Principles of Systems Biology—No. 5

If systems biology is about understanding how links between components yield emergent phenomena, this month's Cell Systems Call (Cell Systems *1*, 307) contains a veritable bounty of examples, showcasing the breadth of the field from systems oceanography to molecular evolution to the influence of cellular niche microenvironments on stem cell development.

From Molecular Oceanography to Ocean Systems Modeling

Samuel Chaffron, VIB, KU Leven and University of Brussels; Lionel Guidi, Sorbonne Universités, UPMC Université Paris 06, CNRS, Laboratoire d'océanographie de Villefranche (LOV), Observatoire Océanologique; Lucie Bittner, Sorbonne Universités, Université Pierre et Marie Curie (UPMC), CNRS, Institut de Biologie Paris-Seine (IBPS), Evolution Paris Seine; and Damien Eveillard, Université de Nantes

Principles

The *Tara* Oceans expedition collected the most comprehensive molecular and oceanographic datasets covering the entire ocean system, opening the exciting possibility of characterizing a key oceanographic process, the biological carbon pump, at the molecular level. In this process, phytoplankton organisms transform CO_2 to organic carbon through photosynthesis, which is then exported through sinking particles, and finally sequestered in the deep ocean. Using a systems biology approach, we constructed ecological networks and identified specific plankton communities and functions strongly associated with carbon export in subtropical oligotrophic oceans (Guidi, et al., Nature *532*, 465–470). We highlighted *Synechococcus* and their phages associated to carbon export, as well as unexpected taxa such as Radiolaria and alveolate parasites. We also identified putative molecular markers to predict carbon export in oligotrophic oceans.

> *"This work integrates genomic and oceanographic knowledge at the global ecosystem level…."*

What's Next?

This work integrates genomic and oceanographic knowledge at the global ecosystem level by combining oceanography, biology, and computer sciences. The next step of this "systems oceanography" approach will be to extend our analyses to two other biological carbon pump components (primary production and remineralization), notably by integrating gene transcription information, to build predictive models of the biological pump, a key component of the carbon cycle in our global ocean systems.

Toward Predicting Evolution

Jianzhi Zhang, Department of Ecology and Evolutionary Biology, University of Michigan

Principles

Evolutionary biology is generally considered a retrospective science, as the relevant data usually result from past events. The consequence is that our understanding of evolutionary principles may be biased, because we have observed only what happened, rather than what could have happened. If all potential evolutionary paths of a species could be compared, we would be able to explain why a particular route was taken and predict where evolution is heading. To achieve this goal, one would need to know the Darwinian fitness of all possible genotypes of a species. While this remains a fantasy, we recently succeeded in quantifying the fitness of >65,000 yeast genotypes that vary only in the sequence of one gene (Li et al., Science *352*, 837–840). The acquired data reveal the fitness effects of all possible point mutations in the gene as well as tens of thousands of combinations of point mutations, allowing prediction of the next few evolutionary steps of this gene.

> *"The acquired data reveal the fitness effects of all possible point mutations in the gene as well as tens of thousands of combinations of point mutations…."*

What's Next?

Fitness also depends on the environment. Measuring fitness in multiple conditions and studying how environmental changes shape evolutionary trajectories will be important. Using the large but limited data to find rules governing the mapping of genotypes to fitness will be extremely valuable.

Calculating Maximal Catalytic Rates of Enzymes inside Cells

Dan Davidi and Ron Milo, Weizmann Institute of Science

Principles

The catalytic rates of enzymes are of central interest in biology and biochemistry, as they determine the rate of most biological processes. Since enzymatic rates can rarely be measured directly in vivo, in vitro k_{cat} values are often used to describe the rate limits of catalysis. However, k_{cat} data are extremely scarce, and most experimentally available k_{cat} values were measured in non-physiological conditions. In our recent study, we answer two fundamental questions (Davidi et al., PNAS *113*, 3401–3406). First, how representative are in vitro k_{cat} values for the maximal catalytic rates of enzymes in vivo? Second, in the era of omics data, can we accelerate the collection of enzyme kinetic constants? We estimate the maximal catalytic rate of enzymes in vivo by combining proteomic, fluxomic, and metabolomic data and find remarkable similarity to in vitro k_{cat} measurements.

> *"We estimate the maximal catalytic rate of enzymes in vivo…and find remarkable similarity to in vitro k_{cat} measurements."*

What's Next?

Our work provides a high-throughput tool to predict the maximal capacity of enzymes inside cells, and therefore it expands the capacity of metabolic modeling techniques. Genome-scale calculation of enzyme kinetic parameters is of interest for metabolic engineering applications, synthetic pathway design, and biophysical analysis of biological systems.

Cell Systems Call

Multi-state Gene Transcription Dynamics

Karen Featherstone, Centre for Endocrinology and Diabetes, University of Manchester; Adam M. Corrigan and Jonathan R. Chubb, MRC Laboratory for Molecular Cell Biology, University College London; Michael R.H. White, Systems Biology Centre, University of Manchester; Julian R.E. Davis, Centre for Endocrinology and Diabetes, University of Manchester

Principles

Gene expression in individual cells is dynamic and apparently stochastic, with dramatic variations over wide-ranging timescales, from seconds to hours. Transcription dynamics have been modeled as either a constant rate process, or a binary (two-state) process, with a single "on" rate, a process referred to as transcriptional "bursting". Recent studies (Featherstone et al., eLife *5*, e08494, and Corrigan et al., eLife *5*, e13051) challenge the perception of simple binary dynamics by identifying continuous distributions of transcription rates in active "on" states. The additional complexity of regulating RNA polII promoter engagement in these new models likely depends on multiple inputs, from fluctuations in transcription factor activity to regulation of chromatin and gene locus architectures.

> *"[Two] recent studies…challenge the perception of simple binary [transcription] dynamics by identifying continuous distributions of transcription rates in active 'on' states."*

What's Next?

Transcriptional and chromatin processes underlying state transitions must be identified. Featherstone et al. characterized transcription dynamics in intact tissue, whose architecture modifies intercellular coordination of timing. Their studies of reporter gene expression were indirect, whereas Corrigan et al. directly measured RNA transcript synthesis, at much shorter timescales. Future questions include (1) what is the dynamic relationship between chromatin remodeling and transcription over broad timescales, (2) are multistate dynamics disturbed in abnormal tissue such as tumors, and (3) how is the spectrum of transcriptional states relevant to different physiological conditions?

RNA Production Delays Shape Transcriptional Response Dynamics

Antti Honkela, University of Helsinki; and Magnus Rattray, University of Manchester

Principles

Understanding the precise timing of gene transcription is essential for modelling gene regulation and expression. We recently reported results of detailed genome-wide dynamical modelling of transcription and expression, highlighting widespread significant delays in RNA production between the completion of Pol-II elongation and accumulation of mature mRNA (Honkela et al., PNAS *112*, 13115–13120). To uncover these delays, we fitted a simple model of transcription to RNA-seq and Pol-II ChIP-seq time course data using Bayesian techniques. Production delays of more than 20 min were observed for 11% of genes. Long delays were more common for genes with short pre-mRNAs. They were also associated with late intron retention in pre-mRNA data, suggesting a link to splicing.

> *"Dynamic models of gene regulation should incorporate signal propagation delays, such as RNA production delays…."*

What's Next?

Dynamic models of gene regulation should incorporate signal propagation delays, such as RNA production delays in our model. Further work is needed to understand the mechanism and regulation behind the observed delays and how they are modulated by the cellular environment. This could be achieved by studying data from different systems or under different stimulations.

Neurons Uses Physics for Receptor Selection

Jianhua Xing, Department of Computational and Systems Biology, School of Medicine, University of Pittsburgh

Principles

With rare exception, each mammalian olfactory sensory neuron stochastically chooses to express one and only one type (actually one allele) of transmembrane olfactory receptor. This is one out of possibly a thousand or more types of receptors, each with about equal probability of expression. Our recent model studies emphasize the utility of treating the system as a dual-objective optimization problem (Tian et al., PNAS, published on line May 9, 2016. http://dx.doi.org/10.1073/pnas.1601722113). Through a histone epigenetic mark switching process, followed by competition between promoters for a limited number of enhancer elements, cells achieve monoallelic expression as the top priority and maximize expression diversity, given that the first requirement has been satisfied. In these two processes, cells use key physics of cooperativity, which we term "richer gets richer" and "intermediate states of a cooperative process tend to be depleted." The model predicts many existing experimental observations, such as counter-intuitive reduction of expression diversity in G9a/GLP knockout mice.

> *"Our recent model studies emphasize the utility of treating the system as a dual-objective optimization problem."*

What's Next?

The olfactory receptor selection problem is attractive for quantitative systems biology studies, especially given that measurements of single-cell temporal dynamics are currently lacking in the field. Integrated efforts between experiments and modeling will reveal fine details of the regulatory mechanism and general implications in gene regulation.

Integrating Entire Metabolic Pathways in *E. coli*

Marcelo C. Bassalo, Department of Molecular, Cellular and Developmental Biology, University of Colorado Boulder; Ryan T. Gill, Department of Chemical and Biological Engineering, University of Colorado Boulder

Principles

The genomic integration of multi-gene constructs is a several-day, low-efficiency process in the vast majority of organisms. By combining CRISPR-Cas9 with lambda-red assisted recombination, we demonstrated a strategy that allows integration of heterologous constructs ranging in size from single genes to entire metabolic pathways at efficiencies high enough (>50%) to remove any requirement for further selection. This strategy enables insertion at any site in the target genome and requires only a single step that can be performed in 1 day (Bassalo et al. ACS Synth. Biol.,published online April 12, 2016. http://dx.doi.org/10.1021/acssynbio.5b00187). Using this strategy, we show integration of a 10 kb isobutanol pathway assembled directly from synthetic oligos at 50% efficiency, thereby providing for construction of a platform production strain directly from synthetic DNA in a few days.

> *"By combining CRISPR-Cas9 with lambda-red assisted recombination, we demonstrated a strategy that allows integration of…entire metabolic pathways at efficiencies high enough (>50%) to remove any requirement for further selection."*

What's Next?

Efficient methods for inserting complex metabolic pathways, regulatory circuits, and multiprotein assemblies will allow for rapid construction of libraries of organisms modified for a broad range of applications across biology and biotechnology. Moreover, increasingly sophisticated multiplex editing technologies can be applied on top of such libraries to further expand the size and scope of designs that can be constructed and tested. Doing so, however, will require further increases in efficiencies in all of the supporting technologies including DNA (oligo) synthesis, assembly, insertion, and editing.

Distinct Vascular Regions Differentially Control Stem Cell Fate Decisions

Tomer Itkin and Tsvee Lapidot, Department of Immunology, Weizmann Institute of Science

Principles

In the bone marrow, peri-vascular niches serve as sites for bone and blood forming stem cell maintenance. However, blood vessels also promote stem cell development and mature blood cell production and form the gateway for leukocyte trafficking. We hypothesized that both processes cannot be performed by the same vascular sites and characterized the architecture and properties of distinct types of less-permeable arterial and leaky sinusoidal blood vessels. We revealed that metabolic activation and cellular trafficking occur exclusively at sinusoidal sites, whereas peri-arterial sites maintain metabolically dormant hematopoietic stem cells. Sinusoidal vascular leakiness allows peripheral blood plasma penetration, activating stem cell metabolism and inducing their migration and development. Our study defines bone marrow stem cell maintenance versus trafficking and differentiation regulation by distinct blood vessel subtypes (Itkin et al., Nature *532*, 323–328).

> *"Our study defines bone marrow stem cell maintenance versus trafficking and differentiation regulation by distinct blood vessel subtypes."*

What's Next?

These findings may improve clinical stem cell transplantation protocols, increasing the yield of mobilized stem cells, enhance transplanted stem cell homing, and expand lodged stem cells to achieve successful and fast engraftment. Follow-up studies will pinpoint sinusoidal "hot-spots" for leukocyte trafficking, will reveal changes in vascular properties and architecture, and will identify cancer stem cell activation and dormancy niches in leukemias and myeloproliferative disorders so diseases may be targeted indirectly via its supportive vasculature.

Notch Links Bone and Vascular Niches

Ralf Adams, Max Planck Institute for Molecular Biomedicine

Principles

Blood vessels in the skeletal system provide niche microenvironments for bone-forming cells and thereby promote osteogenesis. Vascular niches for haematopoietic stem cells (HSCs) have similar crucial roles for lifelong hematopoiesis. Our work shows that Notch signaling in endothelial cells leads to the expansion of HSC niches in bone, which involves increases in CD31-positive capillaries and PDGFRβ-positive perivascular cells, arteriole formation, and elevation of cellular stem cell factor levels (Kusumbe et al., Nature *532*, 380–384). In ageing mice, niche-forming vessels in the skeletal system are strongly reduced but can be restored by activation of endothelial Notch signaling.

> *"…Notch signaling in endothelial cells leads to the expansion of HSC niches in bone…."*

What's Next?

It will be import to explore the full nature of the age-related changes in the skeletal vasculature and the consequences for hematopoiesis and bone formation. Our results suggest that the reactivation of certain signaling pathways in the endothelium of the elderly organism might be useful for improved HSC transplantation. In addition, hematopoiesis in the aged organism is characterized by decreased output of lymphoid and erythroid cells, whereas the myeloid lineage is maintained or increased. It will be exciting to explore if and how these changes are caused by alterations in the ageing bone vasculature, which might enable new therapeutic approaches against immunosenescence.