

6 / Systems Biology

Although molecular biology is considered to have been founded by physicists, this circumstance did not result in a quantitative culture and an accurate, standardised descriptive language characteristic of the hard sciences. On the contrary, with very few exceptions, the biosciences that developed since that time seldom took the opportunity to formalise the mechanisms and functions of living systems with accurate languages and codes, with a nearly complete disregard for metrology (the theory and practice of quantitative measurement). 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, synthetic biology is not just a contemporary update of the recombinant DNA technologies of the past 30 years, along with a 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 on 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 Program figures in the contemporary landscape by developing research lines in environmental genomics, network biology, systemic computation and metabolic engineering.



Molecular environmental microbiology laboratory

88 / SYSTEMS BIOLOGY

PRINCIPAL INVESTIGATOR:

Víctor de Lorenzo

POSTDOCTORAL SCIENTISTS:

Ozlem Akkaya
Yamal Al-Ramahi
Ilaria Benedetti
Belén Calles
David Fernández
Angel Goñi
Juhyun Kim
Esteban Martínez
Pablo I. Nikel

PREDOCTORAL SCIENTISTS:

M^a de los Angeles Hueso
Alberto Sánchez- Pascuala

TECHNICIANS:

Sofía Fraile
Tomás Aparicio

ADMINISTRATIVE ASSISTANT:

Inés Merino

VISITING SCIENTIST:

Ana Zuñiga
(Univ. Adolfo Ibáñez, Chile)

SELECTED
PUBLICATIONS

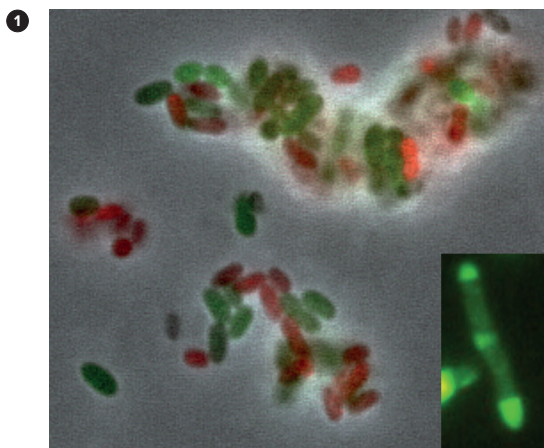
Martínez-García E, Aparicio T, Goñi-Moreno A, Fraile S, de Lorenzo V. SEVA 2.0: an update of the Standard European Vector Architecture for de-/re-construction of bacterial functionalities. *Nucleic Acids Res* 2014; 43: D1183-D1189

Nikel PI, Martínez-García E, de Lorenzo V. Biotechnological domestication of pseudomonads using synthetic biology. *Nat Rev Microbiol* 2014; 12: 368-379

Calles B, de Lorenzo V. Expanding the Boolean logic of the prokaryotic transcription factor XylR by functionalization of permissive sites with a protease-target sequence. *ACS Synth Biol* 2013; 2: 594-603

Pérez-Pantoja D, Nikel PI, Chavarría M, de Lorenzo, V. Endogenous stress caused by faulty oxidation reactions fosters evolution of 2,4-dinitrotoluene-degrading bacteria. *PLoS Genet* 2013; 9: e1003764

The longstanding aim of our team is to produce biological agents for biosensing, remediation and (where possible) transformation of urban and industrial chemical waste that is otherwise dumped into the environment. To this end, we explore and capitalise 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 abilities required of 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 editing the extant genome of this microorganism to enhance desirable properties and reduce drawbacks. We also use surface-display systems to design complex catalytic properties completely separate from the cell metabolism, as well as artificial communities by expression of ectopic adhesins. [ii] Use of camel antibodies as tools for metabolic and regulatory engineering of *P. putida*. Various expression systems allow 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 cells. [iii] Genetic tools for deep refactoring of *P. putida* metabolic properties. The new assets we are developing include a large collection of standardised plasmid and transposon vectors, as well as dedicated reporter systems for parameterisation of the gene expression flow and for switching entire metabolic regimes. [iv] The TOL system borne by plasmid pWWO as a reference for metabolic circuit implantation. The two operons for toluene and *m*-xylene biodegradation encoded in pWWO 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*. Current efforts attempt to develop strains programmed to adopt an entirely anoxic metabolism and/or a genuine glycolytic mode of glucose use. This involves the re-engineering of dozens of genes and metabolic modules.



1 Environmental bacteria offer an outstanding opportunity in synthetic biology for engineering new-to-nature properties in specific members of the community; this includes the possibility of designing consortia with predetermined catalytic activities (which we term “catalytic origami”). The genome of selected strains of *P. putida* must be edited to generate a smooth cell surface, on top of which artificial adhesins can be programmed genetically. Communication between cells of such consortia can also be engineered by repurposing pathway components for the degradation of recalcitrant compounds.



Evolutionary systems

PRINCIPAL INVESTIGATOR:
Susanna Manrubia

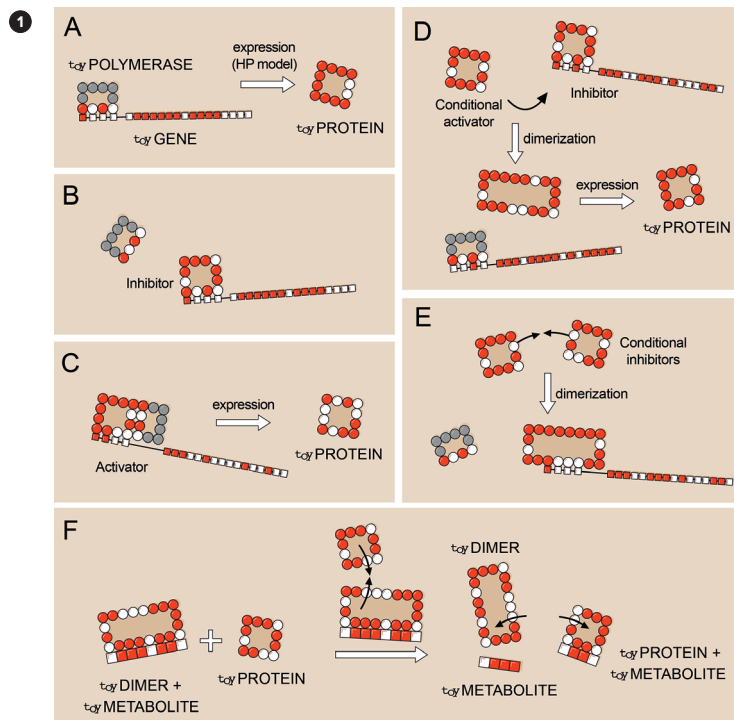
POSTDOCTORAL SCIENTISTS:
Jacobo Aguirre
José A. Capitán



Capitán JA, Aguirre J, Manrubia S. Dynamical community structure of populations evolving on genotype networks. *Chaos Solitons Fractals* 2014; doi:10.1016/j.chaos.2014.11.019

Arias CF, Catalán P, Manrubia S, Cuesta JA. toyLIFE: a computational framework to study the multi-level organisation of the genotype-phenotype map. *Sci Rep* 2014; 4:7549

Our main interest is theoretical investigation of evolutionary systems of different kinds. We develop models inspired by the phenomena observed in natural systems, chiefly populations of RNA sequences and RNA viruses. Our approach addresses the study of general properties of evolving and adapting systems or, at a more specific level, tries to reproduce and predict the response of such populations to environmental changes. We study 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. To understand the limitations of simple genotype-phenotype maps, we recently introduced toyLIFE, a multilevel computational model that proceeds simply but realistically from gene sequence to emerging metabolic network. In a broader scenario, we are also interested in modelling the interaction between agents organised in networks that vie for resources such as food or mates, as competitive interactions that 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 development of analysis techniques with environmental and health purposes. We apply graph theory to antibody microarrays to improve the characterisation of experimental samples, with direct application to allergy control, toxin detection in fresh water ecosystems and planetary sciences. Our studies of viral response to antiviral treatments have determined optimal modes of drug administration to minimise viral load and mutant escape.



1 *ToyLIFE is a simple molecular universe in which genes formed by chains of toyNucleotides are translated into toyProteins. The latter are folded in two dimensions following simple rules of interaction between hydrophobic and hydrophilic toyAminoacids. The same interaction rules are applied between any pair of elements in toyLIFE leading to the emergence of new molecular properties, Boolean networks of interactions between genes, or toyMetabolite catabolism. More information in Arias et al. (Sci Rep 2014; 4:7549).*



Computational systems biology

PRINCIPAL INVESTIGATOR:
Florencio Pazos Cabaleiro

POSTDOCTORAL SCIENTIST:
Mónica Chagoyen Quiles

PREDOCTORAL SCIENTISTS:
Daniel López López
David Ochoa García
Javier López-Ibañez Infante
Guillermo García Martín
Sara Fernández Novo

VISITING SCIENTIST:
David San León Granado
(CSIC, Spain)

In the last two years, we continued work on various aspects of our three main research lines. For prediction of protein-protein interactions, we improved our MIRRORTREE system for co-evolution-based prediction of protein interaction partners in a number of ways. These include the introduction of an estimator of the statistical significance of an observed co-evolutionary score and incorporation of predicted solvent accessibility. For the functional study of biological networks, we continued to improve our pioneering MBROLE system for the functional analysis of metabolomic data, with an exhaustive study of the role of intrinsic protein disorder in shaping molecular networks of *Arabidopsis thaliana*. For the prediction of protein functional sites, we developed a new system for the concomitant prediction of fold, molecular function, and functional sites for structural domains, which is publicly available through the COPRED web server (<http://csbg.cnb.csic.es/copred/>). In addition, we continued to collaborate with experimental groups, helping with their bioinformatics needs, applying the above methodologies and tools.

Detailed information can be found at: <http://csbg.cnb.csic.es/>

SELECTED PUBLICATIONS

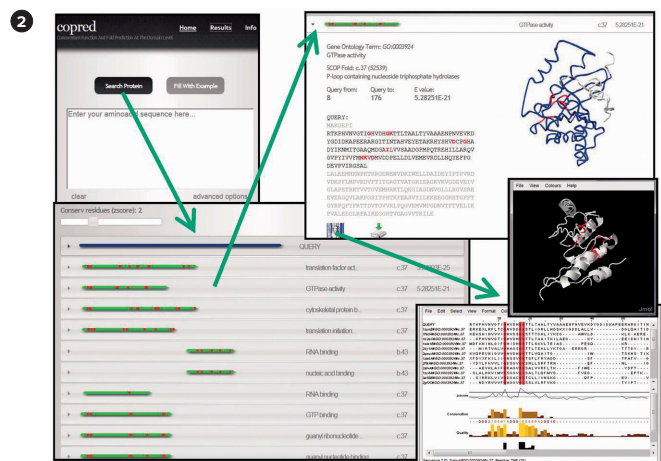
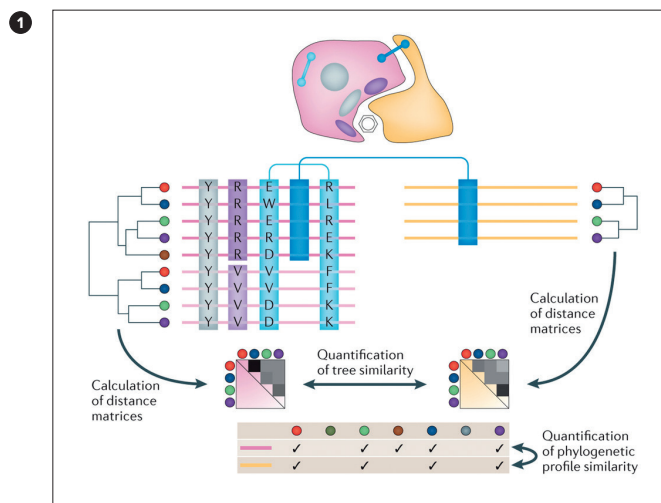
Espaillet A, Carrasco-Lopez C, Bernardo-Garcia N, Pietrosemoli N, Otero LH, Alvarez L, de Pedro MA, Pazos F, Davis BM, Waldor MKI, Hermoso JA, Cava F. Structural basis for the broad specificity of a new family of amino-acid racemases. *Acta Crystallogr D* 2014; D70: 79-90

Lopez D, Pazos F. COPRED: prediction of fold, GO molecular function and functional residues at the domain level. *Bioinformatics* 2013; 29: 1811-1812

Juan D, Pazos F, Valencia A. Emerging methods in protein co-evolution. *Nat Rev Genet* 2013; 14: 249-261

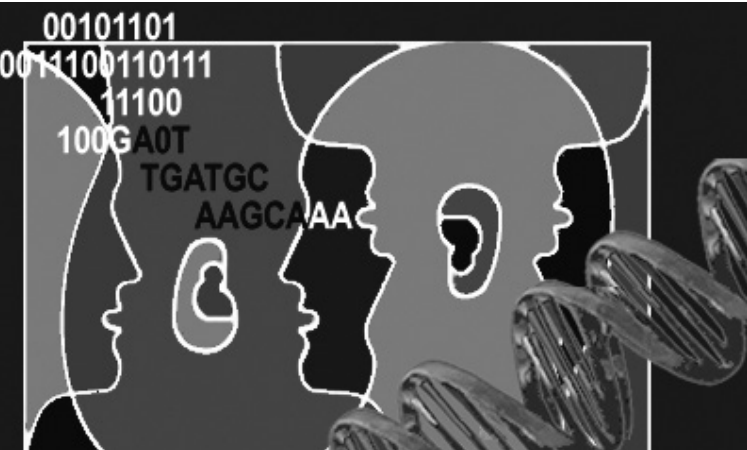
Pietrosemoli N, García-Martín JA, Solano R, Pazos F. Genome-Wide Analysis of Protein Disorder in *Arabidopsis thaliana*: Implications for Plant Environmental Adaptation. *PLoS One* 2013; 8: e55524.

Chagoyen M, Pazos F. Tools for the functional interpretation of metabolomic experiments. *Brief Bioinform* 2013; 14: 737-744



1 Co-evolutionary information extracted from protein multiple sequence alignments indicates protein features related to structure, function and interactions.

2 User interface of the COPRED system for the concomitant prediction of protein function and functional sites.



Logic of genomic systems

PRINCIPAL INVESTIGATOR:
Juan F. Poyatos

PREDOCTORAL SCIENTISTS:
Clara Moreno
Djordje Bajic

VISITING SCIENTIST:
Guillermo Rodrigo
(Universitat Politècnica de València, Spain)



Rodrigo G, Bajic D, Elola I, Poyatos JF. Antagonistic autogenous control of gene expression enhances the dynamic response of a resistance phenotype. *bioRxiv* 2014; doi: 10.1101/008169

Bajic D, Moreno C, Poyatos JF. Rewiring of genetic networks in response to modification of genetic background. *Genome Biol Evol* 2014; 6:3267

Cavaliere M, Poyatos JF. Plasticity facilitates sustainable growth in the commons. *J Royal Soc Interface* 2013; 10:20121006

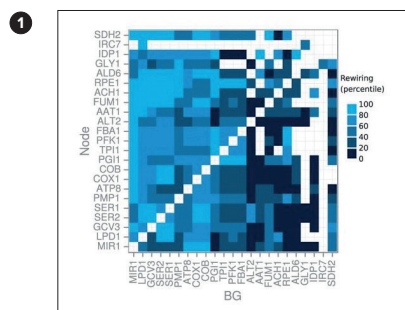
Two main problems were the centre of our research over the last two years. The first refers broadly to the limits of our present representation of biological systems as circuits and networks. The second problem deals with how the behaviour of individual cells determines the collective dynamics of cell populations.

The representation of biological complexity

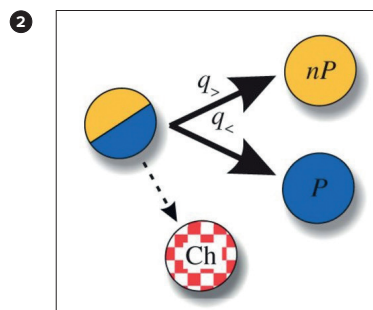
We first analysed the stability of genetic networks. These networks help to map the molecular circuitry within cells by establishing genetic interactions between proteins (two proteins are “genetically” connected when the phenotype of the double mutant differs from that predicted for the combination of single mutations alone). The architecture of genetic networks can be modified with variation of genetic context (other mutations), which we studied by *in silico* modelling and by analysis of high-throughput data. Stability appeared to be closely connected to the mechanistic causes of robustness in the system under study. The second question refers to the assembly of transcriptional networks from its basic building blocks, or network motifs. We examined how the combination of functionally conflicting motifs could lead to trade-offs, and how such trade-offs can be resolved. We combined quantitative experiments and mathematical modelling in a bacterial model system with antagonistic autogenous control, the multiple antibiotic resistance regulatory module in *Escherichia coli*.

Individual behaviour and collective dynamics of cells

Cells commonly cooperate among themselves by the secretion of molecules used by other cells (“public goods”). These interactions help establish cellular communities that are naturally unstable due to the rise of mutant individuals. These individuals avoid the costs associated with production of public goods, which can collapse the community. We analysed the persistence of these systems by examining a situation in which production of public goods by individual cells followed different strategies –constitutive, stochastic or plastic. We developed a mathematical model to quantify how each strategy would function in the presence of the non-contributing mutants, in a model that also controlled for a number of structural parameters. This analysis currently serves for the design and construction of robust bacterial communities in the lab.



1 Rewiring of hubs in response to hub deletions. The only hub with a purely biosynthetic function (IRC7) deviates most from this general pattern.



2 Individual cells as plastic producers. A plastic producer can contribute to the public good (producer state, P) or not (non-producer state, nP). Cheaters (Ch), that is, agents that are permanently in the nP state, arise from plastic producers by mutation.



Microbial community modelling

PRINCIPAL INVESTIGATOR:

Javier Tamames de la Huerta

SENIOR SCIENTIST:

Carlos Pedrós-Alió

TECHNICIAN:

Pablo David Sánchez

VISITING SCIENTISTS:

Ana Suárez-Suárez
(Newcastle University, UK)

Cristina García
(Universidad de Barcelona, Spain)

Elena Toril

(Centro de Astrobiología, INTA-CSIC, Spain)

Laura Sala

(Universidad de Barcelona, Spain)

Tomeu Viver

(IMEDEA-CSIC, Spain)

SELECTED PUBLICATIONS

Guazzaroni ME, Herbst FA, Lores I, Tamames J, Peláez AI, López-Cortés N, Alcaide M, Del Pozo MV, Vieites JM, von Bergen M, Gallego JL, Bargiela R, López-López A, Pieper DH, Rosselló-Móra R, Sánchez J, Seifert J, Ferrer M. Metaproteogenomic insights beyond bacterial response to naphthalene exposure and bio-stimulation. *ISME J* 2013; 7:122-136

López-López A, Richter M, Peña A, Tamames J, Rosselló-Móra R. New insights into the archaeal diversity of a hypersaline microbial mat obtained by a metagenomic approach. *Syst Appl Microbiol* 2013; 36:205-214

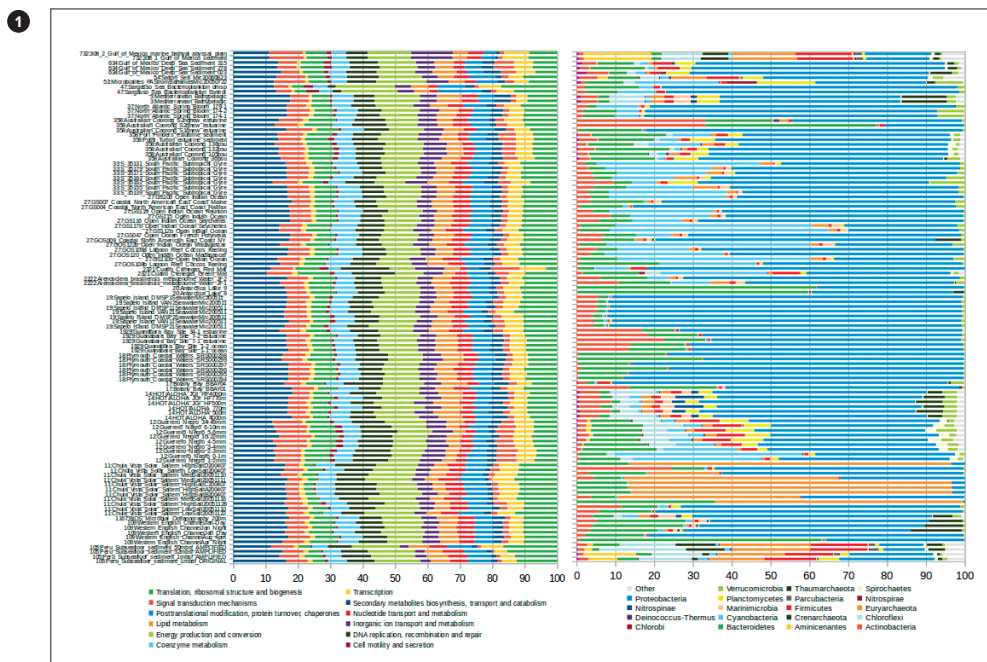
Tamames J, Durante-Rodríguez G. Taxonomy becoming a driving force in genome sequencing projects. *Syst Appl Microbiol* 2013; 36:215-217

Jiménez G, Blanch AR, Tamames J, Rosselló-Mora R. Complete Genome Sequence of *Bacillus toyonensis* BCT-7112T, the Active Ingredient of the Feed Additive Preparation Toyocerin. *Genome Announc* 2013; 1:1

Pascual-García A, Tamames J, Bastolla U. Bacteria dialog with Santa Rosalia: Are aggregations of cosmopolitan bacteria mainly explained by habitat filtering or by ecological interactions? *BMC Microbiol* 2014; 14:284

In these two years, our group has established a solid methodology for the analysis of genomes, metagenomes and metatranscriptomes that has been applied successfully in several instances. We are particularly interested in ecological scenarios such as the detoxification of polluted environments (Guazzaroni *et al.*, 2013), function of hypersaline environments (López-López *et al.*, 2013), and responses to temperature gradients in thermophilic microbial mats (in collaboration with the Pontificia Universidad Católica de Chile, 2014). We also collaborate with several laboratories on the clinical importance of genomes and microbiomes, such as the group of Dr F. Baquero (Hospital Ramón y Cajal, Madrid) in cystic fibrosis, and that of Dr I. Tomás (Universidad de Santiago de Compostela, Spain) in the role of oral microbiomes in periodontal health.

We also study the rules that govern assembly of natural bacterial communities, by defining the relationships between environmental factors and species presence or co-occurrence. This led us to define patterns of functional redundancy in metagenomes, that is, in which cases particular species can be replaced by counterparts with no influence on the overall performance of the microbiome. We also developed predictive systems to determine environmental preferences of individual species based on their genomic compositions. We are now working to extend these tools to predict possible compatibilities and co-occurrences between species in given settings. In the future, we expect to be able to accurately predict the effect of specific perturbations in the performance of the microbiome, and therefore to propose optimal combinations of species for living in defined conditions or fulfilling given roles.



1 Comparison of 100 marine metagenomes showing how functional uniformity (left) corresponds to rather different taxonomic compositions (right). This indicates a high degree of functional redundancy, in which community members can be replaced with no perturbation of the functional profile of the microbiome.