Increasing Confidence of Protein-Protein Interactomes

Limsoon Wong

(Based on work of/with Jin Chen, Kenny Chua, Wynne Hsu, Mong Li Lee, See-Kiong Ng, Rintaro Saito, Wing-Kin Sung)

NUS
National University of Singapore

Outline

- **Reliability of experimental protein-protein interaction data**
- **Identification of false positives**
  - Interaction generality
  - Interaction generality 2
  - Interaction pathway reliability
  - FS Weight
  - Meso-scale network motifs
- **Identification of false negatives**

Copyright 2006 © Limsoon Wong

How Reliable are Experimental Protein-Protein Interaction Data?

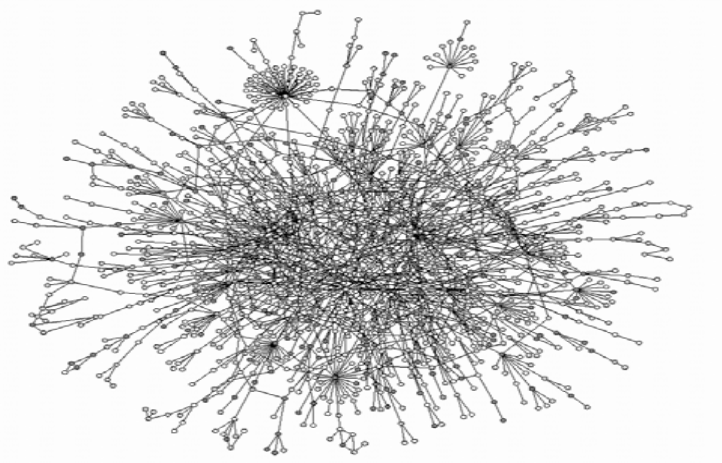


Figure credit: Jeong et al. 2001

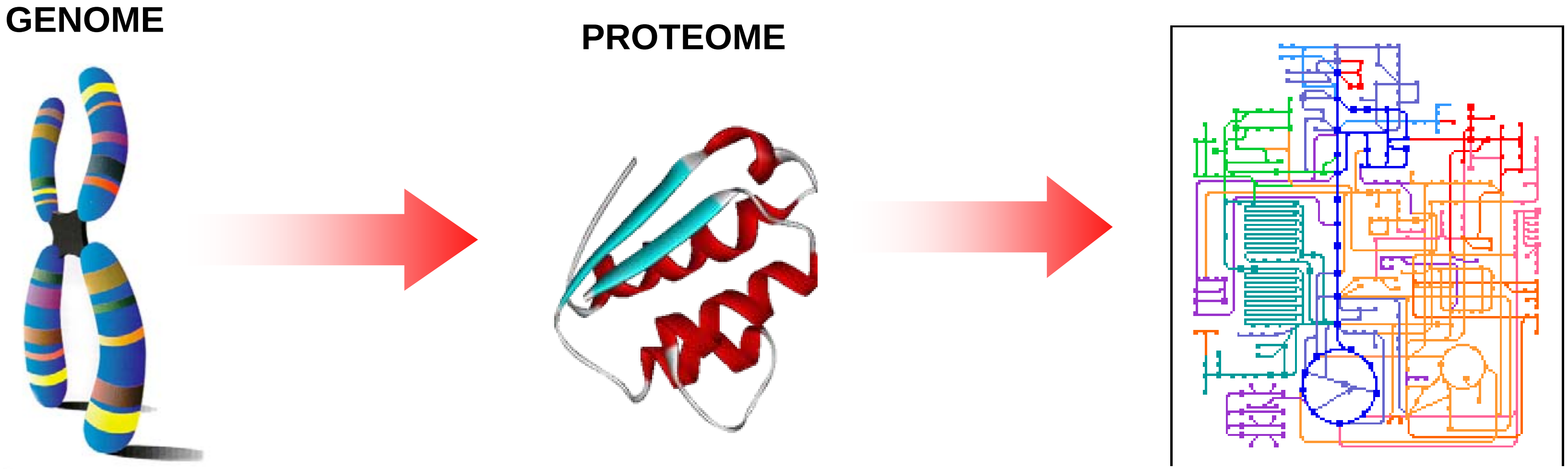
NUS National University of Singapore

Why Protein Interactions?

- Complete genomes are now available
- Knowing the **genes** is not enough to understand how biology **functions**

- **Proteins,** not genes, are responsible for many cellular activities

- Proteins function by **interacting** w/ other proteins and biomolecules



Slide credit: See-Kiong Ng

Copyright 2006 © Limsoon Wong

High-Tech Expt PPI Detection Methods

- **Yeast two-hybrid assays**
- **Mass spec of purified complexes (e.g., TAP)**
- **Correlated mRNA expression**
- **Genetic interactions (e.g., synthetic lethality)**
- **...**

FACT: Generating ***large amounts*** of experimental data about protein-protein interactions can be done with ease.

Slide credit: See-Kiong Ng

Copyright 2006 © Limsoon Wong

Key Bottleneck

- **Many high-throughput expt detection methods for protein-protein interactions have been devised**
- **But ...**

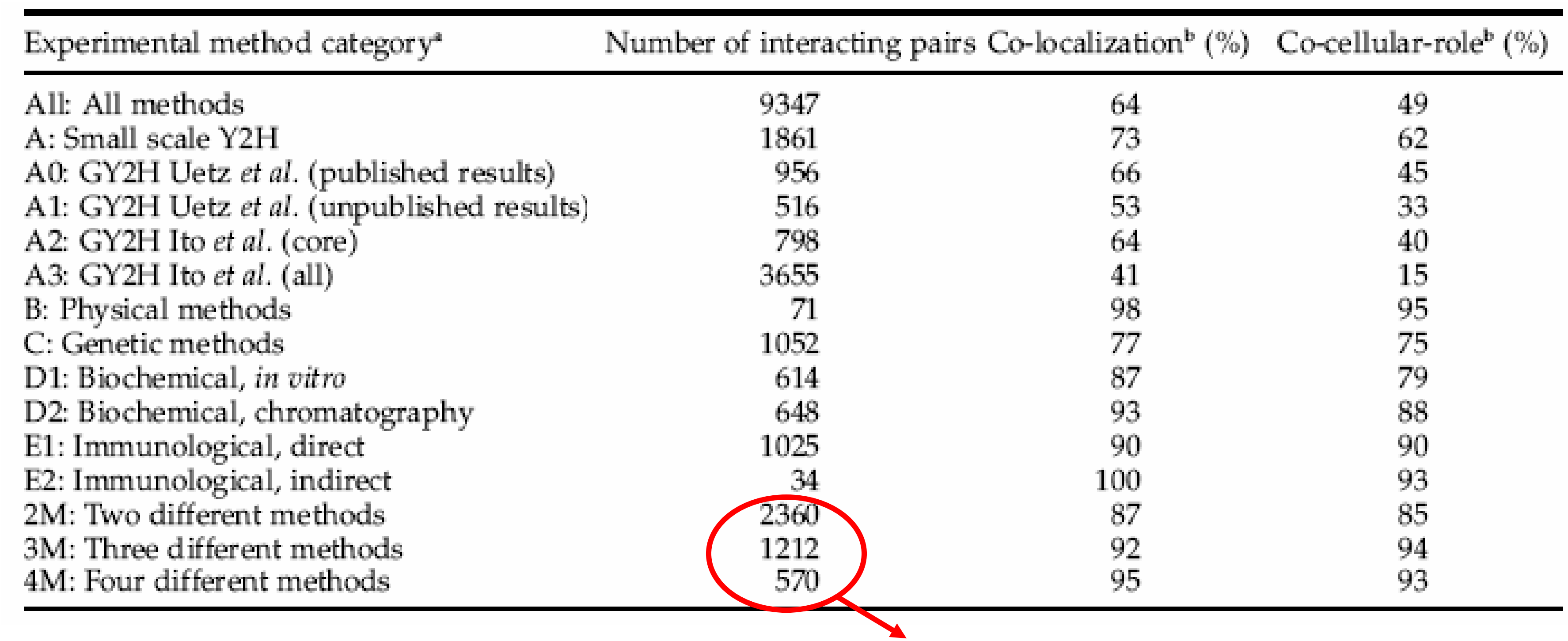
High-throughput approach sacrifice quality for ***quantity***:
(a) limited or biased coverage: ***false negatives***, &
(b) high error rates : ***false positives***

Slide credit: See-Kiong Ng

Copyright 2006 © Limsoon Wong

Some Protein Interaction Data Sets

Sprinzak et al., *JMB*, 327:919-923, 2003



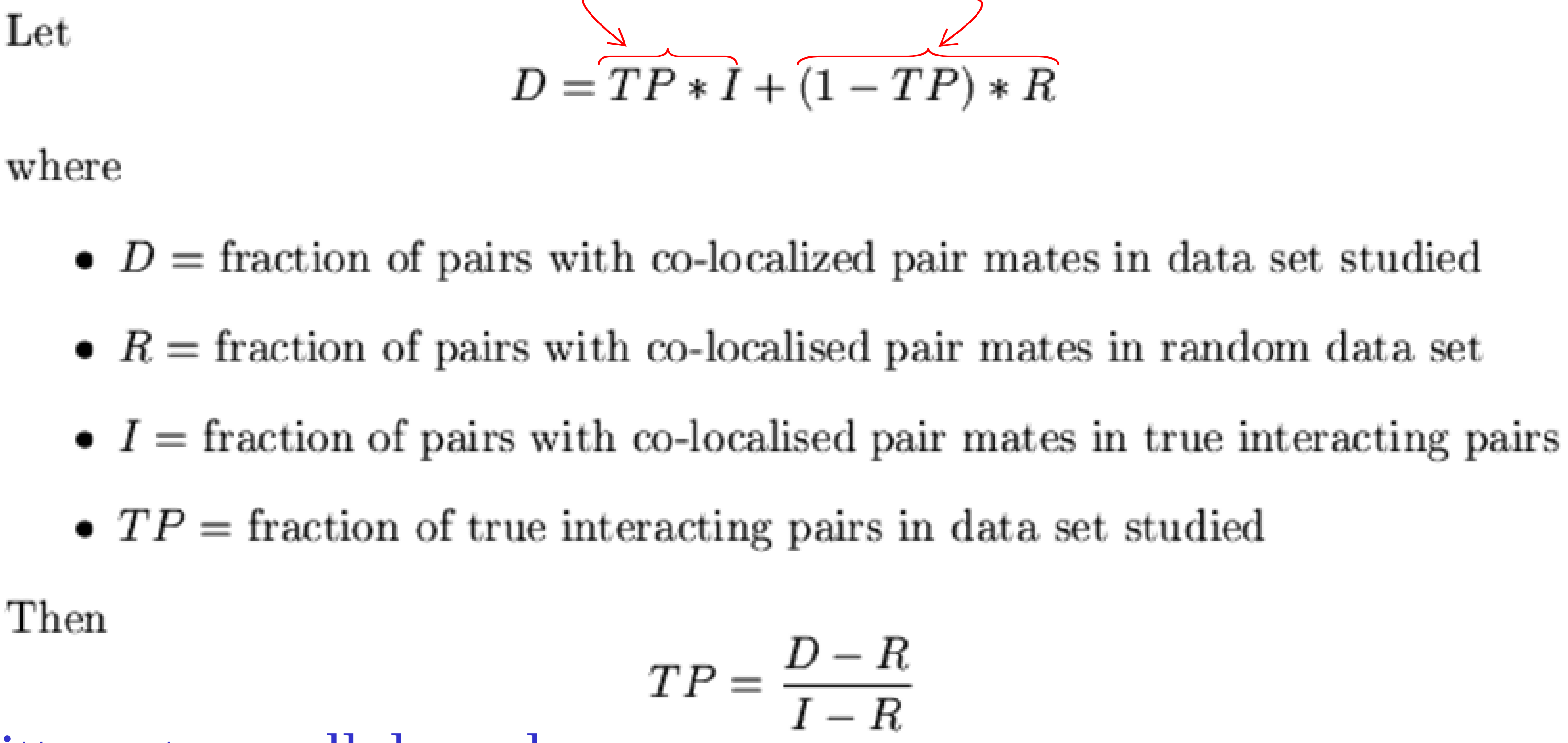
Experimental method category[a]	Number of interacting pairs	Co-localization[b] (%)	Co-cellular-role[b] (%)
All: All methods	9347	64	49
A: Small scale Y2H	1861	73	62
A0: GY2H Uetz *et al.* (published results)	956	66	45
A1: GY2H Uetz *et al.* (unpublished results)	516	53	33
A2: GY2H Ito *et al.* (core)	798	64	40
A3: GY2H Ito *et al.* (all)	3655	41	15
B: Physical methods	71	98	95
C: Genetic methods	1052	77	75
D1: Biochemical, *in vitro*	614	87	79
D2: Biochemical, chromatography	648	93	88
E1: Immunological, direct	1025	90	90
E2: Immunological, indirect	34	100	93
2M: Two different methods	2360	87	85
3M: Three different methods	1212	92	94
4M: Four different methods	570	95	93

Large disagreement betw methods

- **GY2H: genome-scale Y2H**
- **2M, 3M, 4M: intersection of 2, 3, 4 methods**

Quantitative Estimates

Sprinzak et al, *JMB*, 327:919-923, 2003



Let

$$D = \overbrace{TP * I}^{\text{Expected proportion of co-localized pairs among true interacting pairs}} + \overbrace{(1 - TP) * R}^{\text{Expected proportion of co-localized pairs among non true interacting pairs}}$$

where

- D = fraction of pairs with co-localized pair mates in data set studied
- R = fraction of pairs with co-localised pair mates in random data set
- I = fraction of pairs with co-localised pair mates in true interacting pairs
- TP = fraction of true interacting pairs in data set studied

Then

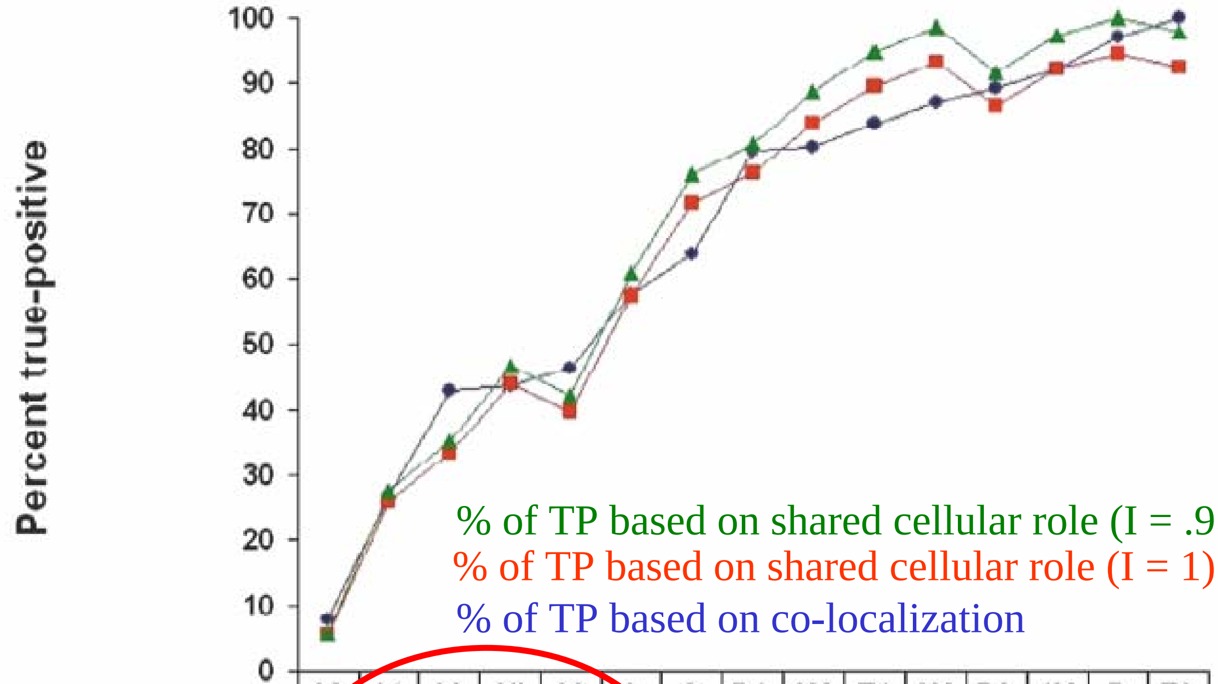
$$TP = \frac{D - R}{I - R}$$

Ditto wrt co-cellular-role

Copyright 2006 © Limsoon Wong

Reliability of Protein Interaction Data

Sprinzak et al, *JMB*, 327:919-923, 2003



% of TP based on shared cellular role (I = .95)

% of TP based on shared cellular role (I = 1)

% of TP based on co-localization

	A3	A1	A2	All	A0	A	C	D1	2M	E1	3M	D2	4M	B	E2
● % TP by Localization, I=1	7.8	26.7	43.0	43.7	46.2	57.7	63.9	79.5	80.2	84.0	87.1	89.0	92.1	97.2	100.
■ % TP by Cellular Role, I=1	5.5	25.9	33.3	43.9	39.7	57.4	71.8	76.3	83.8	89.4	93.2	86.5	92.1	94.5	92.4
▲ % TP by Cellular Role, I=0.95	5.9	27.4	35.3	46.5	42.0	60.8	76.0	80.8	88.7	94.7	98.6	91.6	97.5	100.	97.8

Experimental method category

TP = ~50%

Copyright 2006 © Limsoon Wong

Are We There Yet?

	Coverage	Data quality
DNA genome sequence	**99% of genome sequence**	**99.9% correct**
mRNA profiling	**80-90% of transcripts represented**	**90% of spots are good data**
Protein interaction data	10-30% **of interactions catalogued**	50-70% **of interactions are spurious**

Slide credit: See-Kiong Ng

Copyright 2006 © Limsoon Wong

Objective

- **Some high-throughput protein interaction expts have as much as 50% false positives**

- **Can we find a way to rank candidate interaction pairs according to their reliability?**

- **How do we do this?**
  - Would knowing their neighbours help?
  - Would knowing their local topology help?
  - Would knowing their global topology help?

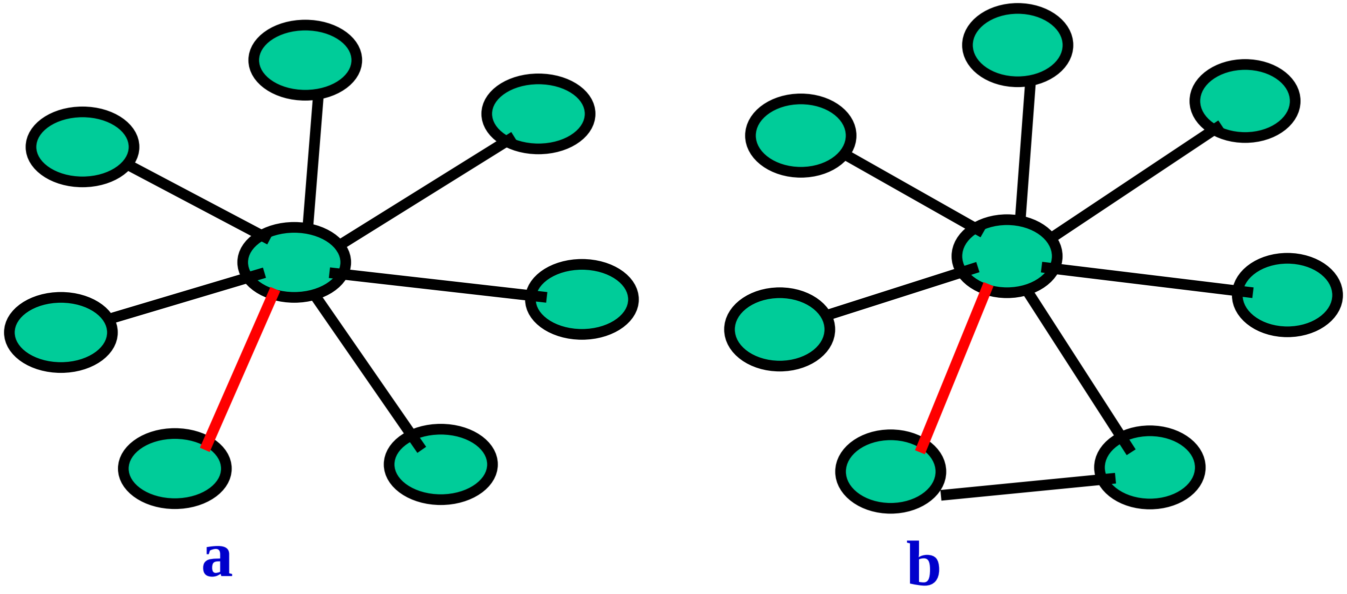
Copyright 2006 © Limsoon Wong

Would knowing their neighbours help?

The story of interaction generality

NUS
National University of Singapore

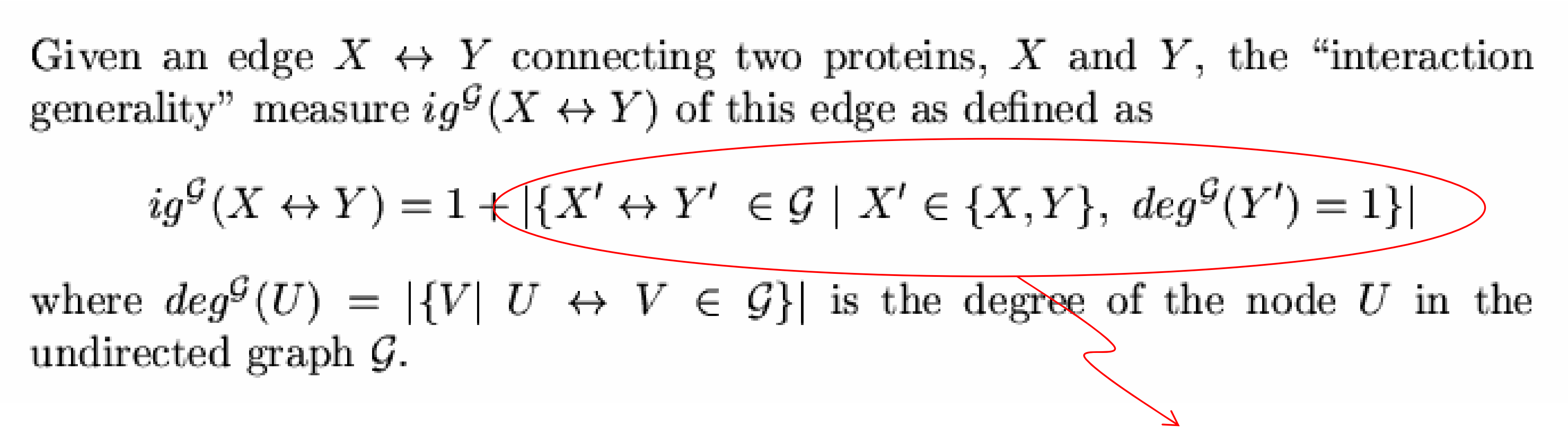
An Observation



- It seems that configuration a is less likely than b in protein interaction networks
- Can we exploit this?

Interaction Generality

Saito et al., *NAR*, 30:1163-1168, 2002

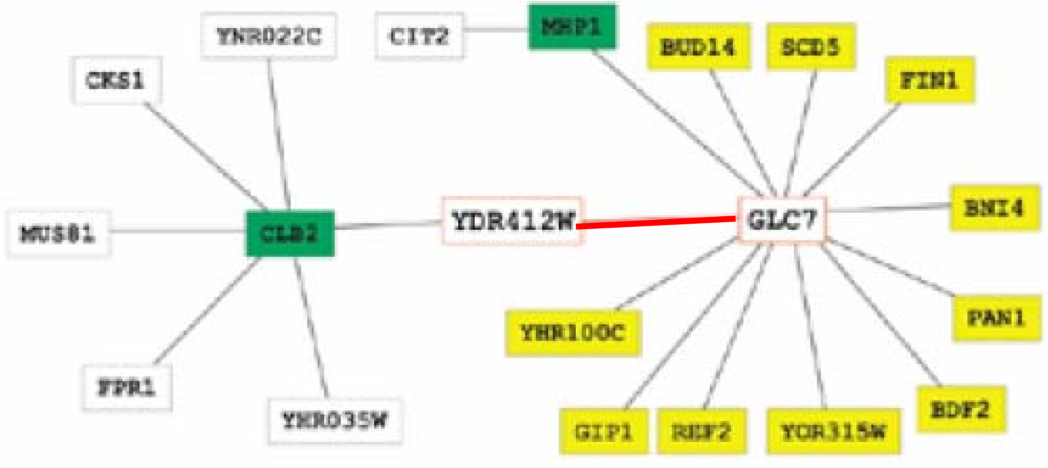


Given an edge $X \leftrightarrow Y$ connecting two proteins, X and Y, the "interaction generality" measure $ig^{\mathcal{G}}(X \leftrightarrow Y)$ of this edge as defined as

$$ig^{\mathcal{G}}(X \leftrightarrow Y) = 1 + |\{X' \leftrightarrow Y' \in \mathcal{G} \mid X' \in \{X,Y\},\ deg^{\mathcal{G}}(Y') = 1\}|$$

where $deg^{\mathcal{G}}(U) = |\{V |\ U \leftrightarrow V \in \mathcal{G}\}|$ is the degree of the node U in the undirected graph $\mathcal{G}$.

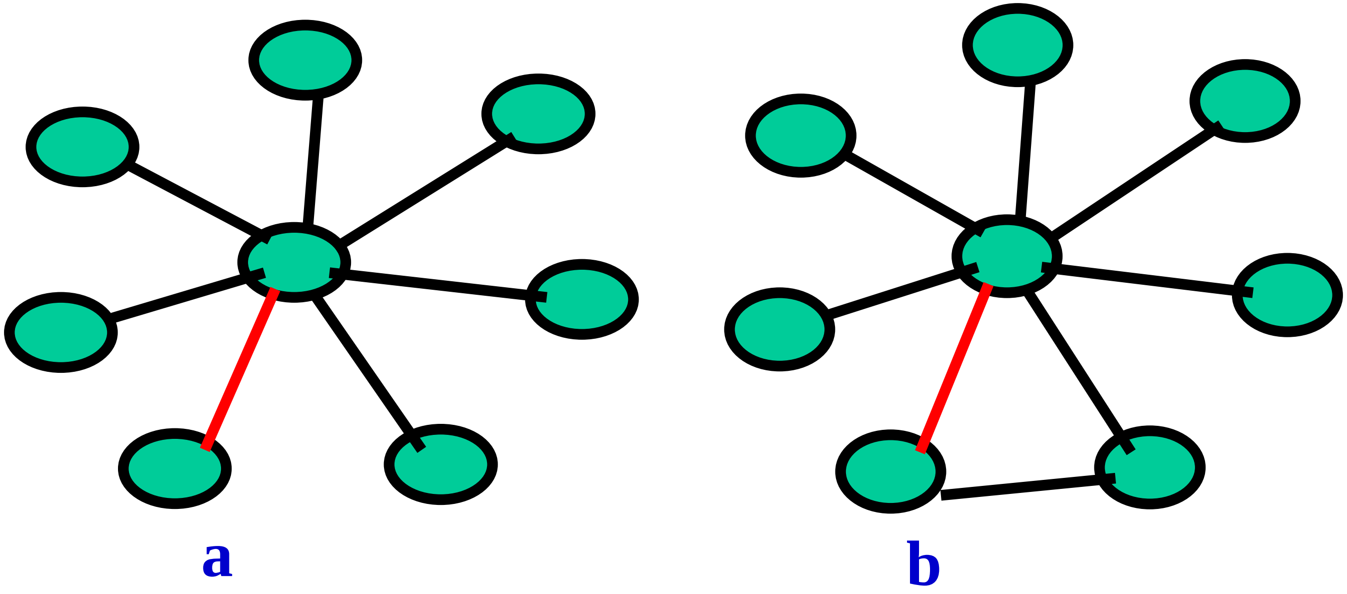
The number of proteins that "interact" with just X or Y, and nobody else



ig(YDR412W↔GLC7) = 1 + # of yellow nodes

Copyright 2006 © Limsoon Wong

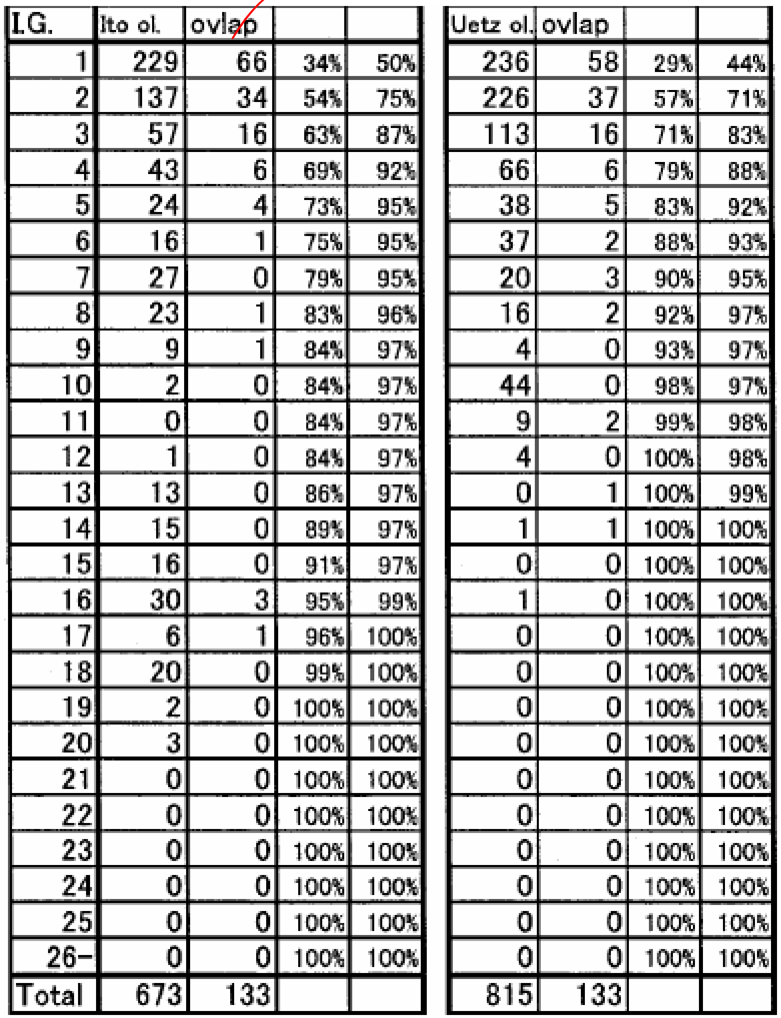
Assessing Reliability Using Interaction Generality



- Recall configuration a is less likely than b in protein interaction networks
- The smaller the "ig" value of a candidate interaction pair is, the more likely that interaction is

Copyright 2006 © Limsoon Wong

Evaluation wrt Intersection of Ito et al. & Uetz et al.



I.G.	Ito ol.	ovlap			Uetz ol.	ovlap		
1	229	66	34%	50%	236	58	29%	44%
2	137	34	54%	75%	226	37	57%	71%
3	57	16	63%	87%	113	16	71%	83%
4	43	6	69%	92%	66	6	79%	88%
5	24	4	73%	95%	38	5	83%	92%
6	16	1	75%	95%	37	2	88%	93%
7	27	0	79%	95%	20	3	90%	95%
8	23	1	83%	96%	16	2	92%	97%
9	9	1	84%	97%	4	0	93%	97%
10	2	0	84%	97%	44	0	98%	97%
11	0	0	84%	97%	9	2	99%	98%
12	1	0	84%	97%	4	0	100%	98%
13	13	0	86%	97%	0	1	100%	99%
14	15	0	89%	97%	1	1	100%	100%
15	16	0	91%	97%	0	0	100%	100%
16	30	3	95%	99%	1	0	100%	100%
17	6	1	96%	100%	0	0	100%	100%
18	20	0	99%	100%	0	0	100%	100%
19	2	0	100%	100%	0	0	100%	100%
20	3	0	100%	100%	0	0	100%	100%
21	0	0	100%	100%	0	0	100%	100%
22	0	0	100%	100%	0	0	100%	100%
23	0	0	100%	100%	0	0	100%	100%
24	0	0	100%	100%	0	0	100%	100%
25	0	0	100%	100%	0	0	100%	100%
26-	0	0	100%	100%	0	0	100%	100%
Total	673	133			815	133		

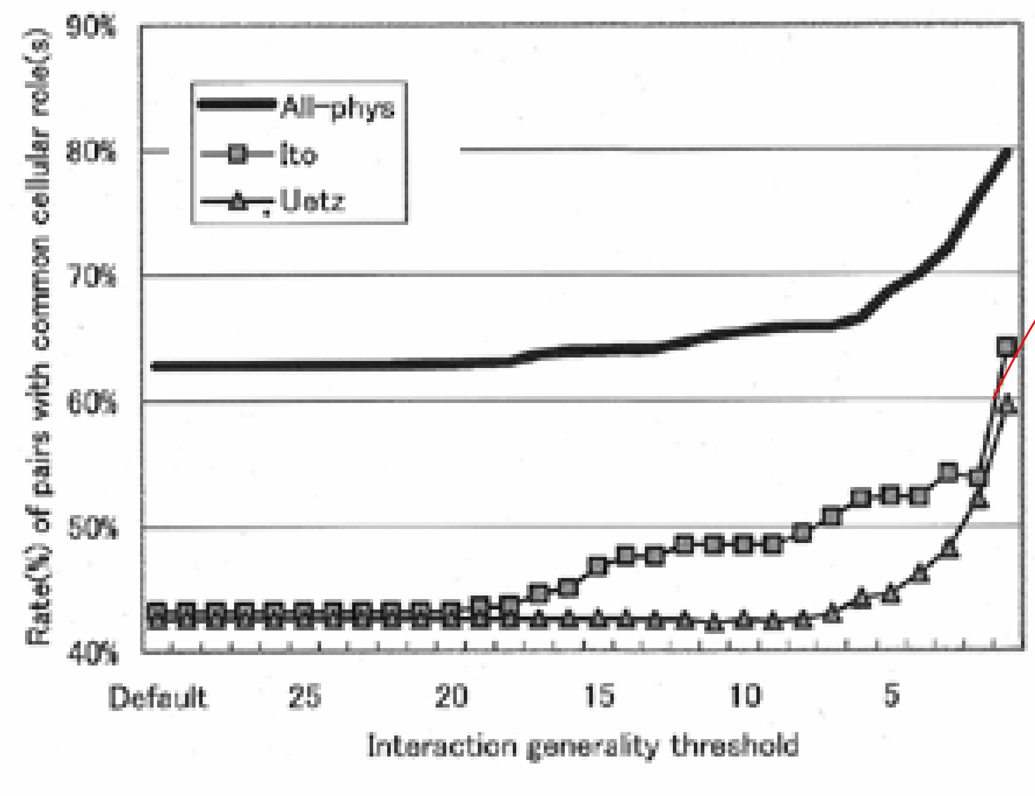
There are 229 pairs
in Ito having ig = 1.
Of these, 66 (or 34%)
are also reported by Uetz

- **Interacting pairs c'mon to Ito et al. & Uetz et al. are more reliable**
- **Also have smaller "ig"**

⇒ "ig" seems to work

Copyright 2006 © Limsoon Wong

Evaluation wrt Co-localization



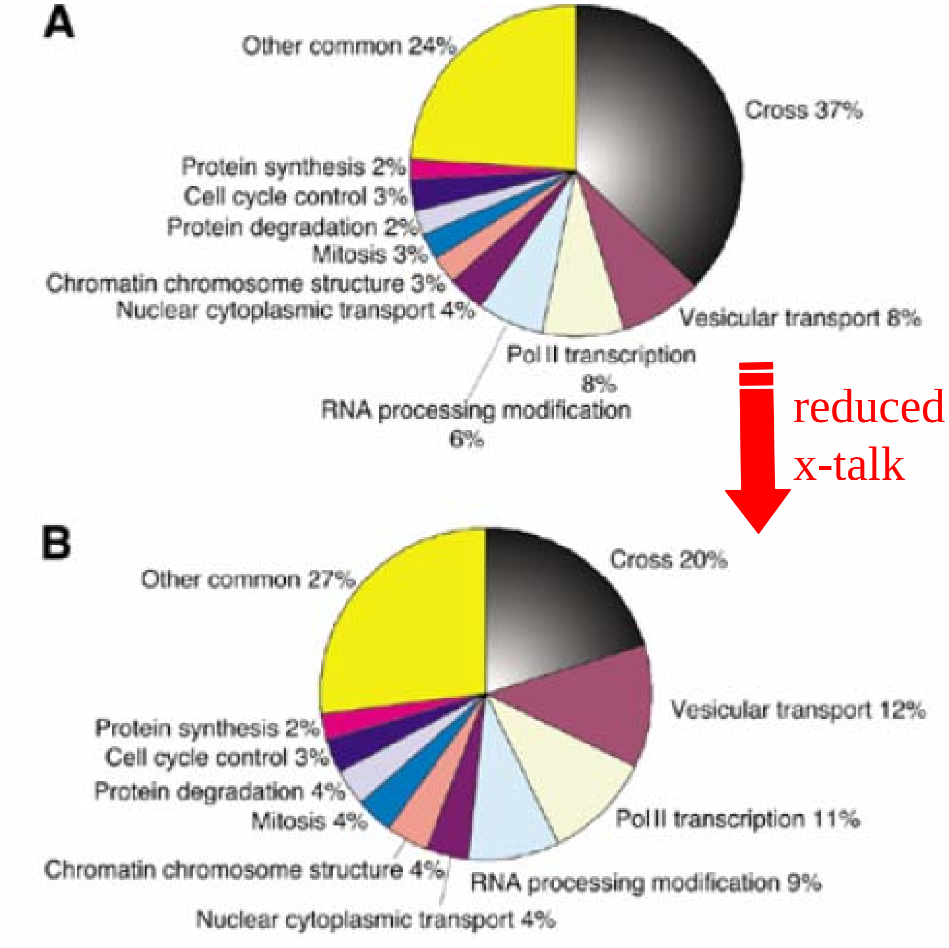
~60% of pairs in
in Ito having ig=1
are known to have
common localization

- Interaction pairs having common cellular localization are more likely
- Also have lower “ig”

⇒ “ig” seems to work

Copyright 2006 © Limsoon Wong

Evaluation wrt Co-cellular Role



reduced x-talk

- **Interaction pairs having common cellular role are more likely**
- **Also have lower "ig"**

⇒ **"ig" seems to work**

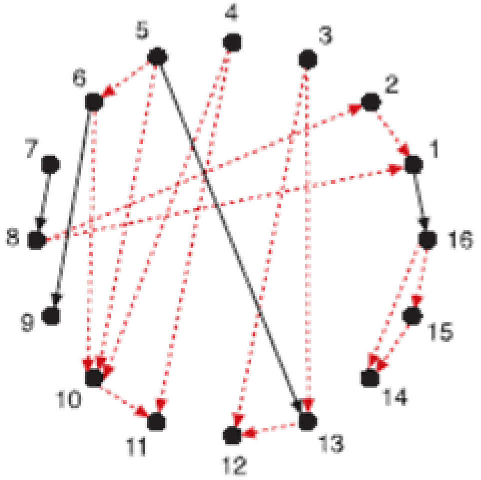
A: before restrict to pairs with "ig = 1"
B: after restrict to pairs with "ig = 1"

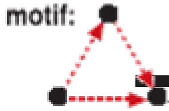
Would knowing their local topology help?

The story of interaction generality 2

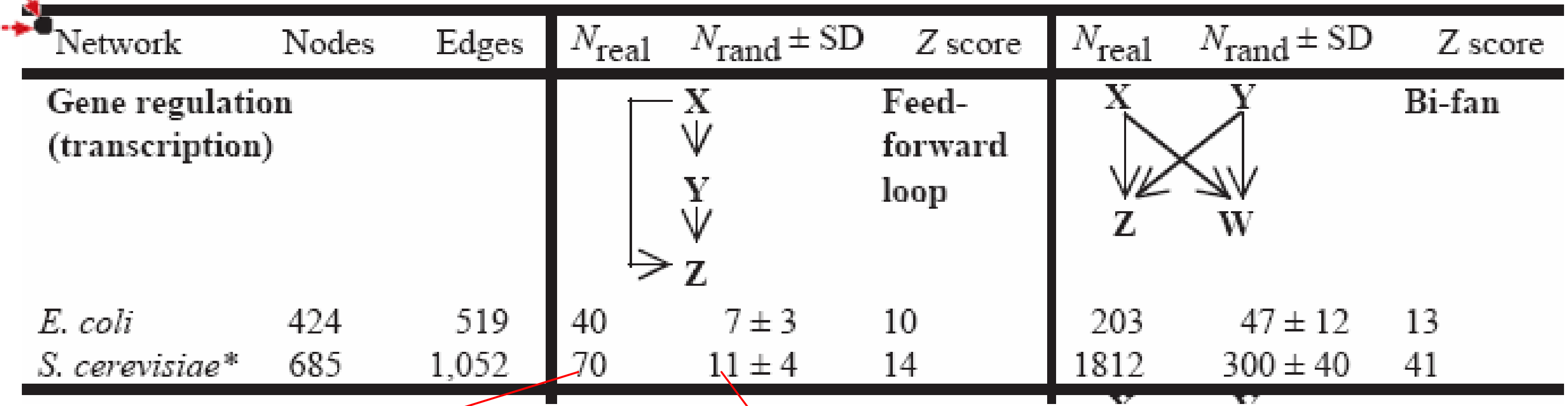
Existence of Network Motifs

Milo et al., Science, 298:824-827, 2002





- **A network motif is just a local topological configuration of the network**
- **“Detected” in gene regulation networks, WWW links, etc.**

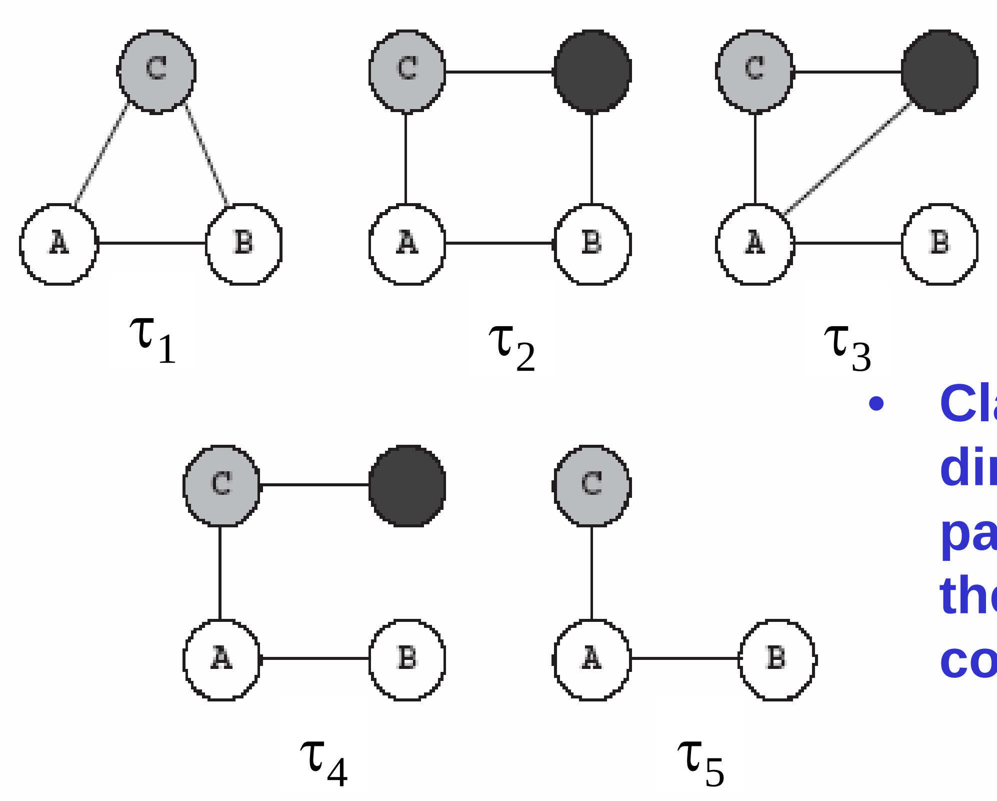


Network	Nodes	Edges	N_{real}	$N_{rand} \pm SD$	Z score	N_{real}	$N_{rand} \pm SD$	Z score
Gene regulation (transcription)				**X** **Y** **Z**	**Feed-forward loop**	**X Y** **Z W**		**Bi-fan**
E. coli	424	519	40	7 ± 3	10	203	47 ± 12	13
*S. cerevisiae**	685	1,052	70	11 ± 4	14	1812	300 ± 40	41

Observed 70 times in S. cerevisiae

Observed ~11 times in random data

5 Possible Network Motifs

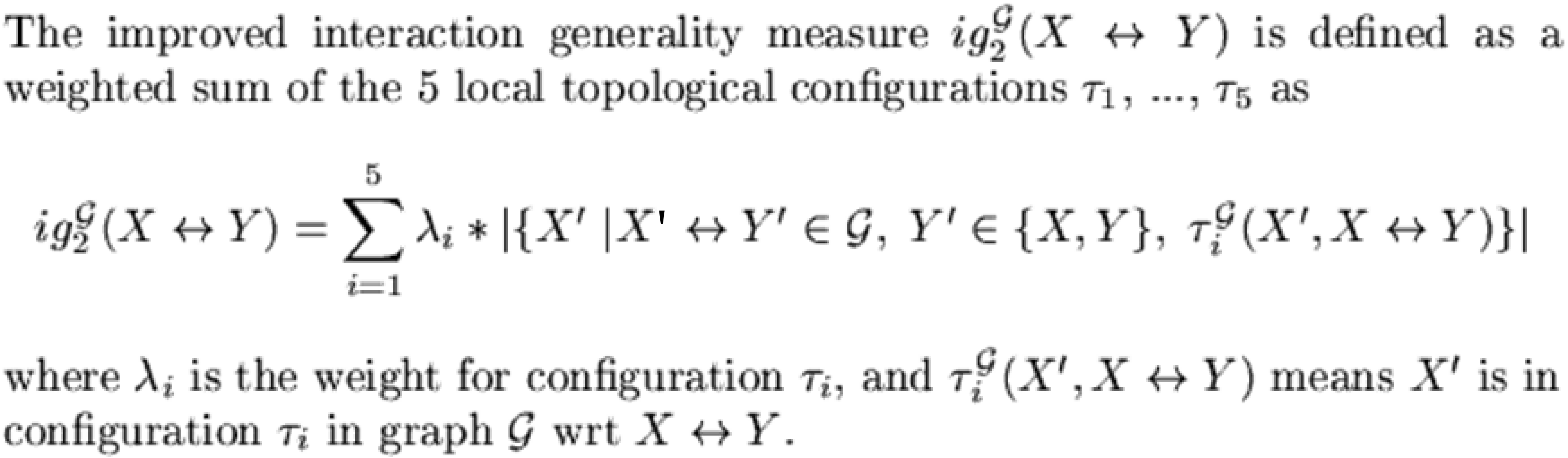


- **Classify a protein C that directly interacts with the pair A↔B according to these 5 topological configurations**

Copyright 2006 © Limsoon Wong

A New Interaction Generality

Saito et al., *Bioinformatics*, 19:756--763, 2003



The improved interaction generality measure $ig_2^{\mathcal{G}}(X \leftrightarrow Y)$ is defined as a weighted sum of the 5 local topological configurations τ_1, ..., τ_5 as

$$ig_2^{\mathcal{G}}(X \leftrightarrow Y) = \sum_{i=1}^{5} \lambda_i * |\{X' \,|X' \leftrightarrow Y' \in \mathcal{G},\ Y' \in \{X,Y\},\ \tau_i^{\mathcal{G}}(X',X \leftrightarrow Y)\}|$$

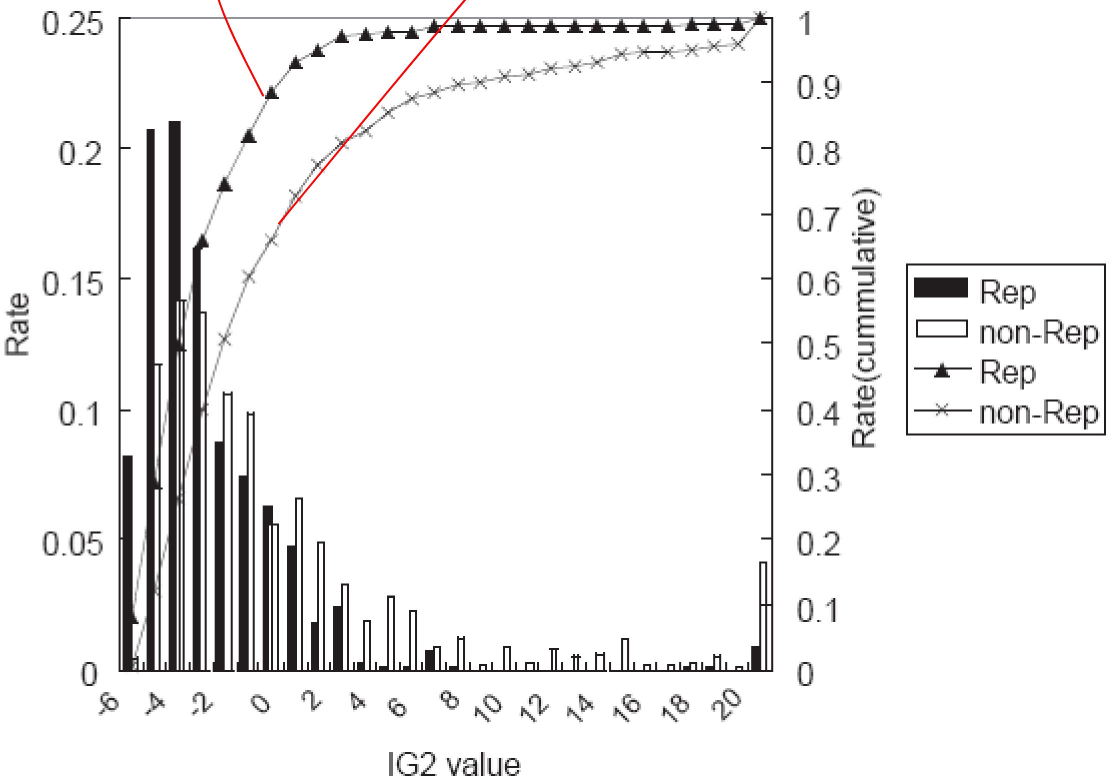
where λ_i is the weight for configuration τ_i, and $\tau_i^{\mathcal{G}}(X',X \leftrightarrow Y)$ means X' is in configuration τ_i in graph $\mathcal{G}$ wrt $X \leftrightarrow Y$.

Copyright 2006 © Limsoon Wong

Evaluation wrt Reproducible Interactions

~90% of pairs in intersection of Ito & Uetz have $ig_2 < 0$.

~60% of pairs not in intersection of Ito & Uetz have $ig_2 < 0$

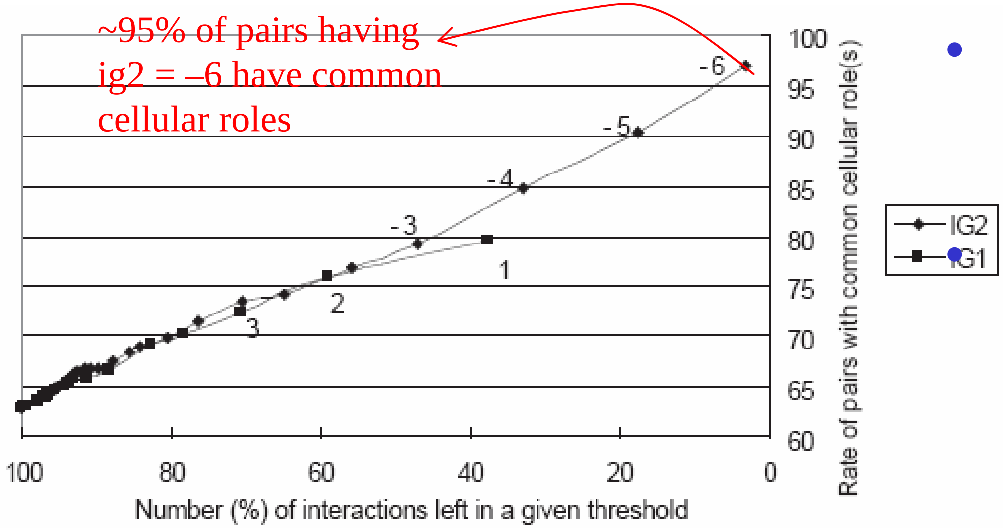


- **"ig_2" correlates to "reproducible" interactions**

⇒ "ig_2" seems to work

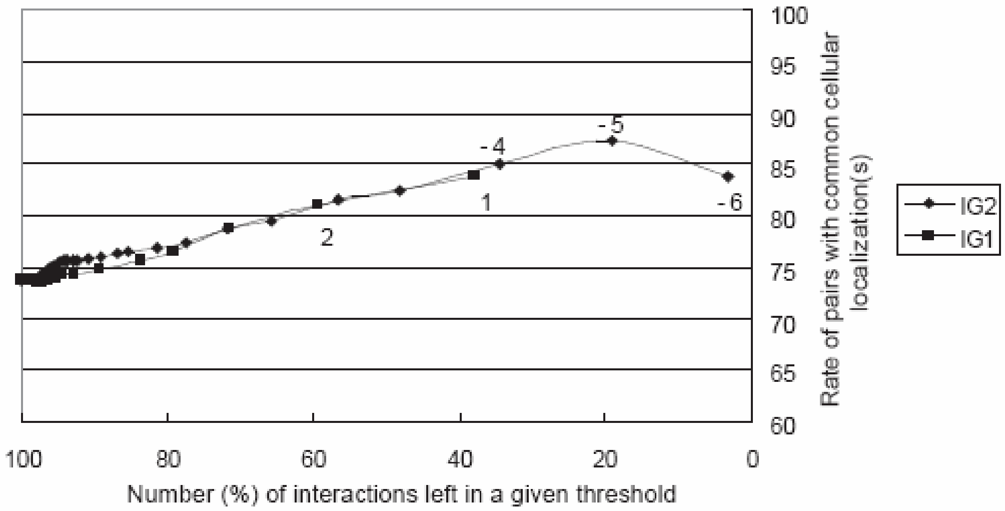
Copyright 2006 © Limsoon Wong

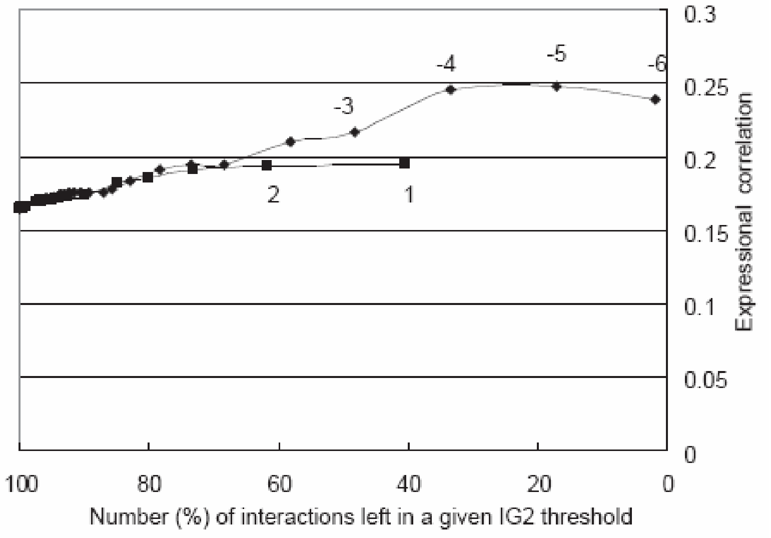
Evaluation wrt Common Cellular Role, etc.


~95% of pairs having ig2 = –6 have common cellular roles

- **“ig$_2$” correlates well to common cellular roles, localization, & expression**

“ig$_2$” seems to work better than “ig”





Copyright 2006 © Limsoon Wong

Would knowing their global topology help?

The story of interaction pathway reliability

Some “Reasonable” Speculations

- A true interacting pair is often connected by at least one alternative path (reason: a biological function is performed by a highly interconnected network of interactions)

- The shorter the alternative path, the more likely the interaction (reason: evolution of life is through “add-on” interactions of other or newer folds onto existing ones)

Copyright 2006 © Limsoon Wong

Therefore...

Conjecture:

"An interaction that is associated with an alternate path of reliable interactions is likely to be reliable."

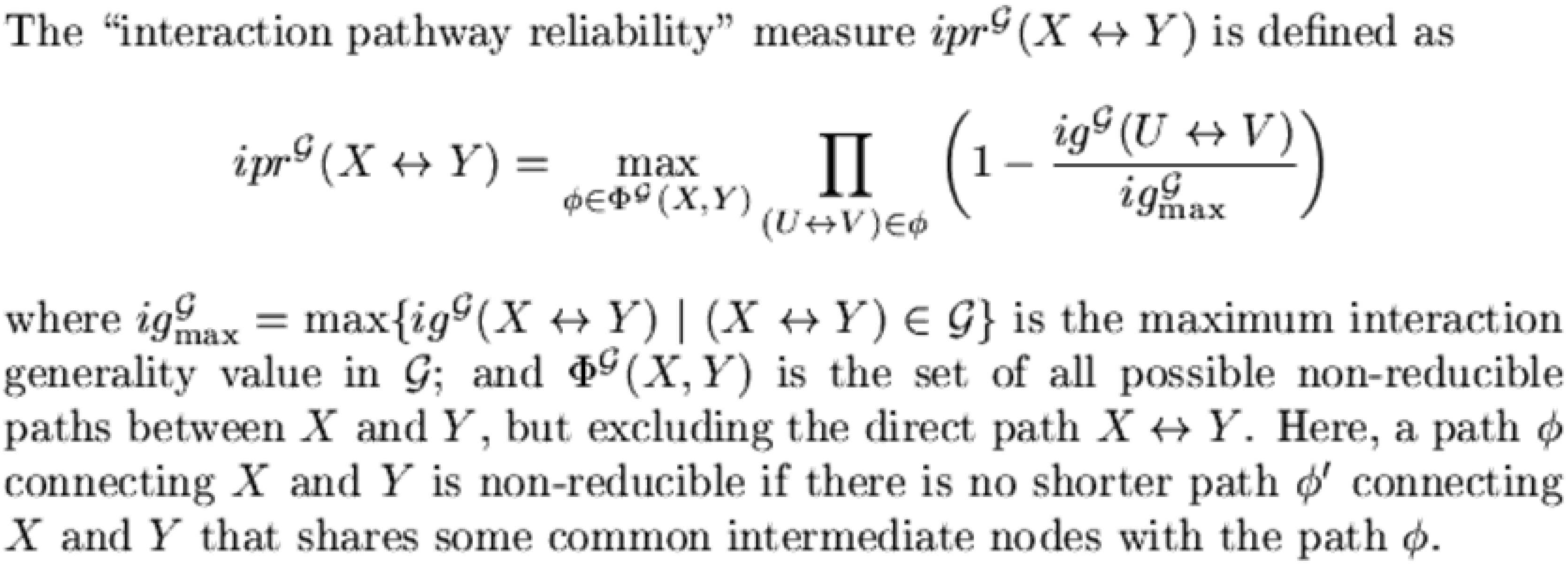
Idea:

Use **alternative interaction paths** as a measure to indicate functional linkage between the two proteins

Slide credit: See-Kiong Ng

Copyright 2006 © Limsoon Wong

Interaction Pathway Reliability



The "interaction pathway reliability" measure $ipr^{\mathcal{G}}(X \leftrightarrow Y)$ is defined as

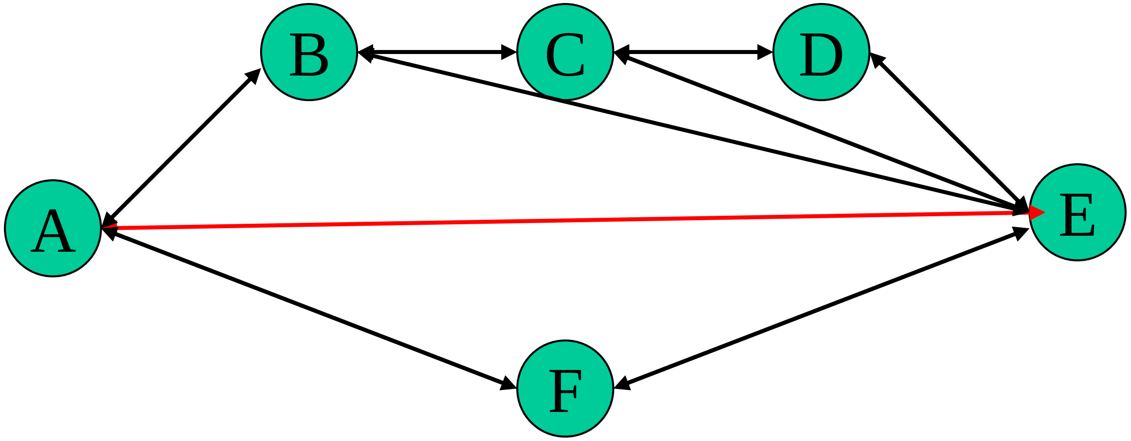
$$ipr^{\mathcal{G}}(X \leftrightarrow Y) = \max_{\phi \in \Phi^{\mathcal{G}}(X,Y)} \prod_{(U \leftrightarrow V) \in \phi} \left(1 - \frac{ig^{\mathcal{G}}(U \leftrightarrow V)}{ig^{\mathcal{G}}_{\max}}\right)$$

where $ig^{\mathcal{G}}_{\max} = \max\{ig^{\mathcal{G}}(X \leftrightarrow Y) \mid (X \leftrightarrow Y) \in \mathcal{G}\}$ is the maximum interaction generality value in $\mathcal{G}$; and $\Phi^{\mathcal{G}}(X, Y)$ is the set of all possible non-reducible paths between X and Y, but excluding the direct path $X \leftrightarrow Y$. Here, a path ϕ connecting X and Y is non-reducible if there is no shorter path ϕ' connecting X and Y that shares some common intermediate nodes with the path ϕ.

IPR is also called IRAP, "Interaction Reliability by Alternate Pathways"

Non-reducible Paths

- **Non-reducible paths are**
  - A←→F←→E
  - A←→B←→E
- **Reducible paths are**
  - A←→B←→C←→D←→E
  - A←→B←→C←→E

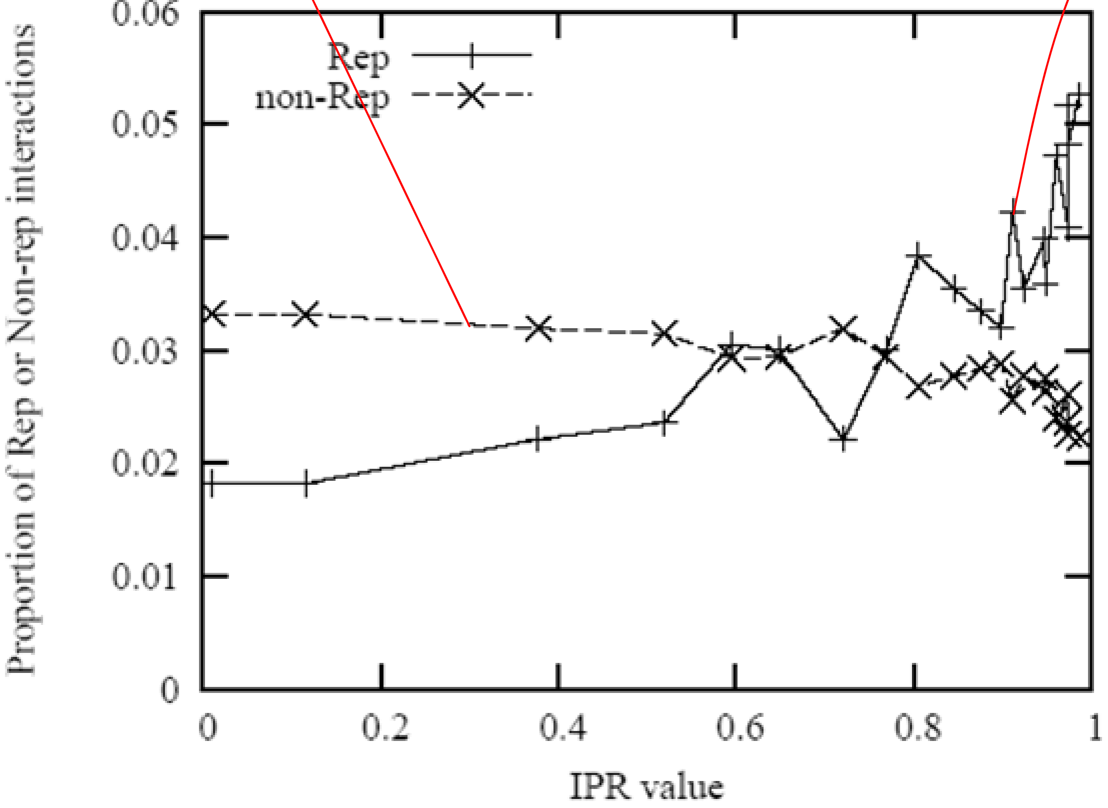


Copyright 2006 © Limsoon Wong

Evaluation wrt Reproducible Interactions

The number of pairs not in the intersection of Ito & Uetz is not changed much wrt the ipr value of the pairs

The number of pairs in the intersection of Ito & Uetz increases wrt the ipr value of the pairs

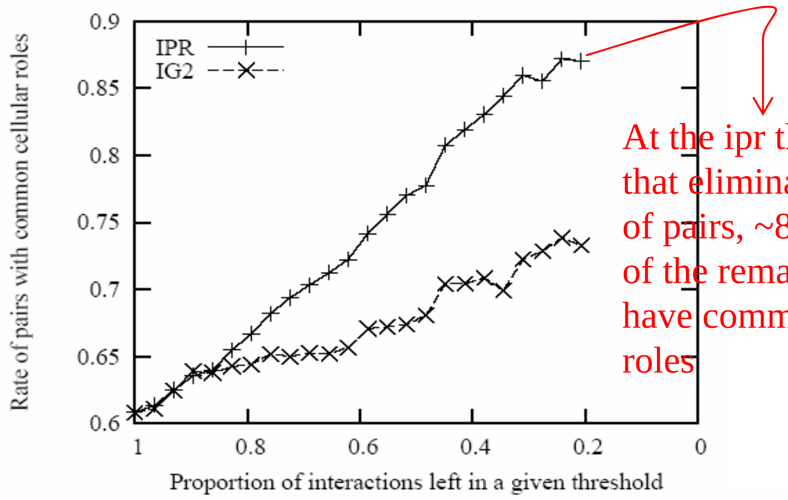


- “ipr” correlates well to “reproducible” interactions

⇒ “ipr” seems to work

Copyright 2006 © Limsoon Wong

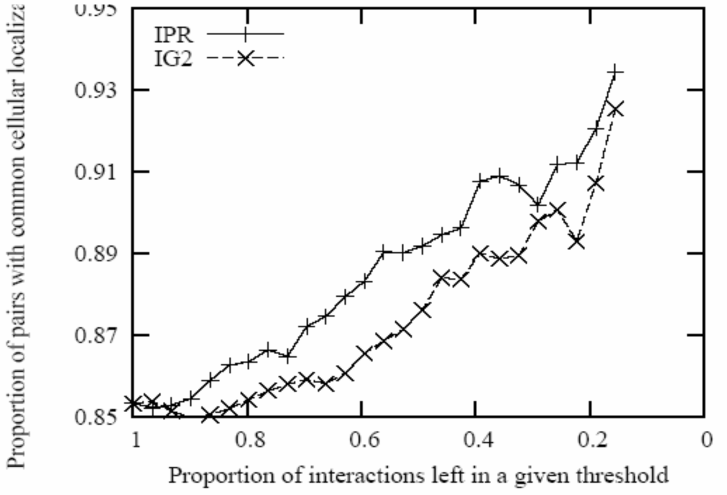
Evaluation wrt Common Cellular Role, etc

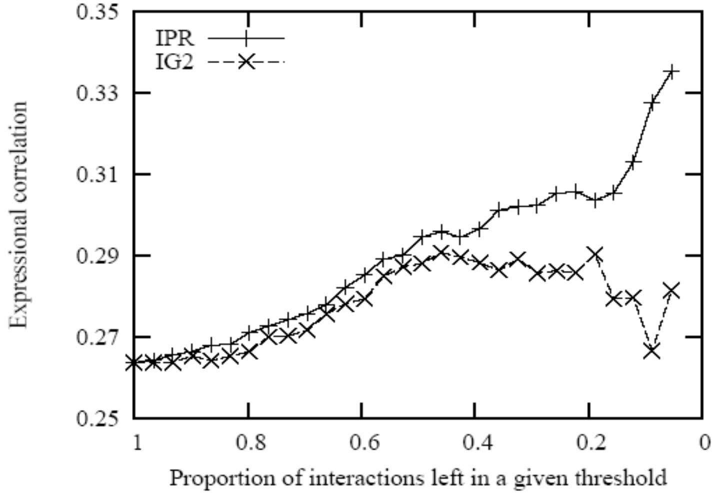


At the ipr threshold that eliminated 80% of pairs, ~85% of the of the remaining pairs have common cellular roles

- “ipr” correlates well to common cellular roles, localization, & expression

⇒ “ipr” seems to work better than “ig$_2$”



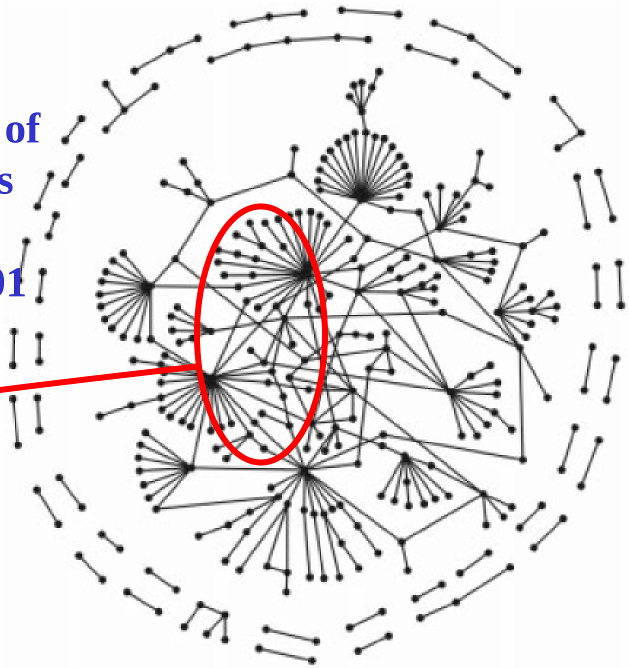


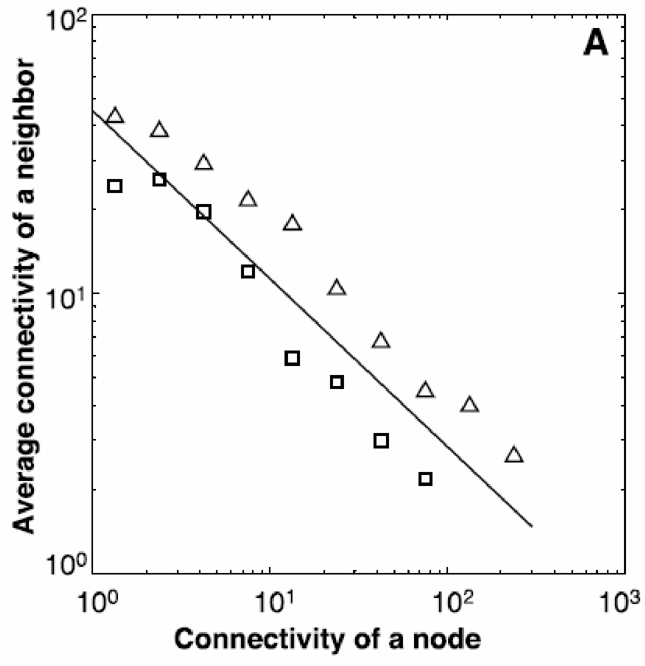
Copyright 2006 © Limsoon Wong

Stability in Protein Networks

Maslov & Sneppen, *Science*, 296:910-913, 2002

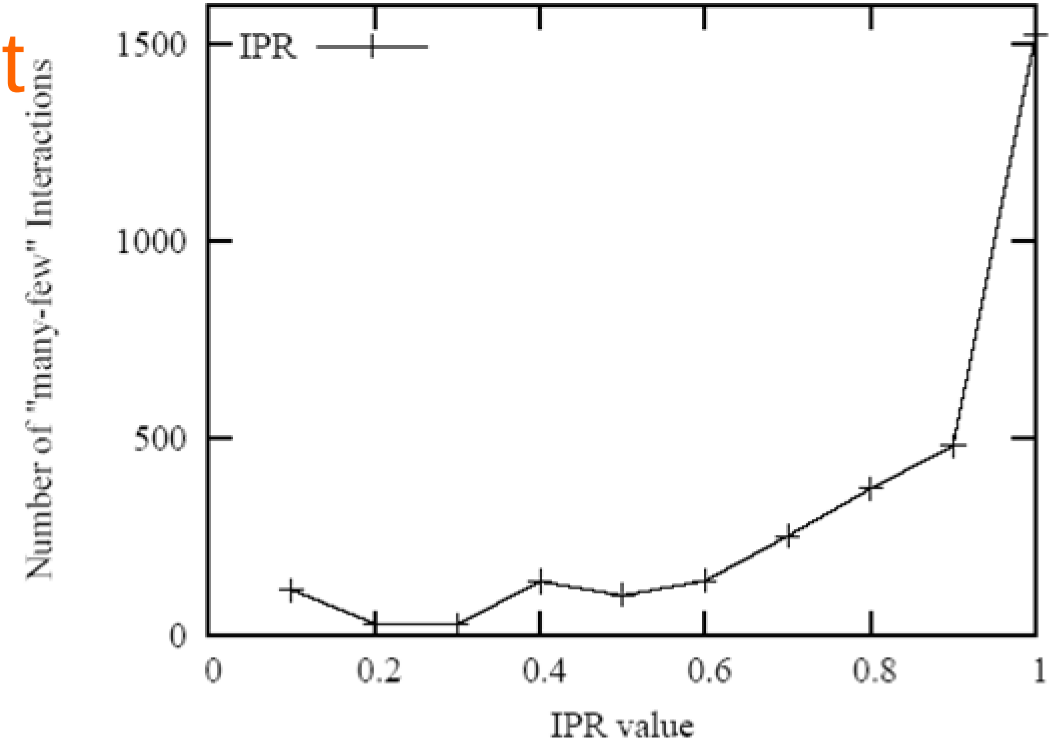
Part of the network of physical interactions reported by Ito et al., PNAS, 2001





- **According to Maslov & Sneppen**
  - Links betw high-connected proteins are suppressed
  - Links betw high- & low-connected proteins are favoured
- **This decreases cross talks & increases robustness**

Evaluation wrt "Many-few" Interactions



- Number of "Many-few" interactions increases when more "reliable" IPR threshold is used to filter interactions
- Consistent with the Maslov-Sneppen prediction

Copyright 2006 © Limsoon Wong

Evaluation wrt “Cross-Talkers”

- **A MIPS functional cat:**
  - | 02 | ENERGY
  - | 02.01 | glycolysis and gluconeogenesis
  - | 02.01.01 | glycolysis methylglyoxal bypass
  - | 02.01.03 | regulation of glycolysis & gluconeogenesis
- **First 2 digits is top cat**
- **Other digits add more granularity to the cat**

⇒ **Compare high- & low- IPR pairs that are not co-localised to determine number of pairs that fall into same cat. If more high-IPR pairs are in same cat, then IPR works**

Copyright (c) 2004 by Limsoon Wong

Copyright 2006 © Limsoon Wong

Evaluation wrt “Cross-Talkers”

- **For top cat**
  - 148/257 high-IPR pairs are in same cat
  - 65/260 low-IPR pairs are in same cat
- **For fine-granularity cat**
  - 135/257 high-IPR pairs are in same cat.
    37/260 low-IPR pairs are in same cat

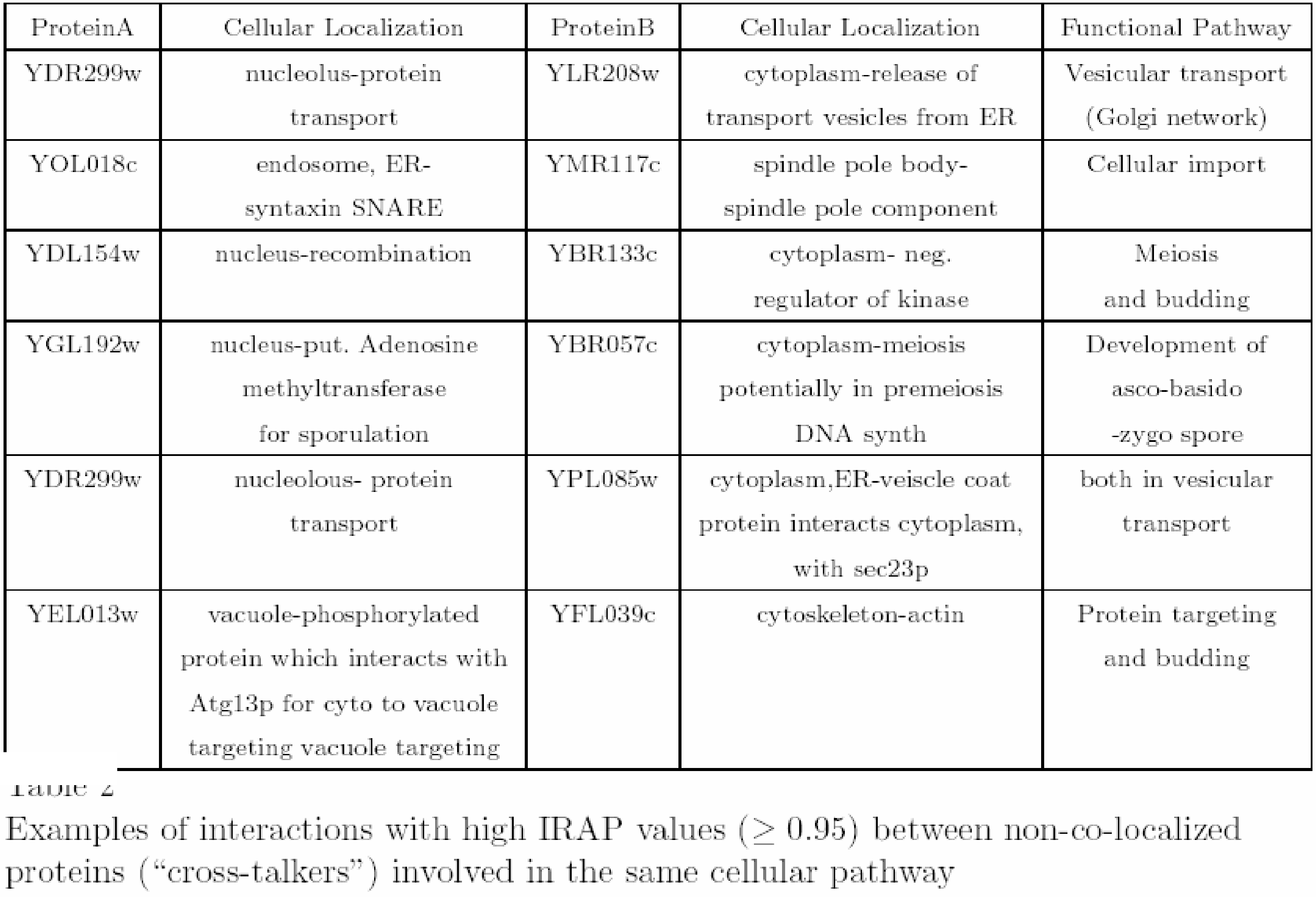
⇒ **IPR works**

⇒ **IPR pairs that are not co-localized are real cross-talkers!**

Copyright (c) 2004 by Limsoon Wong

Copyright 2006 © Limsoon Wong

Example Cross Talkers



ProteinA	Cellular Localization	ProteinB	Cellular Localization	Functional Pathway
YDR299w	nucleolus-protein transport	YLR208w	cytoplasm-release of transport vesicles from ER	Vesicular transport (Golgi network)
YOL018c	endosome, ER-syntaxin SNARE	YMR117c	spindle pole body-spindle pole component	Cellular import
YDL154w	nucleus-recombination	YBR133c	cytoplasm- neg. regulator of kinase	Meiosis and budding
YGL192w	nucleus-put. Adenosine methyltransferase for sporulation	YBR057c	cytoplasm-meiosis potentially in premeiosis DNA synth	Development of asco-basido -zygo spore
YDR299w	nucleolous- protein transport	YPL085w	cytoplasm,ER-veiscle coat protein interacts cytoplasm, with sec23p	both in vesicular transport
YEL013w	vacuole-phosphorylated protein which interacts with Atg13p for cyto to vacuole targeting vacuole targeting	YFL039c	cytoskeleton-actin	Protein targeting and budding

Table 2

Examples of interactions with high IRAP values (≥ 0.95) between non-co-localized proteins ("cross-talkers") involved in the same cellular pathway

Can local topology do better?

The story of FS Weight

Guilt by Association of Common Interaction Partners

- **Two proteins that have a large proportion of their interaction partners in common are likely to directly interact also**

- **In fact, this is a special case of the “alternative paths” used in the IPR index, because length-1 alternative paths = shared interaction partners**

Czekanowski-Dice Distance

- **Functional distance between two proteins** (Brun et al, 2003)

$$D(u,v) = \frac{|N_u \Delta N_v|}{|N_u \cup N_v| + |N_u \cap N_v|}$$

  - **N_k is the set of interacting partners of k**
  - **X Δ Y is symmetric diff betw two sets X and Y**
  - **Greater weight given to similarity**

Is this a good measure if u and v have very diff number of neighbours?

⇒ **Similarity can be defined as**

$$S(u,v) = 1 - D(u,v) = \frac{2X}{2X + (Y + Z)}$$

Functional Similarity Estimate: FS-Weighted Measure

- **FS-weighted measure**

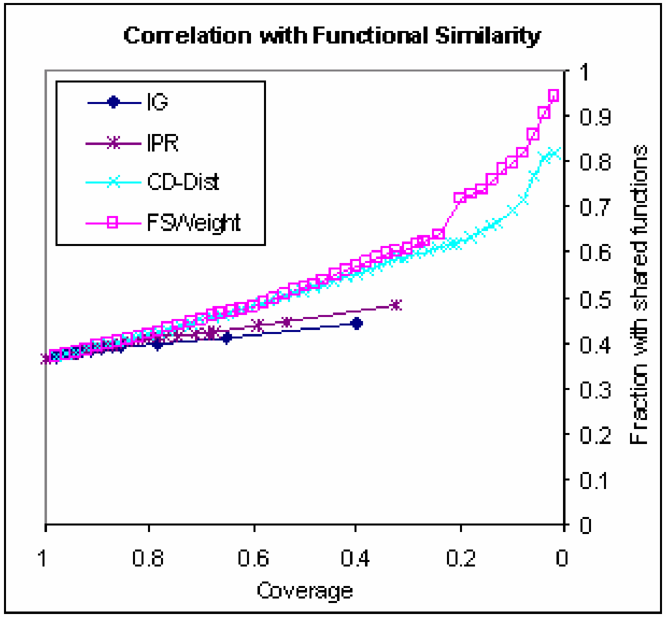
$$S(u,v) = \frac{2|N_u \cap N_v|}{|N_u - N_v| + 2|N_u \cap N_v|} \times \frac{2|N_u \cap N_v|}{|N_v - N_u| + 2|N_u \cap N_v|}$$

  - **N_k is the set of interacting partners of k**
  - **Greater weight given to similarity**

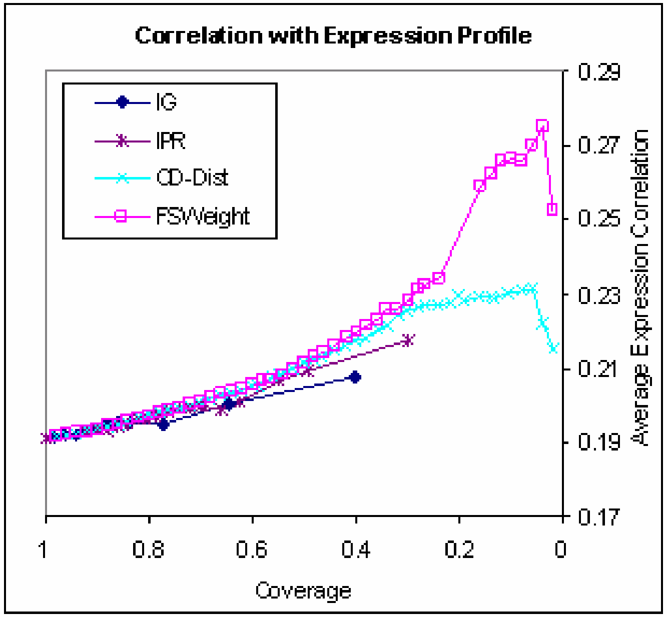
$\Rightarrow$ **Rewriting this as**

$$S(u,v) = \frac{2X}{2X+Y} \times \frac{2X}{2X+Z}$$

Copyright 2006 © Limsoon Wong

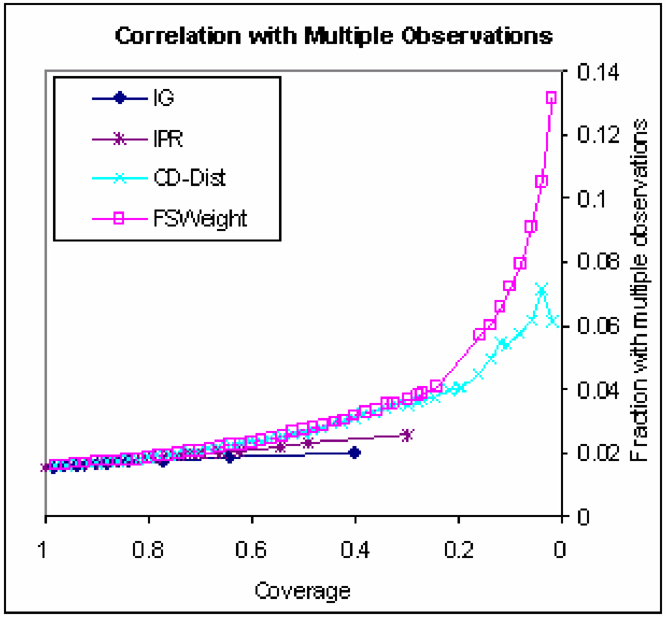


Correlation with Functional Similarity

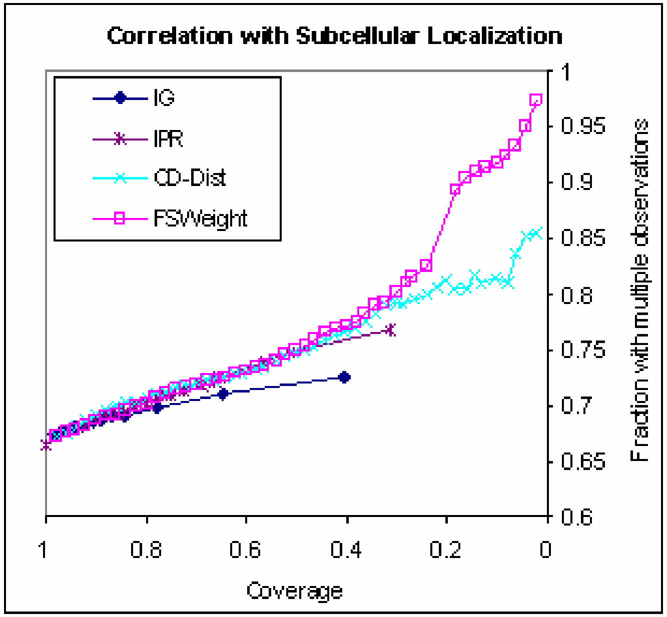


Correlation with Expression Profile

Evaluation wrt Common Cellular Role,etc



Correlation with Multiple Observations



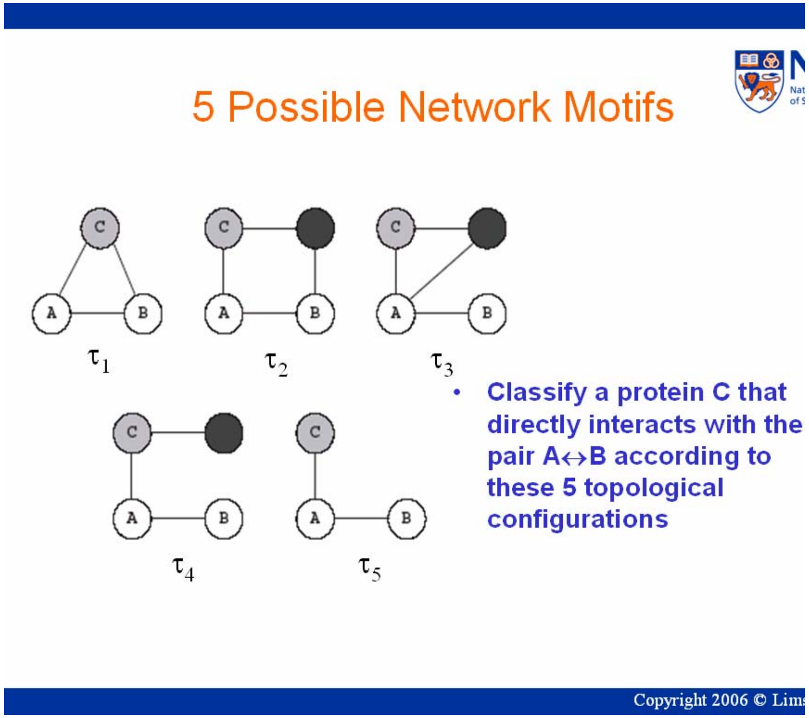
Correlation with Subcellular Localization

Copyright 2006 © Limsoon Wong

Another way to improve using local topology information

The story of meso-scale network motifs

Motivation for “Meso Scale”



5 Possible Network Motifs

- Classify a protein C that directly interacts with the pair A↔B according to these 5 topological configurations

Copyright 2006 © Lims

- **These motifs are very local and very small**
- **Many processes in biological network are ``meso-scale’’ (5-25 proteins)**

⇒ **May be we should also use meso-scale motifs?**

What is a network motif?

- **A network motif g in a PPI network G is a connected unlabelled undirected topological pattern of inter-connections that is repeated and unique in G**
- **Repeated: f_g, the number of occurrences of g in G, is more than threshold F**
- **Unique: s_g, the number of times f_g exceeds $f_{g,rand,i}$ over total number of randomized networks considered, is more than threshold S**

Example

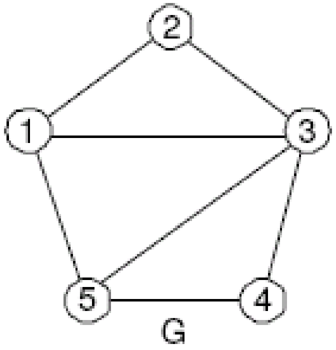


Figure 1: Example graph G.

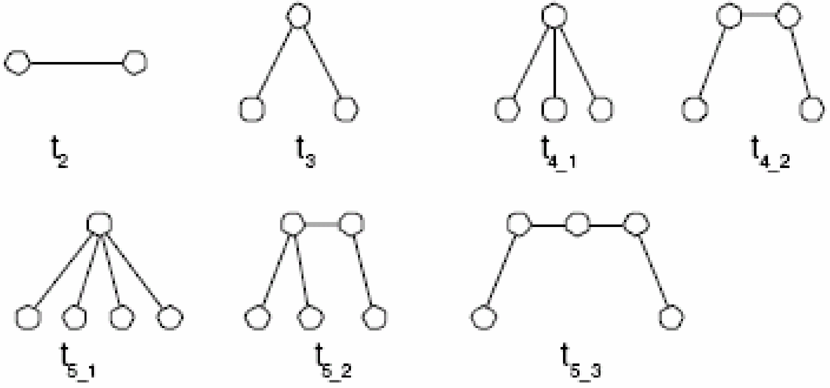


Figure 2: Size 2 to size 5 trees.

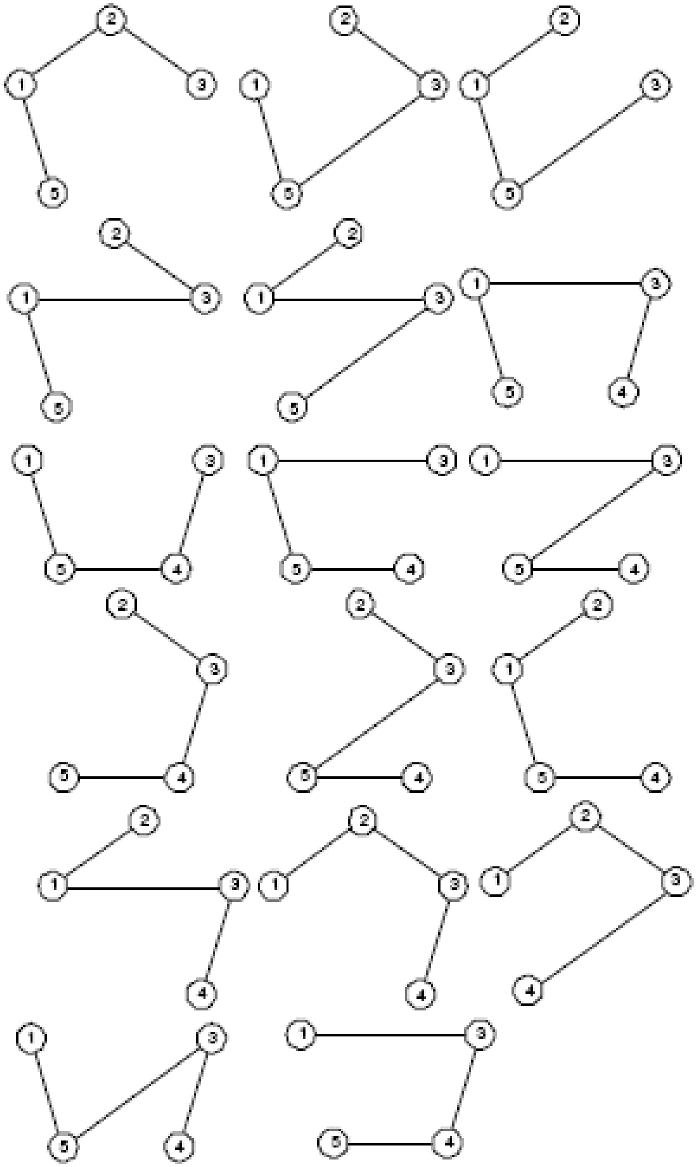
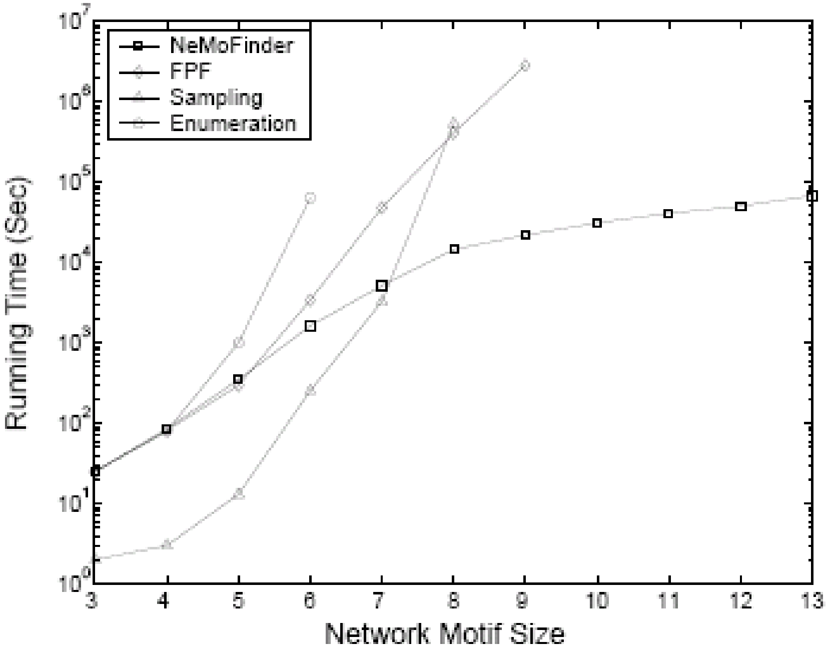
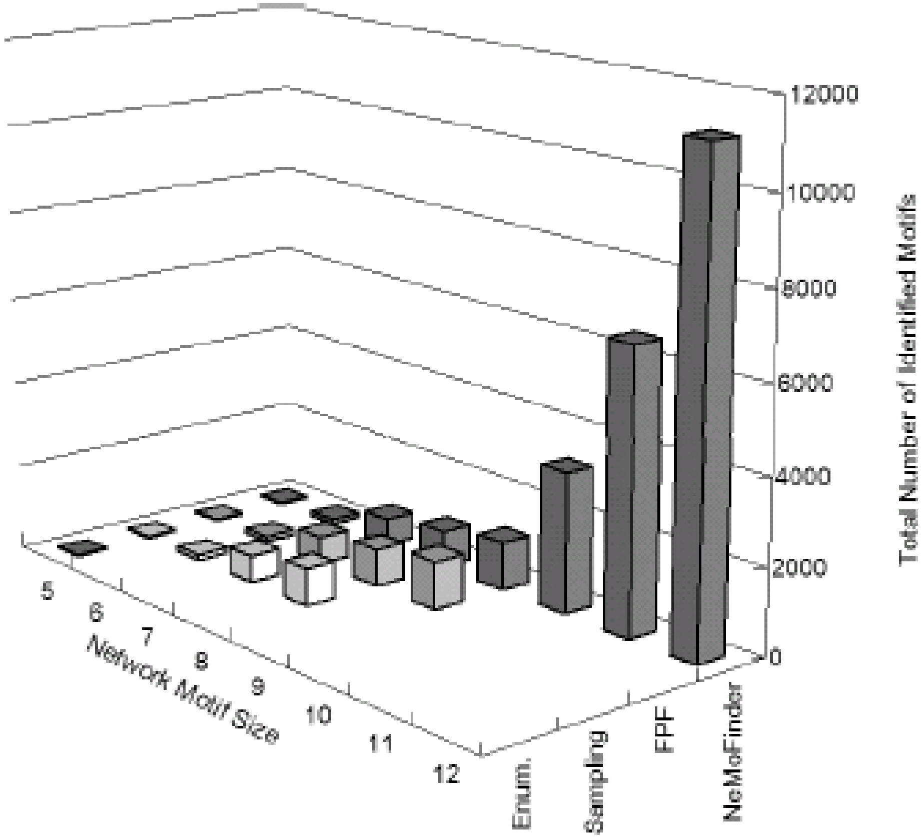


Figure 4: Occurrences of t_{4_2} in G.

Copyright 2006 © Limsoon Wong

NeMoFinder: Discovery of Meso-Scale Motifs





Motif Strength and PPI Reliability

- Strength of a size k motif g is

$$MS^k(g) = \frac{s_g \times f_g}{\max_k}$$

where $\max_k$ is max value of s_g x f_g over all size-k motifs

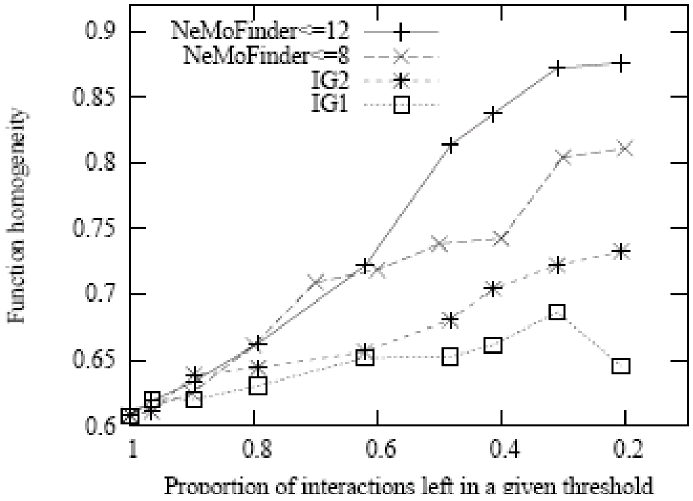
- Motif-strength PPI reliability index is an pair of possibly interacting protein X ↔Y is

$$I(X \leftrightarrow Y) = \sum_{k=2}^{K} \sum_{i=0}^{n} MS^k(g_i) \times k$$

where g_i are motifs involving the edge X ↔Y, and k is size of g_i

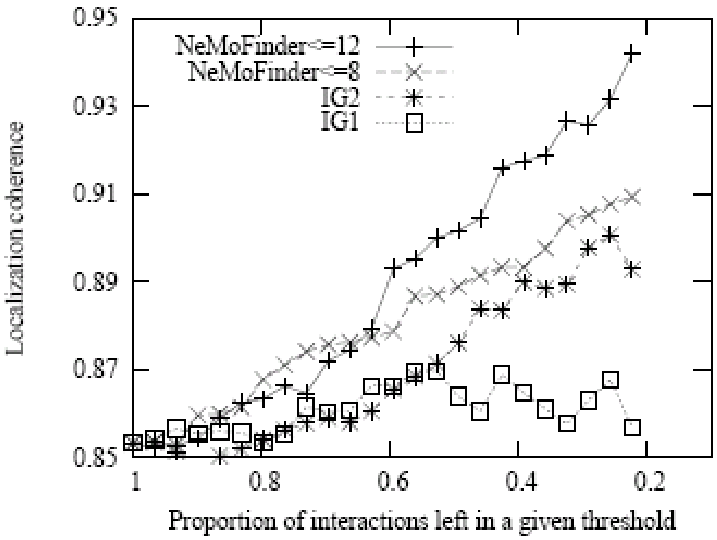
Copyright 2006 © Limsoon Wong

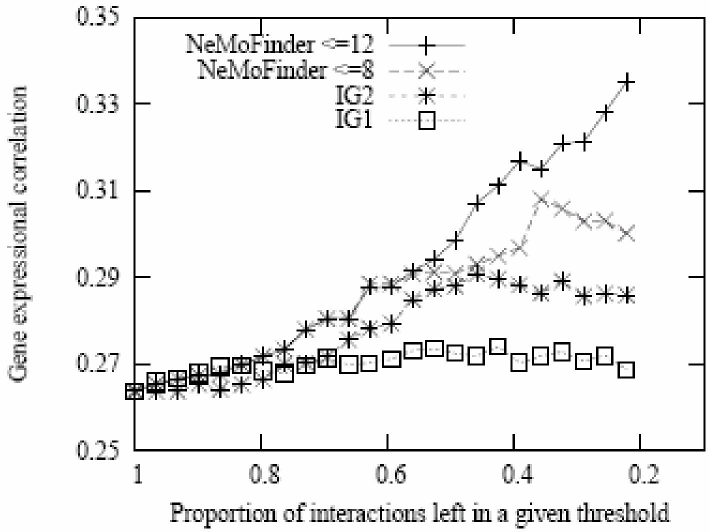
Evaluation wrt Common Cellular Role, etc



- **Motif-strength PPI reliability index correlates well to common cellular roles, localization, & expression**

⇒ works as well as "ipr"





Some Observations

- Meso-scale motifs are more reliable than small local motifs (c.f. “ig_2”)
- Similar performance to “ipr”, but may have advantages if network is sparse (i.e., where few alternate paths are present)

- Btw, this is the first time size-12 network motifs are known to be extracted from yeast PPI network

How about discovering false negatives?

The story of detecting missing information

False Negatives

- **A “false negative” is a failure to detect a real protein-protein interaction**

Copyright 2006 © Limsoon Wong

IPR Detects False Negatives

- **To find out if there is a “missing” interaction between X and Y, we do:**
  - compute ipr value of X↔Y in $G \cup \{X \leftrightarrow Y\}$
  - predict if X↔Y as false negative if “ipr” is high

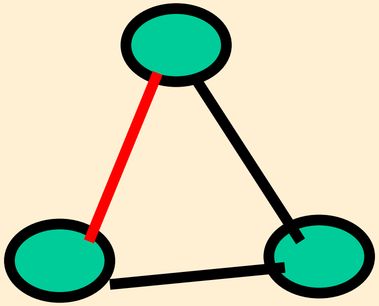
Copyright 2006 © Limsoon Wong

But needs an adjustment …
We call the adjusted index IRAP*

$$ipr^{\mathcal{G}}(X \leftrightarrow Y) = \max_{\phi \in \Phi^{\mathcal{G}}(X,Y)} \prod_{(U \leftrightarrow V) \in \phi} \left(1 - \frac{ig^{\mathcal{G}}(U \leftrightarrow V)}{ig^{\mathcal{G}}_{\max}}\right)$$

replace

“ig” is too generous, it always gives the red “missing” link the best score,



$$1 - \frac{ComNbr^{G}(U \leftrightarrow V)}{ComNbr^{G}_{\max}}$$

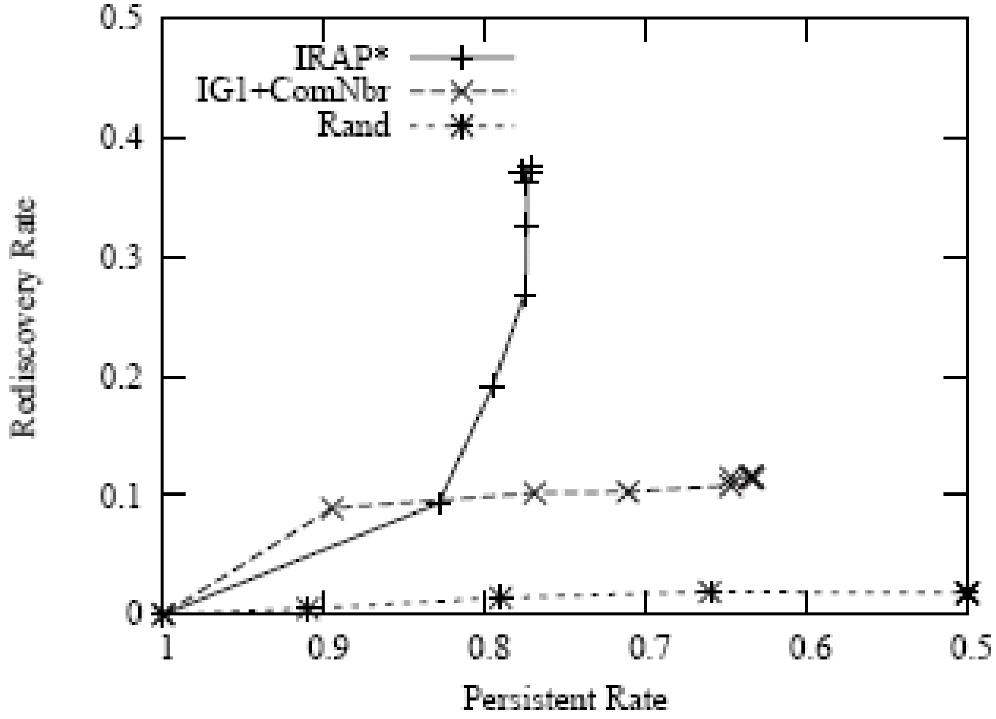
Where $ComNbr^{G}(U \leftrightarrow V)$ is number of common neighbours of U and V in G

Because proteins with a large number of shared partners tend interact themselves

How do we test if this works?

- **To test this, we mimic false negatives by random removal of 50% of high-quality known interactions. Then we check:**
  - – how many removed interactions are rediscovered?
  - – is there diff in rediscovery rates of false negative vs random links?
  - – Is there support in terms of gene expression correlation, common cellular roles, & common cellular locations?

IRAP* Persistence & Rediscovery Rates



- IRAP*: we iterate “ipr” and “irap*” 10 times to remove worst 5% of “false positives” and add best 5% of “false negatives”

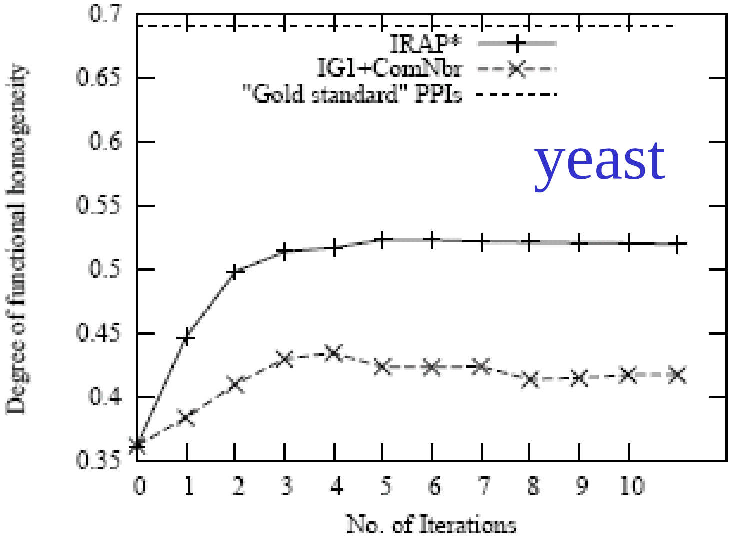
- IG1+ComNbr: we use “ig” to remove “false positives” and “ComNbr” to add “false negatives”, iterated 10 times

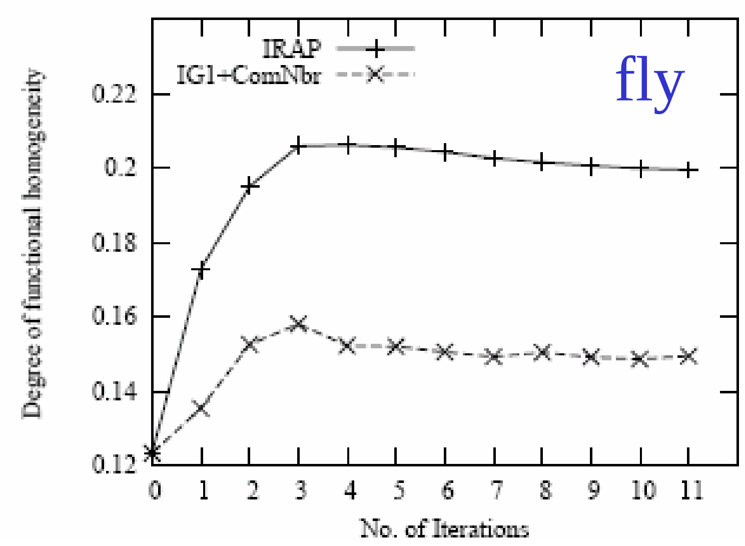
- Rand: randomly add and remove

About 40% of the high-quality “missing” interactions are rediscovered

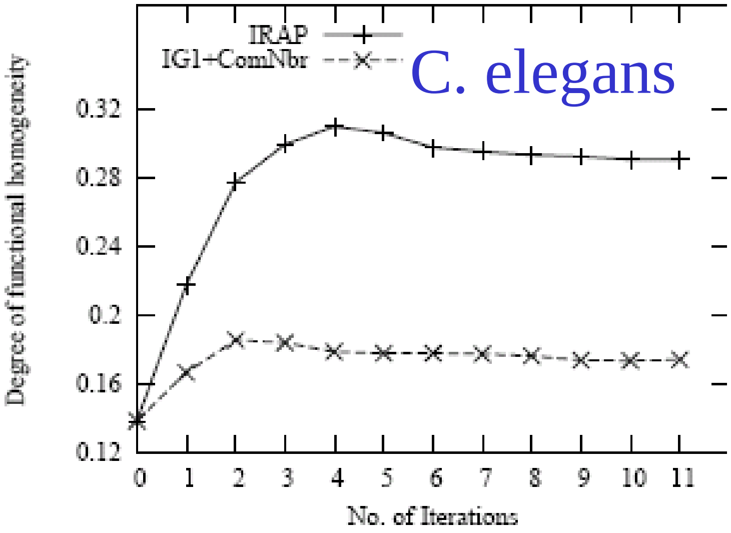
Copyright 2006 © Limsoon Wong

IRAP* Functional Coherence





The "false negatives" detected are functionally coherent.
I.e., IRAP* works



Copyright 2006 © Limsoon Wong

Conclusions

- There are latent local & global network “motifs” that indicate likelihood of protein interactions

- These network “motifs” can be exploited in computational elimination of false positives & false negatives from high-throughput Y2H expt & possibly other highly erroneous interaction data

- IPR & meso-scale motifs are the most effective topologically-based computational measure for assessing the reliability (false positives) of protein-protein interactions detected by high-throughput methods

- IPR/IRAP* can discover new interactions (false negatives) not detected in the expt PPI network

Copyright 2006 © Limsoon Wong

References

- **J. Chen et al,** “Systematic assessment of high-throughput protein interaction data using alternative path approach”, ***Proc. ICTAI 2004*, pages 368—372**
- **J. Chen et al,** “Towards discovering reliable protein interactions from high-throughput experimental data using network topology”, ***Artificial Intelligence in Medicine*, 35:37—47, 2005**
- **J. Chen et al,** “Increasing confidence of protein interactomes using network topological metrics”, ***Bioinformatics*, to appear**
- **J. Chen et al,** “NeMoFinder: Dissecting genome-wide protein-protein interactions with meso-scale network motifs”, ***Proc. KDD 2006***
- **H.N. Chua et al.** “Exploiting indirect neighbours and topological weight to predict protein function from protein-protein interactions”, ***Bioinformatics*, 22:1623—1630, 2006**
- **R. Saito et al,** “Interaction generality, a measurement to assess the reliability of a protein-protein interaction”, ***NAR* 30:1163--1168, 2002**
- **R. Saito et al,** “Construction of reliable protein-protein interaction networks with a new interaction generality measure”, ***Bioinformatics* 19:756--763, 2003**
- **S. Maslov & K. Sneppen,** “Specificity and stability in topology of protein networks”, ***Science* 296:910--913, 2002**
- **E. Sprinzak et al,** “How reliable are experimental protein-protein interaction data?”, ***JMB* 327:919--923, 2003**

Copyright 2006 © Limsoon Wong