



BIOINFORMÁTICA Y BIOLOGÍA COMPUTACIONAL



Universidad
Complutense
Madrid

Curso de la Escuela Complutense de Verano 2007

Redes de Interacciones entre Proteínas

Florencio Pazos (CNB-CSIC)

Florencio Pazos Cabaleiro
Computational Systems Biology Group
Centro Nacional de Biotecnología (CNB-CSIC)
pazos@cnb.uam.es
<http://pdg.cnb.uam.es>



Redes de Interacciones entre Proteínas

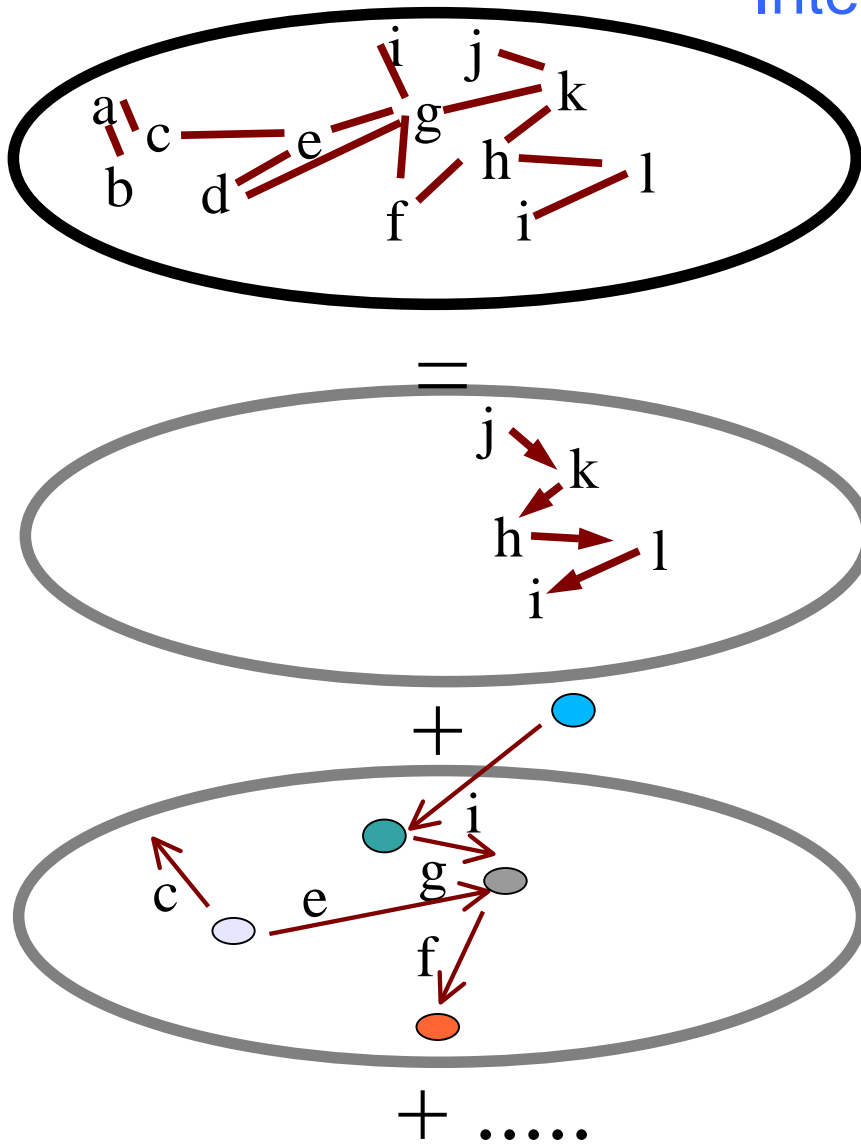
- El Interactoma
 - Determinación experimental masiva del interactoma
 - Estudios globales del interactoma
 - Características topológicas
 - Nodos (proteínas) importantes topologicamente
 - Origen de la topología
 - Motivos topológicos
 - Características funcionales
 - Resumen
 - Calidad de los datos masivos de interacciones
 - Métodos computacionales para predicción de interacciones
 - Conservación de cercanía genómica
 - Fusión génica
 - Perfiles filogenéticos
 - Similitud de árboles filogenéticos
 - Repositorios *on-line* de interacciones
 - Bibliografía
-

Biología de Sistemas

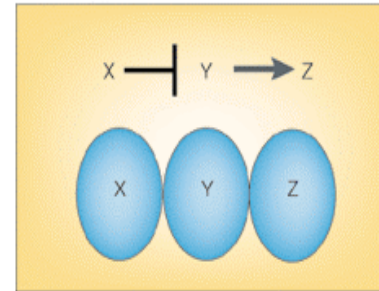
Caracterización masiva de componentes moleculares y sus relaciones

- Genome sequencing (“genome”).
- Transcript characterization (mRNA) (“transcriptome”)
- Characteristics of the protein repertory (“proteome”)
- Cellular localization of the components (“localizome”)
- Gene regulation network (“regulome”)
- **Protein interaction network (“interactome”)**
- Massive gene-phenotype studies (“fenoma”)
- Metabolic networks (“metabolome”)
-

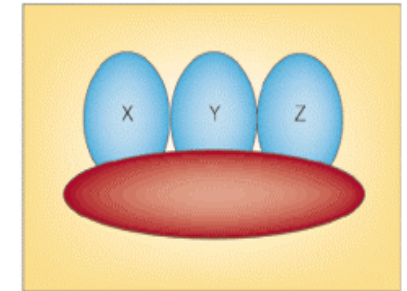
Interactoma



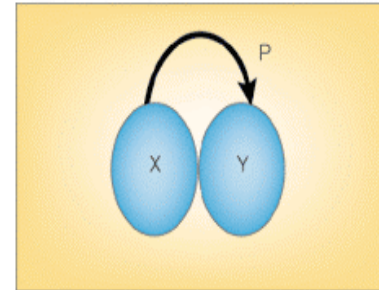
a Genetic pathways



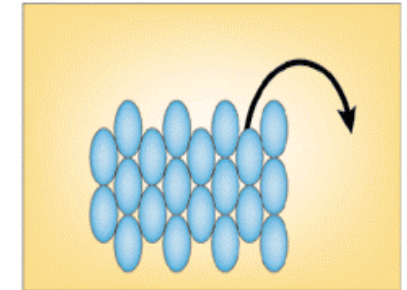
b Pathway scaffolding



c Enzymatic reactions

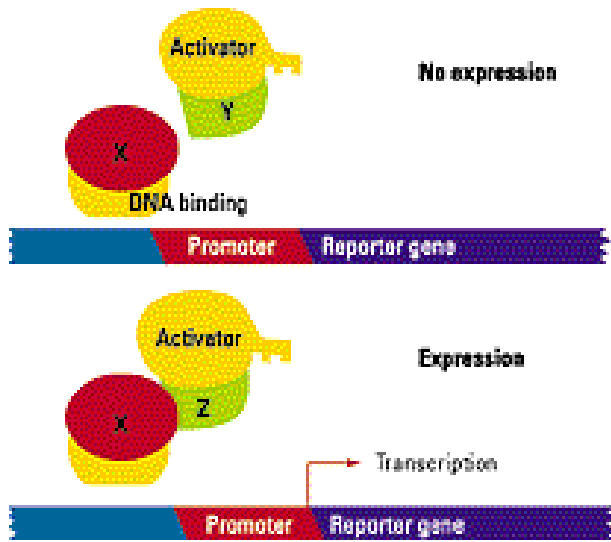


d Molecular machines

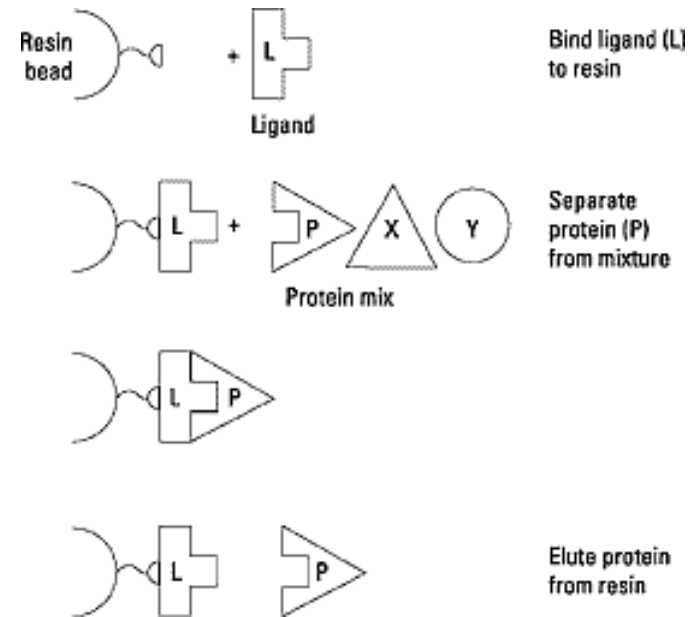


Nature Reviews | Molecular Cell Biology

Determinación Experimental Masiva del Interactoma



Y2H

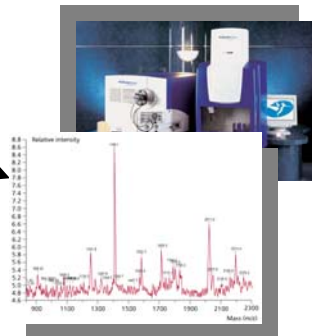
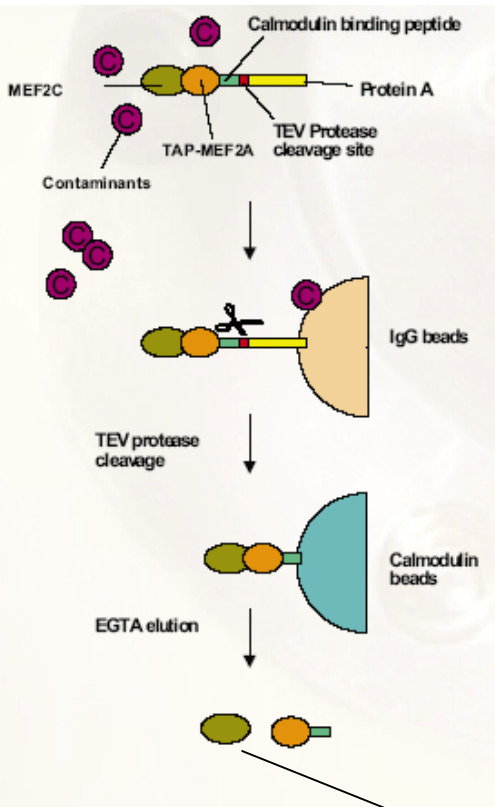


TAP/MS

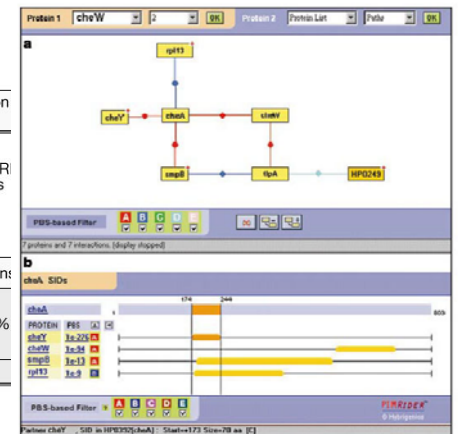
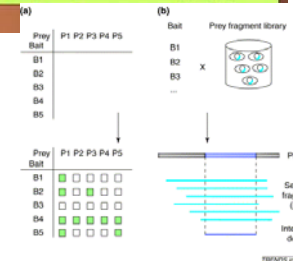
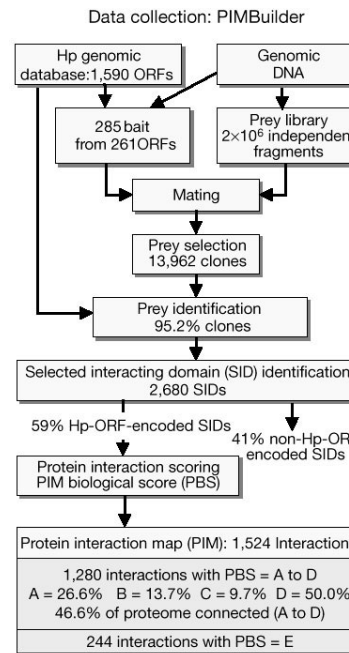
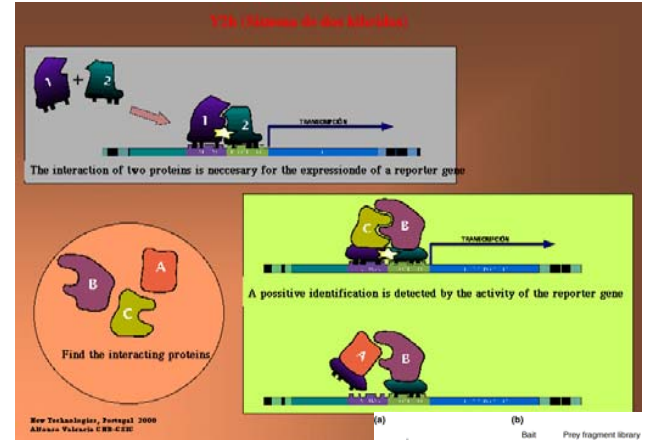
<http://pubs.acs.org/hotartcl/mdd/00/sep/edwards.html>

Determinación Experimental Masiva del Interactoma

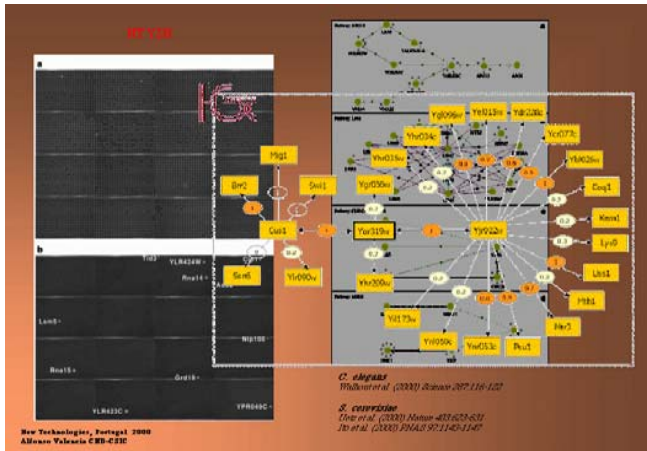
TAP/MS



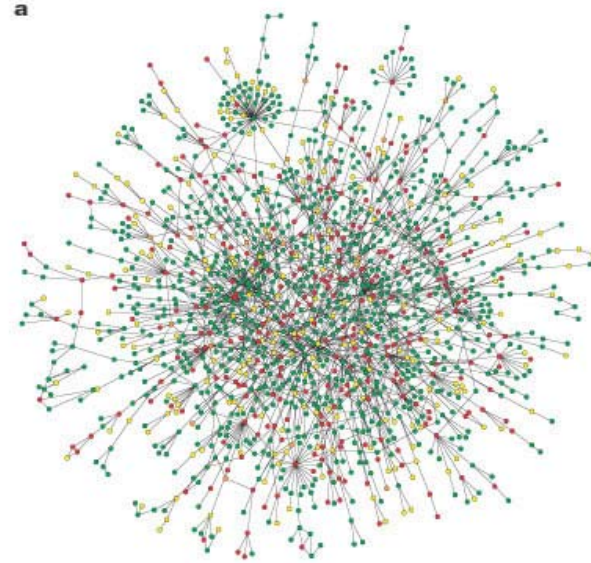
Y2H



Interactomas determinados experimentalmente (*high throughput*)



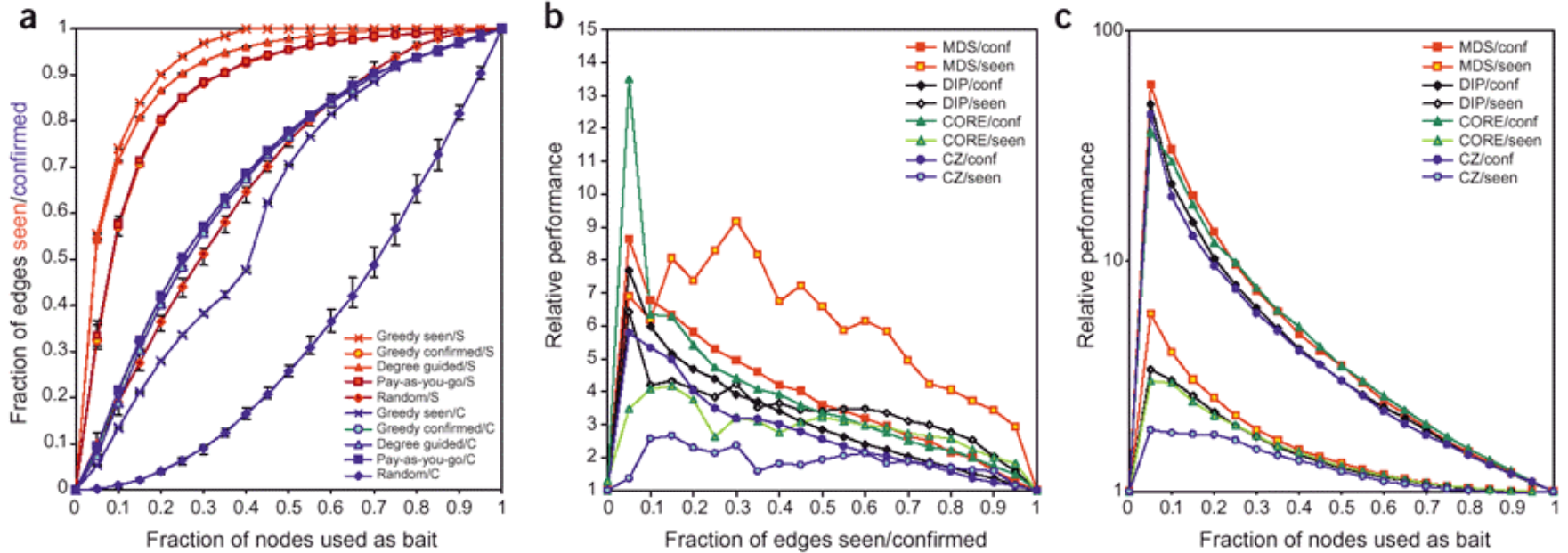
A.Valencia



- Rain, J.C., Selig, L., De Reuse, H., et al. (2001) The protein-protein interaction map of *Helicobacter pylori*. *Nature*, **409**, 211-215.
- Gavin, A.C., et al. (2002) Functional organisation of the yeast proteome by systematic analysis of protein complexes. *Nature*, **415**, 141-147.
- Ho, Y., et al. (2002) Systematic identification of protein complexes in *Saccharomyces cerevisiae* by mass spectrometry. *Nature*, **415**, 180-183.
- Ito, T., et al. (2000) Toward a protein-protein interaction map of the budding yeast: A comprehensive system to examine two-hybrid interactions in all possible combinations between the yeast proteins. *Proc Natl Acad Sci USA*, **97**, 1143-1147.
- Uetz, P., et al. (2000) A comprehensive analysis of protein-protein interactions in *Saccharomyces cerevisiae*. *Nature*, **403**, 623-631.
- Giot, L., Bader, J.S., Brouwer, et al. (2003) A protein interaction map of *Drosophila melanogaster*. *Science*, **302**, 1727-1736.
- Li, S., Armstrong, C.M., Bertin, N., et al. (2004) A map of the interactome network of the metazoan *C. elegans*. *Science*, **303**, 540-543.
- Butland, G., Peregrin-Alvarez, J.M., Li, J., et al. (2005) Interaction network containing conserved and essential protein complexes in *Escherichia coli*. *Nature*, **433**, 531-537.
- Rual, J.F., Venkatesan, K., Hao, T., Hirozane-Kishikawa, T., Dricot, A., Li, N., Berriz, G.F., Gibbons, F.D., Dreze, M., Ayivi-Guedehoussou, N., et al. (2005) Towards a proteome-scale map of the human protein-protein interaction network. *Nature*, **437**, 1173-1178.
- LaCount, D.J., Vignali, M., Chettier, R., Phansalkar, A., Bell, R., Hesselberth, J.R., Schoenfeld, L.W., Ota, I., Sahasrabudhe, S., Kurschner, C., et al. (2005) A protein interaction network of the malaria parasite *Plasmodium falciparum*. *Nature*, **438**, 103-107.
- Uetz, P., Dong, Y.A., Zeretzke, C., Atzler, C., Baiker, A., Berger, B., Rajagopala, S.V., Roupelieva, M., Rose, D., Fossum, E., et al. (2006) Herpesviral protein networks and their interaction with the human proteome. *Science*, **311**, 239-242.

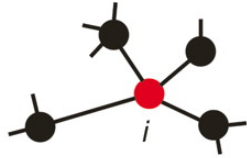
Diseño del experimento

Selección de Anzuelos (*Baits*)



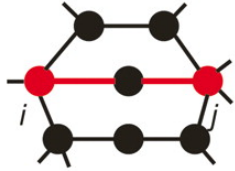
Propiedades Topológicas Globales del Interactoma

Parámetros Topológicos



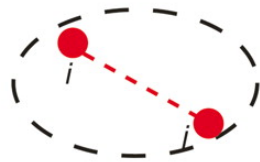
Degree

k_i = number of links connected to node i



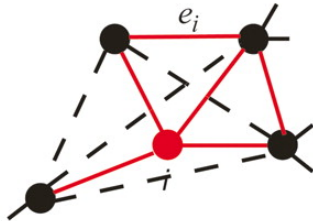
Distance

d_{ij} = shortest path length between node i and j



Diameter

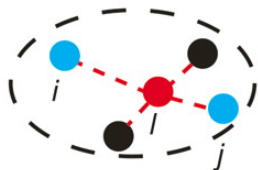
$D = \max \{ d_{ij} \mid i, j \in N \}$ N : all nodes in the network



Clustering Coefficient

$$c_i = \frac{2e_i}{k_i(k_i - 1)}$$

e_i : number of existing links (labeled in red) among the k_i nodes that connect to node i

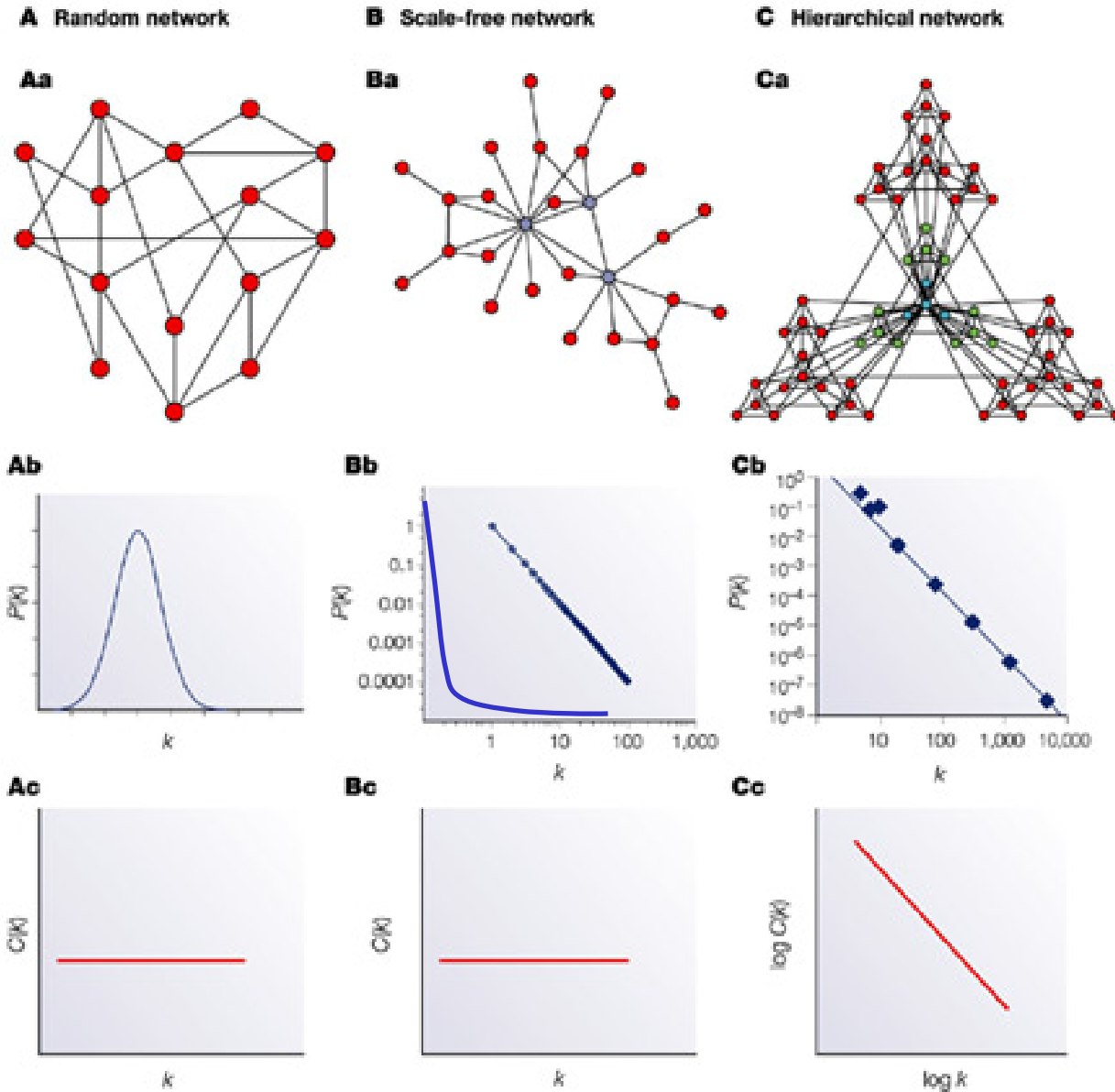


Betweenness

$b_l = \sum_{ij} p_{ij}(l) / p_{ij}$ p_{ij} : number of shortest paths between i and j

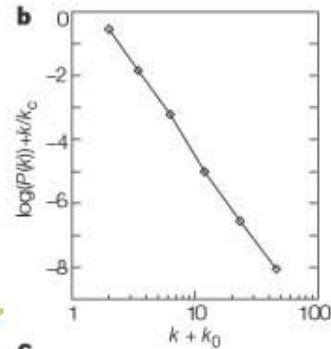
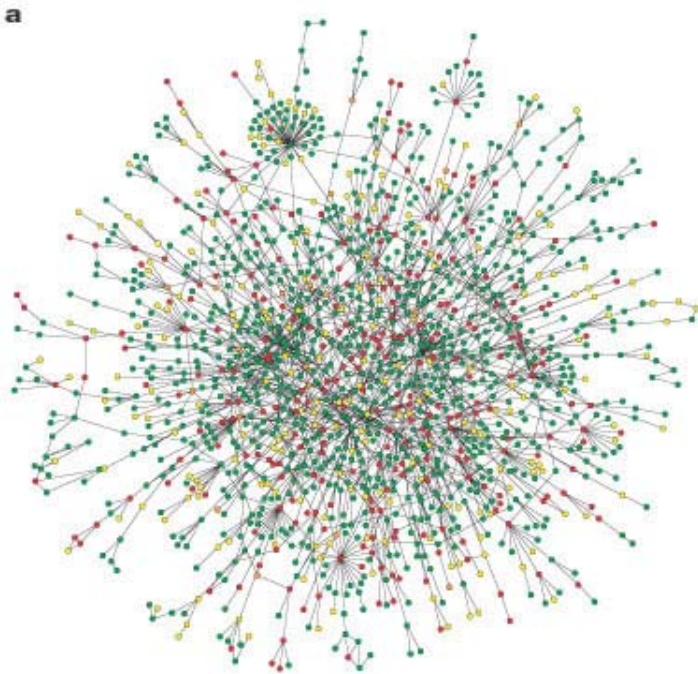
$p_{ij}(l)$: number of shortest paths between i and j going through node l

Propiedades Topológicas Globales del Interactoma

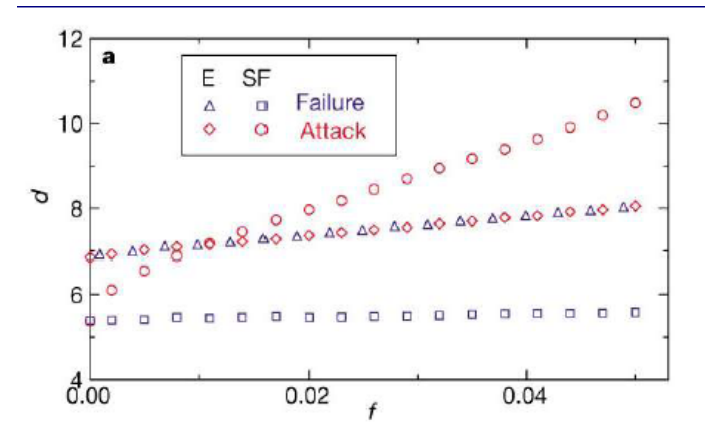


Propiedades topológicas globales del interactoma

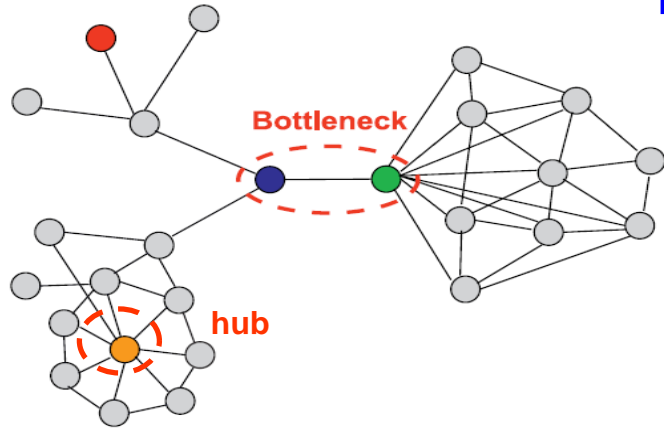
Red *Scale-Free*/Jerárquica



c

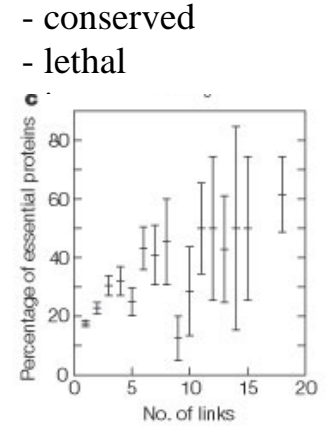
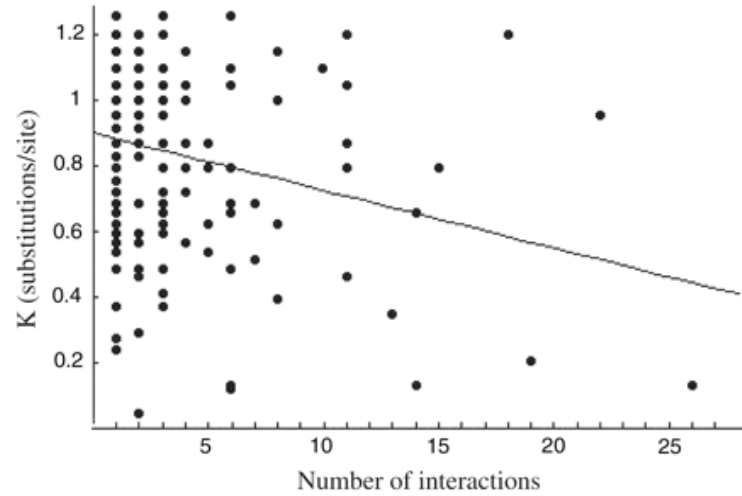


Nodos Importantes

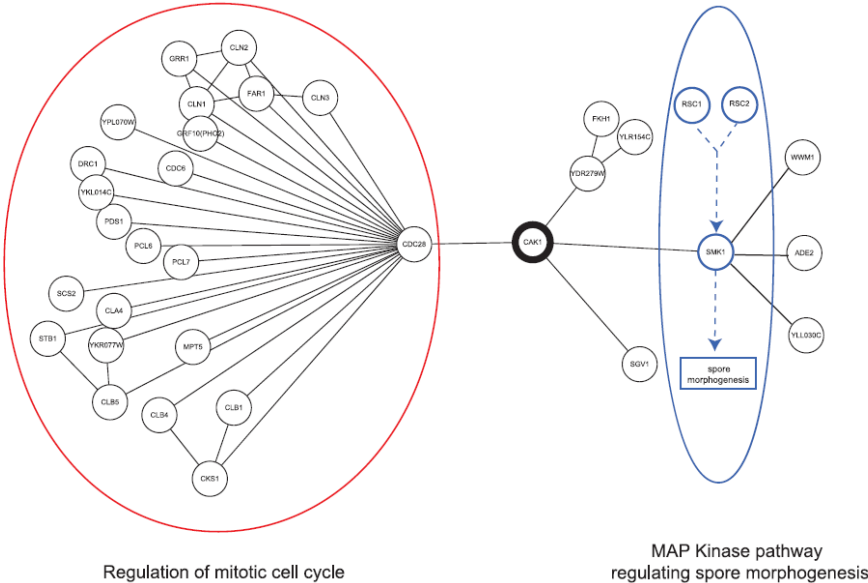


- Hub-bottleneck node
- Non-hub-bottleneck node
- Hub-non-bottleneck node
- Non-hub-non-bottleneck node

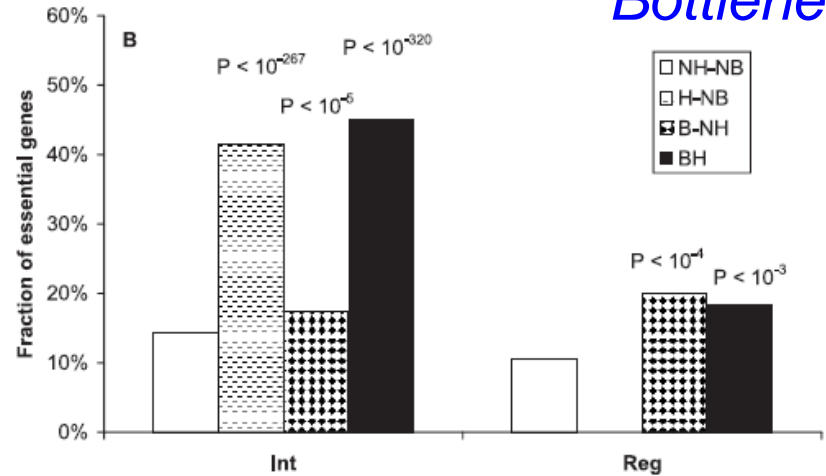
Hubs



Fraser, H.B., Hirsh, A.E., Steinmetz, L.M., Scharfe, C. and Feldman, M.W. (2002) Evolutionary rate in the protein interaction network. *Science*, **296**, 750-752.

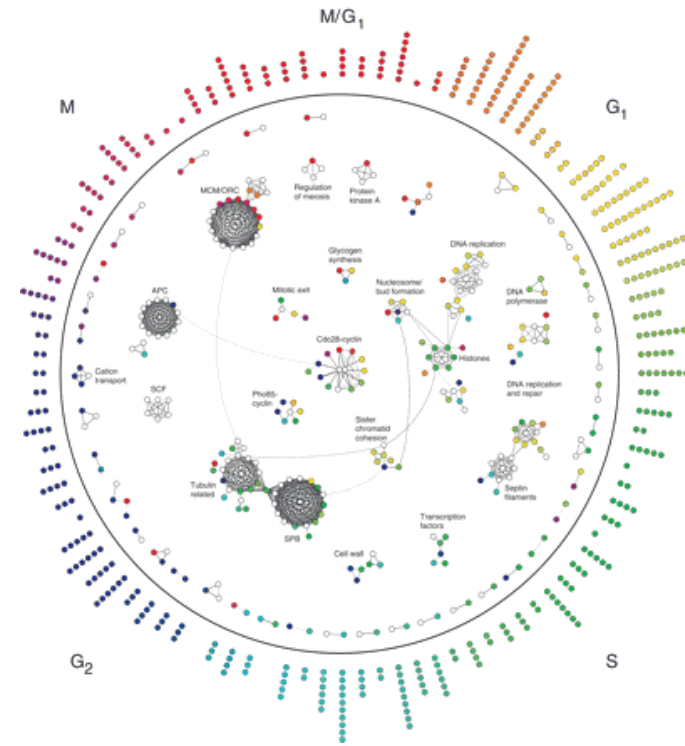
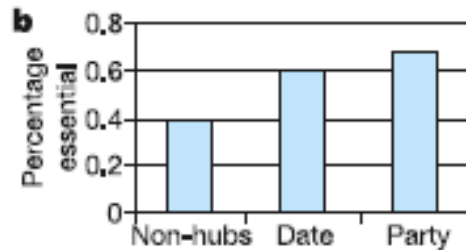
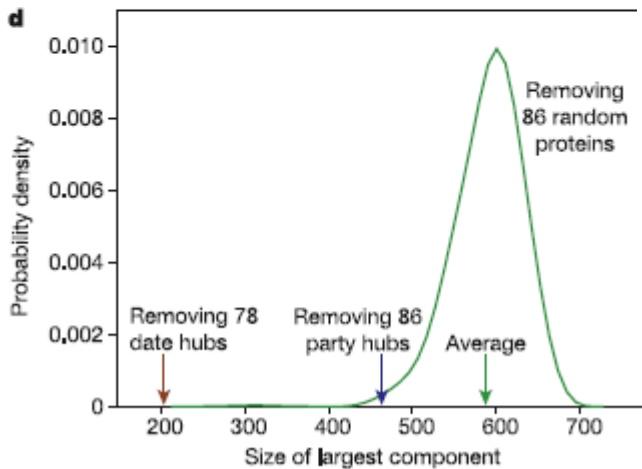
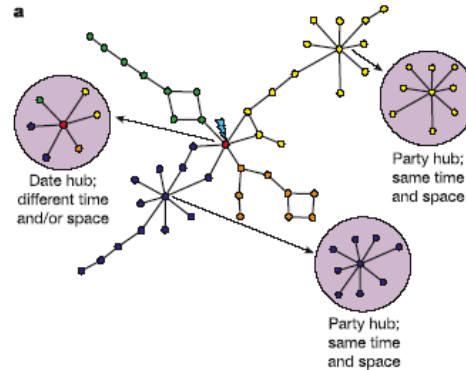
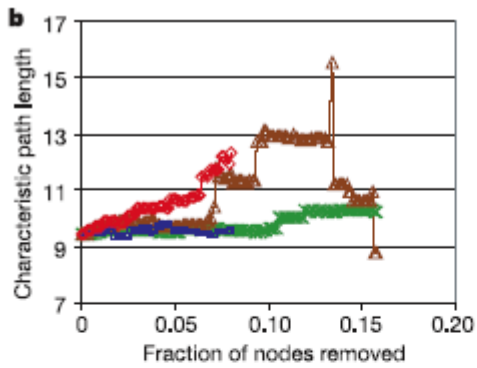


Bottlenecks



Yu, H., Kim, P.M., Sprecher, E., Trifonov, V. and Gerstein, M. (2007) The importance of bottlenecks in protein networks: correlation with gene essentiality and expression dynamics. *PLoS Comput Biol.*, **3**, e59

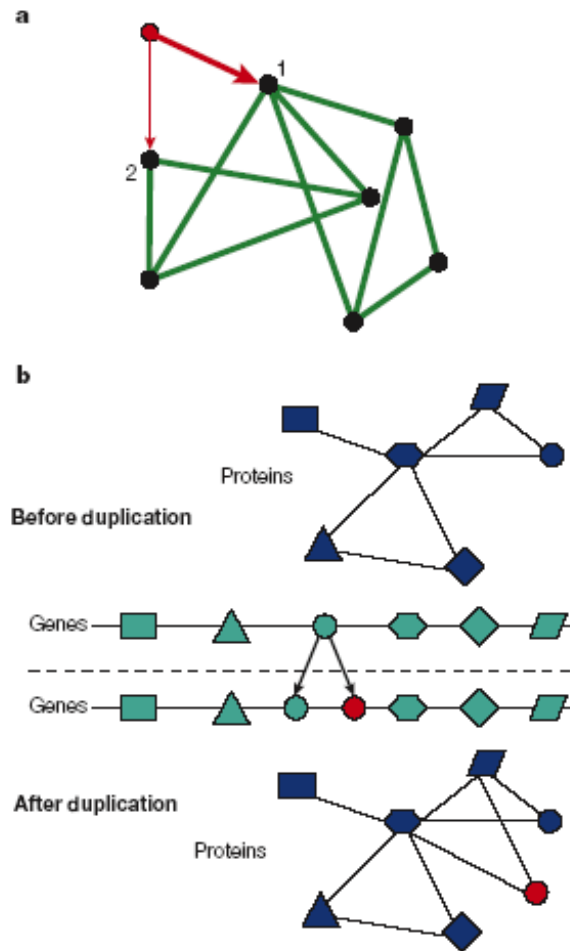
Añadiendo Información Dinámica



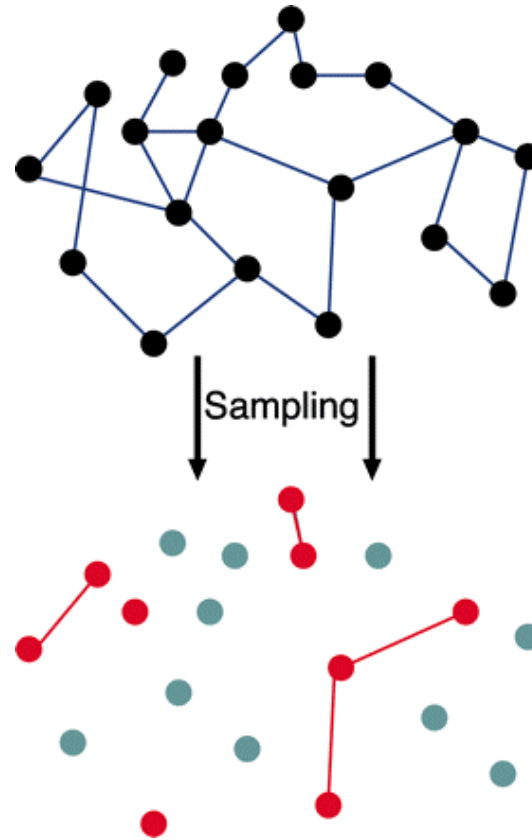
Han, J.D., Bertin, N., Hao, T., Goldberg, D.S., Berriz, G.F., Zhang, L.V., Dupuy, D., Walhout, A.J., Cusick, M.E., Roth, F.P. and Vidal, M. (2004) Evidence for dynamically organized modularity in the yeast protein-protein interaction network. *Nature*, **430**, 88-93. Epub 2004 Jun 2009.

de Lichtenberg U, Jensen LJ, Brunak S, Bork P. (2005). Dynamic complex formation during the yeast cell cycle. *Science*. **307(5710)**:724-727.

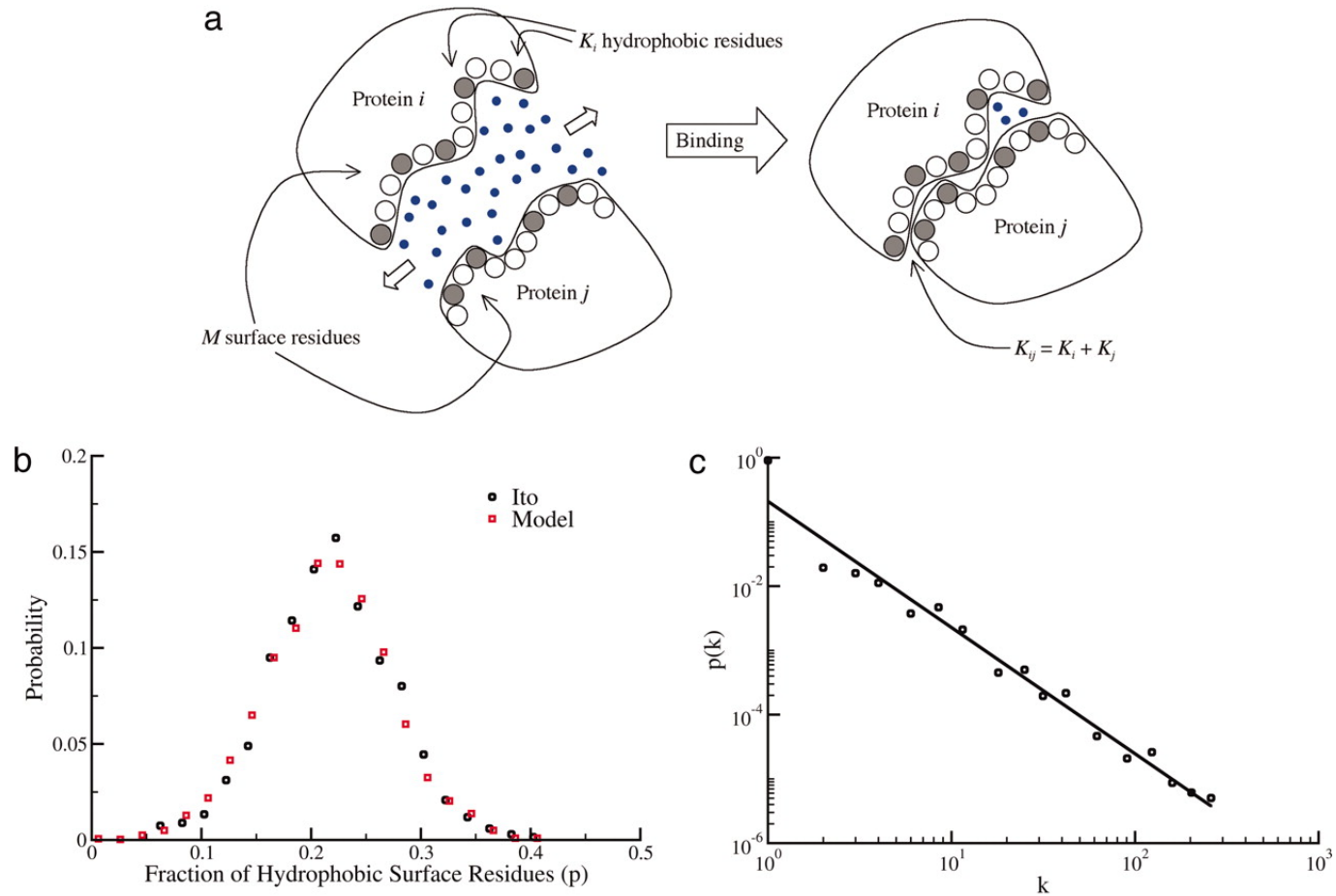
Aparición de la Estructura *Scale-Free*



¿Artefactos debido al muestreo (sampling)?



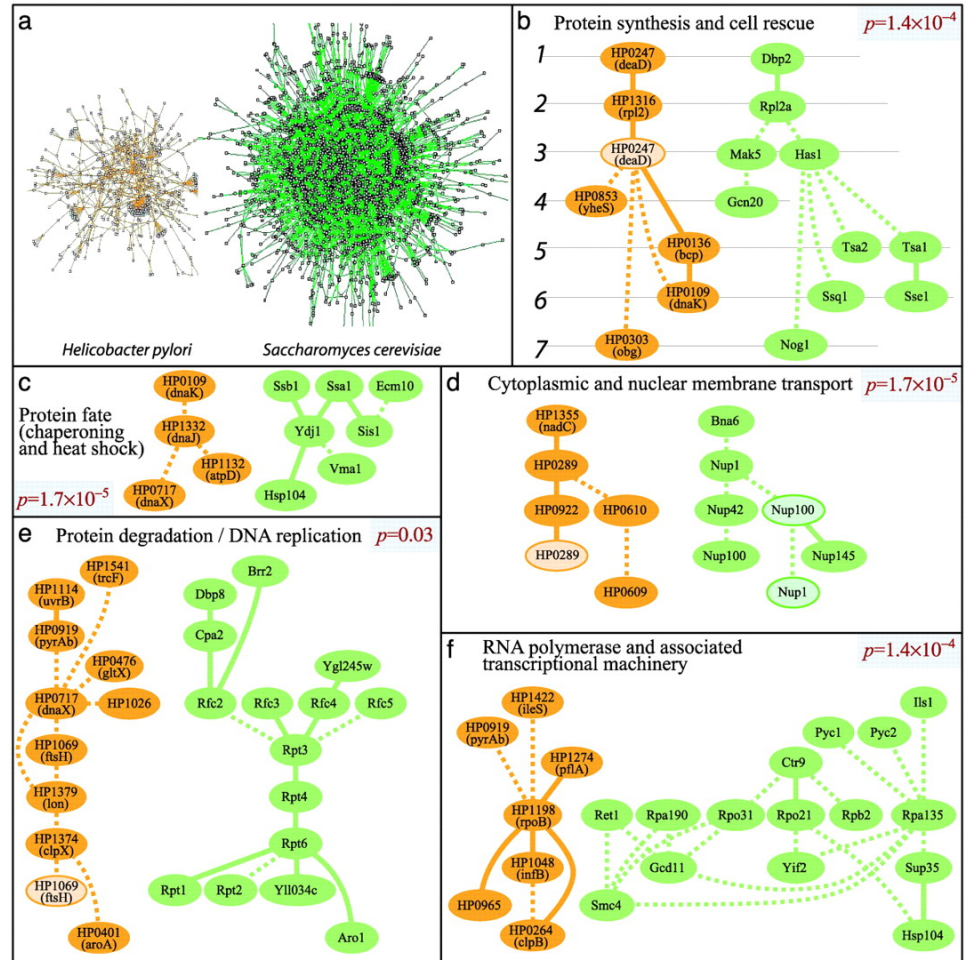
¿Otros Artefactos?



Motivos en la Red de Interacciones

Table 1 Evolutionary conservation of motif constituents

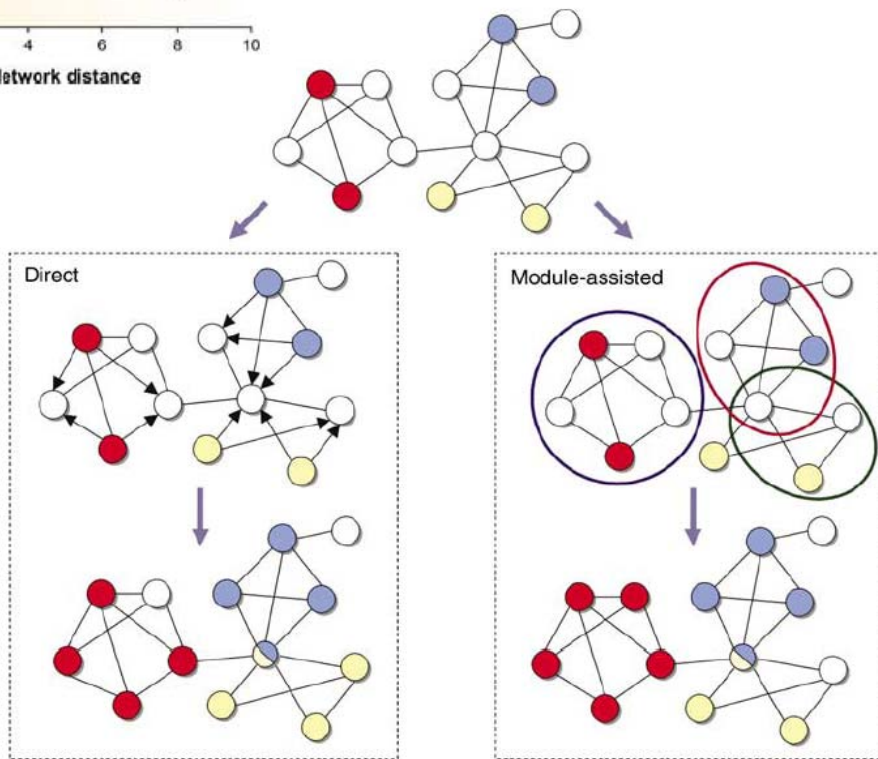
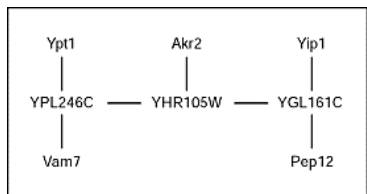
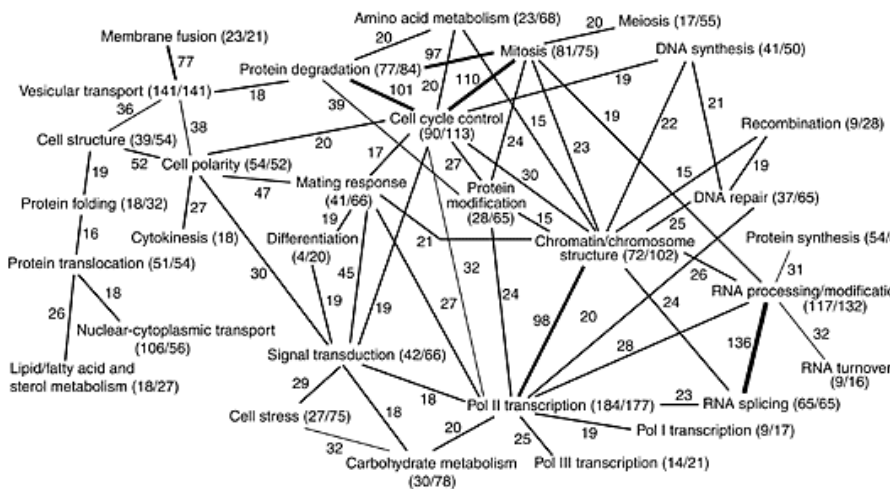
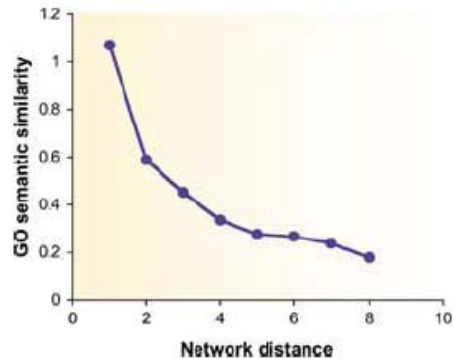
#	Motifs	Number of yeast motifs	Natural conservation rate	Random conservation rate	Conservation ratio
1		9,266	13.67%	4.63%	2.94
2		167,304	4.99%	0.81%	6.15
3		3,846	20.51%	1.01%	20.28
4		3,649,591	0.73%	0.12%	5.87
5		1,763,891	2.64%	0.18%	14.67
6		9,646	6.71%	0.17%	40.44
7		164,075	7.67%	0.17%	45.56
8		12,423	18.68%	0.12%	157.89
9		2,339	32.53%	0.08%	422.78
10		25,749	14.77%	0.05%	279.71
11		1,433	47.24%	0.02%	2,256.67



Wuchty, S., Oltvai, Z.N. & Barabasi, A.L. (2003) Evolutionary conservation of motif constituents in the yeast protein interaction network. *Nat Genet*, **35**, 176-179.

Kelley, B.P., *et al.* (2003) Conserved pathways within bacteria and yeast as revealed by global protein network alignment. *Proc Natl Acad Sci U S A*, **100**, 11394-11399.

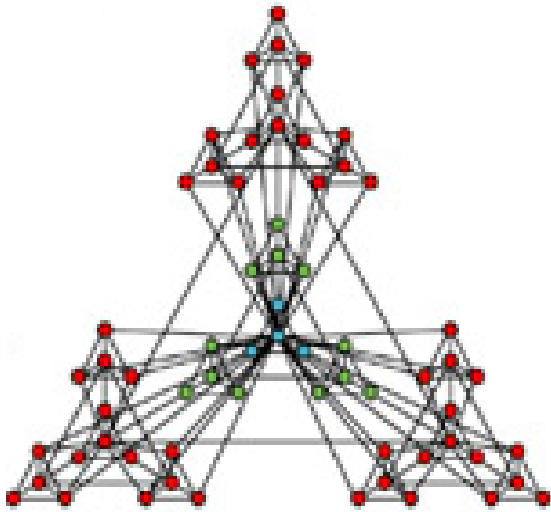
Predicción de Función Basada en el Contexto de Interacciones



Schwikowski, B., Uetz, P. & Fields, S. (2002). A network of protein-protein interactions in yeast. *Nature Biotech* **18**, 1257-1261.

Sharan, R., Ulitsky, I. and Shamir, R. (2007) Network-based prediction of protein function. *Mol Syst Biol.*, **3**, 88.

Estudios globales de la red de interacciones - Resumen



- scale-free / jerárquica
=> resistente a fallos aleatorios; caminos cortos
Topología scale-free explicable por duplicaciones
- hubs: esenciales/conservados
(date/party)
- modulos topológicos <> modulos funcionales
- motivos (pequeños) funcionales conservados

Calidad de los datos de interacciones *high throughput*

Table 1
Large protein interaction screens for eukaryotes

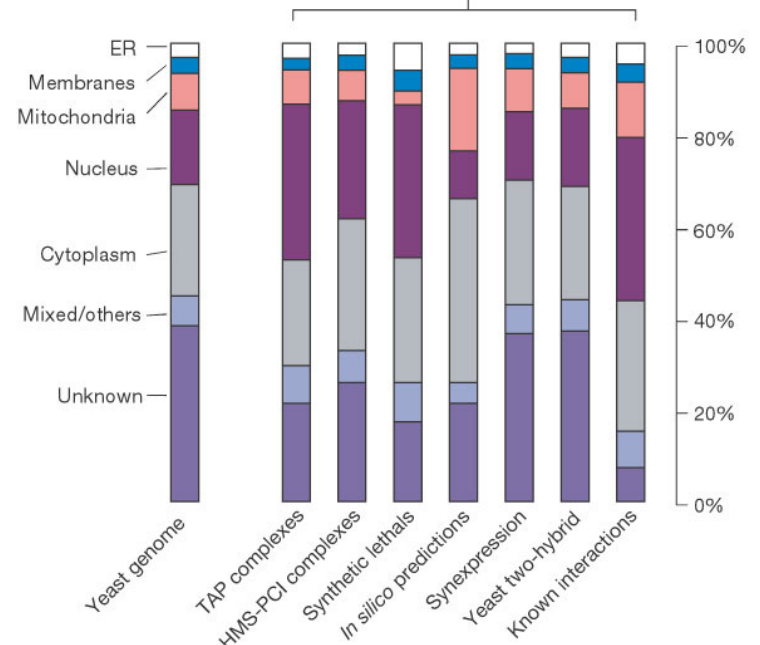
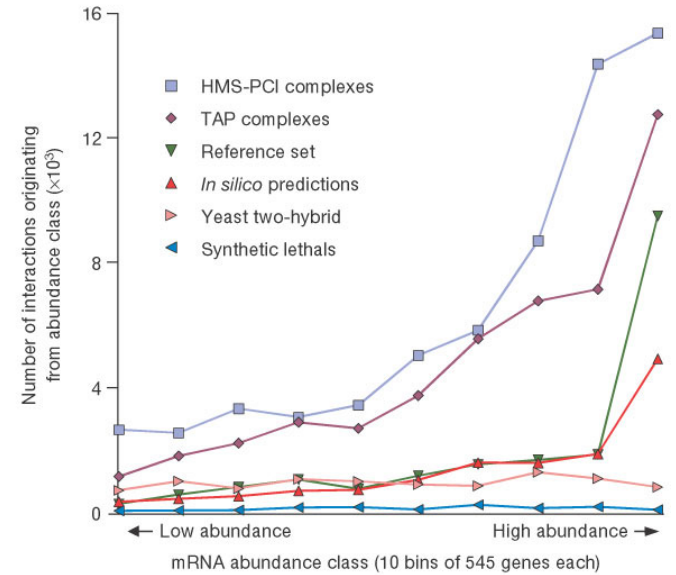
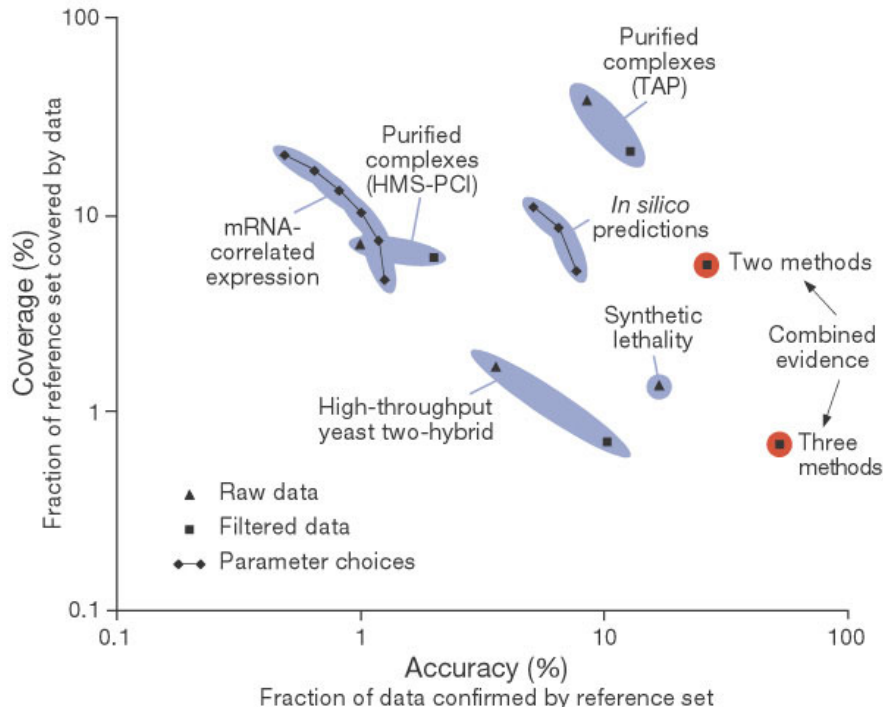
Organism (genes)	Method	Interactions ^a	Proteins	Reference
Yeast (~6000)	Yeast two-hybrid	967	1004	[63]
	Yeast two-hybrid	4549	3278	[13]
	Yeast two-hybrid	420	271	[64,65]
	Co-AP/MS	9421	1665	[66]
	Co-AP/MS	3878	1578	[67]
<i>Drosophila</i> (~14 000)	Yeast two-hybrid	20 405	7048	[49]
	Yeast two-hybrid	1814	488	[14]
Worm (~20 000)	Yeast two-hybrid	4027	1926	[68]

Overlap:
6 int !

^aFor two-hybrid screens, the approximate number of unique binary interactions is shown. For co-AP/MS screens, the approximate number of binary interactions that would result if each bait protein contacted every protein that co-purified with it (the “hub and spoke” model) is shown. Data can be retrieved from one of the databases cited [42–44].

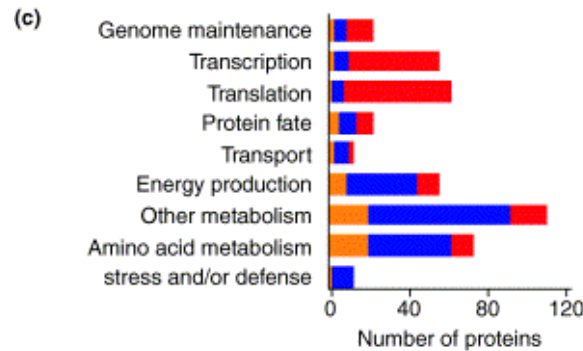
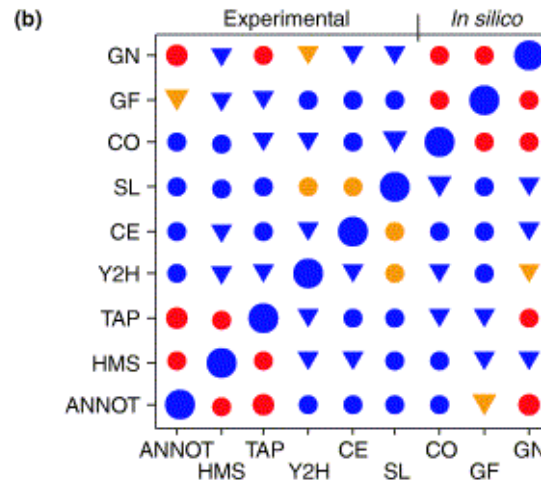
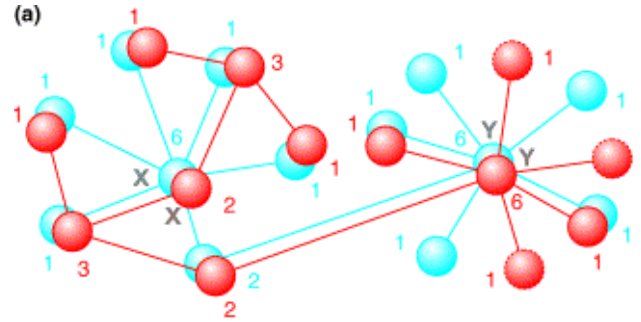
Estimation (yeast): 12.000-40000 (6000)

Calidad de los datos de interacciones *high throughput*

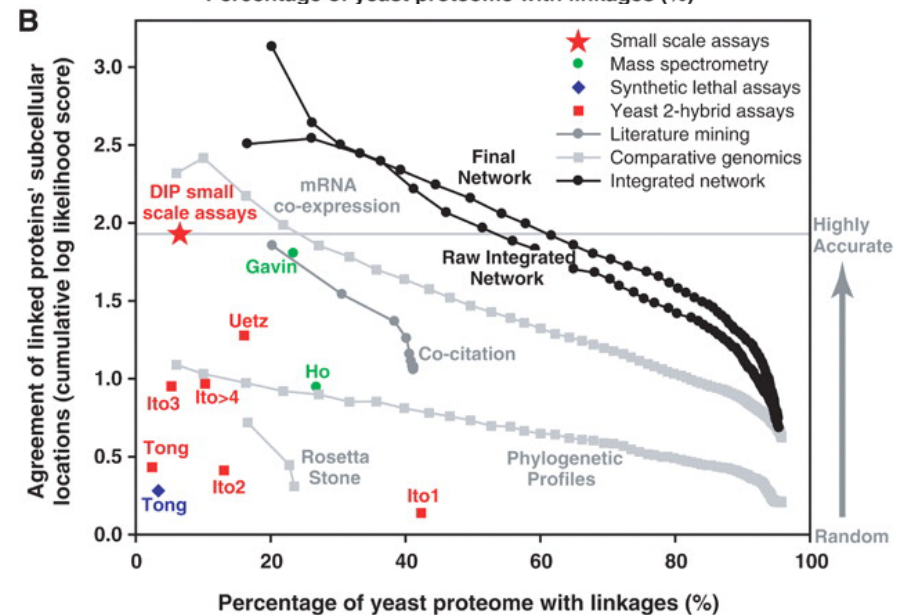
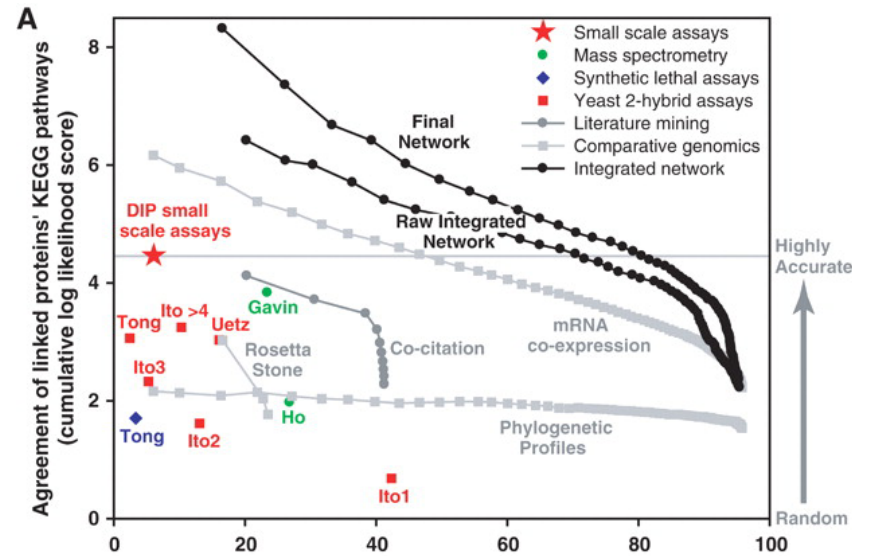
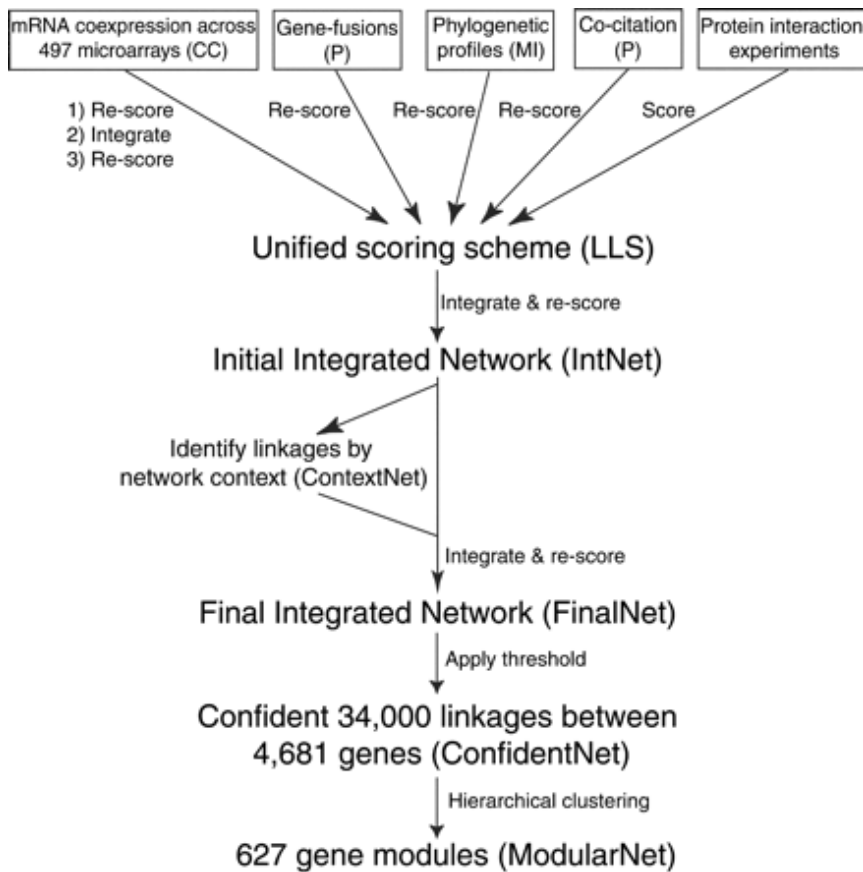


von Mering, C., Krause, R., Snel, B., Cornell, M., Oliver, S.G., Fields, S. and Bork, P. (2002) Comparative assessment of large scale data sets of protein-protein interactions. *Nature*, **417**, 399-403.

Calidad de los datos de interacciones *high throughput*



Combinación con Otras Fuentes de Información para Aumentar Fiabilidad



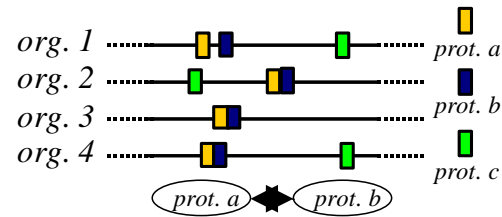
Predicción Computacional de Interacciones entre Proteínas

a) phylogenetic profiles

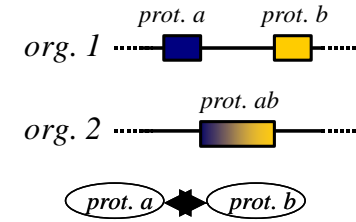
	<i>prot. a</i>	<i>prot. b</i>	<i>prot. c</i>	<i>prot. d</i>
<i>org. 1</i>	1	1	1	1
<i>org. 2</i>	0	1	0	1
<i>org. 3</i>	1	0	1	0
<i>org. 4</i>	1	0	1	1

prot. a
★
prot. c

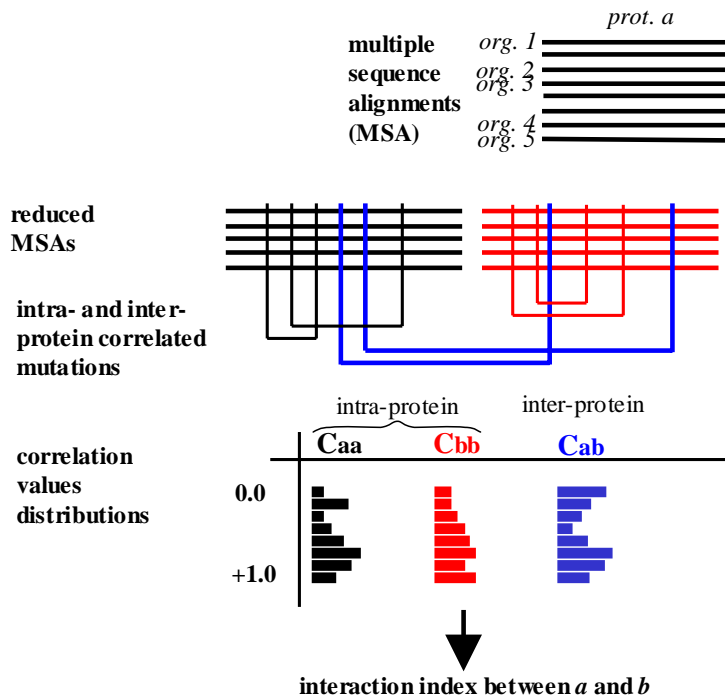
b) conservation of gene neighbouring



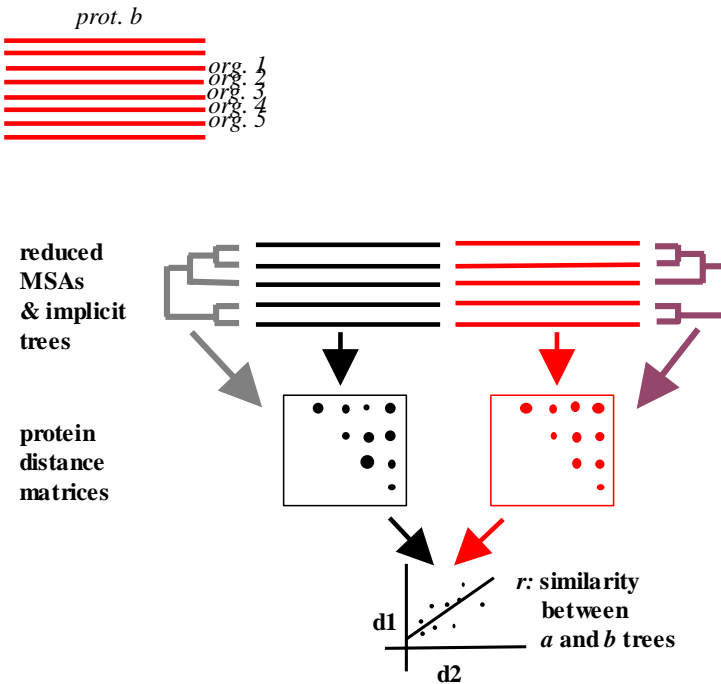
c) gene fusion



e) correlated mutations

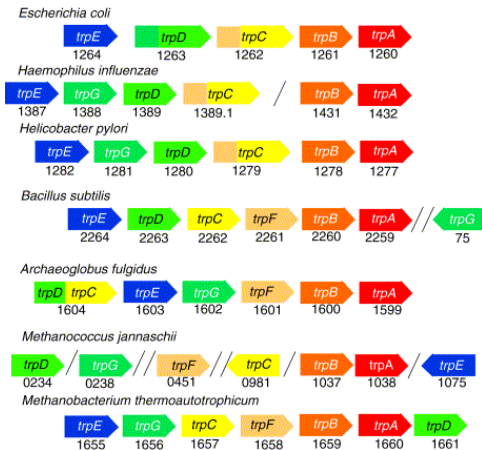
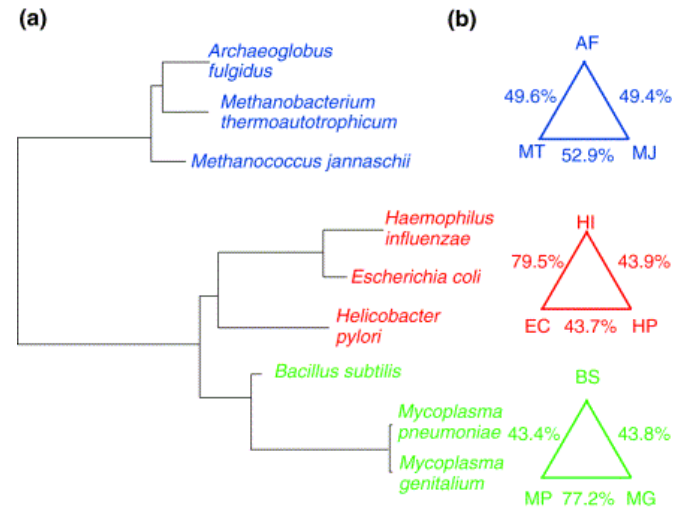
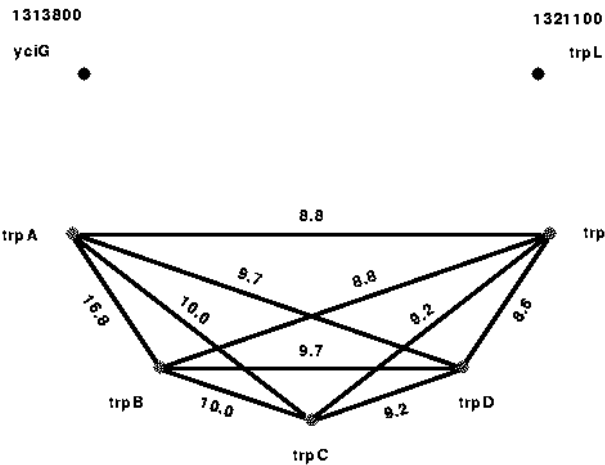
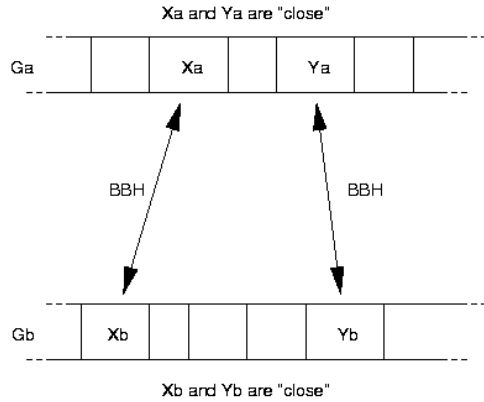


d) similarity of phylogenetic trees



- Huynen, M., Snel, B., Lathe, W. & Bork, P. (2000) Predicting protein function by genomic context: quantitative evaluation and qualitative inferences. *Genome Res*, **10**, 1204-1210.
- Valencia, A. & Pazos, F. (2002) Computational methods for the prediction of protein interactions. *Curr Opin Struct Biol*, **12**, 368-373.
- Salwinski, L. & Eisenberg, D. (2003). Computational methods of analysis of protein-protein interactions. *Curr Opin Struct Biol*. **13**, 377-382.

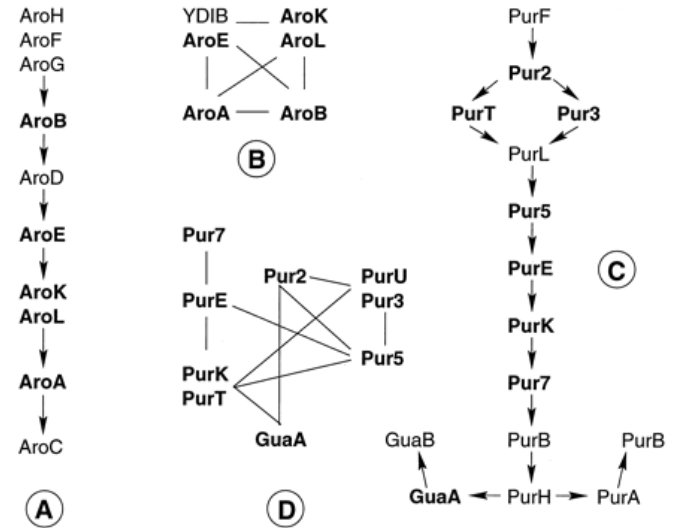
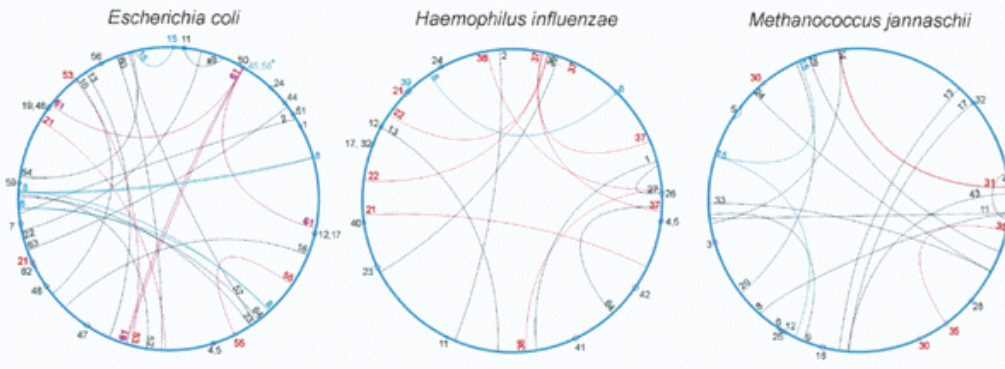
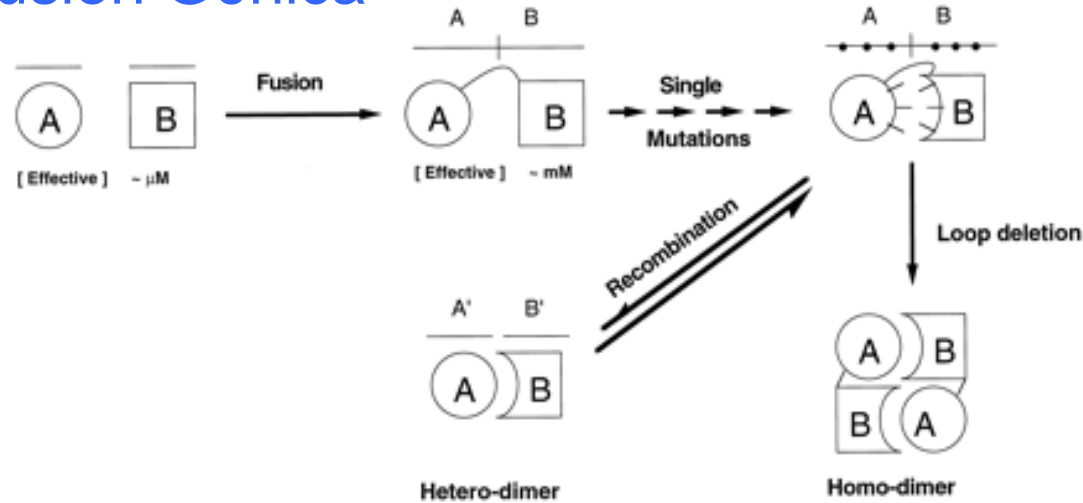
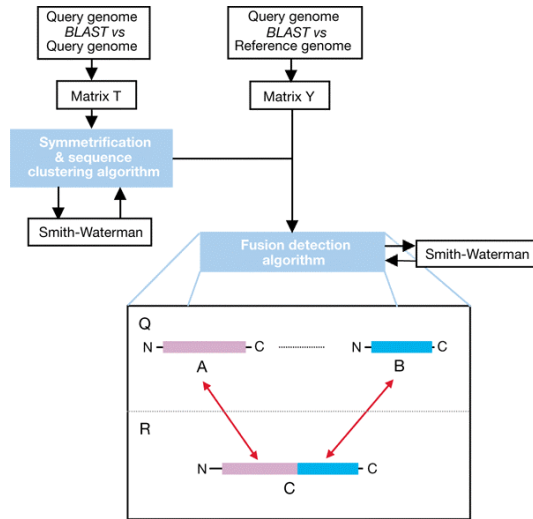
Conservación de Cercanía Genómica



Overbeek, R., Fonstein, M., D'Souza, M., Pusch, G. D. & Maltsev, N. (1999). Use of contiguity on the chromosome to predict functional coupling. *In Silico Biol.* **1**, 93-108.

Dandekar, T., Snel, B., Huynen, M. & Bork, P. (1998). Conservation of gene order: a fingerprint of proteins that physically interact. *Trends Biochem Sci.* **23**, 324-328.

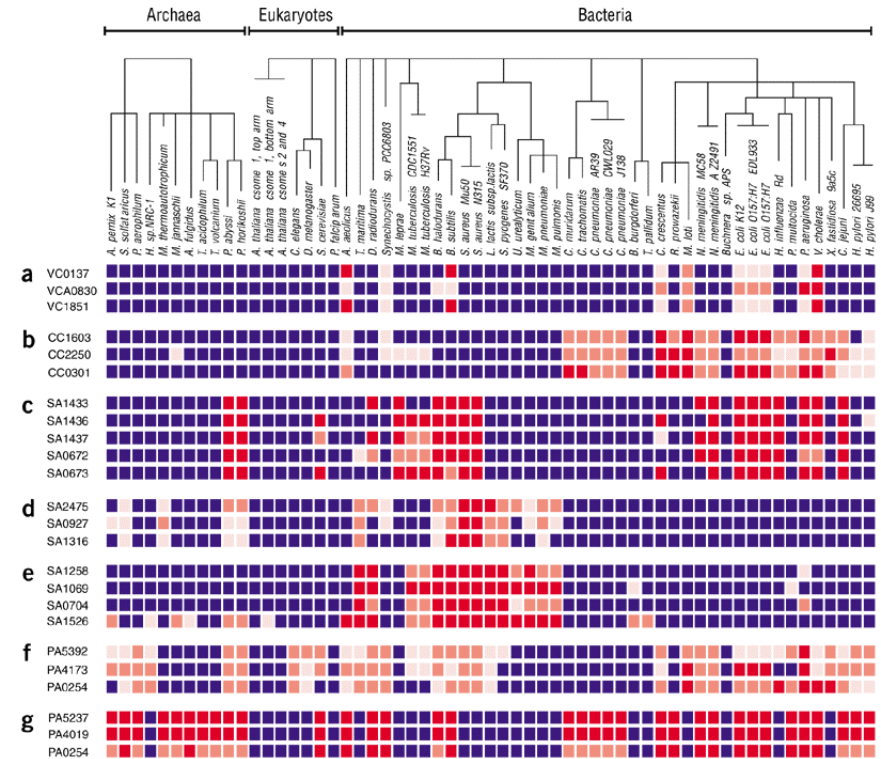
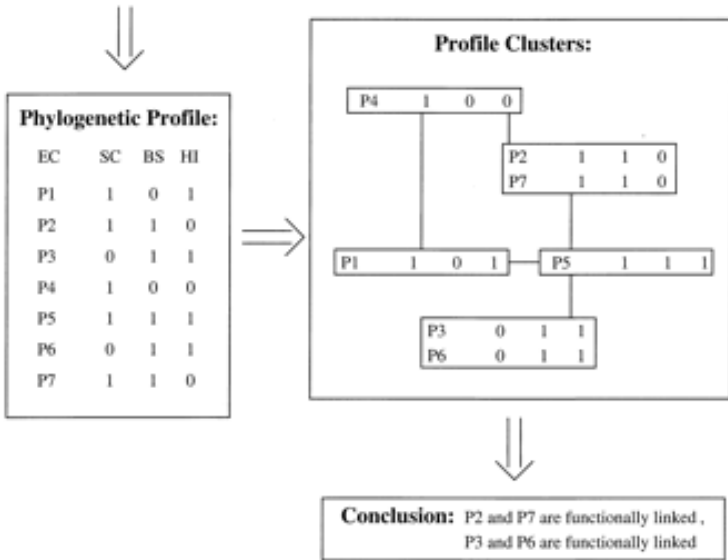
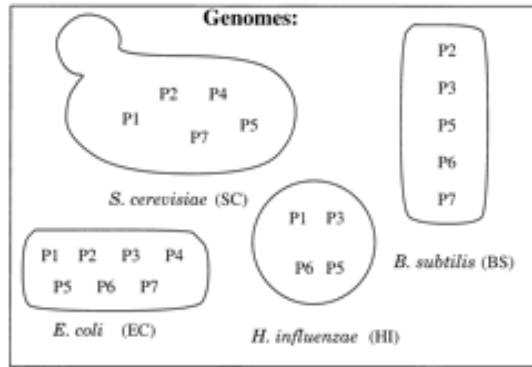
Fusion Génica



Enright, A. J., Iliopoulos, I., Kyripides, N. C. & Ouzounis, C. A. (1999). Protein interaction maps for complete genomes based on gene fusion events. *Nature*. **402**, 86-90.

Marcotte, E. M., Pellegrini, M., Ho-Leung, N., Rice, D. W., Yeates, T. O. & Eisenberg, D. (1999). Detecting protein function and protein-protein interactions from genome sequences. *Science*. **285**, 751-753.

Perfiles Filogenéticos



$$p_{ij} = -1/\log E_{ij}$$

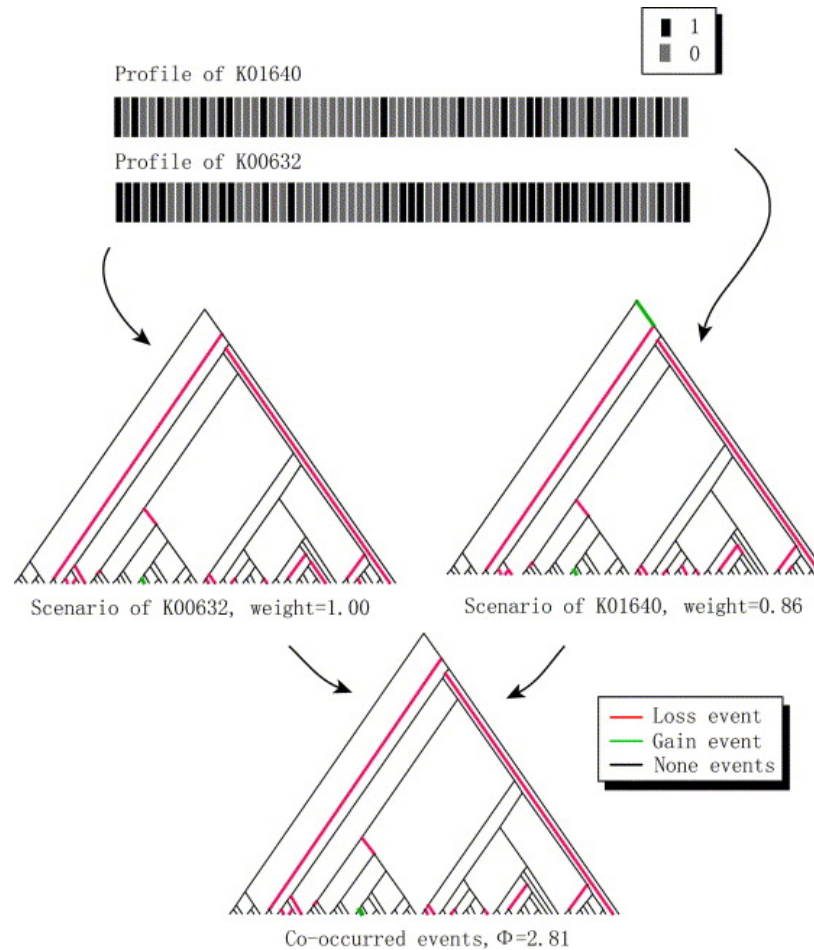
$$H(A) = - \sum p(a) \ln p(a)$$

$$MI(A,B) = H(A) + H(B) - H(A,B)$$

•Pellegrini, M., Marcotte, E. M., Thompson, M. J., Eisenberg, D. & Yeates, T. O. (1999). Assigning protein functions by comparative genome analysis: Protein pylogenetic profiles. *Proc Natl Acad Sci USA*. **96**, 4285-4288.

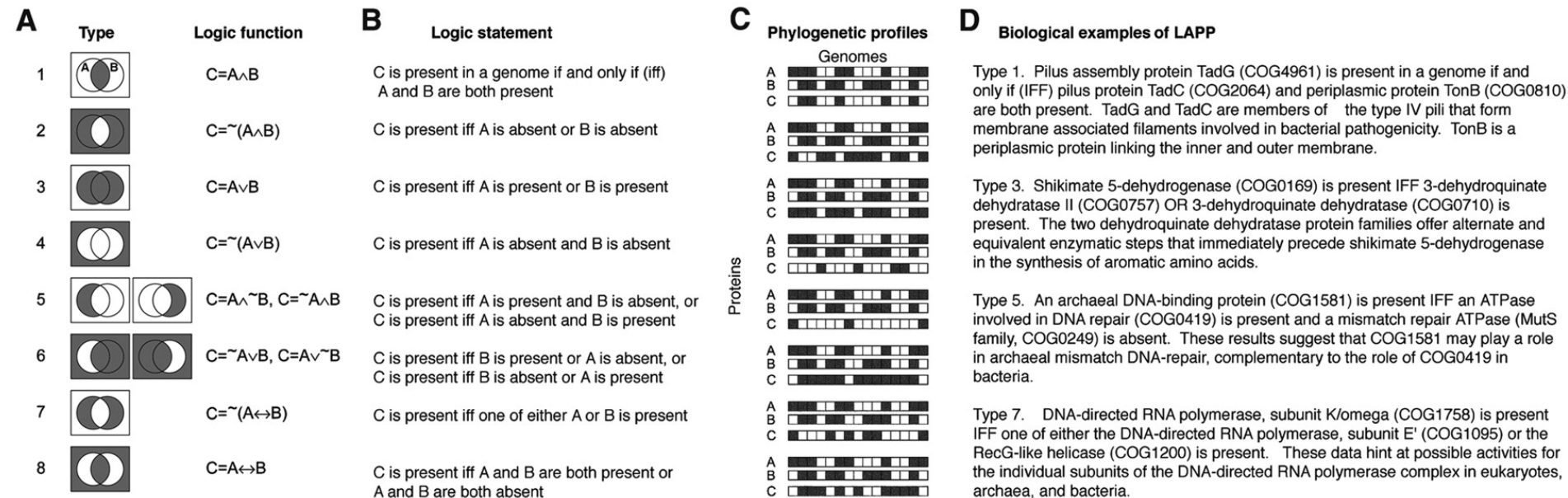
•Date, S. V. & Marcotte, E. M. (2003). Discovery of uncharacterized cellular systems by genome-wide analysis of functional linkages. *Nat Biotechnol*. **21**, 1055-1062.

Perfiles Filogenéticos

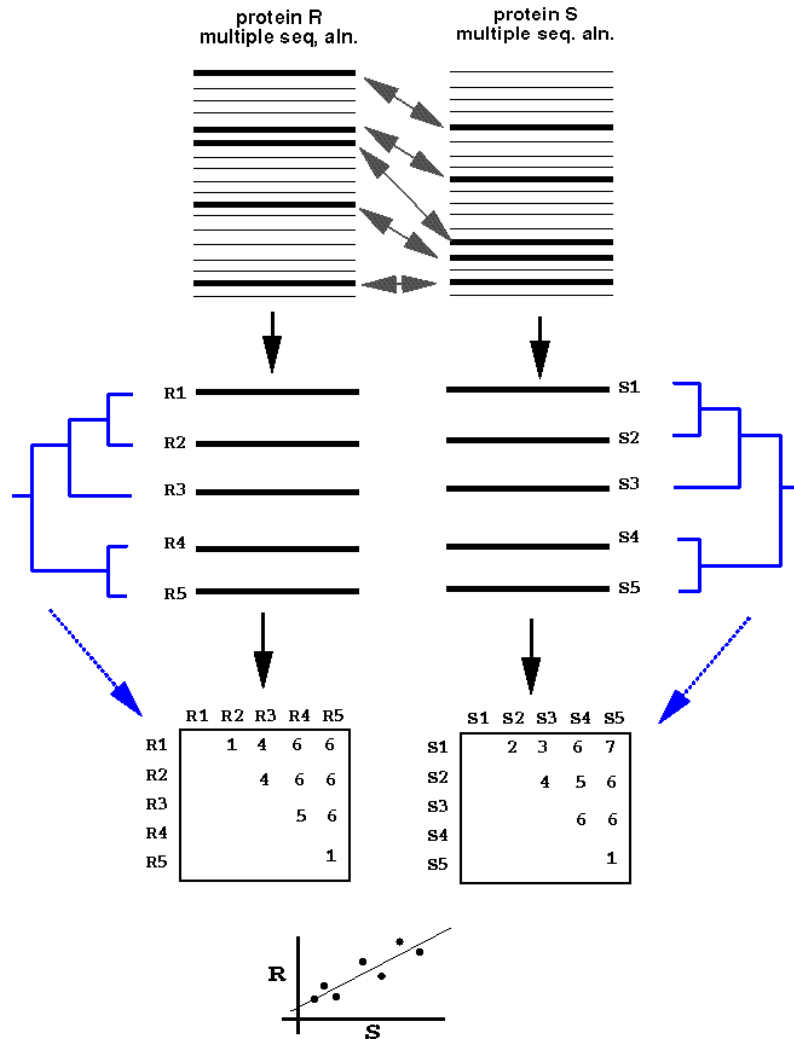


-
- Zhou, Y., Wang, R., Li, L., Xia, X. and Sun, Z. (2006) Inferring functional linkages between proteins from evolutionary scenarios. *J Mol Biol.*, **359**, 1150-1159.
 - Barker, D., Meade, A. and Pagel, M. (2007) Constrained models of evolution lead to improved prediction of functional linkage from correlated gain and loss of genes. *Bioinformatics.*, **23**, 14-20.

Perfiles Filogenéticos



Similitud de Árboles Filogenéticos - *MirrorTree*



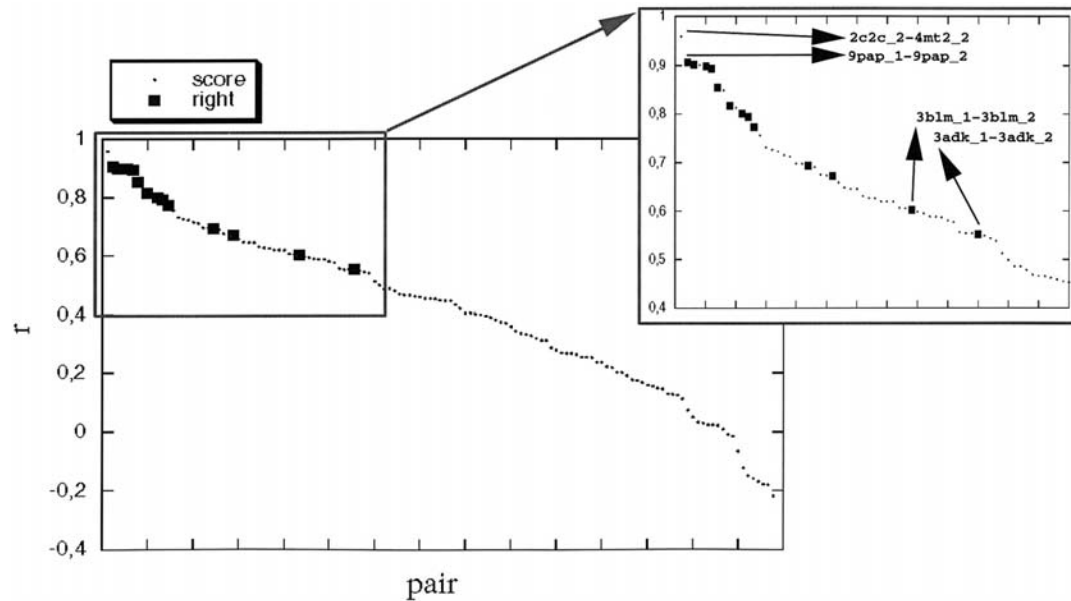
$$r = \frac{\sum_{i=1}^n (R_i - \bar{R}) \cdot (S_i - \bar{S})}{\sqrt{\sum_{i=1}^n (R_i - \bar{R})^2} \cdot \sqrt{\sum_{i=1}^n (S_i - \bar{S})^2}}$$

Goh, C.-S., Bogan, A.A., Joachimiak, M., Walther, D. and Cohen, F.E. (2000) Co-evolution of Proteins with their Interaction Partners. *J Mol Biol*, **299**, 283-293.

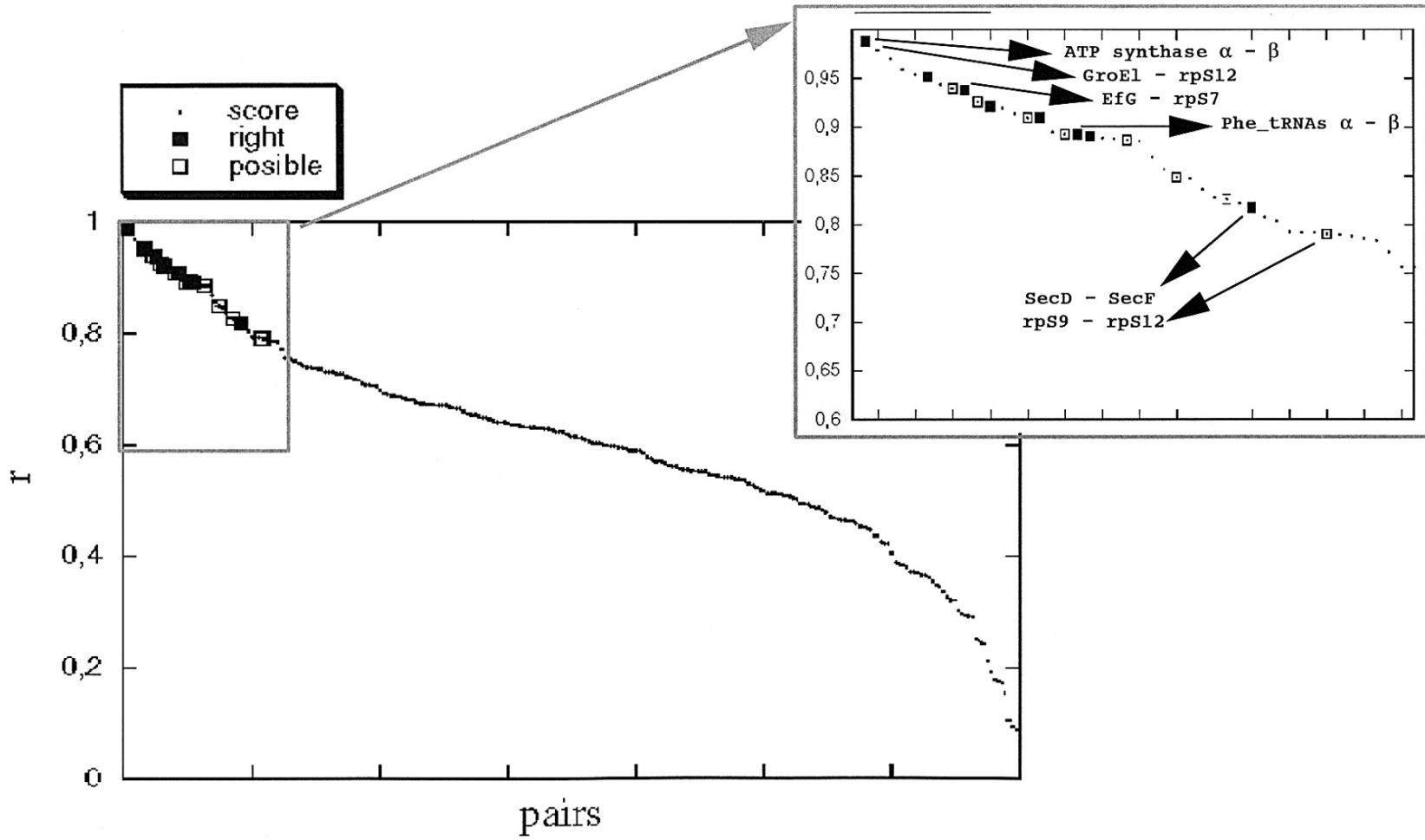
Pazos, F. and Valencia, A. (2001) Similarity of phylogenetic trees as indicator of protein-protein interaction. *Protein Eng*, **14**, 609-614.

MirrorTree

2c2c_2-4mt2_2	0,959	4tnc_1-4mt2_2	0,448
9pap_1-9pap_2 *	0,907	9pap_2-4tnc_1	0,446
2c2c_1-4mt2_2	0,901	1alc_2-4mt2_2	0,446
3pgk_1-3pgk_2 *	0,901	1sgt_2-4mt2_1	0,433
4mt2_1-4mt2_2 *	0,898	3adk_1-4tnc_2	0,421
3trx_1-3trx_2 *	0,894	1rnd_1-4mt2_2	0,405
4tms_1-4tms_2 *	0,854	4tnc_2-4mt2_2	0,405
2c2c_2-4mt2_1	0,849	2c2c_1-3adk_2	0,401
1rnd_1-1rnd_2 *	0,817	1sgt_2-2c2c_1	0,399
2c2c_1-4mt2_1	0,813	4tms_2-3dfr_2	0,394
1alc_1-1alc_2 *	0,801	3adk_1-3dfr_1	0,390
4tnc_1-4tnc_2 *	0,794	1sgt_2-2c2c_2	0,381
2c2c_1-2c2c_2 *	0,773	3adk_2-3dfr_1	0,372
3pgk_1-4tms_1	0,756	1sgt_2-1alc_1	0,371
3pgk_1-4tms_2	0,731	4tms_1-3dfr_2	0,358
2c2c_1-3adk_1	0,726	1sgt_1-4mt2_1	0,343
3pgk_2-4tms_1	0,723	1sgt_1-4mt2_2	0,336
2c2c_2-3pgk_1	0,715	9pap_2-4tnc_2	0,331
1alc_1-1rnd_1	0,712	4tms_1-3dfr_1	0,327
2c2c_2-3pgk_2	0,698	3trx_1-2c2c_2	0,319
1alc_2-1rnd_1	0,697	3trx_1-2c2c_1	0,312
1sgt_1-1sgt_2 *	0,693	1sgt_1-1alc_1	0,312
3pgk_2-4tms_2	0,691	3trx_2-2c2c_2	0,287
3adk_2-3dfr_2	0,675	3trx_2-2c2c_1	0,281
1sgt_2-2pf2_2	0,673	1sgt_1-2c2c_2	0,270
3dfr_1-3dfr_2 *	0,672	1alc_1-4mt2_2	0,268
2c2c_2-9pap_1	0,658	1sgt_1-2c2c_1	0,268
2c2c_1-3pgk_1	0,648	2c2c_1-1rnd_1	0,263
3trx_2-9pap_1	0,646	9pap_1-3adk_2	0,254
1sgt_1-2pf2_2	0,646	2c2c_2-3adk_2	0,254
2c2c_2-3adk_1	0,631	3adk_2-3pgk_1	0,251
3trx_1-9pap_1	0,627	1sgt_1-1rnd_1	0,238
2c2c_2-1alc_2	0,626	3adk_2-3pgk_2	0,238
2c2c_1-3pgk_2	0,620	9pap_2-3adk_2	0,221
3trx_2-9pap_2	0,620	1sgt_2-1alc_2	0,219
1rnd_2-4mt2_1	0,619	2c2c_2-1alc_1	0,203
1alc_2-1rnd_2	0,607	9pap_1-4tnc_1	0,202
1rnd_2-4mt2_2	0,606	1sgt_2-1rnd_1	0,191
3blm_1-3blm_2 *	0,603	1sgt_1-1alc_2	0,178
1alc_1-1rnd_2	0,599	3trx_2-3adk_2	0,175
3trx_1-3pgk_1	0,595	1sgt_1-1rnd_2	0,168
3trx_1-9pap_2	0,589	2pf2_2-1alc_1	0,160
1alc_2-4mt2_1	0,588	2c2c_1-1alc_1	0,155
2c2c_1-1alc_2	0,587	9pap_1-4tnc_2	0,149
2c2c_1-9pap_1	0,581	2c2c_2-1rnd_2	0,146
3trx_1-3pgk_2	0,577	4tms_2-3dfr_1	0,130
4tnc_1-4mt2_1	0,556	3trx_1-3adk_2	0,128
3adk_1-3pgk_1	0,554	2c2c_2-1rnd_1	0,125
		2c2c_1-1rnd_2	0,113



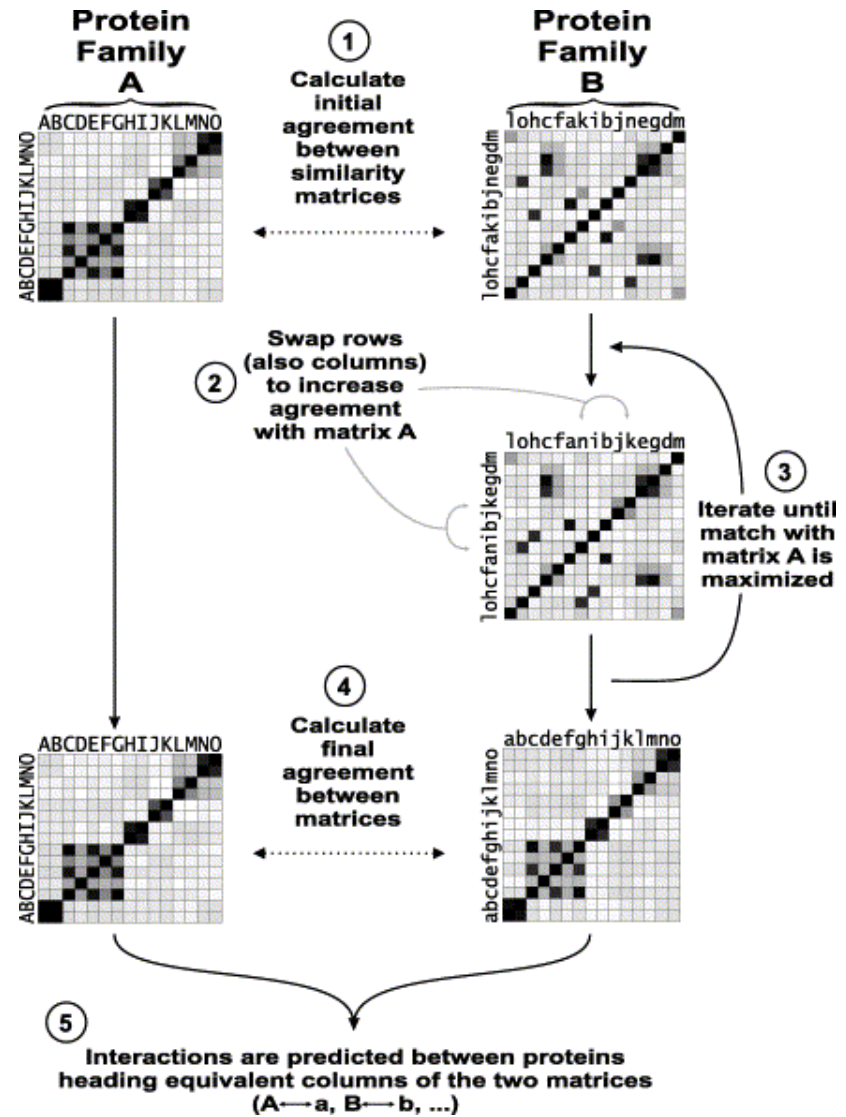
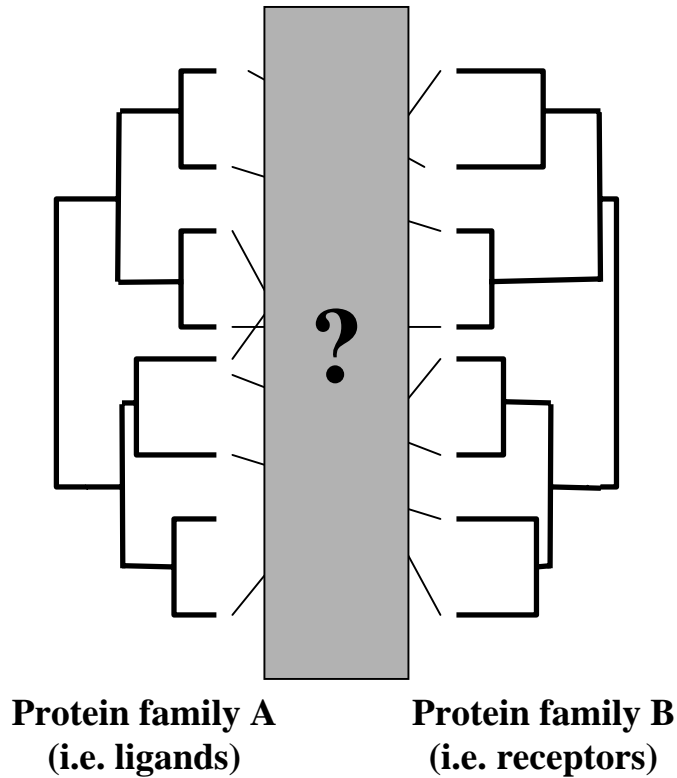
MirrorTree



MirrorTree - Variaciones

- Gertz, J., Elfond, G., Shustrova, A., Weisinger, M., Pellegrini, M., Cokus, S. and Rothschild, B. (2003) Inferring protein interactions from phylogenetic distance matrices. *Bioinformatics*, **19**, 2039-2045.
- Goh, C.S. and Cohen, F.E. (2002) Co-evolutionary analysis reveals insights into protein-protein interactions. *J Mol Biol*, **324**, 177-192.
- Ramani, A.K. and Marcotte, E.M. (2003) Exploiting the co-evolution of interacting proteins to discover interaction specificity. *J Mol Biol*, **327**, 273-284.
- Sato, T., Yamanishi, Y., Horimoto, K., Toh, H. and Kanehisa, M. (2003) Prediction of protein-protein interactions from phylogenetic trees using partial correlation coefficient. *Genome Informatics*, **14**, 496-497.
- Kim, W.K., Bolser, D.M. and Park, J.H. (2004) Large-scale co-evolution analysis of protein structural interlogues using the global protein structural interactome map (PSIMAP). *Bioinformatics*, **20**, 1138-1150. Epub 2004 Feb 1135.
- Tan, S., Zhang, Z. and Ng, S. (2004) ADVICE: Automated Detection and Validation of Interaction by Co-Evolution. *Nucl. Acids. Res.*, **32**, W69-W72.
- Jothi, R., Kann, M.G. and Przytycka, T.M. (2005) Predicting protein-protein interaction by searching evolutionary tree automorphism space. *Bioinformatics*, **21**, i241-i250.
- Mintseris, J. and Weng, Z. (2005) Structure, function, and evolution of transient and obligate protein-protein interactions. *Proc Natl Acad Sci U S A*, **102**, 10930-10935.
- Sato, T., Yamanishi, Y., Kanehisa, M. and Toh, H. (2005) The inference of protein-protein interactions by co-evolutionary analysis is improved by excluding the information about the phylogenetic relationships. *Bioinformatics*, **21**, 3482-3489.
- Tillier, E.R., Biro, L., Li, G. and Tillo, D. (2006) Codep: maximizing co-evolutionary interdependencies to discover interacting proteins. *Proteins.*, **63**, 822-831.
- Jothi, R., Cherukuri, P.F., Tasneem, A. and Przytycka, T.M. (2006) Co-evolutionary Analysis of Domains in Interacting Proteins Reveals Insights into Domain-Domain Interactions Mediating Protein-Protein Interactions. *J Mol Biol.*, **362**, 861-875.
- Tan, S., Zhang, Z. and Ng, S. (2004) ADVICE: Automated Detection and Validation of Interaction by Co-Evolution. *Nucleic Acids Res.*, **32**, W69-W72.

MirrorTree. Variaciones

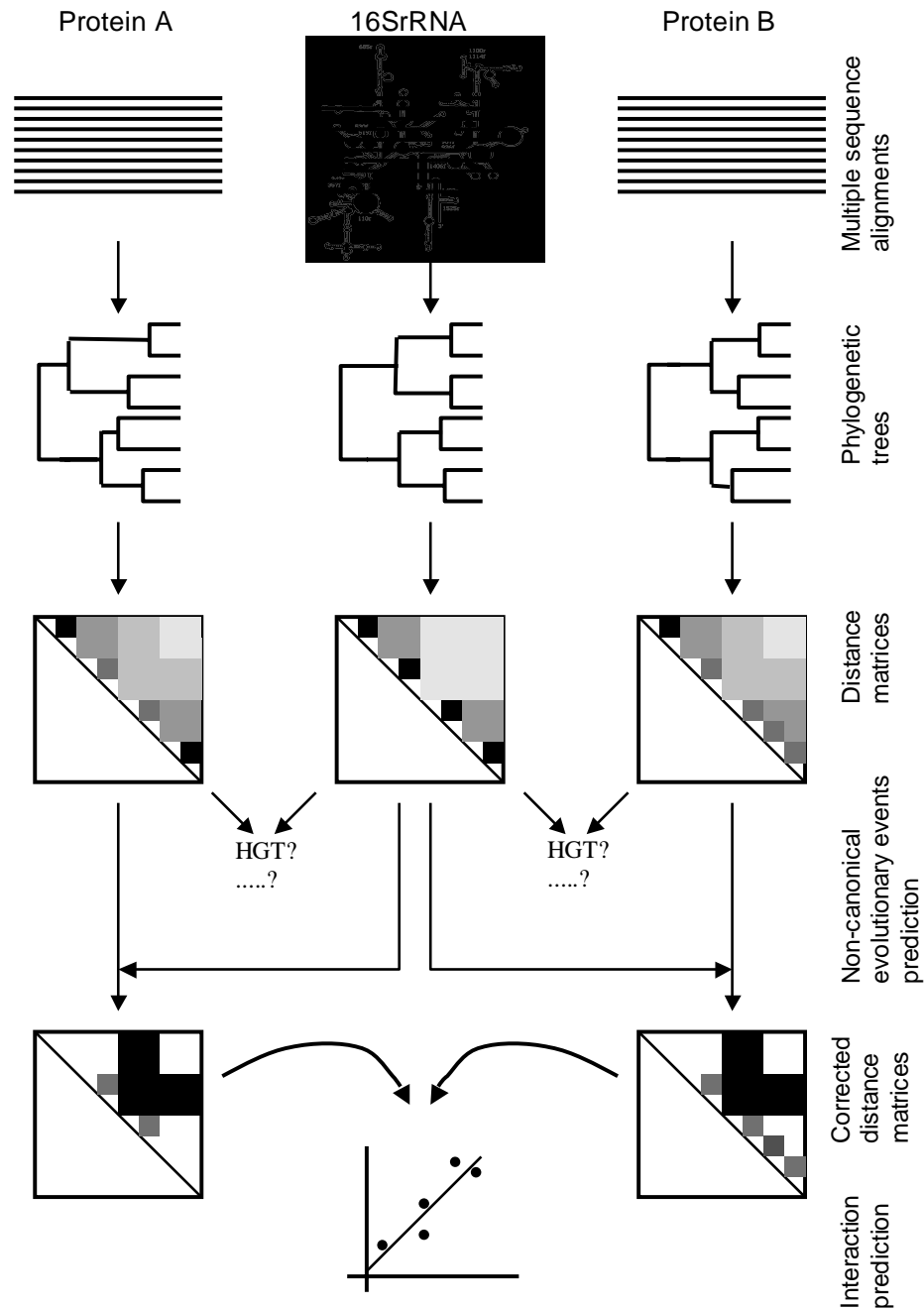


• Ramani, A.K. & Marcotte, E.M. (2003) Exploiting the co-evolution of interacting proteins to discover interaction specificity. *J Mol Biol*, **327**, 273-284.

• Tillier, E.R., Biro, L., Li, G. and Tillo, D. (2006) Codep: maximizing co-evolutionary interdependencies to discover interacting proteins. *Proteins.*, **63**, 822-831.

MirrorTree

tol-mirrortree



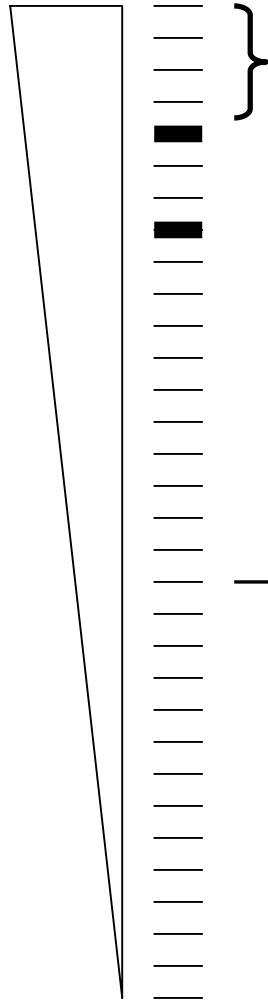
tol-mirrortree

DIP:
516 interactions (*E coli*)

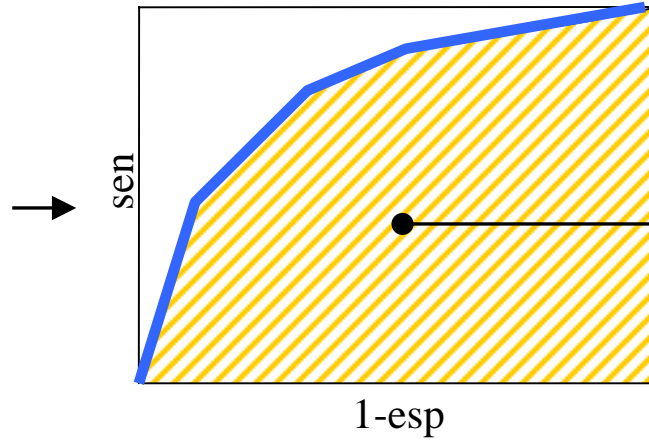
20,087 pairs calculated
(115 true)

118 proteins with
 ≥ 1 calculated true interactor

P00000
List of pairs
sorted by score

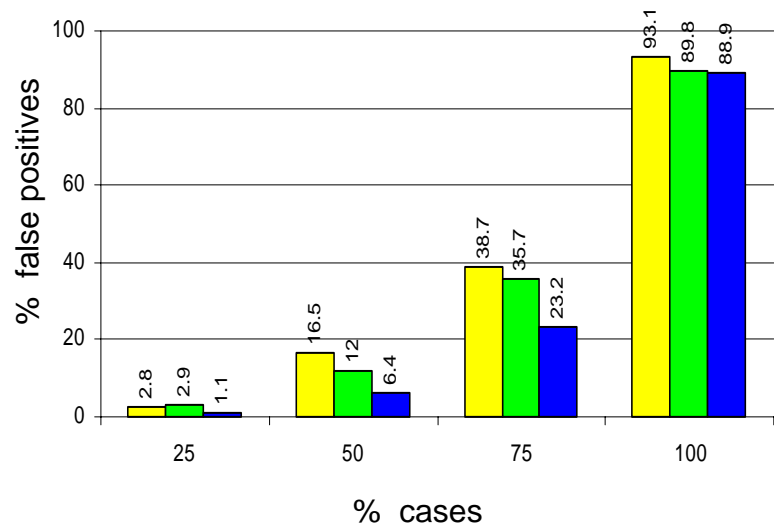
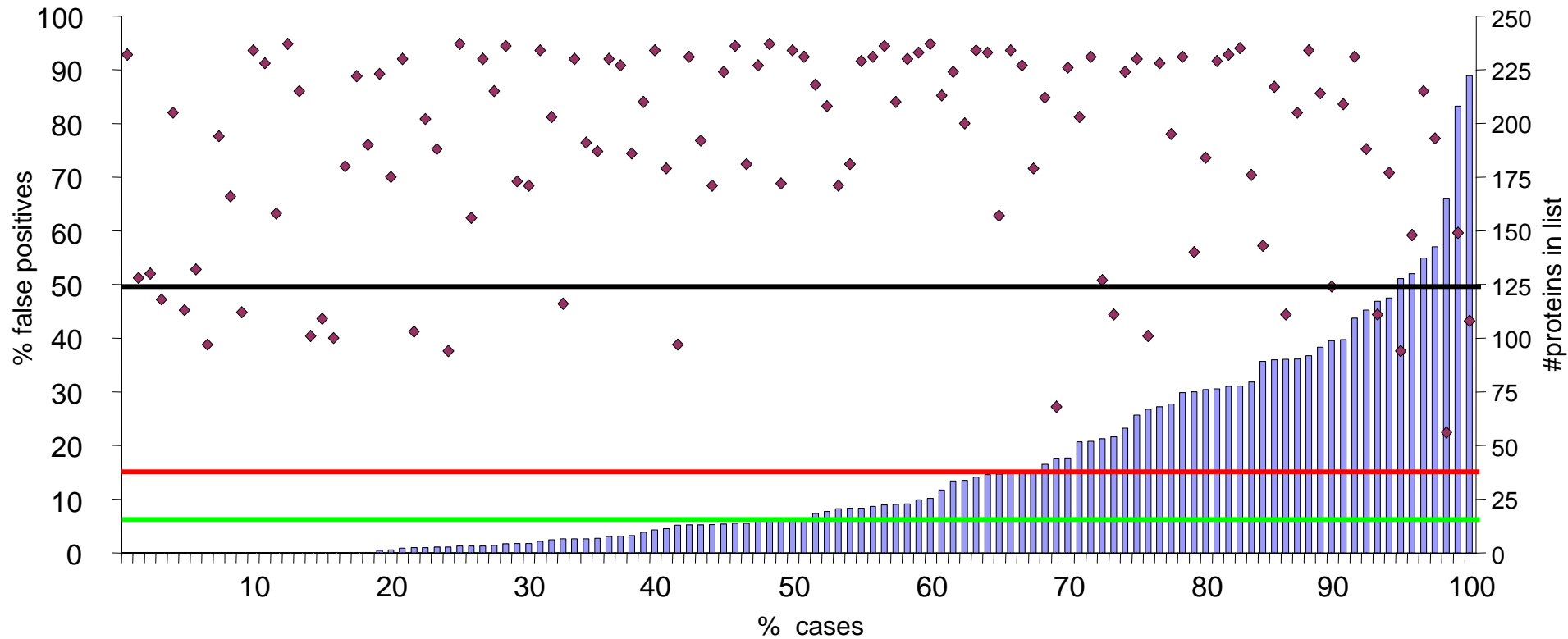


fraction of false positives
0%: perfect
50%: random
(1 int.)



ROC area
1.0: perfect
0.5: random

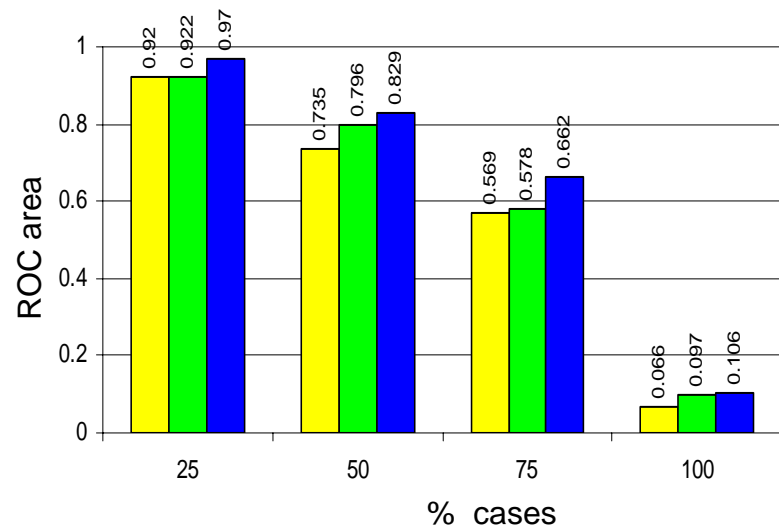
tol-mirrortree



mirrotree

mirrortree
 (tree dist.)

tol-mirrortree



Comparación con versiones anteriores de *mirrortree*

	% false positives	ROC area
<i>mirrortree</i>	23.4	0.71
<i>mirrortree</i> with tree distances	21.9	0.73
<i>tol-mirrortree</i>	14.9	0.79

P(N) values (sign test):

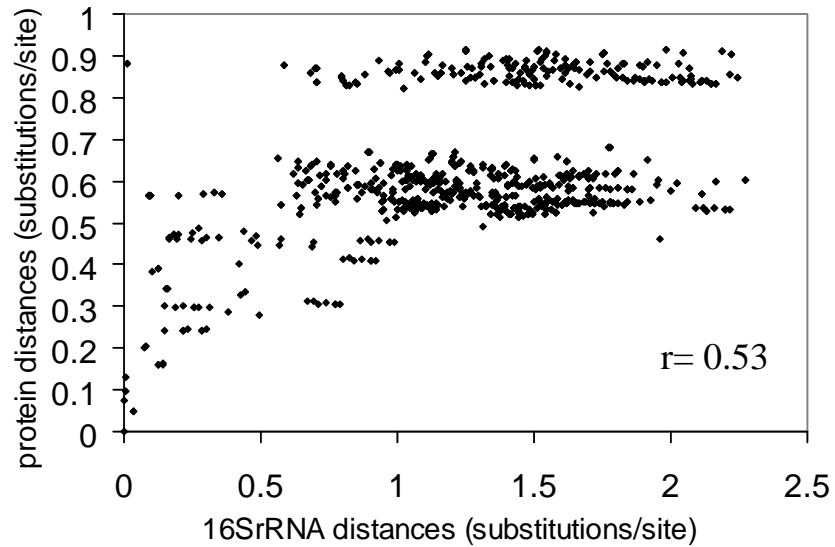
a \ b	<i>mirrortree</i>	<i>mirrortree</i> tree dist.	<i>tol-mirrortree</i>
<i>mirrortree</i>			
<i>mirrortree</i> tree dist.	0.276		
<i>tol-mirrortree</i>	$5.60 \cdot 10^{-6}$	$1.91 \cdot 10^{-5}$	

tol-mirrortree

Detección concomitante de eventos evolutivos no-estándar

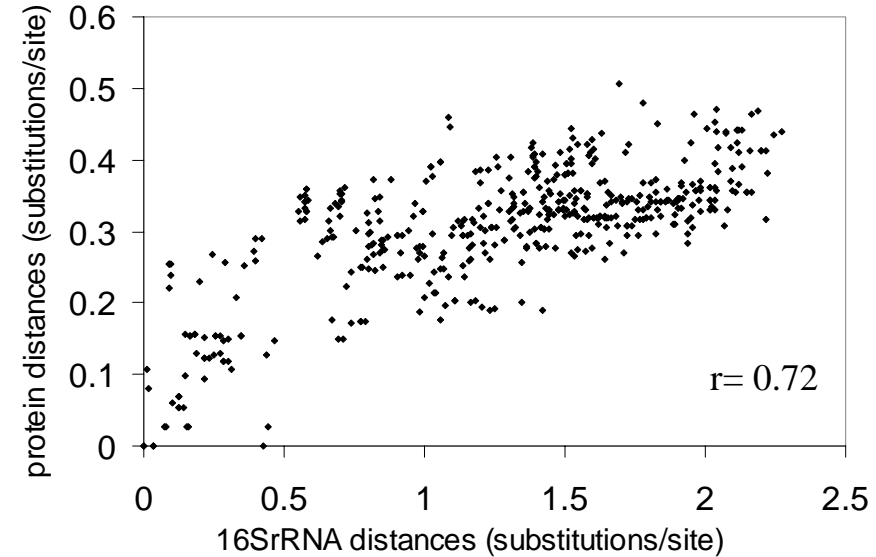
a)

Prolyl-tRNA synthetase

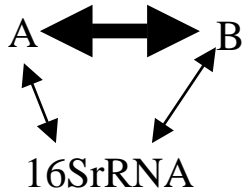


b)

Ribosomal protein L36

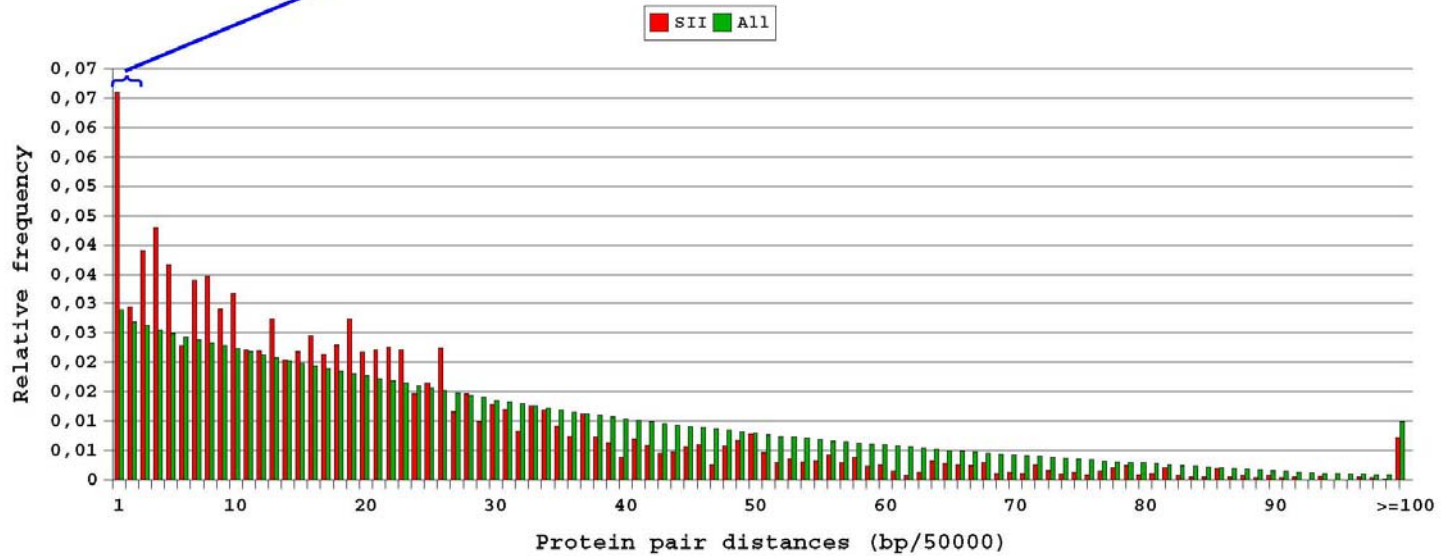
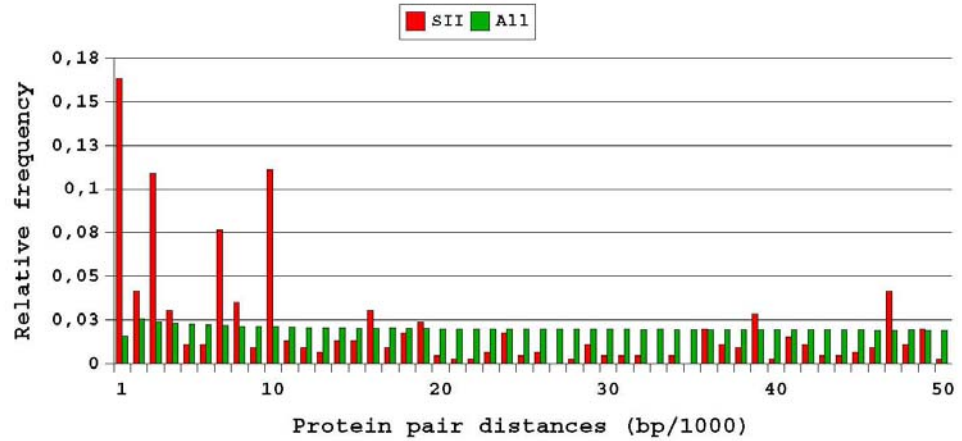


$r \leq 0.5 \rightarrow 25\%$ false pos (vs. 15%)
Excluding them: 13.7% false pos (vs. 15%)



Co-HGT events

SII vs All *E. coli* protein orthologous pair distances

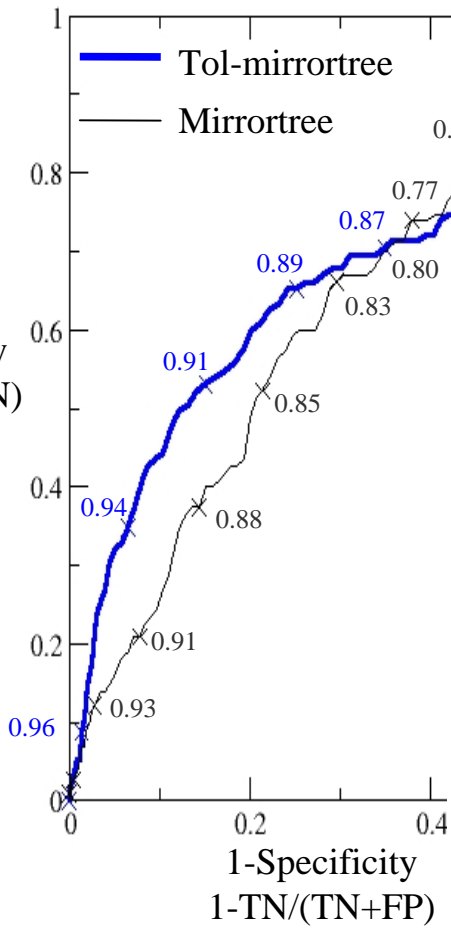


Protein
Multiple Sequences
Organism 1
Organism 2
Organism 3
Organism 4
Organism 5

Protein
Multiple Sequences
Organism 1
Organism 2
Organism 3
Organism 4
Organism 5

Protein Sensitivity
Multiple Sequences
 $TP / (TP + FN)$
Organism 1
Organism 2
Organism 3
Organism 4
Organism 5

Protein
Multiple Sequences
Organism 1
Organism 2
Organism 3
Organism 4
Organism 5



Correlation Coefficient
(ρ^{16S})

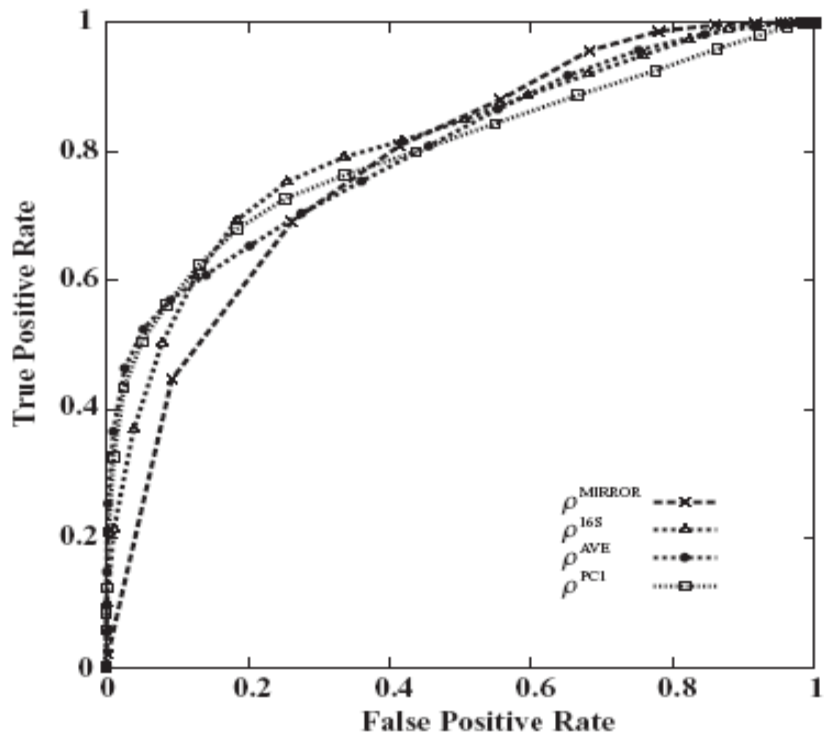
- Protein A - Protein B
- Protein A - Protein C
- Protein A - Protein D
- Protein B - Protein C
- Protein B - Protein D
- Protein C - Protein D

Correlation Coefficient
(ρ^{AVE})

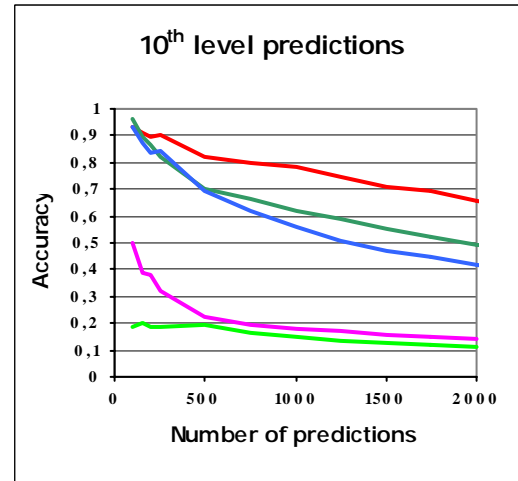
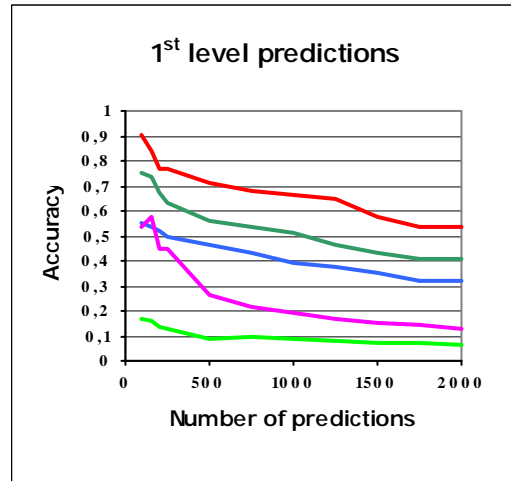
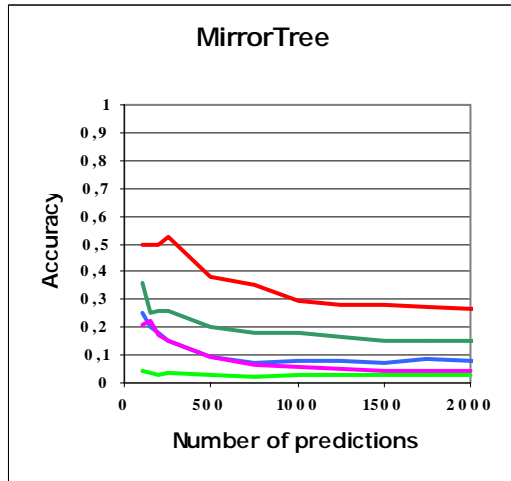
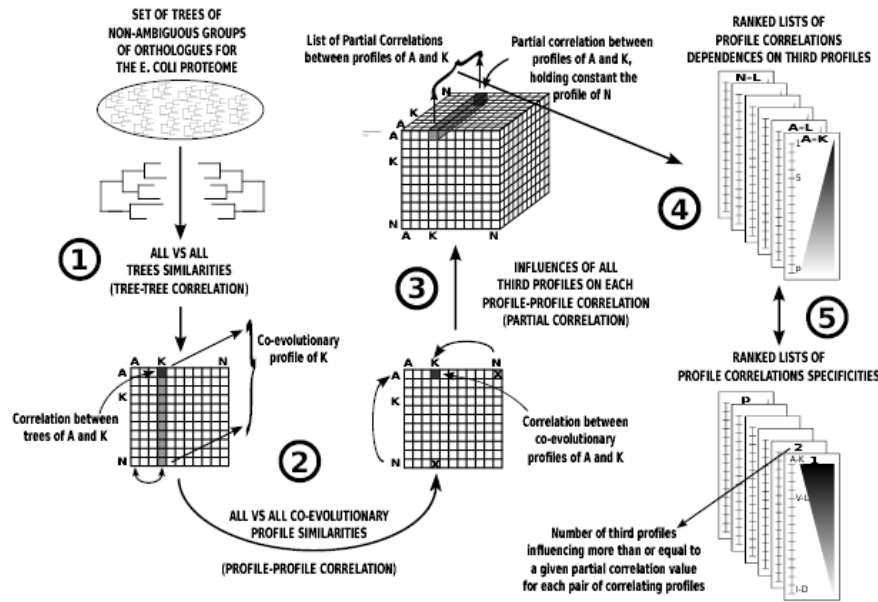
- Protein B
- Protein C
- Protein D
- Protein C
- Protein D
- Protein D

Correlation Coefficient
(ρ^{PC1})

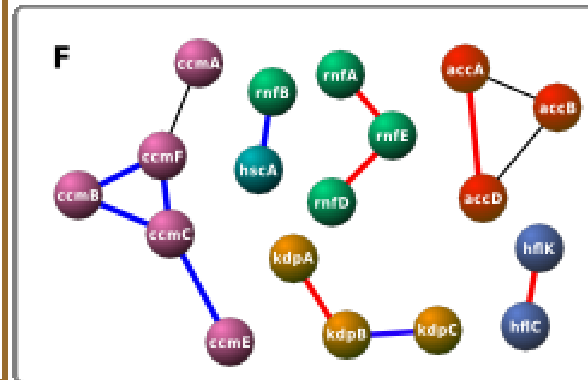
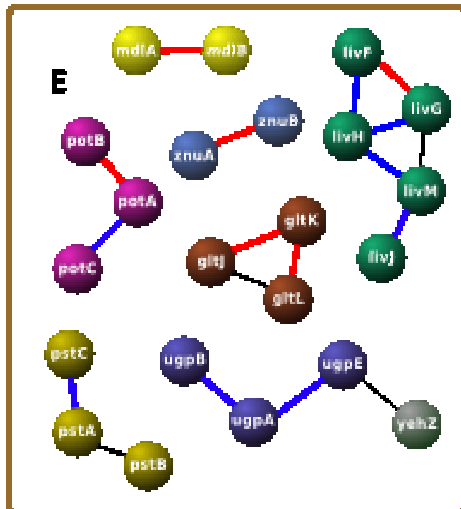
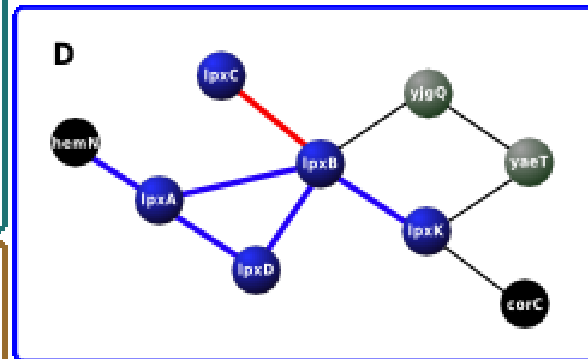
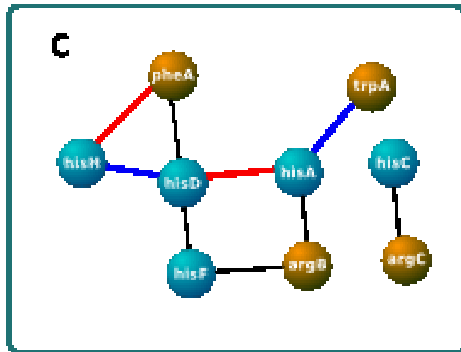
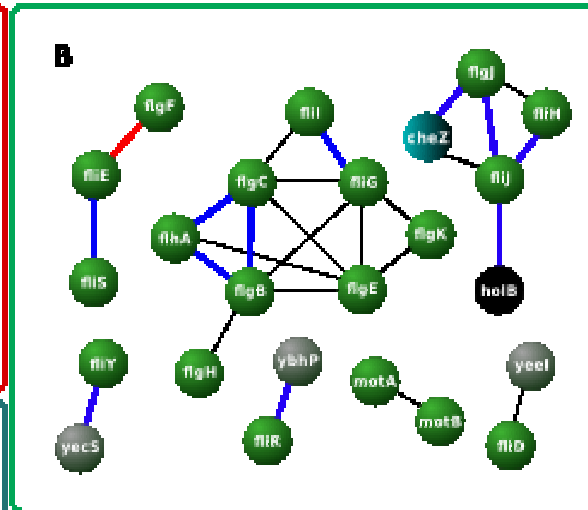
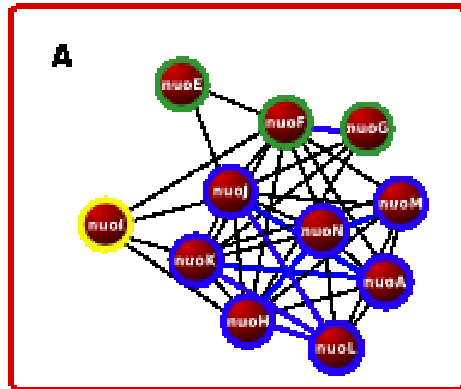
- Protein B
- Protein C
- Protein D
- Protein C
- Protein D
- Protein D



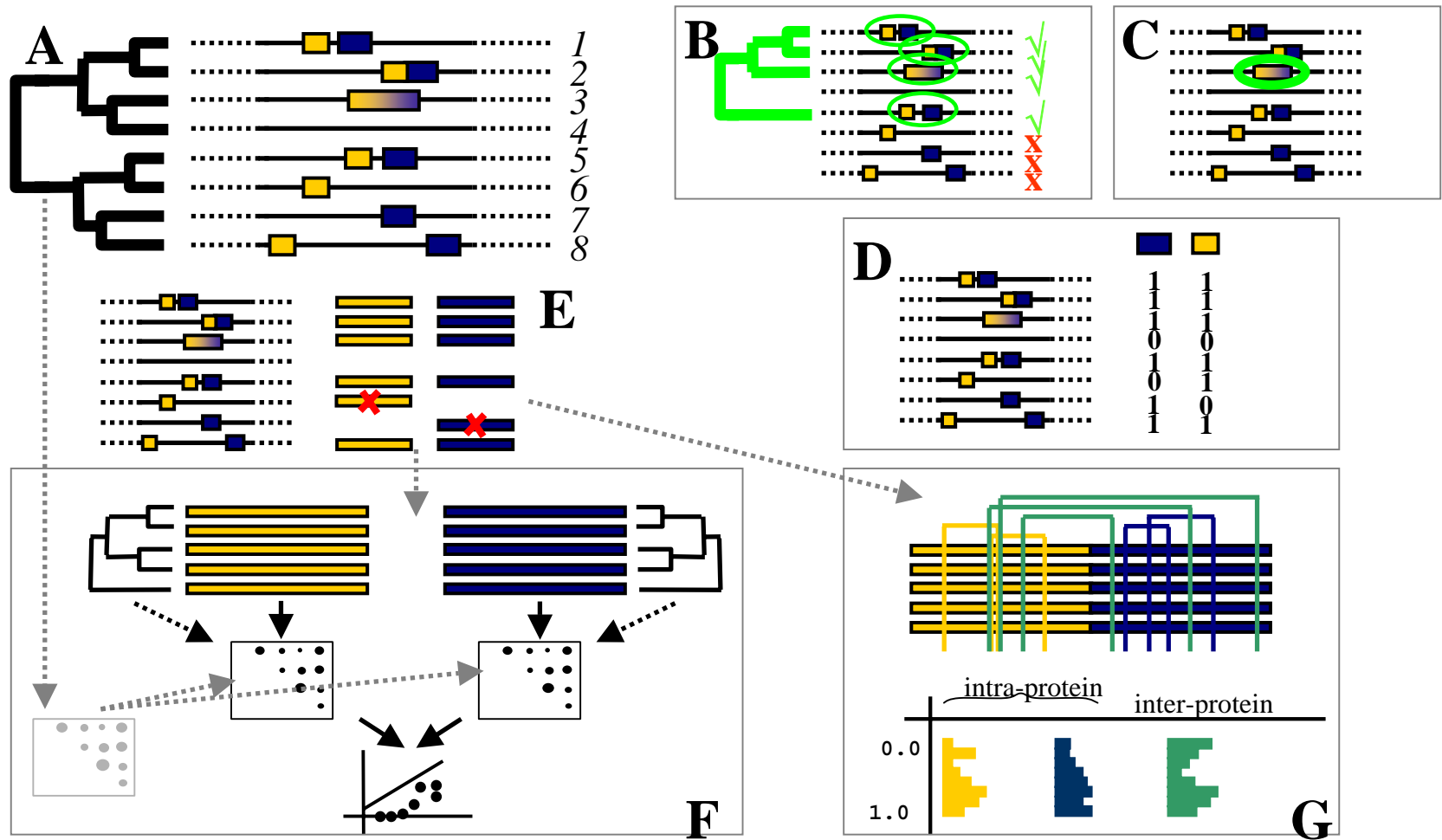
Mirrortree – Usando la Información del Contexto Co-evolutivo



Ejemplos



Métodos Computacionales de Predicción de Compañeros de Interacción



• Alfonso Valencia & Florencio Pazos (2002). Prediction of Protein Interactions with Computational Methods. *Curr Op Str Biol.* **12(3)**: 368-373. [56/67]

• Alfonso Valencia & Florencio Pazos. (2003). Prediction of protein-protein interactions from evolutionary information. *Methods Biochem Anal.* **44**:411-426.

• Florencio Pazos & Alfonso Valencia (2006). Protein Interactions from an Evolutionary Perspective. In "Evolution of Biological Networks". Carsten Wiuf & Michael Stumpf (Eds). Imperial College Press/World Scientific. *In Press*.

Repositorios *on-line* de interacciones predichas

String: functional protein association networks - Mozilla

File Edit View Go Bookmarks Tools Window Help

Back Forward Reload Stop http://string.embl.de/ Search Print

Home Bookmarks SUSE LINUX The Mozilla Org...

Home · Download · Help/Info

STRING

STRING - Search Tool for the Retrieval of Interacting Gen

Enter your gene/protein of interest ...

identifier: e.g. 'trpB', 'ANP1_YEAST', ...
you may also upload a [list](#)

alternatively, paste an amino-acid sequence:

interactors wanted:

What it does ...

STRING is a database of kn protein-protein interactions. The interactions include dir (functional) associations; th sources:

Genomic Context High-throughput Experiments

STRING quantitatively integ these sources for a large n transfers information betw applicable. The database c proteins in 179 species.

References / Info ...

STRING uses orthology information from the excellent COG database (Ref). Up-to-date genomes and proteins are maintained at SWISSPROT and ENSEMBL. STRING references: von Mering et.al. 2005 / Snel et.al. 2000. Miscellaneous: Access Statistics, Robot Access Guide, Supported Browsers.

Home · Download · Help/Info

STRING

conserved genomic neighborhood

Corynebacterium (2 species)
Bifidobacterium longum
Leptospira interrogans
Borrelia burgdorferi
Treponema pallidum

phylogenetic co-occurrence

Cyanobacteria (3 species)
Thermotoga maritima
Deinococcus radiodurans
Fusobacterium nucleatum
Haemifys amelonum
Chlorobium tepidum
Crenarchaeota (4 species)

literature co-occurrence

... inactivation of the epsilon subunit does not alleviate th atp1 (e-1, atp2 (e-2, atp3 (e-3) and atp4 (e-4) mutants. It is suggested that assembly or stability of F₁ in the wild-type, atp2 (e-1) and atp2 (e-2) is suppressed by the atp1 (e-1), atp2 (e-1) and atp2 (e-2) mutants respectively ...

database imports (knowledge)

annotated protein complex (MPS)
ATPA_YEAST ATPB_YEAST ATPG_YEAST ATPQ_YEAST
ATRA_YEAST UCRA_YEAST ATRA_YEAST ATRQ_YEAST
ATRC_YEAST ATRC_YEAST ATRC_YEAST ATRC_YEAST
[... and 52 other proteins]

high-throughput experiments

Chromatin Immunoprecipitation
ATPA_YEAST UCRA_YEAST ATRA_YEAST ATRQ_YEAST
UCR1_YEAST UCR1_YEAST ATRQ_YEAST ATRQ_YEAST
CY1_YEAST UCR1_YEAST UCR1_YEAST [...] and 135 other proteins]

co-expression analysis

Expression
-1.0 0.0 1.0
CONTRAST
UCR2_YEAST
ATP6_YEAST

Your Input:

- ATPA_YEAST ATP synthase alpha chain, mitochondrial precursor (EC 3.6.3.14) [...]
- UCR2_YEAST Ubiquinol-cytochrome C reductase complex core protein 2, mitoch[...]

Predicted Functional Associations:

Association Type	Gene	Function	Neighborhood	Co-occurrence	Coexpression	Experiments	Database	Textmining	Orthology	Score
Green	ATPB_YEAST	ATP synthase beta chain, mitochondrial precursor (EC 3.6.3.14) [...]	●	●	●	●	●	●	●	0.999
Green	ATPG_YEAST	ATP synthase gamma chain, mitochondrial precursor (EC 3.6.3.14) [...]	●	●	●	●	●	●	●	0.999
Green	UCR1_YEAST	Ubiquinol-cytochrome C reductase complex core protein 1, mitoch[...]	●	●	●	●	●	●	●	0.999
Green	UCR1_YEAST	Ubiquinol-cytochrome C reductase iron-sulfur subunit, mitochon[...]	●	●	●	●	●	●	●	0.995
Green	ATPQ_YEAST	ATP synthase oligomycin sensitivity conferral protein, mitochon[...]	●	●	●	●	●	●	●	0.986

Bibliografía

- Alm, E. and Arkin, A.P. (2003) Biological networks. *Curr Opin Struct Biol*, **13**, 193-202.
- Xia, Y., Yu, H., Jansen, R., Seringhaus, M., Baxter, S., Greenbaum, D., Zhao, H. and Gerstein, M. (2004) Analyzing cellular biochemistry in terms of molecular networks. *Annu Rev Biochem*, **73**, 1051-1087.
- Uetz, P. and Finley, R.L., Jr. (2005) From protein networks to biological systems. *FEBS Lett*, **579**, 1821-1827.
- Barabasi, A.L. and Oltvai, Z.N. (2004) Network biology: understanding the cell's functional organization. *Nat Rev Genet*, **5**, 101-113.
- Bork, P., Jensen, L.J., von Mering, C., Ramani, A.K., Lee, I. and Marcotte, E.M. (2004) Protein interaction networks from yeast to human. *Curr Opin Struct Biol*, **14**, 292-299.
- Huynen, M.A., Snel, B., von Mering, C. and Bork, P. (2003) Function prediction and protein networks. *Curr Opin Cell Biol*, **15**, 191-198.
- Valencia, A. & Pazos, F. (2002) Computational methods for the prediction of protein interactions. *Curr Opin Struct Biol*, **12**, 368-373.
- Salwinski, L. & Eisenberg, D. (2003). Computational methods of analysis of protein-protein interactions. *Curr Opin Struct Biol*. **13**, 377-382.