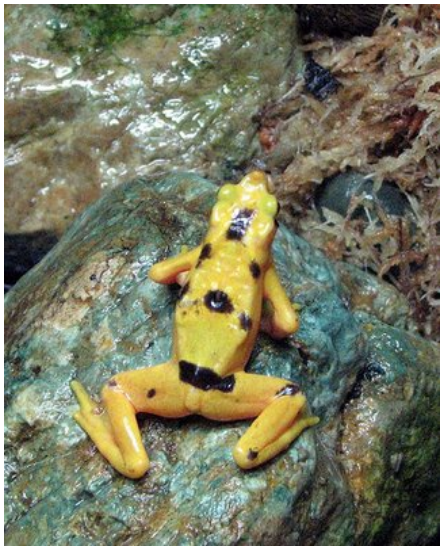
Conclusion #3: Covariation of AA biosynthesis and Import



Why is their fluctuation
in amino acid metabolism?
Is there a feature(s) that
underlies those that are
environmentally-variant
as opposed to those which are not?

[Gianoulis et al., PNAS (in press, 2009)]

Biosensors:
Beyond Canaries in a Coal Mine



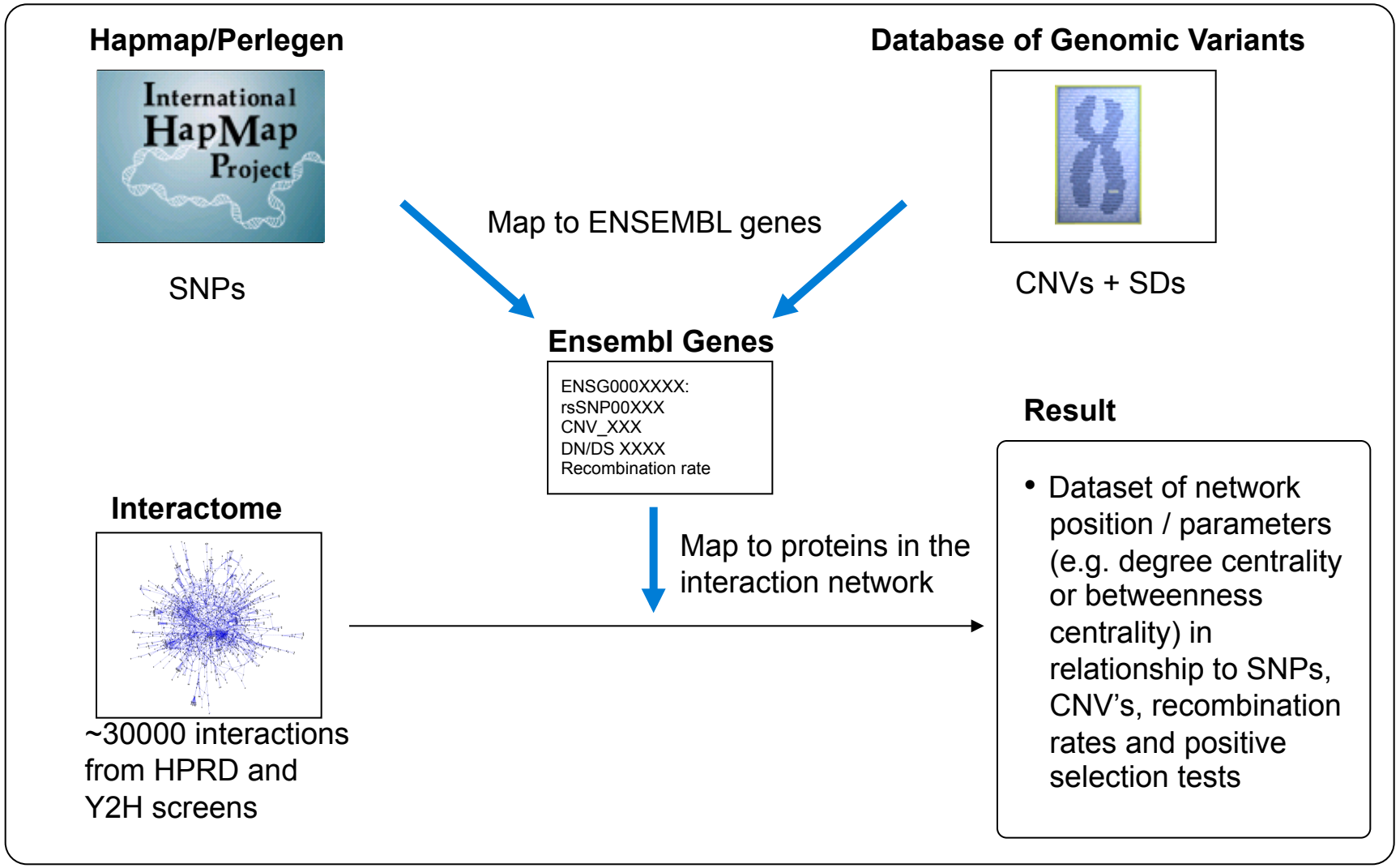
[Gianoulis et al., PNAS (in press, 2009)]

Networks & Variation

Which parts of the network vary most in sequence?
Which are under selection, either positive or negative?

METHODOLOGY: MAP SNP AND CNV DATA ONTO ENSEMBL GENES, AND THEN MAP ENSEMBL GENES TO THE KNOWN INTERACTOME

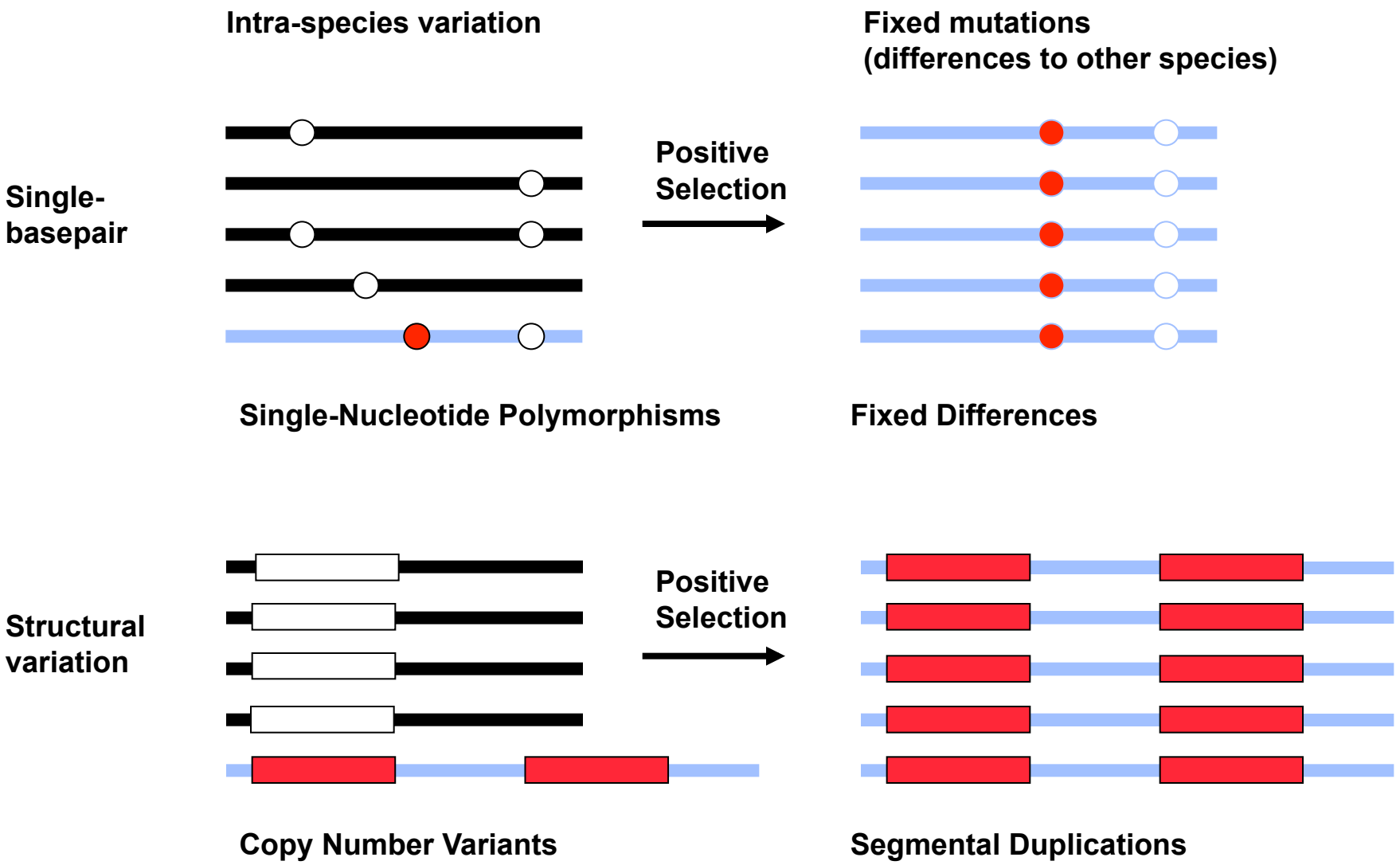
ILLUSTRATIVE



* From Nielsen et al. *PLoS Biol.* (2005) and Bustamante et al. *Nature* (2005)

Source: PMK

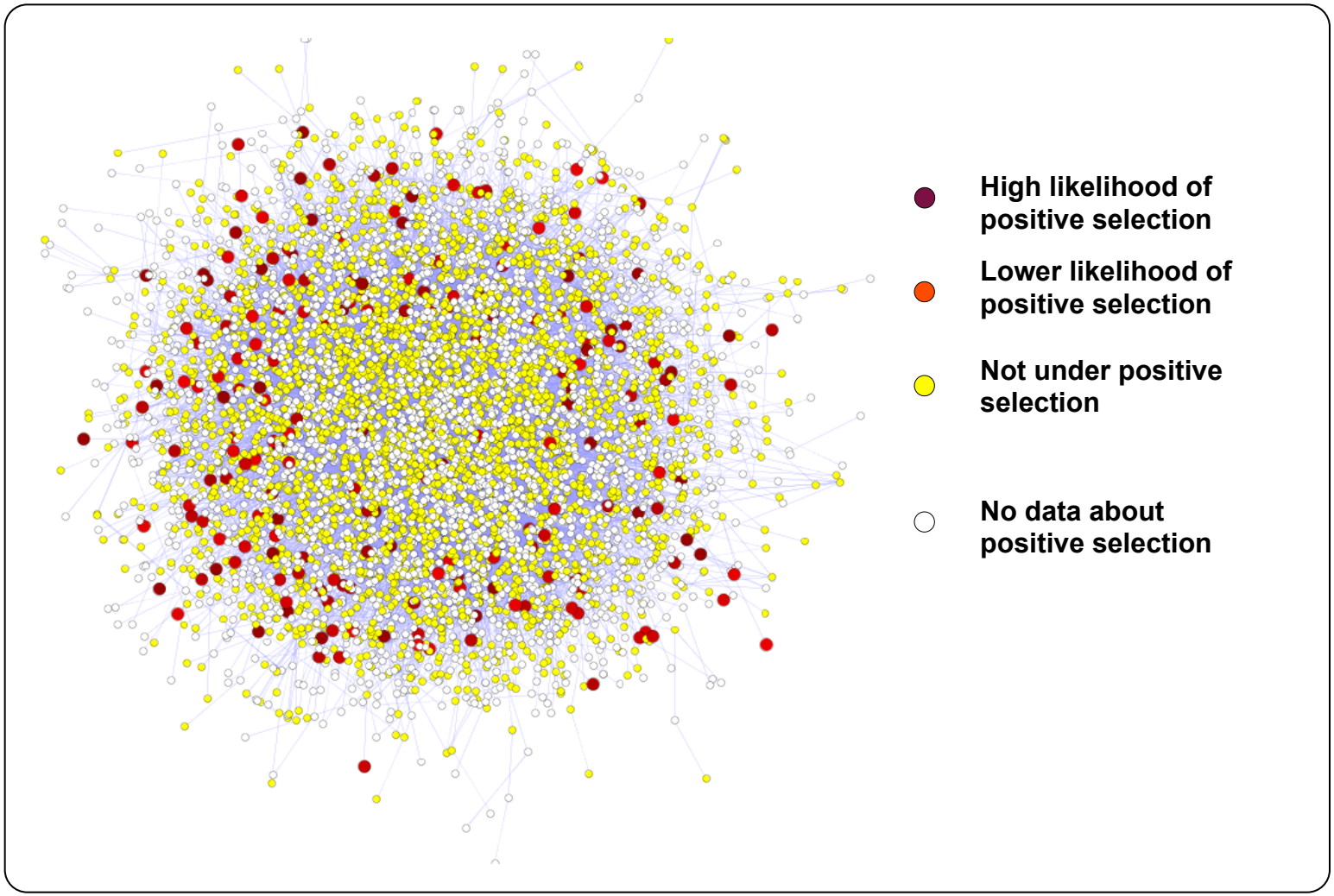
ADAPTIVE EVOLUTION CAN BE SEEN ON TWO DIFFERENT LEVELS



Source: PMK

POSITIVE SELECTION LARGELY TAKES PLACE AT THE NETWORK PERIPHERY

Positive selection in the human interactome

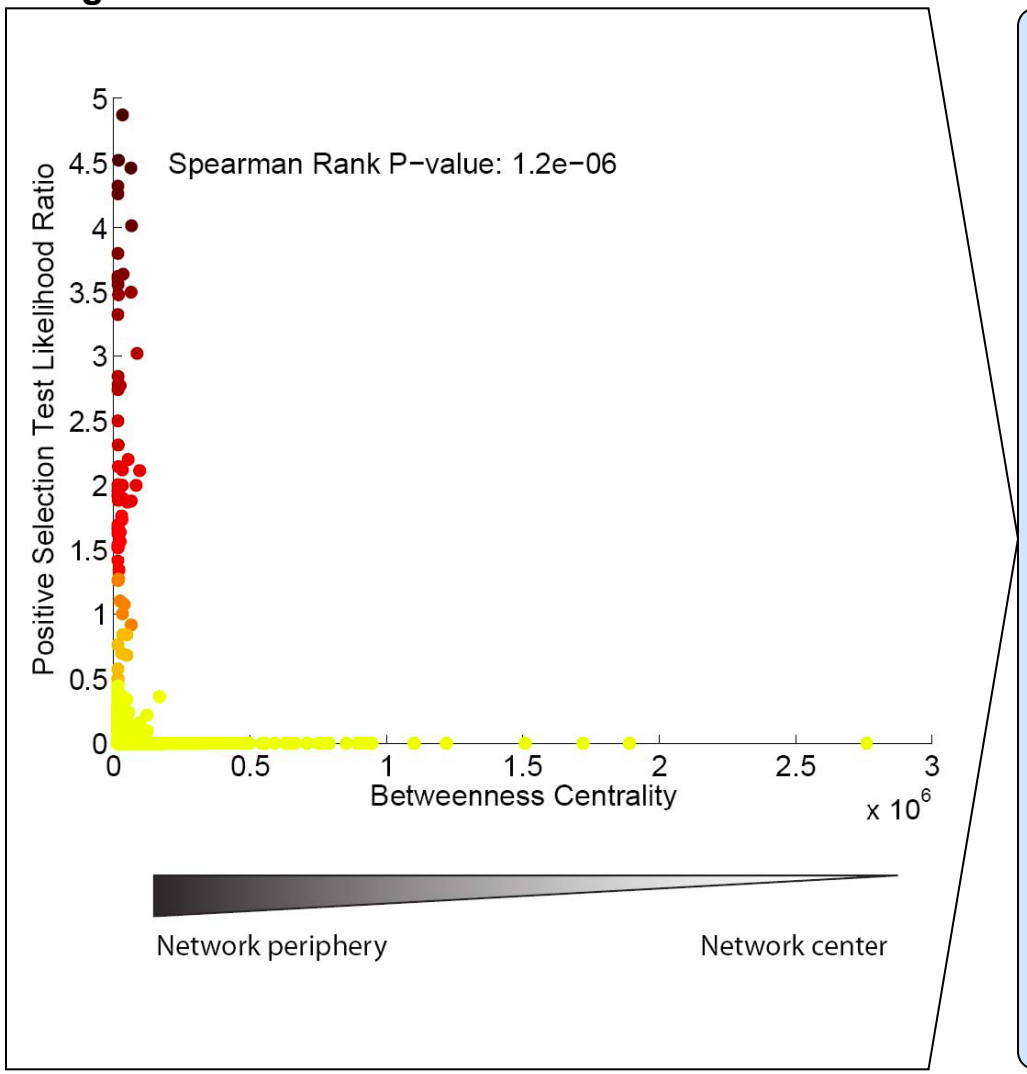


Source: Nielsen et al. *PLoS Biol.* (2005), HPRD, and Kim et al. PNAS (2007)

CENTRAL PROTEINS ARE LESS LIKELY TO BE UNDER POSITIVE SELECTION

Hubs

Degree vs. Positive Selection



Reasoning

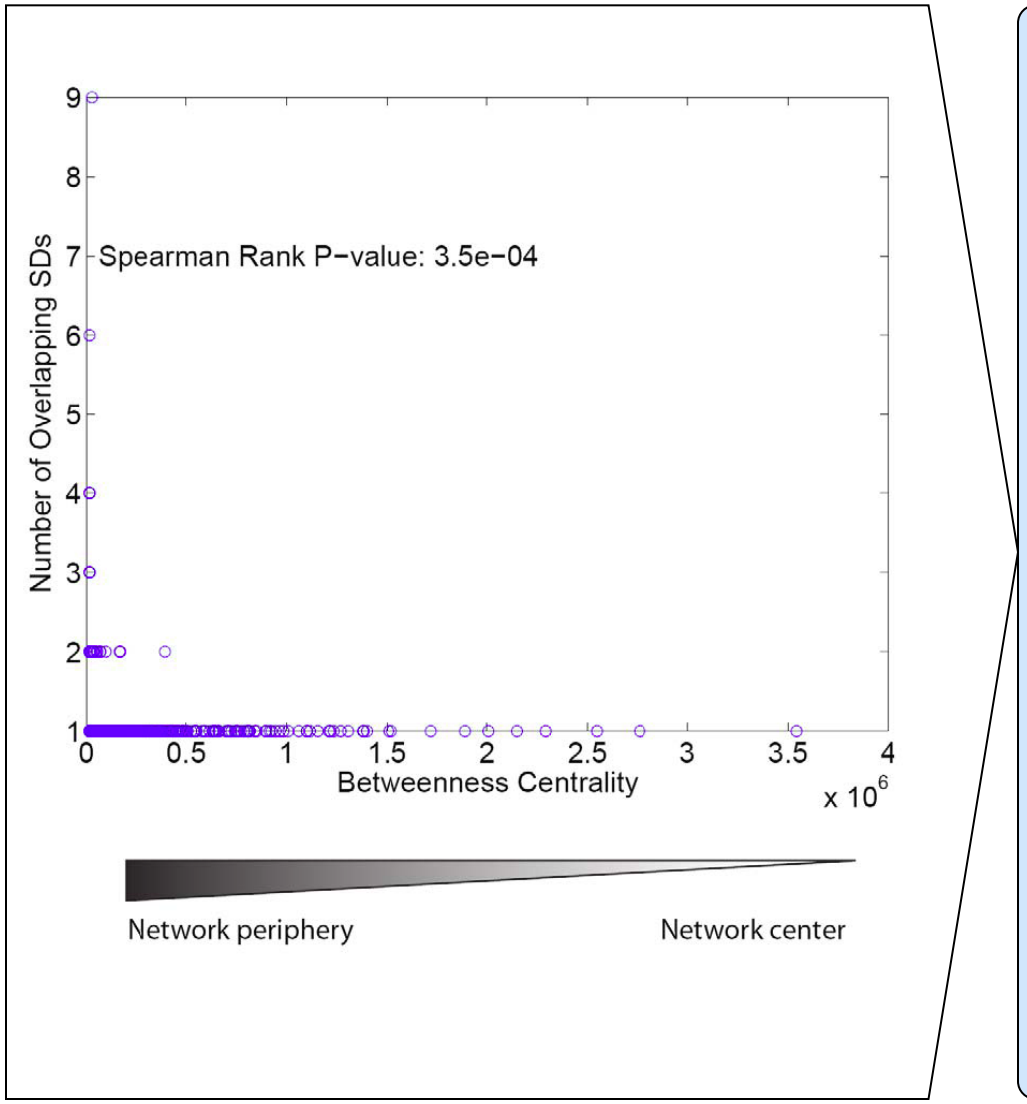
- Peripheral genes are likely to under positive selection, whereas hubs aren't
- This is likely due to the following reasons:
  - Hubs have stronger structural constraints, the network periphery doesn't
  - Most recently evolved functions (e.g. "environmental interaction genes" such as sensory perception genes etc.) would probably lie in the network periphery
- Effect is independent of any bias due to gene expression differences

* With a probability of over 80% to be positively selected as determined by Ka/Ks. Other tests of positive selection (McDonald Kreitmann and LDD) corroborate this result.

Source: Nielsen et al. *PLoS Biol.* (2005), Bustamante et al. *Nature* (2005), HPRD, Rual et al. *Nature* (2005), and Kim et al. PNAS (2007)

CENTRAL NODES ARE LESS LIKELY TO LIE INSIDE OF SDs

Centrality vs. SD occurrence



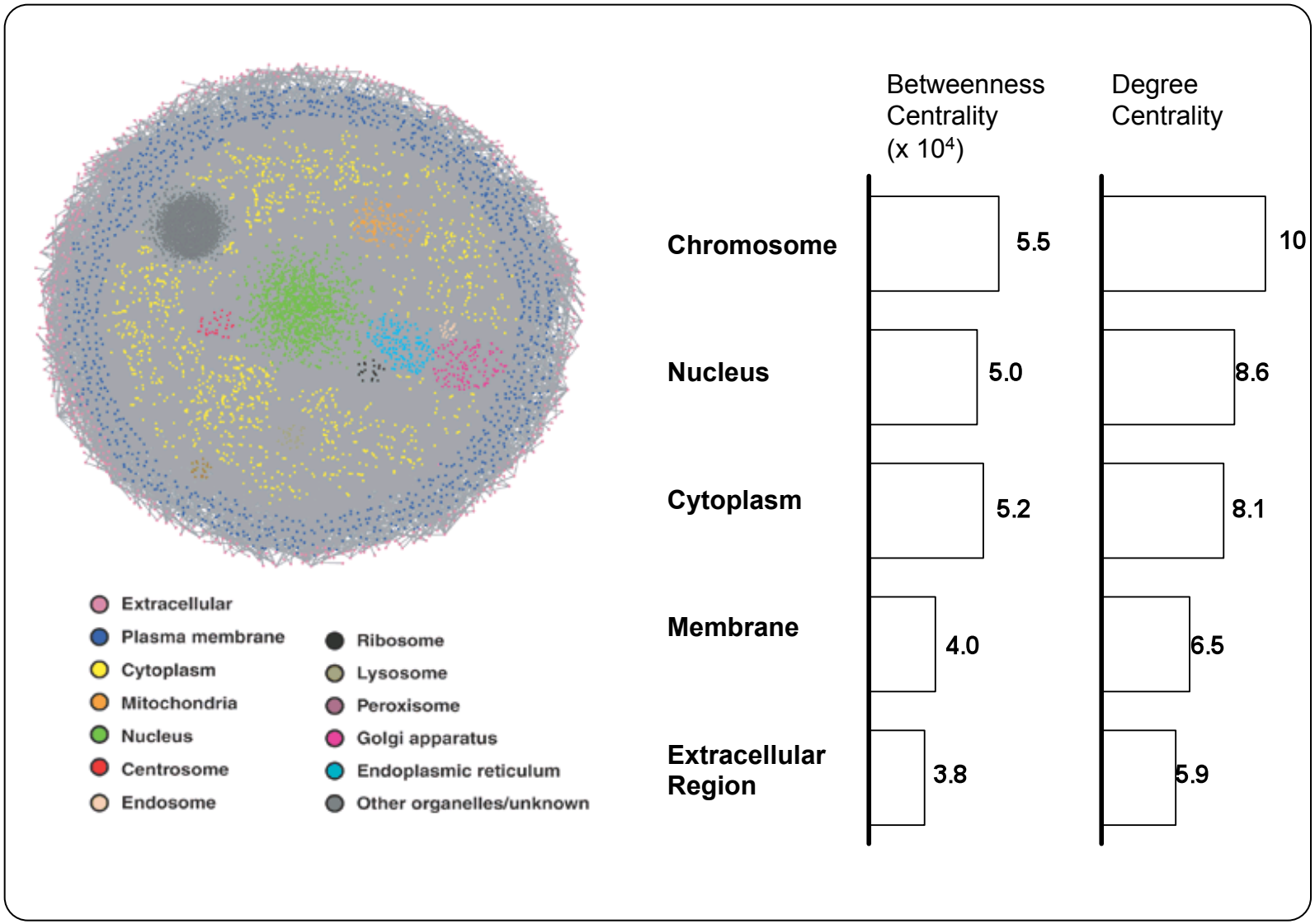
Reasoning

- This result also confirms our initial hypothesis – peripheral nodes tend to lie in regions rich in SDs.
- Since segmental duplications are a different mechanism of ongoing evolution, the less constrained peripheral proteins are enriched in them.
- Note that despite the small size of our dataset for known SD's we get significant correlations. It is to be expected that the correlations will get clearer as more data emerges*

* Specifically, a number of the SDs are likely not fixed, but rather common CNVs in the reference genome

Source: Database of genetic variation, HPRD, Rual et al. *Nature* (2005), and Kim et al. PNAS (2007)

Another explanation: THE NETWORK PERIPHERY CORRESPONDS TO THE CELLULAR PERIPHERY



Source: Gandhi et al. (*Nature Genetics* 2006), Kim et al. PNAS (2007)

IS RELAXED CONSTRAINT OR ADAPTIVE EVOLUTION THE REASON FOR THE PREVALENCE OF BOTH SELECTED GENES AND SDs AT THE NETWORK PERIPHERY?

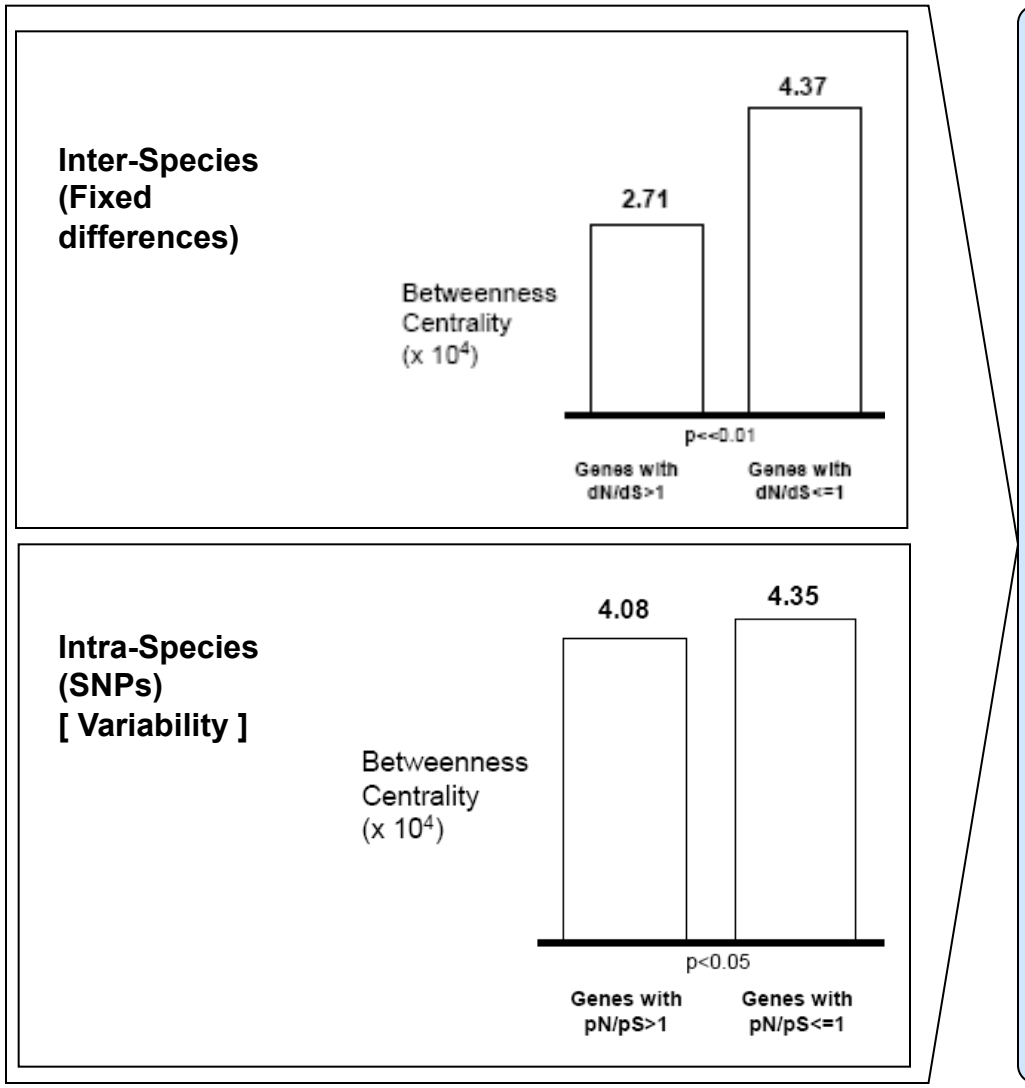
ILLUSTRATIVE

	Relaxed Constraint	Adaptive Evolution
Inter-Species Variation (Fixed differences)	• Increases inter-species variation – more variable loci are under less negative selection • Can be seen in higher Ka/Ks ratio or SD occurrence	• Increases inter-species variation – more variable loci are under less negative selection • Can be seen in higher Ka/Ks ratio or SD occurrence
Intra-Species Variation (Polymorphisms)	• Increases intra-species variation – for the very same reason • Can be seen in both SNPs or CNVs	• Should not have effects on intra-species variation

Source: Kim et al. PNAS (2007)

SOME, BUT NOT ALL OF THE SINGLE-BASEPAIR SELECTION AT THE PERIPHERY IS DUE TO RELAXED CONSTRAINT

Inter vs. Intra-Species Variation in Networks



Reasoning

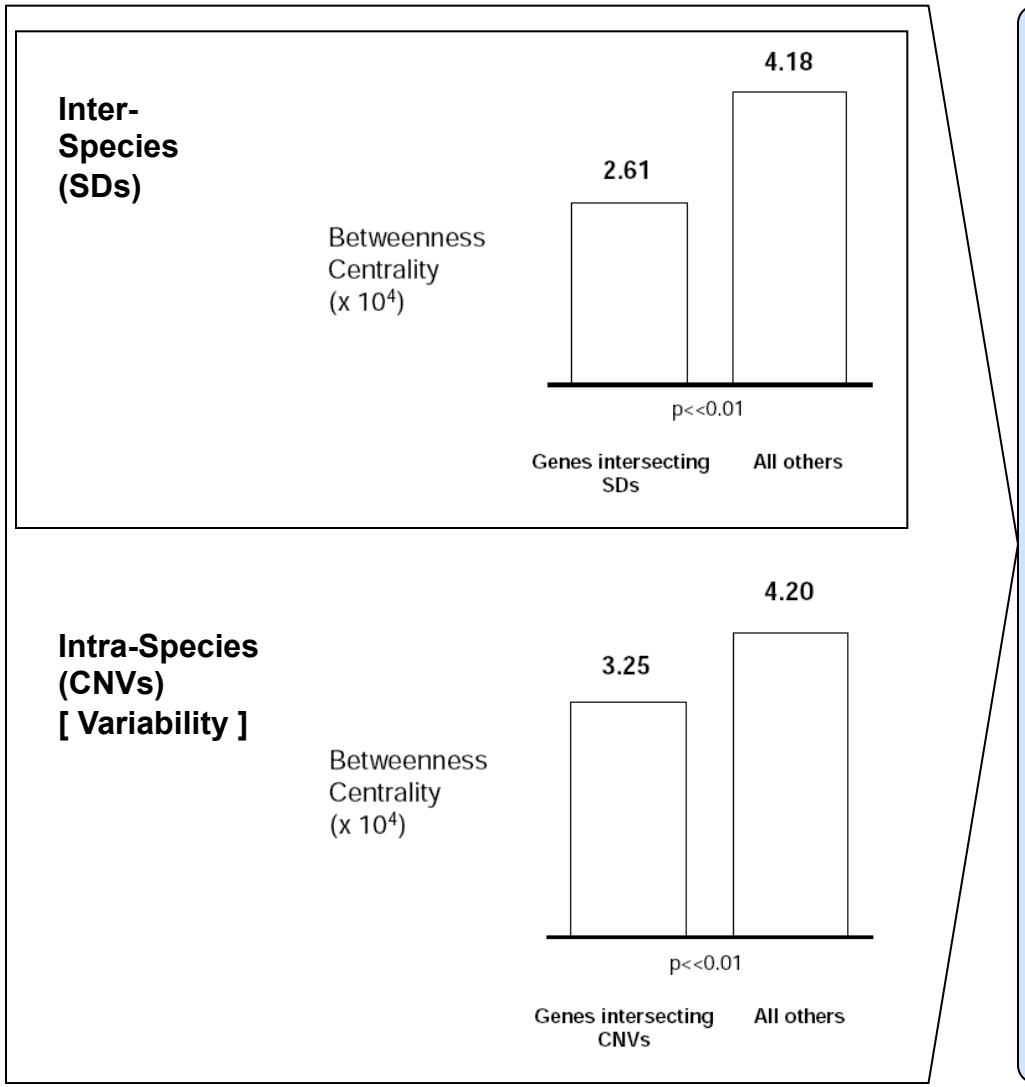
- There is a difference in **variability** (in terms of SNPs) between the network periphery and the center
- However, this difference is much smaller than the difference in **selection**
- This most likely means, that part of the effect we're seeing is due to relaxed constraint (and higher variability)
- But, not the entire effect*

* But it's hard to quantify

Source: Kim et al. (2007) PNAS

Similar Results for Large-scale Genomic Changes (CNVs and SDs)

Inter vs. Intra-Species Variation in Networks



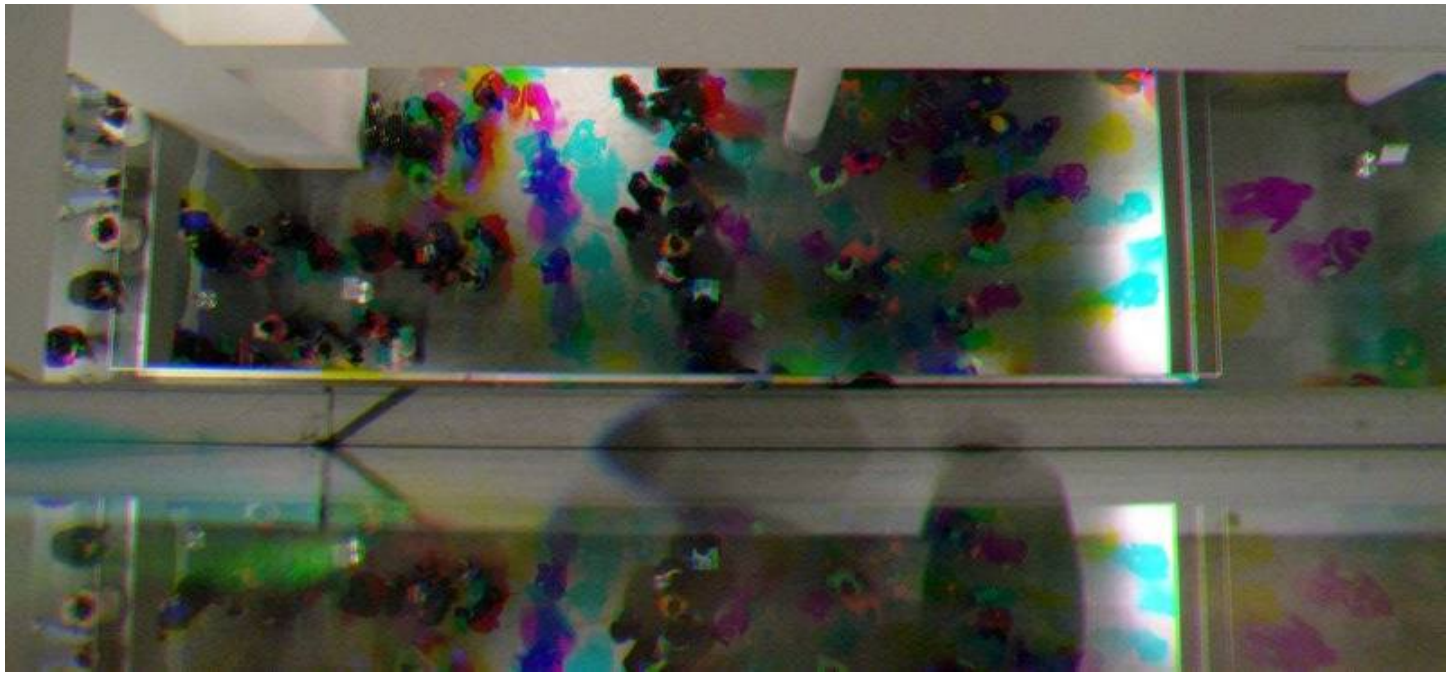
Reasoning

- There a small difference in **variability** (in terms of CNVs) between the network periphery and the center
- But, there is a (as shown before) marked difference in fixed (and hence, presumably, **selected**) SDs at the network periphery and center

Source: Kim et al. (2007) PNAS

Networks & Variation 2

Variation in the miRNA network



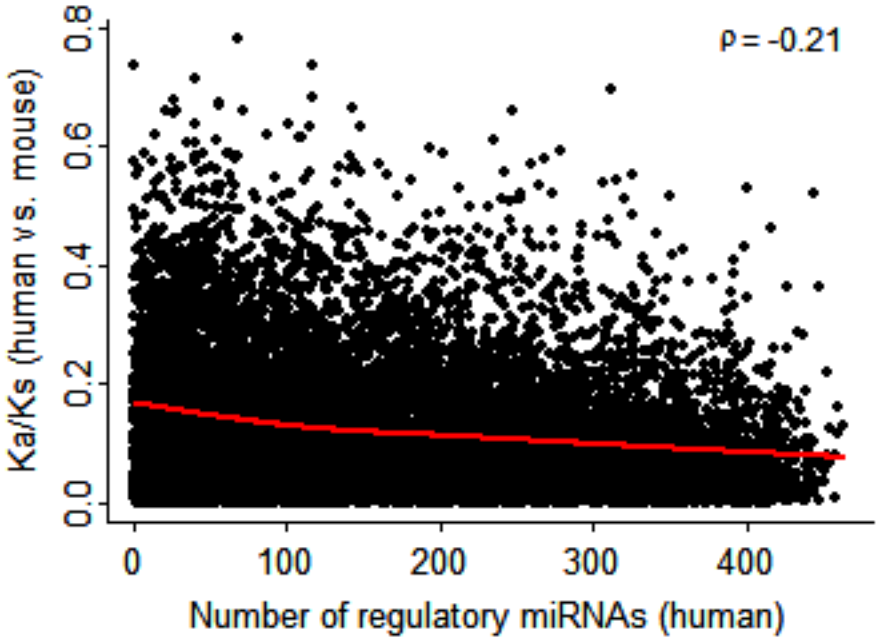
Analyze Regulation in microRNA-target Network

- Relationship between target in degree (number of micro-RNAs that regulate gene) & evolutionary rate of gene?
  - ◊ In deg. related 3' UTR size

- Expectation: more regulation, more constraint

65 Lectures.GersteinLab.org (c) 2009

Relationship between microRNA regulation and protein evolution



Human vs.	Number of genes	Correlation	P-value
chimpanzee	11326	-0.11	2.E-32
mouse	13280	-0.21	7.E-128
rat	12270	-0.20	4.E-107
cow	11683	-0.21	8.E-115
chicken	8061	-0.18	1.E-57

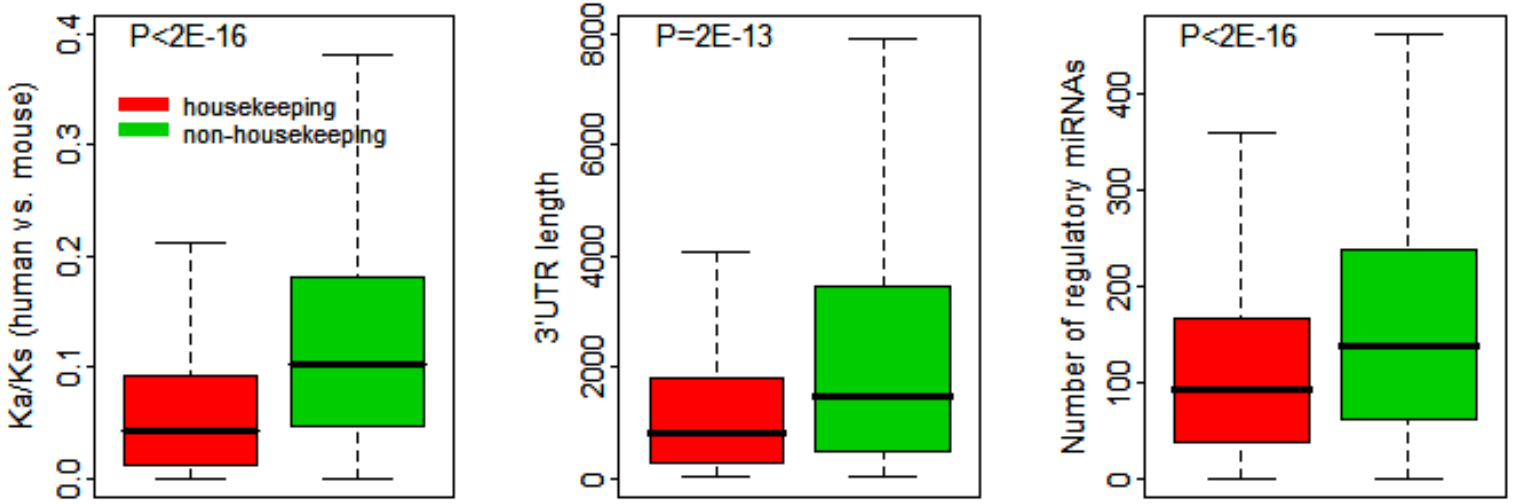
Important genes are regulated more intensively regulated by the microRNAs

[Cheng et al., BMC Genomics, 2009 (in press)]

MicroRNA regulation: a two-way strategy

For non-housekeeping genes, functionally critical genes are intensively regulated by miRNAs and prefer long 3'UTR.

housekeeping genes, however conserved, are selected to have shorter 3'UTRs to avoid miRNA regulation.

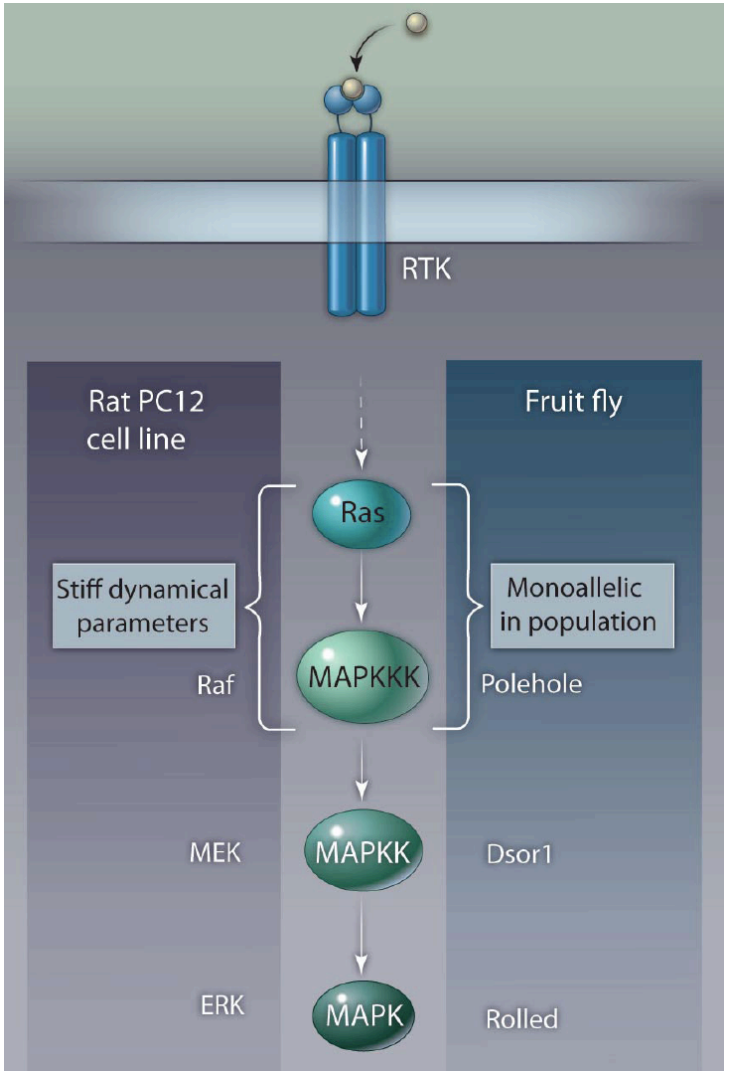


[Cheng et al., BMC Genomics, 2009 (in press)]

Analogies show it reasonable for more variable part of network to be periphery

- Computer Networks
  - Servers in center have much depending on them; thus, can't be frequently updated & patched
  - Servers on periphery often attacked and so need frequent patches
- Social Networks
  - Individuals at center under more constraint (to conform), whereas those at periphery have more freedom to experiment

Speculation: Why more tightly regulated gene might have less variation



Alexander et al. *Sci. Signal.* (2009) 2: pe44

Example: MAP Kinase singaling pathway

Dynamic model:
- ODE model with Michaelis-Menten kinetics
- parameters fit
  to time series data of protein activities
  in response to EGF and NGF
  from rat PC12 cell line

In sensitivity analysis,
stiff parameters cluster around Ras and Raf.

Population study in fruit flies:
- allele variation based on
  PCR of pathway genes

Ras and Raf have less allele variation
than other proteins in the network.

Brown et al. *Phys. Biol.* (2004) 1: 184
Riley et al. *Molec. Ecol.* (2003) 12: 1315

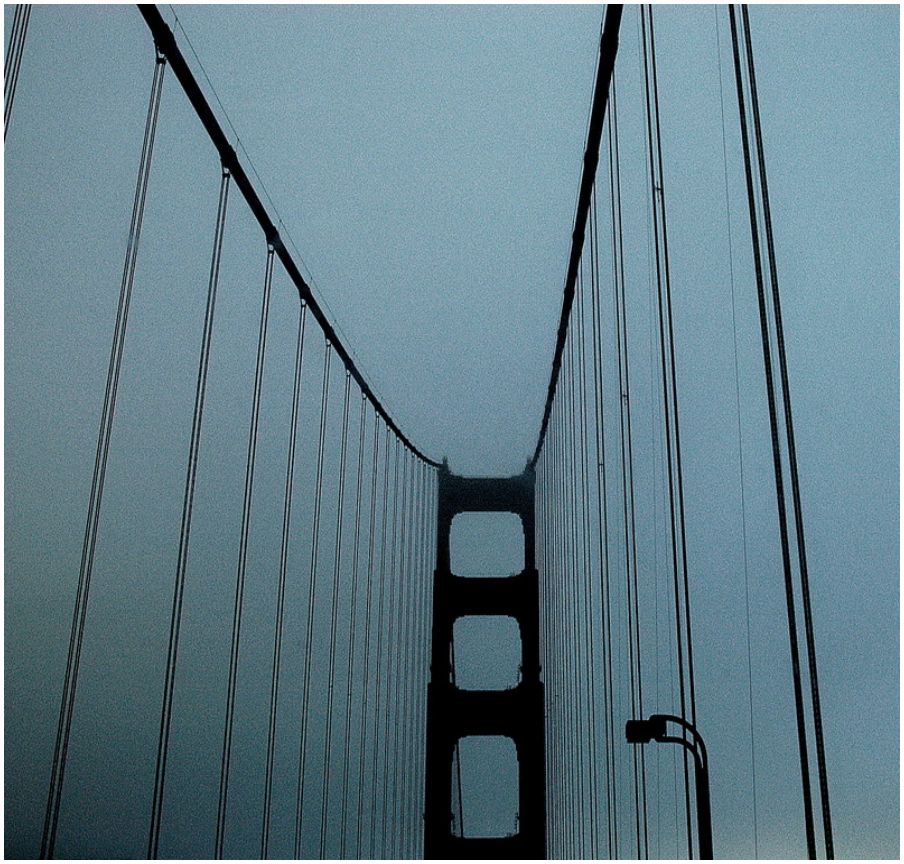
Outline: Molecular Networks

- Why Networks?
- Central Points in Networks
  - ◊ Hubs & Bottlenecks
    (yeast ppi & reg. net)
  - ◊ Tops of Heirarchies
    (yeast reg. net)
  - ◊ Identified by score
    (human miRNA-targ. net)
- Dynamics of Networks
  - ◊ Across environments
    (in prokaryote metab. pathways)
- Protein Networks & Variation
  (human ppi & miRNA-targ. net)

Conclusions:
Analysis of Network Structure

- Centrality Measures in Protein Network
  - ◊ Hubs & Bottlenecks
  - ◊ Importance of later in regulatory networks
- Regulatory Network Hierarchies
  - ◊ Middle managers dominate, sitting at info. flow bottlenecks
  - ◊ Paradox of influence and essentiality
  - ◊ Topmost proteins sit at center of interaction network

Conclusions:
Points of Network Centrality



- RE-score measures degree of (down) regulation of targets vs. non-targets
- Application to miRNA network
- Different RE-score of miRNAs can be used in cancer classification

Conclusions: Networks Dynamics across Environments

- Developed and adapted techniques to connect quantitative features of environment to metabolism.
- Applied to available aquatic datasets, we identified footprints that were predictive of their environment (potentially could be used as biosensor).
- Strong correlation exists between a community's energy conversion strategies and its environmental parameters (e.g. temperature and chlorophyll).
- Suggest that limiting amounts of cofactor can (partially) explain increased import of amino acids in nutrient-limited conditions.

73 Lectures.GersteinLab.org (c) 2009

Conclusions: Connecting Networks & Variation

- We find ongoing evolution (positive selection) at the network periphery.
  - ◊ This trend is present on two levels:
    - On a sequence level, it can be seen as positive selection of peripheral nodes
    - On a structural level, it can be seen as the pattern of SDs that display significantly higher allele frequencies in non-central genes
  - ◊ 2 possible mechanisms for this : adaptive evolution at cellular periphery & relaxation of structural constraints at the network periphery
    - We show that the latter can only explain part of the increased variability

74 Lectures.GersteinLab.org (c) 2009

Conclusions: Connecting Networks & Variation 2

- More highly regulated genes are under more constraint in miRNA-target networks
- Exception for housekeeping genes

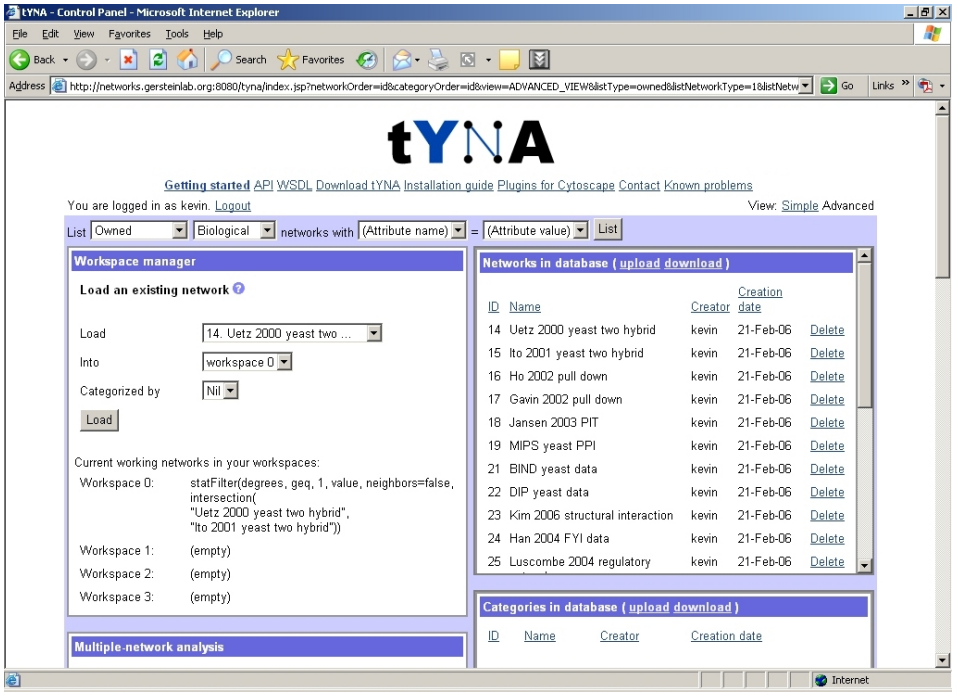
- Speculation as to why variation at periphery is quite reasonable

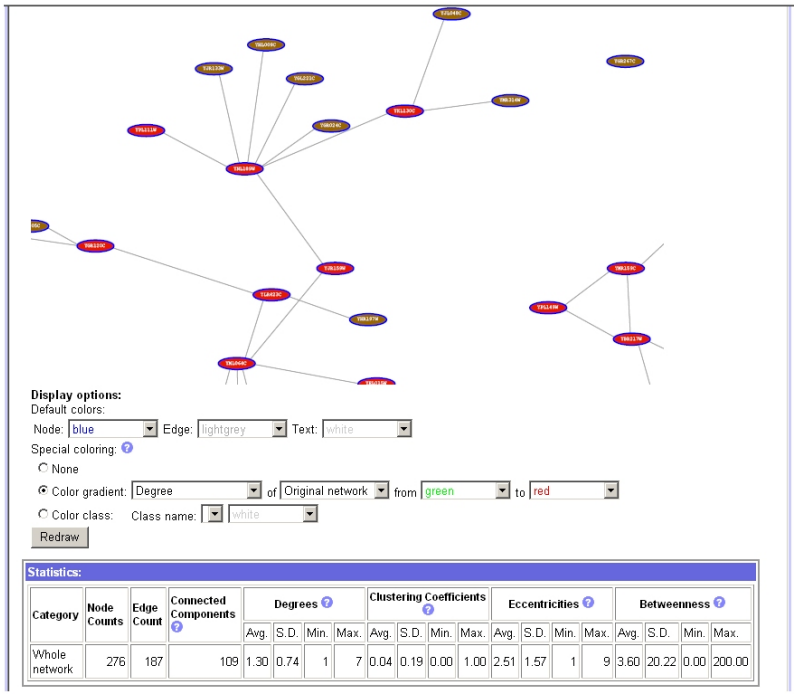
75 Lectures.GersteinLab.org (c) 2009

- an automated web tool

tYNA

(vers. 2 : "TopNet-like Yale Network Analyzer")





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]

Acknowledgements

H Yu
P Kim
K Yip
T Gianoulis
C Cheng

A Paccanaro
P Alves
T Emonet
P Cayting
M Seringhaus
Y Xia
J Korbel
A Sboner
P Patel
P Bork
J Raes
E Franzosa
M Snyder
N Bhardwaj
R Alexander

Networks.GersteinLab.org

Job opportunities currently for postdocs & students

More Information on this Talk

TITLE: Understanding Protein Function on a Genome-scale through the Analysis of Molecular Networks

SUBJECT: Networks

DESCRIPTION:
Network Biology: Understanding metabolic and protein interactions, VIB workshop on the future of proteome research, Ghent, Belgium; 2009.10.08, 9:30-10:10; [I:**VIB**] (Medium networks talk, shortened from [I:**MBINETS**].)

(PPT works on mac & PC and has many photos. Paper references in the talk were mostly from Papers.GersteinLab.org. The above topic list can be easily cross-referenced against this website. Each topic abbrev. which is starred is actually a papers "ID" on the site. For instance,
the topic **pubnet*** can be looked up at
http://papers.gersteinlab.org/papers/pubnet)

PERMISSIONS: This Presentation is copyright Mark Gerstein, Yale University, 2008. Please read permissions statement at http://www.gersteinlab.org/misc/permissions.html . Feel free to use images in the talk with PROPER acknowledgement (via citation to relevant papers or link to gersteinlab.org).

.

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** .