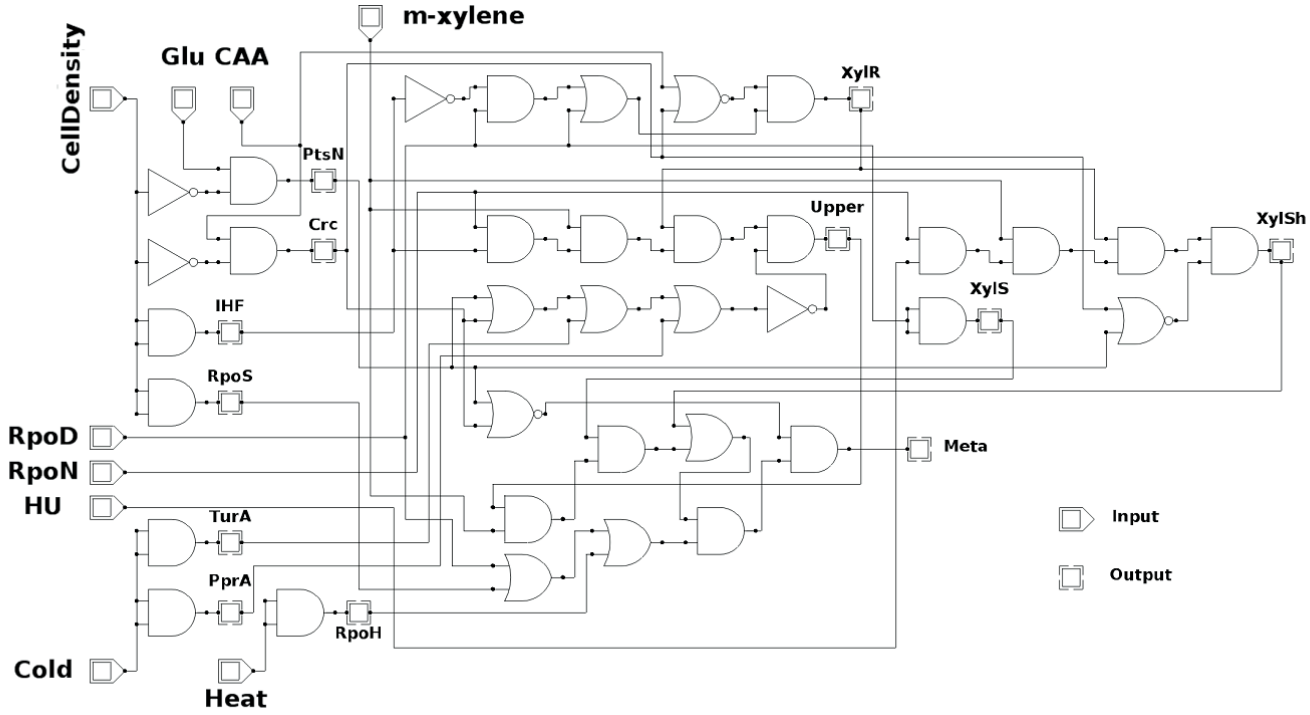
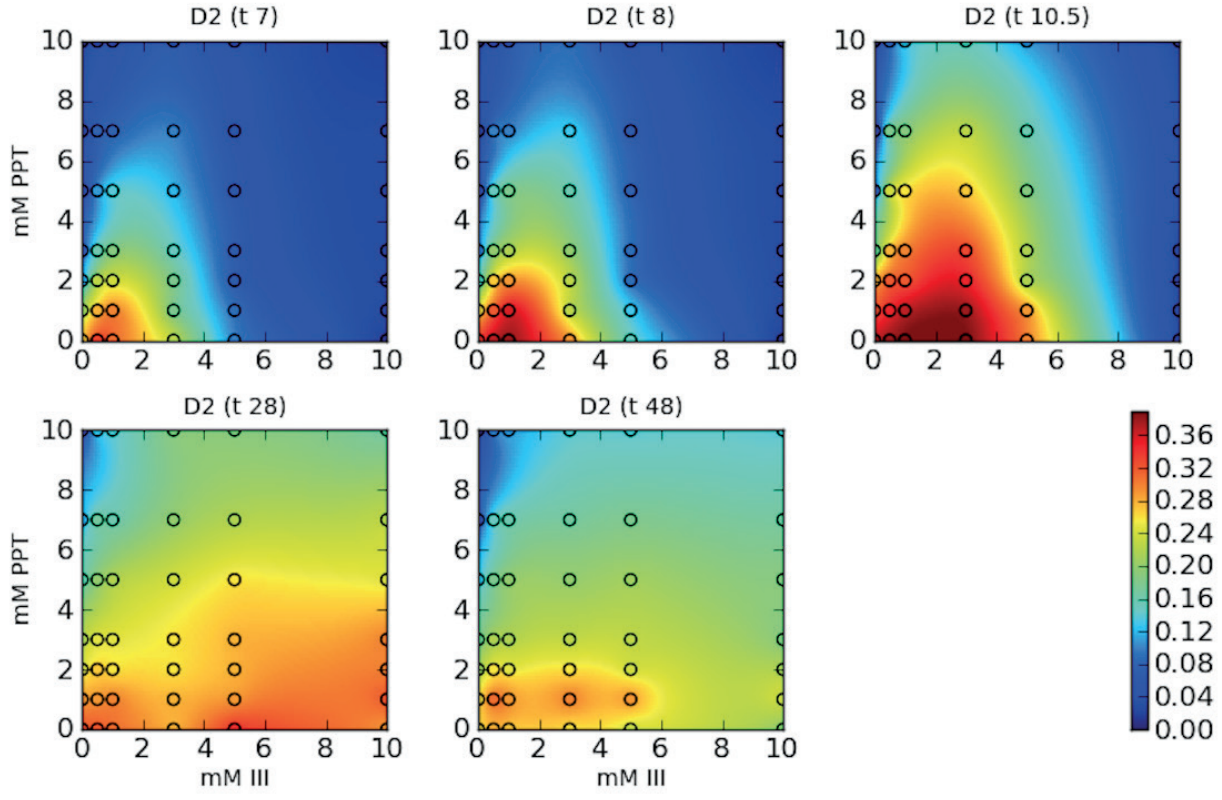

m-xylene
Glu CAA
CellDensity
XylR
PtsN
Crc
Upper
XylSh
XylS
IHF
RpoS
RpoD
RpoN
HU
Meta
TurA
PprA
RpoH
Cold
Heat
Input
Output


D2 (t 7)
D2 (t 8)
D2 (t 10.5)
D2 (t 28)
D2 (t 48)
mM PPT
mM III
0.36
0.32
0.28
0.24
0.20
0.16
0.12
0.08
0.04
0.00

6

SYSTEMS BIOLOGY PROGRAMME

One of the great conceptual novelties of recent years in the life sciences is systems biology. Although molecular biology is considered to have been founded by physicists, this circumstance did not result in a quantitative culture and the accurate, standardized descriptive language characteristic of the "hard sciences". With very few exceptions, the biosciences that have developed since that time seldom took the opportunity to formalize the mechanisms and functions of living systems using accurate languages and codes. Systems biology occupies this niche by analysing biological entities as comprehensible physicochemical objects with a functioning logic that can be modelled, understood and reshaped. By the same token, systems biology is not just a contemporary update of the recombinant DNA technologies of the past 30 years, with descriptive language imported from electrical and industrial engineering. It is also a new interpretative key for living systems as well as a declaration of intent regarding the use and reprogramming of biological objects for human benefit. In the same way that scientific chemistry, as initiated by Lavoisier, evolved into the chemical engineering that is the basis of our industrial society, biology has acquired a transforming potential that could lead to a type of industry and economy very different from the current paradigm. The CNB SysBio Programme figures in the contemporary landscape by developing active research lines in environmental genomics, network biology, systemic computation and metabolic engineering. This framework (which many consider a veritable paradigm shift) seeks to address the complexity of living systems as such, not to divide them into smaller parts (at difference from the reductionism of molecular biology). Systems biology offers remarkable scientific and technological potential for the field of biomedicine and for industrial, agricultural and environmental biotechnology.

HEAD OF PROGRAMME

Víctor de Lorenzo

OUR RESEARCH GROUPS

1. **Molecular environmental microbiology**
   Víctor de Lorenzo
2. **Evolutionary systems**
   Susanna Manrubia
3. **Computational systems biology**
   Florencio Pazos
4. **Logic of genomic systems**
   Juan F. Poyatos
5. **Microbiome analysis**
   Javier Tamames and Carlos Pedrós-Alió

Top: Complex regulatory networks can be recreated as as circuits of logic gates. The image shows the complex genetic device that controls expression of genes for degradation of m-xylene in the environmental bacterium *Pseudomonas putida.*

Bottom: Heat maps show relationship between resistance to the herbicide phosphinothricin and exposure to arsenic in the soil bacterium *P. putida.*

Molecular environmental microbiology

GROUP LEADER

Víctor de Lorenzo

SENIOR SCIENTISTS

Belén Calles
Esteban Martínez

POSTDOCTORAL SCIENTISTS

Yamal Al-Ramahí
David Rodríguez
Tomás Aparicio
Pavel Dvorak

PHD STUDENTS

Alberto Sánchez-Pascuala
M. Ángeles Hueso
Huseyin Tas
Elena Velázquez

TECHNICIANS

Sofía Fraile
Elena Algar

LAB EXECUTIVE ASSISTANT

Inés Merino

VISITING SCIENTISTS

Omar Ordóñez *(CONICET, Arg)*
Jhoanne Hansen *(U. Sao Paolo, Br)*
Stefano Donati *(Wageningen Univ.)*
Estela Ricaurte *(Princeton Univ., USA)*

SELECTED PUBLICATIONS

Nikel PI *et al.* The glycerol-dependent metabolic persistence of *Pseudomonas putida* KT2440 reflects the regulatory logic of the GlpR repressor. mBio 2015; 6: e00340-15

Nikel PI *et al. Pseudomonas putida* KT2440 metabolizes glucose through a cycle formed by enzymes of the Entner-Doudoroff, Embden-Meyerhof-Parnas, and pentose phosphate pathways. J Biol Chem 2015; 290: 25920-25932

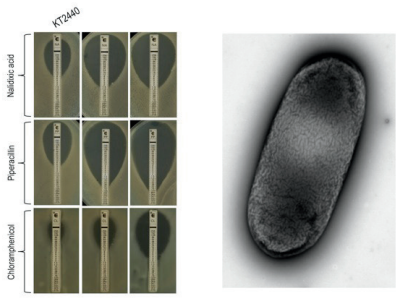
Guantes R *et al.* Transcription factor levels enable metabolic diversification of single cells of environmental bacteria. ISME J 2016; 10: 1122-1133

Goñi-Moreno A *et al.* An implementation-focussed bio/algorithmic workflow for synthetic biology. ACS Synth Biol 2016; 5: 1127-1135

Chavarría M, *et al.* A metabolic widget adjusts the phosphoenolpyruvate-dependent fructose influx in *Pseudomonas putida* KT2440. mSystems 2016; 1: e00154-16

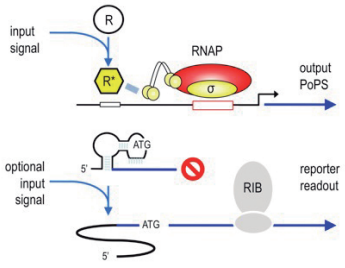
Our longstanding mission is to produce biological agents for biosensing, remediation and (when possible) valorization of chemical waste from urban and industrial activities that is otherwise dumped into the environment. To this end, we explore and capitalize on the interface between environmental microbiology and synthetic biology. Our workhorse is the soil bacterium *Pseudomonas putida*, which combines the ease of genetic programming typical of *Escherichia coli* with the safety, robustness and metabolic capabilities required in whole-cell catalysts for applications in harsh biotechnological settings. Specific activities include [i] development of *P. putida* as a reliable chassis for implantation of genetic and metabolic circuits. This involves profound editing of the extant genome of this microorganism to enhance desirable properties and eliminate drawbacks, as well as exploitation of surface-display systems for designing complex catalytic properties, altogether separated from cell metabolism, and the design of artificial communities by ectopic adhesins. [ii] Development of camel antibodies as tools for metabolic and regulatory engineering of *P. putida*. Various expression systems permit the targeting of camelid VHH fragments to the intracellular compartment, the periplasm or the external medium of the cells. This allows selective perturbations of chosen metabolic routes to enhance desired metabolic and regulatory qualities of bacteria. [iii] Genetic tools for deep refactoring of metabolic properties of *P. putida*. The list of new assets we are developing includes a large collection of standardized plasmid and transposon vectors as well as dedicated reporter systems for parameterization of gene expression flow and switching entire metabolic regimes. [iv] The Tol system borne by plasmid pWW0 as a reference for metabolic circuit implantation. The two operons for toluene and m-xylene biodegradation encoded in pWW0 offer a natural case of expansion of the metabolic repertoire of environmental bacteria through acquisition of new genes. [v] Deep metabolic engineering of *P. putida*. Currents efforts attempt to develop strains that can be entirely programmed to deliver catalytic phenotypes of choice after exposure to and computation of external and internal cues.

1



1 *Pseudomonas putida* strain depleted of certain cell surface structures. Note the spherocylindrical shape and increased sensitivity to some antibiotics.

2



2 Key parameters of the gene expression flow from DNA to proteins. PoPs (RNA polymerase per second passing through a promoter) and RiPs (ribosome per second going through an mRNA).

Evolutionary systems

GROUP LEADER

Susanna Manrubia

POSTDOCTORAL SCIENTISTS

Jacobo Aguirre
Juan Antonio García-Martín

PHD STUDENTS

Adriana Lucía-Sanz
Pilar Somovilla

MASTER'S STUDENT

Pablo Yubero

SELECTED PUBLICATIONS

Manrubia S, Cuesta JA. Evolution on neutral networks accelerates the ticking rate of the molecular clock. J R Soc Interface 2015; 12: 20141010

Capitán JA, Axelsen JB, Manrubia S. New patterns in human biogeography revealed by networks of contacts between linguistic groups. Proc R Soc Lond Ser B-Biol Sci 2015; 282: 20142947

Aguirre J, Manrubia S. Tipping points and early warning signals in the genomic composition of populations induced by environmental changes. Sci Rep 2015; 5: 9664

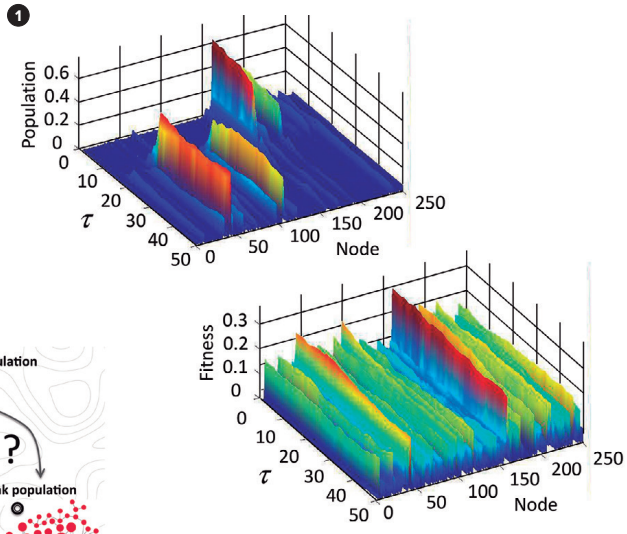
Iranzo J, Buldú JM, Aguirre J. Competition among networks highlights the power of the weak. Nat Commun 2016; 7: 13273

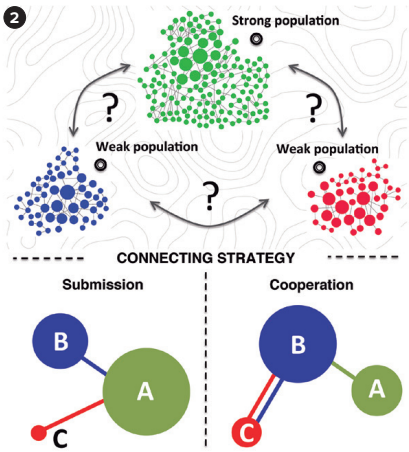
Puente-Sánchez F, Aguirre J, Parro V. A novel conceptual approach to read-filtering in high-throughput amplicon sequencing studies. Nucleic Acids Res 2016; 44: e40

The main interest of our group is the theoretical study of evolutionary systems of different kinds. We develop models inspired by the phenomenology observed in natural systems, chiefly molecular populations, viruses, and interacting agents at the cellular level and up. Our approach addresses the study of fundamental properties of adapting systems –with strong emphasis on their evolutionary origin— or, more specifically, tries to reproduce and eventually, to predict the response of such populations to endogenous and exogenous changes. In this context, we analyse the properties of the genotype-phenotype map through models such as the folded state of RNA sequences, focusing on the topological structure of neutral networks of genotypes and its relevance in adaptation and molecular innovation. Another main goal is to understand the survival strategies of viruses, including the relevance of multipartite genomes or the ecological effect of viral satellites. At a higher organizational level, we are also interested in modelization of the interaction between agents organized in networks that vie for resources, food, or mates, as competitive interactions represent one of the driving forces behind evolution and natural selection in biological systems. Finally, we explore the application of complex systems to biotechnology through the development of analysis techniques with environmental and health purposes. We have applied graph theory to antibody microarrays to improve the characterization of experimental samples, with direct application to allergy control, toxin detection in fresh water ecosystems, and identification of potential biomarkers in space missions. Our studies of viral response to antiviral treatments have determined optimal modes of drug administration to minimize viral load and mutant escape.

1

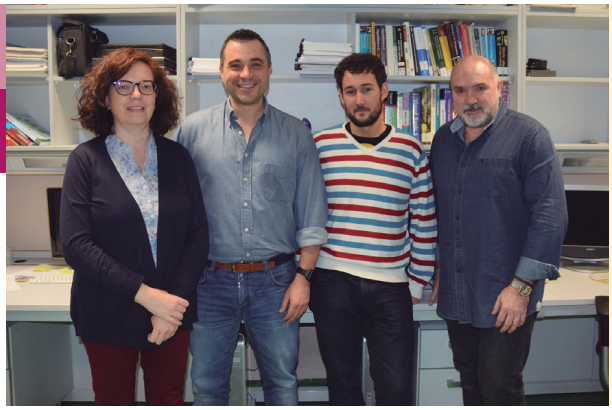
Slow environmental change, fast biological response. The genomic composition of populations (top) can undergo sudden shifts in response to smooth environmental changes (bottom) (modified from Aguirre & Manrubia, 2015).





2

The power of the weak. Cooperative interactions between two a priori weak populations can provide a collective advantage superior to direct alliances with a strong population (modified from Iranzo *et al.*, 2016.



Computational systems biology

GROUP LEADER

Florencio Pazos

SENIOR SCIENTIST

Mónica Chagoyen Quiles

PHD STUDENT

Javier López-Ibáñez Infante

MASTER'S STUDENT

Sara González Pérez

UNDERGRADUATE STUDENTS

Guillermo Martín Serrano
Mª Lourdes Rosano González
Diego Garrido Martín
Sara Fernández Novo

SELECTED PUBLICATIONS

Ochoa D, Juan D, Valencia A, Pazos F. Detection of significant protein co-evolution. Bioinformatics 2015; 31: 2166-2173

López D, Pazos F. Protein functional features are reflected in the patterns of mRNA translation speed. BMC Genomics 2015; 16: 513

López-Ibáñez J, Pazos F, Chagoyen M. MBROLE 2.0-functional enrichment of chemical compounds. Nucleic Acids Res 2016; 44: W201-W204

Chagoyen M, Pazos F. Characterization of clinical signs in the human interactome. Bioinformatics 2016; 32: 1761-1765

Oliveros JC, Franch M, Tabas-Madrid D, San-León D, Montoliu L, Cubas P, Pazos F. Breaking-Cas-interactive design of guide RNAs for CRISPR-Cas experiments for ENSEMBL genomes. Nucleic Acids Res 2016; 44: W267-W271

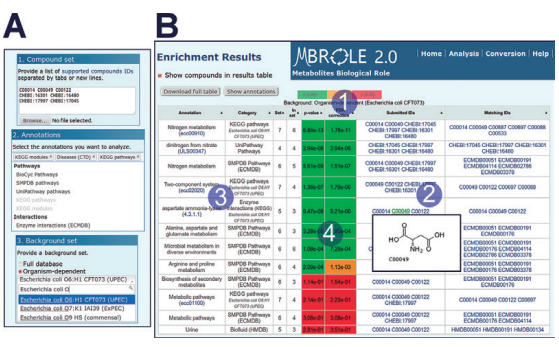
Our group is interested in different aspects of bioinformatics, computational biology and systems biology. Our goal is to obtain new biological knowledge with an *in silico* approach that complements the *in vivo* and *in vitro* methodologies of biology. This mainly involves mining the massive amounts of information stored in biological databases. Besides our lines of scientific research, we also collaborate with experimental groups by providing them with bioinformatics support for their specific needs, and participate in different teaching projects.

We have developed evolutionary-based methods for predicting sites with functional importance in protein sequences and structures. These methods are based on the fact that functional sites are subject to certain evolutionary constraints whose landmarks can be detected on multiple sequence alignments.

The biological functions of many proteins can only be explained in the context of their relationships with others. We have developed evolutionary-based methods for predicting interaction partners, which have been adopted by the scientific community. These methods are based mainly on the hypothesis that interacting or functionally related proteins adapt to each other during the evolutionary process (co-evolution). We try to detect the landmarks that this co-evolutionary process left in the sequences and structures of the proteins.

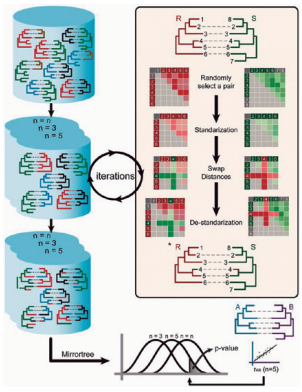
The study of living systems from a network perspective is providing new biological knowledge that could never have been obtained from the study of individual components (genes, proteins, etc.), no matter how detailed. This new strategy is needed to complement the intrinsic limitations of a reductionist approach that cannot keep up with the complexity of living systems. We are studying metabolic networks (central metabolism and biodegradation) and protein interaction networks with this new top-down approach. We also recently began to study human diseases using these new systemic and network-based approaches.

1

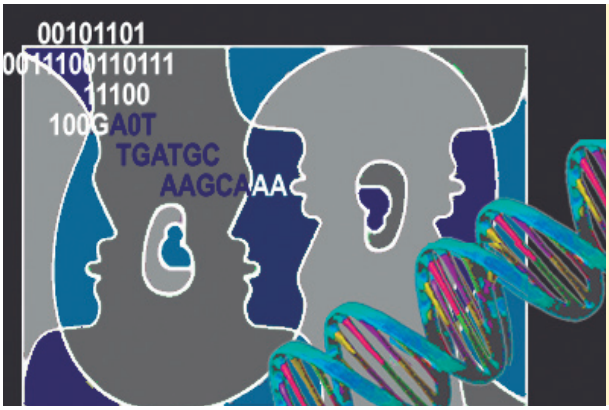


1 Interface of the MBROLE 2.0 system for analysis of metabolomics data

2



2 Scheme of the "p-mirrortree" (pMT) method for statistical evaluation of phylogenetic tree similarity scores. The method is based on generating a null distribution of tree similarity scores for pairs of shuffled trees from which p-values associated to a given score for a pair of real trees can be extracted.



Logic of genomic systems

GROUP LEADER

Juan F. Poyatos

PHD STUDENTS

Djordje Bajic
Clara Moreno-Fenoll

VISITING SCIENTIST

Guillermo Rodrigo
(IBMCP-CSIC, Valencia; iSSB, CNRS, France)

MASTER'S STUDENTS

Alvar Alonso
Ernesto Segredo

SELECTED PUBLICATIONS

Rodrigo G, Majer E, Prakash S, Daròs JA, Jaramillo A, Poyatos JF. Exploring the Dynamics and Mutational Landscape of Riboregulation with a Minimal Synthetic Circuit in Living Cells. Biophys J 2015; 109: 1070-6

Rodrigo G, Poyatos JF. Genetic Redundancies Enhance Information Transfer in Noisy Regulatory Circuits. PLoS Comput Biol 2016; 12: e1005156

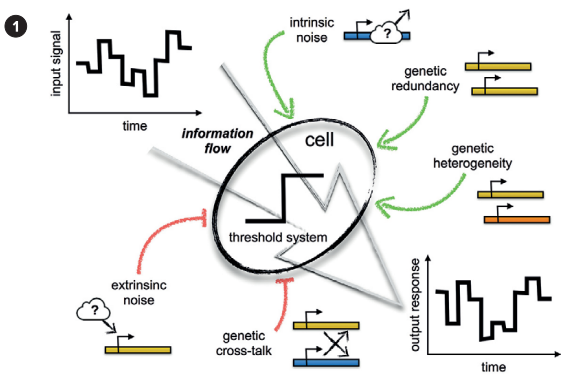
Rodrigo G, Bajic D, Elola I, Poyatos JF. Antagonistic autoregulation speeds up a homogeneous response in *Escherichia coli.* Sci Rep 2016; 6: 36196

We have focused on quantitatively understanding the functional and evolutionary implications of different modes of gene expression control. Our work could further motivate the search for fundamental design principles of gene regulation by non-coding RNAs, demonstrate how the application of concepts from information theory lead to more precise, quantitative understanding of cell systems, and uncover new general dynamic properties of interlinked feedback circuits applicable to bioengineering of synthetic units. Specifically, we studied:

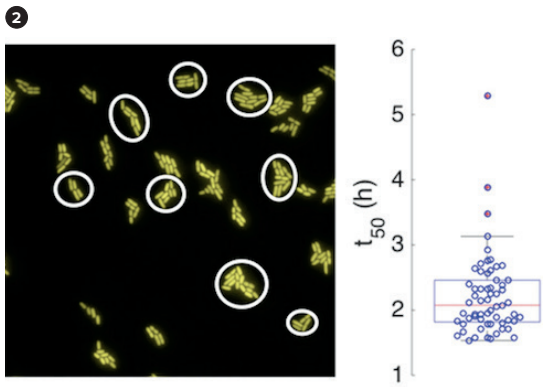
• A simple non-coding RNA circuit, known as riboregulator, to predict its behaviour and mutational landscape. We designed a synthetic riboregulator to then introduce a theoretical framework able to predict expression parameters and the functional consequences of mutations on the corresponding RNA sequences. This framework was tested experimentally by characterization of an RNA switch following the theoretical proposal.

• An original framework based on information theory to examine how multiple copies of a decision-making regulatory circuit turn noise into a beneficial feature, which suggests a novel adaptive function of genetic redundancy (Fig. 1). We tested how genetic redundancies allow information to be maximized, how this effect depends on device specifications, and the interdependence among them. The theory was corroborated by experimental data in several systems.

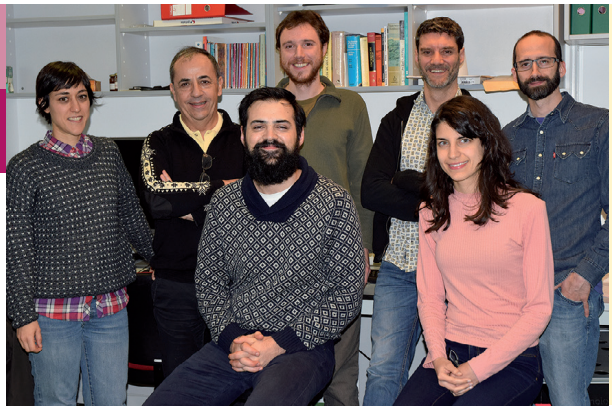
• An operation regime of a regulatory circuit exhibiting antagonistic positive and negative autoregulatory feedback loops. As a model system, we considered the circuit that controls the multiple antibiotic resistance phenotype in *Escherichia coli*. We combined theory with quantitative experiments at the population and single-cell levels (Fig. 2). In this architecture, weak positive feedback serves to increase the transcription rate "on the fly" due to derepression of the system and concomitant expression of the activator. This accelerates the response dynamics and limits the heterogeneity of the induced phenotype within a population.



1 Illustration depicting how intrinsic noise, genetic redundancy, and heterogeneity increase transmission of information by cellular decision-making circuit devices (working as threshold systems). Extrinsic noise and crosstalk among redundant units become limiting factors instead.



2 Subset of *Escherichia coli* lineages with different YFP expression levels after induction of the multiple antibiotic resistance phenotype with 5 mM salicylate. On the right, variation in response time (t50) quantified for 60 different lineages within the population.



Microbiome analysis

GROUP LEADERS

Javier Tamames
Carlos Pedrós-Alió

POSTDOCTORAL SCIENTISTS

Fernando Puente Sánchez
Arturo Marín
Beatriz Cámara

PHD STUDENTS

Marta Cobo Sebastián
Marta Royo-Llonch

SELECTED PUBLICATIONS

Nguyen D, Maranger R, Balagué V, Coll-Lladó M, Lovejoy C, Pedrós-Alió C. Winter diversity and expression of proteorhodopsin genes in a Polar Ocean. ISME J 2015; 9: 1835-1845

Tamames J, Sánchez PD, Nikel PI, Pedrós-Alió C. Quantifying the relative importance of phylogeny and environmental preferences as drivers of gene content in prokaryotic microorganisms. Front Microbiol 2016; 7: 433

Pinhassi J, DeLong EF, Béjà O, González JM, Pedrós-Alió C. Marine bacterial and archaeal ion-pumping rhodopsins: genetic diversity, physiology and ecology. Microbiol Mol Biol Rev 2016; 80: 1-26

Pedrós-Alió C, Manrubia S. The vast unknown microbial biosphere. Proc Natl Acad Sci USA 2016; 113: 6585-6587

We are interested in describing the rules that govern the assembly of microbial communities. Our objective is to achieve predictive capacity of the function of these communities, given their composition and the environmental factors in play. This will allow us to determine the conditions that favour particular combinations of species able to fulfil specific goals in biotechnological, clinical and ecological scenarios.

By combining information on the diversity and function of natural microbial communities, obtained by high-throughput sequencing, with mathematical modelling of the microbial metabolism, we try to understand how specific environmental changes influence the composition and function of the community.

We also aim to understand the ecology of aquatic microorganisms. We want to learn how microorganisms survive in unusual environments and thus derive clues regarding the function of more conventional habitats. Another interest is to comprehend the diversity of microorganisms, and especially the mechanisms that maintain a large variety of rare bacteria species in aquatic ecosystems.

Metagenomics and metatranscriptomics play a pivotal role in our work. Using these approaches, we obtain information on the presence and activity of specific organisms and their functions. We can also evaluate changes in composition, function and activity in response to environmental perturbations. These techniques also allow us to recover quasi-complete genomes that can be used to generate metabolic models for the species involved.

1

1

Sampling at Salar de Ascotán, North of Chile (altitude 3700 m)