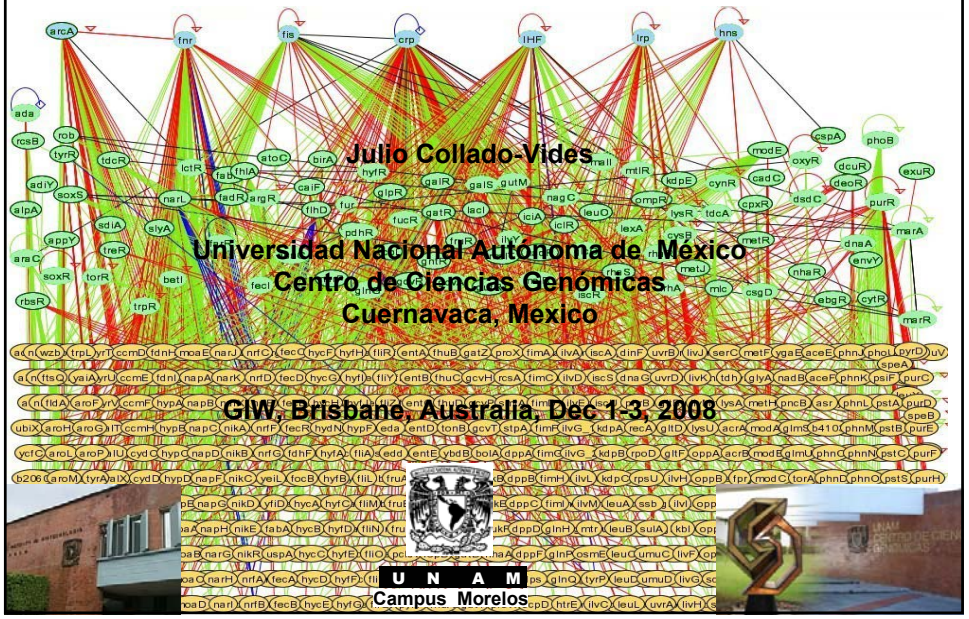
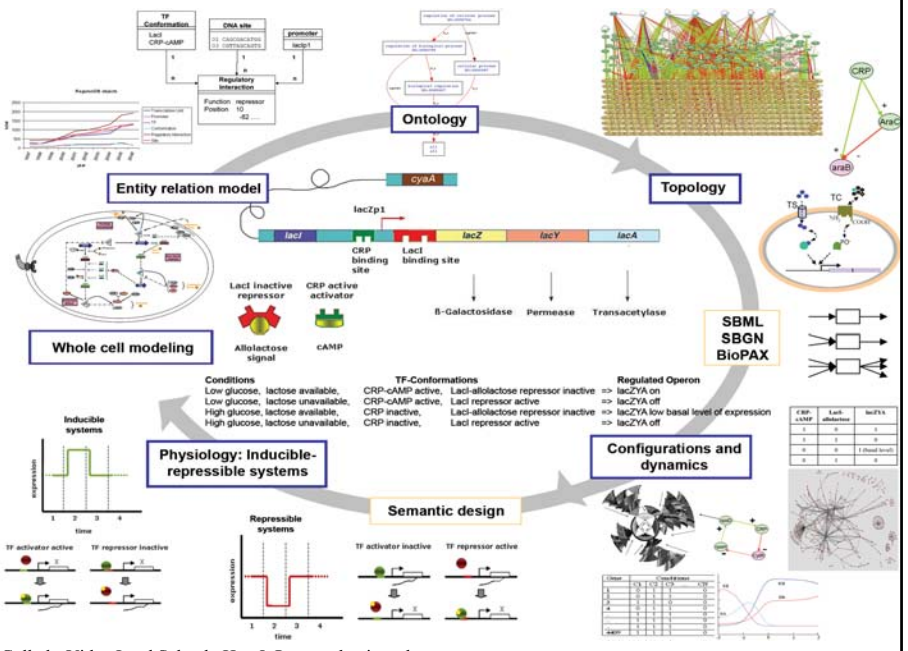
Computational Biology and Modeling of Gene Regulation in Bacteria



The outline:



Collado-Vides J and Salgado H. *J. Bacteriol*. rejected

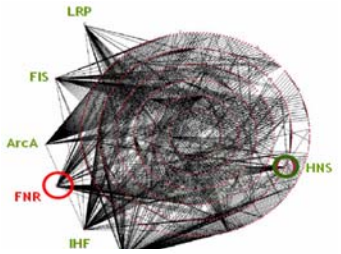
Escherichia coli: Model with knowledge

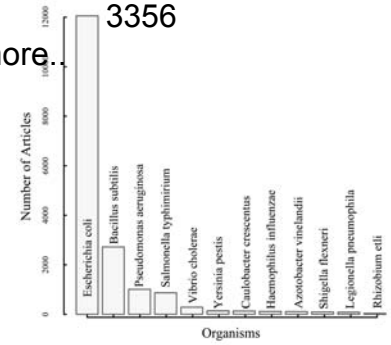
As of July 2008

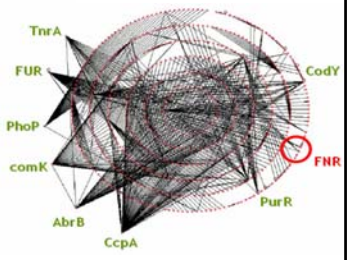
Escherichia coli

Regulons	361
TF BSs	2342
Promoters	1754
TFs	163
TUs	3356

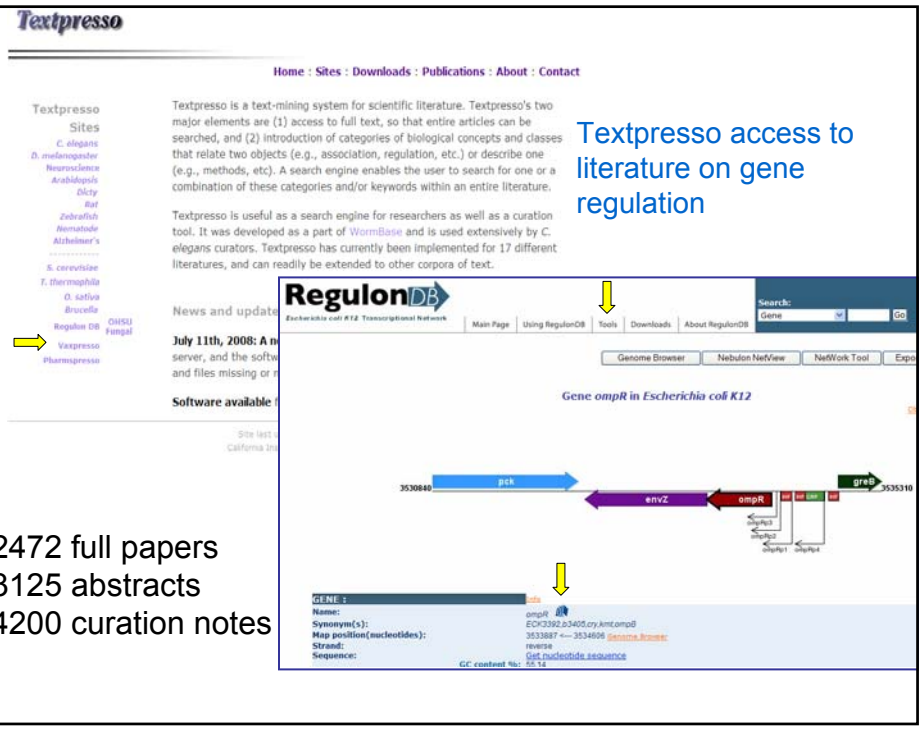
and more…







Bacillus subtilis



Textpresso access to literature on gene regulation

2472 full papers
3125 abstracts
4200 curation notes

View Matches Page

Show Matches

Note: The color scheme only highlights words that are in the selected categories. It does not mean that other specifications for that word are met. In the color scheme, keyword colors supersede category colors, but are superseded by the hyperlink color.

Query: Categories + regulation + effect - Keywords + ompR*

Display page 1 , or previous or next page.

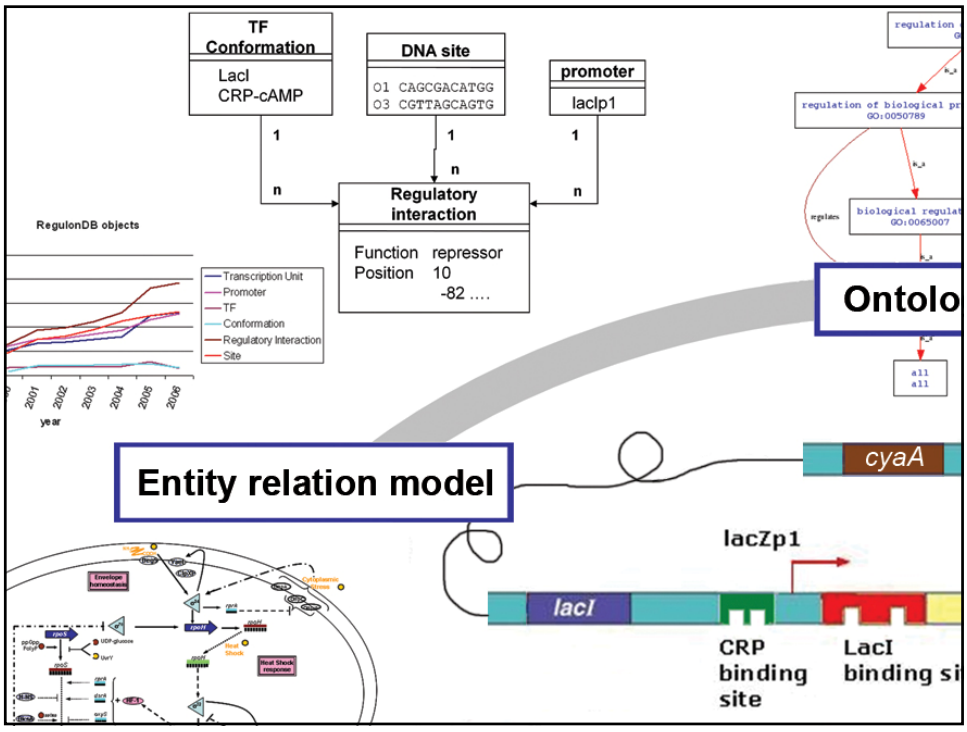
File ID pmid11158569 (Abstract), Sentences 1 to 11: Escherichia coli modulates its porin expression through a histidine kinase , EnvZ , and its cognate response regulator , OmpR . EnvZ is a bifunctional enzyme that possesses both OmpR kinase and phosphorylated OmpR (OmpR-P) phosphatase activities and thus controls the cellular level of OmpR-P In an in vitro -assay system , the addition of OmpR to the reaction mixture consisting of the cytoplasmic domain of EnvZ (EnvZc) and ATP produces a barely detectable amount of OmpR-P because of the dual activities of EnvZ . Here we report that DNA fragments containing the upstream promoter regions of the porin genes (ompF and ompC) can shift the equilibrium between OmpR and OmpR-P dramatically toward OmpR-P Among the four reactions occurring in the mixture , only the EnvZ phosphatase activity was inhibited severely by the specific DNA , in contrast to the previous report by Kenney and her associates that DNA stimulates OmpR phosphorylation by EnvZ [Ames , S K , Frankema , N AND Kenney , L J (1999) Proc . Natl . Acad . Sci . USA 96 , 11792- 11797] . The autophosphorylation of EnvZc and the phosphotransfer from phosphorylated EnvZc to OmpR were not affected by DNA , whereas the autodephosphorylation of OmpR-P was inhibited slightly . We propose that the apparent inhibitory effect of DNA on the EnvZ phosphatase function is caused by sequestrating OmpR-P from the reaction as a result of OmpR-P binding to DNA .

File ID pmid11158569 (Abstract), Sentences 1 to 12: Escherichia coli modulates its porin expression through a histidine kinase , EnvZ , and its cognate response regulator , OmpR . **EnvZ is a bifunctional enzyme that possesses both OmpR kinase and phosphorylated OmpR (OmpR-P) phosphatase activities and thus controls the cellular level of OmpR-P In an in vitro -assay system , the addition of OmpR to the reaction mixture consisting of the cytoplasmic domain of EnvZ (EnvZc) and ATP produces a barely detectable amount of OmpR-P because of the dual activities of EnvZ .** Here we report that DNA fragments containing the upstream promoter regions of the porin genes (ompF and ompC) can shift the equilibrium between OmpR and OmpR-P dramatically toward OmpR-P Among the four reactions occurring in the mixture , only the EnvZ phosphatase activity was inhibited severely by the specific DNA , in contrast to the previous report by Kenney and her associates that DNA stimulates OmpR phosphorylation by EnvZ [Ames , S K , Frankema , N AND Kenney , L J (1999) Proc . Natl . Acad . Sci . USA 96 , 11792- 11797] . The autophosphorylation of EnvZc and the phosphotransfer from phosphorylated EnvZc to OmpR were not affected by DNA , whereas the autodephosphorylation of OmpR-P was inhibited slightly . We propose that the apparent inhibitory effect of DNA on the EnvZ phosphatase function is caused by sequestrating OmpR-P from the reaction as a result of OmpR-P binding to DNA .

File ID pmid11158569 (Abstract), Sentences 1 to 15: Escherichia coli modulates its porin expression through a histidine kinase , EnvZ , and its cognate response regulator , OmpR . EnvZ is a bifunctional enzyme that possesses both OmpR kinase and phosphorylated OmpR (OmpR-P) phosphatase activities and thus controls the cellular level of OmpR-P In an in vitro -assay system , the addition of OmpR to the reaction mixture consisting of the cytoplasmic domain of EnvZ (EnvZc) and ATP produces a barely detectable amount of OmpR-P because of the dual activities of EnvZ . Here we report that DNA fragments containing the upstream promoter regions of the porin genes (ompF and ompC) can shift the equilibrium between OmpR and OmpR-P dramatically toward OmpR-P Among the four reactions occurring in the mixture , only the EnvZ phosphatase activity was inhibited severely by the specific DNA , in contrast to the previous report by Kenney and her associates that DNA stimulates OmpR phosphorylation by EnvZ [Ames , S K , Frankema , N AND Kenney , L J (1999) Proc . Natl . Acad . Sci . USA 96 , 11792- 11797] . The autophosphorylation of EnvZc and the phosphotransfer from phosphorylated EnvZc to OmpR were not affected by DNA , whereas the autodephosphorylation of OmpR-P was inhibited slightly . **We propose that the apparent inhibitory effect of DNA on the EnvZ phosphatase function is caused by sequestrating OmpR-P from the reaction as a result of OmpR-P binding to DNA .**

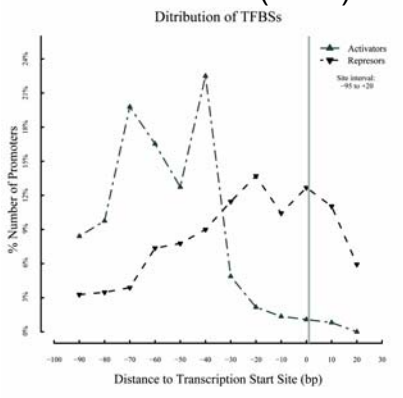
File ID pmid11158569 (Paper), Sentences 1 to 11: The critical role of DNA in the equilibrium between OmpR and phosphorylated OmpR mediated by EnvZ in Escherichia coli Ling Qin , Takeshi Yoshida , and Masayori Inouye Department of Biochemistry , Robert Wood Johnson Medical School , 675 Hoes Lane , Piscataway, NJ 08854 Edited by Carol A Gross , University of California , San Francisco , CA , and approved December 11 , 2000 (received for review August 10

Rodríguez-Penagos, C., Salgado, H., Martínez-Flores, I., and Collado-Vides, J. (2007) "Automatic reconstruction of a bacterial regulatory network using Natural Language Processing" *BMC Bioinformatics* 8: 293.



Proximal sites in σ70 promoters: -65 to +20 (1991)
enable direct contact with RNAP -95 to +20 (2008)

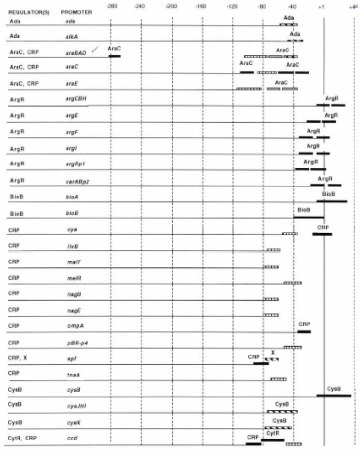
QuickTime™ and a
decompressor
are needed to see this picture.

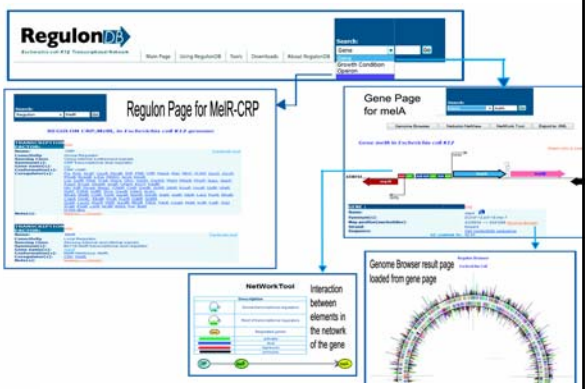


1991: 119 σ70 promoters
2008: 421 σ70 promoters with 1382 TF-BSs
Only 26 promoters (6%) lack a proximal site

Collado-Vides J., Magasanik B. and Gralla J.D. (1991) "Control site location and transcriptional regulation in Escherichia coli" Microbiol. Reviews. 55:371-394
Collado-Vides et al., (2009) "Bioinformatics resources for the study of gene regulation in bacteria" *J. of Bacteriology* (inaugural issue in Computational Biology) -in press

Origins of RegulonDB: 1991 and 1998 -ten years ago





Collado-Vides J., Magasanik B. and Gralla J.D. (1991) "Control site location and transcriptional regulation in Escherichia coli" Microbiol. Reviews. 55:371-394

Huerta A.M.; Salgado H.; Thieffry D., and Collado-Vides J. (1998) "RegulonDB: A Database on Transcription Regulation in *Escherichia coli*" *Nucleic Acids Res.* **26**: 55-60

Different sigma-types of promoter predictions

Sigma Factor	No. of Promoters	Promoters w-Strong Evidence*
Sigma24	61	60
Sigma28	20	5
Sigma32	31	22
Sigma38	81	38
Sigma19	1	1
Sigma54	32	16
Total	226	142

Sigma	# Predictios	# Regions	Cutoff (μ - X σ)	Accuracy (sensitivity, precision)
Sigma24[a]	39	39	-1.9 (0.14 – 2.* 0.98)	0.85 (0.75,0.95)
Sigma28	123	121	7.32 (10.8 – 1.* 3.6)	0.9 (0.9,0.9)
Sigma32	455	412	6.23 (9.56 – 1.* 3.2)	0.6 (0.5,0.6)
Sigma38	1767	1243	4.48 (5.6 – 1.* 2.2)	0.6 (0.6,0.5)
Sigma19	4	4	NA	-
Sigma54	152	147	7.10 (9.7 – 0.5 * 3.2)	0.8 (0.7,1.0)
Total	3,316	2,609	-	-

Araceli Huerta

[a]Rhodius et al. 2006 PLoS

Position Weight matrices

Counts

position	1	2	3	4	5	6	7	8	9	10	Sum
A	2	3	19	0	1	1	1	17	5	3	52
C	0	2	0	16	1	1	17	1	11	3	52
G	3	12	0	1	17	0	0	0	1	9	43
T	14	2	0	2	0	17	1	1	2	4	43

$$w_{i,j} = \ln\left(\frac{f'_{i,j}}{p_i}\right)$$

Weight matrix

position	1	2	3	4	5	6	7	8	9	10	Sum
A	-0.80	-0.43	**1.35**	-2.99	-1.38	-1.38	-1.38	**1.24**	0.05	-0.43	-6.15
C	-3.00	-0.80	-3.00	**1.18**	-1.39	-1.39	**1.24**	-1.39	0.81	-0.43	-8.16
G	-0.43	0.89	-3.00	-1.39	**1.24**	-3.00	-3.00	-3.00	-1.39	0.61	-12.47
T	**1.05**	-0.80	-2.99	-0.80	-2.99	**1.24**	-1.39	-1.39	-0.80	-0.16	-9.03
sum	-3.18	-1.13	-7.64	-4.00	-4.53	-4.53	-4.53	-4.53	-1.33	-0.41	-35.80

Weight of a sequence segment, computed from the weight matrix

residue r	T	G	T	A	A	T	A	A	T	A
W(r)	1.05	0.89	-2.99	-2.99	-1.38	1.24	-1.38	1.24	-0.80	-0.43

Weight -5.558 =SUM[W(r)]

Olivier Sand, Jean Valéry Turatsinze and Jacques van Helden

(a) Annotated TrpR binding sites

Site ID																			Target Operon
ECK120012644	G	T	A	C	T	A	G	T	T	T	G	A	T	G	G	T	A	T	G	aroL-yaiA-aroM
ECK120012187	G	T	A	C	T	A	G	T	T	T	G	A	T	G	G	T	A	T	G	aroL-yaiA-aroM
ECK120012179	G	A	A	C	T	A	G	T	T	A	A	C	T	A	G	T	A	C	G	trpLEDCBA
ECK120012892	G	A	A	C	T	A	G	T	T	A	A	C	T	A	G	T	A	C	G	trpLEDCBA
ECK120012181	G	A	A	C	T	A	G	T	T	A	A	C	T	A	G	T	A	C	G	trpLEDCBA
ECK120012636	G	T	A	C	T	A	G	A	G	A	A	C	T	A	G	T	G	C	A	aroH
ECK120012183	G	T	A	C	T	A	G	A	G	A	A	C	T	A	G	T	G	C	A	aroH
ECK120012185	G	T	A	C	T	C	G	T	G	T	A	C	T	G	G	T	A	C	A	mtr
ECK120012979	G	T	A	C	T	C	G	T	G	T	A	C	T	G	G	T	A	C	A	mtr
ECK120012894	G	T	A	C	T	C	T	T	T	A	G	C	G	A	G	T	A	C	A	trpR

(b) Position specific scoring matrix

A	0	3	10	0	0	7	0	2	0	6	7	2	0	6	0	0	8	0	5
T	0	7	0	0	10	0	1	8	6	4	0	0	9	0	0	10	0	2	0
C	0	0	0	10	0	3	0	0	0	0	0	8	0	0	0	0	0	8	0
G	10	0	0	0	0	0	9	0	4	0	3	0	1	4	10	0	2	0	5

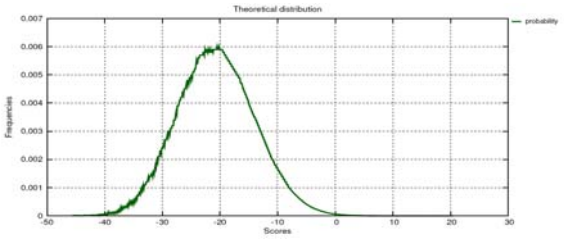
(c) Consensus

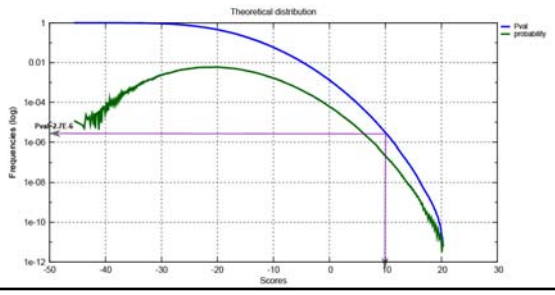
G	w	A	C	T	m	G	t	k	w	r	C	t	r	G	T	r	C	r

(d) Sequence logo

Theoretical ditribution of *scores*

- ***matrix-distrib*** (RSAT) computa la distribución teórica de scores completa.





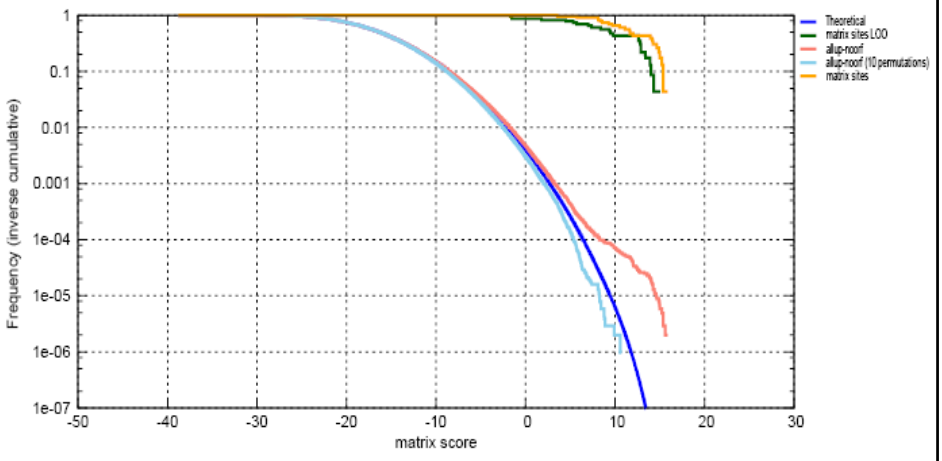
P(W=w|M): *probability of finding a score equal to w.*

P(W >= w|M): Probability of finding by chance a score larger o equal to w. We use this as a p-value.

We expect by random a score equal to 10 every 270kbs

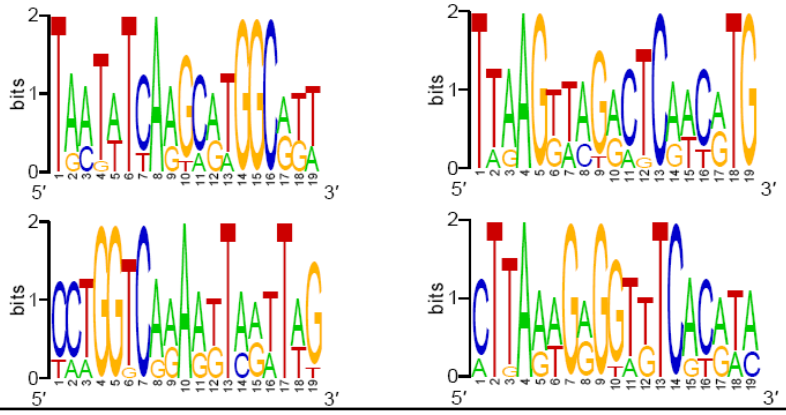
Empirical score distribution

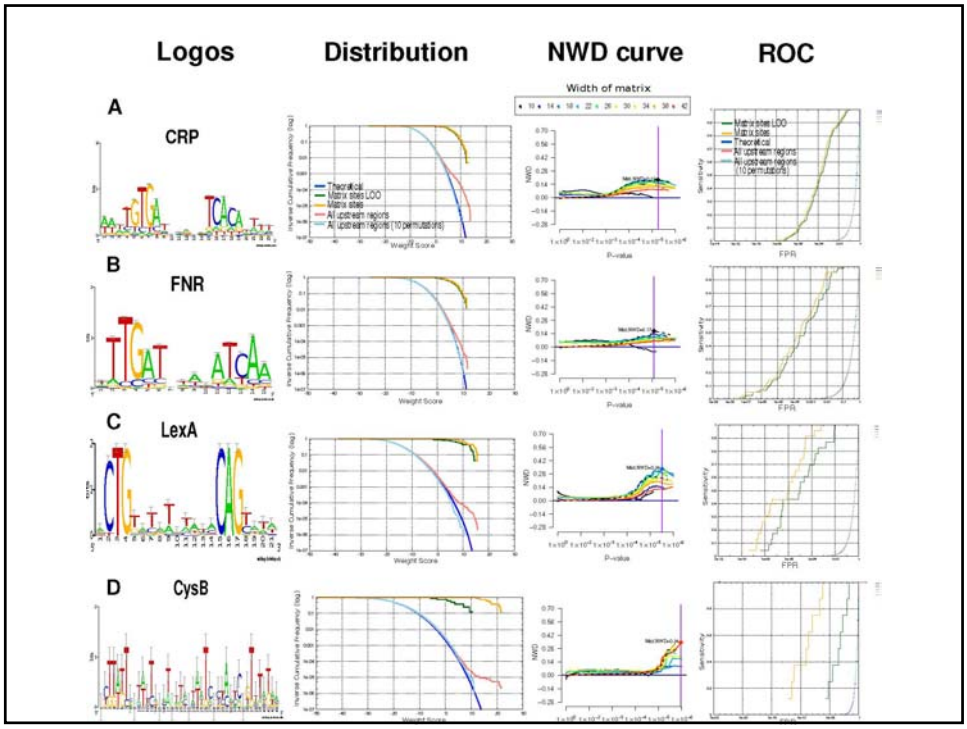
- 10 permutations



Shuffled Matrices

(e) Logos of column-permuted matrices



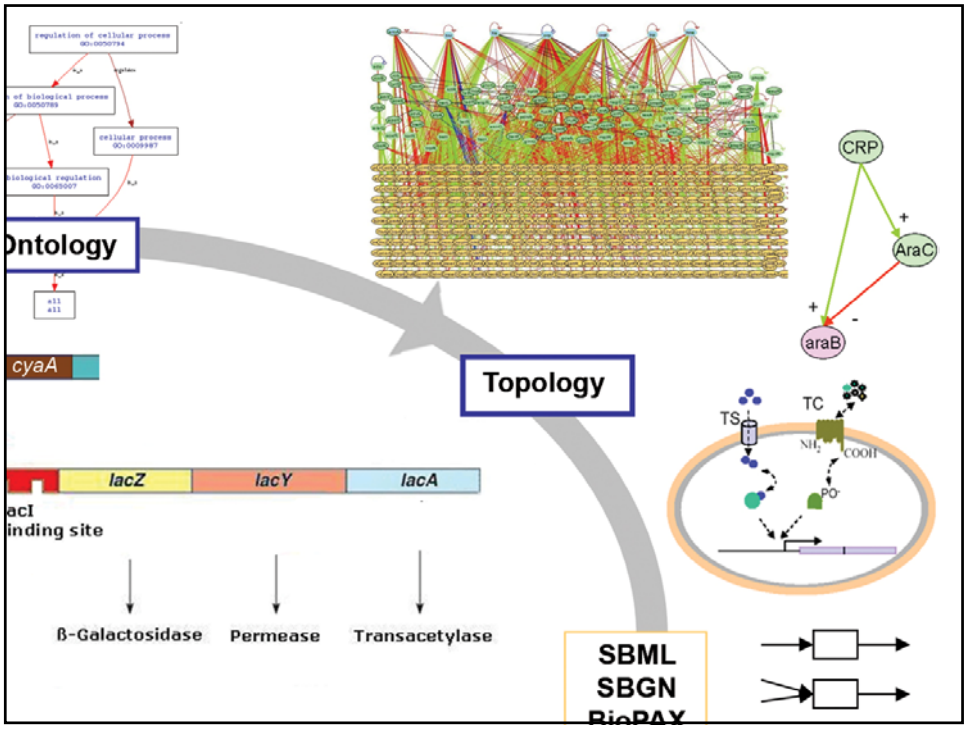


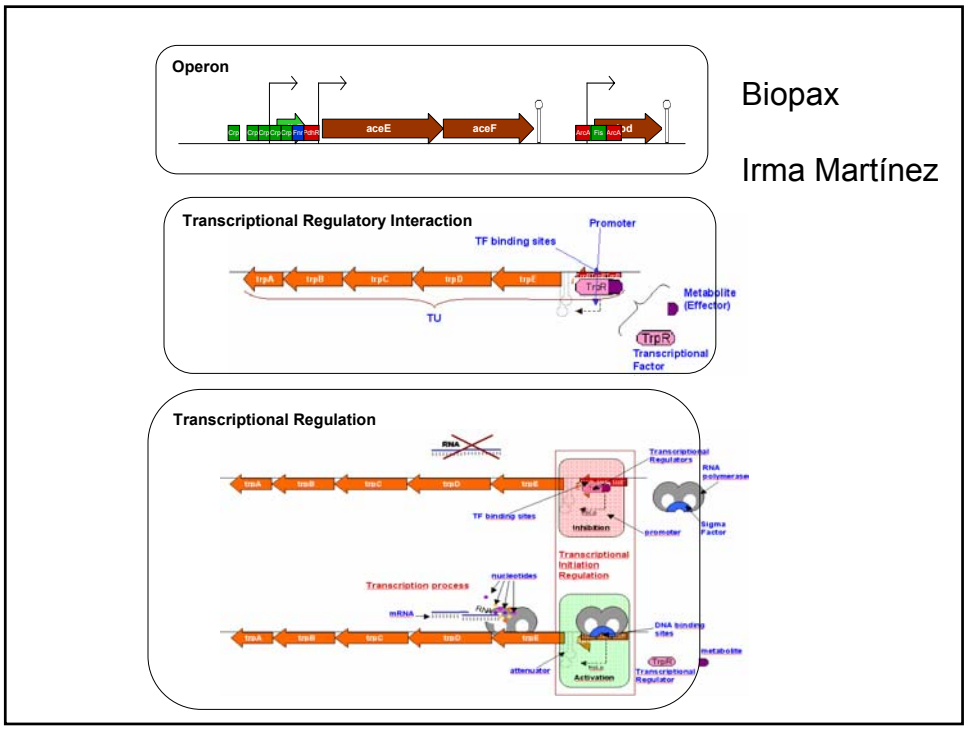
A comprehensive repertoire of PWMs for 54 TFs

- P-values 1.60E-05 or smaller

Medina-Rivera A., Abreu-Goodger C., Salgado H., Collado-Vides J., and van Helden J. "The good, the bad and the ugly: evaluating transcription factor binding motifs in a genome repertoire (submitted)

Factor	AnnotatedSites	PredictedSites	TP	PPV	Sensitivity	Recov. Percent
DgsA	7	10	7	0,7	1,0	100,0
FhlA	4	10	4	0,4	1,0	100,0
GcvA	5	7	5	0,7	1,0	100,0
HipB	4	41	4	0,1	1,0	100,0
MelR	10	42	10	0,2	1,0	100,0
MetR	6	17	6	0,4	1,0	100,0
UlaR	4	12	4	0,3	1,0	100,0
XylR	4	23	4	0,2	1,0	100,0
ArgR	30	70	28	0,4	0,9	93,3
GntR	13	37	12	0,3	0,9	92,3
Rob	6	23	5	0,2	0,8	83,3
NtrC	17	92	14	0,2	0,8	82,4
OxyR	11	42	9	0,2	0,8	81,8
GadE	5	5	4	0,8	0,8	80,0
NsrR	5	10	4	0,4	0,8	80,0
MetJ	25	73	18	0,2	0,7	72,0
ModE	7	12	5	0,4	0,7	71,4
MalT	20	20	14	0,7	0,7	70,0
TyrR	19	50	13	0,3	0,7	68,4
CysB	9	13	6	0,5	0,7	66,7
Fur	76	48	50	1,0	0,7	65,8
LexA	29	77	19	0,2	0,7	65,5
NagC	14	20	9	0,5	0,6	64,3
GlpR	23	63	14	0,2	0,6	60,9
CytR	10	15	6	0,4	0,6	60,0
TorR	8	9	4	0,4	0,5	50,0
TrpR	10	11	5	0,5	0,5	50,0
AgaR	11	12	5	0,4	0,5	45,5
IscR	10	25	4	0,2	0,4	40,0
PhoB	20	69	8	0,1	0,4	40,0
PurR	19	92	7	0,1	0,4	36,8
FNR	79	109	26	0,2	0,3	32,9
IclR	10	18	3	0,2	0,3	30,0
OmpR	22	40	6	0,2	0,3	27,3
ArgP	4	3	1	0,3	0,3	25,0
Nac	12	12	3	0,3	0,3	25,0
NanR	9	5	2	0,4	0,2	22,2
SoxS	18	34	4	0,1	0,2	22,2
RcsAB	10	12	2	0,2	0,2	20,0
MarA	16	21	3	0,1	0,2	18,8
AraC	18	19	3	0,2	0,2	16,7
CRP	241	176	40	0,2	0,2	16,6
DnaA	8	13	1	0,1	0,1	12,5
GalR	8	19	1	0,1	0,1	12,5
ArcA	89	37	9	0,2	0,1	10,1
IHF	100	69	10	0,1	0,1	10,0
FadR	12	15	1	0,1	0,1	8,3
Fis	219	134	5	0,0	0,0	2,3
NarL	80	24	1	0,0	0,0	1,3
CsgD	4	9	0	0,0	0,0	0,0
DeoR	7	7	0	0,0	0,0	0,0
FlhDC	20	25	0	0,0	0,0	0,0
FruR	12	28	0	0,0	0,0	0,0
GalS	7	14	0	0,0	0,0	0,0
Lrp	56	18	0	0,0	0,0	0,0
NarP	18	36	0	0,0	0,0	0,0
PhoP	20	65	0	0,0	0,0	0,0
RhaS	4	35	0	0,0	0,0	0,0

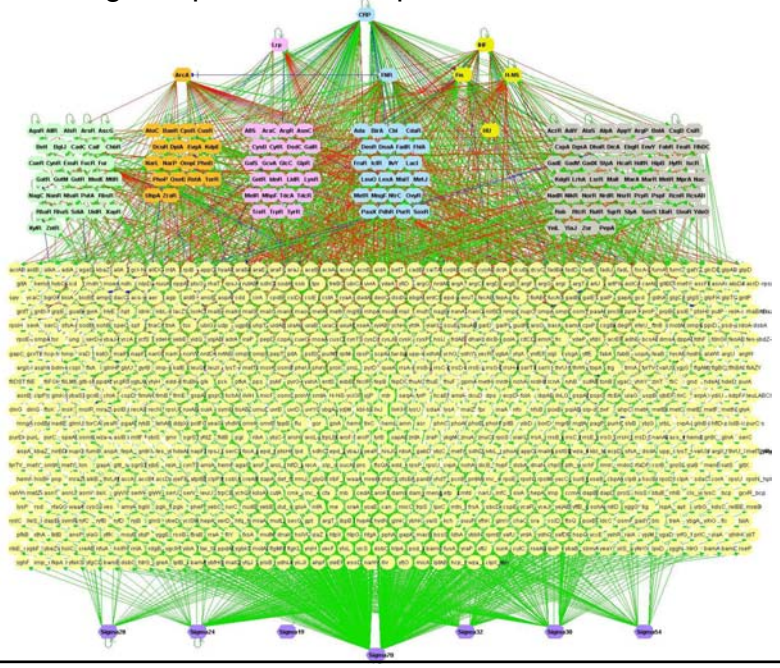




Biopax

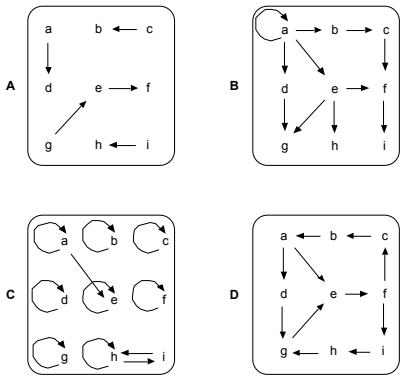
Irma Martínez

TFs and sigma specific transcription initiation



Topology of this updated network ?

Figure 1



Freyre-González J., Alonso-Pavón J.A., Treviño-Quintana L., and Collado-Vides J. "Functional architecture of *Escherichia coli*: new insights provided by a natural decomposition approach" (2008) *Genome Biology* 9: R154

Identifying global TFs by network decomposition (Julio Freyre)

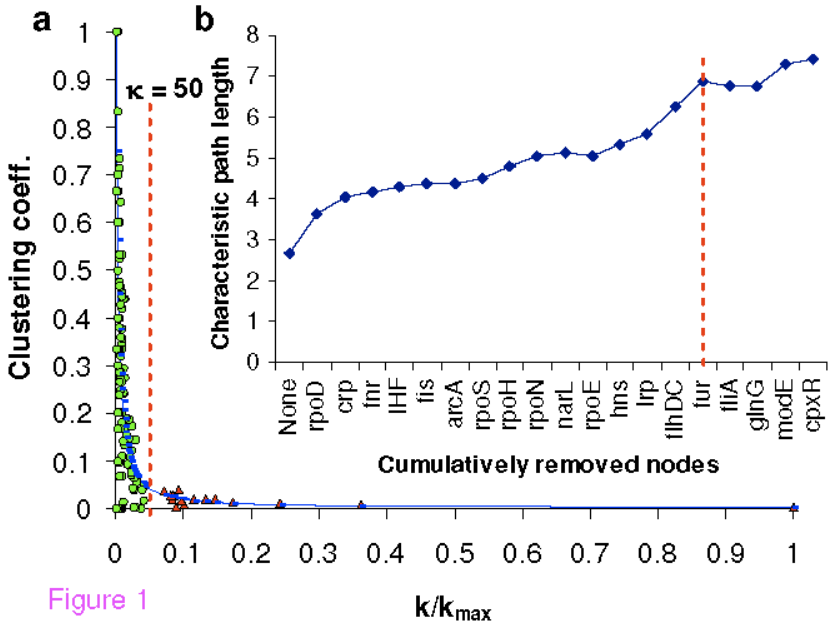
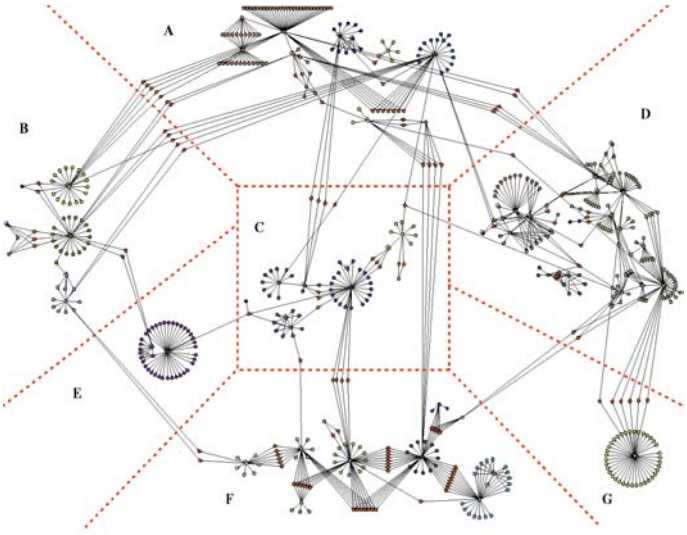
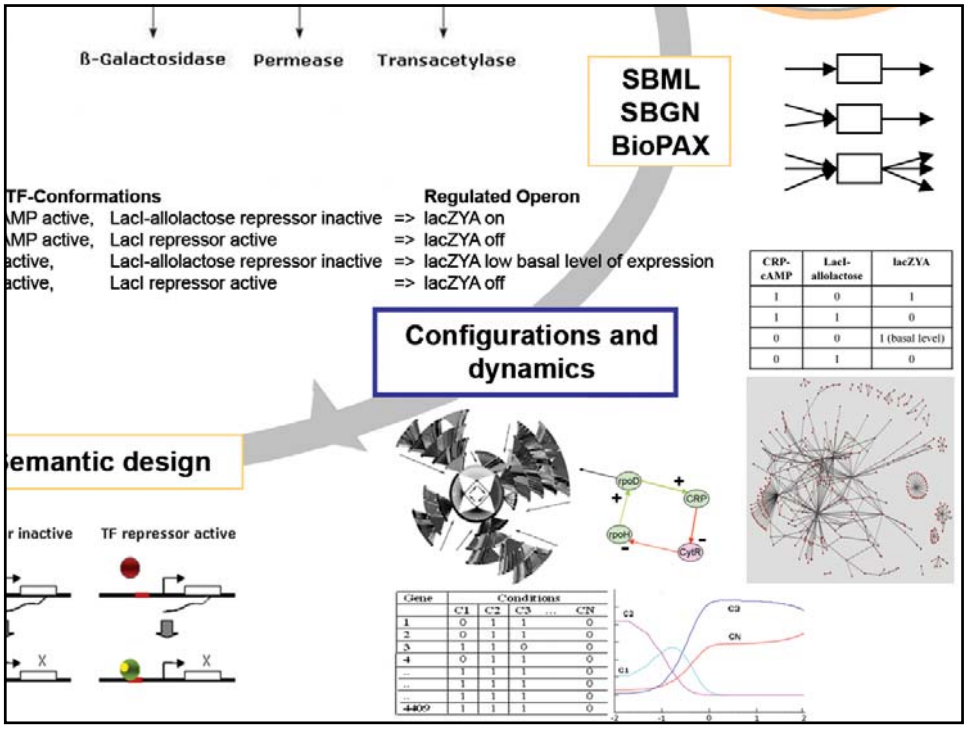


Figure 1



Region	Physiological function	Involved submodules
A	Nitrogen and sulfur assimilation	5.4, 5.5, 5.6, 5.r7, 5.r9, 5.r10, 5.r19
B	Aminoacids, nucleotides, and cofactors biosynthesis	5.8, 5.r5, 5.r14, 5.r15, 5.r24, 5.r25
C	Carbon sources catabolism	5.7, 5.9, 5.11, 5.13, 5.r12, 5.r17
D	Cellular stress response	5.2, 5.3, 5.r1, 5.r2, 5.r3, 5.r6, 5.10, 5.r21, 5.r26
E	Phosphorus assimilation and cell division	5.1
F	Respiration forms and oxidative stress	5.12, 5.r4, 5.r8, 5.r11, 5.r16, 5.r18, 5.r20, 5.r22, 5.r23
G	Motility	5.r13

Figure 2



Feedback loops in the regulatory network: several positive ones

Para ver esta película debe disponer de QuickTime™ y de un descompresor .

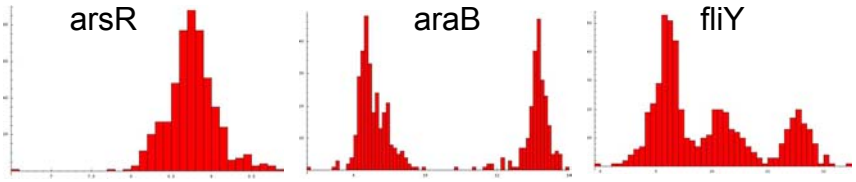
Freyre-González J., Alonso-Pavón J.A., Treviño-Quintana L., and Collado-Vides J. "Functional architecture of *Escherichia coli*: new insights provided by a natural decomposition approach" (2008) *Genome Biology* 9: R154

Phenotypic space: Discretization of expression -dialogue with RegulonDB.

Discretization-1: *An Absolute Reference*

To discretize the expression of a gene across many experiments, we used the complete dynamic range of expression as the reference.

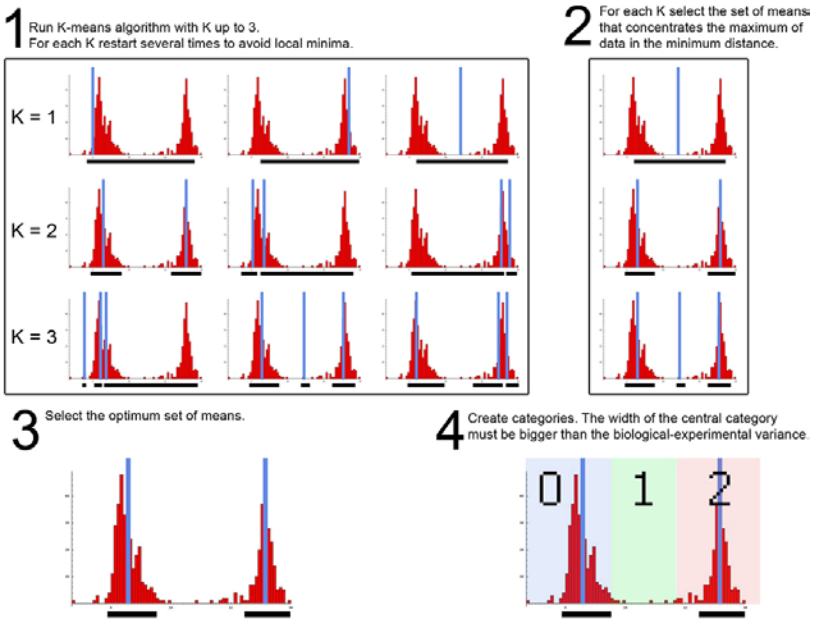
This will allow us to explore the plasticity of a gene's expression.



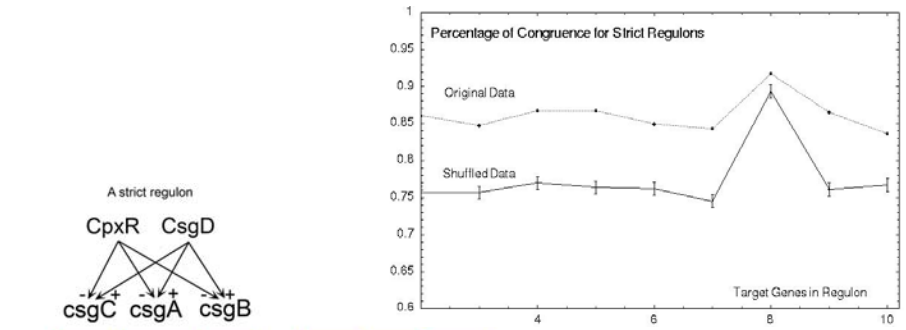
Histograms of the expression level of different genes in different conditions.
The dynamic range and the typical values are easily observed.

1) 380 experiments, 730 chips. Faith JJ et al., *Many Microbe Microarrays Database: uniformly normalized Affymetrix compendia with structured experimental metadata*. Nucleic Acids Research.

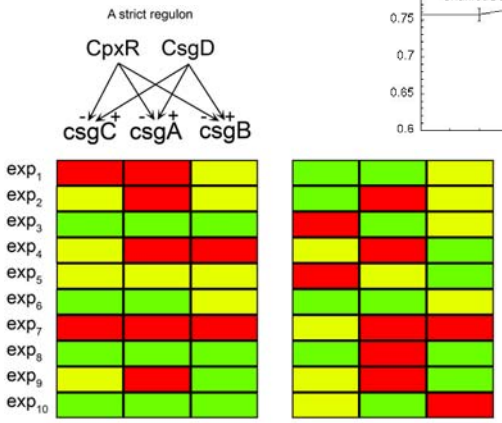
Discretization-2: *The Algorithm*



Discretization-3: *Measuring Congruence to Validate*



Error bars cover 95% of the variability in the shuffled data



Coherence in regulon expression is lost when we shuffle the data.

Discretization-4: *Discretized data will enable..*

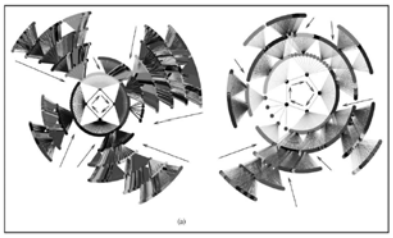
- Infere discrete regulatory functions.

Towards a discrete dynamical model of *E. coli*'s transcriptional regulatory network

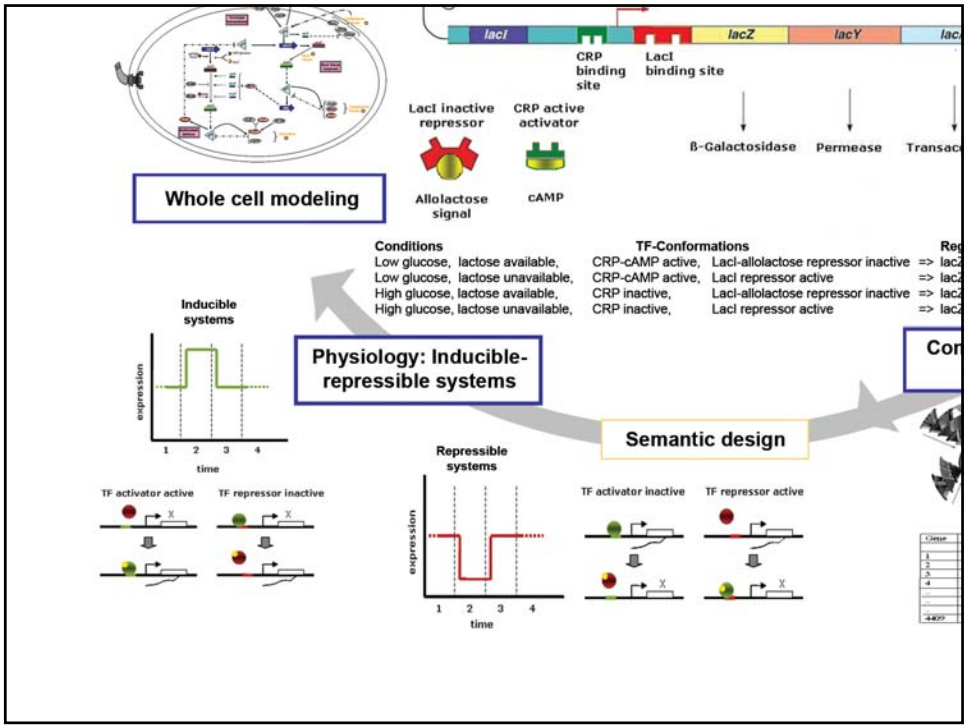
- Estimate the number of functional states (cell types).

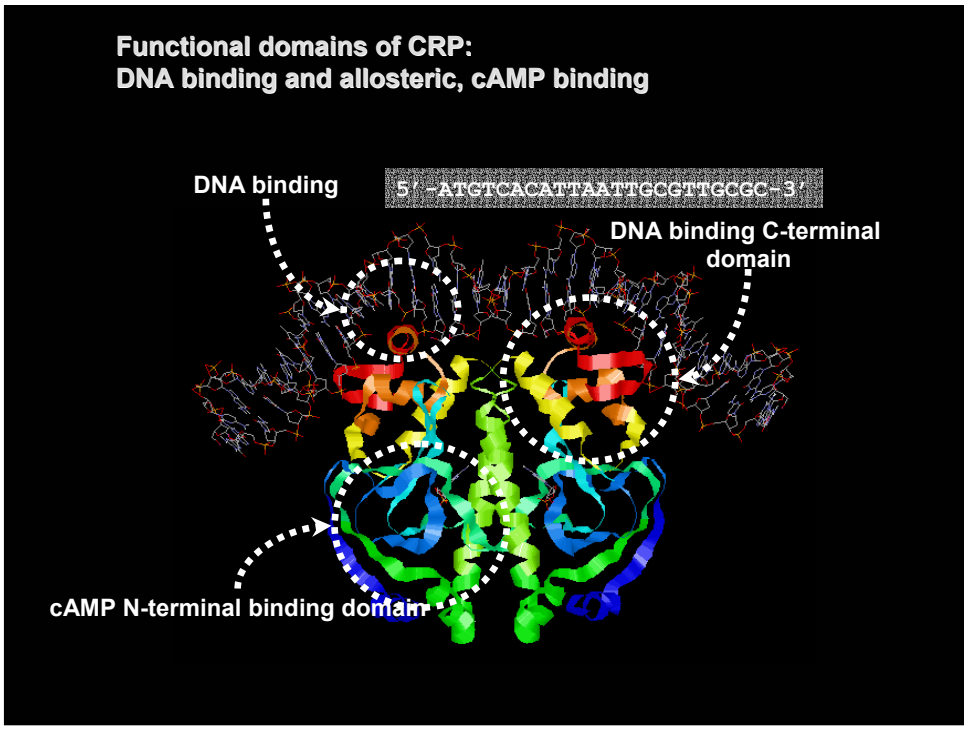
How many essentially different gene expression patterns does *E. coli*'s network encode?

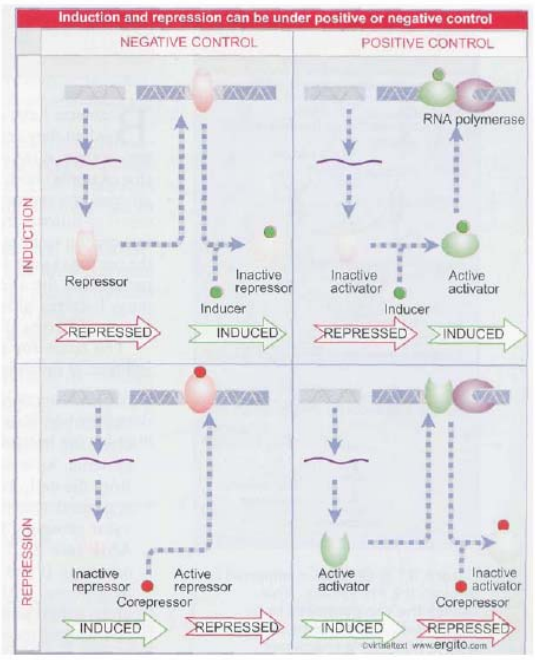
In the simulated *E.coli* network, We expect 3-5 steps to reach an attractor vs a larger number in a chaotic network .



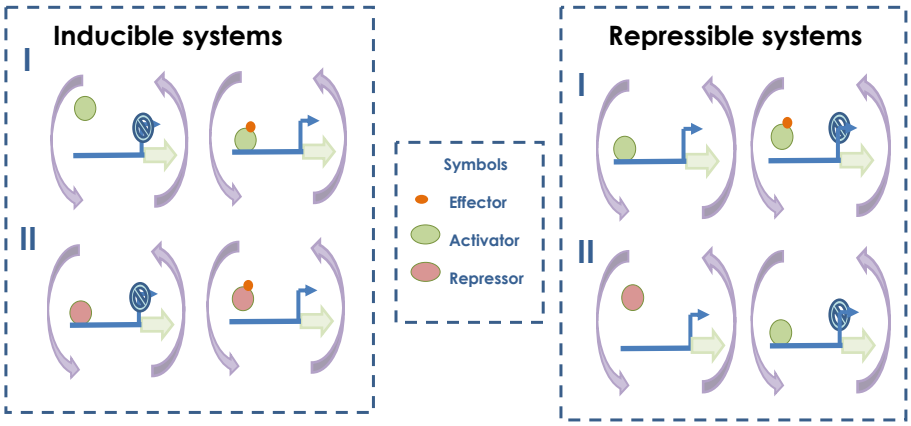
Santiago Sandoval and Maximino Aldana


lacI
lacZ
lacY
CRP binding site
LacI binding site
LacI inactive repressor
CRP active activator
ß-Galactosidase
Permease
Allolactose signal
cAMP
Whole cell modeling
Conditions
TF-Conformations
Low glucose, lactose available, CRP-cAMP active, LacI-allolactose repressor inactive
Low glucose, lactose unavailable, CRP-cAMP active, LacI repressor active
High glucose, lactose available, CRP inactive, LacI-allolactose repressor inactive
High glucose, lactose unavailable, CRP inactive, LacI repressor active
Inducible systems
Physiology: Inducible-repressible systems
Semantic design
Repressible systems
expression
time
TF activator active
TF repressor inactive
TF activator inactive
TF repressor active

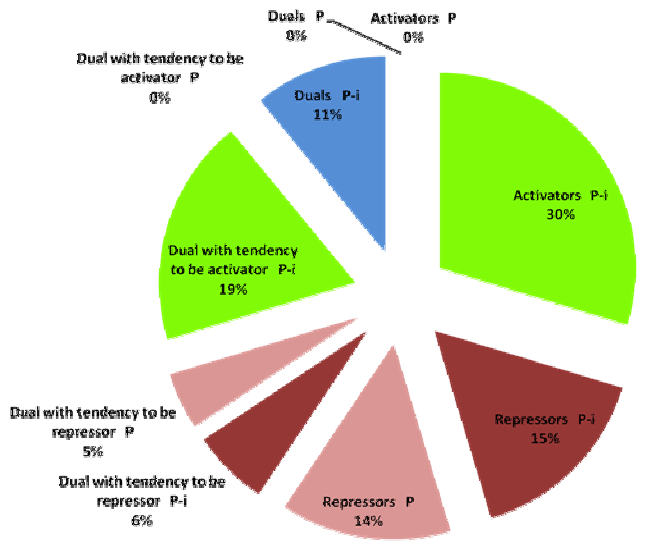

Functional domains of CRP:
DNA binding and allosteric, cAMP binding
DNA binding
5'-ATGTGACATTAATTGCCTTTCCGC-3'
DNA binding C-terminal domain
cAMP N-terminal binding domain


Induction and repression can be under positive or negative control
NEGATIVE CONTROL
POSITIVE CONTROL
INDUCTION
REPRESSION
RNA polymerase
Repressor
Inactive repressor
Inducer
Inactive activator
Active activator
Inducer
REPRESSED
INDUCED
REPRESSED
INDUCED
Inactive repressor
Active repressor
Corepressor
Active activator
Inactive activator
Corepressor
INDUCED
REPRESSED
INDUCED
REPRESSED
www.ergito.com

Lewin

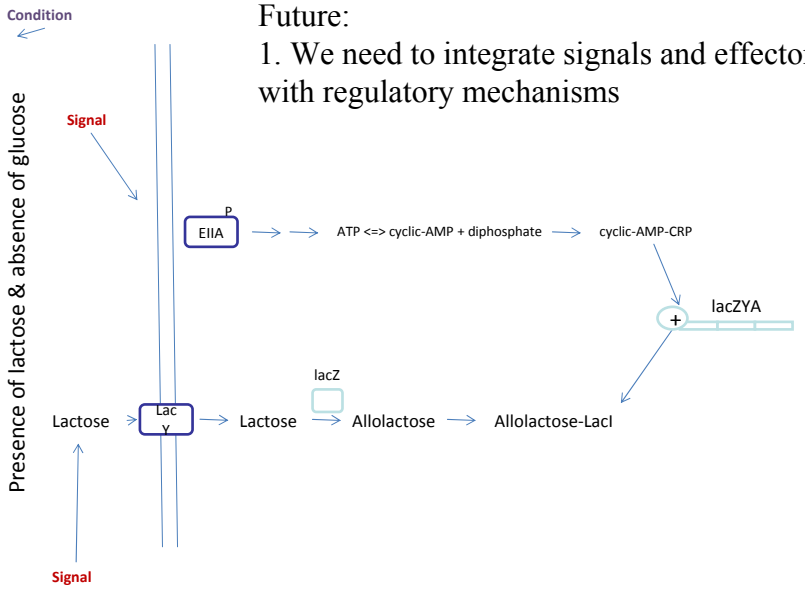

Inducible systems
I
II
Symbols
Effector
Activator
Repressor
Repressible systems
I
II

Preferences in active-DNA-conformation of transcriptional factors in *E. coli* K-12

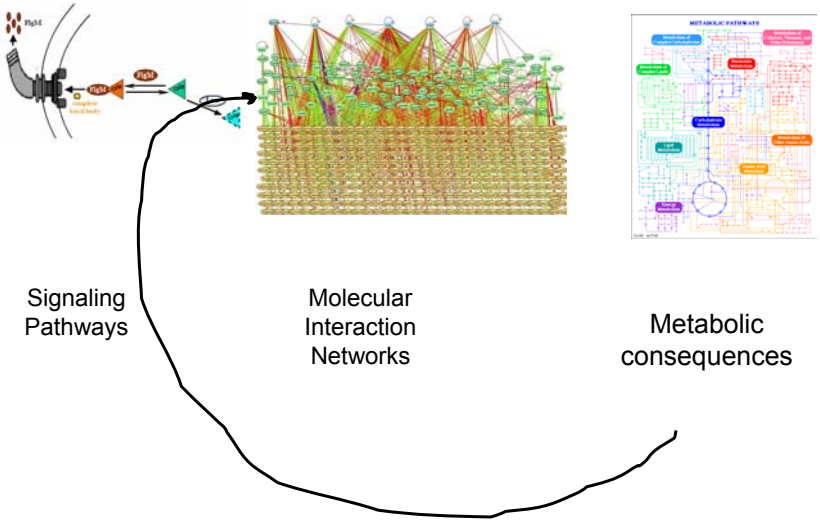


Future:

1. We need to integrate signals and effectors with regulatory mechanisms

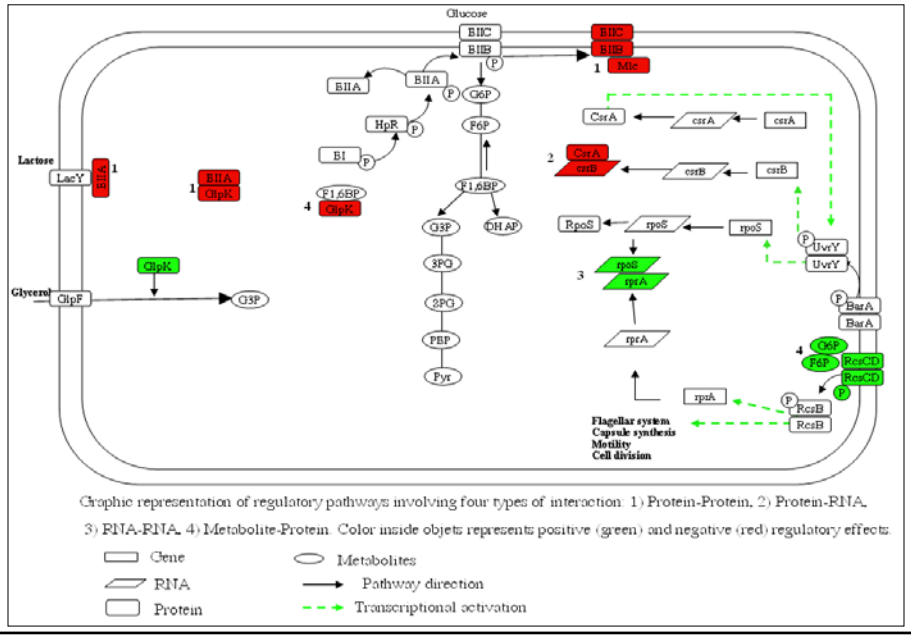


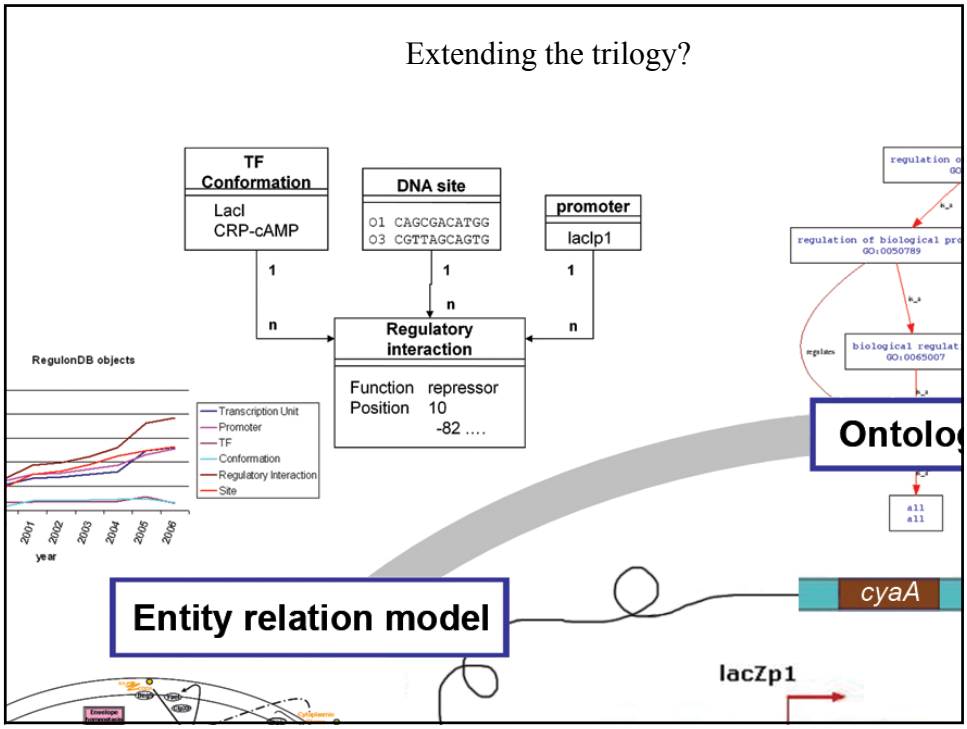
2. integrating signal sensing, regulation, affected metabolic capabilities and feedback.



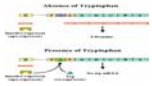
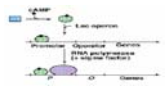
Hierarchy vs circular causality

3. Regulation: a variety of mechanisms, interactions and effects

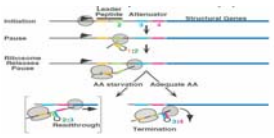
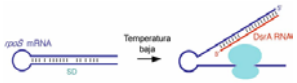


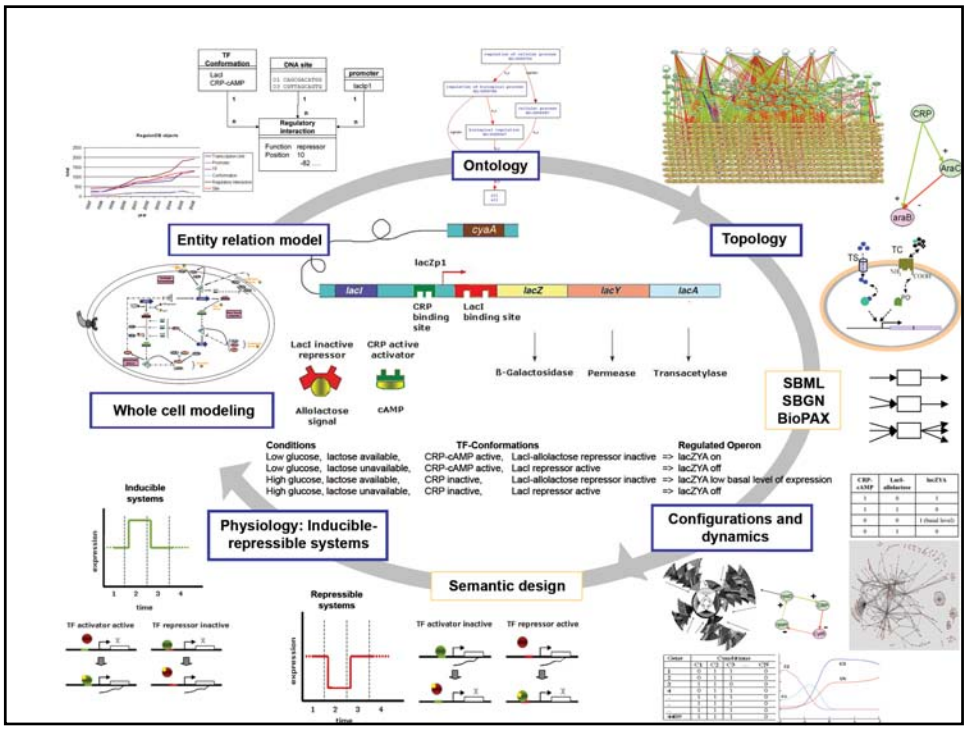


GENETIC REGULATION

Mechanism	Executors	Target site	Effect
Transcription			
Transcription Factors: Protein-DNA **REPRESSION** Example:  **ACTIVATION** Example: 	**Protein** TF: transcription factor	**DNA** TFBS: transcription factor binding sites. Operator sites	Regulates promoter activity.
Structural: DNA-Protein Supercoiling, Methylation of DNA 	**Protein** NAS: nucleoid associated proteins (IHF, HNS, HU)	**DNA** DNA can be altered as a result of interaction of skilled proteins	Modifies accessibility of promoter to DNA regions.
Sigma Factors: Protein-Protein Example: Sigma Factors 	**Protein** Provides specificity of RNAP-promoter	**Protein** RNAP core have to associates with different sigma factors to be activated under different environmental conditions.	Confers functionality and specificity to RNAP fpr promoter recognition

GENETIC REGULATION

Mechanism	Executors	Target Interaction	Effect
Transcription			
Transcriptional Attenuation: Protein-DNA Example: Attenuation in amino acid biosynthetic trp operon 	**RNA** RNA structure formed as stable transcript hairpin followed by a series of U's.	**Protein** The RNA polymerase gets stuck in the termination fork and finally it is released, stopping the transcription.	It affects termination of transcription readthrough at a single discrete site that precedes a regulated gene.
Posttranscriptional Regulation			
Antisense RNA: RNA-RNA **REPRESSION** Example:  **ACTIVATION** Example: 	**RNA** It acts by pairing with their target mRNAs.	**RNA** The target is the mRNA with complementary sequence	Regulation of synthesis of proteins, by affecting mRNA translation



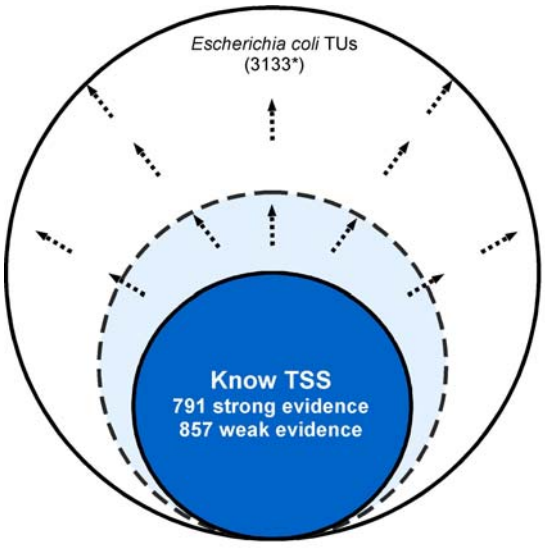
High throughput transcription start sites (TSSs) mapping. Active annotation

Aim: To enrich our understanding of transcriptional regulation by mapping as many TSS as possible to identify:

- Promoters
- Transcription Factors Binding sites
- Operon structure

In collaboration with Enrique Morett, IBT, UNAM

Our limited knowledge of promoters in *E. coli*



Magallanes Project:

Global Transcription Initiation Mapping and Active Promoter Annotation

RegulonDB

Transcriptional units database

Microarray database

Primer design

Experimental mapping procedure:

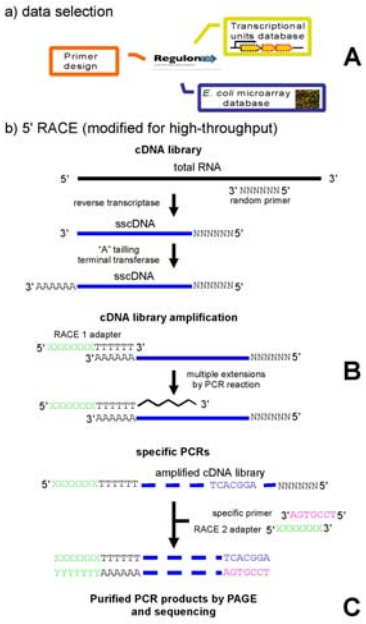
Directed Mapping of Transcriptional Start Sites (DMTSS) by 5' RACE (rapid amplification cDNA ends) modified for high-throughput

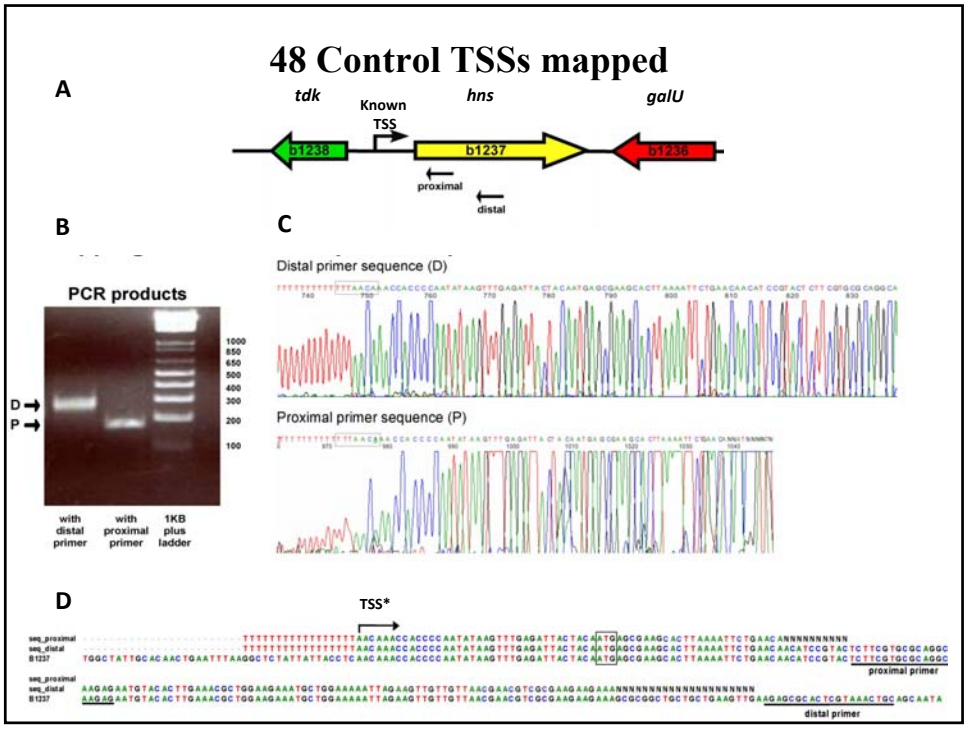
Results and data analysis

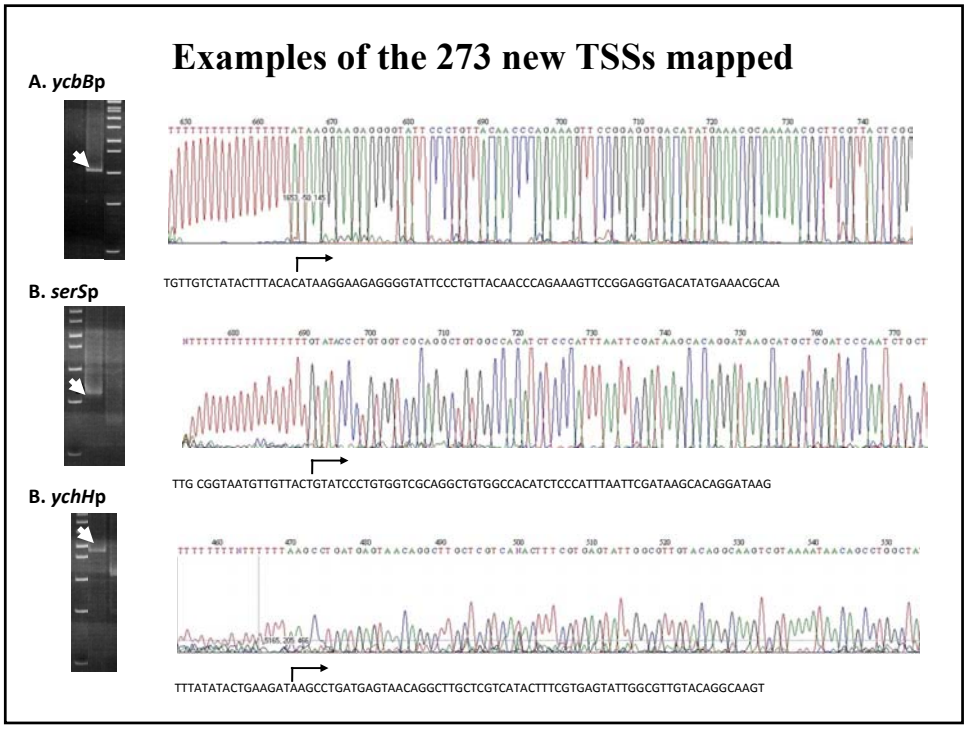
Active annotation in Regulon DB
- *Transcription start site
- *Promoter assignation
- *Sigma factor assignation

Etc.

Strategy to directed mapping of transcription start sites (DMTSS)

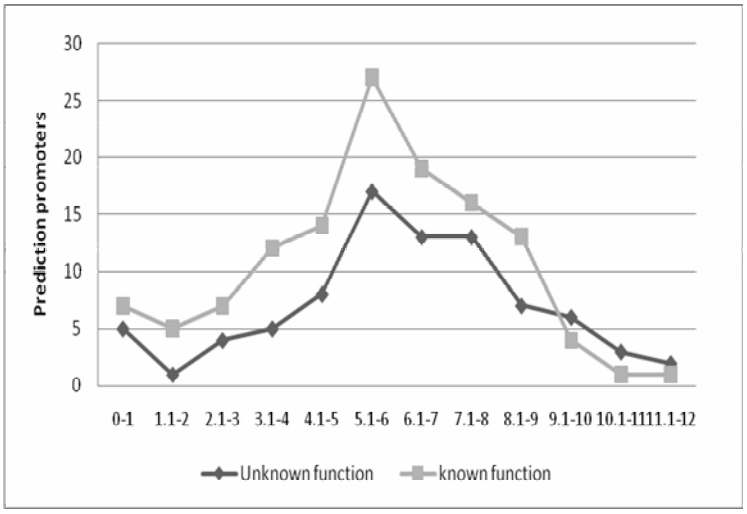



48 Control TSSs mapped
A
tdk
Known TSS
hns
galU
b1238
b1237
b1236
proximal
distal
B
PCR products
D
P
1000
850
650
500
400
300
200
100
with distal primer
with proximal primer
1KB plus ladder
C
Distal primer sequence (D)
Proximal primer sequence (P)
D
TSS*
proximal primer
distal primer


Examples of the 273 new TSSs mapped
A. ycbBp
TGTTGTCTATACTTTACACATAAGGAAGAGGGGTATTCCCTGTTACAACCCAGAAAGTTCCGGAGGTGACATATGAAACGCAA
B. serSp
TTG CGGTAATGTTGTTACTGTATCCCTGTGGTCGCAGGCTGTGGCCACATCTCCCATTTAATTCGATAAGCACAGGATAAG
B. ychHp
TTTATATACTGAAGATAAGCCTGATGAGTAACAGGCTTGCTCGTCATACTTTCGTGAGTATTGGCGTTGTACAGGCAAGT

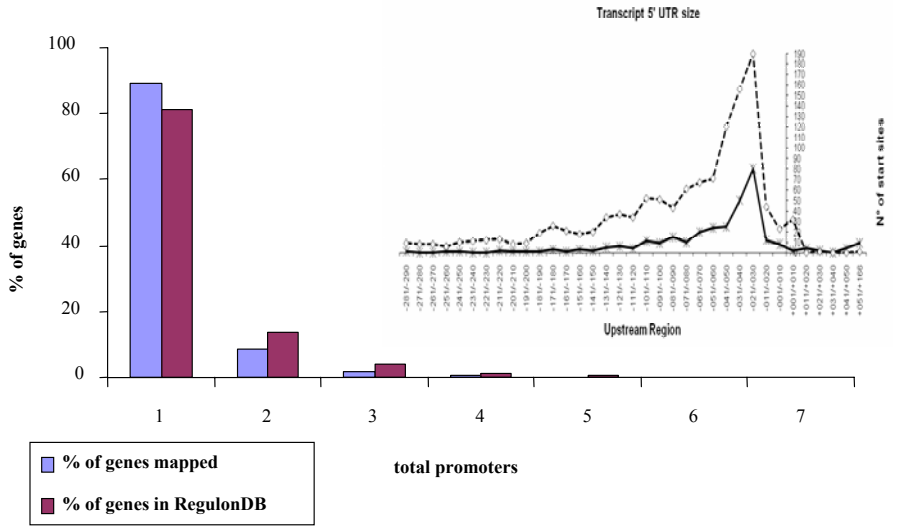
Promoter scores of genes with known or unknown function

About half of the new TSS are for genes with unknown function
Are they the product of real promoters?

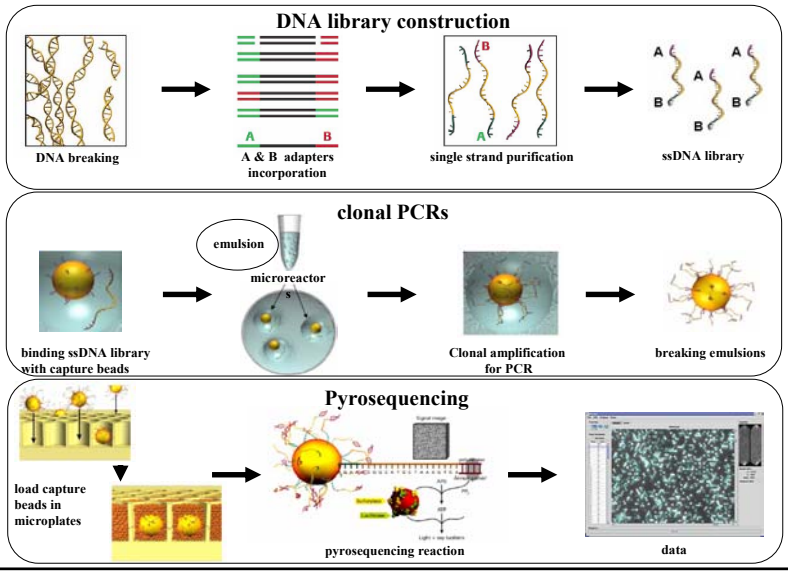


Total promoters by gene

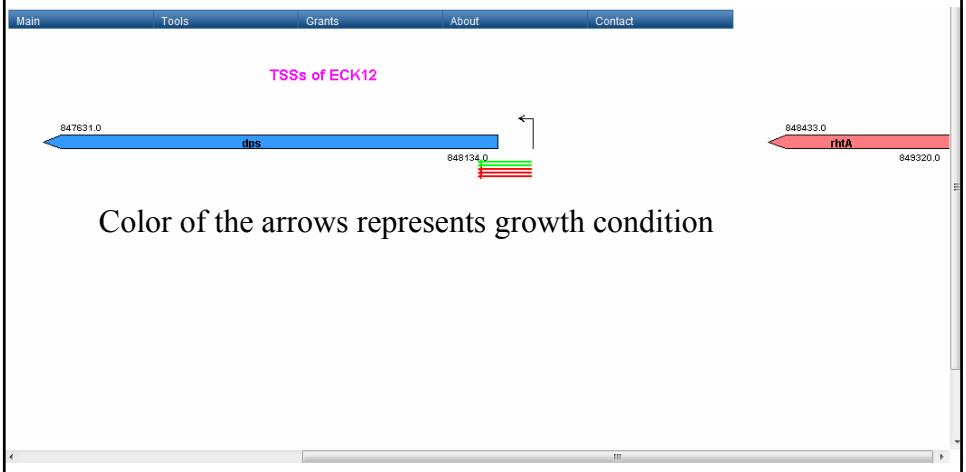
Distance of the TSS from the initial ATG codon



Sequencing with 454 GS20 instrument
Clonal and fast!



Examples of the pyrosequencing results:
Multiple sequences of known TSS.

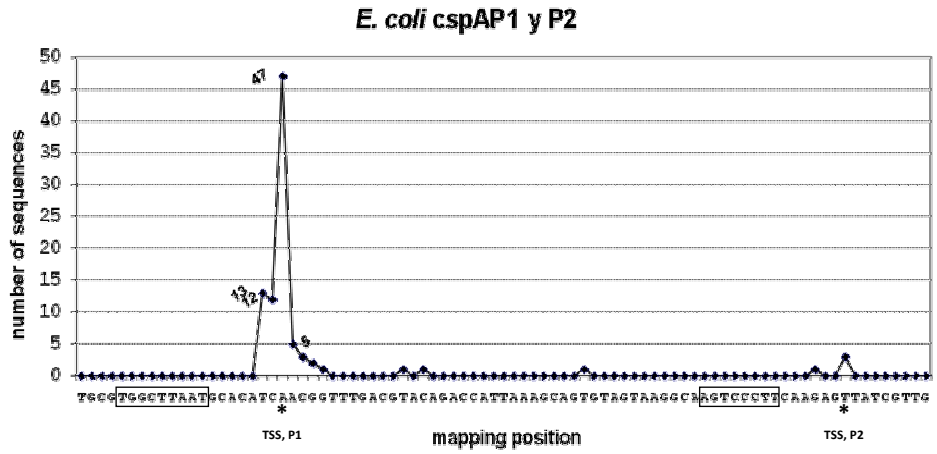


Color of the arrows represents growth condition

Confirmation of the recently mapped *dps* TSS by S. Busby's lab

Transcription not always starts at an exact nucleotide!

cspA



Results of Pyrosequencing

More than 1000 TSSs mapped. Only the high quality ones will be incorporated in RegulonDB

New information about relative expression in different conditions.

New information about mechanism of transcription initiation.

Computational Genomics Program

Funding: NIGMS, UNAM and Conacyt

Undergraduate program in genomics

4.4.5 Mapa curricular del plan de estudios propuesto

MAPA CURRICULAR DE LA LICENCIATURA EN CIENCIAS GENÓMICAS (2008)

SEMESTRE	ASIGNATURAS						TOTAL SEMESTRE
				NIVEL BÁSICO			
1	Matemáticas 1 (8C;4HT)	Principios de Programación (8C;3HT;2HP)	Biología Celular (8C;4HT)	Bioquímica (8C;4HT)	Biología Molecular (8C;4HT)	Seminario 1 (8C;3HT;2HP)	Créditos= 48 Horas Teóricas=22 Horas Prácticas=4 Asignaturas=6
2	Matemáticas 2 (8C;4HT)	Computación (8C;3HT;2HP)	Principios de Estadística (8C;3HT;2HP)	Genética (8C;4HT)	Principios de Evolución (8C;4HT)	Seminario 2 (8C;3HT;2HP)	Créditos= 48 Horas Teóricas=21 Horas Prácticas=6 Asignaturas=6
				NIVEL AVANZADO			
3	Matemáticas 3 (8C;4HT)	Bioinformática y Estadística 1 (8C;3HT;2HP)	Genómica Funcional 1 (8C;4HT)	Genómica Evolutiva 1 (8C;4HT)	Modelos Genómicos (8C;3HT;2HP)	Seminario 3 (8C;3HT;2HP)	Créditos= 48 Horas Teóricas=21 Horas Prácticas=6 Asignaturas=6
4	Matemáticas 4 (8C;4HT)	Bioinformática y Estadística 2 (8C;3HT;2HP)	Genómica Funcional 2 (8C;4HT)	Genómica Evolutiva 2 (8C;4HT)	Genómica Humana (8C;3HT;2HP)	Seminario 4 (8C;3HT;2HP)	Créditos= 48 Horas Teóricas=21 Horas Prácticas=6 Asignaturas=6
				NIVEL INTEGRATIVO			
5	Genómica Integrativa 1 (8C;3HT;2HP)	Genómica Integrativa 2 (8C;3HT;2HP)	Fronteras de la Genómica 1 (8C;4HT)	Fronteras de la Genómica 2 (8C;4HT)	Aplicaciones de la Genómica 1 (8C;4HT)	Aplicaciones de la Genómica 2 (8C;4HT)	Créditos= 48
6	Genómica Integrativa 3 (8C;3HT;2HP)	Genómica Integrativa 4 (8C;3HT;2HP)	Fronteras de la Genómica 3 (8C;4HT)	Fronteras de la Genómica 4 (8C;4HT)	Aplicaciones de la Genómica 3 (8C;4HT)	Aplicaciones de la Genómica 4 (8C;4HT)	
				NIVEL DE INVESTIGACIÓN			
7	Trabajo de Investigación 1 (6C;6HP)	Trabajo de Investigación 2 (6C;6HP)	Trabajo de Investigación 3 (6C;6HP)	Tópico Selecto 1 (10C;5HT)	Tópico Selecto 2 (10C;5HT)	Seminario de Investigación 1 (10C;4HT;2HP)	
8	Trabajo de Investigación 4 (6C;6HP)	Trabajo de Investigación 5 (6C;6HP)	Trabajo de Investigación 6 (6C;6HP)	Tópico Selecto 3 (10C;5HT)	Tópico Selecto 4 (10C;5HT)	Seminario de Investigación 2 (10C;4HT;2HP)	

ensum académico 3632

COLD SPRING HARBOR LABORATORY Ph.D. fellowship

QuickTime™ and a
decompressor
are needed to see this picture.