



6

SYSTEMS BIOLOGY

Systems Biology is a conceptual framework for studying living systems that departs from the reductionism of molecular biology; it pursues the quantitative understanding of complete biological entities rather than the mere comprehension of their parts. One of the key goals of Systems Biology is to reveal the properties embodied in the inner organisation of complete biological objects.

The CNB SysBio Department 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 DEPARTMENT

Florencio Pazos

GROUP LEADER**Saúl Ares****PhD STUDENTS**Pau Casanova
Christian Camilo Cortés García**MASTER STUDENTS**

Gabriel Rodríguez Maroto

**SELECTED PUBLICATIONS**

Castro M, Ares S, Cuesta JA, Manrubia S. The turning point and end of an expanding epidemic cannot be precisely forecast. *Proc Natl Acad Sci USA* 2020; 117: 26190-26196.

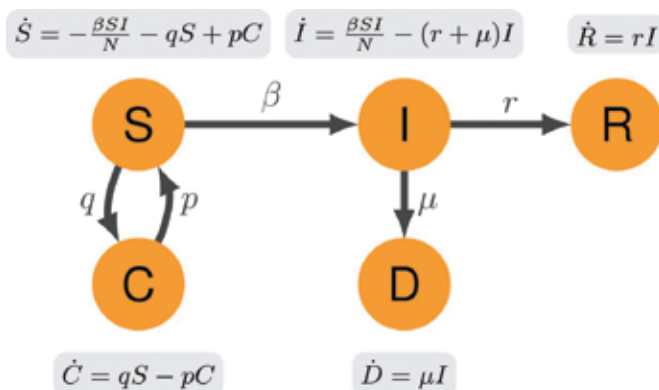
Clocks and rulers in life

We are interested in spatiotemporal phenomena in living systems: oscillations, pattern formation and dynamics of gene expression, using theoretical and computational methods derived from physics and mathematics. In the past year, following the epidemic emergency, our research has focused on epidemic dynamics, where we have presented a new model considering the effect of confinements, calculated the threshold over which lockdown measures inhibit infection spread, and shown that the predictive power of mathematical models of epidemic dynamics is limited by the exponential growth of uncertainties.

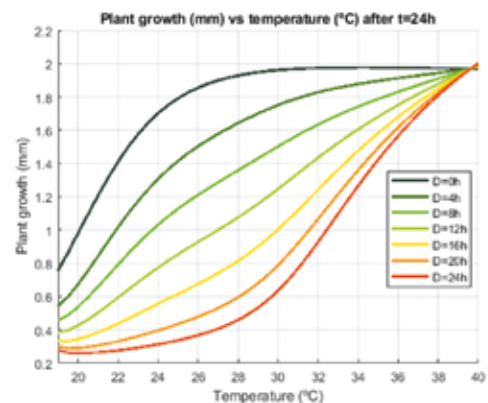
We have also been working on pattern formation in *Anabaena*, a filamentous cyanobacterium that differentiates specialized cells in the absence of fixed nitrogen. Plant research has also been an important topic: we have been working on the effect of light and temperature in plant growth, focusing on the embryonic stem, the hypocotyl, of *Anabaena thaliana*. We have also developed a theory for the regulation of the effect of nitrogen on the tillering of green revolution varieties of rice. Another relevant topic has been bacterial conjugation in Gram-positive bacteria. Conjugation is one of the mechanisms by which bacteria can exchange genetic material, in particular genes necessary to build antibiotic resistance.

Besides pure research, during 2020 we have made a great effort in science popularisation regarding epidemic dynamics, with a great number of TV, radio and newspaper contributions in Spanish and international media, including The Wall Street Journal, the French public radio, Süddeutsche Zeitung or almost all the major Spanish TV channels. On the Twitter account @omeuxeito we discuss and analyse almost on a daily basis epidemic data from the region of Madrid.

1



2



1 Diagram of the epidemic model along with the equations ruling the dynamics. Susceptible individuals (S) can enter and exit confinement (C) or become infected (I). Infected individuals can recover (R) or die (D). N is the total population. Rates for each process are displayed in the figure; q depends on specific measures restricting mobility and contacts, while p stands for individuals that leave the confinement measures (e.g., people working at essential jobs like food supply, health care, or policing), as well as for defection. We fit I to data on officially diagnosed cases, which are automatically quarantined: The underlying assumption is that the real, mostly undetected, number of infections is proportional to the diagnosed cases. From Castro et al. 2020.

2 Model predictions for hypocotyl growth (mm) as a function of temperature ($^{\circ}\text{C}$) after 24 hours for different number of light hours in the day, D . From the Master thesis of Gabriel Rodríguez Maroto.

GROUP LEADER**Víctor de Lorenzo****SENIOR SCIENTISTS**Belén Calles
Eteban Martínez**POSTDOCTORAL SCIENTISTS**Tomás Aparicio
M^o de los Angeles Hueso Gil**TECHNICIANS**Sofía Fraile
Elena Algar
Luis Almonacid**LAB ASSISTANT**

Inés Merino

PHD STUDENTSHuseyin Tas
Elena Velázquez**UNDERGRADUATE STUDENT**

Sherezade Cepeda

VISITING SCIENTISTSJoseph Mougous
(Univ. Washington, USA)
Jillian Silbert
(University of Princeton, USA)
Marcela Villegas
(Universidad del Valle, Colombia)
Natalia Rodríguez Fernández
(Universidad de Almería, España)**SELECTED PUBLICATIONS**

Hueso-Gil A, Nyerges A, Pál C, Calles B, de Lorenzo V. Multiple-site diversification of regulatory sequences enables inter-species operability of genetic devices. *ACS Synth Bio* 2019; 9: 104–114.

Calles B, Goñi-Moreno A, de Lorenzo V. Digitalizing heterologous gene expression in Gram-negative bacteria with a portable on/off module. *Mol Sys Biol* 2019; 15: e8777.

Martínez-García E, Goñi-Moreno A, Bartley B, McLaughlin J, Sánchez-Sampedro *et al.* SEVA 3.0: an update of the Standard European Vector Architecture for enabling portability of genetic constructs among diverse bacterial hosts. *Nucl Acids Res.* 2019; 48: D1164–D1170.

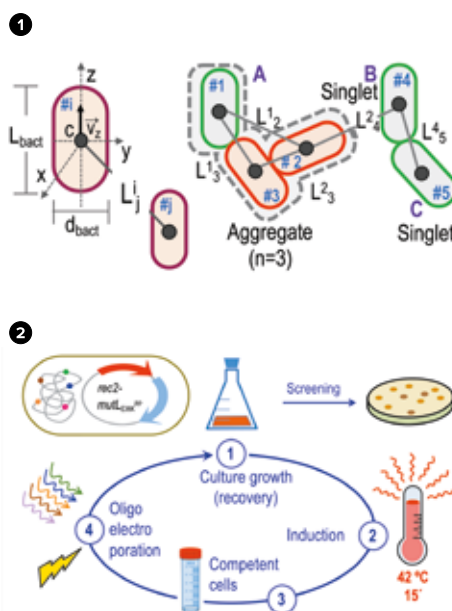
Espeso DR, Martínez-García E, de Lorenzo V. Quantitative assessment of morphological traits of planktonic bacterial aggregates. *Water Res* 2020; 188: 116468.

Dvořák P, Bayer EA, de Lorenzo V. Surface display of designer protein scaffolds on genome-reduced strains of *Pseudomonas putida*. *ACS Synth Biol* 2020; 9: 2749–2764.



Environmental synthetic biology

The longstanding mission our team is the production of biological agents for biosensing, large-scale remediation and valorisation of chemical waste that is otherwise dumped into the Environment by urban and industrial activities. The workhorse to this end is the soil bacterium *Pseudomonas putida*, which combines the ease of genetic programming that is 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 a profound editing of the extant genome of this microorganism for enhancing desirable properties and eliminating drawbacks. Also, the exploitation of surface-display systems for designing complex catalytic properties altogether separated from the cell metabolism and even the design of artificial communities by means of ectopic adhesins. [ii] Genetic tools for deep refactoring of metabolic properties of *P. putida*. The list of new assets that we are developing includes a large collection of standardized plasmid and transposon vectors as well as dedicated reporter systems for parameterization of the gene expression flow and for switching entire metabolic regimes. [iii] The TOL system borne by plasmid pWWO as a natural example of well-nested metabolic circuit implantation. The two operons for toluene and *m*-xylene biodegradation encoded in pWWO offer a case of expansion of the metabolic repertoire of environmental bacteria through acquisition of new genes. [iv] Deep metabolic engineering of *P. putida*. Currents efforts attempt to develop strains that can be entirely programmed to deliver catalytic phenotypes of choice upon exposure and computation of both external and internal cues. This endeavour combines direct rational engineering with fine-tuning of gene expression by means of site-specific diversification of genomic sequences of choice through adaptation to *P. putida* of high-efficacy genome engineering technology.



1 Modelling inter-cell interactions in microbial communities. The figure shows some steps followed by the in house designed computational workflow to identify aggregate clusters within confocal microscopy images. Basically, the starts by obtaining geometrical parameters of individual cells such as the mass center position (c), length (L_{bact}), diameter (d_{bact}) and the axial orientation vector (v_z). Then, in step 2 distances between cell pairs are computational arranged into a distance matrix (L^i_j represents the distance between the bacterial centers of $\#i$ to $\#j$), where each row contains all distance pair combinations of one bacterium to the rest of cells indicated ($L^1_2, L^1_3, \dots, L^1_n$).

2 Scheme of the high-efficacy multiple site genome editing (HEMSE) cycles. The main steps of the procedure are depicted: cultures of *P. putida* EM42 (pSEVA2314-rec2-mutL_{E36K}^{PP}) are grown and induced by a heat-shock; then competent cells are prepared and transformed with recombinering oligonucleotides. After recovery on fresh media cultures enter in the next round of HEMSE by applying the induction step. Screening of allelic replacements within a given cycle is performed after recovery by plating culture dilutions on the appropriate solid media

GROUP LEADER**Susanna Manrubia****POSTDOCTORAL SCIENTISTS**Jacobó Aguirre
Luis F. Seoane**PhD STUDENTS**Adriana Lucía-Sanz
Iker Atienza
Pilar Somovilla**MASTER STUDENT**

Henry P. Secaira

UNDERGRADUATE STUDENT

Fátima González



Evolutionary systems

SELECTED PUBLICATIONS

Catalán P, Elena SF, Cuesta JA, S. Manrubia S. Parsimonious scenario for the emergence of viroid-like replicons de novo. *Viruses* 2019; 11, 425.

Valdano E, Manrubia S, Gómez S, Arenas A. Endemicity and prevalence of multipartite viruses under heterogeneous between-host transmission. *PLoS Comput Biol* 2019; 15, e1006876.

Castro M, Ares S, Cuesta JA, Manrubia S. The turning point and end of an expanding epidemic cannot be precisely forecast. *Proc Natl Acad Sci USA* 2020; 117, 26190-26196.

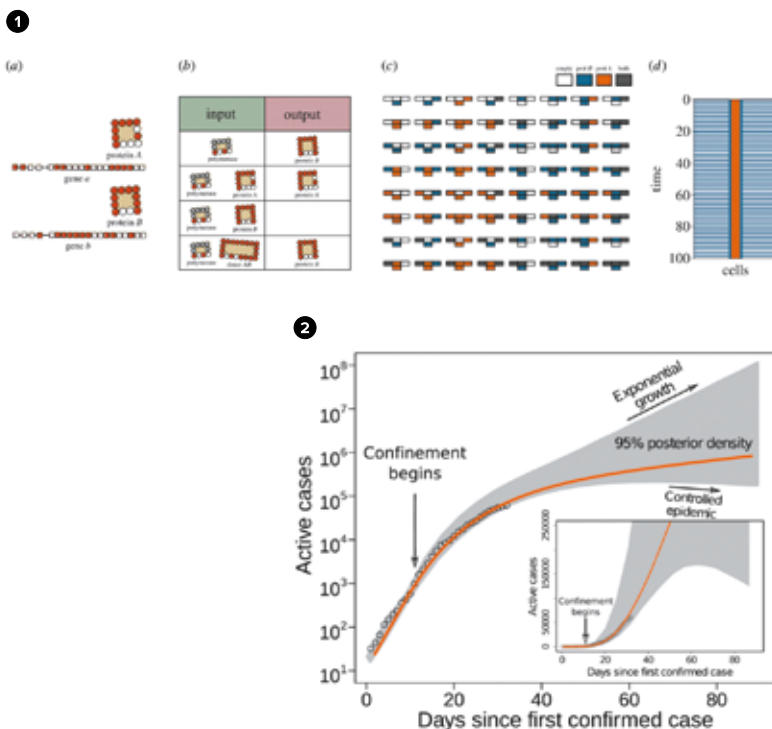
Catalán P, Manrubia S, Cuesta JA. Populations of genetic circuits are unable to find the fittest phenotype as a result of phenotypic bias in a multilevel genotype-phenotype map. *J R Soc Interface* 2020; 17, 20190843.

Zanette DH, Manrubia S., Fat tails and black swans: Exact results for multiplicative processes with resets. *Chaos* 2020, 30, 033104.

The main research topic of the group is the understanding, modelling and analysis of evolutionary mechanisms. For almost two decades, we have investigated the adaptive dynamics of viruses and RNA populations, collaborating closely with experimental groups and addressing broader problems such as the relationship between genotype and phenotype.

In our most recent research, we have been exploring the topological structure that genotype-to-phenotype maps endow in sequence spaces, and its effects in the dynamics of heterogeneous molecular populations. We have uncovered some universal features of sequence spaces topology which are independent of the definition of phenotype and, therefore, of generic consequences for evolution and adaptation. Our results have highlighted, among others, the extent of entropic effects in microscopic evolution, showing that abundant, sufficiently functional phenotypes, might be much more common in nature than highly adapted, but rare ones. A full understanding of microscopic evolution is important to update current evolutionary theories and to derive useful effective models. In this sense, we have questioned the role played by classical metaphors of evolution, such as smooth fitness landscapes, and suggested they must be substituted by network-based representations.

Our research has turned to epidemiology, based on our broad experience with viral evolution and modelling, as a consequence of the crisis caused by COVID-19. Currently, we are exploring the limits of model-based predictions in the face of empirical data, and the effects of the evolution of pathogens and their adaptation to different containment strategies.



1 *toyLIFE* is a multilevel genotype–phenotype map. (a) *toyLIFE* genotypes are binary strings with promoter and coding regions that, when expressed, yields a lattice folded toy protein. (b) Following *toyLIFE*'s interaction rules, we obtain gene regulatory networks (GRNs) in the form of a truth table. (c) Each GRN determines, under some propagation rules, a unique cellular automaton with cells in state empty (white), expressing protein A (orange), expressing protein B (blue) and expressing both proteins (grey). (d) These cellular automata give rise to spatio-temporal patterns of gene expression (Catalán et al., 2020).

2 Fit to data obtained in real time for the daily number of COVID-19 active cases in Spain (from March 1st to March 29th). Despite a reasonable agreement between model and empirical observations in the spreading phase of the pandemic, opposite predictions for the future number of active cases can be derived. The solid line represents the number of infected individuals using best-fit parameters. The vertical arrow denotes March 11th, the day when schools and universities closed. The shaded area represents the 95% predictive confident interval: Its increasing width implies that predictability decays exponentially fast.

GROUP LEADER**Juan Nogales Enrique****POSTDOCTORAL SCIENTISTS**Blas Blázquez Castiñeira
Erika González Álvarez
Igor Martínez Sánchez
Jesús Torres Bacete**PROJECT MANAGEMENT**

Clive Alexander Dove

PHD STUDENTSLucía Agudo Algibe
David San León Granados
María Tsampika Manoli**MASTER STUDENTS**Rodrigo Hedo Berrocal
Francisco Manuel Muñoz López
Iván Martín Martín
Ana del Ramo Galian**VISITING SCIENTISTS**Selma Bennacer
(U. Sorbona, France)
Darwin Carranza Saavedra
(U. Antioquía, Colombia)
América Jazmín Cota Álvarez
(U. Sinaloa, México)
Nahikari López López
(U. Navarra, Spain)
Jaroslaw Poptoński
(U. Wroclaw, Poland)
Rubén Sánchez Clemente
(U. Extremadura, Spain)
Sandra Sordon
(U. Wroclaw, Poland)**SELECTED PUBLICATIONS**García-Jiménez B, Torres-Bacete J, Nogales J. Metabolic modelling approaches for describing and engineering microbial communities. *Comput Struct Biotechnol J* 2020; 19: 226-246.Herencias C, Salgado-Briegas S, Prieto MA, Nogales J. Providing new insights on the biphasic lifestyle of the predatory bacterium *Bdellovibrio bacteriovorus* through genome-scale metabolic modelling. *Plos Comput Biol* 2020; 16(9), e1007646.Goris, T, Pérez-Valero A, Martínez I, Yi D, Fernández-Calleja A, et al. Repositioning microbial biotechnology against COVID-19: the case of microbial production of flavonoids. *Microb Biotechnol* 2020; 13 oct, Online ahead of print.Nogales J, Mueller J, Gudmundsson S, Canalejo FJ, Duque E, et al. High-quality genome-scale metabolic modeling of *Pseudomonas putida* highlights its broad metabolic capabilities. *Environ Microbiol* 2020; (1): 255-269.Lieven C, Beber ME, Olivier BG, Bergmann FT, Ataman M, et al. Memote: A community-driven effort towards a standardized genome-scale metabolic model test suite. *Nat Biotechnol.* 2020; 38 (3), 272-276.

Systems biotechnology

Our foundational aim is the system-level understanding of microbial metabolism as a framework for developing a broad range of novel and non-intuitive biotechnological processes. Taking advantage of metabolic modelling, systems and synthetic biology we are addressing, at different levels, the understanding and full taming of bacterial systems emergence.

Increasing the completeness and scope of metabolic reconstructions

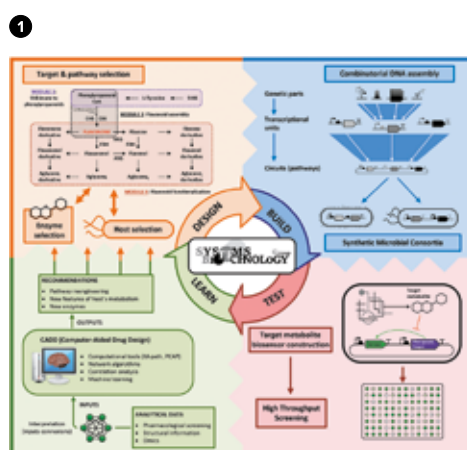
We are involved in the high-quality metabolic modeling of a large set of metabolically diverse bacteria including *P. putida*, *S. elongatus*, *A. platensis*, *Azoarcus CIB*, *S. granuli*, *P. pseudoalcaligenes* and *B. bacteriovorus*. This effort is enabling the system-level analysis of new metabolic processes while providing new computational test-beds for biotechnological applications. We are particularly interested in the inclusion of new metabolic modules such as the generation of reactive oxygen species and the inclusion of underground metabolisms. We are also developing software for the automatic reconstruction of microbial networks.

System-level analysis of Metabolic Robustness in bacteria

The robustness of a system is the property that allows it to maintain its functions despite perturbations. Through the metabolic modeling analysis of *P. putida*, we have identified metabolic cycles providing robustness. By using synthetic biology, ongoing efforts are focused on the rational engineering of such cycles under diverse biotechnological scenarios.

System-level analysis and designing of microbial communities

The division of labor allows an expanded complexity and functionality in bacteria. We are interested in: i) understanding how these expanded capabilities emerge



1 Detail of the Iterative design of DBTL (design–build–test–learn) cycle applied in the lab for addressing complex biotechnological endeavours. Design step includes the selection of target, the *in silico* design of production pathways, optimal pathway segregation, identification of enzymes and the selection of microbial hosts. Build stage is based on combinatorial DNA assembly methods to construct metabolic pathways that will be finally expressed in different components of a synthetic microbial consortium. Test stage includes production of the target compounds and the development high-throughput screening technologies. Learn stage processes the analytical data from the above steps and finds connections between genotype and phenotype and optimized metabolic fluxes to give recommendations to perform subsequent DBTL cycles.

within a community and ii) how we can engineer this community-level functionality towards biotechnological endeavors. To address these two fundamental questions, we have developed a computational platform called FLYCOP for modeling and engineering synthetic microbial consortia. We are applying this technology in the revalorisation of complex polymers such as lignin and plastic waste as well as in the cost effective production of plant-based secondary metabolites such as flavonoids.

GROUP LEADER**Florencio Pazos Cabaleiro****SENIOR SCIENTIST**

Mónica Chagoyen Quiles

PHD STUDENTS

Javier López-Ibáñez Infante

MASTER STUDENTS

Borja Pitarch

Sergio Revert

UNDERGRADUATE STUDENTS

Miguel Fernández

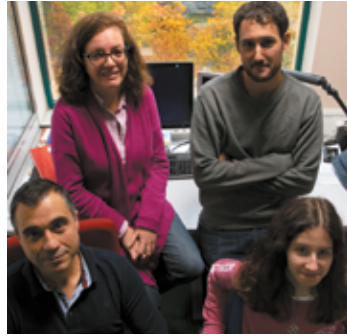
Ángel García de la Torre

Miguel Leal

Ana Mariya Anhel

Jorge Novoa

Gustavo Adolfo Sánchez



Computational systems biology

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 which complements the “in-vivo” and “in-vitro” methodologies of Biology. This mainly involves mining the massive amounts of information stored in biological databases. Within this general goal, we work on different research lines that can be framed in three major areas: prediction of protein functional sites, prediction of protein interaction partners, and functional study of biological networks (with an emphasis on networks related to human diseases). Besides our lines of scientific research, we also collaborate with experimental groups providing them with bioinformatics support for their specific needs, and participate in different teaching projects.

In the past two years we were actively working in deciphering the molecular basis of rare diseases combining data on genomic variations with biological networks. We also perform studies on the ecosystem of web servers supporting molecular biology research, with a focus on those dealing with CRISPR/Cas experiments. Additionally, we continued with our previous work on the prediction of protein binding sites and the prediction of the environmental fate of chemical compounds, finishing a couple of projects along these lines.

SELECTED PUBLICATIONS

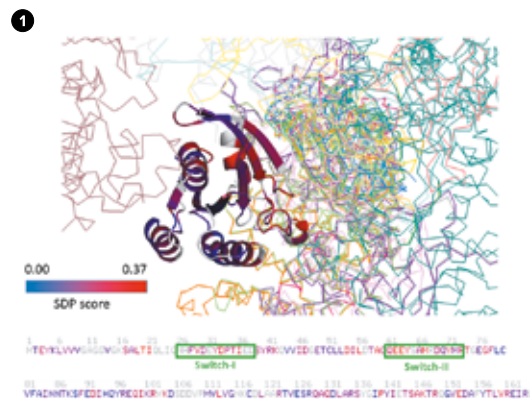
Pazos F, Chagoyen M. Characteristics and evolution of the ecosystem of software tools supporting research in molecular biology. *Brief Bioinform* 2019; 20(4):1329-1336.

Torres-Pérez R, García-Martín JA, Montoliu LI, Oliveros JC, Pazos F. WeReview: CRISPR Tools - live repository of computational tools for assisting CRISPR/Cas experiments. *Bioengineering*. 2019; 6:63.

López-Ibáñez J, Martín LT, Chagoyen M, Pazos F. Bacterial Feature Finder (BaFF) - a system for extracting features overrepresented in sets of prokaryotic organisms. *Bioinformatics*. 2019; 35 (18): 3482-3483.

Jabato FM, Seoane P, Perkins JR, Rojano E, García Moreno A, *et al.* Systematic identification of genetic systems associated with phenotypes in patients with rare genomic copy number variations. *Hum Genet* 2020; doi: 10.1007/s00439-020-02214-7.

Pitarch B, Ranea JAG, Pazos F. Protein residues determining interaction specificity in paralogous families. *Bioinformatics* 2020; btaa934.



1 Predictions of RasH regions involved in controlling effector interaction specificity (Pitarch *et al.*, 2020), mapped on the interaction structural information available for that protein. RasH is shown in ribbon representation, and its 26 crystallized interactors in thin backbone. The method's score for the RasH residues is shown in a color scale, with red representing the highest scores. The same color schema is used in the primary sequence of RasH (below) where the “switch-I” and “switch-II” regions are highlighted. Figure generated with PyMol (www.pymol.org).

2

Tool	Purpose	Score	Organisms	Citations
CRISPR-Cas9	Oligo Designer	x17	12	12
CRISPR-Cas1	Oligo Designer	x17	12	12
CRISPR-Cas2	Oligo Designer	x17	12	12

2 Screenshots of WeReview web interface (Torres-Pérez *et al.*, 2019), a live repository of computational tools for assisting CRISPR/Cas experiments. The main table, together with some of the dialogs for introducing filters in the search for tools are shown.

GROUP LEADER**Juan F Poyatos****PhD STUDENTS**Alvar J. Alonso
Pablo Yubero**UNDERGRADUATE STUDENTS**

Diego Jiménez

SELECTED PUBLICATIONS

Sánchez-Gorostiaga A, Bajić D, Osborne ML, Poyatos JF, Sanchez A. High-order Interactions distort the functional landscape of microbial consortia. *PLoS Biol* 2019; 17 (12): e3000550.

Chagoyen M, Poyatos JF. Complex genetic and epigenetic regulation deviates gene expression from a unifying global transcriptional program. *PLoS Comput Biol* 2019; 15 (9): e1007353.

Kovács K, Farkas Z, Bajić D, Kalapis D, Daraba A, *et al.* Suboptimal global transcriptional response increases the harmful effects of loss-of-function mutations. *Mol Biol Evol* 2020; msaa280.

Poyatos JF. Genetic buffering and potentiation in metabolism. *PLoS Comput Biol* 2020; 16(9): e1008185.

Yubero P, Poyatos JF. The impact of global transcriptional regulation on bacterial gene order. *iScience*. 2020; 23(4):101029.

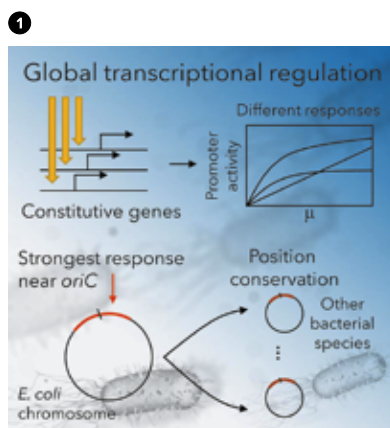


Logic of genomic systems

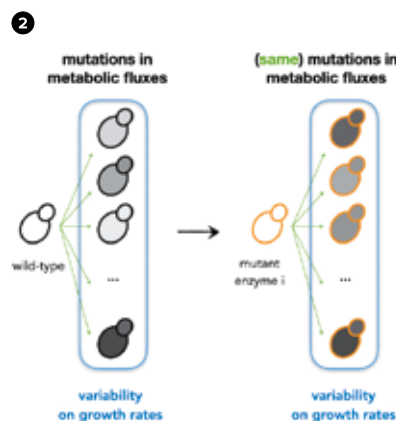
Research at the Logic of Genomic Systems Laboratory searches for design principles in biological systems. During the last years, we examined the global transcriptional program controlling genome-wide gene expression, the link between composition and function in microbial consortia, and the factors that determine the impact of mutations in cellular fitness.

The limited availability of the components that influence gene expression, such as the presence of free RNA polymerases, cofactors, ribosomes, etc., and their differential use at the genomic scale determines *the global transcription control program*. We examined this program by 1) developing a methodology to experimentally characterise on a large scale the response to this program in bacteria, which showed that this response contributes to the bacterial genomic organisation, 2) examining how this program integrates with specific –genetic and epigenetic– regulatory strategies in eukaryotes.

To understand how the interaction between members of a microbial community determines its function, we assembled an artificial consortium of soil bacteria in which function represents starch degradation. Combining theory with experiments, we quantified how the contribution of interactions of different order and type shape the action of the community. Functional robustness to pairwise and higher-order interactions critically affects our ability to predict and engineer function.



1 Bacterial gene expression depends on the allocation of limited transcriptional resources provided a particular growth rate and growth condition. Early studies in a few genes suggested this global regulation to generate a unifying hyperbolic expression pattern. We showed that promoters whose transcriptional response is more dependent on growth rate are preferentially located closer to the origin of replication in the chromosome in *E. coli*, and that the relative location of these genes in other species correlates significantly with their respective growth dynamics, directly related to their habitat.



2 Using a genome-scale metabolic model, we scored the variability of the growth rates in a group of different lines (arrows) in which mutations in the metabolic fluxes are accumulated. We computed this variability in the presence (wild-type) and the absence (mutant metabolism) of a particular enzyme *i*. Here the difference between growth rates and metabolic backgrounds is represented by the colors of the fill and the border of the yeast cartoons, respectively. We then quantify if these very same lines manifest a different variability depending on the absence of the single enzyme. This led us to identify a set of genes acting as buffers and potentiators whose influence depends on the particular working conditions of the metabolism (i.e., type of available nutrients), and the sources of variability considered.

Understanding the impact of mutations is the focus of Genetics, but many questions remain uncertain. We lately examined two: 1) to what extent part of the fitness cost of a mutation comes from improper rewiring of the transcriptome, and 2) how much any enzyme can act as a modifier of the impact of mutations on other enzymes. In 1), we demonstrated that part of the deleterious effects of mutations is indeed caused by such abnormal rewiring. In 2), we showed that any enzyme can buffer or potentiate the impact of mutations, an effect that has implications in particular cancer therapies.

GROUP LEADERS

Javier Tamames de la Huerta
Carlos Pedrós-Alió

POSTDOCTORAL SCIENTISTS

Beatriz Cámara Gallego
 Fernando Puente Sánchez

PhD STUDENTS

Marta Cobo Simón
 Natalia García García
 Diego Jiménez Lalana

MASTER STUDENT

Bárbara Gómez-Tavira
 Álvaro Redondo del Río

VISITING SCIENTISTS

Cinthya Tebes Cayo
(U. Pontificia Católica del Norte, Chile)
 María Fernanda Campos Filgueira
(U. Pontificia Católica del Norte, Chile)

SELECTED PUBLICATIONS

García-García N, Tamames J, Linz AM, Pedrós-Alió C, Puente-Sánchez F. Microdiversity ensures the maintenance of functional microbial communities under changing environmental conditions. *ISME J* 2019; 13: 2969-2983.

González, JM, Hernández L, Manzano I, Pedrós-Alió C. Functional annotation of orthologs in metagenomes: a case study of genes for the transformation of oceanic dimethylsulfoniopropionate. *ISME J* 2019; 13: 1183-1197.

Tamames J, Cobo-Simón M, Puente-Sánchez F. Assessing the performance of different approaches for functional and taxonomic annotation of metagenomes. *BMC Genomics*. 2019; 20 (1): 960.

Díez-Vives, C, [...] Pedrós-Alió C, Gasol JM, Acinas SG. Delineation of ecologically distinct units of marine Bacteroidetes in the Northwestern Mediterranean Sea. *Mol Ecol* 2019; 28: 2846-2859.

Tamames J, Puente-Sánchez F. SqueezeMeta, A highly portable, fully automatic metagenomic analysis pipeline. *Front Microbiol* 2019; Jan 24; 9: 3349.



Microbiome analysis

Microbial communities (microbiomes) are key players in many scenarios, from how the biosphere works to industrial and biotechnological processes, as well as human health and wellness. We study microbiomes of diverse environments trying to learn the rules that govern the assemblage of these microbial communities. This knowledge will help to understand how they function, and to predict the effects of disturbances. Eventually, this will lead to rational design and manipulation of microbiomes.

We focus mostly on marine microbial communities, but we are actively working in many other microbiomes from different environments. We study extreme environments because their microbiotas show fascinating adaptations to the harsh conditions. We work with human-associated microbiomes, such as the gut and the vagina, because of their potential to improve our health. We are also interested in other habitats, such as wastewaters and soils.

We use mostly bioinformatics tools to study the composition and functionality of microbiomes. Metagenomics is the basis of our work, since it provides the basic material: DNA sequences from environmental samples. The analysis of these sequences informs about the presence of diverse organisms and the content of their genomes, and the latter can be linked to functionality. We also carry out experimental work addressing interactions between members of microbiomes.

① *Solar salterns in the Salar de Atacama (Chile). Lithium carbonate precipitation ponds where we are looking for the limits of life at low water activity.*



② *Time series from winter equinox to summer solstice in Cambridge Bay (Canadian Arctic). Copy number of siderophore synthesis genes (left panel) and siderophore transporter genes (right panel) in metagenomes. Pie charts show taxonomic assignments. Gamma proteobacteria (purple) are the main producers of siderophores, while Bacteroidetes (orange) act as cheaters by increasing the number of outer membrane transporters. Webpage: <http://microbiomecnb.com/>*

②

