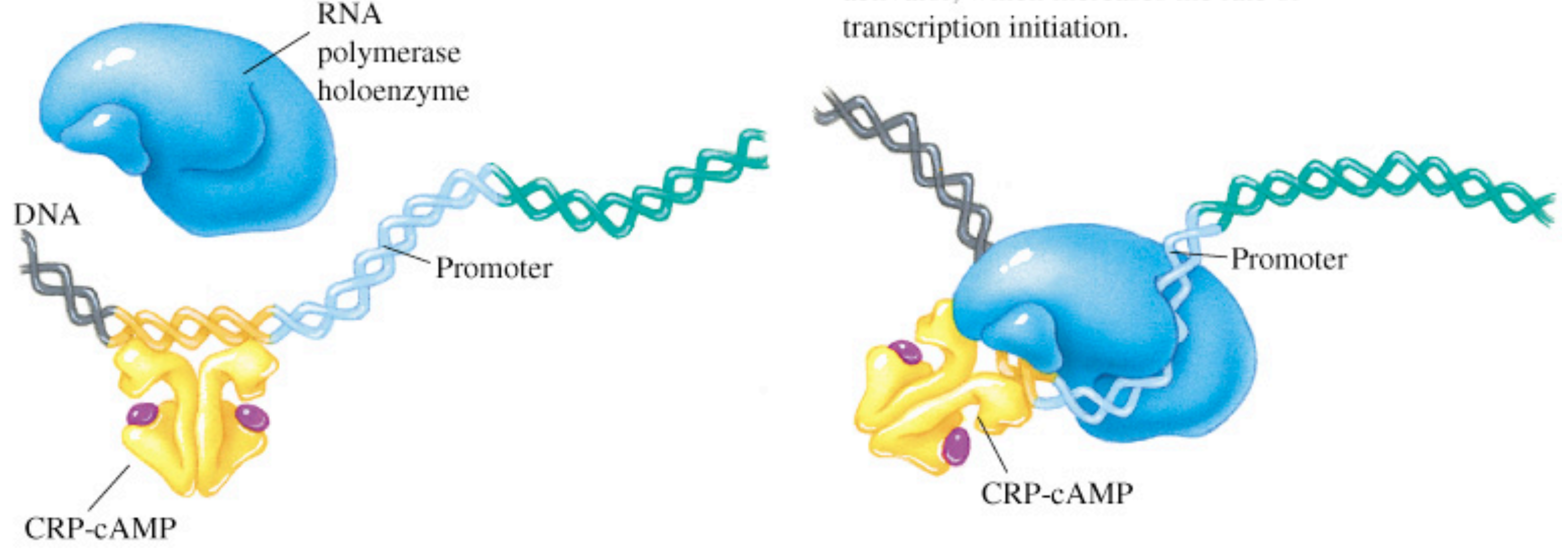
Redes de Regulación Génica

Ildefonso Cases

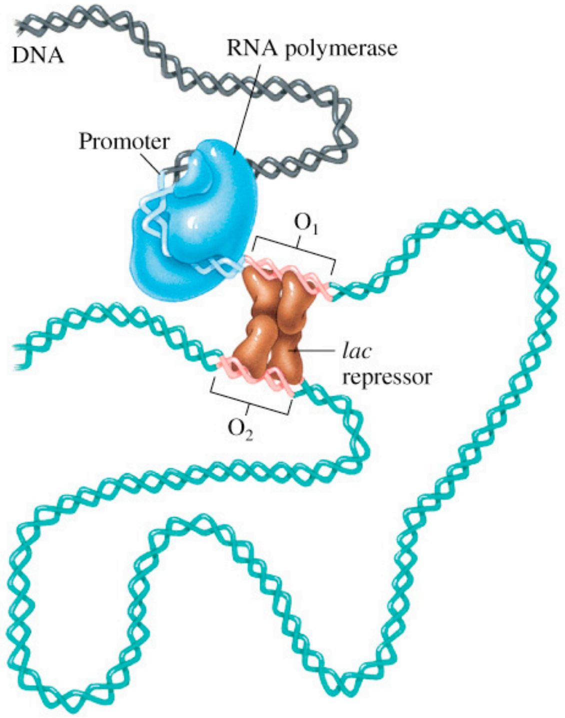
Regulación de la Transcripción

- Resultado de la interacción entre proteínas y DNA.
- El conjunto de proteínas que se unan a su región promotora (directa o indirectamente) va a determinar la expresión de un gen:
  - En que tejidos
  - En que momento del desarrollo
  - Bajo que condiciones ambientales
  - etc.

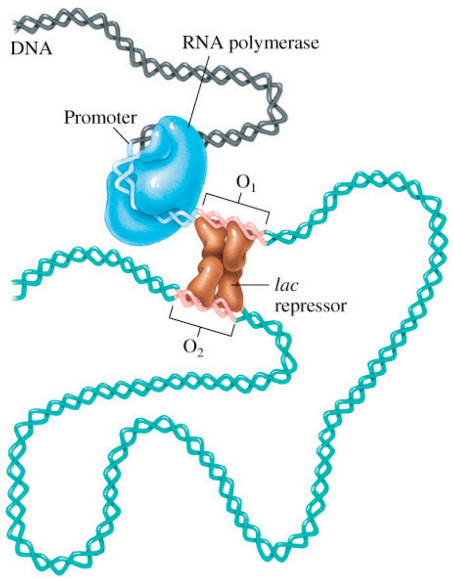


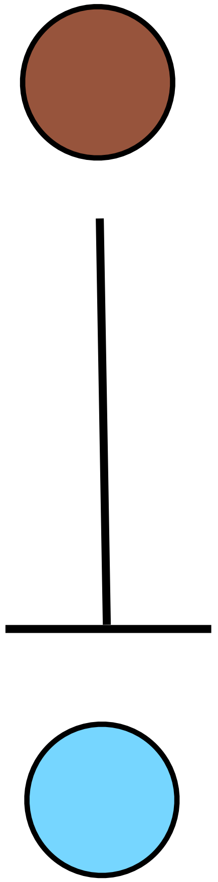
(a) CRP-cAMP binds to a site near the promoter.

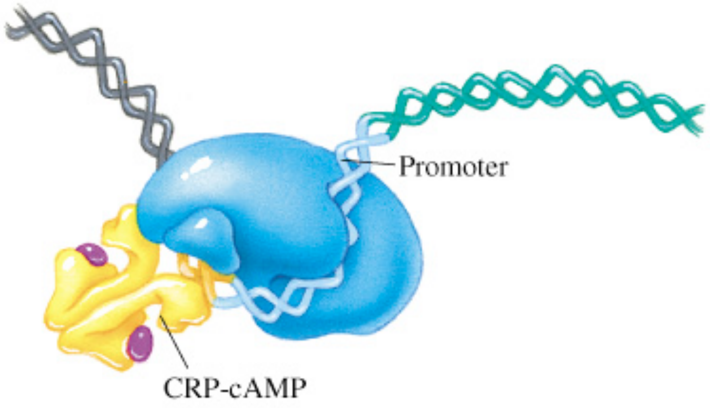
(b) RNA polymerase holoenzyme binds to the promoter and also contacts the bound activator, which increases the rate of transcription initiation.

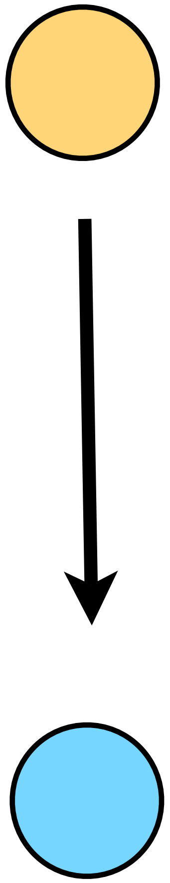


Redes de Regulación

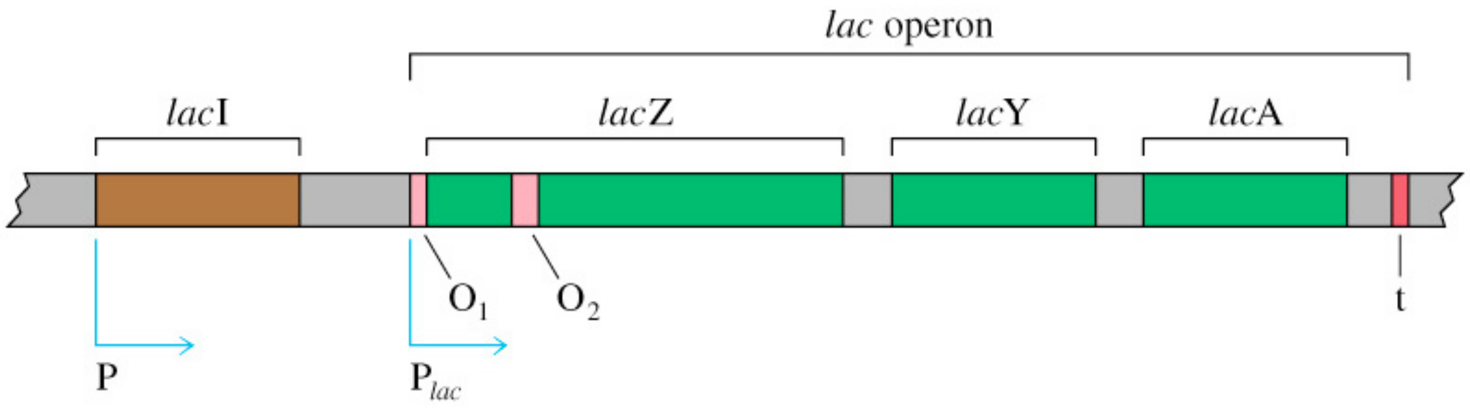


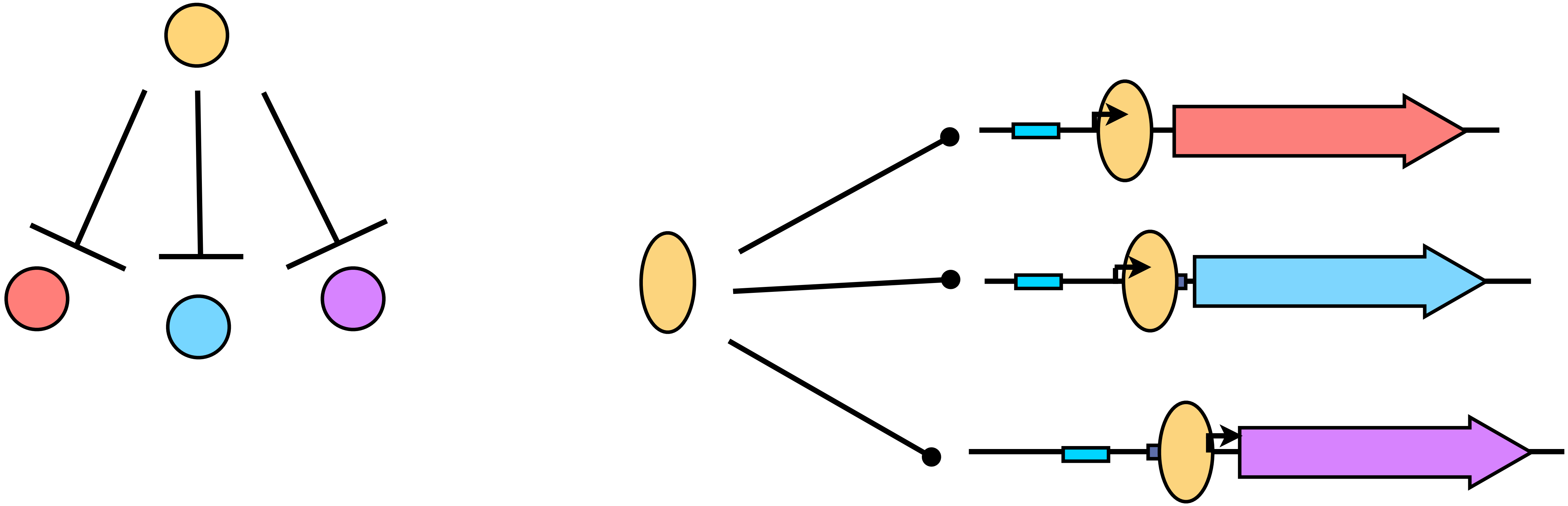




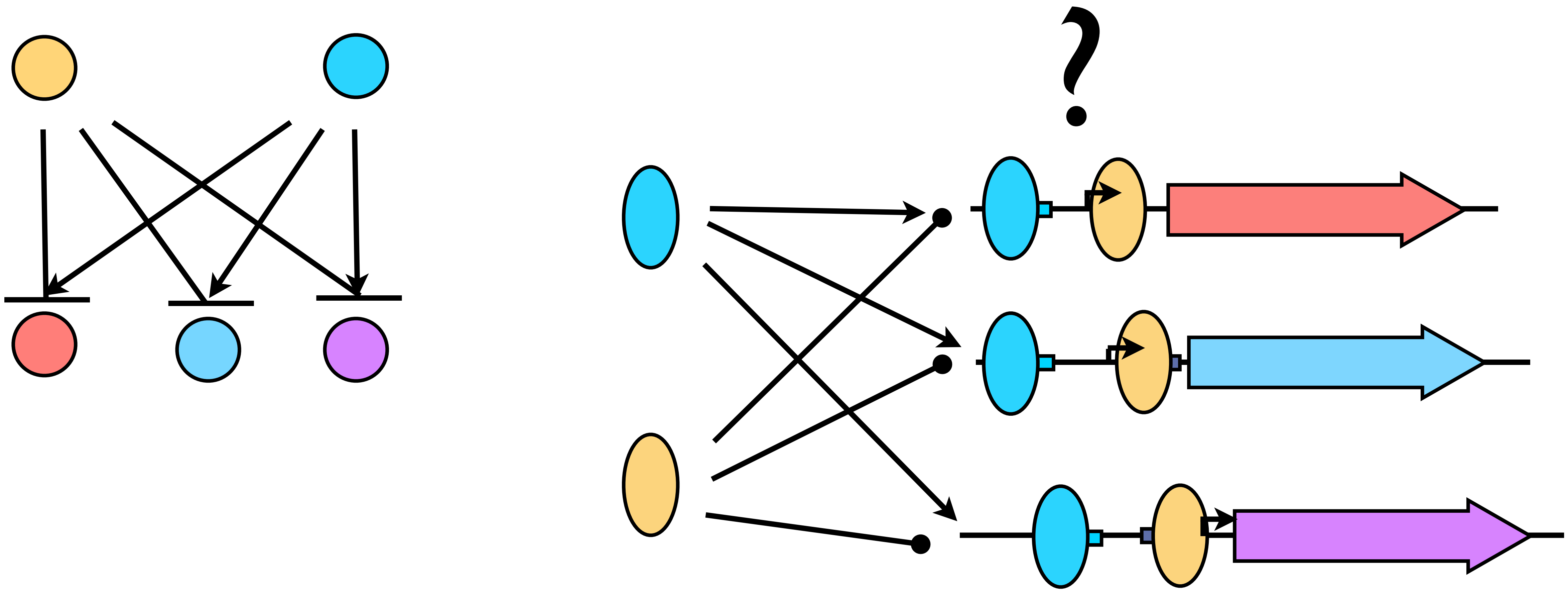


Redes de Regulación

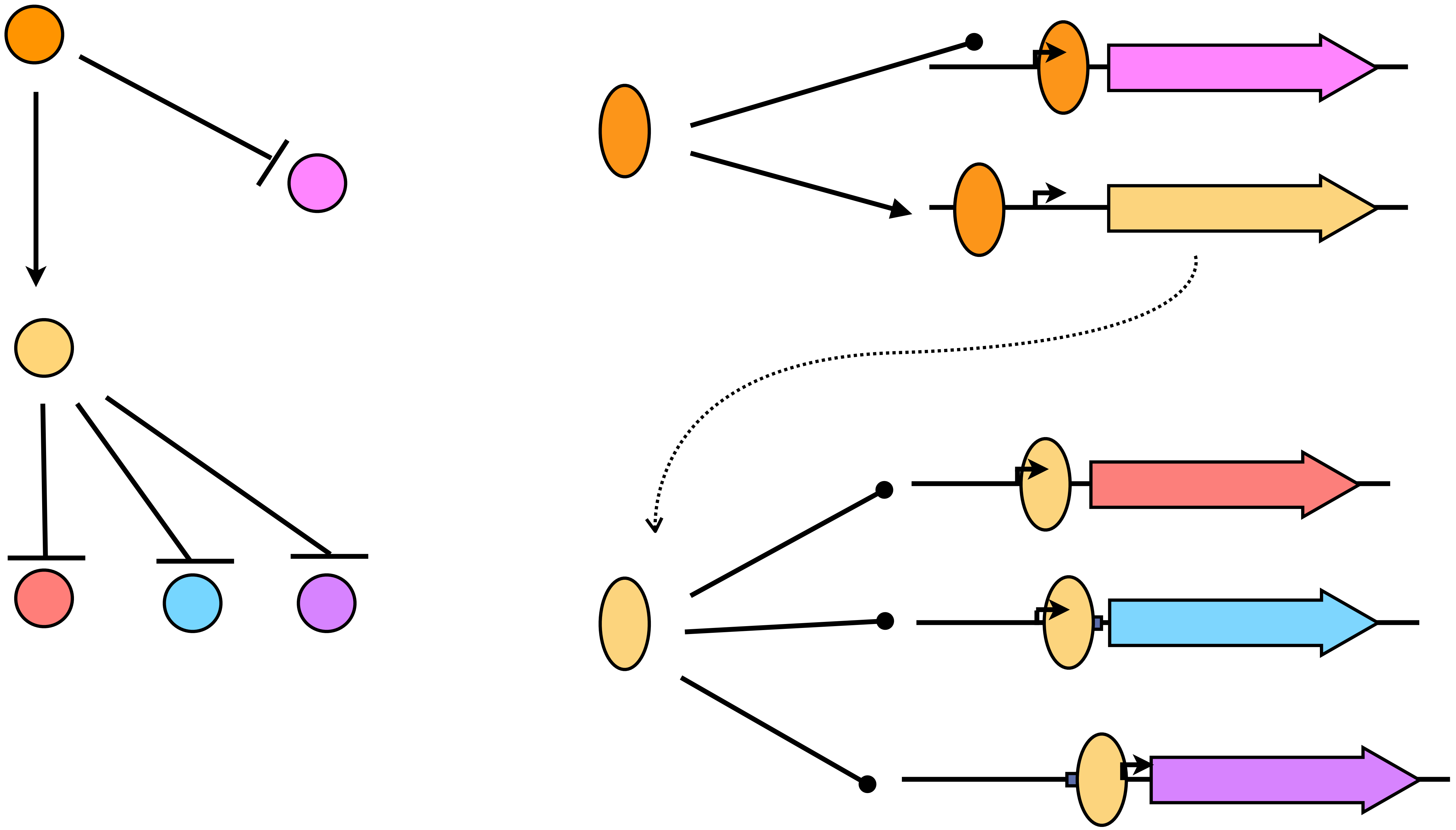


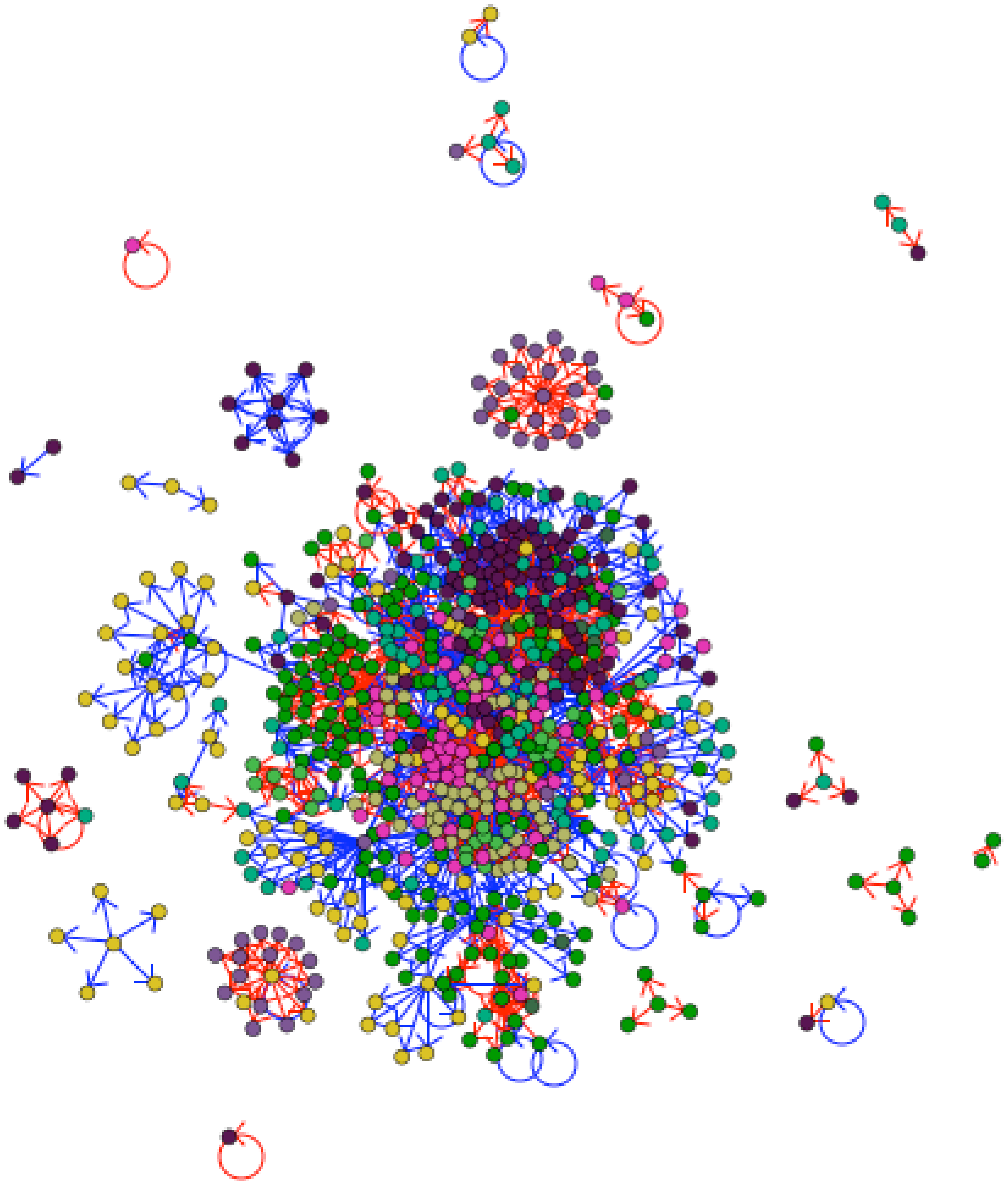


Redes de Regulación



Redes de Regulación

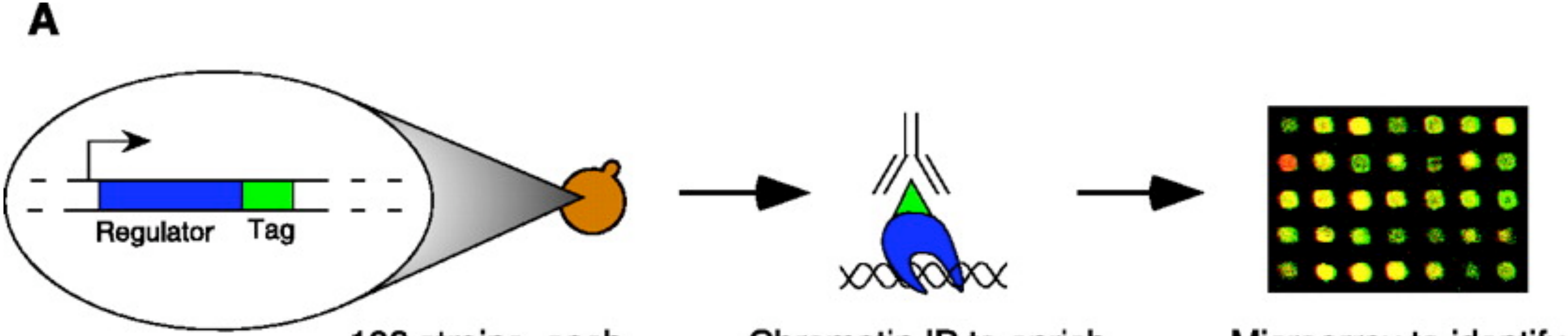




Datos

- Chip-on-Chip
- STAGE/SABE
- DNA-arrays
- Predicción

Chip-On-Chip



106 strains, each with a tagged regulator

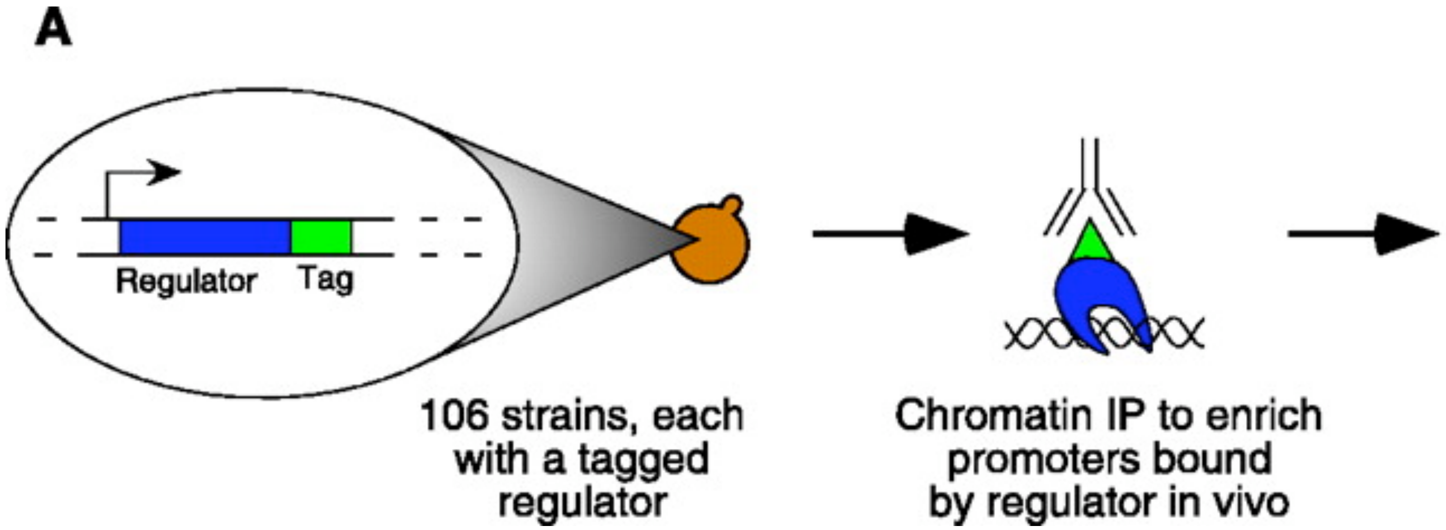
Chromatin IP to enrich promoters bound by regulator in vivo

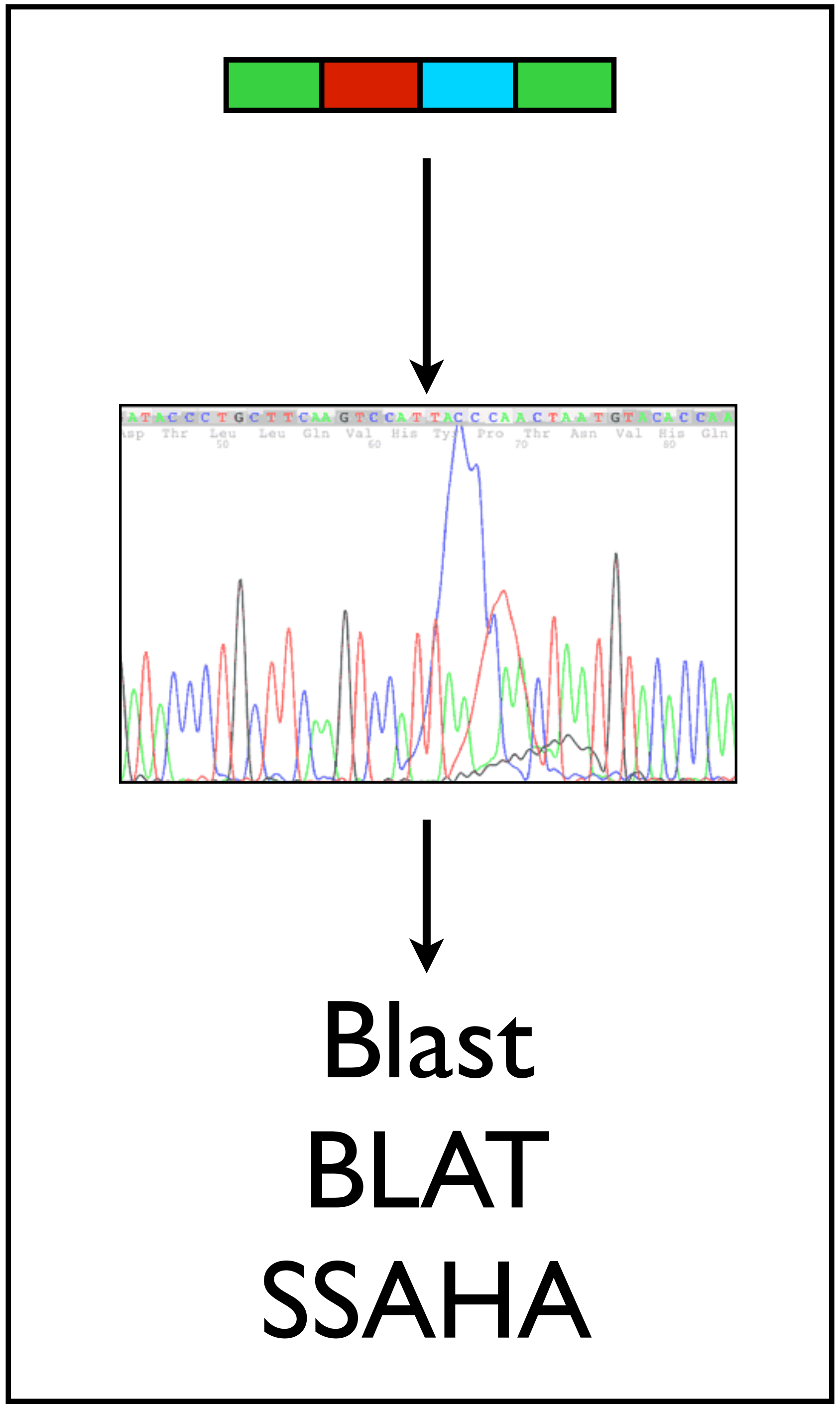
Microarray to identify promoters bound by regulator in vivo

Chip-On-Chip II

- PCR Arrays
  - baja resolucion
- Oligo Array
  - muy caros
- Ambos cubren solo regiones pre-establecidas y cercanas al gen
- "Tile Arrays" muy muy caros!!!

STAGE/SABE





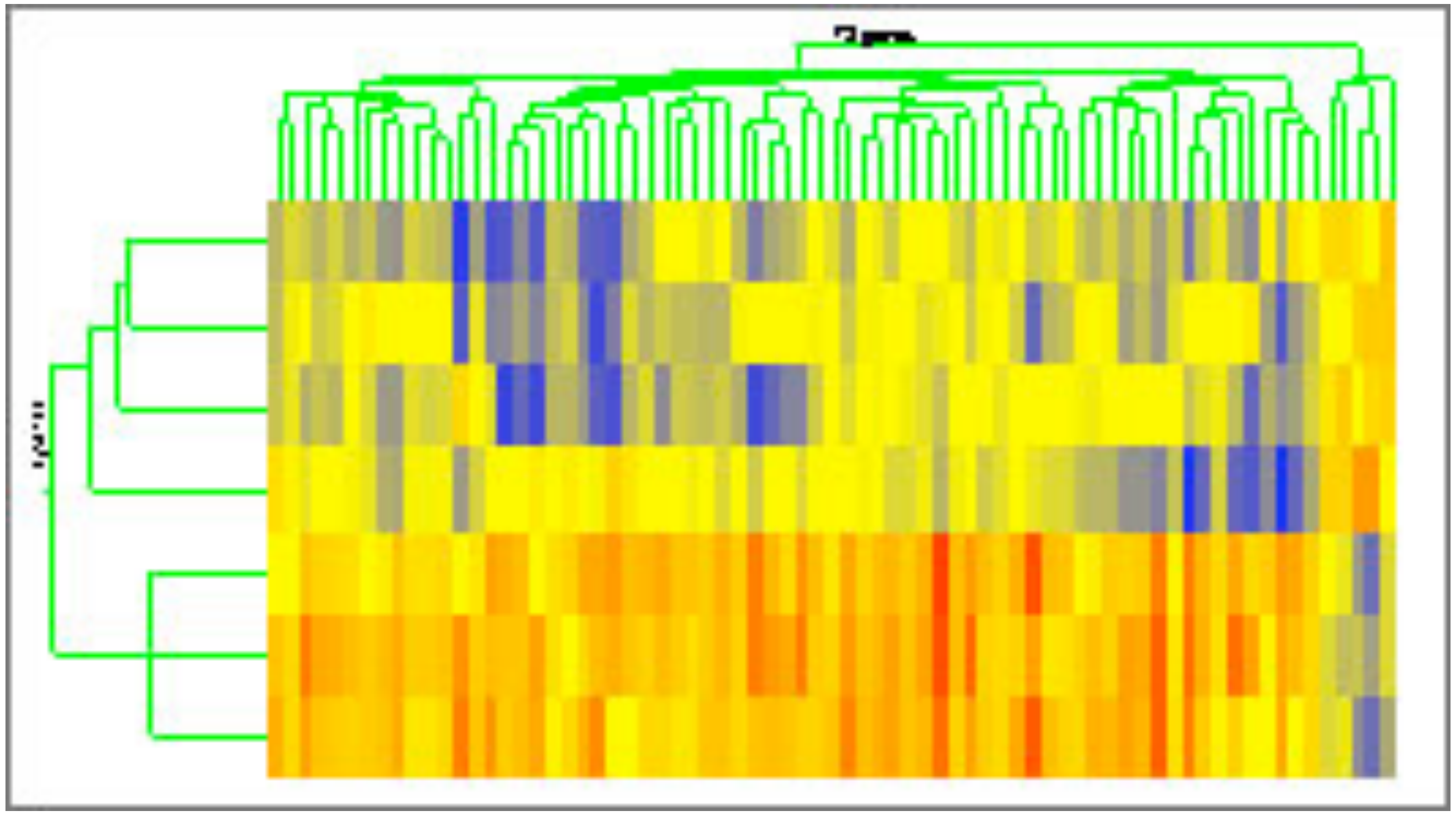
Blast
BLAT
SSAHA

Caro, requiere mucha secuenciación para encontrar todos los sitios.

Problemas

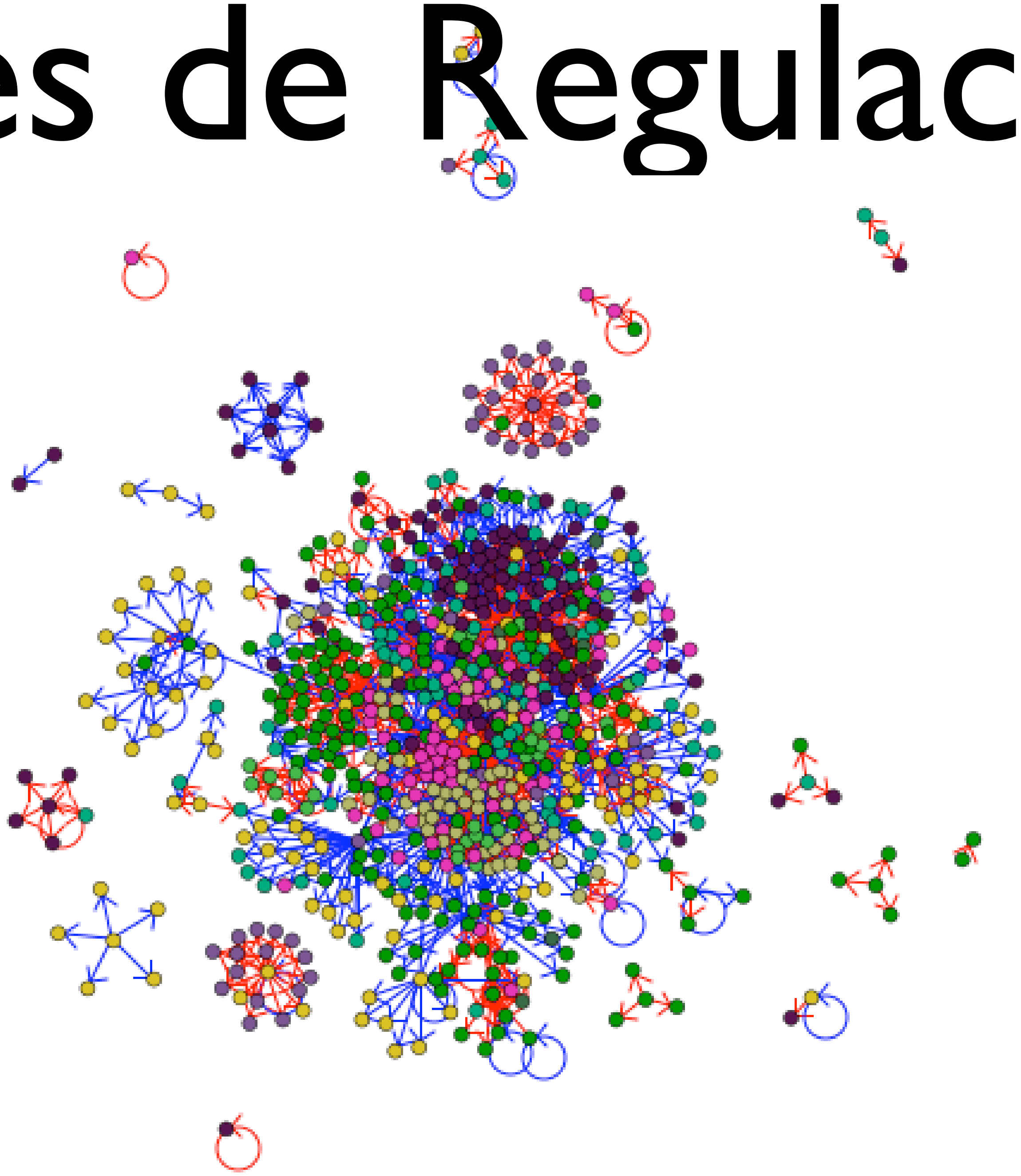
- Solo encontramos los sitios que están unidos en las condiciones del experimento
  - positivo: contexto
  - negativo: no podemos cubrir todas las condiciones
- Sabemos que los TF se unen, pero no si estan activos, ni como actúan sobre el gen: ¿reprimiendo o activando?

DNA arrays



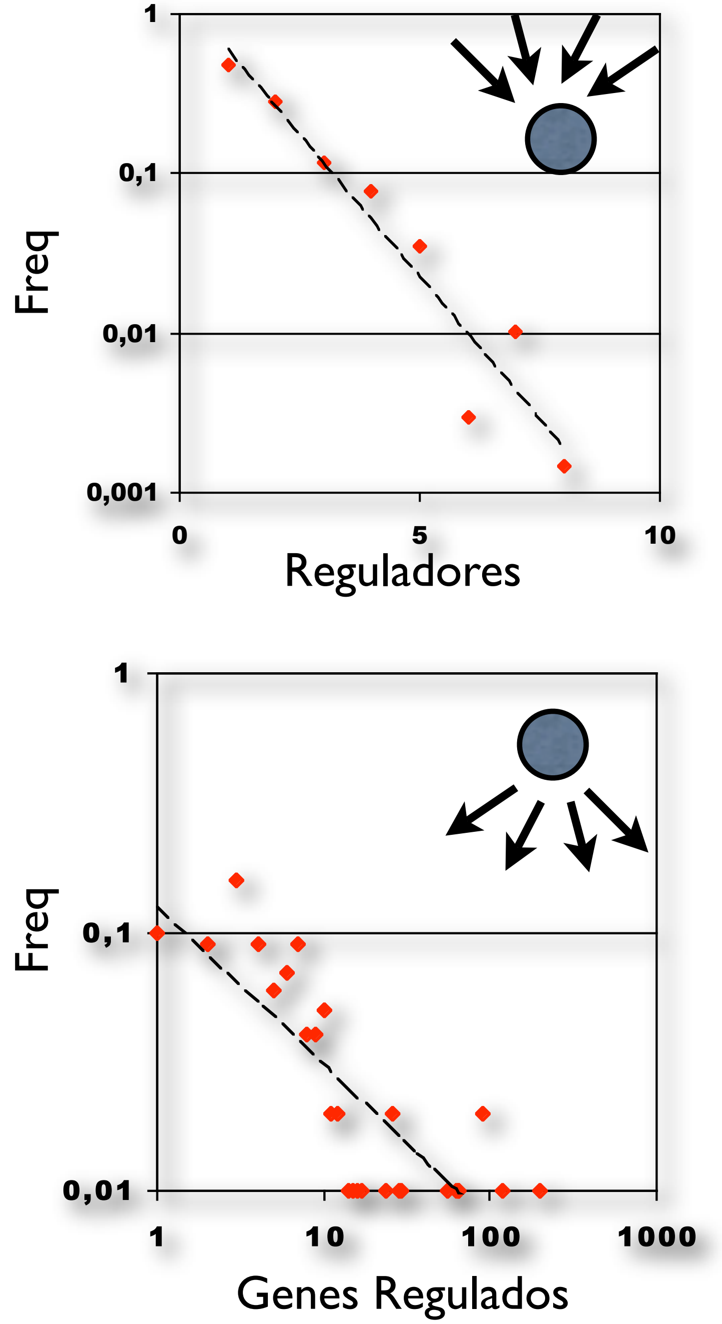
- Relaciones entre la expresión de TF y otros genes:
  - redes bayesianas
- Imposible aplicar a grandes set de genes, solo sub-sistemas
- ambiguos: dan muchas soluciones
- Mejoran con informacion adicional: Chip-on-Chip

Propiedades de las redes de Regulación

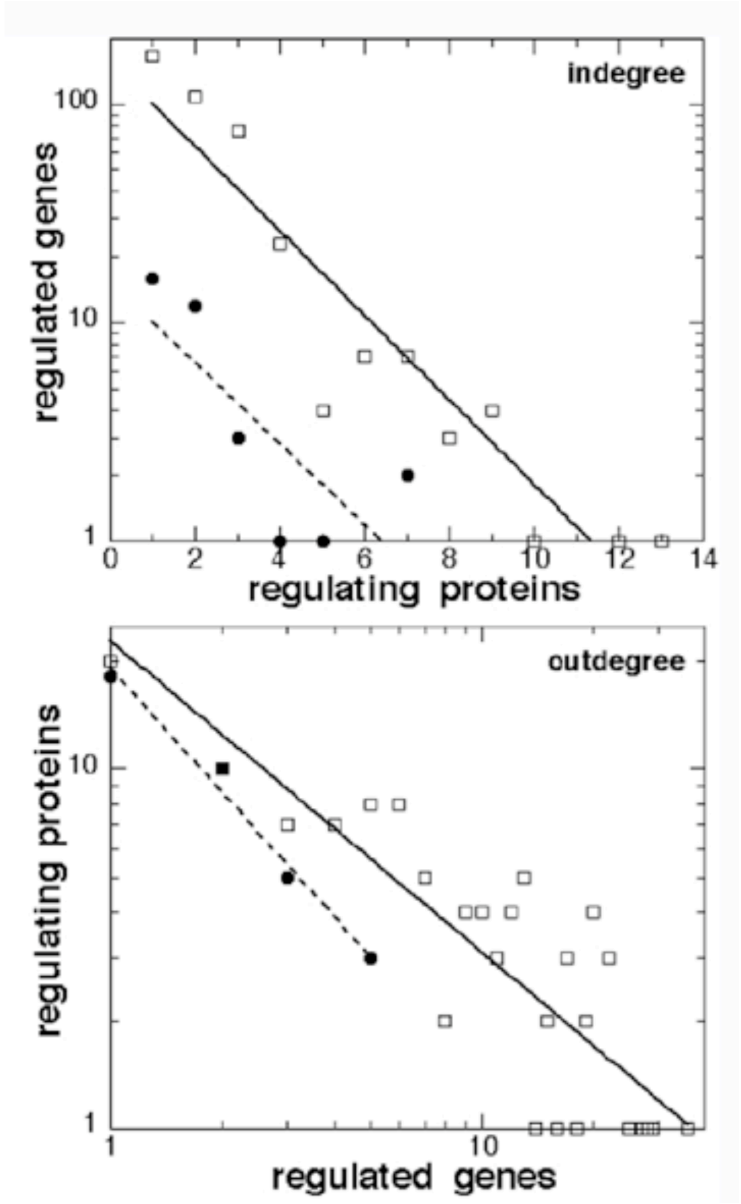


Propiedades Generales

E. coli



S. cerevisiae



Guelzim et al. 2002 Nature Genet. 31:60-63

Distribución exponencial

$$p(k)=Ce^{k}$$

Distribución "Ley de Potencias"

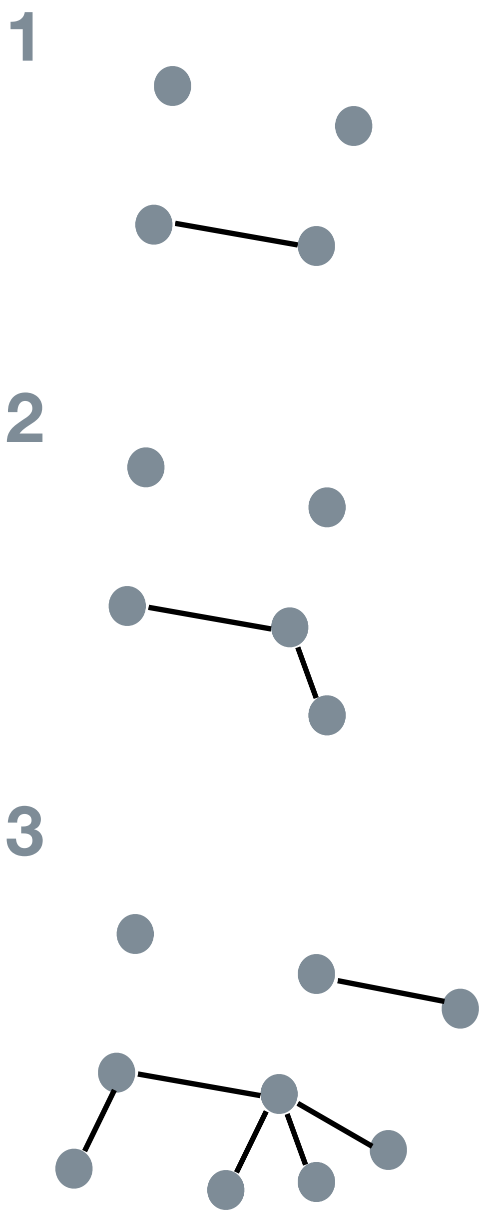
$$p(k)=Ck^{-g}$$

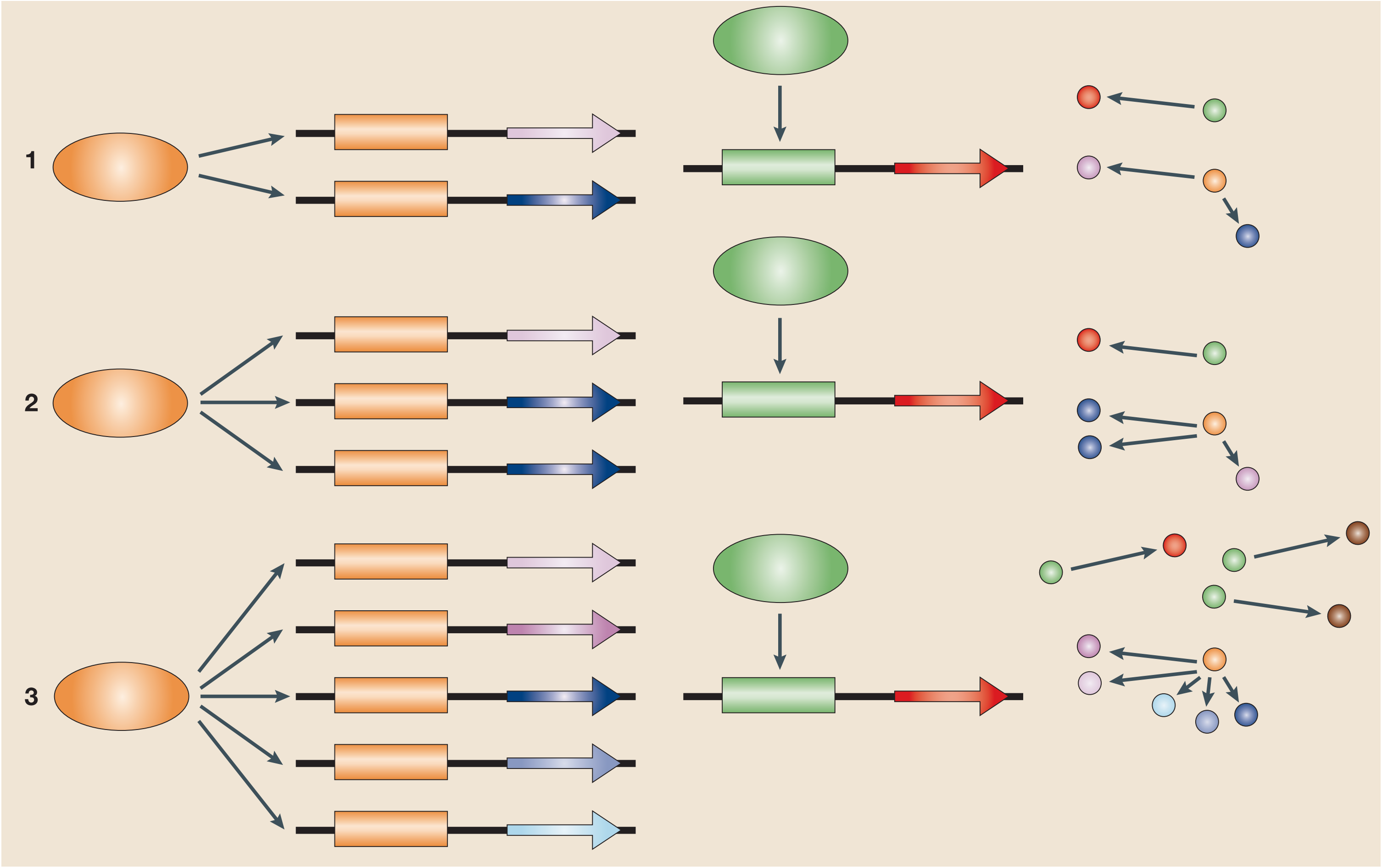
"libre de escala"

robustez

caminos cortos (integracion de señal)

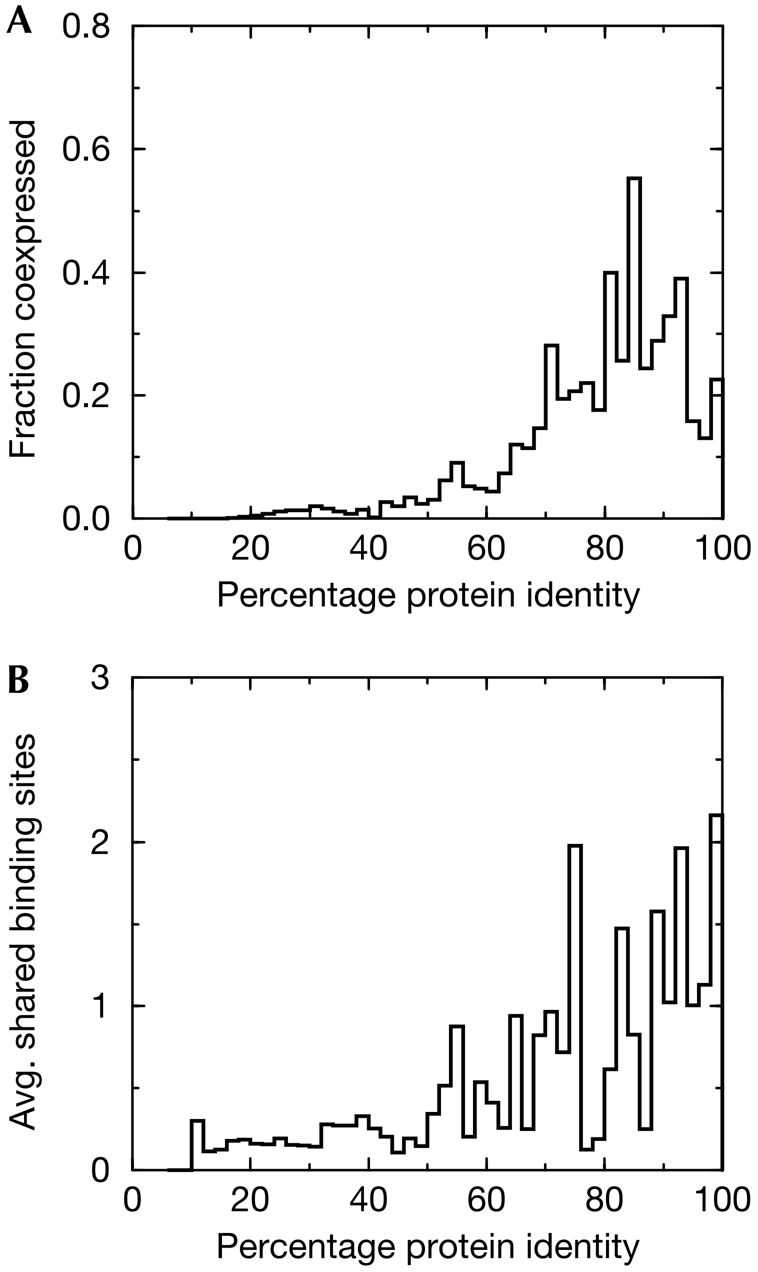
Evolución de Redes



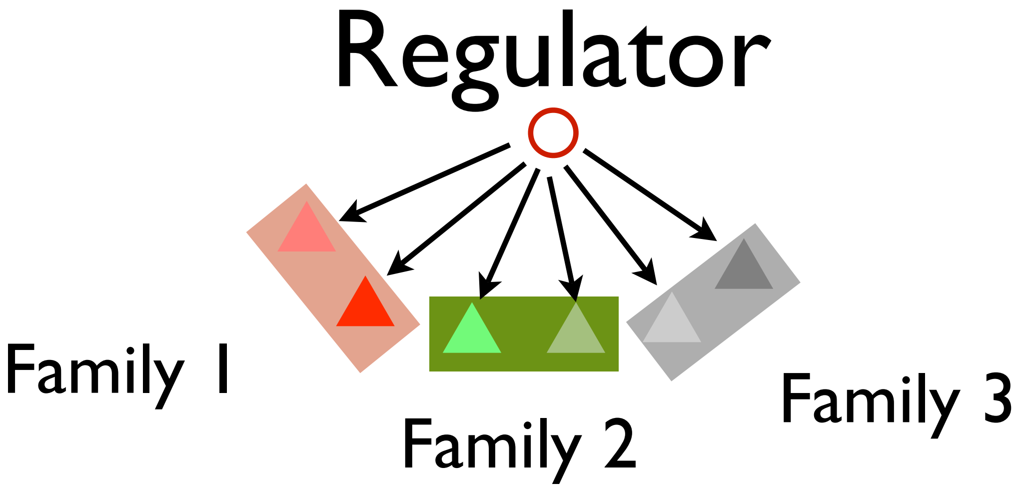


Cases & de Lorenzo 2005 Nature Rev Microb 3:105-118

Duplicación y Evolución



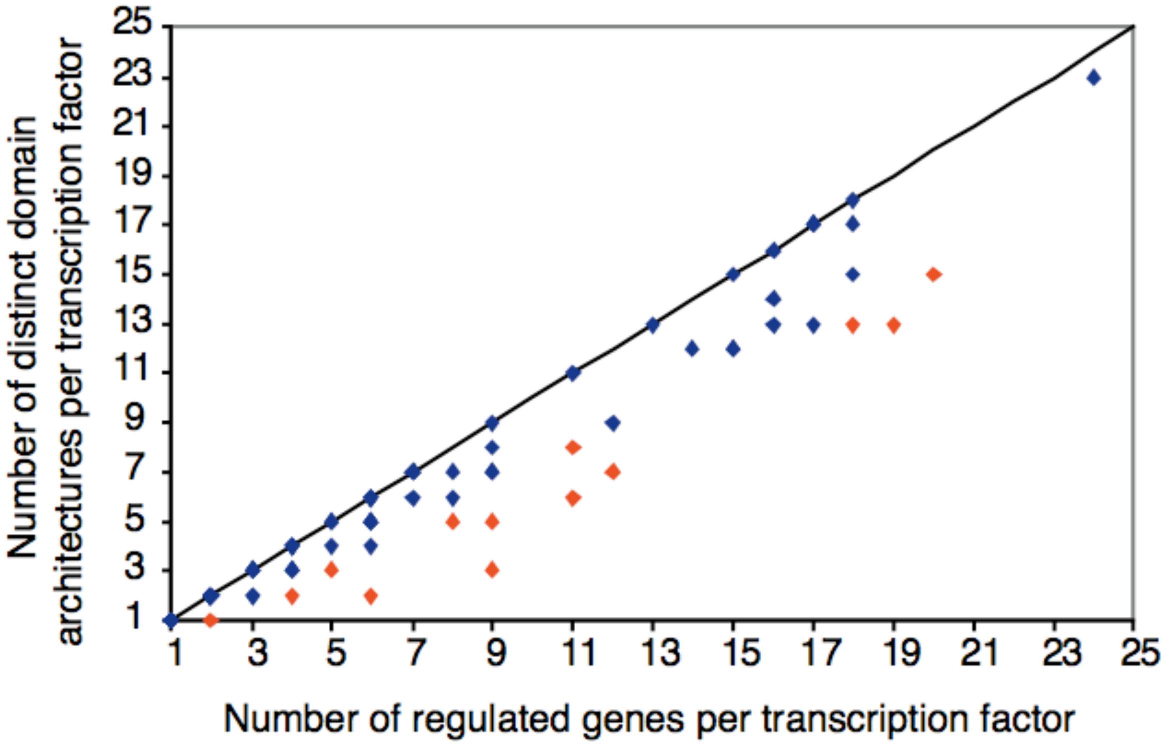
van Noort *et al.*, 2004 EMBO Rep 5(3):280-4


Regulator

Family 1

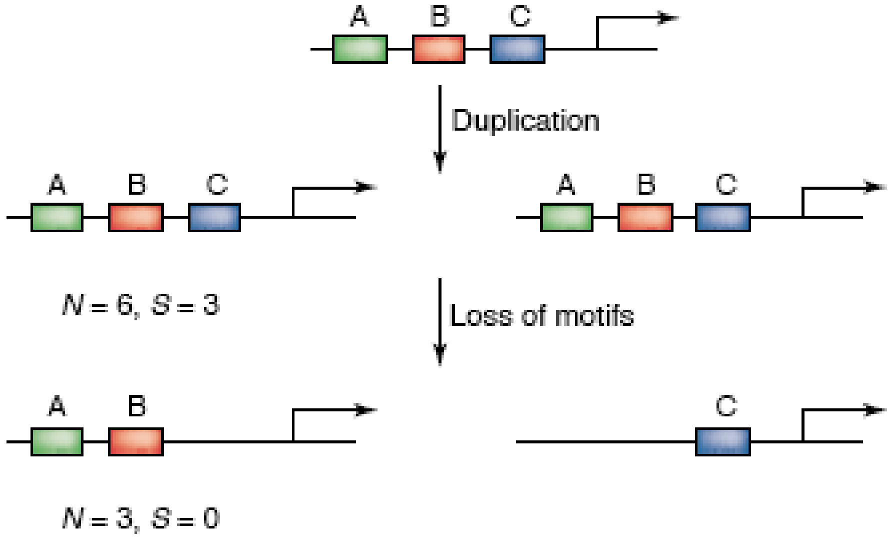
Family 2

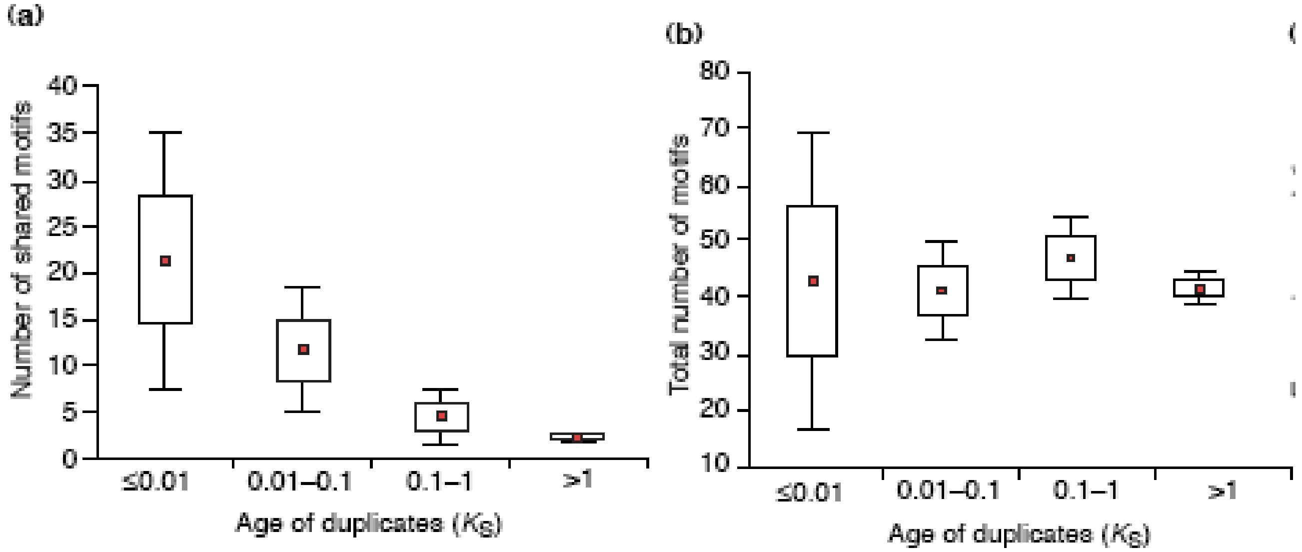
Family 3



Tecihmann & Babu, 2004 Nature Genetics 36(5):492-6

Duplicación de TFBS

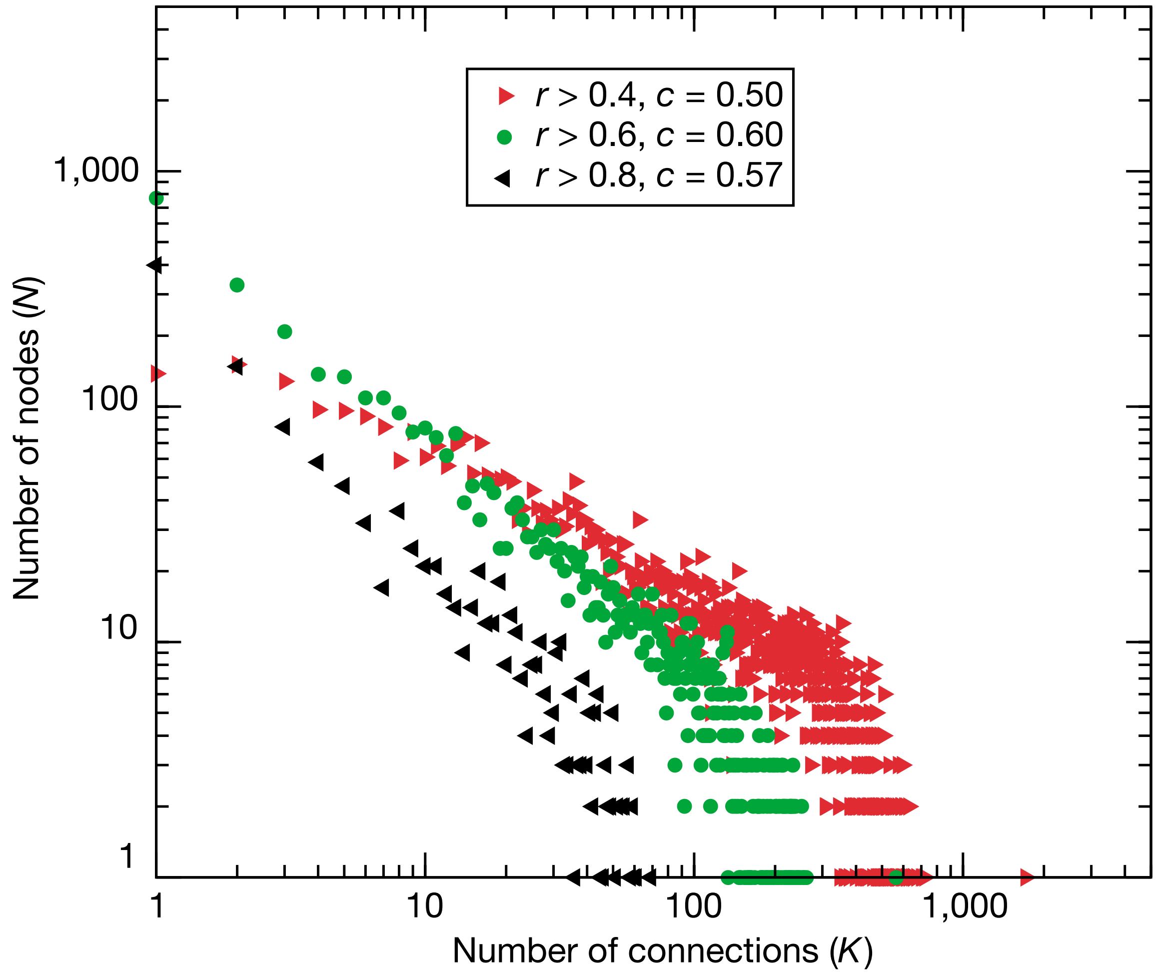




Papp et al,2003. Trends Genet 19:417

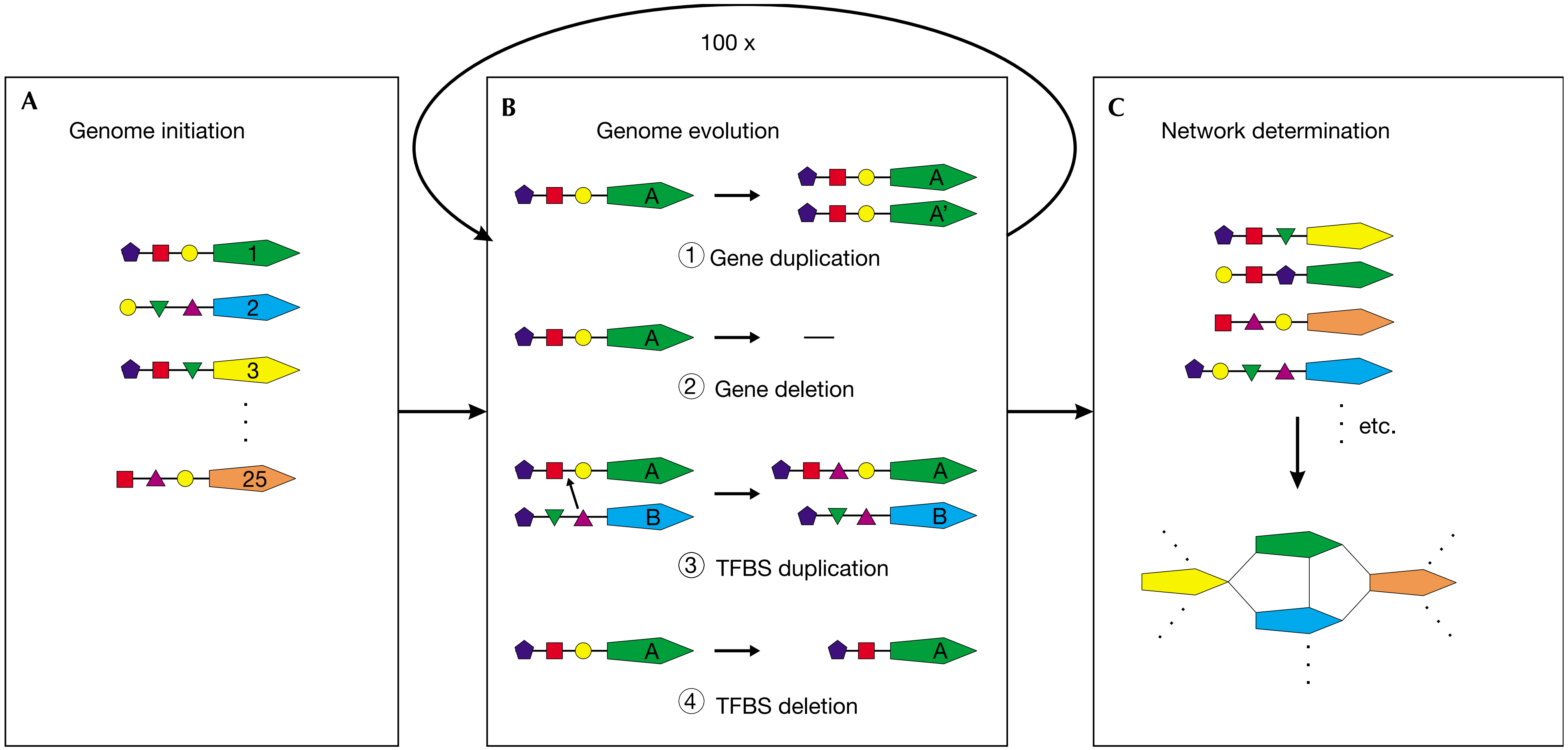
Red de Co-regulación

- gamma≈-1
- c=0.6
- libre de escala
- “mundo pequeño”



van Noort *et al.*, 2004 EMBO Rep 5(3):280-4

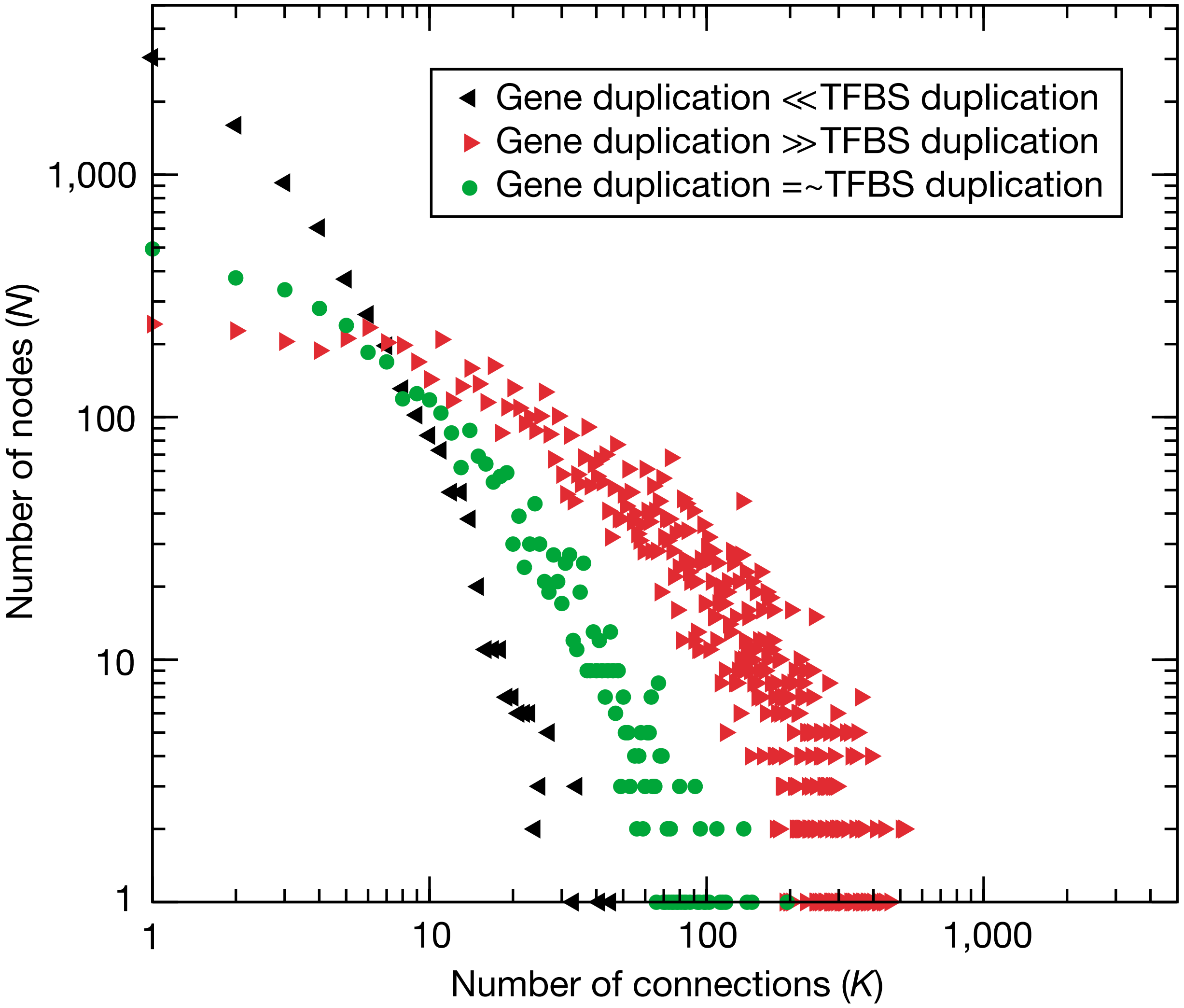
Simulando la evolución



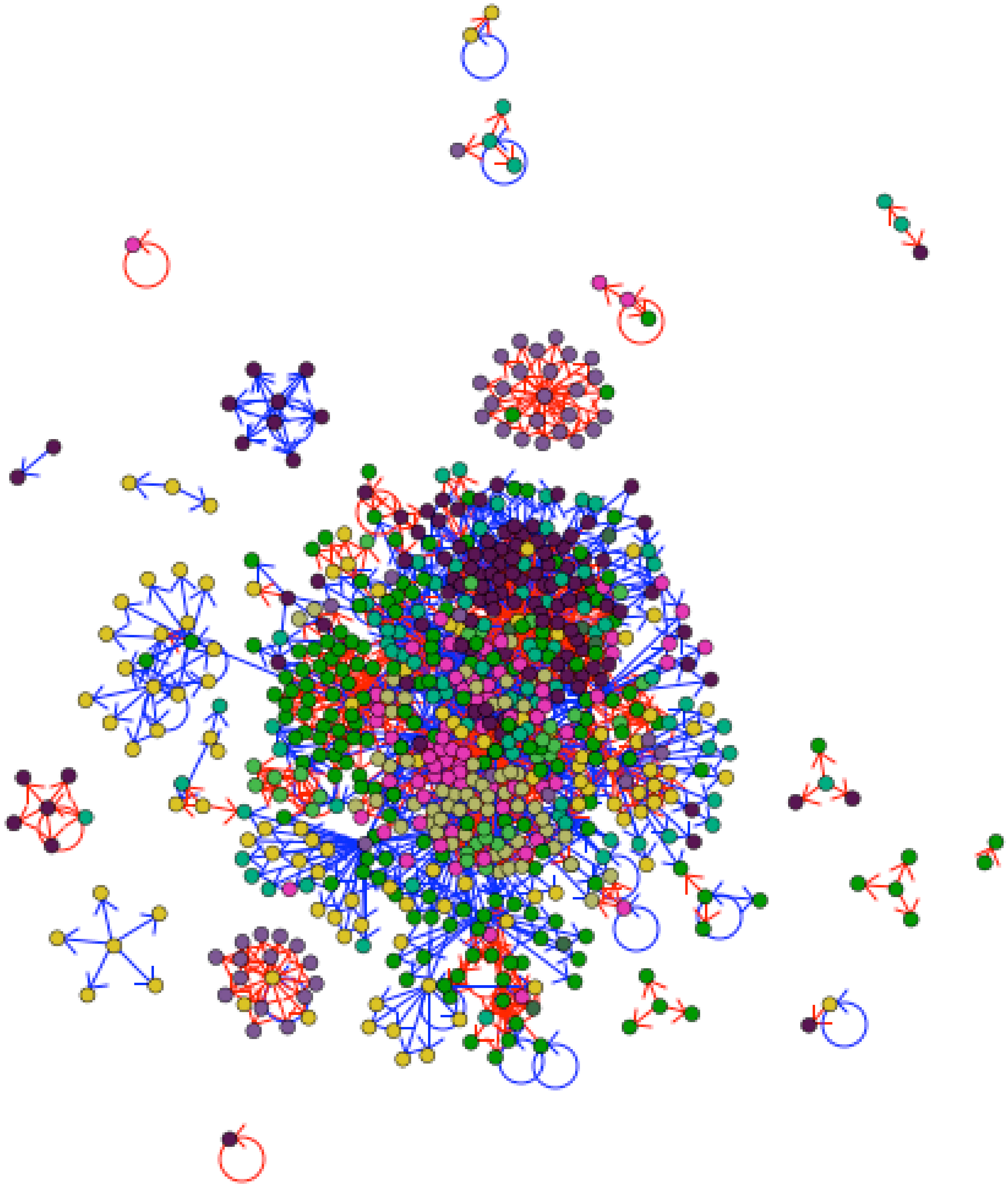
van Noort *et al.*, 2004 EMBO Rep 5(3):280-4

Simulando la Evolución II

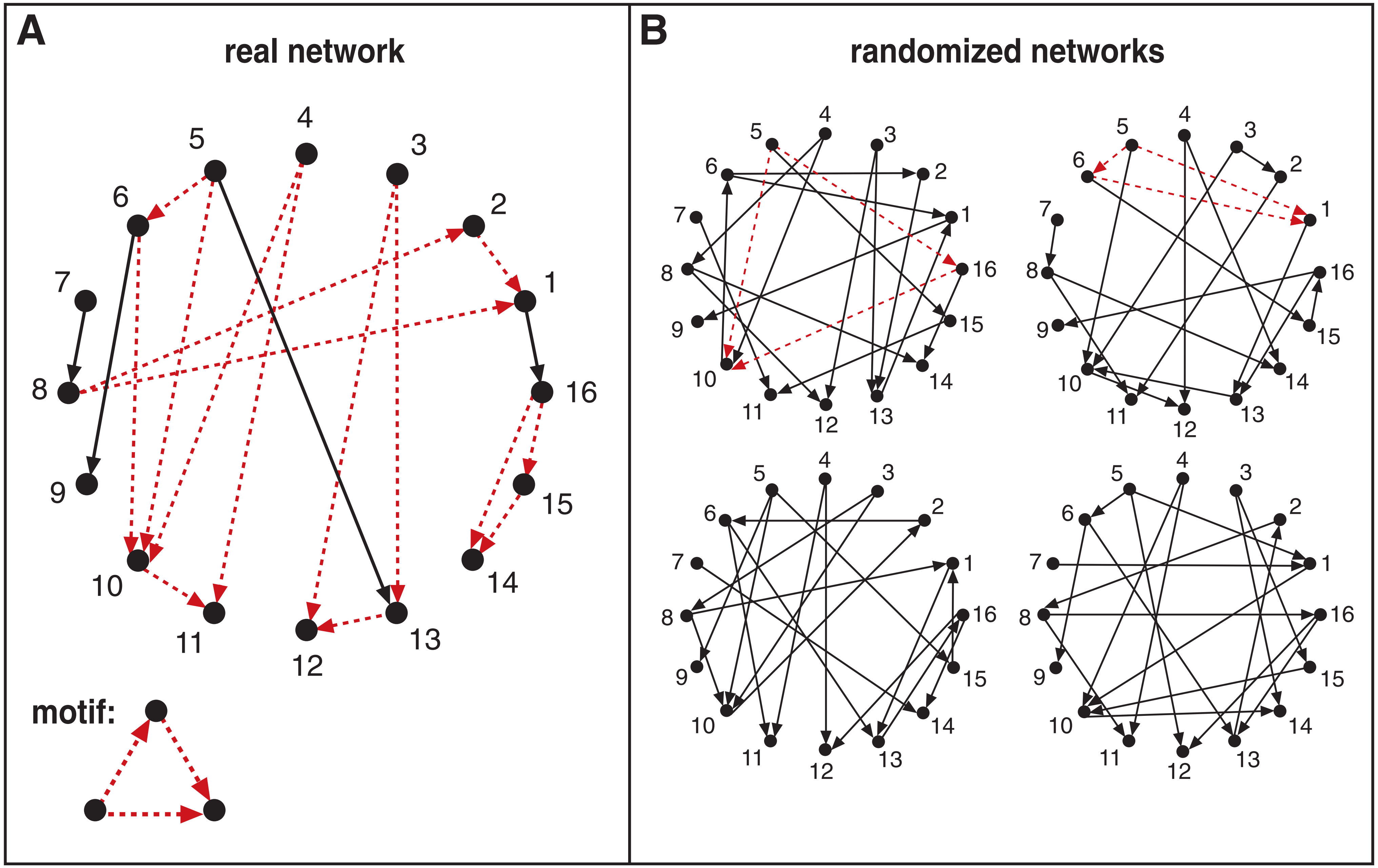
En ausencia de selección pueden aparecer redes de propiedades similares a las reales



van Noort *et al.*, 2004 EMBO Rep 5(3):280-4



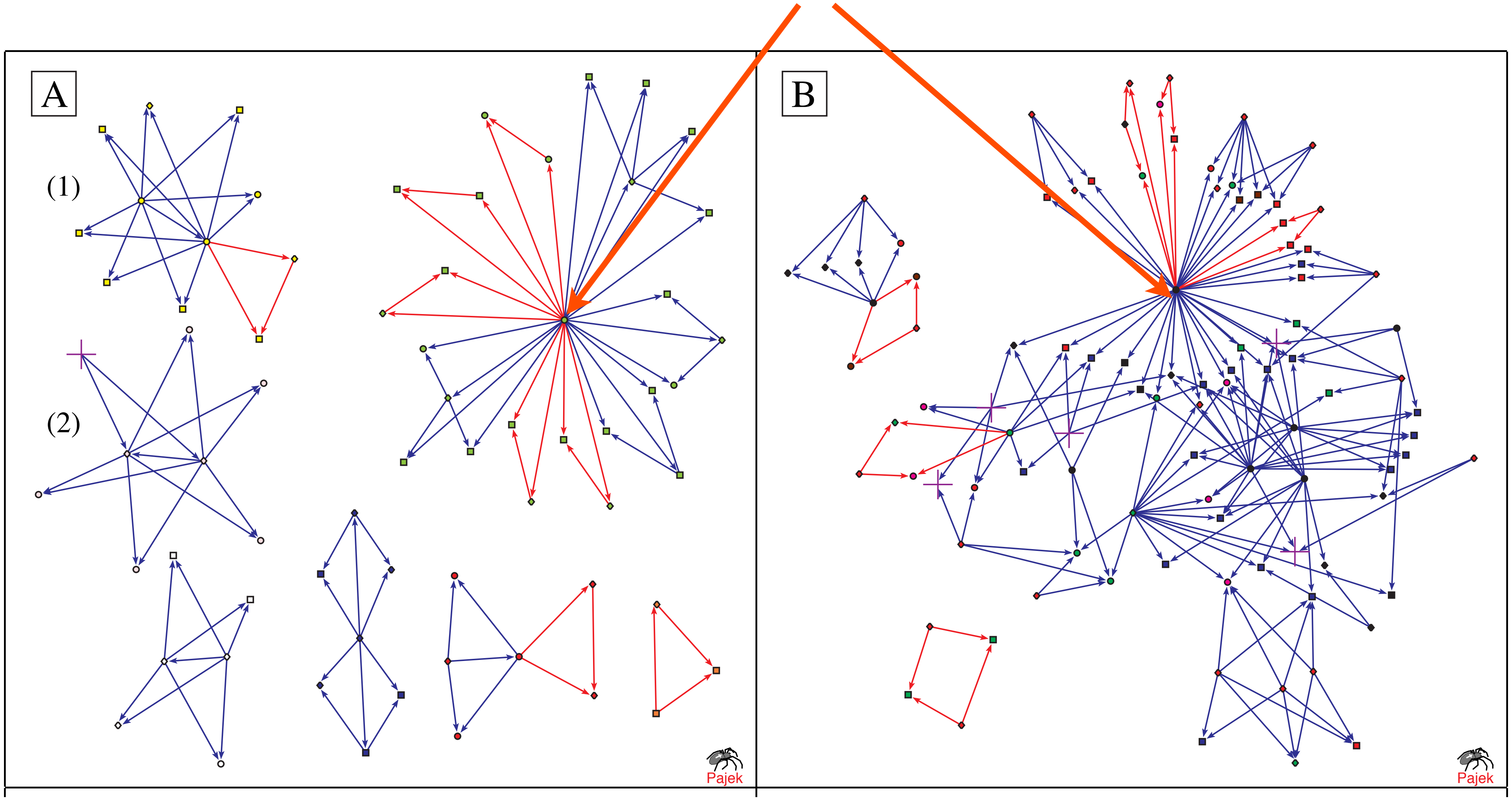
Motivos



Milo *et al*,2002. Science 298:824

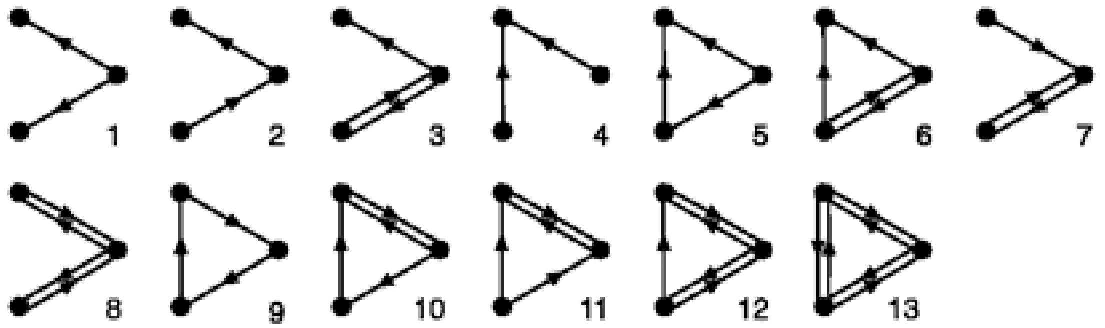
Motivos Solapantes

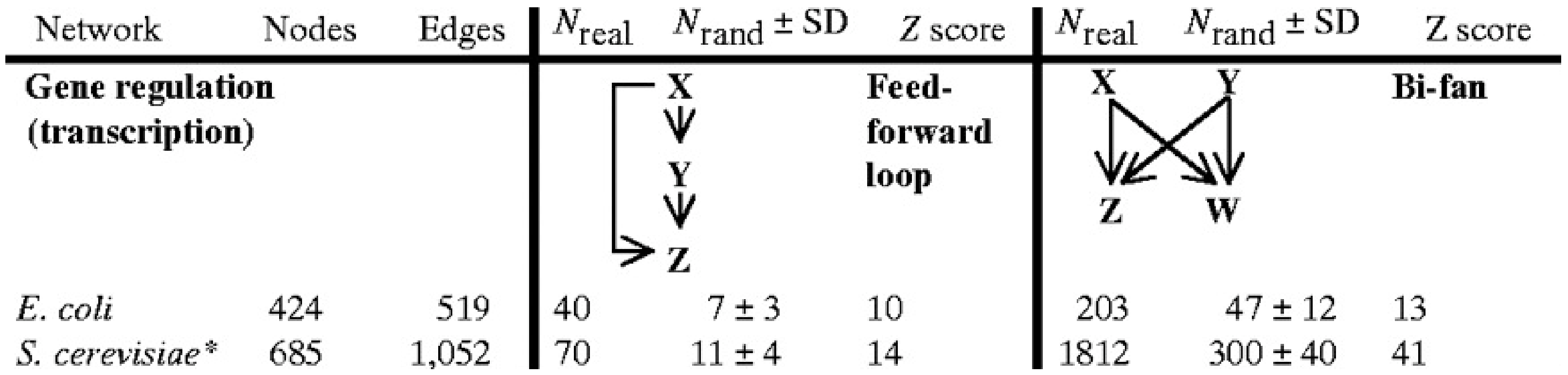
Hubs



Dobrin *et al*,2004. BMC Bioiformatics 5:10

Motivos en Redes de Regulación

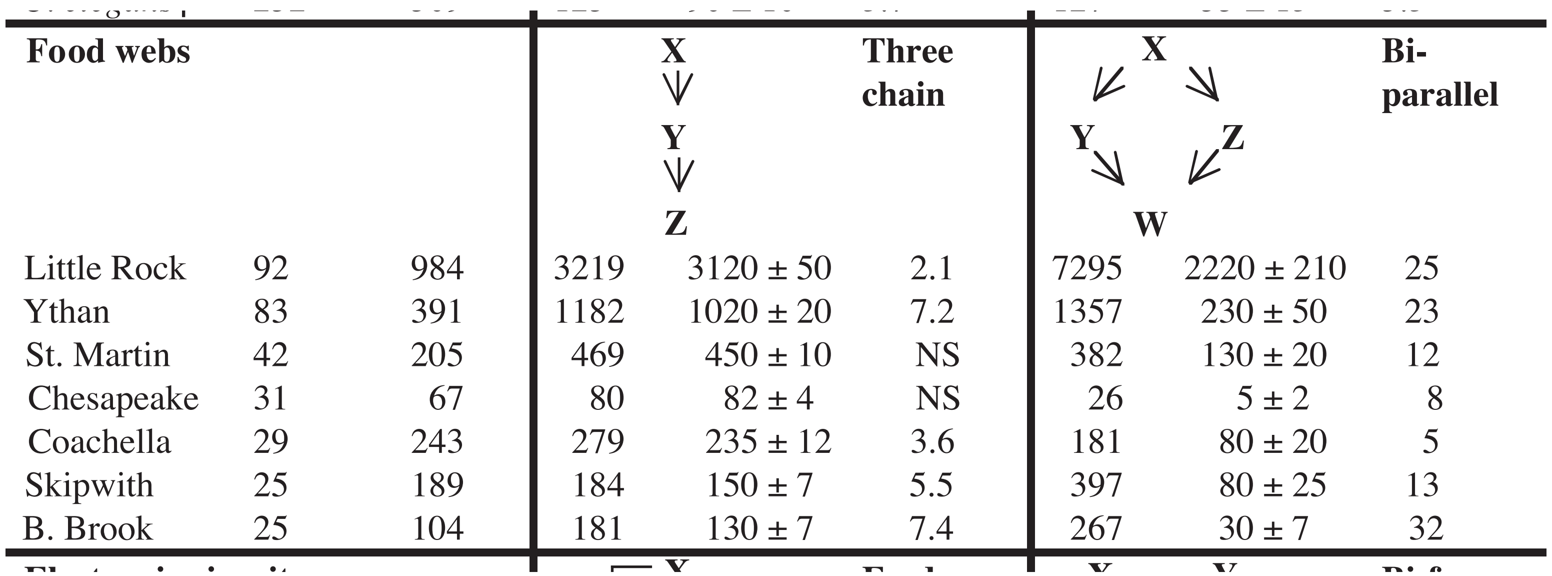




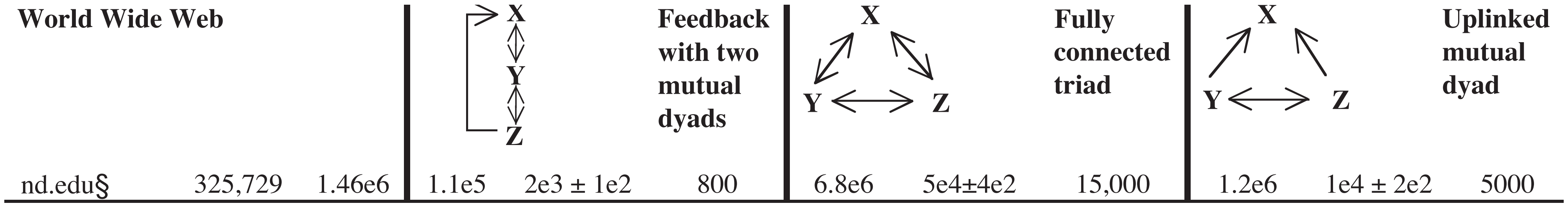
Network	Nodes	Edges	N_{real}	N_{rand} ± SD	Z score	N_{real}	N_{rand} ± SD	Z score
Gene regulation (transcription)				**X** **Y** **Z**	**Feed-forward loop**	**X** **Y** **Z** **W**		**Bi-fan**
E. coli	424	519	40	7 ± 3	10	203	47 ± 12	13
*S. cerevisiae**	685	1,052	70	11 ± 4	14	1812	300 ± 40	41

Milo *et al*,2002. Science 298:824

Motivos en Redes de Regulación II



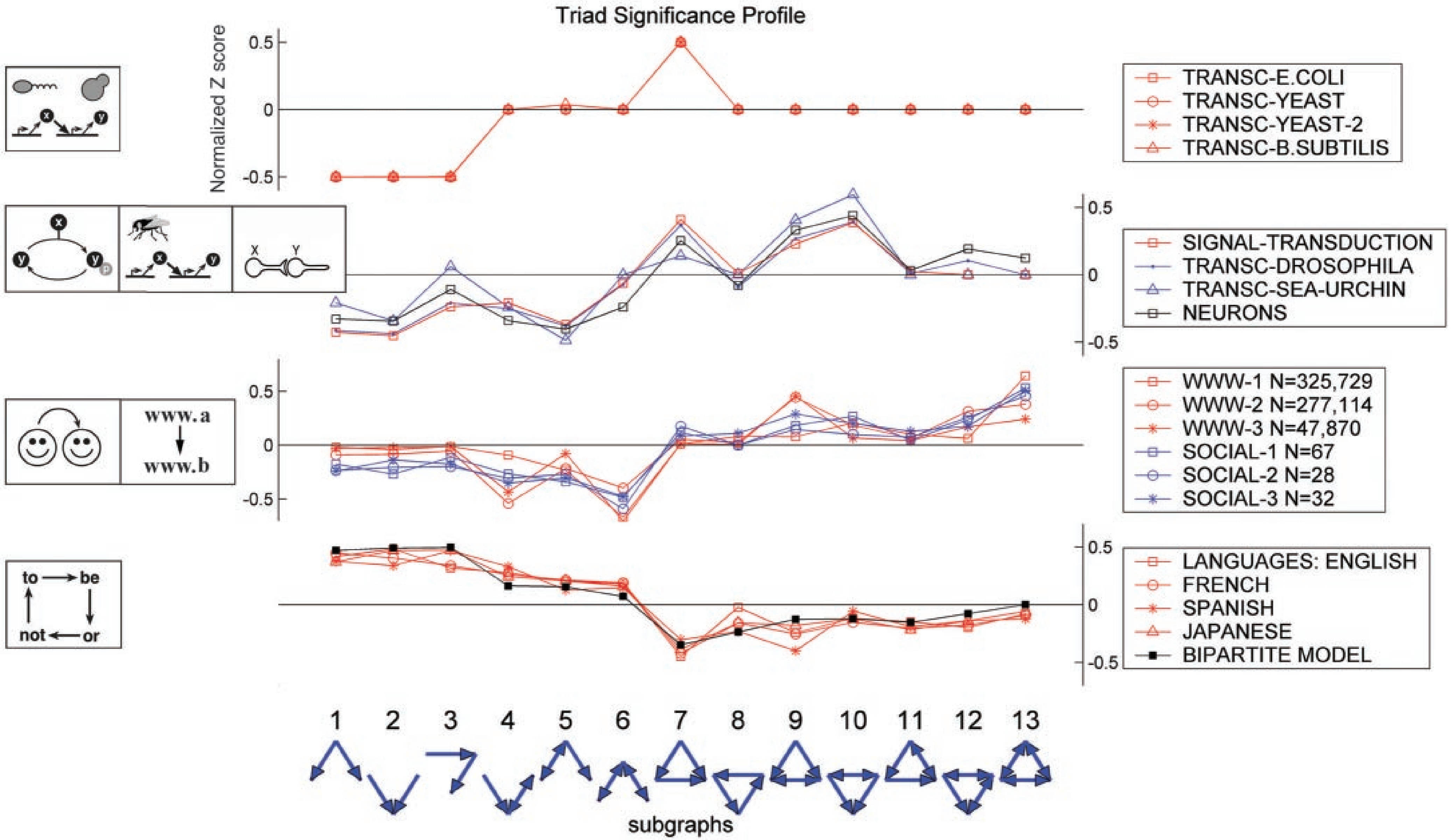
Food webs					Three chain			Bi-parallel
Little Rock	92	984	3219	3120 ± 50	2.1	7295	2220 ± 210	25
Ythan	83	391	1182	1020 ± 20	7.2	1357	230 ± 50	23
St. Martin	42	205	469	450 ± 10	NS	382	130 ± 20	12
Chesapeake	31	67	80	82 ± 4	NS	26	5 ± 2	8
Coachella	29	243	279	235 ± 12	3.6	181	80 ± 20	5
Skipwith	25	189	184	150 ± 7	5.5	397	80 ± 25	13
B. Brook	25	104	181	130 ± 7	7.4	267	30 ± 7	32



World Wide Web					Feedback with two mutual dyads			Fully connected triad			Uplinked mutual dyad
nd.edu§	325,729	1.46e6	1.1e5	2e3 ± 1e2	800	6.8e6	5e4±4e2	15,000	1.2e6	1e4 ± 2e2	5000

Milo *et al*,2002. Science 298:824

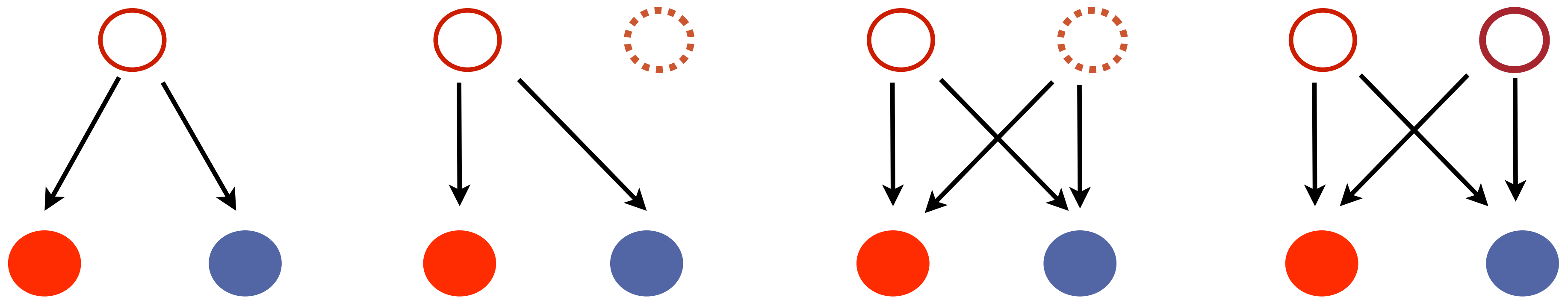
Perfiles



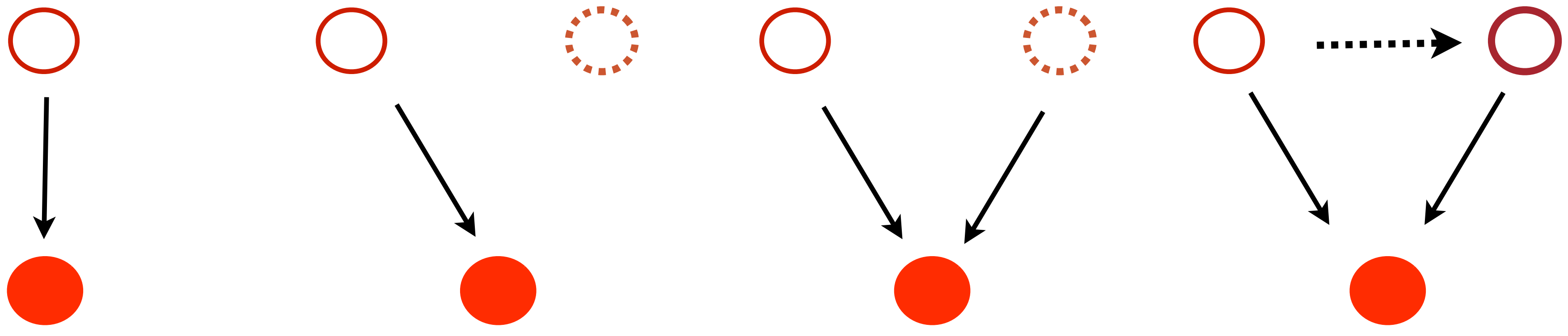
Milo et al. 2004 Science 303:1538-1542

Evolución de Motivos

Bi-fans



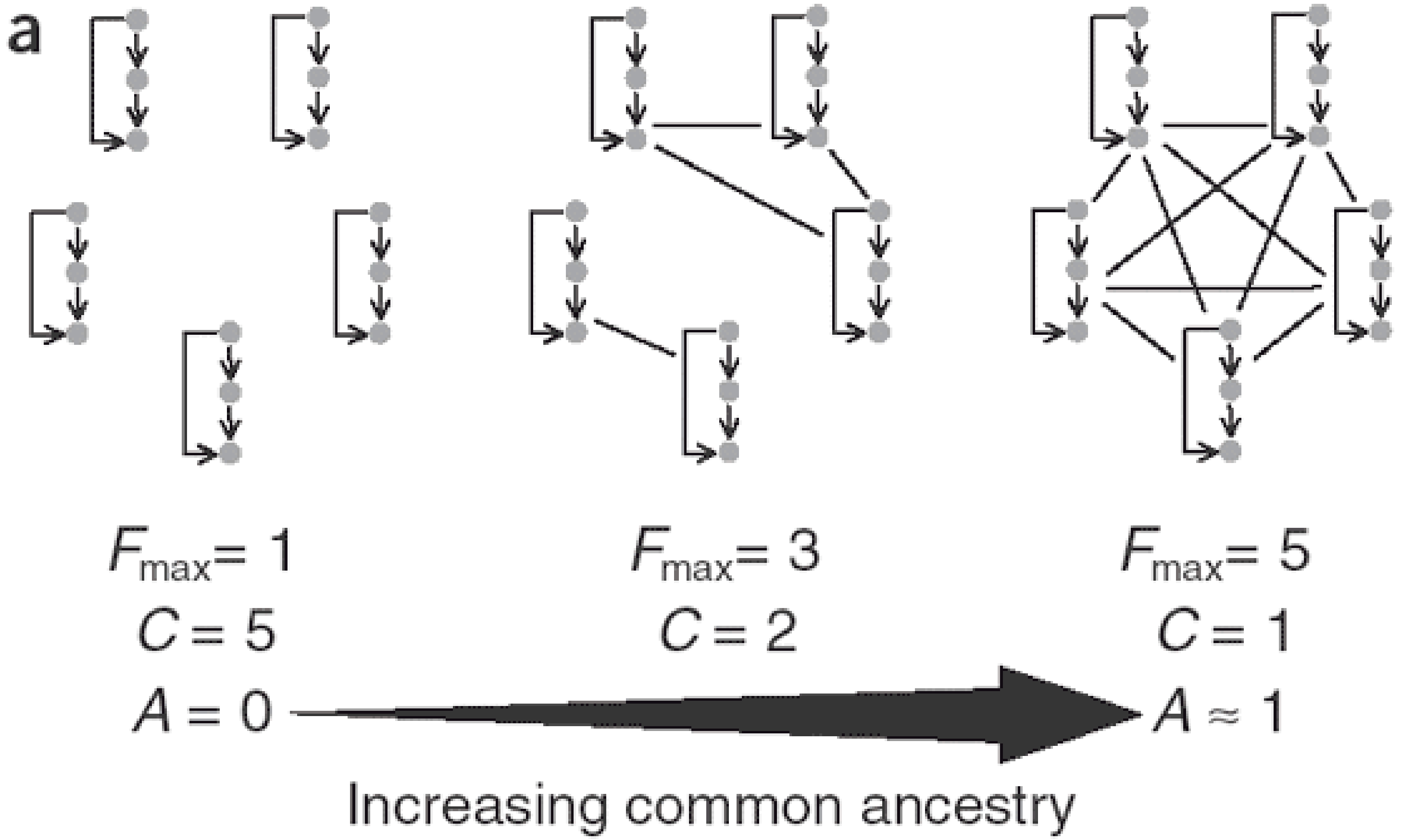
FFL

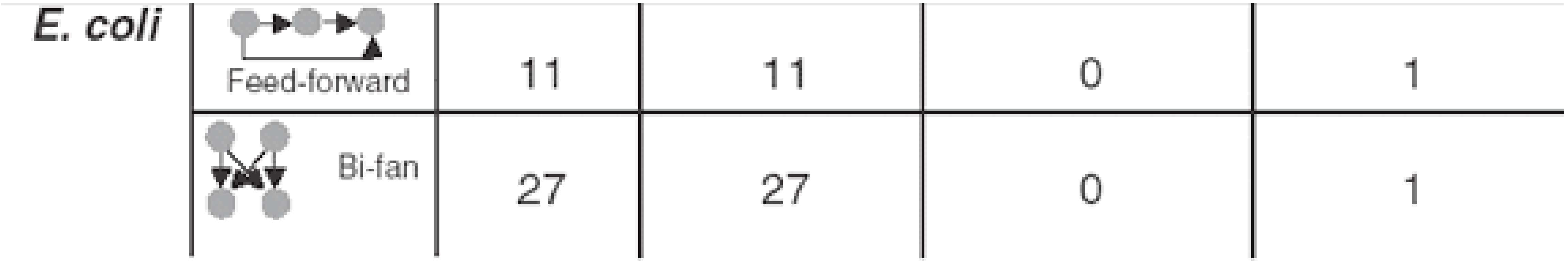


NO: En E. coli no hay reguladores homologos en el mismo motivo

Tecihmann & Babu, 2004 Nature Genetics 36(5):492-6

Evolución de Motivos II

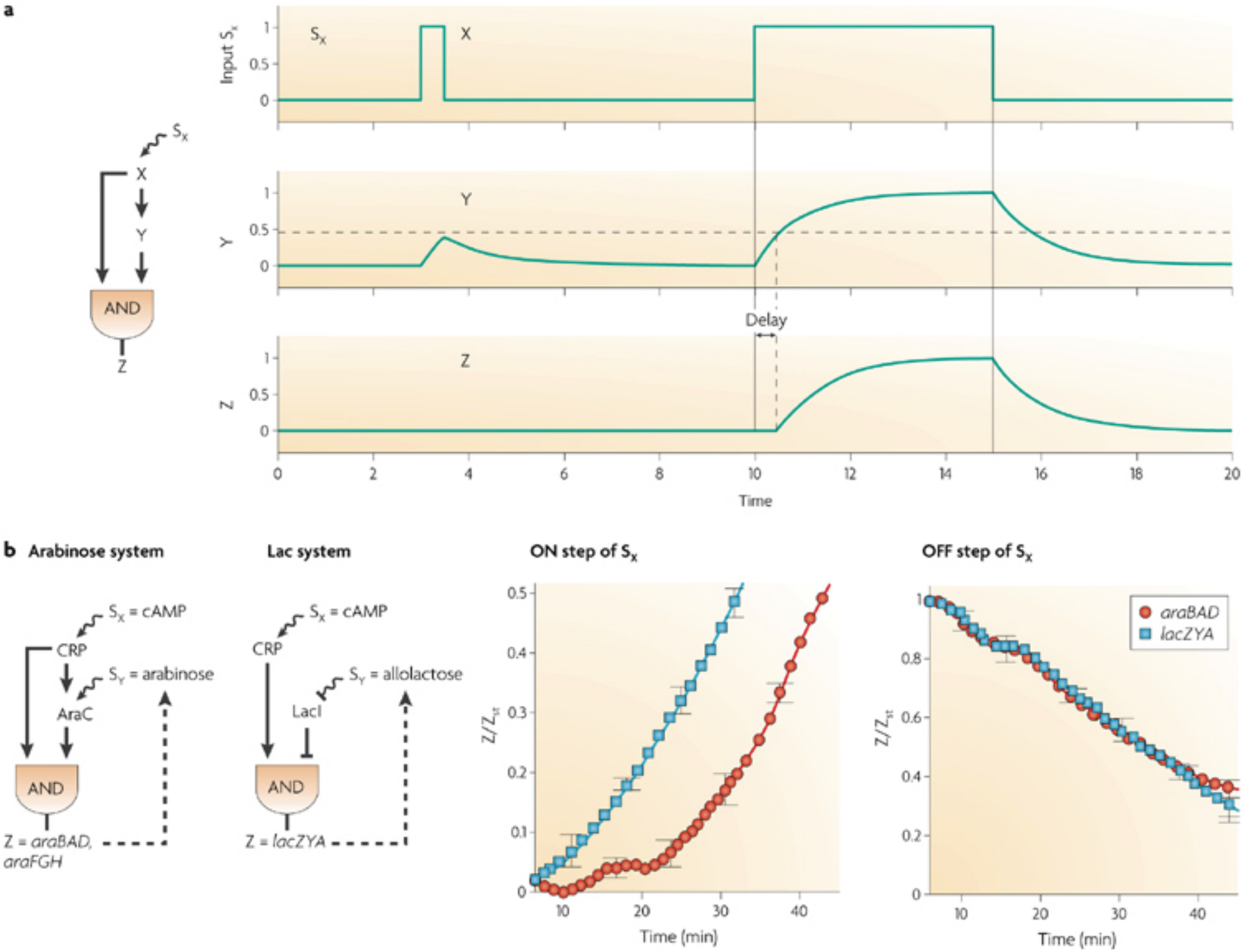




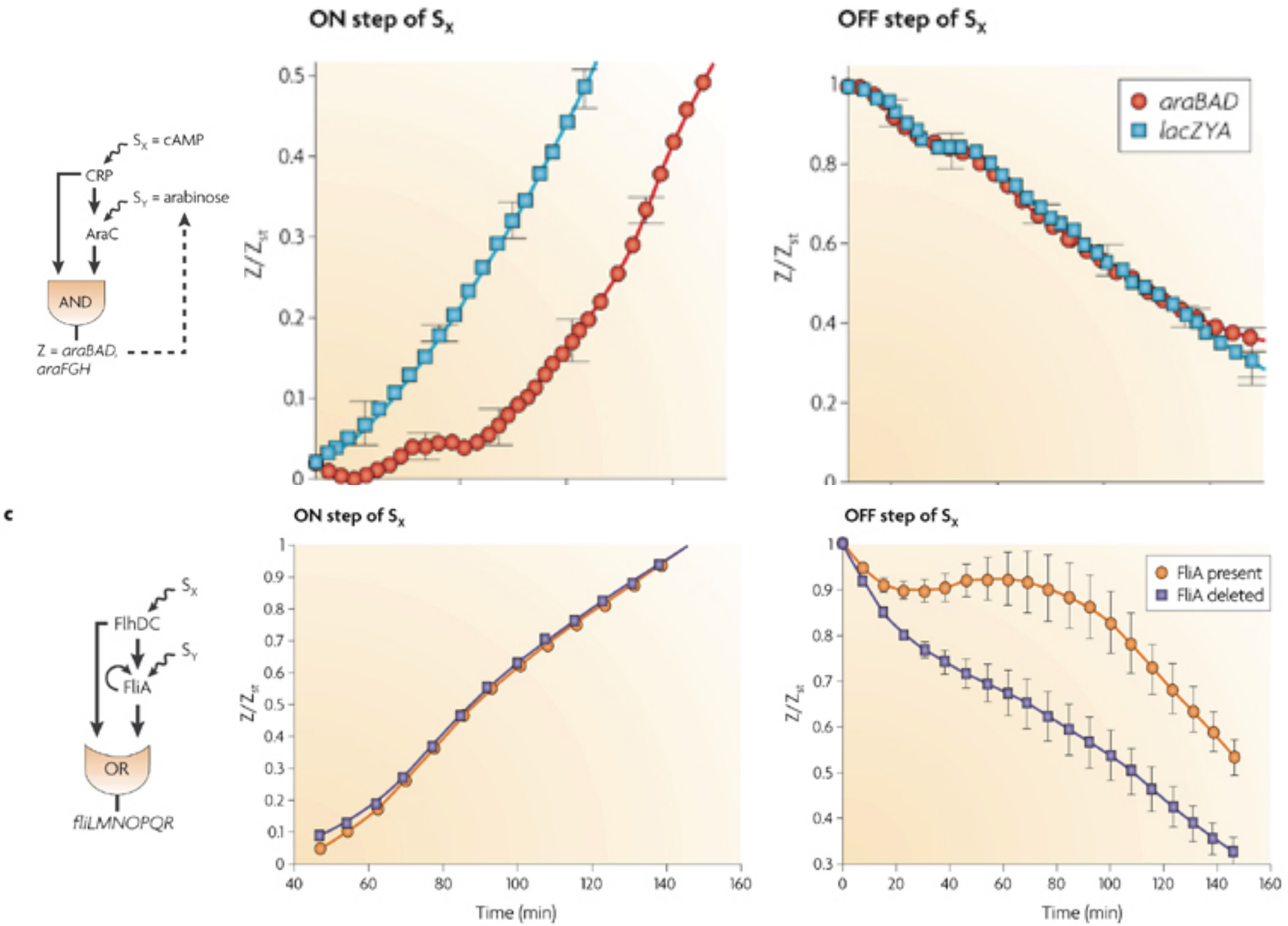
E. coli	Feed-forward	11	11	0	1
	Bi-fan	27	27	0	1

Conant & Wagner,2003. Nat Genet. 34:264

Propiedades de los motivos

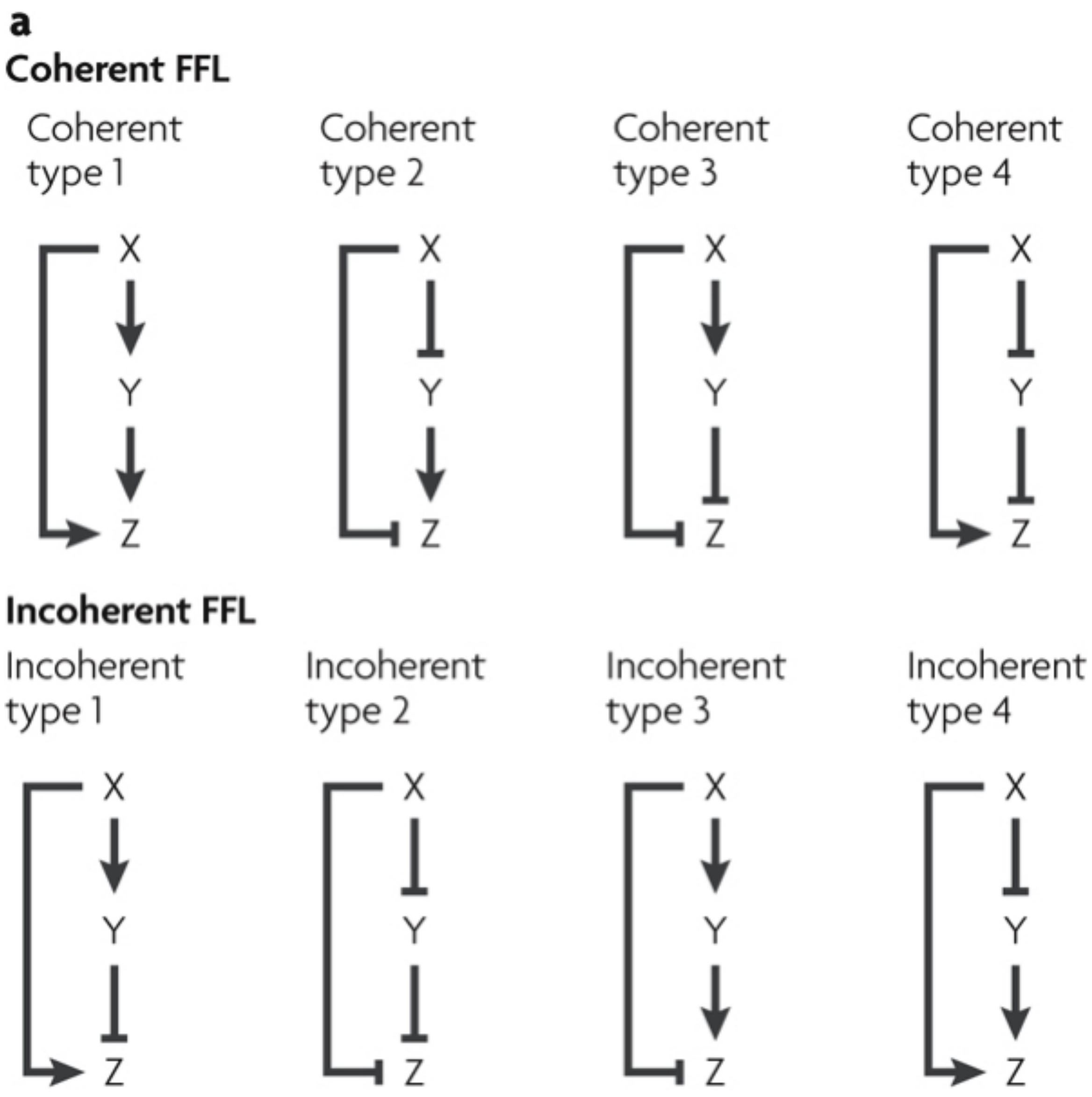


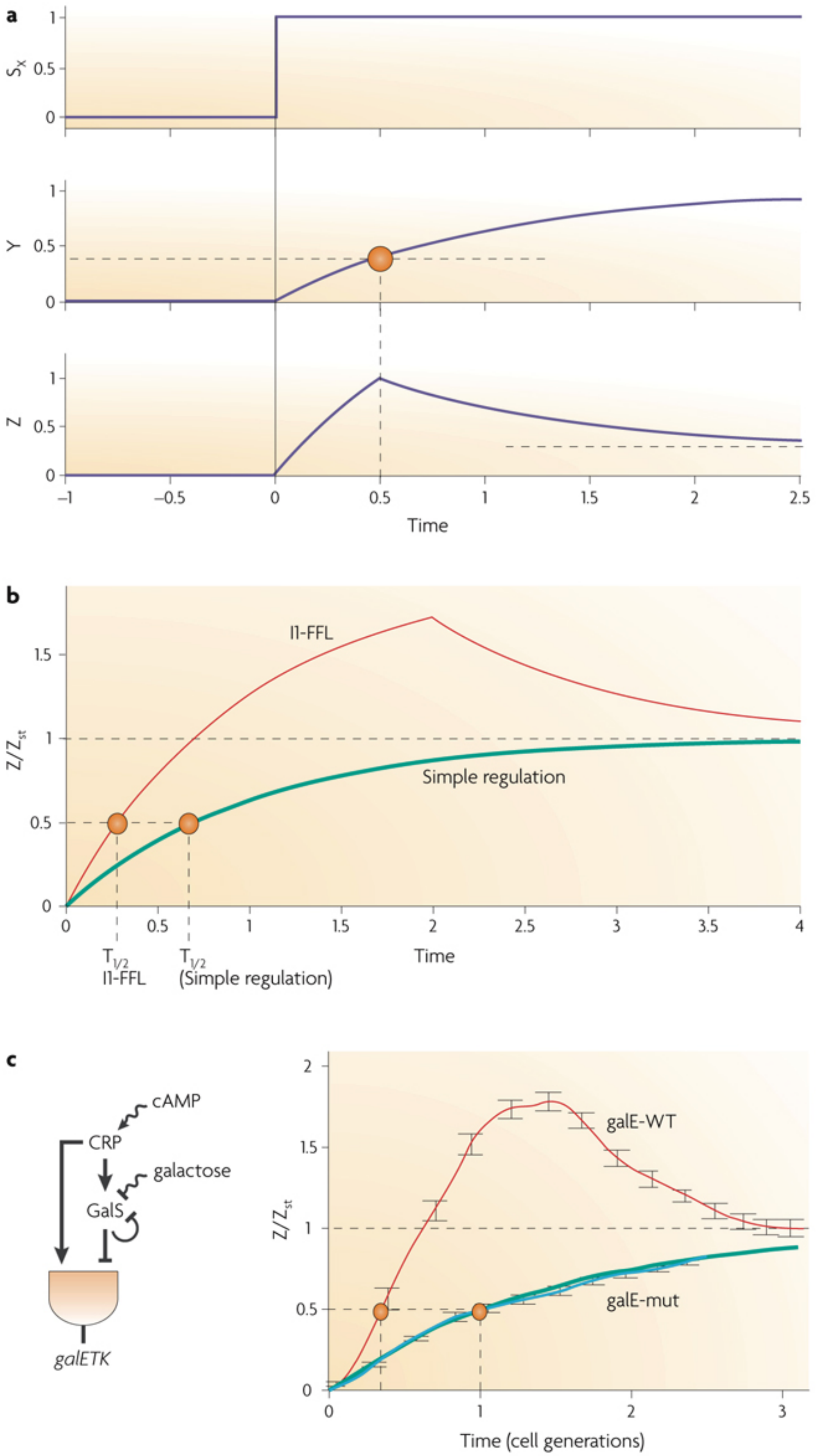
Shen-Orr *et al.*,2002. Nat Genet. 31:64   Alon, U. (2007). Nat Rev. Genet. 8:450:461



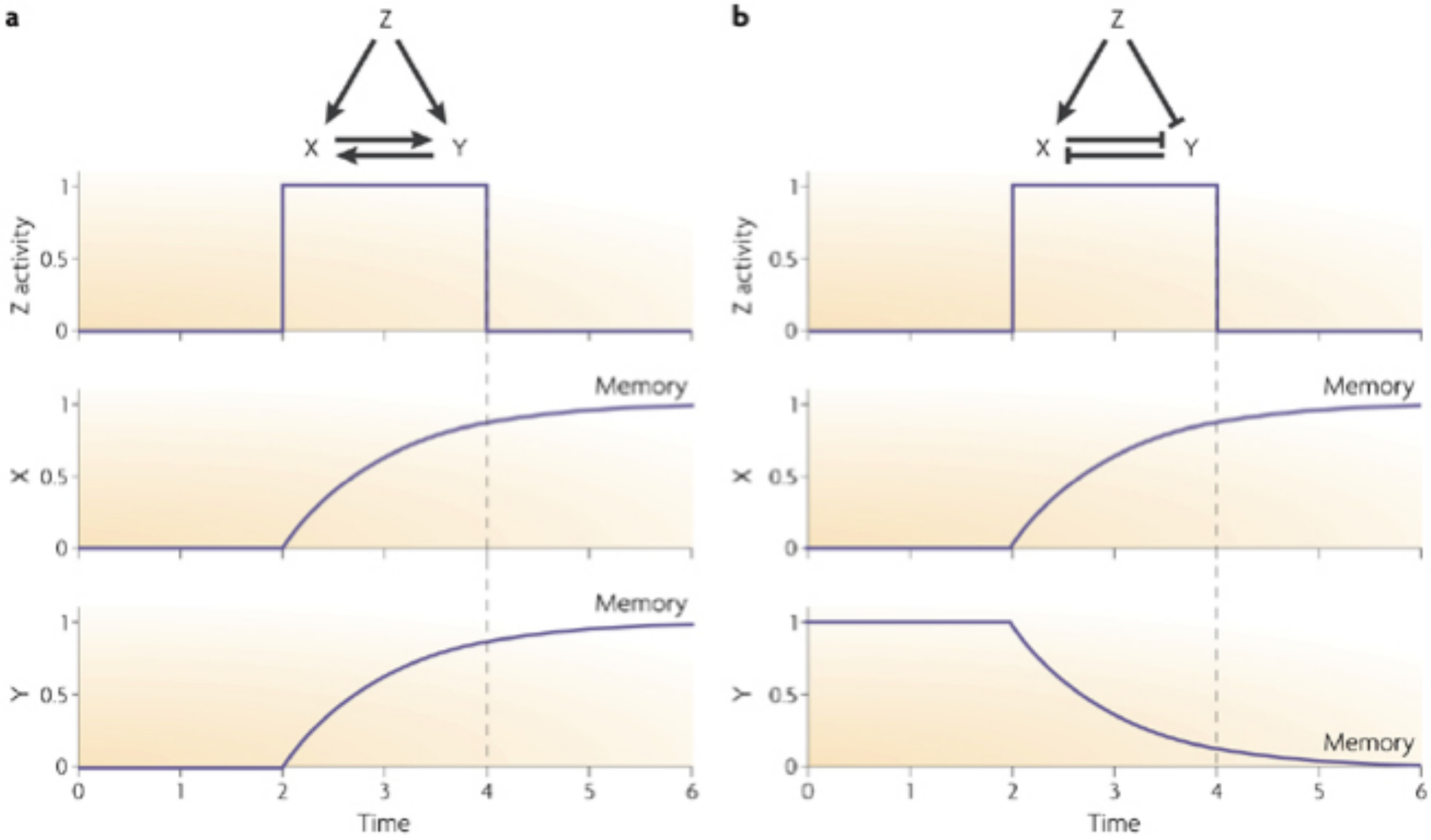
Nature Reviews | Genetics

Alon, U. (2007). Nat Rev. Genet. 8:450:461





Nature Reviews | Genetics

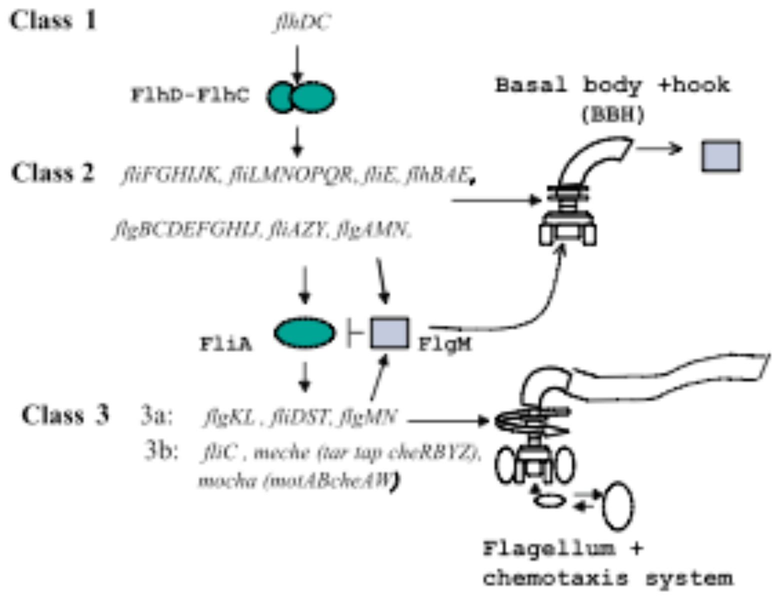


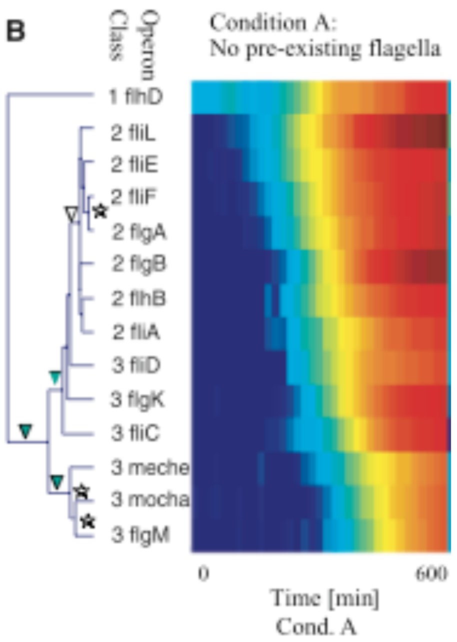
Alon, U. (2007). Nat Rev. Genet. 8:450:461

Pero esto es damasiado general

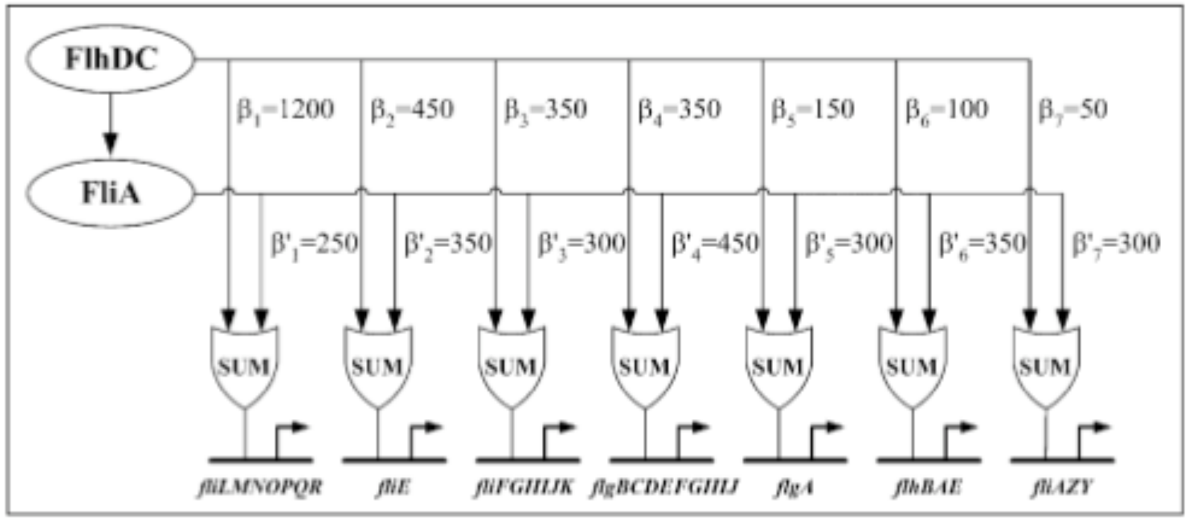
- No todos los reguladores se unen con la misma afinidad al DNA o a su señal.
- No todos los reguladores se unen a todo sus sitios siempre,
- No todos los reguladores estan siempre

Optimizacion Temporal



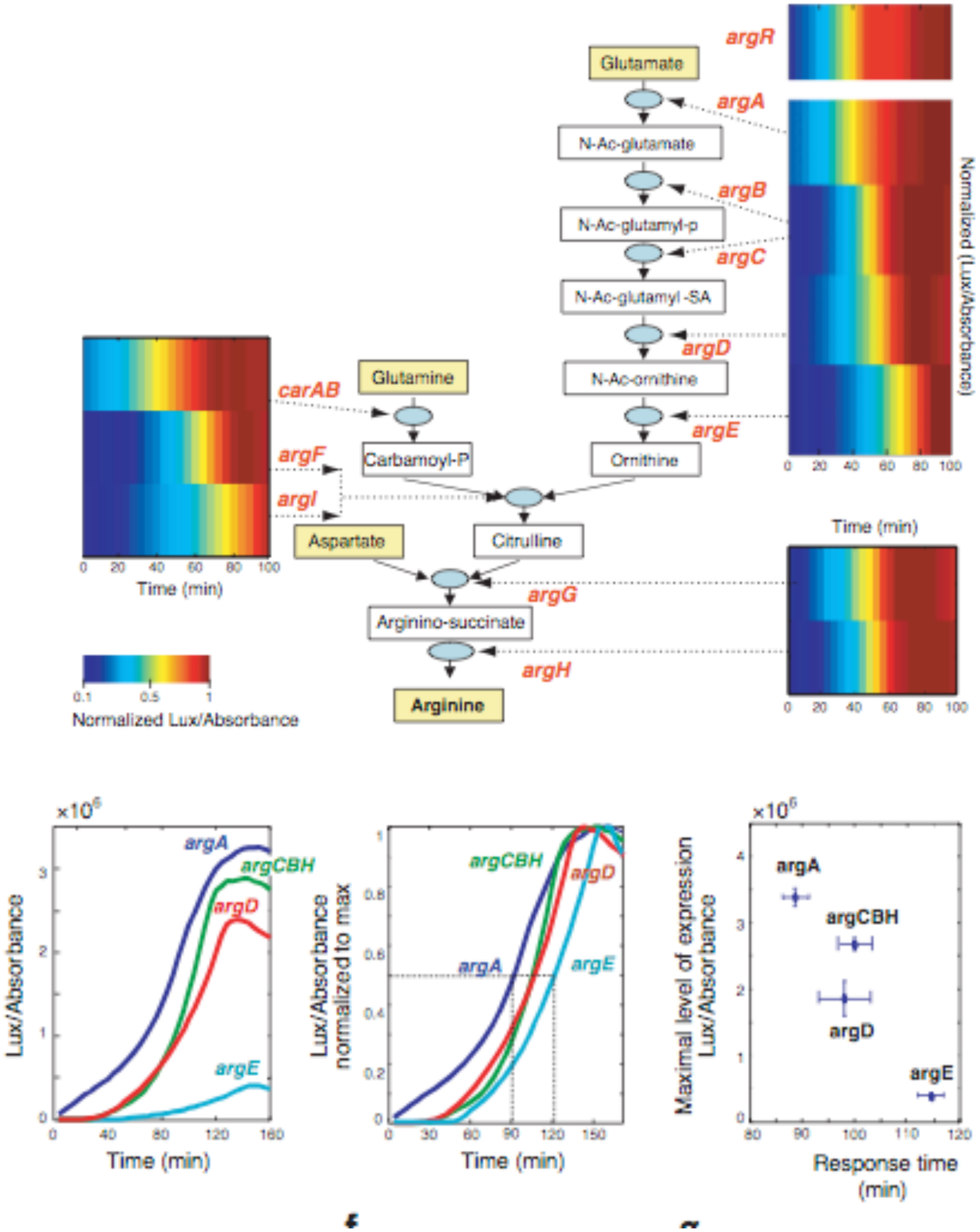


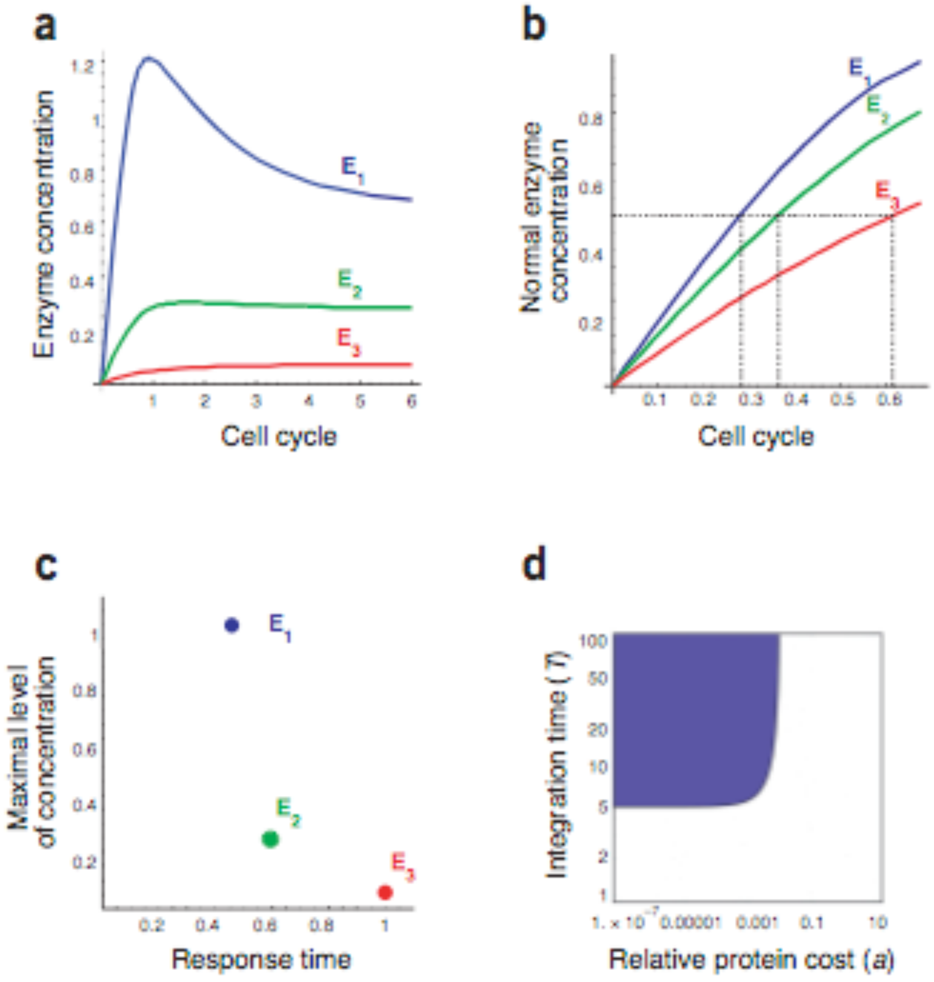
Kalir S, et al (2001) Science 292(5524):2080-3



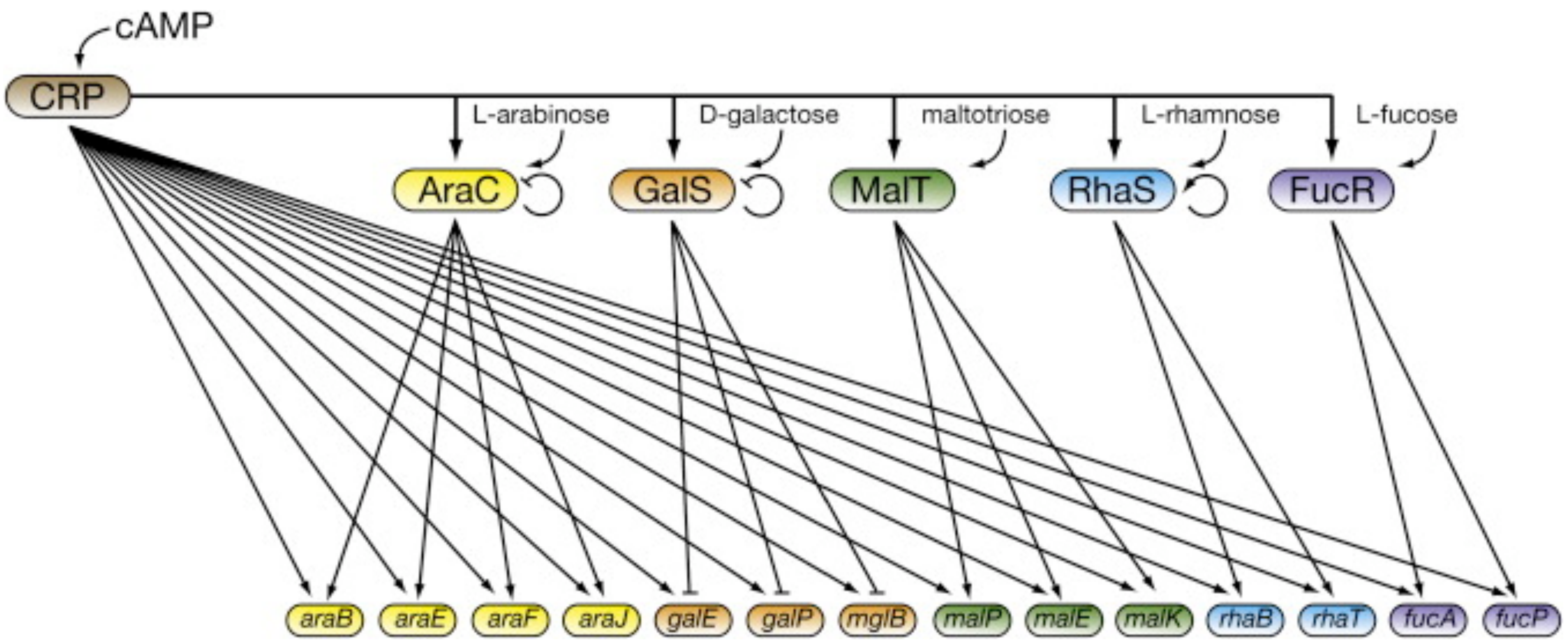
S . Kalir , U . Alon (2004)Cell 117:6 pp 713 - 720

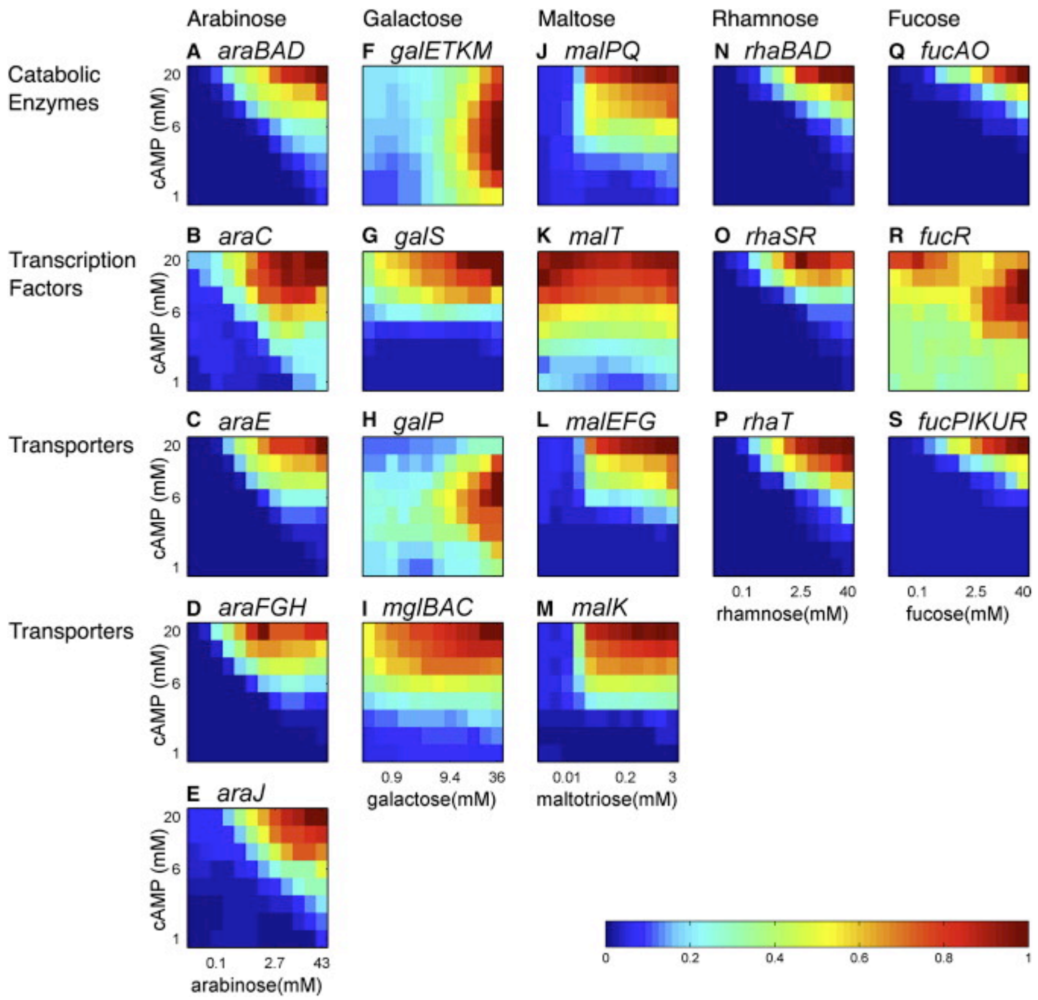
Optimización en Gasto

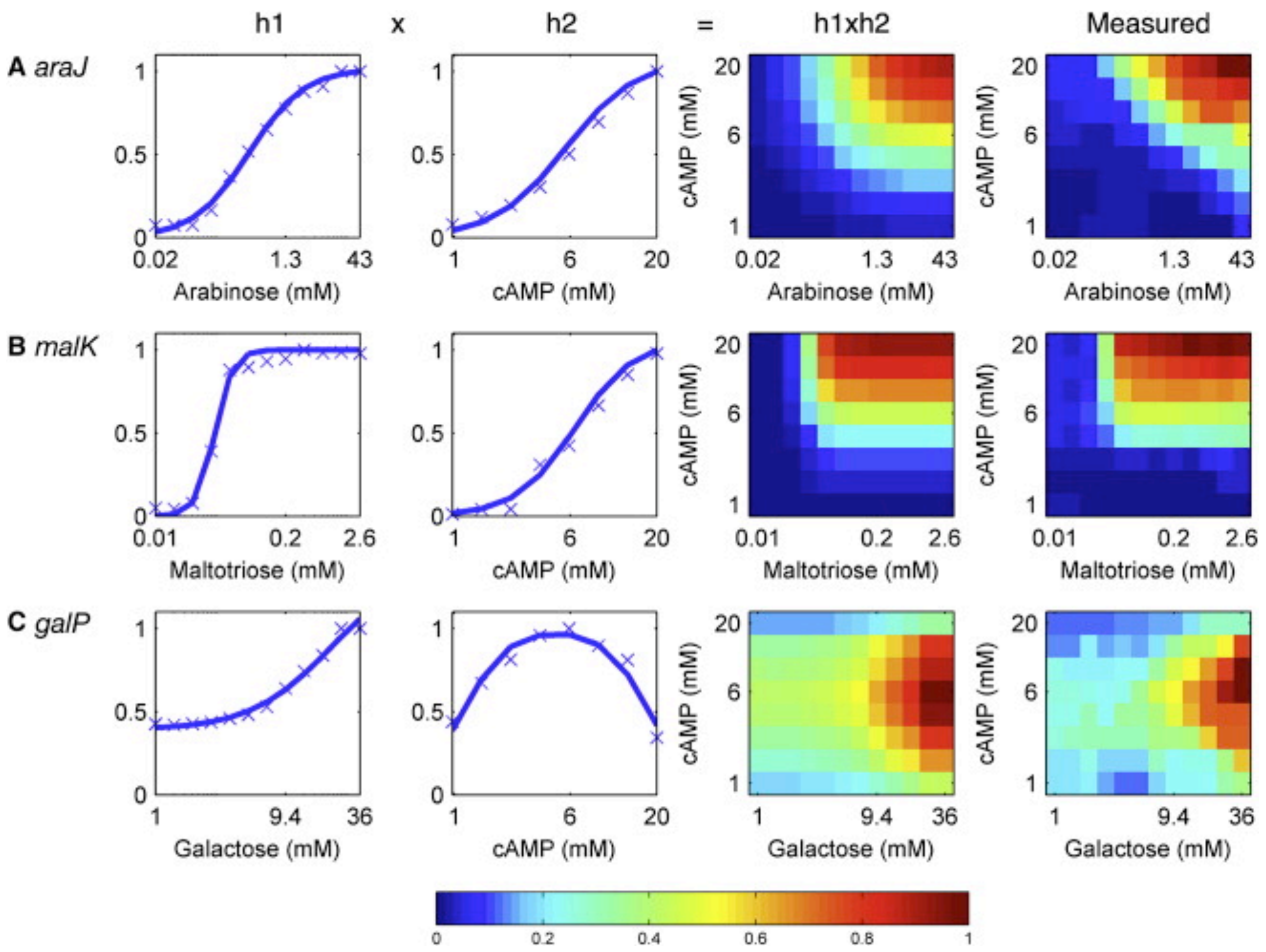




Zaslaver A et al (2004) Nat Gene 36(5):486-91



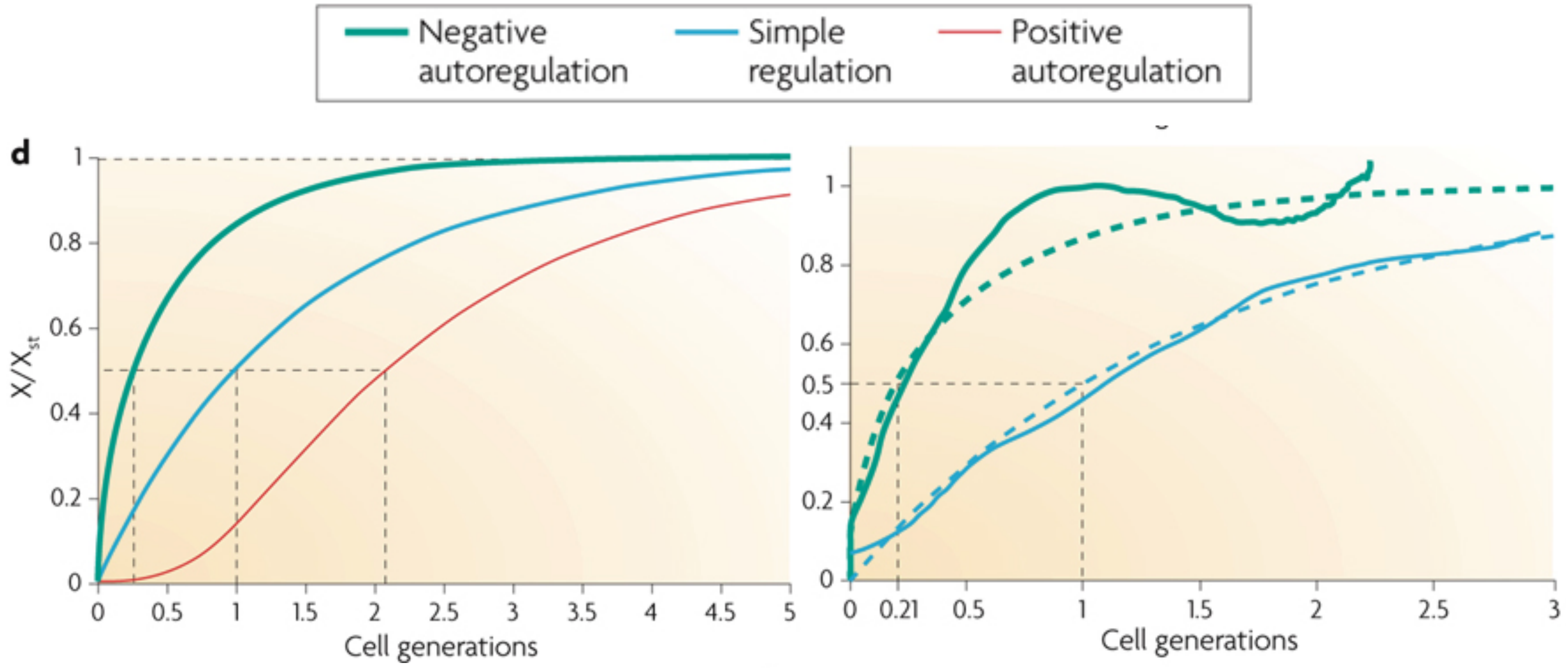




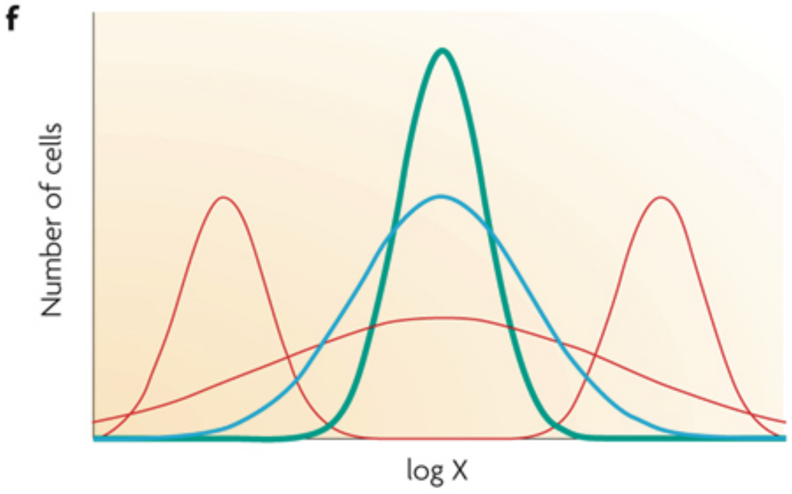
Kaplan et al. (2008). Mol Cell 29:786:792

Autoregulación

- El 70% de los reguladores de E.coli regulan su propia expresión



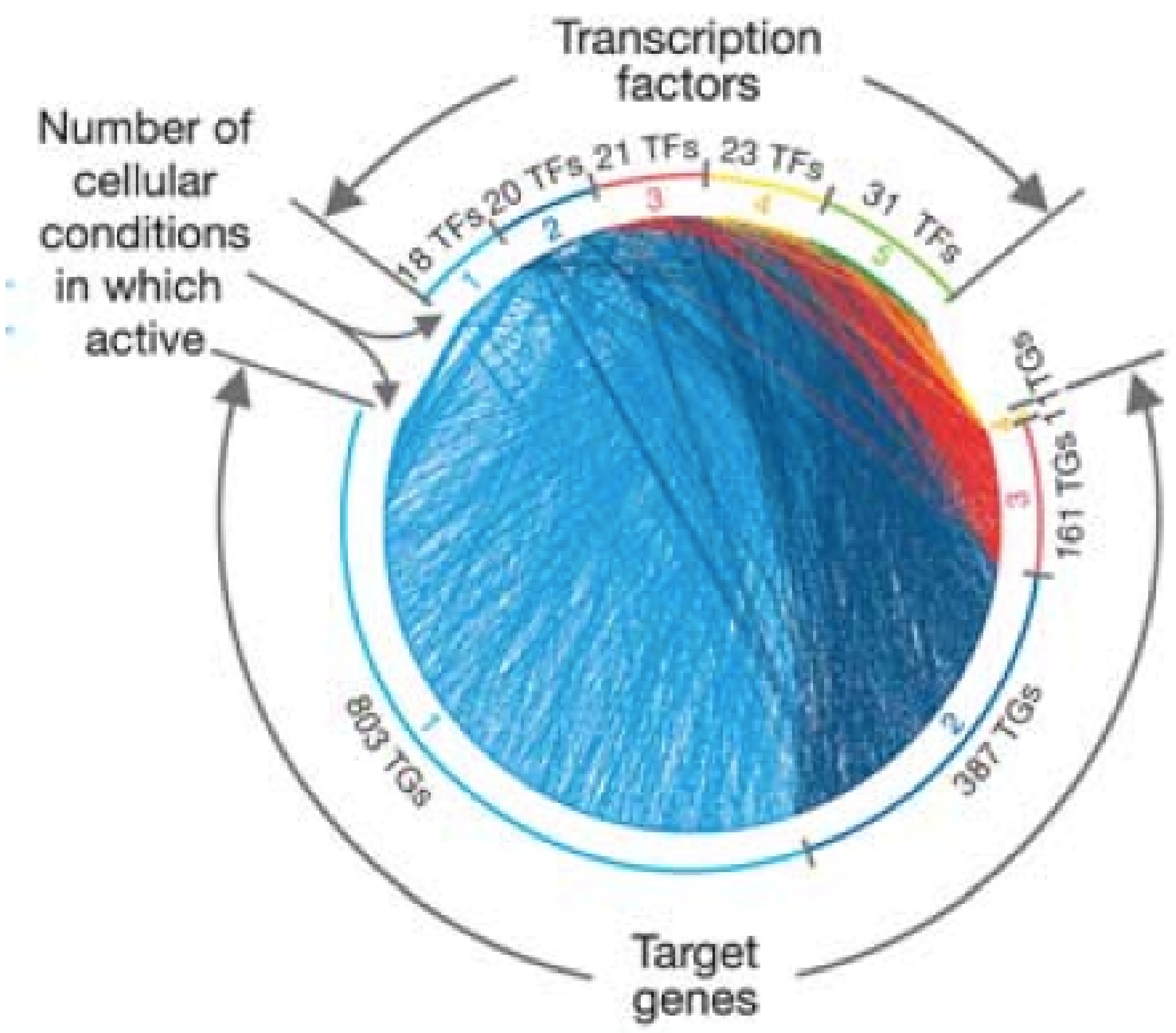
Cinética de Inducción



Estabilidad

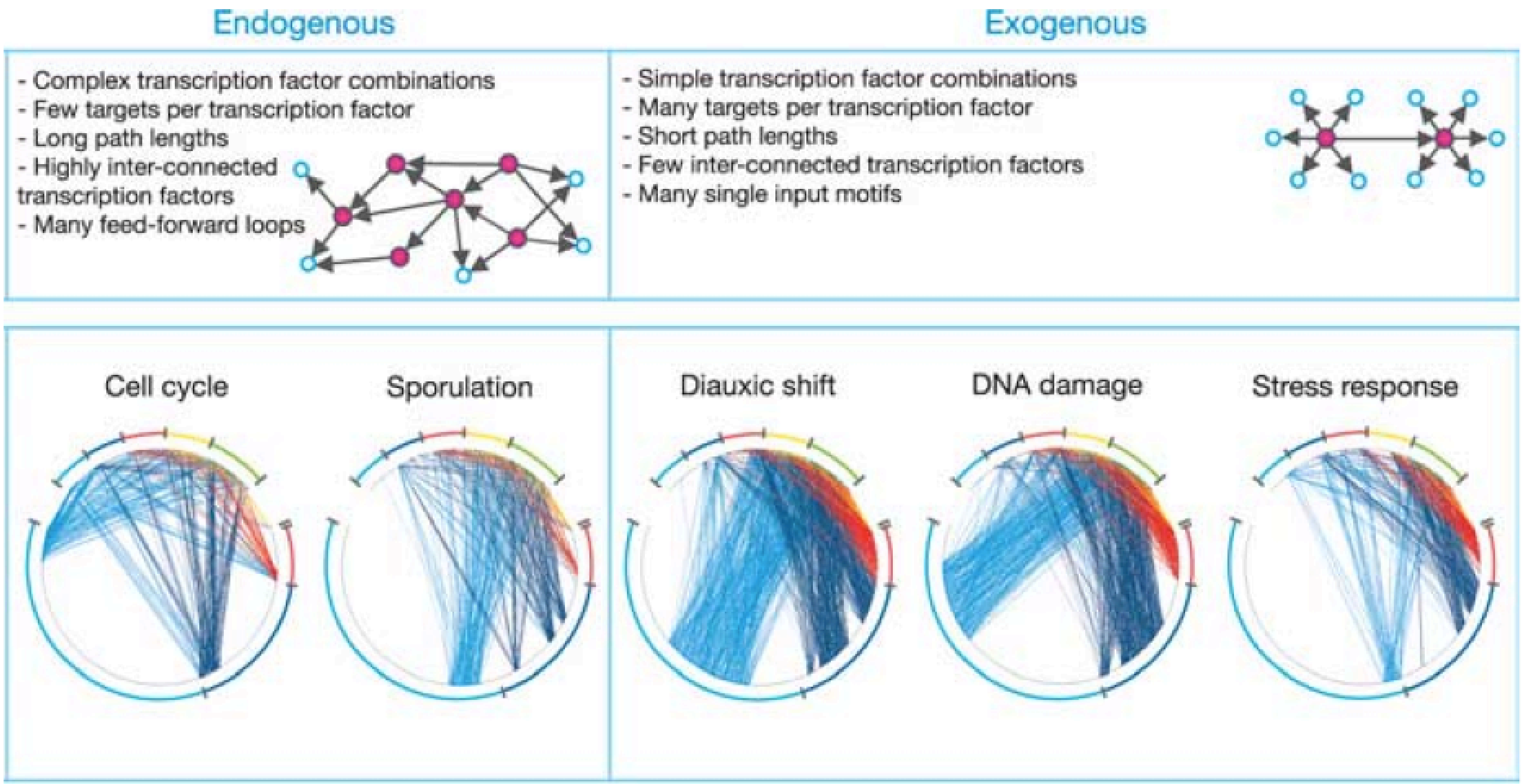
Alon, U. (2007). Nat Rev. Genet. 8:450:461

Dinámica de Redes de Regulación



Luscombe *et al.*, 2004 Nature 431:308

Dinámica de Redes de Regulación



Endogenous

- Complex transcription factor combinations
- Few targets per transcription factor
- Long path lengths
- Highly inter-connected transcription factors
- Many feed-forward loops

Exogenous

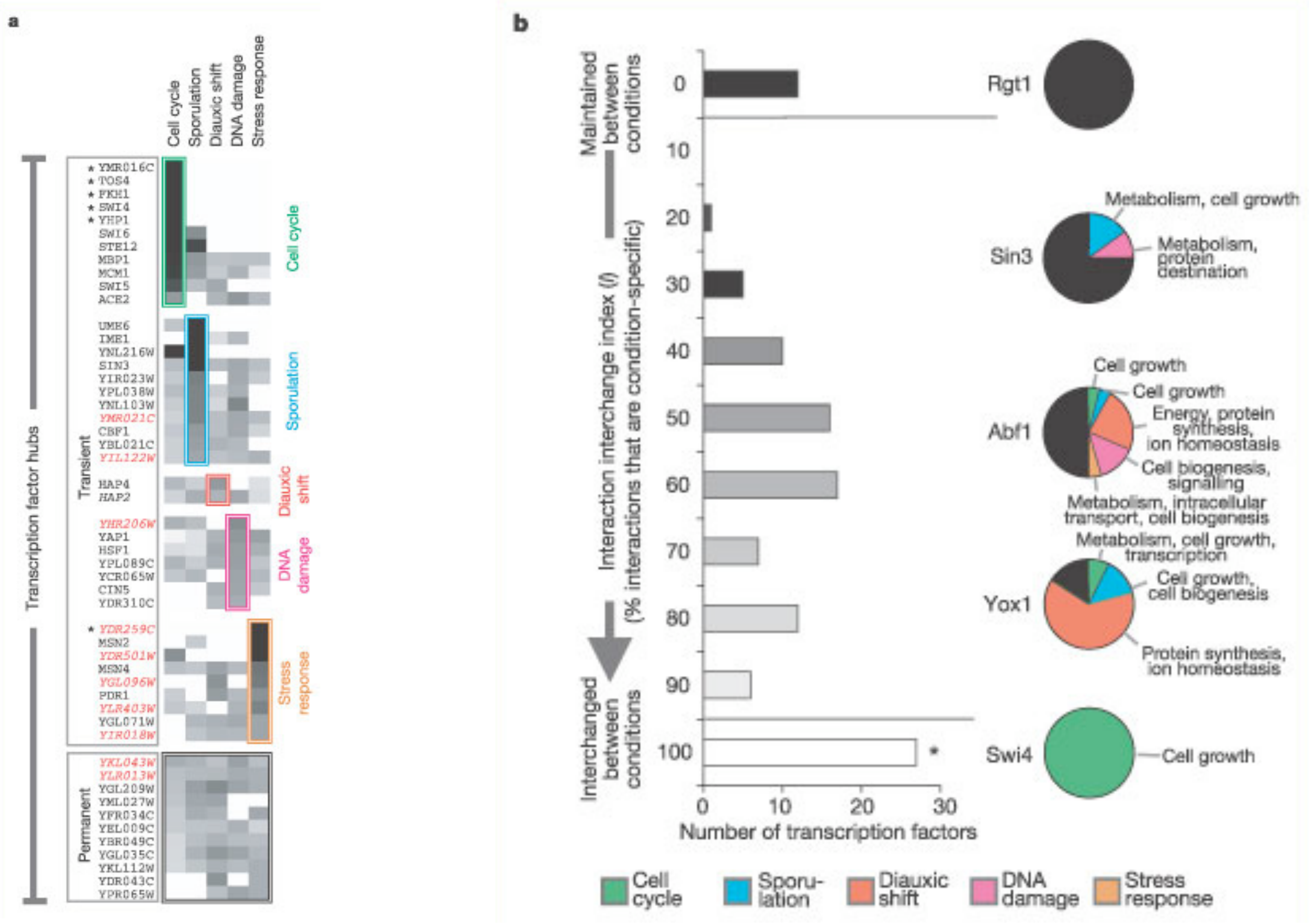
- Simple transcription factor combinations
- Many targets per transcription factor
- Short path lengths
- Few inter-connected transcription factors
- Many single input motifs

Luscombe *et al.*, 2004 Nature 431:308

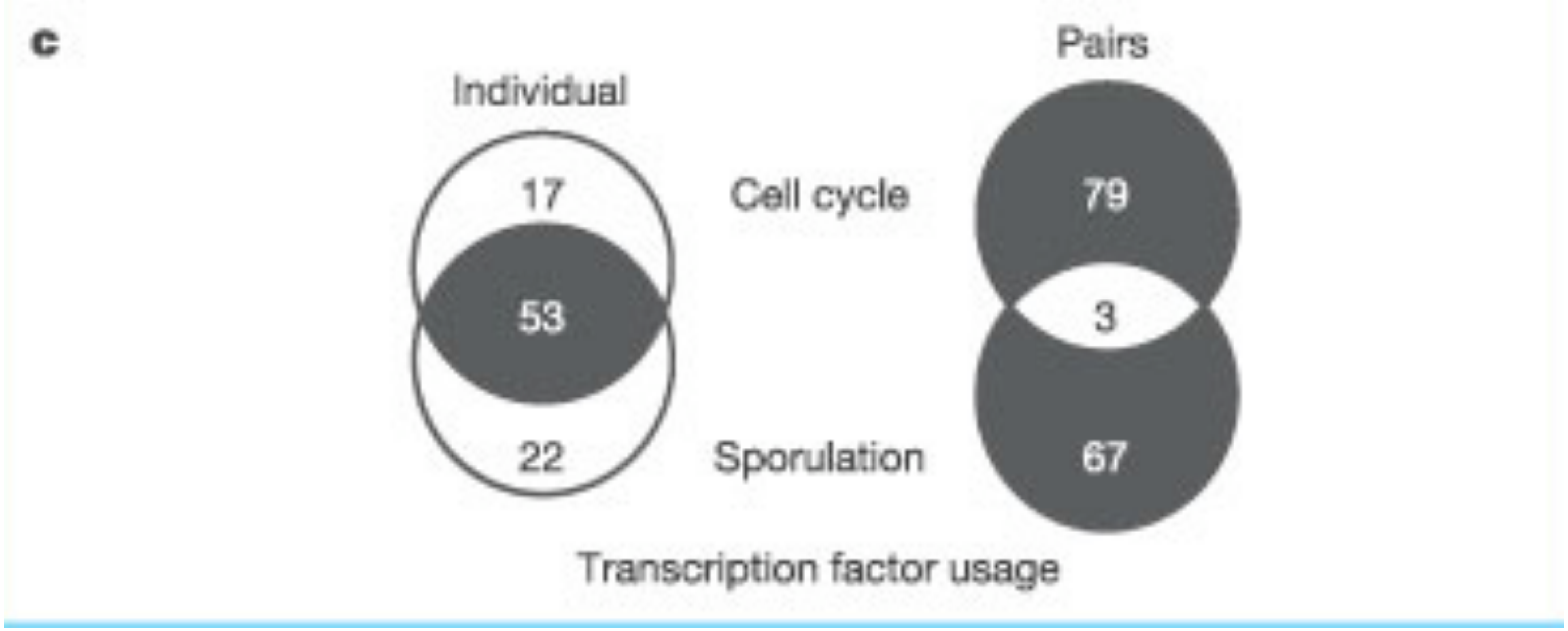
Dinámica de Redes de Regulación

c				Endogenous				Exogenous					
Size	No. of transcription factors	142		70		74		71		72		63	
	No. of target genes	3,420		280		257		748		678		362	
	No. of regulatory interactions	7,074		550		481		1,217		1,082		566	
Topological measures	In-degree ($<k_{in}>$)	2.1		2.0		1.9		1.6		1.6		1.6	
	Out-degree ($<k_{out}>$)	49.8		7.9		6.5		17.1		15.0		9.0	
	Path length ($<l>$)	4.7		4.5		3.4		2.1		2.0		2.2	
	Clustering coefficient ($<c>$)	0.11		0.15		0.14		0.09		0.09		0.08	
Motifs (%)	Single input (SIM)	1,748	(37.6%)	130	(32.0%)	117	(38.9%)	438	(57.4%)	462	(55.7%)	228	(59.1%)
	Multiple input (MIM)	325	(7.0%)	96	(23.7%)	50	(16.6%)	180	(23.6%)	226	(27.3%)	78	(20.2%)
	Feed-forward loop (FFL)	2,581	(55.5%)	180	(44.3%)	134	(44.5%)	145	(19.0%)	141	(17.0%)	80	(20.7%)
	Total	4,654		406		301		763		829		386	

Luscombe *et al.*, 2004 Nature 431:308



Luscombe *et al.*, 2004 Nature 431:308



Luscombe *et al.*, 2004 Nature 431:308

Resumen

- Las redes de regulación son redes dirigidas, y cualificadas.
- El número de genes regulados por un FT sigue una distribución de potencias
- Presenta Motivos específicos FFL y Bi-Fan
- Los motivos se seleccionan por sus propiedades cinéticas

Resumen II

- Los sistemas de regulación están optimizados para la mejor expresión temporal y el menor gasto energético
- Esto se consigue mediante la optimización de la expresión de los FT, su afinidad por el ADN y por su señal
- La red cambia para adaptarse a el tipo de estímulo, y esto se consigue mediante el uso combinatorio de FT.