Departamento Biología Molecular
Universidad Autónoma de Madrid

Redes Biológicas y Biología de Sistemas

21-29 Abril 2008

Coordinación: Florencio Pazos (CNB-CSIC)

Redes Metabólicas

lorencio Pazos Cabaleiro
omputational Systems Biology Group (CNB-CSIC)
azos@cnb.csic.es

Redes Biológicas

Protein-protein physical interaction networks. Here nodes represent proteins, and links represent direct physical contacts between proteins. In addition to direct interaction, two proteins can interact indirectly through other proteins when they belong to the same complex.

Protein-protein genetic interaction networks. In general, two genes are said to interact genetically if a mutation in one gene either suppresses or enhances the phenotype of a mutation in its partner gene.

Expression networks. Large-scale microarray experiments probing mRNA expression levels yield vast quantities of data useful for constructing expression networks. In an expression network, genes that are coexpressed are considered connected.

Regulatory networks. Protein-DNA interactions are an important and common class of interactions. Most DNA-binding proteins are transcription factors that regulate the expression of target genes. Combinatorial use of transcription factors further complicates simple interactions of target genes for a given transcription factor. A regulatory network consists of transcription factors and their targets with a specific directionality to the connection between a transcription factor and its target.

Metabolic networks. These networks describe the biochemical reactions within different metabolic pathways in the cell. Nodes represent metabolic substrates and products, and links represent metabolic reactions.

Signaling networks. These networks represent signal transduction pathways through protein-protein and protein-small molecule interactions. Nodes represent proteins or small molecules, and links represent signal transduction events.

Other... Phosforilation netowrk, ...

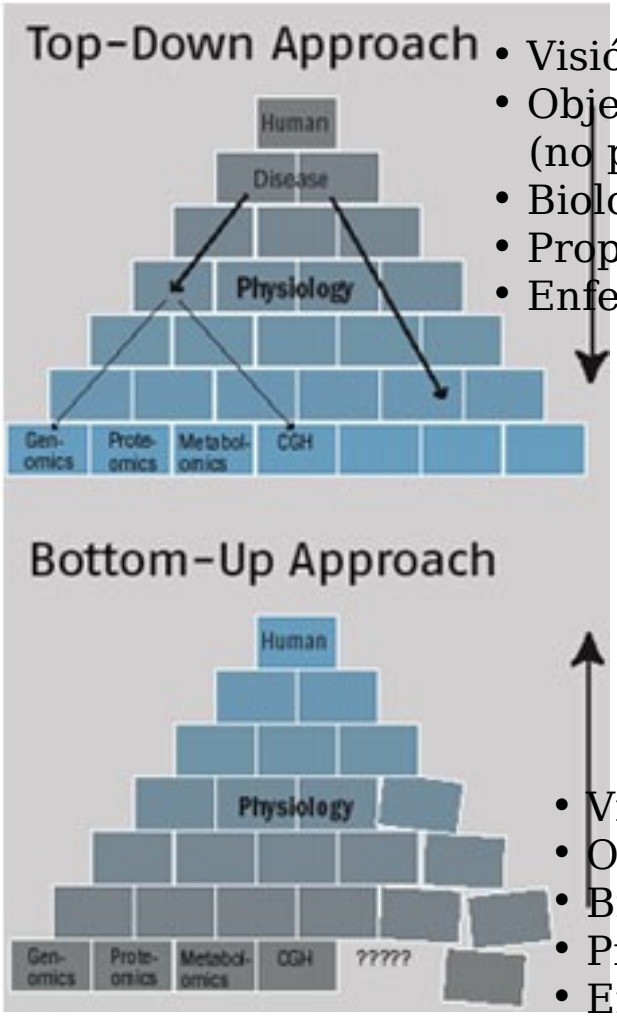
Xia, Y., Yu, H., Jansen, R., Seringhaus, M., Baxter, S., Greenbaum, D., Zhao, H. and Gerstein, M. (2004) Analyzing cellular biochemistry in terms of molecular networks. *Annu Rev Biochem*, **73**, 1051-1087.

Redes Biológicas y Biología de Sistemas

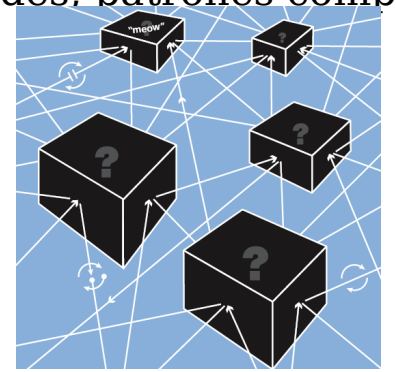
21-29 Abril 2008

L. 21	M. 22	Mi. 23	J. 24	V. 25	S. 26	D. 27
10:00-12:15 Introducción Descripción del curso *F. Pazos (CNB)* **12:30-14:00** Teoría de Grafos (I) *C. Aguirre (UAM)*	**10:00-12:00** Redes de interacciones entre proteínas *D. Juan (CNIO)* **12:00-14:00** Manejo y visualización de redes *D. Juan (CNIO)*	**10:00-12:15** Redes de regulación génica *I. Cases (CNIO)* **12:30-14:00** Teoría de Grafos (II) *C. Aguirre (UAM)*		**10:00-12:00** Bacterias para el medio ambiente: de la Bioremediación a la Biología Sintética *V. Lorenzo (CNB)* **12:30-14:00** Modelado cuantitativo de circuitos biológicos *S. Moreno (CNB)*		
L. 28 **10:00-12:00** Redes metabólicas *F. Pazos (CNB)* **12:00-14:00** Otras redes en Biología Molecular *F. Pazos (CNB)*	**M. 29** **10:00-14:00** Ejercicios prácticos y trabajos de evaluación *F. Pazos (CNB)*	Las clases serán en el Aula de Informática 0 de la Facultad de Ciencias de la UAM. Excepto el primer dia (L.21) que serán en la sala de seminarios "B" del CNB (Hall). **http://pdg.cnb.csic.es/pazos/cursos/bionet_UAM**				

Bioinformática y Sistemas Complejos en Biología
Biología de Sistemas vs. Biología Molecular



- Visión desde el punto de vista de sistemas complejos
- Objeto de estudio: redes, relaciones, propiedades emergentes (no propiedades de componentes individuales).
- Biología de Sistemas
- Propiedades "globales" (emergentes, etc.) -> conocimiento biológico
- Enfermedades (dianas/marcadores = redes, patrones complejos)



?

?

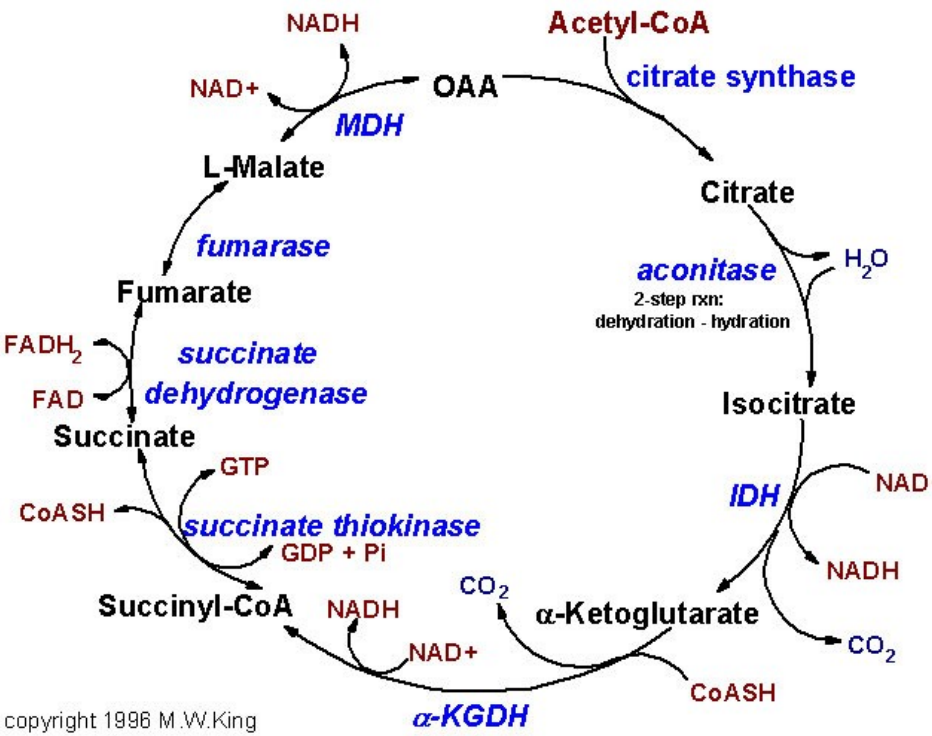
- Visión reduccionista
- Objeto de estudio: componentes (genes proteínas).
- Biología Molecular
- Propiedades de los componentes -> conocimiento biológico
- Enfermedades (dianas/marcadores = proteínas/genes).

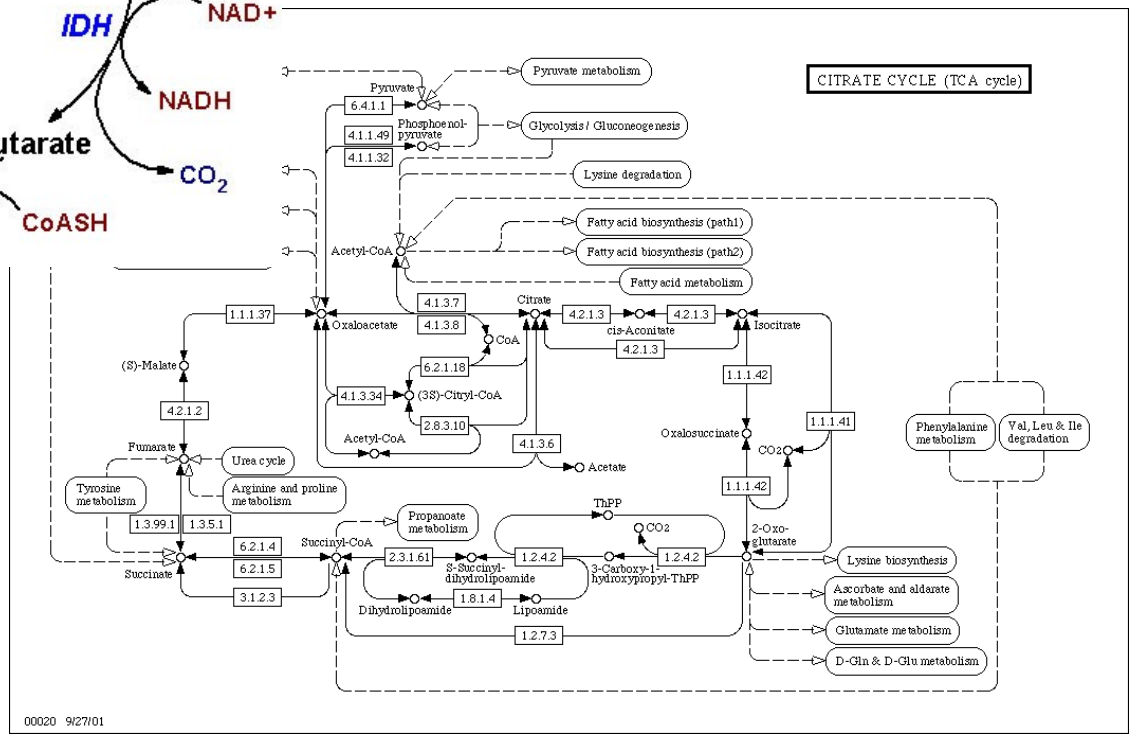
Michael Liebman (www.bioitworld.com)

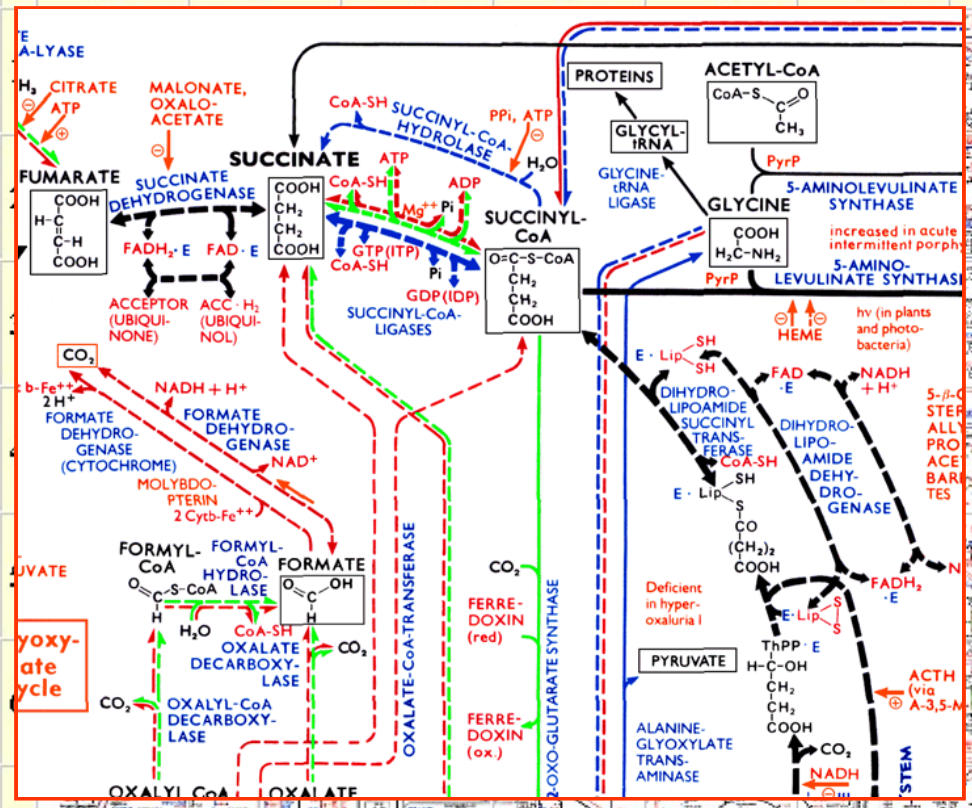
“genomics” / “post-genomics”
Multi-level high-throughput characterization of components

Full-genome sequencing (“genome”).
Characterization of transcripts (mRNA) (“transcriptome”)
Characterization of the protein repertory (“proteome”)
Cellular localization of the components (“localizome”)
Genetic regulation networks (“regulome”)
Protein interaction networks (“interactome”)
High throughput characterization of gene-phenotype relationships (“phenome”
Metabolic networks (“metabolome”)
.....

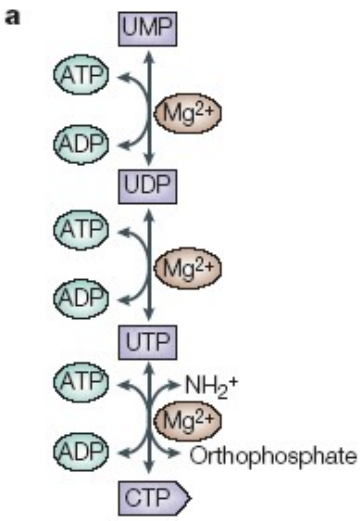
Representación del Metaboloma como Red

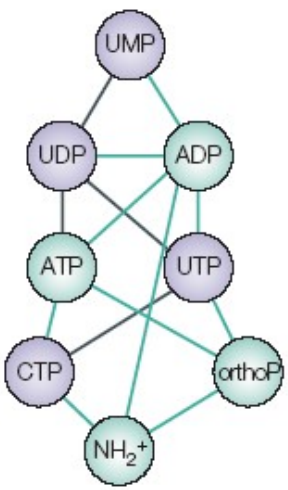


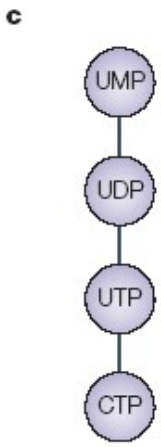



CITRATE
ATP
MALONATE, OXALO-ACETATE
FUMARATE
SUCCINATE DEHYDROGENASE
SUCCINATE
SUCCINYL-CoA-HYDROLASE
PPi, ATP
PROTEINS
ACETYL-CoA
GLYCYL-tRNA
GLYCINE-tRNA LIGASE
GLYCINE
PyrP
5-AMINOLEVULINATE SYNTHASE
SUCCINYL-CoA
5-AMINO-LEVULINATE SYNTHASE
ACCEPTOR (UBIQUI-NONE)
ACC · H2 (UBIQUI-NOL)
SUCCINYL-CoA-LIGASES
GDP(IDP)
HEME
hν (in plants and photo-bacteria)
DIHYDRO-LIPOAMIDE SUCCINYL TRANS-FERASE
DIHYDRO-LIPO-AMIDE DEHY-DRO-GENASE
FORMATE DEHYDRO-GENASE (CYTOCHROME)
MOLYBDO-PTERIN
FORMYL-CoA
FORMYL-CoA HYDRO-LASE
FORMATE
OXALATE DECARBOXY-LASE
OXALYL-CoA DECARBOXY-LASE
OXALATE-CoA-TRANSFERASE
FERRE-DOXIN (red)
FERRE-DOXIN (ox.)
2-OXO-GLUTARATE SYNTHASE
Deficient in hyper-oxaluria I
PYRUVATE
ALANINE-GLYOXYLATE TRANS-AMINASE
ACTH

Conversión de Redes Metabólicas a Grafos

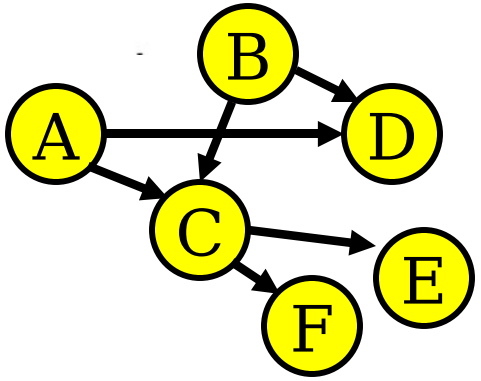






$A + B \rightarrow C + D$

$C \rightarrow E + F$



∀ A, B, C, ... ≠ H20, ions, cofactors, or Nr>X

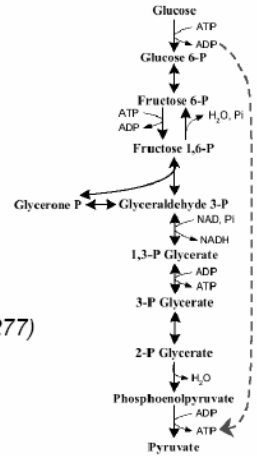
Barabasi, A.L. and Oltvai, Z.N. (2004) Network biology: understanding the cell's functional organization. *Nat Rev Genet*, **5**, 101-113.

Conversión de Redes Metabólicas a Grafos
Exclusión de metabolitos “no informativos”

1) Definición a priori (metabolitos “moneda”)

However, many authors found it necessary to remove the most highly connected metabolites, because they tend to distort the statistics.

(Ma & Zeng 2003, Bioinformatics 17:270-277)



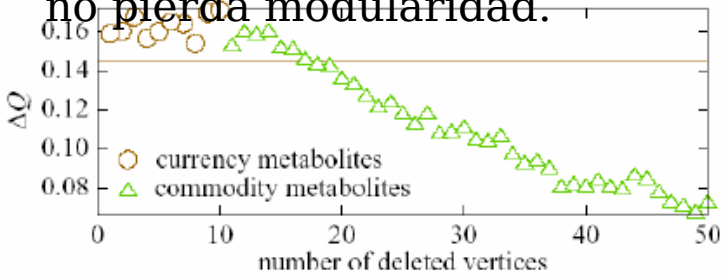
These metabolites are often called **currency metabolites**, since they act as a kind of molecular currency in the cell (for example, ATP). Some investigators call them “current metabolites”, emphasizing their flow through the network. We call the currency metabolites, and by analogy to economics, we call the other metabolites **commodity metabolites**.

Wagner & Fell 2001	Schuster et al. 2002	Ma & Zeng 2003
ATP	ATP	ATP
ADP	ADP	ADP
NADP	NADP	NADP
NADPH	NADPH	
NAD		NAD
NADH		
	P_i	P_i
	H_2O	H_2O
	H^+	
	PP_i	
	CMP	
		CO_2
		O_2
		NH_3

2) Por criterios de conectividad (k>10, 15, ...)

3) Otros criterios topológicos.

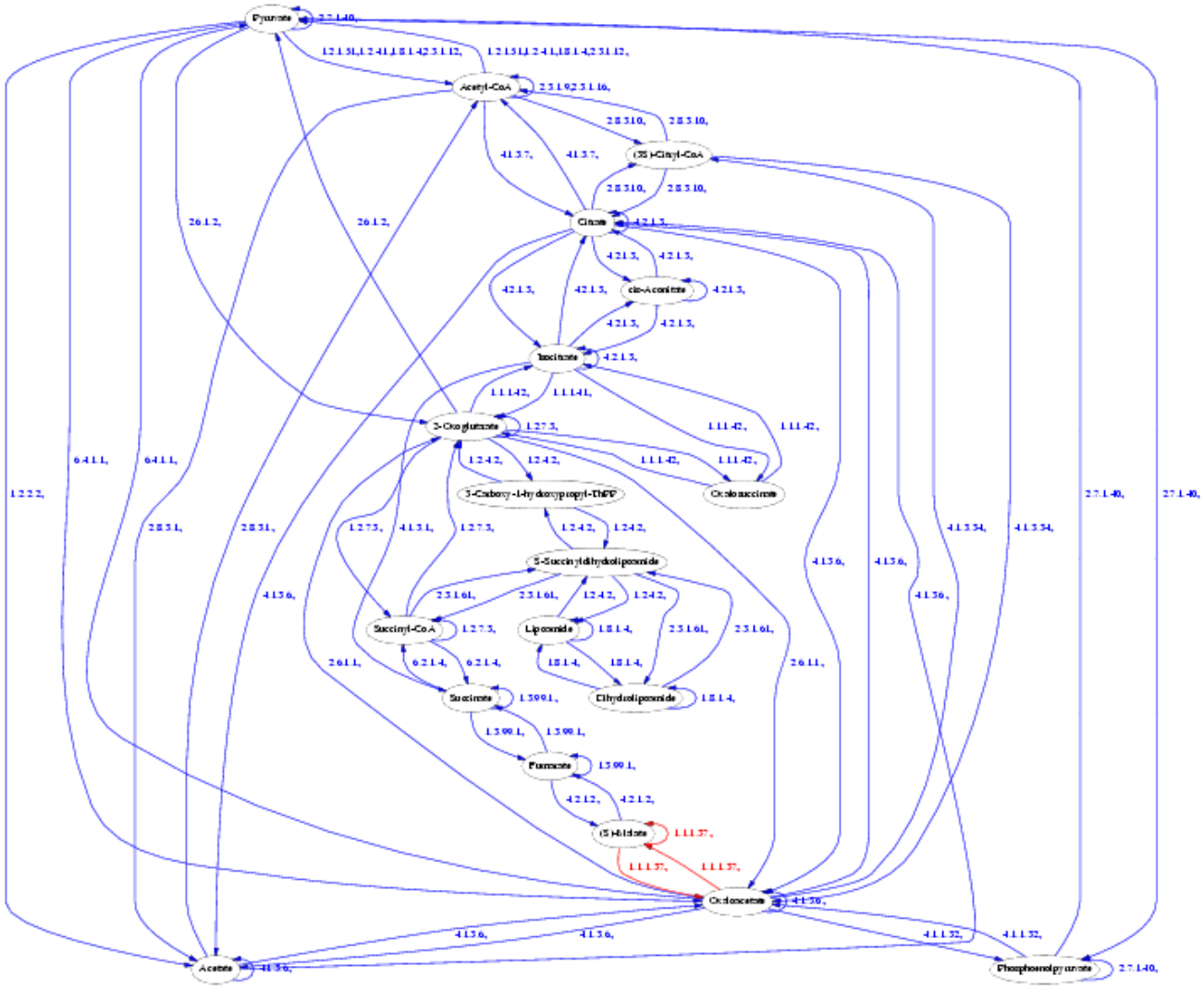
Ej. Metabolitos que conectan muchas veces diferentes módulos: Ir quitando metabolitos muy conectados mientras la red no pierda modularidad.



Wagner & Fell 2001	Schuster et al. 2002	Ma & Zeng 2003
ATP	ATP	ATP
ADP	ADP	ADP
NADP	NADP	NADP
NADPH	NADPH	
NAD		NAD
NADH		
	P_i	P_i
	H_2O	H_2O
	H^+	
	PP_i	
	CMP	
		CO_2
		O_2
		NH_3

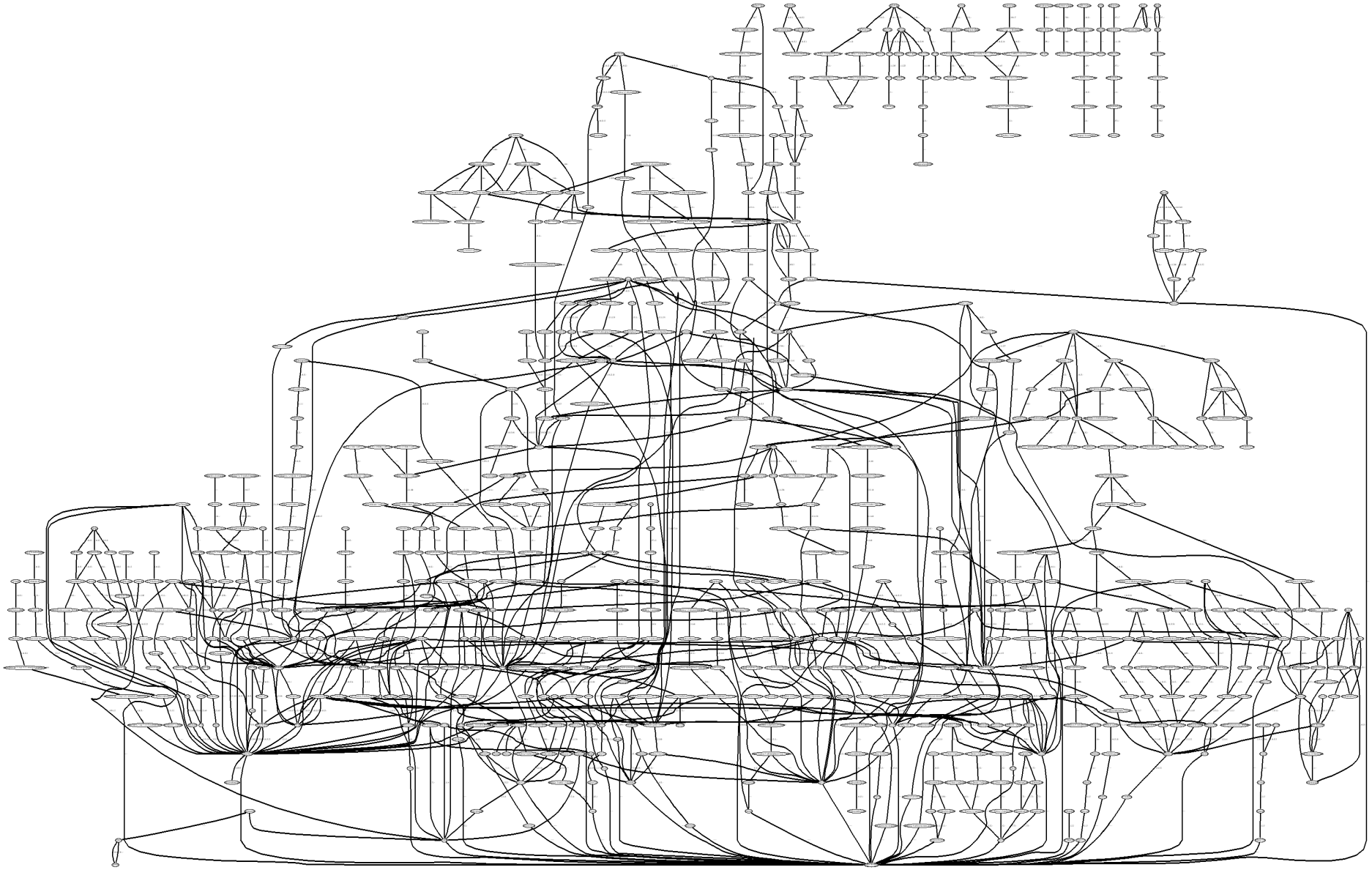
Mikael Huss & Petter Holme

Representación del Metaboloma como Red



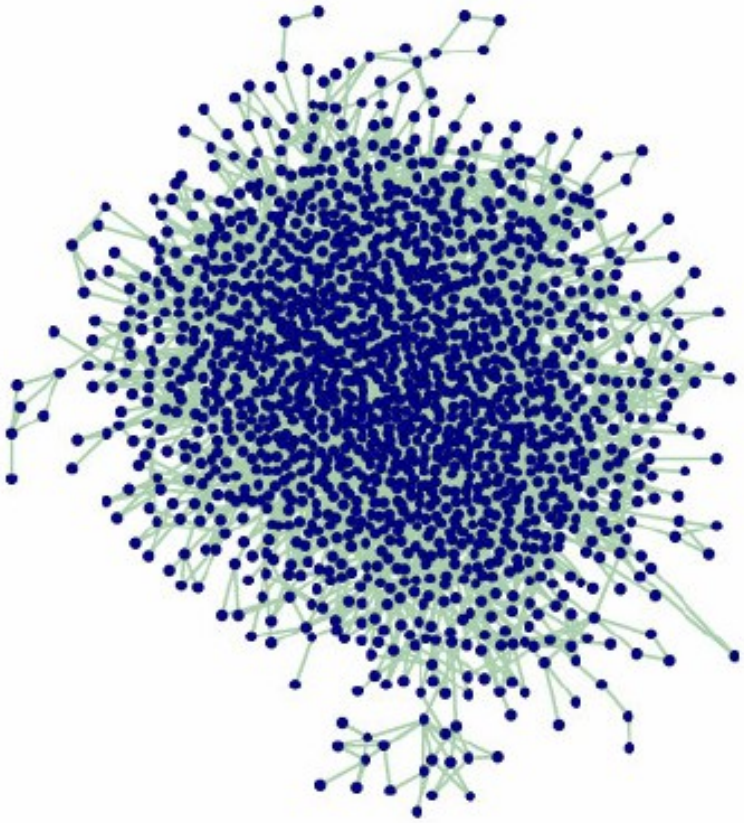
R. Chaleil (IC, London)

Representación del Metaboloma como Red

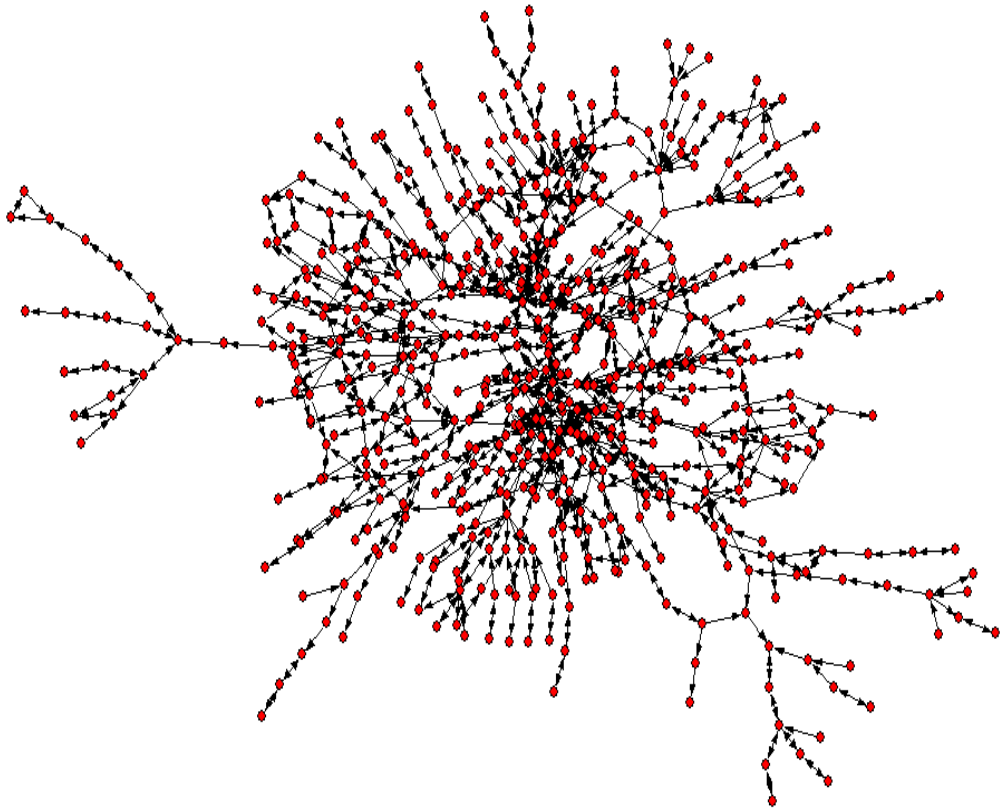


Florencio Pazos, Victor De Lorenzo & Alfonso Valencia. (2003). The organization of the Microbial Biodegradation Network from

Representación del Metaboloma como Red

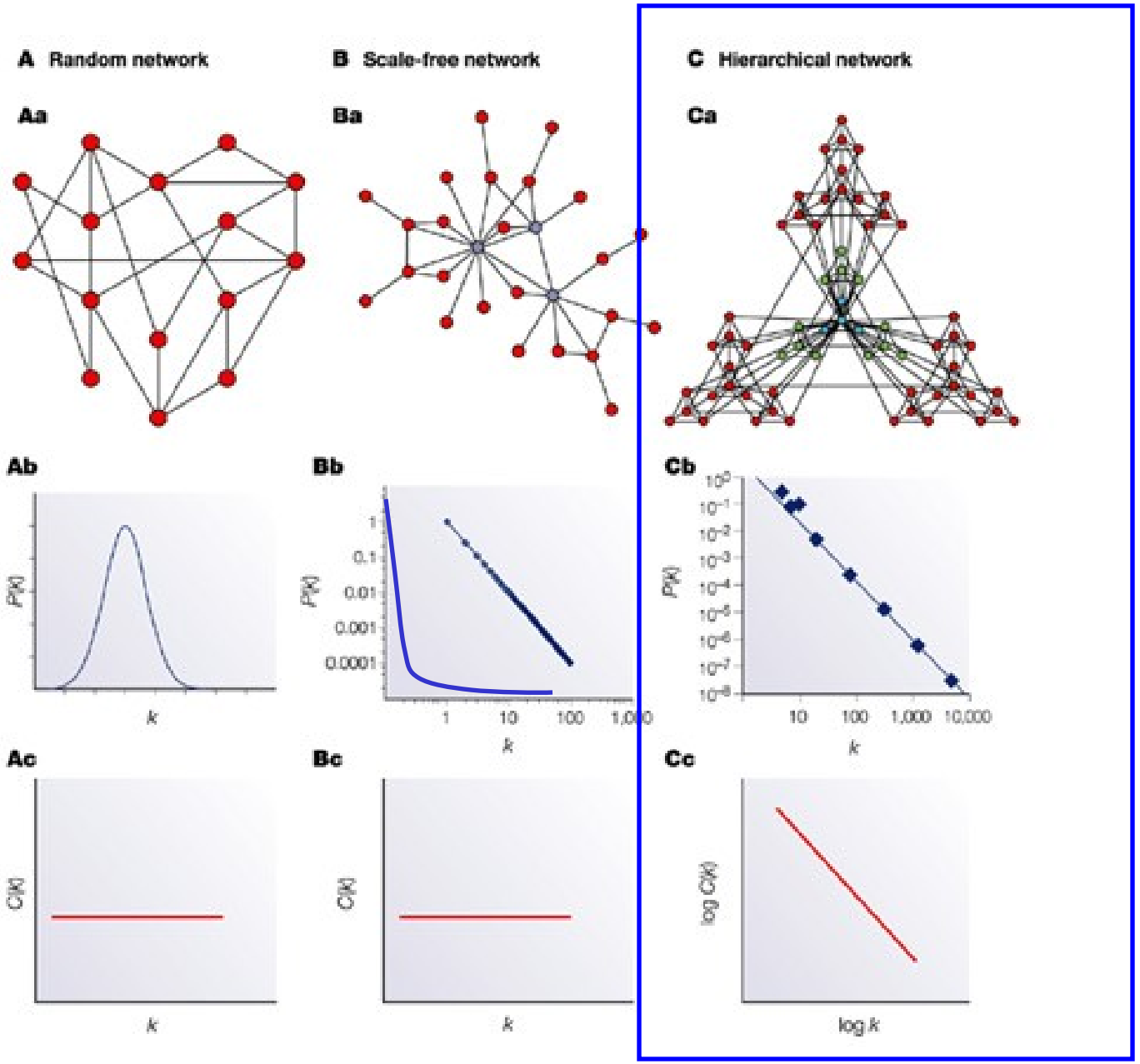


Yeast metabolic network



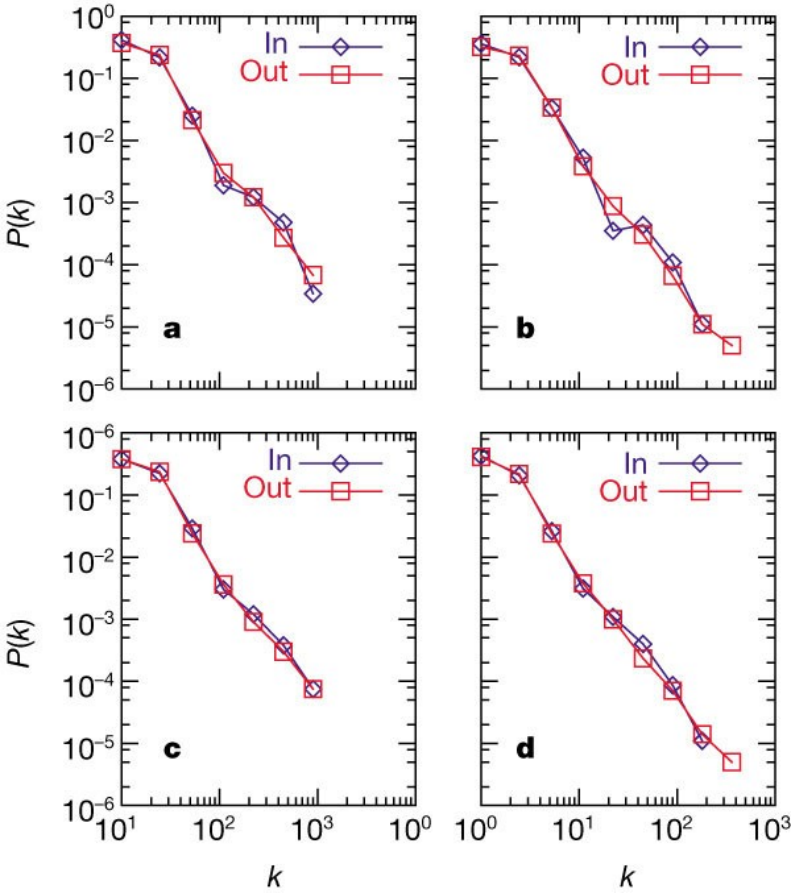
E coli metabolic network

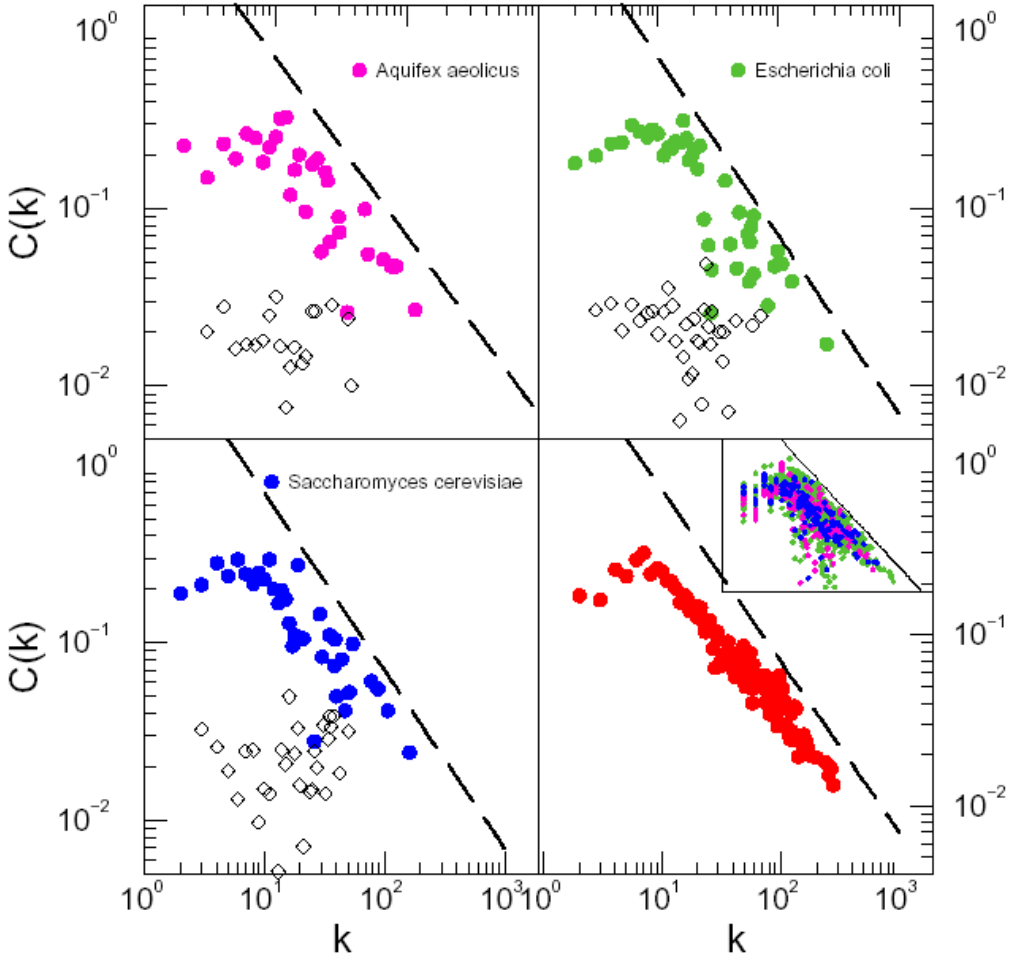
Topological Characteristics of the Metabolic Network



L. and Oltvai, Z.N. (2004) Network biology: understanding the cell's functional organization. *Nat Rev Genet*, **5**, 101-11

Topological Characteristics of the Metabolic Network



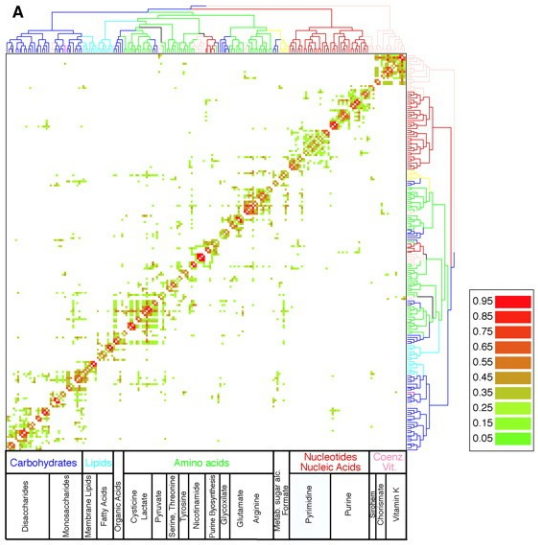


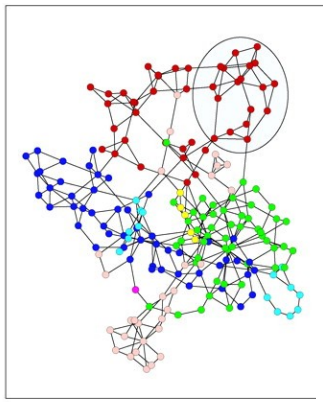
Jeong, H., Tombor, B., Albert, R., Oltvai, Z. N. & Barabasi, A. L. (2000). The large scale organisation of metabolic networks. *Nature* **407**, 651-653

Barabasi, A.L. and Oltvai, Z.N. (2004) Network biology: understanding the cell's functional organization. *Nat Rev Genet*, **5,** 101-113.

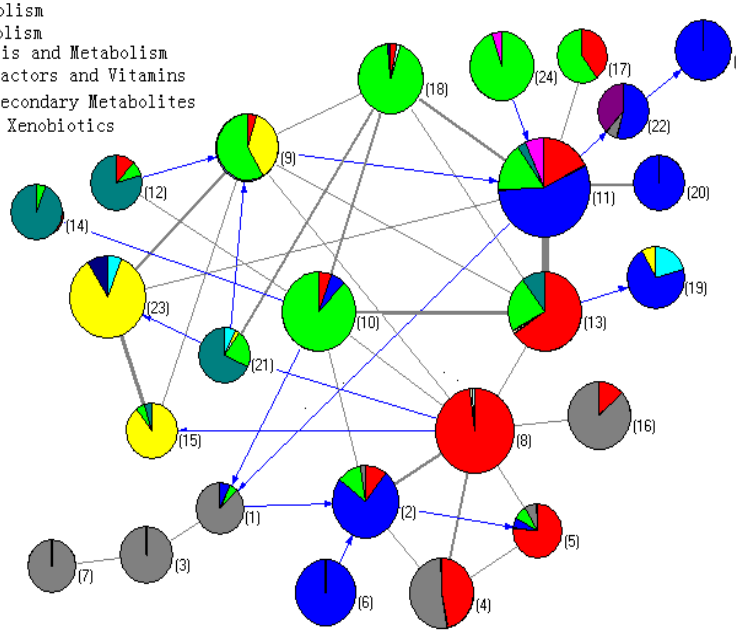
.

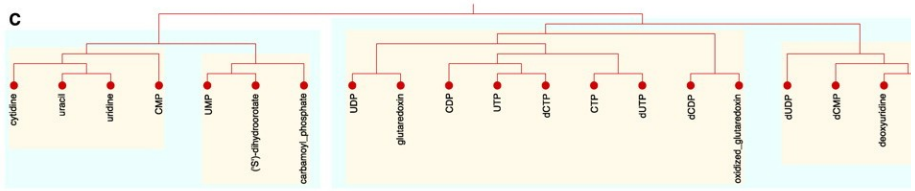
Topological Characteristics of the Metabolic Network

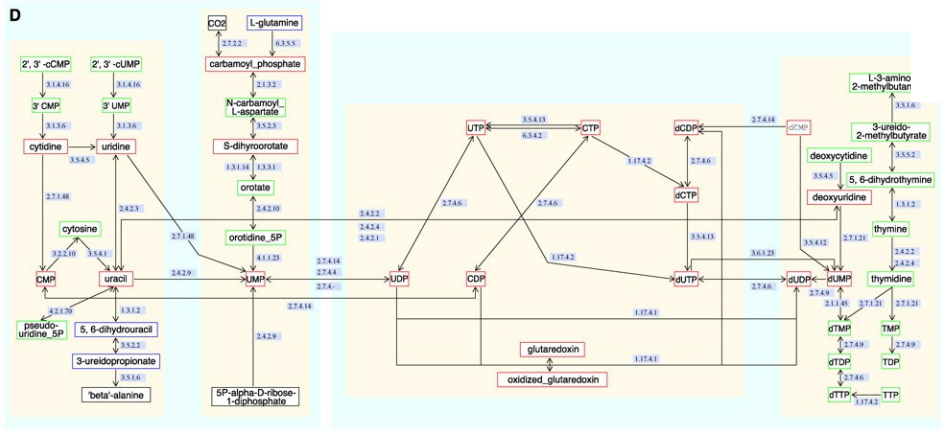




I. Carbohydrate Metabolism
II. Energy Metabolism
III. Lipid Metabolism
IV. Nucleotide Metabolism
V. Amino Acid Metabolism
VI. Glycan Biosynthesis and Metabolism
VII. Metabolism of Cofactors and Vitamins
VIII. Biosynthesis of Secondary Metabolites
IX. Biodegradation of Xenobiotics

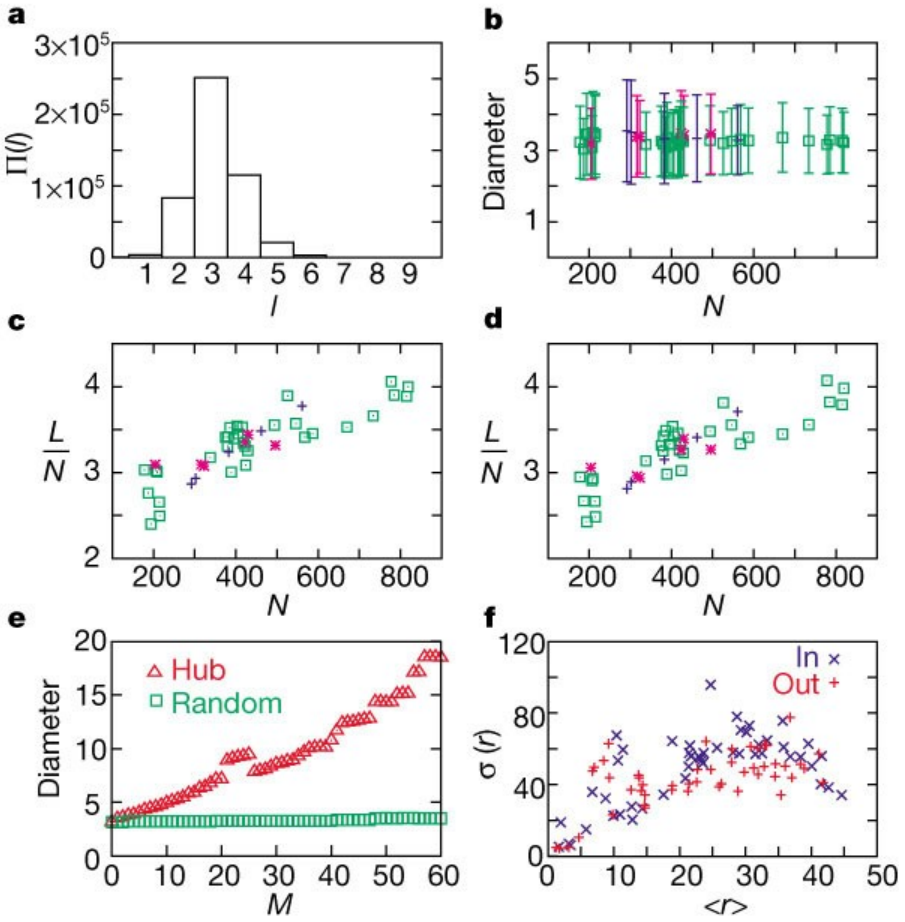






Ravasz, E., Somera, L., Mongru, D.A., Oltvai, Z.N. and Barabási, A.L. (2002) Hierarchical organization of modularity in metabolic networks. *Science*, **297**, 1551-1555.

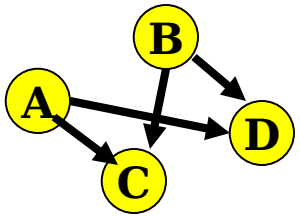
Topological Characteristics of the Metabolic Network



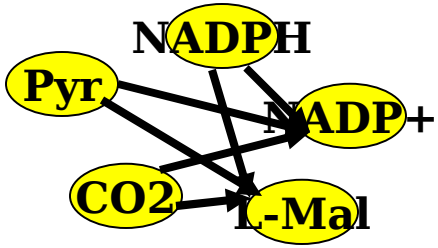
Jeong, H., Tombor, B., Albert, R., Oltvai, Z. N. & Barabasi, A. L. (2000). The large scale organisation of metabolic networks. *Nature* **407**, 651-653.

Small world?

A + B → C + D



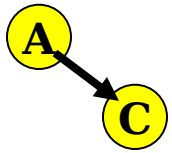
Pyr + NADPH +CO2 → L-Mal + NADP+

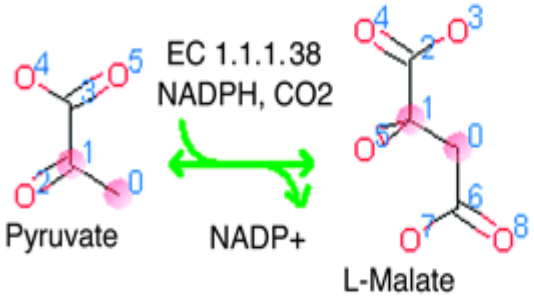


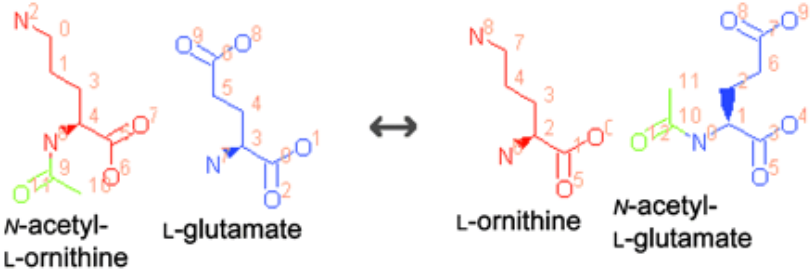
A + B → **C** + D

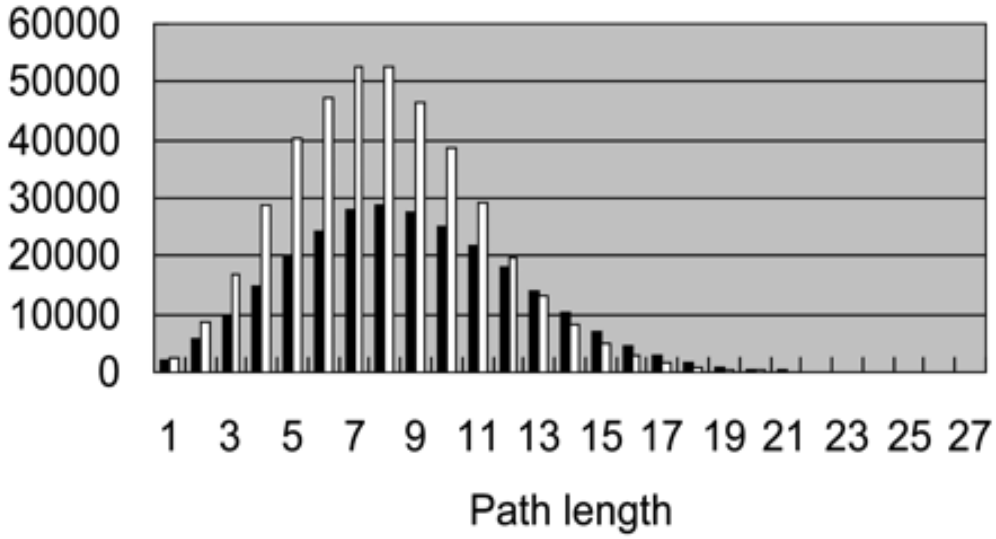
A ⟶ **C**

B D



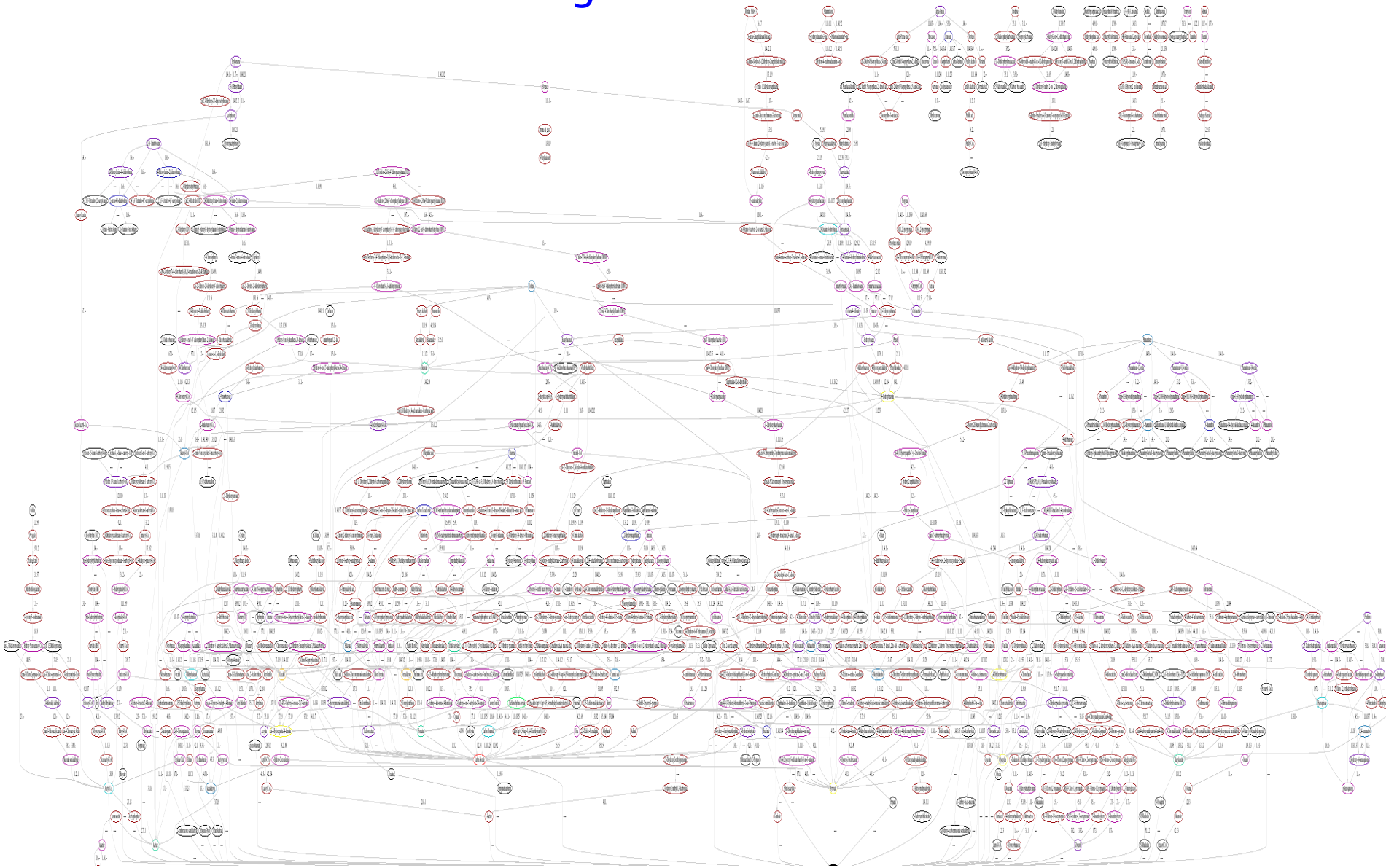






2004). The metabolic world of Escherichia coli is not small. *Proc Natl Acad Sci U S A* **101(6):** 1543-1547.

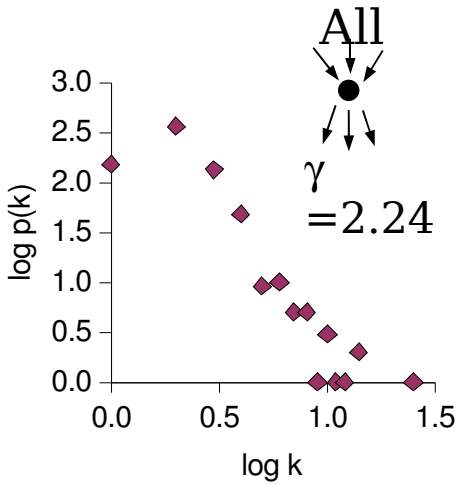
The Biodegradation Network

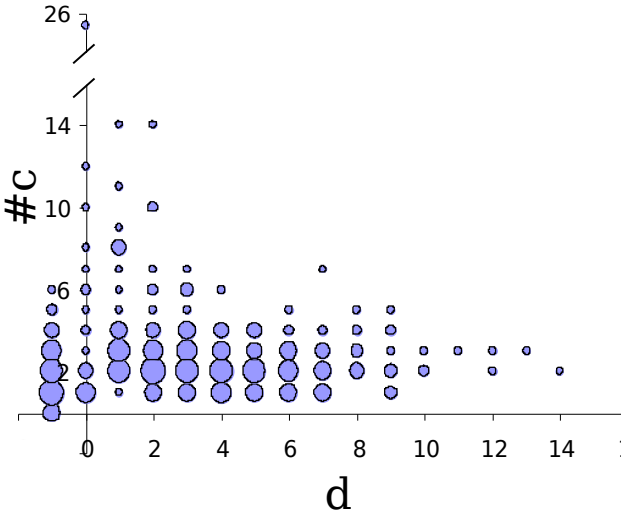


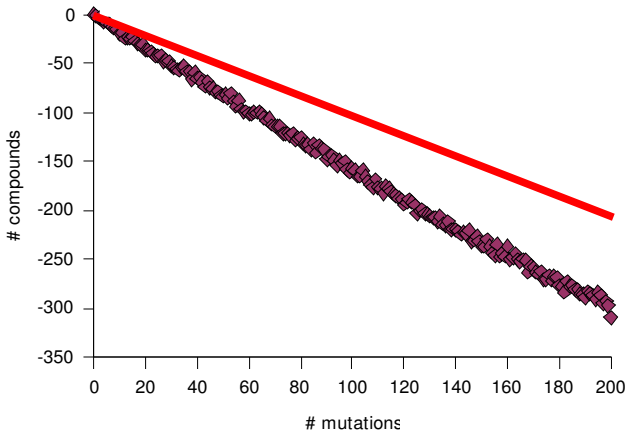
- Florencio Pazos, Victor De Lorenzo & Alfonso Valencia. (2003). The organization of the Microbial Biodegradation Network from a Systems-Biology perspective. *EMBO Rep.* **4(10):**994-999.

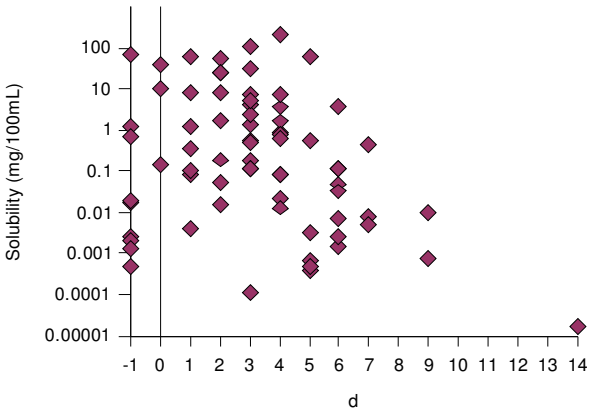
http://pdg.cnb.uam.es/biodeg_net

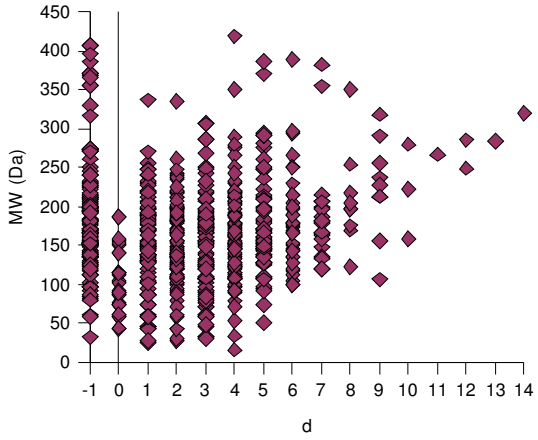
Properties of the Biodegradation Network

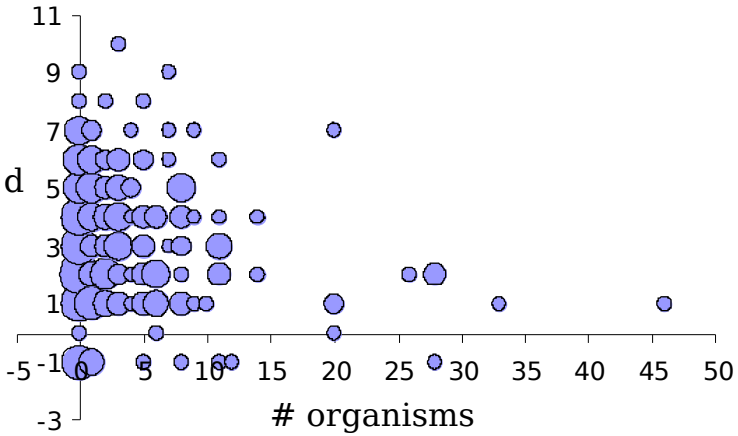






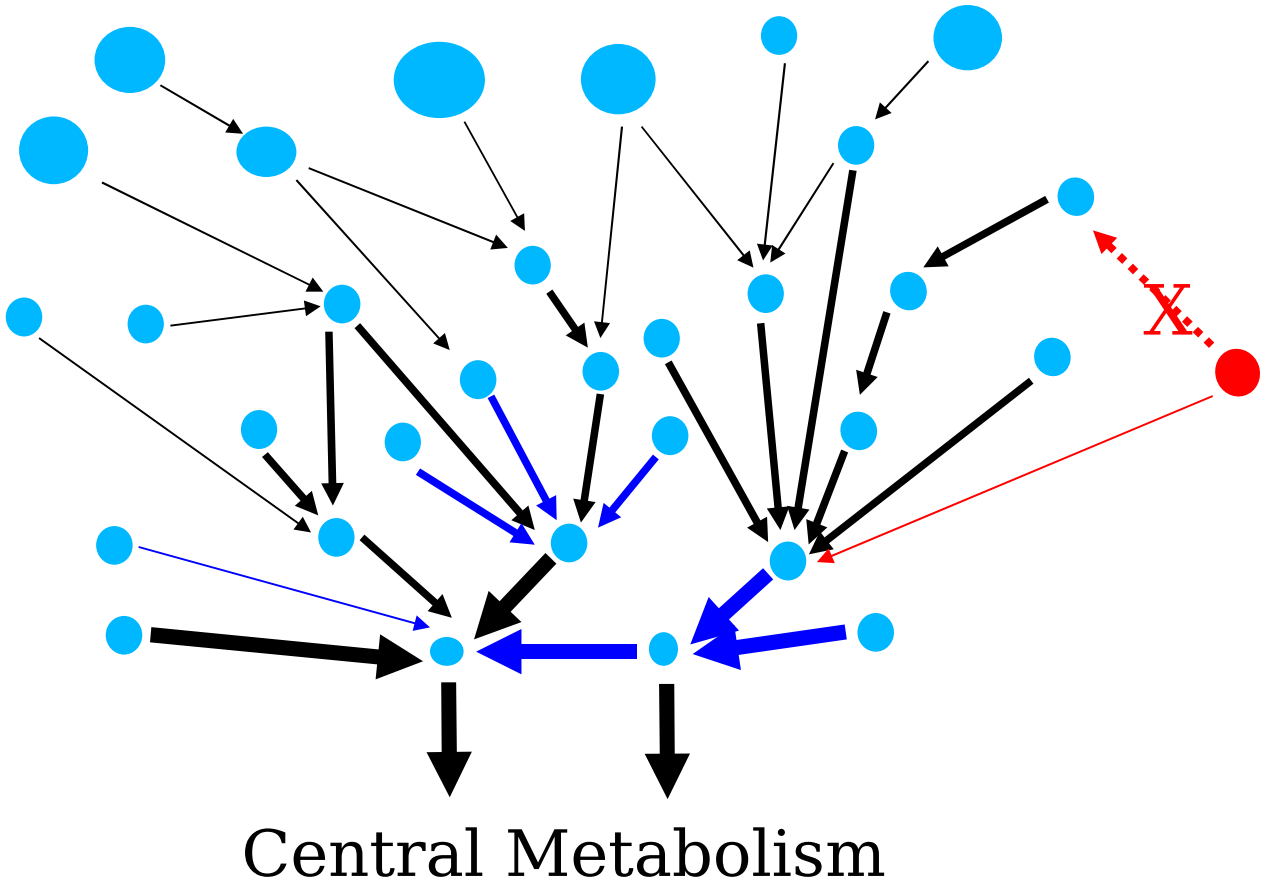






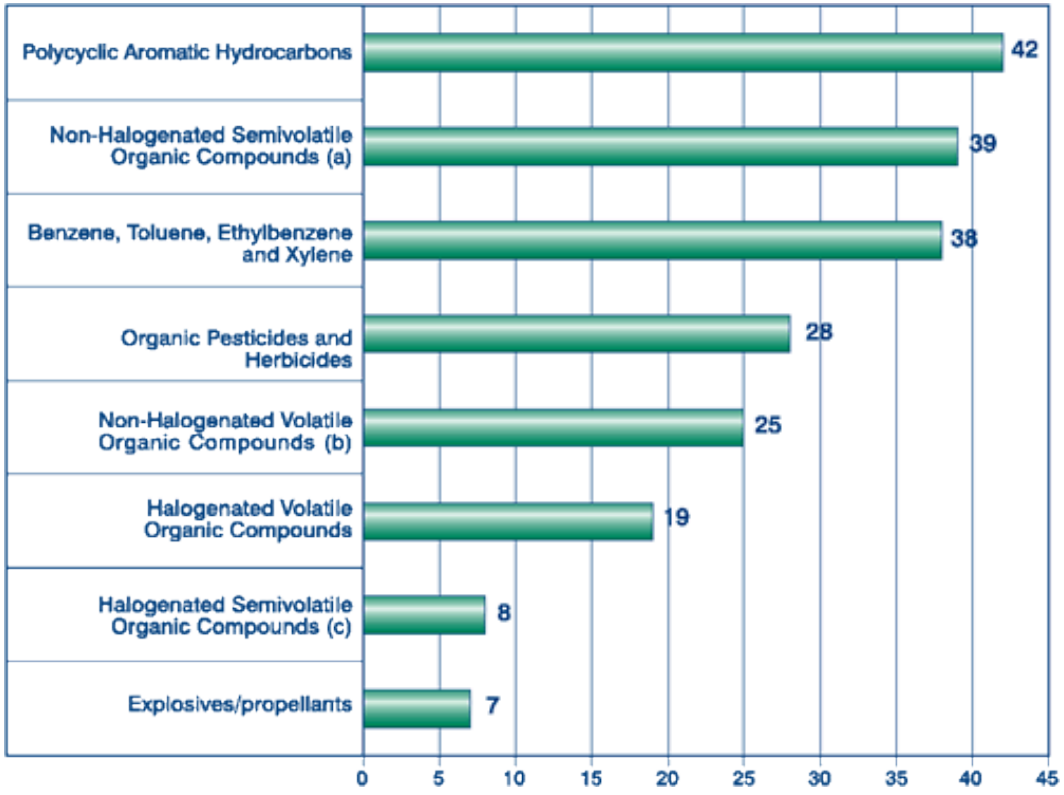
• Florencio Pazos, Victor De Lorenzo & Alfonso Valencia. (2003). The organization of the Microbial Biodegradation Network from a Systems-Biology perspective. *EMBO Rep.* **4(10):**994-999.
http://pdg.cnb.uam.es/biodeg_net

Properties of the Biodegradation Network



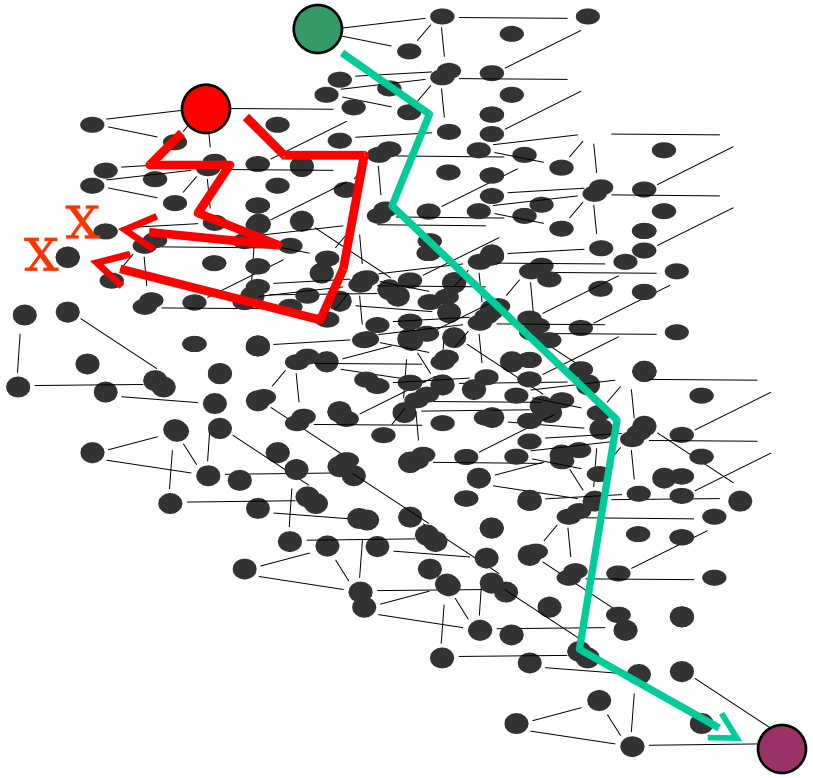
• Florencio Pazos, Victor De Lorenzo & Alfonso Valencia. (2003). The organization of the Microbial Biodegradation Network from a Systems-Biology perspective. *EMBO Rep.* **4(10):**994-999.

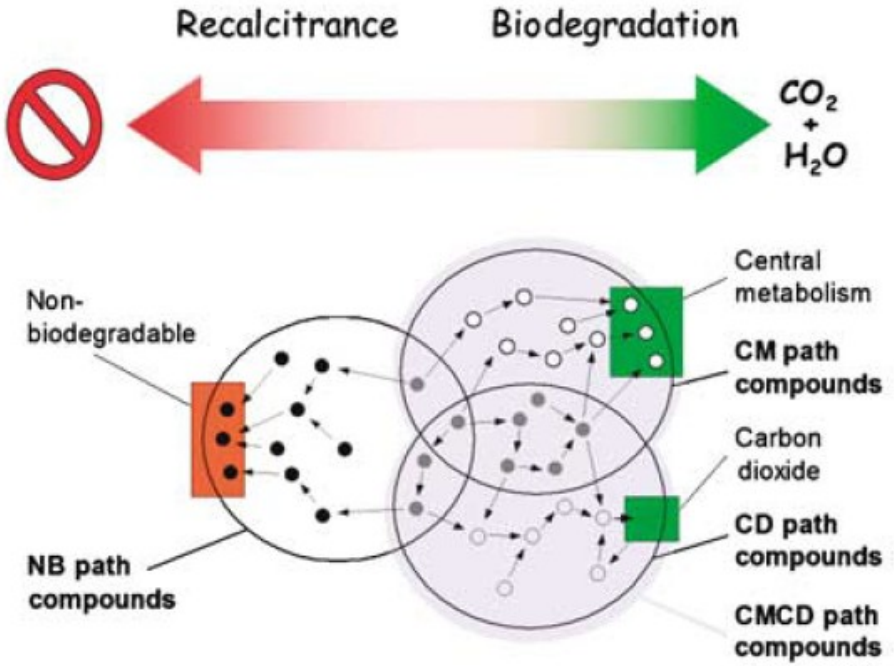
Biodegradación/Bioremediación



Obtaining examples of biodegradable/recalcitrant compounds

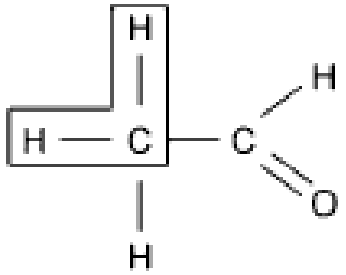
umption: biodegradable compounds are the ones which have a pathway to CM.
Otherwise they are recalcitrant.





Description of the chemical structure

Chemical structure



SMILES

CC=O

Atomic triads

C–C–H	4
C–C=O	1
H–C–H	3
H–C=O	1

MW

44 Da

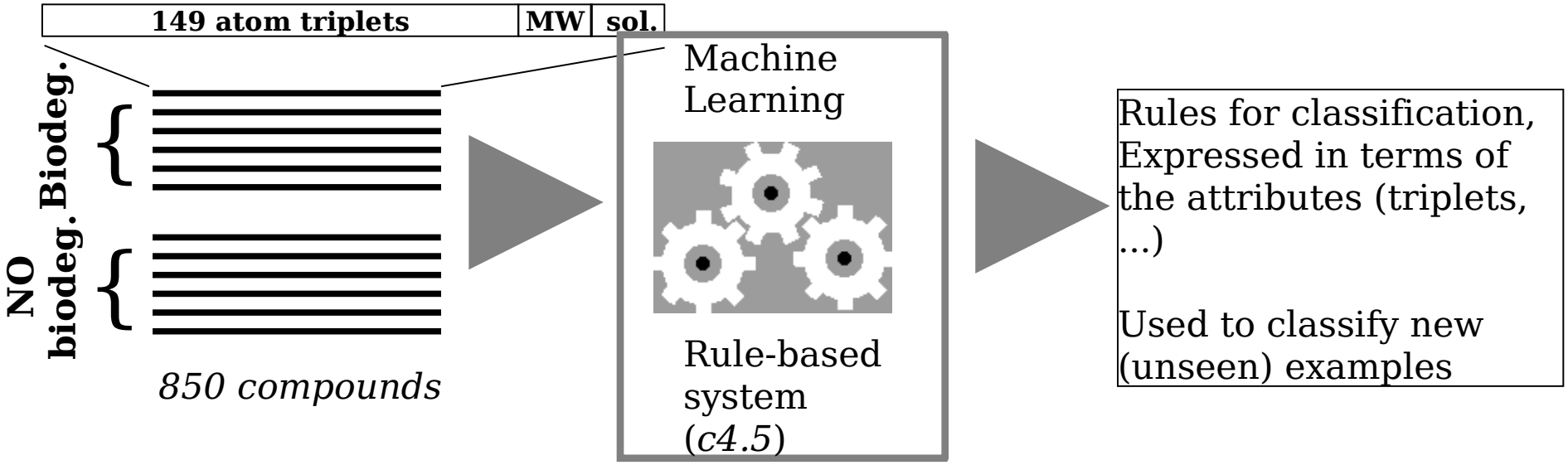
Solubility

0.1 g/100 ml

Compound vector = (0.1, 44 ,4 ,1 ,3 ,1 ,0 ,0 ,.. Vector len= 152 (149 triads + MW + ws)

Gómez, MJ, Pazos F, Guijarro FJ, de Lorenzo V, Valencia A. (2007). The environmental fate of organic pollutants through the global microbial metabolism. *Mol Syst Biol*. **3:**114.

Training the system



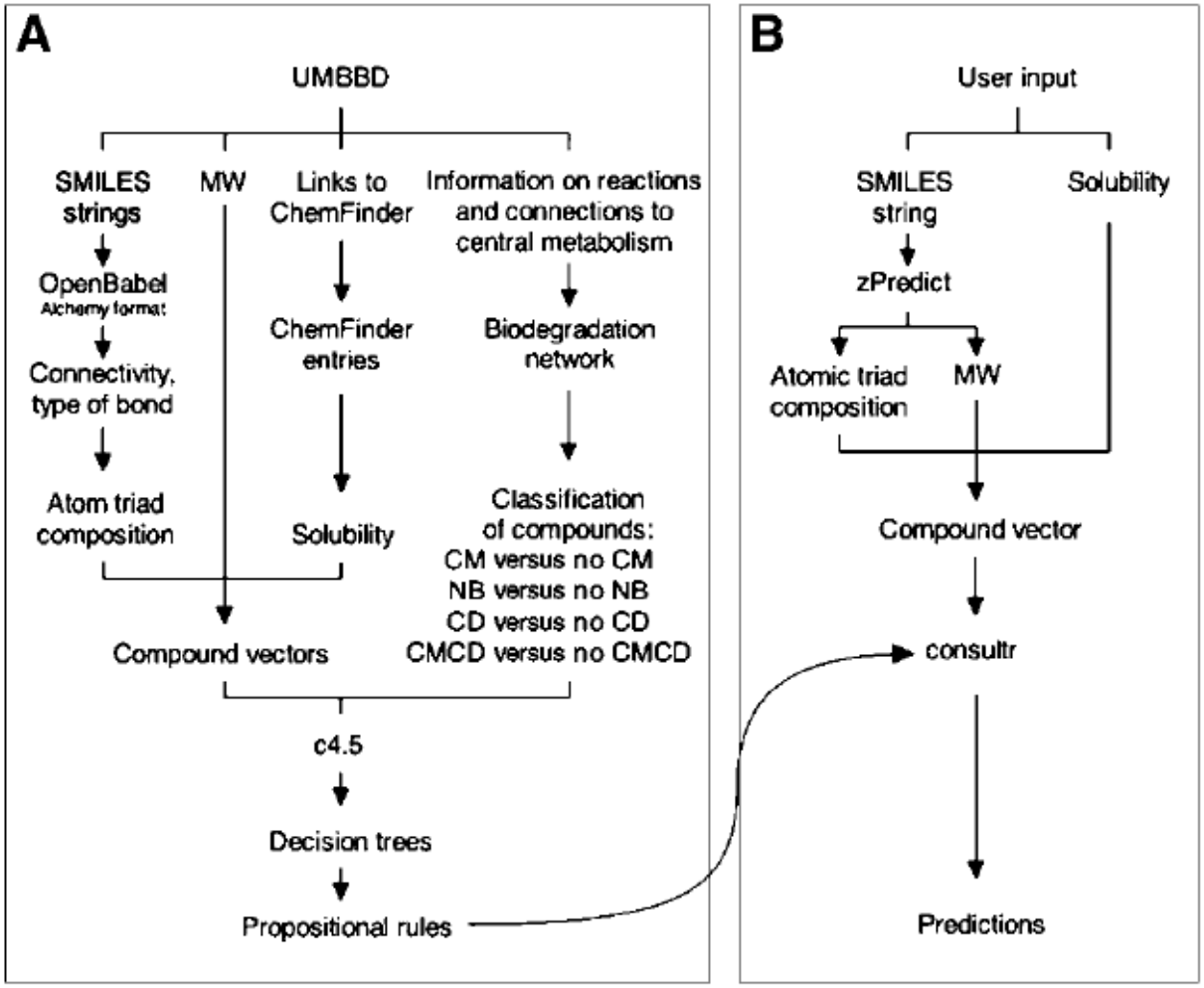
tages of rule-based systems compared with other ML, i.e. NN:

- Explicit rules in human-readable format
- Can handle missing values (solubility)

---

n, JR. (1993). *c4.5: Programs for Machine Learning*. San Mateo, CA. Morgan-Kaufmann

Schema



Gómez, MJ, Pazos F, Guijarro FJ, de Lorenzo V, Valencia A. (2007). The environmental fate of organic pollutants through the global microbial metabolism. *Mol Syst Biol*. **3:**114.

Results

Table I Example of propositional rule generated for the classification compounds in the scheme NB *or* No NB

Rule 55: IF	–C–C–C >19
	–O–C–C >1
	–O–C–C ⩽3

THEN, the compound belongs to the NB class (Confidence 90.6%)
Examples (14 cases)

52 out of the 152 attributes involved in rules

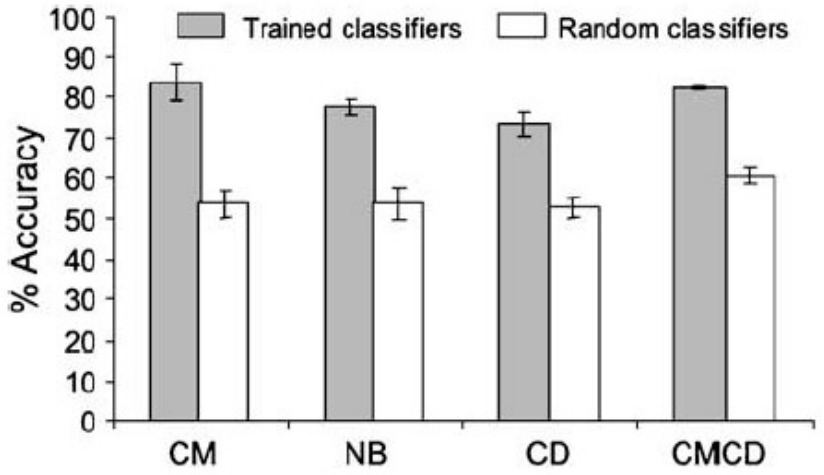


Table II Predictive performance in fivefold cross-validation experiments

Classification scheme[a]	CM or No CM	NB or No NB	CD or No CD	CMCD or No CMCD
Accuracy (%)	87±4	77±4	73±5	82±3
Significance respect to random—P(N)	1.1×10^{-39}	1.2×10^{-24}	2.4×10^{-16}	6.4×10^{-26}
Default class	CM	No NB	No CD	CMCD
Majority class	CM	No NB	No CD	CMCD
No. of cases	533	496	513	634
Sensitivity (%)	93±4	86±4	78±7	92±3
Specificity (%)	83±3	78±2	78±3	85±2
Minority class				
No. of cases	308	345	328	207
Sensitivity (%)	67±5	64±5	66±9	50±11
Specificity (%)	85±8	77±4	66±5	71±5

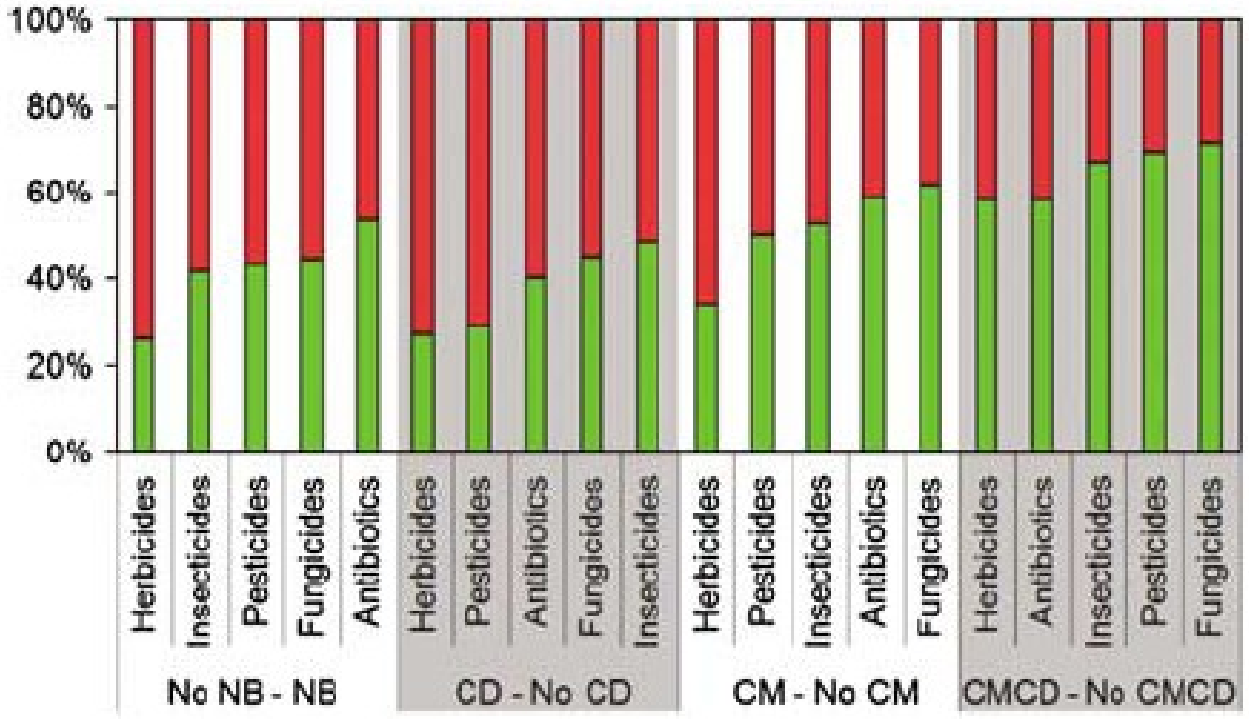
***coli* CM (733 comps.): 90.31% CMCD; 81.44% *No*-NB**

Gómez, MJ, Pazos F, Guijarro FJ, de Lorenzo V, Valencia A. (2007). The environmental fate of organic pollutants through the global microbial metabolism. *Mol Syst Biol*. **3:**114.

Blind predictions

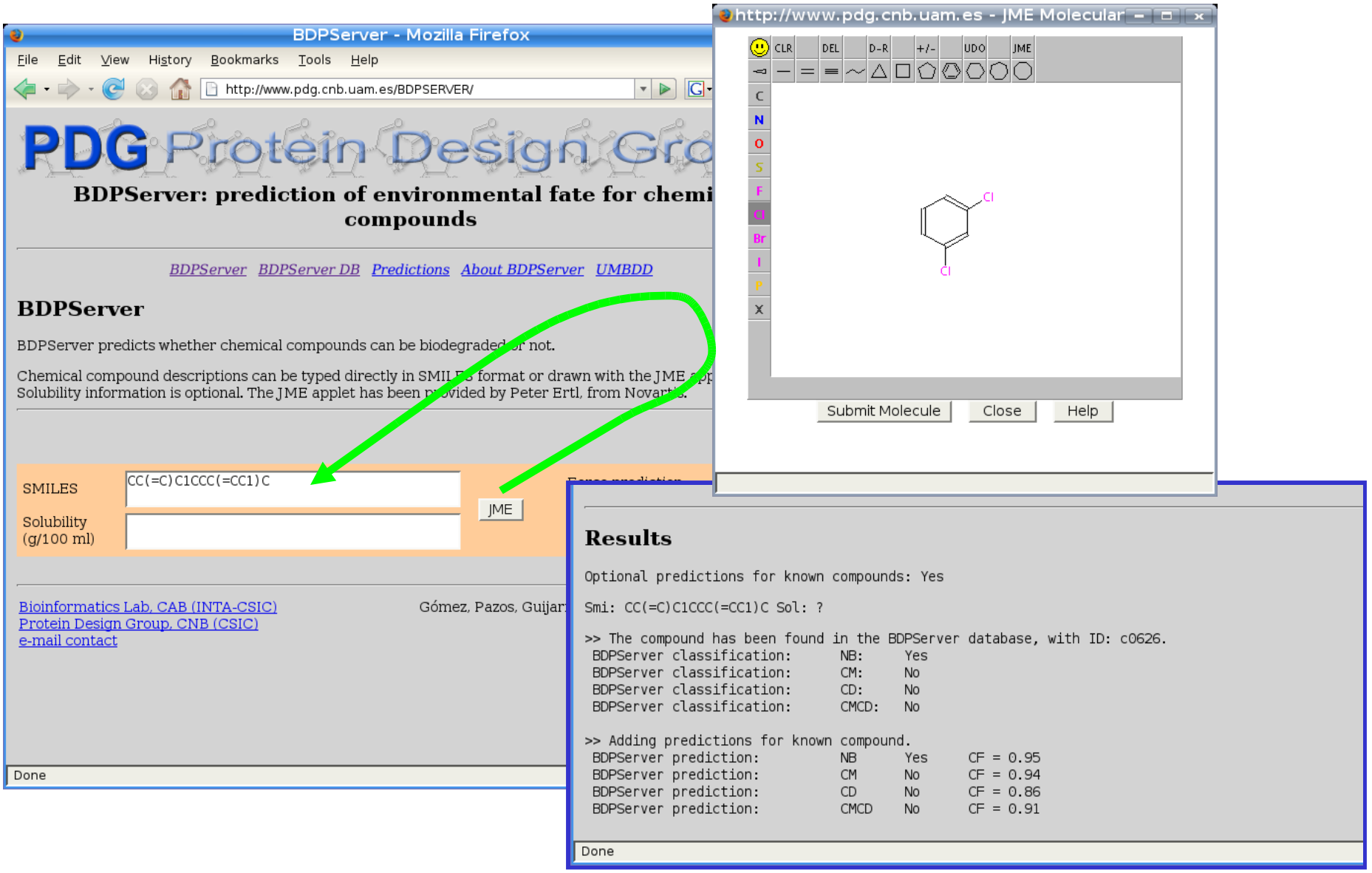
- Chemical compounds regulated by the ECB (HPVCs, LPVCs & *Annex-I*): 70% CMCD

  - PubChem (USA). 3600 comps. aprox.



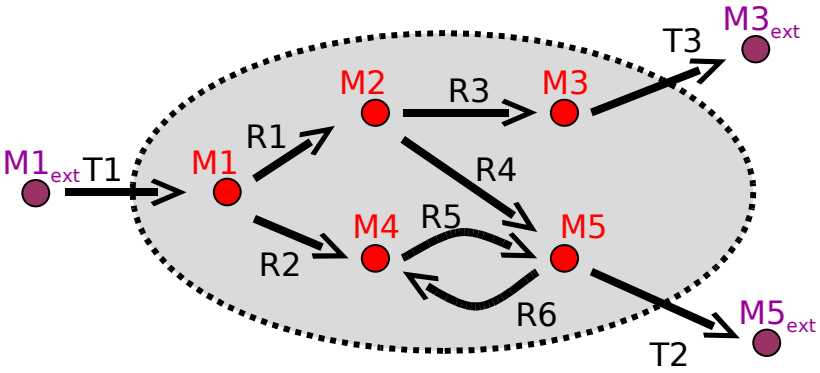
Gómez, MJ, Pazos F, Guijarro FJ, de Lorenzo V, Valencia A. (2007). The environmental fate of organic pollutants through the global microbial metabolism. *Mol Syst Biol*. **3:**114.

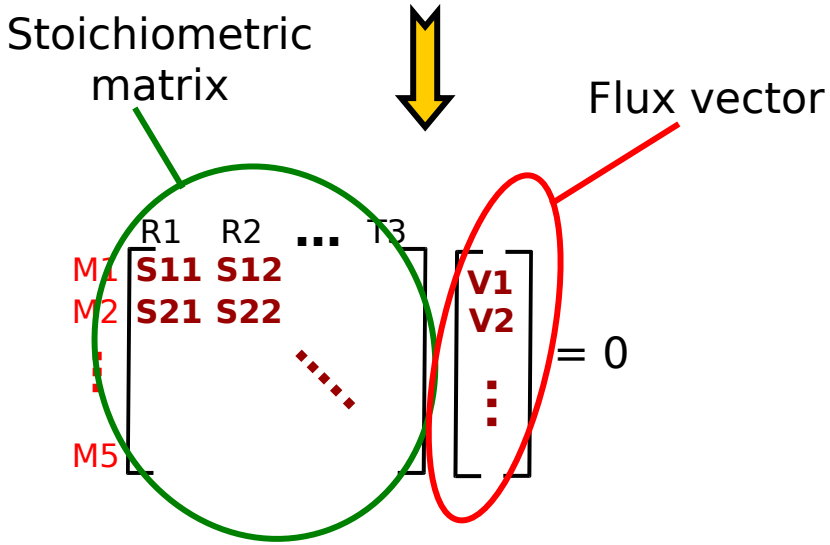
Using the predictor - *BDPServer*



- Gómez, MJ, Pazos F, Guijarro FJ, de Lorenzo V, Valencia A. (2007). The environmental fate of organic pollutants through the global microbial metabolism. *Mol Syst Biol*. **3:**114.
- http://pdg.cnb.csic.es/DBPSERVER/

Análisis de Flujo



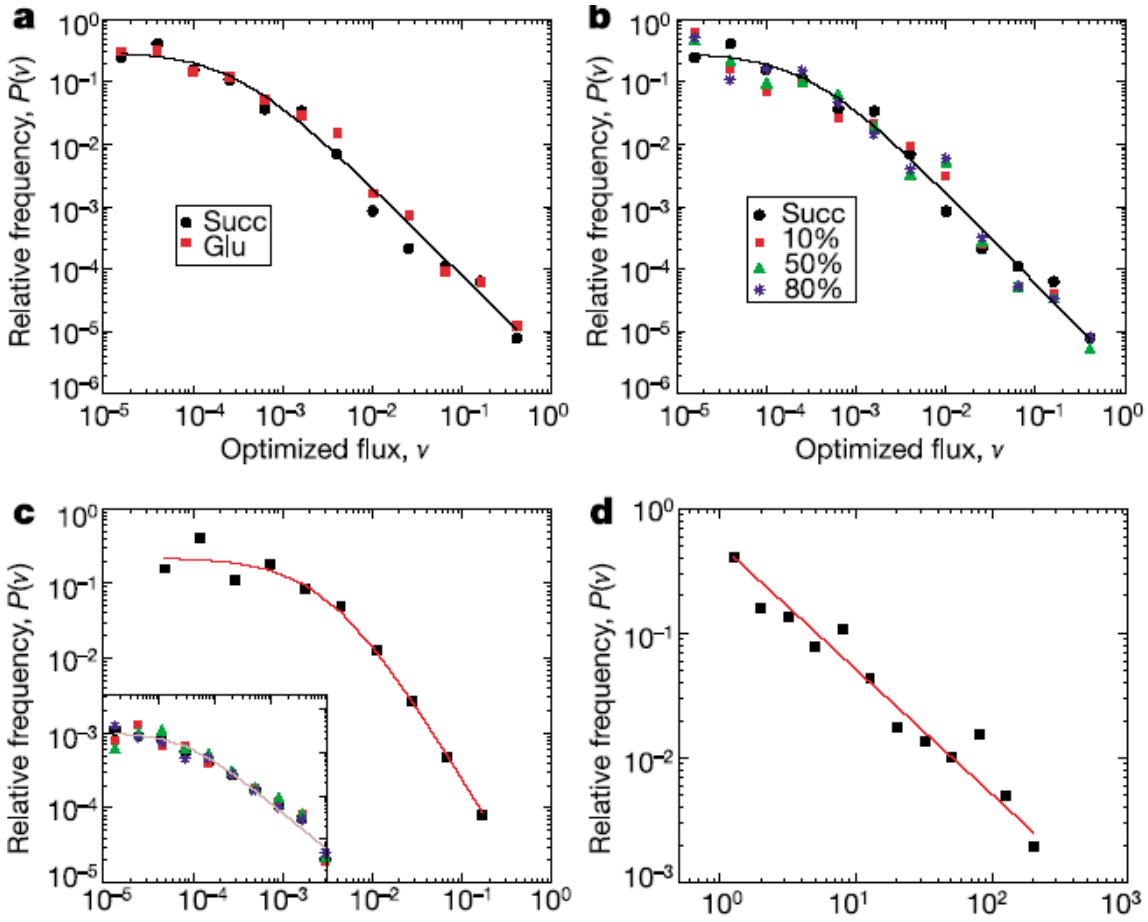


2 M1 → 3 M2

=> S11=2; S21=3

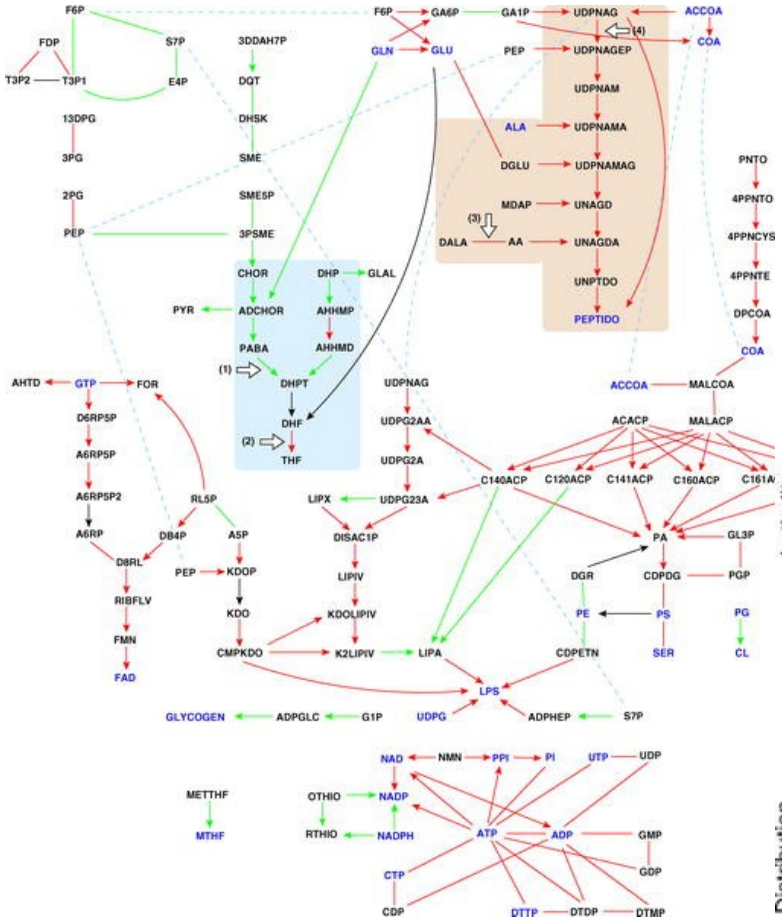
J.S. Edwards & B.O. Palsson, *Proc. Natl. Acad. Sci. USA* **97**, 5528 (2000)
R.U. Ibarra, J.S. Edwards & B.O. Palsson, *Nature* **420**, 186 (2002)
D. Segre, D. Vitkup & G.M. Church, *Proc. Natl. Acad. Sci. USA* **99**, 15112 (2002)

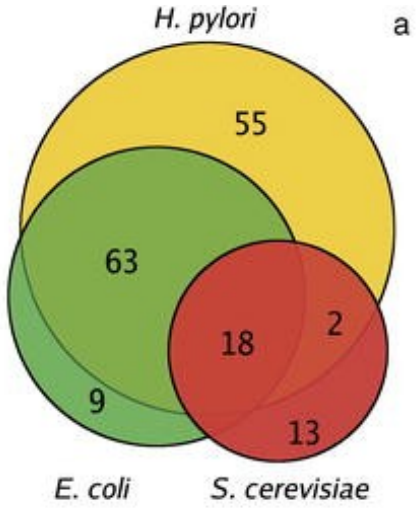
scale-free más alla de la topología

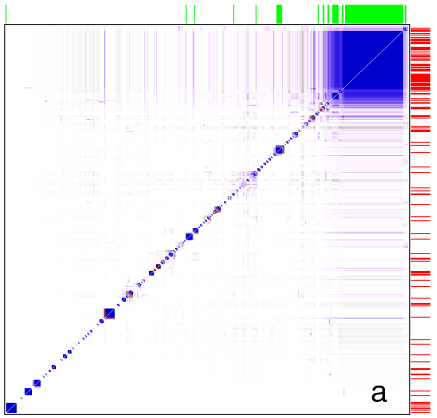


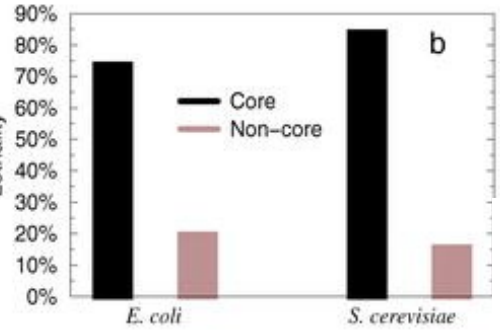
Almaas, E., Kovacs, B., Vicsek, T., Oltvai, Z.N. and Barabasi, A.L. (2004) Global organization of metabolic fluxes in the bacterium Escherichia coli. *Nature.*, **427**, 839-843.

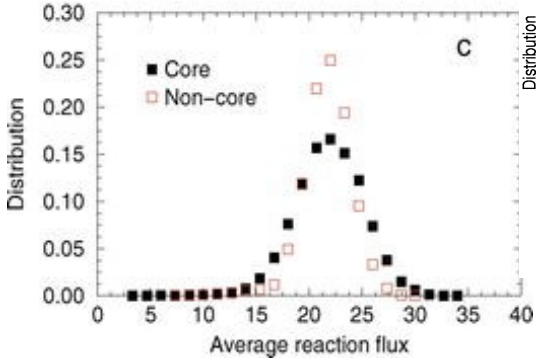
Características del *core* de flujos metabólicos

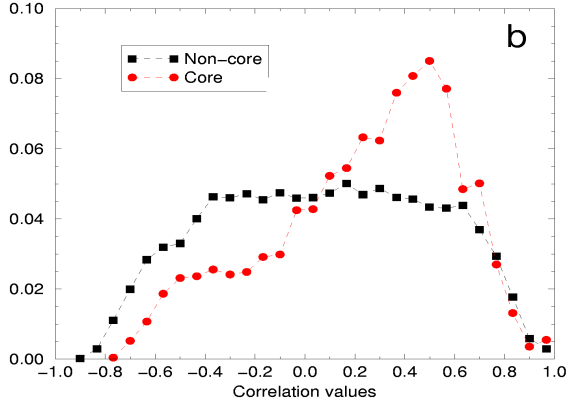








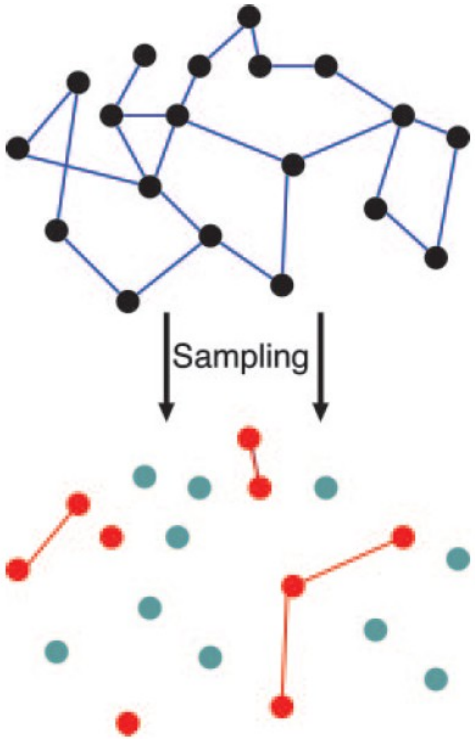




Oltvai, Z.N. and Barabasi, A.L. (2005) The activity reaction core and plasticity of metabolic networks. *PLoS Comput Bi*

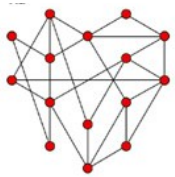
¿Como apareció la estructura *scale-free*?

Muestreo (=>artefacto)



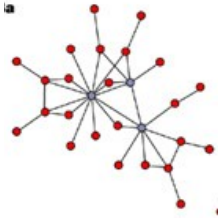
Evolución

Random network

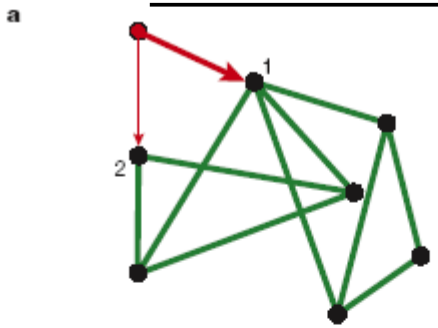


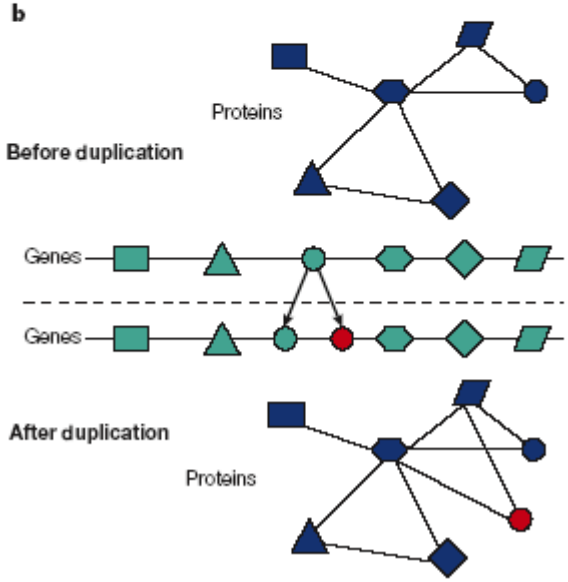
evolution

“scale-free”



Crecimiento



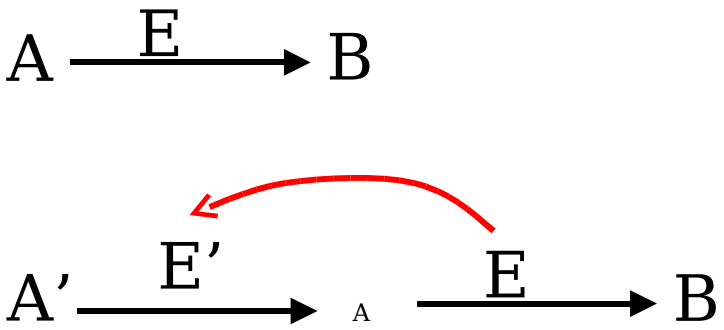


Stumpf, M.P., Wiuf, C. and May, R.M. (2005) Subnets of scale-free networks are not scale-free: Sampling properties of networks. *Proc Natl Acad Sci U S A*, **102**, 4221-4224.

Barabasi, A.L. and Oltvai, Z.N. (2004) Network biology: understanding the cell's functional organization. *Nat Rev Genet*, **5**, 101-113.

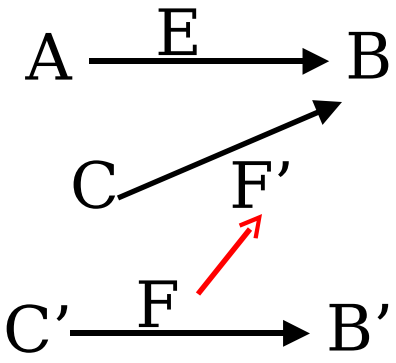
Evolución de la Red Metabólica

Retroevolution



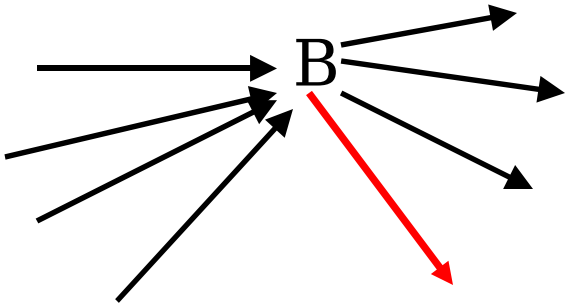
Horowitz, N.H. (1945) On the evolution of biochemical syntheses. *Proc. Natl Acad. Sci. USA,* **31,** 153–157.

Recruitment



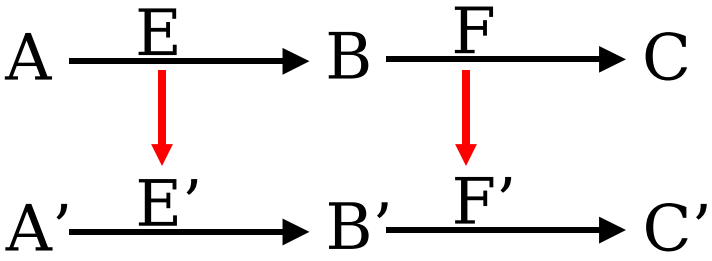
Jensen, R.A. (1976) Enzyme recruitment in evolution of new function. *Annu Rev Microbiol*, **30**, 409-425.

Preferential attachment



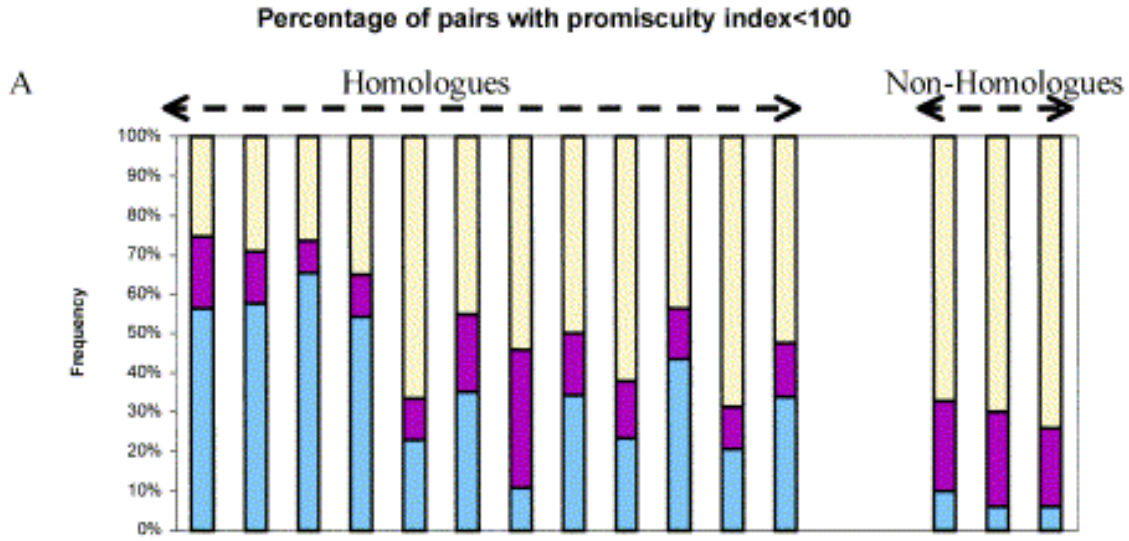
Schmidt, S., Sunyaev, S., Bork, P. and Dandekar, T. (2003) Metabolites: a helping hand for pathway evolution? *Trends Biochem Sci*, **28**, 336-341.

Pathway duplication

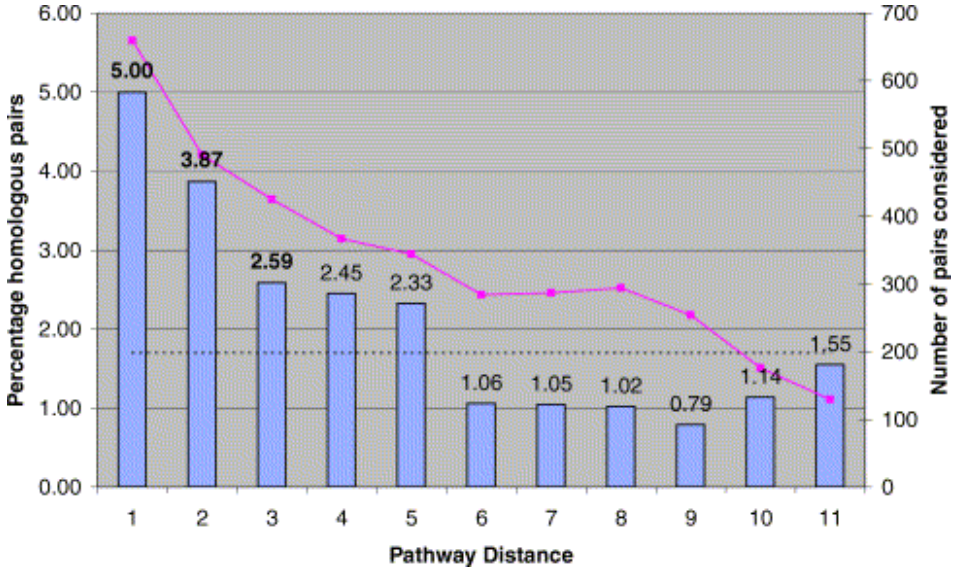


Rison, S.G.C. and Thornton, J.M. (2002) Pathway evolution, structurally speaking. *Curr Opin Struct Biol*, **12**, 374-382.

Evolución de la Red Metabólica

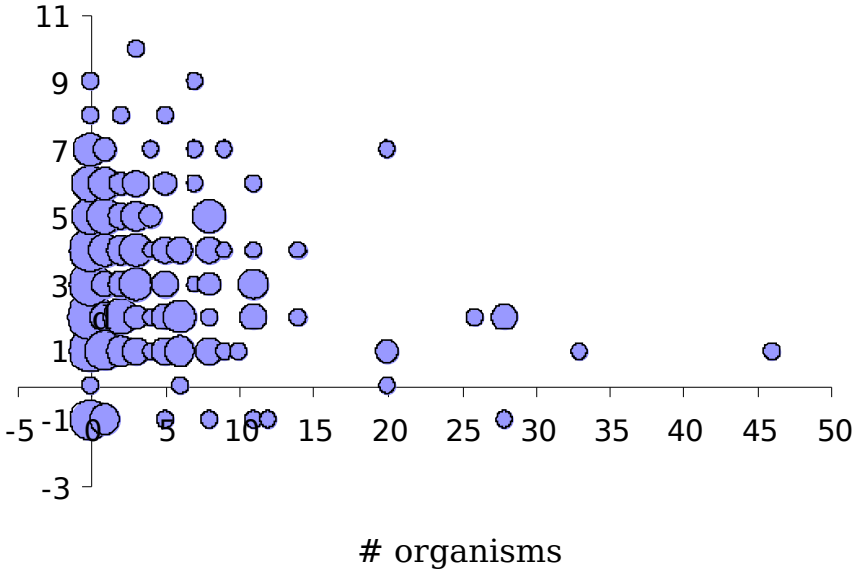


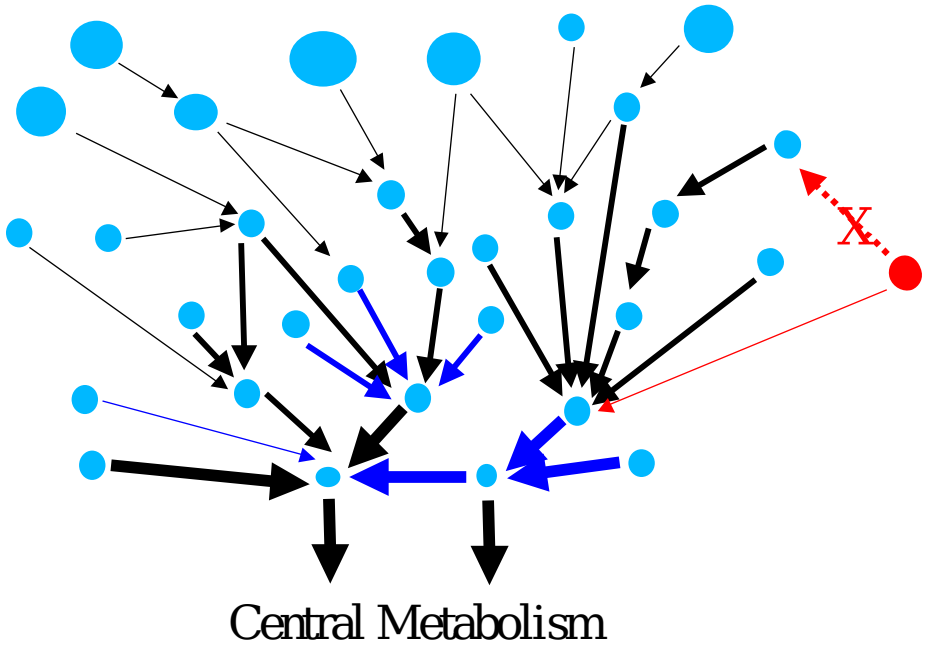
Alves, R., Chaleil, R.A.G. and Sternberg, M.J.E. (2002) Evolution of enzymes in metabolism: a network perspective. *J Mol Biol*, **320,** 751-770.



Rison, S.G.C., Teichmann, S. and Thornton, J.M. (2002) Homology, pathway distance and chromosomal localization of the small molecule metabolism enzymes in *Escherichia coli*. *J Mol Biol*, **318**, 911-932

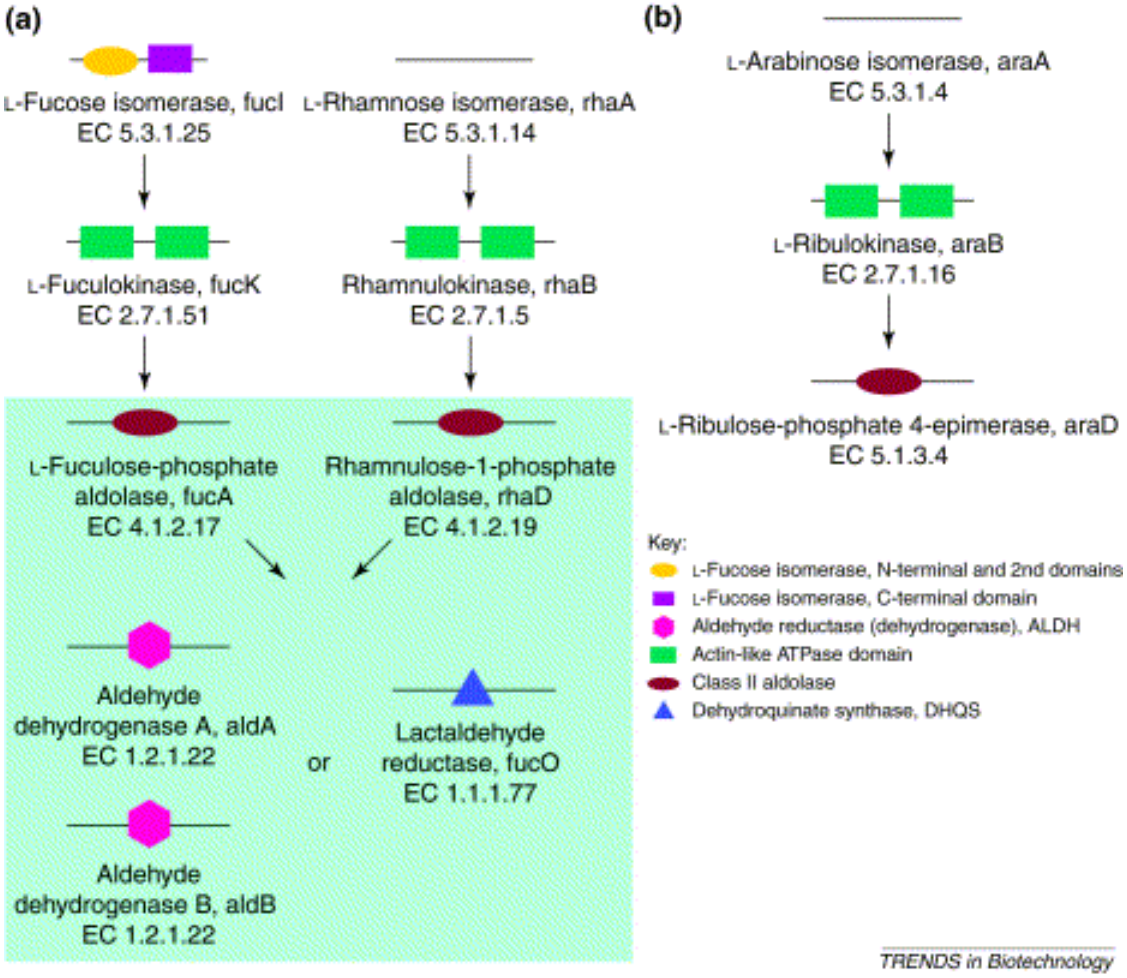
Evolución de la Red Metabólica



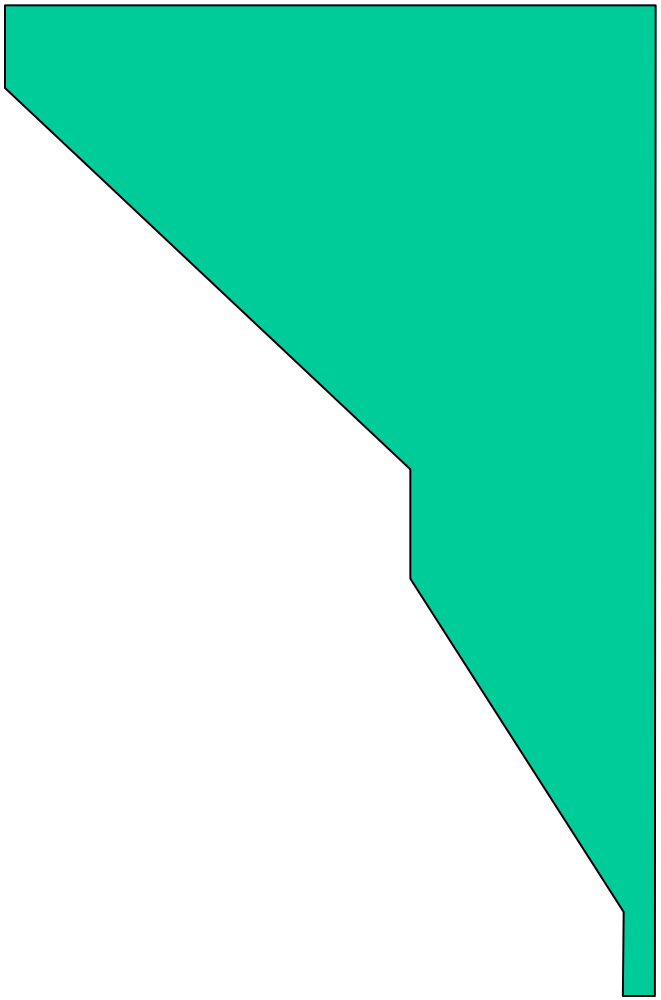


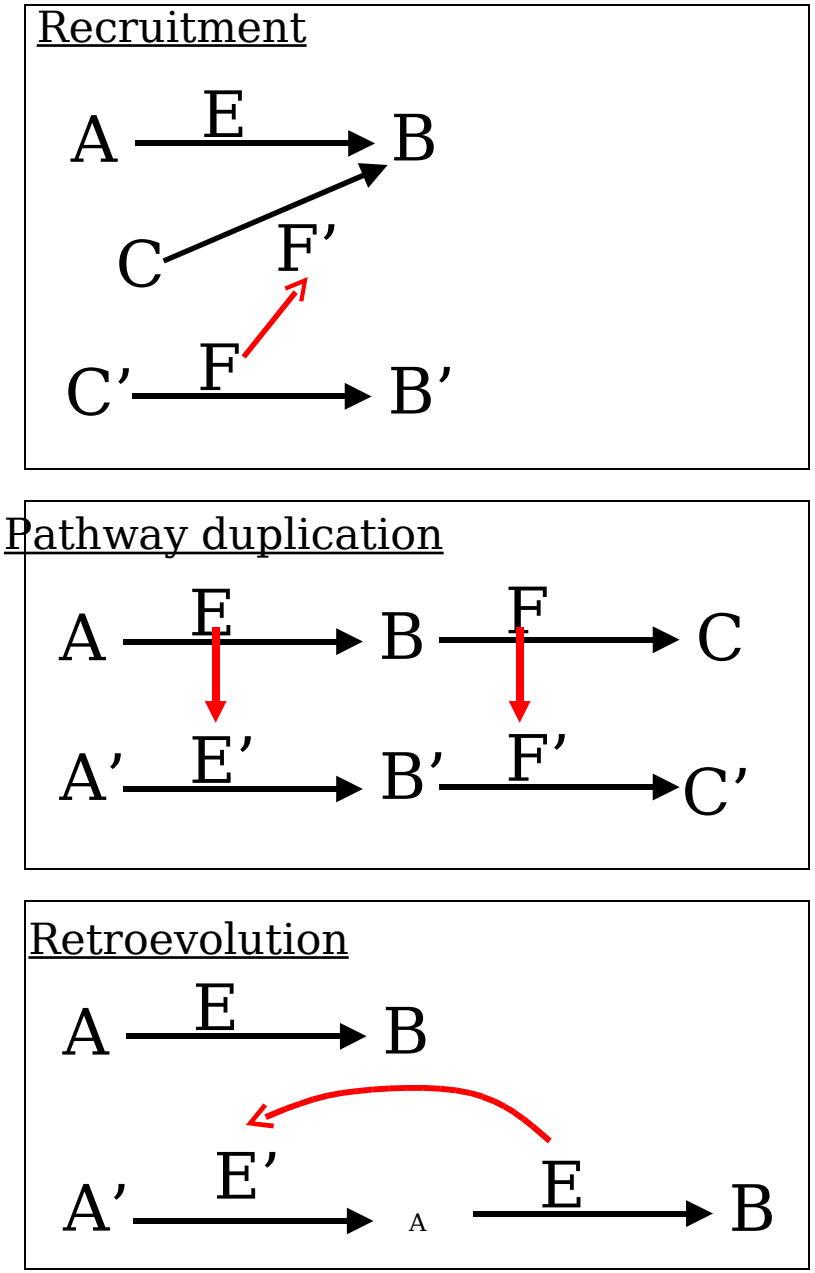
Florencio Pazos, Victor De Lorenzo & Alfonso Valencia. (2003). The organization of the Microbial Biodegradation Network from a Systems-Biology perspective. *EMBO Rep.* **4(10):**994-999.

Evolución de la Red Metabólica



Teichmann, S.A., Rison, S.C., Thornton, J.M., Riley, M., Gough, J. and Chothia, C. (2001) Small-molecule metabolism: an enzyme mosaic. *Trends Biotechnol.*, **19**, 482-486.





Teichmann, S.A., Rison, S.C., Thornton, J.M., Riley, M., Gough, J. and Chothia, C. (2001) Small-molecule metabolism: an enzyme mosaic. *Trends Biotechnol.*, **19**, 482-486.

Principales bases de datos con información metabólica

KEGG – *Metabolismo en general*

http://www.genome.jp/kegg/

Kanehisa, M., Goto, S., Kawashima, S., Okuno, Y. and Hattori, M. (2004) The KEGG resource for deciphering the genome. *Nucleic Acids Res*, **32**, D277-280.

EcoCyc – *Metabolismo de E coli*

http://ecocyc.org

Keseler, I. M., Collado-Vides, J., Gama-Castro, S., Ingraham, J., Paley, S., Paulsen, I. T., Peralta-Gil, M. and Karp, P. D. (2005). EcoCyc: a comprehensive database resource for Escherichia coli. *Nucleic Acids Res.* **33**: D334-337

ENDA – *Centrada en enzimas (incluyendo parámetros cinéticos y termodinámicos)*

p://www.brenda.uni-koeln.de/

Schomburg, I., Chang, A., Ebeling, C., Gremse, M., Heldt, C., Huhn, G. and Schomburg, D. (2004) BRENDA, the enzyme database: updates and major new developments. *Nucleic Acids Res*, **32**, D431-433.

MBBD – *Centrada en biodegradación y bioremediación*

ttp://umbbd.msi.umn.edu/

Ellis, L.B., Hou, B.K., Kang, W. and Wackett, L.P. (2003) The University of Minnesota Biocatalysis/Biodegradation Database: post-genomic data mining. *Nucleic Acids Res*, **31**, 262-265.

---

lorencio Pazos Cabaleiro
omputational Systems Biology Group (CNB-CSIC)
azos@cnb.csic.es
fl. 91.5854669